

Project Report

On

# RAMSEY THEORY

*Submitted*

*in partial fulfilment of the requirements for the degree of*

BACHELOR OF SCIENCE

*in*

MATHEMATICS

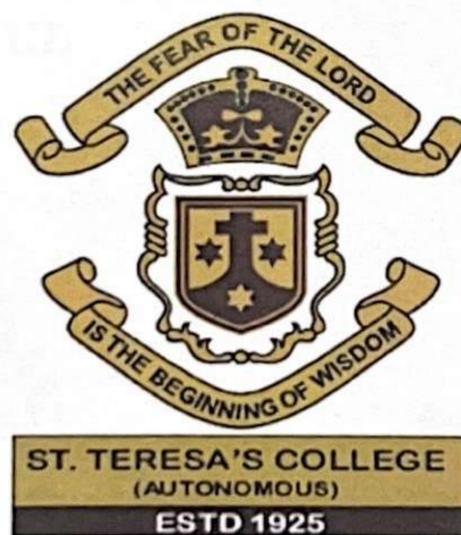
*by*

TANIA N J

(Register No. AB19BMAT057)

*Under the Supervision of*

MRS.NEENU SUSAN PAUL



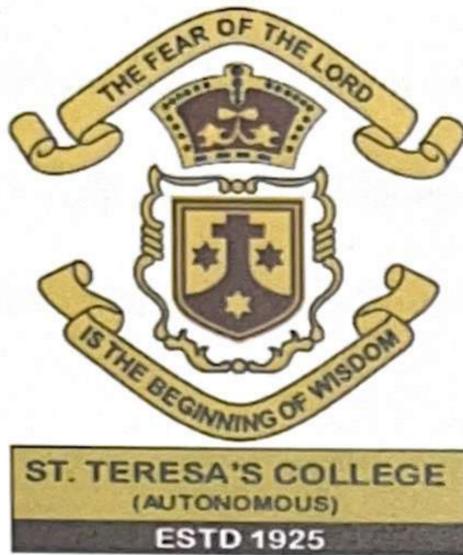
DEPARTMENT OF MATHEMATICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



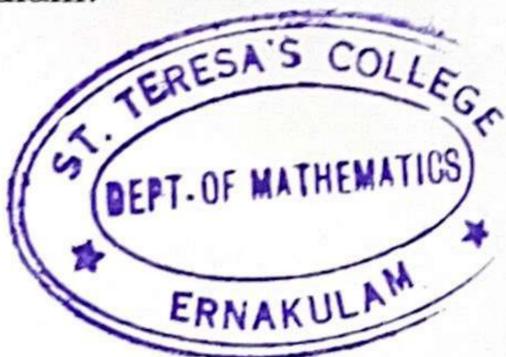
CERTIFICATE

This is to certify that the dissertation entitled, **RAMSEY THEORY** is a bonafide record of the work done by Ms. **TANIA N J** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date:

Place: Ernakulam

*Neenu*  
**Mrs. NEENU SUSAN PAUL**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College (Autonomous),  
Ernakulam.



*Ursala*  
**Dr. Ursala Paul**  
Assistant Professor and Head,  
Department of Mathematics,  
St. Teresa's College (Autonomous),  
Ernakulam.

External Examiners

1: *Neenu*

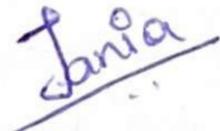
2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Mrs. Neenu Susan Paul, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date:



TANIA N J

AB19BMAT057

# ACKNOWLEDGEMENT

I express my profound gratitude to my internal guide Mrs. Neenu Susan Paul, Assistant Professor, Department of Mathematics, ST. Teresa's college, for her valuable support and guidance through my project.

I would also like to express gratitude for the unending help and support provided by the head of the department throughout the course of project.

I am also thankful to all teaching and non-teaching staffs of Mathematics department, my parents, family members, my group members, friends and all those who have given me immense support and helped me to complete this project into a successful venture.

All above, my sincere gratitude to our almighty god in securing our safety and furnishing us the necessary resources that leads for the accomplishment of this project.

Ernakulam.

Date:

TANIA N J

AB19BMAT057

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<i>INTRODUCTION</i> . . . . .	vii
<i>PRELIMINARIES</i> . . . . .	viii
<b>1 RAMSEY THEORY</b> . . . . .	<b>1</b>
1.1 RAMSEY THEORY-DEFINITION . . . . .	1
1.1.1 EXAMPLE 1 . . . . .	2
1.1.2 EXAMPLE 2 . . . . .	2
1.2 HISTORY OF RAMSEY THEORY . . . . .	2
1.2.1 THE BAUDET-SCHUR-VAN DER WAERDEN THEOREM . . . . .	3
1.2.2 THE GENERALISED SCHUR THEOREM . . . . .	3
1.2.3 THE FRANK PLUMPTON RAMSEY PRINCIPLE . . . . .	4
<b>2 PIGEONHOLE PRINCIPLE, RAMSEY THEOREM AND RAMSEY   NUMBERS</b> . . . . .	<b>5</b>
2.1 THE PIGEONHOLE PRINCIPLE . . . . .	5
2.2 RAMSEY NUMBERS AND RAMSEY THEOREM . . . . .	7
2.3 RAMSEY THEOREM FOR TWO COLORS . . . . .	7
2.4 RAMSEY THEOREM FOR $r$ COLORS . . . . .	8
2.5 RAMSEY NUMBERS . . . . .	9
2.6 RAINBOW-RAMSEY-NUMBERS . . . . .	12

3	THE PARTY PROBLEM	14
3.1	THE PARTY PROBLEM . . . . .	14
3.2	A GENERALISATION OF THE PARTY PROBLEM .	17
3.3	PYTHON PROGRAM . . . . .	21
4	APPLICATIONS OF RAMSEY THEORY	25
4.1	NUMBER THEORY . . . . .	26
4.2	SCHUR'S THEOREM . . . . .	26
4.3	AUTOMATED THEOREM PROVING . . . . .	26
4.4	THE NO-TRIANGLE GAME . . . . .	27
	<i>CONCLUSION</i> . . . . .	30
	<i>REFERENCES</i> . . . . .	31

# INTRODUCTION

Ramsey theory named after the British mathematician and philosopher Frank P Ramsey is a branch of mathematics that focuses on the appearance of order in substructure of a known size. Ramsey was born to mathematician Arthur Stanley Ramsey and Mary Agnes Stanley on 22 February 1903 in Cambridge. His Father was also the President of Magdalene college. He entered Winchester college in 1915 and later on returned to Cambridge to study mathematics at Trinity college.

In 1923, he received his bachelor's degree in mathematics passing his examinations with the result of first class with distinction and was named Senior Wrangler.

Ramsey theory deals with the guaranteed occurrence of specific structures in some part of a large arbitrary structure which has been partitioned into finitely many parts. Ramsey theory is based around the idea that "any structure will necessarily contain an orderly substructure".

A typical result in Ramsey theory starts with some mathematical structures that is then cut into pieces. Problems in Ramsey theory are in a way that typically ask a question of the form how many elements of some structure must there be to guarantee that the particular property will hold?

# PRELIMINARIES

## GRAPH

A graph  $G=(V , E)$  is a set  $V$  of Points called vertices, and a set  $E$  of pairs of vertices called edges.

## SUBGRAPH

Subgraph  $G'=(V',E')$  of a graph  $G=(V , E)$  is a graph such that  $V' \subseteq V$  and  $E' \subseteq E$

## COMPLETE GRAPH

Complete graph on  $n$  vertices, denoted as  $K_n$ , is a graph on  $n$  vertices, with the property that every pair of vertices is connected by an edge.

## ORDER OF A GRAPH

The number of vertices of a graph is called the order of a graph.

## GRAPH COLORING

Graph coloring generally come in two varieties: edge coloring and vertex coloring.

## EDGE-COLORING

Edge - coloring of a graph is assignment of a color to each edge of the graph.

## VERTEX-COLORING

Vertex coloring is the assignment of color to each vertex of the graph.

## MONOCHROMATIC GRAPH

A graph that has been edge colored is called monochromatic graph if all of its edges are same color.

## S-EDGE COLORING

If there are  $S$  different colors used in edge coloring of graph  $G$  then it is called a  $S$  edge coloring of  $G$ .

## Chapter 1

# RAMSEY THEORY

---

### 1.1 RAMSEY THEORY-DEFINITION

Ramsey theory is a branch of mathematics that focuses on the appearance of order in a substructure given a structure of a specific size. The essence of Ramsey theory is that it is a field of mathematics that focuses on items namely graphs of a certain order guaranteeing subgraphs of another order.

The Ramsey type theorems can be seen in many branches of mathematics like number theory, geometry, topology, set theory etc.

Ramsey theory is the study of properties that must occur for sufficiently large structures. We can try to break the structure through partitioning, but a Ramsey property will persist. Ramsey properties exist on integers and graphs which can be seen through the examples given below

### 1.1.1 EXAMPLE 1

Consider positive integer solutions to  $x+y = z$ . Of course,  $1+1 = 2$  and  $2 + 3 = 5$  work if we allow all positive integers. So, let's try to break this by splitting the positive integers into two parts. In Ramsey theory, we typically use colors to describe the partitions, so we will say, red integers and blue integers.

Must we still have a solution to  $x+y = z$ , if we now require the integers to be in the same partition, i.e., the same color? Let's see if we can avoid the property of one part of the partition having a solution to  $x + y = z$ . First, 1 and 2 must be different colors (since  $1+1=2$ ) and consequently, 4 must be the same color as 1 (since  $2 + 2 = 4$ ). Let's say that 1 and 4 are red and 2 is blue. Since  $1+4=5$ , we see that 5 must also be blue and consequently, 3 must be red (since  $2+3=5$ ). But now 1,3, and 4 are all red, so the Ramsey property persists.

### 1.1.2 EXAMPLE 2

Ramsey properties also exist on graphs. For example, if we take  $n \geq 3$  vertices and connect every pair of vertices with an edge, we clearly have a triangle with all edges in the same partition. Can we partition the edges in such a way that we no longer have a triangle with all edges in the same partition? The answer is no, provided we have at least 6 vertices.

## 1.2 HISTORY OF RAMSEY THEORY

As we believe now, David Hilbert's Cube Lemma was the first Ramseyan result. In this work, Hilbert proved the theorem of our interest merely as a tool for his study of irreducibility of rational functions with integral coefficients. But it didn't influence anyone at the time and didn't give birth to Ramsey Theory. Then came

Issai Schur's 1916 Theorem, which could have remained unnoticed too, but Schur was first to realise that he had found something new and striking. It consists of two theorems

### a. THE SCHUR'S THEOREM

The Schur 1916 Theorem states that for any positive integer  $n$  there is an integer  $S(n)$  such that any  $n$ -coloring of the set  $[S(n)]$  contain integers  $a, b, c$  of the same color such that  $a+b = c$ .

### b. STRONGER VERSION OF SCHUR'S THEOREM

Strong version of Schur's Theorem states that for any positive integer  $n$  there is an integer  $S^*(n)$  such that any  $n$ -coloring of the initial positive array  $[S^*(n)]$  distinct integers  $a, b, c$  of the same color such that  $a+b = c$ .

However it didn't reach anywhere.

#### 1.2.1 THE BAUDET-SCHUR-VAN DER WAERDEN THEOREM

Bartel Leendert Van Der Waerden published the proof of the conjecture, this gave us a classic theorem that becomes a root for Ramseyan Mathematics. The Baudet - Schur - Van Der Waerden Theorem states that for any  $k, l$  there is  $W = W(k, l)$  such that any  $k$ -coloring of the set  $[w]$  contains an  $l$ -term monochromatic arithmetic progression.

#### 1.2.2 THE GENERALISED SCHUR THEOREM

Then came the Generalized Schur Theorem which states that for any two positive integers  $k$  and  $l$  there is a positive integer  $[S(k, l)]$  contains a monochromatic arithmetic progression of length  $l$  together with its difference.

### 1.2.3 THE FRANK PLUMPTON RAMSEY PRINCIPLE

The Frank Plumpton Ramsey Principle has two theorems:-

#### a. INFINITE RAMSEY PRINCIPLE

Infinite Ramsey Principle states that for any positive integers  $k$  and  $r$ , if the collection of all  $r$ -element subsets of a infinite set  $S$  is colored in  $k$  colors, then  $S$  contains an infinite subset  $S_1$ , such that all  $r$ -element subsets of  $S_1$ , are assigned the same color.

#### b. FINITE RAMSEY PRINCIPLE

Finite Ramsey Principle states that for any positive integers  $r$ ,  $n$  and  $k$  there is an integer  $m_0 = R(r,n,k)$  such that if  $m \geq m_0$  and the collection of all  $r$ -element subsets of an  $m$ -element set  $S_m$  is colored in  $k$  colors then  $S_m$  contains an  $n$  element subset  $S_n$  such that all  $r$ -element subsets of  $S_n$  are assigned the same color.

## Chapter 2

# PIGEONHOLE

## PRINCIPLE, RAMSEY

## THEOREM AND RAMSEY

## NUMBERS

---

### 2.1 THE PIGEONHOLE PRINCIPLE

The pigeonhole principle can be stated in this way, if more than  $n$  pigeons are put into  $n$  pigeonholes, then some pigeonhole must contain at least two pigeons. In Mathematical language, it can be stated as

**THEOREM 1 :** If an  $n$  - element set is partitioned into  $r$  disjoint subsets where  $n > r$ , then atleast one of the subsets contains more than one element. The above theorem is a special case of the following more general principle.

**THEOREM 2 : ( GENERALISED PIGEONHOLE PRINCIPLE)**

If more than  $mr$  elements are partitioned into  $r$  sets, then some set contains more than  $m$  elements.

**PROOF**

Let  $S$  be a set with  $|s| > mr$ .

Let  $S = S_1 \cup S_2 \cup S_3 \dots S_r$  be any partition of  $S$ .

Assume for a contradiction that  $|S_i| \leq m$  for all  $i = 1, 2, 3 \dots r$ . Then,

$$|S| = \sum_{i=1}^r |S_i| \leq mr \text{ a contradiction.}$$

Hence, for at least one  $i$ , the set  $S_i$  contains more than  $m$  elements,

$$\text{i.e., } |S_i| \geq m + 1$$

we can see of that Theorem 1 is a special case of theorem 2 by taking  $m = 1$ .

**EXAMPLE 1:** For each integer  $n = 1, 2, 3 \dots 200$ , let  $R(n)$  be the remainder when  $n$  is divided by 7. Then some value of  $R(n)$  must occur at least 29 times.

To see this let us think of 200 integers as the pigeons, and the seven possible values of  $R(n)$  as the pigeonholes. Then according to Theorem ( Generalized pigeonhole principle). Since  $200 > 28(7)$ , one of the pigeonholes must contain more than 28 times.

**EXAMPLE 2:** Color each point in the  $xy$  plane having integer coordinates either red or blue. We show that there must be a rectangle with all vertices of the same color. Consider lines  $y=0, y=1,$  and  $y=2$  and their intersections with the lines  $x=i, i=1, 2, \dots, 9$ . On each line  $x=i$  there are three intersection points colored either red or blue. Since there are only  $2^3=8$  different ways to color three points either red or blue, by pigeonhole principle two of the vertical lines say  $x=j$  and  $x = k \neq j$  must have the identical coloring (i.e, the color of  $(j, y)$  is same as the color of  $(k, y)$  for  $y= 0, 1, 2$ ).

Using the pigeonhole principle again, we see that two of the points  $(j,0)$ ,  $(j,1)$  and  $(j, 2)$  must be the same color, say  $(j, y_1)$  and  $(j, y_2)$ . Then the rectangle with vertices  $(j, y_1)$ ,  $(j, y_2)$ ,  $(k, y_1)$ ,  $(k, y_2)$  is the desired rectangle.

Here we used the colors as the pigeonholes.

## 2.2 RAMSEY NUMBERS AND RAMSEY THEOREM

Ramsey's theorem can be considered a refinement of the pigeonhole principle but we are not only guaranteed a certain number of elements in a pigeonhole, but we also have a guarantee of a certain relationship between these elements. It is a theorem that is normally stated in terms of the mathematical concept known as a graph.

The theorem of Ramsey makes a statement about partitioning the  $r$ -element subsets of objects into classes. Roughly put, Ramsey's Theorem says that whenever we partition the  $r$ -sets in a sufficiently large set into  $k$ -classes, there is a  $P$ -subset of  $S$  whose  $r$ -sets all lie in the same class.

A partition is a separation of a set into subsets and the set we want to partition consists of subsets of another set, so for clarity we use the language of coloring instead of the language of partitioning. A  $k$ -coloring of a set is a partition of it into  $k$ -classes. A class or its label is a color. Typically we use  $[K]$  as the set of colors in which case a  $k$ -coloring of  $X$  can be viewed as a function  $f: X \rightarrow [K]$

## 2.3 RAMSEY THEOREM FOR TWO COLORS

Let  $k, l \geq 2$ . There exists a least positive integer  $R = R(k, l)$  such that every edge coloring of  $K_r$  with the colors red and blue, admits either a red  $K_k$  subgraph

or a blue  $K_l$  subgraph.

## PROOF

First note that  $R(k, 2) = k$  for all  $k \geq 2$ , and  $R(2, l) = l$  for all  $l \geq 2$  we proceed via induction on the sum  $k+l$  having taken care of the case when  $k+l=5$ . Hence let,  $k + l \geq 6$  with  $k, l \geq 3$ . We may assume that both  $R(k, l-1)$  and  $R(k-1, l)$  exist. We claim that  $R(k, l) \geq R(k-1, l) + R(k, l-1)$ , which will prove the theorem.

Let  $n = R(k-1, l) + R(k, l-1)$ . Now choose one particular vertex 'v' from  $K_n$ . Then there are  $n-1$  edges from v to other vertices. Let A be the number of red edges and B be the number of blue edges coming out of v. Then either  $A \geq R(k-1, l)$  or  $B \geq R(k, l-1)$ .

Since  $A < R(k-1, l)$  and  $B < R(k, l-1)$ , then  $A+B \leq n-2$  contradicting the fact that  $A+B = n-1$ . We may assume, without loss of generality that  $A \geq R(k-1, l)$ .

Let V be the set of vertices connected to v by a red edge so that  $|V| \geq R(k-1, l)$ .

By inductive hypothesis,  $K_V$  contains either red  $K(k-1)$  subgraph or blue  $K_l$  subgraph. If it contains a blue  $K_l$  subgraph, we are done. If it contains a red  $K(k-1)$  subgraph, then by connecting v to each vertex of this red subgraph we have a red  $K_k$  subgraph (Since v is connected to V by only Red edges) and the proof is complete.

## 2.4 RAMSEY THEOREM FOR r COLORS

For every tuple of integers  $l_1, l_2, \dots, l_r \geq 2$  there is a positive integer R such that the following holds:

For every complete graph G of R vertices and every r-coloring of edges of G, there is a  $i \in \{1, \dots, r\}$  and a complete subgraph of G of order  $l_i$  of which all edges have

color  $i$ .

## 2.5 RAMSEY NUMBERS

The Ramsey number  $R(n, k)$  is the minimum positive integer such that for  $m \geq R(n, k)$ , any red-blue coloring of the edges of  $K_m$  contains either a complete subgraph with blue edges on  $n$  vertices or a complete subgraph with red edges on  $k$  vertices.

### EXAMPLE 1

$R(2, 2) = 2$ . In  $K_2$  there are only two colorings because there is only one edge, the edge is red or blue. In either case, there is a complete monochromatic subgraph of size 2.

### EXAMPLE 2

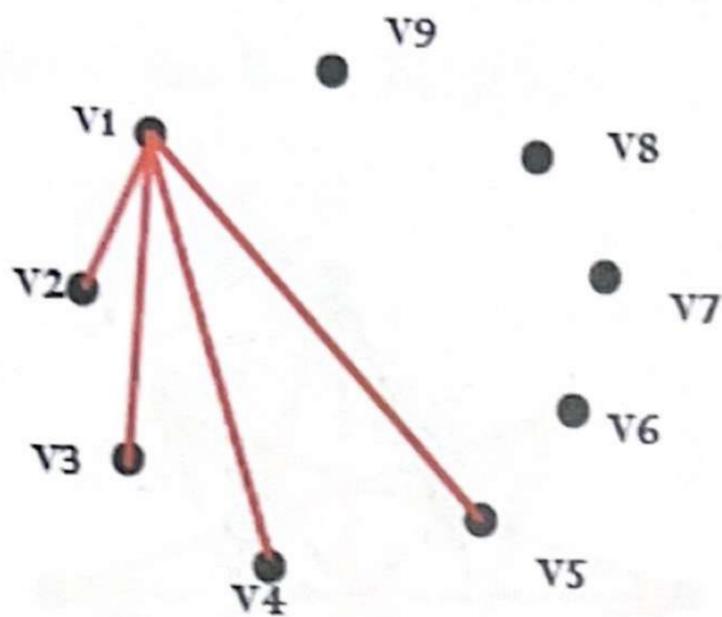
$$R(K_3, K_4) = 9$$

Let there be given a red-blue coloring of  $K_9$ ,

Now we have to show that there is either a red  $K_3$  or a blue  $K_4$ . First observe that it cannot occur that every vertex of  $K_9$  is incident with exactly the three red edges for otherwise the subgraph of  $K_9$  induced by red edges of  $K_9$  is a 3 - regular of order of 9, but there is no such graph. Therefore there are two possibilities.

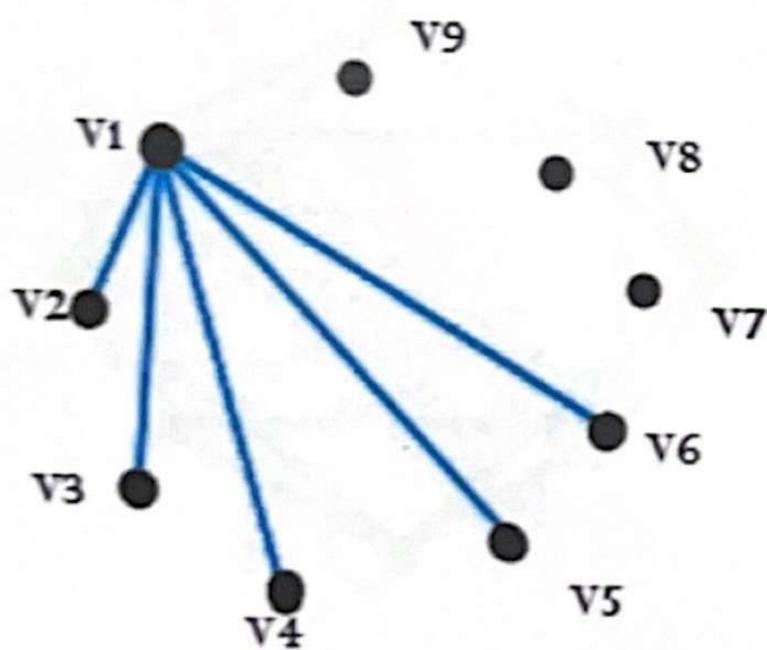
Case 1 :- There exists a vertex  $v_1$  that is incident with 4 red edges. Let

$v_1v_2, v_1v_3, v_1v_4$  and  $v_1v_5$  be the red edges in  $K_9$ . If any two of vertices  $v_2, v_3, v_4$  and  $v_5$  are joined by red edges then  $K_3$  is produced; otherwise every two of the vertices  $v_2, v_3, v_4$  and  $v_5$  are joined by a blue edge producing a blue  $K_4$ .



Case 2 :- There exists a vertex  $v_1$  that is incident with 6 blue edges. Let  $v_1v_2, v_1v_3, v_1v_4, v_1v_5, v_1v_6$  and  $v_1v_7$  be blue edges. Since  $R(K_3, K_3)=6$  the subgraph

$$H = \{v_2, v_3, v_4, v_5, v_6, v_7\} \simeq K_6$$



Contains either a red  $K_3$  or a blue  $K_3$ . If  $H$  contains a red  $K_3$  so does  $K_9$ . If  $H$  contains a blue  $K_3$ , then  $K_9$  contains a blue  $K_4$ .

Therefore  $R(K_3, K_4) \leq 9$ .

Consider the red-blue coloring of  $K_8$  in which the red and blue subgraph of  $K_8$  as

shown in figure (a) and (b) given below respectively. Since there is neither a red  $K_3$  nor a blue  $K_4$ , it follows that  $R(K_3, K_4) \geq 9$ .

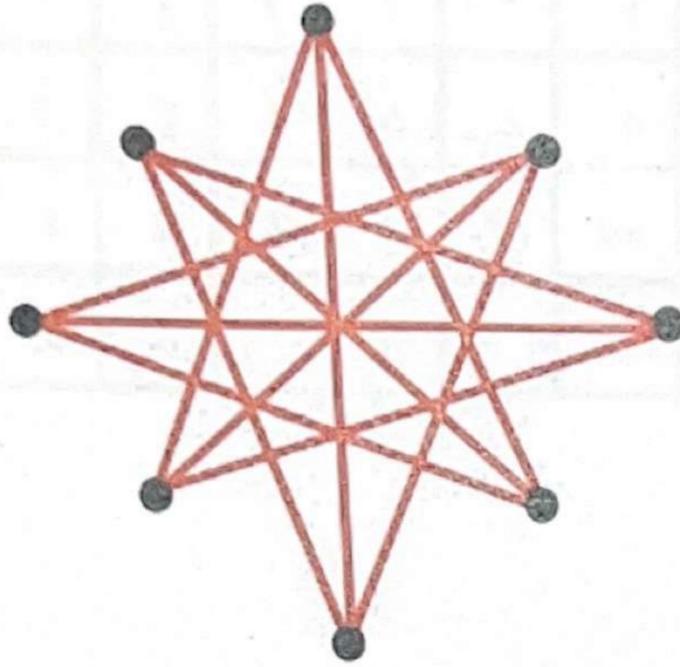


fig (a)

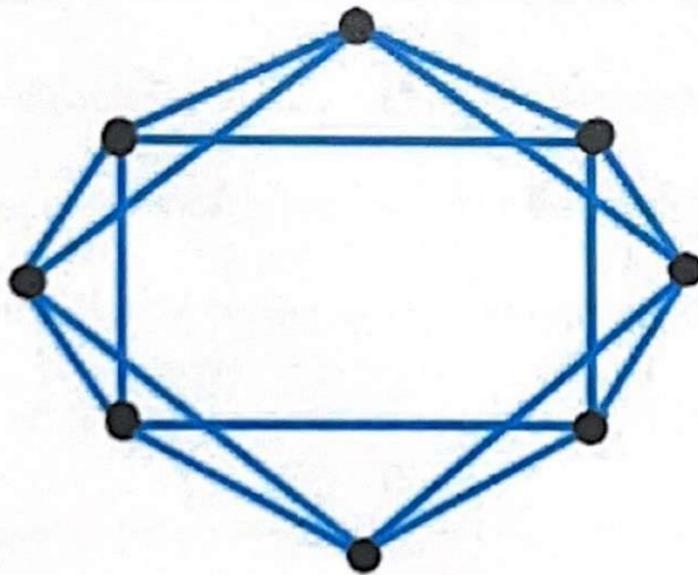


fig (b)

The following table shows all Ramsey numbers  $R(k,l)$  known to date.

k \ l	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7
3	1	3	6	9	14	18	23
4	1	4	9	18			

## 2.6 RAINBOW-RAMSEY-NUMBERS

We have seen that Ramsey numbers originated with the coloring of certain subsets of integers. There is a closely related concept that concerns coloring the integers belonging to certain sequences of integers. By the length of a finite sequence of integers is meant the number of terms in the sequence.

A sequence of integers is monochromatic if every integer in the sequence is colored the same and a sequence is rainbow if no two sequence are colored the same.

If  $F$  is a graph all of whose edges are colored the same, then we refer  $F$  monochromatic  $F$ . If all edges of  $F$  are colored different by, then  $F$  a rainbow  $F$ .

Let the vertex set of a complete graph  $K_n$  be a set of  $n$  positive integers. A coloring of edges of  $K_n$  is called a minimum coloring if two edges  $ij$  and  $kl$  are colored the same if and only if

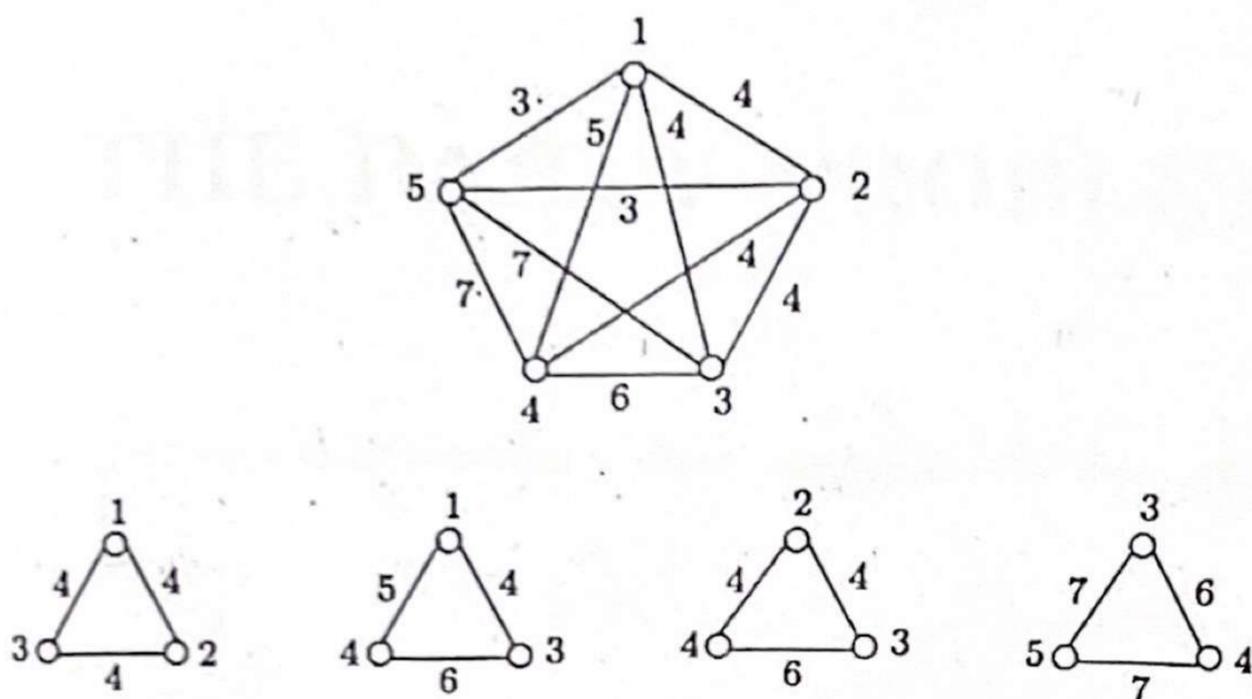
$$\min \{ i,j \} = \min \{ k,l \}$$

While  $n$  coloring of the edges of  $K_n$  is called a maximum coloring if two edges  $ij$

and  $kl$  are colored the same if and only if

$$\max \{i, j\} = \max \{k, l\}$$

A coloring of the edges of  $K_5$  with vertex set  $\{1, 2, 3, 4, 5\}$  is shown below. The graph contains a monochromatic triangle and a rainbow triangle as well as a triangle with minimum coloring and a triangle with a maximum coloring.



Arie Bialostocki and William Voxman defined for a non-empty graph  $F$ , the rainbow Ramsey number  $RR(F)$  as the smallest positive integer  $n$  such that if each edge of a complete graph  $K_n$  is colored from any number of colors, then either a monochromatic  $F$  or a rainbow  $F$  is produced. However Rainbow Ramsey numbers are not always defined.

## Chapter 3

# THE PARTY PROBLEM

---

In this chapter we will be discussing about special case of Ramsey theorem called the party problem.

### 3.1 THE PARTY PROBLEM

The party problem, also known as the maximum clique problem, asks to find the minimum number of guests that must be invited so that at least  $m$  will know each other or at least  $n$  will not know each other. The solutions are known as Ramsey numbers.

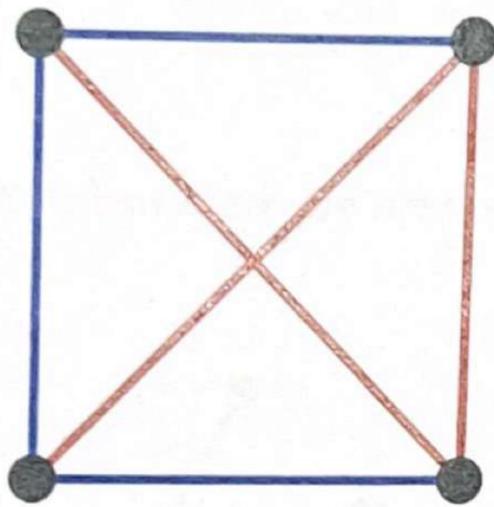
Let us first consider the definition of a graph.

A graph  $G = (V, E)$  is a set  $V$  of points called vertices, and a set  $E$  of pairs of vertices called edges. In the party problem we would like to know how many people must be invited to a party where we can guarantee that there is a group of 3 people who all are either friends or strangers. We want to invite the minimum number of people because we would like to spend as little money as possible.

To clarify the problems we will make some assumptions.

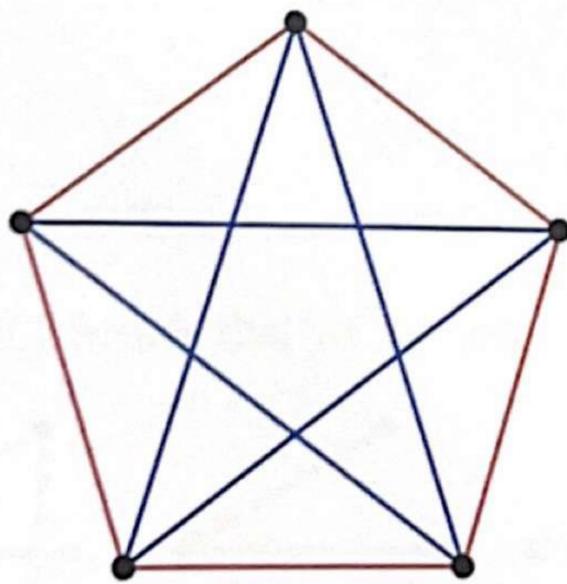
- Every pair of people at the party is a pair of friends or strangers (not both)
- The stranger and friend relationships are symmetrical

To solve this problem, we will represent each person at our party as a vertex on our graph. We will place a red edge between every pair of friends and a blue edge between every pair of strangers.

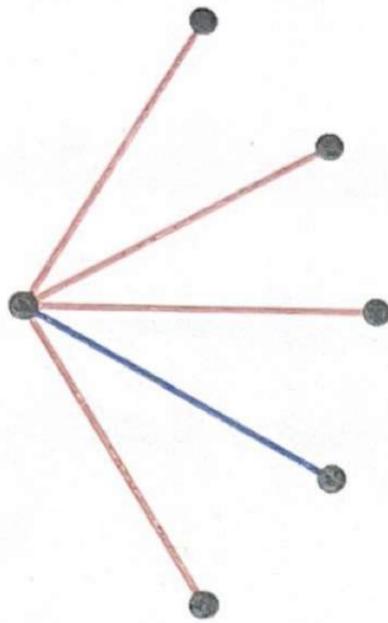


This graph shows us that inviting 4 people is not enough to guarantee a group of 3 mutual friends or a group of 3 mutual strangers.

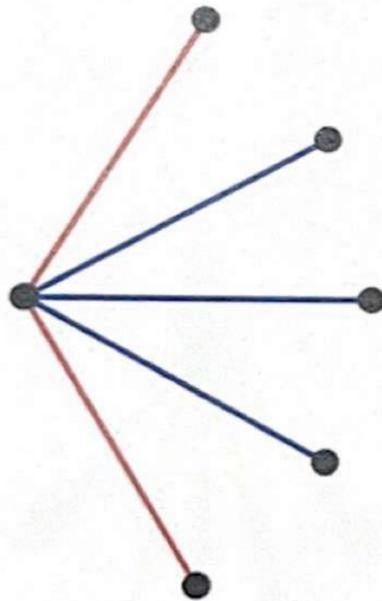
What about 5 people?



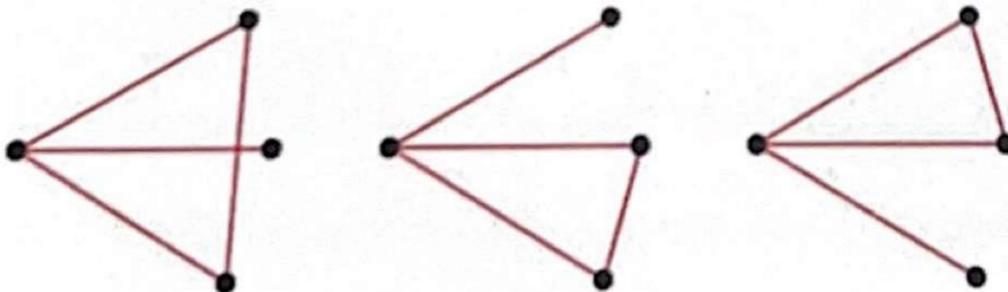
Five is also not enough people, but 6 is enough people and we can prove it by isolating one person.



No matter how we colour the edges there are always atleast 3 edges that are of the same color.



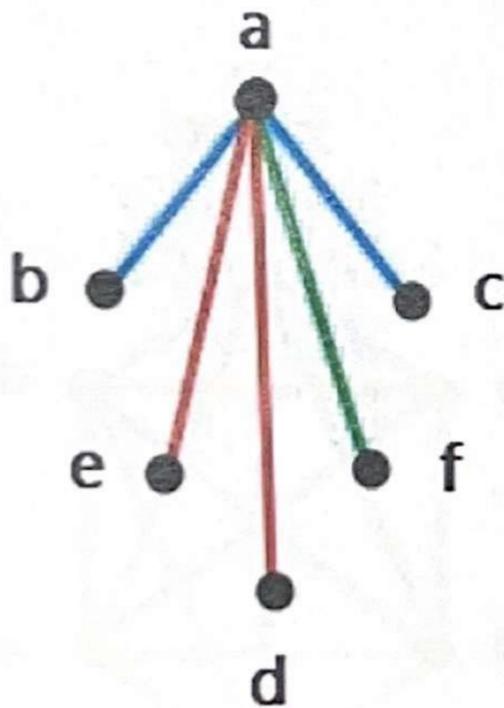
Suppose these 3 edges are red. Then if any of the edges between 3 people are also red, we have the group of 3 friends that we wanted.



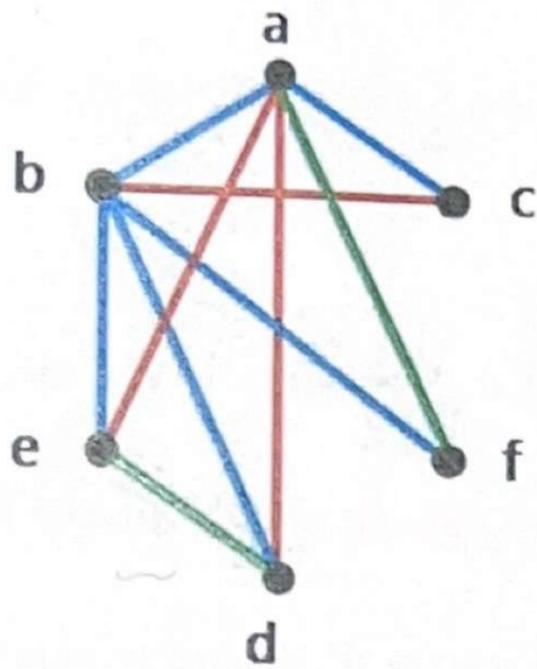
### 3.2 A GENERALISATION OF THE PARTY PROBLEM

Generally we have to take any complete graph  $K_n$  and  $s$  colours. If we color each edge of  $K_n$  arbitrarily in one of these ' $s$ ' colours we might ask if we are forced to have a monochromatic triangle.

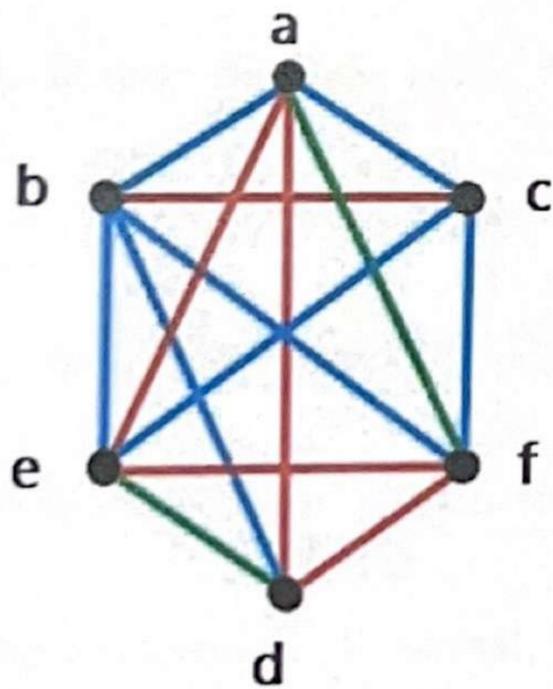
Suppose that the vertices of  $K_6$  are as  $a, b, c, d, e$  and  $f$ . As we have just seen, for our construction we can choose a vertex, say  $a$ , which is incident on three edges of different colours say  $ab$  in blue,  $ad$  is red and  $af$  is green. Now color  $ac$  blue and  $ae$  red, so that all the edges incident with ' $a$ ' are non-colored.



At the next stage of the construction we clearly want to avoid ' $bc$ ' being blue and ' $de$ ' being red. So color ' $bc$ ' in red and ' $de$ ' in green. Now, no matter how we color the other edges incident with the  $b$ , we can not get a monochromatic triangle with ' $b$ ' as a vertex. "So color of all the three edges  $bd, be$  and  $bf$  is blue."



Now look at  $c$ . Coloring  $cd$ ,  $ce$ ,  $cf$  blue gives no monochromatic triangle with  $d$  so we need a little caution. We can't color  $df$  blue. But coloring in red is okay. Finally  $ef$  can't be blue but it can be red. This completes red, blue, green coloring of  $K_6$  which has no monochromatic triangle.



Now to see that if  $n$  is an integer such that every  $s$ -edge coloring of  $K_n$  has a monochromatic triangle then every  $s$ -edge coloring  $K_{n+1}$  also has a monochromatic triangle, simply because any such coloring of  $K_{n+1}$  induces one for  $K_n$ . In

view of the above, this raises the question.

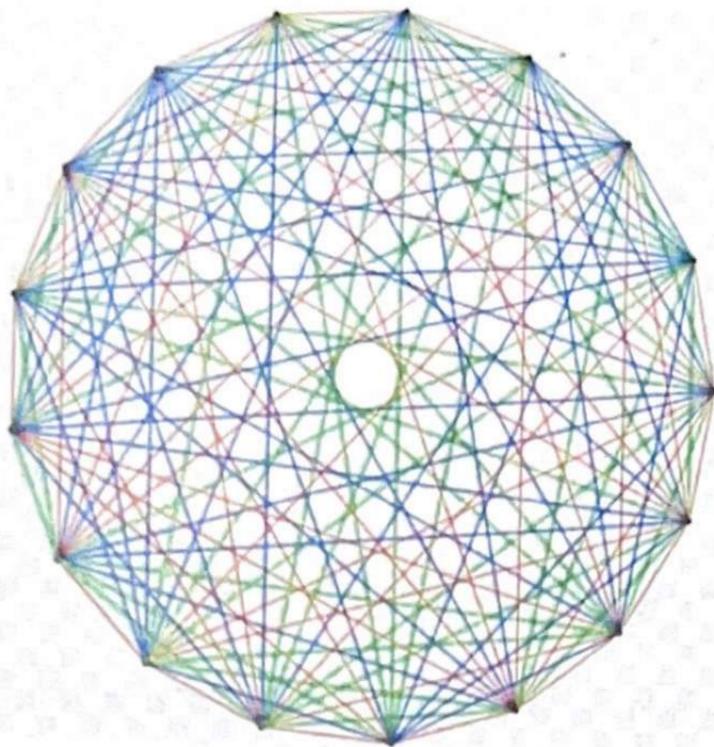
What is the smallest value of 'n' such that every 3-edge coloring of  $K_n$  must have a monochromatic triangle?

One way of tackling this question is to first show that for some 'n' we must have a monochromatic triangle and second to show that we can find a 3-edge coloring of  $K_{n-1}$  which has no monochromatic triangle. If we can find such an 'n' then clearly it must give the smallest  $K_n$  in which every 3-edge coloring has a monochromatic triangle.

Now it turns out to be easy to show that  $K_{17}$ , when 3-edge colored, must have a monochromatic triangle. The proof goes like this

Choose a vertex 'a' of  $K_{17}$ . Since a has degree 16 if we 3-edge color  $K_{17}$ , a must be an end of 6 edges which all have the same color. Without loss of generality, assume that 'a' is the end of the 6 blue edges ab, ac, ad, ae, af and ag. Then if one of the edges joining any pair of the vertices b, c, d, e, f or g is blue, we have a monochromatic triangle and so we are finished. Otherwise the six vertices b, c, d, e, f and g form a  $K_6$  in which the edges are 2-colored, by red and green.

We know that this clique  $K_6$  must contain a monochromatic triangle. Hence  $K_{17}$  must contain a monochromatic triangle when the edges are 3-colored.



### 3.3 PYTHON PROGRAM

The python programme which we developed helps us to find the solution of party problem. In this programme it gives us the minimum number of people to be invited to a party so that the given Ramsey number  $R(n, n)$  is satisfied.

First we will be asked to enter the value for  $n$  (which is the input), then when we enter the value for  $n$  the output will give the minimum number of people.

Then we will be asked, "Do you want to check again" if yes enter 'Yes', if no enter 'No'. If we enter 'Yes' we will be again asked to enter the value for  $n$  and this will continue until we enter 'No'.

For example, Enter the value for  $n$  as 3, then the output will be 6. After the output is given we will be asked, "Do you want to check again" if yes enter 'Yes', if no enter 'No'. Then when we enter Yes, we will be asked to enter the value for  $n$ .

Similarly if we enter the value 4 for  $n$  then the output will be 18. After the output is given we will be asked, "Do you want to check again" if yes enter 'Yes', if no enter 'No'. Then when we enter Yes, we will be asked to enter the value for  $n$ .

Next, if we give the value 5 for  $n$ , then the output will be in the range of 41-48 since it's not yet discovered. After the output is given we will be asked, "Do you want to check again" if yes enter 'Yes', if no enter 'No'.

So the output will give the minimum number of people invited for the given input.

26/02/2022, 12:49

partyProblem.py

```

1 #In the party problem we would like to know how many people must be invited to
  a party where we can guarantee that there is a group of 3 people who either
  all are friends or all strangers. We want to invite the minimum number of
  people because would like to spend as little money as possible.
2 import os
3 def problemdef():
4     print('-----')
5     print(" \n \t \t \t \t \t Party Problem \n")
6     print('-----')
7     print("\n In the party problem we would like to know how many people must
  be invited to a party where we can guarantee that there is a group of 3 people
  who either all are friends or all strangers. We want to invite the minimum
  number of people because would like to spend as little money as possible. \n")
8     print('-----')
9     print("\n Q) What is the minimum number of people that should be invited
  to the party to satisfy the Ramsey number R(n,n)\n")
10 problemdef()
11 y = 1
12 while True:
13     try:
14         n = int(input("\nEnter the value of n : "))
15     except ValueError:
16         print("Type numerical values")
17         continue
18
19     if n <= 1:
20         print('\n->There is no solution for the given condition\n')
21
22     elif n == 2:
23         print("""\n-> Minimum number of people required is "2" to satisfy the
  condition\n""")
24
25     elif n == 3:
26         print("""\n-> Minimum number of people required is "6" to satisfy the
  condition\n""")
27
28     elif n == 4:
29         print("""\n-> Minimum number of people required is "18" to satisfy the
  condition\n""")
30
31     elif n == 5:
32         print("""\n-> Minimum number of people required will range from "43 -
  48" to satisfy the condition\n""")
33
34     else:
35         print("\n->Solution for the given condition is not discovered yet!\n")
36
37     print('-----')
38     a = 0
39     while a == 0:
40         c = input("Do you want to check again. If yes enter 'Yes', if no enter
  'No'\n")
41         if(c == "no") or (c == "No"):
42             a = 1
43             exit()

```

localhost:4649/?mode=python

1/2

26/02/2022, 12:49

```
44     elif c == "yes" or c == "Yes":
45         a = 1
46         os.system('clear')
47         problemdef()
48         continue
49
50
51
52
53
54
```

partyProblem.py

localhost:4649/?mode=python

2/2

## Party Problem

In the party problem we would like to know how many people must be invited to a party where we can guarantee that there is a group of 3 people who either all are friends or all strangers. We want to invite the minimum number of people because would like to spend as little money as possible.

Q) What is the minimum number of people that should be invited to the party to satisfy the Ramsey number  $R(n,n)$

Enter the value of n : 6

->Solution for the given condition is not discovered yet!

Do you want to check again. If yes enter 'Yes', if no enter 'No'  
yes

Q) What is the minimum number of people that should be invited to the party to satisfy the Ramsey number  $R(n,n)$

Enter the value of n : 3

-> Minimum number of people required is "6" to satisfy the condition

Do you want to check again. If yes enter 'Yes', if no enter 'No'  
no

## Chapter 4

# APPLICATIONS OF RAMSEY THEORY

---

In this chapter we will be discussing about some interesting applications of Ramsey theory.

Ramsey-type theorems have roots in different branches of mathematics and the theory developed from them influenced such diverse areas like number theory, set theory, geometric ergodic theory and theoretical computer science. Ramsey-type theorems are showing that if a large enough system is partitioned arbitrarily into finitely many subsystems, atleast one subsystem has a particular property and thus total disorder is impossible.

Ramsey-type theorems scattered around in different fields were put together to form Ramsey theory. Most of the applications are using existing theorem, but there are also papers mostly by Alon where new Ramsey-type theorems are proved at the same time as applications problems are solved.

## 4.1 NUMBER THEORY

Consider the partition  $(\{1, 4, 10, 13\}, \{2, 3, 11, 12\}, \{5, 6, 7, 8, 9\})$  of the set of integers  $\{1, 2, \dots, 13\}$ . We observe that in no subsets of the partition are the integers  $x, y$  and  $z$  (not necessarily distinct) which satisfies the equation.

$$x+y = z \quad (1)$$

Yet no matter how we partition  $\{1, 2, \dots, 13\}$  into 3 subsets, there always exist a subset of the partition which contains solution to equation (1) Schur proved that in general, given any positive integer  $n$ , there exists an integer  $f_n$  such that in partition of  $\{1, 2, \dots, f_n\}$  into  $n$  subsets, there is a subset which contains solution to equation (1).

One of the earliest Ramsey-type results is Schur's theorem in number theory. If  $N$  is partitioned into a finite number of classes, at least one partition class contains a solution to the equation  $x+y=z$ .

There are a number of interesting results produced during the last few years concerning Schur's theorem and its generalization.

## 4.2 SCHUR'S THEOREM

For any  $r$ -colouring of  $Z > 0$  there exist  $x, y, z \in Z > 0$ , all of the same colour, with  $x + y = z$ .

## 4.3 AUTOMATED THEOREM PROVING

The idea of automated theorem proving goes back to at least Hilbert's decision problem. He asked if there exists an algorithm to decide whether a statement in mathematics is true or not, or an algorithm which would find a proof of any

mathematical statement that has a proof. Turing proved that there is no such algorithm in general.

Ramsey's theorem was actually a lemma to a theorem that shows that in a certain special class of first order logic the statements are decidable. Interestingly, using an extension of the finite Ramsey theorem, Paris and Harrington gave the first natural statement unprovable in finite set theory, or equivalently in Peano Arithmetic.

Gödel's incompleteness theorem implies the existence of such statements. The proof is using techniques of mathematical logic. A combinatorial proof was presented by Ketonen and Solovay, then another shorter one by Loeb and Nešetřil.

#### 4.4 THE NO-TRIANGLE GAME

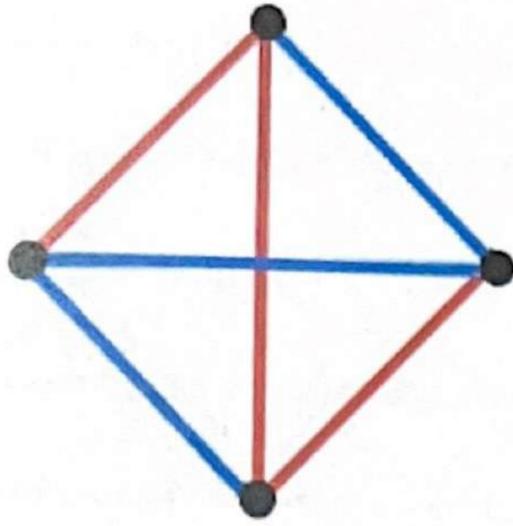
We take the vertices and connect each pair of vertices with an edge. Importantly, edges must be either red or blue. But we can't create any red triangles or blue triangles ie triangles that use these vertices.

Here only one pair of vertices satisfies the no-triangle game.

Let us take  $n$ - number of vertices. Since it's the no triangle game we can't take first 3 vertices because they don't satisfy the no triangle game.

So, let's start with 4 vertices.

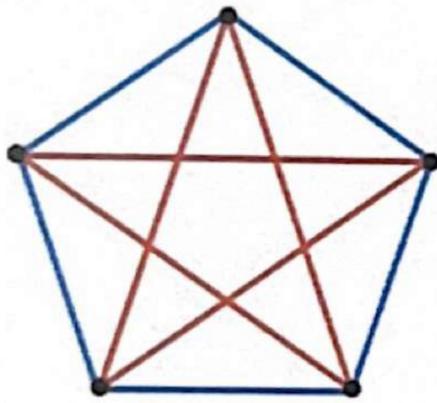
Here are four vertices



Therefore, 4 vertices satisfy the no-triangle game

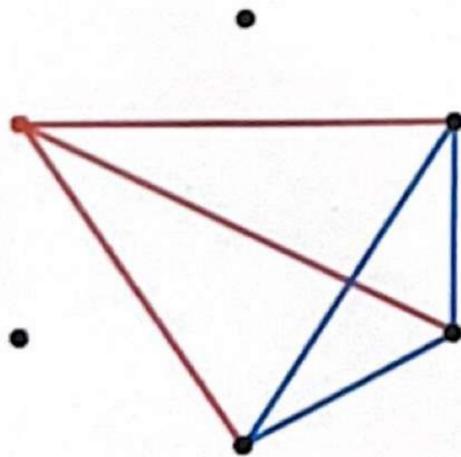
Let's see with 5 vertices.

Here are 5 vertices.



Therefore, 5 vertices also satisfy the game.

Now, let's take 6 vertices.



Here, in every way triangles can be created in both colors.

Therefore it doesn't satisfy the no-triangle game.

We can't satisfy with 6 vertices or any numbers greater than 6.

No-triangle game is based on Ramsey theory. It is a game for two players. Each player in this game chooses a color. Let player A chooses red and player B chooses blue. Red and Blue take turns connecting the six points in the picture. Each pair of points can be connected only once, so if the edge drawn between them is red, it can't be changed.

Players take turns drawing in edges in their colour until one of them creates a monochromatic triangle of red or blue. Whoever creates that triangle first loses.

A version of this puzzle appeared in a math contest in the 1920s, and was written up in more detail by the mathematician Frank Ramsey in a paper published in 1930.

## CONCLUSION

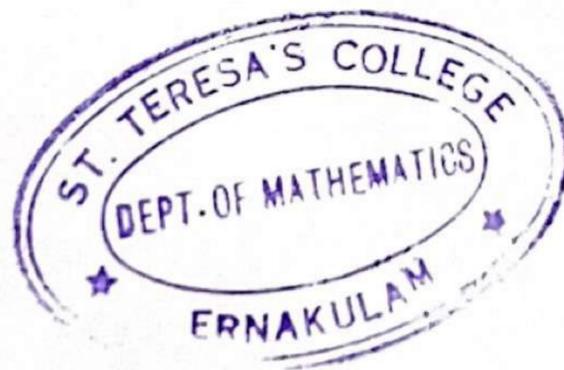
---

In this project we are dealing with Ramsey Theory and its origin. This theory plays a very important role in many fields. In the first chapter we discussed about how Ramsey Theory was developed. In second chapter we discussed about the pigeonhole Principle its generalised form and Ramsey Theorem which is considered as a refinement of the pigeonhole principle. In Ramsey's theorem we discussed about Ramsey's theorem for two colors and for  $r$ -colors. We also discussed about Ramsey numbers and Rainbow Ramsey numbers. In third chapter we discussed about the party problem which is a direct application of Ramsey theorem for 2 colors and we also discussed about a generalisation of the party problem. In 3rd chapter we are also presenting a python programe which we will be able to check the minimum number of people to be invited to a party for the condition to be satisfied. The project is concluded by explaining the applications of Ramsey theory.

## REFERENCES

---

- [1] Douglas. B. West - Introduction to graph theory second edition.
- [2] Gray Chartrand, Ping zhang - Introduction to graph theory.
- [3] J. A. Bondy and U.S.R. Murty - Graph theory with applications.
- [4] Alexander soiefer (Editor) - Ramsey Theory Yesturday, Today and Tomorrow.
- [5] Bruce. M. Landman, Aaron Robertson - Ramsey Theory on the Integers second edition.
- [6] Derek Allan Holton, John Clark - A first Look At Graph Theory.
- [7] Matthew steed- Some Theorems and Applications of Ramsey theory.
- [8] Vera Rosta - Ramsey Theory Applications.
- [9] Jake coulombe - The party Problem An introduction to combinatorics.
- [10] Fred.S Roberts - Applications of Ramsey Theory.
- [11] Lane Barton IV - Ramsey Theory.
- [12] Aaron Robertson - Fundamentals of Ramsey Theory.



Project Report

On

**ANALYSIS ON STARTUP ECOSYSTEM IN  
INDIA**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**MASTER OF SCIENCE**

*in*

**APPLIED STATISTICS AND DATA ANALYTICS**

*by*

**TANIA P R**

(Register No. SM20AS022)

(2020-2022)

*Under the Supervision of*

**MS.SREELAKSHMI M S**



**DEPARTMENT OF MATHEMATICS AND STATISTICS**

**ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



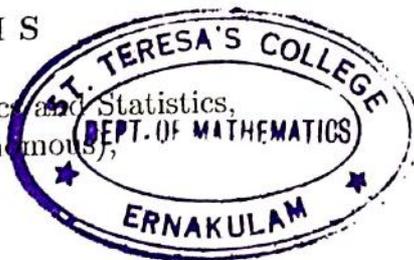
CERTIFICATE

This is to certify that the dissertation entitled, **ANALYSIS ON STARTUP ECOSYSTEM IN INDIA** is a bonafide record of the work done by Ms. **TANIA P R** under my guidance as partial fulfillment of the award of the degree of **Master of Science in Applied Statistics and Data Analytics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 09-05-2022

Place: Ernakulam

*Sreelakshmi*  
Ms. SREELAKSHMI M S  
Assistant Professor,  
Department of Mathematics and Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.



*Shanty B P*

Ms. Shanty B P  
Assistant Professor and HOD,  
Department of Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.

External Examiners

1: *Kumar K*

26/05/2022

2: *m. H. B.*  
26/05/22

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of MS.SREELAKSHMI M S, Assistant Professor, Department of statistics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 09-05-2022

  
TANIA P R

SM20AS022

## ACKNOWLEDGEMENTS

The successful completion of this dissertation owes to the inspiration and constant support that I received from various sources. I avail this opportunity to express my sincere gratitude to all those who helped me directly or indirectly for the completion of work.

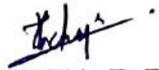
I wish to express my deep gratitude to my guide Ms.Sreelakshmi M S, Assistant Professor, Department of statistics, for the valuable guidance that I have received throughout the period of the study.

I am thankful to Ms.Shanty B P, Head of the Department of Statistics and all the teachers in the department for their supervision, guidance, encouragement and co-operation in improving the quality of the study at various stages.

I sincerely express my gratitude to my teachers, my parents, and my friends and all those who have encouraged me in this endeavor. Above all I am greatly thankful to the grace of Almighty God for the successful completion of my training programme.

Ernakulam.

Date: 09-05-2022

  
TANIA P R  
SM20AS022

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENTS</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 Introduction</b>	<b>1</b>
1.1 What is Startup . . . . .	2
1.2 History of Startups in India . . . . .	3
1.3 Start of Startups Initiative in India . . . . .	4
1.4 Startup Funding Stages . . . . .	5
1.5 Specific Objectives of the Study . . . . .	6
<b>2 Literature Review</b>	<b>7</b>
<b>3 Materials and Methods</b>	<b>13</b>
3.1 Materials . . . . .	13
3.2 Methods . . . . .	14
3.2.1 Exploratory Data Analysis (EDA) . . . . .	14
3.2.2 K-Means Clustering . . . . .	15
3.3 Measure the accuracy or goodness of our clustering technique . . . . .	18
3.3.1 Silhouette Coefficient . . . . .	18
<b>4 Result</b>	<b>20</b>
4.1 How does funding ecosystem changes with time . . . . .	20
4.1.1 Startup Funding Variation Per Month from 2016-2020 . . . . .	21
4.2 Do cities play a major roll in funding? . . . . .	24
4.3 Which industries are favoured by investors for funding . . . . .	25
4.4 Who are the important investors in the Indian Ecosystem . . . . .	26

4.5	How much funds does startups generally get in India . . . . .	29
4.6	Clustering . . . . .	30
4.7	Measure the accuracy of the cluster . . . . .	32
<b>5</b>	<b>Discussion and Conclusion</b>	<b>33</b>
5.1	Python Code . . . . .	35
	<i>REFERENCES</i> . . . . .	37

# Chapter 1

## Introduction

---

1.1 India is a developing south Asian country. It is a most populous and 7th largest country by area. Large population implies a large prospective market in India and puts more pressure for employment in the country. In recent years, Indian youth motivates towards self-employment instead of rely on parents or government or opportunities. Rather, they start to take initiative to look for new challenges. This encourages and promoting self-employment environment creates the startup systems in India. The Indian startup Ecosystem started in the late 1960s with the start of TCS, followed by Infosys in 1981 and others. These startups started as software services serving Indian software needs, and later expanding to exporting software services. The dot-com era was a blessing for many startups, which included marketplaces, ecommerce and vendors. And finally, modern day startups comprised of the latest technology started around 2007-2008, and comprised e-commerce, logistics, marketplaces, and advertising startups. Startups like Flipkart (2007) and IXIGO (2008) were part of these. The Indian startup ecosystem has grown spectacularly during the past 10 years. From a mere 29,000 startups in 2014, the ecosystem is expected to shelter 55,000 startups in 2020 (Inc42, 2020). The ecosystem has received sturdy and constant investment throughout, with venture capital growing 8 times between 2014 and 2019 (Inc42, 2020). The investment enthusiasm remained strong even during Covid-19 when the startup deals came closer to the historic peak of 2017 (Inc42, 2020). The Indian startups are supported by the flagship programme of the government –‘Startup India’, besides receiving support from incubators and accelerators

In addition, startups receive financial assistance from venture capitalists. In the process of receiving financial assistance, startups go through different stages. The seed stage is the first stage of investment where startups receive a modest amount of funds for the development of their product idea. Series A and B stages are growth stages whereas the late-stage is where the startup has made its product commercially available and is in need of investment to acquire more customers. The Indian startup ecosystem is dominated by big late-stage startups who consume a major chunk of investment. The Indian startup ecosystem is transitioning from a young, undeveloped ecosystem to a mature ecosystem. Although reaching full bloom, the ecosystem is characterized by some glaring inequities. Investment is heavily concentrated in a few sectors and a few business models. In addition, a very low proportion of startups are funded. Between 2014 and H12020, only 6 % of the startups were funded (Inc42, 2020). The Indian startup ecosystem is dominated by a few big unicorn1 firms in every sector.

## 1.1 What is Startup

Currently a clear definition of a ‘Startup’ does not exist in the Indian context due to the subjectivity and complexity involved. Considering various parameters pertaining to any business such as the stage of their life cycle, the amount and level of funding achieved, the amount of revenue generated, the area of operations, etc., some conceptual definitions are available in the public domain. The Department of Industrial Policy and Promotion (DIPP) is also working around a clear definition for startups and is expected to make it public in due course.

- A startup is a young company that is beginning to develop and grow, is in the first stages of operation, and is usually financed by an individual or small group of individuals
- A startup is a young company that searches for an unknown business model in order to disrupt existing markets or create new ones.
- A startup is a young, dynamic company built on technology and innovation wherein the founders attempt to capitalize on developing a product or service for which they believe there is a demand.

## **1.2 History of Startups in India**

Indian startups is not just limited to the current century; in fact, it began over four decades ago. Through the 80s, a handful of pioneering IT service companies such as TCS, Infosys and Wipro placed India firmly on the global economic map. These companies invested in technology and leveraged the country's young, English-speaking workforce to provide cost-effective technology services to clients across the globe. The result is an industry that currently employs nearly four million people and generates \$150 bn in revenues.

The 90's witnessed the launch of Airtel, which today is India's largest telecommunications company, with over \$15 bn in revenues and more than 350 million consumers worldwide. Similarly, the 90's also witnessed the launch of ICICI, HDFC and Axis banks, which have gone on to rank amongst the top 10 banks in India as per balance sheet size (as of March 2016). In fact, ICICI and HDFC Bank rank number two and three respectively, after the State Bank of India. Most importantly, they have fundamentally reshaped every element of the Indian banking business model, leading to the betterment of the sector as a whole.

Cut to the present and India has seen a dramatic acceleration in both the quantum and the diversity of startups. Over the last 10 years, India has become one of the top five startup ecosystems in the world, alongside the US, China, the UK, and Israel. The country today is home to a strong angel investor network with the sector witnessing active participation from every major Venture Capitalist (VC) or Private Equity (PE) firm. Over the same period, the startup sector has attracted nearly \$20 bn of capital inflows and today employs about 100,000 people, with that number doubling every two years. Furthermore, over 100 active incubators help entrepreneurs experiment with new ideas, by offering them functional expertise and resources.

However, some of the euphoria has subsided recently. Funding in 2016 has been much harder than the past couple of years. Investors are marking down the value of many startups. They are asking founders to stretch the current funding runway for longer by reducing cash burn, and asking tougher questions around the sustainability of business models as well as profitability. So, does this signal that the Indian startup ecosystem has hit a plateau? Quite the opposite, we think. If

harnessed well, this ecosystem has the potential to change the face of the country over the next two decades.

All successful startups are based on a potentially disruptive consumer value proposition. Think about the convenience of an e-commerce marketplace, or the ability to call a taxi or book a room at the click of a button anywhere, while paying through a mobile wallet. Or consider the benefit of reducing turnaround time in logistics by 50%. What clearly comes through is a wide range of measurable as well as intangible consumer and business benefits. But what is not so obvious is how the startups amplify the economic impact of technology at the grassroots level. For example, e-commerce companies such as Flipkart, Snapdeal and Shopclues etc. have put small businesses on an unimaginable platform. OLA and OYO have helped thousands of taxi drivers and hundreds of small hotel owners achieve significantly higher productivity. Likewise, Rivigo is fundamentally improving the livelihood of its fleet pilots (i.e. truck drivers). It is this ability to inspire transformation at the ground level that gives Indian startups a truly unique character. If this innovation and energy can be directed to areas such as literacy, skill development, healthcare and sanitation, financial inclusion and natural resources, imagine what it can do for India!. Ultimately, the combination of progressive government policies, capital assistance from corporates, guidance from academics and the sheer brilliance of Indian entrepreneurs, cumulatively can create the secret sauce required to unleash the true potential of our startup ecosystem.

While India is a land of many unique problems, its greatest assets include its brainpower and resilience. By harnessing these, it can overcome a wide range of challenges through innovative solutions. Startups provide an excellent environment to nurture talent as well as to leverage the power and scale of latest technologies. In doing so, some of them hold the promise of transforming the country and potentially becoming the next set of Indian technology giants.

### **1.3 Start of Startups Initiative in India**

Honorable Prime Minister Dr. Narendra Modi announced on 15th August, 2015 "Start-up India, Stand up India" to promote Bank Financing for startups and offer incentives to boost entrepreneurship and job creation. Addressing the Nation on the 69th Independence Day, he said, "We are looking at systems for enabling start-

ups. We want to enable start-ups to make India No. 1 in this field. . . . Start-up India; Stand up India.” The Government’s Union Budget allocation of INR 1,000 Cr. towards the Self Employment and Talent Utilization (SETU) scheme is a major boost towards promoting start-ups in the Country.

## 1.4 Startup Funding Stages

There are multiple stages of startup funding: Seed, Series A, Series B, Series C, and so forth. Startups should be conscientious about the funding rounds that they will go through, which are generally based on the current maturity and development of the company. Here’s an overview of the major startup stages.

- **Seed Capital** This is the very first investment of money that is used to start a business, be it for research or development of the prototype of the product or simply the funding that will help you to focus on your project before you take it up full time. This is also where you use your own capital, or take help of FF, that is family and friends

- **Angel Funding** Once you have raised as much as you can from the FFF (Friends, Family, and Fools) you should start looking for Angel Investors. No matter how generous your family and friends are, those finances are limited and the majority of the time, not even enough to help you reach your dreams. We call these investors ‘angel investors’ as these are those people who are wealthy and outside the initial group of your family and friends.

- **Venture Capital (Series A, Series B, Series C, etc.)** Venture Capital or VC comes into the picture when you have already launched your business and have started with the distribution or sales. While the seed and investor money help you launch, they are not bringing you anything in return as you are just spending at that stage and not earning anything in return. At this stage, you should approach a venture capital or VC for funding so as to grow your company and to finally move towards making a profit. There are multiple rounds of Venture Capital financing with each round typically given a letter like A followed by B followed by C, etc.

- **Mezzanine Financing and Bridge Loans** Once the company has progressed and are much more profitable than when they approached VCs for funding. The reason they are looking for bridge loans or mezzanine financing is that they may be eyeing an IPO opportunity or are looking at acquisition and need more funds

or are planning a management buyout and require additional funds. They thus tap into mezzanine financing or bridge financing for extra funds and help. These mezzanine finances are often used six to 12 months before an IPO and once an IPO is acquired they use the IPOs proceeds to pay back the mezzanine financing investors.

- IPO (Initial Public Offering) IPO stands for Initial Public Offering and is used by the company who have finally reached the stage where they can raise money by selling their stocks to the public. The IPO's opening stock price is typically set with the help of investment bankers and financiers who commit to selling X number of the company's shares at Y price so as to raise money for the company. Once the company stock is out for IPO, it is traded through a stock exchange like the Bombay Stock Exchange (BSE) or the National Stock Exchange (NSE) in India. As a company still looking for initial startup funding stages, you might think that Venture Capitals and IPOs are too far away, but they are not.

## **1.5 Specific Objectives of the Study**

1. To analyze the startup ecosystem in India and to extract meaningful insights from data.
2. To understand which Industry the investors are looking forward to invest.

# Chapter 2

## Literature Review

---

An Start-up ecosystem is defined as a set of entrepreneurial factors (both internal and external), entrepreneurial organizations (firms, venture capitalists, business angels, banks) institutions (universities, public sector agencies, incubation centers, financial institutions) and entrepreneurial processes (e.g. business birth rate, number of high growth firms, levels of blockbuster entrepreneurship, number of serial entrepreneurs, degree of sell-out mentality within firms and levels of entrepreneurial ambition), which formally and informally coalesce to connect, mediate and govern the performance within the local entrepreneurial environment (Mason and Brown, 2014).

A start-up ecosystem is formed by people, start-ups in their various stages and various types of organizations in a location, interacting as a system to create new start-up ventures. The ecosystem for start-ups is dynamic. From time to time, ecosystems change in terms of people, organizations and environments (Isenberg, 2011).

Among 4000 successful entrepreneurs, the study of Butler (2017) suggested that the ability to thrive in uncertainty, a passionate desire to author and own project, and unique skills at persuasion as the most distinct character traits. Although these particular characteristics might fit to a 20- year old, the ideal age of startup founder (for otherwise lacking experience) is between 30 and 50 (Kon, Cukier, Melo, Hazzan Yuklea, 2014). Challenging working condition and the pressure of critical decision- making are both part of the rather stressful job of an entrepreneur (Semerci, 2016), which may one day affect his mental or physical health. Therefore, founder with a high tolerance for stress have a valuable assets.

In contrast to common belief, after closely watching over 100 startup companies in the past two decades, (Furr and Ahlstrom, 2011)) discovered that attributes such as passion, vision and determination more often lead to failure than to success. When entrepreneurs invest countless work hours, money and reputation into their project, passion and determination can easily become dogmatism, falling in love with one's product and ignoring honest customer feedback is the reason why most startup fail (Furr Ahlstrom, 2011). Essentially, there should be a beneficial balance between being confident about what you know while at the same time distrusting your knowledge enough to stay eager to learn more (Kelley, Singer Herrington, 2012).

Monitor, G. E., (2005) state that it would be beneficial, especially for complex technology-driven startups, to have a team of founders rather than one single founder. It allows the company to move faster, be more agile to enter a market and more responsive to a change in market conditions. A team also enables opportunities for accelerated and specialized decision making (Eisenhardt Schoonhoven, 1996) as well as faster pace for innovations (Eisenhardt and Tabrizi, 1995). The ideal team of founder and co-founder consist of member with experience in the industry and in leadership. Although these attributes do not need to apply to everyone. In term of education, a heterogeneous team with different background would be preferred over a team of members with the same education (Frank et al 2009). Education itself is regarded as prerequisite for being a successful entrepreneur (Ferrante, 2005) industry experience has a positive impact as well (Walter et al., 2013).

An article was published in the columns of the Economic Times in the wake of the recent changes made in the government policies towards nurturing start-ups. The government of India has taken a positive front towards start-ups. According to the study, the government is considering policy changes to broaden and energize one of PM Narendra Modi's pet initiatives, Start-up India, to include a bigger number of start-ups under the 'medium industry' category that would be eligible for public procurement incentive and preferential benefits. Start-up India was launched with the underlying objective to increase the opportunities of self-employment in India. Under the public procurement policy, central government department and ministries and its central PSUs have to procure at least 20% of

their purchases from micro and small enterprises beginning April 1, 2015. This means that slightly bigger start-ups may miss out on this benefit that could be their initial survival. According to a government official, the governing policy in the matter includes only small and micro enterprises and does not mention medium industry.

Dr. Sabrina Korreck (2019) The paper analyses the current state of the Indian startup ecosystem and has three goals: to provide an understanding of the growth drivers and motivations of Indian startup founders; identify challenges facing these startups; and outline the pillars in place that support them. The analysis uses data collated from semi-structured interviews with startup founders, investors, and representatives of support organizations

Major people considered events of 1991 as the turning points for the Indian Startup ecosystem in India when the country was on the verge of bankruptcy that made the government to rethink its policies. Because of India's ever increasing entrepreneurial performance some analysts consider India as the next Asian miracle (Huang, 2008).

Giurca Vasilescu (2009) suggested that Investors are the main source to provide funding for the developing companies. They provide financial assistance, from the beginning till the company are ready to be on the capital market. Investors also provide managerial support to make companies survive in the competitive market.

Tracxn, a venture capital analytics firm found out in their research that a total amount of \$6.4 billion funding were given to various startups in the first nine month of 2015, out of which \$3.4 billion was invested in online market place in India. (Velayanikal, 2015)

Honorable Prime Minister Dr. Narendra Modi announced the "Start-up India, Stand-up India" on the 15th of August 2015 to promote Bank financing for startups and offer incentives to boost entrepreneurship in India. This is considered as the initial seed to the future of startups in India for the ease of doing business.

The increasing ease of doing business is also bringing in investors in some much-needed but neglected areas. For instance, the focus on the Insolvency and Bankruptcy Code (IBC) has encouraged asset reconstruction companies (CDPQ from Canada and Encore Capital from the US) to invest in India (Hari, 2018).

Crowd funding, which involves raising small amounts of capital from a large number of individuals, is considered to be a major disruption in entrepreneurial financing (Kshetri, 2016). This is because risk taken by investors are less in crowd funding as compared other funding option where risk factor is higher.

Atherton (2012) demonstrated that the financing source of startups (formal or informal) depends on multiple factors and observed an immense disparity between the highly capitalized and under capitalized start-ups

One of the biggest hindrance faced by majority of startup entrepreneurs is finding investment funds to expand or launch a startup, remarked by Berger, Cowan, Frame (2011).

Startup founders first proceed to internal sources of financing (their own funds), followed by use of external financing sources, revealed in a study by Whittam and Wyper (2007).

The enterprises which has been funded from their own sources, friends, and family resources in their start-up, are not any smaller in size as compared to the start-ups funded by bank loans, concluded by Colombo and Grilli (2005) in their study on correlation between the size of start-up companies and potential funding sources.

In another study, Davila, Foster and Gupta (2000) observed that companies which were using Venture funds as the financing source grew faster than those that had used some other financing source.

Maurya (2012) revealed that Bootstrapping and Lean Startups are complementary and both cover techniques for building low-burn startups by eliminating waste through the maximization of existing resources first before expending to acquisition of new or external resources.

Piyush Anand Verma, vikas singal (2018) their paper gave an analytical overview of the boom and potentialities of startup systems in India i.e the progress made by India so far. Therefore, it gave a better understanding of the Investment and financing strategy of entrepreneurial ventures.

Dr suniti chandiok (2016) They concluded that The success of Startup India campaign hinges on initiatives like faster and easier registration of Companies, self-certification for many legal requirements, zero inspection for three years, funding for patents, and speed of patent protection. The study focused on the different

initiatives that taken by government for startups.

Anirudh Garg and Abhishek Krishna Shivam (2017) The study aimed to bring forward the different types of funding used by the start-ups, attitude of government towards growing start-ups and reasons for failure of growing start-ups. Through the study they tried to find out if Start-ups are a bubble about to burst In today's world people can no longer expect large enterprises to guarantee them jobs for life. Individuals are increasingly expected to seek out their own opportunities, actively create value and behave ethically, rather than faithfully follow rules and routines set by others. This leads to the generation of numerous self-employment opportunities. The major chunk of funding market is formed by Venture capital, Angel investment, Government loans and Seed Funding. Crowd funding being new to India still remains a thing of the west.

Madhusudan Narayan, Birajit Mohanty ,Mahesh Kumar (2019) the paper presents the trend and growth pattern of startup ventures, their types and different stages of funding in India. The prime objective of their research was to study growth pattern of the startups and stages of funding received by these startups. The analysis has shown Indian start-up companies prefer primarily funding in four different stages, such as early stage, growth stage, expansion stage and bridge funding. The result has also shown that maximum funding was received in the expansion stage in the financial years. Bridge funding is becoming more and more prevalent, almost 10 per cent of deals.

Shradha Sharma Founder CEO YourStory (2018) Talks about the Indian startup ecosystem and its growth . The report provides an overview of the top 20 ecosystems around the world. Then it provides a comparative analysis on Silicon Valley's, Moscow's and Toronto's ecosystems by describing the differences in the entrepreneurs, market and fund in these ecosystems.

Dharish David, Sasidaran Gopalan, and Suma Ramachandran (2020) focused on The startup environment and funding activities in India.it gave an insight about different startups and its contribution to the country.

Neeraj K Pandey (2018) Developed a regression model for the funding of a startup, which ultimately help to improve the efficiency of an entrepreneurial ecosystem.

BCG article - India's Startup Story: Unlocking True Potential. Discuss about the startup history of India from the beginning to this time.

Surbhi Jain (2018) The paper Growth of startup ecosystems in India. focused on the growth and prospects of Startup systems in India. It says that the great problems in Indian markets are that they are unorganized and fragmented. There is a lack of unambiguous and transparent policy motives, lack of communications sources, lack of knowledge and exposure. Startup systems require a combination of friendly operational, regulatory and taxation issues that affects the working of business environment very much. For building an entrepreneurial environment, the government, big corporate, educational institutions should come forward to provide a culture for startups in India.

# Chapter 3

## Materials and Methods

---

### 3.1 Materials

In the present study, data is collected from trak.in. This dataset contains all startups that got funding between the period January 2016 and April 2021. The dataset contains several different attributes like Location, Industry, Investors, Sub-Vertical, Investment Type, Date, Startup Name, Amount in USD and Website URL

Table 3.1: Data set- Features and description

Features	Description
Sl.No	Unique Serial Nos
Date	Date of investment
Startup Name	Name of the Startup Company
Industry	Industry to which the startup belongs
Sub-Vertical	category of the industry type
Location	Location of the startup
Investor	Name of the Investor
Investment Type	Kind of funding they provided
Amount in USD	Funding Amount in USD
Website URL	Web address

The file indian\_startup\_funding\_keep.csv has the complete information we need about all the startups. The dataset has 2276 rows and 10 columns. The analysis has been done in a Python Jupyter Notebook.

The study focuses on giving the answers to the following questions through exploratory data analysis

- How does the funding ecosystem change with time?

- Do cities play a major role in funding?
- Which industries are favored by investors for funding?
- Who are the important investors in the Indian Ecosystem?
- How much funds does startups generally get in India?

The study also focuses to find the Industry the investors are looking forward to invest.

## 3.2 Methods

### 3.2.1 Exploratory Data Analysis (EDA)

The basic process of data analysis comprises only 4 steps and is a given below and is followed as well.

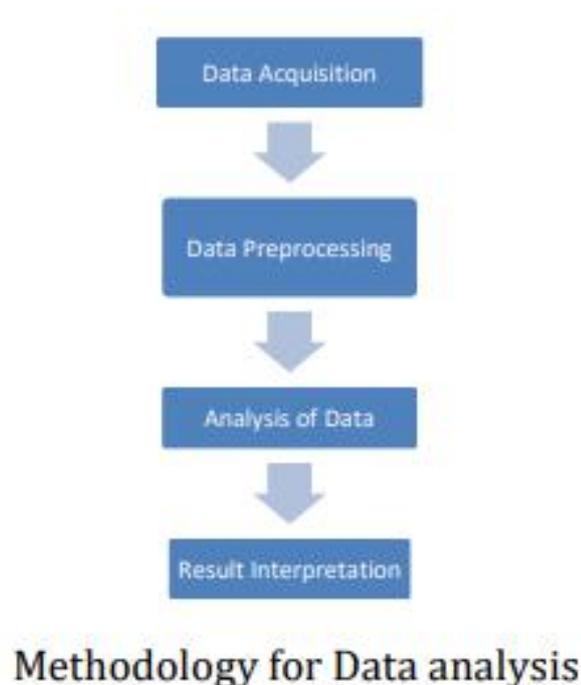


Figure 3.1:

- Data acquisition The data is acquired from data is collected from trak.in. This dataset contains all startups that got funding between the period January 2016 and April 2021. The file is present in a CSV (Comma Separated Value) format.

- Data preprocessing Data preprocessing consists of the following tasks:
  - o Data Cleaning
  - o Data Transformation
  - o Data Integration

- o Data reduction

Here data needs to be cleaned to avoid wrong conclusions. For this task, we have identified the null values in each column. They are as follows:

```
df.isnull().sum()
Sl.no          0
Date           0
Startup Name   0
Industry       0
Sub-vertical   0
Location       0
Investors      0
Investment Type 0
Amount in USD  712
Website URL    292
dtype: int64
```

Figure 3.2:

Since we can see website URL have the most number of null values, 292 null values, and the column is not so important for our study. Therefore we will drop the column since it would not provide any useful information during the analysis. The Amount in USD column also has 721 null values since it is important for our study we use groupby and sum to fill up the null values We would also need to fix some date formats, spelling types and combine alternate names of a company into a single name. Ex: 4/14/2021 to 14/04/2021, Seed Funding to Seed Funding, biju to BYJU'S, etc

### 3.2.2 K-Means Clustering

Machine learning algorithms can be broadly classified into two categories - supervised and unsupervised learning. There are other categories also like semi-supervised learning and reinforcement learning. But, most of the algorithms are classified as supervised or unsupervised learning. The difference between them happens because of presence of target variable. In unsupervised learning, there is no target variable. The dataset only has input variables which describe the data. This is called unsupervised learning.

K-Means clustering is the most popular unsupervised learning algorithm. It is used when we have unlabelled data which is data without defined categories or groups. The algorithm follows an easy or simple way to classify a given data set through a certain number of clusters, fixed apriori. K-Means algorithm works

iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity.

K-Means clustering can be represented diagrammatically as follows:-

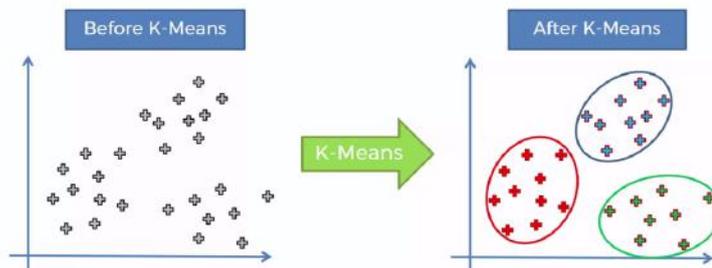


Figure 3.3:

### (i) K-Means Clustering intuition

K-Means clustering is used to find intrinsic groups within the unlabelled dataset and draw inferences from them. It is based on centroid-based clustering.

Centroid - A centroid is a data point at the centre of a cluster. In centroid-based clustering, clusters are represented by a centroid. It is an iterative algorithm in which the notion of similarity is derived by how close a data point is to the centroid of the cluster. K-Means clustering works as follows:- The K-Means clustering algorithm uses an iterative procedure to deliver a final result. The algorithm requires number of clusters K and the data set as input. The data set is a collection of features for each data point. The algorithm starts with initial estimates for the K centroids. The algorithm then iterates between two steps:-

1.Data assignment step Each centroid defines one of the clusters. In this step, each data point is assigned to its nearest centroid, which is based on the squared Euclidean distance. So, if  $c_i$  is the collection of centroids in set C, then each data point is assigned to a cluster based on minimum Euclidean distance.

$$Euclidean\ distance = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (3.1)$$

2.Centroid update step In this step, the centroids are recomputed and updated. This is done by taking the mean of all data points assigned to that centroid's cluster.

The algorithm then iterates between step 1 and step 2 until a stopping criteria is met.

The K-Means intuition can be represented with the help of following diagram:-

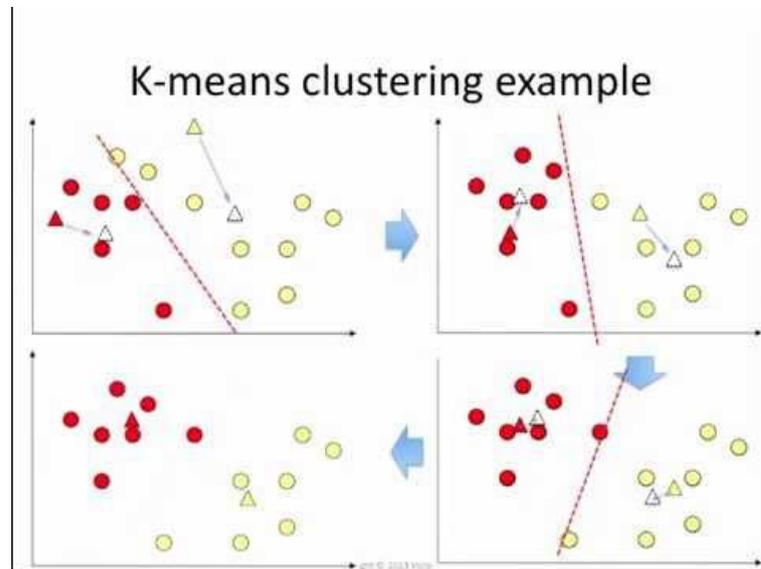


Figure 3.4:

## (ii) Choosing the value of K

The K-Means algorithm depends upon finding the number of clusters and data labels for a pre-defined value of K. To find the number of clusters in the data, we need to run the K-Means clustering algorithm for different values of K and compare the results. So, the performance of K-Means algorithm depends upon the value of K. We should choose the optimal value of K that gives us best performance. There are different techniques available to find the optimal value of K. The most common technique is the elbow method which is described below.

## (iii) The elbow method

The elbow method is used to determine the optimal number of clusters in K-means clustering. The elbow method plots the value of the cost function produced by different values of K. The below diagram shows how the elbow method works:-

We can see that if K increases, average distortion will decrease. Then each cluster will have fewer constituent instances, and the instances will be closer to

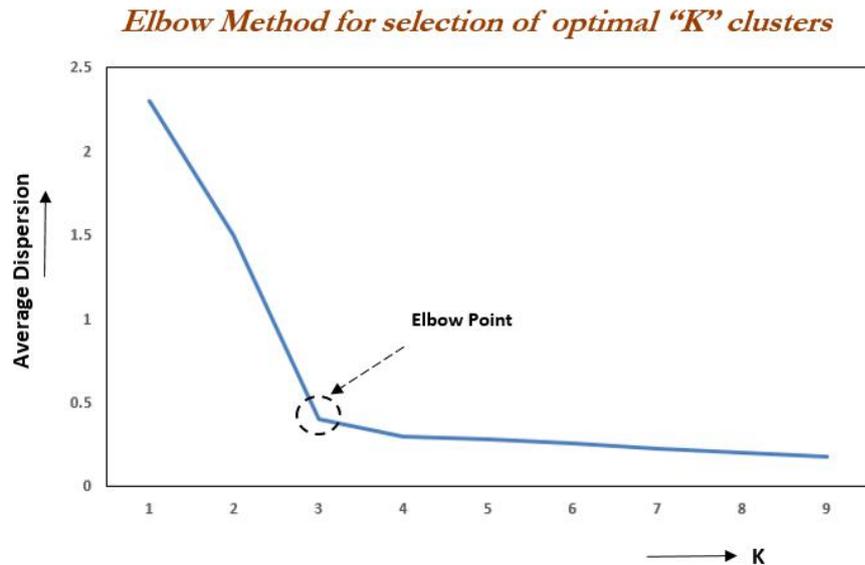


Figure 3.5:

their respective centroids. However, the improvements in average distortion will decline as K increases. The value of K at which improvement in distortion declines the most is called the elbow, at which we should stop dividing the data into further clusters.

### 3.3 Measure the accuracy or goodness of our clustering technique

#### 3.3.1 Silhouette Coefficient

Silhouette Coefficient or silhouette score is a metric used to calculate the goodness of a clustering technique. Its value ranges from -1 to 1.

1: Means clusters are well apart from each other and clearly distinguished.

0: Means clusters are indifferent, or we can say that the distance between clusters is not significant.

-1: Means clusters are assigned in the wrong way.

$$\text{SilhouetteScore} = (b - a) / \max(a, b) \quad (3.2)$$

where,

a= average intra-cluster distance i.e the average distance between each point within a cluster.

b= average inter-cluster distance i.e the average distance between all clusters.

# Chapter 4

## Result

---

### 4.1 How does funding ecosystem changes with time

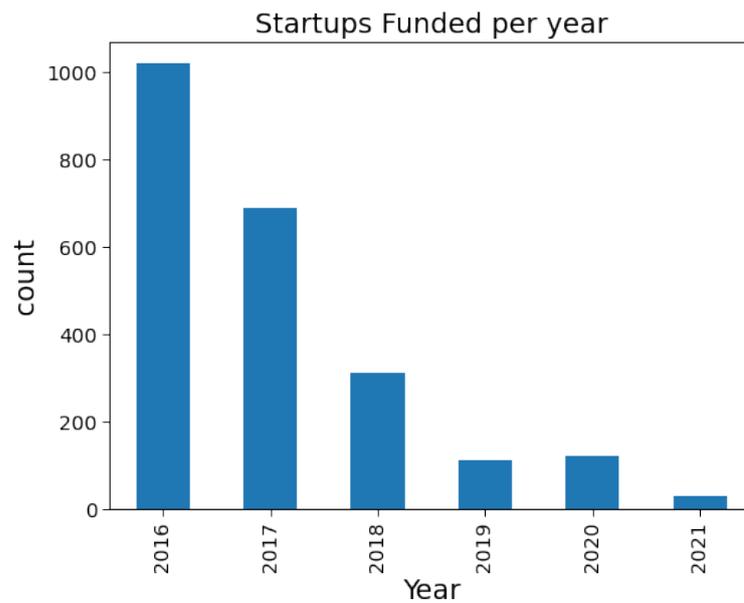


Figure 4.1:

From the above figure 4.1 it is clear that number of startups that have been funded every year for the last six years, with 2016 being the most and 2021 being the least. (Noted that 2021 is least because there are records only for the first month.

The figure 4.2 describe the Startups funded per month by year. The graph clearly pictures the funding count in each month per year.

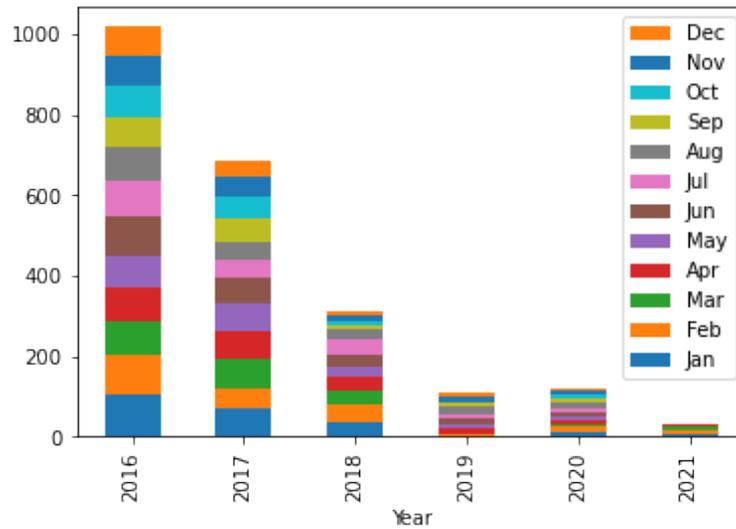


Figure 4.2:

#### 4.1.1 Startup Funding Variation Per Month from 2016-2020

##### Insights of 2016

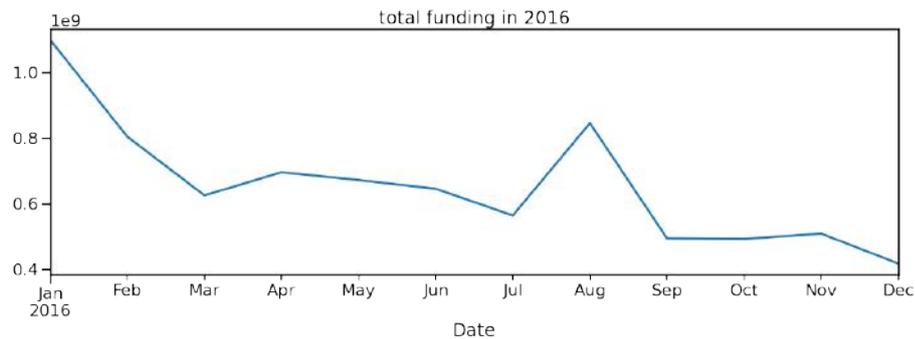
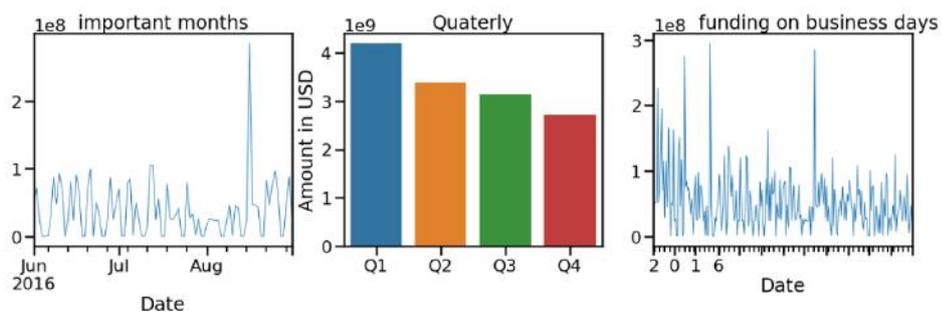


Figure 4.3:

From figure 4.3 it is clear that January and August was most funded. A sudden increase is witnessed in the month of July-August. Further after September there is a uniform low funding due to demonetisation of Indian currency.



Quarter one (Q1) has the most funding

## Insights of 2017

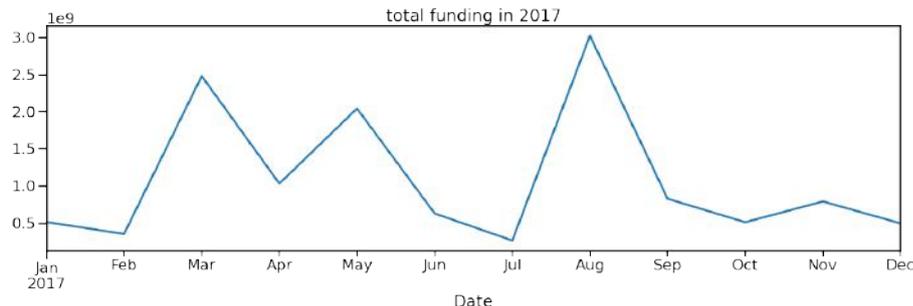
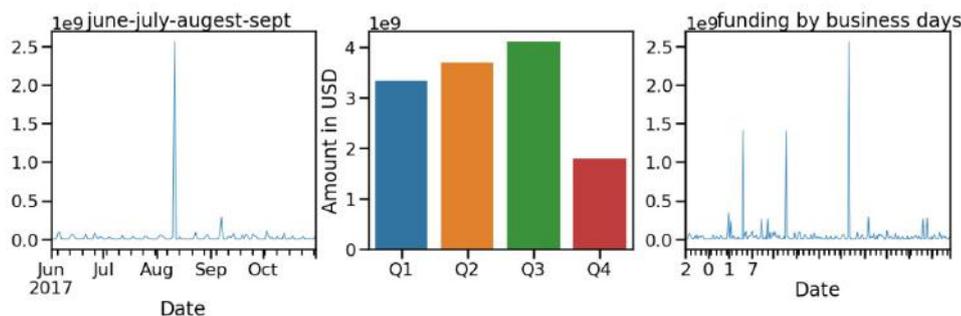


Figure 4.4:

From figure 4.4 it is clear that March, May and August was most funded. A sudden decrease is witnessed in the month of Jun-July. Most funding is done in the month of August.



Quarter three (Q3) has the most funding

## Insights of 2018

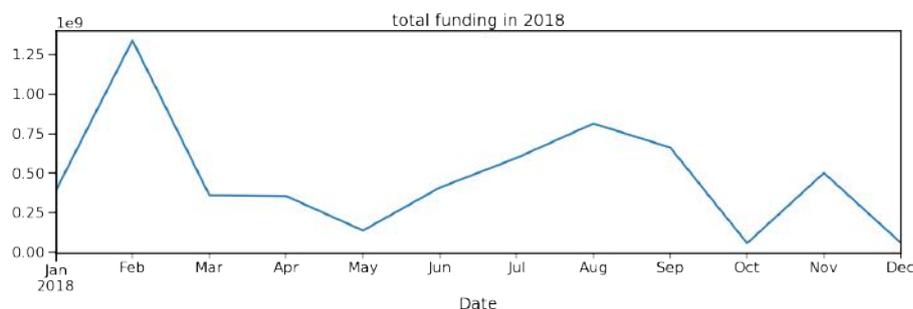
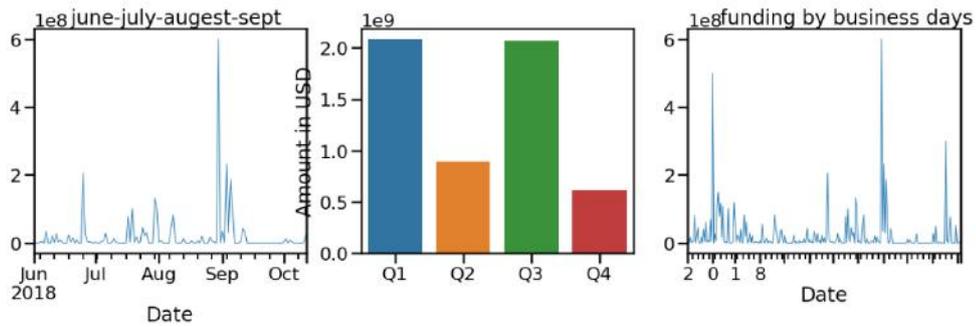


Figure 4.5:

From figure 4.5 it is clear that February and August was most funded. A sudden downfall is seen in May followed by a gradual increase. October has a sudden decrease followed by an increase in November.



Quarter one and three (Q1 and Q3) has the most funding

### Insights of 2019

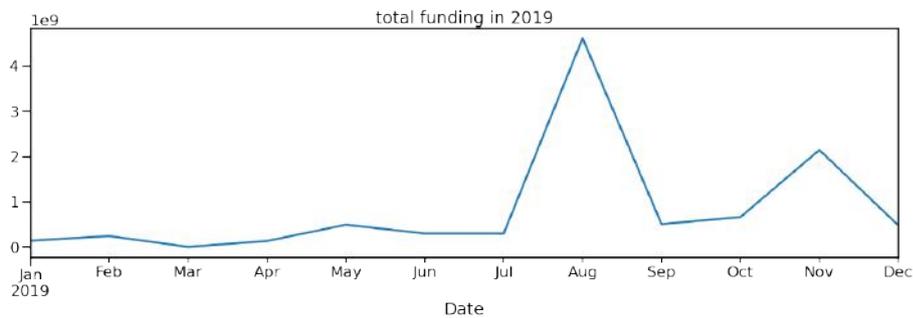
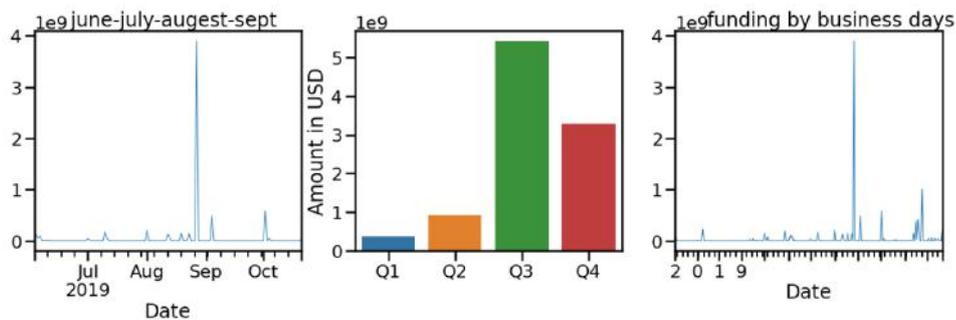


Figure 4.6:

From figure 4.6 it is clear that August and November was most funded. Funding is less in the month of January-July followed by sudden increase in August.



Quarter three (Q3) has the most funding

### Insights of 2020

From figure 4.7 it is clear that February, July and September was most funded. Further from October there is a uniform low funding due to Covid 19 outbreak in India.

4.2. DO CITIES PLAY A MAJOR ROLE IN FUNDING STARTUP ECOSYSTEM IN INDIA

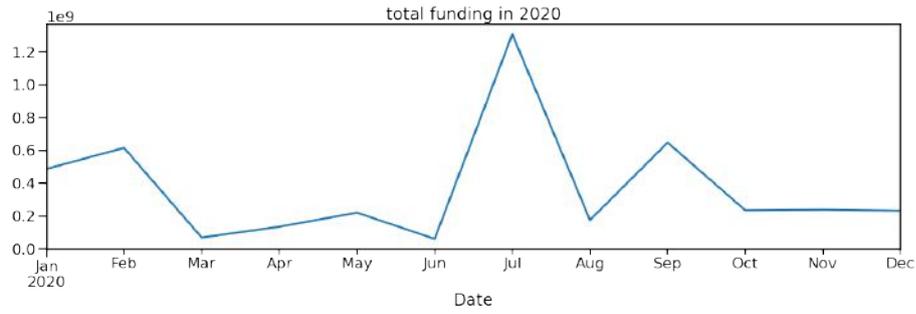
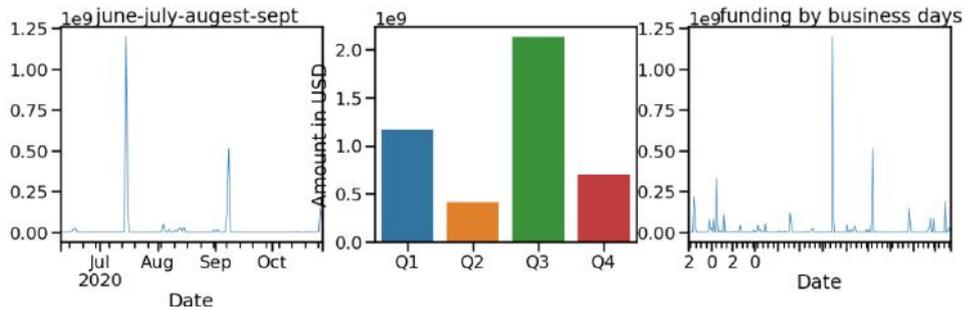


Figure 4.7:



Quarter three (Q3) has the most funding

4.2 Do cities play a major roll in funding?

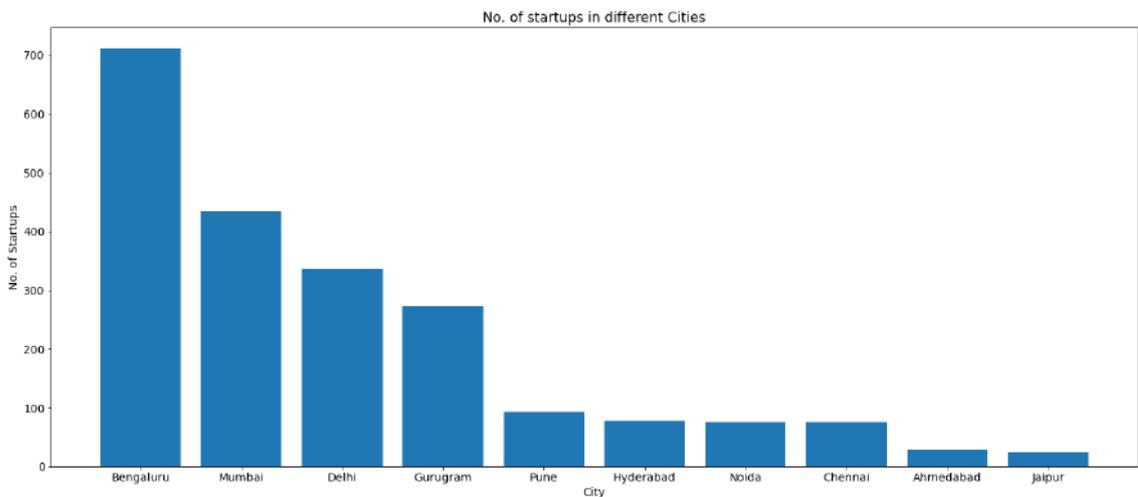


Figure 4.8:

Bangalore has the highest number of startups in the country followed by Mumbai. The figure 4.8 clearly tells us why, Bangalore is called the Silicon Valley states of India

### 4.3 Which industries are favoured by investors for funding

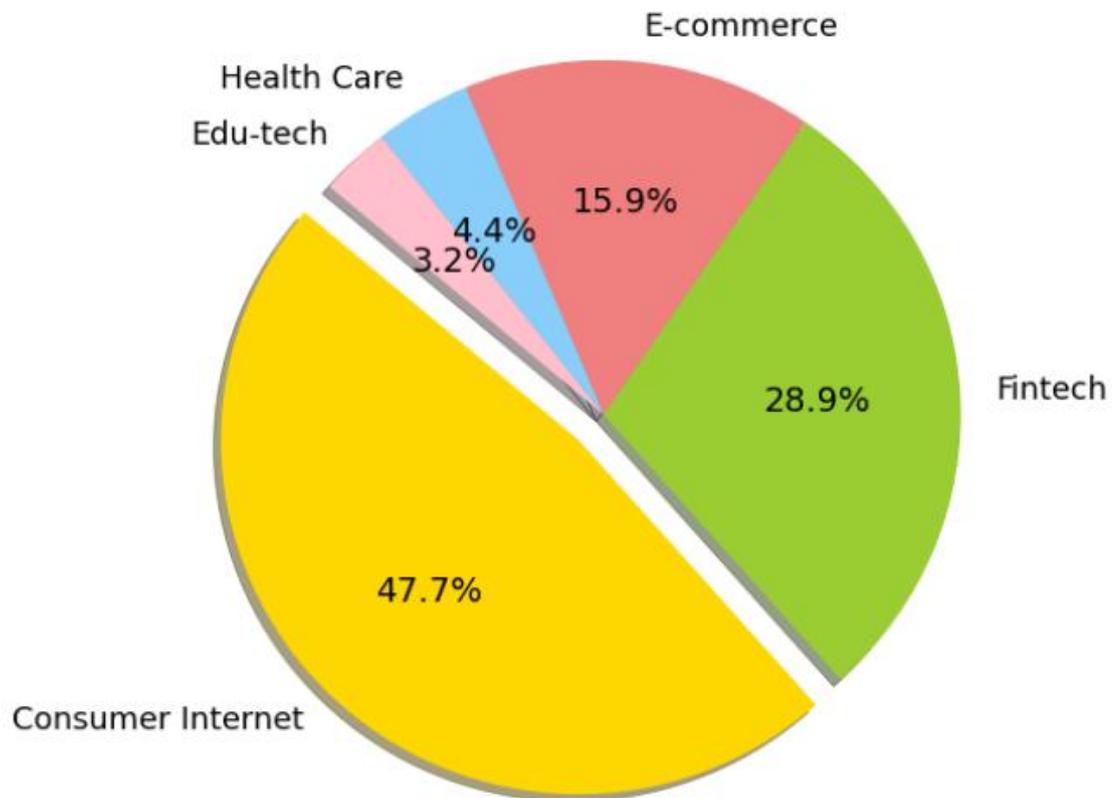


Figure 4.9:

The above pie chart (figure 4.9) describes that, consumer Internet, Fintech and E-commerce are the most popular industry for startups.

## 4.4 Who are the important investors in the Indian Ecosystem

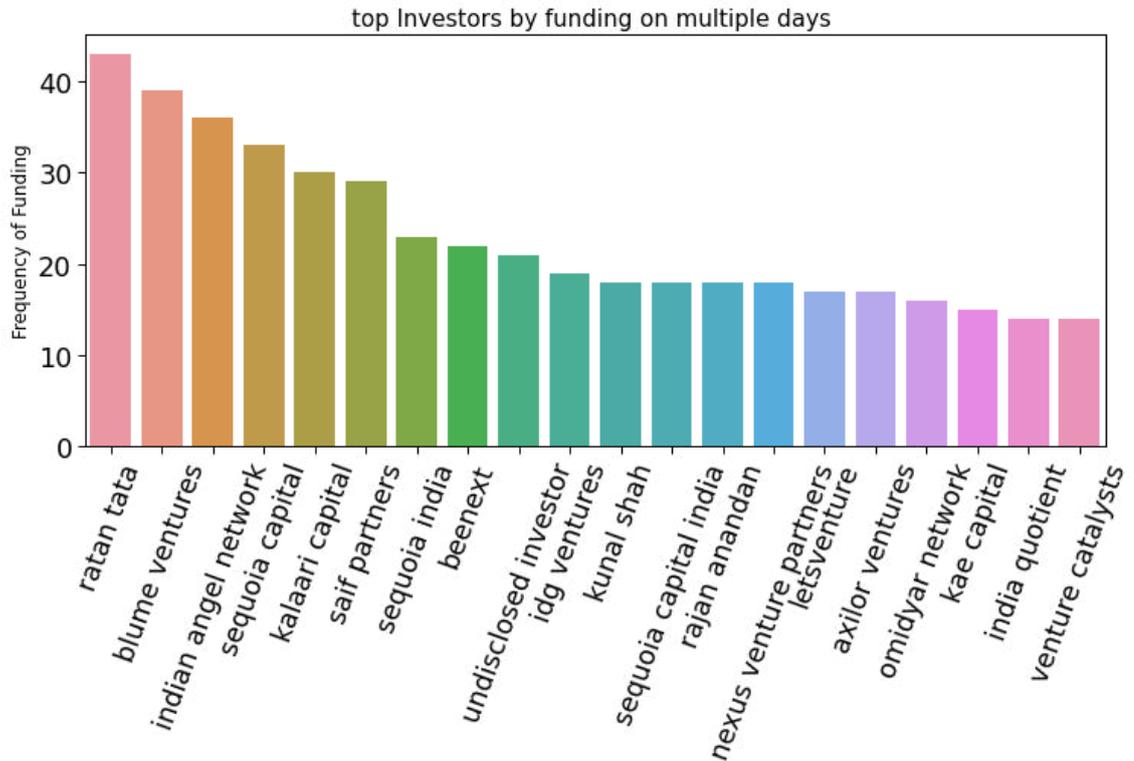


Figure 4.10:

Ratan Tata is the Investors with most funding frequency.

Table 4.1: Investor-Most funding frequency

Ratan Tata	43
Blume Ventures	39
Indian Angel Network	36
Sequoia Capital	33
Kalaari Capital	30
Saif Partners	29
Sequoia India	23
Beenext	22
Undisclosed Investor	21
Idg Ventures	19

Ratan Tata is an Indian industrialist, philanthropist, former chairman of Tata Group and Tata Sons. He is a prolific investor and has made numerous investment in many Startups. His style of investment and funding are revered by many across

the globe. And his investments are known to emerge as giants in their respective sectors with time. Ola Cabs is an example.

### Top 10 Most funded Investors

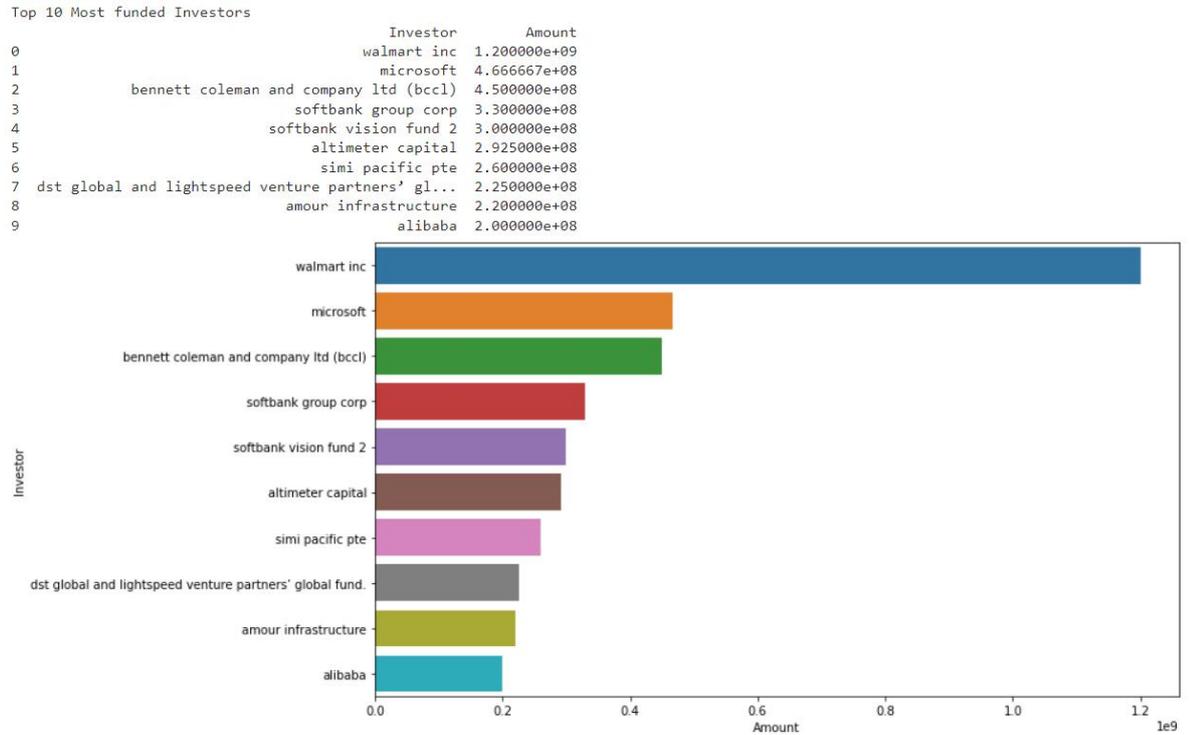


Figure 4.11:

Walmart is a multinational retail corporation that operates several chains of discount department and warehouse stores.

### Startup with most number of investors

It is clear from the histogram that BYJU'S has the most number of investors . BYJU'S is an Indian multinational educational technology company, headquartered in Bangalore.

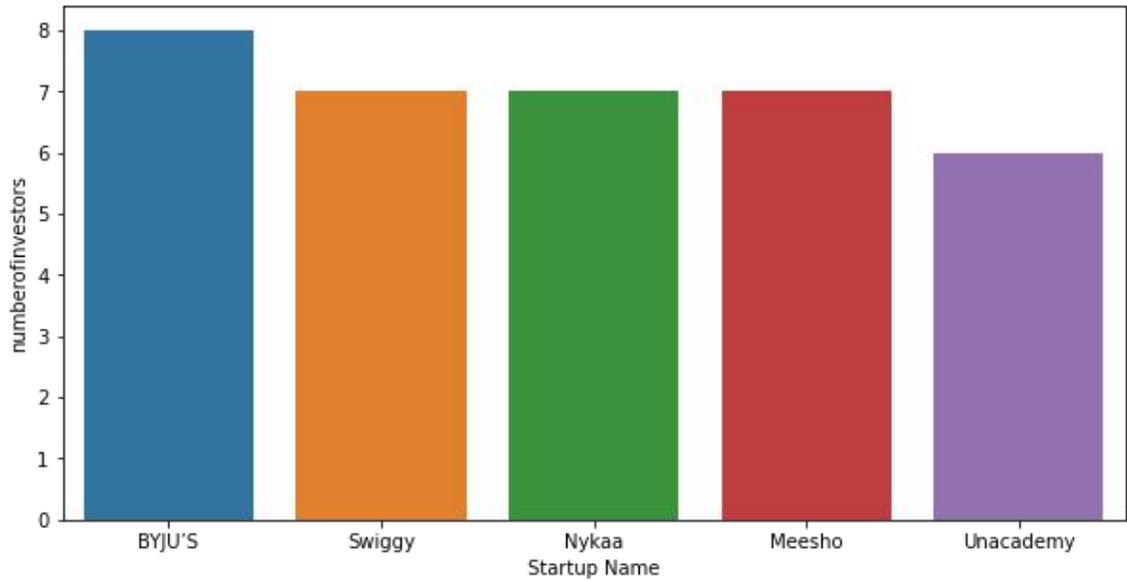


Figure 4.12:

### Most popular type of investment

It is evident from the figure 4.12 that Pre-series funding has been the most popular type of investment among investors. It is typically the time when the founders of the company first get their operations off the ground. The graph makes it clear that the investors want to invest in companies at the early stages of operation.

