

# **A STUDY ON EDUCATIONAL LOANS PROVIDED BY COMMERCIAL BANKS IN ERNAKULAM**

## **PROJECT REPORT**

Submitted by

**AISWARYA BABU (REG. NO. AB19COM004)**

**AKSHARA EDWIN (REG. NO. AB19COM005)**

**ALEENA K.S (REG. NO. AB19COM006)**

Under the guidance of

**Ms. ANN THOMAS KIRIYANTHAN**

In partial fulfillment of requirements for award of the degree of

**Bachelor of Commerce**



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

**COLLEGE WITH POTENTIAL FOR EXCELLENCE**

Nationally Re-Accredited at "A++" Level

Affiliated to

**MAHATMA GANDHI UNIVERSITY**

Kottayam-686560

**MARCH 2022**

# **A STUDY ON EDUCATIONAL LOANS PROVIDED BY COMMERCIAL BANKS IN ERNAKULAM**

## **PROJECT REPORT**

Submitted by

**AISWARYA BABU (REG. NO. AB19COM004)**

**AKSHARA EDWIN (REG. NO. AB19COM005)**

**ALEENA K.S (REG. NO. AB19COM006)**

Under the guidance of

**Ms. ANN THOMAS KIRIYANTHAN**

In partial fulfillment of requirements for award of the degree of

**Bachelor of Commerce**



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

**COLLEGE WITH POTENTIAL FOR EXCELLENCE**

**Nationally Re-Accredited at "A++" Level**

Affiliated to

**MAHATMA GANDHI UNIVERSITY**

**Kottayam-686560**

**MARCH 2022**



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

**COLLEGE WITH POTENTIAL FOR EXCELLENCE**

Nationally re- accredited by A++ Level



**CERTIFICATE**

This is to certify that the project report titled **“A STUDY ON EDUCATIONAL LOANS PROVIDED BY COMMERCIAL BANKS IN ERNAKULAM”** submitted by **AISWARYA BABU, AKSHARA EDWIN and ALEENA K.S** towards partial fulfillment of the requirements for the award of degree of **Bachelor of Commerce** is a record of Bonafide work carried out by them in the academic year 2021-2022.

**Supervising Guide**

**Ms. Ann Thomas Kiriyanthan**

**Assistant Professor**

**Department of Commerce**

**Head of the Department**

**Ms. Ann Thomas Kiriyanthan**

**Assistant Professor**

**Department of Commerce**

**Place: Ernakulam**

**Date: 31.3.2022**



## DECLARATION

We, **AISWARYA BABU, AKSHARA EDWIN, and ALEENA K.S** do hereby declare that this dissertation entitled, “**A STUDY ON EDUCATIONAL LOANS PROVIDED BY COMMERCIAL BANKS IN ERNAKULAM**” has been prepared by us under the guidance of **MS. ANN THOMAS KIRIYANTHAN**, Assistant Professor, Department of commerce, St Teresa’s College, Ernakulam.

We also declare that this dissertation has not been submitted by us fully or partly for the award of any Degree, Diploma, Title or Recognition before.

**Place: Ernakulam**

**Date: 31.3.2022**



**AISWARYA BABU**



**AKSHARA EDWIN**



**ALEENA K.S**

## **ACKNOWLEDGEMENT**

We wish to acknowledge all those who helped us to complete this study. We thank God almighty for helping us and guiding us in the right path and who made all things possible.

We take this opportunity to express our profound gratitude and deep regards to our guide **Ms. Ann Thomas Kiriyanthan**, Assistant professor, St. Teresa's college, Ernakulam for her exemplary guidance, monitoring and constant encouragement throughout the course of this study. The blessings, help and guidance given by her time to time shall carry us a long way in the journey of life on which we are about to embark. She has taken pain to go through the project and make necessary corrections as and when needed.

We express our sincere thanks to the Director **Rev. Sr. Emeline CSST**, Provincial Superior and Manager **Rev. Sr. Dr. Vinitha CSST**, Principal **Dr. Lizzy Mathew**, **Ms. Ann Thomas Kiriyanthan**, Head of the Department of Commerce and to all other faculties of the department for their kind cooperation.

We would like to express our thanks to all respondents and colleagues in developing the project.

We also extend heartfelt thanks to our family for their constant encouragement without which this project would not be possible.

**Aiswarya Babu**

**Akshara Edwin**

**Aleena K.S**

## CONTENTS

SL.NO.	PARTICULARS	PAGE NO.
1	List of tables	-
2	List of figures	-
3	<b>Chapter 1- Introduction</b>	1-6
4	<b>Chapter 2- Review of Literature</b>	7-14
5	<b>Chapter 3-Overview of Banking Industry and Educational Loans Provided by Banks</b>	15-27
6	<b>Chapter 4- Data Analysis and Interpretation</b>	28-48
7	<b>Chapter 5- Summary, Findings, Recommendation and Conclusion</b>	49-52
8	References	-
9	Appendix	-

## LIST OF TABLES

<b>Table No.</b>	<b>TITLE OF TABLES</b>	<b>Page No.</b>
4.1	Demographic data	28
4.2	Type of bank selected	29
4.3	Area of residence	30
4.4	Parental occupation	31
4.5	Annual income of the family	32
4.6	Nature of course taken	33
4.7	Nature of institution undergone	34
4.8	Location of the bank	35
4.9	Source of knowledge about educational loan	36
4.10	Sufficiency of loan taken	37
4.11	Utilization of the loan amount	38

4.12	Adequacy of borrowed loan	39
4.13	Time taken to disburse the loan	40
4.14	Reasons to avail educational loan	41
4.15	Satisfaction level of various aspects regarding educational loan	42
4.16	Nature of security offered	44
4.17	Reasons for approaching the bank for education loan	45
4.18	Problems faced during the pre-loan sanction period	47

## LIST OF FIGURES

<b>Figure No.</b>	<b>TITLE OF FIGURES</b>	<b>Page No.</b>
4.1	Demographic data	28
4.2	Type of bank selected	29
4.3	Area of residence	30
4.4	Parental occupation	31
4.5	Annual income of the family	32
4.6	Nature of course taken	33
4.7	Nature of institution undergone	34
4.8	Location of the bank	35
4.9	Source of knowledge about educational loan	36
4.10	Sufficiency of loan taken	37
4.11	Utilization of the loan amount	38
4.12	Adequacy of borrowed loan	39

4.13	Time taken to disburse the loan	40
4.14	Reasons to avail educational loan	41
4.15	Satisfaction level of various aspects regarding education loan	43
4.16	Nature of security offered	45
4.17	Reasons for approaching the bank for education loan	46
4.18	Problems faced during the pre-loan sanction period	48

**CHAPTER 1**  
**INTRODUCTION**

## **1. 1 INTRODUCTION**

Education is the process of facilitating the learning or acquisition of knowledge, skills, values, morals, beliefs, habits, and self-development. Education was born as a transfer of cultural heritage from one generation to the next. Today, educational goals include more and more new ideas such as learner liberation, critical thinking about the information presented, skills needed for modern society, empathy, and complex professional skills. In today's highly competitive world, the cost of education has increased significantly and continues to rise. This creates problems for both students and parents. When students go for higher education programs at first-class institutions or plan to study abroad, the cost is much higher than parents expect. A student loan is a type of loan designed to help students pay for higher education and related expenses such as tuition, books and supplies, and living expenses. It can differ from other types of loans in that interest rates are quite low and repayment schedules can be delayed while a student is in school. It is also distinguished by strict laws governing amendments and bankruptcy in many countries. Thus the financing of higher education is a matter of great theoretical and empirical debate. The nature, extent and mode of participation of public funds in the education sector involve a long list of arguments put forward by individual scholars and institutions and hence this study.

## **1. 2 SIGNIFICANCE OF THE STUDY**

Education is vital to the human resources and their empowerment in any country. Higher education comes at a higher price. As a result, even though they have merit, it is inaccessible to financially disadvantaged students. In this situation, the role of commercial banks is critical because they provide educational loans to qualified students with financial difficulties. Commercial banks consist of private sector banks, public sector banks and new generation banks. Human capital development is a national responsibility, and it should be everyone's goal to ensure that no deserving student is denied the opportunity to continue higher education due to a lack of financial resources. According to the Finance Ministry, around 18 lakh students all over India are enjoying benefits of educational loans worth ₹ 16000 crores, which have been disbursed to the students who are in need.

Education loan of commercial banks is a key factor and it has a greater significance in the society. This research will be a significant endeavour in promoting a positive atmosphere in the realm of educational financing. The present study is crucial from the point of view of the borrowers in the sense that to understand the needs, difficulties and attitude of the borrowers towards education loan and the benefits of quality education, commercial banks can grant loans for higher education. Education loans should be seen as an investment for economic development and prosperity.

The result of study will also benefit the commercial banks to get an overview regarding the various factors that influence and affect the borrowers while taking education loan thereby they can rectify and improve their services.

### **1. 3 STATEMENT OF THE PROBLEM**

Higher education expenditures are rising in both public and private educational institutions, requiring more money to pay the cost of study and living in colleges and universities which some students cannot afford due to limited financial resources. Student loans are indeed helpful in securing funds for studying within the home country or abroad. However, getting an education loan might not be as easy as you think. While you can easily apply for best student loan scheme available, to have your loan approved can include numerous hurdles. Another important thing to consider is education loan interest rate that is generally very high compared to other types of loan.

Banks collect some form of security against the education loan. This could be bank deposits, post office deposits and various other forms of securities. The list of securities and the amount of margin needed depends from bank to bank. At times a student through parents or guardian cannot raise this money. Perhaps an intervention is needed by authorities and more flexible terms are needed to help students raise money for education.

Students and parents face multiple challenges while planning for their higher education. Some of the issues faced by students include:

- Flexibility in terms of the loans which need to be customized to the requirements of the students and their respective courses.
- Ability to find the required margin money (portion of the fees that is to be paid by students/parents which is not covered by the loan).
- If the required loan amount is more than the value of the collateral security.
- Ability to prove that the University, College and the applied Course are appropriate. (This can be quite challenging if the decision maker is not aware of the thousands of valid and accredited educational choices available in today's fast changing world of education)
- Ability to get the loan approval letter in time to meet the admission and visa related deadlines.
- Ability to get appropriate customer service during the repayment cycle.
- Ability to repay the loan from overseas.
- Ability to get web-based access to the loan account with web enabled transaction processing inclusive of electronic payments, etc.

It's never too easy for bankers as well. In fact, banks have seen non-performing assets rising on account of education loans. They say that especially for students studying abroad, they frequently change their address and it's difficult to trace them for payment. Also, many students get salary that is just about equal to their equated monthly instalment and defaults are high.

#### **1.4 OBJECTIVES**

1. To ascertain the adequacy of educational loans provided to the beneficiaries.
2. To examine the problems faced during the pre-loan sanctioning period.
3. To find out the various sources of awareness about educational loan provided by banks.
4. To explore the level of satisfaction of the borrowers regarding various aspects related to the education loan provided by commercial banks in Ernakulam.
5. To study the various factors that influence the people to avail educational loan.

#### **1.5 SCOPE OF THE STUDY**

This study is confined to the commercial banks in Ernakulam district. The number of students opting for higher education is very high here thus, the need for education loan schemes offered by commercial banks is very important. The purpose of the study is to analyze the adequacy of education loans in Ernakulam district, the limitations of the schemes, to identify various problems faced by the beneficiaries in pre and post loan sanction period and make suggestions to improve their performance. It is conducted among the beneficiaries who have availed education loans and their feedbacks and issues will be recorded. The study will cover the different education loan schemes and the satisfaction level of the borrowers regarding the amount of loan given and time taken for loan disbursement. The current study tries to get an overview about the performance of various educational loans offered by various types of banks in Ernakulam district and to understand the factors considered by the beneficiaries while choosing the education loan from a financial institution. The study also considers the socio-economic profile of the individuals who avail education loans in Ernakulam district. The study also focuses on the perception of the beneficiaries towards the education loans schemes offered by different banks.

## 1.6 METHODOLOGY OF THE STUDY

**UNIVERSE:** Beneficiaries of educational loan of Ernakulam district.

**SAMPLE SIZE:** Sample size is limited to 100 beneficiaries in Ernakulam who took education loan.

**AREA OF STUDY:** Ernakulam.

**RESEARCH DESIGN:** The present study focuses on collecting data from beneficiaries using questionnaire and analysing the data collected in order to ascertain the adequacy of education loan provided to the student beneficiaries.

**COLLECTION OF DATA:** Primary data is used for this study. Primary data is collected by sending questionnaires to the beneficiaries who took education loan. Secondary data is gathered from various websites and researches etc.

**TOOL FOR ANALYSIS:** Suitable statistical tools such as percentages, averages etc are used. Data collected are presented in both tabular and graphical way.

Sources of data	Primary data
Sample size	100 respondents
Area	Residents of Ernakulam
Sampling technique	Convenience sampling
Source of primary data	Structured questionnaire
Tools of analysis	Tabular and graphical presentation are used
Data collected procedure	Primary data is collected by supplying questionnaire

## 1.7 LIMITATIONS OF THE STUDY

1. Time and financial constraints remained a major limitation during the study. The data had to be collected from individuals who have taken education loans within the Ernakulam district in a short period of time.
2. The sample size consists only of 100 respondents.
3. Different banks have different schemes of loans with varied interest rates, so it was difficult to analyse the satisfaction level of the beneficiaries.
4. The details about the beneficiaries were kept confidential by the financial institutions hence it was difficult to collect the data.
5. There is a possibility of sampling errors in the study.

## 1.8 KEYWORDS

- **Education Loan:** It is a sum of money borrowed to finance post-secondary education or higher education-related expenses.
- **Interest Rate:** The proportion of a loan that is charged as interest to the borrower, typically expressed as an annual percentage of loan outstanding.
- **Loan Disbursement:** When the agreed-upon amount is actually paid into borrower's account and is available for use.
- **Beneficiary:** A person who derives advantages from something.

## **1.9 CHAPTERISATION**

### **Chapter 1- Introduction**

This chapter contains a brief introduction regarding the topic, significance, statement of the problem, objectives, methodology, limitations of the study, keywords adapted and chapterisation of the study.

### **Chapter 2- Review of Literature**

This chapter deals with analysing and evaluating the available literature works by researchers concerning with our chosen topic's area.

### **Chapter 3- Overview of Banking Industry and Educational Loans Provided by Banks**

In this chapter it includes introduction and history about banking, types of commercial banks, types of education loans, and interest rates of education loans provided by commercial banks.

### **Chapter 4- Data analysis and interpretation**

This chapter deals with primary data which is collected from the borrowers and analysed using different tables and graphs.

### **Chapter 5- Summary, Findings, Recommendation and Conclusion**

This chapter is the summary of findings, recommendation and it includes conclusion of the study.

**CHAPTER 2**  
**REVIEW OF LITERATURE**

## INTRODUCTION

This chapter documents the important findings, recommendation and conclusions from different thesis, articles and research papers which provided this study with a starting point of what information and studies were documented about educational loan, specifically on certain regions and sections. These studies have been done due to the non-efficiency of educational loan at some point or to understand about it and to familiarize it to the population. The literature review focuses on various aspects of education loan such as significance, adequacy, issues etc. The purpose of this chapter is to establish the background, context and significance of the general field of the study.

**Jandhayala B.G Tilak** (1992) briefly explained the details of the student loan scheme established in India, its strength and weakness. He suggested some marginal improvements that are essential for the better functioning of loans as a means of financing higher education. The main conclusion of the study was that the student loans must be judged more in terms of generating finances for higher education, rather than as a measure to improve access and equity in higher education.

**M.R. Narayana** (2005) made an empirical analysis of the role of student loan by commercial banks in financing the estimated budgetary subsidy to general collegiate education by Government and private aided colleges in Karnataka state. The main objective of this study is to find plausible answers to the policy debate that is (a) on the choice between credit and fiscal instruments to help the student's community in pursuing higher education and (b) in using commercial student loans as an indirect instrument of reduction in budgetary subsidy to higher education in the state. Therefore the results of this study are of special relevance to other states within India as well as in other countries, if their regional governments intend to reduce the budgetary support to higher education through commercial student loan scheme.

**Saumen Chattopadhyay** (2007) critically analysed some of the sources proposed and implemented to raise funds for higher education. The paper focuses on the options that are being carried out by the world over to raise more resources from higher education and feasibility of trying out some of the options in India and discuss the pros and cons of some alternative sources of financing higher education from the point of view of efficiency in allocation of resources and social justice. The study argued that the Government has to play an important role in the higher education system to consolidate India's competitive edge in knowledge generation in a global economy also the need for skilled manpower and to attain social cohesion.

**Hua Shen and Adrian Ziderman (2008)** the study was based on student loans repayment and recovery. They probed two issues, for 44 loans schemes in 39 countries: how much of the original loan is an individual student required to repay (the "repayment ratio") and what percentage of the total costs of loans schemes can the lending body expect to receive back in repayments (the "recovery ratio"). The analysis shows considerable variation in the size of the repayment and recovery ratios across schemes. Moreover, many loans schemes exhibit sizeable built-in subsidies accruing to student borrowers – in over 40 percent of the schemes examined, the repayment ratio is 40 percent or less. Overall loans recovery is considerably lower. Policy implications of these findings are discussed together with a consideration of steps that may be taken to improve the financial outcome of loans schemes.

**Dr.G.Raviselvam and G.Maheswari (2008)** studied the student satisfaction of banking services provided by commercial banks in the city of Thanjavur. It also derived the relationship between service, quality, customer satisfaction and loyalty and also reveals the factors that induce the students to avail loans from commercial banks. The attitude of banks towards the students and the satisfaction level of the students towards the bank was a part of the study. This incorporated the perspective of the student towards funding agency. The study helped both the students and the bankers by providing valuable suggestions to both of them to improve their level of satisfaction. Raviselvam and Maheswari say that students of self-finance college have got the benefits when compared to the students of Government and Aided colleges.

**Dr. Harsha Gandhar (2010)** Attempts to review the growth and performance of education loans in raising the access ratio in higher education during the period 2004-10 through a case study of scheduled commercial banks in Chandigarh (2007) and enlists the deficiencies in the scheme and suggests some policy options in regard. The main conclusion of the study is that the scheme is purely ran commercial basis and does not offer any soft options for the meritorious and the needy. He suggests that the even though the growing magnitude of education loans indicates that the scheme is getting very popular amongst the higher education aspirants in India. But there is strong need to work on the weakness and deficiencies of the scheme.

**Dr. S. Puttaswamaiah** (2010) showed that finances for education are mobilized from different sources like government spending, fees, education loans, and other. Among these, educational loan has been as an alternative way of financing education. In this background the paper analyses the trends and patterns of educational loan in India. The study attempts to understand the pattern of student loans provided by selected commercial banks. The study finds that education loan by public sector is increasing over the years, showing the increasing demand for loans of higher education. Access to higher education should not be limited to few sections of the society which creates inequity.

**Rajesh Tiwari and Bimal Anjum** (2013) studied the role and importance of educational loans for the development of human capital in India. The privatization of higher education and poor budgetary revealed the importance of educational loans for the country. The paper shows the projected trend of population and lower expenditure on education compared to other countries. The study provided some suggestions on enhancing the utility of the educational loans to improve access and employability of students.

**P. Geetha Rani** (2014) made a study to recognize why and how to identify the level of loans need to be converted into grants for addressing the objective of equity and access for indirectly facilitating educational loans to target the affluent and privileged for cost recovery. The secondary data were from various sources such as Reserve Bank of India (RBI), selected educational statistics etc. The primary data from household survey were from the unit's record of the 64<sup>th</sup> round of National Sample Survey Organization (NSSO) on expenditure on education. The article gave detailed explanation of rationale and the mechanisms through which the loans can be converted into grants.

**Sunita Rani** (2014) studied about educational loan in professional higher education, a case study about Punjab. By referring the performances of six regions in educational loan, it's understood that the southern region benefit more than northern region. In that too, public sector participates most than private sector. This extends to both India and Punjab. The percentage of total outstanding accounts and amount advanced to women is 35.94 and 33.12. And it is perfectly made clear that the amount of girls utilizing educational loan is precisely low. The study builds up a strong case for the massive involvement of educational loan and public resources in order to ensure wider access and equity in professional higher education. So people will have to build a social movement on the issue of social concerns.

**Lavanya R (2015)** studies about the problems faced by public sector banks towards retrieval of higher education loan with special references to professional courses in Tamil Nadu. This thesis of bankers survey played a key role which reflected 6 factors such as lack of employment opportunities, marks obtained, low income, emotional burden and natural barriers, Government policies and literacy among parents. A Structural Equation Model (SEM) has been developed for the research study is inter related with each other. Research design was framed based on the study which includes population, sample size, sampling techniques.

**Shinde and Muktaram Baburao (2015)** studies about the significant effect of Educational loan in Maharashtra with special reference to State Bank of India. This study helps to understand the need, importance and role of education loan, to examine the factors driving the growth of Commercial Bank's education loans to make a study of different schemes of promotions of Educations and education loan of Commercial Banks, to study the procedure of education loan of State Bank Of India, to know the impact of State Bank of India's education etc. Challenge in the financing of higher education arises in terms of finding diversified sources of financing and allocating the resources efficiently to meet the objectives of access, equity and quality higher education.

**Panjali N (2015)** studied about the priority sector advances made by commercial banks in UT of Puducherry, especially in regard of educational loan. It is found that the academic course pursued previous academic performance and surety act as major factors of success and failure in securing education. There is a serious perception among the borrowers that the educational loan can be waived off by the politicians, so both the bank and government should take particular steps to change this perception and make them feel comfortable. Regular meetings are to be held in order to bring out proper receiving and giving of loan.

**Arindam Bandyopadhyay (2016)** studied micro-level risk of education loan using large set of historical customer level loan data sample from four major public sector banks in India. The main objective of the study was to examine the performance of loans over time and identify key risk factors of such loans across various geographies and constitutions. The data was collected of performing accounts along with default accounts to compare their behaviours and identify key risk factors. Empirical methodology has been adopted in this study and the results was that the primary objective of the multivariate analysis is to demonstrate the importance of borrower specific characteristics as well as regional and institutional factors in deterring the credit risk of education loans in India.

**Minnu John** (2016) the study is conceptual one which focuses on the Model Education Scheme introduced by Indian Banker's Association. The study shows the present scenario of education loans and the position of the non-performing assets. The study proved the main challenges faced by the students as well as banks; they include customization of loan product to suit the requirements, collateral security sufficient enough to cover the loan amount, appropriateness of the university, timely approval of loan to meet the admission related deadlines. The problems faced by the banks is also taken in the study, the problem of default is worse, in the case of overseas education as the students change their address without informing the bank. But because of political pressures, public sector banks cannot simply stop giving loans or stop building their non-performing assets.

**Dr. Tania Gupta** (2018) studied the accessibility of educational loans in India for financing higher education by percentage analysis. The major finding of the study is that an education loan in India is not a popular cost-sharing device to finance higher education. Only 0.2 percent students out of the total sample had availed an education loan. The project showed that the educational loans should be prompted by the Government of India and the loans should be made available to the needy at lower rates with an extension in repayment period.

**Dr. Manisha** (2018) studied the education loans in Punjab and examined the socio-economic profile of borrowers. The study also extended to the areas such as institutions and their terms and conditions for providing education loans and the status of repayment. The study found out that our system has severely starved of funds in the education sector. The growth of education loans in India is rising as the number of accounts of education loans has been increasing continuously since 2005-06 as the number of accounts of education loans has been increasing. All type of banks public, private, co-operative and gramin banks provide education loan. Maximum number of accounts for education loans is opened by public banks. Major chunk of amount is provided by public banks and outstanding amount is also the highest with these banks. State bank of India has the highest rank in amount outstanding followed by the Canara bank, Punjab National bank, Indian Overseas bank and central bank of India. A very small amount of loans has been provided by the private sector banks mainly, due to higher rate of interest charged by the private sector banks.

**Santhosh Kumar Bojan and A. Saravanakumar (2018)** studied about the education loan features of Canara Bank in Coimbatore. They explained the satisfaction level of providing education loans and the actions taken by Canara Bank towards pending payment dues. The study suggested that as the new procedures are going to be implement for loan sanctioning section in nearby future and the announcement are spreading through the Medias. People are advised to update themselves and get benefit by those schemes. The survey has been reviewed with simple percentage analysis tool and the data interfered has been suggested to the bankers as well as to the government for better solution to students and people are requested to update the upcoming procedures in loan processing.

**Dr. Thulasi Priya and C. Esakkiammal (2018)** studied the bank-wise performance of education loan lent by public sector banks in India. The bank-wise distribution of education loan was analysed by collecting secondary data for a period of five years (2013-2017) from the various annual reports of ministry of finance and various banks. The percentage and rank analysis of setting out of bank wise education loan had resulted that, State Bank of India had contributed more in the disbursement of education loan in the nation. Though SBI, Canara bank, IOB, Indian bank and PNB were the major banks in lending education finance, steps would be taken by RBI along with the government to distribute the education loan scheme by all the public sector banks evenly so that maximum number of students in India would be benefited by the education loan scheme. The study was based on the publications of the Reserve Bank of India, The Indian Banks Association, The University Grants Commission, The Federation of Indian Chamber of Commerce and Industry, Annual Reports of Ministry of Human Resource Development, Ministry of Finance and annual reports of banks.

**S. Uma Priyadharshini and M. Padmaja (2019)** made a study to analyse and evaluate the effectiveness of education loan in empowering rural area students and to identify the level of awareness among the students regarding the procedures of availing educational loans. From the research it is discovered that majority of students from technical background have availed loan and have come from low income groups families. Education loan has not contributed positively towards the increase in socio-economic status of the students even though it is found that education loan is effective in empowering the students socially and psychologically. The study concluded that education loan system will become more effective if the loophole in the system is rectified and will contribute towards the empowerment of rural area students.

**Sandeep M. Khanwalker** (2019) analysed reasons of high default and suggests some steps for Banks and Financial Institutions to obviate slippages. The objectives of the study is to analyse credit low to education loan segment and quantum of irregular or non-performing assets under this category and understanding reasons for education loans accounts turning non-performing assets. The study suggested several measures for bank, student and policy specific.

**Shinta Sebastian** (2019) studied about the educational loan scheme of major commercial banks of selected districts of Kerala. This is a detailed thesis from the perspective of both beneficiaries and bank managers. According to this thesis, borrowers from the urban areas won't bear any difficulties in the procedures of loan disbursement and etc. And the public sector banks own the major share in sanction and disbursement of educational loan in India. It also came to light that the increase in the education loan has increased the non-performing assets of bank considerably. Due to the increased knowledge of the majority of the urban people, they are more aware of educational loan than the rural area. This doesn't differ with the intensity of districts. From the perspective of bank managers, the frequent visit of loaners is the major hindrance faced. But still the urban bank managers are more satisfied than the rural and semi urban bank managers.

**Sajith Kumar** (2019) studied about the socio economic implications of educational loan in Kerala. Educational loan helps all the highly potential students to achieve their dreams. Pursuing a course without educational loan is unimaginable to the poor. But in times, non-ambitious students may result in default payment. This increase the possibility for so many students to achieve loan, so possible steps must be taken to overcome this, as a lot of deserving students are denied of this right. Repayment is a major factor in it. This study reveals inefficiency in the utilization of the scheme and as a result the payback from the educational loan is inefficient. The quality of the course and the institution is a major factor in deciding the future career prospects of the student. So students must be so careful in choosing this. The present outlook of educational loan is as crucial as it is pictured as impossible. Effective measures to be taken to create a positive attitude towards educational loan.

**Sivakumar N** (2019) studies about the student's perception on Educational loan with special reference to Coimbatore district. This thesis based on student education loans procured through banking services, including the function of various banks in providing educational loan services as well as their impact on India's educational system. It is notable that the poor students are dependent on educational loan for higher education. So the bank and government is supposed to impose more and more variety of educational loan schemes to motivate studious students. Both private and public sector gives educational loan and they both shows hindrances like inaccessibility of bank officials, frequent bank visits, lack of guidance and delay in sanctioning and disbursement of loan. Various measures must be put forward for the efficacy of students.

**Esakkiammal C (2019)** studies the performance evaluation of public sector banks educational loan and borrower's perception. Further, this thesis throw the light on the need for the reforms to be taken by the Government in the policy of education loan scheme, which has long term perspective plan like lower interest rate, adequate moratorium period, loan with minimum collateral security and loan to economically weaker section at concessional rates. In order to protect the interest of the borrowers, a specialized committee can be formed by the Government along with members of Indian Banks Association (IBA) to monitor the disbursements of education loan both at national and State level. It also discovers that the entry of Non-Banking Financial Company (NBFC) have slowed down the growth of educational loan. It is because, NBFC attracts more borrowers towards them by express services and the tranquil procedures in advancing the loan, thereby bringing out setback in the education loan considerably for the past three years.

### CONCLUSION

Under the light of these studies, we concluded that a lot of common problems have been around the topic educational loan. We all understood that the private sector bank should is relatively low active in this area due to its higher rates. And also its high time since government has to increase the expenditure on higher education .The default in repayment of loans has led to the increase in non-performing assets. This is due to the carelessness and non-ambitious performance of the students. This reluctance is also due to the unsuccessfulness of certain students in finding particular ways to repay. And this hindrance could only be rectified with borrowers own contribution, this can only be done by rich population. Due the lengthy procedure and the need for security to be to be pledged to the bank to get the loan the poor still cannot benefit from the education loan schemes. And also it's very notable that the urban and semi urban utilizes educational loan, rural is comparatively less active due to less education and low infrastructure.

## **CHAPTER 3**

# **OVERVIEW OF BANKING INDUSTRY AND EDUCATIONAL LOANS PROVIDED BY BANKS**

## INTRODUCTION

Modern banking in India originated in the mid of 18th century. Among the first banks were the Bank of Hindustan, which was established in 1770 and liquidated in 1829–32; and the General Bank of India, established in 1786 but failed in 1791. The largest and the oldest bank which is still in existence is the State Bank of India (SBI). It originated and started working as the Bank of Calcutta in mid-June 1806. In 1809, it was renamed as the Bank of Bengal. This was one of the three banks founded by a presidency Government; the other two were the Bank of Bombay in 1840 and the Bank of Madras in 1843. The three banks were merged in 1921 to form the Imperial Bank of India, which upon India's independence, became the State Bank of India in 1955. For many years, the presidency banks had acted as quasi-central banks, as did their successors, until the Reserve Bank of India was established in 1935, under the Reserve Bank of India Act, 1934.

Banking in India originated in the mid of 18th century. Among the first banks were the Bank of Hindustan, which was established in 1770 and liquidated in 1829–32; and the General Bank of India, established in 1786 but failed in 1791. The largest and the oldest bank which is still in existence is the State Bank of India (SBI). It originated and started working as the Bank of Calcutta in mid-June 1806. In 1809, it was renamed as the Bank of Bengal. This was one of the three banks founded by a presidency Government; the other two were the Bank of Bombay in 1840 and the Bank of Madras in 1843. The three banks were merged in 1921 to form the Imperial Bank of India, which upon India's independence, became the State Bank of India in 1955.

The Indian banking sector is broadly classified into scheduled and non-scheduled banks. The scheduled banks are those included under the 2nd Schedule of the Reserve Bank of India Act, 1934. The scheduled banks are further classified into: nationalized banks; State Bank of India and its associates; Regional Rural Banks (RRBs); foreign banks; and other Indian private sector banks. The SBI has merged its Associate banks into itself to create the largest Bank in India on 1 April 2017. With this merger SBI has a global ranking of 236 on Fortune 500 index. The term commercial bank refers to both scheduled and non-scheduled commercial banks regulated under the Banking Regulation Act, 1949.

## HISTORY

According to the Central Banking Enquiry Committee (1931), money lending activity in India could be traced back to the Vedic period, i.e., 2000 to 1400 BC. The existence of professional banking in India could be traced to the 500 BC. Kautilya's Arthashastra, dating back to 400 BC contained references to creditors, lenders and lending rates. Banking was fairly varied to the credit needs for the trade, commerce, agriculture as well as individuals in the economy, Mr. W.E. Preston, member, Royal Commission on India Currency and finance set up in 1926, observed "it may be accepted that a system of banking that was extremely suited to India's then requirements was in force in that country many countries before the science of banking become an accomplished fact in England". They had their own inland bills of exchange or Hundis which were the major instruments of transactions. The dishonouring of hundies was a rare at that time as most banking worked on mutual trust, confidence and without securities.

The first western bank of a joint stock variety was Bank of Bombay, established 1720 in Bombay. This was followed by bank of Hindustan in Calcutta, which was established in 1770 by an agency house. This agency house and banks were close down in 1932. The first "Presidency Bank" was the Bank of Bengal established in Calcutta on June 2, 1806 with a capital of Rs.50 Lakh. The Government subscribed to 20 percent of its share capital and shared the privilege of appointing directors with voting rights. The bank had the task to discounting the treasury bills to provide accumulation to the Government. The bank was given powers to issue notes in 1823. The Bank of Bombay was the second presidency bank set up in 1840 with a capital of Rs. 52 Lakh, and the Bank of Madras the third Presidency bank established in July 1843 with a capital of Rs. 30 Lakh. The presidency banks were governed by Royal charters. The presidency banks issued currency notes until the passing of the paper currency Act, 1861, when this right to issue currency notes by the presidency banks was taken over and that function was given to the Government. The presidency bank act, which came into existence in 1876, brought the three presidency banks under a common statute and imposed some restrictions on their business. It prohibited them from dealing with risky business of foreign bills and borrowing abroad for lending more than 6 months.

The Imperial Bank of India also functioned as a central bank prior to the establishment of the Reserve Bank in 1935. Thus, during this phase, the Imperial Bank of India performed three set of functions via commercial banking, central banking and the banker to the government. The first Indian owned bank was the Allahabad Bank set up in Allahabad in 1865, the second, Punjab National Bank was set up in 1895 in Lahore, and the third, Bank of India was set up in 1906 in Mumbai. All these banks were founded under private ownership. The Swadeshi Movement of 1906 provided a great momentum to joint stock banks of Indian ownership and many more Indian commercial banks such as Central Bank of India, Bank of Baroda, Canara Bank, Indian Bank, and Bank of Mysore were established between 1906 and 1913.

The World War I (1913-1918) has affected badly the Indian economy and created many problems like high Inflation, low productive of agriculture sector. During the war period, a large number of banks failed. Some banks that failed were also doing trading function with banking function. Most of the banks that failed during war period had low capital base.

At the time of Independence, the Imperial Bank of India and all other commercial banks were urban oriented. Therefore it is the need of the hour, to provide the banking facility to the rural area. It was suggested that the Imperial Bank of India should extent its branches to Taluka or Tehsil to provide the banking services for the neglected area. Imperial Bank of India was taken over by the Government under the State Bank of India, Act, 1955, effective from July 1, 1955. Under the State Bank of India (Subsidiary Banks) Act, 1959, eight state owned/sponsored banks were taken over by State Bank of India as its subsidiaries, now called Associate Banks. With amalgamation of two of them (State Bank of Bikaner and Jaipur), the number of these associate banks has come down to seven. At present, state bank group consists of six banks.

In 1974 the Narasimham Committee examined the financial system and recommended the establishment of Regional Rural Banks (RRB) under the "Regional Rural Banks Act, 1975". Banking in collaboration with central and State Governments, set up Regional Rural Banks in selected regions where the cooperative system was weak and where commercial banks were not very active. On April 15, 1980 six more private sector banks were nationalized, making the number of public sector banks 27. One of the most important policy initiative of this phase was the acceptance and implementation of many recommendations of far reaching implications for the financial sector, made by the Narsimham Committee simultaneously, for strengthening the securities market, Securities and Exchange Board of India was made a statutory body and given sufficient power to deal with various fraudulent practices and scams effectively. The Government of India accepted all major recommendations of Narsimham Reports and started implementing them straightway, despite stiff opposition from banks unions and political parties in the country. It is

primarily because of the financial sector reforms initiated during the last two decades or so that the Indian financial system is acquiring fast the shades of a vibrant, dynamic, globalised, complex system today, creating new opportunities and challenges.

### **TYPES OF COMMERCIAL BANKS**

Commercial banks are the maximum vital additives of the entire banking system. A business financial institution is a profit-primarily based totally economic organization that presents loans, accepts deposits, and gives different economic services, along with overdraft centres and digital switch of funds.

In different words, business banks are economic establishments that take delivery of call for deposits from the overall public, switch price range from the financial institution to another, and earn profit. Commercial banks play an extensive position in pleasing the brief-time period and medium- time period economic necessities of industries. They do now no longer offer, long-time period credit score, in order that liquidity of belongings must be maintained. The price range of business banks belong to the overall public and are withdrawn at a brief notice; therefore, business banks prefers to offer credit score for a brief time period subsidized via way of means of tangible and effortlessly marketable securities. Commercial banks, whilst presenting loans to businesses, keep in mind diverse factors, which include nature and length of business, economic popularity and profitability of the business, and its capacity to pay off loans.

Commercial banks are of 3 types, which might be as follows:

**(a) Public Sector Banks:**

The Central Government entered the banking commercial enterprise with the nationalization of the Imperial Bank of India in 1955. A 60% stake become taken via way of means of the Reserve Bank of India and the brand new financial institution become named State Bank of India. The seven different kingdom banks have become subsidiaries of the brand new financial institution in 1959 whilst the State Bank of India (Subsidiary Banks) Act, 1959 becomes surpassed via way of means of the Union authorities. The subsequent foremost authorities intervention in banking passed off on 19 July 1969 whilst the Indira authorities nationalized a further 14 foremost banks. The overall deposits within side the banks nationalized in 1969 amounted to 50 crores. This flow expanded the presence of nationalized banks in India, with 84% of the overall branches coming beneath authority's control. Public Sector Banks (PSBs) are a prime form of authorities owned banks in India, in which a majority stake (i.e. extra than 50%) is held through the Ministry of Finance of the Government of India or State Ministry of Finance of diverse State Governments of India. The officials running for those entities and their subsidiaries are gazette officials. The

personnel subordinate to the officials running for those respective entities and their subsidiaries also are full-fledged authority's personnel. The stocks of those banks are indexed on inventory exchanges. Their predominant goal is social welfare. This is a sort of business banks which are nationalized through authorities of a rustic. In public region banks, the essential stake is held through the authorities. In India, public region banks function below the pointers of Reserve Bank of India (RBI), which is the relevant bank. Some of the Indian public region banks are State Bank of India (SBI), Corporation Bank, Bank of Baroda, Dena Bank, and Punjab National Bank.

**(b) Private Sector Banks:**

Bank is an economic middleman that accepts deposits and channels the ones deposits into lending activities, both at once via way of means of loaning or in a roundabout way via capital markets. Clients with capital shortfalls and clients with capital surpluses are linked together by a financial institution. Modern banking practice, inclusive of fractional reserve banking and the difficulty of banknotes, emerged within side the seventeenth and 18th centuries. The Bank of England changed into the primary to start the everlasting problem of banknotes in 1695. Banks offer specific fee services, and a financial institution account is taken into consideration crucial via way of means of maximum corporations and individuals.

Banks can create new cash after they make a loan. New loans for the duration of the banking gadget generate new deposits someplace else within side the gadget. Bank can generate sales in quite a few extraordinary methods such as interest, transaction costs and economic advice. The worldwide Private Banking enterprise can appearance returned on 5 very wealthy years, having capitalized on each unheard of increase in Assets below Management and beneficial commercial enterprise economics, consisting of strong sales streams and occasional capital requirements. Despite the current marketplace turmoil, it agree with that wealth swimming pools will hold to increase throughout all regions, specifically in new markets consisting of Asia Pacific, the Middle East, and Latin America. The splendid increase skilled for the duration of the beyond 5 years isn't always sustainable. In the present day worldwide order, in which the arena has end up a large village, clients take a worldwide examine the goods and offerings in phrases of price, quality, transport and after-sale offerings. This fashion has sown the seeds of opposition in each area of financial system and banking area isn't any exception to this event. Banking, the arena over, has been converting at a superb pace. This alternate is because of multifarious elements just like the want to be green in functions, thirst for turning into finance superpowers than mere banks, developing significance of personal banking, the upward push in excessive internet really well worth individuals, etc. the last decade of 90s has witnessed a sea alternate within side the manner banking is executed in India. Private Banks are banks owned through either a character or a well-known partner(s) with restricted partner(s). Private Banks aren't incorporated. In this sort of case, the lenders can appearance to each

the “entirety of the bank’s property” in addition to the whole lot of the sole- proprietor’s/well known-partners’ property. Private Banks can consult with non-authorities owned banks in well known, in comparison to authorities-owned banks, which had been commonplace in communist, socialist and a few social democratic states with inside the twentieth century. Private Banks in India is the banks which just like the public area banks; do now no longer have any authority’s stake. The Indian personal banks can be indexed publicly. Those may be traded on inventory exchanges as well. Private area banks in India keep 18.2 percentage of the whole property of Indian banking enterprise. This refers to a type of business banks wherein essential a part of percentage capital is held through personal agencies and individuals. These banks are registered as businesses with restricted liability. Some of the Indian personal region banks are Vysya Bank, Industrial Credit and Investment Corporation of India (ICICI) Bank, and Housing Development Finance Corporation (HDFC) Bank.

**(c) Foreign Banks:**

The Indian authorities’ securities markets had been widely insulated from the worldwide monetary disaster. There has been no prevalence of agreement failure or default. The muted effect of the worldwide disaster at the Indian authorities’ securities markets may be attributed, inter alia, to the calibrated establishing of the markets to overseas players. Following the intensification of the worldwide monetary disaster in September 2008, the Reserve Bank applied each traditional and unconventional coverage measures a good way to proactively mitigate the unfavourable effect of the worldwide monetary disaster at the Indian economy.

### **TYPES OF EDUCATION LOAN**

Until a few years ago, higher education and quality education was not affordable to some illustrious students because of the financial constraints. There was no alternative but to jump in the job market prematurely which led to untimely end of budding talents and their forceful transformation to the mediocrity. Of course, there were scholarships, but those were so less in numbers that only luckier few could avail them. But now the scene has changed drastically.

The boom in the banking sector has led to release of large amount of funds for education loans. Now, education loans are easily available from various banks in India and this change is encouraging more and more students to take up higher education despite their financial shortcomings. Many nationalized as well as private banks have come up with various educational loan schemes that students can benefit from.

The wave of change could be well gauged from the amusing situation that immediately after the results announcements of CAT or PMT/ IIT JEE, the representatives of the banks queue up for giving education loans to the successful candidate even with very flexible conditions. This scenario is certainly helping the illustrious students to pursue higher education and realize their dreams.

Education loans are basically a form of monetary assistance availed by students to meet the expenses associated with their studies. Education loan can be taken by means of funding, scholarships, financing and rewards, and are granted in cash, which has to be repaid to the lender along with a rate of interest. Students who wish to avail education loans are advised to borrow based on their needs as the repayment periods for these loans can vary to a great extent depending upon the lender and the amount borrowed by the student. Education loans can be classified into different categories based on:

1. Location
2. Course of education
3. Security of collateral or guarantee

#### **ON THE BASIS OF LOCATION**

Education loan can be classified into two categories based on the location

- **Domestic Education Loan**

It's for educational courses within the geographical boundaries of the country. Students who would like to pursue education in India can apply for this loan type. The loan will get approved only if the applicant is admitted to an Indian Educational Institution and meet all other lender criteria. Various banks provide domestic education loans some of the examples are:

HDFC Bank – Up to 30 lakhs

State Bank of India – Up to 15 lakhs

Punjab and Sindh Bank – Up to 15 lakhs

IDBI Bank – Up to 20 lakhs

- **Overseas Education Loan**

Overseas education is for educational courses outside the geographical boundaries of India. Such loans help students pursue their course of their desire in a foreign institution. The loan covers the airfare, accommodation, and tuition fee for students who wish to study abroad. Some of the banks offering overseas education loans are

HDFC Bank – Up to 45 lakhs

State Bank of India – Up to 1.5 crore

Punjab and Sind Bank Up to 20 lakhs

IDBI Bank – Up to 20 lakhs

### **ON THE BASIS OF COURSE OF EDUCATION**

Education loans can be classified into three categories based on the course of education:

- **Undergraduate Education Loans**

A bachelor's degree (also known as a first degree or simply a degree) is a slang term for a degree awarded to a person with a bachelor's degree. In the United States, it is usually offered by higher education institutions such as universities. The most common types of these bachelor's degrees are associate degrees and bachelor's degrees. A bachelor's degree usually lasts at least 3-4 years. In some other education systems, basic education is post-secondary education up to a master's degree. This applies to some science courses in the UK and some long-term medicine courses in Europe. These degrees can be categorized as an undergraduate or first professional degree. Undergraduate student loans can be availed by individuals who have completed secondary education and wish to undertake a three or four yearlong undergraduate degree course to enhance their job opportunities.

- **Graduate loans**

A graduate school is a student who has a bachelor's degree from a university and is advancing a high level of graduate research. A graduate degree usually lasts one year full-time or two years part-time. A graduate diploma, a graduate degree, or a master's degree provides students with a high level of education in the field of study of their choice. The content of the course is designed based on the knowledge acquired in the basic course. Often, the content of the course is inherently more practical in order to better equip students with the skills needed in the workplace. A graduate degree allows students to broaden their career prospects and gain a better understanding of their field of study. Many Graduates would like to continue their education with a post graduate course, usually a 2- year long course in India. It is to get more profound knowledge in the area of interest.

- **Career Education loans**

Many professionals who work for a few years in a corporate job prefer to pause their career and take up professional courses and training to improve their employment prospectus. Such individuals would strive hard to get into reputed business and technical schools. Career education loans are availed by these individuals.

## **ON THE BASIS OF COLLATERAL**

Education loan can be classified into two categories based on the collateral

- **Loan Against Property, Deposits, and Securities**

One can pledge immovable assets such as agricultural land, residential land, flat, house, and others, fixed deposit certificates, recurring deposits, gold deposits, bonds, debentures, and equity shares to get the necessary financing to pursue education.

- **Third-Party Guarantee**

The credit score of the guarantor should be good and he/she is also expected to have sufficient income to cover loan repayment. The age of the guarantor must be above 18 years and he/she should also be a citizen of India. A guarantor is basically a close relative/acquaintance of the borrower. He/she will put some of the assets as collateral and if in case the borrower is not able to pay back the money during the repayment period, he/she owes money to the bank. If the guarantor is also unable to repay the money, then the collateral provided by him will be used to cover up all the remaining costs.

- **Unsecured Education loans**

Unsecured loans are loans provided with no collateral as security. The 2 factors that are considered when providing an education loans without collateral are:

- The co-borrower's and borrower's future income.
- Creditworthiness of the borrower and co-borrower.

There is a processing fee levied for availing education loan without collateral. Depending upon the lender, the processing fee will vary. The interest rate also will be higher for education loans without collateral.

HDFC Bank – For loans up to Rs. 7.5 lakhs no collateral is required.

HSBC – For loans up to 5 lakhs no collateral is required.

State Bank of India – For loans up to 7.5 lakhs no collateral is required.

IDBI Bank – For loans up to 4 lakhs no collateral is required.

## **INTEREST RATES OF EDUCATIONAL LOAN PROVIDED BY COMMERCIAL BANKS**

### **SBI**

The State Bank of India, popularly called SBI, is India's biggest business financial institution with an excellent records of greater than two hundred years. State Bank of India(SBI), Owned through The Government of India, is classified as an Indian Multinational, Public zone banking and Financial offerings corporation, with its headquarters placed in Mumbai, Maharashtra. With greater than 14,000 branches in India, SBI is the most important and one of the top rate banking and economic offerings corporation in India through assets, deposits, profits, branches, customers, and employees. SBI has additionally hooked up and secured its roots globally with 191 overseas places of work unfold throughout 3 countries. The interest rates will be floating for entire period of loan and the effective interest rate is 8.65%.

### **HDFC**

HDFC Bank Limited is an Indian banking and monetary offerings' corporation founded with inside the town of Mumbai, India. It is India's biggest non-public quarter financial institution through property and world's tenth-biggest financial institution through marketplace capitalization as of April 2021. It is the 1/3 biggest corporation through marketplace capitalization of \$122.50 billion at the Indian inventory exchanges. It is likewise the 15<sup>th</sup> biggest corporation in India with almost 120,000 employees. HDFC Bank changed into included in 1994 as a subsidiary of the Housing Development Finance Corporation, with its registered workplace in Mumbai, Maharashtra, India. Its first company workplace and a full-carrier department at San doz House, World have been inaugurated through the then Union Finance Minister, Manmohan Singh. The average interest rate approved is 9% - 13.86%.

## **AXIS**

Axis Bank Limited, previously called UTI Bank, is an Indian banking and economic offerings enterprise founded in Mumbai, Maharashtra. It sells economic offerings to huge and mid-length companies, SMEs and retail businesses. The financial institution turned into based on three December 1993 as UTI Bank, starting its registered workplace in Ahmedabad and a company workplace in Mumbai. The financial institution turned into promoted mutually through the Administrator of the Unit Trust of India (UTI), Life Insurance Corporation of India (LIC), General Insurance Corporation, National Insurance Company, The New India Assurance Company, The Oriental Insurance Corporation and United India Insurance Company. The first department turned into inaugurated on 2 April 1994 in Ahmedabad through Manmohan Singh, then finance minister of India. Its effective rate of interest on educational loan is from 13.70% -15.20%.

## **BANK OF BARODA**

Bank of Baroda (BoB) is an Indian nationalized banking and economic offerings enterprise based within side the metropolis of Vadodara, India. It is the fourth biggest nationalised financial institution in India, with 132 million customers, a complete commercial enterprise of US\$218 billion, and an international presence of one hundred distant places offices. Based on 2019 data, its miles ranked 1145 on Forbes Global 2000 list. The Maharaja of Baroda, Sayajirao Gaekwad III, based the financial institution on 20 July 1908 within side the Princely State of Baroda, in Gujarat. The authorities of India nationalized the financial institution, at the side of thirteen different foremost business banks of India on 19 July 1969 and distinctive as a profit-making public quarter undertaking (PSU). Its effective interest rates start from 6.90% - 10.85%.

## **FEDERAL BANK**

Federal Bank Limited is an Indian non-public region financial institution centered in Aluva, Kochi. The financial institution has 1,272 branches unfold throughout one-of-a-kind states in India. It additionally has consultant workplaces overseas at Abu Dhabi, Qatar, Kuwait, Oman, and Dubai. With a purchaser base of over 10 million, consisting of 1.5 million NRI clients and a huge community of remittance companions across the world, Federal Bank dealt with extra than 15% of India's overall inward remittances of seventy nine billion in 2018. The financial institution has remittance preparations with extra than one hundred ten Banks/Exchange Companies across the world. The financial institution is likewise indexed within side the BSE, NSE and London Stock Exchange and has a department in India's first International Financial Services Centre (IFSC) at GIFT City. Its effective rate of interest starts from 11.50% onwards.

## **ICICI BANK**

ICICI Bank was established in 1994 by the Indian financial institution Industrial Credit and Investment Corporation of India (ICICI) as a wholly owned subsidiary of Vadodara, but its parent company was a joint venture of World Bank and India Public Sector in 1955. It was established as Banking and public sector insurance company sector providing project finance to Indian industry. This bank was established as the Industrial Credit Investment Corporation of the Bank of India before changing its name to ICICI Bank. The parent company later merged with the bank. The Indian Industrial Credit Investment Corporation (ICICI) was founded on January 5, 1955, with Sir Alcott Lamasami Mudaria as the first chairman of ICICI Ltd. ICICI Bank Limited is an Indian multinational banking and financial services company headquartered in Vadodara, India. They provide a wide range of banking products and financial services to corporate and individual customers through a variety of distribution channels and specialized subsidiaries of investment banks, life insurance, non-life insurance, venture capital and wealth management. The bank has 5,275 branches and a network of 15,589 ATMs throughout India and is located in 17 countries. The interest rate for education loan is 10.50.

## **CANARA BANK**

The philanthropist Ammembal Subba Rao Pai founded the Canara Hindu Permanent Fund in Mangaluru, India on July 1, 1906. The bank changed its name to Canara Bank Limited in 1910. It was founded in September 1944 and had three branches when it was taken over on May 20, 1961. The second bank to acquire Canara Bank was Seasia Midland Bank (Alleppey), which was founded on July 26, 1930 and had seven branches at the time of the acquisition. Canara Bank offers a Unified Payment Interface (UPI) app called empower. This app allows customers at Canara Bank and other banks to complete payment and collection transactions using a single mobile app. On November 19, 2017, we released the Canarites (Candi) app, a digital library, a mobile app for field restoration, a retail loan (vehicle) tracking system, and a regulatory orientation tracking system. The bank charges 6.9 percent on their education loans.

## **UCO BANK**

UCO Bank, formerly United Commercial Bank, established in 1943 in Kolkata, is one of the nationalized banks in India. UCO Bank is an Indian commercial bank and government-affiliated company. The bank offers its clients a variety of value-added banking solutions, including NRI's lending schemes, deposit schemes and international banking services for value-added electronic banking solutions. They also own various branches licensed to collect taxes directly in India. The bank has 34 regional offices throughout India. UCO Bank is headquartered in Kolkata. The bank

has 34 regional offices throughout India. UCO Bank offers a wide range of educational loans for research in India and abroad. Loans have affordable interest rates that make higher education available to everyone. Due to the long tenure and moratorium period, repayment is convenient for students who have once found a job. There is also a special education loan program for medical and skill development courses. Their interest rate for education loan is 7.3 percent.

### **IDBI BANK**

IDBI Bank Limited (IDBI Bank or IDBI) is a private bank in India and a subsidiary of Life Insurance Corporation (LIC), providing financial and banking services. The Indian Industrial Development Bank was established in 1964 as a development finance institution that provides financial services to the industrial sector. In 2005, the institution merged with its trading arm, IDBI Bank, to become today's banking company and fall into the category of Other Public Sector Banks. A committee established by the RBI recommends that the Development Banks (IDBI) diversify its activities and reconcile the roles of development finance and banking activities by moving away from the traditional distinction between commercial and development banks. IDBI has changed its role from a development finance institution to a commercial institution in response to reforms in the financial sector. Under the Industrial Development Bank (Business Transfer and Abolition) Act of 2003, IDBI has acquired the status of a limited liability company. IDBI Bank Education Loan aims to provide financial support to students who deserve to pursue higher education in India and abroad. IDBI Bank has a variety of services to choose from and has easy repayment options, so you can get full financial support. IDBI offers education loans at 6.90% rate of interest.

### **PUNJAB NATIONAL BANK**

Punjab National Bank, abbreviated as PNB, is the state-owned bank of India. Headquartered in New Delhi, Delhi, India, it is owned by the Ministry of Finance of the Government of India. Founded in May 1894, the bank is India's second largest state-owned bank in terms of both business volume and network. Punjab National Bank offers a variety of educational loans to students. Loans have attractive interest rates and a term of up to 15 years. However, the repayment period depends on the loan. Higher education (India or abroad), vocational education and training, education at prestigious laboratories, and higher education in Delhi are all covered by an education loan. This loan is also available to people who were born outside India but are pursuing higher education in India. The bank offers provides education loan at an average rate of interest of 6.90 percent.