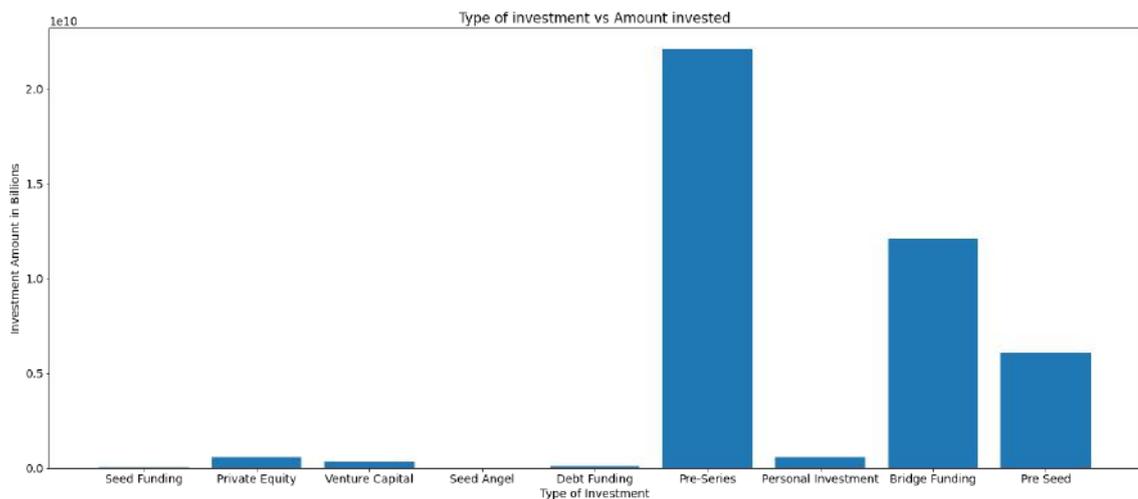


Figure 4.13:

## 4.5 How much funds does startups generally get in India

### Top 10 startups based on amount of funding

Flipkart and Paytm still attract investors due to its position in the market. Flipkart received the top funding in 2017 through private equity.

```

Startup Name
Flipkart          5.209700e+09
Rapido Bike Taxi  3.900000e+09
Paytm             2.468950e+09
BYJU'S           1.809313e+09
Udaan            8.700000e+08
Swiggy           6.230000e+08
Zomato           6.000000e+08
True North       6.000000e+08
GOQii            5.086254e+08
Meesho           4.938874e+08
Name: Amount in USD, dtype: float64

```

Figure 4.14:

### Total number of unique startups funded in each year

Table 4.2: Unique startups from 2016-2021

2021	28
2020	113
2019	106
2018	298
2017	640
2016	959

From Table 4.2 we can say that the year 2016 is the evolution of more unique or new startups in India.

## Total amount funded in each year

Table 4.3: Total Amount funded from 2016-2021

2020	4437057506
2019	3792351220
2018	6164240859
2017	11130104864
2016	13444780072

From Table 4.3 we can say that the year 2016 has got most funding amount in India.

## 4.6 Clustering

### K-Means

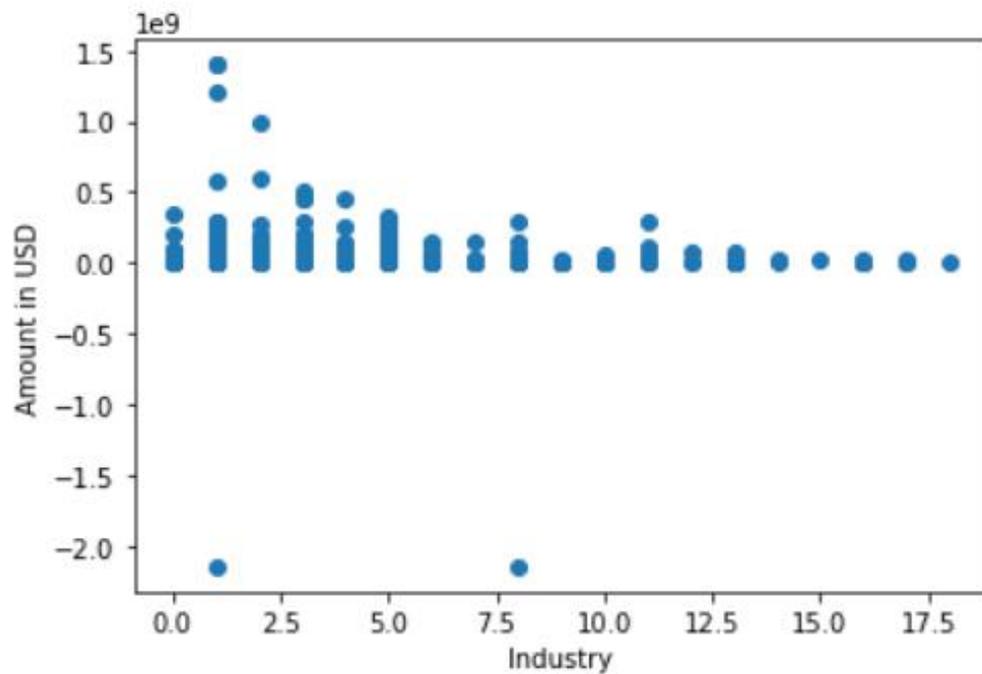


Figure 4.15:

The figure 4.15 shows the initial data set.

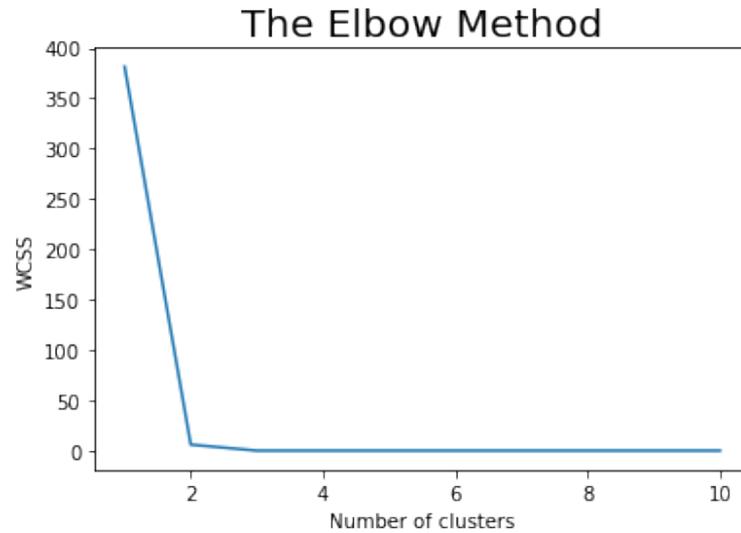


Figure 4.16:

From the above figure 4.16 the number of clusters is determined as three. The

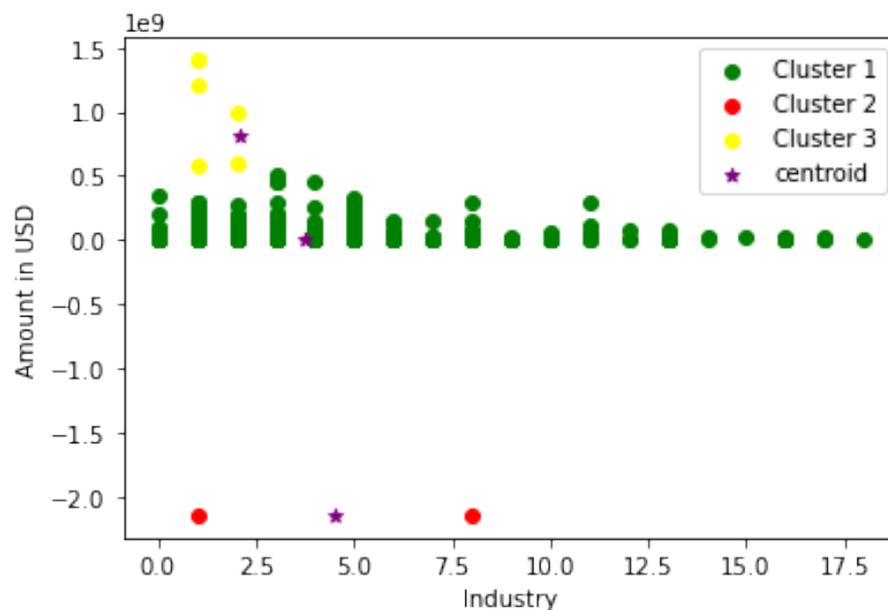


Figure 4.17:

Figure 4.17 shows three clusters.

we need to find what kind of industry different investors are looking forward to invest.

The clusters tells us that investors are interested in cluster 1 which contains Industrys like E-commerce,consumer internrt,edutech,fintech etc.They show less interest in cluster 2 that contains BFSI and others.cluster3 is also preferred by investors but not that much as cluster1.

The investors are not ready to invest a huge amount in a firm they are most interested in Pre-series funding it has been the most popular type of investment among investors. It is typically the time when the founders of the company first get their operations off the ground. So that the investors will not lose a huge amount if the firm fails to establish in the market.

#### **4.7 Measure the accuracy of the cluster**

By using Silhouette Coefficient our cluster has an accuracy of 0.9702154047873705. We can say that the clusters are well apart from each other as the silhouette score is closer to 1

## Chapter 5

# Discussion and Conclusion

---

The conclusion obtained from the analysis and visualization for the ‘Indian Startup Funding’ data is that trends of how the events like Startup India, Indian money demonetization, Covid 19 has influenced the startup ecosystem in recent years. The Indian Startup Ecosystem is changing with time it is clearly visible from the analysis that 2016 being most funded followed by a sudden downfall due to socio-economic factors. It is observed that Bangalore, Mumbai, and Delhi leads the way for maximum investments and is hub to huge number of startups

Unicorns like Flipkart and Paytm still attract investors due to its position in the market. And after demonization Paytm adopted the cashless India policy to boost their revenue.

BYJU’S attracted the most investors in the past years and is planning to join the unicorn clubs. Still today Consumer internet is more favored to investors than any other industry verticals, but to the due advancement of AI, Machine learning, and other advancements in technology, the technology vertical has also attracted a large number of investors. With consumer internet leading the way, its sub-verticals like online pharmacy, online lending platform, online payment gateway, etc have also attracted the most number of investors.

Since India has become a hub for startups, new startup emerges, therefore, Pre-series is important for them. Pre-series is typically the time when the founders of the company first get their operations off the ground. The analysis makes it clear that the investors want to invest in companies at the early stages of operation. It is the most preferred type of investment

An interesting point to take is that individuals like Ratan Tata have taken a keen interest in investing in startups since they invested more than any other individuals.

The investors are not ready to invest a huge amount in a firm they are most interested in Pre-series funding it has been the most popular type of investment among investors. It is typically the time when the founders of the company first get their operations off the ground. So that the investors will not lose a huge amount if the firm fails to establish in the market. They mostly prefer firms like E-commerce, consumer internet, Edu-Tech etc.

With the visualizations, we came to know that various events do affect the funding. Also the cities and industry verticals play an important part in acquiring funding from investors.

## 5.1 Python Code

```

import numpy as np
import pandas as pd
import os
%matplotlib inline

import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

df=pd.read_csv('indian_startup_funding_koop.csv')
df.head()

```

Sl.no	Date	Startup Name	Industry	Sub-vertical	Location	Investors	Investment Type	Amount in USD	Website URL
0	1	4/14/2021	Swiggy	Online Food Delivery	Bengaluru	Amarca Holdings, Carmignac, Falcon Edge Capita...	Series J	343000000	https://www.swiggy.com/
1	2	4/14/2021	Beldara	E-commerce	Mumbai	Hindustan Media Ventures	Venture	7400000	https://beldara.com/
2	3	4/7/2021	Groww	FinTech	Bengaluru	MC Global Edtech, B Capital, Baron, others	Series D	83000000	https://groww.in/
3	4	4/5/2021	Meesho	E-commerce	Bengaluru	SoftBank Vision Fund 2	Series E	300000000	http://www.meesho.com/
4	5	4/1/2021	BYJU'S	Edu.tech	Bengaluru	Innov8 Capital	Series F	460000000	http://www.byjus.com/

Figure 5.1:

```

print("Information of total number of non-empty columns")
print("-----")
print(df.info(null_counts=True))

```

Information of total number of non-empty columns

```

-----
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2276 entries, 0 to 2275
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Sl.no                 2276 non-null   int64
1   Date                 2276 non-null   datetime64[ns]
2   Startup Name         2276 non-null   object
3   Industry             2276 non-null   object
4   Sub-vertical         2276 non-null   object
5   Location             2276 non-null   object
6   Investors            2276 non-null   object
7   Investment Type      2276 non-null   object
8   Amount in USD       1564 non-null   float64
9   Website URL         1984 non-null   object
dtypes: datetime64[ns](1), float64(1), int64(1), object(7)
memory usage: 177.9+ KB
None

```

Figure 5.2:

```

df.groupby(['Investment Type'])['Amount in USD'].mean()

df['Industry'].replace(['Food and Beverage', 'E-commerce', 'Fintech', 'Edu-Tech', 'Health Care', 'Consumer Internet', 'SaaS', 'Agritech',
                        [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18], inplace=True)

plt.scatter(df['Industry'], df['Amount in USD'])
plt.xlabel('Industry')
plt.ylabel('Amount in USD')

from sklearn.cluster import KMeans
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from matplotlib import pyplot as plt
%matplotlib inline

km = KMeans(n_clusters=3)
y_predicted = km.fit_predict(df[['Industry', 'Amount in USD']])
y_predicted

df['cluster'] = y_predicted
df

```

Figure 5.3:

```

from sklearn.cluster import KMeans
wcss = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters = i, init = 'k-means++', random_state = 0)
    kmeans.fit(x)
    wcss.append(kmeans.inertia_)
plt.plot(range(1, 11), wcss)
plt.title('The Elbow Method', fontsize = 20)
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()

df1 = df[df.cluster==0]
df2 = df[df.cluster==1]
df3 = df[df.cluster==2]
plt.scatter(df1.Industry, df1['Amount in USD'], color='green', label = 'cluster 1')
plt.scatter(df2.Industry, df2['Amount in USD'], color='red', label = 'cluster 2')
plt.scatter(df3.Industry, df3['Amount in USD'], color='yellow', label = 'cluster 3')
plt.scatter(km.cluster_centers_[:,0], km.cluster_centers_[:,1], color='purple', markers='*', label='centroid')
plt.xlabel('Industry')
plt.ylabel('Amount in USD')
plt.legend()

from sklearn.metrics import silhouette_score
score = silhouette_score(df, y_predicted, metric='euclidean')
score

print('Silhouette Score: %.3f' % score)

```

Figure 5.4:

## REFERENCES

---

- [1] Ahmad, N. (2019). FUNDING TRENDS FOR STARTUPS IN INDIA. 6(6), 13. Chandio, D. S. (2016). India the world's fastest growing startup ecosystem: A Study. 01(02), 10.
- [2] David, D., Gopalan, S., Ramachandran, S. (n.d.). The Startup Environment and Funding Activity in India. 31.
- [3] Garg, A., Shivam, A. (2017, January 1). Funding to growing startups. Jain, S. (n.d.). Growth of startup ecosystems in India. 4.
- [4] Garg, A., Shivam, A. (2017, January 1). Funding to growing startups. Jain, S. (n.d.). Growth of startup ecosystems in India. 4.
- [5] Kamaluddin, F. A., Sridhar, K. S. (n.d.). Indian Startup Ecosystem: Analysing Investment Concentration and Performance of Government Programmes. 23.
- [6] Korreck, S. (n.d.). The Indian Startup Ecosystem: Drivers, Challenges and Pillars of Support. 44.
- [7] Kumar, M., Narayan, D. (2019). Growth Pattern and Trends in Startup Funding in INDIA. 8, 3721–3724. <https://doi.org/10.35940/ijitee.L2654.1081219>
- [8] Pandey, N. K. (2018). An analysis of startup ecosystem in metropolitan city in India. International Journal of Engineering and Management Research (IJEMR), 8(2), 237–244.
- [9] Paramasivan, D. C. (2016). EMERGING TRENDS IN NEW START-UP TECHNOPRENEURS. Journal Of Business Management, 2(7), 8.
- [10] Prajapati, S., Alam, D. S. (n.d.). START-UP POLICY AND ITS IMPACT ON INDIAN ECONOMY. 1, 2.
- [11] Singh, S., Chauhan, A., Dhir, S. (2019). Analyzing the startup ecosystem of India: A Twitter analytics perspective. Journal of Advances in Management Research, 17(2), 262–281. <https://doi.org/10.1108/JAMR-08-2019-0164>

- [12] Verma, P. A., Singhal, V. (2018). Data Analysis of startups investments and funding trends in India. 05(11), 6.
- [13] Hari, T. N. (2018). The Indian Startup Report 2018. Bengaluru: YourStory Media Pvt Ltd. 2.
- [14] Huang, Y. (2008). The next Asian miracle. Foreign Policy , 32-40. 3.
- [15] Giurca Vasilescu, L. (2009). Business angels: potential financial engines for startups, *Ekonomiska istraživanja*, 22(3), 86-97.
- [16] Keshamoni, S. (2018, 05 30). 4 Start-Ups funded at TiE Startup Heroes challenge. Retrieved 05 30, 2018, from <http://startuphyderabad.com/4-startups-funded-at-tie-startup-heroes-challenge/>
- [17] Kshetri, N. (2016). Fostering Startup Ecosystems in India. *Asian Research Policy* , 94-103.
- [18] Nasscom. (2018, 05 29). NASSCOM START-UP REPORT' 2017. Retrieved 05 29, 2018, from <https://www.investindia.gov.in/sites/default/files/2018-03/Indian%20Startup%20Ecosystem%20-%20Traversing%20the%20maturity%20cycle.pdf>
- [19] O'Hear, S. (2013, 05 26). Anatomy Of An Undisclosed Investment Or Exit. Retrieved 08 26, 2018, from Techcrunch.com: <https://techcrunch.com/2013/05/26/follow-themoney/>
- [20] Velayanikal, M. (2015, 10 23). What are the hottest sectors for venture capital in India this year? Retrieved from Techinasia.com: <https://www.techinasia.com/hottest-sectorsventure-capital-india-infographic>
- [21] <https://trak.in/india-startup-funding-investment-2015/>
- [22] Inc42 (2017). Indian Tech Start-up Funding Report. India: Inc42. [Accessed 26 April 2018] 17 (2020). The State of Indian Startup Ecosystem Report. India: Inc42. [Accessed 3 January 2020]
- [23] <https://www.kaggle.com/code/prashant111/k-means-clustering-with-python#K-Means-Clustering-with-Python>

[24] Wikipedia. (2018,0813). India Startup Ecosystem TimeLine. Retrieved 08 27, 2018, from Wikipedia [https://en.wikipedia.org/wiki/India\\_startup\\_Ecosystem\\_TimeLine](https://en.wikipedia.org/wiki/India_startup_Ecosystem_TimeLine)

Project Report

On

**CORONAVIRUS (COVID-19) IN-DEPTH  
DATA ANALYSIS**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**MASTER OF SCIENCE**

*in*

**APPLIED STATISTICS AND DATA ANALYTICS**

*by*

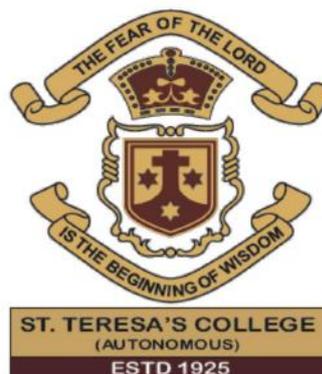
**TANIYA ROSE D'COUTH**

**(Register No. SM20AS023)**

**(2020-2022)**

*Under the Supervision of*

**SREELAKSHMI M S**



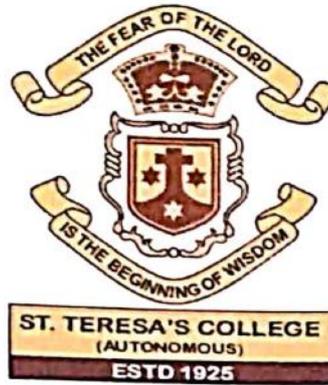
**DEPARTMENT OF MATHEMATICS AND STATISTICS**

**ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

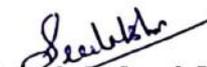
ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

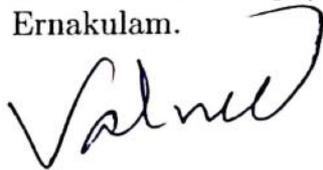


CERTIFICATE

This is to certify that the dissertation entitled, **CORONAVIRUS (COVID 19) IN-DEPTH DATA ANALYSIS** is a bonafide record of the work done by Ms. **TANIYA ROSE D'COUTH** under my guidance as partial fulfillment of the award of the degree of **Master of Science in Applied Statistics And Data Analytics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

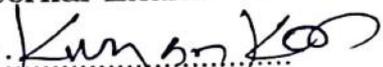
Date: 26/5/22  
Place: Ernakulam

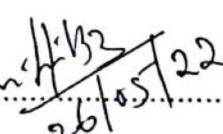
  
**Ms. Sreelakshmi M S**  
Assistant Professor,  
Department of Mathematics And Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.



  
**MS. Shanty P.B**  
Assistant Professor and HOD,  
Department of Mathematics And Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.

External Examiners

1: 

2:   
26/05/22

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of MS.SREELAKSHMI M S, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 09/05/2022



TANIYA ROSE D'COUTH

SM20AS023

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to GOD ALMIGHTY, for his constant love and grace that he has showered upon me to undertake and complete this project.

I am very grateful to my project guide Ms.Sreelakshmi M S for her support throughout my project

I would like to express my special thanks of gratitude to the Head of our department Ms.SHANTY P.B for her suggestions and feedback for the completion of my project.

I acknowledge with thanks to all the teachers of our department, my parents and friends for providing me with the necessities for the project work.

Ernakulam:

Date: 09/05/2022

*Taniya Rose D'Couth*

TANIYA ROSE D'COUTH

SM20AS023

# ABSTRACT

The COVID-19 pandemic is the defining global health crisis of our time and the greatest global humanitarian challenge the world has faced since World War II. The virus has spread widely, and the number of cases is rising daily as governments work to slow its spread. India has moved quickly, implementing a proactive, nationwide, lockdown, with the goal of flattening the curve and using the time to plan and resource responses adequately. This study will give us an idea about the impact of coronavirus globally in terms of confirmed cases, deaths reported, the number of recoveries as well as the active cases also about how India has been affected since the pandemic started in dive into the different states and union territories to learn more about the covid-19 influence and the vaccination status also it deals about the analysis of covid cases in which we do an exploratory data analysis (EDA) and to find the month in which the corona cases had risen to its peak..

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENTS</i> . . . . .	iv
<i>ABSTRACT</i> . . . . .	v
<i>CONTENT</i> . . . . .	vi
<b>1 INTRODUCTION</b>	<b>1</b>
<b>2 LITERATURE REVIEW</b>	<b>3</b>
<b>3 DATA DESCRIPTION AND DATA PREPROCESSING</b>	<b>6</b>
3.1 DATA DESCRIPTION . . . . .	6
3.2 COVID-19 INDIA DATA . . . . .	6
3.3 COVID VACCINE STATEWISE . . . . .	7
<b>4 EXPLORATORY DATA ANALYSIS</b>	<b>9</b>
4.1 ANALYSING COVID-19 CASES IN INDIA . . . . .	9
4.2 WEEKLY CASES . . . . .	10
4.3 MONTHLY CASES . . . . .	11
4.4 COVID CASES 200 DAYS . . . . .	11
4.5 INDIA AGE GROUP WISE DISTRIBUTION . . . . .	12
4.6 PERCENTAGE OF GENDER . . . . .	13
4.7 CRITICALITY OF LOCKDOWN AND CONTAINMENT FOR INDIA . . . . .	14
4.8 5 DAYS MOVING AVERAGE OF CONFIRMED CASES IN TOP 15 STATES . . . . .	15
4.9 TOP 10 STATES IN EACH HEALTH FACILITY . . . . .	16

4.10	URBAN AND RURAL HEALTH FACILITY . . . . .	17
4.11	TESTING STATEWISE INSIGHT . . . . .	18
4.12	ICMR TESTING CENTERS IN EACH STATE . . . . .	19
4.13	ECONOMIC IMPACT OF COVID 19 ON INDIA . . . . .	20
4.13.1	THREE ECONOMIC SCENARIOS MODEL IN- DIA GDP ESTIMATES . . . . .	20
4.14	MOST IMPACTED STATES/UT DURING COVID 19 LOCKDOWN . . . . .	21
<b>5</b>	<b>PREDICTION AND FORECASTING</b>	<b>23</b>
5.1	TIME SERIES FORECASTING: . . . . .	23
5.2	FORECASTING USING LSTM MODEL: . . . . .	23
5.3	MAKING LSTM PREDICTIONS . . . . .	24
5.4	DECOMPOSING THE TIME SERIES. . . . .	25
5.5	PLOTTING AUTO-CORRELATION FUNCTION . . . . .	26
5.6	DICKEY FULLER TEST (ADF TEST) . . . . .	26
5.7	DIVIDING TIME SERIES INTO TRAIN AND TEST FOR PREDICTIONS . . . . .	26
<b>6</b>	<b>CONCLUSION</b>	<b>29</b>
<b>7</b>	<b>REFERENCE</b>	<b>31</b>

# Chapter 1

## INTRODUCTION

---

The new coronavirus strain (COVID-19), emerged in Wuhan, China. On 2/11/2020, the World Health Organization designated the name COVID-19 for the disease. This may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). As of May 8th, 2020, in India, 56,342 positive cases have been reported. India, with a population of more than 1.34 billion—the second largest population in the world—will have difficulty in controlling the transmission of severe acute respiratory syndrome coronavirus 2 among its population. Multiple strategies would be highly necessary to handle the current outbreak; these include computational modeling, statistical tools, and quantitative analyses to control the spread as well as the rapid development of a new treatment. The Ministry of Health and Family Welfare of India has raised awareness about the recent outbreak and has taken necessary actions to control the spread of COVID-19. The central and state governments are taking several measures and formulating several wartime protocols to achieve this goal. Moreover, the Indian government implemented a 55-days lockdown throughout the country that started on March 25th, 2020, to reduce the transmission of the virus. This outbreak is inextricably linked to the economy of the nation, as it has dramatically impeded industrial sectors because people worldwide are currently cau-

tious about engaging in business in the affected regions. COVID-19 has turned the world upside down. Everything has been impacted. How we live and interact with each other, how we work and communicate, how we move around and travel. Every aspect of our lives has been affected. Although the world is in lockdown, governments, epidemiologists, school principals, entrepreneurs, and families around the world are already planning the next steps: how to safely reopen schools and businesses, how to commute and travel without transmitting or contracting infection, how to support those most affected by the crisis – the millions who have lost their livelihoods or their loved ones, how to ensure the already serious inequalities don't deteriorate further. Most people who fall sick with COVID-19 will experience mild to moderate symptoms and recover without special treatment. However, some will become seriously ill and require medical attention. The virus can spread from an infected person's mouth or nose in small liquid particles when they cough, sneeze, speak, sing or breathe. These particles range from larger respiratory droplets to smaller aerosols. You can be infected by breathing in the virus if you are near someone who has COVID-19, or by touching a contaminated surface and then your eyes, nose, or mouth. The virus spreads more easily indoors and in crowded settings. An easy way to decrease SARS-CoV-2 infection rates is to avoid virus exposure. People from India should avoid traveling to countries highly affected by the virus, practice proper hygiene, and avoid consuming food that is not home-cooked. Necessary preventive measures, such as wearing a mask, regular hand washing, and avoiding direct contact with infected persons, should also be practiced.

## Chapter 2

# LITERATURE REVIEW

---

Vaccine Effectiveness against Omicron in Children and Adolescents by A.M. Price and Others. In this study evaluating BNT162b2, vaccine effectiveness against hospitalization for Covid-19 in the delta-predominant period among adolescents 12 to 18 years of age was more than 90%. Post-COVID-19 global health strategies: the need for an interdisciplinary approach by Agostino Gemelli” IRCCS. This article aims to describe the importance of the interdisciplinary approach—coordinated by geriatricians—to cope with the potential post-acute care needs of recovered COVID-19 patients. Current Situation of Coronavirus Disease by Albaraa A Milibari, this article discussing the pandemic situation in 2020 says that All ages are at risk of getting the illness. This is because the ailment is transmitted through large droplets that result from coughing and sneezing by symptomatic individuals. In some instances, the infection can happen in asymptomatic individuals before the beginning of symptoms. This paper provides in-depth information on COVID-19 as it discusses the disease epidemiology, transmission, clinical features, diagnosis, treatment, and prevention.

Novel coronavirus disease (COVID-19) pandemic by Muhammad Fayyaz ur Rehman, this article says about the current status of the COVID-19, epidemiology, an overview of phylogeny, mode of action, diagnosis, and possible treatment methods and vaccines, and briefly highlights the history, phylogeny, genomics, epidemiology, mode of action, disease symptoms, diagnosis, and possible treatment methods of COVID-19 and the

research progress in the development of vaccines against SARS-Cov-2. Global impact of Covid-19 Pandemic by Prof. Dr. İsmail Hakki. This report provides a critical assessment of the global situation of the COVID-19 pandemic. The statistical measures such as regression, variance, distributions, t-test, and ANOVA test have been conducted to understand and explore the ongoing situations and to predict the forecoming scenario, rate of growth, and possible risk assessment of the transmission

District level correlates of COVID-19 pandemic in India during March-October 2020 by Vandana Tamrakar. COVID-19 is affecting the entire population of India. Understanding district-level correlates of the COVID-19 infection ratio (IR) is essential for formulating policies and interventions. The present study aims to investigate the district-level variation in COVID-19 from March-October 2020. The present study also examines the association between India's socioeconomic and demographic characteristics and the COVID-19 infection ratio at the district level. We identified hotspot and cold spot districts for COVID-19 cases and infection ratio. We have also carried out two sets of regression analyses to highlight the district level demographic, socioeconomic, household infrastructure facilities, and health-related correlates of the COVID-19 infection ratio. The situation of India in the COVID-19 Pandemic: India's Initial Pandemic Experience by Azizah F. Siddiqui. In this article, the author discussed the impact of COVID-19 through screening and surveillance methods adopted in India, as well as the potential health system, and social, political, and economic consequences. Evaluation and prediction of COVID-19 in India: A case study of worst-hit states by Danish Rafiq. In this manuscript, system modeling and identification techniques are applied in developing a prognostic yet deterministic model to forecast the spread of COVID-19 in India. The model is verified with the historical data and a forecast of the spread for 30-days is presented in the 10 most-affected states of India. The major results suggest that our model can very well capture the disease variations with high accuracy. The results also show a steep rise in the

**total cumulative cases and deaths in the coming weeks**

## Chapter 3

# DATA DESCRIPTION AND DATA PREPROCESSING

---

### 3.1 DATA DESCRIPTION

The data used for the study are secondary. The data is collected from Kaggle. Two datasets are used for the study

### 3.2 COVID-19 INDIA DATA

The data has observations from 30-01-2020 - to 11-08-2021. There are 18110 observations and 9 columns. This dataset has daily level information on the number of affected cases, deaths, and recovery from the 2019 novel coronavirus. This is time-series data and so the number of cases on any given day is the cumulative number. we have 2 columns named Confirmed Indian Nationals and Confirmed Foreign Nationals, that we won't be using. So dropping these two columns.

Date = Date which the coronavirus cases are observed. Time = A particular time where data is collected. State/UnionTerritories = State which covid cases reported. Cured = Cumulative number of cured cases. Death = Cumulative number of Deaths. Confirmed = Cumulative number of Confirmed cases.

3.3. COVID VACCINE STATEWISE

Sno	Date	Time	State/Unic	Confirmed	Confirmed	Cured	Deaths	Confirmed
1	#####	6:00 PM	Kerala	1	0	0	0	1
2	#####	6:00 PM	Kerala	1	0	0	0	1
3	#####	6:00 PM	Kerala	2	0	0	0	2
4	#####	6:00 PM	Kerala	3	0	0	0	3
5	#####	6:00 PM	Kerala	3	0	0	0	3
6	#####	6:00 PM	Kerala	3	0	0	0	3
7	#####	6:00 PM	Kerala	3	0	0	0	3
8	#####	6:00 PM	Kerala	3	0	0	0	3
9	#####	6:00 PM	Kerala	3	0	0	0	3
10	#####	6:00 PM	Kerala	3	0	0	0	3
11	#####	6:00 PM	Kerala	3	0	0	0	3
12	#####	6:00 PM	Kerala	3	0	0	0	3
13	#####	6:00 PM	Kerala	3	0	0	0	3
14	#####	6:00 PM	Kerala	3	0	0	0	3
15	#####	6:00 PM	Kerala	3	0	0	0	3
16	#####	6:00 PM	Kerala	3	0	0	0	3
17	#####	6:00 PM	Kerala	3	0	0	0	3
18	#####	6:00 PM	Kerala	3	0	0	0	3
19	#####	6:00 PM	Kerala	3	0	0	0	3
20	#####	6:00 PM	Kerala	3	0	0	0	3
21	#####	6:00 PM	Kerala	3	0	0	0	3
22	#####	6:00 PM	Kerala	3	0	0	0	3
23	#####	6:00 PM	Kerala	3	0	0	0	3
24	#####	6:00 PM	Kerala	3	0	0	0	3
25	#####	6:00 PM	Kerala	3	0	0	0	3
26	#####	6:00 PM	Kerala	3	0	0	0	3
27	#####	6:00 PM	Kerala	3	0	0	0	3

Figure 3.1: COVID-19 India Data set

3.3 COVID VACCINE STATEWISE

The data has observations from 16/01/2021 - to 15/08/2021. There are 7845 rows and 24 columns. This dataset has information on Total Doses Administered, Sessions, Sites, First Dose Administered, Second Dose Administered, Male (Dose Administered), Female (Dose Administered), Transgender (Dose Administered), Covaxin (Dose Administered), Covishield (Dose Administered), Sputnik V, AEFI, 18-44 Years (Dose Administered), 45-60 Years (Dose Administered), 60+ Years (Dose Administered), 18-44 Years (Individuals Vaccinated), 45-60 Years (Individuals Vaccinated), 60+ Years (Individuals Vaccinated), Male

3.3. COVID VACCINE STATEWISE

(Individuals Vaccinated), Female (Individuals Vaccinated), Transgender (Individuals Vaccinated), Total Individuals Vaccinated. The variables used in the dataset are; Vaccinated on: Vaccinated date State: The state where vaccination is done. Sessions Sites First Dose Administered Second Dose Administered Male (Dose Administered) Female (Dose Administered) Transgender (Dose Administered) Covaxin (Dose Administered) Covishield (Dose Administered) Sputnik V AEFI 18-44 Years (Dose Administered) 45-60 Years (Dose Administered) 60+ Years (Dose Administered) 18-44 Years (Individuals Vaccinated) 45-60 Years (Individuals Vaccinated) 60+ Years (Individuals Vaccinated) Male (Individuals Vaccinated) Female (Individuals Vaccinated) Transgender (Individuals Vaccinated) Total Individuals Vaccinated.

Updated C State	Total Dose Sessions	Sites	First Dose	Second Dc Male (Dose Administered)	Female (Dose Administered)	Transgender (Dose Administered)	Covaxin (Dose Administered)	Covishield (Dose Administered)	Sputnik V (Dose Administered)	AEFI (Dose Administered)	18-44 Year (Individuals Vaccinated)	45-60 Year (Individuals Vaccinated)	60+ Year (Individuals Vaccinated)	Total (Individuals Vaccinated)	Male (Individuals Vaccinated)	Female (Individuals Vaccinated)	Transgender (Individuals Vaccinated)
##### India	48276	3455	2957	48276	0	0	579	47697	0	0	23757	24517	2	2	23757	24517	2
##### India	58604	8532	4954	58604	0	0	635	57969	0	0	27348	31252	4	4	27348	31252	4
##### India	99449	13611	6583	99449	0	0	1299	98150	0	0	41361	58083	5	5	41361	58083	5
##### India	195525	17855	7951	195525	0	0	3017	192508	0	0	81901	113613	11	11	81901	113613	11
##### India	251280	25472	10504	251280	0	0	3946	247334	0	0	98111	153145	24	24	98111	153145	24
##### India	365965	32226	12600	365965	0	0	5367	360598	0	0	132784	233143	38	38	132784	233143	38
##### India	549381	36988	14115	549381	0	0	8128	541253	0	0	193899	355402	80	80	193899	355402	80
##### India	759008	43076	15605	759008	0	0	11192	747816	0	0	267856	491049	103	103	267856	491049	103
##### India	835058	49851	18111	835058	0	0	13156	821902	0	0	296283	538647	128	128	296283	538647	128
##### India	1277104	55151	19682	1277104	0	0	18858	1258246	0	0	444137	832766	201	201	444137	832766	201
##### India	1293784	60821	21467	1293784	0	0	19604	1274180	0	0	449119	844448	217	217	449119	844448	217
##### India	1726490	69495	23737	1726490	0	0	27377	1699113	0	0	586081	1140137	272	272	586081	1140137	272
##### India	2295491	78523	25610	2295491	0	0	36921	2258570	0	0	771229	1523939	323	323	771229	1523939	323
##### India	2814803	83664	26219	2814803	0	0	43604	2771199	0	0	939069	1875368	366	366	939069	1875368	366
##### India	3067736	87822	26643	3067736	0	0	48300	3019436	0	0	1022380	2044950	406	406	1022380	2044950	406
##### India	3127107	91593	27011	3127107	0	0	58890	3068217	0	0	1061307	2065391	409	409	1061307	2065391	409
##### India	3350265	97432	27751	3350265	0	0	69372	3280893	0	0	1152344	2197431	490	490	1152344	2197431	490
##### India	3527971	106461	29522	3527971	0	0	76794	3451177	0	0	1218507	2308898	566	566	1218507	2308898	566
##### India	3825835	116568	31167	3825835	0	0	86001	3739834	0	0	1324273	2500887	675	675	1324273	2500887	675
##### India	4314304	126714	32505	4314304	0	0	99015	4215289	0	0	1504527	2809042	735	735	1504527	2809042	735
##### India	4765924	137110	33947	4765924	0	0	133911	4632013	0	0	1710466	3054631	827	827	1710466	3054631	827
##### India	5111827	145121	34814	5111827	0	0	195504	4916323	0	0	1896705	3214244	878	878	1896705	3214244	878
##### India	5168099	153231	35778	5168099	0	0	206130	4961989	0	0	1928947	3238273	879	879	1928947	3238273	879
##### India	5615489	164116	36740	5615489	0	0	290785	5324714	0	0	2211020	3403576	903	903	2211020	3403576	903
##### India	6002474	175556	38131	6002474	0	0	363692	5638782	0	0	2458329	3543208	937	937	2458329	3543208	937
##### India	6462270	187224	39339	6462270	0	0	480760	5981510	0	0	2785194	3676114	962	962	2785194	3676114	962
##### India	6958553	197650	40332	6958553	0	0	609108	6349445	0	0	3153649	3803908	996	996	3153649	3803908	996

Figure 3.2: Covid vaccine statewise dataset

# Chapter 4

## EXPLORATORY DATA ANALYSIS

---

### 4.1 ANALYSING COVID-19 CASES IN INDIA

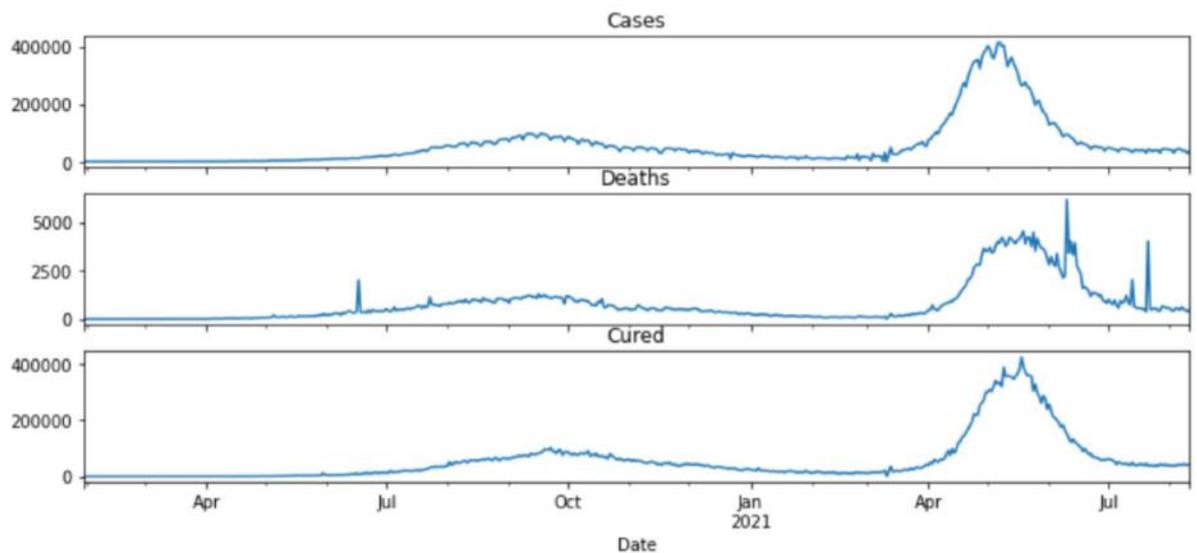


Figure 4.1: CURED, DEATH, CONFIRMED

Plotting time series of 3 variables, Cases, Death and Cured, we can understand that from the beginning of covid 19 virus to about 2020 July there is a flattened graph in case of covid cases and the cured, and at the end of July, there is a slight change in the case of covid deaths. And from 2020 August to about 2021 January, there is a slight increase in covid cases so that the number of covid deaths also increases and

4.2. WEEKLY CASES

the relief is that the number of cured cases also increases. From 2021 April, we can see that the three cases increases and decreased by about August 2021.

4.2 WEEKLY CASES

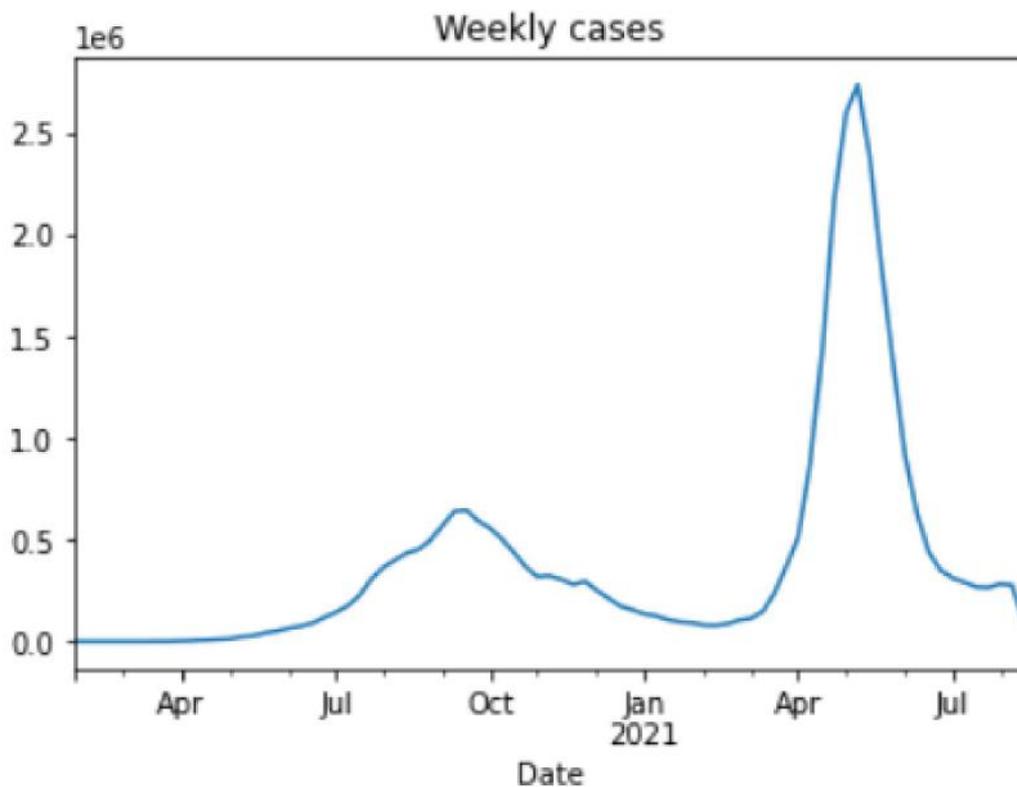


Figure 4.2: weekly cases

This graph represents the variation in weekly covid cases from February to August. It shows only the least number of covid cases during the month of February, March, and April. Then it increases slowly and reaches 1.0 in the month of October. Then the graph decrease till the end of March 2021. Then there is a rapid increase in covid cases and it reaches its steep in may, and with the rapid increase, there is also a sudden rapid decrease in the covid cases.

### 4.3 MONTHLY CASES

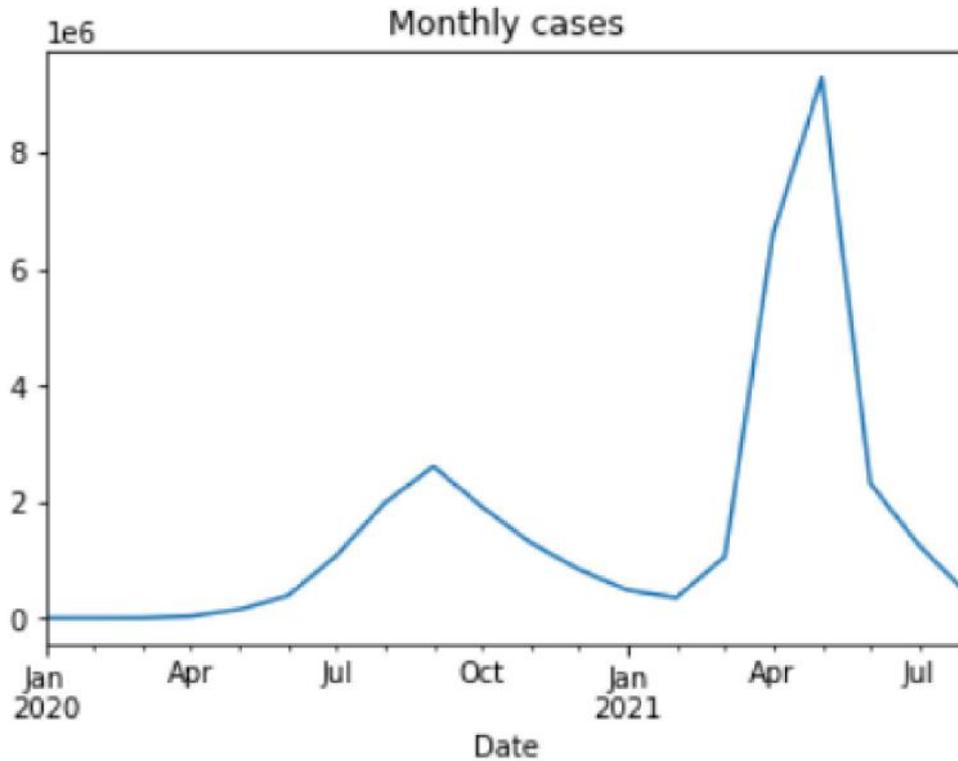


Figure 4.3: monthly cases

This graph represents the variation in monthly covid cases from January to August. It shows only the least number of covid cases during the month of January, February, March, and April. Then it increases slowly and reaches 3 in the month of september. Then the graph decreases till the end of february 2021. Then there is a rapid increase in covid cases and it reaches its step in may, and with the rapid increase, there is also a sudden rapid decrease in the covid cases in the month of june and july.

### 4.4 COVID CASES 200 DAYS

The graph represents the variation in covid cases for 200 days from 2021 February to August. The graph goes steady from February to March

4.5. INDIA AGE GROUP WISE DISTRIBUTION

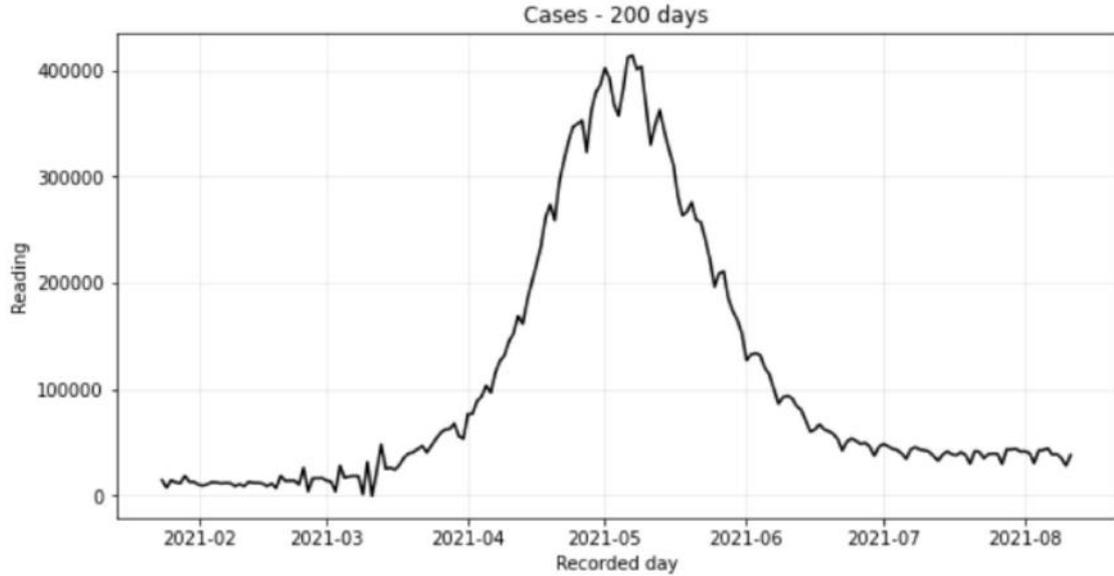


Figure 4.4: cases 200-days

with least number of covid cases as compared to other months. Then it shows a variation in the month of March from where there is a rapid increase in the covid cases and it reaches its steep in the month of May, even though it shows slight variation but it kept increasing. Later on the graph decreases rapidly till June then slowly it keep on decreasing till May from where the graph goes steadily.

4.5 INDIA AGE GROUP WISE DISTRIBUTION

We could see that the age group  $\geq 40$  is the most affected which is against the trend which says elderly people are more at risk of being affected. Only 17 percentage of people  $\geq 60$  are affected. the mostly affected people are from the age group of 20-29 with the highest percentage of 24.9 percentage

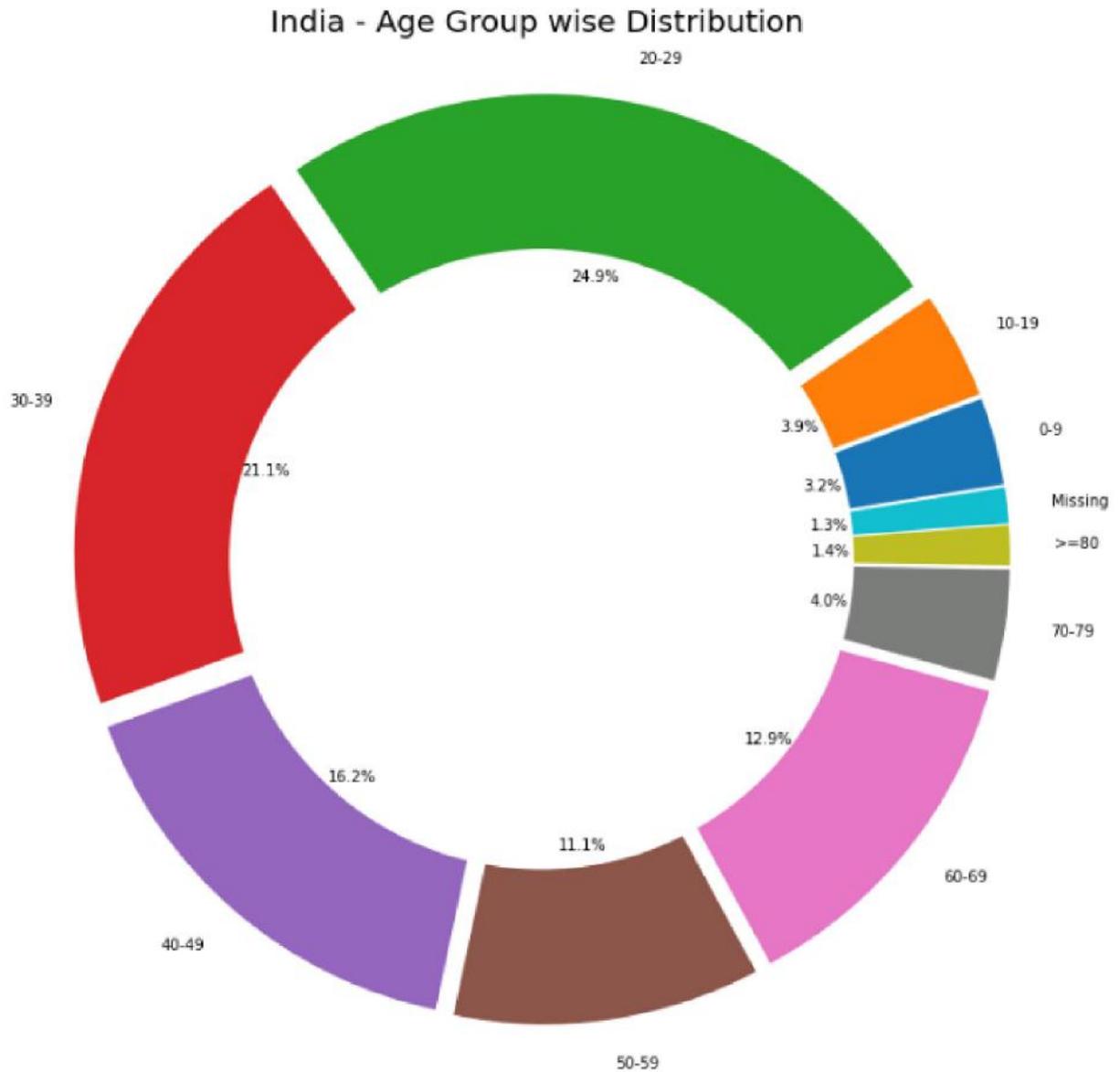


Figure 4.5: India age group wise distribution

## 4.6 PERCENTAGE OF GENDER

Men are the most affected accounting to 67 percentage. But, remember we have 80 percentage data missing.

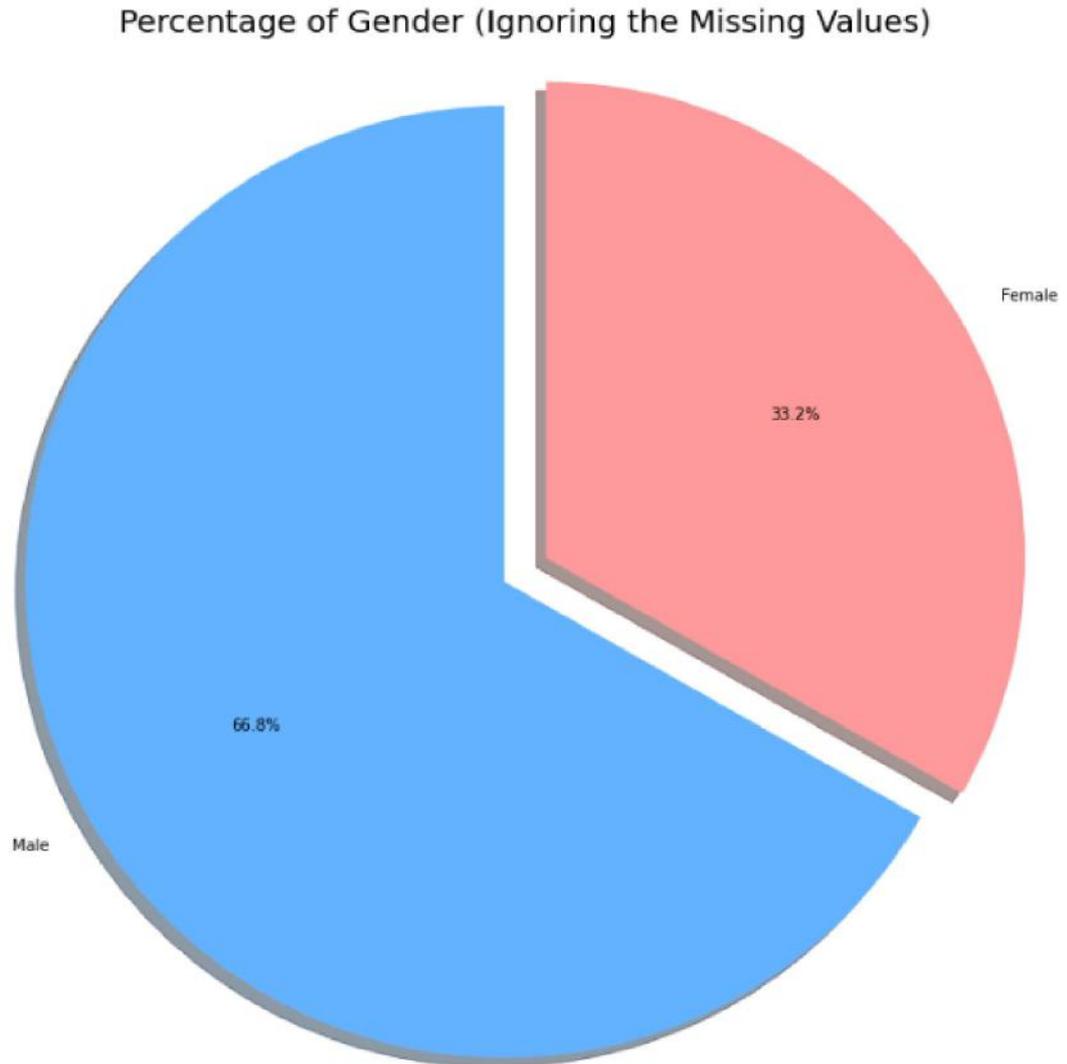


Figure 4.6: percentage of gender

## 4.7 CRITICALITY OF LOCKDOWN AND CONTAINMENT FOR INDIA

The graph for criticality of lockdown and containment for india shows that during the nationwide lockdown which was on 25 th march, the number of covid cases were at its lowest. During the time of 15 th april, where India was without lockdown and containment measures, the covid cases had increased by 8.2 lakhs. The covid cases reached 1.2 lakh when there was a containment zone, but no lockdown. When nationwide lockdown and containment measures were undertaken, the number decreased to 7447 confirmed cases.

### Criticality of Lockdown & Containment for India

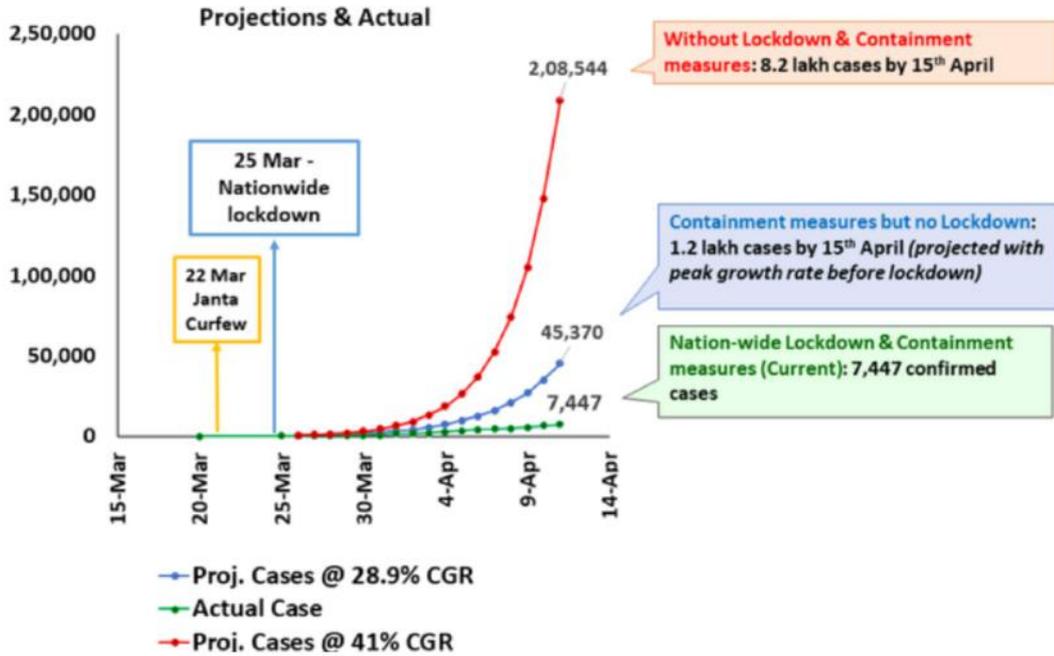


Figure 4.7: criticality of lockdown and containment for india

### 4.8 5 DAYS MOVING AVERAGE OF CONFIRMED CASES IN TOP 15 STATES

From the 15 graph it is clear that, the 7 states Rajasthan, Tamil nadu, Gujarat, west bengal, Uttar pradesh, Punjab and Maharashtra has the highest covid 19 cases

4.9. TOP 10 STATES IN EACH HEALTH FACILITY



Figure 4.8: 5 days moving average of confirmed cases in top 15 states

4.9 TOP 10 STATES IN EACH HEALTH FACILITY

This 4 graph shows information regarding Primary health centers, community health centers, district hospitals and Total public health facilities of India. In which Uttar Pradesh stands at the top in case of health facility among the 4 graphs. Where west bengal stands at the lowest position among the top 10 states in case of Primary health centers and karnataka stands in the lowest position in case of Community health centers, District hospitals and Total public health facilities. .

4.10. URBAN AND RURAL HEALTH FACILITY

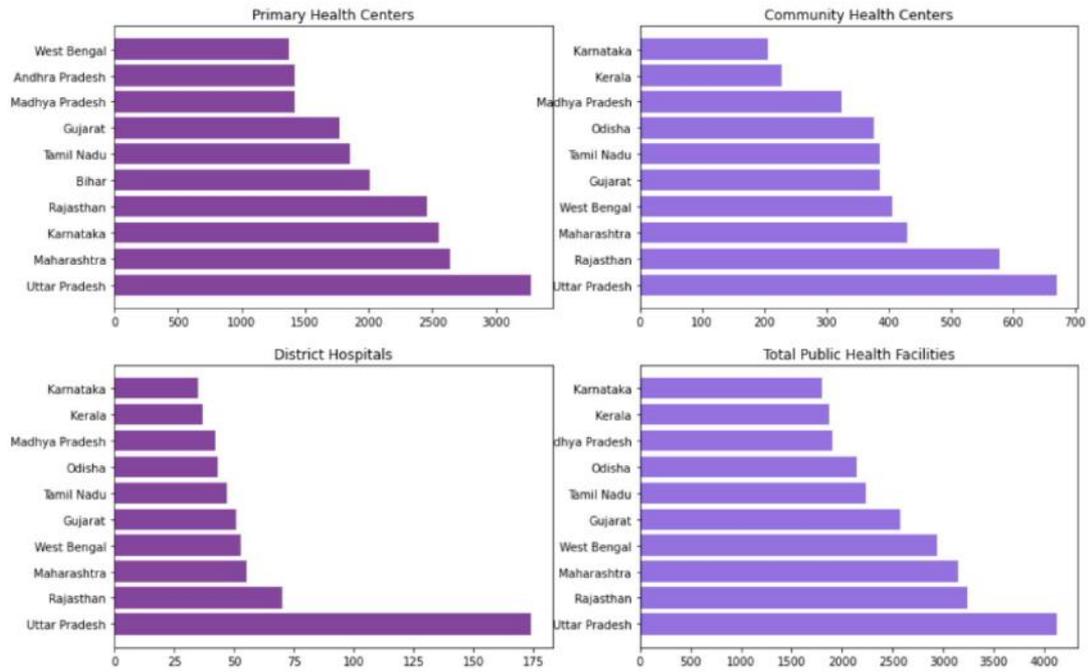


Figure 4.9: top 10 states in each health facility

4.10 URBAN AND RURAL HEALTH FACILITY

The 4 graphs which are shown above give us the information regarding Urban and Rural health facilities. Uttar Pradesh stands front in case of Rural hospital and Tamil nadu in case of Rural beds, where Tamil nadu and Madhya Pradesh being the lowest in case of Rural hospital and rula beds. In case of Urban hospitals, Tamil nadu stands first and West bengal in case of urban beds,where Gujarat and Andhra pradesh being the lowest in case of Urban hospital and Urban beds.

4.11. TESTING STATEWISE INSIGHT

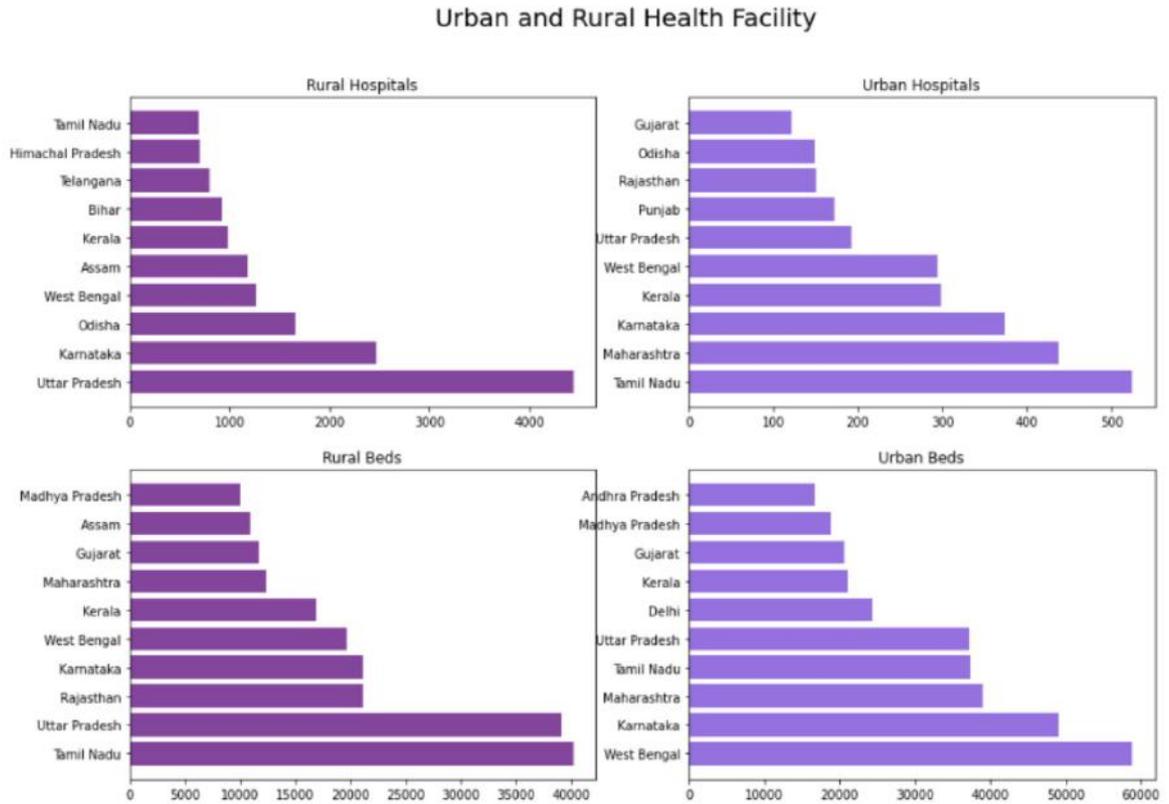


Figure 4.10: urban and rural health facility

4.11 TESTING STATEWISE INSIGHT

This is the information about statewise covid testing. Where Tamil nadu tests about 3 lakh peoples and among them a large majority was negative and some samples were positive.andhra pradesh tests about 230000 samples and less than 1000 samples are positive. Maharashtra tests about 250000 samples and about 4000 samples are positive. Arunachal pradesh, Assam , chandigarh, chhattisgarh, Haryana, Himachal pradesh etc shows very less positive samples.among the states in india, Maharashtra have the highest positive samples.

4.12. ICMR TESTING CENTERS IN EACH STATE

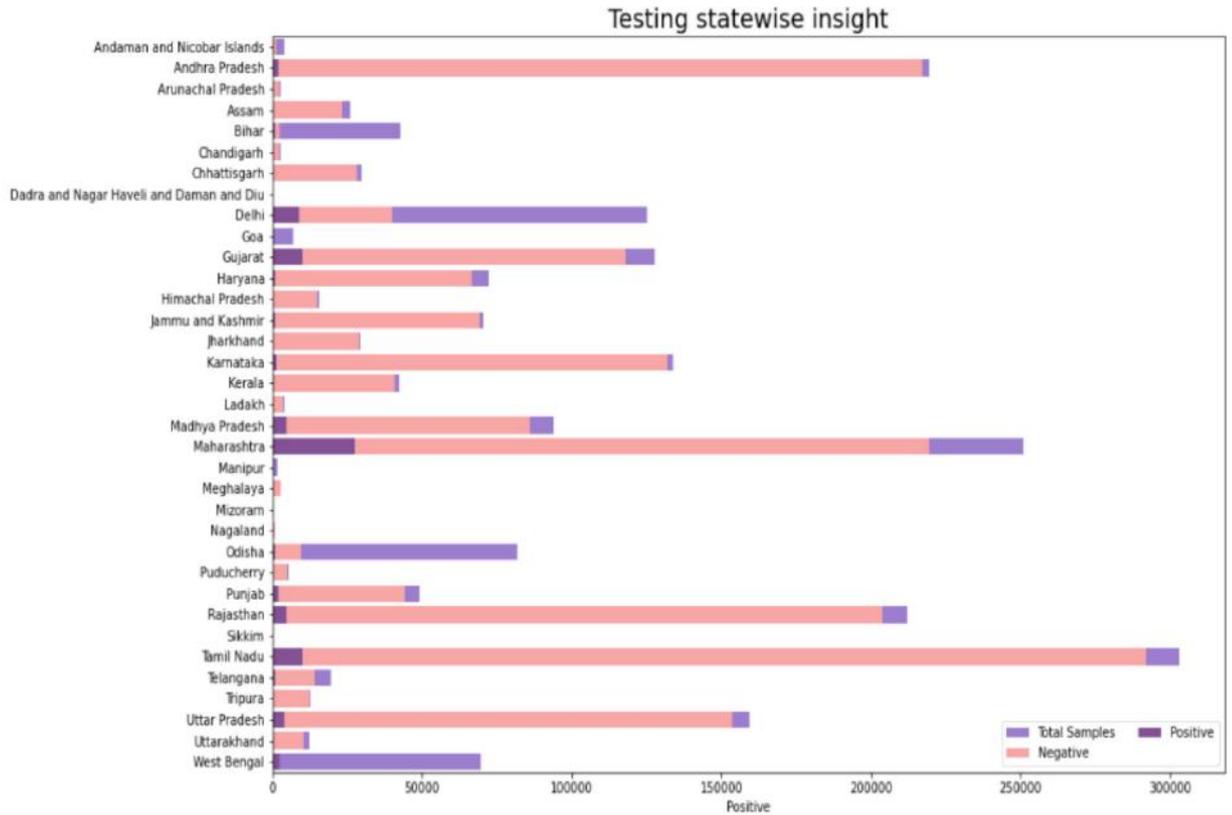


Figure 4.11: testing statewise insight

4.12 ICMR TESTING CENTERS IN EACH STATE

From this we can understand that maharashtra has the large number of testing centers as compared to other states and the states like Mizoram, Goa, Puducherry, Meghalaya, Arunachal pradesh, Tripura, Andaman and Nicobar islands, Sikkim, Ladakh has the least number of testing centers in india.

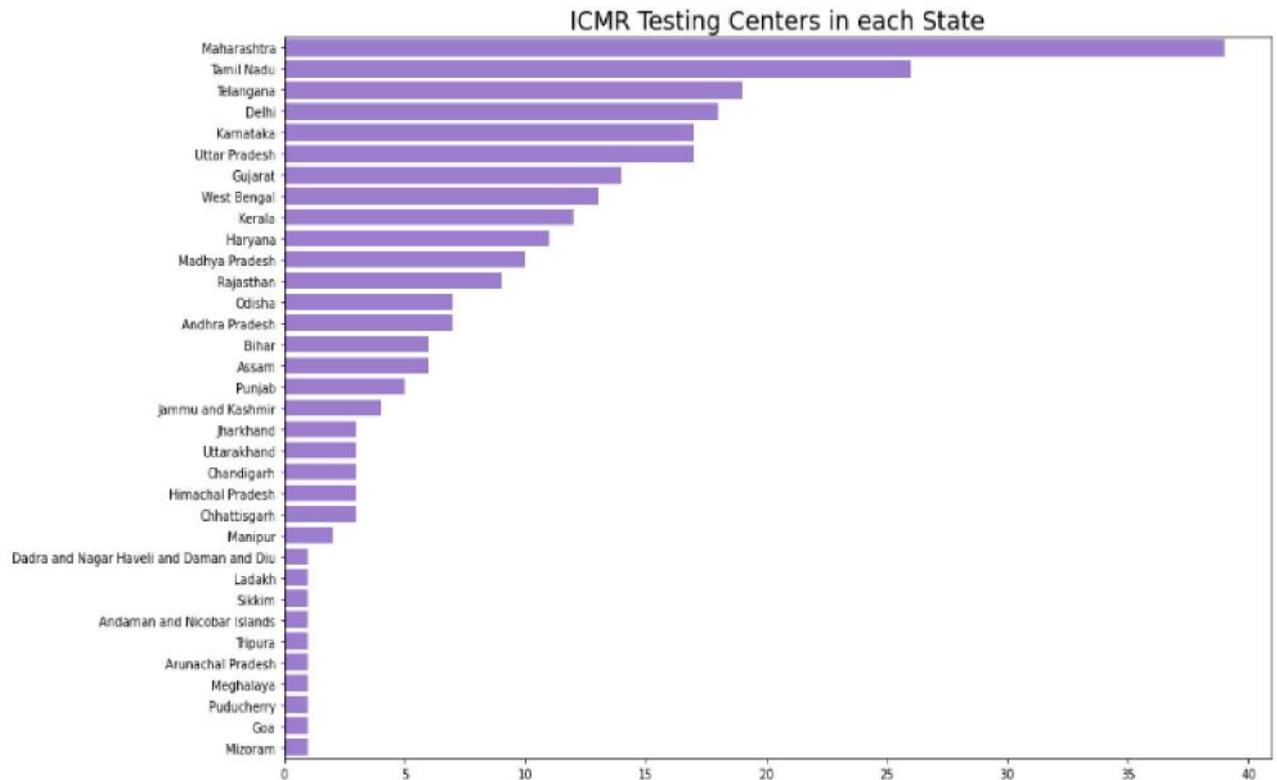


Figure 4.12: ICMR testing centers in each state

### 4.13 ECONOMIC IMPACT OF COVID 19 ON INDIA

The coronavirus is having a growing impact on the indian economy. Here is an insight on how an extension of the lockdown would affect the economy.

#### 4.13.1 THREE ECONOMIC SCENARIOS MODEL INDIA GDP ESTIMATES

In scenario 1, the nationwide lockdown was lifted on April 15, 2020 (end of 21-day deadline);prior relaxation for select areas(eg, logistics). Back to work in “Save lives and livelihoods” mode,with strong protection protocols. It’s a support to households, Corporations, and banking systems with fiscal and monetary stimuli. In scenario 2, Lockdown continues until mid-May 2020; moderate relaxation after April 15,2020(end of 21-day deadline); restarting supply chains and normalizing production and consumption and takes 3-4 months. Stabilization and stimulus

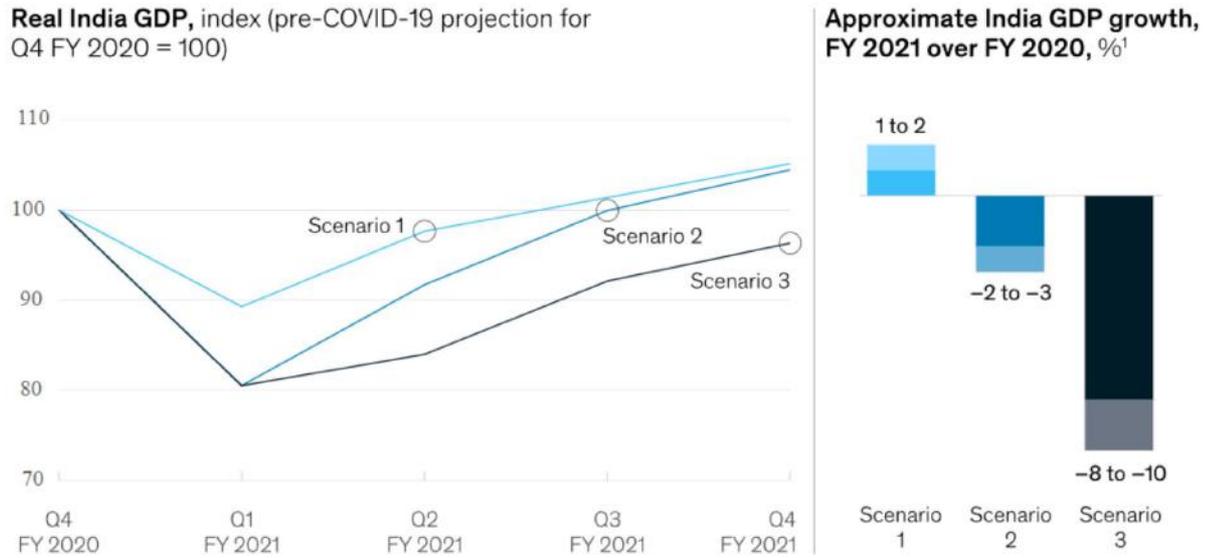


Figure 4.13:

package, broader than in scenario 1. In scenario 3, Lockdown as in scenario 2 with additional 2-3 weeks lockdown in Q2 and Q4 FY 2021 because of virus resurgence. There is low labor availability because of limited reverse migration. Stabilization and stimulus package, even broader than in scenario 2.

#### 4.14 MOST IMPACTED STATES/UT DURING COVID 19 LOCKDOWN

During COVID-19 lockdown every state/UT faced the crisis of unemployment. Top five most impacted states/UT during COVID-19, Puducherry Jharkhand Bihar Haryana Tripura

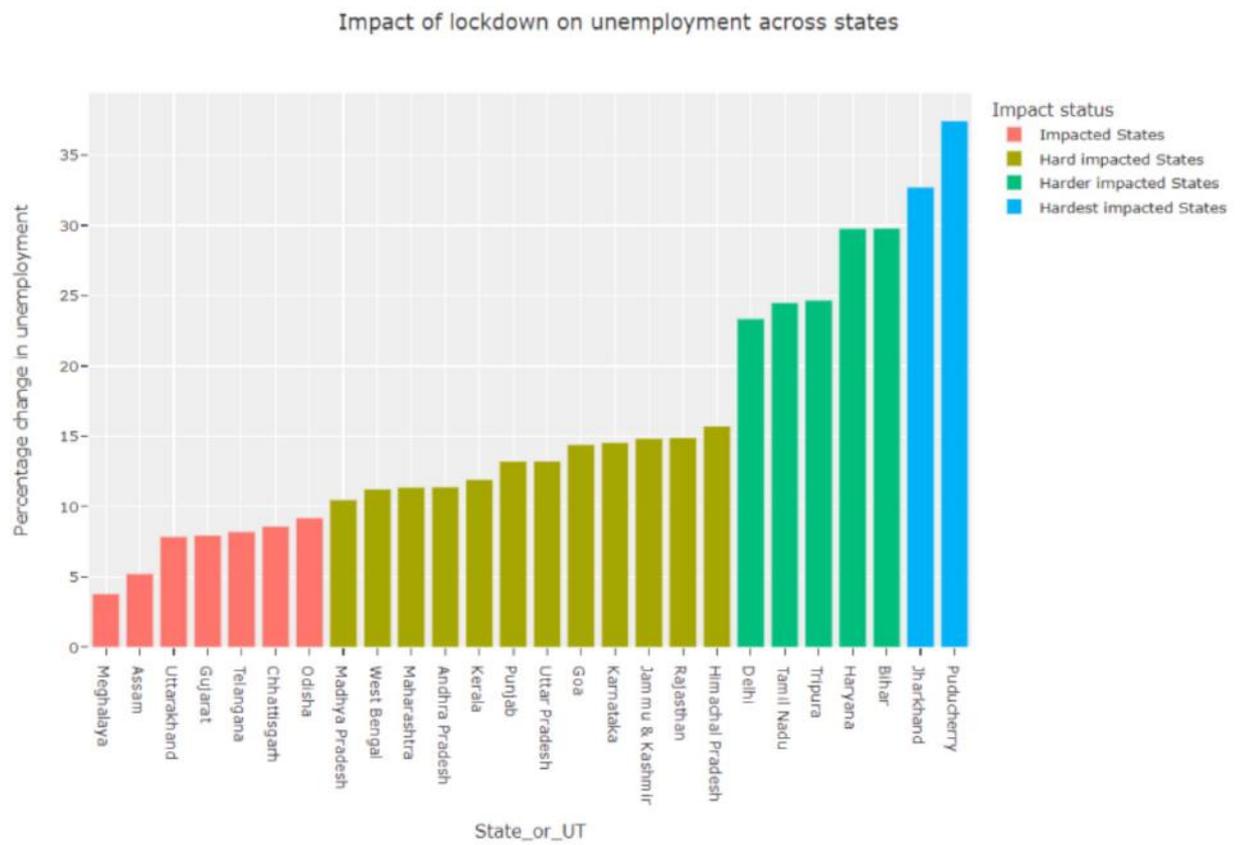


Figure 4.14:

# Chapter 5

## PREDICTION AND FORECASTING

---

### 5.1 TIME SERIES FORECASTING:

Time series forecasting is the process of analyzing time series data using statistics and modeling to make predictions and inform strategic decision-making. It's not always an exact prediction, and the likelihood of forecasts can vary wildly—especially when dealing with the commonly fluctuating variables in time series data as well as factors outside our control. However, forecasting insight about which outcomes are more likely—or less likely—to occur than other potential outcomes.

Time series forecasting occurs when you make scientific predictions based on historical time stamped data. It involves building models through historical analysis and using them to make observations and drive future strategic decision-making.

### 5.2 FORECASTING USING LSTM MODEL:

The recurrent neural network uses long short-term memory blocks to provide context for the way the program receives inputs and creates outputs. The long short-term memory block is a complex unit with various components such as weighted inputs, activation functions, inputs from previous blocks and eventual outputs.

The unit is called a long short-term memory block because the program is using a structure founded on short-term memory processes to create longer-term memory. These systems are often used, for example, in natural language processing. The recurrent neural network uses the long short-term memory blocks to take a particular word or phoneme, and evaluate it in the context of others in a string, where memory can be useful in sorting and categorizing these types of inputs. In general, LSTM is an accepted and common concept in pioneering recurrent neural networks.

### 5.3 MAKING LSTM PREDICTIONS

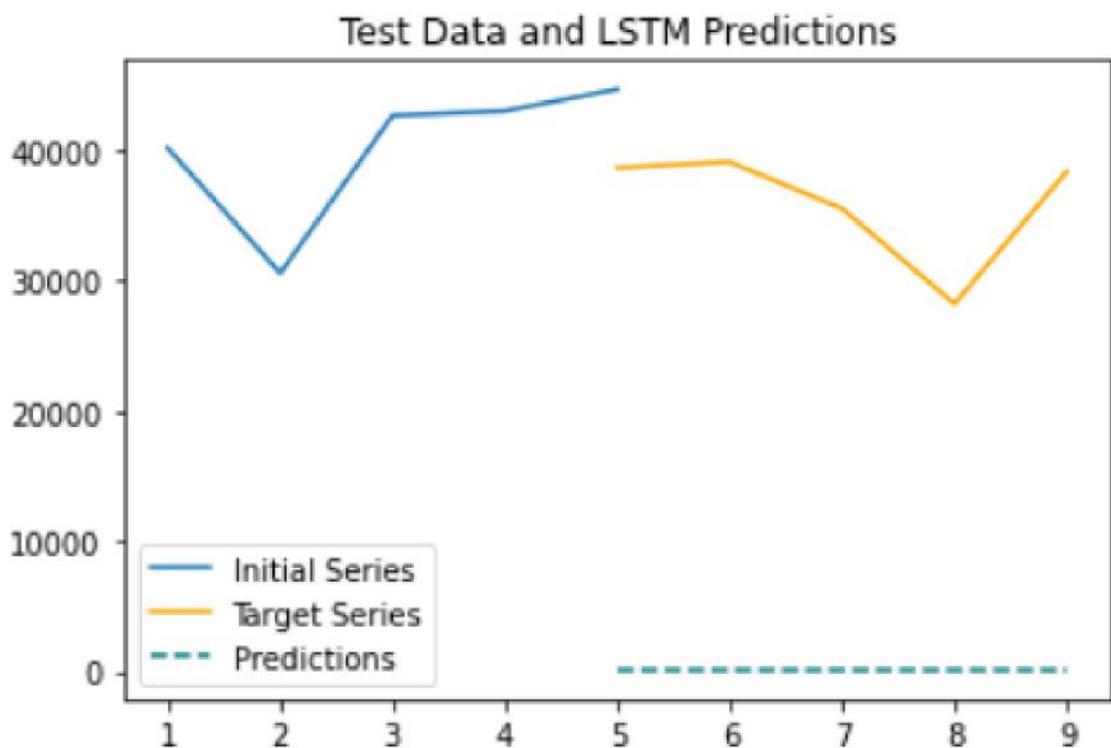


Figure 5.1:

The deep learning model fails to learn due to a small number of training instances.

## 5.4 DECOMPOSING THE TIME SERIES.

Any time series has 3 components associated with it:

1. Trend
2. Seasonality
3. Residual

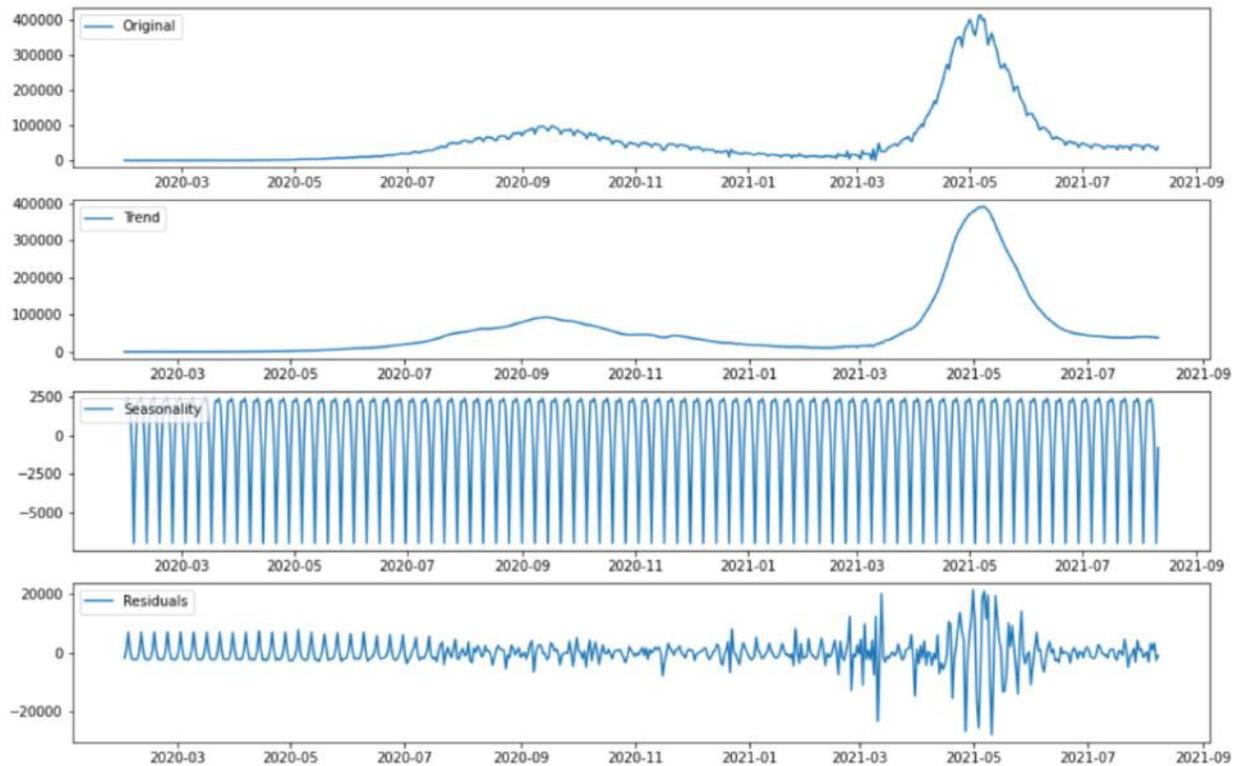


Figure 5.2:

From the 3 time series components, it is clear that the covid cases are low during the period from march 2020 to about march 2021. But there is a slight increase in corona cases during september 2020. And the highest corona cases were reported in the month of may 2021 as compared to other months.

## 5.5 PLOTTING AUTO-CORRELATION FUNCTION

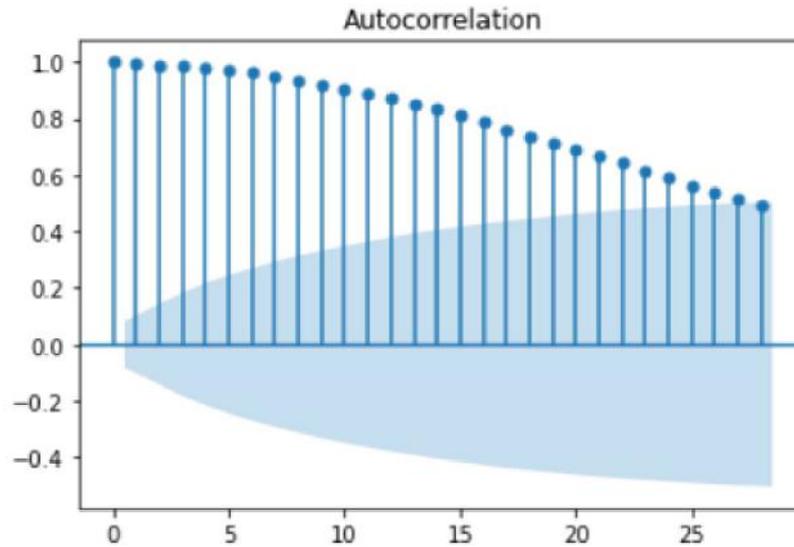


Figure 5.3:

## 5.6 DICKEY FULLER TEST (ADF TEST)

Augmented Dickey Fuller test (ADF Test) is a common statistical test used to test whether a given Time series is stationary or not. from this test we get Pvalue:0.022928166916663568 p-value obtained is greater than significance level. Hence we cannot reject the null hypothesis. Therefore, We conclude the Time series is non-stationary. and then by doing this, we get p-value: 0.0008932948353244133 so we can say that the time series is stationary

## 5.7 DIVIDING TIME SERIES INTO TRAIN AND TEST FOR PREDICTIONS

We will be making predictions for 30 days. we are dividing the time series into train and test data.30 data are train data and 30 data are test data. so,at first we are doing simple exponential smoothing.

From this test ,it is clear that the test data and predicted values are same.And the MSE: 664556681.3479285

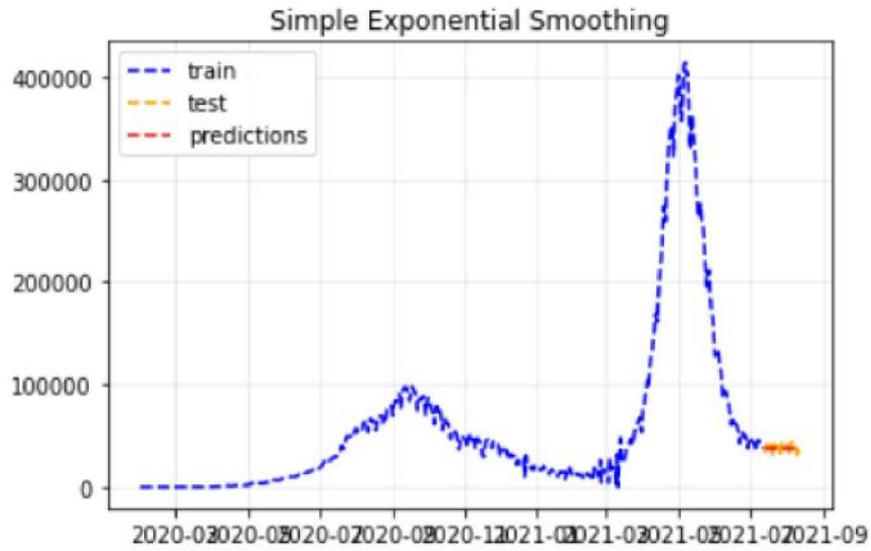


Figure 5.4: simple exponential smoothing

Next is double exponential smoothing

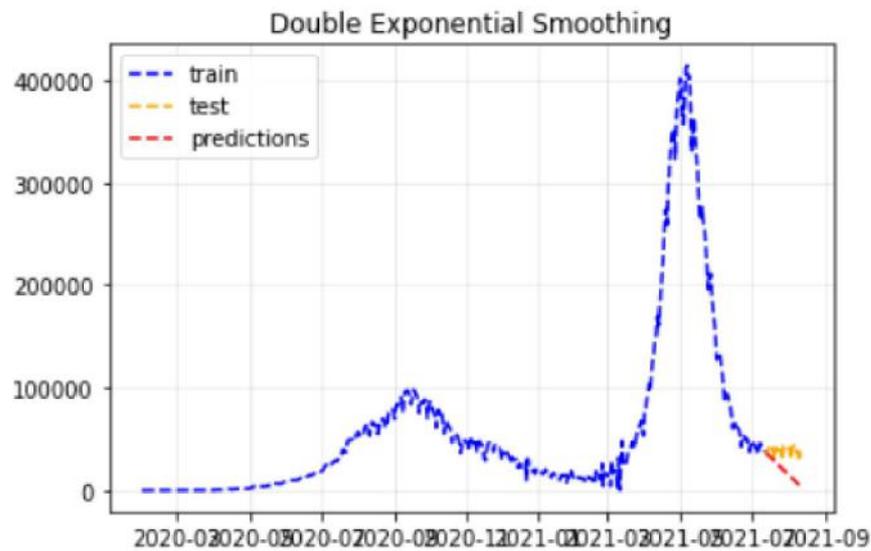


Figure 5.5: DOUBLE EXPONENTIAL SMOOTHING

From this test ,it is clear that the test data and the predicted values are different and the MSE: 13456167365.577461

Next is Triple exponential smoothing

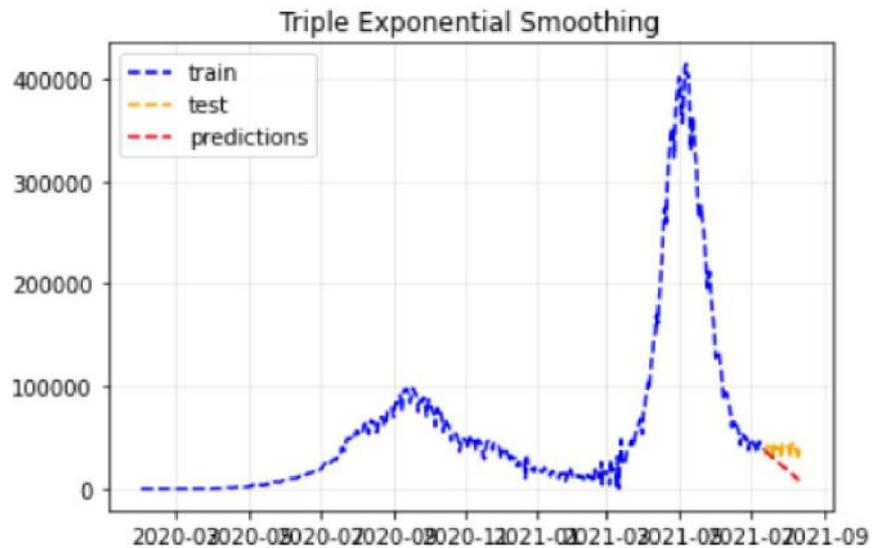


Figure 5.6: TRIPLE EXPONENTIAL SMOOTHING

From this test ,it is clear that the test data and the predicted values are different and the MSE: 10736586421.807472

Comparing the results of the 3 statistical models.

Single MSE :664556681.3479285

Double MSE :13456167365.577461

Triple MSE :10736586421.807472

## Chapter 6

# CONCLUSION

---

The COVID-19 pandemic is the defining global health crisis of our time and the greatest global humanitarian challenge the world has faced since World War II. The virus has spread widely, and the number of cases is rising daily as governments work to slow its spread. So as mentioned above the covid cases, deaths and cured are increased during the month of may 2021 and the age group  $\geq 40$  is the most affected. Men are the most affected accounting to 67 Percent. When nationwide lockdown and containment measure was undertaken, so that the number decreased to 7447 confirmed cases. Rajasthan, Tamil nadu, Gujarat, west bengal, Uttar pradesh, Punjab and Maharashtra has the highest covid 19 cases. Uttar Pradesh stands at the top in terms of health facilities. Maharashtra has the highest number of confirmed cases. Tamil nadu tests about 3 lakh peoples and among them a large majority was negative and some samples were positive. andhra pradesh tests about 230000 samples and less than 1000 samples are positive. Maharashtra has the large number of testing centers as compared to other states and the states like Mizoram, Goa, Puducherry, Meghalaya, Arunachal pradesh, Tripura, Andaman and Nicobar islands, Sikkim, Ladakh has the least number of testing centers in India. Puducherry, Jharkhand, Bihar, Haryana, Tripura, these were the top five states which suffered unemployment. From this the 3 time series components- Trend, Seasonality and residual, it is clear that the covid cases is low during the period from march 2020 to about march 2021. But there is a slight in-

crease in corona cases during september 2020. And the highest corona cases were reported in the month of may 2021 as compared to other months.

# Chapter 7

## REFERENCE

---

- [1] Lu H., Stratton C.W., Tang Y.W. Outbreak of pneumonia of unknown etiology in Wuhan China: the mystery and the miracle. *J Med Virol.* 2020 [PMC free article] [PubMed] [Google Scholar]
- [2] Hui D.S., E I.A., Madani T.A., Ntoumi F., Kock R., Dar O. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health – the latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis.* 2020;91:264–266. [PMC free article] [PubMed] [Google Scholar]
- [3] Gorbalenya A.E.A. Severe acute respiratory syndrome-related coronavirus: the species and its viruses – a statement of the Coronavirus Study Group. *BioRxiv.* 2020 doi: 10.1101/2020.02.07.937862. [CrossRef] [Google Scholar]
- [4] Burki T.K. Coronavirus in China. *Lancet Respir Med.* 2020 [PMC free article] [PubMed] [Google Scholar]
- [5] NHS press conference, February 4, 2020. Beijing, China. National Health Commission (NHC) of the People’s Republic of China, <http://www.nhc.gov.cn/xcs/xwbd/202002/235990d202056cfcb202043f202004a202070d202007>
- [6] COVID-19: a timeline of the coronavirus outbreak. Devex., <https://www.devex.com/news/covid-19-a-timeline-of-the-coronavirus-outbreak-96396> (Accessed on 30 May 2020).Google Scholar
- [7] Coronavirus pandemic (COVID-19). Our World In Data., <https://ourworldindata.org/coronavirus> (accessed 30 May 2020).Google Scholar
- [8] Coronavirus disease (COVID-19) outbreak. World Health Organization., <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-COVID-19/novel-coronavirus-2019-ncov> (Accessed on 27 May 2020).Google Scholar

- [9] 1st known case of coronavirus traced back to November in China. Live Science., <https://www.livescience.com/first-case-coronavirus-found.html> (accessed 27 May 2020).Google Scholar
- [10] 2019-nCoV outbreak is an emergency of international concern. World Health Organization., <http://www.euro.who.int/en/health-topics/health-emergencies/international-health-regulations/news/news/2020/2/2019-ncov-outbreak-is-an-emergency-of-international-concern> (accessed 27 May 2020).Google Scholar
- [11] Situation update worldwide, as of 30 may 2020 epidemiological update. European Centre for Disease Prevention and Control., <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases> (Accessed on 30 May 2020).Google Scholar
- [12] Zhou P , Yang X , Wang X , et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;579:270–3 doi: 10.1038/s41586-020-2012-7.CrossRefPubMedGoogle Scholar
- [13] Sommer IE , Bakker PR . What can psychiatrists learn from SARS and MERS outbreaks? *Lancet Psychiatry* 2020;18.Google Scholar
- [14] Zhang T , Wu Q , Zhang Z . Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Current Biol* 2020;30:1346–51 doi: 10.1016/j.cub.2020.03.022.CrossRefPubMedGoogle Scholar
- [15] Liu P , Chen W , Chen J-P . Viral metagenomics revealed sendai virus and coronavirus infection of Malayan Pangolins (*Manis javanica*). *Viruses* 2019;11:979 doi: 10.3390/v11110979.Google Scholar
- [16] Li X , Zai J , Zhao Q , et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J Med Virol* 2020;1–10.Google Scholar
- [17] Coronavirus and the laboratories in Wuhan. *The Wall Street Journal*. 21 April 2020., <https://www.wsj.com/articles/coronavirus-and-the-laboratories-in-wuhan-11587486996> (accessed 28 May 2020).Google Scholar
- [18] Concealing the truth' Chinese whistleblower speaks out on COVID-19 outbreak in Wuhan. *Express* 2 April 2020., <https://www.express.co.uk/news/world/1263227/china-coronavirus-whistleblower-covid19-wuhan-seafood-market-spt> (accessed 28 May 2020).Google Scholar

- [19] Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). World Health Organization. 28 February 2020.,  
[https://www.who.int/publications-detail/report-of-the-who-China-joint-mission-on-coronavirus-disease-2019-\(COVID-19\)](https://www.who.int/publications-detail/report-of-the-who-China-joint-mission-on-coronavirus-disease-2019-(COVID-19)) (accessed 28 May 2020).Google Scholar
- [20] Coronavirus disease 2019. Wikipedia.,  
[https://en.wikipedia.org/wiki/Coronavirus\\_disease\\_2019](https://en.wikipedia.org/wiki/Coronavirus_disease_2019)(accessed 2 May 2020).Google Scholar



## CERTIFICATE

This is to certify that the project entitled “Effect of processing of rice on quality of unniyappam”, submitted by Ms. Tanya Baby (Reg no: VM20FPT014) during February 2022 – July 2022, in the partial fulfilment for award of the degree of Master of Vocational Studies in Food Processing Technology, Kerala, is the result of study carried out by her in department of Traditional Foods and Applied Nutrition , at CSIR-Central Food Technology Research Institute, Mysore, under the guidance of Dr. Chetana Ramakrishna

**Dr. AS CHAUHAN**  
Senior Principal Scientist  
Head of Department  
Department of Traditional Foods and Applied Nutrition  
CSIR-CFTRI, Mysore

**Dr. CHETANA RAMAKRISHNA**  
Principal Technical Officer

**Place: Mysore**

**Date: 21-07-2022**

**Co-ordinator**  
**HRD Department**  
**CSIR-CFTRI MYSURU-570 020**

प्रमुख / HEAD  
मानव संसाधन विकास विभाग  
Human Resource Development  
सीएसआईआर-सीएफटीआरआई / CSIR-CFTRI  
मैसूरु / Mysuru-570 020



# **EFFECT OF PROCESSING OF RICE ON QUALITY OF UNNIYAPPAM**

**SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE OF  
MASTER OF VOCATIONAL STUDIES IN FOOD PROCESSING TECHNOLOGY**

**SUBMITTED BY  
TANYA BABY**

**VM20FPT014**



**St. TERESA'S COLLEGE (AUTONOMOUS) ERNAKULAM  
Affiliated to Mahatma Gandhi University, Kottayam, Kerala**

**UNDER THE GUIDANCE OF  
Dr. CHETANA RAMAKRISHNA  
Principal Technical Officer**



**Department of Traditional Foods and Applied Nutrition  
CSIR-CENTRAL FOOD TECHNOLOGICAL RESEARCH INSTITUTE  
Mysore-570020, Karnataka (INDIA)**

Project Report

On

# FUZZY LOGIC AND IT'S APPLICATION

*Submitted*

*in partial fulfilment of the requirements for the degree of*

BACHELOR OF SCIENCE

*in*

MATHEMATICS

*by*

NIA NELSON

(Register No. AB19AMAT021)

*Under the Supervision of*

DR. URSALA PAUL



DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)  
ERNAKULAM, KOCHI - 682011

APRIL 2022

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



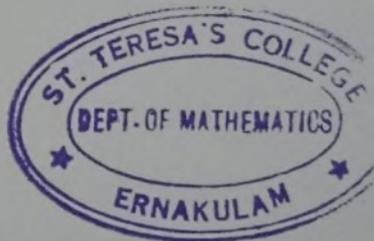
CERTIFICATE

This is to certify that the dissertation entitled, **FUZZY LOGIC AND IT'S APPLICATIONS** is a bonafide record of the work done by Ms. **NIA NELSON** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 04/03/2022

Place: Ernakulam

*Ursala*  
**Dr. Ursala Paul**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



*Ursala*  
**Dr. Ursala Paul**  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

**External Examiners**

1:.....  
*Jeeva*

2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Dr. Ursala Paul, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam

Date: 04/03/2022

NIA NELSON

AB19AMAT021

## ACKNOWLEDGEMENT

I express my deep sense of gratitude to Dr. Ursala Paul, our faculty incharge and HOD of the Department of Mathematics, St Teresa's College (Autonomous) Ernakulam for her excellent guidance, suggestions and assistance which helped me to complete this project successfully.

I would like to thank all other teachers of the department and all those who gave me moral support and helped me to complete my project. My deepest gratitude to my parents and friends for their valuable support and cooperation all throughout this work.

Thank you All

Ernakulam

Date: 15/02/2022



NIA NELSON

AB19AMAT021

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 OBJECTIVES OF THE PROJECT . . . . .	1
1.2 METHODOLOGY AND APPROACH . . . . .	1
<b>2 EMERGENCE OF FUZZY LOGIC</b>	<b>2</b>
2.1 HARD COMPUTING . . . . .	2
2.2 SOFT COMPUTING . . . . .	2
2.2.1 ARTIFICIAL NEURAL NETWORK . . . . .	3
2.2.2 FUZZY LOGIC . . . . .	4
2.2.3 EVOLUTIONARY COMPUTATION . . . . .	4
2.3 ORIGIN OF FUZZY LOGIC . . . . .	4
<b>3 FUZZY LOGIC</b>	<b>6</b>
3.1 CRISP LOGIC v/s FUZZY LOGIC . . . . .	6
3.2 FUZZY PROPOSITION . . . . .	7
3.2.1 UNCONDITIONAL AND UNQUALIFIED PROPOSITIONS	7
3.2.2 UNCONDITIONAL AND QUALIFIED PROPOSITIONS .	8
3.2.3 CONDITIONAL AND UNQUALIFIED PROPOSITIONS .	8
3.2.4 CONDITIONAL AND QUALIFIED PROPOSITIONS . . .	8
3.3 FUZZY QUANTIFIERS . . . . .	8
3.3.1 FUZZY QUANTIFIERS OF FIRST KIND . . . . .	9
3.3.2 FUZZY QUANTIFIERS OF SECOND KIND . . . . .	10

3.4	FUZZY RULES OF INFERENCE . . . . .	10
3.4.1	MODUS PONENS . . . . .	10
3.4.2	MODUS TOLLENS . . . . .	10
3.4.3	DECOMPOSITION . . . . .	11
3.4.4	AGGREGATION . . . . .	12
4	FUZZY LOGIC CONTROLLER AND ITS APPLICATION	14
4.1	Fuzzy Logic Controller . . . . .	14
4.1.1	Approaches . . . . .	15
4.2	APPLICATION . . . . .	16
5	CONCLUSION	21
	REFERENCES . . . . .	22

# Chapter 1

## INTRODUCTION

---

Fuzzy means vagueness or ambiguity. Fuzzy Mathematics is a branch of mathematics which includes ambiguous set theory and ambiguous logic that deals with the partial incorporation of elements into a set as opposed to including a simple binary "yes" or "no" in one spectrum. In this project we focus on ambiguous logic - a multi-valued logic that can incorporate true values into any real number lying in  $[0,1]$ , and its application in a washing machine.

### 1.1 OBJECTIVES OF THE PROJECT

- To understand the importance of Fuzzy Logic.
- To know how it is different from Crisp Logic and hence its wide acceptance and applications.
- To understand the whole working process of a Fuzzy Logic Controller.
- To study its application in Washing Machine.

### 1.2 METHODOLOGY AND APPROACH

After getting to know about the basic definitions and functioning of fuzzy logic controllers, we have made use of the Mamdani approach of aggregating the rules, by disjunctive method. We have also plotted the membership functions and derived corresponding degree of memberships, which are used for calculations.

## Chapter 2

# EMERGENCE OF FUZZY LOGIC

---

The process of using computer technology to complete a given goal-oriented task is called Computing. Briefly it is the process of giving an input to get back an output using certain control actions.

Computing can be classified as:

1. Hard computing
2. Soft computing

### 2.1 HARD COMPUTING

Hard computing is a traditional computing method. It is the best for solving the mathematical problems. It relies basically on binary logic and crisp system. The input data for hard computing should be exact and clear. The output will be precise and unambiguous. Hard computing uses two-valued logic. It makes use of proper mathematical models or algorithms to perform computing, thus giving precise (not approximate) results. Its control action must be unambiguous and formally defined and there must exist any given mathematical models.

### 2.2 SOFT COMPUTING

Soft computing is a set of algorithms. It includes neural network, fuzzy logic and evolutionary computing. It is an approach where we compute solutions to the existing complex problems, where output results are imprecise or fuzzy in

nature. It does not require any mathematical modelling to find the solution for a given problem. It makes use of algorithms which are not precise, uncertain, partially true and approximate. It is used in solving real-world problems as it has a random probability distribution or pattern that may be analysed statistically but may not be predicted precisely. Soft computing relies basically on formal logic. Soft computing use multivalued logic and is adaptive in nature.

Soft computing has three main branches:

- i) Artificial neural network
- ii) Fuzzy logic
- iii) Evolutionary computation

### 2.2.1 ARTIFICIAL NEURAL NETWORK

An artificial neural network is similar to the network of neurons in human brain. Using this the computer will be able to learn things and make decisions like a human. Regular computers are programmed such that they behave like interconnected brain cells. It communicates using electrical signals. Some of its uses are classification of information, clustering data and prediction of outcomes.

Eg:- If we take a handwritten character – say alphabet 'A' – it can have different forms when it is written in English by many people.

As shown below, whichever way they write we can understand the character, because one already knows how the character looks like. This concept can be compared to our neural network system.

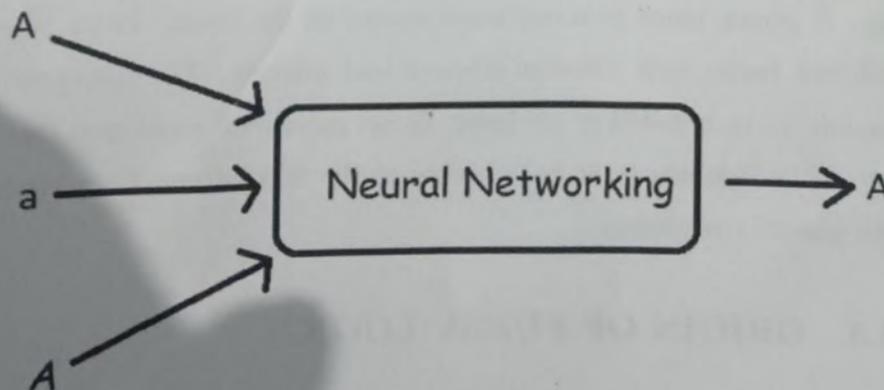


Figure 2.1: Example of Neural Networking

### 2.3. ORIGIN OF FUZZY LOGIC

How is neural network related to mathematics? The main elements of neural network are neurons and synapses, both in charge of computing mathematical operations. Neural networks are nothing but a series of mathematical computations. Each synapse holds a weight, while each neuron computes a weighted sum using input data and synapses weights.

#### 2.2.2 FUZZY LOGIC

Fuzzy logic is a multivalued logic system in which the truth value of variables will be any real number in  $[0,1]$ . It is the computing approach based on degrees of truth rather than the usual true or false method (1 or 0). Fuzzy logic has intermediate values that present partial truth or false. It can have the truth values like partially true/false, rather than completely true/false values. It considers all available information, which can be certain or uncertain information and make the best possible output from it.

#### 2.2.3 EVOLUTIONARY COMPUTATION

Computational systems that study ideas from natural evolution is called Evolutionary computation. It's actually a bunch of troubleshooting techniques with principles based on genetic inheritance and natural selection. Evolutionary computation is usually implemented on computer systems that are used to solve problems, implementing techniques such as evolutionary algorithms, differential evolution, genetic algorithms and harmony search.

The mathematics behind the working of evolutionary computing is its sequence representation as every chromosome is a sequence of genes.

Eg:- A person plans to invest some money in the bank. There is availability of different banks with different schemes and policies. The individual decides the amount to be invested in the bank, to get maximum profit and there are certain criteria for that which can be overcome by the "Evolutional Computing" algorithm like genetic computing.

### 2.3 ORIGIN OF FUZZY LOGIC

There have been a lot of discussions about the uncertainty of events and occurrences, the imprecision of outcomes and they are also much into philosophical

debates. Boolean logic was insufficient to cope with this ambiguity. Hence there was a necessity of something which was beyond the two-valued logic. The search and discovery began in the early beginning of 20th century. Emil Leon Post, in 1921, had introduced more than 2 truth logical degrees. Later, Jan Lukasiewicz Clarence Irving Lewis presented the 3 valued logic. In 1930, Lukasiewicz and Alfred Tarski had introduced the 0-valued version, sometimes called the Lukasiewicz-Tarski logic. Later in 1945, Grigore Constantin Moisil reformulated it in an axiomatic algebraic form, and extended to n-valued logic.

The term fuzzy logic was coined in 1965. The suggestion of the ambiguous set theory was by the great mathematician and computer scientist Lotfi Zadeh. However, vague logic has been studied since the 1920s as a logic of infinite value - especially Lukashevich and Tarski. The vague logic is based on people's observation make decisions based on inaccuracies and numericities in formation. These models have the ability to identify, represent, manage, interpret use and obscure data and information.

## Chapter 3

# FUZZY LOGIC

---

### 3.1 CRISP LOGIC v/s FUZZY LOGIC

The study of principles and methods of reasoning is called *Logic*. Propositions are statements that must either be true or false. And the logic behind such propositions is called classical logic/crisp logic. It is the same as Boolean logic, which gives the output either 0 or 1. The statement is either true (1) or false (0). We have a part of this logic, called propositional logic, where there are combinations of variables – called logic variables instead of arbitrary propositions and these variables assume true/false value only when it takes particular proposition. Every proposition has a negation. So, the negation of truth value "true" is false (and vice-versa).

But at times, the two-valued logic need not be sufficient to give a satisfactory truth value to a given statement. Hence an *Intermediate* value was introduced to the two-valued logic.

Eg:- We can take true = 1, false = 0 and the intermediate value =  $\frac{1}{4}$ . And the negation values can be taken as:  $\text{neg}(0) = 1$ ,  $\text{neg}(1) = 0$  and  $\text{neg}(\frac{1}{4}) = \frac{3}{4}$ .

Later on, n-valued logic was derived, which had n truth values, where  $n \geq 2$  and lie on [0,1]. Such a logic is called *Fuzzy Logic*. It is a continuous valued logic with a membership function.

## 3.2 FUZZY PROPOSITION

A statement 's' which acquires a fuzzy truth value  $T(s)$  is called a *Fuzzy proposition*. The truth value of fuzzy proposition depends on its degree of truth and its value lies in  $[0,1]$ .

Eg:- Let 's' be a statement: The weather is fine.

Then the values of  $T(s)$  can be

$T(s) = 1$ , if s is absolutely true.

$T(s) = 0.7$ , if s is partially true.

$T(s) = 0$ , if s is totally false.

This range of the truth values is actually the main difference between classical propositions and fuzzy propositions. Because, the classical propositions can either be true or false, but for fuzzy propositions, the true or false value is a matter of degree. In other words, the truth and falsity of classical propositions are expressed by - 1 and 0, respectively, whereas in fuzzy propositions, the degree of truth is expressed by a number in the unit interval  $[0, 1]$ . Fuzzy propositions are mainly of four types:

1. Unconditional and unqualified propositions
2. Unconditional and qualified propositions
3. Conditional and unqualified propositions
4. Conditional and qualified propositions

### 3.2.1 UNCONDITIONAL AND UNQUALIFIED PROPOSITIONS

The canonical form of this type of fuzzy proposition is -

$$p: T \text{ is } F$$

Where, T is a variable which takes value t from a universal set U. F is a fuzzy set on U that represents a given inaccurate predicate such as fast, low, tall etc.

Eg:- Let 's' be a statement: The tree is tall.

Then the values of  $T(s)$  can be -

$T(s)=1$ , if s is absolutely true

$T(s)=0.6$ , if s is partly true

$T(s)=0$ , if s is totally false

where,  $T(s) = \mu_F(s)$  - the membership grade function that indicates the degree of truth of  $v$  belongs to  $F$ . Its value ranges from 0 to 1.

### 3.2.2 UNCONDITIONAL AND QUALIFIED PROPOSITIONS

The canonical form of this type of fuzzy proposition is -

$$p: V \text{ is } F \text{ is } S$$

where,  $V$  and  $F$  have the same meaning and  $S$  is a fuzzy truth qualifier.

Eg:- We can have a statement  $p$ : The tree is tall is very true.

### 3.2.3 CONDITIONAL AND UNQUALIFIED PROPOSITIONS

The canonical form of this type of fuzzy proposition is -

$$p: \text{if } X \text{ is } A, \text{ then } Y \text{ is } B$$

where,  $X, Y$  are variables in universes  $U_1$  and  $U_2$  and  $A, B$  are fuzzy sets on  $X, Y$ .

Eg:-  $p$ : If the tree is tall then risk is low.

### 3.2.4 CONDITIONAL AND QUALIFIED PROPOSITIONS

The canonical form of this type of fuzzy proposition is -

$$p: (\text{if } X \text{ is } A, \text{ then } Y \text{ is } B) \text{ is } S$$

where, all variables have same meaning as previously declared.

Eg:-  $p$ : If the tree is tall than risk is low is true.

## 3.3 FUZZY QUANTIFIERS

The scope of fuzzy propositions can be extended using fuzzy quantifiers. *Fuzzy Quantifiers* can be defined as the fuzzy numbers which take part in fuzzy propositions and they make use of characteristic linguistic terms.

There are two different types of fuzzy quantifiers: -

1. Fuzzy quantifiers of first kind
2. Fuzzy quantifiers of second kind

### 3.3.1 FUZZY QUANTIFIERS OF FIRST KIND

These are defined on the set of real numbers. The characteristic linguistic terms they make use are - "about 20", "much more than 80", "at least about 6", etc. These are also known as Absolute Fuzzy Quantifiers.

There are 2 types of propositions which involve the fuzzy quantifiers of first kind:

i. The first type of propositions will be of the form -

p: There are Q i's in I such that  $V(i)$  is F.

Here Q - is a first kind quantifier, i - is an individual from a given set I,  $V(i)$  - is a variable associated to the individual I which takes values from a universe - E, and F - is a fuzzy set defined over the universe E.

Eg: - Let us take a statement -

p: There are about 7 athletes in the ground, whose running speed is very high.

In the given proposition,

- Q - fuzzy number capturing the linguistic term - about 7.
- i - the athletes
- $V(i)$  - running speed of athlete i.
- F - Fuzzy set that represents a high speed in running.

ii. The second type of propositions will be of the form -

p: There are Q i's in I such that  $V_1(i)$  is  $F_1$  and  $V_2(i)$  is  $F_2$ .

Here  $V_1, V_2$  - are variables that take values from sets  $E_1, E_2$  respectively, I - is an index set by which distinct measurements of variables  $V_1, V_2$  are identified, Q - is a fuzzy number on real numbers, and  $F_1, F_2$  - are fuzzy sets on  $E_1, E_2$  respectively.

Eg:- Let us take a statement

p: There are about 7 athletes in the ground, whose running speed is very high and whose fitness is perfect.

In the given proposition,

- Q - fuzzy number capturing the linguistic term - about 7.
- i - the athletes

- $I$ : the index set which labels the athletes in the - ground
- $V_1(i)$  - running speed of athlete  $i$ .
- $V_2(i)$  - fitness of athlete  $i$ .
- $F_1$  - Fuzzy set that represents a high speed in running.
- $F_2$  - Fuzzy set that represents perfect fitness of athlete.

### 3.3.2 FUZZY QUANTIFIERS OF SECOND KIND

These are defined on the interval  $[0, 1]$ . The characteristic linguistic terms they make use are - "almost all", "about half", "most", etc.

These are also known as Relative Fuzzy Quantifiers.

## 3.4 FUZZY RULES OF INFERENCE

We make use of *fuzzy rules* in fuzzy logic systems to infer an output based on the input variables. It is a conditional statement of the form:

If  $P$  then  $Q$

Where  $P = x$  is  $A$ , which is the antecedent part and  $Q = y$  is  $B$ , which is the consequent part. Here  $A$  and  $B$  are the linguistic values which may be defined on the fuzzy sets of  $X$  and  $Y$  respectively - which are universe of discourse.

Eg:- If you study well, then you will pass the exam.

The most important rules of inference are

- i. Modus ponens
- ii. Modus tollens

### 3.4.1 MODUS PONENS

A modus ponens rule is in the form:

Premise:  $x$  is  $A^*$ .

Implication: IF  $x$  is  $A$  THEN  $y$  is  $B$ .

Conclusion:  $y$  is  $B^*$ .

### 3.4.2 MODUS TOLLENS

A modus tollens rule is in the form:

Premise:  $x$  is  $B^*$ .

Implication: IF  $x$  is  $A$  THEN  $y$  is  $B$ .

Conclusion:  $y$  is  $A^*$ .

The fuzzy rules can either be decomposed or aggregated.

### 3.4.3 DECOMPOSITION

The compound fuzzy rules are decomposed to  $n$  number of simple canonical rules.

There are 4 methods of decomposition:

- i. Multiple Conjunctive Antecedent
- ii. Multiple Disjunctive Antecedent
- iii. Conditional Statement
- iv. Nested If-Then Rules

#### MULTIPLE CONJUNCTIVE ANTECEDENT

Consider the fuzzy rule of the form:

If  $x$  is  $A_1, A_2, A_3, \dots, A_n$ , then  $y$  is  $B$ .

When we have such a compound rule, we apply the method of *Multiple Conjunctive Antecedent* and take the conjunction of each antecedent.

i.e.

$$A_m = A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n \quad (3.1)$$

We know that conjunction  $\rightarrow$  taking the minimum of membership value.

$$\mu_{A_m}(x) = \min [\mu_{A_1}(x), \mu_{A_2}(x), \mu_{A_3}(x) \dots \mu_{A_n}(x)]$$

Then we substitute the corresponding antecedent value, get the consequent and hence decompose the fuzzy rule as:

If  $x$  is  $A_m$  then  $y$  is  $B_m$ .

#### MULTIPLE DISJUNCTIVE ANTECEDENT

Consider the fuzzy rule of the form: If  $x$  is  $A_1, A_2, A_3, \dots, A_n$ , then  $y$  is  $B$ .

When we have such a compound rule, we apply the method of *Multiple Disjunctive*

*Antecedent* and take the disjunction of each antecedent.

i.e.

$$A_m = A_1 \cap A_2 \cap A_3 \cap \dots A_n \quad (3.2)$$

We know that disjunction  $\rightarrow$  taking the maximum of membership value.

$$\mu A_m (x) = \max [\mu A_1 (x), \mu A_2 (x), \mu A_3 (x) \dots \mu A_n (x)]$$

Then we substitute the corresponding antecedent value, get the consequent and hence decompose the fuzzy rule as:

If  $x$  is  $A_m$  then  $y$  is  $B_m$ .

#### CONDITIONAL STATEMENTS

Consider the statement:

If  $A_1$  then  $B_1$  else  $B_2$ .

This form of fuzzy rule implies that if a given antecedent  $A_1$  is true, then a given consequent  $B_1$  is true or else the consequent  $B_2$  is true. This compound rule is a *Conditional Statement*. The given compound rule can be decomposed to a simple canonical form using an OR operator as:

If  $A_1$  then  $B_1$  OR If not  $A_1$  then  $B_2$ .

#### NESTED IF-THEN RULES

Consider the statement:

If  $A_1$  then (if  $A_2$  then  $B_1$ ).

This form of fuzzy rule implies that if a given antecedent  $A_1$  is true, and then if a given antecedent  $A_2$  is true, then the consequent  $B_2$  is true. This compound rule is a *Nested If-Then Rules*. The given compound rule can be decomposed to a simple canonical form using an AND operator as:

If  $A_1$  AND  $A_2$  then  $B_1$ .

#### 3.4.4 AGGREGATION

Aggregation is in others words the process of combining. When we have complex fuzzy rules, we aggregate or combine and generate the output rule. The given system can be designed in 2 different ways:

- i) Conjunctive System of Rules
- ii) Disjunctive System of Rules

#### CONJUNCTIVE SYSTEM

Here the system is designed in such a way that it gives equal importance to all the rules and hence all of them must be satisfied for the system to work. Consider that we have  $n$  fuzzy rules, and  $y_1, y_2, y_3, \dots, y_n$  are the  $n$  consequences of those rules. Then we take

$$y_m = y_1 \cap y_2 \cap y_3 \dots y_n \quad (3.3)$$

and hence  $y_m$  is the aggregate output. We can also find the membership value for that particular aggregate output by:

$$\mu_{y_m}(y) = \min [\mu_{y_1}(y), \mu_{y_2}(y), \mu_{y_3}(y) \dots \mu_{y_n}(y)]$$

as we know that conjunction  $\rightarrow$  taking the minimum of membership value.

#### DISJUNCTIVE SYSTEM

Here the system is designed in such a way that equal importance is not given to all the rules and hence all of them need not be satisfied for the system to work. Consider that we have  $n$  fuzzy rules, and  $y_1, y_2, y_3, \dots, y_n$  are the  $n$  consequences of those rules. Then we take

$$y_m = y_1 \cup y_2 \cup y_3 \dots y_n \quad (3.4)$$

and hence  $y_m$  is the aggregate output. We can also find the membership value for that particular aggregate output by:

$$\mu_{y_m}(y) = \max [\mu_{y_1}(y), \mu_{y_2}(y), \mu_{y_3}(y) \dots \mu_{y_n}(y)]$$

as we know that disjunction  $\rightarrow$  taking the maximum of membership value.

## Chapter 4

# FUZZY LOGIC CONTROLLER AND ITS APPLICATION

---

### 4.1 Fuzzy Logic Controller

*Fuzzy controllers* work by utilizing knowledge drawn from human operators. There are four modules in fuzzy controller: a fuzzy rule base, a fuzzy inference engine, fuzzification and defuzzification modules. The main steps followed by a Fuzzy controller are:

- i. Taking the values of all the variables representing the relevant conditions of the controlled process.
- ii. Fuzzification - that is the conversion of these values to fuzzy values so that they represent the uncertainties.
- iii. Evaluation of control rules in the fuzzy rule base, by the inference engine, using these fuzzified values.
- iv. The result obtained is a fuzzy set, defined on a universe  $f$ .
- v. Defuzzification - that is the conversion of these values to a single crisp value.

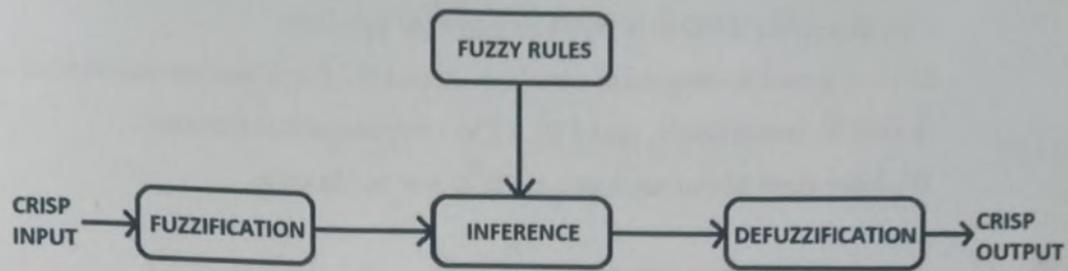


Figure 4.1: Fuzzy Logic Controller

### 4.1.1 Approaches

There are mainly 2 types of inference approaches:

- i. Mamdani Approach
- ii. Takagi Sugeno's Approach

#### MAMDANI APPROACH

It was used in 1975, by Professor Ebrahim Mamdani to control a steam engine and boiler combination. It is one of the most commonly used inference approaches and it's a linguistic fuzzy modelling approach where the values are words/sentence type.

Eg:- very true, partially true, untrue etc.

It is processed in 4 steps:

- i. Fuzzification of input variables.
- ii. Rule evaluation.
- iii. Aggregation of the rule outputs.
- iv. Defuzzification.

#### TAKAGI SUGENO'S APPROACH

It was introduced by Michio Sugeno. Sugeno fuzzy inference is similar to the Mamdani approach. The change in this approach is the change in rule consequent (resultant). It is a precise fuzzy modelling where more numerical terms are used compared to Mamdani approach, and hence giving better results. That is, in the

fuzzy rule, a function of the input variables is being used, instead of the fuzzy set.  
Eg:- If  $x$  is  $A$ , AND  $y$  is  $B$ , Then  $z$  is  $f(x, y)$ .  
Here  $x, y$  and  $z$  - linguistic variables,  $A$  and  $B$  - fuzzy sets on universe of discourse  $X$  and  $Y$ , respectively; and  $f(x, y)$  is a mathematical function.  
We have used Mamdani's approach in our application.

## 4.2 APPLICATION

Washing machines are common household appliances in India. The washing machines that efficiently control the wash time is considered vital. Conventional, proportional, integral and differential [PID] controllers have proven to be less capable. Therefore, fuzzy logic enables designers to control complex systems more effectively than traditional approaches. The users of washing machines have been facing the problem of selecting the duration of wash time based on the amount of dirtiness on clothes. In order to overcome these problems, fully-automatic washing machines, which are based on fuzzy logic, offers the advantage of performance, simplicity and less cost.

We have 5 steps in the application:

- i. Identifying Linguistic input and output variables.
- ii. Defining Membership functions for input and output variables.
- iii. Creating a Rule base.
- iv. Rule Evaluation.
- v. Defuzzification.

Here, we have 4 input variables and 5 output variables to get correct wash time. We consider our input as dirt and output as wash-time. So, the input variables will be:

Very Low Dirt (VLD), Low Dirt (LD), High Dirt (HD), Very High Dirt (VHD) and similarly the output variables will be Very Short (VS), Short (S), Medium (M), High (H), Very High (VH).

For Dirt:

$$\mu_{VLD}(x) = \frac{(30 - x)}{30} \quad \text{when } 0 \leq x \leq 30 \quad (4.1)$$

$$\mu_{LD}(x) = \frac{(x)}{30} \quad \text{when } 0 \leq x \leq 30 \quad (4.2)$$

$$\mu_{LD}(x) = \frac{(60 - x)}{30} \quad \text{when } 30 \leq x \leq 60 \quad (4.3)$$

$$\mu_{HD}(x) = \frac{(x - 30)}{30} \quad \text{when } 30 \leq x \leq 60 \quad (4.4)$$

$$\mu_{HD}(x) = \frac{(90 - x)}{30} \quad \text{when } 60 \leq x \leq 90 \quad (4.5)$$

$$\mu_{VHD}(x) = \frac{(x - 60)}{30} \quad \text{when } 60 \leq x \leq 90 \quad (4.6)$$

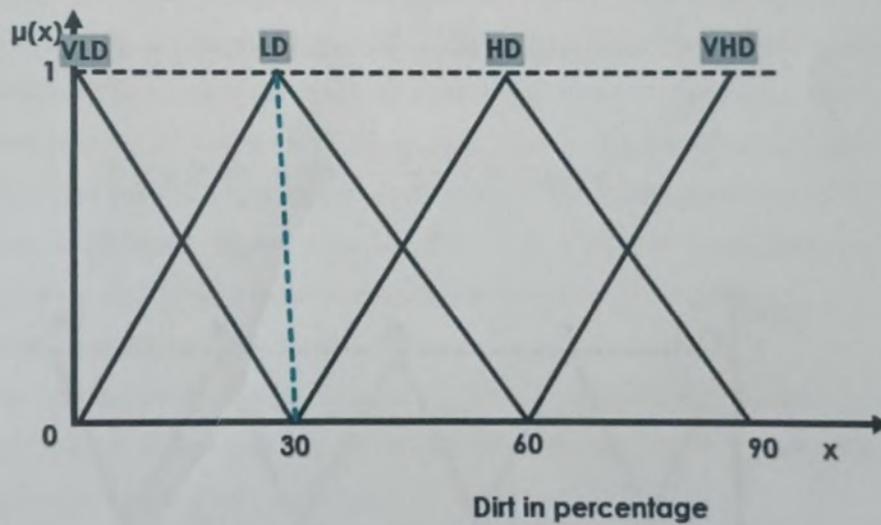


Figure 4.2: Dirt percentage graph

For Wash-time:

$$\mu_{VS}(y) = \frac{(20 - y)}{20} \quad \text{when } 0 \leq y \leq 20 \quad (4.7)$$

$$\mu_S(y) = \frac{(y)}{20} \quad \text{when } 0 \leq y \leq 20 \quad (4.8)$$

$$\mu_S(y) = \frac{(40 - y)}{20} \quad \text{when } 20 \leq y \leq 40 \quad (4.9)$$

$$\mu_M(y) = \frac{(y - 20)}{20} \quad \text{when } 20 \leq y \leq 40 \quad (4.10)$$

$$\mu_M(y) = \frac{(60 - y)}{20} \quad \text{when } 40 \leq y \leq 60 \quad (4.11)$$

$$\mu_H(y) = \frac{(y - 40)}{20} \quad \text{when } 40 \leq y \leq 60 \quad (4.12)$$

$$\mu_H(y) = \frac{(80 - y)}{20} \quad \text{when } 60 \leq y \leq 80 \quad (4.13)$$

$$\mu_{VH}(x) = \frac{(y - 60)}{20} \quad \text{when } 60 \leq y \leq 80 \quad (4.14)$$

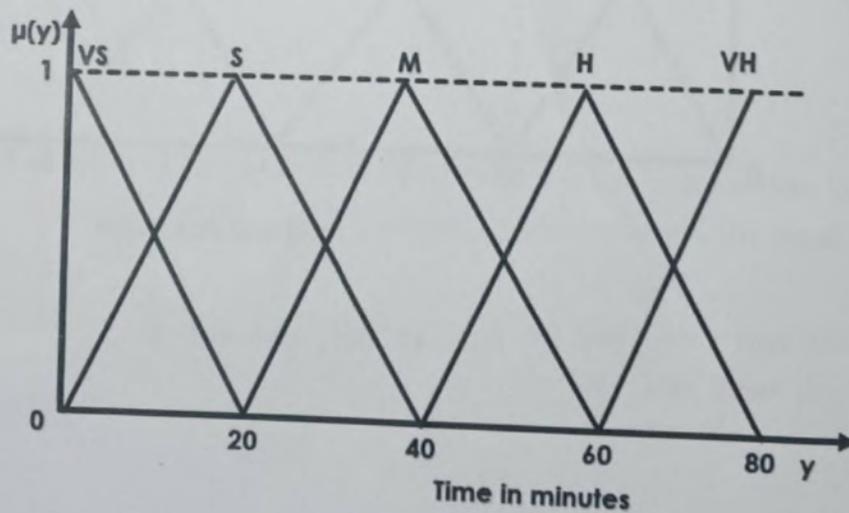


Figure 4.3: Wash-time graph

We could clearly get the membership functions of both the dirt and wash-time from the above graphs. We have also framed the following fuzzy rule base –

Table 4.1: Fuzzy Rule Table

DIRT	WASH-TIME
VLD	VS/S
LD	M
HD	H
VHD	H/VH

Rule 1: If the Dirtiness of cloth is Very Low then the Wash-time is Very Short or Short.

Rule 2: If the Dirtiness of cloth is Low then the Wash-time is Medium.

Rule 3: If the Dirtiness of cloth is High then the Wash-time is High.

Rule 4: If the Dirtiness of cloth is Very High then the Wash-time is High or Very High.

We have imprecise rules defined and hence they too are not crisp but fuzzy values. The 4 input parameters are detected by the sensors and then fuzzified as per the membership function of the respective variables. The corresponding rules are then evaluated and our system is a disjunctive system. So, the Mamdani approach is applied and rules are aggregated disjunctively. From this we defuzzify by substituting the obtained degree of membership value in the corresponding equations of wash-times and hence obtain the duration of wash-time required.

Now let us consider an example:

Let us assume 40 % of dirt. Comparing to the membership functions, we understand that 40 % lies between 30 % and 60 % of dirt. Taking the corresponding degrees of memberships, we have:

$$\mu_{LD}(x) = \frac{(60 - x)}{30} \quad \text{when } 30 \leq x \leq 60 \quad (4.15)$$

$$\mu_{HD}(x) = \frac{(x - 30)}{30} \quad \text{when } 30 \leq x \leq 60 \quad (4.16)$$

Substituting  $x = 40$ , for low dirt and high dirt values in the equations of degrees

of memberships, we get:

$$\mu_{LD}(x) = \frac{(60 - 40)}{30} = \frac{2}{3} \quad (4.17)$$

$$\mu_{HD}(x) = \frac{(40 - 30)}{30} = \frac{1}{3} \quad (4.18)$$

From the fuzzy rule base, we know that

Low Dirt  $\rightarrow$  Medium Wash-time AND High Dirt  $\rightarrow$  High Wash-time.

With Mamdani approach, we aggregate the rules, by disjunctive method:

Rule 2: If the Dirtiness of cloth is Low then the Wash-time is Medium( $y_1$ ).

Rule 3: If the Dirtiness of cloth is High then the Wash-time is High( $y_2$ ).

$$y_m = y_1 \cup y_2 \quad (4.19)$$

$$\mu_{y_m}(y) = \max [\mu_{y_1}(y), \mu_{y_2}(y)] = \max \left[ \frac{2}{3}, \frac{1}{3} \right] = \frac{2}{3} \quad (4.20)$$

$\rightarrow$  which is the degree of membership of low dirt.

Therefore, it requires medium wash-time duration.

Now substituting

$$\mu_M(y) = \frac{2}{3}, \text{ in the equations:} \quad (4.21)$$

$$\mu_M(y) = \frac{(y - 20)}{20} \text{ when } 20 \leq y \leq 40 \quad (4.22)$$

$$\frac{(y - 20)}{20} = \frac{2}{3} \rightarrow y = \frac{160}{3} \quad (4.23)$$

$$\mu_M(y) = \frac{(60 - y)}{20} \text{ when } 40 \leq y \leq 60 \quad (4.24)$$

$$\frac{(60 - y)}{20} = \frac{2}{3} \rightarrow y = \frac{140}{3} \quad (4.25)$$

Taking the average, we get

$$\frac{\frac{160}{3} + \frac{140}{3}}{2} = 50 \text{ minutes} \quad (4.26)$$

## Chapter 5

# CONCLUSION

---

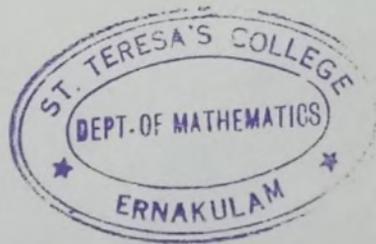
We have seen the application of Fuzzy Logic in washing machines. In the washing machine, various sensors are used to assess the different conditions inside the obscure logic machine and adjust its operation accordingly.

Similarly, it has numerous applications in facial pattern recognition, air conditioners, vacuum cleaners, antiskid braking systems, transmission systems, control of subway systems and unmanned helicopters and a number of knowledge-based applications for multi-objective optimization of power systems based on certain pre-defined conditions. Fuzzy logic is useful for many people involved in research and development. It can also be found in many engineering and scientific works.

Fuzzy logic will be applied to a wide range of systems and products in the future. Due to technological advancement and digital transformation, there is a large scope for fuzzy logic in the world.

## REFERENCES

- [1] George J. Klir/Bo Yuan; Fuzzy sets and Fuzzy Logics – Theory and applications; Published by Prentice Hall PTR (1995)
- [2] Timothy J. Ross; Fuzzy Logic with Engineering Applications, Third Edition; John Wiley Sons, Limited (2010)
- [3] Rajjan Shinghal; Introduction to Fuzzy Logic; Published by PHI learning Private Limited (2013)
- [4] Hung T. Nguyen, Carol L. Walker, Elbert A. Walker; First Course in Fuzzy Logic, Fourth Edition; Published by Chapman and Hall/CRC (November 13, 2018)
- [5] D. S Hooda, Vivek Raich; Fuzzy Logic Models and Fuzzy Control - An Introduction; Alpha Science International Limited (2017)
- [6] <https://youtube.com/playlist?list=PLYwpaLSFmcCPU18mAnb4g1oExKd0n4Gv>
- [7] <https://en.m.wikipedia.org/wiki/Fuzzylogic>
- [8] <https://en.m.wikipedia.org/wiki/>
- [9] <https://youtu.be/AtrWDJQQ72A>



Project Report

On

**MATHEMATICS IN CRIME SOLVING AND  
CRIME CONTROL**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

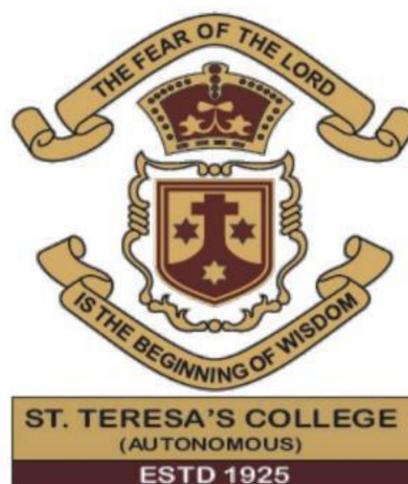
*by*

**LIDHIA K.S**

(Register No. AB19BMAT038)

*Under the Supervision of*

**DR .SUSAN MATHEW PANAKKAL**

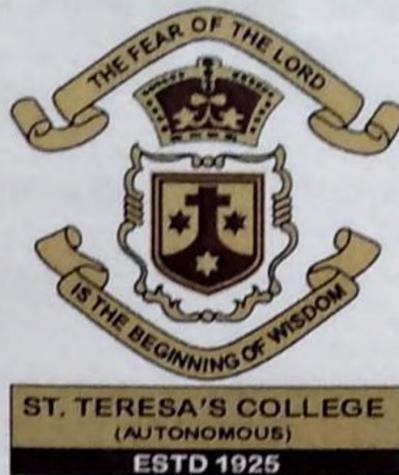


**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**FEBRUARY 2022**

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



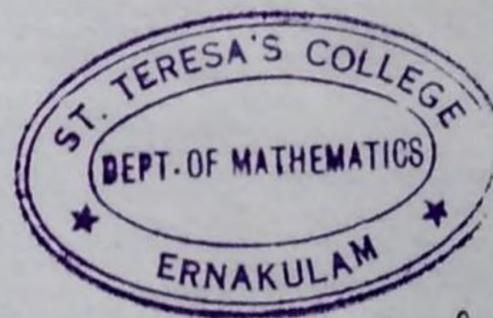
CERTIFICATE

This is to certify that the dissertation entitled, **MATHEMATICS IN CRIME SOLVING AND CRIME CONTROL** is a bonafide record of the work done by Ms. **LIDHIA K.S** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date:

Place: Ernakulam

*Susan*  
Dr.Susan Mathew Panakkal  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



*Ursala*  
Dr.Ursala Paul  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1: *Jeenu*

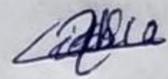
2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Dr. Susan Mathew Panakkal, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 08/03/2022



LIDHIA K.S

AB19BMAT038

## ACKNOWLEDGEMENT

Firstly I would like to thank god almighty for giving me grace to execute this project. I express my gratitude to our project guide **Dr.Susan Mathew Panakkal**,Department of Mathematics and Statistics, St Teresa's College, Ernakulam.I would like to mention the support and help offered by **Dr.Ursala Paul**,HOD,Department of Mathematics,St Teresas's College,Ernakulam.

I would also like to thank all other teachers of the Department of Mathematics and Statistics for their immense support throughout the course of completing this project.

I also thank my parents, friends, especially **My group Members** and all those who have given me the moral support and guidance which helped me to complete my project. This project would have been not possible without help of the above mentioned people.

Ernakulam

Date: 08 | 03 | 20 22

LIDHIA K.S

AB19BMAT038

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 PREREQUISITES</b>	<b>1</b>
1.1 MATHEMATICAL MODELLING . . . . .	1
1.1.1 INTRODUCTION . . . . .	1
1.1.2 DEFINITION . . . . .	1
1.1.3 ADVANTAGE AND APPLICATION . . . . .	2
1.1.4 STEPS INVOLVED . . . . .	2
1.2 GRAPH THEORY . . . . .	4
1.2.1 INTRODUCTION . . . . .	4
1.2.2 DEFINITION . . . . .	4
1.2.3 GRAPH THEORY IN CRIME . . . . .	4
1.2.4 APPLICATION OF GRAPH COLOURING . . . . .	5
1.3 FORENSIC STATISTICS . . . . .	6
1.3.1 INTRODUCTION . . . . .	6
1.3.2 CONCEPTS INVOLVED . . . . .	6
<b>2 CRIME SOLVING USING MATHEMATICAL MODELLING</b>	<b>11</b>
2.1 ROSSMO MODEL . . . . .	11
2.1.1 INTRODUCTION . . . . .	11
2.1.2 ROSSMO FORMULA . . . . .	12
2.1.3 DRAWBACKS AND SUGGESTED MODIFICATION . . . . .	14

<b>3</b>	<b>CRIME PROBLEM USING GRAPH THEORY AND FORENSIC STATISTICS</b>	<b>15</b>
3.1	GRAPH THEORY . . . . .	15
3.2	FORENSIC STATISTICS . . . . .	19
<b>4</b>	<b>CONTROLLING CRIME BY ANALYSING DATA MANIPULATION</b>	<b>23</b>
4.1	BENFORD'S LAW . . . . .	23
4.1.1	INTRODUCTION . . . . .	23
4.1.2	THEORY . . . . .	24
4.1.3	BENFORD'S LAW CONDITIONS . . . . .	24
4.1.4	DISTRIBUTION OF FIRST DIGIT ACCORDING TO BENFORD'S LAW . . . . .	25
4.2	APPLICATIONS . . . . .	27
4.2.1	OTHER APPLICATIONS . . . . .	28
	<i>REFERENCES</i> . . . . .	29

# Chapter 1

## PREREQUISITES

---

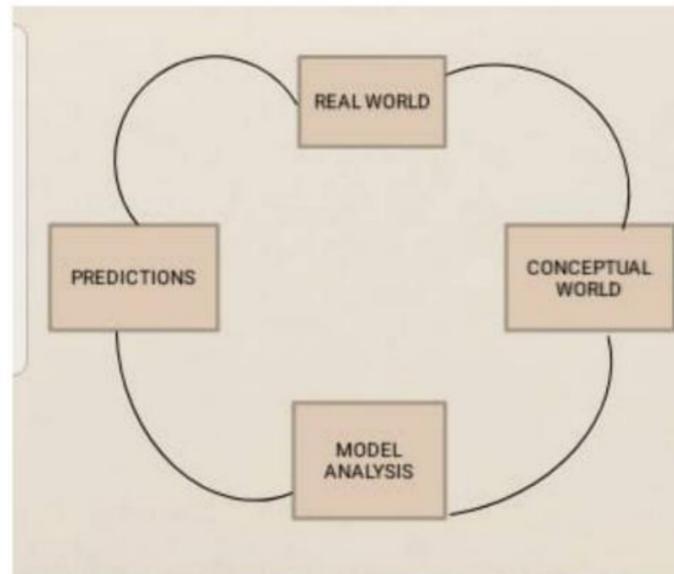
### 1.1 MATHEMATICAL MODELLING

#### 1.1.1 INTRODUCTION

Mathematical Modelling is an important branch of mathematics. 'Model' is basically an imitation of real world objects. Modelling is a miniature representation of some object or a demonstration for reference of an object to be made. There are two types of models prototypic(physical model) and schematic(charts,graph,diagram etc). The manipulation of modelling and implementing it in mathematics helped in the betterment of various fields.

#### 1.1.2 DEFINITION

Representation of real objects and phenomenon in mathematical language is stated as mathematical modelling. The subjects that cannot be studied or is difficult to study using normal theoretical methodology can be done using mathematical modelling



### 1.1.3 ADVANTAGE AND APPLICATION

Mathematical modelling comes to light when the theoretical method become :

- Time consuming
- Expensive
- Often dangerous
- Simply impossible

Mathematical modelling basically deals with analysing, predicting and insight of real world knowledge. As mentioned earlier it's application is beyond limit it is applicable in the field of Epidemiology, biological transpore , vehicular traffic etc.Consider the current situation of covid 19, it forced us to trace multiple things and track down and even predict the spread and new waves . Evidently, this won't be possible by simple experimental method as the real world object that we are dealing with is quite large.Hence, modelling is an unavoidable and powerful tool.

### 1.1.4 STEPS INVOLVED

- A problem from the real world is taken in account and is studied to make certain assumptions involving the parameters affecting the considered problem.
- Analytical and numerical analysis of the assumptions are made.
- A model is predicted for the same

- Both experimental and Intuitive validation is done.
- If valid, the model is accepted or else the process is repeated

Mathematical modelling is used in the project to study an existing model named Rossmo Model introduced by **Dr.Kim Rossmo**. The basic mathematical concept used in the generation of this model is **probability**. This was used to hunt down a serial rapist bringing Dr.Rossmo to limelight.

## 1.2 GRAPH THEORY

### 1.2.1 INTRODUCTION

In mathematics, Graph theory is a special branch of mathematics that deals with the pairwise relation of one object to other. A graph consists of vertices also known as nodes which are interconnected by edges. Graphs is a crucial field of study in mathematics.

### 1.2.2 DEFINITION

Graph Theory is used to study the relationship between different objects. A Graph  $G(V,E)$  denotes the collection of graph with vertex  $V(G)$  and edge  $E(G)$ .

### 1.2.3 GRAPH THEORY IN CRIME

We can analyse a crime using three elements of graph theory such as the vertex set, the edge set, and the incidence function that relates edges to vertices. The mathematics of graph theory explains that an edge can have an interconnection from one vertex to another if there is any connection between the entities. The prior step in crime solving is to locate different elements involved in the incident. These elements will be represented as vertices. When any two elements have any connection that connection is represented as an edge. The edge not only connects the vertices but acts as the initiator of that connection which results in graph with parallel edges.

#### 1.2.4 APPLICATION OF GRAPH COLOURING

In graph theory, graph colouring is used to label the graph. The vertices, Edges and faces are denoted using colours without the adjacent of the above being labelled with same colours. This method of colouring is stated as *vertex colouring*, *edge colouring* and *face colouring* respectively. In crime solving this method can be used to interpret the connection between the collected evidences. An illustration of the same is explained later in this paper.

## 1.3 FORENSIC STATISTICS

### 1.3.1 INTRODUCTION

Forensic statistics refers to the application of probability, mathematical techniques and statistical concepts to the forensic data and the scientific evidence collected. Mathematics is an unavoidable part of Forensic science. The application of mathematics and statistics makes the scientific analysis of evidences methodical. Forensic investigators collect the evidences, analyse and document it using statistical methods.

Biological evidences such as DNA evidence, blood samples, hair samples, fingerprints, etc found at the crime scene are examined using various probability models to draw inferences from them. DNA profiling, fingerprint analysis, blood sample analysis, and statement analysis are major functionalities of forensic science that are deterministic in solving a crime. Forensic Statistics is widely used by Forensic Investigators and Legal practitioners to analyse statements and test the significance of evidences.

### 1.3.2 CONCEPTS INVOLVED

There are different mathematical and statistical techniques used to solve crimes. Some of the main mathematical principles used in crime solving are:-

- TRIGONOMETRY
- EXPONENTIAL & LOGARITHMIC FUNCTIONS
- CONDITIONAL PROBABILITY
- BAYES' THEOREM
- LIKELIHOOD RATIO

## TRIGONOMETRY

Trigonometry studies the relationship between the angles and the length of sides of a right angled triangle. In Crime solving trigonometry is mainly used in Blood Spatter Analysis and it also helps in Ballistic Analysis.

## BLOOD SPATTER ANALYSIS

Blood Spatter Analysis deals with the study of bloodstains. The bloodstains found at the crime scene are analysed by experts. Bloodstains develop a pattern due to the impact of multiple bloodstains originating from a single source, that is, from a gunshot or a stab wound. Working on these patterns, the investigators can trace back to the position of the victim when the crime occurred or the nature of the blood source. Blood Spatter Analysis is conducted in three steps using trigonometry.

1. Calculate the angle of impact after identifying a proper set of bloodstains.
2. The point of convergence of the bloodstains within the plane of the pattern is determined.
3. The point of convergence of the source is determined.

There are two scenarios under consideration to analyse the shape of the blood stain and calculate the angle of impact; first, the impact from a stationary source and second, the impact from a moving source.

**STATIONARY SOURCE:** Bloodstains take a spherical shape under a perpendicular impact. If the angle of impact is less than  $90^\circ$  then the blood drop is going to be elongated and elliptical in shape. The angle of impact  $\theta$  is calculated using the formula

$$\sin \theta = \frac{\text{width}}{\text{length}}$$

**MOVING SOURCE:** Blood drops falling from a moving source like blood falling from the assailant or the moving wounded victim hits the ground with an impact

less than  $90^\circ$ . The components of velocity affect the angle of impact. The effective angle of impact  $\theta$  is given by,

$$\tan \theta = \frac{V_g}{V_h}$$

Where,  $V_h$  is the horizontal component of velocity due to speed.  $V_g$  is the velocity component due to gravitational acceleration and is given by,  $V_y = \sqrt{2gh}$  when  $g$  is the acceleration due to gravity and  $h$  is the drop height.

### **BALLISTIC ANALYSIS**

Firearms involved in a crime are examined to discover evidences in forensic ballistic analysis. A variety of instruments and devices are also used in Ballistic Analysis. Trigonometry appears in the non-mechanical part of Ballistic Analysis. Bullet trajectories refer to the path travelled by the bullet after firing a gun. Studying the bullet trajectories the investigators can determine the position of the criminal at the time of crime and the height of the criminal.

## EXPONENTIAL & LOGARITHMIC FUNCTIONS

An Exponential function  $f(x) = a^x$  shows the relationship between 2 variables where a constant change in the independent variable induces a proportional change in the other variable. Logarithmic functions denoted by  $f(x) = \log_a x$  are the inverse of exponential functions.

The exponential and logarithmic functions are two mathematical functions that play a significant role in forensic science. Rates of heating or cooling, or the rate of metabolizing of alcohol and drugs help to reach conclusions about the time elapsed since death.

## CONDITIONAL PROBABILITY

The probability of an outcome or event happening, given another event has already occurred is defined as Conditional Probability. It is simply the relationship between the occurrence of two events. The theory of conditional probability studies the odds of an event occurring when another event has happened and the chances that both the events are connected. The formula for calculating the conditional probability of two events A and B is,

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$$

$P\left(\frac{A}{B}\right)$  is pronounced as Probability of the occurrence of event A when B has already happened.

The theory of Conditional Probability plays a crucial role in Crime Solving. It has a wide range of application including studying the Witness reliability, Statement analysis and many more. Misinterpreting conditional probabilities can lead to statistical fallacies.

## BAYES' THEOREM

Bayes' theorem is a special case or an application of Conditional Probability. Named after Thomas Bayes, the theorem explains the odds of an event happening in relation to any given conditions. The Bayes' formula can be used to solve complex problems in a well constructed and precise way whereas the formula for conditional probability can be used only for relatively simpler problems. Bayes' formula is given by,

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right) P(A)}{P(B)}$$

Bayes Formula helps to reach precise conclusions regarding the efficiency of evidences, strength of statements provided by the witness, and to check the validity of the assumptions used in solving a case.

## LIKELIHOOD RATIO

In crime solving there are a few terms to be familiar of to understand Likelihood ratio. H is called Hypothesis, mainly two types here;  $H_p$  or Prosecution Hypothesis, and  $H_d$  or Defence Hypothesis and E is the evidence. Likelihood Ratio is defined as the ratio of the probability of E occurring when  $H_p$  is true, to the probability of occurrence of E given  $H_d$ .

$$LR = \frac{P\left(\frac{E}{H_p}\right)}{P\left(\frac{E}{H_d}\right)}$$

If  $LR > 1$ , it is in support of Prosecution Hypothesis,  $LR < 1$ , it is in support of Defence Hypothesis and when  $LR = 1$  it is considered inadmissible. One of the major errors caused due to misinterpreting LR and Conditional probability is Prosecutor's Fallacy.

## Chapter 2

# CRIME SOLVING USING MATHEMATICAL MODELLING

---

### 2.1 ROSSMO MODEL

#### 2.1.1 INTRODUCTION

Rossmo Model is developed by Dr. Kim Rossmo and is an application of mathematical modelling. The mathematical concept involved in the prediction is Probability. This is a blend of two entirely different subjects, mathematics and criminology. Here, the real world problem that we are dealing with is complex i.e., we have to hunt down a serial killer which is a social threat where the simple methods can end up in vain. A criminology concept lies the foundation on which the model is built, which is:

*Criminals neither go too far to commit crime and also they never commit crime near their locality.*

With this assumption he developed a formulae that sums up the probabilities of where the criminal is likely to live. A heat map was produced based on this inference which forms a heat zone of the probable residence of prime suspect. This was used and was succeeded in early 90's to hunt down a serial rapist. Since, we are challenging human being with a theoretical concept it has its drawbacks.

2.1.2 ROSSMO FORMULA

$$P_{ij} = k \sum_{n=1}^{(totalcrimes)} \left[ \frac{\Phi_{ij}}{(|X_i - x_n| + |Y_j - y_n|)} + \frac{(1 - \phi_{ij})(B^{(g-f)})}{(2B - |X_i - x_n| - |Y_j - y_n|)^g} \right]$$

$P_{ij}$  : Probability of the culprit to reside in the considered square

f,g : Constants chosen from past crime to work better

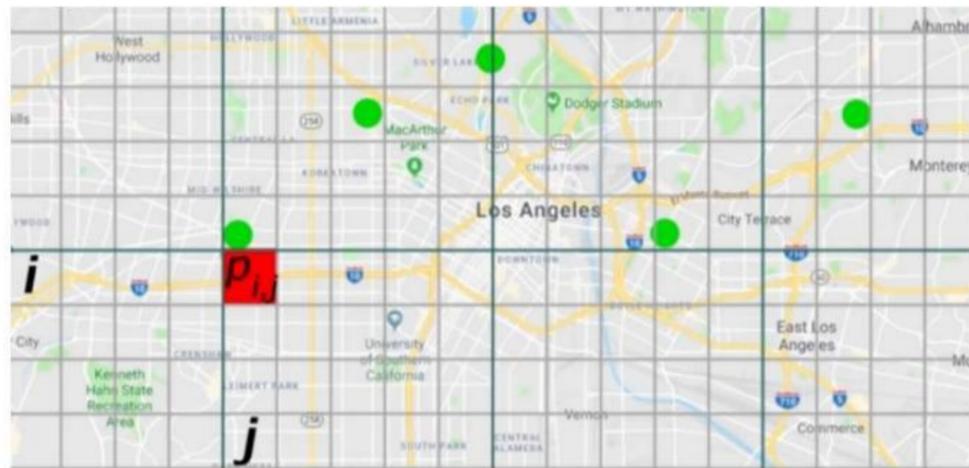
2B : Buffer zone formulated from past crimes.

B : Radius of the buffer zone.

$\phi_{ij}$  : Constant that is used to add weight to the formulae.

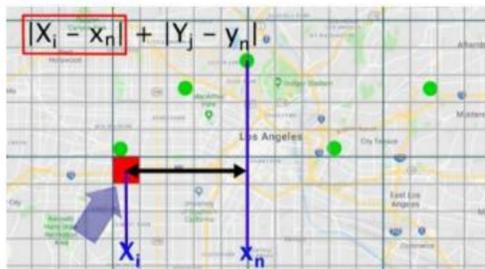
K : Empirically determined constant.

A grid map is taken for study. Here, the map is of Los Angeles. The crime spots are then marked in green. Assume a square in the grid as the residence of the culprit and mark it in Red. The column where the assumed square of the grid is present is marked as 'j' and the row in which it lies is marked as 'i'.  $P_{ij}$  is the probability of criminal to reside in the assumed square.

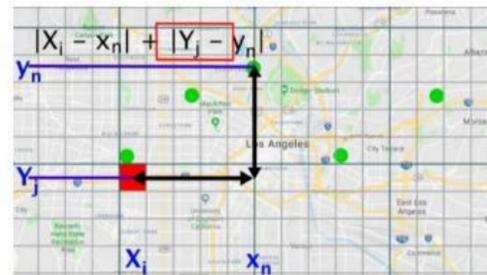


source:<http://brilliant.org>

The Rossmo formula is the summation is of N past crimes that occurred in the  $(x_n, y_n)$  coordinates in the past.



source:<http://brilliant.org>



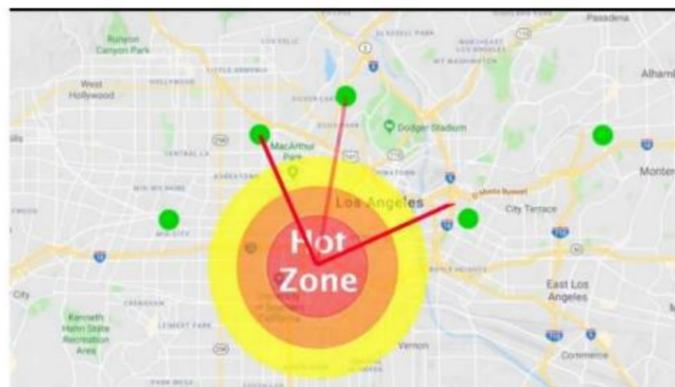
source:<http://brilliant.org>

### Necessity of the two terms in the formula:

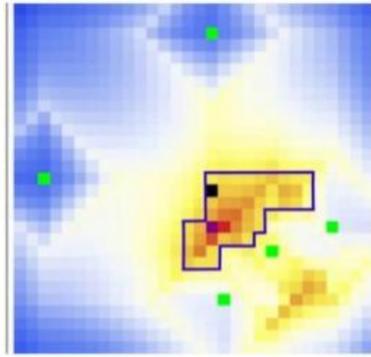
By criminology concept we already assumed that the criminal will not commit crime too close nor too far to the place where he is residing. So, a balance should be maintained in the formula.

- $|X_i - x_n|$  : distance of crime scene from square considered in accordance with x axis.
- $|Y_i - y_n|$  : distance of crime scene from square considered in accordance with y axis.

Let us consider the denominator of the first term in right hand side. This term is basically the distance between the assumed square and the crime sport. In this term as the distance increases the probability decreases. When we consider the second term, the distance is subtracted from the buffer zone. Therefore as denominator decrease probability of residence increase. This variation maintains a balance in accordance with the assumption made. The steps are repeated for every square in the grid. A heat map is developed with the help of the obtained data.



source:<http://brilliant.org>



source:<http://brilliant.org>

Here, a heat map is given which is developed using Rossmo Formula. This helped in finding out the residence of a 70's serial killer Richard Trenton Chase.

### 2.1.3 DRAWBACKS AND SUGGESTED MODIFICATION

Challenging human beings is indeed a big deal Here a theoretical concept is manipulated to do the same. Therefore it will have drawbacks. The aspects of human behaviour is complex and uncertain . Rossmo formula is used to track the residing place of the killer. We cannot be cent percent sure about the person to be present there. This increased the chances of modification.

Modification was proposed by Feroz shah Syed , Didong Li , Xun Zhang and Zhenhong Guo. This was about calculating **escape route of criminal** and the **Time and location of next crime**. This could help in strengthening of existing model.

## Chapter 3

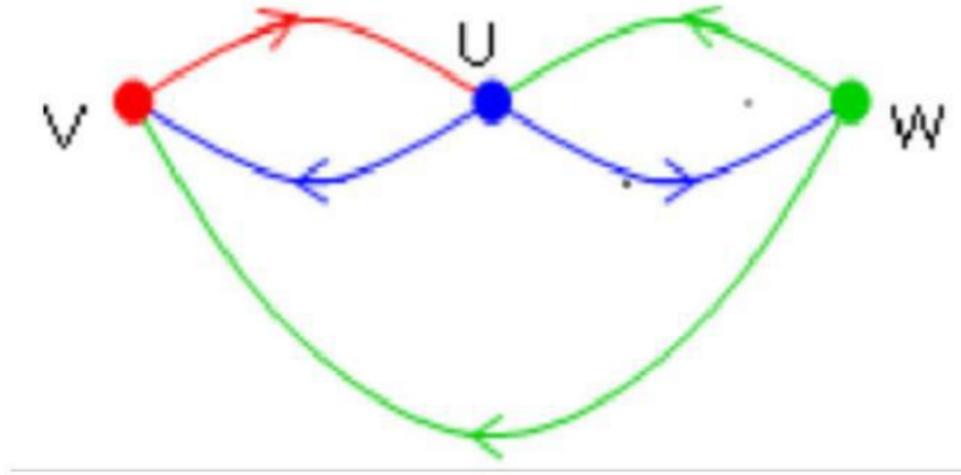
# CRIME PROBLEM USING GRAPH THEORY AND FORENSIC STATISTICS

---

### 3.1 GRAPH THEORY

Consider the scenario involving a crime where a jewellery shop has been subjected to robbery .Let us assume the suspects U , V and W as vertices . Assign the colours blue , red and green to the vertices. Suspect U says he is not the robber then lines to be drawn from U to V and W and assign colour blue . Suspect V says U is the robber then there is a line connecting from V to U and assign colour red . Suspect w says he is not the robber then lines to be drawn from W to U and V and assign colour green.

SUSPECTS	INVESTIGATION
U	I am not the robber
V	U is the robber
W	I am not the robber



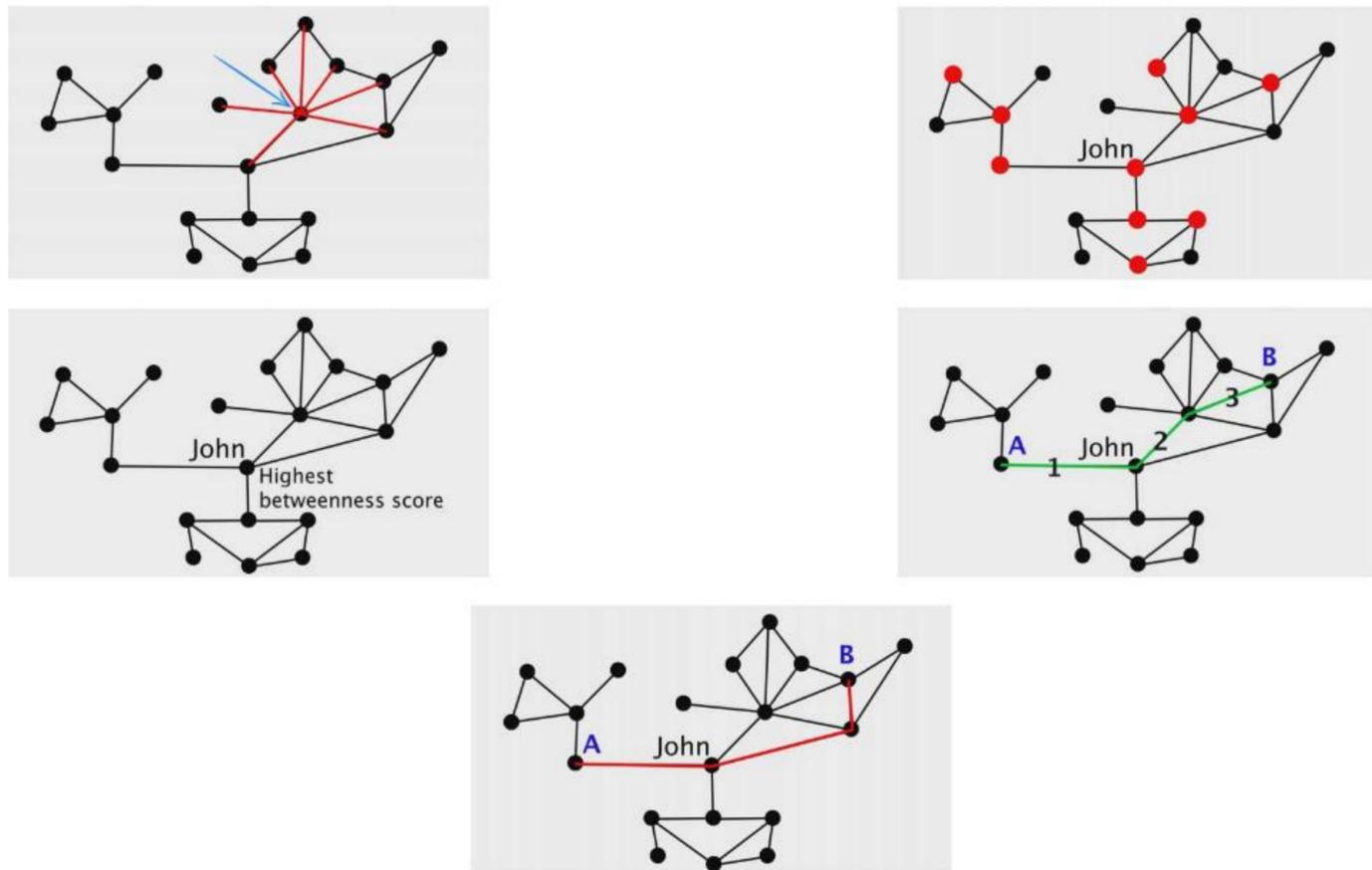
source:international journal of recent technology and engineering

This is the graph showing the connection between the attackers of the 9/11 attack through FBI investigation. What they did to prioritize people by law enforcement can be expressed in a simpler way using graphs.



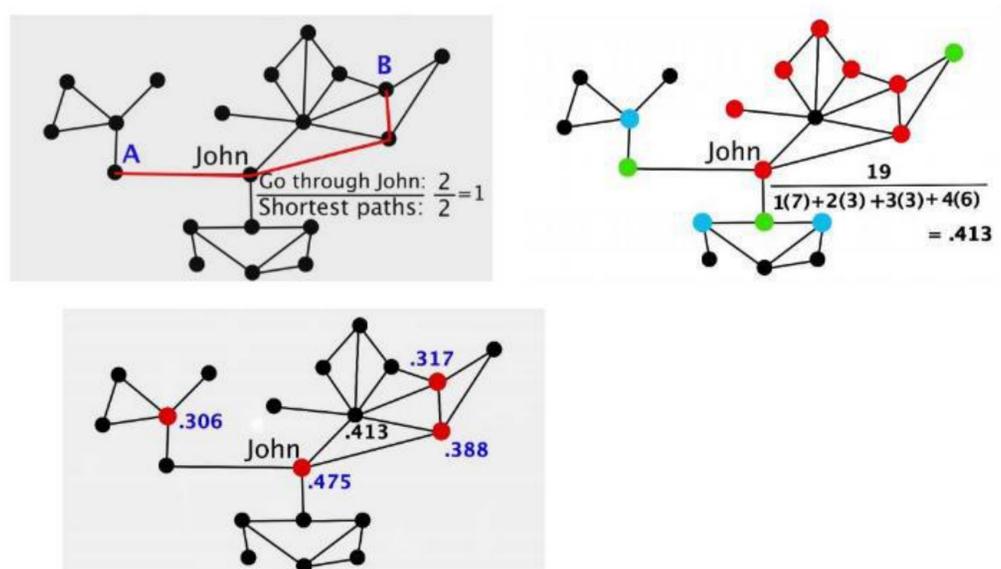
source:<http://brilliant.org>

Let's take the example of a school. All the students are connected in one way or the other whether it be social media, classes they share, close friendship and so on. For any two people let's say they are connected if they are close friends or in close contact with each other.



source:<http://brilliant.org>

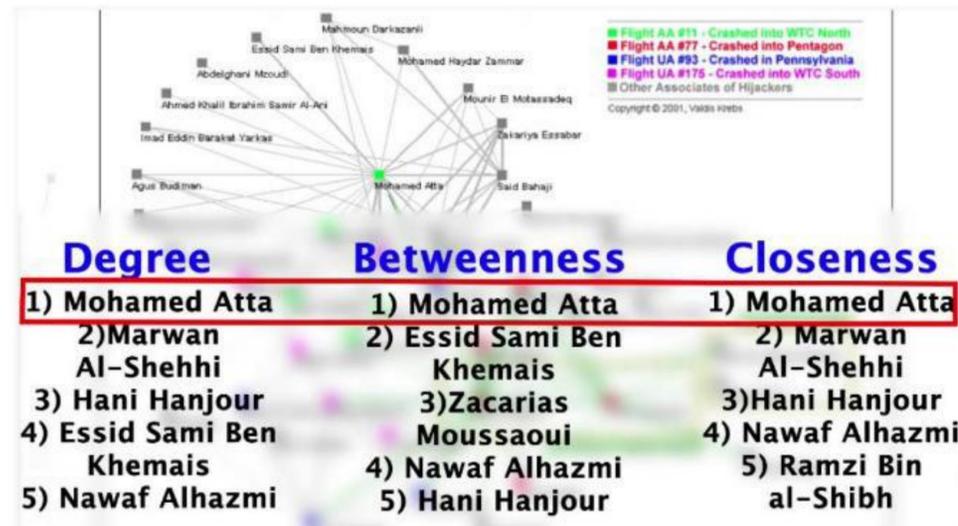
The dot to which the arrow is pointed in the first figure has most connections. But through further analysis we can see that John is the prime connection between the groups, without John no information can be passed. Here comes the importance of highest betweenness score and closeness centrality. Highest betweenness score is the score received to each nodes based on the number of shortest paths that passes through the node. consider the points A and B, the information from A reaches B through John by a number of ways among which two number of ways which is the shortest path with 3 edges. So the betweenness score of John is  $\frac{\text{Gothroughjohn}}{\text{Shortestpath}} = \frac{2}{2} = 1$ .



source:<http://brilliant.org>

In a connected graph closeness centrality of a node is a measure of centrality in a network calculated as a reciprocal of the sum of the length of the shortest path

between the node and all other nodes in the graph. Thus more central a node is the closer it is to all other nodes. When considering the given graphs the closeness centrality of John is greater compared to others. So from John we trace back to find out from where a news started spreading. Similarly, During 9/11 attack Mohammed Atta was found to have the highest degree, betweenness score and closeness centrality. So, he was the prime link in the attack.



source: <http://brilliant.org>

### 3.2 FORENSIC STATISTICS

1) A wounded criminal walks away from the crime scene with blood dripping from a wound on his arms forming elongated, elliptical bloodstain of width  $7.2mm$  and length  $9mm$ . What is the walking speed of the criminal?

SOLUTION:

Angle of Impact  $\theta$  is given by,

$$\sin \theta = \frac{\text{width}}{\text{length}}$$

$$\theta = \sin^{-1} \left( \frac{\text{width}}{\text{length}} \right)$$

$$\theta = \sin^{-1} \left( \frac{7.2}{9} \right)$$

$$\theta = \sin^{-1}(0.8) = 53^\circ$$

As the wound is on his arms we assume the drop height to be  $1m$ .

Then the vertical impact speed  $V_g$  is given by,

$$V_g = \sqrt{2gh}$$

given,  $g = 9.8m/s^2$ ,  $h = 1m$

$$V_g = \sqrt{2 \times 9.8 \times 1} = \sqrt{19.6} = 4.43m/s^2$$

when  $\theta = 53^\circ$

$\tan \theta = \tan 53^\circ = 1.33$

$$\tan \theta = \left( \frac{V_g}{V_h} \right)$$

The horizontal speed can be calculated by,

$$V_h = \left( \frac{V_g}{\tan \theta} \right)$$

$$V_h = \left( \frac{V_g}{\tan 53^\circ} \right)$$

$$V_h = \frac{4.43}{1.33}$$

$$V_h = 3.33m/s^2$$

2) Magi's DNA profile matched with the DNA evidence discovered at the crime scene. Her identical twin Grace is the defendant. A DNA profile is of a type that is found in only 1 out of a billion people. Check the admissibility of the evidence and evaluate which hypothesis it is in support of.

SOLUTION:

$H_p$ : The DNA source is Grace (Prosecution Hypothesis)

$H_d$ : The DNA source is Magi (Defence Hypothesis)

E: The DNA evidence found matched with Magi's DNA profile.

The chances of a DNA match in siblings is one out of 10,0 whereas identical twins have the same DNA profile. Therefore Grace's DNA profile matches with Magi's DNA profile.

Hence,

$$P\left(\frac{E}{H_p}\right) = P\left(\frac{E}{H_d}\right)$$

Likelihood Ratio

$$LR = \frac{P\left(\frac{E}{H_p}\right)}{P\left(\frac{E}{H_d}\right)}$$

As both the probabilities are equal,

$$LR = 1$$

Since  $LR = 1$ ,

- The evidence is inadmissible concerning the hypotheses.
- It is neutral and supports neither of the hypotheses.

3) Jim is accused of a robbery. The probability that he is innocent is 59%. Timothy, Jim's friend who runs a shop near the crime scene is called on as the prime witness. Timothy's probability of saying the truth is 0.73. The probability of Timothy saying the truth when Jim is guilty is 0.42. What is the probability that Jim is actually guilty if Timothy lied?

SOLUTION:

Defining all the events involved in this problem,

G: Jim is guilty.

$G'$ : Jim is innocent.

T: Timothy says the truth.

$T'$ : Timothy lies.

$\frac{T}{G}$ : Timothy says the truth when Jim is guilty.

$\frac{T'}{G}$ : Timothy lies when Jim is guilty.

Here, we need to evaluate  $P\left(\frac{G}{T'}\right)$ ,

$P\left(\frac{G}{T'}\right)$  is the probability of Jim being guilty when Timothy gives a false statement.

Given,

$$P(G) = 1 - 0.59 = 0.41$$

$$P(G') = 0.59$$

$$P(T) = 0.73$$

$$P(T') = 1 - 0.73 = 0.27$$

$$P\left(\frac{T}{G}\right) = 0.42$$

$$\begin{aligned} P\left(\frac{T'}{G}\right) &= 1 - P\left(\frac{T}{G}\right) \\ &= 1 - 0.42 = 0.58 \end{aligned}$$

$$P\left(\frac{G}{T'}\right) = \frac{P\left(\frac{T'}{G}\right) \times P(G)}{P(T')}$$

Substituting the values, we get,

$$P\left(\frac{G}{T'}\right) = \frac{0.58 \times 0.41}{0.27}$$

$$P\left(\frac{G}{T'}\right) = \frac{0.23}{0.27}$$

$$P\left(\frac{G}{T'}\right) = 0.88$$

Hence, the probability that Jim is guilty given Timothy says the truth is 0.88.

## Chapter 4

# CONTROLLING CRIME BY ANALYSING DATA MANIPULATION

---

In this fast moving world data plays pivotal role in day to day human activities. Data is considered as a resource, crime containing data manipulation must be strictly monitored and stringent actions must be taken against those fraudulent activities. As data is not merely a group of things or numbers but an asset.

### 4.1 BENFORD'S LAW

#### 4.1.1 INTRODUCTION

In 1981, Canadian-American astronomer Simon Newcomb found out that the former pages of logarithmic table were much worn when compared to other pages. He published a result that the probability of a single number  $N$  being the leading digit of a number is  $\log(N+1) - \log(N)$  based on his observation.

This was noticed by Frank Benford in 1938, a physicist and he tested the data from 20 different domains including surface area of 335 rivers, 104 physical constants, 1800 molecular weights, 308 numbers contained in an issue of Reader's digest etc. Total 20229 observations were used. Hence this discovery was named after Benford.

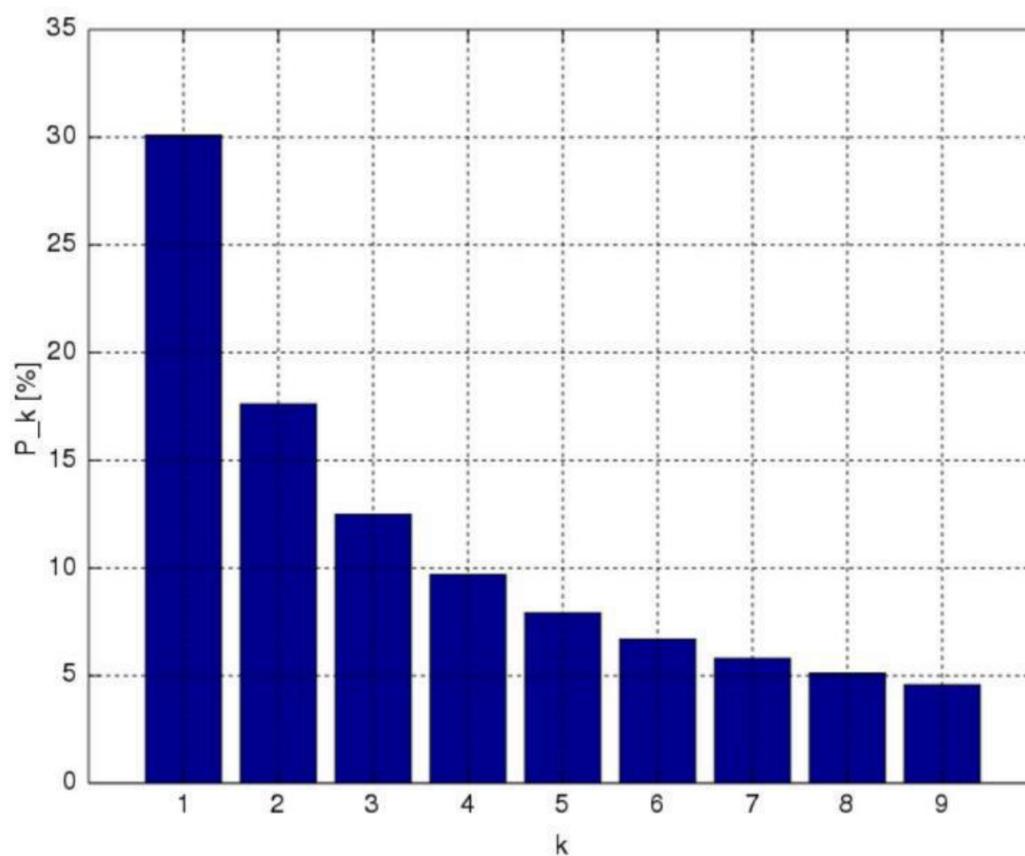
### 4.1.2 THEORY

Benford's law also known as Newcomb-Benford's law, law of Anomalous number or First digit law. Benford's law states that in a set of natural numbers the leading digit is more likely to be the number 1 the likelihood of leading digit be any other number decreases as it gets closer to number 9. Benford's law predicts the occurrence of digits in large sets of data. This law maintains that we can expect some digits to occur more often than other digits. For sets obeying Benford's law the number 1 appears as leading digit about 30% while the number 9 appears as leading digit less than 5%.

### 4.1.3 BENFORD'S LAW CONDITIONS

- The numbers in the data set should be the same objects -We cannot apply Benford's law when two different objects data is combined as one data set.
- There should be no built-in maximum or minimum to the numbers -We cannot apply Benford's law when the number is having a maximum or minimum limit.
- The numbers should not be assigned - In case of phone number, invoice number, car numbers Benford's law cannot be applied as they are assigned numbers.
- Does not apply for Uniform Distribution -In uniform distribution the occurrence of each number is about 11%

#### 4.1.4 DISTRIBUTION OF FIRST DIGIT ACCORDING TO BENFORD'S LAW



source:<http://wikipedia.org>

- Each bar represents digits from 1 to 9
- Height of the bar is the percentage of the numbers that start with the digit.

$d$	$P(d)$	Relative size of $P(d)$
1	30.1%	
2	17.6%	
3	12.5%	
4	9.7%	
5	7.9%	
6	6.7%	
7	5.8%	
8	5.1%	
9	4.6%	

source:<http://wikipedia.org>

A set of numbers is said to satisfy Benford's Law if the leading digit  $[d=1,2,\dots,9]$  occurs with probability

$$p(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}[(d+1)/d] = \log_{10}(1 + 1/d)$$

## 4.2 APPLICATIONS

### Live case study

From NSE 13th December data is taken. We will look into the total traded value, applied Benford’s law then we will see that the real data set is exactly falling into these percentage. And a comparison between what Benford is saying and actual data is done.

	A	J	N	O	P	Q	R	S	T	U
	SYMBOL	TOTTRDVAL	1st Digit		Digit	Actual	Actual %	Benford	Diff	
2	20MICRONS	40,47,865	4			1	645	29.45%	30.10%	-0.07%
3	21STCENMGM	5,71,621	5			2	384	17.53%	17.61%	-0.08%
4	3IINFOLD	38,11,94,542	3			3	279	12.74%	12.49%	0.25%
5	3MINDIA	2,98,51,519	2			4	207	9.45%	9.69%	-0.24%
6	3PLAND	4,32,727	4			5	192	8.77%	7.92%	1%
7	5PAISA	1,21,21,324	1			6	153	6.99%	6.69%	0.30%
8	610GS2031	50,000	5			7	133	6.07%	5.80%	0.27%
9	63MOONS	65,71,181	6			8	90	4.11%	5.12%	-1.01%
10	667GS2050	69,700	6			9	107	4.89%	4.58%	0.31%
11	676GS2061	1,91,910	1				2,190.0			
12	716GS2050	218	2							
13	762GS2039	10,158	1							
14	772GS2055	44,440	4							
15	795GS2032	110	1							
16	A2ZINFRA	28,48,616	2							
17	AAATECH	1,89,000	1							
18	AAKASH	3,18,06,428	3							
19	AAREYDRUGS	14,22,431	1							
20	AARON	16,72,841	1							
21	AARTIDRUGS	56,89,59,966	5							
22	AARTIIND	1,15,71,37,035	1							
23	AARTISURF	1,24,06,141	1							
24	AARVEEDEN	4,51,112	4							

### Inference

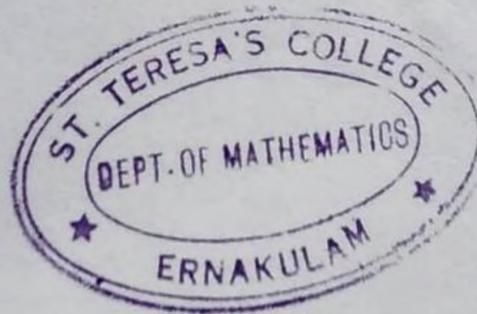
The difference is matching (not even 0.5%). When the difference is more than 1% or 2% .It doesn't mean a fraud ,it could alarm the auditor.

### Spreadsheet functions

- Extract the 1st digit out of the database [=LEFT(J2)] click enter
- Column for digits 1 to 9
- Actual frequency column (how many times the digit appearing in the beginning) [COUNTIF] Actual percentage (Actual/Total)  $\times$  100 [Q2/Q11(F4)]
- Benford's percentage [=log10(1/P2+1)]
- Difference (Actual - Benford)

#### 4.2.1 OTHER APPLICATIONS

- Accounting Fraud detection
- Use in Criminal Trails.
- Election Data
- Genome Data
- Scientific Fraud Detection economic Data Digit Analysis



## REFERENCES

---

- [1] A Forensic Formula for Solving Crimes
- [2] An Introduction to Statistical thinking for Forensic Practitioners.
- [3] Conditional Probability of Crimes
- [4] Mathematical Modeling Basics', TU Delft OCW |<https://ocw.tudelft.nl/courses/mathematical-modeling-basics/> [accessed 15 September 2021].
- [5] Mathematical modelling in criminology , feroz shah syed , didong li , xun zhang and zhenhong guo
- [6] Problems with Likelihood ratio method – Bayes Knowledge
- [7] Teaching Statistics using Forensic examples.
- [8] <https://en.m.wikipedia.org/wiki/Benford>
- [9] <https://youtu.be/gRAga1V12aM>
- [10] <https://en.wikipedia.org/wiki/Graphtheory>

Project Report

On

# TELECOM CHURN DATA ANALYSIS

*Submitted*

*in partial fulfilment of the requirements for the degree of*

MASTER OF SCIENCE

*in*

APPLIED STATISTICS AND DATA ANALYTICS

*by*

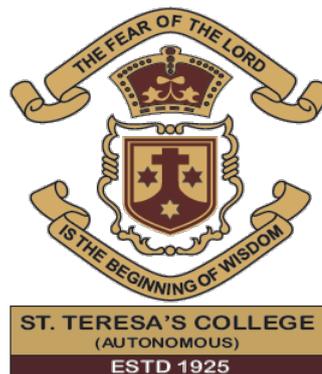
ATHULYA B PANICKER

(Register No. SM2OAS007)

(2020-2022)

*Under the Supervision of*

VRINDA MURALEEDHARAN



DEPARTMENT OF MATHEMATICS AND STATISTICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



CERTIFICATE

This is to certify that the dissertation entitled, **TELECOM CHURN DATA ANALYSIS** is a bonafide record of the work done by Ms. **ATHULYA B PANICKER** under my guidance as partial fulfillment of the award of the degree of **Master of Science in Applied Statistics and Data Analytics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 26/5/22  
Place: Ernakulam

*VRINDA*

Ms. **VRINDA MURALEEDHARAN**  
Assistant Professor,  
Department of Mathematics and Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.



*Shanty B P*

Ms. **SHANTY B P**  
Assistant Professor and HOD,  
Department of Statistics,  
St. Teresa's College (Autonomous),  
Ernakulam.

External Examiners

1: *Kumar*

*26/05/2022*

2: *m. H. B. P.*  
*26/05/22*

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **VRINDA MURALEEDHARAN**, Assistant Professor, Department of Mathematics and Statistics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 09/05/2022

*Athulya B. Panicker*

**ATHULYA B PANICKER**

**SM20AS007**

## ACKNOWLEDGEMENT

I express my deep sense of gratitude to Ms. Vrinda Muraleedharan, Assistant Professor of Department of Mathematics and Statistics, St Teresa's College(Autonomous), Ernakulam for her valuable guidance, constant suggestion and encouragement throughout the course of this project. I am also extremely grateful to the HOD Ms. Shanty B P for her continuous support and all the teachers of the department who provided us with the necessities for the completion of the project.

I also wish to thank my colleagues and all others who have given me the moral support to complete this project and made this venture a success. Above all, I thank God Almighty and my parents for giving me the blessings to take over this project.

Ernakulam.

Date: 09/05/2022

*Neeya*

ATHULYA B PANICKER

SM20AS007

# ABSTRACT

Customer churn is the rate at which customers stop doing business with an organisation or an entity. If we could puzzle out why a customer leaves, it would help the entity to find their various control initiatives. In this study, to find out the reasons of why customers are churning from a telecom company, various factors for customer churning are analyzed, supervised machine learning algorithms such as Logistic Regression, Random Forest, K-Nearest Neighbours, Decision Tree and Naive Bayes are used for modelling the dataset. The modelling results proved that Logistic Regression is the optimal model of choice for the dataset used as it has relatively the highest combination of precision, recall and F scores. Then K-fold cross validation is used to evaluate the chosen model and the results are visualised on Confusion Matrix.

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>ABSTRACT</i> . . . . .	v
<i>CONTENTS</i> . . . . .	vi
<b>1 INTRODUCTION AND LITERATURE REVIEW</b>	<b>1</b>
1.1 INTRODUCTION . . . . .	1
1.2 RELATED WORKS . . . . .	2
<b>2 DATA DESCRIPTION AND PREPROCESSING</b>	<b>6</b>
2.1 DATA DESCRIPTION . . . . .	6
2.2 DATA PREPROCESSING . . . . .	8
<b>3 DATA EVALUATION</b>	<b>10</b>
3.1 EXPLORATORY DATA ANALYSIS . . . . .	10
<b>4 MODEL SELECTION AND EVALUATION</b>	<b>23</b>
4.1 MODEL SELECTION . . . . .	23
4.2 MODEL EVALUATION . . . . .	25
<b>5 CONCLUSION</b>	<b>29</b>
<i>REFERENCES</i> . . . . .	30

# Chapter 1

## INTRODUCTION AND LITERATURE REVIEW

---

### 1.1 INTRODUCTION

Over the last few decades, the telecommunication industry (TCI) has seen extensive growth and development in terms of technology, level of competition, new products and services and so on. However, because of enormous competition, active environment, saturated markets, attractive offers etc TCI faces serious customer churn issues, which is considered to be an intimidating problem. In a driven market, where customers have countless choice of service providers, they can simply switch services and even service providers. Such customers are referred to as churned customers and their switching is termed as customer attrition.

To generate more revenues in an industry the three common strategies used are: (i) To increase the control period of customers. (ii) To bring in new customers. (iii) To promote the existing customers being the other two.

In point of fact, the most profitable strategy among the above three is the customer retention, since customer revenue severely hits the company's income and its marketing expenses.

When a customer has a long term dissatisfaction over the company's services, churn happens to be the unavoidable result. A customer's

complete discontinuation from a service or a service provider does not happen in a day, somewhat the dissatisfaction of the customer, grown over time and intensified by the lack of attention by the service provider, results in such a raging gesture by the customer. To avoid this, the service provider must work on constraints which are discerned by the customers in its services to retain the resentful customers. Therefore, it is highly favourable for a service provider to be able to identify a customer as a future churned customer. In this circumstance, non-churn customers are those who are unwilling to move from one service provider to another in contrast to churn customers.

Suppose a telecom company can predict that a customer is likely to churn, then it can possibly serve targeted benefaction to that customer to reduce his dissatisfaction, increase his commitment and thus potentially retain him or her. This clearly has a positive impact on the turnover of the telecom company. Also, the branding and fame of the company are also adversely affected by the customer attrition. Specifically, churn prediction is a very relevant task particularly in the telecom sector. Finally, telecom companies generally maintain a complete regular report of the customers to understand their standing and to predict their duration in continuing the services. Also, telecom companies nowadays mainly focus on retaining their long term customers rather than getting new ones because expense of getting new customers is very high. Thus, churn prediction become more important in the telecom sector. In this project, the customer churn prediction problem is considered as a data problem to analyze the different factors and the dataset is modelled using various supervised machine learning algorithms.

## 1.2 RELATED WORKS

Samah Wael Fujo, Suresh Subrahmanian and Moaiad Ahmad Khder published Customer Churn Prediction in Telecommunication Industry using Deep Learning on 1 Jan 2022. Deep-BP-ANN implemented in this study using two feature selection methods, Variance Thresholding and

Lasso Regression. To stop training at the right time and prevent overfitting, the model is strengthened by early stopping technique. Efficiency of minimizing overfitting between drop out and activity regularisation strategies for two real datasets are compared. Different evaluation approaches such as Holdout and 10-fold cross-validation are used here to evaluate the model's efficiency. In predicting customer churn, the findings outperform ML techniques: XG Boost, Logistic Regression, Naive Bayes and KNN. Moreover, the Deep-BP-ANN model's accuracy outperforms the existing deep learning techniques that use Holdout or 10-fold CV for the same datasets.

A study on data transformation based optimized customer churn prediction model for the telecommunication industry was done by Joydeb Kumar Sana, Mohammad Zoynul Abedin, M Sohel Rahman and M Saifur Rahman. In this, they have investigated various data transform methods to improve the prediction performance. This study is conducted in a customer churn prediction context in the telecommunication industry, where customer attrition is a common phenomenon. They have proposed a novel approach of combining data transformation methods with the machine learning models for the CCP problem. They have also presented comprehensive comparisons to affirm the effect of the transformation methods. The comparison results and the statistical test proved that most of the proposed data transformation based optimized models improve the performance of CCP significantly. Overall, an efficient and optimized CCP model for the telecommunication industry has been presented through this manuscript.

Customer churn analysis in telecom industry is a paper published by Kiran Dahiya and Surbhi Bhatia. In order to meet the need of surviving in the competitive environment, the retention of existing customers has become a huge challenge. In this paper, they have proposed a framework for the churn prediction model and implements it using the WEKA data mining software. Also they have compared the efficiency and performance of two supervised machine learning algorithms Decision Tree and Logistic Regression. And they have found out that

Decision Tree is far much higher than the Logistic Regression technique which clearly states that Decision Tree is an efficient technique in this case.

Vishal Mahajan, Dr Richa Misra and Dr Renuka Mahajan presented a paper on Review of Data Mining Techniques for Churn Prediction in Telecom in the year 2015. This paper provides a review of around 100 recent journal articles starting from year 2000 to present the various data mining techniques used in multiple customer based churn models. Then it summarizes the existing telecom literature by highlighting the sample size used, churn variables employed and the findings of different DM techniques. Finally, the paper lists the most popular techniques for churn prediction in telecom as Decision Trees, clustering and regression analysis.

A Study on Customer Churn Prediction in Telecommunication is done by Nabgha Hashmi, Naveed Anwer Bhatt and Dr Muddesar Iqbal. Data mining techniques enable the telecommunication companies to be equipped with effective methods for reducing churn rate. This paper reviews 61 journal articles to survey the pros and cons of renowned data mining techniques used to build predictive customer churn models in the field of telecommunication.

Irfan Ullah, Basit Raza, Ahmad Kamran Malik, Muhammad Imran, Saif Ul Islam and Sung Won Kim published the paper ‘A Churn Prediction Model using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector’ on May 6, 2019. The paper proposes a churn prediction model that uses classification, as well as, clustering techniques to identify the churn customers and provides the factors behind the churning of customers in the telecom sector. Here, feature selection is performed by using information gain and correlation attribute ranking filter. This paper also identified churn factors that are essential in determining the root causes of churn. And the proposed model is evaluated using metrics such as accuracy, precision, recall, F-measure and ROC curve. The results reveal that the proposed churn prediction model produced bet-

ter churn classification using Random Forest algorithm and customer profiling using K-Means clustering. The paper also provides the factors behind the customer attrition through the rules generated by using the attribute-selected classifier algorithm.

# Chapter 2

## DATA DESCRIPTION AND PREPROCESSING

---

### 2.1 DATA DESCRIPTION

The data has a total of 7043 observations with 19 explanatory variables and 1 response variable.

graphicx

	customer	gender	SeniorCiti	Partner	Depender	tenure	PhoneSer	MultipleLi	InternetSv	OnlineSec	OnlineBac	DevicePrc	TechSupp	Streaming	Streaming	Contract	Paperless	PaymentM	MonthlyC	TotalChar	Churn
0	7590-VHV	Female	0	Yes	No	1	No	No phone	DSL	No	Yes	No	No	No	No	Month-to	Yes	Electronic	29.85	29.85	No
1	5575-GNV	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mailed ch	56.95	1889.5	No
2	3668-QPY	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to	Yes	Mailed ch	53.85	108.15	Yes
3	7795-CFO	Male	0	No	No	45	No	No phone	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank tran	42.3	1840.75	No
4	9237-HQJ	Female	0	No	No	2	Yes	No	Fiber opti	No	No	No	No	No	No	Month-to	Yes	Electronic	70.7	151.65	Yes
5	9305-CDS	Female	0	No	No	8	Yes	Yes	Fiber opti	No	No	Yes	No	Yes	Yes	Month-to	Yes	Electronic	99.65	820.5	Yes
6	1452-KIO	Male	0	No	Yes	22	Yes	Yes	Fiber opti	No	Yes	No	No	Yes	No	Month-to	Yes	Credit car	89.1	1949.4	No
7	6713-OKO	Female	0	No	No	10	No	No phone	DSL	Yes	No	No	No	No	No	Month-to	No	Mailed ch	29.75	301.9	No
8	7892-POO	Female	0	Yes	No	28	Yes	Yes	Fiber opti	No	No	Yes	Yes	Yes	Yes	Month-to	Yes	Electronic	104.8	3046.05	Yes
9	6388-TAB	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No	No	No	One year	No	Bank tran	56.15	3487.95	No
10	9763-GRS	Male	0	Yes	Yes	13	Yes	No	DSL	Yes	No	No	No	No	No	Month-to	Yes	Mailed ch	49.55	587.45	No
11	7469-LKB	Male	0	No	No	16	Yes	No	No	No intern	Two year	No	Credit car	18.95	326.8	No					
12	8091-TTV	Male	0	Yes	No	58	Yes	Yes	Fiber opti	No	No	Yes	No	Yes	Yes	One year	No	Credit car	100.35	5681.1	No
13	0280-XJG	Male	0	No	No	49	Yes	Yes	Fiber opti	No	Yes	Yes	No	Yes	Yes	Month-to	Yes	Bank tran	103.7	5036.3	Yes
14	5129-JLP	Male	0	No	No	25	Yes	No	Fiber opti	Yes	No	Yes	Yes	Yes	Yes	Month-to	Yes	Electronic	105.5	2686.05	No
15	3655-SNQ	Female	0	Yes	Yes	69	Yes	Yes	Fiber opti	Yes	Yes	Yes	Yes	Yes	Yes	Two year	No	Credit car	113.25	7895.15	No
16	8191-XWS	Female	0	No	No	52	Yes	No	No	No intern	One year	No	Mailed ch	20.65	1022.95	No					
17	9959-WOF	Male	0	No	Yes	71	Yes	Yes	Fiber opti	Yes	No	Yes	No	Yes	Yes	Two year	No	Bank tran	106.7	7382.25	No
18	4190-MFL	Female	0	Yes	Yes	10	Yes	No	DSL	No	No	Yes	Yes	No	No	Month-to	No	Credit car	55.2	528.35	Yes
19	4183-MYF	Female	0	No	No	21	Yes	No	Fiber opti	No	Yes	Yes	No	No	Yes	Month-to	Yes	Electronic	90.05	1862.9	No
20	8779-QRD	Male	1	No	No	1	No	No phone	DSL	No	No	Yes	No	No	Yes	Month-to	Yes	Electronic	39.65	39.65	Yes
21	1680-VDC	Male	0	Yes	No	12	Yes	No	No	No intern	One year	No	Bank tran	19.8	202.25	No					
22	1066-JKSG	Male	0	No	No	1	Yes	No	No	No intern	Month-to	No	Mailed ch	20.15	20.15	Yes					
23	3638-WEA	Female	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No	Yes	No	No	Two year	Yes	Credit car	59.9	3505.1	No
24	6322-HRP	Male	0	Yes	Yes	49	Yes	No	DSL	Yes	No	Yes	No	No	No	Month-to	No	Credit car	59.6	2970.3	No
25	6865-JZNX	Female	0	No	No	30	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to	Yes	Bank tran	55.3	1530.6	No

The following are the explanatory variables alongside the response variable or dependent variable ‘Churn’.

- **GENDER:** Gender of a customer.
- **SENIOR CITIZEN:** Whether a particular customer is above 60 years of age.
- **PARTNER:** Denotes whether a particular customer has a partner.
- **DEPENDENTS:** Denotes whether a customer has dependent members living with him/her.
- **TENURE:** Total number of months a particular person has been a customer of the company.
- **PHONE SERVICE:** Whether a particular customer has/had opted for phone service.
- **MULTIPLE LINES:** Denotes whether the connection has got multiple lines (i.e, multiple users using one account).
- **INTERNET SERVICE:** Denotes whether a customer has/had opted for internet service.
- **ONLINE SECURITY:** Denotes whether a customer has/had opted online security feature.
- **ONLINE BACKUP:** Denotes whether a customer has/had opted for online backup feature.
- **DEVICE PROTECTION:** Denotes whether a customer has/had opted for Device Protection Program (such as insurance for the device from the company).
- **TECH SUPPORT:** Denotes whether a customer has/had opted for technical support of the company.
- **STREAMING TV:** Denotes whether a customer has/had opted for TV streaming service from the company.
- **STREAMING MOVIES:** Denotes whether a customer has/had opted for Movie Streaming Service from the company.
- **CONTRACT:** Whether a customer has/had signed contract for a duration of time.
- **PAPERLESS BILLING:** Whether a particular customer has/had chosen to pay bills online.

- **PAYMENT METHOD:** The way via which the payment is made (such as credit card, bank transfer).
- **MONTHLY CHARGES:** Amount charged to a particular customer the month before the data were collected.
- **TOTAL CHARGES:** Total amount charged during the whole period of a particular customer.

## 2.2 DATA PREPROCESSING

One of the main stages of information discovery activities is data pre-processing, which plays a significant role. It requires a variety of steps, such as transformation of data and reduction of data. If raw data is converted into low- quality data, learning algorithm's efficiency and accuracy will be compromised. So the collected data can be correctly analyzed by performing proper data pre-processing steps and choosing suitable learning algorithms. Telecom datasets have several issues to address, such as missing values, non-numeric features, inconsistent scales of features etc. Therefore, before implementing a learning model, it is important to pre-process the data. In this project, the collected dataset is pre-processed using almost the same strategies. The pre-processing phase includes handling the missing data, identifying unique values if any, encoding the categorical variables into two methods: 'label encoding' and 'one-hot encoding'.

- **MISSING DATA:** Missing data influences statistical analysis through knowledge loss and data pattern irregularities. But the data here has no missing values and hence this step is not required for the chosen dataset.

- **IDENTIFY UNIQUE VALUES:** The two categorical variables in the dataset are Contract and Payment Methods. In each of these variables, customers are either on a month-to-month rolling contract or on a fixed contract for one or two years. Also, they are paying bills via credit card, bank transfer or electronic checks.

- **CHECK RESPONSE VARIABLE DISTRIBUTION:** This step is done to see if the dataset upholds any imbalance issues. Here, the

chosen dataset is imbalanced with a high proportion of active customers compared to their churned counterparts.

- **LABEL ENCODE BINARY DATA:** Label encoding is a basic technique used to map categories as continuous integers in nominal attributes. ML algorithms can typically only have numerical values as their explanatory variables. Hence, label encoding is quite pivotal as they encode categorical variables with appropriate numerical values. Here, all the categorical variables having only two unique values are label encoded.

- **ONE-HOT ENCODING (OHE OR DUMMY):** It is a commonly used approach based on a binary encoding that uses a 1 to denote a group's presence and a 0 to denote the absence. This technique does not harm the raw information. Here, any categorical variables that have more than two unique values are dealt with one-hot encoding in the subsequent section.

## Chapter 3

# DATA EVALUATION

---

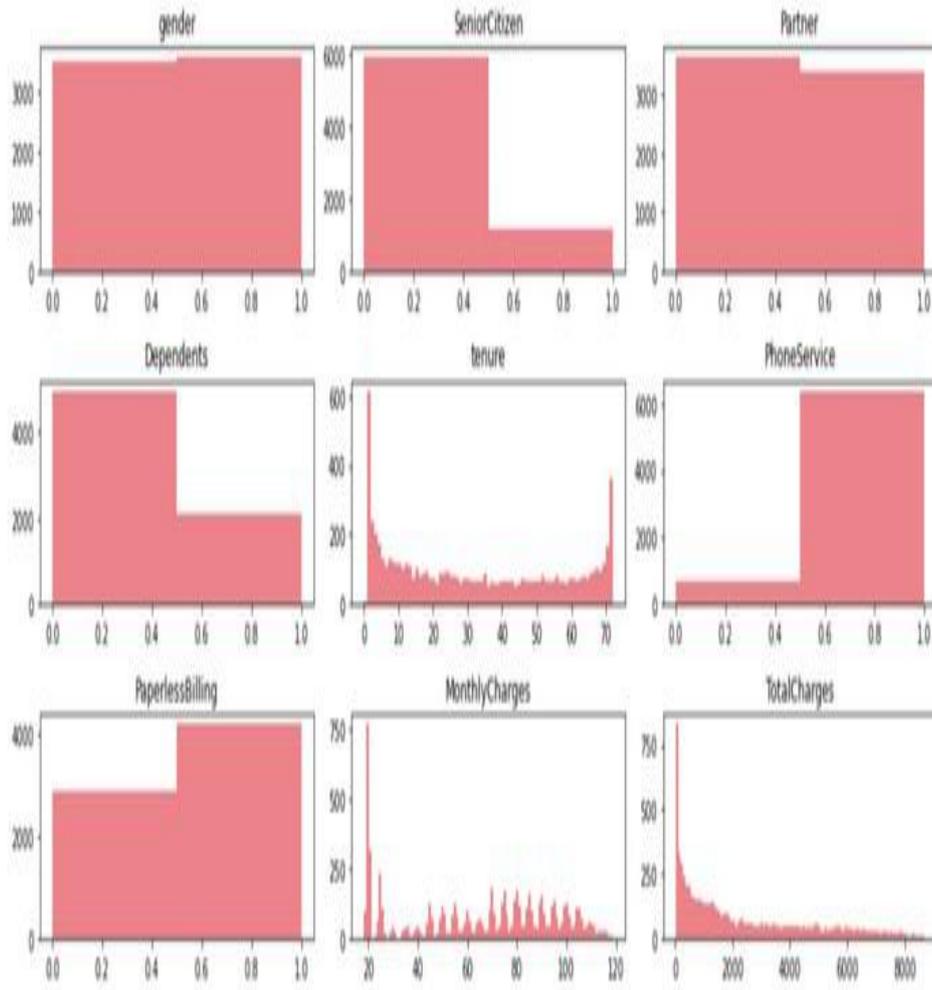
### 3.1 EXPLORATORY DATA ANALYSIS

One of the most important challenges in developing a mobile telecom service industry is retaining its customers. Exploratory Data Analysis (EDA) enables a service provider to track the product service that influence the customers to churn and suggest the best strategy to attract and retain the customers. To get a better understanding of the chosen data, EDA has applied here.

#### HISTOGRAM OF NUMERIC COLUMNS

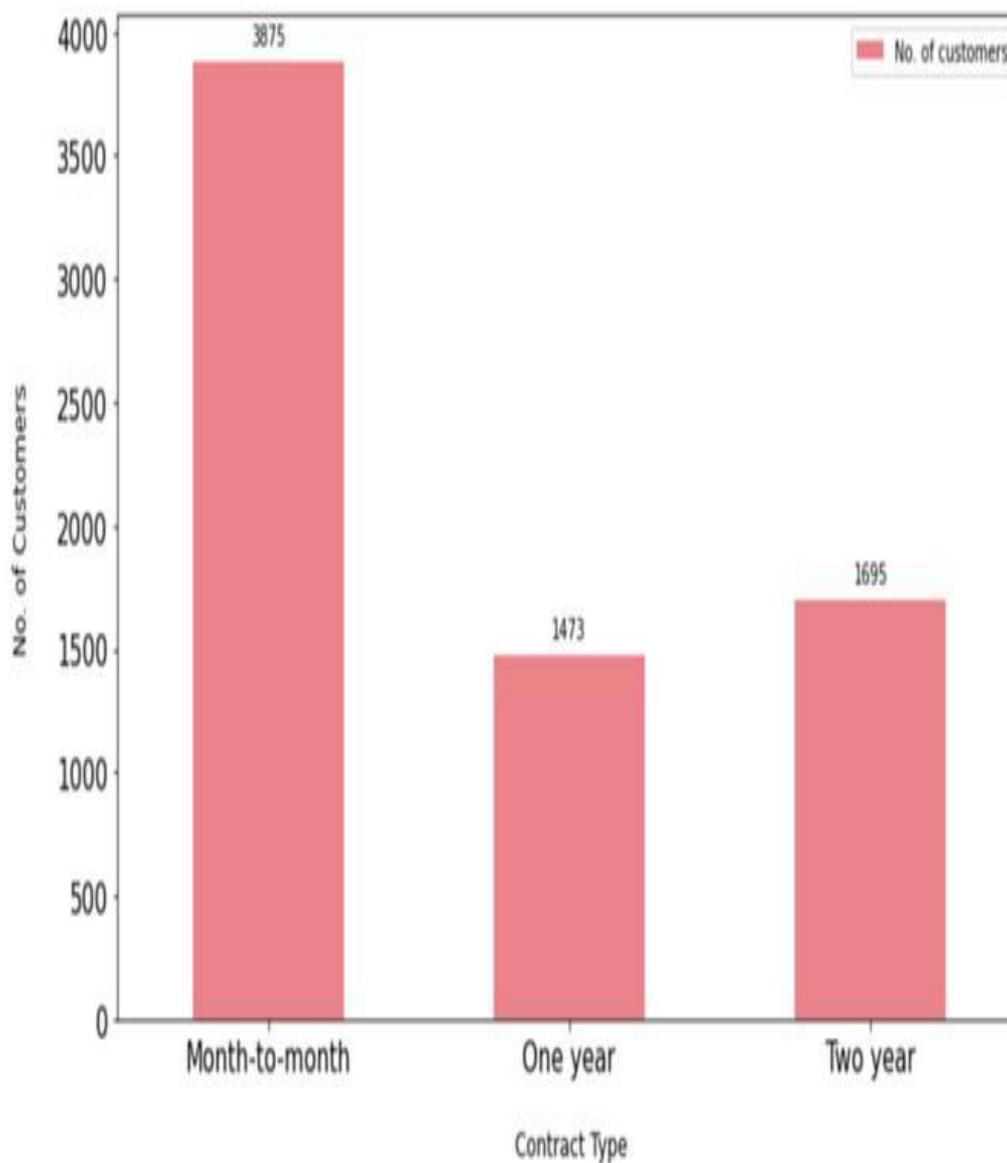
A few insights can be made based on the histograms for numerical variables:

- In gender distribution, the dataset features a relatively equal proportion of male and female customers.
- Most of the customers in the dataset are younger people.

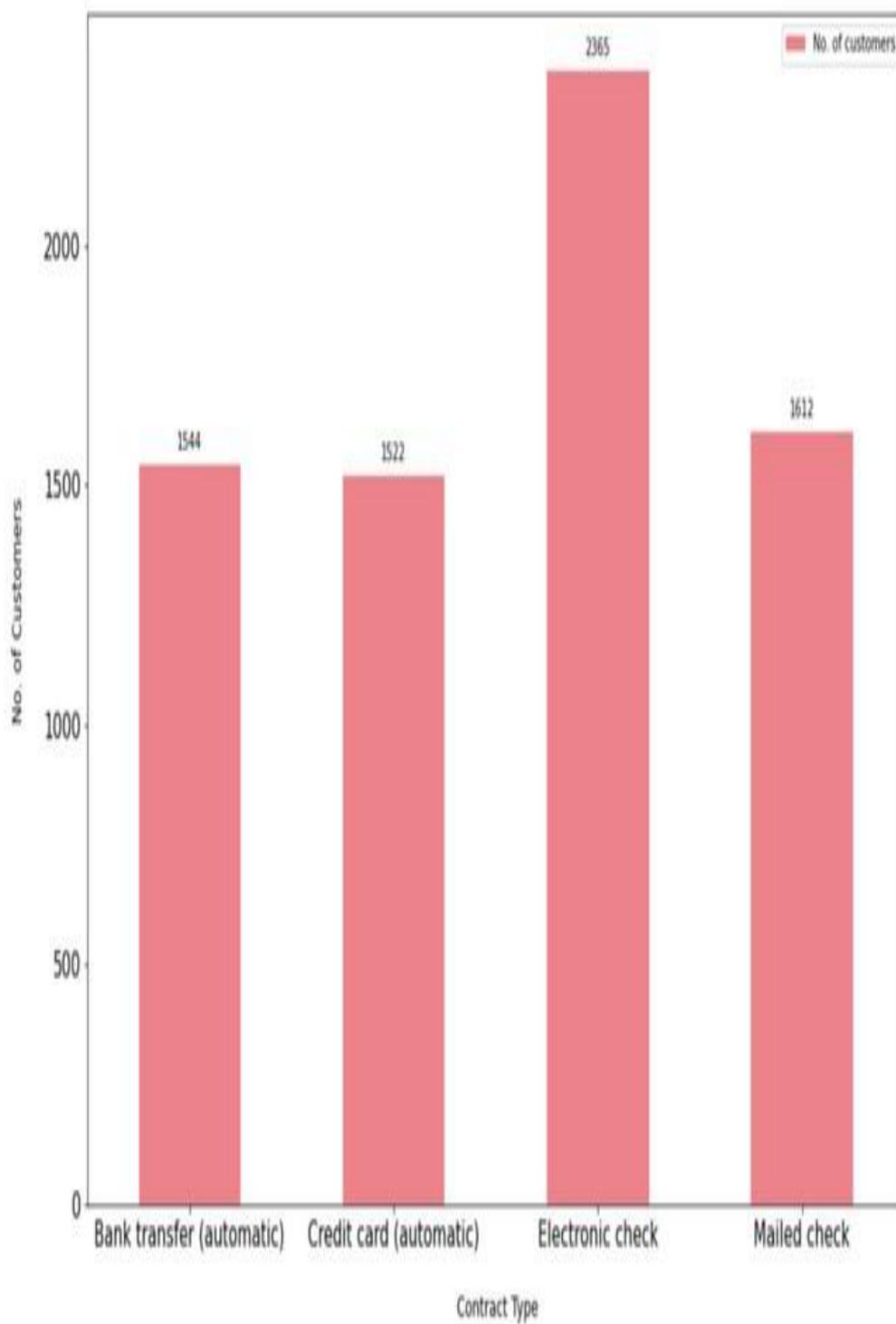


- Not many customers seem to have dependents but almost half of the customers have a partner.
- Most of the customers seem to have phone service and 3/4th of them have opted for paperless billing.

## ANALYZE THE DISTRIBUTION OF CATEGORICAL VARIABLES

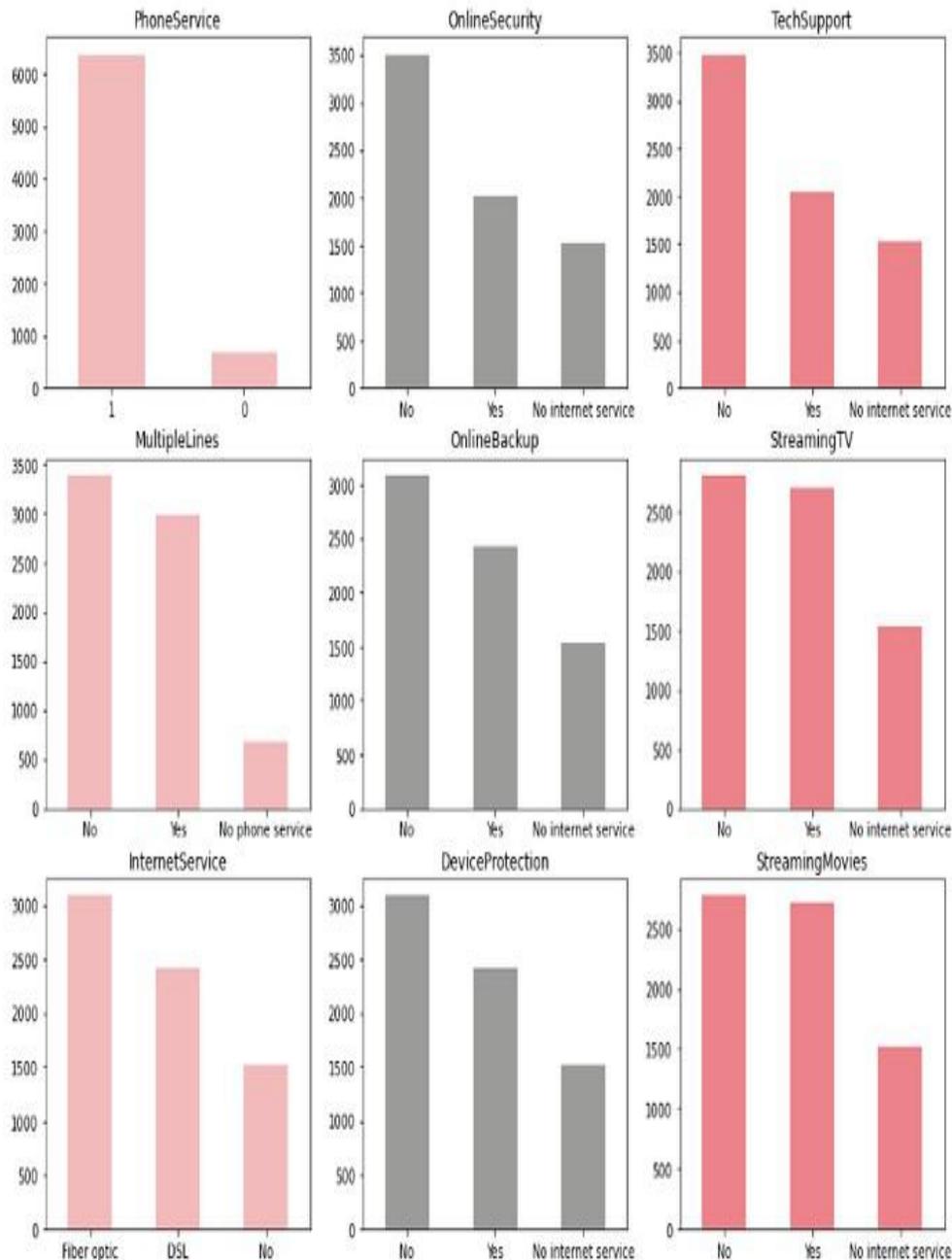


Above figure shows the Distribution of Contract Type. In the one year and two year contracts, there are a more or less equal proportion of customers. But most of the customers seem to have a prepaid connection with the telecom company.



Above figure shows the Distribution of Payment Method Type. Here, it indicates that most of the customers prefer to pay their bills electronically followed by bank transfer, credit card and mailed checks.

## DISTRIBUTION OF LABEL ENCODED CATEGORICAL VARIABLES



o Here, most of the customers have phone service out of which almost half of the customers have multiple lines.

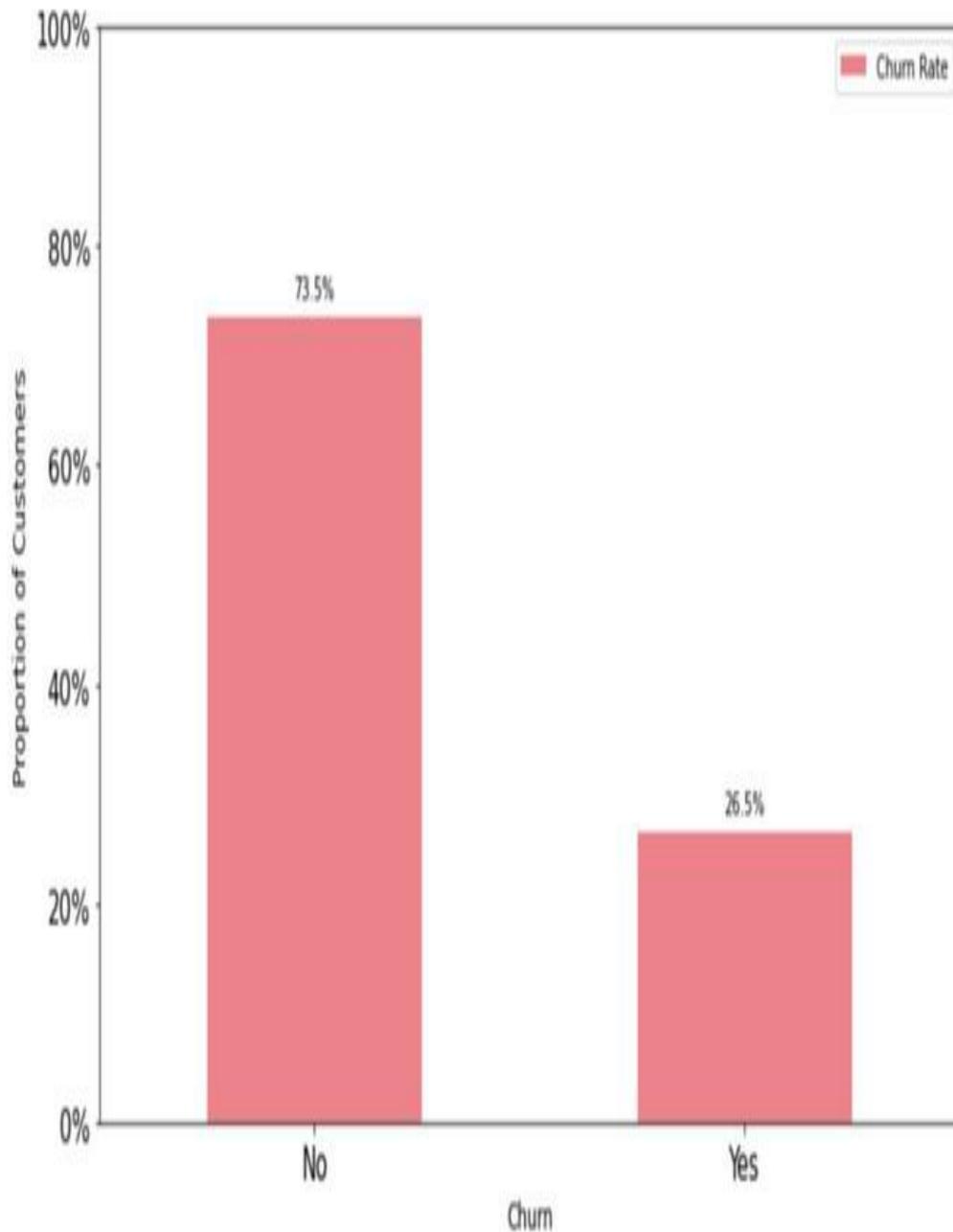
o Around 3/4th of the customers have opted for internet service via Fibre Optic and DSL connections.

o Almost half of the internet users have subscribed to Streaming TV and Movies.

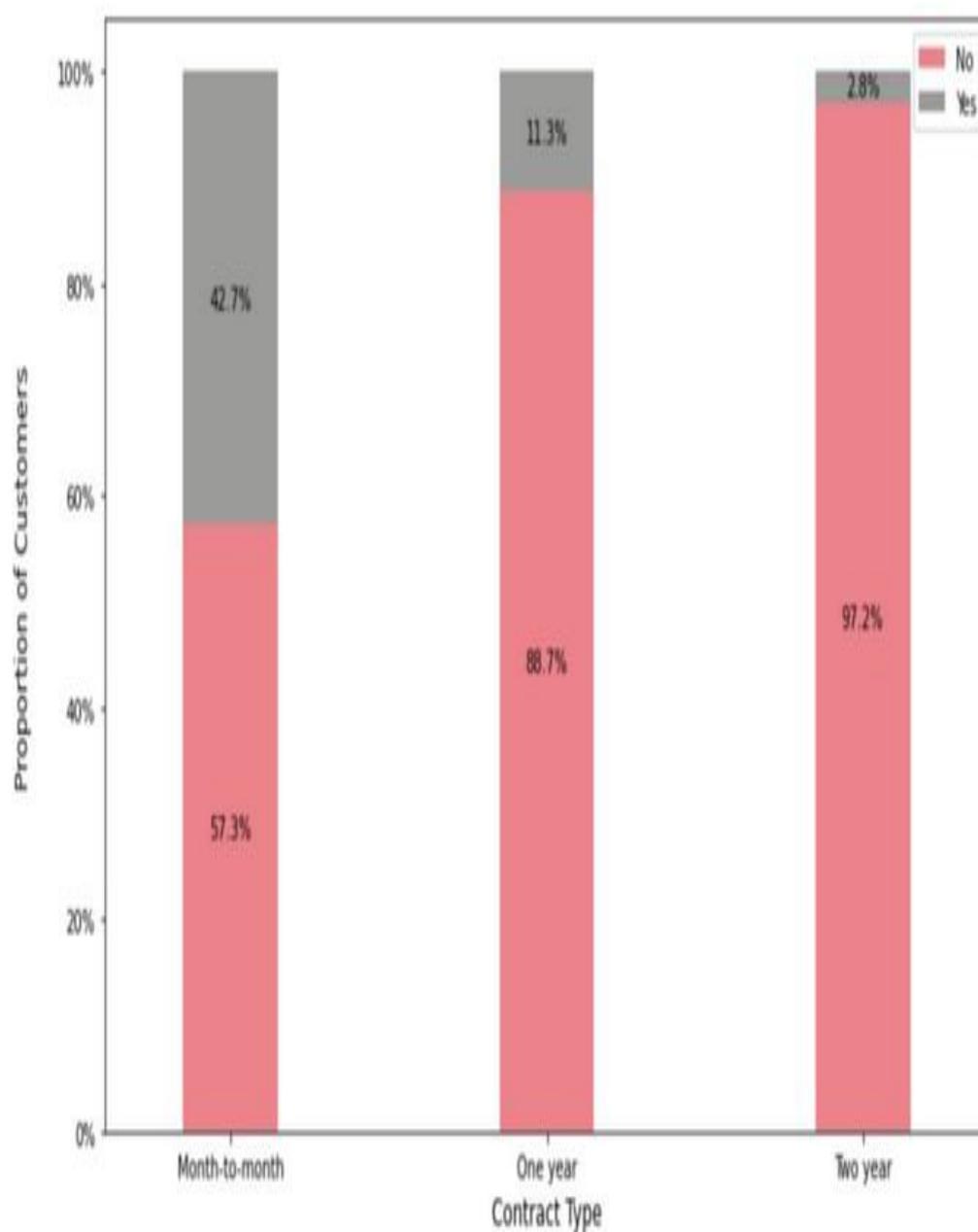
o Customers who have availed Online Backup, Device Protection,

Technical Support and Online Security features are a minority.

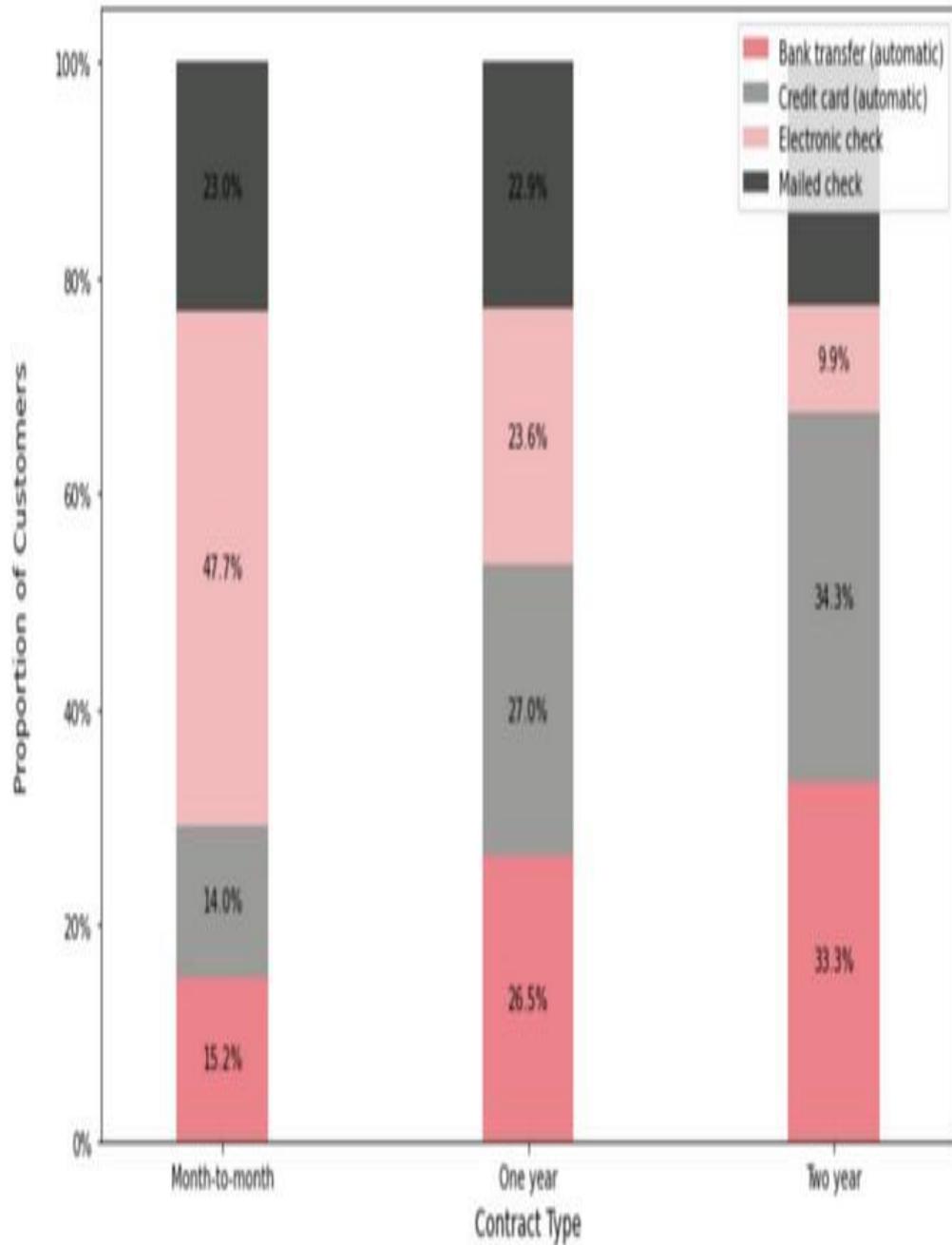
#### ANALYZE THE CHURN RATE BY CATEGORICAL VARIABLES



Above figure shows the Overall churn rate. It shows that around 74 percentage of the customers are active which in turn implies that the dataset is skewed.

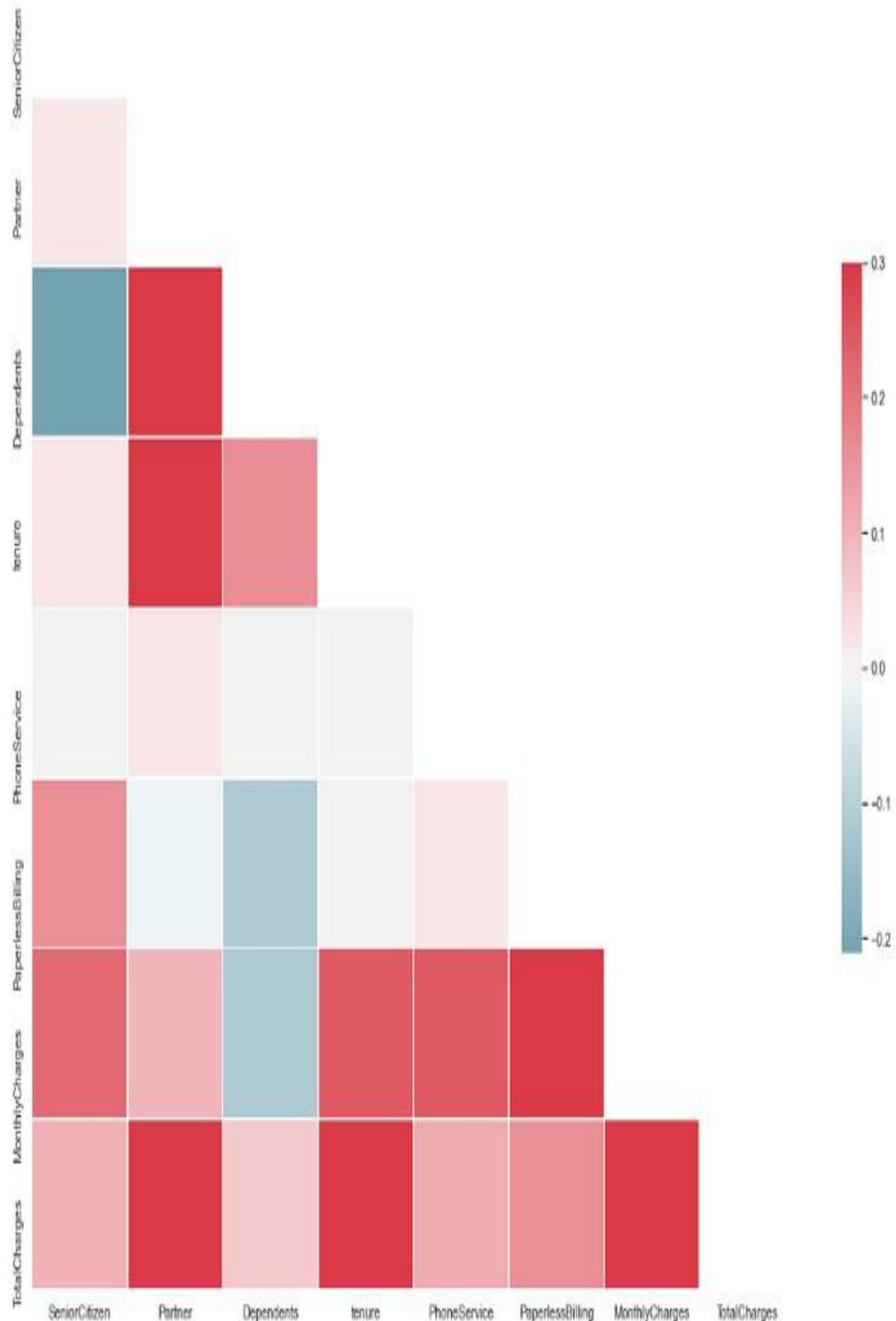


Above figure shows the Churn Rate by Contract Type. It shows a very high probability of customers with a Month-to-Month contract to churn when compared to customers on One or Two years contracts.