**CHAPTER 4**  
**DATA ANALYSIS AND INTERPRETATION**

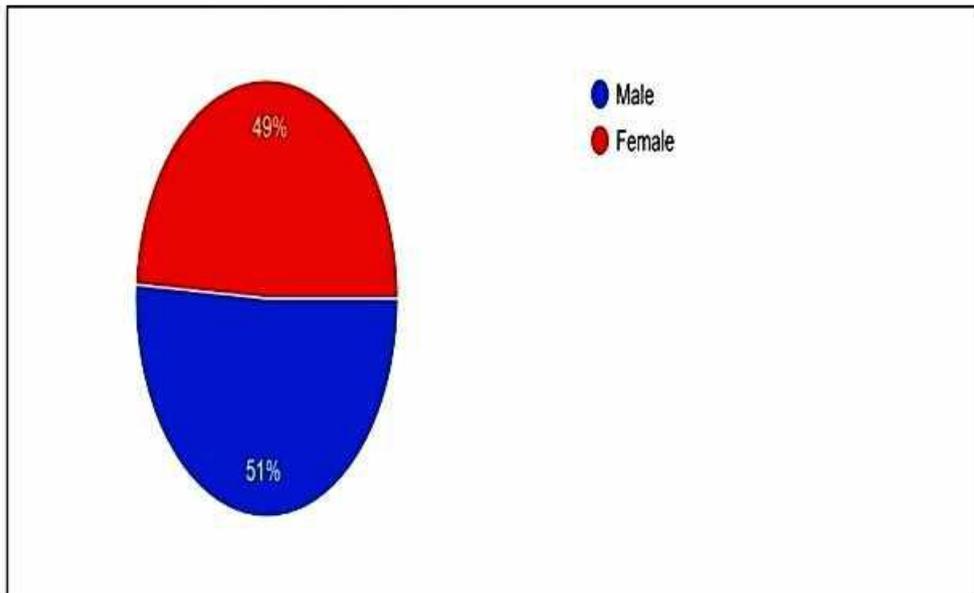
## DATA INTERPRETATION

**Table 4.1 DEMOGRAPHIC DATA:**

<b>GENDER</b>	<b>FREQUENCY</b>	<b>PERCENTAGE</b>
Female	49	49%
Male	51	51%
Others	0	0
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.1**



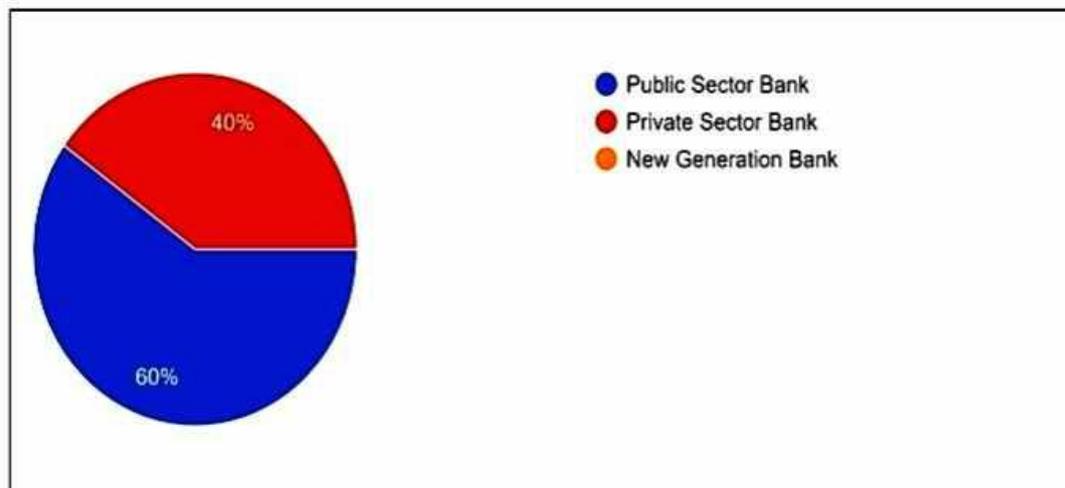
**Interpretation:** Among the total responses, male responses surpassed female responses. Male responses accounted for 51%, while female responses accounted for 49%. The others category was NIL.

**Table 4.2 TYPE OF BANK SELECTED:**

<b>BANKS</b>	<b>FREQUENCY</b>	<b>PERCENTAGE</b>
Public sector	60	60%
Private sector	40	40%
New Generation	0	0
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.2**



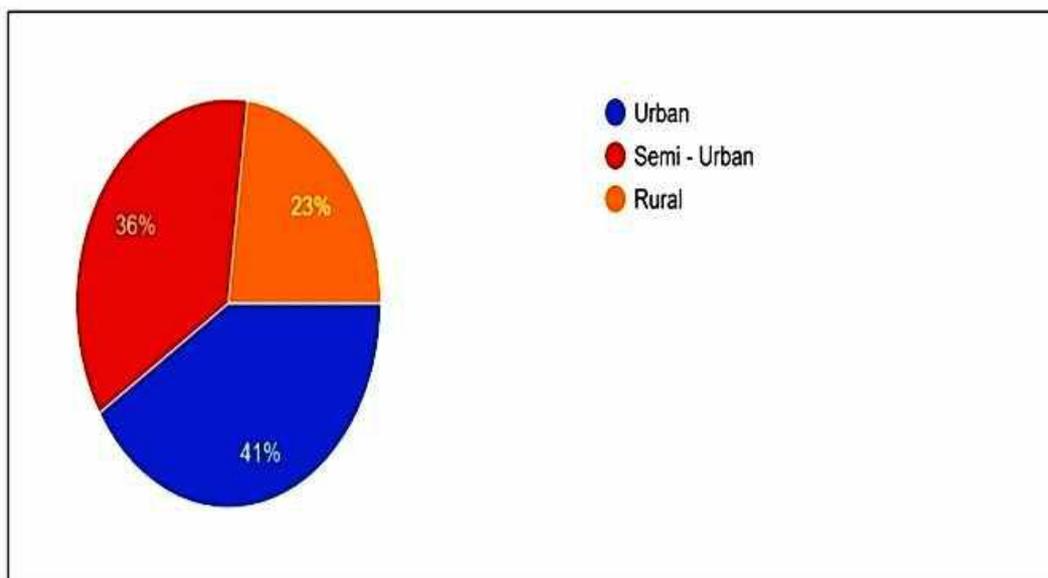
**Interpretation:** Majority of the users prefer public sector bank. 60% of the respondents have taken educational loan from public sector banks. Only 40% of the respondents chose private sector bank. Responses for new generation bank was NIL.

**Table 4.3 AREA OF RESIDENCE:**

RESIDENCE	FREQUENCY	PERCENTAGE
Urban	41	41%
Semi-Urban	36	36%
Rural	23	23%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.3**



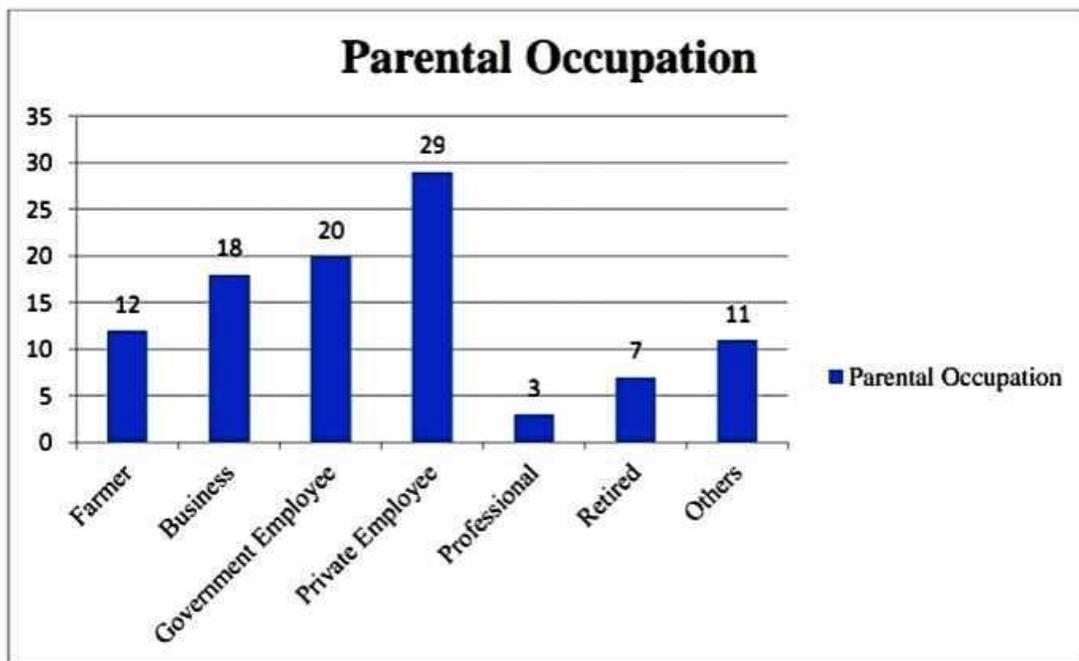
**Interpretation:** From the above table 4.3, 41% of the total responses were from the students belonging to the urban area, 36% from rural area and only 23% from the rural category.

**Table 4.4 PARENTAL OCCUPATION:**

OCCUPATION	FREQUENCY	PERCENTAGE
Farmer	12	12%
Business	18	18%
Government Employee	20	20%
Private Employee	29	29%
Professional	3	3%
Retired	7	7%
Others	11	11%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.4**



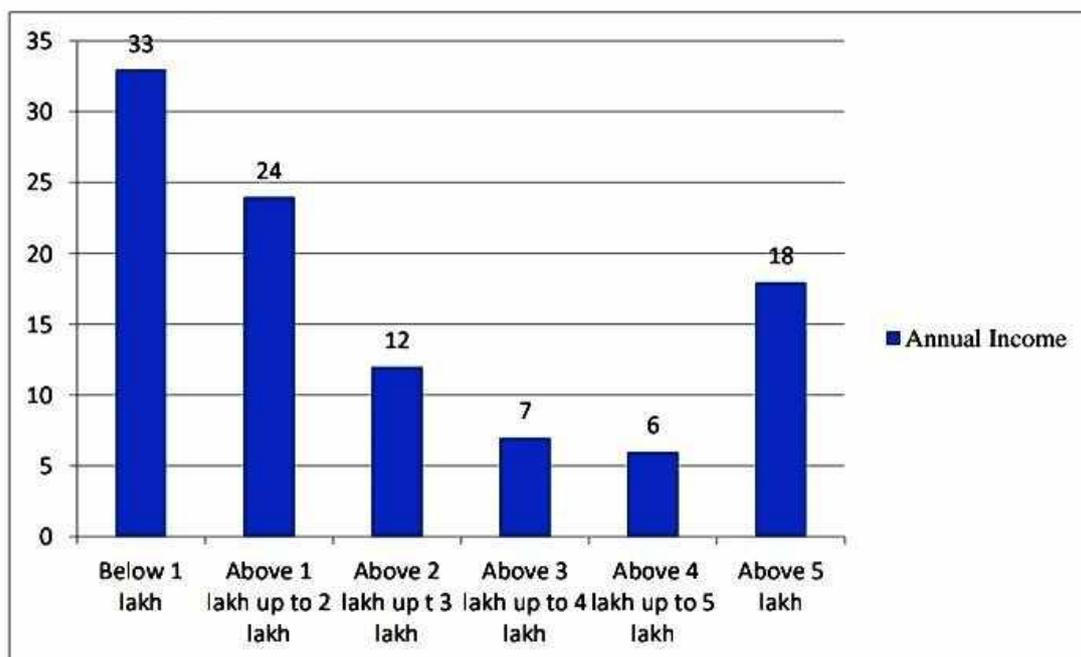
**Interpretation:** The above table 4.4 conveys that almost all the individuals having different occupation take education loans for their children. However, the private employees hold the biggest share. 29% of the respondents were them, the lowest share was owned by the professionals they were only 3% of the total responses. 20% of the responses were government employees followed by businessman with 18%, 12% belonged to the farmers and 7% by the retired employees.

**Table 4.5 ANNUAL INCOME OF THE FAMILY:**

ANNUAL INCOME	FREQUENCY	PERCENTAGE
Below 1 lakh	33	33%
Above 1 lakh up to 2 lakh	24	24%
Above 2 lakh up to 3 lakh	12	12%
Above 3 lakh up to 4 lakh	7	7%
Above 4 lakh up to 5 lakh	6	6%
Above 5 lakh	18	18%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.5**



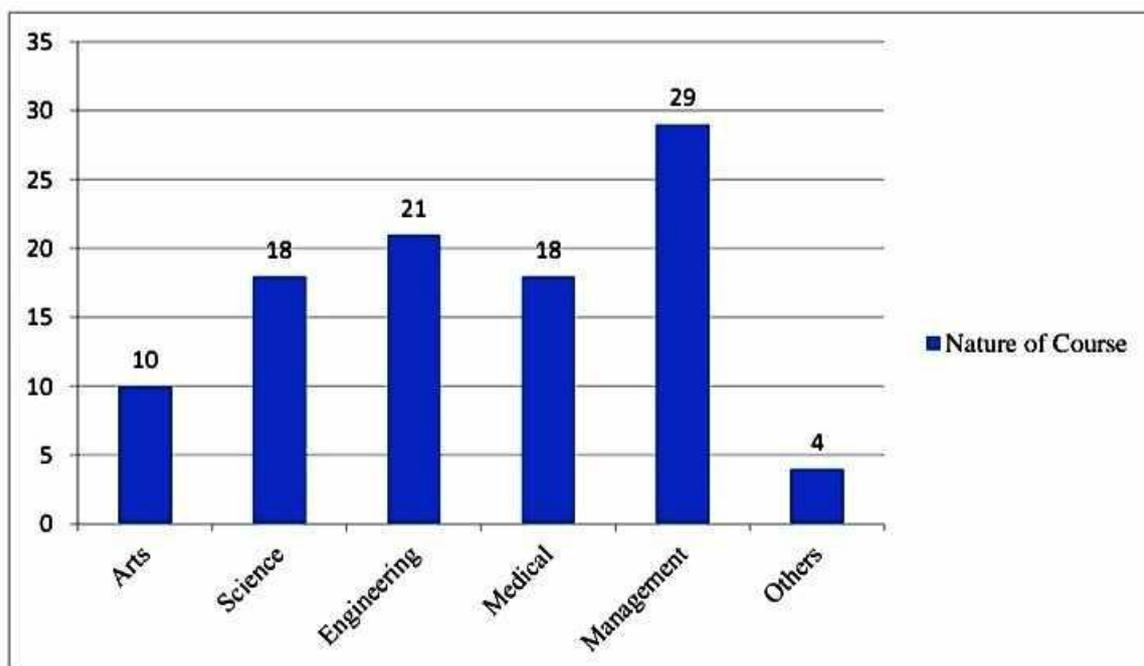
**Interpretation:** Out of the total responses, 33% was from the individuals with annual income below one lakh. 24% was from the individuals whose annual income is above one lakh but not above two lakhs. Individuals belonging to the category of annual income above five lakhs covered 18% out of 100 responses. 12% from those with an annual income of more than 2 lakhs but less than 3 lakhs. 7% was from the individuals with annual income above three lakhs but not above four lakhs and 6% with annual income above four lakhs above five lakhs.

**Table 4.6 NATURE OF COURSE TAKEN:**

COURSE	FREQUENCY	PERCENTAGE
Arts	10	10%
Science	18	18%
Engineering	21	21%
Medical	18	18%
Management	29	29%
Others	4	4%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Source

**Figure 4.6**



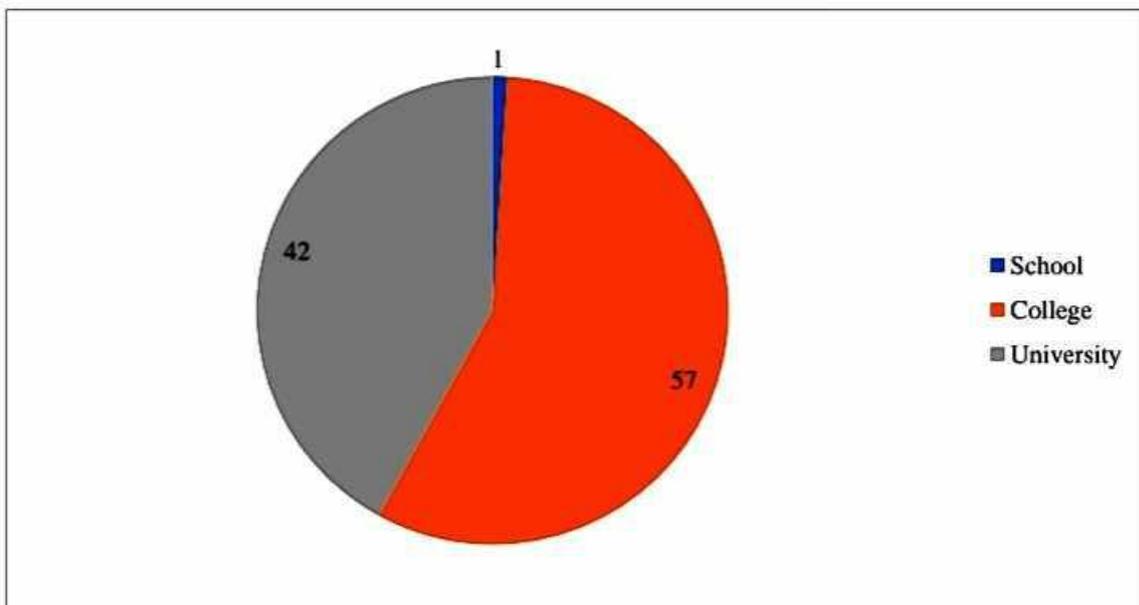
**Interpretation:** Majority of the respondents of 29% had availed education loan for management courses. Engineering courses came in second with 21% of the total responses. Both Science and medical courses is holding 18%. 10% out of the total responses was from the individuals pursuing Arts courses. 4% was from the other category.

**Table 4.7 NATURE OF INSTITUTION UNDERGONE:**

INSTITUTE	FREQUENCY	PERCENTAGE
School	1	1%
College	57	57%
University	42	42%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.7**



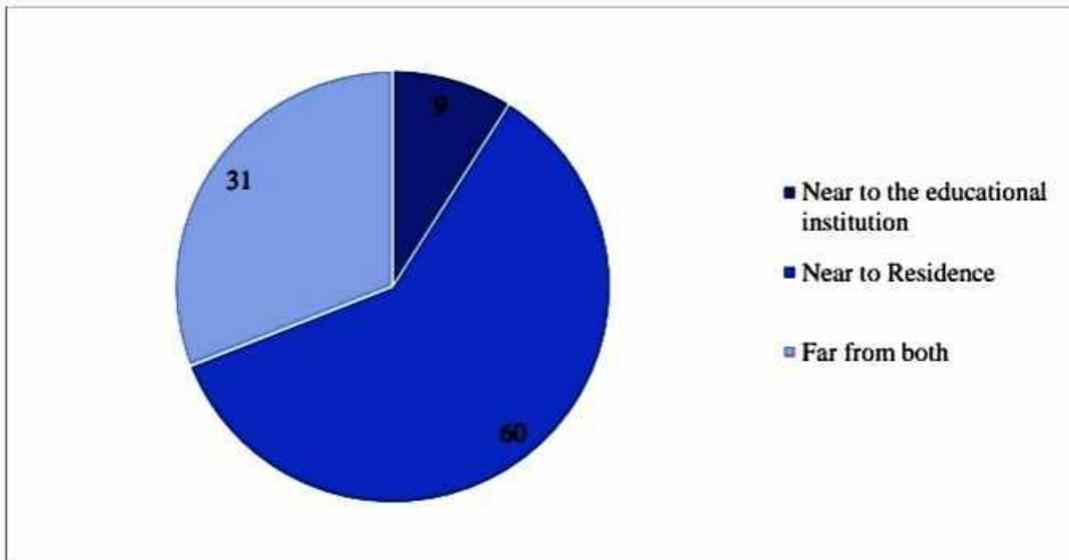
**Interpretation:** It is evident from the table that majority of the individuals had taken education loans for college education. They hold 57% of the responses. There is not much difference with the university courses as 42% of the total responses are from this category. Only 1% out of 100 responses had availed education loan for their school education.

**Table 4.8 LOCATION OF THE BANK:**

LOCATION	FREQUENCY	PERCENTAGE
Near to the educational institution	9	9%
Near to residence	60	60%
Far from both	31	31%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.8**



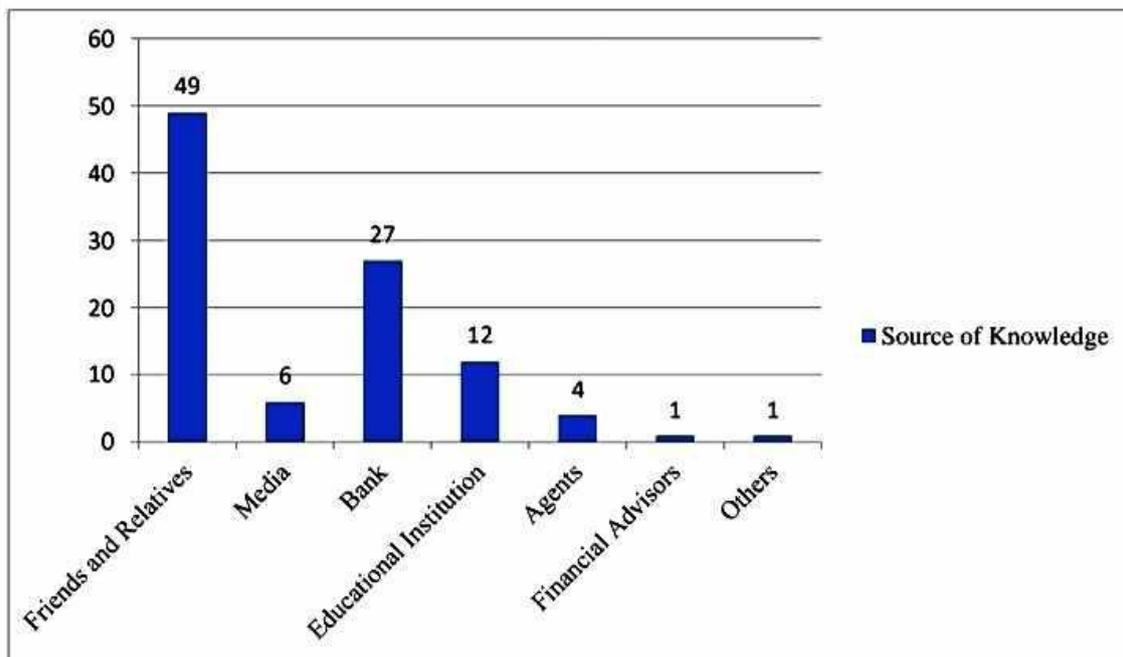
**Interpretation:** Among the total responses, 60% of the individuals had taken education loans from the banks near to their residence. 9% out of the total responses had chosen banks near to their educational institution and 31% of the respondents took education loans from banks far from both their residence and educational institution.

**Table 4.9 SOURCE OF KNOWLEDGE ABOUT EDUCATIONAL LOAN:**

SOURCE	FREQUENCY	PERCENTAGE
Friends and relatives	49	49%
Media	6	6%
Bank	27	27%
Educational institution	12	12%
Agents	4	4%
Financial Advisors	1	1%
Others	1	1%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.9**



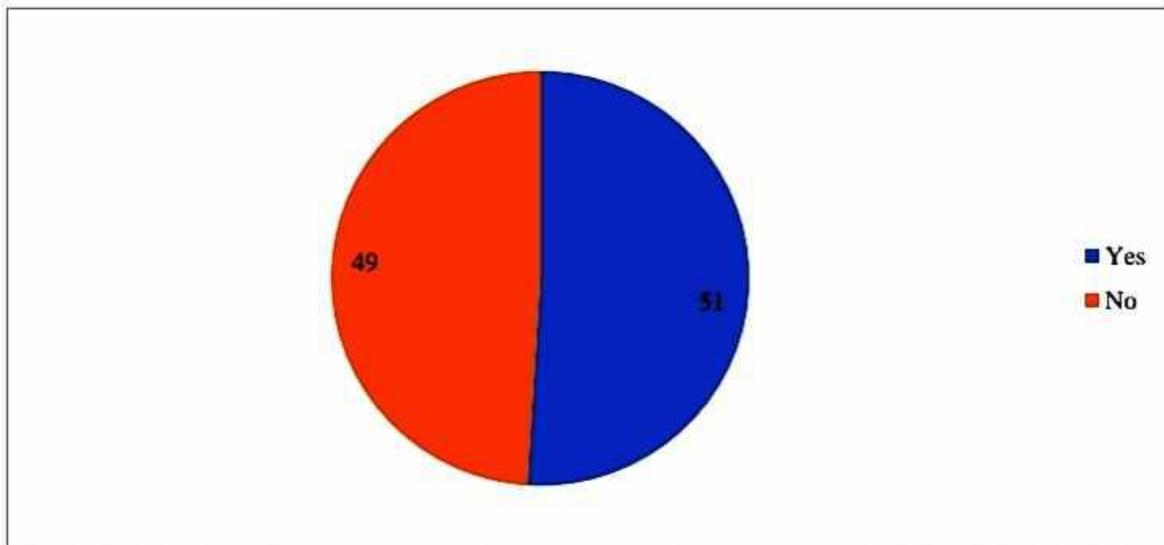
**Interpretation:** Among the total responses, 49% of the individuals came to know about educational loan from friends and relatives, 6% from media, 27% from banks, 12% from educational institution, 4% from agents, 1% from financial advisors and 1% from other sources.

**Table 4.10** SUFFICIECNCY OF LOAN TAKEN:

RESPONSE	FREQUENCY	PERCENTAGE
Yes	51	51%
No	49	49%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.10**



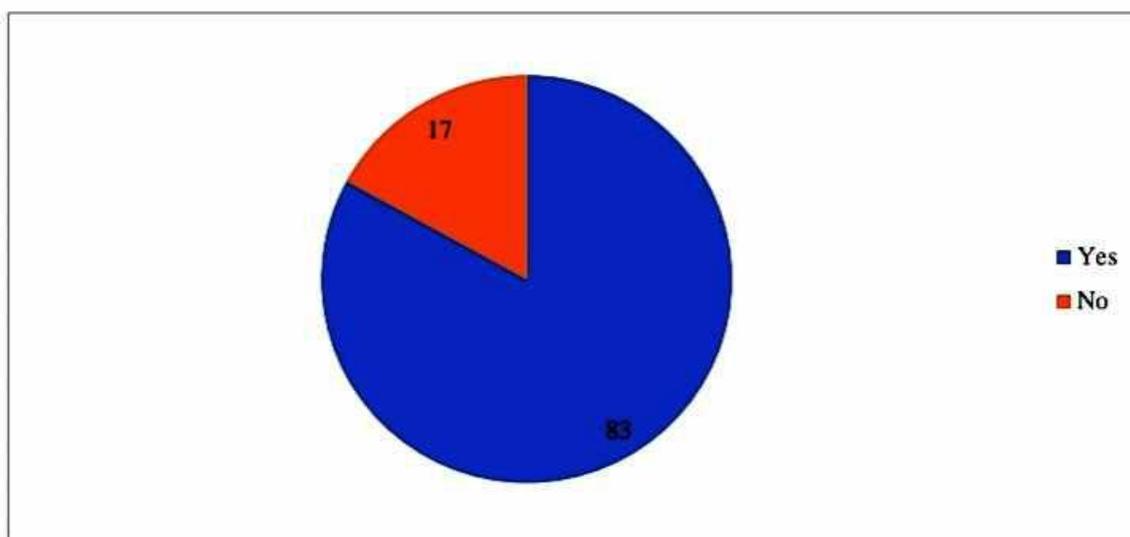
**Interpretation:** Among the responses, 51% of the respondents found the amount of education loan received sufficient and 49% of the respondents did not find the amount sufficient.

**Table 4.11 UTILISATION OF THE LOAN AMOUNT:**

RESPONSE	FREQUENCY	PERCENTAGE
Yes	83	83%
No	17	17%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.11**



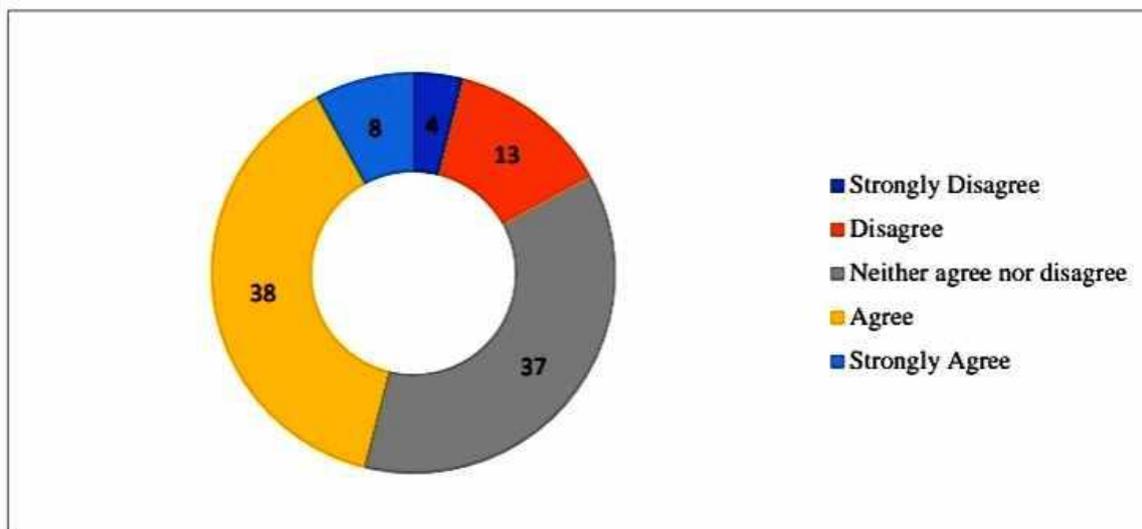
**Interpretation:** Among the total responses, 83% of the responses fully utilised the borrowed amount. Only 17% of individuals were unable to fully utilise the borrowed amount from the bank.

**Table 4.12 ADEQUACY OF BORROWED LOAN:**

RESPONSE	FREQUENCY	PERCENTAGE
Strongly Disagree	4	4%
Disagree	13	13%
Neither agree nor disagree	37	37%
Agree	38	38%
Strongly Agree	8	8%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.12**



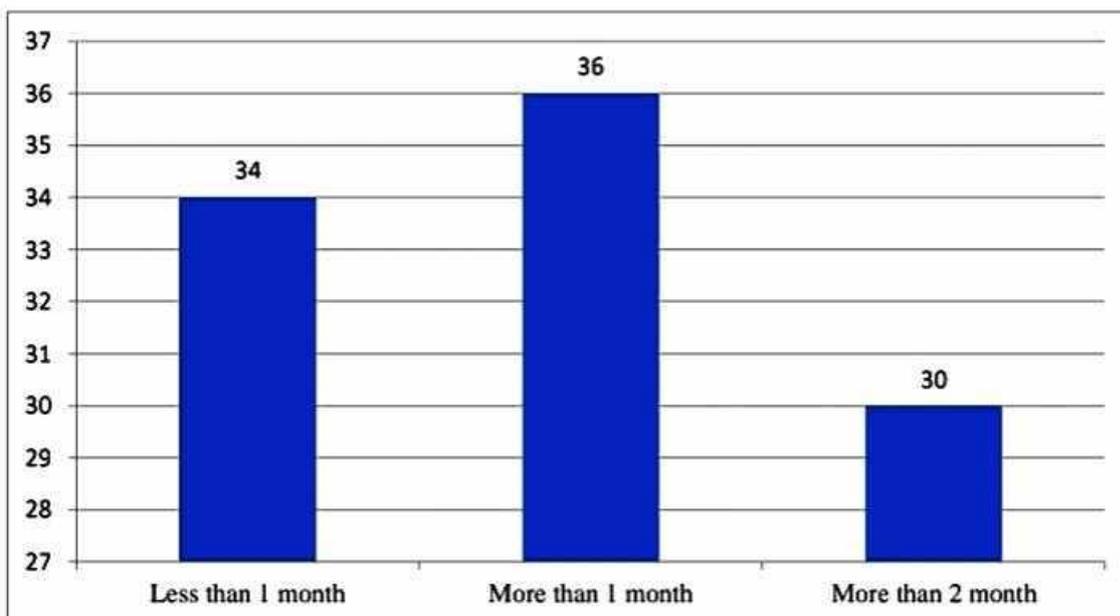
**Interpretation:** Among the total responses, 38% of the individuals agree about the adequacy of borrowed amount, 37% of individuals stated that they neither agree nor disagree about the adequacy of the loan amount and 13% of the individuals disagree about the adequacy of the borrowed amount, 8% strongly agree and 4% strongly disagree about the adequacy of the amount borrowed.

**Table 4.13** TIME TAKEN TO DISBURSE THE LOAN

DURATION	FREQUENCY	PERCENTAGE
Less than 1 month	34	34%
More than 1 month	36	36%
More than 2 month	30	30%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.13**



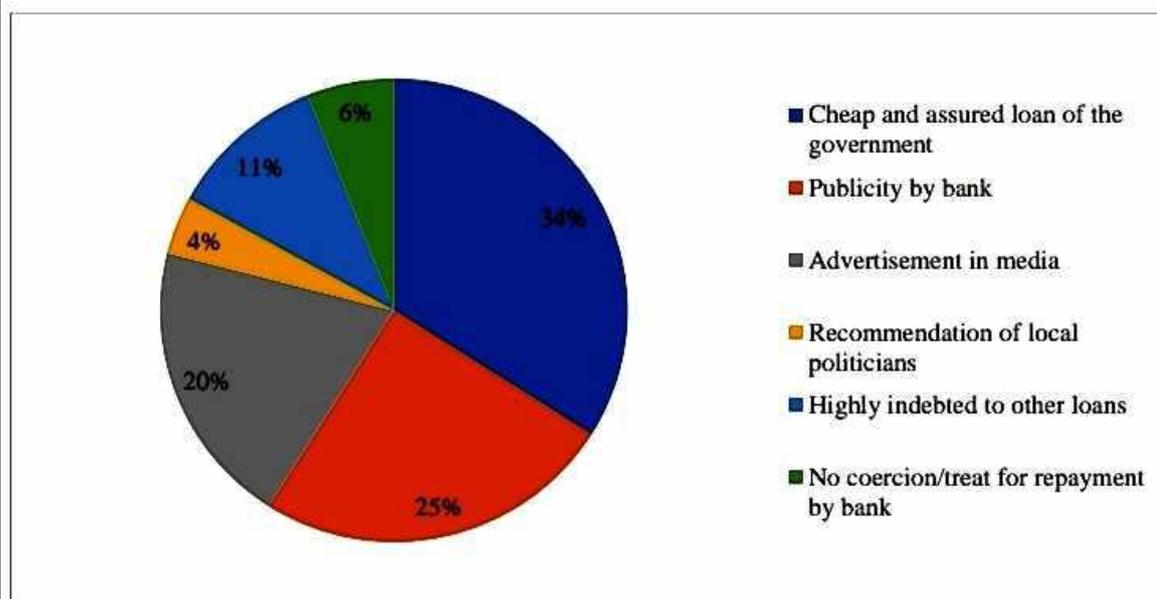
**Interpretation:** Among the total responses, 34% of the individuals state that the time taken for the disbursement of the loan is less than 1 month. 36% the individual states that it takes more than 1 month. 30% of the individuals states that it takes more than 2 months.

**Table 4.14 REASONS TO AVAIL EDUCATIONAL LOAN:**

<b>RESPONSE</b>	<b>FREQUENCY</b>	<b>PERCENTAGE</b>
Cheap and assured loan of the government	34	34%
Publicity by bank	25	25%
Advertisement in media	20	20%
Recommendation of local politicians	4	4%
Highly indebted to other loans	11	11%
No coercion/treat for repayment by bank	6	6%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.14**



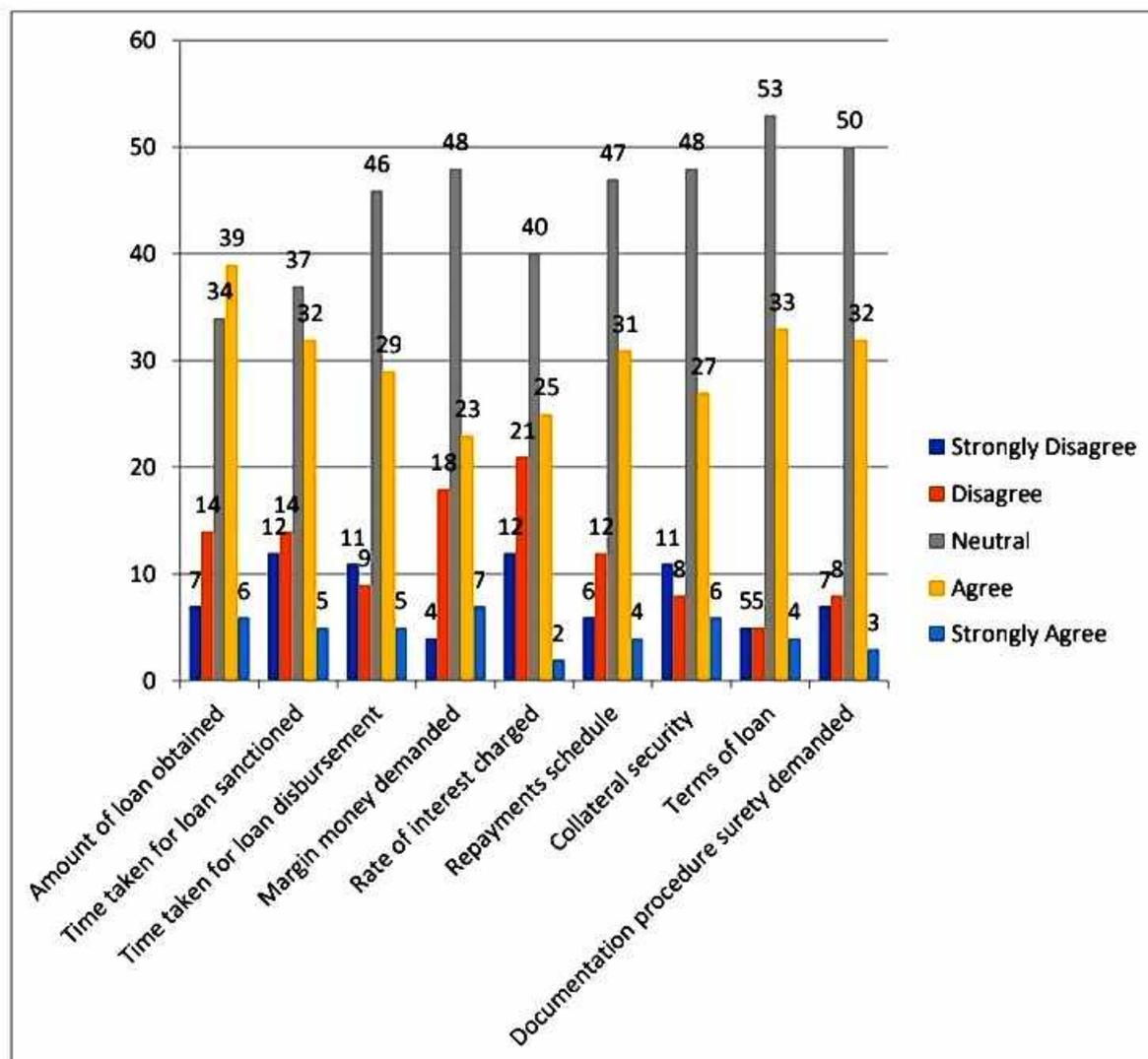
**Interpretation:** Among the total responses, 34% of the individuals took loan due to cheap and assured loan of the government, 25% of them, due to the publicity by banks, 20% of them from the exposure of advertisement in media, 4% due to the recommendation of local politicians, 11% of them were indebted to other loans and 6% of them due to no coercion/treat for repayment by bank.

**Table 4.15 SATISFACTION LEVEL OF VARIOUS ASPECTS REGARDING EDUCATIONAL LOAN:**

<b>ISSUES</b>	<b>STRONGLY DISAGREE</b>	<b>DISAGREE</b>	<b>NEUTRAL</b>	<b>AGREE</b>	<b>STRONGLY AGREE</b>	<b>TOTAL</b>
Amount of loan obtained	7	14	34	39	6	100
Time taken for loan sanctioned	12	14	37	32	5	100
Time taken for loan disbursement	11	9	46	29	5	100
Margin money demanded by the bank	4	18	48	23	7	100
Rate of interest charged	12	21	40	25	2	100
Repayments schedule	6	12	47	31	4	100
Collateral security	11	8	48	27	6	100
Terms of loan	5	5	53	33	4	100
Documentation procedure surety demanded by the bank	7	8	50	32	3	100

Source: Primary Data

**Figure 4.15**



**Interpretation:** Table 4.15 shows the satisfaction level on various aspects related to education loan. Analysis of the data collected on the satisfaction level regarding the amount of loan received by the beneficiaries shows that 39% of the respondents agreed that they are satisfied about the amount of loan, 34% are neutral on it, 14% disagreed, 7% strongly disagreed and 6% strongly agreed. Data on time taken for loan sanctioned shows that 37% of the respondents are neutral, 32 agreed about the time taken for loan sanctioning, 14% disagreed, 12% strongly disagreed and 5% strongly agreed. Satisfaction level of the beneficiaries regarding time taken for loan disbursement shows that 46% are neutral about it, 29% agreed that they are satisfied with the time taken for loan disbursement, 11% strongly disagreed, 9% disagreed and 5% strongly agreed. Satisfaction level regarding margin money demanded by the bank indicates that 48% were neutral on it, 23% agreed, 18% disagreed, 7% strongly agreed that they are satisfied with the margin money demanded by the bank and 4% strongly disagreed. Data collected on the satisfaction level regarding rate of interest charged reveals

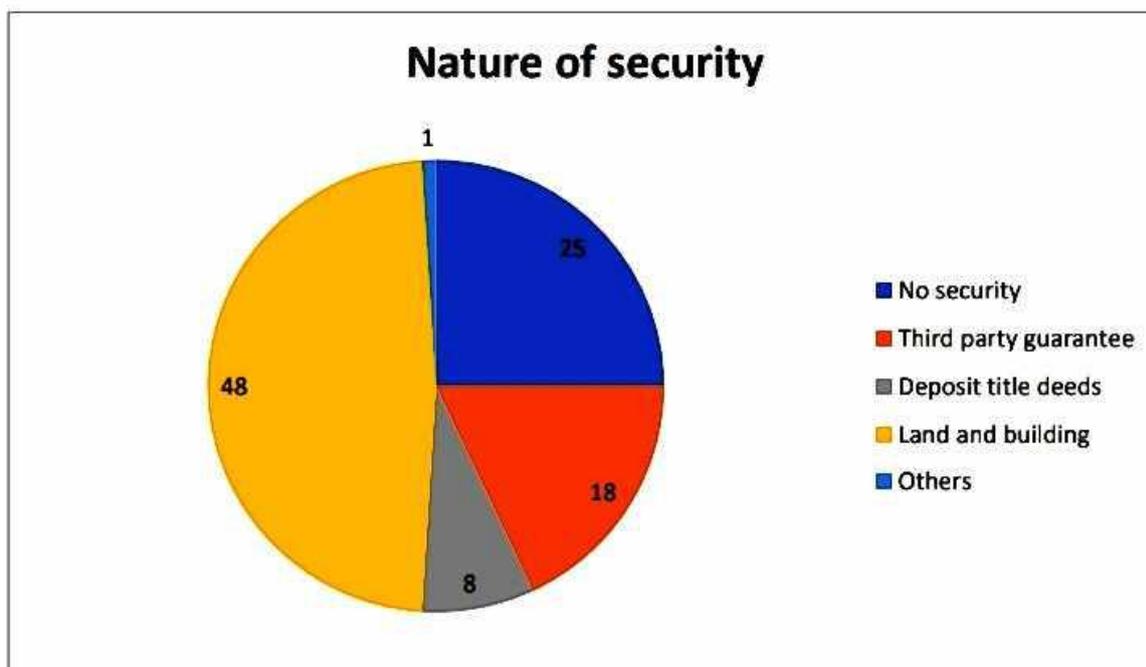
that 40% of the respondents were neutral on it, 25% agreed, 21% disagreed, 12% strongly disagreed and 2% strongly agreed. Satisfaction level regarding repayments schedule from the data collected shows that 47% were neutral on it, 31% of the respondents agreed, 12% disagreed, 6% strongly disagreed and 4% strongly agreed. Data collected on satisfaction level regarding collateral security indicates that 48% respondents were neutral on it, 27% agreed, 11% strongly disagreed, 8% disagreed and 6% strongly agreed. Satisfaction level regarding the data collected on terms of loan provided by the bank shows that 53% of the respondents were neutral on it, 33% agreed, 5% strongly disagreed, 5% disagreed and 4% strongly agreed. Data collected on the satisfaction level of beneficiaries regarding documentation procedure surety demanded by the bank shows that 50% of the respondents were neutral on it, 32% agreed, 8% disagreed, 7% strongly disagreed and 3% strongly agreed.

**Table 4.16 NATURE OF SECURITY OFFERED**

<b>SECURITY</b>	<b>FREQUENCY</b>	<b>PERCENTAGE</b>
No security	25	25%
Third party guarantee	18	18%
Deposit title deeds	8	8%
Land and building	48	48%
Others	1	1%
<b>Total</b>	<b>100</b>	<b>100%</b>

Source: Primary Data

**Figure 4.16**



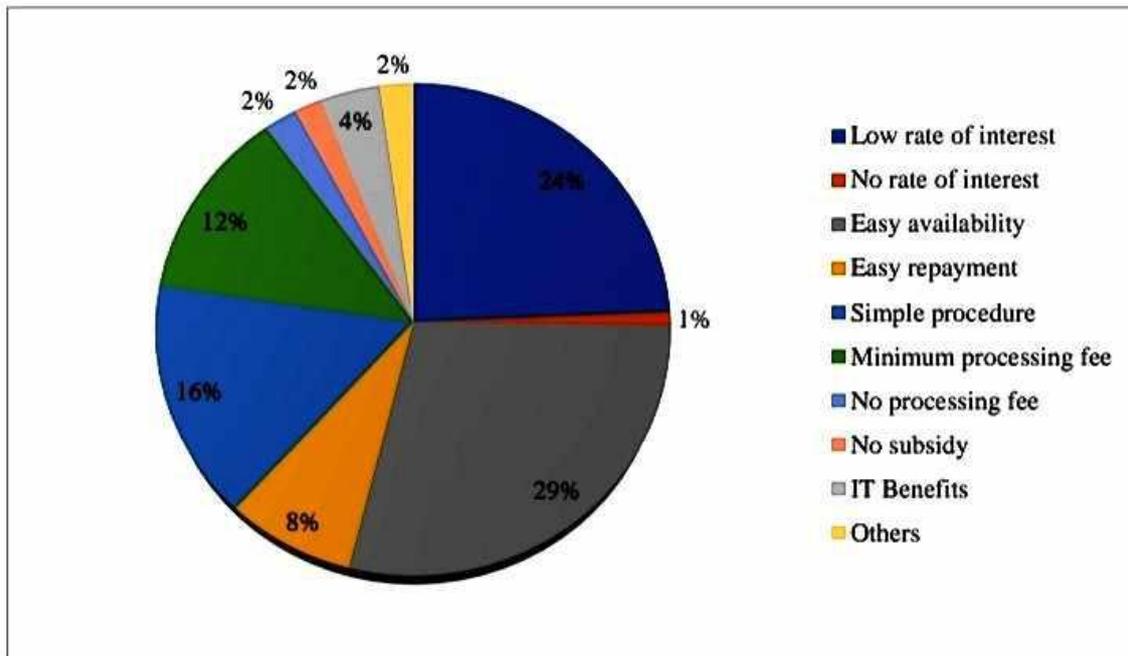
**Interpretation:** From the above table 4.16 shows that 48% of the respondents have pledged land and buildings for availing education loan, 25% of the beneficiaries have not pledged any security, 18% beneficiaries had third party guarantee, 8% had deposit title deeds and 1% from others.

**Table 4.17 REASONS FOR APPROACHING THE BANK FOR EDUCATION LOAN**

REASONS	FREQUENCY	PERCENTAGE
Low rate of interest	56	56%
No rate of interest	2	2%
Easy availability	66	66%
Easy repayment	18	18%
Simple procedure	36	36%
Minimum processing fee	29	29%
No processing fee	5	5%
No subsidy	4	4%
IT Benefits	9	9%
Others	5	5%

Source: Primary Data

Figure 4.17



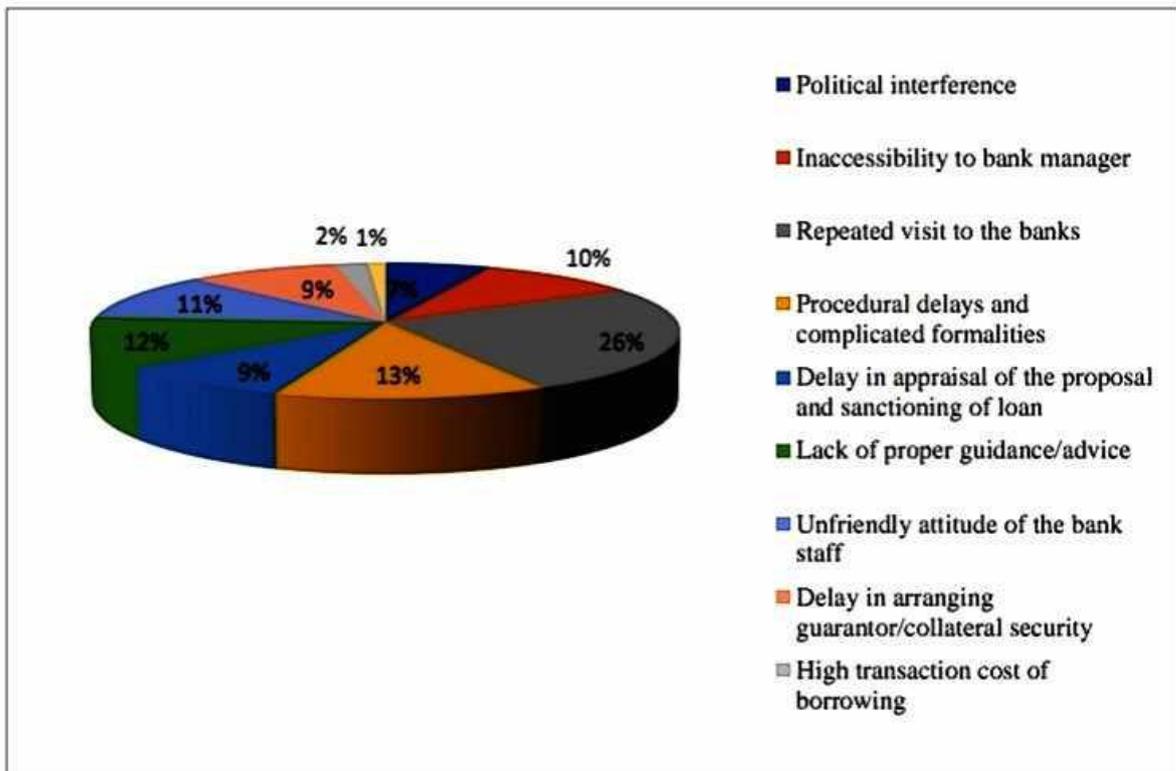
**Interpretation:** From the above table 4.17, 66% of respondents opted for educational loan because of easy availability of loan, 56% opted because of low rate of interest provided by the bank, 36% took loan due to simple procedure, 29% chose since it has only minimum processing fee, 18% of people opted for loan due to easy repayment of loan, 9% of the borrowers opted because of the IT benefits they will receive, 5% from both no processing fee and others and 4% opted because of no subsidy.

**Table 4.18 PROBLEMS FACED DURING THE PRE-LOAN SANCTION PERIOD**

<b>PROBLEMS</b>	<b>FREQUENCY</b>	<b>PERCENTAGE</b>
Political interference	18	18%
Inaccessibility to bank managers	28	28%
Repeated visits to the banks	71	71%
Procedural delays and complicated formalities	37	37%
Delay in appraisal of the proposal and hence sanctioning of loan	25	25%
Lack of proper guidance or advice by the field officer	33	33%
Unfriendly attitude of the bank staff	31	31%
Delay in arranging guarantor or proving collateral security	26	26%
High transaction cost of borrowing	6	6%
Others	3	3%

Source: Primary Data

Figure 4.18



**Interpretation:** Among the total response, 71% of the respondents have faced issue regarding repeated visit to the banks, 37% of the people had issues concerning with procedural delays and complicated formalities, 33% had issues regarding lack of proper guidance or advice, 31% dealt with unfriendly attitude of the banking staff, 28% had issues with inaccessibility to bank managers, 26% had difficulties in arranging guarantor or collateral security, 25% faced delay in appraisal of the proposal and hence sanctioning of loan, 18% of the borrowers faced political interference, 6% had issues with high transaction cost of borrowing and 3% out of the total responses had specified in the other category that they had no issues during the pre-sanctioning of the loan.

**CHAPTER 5**  
**SUMMARY, FINDINGS, RECOMMENDATIONS AND**  
**CONCLUSION**

## **5.1 SUMMARY**

The number of students pursuing higher education is increasing day by day thereby the importance of education loan has also increased. In this scenario our project studies the perception of the beneficiaries of educational loan. This research conducted a survey of 100 borrowers, found various information regarding the satisfaction level of the borrowers, adequacy of the education loan provided to the beneficiaries, which sector of financing bank they prefer and the reason they availed educational loan among many other information.

## **5.2 FINDINGS**

1. From the light of study, it is evident that beneficiaries who have taken education loan are more males compared to females. But with the passage of time, this trend is changing. Women and men comparatively show equal participation. And till now, other categories haven't showed any participation.
2. Comparatively public sector is the epitome of lending educational loans. Private sector banks are seeing low. This trend hasn't changed yet.
3. In the case of area of residence, urban areas precedes its participation. This is due to the lack of education and awareness in the rural area. Urban areas are rich in amenities and facilities and know the importance of education.
4. Educational loan is majorly taken by children of private employees. This indirectly shows the low income from private companies. But neither a child of government employees avoid educational loan.
5. The study shows that the students with low annual income that is below 1 lakh prone to avail educational loans.
6. Students who opt management and engineering courses shows high interest in educational loan. This is due to the high expenses incurred during the course time period.
7. According to the survey, the majority of the borrowers have taken education loan for attending colleges.

8. From the analysis it's apparent that the borrowers have taken loan from bank which is near to their residence.
9. The study shows that majority of the beneficiaries obtained knowledge about educational loans through their friends and relatives and through banks.
10. The analysis shows that most of the respondents received the loan they originally requested.
11. According to the survey, the majority of recipients made full use of the education loan for their higher studies expenditure.
12. The study shows that most of the respondents agree that the loan amount is enough to cover the cost of studying. However, some of them neither agree nor disagree.
13. It was found that it takes a minimum of 1 month period for the disbursement of the loan amount in most of the cases. This can badly affect students because they have to make the course payments within in the deadlines
14. Public sector banks have succeeded in gaining trust of the people. Most of the borrowers chose education loans from commercial banks due to the cheap and assured loan schemes offered by the commercial banks, especially the public sector commercial banks.
15. Satisfaction level of the beneficiaries on various aspects related to education loan were mostly satisfied and majority neither agree nor disagree.
16. The public sector banks provide education loans up to a certain amount without any collateral security. But for higher amounts land, building, gold etc is given as security. The individuals who are not able to provide such collateral security due to their poor economic conditions are unable to get education loans and thereby lose their opportunity for better education
17. Majority of the people depend on education loans for their higher education expenses. Education loans are easily available compared to other sources. And the rate of interest is not much high.
18. The major problem faced by the borrowers in the pre- loan sanction period is the repeated visits to the bank. The procedure for obtaining loan is so complicated that the beneficiaries had to repeatedly visit the banks.