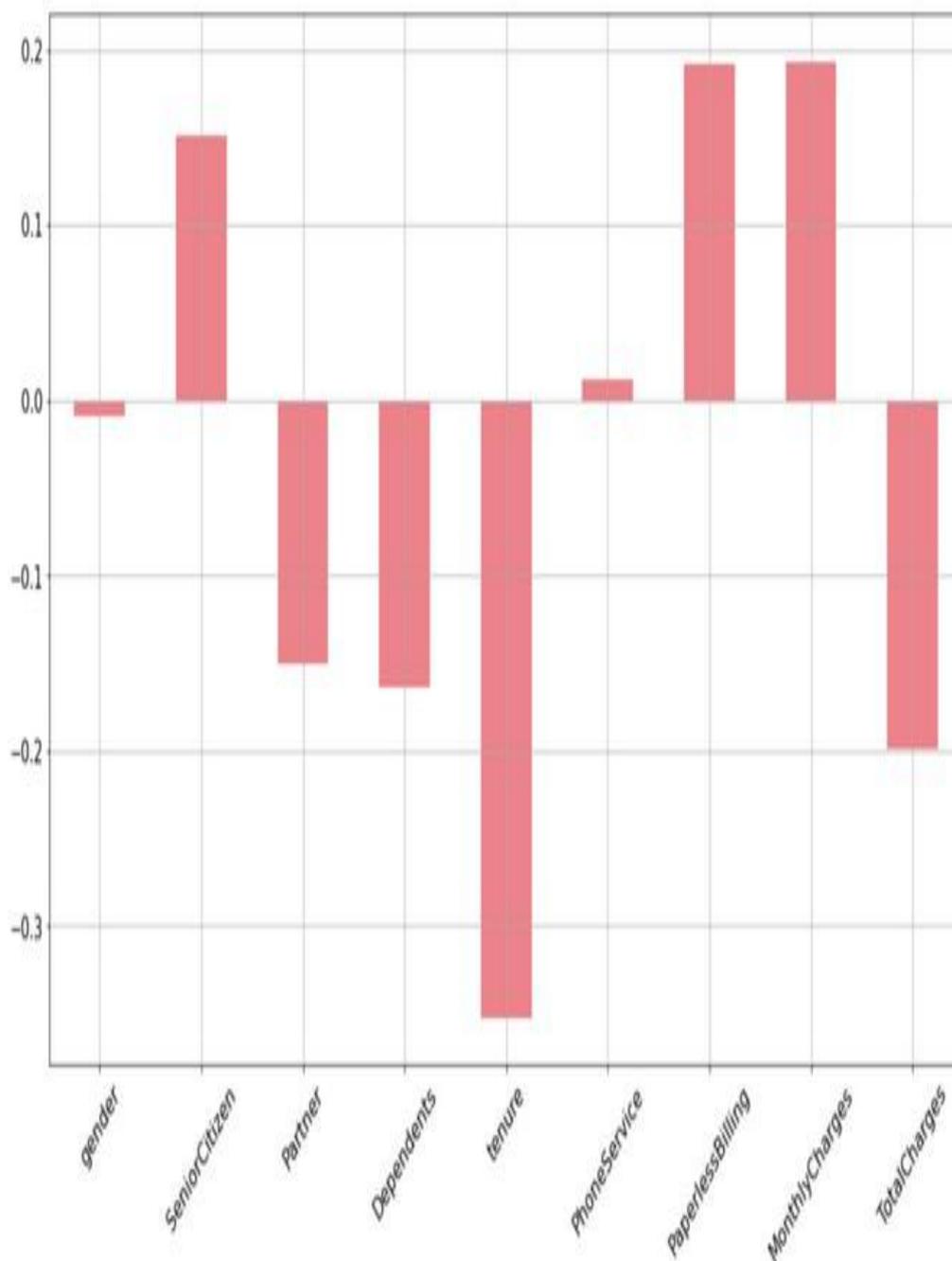
Above figure shows the Churn Rate by Payment Method Type. Among all the payment method segments, customers who pay via bank transfers seem to have the lowest churn rate.

## CORRELATION MATRIX



It is simply a table which displays the correlation. Each cell in the table shows correlation between the two variables. The matrix helps to discover the bi-variate relationship between the independent variables.

## POSITIVE AND NEGATIVE CORRELATIONS



From the above plot, it is clear that partner, dependents, tenure and total charges seem to be negatively related to churn. In contrast, it increases with age, monthly charges and paperless billing.

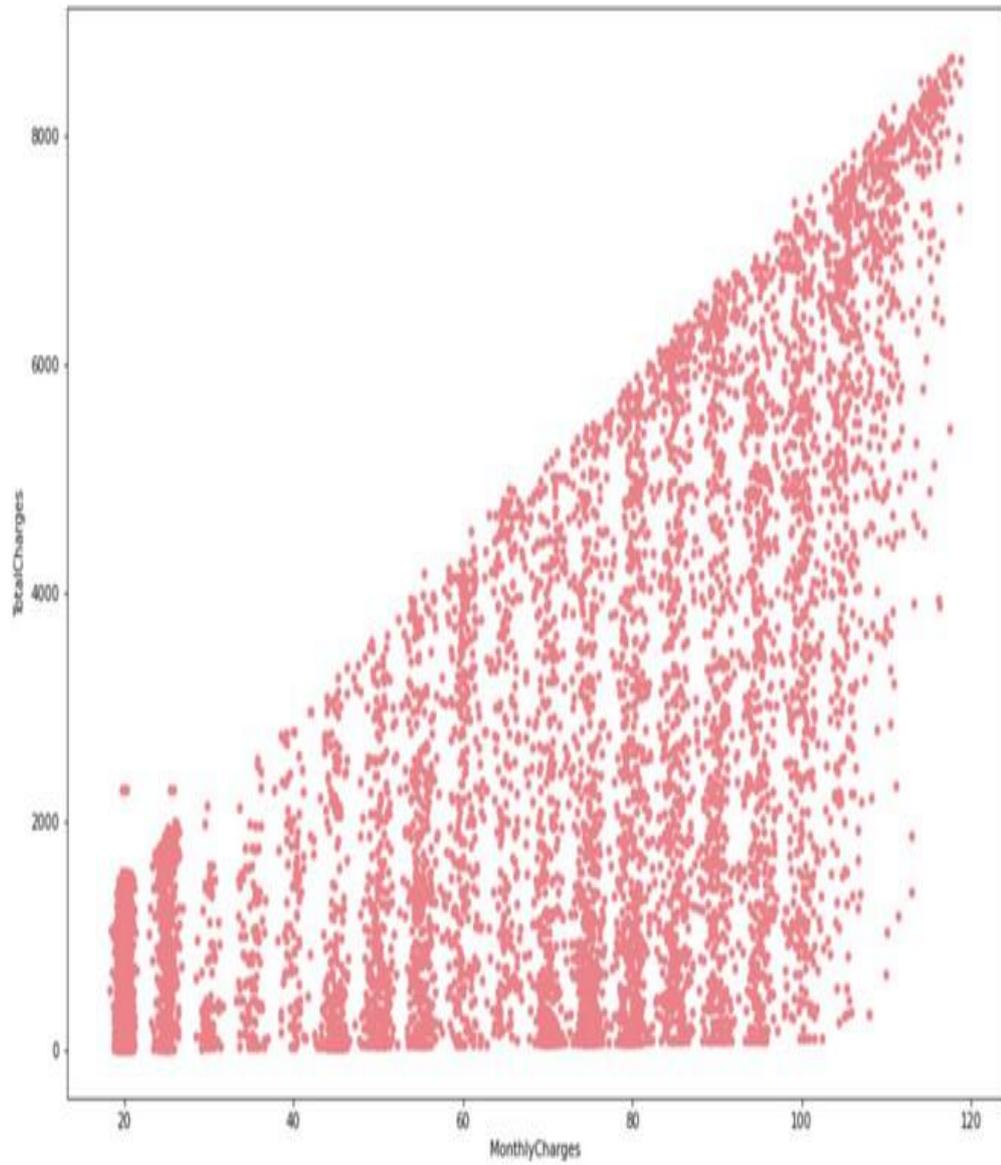
## MULTICOLLINEARITY USING VARIABLE INFLATION FACTORS(VIF)

Unlike correlation matrix, VIF determines the strength of the correlation of a variable with a group of other independent variables in a

dataset. VIF starts usually at 1 and anywhere exceeding 10 indicates high multicollinearity between the independent variables.

VARIABLES	VIF
Gender	1.921286
Senior Citizen	1.327766
Partner	2.815272
Dependents	1.921208
Tenure	10.549667
Phone Service	7.976386
Paperless Billing	2.814160
Monthly Charges	13.988649
Total Charges	12.570269

Here, tenure, monthly charges and total charges have VIF value greater than one. But when compared monthly charges and total charges have higher VIF value. The below graph shows the collinearity of monthly charges and total charges. In order to bring down the multicollinearity between the correlated features, one of the below correlated features have to be dropped.



So

here, total charges is dropped.

VARIABLES	VIF
Gender	1.879536
Senior Citizen	1.323089
Partner	2.814574
Dependents	1.908533
Tenure	3.287603
Phone Service	5.963240
Paperless Billing	2.745897
Monthly Charges	7.453993

Here, it is clear that after dropping the total charges, VIF values for all the independent variables have decreased to a considerable extent. All the values are between 1 and 10.

#### KEY FINDINGS FROM THE EDA

- The dataset does not have any missing values.
- Features such as monthly charges and age shows positive correlation while partner, dependents and tenure has negative correlation.
- Since majority of customers are active, the dataset is imbalanced.
- Customers with a month-to-month connection have a very high probability to churn that too if they have subscribed to pay via electronic checks.
- There is multicollinearity between monthly charges and total charges. In order to decrease the VIF values, total charges is dropped.
- Most of the customers in the dataset are younger people.

## Chapter 4

# MODEL SELECTION AND EVALUATION

---

The dataset is split randomly into two; one for training and the other one for testing in the ratio 4:1. From both the data frames Customer ID is separated. As part of feature scaling, the variables are normalized so that all the test and training variables are scaled within a range of 0 to 1. Now the classification algorithms are modelled over the training dataset and their accuracy, precision, recall, F1 and F2 scores are evaluated. The algorithms used here are Logistic Regression, Random Forest, K-Nearest Neighbors, Decision Tree and Naive Bayes classifier.

### 4.1 MODEL SELECTION

#### LOGISTIC REGRESSION

It is a type of statistical analysis which is often used for predictive analytics and modelling, and extends to application in machine learning. It is also known as logit model. It is used in statistical software to understand the relationship between the dependent variable and one or more independent variables by estimating probabilities using a logistic regression equation.

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
LOGISTIC REGRESSION	0.803407	0.652038	0.55615	0.600289	0.573003

**RANDOM FOREST ALGORITHM**

It is a supervised machine learning algorithm that is used widely in classification and regression problems. One of the most important features of random forest algorithm is that it can handle the dataset containing continuous variables as in the case of regression and categorical variables as in the case of classification. It performs better results for classification problems.

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
RANDOM FOREST	0.779276	0.617100	0.443850	0.516330	0.470255

**K-NEAREST NEIGHBORS ALGORITHM**

It is one of the simplest machine learning algorithms based on supervised learning technique. K-NN algorithm can be used for regression as well as for classification but mostly it is used for the classification problems. It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification it performs an action on the dataset.

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
K-NN	0.768630	0.570175	0.521390	0.544693	0.530468

**DECISION TREE ALGORITHM**

Decision tree is a supervised learning technique that can be used for both classification and regression problems. The name itself suggests that it uses a flowchart like a tree structure to show predictions that result from a series of feature based splits. It starts with a root node and ends with a decision made by leaves.

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
DECISION TREE	0.739532	0.508997	0.529412	0.519004	0.525199

**NAIVE BAYES CLASSIFIER**

It is a classification technique based on Baye's theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes

classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Naive Bayes model is easy to build and particularly useful for very large datasets. It is known to outperform even highly sophisticated classification methods.

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
NAIVE BAYES	0.703336	0.467359	0.842246	0.601145	0.725806

#### COMPARISON OF ABOVE RESULTS

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	F2 SCORE
LOGISTIC REGRESSION	0.803407	0.652038	0.55615	0.600289	0.573003
RANDOM FOREST	0.779276	0.617100	0.443850	0.516330	0.470255
K-NN	0.768630	0.570175	0.521390	0.544693	0.530468
DECISION TREE	0.739532	0.508997	0.529412	0.519004	0.525199
NAIVE BAYES	0.703336	0.467359	0.842246	0.601145	0.725806

The above comparison shows the accuracy, precision, recall and F scores of the modelled classification algorithms. Clearly, it can be concluded that Logistic Regression is an optimal model of choice for the given dataset as it has relatively the highest combination of accuracy, precision, recall and F scores.

## 4.2 MODEL EVALUATION

The optimal model of choice for the given dataset is Logistic Regression. The selected model (here it is Logistic Regression) is fitted on the training dataset and the results are evaluated. So here, the model evaluation is done through K-Fold Cross Validation technique that primarily helps to fix the variance.

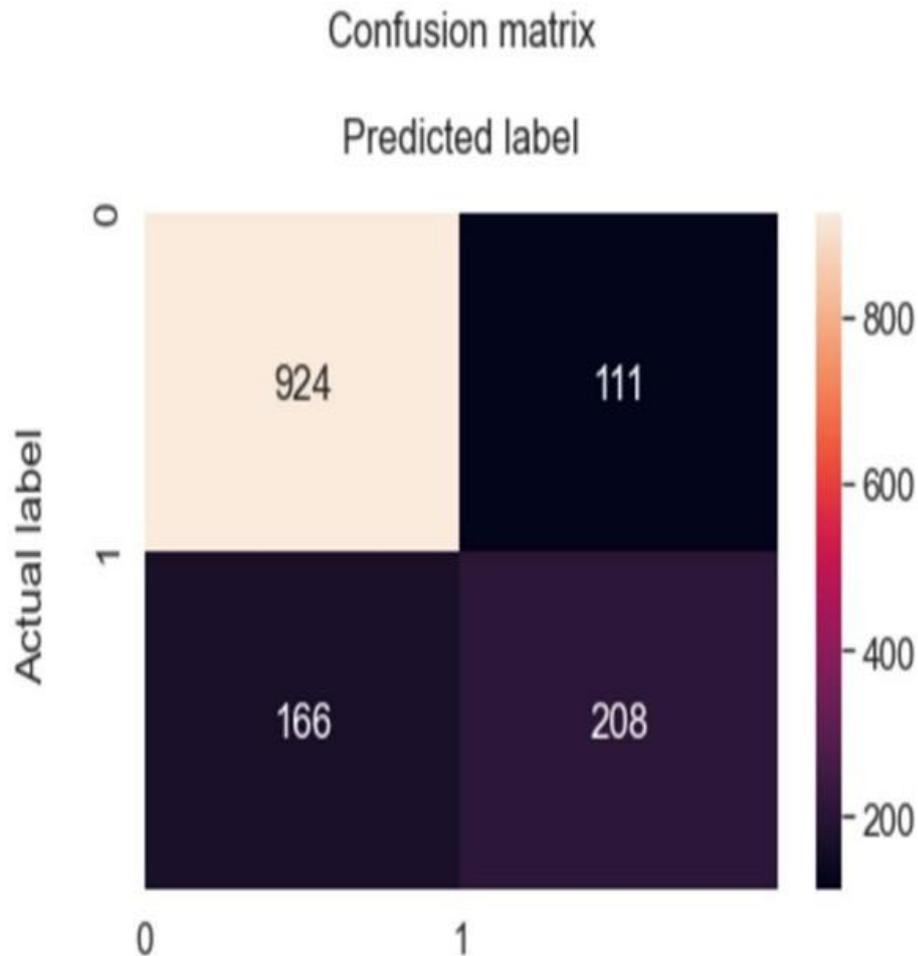
#### K-FOLD CROSS VALIDATION

Cross- validation is a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single

parameter called  $k$  that refers to the number of groups that a given data sample is to be split into. Here, there is a chance for variance problem because, for Logistic Regression the accuracy is very high (when compared to other classification algorithms) while running on training and test datasets but then the accuracy might look different when the model is run on another test set. In order to fix the variance problem  $k$ -fold cross validation technique split the training dataset into ten folds or subsets and train the model on nine folds (subsets) before testing it on the test fold. So, to train the model on all ten combinations of nine folds this gives a flexibility. After  $k$ -fold cross validation, it is found that Logistic Regression classifier accuracy is 0.80 (+/- 0.04). This implies that, it has an accuracy anywhere between 76 percentage to 84 percentage while running this model on any test set. Then, the results are visualised on confusion matrix.

#### CONFUSION MATRIX

A confusion matrix is a table that is often used to describe the performance of a classification model (or classifier) on a set of test data for which the true values are known. It is a performance measurement for machine learning classification. In predictive analytics, a table of confusion (sometimes also called a confusion matrix) is a table with two rows and two columns that reports the number of true positives, false negatives, false positives and true negatives. This allows more detailed analysis than simply observing the proportion of accuracy.



The confusion matrix indicates that there are (208+924) correct predictions and (166+111) incorrect predictions.

Accuracy rate = (Number of correct predictions / Total predictions) \* 100 = (1132 / 1409) \* 100 = 80.34

Error rate = (Number of wrong predictions / Total Predictions) \* 100 = (277 / 1409) \* 100 = 19.65

So, the model got an accuracy of 80 percentage which clearly shows the characteristics of a reasonably good model.

#### PREDICTING THE FEATURE IMPORTANCE

Logistic Regression is the optimal model for fitting the chosen dataset. Logistic Regression helps to determine the key features that have significance in predicting the target variable here it is, Churn. So, the logit model predicts a positive correlation with churn if any customer has subscribed to Month-to-Month contract, Electronic Checks, Optic

**Fibre Internet Service, Tech support and Absence of Payment Security.** On the other hand, the model predicts that the churn rate would decrease negatively with **Online Security, One-year contract and mailed checks** as the customer's payment medium.

## Chapter 5

# CONCLUSION

---

Predicting customer churn is one of the most important factors in business planning in telecommunication companies. So, in this domain, the researchers have been looking at the key factors of churn to retain customers. In this study, I have used five different machine learning algorithms which are Logistic Regression, Random Forest, K-Nearest Neighbor, Decision Tree and Naive Bayes classifier to model the dataset. After comparing the accuracy, precision, recall and F scores of the machine learning algorithms used, it is found that Logistic Regression is the optimal model of choice for the given dataset as it has relatively the highest combination of accuracy, precision, recall and F scores. After that, the logit model is evaluated using K-Fold Cross validation and is visualised with the help of Confusion Matrix. The obtained results show that the logit model has an accuracy of 80 percentage which clearly shows the characteristics of a good model. Also, the model helped to determine the key features that have significance in predicting the target variable Churn. .

## REFERENCES

---

- [1] Samah Wael Fujo, Suresh Subramanian, Moaiad Ahmad Khder. (2022). Customer Churn Prediction Telecommunication Industry Using Deep Learning.
- [2] Joydeb Kumar Sana, Mohammad Zoynul Abedin, M. Sohel Rahman, M. Saifur Rahman. (2022). Data Transformation based optimised customer churn prediction model for the telecommunication industry.
- [3] Susan Li. (2016). Predict Customer Churn with R. <https://towardsdatascience.com/predict-customer-churn-with-r-9e62357d47b4>
- [4] Lawrence Lo. (2012). Telecom Churn Analysis. [http://rstudio-pubs-static.s3.amazonaws.com/443094\\_b2c15d74e7e4b7b96d8fc95f3162b08.html](http://rstudio-pubs-static.s3.amazonaws.com/443094_b2c15d74e7e4b7b96d8fc95f3162b08.html)
- [5] Vishal Mahajan, Dr. Richa Misra, Dr. Renuka Mahajan. (2015). Review of Data Mining Techniques for Churn Prediction in Telecom.
- [6] Irfan Ullah, Basit Raza, Ahmad Kamran Malik, Muhammad Imran, Saif Ul Islam, Sung Won Kim. (2019). A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector.
- [7] Praveen Lalwani, Manas Kumar Mishra, Jasroop Singh Chadha, Pratyush Sethi. (2021). Customer churn prediction system: a machine learning approach.
- [8] Ozer Celik, Usame O. Osmanoglu. (2019). Comparing to Techniques used in Customer Churn Analysis.
- [9] Afaq Alam Khan, Sanjay Jamwal, M.M. Sepehri. (2010). Applying Data Mining to Customer Churn Prediction in an Internet Service Provider.

- [10] Teemu Mutanen. (2006). Customer churn analysis - a case study.
- [11] Nabgha Hashmi, Naveed Anwer Butt, Dr. Muddesar Iqbal. (2013). Customer Churn Prediction in Telecommunication A Decade Review and Classification.
- [12] Kiran Dahiya, Surbhi Bhatia. (2015). Customer Churn Analysis in Telecom Industry.

# ARCTIC PLANTS : HOW DO THEY SURVIVE?

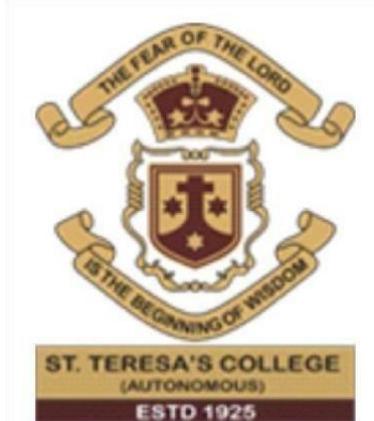
A REVIEW

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF BACHELOR OF  
SCIENCE IN BOTANY

BY

THAZNEEM HANEEF

REG.NO: AB19BOT015



DEPARTMENT OF BOTANY

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM – KERALA

2019-2022

# CERTIFICATE

This is to certify that the dissertation entitled 'ARCTIC PLANTS :HOW DO THEY SURVIVE?' submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Botany is an authentic record of the work carried out by Thazneem Haneef during her B.Sc. course from 2019-2022 under the guidance and supervision of Dr. Elsam Joseph , Department of Botany, St. Teresa's College (Autonomous), Ernakulam.



Dr. Liza Jacob

Head of the Department

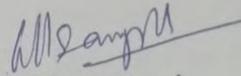
Department of Botany

St. Teresa's College(Autonomous)

Ernakulam

Place : Ernakulam

Date : 26-04-2022



Dr. Elsam Joseph

Associate Professor

Department of Botany

St. Teresa's College (Autonomous)

Ernakulam

**Dr. ELSAM JOSEPH**

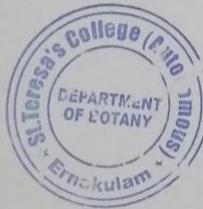
Associate Professor

Research Guide

Department of Botany

St. Teresa's College (Autonomous)

Ernakulam - 682 035



External Examiners

1. Dr. Anila N. Anish 9/5/22
2. Dr. Anisha S. Anish 9/5/22

## DECLARATION

I, Thazneem Haneef declare that the dissertation entitled 'ARCTIC PLANTS :HOW DO THEY SURVIVE?' is an authentic research work carried out by me under the supervision and guidance of Dr. Elsam Joseph, Associate Professor , Department of Botany, St. Teresa's College (Autonomous), Ernakulam, in partial fulfilment of requirements for the award of B.Sc Degree in Botany and no part of it has previously formed the basis for the award of any degree, diploma or associateship in any institution.



THAZNEEM HANEEF

Place : ERNAKULAM

Date : 26-04-2022

## ACKNOWLEDGEMENTS

This is a golden opportunity for me to convey my sincere regards to all those who helped to accomplish my dissertation work successfully.

First of all I thank the Almighty God for the numerous favours and mercies bestowed upon me, particularly during the project's duration.

It's my pleasure to thank Dr. Elsam Joseph, my project guide, for the guidance, whole hearted support and encouragement which she has rendered throughout the course of work. I am indebted for her continuous inspiration which gave me a golden opportunity to do this wonderful project smoothly.

I am grateful to Dr. Liza Jacob, Head of the Department of Botany, St. Teresa's College (Autonomous), Ernakulam, for her invaluable and unselfish aid and encouragement throughout the dissertation process. This job would not have been a success without her assistance and moral support.

I express my gratitude to all Non-teaching staff of Department of Botany, St. Teresa's College (Autonomous), Ernakulam for their kind assistance and support throughout this project.

Last but not least, I express my heartfelt thanks to all my friends and well-wishers who assisted me in completing this project.



THAZNEEM HANEEF

# CONTENTS

1.	INTRODUCTION.....	1
2.	REVIEW OF LITERATURE.....	5
3.	DATA OF DIFFERENT ARCTIC PLANTS.....	10
	<i>Saxifraga oppositifolia</i> .....	10
	<i>Papaver radicum</i> .....	16
	<i>Salix arctica</i> .....	21
	<i>Cassiope tetragona</i> .....	26
	<i>Picea pungens</i> .....	32
	<i>Viola x wittrockiana</i> .....	43
4.	CONCLUSION.....	50
5.	REFERENCES.....	54

# INTRODUCTION

The Arctic is a polar region located at the northernmost part of Earth. The Arctic consists of the Arctic Ocean, adjacent seas, and parts of Canada, Denmark (Greenland), Finland, Iceland, Norway, Russia, Sweden and the United States (Alaska). Land within the Arctic region has seasonally varying snow and ice cover, with predominantly treeless permafrost (permanently frozen underground ice) containing tundra. Arctic seas contain seasonal sea ice in many places. The Arctic region is a unique area among Earth's ecosystems. The cultures in the region and the Arctic indigenous peoples have adapted to its cold and extreme conditions. Life in the Arctic includes zooplankton and phytoplankton, fish and marine mammals, birds, land animals, plants and human societies. Arctic land is bordered by the subarctic (Seval H.C,2019).

Approximately 1,700 species of plants live on the Arctic tundra, including flowering plants, dwarf shrubs, herbs, grasses, mosses, and lichens. They have the ability to withstand extremely cold temperatures in the winter (winter hardiness), and grow and reproduce in summer conditions that are quite limiting. The tundra is characterized by permafrost, a layer of soil and partially decomposed organic matter that is frozen yearround. Only a thin layer of soil, called the active layer, thaws and refreezes each year. This makes shallow root systems a necessity and prevents larger plants such as trees from growing in the Arctic. The cold climate and short growing season also prevent tree growth. Trees need a certain amount of days above 50 degrees F, 10 degrees C, to complete their annual growth cycle (Billings W.D & Mooney H. A, 1968).

Arctic summers are short, cold and unpredictable, but plants live even in the most barren places and the region is greener than you might expect. Plant richness is lower in the Arctic than further south, but you will find an impressive diversity of species, fascinating adaptations and ecosystem variations, many of which are easy to observe when walking on the tundra. Plants do not exist to be beautiful or fascinating; however, they play a critical role in carbon cycling and energy balance of Arctic ecosystems. Some species are common and circumpolar, whilst specialists and rare species are found in small, defined areas.

Arctic plant life, or flora, includes vascular plants like bushes, herbs, grass and sedges, but also mosses, lichens and fungi. Trees can be found in a few places in the Low Arctic, including Russia and North America. Some plants have colorful flowers, while others appear like a green or dusty cover on the ground. A major part of plant biomass is below ground, like roots, bacteria, and fungi hyphae. More than 2200 species of vascular plants live in the Arctic. Diversity varies among regions, e.g. Svalbard has 170 vascular plant species whereas about 700 species live around the Bering Strait. Arctic plants are adapted to cold climate and short growing seasons. Most Arctic plants only grow a few cm tall. Trees and bushes can only survive in the southernmost parts. Despite slow growth rates, some single plants grow huge and very old. Genetic studies reveal ages of some sedge plant mats more than 3000 years old. Plants with similar ecological requirements grow together and make up the vegetation. Vegetation can be a continuous plant cover or only scattered plants (Matveyeva .N. & Chernov. Y,2000).

Diversity of Arctic vegetation ranges from barren land with only lichens on boulders, to lush and vigorous spots with high biomass and numerous species. Arctic vegetation types occur in mosaic and can change completely within only a few centimeters.

Vegetation shifts along ecological gradients like moisture (dry to wet), topography (exposed ridges to protected snow beds), nutrient availability (low to high) and soil texture (fine to coarse-grained). Wherever you are in the Arctic, vegetation is distributed and identified along such gradients. Heath vegetation is dry or moist and dwarf shrubs, ericaceous species and lichens are common. Mosses, grasses and sedges dominate wetlands. Nitrogen is an essential nutrient and a main limiting factor to plant production in the Arctic, due to low temperatures and slow decomposition. On and underneath Arctic bird cliffs, guano from birds adds plenty of nutrients, and bird cliff vegetation is highly productive and spectacular in their shocking green appearance. Similar green colors occur by cultural remains like slaughtering places or hunting cabins, indicating increased level of nutrients over the course of centuries ( Körner. C, 2021).

Plants have developed impressive and fascinating adaptations to survive in the Arctic. Growing close to the ground and close to a neighbor protects plants against low temperatures and mechanical damage from ice, snow and dust. Hairy leaves and stalks are striking in many species. These features protect plants from mechanical damage, and capture moisture and sunlight. Arctic plants can photosynthesize at extremely low temperatures, some even under a snow cover. Several Arctic plants prefabricate flower buds years in advance allowing them to bloom rapidly when snow melts. Most vascular plants are wind pollinated, including grasses, sedges and many herbs. Self-fertilization is frequent as many species are hermaphrodites. Many Arctic plants produce

impressive colored flowers despite limited nutrient resources, extreme climate and few pollinators. Flies are not efficient pollinators but important in many common white and yellow flowers. Yellow flowers reflect ultraviolet light, making them look purple to insects. Blue and purple flowers are more common in Arctic regions with more pollinators and specialist pollinators (Gellhorn J, 2002).

A changing Arctic climate affects plant species and vegetation. Warmer and wetter summers, and longer growing season allows some common native plants to grow taller and expand, while some rare high-arctic species struggle under the new conditions. Non-native species have higher chances of survival and expansion in warmer climate and have the potential to change native plant life. Preventing the introduction of nonnatives into the Arctic is ever more important in a changing climate.

# REVIEW OF LITERATURE

About 1,702 species of plants live on the Arctic tundra, including flowering plants, short shrubs, herbs, grasses, mosses, and lichens. These plants are adapted to short, cold growing seasons. They have the ability to withstand extremely cold temperatures in the winter (winter hardiness), and grow and reproduce in summer conditions that are quite limiting. As of 2005, arctic vegetation covered approximately  $5 \times 10^6$  km<sup>2</sup> ( $1.9 \times 10^6$  sq mi) of land. The area of arctic vegetation decreased by approximately 1.4 million square kilometres ( $0.54 \times 10^6$  sq mi) from 1980 to 2000, with a corresponding increase in the boreal forest. This decrease is linked to the warming of the Arctic due to climate change.

Arctic plants have a number of adaptations to the compressed growing season and low temperatures:

- They initiate growth rapidly in the spring, and flower and set seed much sooner than plants that grow in warmer conditions.
- Their peak metabolic rate occurs at a much lower temperature than plants from farther south, but only peaks for a short growing season.
- Some Arctic plants grow close to the ground as cushion plants, which keep the plants close to the warm soil and shield the tender central growing shoot.

- Arctic plants limit their height to be below the snow level. Plants that protrude above the snow are subject to strong winds, blowing snow, and being eaten by caribou, muskox, or ptarmigan.
- Arctic plants can survive very low temperatures because of high concentration of soluble carbohydrates, such as raffinose.
- Reproduction by vegetative propagation is common.
- Mosses and lichens are common in the Arctic. These plants have the ability to stop growth at any time and resume it promptly when conditions improve. They can even survive being covered by snow and ice for over a year.

J. D. Hooker never visited the Arctic region but he worked intensively at many collections made in various parts of the Arctic in the first six decades of the 19th century and now at Kew. The results of his studies are summarized in his paper “Outlines of the Distribution of Arctic Plants”, published in *Trans. Linn. Soc. Lond.* 23: 251–348 (1862). Apparently Hooker defined the “Arctic” very definitely by the Arctic Circle, at least the map accompanying his paper suggests this as does also the sentence “The arctic flora forms a circumpolar belt of 10° to 14° latitude, north of the arctic circle.” Hooker estimates the flora within the Arctic Circle to consist of 925 Cryptogams and 762 Phanerogams (214 Monocots. and 548 Dicots.).

. Only a few surveys of arctic, subantarctic or alpine mycorrhiza are available, e.g. studies by Russian scientists in the low Arctic of Siberia (Katenin 1964(Katenin , 1972, and by American scientists in Alaska (Laursen and Chmielewski 1982). In the high

Arctic, studies were made of the Canadian Devon and Ellesmere Islands (Stutz 1972;Kohn and Stasovski 1990).

The occurrence of mycorrhiza in *S. polaris* and *D. octopetala* agrees with earlier investigations (for references, see Table 3), but in contrast to most investigations we did not detect any ectomycorrhiza in *Polygonum viviparum* (Hesselman 1900;Katenin 1964Katenin , 1972Fontana 1977;Read and Haselwandter 1981;Lesica and Antibus 1986). Since Hesselman (1900) also examined mycorrhiza in Spitsbergen, *P. viviparum* is obviously ectomycorrhizal in some circumstances, e.g. when growing in the vicinity of other ectomycorrhizal plants. ...

... Many authors found the same herbaceous groups to be either VA mycorrhizal or non-mycorrhizal at alpine or arctic sites. For example, *Trisetum spicatum* is reported to be VA mycorrhizal by Lesica and Antibus (1986) but non-mycorrhizal by Katenin (1964), in common with the present findings. A general feature of these descriptions is the relative decrease in the number of plants colonized by VAM fungi and in the infection percentage towards polar sites and higher altitudes. ...

A study was done by Andrews C J. (1996). on ' How Plants Survive Ice?' and found that Plant species have had to adapt to freezing and the presence of ice in many climatic zones. Annual plants avoid ice by seed dispersal but, for biennials and perennials to survive they must cope with ice in various forms. Most plants that are regularly exposed to ice during their life cycles have acquired a dormant or quiescent winter period, when they are more tolerant to freezing temperatures. This Botanical Briefing explores some associations between plants and ice, with an emphasis on processes in plants that alleviate stress imposed by ice cover. Examples are taken from winter cereals which

must reach an equilibrium both with ice and with freezing temperatures for survival and economic productivity.

A study on ' How Plants perceive and respond to temperature stress ' was conducted by Ding, Yanglin and Yang, Shuhua and found that The dramatic temperature fluctuations spurred by climate change inhibit plant growth and threaten crop productivity. Unraveling how plants defend themselves against temperature stress-induced cellular impairment is not only a crucial fundamental issue but is also of critical importance for agricultural sustainability and food security. Here, we review recent developments in elucidating the molecular mechanisms used by plants to sense and respond to cold and heat stress at multiple levels. We also describe the trade-off between plant growth and responses to high and low temperatures. Finally, we discuss possible strategies that could be used to engineer temperature stress-tolerant, high yielding crops

In 1981, a journal called ' Pollination ecology' was published by the author Evgeniy A. Tikhmenev. The journal has the results of the observations on the blooming and pollination of plants of the Arctic tundra are discussed. High levels of pollen viability, prolonged functioning of flowers, a combination of different ways to contribute to the productivity of plants pollination in the specific context of high-latitude Arctic.

In 2018, Bokhorst S, Jaakola L, Karppinen K, Edvinsen G. K, Bjerke J. W published a work on 'Contrasting survival and physiological responses of sub - Arctic plant types to extreme winter warming and nitrogen. Planta. They found that Evergreen plants are more vulnerable than grasses and birch to snow and temperature variability in the

subArctic. Most Arctic climate impact studies focus on single factors, such as summer warming, while ecosystems are exposed to changes in all seasons. Through a combination of field and laboratory manipulations, we compared physiological and growth responses of dominant sub-Arctic plant types to midwinter warming events (6 °C for 7 days) in combination with freezing, simulated snow thaw and nitrogen additions. We aimed to identify if different plant types showed consistent physiological, cellular, growth and mortality responses to these abiotic stressors. Evergreen dwarf shrubs and tree seedlings showed higher mortality (40-100%) following extreme winter warming events than *Betula pubescens* tree seedlings and grasses (0-27%). All species had growth reductions following exposure to - 20 °C, but not all species suffered from - 10 °C irrespective of other treatments. Winter warming followed by - 20 °C resulted in the greatest mortality and was strongest among evergreen plants. Snow removal reduced the biomass for most species and this was exacerbated by subsequent freezing. Nitrogen increased the growth of *B. pubescens* and grasses, but not the evergreens, and interaction effects with the warming, freezing and snow treatments were minor and few. Physiological activity during the winter warming and freezing treatments was inconsistent with growth and mortality rates across the plants types. However, changes in the membrane fatty acids were associated with reduced mortality of grasses. Sub-Arctic plant communities may become dominated by grasses and deciduous plants if winter snowpack diminishes and plants are exposed to greater temperature variability in the near future.

# **Saxifraga oppositifolia**

**Kingdom : Plantae**  
**Phylum : Tracheophytes**  
**Clade : Angiosperms**  
**Clade : Eudicots**  
**Class : Magnoliopsida**  
**Order : Saxifragales**  
**Family : Saxifragaceae**  
**Genus : Saxifraga**  
**Species : S. oppositifolia**



## ***HABIT***

Contributing a splash of colour to otherwise rocky gray surroundings, this low, matted plant blooms with many beautiful lilac-coloured or magenta flowers. It is a lowgrowing, densely or loosely matted plant growing up to 5 cm (2.0 in) high, with somewhat woody branches of creeping or trailing habit close to the surface. The leaves are small, rounded, scale-like, opposite in four rows with ciliated margins. An arctic alpine cushion plant, this perennial has many tiny, overlapping, gray-green, scale-like leaves, arranged in ranks of four. The leaves are only 2 mm – 4 mm (1/10 – 1/5 in.) long, but they completely cover the plant's short stems. If you look closely with a hand lens, tiny rigid hairs can be seen at the leaf edges.

The purple (rarely, white), star-i flowers are large, in comparison to the leaves (1 cm – 1.5 cm [1/5 - 3/5 in.]). They grow on short stalks, often in great profusion, above the cushion of leaves, with five sepals and five petals. Male and female flowering parts are purple like the petals, but when the anthers open, it's easy to see the bright orange pollen. The fruit is a two-beaked capsule. The root system is extensive, with a main tap root about 50 cm (20 in.) long, and many side branches.

Purple saxifrage is one of the earliest flowering plants in its arctic or alpine habitat (e.g. April in the mountains; June in the Arctic). It begins to flower very soon after the snow melts, often right beside a snow bank. The timing of flowering is dependent on the time of melting – the temperature increase following snowmelt triggers flowering. This plant can be found blooming in cool, shady spots as late as the end of July. Most of the flowers in a cushion or colony open at the same time.

This plant is one of several species studied as part of the International Tundra Experiment (ITEX), which explores the impact of climate change on the growth, phenology and reproductive success of arctic plants. The greatest change in climate and plant response is predicted for high latitude areas.

## ***HABITAT***

*Saxifraga oppositifolia*, the purple saxifrage or purple mountain saxifrage, is a species of plant that is very common in the high Arctic and also some high mountainous areas further south, including northern Britain, the Alps and the Rocky Mountains.

It grows in all kinds of cold temperate to Arctic habitats, usually found from sea level up to 1,000 m (3,300 ft), in many places colouring the landscape. Its native habitats include tundra, arctic coastal bluffs, alpine scree, and rock crevices. It is seen in Fellside snow-bed sites, stream banks, wet rock surfaces, Lappish rock faces, rocky places, gravels, fell heaths, etc.

Swiss botanist Christian Körner found the plant growing at an elevation of 4,505 m (14,780 ft) in the Swiss alps, making it the highest elevation angiosperm in Europe. As purple saxifrage grows broadly also in more southern mountainous areas, it has been chosen for international monitoring to help study the climate and its relation to the plant's flowering time. A rapid warming of the climate can be especially destructive to fell plants. Purple saxifrage grows in the highest parts of the fell zone, where only a few

flowers survive. As Finland's only purple-lowered saxifrage, this species cannot be confused with any other saxifrage. It bears closest resemblance to moss campion (*Silene acaulis*), although its leaves are thin and needle-like. It is even known to grow on Kaffeklubben Island in north Greenland, at 83°N 40°E, the most northerly plant locality in the world.

## ***LIFE CYCLE & ADAPTATIONS***

Paleobotanists suggest that purple saxifrage evolved in a high alpine region, so it was well adapted to spread successfully into the arctic. This plant has a very long life span. Its habit of growing low to the ground protects it against evaporation and abrasion by wind-driven sand or snow. The foliage is low and crowded, and living parts don't project very high above the ground surface. Plants put out about two pairs of new leaves each growing season, so a dense cushion of this plant only 10 cm (4 in.) wide is probably many decades old.

Purple saxifrage is also adapted to the very short growing season of arctic and alpine environments. The flowering buds overwinter in an advanced stage (ready for blooming), protected by the foliage. Once the snow cover melts, flowering occurs in about 5-16 days. Individual flowers last about 12 days. The plant requires less than two months (about 54 days) from first flowering for seeds to ripen, and this timespan is even shorter if flowering starts later in the season.

Purple saxifrage can propagate by either self-fertilization or cross-fertilization. The second method is better in the long run, for the major key to long-term survival in the natural world is genetic variation through cross-fertilization, where flowers receive pollen from the flowers of a different plant. Cross-fertilization is normally assisted by insects. The purple flowers of saxifrage are attractive to bumble bees, moths and butterflies in the early part of the flowering season, and to small flies later on. However, in the harsh arctic or alpine climate, insect pollinators may be uncommon or, on cold days, inactive. The female parts of the flower (the stigma) mature first, allowing pollination with pollen from another open flower if insects are available. If crosspollination does not occur within two to four days, the male flower parts (anthers) lengthen and curve inward to facilitate self-fertilization by the first gust of wind. However, in purple saxifrage, self-pollination produces many fewer seeds, some of which are infertile. No matter what method of fertilization occurs, most seeds are not released from the seed capsule until some snow covers the ground. Only then can the wind carry the seeds easily away over the smooth, frozen surface.

Saxifrage often grows in the same environments as white dryad and other low-growing plants that can withstand harsh conditions. These mountainous or tundra areas are very sensitive to trampling, so in order to protect these beautiful and sensitive plants, hikers should stay on trails. If the saxifrage plant is growing on limestone rock, watch for flecks of calcium carbonate on the leaf tips. The minerals are extruded through a pore called a “hydatode,” usually visible with a hand lens. This species is specifically adapted to conditions of low nitrogen, and can grow in very poor soil. Many arctic species have evergreen leaves – an important adaptation to the short growing season. Purple saxifrage, *Saxifraga oppositifolia*, utilizes its evergreen leaves during the first

days of sunshine after the long winter. This plant is so efficient that it is one of the very first to bloom in the Arctic – often in early June, even before all the snow has melted. After the snow melts in alpine or arctic regions the flowers of purple saxifrage appear almost as soon as the ground is warm. Once in bloom, the flowers usually last 10 to 14 days. In areas where the snow is quite deep, and the ground does not thaw until late in the summer, purple saxifrage may not be present at all. Flowering abundance in saxifrage is strongly related to the intensity of the sun from July to September of the preceding season. Because much of its floral bud development occurs in the fall of the previous year, the amount of time it is not covered in snow, and can continue developing, has a large impact on the abundance of flowers the next year. Consequently, overall solar radiation is an important factor, because the snow-free period is mainly determined by the amount of sun between July and September.

Experiment designed to examine effect of increased temperature on the growth of purple saxifrage in the wild showed that this plant would not be able to compete with other plants after colonization of the "warmed" Arctic. That's why climate changes represent the major threat for the survival of purple saxifrage in the wild.

# **Papaver radicum**

**Kingdom : Plantae**  
**Phylum : Tracheophytes**  
**Clade : Angiosperms**  
**Clade : Eudicots**  
**Class : Magnoliopsida**  
**Order : Ranunculales**  
**Family : Papaveraceae**  
**Genus : Papavar**  
**Species : P.radicatum**

**Genus**



## ***HABIT***

*Papaver radicatum* is a species of poppy known by the common names Arctic poppy, rooted poppy, and yellow poppy. It is a flowering plant in the family Papaveraceae. They have a self-supporting growth form. They have simple, broad leaves and capsule fruit. The plant is known for appearing in white and yellow colorations. The yellow variant is much more widely distributed, occurring out in the open, compared to the white flowers typically only found in more remote locations.

*Papaver Radicatum* has hair like structures covering the stem and sepals that are used to retain heat to act as insulation. The plant also shows heliotropism where the flower grows to face the sunlight, and the petals reflect light towards the pistil. The petals grow in a cup shape that are used to trap light and air inside to be used as warmth for the ovaries. The leaves of Arctic poppy grow in a dense rosette around the base of the stem are lance-shaped and pinnate. The flowering stems of the Arctic poppy are around ten centimetres long, but the rest of the plant is relatively low-growing.

## ***HABITAT***

Arctic poppies are one of the most northerly growing plants in the world. Covered in black hair they are hardy and tough, with some surprisingly delicate yellow or white petals. These flowers continually turn to face the sun, tracking its progress across the sky, and attracting insects to the centre of the bloom. Arctic poppies are very rare and uncommon globally. Some species are abundant in certain areas but within a narrow range. It is circumpolar in distribution and grows in arctic and alpine zones in Europe, North America, and Asia.

It grows at a latitude of 83°40'N on Kaffeklubben Island, making it one of the northernmost plants in the world. It appears on the Coat of arms of Nunavut. Arctic poppies are found growing in meadows, mountains and dry river beds. They thrive among stones that both absorb the sun's heat and provide shelter for the roots. Within their Arctic range they are most common in the Nunavut region of Canada.

The Arctic poppy (*Papaver laestadianum*) is a rare, endemic, perennial plant species found only in the harsh, Arctic conditions of the northernmost parts of Scandinavia. The solitary flower head consists of four vibrant yellow petals, supported by an erect stem which is long and narrow and has thick, spreading black hairs. Within the cupshaped flower are five stigma protruding from a flat disk. The arctic poppy likes to hang out in Arctic meadows. It also likes to grow in the mountains, in the gravel beds of dry rivers, or in rocky high plains. It likes to grow among the rocks because the stones absorb heat from the sun and provide moist shelter for the roots. Arctic poppies are very rare and uncommon globally. Some species are abundant in certain areas but within a narrow range.

## ***LIFE CYCLE & ADAPTATIONS***

The Arctic Poppy is about 10 to 15 cm tall at the end of their growing. The weight of it is very light since it is a plant. Some colours it comes in are yellow and white. The Arctic Poppy is supposed to be very tough except for the petals of it because it is very delicate. The poppy is covered with black hairs. The flower is formed from four petals and is then created into a cup like form. The kingdom is Plantae, the Phylum is

Magnoliophyta, the Class is Magnoliopsida, Order is Papaverales, Family is Papaveraceae, Genus is Papaver, and species is papaver radicum. The gestational period of a poppy is: first it is a seed, then it starts to sprout some roots, then the roots get bigger and a stem starts to form, finally it gets bigger and bigger until it is a fully grown Arctic Poppy.

Sowing seeds is one of the easiest and best ways to propagate poppies. In a wild environment, poppies reproduce rapidly by self-seeding. To sow poppy seeds, simply scatter the seeds over a garden bed or meadow. Sow poppy seeds in fall or early spring.

The Arctic poppy has five distinctive adaptations to live in the Arctic. These characteristics help them to survive in cold and windy conditions.

- The whole plant is covered in black hairs and that could be an adaptation because the hairs could be a warning sign that tells predators not to come near it and to retain heat.
- The plant had a low toxicity and that is probably an adaptation that they got over time so animals wouldn't eat it.
- The white colour of the white arctic poppies is an adaptation because it camouflages with the arctic environment that it lives in.

- Arctic poppies are heliotropic. This means that the poppy turns its face to the sun so practically its face follows the sun to survive. This is an adaptation for this plant so it could get heat from the sun.
- The poppy adapted to growing on rocks and now it adapted so much that the rock allows the roots of the poppy to be moist. It is low lying to protect it from cold winds and has thin leaves to reduce water loss by transpiration.

# **Salix arctica**

**Kingdom** : Plantae  
**Phylum** : Tracheophytes  
**Clade** : Angiosperms  
**Clade** : Eudicots  
**Clade** : Rosids  
**Class** : Dicotyledonae  
**Order** : Malpighiales  
**Family** : Salicaceae  
**Genus** : Salix  
**Species.** : S. arctica



## ***HABIT***

*S. arctica* is typically a low shrub growing to only 15 cm (6 in) in height (rarely to 25 cm (10 in) high), but in the Pacific Northwest, it may reach 50 cm (20 in) in height. It has many different shapes, but sometimes has long trailing branches that root where they touch the surface. Rock willows/Arctic willows grow prostrate, shrub, and carpet.

The leaves are oval shaped with pointed tips, wedge shaped bottoms, and have little stalks. These leaves are 15-50 mm in length. The leaves are dark green on the bottom and a lighter green on the top; they are pubescent, with long, silky, silvery hairs. Like the rest of the willows, Arctic willow is dioecious, with male and female catkins on separate plants. As a result, the plant's appearance varies; the female catkins are redcoloured, while the male catkins are yellow-coloured. It has many different shapes, but sometimes has long trailing branches that root where they touch the surface. The flowers of the *Salix arctica* are upright scaly spikes that are unisexual flowers with no petals. The *Salix arctica* blooms in the spring. They are 5 cm and are dark brown or sparkling pink. There's no fruit on the plant, only seeds. There are no taproots on this plant. The lateral roots are shallow due to the frozen ground underneath the permafrost. Despite its small size, it is a long-lived plant, growing extremely slowly in the severe arctic climate; one in eastern Greenland was found to be 236 years old.

## ***HABITAT***

The Arctic willow grows in tundra and rocky moorland, and is the northernmost woody plant in the world, occurring far above the tree line to the northern limit of land on the north coast of Greenland. Its distribution is circumpolar. It occurs in Canada in the mainland northern territories and in the Arctic Archipelago all the way up to Ellesmere Island alongside Greenland, and in northern Quebec and Labrador, as well as in northern Iceland, Fenno-Scandinavia, northern Russia and northern Alaska.

Arctic willow or rock willow can also be found in the North American tundra. The North American tundra consists of Northern Alaska and Northern Canada. It stretches from 52° N to 65° N and from 60° W to 165° W. It also occurs further south in North America on high-altitude alpine tundra, south to the Sierra Nevada in California and the Rocky Mountains in New Mexico, to Xinjiang in China in Asia.

The Tundra can be described as a treeless plain with uneven ground. Each year this treeless plain gets 15-35 cm (6-14 in) of precipitation. There are two seasons in the tundra. They are winter and summer. *Salix arctica* prefers to live in dry, cold, open places, hummocks in wet sphagnum bogs, sedge meadows, margins of pools, and muddy salt flats in coarse sandy soil. The Arctic willow likes cold climates, which makes the tundra a perfect place for it because the average temperatures range from 70 degrees F to 20 degrees F.

## ***LIFE CYCLE & ADAPTATIONS***

The Arctic willow or *Salix arctica* is a resident of the North American Tundra and has adapted to the weather conditions of this region. The Arctic willow forms a pesticide to keep the predator insects (such as the Arctic woolly bear) away. Since rainfall is extremely low (6-14 inches annually) in these regions, the Arctic willow survives by

possessing a shallow root system, which allows it to absorb soil moisture at shallow depths. Such a shallow root system also helps the growth of these plants in permafrost. To protect itself against the cold weather, the Arctic willow has adapted by growing long fuzzy hairs and growing close to the ground. Such adaptations protect the plant from winds. These plants grow like a carpet gaining the heat energy from the ground and protecting themselves from the cold weather in tundra regions.

Over a long period of time, the arctic willow has made many remarkable adaptations living in the tundra which have allowed success in their living. One of their many abilities are to produce a pesticide for itself to protect insects from damaging it. It now grows very short with long, fluffy hair which prevents them from cold temperatures. The arctic willow only has shallow roots, not long roots, to stop them from growing in permanently frozen soil, however, shallow roots can be damaged more easily. Thus, small changes or alterations made in the tundra may drastically affect the arctic willow's existence.

The *Salix arctica* is very plentiful in the wild, but is still vulnerable because of the delicate habitat it lives in. It has a shallow root depth, which makes them susceptible to root damage. It also has a short growing season. There is also a limited food supply for herbivorous insects. There is also a low N.P.P. (nitrogen/ phosphorus/ potassium) in the North American tundra. This means that the chemical fertility of the soil is low. There are many studies being done on the *Salix arctica*. Some of the organizations involved are I.T.E.X. (International Tundra Experiment), S.A.G.E. (Sustainable Arid Grassland Ecosystems). International Tundra Experiment placed open top chambers

(O.T.C) in order to raise the average temperature a couple of degrees to find out how it affected the plants, insects, etc. Sustainable Arid Grassland Ecosystems studied the arctic grasslands and plants.

The Tundra is a delicate place where tire tracks can last for years. The *Salix arctica* has adapted well in these frigid non-fertile conditions. In such conditions small changes could drastically affect the *Salix arctica*.

# **Cassiope tetragona**

**Kingdom : Plantae**  
**Phylum : Tracheophytes**  
**Clade : Angiosperms**  
**Clade : Eudicots**  
**Clade : Asterids**  
**Class : Magnoliopsida**  
**Order : Ericales**  
**Family : Ericaceae**  
**Genus : Cassiope**  
**Species : C. Tetragona**



## ***HABIT***

White arctic mountain heather is a mat-forming perennial sub-shrub. It is hermaphroditic (containing both male and female reproductive parts). The stems are decumbent (on the ground) to erect, forming a mat, and are covered in overlapping evergreen leaves that form a scaly pattern around the stem (imbricate). The leaves are simple, 2-6 mm long, narrowly triangular, and arranged in rows of 4 around the stem.

The flowers arise from pedicels emerging from between the leaves. The pedicels are about twice the length of the leaves and have one flower per pedicel. There may be multiple flowers per stem. The flower has 5 yellow or green sepals, about 2 mm long. The white petals are connate with spreading tips. The petals are united, forming a bellshaped corolla. The corolla contains 5 ovaries and 10 stamens.

There are two subspecies found in North America. *Cassiope tetragona* ssp. *tetragona* with pedicels that extend beyond the branch tips and with larger flowers and *Cassiope tetragona* ssp. *saximontana* with pedicels not extending beyond the branch tips and smaller corollas (4-6 mm).

Arctic white heather is a familiar sight to visitors to Halti because it grows in many places as the main dwarf shrub and its dark green patches stand out in the landscape. Even flowerless it is easy to identify from its stiff form and scaly leaves, which are

clearly grouped into four rows. The shoots look as if they are braided, and the Saami name for the species is “plait-plant”. Wide, dense stands of Arctic white heather usually flower abundantly and are very impressive, and the flowers are somewhat reminiscent of lily-of-the-valley. In arctic areas the plant was commonly used as fuel

– and Finnish reindeer herders and ramblers have also used it for this purpose. Nowadays the Lapland visitor is better using a portable stove to cook his meals and leave the slowly-regenerating arctic white heather to grow in peace. Shoots can reach up to 20 years old.

## ***HABITAT***

*Cassiope tetragona* has a widespread circumboreal distribution. It’s a very common plant in Alaska, northern Canada, Scandinavia, and northern Russia.

Moderately thermophilous. White arctic mountain heather frequently grows in alpine tundra and on ridges. A dominant species in heaths where frost movements of the ground is limited. Often growing on slopes with good drainage and some snow protection during winter. The range suggests that it is slightly basiphilous, not present in the areas in the south and west with the most acidic substrates.

In the middle arctic tundra zone and the clearly and weakly continental sections, slightly transgressing into the transitional section. *Cassiope tetragona* is present only on Spitsbergen and only north of Van Mijenfjorden. It is rather common in the areas

around Isfjorden, Kongsfjorden, Liefdefjorden, Woodfjorden, and Wijdefjorden, with a few sites at Sorgfjorden and Lomfjorden (NE Spitsbergen, Ny-Friesland).

Globally, this is one of the more widespread of arctic plants, present and usually common on all northern continents, in Europe reaching south to N Fennoscandia.

### ***LIFE CYCLE & ADAPTATIONS***

Cassiope tetragona are Perennial and they are potentially very long-lived. They are Mat-forming dwarf shrub with extensive below and above ground branching resulting in large, often very dense stands up to 10–20(30) cm tall. Branching is potentially at opposite leaves but often only one branch developed. Their Aerial branches densely covered by leaves.

#### **Leaf :**

Leaves are opposite, 4–5 × 1.3–1.7 mm, evergreen, lasting for 4–5 years or more, appressed and densely imbricate on branches in four rows, thick, boat-shaped, with a very short petiole (ca. 1 mm) hidden inside the leaf cover, oblong to triangular, obtuse, with the mid vein impressed in a groove along the back, dark green, densely pubescent with very short hairs, young leaves sticky.

#### **Inflorescence :**

Flowers 2–4 together at the end of previous year's shoot, overtopped by the emerging current year's shoot.

#### **Flower :**

Flowers nodding on a 7–14(16) mm, slender, red or yellowish pedicel curved at the end.

Flowers radially symmetric with 5 free sepals and 5 fused petals. Sepals 1.8–2.2 × 0.8–1.2 mm, oblong, obtuse or slightly dentate at apex, whitish or pink. Corolla 6–8 × 6–8 mm, bell-shaped, fused for 2/3–3/4 of its length with 5 triangular petal lobes revolute at their ends, cream coloured. Stamens 5, ending ca. 5 mm inside the corolla; filaments 2–2.5 mm; anthers almost orbicular, ca. 1 mm, with two long spurs opposite to the two pores opening for pollen dispersal. Gynoecium of 5 carpels, with a 2 mm long, straight, slender style and stigma exposed near the mouth of the corolla.

**Fruit :**

Fruit a nearly globular capsule, 2.5–4 × 2.5–4 mm, becoming erect in fruit stage, opening by 5 apical slits into 5 valves, with numerous, very small seeds. The style remains on the capsule until seeds are dispersed.

**Reproduction :**

Sexual reproduction by seeds; very restricted and local vegetative reproduction by layering. Flowers are strictly adapted to insect pollination. The spurs on the anthers function as levers; the anthers only empty pollen through the pores when an insect moves the spurs. Seed set is regular in Svalbard, vouching for efficient insect pollination. Seeds germinate to ca. 8 % in an experiment (Alsos et al. 2013); they probably depend on some interaction with fungi for efficient germination (mycotrophic). Seeds are very small and light and dispersed by wind.

The hairs on the stems of many tundra plants , help to trap heat near the plant and act as protection from the wind. Plants adapted to the tundra have small waxy leaves to

prevent the loss of precious water in this dry environment. Plants also have adapted to the Arctic tundra by developing the ability to grow under a layer of snow, to carry out photosynthesis in extremely cold temperatures, and for flowering plants, to produce flowers quickly once summer begins. A small leaf structure is another physical adaptation that helps plants survive.

# **Picea pungens**

**Kingdom : Plantae**  
**Phylum : Tracheophytes**  
**Clade : Gymnosperms**  
**Division : Pinophyta**  
**Class : Pinopsida**  
**Order : Pinales**  
**Family : Pinaceae**  
**Genus : Picea**  
**Species : P. pungens**



## ***HABIT***

The Colorado Blue Spruce is a natural inhabitant of forests and mountain valleys in the western United States of America. A popular tree (not a poplar tree!) used for landscaping worldwide and used as a windbreak, *Picea pungens* is a beautiful looking Conifer.

*Picea* means spruce, and *pungens* means pointed, in reference to its upright, needlelike leaves. These are covered in a glaucous wax, which gives the tree its characteristic bluegreen colour. The cylindrical cones are green, turning pale brown as they ripen. The tiers of horizontal branches form a conical shape that becomes more columnar as the tree grows.

Blue Spruce trees at maturity range anywhere from 60 to 100 feet tall, some reaching as tall as 150 feet! Colorado Blue Spruce used in landscaping From top to bottom this Spruce has stiff horizontal branches spreading 10 to 20 feet in diameter and the trunk, up to 5 feet in diameter. The Colorado Blue Spruce, *P. pungens*, is a slow growing tree but with the right conditions and conservation can grow quite large. Spruce trees are used all around the country and even grown internationally. Harvested for wood in the western portion of the nation, the trees main commercial use is for the sale of Christmas trees.

The needles of the Colorado Spruce are primarily concentrated above the stem, approximately 1 inch in length and brutally sharp. Blue Spruce needles They consist of 4 sides with 6 stomatal lines on each side allowing for gas exchange . These needles are blue, blue-green in color sometimes with almost a silver tint to them and very rigid.

The Colorado Blue Spruce provides homes for many creatures. Grouse living among the floor of a forest Spruce Grouse or *Falcipennis canadensis* lives strictly in conifer forests, taking refuge and nesting in spruce trees. These fowl rely on the Spruce for safety, protection, occasionally food, and reproduction (taking care of their eggs). Spruce trees get nothing in return and therefore this relationship is considered commensal.

The blue spruce will thrive on any fertile, moisture-retentive but free draining soil with a neutral to acidic pH. They grow in a sunny spot. It does not need pruning. Feed with a balanced fertiliser in spring to encourage healthy growth. The blue spruce does not need regular pruning. You could remove any dead, damaged or crossing shoots in spring. Keep well watered after planting while the roots are establishing; mature plants should get all the moisture they need from rainfall. Feed with a balanced fertiliser in spring to encourage healthy growth of the seed. Colorado Blue Spruce with new male cones The roots of the tree on the other hand display negative phototropism (they move away from the light) and positive gravitropism (they move toward gravity). Using these stimuli the Blue Spruce tree is able to sense the environment around it and react however it deems necessary. The Colorado Spruce is found growing in mountain valleys, but chiefly found near streams, where moisture levels of the soil are high. The strong, adaptable roots of the Spruce can even withstand changes in pH and even drought (up to a certain point).

Interactions between *P. pungens* and other species are numerous and come in all varieties (mutualism, commensalism, etc.). As review, mutualism benefits both organisms, commensalism is a symbiotic relationship where one species benefits while the other is unaffected, and in parasitism one species benefits and the other is harmed.

It has been estimated that upwards of 70% of all land plants form mutualistic relationships with mycorrhizal fungi, one of the most important associations the plant has. In the pine family the majority of the fungi are classified as ectomycorrhizae. plants and fungi share nutrientsThe relationship that takes place beneath the soil, where ectomycorrhizae surround the cells in the roots of their spruce substrate with hyphae, proves advantageous for both sides. The attached mycorrhizae increase the surface area of the roots of the plant, and also allow for transportation of nutrients and minerals from plant to fungus.

## ***HABITAT***

The blue spruce (*Picea pungens*), also commonly known as green spruce, white spruce, Colorado spruce, or Colorado blue spruce, is a species of spruce tree. It is native to North America, and is found in USDA growing zones 1 through 7. It is found naturally in Arizona, Colorado, Idaho, New Mexico, Utah and Wyoming. It has been widely introduced elsewhere and is used as an ornamental tree in many places far beyond its native range. The blue spruce has blue-green colored needles and is a coniferous tree.

They grow on any fertile, moisture-retentive but free draining soil with a neutral to acidic pH, in a sunny spot. The blue spruce looks particularly good planted with other contrasting conifers, winter heathers and other plants with winter interest, such as dogwoods. The blue spruce is drought tolerant when established and is suitable for exposed or windy sites.

*Picea pungens* is most commonly found growing along streamsides in mountain valleys, where moisture levels in the soil are greater than the often low rainfall in the

area would suggest. Large conifer forests in the western U.S. hold the majority of wild Blue Spruce, although you may see Spruce populations nearly anywhere in that area. The Rocky Mountains in Colorado, home to many Colorado Blue Spruce The Rocky Mountains of Colorado are a great place to find a Blue Spruce. Most often found wild in Utah and Colorado, occasionally making appearances in New Mexico and Wyoming. Native grown Spruce tend to live in higher altitudes from 1,750-3,000 m but nonnative grown Spruce can inhabit any range of landscape due to their ability to adapt. *P. pungens* is a species of spruce native to western North America although it is cultivated in all 50 states and commonly cultivated in Europe. The Blue Spruce, despite its limited natural range, is able to grow under a wide variety of conditions, and it is considered highly desirable as a landscape plant due to the unusual blue-gray color of its foliage.

Conifer forests cover large areas but have few plant and animal species. Other conifers, and even other Spruces, such as the Norway Spruce in natural habitat Norway Spruce (*Picea abies*) dwell in these areas. Animals mentioned before, such as the spruce grouse and many insect and worm species reside among the trees and in the soil. Other birds may nest in these trees and contribute to the biome and ecosystem.

The Colorado Blue Spruce plays a crucial role in this habitat. The ecological niche of *Picea pungens* growing in the conifer forests of mountain valleys includes absorbing

light, water and nutrients (for photosynthesis), providing shelter for birds, small game and insects and providing food for organisms such as squirrels and spruce grouse, as well as giving off oxygen into the atmosphere.

## ***LIFE CYCLE & ADAPTATIONS***

The Colorado Blue Spruce is a very adaptable species that can withstand quite drastic changes within its environment. The Spruce displays positive phototropism as well as negative gravitropism (or geotropism) in its stem as it reaches towards the light out of the seed. Colorado Blue Spruce with new male cones The roots of the tree on the other hand display negative phototropism (they move away from the light) and positive gravitropism (they move toward gravity). Using these stimuli the Blue Spruce tree is able to sense the environment around it and react however it deems necessary. The Colorado Spruce is found growing in mountain valleys, but chiefly found near streams, where moisture levels of the soil are high. The strong, adaptable roots of the Spruce can even withstand changes in pH and even drought (up to a certain point).

Interactions between *P. pungens* and other species are numerous and come in all varieties (mutualism, commensalism, etc.). As review, mutualism benefits both organisms, commensalism is a symbiotic relationship where one species benefits while the other is unaffected, and in parasitism one species benefits and the other is harmed. Blue Spruce tree

It has been estimated that upwards of 70% of all land plants form mutualistic relationships with mycorrhizal fungi, one of the most important associations the plant has. In the pine family the majority of the fungi are classified as ectomycorrhizae.

plants and fungi share nutrients. The relationship that takes place beneath the soil, where ectomycorrhizae surround the cells in the roots of their spruce substrate with hyphae, proves advantageous for both sides. The attached mycorrhizae increase the surface area of the roots of the plant, and also allow for transportation of nutrients and minerals from plant to fungus.

Like all members of Kingdom Plantae, *P. pungens* is autotrophic, meaning it can create its own energy; it is unnecessary for it to consume other species. This organism is more specifically known as being photosynthetic. Photosynthesis is the process by which plants use energy from the sun to produce sugar. The reaction uses water (H<sub>2</sub>O), carbon dioxide (CO<sub>2</sub>) and sunlight to produce sugar (C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) and oxygen (O<sub>2</sub>) seen here:



The tree gets the necessary water from the ground, which it absorbs through its roots. Carbon dioxide is taken from the atmosphere by its leaves. Plants undergo photosynthesis and light is necessary. Even more specifically, the CO<sub>2</sub> gas is transported across the cell wall by the stomata (or stomatal lines). Stomata allow for gas exchange in leaves. Finally the energy required for the reaction comes from the sun in the form of solar power. To take in light from the sun, plants use chloroplasts with specialized green pigments called chlorophyll to convert unusable energy from the sun into chemical energy that they can use to fuel the photosynthesis reaction. This produces oxygen (what we breathe) and glucose (sugar) for the plant. Plants take this sugar and through cellular respiration generate energy (ATP). All plants, including the Blue Spruce store their food as starch.

Specialized structures within the spruce's trunk allow for transport of this sugar, glucose, as well as water and many other substances. Vascular tissues called xylem and phloem, usually found together, help move sugars, water and other nutrients up, down, and to every part of the tree. To generalize the structures of each, the xylem moves water and nutrients up the trunk to the top of the tree and everywhere else necessary. The phloem moves sugars down to the roots and places in between. Moving water up, against the force of gravity is not an easy task. Here are a few properties of water that are vital to the UPWARD transport through the tree: water moves from an area of high water potential (low concentration) to low water potential (high concentration), water molecules are highly polarized and have a tendency to attach to one another (cohesion), and water molecules have a high affinity for other polar molecules (adhesion). Now, the question is, how does water get to the tops of trees? Water evaporating from leaves (transpiration) creates a lowered water potential of the mesophyll cells of the leaf. Water is transported to the top of the tree and to the tips of leaves. Due to the increased concentration of ions in the leaves, water from the high potential xylem moves into the low potential mesophyll cells. This reduces pressure in the xylem, which in turn draws water into the vessels. Adhesion and cohesion keep the chain from breaking and do not allow water molecules to detach from one another as they move up the trunk. Roots soak up water from the soil using this very same idea of water potential.

The Colorado Blue Spruce is quite tolerant to adverse conditions. It prefers moist, acidic soils, but can withstand dry soil, variant pH, and even winter salt spray, which makes it desirable for growth in colder climates.

The Blue Spruce serves as primary host to many organisms. Some of these include the engraver beetle, *Ips hunteri* an insect that feeds on the bark of the Colorado Spruce, the Leichhardt's grasshopper, as well as numerous bacteria and mycorrhizal fungus species.

Conifers are monoecious. This means that there are male cones and female cones on the same tree. Male cones, called staminate cones, give off pollen while their female counterparts, ovulate cones, catch the pollen blowing in the wind. This allows for the possibility of cross-fertilization which in turn increases the variation in the genes of offspring. Pollen (fertilizing haploid cell,  $n$ ) and the megaspore (receiving haploid cell,  $n$ ), once combined form a diploid cell. This diploid cell ( $2n$ ) then grows into a zygote (baby conifer) that eventually lives in a seed.

Male (red) and female (tan/brown) cones of the Colorado Spruce Male strobili (cones) develop all over the tree. Female strobili on the other hand generally grow on the upper 10% of the Spruce. Most male strobili of the Blue Spruce are a brilliant shade of red when they emerge from the buds and contain hundreds of thousands of grains of pollen. Female cones turn from a pale green to a deep bright red during peak receptivity and have the potential to produce upwards of 350 to 450 seeds each! Cones mature in August and seed shed begins around mid-September and continues into the winter. Mature Blue Spruce cone Germination of seeds doesn't take place until the spring or early summer. Optimum conditions for this germination in the habitat of the Blue Spruce are confined to open mineral soil with a good ratio of shade and direct sunlight.

The life cycle of *P. pungens* follows the normal life cycle of all coniferophyta dominated by the sporophyte generation (the tree). When the sporophyte tree is mature, it produces diploid (2n) male and female cones or strobili as mentioned before. These diploid cells undergo meiosis to become haploid gametophytes. These gametophytes are the microspores of the staminate cone and the megaspores of the ovular cone. Once these haploid spores pollinate (pollen meets megaspore) they go through fertilization to create a diploid zygote. The subsequent embryo develops within a naked seed and upon receiving the proper hormonal and environmental cues, germinates and develops a diploid sporophyte, or tree. The Colorado Spruce doesn't start producing seeds until it is about 20, and once the tree is over 150 seed production declines.

The blue spruce is also sold as a Christmas tree and its foliage is useful for making a Christmas wreath.

The Colorado Blue Spruce is a very adaptable species that can withstand quite drastic changes within its environment. The Spruce displays positive phototropism as well as negative gravitropism (or geotropism) in its stem as it reaches towards the light out of the seed. Colorado Blue Spruce with new male cones The roots of the tree on the other hand display negative phototropism (they move away from the light) and positive gravitropism (they move toward gravity). Using these stimuli the Blue Spruce tree is able to sense the environment around it and react however it deems necessary. The Colorado Spruce is found growing in mountain valleys, but chiefly found near streams, where moisture levels of the soil are high. The strong, adaptable roots of the Spruce can even withstand changes in pH and even drought (up to a certain point).

# **Viola × wittrockiana**

**Kingdom : Plantae**  
**Phylum : Tracheophytes**  
**Clade : Angiosperms**  
**Clade : Eudicots**  
**Clade : Rosids**  
**Class : Dicotyledonae**  
**Order : Malpighiales**  
**Family : Violaceae**  
**Genus : Viola**  
**Species : V. × wittrockiana**



## ***HABIT***

The garden pansy (*Viola* × *wittrockiana*) is a type of large-flowered hybrid plant cultivated as a garden flower. It is derived by hybridization from several species in the section *Melanium* ("the pansies") of the genus *Viola*, particularly *Viola tricolor*, a wildflower of Europe and western Asia known as heartsease. It is sometimes known as *Viola tricolor* var. *hortensis*, but this scientific name is ambiguous. While *Viola tricolor* var. *hortensis* Groenland & Rümpler is a synonym of *Viola* × *wittrockiana*, *Viola tricolor* var. *hortensis* DC. refers to a horticultural variety of wild pansy (*Viola tricolor* without interspecific hybridization) that had been illustrated in *Flora Danica* in 1777 before the existence of *Viola* × *wittrockiana*.

The chromosome number of *Viola* × *wittrockiana* is  $2n = 44-52$ , with most cultivars being  $2n = 48$ . The flower is 5 to 8 centimetres (2 to 3 in) in diameter and has two slightly overlapping upper petals, two side petals, and a single bottom petal with a slight beard emanating from the flower's center. These petals are usually white or yellow, purplish, or blue. The plant may grow to 23 cm (9 in) in height, and prefers sun to varying degrees and well-draining soils.

A favorite garden flower, the pansy is one of the oldest cultivated flowering plants. The markings of the petals often make the blossoms look like upturned faces. A bed of lowgrowing pansies—purple, blue, lilac, white, brown, orange, yellow—is sometimes used to give a garden border a brilliant accent of color. There are some 400 to 600 species of pansies. They all belong to the genus *Viola*, in the violet family, *Violaceae*.

The wild pansy (*Viola tricolor*) is native to cool parts of Europe but has been widely naturalized in North America. Old folk names for this pansy are johnny-jump-up, heartsease, and love-in-idleness. It is an annual or a short-lived perennial that grows about 6 to 12 inches (15 to 30 centimeters) tall. It has heart-shaped or rounded leaves at the base and oblong or oval leaves growing from the stems. The velvety flowers are usually purple and yellow, measure less than 1 inch (2.5 centimeters) across, and have five petals. Four petals are in pairs; the fifth has a short spur. The plant grows best in rich soil in a damp, cool climate.

## ***HABITAT***

Pansies prefers a cool moist well-drained humus-rich soil in partial or dappled shade and protection from scorching winds. It can tolerate sandstone and limestone soils but becomes chlorotic if the pH is too high. Prefers a pH between 6 and 6.5. A hybrid species of garden origin, there are many named forms selected for their ornamental value. There is at least one cultivar selected for its edible qualities - 'Gourmet Brand Salad Mixture' is a large-flowered, early-blooming mixture, selected to provide a large number of flowers for salads over the longest harvest season. All members of this genus have more or less edible leaves and flower buds, though those species with yellow flowers can cause diarrhoea if eaten in large quantities.

Pansies are surprisingly hearty in cold weather, though. They'll survive a frost, bouncing back from even single digit temperatures. If the blooms wither in the cold, the plants will often stay alive to bloom again, which makes them a great flowering plant for fall and early winter color. It flowers from April to September (in the Northern Hemisphere). The plants are hermaphrodite and self-fertile, pollinated by bees.

## LIFE CYCLE AND ADAPTATIONS

Pansies are one of the spring's charmers. Their sunny little "faces" and wide variety of colors elects them as one of the most popular bedding and container flowers. Pansies are actually quite hardy, but they bloom in cooler weather and hot temperatures can reduce flowering and make them leggy and unsightly. Pansies are biennials, meaning they have a two-year life cycle. They grow from seed but only produce leaves and stems in their first year. In the second year, pansies produce flowers and seeds and the plants then die. Garden centers sell 1-year-old pansies as bedding plants for spring or fall color. By the time you purchase them blooming, they are in their second year. Most commercially sold plants are hybrids and do not have cold hardiness or longevity. That being said, you can get pansies to survive into future years in temperate climates.

In the Pacific Northwest, pansies will often come back the next year or their prolific seedlings will provide year after year of color. Gardeners in the Midwest and South should assume their plants are annuals. So pansies are perennials but only in areas with short freezes, cool summers and moderate temperatures. The rest of us should treat them as welcome but short lived annuals. Even in regions where the plants can be used as perennials, they are short lived. The average pansy lifespan is only a couple of years. The good news is that a wide variety of the plants are offered as easy to grow seeds and, in some areas, they will naturally reseed themselves. That means the flowers may reappear the next year but just as second generation volunteers.

The female parts of the flower are in the pistil in the center of the flower where the petals meet. The pistil contains the stigma, style and ovary. It is shaped somewhat like

a vase, with the stigma the top of the vase, the style the neck of the vase and the ovary the wide vase bottom. The seeds of the pansy develop in an ovule inside the ovary. When insects pollinate the flower, the pollen is placed on the stigma to germinate when the insect reaches for the nectar that collects on the lower petal, directing the mouth of the insect toward the stigma

Pansy flowers have five petals. There are two overlapping, side-by-side upper petals. The lower half of the flower consists of three petals -- a large center spurred petal and two side petals. Each petal is covered with tiny hairs that get longer toward the center of the flower. These hairs give the petals a velvety feel. Pansies are grown in a wide variety of colors including solid, bicolor or tricolor in shades of white, yellow, orange, purple, brown, red and pink, among others. *Viola tricolor* is often white, yellow and purple. Some cultivars include "Hohenfeuer," a solid orange, "Moon Moth," a solid white and "Sunny Boy" which is yellow with a black face. Colorful petals attract pollinators while protecting the reproductive parts.

The male parts of the flower surround the stigma like a cone in the center of the flower. The stamen, consisting of a filament topped by an anther, produces the pollen. There are five anthers in each flower. Each anther is covered with a dusting of yellow pollen, which germinates when transferred to the stigma, fertilizing the ovule to create seeds. The stamens secrete honey that attracts insects. As the insects reach for the honey they pass through the pollen transferring it to the stigma.

At the top of the flower stem, or peduncle, are five sepals. The sepals are green and leafy, surrounding and protecting the flower bud as it develops. The sepals attach to the receptacle, which is the rounded part at the top of the stem. This is also where the parts

of the flower and petal structure attach to the stem when they develop. Pansy flowers are zygomorphic, meaning the right and left halves of the flowers mirror each other, but they have no other type of symmetry.

Pansies have a tolerance to low-temperatures and actually thrive in cool weather (40°60°F). They are capable of surviving temperatures down to the single digits, but when the air temperature drops below 25°F, pansy foliage will wilt and turn a graygreen color. This is a normal defense response to cold weather, and they usually bounce back with vigor when warm weather returns.

While Pansies can tolerate some cold, they absolutely do not like excessive heat and humidity. Heat causes pansies to become leggy and stop blooming. When summer warmth begins to get the upper hand, go ahead and remove pansies to make way for your summer annuals. However with the right conditions, you can try to get your pansies to survive the heat of summer. If you want to try to “over-summer” your pansies, plant them in a somewhat shady area.

Pansies adapt by using their roots to pretty much eat the nutrients and send it up the stalk to the leaf stalks then the flowers are able to survive and so is the rest of the plant. That is the roots and root hairs job. Then the stems job is to keep the plant standing, the leaf stalks job is to spread the flowers apart so that the flowers have the same amount of sun. Then the flowers have to sweat the water slowly from the sun.

# CONCLUSION

The importance of plants in our life is inexhaustible. They are very important for survival of every specie living on Earth. They provide us food, shelter, habitat and medicines. Just a single tree can provide food and shelter to many species like insects, birds and reptiles. These green plants produce food for all organisms and these foods are rich in vitamins and carbohydrates that are necessary for our health. We get clean water because of plants. They move water from soil to atmosphere through a process called transpiration. Without plants, our water cycle will be disrupted and our oxygen supply will be cut short because they improve oxygen level and reduce carbon dioxide level. In short, we will not be able to survive without plants.

Plants maintain balance in ecosystem. They also help to purify the air that we breathe. They help to fight global warming by maintaining the global temperature to save us from scorching heat. They also cool down the air as they lose moisture and reflect heat upward through their leaves. They maintain a balance in temperature by reflecting harmful radiations of sun. They play an important role to convert carbon dioxide to oxygen through a process called 'photosynthesis'.

So, more trees mean more oxygen and less carbon dioxide in atmosphere. If there are less plants that means there will be more carbon dioxide in air and with more CO<sub>2</sub>, more of sun's radiations will be reflected to earth, instead of space. The importance of plants is not just limited to being a source of food or oxygen. Plants are also used for various medicinal purposes. These medicines are safer because of lower side effects. There is a therapy called herbal therapy that is one of the oldest types of treatments this world has ever known. Herbal remedies have proven to be genuine and more effective.

About 1,702 species of plants live on the Arctic tundra, including flowering plants, short shrubs, herbs, grasses, mosses, and lichens. These plants are adapted to short, cold growing seasons. They have the ability to withstand extremely cold temperatures in the winter (winter hardiness), and grow and reproduce in summer conditions that are quite limiting.

As of 2005, arctic vegetation covered approximately  $5 \times 10^6$  km<sup>2</sup> ( $1.9 \times 10^6$  sq mi) of land. The area of arctic vegetation decreased by approximately 1.4 million square kilometres ( $0.54 \times 10^6$  sq mi) from 1980 to 2000, with a corresponding increase in the boreal forest. This decrease is linked to the warming of the Arctic due to climate change.

Arctic plants have a number of adaptations to the compressed growing season and low temperatures:

- They initiate growth rapidly in the spring, and flower and set seed much sooner than plants that grow in warmer conditions.
- Their peak metabolic rate occurs at a much lower temperature than plants from farther south, but only peaks for a short growing season.
- Some Arctic plants grow close to the ground as cushion plants, which keep the plants close to the warm soil and shield the tender central growing shoot.
- Arctic plants limit their height to be below the snow level. Plants that protrude above the snow are subject to strong winds, blowing snow, and being eaten by caribou, muskox, or ptarmigan.
- Arctic plants can survive very low temperatures because of high concentration of soluble carbohydrates, such as raffinose.
- Reproduction by vegetative propagation is common.

- Mosses and lichens are common in the Arctic. These plants have the ability to stop growth at any time and resume it promptly when conditions improve. They can even survive being covered by snow and ice for over a year.

The cold is a restrictive factor to plant growth. It can be caused for two main reasons: height and high latitudes. When the height raises, the cold also does it; for each 100 meters of height, temperature gets down 1°C. And in high latitudes, cold is caused by low insolation (only a little amount of sun's heat is received). Plants can live until certain limits in high mountains, originating the alpine biome, and even become an ecosystem above the polar circle in the northern hemisphere, forming the tundra biome. Therefore, plants can survive in these cold ecosystems somehow.

The trees' growth is very restricted in both biomes. Indeed, trees are missing in tundra and only can be found in subalpine zone in the high mountains, between 1.600m and 2.400m; even so, the biggest height where trees can occur depends of different climatic factors and of the topographic relief. On the other hand, shrubs are uncommon in both biomes, being the most of them smaller and creeping. That way, they can protect themselves against heavy frosts and cold winds, because they get covered of snow during the unfavourable period. Cranberry bush (*Vaccinium vitis-idaea*) is a good example of this kind of shrubs. The herbs, bryophytes (e.g. mosses) and lichen together, are the most dominant of these two biomes, because they are the most abundant.

Plants from climates with cold winters have evolved to survive winter by going dormant. That means not just dropping leaves and slowing or stopping growth, but also reducing the amount of water in branch and root tissues. The lowered concentration of water in a plant's tissue acts like a natural antifreeze: It means it takes deeper cold to form ice inside them.

Plants are very important for survival of living beings. It is because of plants that we breathe fresh air. It is our responsibility to take care of plants because our lives depend on their survival. Without them, every specie will die because of lack of oxygen on planet.

Unfortunately, people cut down these trees to make furniture and other wooden materials. We need to understand that plants also have a life. Therefore, it is very important to have awareness about plants and taking necessary measures to promotes tree plantation.

## REFERENCES

- Billings, W. D., & Mooney, H. A. (1968). The ecology of arctic and alpine plants. *Biological reviews*, 43(4), 481-529.
- Gellhorn J. (2002). *Song of the alpine: The Rocky Mountain tundra through the seasons*. Big Earth Publishing.

[https://arcticbioscan.ca/wiki/w/White\\_Heather](https://arcticbioscan.ca/wiki/w/White_Heather)

- [http://bioweb.uwlax.edu/bio203/s2009/loch\\_spen/History%20and%20Interesting%20Facts.htm](http://bioweb.uwlax.edu/bio203/s2009/loch_spen/History%20and%20Interesting%20Facts.htm)

<https://blueplanetbiomes.org>

- <https://en.m.wikipedia.org>
- <https://en.m.wikipedia.org/wiki/Arctic>
- [https://en.m.wikipedia.org/wiki/Arctic\\_vegetation](https://en.m.wikipedia.org/wiki/Arctic_vegetation)
- [https://en.m.wikipedia.org/wiki/Blue\\_spruce](https://en.m.wikipedia.org/wiki/Blue_spruce)
- [https://en.m.wikipedia.org/wiki/Dryas\\_octopetala](https://en.m.wikipedia.org/wiki/Dryas_octopetala)
- <https://homeguides.sfgate.com/parts-pansy-flower-70070.html>
- <https://kids.britannica.com/students/article/pansy/276269>
- <https://luontoportti.com/en/t/1325/mountain-avens>

- <http://svalbardflora.no/oldsite/index.php?id=404>
- <http://svalbardflora.no/oldsite/index.php?id=589>
- <https://www.aeco.no>
- <https://www.britannica.com/plant/buttercup>
- [https://www.flora.dempstercountry.org/o.Site.Folder/Species.Program/Species.php?species\\_id=Cassi.tetra](https://www.flora.dempstercountry.org/o.Site.Folder/Species.Program/Species.php?species_id=Cassi.tetra) <https://www.gardenersworld.com/how-to-grow-plants/how-to-grow>
- <https://www.gardeningknowhow.com/ornamental/bulbs/ranunculus/growing-ranunculus-buttercups.htm>
- <https://www.grida.no>
- <https://www.gvsu.edu/itex-aon/snow-buttercup-42.htm>
- <https://www.inform.umd.edu/Edles/colleges/Lfsc/life-sciences>
- <https://www.lwpetersen.com/alaska-wildflowers/white-arctic>
- <https://www.lwpetersen.com/alaska>
- Körner, C. (2021). Alpine plant life: functional plant ecology of high mountain ecosystems. Springer Nature.
- Matveyeva, N., & Chernov, Y. (2000). Biodiversity of terrestrial ecosystems. The arctic: Environment, people, policy, 233-274.

- Seval, H. F. (2019). Arktik bölge'de uluslararası siyasi düzen: teorik bir yaklaşım. Akdeniz *İİBF* Dergisi, 19, 1-24.

FEMINIST IDENTITY AND PSYCHOLOGICAL WELL-BEING

**PROJECT REPORT**

**A STUDY ON THE RELATIONSHIP BETWEEN FEMINIST IDENTITY  
DEVELOPMENT AND PSYCHOLOGICAL WELL-BEING**

Submitted by:

THERESA JAISON

Register No.:

SB19PSY039

Under the guidance of

MS. ANJITHA VENUGOPAL

In partial fulfilment of the requirement for award of the degree of

**B.Sc. PSYCHOLOGY**



**ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM**

Nationally Re-accredited at 'A++' level (4<sup>th</sup> cycle)

Affiliated to: Mahatma Gandhi University

**MARCH 2022**

## CERTIFICATE

This is to certify that the project report entitled, “THE RELATIONSHIP BETWEEN FEMINIST IDENTITY DEVELOPMENT AND PSYCHOLOGICAL WELL-BEING”, is a bonafide record submitted by MS. THERESA JAISON, Reg. No. SB19PSY039, in partial fulfilment of the requirements for the award of the Degree of Bachelor of Psychology during the academic year 2019-2022.



*Bindu John*

Ms. Bindu John

Head of the Department

Department of Psychology

St. Teresa's College, Ernakulam

*Anjitha Venugopal*

Ms. Anjitha Venugopal

Assistant Professor

Department of Psychology

St. Teresa's College,  
Ernakulam

External Examiner:

*Jinson*  
10/10/22

**DECLARATION**

I, Theresa Jaison, hereby declare that the study presented in the dissertation entitled, “The Relationship Between Feminist Identity Development and Psychological Well-being”, which is submitted to the Department of Psychology, St. Teresa’s College, Ernakulam is a bonafide record of the research work carried out by me, under the supervision and guidance of Ms. Anjitha Venugopal, Assistant Professor, Department of Psychology, St. Teresa’s College, Ernakulam, in partial fulfillment of the requirements for the degree of Bachelor of Science in Psychology and has not previously formed the basis for the award of any degree, diploma, fellowship, title or recognition before.

Place: Ernakulam

Theresa Jaison

Date: 05/05/2022

### **ACKNOWLEDGEMENT**

Undertaking a research project for the first time requires constant assistance and guidance. Hence, I would like to express my gratitude to the Department of Psychology, St. Teresa's College, Ernakulam for providing me with this opportunity.

I acknowledge my gratitude to my research guide, Ms. Anjitha Venugopal, Assistant Professor, Psychology, for encouraging and guiding us throughout all the phases of our research.

I extend my sincere thanks to my parents, teachers and my research partners who supported me throughout this time. I am grateful to everyone who gave me guidance, encouragement, suggestions and constructive criticisms which immensely improved this project.

Above all, I thank God Almighty for blessing me in all the stages of the project and for helping me complete the project successfully.

Thanking you

Theresa Jaison

**CONTENTS**

<b>SL. NO.</b>	<b>TOPIC</b>	<b>PAGE NO.</b>
1.	CHAPTER I: INTRODUCTION	6
1.1	Background of the Study	7
1.2	Problem Statement	8
1.3	Need and Significance of the Study	8
1.4	Scope of the study	9
1.5	Objectives of the study	9
1.6	Limitations of the Study	9
2.	CHAPTER II. REVIEW OF LITERATURE	10
	Literature Review	11
3.	CHAPTER III: THEORETICAL FRAMEWORK	16
4.	CHAPTER IV: RESEARCH AND METHODOLOGY	25
4.1	Objectives	26
4.2	Hypotheses	26
4.3	Research Design	27
4.4	Sources of Data	27
4.5	Sample Design	28
4.6	Sample Size	28

4.7	Sampling Method	28
4.8	Method of Data Collection	29
4.9	Drafting a Questionnaire	29
4.10	Data Analysis Techniques	30
5.	CHAPTER V: DATA ANALYSIS	33
6.	CHAPTER VI: FINDINGS	36
7.	CHAPTER VII: RECOMMENDATIONS	58
8.	CHAPTER VIII: CONCLUSIONS	60
	REFERENCES	62
	APPENDIX A	71
	APPENDIX B	72

**CHAPTER I**  
**INTRODUCTION**

## INTRODUCTION

### 1.1 BACKGROUND OF STUDY

One of the most famous social and political movements in the world, The Women's Rights Movement arose from the displeasure of a handful of women and went on to better women's lives for centuries, starting from a single spark of rebellion against the institution of patriarchy. Though the major goals of the feminist movement involve reform on issues like equal pay, reproductive rights, and sexual harassment, the priorities vary among communities and individuals. Whether the opposition is against domestic violence or genital mutilation, the idea of feminism is still looked at with contempt and derision.

Despite holding feminist ideas some women still seem to fear wearing the label of a feminist. The word has been twisted and misinterpreted throughout history. When feminists like Laura Cereta published *Epistolae familiares* (Collected Letters of a Renaissance Feminist) in the 15th century containing a panoply of letters from women about the denial of education and marital oppression to the absurdity of women's garments, they would not have had the faintest idea that six centuries down the lane, the world is still dealing with the same kind of problems. From a pro-sati society, India has also grown along with the world. The same cannot be said about the rural and interior parts of the country for they still deal with female infanticides and child marriages because daughters are a "burden".

The progress, if any, is that today women are more opinionated and assertive in their views on feminism. However, living in a society that continues to propagate traditionalist views, the effect of such feminist ideals on their well-being and as a result, their ability to fully function as a member of the society is yet to be fully uncovered.

## **1.2 PROBLEM STATEMENT**

The following research seeks to study the relationship between feminist identity and psychological well-being in Indian Women. The study also aims to ascertain the reliability of the Feminist Identity Composite Scale (FIC) in Indian women.

## **1.3 NEED AND SIGNIFICANCE OF THE STUDY**

When a photo of the Twitter CEO holding a placard ‘Smash Brahmanical Patriarchy’ hit social media platforms in 2018, multiple groups took offence and deemed it “hate speech” and an attempt to “malign the community”. More recently Indian men trended #MarriageStrike on Twitter to express their disapproval of criminalising marital rape, which currently is legal in India. Some deemed it a win-win for the feminist movement in the country as the men who trended the hashtag refused to marry if their future wives are given the power to “consent” to sexual relations with the husband. In India, forced sexual relations are not considered a criminal offence if it happens under the institution of legalised marriage. Kerala boasts of having a literacy rate of 96.2% but the cases of dowry deaths in the second half of 2021 shocked the entire state. A large part of the educated population still holds traditionalist views that are anything but progressive. When the court decided to open Sabarimala Temple’s gates to women of menstruating ages, the majority of the Hindu population, men and women alike, protested against the change. While the entrance of women is allowed legally, it is far from reality. Most often they are stopped at the footsteps of the temple by temple staff who deem it disruptive to the peace on the temple grounds.

Regardless of its source, the experience of oppression has an effect on the development and well-being of an individual. With increased patriarchal practices, women and other minorities have come to face adversities in various aspects of their lives. Even today identities of women are often linked with the men involved in their lives in collectivistic patrilineal societies like India. Women are also more likely to have depression, anxiety and somatic complaints. In the 21st century, women still are at a social disadvantage in India. There is a need to understand

how much of this disadvantage is tolerated by women today and how its acceptance or resistance is affecting their well-being and in the long term, their ability to function competently as a member of society.

#### **1.4 SCOPE OF THE STUDY**

This study adds to the growing research literature on gender studies and sheds light on how institutionalised injustice can be detrimental to mental health. At its core, feminism remains a movement for full economical, social and political equality for women. Feminism today is an umbrella term that advocates for equality among people regardless of their gender. It seeks to interrogate and eradicate the inequalities along the lines of sexuality, gender, sex, race, ability and class. Feminists all over the world seek to explore the inequities along these intersectional lines and eliminate them. Since feminist ideologies have not garnered enough positive attention in the mainstream Indian culture, this study aims to explore the development of feminist identity and the impact it can have on well-being. The study attempts to ensure the reliability of the Feminist Identity Composite scale in an Indian population for future reference.

#### **1.5 OBJECTIVES OF THE STUDY**

- To find a relationship between feminist identity and psychological well-being in Indian women.
- To ascertain the reliability of the Feminist Identity Composite Scale in Indian women.

#### **1.6 LIMITATIONS OF THE STUDY**

This study was carried out diligently to the best of our ability. However, there are a few limitations that may have a possible influence on the results. Due to the COVID-19 pandemic, the data was primarily collected via Google Forms. The majority of the women under the sample were South Indians/Keralites. Hence, a large scale study coupled with personal interviews of each woman in the sample is required to completely understand the relationship between feminist identity and psychological well-being

**CHAPTER II**  
**REVIEW OF LITERATURE**

## LITERATURE REVIEW

Feminism is best understood as that which “emphasizes the validity of women’s own interpretations of their lived experiences and needs, protests against the institutionalized injustice perpetrated by men as a group against women as a group, and advocates the elimination of that injustice by challenging the various structures of authority or power that legitimate male prerogatives in a given society” (Yakushko, 2007; Offen, 1988). In this male-dominated society, gender and gender discrimination are observed to have a significant impact on mental health, as women have a less social advantage, and it has been noted that the scarcity of protective factors like feeling respected or valued increases this risk (Basu, 2012).

Empowerment and solidarity among women, promotion of gender equality (Harlan, 1998), and awareness of the social and personal impacts of living in a patriarchal culture can lead to the advancement of the women’s movement and better living conditions for women. According to feminist theorists, such awareness can facilitate distinguishing healthy and socially ingrained behaviour, thereby promoting personal freedom in making healthy life decisions (Saunders & Kaushbeck-West, 2006; Prochaska & Norcross, 1999). In this manner, from what was earlier limited to equal rights, feminism of today has a broader area of coverage. As a result of intersectionality, various types of feminism emerge, such as socialist feminism, liberal feminism, and lesbian feminism to name a few (Henley, Meng, O’Brien, McCarthy, & Sockloskie, 1998; Szymanski, 2004; DeBlaere et al., 2019), each pioneering a diverse definition of the feminist ideology.

The multiracial feminist theory hypothesises that feminist identification, the salience of feminism in women’s lives and what women understand feminism to be, may differ across racial and ethnic groups (Harnois, 2005; Robnett and Anderson, 2017; Liu & Zheng, 2019). In the US, Caucasian females look for equal academic, employment, and political opportunities whereas African-American females seek ways to seamlessly blend family life along with their careers while trying to reduce the effect of racism in their lives. Moving over to collectivistic Asian

culture, feminist reforms have diverse goals and only lately involved diverse individuals from the grassroots level at the forefront, instead of the state and intellectuals. Additionally, the fourth wave of feminism in China has ideals aligning with international feminist goals such as security from sexual misconduct and addressing LGBTQ issues along with an equal division of family and work roles (Liu & Zheng, 2019). Marginalized women with their unique social locations due to their multiple identities develop feminist identities differently from their nonmarginalized counterparts. DeBlaere and Bertsch in 2013, explored the relationship between perceived sexist events and distress in sexual minority women of colour and found that as women better understand their multiple identities, they learn to better identify their experiences with discrimination.

In psychological research, approaches that have dominated the feminist literature involve a developmental model that arose from counselling psychology and a social psychology model. Measures of feminist identity and its correlates have mostly been based on Downing and Roush's theoretical model developed in 1985 based on Cross' (1971; Cross & Vandiver, 2001) developmental model of politicized Black identity. This approach identified five feminist identity stages through which an individual progresses and clearly defines how one moves from unquestioned acceptance of a dominant group's beliefs about oneself to consciousness and activism regarding one's disadvantaged position in society. The stages include Passive Acceptance, Revelation, Embeddedness-Emanation, Synthesis, and Active Commitment, which are further elaborated in the theoretical framework. Feminist identity as politicized group identity is an extensively studied area in social psychology. To describe the constituents of group consciousness, three broad categories are utilised Identity, Injustice, and, Efficacy.

Based on the developmental model, Feminist Identity Scale (FIS) (Rickard, 1989) and Feminist Identity Development Scale (FIDS) (Bargad and Hyde, 1991) were constructed first. The Feminist Identity Composite (FIC; Fischer et al., 2000) includes the best items from its two predecessors. Rooted in the social identity approach, Feminist Perspectives Scale (FPS; Henley et al., 1998) and Attitudes Toward Feminism and the Women's Movement Scale (FWM;

Fassinger, 1994) are widely used. Interestingly, there are significant overlaps between the developmental and social models of feminist identity development, for instance, the Revelation and Injustice are based on the awareness of social inequalities and Active Commitment and Efficacy are rooted in the behavioural manifestation of the feminist identity. Thus, self-labelling or identifying as a feminist or non-feminist has been linked to the extent of a woman's feminist actions.

However, feminist identities may differ depending on how questions are asked. Around 81% of women who did not consider themselves feminists agreed with some or all of the goals of feminism. Adopting the label of "feminist" can cause unease due to stereotyping in many women, who instead choose to identify with their specific beliefs (Liss et al., 2001). Indicating the broader implications of the development of feminist identity, Szymanski et al. (2009) suggest that there is a need to focus on internalized misogyny, the subconscious devaluation or mistrust of other women and belief in male superiority, particularly in therapy. Internalized misogyny can aggravate the relationship between sexist events and psychological distress in heterosexual women. Since women have much less economic and political power than men in the patriarchal structure of the world, sociopolitical roots play a major role in the aetiology and maintenance of many mental health struggles faced by women (McNamara & Rickard, 1989). Thus, feminist identity and its development is a powerful tool in therapy for women.

In a study to explore the well-being among university students with relatively homogenous backgrounds alongside their gender roles, September et al. (2001) administered a test that clearly emphasized the distinction between biological sex and gender roles without explicitly mentioning the terms *feminine* and *masculine*. The results indicated that the majority of women were classified as androgynous, suggesting that they are more comfortable with defying stigmatized gender roles perhaps due to the rising heat of feminism. Women who are in the later stages of feminist identity are more likely to identify as bisexual or lesbian (Yakushko, 2007; Simoni et al. 1999). The study also sheds light on the fact that women with feminist identities were more likely to have graduate degrees, work outdoors and have more income. Ryff (1995)

defines psychological well-being as feeling good about one's self and one's life, the sense that one is continuing to grow and develop as a person, the belief that life has meaning and purpose, having good relationships with others, the ability to manage life effectively, and a sense of self-determination. A theoretical model devised based on this suggested six components namely, Self-Acceptance, Positive Relations, Autonomy, Environmental Mastery, Purpose in Life, and Personal Growth each plays a role in determining the well-being of a person (Ryff, 1995; September et al., 2001). Researchers also found demographic differences in various aspects of well-being. Holding strong feminist beliefs has been linked with high self-esteem (Fischer and Good 1994), self-efficacy (Eisele and Stake 2008), academic achievement (Valenzuela 1993), rejection of feminine norms for thinness and appearance (Hurt et al. 2007), sexual well-being (Schick et al. 2008), and sexual openness (Bay-Cheng and Zucker 2007).

Exposure to feminist ideology is said to have a negative correlation with depressive episodes in females (Mauthner, 1998; Weitz, 1982). Also, Yakushko in 2007 found that women who held Moderate and Feminist Values as opposed to Traditional Values had a significantly higher measure of overall well-being. Since FIC has been proved to be the most appropriate scale so far to measure feminist identity development, over 20 studies have used it. The FIC subscales have been related to optimism (Peterson et al., 2008), autonomy (Yoder, Snell, & Tobias, 2012), racial acceptance and awareness of White privilege (Wolff & Munley, 2012), multiple psychological distress variables (e.g., negative affect, depression, phobic anxiety, body image concerns; Blue & Berkel, 2010; Clarke, Murnen, & Smolak, 2010; Fischer & Good, 2004; Murnen & Smolak, 2008; Sabik & Tylka, 2006), and indicators of well-being (e.g., self-esteem, self-efficacy, personal growth, self-acceptance, personal agency, self-empowerment; Peterson et al., 2008; Yakushko, 2007; Yoder et al., 2012). However, the majority of the FIC studies have been done with undergraduate students and focus on the experiences of heterosexual women (DeBlaere et al., 2019). When Szymanski (2004, 2005) used the scale with sexual minority women in two studies, one yielded results that align with the heterosexual sample whereas the other one produced contradictory results. Moreover, Caucasian American females have been the focal samples of such studies. The internal consistency reliability estimates for each stage of FIC

are 0.74 (passive acceptance), 0.75 (revelation), 0.86 (embeddedness-emanation), 0.71 (synthesis) and 0.81 (active commitment) in these studies. However, when administered to African-American students, Cronbach's alpha was lower for all five subscales (Blue & Berkel, 2010) and in the case of women from mainland China, four items of the FIC had to be dropped and two others had poor factor loadings (Liu & Zheng, 2019). Similarly, the convergent, discriminant and construct validity only reached the recommended cut-offs in the Caucasian females but had variations in the marginalised samples. These studies suggest that feminist identity development is closely tied to the socio-cultural experiences of a woman and this model may not apply universally, especially to the individuals belonging to marginalized groups.

In a country like India, where people are much more insular and community-oriented, feminism is often misunderstood and therefore frowned upon. Through the handful of Indian studies regarding gender or feminist identity development, the consensus reached was that problems women face do not stem from internal personal deficiencies but rather from the society itself (eg. sexism, racism) (Srivastava, K, 2007). With globalisation, waves of Western ideals of individualism swept through the country, especially influencing the millennial and Gen-Z women. The two generations of women are perhaps much more open to declaring themselves as feminists. There, however, remains great uncertainty in the well-being of such women in a largely collectivistic country like India. Further research among women is required to ascertain how strong their ideals are and how much is it affecting their well-being in the country. There is also an urgency to assess psychometric properties like reliability and validity of the FIC in diverse samples (eg. Indian women) for future revisions.

**CHAPTER III**  
**THEORETICAL FRAMEWORK**

## **THEORETICAL FRAMEWORK**

In light of the existing literature, there is a lack of authentic research on the psychological well-being of Indian women with respect to their feminist identity development. To bridge this gap, this research aims to gather and analyse data from Indian women to find possible correlations between their feminist identity and psychological well-being.

Feminism can be explained as “a concept that can encompass both an ideology and a movement for sociopolitical change based on a critical analysis of male privilege and women's subordination within any given society” (Offen, 1988). The American Psychological Association defines it as “any of a number of perspectives that take as their subject matter the problems and perspectives of women or the nature of biological and social phenomena related to gender”. The concept of identity is also outlined by the APA, as involving “a sense of continuity, or the feeling that one is the same person today that one was yesterday or last year.” Erikson defines it as a “fundamental organising principle which develops constantly throughout the lifespan.” This process of identity formation uses all levels of mental functioning, with self-reflection and observation (Kaur, 2021; Erikson, 1974).

Feminist identity may be described as what a young woman forms in late adolescence as the foundation for her adult identity (Kaur, 2021; Josselson, 1987). Eisele and Stake (2008) define it as “a woman’s collective or social identity that involves adopting feminist attitudes and identifying as a feminist”. It may also be defined as “a woman’s recognition of gender-based boundaries and inequality and a drive toward their demise” (Colaner & Rittenour, 2015; Allen, 2000).

A theory of feminist identity development that supports the current study is the model proposed by Downing and Roush (1985). This theory has been based on Cross’ (1971) theory of positive Black identity development, comprising five stages: pre-encounter, encounter, immersion-emersion, internalisation and internalisation-commitment.

Downing and Roush's model also describes five stages—passive acceptance, revelation, embeddedness-emanation, synthesis and active commitment. Women may go through these stages at different speeds, cycle through them over and over, stagnate at certain points or revert to earlier stages.

The first stage, Passive Acceptance (PA), involves a state in which women are not aware of, or in denial of the individual and institutional prejudice against them. Women in this stage are accepting of the male dominating system and its perspectives like traditional gender roles and the idea that men are superior. They may not engage with people who disturb the balance they have created for themselves. During the transition period to the next stage, women may become 'ready' to change or take a risk (Downing & Roush, 1985; Erickson, 1950).

The second stage, Revelation (R), is activated when women experience one or more crises that they are no longer able to ignore, such as an unhappy marriage, inability to pursue an education or career, witnessing discrimination against their daughters, encounters with activists for the women's rights movement or gender-based violence or abuse. This change may be sudden, or more often, gradual, and may be made difficult by the perceptual distortions that are typical of women in the passive acceptance stage. During the revelation stage, women feel both anger, as feeling betrayed or duped by the universe (Downing & Roush, 1985; Avery, 1977), as well as guilt for their participation in their own oppression and lack of wanting to change things. Dualistic thinking of all men as negative and all women as positive is also seen.

The third stage, Embeddedness-Emanation (EE), is in particular, more difficult for women, as the dominant culture is so deeply ingrained in them in all aspects of their lives. Women may feel the need to withdraw from their other circles and engross themselves in creative activities such as art, music or drama that show how oppressed women are. A close emotional connection is created with other women, giving them chances to strengthen their identities and release their anger into a support system. In the second half of this stage, women become open to different viewpoints and a more relativistic perspective, while still exerting

caution around men. They gradually realise that simply raging about oppression is not dismantling the dominant culture and come to an awareness that their beliefs in the embeddedness phase are as rigid as in the passive acceptance phase.

The fourth stage, Synthesis (S), involves an integration of positive female qualities with one's own personality to create a realistic self-concept that surpasses gender roles and is capable of evaluating men individually rather than with collective stereotypes. Women in this stage have "struck a flexible truce" (Downing & Roush, 1985; Avery, 1977) with the world. They are able to direct their energy in a productive manner and respond to discrimination appropriately.

The fifth stage, Active Commitment (AC), is concerned with women learning to put this newfound identity into meaningful action to make a change in society. This may be done by using their unique talents to choose issues that are relevant to them, and to which they can contribute. Only a few women progress to this stage; many women who externally appear to be working for their rights may be doing so out of some unfulfilled need leftover from previous stages like revelation or embeddedness.

For further studies to be carried out on feminist identity development, there must be measures to assess these levels in women. To evaluate feminist identity development, three scales have been created on the basis of this model.

The Feminist Identity Scale (FIS) was developed by Rickard in 1987. This scale, though based on the Downing and Roush model, dropped the dimension of Active Commitment as it was thought to be a "behavioural manifestation of S-level integration" (Rickard, 1989), and less a separate identity. A study later conducted (Rickard, 1989) to assess the relationship between dating behaviours and feminist identity level attempted to validate the FIS. Liss et al. (2001) assessed the factors that predicted and hindered feminist social identity in college women, using the FIS to assess feminist identity development.

The Feminist Identity Development Scale (FIDS) was developed by Bargad and Hyde in 1991 to assess the effect of educational courses in women's studies on the feminist identity development of women who attended them. While other studies measured attitude changes, self-esteem changes, career aspirations, and generalized beliefs of social roles due to women's studies, there were none assessing personal conceptions of feminism at the time. The accompanying study using the developed scale showed that women's studies courses did in fact contribute to the development of a feminist identity (Bargad & Hyde, 1991). Moradi and Subich (2002) investigated feminist identity development and perceived sexist events along with their relation to psychological distress, using the FIDS to assess feminist identity.

The Feminist Identity Composite Scale (FIC) was developed by Fischer et al. (2000) in response to the need for more sophisticated assessments of feminist identity development. It intended to operationalize the Downing and Roush model, retaining the five dimensions. The above scales of FIS and FIDS were examined and a composite measure was derived "using their best items" (Fischer et al., 2000) with better psychometric properties, consistency and validity. As the feminist identity development of a population of Indian women is being evaluated for the first time, it is fair to believe that the FIC would be a more appropriate scale in the current study.

Moradi and Subich (2002) compared the reliability and validity of the three scales assessing feminist identity development described above and found that only the Feminist Identity Composite scale had acceptable internal consistency reliability for all subscales. Content validity was found to be best for the Feminist Identity Development Scale; overall trends indicated that the FIC was superior. Hyde (2002) commented on this evaluation, stating that there were inconsistencies and that both the FIDS and FIC are comparable measures

Nevertheless, several studies have used the FIC to examine feminist identity. Fischer and Good (2004) found that the dimension of Revelation was linked with greater psychological distress. A study by DeBlaere in 2013 on the moderating role of womanism between perceived sexism and psychological distress, showed comparable results that womanism correlated

positively with psychological distress. Saunders and West (2006) investigated the relationships between feminist identity development, gender role orientation and psychological well-being; a more developed feminist identity as measured by the FIC, related positively to psychological well-being. Parallel to these findings, Yakushko (2007) reported that women who had more feminist values scored higher in overall well-being than those with traditional values.

As Ryff (1995) defined, psychological well-being includes “feeling good about one’s self and one’s life, the sense that one is continuing to grow and develop as a person, the belief that life has meaning and purpose, having good relationships with others, the ability to manage life effectively and a sense of self-discrimination.” The concept of well-being can also be described as, “optimal psychological functioning and experience” (Ryan & Deci, 2001). Two traditions can typically be observed in the study of well-being, that is, the hedonic view and the eudaimonic view.

The hedonic view equates pleasure or happiness with well-being. This approach has historically been seen in many forms and its focus may range from bodily pleasures to broader self-interests (Ryan & Deci, 2001). More modern concepts of hedonism have included pleasures of both the mind and body (Ryan & Deci, 2001; Kubovy, 1999) and the belief that well-being comprises subjective happiness and experiences of pleasure and displeasure. Hedonic contents of well-being include satisfaction and comfort and are associated with self-focus and the ‘here and now’ (Huta, 2016).

The eudaimonic view suggests that not all desires a person has will create a sense of well-being when they are achieved. Not all pleasure producing outcomes promote wellness. Eudaimonia is thought to occur when the life activities of people are most congruent with their deeply held values (Ryan & Deci, 2001; Waterman, 1993), and in this case, people would feel most authentic and engage in activities that led to personal growth and development. Eudaimonic contents of well-being would include personal maturity, ethics, authenticity, autonomy and broader meaning to life, as well as associations with balancing the self and others, balancing the

present and future and creating one's own vision (Huta, 2016). This approach has been adopted by Ryff, as evaluated in 1989, in which it was found that many theories of well-being overlap with some similar features. These make up the core dimensions of Ryff's six-factor theory of psychological well-being, detailed as follows.

The first dimension is *self-acceptance*, regarded as the most recurrent criterion of psychological well-being. Being able to accept oneself is a defining feature of good mental health, maturity, ideal functioning and reaching one's full potential. The second dimension involves *positive relations with others*, or the ability to love and be loved, and to have trusting interpersonal relationships. People who have reached self-actualization are described as having more empathy for humans and being capable of deeper relationships. The third dimension is *autonomy*, which emphasises independence and self-control. A fully functioning person is said to have an internal locus of evaluation, using personal standards to assess themselves rather than seek approval from others to feel self-worth. A person should ideally have a sense of freedom of choice without needing to rely on collective fears or beliefs of society.

The fourth dimension is *environmental mastery*, that is, the ability of a person to create environments that are suited for their mental state. Many theories of well-being stress the necessity of active participation, advancing in the world and changing it with various physical or mental activities, also see it as a sign of maturity. The fifth dimension, *purpose in life*, is included in definitions of mental health and maturity. A person should have a sense of direction or meaning in life, as well as certain goals to achieve. The last dimension, *personal growth*, suggests that a person should continue to grow and expand themselves, to increase their potential throughout their life. Instead of stopping at a state wherein, they have no problems, they may try to remain open to new experiences. This dimension comes closest in meaning to the notion of eudaimonia.

In an attempt to operationalize these dimensions for future examination, a scale for psychological well-being was created (Ryff, 1989), by defining each of the six aspects of positive functioning.

For example, a high scorer in self-acceptance has a positive attitude towards themselves and acknowledges their good as well as bad qualities, while a low scorer is not satisfied with themselves and wishes to be someone different. A high scorer in autonomy is self-determining and independent, while a low scorer conforms to peer pressure and needs others to make decisions for them.

Ryff's Psychological Well-being scales were used in a study assessing the impact of self-labelling as a feminist and feminist beliefs on women's well-being, egalitarianism and activism (Yoder et al., 2011). The scales were adjusted for a single measure of overall well-being, and it was found that self-labelling was not related to personal well-being, but feminist beliefs were. Another study by Yoder et al. (2012) used three of the six independent scales of Ryff's well-being measure, namely, autonomy, self-acceptance and personal growth. It was discovered that feminist beliefs and well-being with liberation are complexly and significantly related.

September et al. (2001) examined this model of well-being in young adults using Ryff's scales in a study of the relation between well-being, impostor feelings and gender role orientation. From the study, it was found that well-being derived from positive interpersonal relationships is higher in people with more expressive traits, and well-being derived from autonomy is higher in those with more instrumental traits. Saunders and West (2006) also used Ryff's measure to determine the relationships between feminist identity, gender roles and psychological well-being. Here, instrumentality, expressiveness and feminist identity were positively associated with well-being. Yakushko in 2007 had comparable findings, also using Ryff's scales to measure six domains of subjective well-being.

Ryff's Psychological Well-being scales have become a very widely used measure of positive psychological functioning, made evident by the abundant publications utilising this scale. As stated above, this measure has been used in studies of feminist identity and well-being in other countries. Based on the existing literature, it is convincing that this scale would be the best fit for the current study.

**CHAPTER IV**  
**RESEARCH AND METHODOLOGY**

## RESEARCH AND METHODOLOGY

### 4.1 OBJECTIVES

- To find a relationship between feminist identity and well-being in Indian women.
- To ascertain the reliability of the Feminist Identity Composite Scale in Indian women.

### 4.2 HYPOTHESES

The primary hypotheses are,

1. H<sub>0</sub>: There is no significant relationship between subscales of the Feminist Identity Composite Scale and subscales of Ryff's Scales of Psychological Well-being.  
  
H<sub>1</sub>: There is a significant relationship between subscales of the Feminist Identity Composite Scale and subscales of Ryff's Scales of Psychological Well-being.
2. H<sub>0</sub>: The Feminist Identity Composite Scale is not reliable for the chosen sample population.  
  
H<sub>1</sub>: The Feminist Identity Composite Scale is reliable for the chosen sample population.

The secondary hypotheses are,

3. H<sub>0</sub>: There is no significant relationship between age, Feminist Identity and Psychological Well-being.  
  
H<sub>1</sub>: There is a significant relationship between age, Feminist Identity and Psychological Well-being.

4. H<sub>0</sub>: There is no significant association between identifying as a Feminist and Psychological Well-being.

H<sub>1</sub>: There is a significant association between identifying as a Feminist and Psychological Well-being.

### **4.3 RESEARCH DESIGN**

The study done is based on a quantitative research design aimed at understanding the relationship between feminist identity and well-being in Indian women. Non-probability snowball sampling was used to collect responses from the participants. Consent to participate in the study and the demographic details like name, age, employment status and place of residence were collected as well. The questionnaires circulated via Google Forms yielded initial responses from 345 women of which 5 fell under the exclusion criteria. Therefore, the final sample size was 340. The Feminist Identity Composite Scale and The Ryff Scales of Psychological Well-being (18-item version) were the tools utilized for the same. The data was organised using Microsoft Excel and the statistical analysis was done using IBM SPSS software version 28.0.1.1 (14). Pearson's Correlation test, Regression Analysis and Chi-square test were run and probable associations between variables were found. Cronbach's Alpha was found using the same software to check the reliability of the questionnaires among the Indian population.

### **4.4 SOURCES OF DATA**

The data was collected by circulating Google Forms online. Non-probability snowball sampling was used to collect the same. The participants were Indian women who are of the age 18 or above.

#### **4.5 SAMPLE DESIGN**

The sample design used was non-probability snowball sampling. Sampling is a technique of selecting a representative part of the population such that it has all the characteristics of the population. From this sample, statistical inferences are made and generalised to the population. In non-probability sampling, the selection of the sample is done randomly. As opposed to probability sampling, every element in the population does not have an equal chance of being selected. Snowball sampling is used when direct access to the sample is limited (eg. due to COVID-19). Here the researchers pass on the questionnaire to a few participants who in turn pass it on to others who fit the criteria.

#### **4.6 SAMPLE SIZE**

The total sample size was 345, of which 5 responses were excluded due to evident response bias or not fitting the inclusion criteria. The sample was predominated by women from Kerala although it included women from other states of India as well, such as Delhi, Karnataka, Tamil Nadu, Maharashtra, Telangana, Meghalaya, Andhra Pradesh, Uttar Pradesh, West Bengal and Goa.

#### **4.7 SAMPLING METHOD**

For this study, the sample was chosen based on two criteria: (i) the participant has to be of Indian nationality and (ii) the participant has to be of the age 18 or above at the time of responding. Making use of snowball or network sampling, the responses were collected to meet a controlled quota. The researchers contacted a few individuals who met the selection criteria and the sample grew like a snowball rolling downhill to include a sizable population with these characteristics. The response collection carried on until the specified quota criteria were met and 345 responses were obtained. From this, 5 were excluded due to evident response bias and failure to meet the sampling criteria, leaving 340 valid responses.

## **4.8 METHOD OF DATA COLLECTION**

The data collection was primarily via Google Forms. The questionnaire was drafted consisting of asking the demographic details like name, age, state of residence, socioeconomic status and whether they identify as a feminist. This was followed by the Feminist Identity Composite scale and Ryff's Psychological Well-being scale. The data was organised using Microsoft Excel. IBM SPSS software version 28.0.1.1 (14) was used to perform the statistical analyses.

## **4.9 DRAFTING A QUESTIONNAIRE**

### **Demographic Details**

The demographic information of the participants was collected using a brief questionnaire asking for the participant's age, work status, the state in which they currently reside and self-perceived socioeconomic status. An additional question asked if they considered themselves to be feminists.

### **Feminist Identity Composite Scale**

The participant's level of feminist identity development was measured using the Feminist Identity Composite Scale (FIC; Fisher et al., 2000). Formed from the Feminist Identity Scale (FIS; Rickard, 1987) and the Feminist Identity Development Scale (FIDS; Bargad & Hyde, 1991), this scale consists of 33 positively phrased statements that are intended to assess an individual's positive feminist identity. Each question is scored with a 5-point Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree). This scale follows the Downing and Roush model (1965) of feminist identity development, spanning five dimensions or subscales, Passive Acceptance, Revelation, Embeddedness-Emanation, Synthesis and Active Commitment. The subscale scores were calculated by finding means across the items in each subscale and higher means indicated a higher level of agreement with that stage.

The FIC is found to have acceptable Cronbach's alpha scores, primarily when the sample population consisted of white women. Fisher et al. (2000) found alpha coefficients from 0.68-0.84 for the subscales. The internal consistency reliability estimates for each stage of FIC are 0.74 (passive acceptance), 0.75 (revelation), 0.86 (embeddedness-emanation), 0.71 (synthesis) and 0.81 (active commitment) in the mentioned studies. Blue and Berkel (2010), however, conducted a study with African women, resulting in relatively lower reliability. An assessment of the reliability of the FIC in an Indian population is yet to be done.

### **Ryff's Psychological Well-being Scale**

In this study, the short version of Ryff's measure of psychological well-being (1989) was used. This scale contains 18 statements, assessing six dimensions of psychological well-being, which are autonomy (independence and self-control), environmental mastery (capacity to manage environments), personal growth (continuing to grow and develop oneself), positive relations with others (having warm and trusting relationships), purpose in life (a sense of direction and meaning) and self-acceptance (being positive towards oneself). Agreement with each statement was indicated using a 7-point Likert-type scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Reliability measures were found to range from 0.8 to 0.9 on each subscale.

## **4.10 DATA ANALYSIS TECHNIQUES**

### **Correlation**

Correlation is measuring the degree of the relationship between two variables. Karl Pearson's coefficient of correlation otherwise known as simple correlation is the most widely used method. The value  $r$  is known as the correlation coefficient, whose values lie between -1 and +1. All positive values indicate positive correlations, that is, when one variable increases in value, the other also increases, and negative values indicate negative correlations, that is, when the value of one variable increases, the other decreases. If the value of  $r$  is 0, that means there is no correlation. An  $r$  value of -1 is indicative of a perfect negative correlation and a value of +1 is a perfect positive correlation. This means that variations in the independent variable account for

100% of the variations in the dependent variable, and there is a constant change in the dependent variable for each unit change in the independent variable. The closer the  $r$  value is to 1, the higher the correlations.

### **Regression**

A set of statistical procedures to find out trends in data or to estimate relationships between a dependent and a series of other independent variables is called regression analysis. Regression enables researchers to estimate how dependent variables are affected as the independent variable changes. Such descriptions of variables are obtained by fitting a line to the observed data. The method attempts to quantify the strength and character of this relationship. Regression can be mainly of three types linear, multiple linear and nonlinear. The former two are the most commonly used ones and the latter is used when there is a requirement to interpret complicated data sets that have a nonlinear relationship.

#### **Multiple Linear Regression**

This technique is used in estimating the relationship between one dependent variable and two or more independent variables. Here, the dependent variable shows either a linear relationship with the independent variable(s) or a non-linear relationship. The technique allows researchers to understand the relative contribution of each independent variable in the total variance. Researchers make use of this technique when,

- the strength of the relationship between the independent variable(s) and the dependent variable is to be determined
- the values of the dependent variable at specific values of the independent variable are to be found out

### **Reliability**

The consistency of a research study or a psychological tool is called its reliability in psychological research. When a measure produces similar results under consistent conditions it is said to have high reliability. The primary idea, therefore, is to be able to replicate significant

results when the study is performed under similar conditions. The reliability of a psychological tool can be measured or assessed by checking the consistency of results obtained across time, observers and parts of the test itself.

Reliability may be any of the following types,

- Test-retest Reliability – The consistency is checked in the results of the same population across time.
- Inter-rater (Inter-observer or Inter-coder) Reliability – The consistency is checked when raters or observers administering the test are different.
- Internal Consistency – The consistency is checked within different parts of the test.

Cronbach's alpha

Developed by Lee Cronbach in 1951, Cronbach's alpha or coefficient alpha measures the internal consistency or the reliability of a scale. It can be thought of as the mean of all possible split-half coefficients, corrected by the Spearman-Brown formula. In this method, assessment of reliability is done by comparing the covariance among the items of an instrument with the overall variance. Relative to this overall variance, a reliable instrument or test would have a great deal of covariance. High covariance often indicates a strong relationship between the variables.

**Chi-Square Test**

A Chi-square ( $\chi^2$ ) test is used to assess if there is a significant relationship between the two nominal variables. A set of observed values are compared against the expected values to ascertain the existing relationship between them. The two commonly used  $\chi^2$  tests are the  $\chi^2$  goodness of fit test and the  $\chi^2$  test of independence. For a  $\chi^2$  test of independence, a p-value that is less than or equal to the significance level indicates there is sufficient evidence to conclude that the observed distribution is not the same as the expected distribution.

**CHAPTER V**  
**DATA ANALYSIS**

## **DATA ANALYSIS**

### **Correlation**

The relationship between the main variables feminist identity measured using the FIC and psychological well-being was found by using Pearson's Correlation in IBM SPSS software. Pearson's correlation is the most commonly used measure of correlation in statistics. Also called Pearson's Product Moment Correlation, this allows researchers to find out whether there exists a linear relationship between two numerical data sets. Here, the subscale and total scores of FIC were correlated with the subscale and total scores of Ryff's Psychological Well-being to understand whether there exists a relationship between the variables. Correlation was also used to evaluate the relationship between age with Feminist Identity scores and age and well-being scores.

### **Reliability**

The reliability of the FIC scale was also tested in Indian women using Cronbach's alpha on SPSS. This measure assesses internal consistency or the consistency within different parts of the test. This measure is most commonly employed to check the reliability of the questionnaire when there are multiple Likert items in it. The FIC is scored on a 5-point Likert scale with responses ranging from Strongly Disagree to Strongly Agree.

### **Regression**

Multiple linear regression, an extension of simple linear regression, is used to estimate the relationship between one dependent variable and two or more independent variables. Using the SPSS software this tool was used to identify the relationship between the subscales of Feminist Identity (independent variable) with overall well-being and with each subscale of Ryff's Psychological Well-being test.

**Chi-square Test**

This test evaluates the association between two categorical variables. The relationships between the subscales and total scores of the FIC and whether the participants considered themselves a feminist were examined. The subscale and total values were classified into Low, Average and High categories based on the quartile values—values up to the first quartile fell in ‘Low’, values up to the third quartile fell in ‘Average’ and values above this fell in ‘High’. The relationships between these were found using chi-square tests on IBM SPSS.

**CHAPTER VI**  
**FINDINGS**

## FINDINGS

Prior to the data analysis, exclusion criteria were applied and five responses were dropped due to evident response bias and failure to meet the sampling criteria, as mentioned in the previous chapter. Out of 345 responses, 340 were valid and used in the data analysis. Of the 340 participants, based on work status 47.1% were students, 10% were not employed, and 42.9% were employed. 53.53% fall between the ages of 18 and 25, 23% fall between the ages of 25 and 40, and 23.3% are 40 and above. 85.59% of the participants are residents of Kerala, 2.64% Tamil Nadu, 2.35% Karnataka, 1.76% Maharashtra, 1.17% Abroad and the remaining from other states in India. The participants were asked to describe their socioeconomic status and 73.2% are upper-middle class, 26.2% lower-middle class, and 0.6% high.

The participants were considerably varied in whether they considered themselves feminists, with 45% replying with 'Yes' to the question 'Are you a feminist?', 28.8% replying with 'Maybe' and 26.2% replying with 'No.' 95.5% of the participants who responded with 'No' to this question had higher scores in the Embeddedness, Synthesis and Active Commitment subscales, which may imply a misconception or prejudice associated with the term 'feminist', leading to some reluctance to identify with it.

It was found that the majority of the participants fell in the Synthesis stage, coming up to 42.65%. 1.18% of participants belonged to the Passive Acceptance stage, 4.71% fell in Revelation, 25.29% in Embeddedness and 26.18% in Active Commitment. If participants had similar scores in more than one subscale, as a general rule, the higher one was chosen. 77.2% of participants in the Synthesis stage belong to the upper-middle class or high socioeconomic class, but there is no significant relationship between the stages and the socioeconomic status.

The average feminist identity development score was obtained from the mean total scores of the same, as 3.53 (SD=0.36) of a maximum possible 5. On a 5-point response scale, this

indicates a neutral level of feminist identity. The average psychological well-being score was obtained from the mean total scores in a similar procedure, as approximately 5 (SD=0.75) of a maximum possible 7. On a 7-point response scale, this indicates a moderately high level of well-being.

The mean subscale scores and mean total scores of the Feminist Identity Composite Scale (FIC) and Ryff's Psychological Well-being Scale, as well as their respective standard deviations, are listed in Tables 1.1 and 1.2.

**Table 1.1**

*Means and standard deviations of FIC scores*

<b>Scores</b>	<b>Mean</b>	<b>Standard Deviation</b>
Overall FIC	3.53	0.37
Passive Acceptance	2.36	0.73
Revelation	3.37	0.74
Embeddedness	3.78	0.86
Synthesis	4.19	0.49
Active Commitment	4.08	0.55

**Table 1.2**

*Means and standard deviations of Well-being scores*

<b>Scores</b>	<b>Mean</b>	<b>Standard Deviation</b>
Overall Well-being	4.99	0.75
Autonomy	4.89	0.94
Environmental Mastery	4.85	1.04
Personal Growth	5.65	1.00
Positive Relations	4.89	1.32
Purpose in Life	4.71	1.14
Self-Acceptance	5.24	1.21

As observed from the table, participants scored relatively similarly on the six subscales of Ryff's Psychological Well-being scale.

*Relationship between subscales of the Feminist Identity Composite Scale and subscales of Ryff's Scales of Psychological Well-being*

The relationships between the main variables of interest—feminist identity development and psychological well-being were examined. Using Pearson's correlation, it was found that the subscales and total scores of the FIC are significantly correlated with the subscales and total scores of Ryff's psychological well-being scale. The following table displays these values.

**Table 2**  
*Correlations between FIC scores and Well-being scores*

	PA	R	EE	S	AC	FIC	AU TO	EN V	PG	PR	PL	SA	WB
<b>PA</b>	1	-0.108*	-0.108*	-0.170*	-0.235*	0.214**	-0.176*	0.086	-0.267*	-0.090	-0.199*	-0.038	-0.141*
		0.047	0.046	*0.002	*<0.001	<0.001	*0.001	0.119	*<0.001	0.101	*<0.001	0.485	*0.009
<b>R</b>		1	0.376**	0.221**	0.343**	0.724**	-0.062	-0.269*	-0.007	-0.240*	-0.132*	-0.234*	-0.252*
			<0.001	<0.001	<0.001	<0.001	0.254	*<0.001	0.895	*<0.001	0.015	*<0.001	*<0.001
<b>EE</b>			1	0.220**	0.325**	0.589**	-0.010	-0.163*	0.053	-0.058	-0.005	-0.024	-0.067
				<0.001	<0.001	<0.001	0.860	*0.003	0.333	0.290	0.921	0.659	0.218
<b>S</b>				1	0.697**	0.605**	0.162**	-0.106	0.255**	0.103	0.222**	0.195**	0.260**
					<0.001	<0.001	0.003	0.053	<0.001	0.060	<0.001	<0.001	<0.001
<b>AC</b>					1	0.678**	0.142**	0.063	0.220**	0.051	0.212**	0.193**	0.201**
						<0.001	0.009	0.242	<0.001	0.350	<0.001	<0.001	<0.001

FEMINIST IDENTITY AND PSYCHOLOGICAL WELL-BEING

<b>FIC</b>	1	-0.0 16	-0.0 91	0.04 1	-0.1 27*	-0.0 20	-0.0 20	-0.0 65
		0.77 1	0.09 9	0.44 9	0.02 0	0.71 9	0.71 5	0.23 5
<b>AU TO</b>	1	0.25 6**	0.22 2**	0.14 9**	0.00 5	0.22 8**	0.45 6**	
		<0.0 01	<0.0 01	0.00 7	0.93 0	<0.0 01	<0.0 01	
<b>EN V</b>			1	0.30 1**	0.34 2**	0.11 4*	0.47 6**	0.64 7**
				<0.0 01	<0.0 01	0.03 8	<0.0 01	<0.0 01
<b>PG</b>				1	0.34 7**	0.36 9**	0.33 3**	0.62 5**
					<0.0 01	<0.0 01	<0.0 01	<0.0 01
<b>PR</b>					1	0.43 6**	0.41 5**	0.72 0**
						<0.0 01	<0.0 01	<0.0 01
<b>PL</b>						1	0.25 0**	0.53 8**
							<0.0 01	<0.0 01
<b>SA</b>							1	0.74 6**
								<0.0 01
<b>WB</b>								1

\*\* Correlation significant at 0.01 level (2-tailed).

\* Correlation significant at 0.05 level (2-tailed).

From Table 2, it is evident that Passive Acceptance of the FIC is significantly negatively associated with other FIC subscales, as well as with Autonomy, Personal Growth, Purpose in

Life and overall Well-being. It is significantly positively correlated to total FIC scores. Other relationships are not significant.

Revelation subscale of FIC is significantly positively associated with Embeddedness, Synthesis, Active Commitment and total FIC scores; and negatively associated with Environmental Mastery, Positive Relations, Purpose in Life, Self-Acceptance and overall Well-being. Other relationships are negative but not significant.

Embeddedness-Emanation subscale of FIC is significantly positively associated with Synthesis, Active Commitment and total FIC scores, and negatively associated with Environmental Mastery. Other relationships are mainly negative but not significant. It has no significant association with overall Well-being.

Synthesis subscale of FIC is significantly positively associated with Active Commitment and total FIC scores, as well as with Autonomy, Personal Growth, Purpose in Life, Self-Acceptance and overall Well-being. Other relationships are positive but not significant.

Active Commitment subscale of FIC is significantly positively associated with total FIC scores, as well as with Autonomy, Personal Growth, Purpose in Life, Self-Acceptance and overall Well-being. Other relationships are positive but not significant.

The subscales of Ryff's psychological well-being are all significantly positively associated with each other and with overall well-being, apart from Autonomy with Purpose in Life.

In short, Passive Acceptance and Revelation subscales are negatively associated with well-being, Synthesis and Active Commitment are positively associated with well-being and Embeddedness has no significant relationship with well-being. Before conducting the regression analysis to understand the effect of the subscales of FIC on the subscales and overall well-being, the skewness of the dependent variables was found, as depicted in the following table.

**Table 3***Skewness statistics of the dependent variables for regression analysis*

<b>Dependent Variables</b>	<b>Skewness Statistic</b>
Autonomy	-0.501
Environmental Mastery	-0.498
Personal Growth	-0.697
Positive Relations	-0.172
Purpose in Life	-0.094
Self-Acceptance	-0.624
Overall Well-being	-0.209

The variables are slightly negatively skewed. The regression model in the study uses the overall well-being score as well as the six subscale scores as the dependent variables, and the five subscales of the FIC as the independent variables.

The following table shows a regression analysis with overall well-being as the dependent variable and the five subscales of the FIC as the independent variables.

**Table 4.1***Regression analysis with overall well-being scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	4.322	0.395		10.955	<0.001
<b>PA</b>	-0.113	0.052	-0.109	-2.163	0.031
<b>R</b>	-0.353	0.056	-0.348	-6.358	<0.001
<b>EE</b>	-0.041	0.048	-0.047	-0.867	0.387
<b>S</b>	0.336	0.105	0.218	3.186	0.002
<b>AC</b>	0.216	0.100	0.158	2.163	0.031

$R^2=0.195$ ,  $F=16.15$ ,  $p \text{ value}<0.001$

In this regression,  $R^2$  is found to be 0.195, indicating that 19.5% of the variation in total well-being can be explained by the subscales of the FIC. The regression is significant at 1% level of significance ( $F=16.15$ ,  $p\text{-value}<0.001$ ). The B coefficients show that Passive Acceptance,

Revelation and Embeddedness are negatively related to well-being, and Synthesis and Active Commitment are positively related. The beta values show that the highest negatively related variable is Revelation (-0.348), and the highest positive one is Synthesis (0.218). The significance or p-values show that Passive Acceptance, Revelation, Synthesis and Active Commitment significantly affect overall well-being.

When adding age as an independent variable, the  $R^2$  value increases to 0.226, that is, 22.6% of the variation in total well-being can be explained by the subscales of the FIC and age. The regression shows that age positively influences well-being.

Similar regression analyses are conducted for each subscale of Ryff's psychological well-being. The following table shows a regression analysis with Autonomy as the dependent variable and the FIC subscales as the independent variables.

**Table 4.2**

*Regression analysis with Autonomy scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	4.349	0.580		7.498	<0.001
<b>PA</b>	-0.182	0.077	-0.129	-2.365	0.019
<b>R</b>	-0.174	0.082	-0.127	-2.134	0.034
<b>EE</b>	-0.020	0.070	-0.017	-0.283	0.777
<b>S</b>	0.278	0.155	-0.133	1.797	0.073
<b>AC</b>	0.104	0.147	0.056	0.708	0.480

$R^2=0.057$ ,  $F=4.055$ ,  $p\text{ value}=0.001$

Here, the  $R^2$  value is 0.057, which means that 5.7% of the variation in Autonomy is explained by the subscales of the FIC. The independent variables of Passive Acceptance and Revelation are significantly negatively related to Autonomy, while the Synthesis and Active Commitment subscales are positively related but not significant.

For the Environmental Mastery subscale, a regression analysis was done with independent variables as the FIC subscales, as shown in the following table.

**Table 4.3**

*Regression analysis with Environmental Mastery scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	4.131	0.641		6.445	<0.001
<b>PA</b>	0.176	0.085	0.110	2.074	0.039
<b>R</b>	-0.394	0.090	-0.252	-4.365	<0.001
<b>EE</b>	-0.160	0.077	-0.119	-2.071	0.039
<b>S</b>	0.268	0.171	0.113	1.566	0.118
<b>AC</b>	0.252	0.162	0.119	1.554	0.121

$R^2=0.108$ ,  $F=8.090$ ,  $p \text{ value}<0.001$

The  $R^2$  value is 0.108, which means that 10.8% of the variation in Environmental Mastery is explained by the subscales of the FIC. Here, the Passive Acceptance, Revelation and Embeddedness subscales are significantly related to Environmental Mastery; Passive Acceptance positively and the others negatively. Synthesis and Active Commitment are positively related, but not significant.

For the Personal Growth subscale, a regression analysis was done with independent variables as the FIC subscales, as shown in the following table.

**Table 4.4**

*Regression analysis with Personal Growth scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	4.617	0.565		8.177	<0.001
<b>PA</b>	-0.324	0.075	-0.228	-4.320	<0.001
<b>R</b>	-0.119	0.079	-0.086	-1.502	0.134
<b>EE</b>	0.015	0.068	0.013	0.227	0.821
<b>S</b>	0.401	0.151	0.190	2.658	0.008
<b>AC</b>	0.109	0.143	0.058	0.764	0.446

$R^2=0.122$ ,  $F=9.283$ ,  $p \text{ value}<0.001$

The  $R^2$  value is 0.122, which means that 12.2% of the variation in Personal Growth is explained by the subscales of the FIC. Here, the Passive Acceptance and Synthesis are significantly related. Synthesis is positively related, while Passive Acceptance is negatively related.

For the Positive Relations subscale, a regression analysis was done with independent variables as the FIC subscales, as shown in the following table.

**Table 4.5**

*Regression analysis with Positive Relations scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	5.527	0.760		7.271	<0.001
<b>PA</b>	-0.163	0.101	-0.086	-1.614	0.107
<b>R</b>	-0.551	0.107	-0.299	-5.151	<0.001
<b>EE</b>	-0.026	0.092	-0.016	-0.280	0.780
<b>S</b>	0.320	0.203	0.114	1.575	0.116
<b>AC</b>	0.080	0.192	0.032	0.418	0.676

$R^2=0.096$ ,  $F=7.090$ ,  $p \text{ value}<0.001$

Here, the  $R^2$  value is 0.096, which means that 9.6% of the variations in Positive Relations are explained by the subscales of the FIC. While Passive Acceptance, Revelation and Embeddedness are negatively related and Synthesis and Active Commitment are positively related, only Revelation is significant.

For the Purpose in Life subscale, a regression analysis was done with independent variables as the FIC subscales, as shown in the following table.

**Table 4.6**

*Regression analysis with Purpose in Life scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	3.777	0.630		5.991	<0.001
<b>PA</b>	-0.242	0.084	-0.153	-2.895	0.004
<b>R</b>	-0.344	0.089	-0.222	-3.877	<0.001
<b>EE</b>	-0.018	0.076	-0.014	-0.239	0.811
<b>S</b>	0.324	0.168	0.138	1.925	0.055
<b>AC</b>	0.334	0.160	0.160	2.094	0.037

$R^2=0.121$ ,  $F=9.200$ ,  $p \text{ value}<0.001$

Here, the  $R^2$  value is 0.121, which means that 12.1% of the variations in Purpose in Life are explained by the subscales of the FIC. Both Passive Acceptance and Revelation are significantly negatively related to this subscale, while Active Commitment is significantly positively related.

For the Self-Acceptance subscale, a regression analysis was done with independent variables as the FIC subscales, as shown in the following table.

**Table 4.7**

*Regression analysis with Self-Acceptance scores as the dependent variable*

<b>Variable</b>	<b>B coefficient</b>	<b>Std. Error</b>	<b>Beta coefficient</b>	<b>t</b>	<b>p value</b>
<b>(Constant)</b>	3.535	0.700		5.050	<0.001
<b>PA</b>	0.055	0.093	0.031	0.590	0.556
<b>R</b>	-0.536	0.099	-0.309	-5.442	<0.001
<b>EE</b>	-0.039	0.084	-0.026	-0.464	0.643
<b>S</b>	0.423	0.187	0.161	2.263	0.024
<b>AC</b>	0.416	0.177	0.178	2.349	0.019

$R^2=0.134$ ,  $F=10.341$ ,  $p \text{ value}<0.001$

Here, the  $R^2$  value is 0.134, which means that 13.4% of the variations in Self-Acceptance are explained by the subscales of the FIC. Here, Revelation is significantly negatively related to Self-Acceptance, while both Synthesis and Active Commitment are positively and significantly related.

In general, the regression analysis shows that Synthesis, Active Commitment and age are significantly positively related to well-being and its subscales, while Passive Acceptance, Revelation and Embeddedness are rather negatively associated with well-being and its subscales.

From the correlation and regression analyses, it is evident that the first hypothesis  $H_0$  has been rejected, so there is a significant relationship between the subscales of the Feminist Composite Identity scale and Ryff's Psychological Well-being scale.

#### *Reliability of Feminist Identity Composite Scale*

Reliability analyses have been done using Cronbach's alpha, a reliability measure for which if values are above 0.6, the reliability is considered good and 0.4-0.6 are average. As the focus of the current study and the second hypothesis is regarding the reliability of the Feminist Identity Composite, the Cronbach's alpha values for the FIC and subscales are shown in the table below. The reliability of Ryff's Psychological Well-being measure has also been found.

**Table 5**

*Cronbach's alpha values for the FIC and subscales*

<b>Scale/Subscale</b>	<b>Cronbach's Alpha</b>
Feminist Identity Composite scale	0.82
Passive Acceptance subscale	0.77
Revelation subscale	0.84
Embeddedness-Emanation subscale	0.91
Synthesis subscale	0.80
Active Commitment subscale	0.83

The reliability of all subscales of the FIC as well as the total scores of FIC are very good. The reliability of Ryff's Psychological Well-being scale is slightly less, but also above average, at 0.76. From this result, the second hypothesis that the Feminist Identity Composite Scale is not reliable for the chosen sample population of Indian women can be rejected.

*Relationship between age, Feminist Identity and Psychological Well-being*

In the next section of data analysis, correlations between any demographic variables (age, socioeconomic status and self-identified feminism) and main variables of interest were investigated.

Of these, age was slightly negatively correlated with the mean total FIC score at a 0.01 significance level, as well as with Revelation and Embeddedness-Emanation subscales; and slightly positively correlated with Passive Acceptance subscale. There were no significant correlations between age and Synthesis and Active Commitment subscales. Age was slightly positively correlated with the mean total psychological well-being score at 0.01 significance level, as well as with Environmental Mastery, Positive Relations and Self-Acceptance subscales. There were no significant correlations with Autonomy, Personal Growth and Purpose in Life subscales. Pearson's correlation was used. These correlations are given in Tables 6.1 and 6.2

**Table 6.1**

*Correlations between age and FIC scores*

	Age	PA	R	EE	S	AC	Overall FIC
<b>Age</b>	1	0.213** <0.001	-0.318** <0.001	-0.328** <0.001	-0.038 0.480	-0.070 0.199	-0.193** <0.001
<b>PA</b>		1	-0.108* 0.047	-0.108* 0.046	-0.170** 0.002	-0.235** <0.001	0.214** <0.001
<b>R</b>			1	0.376** <0.001	0.221** <0.001	0.343** <0.001	0.724** <0.001

<b>EE</b>	1	0.220** <0.001	0.325** <0.001	0.589** <0.001
<b>S</b>		1	0.697** <0.001	0.605** <0.001
<b>AC</b>			1	0.678** <0.001

\*\* Correlation significant at 0.01 level (2-tailed).

\* Correlation significant at 0.05 level (2-tailed).

**Table 6.2**

*Correlations between age and Well-being scores*

	<b>Age</b>	<b>AUTO</b>	<b>ENV</b>	<b>PG</b>	<b>PR</b>	<b>PL</b>	<b>SA</b>	<b>Overall Well-being</b>
<b>Age</b>	1	0.106 0.051	-0.289* * <0.001	-0.005 0.932	0.226** <0.001	0.103 0.059	0.178** 0.001	0.243** <0.001
<b>AUTO</b>		1	0.343** <0.001	0.262** <0.001	0.109* 0.045	-0.035 0.523	0.310** <0.001	0.487** <0.001
<b>ENV</b>			1	0.282** <0.001	0.343** <0.001	0.086 0.111	0.533** <0.001	0.677** <0.001
<b>PG</b>				1	0.286** <0.001	0.340* * <0.001	0.331** <0.001	0.628** <0.001
<b>PR</b>					1	0.418* * <0.001	0.438** <0.001	0.713** <0.001
<b>PL</b>						1	0.245** <0.001	0.544** <0.001
<b>SA</b>							1	0.763** <0.001

\*\* Correlation significant at 0.01 level (2-tailed).

\* Correlation significant at 0.05 level (2-tailed).

*Association between identifying as a Feminist and Psychological Well-being.*

All five subscales and the total FIC scores were significantly associated at 0.01 level, indicating that self-identified feminists had higher scores in the FIC.

The relationships between the subscales and total scores of psychological well-being, and whether participants considered themselves to be feminists were also found, as per the secondary hypothesis. The subscales of Autonomy, Environmental Mastery and Purpose in Life were significantly associated with whether they identified as feminists, but Personal Growth, Positive

Relations and Self-Acceptance were not significant. Overall well-being was also not significantly associated, indicating that self-identified feminists did not necessarily have higher well-being. The following tables show these associations.

**Table 7.1***Associations between Autonomy subscale scores and identifying as a feminist*

		Autonomy Scores			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	36 23.4%	81 52.6%	37 24.0%	154 100%
	Maybe	42 43.3%	42 43.3%	13 13.4%	97 100%
	No	27 30.3%	43 48.3%	19 21.3%	89 100%
Total		105 30.9%	166 48.8%	69 20.3%	340 100%

Chi-square value=12.1, df=4, p-value=0.017

**Table 7.2***Associations between Environmental Mastery subscale scores and identifying as a feminist*

		<b>Environmental Mastery Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	54 35.1%	72 46.8%	28 18.2%	154 100%
	Maybe	23 23.7%	53 54.6%	21 21.6%	97 100%
	No	10 11.2%	47 52.8%	32 36%	89 100%
Total		87 25.6%	172 50.6%	81 23.8%	340 100%

Chi-square value=21.3, df=4, p-value&lt;0.001

**Table 7.3***Associations between Personal Growth subscale scores and identifying as a feminist*

		<b>Personal Growth Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	37 24.0%	86 55.8%	31 20.1%	154 100%
	Maybe	35 36.1%	50 51.5%	12 12.4%	97 100%
	No	33 37.1%	46 51.7%	10 11.2%	89 100%
Total		105 30.9%	182 53.5%	53 15.6%	340 100%

Chi-square value=8.34, df=4, p-value=0.080

**Table 7.4***Associations between Positive Relations subscale scores and identifying as a feminist*

		<b>Positive Relations Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	40 26%	87 56.5%	27 17.5%	154 100%
	Maybe	30 30.9%	47 48.5%	20 20.6%	97 100%
	No	20 22.5%	44 49.4%	25 28.1%	89 100%
Total		90 26.5%	178 52.4%	72 21.2%	340 100%

Chi-square value=5.20, df=4, p-value=0.267

**Table 7.5***Associations between Purpose in Life subscale scores and identifying as a feminist*

		<b>Purpose in Life Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	43 27.9%	88 57.1%	23 14.9%	154 100%
	Maybe	45 46.4%	35 36.1%	17 17.5%	97 100%
	No	32 36%	42 47.2%	15 16.9%	89 100%
Total		120 35.3%	165 48.5%	55 16.2%	340 100%

Chi-square value=11.53, df=4, p-value=0.021

**Table 7.6***Associations between Self-Acceptance subscale scores and identifying as a feminist*

		<b>Self-Acceptance Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	44 28.6%	85 55.2%	25 16.2%	154 100%
	Maybe	32 33%	52 53.6%	13 13.4%	97 100%
	No	16 18%	49 55.1%	24 27%	89 100%
Total		92 27.1%	186 54.7%	62 18.2%	340 100%

Chi-square value=9.44, df=4, p-value=0.051

**Table 7.7***Associations between overall well-being scores and identifying as a feminist*

		<b>Overall Well-being Scores</b>			
		Low	Average	High	Total
<b>Are you a feminist?</b>	Yes	38 24.7%	76 49.4%	40 26%	154 100%
	Maybe	30 30.9%	47 48.5%	20 20.6%	97 100%
	No	21 23.6%	43 48.3%	25 28.1%	89 100%
Total		89 26.2%	166 48.8%	85 25%	340 100%

Chi-square value=2.355, df=4, p-value=0.671

Socioeconomic status responses were divided into lower-middle class and upper-middle class/high class and compared against the mean subscales of the FIC as well as with the mean total, in a similar process. The subscales of Passive Acceptance, Revelation, Synthesis and Active Commitment, as well as total FIC scores, were not found to have a significant association with socioeconomic status; however, the Embeddedness subscale was found to be significant at a 0.05 significance level. Overall, there is no significant relationship.

The mean subscales and total scores of psychological well-being were classified similarly into Low, Average and High. The subscales of Autonomy, Environmental Mastery, Personal Growth, Purpose in Life and Self-Acceptance were not found to be significantly associated with socioeconomic status. Positive Relations and overall well-being were related to socioeconomic status at a 0.05 significance level.

## **Discussion**

Feminism is a range of socio-political ideologies and movements aiming to define and establish the equality of sexes on all grounds. The origin of the idea dates back to 3rd century BCE Rome, where women blocked entrances to the Capitoline Hill to express their disapproval of the laws limiting them to use expensive goods. From what might have started as a disagreement that was discussed at tea among the womenfolk, feminism has since grown into a wave of movements that have questioned and continues to question ethos across the globe. However, in this male-dominated world, the struggle for equality still continues.

The feminist label has long been viewed by the traditional communities as misleading. Collectivistic communities, especially ones like India, consider feminism to be something 'against their culture'. Among the dissenters, some people expect women to stick to traditional gender roles, while others claim to be advocates for women's rights but are hesitant to label themselves as feminists. Perhaps owing to these conditions, as much as a quarter of the sample identified as non-feminists. Ironically, nearly all the sceptics had scores that were suggestive of feminist identity. This indicates that more than being unsupportive of the cause of feminism, it is the discredited label that seems to be the issue. On the other end, self-identified feminists also

had higher scores in the FIC, indicating that while some women are influenced by society's take on the feminist label, there exist other women who understand the true reason for what the movement stands for.

Most women fell in the Synthesis stage of Feminist Identity. This stage is the fourth stage that involves the integration of positive feminine qualities with one's own personality to create a realistic self-concept that surpasses gender roles and the capability of evaluating men individually rather than with collective stereotypes. This suggests that most women in the sample are aware of the defects in the system and they are able to direct their energy to respond appropriately to discrimination. This is the highest level to which most women progress. They fight their own battles but seldom do they commit to bringing change for the larger good.

Reliability analyses done for the FIC scale using Cronbach's alpha indicated that the scale is reliable for the sample of women chosen. Suggesting that FIC is a reliable scale to measure the feminist identity of Indian women, these findings can aid future research. However, an analysis of reliability and validity on a larger and more diverse sample can support this finding.

Analyses showed that psychological well-being and feminist identity development are significantly correlated. More importantly, the Passive Acceptance and Revelation subscales are negatively related to well-being. This suggests that the internalisation of and conformity to misogynistic ideas influence a woman's mental health significantly. The inequality and injustice that women face in a male-dominated world give rise to discomfort and stress, which may even escape the notice of the oppressed individual, and reduce women's psychological well-being. However, once these individuals are able to identify the root of their distress, they move a step ahead in their feminist identity but this does not improve their mental well-being yet. Perhaps, recognizing all the ways in which they are disadvantaged and discriminated against adds to the existing factors negatively impacting the women's psychological health. Unsurprisingly, the subscales Synthesis and Active Commitment are positively associated with well-being. As the final stages in the development of feminist identity, the positive association of these subscales

with overall well-being indicate that once women reach these levels in their lives, they are bound to experience a better mental health status. The inclusion of stereotypically feminine values or qualities into one's personality, not as an internalisation of social norms, but as a conscious choice of self-acceptance leads women to have a better self-concept. This integration allows the individuals to look beyond gender roles and enable them to evaluate their counterparts as individuals and not in the light of stereotypes. Being able to set aside stereotypic appraisals of people, women with fully developed feminist identities also feel the urge to make the world a better place for everyone. These women have the resources and ideas to act in congruence with their beliefs. Being an instrument of change has many positive effects on the mental well-being of a person; making even the slightest of impacts on the world can be greatly satisfying for a woman who is operating to change the world for the better. Regression analysis yields that Passive Acceptance, Revelation, Synthesis and Active Commitment significantly affect overall well-being, where Revelation is the highest negatively related and Synthesis is the highest positively related.

For this chosen sample, findings revealed that age and Passive Acceptance are slightly positively correlated and the same is slightly negatively correlated with Revelation and Embeddedness-Emanation subscales. This might indicate the effect of social conditioning throughout women's lives. As a woman grows older, perhaps the initial motivation to not conform to the system decreases due to the numerous forms of discrimination and harassment she faces. This might make women regress to a state of Passive Acceptance, where they are in denial of the injustices that are happening to them. They assume traditional gender roles, internalise these values, and believe that "this is how the system is supposed to be". Such beliefs might help them cope with the feelings of being unfairly treated. Age was also seen to be slightly positively correlated with overall psychological well-being. Assumption of traditional gender roles might bring forth much applause and acceptance from people around them which might act as a reinforcing factor for such beliefs. The association is seen only in subscales Environmental Mastery, Positive Relations and Self-Acceptance. Correlations with Autonomy, Personal Growth and Purpose in Life subscales showed no significant relations. In the same breath, higher scores

of feminist identity, Synthesis and Active Commitment, were obtained by women of all ages in the sample indicating the influence of age on being a feminist is only as strong as any other factor.

While the socioeconomic status of the women did not have a significant relationship with the overall development of feminist identity, it was considerably associated with the Embeddedness subscale. This connection can be attributed to the fact that the lack of financial resources and social mobility in lower socioeconomic classes limits the creative exploration that is characteristic of this dimension. Women in this stratum of society may adhere to the patriarchal values because, in their family, a male has a better chance at education and thus the advantage of being an earning member. Women with higher socioeconomic status have the resources to support themselves without relying on others and this autonomy allows these women to accept themselves for who they are and have healthier relationships with those around them. This relation is proved by the association of the Positive Relations and Self-Acceptance subscales and overall well-being scores with socioeconomic class.

Considering the relation between social factors and well-being, claiming to hold feminist ideals and expressing them is bound to have an influence on the psychological well-being of an individual. There were correlations between the subscales of Autonomy, Environmental Mastery and Purpose in Life with self-identification as a feminist. Women who readily exert their independence and function to meet their own principles rather than seek approval from others experience a freedom of choice that improves the state of their mental health. This freedom of choice also enables women to create an environment for themselves that suits their own needs without conforming to the collective misogynistic ideals. Perhaps, as a result of this, women who identify as feminists have a direction in life that is not dictated by others and display a brand of maturity that seems to advance themselves and the people around them.

**CHAPTER VII**  
**RECOMMENDATIONS**

### **RECOMMENDATIONS**

Since there is a scope of error when people fill in data on Google Forms, it is recommended to conduct the study offline through surveys in the presence of an administrator. Suggestions also include making use of the 42-item version of Ryff's Psychological Well-being Test as opposed to the 18-item version used in the present study. The sample size can also be increased to include diverse individuals with respect to age, socioeconomic class, education and region. The reliability of FIC may be confirmed among Indian women by increasing the sample size and the regional and socioeconomic variety in the sample.

**CHAPTER VIII**  
**CONCLUSION**

## **CONCLUSION**

The present study aimed to find a relationship between Feminist Identity and Psychological Well-being in a sample of Indian women. Women who believe in abiding by traditional gender roles have less psychological well-being than women who are on the higher levels of feminist identity. The Feminist Identity Composite Scale was also found to be reliable in the chosen sample of Indian women. The data also revealed that older women in the chosen sample are more likely to be in the Passive Acceptance stage. It was also found that self-identified feminists had higher scores in the FIC. Data indicates that self-identified feminists did not necessarily have higher well-being

**REFERENCES**

- (N.d.-a). Yale.Edu. Retrieved April 30, 2022, from  
<http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm>
- (N.d.-b). Egyankosh.Ac.In. Retrieved April 30, 2022, from  
<https://egyankosh.ac.in/bitstream/123456789/20900/1/Unit-2.pdf>
- A Refresher on Regression Analysis*. (2015, November 4). Harvard Business Review.  
<https://hbr.org/2015/11/a-refresher-on-regression-analysis>
- Abbott, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., & Croudace, T. J. (2010). An evaluation of the precision of measurement of Ryff's Psychological Well-being scales in a population sample. *Social Indicators Research*, *97*(3), 357–373.  
<https://doi.org/10.1007/s11205-009-9506-x>
- Anagol, P. (2017). *The emergence of feminism in India, 1850-1920* (1st Edition). Routledge. <https://doi.org/10.4324/9781315240022>
- Bargad, A., & Hyde, J. S. (1991). Women's studies: A study of feminist identity development in women. *Psychology of Women Quarterly*, *15*(2), 181–201.  
<https://doi.org/10.1111/j.1471-6402.1991.tb00791.x>
- Basu, S. (2012). Mental health concerns for Indian women. *Indian Journal of Gender Studies*, *19*(1), 127–136. <https://doi.org/10.1177/097152151101900106>
- Beasley, C. (1999). *What is Feminism?: An Introduction to Feminist Theory*. SAGE Publications.
- Beers, B. (2022, February 8). *Regression Definition*. Investopedia.  
<https://www.investopedia.com/terms/r/regression.asp>

Bevans, R. (2020, February 20). *Multiple linear regression*. Scribbr.

<https://www.scribbr.com/statistics/multiple-linear-regression/>

Burkett, E., & Brunell, L. (2021). feminism. In *Encyclopedia Britannica*.

<https://www.britannica.com/topic/feminism>

Carlyle, M. S. (2017). *Postfeminist distortions: complicated discourses of feminist identity, choice and sexuality*. <https://keep.lib.asu.edu/items/155646>

Colaner, C. W., & Rittenour, C. E. (2015). “Feminism begins at home”: The influence of mother gender socialization on daughter career and motherhood aspirations as channeled through daughter feminist identification. *Communication Quarterly*, 63(1), 81–98. <https://doi.org/10.1080/01463373.2014.965839>

*Correlation Coefficient formula*. (2019, February 28). BYJUS; BYJU’S.

<https://byjus.com/correlation-coefficient-formula/>

*Correlation coefficient: Simple definition, formula, easy steps*. (2021, May 24). Statistics How To.

<https://www.statisticshowto.com/probability-and-statistics/correlation-coefficient-formula/>

*Cronbach’s Alpha in SPSS Statistics - procedure, output and interpretation of the output using a relevant example*. (n.d.). Laerd.Com. Retrieved May 1, 2022, from

<https://statistics.laerd.com/spss-tutorials/cronbachs-alpha-using-spss-statistics.php>

*Cronbach’s alpha: Definition, interpretation, SPSS*. (2021, February 28). Statistics How To.

<https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/cronbachs-alpha-spss/>

- Dash, P. P. (2021). Feminism and its discontents: Punishing sexual violence in India. *Indian Journal of Gender Studies*, 28(1), 7–28.  
<https://doi.org/10.1177/0971521520974843>
- DeBlaere, C., & Bertsch, K. N. (2013). Perceived sexist events and psychological distress of sexual minority women of color: The moderating role of womanism. *Psychology of Women Quarterly*, 37(2), 167–178.  
<https://doi.org/10.1177/0361684312470436>
- DeBlaere, C., Chadwick, C. N., Zelaya, D. G., Bowie, J.-A., Bass, M. F., & Finzi-Smith, Z. (2017). The Feminist Identity Composite: An examination of structural validity with sexual minority women. *Psychology of Women Quarterly*, 41(1), 20–31.  
<https://doi.org/10.1177/0361684316676046>
- Devika, J. (2020). The defence of aachaaram, femininity, and Neo-Savarna power in Kerala. *Indian Journal of Gender Studies*, 27(3), 445–470.  
<https://doi.org/10.1177/0971521520939283>
- Devika, J., & Thampi, B. V. (2007). Between ‘empowerment’ and ‘liberation’: The Kudumbashree initiative in Kerala. *Indian Journal of Gender Studies*, 14(1), 33–60. <https://doi.org/10.1177/097152150601400103>
- Downing, N. E., & Roush, K. L. (1985). From passive acceptance to active commitment: A model of feminist identity development for women. *The Counseling Psychologist*, 13(4), 695–709. <https://doi.org/10.1177/0011000085134013>
- Duncan, L. E., Garcia, R. L., & Teitelman, I. (2021). Assessing politicized gender identity: Validating the Feminist Consciousness Scale for men and women. *The Journal of Social Psychology*, 161(5), 570–592.  
<https://doi.org/10.1080/00224545.2020.1860883>
- Education, & Careers Desk. (2021, September 8). *International literacy day 2021: 10 Indian states, union territories with highest literacy rates*. News18.

<https://www.news18.com/news/education-career/international-literacy-day-2021-10-indian-states-union-territories-with-highest-literacy-rates-4175855.html>

Eisele, H., & Stake, J. (2008). The differential relationship of feminist attitudes and feminist identity to self-efficacy. *Psychology of Women Quarterly*, 32(3), 233–244. <https://doi.org/10.1111/j.1471-6402.2008.00432.x>

Fassinger, R. E. (1994). Development and testing of the Attitudes Toward Feminism and the Women's Movement (FWM) scale. *Psychology of Women Quarterly*, 18(3), 389–402. <https://doi.org/10.1111/j.1471-6402.1994.tb00462.x>

*Feminist Identity Theory Downing and Roush's Model of Positive Feminist Identity Development* Wendy S. (n.d.).

Fischer, A. R., & Good, G. E. (2004). Women's feminist consciousness, anger, and psychological distress. *Journal of Counseling Psychology*, 51(4), 437–446. <https://doi.org/10.1037/0022-0167.51.4.437>

Fischer, A. R., & Holz, K. B. (2007). Perceived discrimination and women's psychological distress: The roles of collective and personal self-esteem. *Journal of Counseling Psychology*, 54(2), 154–164. <https://doi.org/10.1037/0022-0167.54.2.154>

Fischer, A. R., Tokar, D. M., Mergl, M. M., Good, G. E., Hill, M. S., & Blum, S. A. (2000). Assessing women's feminist identity development: Studies of convergent, discriminant, and structural validity. *Psychology of Women Quarterly*, 24(1), 15–29. <https://doi.org/10.1111/j.1471-6402.2000.tb01018.x>

Hayes, A. (2022, January 3). *Multiple Linear Regression (MLR)*. Investopedia. <https://www.investopedia.com/terms/m/mlr.asp>

*How to perform a multiple regression analysis in SPSS Statistics*. (n.d.). Laerd.Com. Retrieved May 1, 2022, from

<https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php>

Huppert, F. A. (2009). Psychological Well-being: Evidence Regarding its Causes and Consequences. *Applied Psychology. Health and Well-Being*, 1(2), 137–164.

<https://doi.org/10.1111/j.1758-0854.2009.01008.x>

Huta, V. (2016). *The Routledge handbook of media use and well-being: International perspectives on theory and research on positive media effects* (L. Reinecke & M. B. Oliver, Eds.). Routledge.

Hyde, J. S. (2002). Feminist identity development: The current state of theory, research, and practice. *The Counseling Psychologist*, 30(1), 105–110.

<https://doi.org/10.1177/0011000002301007>

Jackson, L. A., Fleury, R. E., & Lewandowski, D. A. (1996). Feminism: Definitions, support, and correlates of support among female and male college students. *Sex Roles*, 34–34(9–10), 687–693. <https://doi.org/10.1007/bf01551502>

Kaur, P. (2021). Feminist Identity Development in relation with Emotional Competence of Indian Educated Women Aayushi International Interdisciplinary. *Interdisciplinary Research Journal*, 8.

Kothari, C. R. (2004). *Research methodology: Methods and techniques* (2nd ed.). New Age International.

Liss, M., O'Connor, C., Morosky, E., & Crawford, M. (2001). What makes a feminist? Predictors and correlates of feminist social identity in college women. *Psychology of Women Quarterly*, 25(2), 124–133. <https://doi.org/10.1111/1471-6402.00014>

Liu, Y., & Zheng, Y. (2019). Reliability and validity of Feminist Identity Composite in Chinese women. *Frontiers in Psychology*, 10, 2842.

<https://doi.org/10.3389/fpsyg.2019.02842>

- Marecek, J., & Kravetz, D. (1977). Women and mental health: a review of feminist change efforts. *Psychiatry*, *40*(4), 323–329.  
<https://doi.org/10.1080/00332747.1977.11023945>
- McNAMARA, K., & Rickard, K. M. (1989). Feminist identity development: Implications for feminist therapy with women. *Journal of Counseling and Development: JCD*, *68*(2), 184–189. <https://doi.org/10.1002/j.1556-6676.1989.tb01354.x>
- Moradi, B., & Subich, L. M. (2002a). *Feminist Identity Development Measures: Comparing the Psychometrics of Three Instruments The Counseling Psychologist*. *30*, 66–86.
- Moradi, B., & Subich, L. M. (2002b). Perceived Sexist Events and Feminist Identity Development Attitudes. *Links to Women's Psychological Distress The Counseling Psychologist*, *30*, 44–65.
- Moradi, B., Subich, L. M., & Phillips, J. C. (2002). Revisiting feminist identity development theory, research, and practice. *The Counseling Psychologist*, *30*(1), 6–43. <https://doi.org/10.1177/0011000002301002>
- Offen, K. (1988). Defining feminism: A comparative historical approach. *Signs*, *14*(1), 119–157. <https://doi.org/10.1086/494494>
- Pearson correlation coefficient - quick introduction*. (n.d.). Spss-Tutorials.Com. Retrieved May 1, 2022, from <https://www.spss-tutorials.com/pearson-correlation-coefficient/>
- Purnell, L. M. (n.d.). *I think I am a feminist: A study of feminist identity development of undergraduate college women*. Retrieved April 30, 2022, from [https://core.ac.uk/display/213815143?utm\\_source=pdf&utm\\_medium=banner&utm\\_campaign=pdf-decoration-v1](https://core.ac.uk/display/213815143?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1)

*Regression analysis*. (2020, February 12). Corporate Finance Institute.

<https://corporatefinanceinstitute.com/resources/knowledge/finance/regression-analysis/>

*Regression analysis: Step by step articles, videos, simple definitions*. (2013, September 27). Statistics How To.

<https://www.statisticshowto.com/probability-and-statistics/regression-analysis/>

Rickard, K. M. (1989). The relationship of self-monitored dating behaviors to level of feminist identity on the feminist identity scale. *Sex Roles, 20*(3–4), 213–226.

<https://doi.org/10.1007/bf00287993>

Roy, R. E., Weibust, K. S., & Miller, C. T. (2007). Effects of stereotypes about feminists on feminist self-identification. *Psychology of Women Quarterly, 31*(2), 146–156.

<https://doi.org/10.1111/j.1471-6402.2007.00348.x>

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*(1), 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>

Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*(4), 719–727.

<https://doi.org/10.1037//0022-3514.69.4.719>

Ryff, Carol D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*(6), 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>

<https://doi.org/10.1037/0022-3514.57.6.1069>

Ryff, Carol D. (1995). Psychological well-being in adult life. *Current Directions in Psychological Science, 4*(4), 99–104.

<https://doi.org/10.1111/1467-8721.ep10772395>

- Ryff, Carol D. (2014). Psychological well-being revisited: advances in the science and practice of eudaimonia. *Psychotherapy and Psychosomatics*, 83(1), 10–28. <https://doi.org/10.1159/000353263>
- Saunders, K. J., & Kashubeck-West, S. (2006). The relations among feminist identity development, gender-role orientation, and psychological well-being in women. *Psychology of Women Quarterly*, 30(2), 199–211. <https://doi.org/10.1111/j.1471-6402.2006.00282.x>
- September, A. N., McCarrey, M., Baranowsky, A., Parent, C., & Schindler, D. (2001). The relation between well-being, impostor feelings, and gender role orientation among Canadian university students. *The Journal of Social Psychology*, 141(2), 218–232. <https://doi.org/10.1080/00224540109600548>
- Siegel, J. A., & Calogero, R. M. (2021). Measurement of feminist identity and attitudes over the past half century: A critical review and call for further research. *Sex Roles*, 85(5–6), 248–270. <https://doi.org/10.1007/s11199-020-01219-w>
- Springer, K. W., & Hauser, R. M. (2006). An assessment of the construct validity of Ryff's Scales of Psychological Well-Being: Method, mode, and measurement effects. *Social Science Research*, 35(4), 1080–1102. <https://doi.org/10.1016/j.ssresearch.2005.07.004>
- SPSS chi-square independence test - beginners tutorial*. (n.d.). Spss-Tutorials.Com. Retrieved May 1, 2022, from <https://www.spss-tutorials.com/spss-chi-square-independence-test/>
- Srivastava, K. (2012). Women and mental health: Psychosocial perspective. *Industrial Psychiatry Journal*, 21(1), 1–3. <https://doi.org/10.4103/0972-6748.110938>
- Szymanski, D. M., Gupta, A., Carr, E. R., & Stewart, D. (2009). Internalized misogyny as a moderator of the link between sexist events and women's psychological distress. *Sex Roles*, 61(1–2), 101–109. <https://doi.org/10.1007/s11199-009-9611-y>

*Types of sampling: Sampling methods with examples.* (2018, April 9). QuestionPro.

<https://www.questionpro.com/blog/types-of-sampling-for-social-research/>

Walters, V. (1993). Stress, anxiety and depression: women's accounts of their health problems. *Social Science & Medicine* (1982), 36(4), 393–402.

[https://doi.org/10.1016/0277-9536\(93\)90401-o](https://doi.org/10.1016/0277-9536(93)90401-o)

Yakushko, O. (2007). Do feminist women feel better about their lives? Examining patterns of feminist identity development and women's subjective well-being. *Sex Roles*, 57(3–4), 223–234. <https://doi.org/10.1007/s11199-007-9249-6>

Yoder, J. D., Perry, R. L., & Saal, E. I. (2007). What good is a feminist identity?: Women's feminist identification and role expectations for intimate and sexual relationships. *Sex Roles*, 57(5–6), 365–372.

<https://doi.org/10.1007/s11199-007-9269-2>

Yoder, J. D., Snell, A. F., & Tobias, A. (2012). Balancing multicultural competence with social justice: Feminist beliefs and optimal psychological functioning. *The Counseling Psychologist*, 40(8), 1101–1132.

<https://doi.org/10.1177/0011000011426296>

Yoder, J. D., Tobias, A., & Snell, A. F. (2011). When Declaring “I am a Feminist” Matters: Labeling is Linked to Activism. *Sex Roles*, 64(1–2), 9–18.

<https://doi.org/10.1007/s11199-010-9890-3>

Zucker, A. N. (2004). Disavowing Social Identities: What it Means When Women Say, “I’m not a Feminist, But ....” *Psychology of Women Quarterly*, 28(4), 423–435.

<https://doi.org/10.1111/j.1471-6402.2004.00159.x>

## APPENDIX A

### **Relation Between Feminist Identity And Psychological Well-being**

Hello, ladies!

We are final year students of B.Sc. Psychology at St. Teresa's College (Autonomous), Ernakulam. Through this survey, we aim to collect the data to measure the relation between feminist identity development and psychological well-being in Indian women who are 18 and above.

We appreciate you taking the time to fill out this form. We request you to be honest while answering the questions and to complete it to the best of your knowledge.

Thank you!

### **Informed Consent**

I volunteer to participate in this research project, "Relation Between Feminist Identity And Psychological Well-being". I understand that this project is designed to gather information for non-funded academic research work. I understand that I will not be paid for my participation. I may withdraw and discontinue participation at any time without penalty. I understand that the researchers will not identify me by my name in any reports using information obtained from the study, and that my confidentiality as a participant in the study will remain secure. I also understand that the information obtained from this research will be used for academic purposes only.

By clicking on the 'I Agree' option, I confirm that I have read the above information and give my full consent to participate in this study.

- I agree

## **APPENDIX B**

### **A. Demographic Details**

1. Email ID
2. Name (Initials Only)
3. Age
4. Work Status
  - Employed
  - Not Employed
  - Student
5. State of Current Residence
6. How would you describe your socio-economic status?
  - High
  - Upper Middle Class
  - Lower Middle Class
  - Poor
7. Are you a Feminist?
  - Yes
  - No
  - Maybe

**B. Feminist Identity Composite Scale**

Each item was required to be rated according to the following scale,

- a. Strongly disagree
- b. Disagree
- c. Neither agree nor disagree
- d. Agree
- e. Strongly agree

1. I am very committed to a cause that I believe contributes to a more fair and just world for all people.
2. I want to work to improve women's status.
3. I am willing to make certain sacrifices to effect change in this society in order to create a nonsexist, peaceful place where all people have equal opportunities.
4. It is very satisfying to me to be able to use my talents and skills in my work in the women's movement.
5. I care very deeply about men and women having equal opportunities in all respects.
6. I choose my "causes" carefully to work for greater equality of all people.
7. I feel that I am a very powerful and effective spokesperson for the women's issues I am concerned with right now.
8. On some level, my motivation for almost every activity I engage in is my desire for an egalitarian world.
9. I owe it not only to women but to all people to work for greater opportunity and equality for all.
10. I feel like I have blended my female attributes with my unique personal qualities.
11. I am proud to be a competent woman.
12. I have incorporated what is female and feminine into my own unique personality.
13. I enjoy the pride and self-assurance that comes from being a strong female.

14. As I have grown in my beliefs I have realised that it is more important to value women as individuals than as members of a larger group of women.
15. Gradually, I am beginning to see just how sexist society really is.
16. I feel angry when I think about the way I am treated by men and boys.
17. Men receive many advantages in society and because of this are against equality for women.
18. I never realized until recently that I have experienced oppression and discrimination as a woman in society.
19. I feel like I've been duped into believing society's perceptions of me as a woman.
20. My female friends are like me in that we are all angry at men and the ways we have been treated as women.
21. In my interactions with men, I am always looking for ways I may be discriminated against because I am female.
22. Regretfully, I can see ways in which I have perpetuated sexist attitudes in the past.
23. I am very interested in women writers.
24. I am very interested in women musicians.
25. I am very interested in women artists.
26. I am very interested in women's studies.
27. I don't see much point in questioning the general expectation that men should be masculine and women should be feminine.
28. One thing I especially like about being a woman is that men will offer me their seat on a crowded bus or open doors for me because I am a woman.
29. I like being a traditional female.
30. I think that men and women had it better in the 1950s when married women were housewives and their husbands supported them.
31. If I were married to a man and my husband was offered a job in another state, it would be my obligation to move in support of his career.
32. I think that most women will feel most fulfilled by being a wife and a mother.
33. I think it's lucky that women aren't expected to do some of the more dangerous jobs that men are expected to do, like construction work or race car driving.

**C. Ryff's Psychological Well-being Scale**

Each item was required to be rated according to the following scale,

- a. Strongly agree
- b. Somewhat agree
- c. A little agree
- d. Neither agree nor disagree
- e. A little disagree
- f. Somewhat disagree
- g. Strongly disagree

1. "I like most parts of my personality."
2. "When I look at the story of my life, I am pleased with how things have turned out so far."
3. "Some people wander aimlessly through life, but I am not one of them."
4. "The demands of everyday life often get me down."
5. "In many ways I feel disappointed about my achievements in life."
6. "Maintaining close relationships has been difficult and frustrating for me."
7. "I live life one day at a time and don't really think about the future."
8. "In general, I feel I am in charge of the situation in which I live."
9. "I am good at managing the responsibilities of daily life."
10. "I sometimes feel as if I've done all there is to do in life."
11. "For me, life has been a continuous process of learning, changing, and growth."
12. "I think it is important to have new experiences that challenge how I think about myself and the world."
13. "People would describe me as a giving person, willing to share my time with others."
14. "I gave up trying to make big improvements or changes in my life a long time ago"
15. "I tend to be influenced by people with strong opinions"
16. "I have not experienced many warm and trusting relationships with others."

17. "I have confidence in my own opinions, even if they are different from the way most other people think."

18. "I judge myself by what I think is important, not by the values of what others think is important."

**ST. TERESA'S COLLEGE**  
**(AUTONOMOUS)**  
**AFFILIATED TO MAHATMA GANDHI UNIVERSITY**



**CHILD VACCINATION DRIVE**  
**PROJECT REPORT**

**In partial fulfilment of the requirements for the award of the degree of**

**BACHELOR OF VOCATION IN**  
**SOFTWARE DEVELOPMENT**

**By**

**THYBA VZ**  
**III BVOC SWD**  
**Register No: VB19SWD017**

**Under the guidance of**  
**Mrs DHANYA R**  
**DEPARTMENT OF COMPUTER APPLICATIONS**  
**2021-2022**

# CERTIFICATE



This is to certify that the project report entitled "CHILD VACCINATION DRIVE", a bonafide record of the work done by THYBA VZ during the year 2019-22 and submitted in partial fulfilment of the requirements for the degree of Bachelor of vocation in Software development under Mahatma Gandhi University.

Head of the Department



Internal Examiner

External Examiner

# CERTIFICATE



This is to certify that the project report entitled “CHILD VACCINATION DRIVE “, a bonafide record of the work done by THYBA VZ during the year 2019-22 and submitted in partial fulfilment of the requirements for the degree of Bachelor of vocation in Software development under Mahatma Gandhi University.

**Head of the Department**

**Internal Examiner**

**External Examiner**

April 0

**To Whom So Ever It May Concern**

*This is to certify that, **Thyba V Z**, 6<sup>th</sup> semester **B.Voc Software Deve**  
student of **St.Teresa's College (Autonomous)** has successfully comp  
project entitled "**Child Vaccination Drive**" from our organization.*

-----

*We wish her all the best for her bright future endeavors.*

**DIRECTOR**



**ALAN DANIEL**



# DECLARATION

I.THYBA VZ (Register no: VB19SWD017), Bvoc software development final year student of St. Teresa's College (Autonomous), Ernakulam, hereby declare that the project submitted named a CHILD VACCINATION DRIVE for the Bachelor's Degree in Bvoc software development is my original work. I further declare that the said work has not previously been submitted to any other university or academic body.

**Place : Ernakulam**

**THYBA VZ**

**Date :**

# ACKNOWLEDGEMENT

In this humble endeavor I have received a great deal of support and guidance from different quarters. First and foremost I thank the God almighty, for bestowing upon me abundance of grace, wisdom and power throughout the study and making it a success.

I am extremely grateful to my Director Rev. Dr. Sr. Vinitha (CSST) and principle Dr. Lizzy Mathew for giving me this opportunity.

I take this opportunity to express my profound sense of gratitude and indebtedness to the Head of the Department of Computer Applications as well as my project guide.

Mrs DHANYA R. for the encouragement and guidance in accomplishing the work.

I extend my gratitude to Mr. ALAN DANIEL, faculty at Explorics solutions, Ernakulam under whose guidance and support this project work has been carried out.

Last but not the least, I would like to thank my parents and friends for motivating me and providing me the right environment for making this project work a great success.

**THYBA VZ**

## **SYNOPSIS**

The CHILD VACCINATION DRIVE is a software application which avoids more manual hours that need to spend in record keeping and generating reports.

Parents are constantly concerned about the health and safety of their children. Therefore, they take many steps in order to prevent their children from catching a disease. One of the options is vaccination. Vaccine works to protect infants, children and even adults from illnesses and death caused by many infectious diseases. Vaccination has its own time, period and schedule. The system being developed is web based application that helps the parents to know the vaccination details that is to be given to the child, its importance and timings. Parents can also check the hospitals that provide the vaccinations for each age of child and can also know the date and timing. The system provides complete information regarding the vaccination.



## Contents

<b>1.INTRODUCTION.....</b>	<b>3</b>
<b>1.1 About Project.....</b>	<b>3</b>
<b>1.2 About Organization.....</b>	<b>3</b>
<b>1.3 Our Services .....</b>	<b>4</b>
<b>1.4 Objectives of the project.....</b>	<b>4</b>
<b>2.SYSTEM ANALYSIS.....</b>	<b>5</b>
<b>2.1 Introduction .....</b>	<b>5</b>
<b>2.2 Existing system.....</b>	<b>5</b>
<b>2.3 Proposed system.....</b>	<b>6</b>
<b>3.System Specification.....</b>	<b>6</b>
<b>3.1Operating System .....</b>	<b>7</b>
<b>3.2 Language.....</b>	<b>7</b>
<b>3.3Hardwarespecification.....</b>	<b>8</b>

<b>4.System Design .....</b>	<b>8</b>
<b>4.1 Introduction.....</b>	<b>8</b>
<b>4.2 ER diagram.....</b>	<b>9</b>
<b>5.Data Flow diagram.....</b>	<b>10</b>
<b>5.1 Data Dictionary.....</b>	<b>14</b>
<b>5.2 database design.....</b>	<b>14</b>
<b>6.System development .....</b>	<b>16</b>
<b>6.1 Introduction.....</b>	<b>16</b>
<b>6.2 System implementation.....</b>	<b>18</b>
<b>6.3 debugging.....</b>	<b>18</b>
<b>6.4 System Security.....</b>	<b>18</b>
<b>6.5 Scope for future enhancement.....</b>	<b>18</b>
<b>6.Source code.....</b>	<b>19</b>
<b>6.1 Login Module.....</b>	<b>19</b>
<b>7.Conclusion.....</b>	<b>27</b>
<b>8. Bibliography.....</b>	<b>27</b>
<b>9. Appendix.....</b>	<b>28</b>
<b>9.1 Sample input and output design.....</b>	<b>28</b>

## **1.INTRODUCTION**

### **1.1 About Project**

Prevention of the disease is the key to public health. It is a general saying that “prevention is always better than cures”. Vaccines protect people from catching specific diseases. Vaccines also help preventing the Spread of infectious diseases in a country. Such diseases include polio, whooping cough, diphtheria, measles, rubella (German measles), mumps, Haemophilus influenza type b (Hib) and tetanus (Malone &Hinman, 2003).

Parents are constantly concerned about the health and safety of their children. Therefore, they take many steps in order to prevent their children from catching a disease. One of the options is vaccination. Vaccine works to protect infants, children and even adults from illnesses and death caused by many infectious diseases. Vaccination has its own time, period and schedule. The system being developed is web based application that helps the parents to know the vaccination details that is to be given to the child, its importance and timings. Parents can also check the hospitals that provide the vaccinations for each age of child and can also know the date and timing. The system provides complete information regarding the vaccination.

### **1.2 About Organization**

Exploric Solutions situated in Kochi, began its journey in December 2013, with an intention to provide the best, genuine and affordable services in the field of Information Technology. A solution for all business communication needs by helping you to build your brand through the online media and to generate more sales.

Exploric Solutions always keen to produce online exposures for our clients, making it easy for them to market and manage their products or services. Spending sufficient time with each client in order to carefully design & develop the best website in accordance to his/her requirements. Every website we create is a custom design that consists of visual quality, ease of navigation and market impact. Exploric Solutions is fully committed in providing the best and genuine solutions, adding value to the services and products that are delivered and

assures you of the highest quality standards. Explorianz - our young and energetic team has the passion, and dares to explore new shores of innovation to bring out the best services and products, along with an exploring mind to share the technical ideas and knowledge to build up our young generation, with innovative and creative ideas, to be the best in this competitive world.

### **1.3 Our Services**

Web Designing (Static & Dynamic), Web Development, SEO/SMO Service, Digital Marketing, Mobile Application, Logo Designing and Animation, Visiting Card Designing, Brochure and Poster Designing, Software Training and Internship Programs (with Placement Assistance), 2D Animation.

### **1.4 Objectives of the project**

- A user friendly interface which can be operated by anyone with a minimum knowledge of the computer system.
- It should maintain a well-organized database for storing the resources that are provided by this system.
- Eliminate the entering of invalid data.
- The system should cover almost all the functional areas of this system.
- Admin has the full authority towards the system whereas customer accesses this software through login with their id's and passwords.
- The administrator should have the rights for overall management.

## **2. SYSTEM ANALYSIS**

### **2.1 Introduction**

System analysis is concerned with becoming aware of the problem, identifying the relevant and most decisional variables, analyzing and synthesizing the various factors and determining an optional or at least a satisfactory solution. During this a problem is identified, alternate system solutions are studied and recommendations are made about committing the resources used to the system.

System Analysis is the detailed study of the various operations performed by the system and their relationships within and outside the system. Analysis is the process of breaking something into its parts so that the whole may be understood.

- Understand the existing system and problems
- Analyze & Develop the proposed system
- Maintain proper databases and user friendly design

### **2.2 Existing system**

The existing system uses manual method for the whole process such as finding out the which all vaccines to be given to the child etc. This requires a lot of hard work and time consumption to complete the task. This may include human errors. In the existing system, it is difficult to retrieve some particular information. In spite of all the efforts undertaken, the destruction of data may happen often. Hence, the computerization of the system in finding out or predicting the disease to an extent helps a lot.

### **2.3 Proposed system**

In the proposed system the parent can add their child details with date of birth based on which the system will display the vaccines to be given to child. Parent has an option to set the remainder for the vaccine.

All the operations and activities related to the “Vaccination Remainder System” can be carried out efficiently. The project maintains well organized database for storing the

resources that they are provided by the admin. This helps us to eliminate the entering of invalid data. Most problems of manual system can be solved by this system.

### **3. System Specification**

#### Development Configuration

- ✓ Operating System – Windows 7 or above
- ✓ Software – Visual Studio Code
- ✓ Database and Server – XAMPP

#### Implementation Configuration

- ✓ Client machine- Windows 7 or above, Chrome/Microsoft Edge / Mozilla Firefox
- ✓ Server Machine- Windows 7 or above, WAMP or XAMP and MySQL

### **3.1 Operating System**

#### **WINDOWS 8**

Windows 8 is an operating system that was produced by Microsoft, released as part of the Windows NT family of operating systems. The product was released to manufacturing on August 1, 2012, and generally to retail on October 26 of the same year.

Windows 8 introduced major changes to the operating system's platform and user interface to improve its user experience on tablets, where Windows was now competing with mobile operating systems, including Android and iOS.[7] In particular, these changes included a touch-optimized Windows shell based on Microsoft's "Metro" design language and the Start screen (which displays programs and dynamically updated content on a grid of tiles), a new platform for developing "apps" with an emphasis on touchscreen input, integration with online services (including the ability to synchronize apps and settings between devices), and Windows Store, an online distribution for downloading and purchasing new software. Many of these features were adoptions from Windows Phone. Windows 8 added support for USB 3.0, Advanced Format hard drives, near field communications, and cloud computing. Additional security features were introduced,

such as built-in antivirus software, integration with Microsoft SmartScreen phishing filtering service and support for UEFI Secure Boot on supported devices with UEFI firmware, to prevent malware from infecting the boot process.

### **3.2 Language**

#### **JAVA**

Java is a programming language originally developed by James Gosling at Sun Microsystems (now a subsidiary of Oracle Corporation) and released in 1995 as a core component of Sun Microsystems' Java platform. The language derives much of its syntax from C and C++ but has a simpler object model and fewer low-level facilities. Java applications are typically compiled to byte code (class file) that can run on any Java Virtual Machine (JVM) regardless of computer architecture. Java is a general-purpose, concurrent, class-based, object-oriented language that is specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere." Java is currently one of the most popular programming languages in use, particularly for client-server web applications.

The original and reference implementation Java compilers, virtual machines, and class libraries were developed by Sun from 1995. As of May 2007, in compliance with the specifications of the Java Community Process, Sun relicensed most of its Java technologies under the GNU General Public License. Others have also developed alternative implementations of these Sun technologies, such as the GNU Compiler for Java and GNU Class path.

### **3.3 Hardware specification**

Processor : i3 or more

RAM : 4GB of RAM

Hard disk space : 500 GB free hard disk space

Input Devices: mouse, keyboard

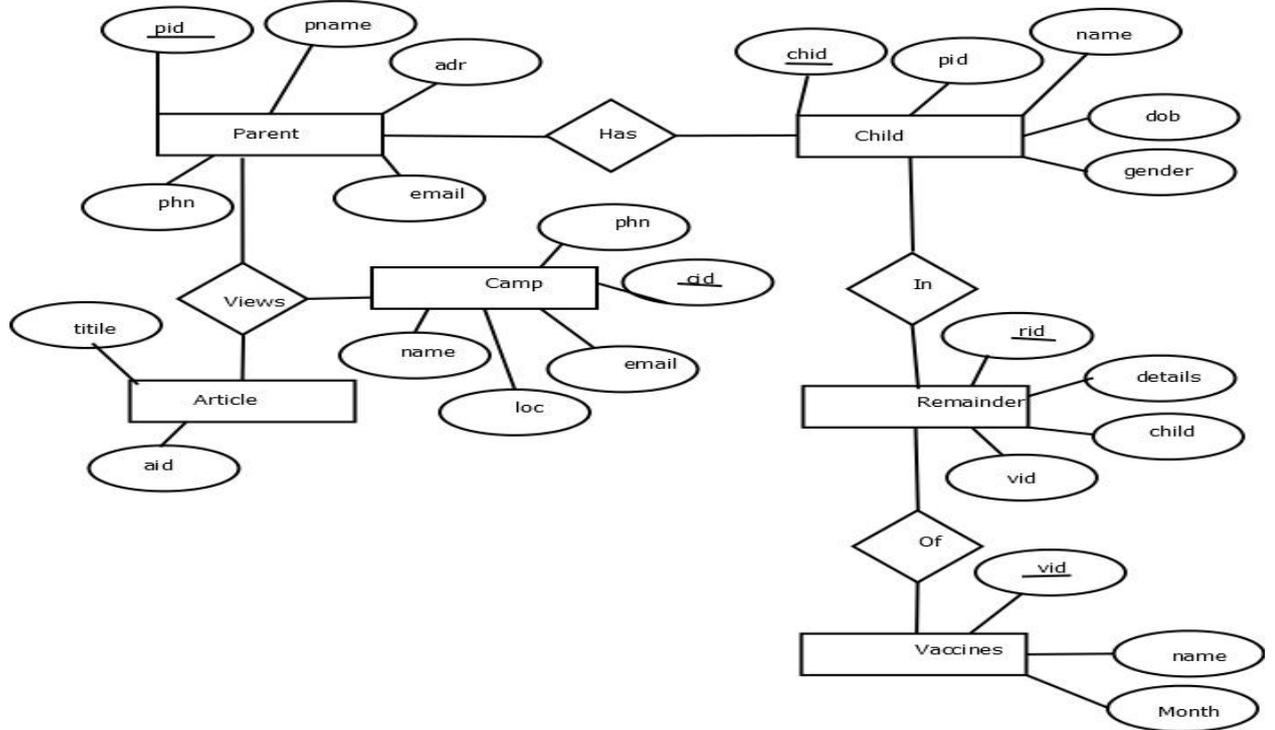
Output devices : printer, Monitor

## **4. System Design**

## 4.1 Introduction

**System Design** is the process of designing the architecture, components, and interfaces for a system so that it meets the end-user requirements.

## 4.2 ER diagram



## 5.Data Flow diagram

A DFD also known as ‘bubble chart’ has the purpose of clarifying system requirements and identifying major transformations. It shows the flow of data through a system. It is a graphical tool because it presents a picture. The DFD may be partitioned into levels that represent increasing information flow and functional detail. Four simple notations are used to complete a DFD. These notations are given below:-

### DATA FLOW

The data flow is used to describe the movement of information from one part of the system to another part. Flows represent data in motion. It is a pipe line through which information flows. Data flow is represented by an arrow.

### PROCESS



A circle or bubble represents a process that transforms incoming data to outgoing data. Process shows a part of the system that transforms inputs to outputs.

### EXTERNAL ENTITY

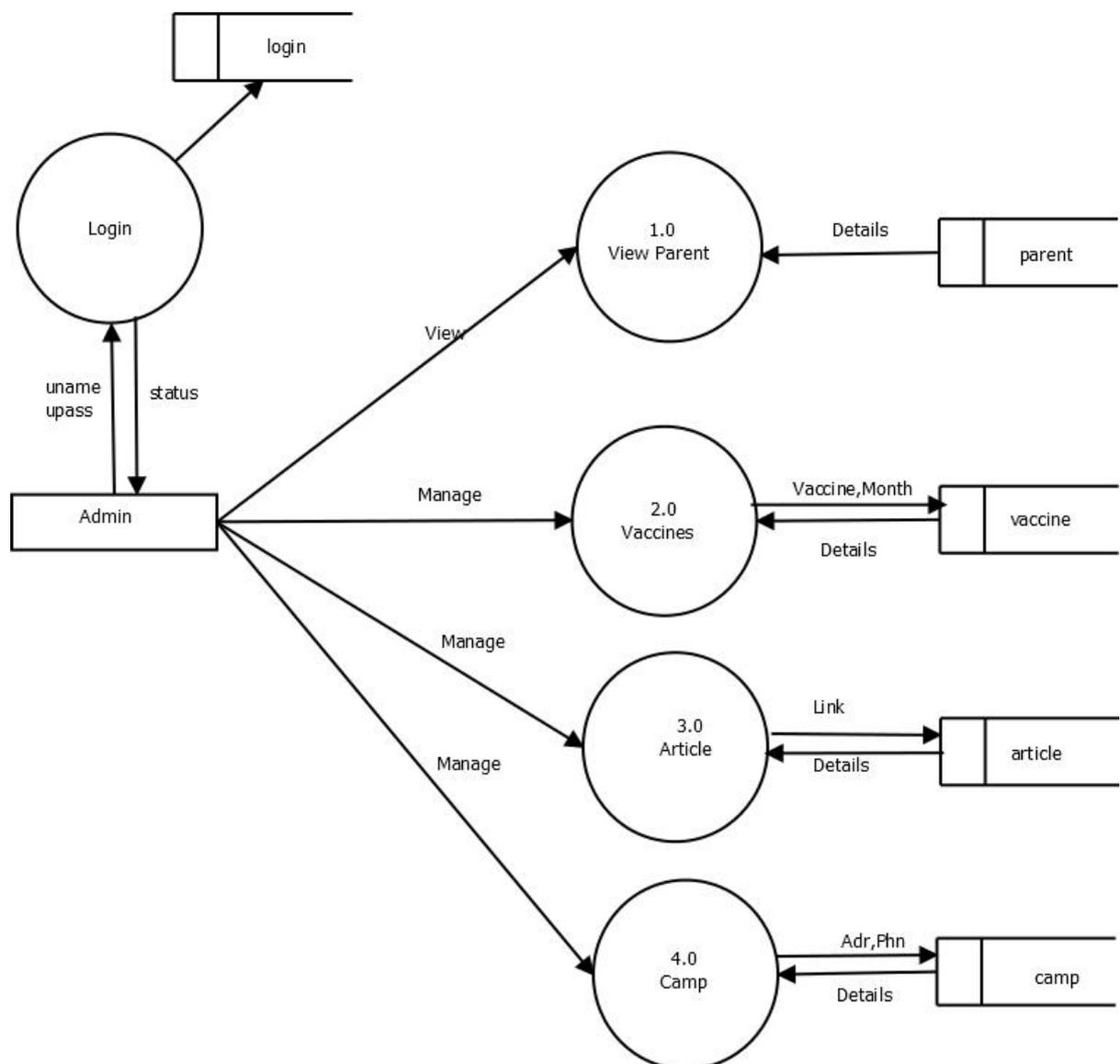


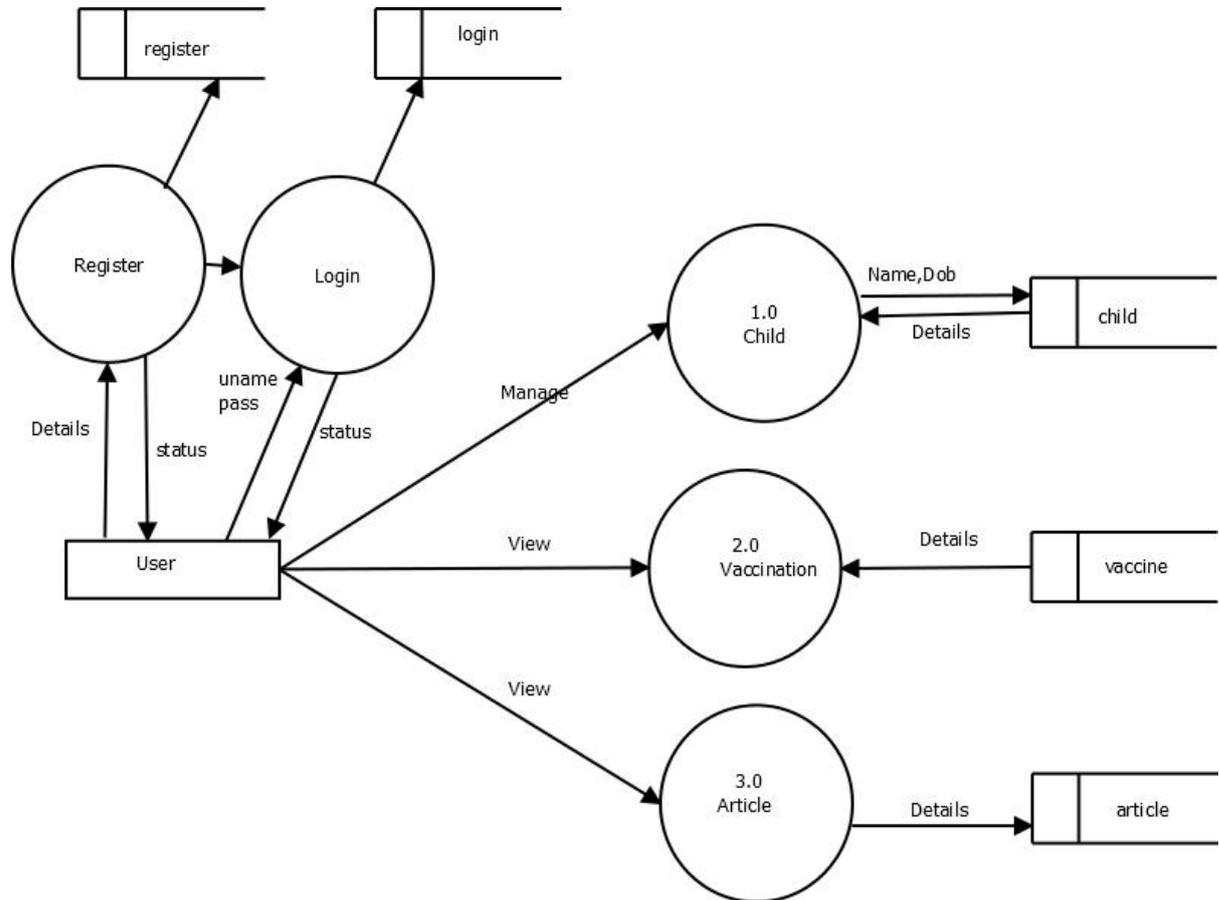
A square defines a source or destination of system data. External entities represent any entity that supplies or receive information from the system but is not a part of the system.

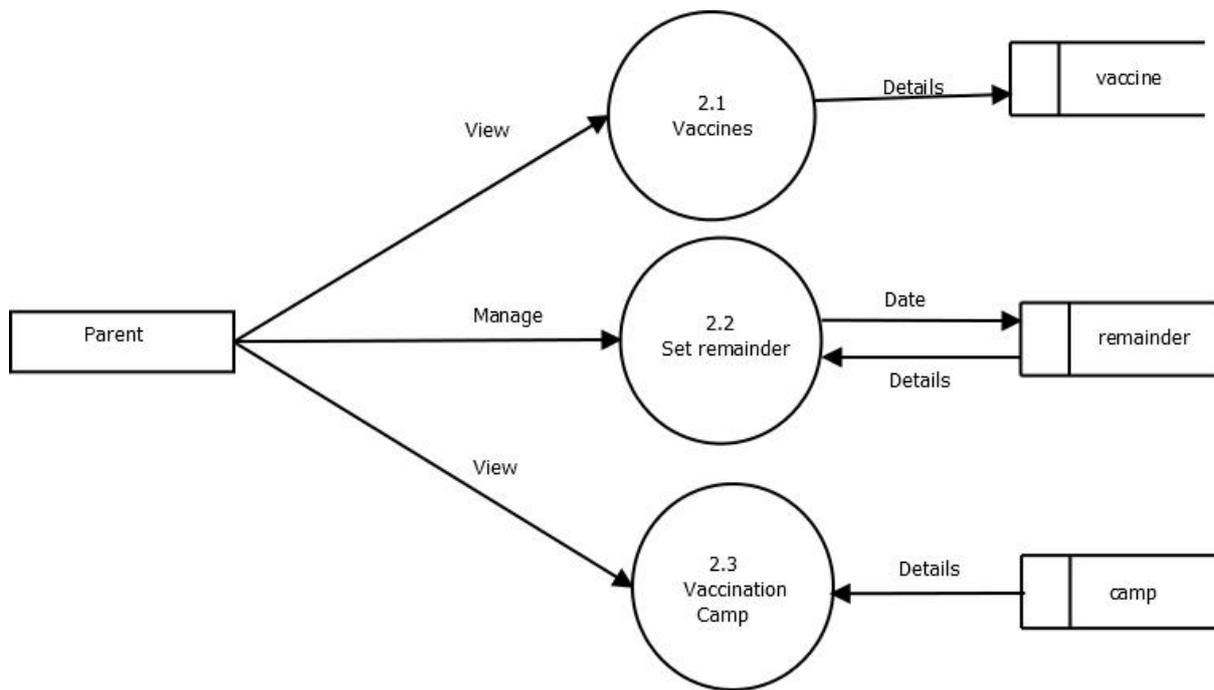
### DATA STORE

The data store represents a logical file. A logical file can represent either a data store symbol which can represent either a data structure or a physical file on disk. The data store is used to collect data at rest or a temporary repository of data. It is represented by open rectangle.



**Level 1 Admin**

**Level 1 Parent****Level 2 parent**



## 5.1 Data Dictionary

A data dictionary contains metadata i.e data about the database. The data dictionary is very important as it contains information such as what is in the database, who is allowed to access it, where is the database physically stored etc. The users of the database normally don't interact with the data dictionary, it is only handled by the database administrators.sss

## 5.2 database design

Database design can be generally defined as a collection of tasks or processes that enhance the designing, development, implementation, and maintenance of enterprise data management system. Designing a proper database reduces the maintenance cost thereby improving data consistency and the cost-effective measures are greatly influenced in terms of disk storage space. Therefore, there has to be a brilliant concept of designing a database. The designer should follow the constraints and decide how the elements correlate and what kind of data must be stored.

### Article

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Aid	Int	Primary Key	Attendance Id
Topic	Varchar(1000)	Not Null	Topic
Wlink	Varchar(1000)	Not Null	Weblink

### Camp

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
cid	Int	Not Null	Camp Id
Cname	Varchar(20)	Not Null	Camp Name
Adr	Varchar(20)	Not Null	Address
City	Varchar(20)	Not Null	City
Dist	Varchar(20)	Not Null	District
Pin	Int(10)	Not Null	Pin code
Phn	Varchar(20)	Not Null	Phone Number
Em	Varchar(20)	Not Null	Email

### Child

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Chid	Int	Primary Key	Child Id
Pid	Int	Not Null	Parent Id

Chname	Varchar(20)	Not Null	Child Name
Dob	Date	Not Null	Date Of Birth
Gen	Varchar(10)	Not Null	Gender

## Feedback

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Fid	Int	PRIMARY KEY	Feedback Id
Fdbk	Varchar(1000)	Not Null	Feedback

## Login

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Uid	Int	Not Null	UserId
Uname	Varchar(20)	Not Null	User name
Upass	Varchar(20)	Not Null	User Password
Utype	Varchar(20)	Not Null	User type

## Parent

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Pid	Int	PRIMARY KEY	Parent Id
Fname	Varchar(20)	Not Null	First Name
Lname	Varchar(20)	Not Null	Last Name
Adr	Varchar(100)	Not Null	Address
City	Varchar(20)	Not Null	City
Dist	Varchar(20)	Not Null	District
Pin	Int	Not Null	Pin Code
Phno	Varchar(20)	Not Null	Phone Number
Em	Varchar(20)	Not Null	Email

## Reminder

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Rid	Int	PRIMARY KEY	Reminder id
Vid	Int	Not Null	Vaccination Id
Pid	Int	Not Null	Parent Id

Chid	Int	Not Null	Child Id
Dte	Date	Not Null	Date

## Vaccine

FIELD NAME	DATATYPE	CONSTRAINTS	DESCRIPTION
Vid	Int	PRIMARY KEY	Vaccination Id
Vname	Varchar(50)	Not Null	Vaccination name
Descp	Varchar(2000)	Not Null	Description
Vmonths	Int	Not Null	Vaccination Months

### 6.System development

System testing is an essential step for the development of a reliable and error-free system. Once source code has been generated, software must be tested to uncover and correct as many errors as possible before delivery to your customer. Your goal is to design a series of test cases that have a high likelihood of finding errors but how, there are different methods that provide a systematic guidance for design.

#### 6.1 Introduction

Testing is a process of executing a program with the interest of finding error. A good test is one that has a high probability of finding the yet undiscovered error. Testing should systematically uncover different classes of error in a minimum amount of time with a minimum amount of effort. Two classes of inputs are provided to test the process.

1. A software configuration that includes a software requirements specification, a design specification and source code.

2. A software configuration that includes a test plan and procedure, any testing tool and test cases and their expected result.

Testing is divided into several distinct operations.

##### 1. Unit Testing

Unit testing is the first level of testing. In this process the code produced during the coding phase is verified. The goal is to test the internal logic of the modules. This is also known as

“module testing”. The modules are tested separately; this testing is carried out during programming stage itself.

This testing is carried out during the programming stage itself. In this testing step each module is found to be working satisfactorily as regards to the expected output from the module. Using a method called white box testing in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose and where each module or component of the software is tested individually. In the unit test case we will be testing the separate modules of the Software. We will test the components by passing data through it and we will be monitoring data to find the errors. We will be looking for entry and exit conditions of the data. We will make sure that all the components work without any troubles.

## **2. Integration Testing**

Integration testing is performed to check the correctness of the interface between the modules. The goal is to see if the modules can be integrated properly. In this project all the modules are combined and then entire program is tested as whole, thus in the integration step, all the errors uncovered for the next testing steps

After splitting the program into units, the units were tested together to see the defects between each module and function. It is testing two or more modules or functions together with the intent of finding interface defects between the modules or functions. Testing completed at as part of unit or functional testing, and sometimes, becomes its own standalone test phase. On a larger level, integration testing can involve putting together of groups of modules and functions with the goal of completing and verifying that the system meets the system requirements.

## **3. Validation Testing**

After validation testing , software is completely assembled as a package, interfacing errors that have been uncovered and corrected and the final series of software test; the validation test begin.

Steps taken during software design and testing can greatly improve the probability of successful integration in the larger system. System testing is actually a series of different test whose primary purpose is to fully exercise the compute-based system.

## 6.2 System implementation

Implementation is a process of ensuring that the information system is operational. It involves –

- Constructing a new system from scratch
- Constructing a new system from the existing one.

Implementation allows the users to take over its operation for use and evaluation. It involves training the users to handle the system and plan for a smooth conversion.

## 6.3 debugging

1. To launch an application into the market, it is very necessary to cross-check it multiple times so as to deliver an error-free product.
2. When we talk about delivering a bug-free product, then our main concern is all about customer satisfaction because if you are application is not up to the mark, then eventually it will demolish the company's reputation in the market.
3. In this article, we are going to see what makes debugging stand out of the queue and how it is different from software testing.

## 6.4 System Security

System security refers to protecting the system from theft, unauthorized access and modifications, and accidental or unintentional damage. In computerized systems, security involves protecting all the parts of computer system which includes data, software, and hardware. Systems security includes system privacy and system integrity.

## 6.5 Scope for future enhancement

Any system which has been in use for a number of years gradually decays and become less effective because of change in environment to which it has to be adapted. For the time being it is possible to overcome problems by amendments and minor modifications to acknowledge the need of fundamental changes.

“Vaccination Remainder System” satisfies the requirements of the management. The system is developed in a user friendly manner. It has one module for manipulating the database. The application can be enhanced in the future with the needs of the management. The database and the information can be updated to the latest coming versions. There are also possibilities for enhancing and further developing the project with the latest information and needs of the management, since the coding are in procedural block formats, altering the code is also made easy.

## 6.Source code

### 6.1 Login Module

```

<!DOCTYPE html>

<html>

<head>

<title>Clinical Lab a Medical Category Flat Bootstrap Responsive Website Template |
Contact :: w3layouts</title>

<!-- for-mobile-apps -->

<meta name="viewport" content="width=device-width, initial-scale=1">

<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />

<meta name="keywords" content="Clinical Lab Responsive web template, Bootstrap Web
Templates, Flat Web Templates, Android Compatible web template,

Smartphone Compatible web template, free webdesigns for Nokia, Samsung, LG,
SonyEricsson, Motorola web design" />

<script type="application/x-javascript"> addEventListener("load", function() {
setTimeout(hideURLbar, 0); }, false);

function hideURLbar(){ window.scrollTo(0,1); } </script>

<!-- //for-mobile-apps -->

<link href="css/bootstrap.css" rel="stylesheet" type="text/css" media="all" />

<link href="css/style.css" rel="stylesheet" type="text/css" media="all" />

<!-- js -->

<script type="text/javascript" src="js/jquery-2.1.4.min.js"></script>

<!-- //js -->

<link href='//fonts.googleapis.com/css?family=Poiret+One' rel='stylesheet' type='text/css'>

<link
href='//fonts.googleapis.com/css?family=Lato:400,100,100italic,300,300italic,400italic,700,7
00italic,900,900italic' rel='stylesheet' type='text/css'>

<!-- start-smoth-scrolling -->

```

```

<script type="text/javascript" src="js/move-top.js"></script>
<script type="text/javascript" src="js/easing.js"></script>
<script type="text/javascript">
jQuery(document).ready(function($) {
$(".scroll").click(function(event){
event.preventDefault();
$('html,body').animate({scrollTop:$(this.hash).offset().top},1000);
});
});
</script>
<!-- start-smoth-scrolling -->
<script src="js/responsiveslides.min.js"></script>
</head>
<body>
<!-- header -->
<div class="header_w3l">
<div class="container">
<nav class="navbar navbar-default">
<div class="navbar-header">
<button type="button" class="navbar-toggle" data-toggle="collapse" data-target="#bs-example-navbar-collapse-1">
<span class="sr-only">Toggle navigation</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
<h1><a href="index.html">Vaccine</span></a></h1>

```

```
</div>
<!-- top-nav -->
<div class="collapse navbar-collapse" id="bs-example-navbar-collapse-1">
<ul class="nav navbar-nav navbar-right">
<li><a href="index.html" class="active">Home</a></li>

<li><a href="login.jsp">Sign In</a></li>
</ul>
<div class="clearfix"> </div>
</div>
</nav>
</div>
</div>
<!-- header -->
<!-- banner -->
<div class="banner_w3ls page_head">

</div>
<!-- //banner -->
<div class="map all_pad">
<div class="container">

<h3 class="title agile">Register Here</h3>
<center><form action="pregaction.jsp" onsubmit="return check();">
First name <br/><input type="text" name="fn" id="fn" required /><br>
Last name <br/><input type="text" name="ln" id="ln" required /><br>
```

```
Address <br/><textarea name="ad" id="ad" required></textarea><br>
City <br/><input type="text" name="ct" id="ct" required /><br>
District <br/><input type="text" name="ds" id="ds" required /><br>
Pincode <br/><input type="text" name="pc" id="pc" maxlength="6" required /><br>
Phone number <br/><input type="text" name="pn" id="pn" required maxlength="10" /><br>
Email id <br/><input type="text" name="em" id="em" required /><br>
Password <br/><input type="text" name="pw" id="pw" required /><br>
<input type="submit" value="Register"/>
</form>
<br/>
<br/>
</center>
</div>
</div>
<!-- contact -->

<!-- smooth scrolling -->
<script type="text/javascript">
$(document).ready(function() {
/*
var defaults = {

containerID: 'toTop', // fading element id

containerHoverID: 'toTopHover', // fading element hover id
```

```

scrollSpeed: 1200,

easingType: 'linear'
        };

*/

$.UItoTop({ easingType: 'easeOutQuart' });
});
</script>

<a href="#" id="toTop" style="display: block;"> <span id="toTopHover" style="opacity:
1;"> </span></a>

<!-- //smooth scrolling -->

<script type="text/javascript" src="js/bootstrap-3.1.1.min.js"></script>

</body>

<script type="text/javascript">
    var letters=/^[a-z A-Z]+$/;
    var numbers=/^[0-9]+$/;
    function check()
    {
        if(!document.getElementById("fn").value.match(letters))
        {

            alert('Please input alphabet characters only,enter first name');
            return false;
        }

        else if(!document.getElementById("ln").value.match(letters))

```

```
{  
  
    alert('Please input alphabet characters only,enter last name');  
    return false;  
}  
  
else if(!document.getElementById("ct").value.match(letters))  
{  
  
    alert('Please input alphabet characters only,enter city');  
    return false;  
}  
  
    else if(!document.getElementById("ds").value.match(letters))  
{  
  
    alert('Please input alphabet characters only,enter district');  
    return false;  
}  
  
else if(!document.getElementById("pn").value.match(numbers))  
{  
  
    alert('Please input numeric characters only,enter phone number');  
    return false;  
}  
  
    else if(document.getElementById("pn").value.length<10)  
{  
  
    alert('Enter phone number with minimum lebgth of 10 numbrers');
```

```
return false;
    }
    else if(document.getElementById("pc").value.length<6)
{
alert('Enter pin with minimum lebgth of 6 numbrers');
return false;
    }
    else if(!document.getElementById("pc").value.match(numbers))
{
alert('Please input numeric characters only,enter pin code');
return false;
}

    else if(document.getElementById("pass").value.length<8)
{
alert('Enter password with minimum lebgth of 8 characters');
return false;
}
else
{
return true;
}
}
```

</script>

## 7.Conclusion

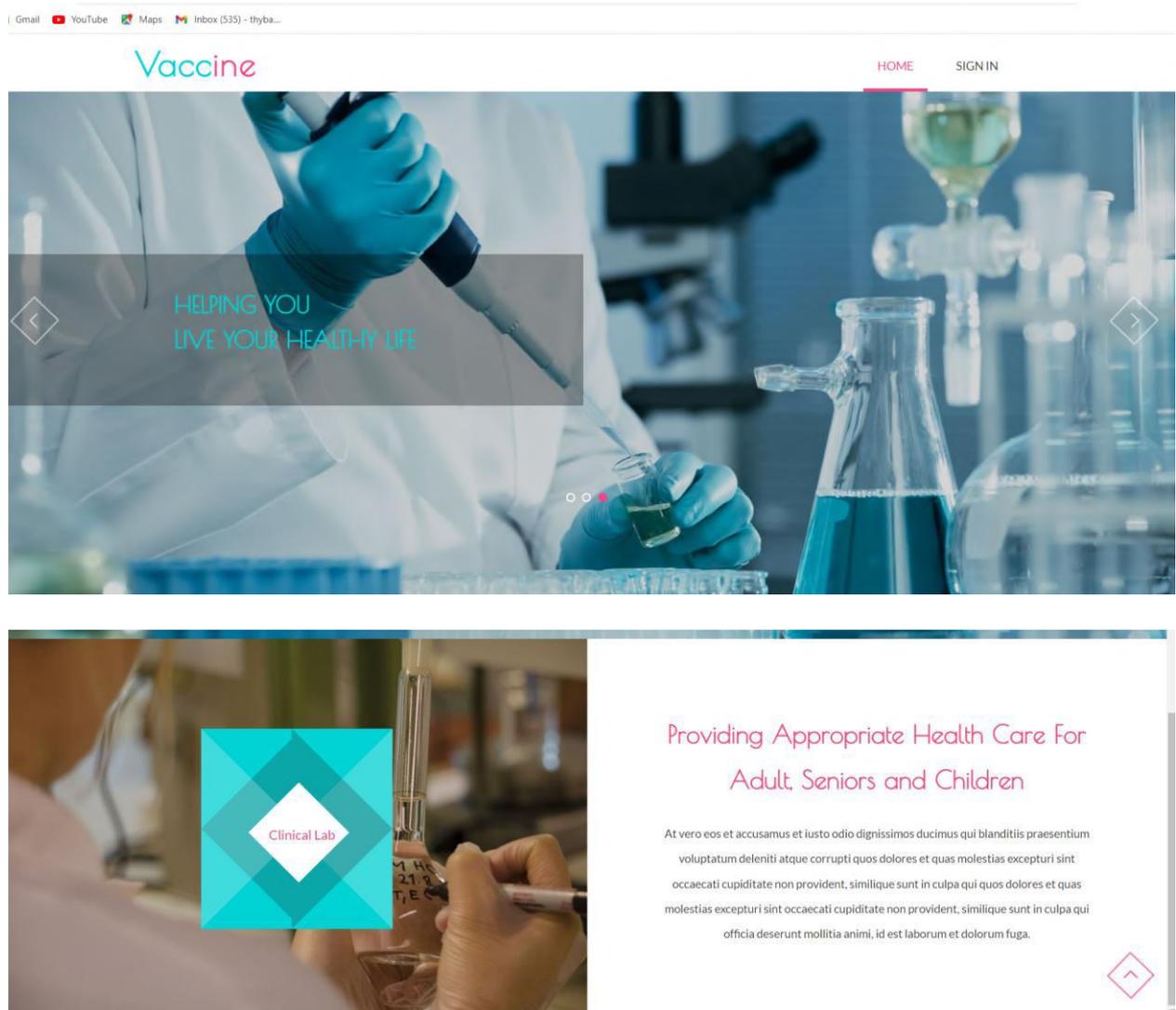
Parents are constantly concerned about the health and safety of their children. Therefore, they take many steps in order to prevent their children from catching a disease. One of the options is vaccination. Vaccine works to protect infants, children and even adults from illnesses and death caused by many infectious diseases. Vaccination has its own time, period and schedule. The system being developed is web based application that helps the parents to know the vaccination details that is to be given to the child, its importance and timings. Parents can also check the hospitals that provide the vaccinations for each age of child and can also know the date and timing. The system provides complete information regarding the vaccination.

## 8.Bibliography

- \*Elias. M. Awad, “*System Analysis and Design*”, Galgoeia publications – Second Edition 2005.
  - \*Henry Korth,” *Data Base Design Concept*”, Mc grew hill – Fifth Edition 2001.
  - \*Roger.S.Pressman,” *Software Engineering* “,Mc grew hill Fifth Edition 2006.
- [www.stackoverflow.com](http://www.stackoverflow.com)
- [www.javatpoint.com](http://www.javatpoint.com)

## 9. Appendix

### 9.1 Sample input and output design



### Article Entry

Topic  Weblink

Topic Name	Web Link	
asd	<a href="https://health.economictimes.indiatimes.com/tag/vaccination">https://health.economictimes.indiatimes.com/tag/vaccination</a>	delete
tgfhu	efefgh	delete
vaccine	<a href="https://www.cdc.gov/vaccines/schedules/hcp/imz/child-adolescent.html">https://www.cdc.gov/vaccines/schedules/hcp/imz/child-adolescent.html</a>	delete

### Vaccine Entry

Vaccine Name   
 Description   
 Months

Vaccine ID	Vaccine Name	Description	Vaccination Months	
3	polio	polio	2	delete
4	measles 2	regtrg	2	delete

### Parents Registered

First Name	Last Name	Address	City	District	Pincode	Phone Number	Email ID	
xcv	xcvxcv	xcvxcvxcv	xcvxcv	daf	324	3423423423	xcv@xcv.xcv	Delete

### Camp Entry

Camp name

Address

City

District

Pincode

Phone number

Email id

Add

Camp Name	Address	City	District	Pincode	Phone Number	Email ID	
kollam	ybnhh	Ernakulam	ernakulam	682028	7025596077	dfg@gmail.com	delete
thrissur	nmnb	thrissur	thrissur	682028	6733905620	var@gmail.com	delete

### Feedback

Feedback ID	Feedback	
3	helpfull	delete
4	nice	delete
5	good reminder	delete
6	remind me	delete

### Child

Child name

Date of Birth

Gender  
 Male  Female

Add

Child Name	
gouri	Details
anju	Details
kevin	Details
meera	Details
arjun	Details
alan	Details
hañz	Details
avantika	Details



## Camp

Camp Name	Address	City	District	Pincode	Phone Number	Email ID
kollam	vbnhh	Ernakulam	ernakulam	682028	7025596077	dfg@gmail.com
thrissur	nmnb	thrissur	thrissur	682028	6733905620	var@gmail.com

**GROWTH PERFORMANCE OF *ABELMOSCHUS ESCULENTUS* L.GROWN IN SOILS CONTAINING DIFFERENT FERTILIZER COMBINATIONS**

DISSERTATION  
SUBMITTED IN PARTIAL FULLFILLMENT OF THE  
REQUIREMENTS FOR THE AWARD OF THE DEGREE OF  
BACHELOR OF SCIENCE

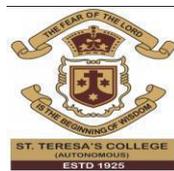
IN

BOTANY

BY

TINCY. K.P

REG NO : AB19BOT017



DEPARTMENT OF BOTANY AND CENTRE FOR  
RESEARCH  
ST.TERESA'S COLLEGE (AUTONOMOUS)  
ERNAKULAM  
2022

**GROWTH PERFORMANCE OF *ABELMOSCHUS ESCULENTUS* L.GROWN IN SOILS CONTAINING DIFFERENT FERTILIZER COMBINATIONS**

DISSERTATION  
SUBMITTED IN PARTIAL FULLFILLMENT OF THE  
REQUIREMENTS FOR THE AWARD OF THE DEGREE OF  
BACHELOR OF SCIENCE

IN

BOTANY

BY

TINCY. K.P

REG NO : AB19BOT017



DEPARTMENT OF BOTANY AND CENTRE FOR  
RESEARCH  
ST.TERESA'S COLLEGE (AUTONOMOUS)  
ERNAKULAM  
2022

## CERTIFICATE

This is to certify that the dissertation entitled "Growth performance of *Abelmoschus esculentus* L.grown in soil containing different fertilizers".Submitted in partial fulfillment of the requirements for the award of the Degree of Bachelor of science in Botany is an authentic work carried out by TINCY K.P (AB19BOT017) under the supervision and guidance of of Dr. Liza Jacob.



Dr.LIZA JACOB

Head of the Department of Botany and centre for research

Supervising Teacher

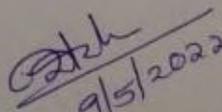
St.Teresa's College(Autonomous)

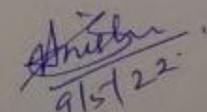
Ernakulam



Place : Ernakulam

Date : 31 - 03 - 22

  
9/5/2022

Anisha S   
9/5/22

Anite N

## ACKNOWLEDGMENT

I thank God almighty for the blessings bestowed upon me that has helped me to complete the project successfully. I express my heartfelt gratitude and sincere thanks to Dr. Liza Jacob (Head of the Department) for the valuable guidance and support throughout the project.

I also express my sincere thanks to Department of Botany and all my teachers for the valuable suggestions and help. My special thanks to all the non-teaching staff and my dear friends for their moral support.



Place : Ernakulam

Date :

TINCY K.P

## CONTENTS

Chapter No:	Title	Page no:
1	Introduction	1
2	Review of literature	3
3	Materials and methods	6
4	Observation and Result	10
5	Discussion	22
6	Summary and Conclusion	24
7	Bibliography	25

## INTRODUCTION

Lady's finger or okra (*Abelmoschus esculentus* L.) is one of the most commonly grown and consumed species of the family Malvaceae. It is a nutritious vegetable cultivated in tropical to subtropical and warm temperate regions of the world. Okra is an affordable source of protein, carbohydrates, minerals and vitamins, dietary fiber, and other phytonutrients with physiological benefits. Okra seeds contain a considerable amount of edible oil, which is rich in unsaturated fatty acids. Okra is a tropical plant which grows best in warm climates. It is available yearround, with a peak season during the summer months. The pods grow rapidly, being ready for harvest in about 60 days of summer weather, when grown from seed. They must be picked about 4 to 5 days after flowering, when 4 inches or so in length. Okra comes in varying shades of green (there is also a new red variety), and can be smooth or have a ribbed surface. Okra is a good source of vitamin C and A, also B complex vitamins, iron and calcium. It is low in calories, a good source of dietary fiber, and is fat-free. In West Africa, leaves, buds, and flowers of okra are also consumed. The dried seeds provide oil, protein, vegetable curd, and a coffee additive or substitute. Okra dry seeds contain 18–20% oil and 20–23% crude protein. Foliage can be used for biomass, and the dried stems serve as a source of paper pulp or fuel. To a limited extent, okra is used in canned, dehydrated, or frozen forms. It has an average nutritive value of 3.21, which is higher than tomato, eggplant, and most cucurbits except bitter melon. The cultivation of okra extends throughout the tropics and warmer parts of temperate Asia. It is commercially grown in India, Turkey, Iran, West Africa, Yugoslavia, Bangladesh, Afghanistan, Pakistan, West Bengal, Burma, Japan, Malaysia, Brazil, Ghana, Ethiopia, Cyprus, and the southern USA. Okra is an important crop in temperate and tropical climates. Okra packs a good amount of antioxidants which are known to eliminate the damage done by free radicals. These antioxidants in okra include polyphenols, isoquercetin, flavonoid, etc. Apart from helping with the harmful molecules, polyphenols also improve heart health and protect the Research in animals has shown that bhindi or bhindi extracts might help lower blood sugar levels. Furthermore, the study suggested that the vegetable lowered the absorption of sugar in the digestive tract, assisting with maintaining a stable sugar level. Okra is offered by the gel-like substance present in the vegetable known as mucilage. This substance binds to the cholesterol molecules during the process of digestion and is excreted out with stools

instead of getting absorbed in the body. Moreover, the benefits of poly phenol in bhindi also assist with keeping your heart healthy. Warm humid tropical conditions are ideal for luxurious growth and high yield of okra. It grows best within a temperature range of 24-27°C. It is highly tolerant to high temperature and drought condition but, highly sensitive to frost and temperature below 12°C. Seeds germinate poorly at ground temperature of 20 °C or less. The crop can be successfully grown in rainy season even in heavy rainfall area. In India, it is grown in summer months and during the rainy season. Okra can be grown in a wide range of soils. However, it grows best in loose, friable, well-drained sandy loam soils rich in organic matter. It also gives good yield in heavy soils with good drainage. A pH range of 6.0-6.8 is considered as optimum. Alkaline, saline soil and sand soils with poor drainage are not good for this crop. The main planting seasons for okra under Indian conditions are February-March and June-July. However, optimum time of seed sowing varies greatly depending upon climate, varieties and their temperature requirement for growth. In north Indian plains, spring summer crop is sown in February- March, where as rainy season crop is sown in the month of June-July. For Kharif season seeds are sown in the month of June - July. The seed rate for rainy season is 8 to 10 kg. For summer season it is sown in the month of February - March.

The characteristics of soil play a major role in the plants ability to absorb water and nutrients. Plants cannot survive without essential nutrients such as N,P,K,Ca,Mg,S etc. Each soil is different from one another due to the difference in the proportion of components present in them. So that the growth of the plants also varies with the components in the soil in which they grow. The present project is a comparative study to understand the growth performance of *Abelmoschus esculentus* L. grown in soils amended with NPK fertilizer, cow-dung and egg shell. Studies were conducted based on the observations made during different growth stages of the plants.

## REVIEW OF LITERATURE

A review of literature in the related field was conducted in order to have an understanding of the work done on the subject.

Khalofah et al. (2022) studied the impact of NPK fertilizer on growth and nutrient accumulation in juniper revealed that increasing fertilizer doses improved the growth and nutrient acquisition and application of fertilizer on intact soil recorded the highest values of growth rates.

According to the study conducted by Du Q-J et al. (2021) noticed that the fertilizer rate at 840 kg/ha has not only maintained the productivity of soil but also tomato growth and quality of fruit which makes the non-pressure gravity irrigation a potential and cost-effective way for fertilizer application.

Adam truax et al. (2006) analyzed “how soil composition relates to plant growth and diversity”. They hypothesized that soils with higher quantities of nitrogen, phosphorous, potassium and neutral pH will in turn have a higher diversity of plant life than those soils that are more nutrient poor.

In the study by Passioura (1991) on the soil structure and plant growth he noticed that root grow rapidly in friable soil and to supply the leaf with water and nutrients.

Sanni (2016) studied the effect of compost, cow dung and NPK 15-15-15 fertilizer on growth and yield performance of Amaranth shows that the use of cow dung and composts has potential to improve the growth and yield of amaranth and improves the soil physio chemical properties in the study area.

Serenella Nardi (2001) studied physiological effect humic substance on higher plants and found that humic substance exhibit positive influence on nutrient uptake, plant cell growth and development.

Petra Kidd and John Proctor (2001) investigated “why plants grow poorly on very acid soil are ecologist missing the obvious ?” In the grass *Holcus lanatus* L. racial difference are reported in plant growth response to increasing acidity. The fact of plant adaptation to H<sup>+</sup> toxicity supports the view that this is a important factor in acid soil.

The study on Relationships between pH values of organic soils and availabilities of 12 plant nutrients by Leucas and Davis (1981) shows acidity in soil decrease the nitrogen uptake by plants which is essential for satisfactory growth.

Ezzo (2010) observed the response of sweet pepper, made grown in sandy soil and clay. This results the increases in plant length, stem diameter.

Zhong Kato (1988) studied on “The effects soil moisture on the growth and yield of solanaceous fruit”.

The study conducted for evaluation of eggshell as organic fertilizer on sweet basil by Wijaya and Teo (2019) shows that eggshell has significant effect to plants growth.

Satyanarayana et al. (Year?) studied on influence of integrated use of farmyard manure and inorganic fertilizers on yield components of irrigated lowland rice show greater yield.

Molik et al. (2016) experiment on to evaluate the effects of organic and inorganic fertilizers on the yield components of okra result that organic manure favoured the yield okra than inorganic fertilizers.

Rafiqul Islam (2002) analysed the effect of Different Levels of Chemical and Organic Fertilizers on Growth, yield contributing characters and protein content of wheat increased with increase in the levels inorganic fertilizers, addition of composts with fertilizer and also by increasing split application of nitrogen.

Tripathy (1978) studied the effect of chemical and fertilizer factor effluent on mineral composition in wheat and he observed a reduction in the higher concentration.

Kyi MoKumudra et al. (2017) analysed the combined effect of Organic Manures and Inorganic Fertilizers on the Growth and Yield of Hybrid Rice.

Application of dilute paper mill effluent increased the organic carbon, cation exchange capacity, available N, P, K, and micronutrients in the soil. The amount of manganese and iron content increased with the addition of spent wash to the soil (Sharma and Aggarwal, 2003)

Ano1 and Ubochi (2007) studied the mechanism of reaction leading to neutralization of soil acidity by animal manures was studied in the laboratory. Animal manures significantly increased the soil pH from 4.6 to values above 5.6 and also reduced exchangeable acidity from 3.00 cmol kg<sup>-1</sup> to values below 0.35 cmol kg<sup>-1</sup>.

Ayoola and Adeniyani (2021) to determine the effects of NPK fertilizer and poultry manure on the yield and yield components in cassava/maize/melon systems. Crop yields were statistically the same under NPK alone and NPK + poultry manure but significantly higher than both poultry manure alone and control in both locations.

Organic fertilizers as a route to controlled release of nutrients. The improper use of organic fertilizers leads to over fertilization or nutrient deficiency in the soil. Hence, controlled release of organic fertilizers is an effective and advanced way to overcome these impacts and maintain sustainable agriculture yield (Hitha Shaji and Linu Mathew 2021).

Prakash et al. did field experiments during kharif 2017 and 2018 to find out the best level of NPK fertilizer for rice with two varieties of rice of same duration (130-135 days) and concluded that yield, nutrient uptake, nutrient requirement and nutrient use efficiency of the rice were significantly influenced by levels of CF.

Application of diluted paper mill effluent increased the organic carbon, cation exchange capacity, available N, P, K and micro nutrients in the soil. The amount of manganese and iron content increased with the addition of solvent wash to the soil (Sharma and Aggarwal, 2003)

## MATERIALS AND METHODS

The study was conducted to find out the growth performance of *Abelmoschus esculentus* L. grown in soils containing NPK fertilizer, cow-dung and eggshell. The materials and method of present study as follows:

Plant used for study : *Abelmoschus esculentus* L.