### **5.3 RECOMMENDATIONS**

1. Time lag and procedural delay in educational loan disbursement is to be avoided by adopting speedy and systematic procedure especially in commercial banks of semi-urban and rural areas.
2. The collateral security requirements should be reduced for the people belonging to the backward category only then education loans will be made available to people who are really in need of it.
3. The rate of interest for education loans should be reduced in the private sector banks.
4. The amount given as loan should be based on the expenses incurred on each course because different courses have different fee structure.
5. The procedural constraints should be reduced because the students have to pay their course fees within the deadlines. Hence to reduce the time taken for loan disbursement is to be reduced by making the loan application procedure simpler.
6. The repeated visits to the bank for loan sanction can be avoided by encouraging the borrowers to use online mode of application.

## 5.4 CONCLUSION

Students in higher education require more money to cover the cost of study and living in universities as higher education costs rise in both public and private educational institutions. Students may be struggling as a result of limited financial resources and rising education costs. They face financial difficulties in their daily lives. This study helps people to understand about educational loan. A statistical analysis has been taken place. In light of 100 responses, we came to an end where public sector shows more participation than private sector, that too in urban areas. School students taking loan is somewhat small as school doesn't incur much of an expense. Due to the over indulgence of bank in their marketing field, commercial banks bags a lot of awareness and details. Beneficiaries share the same interest in finding cheap and discount. Clients chose commercial banks for its cheap and assured loan, that too from government. Majority of people chose land and buildings as collateral or security. This is due to the reduced procedure in availing educational loan from commercial banks. Easy availability of funds for higher education is the greatest merit of educational loan, and then comes low interest rates. Any other lending forms carry high amount of interest rates. Whereas educational loan interest rates are feasible and accountable. Some of the disastrous demerits of educational loan is its repeated visits to banks and lengthy procedures. This happens with the attitude of banks and its employees. This discourages people from taking education loans. Inaccessibility to the needed service and knowledge lead the clients' nauseas and unsatisfied. Even in urban areas, inaccessibility to services is very common. Majority of students were satisfied with the given amount. And majority of students were able to utilize the lend loans for which it was taken. This study encounters about the basic and crucial data of beneficiaries, and helps all of the students who have a natural instinctive image of the future.

## **REFERENCES**

## **JOURNALS:**

Tilak, J. B. G. (1992). Student loans in financing higher education in India. *Higher Education*, 23(4), 389–404.

Chattopadhyay, S. (2007). Exploring Alternative Sources of Financing Higher Education. *Economic and Political Weekly*, 42(42), 4251–4259.

Dr.G.Raviselvam., & G.Maheswari (2008). A study on student satisfaction towards educational loan from commercial banks with special reference to Thanjavur City. *International Journal of Creative Research Thoughts (IJCRT)*, Volume 6, Issue 2, ISSN: 2320-2882, April 2018.

Shen, H., & Ziderman, A. (2009). Student loans repayment and recovery: international comparisons. *Higher education*, 57(3), 315-333.

Tiwari, R., & Anjum, B. (2013). Role of Education Loan in Indian Higher Education (SSRN Scholarly Paper ID 3833834). *Social Science Research Network*.

Rani, P. G. (2014). Education loans and financing higher education in India: Addressing equity. *Higher Education for the Future*, 1(2), 183-210.

Rani, S (2014). Education loans in professional higher education: A case study of Punjab.

Paramanandam, D. A., & Packirisamy, P. (2015). An empirical study on the impact of micro enterprises on women empowerment. *Journal of Enterprising Communities: People and Places in the Global Economy*.

Shinde, M. B (2015). A Study of Significant Impact of Education Loan in Maharashtra with Special Reference to State Bank of India.

Panjali, N. (2015). Perception of Bankers and Beneficiaries towards Education Loan (Doctoral dissertation).

Ramachandran, S.(2015) An analysis of problems faced by public sector banks towards retrieval of higher education loan with special references to professional courses in Tamilnadu.

Bandyopadhyay, A. (2016). Studying borrower level risk characteristics of education loan in India. *IIMB Management Review*, 28(3), 126–135.

John, M. (2016). Current Scenario in Education Loans in India. *IRA-International Journal of Management & Social Sciences*, 3(1), 96-109.

Kumar, S. S. (2018). Education loan and non-performing assets: banker's perspectives. *TRANS Asian Journal of Marketing & Management Research (TAJMMR)*, 7(1), 134-143.

Sebastian, S. (2018). An evaluation of education loan scheme of major commercial banks in the select districts of Kerala.

Dr. Thulasi, Priya., & C. Esakkiammal (2018). A study on bank-wise performance of education loan lent by public sector banks in India. *International Journal of Pure and Applied Mathematics Volume 119 No. 18 2018*, 3081-3090.

Dr. Tania Gupta (2018). A study on accessibility of educational loans in India for financing higher education. *Impact factor, UGC approved journal No. 48514, ISSN: 2249-894X, Vol-7, Issue-12, September 2018*.

Dr. Manisha (2018). Education loans in Punjab- An analysis. *Journal of advanced studies in Education and Management Winter, Vol. 1, No.4, ISSN No. 2350-0492*.

Santhosh, K.B., & A. Saravanakumar (2018). Preference of Education Loan- A study with special reference to Canara Bank in Coimbatore. *International Journal for Research in Engineering Application & Management (IJREAM) ISSN: 2454-9150 Vol-04, Issue-03, June 2018*.

Kaur, J., & Arora, S. (2019). Indian students' attitude toward educational debt: Scale development and validation. *Quality Assurance in Education*, 27(4), 361–383.

Khanwalker, S. M. (2019). Education Loan in India-A Review. *International Journal of banking, Risk and Insurance*, 7(1), 48.

Sivakumar N (2019). Students Perception on Educational Loan with special reference to Coimbatore District.

Esakkiammal C (2019). Performance Evaluation of Public Sector Banks Education Loan and Borrowers Perception an Empirical Study.

**WEBSITES:**

[www.google.com](http://www.google.com)

[www.shodhganga.inflibnet.ac.in](http://www.shodhganga.inflibnet.ac.in)

[www.researchgate.com](http://www.researchgate.com)

[www.booksc.com](http://www.booksc.com)

## **APPENDIX**

# **QUESTIONNAIRE**

## **QUESTIONNAIRE TO THE BENEFICIARIES**

As a part of our B.COM final year project, we the students of St. Teresa's College, Ernakulam are conducting a survey on "A study on education loans provided by commercial banks in Ernakulam". Please be kind enough to spend few minutes to fill in the questionnaire if you have availed education loan. Any information provided by you will be used for academic purpose only.

1. District:

2. Gender:  Male  Female  Others

3. Financing Bank:  Public Sector bank  Private Sector Bank  New Generation Bank

4. Area of residence:  Urban  Semi-urban  Rural

5. Parental occupation:

Farmer

Business

Government Employee

Private employee

Professional

Retired

Others (Please Specify): \_\_\_\_\_

6. Annual income of the Family:

- Below 1 Lakh
- Above 1 Lakh up to 2 Lakh
- Above 2 Lakh up to 3 Lakh
- Above 3 Lakh up to 4 Lakh
- Above 4 Lakh up to 5 Lakh
- Above 5 Lakh

7. Nature of course for which loan is obtained:

- Arts
- Science
- Engineering
- Medical
- Management
- Others (Please specify): \_\_\_\_\_

8. Nature of institution where education is undergone:

- School
- College
- University

9. The bank from which you have availed loan is near to:

- Near to the educational institution
- Near to residence
- Far from both residence and place of study

10. How did you come to know about Educational loan?

- Friends & Relatives
- Media
- Bank
- Educational institution
- Agents
- Financial Advisors
- Others (Please specify): \_\_\_\_\_

11. Did you get the total amount of the loan that you have applied for?

- Yes
- No

12. Do you utilize the loan fully for the purpose for which applied?

- Yes
- No

13. The borrowed loan was adequate to meet education expenditure

- Strongly Disagree
- Disagree
- Neither agrees nor disagrees
- Agree
- Strongly Agree

14. How much time was taken by the bank to disburse the sanctioned loan?

- Less than 1 month
- More than 1 month
- More than 2 months

15. What prompted you to avail educational loan from commercial bank?

- Cheap and assured loan of government
- Publicity by bank
- Advertisements in Media
- Recommendation of local politicians
- Highly indebted to other loans
- No coercion/treat for repayment by bank

16. Rate your level of satisfaction towards the following issues relating to education loan by Strongly Disagree, Disagree, Neutral, Strongly Agree and Agree.

SL.NO.	Statements	Level of Satisfaction
i.	Amount of loan obtained:	
ii.	Time taken for loan sanctioned:	
iii.	Time taken for loan disbursement:	
iv.	Margin money demanded by the bank:	
v.	Rate of Interest charged:	
vi.	Repayments schedule:	
vii.	Collateral Security:	
viii.	Terms of Loan:	
ix.	Documentation procedure Surety demanded by the bank:	

17. Nature of security offered:

- No security
- Third party guarantee
- Deposit title deeds
- Land and Building
- Others (Please specify): \_\_\_\_\_

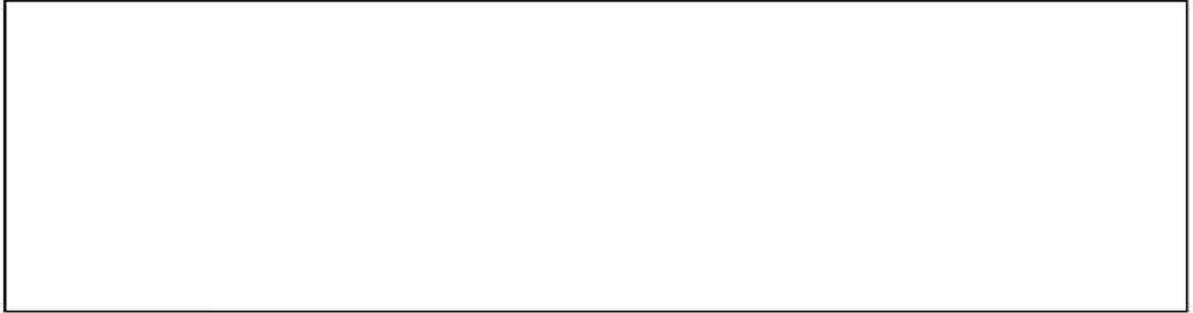
18. Select the reasons for having approached the bank for Educational Loan:

- Low Rate of Interest
- No Rate of Interest
- Easy Availability
- Easy Repayment
- Simple Procedure
- Minimum Processing fee
- No processing fee
- For subsidy
- IT Benefits
- Others (Please specify): \_\_\_\_\_

19. Select the Problems faced during the pre loan sanction period:

- Political interference
- Inaccessibility to bank manager
- Repeated visit to the banks
- Procedural delays and complicated formalities
- Delay in appraisal of the proposal and hence sanctioning of loan
- Lack of proper guidance/ advice by the field officer
- Unfriendly attitude of the bank staff
- Delay in arranging guarantor/ proving collateral security
- High transaction cost of borrowing
- Others (Please specify): \_\_\_\_\_

20. Suggest measures to improve the positive perception of customers with regards to Education loan:



# **A STUDY ON THE IMPACT OF ADVERTISEMENTS ON CONSUMERS WITH SPECIAL REFERENCE TO COSMETICS INDUSTRY**

## **Project Report**

Submitted by

**EDWINA P EDISON (REG.NO.AB19COM043)**

**JOMOL A J (REG.NO.AB19COM044)**

**MALAVIKA K S (REG. NO.AB19COM045)**

Under the guidance of

**Ms. SHANA XAVY**

In partial fulfillment of requirements for award of the degree of

**Bachelor of Commerce**



**Nationally re- accredited by A ++ Level**

**Affiliated to**

**MAHATMA GANDHI UNIVERSITY**

**Kottayam – 686560**

**MARCH 2022**

**ST TERESA'S COLLEGE (AUTONOMOUS) ERNAKULAM**  
**COLLEGE WITH POTENTIAL FOR EXCELLANCE**

**Nationally re- accredited by A ++ Level**



**CERTIFICATE**

This is to certify that the project report titled “**A STUDY ON THE IMPACT OF ADVERTISEMENTS ON CONSUMERS WITH SPECIAL REFERENCE TO COSMETICS INDUSTRY**” submitted by **EDWINA P EDISON, JOMOL A J, MALAVIKA K S** towards partial fulfillment of the requirements for the award of degree of **Bachelor of Commerce** is a record of bonafide work carried out by them in the academic year 2021 – 2022

**Supervising Guide**

**Ms. Shana Xavy**

**Assistant Professor**

**Department of Commerce**

**Head of the Department**

**Ms. Ann Thomas Kiriyanthan**

**Assistant Professor**

**Department of commerce**

**Place: Ernakulam**

**Date: 31.3.2022**



## DECLARATION

We, Malavika k s, Edwina P Edison, Jomol A J hereby declare that this dissertation entitled, “**A STUDY ON THE IMPACT OF ADVERTISEMENTS ON CONSUMERS WITH SPECIAL REFERENCE TO COSMETICS INDUSTRY**” has been prepared by us under the guidance of Mrs. SHANA XAVY, Assistant Professor, Department of Commerce, St Teresa’s College, Ernakulam.

We also declare that this dissertation has not been submitted by us fully or partly for the award of any Degree, Diploma, Title or Recognition before.

Place: ERNAKULAM

MALAVIKA KS



Date: 31.3.2022

EDWINA P EDISON



JOMOL A J



## ACKNOWLEDGEMENT

We wish to acknowledge all those who helped us to complete this study. We thank God almighty for helping us and guiding us in the right path and who made all things possible

We take this opportunity to express our profound gratitude and deep regards to our guide Ms. Shana Xavy, Assistant professor, St. Teresa's college, Ernakulam for her exemplary guidance, monitoring and constant encouragement throughout the course of this study. The blessings, help and guidance given by her time to time shall carry us a long way in the journey of life on which we are about to embark. She has taken pain to go through the project and make necessary corrections as and when needed. We express our sincere thanks to the Director Rev. Sr. Emeline CSST, Rev. Dr. Sr. Vinitha CSST (Provincial superior and manager), Principal Dr. Lizzy Mathew and Ms. Ann Thomas Kiriyanthan , Head of the Department of Commerce ,to all other faculties of the department of commerce, St. Teresa's College, for their support and valuable suggestions.

We would like to express our thanks to all respondents and colleagues in developing the project.

We also extend heartfelt thanks to our family for their constant encouragement without which this project would not be possible

MALAVIKA KS

EDWINA P EDISON

JOMOL A J

# CONTENTS

Sl. No:	Particulars	Page No:
1	List of tables	
2	List of figures	
3	Chapter 1 –Introduction	
4	Chapter 2 – Review of Literature	
5	Chapter 3 – Theoretical Framework	
6	Chapter 4 -Data Analysis and Interpretation	
7	Chapter 5 -Summary, Findings, Recommendations and Conclusion	
8	Bibliography	
9	Appendix-Questionnaire	

## **LIST OF TABELS**

Table No	Title	Page No.
4.1	Gender of Respondents	
4.2	Age of Respondents	
4.3	Category of Respondents	
4.4	Use of cosmetic products	
4.5	Preference on the type of products	
4.6	Mode of purchasing	
4.7	Frequency in purchase	
4.8	Monthly income spend	
4.9	Buying products after watching advertisement	
4.10	Factors influencing in choosing products	
4.11	Influence of celebrity endorsement	
4.12	Factors motivating the consumers	
4.13	Advertisement that has long last impact	
4.14	Expensive or Cheaper	

4.15	Believe in promised effect of products	
4.16	Quality and standards of advertisement	
4.17	Use of Sample products	
4.18	Changing perception on the basis of quality advertisement	
4.19	Preference of same brand of products	
4.20	Switching the brands due to advertisement	
4.21	Belief of information through advertisement	
4.22	Use in day to day life	
4.23	Availability at fair rate	

## LIST OF FIGURES

Figure No	Title	Page No.	
4.2	Age of Respondents		
4.3	Category of Respondents		
4.4	Use of cosmetic products		
4.5	Preference on the type of products		
4.6	Mode of purchasing		
4.7	Frequency in purchase		
4.8	Monthly income spend		
4.9	Buying products after watching advertisement		
4.10	Factors influencing in choosing products		
4.11	Influence of celebrity endorsement		
4.12	Factors motivating the consumers		
4.13	Advertisement that has long last impact		

4.14	Expensive or Cheaper		
4.15	Believe in promised effect of products		
4.16	Quality and standards of advertisement		
4.17	Use of Sample products		
4.18	Changing perception on the basis of quality advertisement		
4.19	Preference of same brand of products		
4.20	Switching the brands due to advertisement		
4.21	Belief of information through advertisement		
4.22	Use in day to day life		
4.23	Availability at fair rate		

**Chapter1**  
**INTRODUCTION**

# A STUDY ON THE IMPACT OF ADVERTISEMENTS ON CONSUMERS WITH SPECIAL REFERENCE TO COSMETICS INDUSTRY

## 1.1 INTRODUCTION

Cosmetic advertising is the promotion of cosmetics and beauty products by the cosmetic industry through a variety of media. The advertising campaigns are usually aimed at women wishing to improve their appearance commonly to increase physical attractiveness and reduce the sign of ageing.

Cosmetics are constituted mixtures of chemical compounds derived from either natural sources or synthetically created ones. Cosmetics have various purposes. Those created for personal care and skin care used to cleanse or protect the body or skin. Cosmetics designed to enhance ones appearance can be used to conceal blemishes, enhance ones natural features (such as eyebrow and eyelashes) add color to person's face and can be used to change the appearance of the face entirely to resemble a different person, creature or object. Cosmetics can be also designed to add fragrance to body.

The manufacture of cosmetics is dominated by the small number of multinational corporations that originated in the early 20<sup>th</sup> century, but the distribution and sales of cosmetics is spread among a wide range of businesses. The world's largest cosmetics companies are L'Oreal, Procter & Gamble, Unilever, Shiseido, and Estee lauder. In 2005, the market volume of the cosmetic industry in the US, Europe and Japan was about EUR 70 billion a year. In Germany cosmetic industry generated 12.6 billion of retail sales in 2008, which makes the German cosmetic industry the third largest in the world, after Japan and US. German exports of cosmetic reached 5.8 billion in 2008, whereas imports of cosmetics totaled 3 billion.

## **1.2 SIGNIFICANCE OF THE STUDY**

The use of cosmetics and the impact of advertisements in influencing the society in using cosmetics and their result in cosmetics industry particularly in this 21<sup>st</sup> century has been dealt in this study. The changing attitude of people regarding beauty has enhanced the need for this study. The study analyse the effect of ads in influencing and encouraging the target group in purchasing the product. More over it also emphasis on the benefits and drawbacks suffered by the cosmetic industry through these ads .

## **1.3 SCOPE OF STUDY**

Now a days cosmetics has become the part of man's day to day life. Even men doesn't stand back in from using cosmetics for their grooming and make over. The barriers of gender is also ignored in this. After the penetration of more cosmetic brands the competition has increased to next level. The fittest will survive rule is applied in the market. The scope of the study is to understand and evaluate the measure and impact that the advertisements make in this sector and its effects on cosmetics companies. The area under study is the people of Ernakulam district. It also studies the future growth of the industry and creativity in ads.

## **1.4 Problem Statement**

Ever since the attitude of the people has changed regarding beauty the need for change in cosmetics industry was evident. The growing popularity for being original and natural compelled the industry to change from the traditional approach of beauty to the new era of beauty. The study aims to find the various modes through which ad agencies help the cosmetic industry to get success in their attempt to adapt themselves to the changes and also the ill effects of relying ads in this process.

## **1.5 METHODOLOGY**

### **1.6.1 RESEARCH DESIGN**

The research mainly focus on collecting Data from the respondent using questionnaire and analysis of data in order to arrive at a conclusion .The data was analyzed using basic statistical tool. The data collection was primary and secondary in nature. The total sample size for this research study was respondents. In this research paper the data collection was done through primary and secondary medium.

### **1.6.2 COLLECTION OF DATA**

☐ Primary data- Primary data for the project are collected by supplying questionnaires among the respondents within the study ☐ Secondary data- Secondary data are collected from books, internet, and newspapers.

### **1.6.3 SAMPLE DESIGN**

Sampling technique used for this study is convenience sampling. Convenience sampling is a type of non- random sampling used for collecting samples of the target population that meet certain practical criteria like geographical location, easy accessibility or the willingness to participate are included for the purpose of study.

Population- The population considered under this study is limited to Ernakulum district

Sample Size- A sample of 138 respondents was taken by supplying the questionnaire

The methodology of the study has been summarized below:

Sources of data Primary and secondary data

Sample Size :138 RESPONSES

Sample area :Residence of Ernakulum district

Sampling design: Convenient sampling

Sources of primary data: Questionnaires

Sources of secondary data : Internet, journals, newspapers

#### **1.6.4 TOOLS FOR ANALYSIS:**

The collected data are analysed with the help of statistical tools like Percentage analysis is the method to represent raw data as a percentage for better understanding of collected data. Percentage refers to a special kind of ratio.

#### **1.7 LIMITATIONS**

- \* Out of total number of population only 138 number of population were able to be included in the study .
- \*Sometimes, the busy schedules of people restricted the data collection
- \* Lack of previous studies in the research area
- \* The study was conducted based on the responses of the people so it may contain sampling errors
- \* Response of the people is not enough to get a inclusion and they were not co-operate enough for the study
- \* Constraints of a particular location

#### **1.8 KEYWORDS**

**Natural cosmetics:** They are products that contain only natural raw materials – mineral resources and ingredients of plant or animal origin. In addition, they should be obtained as a result of such processes as: filtration, extraction, drying, distillation, pressing, milling, lyophilization and sieving.

**Artificial cosmetics:** Synthetic chemicals are mainly present in the form of preservatives and fragrance agents. It is true that not all of the synthetic chemical present in the personal care products cause adverse health effects but these might have some dangerous ingredients that are being classified as carcinogen i.e. it might cause cancer in consumer's body.

**Niche Marketing:** It is the subset of the market on which a specific product is focused.

**Green washing:** It is the process of conveying a false impression or providing misleading information about how a company's products are more environmentally sound.

## **OBJECTIVES**

- \*To study about the current situation of cosmetic industry.
- \*To study the change in behavior and attitude of people towards new era beauty.
- \*To understand the change in demand of the people and the ability of the manufacturer to cater the demands.
- \*To study the threats faced by the industry and its shortcomings.
- \*To examine the consumer behavior towards cosmetics in Ernakulam district.
- \*To evaluate the effectiveness of cosmetic advertisements and other promotional tools among consumers in Ernakulam district.

## **CONTROVERSIES**

\*Ingredients used in cosmetics: The key ingredients present in most cosmetics include water emulsifiers, preservatives, thickness, moisturizers, colors and fragrances which can cause harmful effects on skin. Higher rates of cosmetics are also a reason that gave rise to controversies. It affects on normality of life causing cancer, birth defects, developmental and reproductive impairments.

## **1.9 CHAPTERISATION**

**Chapter 1-** Introduction: This chapter contains brief introduction of the topic, its scope and significance, problem statement, methodology adopted, keywords, limitations and chapterisation of the study.

**Chapter 2-** Review of Literature: This chapter includes review of earlier studies related to startups.

**Chapter 3-** Theoretical Framework: This chapter consists of concepts and, together with their definitions and reference to relevant scholarly literature, existing theory that is used for the particular study.

**Chapter 4 -** Data analysis and interpretations: In this chapter, datas are analysed and interpreted based on various observations from the respondents.

**Chapter 5-** Summary, Findings and Recommendations: It is the summary of the findings and recommendations. It contains the final conclusion of the study.

**CHAPTER 2**  
**LITERATURE REVIEW**

## **2.1 INTRODUCTION**

In any research, it is necessary to conduct a comprehensive literature survey to identify the scope and gaps of the study. Reviewing the previous study on the same topic will be helpful in giving ideas about the collection of data and its interpretation in the study. The research is conducted to know the impact of advertisements on cosmetic industry in Ernakulam city. Several researches were conducted in this topic by other researchers. Some of them are used in the study.

## **2.2 LITERATURE REVIEW**

**Kapil Mehtre, G Kavya Reddy, N Manaswini Reddy, Rohan Rajput,  
Nadiminti Rajesh Kumar**

**Journal of Marketing Vistas 11 (1), 20-35, 2021**

One of the major promoting tools is the product among the target population is the social media (SM). There are numerous factors that make social media platform as the best alternative compared to other promotion tools. Social media is cost effective and facilitates direct contact with the target customers. So, considering all the above aspects regarding the advertorial types social media has tremendous versatility in use and scope of study. Therefore the following study is being conducted to get substantial information regarding the effectiveness of various social media advertorial factors affecting cosmetics usage.

**Alubala A Agneta**

**University of Nairobi, 2018**

Facebook advertising is the new mantra for several brands. This study sought to investigate the effect of Facebook advertising on cosmetic sales and purchase by women in Kenya. Specifically, this investigation was directed by the following specific objectives: to determine the extent to which Facebook is used to advertise cosmetics products in Kenya, to determine the effect of Facebook advertising of cosmetics on attitude and belief of Kenyan women about the product and finally

examine the effect of Facebook advertising on the buying decision of Kenyan women on cosmetics product. The study recommended that growth of social media advertisement awareness to be created among women at all levels in order to increase their attitude of purchasing cosmetics through Facebook and improved the quality of advertisements on Facebook.

**R Sunderaraj**

**ICTACT Journal on Management Studies 4 (3), 800-807, 2018**

The objective of advertisement is to study the Impact of advertisement on buying behavior. It is found that there is no relationship between age of the respondents and level of impact of advertisement and there is no relationship between income and satisfaction with advertised product at the time of using.

**Alica Radtke says:**

The extensive literature review provides first evidence that the social media usage and cosmetics shopping behaviour differs between the three generations. These results show that advertisements are very useful in creating the awareness among the people but they are failed to build strong perceptions in the mind of consumers.

**Samar Fatima, Samreen Lodhi says:**

International journal of management sciences and business research, 2015 Advertisements are helpful in creating the awareness and perception among the customers of cosmetic products. This particular research was conducted on the 200 young male or female who use different brands of cosmetics to check the influence of advertisement on their buying behavior while creating the awareness and building the perceptions.

**TIGRAN HARUTYUNYAN, DONARA GRIGOR;** Social media advertising has become an important marketing channel for cosmetics companies. However, the social media advertising strategies of the main cosmetics manufacturers focus on the Millennials, although older customers generate a higher income. After a detailed analysis the results offer important implications for cosmetics companies: each generation has to be targeted differently when considering social media advertisement.

**Charbel Abi-Ramia says:**

Celebrity endorsement in the cosmetic industry has been vastly growing nowadays. This research focuses on the impact of involving a well-known celebrity in advertising campaigns, and the influence it has on consumers' purchase decisions.

**Jianan Wang**

**CONVERTER 2021 (7), 402-407, 2021**

With the rapid development of the economy and e-commerce, and the improvement of per capita consumption levels, more and more consumers are looking for personalized products, and consumers are increasingly demanding cosmetics. In addition, cosmetics consumption in emerging markets in the Asia-Pacific region, led by China, is proliferating. In 2013, China's cosmetics market surpassed Japan, becoming the second-largest cosmetics consumer in the world, accounting for 14% in 2019. In addition, the packaging cost of cosmetics accounts for a large part of its operating cost, and according to Estee Lauder's successful case, its reverse logistics system saves the company significant cost and improves its logistics efficiency. Therefore, by analyzing 211 sets of data of 21 domestic cosmetics enterprises in 10 years, this paper establishes an evaluation system of domestic cosmetics reverse logistics to effectively evaluate the logistics efficiency of China's cosmetics industry.

**M Adam, Nazish Hussain**

**British Journal of Marketing Studies 5 (3), 79-121, 2017:** This research is conducted on "impact of celebrity endorsement on consumer's buying behavior"; this behavior judges the purchase intention of consumers in the field of cosmetics, this research only focuses on the buying intention of the female segment. Celebrity endorsement is made up of four main pillars that are credibility, attractiveness, product matchup and lastly meaning transfer. All these four elements combine to make up celebrity endorsement. The findings of this research report shows that the element of credibility (which is composed of expertise and trustworthiness) of the celebrity positively affects consumers intention to purchase cosmetics. Celebrities also transfer meanings to the product that positively influences consumer's intention to buy cosmetics. This research also

shows that people get more attracted towards celebrity endorsed advertisements than the ones that doesn't have celebrities into them which ultimately leads them to recall the products (cosmetics) much easier because celebrities appeared into those advertisements.

### **Impact of using Celebrity and Non-Celebrity in an Advertisement**

**Akshay Khurana, Aman Aggarwal, Prabhsimran Kaur, Saloni Garg, Vipul Garg, Yash Aggarwal**

**Available at SSRN 3908567, 2020**

In the current time of data blast and media impact, promotions assume a significant part in changing insight or considering buyer. Celebrities have been utilized for a wide assortment of cosmetics brands. There is restricted investigation on buyer impression of celebrity versus non-celebrity for the cosmetics items. Advertisers use superstar support to help in simple brand review particularly during buying circumstances. Mainstream big names like cricketers and film stars are considered as god by their sweet hearts, however it is fundamental for advertisers to build up the connection between cosmetics item and superstar by thinking about the kind of cosmetics item. The work of individuals who partner with promotion office to sort out which kind of and which medium is the awesome the most plausible for their image. Organization needs to arrive at expected advantages of endorsers either big name or average person to guarantee that they are not in clash with an association's worth framework or liable to produce contrary exposure. The motivation behind this paper is to study the factors like credibility, expertise, attractiveness, decision making behavior to study impact of celebrity and non-celebrity endorsers in cosmetics commercials on buying goals of clients.

**Morag Martin's history of the cosmetic industry in France examines the evolution of popular tastes and standards of beauty during the late 18th and early 19th centuries.**

Martin introduces readers to the social and economic world of cosmetic production and consumption, recounts criticisms against the use of cosmetics from

a variety of voices, and examines how producers and retailers responded to quickly evolving fashions the newfound popularity of cosmetics raised serious questions. Critics—from radical philosophes to medical professionals—complained that the use of cosmetics was a threat to social morals and questioned the healthfulness of products that contained arsenic, mercury, and lead. Cosmetic producers embraced these withering criticisms, though, skillfully addressing these concerns in their marketing campaigns, reassuring consumers of the moral and physical safety of their products

### **Impact of advertising strategies on the cognitive and behavioral component of attitude of women consumers**

**Harminderjit Kaur, Bikramjit Singh Hundal**

**Journal of Asia Business Studies, 2017**

The literature was scant in examining the influence of the mostly used traits on the purchase behaviour and the switching behaviour of consumers.

The results indicate that repeated exposure, comparison of products and sexual appeals has a significant impact on the mind of consumers which determines the influence of advertising tactics. Due to the changing behaviour and the attitude, income level and media, the response of the consumers in the present study may not be relevant in the future period.

### **A Study on Role of Television Advertisement on Cosmetic Purchase Among Youth**

**Radha Vyas, Gautama Parmar**

**Global Journal of Research in Management 9 (1), 29, 2019**

Television advertisements are the one of the important source for product promotion, with increase in the penetration of television and DTH (direct to home), it is become easy tool to reach to large number of population. Television advertising in India is the most popular advertising option. The present study was conducted to investigate role of television advertisement on cosmetic purchase among youth.

## **Facebook advertising as a marketing tool: Examining the influence on female cosmetic purchasing behaviour**

**Barween Hikmat Al Kurdi, Muhammad Turki Alshurideh**

**International Journal of Online Marketing (IJOM) 11 (2), 52-74, 2021**

Social media platforms are widely used these days for the advertising and marketing of products. Facebook is considered one of the main social media platforms used by users these days. Currently, there are limited studies investigating the use of Facebook as an advertising communication platform, especially for the purchase of cosmetic products. This study targeted female consumers to ascertain to what extent Facebook advertising influenced their cosmetic buying behaviour through using a set of factors that were selected, namely, advertisement quality, advertisement design, message strength, advertisement repetitiveness, and message content. Smart PLS was used to assess the study model and to test the study's hypotheses. The study found that the main factors affecting consumer behaviour were advertisement quality and advertisement repetitiveness. The paper discusses the study's findings by presenting a set of implications and making recommendations.

**CHAPTER 3**  
**THEORITICAL FRAME WORK**

## **3.1 DEFINITION**

### **COSMETICS**

Cosmetics are defined as “items with mild action on the human body for the purpose of cleaning, beautifying, adding to the attractiveness, altering the appearance, or keeping or promoting the skin or hair in good condition”. Cosmetics are constituted mixtures of chemical compounds derived from either natural sources or synthetically created ones.

#### **3.1.1 COSMETIC INDUSTRY**

The cosmetic industry describes the industry that manufactures and distributes cosmetic products. These include color cosmetics, like foundation and mascara, skincare such as moisturizers and cleansers, haircare such as shampoos, conditioners and hair colors, and toiletries such as bubble bath and soap. The largest cosmetic companies are Johnson & Johnson, L'Oreal Paris, Gillette, Neutrogena, Nivea and Chanel, Inc. The market volume of the cosmetics industry in Europe and the United States is about EUR €70b per year, according to a 2005 publication

#### **3.1.2 ADVERTISEMENTS**

An advertisement (often shortened to advert or ad) is the promotion of a product, brand or service to a viewership in order to attract interest, engagement and sales. Advertisements come in many forms, from copy to interactive video. An advertisement is different from other types of marketing because it is paid for, and because the creator of an advert has total control over the content and marketplace.

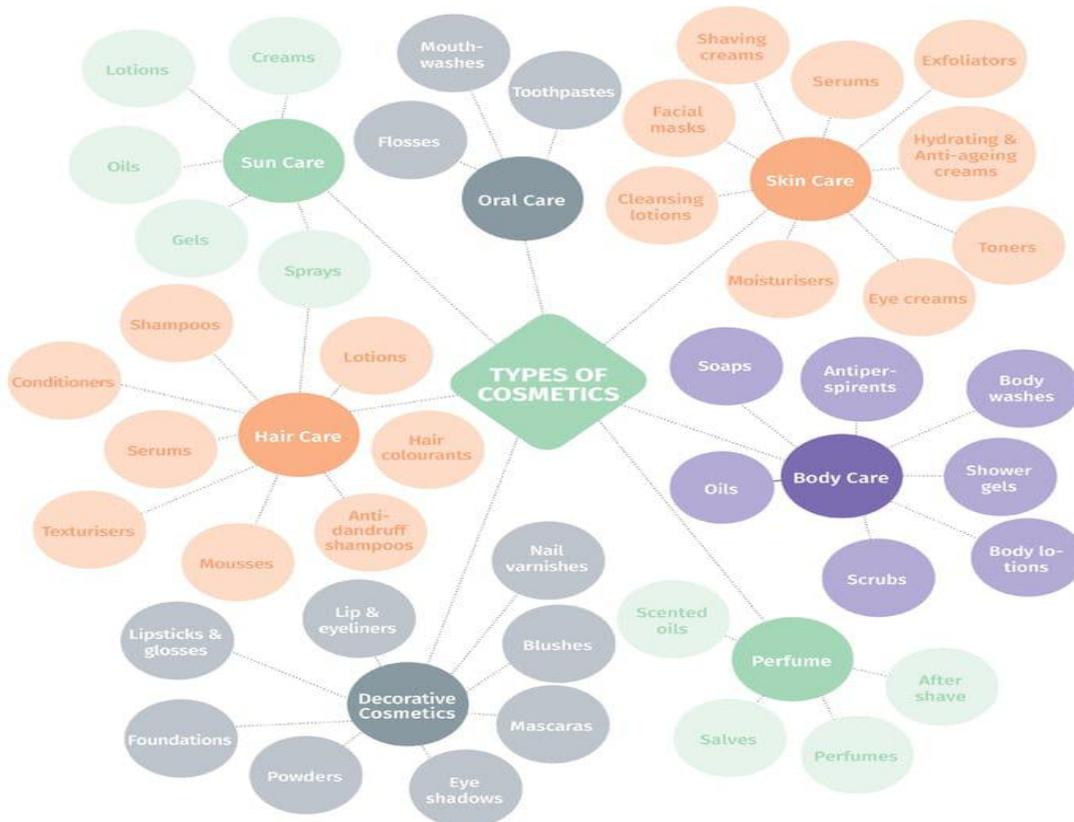
#### **3.1.3 COSMETIC ADVERTISEMENTS**

Advertisement of beauty products by the cosmetics industry through Cosmetic advertising is the promotion of cosmetics through a variety of media. The advertising campaigns are usually aimed at women wishing to improve their appearance, commonly to increase physical attractiveness and reduce the signs of ageing.

### 3.2 Category of cosmetics

Specifically, there are seven categories of cosmetics and personal care products - oral care, skin care, sun care, hair care, decorative cosmetics, body care and perfumes. Further examples are listed in the diagram below.

COSMETICS ARE AN IMPORTANT PART OF PEOPLE'S EVERYDAY LIFE



### 3.3 COSMETIC ADVERTISEMENTS;

Cosmetics advertising works by using several different techniques that encourage its target market to buy specific cosmetics and skincare products. The effectiveness of the advertising typically depends on convincing people that the product being advertised actually can improve their appearance and that, the audience, are in need of appearance enhancement.

Depending on regulations regarding advertising claims in the jurisdiction where the advertising is featured, the cosmetics manufacturers and their advertising producers may use several methods to persuade individuals to make a purchase.

These can include "angel dusting" and medical claims, as well as promises of general lifestyle enhancement.

### **3.3.1 SOME OF THE STRATEGIES USED FOR COSMETICS ADVERTISEMENTS PURPOSE ARE:**

\*Use Brand Ambassadors to Promote Products

\*Provide How-to Videos and Images

This is a lateral, "we can help you" ad strategy. Use it to express the benefit of your product or service.

\*"Before and After" Strategy, Use this strategy to show life without the product and then life with the product.

\*Empathy Strategy, Use this strategy to develop the audience's empathy towards a targeted audience.

\*Testimonial Strategy. This approach leverages a famous or non-famous consumer to support the quality or benefits of a brand's product.

\* "Demonstration" Strategy. This type of ad strategy focuses solely on ways to explicitly show the product benefit via a "demonstration.

\* "Negative to Positive" Strategy. Use this strategy to turn a boring or negative perception about a product, brand, or market into a positive one by adding an extra or clever argument.

\* "Competitive" or "Comparison" Strategy. Use this strategy to make either an overt or subtle reference to any differences between your product and a competitor's product.

THESE ADVERTISEMENT TECHNIQUES CREATES A SENSE OF INSECURITY AND A CRAVING FOR CHANGE AND IMPROVEMENT IN THE MINDS OF TARGET GROUPS.

THIS IS THE MAJOR IMPACT AND RESULT OF ADVERTISEMENTS IN THE FIELD OF HEALTH AND BEAUTY.

### **3.3.2 How Cosmetics Companies are Utilizing Paid Social Media to Grow Their Brands?**

Cosmetics brands look to build brand loyalty and awareness through key representatives whether in the entertainment industry or through rising social media influencers. Certain icons ruling social media today bring key consumer bases closer to the brands they associate with.

Audience targeting strategies across various social media channels help build brand awareness and help lead to a growing number of sales.

### **3.3.3 Cosmetics company's exploited by ad agencies**

1. **Excessive Thinness:** Unhealthy body images in advertising regardless of whether they are used to sell weigh-loss products or something else project an unrealistic image of women's body weight, and according to HealthyPlace.com, can contribute to the development of anorexia, bulimia and other serious eating disorders in women.

2. **Sexual Exploitation:** Sexuality can be a powerful motivator, and many advertisements use sexually explicit imagery to help sell their products. While sexiness in and of itself isn't harmful, reports from Jean Kilbourne and other activists cite an alarming trend of objectification and exploitation in advertising. Women often appear wanton, passive and child-like in advertisements, sending a message that such qualities are normal and even desirable in women. Even more disturbing, men receive the message that women should act submissive and wanton, and come to expect that in their relationships with the opposite sex.

3. **Ageism:** Along with body weight issues, advertisements often depict very young and impossibly beautiful women. According to the "New Yorker" and others, advertisers covert young demographics since they often have more spending money. That can result in ageism, particularly against women, who may perceive

images of unattainable youth as ideals to strive for. As women get older, they feel pressure to look younger, ignoring the natural beauty of a 50- or 60-year-old body in an impossible effort to retain a 20- or 30-year-old one.

**4. Consequence-Free Fantasies:** Many ads present consequence-free scenarios with beautiful people enjoying, say, junk food or video games. Such images imply that one can live a healthy lifestyle by purchasing such products, when in truth, the opposite is often true. When coupled with other images presented to women -- thinness, youth and sexuality -- it creates a dangerous impression that such a lifestyle is attainable, when in fact it can often result in low-self esteem and damaging habits.

### **3.4 LEGAL REGULATIONS ON ADVERTISEMENTS OF COSMETICS BY COSMETIC INDUSTRY**

As mentioned above, it is legally possible to advertise cosmetic products. However, as is the case with the advertisement of any product, it is particularly subject to restrictions in the following regulations:

#### **Law on the Protection of the Consumer Numbered 6502**

#### **Commercial Advertising and Unfair Commercial Practices Regulation**

#### **Cosmetics Law**

#### **Cosmetic Regulation**

#### **Regulation on Health Declaration and Health Declaration of the Products Offered for Sale**

Especially as mentioned in the Regulation on Health Declaration and Health Declaration of the Products Offered for Sale, in the advertisement of cosmetic products;

Phrases or implications for treating or preventing any disease, helping to cure, diagnosing, or correcting, regulating or modifying a physiological function,

Statements that claim or imply that, as a result of pharmacological, immunological or metabolic effects, they renew, correct or change physiological functions,

Statements referring to the effect of human medicinal product cannot be used.

Ads for beauty-enhancing products seem to make consumers feel that their current attractiveness levels are different from what they would ideally be. "Consumers seem to 'compare' themselves to the product images in advertisements, even though the advertisement does not include a human model," the authors write.

**"Exposure to beauty-enhancing products in advertisements lowered consumers'"**self-evaluations, in much the same way as exposure to thin and attractive models in advertisements has been found to lower self-evaluations," the authors conclude.

### **3.5 BENEFITS OF COSMETICS;**

Cosmetics are products designed to cleanse, protect and change the appearance of external parts of our bodies. The key ingredients present in most cosmetics include water, emulsifiers, preservatives, thickeners, moisturizers, colors and fragrances.

**1 Emotional /mental benefits:** Beyond physical health, cosmetics can help to improve our mood, enhance our appearance and boost our self-esteem. They can also help to exhibit personal style and, as such, are an important means of social expression. In a study by FEBEA, over 60% of respondents claimed that cosmetics have a positive impact on well-being, image, self-confidence and mood, with a large proportion (+40%) also identifying benefits in terms of social life, love life, family life, professional life and health.

## 2 **Functional benefits:**

Cosmetics contribute to wellbeing and healthy lifestyles. Our hands carry pathogens from contaminated sources; so simple tasks such as washing hands with soap can help prevent serious illness. Indeed, multiple studies have shown that the leading causes of child mortality in developing countries, diarrhoea and respiratory infections, can be prevented by hand washing with soap. The use of toothpaste, particularly when containing fluoride, reduces the prevalence of dental caries. Toothpaste reduces plaque and tartar, which can lead to tooth damage and gum disease.

## 3 **Economic benefits:**

Beyond health, there are economic advantages to dental care: there is strong evidence that the benefits of preventing tooth decay far exceed the costs of treatment. Indeed, if we assume that, without toothpaste, total expenditure on oral health would be 5% higher, the total benefits of using toothpaste (in terms of avoided costs) would be approximately €26.5 billion by 2020-21.

## 4 **Innovational benefits:**

Innovation in our industry is never static. Constantly evolving consumer expectations with regards to product attributes and safety means new products or iterations are always in development, increasingly geared towards personalized solutions for individual skin and hair types, for instance.

### **3.6 BROAD HEADS UNDER COSMETICS AND ROLE OF ADS IN THEIR PROMOTION**

Basically the cosmetics are of **two types ;natural and synthetic**

#### **3.6.1 Natural cosmetics? their composition?**

Natural cosmetics are products that contain only natural raw materials – mineral resources and ingredients of plant or animal origin. In addition, they should be obtained as a result of such processes as: filtration, extraction, drying, distillation,

pressing, milling, lyophilization and sieving. Obtaining natural raw materials with the use of the aforementioned physical methods, which process the material to a small extent, allows to preserve the nutritional values, vitamins and beneficial substances found in the raw materials. Ingredients of animal origin can be used to produce natural cosmetics only when doing so is not harmful to the health or life of animals. Such raw materials are, among others, honey, propolis, lanolin and milk. In the production of natural cosmetics, it is also allowed to use cosmetic raw materials obtained in biotechnological processes, i.e. microbiological or enzymatic. Hyaluronic acid, among others, is obtained this way.

In addition to raw materials of plant and animal origin, in natural cosmetics you can also find different kind of sea components (e.g. algae) essential oils, inorganic salts and oxides derived from minerals, natural emollients, dyes, emulsifiers (e.g. sucrose esters), naturally occurring active substances and preservatives (benzoic acid and its salts, benzyl acid, salicylic acid).

### **3.6.2 Role of ads in promotion of natural cosmetics ;**

The ads in social medias and other printed medias helps to popularize the effects of these natural cosmetics by describing its benefits and ingredients which is from nature .They usually make use of natural backgrounds for these purpose and also the support of evidences from Vedas of ancient times are also taken by the ad agencies for the testimonials .Usually people have tendency to believe in Vedic medicinal evidence which is exploited by the advertising companies to market their client products eg. fair and lovely Ayurvedic skin care products ads ,Patanjali natural skin care ads

### **3.6.3 Synthetic or artificial cosmetics and its components ?**

Artificial cosmetics

Synthetic chemicals are mainly present in the form of preservatives and fragrance agents. It is true that not all of the synthetic chemicals present in the personal care products cause adverse health effects but these might have some dangerous ingredients that are being classified as carcinogen i.e. it might cause cancer in

consumer's body .Some of the ingredients can be neurotoxins, reproductive toxins that have been proven to affect the brain development and reproduction. Some ingredients can be cytotoxic that cause impairment to the cell existence which might result in immediate cell death by losing cell membrane integrity (necrosis) or the cell might undergo apoptosis. With all benefits, cosmetics and personal care products come as a parcel of never ending list of allergies. It can be skin allergy (permanent discoloration of skin), breathing allergy (damage to nose and nasal passage) and hair allergy (redness in scalp, hair fall, excessive dandruff, thinning of hair). Some agents can be labelled as hypo-allergic, even though they might have potential carcinogenic property. Many studies have been conducted on patient's cosmetic products to study the undesirable effects caused by them to human beings.

#### **3.6.4 Role of ads in promotion of artificial ads**

Now a days people always have an argue to use chemicals for more effective result in a short span of time .This tendency of human beings are used by the ad agencies and cosmetics in dusty for their marketing .Usually modern and color full ads are used for this purpose to attract target group by using celebrities for promotion. These types of ads have a secret nature as it tries to hide some chemical ingredients and their quantities from the view of people who use them. The market demand for chemical cosmetics are more as it gives more shine and impact in less time. More over most of the makeup products are of artificial in nature.

#### **3.6.5 Natural v/s synthetic cosmetics**

As per the latest survey, women in the age group of 35-54 are increasingly inclined to go green with 69% favoring natural products. 6 in 10 women now read the product labels prior to purchase! The question is whether we are right in being wary of chemicals in our skin care. Are we blindly buying into natural brands or there is some scientific reason behind it? – everything is chemical and there is nothing like chemical-free! Everything including water is chemical. Better term is “synthetic” which generally means that a natural derived ingredient has been molecularly replicated in a lab to provide the same effect.

Natural products are not always good! They might not be ethically sourced or organic. Mineral oil is natural but we would not like it on our skin and non-natural palm oil has a devastating impact on the wild life, hence again not preferable. In a nut-shell, it is not enough that your natural organic skincare product is “natural” but it has to be something more and that is “organic”.

### **3.7 MOTIVATION FACTORS;**

Factors that motivates people to purchase cosmetics

**1.CONSUMER BUYING BEHVAIOUR:** It is a decision process as well as an attitude of the people involved in purchasing and using products. Consumers make purchase decisions for buying small as well as large products. After recognizing a need or a want, consumers begin searching for products or services that fit their requirements.

**2.Brand:** The Brand is considered as a name and symbol of a company. Brand plays an important role in creating a positive image among customers. Brand name has the ability to create loyal customers as well as to retain market share of the company. Loyal customers always remain faithful or loyal to a brand, they repurchase it and through word of mouth they recommend it to others.

**3.Quality:** High quality cosmetics help to build and maintain the confidence of targeted customers and convince them to purchase it. Customers are concerned about the quality of cosmetic products before deciding to purchase them. To remain competitive among the competitor’s cosmetic firms should increase and concentrate more on improving the quality of products.

**4.Packaging:** product labeling and packaging is a representative of manufacturer and act as the carrier of advertising messages and company slogan. Yang (2004), states that proper packaging create distinctiveness among the companies. Companies know about the needs of proper packaging and how a slight change in packaging results in their profitability. Therefore, now companies are completely focused on different aspects of packaging and thereby to reap high productivity and profitability.

**5.Price:** Hermann et al (2007) research shows a relation between price and customer satisfaction. For purchasing a product price plays an important part and has a strong impact on the satisfaction level of customers. In another research, Lee et al (2010) founded the relation between price and its impact on purchase decision. Khraim (2011), state that for a loyal customer price doesn't matter and it effect on the purchase decision of customers.

**6.Advertising:** Advertising can be considered as a subset of marketing mix (4P) that are place, promotion, price, and product. For promotion of a product, one of the main strategies used is advertising. In order to make awareness of a product in the mind of potential customers, advertising is an important tool and it influences the customer's decision to buy a product. If advertising is attractive, then customers pays more focus on that and thereby creates a feeling towards the product and that leads to the way of brand promotion. Those who have loyal feeling show a positive attitude towards a brand.

### **3.8 Gender and use of cosmetics**

It's no secret women spend a lot of time and money on beauty. Today and AOL survey found that women spend about 55 minutes a day on beauty prep—that's almost two full weeks a year. Financially, a survey from beauty e-tailorSkin store estimated that women spend \$300,000 in their lifetime just on their faces (although they surveyed their own customers who are probably more product-obsessed than the average woman).The concept of beauty has always been ever-changing. As society continues to evolve, we have seen many trends come and go. Now a days the concept of beauty care is not limited to women .Today men are also very conscious about their looks and also never back out from grooming their self to look much better.

Many industries have come under scrutiny for their uneven gender representation in upper management, but the beauty industry has largely been absent from this conversation.

However, one increasing market demand seems to stay: the need for gender equality in the beauty industry. Men are known to have used cosmetics in Roman times, although it was frowned upon by society. Aside from traditional use and use in the arts, research shows that young men who are between 18–34 years old are more likely to use cosmetics. During the reign of Queen Elizabeth I, cosmetics were very popular among men, as they valued ghostly, powdered skin.[1] Male cosmetics were originally targeted towards homosexual men, however, market research revealed that only a third of male cosmetic consumers were gay. Some men use beauty products to cover perceived flaws on their faces, such as acne marks and freckles. Additionally, some men use cosmetics to boost their physical appearance. Makeup has evolved through out the years . Men have found a way to use makeup very simply and organically, to create everyday looks for themselves without any stigmas behind it. The packaging of male cosmetics is generally simple. The colors are mainly blue, green, grey, white or black. Compared to women's cosmetics, there are fewer bright colors such as pink, red and purple . Male cosmetics include all cosmetic products intended for use by men, such as makeup, skincare products, hair care products, body care products, sun care products, perfumes, and other decorative cosmetics . The global beauty industry is worth over \$500 billion, with a steadily growing 8% year-to-date (as of July 2018) in the men's grooming market.

Even the American Academy of Dermatology backs up the market statistics by stating that the basics of daily skincare for men and women are the same.

From female perspective , cosmetics and makeup are the essential part of their day to day life. it is also a custom to believe that a women should wear cosmetics to be a women. Human beauty is a cultural construct that involves facial and body attractiveness. One modification to appearance that is pervasive is the use of cosmetics. This paper explores the consumption of cosmetic products among women as a form of ritualized social comparison with others and as a means of giving women individual control. First, the findings of an in-store exploratory study of 77 female cosmetics consumers are related to social comparison and the process of identity construction. Second, the results of an in-depth observation interview study of 12 female cosmetics consumers are used to elucidate the ritual

and meaning of cosmetics consumption and its relation to individual identity. There is strong evidence to suggest that the ritualized acts of purchasing and using cosmetics serve to enhance personal identity construction and to give women a greater sense of cultural power and social agency in a postmodern world.

### **3.8.1 Role of ads in balancing gender in cosmetics ;**

In today's world gender discrimination is a serious taboo . The advertisement agencies cautiously strive toward making their ads gender friendly. Now a days the gender gap in beauty industry especially cosmetics and make up are growing down. Mealso interested in make up and beauty care ,hence ads for men grooming is of serious importance .brands like garnier,nivea,fair and handsome [from fair and lovely ],beardo,the ustraa are some of the most commonly advertised cosmetics brands for men grooming. Many unisex brands and products are also in the market. Advertisements play a major role in increasing the demand for these brand products .

## **3.9 ATTITUDE OLD PEOPLE TOWARDS CONSUMPTION OF COSMETICS**

**According to Fishbein and Aizen (1975),** refers attitude, "it is a learned predisposition to react to favourable or unfavourable manner with respect to a given object.

**"Krech (1962), states,** "an attitude as a person's favorable or unfavorable evaluations, emotions feelings &action ways some objects or ideas, which leads to behave in a fairly constant way towards similar objects.

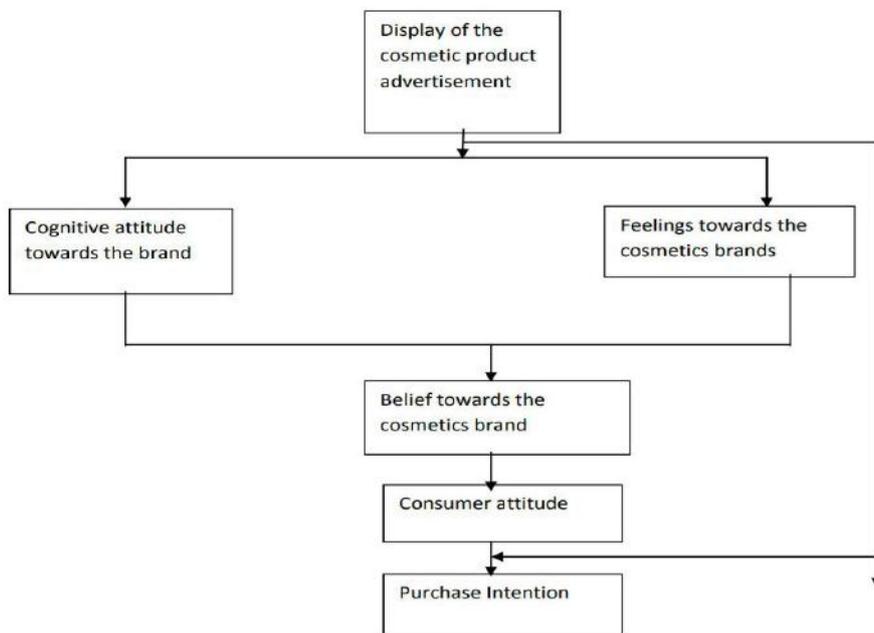
Consumer buying behavior mainly depends on his attitudes.