Materials required : Okra seeds, grow bags, NPK fertilizer, Cow-dung, eggshell, soil, sand.

### Method 1

For doing the experiment, the seeds of okra were sown in six different combinations of soil were taken such as soil with sand and NPK fertilizer, soil with sand and cow-dung, soil with sand and slightly broken eggshell and soil with powdered egg shell. The growth and development were studied with the help of following experiment and observations. The grow bags were prepared in the ratio 1:1:1. In the first grow bag soil with sand and NPK fertilizer were mixed. In second grow bag soil with sand and cow-dung were mixed. In third grow bag soil with sand and eggshell were mixed. Three grow bags of the same combination was also prepared. The okra seeds were sown in the sandy soil and allowed to germinate. After seedling attain two leaves the healthy seedlings were selected and transferred to grow bags. One plantlet is planted on each grow bag.

### Method 2

First the seeds are dipped in water for one night then it can be tremor into the soil. The seed will germinate within 6-8 days the tiny sproutout will be visible then it can be visible. Then it can be converge it into the grow bag filled with soil three of them are taken for the analysis. One plant is kept in the mixture of egg shell cow dung NPK. Egg shell and cow dung mixture is put in the second plant. Third one is treated in the mixture of egg shell and npk. Plant growth height are taken and analyse the effect and quality of fertilizer.

### **Method 3**

For doing the experiment , the seed of okra were sows in different conditions. They are soil - sand, soil with NPK and cowdung, soil- whole egg shell.

The growth and development were studied with the help of following experiment and observation. The grow bags were prepared in the ratio 1:1:1. In the first grow bag only sand - soil is used. In second grow bag sand - soil with NPK and cowdung is mixed. And the third grow bag sand - soil with egg shell whole is mixed. In these 3 combination with two bags of one type is prepared. The okra seed attains more than 2 leaves the healthy seedlings were selected and transferred in the grow bags. One plantlet was planted in each grow bag.

## **GROWTH PARAMETERS**

Various growth parameters considered for the study were:

### **1. MORPHOLOGICAL PARAMETER**

- a. Shoot length

### **2. PLANT BIOMASS**

- a. Fresh weight of leaf,stem,root
- b. Dry weight of a leaf,stem,root

### **3. PLANT PIGMENT**

- a. Chlorophyll
- b. Carotenoid

## **CHLOROPHYLL AND CAROTENOID CONTENT**

The method of Arnon (1949) was employed for the quantitative estimation of chlorophyll and carotenoid content.

80% acetone was prepared. A pre weighed (250mg) quantity of fresh leaf material was ground into fine paste. 10ml of 80% acetone was added into it. The extract was centrifuged repeatedly till the leachate become colorless. The supernatant was taken together was taken together and was made up to 25ml with 80% acetone. The extract was kept away from direct sunlight. The optical density of the extract was read at 470,490,580,610,700 wavelengths. The samples were analyzed in duplicates.

## OBSERVATION AND RESULTS

### SHOOT LENGTH (Table 2)

Growth of *Abelmoschus esculentus* L. in NPK mixed with soil and sand

- The shoot length observed on 1<sup>st</sup> week was 6cm
- The shoot length observed on 2<sup>nd</sup> week was 7cm.
- The shoot length observed on 3<sup>rd</sup> week was 9cm.
- The shoot length observed on 4<sup>th</sup> week was 9cm.
- The shoot length observed on 5<sup>th</sup> week was 11cm.
- The shoot length observed on 6<sup>th</sup> week was 13cm.
- The shoot length observed on 7<sup>th</sup> week was 13cm.
- The shoot length observed on 8<sup>th</sup> week was 15cm.
- The shoot length observed on 9<sup>th</sup> week was 17 cm.
- The shoot length observed on 10<sup>th</sup> week was 20cm.

Growth of *Abelmoschus esculentus* L.in Cow dung mixed with sand and soil.

- The shoot length observed on 1<sup>st</sup> week was 6cm
- The shoot length observed on 2<sup>nd</sup> week was 8cm.
- The shoot length observed on 3<sup>rd</sup> week was 9cm.
- The shoot length observed on 4<sup>th</sup> week was 10cm.
- The shoot length observed on 5<sup>th</sup> week was 12cm.
- The shoot length observed on 6<sup>th</sup> week was 14cm.
- The shoot length observed on 7<sup>th</sup> week was 16cm.
- The shoot length observed on 8<sup>th</sup> week was 18cm.
- The shoot length observed on 9<sup>th</sup> week was 21cm.
- The shoot length observed on 10<sup>th</sup> week was 23cm.

Growth of *Abelmoschus esculentus* L.in NPK mixed with sand and soil.

- The shoot length observed on 1<sup>st</sup> week was 5cm
- The shoot length observed on 2<sup>nd</sup> week was 8cm.

- The shoot length observed on 3<sup>rd</sup> week was 11cm.
- The shoot length observed on 4<sup>th</sup> week was 12cm.
- The shoot length observed on 5<sup>th</sup> week was 12 cm.
- The shoot length observed on 6<sup>th</sup> week was 13cm.
- The shoot length observed on 7<sup>th</sup> week was 15cm.
- The shoot length observed on 8<sup>th</sup> week was 17cm.
- The shoot length observed on 9<sup>th</sup> week was 19 cm.
- The shoot length observed on 10<sup>th</sup> week was 21cm.

Growth of *Abelmoschus esculentus* L. in soil mixed with cow dung and NPK

- The shoot length observed on 1<sup>st</sup> week was 17cm
- The shoot length observed on 2<sup>nd</sup> week was 21cm
- The shoot length observed on 3<sup>rd</sup> week was 25cm
- The shoot length observed on 4<sup>th</sup> week was 28cm
- The shoot length observed on 5<sup>th</sup> week was 31cm
- The shoot length observed on 6<sup>th</sup> week was 34cm
- The shoot length observed on 7<sup>th</sup> week was 37cm
- The shoot length observed on 8<sup>th</sup> week was 41cm
- The shoot length observed on 9<sup>th</sup> week was 45cm
- The shoot length observed on 10<sup>th</sup> week was 51cm

Growth of *Abelmoschus esculentus* L. in soil mixed with whole egg shell

- The shoot length observed on 1<sup>st</sup> week was 18cm
- The shoot length observed on 2<sup>nd</sup> week was 23cm
- The shoot length observed on 3<sup>rd</sup> week was 28cm
- The shoot length observed on 4<sup>th</sup> week was 33cm
- The shoot length observed on 5<sup>th</sup> week was 37cm
- The shoot length observed on 6<sup>th</sup> week was 40cm
- The shoot length observed on 7<sup>th</sup> week was 45cm
- The shoot length observed on 8<sup>th</sup> week was 49cm

- The shoot length observed on 9th week was 54cm
- The shoot length observed on 10th week was 58cm

Growth of *Abelmoschus esculentus* L. in soil and sand

- The shoot length observed on 1st week was 12 cm
- The shoot length observed on 2nd week was 15cm
- The shoot length observed on 3 rd week was 19cm
- The shoot length observed on 4th week was 22cm
- The shoot length observed on 5th week was 25cm
- The shoot length observed on 6th week was 28cm
- The shoot length observed on 7th week was 31cm
- The shoot length observed on 8th week was 35cm
- The shoot length observed on 9th week was 38cm
- The shoot length observed on 10th week was 39cm

Growth of *Abelmoschus esculentus* L. Eggshell mixed with NPK and cow dung

- The shoot length was observed on 1st week 12 cm
- The shoot length was observed in 2nd week 15 cm
- The shoot length was observed on 3rd week 22 cm
- The shoot length was observed on 4th week 27 cm
- The shoot length was observed on 5th week 32 cm
- The shoot length was observed on 6th week 36 cm
- The shoot length was observed on 7th week 40 cm
- The shoot length was observed on 8th week 44 cm
- The shoot length was observed on 9th week 49 cm
- The short length was observed on 10th week 53 cm

Growth of *Abelmoschus esculentus* L. Egg shell mixed with cow dung

- The short length was observed on 1st week 19 cm
- The short length was observed on 2nd week 26 cm
- The short length was observed on 3rd week 30cm
- The short length was observed on 4th week 34 cm
- The short length was observed on 5th week 38 cm
- The short length was observed on 6th week 42 cm
- The short length was observed on 7th week 46 cm
- The short length was observed on 8th week 51 cm
- The short length was observed on 9th week 55 cm
- The short length was observed on 10th week 55 cm

Growth of *Abelmoschus esculentus* L. Egg shell mixed with NPK

- The short length was observed on 1st week 13 cm
- The short length was observed on 2nd week 16 cm
- The short length was observed on 3rd week 19 cm
- The Short length was observed on 4th week 22 cm
- The short length has observed on 5th week 26 cm
- The short length has observed on 6th week 31cm
- The short length was observed on 7th week 35 cm
- The short length was observed on 8th week 40 cm
- The short length was observed on 9th week 46 cm
- The short length was observed on 10th week 51cm

**TABLE 1 : SHOOT LENGTH OF *ABELMOSCHUS ESCULENTUS* L. OBSERVED WEEKLY**

SL NO:	Combinations	SHOOT LENGTH IN WEEKS (cm)									
		1	2	3	4	5	6	7	8	9	10
1	Cowdung	6	8	9	10	12	14	16	18	21	23
2	NPK	5	8	11	12	12	13	15	17	19	21
3	Eggshell	6	7	9	9	11	13	13	15	17	20
4	Cowdung and NPK	17	21	25	28	31	34	37	41	45	51
5	Whole eggshell	18	23	28	33	37	40	45	49	54	58
6	Soil and sand	12	15	19	22	25	28	31	35	38	39
7	Eggshell,NPK and cowdung	12	15	22	27	32	36	40	44	49	53
8	Eggshell with cowdung	19	26	30	34	38	42	46	51	55	55
9	Eggshell with NPK	13	16	19	22	26	31	35	40	46	51

## **HEIGHT AND WEIGHT OF ROOT,SHOOT,WHOLE PLANT AND NUMBER OF LEAVES AND FRUITS (Table 2)**

Plant in cowdung + soil and sand

- Height of the plant = 23 cm
- Number of leaves = 9
- Number of fruits = 2
- Weight of whole plant = 5.89g
- Weight of root = 0.43g
- Weight of shoot = 3.27g

Plant in eggshell + soil and sand

- Height of the plant = 20 cm
- Number of leaves = 11
- Number of fruits = 4
- Weight of whole plant = 4.83g
- Weight of root = 0.34g
- Weight of shoot = 4.27g

Plant in cowdung with NPK

- Height of the plant = 55 cm
- Number of leaves = 13
- Number of fruits = 5
- Weight of whole plant = 29.87g
- Weight of root = 2.99g
- Weight of shoot = 28.97g

Plant in whole eggshell

- Height of the plant = 62 cm
- Number of leaves = 12
- Number of fruits = 10
- Weight of whole plant = 54.52g

- Weight of root = 5.04g
- Weight of shoot = 47.76g

Plant in soil and sand

- Height of the plant = 41 cm
- Number of leaves = 18
- Number of fruits = 6
- Weight of whole plant = 14.82g
- Weight of root = 1.95g
- Weight of shoot = 11.36g

Plant in eggshell ,cowdung and NPK

- Height of the plant = 55 cm
- Number of leaves = 15
- Number of fruits = 6
- Weight of whole plant = 13.86g
- Weight of root = 0.89g
- Weight of shoot = 12.77g

Plant in eggshell and cowdung

- Height of the plant = 51 cm
- Number of leaves = 3
- Number of fruits = 8
- Weight of whole plant = 14.08g
- Weight of root = 2.25g
- Weight of shoot = 16.00g

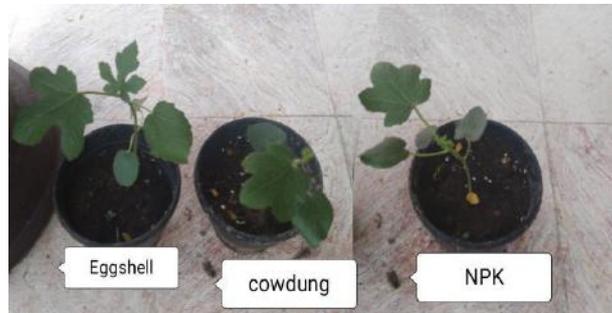
**TABLE 2 : HEIGHT AND WEIGHT OF  
ROOT,SHOOT,WHOLE PLANT AND NUMBER OF  
LEAVES AND FRUITS**

SL NO ;	COMBINATI ONS	HEIG HT ( in cm)	NO:O F LEAV ES	NO:O F FRUI TS	WEIGHT OF WHOLE PLANT (in gm)	WEIGH T OF ROOT (in gm)	WEIGH T OF SHOOT (in gm)
1	Cowdung	23	9	2	5.89	0.43	3.27
2	Eggshell	20	11	4	4.83	0.34	4.27
3	Cowdung and NPK	55	13	5	29.87	2.99	28.97
4	Whole eggshell	62	12	10	54.52	5.04	47.76
5	Soil and sand	41	18	6	14.82	1.95	11.36
6	Eggshell,NPK and cowdung	55	15	6	13.86	0.89	12.77
7	Eggshell with cowdung	51	3	8	14.08	2.25	16.00

### PIGMENT CONTENT OF THE LEAVES:(Table 3)

Treatment	Chlorophyll a(mg/g)	Chlorophyll b(mg/g)	Carotenoid (mg/g)
Eggshell, soil and sand	3.43	3.33	1.55
Cow dung, soil and sand	3.12	3.01	1.42
NPK. Soil and sand	3.42	3.22	1.49
Egg shell, cow dung and NPK	3.37	3.20	1.54
Egg shell and Cow dung	3.41	3.30	1.55
Egg shell, NPK	3.36	3.28	1.59
Cow dung and NPK	2.97	2.85	1.43
Soil and sand	1.87	1.54	0.76
Egg shell and soil	3.49	3.47	1.78

*ABELMOSCHUS ESCULENTUS* L.grown in Eggshell  
cow-dung  
NPK



*ABELMOSCHUS ESCULENTUS* L.grown in  
 Cowdung and NPK  
 Whole eggshell  
 Soil and sand



soil and sand    Cowdung and NPK    Whole eggshell



soil and sand    Whole eggshell    Cowdung and NPK



soil and sand    Cowdung and NPK    Whole eggshell

*ABELMOSCHUS ESCULENTUS* L.grown in  
Eggshell,NPK and cowdung  
Eggshell with cowdung  
Eggshell with NPK

Eggshell -Cow-dung-NPK

Eggshell-Cow-dung

Eggshell -NPK



## DISCUSSION

The present work was carried out to study growth performance of *Abelmoschus esculentus* grown in soil containing cow-dung, eggshell and NPK fertilizers. The study revealed that there was a considerable variation in growth and productivity of plants.

The study of Abdul maliq and Abayomi(2017) have reported that application of cow-dung as organic manure enhanced the growth and yield of okra when compared with inorganic fertilizer. Similarly the plant grown in cow-dung show considerable growth compared to the plants grown in NPK fertilizer.

According to the study of Premsekhar and Rajashree(2009) on influence of organic manures on growth, yield and quality of okra revealed that organic manure application improve physical and biological properties of soil resulting in better supply of nutrients to plants. Also the plants we planted on eggshell and cow-dung shows improved quality in okra production.

In the study of Vishaw Vikas and Amitesh sharma(2019) on application of organic manures and their influence on okra growth shows that there is a significant impact on the per fruit weight of okra, fruit, weight per plant, crop yield, seeds per fruit, total seed yield etc. The plants we grow on eggshell show significant increase on crop yield.

According to the study conducted by Savci S.(2012) An agriculture pollutant : chemical fertilizer. This study demonstrate the combined application of inorganic and organic manures has the potential to reduce chemical fertilizer usage without decreasing the yield of hybrid rice, and can enhance the growth , yield .

According to the study of Afe and Oluleye(2017) use of inorganic fertilizer influenced growth and fruit yield of okra also in our study the application of NPK influence the growth and fruit yield.

Based on the study of the estimation of total chlorophyll effects of nitrogen deficiency on photosynthetic traits of maize hybrids released in different years .in this study aims the new maize hybrids outperformed old one even at reduced N Rates .

Understanding the mechanisms of the difference in performance between newer and older hybrid under N deficiency ,and reduce environmental pollution caused by N fertilizers.

Nihort(2017)conducted the study about the effect of NPK fertilizer rates and method of application on growth and yield of okra at Ado-Ekiti Southwestern,nigeria the study was conducted to determine the effect of NPK fertilizer application on growth and yield of okra.

In the present work it was observed that the growth and development of *Abelmoschus esculentus* L. When treated with eggshell and cow dung was better compared to control and Npk. This efficiency may be due to the presence of the organic matter and minerals.

The biofertilizers enhance the growth and productivity of the plant. But the chemical fertilizer negative affect the growth and productivity of the plant. From the result of the present study, it may be concluded that biofertilizers like eggshell, cowdung help to enhance the growth and productivity of the plant.Such agricultural practices will also provide a proper health and economy. More studies are needed in this direction.

## SUMMARY AND CONCLUSION

The present work was carried out to study growth performance of *Abelmoschus esculentus* grown in soil containing cow-dung, eggshell and NPK fertilizers. The study revealed that there was a considerable variation in growth and productivity of plants.

The plant okra is germinated using water by kept it under water overnight. Then it become planted in different fertilizers. The natural and artificial fertilizers are taken for analysis. The natural fertilizers taken were egg shell, cow-dung and artificial fertilizer NPK are taken for the analysis. The plant growth observed is different in different fertilizers. The healthy plants are grown in the natural or organic fertilizers than the artificial medium.

The organic medium consisting of egg shell and cow-dung is good and far better than compared to NPK fertilizer because the egg shell contains calcium which moderates the soil acidity while providing nutrients and cow dung makes improve the fertility of the soil and increase the soil organic matter and enhances the microbial activity. By using NPK which directly effect the plant growth by feeding the plant, it often take longer time to helps the plants grow. And some of the plants that directly planted in high amount of NPK fertilizer cause wilting and destruction of plants. Hence natural fertilizer make the plant healthier than using artificial fertilizer. Also it is important to use organic manures in agriculture as they cause no problems to the environment.

## BIBLIOGRAPHY

Adepoju, A.Y., (1995). Comparative effects of the locally NPK compound fertilizer on maize yield. *Nig. J. Soil Sci.*, 11: 130-138.

Aliyu.L.(2000).Effect of organic and mineral fertilizers on growth yield and composition of pepper (*Capsicum annum L.*).*Biol.Agric.Hort.*18,(pp.29-36).

D. Arnon(1949).Estimation of total chlorophyll.*Plant Physiology*,24(1)1–15. 1949.

Ebaid and I. S.,(2007).Utilization of rice husk as an organic fertilizer to improve productivity and water use efficiency in rice fields. *African Crop Science Conference Proceeding*.8, pp.1923-1928.

Ezzo,M. I.,Gala.A. A.,Habib,H.A.M.,Helay,A. A. (2010).Response of sweet pepper grown in sandy and clay soil lysimeters to water regimes.*American-Eurasian journal of agricultural and enviornmental science*,8(1),18-26.

Froning,G.W.,D.,Bergquist.(1990).The effects of coffee and inedible eggshells and technical albumen using eggshells on fertilizer for red clover and barely extrusion technology.*Poultry Science*.69(205-252).

Kayode,S.A.,J.F.,(2015).Effect on growth and yield of okra in an ultisol.*Journal of agricultural studies*,4(1),73.

Kidd,P.S.,Proctor,J.(2001).Why plants grow poorly on very acid soils:are ecologists missing the obvious?..*Journal of Expermental Botany*,52(357),791-799.

King'ori,A.M,(2011).A Review of the Uses of Poultry Eggshells and Shell Membranes.*International Journal of Poultry science*.10(11),908-912.

Mani,S. and Ramanathan,M.K.(1980)Effect of nitrogen and phosphorus on yield of bhendi fruits.*South Indian Hort.*20,136-138.

Nardi,S.,Pizzeghello,D.,Muscolo,A.,&Vianello,A.(2002).

Nihort ( 1985). Effect of spacing and different rates of nitrogen fertilizer on seed yield of okra . *Nihort Vegetable Programme Annual Report*, 45-47.

Obi,M.E and Ebo,P.O.(1995).Effects of organic and inorganic amendments on soil physical properties and maize production in a severely degraded soil in southern Nigeria.*Bioresource Technol.*51,117-123.

Ogunlela ,V.b.,Ahmed,M.K.and Majanbu,I.S.(1989).Leaf characteristics and mineral element concentration of okra as influenced by fertilizer application,variety and plant age in the Nigerian savanna.*J.Agric.Trop.Subtrop.*90,127-135.

Omotoso,O.S.(2007).Effect of NPK fertilizer rates and method of Application on Growth and Yield of Okra.*International journal of Agricultural Research*,2(7),614-619.

Passioura,J.B.(1983).Roots and drought resistance.*Agricultural water management*,7(1-3),265-280.

Physiological effects of humic substances non higher plants.*Soil Biology and Biochemistry*.

S. Savci.(2012).An agricultural pollutant: chemical fertilizer.*International Journal of Environmental Science and Development*, 3(1),77–80.

**ST. TERESA'S COLLEGE**  
**(AUTONOMOUS)**  
**AFFILIATED TO MAHATMA GANDHI UNIVERSITY**



**FIX IN TO DO**  
**PROJECT REPORT**

In partial fulfilment of the requirements for the award of the degree of

**BACHELOR OF SCIENCE IN**  
**COMPUTER APPLICATIONS**  
**[TRIPLE MAIN]**

By

**Treesa Preeshma T P**

**III B.Sc. Computer Applications [Triple main]**

**Register No: SB19CA028**

**Under the guidance of**

**Mrs. Harsha K M**

**DEPARTMENT OF COMPUTER APPLICATIONS**

**2019-2022**



## CERTIFICATE

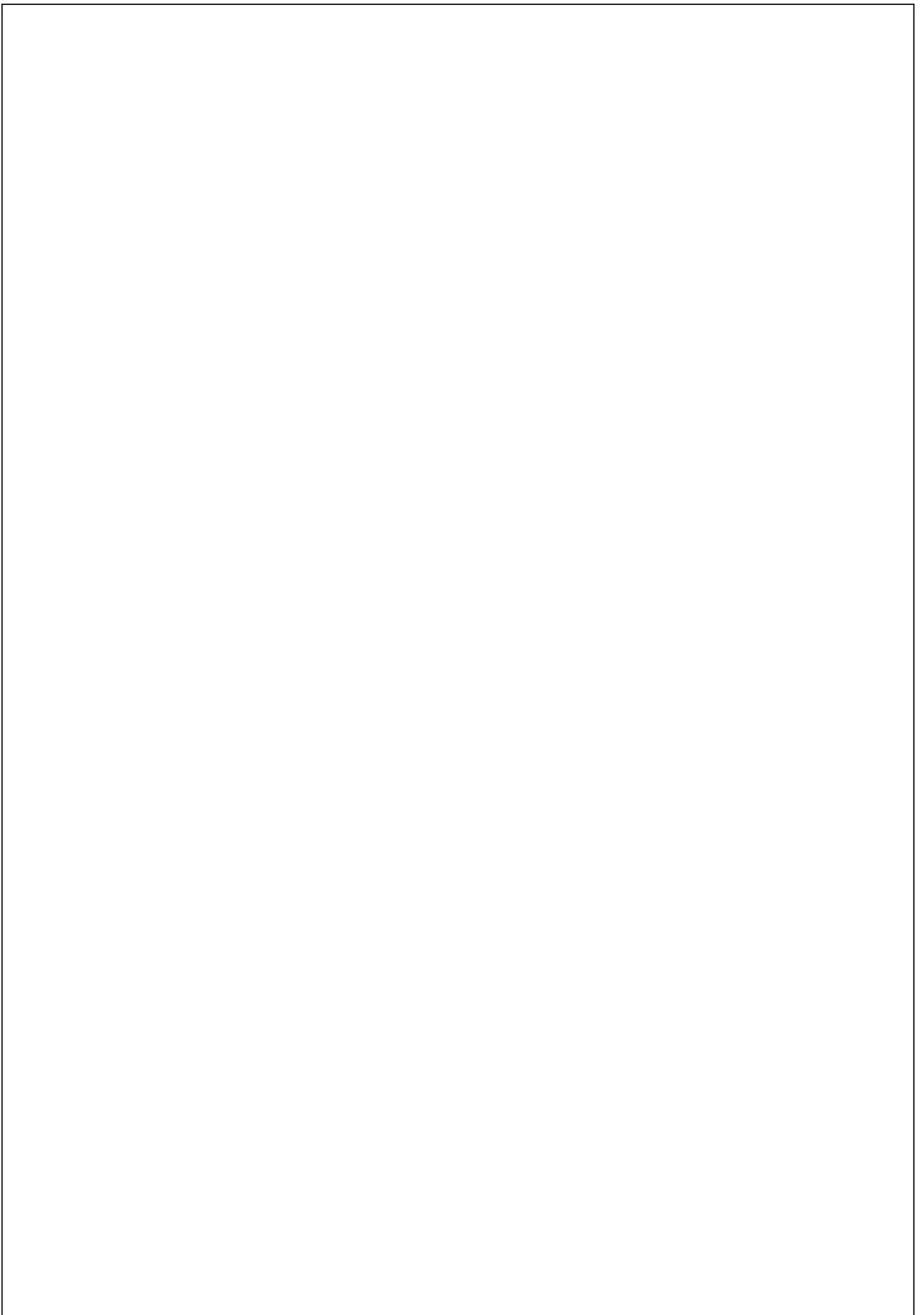
This is to certify that **Ms. TREESAPREESHMA T P.** (Reg. No: SB19CA028) Bachelor of Science In Computer Applications (Triple main) VI semester student of **ST. TERESA'S COLLEGE ERNAKULAM** affiliated to Mahatma Gandhi university, has done project work entitled "**FIX IN TO DO**" in PYTHON+ANDROID under the guidance of our senior faculties towards the fulfillment of the award of "Bachelor of Science In Computer Applications (Triple Main)" during the period of October 2021 to March 2022.

She successfully completed the project and during the period she was methodical and hardworking.

For **RISS TECHNOLOGIES**

Chief Executive Officer



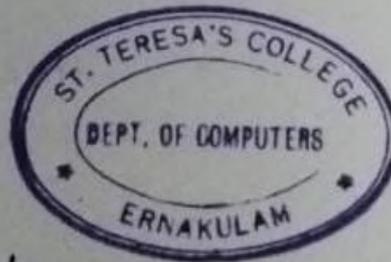




## CERTIFICATE

This is to certify that the project report entitled "FIX IN TO DO ", a bona fide record of the work done by TREESA PREESHMA T P during the year 2021-22 and submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Applications (Triple main) under Mahatma Gandhi University.

Head of the Department



Internal Examiner :

Date: 06/04/2022

External Examiner :

## **DECLARATION**

I, Treesa Preeshma T P , BSc Computer Application [Triple main] final year student of St. Teresa's College (Autonomous), Ernakulam, Register No. SB19CA028, hereby declare that the dissertation submitted for the Bachelor's Degree in Computer Applications is my original work. I further declare that the said work has not previously been submitted to any other university or academic body.

Date:

Treesa Preeshma T P

Place: Ernakulam

# ACKNOWLEDGEMENT

First and foremost I thank God Almighty for his blessings. I take this opportunity to express my gratitude to all those who helped me in completing this project successfully. I wish to express my sincere gratitude to the Manager **Rev. Dr. Sr. Vinitha CSST**, the Director **Rev. Sr Emeline CSST** and the Principal **Dr. Lizzy Mathew** for providing all the facilities.

I express my sincere gratitude towards the Head of the Department **Mrs. Raji S Pillai** and the Course Coordinator **Mrs. Sheeba Emmanuel** for the support. I deeply express my sincere thanks to my guide **Mrs. Harsha K M** for her proper guidance and support throughout the project work. I am highly indebted to my external guide at RISS Technologies, **Mr. Akash** for training me well to develop this project up to my level of expectation. I thank him for giving me such attention and time despite the hectic schedule.

My deepest appreciation goes to my beloved teachers whose cooperation and suggestion throughout the project helped me a lot. I thank all my friends and classmates for their support. I convey my hearty thanks to my parents for their moral support, suggestion and encouragement.

# SYNOPSIS

There has been a growing need of handyman services in most countries around the world which is attributed to several contributing factors which include, local demand, market influence, having second homes, commercial property maintenance and people's lives becoming busier and more hectic with more people now looking for help with odd-jobs around their home, like changing light fittings, plumbing, painting, house made, putting up shelves. The existing handyman services solutions striving to solve this problem are defragmented offering contacts scattered in the web and there is no structured method of determining the Location and quality of service provided by these handymen.

The proposed model is a mobile application for locating handyman services within a locality to help in streamlining this process and provide a structured approach for determining location and quality of service to be provided by the handymen. The application was developed in android operating system because of its popularity among many mobile users.



## TABLE OF CONTENTS

<b>1. INTRODUCTION .....</b>	<b>2</b>
1.1 ABOUT PROJECT .....	3
1.2 ABOUT ORGANIZATION .....	3
1.3 OBJECTIVES OF THE PROJECT AND THE ORGANIZATION .....	3
<b>2. SYSTEM ANALYSIS .....</b>	<b>4</b>
2.1 INTRODUCTION .....	5
2.2 EXISTING SYSTEM .....	5
2.3 PROPOSED SYSTEM.....	5
2.4 SYSTEM SPECIFICATION .....	5
2.5 OPERATING SYSTEM .....	5
2.6 LANGUAGES AND SOFTWARE PACKAGES.....	6
2.7 HARDWARE AND SOFTWARE SPECIFICATIONS .....	7
<b>3. SYSTEM DESIGN.....</b>	<b>9</b>
3.1 INTRODUCTION .....	10
3.2 DATA FLOW DIAGRAM.....	10
3.3 DATA DICTIONARY.....	13
3.4 DATABASE DESIGN .....	13
<b>4. SYSTEM DEVELOPMENT.....</b>	<b>18</b>
4.1 INTRODUCTION .....	19
4.2 PROCESS DESCRIPTION.....	19
4.3 CODE DESIGN .....	20
<b>5. SYSTEM TESTING AND IMPLEMENTATION .....</b>	<b>25</b>
5.1 INTRODUCTION .....	26
5.2 IMPLEMENTATION .....	26
5.3 DEBUGGING .....	26
5.4 SYSTEM SECURITY.....	27
5.5 SCOPE FOR FUTURE ENHANCEMENT .....	27
<b>6. CONCLUSION .....</b>	<b>28</b>
<b>7. APPENDIX .....</b>	<b>30</b>
<b>8. BIBLIOGRAPHY.....</b>	<b>41</b>
<i>Reference Website.....</i>	<i>41</i>



# 1. INTRODUCTION



## **1.1 About Project**

The Fix in To Do application can help the customer to get the service without going anywhere .the search engine provides an easy and convenient to search and select to find the nearby handyman workers and make the task easier for the customers .This application serve the list of benefits to not only one or two industries but many such as plumbing , painting, electricians ,carpentry ,car care and many more.....

## **1.2 About Organization**

RISS TECHNOLOGIES is a rapidly growing company that provides professional IT services. They are one of the largest and Best software development companies in Kerala with focus on .Net, PHP, Java, Software testing, SEO and Web Design

## **1.3 Objectives of the Project and the Organization**

The main objective of the project is to make searching , viewing, selecting of a workers like painting ,plumbing, car care, house helper (handyman) for the user needs ,over internet to make task easier.

The main objective of the organisation is to continuously optimize their customers' business through our world-class solutions; services and products. They ensure the success of the company by constantly and consistently satisfying the customers, shareholders and employees.



## 2. SYSTEM ANALYSIS SYSTEM ANALYSIS



## **2.1 Introduction**

System Analysis is the complete study of the system and identifying its objectives mainly for problem solving purposes. Each and every modules of the system are evaluated. Inferences are made from these studies to ensure that all the components of the system is working efficiently.

System Analysis involves gathering information related to the system and developing the accurate tools for analysis. Studying and analysing the existing system is important for system analysis. Identifying the drawbacks in the existing system and how it is been rectified in the proposed system is one of the main aim.

## **2.2 Existing System**

Before the implementation of proposed system another system is their existence called existing system. in earlier finding of a handyman worker was a big task and ne need to go out and find them for our needs and sometimes we should search them for days to make the service

## **2.3 Proposed System**

The Fix in To Do application can help the customer to get the service without going anywhere .the search engine provides an easy and convenient to search and select to find the nearby handyman workers and make the task easier for the customers .This application serve the list of benefits to not only one or two industries but many such as plumbing , painting, electricians ,carpentry ,car care and many more.....

## **2.4 System Specification**

System specification specifies the hardware and software configuration of the new system. It helps to define the operational and performance guidelines of the system.

## **2.5 Operating System**

An Operating System (OS) is an interface between computer user and computer hardware. It is a software which performs all the basic tasks like file management, memory management, process management, handling input and output and controlling peripheral devices such as disk drives and printers. The operating system required for proper execution of the system is Windows 10 or above. System specifications for the app to run:

## **2.6 Languages and Software Packages**

### ✦ Python

Python is a widely used high-level programming language for general-purpose programming, created by Guido van Rossum and first released in 1991. An interpreted language, Python has a design philosophy that emphasizes code readability and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++ or Java. The language provides constructs intended to enable writing clear programs on both a small and large scale. Python features a dynamic type system and automatic memory management and supports multiple programming paradigms, including object-oriented, imperative, functional programming, and procedural styles. Python interpreters are available for many operating systems, allowing Python code to run on a wide variety of systems.

### ✦ MySQL

MySQL is an open-source relational database management system (RDBMS); it is the world's second most widely used RDBMS, and the most widely used open-source client-server model RDBMS. The SQL acronym stands for Structured Query Language. The MySQL development project has made its source code available under the terms of the GNU General Public License, as well as under a variety of proprietary agreements. MySQL is a popular choice of database for using web applications, and is a central component of the widely used LAMP open-source web application software stack (and other "AMP" stacks). LAMP is an acronym for "Linux, Apache, MySQL, Pearl, PHP, and Python ". Free-software open-source projects that require a full-featured database management system often use MySQL.

### ✦ HTML

**Hypertext Markup Language**, commonly referred to as HTML, is the standard markup language used to create web pages. Along with CSS, and JavaScript, HTML is a cornerstone technology used to create web pages, as well as to create user interfaces for mobile and web applications. Web browsers can read HTML files and render them into visible or audible web pages. HTML describes the structure of a website semantically along with cues for presentation, making it a markup language, rather than a programming language.

HTML can embed scripts written in languages such as JavaScript which affect the behavior of HTML web pages. HTML markup can also refer the browser to Cascading Style Sheets (CSS) to define the look and layout of text and other material.

✦ **ANDROID**

Android is a mobile operating system developed by Google, based on the Linux kernel and designed primarily for touchscreen mobile devices such as smartphones and tablets. Android's user interface is mainly based on direct manipulation, using touch gestures that loosely correspond to real-world actions, such as swiping, tapping and pinching, to manipulate on- screen objects, along with a virtual keyboard for text input. In addition to touchscreen devices, Google has further developed Android TV for televisions, Android Auto for cars, and Android Wear for wrist watches, each with a specialized user interface. Initially developed by Android Inc., which Google bought in 2005, Android was unveiled in 2007, along with the founding of the Open Handset Alliance – a consortium of hardware, software, and telecommunication companies devoted to advancing open standards for mobile devices.

## **2.7 Hardware and Software Specifications**

### **Software Requirements**

A software requirement specification (SRS), a requirements specification for a software system, is a complete description of the behaviour of a system to be developed and may include a set of use cases that describe interactions the users will have with the software. In addition it also contains non-functional requirements. Non-functional requirements impose constraints on the design or implementation (such as performance engineering requirements, quality standards, or design constraints) the software requirements specification document enlists all necessary requirements that are required for the project development. To derive the requirements we need to have clear and thorough understanding of the products to be developed. This is prepared after detailed communications with the project team and customer.

Operating System: Windows 10 or above

Front End: HTML,Python(for web application) andANDROID(for mobile application)

Back End: MySQL

Software:sublime Text, WAMP,Android studio

Web Browser:Internet Explorer/Google Chrome/Firefox

Developing tool: Python

### **Hardware Requirements**

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. A hardware requirements list is often accompanied by a hardware compatibility list (HCL), especially in case of operating systems. An HCL lists tested, compatible, and sometimes incompatible hardware devices for a particular operating system or application.

Processor: Intel Pentium or above

RAM: 8 GB or above

Hard Disk: 145GB

## **3. SYSTEM DESIGN**



### **3.1 Introduction**

It is a process of planning a new business system or replacing an existing system by defining its components or modules to satisfy the specific requirements. Mainly focuses on how to accomplish the objectives of the system.

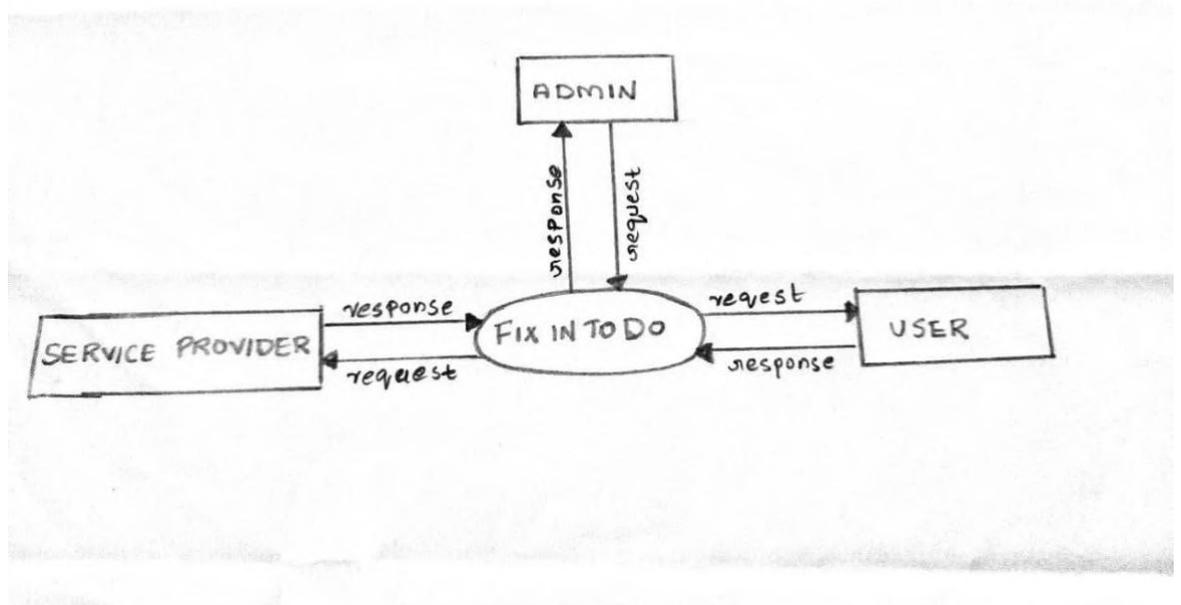
### **3.2 Data Flow Diagram**

A data flow diagram (DFD) is a graphical representation of the flow of data through an information system. A DFD is often used as a primary step to create an overview of the system, which can later be elaborated.

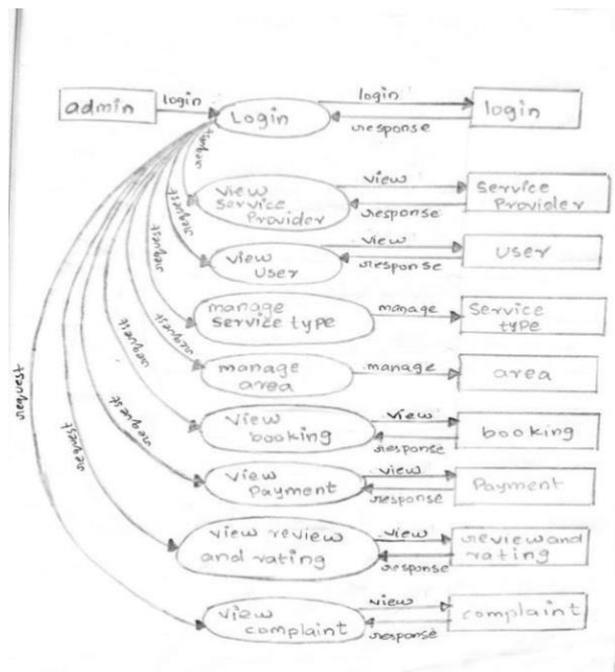
A DFD shows what will be the input of the system as well as the output. It clearly represents where the data will come from and go to, and where the data will be stored.



**DFD level Zero**

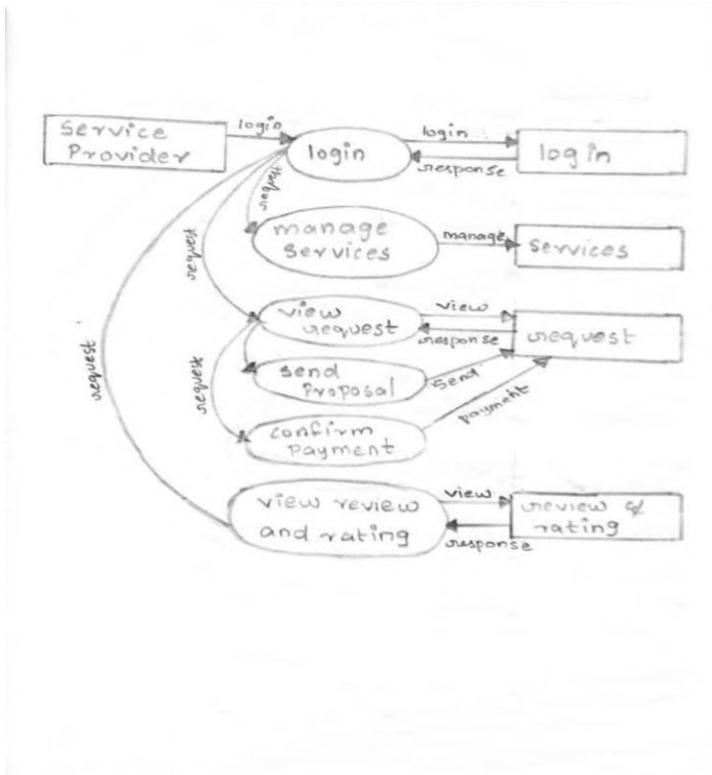


**DFD Level One**

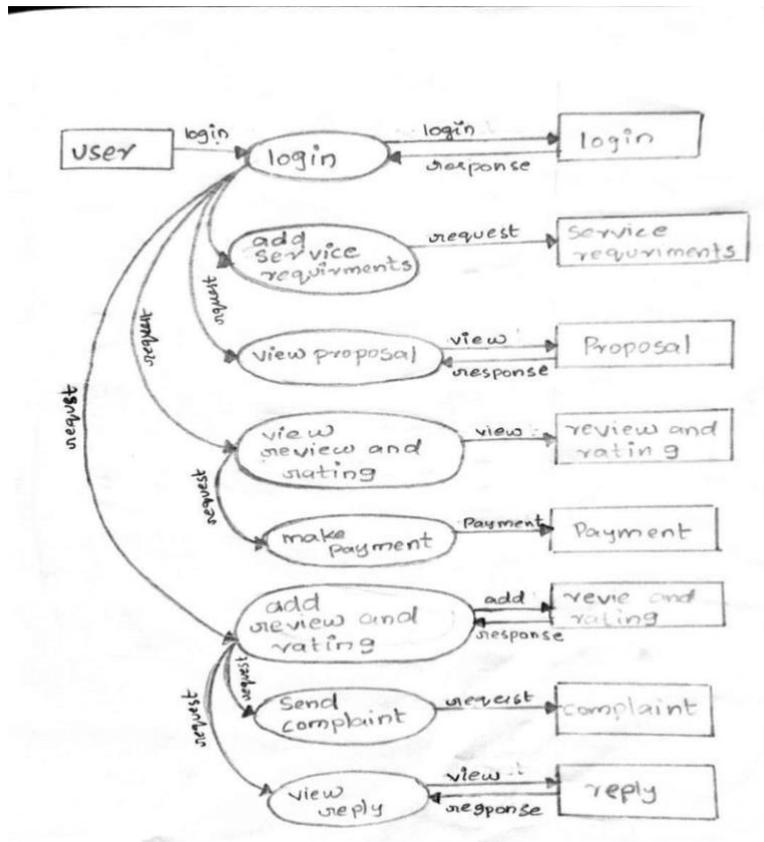




**DFD Level One**



**DFD Level Two**





### 3.3 Data Dictionary

A data dictionary contains metadata. The data dictionary is very important as it contains information such as what is in the database, who is allowed to access it, where is the database physically stored etc. The users of the database normally don't interact with the data dictionary, it is only handled by the database administrators.

### 3.4 Database Design

**Database Design** is a collection of processes. The main aim of database designing is to produce logical and physical design models for the suggested database system.

The logical model focuses on the data requirements and the data to be stored independent of physical components.

The physical data design model translates the logical design of the database onto physical media using hardware resources and software systems.

#### **Login**

Column	Type	Null	Default
login_id	Int(11)	No	
username	Varchar(55)	Yes	
password	Varchar(55)	Yes	
usertype	varchar(55)	Yes	

#### **Service provider**

Column	Type	Null	Default
Provider_id	int(11)	No	



Login_id	int(11)	Yes	
Fname	varchar(55)	Yes	
Lname	varchar(55)	Yes	
Area_id	int(11)	Yes	
Place	varchar(55)	Yes	
Pincode	varchar(55)	Yes	
phone	varchar(55)	Yes	
email	varchar(55)	Yes	

## User

Column	Type	Null	Default
User_id	int(11)	No	
login_id	int(11)	Yes	
Area_id	Int(11)	Yes	
Fname	Varchar(55)	Yes	
Lname	Varchar(55)	Yes	
House_name	Varchar(55)	Yes	
Place	Varchar(55)	Yes	
pincode	Varchar(55)	Yes	
District	Varchar(55)	Yes	
phone	Varchar(55)	Yes	
email	Varchar(55)	Yes	

## Service type

Column	Type	Null	Default
Servicer_type_id	int(11)	No	
Type_name	varchar(55)	Yes	
description	varchar(55)	Yes	

**Provider service**

Column	Type	Null	Default
Provider_service_id	int(11)	No	
Service_type_id	int(11)	Yes	
Provider_id	int(11)	Yes	

**Area**

Column	Type	Null	Default
area_id	int(11)	No	
Area_name	varchar(55)	Yse	
Area_description	varchar(55)	Yes	

**Booking**

Column	Type	Null	Default
Booking_id	Int(11)	No	
User_id	Int(11)	Yes	
Service_type_id	Int(11)	Yes	
Provider_id	Int(11)	Yes	
Booking_date	Varchar(55)	Yes	
Booking_description	Varchar(55)	Yes	
Work_image_1	Varchar(55)	Yes	
Work_imager_2	Varchar(55)	Yes	
Booking_status	Varchar(5)	Yes	

## Proposal

Column	Type	Null	Default
Proposal_id	Int(11)	No	
Booking_id	Int(11)	Yes	
Estimate_amount	Int(11)	Yes	
Provider_id	Int(11)	Yes	
Proposal_desc	Varchar(55)	Yes	
Proposal_date	Varchar(55)	Yes	
Proposal_status	Varchar(55)	Yes	

## Payment

Column	Type	Null	Default
Pay_id	Int(11)	No	
Proposal_id	Int(11)	Yes	
Amount_paid	Varchar(55)	Yes	
Pay_description	Varchar(55)	Yes	
Pay_date	Varchar(55)	Yes	
Pay_type	Varchar(55)	Yes	

## Reviews

Column	Type	Null	Default
Review_id	Int(11)	No	
User_id	Int(11)	Yes	
Provider_id	Int(11)	Yes	
Review_desc	Varchar(55)	Yes	
Rating	Varchar(55)	Yes	
Review_date	Varchar(55)	Yes	

## Complaint

Column	Type	Null	Default
Compl_id	Int(11)	No	
User_id	Int(11)	Yes	
Comp_desc	Varchar(55)	Yes	
Reply_desc	Varchar(55)	Yes	
Comp_date	Varchar(55)	Yes	

## **4. SYSTEM DEVELOPMENT**



## **4.1 Introduction**

Software Development is the process of analysing, designing, testing, implementation and maintenance. It is called Software Development Life Cycle (SDLC). Different SDLC include waterfall, prototyping, iterative, incremental, spiral development, rapid application development and agile methodology.

## **4.2 Process Description**

There are three modules :

### **Admin**

- Login
- View service provider
- View users
- Manage service types
- Manage areas
- View bookings
- View payment
- View review and rating
- View complaint ○ send reply

### **Service Provider**

- Register
- Login
- Manage services
- View requests ○ Send proposal ○ Confirm payment • View review and rating

### **Users**

- Register
- Login
- Add a service requirement ○ View proposals ○ View review and rating ○ Make payment
- Add review and rating
- Send complaint ○ View reply

### **4.3 Code Design**

**Main.py:**

```
from flask import *

from public import public from admin import admin from provider import provider

from api import api

app=Flask(_name_)

app.secret_key="jmm"

app.register_blueprint(public) app.register_blueprint(admin,url_prefix='/admin')

app.register_blueprint(provider,url_prefix='/provider') app.register_blueprint(api,url_prefix='/api')

app.run(debug=True, host="192.168.23.53", port=5020)

# app.run(debug=True, port=5020)
```

**LOGIN:**

```
package com.example.basicservices;

import org.json.JSONArray; import org.json.JSONObject;

import android.os.Bundle; import android.preference.PreferenceManager; import android.app.Activity; import
    android.content.Intent; import android.content.SharedPreferences; import
    android.content.SharedPreferences.Editor
;
import android.util.Log; import android.view.Menu; import android.view.View; import android.widget.Button;
import android.widget.EditText; import android.widget.TextView; import android.widget.Toast;

public class Login extends Activity implements JsonResponse {
    EditText e1,e2;

    Button b1;
    String username,password;

    TextView t1;
    SharedPreferences sh;
```

@Override protected void onCreate(Bundle

```

savedInstanceState) {

                                                                    super.onCreate(savedInstanceState);

                                                                    setContentView(R.layout.activity

_login);

e1=(EditText)findViewById(R.id.uname)
;

e2=(EditText)findViewById(R.id.pass);

b1=(Button)findViewById(R.id.login);

sh=PreferenceManager.getDefaultSharedPreferences(getApplicationContext());

t1=(TextView)findViewById(R.id.textView2);
    t1.setOnClickListener(new View.OnClickListener() {

@Override
                                                                    public void
onClick(View arg0) {
                                                                    // TODO
Auto-generated method stub

startActivity(new

```

Intent(getApplicationContext()),

```

Registration.class));
                                }
                                });

b1.setOnClickListener(new View.OnClickListener() {

@Override
                                public void
onClick(View arg0) {
                                // TODO
Auto-generated method stub

username=e1.getText().toString();

                                password=e2.getText().toString();

if(username.equalsIgnoreCase(""))
){

                                Toast.makeText(getApplicationContext(), "Enter All The Fields",
Toast.LENGTH_LONG).show();

                                e1.setError("");

                                e1.setFocusable(true); } else

if(password.equalsIgnoreCase("")){

Toast.makeText(getApplicationContextC

```

```
ontext(), "Enter All The Fields", Toast.LENGTH_LONG).show();
```

```
        e2.setError("");

        e2.setFocusable(true);
    }

else { JsonRequest jr=
new JsonRequest();

jr.json_response=(JsonResponse)
Login.this;

String
q="/login/?username="+username+"&password="+password;

        q.replace(" ", "%20");

        jr.execute(q);

    }
}
});
}
```



## **5. SYSTEM TESTING AND IMPLEMENTATION**



## **5.1 Introduction**

Software testing is defined as a process to check whether the actual results match the expected results and to ensure that the software system is error free. Software testing also helps to identify defects, gaps or missing requirements in contrary to the actual requirements. It can be done manually or using automated tools.

## **5.2 Implementation**

Implementation is the action that must follow any preliminary thinking in order for something to actually happen. Software/hardware implementations should always be designed with the end user in mind and the implementation process usually benefits from user involvement and support from managers and other top executives in the company. If users participate in the design and implementation of the system, ideally it will serve their business objectives more accurately and reflect their priorities and the ways in which they prefer to work.

## **5.3 Debugging**

Debugging is the process of finding and resolving defects or problems within a system that prevent the proper functioning of the system. Different types of debugging methods used in this system are:

- **Unit Testing**

It is the first level of testing. Each module is tested individually and focus is given for finding errors limited to each individual module and correcting them. The different modules of the system are tested individually and corrected all errors. Each module is focused to work satisfactorily with regard to the expected output from the module. Validation checks for fields are also done here. Each process was done individually and tested separately. Errors discovered were corrected. After unit testing, modules were integrated to form the complete system.

- **Integration Testing**

The tested modules are combined into subsystems and these are tested again. It is the second-level of testing. When modules are integrated, problems can arise at boundaries like incompatibility of data type of value being passed across the interface or some unexpected problems which appeared only after integration. Integration test was carried out while each module was integrated, interfaces were tested and corrected errors.

- **System Testing**

The tested modules are combined into subsystems and these are tested again. It is the second-level of testing. When modules are integrated, problems can arise at boundaries like incompatibility of data type of value being passed across the interface or some unexpected problems which appeared only after integration. Integration test was carried out while each module was integrated, interfaces were tested and corrected errors.

- **Validation Testing**

For each input forms validation testing are done to ensure that only allowed values will be entered. Entering incorrect values does the validation testing and it is checked whether the errors are being considered. Incorrect values are to be discarded. The errors are rectified.

## **5.4 System Security**

Password encryption is used to protect each user's details.

## **5.5 Scope for Future Enhancement**

The current system is flexible and can be modified in the future.

## **6. CONCLUSION**



## **COCLUSION**

The software developed has fulfilled the necessary requirements as required by the user. It is ensured that all the programs are working properly in the “Fix in to do” application. The application used to reduce the human effort to make task easier, it helps us to search the handyman workers for the users need, through internet The system is used to operate in a user- friendly manner. Proper documentation done from different areas provides smooth running of all the operations without difficulty. The system that has been developed in android studio is to improve the user interactivity. This project avoids errors.. The project has been implemented and tested.

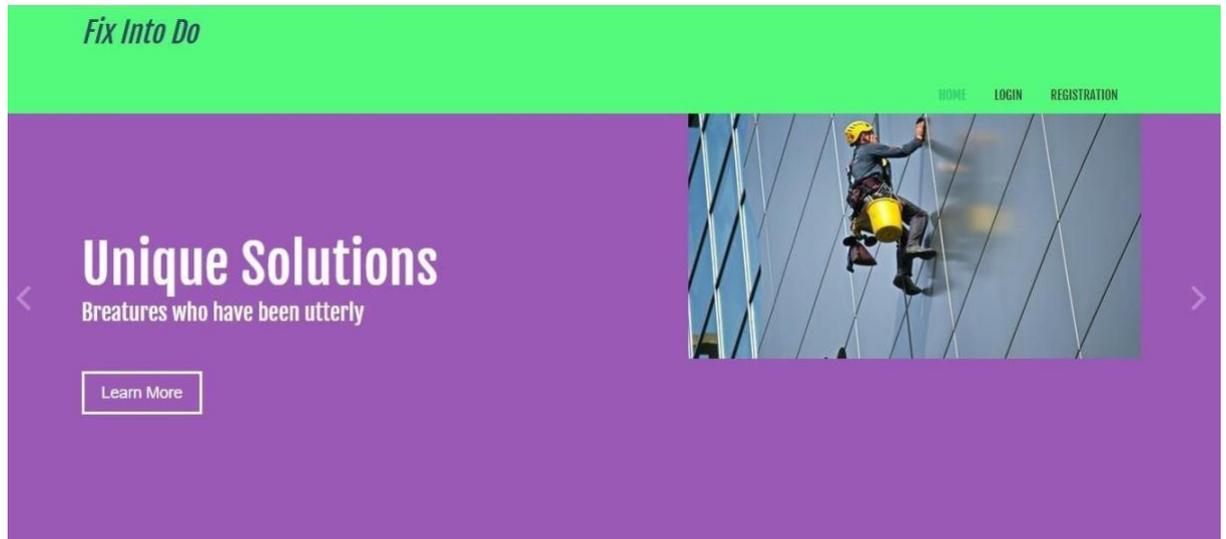


## **7. APPENDIX**



## **Interface**

## HOME PAGE



## LOGIN PAGE

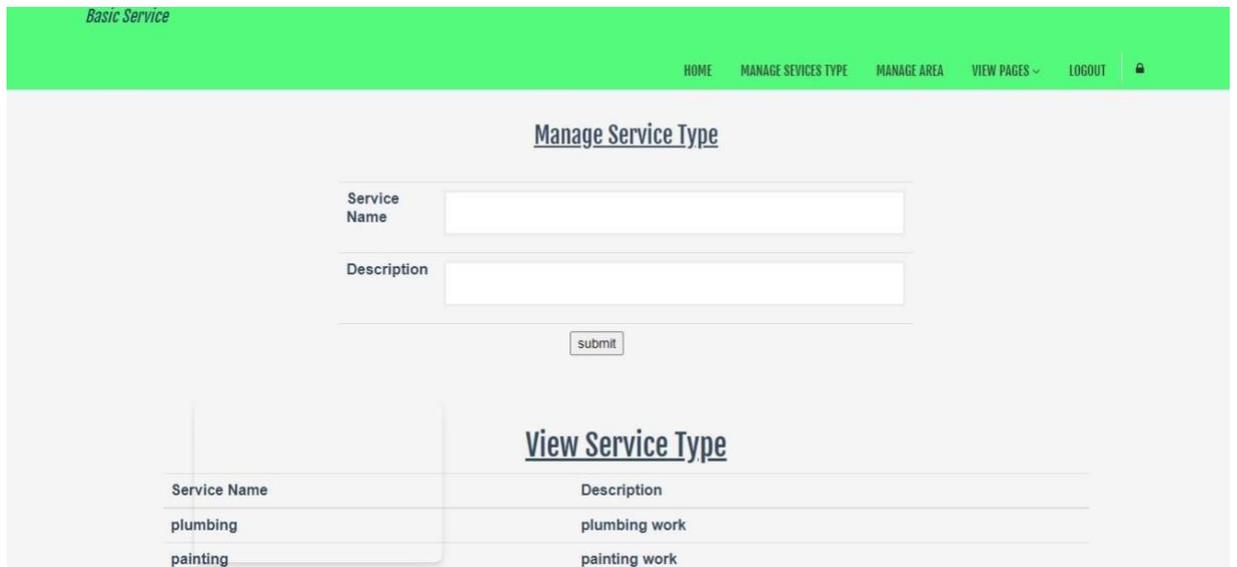




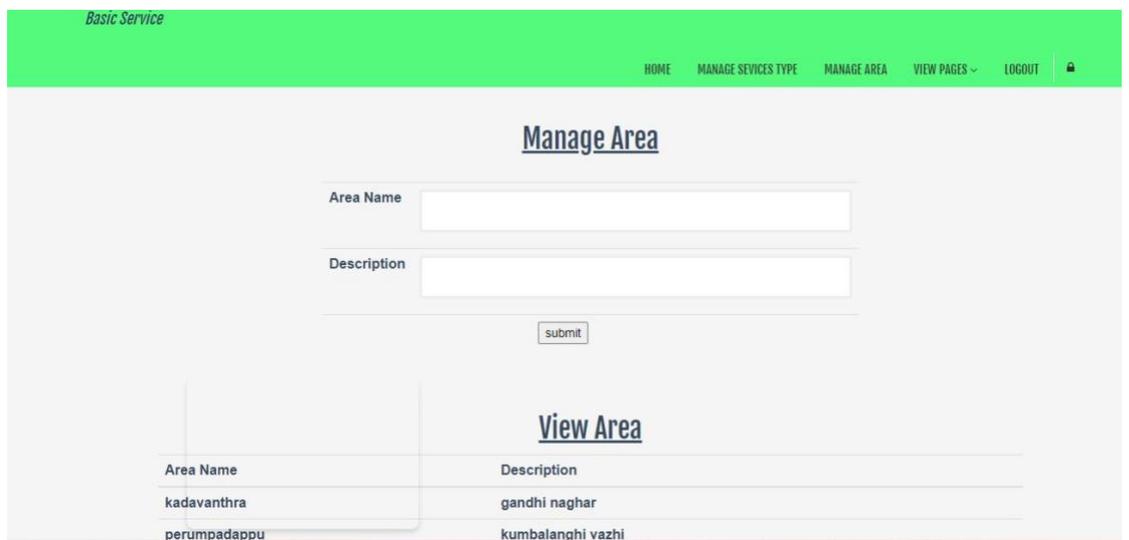
## **LOGIN FORM**



### ADMIN: MANAGE SERVICE TYPE



### ADMIN: MANAGE AREA





**ADMIN: VIEW SERVICE PROVIDER**

Basic Service

HOME MANAGE SERVICES TYPE MANAGE AREA VIEW PAGES LOGOUT

### View Service Provider

Name	Area Name	Place	Pincode	Phone	Email
arun raj	thevara	ernakulam	682025	9562074945	arun@gmail.com
sajin sakeer	fortkochi	ernakulam	682021	9824576379	sajin@gmail.com
jefin jacob	kannamali	ernakulam	682008	9048136273	jefin@gmail.com

**ADMIN: VIEW USER**

Basic Service

HOME MANAGE SERVICES TYPE MANAGE AREA VIEW PAGES LOGOUT

### View User

Name	Area Name	House Name	Place	Pincode	District	Phone	Email
Preeshma T P	kadavanthra	thekkepadath house	vathuruthy	682029	ernakulam	8943012635	tpreeshma@gmail.com
ashika tinu	kadavanthra	palliparambil house	kadavanthra	682020	ernakulam	9400158377	ashikatinu@gmail.com
philomina sabu	perumpadappu	palliparambil house	perumpadappu	682020	ernakulam	7591923659	philominasabu@gmail.com
Lefiya P.F	kadavanthra	palliparambil house	kannamaly	682008	ernakulam	8281519621	lefiyapf@gmail.com

**ADMIN: VIEW BOOKING**

Basic Service

HOME MANAGE SERVICES TYPE MANAGE AREA VIEW PAGES LOGOUT

### View Booking

User name	Service Provider Name	Date	Description	Image1	Image2
-----------	-----------------------	------	-------------	--------	--------



**ADMIN: VIEW PAYMENT**

*Basic Service*

HOME    MANAGE SERVICES TYPE    MANAGE AREA    VIEW PAGES ~    LOGOUT    🔒

### View Payment

User Name	Proposal description	amount	payment description	Date	pay_type
philomina sabu	plumbing	250	offline	20/4/2022	cash

**ADMIN: VIEW REVIEW AND RATING**

*Basic Service*

HOME    MANAGE SERVICES TYPE    MANAGE AREA    VIEW PAGES ~    LOGOUT    🔒

### View Review and Rating

User Name	Service Provider Name	Description	Rating	Date
Preeshma T P	arun raj	good	7.5	2/3/2022

**ADMIN: VIEW COMPLAINT**

*Basic Service*

HOME    MANAGE SERVICES TYPE    MANAGE AREA    VIEW PAGES ~    LOGOUT    🔒

### View Complaint

User Name	Description	Date	reply
Preeshma T P	not perfect	2022-03-02	<input type="text"/> <input type="button" value="Submit"/>
philomina sabu	slower	2022-04-01	<input type="text"/> <input type="button" value="Submit"/>
ashika tinu	nill	2022-04-01	<input type="text"/> <input type="button" value="Submit"/>



**SERVICE PROVIDER: REGISTRATION**

*Fix Into Do* HOME LOGIN REGISTRATION

### Registration

First Name

Last Name

Area Name

Place

Pincode

Phone

aadhar

Email

UserName

Password

**SERVICE PROVIDER: ADD SERVICE TYPE**

*Basic Service* HOME MANAGE SERVICES VIEW REQUEST VIEW MY PAYMENTS VIEW REVIEW AND RATING LOGOUT

### Add Service Type

Service type Name

### View Service Type

Sl.no	Service type Name
1	plumbing
2	painting



**SERVICE PROVIDER: VIEW REQUEST**

*Basic Service*

HOME    MANAGE SERVICES    VIEW REQUEST    VIEW MY PAYMENTS    VIEW REVIEW AND RATING    LOGOUT

### View Request

Sl.no	User Name	Sefrvce	Date	Description	Booking Status	
1		plumbing	2022-04-01	plumbing	pending	<a href="#">Add Proposal</a>
2		painting	2022-04-01	painting	pending	<a href="#">Add Proposal</a>
3		plumbing	2022-04-02	plumbing	accept	
4		plumbing	2022-04-04	plumbing	pending	<a href="#">Add Proposal</a>
5		plumbing	2022-04-04	plumbing	pending	<a href="#">Add Proposal</a>

### SERVICE PROVIDER: VIEW PAYMENT

*Basic Service*

HOME    MANAGE SERVICES    VIEW REQUEST    VIEW MY PAYMENTS    VIEW REVIEW AND RATING    LOGOUT

### VIEW PAYMENT

user name	Proposal description	amount	Payment Description	Date
	plumbing	250	offline	20/4/2022

### SERVICE PROVIDER: VIEW TRVIWE AND RATING

*Basic Service*

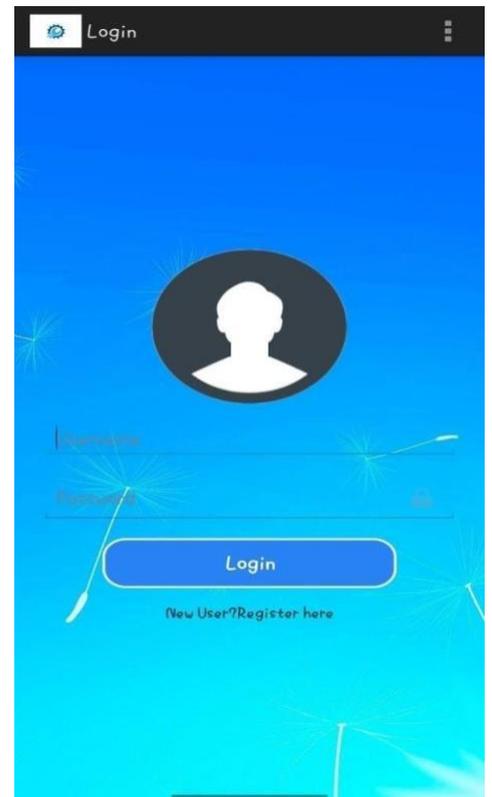
HOME    MANAGE SERVICES    VIEW REQUEST    VIEW MY PAYMENTS    VIEW REVIEW AND RATING    LOGOUT

### View Review And Rating

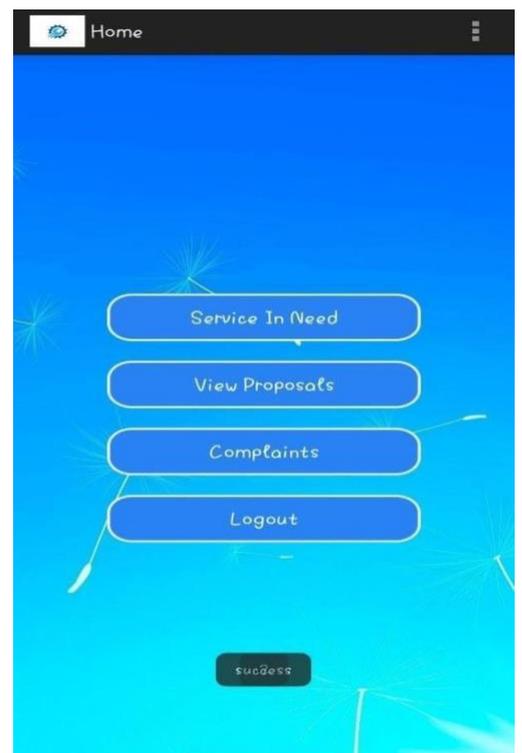
User Name	Description	Rating	Date
Preeshma T P	good	7.5	2/3/2022



**USER:LOGIN FORM**

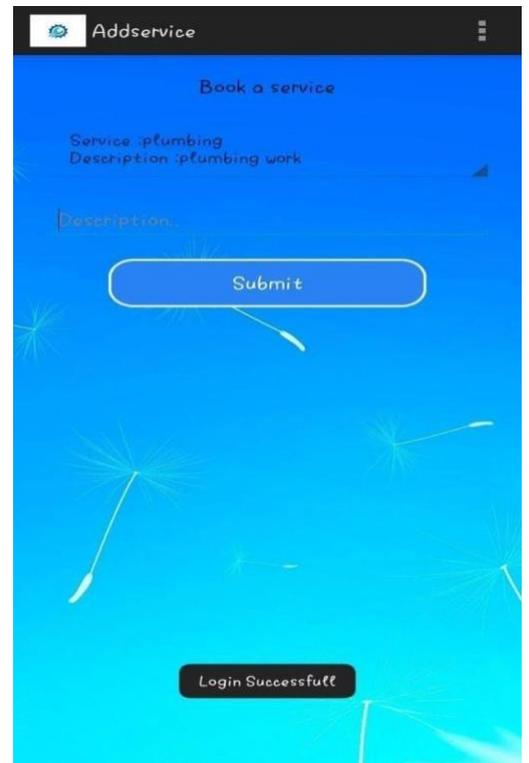


### USER:LOGIN BOOKING

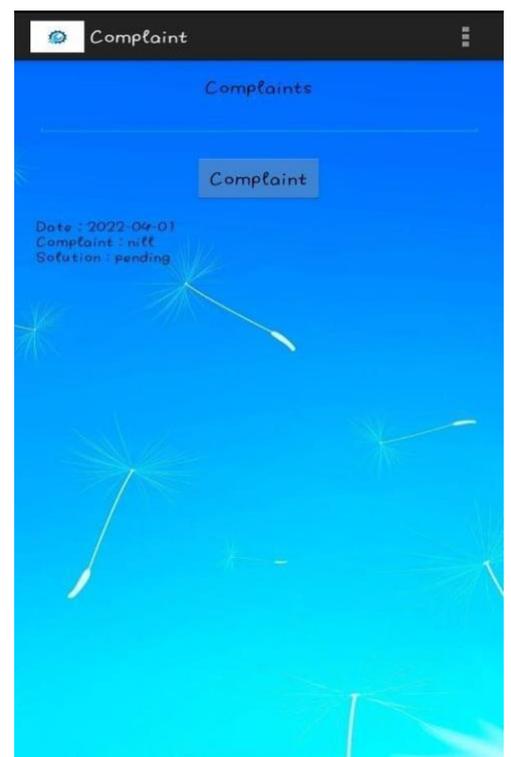




**USER: BOOKIG A SERVICE**



**USER: COMPLAINT**

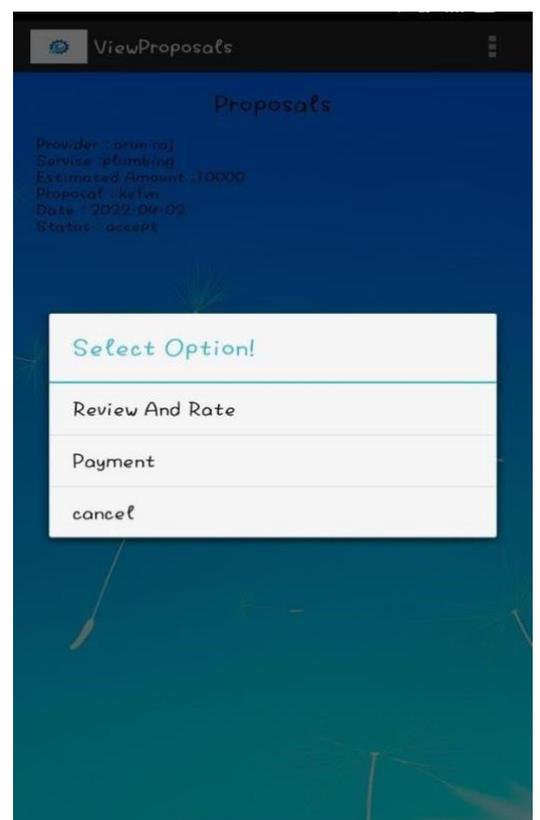




**USER: PROPOSAL**

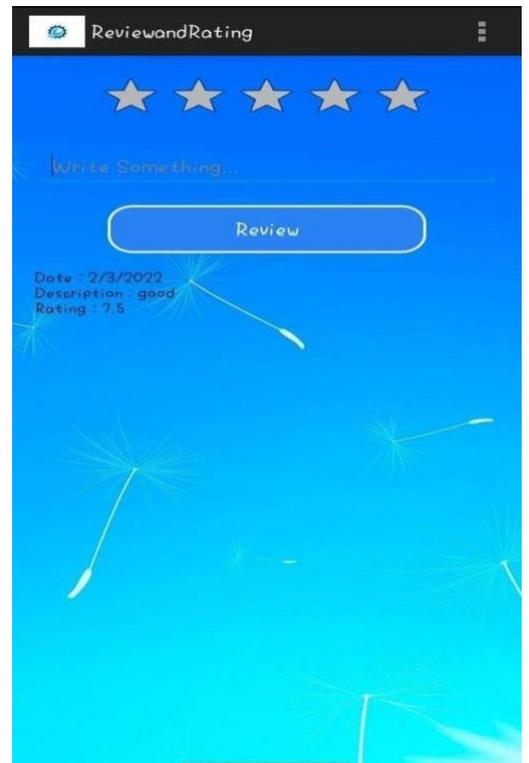


**USER: SELECT OPTION**

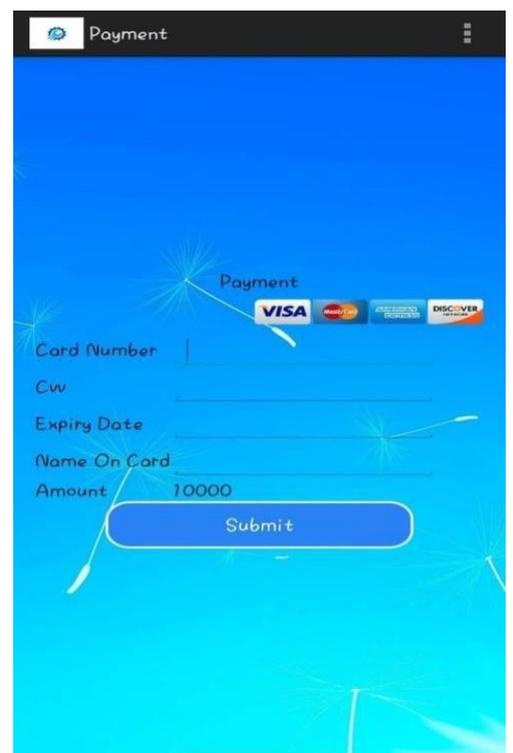




**USER: REVIEW AND RATING**



## USER:PAYMENT





## **8. BIBLIOGRAPHY**

## **Reference Website**

- <https://stackoverflow.com/>
- <https://www.tutorialspoint.com/>
- <https://www.technotalkative.com/>
- <https://codinginflow.com/tutorials/android/>
- <https://gist.github.com/mstfldmr/fc4fa436f2e553b10865>
- <https://www.javatpoint.com/>
- <https://www.mysqltutorial.org/>

**PROJECT REPORT**

**THE RELATIONSHIP BETWEEN  
SOCIAL PHOBIA AND STRESS AMONG  
UNDERGRADUATE STUDENTS**

Submitted by:

TRESSA VARSHA RAJESH

Register No:

SB19PSY040

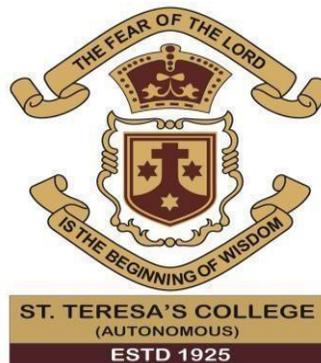
Under the guidance of

Ms. ASHYA K SALIM

In partial fulfilment of requirement for

award of the degree of

B.Sc. PSYCHOLOGY



**ST. TERESA'S COLLEGE  
(AUTONOMOUS), ERNAKULAM**

Nationally Re-accredited at 'A++' level (4<sup>th</sup> cycle)

Affiliated to Mahatma Gandhi University

**MARCH 2022**

## CERTIFICATE

This is to certify that the project report entitled, “THE RELATIONSHIP BETWEEN SOCIAL PHOBIA AND STRESS AMONG UNDERGRADUATE STUDENTS”, is a bonafide record submitted by MS. TRESSA VARSHA RAJESH, Reg.no. SB19PSY040, in partial fulfilment of the requirements for the award of the Degree of Bachelor of Psychology during the academic year 2019-2022.



Ms. Bindu John

Head of the Department

Department of Psychology

St. Teresa's College, Ernakulam

Ms. Ashya K Salim

Assistant Professor

Department of Psychology

St. Teresa's College, Ernakulam

External Examiner:

## **DECLARATION**

I, Tressa Varsha Rajesh, hereby declare that the study presented in the dissertation entitled, “The relationship between social phobia and stress among undergraduate students”, which is submitted to the Department of Psychology, St. Teresa’s College, Ernakulam is a bonafide record of the research work carried out by me, under the supervision and guidance of Ms. Ashya K Salim, Assistant Professor, Department of Psychology, St. Teresa’s College, Ernakulam, in partial fulfilment of the requirements for the degree of Bachelor of Science in Psychology and has not previously formed the basis for the award of any degree, diploma, fellowship, title or recognition before.

Place: Ernakulam

Tressa Varsha Rajesh

Date:

## **ACKNOWLEDGEMENT**

It is not possible to prepare a project report without the assistance and encouragement of other people. This one is certainly no exception. I would like to express my deep heartfelt gratitude to the Department of Psychology, St. Teresa's College, Ernakulam for providing me with the opportunity to undertake the research.

I acknowledge my indebtedness and deep sense of gratitude to my research guide, Ms. Ashya K Salim, Assistant Professor, Psychology, for encouraging and guiding me throughout all the phases of my research.

I extend my sincere thanks to my parents, teachers and my friends who all have supported me throughout the time. I am grateful to each and every one who has given me guidance, encouragement, suggestions and constructive criticisms which has contributed immensely for this project.

Above all, I thank God Almighty for blessing me in all the stages of the project and for helping me complete the project successfully.

Thanking you

Tressa Varsha Rajesh

## CONTENT

<b>SL.NO</b>	<b>TOPIC</b>	<b>PAGE NO</b>
1	CHAPTER I: INTRODUCTION	8
1.1	Background of study	9
1.2	Problem Statement	10
1.3	Need and Significance of the study	10
1.4	Scope of the study	11
1.5	Objectives of the study	11
1.6	Limitations of the study	11
2	CHAPTER II: REVIEW OF LITERATURE	12
2.2	Literature Review	13
3	CHAPTER III: THEORETICAL FRAMEWORK	18
4	CHAPTER IV: RESEARCH AND METHODOLOGY	24
4.1	Objectives	25
4.2	Hypothesis	25
4.3	Research Design	25
4.4	Sources of Data	26
4.5	Sample Size	26
4.6	Sampling Method	26
4.7	Method of Data Collection	26
4.8	Drafting a Questionnaire	26
4.9	Data Analysis Techniques	28
5	CHAPTER V: DATA ANALYSIS	30

6	CHAPTER VI: FINDINGS	34
7	CHAPTER VII: RECOMMENDATIONS	35
8	CHAPTER VIII: CONCLUSIONS	36
	REFERENCES	37
	APPENDIX	42

## **ABSTRACT**

“The relationship between social phobia and stress among undergraduate students”

The current study investigated the relationship between social phobia and stress among undergraduate students. The sample consisted of 150 Indian adults. The data was collected using Social Phobia Inventory (SPIN) (Connor et al., 2000) and Perceived Stress Scale (PSS-10) (Cohen, Kamarck and Mermelstein, 1983). Kolmogorov test was used to test normality of the data. Spearman’s rank correlation and regression analysis were used for data analysis. The results revealed that there is a significant positive correlation between social phobia and stress. Social anxiety occurs when individuals fear social situations in which they anticipate negative evaluations by others or perceive that their presence will make others feel uncomfortable (Philip Jefferies & Michael Ungar, 2020). Perceive stress is the feelings or thoughts that an individual has about how much stress they are under at a given point in time or over a given time period (Anna C. Philips, 2013). Individuals with Social phobia may have a heightened sensitivity to social stress as a result of maladaptive cognitive processes in the context of interpersonal events with potential for negative evaluation (Antonia S. Farmer and Todd Kashdan, 2015). Spearman’s rank correlation and regression analysis were used for data analysis. The results revealed that there is a significant positive correlation between social phobia and stress.

Keywords: Social phobia, Stress

**CHAPTER 1**  
**INTRODUCTION**

Phobia is a persistent and disproportionate fear of some specific object or situation that presents little or no actual danger and yet leads to a great deal of avoidance of these feared situations. The three main categories of phobia in DSM-5 are specific phobia, social phobia and agoraphobia. A person is diagnosed as having a specific phobia if the person shows strong and persistent fear that is triggered by the presence of a specific object or situation. The DSM-5 describes social phobia as characterized by disabling fears of one or more specific social situations. In agoraphobia, the most commonly feared and avoided situations include streets and crowded places such as shopping malls, stores and movie theatres.

In social phobia, in the specific social situations a person fears that he or she may be exposed to the scrutiny and potential negative evaluation of others or that he or she may act in an embarrassing or humiliating manner. Because of their fears, individuals with social phobias either avoid these situations or endure them with great distress. The single most common types of social phobia is the fear of public speaking.

Stress is a feeling of emotional or physical tension which can arise from any event or thought that make you feel frustrated, angry or nervous. It is the bodily reaction to a challenge or demand. Although stress and anxiety share a similar pattern of physiological reactions, they differ in the interpretation of the situation. Stress emerges when an organism is confronted with overstaining demands, whereas anxiety is an emotional consequence of perceived threat. As anxiety is the most important motivation for avoidance behavior and stress is common in many situations in daily life, it is of special interest to investigate both states and how social phobia or social anxiety is related to stress.

## **BACKGROUND OF STUDY**

The topic “The Relationship Between Social Phobia and Stress Among Undergraduate Students” was taken to assess if social phobia and stress are related and to see if the presence of social phobia results in an increase in stress. Social phobia, also called social anxiety disorder is a long term and overwhelming fear of social situations. It is a common problem that usually start during teenage years and can be very distressing and have a big impact on the individual’s life. Feeling nervous in some social situations is normal but in social phobia, everyday interactions cause significant anxiety, self-

consciousness and embarrassment because you fear being scrutinized or being judged negatively by other. In social phobia, fear and anxiety lead to avoidance that can disrupt your life. Severe stress can affect your relationship, daily routines, work, schools or other activities. The symptoms of social phobia or social anxiety disorder can change overtime and may flare up if the individual is facing a lot of changes, stress or demands in your life. Stress can cause feelings of anxiety and depression and can reduce a person's ability to cope in particular situations.

## **PROBLEM STATEMENT**

The current research investigates the relationship between social phobia and stress among undergraduate students.

## **NEED AND SIGNIFICANCE**

It is normal to feel self-conscious, nervous, or shy in front of others at times and most people get through these moments when they need to. Whereas for some individuals, the anxiety that accompanies with feeling shy or self-conscious can be extreme and when that happens the individuals feel intense fear and anxiety that it prevents them from speaking up or socializing. Social phobias like other phobias are a fear reaction to something that isn't actually dangerous although the body and mind react as if the danger is real. It can affect various parts of an individual's life such as work, school and other daily activities. Hans Selye used the word stress to describe the difficulties and strains experienced by living organisms as they struggle to cope with and adapt to changing environmental conditions. Stress effects a personal mentally and physiologically. There are very few studies that assess the relationship between social phobia and stress among undergraduate students. A students college experience is a time that shapes them to who they are and determines what they become and so it is an important stage in a person's life. But there are individuals who have excessive anxiety about specific social situations such as public speaking which is most of the time required to do by college. An undergraduate student with social anxiety would be prevented from doing things like meeting up with friends, going to college or trying something new and cause issues as chronic stress

## **SCOPE OF THE STUDY**

Social anxiety is a concern for young adults around the world, many of whom do not recognize the difficulties they may experience. A large number of young people may be experiencing substantial disruption in functioning and well-being which may be ameliorable with appropriate education and intervention. The impact of social anxiety is widespread, affecting functioning in various domains of life which may lead to lowering general mood and wellbeing (Mineka S, Watson D & Clark LA, 1998). Individuals experiencing social anxiety are more likely to be victims of bullying and are at greater risk of leaving school early with poor qualification all of which can cause severe stress in the individual affecting more aspects of the individual (Philip Jefferies & Michael Ungar, 2020). By carrying out this study, it would be understood how social phobia or social anxiety is related to stress and can be used for further research to find out proper ways to help the individuals with the difficulties they face.

## **OBJECTIVES**

- To analyze the relationship between social phobia and stress among undergraduate students.
- To examine the role of social phobia on stress.

## **LIMITATIONS OF STUDY**

The current research has various limitations. The survey was conducted through online platform using self-report questionnaire so there is a probability that the participants were not being fully honest in their answers. The population from which the sample has been chosen is Indian adults. Therefore, the results of the research cannot be generalized to other countries. Additionally, the sample gotten were not of equal gender distribution and did not assess the various social situations causing the individuals with social phobia to experience increased stress.

**CHAPTER 2**  
**REVIEW OF LITERATURE**

## **LITERATURE REVIEW**

A prevalence study conducted by Philip Jefferies and Michael Ungar (2020) in seven countries using a self-report survey, the global prevalence of social phobia was found to be significantly higher with more than 1 in 3 respondents meeting the threshold criteria for having Social Anxiety Disorder (SAD). It was also found that prevalence and severity of the symptoms did not differ between sexes but varied as a function of age, country, work status, level of education and whether an individual lived in urban or rural location. The data of the study indicated that the social phobia is a concern for young adults around the world, many of whom do not recognize the difficulties they may experience. The findings suggest that levels of social anxiety may be rising among young people.

A cross-sectional study conducted by Hajure M and Abdu Z (2020) showed high prevalence of social phobia among the university students and its significant negative effects on quality of life which require prompt identification and treatment. It was found that the majority of the students have mild forms of social anxiety disorder. The study also shows significant association with higher prevalence of SAD among students in the age group of 18-20 years as compared to older groups. The study concluded that it shows high prevalence of social phobia among the university students and its significant negative effects on quality of life.

In a study by Graham Russell and Phil Topham (2012), the findings are consistent with previous research on social anxiety and suggest that for a significant minority of students, social anxiety is a persistent, hidden disability that impacts on learning well-being. The findings highlighted the need for enhanced pedagogic support for students with social anxiety.

Social Anxiety Disorder (SAD) is a persistent fear of situations where a person is exposed to a possibility of being under scrutiny which may lead to decreased academic performance which in turn can lead to depression and suicidal tendencies. In the cross-sectional descriptive study conducted by Preeti and Parnava Das (2019) it was found that 12.62% of the study participants were having social phobia and 5.95% were having social anxiety. It was concluded that early detection and appropriate intervention needs to be done among the students as they are the most vulnerable and most malleable.

Rakhi Gupta (2019) conducted comparative research in which the hypothesis that it is observed that males suffer with high amount of social anxiety as compared to female and more are the chances of getting prone to other anxiety disorders in future tend to stand correct, which was high in the level of social anxiety but not much difference was found when compared to that of females. It is verified that people that are actually facing severe problems of social anxiety, which can turn into a dangerous disorder, and is important for every individual to take care of themselves and their family members by giving them the freedom of speech, from early stage of development and to open up their shyness, and stage fear from childhood, and mold their childhood in a way that they build up a bold personality in future, and keep away from the problems of anxiety.

Tore Aune, Hans M. Nordahl and Deborah C. Beidel (2022) conducted a study to demonstrate that SAD point prevalence among adolescents can vary from 2.0% to 5.7% based on the diagnostic threshold used in the DSM-5, the results demonstrate that assessing three general social situations reveals an overall prevalence of 5.7%. The population-based study found that among adolescents, the prevalence of SAD varies from 2.0% to 5.7% depending upon the diagnostic threshold. Approximately 3.5% of the individuals diagnosed with SAD had the DSM-5 performance only subtype.

A study conducted (2012) with the aim to evaluate the prevalence of social phobia in a large sample of Brazilian college students and to examine the academic impact of this disorder using the Social Phobia Inventory (SPIN) and the MINI-SPIN, the results showed the prevalence of social phobia among the university students as 11.6% and women with social phobia had significantly lower grades than those without the disorder. It was found that fear of public speaking was the most common social fear and only two out of the 237 students with social phobia (0.8%) had previously received a diagnosis of social phobia and were under treatment. It was concluded that preventive strategies are recommended to reduce the under-recognition and the adverse impact of social phobia on academic performance and overall quality of life of university students.

Maurice M. Ohayon and Alan F. Schatzberg (2010) conducted a study to estimate the co-occurrence of social phobia with major depressive disorder and to analyse their interaction showed the point prevalence for social phobia was 4.4% of the sample which included individuals aged 15 years or older. It was found to be higher in women and decreased with age. Major depressive disorders were found in 19.5% of participants with

social phobia and the odds of developing a major depressive episode 2 years after the appearance of the social phobia was 5.74.

The result of the study conducted to investigate the relationship between perceived stress and social anxiety with psychological cohesion in pandemic crisis, with the emp on covid-19 patients discharged from Sabzevar Vasei Hospital showed that there is a significant relationship between perceived stress and social anxiety and its components with the feeling of psychological cohesion in a negative direction and among the components of social anxiety, the share of social status anxiety is higher (Shakerinasab M, Sani M Z, Farahani Z K, Ghalehnovi Z, 2021)

Dawit Yikealo. Werede Tareke and Ikali Karvinen (2018) conducted a study which aimed to investigate the level of stress among the College of Education (CoE) students in Eritrea Institute of Technology. Descriptive research was undertaken to assess the student's level of stress. A randomly drawn participants completed a self-developed questionnaire assessing their level of stress on five domains. The results reveals that there was a moderate level of stress among the college students. It was found that out of the five domains, academic and environmental stressors were found contributing most to the level of stress among the students. Besides, the students' levels of stress were found to have statistically significant association with their gender and grade point average.

The study conducted by K. Jayasankara Reddy, Karishma Rajan Menin and Anjana Thattil (2018) aimed to understand the sources and impact of academic stress in order to derive adequate and efficient intervention strategies by employing a quantitative research design where participants were screen using Academic Stress Scale. The main objectives of the study were to understand the level of academic stress faced by the students and the different sources that contributed to the same. It was also conducted to understand if there was any gender wise and stream wise differences in academic stress reported by the participants. The study brought into light that academic stress still continues to be devastating problem affecting a student's mental health and well-being. Stream wise difficulties in the experience of stress were also highlighted.

The objective of the cross-sectional study conducted by Vivek, Girish, Yugantara and Alka (2013) was to assess stress among students of various professional colleges and its association with various academic, social and health-related factors with the help of a

pretested self-administered questionnaire for data collection. It was concluded that students from all three fields (medical, dental and engineering) were exposed to stress.

Tausif Islam, M Sabrina Moonajilin and Rajib-ul-islam (2018) conducted a cross sectional study with the purpose to find out the causes of stress among university students by distributing structured questionnaires to the students of Jahangirnagar University. Convenience sampling technique was adapted for the selection of respondents for data collection and a quantitative method was used in gathering and analysing the data. The results obtained showed that according to majority of the respondents, future worry was the most effective factor causing stress.

Henry Stoddard (2017) conducted a study in which university students and various mental health professional were surveyed and interviewed. The results showed no significant differences in demographic groups and perceived stress levels.

From the analysis and interpretations, the study conducted by Nayana, Priyanka and Vidya (2020) to know the academic stress among college students, a clear-cut information about the frequency of academic stress among male and female, factors which cause academic stress, problems caused by stress, the techniques implemented by the college was put forth. The findings concluded that the major problems caused by academic stress in both male and female students are sleep disturbance, poor concentration and headache.

A study conducted by Christopher J. Graver (2007) assessed the influence of stress on neuropsychological functioning was assessed in socially phobic (SP), comorbid socially phobic/major depression (CM), and asymptomatic control subjects (AC) under baseline and stressor conditions.

The finding of a study (Anke W. Blote, Miers A. C., Westenberg P. M, 2021) revealed that adolescents self-perceived and social anxiety and stress responses are linked both concurrently as well as over time. The findings of the study that adolescents with higher levels of social anxiety relatively more often develop adaptive as well as maladaptive stress responses later on, is consistent with the Wright et al. (2010) study.

## **RESEARCH GAP**

The study revolved around social phobia and stress and its relationship among undergraduate students. Life as a socially-anxious student can be challenging as experiencing a constant state of anxiety can hamper studies by affecting the student's ability to focus on what the teachers and lectures are saying. There are various studies conducted on social phobia and stress separately but very few on the relationship between both especially among undergraduate students. Fewer investigations of the relationship between social phobia and stress among undergraduate students in Indian context call out the need for such studies in India. A positive correlation of the variables in the current study can highlight the importance of providing help to students struggling with social phobia. The current study could be enhanced by identifying the various factors that could cause social phobia among undergraduate students and research on the various methods that could be adopted to help students with social phobia and experiencing increased stress to perform better in college.

**CHAPTER - 3**  
**THEORETICAL FRAMEWORK**

Social phobia as the DSM-5 describes it, is characterized by disabling fears of one or more specific social situations such as public speaking, urinating in a public bathroom or eating or writing in public. In such situations, an individual fears that he or she may be exposed to the scrutiny and potential negative evaluation of others of that they may act in an embarrassing or humiliating manner. Because of their fears, individuals with social phobia either avoid such situations or endure them with great distress. The most common type of social phobia is the intense fear of public speaking where studies have found that nearly nine out of ten people with social anxiety are frightened to speak in front of others. In higher educational institutions where speaking in front of others is essential academic skills and a key to participate in social activities, having social phobia will impose a significant burden on student's academic performance and social life.

Social phobias typically begin somewhat later, during early or middle adolescence or certainly by early adulthood. Fear is an intense biological response to immediate danger and anxiety is an emotion regarding things we think may happen. Many studies have indicated that social anxiety is a prevalent disorder among university students. For example, studies from Sweden and India have reported the prevalence of SAD among university students to be 61% and 19.5% respectively. The circumstances where social phobia might occur include restaurant, cafeterias dining halls, seminars, meetings and other places where the person feels observed by others. The individual experiences anxiety symptoms that make it uncomfortable and shameful. Frequently, a variety of somatic and cognitive signs and symptoms occur because of being embarrassed or judged pessimistically by others, which is typically characterized by autonomic excitement, including symptoms like increased sweating, apnea, tremors, tachycardia and nausea.

### **Theories of Social Phobia**

#### **1. Clark & Wells' model of social phobia**

Clark & Wells' model of social phobia, published in 1995, provides a cognitive behavioral formulation of social anxiety. Clark (2001) describes how the model attempts to solve the 'puzzle' of why social anxiety persists despite regular exposure to feared social situations. The model proposes that entering a feared situation activates a set of beliefs and assumptions that have been shaped by one's earlier experiences. These beliefs and assumptions concern both the individual, and how they think they should behave in social situations. Holding these assumptions predisposes socially anxious individuals to appraise

particular social situations as dangerous, and to make predictions that they will not meet their own (often high) standards for performance. Once a situation has been appraised in this manner, Clark & Wells propose that an ‘anxiety program’ is activated automatically. This program leads to automatic changes in affective, attentional, behavioral, cognitive, and somatic processing which are intended to protect the individual from harm, but which are accompanied by unintended consequences that serve to maintain the social anxiety. An important insight of the model is that when people with social anxiety enter feared situations their focus of attention changes. They become preoccupied with highly detailed monitoring and observation of themselves, and they “use the internal information made accessible by self-monitoring to infer how they appear to other people and what other people are thinking about them” (Clark, 2001). Clark argues that this results in a vicious cycle whereby “most of the evidence for their fears is self-generated and disconfirmatory evidence (such as other people’s responses) becomes inaccessible or is ignored”.

## 2. Cognitive model of Albert Bandura

The cognitive model of Bandura is another theoretical model that explains the cause of social anxiety. According to Albert Bandura, individuals learn new ways of thinking and behaviors by observing how other people think and behave. According to social learning theory, people with anxiety disorders may have learned to be anxious through prior contact with other people, other people may have communicated by their actions or information they provided that certain situations or objects are dangerous and subsequently must avoid at all cost. Through exposure to early learning experiences people may mimic the anxious behavior of others. Social learning theory suggest that this learning can take place simply through observation. Therefore, individuals may learn to avoid certain objects or situations without having any experience or knowledge. The main construct to this model of explaining social anxiety is self-efficacy. It refers to the perceive ability of the individual to perform a desired action. In Bandura’s view of social anxiety, when people consider themselves without the necessary authority to dealt with potentially hazardous and life-threatening events, they feel very stressed.

### 3. Psychodynamic theory

From the psychodynamic perspective, social anxiety is believed to be part of a larger problem that develops during childhood. Scientists with this psychodynamic perspective view anxiety as a disorder of childhood origin and therefore see social anxiety as resulting from early experience and attachments to your caregivers and other important people in the person's life. Although that is no comprehensive psychoanalytic theory of social phobia, there are several beliefs about the origins of social anxiety. Each of the conflicts is believed to result in shame, insecurity, withdrawal and low self-esteem. According to this the origins may be an expectation of shame and humiliation elicited by critical or harsh parents, a conflict between the individual's need to achieve and succeed and fear of success, a conflict between an individual's need for independence and fear of rejection or abandonment by parents, or narcissistic fear of being unable to make a perfect impression.

### 4. Meta-cognitive model

Another model of social anxiety is the meta-cognition model which can be considered as an aspect of public knowledge that plays a role in all cognitive activities. Meta-cognition is the aspect part of information processing systems that are responsible for monitoring, interpreting, evaluating the content and self-organizing processes of the information processor system. Researchers such as Wells have focused on identifying the meta-cognitive beliefs involved in the information of social anxiety and he has referred to cognitive-attention syndrome in people who have problem controlling their repetitive thoughts, which will appear in the form of anxiety, mental rumination and self-reviewing behaviour. In the meta-cognitive model, Wells proposed that most of the negative general assessments or conventional beliefs are influenced and controlled by meta-cognitive processes.

## **Theories of Stress**

Stress is generally known as patterns of disruptive physiological and psychological reaction to events that threaten the ability to cope. Any circumstances that threaten or are perceived to threaten one's well-being and thereby tax one's coping abilities can lead to stress. There are many approaches to the study of stress. In one approach stress is seen as a stimulus and focus on the impact of stressors, in another stress is treated as a response and examines the strains that the stressors produce the physiological consequences and the third

approach views stress as a process that involves continuous interaction and adjustment or transaction between the person and the environment.

### 1. Stress as a response model

Stress as a response model which was initially introduced by Hans Selye (1956) describes stress as a physiological response pattern and was captured with his general adaptation syndrome (GAS) model that describes stress as a dependent variable and includes the three concepts; stress is a defence mechanism, stress follows stages of alarm, resistance and exhaustion and finally if the stress is prolonged or severe, it could result in diseases of adaptation or even death. Later, in the stress concept: past, present and future (1983), he introduced the idea that the stress response could result in positive or negative outcome based on the cognitive interpretations of the physical symptoms or physiological experience. According to this, stress could be experienced as eustress or distress. However, Selye considered stress to be a physiological response. Gradually, the thinking on stress to include and involve psychological concepts earlier in stress model was expanded by other researchers. The response model of stress includes coping within the model itself. When confronted with a negative stimulus, the alarm response initiates the sympathetic nervous system to fight or avoid the stressor, then the resistance response the initiated the physiological systems with a flight or fight reaction to the stressor, returning the system to homeostasis, accommodating the stressor.

### 2. Stress as a stimulus

The theory of stress as a stimulus which was introduced in the 1960s viewed stress as a significant life event or change that demands response, adjustment or adaptation. Holmes and Rahe theorized that stress was an independent variable in the health-stress-coping equation while some other correlations emerged between the social readjustment rating scale scores and illness. There were some problems with stress as a stimulus theory as it assumes change is inherently stressful, life events demand the same level of adjustment across the population and that there is a common threshold of adjustment beyond which illness will result.

### 3. Transactional theory of stress

Richard Lazarus developed the transactional theory of stress and coping in the attempt to explain stress as more of a dynamic process. This model presents stress as a

product of a transaction between a person (including multiple systems: cognitive, physiological, affective, psychological, neurological) and his or her complex environment. Dr. Susan Hobasa introduced stress as a transaction model with the most impact where the concept of hardiness was first used. Hardiness refers to the pattern of personality characteristics that distinguishes individuals who remain healthy under life stress compared to those individuals who develop health problems. Multiple variables were introduced to the stress transaction model by researchers, expanding and categorizing various factors to account for the complex systems involved in experiencing a stressor.

**CHAPTER – 4**  
**RESEARCH AND METHODOLOGY**

## **OBJECTIVES**

- To analyze the relationship between social phobia and stress among undergraduate students
- To examine the role of social phobia on stress.

## **HYPOTHESIS**

1. There is a positive correlation between social phobia and stress among undergraduate students.
2. Social phobia significantly predicts stress.

## **RESEARCH DESIGN**

Descriptive and correlational research design was used to help understand the relationship between social phobia and stress. Descriptive research design is a type of research design that aims to obtain information to systematically describe a phenomenon, situation, or population. It involves the collection of quantifiable and systematic data that can be used for the statistical analysis of the research problem. The data collected in descriptive research provides a base for further research as it helps obtain a comprehensive understanding of the research question so that it can be answered appropriately. The correlational research is a technique that help researchers establish a relationship between two closely connected variables. The aim of correlational research is to identify variables that have some sort of relationship to the extent that a change in one creates some change in the other. There are essentially three main types of correlational research which are positive correlational which involves two variables that are statistically corresponding where an increase in one variable creates a like change in the other, negative correlational which involves two variables that are statistically opposite where an increase in one of the variables creates an alternate effect or decrease in the other variable and zero correlational research which involves two variables that are not necessarily statistically connected and a change in one of the variables may not trigger a corresponding or alternate change in the other variable.

## **SOURCES OF DATA**

- Primary sources of data: the primary source of data was from the survey conducted on undergraduate students
- Secondary sources of data: the secondary sources of data include references from websites and books

## **SAMPLE SIZE**

The study conducted has a sample size of 150, consisting undergraduate students (18-21 years of age). Out of the total number of participants, 101 were females and 49 were males. There were 33 individuals of 18 years of age, 41 individuals of 19 years of age, 45 people of 20 years of age and 31 individuals of 21 years of age.

## **SAMPLING METHOD**

Convenient sampling method was used for administering the research study. Convenient sampling is the most common form of non-probabilistic sampling method where the sample is taken from a group of people easy to contact or to reach.

## **METHOD OF DATA COLLECTION**

The data required for the research study were collected in the form of an online questionnaire, sending the questionnaire through Google form to the participants that fit the criteria.

## **DRAFTING QUESTIONNAIRE**

The scales used in the study are Social Phobia Inventory (SPIN) and Perceived Stress Scale-10 (PSS-10). The Social Phobia Inventory (abbreviated as SPIN) is a 17-item questionnaire for screening and measuring severity of Social Anxiety Disorder (Social Phobia – SP). It was developed in 2000 by Connor et al. at the Psychiatry and Behavioral Sciences Department, Duke University, USA. SPIN assess different aspects related to social phobia – fear, avoidance, and physiologic symptoms. Each item is measured on a 5-point Likert scale, ranging from 0 (not at all) to 4 (extremely).

Respondents indicate how much each item bothered them during the past week. A categorical interpretation is suggested, where scores less than 20 are considered no SP, 21-30 mild SP, 31-40 moderate SP, 41-50 severe SP, and 51 or higher very severe SP.

Results from the original validation study suggest that the SPIN possesses strong internal consistency (Full scale  $\alpha$  0.94, subscales -0.80-0.91), adequate test-retest reliability ( $r = 0.78$ ), significant convergent validity with Brief Social Phobia Scale ( $r = 0.57$ , sub-scale  $r = 0.47$ -0.66), divergent validity across 3 measures ranged from  $r = 0.01$ - 0.34, adequate construct validity – was able to differentiate between subjects with and without social phobia (cut off score 19), also different levels of social phobia were reflected by different levels of SPIN scores. Another study conducted with Brazilian university students also found SPIN to have excellent psychometric properties.

The Perceived Stress Scale (PSS) is a measure of the degree which situations in an individual's life is evaluated as stressful developed by Cohen, Kamarck and Mermelstein, published in 1983. It is a well-established self-report measure based on the conceptualization of stress. The scale assesses “the degree to which situations in one's life are appraised as stressful”. The 14-item scale (PSS14) was developed in 1983 and there have been no major revisions since. A four-item and 10-item (PSS10) versions of the scale has also been validated. The PSS10 allows the assessment of perceived stress without any loss of psychometric quality. The shortened 10-item scale (PSS-10) showed slightly improved reliability (Cronbach alpha = .78 vs. Cronbach alpha = .75) and equivalent validity and has therefore been recommended for epidemiological and clinical research. The Perceived stress scale is a classic stress assessment instrument. The Perceived Stress Scale (PSS) is the most widely used psychological instrument for measuring the perception of stress. The PSS was designed for use in community samples with at least a junior high school education. PSS scores are obtained by reversing responses (e.g., 0 = 4, 1 = 3, 2 = 2, 3 = 1 & 4 = 0) to the four positively stated items (items 4, 5, 7, & 8) and then summing across all scale items. A short 4 item scale can be made from questions 2, 4, 5 and 10 of the PSS 10 item scale.

## DATA ANALYSIS TECHNIQUE

The data used was qualitative and was analysed using Statistical Package for Social Sciences (SPSS Statistics 28) and Excel. SPSS is a statistical software developed by IBM for data management, analysis and investigation. The variables were social phobia and stress. The reliability of questionnaire's was assessed using SPSS and the Cronbach's alpha value was found to be 0.981 for SPIN and 0.88 for PSS-10. The sample size was greater than 50, thus the Kolmogorov test was used to find the normality. The test of normality was conducted and the Kolmogorov-Smirnov significance value for both SPIN and PSS were found to be  $<.001$  which shows that the data is non normally distributed and hence non-parametric test is used. The data was collected using a convenient sampling method which is a type of non-probability sampling method where the sample is taken from a group of people easy to contact or to reach. To find the relationship between social phobia and stress, Spearman's rank correlation test was used and Regression analysis was used to analyze the data and predict the outcome.

Spearman's Rank Correlation is a nonparametric measure of rank correlation (statistical dependence between the rankings of two variables). It assesses how well the relationship between two variables can be described using a monotonic function. The Spearman correlation between two variables is equal to the Pearson correlation between the rank values of those two variables; while Pearson's correlation assesses linear relationships, Spearman's correlation assesses monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman correlation of  $+1$  or  $-1$  occurs when each of the variables is a perfect monotone function of the other.

Intuitively, the Spearman correlation between two variables will be high when observations have a similar (or identical for a correlation of 1) rank (i.e., relative position label of the observations within the variable: 1st, 2nd, 3rd, etc.) between the two variables, and low when observations have a dissimilar (or fully opposed for a correlation of  $-1$ ) rank between the two variables. Spearman's coefficient is appropriate for both continuous and discrete ordinal variables.

Regression analysis is a set of statistical processes for estimating the relationships between a dependent variable (often called the 'outcome variable') and one or more independent variables (often called 'predictors', 'covariates', or 'features'). The most common

form of regression analysis is linear regression, in which one finds the line (or a more complex linear combination) that most closely fits the data according to a specific mathematical criterion. Simple linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable.

Raw data in this study is collected using scales, Social Phobia Inventory (SPIN) and Perceived Stress Scale-10 (PSS-10). SPIN has 17 questions and is a Likert-type scale with five-point: Not at all, A little bit, Somewhat, Very much and Extremely scored as 0,1,2,3 and 4 respectively. The total SPIN score is calculated by adding up all the items. A score above 19 appears to suggest a possibility of social anxiety. The possible range is between 0 and 68. PSS-10 has 10 questions, out of which questions 4,5,7 and 8 are reverse scored. It is a Likert-type scale with five points: Never, Almost Never, Sometimes, Fairly Often and Very Often scored as 0,1,2,3 and 4. PSS scores are obtained by reversing the four positively stated items (items 4,5,7 and 8) and then summing across all scale items.

**CHAPTER – 5**  
**DATA ANALYSIS**

The study was conducted among 150 undergraduate students who are currently pursuing their bachelor's degree. The objectives of the study are 1) To examine the relationship between social phobia and stress among undergraduate students. To meet the objectives of the study the following hypothesis has been put forth which is 1) There is a significant positive correlation between social phobia and stress. Correlational research design was used to meet the objectives of the study. The data was analysed using Spearman's rank correlation coefficient analysis. Hierarchical regression was used to check the moderating impact of social phobia on stress.

**Table 1**

*Reliability Of Instruments Obtained On The Sample*

S.no	Instruments	N	Cronbach' $\alpha$
1.	SPIN	150	.981
2.	PSS-10	150	.088

The table shows the reliability of the instruments using Cronbach's  $\alpha$  and number of values in the population (N). Cronbach's alpha reliability coefficient normally ranges between 0 and 1, the closer the coefficient is to 1.0, the greater is the internal consistency of the items in the scale. The instruments used in the study are Social Phobia Inventory (SPIN) and Perceived Stress Scale-10 (PSS-10), both of which had Cronbach's alpha coefficient of 0.981 and 0.088 respectively. Both the coefficients are closer to 1 hence it can be said that SPIN and PSS-10 have good reliability.

**Table 2***Result of Kolmogorov-Smirnov test of normality of Social Phobia and Stress.*

	Statistic	df	Sig.
Social phobia	.126	150	.001
Stress	.120	150	.001

The Kolmogorov- Smirnov test of normality pointed out that the significance level of social phobia and stress is less than 0.05, and hence they are not normally distributed.

**Table 3***Summary of Spearman's Rank Correlation between Social phobia and stress among undergraduate students*

Variable	Stress
Social phobia	.581**

\*\*=  $p < 0.01$

The above table examines the relationship between social phobia and stress among undergraduate students. With reference to the table, it can be noted that there exists a significant positive correlation between social phobia and stress. Hence the hypothesis “There is a significant positive correlation between social phobia and stress” is accepted.

**Table 4***Regression analysis predicting the Stress (DV) by Social phobia (IV)*

Variable	$R^2$	Adjusted $R^2$	df	F	B	T	Sig.
Social phobia	.263	.258	149	52.738	.102	7.262	<.001

This table shows the results of regression analysis for predicting the impact of social phobia on stress.  $R^2$  shows the total variation for the dependent variable that could be explained by the independent variable. Since the value is greater than 0.5, it shows that the model is effective enough to determine the relationship. Adjusted  $R^2$  shows the generalization of the results, i.e., the variation of the sample from the population in multiple regression. It is required to have a difference between  $R^2$  and Adjusted  $R^2$  minimum. In this case since the value is .258, which is not far off from .263, so it is good. Since the significance value is less than 0.05, the result is significant. F represents an improvement in the prediction of the variable by fitting the model after considering the inaccuracy present in the model. Since the value is greater than 1, it is an efficient model. A study conducted (Antonina S, Farmer and Todd B. Kashdan, 2015) showed that people with social phobia displayed greater stress sensitivity, particularly in negative emotion reactions to stressful social events, compared to healthy controls.

## **CHAPTER – 6**

### **FINDINGS**

In the current study which aimed to assess the relationship between social phobia and stress among undergraduate students, it was found that there is a significant positive correlation between social phobia and stress among undergraduate students. A study conducted (Antonina S, Farmer and Todd B. Kashdan, 2015) showed that people with social phobia displayed greater stress sensitivity, particularly in negative emotion reactions to stressful social events, compared to healthy controls. The hypothesis that there is significant positive correlation between social phobia and stress has been proved by the spearman rank correlation and according to the results of the regression analysis it has been found that when an individual experiences increased social phobia or scores high on social phobia, he or she would have increased perceived stress as well. It has been proved that social phobia significantly predicts stress among undergraduate students.

## **CHAPTER – 7**

### **RECOMMENDATIONS**

The results of the statistical analysis conducted to assess the relationship between social phobia and stress revealed that there is a positive correlation between the variables social phobia and stress. The findings of the study show that there is a significant relationship between social phobia and stress. In light of the limitations encountered during the research process, the following recommendations could be taken into consideration for further research on relationship between social phobia and stress in a correlational research study. Firstly, the sample used in this study was not of equal gender distribution so taking a sample with equal gender distribution can help gain insight about how gender differences might affect both variables. Secondly, the study could be more focused one aspect of social phobia so that the relationship could be better understood. Additionally, better understanding can be made about ways to reduce stress in individuals who experience social phobia by conducting future research by including research about the cause and effect of increased perceived stress experienced by individuals with social phobia.

## **CHAPTER – 8**

### **CONCLUSION**

The study focused on analyzing the relationship between social phobia and stress among undergraduate students. The data was collected using convenient sampling method and consisted of 101 female and 49 male undergraduate students in India. The results were assessed using Spearman rank correlation and regression analysis carried out in SPSS software. The results of the Spearman's rank correlation were positive correlation, which shows that there is a significant relationship between social phobia and stress. The results of the regression analysis imply that individuals who experience increased social phobia would also experience increased stress under the social situation which causes the individual to experience increased fear and anxiety. It can be inferred from the current study that there is a significant positive correlation between the two variables, social phobia and stress among undergraduate students and that when social phobia increases, stress increases along with it.

**CHAPTER - 9**  
**REFERENCES**

## Books

Butcher, J. N., Hooley, J. M., & Mineka, S. M. (2013). *Abnormal psychology (16th edition)* (16th ed.). Pearson.

## Journals

Aune, T., Nordahl, H. M., & Beidel, D. C. (2022). Social anxiety disorder in adolescents: Prevalence and subtypes in the Young-HUNT3 study. *Journal of Anxiety Disorders*, 87, 102546. <https://doi.org/10.1016/j.janxdis.2022.102546>

Baptista, C. A., Loureiro, S. R., de Lima Osório, F., Zuardi, A. W., Magalhães, P. V., Kapczinski, F., Filho, A. S., Freitas-Ferrari, M. C., & Crippa, J. A. S. (2012). Social phobia in Brazilian university students: Prevalence, under-recognition and academic impairment in women. *Journal of Affective Disorders*, 136(3), 857–861. <https://doi.org/10.1016/j.jad.2011.09.022>

Blöte, A. W., Miers, A. C., & Westenberg, P. M. (2021). Concurrent and Prospective Associations Between Social Anxiety and Responses to Stress in Adolescence. *Research on Child and Adolescent Psychopathology*, 50(5), 659–668. <https://doi.org/10.1007/s10802-021-00880-3>

Farmer, A. S., & Kashdan, T. B. (2015). Stress sensitivity and stress generation in social anxiety disorder: A temporal process approach. *Journal of Abnormal Psychology*, 124(1), 102–114. <https://doi.org/10.1037/abn0000036>

Graver, C. J., & White, P. M. (2007). Neuropsychological effects of stress on social phobia with and without comorbid depression. *Behaviour Research and Therapy*, 45(6), 1193–1206. <https://doi.org/10.1016/j.brat.2006.08.002>

Gupta, R. (2019). Social anxiety: A comparative study. *The International Journal of Indian Psychology*, 7(1). <https://doi.org/10.25215/0701.114>

- Hajure, M., & Abdu, Z. (2020). <p>social phobia and its impact on quality of life among regular undergraduate students of mettu university, mettu, ethiopia</p> *Adolescent Health, Medicine and Therapeutics, Volume 11, 79–87.* <https://doi.org/10.2147/ahmt.s254002>
- Hakami, R. M., Mahfouz, M. S., Adawi, A. M., Mahha, A. J., Athathi, A. J., Daghreeri, H. H., Najmi, H. H., & Arishi, N. A. (2018). Social anxiety disorder and its impact in undergraduate students at Jazan University, Saudi Arabia. *Mental Illness, 9*(2). <https://doi.org/10.4081/mi.2017.7274>
- Islam, T., Moonajilin, M. S., & R.--. (2019). A study on stress among University students, bangladesh. *International Journal of Academic Health and Medical Research (IJAHMR), 2*(10), 10–17. <https://philarchive.org/archive/ISLASO-2>
- Jayasankara Reddy, K., Rajan Menon, K., & Thattil, A. (2018). Academic Stress and its Sources Among University Students. *Biomedical and Pharmacology Journal, 11*(1), 531–537. <https://doi.org/10.13005/bpj/1404>
- Jefferies, P., & Ungar, M. (2020). Social anxiety in young people: A prevalence study in seven countries. *PLOS ONE, 15*(9). <https://doi.org/10.1371/journal.pone.0239133>
- Musa, C., & Lépine, J. (2000). Cognitive aspects of social phobia: a review of theories and experimental research. *European Psychiatry, 15*(1), 59–66. [https://doi.org/10.1016/s0924-9338\(00\)00210-8](https://doi.org/10.1016/s0924-9338(00)00210-8)
- Ohayon, M. M., & Schatzberg, A. F. (2010). Social phobia and depression: Prevalence and comorbidity. *Journal of Psychosomatic Research, 68*(3), 235–243. <https://doi.org/10.1016/j.jpsychores.2009.07.018>
- P., & Das, P. (2019). Prevalence of social anxiety disorder and its determinants among undergraduate medical students of east delhi. *International Journal Of Community*

*Medicine And Public Health*, 6(3), 1335. <https://doi.org/10.18203/2394-6040.ijcmph20190636>

Russell, G., & Topham, P. (2012). The impact of social anxiety on student learning and well-being in higher education. *Journal of Mental Health*, 21(4), 375–385. <https://doi.org/10.3109/09638237.2012.694505>

SHAHBAZIRAD, A., GHADAMPOUR, E., GHAZANFARI, F., & MOMENI, K. (2016). MODEL OF SOCIAL ANXIETY DISORDER: A STRUCTURAL EQUATION MODELING OF COGNITIVE, META-COGNITIVE, BEHAVIORAL ASPECTS. *PSYCHOLOGIA*, 59(4), 202–214. <https://doi.org/10.2117/psysoc.2016.202>

Shakerinasab, M., Sani, M. Z., Farahani, Z. K., & Ghalehnovi, Z. (2021). The relationship between perceived stress and social anxiety with psychological cohesion in pandemic crisis, with emphasis on Covid-19 patients discharged from Sabzevar Vasei Hospital. *Journal of Clinical Psychology*, 13(2), 87–96. <https://doi.org/10.22075/JCP.2021.22641.2083>

Waghachavare, V. B., Dhumale, G. B., Kadam, Y. R., & Gore, A. D. (2013). A Study of Stress among Students of Professional Colleges from an Urban Area in India = دراسة عن التوتر بين طلاب الكليات المهنية من منطقة حضرية في الهند. *Sultan Qaboos University Medical Journal*, 13(3), 422–429. <https://doi.org/10.12816/0003266>

## Websites

Barrell, A. (2020, April 24). *Stress vs. anxiety: How to tell the difference*. Medical News Today. <https://www.medicalnewstoday.com/articles/stress-vs-anxiety#when-to-see-a-doctor>

Dent, I. (2014). *Theories of stress*. Slideshare. Retrieved 03–17, from <https://www.slideshare.net/UDDent/theories-of-stress>

- Hengen, K. M. (2021). *Stress Makes the Difference: Social Stress and Social Anxiety in Decision-Making Under Uncertainty*. *Frontiers*.  
<https://www.frontiersin.org/articles/10.3389/fpsyg.2021.578293/full>
- How Psychoanalysis Is Used in Treating Social Anxiety Disorder*. (2020, September 24). Verywell Mind. <https://www.verywellmind.com/psychoanalysis-for-social-anxiety-disorder-3024911>
- Psychology Tools. (2022, January 28). *Cognitive Behavioral Model Of Social Phobia (Clark, Wells, 1995)*. <https://www.psychologytools.com/resource/cognitive-behavioral-model-of-social-phobia-clark-wells-1995/>
- Social Anxiety Disorder: More Than Just Shyness*. (2022). National Institute of Mental Health (NIMH). <https://www.nimh.nih.gov/health/publications/social-anxiety-disorder-more-than-just-shyness>
- Social anxiety disorder (social phobia) - Symptoms and causes*. (2021, June 19). Mayo Clinic. <https://www.mayoclinic.org/diseases-conditions/social-anxiety-disorder/symptoms-causes/syc-20353561>
- Social Explanations of Anxiety Disorders*. (2020). Mentalhelp.Net. <https://www.mentalhelp.net/anxiety/social-explanations/>
- Social Phobia (for Teens) - Nemours KidsHealth*. (2018). Nemours Teen Health. <https://kidshealth.org/en/teens/social-phobia.html>
- The Cognitive Theory of Social Anxiety*. (2012, December 7). Andrew Kukes Foundation for Social Anxiety. <https://akfsa.org/research/the-cognitive-theory-of-social-anxiety/#:%7E:text=THE%20MAIN%20POINT%3A,and%20underestimating%20the%20positive%20aspects>
- Walinga, J. (2019, June 28). *12.3 Stress and Coping – Introduction to Psychology*. Pressbooks. <https://openpress.usask.ca/introductiontopsychology/chapter/stress-and-coping/>

## **CHAPTER – 10**

### **APPENDIX**

## APPENDIX A

### Informed consent

I understand that my participation in this study is voluntary and I can decline my participation without giving any reason. I understand that this study requires demographic details and questionnaires to be filled out. I have been given a sufficient amount of time to review the information and have sought the required clarification too.

By clicking on the "I agree" option, I give my full consent to participate in this study.

I agree

I disagree

### Socio-demographic details

1. Name:
2. Age
3. Gender
4. Education

## APPENDIX B

### Social Phobia Inventory (SPIN)

Connor et al. (2000)

	Not at all	A little bit	Sometimes	Very much	Extremely
1. I am afraid of people in authority					
2. I am bothered by blushing in front of people					
3. Parties and social events scare me					

4. I avoid talking to people I don't know					
5. Being criticized scares me a lot					
6. I avoid doing things or speaking to people for fear of embarrassment					
7. Sweating in front of people causes me distress					
8. I avoid going to parties					
9. I avoid activities in which I am the center of attention					
10. Talking to strangers scares me					
11. I avoid having to give speeches					
12. I would do anything to avoid being criticized					
13. Heart palpitations bother me when I am around people					
14. I am afraid of doing things when people might be watching					
15. Being embarrassed or looking stupid are among my worse fears					
16. I avoid speaking to anyone in authority					
17. Trembling or shaking in front of others is distressing to me					

## APPENDIX C

### Perceived Stress Scale-10 (PSS-10)

Cohen, Kamarck and Mermelstein (1983)

	Never	Almost never	Sometimes	Fairly often	Very often
1. In the last month, how often have you been upset because of something that happened unexpectedly?					
2. In the last month, how often have you felt that you were unable to control the important things in your life?					
3. In the last month, how often have you felt nervous and “stressed”?					
4. In the last month, how often have you felt confident about your ability to handle your personal problems?					
5. In the last month, how often have you felt that things were going your way?					
6. In the last month, how often have you found that you could not cope with all the things that you had to do?					
7. In the last month, how often have you been able to control irritations in your life?					
8. In the last month, how often have you felt that you were on top of things?					
9. In the last month, how often have you felt that you were on top of things?					
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?					



**Project Report**

**On**

**A STUDY ON THE EFFECT OF COVID-19  
ON SMALL BUSINESSES**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**SARAH MARIA SAJU**

**(Register No. AB19BMAT003)**

*Under the Supervision of*

**DR. ELIZABETH RESHMA M T**



**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

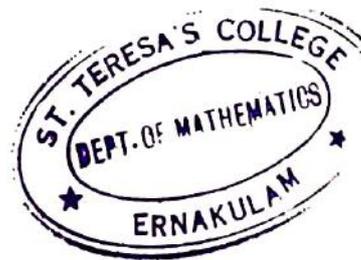


CERTIFICATE

This is to certify that the dissertation entitled, **A STUDY ON THE EFFECT OF COVID-19 ON SMALL BUSINESSES** is a bonafide record of the work done by Ms. **SARAH MARIA SAJU** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 08-03-2022  
Place: Ernakulam

**Dr. Elizabeth Reshma M.T**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



**Dr. Ursala Paul**  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1: *Jeema Saxe* *June*  
6/5/2022

2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Dr. Elizabeth Reshma M.T, Assistant Professor, Department of Mathematics, St. Teresa's College (Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 08-03-2022



SARAH MARIA SAJU

AB19BMAT003

## ACKNOWLEDGEMENT

Firstly, I thank God Almighty for His grace for being able to complete this project work successfully. I express my deep sense of gratitude to our guide Dr. Elizabeth Reshma M.T (Assistant Professor), Department of Mathematics and Statistics, St. Teresa's College, Ernakulam, for her valuable guidance and suggestions. I would like to mention the unending help and support provided by Ms. Shanty B.P (H.O.D of Statistics), Ms. Reshmy S (Assistant professor) and Ms. Rosmin Raju throughout the course of project. I also thank Dr. Ursala Paul ( H.O.D of Mathematics), other teachers of the department, parents, friends, especially my group members and all those who gave me the moral support. This project would not have been possible without the support of the people mentioned above.

Ernakulam.  
Date: 08-03-2022

  
SARAH MARIA SAJU  
AB19BMAT003

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 SIGNIFICANCE OF STUDY . . . . .	2
1.2 OBJECTIVES . . . . .	2
<b>2 DATA DESCRIPTION</b>	<b>3</b>
<b>3 METHODOLOGY</b>	<b>4</b>
3.1 PERCENTAGE CHANGE . . . . .	5
3.2 TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION . . . . .	5
3.2.1 ASSUMPTIONS . . . . .	5
3.2.2 DEFINITION . . . . .	5
3.2.3 USES . . . . .	6
3.3 CHI SQUARE TEST . . . . .	6
3.3.1 ASSUMPTIONS . . . . .	6
3.3.2 PROCEDURE . . . . .	7
3.3.3 APPLICATIONS . . . . .	7
<b>4 DATA ANALYSIS</b>	<b>8</b>
4.1 GRAPHICAL ANALYSIS . . . . .	8
4.2 STATISTICAL ANALYSIS . . . . .	13

5	CONCLUSION	17
	<i>REFERENCES</i> . . . . .	18
	<i>ANNEXURE</i> . . . . .	19

# Chapter 1

## INTRODUCTION

---

COVID-19 put an immediate halt to many business activities across the globe, as several countries had shut down their ports, airports and domestic transportation while imposing nation-wide lockdowns, leading to a disturbance in business and civil life. The enormity of the lockdown affected manufacturing activities and supply chains alike, disrupting the overall economy of India. Like any other major sector, the Micro, Small and Medium Enterprise (MSME) sector also witnessed a considerable decline in economic activities and loss of jobs due to the nationwide lockdown. Declining output is not a good sign for the economy as the sector provides gainful non- farm employment to millions, especially in rural areas. MSME sector has emerged as a very important sector of the Indian economy, contributing significantly to employment generation, innovation, exports and inclusive growth of the economy. MSMEs are the backbone of the socio-economic development of our country. It also accounts for 45 percentage of the total industrial production, 40 percent of total exports and contributes very significantly to the GDP. Manufacturing segment within the MSME contributes to 7.09 percent of GDP. MSME also contributes to 30 percent of services. The total contribution of MSME to the GDP is 37.54. The announcement of country wide lockdown dragged MSME owners in unexpected times, where no one had experience to handle this kind of situation. Extended lockdown had negative impact on supply of finished goods, procurement of raw material and availability of employees to work in production and supply processes. During April to June 2020, sector faced challenges related to debt repayments, wages/salaries, statutory dues, etc. Reports have shown

that disruptions caused by the Covid-19 pandemic have impacted MSMEs earnings by 20-50 percent, micro and small enterprises faced the maximum heat, mainly due to liquidity crunch. Enterprises working in essential commodity business were better off in terms of interrupted but predictable cash flows. This study is an effort to examine how covid 19 affected the small business sectors in Ernakulam , with the help of statistical methods. Thus, the study provides deeper insights into the impacts of COVID - 19 pandemic on small business.

**Definition of Small Business :** A business which functions on a small scale level involves less capital investment, less number of labour and fewer machines to operates is known as a Small Business.

## 1.1 SIGNIFICANCE OF STUDY

The statistical study is relevant as it tries to address the significant changes in the small business sectors before and during the pandemic. We also get an insight about how some enterprises innovated their ways by shifting focus from non essential commodities towards essential commodities; like production of hand sanitizers, toiletries, PPE kits, reusable masks etc. are able to survive in tough time.

## 1.2 OBJECTIVES

1. To study the impact of COVID-19 on the number of small business owners.
2. To compare the sales volume of small businesses before and during COVID-19.
3. To compare the annual profit of small businesses before and during COVID-19.

## Chapter 2

# DATA DESCRIPTION

---

The data used in this study is the primary data collected directly from the small business owners/employees through an online form and also by conducting in-person surveys.

The variables under consideration are

- Period at which the business started (before /during pandemic)
- Area of business (urban/rural)
- Mode of taking orders (directly/through social media/ others)
- Whether the business closed down during pandemic.
- Monthly profit (in rupees)
- Number of orders per month
- Factors (availability of raw materials/delivery and transportation/online marketing/government restrictions) affecting the business.
- Degree of impact (Affected positively/Affected negatively/ Not affected)

## Chapter 3

# METHODOLOGY

---

This study is done with the help of primary data.

The primary data is collected through surveys. A well structured questionnaire is to be formed and circulated among the small business owners in Ernakulam using online form and door to door data collection.

The target population of the survey was the rural and urban population of the Ernakulam district. In the present situation the use of online mode for conducting surveys are found to be a great help.

Questionnaire concerned with the increase and decrease in profit and number of orders before and during the pandemic along with the factors which may be effected the business are included in the questionnaire. From the collected data the sales volume and annual income are calculated.

The survey questionnaire which was circulated among the small business owners in the Ernakulam district received 107 responses out of which 27 were found to be inappropriate and were excluded further. Collected data was interpreted using bar charts and tables and analyzed using Z-test and Chi square test.

### 3.1 PERCENTAGE CHANGE

A percentage analysis shows that how two items changed as percentage from one period to another period.

$\frac{\text{change in the performance}}{\text{the base value}} \times 100 \rightarrow \text{denotes the percentage change}$

### 3.2 TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION

#### Z-TEST

##### 3.2.1 ASSUMPTIONS

1. Sampling distribution of the test statistic is normal.
2. Sample statistics are close to the population parameter and therefore finding standard error, sample statistics are used in place where population parameter are to be used.

##### 3.2.2 DEFINITION

A Z-test is a statistical test to determine whether two population means are different when the variances are known and the sample size is large. A Z-test is a hypothesis test in which the Z-statistics follows a distribution. A Z-statistics, or Z-score is a number representing the result from the Z-test.

The formula is given by;

$$Z = \frac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1$  and  $\bar{x}_2$  are the means of the two samples,  $\Delta$  is the hypothesized difference between the population means (0 if testing for equal means), and  $s_1$  and  $s_2$  are the standard deviations of the two samples, and  $n_1$  and  $n_2$  are the sizes of the two samples.

Null Hypothesis  $\rightarrow H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $\rightarrow H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

### **3.2.3 USES**

1. To test the given population mean when the sample is large or when the population SD is known.
2. To determine whether two population means are different when the variances are known and the sample size is large.
3. To test the equality of two samples standard deviations when the samples are large or when the population standard deviations are known.
4. To test population proportions.
5. To test the equality of two sample proportions.
6. To test the population SD when the sample is large.
7. To test the equality of the correlation coefficients.

## **3.3 CHI SQUARE TEST**

The test we use to measure the differences between what is observed and what is expected based on an assumed hypothesis is called the **Chi-Square Test**.

### **3.3.1 ASSUMPTIONS**

Both variables are catagorical:

It is assumed that both both variables are categorical. That is, both variables take on value that are names or lables.

### 3.3.2 PROCEDURE

1. First of all set the hypotheses that the variables are independent as  $H_0$  and that there is a relationship between the variables as alternative hypothesis  $H_1$ .
2. Tabulation of observed values and calculation of expected values by the formula

$$E_{ij} = \frac{i^{\text{th}} \text{ row total} \times j^{\text{th}} \text{ column total}}{\text{Grand total}}$$

where  $E_{ij}$  is the expected value in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column.

3. Then the test statistic for the Chi-Square Test of Independence denotes as  $\chi^2$  is computed as:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where O is the observed value and E is the expected value.

4. Finally the calculated value of  $\chi^2$  is compared to the critical value from the  $\chi^2$  distribution table with degrees of freedom  $df=(R-1)(C-1)$  (where R is the number of rows and C is the number of columns) and chosen confidence level. If the calculated  $\chi^2$  value is greater than the critical  $\chi^2$  value, then we reject the null hypothesis.

### 3.3.3 APPLICATIONS

1. To test the goodness of fit of distributions.
2. To test the independence of attributes.
3. Test of Homogeneity.

## Chapter 4

# DATA ANALYSIS

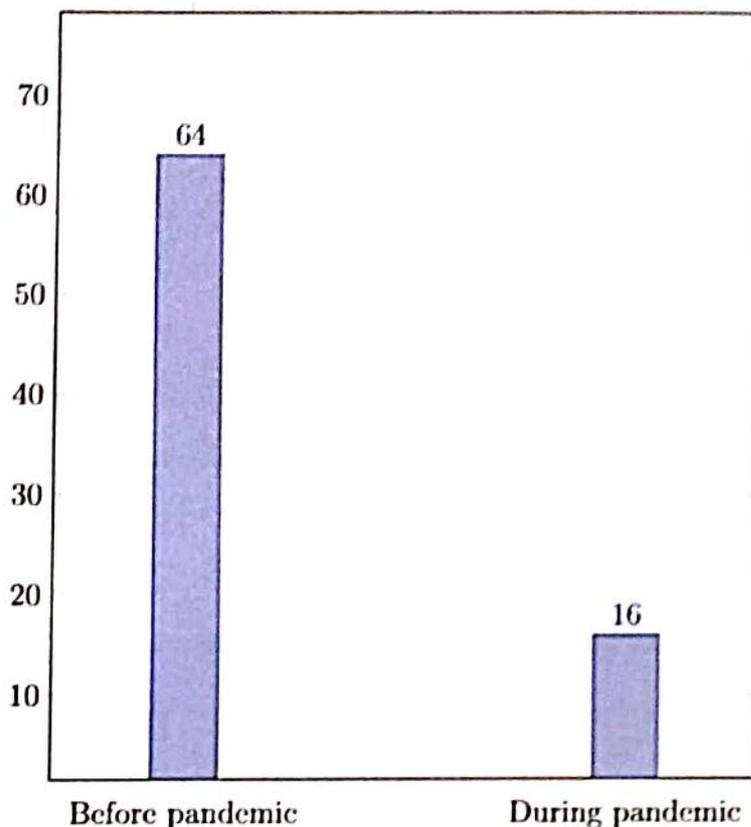
---

### 4.1 GRAPHICAL ANALYSIS

#### **Impact on the Number small Business Owners :**

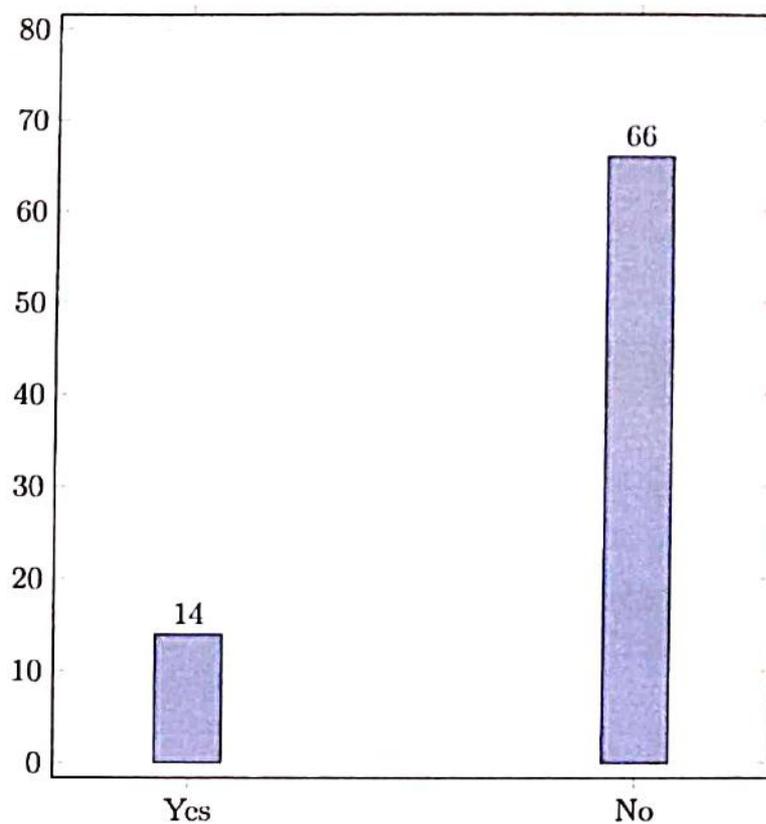
The Number of small business owners varies before and during Pandemic. It is analysed from 80 responses that, 64 small businesses started before Pandemic period and 16 small businesses started during the Pandemic period.

When was the business started	Number
Before Pandemic	64
During Pandemic	16
Total	80



The respondents were asked whether their business closed down during the pandemic period. It is observed that from 80 respondents, 14 respondents' business closed down during the pandemic period.

Did your business closed down ?	Number
Yes	14
No	66
Total	80

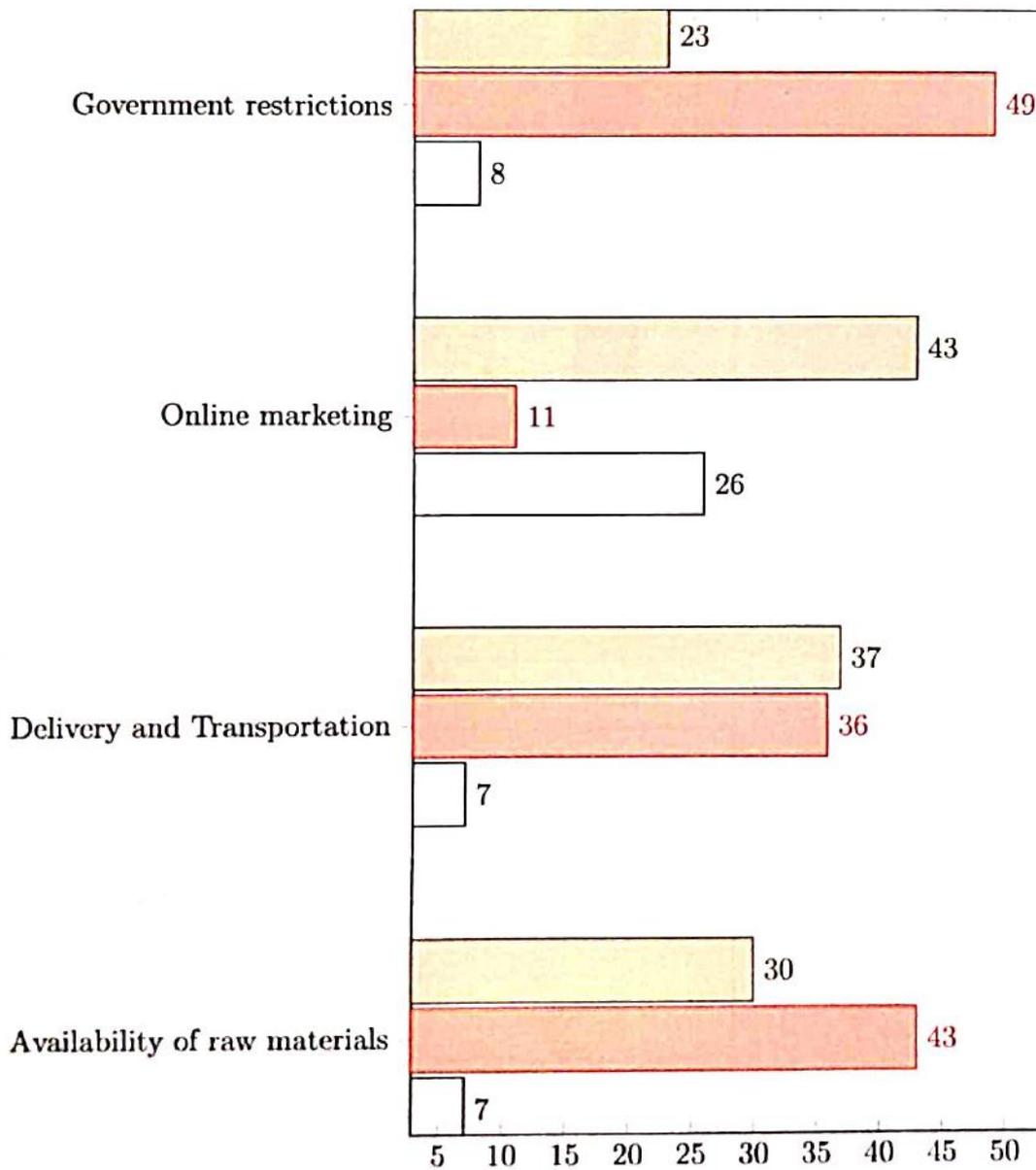


**Percentage of change in Number of Owners :**

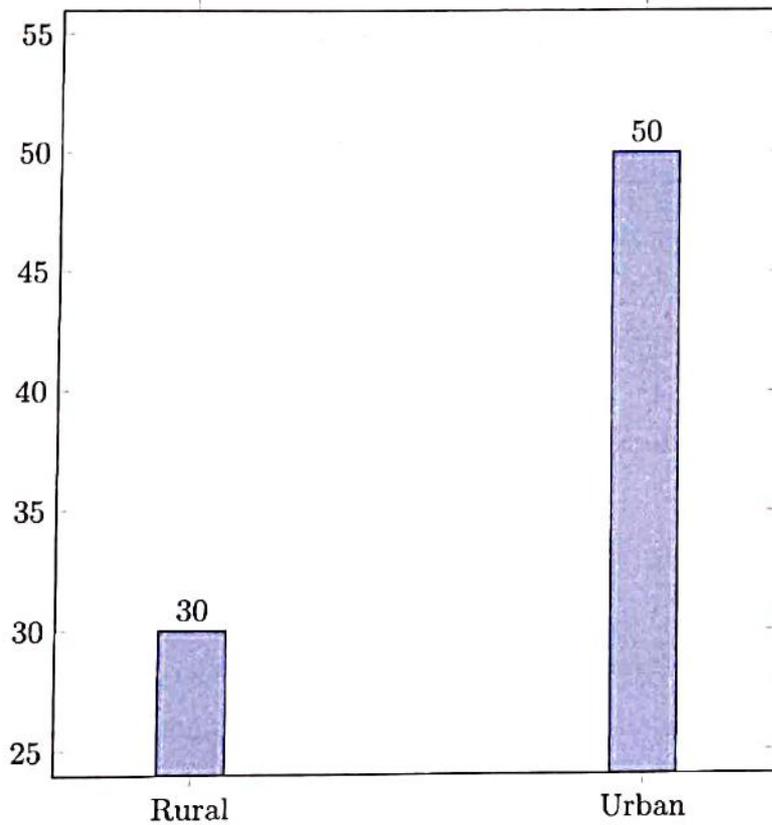
Percentage change = 17.5 %

Hence, out of 80 small businesses, 82.5 % survived during pandemic.

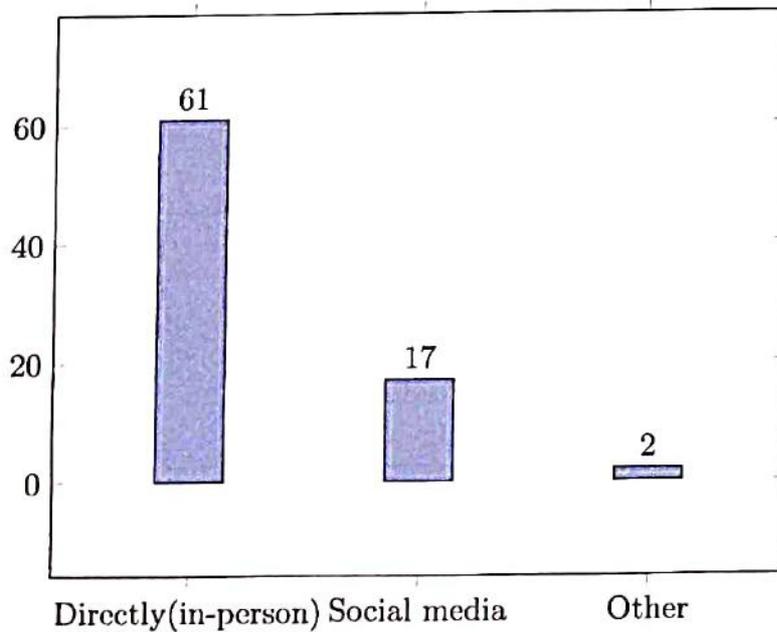
**Effect of various Factors on small Business:**



**Region of Business :**



**Mode of taking Orders :**



More than 60 % of the small business owners responded sell their products directly only.

## 4.2 STATISTICAL ANALYSIS

### Comparing the Sales Volumes :

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(x_1 - x_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 4492.38$ ,  $\bar{x}_2 = 1244.25$ ,  $\Delta = 0$ ,  $s_1^2 = 49505032.11^2$ ,  $s_2^2 = 4466765.51^2$ ,  $n_1 = 63$  and  $n_2 = 80$

Substituting,

$$Z = \frac{(4492.38 - 1244.25) - 0}{\sqrt{\frac{49505032.11^2}{63^2} + \frac{4466765.51^2}{80^2}}} = 3.54$$

Level of significance,  $\alpha = 0.05$

P ( $Z = < z$ ) one tail = 0.0002

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Sales volume during pandemic.

**Comparing the Annual Profits :**

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 293898.3$ ,  $\bar{x}_2 = 147309$ ,  $\Delta = 0$ ,  $s_1^2 = 53289861751^2$ ,  $s_2^2 = 25309730834^2$ ,  $n_1 = 63$   
and  $n_2 = 80$

Substituting,

$$Z = \frac{(293898.3 - 147309) - 0}{\sqrt{\frac{53289861751^2}{63^2} + \frac{25309730834^2}{80^2}}} = 3.75$$

Level of significance,  $\alpha = 0.05$

$P(Z = < z)$  one tail = 0.00009

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Annual profit during pandemic.

### Analysing the Effect of various factors on small business :

The following factors were studied in relation to the degree of impact they had on small Businesses during the Pandemic. The data is interpreted using a contingency table and examined using Chi-square Test.

Observed values( $O_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	7	43	30	80
Delivery and Transportation	7	36	37	80
Online Marketing	26	11	43	80
Government restrictions	8	49	23	80
Total	48	139	133	320

Expected values( $E_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	12	34.75	33.25	80
Delivery and Transportation	12	34.75	33.25	80
Online Marketing	12	34.75	33.25	80
Government restrictions	12	34.75	33.25	80
Total	48	139	133	320

$(O_i - E_i)^2 / E_i$ values			
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected
Availability of Raw materials	2.0833	1.9586	0.3177
Delivery and Transportation	2.0833	0.0449	0.4229
Online Marketing	16.3333	16.2320	2.8590
Government restrictions	1.3333	5.8435	3.1598

From the Table, we get

$$\chi^2 = 52.67186852$$

$$df = 6$$

$$p\text{-value} = 1.36605 \times 10^{-9}$$

Therefore, there is a relation between factors and degree of impact. Thus the shortage of raw materials and transportation restrictions due to lockdowns have brought the small businesses down during the pandemic.

## Chapter 5

# CONCLUSION

---

The sudden rise for the price and less availability of raw materials affected the small business owner's financial recession and also lead to the close down of their businesses during the pandemic period. For the owners who had enough stocks for their business even during the pandemic time, they could not sell off their products to their usual target level as people depended online business modes more during the pandemic time.

Expecting such pandemic situations again, small businesses shall be equipped with precautions to withstand them. Expanding business to online mode and using the advantages of online marketing will be lead to development in sales. Adapting shift in business strategy (i.e., including relevant products into the existing business) might help in improving the business in difficult times.

## REFERENCES

---

- [1] Indrakumar, D. COVID-19 and its impact on micro, small and medium enterprises in India, Manpower journal, Vol. LIV, No. 34, p. 14, 2020
- [2] <https://www.smechamberofindia.com/about-msme-in-india.php>.
- [3] [https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on\\_x0002\\_micro-small-and-medium-enterprises-msmes/](https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on_x0002_micro-small-and-medium-enterprises-msmes/).
- [4] [https://www.mondaq.com/india/operational-impacts-and\\_x0002\\_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in\\_x000\\_india](https://www.mondaq.com/india/operational-impacts-and_x0002_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in_x000_india).

## ANNEXURE

Questionnaire for the Study :

1. What product are you selling ?

2. When was the business started ?

- a. Before pandemic
- b. During pandemic

3. Area of business :

- a. Rural
- b. Urban

4. How do you take orders ? (multiple option selection given)

- a. Directly (in-person)
- b. Through social media (Whatsapp, Facebook, Instagram)
- c. Other (specify, if any):

5. Did your business close down during pandemic ?

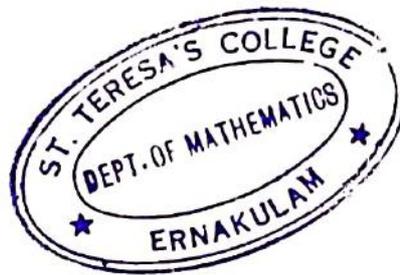
- a. Yes
- b. No

6. Specify the number: (Please only enter numbers. Put '0' if not applicable)

	Before pandemic	During pandemic
Number of orders (per month)		
Monthly profit		

7. How did the following factors affect your business during pandemic?

	Affected positively	Affected negatively	Partially affected positively	Partially affected negatively	Not affected
Availability of raw materials					
Delivery and Transportation					
Online marketing					
Government restrictions					



Project Report

On

**A STUDY ON THE EFFECT OF COVID-19  
ON SMALL BUSINESSES**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**SHABANA N**

(Register No. AB19BMAT004)

*Under the Supervision of*

**DR. ELIZABETH RESHMA M T**



**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)  
ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

Project Report

On

**A STUDY ON THE EFFECT OF COVID-19  
ON SMALL BUSINESSES**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**SHABANA N**

(Register No. AB19BMAT004)

*Under the Supervision of*

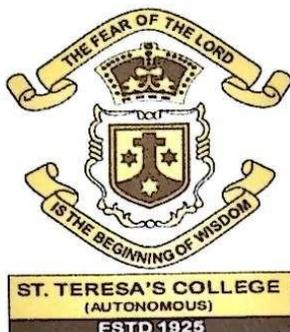
**DR. ELIZABETH RESHMA M T**



**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)  
ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

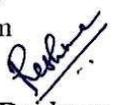
ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

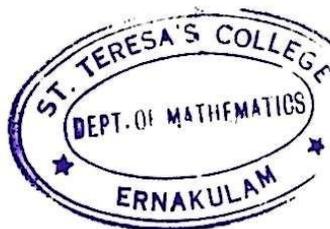


CERTIFICATE

This is to certify that the dissertation entitled, **A STUDY ON THE EFFECT OF COVID-19 ON SMALL BUSINESSES** is a bonafide record of the work done by Ms. **SHABANA N** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

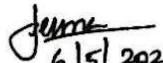
Date: 08-03-2022  
Place: Ernakulam

  
**Dr. Elizabeth Reshma M.T**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



  
**Dr. Ursala Paul**  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1:..........  
6/5/2022

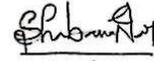
2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Dr. Elizabeth Reshma M.T, Assistant Professor, Department of Mathematics, St. Teresa's College (Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 08-03-2022



SHABANA N

AB19BMAT004

## ACKNOWLEDGEMENT

Firstly, I thank God Almighty for His grace for being able to complete this project work successfully. I express my deep sense of gratitude to our guide Dr. Elizabeth Reshma M.T (Assistant Professor), Department of Mathematics and Statistics, St. Teresa's College, Ernakulam, for her valuable guidance and suggestions. I would like to mention the unending help and support provided by Ms. Shanty B.P (H.O.D of Statistics), Ms. Reshmy S (Assistant professor) and Ms. Rosmin Raju throughout the course of project. I also thank Dr. Ursala Paul ( H.O.D of Mathematics), other teachers of the department, parents, friends, especially my group members and all those who gave me the moral support. This project would not have been possible without the support of the people mentioned above.

Ernakulam.

Date: 08-03-2022

SHABANA N

AB19BMAT004

# Contents

<i>CERTIFICATE</i> .....	ii
<i>DECLARATION</i> .....	iii
<i>ACKNOWLEDGEMENT</i> .....	iv
<i>CONTENT</i> .....	v
1 INTRODUCTION .....	1
1.1 SIGNIFICANCE OF STUDY .....	2
1.2 OBJECTIVES .....	2
2 DATA DESCRIPTION .....	3
3 METHODOLOGY .....	4
3.1 PERCENTAGE CHANGE .....	5
3.2 TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION .....	5
3.2.1 ASSUMPTIONS .....	5
3.2.2 DEFINITION .....	5
3.2.3 USES .....	6
3.3 CHI SQUARE TEST .....	6
3.3.1 ASSUMPTIONS .....	6
3.3.2 PROCEDURE .....	7
3.3.3 APPLICATIONS .....	7
4 DATA ANALYSIS .....	8
4.1 GRAPHICAL ANALYSIS .....	8
4.2 STATISTICAL ANALYSIS .....	13

5 CONCLUSION	17
REFERENCES . . . . .	18
ANNEXURE . . . . .	19

## Chapter 1

# INTRODUCTION

The first paragraph of the introduction discusses the importance of the study and the objectives of the research. It highlights the need for a comprehensive understanding of the subject matter and the role of the researcher in this process. The text emphasizes the significance of the findings and the potential impact on the field of study.

The second paragraph provides a brief overview of the methodology used in the study. It describes the data collection methods, the sample size, and the analytical techniques employed. The author explains how these methods were chosen to address the research objectives and ensure the validity and reliability of the results.

The third paragraph discusses the structure of the report and the organization of the chapters. It outlines the main sections of the study, including the literature review, the theoretical framework, the empirical analysis, and the conclusions. The author provides a clear roadmap for the reader, indicating the flow of the argument and the key points to be covered in each chapter.

The fourth paragraph concludes the introduction by summarizing the main points and reiterating the significance of the study. It expresses the author's hope that the findings will contribute to the existing knowledge and provide a solid foundation for further research in the field.

## Chapter 1

# INTRODUCTION

---

COVID-19 put an immediate halt to many business activities across the globe, as several countries had shut down their ports, airports and domestic transportation while imposing nation-wide lockdowns, leading to a disturbance in business and civil life. The enormity of the lockdown affected manufacturing activities and supply chains alike, disrupting the overall economy of India. Like any other major sector, the Micro, Small and Medium Enterprise (MSME) sector also witnessed a considerable decline in economic activities and loss of jobs due to the nationwide lockdown. Declining output is not a good sign for the economy as the sector provides gainful non- farm employment to millions, especially in rural areas. MSME sector has emerged as a very important sector of the Indian economy, contributing significantly to employment generation, innovation, exports and inclusive growth of the economy. MSMEs are the backbone of the socio-economic development of our country. It also accounts for 45 percentage of the total industrial production, 40 percent of total exports and contributes very significantly to the GDP. Manufacturing segment within the MSME contributes to 7.09 percent of GDP. MSME also contributes to 30 percent of services. The total contribution of MSME to the GDP is 37.54. The announcement of country wide lockdown dragged MSME owners in unexpected times, where no one had experience to handle this kind of situation. Extended lockdown had negative impact on supply of finished goods, procurement of raw material and availability of employees to work in production and supply processes. During April to June 2020, sector faced challenges related to debt repayments, wages/salaries, statutory dues, etc. Reports have shown

that disruptions caused by the Covid-19 pandemic have impacted MSMEs earnings by 20-50 percent, micro and small enterprises faced the maximum heat, mainly due to liquidity crunch. Enterprises working in essential commodity business were better off in terms of interrupted but predictable cash flows. This study is an effort to examine how covid 19 affected the small business sectors in Ernakulam , with the help of statistical methods. Thus, the study provides deeper insights into the impacts of COVID - 19 pandemic on small business.

**Definition of Small Business :** A business which functions on a small scale level involves less capital investment, less number of labour and fewer machines to operates is known as a Small Business.

## **1.1 SIGNIFICANCE OF STUDY**

The statistical study is relevant as it tries to address the significant changes in the small business sectors before and during the pandemic. We also get an insight about how some enterprises innovated their ways by shifting focus from non essential commodities towards essential commodities; like production of hand sanitizers, toiletries, PPE kits, reusable masks etc. are able to survive in tough time.

## **1.2 OBJECTIVES**

1. To study the impact of COVID-19 on the number of small business owners.
2. To compare the sales volume of small businesses before and during COVID-19.
3. To compare the annual profit of small businesses before and during COVID-19.

## Chapter 2

# DATA DESCRIPTION

---

The data used in this study is the primary data collected directly from the small business owners/employees through an online form and also by conducting in-person surveys.

The variables under consideration are

- Period at which the business started (before /during pandemic)
- Area of business (urban/rural)
- Mode of taking orders (directly/through social media/ others)
- Whether the business closed down during pandemic.
- Monthly profit (in rupees)
- Number of orders per month
- Factors (availability of raw materials/delivery and transportation/online marketing/government restrictions) affecting the business.
- Degree of impact (Affected positively/Affected negatively/ Not affected)

## Chapter 3

# METHODOLOGY

---

This study is done with the help of primary data.

The primary data is collected through surveys. A well structured questionnaire is to be formed and circulated among the small business owners in Ernakulam using online form and door to door data collection.

The target population of the survey was the rural and urban population of the Ernakulam district. In the present situation the use of online mode for conducting surveys are found to be a great help.

Questionnaire concerned with the increase and decrease in profit and number of orders before and during the pandemic along with the factors which may be effected the business are included in the questionnaire. From the collected data the sales volume and annual income are calculated.

The survey questionnaire which was circulated among the small business owners in the Ernakulam district received 107 responses out of which 27 were found to be inappropriate and were excluded further. Collected data was interpreted using bar charts and tables and analyzed using Z-test and Chi square test.

### 3.1 PERCENTAGE CHANGE

A percentage analysis shows that how two items changed as percentage from one period to another period.

$\frac{\text{change in the performance}}{\text{the base value}} \times 100 \rightarrow$  denotes the percentage change

### 3.2 TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION

#### Z-TEST

##### 3.2.1 ASSUMPTIONS

1. Sampling distribution of the test statistic is normal.
2. Sample statistics are close to the population parameter and therefore finding standard error, sample statistics are used in place where population parameter are to be used.

##### 3.2.2 DEFINITION

A Z-test is a statistical test to determine whether two population means are different when the variances are known and the sample size is large. A Z-test is a hypothesis test in which the Z-statistics follows a distribution. A Z-statistics, or Z-score is a number representing the result from the Z-test.

The formula is given by;

$$Z = \frac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1$  and  $\bar{x}_2$  are the means of the two samples,  $\Delta$  is the hypothesized difference between the population means (0 if testing for equal means), and  $s_1$  and  $s_2$  are the standard deviations of the two samples, and  $n_1$  and  $n_2$  are the sizes of the two samples.

Null Hypothesis  $\rightarrow H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $\rightarrow H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

### 3.2.3 USES

1. To test the given population mean when the sample is large or when the population SD is known.
2. To determine whether two population means are different when the variances are known and the sample size is large.
3. To test the equality of two samples standard deviations when the samples are large or when the population standard deviations are known.
4. To test population proportions.
5. To test the equality of two sample proportions.
6. To test the population SD when the sample is large.
7. To test the equality of the correlation coefficients.

## 3.3 CHI SQUARE TEST

The test we use to measure the differences between what is observed and what is expected based on an assumed hypothesis is called the **Chi-Square Test**.

### 3.3.1 ASSUMPTIONS

Both variables are categorical:

It is assumed that both both variables are categorical. That is, both variables take on value that are names or lables.

### 3.3.2 PROCEDURE

1. First of all set the hypotheses that the variables are independent as  $H_0$  and that there is a relationship between the variables as alternative hypothesis  $H_1$ .
2. Tabulation of observed values and calculation of expected values by the formula

$$E_{ij} = \frac{i^{\text{th}} \text{ row total} \times j^{\text{th}} \text{ column total}}{\text{Grand total}}$$

where  $E_{ij}$  is the expected value in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column.

3. Then the test statistic for the Chi-Square Test of Independence denotes as  $\chi^2$  is computed as:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where O is the observed value and E is the expected value.

4. Finally the calculated value of  $\chi^2$  is compared to the critical value from the  $\chi^2$  distribution table with degrees of freedom  $df = (R-1)(C-1)$  (where R is the number of rows and C is the number of columns) and chosen confidence level. If the calculated  $\chi^2$  value is greater than the critical  $\chi^2$  value, then we reject the null hypothesis.

### 3.3.3 APPLICATIONS

1. To test the goodness of fit of distributions.
2. To test the independence of attributes.
3. Test of Homogeneity.

## Chapter 4

# DATA ANALYSIS

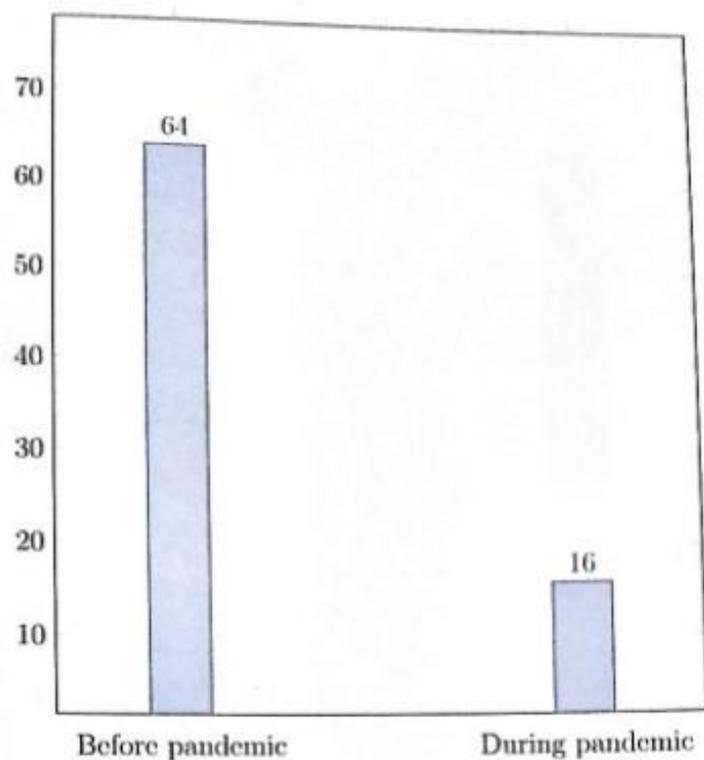
---

### 4.1 GRAPHICAL ANALYSIS

#### Impact on the Number small Business Owners :

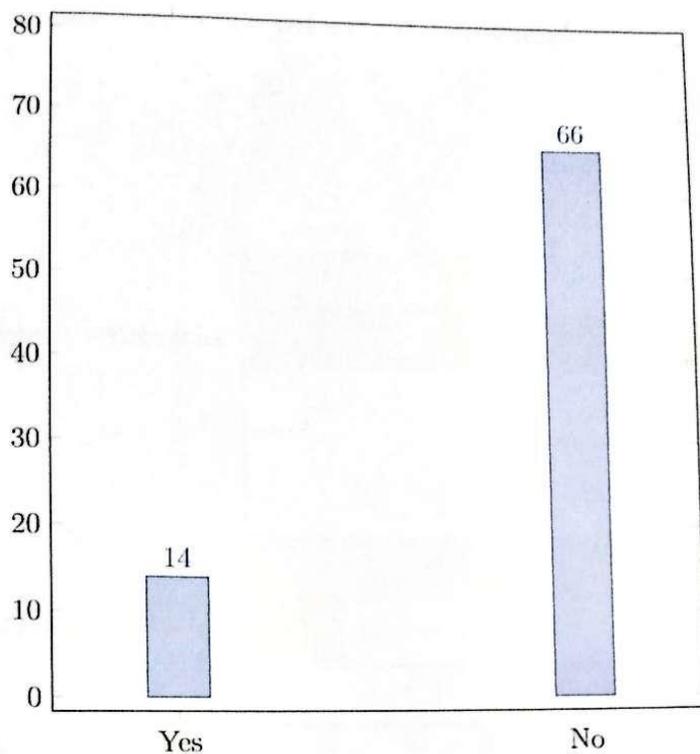
The Number of small business owners varies before and during Pandemic. It is analysed from 80 responses that, 64 small businesses started before Pandemic period and 16 small businesses started during the Pandemic period.

When was the business started	Number
Before Pandemic	64
During Pandemic	16
Total	80



The respondents were asked whether their business closed down during the pandemic period. It is observed that from 80 respondents, 14 respondents' business closed down during the pandemic period.

Did your business closed down ?	Number
Yes	14
No	66
Total	80

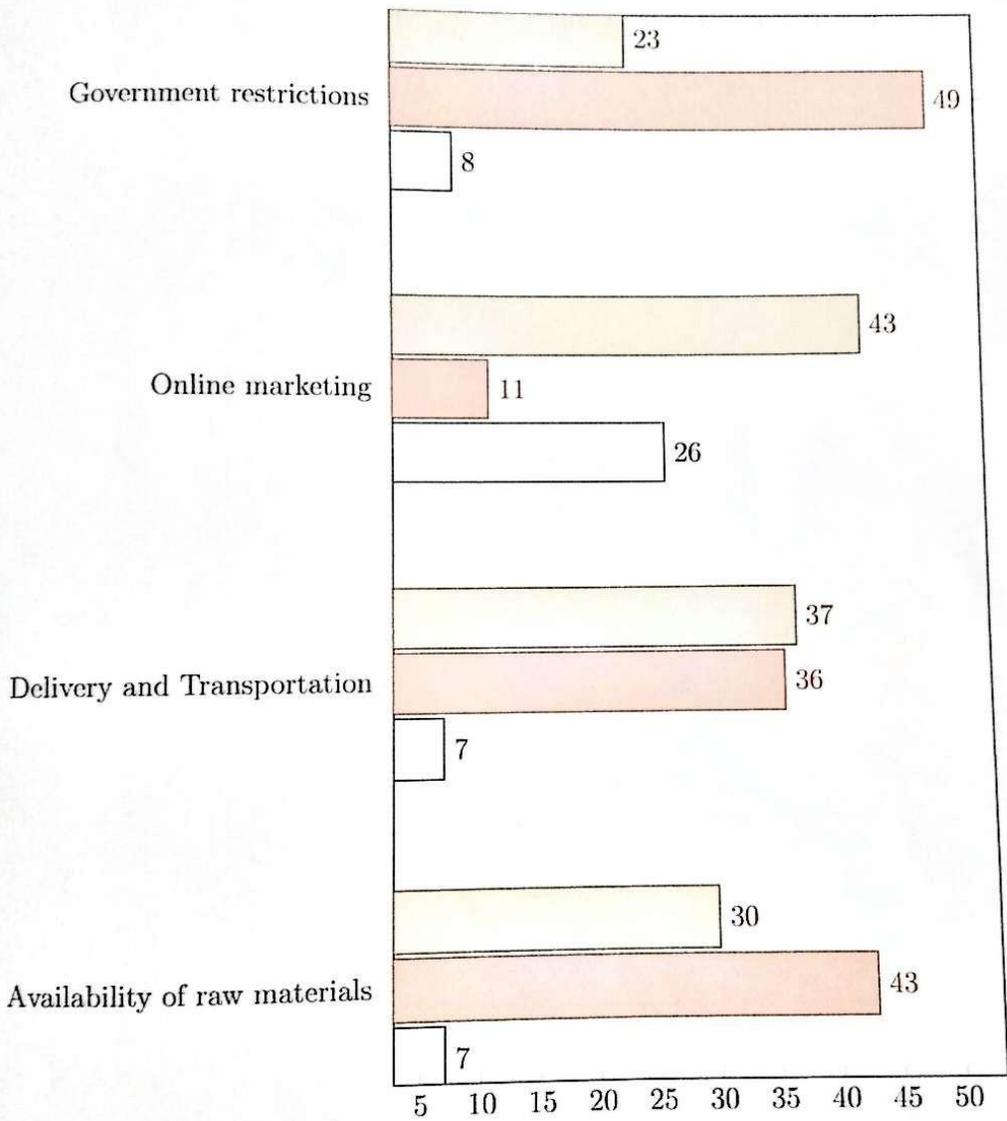


**Percentage of change in Number of Owners :**

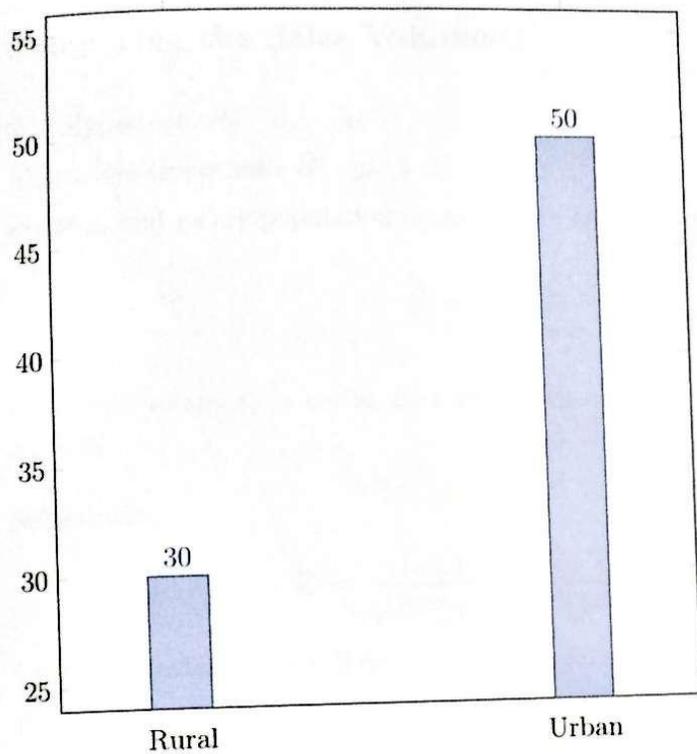
Percentage change = 17.5 %

Hence, out of 80 small businesses, 82.5 % survived during pandemic.

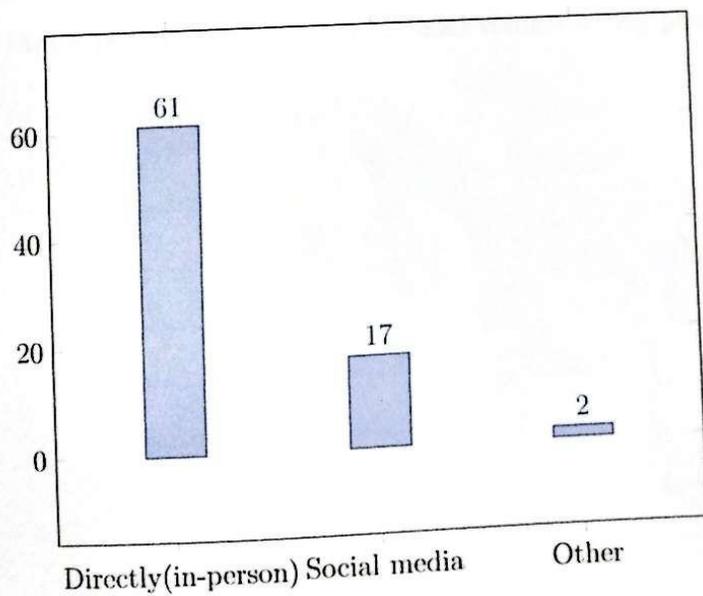
**Effect of various Factors on small Business:**



**Region of Business :**



**Mode of taking Orders :**



More than 60 % of the small business owners responded sell their products directly only.

## 4.2 STATISTICAL ANALYSIS

### Comparing the Sales Volumes :

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(x_1 - x_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 4492.38$ ,  $\bar{x}_2 = 1244.25$ ,  $\Delta = 0$ ,  $s_1^2 = 49505032.11^2$ ,  $s_2^2 = 4466765.51^2$ ,  $n_1 = 63$  and  $n_2 = 80$

Substituting,

$$Z = \frac{(4492.38 - 1244.25) - 0}{\sqrt{\frac{49505032.11^2}{63^2} + \frac{4466765.51^2}{80^2}}} = 3.54$$

Level of significance,  $\alpha = 0.05$

$P(Z = < z)$  one tail = 0.0002

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Sales volume during pandemic.

### Comparing the Annual Profits :

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 293898.3$ ,  $\bar{x}_2 = 147309$ ,  $\Delta = 0$ ,  $s_1^2 = 53289861751^2$ ,  $s_2^2 = 25309730834^2$ ,  $n_1 = 63$   
and  $n_2 = 80$

Substituting,

$$Z = \frac{(293898.3 - 147309) - 0}{\sqrt{\frac{53289861751^2}{63^2} + \frac{25309730834^2}{80^2}}} = 3.75$$

Level of significance,  $\alpha = 0.05$

$P(Z = < z)$  one tail = 0.00009

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Annual profit during pandemic.

### Analysing the Effect of various factors on small business :

The following factors were studied in relation to the degree of impact they had on small Businesses during the Pandemic. The data is interpreted using a contingency table and examined using Chi-square Test.

Observed values( $O_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	7	43	30	80
Delivery and Transportation	7	36	37	80
Online Marketing	26	11	43	80
Government restrictions	8	49	23	80
Total	48	139	133	320

Expected values( $E_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	12	34.75	33.25	80
Delivery and Transportation	12	34.75	33.25	80
Online Marketing	12	34.75	33.25	80
Government restrictions	12	34.75	33.25	80
Total	48	139	133	320

Factors / Degree of Impact	$(O_i - E_i)^2 / E_i$ values		
	Affected positively	Affected negatively	Not affected
Availability of Raw materials	2.0833	1.9586	0.3177
Delivery and Transportation	2.0833	0.0449	0.4229
Online Marketing	16.3333	16.2320	2.8590
Government restrictions	1.3333	5.8435	3.1598

From the Table, we get

$$\chi^2 = 52.67186852$$

$$df = 6$$

$$p\text{-value} = 1.36605 \times 10^{-9}$$

Therefore, there is a relation between factors and degree of impact. Thus the shortage of raw materials and transportation restrictions due to lockdowns have brought the small businesses down during the pandemic.

## Chapter 5

# CONCLUSION

---

The sudden rise for the price and less availability of raw materials affected the small business owner's financial recession and also lead to the close down of their businesses during the pandemic period. For the owners who had enough stocks for their business even during the pandemic time, they could not sell off their products to their usual target level as people depended online business modes more during the pandemic time.

Expecting such pandemic situations again, small businesses shall be equipped with precautions to withstand them. Expanding business to online mode and using the advantages of online marketing will be lead to development in sales. Adapting shift in business strategy (i.e., including relevant products into the existing business) might help in improving the business in difficult times.

## REFERENCES

---

- [1] Indrakumar, D. COVID-19 and its impact on micro, small and medium enterprises in India, Manpower journal, Vol. LIV, No. 34, p. 14, 2020
- [2] <https://www.smechamberofindia.com/about-msme-in-india.php>.
- [3] [https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on\\_x0002\\_micro-small-and-medium-enterprises-msmes/](https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on_x0002_micro-small-and-medium-enterprises-msmes/).
- [4] [https://www.mondaq.com/india/operational-impacts-and\\_x0002\\_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in\\_x000\\_india](https://www.mondaq.com/india/operational-impacts-and_x0002_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in_x000_india).

## ANNEXURE

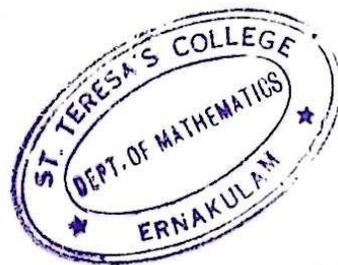
Questionnaire for the Study :

1. What product are you selling ?
  
2. When was the business started ?
  - a. Before pandemic
  - b. During pandemic
  
3. Area of business :
  - a. Rural
  - b. Urban
  
4. How do you take orders ? (multiple option selection given)
  - a. Directly (in-person)
  - b. Through social media (Whatsapp, Facebook, Instagram)
  - c. Other (specify, if any):
  
5. Did your business close down during pandemic ?
  - a. Yes
  - b. No
  
6. Specify the number: (Please only enter numbers. Put '0' if not applicable)

	Before pandemic	During pandemic
Number of orders (per month)		
Monthly profit		

7. How did the following factors affect your business during pandemic?

	Affected positively	Affected negatively	Partially affected positively	Partially affected negatively	Not affected
Availability of raw materials					
Delivery and Transportation					
Online marketing					
Government restrictions					



Project Report

On

# SIR MODEL AND IT'S APPLICATION

*Submitted*

*in partial fulfilment of the requirements for the degree of*

BACHELOR OF SCIENCE

*in*

MATHEMATICS

*by*

SREELAKSHMI P B

(Register No. AB19BMAT009)

*Under the Supervision of*

Dr. Ursala Paul



DEPARTMENT OF MATHEMATICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



CERTIFICATE

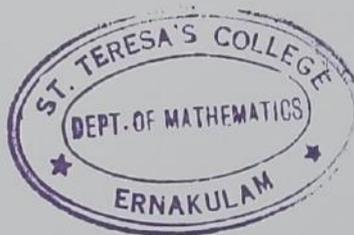
This is to certify that the dissertation entitled, **SIR MODEL AND IT'S APPLICATION** is a bonafide record of the work done by Ms. **SREELAKSHMI P B** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 08/03/2022

Place: Ernakulam

**Dr. Ursala Paul**

Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



**Dr. Ursala Paul**

Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1: Jeema Jon June  
6/5/2022

2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Dr. Ursala Paul, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

**SREELAKSHMI P B**

Date: 08/03/2022

**AB19BMAT009**

## ACKNOWLEDGEMENT

Firstly I thank God Almighty for giving me his grace to execute this project work successfully. I express my deep sense of gratitude to our guide Dr. Ursala Paul, H.O.D of the Department of Mathematics and Statistics, St. Teresa's College, Ernakulam, for her valuable guidance and suggestions. I would like to mention the unending help and support provided by Smt. Reshmy S and Dr. Susan Mathew Panakkal of the Department of Mathematics and Statistics throughout the course of the project.

I also thank other teachers of the Department, parents, friends especially my group members and all those who give me the moral support and helped me to complete my project. This project would not have been possible without the support of the people mentioned above.

Ernakulam.

Date: 08/03/2022

SREELAKSHMI P B

AB19BMAT009

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 INTRODUCTION</b>	<b>1</b>
<b>2 SIR MODEL</b>	<b>2</b>
2.1 What is SIR Model? . . . . .	2
2.2 Differential equations for sir model . . . . .	4
2.3 Will the disease spread? . . . . .	6
2.4 What is the maximum number of infectives at any given period of time? . . . . .	8
2.5 At the end of the Pandemic . . . . .	10
2.6 2.6.What does this mean for Covid 19 ? . . . . .	12
<b>3 APPLICATION</b>	<b>13</b>
<b>4 CONCLUSION</b>	<b>20</b>
<b>5 REFERENCES</b>	<b>22</b>

# Chapter 1

## INTRODUCTION

---

Mathematical modelling of infectious diseases was initiated by Bernoulli in 1760. The work of Kermack and McKendrick, published in 1927, had a major influence on the modelling framework. The SIR model aims to predict the number of individuals who are susceptible to infection, are actively infected, or have recovered from infection at any given time. SIR is a model, a technique, fundamental to the science of epidemiology — the branch of medicine that investigates the start, spread and control of diseases. It's medicine, but not a clinical pursuit like surgery or ophthalmology; instead, it delves into the data about a disease and looks for patterns. Epidemiologists do this by mathematically modelling the disease. One way to do so is to “compartmentalize” the population that's been affected by the disease. The idea is that each compartment has similar characteristics as far as the disease is concerned, and we can draw conclusions about each one, as well as about how they relate to the others. Epidemiologists have used compartment models for years to make vital predictions about epidemics like measles and polio. This is how we can estimate how many people will be infected, how quickly the infections will spread through a population, how long the pandemic will last, what effect various measures — vaccination, hospitalization, etc — will have on the disease, and more. Understanding the disease through such data and models helps clarify the best way we can use always limited resources to fight the disease.

## Chapter 2

# SIR MODEL

---

### 2.1 What is SIR Model?

SIR is really a modification of a simpler model, SIR. SIR divides the population into three compartments. “S”, for those who are susceptible to the disease, likely a significant chunk of the population. “I” refers to people who are infectious; meaning they can pass on the infection to someone else. “R” is the set of people who have either recovered from the disease or have died, and thus can no longer spread the disease like their compatriots in “I” can. (Thus “R” is sometimes said to stand for “removed”, meaning people who have been removed from the infectious population). Of course, these are not fixed numbers. Instead, they vary as the disease progresses. It’s that very variation that the model seeks to simulate, though we can say some things about it simply by observing and inferring. In fact, that’s exactly how we flesh out the before a disease breaks out in a country, “S” is the whole population, because we can assume everyone is susceptible. Then one person gets infected. If she in turn infects every person she meets as she moves about, you can imagine how quickly the disease spreads — because each of those infected people will also move about and infect others. What this means is, as the disease spreads, “S” decreases rapidly, because susceptible people get infected and move from the “S” basket to the “I” basket. Thus “I”’s numbers rise dramatically. But of course this cannot carry on indefinitely, because those who get infected will eventually recover. (And some will probably die). If they have developed an immunity to the disease, they cannot return to the “I” basket. So “I” eventually reaches a peak and starts

decreasing. In tandem, "R"'s numbers grow until most of the population is in that compartment. At that point, the disease has run its course.

Mathematical model from its simple beginnings. Mathematicians consider this basic description of an epidemic's spread and set out to model it. They will seek to capture the change in the numbers in each compartment as a function of elapsed time, but also as it relates to the change in other compartments. After all, the decrease in "S" is a direct consequence of the increase in "I"; and the rise in "R" has consequences for both "I" and "S". Taking all this into account produces a set of what mathematicians call "differential equations". This set is the SIR model.

## 2.2 Differential equations for sir model

Let's make assumptions that:

1. The epidemic is sufficiently short or doesn't last that long. Total population remains constant.
2. The ways in which disease are transmitted. The rate of increase in I is proportional to the contact between susceptibles and infectives.
3. Infectives recover or die at a constant rate.

Rate of change of susceptible with time,

$$\frac{dS}{dt} = -rIS$$

where,

r = Rate of contact

I = Number of infectives

S = Number of susceptibles

Also, the rate of change of susceptibles decreases as more and more people gets infected.

Rate of change of I with time,

$$\frac{dI}{dt} = rIS - aI$$

where, a = constant rate

Rate of change of I with time will grow according to people moving from susceptibles to infectives. From infective, they move to the next compartment, i.e., recovered or death.

Rate of change of removed or recovered people with time,

$$\frac{dR}{dt} = aI$$

gain people from infectives.

## 2.2. DIFFERENTIAL EQUATIONS FOR SIR MODEL MODEL AND IT'S APPLICATION

Initially, the values of I and S will be  $I_0$  and  $S_0$ , respectively. The value of R will be 0

$$\begin{aligned}S + R + I &= S_0 + I_0 + 0 \\ &= S_0 + I_0\end{aligned}$$

Since we're taking constant population,

$$\frac{d(S+R+I)}{dt} = 0$$

### 2.3 Will the disease spread?

To know the spread of disease, we have to see if the number of infectives starts to increase from its initial value,  $I_0$ , and if it does then there will be a spread of disease through the population.

As the disease spread to more and more people in the population, the number of people in susceptible keeps decreasing, i.e., the number of susceptibles will be less than the initial value.

$$S \leq S_0$$

In the rate of change of susceptibles equation, the values of  $r$ ,  $S$  and  $I$  all are positive. But, since the number of susceptibles keeps decreasing with time, there is a negative sign.

$$\frac{dS}{dt} = -rIS$$

Now, substitute this inequality in rate of change in infectives equation, we get,

$$\frac{dI}{dt} = rIS - aI$$

$$\frac{dI}{dt} = I(rS - a)$$

$$\frac{dI}{dt} < I(rS_0 - a)$$

The spread of the disease depends on the sign of  $(r S_0 - a)$ .

If its negative, we can conclude that there's no spread of the disease. But if the sign is positive, then we can conclude that the disease is spreading rapidly. That is,

$$S_0 > \frac{a}{r}$$

$$S_0 > \frac{1}{q}$$

where,  $q = \frac{r}{a}$  and is called as contact ratio.

Contact ratio is defined as the fraction of population that comes in contact with an infected individual. If  $q$ , the contact ratio, increases number of infectives also increases. In order to reduce the number of infectives we have to reduce the contact ratio,  $q$ .

Then, rearranging the equation, we get a new parameter,  $R_0$ .

$$R_0 = \frac{r}{a} S_0$$

The new parameter is known as the Basic Reproductive number or Basic Reproductive ratio. If the ratio is  $> 1$ , then we are having an epidemic. The ratio represents the number of secondary infections in the population caused by one initial primary infection. That is, if one person has the disease then  $R_0$  will tell you how many infections, on average, will be caused in the population by that person.

For seasonal flu, the  $R_0$  is between 1.5 – 2. The ratio of the ongoing pandemic is estimated to be between 3 – 4, but it can be greater than that.

## 2.4 What is the maximum number of infectives at any given period of time?

Here we are going to find the maximum number of infectious at any given period of time.

We have,

$$\begin{aligned} \frac{dI}{ds} &= \frac{(rIS - aI)}{(-rIS)} \\ &= -1 + \left(\frac{1}{qS}\right) \end{aligned}$$

By integrating on both sides, we get,

$$I + S - \frac{1}{q} \ln(S) = I_0 + S_0 - \frac{1}{q} [\ln S_0]$$

Using the maximum and minimum function,

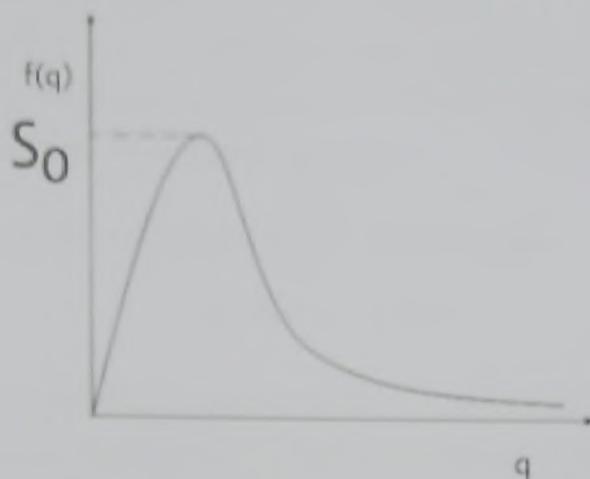
$$I_{(max)} = I_0 + S_0 - \frac{1}{q} [1 + \ln(qS_0)]$$

So the maximum people that will get infected will be,

Total population  $[(S_0 + I_0)]$  - a positive value  $[\left(\frac{1}{q} - (1 + \ln qS_0)\right)]$

Let the positive value be a function,  $f(q) = \frac{1}{q}(1 + \ln qS_0)$

From the above equation, we can see that the contact ratio (q) and the function f(q) depend on each other and is inversely proportional. That is, if q is very large, the function f(q) will be very small. So, we can understand that when maximum number of people are infected at any given time, the total number of population - the function, f(q), is quite small. The maximum number of people that can catch the disease at one point of time is actually the most of the population.



Here, we have a very sharp increase initially and then the function starts to decrease. The peak up here is around  $S_0$ . As contact ratio ' $q$ ' which is a large value increases, the function is going to be quite small. So maximum number of infectives remains the same as we're not subtracting much from the total population.

## 2.5 At the end of the Pandemic

We had assumed in the beginning that the total population is a constant. The disease to end means that the total number of population that caught the disease. At the end of the pandemic, the number of infectives has to decrease to zero. The people who died or recovered from the disease are the people in R compartment. The size of R of the removed compartment at the end of the outbreak will give the total number of people that caught the disease.

We know that the total population is given by the equation,

$$R + S + I = I_o + S_o$$

where,  $I_o$  is the number of infectives and  $S_o$  is the susceptible population at the beginning of the pandemic. Here  $R_o$  is 0 because recovery rate at the beginning is zero.

Rearrange to find the size of R, the recovered component at the end of the outbreak.

$$R_{end} = I_o + S_o - S_{end}$$

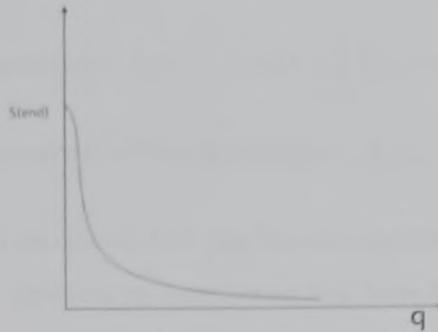
From the above equation,

$$I + S - \frac{1}{q}[1 + \ln(qS_0)] = I_o + S_o - \frac{1}{q}[\ln S_0]$$

$$\text{we get, } S_{(end)} + \frac{1}{q}[\ln(S_{end})] = I_o + S_o - \frac{1}{q}[\ln S_0]$$

Solving this equation we get the value of  $S_{(end)}$ . Substituting this value in the former equation we get the number of removed people or the size of the removed population at the end of the pandemic.

plot a graph with  $S_{end}$  as Y-axis and  $q$  as X-axis



' $q$ ' which is the contact ratio is really a key in controlling the disease outbreak. At far end of the graph, have a very very small value. Since ' $q$ ' is a large value  $S_{(end)}$  is actually going to be quite small. This is a bad news.

$R_{(end)}$  which is the total number of people who catch disease is (total population) -  $S_{(end)}$  remains the same. Because we're not subtracting much from total population as  $S_{(end)}$  is quite small. So majority of the population will catch the disease, if ' $q$ ' is sufficiently large.

## 2.6 2.6.What does this mean for Covid 19 ?

1. Spread  $R_0 = qS_0 > 1$

2.  $I_{(max)} = \text{Total population} - f(q)$        $[f(q) = 1 \frac{1}{q}(1 + \ln q S_0)]$

3. Total removed/recovered = Total population -  $S_{(end)}$

If  $R_0$  is greater than 1 we can say that the disease is spreading. The maximum number of infected people is obtained by subtracting  $f(q)$  from the total population.

That is  $q$  is large then the disease will spread and epidemic will occur

## Chapter 3

# APPLICATION

---

This is an attempt to test the SIR Epidemic Model for COVID-19 spread in India during this pandemic period. The existing COVID-19 data of India were collected from open data base of various websites to fit in the model. The projected curve was drawn based on the realistic data from various open resources till 31 May 2020. The projected curves are changing rapidly because of changing scenario of virus spread and sample testing. The effects of various parameters like testing uncertainty, social-isolation and lockdown effect were taken in consideration before drawing projection curve. The peak value of pandemic is expected in the first week of October 2020. The pandemic may last in the country till the end of March 2021. Strict guidelines are also required to avoid the second wave of COVID-19 after the withdrawal

The SIR (Susceptible-Infected-Removed) model is a simple model of an epidemic of an infectious disease in a large population developed by Kermack and McKendrick [1]. It is based on the assumption that the population consists of three types of individuals:

*Susceptible (S):* Who are not infected but could become infected.

*Infected (I):* These individuals have the disease and can transmit it to the susceptible.

*Removed (R):* These individual may or may not have the disease, but they can't become infected and they can't transmit the disease to others. They may have a natural immunity, or they may have recovered from the disease and are immune

from getting it again, or they may have the disease but are incapable of transmitting it (e.g. persons placed in isolation), or they may have died. The deaths other than deaths from this disease can be neglected. All these are functions of the time,  $t$ , and they change according to a system of differential equations.

New infections occur as a result of contact between infected and susceptibles. In this simple model the rate at which new infections occur is  $\beta IS$ , where  $\beta$  is a constant. When a new infection occurs, the infected individual moves from the susceptible class to the infective class. In this simple model, there is no other way individuals can enter or leave the susceptible class, so the first differential equation is:

$$\frac{ds}{dt} = -\beta IS$$

The other process that can occur is that infective individuals can enter the removed class. It is assumed that this happens at the rate  $\gamma I$  for some positive constant. Thus other two differential equations are

$$\frac{dI}{dt} = \beta IS - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

Here the total population at a given time  $[S(t)+I(t)+R(t)]$  is constant because

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = -\beta IS + (\beta IS - \gamma I) + \gamma I = 0$$

If  $I = 0$ , i.e. there are no infected, the right sides of all are 0, so nothing changes. The model is applicable only if there are some infected. Let us start with a population of almost all susceptibles plus a small number of infected. It is important to find if the number of infected increases substantially, producing an epidemic. If it is an epidemic, how long will it last? The contagiousness of an epidemic is described by reproduction rate,  $R_0$ . If  $R_0 > 1$ , the disease will be transmitted between people, and there may be an outbreak. If  $R_0 = 1$ , the disease is endemic, and where  $R_0 < 1$ , the disease will decline and eventually die out. The potential size of an epidemic is mainly based on the magnitude of the  $R_0$  value for that event. The  $R_0$  value only applies when everyone in a population is completely vulnerable to the disease.

## Model calculations

World Health Organization (WHO) announced COVID-19 outbreak as Pandemic on 12 April 2020 [2]. It has now spread in 215 countries with total cases more than ten millions as on 30 June 2020 [3]. The spread of COVID-19 is also increasing in India. As per Ministry of Health and Family Welfare, Govt of India, total confirmed cases of COVID-19 as on 31 July 2020 are 16,97,054 [4-6]. An effort is made here to test the SIR epidemic model from Indian prospective. Following parameters are taken to test the model based on the realistic data from various open resources:

1. Population of India in 2020 = 1,380,004,000 [7]
2. First COVID-19 case reported in India = 30 January 2020 [8]
3. Total number of infected on first day = 1
4. First infected patient recovered on the date = 20 February 2020 [9]
5. Total number of Infected persons on 20 February 2020 = 3
6. Recovery rate on the day of first recovered person = 0.33
7. Reproduction rate,  $R_0$ , as estimated empirically from existing data upto 25 March 2020 = 1.52
8. Reproduction rate,  $R_0$ , as estimated empirically from existing data for lockdown period from 25 March 2020 to 31 May 2020 = 1.30
9. Total number of infected persons on 31 July 2020 = 16,97,054 [4, 5]

A number of organizations like John-Hopkins University, Massachusetts Institute of Technology, University of Geneva, Singapore University of Technology and Design, Indian Council of Medical Research, etc [10-14] developed the models projecting the spread of COVID-19 through India based on SIR/SEIR epidemic models. Most of these models were used to project the spread of COVID-19 globally. A model focussed on a particular country with single nationwide policy be applied uniformly across the country is always better than that for globally. It implies that the local models are more useful than global models. In present study, SIR Model is used to estimate the pandemic life cycle curves and expected ending dates as per available data from open resources.

The projection based on the data before lockdown (i.e. 25 March 2020) in the country is shown in Fig. 1. The number of infected persons may be very low because of inadequate testing of COVID-19 samples in the beginning. A correction factor is applied with the assumption that the testing should have increased by about five times during the period of pandemic period of about four months. The estimated average value of reproduction rate,  $R_0$ , from 30 January 2020 to 24 March 2020 was 1.52. The peak value of pandemic was expected in the first week of June 2020 without lockdown effect.

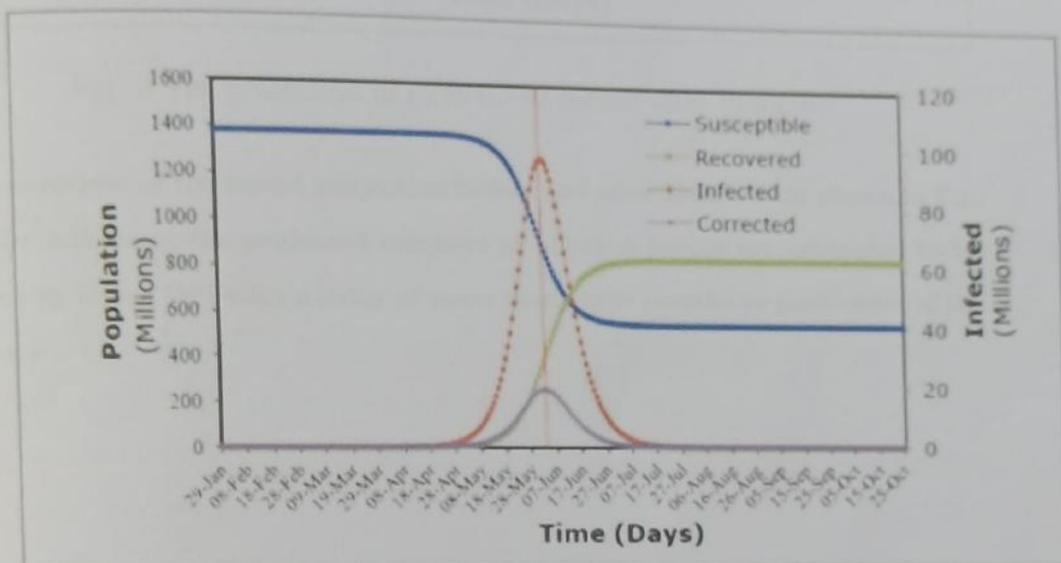


Fig. 1 The projection of COVID-19 spread before lockdown

Fig. 2 shows the estimation of the pandemic life cycle curves and expected ending date with lockdown effect begun from 25 March 2020. The estimated average value of reproduction rate,  $R_0$ , from 25 March 2020 to 20 May 2020 is 1.30. The peak value of pandemic is expected in the first week of October 2020

The projection of COVID-19 spread was estimated based on the actual data available upto 30 June 2020. The lockdown effect was included in estimating the growth of COVID-19 cases on 31 May 2020. The unlock process has started in the country from 01 June 2020 and no further change was made so far.

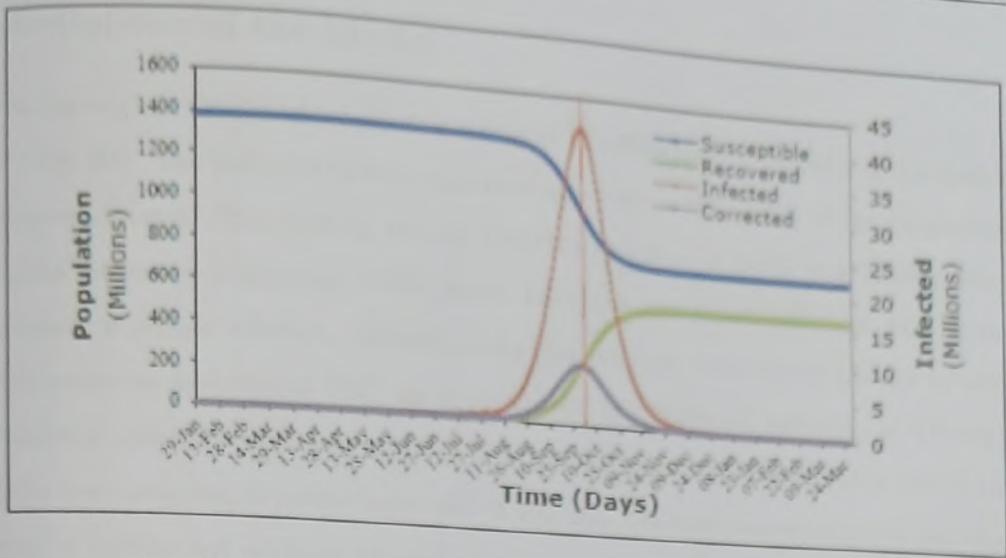


Fig. 2 The projection of COVID-19 spread after lockdown

The comparison of the model projection before and after lockdown is shown in Fig. 3. After lockdown, the projected numbers of infected people are estimated to be reduced by about 50% with a delay of more than three months in peak value of the pandemic.

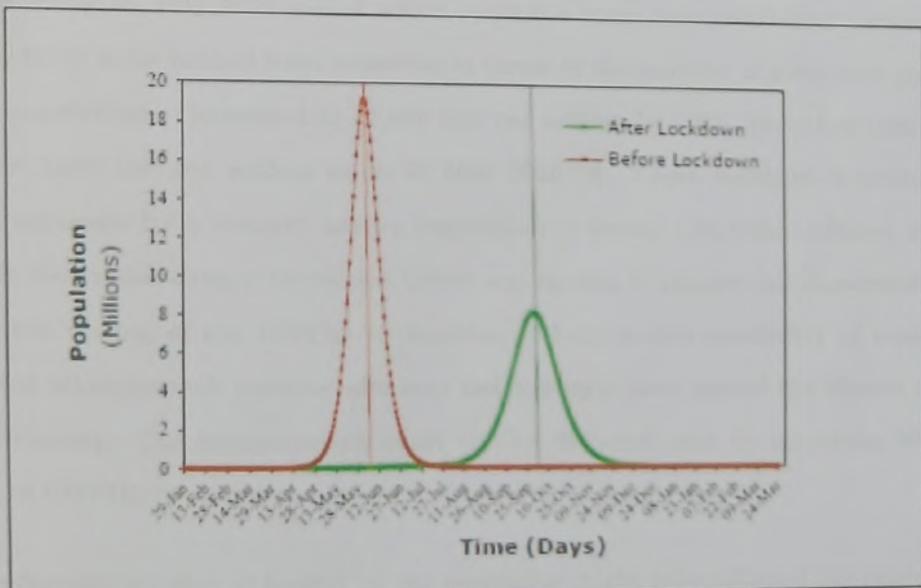


Fig. 3 The comparison of the projected COVID-19 spread before and after lockdown

estimation.

Another important factor is that the interaction rate and infection rate are very dynamic i.e. subject to changes, as quarantine and lock-down measures are implemented. The partial withdrawal of lockdown and mass movement of people during lockdown period have also increased the uncertainty in estimation of interaction rate. The asymptomatic persons may spread the infection during their movement. These parameters collectively may have affected the model calculation. Under these circumstances, the daily counts extracted from the model may not exactly match with the reported cases but the trend of spread may give some valuable information on pandemic life cycle and expected ending dates of COVID-19 in the country. This is now well accepted that the COVID-19 vaccine is the only thing that can bring back the normalcy in life. The lockdown and social distancing may delay the spread of COVID-19 till the development of the vaccine. Strict guidelines are also required to avoid the second wave of COVID-19 after the withdrawal of lockdown.

## Chapter 4

# CONCLUSION

---

Models can be used to predict and understand how an infectious disease spreads in the world and various factors affect the dynamics. Even if the predictions are inaccurate, it has been clear to scientists from many decades that quarantine, social distancing and adoption of very strict health and safety standards are essential to stop the spread of the virus. Quarantine was even implemented in medieval times to fight the black death before there was knowledge of the existence of virus. In this sense, this pandemic reveals the failure of policy makers, since it is well-known from basic modeling results that earlier adoption of those measures can save thousands of lives and even prevent the pandemic. The interface of science, society and politics is still uneasy, even in highly developed countries, revealing a disregard for scientific evidence.

A high number of secondary COVID-19 infections can take place when an infected individual is introduced into a community. It is essential to simulate the process of infection (and death) in advance so as to apply adequate control measures and mitigate the risk of virus diffusion. One of the most commonly used mathematical algorithms to describe the diffusion of an epidemic disease is the SEIR model, which we have applied to compute the number of infected, recovered, and dead individuals on the basis of the number of contacts, probability of disease transmission, incubation and infectious periods, and disease fatality rate.

Application shows that the SIR model is more suitable to predict the epidemic trend due to the spread of the disease as it can accommodate surges and be adjusted to

the recorded data. By comparing the published data with predictions, it is possible to predict the success of government interventions.

The SIR model is one of several types of models that can be used to model an infectious disease epidemic. Although no model can perfectly predict the future, a good model provides an approximation that is accurate enough to be useful for informing public policy.

The application used in this study represents the spread Covid-19 in India . Here we had seen the projection of covid 19 before lockdown and after lockdown based on the actual data .After lockdown, the projected number of infected people were estimated to be reduced .Thus we can realize that contact ratio causes increase in spreading. So through lockdown process contact ratio decreased.This lead to the reduction in number of infectives. So we have to follow some restrictions or methods proposed by our government in order to reduce the spreading. Methods include quarantines; travel restrictions; and the closing of schools, workplaces, stadiums, theatres, or shopping centres. Individuals may apply social distancing methods by staying at home, limiting travel, avoiding crowded areas, using no-contact greetings, and physically distancing themselves from others.Each and everyone should be aware of self protection measures including frequent hand washing or using hand sanitizer ,avoid direct contact with sick people ,use mask properly to cover nose and mouth, use tissues to cover cough and sneeze and clean the touched items using sanitizer regularly.

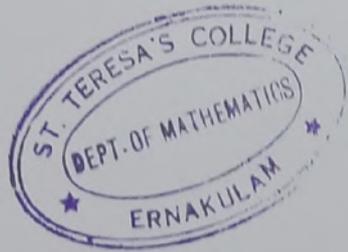
The SIR model used here is only a simple one and thus, the predictions that come out might not be accurate enough, something that also depends on the published data and their trustworthiness. However, as the model data show, one thing that is certain is that COVID-19 is not going to go way quickly or easily

## Chapter 5

# REFERENCES

---

- Ramola, R. C. (2020). Projected Spread of COVID-19 through India Based on SIR Model. *Reproduction*, 1-52.
- Tolles, J., Luong, T. (2020). Modeling epidemics with compartmental models. *Jama*, 323(24), 2515-2516.
- Smith, D., Moore, L. (2004). The SIR model for spread of disease-the differential equation model. *Convergence*.
- Cooper, I., Mondal, A., Antonopoulos, C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons Fractals*, 139, 110057.
- Expanding SIR Model; <https://youtu.be/NKMHhm2Zbkw>



Project Report

On

**A STUDY ON THE EFFECT OF COVID-19  
ON SMALL BUSINESSES**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**ROSHNI UNNIKANNAN**

(Register No. AB19BMAT002)

*Under the Supervision of*

**DR. ELIZABETH RESHMA M T**



**DEPARTMENT OF MATHEMATICS**

**ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

Project Report

On

A STUDY ON THE EFFECT OF COVID-19  
ON SMALL BUSINESSES

*Submitted*

*in partial fulfilment of the requirements for the degree of*

BACHELOR OF SCIENCE

*in*

MATHEMATICS

*by*

ROSHNI UNNIKANNAN

(Register No. AB19BMAT002)

*Under the Supervision of*

DR. ELIZABETH RESHMA M T



DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

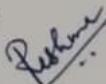
ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

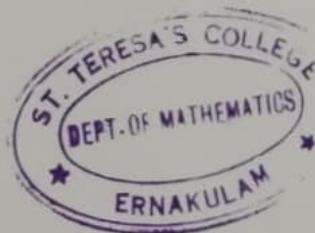


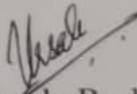
CERTIFICATE

This is to certify that the dissertation entitled, **A STUDY ON THE EFFECT OF COVID-19 ON SMALL BUSINESSES** is a bonafide record of the work done by Ms. **ROSHNI UNNIKANNAN** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

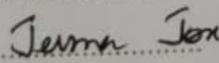
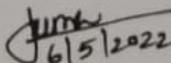
Date: 08-03-2022  
Place: Ernakulam

  
**Dr. Elizabeth Reshma M.T**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



  
**Dr. Ursala Paul**  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1:.....   6/5/2022

2: .....

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **Dr. Elizabeth Reshma M.T.**, Assistant Professor, Department of Mathematics, St. Teresa's College (Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

Date: 08-03-2022



**ROSHNI UNNIKANNAN**

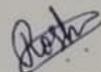
**AB19BMAT002**

## ACKNOWLEDGEMENT

Firstly, I thank God Almighty for His grace for being able to complete this project work successfully. I express my deep sense of gratitude to our guide Dr. Elizabeth Reshma M.T (Assistant Professor), Department of Mathematics and Statistics, St. Teresa's College, Ernakulam, for her valuable guidance and suggestions. I would like to mention the unending help and support provided by Ms. Shanty B.P (H.O.D of Statistics), Ms. Reshmy S (Assistant professor) and Ms. Rosmin Raju throughout the course of project. I also thank Dr. Ursala Paul ( H.O.D of Mathematics), other teachers of the department, parents, friends, especially my group members and all those who gave me the moral support. This project would not have been possible without the support of the people mentioned above.

Ernakulam.

Date: 08-03-2022



ROSHNI UNNIKANNAN

AB19BMAT002

# Contents

	<i>CERTIFICATE</i> . . . . .	ii
	<i>DECLARATION</i> . . . . .	iii
	<i>ACKNOWLEDGEMENT</i> . . . . .	iv
	<i>CONTENT</i> . . . . .	v
1	INTRODUCTION . . . . .	1
1.1	SIGNIFICANCE OF STUDY . . . . .	2
1.2	OBJECTIVES . . . . .	2
2	DATA DESCRIPTION . . . . .	3
3	METHODOLOGY . . . . .	4
3.1	PERCENTAGE CHANGE . . . . .	5
3.2	TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION . . . . .	5
3.2.1	ASSUMPTIONS . . . . .	5
3.2.2	DEFINITION . . . . .	5
3.2.3	USES . . . . .	6
3.3	CHI SQUARE TEST . . . . .	6
3.3.1	ASSUMPTIONS . . . . .	6
3.3.2	PROCEDURE . . . . .	7
3.3.3	APPLICATIONS . . . . .	7
4	DATA ANALYSIS . . . . .	8
4.1	GRAPHICAL ANALYSIS . . . . .	8
4.2	STATISTICAL ANALYSIS . . . . .	13

5	CONCLUSION	17
	<i>REFERENCES</i> . . . . .	18
	<i>ANNEXURE</i> . . . . .	19

## Chapter I

# INTRODUCTION

# Chapter 1

## INTRODUCTION

---

COVID-19 put an immediate halt to many business activities across the globe, as several countries had shut down their ports, airports and domestic transportation while imposing nation-wide lockdowns, leading to a disturbance in business and civil life. The enormity of the lockdown affected manufacturing activities and supply chains alike, disrupting the overall economy of India. Like any other major sector, the Micro, Small and Medium Enterprise (MSME) sector also witnessed a considerable decline in economic activities and loss of jobs due to the nationwide lockdown. Declining output is not a good sign for the economy as the sector provides gainful non- farm employment to millions, especially in rural areas. MSME sector has emerged as a very important sector of the Indian economy, contributing significantly to employment generation, innovation, exports and inclusive growth of the economy. MSMEs are the backbone of the socio-economic development of our country. It also accounts for 45 percentage of the total industrial production, 40 percent of total exports and contributes very significantly to the GDP. Manufacturing segment within the MSME contributes to 7.09 percent of GDP. MSME also contributes to 30 percent of services. The total contribution of MSME to the GDP is 37.54. The announcement of country wide lockdown dragged MSME owners in unexpected times, where no one had experience to handle this kind of situation. Extended lockdown had negative impact on supply of finished goods, procurement of raw material and availability of employees to work in production and supply processes. During April to June 2020, sector faced challenges related to debt repayments, wages/salaries, statutory dues, etc. Reports have shown

that disruptions caused by the Covid-19 pandemic have impacted MSMEs earnings by 20-50 percent, micro and small enterprises faced the maximum heat, mainly due to liquidity crunch. Enterprises working in essential commodity business were better off in terms of interrupted but predictable cash flows. This study is an effort to examine how covid 19 affected the small business sectors in Ernakulam , with the help of statistical methods. Thus, the study provides deeper insights into the impacts of COVID - 19 pandemic on small business.

**Definition of Small Business :** A business which functions on a small scale level involves less capital investment, less number of labour and fewer machines to operates is known as a Small Business.

## 1.1 SIGNIFICANCE OF STUDY

The statistical study is relevant as it tries to address the significant changes in the small business sectors before and during the pandemic. We also get an insight about how some enterprises innovated their ways by shifting focus from non essential commodities towards essential commodities; like production of hand sanitizers, toiletries, PPE kits, reusable masks etc. are able to survive in tough time.

## 1.2 OBJECTIVES

1. To study the impact of COVID-19 on the number of small business owners.
2. To compare the sales volume of small businesses before and during COVID-19.
3. To compare the annual profit of small businesses before and during COVID-19.

## Chapter 2

# DATA DESCRIPTION

---

The data used in this study is the primary data collected directly from the small business owners/employees through an online form and also by conducting in-person surveys.

The variables under consideration are

- Period at which the business started (before /during pandemic)
- Area of business (urban/rural)
- Mode of taking orders (directly/through social media/ others)
- Whether the business closed down during pandemic.
- Monthly profit (in rupees)
- Number of orders per month
- Factors (availability of raw materials/delivery and transportation/online marketing/government restrictions) affecting the business.
- Degree of impact (Affected positively/Affected negatively/ Not affected)

## Chapter 3

# METHODOLOGY

---

This study is done with the help of primary data.

The primary data is collected through surveys. A well structured questionnaire is to be formed and circulated among the small business owners in Ernakulam using online form and door to door data collection.

The target population of the survey was the rural and urban population of the Ernakulam district. In the present situation the use of online mode for conducting surveys are found to be a great help.

Questionnaire concerned with the increase and decrease in profit and number of orders before and during the pandemic along with the factors which may be effected the business are included in the questionnaire. From the collected data the sales volume and annual income are calculated.

The survey questionnaire which was circulated among the small business owners in the Ernakulam district received 107 responses out of which 27 were found to be inappropriate and were excluded further. Collected data was interpreted using bar charts and tables and analyzed using Z-test and Chi square test.

### 3.1 PERCENTAGE CHANGE

A percentage analysis shows that how two items changed as percentage from one period to another period.

$\frac{\text{change in the performance}}{\text{the base value}} \times 100 \rightarrow \text{denotes the percentage change}$

### 3.2 TESTING EQUALITY OF MEANS BASED ON PAIRED OBSERVATION

#### Z-TEST

##### 3.2.1 ASSUMPTIONS

1. Sampling distribution of the test statistic is normal.
2. Sample statistics are close to the population parameter and therefore finding standard error, sample statistics are used in place where population parameter are to be used.

##### 3.2.2 DEFINITION

A Z-test is a statistical test to determine whether two population means are different when the variances are known and the sample size is large. A Z-test is a hypothesis test in which the Z-statistics follows a distribution. A Z-statistics, or Z-score is a number representing the result from the Z-test.

The formula is given by;

$$Z = \frac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1$  and  $\bar{x}_2$  are the means of the two samples,  $\Delta$  is the hypothesized difference between the population means (0 if testing for equal means), and  $s_1$  and  $s_2$  are the standard deviations of the two samples, and  $n_1$  and  $n_2$  are the sizes of the two samples.

Null Hypothesis  $\rightarrow H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $\rightarrow H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

### 3.2.3 USES

1. To test the given population mean when the sample is large or when the population SD is known.
2. To determine whether two population means are different when the variances are known and the sample size is large.
3. To test the equality of two samples standard deviations when the samples are large or when the population standard deviations are known.
4. To test population proportions.
5. To test the equality of two sample proportions.
6. To test the population SD when the sample is large.
7. To test the equality of the correlation coefficients.

## 3.3 CHI SQUARE TEST

The test we use to measure the differences between what is observed and what is expected based on an assumed hypothesis is called the **Chi-Square Test**.

### 3.3.1 ASSUMPTIONS

Both variables are catagorical:

It is assumed that both both variables are categorical. That is, both variables take on value that are names or lables.

### 3.3.2 PROCEDURE

1. First of all set the hypotheses that the variables are independent as  $H_0$  and that there is a relationship between the variables as alternative hypothesis  $H_1$ .
2. Tabulation of observed values and calculation of expected values by the formula

$$E_{ij} = \frac{i^{\text{th}} \text{ row total} \times j^{\text{th}} \text{ column total}}{\text{Grand total}}$$

where  $E_{ij}$  is the expected value in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column.

3. Then the test statistic for the Chi-Square Test of Independence denotes as  $\chi^2$  is computed as:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where O is the observed value and E is the expected value.

4. Finally the calculated value of  $\chi^2$  is compared to the critical value from the  $\chi^2$  distribution table with degrees of freedom  $df=(R-1)(C-1)$  (where R is the number of rows and C is the number of columns) and chosen confidence level. If the calculated  $\chi^2$  value is greater than the critical  $\chi^2$  value, then we reject the null hypothesis.

### 3.3.3 APPLICATIONS

1. To test the goodness of fit of distributions.
2. To test the independence of attributes.
3. Test of Homogeneity.

## Chapter 4

# DATA ANALYSIS

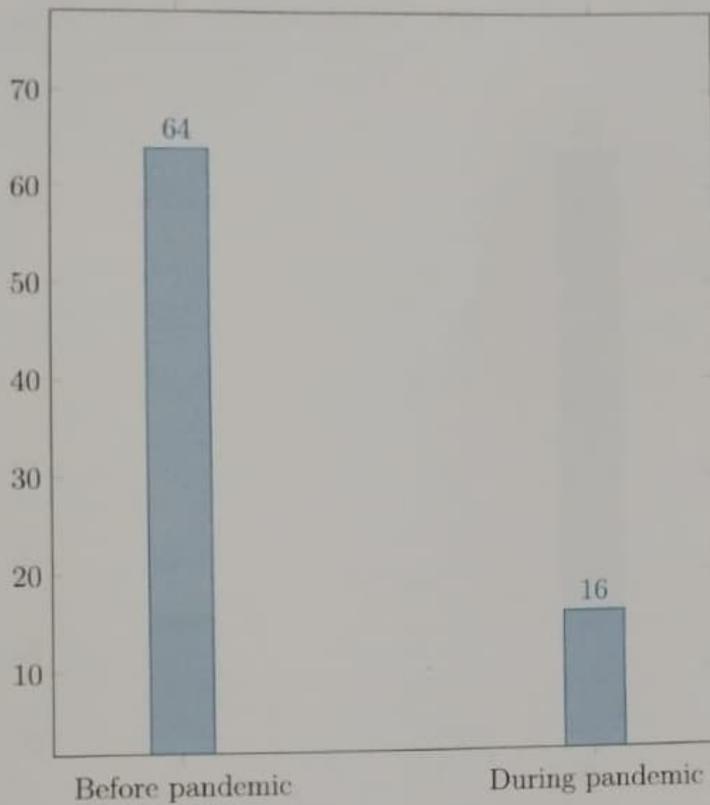
---

### 4.1 GRAPHICAL ANALYSIS

#### Impact on the Number small Business Owners :

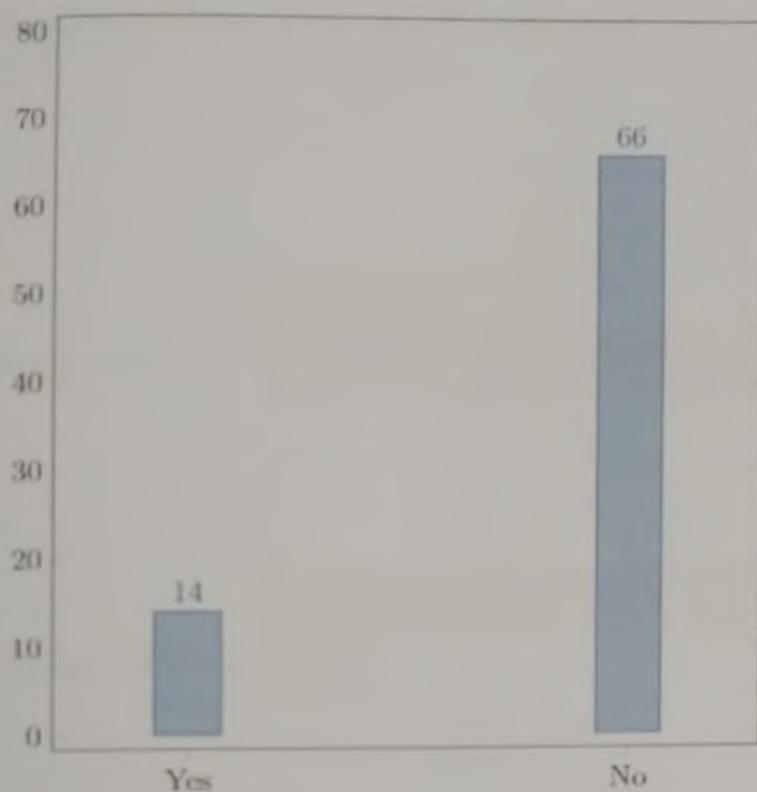
The Number of small business owners varies before and during Pandemic. It is analysed from 80 responses that, 64 small businesses started before Pandemic period and 16 small businesses started during the Pandemic period.

When was the business started	Number
Before Pandemic	64
During Pandemic	16
Total	80



The respondents were asked whether their business closed down during the pandemic period. It is observed that from 80 respondents, 14 respondents' business closed down during the pandemic period.

Did your business closed down ?	Number
Yes	14
No	66
Total	80

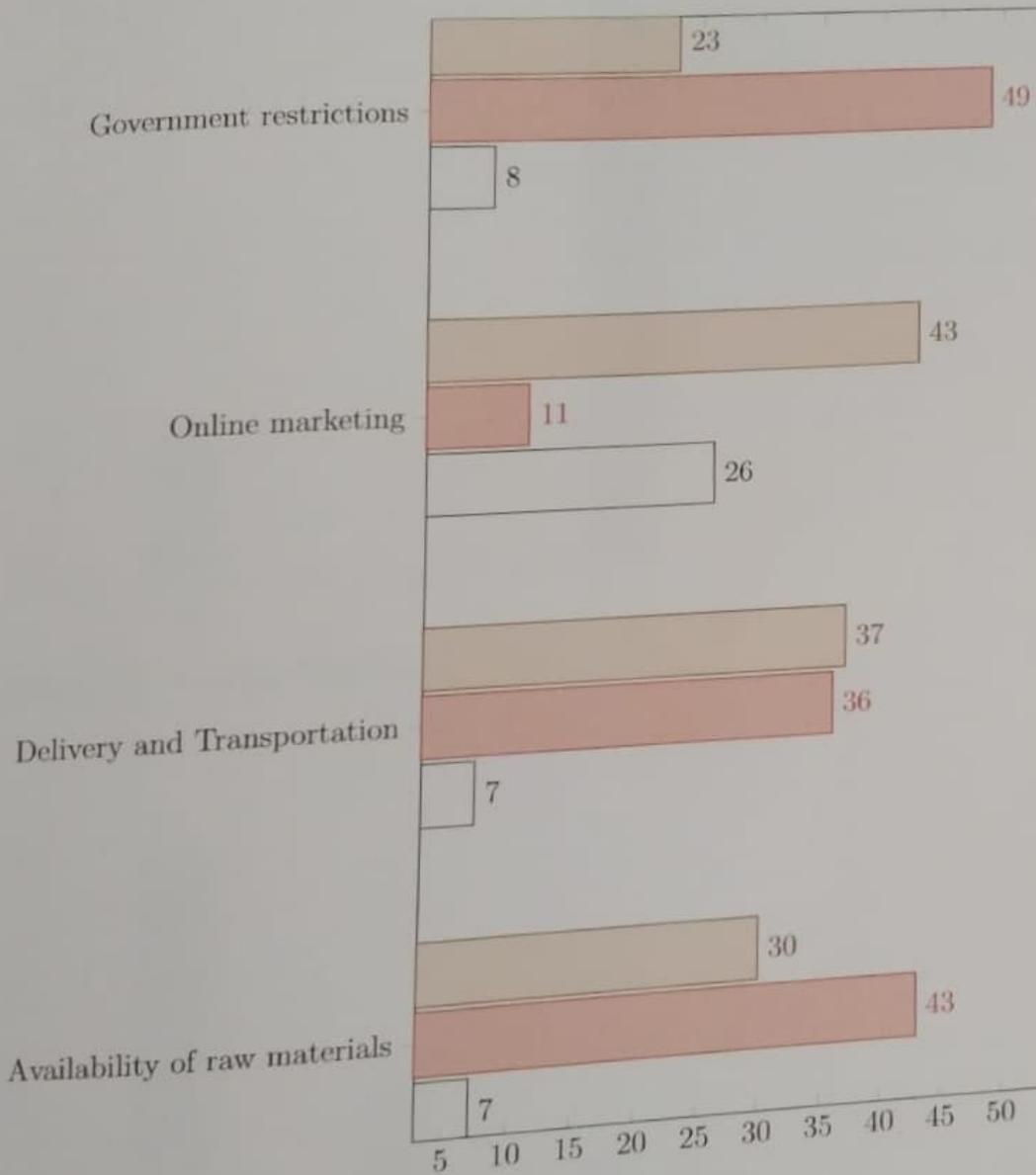


Percentage of change in Number of Owners :

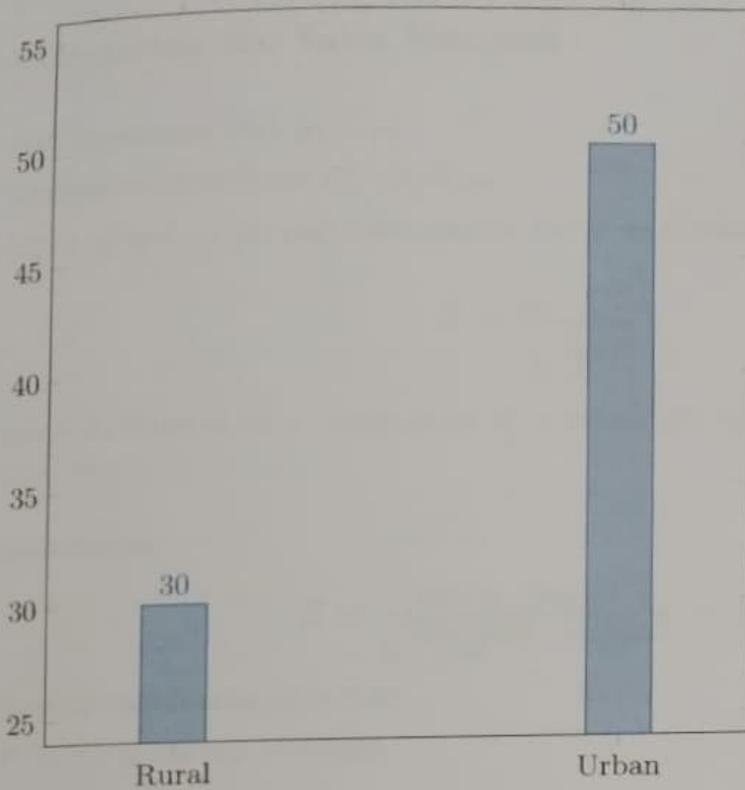
Percentage change = 17.5 %

Hence, out of 80 small businesses, 82.5 % survived during pandemic.

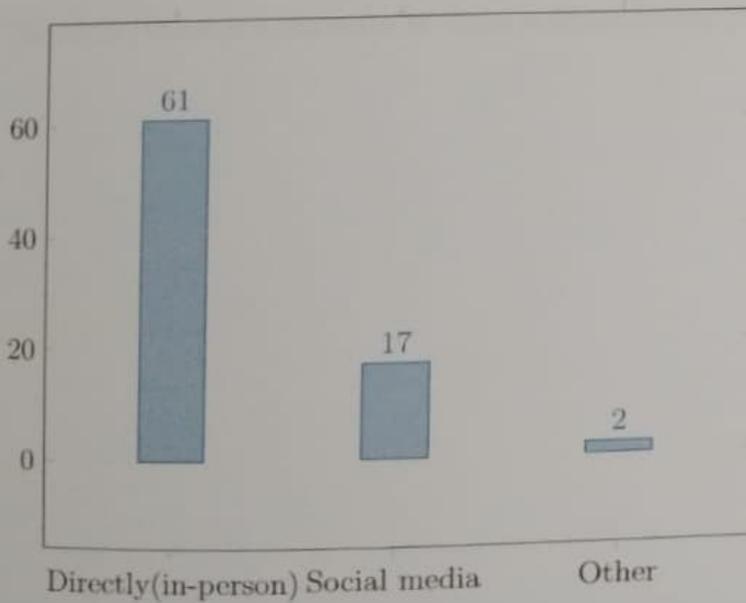
Effect of various Factors on small Business:



**Region of Business :**



**Mode of taking Orders :**



More than 60 % of the small business owners responded sell their products directly only.

## 4.2 STATISTICAL ANALYSIS

### Comparing the Sales Volumes :

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 4492.38$ ,  $\bar{x}_2 = 1244.25$ ,  $\Delta = 0$ ,  $s_1^2 = 49505032.11^2$ ,  $s_2^2 = 4466765.51^2$ ,  $n_1 = 63$  and  $n_2 = 80$

Substituting,

$$Z = \frac{(4492.38 - 1244.25) - 0}{\sqrt{\frac{49505032.11^2}{63^2} + \frac{4466765.51^2}{80^2}}} = 3.54$$

Level of significance,  $\alpha = 0.05$

$P(Z = < z)$  one tail = 0.0002

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Sales volume during pandemic.

### Comparing the Annual Profits :

Null Hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative Hypothesis  $H_1 : \mu_1 > \mu_2$

where  $\mu_1$  and  $\mu_2$  are population means before and during pandemic respectively.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where  $\bar{x}_1 = 293898.3$ ,  $\bar{x}_2 = 147309$ ,  $\Delta = 0$ ,  $s_1^2 = 53289861751^2$ ,  $s_2^2 = 25309730834^2$ ,  $n_1 = 63$

and  $n_2 = 80$

Substituting,

$$Z = \frac{(293898.3 - 147309) - 0}{\sqrt{\frac{53289861751^2}{63^2} + \frac{25309730834^2}{80^2}}} = 3.75$$

Level of significance,  $\alpha = 0.05$

$P(Z = < z)$  one tail = 0.00009

Therefore,  $H_0$  is rejected.

Hence  $\mu_1 > \mu_2$

Therefore, there is a decrease in Annual profit during pandemic.

### Analysing the Effect of various factors on small business :

The following factors were studied in relation to the degree of impact they had on small Businesses during the Pandemic. The data is interpreted using a contingency table and examined using Chi-square Test.

Observed values( $O_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	7	43	30	80
Delivery and Transportation	7	36	37	80
Online Marketing	26	11	43	80
Government restrictions	8	49	23	80
Total	48	139	133	320

Expected values( $E_i$ )				
Factors / Degree of Impact	Affected positively	Affected negatively	Not affected	Total
Availability of Raw materials	12	34.75	33.25	80
Delivery and Transportation	12	34.75	33.25	80
Online Marketing	12	34.75	33.25	80
Government restrictions	12	34.75	33.25	80
Total	48	139	133	320

Factors / Degree of Impact	$(O_i - E_i)^2 / E_i$ values		
	Affected positively	Affected negatively	Not affected
Availability of Raw materials	2.0833	1.9586	0.3177
Delivery and Transportation	2.0833	0.0449	0.4229
Online Marketing	16.3333	16.2320	2.8590
Government restrictions	1.3333	5.8435	3.1598

From the Table, we get

$$\chi^2 = 52.67186852$$

$$df = 6$$

$$p\text{-value} = 1.36605 \times 10^{-9}$$

Therefore, there is a relation between factors and degree of impact. Thus the shortage of raw materials and transportation restrictions due to lockdowns have brought the small businesses down during the pandemic.

## Chapter 5

# CONCLUSION

---

The sudden rise for the price and less availability of raw materials affected the small business owner's financial recession and also lead to the close down of their businesses during the pandemic period. For the owners who had enough stocks for their business even during the pandemic time, they could not sell off their products to their usual target level as people depended online business modes more during the pandemic time.

Expecting such pandemic situations again, small businesses shall be equipped with precautions to withstand them. Expanding business to online mode and using the advantages of online marketing will be lead to development in sales. Adapting shift in business strategy (i.e., including relevant products into the existing business) might help in improving the business in difficult times.

## REFERENCES

---

- [1] Indrakumar, D. COVID-19 and its impact on micro, small and medium enterprises in India, Manpower journal, Vol. LIV, No. 34, p. 14, 2020
- [2] <https://www.smechamberofindia.com/about-msme-in-india.php>.
- [3] [https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on\\_x0002\\_micro-small-and-medium-enterprises-msmes/](https://timesofindia.indiatimes.com/blogs/agyeya/covid-19-affect-on_x0002_micro-small-and-medium-enterprises-msmes/).
- [4] [https://www.mondaq.com/india/operational-impacts-and\\_x0002\\_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in\\_x0002\\_india](https://www.mondaq.com/india/operational-impacts-and_x0002_strategy/1086172/impact-of-covid-19-on-the-manufacturing-sector-in_x0002_india).

# ANNEXURE

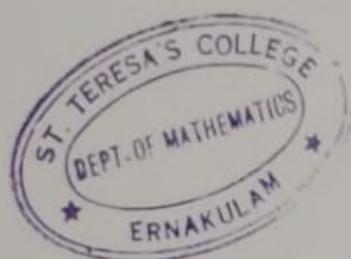
Questionnaire for the Study :

1. What product are you selling ?
  
2. When was the business started ?
  - a. Before pandemic
  - b. During pandemic
  
3. Area of business :
  - a. Rural
  - b. Urban
  
4. How do you take orders ? (multiple option selection given)
  - a. Directly (in-person)
  - b. Through social media (Whatsapp, Facebook, Instagram)
  - c. Other (specify, if any):
  
5. Did your business close down during pandemic ?
  - a. Yes
  - b. No
  
6. Specify the number: (Please only enter numbers. Put '0' if not applicable)

	Before pandemic	During pandemic
Number of orders (per month)		
Monthly profit		

7. How did the following factors affect your business during pandemic?

	Affected positively	Affected negatively	Partially affected positively	Partially affected negatively	Not affected
Availability of raw materials					
Delivery and Transportation					
Online marketing					
Government restrictions					



**RICHARD WRIGHT'S BLACK BOY: A PORTRAIT OF RACISM AND SUBJUGATION**



*Project submitted to St. Teresa's College (Autonomous) in partial fulfilment of the requirement for the degree of BACHELOR OF ARTS in English Language And Literature*

*By*

**SOJA M.S**

**Register No. AB19ENG041**

**III B.A English Literature**

**St.Teresa's College (Autonomous)**

**Ernakulam**

**Cochin - 682 011**

**Kerala**

*Supervisor*

**Ms. Lakshmipriya P Santhosh**

**Department Of English**

**St.Teresa's College (Autonomous)**

**Ernakulam**

**Kerala**

**February 2022**

---

## DECLARATION

I hereby declare that this project entitled' *Richard Wright's Black Boy: A Portrait Of Racism and Subjugation* ' is the record of bonafide work done by me under the guidance and supervision of Ms. Lakshmipriya P Santhosh, Assistant Professor, Department Of English.

Ernakulam  
February 2022

SOJA M.S  
Register Number AB19ENG041  
III B.A ENGLISH LANGUAGE AND LITERATURE  
St.Teresa's College (Autonomous)  
Ernakulam

## CERTIFICATE

I hereby declare that this project entitled '*Richard Wright's Black Boy : A Portrait Of Racism And Subjugation*' by SOJA M.S is a record of bonafide work carried out by her under my supervision and guidance.

Ms.Lakshmipriya P Santhosh  
Department Of English  
St.Teresa's College (Autonomous)  
Ernakulam

Ernakulam

February 2022

# **RICHARD WRIGHT'S BLACK BOY: A PORTRAIT OF RACISM AND SUBJUGATION**

*By*

SOJA M.S

B.A English Language And Literature

St. Teresa's College (Autonomous), Ernakulam

Register No: AB19ENG041

2019-2022

Supervising Teacher: Ms.Lakshmipriya P Santhosh

The Study titled 'Richard Wright's Blackboy: A Portrait of Racism and Subjugation' examines about the ideology underlying racism can manifest in many aspects of social life. Such aspects are described in this section, although the list is not exhaustive. It portrays the theme of Racism and it's effects, then mainly the problems faced in white neighborhood. Narrator is very much fallen to the deep sensual observation of the character who has been deliberately focusing in the racial identity. He looks for unity and human connection, but is often frustrated in search. He finds these circumstances generally unjust and fights attempts to quell his intellectual curiosity and potential as he dreams of moving north and becoming a writer.