Specifically, the attitudes of consumers can have a critical role on beauty cosmetics buying behavior. The middle aged people have positive attitude towards beauty cosmetic products compared to young aged people. Through the beauty cosmetic products they try to keep young or give young appearance. And, Middle aged people almost they settled in their life. So they have the self spending power

to buy beauty cosmetics products. Hence, they have positive attitude towards beauty cosmetic products compared to young aged people.

Monthly family income does not have any influence on the attitude towards beauty cosmetic products. All categories of people want to maintain their self image regarding physical appearance among the society. Regarding occupation and marital status, housewives and married people have positive degree of attitude towards beauty cosmetic products. Compared to privately employed people and those doing business, house wives having time to take care about themselves. Married people have crossed certain age limit, for that reason they need to take care about their physical appearance. Now-a-days married woman have conscious about skin care and personal care.

Hence, married people have positive attitude towards beauty cosmetic products compared to unmarried people. Apart from the above factors, consumers are having more conscious on quality, money value and brands of a beauty products.



A framework: Consumer attitude towards advertisement leads to purchase intention

## **3.10 PERSUASION AND CONTROVERSIES ON COSMETIC INDUSTRY AND ADVERTISEMENTS**

### **3.10.1 PERSUASIONS /ADVANTAGES**

The obsession with youth and beauty sends thousands of people to stores searching for a quick fix. Looking into cosmetics retailers, teenagers are trying on makeup to make them look more mature while middle-aged women are sourcing for the best anti-aging cream for her skin. The beauty industry relies on all kinds of persuasion techniques to sell products. Brands like to rely on celebrities to sell their beauty products. "If the product is good enough for her, it's good enough for me." This philosophy is often the impetus behind advertisements for makeup, skin products, and hair products. This causes consumers spend extra money on items that might not be suitable or workable on themselves. Celebrities can also help to build brand awareness through promoting products as consumers believe the product is used by the celebrity and this can affect consumer's decision making. People not only expect to look good, but feel good too. Consumers expect the products will make them look better than they did without the products or even believe these products will make them more appealing like some of the models who advertise the products. Therefore, Advertisers use those buzzwords to appeal to a wider range of consumers and to wider their marketing strategy by targeting new generations and gain the attention of diverse consumers.

### **3.10.2 SOME OF THE IMPORTANT ADVANTAGES OR BENEFITS OF COSMETIC ADVERTISEMENTS**

- 1. Target the relevant audience:** When you start a marketing strategy, it is always good to know who you want to reach. Without a clear target group, marketing strategy is not going to be a success.
- 2. Creates brand recognition:** Achieving brand recognition is one of the most significant promoting goals for any business.. Fortunately, social media allows a simple yet powerful brand building. While comparing with traditional marketing,

social media marketing has more benefits as it can get your image in front of customers substantially and effectively.

**3. Influencers strength:** If you want to set up social media marketing services in the beauty industry, it is good to keep part of your budget free for collaborations. It is clear that more and more brands are looking for influencers themselves. For example, the American makeup brand Too Faced worked with the Indian YouTuber Tutorials on a makeup palette that was sold out in no time.

**4. Niche Marketing:** Owning a small cosmetics business makes it easier for you to target specific niches, compared to larger companies that may feel the profit margin is too low to sell specialty cosmetics. For instance, you may want to focus on creating a line of makeup and skincare products for people with allergies to ingredients commonly found in most cosmetics. This helps you build a brand people remember when they need more cosmetic products or want to refer people who experience similar problems and need these types of special cosmetic products.

**5. Personalized Customer Service:** Providing personalized customer service is a key advantage of owning a small cosmetics company. For instance, offering instruction on how to use the cosmetics and helping customers choose the best ones to match their skin type and tone gives you a way to build customer loyalty.

**6. Free Services:** If you want to create your own product line, cosmetic labs may help you for free in hopes of getting the order when you're ready to manufacture your line of products, according to Inc. Plus, new cosmetic lines that offer unique features, such as using organic products or being designed by celebrities, tend to get free coverage since publications and online sites are always looking for the latest and greatest in the cosmetic world.

**7. Easy to Showcase:** Unlike some businesses, the cosmetic industry is all about how a product looks and works on someone's skin. The advantage of this concept is that it's easy to showcase the product in multiple ways, such as in its product packaging as well as in before and after photos showing people who use it. These

images can appear on your website, social networking sites, in print and in television ads.

### 3.10.3 CRITICISM/DISADVANTAGES

Many campaigns have come under fire through their alleged use of pseudoscience and their promotion of unrealistic goals. Moreover, many campaigns are accused of inducing harmful habits on people (such as bulimia and anorexia), and leading to destructive plastic surgery practices. In addition, cosmetic advertising is often accused of excessively using photo manipulation to enhance the appearance of models, instead of using the cosmetics themselves, creating an unrealistic image of the product's benefits, Este Lauder for example .Cosmetics are a major expenditure for many women, with the cosmetics industry grossing around 7 billion dollars a year, according to a 2008 YWCA report. Cosmetic retailers design advertising to alter women's attitudes toward cosmetics, encouraging them to buy more products. Many advertisers shape this attitude by encouraging women to feel dissatisfied with their appearance. The flawlessness of advertising woman is, in fact, an illusion created by makeup artists, photographers, and photo retoucher. Each image is painstakingly worked over: teeth and eyeballs are bleached white; blemishes, wrinkles, and stray hairs are airbrushed away. Media images convey normative information as to what an attractive body looks like that prompts women to evaluate their own body against this normative standard. More over; it adds to the cost of production and price of product which leads to price war and increase a unhealthy competition in market.

Deceptive Advertisements creates a bad image about the product and also affects the sales and distribution of the product which in turn leads to the collapse of the product and brand.

Leads to Unequal Competition and Creates a Monopolistic Market creates unnecessary demand for useless products.

### **3.11 Cosmetic industry today**

The impact of the COVID-19 pandemic on consumers' perspectives of beauty and individual cosmetic product were huge. Consumers' perspectives of beauty are impacted by a pandemic. Global consumers perceive skincare as an important aspect during the pandemic, while the importance of makeup fell after the outbreak. The awareness of skincare and makeup products has changed. The spread of the pandemic (SOP) has a positive impact on skincare products, but a negative impact on makeup products, except for eye makeup products, which was positive. Finally, the SOP was not significant in terms of consumers' interest in masks. Fifth, interest in masks showed a positive relationship with interest in skincare products, such as cleansing products, while a negative relationship was observed with interest in makeup products. Overall, this study concludes that pandemics certainly have an impact on global consumers' perspectives. As a pandemic spread, interest in skincare products increases, while interest in makeup products decreases. This study has academic significance in that it investigates the effects of consumption of cosmetic products during the stay-at-home rules. It can be used as standard information for setting marketing strategies in pandemic-like situations in the future.

#### **3.11.1 OPPURTUNITIES AND THREATS FACED BY THE INDUSTRY IN THE NEW ERA;**

##### **Opportunities ;**

**Increasing Availability of Small and Affordable Products:** The desire to look fresh and alluring throughout the day among the Indian consumers is propelling the companies to introduce pocket-friendly sizes of their products, which can be easily carried everywhere. Currently, the market is flush with products in small packs or sachets. The launch of such products is important in small cities and rural areas. Availability of these pocket-friendly sizes has been turning out to be a major profit cherner for the companies. Premium hair shampoo and conditioner brand such as TRESemmé has launched its range of shampoos and conditioners in sachets to increase the affordability of its products as well as to capture a broader set of women consumers.

**Rising Women Participation in Workforce:** Young working women in the country are willing to spend on cosmetics to enhance their beauty. Moreover, the launch of new and innovative women's cosmetics is luring them to purchase more products.

**Introduction of Innovative Products with Multiple Benefits:** Products with multiple benefits such as anti-ageing, moisturizing and SPF protection are gaining prominence in the cosmetics industry. Foundation creams that are being introduced in the market include SPF protection moisturizers, are oil-free and do not clog pores. Such products with multiple benefits are in high demand, especially among young women.

**Changing Distribution Strategies:** Increasing internet penetration is driving sales of cosmetic products through online sales channels over traditional retail formats. Rapid urbanization and changing lifestyles are encouraging consumers to purchase cosmetic products through online sales channels instead of traditional formats

**Growing Consumer Interest in Cosmeceutical Products:** The raw materials used in the production of organic cosmetics are also produced without using any chemicals or pesticides, they are perceived to be completely safe by the consumers. Hence, consumers are switching to organic, natural cosmetics and Cosmeceuticals that are free from harmful chemicals. Moreover, the increasing number of health issues due to chemicals in cosmetics and rising levels of pollution in major cities is driving people to switch to cosmeceuticals, organic or natural cosmetics instead of standard chemical based cosmetics.

These are some of the opportunities or benefits for cosmetics industry to sustain in the market .

### **Threats or challenges**

The beauty industry shapes girls, boys, men and women around the world. From putting make-up on Barbie at a young age to picking out the blush you want to wear at your wedding, the beauty industry is in every stage of life. There are problems in the beauty industry ranging from representation to idealization and

lack of regulation. Lack of regulation is the culprit of many of the problems with the beauty industry, especially in the United States. According to the FDA, “The law does not require cosmetic products and ingredients, other than color additives, to have FDA approval before they go on the market.

### **Green washing**

Green washing occurs when a company spends more money marketing their brand and products as environmentally friendly and safe rather than spending money to actually make their brand and products meet these qualifications.

### **Chemical Catastrophe**

Chemical ridden products, vague ingredient labels, and hidden chemicals are three of the major problems with the beauty industry. All these problems stem from one major source: lack of regulation.

### **Marketing Myths**

Bloggers, TikTok, beauty magazines, influencers, and companies are all responsible for marketing myths in the industry. And trust me, there are a lot of them. Unless you have a medical background or are extremely knowledgeable in beauty and health, the myths are hard to debunk on your own.

Difficult Access to Beauty Education Learning about the ingredients in your products is quite the hassle. It involves deep dives on google and lots of clicking to find trustworthy sources.

### **Idealized Beauty and Poor Mental Health;**

Social media, beauty stores, and magazines showcase unrealistic images of beauty in their advertising and marketing efforts. Professional photo shoots and editing portray beauty as flawless, airbrushed skin with no blemishes or scars. Thick lips, thick lashes and thick eyebrows. Voluminous shiny hair. I could go on and on.

## **Revolving Body Expectations**

It feels like there are new diets and new workouts to partake in every year in order to maintain revolving body expectations. One year society says you need thick thighs and a big butt, the next society says you need to be stick thin. Working to maintain these expectations is draining and potentially dangerous to our nutrition and mental health.

All of these are challenges faced by the beauty industry in the new era. The perception about beauty had changed a lot in this century. People believe to be natural and ordinary than artificial .The whole concept of fairness is changed and the myth about fairness or whitening is beauty is thrown out from the industry. The industry is showcasing more innovation in thoughts and products, but the fact that competition in the area is high and vigorous cannot be ignored. Thus advertisements and promotion techniques plays a vital role.

**CHAPTER 4**  
**DATA ANALYSIS**

## DATA ANALYSIS AND INTERPRETATION

Analysis and interpretation of data is necessary to understand the problems and arrive at a conclusion. Data analysis transforms the data into information for the research. Data interpretation draws the conclusion from the data collected.

This chapter includes the response obtained from 132 respondents of Ernakulam city, in the form of tables, graphs, and diagrams.

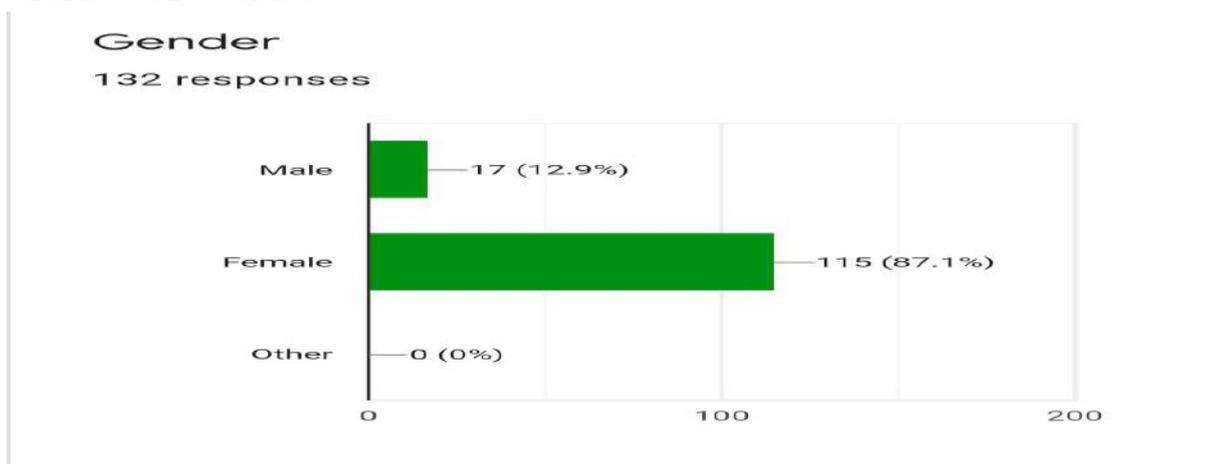
**TABLE 4.1**

### GENDER OF RESPONDENTS

GENDER	NO.OF RESPONDENTS	PERCENTAGE
MALE	17	12.9
FEMALE	115	87.1
TOTAL	132	100

**FIGURE 4.1**

### GENDER OF RESPONDENTS



Inference : Among the respondents 12.9 % and 87.1% are female .From the above table it is clear that majority of the respondents are female.

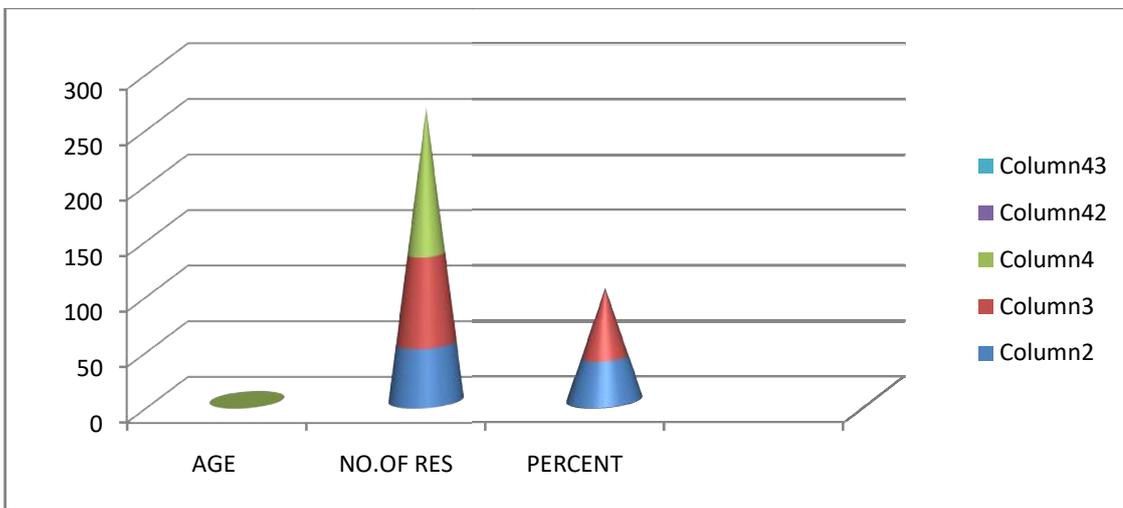
**TABLE 4.2**

**AGE OF RESPONDENTS**

AGE RANGE	NO.OF RESPONDENTS	PERCENTAGE
BELOW 25	52	39.39
ABOVE 25	80	60.60
TOTAL	132	100

**FIGURE 4.2**

**AGE OF RESPONDENTS**



Inference : Classification on the basis of age of respondents shows that majority of the respondents;60.60% belongs to the age group of above 25 that is youth. Remaining 39.39 % of the respondents belongs to the age group below 25 ; that is teenage.

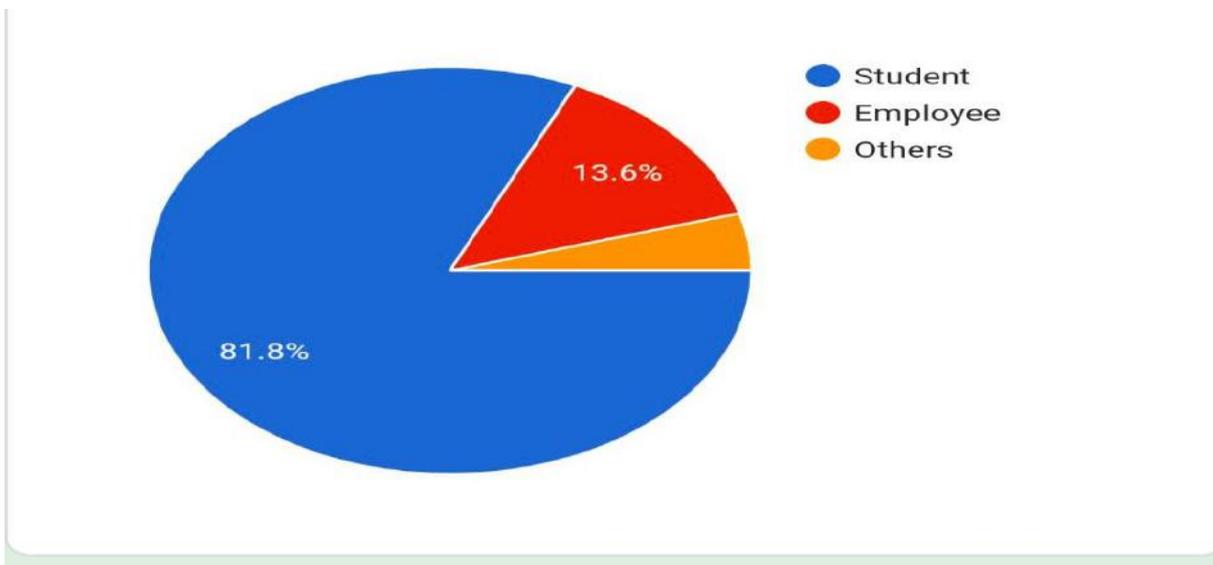
**TABLE 4.3**

**CATEGORY OF RESPONDENTS**

CATEGORY	NO.OF RESPONDENTS	PERCENTAGE
STUDENTS	108	81.8
EMPLOYEES	18	13.6
OTHERS	6	4.6
TOTAL	132	100

**FIGURE 4.3**

**CATEGORY OF RESPONDENTS**



Inference :Among the respondents almost major part ;81.8% are students ,and 13.6% are employed ones and .4.6% are unemployed or others.Thus majority of the respondents are students.

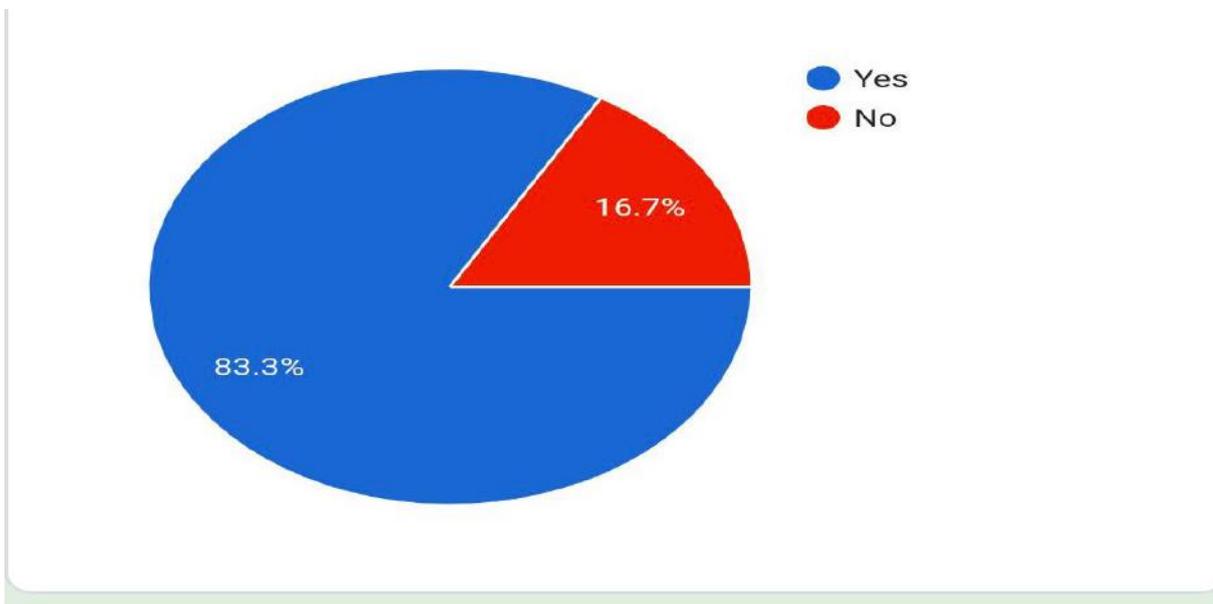
**TABLE 4.4**

**NO.OF RESPONDENTS USING COSMETIC PRODUCTS**

USING	110	83.3
NOT USING	22	16.7
TOTAL	132	100

**FIGURE 4.4**

**NO.OF RESPONDENTS USING COSMETIC PRODUCTS**



Inference : From the above table and chart it is observed that 83.3 % of the respondents are using cosmetics in their day to day life and rest of the respondents which is 16.7% are not a user of cosmetics. So major fraction of the respondents are using cosmetics in their daily life.

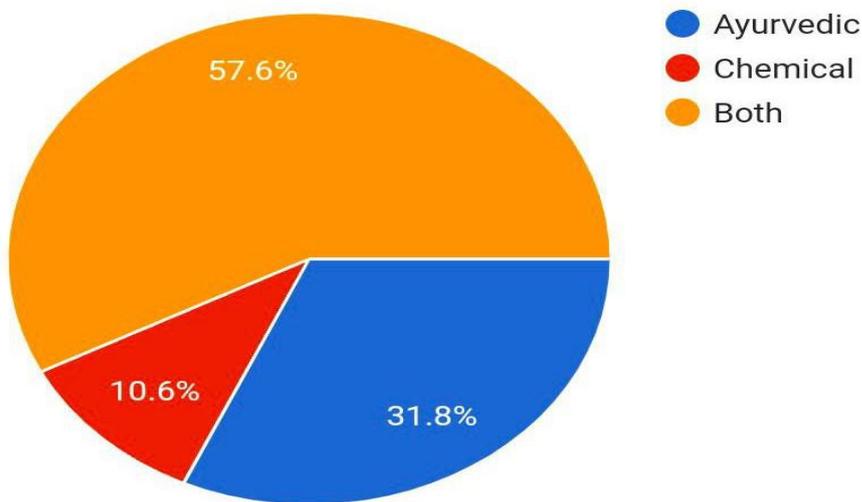
**TABLE 4.5**

**PREFERENCE OF RESPONDENTS ON THE TYPE OF COSMETICS**

TYPE	NO.OF USERS	PERCENTAGE
AYURVEDIC	42	31.8
CHEMICAL	14	10.6
BOTH	76	57.6
TOTAL	132	100

**FIGURE 4.5**

**PREFERENCE OF RESPONDENTS ON THE TYPE OF COSMETICS**



Inference :From the above table and graph its clear that majority ie.57.6% of respondents are using both ayurvedic and chemical cosmetics ,and 31.8% make use of ayurvedic cosmetics and remaining 10.6% are the users of chemical cosmetics.

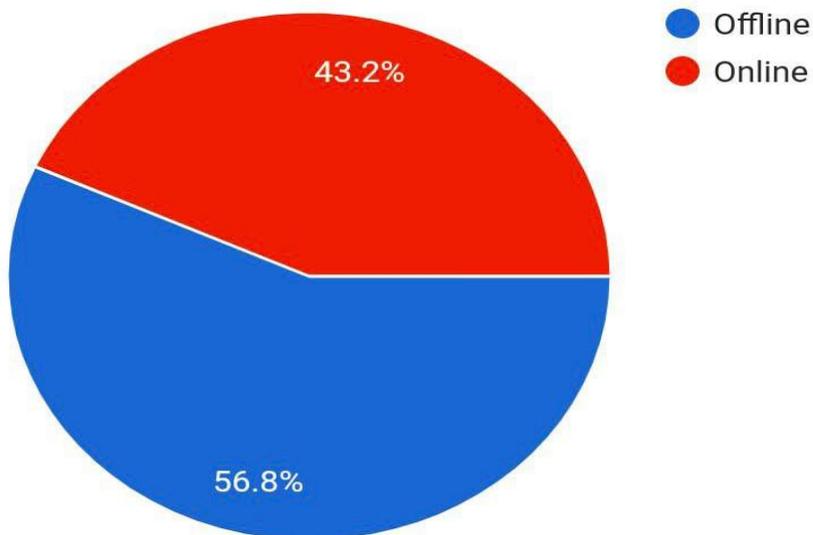
**TABLE 4.6**

**MODE OF PURCHASING COSMETICS**

MODE	NO.OF RESPONDENTS	PERCENTAGE
OFFLINE	74	56.8
ONLINE	58	43.2
TOTAL	132	100

**FIGURE 4.6**

**MODE OF PURCHASING COSMETICS**



Inference :From the analysis most of the respondents that is 56.8 percentage are offline procurers of cosmetics and remaining 43.4 percentage of respondents purchases cosmetics online .Thus it is clear that even in this e-world people prefer to buy cosmetics via physical shopping .

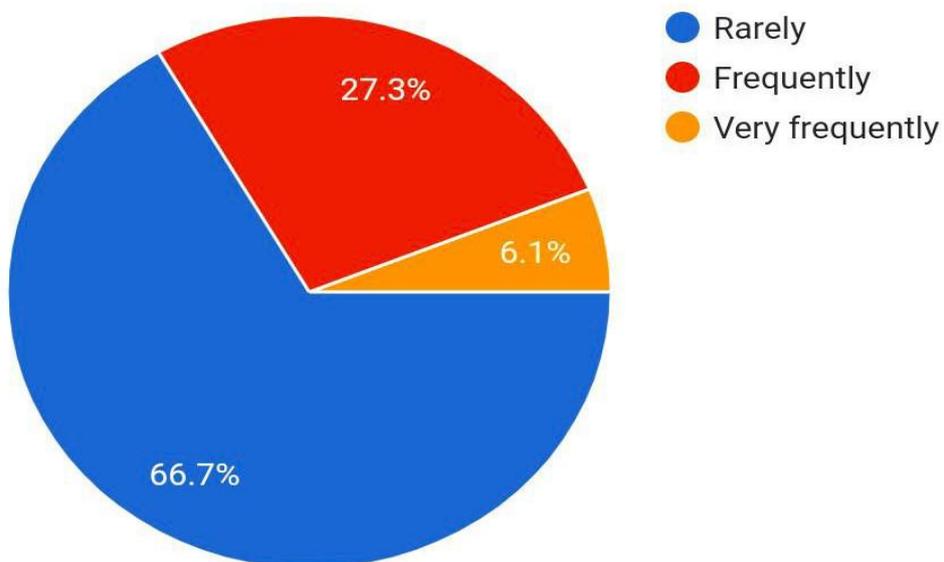
**TABLE 4.7**

**THE FREQUENCY IN PURCHASE OF COSMETICS**

FREQUENCY	NO.OF CONSUMERS	PERCENTAGE
RARELY	88	56.7
FREQUENTLY	36	27.3
VERY FREQUENTLY	8	6.1
TOTAL	132	100

**FIGURE 4.7**

**THE FREQUENCY IN PURCHASE OF COSMETICS**



Inference: From the table and graphical representation we can understand that most of the respondents that is 66.7% rarely buy cosmetics products and 27.3% are frequent buyers whereas 6.1% are very frequent buyers of cosmetics.

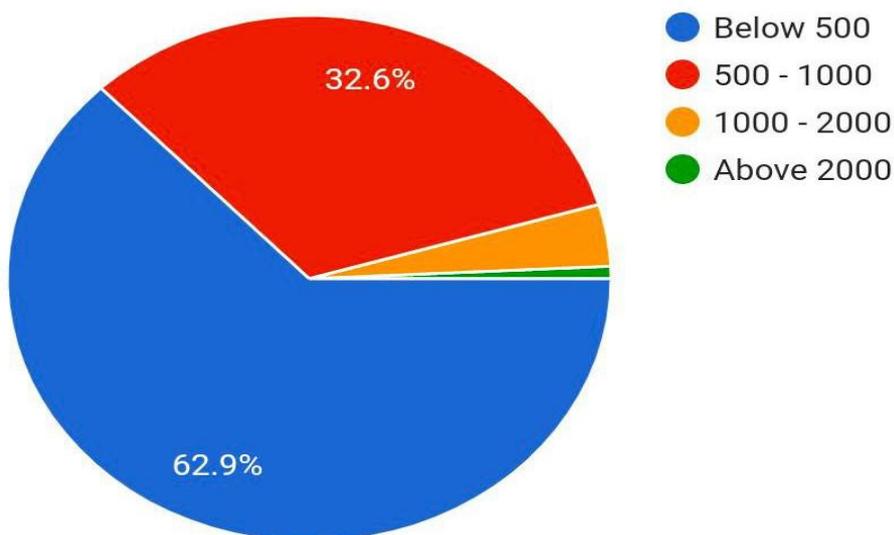
**TABLE 4.8**

**MONTHLY INCOME SPENT ON COSMETICS**

MONTHLY INCOME SPENT	RESPONDENTS	PERCENTAGE
BELOW 500	83	62.9
500 TO 2000	43	32.6
ABOVE 2000	6	4.5
TOTAL	132	100

**FIGURE 4.8**

**MONTHLY INCOME SPENT ON COSMETICS**



Inference :It is observed that most of the respondents ;62.9 percent of them spent below 500 of their income on cosmetics almost 32.6 percent respondents spent 500 to 2000 of their income on cosmetics .The remaining part that is 4.5 percent spends above 2000 for cosmetics from their income

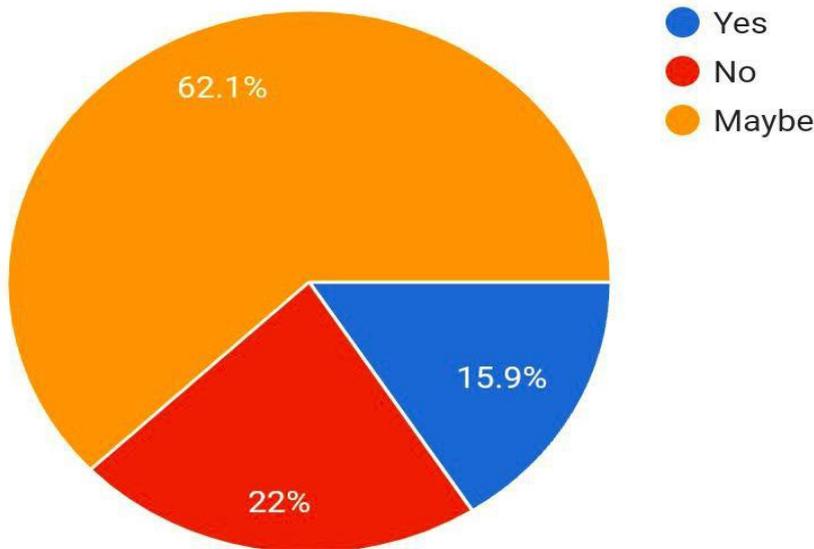
**TABLE 4.9**

**REPONSE ON BUYING COSMETICS AFTER WATCHING ITS ADVERTISEMENTS**

OPINION	RESAPONDENTS NO.	PERCENTAGE
YES	21	15.9
NO	29	22
MAYBE	82	62.19
TOTAL	132	100

**FIGURE 4.9**

**REPONSE ON BUYING COSMETICS AFTER WATCHING ITS ADVERTISEMENTS**



INFERENCE: As per the analysis of table and chart most of the respondents that is 62.1 percentage are 'may be 'willing to purchase cosmetics after watching ads. 22 percentage respondent are not ready to buy cosmetics under ad influence only 15.9 percent is clearly ready to buy cosmetics under ad influence.

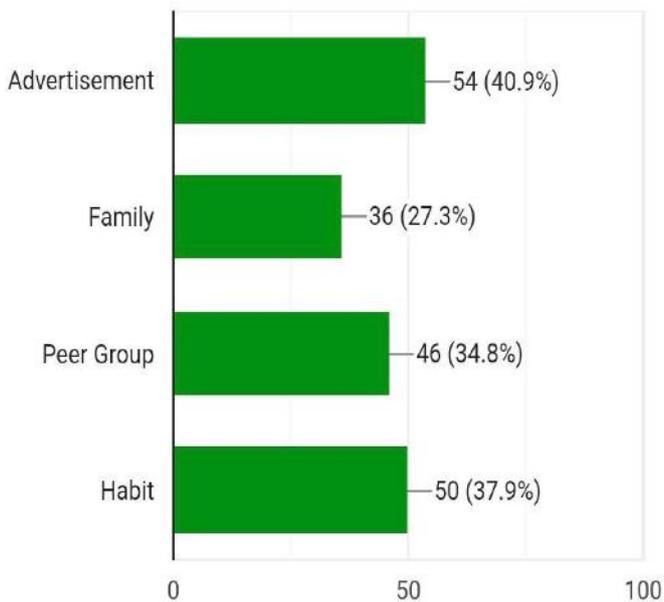
**TABLE 4.10**

**FACTORS INFLUENCING WHILE CHOOSING COSMETICS PRODUCTS**

FACTORS	RESPONDENTS NO.	PERCENTAGE
ADVERTISEMENT	54	40.9
FAMILY	36	27.3
PEER GROUP	46	34.8
HABIT	50	37.9

**FIGURE 4.10**

**FACTORS INFLUENCE WHILE CHOOSING COSMETIC PRODUCTS**



INFERENCE:As per the analysis most of them are influenced by advertisement with 40 percentage. 27 percentage are influenced by family 34 percentage are influenced by peer group and only 37 percentage are influenced by habit. It clearly state that advertisement and family play an important role in choosing the products.

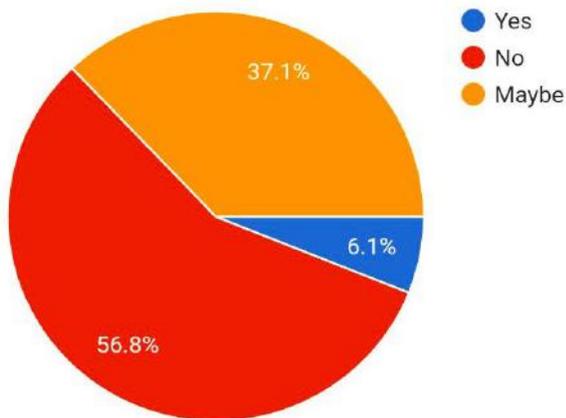
**TABLE 4.11**

**INFLUENCE OF CELEBRITY ENDORSEMENT TO BUY A PRODUCT**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	9	6.1
NO	75	56.8
MAY BE	48	37.1
TOTAL	132	100

**FIGURE 4.11**

**INFLUENCE OF CELEBRITY ENDORSEMENT TO BUY A PRODUCT**



INFERENCE: By observing the chart we can analyze the most of the persons with 56.8% is not influenced by any celebrity endorsement, and only 6.1% of people is only influenced by them. 37.1% of people were may be influenced by them. Thus we can conclude that celebrity endorsement does not have much influence on people.

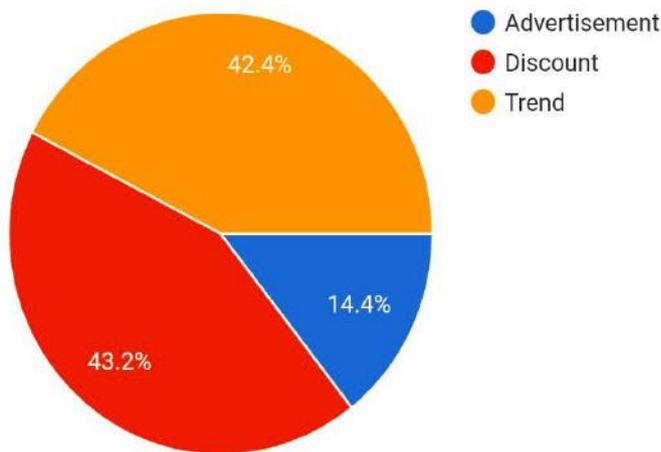
**TABLE 4.12**

**FACTORS THAT MOTIVATES THE CONSUMER TO PURCHASE COSMETIC PRODUCTS**

FACTORS	RESPONDENTS NO.	PERCENTAGE
ADVERTISEMENT	20	14.4
DISCOUNT	57	43.2
TREND	55	42.4

**FIGURE 4.12**

**FACTORS THAT MOTIVATES THE CONSUMER TO PURCHASE COSMETIC PRODUCTS**



INFERENCE: As per the analysis 43.2% of people are influenced by discount in purchasing cosmetic products. 42.4% of people were influenced by trend in purchasing cosmetic products and only 14.4% of people were influenced by advertisement. Discount has significant role in motivating people for purchase.

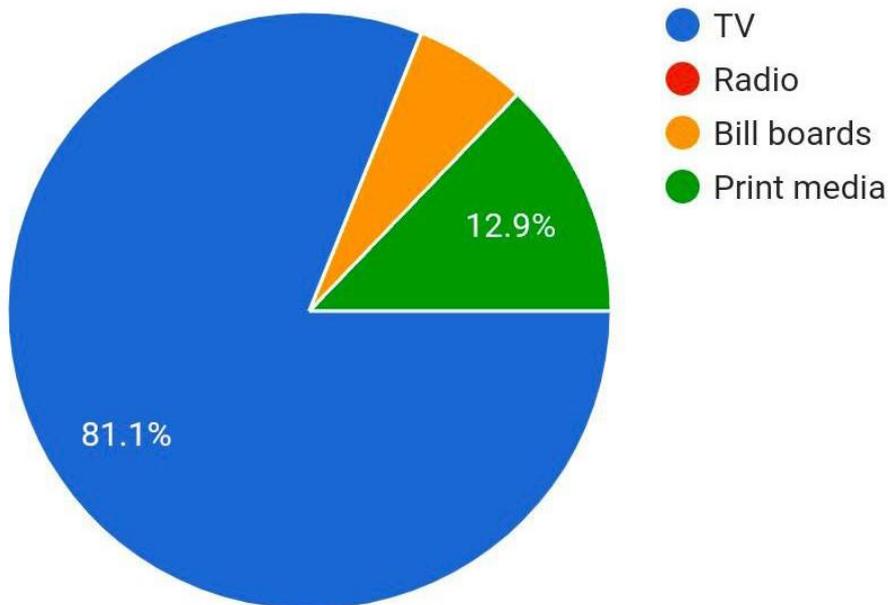
**TABLE 4.13**

**ADVERTISEMENT THAT HAS LONG LASTING IMPACT ON COSMETIC PRODUCTS**

ADVERTISEMENT	RESPONDENTS NO.	PERCENTAGE
TV	107	81.1
RADIO	0	0
BILL BOARDS	8	2
PRINT MEDIA	17	12.9

**FIGURE 4.13**

**ADVERTISEMENT THAT HAS LONG LASTING IMPACT ON COSMETIC PRODUCTS**



INFERENCE: From the above table and graphical representation it is clear that majority with 81.1%, TV advertisement has a great impact on cosmetic products. And print media makes a 12.9% impact. Radio does not making any impact and bill board has only 2% impact.

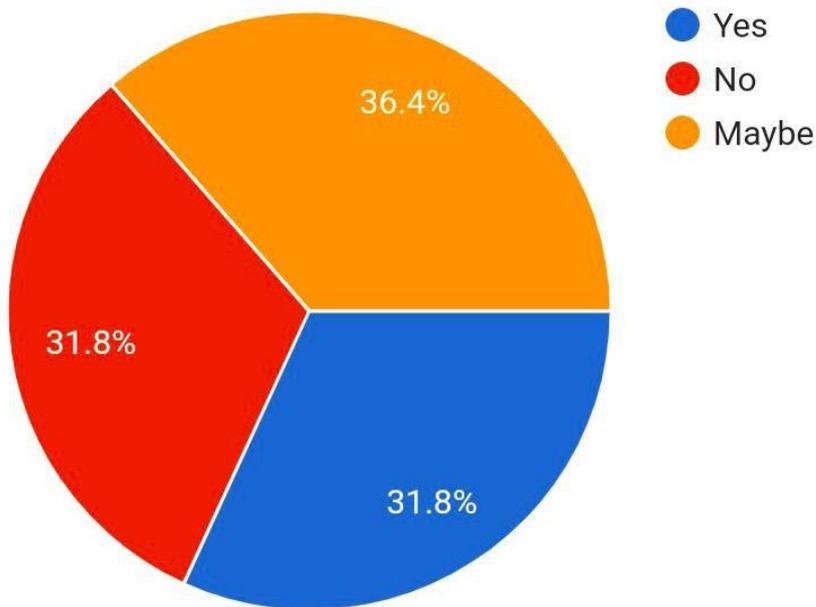
**TABLE 4.14**

**RESPONSE ON EXPENSIVE OR CHEAPER PRODUCT IS BETTER**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	42	31.8
NO	42	31.8
MAY BE	48	36.4
TOTAL	132	100

**FIGURE 4.14**

**RESPONSE ON EXPENSIVE OR CHEAPER PRODUCTS IS BETTER**



INFERENCE: By observation it is clear that 31.8% of people are agreeing and not agreeing with the fact that expensive products are better than cheaper products. And 36.4% of people are somewhat agreeing with the statement.

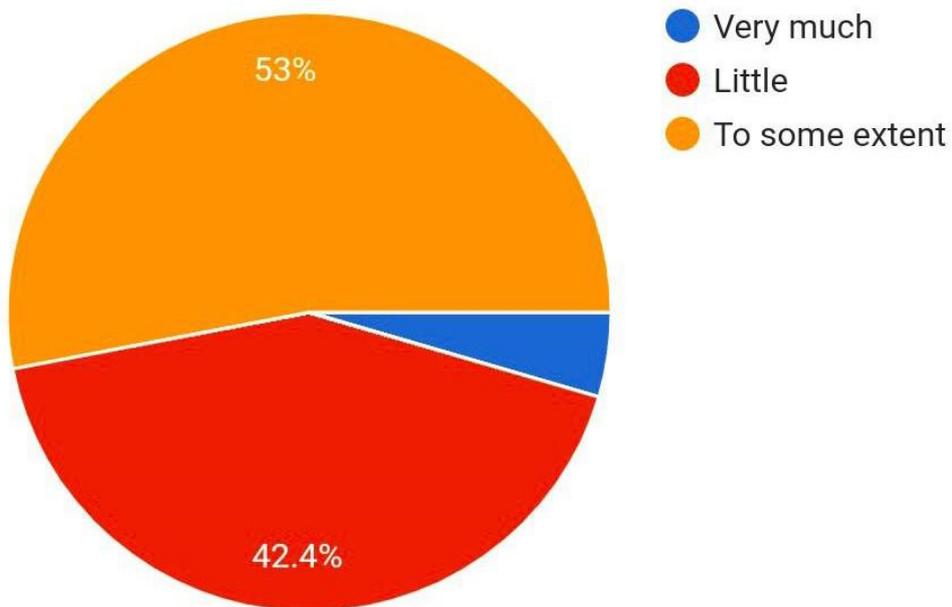
**TABLE 4.15**

**RESPONSE ON BELIEVE IN PROMISED EFFECT OF COSMETIC PRODUCTS**

OPINION	RESPONDENTS NO.	PERCENTAGE
VERY MUCH	8	4.6
LITTLE	55	42.4
TO SOME EXTEND	69	53

**FIGURE 4.15**

**RESPONSE ON BELIEVE IN PROMISED EFFECT OF COSMETIC PRODUCTS**



INFERENCE: From the analysis only 4.6% people were very much believe in promised effect of cosmetic products.42.4% of people has little believe and 53% of people were somewhat believe in the promised effect of cosmetic products.

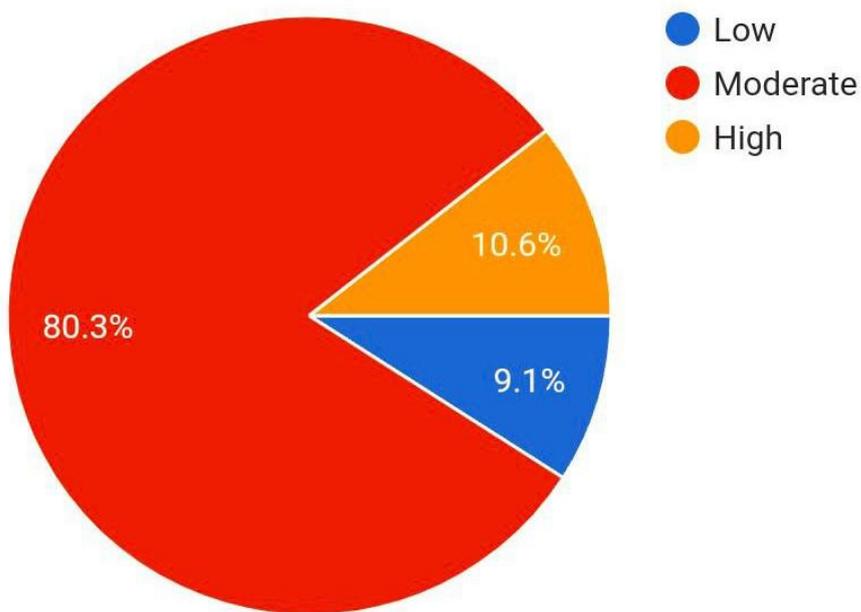
**TABLE 4.16**

**RESPONSE ON QUALITY AND STANDARDS OF COSMETIC ADS**

OPINION	RESPONDENTS NO.	PERCENTAGE
LOW	12	9.1
MODERATE	105	80.3
HIGH	15	10.6
TOTAL	132	100

**FIGURE 4.16**

**RESPONSE ON QUALITY AND STANDARDS OF COSMETIC ADS**



INFERENCE: From the above table and chart representation it is found that majority of the respondents that is 80.3percentage moderately agree on the quality and standards of cosmetic products.10.6 percentage of respondents highly agree and remaining 9.6percentage of respondents not so agreeing to the quality and standards of product.

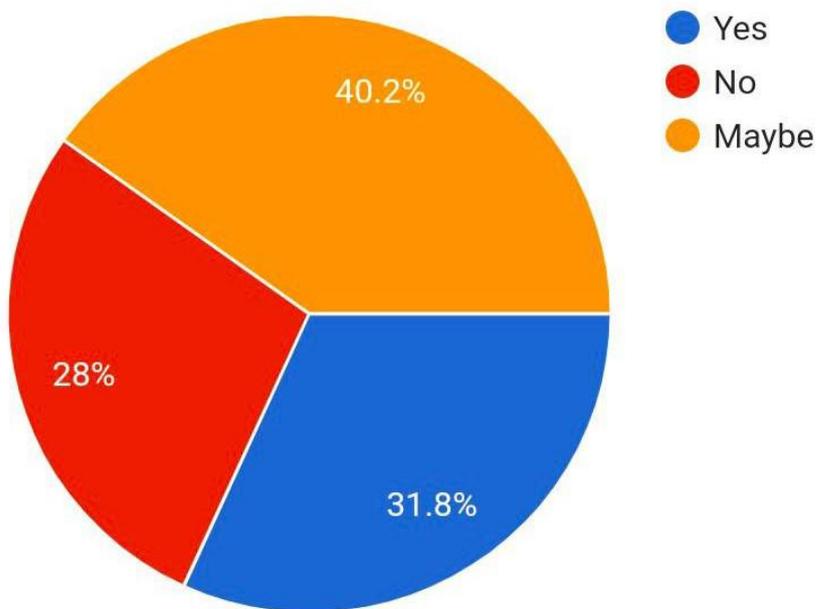
**TABLE 4.17**

**RESPONSE ON USE OF SAMPLE PRODUCTS OF COSMETICS**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	41	31.8
NO	36	28
MAY BE	55	40.2
TOTAL	132	100

**FIGURE 4.17**

**RESPONSE ON USE OF SAMPLE PRODUCTS OF COSMETICS**



INFERENCE: It is observed that 40.2 percentage of respondents are somewhat using sample products of cosmetics. 31.8 percentage of respondents are using sample products and only 28 percentage of respondents were not using any kind of sample products.

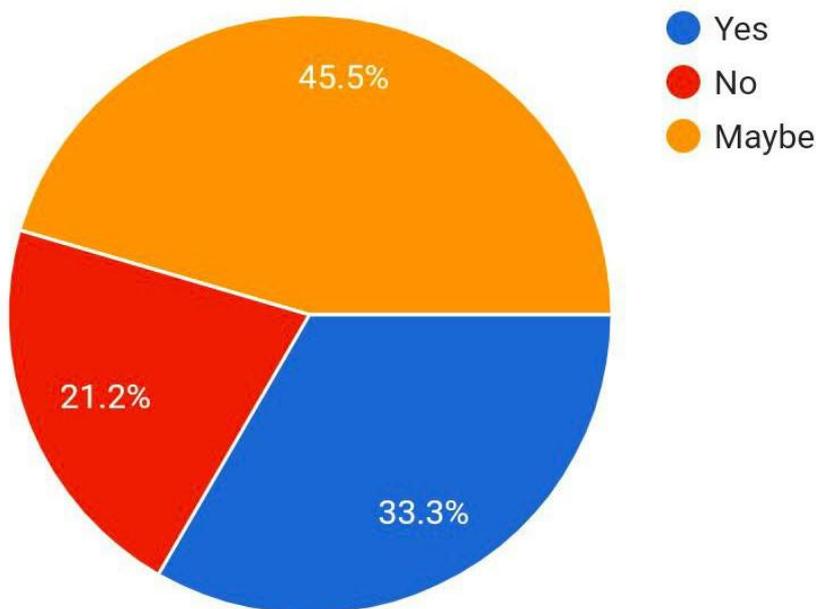
**TABLE 4.18**

**CHANGING PERCEPTION ON THE BASIS OF QUALITY ADVERTISEMENT**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	45	33.3
NO	27	21.2
MAY BE	60	45.5
TOTAL	132	100

**FIGURE 4.18**

**CHANGING IN PERCEPTION ON THE BASIS OF QUALITY ADVERTISEMENT**



INFERENCE: From the analysis it is evident that 33.3 percentage of respondents change their perception according to the quality of advertisement. 45.5 percentage of respondents somewhat change their perception and 21.2 percentage of respondents were not affected by any change in quality of advertisement.

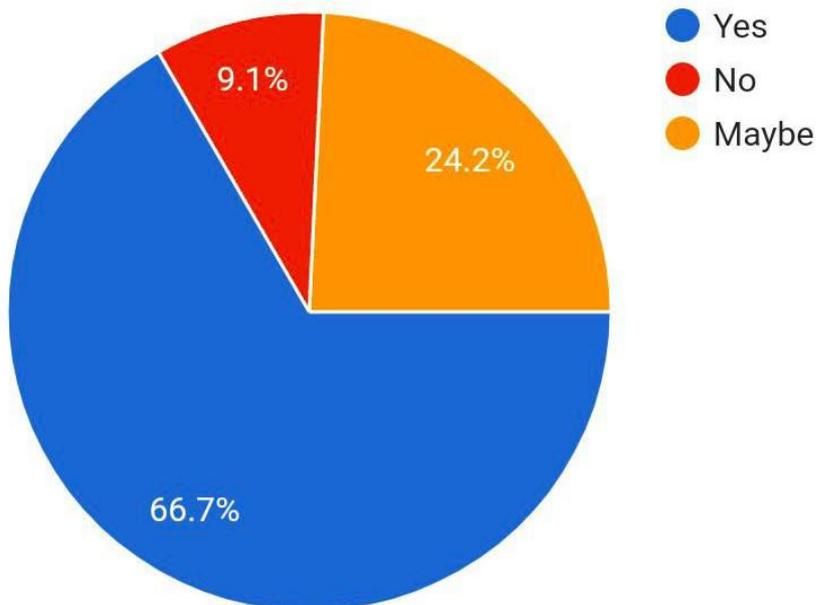
**TABLE 4.19**

**RESPONSE ON THE PREFERENCE OF SAME BRAND OF PRODUCTS REPEATEDLY**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	88	66.7
NO	12	9.1
MAYBE	32	24.2
TOTAL	132	100

**FIGURE 4.19**

**RESPONSE ON THE PREFERENCE OF SAME BRAND OF PRODUCTS REPEATEDLY**



INFERENCE: Majority of the respondents with 66.7 percentage prefer same brand of products repeatedly. 24.2 percentage of respondents maybe prefer the same brand of products and 9.1 percentage of respondents were not concerned about the brand.

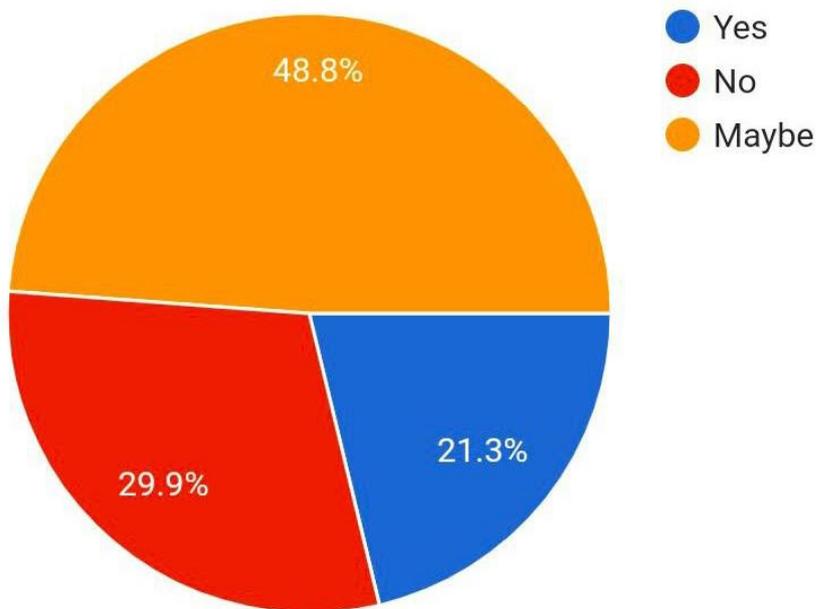
**TABLE 4.20**

**RESPONSE ON SWITCHING THE BRANDS DUE TO ADVERTISEMENT**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	28	21.3
NO	40	29.9
MAYBE	64	48.8
TOTAL	132	100

**FIGURE 4.20**

**RESPONSE ON SWITCHING THE BRANDS DUE TO ADVERTISEMENT**



INFERENCE: From the above analysis 48.8 percentage of respondents were somewhat wish to switch the brand of cosmetic products on the basis of advertishment.21.3 percentage of respondents were willing to switch their brands and 29.9 percentage of respondents were not willing to switch their products.