## **ACKNOWLEDGEMENT**

I take this opportunity to thank God Almighty for showering his abundant blessings and grace upon me during the course of my project.

I would like to extend my sincere gratitude towards Dr. Lizzy Mathew, Principal, St Teresa's College (Autonomous) for her support.

I would like to express my sincere gratitude to my supervisor Ms. Lakshmi Priya P Santhosh for the constant support and patience. I would have never completed this project without her help and guidance.

I am greatly thankful for Dr. Latha Nair, Head of Department of English, St Teresa's College (Autonomous) for her constant encouragement and motivation.

I extend my sincere gratitude to all other faculty members of the department for their help and encouragement.

SOJAM.S

## CONTENTS

	Introduction	1
Chapter 1:	Human Rights Racism	3
Chapter 2:	The projection of Racism in Richard Wright's Blackboy	11
	Conclusion	19
	Work cited	22

## Introduction

Racism is the belief that different races possess distinct characteristics, abilities, or qualities, especially so as to distinguish them as inferior or superior to one another. Racism can be most simply understood as someone behaving differently to another person based on the colour of their skin and culture.

In this project, you are going to find out what racism is and where it comes from and what you can do if you see or suffer from racism. Racism includes picking up on people who are from a different country too. It is also felt by lots of different groups.

For example:- Jewish people have been persecuted (This is called anti-Semitism).

With protests and demonstrations taking place around the globe in recent weeks and people talking about removing old statues, it's likely you've heard the word 'racism' a lot recently.

Racism is where someone treats another person differently because their skin colour is not the same as theirs, they speak a different language or have different religious beliefs, for example.

In this guide, you're going to find out what racism is, where it comes from and what you can do if you see or suffer from racism. First, racism is one of many expressions of our evolved capacity to live and work in groups. The very human tendency to identify with an “us” defines the broader “them”.

The histories of slavery and racism in the United States have never been more pertinent. This is also the case for the comparatively understudied history of race as a concept, without which it is impossible to understand how Europeans and their colonial “descendants” in the United States engineered the most complete and enduring dehumanization of a people in history.

The logic behind the history of race initially seems deceptively clear: to justify the forced deportation of 400,000 Black Africans to North America (and another eleven million to other parts of the Americas between 1525 and 1866), Europeans and their American heirs found it necessary to debase and revile their captives. Yet today's racism is more than a malignant byproduct of the 19th-century American plantation system; it also grew out of an elaborate and supposedly “scientific” European conception of the human species that began during the Enlightenment.

In this vein I have enjoyed catching up with some of the most recent research on the evolution of race and racism. Two of the most important reads are an Article On The Roots Of Racism by Elizabeth Culotta and Colleagues On The Neuroscience Of Racism by Jennifer Kubotta.

Of the many topics in the literature and practice of international development, racism seems to be one of the most relevant and least covered. Racism is a powerful, violent and complex system that will not be given justice in this short post, but I hope to advance the conversation because I perceive it to be a fundamental issue in international development. One of my biggest critiques of development is the apparent lack of concern and urgency to question assumptions.

The organization Dismantling Racism Works provides important frameworks and resources to enable effective dialogue on racism. One of their definitions is: Racism = race prejudice + social and institutional power. “Racism involves one group having the power to carry out systematic discrimination through the institutional policies and practices of the society and by shaping the cultural beliefs and values that support those racist policies and practices.

Progressive thinkers, abolitionists and, eventually, formerly enslaved people including the writer Olaudah Equiano began critiquing the roots and effects of racial prejudice as early as the 1770s. And yet, even as scientific research has confirmed just how wrong Enlightenment theories of race were, many of the most rearguard and unscientific European notions regarding race have remained deeply embedded in the American psyche, not to mention in the arsenal of the Alt Right. Some of the major

racist writers like Enid Blyton, Rudyard Kipling, Charles Dickens, Dav pilkey and Richard Wright hasplayed a key role which truly addresses about the racist ideas and it's development.

So I choose Richard wright who really experienced and exposed about the matters deepness.Richard Nathaniel Wright (September 4, 1908 - November 28, 1960) was an American author of novels, short stories, poems, and non-fiction. Much of his literature concerns racial themes, especially related to theplight of African Americans during the late 19th to mid-20th centuries suffering discrimination and violence. Literary critics believe his work helped change race relations in the United States in the mid-20th century.Wright insisted that young communist writers be given the space to cultivate their talents .And also Wright's work expresses a realistic and brutal portrayal of white society's oppression of African Americans.Anger and protest served as a catalyst for literature intended to promote social

change by exposing the injustices of racism, economic exploitation, and imperialism. Through his art, Wright turned the torment of alienation into a voice calling for human solidarity and racial advancement.

# Chapter 1

## Human Rights Racism

Racism and racist appear to be words of recent origin, with no citations currently known that would suggest these words were in use prior to the early 20th century. But the fact that the words are fairly new does not prove that the concept of racism did not exist in the distant past. An emerging body of research indicates that racism is a potential source of acute and chronic stress for many ethnic group members.

People often define racism as disliking or mistreating others on the basis of race. That definition is wrong. Racism is a system of advantage based on race. It is a hierarchy. It is a pandemic. Racism is so deeply embedded within U.S. minds and U.S. society that it is virtually impossible to escape.”

Research also shows that children are more attuned to faces of the racial majority group. That is, Black children are better at recognizing White faces than White children are at recognizing Black faces. This disparity can have tragic real-world consequences. In a criminal lineup, for instance, not being able to recognize Black faces, paired with biased preferences and beliefs, increase the odds that an innocent Black suspect will be misidentified as the perpetrator of a crime.

International human rights law seeks to eliminate racial discrimination in the world through treaties that bind and norms that transform. Yet law's impact on eradicating racism has not matched its intent. Racism, in all of its forms, remains a massive cause of discrimination, indignity, and lack of equality for millions of people in the world today. This Article investigates why. Applying a critical race theory analysis of the legal history and doctrinal development

of race and racism in international law, Professor Spain Bradley identifies law's historical preference for framing legal protections around the concept of racial discrimination. She further exposes that international law has neither explicitly defined nor prohibited racism. In response, Professor Spain Bradley advances a long-overdue claim: racism should be affirmatively and explicitly recognized as a human rights violation under international law. She argues that addressing racism in the world today requires understanding how human rights are violated by racial ideologies in addition to discriminatory acts. Insights from neuroscience about racial bias deepen these understandings. By naming "human rights racism" as the central challenge, this Article calls upon the international community to affirmatively recognize racism's extensive harm and to take more seriously its eradication.

Richard Wright is one of the most prominent African-American writers who deals with the

crucial threads like racism and subjugation of the black people. The Black Boy deserves to be called as the masterpiece for highlighting the pathetic portrait of the life of black people in America. Richard Wright rightly observes and expresses the minute happenings took place with him and the Black Boy is nothing but the portrait of all such happenings which appropriately underline the discrimination, racism, harassment and subjugation of the black people in the American society. The Black Boy is a memoir which deals with Richard's journey of life from his childhood and presents his typical perception of looking towards the white as well as black people or African American people in the society of America. The present paper is an attempt to explore this journey of Richard and analysing his experiences as an African American being in the lights of discrimination, racism, suppression, oppression, harassment, marginalization and typical approach of the white people or Euro-Americans of looking towards the black or African American people.

So that Richard Wright's *Blackboy* is also a connection of human rights racism. We can understand the threats they have dealt with. The theme also contains other factors too, but also we can point out the problems that differentiate humans. Racism itself is just a fact in it, there many other things that includes based on violence, subjugation etc.

One of the great successes of the past century has been the popularity of the idea of universal human rights. But the full meaning of human rights is often not fully understood, namely equal human rights: not just rights for me and people 'like me', but for each and every one of us, whether you are like me or not. The real message of the Universal Declaration of Human Rights is that everyone is entitled to the same human rights and to equal human dignity. No one can be denied their human rights because

they are different from others, whether by sex, race or ethnicity, work or descent, caste, culture, religion, skin colour or other grounds. The struggle to ensure equality of treatment for everyone is thus at the centre of all efforts to promote the universal protection of human rights.

Richard whose life is lived in grinding poverty and amid terrible racial suppression and violence. They allow Richard to develop as an individual, and provide windows onto different parts of the world, places Richard can only dream of visiting one day, after he has left the South.

Characters respond to this overwhelming racism in different ways. Some black families, like Granny's, find solace in religion, but Wright does not have any "feeling for God," and rejects the stern discipline

(symbolized by the “switch” used for beatings) that some black families impose on themselves. Some black workers, like Shorty in Memphis, act as “clowns” for white men to gain favors and make extra money, but Wright is unwilling to act submissively for white men's benefit, and he knows that Shorty will never save enough money to leave. Wright's only solace, and eventually his salvation, comes in the form of books. He begins a serious effort in self-education in Memphis, and reads enough to gain some knowledge of the world beyond the American South. This reading does not always help Wright overcome racial violence, but it gives him the confidence to try his luck in the broader world. Richard sees that some people have lighter skin, and other people darker skin, but he only understands what these distinctions mean, culturally and politically, after observing the bigotry of whites and the fear with which black families live. *Black Boy* shows in brutal detail the consequences of Southern racism.

Richard Wright's brilliance in writing about such literary facts are mostly get into new construction of thoughts; The goal was to estrange all blacks from the entire arenas of life not merely political arena.

Finally, it is Richard's education in literature that allowshim to write *Black Boy* itself. The memoir describes Richard's young life and Southern upbringing asforming his intellectual path, and included in the book are the events, and texts, that have shaped him. By the memoir's end, Richard is writing his own material—he is contributing to the “conversation” between books thathas captured his attention since childhood.

## Chapter 2

### The Projection of Racism in Richard Wright's Blackboy

The study lays an emphasis of an issue of violence which occurs abundantly in blackboy. Violence is the prominent theme in the novel. The story in the novel blackboy mainly recollects the life of wright and more importantly the interactions with the whites, and his neighbours and his own family who were scared and very pious during Jim Crow era. The life that Richard wright faced in his childhood and early adulthood was hurtful. But due to his endless efforts and intelligence and desire to fight against injustice, he was able to achieve considerable success as a black despite living under the Jim crow laws.

From the beginning of his life he had some quest for knowledge. At every situation he was curious to know the reason behind every fact, as he grows up he faces many problems ,poverty , hunger for food , as well as for knowledge , family segregation. He observes that some

people are different from each other .That is colour different and some are lighter in colour and some are dark and then light Colour people are superior than black and they are given more respect than darks. Another thing he observes that for black people there is separate place to live in the society, There is not enough even white people children studies in their own schools and darks are not allowed. The novel blackboy reflects the racism and oppression of the black people at every place. Even women were not spared from this suffering. The novel is about the struggle of an individual to survive in the atmosphere of oppressors and cope up with the environment . Richard was able to break away the racial system of the society he discovers his potential, self worth and as a writer.

By the end of the memoir, Wright sees race as a problem to be addressed, if not entirely overcome, by the conscience of every individual , attempting to make sense of the society in which they live.

As it is true that the condition of African American people was very pathetic in. This paper tries to explore racism and oppression of black American people , who suffered discrimination and brutality by the whites. They were oppressed by the whites, being an African American Richard himself faced many problems while growing up. This paper reflects the condition of African American through the novels Blackboy.

Black Boy shows in brutal detail the consequences of Southern racism. It also demonstrates that racial distinctions are not “inherent” or “biological,” It is obvious that violence was deeply rooted in American history. Although, Americans were excited about the establishing of the American Colonies, and then about the United States, since the country promoted the search for liberty including political and economic freedom, and freedom of religion. However, from the beginning, the American society was simultaneously founded on atrocious forms, of oppression and injustice that implied the complete rejection of freedom for slaves. And this

consequence can be regarded as a fundamental paradox of American history. In order to be more specific concerning the issue of violence in America, it is necessary to explore the living conditions of the black community during the implementation and reinforcing of those laws in Southern America, more specifically in Mississippi where Richard Wright's *Black Boy* was set.

The Constitution of America plainly states that "All men and women are created equal", however, black people were constantly oppressed during the beginning of twentieth century in the South according to the Jim Crow Laws (Zinn 688). The whites had implemented plenty of strategies to restrain blacks from receiving citizenship in America. Furthermore, they created murdering groups for eliminating the African-Americans. The blacks were considered as inferior and were dealt with in a degrading manner.

Richard was dwelling in a typical phase of his life and he was unable to receive the proper support and

encouragement from his mother. He has developed his own mind set about the Euro Americans but fails in its execution. Once, when he was of the age of ten only, he has been hit by the white guys while quarreling for the issue of roundhouse. Richard scares to show it to his mother but medical assistance was needed so he approaches to his mother and tells her. In his words, “She rushed me to a doctor who stitched my scalp; but when she took me home she beat me, telling me that I must never fight white boys again, that I might be killed by them...”

Richard’s innocent but analytical mind raises many questions about racism in the book. Many times he has been convinced by his mother but in certain situations his mother also becomes helpless to answer his questions. Richard’s mind was unable to accept this discrimination and harassment of the African Americans and he was powerless to deal with this issue of racism. This stream of racism or discrimination or subjugation of the African Americans has been flowing since long and Richard childhood itself becomes the witness of it. It is already said

that the childhood's marks rarely remove and it happens with Richard. Once Richard finds the black prisoners who were being observed by the white wardens and he asks his mother why is it so? Why all black people are looking like prisoners and white people are observing them. The dialogues between them show the psychology of Richard and his curiosity to know the deeds of the white people. He astonishes to see the large population of African American who has been suffering and facing many difficulties and there were away from all the infrastructural access of the society. They were human beings for their own sake and they have been treated in a marginalized way by the so called core element of the society, Euro Americans. The way he expresses his feelings about the children he came across is quite touching.

Richard highlights one more crucial instance of the racism and how African American people were being tortured psychologically and economically.

Eventually he came to the point at fighting and this is the result of the domination and

exploitation of the African Americans. Although he is a child but his psychology was developing with a great tempo and unconsciously his mother also realises the same but she was totally helpless and trying to satisfy him with some kinds of answers which may create positive and practical approach in the psyche of Richard. Moreover, the story in the novel *Black Boy* mainly recollects the life of Wright and more importantly his interactions with the whites, his neighbors and his own family who were scared and very pious during Jim Crow Era. There had also been several incidents where the black boy reacted to abuse by similarly calling their abusers names not fit for a civilized society. Such names were very hurtful and included the likes of "white bastard", "sons of bitches", and so on. This means that, when one reads this part of the book for the first time, one must ponder how Wright's mother could be so strict, inflexible and uncaring. It seemed that she was not troubled for her son.

To conclude, it has been widely acknowledged that the racism, oppression, suppression, subjugation, marginalization, etc. are quite well acquainted threads in literature as far as the Blacks or African Americans are concerned. Although American Constitution rightly states that all men and women are created equal, the black or African American people have been treated badly and oppressed in the couple of decades of the twentieth century in the South as far as Richard's portrait depicted in his *Black Boy* is concerned. Richard appropriately states in his book that such ideology and antisocial thinking can be changed only with the help of reading and writing. It is the only medium through which the mass can be aware of the equality, prosperity and beauty of life.

## Conclusion

I think it is unsurprising that we often talk about centring local interests while rarely even mention racism. While truly addressing racism, I think we can do a better job of making racism one of the foundational lenses through which projects must be designed and assessed. I'm actually honoured to share this extremely important and relevant topic before you. Development in itself thrives from both the production and consumption of racist ideas that is described. It is frustrating that so many facets of development are intertwined with these ideas that it requires effort to call out and fight against the racist rhetoric we can encounter often. I agree that being equipped with the tools and knowledge to oppose racism is very important and necessary. This is something that I wish was taught, acknowledged and raised front and center in more classrooms and discussions. Those who work in

development in any capacity should work hard to call out racism when they see it.

So far from chapter one we have gone through the simplest form of human rights racism that is affected harshly and we analysed by comparing the work of Richard Wright's *Black Boy*. Racism as a problem among individuals is a familiar topic in literature. *Black Boy*, however, explores racism not only as an odious belief held by odious people but also as an insidious problem knit into the very fabric of society as a whole.

And in Chapter 2 it is finalized about the Projection of racism and its controversies. For a black character to be acceptable within a white framework, ordinarily the only framework available to him, he must to a certain extent be brutalized. To become respectable within that framework is to become subhuman, and if one is judged to be subhuman, it follows that his capabilities and aspirations will be treated as subhuman as well.

Wright never retreated from this theoretical position, but in his own works, he rarely measured up to its requirements. So that from this theoretical framework, this sensation furthers exists to know more and standing exactly as a great topic among others. Eventually we are able to receive a consistent formal knowledge about racism and subjugation, not only through the theories and the affects but through a prominent writer Richard Wright who specifically puts a key to clarify more about this project.

## Works Cited

Bowser, Benjamin P. (2017). "Racism: Origin and Theory". *Journal of Black Studies*. 48 (6): 572–590. doi:10.1177/0021934717702135. S2CID 151511329

"Definition of racism" Merriam-Webster. Retrieved January 3, 2018.

Helms, Janet (1993). *Black and White Racial Identity: Theory, Research, and Practice*. New York: Praeger. ISBN 978-0275946128.

Isaac, Benjamin H. (2004). *The Invention of Racism in Classical Antiquity*. Princeton NJ: Princeton University Press.

Joyce, Joyce Ann. "Wright, Richard (1908–1960)." *African American Writers*, edited by Valerie Smith, 2nd ed., vol. 2, Charles Scribner's Sons, 2001, pp. 875-894. Literature Resource Center, [http://link.galegroup.com/apps/doc/CX1387200066/LitRC?u=was\\_h43584&sid=LitRC&xid=8e2fb644](http://link.galegroup.com/apps/doc/CX1387200066/LitRC?u=was_h43584&sid=LitRC&xid=8e2fb644). Accessed 3 Apr. 2019.

Race, history and culture – Ethics – March 1996 Extract of two articles by Claude Lévi-Strauss.

Racisms: An Introduction Sage. Archived from the original on 2019-04-01. Retrieved 2017-06-21.

"Racism, n." OED Online. September 2013. Oxford University Press. Accessed December 3, 2013.

Richards, G. (1997). *Race, Racism, and Psychology: Towards a Reflexive History*. New York: Routledge.

Richard T. (ed.). *Encyclopedia of Race, Ethnicity, and Society*. SAGE. pp. 1113–1115. ISBN 978-1-4129-2694-2.

"Richard Wright" Mississippi Encyclopedia. Retrieved October 29, 2019.

"The Rise and Fall of Jim Crow . Jim Crow Stories . Segregation in the U. S. Government".

Wellman, David T. (1993). *Portraits of White Racism*. New York:Cambridge University Press.  
p. x.

Wilson, Charles, E., *Race and Racism in Literature*, GreenwoodPress, London, 2005.print.

Wright, Richard. *Black boy: (American hunger): a record of childhood and youth*, New York:  
Perennial Classics, 1998.

Wright, Richard, 1908-1960 (1998). *Black boy : (American hunger) : a record of childhood  
and youth (1st Perennial Classicsed.)*. New York. ISBN 978-0060929787. OCLC 39339337.

Wright, Richard (1966). *Black Boy*. New York: Harper and RowPublishers. ISBN 0-06-  
083056-5.

Wright, Richard (1993). *Black Boy*. New York: Harper Collins.pp. 455–459. ISBN 0-06-  
081250-8.

Wright, Richard, *Black Boy*, Vintage Publication, London, 2000.Print.







Project Report

On

THE EFFECT OF COVID-19 ON ONLINE  
FOOD DELIVERY SERVICES.

*Submitted*

*in partial fulfilment of the requirements for the degree of*

BACHELOR OF SCIENCE

*in*

MATHEMATICS

*by*

ANEETA SUNISH

(Register No. AB19AMAT004)

*Under the Supervision of*

NEENU SUSAN PAUL



DEPARTMENT OF MATHEMATICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

Project Report

On

**THE EFFECT OF COVID-19 ON ONLINE  
FOOD DELIVERY SERVICES.**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

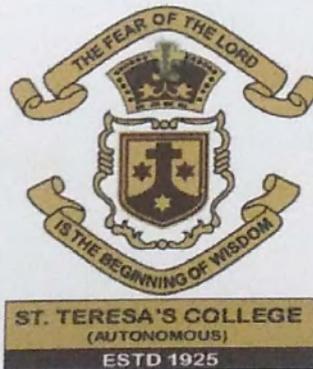
*by*

**ANEETA SUNISH**

(Register No. AB19AMAT004)

*Under the Supervision of*

**NEENU SUSAN PAUL**



**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

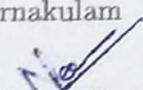


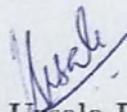
CERTIFICATE

This is to certify that the dissertation entitled, **THE EFFECT OF COVID-19 ON ONLINE FOOD DELIVERY SERVICES**. is a bonafide record of the work done by Ms. **ANEETA SUNISH** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

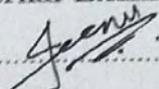
Date:

Place: Ernakulam

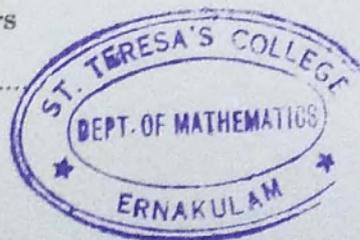
  
Neenu Susan Paul  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

  
Dr. Ursala Paul  
Assistant Professor and Head ,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

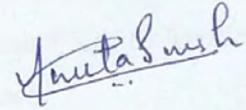
1: 

2: .....



## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Neenu Susan Paul, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.



Ernakulam.

**ANEETA SUNISH**

Date: 04 - 03 - 2022

**AB19AMAT004**

## ACKNOWLEDGEMENT

When we set goals for ourselves, there are always obstacles in the way that may deter us from accomplishing the goals. There are also people in our lives that are aware of those goals, and encourage us and also support us to continue regardless of the obstacles, it is now that I can formally thank those people for doing just that for me, before thanking anyone on this earth, I must first thank God for being at my side during this challenging time of my life. The spiritual support has helped to keep me focussed. Thanks to my Guide Neenu Susan Paul, St. Teresa's College for the support and encouragement through this process.

Ernakulam.

ANEETA SUNISH

Date:

AB19AMAT004

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENT</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Background of the Study . . . . .	2
1.2 Literature Review. . . . .	3
1.3 E-Commerce . . . . .	4
1.4 Objective . . . . .	5
<b>2 METHODOLOGY</b>	<b>6</b>
2.1 Data Collection . . . . .	6
2.2 Questionnaire . . . . .	7
2.3 Chi-Square Test . . . . .	7
<b>3 DATA ANALYSIS</b>	<b>9</b>
3.1 Survey Analysis . . . . .	9
3.2 Data Analysis . . . . .	20
<b>4 RESULT AND CONCLUSION</b>	<b>27</b>
4.1 Conclusion . . . . .	27
4.2 Result . . . . .	30
<i>REFERENCES</i> . . . . .	32
<i>Annezure</i> . . . . .	33

# Chapter 1

## Introduction

---

The COVID-19 pandemic has disrupted nearly every area of people's lives, including their capacity to buy products. Consumers have been confined at home as a result of government-imposed lockdowns, inhibiting regular purchasing habits, and many brick-and-mortar companies have closed. Pharmacies and supermarkets, for example, have stayed open albeit with shortened hours. Many eateries have shuttered or resorted to takeout in order to stay afloat. For many clients, home delivery has given a solution to some of COVID-19's issues. Employees that are required to work remotely, as well as a range of other groups, such as parents who must combine work and parental responsibilities, or individuals who are at risk of serious COVID-19 health concerns, may find e-commerce and home delivery to be a useful option. Despite the difficulties people have had in ordering food online as a result of the pandemic, such as a lack of food inventories, a lack of public transportation, and fewer hours to work at convenience stores and supermarkets, a large percentage of people have been able to obtain sufficient amounts of food stocks. All of this can be attributed to the rapid increase in the number of online food apps, which encourage people to order their meals online and have them delivered to their homes or picked up by themselves without having to enter the restaurant. Programs that encourage grocery shopping and food ordering online can also help to reduce the spread of the Covid-19 virus by reducing interaction between customers and sellers, thereby creating a safe environment in and

of itself. Considering all of the facts, the fact that customers are shifting to ordering food and shopping for groceries online as a result of the pandemic has implications for the retail market as well.

### 1.1 Background of the study

The early stages of the covid pandemic, as well as the strict lockdown, harmed India's online food delivery companies, which were dominated by Zomato and Swiggy. Consumers had resisted online food ordering due to safety concerns and a general preference for home cooked food. However, things improved later on, and recovery was quicker. According to analysts, the pandemic has assisted companies in improving unit economics and adding more customers, as the frequency of dining out has decreased.

As we stay at home to mitigate the impact of virus many of us have turned to delivery services for meals and groceries for the first time. Before the advent of the Covid nineteen, young people were increasingly buying food online. But now people of all ages are buying food online equally. In this way the spread of the covid is greatly reduced. In the past, large restaurants were mostly using online food delivery services but now it has changed and even smaller hotels started to use online food delivery platforms. The hospitality industry has adapted to make it easier for everyone to support a favourite local restaurant, avoid going out, and simply find a relaxing respite during these stressful times. Because of the nature of how Covid-19 spreads, close contact with others may pose the greatest risk of infection. Many restaurants now allow customers to pay ahead of time, either by phone or online, eliminating the need for physical cash transfers or credit card handling, both of which can potentially harbor the virus.

During the pandemic, diners became accustomed to ordering delivery, and the habit may persist long after dining rooms reopen. However, restaurants and delivery companies continue to be uneasy part-

ners, haggling over fees and struggling to make the service profitable for both parties. Companies such as Zomato and Uber Eats assisted many restaurants in remaining open during lockdowns by allowing diners to stay in and still order out. However, the convenience came at a cost delivery companies can charge commission fees of 30% or more per order, reducing restaurants already meager profits. Delivery was already increasing prior to the pandemic, but it sky rocketed during lockdowns around the world.

Prior to the outbreak of Covid-19, many young people preferred online food delivery. People belonging to other age categories mostly prefer home-cooked meals because they are safe and healthy. During the Covid-19 scenario, everyone was confined to their homes. There were no stores or hotels open, and all local businesses began delivering products through online platforms. And because it was convenient and time-saving, everyone began ordering food online. Online delivery was also safe and reduced the spread of the virus.

## 1.2 Literature Review

The impact of COVID-19 on restaurant meal ordering via apps 2020: This article was published on April 22, 2021 by Statista's Research Department. According to a Local Circles poll conducted in May 2020, approximately 65 percent of respondents stated that they would not order restaurant meals for delivery within 30 days of the corona virus lockdown being lifted. Approximately 3% of those polled said they would order more than four times during this time period.

According to Priyadharshini (2017), India has more people between the ages of 10 and 24, making it the world's largest young population. With more young people joining the workforce every day, economic growth, increased female labour power, and increased consumer mobility, the traditionally difficult Indian market has evolved and is in need of a more diverse menu.

Samsudin et al (2011) points out that alongside client feedback for an eatery, a plan and execution of wireless food ordering framework was completed. It empowers cafes proprietors to setup the framework in wireless environment and update menu presentations effectively. Advanced mobile has been coordinated within the adaptable wireless food ordering system requesting framework with continuous client criticism execution to encourage ongoing correspondence between eatery proprietors and clients.

Rathore et al (2018) states that 50.8% of consumers use a food delivery service because they don't want to cook because it allows them to have food delivered to their home or office in under 60 minutes.

According to Pathan et al (2017), an online food ordering system can be used to build up a restaurant and mess menu, and customers can quickly place orders. Also, with an online food menu, orders can be readily traced, the client database can be maintained, and the meal delivery business may grow. Restaurants and mess may quickly update their online restaurant menus and upload photographs. Potential consumers can quickly examine a restaurant menu on the internet and place orders at their leisure. As a result, an automated food ordering system with feedback and wireless communication is shown.

### 1.3 E-Commerce

According to Garret (1996), electronic commerce or e-commerce is the exchange of goods and services via the internet or other computer networks. Buyers and sellers conduct business via networked computers in e-commerce. Electronic commerce also includes the exchange of business information, the maintenance of business relationships, and the conduct of business transactions via communication networks. It contains the relationship between companies (business-to-business), customers (customer-to-customer), and companies and cus-

tomers (business- to-customer). Currently, the business to business sector dominates e-commerce, while the customer-oriented segment lags far behind, accounting for fewer than 10 even though they are all experiencing an exponential growth (Vladimir, 1998).

Buyers like the convenience that e-commerce provides. They can compare prices and make purchases without leaving the house by visiting the World Wide Web (www) sites of numerous suppliers 24 hours a day, seven days a week. For sellers, e-commerce offers a way to cut costs and expand their markets. They do not need to hire staff or maintain a store or distribute mail order catalogs. For retailers, e-commerce allows them to save money while also expanding their customer base. They don't have to hire people, keep a store running, or print and distribute catalogues. Sellers have the ability to market their products or services globally because they sell over the global internet and are not limited by the physical location of a store.

There are several drawbacks to e-commerce. Some customers are hesitant to make online purchases. Customers want to test the comfort of an expensive item, thus online furniture enterprises, for example, have failed for the most part. Many people consider shopping to be a social experience; for example, they may enjoy going to a shops or a shopping mall with family and friends, which they cannot get online. Customers Furthermore, must be guaranteed that credit card transactions are safe and secure. Their personal information is kept confidential. E-commerce not only expands the range of products available to customers, but it also makes it easier for them to find what they want, not only in terms of products and services, but also in terms of attracting new customers and retaining existing ones.

#### 1.4 Objective

- To interpret and to find if there is any relation between place of residence and the mode of preference of food delivery services.

- To interpret and to find if there is any relation between age group and variation in the usage of online food delivery services.
- To interpret and to find if there is any relation between age group and amount of money spend on online food delivery services.
- To interpret and to find if there is any relation between age group and reasons for choosing on online food delivery services.

## Chapter 2

# METHODOLOGY

---

### 2.1 Data Collection

In order to meet the research objectives, it is critical that the data collected is accurate. All data sources can be divided into two categories: Primary data is gained by direct observation or data collected by the researcher. It refers to data that has been acquired for a specific purpose from a field of inquiry and is of a unique type. Primary data for the project was acquired primarily using the survey approach, utilising the tool questionnaire. Secondary data are ones that have already been acquired by others for a specific reason and are then used in a variety of situations. It is second hand information on an incident that the researchers have not personally witnessed.

Customers who order food online are considered the study's target group. The data was taken between November 23rd and January 1st, 2022. India's statewide lockdown began on March 25, 2020, in order to restrict population movement. The government, on the other hand, allowed e-commerce businesses to continue operating during this time. A well-structured online questionnaire was created using the Google forms and distributed to the responders. During the lockdown, an online-based survey is a viable option for data collection to safeguard the safety of respondents and researchers. Students, employed, and unemployed citizens were among those who responded. We distribute the survey via WhatsApp and social media sites. This study was carried out

with the permission of all participants, and no personal information was gathered. The samples were collected from Ernakulam district, Kerala. Numerous Indian state governments did not allow the operation of on-line food delivery during the statewide lockdown, although many well-established online food delivery businesses like Zomato and Swiggy were fully active in Ernakulam during the nationwide shutdown.

## 2.2 Questionnaire

The Questionnaire is used to collect data from responders, and it is created using Google Forms. It consists of a sequence of questions that the investigators are expected to ask and the respondents are supposed to choose an alternative for each individual enquiry. Questions are in the form of multiple choice questions. There is no personal information collected. Customers that order food online and live in the Ernakulam city area were chosen at random for primary data. Data was collected using a standardised questionnaire. The sample size was determined by taking 500 respondents from the total population of respondents in Ernakulam.

## 2.3 Chi-Square Test

A chi-square statistic is a measure of the difference between the observed and expected frequencies of the outcomes of a set of events or variables. Chi-Square depends on the size of the difference between actual and observed values, the degrees of freedom, and the sample size. Chi-Square can be used to test whether two variables are related or independent from one another. It can also be used to test the goodness-of-fit between an observed distribution and a theoretical distribution of frequencies.

When the chi-square test is used as a test of independence, it allows the researcher to test whether the two attributes being tested are associated or not. For this test, a null and alternative hypothesis is for-

mulated where the *null hypothesis* that the two attributes are not associated, and the *alternative hypothesis* is that the attributes are associated.

From the given data, the expected frequencies are then calculated i.e.,

$$\text{ExpectedFrequency} = \text{RowTotal} \times \text{ColoumnTotal} \div \text{GrandTotal} \quad (2.1)$$

followed by the calculation of chi-square value. The null or alternative hypothesis is accepted based on the calculated chi-square value. If the calculated chi-square value is less than the value in the table at the given level of significance, the null hypothesis is accepted, indicating that no relationship exists between the two attributes. If the calculated chi-square value is greater than the value in the table, the alternative hypothesis is accepted, indicating that there is a relationship between the two attributes.

The *p-value* or the calculated probability is the best probability to provide the smallest level of significance at which the null hypothesis is not true. If the *p-value is small (less than 0.05)*, it indicates a piece of strong evidence against the null hypothesis. As a result, the null hypothesis is rejected and the alternative hypothesis is accepted. This means that the results of the research study are statistically significant. If the *p-value is large (greater than 0.05)*, it indicates weak evidence against the null hypothesis. As a result, the null hypothesis is not rejected and the alternative hypothesis is not accepted. This means that the results of the research study are not statistically significant.

In order to find the p-value from the chi-square test, at first, the chi-square test is to be performed to obtain the chi-square value. While performing the test, the *degree of freedom* is also calculated by the formula,  $d.f = (c-1)(r-1)$  where *c* is the number of columns and *r* is the number of rows. Now the chi-square distribution table is entered, with the obtained degree of freedom, and the value of the chi-square is found in the table.

## Chapter 3

# DATA ANALYSIS

---

### 3.1 Survey Analysis

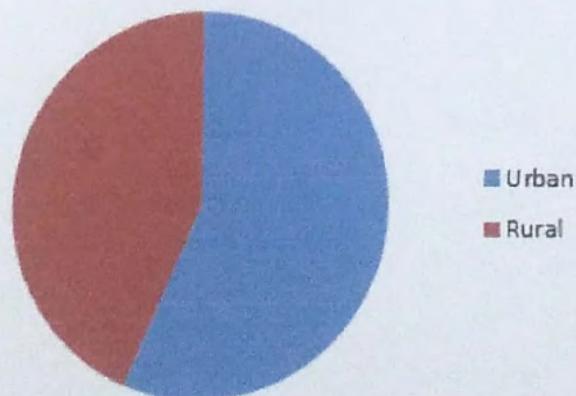
#### 1.PLACE OF RESIDENCE

About 55.2% of responses were from urban area while 44.8% were from rural area.

Area	No Of Responses
Urban	276
Rural	224

Table 3.1: Frequency

#### Place of residence.

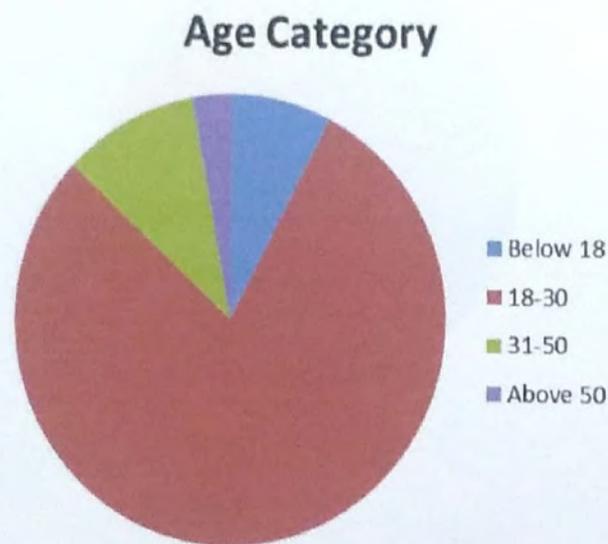


## 2) AGE CATEGORY

About 8.8% of the respondents belong to below 18 category, 74.2% belongs to 18- 30 category, 12.6% belongs to 30-50 category and 4.4% belongs to Above 50 category.

Area Group	No Of Responses
Below18	44
18-30	371
30-50	63
Above50	22

Table 3.2: Frequency

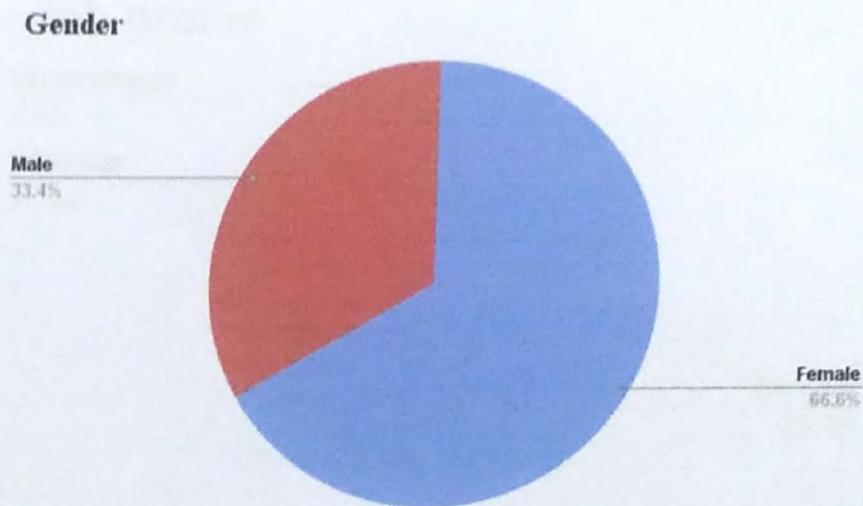


### 3) GENDER

About 66.6% responses were received from female category, about 33.4% responses were received from male category.

Gender	No Of Responses
Male	167
Female	333

Table 3.3: Frequency

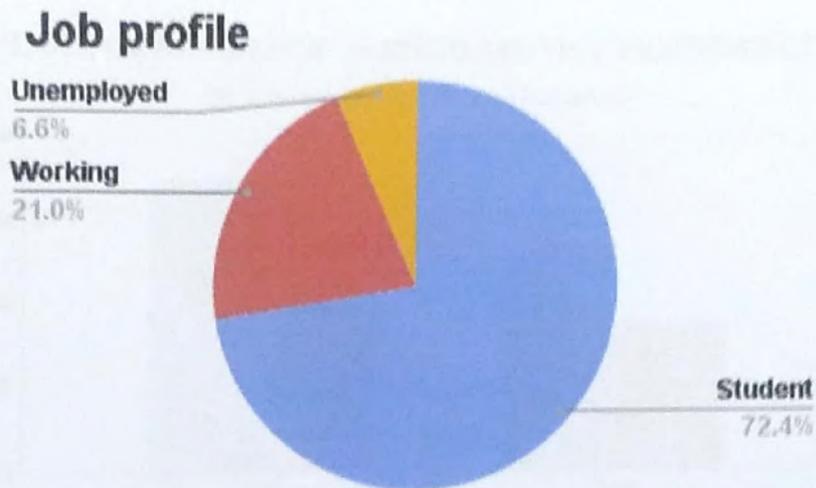


#### 4) JOB PROFILE

About 72.4 % of respondents are students, 21 % of respondents are working and about 6.6 % of respondents are unemployed.

Job Profile	No Of Respondents
Students	362
Working	105
Unemployed	33

Table 3.4: Frequency



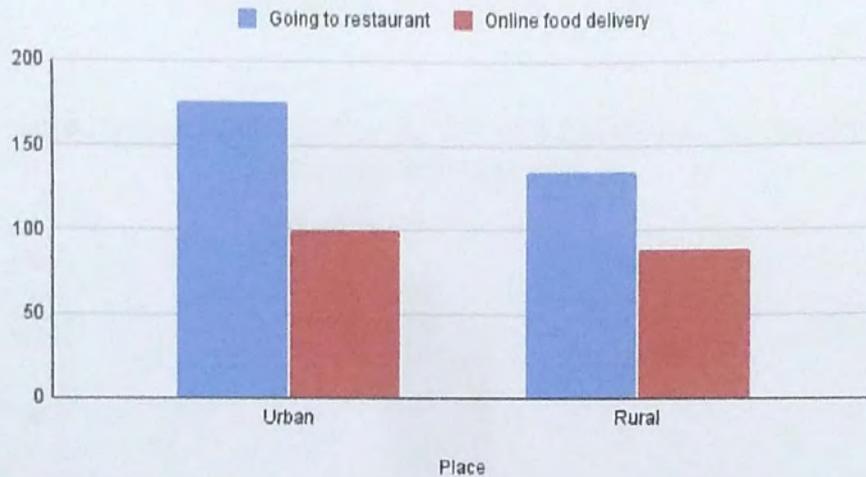
## 5) PLACE OF RESIDENCE v/s FOOD SERVICE PREFERENCES.

Place	Going To Restaurants	Online Food Delivery
Urban	176	100
Rural	135	89

Table 3.5: Frequency

Place	Going to restaurant	Online food delivery
Urban	56.6%	52.91%
Rural	43.4%	47.09%

## PLACE OF RESIDENCE v/s FOOD SERVICE PREFERENCES.



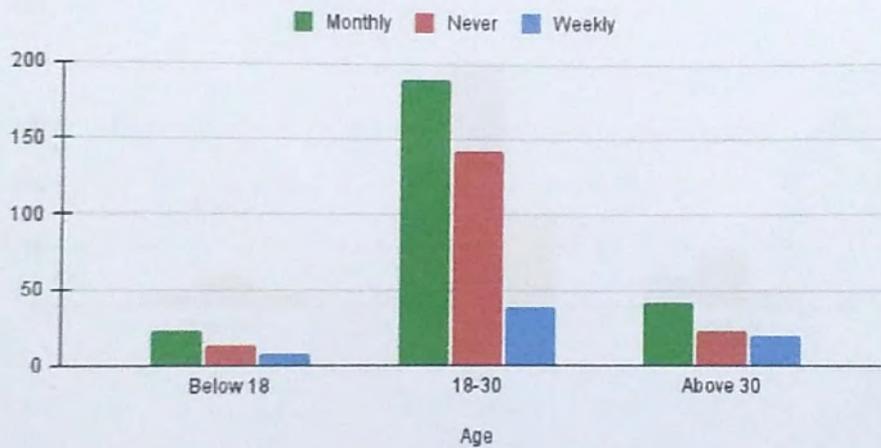
### 6. AGE CATEGORY V/S USAGE OF ONLINE FOOD DELIVERY SERVICES BEFORE COVID

Age	Weekly	Monthly	Never
Below 18	8	23	13
18-30	39	190	142
Above 30	20	42	23

Table 3.6: Frequency

Age	Weekly	Monthly	Never
Below 18	11.94%	9.01%	7.3%
18-30	58.21%	74.5%	79.78%
Above 30	29.85%	16.47%	12.92%

### AGE CATEGORY V/S USAGE OF ONLINE FOOD DELIVERY SERVICES BEFORE COVID.



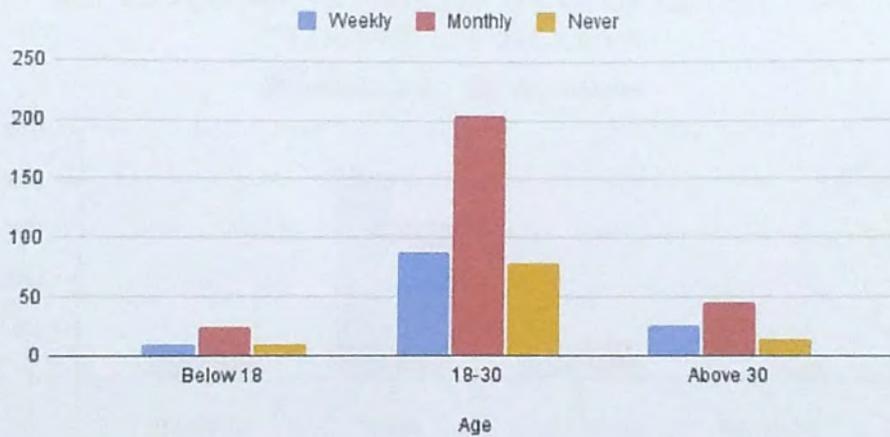
### 7. AGE CATEGORY V/S USAGE OF FOOD DELIVERY SERVICES DURING COVID

Age	Weekly	Monthly	Never
Below 18	10	24	10
18-30	88	204	79
Above 30	26	45	14

Table 3.7: Frequency

Age	Weekly	Monthly	Never
Below 18	8%	8.8%	9.7%
18-30	71%	75.72%	76.7%
Above 30	21%	16.48%	13.6%

### AGE CATEGORY V/S USAGE OF ONLINE FOOD DELIVERY SERVICES DURING COVID.



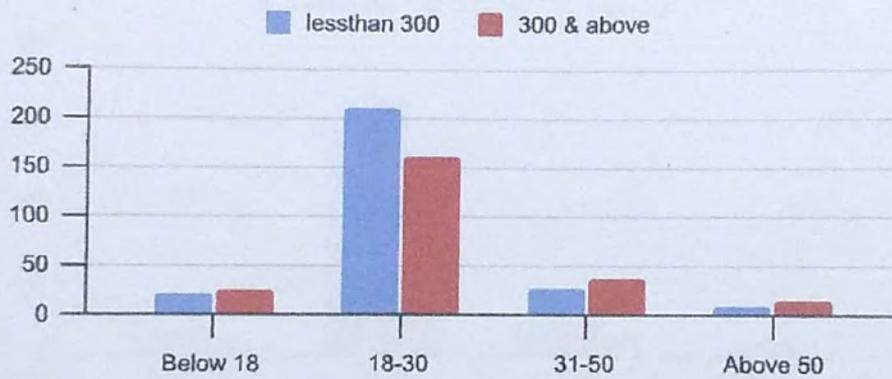
### 8. AGE CATEGORY V/S AMOUNT SPEND ON ONLINE FOOD DELIVERY BEFORE COVID

Age	Less than 300	300 and above
Below 18	20	24
18-30	211	160
31-50	27	35
Above 50	8	14

Table 3.8: Frequency

Age	Less than 300	300 and above
Below 18	7.51%	10.25%
18-30	79.32%	68.37%
31-50	10.15%	15.38%
Above 50	3%	6%

### AGE CATEGORY VS AMOUNT SPEND ON ONLINE FOOD DELIEVRY BEFORE COVID



### 9. AGE CATEGORY V/S AMOUNT SPEND ON ONLINE FOOD DELIVERY DURING COVID

Age	Less than 300	300-700	Above 700
Below 18	19	18	7
18-30	139	169	63
31-50	23	26	14
Above 50	5	11	6

Table 3.9: Frequency

Age	Less than 300	300-700	Above 700
Below 18	10.21%	8.03%	7.78%
18-30	74.73%	75.44%	70%
31-50	12.36%	11.6%	15.55%
Above 50	2.68%	4.91%	6.67%

### AGE CATEGORY V/S AMOUNT SPEND ON ONLINE FOOD DELIVERY DURING COVID



### 10. AGE CATEGORY v/s REASONS FOR CHOOSING ONLINE FOOD DELIVERY SERVICES BEFORE COVID.

Age	Convenient	Fast delivery	Time saving	All of the above
Below 18	8	8	6	22
18-30	114	22	55	180
Above 30	25	7	19	34

Table 3.10: Frequency

Age	Convenient	Fast delivery	Time saving	All of the above
Below 18	5.44%	21.62%	7.5%	9.32%
18-30	77.55%	59.45%	68.75%	76.27%
Above 30	17%	18.91%	23.75%	14.4%

#### AGE CATEGORY v/s REASONS OR CHOOSING ONLINE FOOD DELIVERY BEFORE OVID



### 11. AGE CATEGORY v/s REASONS FOR CHOOSING ONLINE FOOD DELIVERY SERVICES AFTER COVID

Age	No restaurants nearby	Safety	To avoid transportation	None of the above
Below 31	30	138	173	74
Above 30	12	28	28	17

Table 3.11: Frequency

Age	No restaurants nearby	Safety	To avoid transportation	None of the above
Below 31	71.42%	83.13%	86.06%	81.31%
Above 30	28.57%	16.86%	13.93%	18.68%

#### AGE CATEGORY VS REASONS FOR CHOOSING ONLINE FOOD DELIVERY DURING COVID



## 3.2 Data Analysis

## 1. MODE OF PREFERENCE V/S PLACE OF RESIDENCE

Place	Going to restaurant	Online food delivery	Total
Urban	176	100	276
Rural	135	89	224
Total	311	189	500

Table 3.12: Observed frequency

$H_0$ : There is no relation between place of residence and mode of preference.

$H_1$ : There is relation between place of residence and mode of preference.

Significance Value: 0.05

Place	Going to restaurant	Online food delivery	Total
Urban	171.672	104.328	276
Rural	139.328	84.672	224
Total	311	189	500

Table 3.13: Expected frequency

$\chi^2$  value :  $P = 0.422 > 0.05$

$H_0$  is accepted

Hence there is no relation.

## 2.USAGE OF FOOD DELIVERY BEFORE COVID V/S AGE CATEGORY

Age	Weekly	Monthly	Never	TOTAL
Below 18	8	23	13	44
18-30	39	190	142	371
Above 30	20	42	23	85
Total	67	255	178	500

Table 3.14: Observed frequency

$H_0$  There is no relation between age category and usage of food delivery services before covid

$H_1$  There is relation between age category and usage of food delivery services before covid

Significance Value: 0.05

Age	Weekly	Monthly	Never	Total
Below 18	5.896	22.44	15.664	44
18-30	49.714	189.21	132.076	371
Above 30	11.39	43.35	30.26	85
Total	67	225	178	500

Table 3.15: Expected frequency

$\chi^2$  value :  $P = 0.00714 < 0.05$

$H_0$  is rejected

Hence there is relation

### 3.USAGE OF FOOD DELIVERY DURING COVID V/S AGE CATEGORY

Age	Weekly	Monthly	Never	Total
Below 18	10	24	10	44
18-30	88	204	79	371
Above 30	26	45	14	85
Total	124	273	103	500

Table 3.16: Observed frequency

$H_0$ : There is no relation between age category and usage of food delivery services after covid

$H_1$ : There is relation between age category and usage of food delivery services after covid

Significance Value: 0.05

Age	Weekly	Monthly	Never	Total
Below 18	10.912	24.024	9.064	44
18-30	92.008	202.566	76.426	371
Above 30	21.08	46.41	17.51	85
Total	124	273	103	500

Table 3.17: Expected frequency

$\chi^2$  value :  $P = 0.24 > 0.05$

$H_0$  is accepted

Hence there is no relation

## 4. AMOUNT SPEND ON FOOD DELIVERY BEFORE COVID V/S AGE CATEGORY

Age	Less than 300	300 and Above	Total
Below 18	20	24	44
18-30	211	160	371
31-50	27	36	63
Above 50	8	14	22
Total	266	234	500

Table 3.18: Observed frequency

$H_0$ : There is no relation between age category and the amount spend on food delivery before covid.

$H_1$ : There is relation between age category and the amount spend on food delivery before covid

Significance Value: 0.05

Age	Less than 300	300 Above	Total
Below 18	23.408	20.592	44
18-30	197.372	173.628	371
31-50	33.516	29.484	63
Above 50	11.704	10.296	22
Total	266	234	500

Table 3.19: Expected frequency

$\chi^2$  value :  $P = 0.0405 < 0.05$

$H_0$  is rejected

Hence there is a relation.

### 5. AMOUNT SPEND ON FOOD DELIVERY DURING COVID V/S AGE CATEGORY

Age	Less than 300	300-700	Above 700	Total
Below 18	19	18	7	44
18-30	139	169	63	371
31-50	23	26	14	63
Above 50	5	11	6	22
Total	186	224	90	500

Table 3.20: Observed frequency

$H_0$ : There is no relation between age category and the amount spend on food delivery during covid

$H_1$ : There is relation between age category and the amount spend on food delivery during covid

Significance Value: 0.05

Age	Less than 300	300-700	Above 700	Total
Below 18	16.368	19.712	7.92	44
18-30	138.012	166.208	66.78	371
31-50	23.436	28.224	11.34	63
Above 50	8.184	9.856	3.96	22
Total	186	224	90	500

Table 3.21: Expected frequency

$\chi^2$  value :  $P = 0.652 > 0.05$

$H_0$  is accepted

Hence there is no relation

### 6. AGE CATEGORY V/S REASONS FOR CHOOSING ONLINE FOOD DELIVERY SERVICES BEFORE COVID

Age	Convenient	Fast delivery	Time saving	All of the above	TOTAL
Below 18	8	8	6	22	44
18-30	114	22	55	180	371
Above 30	25	7	19	34	85
TOTAL	147	37	80	236	500

Table 3.22: observed frequency

$H_0$ : There is no relation between age category and the reasons for choosing online food delivery before covid

$H_1$ : There is relation between age category and the reasons for choosing online food delivery before covid

Significance Value: 0.05

Age	Convenient	Fast delivery	Time saving	All of the above	TOTAL
Below 18	12.936	3.256	7.04	20.768	44
18-30	109.074	27.454	59.36	175.112	371
Above 30	24.99	6.29	13.6	40.12	85
TOTAL	147	37	80	236	500

Table 3.23: Expected frequency

$\chi^2$  value :  $P = 0.0302 < 0.05$

$H_0$  is rejected

Hence there is relation

### 7. AGE CATEGORY V/S REASONS FOR CHOOSING ONLINE FOOD DELIVERY SERVICES DURING COVID.

Age	No restaurants nearby	Safety	To avoid transportation	None of the above	TOTAL
Below 31	30	138	173	74	415
Above 30	12	28	28	17	85
TOTAL	42	166	201	91	500

Table 3.24: observed frequency

$H_0$  : There is no relation between age category and reason for choosing online food delivery during Covid.

$H_1$  : There is a relation between age category and reason for choosing online food delivery during Covid.

Significance level: 0.05

Age	No restaurants nearby	Safety	To avoid transportation	None of the above	TOTAL
Below 30	34.86	137.86	166.83	75.53	415
Above 30	7.14	28.22	34.17	15.47	85
TOTAL	42	166	201	91	500

Table 3.25: Expected Frequency

$\chi^2$  value :  $P = 0.137 > 0.05$

$H_0$  is accepted.

Hence there is no relation.

## Chapter 4

# RESULT AND CONCLUSION

---

### 4.1 Conclusion

#### 1.PLACE OF RESIDENCE V/S FOOD SERVICE PREFERENCES.

In the case of online food delivery, people living in both rural and urban areas have equal accessibility and availability to online food services. Restaurants are growing at a fast rate in both rural and urban areas. People can directly go to the restaurants according to their needs. Hence there is no relation between place of residence and food service preferences.

#### 2.USAGE OF FOOD DELIVERY SERVICES V/S AGE CATEGORY.

There is a relation between age categories and usage of food delivery services before covid. According to our survey, people belonging to the age category between 18-30 are most likely to order food online as most of them are working class and they do not have enough time to cook food. So it would be convenient for them to order food online. These are the ones who use smartphones more frequently and they are more familiar with food ordering applications. As we can see from the above conclusion, people belonging to the category above 50 are less likely to order food online because they are less aware of modern technologies. So they prefer to have healthy homemade foods. Hence here we can

conclude that age affect the usage of food delivery.

There is no relation between age categories and usage of food delivery services during covid. People have been held at home due to the govt imposed lockdowns, preventing regular shopping habits. Many restaurants have closed, thus forcing the consumers to buy food online irrespective of their preferences. Covid-19 hit everyone so badly that the working industry had to change to work from home. The people belonging to this sector didn't have enough time to cook, forcing them to order food online. All categories started showing a slight increase in the usage of online food delivery, because of the less accessibility of transportation services and safety concerns. All age groups faced equal difficulty during Quarantine period. Online food delivery platforms played a vital role in the lives of all people. Thus irrespective to their gender and age all people started using online food delivery services. Hence here we can conclude that age does not affect the usage of online food delivery services.

### 3. AMOUNT SPEND ON ONLINE FOOD DELIVERY SERVICES V/S AGE CATEGORY.

There is a relation between age categories and the amount spent on online food delivery services before covid-19. The people belonging to the age categories between 18-30 are usually the ones that spend more money on ordering food online. They love to live their lives to the fullest and also they love to try more new dishes. Before corona most of them were financially stable and it was affordable for them to order food. And also because they are hesitant to cook and they prefer to order food instead. On the other hand people of other age groups are very much self aware of their income and preferably would like to save money rather than spending more on eating junk food. Hence here we can conclude that age affects the amount spend on food delivery before covid.

There is no relation between age categories and the amount spent on online food delivery services during covid-19. Food is essential to human life since it supplies energy for all activities, growth, and development. People of all ages require food. During Covid-19, more individuals, particularly those in quarantine, had no choice but to purchase food online. People are willing to order food online even if it means spending more money for delivery because of food safety, quick delivery, and the availability of a wide variety of food items. As a result, when the pandemic began, people were not interested in purchasing food through online platforms. Because the population's fear and anxiety levels have increased, the perception of risk can also include the fear of contagion via food, packaging, and contact with the delivery person at the time of delivery. However, people's perspectives have gradually improved. All categories have begun to buy meals through these online platforms because they restrict customer contact with restaurant employees and let customers to enjoy their favourite restaurant food at home during a pandemic like COVID-19. Hence here we can say that age does not affects the amount spend on online food delivery services during covid.

#### 4.REASONS FOR CHOOSING ONLINE FOOD DELIVERY SERVICES V/S AGE CATEGORY.

There is a relation between age categories and the reasons for choosing online food delivery services before covid. Because even before covid 19 pandemic people belonging to different age groups had different opinions. Nowadays people don't have enough time to go out and eat, so they started taking more advantage of the online food delivery platforms. Some might order food online as to reduce transportation, hesitation to cook (especially for those who find it difficult to cook well) or to save time. People are generally more satisfied with online delivery because it is done at their convenience and avoids the stress of waiting, standing in line, takeout, and so on. The waiting time is well spent by

doing something else at home or at work. The people belonging to the age category 18-30 are more likely to order food online and older adults are also making use of this online food delivery platforms. Therefore each of them have their own reasons. Hence here we can conclude that age affects the reasons for choosing online food delivery services.

There is no relation between age categories and the reasons for choosing online food delivery services during covid. Because majority of the respondents across all age groups started ordering food online as they practiced social distancing and safety as to minimise the spread of covid. Home delivery of food items helped people to avoid transportation. Hence, there was no any shortage or unavailability of food items occurred during the pandemic. Because of how covid-19 spreads, the greatest risk of infection could occur through intimate contact with others. Many restaurants now allow customers to pay in advance, either via phone or online, reducing the need for physical cash transfers or credit card processing, both of which might carry the virus. Hence here we can conclude that age doesnot affect the reasons for choosing online food delivery services.

## 4.2 Result

The sudden onset of the Covid-19 crisis has surprised consumers, companies and government agencies. This study focuses on the effect of Covid-19 pandemic on online food delivery services. Based on the data, we analysed online food delivery platform usage, how it has helped during the Covid-19 pandemic and whether current changes could be long lasting.

Overall, Covid-19 has had a net positive impact on frequency and spending on online food delivery. For those not cooking at home, online delivery service from restaurants is one way to limit the number of trips outside the home. From the survey, we can conclude that usage of food delivery services varies with age category before Covid-19. People belonging to age category 18-30 are likely to order food more often. And

also these people spend more money on food delivery. Despite the fact that a large section of the population in Ernakulam uses online food delivery apps, there are still people who do not use food apps owing to health and quality concerns.

In a nutshell, the majority of users are students and working people, indicating that the online food ordering system is gaining popularity among young people. The changing lifestyles of consumers, as well as the growth of online activity in India, have unquestionably changed the trends in online food ordering.

## REFERENCES

---

- [1] Complementary Statistics : Author-Ramakrishna Pillai.
- [2] <https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/chi-square/>
- [3] R., Priyadarshini. (2017). Consumer perception towards MNC fast food outlets in Coimbatore. International Journal of Applied Research,3(3).Retrieved from <http://www.allresearchjournal.com/archives/2017/vol3issue3/PartD/3-3-54-357.pdf>
- [4] Samsudin, N. A., Khalid, S. K., Kohar, M. F., Senin, Z., and Ihkasan, M. N.(2011). A customizable wireless food ordering system with real time customer feedback. 2011 IEEE Symposium on Wireless Technology and Applications (ISWTA). doi:10.1109/iswta.2011.6089405
- [5] Rathore, S. S., and Chaudhary, M. (2018). Consumer's Perception on Online Food Ordering. International Journal of Management and Business Studies,8(4). Retrieved from <http://www.ijmbs.com/Vol8/issue4/2-suryadev-singh-rathore.pdf>
- [6] R., A., Singh, A., Pathan, S., and Kanade, V. (2017). Online Food Ordering System. International Journal of Computer Applications,180(6), 22-24.doi:10.5120/ijca2017916046

## Annexure

1. What age category do you belong to?

- Below 18
- 18-30
- 31-50
- Above 50

2. Gender

- Male
- Female
- Others

3. Place of residence

- Urban
- Rural

4. Job profile

- Working
- Unemployed
- Student

5. What do you prefer?

- Going to restaurants
- Online food delivery

6. How often do you order food online before covid-19?

- Daily
- Weekly
- Monthly

- Never

7. How often do you order food online during covid-19?

- Daily
- Weekly
- Monthly
- Never

8. How much do you spend on online food delivery before covid-19?

- Less than 300
- 300-700
- Above 700

9. How much do you spend on online food delivery during covid-19?

- Less than 300
- 300-700
- Above 700

10. Why do you prefer online food delivery before covid-19?

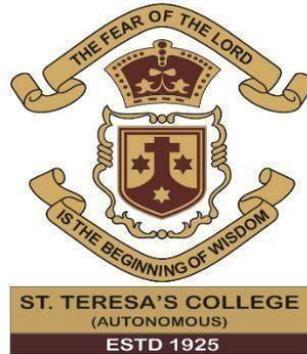
- Fast delivery
- Convenient
- Time saving
- All of the above

11. Why do you prefer online food delivery during covid-19?

- Safety
- To avoid transportation
- No restaurant nearby
- None of the above



**A FEMINIST ANALYSIS OF ANITA NAIR'S *LADIES COUPE***



*Project submitted to St. Teresa's College (Autonomous) in partial fulfilment  
of the requirement for the degree of BACHELOR OF ARTS in English  
Language and Literature*

*By*

**Ummuhani Nazreen  
Register No: AB19ENG042  
B.A. English Literature  
St. Teresa's College  
(Autonomous)  
Ernakulam  
Cochin-6  
82 011  
Kerala**

*Supervisor*

**Lakshmi Priya P  
Santhosh  
Department of  
English  
St. Teresa's College  
(Autonomous)  
Ernakulam**

**Kerala**  
**March 2022**

**DECLARATION**

I hereby declare that this project entitled “A Feminist Analysis of Anita Nair’s *Ladies Coupe*” is the record of bona fide work done by me under the guidance and supervision of Lakshmipriya P Santhosh, Assistant Professor, and Department of English.

Ernakulam  
March 2022

Ummuhani Nazreen  
Register No: AB19ENG042  
B.A. English Literature  
St. Teresa’s College  
(Autonomous) Ernakulam

## CERTIFICATE

I hereby declare that this project entitled “A Feminist Analysis of Anita Nair’s *Ladies Coupe*” is the record of bona fide work done by me under the guidance and supervision of Lakshmipriya P Santhosh, Assistant Professor, and Department of English.

Lakshmipriya P Santhosh

Department of English

Ernakulam

St. Teresa’s College (Autonomous)

March 2022

Ernakulam

**An Abstract of the Project entitled**  
**A Feminist Analysis of Anita Nair's *Ladies Coupe***

**By**

**Ummuhani Nazreen**

**BA English Language and Literature**

**St Teresa's College (Autonomous)**

**Ernakulam**

**Register Number: AB18ENG042 (2019-22)**

**March 2022**

**Supervising Teacher: Ms. Lakshmi Priya Santhosh**

Sexual harassment, domestic violence, salary disparities between men and women, eating disorders and body image, sexual and reproductive rights, honor crimes, and female genital mutilation are among the primary reasons and the problems that feminists are concerned about. Feminism refers to a broad spectrum of political groups, philosophies, and social movements with the purpose of defining, establishing, and achieving gender equality in political, economic, personal, and social matters. Feminism's basic assumption is to promote women's equality and fairness in all aspects of life, as well as to provide opportunities for women to have equal access to resources that are ordinarily freely available to males. Since then, feminism has expanded to include a wide range of topics, including the history of women's oppression and ways to overcome the "anxiety of authorship" by creating their own literary canon. Because there are multiple psychosocial and cultural conceptions of femineity, feminism is a serious endeavor to analyze, explain, and clarify.

Feminist theory encompasses more than just women's rights and gender equality. It also includes guiding principles that can be applied to a variety of situations. This research undertakes using by using the “qualitative” ethnographic method. This study demonstrates how the postpartum depression self-help movement exemplifies five feminist methodology features: a gender perspective, a focus on women's experiences, reflexivity, participatory techniques, and social action.

## ACKNOWLEDGEMENT

I would like to thank God for giving me proper guidance throughout the process of crafting this project.

I would like to extend my gratitude towards Dr. Lizzy Mathew, Principal, St Teresa's College (Autonomous) for her support.

I am deeply indebted to my guide, Lakshmipriya P Santhosh, Department of English, St Teresa's College (Autonomous), who has acted as a constant pillar of support from the very beginning; without whose guidance and constructive feedback I would have never completed this project.

I am greatly thankful for Dr. Latha Nair, Head of Department of English, St Teresa's College (Autonomous) for her constant encouragement and motivation.

I am grateful for Dr. Tania Mary Vivera, who has given us a strong base on Research Methodology and all other faculty members of the department for their help and encouragement.

Ummuhani Nazreen

## CONTENTS

Introduction	1
Chapter 1: Understanding Feminism	3
Chapter 2: A Feminist Study of <i>Ladies Coupe</i>	8
Conclusion	13
Works Cited	14

## Introduction

Feminism is a set of socio-political movements and philosophies aimed at defining and establishing gender equality in political, economic, personal, and social terms. Feminism is the belief that cultures priorities men's viewpoints and that women are treated unfairly within those communities. Efforts to alter this include combating gender stereotypes and ensuring that women have similar educational, professional, and interpersonal opportunities and outcomes as men.

*Ladies Coupé* by Anita Nair covers the journey of five middle-aged women, including the protagonist, Akhila, who is looking for freedom. This book was well-received when it was released in 2001. It depicts the lives of women in India who are defeated by fate and society, resulting in their life's consequences.

She is the author of the bestseller novels *The Better Man* and *Ladies Coupe*, as well as poetry collections *Malabar Mind* and *Satyr of Subway* and *Eleven Other Stories*, and two children's books titled *The Puffin Book of World Myths and Legends*. She is also the editor of *Where the Rain Is Born*, an anthology of literature about Kerala. When Anita Nair was working as the creative director of an advertising agency in Bangalore, she published her first book, *Satyr of the Subway*, a collection of short stories that she sold to Har-Anand Press. The Virginia Center for the Creative Arts awarded her a fellowship for the book. *Ladies Coupé* was named one of India's top five books in 2002. The work is about women's experiences in a male-dominated society, and it is conveyed with remarkable insight, solidarity, and wit. *Ladies Coupe* (2001) was named one of the top five books of the year in 2002 and has been translated into over twenty-five languages worldwide. *Mistress* (2003), *Adventures of Nonu*, the *Skating Squirrel* (2006), *Living Next Door to Alice* (2007), and *Magical Indian Myths* (2007) are among Nair's other works

(2008). Many travelogues are among Nair's writings. Anita Nair, a best-selling novelist, has debuted as a playwright with her play *Nine Faces of Being*. The plot is based on Nair's novel *Mistress*. In May 2012, she received the Kerala Sahitya Academy Award for her contribution to Literature and Culture, as well as the Arch of Excellence Award for Literature from the All-India Achievers' Conference in New Delhi.

Anita Nair is one of India's most well-known female authors. With fun and humor, she depicts the current state of women in society through her expression. Story telling is an art form, and she excels at it. She delves into the human mind and takes the reader on a fantastic journey by telling engrossing stories about colorful and distinctive characters. She invokes real-life events, and her style often leads readers to believe that the character is one of their neighbors or a friend. The book's title itself makes a comment about how women are treated in Indian society. A women only queue in a ticket reservation counter, or a 'ladies' special bus, or even a *Ladies Coupe*, is prevalent. Why do Indian women get such treatment? It's because it's assumed that a woman can only thrive in an all-environment, women's which we know only exists on Venus, and that a woman requires a man to survive in the world we live in. Every Indian woman wants to avoid this stereotype, but can a woman truly thrive without the protection of a man? This study is all about the feminism that is being happened in our society. And due to this how the women are being depressed in our society and how they feel alone. In feminism itself there are many branches where I look at the liberal feminism precisely. The book has eleven chapters in it. From this Novel it is trying to prove that how women's can be independent without a partner and become successful in life. The next topic of this project leads to the liberal feminism and further on to all the other topics which is necessary for the completion of this project.



## Chapter 1

### Understanding Feminism

Feminism is a set of socio-political movements and philosophies aimed at defining and establishing gender equality in political, economic, personal, and social terms. Feminism is the belief that cultures prioritize men's viewpoints and that women are treated unfairly within those communities. Efforts to remedy this include combating gender stereotypes and ensuring that women have equal educational, professional, and interpersonal opportunities and outcomes as men. There are four types of feminism such as radical, Marxist, liberal and difference. Liberal feminism owes its origins to classical liberalism and has been associated with social liberalism since the late 1800s. Beginning in the late 18th century, liberal feminists sought suffrage for women in the hopes of gaining individual independence.

Liberal feminism sometimes known as mainstream feminism is a subset of feminism that focuses on achieving gender equality through political and legal reform within a liberal democratic framework. It works within the mainstream society's structure to integrate women into it. Liberal feminism emphasizes the public sphere, particularly laws, political institution, education and jobs, and sees the denial of equal legal and political rights as the primary barrier to equality. This is related with centrism and is inclusive and socially progressive while broadly accepting current institutions of power in liberal democratic states. It actively encourages men to participate in feminism, and both women and men have always been active participants in the movement, progressive men played a crucial role alongside women in the fight for equality.

Liberalism is a set of ideologies that emphasize the importance of liberty and believe that a just state ensures individual liberty. Liberal feminists support this ideal and function for the state, as well as women's freedom, Liberalism is divided into two types due to a disagreement over how freedom should be defined this disagreement also splits liberal feminism. Some liberals consider freedom to be the absence of coercion. In the literature, such people are referred to as classical liberals. Classical liberals were committed to individualism, liberty, and equal rights. They believed these goals required a free economy with minimal government interference. This is appropriate given the historical perspective they hold. There are, however, current classical liberals, including contemporary classical liberal feminists, as readers shall see. Classical liberal feminism and egalitarian liberal feminism are both doctrinal families with substantial internal distinctions, which this essay aims to explain. Nonetheless, the difference in thinking about freedom between classical and egalitarian liberal feminists has significant implications for how each specifies the content of a liberal feminist agenda, and what role the state is assigned.

Some egalitarian liberal feminists believe that personal social arrangements should not only be freely selected, but also be marked by fairness or justice. To demonstrate one way in which heterosexual personal relationships frequently fail to be fair or just, Jean Hampton draws on the contractually tradition in moral and political philosophy. According to Hampton, a personal relationship is only fair if both parties could rationally accept the cost and benefit distribution if it were the topic of an educated, unforced agreement in which we think of ourselves as driven only by self-interest. Of course, many women enter or stay in relationships for a variety of reasons, including affective benefits; for example, women often find satisfaction in pleasing others or carrying out a duty. Hampton does not dismiss them because she believes

that a woman's affective nature is unrelated to her essential self. She also doesn't dismiss them because she believes this aspect of a woman's nature is unimportant. Because the affective benefits of partnerships are not received from the other, but rather from one's own character,<sup>5</sup> her test rules them out. While they may, and probably should, play a role in a woman's overall decision to enter or stay in a relationship, Hampton believes they should not play a role in determining whether a relationship is fair. A partnership is fair or just, as Linda Radzik says in her defense of Hampton, if the benefits that flow from one to the other are equal, that is, if one provides as much as she receives. When one party receives much more from the other than he offers, he is rejecting the other's lawful right to reciprocity.

This test formalizes a key insight of the feminist movement, personal relationships, particularly traditional heterosexual relationships, are frequently unfair to women, and even exploit women's proclivity for caring about others. This type of injustice is not uncommon. As a result, Hampton's test allows critique of a broad range of human social behavior. Hampton, on the other hand, does not advocate for women to stop valuing other people's happiness or obligation fulfillment. Instead, she urges the feminist movement to foster sensitivity and aversion to injustice in both women and men, as well as offer solutions.

Relationships do not have to be just in the way Hampton suggests in procedural accounts of personal autonomy. According to procedural accounts, the decision to enter or remain in a personal relationship in which one contributes more than she receives can be autonomous. As a result, the emphasis should be on ensuring that women are not forced into or unable to leave them. To be sure, Hampton's description of fairness in personal relationships can help women and men think about their own choices. It prompts thought about how one's own preferences

influence the distribution of rewards and burdens in a partnership. Furthermore, moral criticism of relationships that take advantage of women's preference serves as a reminder that relationships might be different. By expanding the imagination, this reminder increases personal agency. As a result, procedural theories of personal autonomy might contain Hampton's test, not as a definitive of the acceptability of social arrangements, but as a contribution to the kind of contemplation on the good life on which individual autonomy is based.

Liberal feminism's critics object to a lack of critique of basic gender connections, a reliance on state action that links women's interests to the powerful, a lack of class or race analysis, and a lack of understanding of how women vary from men. Liberal feminism is frequently accused by critics of assessing women and their accomplishments by masculine norms.

White feminism is a type of liberal feminism that assumes that the problems that white women suffer are the same problems that all women face, and that unity around liberal feminist ideals is more essential than racial equality and other such aims. Intersectionality is a philosophy that was formed in response to liberal feminism's prevalent racial blind spot.

In recent years, liberal feminism has been associated with a type of libertarian feminism known as equality feminism or individual feminism. Individual feminism frequently rejects legislative or state action, preferring instead to focus on improving women's talents and abilities so that they can compete more effectively in the world as it is. This feminism is opposed to legislation that provide men or women benefits or privileges.

Liberal feminism, according to Jagger, who is an English-born American feminist philosopher. She is a Distinguished Research Professor at the University of Birmingham in the

United Kingdom and a College Professor of Distinction in the Philosophy and Women and Gender Studies departments at the University of Colorado, Boulder. She was one of the first philosophers to bring feminist concerns into the discussion. It “thought and work that focuses more on concerns including as equality in the workplace, education and political rights. Liberal feminism also considers how one’s personal life affects or improves public equality.”<sup>7</sup>

As a result, liberal feminist favor equal marriage and increased male engagement in childcare. Support for abortion and other reproductive rights is linked to a sense of control and autonomy over one’s own life. Putting an end to domestic abuse and sexual harassment removes barriers to women gaining equality to men.

Gender equality in the public domain such as equal access to school, equal pay, ending employment sex segregation and better working conditions, is the core goal of liberal feminism. From this perspective, legal change would enable these objectives to be realized. Issues in the private realm are of significance because they influence or obstruct equality in the public arena. Gaining equal access to traditionally male dominated occupations as well as being compensated and promoted similarly, is a critical goal.

## Chapter 2

### A Feminist Study of *Ladies Coupe*

Anita Nair is the most promising of the newcomers and a force to be reckoned with. Her first book. Her first novel, *The Better Man*, has established her as one of India's most self-aware novels, and her second, *Ladies Coupe*, is in some respects even better than her first, but it is impossible to compare the two because they are so dissimilar. In *Ladies Coupe*, Anita Nair explores the topic of patriarchy and how it represents an unequal relationship. The story is an attempt to demonstrate how, in real life, oppression and suppression can take many forms, including love, protection, and the guarantee of security. Patriarchy is a frequent concept in every woman's life, but Anita Nair illustrates the diversity within each woman with care, since she does not want to reduce women's lives to a single ideal.

*Ladies Coupe* is the stories of six women who meet by coincidence on a train journey. The protagonist, Akhilandeswari, listens to the stories of five other women in the compartment and tells her own experience in parts and pieces, hoping to find an answer to the dilemma that has plagued her entire life: Can a woman be happy without a male or does she need a man to feel complete? The narrative jumps back and forth between the past and the present, so we see women who are humiliated and debased in addition to the five ladies in the compartment. Sunita Sinha says "Nair's India suffers from a patriarchal system that has sought in various ways to restrict, degrade, and debase women. The novel's question not only rattles the conceptual ground of man's patriarchal role in our traditional society, but also implies the presence of an alternate world". (149)

Akhila is a forty-five-year-old spinster who fills the roles of daughter, sister, aunt, and family breadwinner. As her father passes away unexpectedly, she is forced to face the entire family's sorrow. Indra Devi says "Anita Nair undoubtedly hints at the family's comfortable acceptance of her as the family's head of household on a spot typically reserved for the patriarch in both the colonial and post-colonial periods," (220). When Akhila sees a man at the train station surrounded by his entire family of uncles, aunts, cousins, and grandparents, she recognizes a resemblance between him and herself. Akhila stared at the man who carried the weight of other people's dreams on his shoulders. That was something she was aware of. That she could realize. "Akhila felt the iron bands around her chest begin to loosen: Dare I breathe again? Akhila felt the iron bands around her chest begin to loosen when Narsi her brother became the first graduate and secured a teaching job, while Narayan her other brother entered the tank factory as a machinist. Do I dare to dream once more?"(77 Nair's). Can I start feeling like a woman again now that the boys are men? Even though she was the family's breadwinner, she was not seen as the family's leader, and her needs and desires were ignored. Narsi did not ask anyone's permission to marry because he was a guy; instead, he "decided" to marry, "Narsi decided he wanted to get married" (77 Nair's). No one could fault his choice, and there was nothing anyone could say save perhaps - Don't you think you should wait for your elder sister to get married before you think about a bride and a family? He said when he told his family he was planning to marry the principal's daughter. However, who was it who delivered this rebuke? Both Narayan and Narsi married in the same hall, on the same day and at the same time. What about you? Akhila waited for Amma or her brothers to question about her marriage, but they never did.

Since Appa's death, you've been the family's leader. "Don't you want a husband, kids, and a place to call home"(77 Nair's)? Even though Akhila had performed all the tasks of a family head <sup>10</sup> her brothers and sister, she was not recognized as the true head simply because she was a woman. "Perhaps you should ask your brothers for permission first" (150 Nair's), Amma replies, implying that she expected her to acquire permission from her brothers, the family's men, before going on an office tour. "You might be older, but you are a woman, and they are the men of the family (150 Nairs)", Amma simply explains when Akhila argues that she is their elder sister and that she should not have to ask their permission. Akhila's interaction with a classmate inspired her to consider living a life of her own. However, patriarchal dominance emerged even then. "Do you think the men will accept to this (204 Nair's) "? Akhila brazenly asks Padma after telling her about her plan to live alone. "Do you believe they'll let you live alone? (204 Nair's)" says the narrator. "For heaven's sake, I don't need anyone's approval, (204 Nair's)" Akhila declares defiantly, and Padma mocks her by saying, They are the family's men. Padma, Narsi, and Narayan, for example, were deeply established in the patriarchal society and could not handle the concept of a woman living alone. Akhila's defiance, on the other hand, was so powerful that she boarded the train to Kanyakumari.

Anita Nair has used the role of Karpagam to raise awareness of women's aspirations and the need for self-expression in the society. Karpagam is portrayed as a strong woman attempting to define herself in a patriarchal society. She is a widow, but unlike most widows, she dresses in Kumkum and bright colors. She is a brave woman who defies patriarchy by declaring, I don't care what my family or anyone else thinks. I'm who I'm supposed to be. And I have the same right as everybody else to live my life as I see fit. Tell me, as young ladies, didn't we wear bright

colors, jewelry, and a bottu? It makes no difference if she is married or not, or if her husband is living or dead. Who made these laws in the first place? Some man who couldn't face the thought that his wife was still attractive to other men after he died. When she says, I live alone, she demonstrates incredible defiance. I've been doing it for a long time. Akhila, we are a formidable force. Whatever it is that you believe you want to do. Live on your own. Make a life for yourself that prioritizes your needs. Karpagam, are you genuine or are you some goddess who had come here to lead me out of this? Akhila wonders. Karpagam, are you real or are you some goddess who had come here to guide me out of this? As a result, Anita Nair portrays Karpagam as a bold woman who defies patriarchy's wider framework that denies women their personal independence.

Anita Nair utilizes characters such as Akhila's mother to demonstrate how women are staunch supporters of the patriarchal establishment that has imposed stringent social, political, and economic restrictions on women. Akhila's mother is a traditional and conservative woman, a loving wife who believes that a wife is always inferior to her husband. She is the type of lady who never makes her own decisions, instead relying on her spouse because she feels, He knows best. She says, a goodwife learned to put her husband's interests ahead of everyone else's, including her father's. A good wife heeded her husband's instructions and carried them out. It's preferable to accept the fact that the woman isn't as good as the husband. There will be no conflict or discord in this manner. Accepting one's place in life and living accordingly is much easier and simpler. I don't approve of what Karpagam's mother is doing, Akhila says when her mother wants to take music lessons because Karpagam's mother teaches dancing. I want my wife to take care of my children and me, she tells Akhila, recalling what her father told her when they

were first married. I don't want her to be so preoccupied with her job that she neglects the house or neglects my needs. Even after her father's death, Akhila's mother allows her eldest daughter Akhila to shoulder the family's obligations, relying on her sense of duty to keep them safe, secure, and comfortable. Sarasa mami's family is brought in by Anita Nair, who is dealing with a similar predicament. Sarasa mami goes to every door, begging to be taken as a servant - maid, as Subramani Iyer, Sarasa mami's husband, dies. However, every community treats her as if she were a beggar, handing her a handful of grains, forcing her to sell her daughter Jaya to make ends meet. This family is chosen by Anita Nair as a total contrast to Akhila's family in order to portray how women's moral difficulties are caught in social and emotional situations, fighting oppression and fate.

Akhila arrives at Kanyakumari as a powerful lady to rediscover her "self" after listening to the lives of various women in the car. The more she wants to be free of the life she has lived for others, the more she desires more of her own life, which is more of Hari, and she follows through on her plan to reconnect with him over the phone. In the end, she triumphs in her resistance of patriarchy. She defies patriarchal ideologies that have enslaved her and prevented her from discovering her "self."

Anita Nair has given an increasingly popular concept of patriarchy in her work *Ladies Coupe*, in which a woman is limited by tradition to be dependent on men, handicapped to see her own strength. She has shown her women struggling side by side due to patriarchy, but she has also offered them a sign of rebellion against patriarchy at the end. Her women have been portrayed as bright, inquisitive women who are fed up with the injustice and revolt that has been perpetrated against them. As a result, Anita Nair's women highlight the issue of patriarchy's

consolidation of their way of life, seeing it not only as a site of oppression at home and in society, but also as a battleground to defeat their oppressors.

## Conclusion

*Ladies Coupe* by Anita Nair explores the various realities of women as subalterns and highlights the diversity within and among women. It's about Akhila, Janaki, Margaret, Sheela, Prabha Devi, and Marikolanthu, six women of diverse ages, classes, and experiences, who happen to meet in the *Ladies Coupe*, a second-class compartment on Indian railways. The coupe represents their existence as well as the place they held in life and society. They tell their stories to help Akhila, the heroine, find a solution to the question of whether a woman in a male-dominated society can live a life free of men. They take distinct approaches to the problem of women's subordination, and they propose to Akhila equally diverse yet effective alternatives.

The plot of *Ladies Coupe* is based on a trip, with a storyline that travels backwards and forwards in time to determine the future. This is a voyage of self-discovery and recognizing one's value as a member of society. The title itself is a metaphor for a woman's journey from birth to death. *Ladies Coupe* is a compartment reserved exclusively for women, which can be compared to their compact world, where they can share their smiles, tears, marital life, lovers, and children – the most private and special moments of their lives – without fear of being exposed, because they are all strangers to each other and are unlikely to meet again. The novel's appeal stems from its vivid description and orderly sequence of events. "It is a work in which fiction combines with reality and where female voices are authentic," Clara Nubile explains. Indeed, it is a gendered tale that inspires all women with hope and strength."

Anita Nair depicts how modern Indian women struggle to overcome numerous difficulties in terms of sexuality, economics, and domesticity. In her works, male and female

14  
protagonists combat their interpersonal problems without regard for success, and they frequently achieve some type of harmony. Novelists, like readers, are aware that emotion cannot be analyzed logically. They frequently emphasize on the man-woman relationship's lack of emotional fulfilment. Considering this, Anita Nair authored *Ladies Coupe*, a compelling story that delineates feminine sensibility. The experiences of six female characters, Janaki, Sheela, Margaret Shanti, Prabha Devi, Marikolunthu, and Akhila, are primarily projected in this delineation.

*Ladies Coupe* deconstructs the religious, traditional, and ideological, all of which are taken for granted. Akhila's family does not allow her to marry and start a family; instead, she is expected to provide for them. Anita Nair presumably hints at the family's comfortable acceptance of her as the family's head of household on a spot typically allocated for the patriarch in both the colonial and post-colonial periods.

Akhila's outlook changes dramatically as a result of the personal stories told by these five female characters. She realizes how essential desire and fulfillment were in their life. This reminds her of her own brief meeting with Hari, and she is compelled to form a physical and emotional link with him to quench her soul's thirst. A young kid falls for her in Kanyakumari, and she enjoys a night's relationship with him before dialing Hari's number. She realizes that, even at the age of 45, she has a right to love. Her choice is a statement of defiance against society. She has undergone a spiritually and emotionally liberating trip, and Akhila shapes herself and discovers life. As a result, the novel redefines women's lives, and a feminist voice can be heard throughout.



## Works Cited

- “Anita Nair's Ladies Coupe - Core Chapter.” *Blogger*, 23 Mar. 2020,  
4aprilrose24.blogspot.com/2020/03/anita-nairs-ladies-coupe-core-chapter.html?m=1.  
Accessed 3 Jan. 2021.
- Baehr, Amy R. “Liberal Feminism.” *Stanford Encyclopedia of Philosophy*, Stanford University,  
31 Dec. 2020, <https://plato.stanford.edu/entries/feminism-liberal/>.
- Biradar, Basavaraj. “Critical Analysis of Ladies Coupe by Prof.Imran Mulla.” *Blogger*, 12 Mar.  
2017,  
<http://basava097.blogspot.com/2017/05/critical-analysis-ofladies-coupeby.html?m=1>.  
Accessed 3 Jan. 2021.
- Flavia, Silvia. “The Concept of Patriarchy and Female Defiance in Anita Nair’s Ladies Coupe.”  
*The Journal For English Language and Literary Studies*,  
[www.tjells.com/article/138\\_SYLVIA%20final.pdf](http://www.tjells.com/article/138_SYLVIA%20final.pdf) .
- Lewis, Jone Johnson. "Liberal Feminism." ThoughtCo, Feb. 16, 2021,  
[thoughtco.com/liberal-feminism-3529177](http://thoughtco.com/liberal-feminism-3529177).
- Mason, Susan A. “Qualitative versus Quantitative Approaches.” *SpringerLink*, Springer, New  
York, NY, 1 Jan. 1970,  
[https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-1698-3\\_167](https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-1698-3_167).
- Nair, Anita. *Ladies Coupe*. New Delhi: Penguin Books, 2001.



Project Report

On

**A PRIOR ON USING MATHEMATICAL  
MODELLING TO UNDERSTAND COVID 19  
SCENARIO**

*Submitted*

*in partial fulfilment of the requirements for the degree of*

**MASTER OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**UTHARA M S**

**(Register No. SM20MAT016)**

**(2020-2022)**

*Under the Supervision of*

**Smt VEENA V S**



**DEPARTMENT OF MATHEMATICS  
ST. TERESA'S COLLEGE (AUTONOMOUS)**

**ERNAKULAM, KOCHI - 682011**

**APRIL 2022**

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



CERTIFICATE

This is to certify that the dissertation entitled, **A PRIOR ON USING MATHEMATICAL MODELLING TO UNDERSTAND COVID 19 SCENARIO** is a bonafide record of the work done by Ms. **UTHARA M S** under my guidance as partial fulfillment of the award of the degree of **Master of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 27-05-2022.

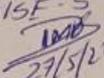
Place: Ernakulam

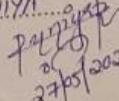
  
Smt **VEENA V S**  
Assistant Professor,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.



  
Dr. Ursala Paul  
Assistant Professor and HoD of the department,  
Department of Mathematics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1: **DIVYA... MARY... DAISE S**  
  
27/5/2022

2: **MARY... RUVIYA**  
  
27/5/2022

## DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **Smt VEENA V S**, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Place: Ernakulam.

Date: 27-05-2022



UTHARA M S

SM20MAT016

## ACKNOWLEDGEMENTS

Though I have taken efforts in this project, it would not have been possible without the kind support and help of many individuals. I would like to extend my sincere thanks to all of them. I am highly indebted to **Smt NISHA OOMMEN** and **Smt DHANALAKSHMI O M** for their guidance and constant supervision as well as for providing necessary information regarding the project and also for their support in completing the project. I would like to express my gratitude to my parents for their kind cooperation and encouragement which help me in completion of this project.

I am very grateful to my project guide **Smt VEENA V S** for the immense help during the period of work . Her constant guidance and willingness to share her vast knowledge made us understand this project and its manifestations in great depths ad helped us to complete the assigned tasks on time.

In addition, very energetic and competitive atmosphere of the department had much to do with this work. I acknowledge with thanks to faculty, teaching and non-teaching staff of the department and colleagues.

I am also very thankful to HoD Dr.Ursala Paul for her valuable suggestions, critical examination of work during the progress.

Place: Ernakulam.

Date: 27/05/2022

  
UTHARA M S

SM20MAT016

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENTS</i> . . . . .	iv
<i>CONTENT</i> . . . . .	v
<b>1 Introduction</b>	<b>1</b>
<b>2 MODEL ON THE DISEASE</b>	<b>5</b>
2.1 THE MATHEMATICAL MODEL . . . . .	7
2.2 METHODOLOGY . . . . .	8
2.2.1 FEASIBLE SOLUTION . . . . .	8
2.3 EQUILIBRIA . . . . .	9
2.4 GENERAL REPRODUCTION NUMBER . . . . .	11
<b>3 MODEL ON LOCK DOWN</b>	<b>12</b>
<b>4 MODEL ON COVID VACCINATION</b>	<b>14</b>
4.1 VACCINATION MODEL . . . . .	16
<b>5 Conclusion</b>	<b>18</b>
<i>REFERENCES</i> . . . . .	20

# Chapter 1

## Introduction

---

The novel coronavirus (COVID-19) pandemic that emerged from Wuhan city in December 2019 overwhelmed health systems and paralyzed economies around the world. It became the most important public health challenge facing mankind since the 1918 Spanish flu pandemic. The disease started as an outbreak of pneumonia of unknown cause. It rapidly became a devastating pandemic, spreading to every country on earth, and inflicting severe public health and socio-economic burden globally. By the end of August 2020, the pandemic caused over 25.6 million confirmed cases and 854, 000 deaths globally. China was the first epicenter of COVID-19, followed by Europe.

COVID-19 is transmitted from human-to-human through direct contact with contaminated objects or surfaces and through inhalation of respiratory droplets from both symptomatic and -infectious humans. There is also limited evidence that the virus can be exhaled through normal breathing. The incubation period of COVID-19 ranges from 2 to 14 days, and most infections (over 80%) show mild or no clinical symptoms of the disease. The common symptoms of the disease include fever, coughing and shortness of breath for mild cases, and pneumonia for severe cases. Although data clearly shows that most of the COVID-19 related deaths and severe cases typically occur in the elderly and people with comorbidities, such as people with diabetes, hypertension, obesity, kidney disease and other conditions that suppress or compromise the

immune system, younger people and front line healthcare workers are also at high risk of acquiring COVID-19 .

Corona viruses (CoVs) are a major group of RNA viruses that cause diseases in mammals and birds. In humans, these viruses are associated with multiple respiratory diseases of varying severity. For instance, the mild form of coronavirus infections causes diseases such as the common cold, while the severe form can cause diseases such as the severe acute respiratory syndrome (SARS-CoV), middle eastern respiratory syndrome (MERS) and COVID-19 (caused by SARS-CoV-2). “Coronavirus” is derived from the Latin word “Corona”, meaning crown or wreath, which is related to the characteristic appearance of the virions of the virus. Zoonotic scientists estimated that there are millions of coronaviruses in the wild, thereby making humans vulnerable to coronavirus pandemics periodically (since, owing to their genetic makeup, human coronaviruses are rated among the most rapidly evolving human viruses). It is believed that human coronaviruses have their origins in bats and rodents, and some human activities, such as urbanization and poultry farming, help in expediting their evolution.

Two pandemics of coronaviruses have occurred in recent years. These include the 2002/2003 pandemic of severe acute respiratory syndrome (SARS-CoV), a highly transmissible disease which started in the Guangdong province of China and spread to 29 countries (causing 8000 cases and 744 fatalities globally). Palm civet and bats were the natural reservoirs of SARS-CoV, which has a mortality rate of 10% . In 2012, a pandemic of the middle eastern respiratory syndrome (MERS-CoV) started out of Saudi Arabia and spread to 27 countries, causing 2519 cases and 866 deaths by January 2020. Over 80% of MERS-CoV occurred in Saudi Arabia. MERS-CoV, which was believed to have originated from bats and then likely spread from infected dromedary (Arabian) camels to humans, has a mortality rate of about 35%. SARS-CoV and MERS-CoV have similar clinical symptoms.

Consequently, control and mitigation efforts against COVID-19 are fo-

cused on the implementation of non-pharmaceutical interventions (NPIs), such as community lockdown, maintaining social (physical)-distancing, using face masks in public, quarantine of suspected cases, isolation and hospitalization of confirmed cases, surveillance and serology testing and contact tracing. This study introduces some of the basic principles and methodologies for using mathematical modeling, backed by rigorous analysis and statistical data analytics, to gain insight into the transmission dynamics and control of infectious diseases, such as COVID-19, in human population.

Mathematical models have the potential to trace and predict the epidemic trajectory under different scenario. Various mathematical, statistical models have been proposed to understand the dissemination trajectory for a pandemic. Among these models, Susceptible(S)-Infected (I) -Recovered(R) model (SIR model) has been frequently used in past to predict the influence of HIV virus. Recently, SIR model has also been applied for prediction of COVID-19 trajectory and its epidemic peak. However, such studies have been carried out at the very earlier stage of pandemic. Moreover, such studies are primarily focused on the COVID-19 spread tracing under normal circumstances i.e. containment strategies (policies) have not been considered in such studies for COVID-19 prediction

In the absence of a ready-to-use vaccine, and besides medical and biological research, mathematical models can play an important role in understanding and predicting disease transmission. Moreover, it helps to implement appropriate measures and efficient strategies to control the pandemic's spread and mitigate its impact.

The objectives of the present study are:

1. To analyze the transmission dynamics of COVID-19 among humans via mathematical modeling.

2. To investigate the impact of control strategies such as lockdown, quarantine and isolation to control the spread of the global pandemic (COVID-19).
3. The impact of lockdown/quarantine of susceptible individuals. What would be the outcomes if there is partial/full lockdown?
4. The impact of vaccine.

In the study I firstly investigate about the model of COVID 19. I introduced a flow chart to observe the flow of the disease. Then next I have studied about the lockdown and its impact. Finally I have studied about the vaccination model and have made a flow chart to know the effect of vaccine. The flow chart is helpful to understand the models.

## Chapter 2

# MODEL ON THE DISEASE

---

In this work, we shall study the transmission mechanism of COVID-19 using a deterministic compartmental model.

Here, the population is divided into four compartments:

$S$  is the compartment of susceptible individuals;

$E$  is the exposed individuals (Infected but not infectious)

$I$  is the infected individuals

$R$  is the removed individuals

$N$  is the total population

Then we can say that

$$N = S + E + I + R \tag{2.1}$$

Firstly a person will be in  $S$  then after getting the virus the person will move to  $E$  and after the latency period the person move to sub-population  $I$ . Then after getting treated or by death the person will move to the  $R$  sub-population. This is the step 1 for the modelling of the disease.

The figure 2.1 gives the pictorial representation of the flow of virus.

NOTE: A removed one can become susceptible again.

There are some factors that are responsible for the flow of the virus.

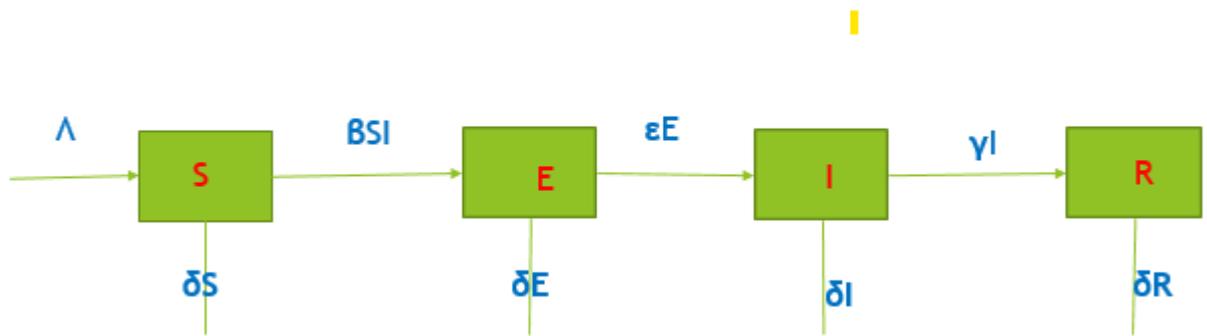


Figure 2.1: flow of the proposed model

1. The susceptible population  $S$  we can have new peoples coming into it (like giving birth). We can call this by birth rate .We represent this factor by the symbol  $\lambda$  .

2. In each subpopulation we can see some person removed by death (death caused by natural causes other than this disease).

That is the death rate of each subpopulation  $S, E, I, R$  are  $\delta S, \delta E, \delta I, \delta R$

NOTE: The largest among the death rates is  $\delta R$

3. Now a person move from subpopulation  $S$  to  $E$  by a factor called mass effect or bilinear incidence. This happen when a person get interact with a person having the corona virus. We denote this by the symbol  $\beta SI$

Here all the population is mixed together. So a susceptible can get contact with an in infected on this parameter is the reason for the need of quarantine.

Therefore,

The relative contact of infected people with susceptible population  
 $= \beta SI$

4. A person in  $E$  have latency period to get into the sub population  $I$ . So here the factor responsible is the latency period. We denote it

by  $\epsilon E$

.

5. The parameter that lead a person to move from infectious to removed is the average period of infection. We denote it by  $\gamma I$ .

The recovery rate =  $\gamma$

This is how a virus travel in the transport diagram.

## 2.1 THE MATHEMATICAL MODEL

Now we have to study about the change that is happening to the population over time due to these factors. In general what we need in a model is  $S$  to be large and  $I$  to be very small. From the figure 2.1 we arrive at a system of differential equations that tells how each population is changing over time.

The system of differential equation is as follows:

$$dS/dt = \gamma I - \beta SI - \delta S \dots \dots (2.2)$$

$$dE/dt = \beta SI - \epsilon E - \delta E \dots \dots (2.3)$$

$$dI/dt = \epsilon E - \delta I - \gamma I \dots \dots (2.4)$$

$$dR/dt = \gamma I - \delta R \dots \dots (2.5)$$

Here there is four parameters so we arrive into a system of differential equation with four equations.

## 2.2 METHODOLOGY

An aim of a model is to create a positive solution. To find this we need to find a feasible solution where we would have a solution or non-negative solution.

### 2.2.1 FEASIBLE SOLUTION

For a feasible solution the susceptible population to be positive and the others need to be zero. The total population will always be positive So that there is no disease. We denote the feasible region as  $\Sigma$ .

Since,

the total population  $N$  is always positive its rate of change will also be positive.

$$dN/dt \geq 0$$

We know that

$$dN/dt = d(S+E+I+R)/dt$$

For each terms we have expressions from the system of linear equations

That is,

$$\begin{aligned} \frac{dN}{dt} &= \gamma I - \beta SI - \delta S + \beta SI - \epsilon E - \delta E + \epsilon E - \delta I - \gamma I + \gamma I - \delta R \\ &= \lambda - \delta S - \delta E - \delta I - \delta R \\ &= \lambda - \delta(S+E+I+R) \\ &= \lambda - \delta N \\ &\Rightarrow \lambda - \delta N \geq 0 \\ &\Rightarrow \lambda \geq \delta N \end{aligned}$$

$$\Rightarrow \frac{\lambda}{\delta} \geq N \dots \dots (2.6)$$

$$\Rightarrow \frac{\lambda}{\delta} \geq S+E+I+R$$

This gives the feasible region. Thus the feasible region and the subset of  $R^4$  and is the ratio of birth rate and death rate.

### 2.3 EQUILIBRIA

This model give rise to two equilibrium's:

1. Disease free equilibrium
2. Endemic equilibrium

#### 1. Disease free equilibrium

It is the case where there is no disease. That is, the  $E, I, R$  sub populations eventually becomes zero.

#### 2. Endemic equilibrium

It is the case when the disease comes and the corona virus is going to stay in the population where none of the rates changes anymore.

\*To find the disease free equilibrium

Disease free equilibrium happens when  $E = 0$ . Now we know that if  $E$  is zero then eventually  $I$  and  $R$  becomes zero.

We know that

$$\frac{dI}{dt} = \epsilon E - \delta I - \gamma I$$

$$\Rightarrow \epsilon E = (\delta + \gamma)I$$

$$I = \epsilon / (\delta + \gamma) E$$

Also,

$$\begin{aligned} dR/dt &= \gamma I - \delta R \\ \gamma I - \delta R &= 0 \\ R &= (\gamma/\delta)I \\ \Rightarrow R &= (\gamma/\delta)\epsilon/(\delta + \gamma)E \end{aligned}$$

Also,

$$\begin{aligned} dS/dt &= \lambda - \beta SI - \delta S \\ 0 &= \lambda - \beta SI - \delta S \\ \Rightarrow \frac{\lambda}{\delta} &= S \end{aligned}$$

Therefore the disease free equilibrium  $P_0$  is  $(S, 0, 0, 0) \dots \dots (2.7)$

*That is,  $P_0 = (\lambda/\delta, 0, 0, 0)$*

*\*To find the endemic equilibrium*

*Endemic equilibrium happens when  $E \neq 0$*

*This is the second phase of modelling. When  $E^* \neq 0$   $I^* \neq 0$*

*$R^* \neq 0$*

*Then*

*(2.3) + (2.4)*

$$\begin{aligned} \Rightarrow \beta SI - \epsilon E - \delta E + \epsilon E - \delta I - \gamma I \\ \Rightarrow \beta SI - (\delta + \gamma)I - \delta E = 0 \end{aligned}$$

Since we know the value of I , we get

$$\begin{aligned} \beta S\epsilon/(\delta + \gamma)EI - (\epsilon + \delta)E &= 0 \\ \Rightarrow [\beta S(\epsilon/(\delta + \gamma)) - (\epsilon + \delta)]E &= 0 \end{aligned}$$

There fore the endemic equilibrium is  $S^* = (\epsilon + \delta)(\delta + \gamma)/\beta\epsilon \dots \dots (2.8)$

## 2.4 GENERAL REPRODUCTION NUMBER

This is last phase of disease modelling.

**GENERAL REPRODUCTION NUMBER** is the number of new case any single individual is going to create. We denote this by  $R^0$ .

We know that ,

$$\begin{aligned} S &\geq \lambda/\delta \\ \Rightarrow (\epsilon + \delta)(\delta + \gamma)/\beta\epsilon &\geq \lambda/\delta \\ \text{iff } 1 &\geq \lambda\beta\epsilon/\delta(\epsilon + \delta)(\delta + \gamma) \end{aligned}$$

That is,

- $R^0 \leq 1$  is the disease free equilibrium
- $R^0 \geq 1$  is the endemic equilibrium

By the help of this study we can control the disease by using proper parameters such as quarantine , lockdown to make  $R^0$  less than 1 and this we will have a disease free equilibrium. This is how we model the corona virus disease

## Chapter 3

# MODEL ON LOCK DOWN

---

Due to the pandemic, where the worldwide infection and decease rate are really very alarming, the world needs quick recovery and for that, a proper prediction regarding the transmission range, transmission trend and decision making of lockdown measures are needed. From the existing literature, it is clear that there have been many attempts to predict the diseasespreading all over the world.However , the isolation and social distancing guidelines is required to analyse the decisions .The analysis for the effect of a lockdown was performed without the influence of the other control measures, like social distancing and mask wearing, to quantify its absolute effect. Hypothetical lockdown timing was shown to be the critical parameter in ameliorating pandemic peak incidence. More importantly, we found that well-timed lock downs can split the peak of hospitalizations into two smaller distant peaks while extending the overall pandemic duration. The timing of lock downs reveals that a “tunneling” effect on incidence can be achieved to bypass the peak and prevent pandemic caseloads from exceeding hospital capacity. The resolved factor R depends on S and I,so here we have restricted our implementation to SI. With the progress of disease spreading over a longer time, the new births have been considered. In this scenario, the system can be represented by

$$dS/dt = \lambda - (1 - \alpha)r(1 - \mu)SI - \alpha S \dots \dots (3.1)$$

$$dI/dt = (1 - \alpha)r(1 - \mu)SI - aI - \mu I \dots \dots (3.2)$$

This is the lockdown model .

Here,

$a$  = recovery rate

$\alpha$  = lockdown rate of susceptible

$\mu$  = isolation rate of infectious

To detect the effectiveness of social distancing approach we have divided the infected population into two categories by extending the conventional SEIR model. The first one is called Type I who are generally the detectable infected persons and the second one is called Type II who are undetectable infected person. If the probability of an infected people becoming detectable is denoted by  $p^1$  and becoming undetectable is  $p^2$ , then

$$p^1 + p^2 = 1 \dots\dots (3.3)$$

If the transmission rates among Type I and Type II people are denoted by  $r_1$  and  $r_2$  respectively and the recovery rates for the same area 1 ( $a_1$ ) and area 2 ( $a_2$ ), then

$$L^0 = p^1 (r_1/a_1) + p^2 (r_2/a_2) \dots\dots (3.4)$$

In practical scenario, Type II population has a higher transmission rate than that of a Type I population. The controllable measure of the disease is indicated by spectral radius which in turn shows that if the basic reproduction number  $L^0 \geq 1$  then there is some outbreak, and  $L^0 \leq 1$  then no outbreaks there.

Allowing all to keep their interpersonal contacts up-to a fraction of normal contacts and cancelling mass gatherings, these two approaches of maintaining social distance have been considered here.

The lockdown in the whole country announced by government has definitely acted as infection controller and helped to face the challenge of COVID-19 in the desired form.

## Chapter 4

# MODEL ON COVID VACCINATION

---

Vaccines can prevent infectious diseases. Vaccines do prevent measles, polio, hepatitis B, influenza and many others. When most people in a community are protected by vaccination, the ability of the pathogen to spread is limited. This is called ‘herd’ or ‘indirect’ or ‘population’ immunity. When many people have immunity, this also indirectly protects people who cannot be vaccinated, such as those who have compromised immune systems.

To bring this pandemic to an end, a large share of the world needs to be immune to the virus. The safest way to achieve this is with a vaccine. Vaccines are a technology that humanity has often relied on in the past to bring down the death toll of infectious diseases. Within less than 12 months after the beginning of the pandemic, several research teams rose to the challenge and developed vaccines that protect from SARS-CoV-2.

Vaccines greatly reduce the risk of infection by training the immune system to recognize and fight pathogens such as viruses or bacteria. Most research on COVID-19 vaccines involves generating responses to all or part of the spike protein that is unique to the virus that causes COVID-19. When a person receives the vaccine, it will trigger an immune response. If the person is infected by the virus later on, the immune system recognizes the virus and, because it is already pre-

pared to attack the virus, protects the person from COVID-19.

★How safe are the COVID-19 vaccines?

The safety requirements for COVID-19 vaccines are the same as for any other vaccine and will not be lowered in the context of the pandemic. Safety trials begin in the lab, with tests and research on cells and animals first, before moving onto human studies. The principle is to start small and only move to the next stage of testing if there are no safety concerns. Clinical trials are evaluating COVID-19 vaccines in tens of thousands of study participants to generate the scientific data and other information needed to determine safety and effectiveness. These clinical trials are being conducted by manufacturers according to rigorous standards. The COVID-19 vaccines are tested in a broad population of people – not only young, physically fit. volunteers, but also older people and people with underlying health conditions. After deployment, the vaccines will continue to be carefully monitored for safety and effectiveness.

★What is done to fast-track vaccine development in a public health emergency?

Continuous dialogue between developers and regulatory experts and early scientific advice helps speed up vaccine development. Advising companies on regulatory requirements helps ensure that standards of safety and efficacy are embedded early in the process and are not compromised by fast-track development. Resource mobilization for COVID-19 vaccines is done simultaneously which allows for accelerated development and manufacturing of vaccines. Companies are expanding manufacturing capacity and large-scale production, to facilitate vaccine deployment without delay once approved.

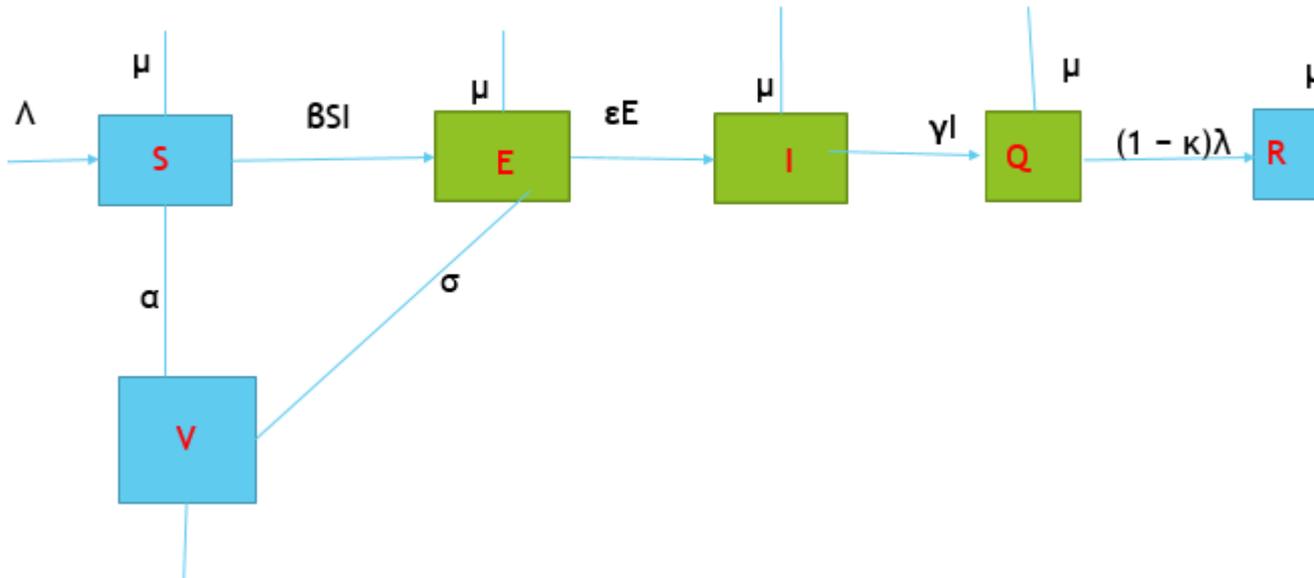


Figure 4.1: flow of vaccination model

## 4.1 VACCINATION MODEL

This work therefore aims to study the impact of vaccination on the COVID-19 spread. An extended SEIR model comprising of seven compartments—susceptible, exposed, infectious, quarantined, recovered, deaths, and vaccinated—is then proposed. First, we conduct a mathematical analysis to illustrate the non-negativity, boundedness, epidemic equilibrium, existence, and uniqueness of the endemic equilibrium, and the basic reproduction number of the mode .

We extend the SEIR model to seven compartments to simulate the epidemic of COVID19. Seven state variables are considered within a population, that is,  $S(t)$ ,  $E(t)$ ,  $I(t)$ ,  $Q(t)$ ,  $R(t)$ ,  $D(t)$ , and  $V(t)$ , denoting the number of susceptible, exposed (infected, but not yet infectious), infectious (not yet quarantined), quarantined (confirmed and infected), recovered, dead, and vaccinated cases, respectively. The disease transmission flow of the proposed model is sketched in Figure 4.1

The model is then governed by the following set of nonlinear ordinary differential equations:

Here,

$$\begin{aligned}
 \frac{dS(t)}{dt} &= \Lambda - \beta S(t)I(t) - \alpha S(t) - \mu S(t), \\
 \frac{dE(t)}{dt} &= \beta S(t)I(t) - \gamma E(t) + \sigma\beta V(t)I(t) - \mu E(t), \\
 \frac{dI(t)}{dt} &= \gamma E(t) - \delta I(t) - \mu I(t), \\
 \frac{dQ(t)}{dt} &= \delta I(t) - (1 - \kappa)\lambda Q(t) - \kappa\rho Q(t) - \mu Q(t), \\
 \frac{dR(t)}{dt} &= (1 - \kappa)\lambda Q(t) - \mu R(t), \\
 \frac{dD(t)}{dt} &= \kappa\rho Q(t), \\
 \frac{dV(t)}{dt} &= \alpha S(t) - \sigma\beta V(t)I(t) - \mu V(t),
 \end{aligned}$$

$Q(t)$  = the number of quarantined cases

$V(t)$  = the number of vaccinated cases

$\kappa$  = average quarantine time

In this vaccination modelling the original population  $N$  is again subdivided into 2. That is to  $Q$  and  $V$ .

Thus,

$$N = S + E + I + R + Q + V \tag{4.1}$$

In the transport diagram of vaccination model one person travel from  $I$  to  $Q$  from the and from  $Q$  to  $R$  by the recovery time.

The coefficients  $\lambda, \beta, \alpha, \mu, \gamma, \sigma, \delta, \kappa,$  and  $\rho$  represent the new births and new residents transmission rate divided by  $N$ , vaccination rate (rate of people who are vaccinated), average latent time, average quarantine time, mortality rate, average days until recovery days until death, respectively.

$\sigma$  is the vaccine inefficiency and  $(1-\sigma)$  is the vaccine efficiency.

If  $\sigma = 0$  the vaccine offers 100% protection against the disease.

# Chapter 5

## Conclusion

---

We have developed a mathematical model that is able to simulate the impact of SARS-CoV-2 variants and vaccines on the spread of COVID-19. It is based on the models proposed of disease. We were able to simulate other main mechanisms influencing the disease spread. In chapter 2 I discussed about the model. It is to be noted from this analysis that one of the very key parameters is the disease transmission coefficient  $\gamma$  which plays a significant role in determining the basic reproduction number  $R^0$ .

COVID-19 has been declared as pandemic by WHO and is currently become a major global threat. Prediction of a disease may help us to understand the factors affecting it and the steps that we can take to control it. The Government of India has taken preventive measures such as complete lockdown in the very early stage of disease, physical distancing and case isolation. The most important issue is that many healthcare professionals are visiting each and every household in the hot spot area across the country to trace and isolate infected persons to curtail the spread of disease.

In this work on chapter 3, we presented that isolation of the infected human overall can reduce the risk of future COVID-19 spread. Our model shows that the coronavirus spreads through contact and describes how fast something changes by counting the number of people who are infected and the likelihood of new infections. Those new infections are what induce the epidemic. For this reason, we think that this research

may lead to better guessing of the spread of this pandemic in the future. This paper is devoted to implement the coronavirus mathematical model containing isolation class. The reproductive number-related stability is discussed.

In future, these shortcomings will be attempted to overcome for getting improved result of the scenario and experimental results show that our proposed approach is well suited in its concerned domain. However, the prediction will change based on some external factors like government decisions and human actions, so we can't train the data as of today to be true for a different situation in future. Our modified SEIR model is not only productive than the traditional one, but also more capable of handling and adopting the current situation.

In the fourth chapter the vaccination models help us to know the effect of covid vaccine in the society. The flowchart proposed in the vaccination model and the disease modelling can help to figure out the flow of the virus due to the pandemic.

Thus, its forecasts are as reliable and can capture the dynamics of the pandemic. Due to real time change in data daily, the predictions will accordingly change.

*"if we control this contact rate, the control of the current disease is possible, otherwise it will be the worse"*

## REFERENCES

---

- [1] Abba B. Gumel, , Calistus N, Enahoro A. Iboi b. - Infectious Disease Modelling
  
- [2] Ahona Ghosh · Sandip Roy · Suparna Biswas - Mathematical Modelling For Decision Making of Lockdown during COVID-19
  
- [3] Anwar Zeb , Ebraheem Alzahrani and d Gul Zaman - Mathematical modeling of COVID-19 transmission: the roles of intervention strategies and lockdown
  
- [4] Brijesh P. Singh<sup>1</sup> Gunjan Singh<sup>2</sup> <sup>1</sup> Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005 <sup>2</sup> School of Applied Sciences, Amity University, Lucknow-226028 - Modeling Tempo of COVID-19 Pandemic in India and Significance of Lockdown
  
- [5] Central university of Rajasthan - Mathematical Modelling of COVID-19 transmission : the roles of intervention strategies and lock down
  
- [6] Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... Xing, X. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*.
  
- [7] Rabih Ghostine, Mohamad Gharamti - Mathematical modelling of Covid-19 with the effect of vaccine

- [8] Sarita Bugalia , Vijay Pal Bajiya - Mathematical modeling of COVID-19 transmission: the roles of intervention strategies and lockdown
- [9] Vipin Tiwari, Namrata Deyal , Nandan S. Bisht<sup>1</sup> <sup>1</sup>Department of Physics, KU, SSJ campus Almora (263601), Uttarakhand, India - Mathematical modelling based study and prediction of COVID-19 epidemic dissemination under the impact of lockdown in India
- [10] World Health Organization: Updated WHO advice for international traffic in relation to the outbreak of the COVID-19. WHO, Geneva. [https://www.who.int/ith/ COVID-19adviceforinternationaltraffic/en/](https://www.who.int/ith/COVID-19adviceforinternationaltraffic/en/). Accessed 11 Dec 2021