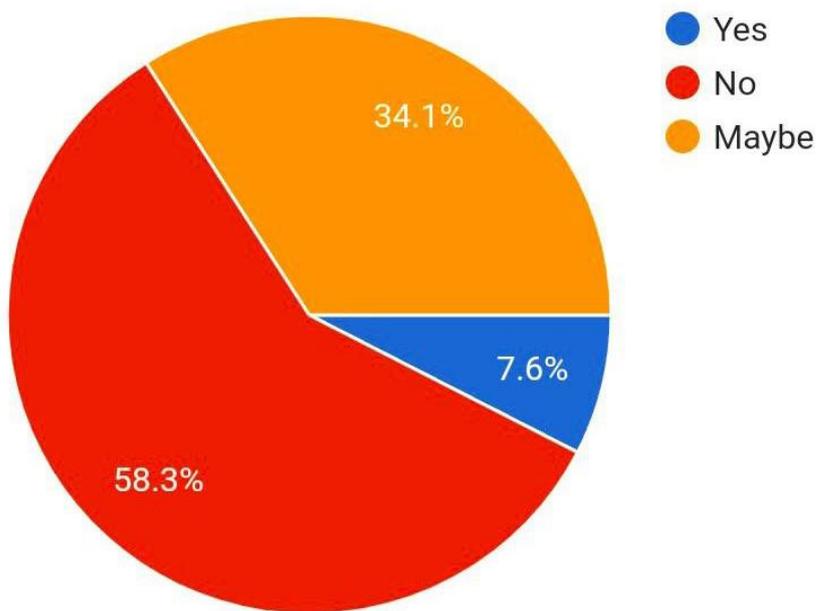
**TABLE 4.21**

**RESPONSE ON THE BELIEF OF INFORMATION GIVEN THROUGH ADVERTISEMENT**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	10	7.6
NO	77	58.3
MAYBE	45	34.1
TOTAL	132	100

**FIGURE 4.21**

**RESPONSE ON THE BELIEF OF INFORMATION GIVEN THROUGH ADVERTISEMENT**



INFERENCE: As per the analysis 58.3 percentage of respondents are considered that the information provided through advertisement are not enough to buy the cosmetic products. 34.1 percentage of respondents may be considered that the information is enough and for remaining 7.6percentage of respondents, information given through advertisement is enough.

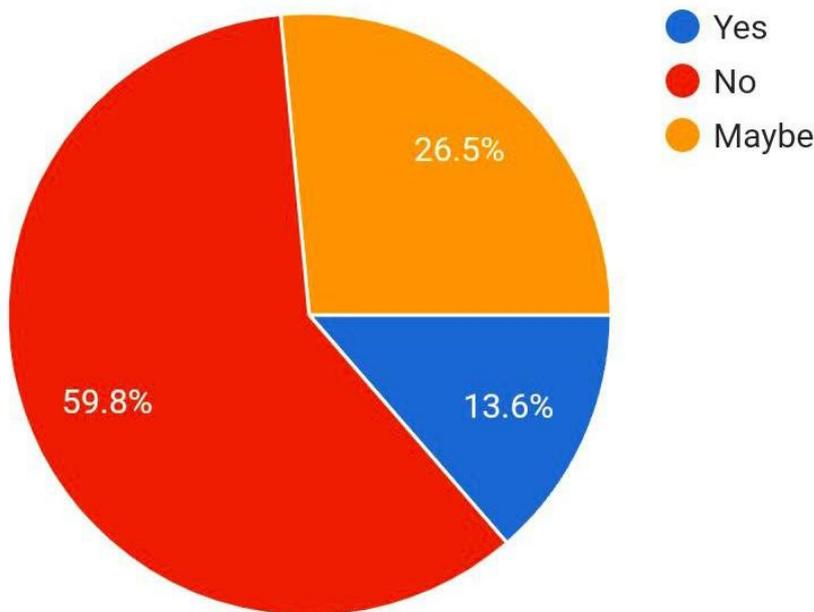
**TABLE 4.22**

**RESPONSE ON THE USE OF COSMETIC PRODUCTS IN DAY TO DAY LIFE**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	18	13.6
NO	80	59.8
MAYBE	34	26.5
TOTAL	132	100

**FIGURE 4.22**

**RESPONSE ON THE USE OF COSMETIC PRODUCTS IN DAY TO DAY LIFE**



Inference: Most of respondents with 59.8 percentage are not agreeing with the fact that cosmetic products are needed for day to day life. 26.5 percentage of respondents are somewhat agreeing to this statement and 13.6 percentage of respondents are considered that cosmetic products are necessary for day to day life.

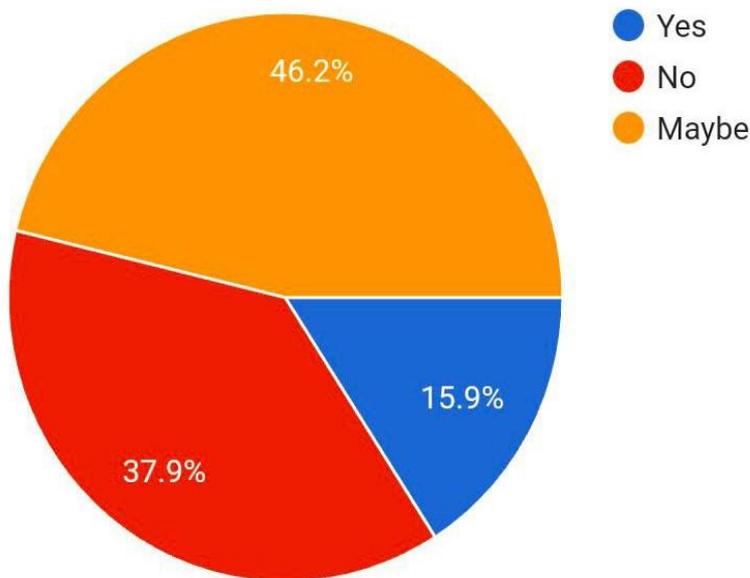
**TABLE 4.23**

**RESPONSE ON THE AVAILABILITY OF COSMETICS AT FAIR RATE**

OPINION	RESPONDENTS NO.	PERCENTAGE
YES	22	15.9
NO	50	37.9
MAYBE	60	46.2
TOTAL	132	100

**FIGURE 4.23**

**RESPONSE ON THE AVAILABILITY OF COSMETICS AT FAIR RATE**



Inference: From the above table and chart it is evident that most of the respondents with 46.2percentage are may be feel that cosmetics are available at fair rate.37.9 percentage of respondents does feel that products are not available at fair rate and 15.9 percentage of respondents are agreeing that products available at fair rate.

## **CHAPTER 5**

# **SUMMARY, FINDINGS, SUGGESTIONS, CONCLUSIONS**

## **5.1 SUMMARY**

Survey was conducted to analyse the impact of advertisement on consumers with special reference to cosmetics industry. The population selected was from Ernakulam district.

From the analysis we find that advertisement is the major factor which influence consumers for buying cosmetic products. They believe that the information generated out of advertisement are enough purchase the cosmetic products. So, the growth of cosmetic industry depends on one of the main factor advertisement. A good advertisements leads the cosmetic industry into a positive atmosphere.

This chapter deals with the findings, suggestions and conclusion through data analysis.

## **5.2 FINDINGS**

As per the analysis we can find that most of the consumers purchase cosmetic through online and the factor which influence to buy cosmetics is advertisement . They believe that the information given in advertisements are enough to buy cosmetics products. The concern for purchase cosmetics by the consumers are quality, quantity, price, advertisement, long-lasting effect, health issue, effectiveness, public reviews etc.

1. Majority of the respondents are belongs to the age group above 25.
2. We can find that female respondents(87.1%) are more than male respondents.(12.9%)
3. Majority of the respondents are students.(81.8%)
4. Majority of the respondents are using cosmetics products in their day to day life.(83.3%)
5. Most of the consumers are preferred to use both Ayurvedic and chemical products.(57.6%)
6. Majority of the consumers prefer to purchase cosmetic products through offline mode.(56.8%)
7. As per our findings , Majority of the consumers buy cosmetic products rarely.(66.7%)
8. Majority of the consumers spend below 500 from their monthly income for purchasing cosmetics products.(62.9%)
9. There is equal chance in buying or not to buy the cosmetic products after watching it's advertisements.(62.1%)
10. Majority of the consumers are influenced by the factor advertisement for choosing cosmetic products.(40.9%)
11. For purchasing cosmetic products the factor celebrity endorsement is not influenced by the customers to a great extent.(56.8%)
12. Most of the consumers are motivated to purchase is it's discount that avail the cosmetic products and not its advertisement or trend.(43.2%)

13. Majority of the consumers consider that the advertisement that shown on media like TV creates long-lasting impact in the minds of customers in terms of cosmetic products.(81.1%)
14. Most of the customers believe in the promised effect of cosmetic products only to some extent.(53%)
15. Majority of customers think that the quality and standards of cosmetics products that shown in advertisements are only a moderate to purchase.(80.3%)
16. Most of the consumers agreed that their perception about the products can change if their advertisements maintain quality standards.(45.5%)
17. A large number of customers are prefer to use same brand products repeatedly if they are satisfied with that brand products.(66.7%)
18. Majority of consumers believe that the information given through advertisements are not enough to buy it.(58.3%)
19. A large number of customers think that it is not necessary to use cosmetics in day to day life.(58.6,%)
20. Most of the consumers feels that the cosmetic products are available at fair rate sometimes only. (46.2%)
21. The concerns for purchasing cosmetic products by majority of consumers are its quality, price, advertisement, long-lasting effect, health, effectiveness, public reviews etc.
22. Now a days the view of society towards beauty has changed. They doesn't believe that fairness is beauty, in fact wellness is beauty .
23. The change of advertisements with the change of societal attitude it also considerably noted.
24. The consumers changing attitude resulted in innovations in the cosmetics industry.
25. The popularity and growth of industry can be noted. More over new products entry is also marked.

### **5.3 SUGGESTIONS**

1. Increase local people review of the product along with advertisement.
2. Make the online website of selling cosmetic products more attractive, unique, useful, and professional.
3. Advertisements should project the brand logo in unique precise and attractive form.
4. Try to reach the local people through flyers, a leaflet which give relevant information about cosmetic products.
5. Use more expert people and celebrities testimonial and consistent image along with your advertisement.
6. Use famous celebrities photos along with other details in the packing box.
7. Make use of user generated content for advertising the products rather than artificial one.
8. Make use of visual social media platforms like Instagram for the product advertisement.

## **5.4 CONCLUSION**

The study on “ the impact of advertisements on cosmetics industry “aims at identifying the impact of commercial ads of cosmetics on consumers .Cosmetics include all those chemical and natural products which promises to enhance the beauty and wellness of people. Now a days using cosmetics has become one of the integral part of human life. People use cosmetics to improve their physical and mental satisfaction. The cosmetics industry break the chains of gender discrimination by making products which both men and women can make use of. More over it also break the taboo of using cosmetics by high status people of the society to common people through its reasonable rates and appealing ads.

Through this study the role of ads in making the cosmetics purchase was evidently found out. It help us to understand the various techniques used by ad agencies to promote the products. Apart from that the study pointed out the ill effects of over promotion and exploitations faced by the cosmetics industry from ad companies .

The data analysis conducted helped us to know the preference of the people regarding the type of cosmetics ,its usage ,the income they spend on it ,the influence of advertisements etc...The sampling technique used was questionnaire collection ,and the technique was very useful too. Response was collected from 132 people among them some where students and some where employees which belong to the target group focused.

From the study we were able to conclude that most people consider cosmetics as a part of their day to day life ,and ads plays a major role in its purchase and use.

Thus we conclude the study by stating that advertisements are necessary promotional tool for cosmetic industry and it has its impact on the growth and profitability of cosmetic industry

# **BIBLIOGRAPHY**

## **REFERENCES;**

### **Online articles :**

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2740920](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2740920) Impact of Advertisement on Buying Behavior of the Consumers: Study of Cosmetic Industry in Karachi City.

<http://j.vidhyayanaejournal.org/index.php/journal/article/download/132/134> Impact of Advertisement on Buying behavior of the consumers: Study of Cosmetic Industry in Ahmedabad City.

[https://ajmjournal.com/HTML\\_Papers/Asian%20Journal%20of%20Management\\_PID\\_2020-11-4-12.html](https://ajmjournal.com/HTML_Papers/Asian%20Journal%20of%20Management_PID_2020-11-4-12.html) Impact of Advertisement on Consumer buying pattern in Cosmaceutical Segment.

[http://ictactjournals.in/paper/IJMS\\_Vol4\\_Iss3\\_Paper4\\_800\\_808.pdf](http://ictactjournals.in/paper/IJMS_Vol4_Iss3_Paper4_800_808.pdf) Impact of Advertisement on Buying behavior of Consumers in Sivakasi.

<http://www.ijmtbr.com/PDF/IJMTBR-2019-07-03-10.pdf> TO STUDY THE INFLUENCE OF ATTRACTIVE ADVERTISEMENT ON BUYING BEHAVIOR OF YOUNG GIRL STUDENTS OF MANAGEMENT INSTITUTES WITH ...

<https://www.academia.edu/download/63006486/IJRCL4Nov19-835520200418-107791-15txxyq.pdf> IMPACT OF CELEBRITY ENDORSEMENT ON BRAND IMAGE A STUDY ON COSMETIC INDUSTRY.

<https://boa.unimib.it/bitstream/10281/150298/1/7th.ISF.Oxford%202017.pdf#page=16> Advertising for Natural Beauty Products: The Shift in Cosmetic Industry.

<https://sourceessay.com/positive-and-negative-effects-of-advertisements/> POSITIVE AND NEGATIVE SIDES OF ADVERTISEMENT

<https://web.archive.org/web/20130312162359/https://www.womenshealth.gov/publications/our-publications/fact-sheet/cosmetics-your-health.cfm> Cosmetics and health fact sheet.

<https://www.washingtonpost.com/archive/opinions/2002/02/03/when-celebrity-endorsers-go-bad/260776e6-d38c-4319-b683-eb466c499dce/>

[https://en.m.wikipedia.org/wiki/Advertising#cite\\_ref-20](https://en.m.wikipedia.org/wiki/Advertising#cite_ref-20)

### **Journals :**

▪ "Cosmetics Overview". U.S. Food & Drug Administration. Archived from the original on 30 March 2019. Retrieved 30 March 2019.

▪ Riordan, Teresa (2004-12-12). "'Inventing Beauty'". The New York Times. ISSN 0362-4331. Archived from the original on 2021-08-24. Retrieved 2021-08-24.

▪ JOURNAL OF COSMETIC SCIENCE. The Official Journal of the Society of Cosmetic

**Chemists William J. Stanton. Fundamentals of Marketing. McGraw-Hill (1984).**

▪ **Kalender, G. I. (2021). The Semiotic Analysis of Cosmetic Advertisements on Facebook. Advances in Social Sciences Research Journal, 7(12), 658–671. <https://doi.org/10.14738/assrj.712.9528>**

**Books :**

▪ **Plant Extracts in Skin Care Products. Editors: M. Beatriz P.P. Oliveira, Francisca Rodrigues.**

▪ **Arens, William, and Michael Weigold. Contemporary Advertising: And Integrated Marketing Communications (2012)**

▪ **Belch, George E., and Michael A. Belch. Advertising and Promotion: An Integrated Marketing Communications Perspective (10th ed. 2014)**

▪ **Biocca, Frank. Television and Political Advertising: Volume I: Psychological Processes (Routledge, 2013)**

▪ **Chandra, Ambarish, and Ulrich Kaiser. "Targeted advertising in magazine markets and the advent of the internet." Management Science 60.7.**

**Article :**

▪ **Developing Eco-Friendly Skin Care Formulations with Microemulsions of Essential Oil by Sie Huey Lee \*,Pui Shan Chow and Chetan Kantilal Yagnik.**

# APPENDIX

# **QUESTIONNAIRE**

A study on the impact of advertisement on consumers with special reference to cosmetics industry

Dear Sir/Madam,

We are the students of Department of commerce (Regular), St Teresa's college, Ernakulam. As a part of project for the academic year 2019-22, we intend to have a study on "impact of advertisement on consumers with special reference to cosmetics industry".

We request you to be kind enough to fill up the questionnaire so as to enable us to obtain information about the impact of advertisement on cosmetics industry.

1. **Name\*:** \_\_\_\_\_
2. **Age\* :** \_\_\_\_\_
3. **Gender \***
  - Male
  - Female
  - Others
4. **Category\***
  - Student
  - Employee
  - Others
5. **Are you a consumer of cosmetic products?\***
  - Yes
  - No

- 6. Which type of cosmetics do you prefer you most?\***
  - Ayurvedic
  - Chemical
  - Both
- 7. Which mode of purchasing do you prefer?\***
  - Offline
  - Online
- 8. How often do you buy cosmetics products?\***
  - Rarely
  - Frequently
  - Very frequently
- 9. What part of your monthly income is spend on cosmetics?\***
  - Below 500
  - 500 to 2000
  - Above 2000
- 10. Do you buy the cosmetics products after watching its advertisements?\***
  - Yes
  - No
  - Maybe
- 11. Which factors influence you the most while choosing your cosmetic products?\***
  - Advertisements
  - Peer group
  - Family
  - Habit
- 12. Does celebrity endrosemment influence you to buy a cosmetic product?\***
  - Yes
  - No
  - Maybe
- 13. What motivates your purchasing process in terms of cosmetics?\***
  - Advertisement
  - Discount
  - Trend

- 14. According to you what kind of advertisement has long lasting impact in terms of cosmetic products?\***
- TV
  - Radio
  - Billboards
  - Print Media
- 15. Do you think expensive cosmetic products are better than cheaper products?\***
- Yes
  - No
  - Maybe
- 16. How much do you believe in the promised effect of cosmetic products?\***
- Very much
  - Little
  - To some extent
- 17. What do you think about the quality and standards of cosmetics advertisement?\***
- Low
  - Moderate
  - High
- 18. Do you use sample products before buying cosmetics products?\***
- Yes
  - No
  - Maybe
- 19. 19. Do quality advertisements change your perception about the products?\***
- Yes
  - No
  - Maybe
- 20. Do you prefer to use same brand of products repeatedly ?\***
- Yes
  - No
  - Maybe

- 21. Do you wish to switch the brand of the cosmetic products ?\***
- Yes
  - No
  - Maybe
- 22. Do you believe that the information given through advertisements are enough to buy it?\***
- Yes
  - No
  - Maybe
- 23. Do you think it is necessary to use cosmetics in day to day life?\***
- Yes
  - No
  - Maybe
- 24. Do you feel that cosmetics are available at fair rate?\***
- Yes
  - No
  - Maybe
- 25. What are your concerns for selecting a cosmetic product?\***

Project Report

On

# STROKE PREDICTION

*Submitted*

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

MASTER OF SCIENCE

*in*

APPLIED STATISTICS AND DATA ANALYTICS

*by*

ELIZA IVAN THEODORE

(Register No. SM20AS011)

(2020-2022)

*Under the Supervision of*

VRINDA MURALEEDHARAN



DEPARTMENT OF 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, **STROKE PREDICTION** is a bonafide record of the work done by Ms. **ELIZA IVAN THEODORE** 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*  
**Vrinda Muraleedharan**  
Assistant Professor,  
Department Statistics,  
St. Teresa's College(Autonomous),  
Ernakulam.

*Shanty B P*

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

External Examiners

1: *Komen*

2: *m. H/12*  
26/05/22

26/05/2022

# DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of VRINDA MURALEEDHARAN, 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.

Elizatheodore

ELIZA IVAN THEODORE

Date: 09/05/2022

SM20AS011

## ACKNOWLEDGEMENTS

I must mention several individuals who encouraged me to carry this work. Their continuous invaluable knowledgeable guidance throughout the course of this study helped me to complete the work up to this stage

I am very grateful to my project guide Vrinda Muraleedharan for the immense help during the period of work

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'm also very thankful to HoD for their valuable suggestions, critical examination of work during the progress.

Ernakulam.

Date: 09/05/2022

ELIZA IVAN THEODORE

SM20AS011

# ABSTRACT

The main objective of this project is to study and construct a prediction model for predicting stroke and also to assess which model is more accurate. In this project we will explore seven different models to see which produces more accurate and repeatable results. The models used are: Decision Tree, Logistic Regression, Random Forest, Support Vector Machine, K Nearest Neighbour, Naive Bayes and KMeans Clustering. After using these models we come to an accurate prediction. From the prediction outcome of the models, the best performance model will undergo the cross validation process to evaluate its repeatability.

# Contents

<i>CERTIFICATE</i> . . . . .	ii
<i>DECLARATION</i> . . . . .	iii
<i>ACKNOWLEDGEMENTS</i> . . . . .	iv
<i>ABSTRACT</i> . . . . .	v
<i>CONTENT</i> . . . . .	vi
<b>1 INTRODUCTION AND LITERATURE REVIEW</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Literature Review . . . . .	2
<b>2 DATA DESCRIPTION AND PREPROCESSING</b>	<b>7</b>
2.1 Data Source: . . . . .	7
2.2 ATTRIBUTE INFORMATION . . . . .	8
2.3 DATA PREPROCESSING . . . . .	9
2.4 ID . . . . .	9
2.5 Gender . . . . .	9
<b>3 DATA EVALUATION</b>	<b>10</b>
3.1 EXPLORATORY DATA ANALYSIS . . . . .	10
3.2 Numerical Feature Analysis . . . . .	12
3.3 Multicollinearity Analysis . . . . .	14
3.4 Final Preprocessing . . . . .	14
<b>4 MODEL SELECTION</b>	<b>16</b>
4.1 Decision Tree: . . . . .	16
4.2 LOGISTIC REGRESSION MODEL . . . . .	17
4.3 Random Forest Algorithm . . . . .	18

4.4	Support Vector Machine Algorithm . . . . .	18
4.5	K-Nearest Neighbor(KNN) Algorithm . . . . .	19
4.6	Naive Bayes Classifier Algorithm . . . . .	19
4.7	K-Means Clustering Algorithm . . . . .	20
<b>5</b>	<b>MODEL BUILDING AND PREDICTION</b>	<b>22</b>
5.1	PREDICTION: . . . . .	22
<b>6</b>	<b>Cross Validation and Conclusion</b>	<b>24</b>
6.1	Cross Validation . . . . .	24
6.2	Conclusion . . . . .	24
	<i>REFERENCES</i> . . . . .	25

# Chapter 1

## INTRODUCTION AND LITERATURE REVIEW

---

### 1.1 Introduction

Stroke is a medical condition that can lead to the death of a person. It's a severe condition and if treated on time we can save one's life and treat them well. There can be a number of factors that can lead to strokes and in this project blog, we will try to analyze a few of them. I have taken the dataset from Kaggle. It has 11 variables and 5110 observations. According to World Health Organisation (WHO), stroke are the second leading cause of death and the third leading cause of disability globally. Stroke is the sudden death of some brain cells due to lack of oxygen when the blood flow to the brain is lost by blockage or rupture of an artery to the brain, it is also a leading cause of dementia and depression. Nearly 800,000 people in the United States suffer from a stroke each year, with about three in four being first-time strokes. 80 strokes can be prevented, so putting in place proper education on the signs of stroke is very important. The objective of this study is to construct a prediction model for predicting stroke and to assess the accuracy of the model. We will explore seven different models to see which produces reliable and repeatable results. The models are: Decision Tree, Logistic Regression, Random Forest, Support Vector Machine, K Nearest Neighbour, Naive Bayes and K Means Clustering.

## 1.2 Literature Review

### 1. Stroke Prediction : A Systematic Review By - (Jeena R S , Dr Sukesh Kumar A)

Stroke is a cerebro-vascular ailment affecting the normal blood supply to the brain. Without proper supervision, it leads to death or long term disability. Stroke can be either ischemic or hemorrhagic. There is a possibility for the cooccurrence of both ischemic and hemorrhagic strokes. Stroke[1] caused due to a clot in the blood vessel is known as Ischemic stroke and that due to a rupture of blood vessel is referred to as Hemorrhagic stroke. The deficiency of oxygen to the cerebral nervous system is referred to as ischemia which finally leads to their death. About 85 percentage of all strokes belong to ischemic category. Its frequency is accelerating in developing countries like India due to unhealthy lifestyles. Stroke is largely driven by demography and enhanced by increasing frequency of modifiable risk factors. Several risk factors contribute to the risk of stroke. Lifestyle risk factors include diet, cigarette smoking [3], obesity, physical inactivity, alcohol consumption, family and genetic factors, age, sex, drug usage , race, geographic location, climate and socioeconomic factors whereas medical conditions include Atrial fibrillation, Blood pressure [2], Diabetes mellitus, Cholesterol, Mitral valve disease, C reactive protein, Sickle cell disease, Hyperlipidemia, and Transient ischemic attack (TIA). High blood pressure, heart disease and diabetes are the highest risk factors of stroke. But, they often do not cause symptoms in their early stages. Total count of risk factors is directly related to probability of stroke occurrence. Research work shows that physiological parameters can be considered as risk factors for predicting the near-term occurrence of stroke. The aim of this paper is to review the presently existing and widely used cardiovascular risk assessment models and to examine the evidence available on new biomarkers and the nonclinical measures in improving the risk prediction. Identification of individuals at risk of cardiovascular disease (CVD) is of main concern. Cardiovascular risk

prediction has progressed with the development and refinement of risk prediction models based upon well established clinical factors, discovery of biomarkers and social factors may give extra details on the risk of disease. Risk charts and risk score, based on global absolute risk are the key tools for CVD risk- assessment and determine the likelihood of developing the disease .

## 2. Stroke

By : Graeme J Hankey In the past decade, the definition of stroke has been revised and major advances have been made for its treatment and prevention. For acute ischaemic stroke, the addition of endovascular thrombectomy of proximal large artery occlusion to intravenous alteplase increases functional independence for a further fifth of patients. The benefits of aspirin in preventing early recurrent ischaemic stroke are greater than previously recognised. Other strategies to prevent recurrent stroke now include direct oral anticoagulants as an alternative to warfarin for atrial fibrillation, and carotid stenting as an alternative to endarterectomy for symptomatic carotid stenosis. For acute intracerebral haemorrhage, trials are ongoing to assess the effectiveness of acute blood pressure lowering, haemostatic therapy, minimally invasive surgery, antiinflammation therapy, and neuroprotection methods. Pharmacological and stem-cell therapies promise to facilitate brain regeneration, rehabilitation, and functional recovery. Despite declining stroke mortality rates, the global burden of stroke is increasing. A more comprehensive approach to primary prevention of stroke is required that targets people at all levels of risk and is integrated with prevention strategies for other diseases that share common risk factors.

## 3. An integrated machine learning approach to stroke prediction

By : (A Khosla, Y Cao, CCY Lin, HK Chiu,)

Stroke is the third leading cause of death and the principal cause of serious long-term disability in the United States. Accurate prediction of stroke is highly valuable for early intervention and treatment. In this study, we compare the Cox proportional hazards model with a machine learning approach for stroke prediction on the Cardiovascular Health

Study (CHS) dataset. Specifically, we consider the common problems of data imputation, feature selection, and prediction in medical datasets. We propose a novel automatic feature selection algorithm that selects robust features based on our proposed heuristic: conservative mean. Combined with Support Vector Machines (SVMs), our proposed feature selection algorithm achieves a greater area under the ROC curve (AUC) as compared to the Cox proportional hazards model and L1 regularized Cox feature selection algorithm. Furthermore, we present a margin-based censored regression algorithm that combines the concept of margin-based classifiers with censored regression to achieve a better concordance index than the Cox model. Overall, our approach outperforms the current state-of-the-art in both metrics of AUC and concordance index. In addition, our work has also identified potential risk factors that have not been discovered by traditional approaches. Our method can be applied to clinical prediction of other diseases, where missing data are common and risk factors are not well understood.

4. A stroke prediction score in the elderly By : (T Lumley, RA Kronmal, M Cushman, TA Manolio... - Journal of clinical)

The objective of this study was to construct a prediction model for predicting stroke in an elderly U.S. population, and to assess the accuracy in this population of other previously published prediction models. The subjects were participants in the Cardiovascular Health Study: 2,495 men and 3,393 women age 65 years and older at baseline, and followed for 6.3 years. Among 5,711 participants free of baseline stroke, 399 strokes occurred. Sex-specific prediction equations were constructed using study variables that were most importantly related to incident stroke: age, systolic blood pressure, diabetes, ECG diagnosis of atrial fibrillation or left ventricular hypertrophy, confirmed history of cardiovascular disease, diabetes, time to walk 15 ft, and serum creatinine. The prediction rule was implemented as a risk score and in a Web-based interactive Java applet. Overall, the model predicted 5-year stroke risks ranging from less than 1 to 59 percentage. The 20 percentage of subjects in the highest predicted risk group had a 5-year actual

stroke incidence rate of 15 percentage, while the 20 percentage lowest risk group had a 1 percentage incidence. Risk scores from two other studies performed well in these study participants. Effective discrimination between low and high stroke risk in the elderly was possible in this cohort with data that are easy to obtain. Evaluation of the generalizability and clinical usefulness of this prediction model requires further research.

5. Performance analysis of machine learning approaches in stroke prediction ( By : Minhaz Uddin Emon, Maria Sultana Keya, Tamara Islam Meghla)

Most of strokes will occur due to an unexpected obstruction of courses by prompting both the brain and heart. Early awareness for different warning signs of stroke can minimize the stroke. This research work proposes an early prediction of stroke diseases by using different machine learning approaches with the occurrence of hypertension, body mass index level, heart disease, average glucose level, smoking status, previous stroke and age. Using these high features attributes, ten different classifiers have been trained, they are Logistics Regression, Stochastic Gradient Descent, Decision Tree Classifier, AdaBoost Classifier, Gaussian Classifier, Quadratic Discriminant Analysis, Multi layer Perceptron Classifier, KNeighbors Classifier, Gradient Boosting Classifier, and XGBoost Classifier for predicting the stroke. Afterwards, results of the base classifiers are aggregated by using the weighted voting approach to reach highest accuracy. Moreover, the proposed study has achieved an accuracy of 97 percentage, where the weighted voting classifier performs better than the base classifiers. This model gives the best accuracy for the stroke prediction. The area under curve value of weighted voting classifier is also high. False positive rate and false negative rate of weighted classifier is lowest compared with others. As a result, weighted voting is almost the perfect classifier for predicting the stroke that can be used by physicians and patients to prescribe and early detect a potential stroke.

6. Stroke prediction using artificial intelligence

By :( MS Singh, P Choudhary )

A stroke occurs when the blood supply to a person's brain is interrupted or reduced. The stroke deprives person's brain of oxygen and nutrients, which can cause brain cells to die. Numerous works have been carried out for predicting various diseases by comparing the performance of predictive data mining technologies. In this work, we compare different methods with our approach for stroke prediction on the Cardiovascular Health Study (CHS) dataset. Here, decision tree algorithm is used for feature selection process, principle component analysis algorithm is used for reducing the dimension and adopted back propagation neural network classification algorithm, to construct a classification model. After analyzing and comparing classification efficiencies with different methods and variation models accuracy, our work has the optimum predictive model for the stroke disease with 97.7 percentage accuracy.

## Chapter 2

# DATA DESCRIPTION AND PREPROCESSING

---

### 2.1 Data Source:

A population of 5110 people are involved in this study with 2995 females and 2115 males. The dataset for this study is extracted from Kaggle data repositories (<https://www.kaggle.com/datasets>) to predict whether a patient is likely to get stroke based on the following information:

With;

Row Count: 5110

Column Count: 12

Female: 2995

Male: 2115

A	B	C	D	E	F	G	H	I	J	K	L	M
id	gender	age	hypertensi	heart_dise	ever_marr	work_type	Residence	avg_glucose	bmi	smoking_s	stroke	
9046	Male	67	0	1	Yes	Private	Urban	228.69	36.6	formerly s	1	
51676	Female	61	0	0	Yes	Self-emplc	Rural	202.21	N/A	never smo	1	
31112	Male	80	0	1	Yes	Private	Rural	105.92	32.5	never smo	1	
60182	Female	49	0	0	Yes	Private	Urban	171.23	34.4	smokes	1	
1665	Female	79	1	0	Yes	Self-emplc	Rural	174.12	24	never smo	1	
56669	Male	81	0	0	Yes	Private	Urban	186.21	29	formerly s	1	
53882	Male	74	1	1	Yes	Private	Rural	70.09	27.4	never smo	1	
10434	Female	69	0	0	No	Private	Urban	94.39	22.8	never smo	1	
27419	Female	59	0	0	Yes	Private	Rural	76.15	N/A	Unknown	1	
60491	Female	78	0	0	Yes	Private	Urban	58.57	24.2	Unknown	1	
12109	Female	81	1	0	Yes	Private	Rural	80.43	29.7	never smo	1	
12095	Female	61	0	1	Yes	Govt_job	Rural	120.46	36.8	smokes	1	
12175	Female	54	0	0	Yes	Private	Urban	104.51	27.3	smokes	1	
8213	Male	78	0	1	Yes	Private	Urban	219.84	N/A	Unknown	1	
5317	Female	79	0	1	Yes	Private	Urban	214.09	28.2	never smo	1	
58202	Female	50	1	0	Yes	Self-emplc	Rural	167.41	30.9	never smo	1	
56112	Male	64	0	1	Yes	Private	Urban	191.61	37.5	smokes	1	
34120	Male	75	1	0	Yes	Private	Urban	221.29	25.8	smokes	1	
27458	Female	60	0	0	No	Private	Urban	89.22	37.8	never smo	1	
25226	Male	57	0	1	No	Govt_job	Urban	217.08	N/A	Unknown	1	
70630	Female	71	0	0	Yes	Govt_job	Rural	193.94	22.4	smokes	1	
13861	Female	52	1	0	Yes	Self-emplc	Urban	233.29	48.9	never smo	1	
68794	Female	79	0	0	Yes	Self-emplc	Urban	228.7	26.6	never smo	1	
64778	Male	82	0	1	Yes	Private	Rural	208.3	32.5	Unknown	1	
4219	Male	71	0	0	Yes	Private	Urban	102.87	27.2	formerly s	1	
70822	Male	80	0	0	Yes	Self-emplc	Rural	104.12	23.5	never smo	1	
38047	Female	65	0	0	Yes	Private	Rural	100.98	28.2	formerly s	1	

Figure 2.1: Figure in Latex

## 2.2 ATTRIBUTE INFORMATION

```

1. id           : unique identifier
2. gender       : "Male", "Female" or "Other"
3. age         : age of the patient
4. hypertension : 0 if the patient doesn't have hypertension, 1
if the patient has hypertension
5. heart_disease : 0 if the patient doesn't have any heart
diseases, 1 if the patient has a heart disease
6. ever_married : "No" or "Yes"
7. work_type    : "children", "Govt_jov", "Never_worked",
"Private" or "Self-employed"
8. Residence_type : "Rural" or "Urban"
9. avg_glucose_level : average glucose level in blood
10. bmi         : body mass index
11. smoking_status : "formerly smoked", "never smoked", "smokes" or
"Unknown"
12. stroke      : 1 if the patient had a stroke, 0 the patient
do not have a stroke

```

## 2.3 DATA PREPROCESSING

On performing data cleaning we get that there are 201 missing values in BMI feature. A simple way to deal with the missing values is to remove the rows with null values, this may potentially remove data that aren't null. Thus, we will substitute missing values with mean of bmi and check if imputations are done.

## 2.4 ID

The total number of unique id is same as row count. We do not need another identifier. Thus, we will drop this column.

## 2.5 Gender

```
Female    2994
Male      2115
Other      1
Name: gender, dtype: int64
```

Gender needs to be categorized as binary variable. And from the analysis perspective, it will be tedious to create another variable for one row value ('Others'). Hence, we will impute this single value with mode in this column. Then we get the result as:

```
Female    2995
Male      2115
Name: gender, dtype: int64
```

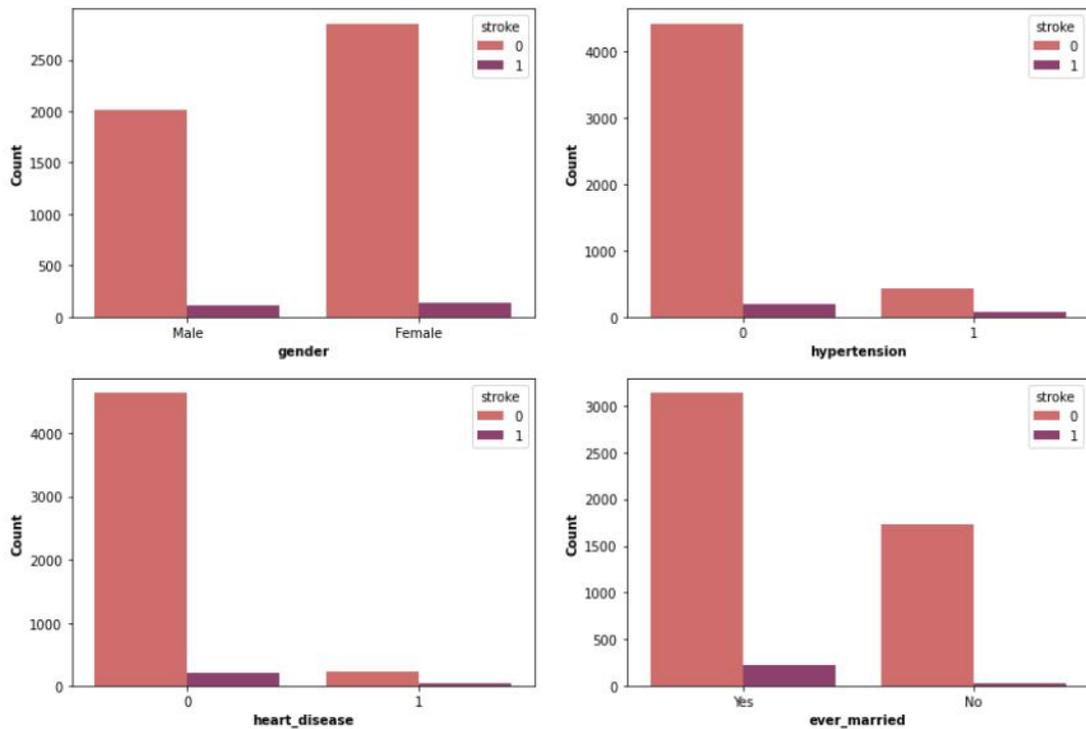
# Chapter 3

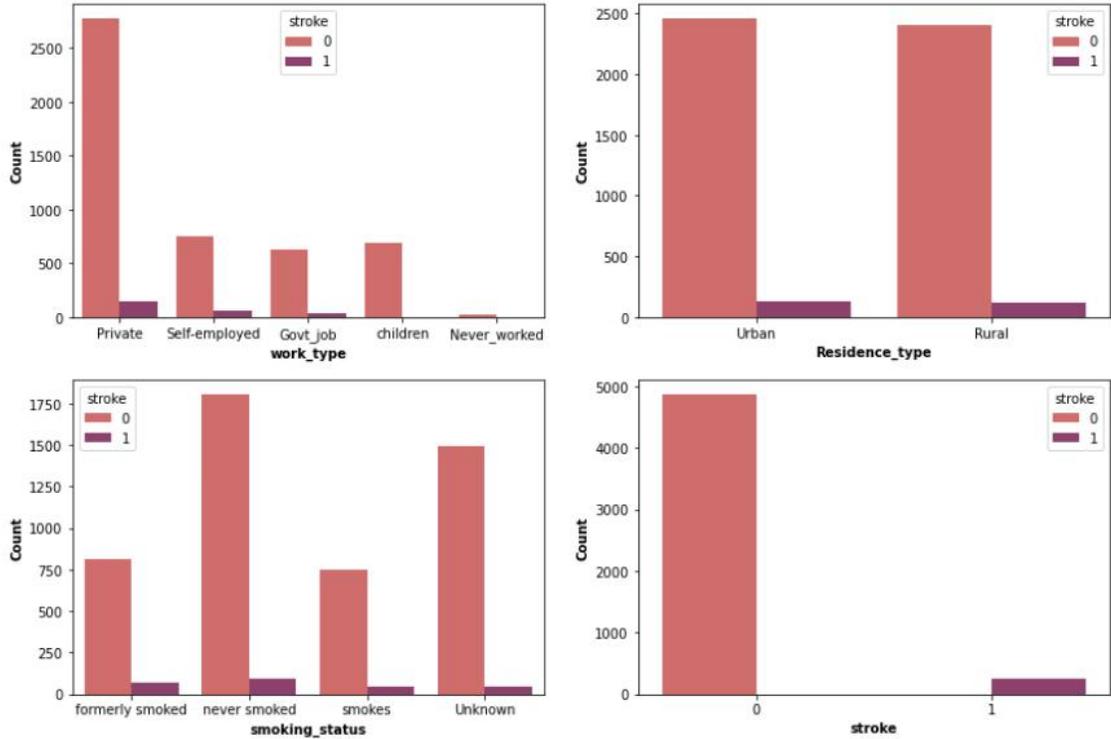
## DATA EVALUATION

---

### 3.1 EXPLORATORY DATA ANALYSIS

#### 1. Categorical Feature Analysis

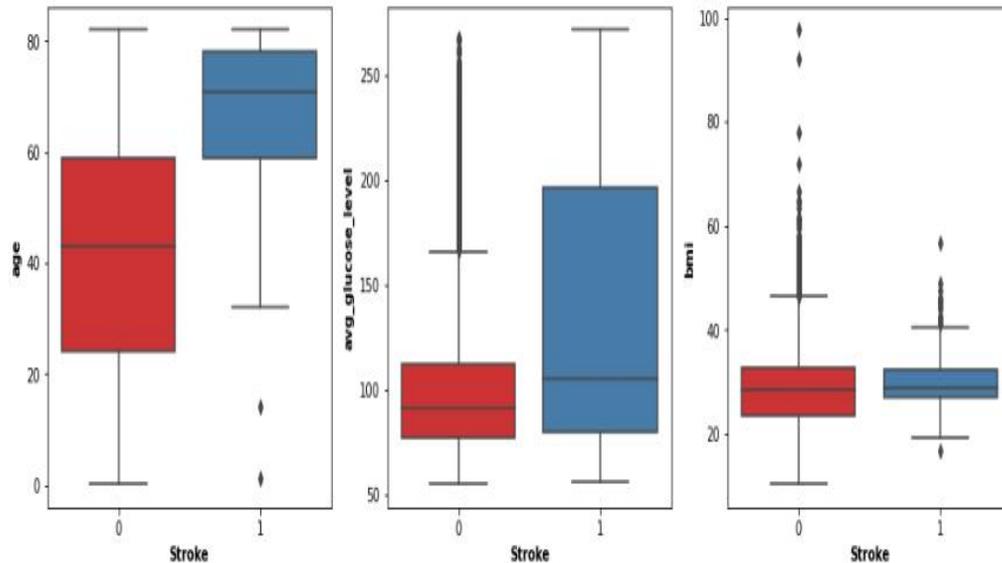




From the above count plot, some observations can be drawn:

- **hypertension:** Subjects that previously diagnosed with hypertension have highly risk of having stroke.
- **heart disease:** Subjects that previously diagnosed with heart disease have highly risk of having stroke.
- **ever married:** Subjects that ever married have highly risk of having stroke.
- **work type:** Subjects that have any work experience and in government related work have highly risk of having stroke while those with no work experience barely experienced stroke.
- **Residence type:** No obvious relationship with likelihood of experiencing stroke.
- **smoking status:** Being a smoker or former smoker increases risk of having a stroke.

### 3.2 Numerical Feature Analysis



From the above boxplot, some observations can be drawn:

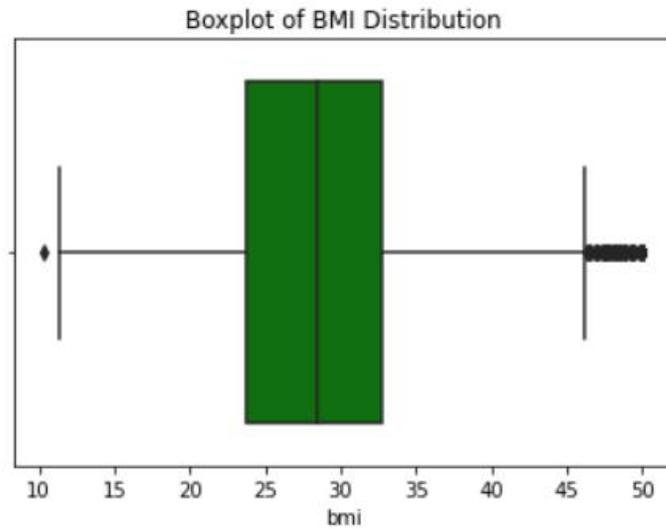
- age: Subjects with stroke tends to have higher mean age.
- ave glucose level: Subjects with stroke tends to have higher average glucose level.
- bmi: bmi index does not give much indication on the likelihood of experiencing stroke. bmi index for super obesity is 50. Outliers in this feature should be replaced to its highest limit (50).

The important results drawn are:

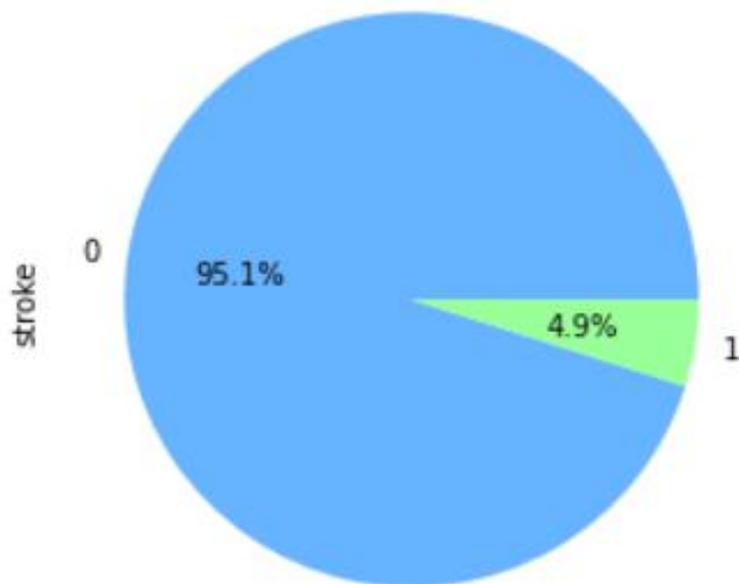
There are total 79 counts of outliers detected.

Replace values more than 50 in bmi column to 50.

By plotting boxplot of bmi column to confirm alterations are made.



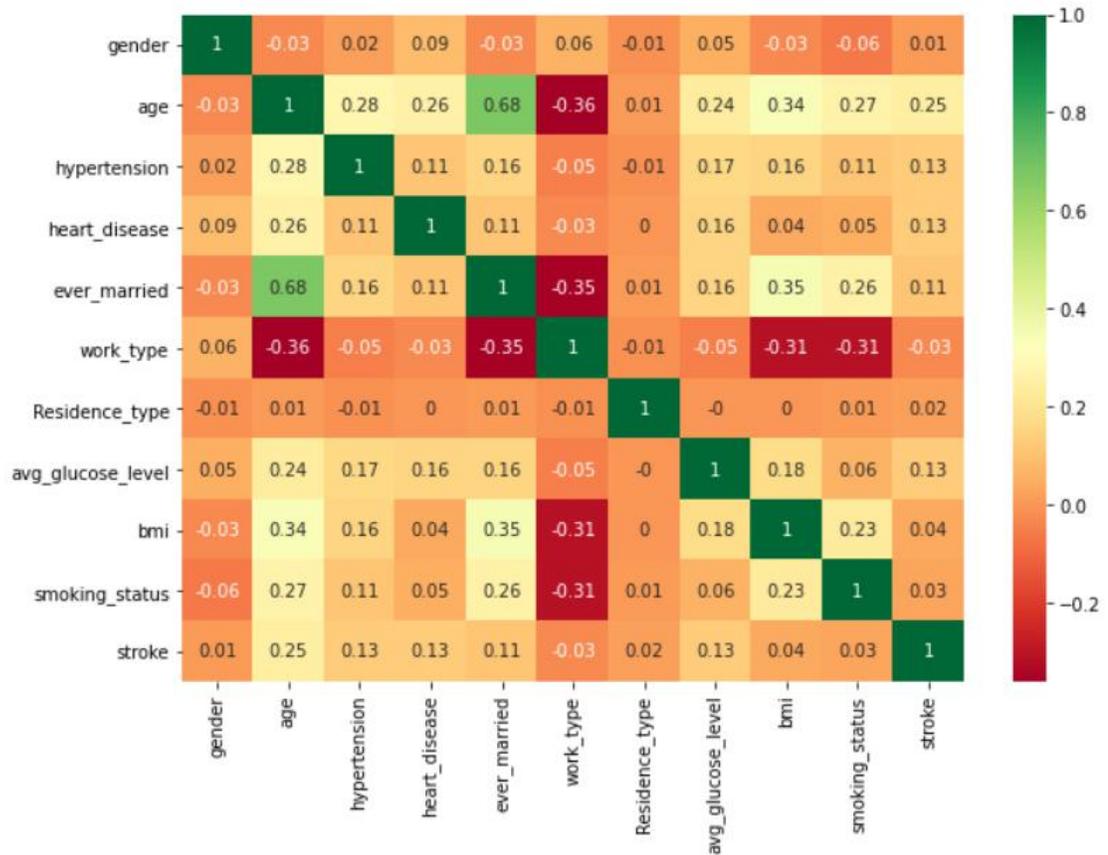
Pie Chart of Stroke Status



From the above pie chart we get that 4.9 percentage of the population in this dataset is diagnosed with stroke.

### 3.3 Multicollinearity Analysis

Since correlation check only accept numerical variables, preprocessing the categorical variables becomes a necessary step, we need to convert these categorical variables to numbers encoded to 0 or 1.



### 3.4 Final Preprocessing

Variables that are measured at different scales do not contribute equally to model fitting and might end up creating a bias. Thus, to deal with this potential problem feature standardization is usually used prior to

model fitting.

	<b>avg_glucose_level</b>	<b>bmi</b>	<b>age</b>
<b>0</b>	2.706375	1.066746	1.051434
<b>1</b>	2.121559	0.013363	0.786070
<b>2</b>	-0.005028	0.506346	1.626390
<b>3</b>	1.437358	0.766044	0.255342
<b>4</b>	1.501184	-0.655458	1.582163

	<b>gender</b>	<b>hypertension</b>	<b>heart_disease</b>	<b>work_type</b>	<b>Residence_type</b>	<b>smoking_status</b>	<b>stroke</b>	<b>avg_glucose_level</b>	<b>bmi</b>	<b>age</b>
<b>0</b>	1	0	1	2	1	1	1	2.706375	1.066746	1.051434
<b>1</b>	0	0	0	3	0	2	1	2.121559	0.013363	0.786070
<b>2</b>	1	0	1	2	0	2	1	-0.005028	0.506346	1.626390

## Chapter 4

# MODEL SELECTION

---

In this project we are using seven different algorithms to build the model. The 7 algorithms are:

- Decision Tree
- Logistic Regression
- Random Forest
- Support Vector Machine
- K Nearest Neighbour
- Naive Bayes
- KMeans Clustering

### 4.1 Decision Tree:

Decision tree is the most powerful and popular tool for classification and prediction. A Decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

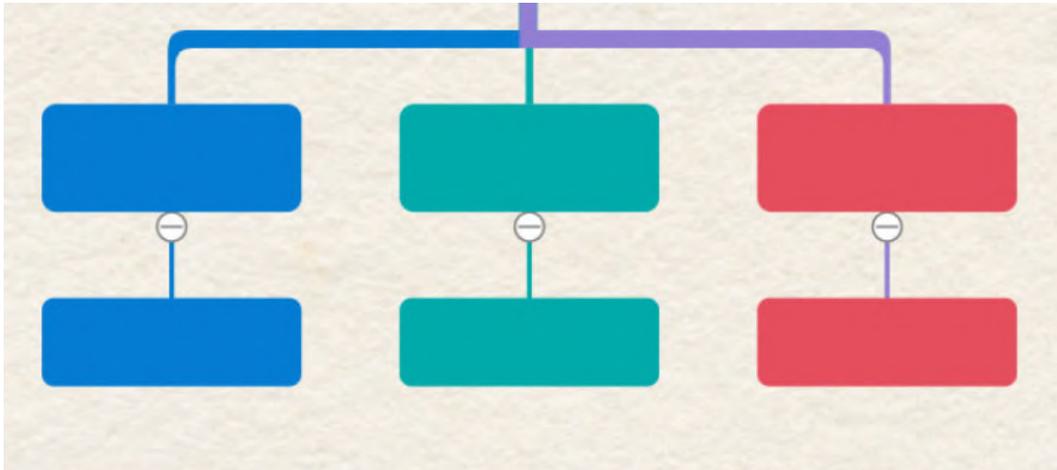
Decision tree is a flowchart-like structure in which each internal node represents a "test" on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules. In decision analysis, a decision tree and the closely related influence diagram are used as a visual and analytical decision support tool, where

the expected values (or expected utility) of competing alternatives are calculated.

A decision tree consists of three types of nodes:

1. Decision nodes – typically represented by squares
2. Chance nodes – typically represented by circles
3. End nodes – typically represented by triangles

Illustration of a decision tree

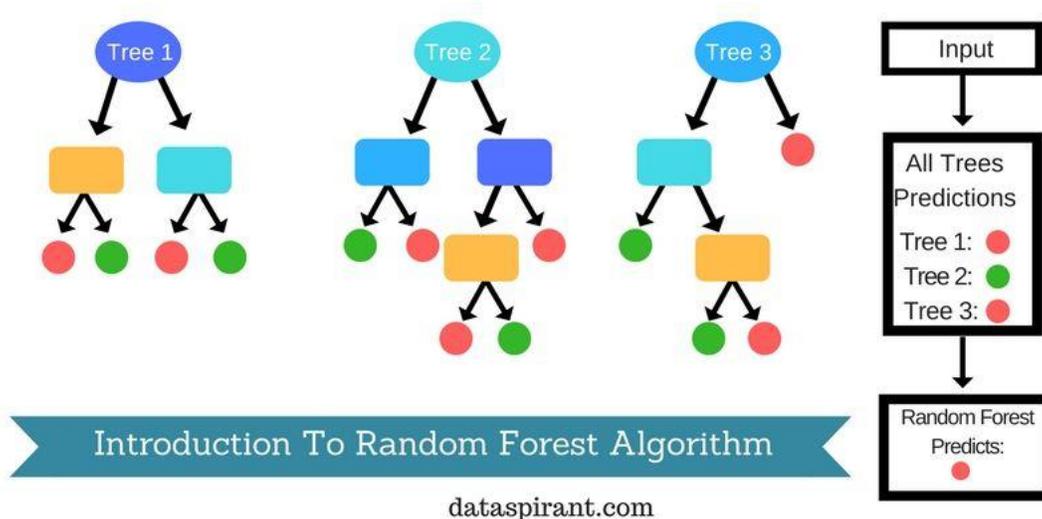


## 4.2 LOGISTIC REGRESSION MODEL

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables. Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1. Logistic Regression is much similar to the Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems.

### 4.3 Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting. The given figure is a diagrammatic representation of random forest



### 4.4 Support Vector Machine Algorithm

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems

in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.

#### 4.5 K-Nearest Neighbor(KNN) Algorithm

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique. K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems. K-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data. 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. KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

#### 4.6 Naive Bayes Classifier Algorithm

Naive Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional

training dataset. Naive Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object. Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles

**Naive:** It is called Naive because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other. **Bayes:** It is called Bayes because it depends on the principle of Bayes' Theorem.

**Bayes' Theorem:**

- Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.
- The formula for Bayes' theorem is given as:

$$P(A/B) = P(B/A)P(A) \div P(B)$$

Where,

$P(A/B)$  is Posterior probability: Probability of hypothesis A on the observed event B.

$P(B/A)$  is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

$P(A)$  is Prior Probability: Probability of hypothesis before observing the evidence.

$P(B)$  is Marginal Probability: Probability of Evidence.

## 4.7 K-Means Clustering Algorithm

K-Means Clustering is an unsupervised learning algorithm that is used to solve the clustering problems in machine learning or data science.

In this topic, we will learn what is K-means clustering algorithm, how the algorithm works, along with the Python implementation of k-means clustering. K-Means Clustering is an Unsupervised Learning algorithm, which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if  $K=2$ , there will be two clusters, and for  $K=3$ , there will be three clusters, and so on.

## Chapter 5

# MODEL BUILDING AND PREDICTION

---

All the predictor variables will be mapped to an array  $x$  and the target variable to an array  $y$ . The target variable is 'stroke' column.

### 5.1 PREDICTION:

From the above accuracy summary, Logistic Regression, Random Forest and KNN models all gives high accuracy score of 0.95. However, it is also important to consider the error type and recall value of each model. Models with 0.95 accuracy score generally have high false negative as shown in the confusion matrix. High false negative indicates type 2 error. For our study on stroke prediction, we want to avoid type 2 error as it means that we fail to identify subjects that has stroke and deem them stroke free instead. Inspecting from the classification report above, Naive Bayes Model has fit our objective although the accuracy is 0.87.

```
Decision Tree Model: 0.9159
Logreg Model: 0.955
Random Forest Model: 0.9543
Support Vector Machine Model: 0.9243
kNN Model: 0.9524
Naive Bayes Model: 0.8728
KMeans Model: 0.2172
```

## Chapter 6

# Cross Validation and Conclusion

---

### 6.1 Cross Validation

- Using the mean cross-validation, we can conclude that we expect the model to be around 87.31 percentage accurate on average.
- Our original model accuracy is 0.8728, but the mean crossvalidation accuracy is 0.8731. So, the 10-fold cross-validation accuracy does result in performance improvement for this model.

### 6.2 Conclusion

1. Various model was used to predict whether a person is subjected to stroke. Naive Bayes model yields a very good performance as indicated by the model accuracy which was found to be 87.28 percentage
2. Using the mean cross-validation, we can conclude that we expect the model to be around 87.31 percentage accurate on average.
3. Our original model accuracy is 87.28 percentage and the mean cross-validation accuracy is 87.31 percentage
4. Thus from the predictive analysis we can conclude that Naïve Bayes algorithm gives the most accurate value.

## REFERENCES

---

- [1] 1.Sonoda, Shigeru, et al. "Stroke outcome prediction using reciprocal number of initial activities of daily living status." *Journal of Stroke and Cerebrovascular Diseases* 14.1 (2005): 8-11.
- 2.Ridwan, Amiruddin, et al. "Primary prevention of stroke through development of mobile health application." *Enfermería Clínica* 30 (2020): 133-139.
- 3.Emon, Minhaz Uddin, et al. "Performance analysis of machine learning approaches in stroke prediction." 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2020.
- 4.Singh, M. Sheetal, and Prakash Choudhary. "Stroke prediction using artificial intelligence." 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON). IEEE, 2017.
- 5.Bhattacharya, Sweta, et al. "Antlion re-sampling based deep neural network model for classification of imbalanced multimodal stroke dataset." *Multimedia Tools and Applications* (2020): 1-25.

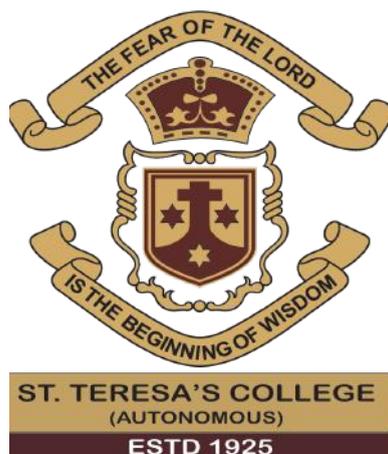
**Phytochemical screening, Functional groups and Elemental analysis of *Ipomoea cairica* (L.) subjected to abiotic stress in the industrial area of Eloor, Ernakulam district of Kerala**

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT  
FOR THE AWARD OF DEGREE OF  
“BACHELOR OF SCIENCE” IN

**BOTANY**

By

**NAME (REG.NO)**



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

## CERTIFICATE

This is to certify that the dissertation entitled " Phytochemical screening, Functional groups and Elemental analysis of *Ipomoea cairica* (L.) subjected to abiotic stress in the industrial area of Eloor, Ernakulam district of Kerala " submitted by Elizabeth Rithu Sunil (Reg.No Ab19bot005) as a part of the B.Sc degree course for the year 2021-2022 has been carried out under the supervision and guidance of Dr. Asha D , Department of Botany, St. Teresa's college Ernakulam.

Dr.Asha.D  
Supervising Teacher  
St. Teresa's college ,Ernakulam



Dr.Liza Jacob  
Head of the Department  
St. Teresa's college ,Ernakulam

Place: Ernakulam

Date: 9-5-2022

Examiner: Dr Anula N 9/5/22.

Anisha S 9/5/22.

## ACKNOWLEDGEMENT

First and foremost, I thank the God almighty for endowing me with an immense blessing which helped me to overcome the hurdles, paving the way for a successful completion of the study.

I wish to express my gratitude to Dr Liza Jacob, Head of the Department of Botany, St Teresa's College (Autonomous), Emakulam for providing all the necessary facilities for the work.

I thank my teacher and guiding light, Dr. Asha D. without whose guidance I would not be able to complete this project work.

I am also indebted to Dr. Lizzy Mathew, Principal of St. Teresa's College (Autonomous), Emakulam for providing all facilities for this project work.

I am greatly thankful to Dr. Alphonsa Vijaya Joseph, Vice Principal of St. Teresa's College (Autonomous), Emakulam for her constant support and help.

I also acknowledge my sincere thanks to the Sophisticated Test and Instrumentation Centre, CUSAT, Kochi, for the valuable help rendered throughout completion of the study.

I use this opportunity to convey my special thanks to all teaching and non-teaching staff of the Department of Botany, St. Teresa's College (Autonomous), Ernakulam, for their valuable suggestions and guidance during the course of work.

I acknowledge my sincere gratitude to my parents, and all my friends for their co-operation and encouragement.

Place : Ernakulam

NAME

Date :

## CONTENTS

<b>CHAPTE R</b>	<b>TITLE</b>	<b>PAGE NUMBER</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
<b>2</b>	<b>REVIEW OF LITERATURE</b>	<b>7</b>
<b>3</b>	<b>MATERIALS &amp; METHODS</b>	<b>9</b>
<b>4</b>	<b>RESULTS</b>	<b>15</b>
<b>5</b>	<b>DISCUSSION</b>	<b>21</b>
<b>6</b>	<b>SUMMARY &amp; CONCLUSION</b>	<b>23</b>
<b>7</b>	<b>REFERENCES</b>	<b>24</b>

## INTRODUCTION

Knowledge of herbs has been handed down from generation to generation for thousands of years. The revival of interest in natural drugs started in last decade mainly because of the widespread belief that green medicine is healthier than synthetic products. In the recent past, there has been a tremendous increase in the use of plant-based health products in developing as well as developed countries resulting in an exponential growth of herbal products globally.

According to the WHO, about 80% of the population in the world rely on the traditional medicine for the treatment of various diseases (Padmaa, et al., 2010). However, due to over population, urbanization, and continuous exploitation of these herbal reserves, the natural resources along with their related traditional knowledge are depleting day by day (Pande et al., 2007).

Most of the medicinal plant parts are used as raw drugs and they possess varied medicinal properties (Mahesh and Satish 2008). Medicinal plants are easily available, less expensive and also have no side effects (Cathrine et al., 2011). Medicinal plants represent a rich source of the antimicrobial agent (Mahesh et al., 2008).

In the present era of drug development and discovery of newer drug molecules, many plant products are evaluated on the basis of their traditional uses. In this regard, one of the many plants which are being evaluated for their therapeutic efficacies is *Ipomoea cairica* (Convolvulaceae) which is a perennial herb of unknown origin, possibly tropical Africa and Asia (Austin and Huaman, 1996); widely cultivated, now distributed nearly pantropically. Grows in disturbed sites, such as roadsides and waste-ground in urban areas, and is invasive of natural habitats, especially along river banks and coastal dunes.

Environment can be defined as a sum total of all the living and non-living elements and their effects that influence human life. While all living or biotic elements are animals,

plants, forests, fisheries, and birds, non-living or abiotic elements include water, land, sunlight, rocks, and air. Environmental science is an interdisciplinary academic field that integrates physical, biological, and geography (including ecology, biology, physics, chemistry, plant science, zoology, mineralogy, oceanography, limnology, soil science, geology and physical geography, and atmospheric science) to the study of the environment, and the solution of environmental problems. Environmental science emerged from the fields of natural history and medicine during the Enlightenment. Today it provides an integrated, quantitative, and interdisciplinary approach to the study of environmental systems.

Nature is a synthesis of the five elements air, water, soil, fire and earth and man has been living in balanced harmony with nature since time immemorial. But the changes which man has brought about in his environment were largely determined by his necessities, his knowledge and his value. (Ajantha et al., 1989). These are undoubtedly being polluted by the dust, smoke, sewage, sullage, industrial discharge and city refuse. (Arun, 1999).

As a result of rain, dew or fog, leaves and other aerial organs of plants often become wet. In most species, the turgidity of leaves is restored from the wilted condition after immersion of water for 24 hour or less. Water largely entered through the epidermal cells, through hairs and specialized epidermal cells can provide for such entry in some cases. On prolonged immersion the leaves of many species show opening of stomata and some time also the injection the intercellular spaces with liquid water under certain condition, leaves and other aerial parts are capable of absorbing water from the air, it has been reported that pine seedlings of *Pinus ponderosa* in dry soil can absorb water vapour from the near saturated or unsaturated atmosphere (Subhash, 1980).

Human population size has grown enormously over the last hundred years. This means an increase in demand for food, water, home, electricity, roads, automobiles and numerous other commodities. These demands are exerting tremendous pressure on our natural resources and are also contributing to pollution of air, water and soil. The need of the hour is to check the degradation and depletion of our precious natural resources and pollution without halting the process of development.

## **Macronutrients and Micronutrients in Plants**

The essential elements can be divided into macronutrients and micronutrients. Nutrients that plants require in larger amounts are called macronutrients. About half of the essential elements are considered macronutrients: carbon, hydrogen, oxygen, nitrogen, phosphorus, potassium, calcium, magnesium, and sulphur. The next-most-abundant element in plant cells is nitrogen (N); it is part of proteins and nucleic acids. Nitrogen is also used in the synthesis of some vitamins. Hydrogen and oxygen are macronutrients that are part of many organic compounds and also form water. Oxygen is necessary for cellular respiration; plants use oxygen to store energy in the form of ATP. Phosphorus (P), another macromolecule, is necessary to synthesise nucleic acids and phospholipids. As part of ATP, phosphorus enables food energy to be converted into chemical energy through oxidative phosphorylation. Light energy is converted into chemical energy during photophosphorylation in photosynthesis; and into chemical energy to be extracted during respiration. Sulphur is part of certain amino acids, such as cysteine and methionine, and is present in several coenzymes. Sulphur also plays a role in photosynthesis as part of the electron transport chain where hydrogen gradients are key in the conversion of light energy into ATP. Potassium (K) is important because of its role in regulating stomatal opening and closing. As the openings for gas exchange, stomata help maintain a healthy water balance; a potassium ion pump supports this process.

Magnesium (Mg) and calcium (Ca) are also important macronutrients. The role of calcium is twofold: to regulate nutrient transport and to support many enzyme functions. Magnesium is important to the photosynthetic process. These minerals, along with the micronutrients, also contribute to the plant's ionic balance.

In addition to macronutrients, organisms require various elements in small amounts. These micronutrients, or trace elements, are present in very small quantities. The seven main micronutrients include boron, chlorine, manganese, iron, zinc, copper, and molybdenum. Boron (B) is believed to be involved in carbohydrate transport in plants; it also assists in metabolic regulation. Boron deficiency will often result in bud dieback. Chlorine (Cl) is necessary for osmosis and ionic balance; it also plays a role in photosynthesis. Copper (Cu) is a component of some enzymes. Symptoms of copper deficiency include browning of leaf tips

and chlorosis (yellowing of the leaves). Iron (Fe) is essential for chlorophyll synthesis, which is why an iron deficiency results in chlorosis. Manganese (Mn) activates some important enzymes involved in chlorophyll formation. Manganese-deficient plants will develop chlorosis between the veins of its leaves. The availability of manganese is partially dependent on soil pH. Molybdenum (Mo) is essential to plant health as it is used by plants to reduce nitrates into usable forms. Some plants use it for nitrogen fixation; thus, it may need to be added to some soils before seeding legumes. Zinc (Zn) participates in chlorophyll formation and also activates many enzymes. Symptoms of zinc deficiency include chlorosis and stunted growth.

Deficiencies in any of these nutrients, particularly the macronutrients, can adversely affect plant growth. Depending on the specific nutrient, a lack can cause stunted growth, slow growth, or chlorosis. Extreme deficiencies may result in leaves showing signs of cell death. (Boundless, 2021).

### **Medicinal plant selected for the study**

*Ipomoea* is the largest genus in the flowering plant family Convolvulaceae. This family comprises of plants with high industrial, pharmaceutical, scientific, and cultural significance. It is distributed across continents and has its presence in Senegal, Nigeria, tropical Africa and into Asia, Mascarene Island, Malaysia and Australasia. Typically, *Ipomoea* is used in folk medication for the treatment of hemorrhoids, diabetes, bronchitis and arthritis (Shubhangi and Patil, 2004). The most widespread common name is morning glories. The genus includes food crops; the tubers of sweet potatoes (*Ipomoea batatas*) and the leaves of water spinach (*I. aquatica*) are commercially important food items and have been for millennia. The water spinach (*I. aquatica*) commonly found in ponds, cultivated for stem and leaves used as vegetables.

### **Habit and habitat**

Perennial twiner with tuberous root stock. Terrestrial, common among bushes, hedges of gardens, waste places and outskirts of forests. Also cultivated occasionally in the gardens, parks and railway platforms.

## **Morphological characters**

*Ipomoea cairica* comes under the family Convolvulaceae. Perennial twiner with tuberous root-stock; The slender stems are hairless; Leaves palmately 5 to 7-partite; segments elliptic - obovate or lanceolate, narrowed at both ends, retuse, mucronate at apex, glabrous; Flowers in 1 to 3-flowered cymes; Calyx-segments unequal, ovate, mucronulate, tuberculate on the back of outer ones; Corolla 6 -7 cm long, white or purple; Capsules 2 -celled, 4 -valved; Seeds pubescent.

This plant reproduces vegetatively by rooting along its stems and also produces seeds. Stem fragments and seeds are often dispersed in dumped garden waste and can also be spread by water (Barker, 2005).

Some are simply weeds, whereas others are economically important, viz., sweet potatoes and still others are ornamental plants, such as the morning glories *I. purpurea* and *I. cairica*, respectively. The genus *Ipomoea* including *Ipomoea cairica* and *Ipomoea palmate* has been reported to have many biological activities (Teow et al., 2007). Several species of *Ipomoea* have been used as medicines. Studies have shown the potential of the genus as a source of therapeutic agents. The glycoresins within them constitute one important chemotaxonomic marker of this family (Wagner, 1993).

## **Locality of Sample Collection**

Eloor is a suburb of Kochi and a municipality in Paravur Taluk, Ernakulam District in the Indian state of Kerala, India. It is an industrial area situated around 13 kilometres (8.1 miles) north of the city centre. It is an island of 14.21 km<sup>2</sup> formed between two distributaries of river Periyar and is the largest industrial belt in Kerala. The neighbouring places of Eloor are Kalamassery industrial hub, Aluva, Cheranalloor and Paravur. There are various companies of different kinds along the industrial belt including Fertilisers and Chemicals Travancore (FACT), Indian Rare Earths Limited, Hindustan Insecticides Limited and many others manufacturing a range of products like chemical-petrochemical products, rare-earth elements, rubber-processing chemicals, fertilizers, zinc/chromium compounds and leather products. The industrial belt of Eloor in Kerala is one of the world's 'top toxic hot spots', according to international environment group Greenpeace.

The aim of the study was to conduct phytochemical study, functional group and elemental analysis of *Ipomoea cairica* growing in an industrial area.

## **OBJECTIVES**

To detect the phytochemical constituents present in the ethanol extract of *Ipomoea cairica* by qualitative phytochemical analysis.

To evaluate the functional groups present in the species using FTIR technique.

To analyse the elemental composition of *Ipomoea cairica* growing in a locality experiencing industrial pollution using SEM-EDX.

## REVIEW OF LITERATURE

- o Calixto (2005) collected the information from WHO and reported that 65% - 80% of the world's population in developing countries depend on the medicinal plants for their primary health care due to the poverty and lack of access to modern medicine.
- o *Ipomoea cairica* is a perennial fast-growing vine. This species is adapted to grow in area with extreme seasonal fluctuations and because of its tuberous roots, it has been observed that the top part of the plant may die and later resprouts when environmental conditions are favourable (Weber, 2003; Maimela and Gumede, 2019).
- o Eames and Mac (1947) did a study on the number of stomata per unit area of epidermis variously greatly.
- o Jyoti et al., (2008) analysed the amount of Pb, Cd, Cr, and Ni in soil and plant samples of *Abutilon indicum*, *Calotropis procera*, *Euphorbia hirta*, *Peristrophe bycaliculata*, and *Tinospora cordifolia* collected from 3 environmentally different sites.
- o Sutapa et al., (2015) investigated the methanol extracts of leaves and flowers of *Ipomea cairica* for antimicrobial and antioxidant effects. Results showed very good activity against all bacterial and fungal strains tested. The extract also showed remarkable antioxidant activity, with DPPH scavenging activities correlating with its reductive potential.
- o A Study of *Ipomoea cairica* extract by Shefali et al. (2013) showed dose dependent reduction of response in the formalin test inflammatory phase in mice. A possible explanation for the antinociceptive activity is the release of pro-nociceptive mediators unrelated to carrageenan-induced edema. The isolated caffeoylquinic acids could explain, in part, the antinociceptive effect.
- o Flavio et al (2006) evaluated the in-vitro antioxidant activities of methanolic extracts of leaves and flowers by DPPH radical inhibition. The methanolic extract of *I. cairica* leaves showed maximum antioxidant activity (83.52%)

while the methanolic extract of flowers showed maximum antioxidant activity at 81.85% at 500 µg/ml concentrations.

- o The methanol extract of (MEIP) flowering tops showed antioxidant activity by inhibiting DPPH and hydroxyl radical, nitric oxide and super oxide anion scavenging, hydrogen peroxide scavenging, and reducing power activities (Dudharejia and Shah, 2009).
- o Aqueous methanol extract of *Ipomoea cairica* leaves possess a strong anti-inflammatory activity (Mohamed and Karawya, 2010).
- o The major constituents of the extract were the coumarins, scopoletin and umbelliferone and the lignans, arctigenin, matairesinol and trachelogenin (Olga *et al.*, 1997; Sharda and Kokate, 1979). Indole alkaloids were isolated from the leaves of this species (Mohamed and Karawya, 2010).
- o From the aerial parts of *Ipomoea cairica* the coumarins umbelliferone and scopoletin, and the dibenzyl-g-butyrolactone lignans arctigenin, matairesinol and trachelogenin, were isolated along with  $\beta$ -sitosterol and fatty acids (Singh *et al.*, 2013).
- o (+)-(8*R*,8'*S*)-thujaplicatin methyl ether, arctigenin, matairesinol, *trans*-2,3-dibenzyl butyrolactone, vanillic acid, hydroxybenzoic acid, methoxybenzoic acid, methylparaben, stearic acid, palmitic acid, oleic acid, friedelinol and a mixture of  $\beta$ -sitosterol and stigmasterol were obtained from the methanolic extract of the *Ipomoea cairica* (Ralte, 2014).
- o The crystallite domain size was calculated from the width of the XRD peaks, assuming that they are free from non-uniform strains, using the Debye-Scherrer formula (Cullity, 1978). Functional groups present in the bioengineered nanoparticles and interactions with protein were analysed by transform infrared (FTIR) analysis (Najitha and Balasubramanian 2014b).
- o The hydroxyl and carboxyl groups present in the plants were able to bind with the metals and the flavonoids and phenols have exclusive power to warp nanoparticles to avoid the agglomeration (Ahmad *et al.*, 2010).

## MATERIALS AND METHODS

### Plant selected for the study

*Ipomoea cairica* is a vining, herbaceous, perennial plant with palmate leaves and large, showy white to lavender flowers. A species of morning glory, it has many common names, including mile-a-minute vine, Messina creeper, Cairo morning glory, coast morning glory and railroad creeper.

### Scientific Name

*Ipomoea cairica* (L.) Sweet

### Synonyms

*Convolvulus cairicus* L.

*Convolvulus pendulus* (R. Br.) Spreng.

*Ipomoea palmata* Forssk.

*Ipomoea pendula* R. Br.

*Ipomoea tuberculata* (Desr.) Roem. & Schult.

### Systematic Position

- Kingdom: Plantae
- Division : Phanerogamae
- Class: Dicotyledons
- Subclass : Gamopetalae
- Series : Bicarpellatae
- Order: Solanales
- Family: Convolvulaceae
- Genus: Ipomoea
- Species : *Ipomoea cairica*

**Common names :** 'Mile-a-minute vine', 'Messina creeper', 'Cairo Morning glory', 'Coast Morning glory', 'Railroad creeper'.

**Malayalam Name:** Kolambi Poo

### **Botanical Characteristics:**

One of the commonest yet most useful of the evergreen creepers, refreshing the eye in the hottest weather with its clear, green leaves and delicate, mauve blooms, the Railway Creeper is found in gardens, villages, and on practically every railway station, thus earning for itself its nickname. This morning glory vine is beautiful, climbing on to whatever it finds - the purple flower studded vine wrapped around bending bamboo stems, is a pleasing sight. Its stem is hairless, readily set roots when in touch with the earth. Flowers purple, pink or rarely pinkish white, to 8cm across, solitary or in groups of 2-3. Fruit a 4-valved capsule, about 1cm across, each valve with 1 seed. Seed with wispy hairs attached. Spread by wind, water and humans.

### **Plant collection**

The whole plant of *Ipomoea cairica* (L.) used for the investigation was obtained from Eloor industrial area, Ernakulam, Kerala. The plant specimen was authenticated using the Flora of presidency of Madras. Fresh plant material was washed under running tap water, then air dried and powdered.

### **Extraction**

The plant powder was extracted with ethanol using cold extraction. Totally 30 g of dried plant powder was extracted in 250 ml of ethanol for 24 hours in occasional shaking at room temperature. The supernatant was collected and evaporated to make the final volume one-fifth of the original volume. It was stored at 4°C in airtight bottles for further studies.

## Phytochemical analysis

The ethanolic extract of *Ipomoea cairica* was subjected to preliminary phytochemical screening for their presence or absence of active constituents utilising a standard method of analysis.

### 1. Test for Alkaloids

**Mayer's Test:** 2 ml test sample was treated with few drops of Mayer's reagent along sides of the test tube and the formation of white or creamy precipitate indicated the presence of alkaloids.

(Mayer's Reagent- 1.358g of  $\text{HgCl}_2$  was dissolved in 60ml of water and poured into a solution of 5g of potassium iodide in 10ml of distilled water and made upto 100 ml of reagent).

**Dragendorff's Test:** Test sample was treated with 1-2 ml of Dragendorff's reagent and the formation of prominent reddish brown precipitate indicated the presence of alkaloids.

(Dragendorff's Stock reagent- Solution (A): 0.85g basic bismuth nitrate dissolved in 10ml glacial acetic acid and 40 ml water under heating. Solution (B): 8g potassium iodide dissolved in 30 ml water. Stock solution: A and B mixed in 1:1 ratio).

**Wagner's Test:** Test sample was treated with few drops of Wagner's reagent along sides and the formation of reddish brown precipitate indicated the presence of alkaloids.

(Wagner's reagent - 2 g Iodine and 6 g potassium Iodide dissolved in 5 ml distilled water and made upto 100 ml.)

### 2. Test for Flavonoids

**Sulphuric acid test:** A fraction of the extract was treated with concentrated  $\text{H}_2\text{SO}_4$  and observed for the formation of orange colour.

**NaOH test:** A small amount of extract was treated with aqueous NaOH and HCl, observed for the formation of yellow orange colour.

### 3. Test for Amino acids

**Ninhydrin Test:** Extract solution was treated with Ninhydrin (Triketohydrindene hydrate) at the pH range of 4 - 8. Development of purple colour indicated the positive response for amino acids.

### 4. Test for Reducing sugars

**Fehling's test for free reducing sugar:** About 0.5 g of extract was dissolved in distilled water and filtered. The filtrate was heated with 5 ml of equal volumes of Fehling's solution A and B. Formation of a red precipitate of cuprous oxide was an indication of the presence of reducing sugars.

**Benedict's Test:** To 5 ml of the extract solution, 5 ml of Benedict's solution was added in a test tube and boiled for a few minutes. Development of brick red precipitate confirmed the presence of reducing sugars.

### 5. Test for Anthraquinones

**Modified Borntrager's Test:** 5 ml of extract solution was hydrolyzed with dilute sulphuric acid and extracted with benzene. 1 ml of dilute ammonia was added to it. Rose pink coloration suggested the positive response for anthraquinones.

### 6. Test for Saponins

**Foam Test:** A small amount of extract was shaken with water and looked for the formation of persistent foam.

### 7. Test for Sterols

**Liebermann-Burchard test:** One ml extract was treated with chloroform, acetic anhydride and added drops of H<sub>2</sub>SO<sub>4</sub> and observed for the formation of dark pink or red colour.

**Sulphuric acid test:** The fraction of extract was treated with ethanol and H<sub>2</sub>SO<sub>4</sub> and observed for the formation of violet blue or green colour.

## 8. Test for Tannins

**FeCl<sub>3</sub> Test:** 5 ml of extract solution was allowed to react with 1 ml of 5% ferric chloride solution. Greenish black colouration indicated the presence of tannins.

**Potassium Dichromate Test:** 5 ml of the extract was treated with 1 ml of 10% aqueous potassium dichromate solution. Formation of yellowish brown precipitate suggested the presence of tannins.

**Lead acetate Test:** 5 ml of the extract was treated with 1 ml of 10% lead acetate solution in water. Yellow colour precipitation gave the test for tannins.

## 9. Test for Phenols

**Ferric chloride test:** A fraction of extract was treated with 5% ferric chloride, formation of deep blue colour confirms the presence of phenol

**Liebermann's test:** The extract was heated with sodium nitrite, add H<sub>2</sub>SO<sub>4</sub> solution diluted with water and add excess of dilute NaOH and observed for the formation of deep red or green or blue colour.

## 10. Test for Glycosides

**Legal's test:** Dissolved the extract (0.1 g) in pyridine, added sodium nitroprusside reagent and made alkaline with NaOH solution. Pink to red colour solution indicates the presence of glycosides.

**Borntrager's test:** The extract is hydrolyzed with concentrated HCl for 2 hours on a water bath and filtered and few ml of above filtrate was shaken with chloroform, chloroform layer was separated and added 10 % ammonia, formation of pink colour indicates the presence of glycosides

## 11. Test for Gums

**Molisch's Test:** 2 ml of concentrated sulphuric acid was added to 2 ml of extract solution. Then it was treated with 15%  $\alpha$ -naphthol in ethanol (Molisch's reagent).

Formation of a red violet ring at the junction of two layers indicated the positive test for gums.

## 12. Test for Terpenoids

**Chloroform test:** The plant extract was taken in a test tube with few ml of chloroform and add concentrated sulphuric acid carefully to form a layer and observed for presence of reddish brown colour.

**Liebermann-Burchard test:** 1 ml extracts was treated with chloroform, acetic anhydride and added drops of H<sub>2</sub>SO<sub>4</sub> and observed for the formation of dark green colour.

## FTIR spectrum analysis

The ethanol extract of *Ipomoea cairica* (L.) was mixed with KBr salt, using a mortar and pestle, and compressed into a thin pellet. Infrared spectra were recorded on a Shimadzu FTIR Spectrometer 8000 series, between 4,000-400 cm<sup>-1</sup>.

## Scanning Electron Microscopy (SEM) with Energy Dispersive X-Ray Analysis (EDX)

The ethanol extract derived from plant sample of *Ipomoea cairica* (L.) was subjected to the elemental analysis using Scanning Electron Microscope (SEM) with an energy dispersive x-ray spectrometer (EDX). For doing the elemental analysis, Scanning Electron Microscope (with EDX) was used with a model number Jeol 6390LV having accelerating voltage: 0.5 kV to 30 kV and resolution at 4 nm (30 kV) which provides images at magnification: 300,000. SEM provides detailed high resolution images of the sample by rastering a focussed electron beam across the surface and detecting secondary or backscattered electron signal. An Energy Dispersive X-Ray Analyzer (EDX or EDA) is also used to provide elemental identification and quantitative compositional information. SEM provides images with magnifications up to ~X50,000 allowing sub micron-scale features to be seen i.e. well beyond the range of optical microscopes. It is very rapid, high resolution imaging with identification of elements present in the plant samples.

## RESULTS

### Phytochemical screening of plant materials

The phytochemical screening of *Ipomoea cairica*(L.) is displayed in Table 1. Preliminary phytochemical screening of twelve different secondary metabolites were done in the ethanol extract of *Ipomoea cairica*. The ethanol extract showed the presence of the maximum number of secondary metabolites in the plant.

### Fourier transform infrared (FTIR) fingerprinting analysis

Figure 2 shows the results of FTIR spectroscopic studies which revealed the presence of various functional groups in ethanol leaf extract of *Ipomoea cairica* (L.). The peak at 3989 and 3385.07  $\text{cm}^{-1}$  is corresponding to hydrogen - bonded O-H stretching frequency respectively. The peak at 3385.07  $\text{cm}^{-1}$  that peak covers the entire region with a very broad peak. The peak at 2929.87  $\text{cm}^{-1}$  to assign H-C-H stretch. The peak at 2223.92  $\text{cm}^{-1}$  to assign  $\text{C}\equiv\text{N}$  stretch. The absorptions can be seen as several distinct peak in this region. The more intense bands occurring at 2200.78, 2181.49, 2152.56, 2133.27 corresponding to  $\text{C}\equiv\text{C}$  stretching indicate the presence of alkynes.

The peak at 1710.86  $\text{cm}^{-1}$  to assign  $\text{C}=\text{O}$  stretching indicate the carboxylic acid. The peak at 1627.97  $\text{cm}^{-1}$  to assign  $\text{C}-\text{C}=\text{C}$  stretching indicate the alkenes. The peak at 1263.37, 1055.06  $\text{cm}^{-1}$  assigned to the C-O stretching vibration and indicate the esters.

## Energy dispersive X-ray spectroscopy

Table 2 shows the results of the elemental composition of *Ipomoea cairica*(L.) using SEM and EDX technique. The SEM - EDX spectra of the ethanol extract of the *Ipomoea cairica*(L.) is shown in figure 3.

The topography of elements was confirmed through SEM images which is displayed in figure 4. Oxygen, Silicon, Chloride, Potassium and Calcium were present in the sample. Many standards were used like  $\text{CaCO}_3$ ,  $\text{SiO}_2$ , KCL, MAD, Ca wollastonites. In all these elements, Carbon and Oxygen were present as high concentrations while Potassium was seen as moderate amount. But P, Si, Cl and S were present only in trace quantities. Trace elements are estimated by determining the percentage abundance (%) of elements in the sample.

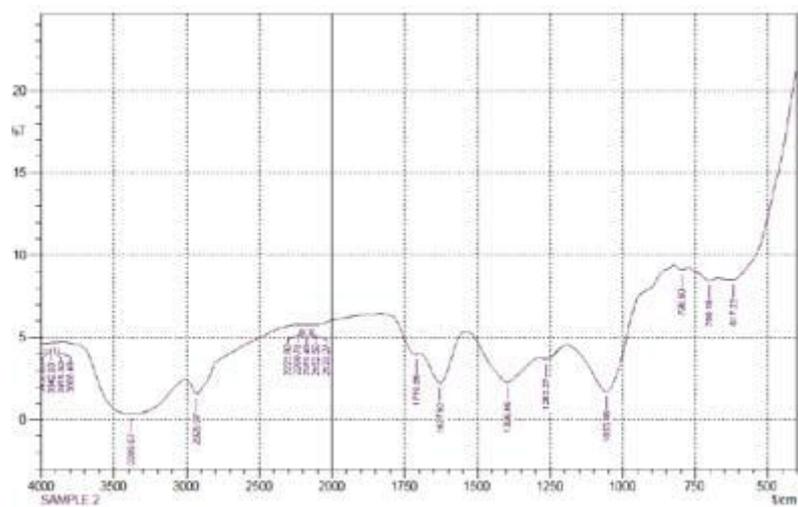


**Fig. 1: Medicinal plant *Ipomoea cairica* (L.) collected from Industrial area, Eloor**

<b>Plant Constituents</b>	<b>Test/ Reagent</b>	<b><i>Ipomoea cairica</i> (L.)</b>
<b>Alkaloids</b>	<b>Mayer's Test</b>	+
	<b>Dragendorff's Test</b>	+
	<b>Wagner's Test</b>	
<b>Phenols</b>	<b>Ferric chloride test</b>	+
	<b>Liebermann's test</b>	+
<b>Tannins</b>	<b>FeCl<sub>3</sub> Test</b>	+
	<b>Potassium Dichromate Test</b>	+
	<b>Lead acetate Test</b>	+
<b>Flavonoids</b>	<b>Sulphuric acid test</b>	+
	<b>NaOH test</b>	+
<b>Anthraquinones</b>	<b>Modified Borntrager's Test</b>	+
<b>Saponins</b>	<b>Foam Test</b>	+
<b>Steroids</b>	<b>Liebermann-Burchard test</b>	+
	<b>Sulphuric acid test</b>	+
<b>Terpenoids</b>	<b>Chloroform test</b>	+
	<b>Liebermann-Burchard test</b>	+
<b>Glycosides</b>	<b>Legal's test</b>	-
	<b>Borntrager's test</b>	-
<b>Reducing sugars</b>	<b>Benedict's Test</b>	-
	<b>Fehling's test for free reducing sugar</b>	-
<b>Amino acids</b>	<b>Ninhydrin Test</b>	-
<b>Gums</b>	<b>Molisch's Test</b>	-

**Presence: +, Absence: -**

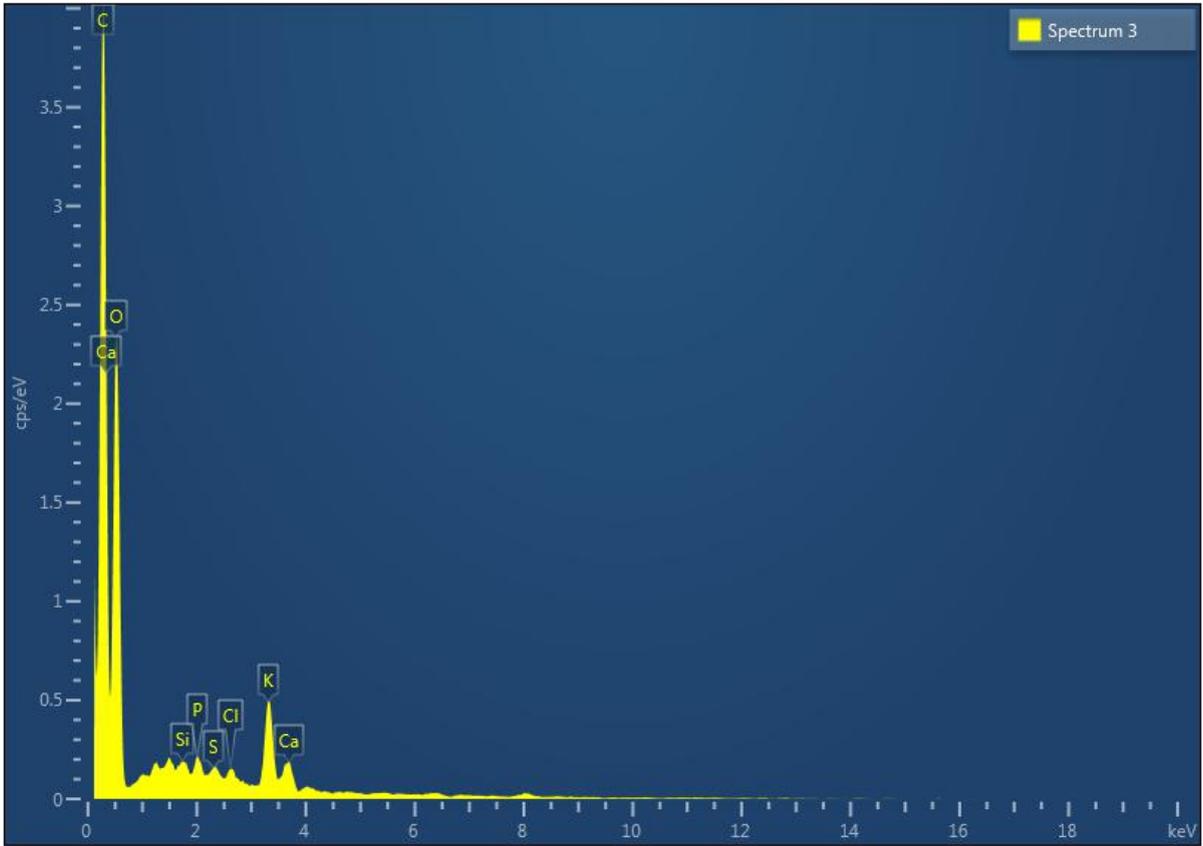
**Table 1: Phytochemical screening of leaf extract of *Ipomoea cairica* (L.)**



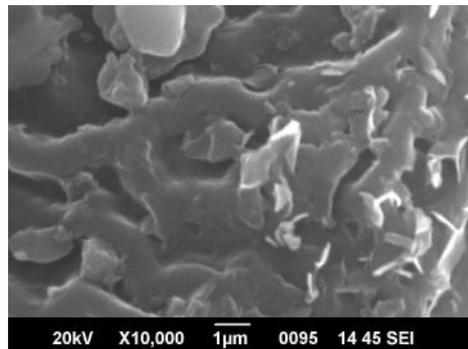
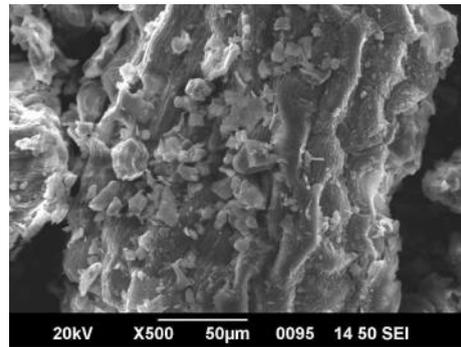
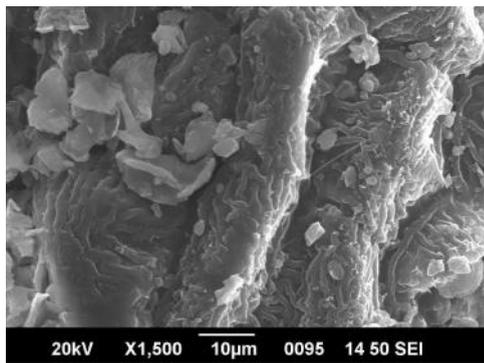
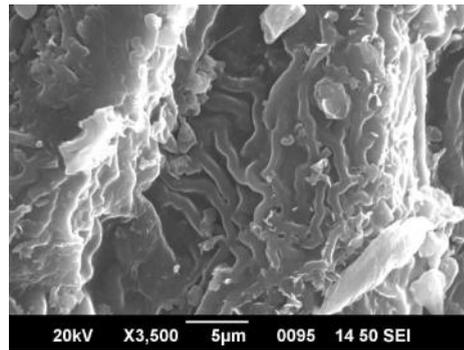
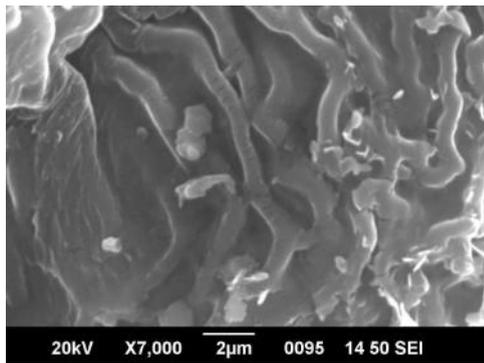
**Fig. 2: FTIR Spectrum analysis of Ethanol leaf extract of *Ipomoea cairica* (L.)**

Element	Line Type	Weight %	Atomic %
C	K series	57.57	65.39
O	K series	39.18	33.41
Si	K series	0.17	0.08
P	K series	0.34	0.15
S	K series	0.15	0.06
Cl	K series	0.23	0.09
K	K series	1.81	0.63
Ca	K series	0.55	0.19
TOTAL		100	100

**Table 2: The percentage of trace elements present in ethanol leaf extract of *Ipomoea cairica* (L)**



**Fig. 3: The SEM EDX spectra of ethanol leaf extract of *Ipomoea cairica* (L.)**



**Fig. 4:** The SEM EDX images of ethanol leaf extract of *Ipomoea cairica* (L.)

## DISCUSSION

*Ipomea cairica* is commonly known as 'Railway creeper' or Morning glory. The plants are medicinally used as an antioxidant, anti-inflammatory, antiviral, and highly potent against malaria (Arora *et al.*, 2013). Morning Glory is a tropical plant and is commonly found in tropical regions of Africa and Asia, including India. Most parts of the *Ipomoea cairica* have been recorded to be edible; the leaves are eaten when still young and roots are cooked before consuming. In some areas this plant is considered to have antibiotic properties. *Ipomoea cairica* is often used as a screening plant in the garden by allowing the vine to twist and climb along a trellis or fence, giving a beautiful display.

*Ipomoea cairica* (Coastal morning glory) is considered as a significant environmental weed commonly seen in India. This species is capable of very rapid growth and can completely smother trees and under-storey plants, but it will creep along the ground in the absence of supporting vegetation. Significant infestations may lead to a reduction in biodiversity through the replacement of native vegetation and the displacement of certain native animals (Whistler, 1988)..

It also commonly invades rainforest margins, where it grows over larger trees and smothers tree saplings and under-storey shrubs, and is a major problem in littoral rainforest remnants. However, it is also a weed of sandy beach fronts and other coastal environments, drier forests, wetlands, and limestone cliffs.

### **Phytochemical analysis**

Phytochemical screening is of paramount importance in identifying new source of therapeutically and industrially valuable compound having medicinal significance, to make the best and judicious use of available natural wealth (Ambasta *et al.*, 1986; Kokate *et al.*, 1998). Phytochemical investigation of the ethanolic extracts of *Ipomoea cairica* (L.) revealed the presence of various phytochemicals such as alkaloids, phenols, tannins, flavonoids, anthraquinones, saponins, steroids & terpenoids. The presence of alkaloids in *Ipomoea cairica*(L.) may be associated with their use by traditional medicines for the treatment of different diseases. Phenolic compounds present in it have attracted a great attention in relation to their potential for beneficial effects on health (Narayana *et al.*, 2001).

phytochemical analysis of plants for the presence of saponins are widely well known to have expectorant and anti tissue activity (Rao et. al., 1984; Sharama et al., 1984 ). FTIR spectroscopy data analysis helps in understanding the chemical functionality of the compound in the plant sample and when run under IR region in the range of 400-4000  $\text{cm}^{-1}$  there was a variation in the peak in both the plant samples (Thenmozhi et al., 2011; Kalaiselvi et al., 2012).

The bands between 3000 and 2800  $\text{cm}^{-1}$  represent C-H stretching vibrations that are mainly generated by lipids (Wolkers & Hoekstra, 1995; Wei et al., 2009). The weak absorption band of 796.60, 700.16, 617.22  $\text{cm}^{-1}$  indicate the presence of chloride in the plant extract (Muruganatham, 2009). The O-H stretching indicate the phenolic compound, that have excellent antioxidant properties (Shirwaikar, 2003).

FTIR allows detecting whole range of infrared spectrum simultaneously providing speed and accuracy in measurements of biological specimens (Griffiths & Haseth, 1986). The presence of characteristic functional groups phenols, carboxylic acid, alkanes, alkenes, esters, alkynes, fatty acid and lipids are responsible for various medicinal properties of *Ipomoea cairica* (L.) Based on the functional group analysis, *Ipomoea cairica* (L.) doesn't contain any toxic compounds.

In the present study, results of EDX showed the presence of trace elements such as Si, Cl, S, K and Ca in *Ipomoea cairica* (L.). Deficiency of these trace elements in human subjects can occur under the most practical dietary conditions and in much diseased status (Udayakumar & Begum, 2004). Trace elements play both curative and preventive role in combating diseases. Calcium is needed in the development of bone and teeth and it regulate heart rhythm, help in normal blood clotting, maintain proper nerve and muscle functions and lower blood pressure (Bibi et al., 2006). Potassium is essential for the transport of nutrients inside the cell. Without potassium, nutrients could not able enter into the cell that lead cell death. Silicon is also another important element to prevent the hardening of veins and arteries. Chloride works with sodium and potassium carry an electrical charge when dissolved body fluids and to regulate the pH in the body. Chloride is also important for digest the food properly and absorb many elements. The presence of these trace elements in *Ipomoea cairica* (L.) marks its use in therapeutic purpose.

## SUMMARY & CONCLUSION

Plant based drug has been used worldwide in traditional medicines for treatment of various disease. Phytochemical studies have attracted the attention of plant scientists due to the development of new and sophisticated techniques. These techniques played a significant role in the search for additional resources of raw material for pharmaceutical industry [2]. *Ipomoea cairica* (L.) Commonly known as 'Railway creeper' or Morning glory' belongs to the family Convolvulaceae. It has also ornamental value as climber with attractive flowers. This plant also included as plants affecting central nervous system [5], and also actively used as an antioxidant [6]. The main objective of the study was to screen the ethanol leaf extract of *Ipomoea cairica*(L.) for its phytochemical constituents, and to evaluate the functional groups using FITR and elemental analysis through EDX analysis growing in an industrial area.

The plant showed the presence of many phytochemicals which are responsible for the various pharmacological medicinal properties. The results obtained from the present study revealed that, the ethanol extracts of *Ipomoea cairica* (L.) showed the presence of most of the secondary metabolites like alkaloids, phenols, tannins, flavonoids, anthraquinones, saponins, steroids & terpenoids in the leaves. The FTIR study revealed the presence of functional groups of trace elements. The SEM-EDX study of the ethanol extract showed the presence of trace elements. However, the species was devoid of any heavy metals even though they were collected from an industrial area. The intensive study on the out coming active constituents of *Ipomoea cairica* (L.) can lead to the discovery of new botanical - drug.

The present study revealed that locally available non-economical weed plant *Ipomea cairica* commonly found in waste lands has a great pharmaceutical potential. Based on the result of this study it can be said that *Ipomoea cairica* has a leading capacity for the development of new good efficacy drugs in future and can be effective source to treat and control many diseases due to the presence of phytochemicals. Thus, it may conclude that the present study will add some specific criteria for the authentication of *Ipomoea cairica* and will be helpful to increase the economic potentiality of this plant.

## REFERENCE

1. Ambasta SP, Ramachandran K, Kashyapa K, Chand R. Useful plants of India. Publication and information directorate. Council of Scientific and Industrial Research: New Delhi;1986. p. 443-7.
2. Bibi S, Dastagir G, Hussain F, Sanaullah P. Elemental composition of *Viola odorata* Linn. Pak J Pl Sci 2006;12:141-3.
3. Flavio Amaral et al. Antinociceptive effect from *Ipomoea cairica* extract. Journal of Ethnopharmacology, 05/2006; 105(1-2):148-53. / DOI: 10.1016/j.jep.2005.10.012
4. Griffiths PR, De Haseth JA. Fourier transform infrared spectroscopy. John Wiley and Sons: New York;1986. p. 656.
5. Kalaiselvi M, Gomathi D, Vidya B, Uma C. Evaluation of Antioxidant potential and fourier transform infrared spectroscopy analysis of *Ananas comosus* (L.) Merr peel. Int Res J Pharm 2012;3:237-42.
6. Kokate CK, Purohit AP, Gokhale SB. Practical pharmacognosy. 1<sup>st</sup>Ed. Vallabh prakashan: Delhi;1998.
7. Lin HX, Sun SQ, Lv GH, Chan KK. Study on Angelica and its different extracts by Fourier transform infrared spectroscopy and two-dimensional correlation IR spectroscopy. Mol Biomol Spec 2006;64:321-6.
8. Muruganatham S, Anbalagan G, Ramamurthy N. FT-IR and sem-eds comparative analysis of medicinal plants, *Eclipta alba* HASSK and *Eclipta prostrata* linn. Rom J Biophys 2009;19;285-94.
9. Narayana KR, Reddy MS, Chaluvadi MR, Krishna DR. Bioflavonoids classification, pharmacology, biochemical effects and therapeutic potential. Ind J Pharmacol 2001;33:2-16.
10. Rao UP, Brahman M, Saxena HO. Phytochemical survey of Marurbhanj, Ganjam and Puri Dist. (Orissa) for tannins, Saponins, Flavonoids. Ind Drug 1984;22:503-7.
11. Sharma SD, Chishti AM, Koul MK. Phytochemical survey of plants from Kashmir-II. Indian Drugs 1984;22:187-95.
12. Shefali Arora, Deepak Kumar And Shiba. Phytochemical, Antimicrobial and Antioxidant Activities of Methanol Extract of Leaves and Flowers of *Ipomoea cairica*. International Journal of Pharmacy and Pharmaceutical Sciences, 2013 Vol 5, Issue 1(8)

13. Shirwaikar A, Malini S, Kumari SC. Protective effect of *Pongamia pinnata* flowers against cisplatin and gentamicin induced nephrotoxicity in rats. *Indian J Exp Biol* 2003;1:58-62.
14. Sutapa Choudhury, Chowdhury Habibur Rahaman, Sudhendu Mandal. Studies on *Ipomoea cairica* (L.) Sweet - A Promising Ethnomedicinally Important Plant . *Journal in Innovations in Pharmaceutical and Biological Sciences*, 2015; 2(4): pp 378-395
15. Thenmozhi M, Bhavya PK, Rajeshwari S. Compound Identification Using HPLC and FTIR In *Eclipta alba* and *Emilia sonchifolia*. *Int J Eng Sci Tech* 2011;3:292-8.
16. Udayakumar R, Begum VH. Elemental analysis of medicinal plants used in controlling infectious diseases. *Hamdard Med* 2004;67:35-6.
17. Vanlalhruii Ralte. Evaluation of phytochemical contents of *Ipomoea cairica* (L) Sweet - a Qualitative approach. 2014; 14(3):146-149
18. Wei ZL, Dong L, Tian ZH. Fourier transform infrared spectrometry study on early stage of cadmium stress in clover leaves. *Pak J Bot* 2009;41:1743-50.
19. Wolkers WF, Hoekstra AF. Ageing of dry desiccation-tolerant pollen does not affect protein secondary structure. *Plant Physiol* 1995;109:907-15.

Project Report

On

# COVID-19 COMORBIDITIES AND RISK FACTOR ANALYSIS

*Submitted*

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

MASTER OF SCIENCE

*in*

APPLIED STATISTICS AND DATA ANALYTICS

*by*

ELSA PHILOMON

(Register No. SM20AS012)

(2020-2022)

*Under the Supervision of*

MS.SANGEETHA CHANDRAN



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, **COVID-19 COMORBIDITIES AND RISK FACTOR ANALYSIS** is a bonafide record of the work done by Ms. **ELSA PHILOMON** 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

For *Amal*

**SANGEETHA CHANDRAN**

Assistant Professor,  
Department of Computer Applications,  
St. Teresa's College (Autonomous),  
Ernakulam.



*Shanty B P*

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

External Examiners

1: *James K S D*

*26/05/2022*

2: *Shanty 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 MS.SANGEETHA CHANDRAN, Assistant Professor, Department of Computer Applications, 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

  
ELSA PHILOMON

SM20AS009

## 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.Sangeetha Chandran, Assistant Professor, Department of Computer Applications, 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*

  
ELSA PHILOMON  
SM20AS012

# ABSTRACT

In this study the comorbidities and risk factors that leads to severe outcomes in covid positive patients are being analysed. The COVID-19 outbreak has severely affected human health and economic development worldwide and is still causing threat to the people's lives. To better protect and treat higher risk patients and to lessen the burden on healthcare systems, it is important to identify the risk factors of COVID-19.

The COVID-19 pandemic has resulted in lots of deaths within the United Kingdom. The study used detailed information on 582 COVID-19-positive inpatients in Bradford and Calderdale between February-August 2020 that were extracted from Electronic Health Records.

It was studied if the presence of comorbidities such as diabetes, hypertension, cardiovascular disease, asthma, chronic obstructive pulmonary disease, cancer, renal disease and obesity were associated with serious health consequences during COVID-19 illness, including death. The study also aims to find the association of factors like age and gender with the death of a covid positive patient.

Machine learning techniques such as logistic regression and correlation methods were used to study the relationship between pre-existing comorbidities with mortality. The important features were extracted using ExtraTreesClassifier. Those patients with pre-existing comorbidities have been found to be at increased risk of mortality. Patients with diabetes, hypertension, cardiovascular disease or renal disease or a combination of any of these were most likely to experience severe outcomes of the infection. Higher rate of mortality was seen in people above the age of 50 years. The death rate was higher among males as compared to females.

Thus, identifying the risk factors helps in providing appropriate medical support to the patients and hence to reduce severe mortality.

# 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 Objectives</b>	<b>3</b>
<b>3 Literature Review</b>	<b>4</b>
<b>4 Materials and Methods</b>	<b>9</b>
4.1 Sample . . . . .	9
4.2 Data availability . . . . .	9
4.3 Introduction to Supervised Learning . . . . .	10
4.4 Logistic Regression . . . . .	11
4.4.1 Assumptions of Logistic Regression . . . . .	12
4.4.2 Implementing Logistic Regression in Python . . . . .	12
4.5 Analysis of Data . . . . .	13
4.5.1 Outcome Distribution . . . . .	13
4.5.2 Gender Distribution . . . . .	13
4.5.3 Age Category Distribution . . . . .	14
4.5.4 Different Characteristics, Comorbidities and Outcomes . . . . .	15
<b>5 Results</b>	<b>17</b>
5.1 Age Category and Outcome Distribution . . . . .	17
5.2 Gender and Outcome Distribution . . . . .	18

5.3	BMI Category and Outcome Distribution . . . . .	18
5.4	Ethnicity and Outcome Distribution . . . . .	19
5.5	Socioeconomic status and Outcome Distribution . . . . .	19
5.6	Correlation . . . . .	20
5.7	Important Features . . . . .	20
5.8	Confusion Matrix . . . . .	21
5.9	Model Accuracy . . . . .	21
<b>6</b>	<b>Python Codes</b>	<b>22</b>
<b>7</b>	<b>Discussion and Conclusion</b>	<b>26</b>
	<i>REFERENCES</i> . . . . .	27

# Chapter 1

## Introduction

---

The coronavirus disease 2019 (COVID-19) is an acute infectious pneumonia caused by a severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) infection previously unknown to humans. Spreading of COVID-19 mainly occurs when people are exposed to virus containing respiratory droplets and airborne particles exhaled by an infected person. The virus causes mild symptoms in the majority of cases, the most common being: fever, dry cough, and fatigue.

The disease has the characteristics of fast transmission and strong infectivity. Since the outbreak in early December 2019 in Wuhan, China, it has rapidly developed into a worldwide pandemic, with more than 3 million patients confirmed to have been diagnosed with the disease in more than 200 countries, and the number of infected people is probably much higher. As of April 30, 2020, 217,769 people died of COVID-19 infection. Despite the public health responses aimed at containing the disease and delaying its spread; during the courses of treatment, due to the large increase in the demand for hospital beds and the shortage of medical equipment, coupled with the lack of specific medicine, patients with basic diseases or old age are more likely to progress to severe disease, leading to death. Reports show that 14.1–33.0 percent of COVID-19 infected patients are prone to develop into severe cases, and the mortality rate of critical cases is 61.5 percent, increasing sharply with age and underlying comorbidities. Furthermore, medical staff may also be infected, which makes many countries face critical care crisis. COVID-19 poses an important and urgent threat to global health.

This pandemic continues to challenge medical systems worldwide in many aspects, including sharp increases in demands for hospital beds and critical shortages in medical equipment, while many healthcare workers have themselves been infected. Thus, the capacity for immediate clinical decisions and effective usage of healthcare resources is crucial.

Effective screening enables quick and efficient diagnosis of COVID-19 and can mitigate the burden on healthcare systems. Prediction models that combine several features to estimate the risk of infection have been developed in hopes of assisting medical staff worldwide in triaging patients when allocating limited healthcare resources

# Chapter 2

## Objectives

---

1. To study the association of pre-existing comorbidities with the death of a covid patient.

That is, to find if pre-existing comorbidities like diabetes, hypertension, cardiovascular disease, asthma, chronic obstructive pulmonary disease, cancer and renal disease causes severe outcomes in a covid positive patient.

2. To find the association of factors like:

Age,

Gender,

Socioeconomic status and

BMI (body mass index), with the outcome of the hospitalised covid-19 patients.

## Chapter 3

# Literature Review

---

In the article, Ethnicity, pre-existing comorbidities, and outcomes of hospitalised patients with COVID-19 by Gillian Santorelli, Michael McCooe, Trevor A. Sheldon, John Wright and Tom Lawton, it was studied if patients with existing comorbidities and minority ethnic groups are at increased risk of mortality. They wished to determine if there were any differences in intensive care unit (ICU) admission and 30-day hospital mortality in a city with high levels of deprivation and a large community of people of South Asian heritage. Detailed information on 582 COVID-19-positive inpatients in Bradford and Calderdale between February-August 2020 were extracted from Electronic Health Records. Logistic regression and Cox proportional hazards models were used to explore the relationship between ethnicity with admission to ICU and 30-day mortality, respectively accounting for the effect of demographic and clinical confounders. The sample consisted of 408 (70 percent) White, 142 (24 percent) South Asian and 32 (6 percent) other minority ethnic patients. Ethnic minority patients were younger, more likely to live in deprived areas, and be overweight/obese, have type 2 diabetes, hypertension and asthma compared to white patients, but were less likely to have cancer (South Asian patients only) and COPD. Male and obese patients were more likely to be admitted to ICU, and patients of South Asian ethnicity, older age, and those with cancer were less likely. Being male, older age, deprivation, obesity, and cancer were associated with 30-day mortality. The risk of death in South Asian patients was the same as in white patients HR 1.03 (0.58, 1.82). It was found that, despite South Asian patients being less likely to be admitted to ICU and having a higher prevalence of diabetes and obesity, there was no difference in the

risk of death compared to white patients. This contrasts with other findings and highlights the value of studies of communities which may have different ethnic, deprivation and clinical risk profiles

Bradford Teaching Hospitals Foundation Trust (BTHFT) and Calderdale and Huddersfield Foundation Trust (CHFT) are large hospital trusts in West Yorkshire in the UK which consist of three separate acute hospitals and together serve a population of approximately one million people. Bradford is a deprived and ethnically diverse city, with high rates of health problems such as obesity and cardiovascular disease as well as the highest prevalence of diabetes in the UK. Calderdale and Huddersfield, by comparison have a smaller proportion of residents from black, south Asian and minority ethnic backgrounds yet have many areas that suffer poor health outcomes and significant deprivation. The current pandemic has exposed and resulted in inequalities between UK ethnic groups .

On the basis of early evidence from China and Italy, Bradford teaching hospitals, chose to adopt the wide- spread early use of enhanced respiratory support in the form of CPAP (Continuous Positive Airway Pressure) and self-proning in the management of more severe COVID-19, with the aims of improving patient outcomes and controlling critical care demand. Preliminary analysis of patients receiving CPAP and mechanical ventilation at BTHFT did not demonstrate an increased mortality rate amongst those of South Asian ethnicity and this observation provided impetus for this multi-centre review of inpatients with a diagnosis of COVID 19. The overall aim of this study was to examine the ethnic, demographic, socio-economic and clinical risk factors associated with outcomes of hospital inpatients who tested positive for COVID-19. The endpoints were ICU admission and 30-day in-hospital mortality.

In the article, *The Impact of Pre-existing Comorbidities and Therapeutic Interventions on COVID19* by Lauren A. Callender , Michelle Curran , Stephanie M. Bates, Maelle Mairesse , Julia Weigandt and Catherine J. Betts, it was said that, evidence from the global outbreak of SARS-CoV-2 has clearly demonstrated that individuals with pre-existing comorbidities are at a much greater risk of dy-

ing from COVID-19. This is of great concern for individuals living with these conditions, and a major challenge for global healthcare systems and biomedical research. Not all comorbidities confer the same risk, however, many affect the function of the immune system, which in turn directly impacts the response to COVID-19. Furthermore, the myriad of drugs prescribed for these comorbidities can also influence the progression of COVID-19 and limit additional treatment options available for COVID-19. Here, they reviewed immune dysfunction in response to SARS-CoV-2 infection and the impact of pre-existing comorbidities on the development of COVID-19. They explored how underlying disease etiologies and common therapies used to treat these conditions exacerbate COVID-19 progression. Moreover, they discussed the long-term challenges associated with the use of both novel and repurposed therapies for the treatment of COVID-19 in patients with pre-existing comorbidities.

From this study, it was found that patients with pre-existing comorbidities are at a greater risk of dying from COVID-19. However, not all comorbidities confer the same risk. By exploring the underlying disease etiologies and common therapies used to treat these conditions, they have discussed their impact on COVID-19. Comorbidities closely associated with age, chronic inflammation and dysregulated metabolism such as hypertension, cardiovascular disease, and diabetes are the most prevalent comorbidities. However, many of these comorbidities are strongly associated with each other. Consequently, many patients will have multiple comorbidities and therefore while they have discussed these individually, the reality is that a combination of factors will be at play. Furthermore, as multiple drug use is inevitable for patients with pre-existing comorbidities, the impact of overlaying drugs on an already compromised state and the possibility of DDI leading to adverse events needs to be carefully considered. As the scale of this pandemic continues to accelerate globally, they hope this review provides healthcare professionals and biomedical researchers with a more comprehensive understanding of the impact of pre-existing comorbidities on COVID-19 development and treatment.

The article, Presenting Characteristics, Comorbidities, and Outcomes Among

5700 Patients Hospitalized With COVID-19 in the New York City Area by Safiya Richardson, Jamie S. Hirsch, Mangala Narasimhan, James M. Crawford, Thomas McGinn, Karina W. Davidson, and the Northwell COVID-19 Research Consortium, aimed to describe the clinical characteristics and outcomes of patients with COVID-19 hospitalized in a US health care system. The study included all sequentially hospitalized patients with COVID-19 admitted to 12 hospitals in New York City, Long Island, and Westchester County, New York, within the Northwell Health system between March 1, 2020, and April 4, 2020, inclusive of these dates. Confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection by positive result on polymerase chain reaction testing of a nasopharyngeal sample among patients requiring admission.

Clinical outcomes during hospitalization, such as invasive mechanical ventilation, kidney replacement therapy, and death. Demographics, baseline comorbidities, presenting vital signs, and test results were also collected. A total of 5700 patients were included (median age, 63 years [interquartile range IQR, 52-75; range, 0-107 years]; 39.7 percent female). The most common comorbidities were hypertension (3026; 56.6 percent), obesity (1737; 41.7 percent), and diabetes (1808; 33.8 percent). At triage, 30.7 percent of patients were febrile, 17.3 percent had a respiratory rate greater than 24 breaths/minute, and 27.8 percent received supplemental oxygen. The rate of respiratory virus co-infection was 2.1 percent. Outcomes were assessed for 2634 patients who were discharged or had died at the study end point. During hospitalization, 373 patients (14.2 percent) (median age, 68 years [IQR, 56-78]; 33.5 percent female) were treated in the intensive care unit care, 320 (12.2 percent) received invasive mechanical ventilation, 81 (3.2 percent) were treated with kidney replacement therapy, and 553 (21 percent) died. As of April 4, 2020, for patients requiring mechanical ventilation ( $n = 1151$ , 20.2 percent), 38 (3.3 percent) were discharged alive, 282 (24.5 percent) died, and 831 (72.2 percent) remained in hospital. The median post discharge follow-up time was 4.4 days (IQR, 2.2-9.3). A total of 45 patients (2.2 percent) were readmitted during the study period. The median time to readmission was 3 days (IQR, 1.0-4.5) for readmitted patients. Among the 3066 patients who remained hospitalized at the final study follow-up date (median age, 65 years [IQR, 54-75]), the median follow-up at time

of censoring was 4.5 days (IQR, 2.4-8.1). This case series provides characteristics and early outcomes of sequentially hospitalized patients with confirmed COVID-19 in the New York City area.

In the article, Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning by Wan Xu, Nan-Nan Sun, Hai-Nv Gao, Zhi-Yuan Chen, Ya Yang, Bin Ju, and Ling-Ling Tang, it is said that, acute respiratory distress syndrome (ARDS) is one of the common clinical manifestations of severe COVID-19 and it is also responsible for the current shortage of ventilators worldwide. This study aims to analyze the clinical characteristics of COVID-19 ARDS patients and establish a diagnostic system based on artificial intelligence (AI) method to predict the probability of ARDS in COVID-19 patients. They collected clinical data of 659 COVID-19 patients from 11 regions in China. The clinical characteristics of the ARDS group and no-ARDS group of COVID-19 patients were elaborately compared and both traditional machine learning algorithms and deep learning-based method were used to build the prediction models. Results indicated that the median age of ARDS patients was 56.5 years old, which was significantly older than those with non-ARDS by 7.5 years. Male and patients with BMI greater than 25 were more likely to develop ARDS. The clinical features of ARDS patients included cough (80.3 percent), polypnea (59.2 percent), lung consolidation (53.9 percent), secondary bacterial infection (30.3 percent), and comorbidities such as hypertension (48.7 percent). Abnormal biochemical indicators were strongly related to the aggravation of ARDS. Furthermore, through various AI methods for modeling and prediction effect evaluation based on the above risk factors, decision tree achieved the best AUC, accuracy, sensitivity and specificity in identifying the mild patients who were easy to develop ARDS, which undoubtedly helped to deliver proper care and optimize use of limited resources.

These literatures gave an idea about the risk factors of covid-19 and thus helped to understand more about the topic.

# Chapter 4

## Materials and Methods

---

### 4.1 Sample

This prospective cohort study includes all patients admitted to Bradford Teaching Hospitals Foundation Trust (BTHFT) between 17/02/2020 – 20/06/2020 and Calderdale and Huddersfield Foundation Trust (CHFT) between 01/03/2020 – 08/08/2020, and who tested positive for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) using RT-PCR on admission or during their stay. Bradford has a large South Asian population (24.9 percent, 20.4 percent of who are of Pakistani origin; 67.4 percent white), whereas Calderdale and Kirklees (the Metropolitan Borough in which Huddersfield resides) have a larger white population (89.6 percent and 79.1 percent respectively).

Ethics approval: The recent Health Service (Control of Patient Information) Regulations 2002 notices requires NHS Trusts and others to process confidential patient information without consent for COVID-19 public health, surveillance and research purposes, thus no ethical approval or consent for publication was required.

### 4.2 Data availability

The dataset was taken from the site [dataverse.harvard.edu](https://dataverse.harvard.edu).

Harvard Dataverse: Replication Data for: Ethnicity, pre-existing comorbidities, and outcomes of hospitalised patients with COVID-19.

This project contains the following underlying data:

- DatasetForPublication.xls

Detailed information on ethnicity, demographic characteristics, hospital stay, and pre-existing comorbidities was manually extracted from hospital electronic patient records (EPR) for BTHFT patients; the data for those admitted to CHFT was extracted directly from the data warehouse. Self-defined ethnicity was categorised to reflect the majority ethnic groups in the region: White (British, Irish, any other White), South Asian (Indian, Pakistani, Bangladeshi), and other ethnic minority; patients with missing ethnicity were excluded from the analysis. Potential risk factors for severe COVID-19 were identified a priori: age, sex, obesity, and pre-existing comorbidities (diabetes, hypertension, coronary heart disease, asthma, chronic obstructive pulmonary disease (COPD), cancer, chronic renal disease). BMI categories were defined as healthy ( $<25$  kg/m<sup>2</sup>) overweight (25–29.9 kg/m<sup>2</sup>) and obese ( $\geq 30$  kg/m<sup>2</sup>) based on BMI at hospital admission. Socioeconomic status was estimated using the English Indices of Multiple Deprivation (IMD) quintiles derived from residential postcode. Due to small numbers in the least deprived quintile, IMD 4 and 5 were combined.

### 4.3 Introduction to Supervised Learning

Supervised machine learning algorithms derive insights, patterns, and relationships from a labeled training dataset. It means the dataset already contains a known value for the target variable for each record. It is called supervised learning because the process of an algorithm learning from the training dataset is like an instructor supervising the learning process. You know the correct answers, the algorithm iteratively makes predictions on the training data and the instructor corrects it. Learning ends when the algorithm achieves the desired level of performance and accuracy.

Supervised learning problems can be further classified into regression and classification problems.

- **Classification:** In a classification problem, the output variable is a category, such as “red” or “blue,” “disease” or “no disease,” “true” or “false,” etc.
- **Regression:** In a regression problem, the output variable is a real continuous

value, such as “dollars” or “weight.”

## 4.4 Logistic Regression

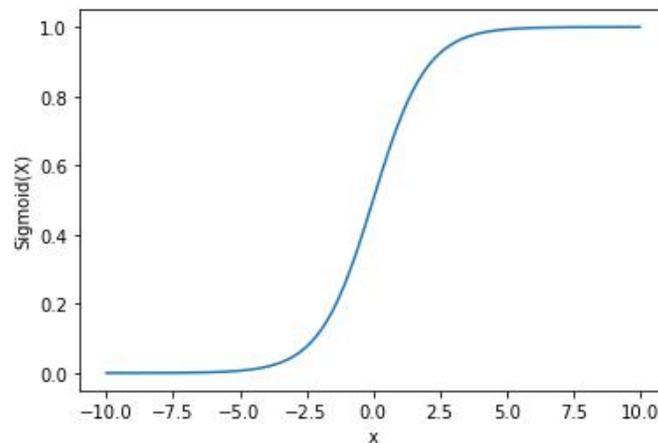
Logistic regression is a statistical method that is used for building machine learning models where the dependent variable is dichotomous: i.e, binary. Logistic regression is used to describe data and the relationship between one dependent variable and one or more independent variables. The independent variables can be nominal, ordinal, or of interval type.

The name “logistic regression” is derived from the concept of the logistic function that it uses. The logistic function is also known as the sigmoid function. The value of this logistic function lies between zero and one.

Below is the function for logistic regression:

$$\text{Sig}(x) = \frac{1}{1 + e^{-x}}$$

X is the numerical value that needs to be transformed.



If we feed an output value to the sigmoid function, it will return the probability of the outcome between 0 and 1. If the value is below 0.5, then the output is returned as No. If the value is above 0.5, then the output is returned as Yes.

### 4.4.1 Assumptions of Logistic Regression

- Independent variables show a linear relationship with the log of output variables.
- Non-Collinearity between independent variables. That is, independent variables are independent of each other.
- Output variable is binary

### 4.4.2 Implementing Logistic Regression in Python

First, the Libraries needed for creating the model are imported.  
Then the dataset is imported to the notebook.

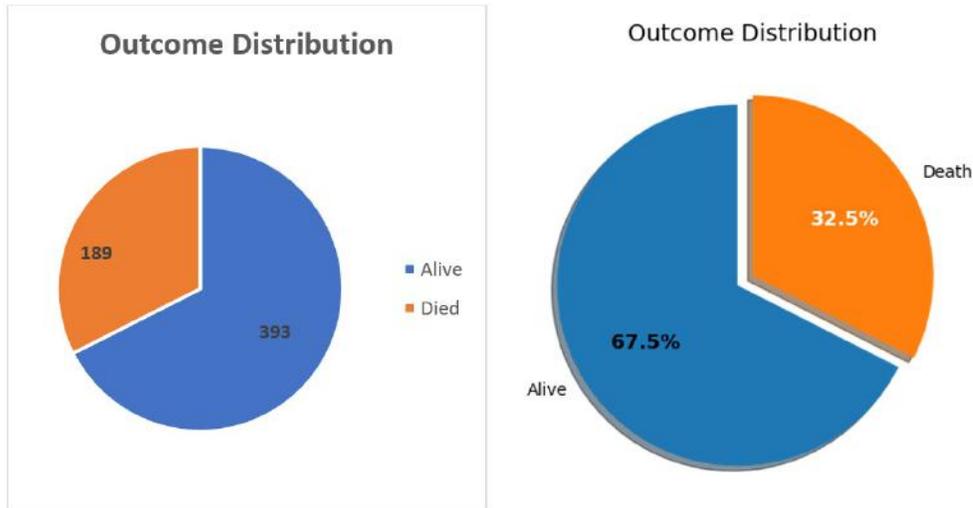
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	id	sex	agecat	ethnicity3c	imd	bmicat	diabetes1	diabetes2	hypertensi	cvd	asthma	copd	cancer	renaldis	icu	died30day	timeat	risk
2	1	Female	70-79	White	IMD 1	Overweigh	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	6
3	2	Male	60-69	White	IMD 1	Overweigh	No	Yes	No	No	No	No	No	No	No	No	No	17
4	3	Male	70-79	White	IMD 1	Healthy wt	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	8
5	4	Male	80+	White	IMD 4/5	Healthy wt	No	No	No	Yes	No	No	No	No	No	Yes	No	13
6	5	Male	60-69	White	IMD 3	Healthy wt	No	No	No	No	No	No	No	No	No	No	No	7
7	6	Female	80+	White	IMD 1	Healthy wt	No	No	No	Yes	No	No	No	No	No	No	No	29
8	7	Female	80+	White	IMD 3	Overweigh	No	No	No	Yes	No	Yes	No	No	No	No	No	8
9	8	Male	70-79	White	IMD 4/5	Healthy wt	No	No	No	No	No	No	No	No	No	No	No	8
10	9	Male	80+	White	IMD 3	Healthy wt	No	No	Yes	No	No	No	No	No	No	Yes	No	6
11	10	Male	50-59	White	IMD 2	Healthy wt	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	No	3
12	11	Male	50-59	White	IMD 4/5	Obese	No	No	No	No	Yes	No	No	No	Yes	No	No	8
13	12	Female	18-49	South Asit	IMD 1	Obese	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	24
14	13	Female	18-49	White	IMD 1	Obese	No	No	No	No	No	No	No	No	No	No	No	11
15	14	Female	18-49	White	IMD 3	Overweigh	No	Yes	No	No	No	No	No	No	Yes	No	No	30
16	15	Female	18-49	South Asit	IMD 1	Obese	No	No	No	No	No	No	No	Yes	No	No	No	5
17	16	Male	18-49	South Asit	IMD 2	Obese	No	No	No	No	No	No	No	No	Yes	No	No	7
18	17	Male	50-59	White	IMD 1	Overweigh	No	No	No	No	No	No	No	No	No	No	No	25
19	18	Male	18-49	Other	IMD 2	Overweigh	No	No	No	No	No	No	Yes	No	No	No	No	18
20	19	Male	18-49	South Asit	IMD 2	Healthy wt	No	Yes	Yes	No	No	No	No	No	No	No	No	6
21	20	Female	18-49	South Asit	IMD 2	Obese	No	Yes	No	No	No	No	No	No	No	No	No	5
22	21	Male	18-49	Other	IMD 1	Obese	No	No	Yes	No	Yes	No	No	No	No	No	No	17
23	22	Female	18-49	South Asit	IMD 2	Obese	No	No	No	No	No	No	No	No	No	No	No	2
24	23	Female	18-49	South Asit	IMD 1	Obese	No	No	No	No	Yes	No	No	No	No	No	No	14
25	24	Male	50-59	White	IMD 1	Overweigh	No	No	No	No	No	No	No	No	No	No	No	4
26	25	Female	50-59	White	IMD 2	Obese	No	No	No	No	No	No	No	No	No	No	No	11
27	26	Female	50-59	White	IMD 2	Healthy wt	No	No	No	Yes	No	No	No	No	No	No	No	29

Independent variables are not influenced by any factors and impact the outcome variable. Here,sex,agecat,ethnicity,imd,bmicat,diabetes1, diabetes2,hypertension,cvd,asthma,copd,cancer,r and timeatrisk are the independent variables.

Dependent variables are influenced or dependent on the independent variables.Here,‘died30days’ is the dependent variable.

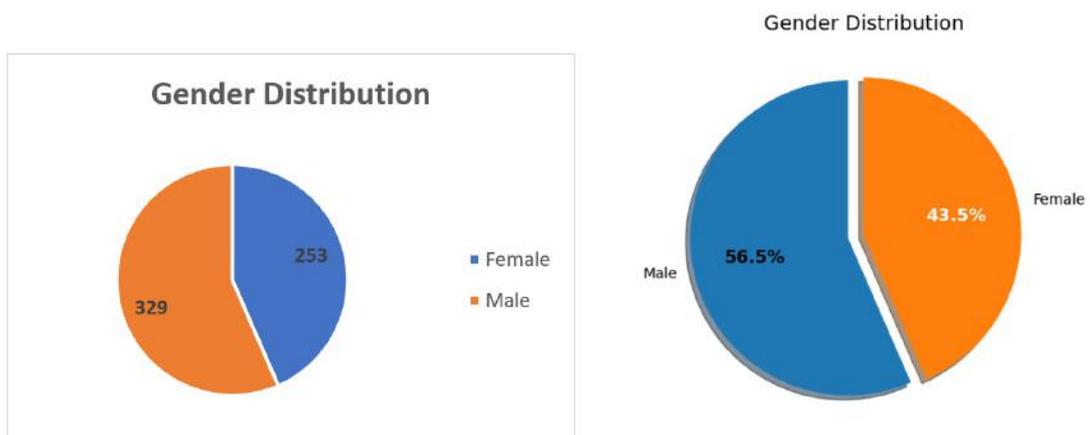
## 4.5 Analysis of Data

### 4.5.1 Outcome Distribution



The above figure shows the Outcome Distribution. Out of the 582 patients, 189 patients died and 393 were alive, i.e, 32.5 percent patients died and 67.5 percent were alive.

### 4.5.2 Gender Distribution

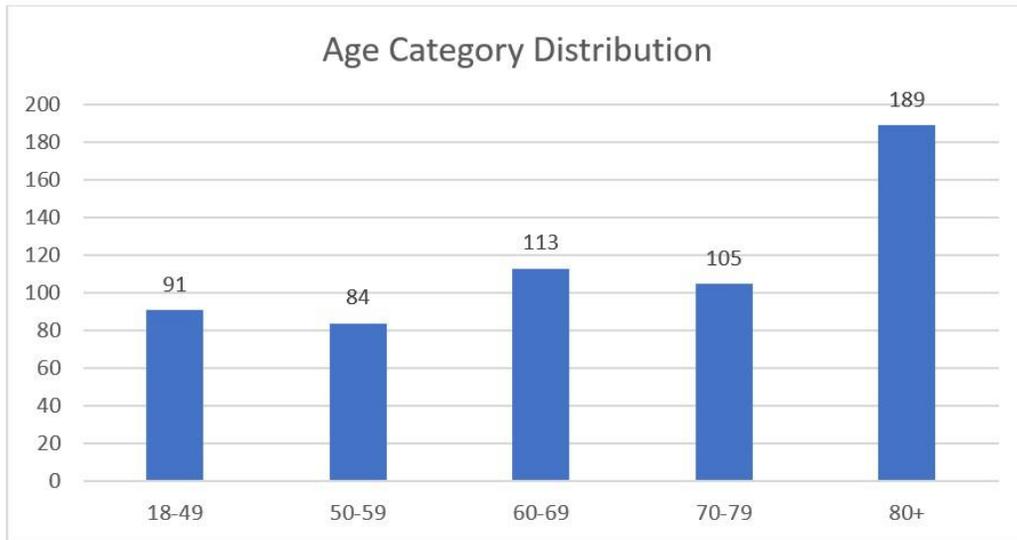


The above figure shows the Gender Distribution. Out of the 582 patients, 329 patients were males and 253 females, i.e, 56.5 percent patients were males and 43.5

percent females.

### 4.5.3 Age Category Distribution

The figure below shows the Age Category Distribution. The patients were divided into 5 age categories:



4.5.4 Different Characteristics, Comorbidities and Outcomes

Table showing different characteristics and outcomes

Characteristic		Died				Total n=582	
		Yes (n=189)		No (n=393)		n	
		n	%	n	%		
Gender	Male	126	66.67	203	51.65	329	56.53
	Female	63	33.33	190	48.35	253	43.47
Age Category	18-49	3	1.59	88	22.39	91	15.64
	50-59	13	6.88	71	18.07	84	14.43
	60-69	32	16.93	81	20.61	113	19.42
	70-79	42	22.22	63	16.03	105	18.04
	80+	99	52.38	90	22.90	189	32.47
Ethnicity	South Asian	28	14.82	114	29.0	142	24.4
	White	154	81.48	254	64.631	408	70.1
	Others	7	3.70	25	6.361	32	5.5
Body Mass Index(BMI)	Healthy weight	98	51.852	182	46.31	280	48.11
	Obese	40	21.16	114	29.01	154	26.46
	Overweight	51	26.98	97	24.682	148	25.43

Table showing different comorbidities and outcomes

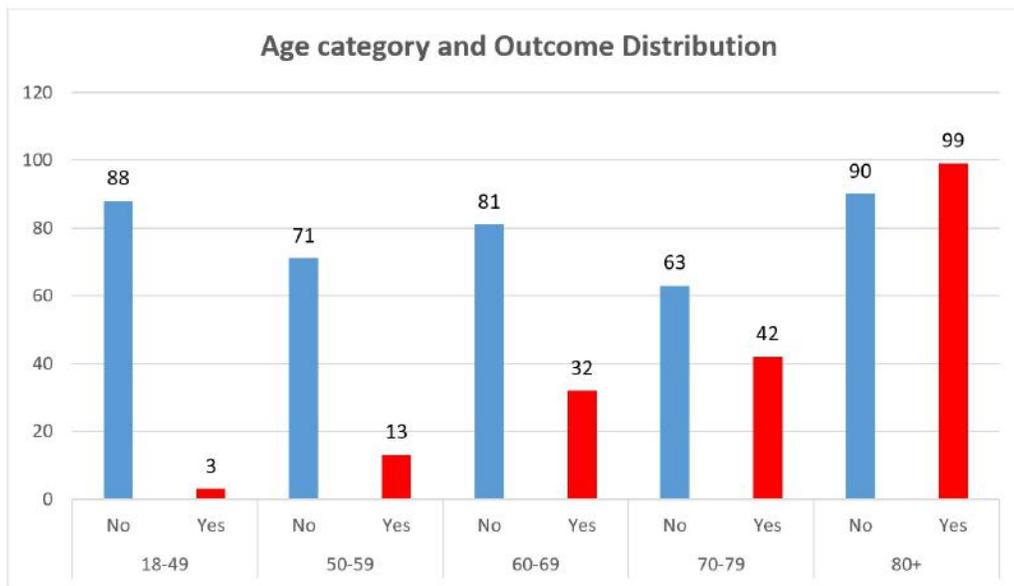
Comorbidities	Died				Total n=582	
	Yes (n=189)		No (n=393)		n	%
	n	%	n	%		
<b>Diabetes 1</b>						
Yes	3	1.59	7	1.78	10	1.72
No	186	98.41	386	98.22	572	98.28
<b>Diabetes 2</b>						
Yes	76	40.2	131	33.33	207	35.57
No	113	59.8	262	66.67	375	64.43
<b>Hypertension</b>						
Yes	94	49.74	176	44.78	270	46.39
No	95	50.26	217	55.22	312	53.61
<b>Cardiovascular disease</b>						
Yes	83	43.92	93	23.66	176	30.24
No	106	56.08	300	76.34	406	69.76
<b>Asthma</b>						
Yes	23	12.17	68	17.3	91	15.64
No	166	87.83	325	82.7	491	84.36
<b><u>copd</u></b>						
Yes	34	17.99	44	11.2	78	13.4
No	155	82.01	349	88.8	504	86.6
<b>Cancer</b>						
Yes	26	13.76	21	5.34	47	8.08
No	163	86.24	372	94.66	535	91.92
<b>Renal Disease</b>						
Yes	72	38.1	74	18.83	146	25.1
No	117	61.9	319	81.17	436	74.9

# Chapter 5

## Results

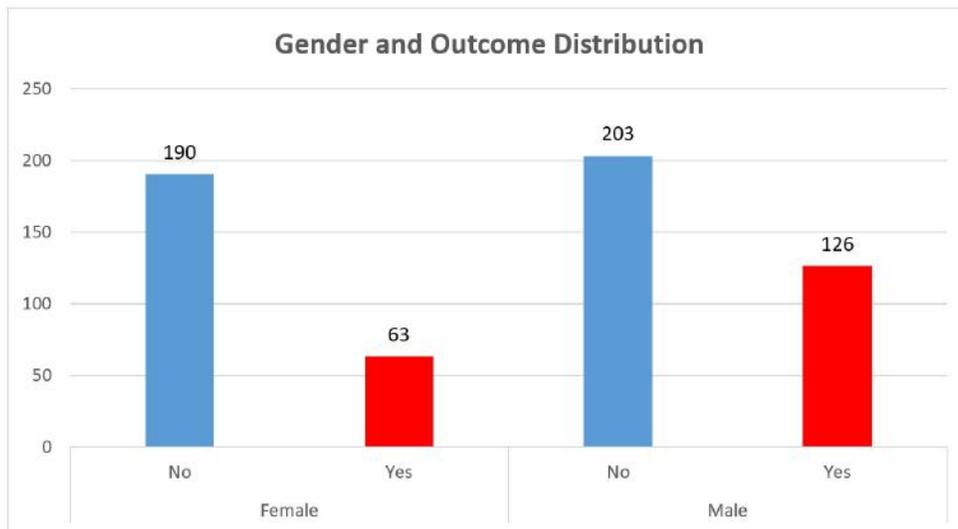
---

### 5.1 Age Category and Outcome Distribution



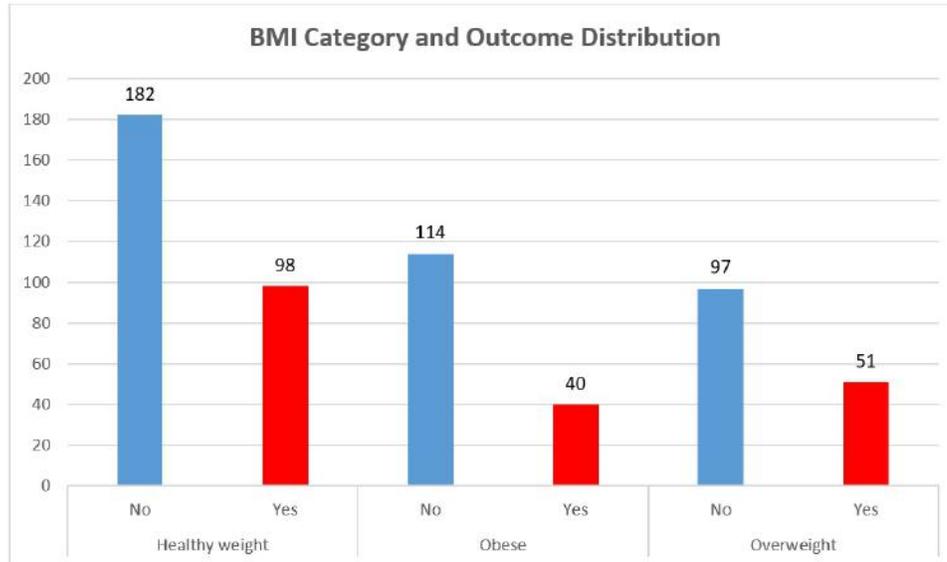
From the above chart we can say that, as the age is a major risk factor for covid positive patients. As the age increases, the number of deaths is also seen increasing.

## 5.2 Gender and Outcome Distribution



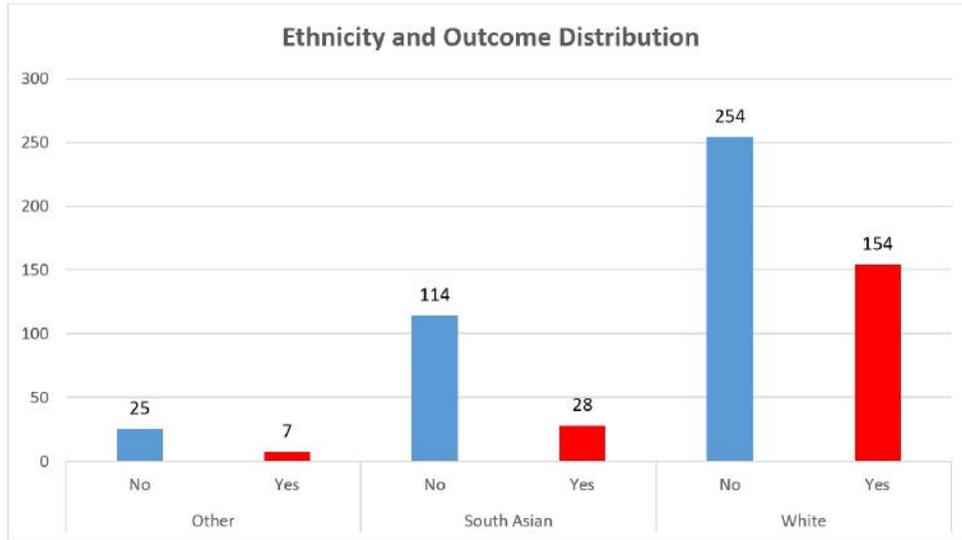
Higher number of deaths are seen in males as compared to females.

## 5.3 BMI Category and Outcome Distribution

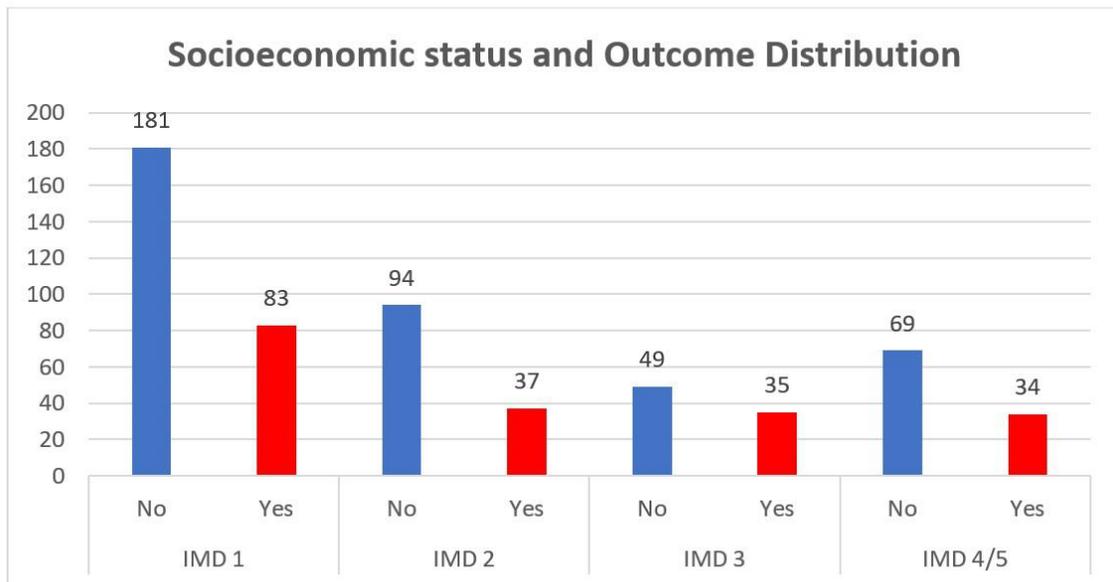


We can see a comparatively higher death count in patients having overweight.

### 5.4 Ethnicity and Outcome Distribution



### 5.5 Socioeconomic status and Outcome Distribution



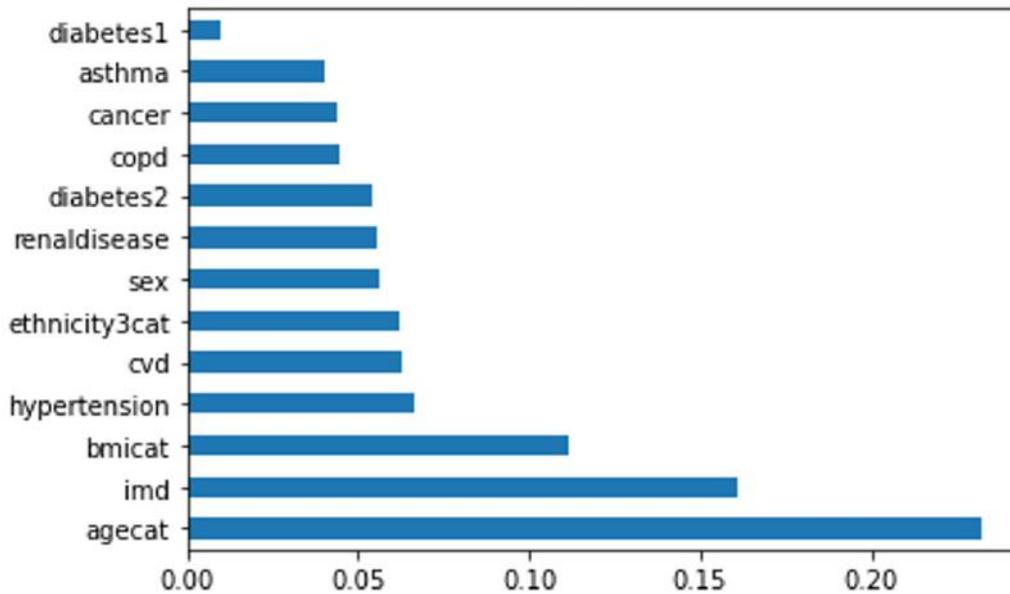
## 5.6 Correlation

Correlation matrix showing the relationship between different features

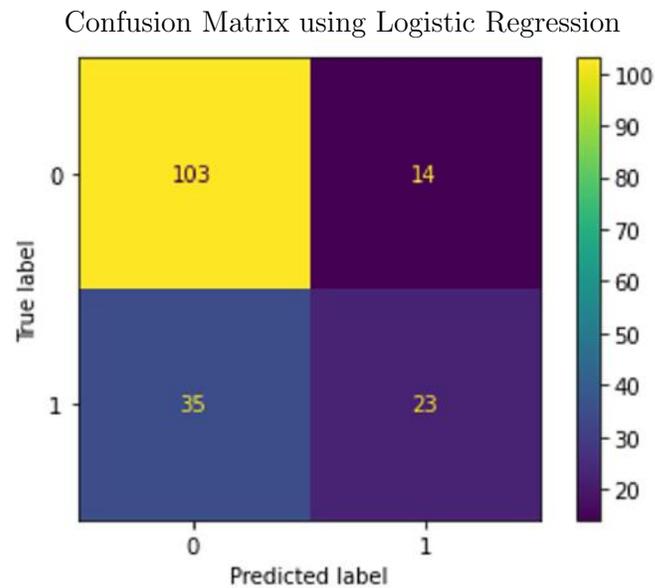


## 5.7 Important Features

Important Features identified using ExtraTreesClassifier



## 5.8 Confusion Matrix



## 5.9 Model Accuracy

The Logistic Regression Model Accuracy is 0.72

	<b>precision</b>	<b>recall</b>	<b>f1-score</b>	<b>support</b>
<b>0</b>	<b>0.75</b>	<b>0.88</b>	<b>0.81</b>	<b>117</b>
<b>1</b>	<b>0.62</b>	<b>0.40</b>	<b>0.48</b>	<b>58</b>
<b>accuracy</b>			<b>0.72</b>	<b>175</b>
<b>macro avg</b>	<b>0.68</b>	<b>0.64</b>	<b>0.65</b>	<b>175</b>
<b>weighted avg</b>	<b>0.71</b>	<b>0.72</b>	<b>0.70</b>	<b>175</b>

# Chapter 6

## Python Codes

---

In [1]:

```
#importing Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sn
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
from sklearn.metrics import mean_absolute_error
```

```
#to get information about the data
data.info()
```

```
0  id          582 non-null  int64
1  sex         582 non-null  object
2  agecat      582 non-null  object
3  ethnicity3cat 582 non-null  object
4  imd         582 non-null  object
5  bmicat      582 non-null  object
6  diabetes1   582 non-null  object
7  diabetes2   582 non-null  object
8  hypertension 582 non-null  object
9  cvd         582 non-null  object
10 asthma     582 non-null  object
11 copd       582 non-null  object
12 cancer     582 non-null  object
13 renaldisease 582 non-null  object
14 icu        582 non-null  object
15 died30days 582 non-null  object
16 timeatrisk 582 non-null  int64
dtypes: int64(2), object(15)
memory usage: 77.4+ KB
```

```
xfeatures=data[['sex', 'agecat', 'ethnicity3cat', 'imd', 'bmicat', 'diabetes1',
               'diabetes2', 'hypertension', 'cvd', 'asthma', 'copd', 'cancer',
               'renaldisease']]
ylabels=data['died30days']
```

```
from sklearn.ensemble import ExtraTreesClassifier
```

In [182]:

```
et_clf = ExtraTreesClassifier()  
et_clf.fit(xfeatures,ylabels)
```

Out[182]:

```
ExtraTreesClassifier()
```

In [183]:

```
print(et_clf.feature_importances_)
```

```
[0.05612079 0.23215654 0.06234147 0.16129717 0.11179223 0.00981666  
 0.05395737 0.06639243 0.06283838 0.03984124 0.04432223 0.04390055  
 0.05522294]
```

In [184]:

```
feature_importance_df=pd.Series(et_clf.feature_importances_,index=xfeatures.columns)
```

In [185]:

```
feature_importance_df
```

Out[185]:

```
sex          0.056121  
agecat      0.232157  
ethnicity3cat 0.062341  
imd         0.161297  
bmicat      0.111792  
diabetes1   0.009817  
diabetes2   0.053957  
hypertension 0.066392  
cvd         0.062838  
asthma      0.039841  
copd        0.044322  
cancer      0.043901  
renaldisese 0.055223  
dtype: float64
```

In [204]:

```
accuracy_score(y_test,logreg.predict(x_test))
```

Out[204]:

0.72

In [205]:

```
#building model  
model_logit=LogisticRegression()  
model_logit.fit(x_train_b,y_train_b)
```

Out[205]:

LogisticRegression()

In [206]:

```
#model accuracy  
# method 1  
model_logit.score(x_test_b,y_test_b)
```

Out[206]:

0.72

In [244]:

```
from sklearn.metrics import confusion_matrix
```

In [245]:

```
y_pred=model_logit.predict(x_test_b)
```

---

In [246]:

```
#accuracy score  
accuracy_score(y_test,y_pred)
```

Out[246]:

0.72

In [247]:

```
#using confusion matrix  
confusion_matrix(y_test,y_pred)
```

Out[247]:

```
array([[103, 14],  
       [ 35, 23]], dtype=int64)
```

In [248]:

```
#plot  
from sklearn.metrics import plot_confusion_matrix
```

In [249]:

```
plot_confusion_matrix(model_logit,x_test_b,y_test_b)
```

In [259]:

```
#building model  
logregdis=LogisticRegression()  
logregdis.fit(x_train1,y_train1)
```

Out[259]:

```
LogisticRegression()
```

In [260]:

```
#model accuracy  
# method 1  
logregdis.score(x_test1,y_test1)
```

Out[260]:

```
0.72
```

In [261]:

```
#building model  
model_logit=LogisticRegression()  
model_logit.fit(x_train_d,y_train_d)
```

Out[261]:

```
LogisticRegression()
```

In [262]:

```
#model accuracy  
# method 1  
model_logit.score(x_test_d,y_test_d)
```

Out[262]:

```
0.7142857142857143
```

## Chapter 7

# Discussion and Conclusion

---

We can view the progression of a deteriorating patient with COVID-19 along a pathway: they must become infected, symptomatic, and then deteriorate and be appropriate for hospital treatment. To receive ICU treatment, they must be normally well enough that it is felt they will benefit, but also acutely unwell enough that it is necessary. It is clear that in Bradford at least, patients who were not felt appropriate for ICU represented the vast majority of deaths . Analyses of mortality risk in COVID-19 must therefore pay attention to the point at which the patient becomes at-risk.

This study agrees with other work that the risk of hospital mortality rises with increasing age and male sex, as well as obesity and a variety of other comorbidities. Patients with diabetes, hypertension, cardiovascular disease or renal disease or a combination of any of these were most likely to experience severe outcomes of the infection. The same factors seem to relate to a risk of ICU admission, except that there is a more complex relationship in that people with some comorbidities or higher age may be considered too frail to benefit from intensive treatment despite severe disease.

## REFERENCES

---

- [1] Lauren A. Callender, Michelle Curran, et al. “The Impact of Pre-existing Comorbidities and Therapeutic Interventions on COVID-19 ”, *Front. Immunol.*, 11 August 2020.
- [2] Safiya Richardson, Jamie S. Hirsch, Mangala Narasimhan, et al “Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area ” *JAMA*. 2020;323(20):2052–2059. doi:10.1001/jama.2020.6775 , April 22, 2020.
- [3] Xu, Wan et al. “Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning” *Scientific reports* vol. 11,1 2933. 3 Feb. 2021, doi:10.1038/s41598-021-82492-x.
- [4] Santorelli, Gillian et al. “Ethnicity, pre-existing comorbidities, and outcomes of hospitalised patients with COVID-19 ” *Wellcome Open Research* vol. 6 32. 30 Jun. 2021, doi:10.12688/wellcomeopenres.16580.2.
- [5] Paul M. McKeigue, Amanda Weir, et al. “Rapid Epidemiological Analysis of Comorbidities and Treatments as risk factors for COVID-19 in Scotland (REACTSCOT): A population-based case-control study”, October 20, 2020 .
- [6] Mohammad Rahanur Alam, et al. “Comorbidities might be a risk factor for the incidence of COVID-19: Evidence from a web-based survey ”, *Preventive Medicine Reports* Volume 21, March 2021, 101319 .
- [7] Stephanie L. Harrison, et al. “Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis ”, September 10, 2020 <https://doi.org/10.1371/journal.pmed.1003321>.
- [8] Juhyun Song, et al. “Clinical course and risk factors of fatal adverse outcomes in COVID-19 patients in Korea: a nationwide retrospective cohort study ”, 12 May 2021.

**USING PSYCHOLOGICAL CAPITAL TO PREDICT MENTAL HEALTH  
AMONG DEGREE COLLEGE STUDENTS**

Dissertation submitted in partial fulfilment of the requirements for the award of  
Master of Science in Psychology

By

**ELSA JOHN**

**Register No: SM20PSY006**

Under the guidance of

Ms. Bindu John

Assistant Professor, Department of Psychology



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

Nationally Re-accredited at 'A++' level (4th cycle)

Affiliated to: Mahatma Gandhi University

**MARCH 2022**

**CERTIFICATE**

This is to certify that the dissertation entitled, “**USING PSYCHOLOGICAL CAPITAL TO PREDICT MENTAL HEALTH AMONG DEGREE COLLEGE STUDENTS**”, is a bonafide record submitted by Ms. Varsha S, Reg.no. SM20PSY022, of St. Teresa’s College, Ernakulam under my supervision and guidance and that it has not been submitted to any other university or institution for the award of any degree or diploma, fellowship, title or recognition before.

Ms. Bindhu John  
Head of the Department  
Department of Psychology  
St. Teresa’s College, Ernakulam

Date: 30/05/2022

External Examiner 1: .....

External Examiner 2: .....

Internal Examiner: .....

## **DECLARATION**

I, Elsa John, do hereby declare that the work represented in the dissertation embodies the results of the original research work done by me in St. Teresa's College, Ernakulam under the supervision and guidance of Ms. Bindu John, Assistant Professor, Department of Psychology, St. Teresa's College, Ernakulam, it has not been submitted by me to any other university or institution for the award of any degree, diploma, fellowship, title or recognition before.

Elsa John

Place: Ernakulam

Date: 3/05/2021

## **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. Bindu John, 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  
Elsa John

**CERTIFICATE**

This is to certify that this project is an authentic record of research carried out by Elsa M John under my guidance and supervision, and that no part of it has been presented before for any other degree, diploma, or title

Date: 30/05/2022

Place: Ernakulam

Ms. Bindu John

Head of the Department

Ms. BINDU JOHN

M.A (Psy), B.Ed AIMA Accredited

Management Teacher- Human

Resource Management

Faculty Guide

**DECLARATION**

I, Elsa M John, hereby declare that the project entitled, “Using Positive Psychological Capital to predict Mental Health among degree college students.” has not been submitted by me for the award of any degree, title or recognition, before.

Place: Ernakulam

Ms. Elsa M John

Date: 30/05/2022

<b>SL NO</b>	<b>CONTENTS</b>	<b>PAGE NO</b>
1.	INTRODUCTION	11
2.	REVIEW OF LITERATURE	21
3.	METHODOLOGY	29
4.	RESULT AND DISCUSSION	41
5.	CONCLUSION	40
6.	REFERENCES	42
7.	APPENDICES	50

**LIST OF TABLES**

<b>Table No</b>	<b>List of tables</b>	<b>Page no</b>
1.	Mean, and Standard Deviation of Mental Health and Psychological Capital over College life and Overall life	31
2.	Correlation between Mental Health and Psychological Capital	32
3.	Multiple comparisons table	33
4.	ANOVA of the categories of Mental Health	34
5.	Linear regression of Overall-life Psychological Capital on Mental Health	35
6.	Linear regression of College-life Psychological Capital on Mental Health	36

**LIST OF FIGURES**

<b>Figure No</b>	<b>List of figures</b>	<b>Page No</b>
1.	Correlation between Mental Health and Psychological Capital	32
2.	Mean plot of categories of Psychological Capital	34
3.	Linear regression of Overall-life Psychological Capital on Mental Health	35
4.	Linear regression of College-life Psychological Capital on Mental Health	36

**LIST OF APPENDICES**

<b>SL NO</b>	<b>APPENDICES</b>	<b>Page No</b>
1.	DEMOGRAPHIC DETAILS	50
2.	MENTAL HEALTH CONTINUUM- SHORT FORM	52
3.	ACADEMIC PSYCHOLOGICAL CAPITAL QUESTIONNAIRE	54

## **CHAPTER 1: INTRODUCTION**

Mental health problems among college students have been increasing in prevalence and severity during recent times (Blanco et al., 2008; Drum, Brownson, Burton Denmark, & Smith, 2009; Eisenberg, Hunt, & Speer, 2013; Fink, 2014; Gallagher, 2008; Hunt & Eisenberg, 2010). Since the last decade of twentieth century, mental illness has been considered as a public health issue alongside other physical illness (Keyes, 2014). Rather than focusing on the promotion of positive mental health, policy makers and scholars are still focused on addressing mental illness and instituting risk reduction programs. Efforts and initiatives to shift the conception of mental health from the mere “absence of mental illness” to something complete and positive, or to the presence of “well-being” have not yet been significantly effective. (Keyes, 2002)

The WHO emphasizes the need to promote positive mental well-being by defining good mental health as “... a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community” (Seligman, 2012; World Health Organization [WHO], 2004) (WHO, 2004). Furthermore, the Keyes’ mental health continuum model is exemplified in the WHO’s perspective on mental health—as a complete state that represents not only absence of disease, but also consists of the presence of positive states of human capacities and functioning categorized along a continuum by levels of flourishing, moderate, and languishing. (Keyes,2002, 2007, 2013).

## **1.1 BACKGROUND OF THE STUDY**

In the past few years, mental health issues such as depression, anxiety, suicidal ideation, and self-injury among college students are growing at an alarming rate. Since the last decade of twentieth century, mental illness has been considered a public health issue alongside other physical illness. Recent studies indicate that there is an upward trend in the number of students with severe psychological problems which makes the mental health profile of college students a growing concern. Attending college has a dual impact on an individual. It marks a significant and stressful developmental period of an individual's life. On the other hand, it is at this phase of life where various psychological disorders have their first onset. It is imperative that mental health professionals at the college level consider the problems associated with identity formation and the gradual manifestations of mental illnesses during these high-risk early adult years because the prevalence of developmental and mental health issues in college student communities is becoming more diverse and complex than ever in the history of higher education. (Christine S Bhat., 2015)

Also, the concept of Psychological Capital which is not much familiar among all, is strongly linked to increased well-being, work and life satisfaction. Therefore, it's necessary that the upcoming future, i.e.; the college students' mental health and PsyCap are assessed to help students persevere in their studies in a psychologically healthier manner. As PsyCap and its dimensions are malleable, (Luthans et al., 2007; Luthans et al.,2012) they can be developed through training interventions along with various positive psychology exercises that focuses on enhancing well-being in conjunction with PsyCap. (Luthans, Avey, & Patera, 2008; Luthans, Avey, Avolio, Norman, & Combs, 2006; Luthans, Avolio, Avey, & Norman, 2007; Luthans, Youssef, & Avolio, 2007). Exploring further in this regard revealed a variety of potent, low- cost approaches that impact mental health through development of positive strengths such as hope, efficacy, resilience, and optimism. (Luthans et al., 2008)

## **1.2 PROBLEM STATEMENT**

The mental health of college students is a growing concern as there has been a significant hike in the number of youngsters with mental health issues. (Blanco et al., 2008; Drum, Brownson, Burton Denmark, & Smith, 2009; Eisenberg, Hunt, & Speer, 2013; Fink, 2014; Gallagher, 2008; Hunt & Eisenberg, 2010). Also, individuals are prone to develop mental health

problems during this age period and health professionals need to pay closer attention to college students. The concept of Psychological Well-being is important since it is strongly linked to increased well-being, work and life satisfaction. Therefore, it's necessary that the upcoming future, i.e.; the college students' mental health and Psychological Capital are assessed and help students persevere in their educational studies in a psychologically healthier manner. As PsyCap dimensions are malleable, they can be developed through trainings interventions.

### **1.3 NEED AND SIGNIFICANCE OF THE STUDY**

By drawing connections among a management/organizational concept and an overarching positive psychology approach of mental health among within student population, the study challenges the conventional conceptual frameworks for mental health by enlightening the positive aspects of human functioning and wellness, as well as by emphasizing the significance of individual positive strengths. In doing so, this study addressed knowledge gap in the literature and added implications to the counseling and higher education body of research.

In order to attain optimal mental health (flourishing), the findings of this study can be helpful to professionals in college-settings to redirect their focus on developing positive psychological resources such as the hope, efficacy, resiliency, and optimism. The findings of this research study also help students persevere in their educational studies in a psychologically healthier manner, making a case of integrating PsyCap into academic curriculum. As PsyCap and its dimensions are malleable (Luthans et al., 2007; Luthans et al., 2012), they can be developed through training interventions and various positive psychology exercises that focuses on enhancing well-being in conjunction with PsyCap (Luthans, Avey, & Patera, 2008; Luthans, Avey, Avolio, Norman, & Combs, 2006; Luthans, Avolio, Avey, & Norman, 2007; Luthans, Youssef, & Avolio, 2007). Exploring further in this regard revealed a variety of potent, low-cost approaches that impact mental health through development of positive strengths such as hope, efficacy, resilience, and optimism. (Luthans et al., 2008)

This study proposes to explore the role of PsyCap and additional correlates (academic achievement, and socio-demographics such as gender, age, race/ ethnicity, sexual orientation,

current educational degree) in predicting mental health to bring a positive light on mental health and psychological resources of college students. This study utilizes the overarching positive psychology framework to investigate various positive psychological constructs among college student population (Seligman, 2002, 2003; Seligman & Csikszentmihalyi, 2000; Seligman et al., 2005). The rationale for developing various programs and initiatives that ensure fostering higher levels of positive mental health (flourishing) in college campuses will be provided by the findings of this study.

#### **1.4 SCOPE OF THE STUDY**

The scope of this study is limited only to degree students from Kerala.

## **CHAPTER -2: REVIEW OF LITERATURE**

## **2.1 REVIEW OF LITERATURE**

In 2010, Braithwaite et al. used the National College Health Assessment on a population of 1621 college students to compare the difference in the degree of well-being between single college students and college students in committed romantic relationships. The results indicated that students in committed romantic relationships had fewer mental health problems than the single students. Chen & Lucock in 2022, studied the impact of the COVID-19 pandemic on the Levels of mental health and quality of life in 1173 university students during the COVID-19 pandemic using Patient health questionnaire (PHQ-9), Generalized anxiety disorder questionnaire (GAD-7), Brief resilience scale (BRS), EQ-5D-5L, COVID-19 related questions. Results indicated that high levels of anxiety and depression, with more than 50% experiencing levels above the clinical cut offs, and females scoring significantly higher than males. It also suggested relatively low levels of resilience, Higher levels of distress were associated with lower levels of exercising, higher levels of tobacco use, and a number of life events associated with the pandemic and lockdown. In the year 2021, Prasath et al., conducted a study on the relationships between psychological capital (PsyCap), coping strategies, and well-being among 609 university students using self-report measures. Results revealed that well-being was significantly lower during COVID-19 compared to before the onset of the pandemic. It also indicated that PsyCap predicted well-being mediating role of coping strategies between PsyCap and well-being. In 2019, Gujar & Ali studied the effects of psychological capital and self-esteem on emotional and behavioral problems among 387 adolescents using PsyCap Questionnaire, Rosenberg Self-Esteem Scale, and Strengths and Difficulties Questionnaire (self-report). Results showed that Emotional and behavioral problems of the adolescent population are closely related to self-esteem and PsyCap. Son et al., in 2020, conducted a study on the effects COVID-19 on 195 college students using Mental Health Questionnaire, Psychological well-being scale. Results showed that Pandemic has a great effect on the Mental Health and Well-being of students. In the year 2019, Sarshar et al., conducted a study on 261 high school students to investigate the relationship between psychological capital and perfectionism using PsyCap Questionnaire (PCQ)

& Frost Multidimensional Perfectionism Scale. Results showed that Four aspects of psychological capital affect the negative perfectionism. Yan et al., in 2016 conducted a study to explore the influences of coping styles on interpersonal relationships and academic performances when PsyCap acts as the mediating variable, and to explore the influences of family relations on PsyCap when coping styles acts as the mediating variable on 832 college graduates using PsyCap Questionnaire (PCQ) & Simplified Coping Style Questionnaire (SCSQ). Results showed that Coping styles of college graduates have a significant impact on interpersonal relationships and academic performances with PsyCap play a mediating role. In 1988, Warr et al., conducted a study on the Effect of unemployment on mental health on 3077 unemployed people using General Health Questionnaire. Results indicated that there is Significant relation between unemployment and mental health. Truong & Ma, in 2006 Study on the Relation between Neighborhoods and Mental Health on 1600 individuals using General Health Questionnaire. Results revealed that There is Significant relation between Mental Health and neighborhood characteristics. In the year 2019, Hernández-Varas et al., studied to examine whether psychological capital, work satisfaction and health self-perception are able to predict psychological wellbeing in a military population on 492 Spanish soldiers using PsyCap questionnaire, Minnesota Job Satisfaction Questionnaire, SF- 36D Health Survey Questionnaire. Results showed a positive significative correlation detected between the variables psychological capital, work satisfaction and health self-perception and psychological wellbeing. Avey et al., in 2010 studied the Impact of PsyCap on employee well-being overtime on 280 employees from various occupations using PCQ-24, Index of PWB and General Health Questionnaire and the following results showed that there is Significant relation between PsyCap and Well-being. Alzyoud & Mert, in the year 2019 studied to find whether PsyCap is positively correlated with Organisational identity on Chinese workers using Fred Luthans PsyCap Questionnaire and Milleretal Organisational Identification Scale. Results showed that PsyCap is positively related to Organisational Identity. In 2020, Santisi et al., studied the relationships between psychological capital and two dimensions of quality of life on 807 Italian workers using Psychological Capital Questionnaire, Courage Measure, Satisfaction with Life Scale, and Flourishing Scale. results indicated that psychological capital predicts life satisfaction and flourishing, with an indirect mediation effect of courage only on flourishing. Hashmi et al., in 2022 Study to determine the interactive effects of sexual harassment and psychological capital on victims' burnout using the

evidence from the post-#MeToo movement era on 304 Female Employees using Sexual Harassment and Gender Harassment related questions, PsyCap Questionnaire. Results found out that Gender harassment leads to burnout causing emotional exhaustion, depersonalization and reduced personal accomplishment among the victims. However, psychological capital helps victims cope with these adverse effects of gender harassment. In 2020, Tian et al., Study to explore the association of occupational stress with fatigue and to examine the mediating role of psychological capital (PsyCap) among 1104 Chinese physicians using 14-item Fatigue Scale (FS-14), the Effort-reward Imbalance questionnaire (ERI), the Psychological Capital Questionnaire (PCQ). Results showed that Occupational stress and PsyCap were significantly associated with fatigue. PsyCap significantly mediated the association. Dhiman & Arora, in 2018 Study to determine the significant negative relationship between psychological capital and turnover intentions of 230 paramedical & nursing staff using PCQ scale & turnover intention was measured using a three-item measure based on Mobley, Horner & Hollingsworth theory (1978). Results showed that psychological capital has a significant negative relationship with turnover intentions. Kazemian Moghadam & Haroon Rashidi, in the year 2020 conducted a study to determine the significant relationship of psychological capital and spirituality at work with psychological well-being in 123 nurses using PsyCap questionnaire (Luthans, 2007), psychological well-being scale (Ryff, 1989), and spirituality at workplace questionnaire (Kinjerski and Skrypnek, 2006). Results revealed that psychological capital and spirituality in work are considered in programs to promote mental health in nurses. In 2019, Kawalya et al., studied to determine the significant and positive relationship between PsyCap and flow experience on 800 professional nurses using PsyCap Questionnaire (Luthans). Results showed that there is a positive and significant relationship between PsyCap and flow experience among Nurses. In 1992, Norris et al., conducted a study to determine the significant effects of indoor and outdoor exercise sessions on acute wellbeing and stress levels on 140 collegiate sports participants using Mental Health Questionnaire. Results showed that Engagement in outdoor exercise did not lead to more beneficial changes than engagement in exercise sessions indoors, outdoor exercise environments were perceived as more calming and exercise sessions in more calming environments were associated with more stress-reductive effects. Robinson et al., in 2021, Studied to find whether there is a significant extent to which mental health symptoms changed during the pandemic in 2020 on 200 samples using Mental Health Questionnaire.

Results showed that compared to pre-pandemic outbreak, there was an overall increase in mental health symptoms observed during March-April 2020. In the year 2021, Munawer et al., conducted a study to determine whether PsyCap is significantly correlated with mental health on 400 Graduate Students using Academic Psychological Capital Questionnaire, Mental Health Continuum-Short Form (MHC –SF). Results showed that the dimensions of PsyCap(HERO) were correlated with the mental health of the graduate students, and there was a positive predictive relationship between PsyCap and Mental Health. Sarif et al., in the year 2021, Studied to determine significant relationship between Mental Health and PsyCap of Digital Learning students amidst Covid 19 on 200 students using PsyCap Questionnaire (Luthans), Mental health Continuum. Results indicated that there are effects on mental health of digital learning students and PsyCap can be used as a coping strategy. Younas et al., conducted a study to examine the relationship between mental health and psychological capital among young adults in the year 2020 on 200 University students using Strength and Difficulty Questionnaire (SDQ), Psychological Capital Questionnaire (PsyCap). Results showed that PsyCap among young adults leads towards better mental health as it is negatively associated with internalizing (emotional problems) and externalizing (pro-social) behavior. In 2015, Krasikova et al., did a Study to find whether psychological capital is negatively related to diagnoses for mental health problems (PTSD, anxiety, and depression), Psychological capital is negatively related to diagnoses for substance abuse on 1,889 U.S. Army soldiers using Global Assessment Tool, Post Deployment Health Questionnaire, PCQ-12. Results showed that PsyCap affected substance abuse diagnoses both directly and indirectly. Razaq & Latif, n.d. conducted a study to determine the relationship between Psychological Capital, Mental Health and Burnout among Professionals working in Autism Centers in Punjab Pakistan in the year 2021, on 450 Professionals who handle children with autism using Psychological Capital Questionnaire, General Health Questionnaire and Teacher Burnout Inventory. Results showed that There is a significant positive relationship between Psychological Capital and Mental Health. In 2017, Haleem et al., conducted a study to determine the significant relationship between psychological capital and mental health of rescue workers on 502 male rescue workers with their age ranging from 22 to 44 years using Urdu versions of Mental Health Inventory & Psychological Capital Questionnaire. Results found out that there is significant relationship between psychological capital and mental health of rescue workers. Estiri et al., in 2016, conducted a study on the Impact of psychological capital on

mental health by investigating the mediating effects of job burnout on this relationship on 384 Iranian Nurses using PCQ-24, Burnout Inventory (MBI), General Health Questionnaire. Results proved that there is a significant relationship between psychological capital, job burnout and mental health; also, there is a significant negative relationship between psychological capital and job burnout, and a significant positive relationship between psychological capital and mental health.

## **2.2 RESEARCH GAP**

In recent years' mental health issues such as depression, anxiety, suicidal ideation, and self-injury among college students has been showing an upward trend. Since the last decade of twentieth century, mental illness has been considered a public health issue alongside other physical illness. The mental health profile of college students is a growing concern as there is an upward trend in the number of students with severe psychological problems being reported in recent studies. y (Blanco et al., 2008; Drum, Brownson, Burton Denmark, & Smith, 2009; Eisenberg, Hunt, & Speer, 2013; Fink, 2014; Gallagher, 2008; Hunt & Eisenberg, 2010). The prevalence of developmental and mental health issues in college student communities is currently becoming more diverse and complex than ever in the history of higher education. Attending college marks a significant and stressful developmental period of an individual's life. Various psychological disorders have their first onset during this period. It is therefore imperative that mental health professionals at the college level consider the problems associated with identity formation and the gradual manifestations of mental illnesses during these high-risk early adult years.

It is necessary that the upcoming future, i.e.; the college students' mental health and PsyCap are assessed to help students persevere in their educational studies in a psychologically healthier manner. Finding the relationship of sociodemographic variables with mental health would pave the way for researchers to further examine distinct and 'at-risk' sections of the population. Furthermore, the findings of this study will have the potential to guide and frame student development programs for college students in both overall life and academic areas. The results may be applicable for individuals' as well as for institutional development.

## **CHAPTER-3: METHODOLOGY**

### **3.1 AIM**

To use Positive Psychological Capital to predict Mental Health among Degree College Students.

### **3.2 STATEMENT PROBLEM**

Growing concern regarding the Mental Health of college students.

### **3.3 OBJECTIVES**

The goal of the current research is to test the following research objectives:

1. To describe the level of psychological capital and mental health of college students using their PsyCap and Mental Health scores.
2. To explore the relationships between psychological capital and mental health of college students in the studied sample
3. To determine if differences exist in PsyCap of college students whose scores place them in one of three mental health continuum categories languishing, moderately mentally healthy, and flourishing.
4. To determine the extent to which PsyCap HERO dimensions within the Overall-life and the college-work categories predict the variability in mental health among college students.

### **3.4 HYPOTHESES**

1. There will be a significant relationship between Psychological Capital and Mental Health.

2. There will be a significant difference between the Psychological Capital of college students who fall in the Languishing, Moderately Mentally Healthy and Flourishing categories.

### **3.5 OPERATIONAL DEFINITIONS**

**3.5.1 Mental health-** Keyes (2002) defines mental health as “a syndrome of symptoms of positive feelings and positive functioning in life” (p. 207). The three categories of mental health on the continuum are languishing, moderately mentally healthy, and flourishing (Keyes, 2002, p. 210). Keyes’ describes dimensions of subjective well-being that accounts for mental health symptoms, broadly are of three clusters- (i) Emotional Well-Being; (ii) Psychological well-being; and (iii) Social well-being (Keyes, Shmotkin, & Ryff, 2002, p. 1007).

➤ **Emotional well-being.** Emotional well-being is a specific dimension of subjective well-being and a subset of symptoms used to diagnose states of mental health, which reflects on the degree to which individuals self-report the experience of symptoms of positive and negative affect (Keyes, 2000, p. 71).

➤ **Psychological well-being.** *Psychological well-being* is said to result when an individual is finding meaning and purposeful direction in life, accepting oneself, seeking continued personality development, acting and thinking anonymously and establishing potential

relationships with others (Ryff & Keyes, 1995; p. 722). Also referred as *Eudaimonic well-being* (Ryan & Deci, 2001).

➤ **Social well-being.** Keyes' (1998) states that, "individuals are mentally healthy when they view special life as meaningful and understandable, when they see society as possessing potential for growth, and when they feel they belong in their communities, are able to accept all parts of society, and when they see their lives as contributing to society".

**3.5.2 Positive psychological capital-** Luthans, Youssef, and Avolio (2007) defines PsyCap as a positive psychological strength or resource "characterized by:

(a) having confidence (self-efficacy) to take on and put in the necessary effort to succeed at challenging tasks;

(b) making a positive attribution (optimism) about succeeding now and in the future;

(c) persevering toward goals and, when necessary, redirecting paths to goals (hope) in order to succeed; and

(d) when beset by problems and adversity, sustaining and bouncing back and even beyond (resilience) to attain success".

## **3.6 SAMPLE**

The size of the sample was 195 undergraduate students above the age of 18 from different colleges of the Kerala State.

### **3.6.1 POPULATION**

The population considered to draw conclusions from was undergraduate students within the age group of 18-22.

### **3.6.2 SAMPLING DESIGN**

Descriptive Cross Sectional Research design was used to study the relationship between Psychological Capital and Mental Health of college going undergraduate students.

- Inclusion Criteria-
  - Undergraduate students
  - Students within the age group of 22
  - Students from colleges in Kerala.
  
- Exclusion Criteria
  - Post graduate, Mphil or PhD Students.

### **3.7 TESTS OR TOOLS USED FOR DATA COLLECTION**

1) Mental health is assessed using Keyes' Mental Health Continuum- Short Form [MHCSF] by Corey L M Keyes. The three well-being clusters assessed by the MHC-SF are emotional well-being, social well-being, and psychological well-being.

Purpose- MHCSF is used to assess the Mental Health of the students.

Scoring- Items are summed, yielding a total score ranging from 0 to 70. Subscale scores range from 0 to 15 for the emotional (hedonic) well-being, from 0 to 25 for social well-being, and from 0 to 30 for psychological well-being. Flourishing mental health is defined by reporting  $\geq 1$  of 3 hedonic signs and  $\geq 6$  of 11 eudaimonic signs (social and psychological subscales combined) experienced "every day" or "5-6 times a week." Higher scores indicate greater levels of positive well-being.

Validity- The MHC-SF has been validated (discriminant validity) and has shown excellent internal consistency ( $> 0.80$ ) in adolescents of 12 through 18 years and adults in the United States (Keyes, 2009). In addition to the U.S, this instrument has been extensively tested-retested for assessing reliability and validity in Netherlands and South Africa population.

Reliability- The three successive three-month periods test-retest reliability of the instrument averaged 0.68, additionally the 9-month test-retest averaged 0.65 (Keyes, 2007; Lamers et al., 2011). Additionally, the test-retest reliability for the Long-Form was over a time period of 4 weeks and estimated as 0.57 for the overall psychological well-being domain, 0.64 for the overall emotional well-being domain, and 0.71 for the overall social well-being domain (Keyes, 2009; Robitschek & Keyes, 2009).

Thus, the MHC-SF is an empirically tested, highly reliable and valid instrument to measure positive mental health cross-culturally and over a period of time.

II) PsyCap in this study is measured using Academic PsyCap Questionnaire (A-PCQ) modified and adapted by Luthans, Luthans, and Jensen (2012). PsyCap score in A-PCQ are assessed using two distinct domains: Overall-life (OL) and School-work (SW), and an overall PsyCap score.

Purpose- Academic Psychological Capital Questionnaire is used to assess positive Psychological Capital of the students.

Scoring- PCQ scales are scored using a 6-point Likert scale. Each dimension has 6 items respectively. Each scale measures its own unique dimension (HERO) of PsyCap. An overall PsyCap score is calculated by taking the mean of all the items in the PCQ. Items 13, 20 and 23 are reverse scored.

Validity-

***Discriminant/convergent validity.*** Youssef and Luthans (2007) report that each of the four constituting positive constructs (HERO) shows empirically based discriminant validity in addition to previous studies of Bryant and Cvangros (2004), Carifio and Rhodes (2002), and Magaletta and Oliver (1999). PsyCap was found to be not related to age or education demographics and personality dimensions of agreeableness or openness (Luthans, Avolio, et al., 2007). However, in the same study, PsyCap was strongly associated to core self-evaluations (0.60) and moderately related to extraversion (0.36) and conscientiousness (0.39). In the studies with PsyCap, the regression model without the PsyCap composite was found to

be significant ( $R_2 = 0.13, p < 0.001$ ), however the change in  $R_2$  was also significant ( $\Delta R_2 = 0.04, p < 0.001$ ) (Luthans et al., 2007; Luthans, Avolio et al., 2007). This indicated that PsyCap predicted unique variance in job satisfaction which was beyond the two personality traits and core self-evaluations. PsyCap was confirmed to be the greater contributor in predicting affective organizational commitment as the beta weight for PsyCap was largest in the regression model (Luthans et al., 2007).

**Criterion validity.** Study by Luthans, Avolio et al. (2007) found that PsyCap had a stronger relationship to job satisfaction ( $p < 0.01$ ) than conscientiousness and extraversion, however, not as much as core-evaluations and affective organizational commitment had ( $p < 0.001$ ) with PsyCap. Studies have reported that impact of PsyCap is greater in studies based in the US in comparison to the studies outside of the US (Avey, Wernsing, & Mhatre, 2011). Additionally, PsyCap's impact varied based on the sample, for example, the service industry versus manufacturing industry (Luthans et al., 2007). However, it is empirically validated that PsyCap significantly predicted variances in a wide array of attitudes and behaviors that comprised various demographics characteristics and personality traits (Luthans et al., 2007).

**Reliability-** The Cronbach's alpha reliability for A-PCQ in the preliminary study with college students was 0.90 (Luthans et al., 2012) and for the study is 0.95. Dr. Brett Luthans, developer of A-PCQ indicates that all the psychometric properties of the instrument remain the same as that of original PCQ. In calculating reliability estimate for the total PsyCap and each adapted measure from four sample populations, Luthans et al. (2007) found that reliability of the overall PsyCap measure in all samples was consistently above conventional standards (Luthans, Avolio, Avey, & Norman, 2007). The Cronbach alphas as mentioned in Luthans et al. (2007) were: overall PsyCap (0.88, 0.89, 0.89, 0.89); hope (0.71, 0.75, 0.80, 0.76); efficacy (0.75, 0.84, 0.85, 0.75); resiliency (0.71, 0.71, 0.66, 0.72); optimism (0.74, 0.69, 0.76, 0.79). Internal consistency reliability for optimism and resilience were found to be consistently lower than self-efficacy and hope domains (Dawkins, Martin, Scott, & Sanderson, 2013). However, Luthans believed that this difference is because of the reverse-scored items in resilience and optimism scales that reduce scale reliability (Schmitt & Stults, 1985).

### **3.8 PROCEDURE**

The data was collected using the Questionnaire method of Qualitative Data Collection which comes under the Primary data collection method which is the collection of data directly from the source or the sample. A questionnaire involves reading of each question and responding to them as required.

The sample for this study was undergraduate students who are above the age of 18. 195 was the sample size and the data was collected by distributing the Mental health questionnaire, PsyCap questionnaire, demographic details along with the consent form. The questionnaires were printed and distributed at All Saints college Trivandrum, Mother Teresa College Trivandrum, Mar Ivanios College Trivandrum, Catholicate College Pathanamthitta, Pamba college Thiruvalla and Kristu Jyoti College Changanaserry.

### **3.9 DATA ANALYSIS**

- Data analysis is done using SPSS version 26.0.
- Descriptive statistics is used.
- Pearson Correlation
- Homogeneity of Variance found using Levene's Test
- ANOVA

## **CHAPTER-4: RESULTS AND DISCUSSION**

## **RESULT**

In the past few years, mental health issues such as depression, anxiety, suicidal ideation, and self-injury among college students are growing at an alarming rate. Since the last decade of twentieth century, mental illness has been considered a public health issue alongside other physical illness. Recent studies indicate that there is an upward trend in the number of students with severe psychological problems which makes the mental health profile of college students a growing concern y (Blanco et al., 2008; Drum, Brownson, Burton Denmark, & Smith, 2009; Eisenberg, Hunt, & Speer, 2013; Fink, 2014; Gallagher, 2008; Hunt & Eisenberg, 2010).

Attending college has a dual impact on an individual. It marks a significant and stressful developmental period of an individual's life. On the other hand, it is at this phase of life where various psychological disorders have their first onset. It is imperative that mental health professionals at the college level consider the problems associated with identity formation and the gradual manifestations of mental illnesses during these high-risk early adult years because the prevalence of developmental and mental health issues in college student communities is becoming more diverse and complex than ever in the history of higher education.

Also, the concept of Psychological Capital which is not much familiar among all, is strongly linked to increased well-being, work and life satisfaction. Therefore, it's necessary that the upcoming future, i.e.; the college students' mental health and PsyCap are assessed to help students persevere in their studies in a psychologically healthier manner. As PsyCap and its dimensions are malleable, they can be developed through training interventions along with various positive psychology exercises that focuses on enhancing well-being in conjunction with PsyCap (Luthans, Avey, & Patera, 2008; Luthans, Avey, Avolio, Norman, & Combs, 2006; Luthans, Avolio, Avey, & Norman, 2007; Luthans, Youssef, & Avolio, 2007). Exploring further in this regard revealed a variety of potent, low- cost approaches that impact mental health through development of positive strengths such as hope, efficacy, resilience, and optimism.

The primary purpose of this cross-sectional, exploratory study was to examine the relationships and predictions that exist between current levels of mental health and psychological capital [PsyCap] among college students as measured by the Keyes' Mental Health Continuum Short Form [MHC-SF] and Academic Psychological Capital Questionnaire [A-PCQ]. Mental health was measured by MHC-SF and participants were categorized into three groups languishing, moderately mentally healthy, and flourishing. Similarly, PsyCap was measured using A-PCQ and participants were measured for their current level of Hope, Efficacy, Resilience, and Optimism [HERO]. The main aim of the study was to establish the associations between the studied variables and to test psychological capital as a predictor of mental health.

### Research Objective 1

The first objective was to describe the level of psychological capital and mental health of college students using their PsyCap and Mental Health scores. This objective focused on describing the current level of PsyCap and mental health of the studied sample using basic descriptive statistics.

*Table 1.*

*Mean, and standard deviation of Mental Health and Psychological Capital over College life and Overall Life.*

Variables	N	Mean	Std. Deviation
Mental Health	195	46.15	12.741
Overall Life	195	96.72	19.099
College life	195	98.01	17.976
Psychological Capital	195	194.73	35.825

Table 1 gives the Mean and Standard Deviation of Mental Health variable and Psychological Capital variable. The Mean and Standard deviation for the Mental Health variable was found to be 46.15 and 12.741 respectively. The range of scores of the sample lies between 9 and 70. The Mean and Standard Deviation of Psychological Capital was found to be 194.73 and 35.825 respectively. The score range of the sample lies between 86 and 277.

### Research Objective 2

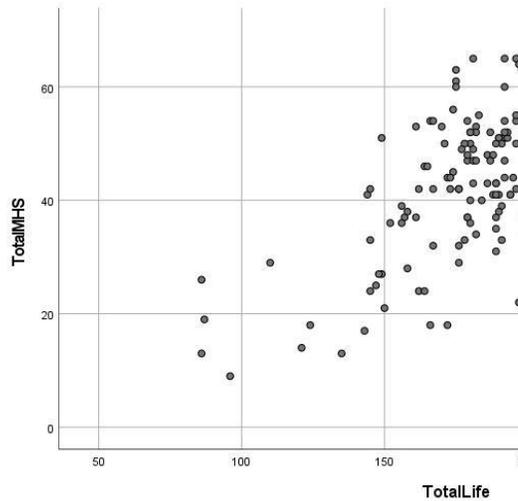
The second research objective was to explore the relationship between psychological capital and mental health of college students in the studied sample. The relationship between PsyCap scores and mental health scores were analyzed using Pearson correlation coefficient.

*Table 2*

*Correlation between Mental Health and Psychological Capital*

		Mental Health
Psychological Capital	Pearson Correlation	.650**
	Sig. (2-tailed)	.000
	N	195

\*\* . Correlation is significant at the 0.01 level (2-tailed).



The correlation coefficient between Mental Health and Psychological Capital was found to be statistically significant with the p value less than 0.05 ( $r=0.650$ ,  $p < 0.01$ ). Hence, we can accept the hypothesis that there is a relationship between Mental Health and Psychological Capital. This clearly indicates a significant moderate positive correlation between Mental Health and Psychological Capital.

### Research Objective 3

The third research objective is to determine if differences exist in PsyCap of college students whose scores place them in one of three mental health continuum categories languishing, moderately mentally healthy, and flourishing.

*Table 3.*

#### *Multiple Comparisons*

*Dependent Variable: Total Life*

<b>1. Category</b>	(J) Category	Mean Difference (I-J)	Std. Error	Sig.
<b>1</b>	2	-71.451*	16.794	.000

	3	-106.250*	18.831	.000
<b>2</b>	1	71.451*	16.794	.000
	3	-34.799*	9.221	.001
<b>3</b>	1	106.250*	18.831	.000
	2	34.799*	9.221	.001

\* The mean difference is significant at the 0.05 level.

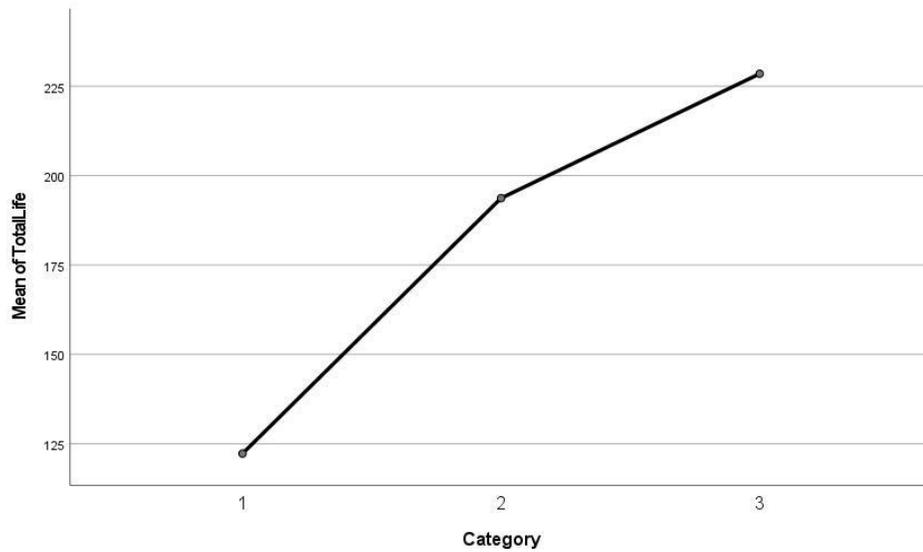
The normality of the data was ensured by performing Levene's Test of Homogeneity of variances. A p value of more than 0.05 was obtained which led to reject the null hypothesis and accept the Homogeneity of variances. Hence, ANOVA was performed on three groups to compare them based on the variable of PsyCap.

Table 4.

*ANOVA of the categories of Mental health*

Mental health	N	Mean	Std. Deviation	F	Sig.
Languishing	4	122.25	39.331		
Moderately Mentally Healthy	177	193.70	32.951	16.845	.001
Flourishing	14	228.50	35.163		

Table 4 shows the result of the ANOVA Test conducted on categories of Mental health. The F- value obtained is 16.845 and the corresponding p-value is .001. This clearly shows that there is a significant difference between the PsyCAP of college students whose scores place them in languishing, moderate and flourishing. Hence, we accept the hypothesis that there is a significant difference between the PsyCAP of college students whose scores place them in languishing, moderately mentally healthy and flourishing. The mean value of Mental health was 122.25 (SD= 39.331) for the Languishing group, 193.70 (SD= 32.951) for the Moderately Mentally Healthy group and for the Flourishing group 228.50 (SD= 35.163).



#### Research Objective 4-

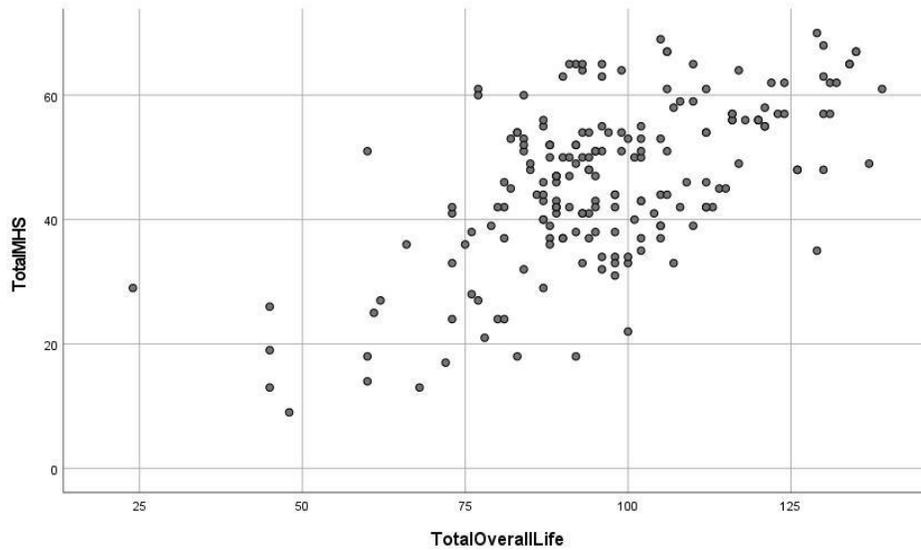
To determine the extent to which PsyCap HERO dimensions within the Overall-life and the college-work categories predict the variability in mental health among college students.

*Table 5.*

*Linear Regression of Overall Life Psychological Capital on Mental health*

	R	R Square	Adjusted R Square	Standardized Coefficients Beta	F	Sig.
Overall Life	.605	.366	.363	.605	111.532	.001 <sub>b</sub>

- a. Predictors: (Constant), Overall Life  
 b. Dependent Variable: Mental Health



The correlation value ( $R = .605$ ) shows a significantly moderate positive relationship between Overall Life Psychological Capital and Mental health. The F value ( $p = .001$ ,  $p < .01$ ) suggests a significant influence of Overall Life Psychological Capital on Mental health of college students. For each unit increase of Overall Life Psychological Capital 0.605 ( $\beta$ ) increase in Mental health. There is 36.6% ( $R^2$ ) of accounted variance for Overall Life Psychological Capital on Mental health.

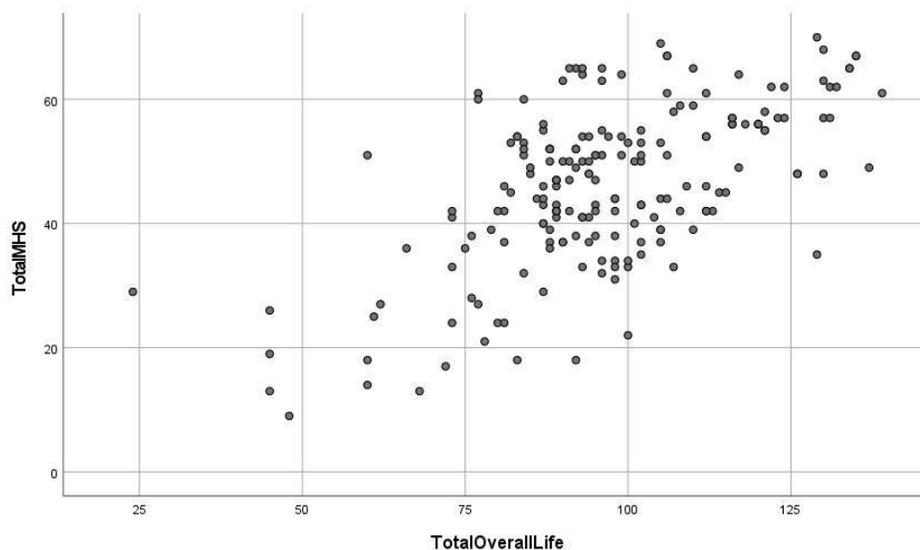
Table 6.

*Linear Regression of College Life Psychological Capital on Mental health*

	R	R Square	Adjusted R Square	Standardized Coefficients Beta	F	Sig.
College Life	.653 <sup>a</sup>	.426	.423	.653	143.234	.000 <sup>b</sup>

a. Predictors: (Constant), College Life

b. Dependent Variable: Mental Health



The correlation value ( $R = 0.653$ ) shows a significantly moderate positive relationship between College Life Psychological Capital and Mental health. The F value ( $p = .001$ ,  $p < .01$ ) suggests a significant influence of College Life Psychological Capital on Mental health. For each unit increase of College Life Psychological Capital, 0.653 ( $\beta$ ) increases in Mental health. There is 42.6% ( $R^2$ ) of accounted variance for College Life Psychological Capital on Mental health. Therefore, we accept the hypothesis that PsyCap can be used to predict the variability in Mental Health within the college population.

## **DISCUSSION**

In the past few years, mental health issues such as depression, anxiety, suicidal ideation, and self-injury among college students are growing at an alarming rate. Since the last decade of twentieth century, mental illness has been considered a public health issue alongside other physical illness. Recent studies indicate that there is an upward trend in the number of students with severe psychological problems which makes the mental health profile of college students a growing concern. Attending college has a dual impact on an individual. It marks a significant and stressful developmental period of an individual's life. On the other hand, it is at this phase of

life where various psychological disorders have their first onset. It is imperative that mental health professionals at the college level consider the problems associated with identity formation and the gradual manifestations of mental illnesses during these high-risk early adult years because the prevalence of developmental and mental health issues in college student communities is becoming more diverse and complex than ever in the history of higher education.

Also, the concept of Psychological Capital which is not much familiar among all, is strongly linked to increased well-being, work and life satisfaction. Therefore, it's necessary that the upcoming future, ie; the college students' mental health and PsyCap are assessed to help students persevere in their studies in a psychologically healthier manner. As PsyCap and its dimensions are malleable, they can be developed through training interventions along with various positive psychology exercises that focuses on enhancing well-being in conjunction with PsyCap. Exploring further in this regard revealed a variety of potent, low- cost approaches that impact mental health through development of positive strengths such as hope, efficacy, resilience, and optimism.

The primary purpose of this cross-sectional, exploratory study was to examine the relationships and predictions that exist between current levels of mental health and psychological capital [PsyCap] among college students as measured by the Keyes' Mental Health Continuum Short Form [MHC-SF] and Academic Psychological Capital Questionnaire [A-PCQ]. Mental health was measured by MHC-SF and participants were categorized into three groups languishing, moderately mentally healthy, and flourishing. Similarly, PsyCap was measured using A-PCQ and participants were measured for their current level of Hope, Efficacy, Resilience, and Optimism [HERO]. The main aim of the study was to establish the associations between the studied variables and to test psychological capital as a predictor of mental health.

The first objective was to describe the level of psychological capital and mental health of college students using their PsyCap and Mental Health scores. This objective focused on describing the current level of PsyCap and mental health of the studied sample using basic descriptive statistics. The baseline characteristics of the study population were described using measures of central tendency.

The second objective was to explore the relationship between psychological capital and mental health of college students in the studied sample. The relationship between PsyCap scores

and mental health scores were analyzed using correlation coefficients. From the analysis, it has been found out that there exists a significant relation between the two variables meaning that Mental Health definitely has an effect on the Psychological Capital of the individual. A person categorized as Flourishing or Moderately Mentally Healthy seems to have a better PsyCap score when compared with the scores of a Languishing person. This implies that a mentally well person will be hopeful, self-efficient, resilient and optimistic.

The third objective was to determine if differences exist in PsyCap of college students whose scores place them in one of three mental health continuum categories languishing, moderately mentally healthy, and flourishing. The results of the analysis show that there is a significant difference between the PsyCap of college students whose scores place them in languishing, moderate and flourishing. Students coming in the Languishing category of Mental Health Continuum has a lower PsyCap score whereas the students in the Moderately Mentally Healthy category have a better score and lastly the ones in the Flourishing category of Mental Health Continuum have the best PsyCap score. Hence, this proves that there is a significant difference between the PsyCap scores of students coming under the three categories of Mental Health Continuum. It also implies that there a significant positive correlation between the two variables of the study as proved by the second objective.

The fourth and final objective of this study was to determine the extent to which PsyCap HERO dimensions within the Overall-life and the college-work categories predict the variability in mental health among college students. From the analysis, it is clear that PsyCap and Mental Health has a moderately positive correlation. It also proves that the dimensions of the Overall-life and College-life categories can predict the variability in Mental Health of College students. The two hypotheses of this Research; were therefore accepted.

Taken together, the results of this study provide an essential contribution to theory on mental health continuum and psychological capital. In this study, majority of the respondents were Moderately Mentally Healthy. There is no conclusion that claims that PsyCap is a sufficient condition for mental health, nor would the researcher choose to evaluate individual's lives solely on the basis of whether they are flourishing; researcher values additional characteristics. However, this study attempts to bridge a gap by establishing additional empirical evidence to connect individuals' positive psychological strengths and mental health—in order to promote the

best among the college students. Positive mental health and well-being have gained greater momentum as a serious alternative to the numerous mental illness prevention strategies in past few decades. Focusing on building positive resources in college students may be helpful in propelling students from a languishing mental health state towards a sustainable state of positive and complete mental health. The results highlight the value of further investigation into positive psychological strengths as pivotal components of college students' positive mental health, and have significant implications for the development and content of strategies to promote mental health in college students.

## **CHAPTER-5: CONCLUSION**

### **5.1 FINDINGS**

- The two variables of this study; Mental Health and Psychological Capital have a significant relationship between them. Mental health tends to influence the Psychological Capital of the individual and vice versa.
- It has been determined that there exist differences in Psychological Capital of college students whose scores place them in one of the three mental health continuum categories: languishing, moderately mentally healthy and flourishing.
- Another finding was Psychological Capital HERO dimensions within the Overall-life and College-life categories predict the variability in Mental Health among college students.

### **5.2 LIMITATIONS**

- The study employs self-reported, cross-sectional, survey data. This study is a quantitative study; where examination of the prevalence, relationships, and predictions occur rather than descriptive, in-depth, and deeper understanding of phenomenon that comes with qualitative data. Additionally, as it is non-experimental research, the results would

provide only the existence or non-existence of relationship between variables and not the causal factor.

- A more robust comparison between the mental health continuum categories could have been achieved if all the three cohorts had comparable numerical strengths.

### **5.3 SUGGESTIONS FOR FURTHER RESEARCH**

- A much larger sample can be used for further research so that a much comparable numbers across all categories can be obtained.
- The sample of this present study was undergraduate students alone, all students doing their college; be it Post Graduate or Doctoral programs can be included in the further research because students are all the same and they are more likely to be vulnerable towards severe Mental Health issues including stress, burnout, etc. due to their workloads or any other known factors.

### **REFERENCES**

- Adams, T. B., Bezner, J. R., Drabbs, M. E., Zambarano, R. J., & Steinhardt, M. A. (2000). Conceptualization and measurement of the spiritual and psychological dimensions of wellness in a college population. *Journal of American college health, 48(4)*, 165-173.
- Alzyoud, M., & Mert, I. (2019). Does employees' psychological capital buffer the negative effects of incivility? *EuroMed Journal of Business*, ahead-of-print.  
<https://doi.org/10.1108/EMJB-03-2018-0021>
- Avey, J. B., Luthans, F., & Jensen, S. M. (2009). Psychological capital: A positive resource for combating employee stress and turnover. *Human Resource Management, 48(5)*, 677-693.

- Avey, J. B., Luthans, F., Smith, R. M., & Palmer, N. F. (2010). Impact of positive psychological capital on employee well-being over time. *Journal of occupational health psychology, 15*(1), 17.
- Avey, J. B., Reichard, R. J., Luthans, F., & Mhatre, K. H. (2011). Meta-analysis of the impact of positive psychological capital on employee attitudes, behaviors, and performance. *Human resource development quarterly, 22*(2), 127-152.
- Avey, J. B., Wernsing, T. S., & Mhatre, K. H. (2011). A longitudinal analysis of positive psychological constructs and emotions on stress, anxiety, and well-being. *Journal of Leadership & Organizational Studies, 15*48051810397368.
- Alzyoud, M., & Mert, I. (2019). Does employees' psychological capital buffer the negative effects of incivility? *EuroMed Journal of Business*, ahead-of-print.  
<https://doi.org/10.1108/EMJB-03-2018-0021>
- Braithwaite, S. R., Delevi, R., & Fincham, F. D. (2010). Romantic relationships and the physical and mental health of college students. *Personal Relationships, 17*(1), 1–12. <https://doi.org/10.1111/j.1475-6811.2010.01248.x>
- Carifio, J., & Rhodes, L. (2002). Construct validities and the empirical relationships between optimism, hope, self-efficacy, and locus of control. *Work: A Journal of Prevention, Assessment and Rehabilitation, 19*(2), 125-136.
- Chen, T., & Lucock, M. (2022). The mental health of university students during the COVID-19 pandemic: An online survey in the UK. *PLOS ONE, 17*(1), e0262562.  
<https://doi.org/10.1371/journal.pone.0262562>

DIGITAL LEARNING, STUDENTS' MENTAL HEALTH AND ROLE OF  
PSYCHOLOGICAL CAPITAL AMIDST COVID-19 PANDEMIC IN INDIA |

Ensemble. (n.d.). Retrieved May 26, 2022, from

<https://www.ensembledrms.in/articles/ensemble-2021-sp1-a013/>

Estiri, M., Nargesian, A., Dastpish, F., & Sharifi, S. M. (2016). The impact of psychological capital on mental health among Iranian nurses: Considering the mediating role of job burnout. *SpringerPlus*, 5(1), 1377.

<https://doi.org/10.1186/s40064-016-3099-z>

Gujar, N. M., & Ali, A. (2019). Effects of psychological capital and self-esteem on emotional and behavioral problems among adolescents. *Journal of Mental Health and Human Behaviour*, 24(2), 85. [https://doi.org/10.4103/jmhbb.jmhbb\\_59\\_1](https://doi.org/10.4103/jmhbb.jmhbb_59_1)

Gupta, P., De, N., Hati, S., Saikia, C., & Karmakar, R. (2019). The Relationship between Positive Psychological Capital and Coping Styles: A Study on Young Adults. *Psychology*, 10(12), 1649–1662.

<https://doi.org/10.4236/psych.2019.1012109>

Hashmi, S. D., Shahzad, K., & Abbas, F. (2022). The interactive effects of sexual harassment and psychological capital on victims' burnout: Evidence from the post -#MeToo movement era. *Gender in Management: An International Journal*, 37(4), 509–523. <https://doi.org/10.1108/GM-04-2020-0136>

Haleem, M., Masood, S., Aziz, M., & Jami, H. (2017). Psychological capital and mental health of rescue workers. *Pakistan Journal of Psychological Research*, 32, 429–447.

Hernández-Varas, E., Labrador Encinas, F. J., & Méndez Suárez, M. (2019).

Psychological capital, work satisfaction and health self-perception as predictors of psychological wellbeing in military personnel. *Psicothema*, 31(3), 277–283.

<https://doi.org/10.7334/psicothema2019.22>

Kawalya, C., Munene, J., Ntayi, J., Kagaari, J., Mafabi, S., & Kasekende, F. (2019).

Psychological capital and happiness at the workplace: The mediating role of flow experience. *Cogent Business & Management*, 6.

<https://doi.org/10.1080/23311975.2019.1685060>

Keyes, C. L. & Haidt, J (eds.) (2003). *Flourishing: Positive psychology and the life well lived*.

Keyes, C. L. (1998). Social well-being. *Social Psychology Quarterly*, 61, 121–140.

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of health and social behavior*, 207-222.

Keyes, C. L. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of consulting and clinical psychology*, 73(3), 539.

Keyes, C. L. (2006). The subjective well-being of America's youth: Toward a comprehensive assessment. *Adolescent & Family Health*.

Keyes, C. L. (2007). Promoting and protecting mental health as flourishing: a

complementary strategy for improving national mental health. *American Psychologist*, 62(2), 95.

Keyes, C. L. (2013). *Mental well-being: International contributions to the study of positive mental health*. Springer. 166

Keyes, C. L. (2014). Mental Health as a Complete State: How the Salutogenic Perspective Completes the Picture. In *Bridging Occupational, Organizational and Public Health* (pp. 179-192). Springer Netherlands.

Keyes, C. L., & Lopez, S. J. (2002). Toward a science of mental health: Positive directions in diagnosis and interventions. In C. R. Snyder & S. J. Lopez (Eds), *Handbook of positive psychology* (pp. 45–59). New York: Oxford University Press.

Keyes, C. L., & Lopez, S. J. (2002). Towards a science of mental health. In CR Snyder & JL Lopez (Eds) *Handbook of Positive Psychology*. New York: Oxford University Press.

Keyes, C. L., & Simoes, E. J. (2012). To flourish or not: Positive mental health and all cause mortality. *American journal of public health*, 102(11), 2164-2172.

Keyes, C. L., Dhingra, S. S., & Simoes, E. J. (2010). Change in level of positive mental health as a predictor of future risk of mental illness. *American Journal of Public Health*, 100(12), 2366.

Keyes, C. L., Eisenberg, D., Perry, G. S., Dube, S. R., Kroenke, K., & Dhingra, S. S. (2012). The relationship of level of positive mental health with current mental disorders in predicting suicidal behavior and academic impairment in college

students. *Journal of American College Health*, 60(2), 126-133. 167

Keyes, C. L., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: the empirical encounter of two traditions. *Journal of personality and social psychology*, 82(6), 1007.

Krasikova, D. V., Lester, P. B., & Harms, P. D. (2015). Effects of Psychological Capital on Mental Health and Substance Abuse. *Journal of Leadership & Organizational Studies*, 22(3), 280–291. <https://doi.org/10.1177/1548051815585853>

Luthans, B. C., Luthans, K. W., & Jensen, S. M. (2012). The impact of business school students' psychological capital on academic performance. *Journal of Education for Business*, 87(5), 253-259.

Luthans, F. (2002a). The need for and meaning of positive organizational behavior. *Journal of organizational behavior*, 23(6), 695-706.

Luthans, F. (2002b). Positive organizational behavior: Developing and managing psychological strengths. *The Academy of Management Executive*, 16(1), 57-72.

Luthans, F., & Jensen, S. M. (2002). Hope: A new positive strength for human resource development. *Human Resource Development Review*, 1(3), 304-322.

Munawar, I., Jahan, M., & Khan, A. (2021). Exploring Psychological Capital as Predictor of Mental Health among Graduate-Level Students. *International Journal of Indian Psychology*, 9(2). <https://doi.org/10.25215/0902.072>

Norris, R., Carroll, D., & Cochrane, R. (1992). The effects of physical activity and

exercise training on psychological stress and well-being in an adolescent population. *Journal of Psychosomatic Research*, 36(1), 55–65.

[https://doi.org/10.1016/0022-3999\(92\)90114-h](https://doi.org/10.1016/0022-3999(92)90114-h)

Prasath, P. R., Mather, P. C., Bhat, C. S., & James, J. K. (2021). University Student Well-Being during COVID-19: The Role of Psychological Capital and Coping Strategies. *Professional Counselor*, 11(1), 46–60.

Rafaq, F., & Latif, S. (n.d.). PSYCHOLOGICAL CAPITAL AMONG PROFESSIONALS IN PAKISTAN. 1.

Robinson, E., Sutin, A., Daly, M., & Jones, A. (2021). A systematic review and meta-analysis of longitudinal cohort studies comparing mental health before versus during the COVID-19 pandemic in 2020. *Journal of Affective Disorders*, 296.

<https://doi.org/10.1016/j.jad.2021.09.098>

Santisi, G., Lodi, E., Magnano, P., Zarbo, R., & Zammitti, A. (2020). Relationship between Psychological Capital and Quality of Life: The Role of Courage. *Sustainability*, 12(13), 5238. <https://doi.org/10.3390/su12135238>

Sarshar, Z., Eskafi-Noghani, M., & Mohammadpour, H. (2019). Relationship between Psychological Capital and Perfectionism among Female High School Students in the City of Gonabad. *Journal of Research and Health*, 9(7), 633–638. <https://doi.org/10.32598/JRH.1368.2>

Son, C., Hegde, S., Smith, A., Wang, X., & Sasangohar, F. (2020). Effects of

COVID-19 on College Students' Mental Health in the United States: Interview Survey Study. *Journal of Medical Internet Research*, 22(9), e21279.

<https://doi.org/10.2196/21279>

Truong, K. D., & Ma, S. (2006). A systematic review of relations between neighborhoods and mental health. *The Journal of Mental Health Policy and Economics*, 9(3), 137–154.

Tian, F., Shu, Q., Cui, Q., Wang, L., Liu, C., & Wu, H. (2020). The Mediating Role of Psychological Capital in the Relationship Between Occupational Stress and Fatigue: A Cross-Sectional Study Among 1,104 Chinese Physicians. *Frontiers in Public Health*, 8, 12. <https://doi.org/10.3389/fpubh.2020.00012>

World Health Organization (2004). *Promoting Mental Health: Concepts, Emerging Evidence, and Practice*. Geneva, Switzerland: World Health Organization.

Younas, S., Tahir, F., Sabih, F., Hussain, R., Hassan, A., Sohail, M., Hafsa, S., Munawar, A., Kanwal, R., Tanvir, M., & Zahra, M. (2020). PSYCHOLOGICAL CAPITAL AND MENTAL HEALTH: EMPIRICAL EXPLORATION IN PERSPECTIVE OF GENDER. *PONTE International Scientific Researchs Journal*, 76. <https://doi.org/10.21506/j.ponte.2020.1.11>

**APPENDICES**

**APPENDIX I**

**Demographic Details-**

**1. Name in Initials -**

**2. Age -**

**3. Date of Birth -**

**4. Gender -**

**5. Religion -**

**6. Place -**

**7. Annual Income of Family (approx) -**

**8. How do you identify your sexual orientation?**

**Bisexual**

**Lesbian or Gay**

**Straight or Heterosexual**

**Other \_\_\_\_\_(Specify)**

**9. What is your current relationship status?**

**In a relation**

**Single**

**10. What is your enrolled degree?**

**11. What was your grade point average (GPA) last semester?**

\_\_\_\_\_(Provide exact point score if possible)

**12. To what extent do you consider yourself to be a religious person?**

**Not at all   Slightly   Moderately   Very Extremely**

**13. To what extent do you consider yourself to be part of a religious community?**

**Not at all   A little   Moderately   Mostly   Completely**

**If so, which religious community are you a part of?**

**14. To what extent do you have spiritual beliefs?**

Not at all   Slightly   Moderately   Very Extremely

**15. Which of the following are the most significant social support system (s) you had/have (please check the appropriate)**

- Immediate family (Spouse, Parents or Siblings)
- Relatives
- Friends
- Teachers/Professors
- Mentors/Coaches/Consultants
- Mental Health Providers (Coursellors, Psychologists, Social Workers etc.)
- Human Services Agencies
- Clergy or members of religious/spiritual organisations(church, mosque, temple, synagogue, cathedral etc.)
- Student Organisations
- Community based organisations

Other (Specify) \_\_\_\_\_

## APPENDIX II

### Mental Health Continuum-Short Form (MHC-SF; Keyes, 2009)

**General Instructions-** Please answer the following questions are about how you have been feeling during the past month. Place a check mark in the box that best represents how often you have experienced or felt the following:

During the past month, how often did you feel ...	EVER	ONCE OR TWICE	ABOUT ONCE A WEEK	ABOUT 2 OR 3 TIMES A WEEK	ALMOST EVERY DAY	EVERY DAY
1. Happy						

2. Interested in life						
3. Satisfied with life						
4. That you had something important to contribute to society						
5. That you belonged to a community (like a social group, or your neighbourhood)						
6. That our society is a good place, or is becoming a better place, for all people						
7. That people are basically good						
8. That the way our society works makes sense to you						
9. that you liked most parts of your personality						
10. good at managing the responsibilities of your daily life						

11. that you had warm and trusting relationships with others						
12. that you had experiences that challenged you to grow and become a better person						
13. confident to think or express your own ideas and opinions						
14. that your life has a sense of direction or meaning to it						

**APPENDIX III**

**Academic Psychological Capital Questionnaire; Academic PsyCap (A-PCQ; Luthans, Luthans, & Jensen, 2012)**

General Instructions- Below are a series of statements that describe how you may think about yourself RIGHT NOW. We are asking you to consider each question relative to your overall life and work aspects. Use the scale below to indicate your level of agreement or disagreement with each statement.

<b>Strongly Disagree</b>	<b>I disagree</b>	<b>Somewhat disagree</b>	<b>Somewhat agree</b>	<b>Agree</b>	<b>Strongly Agree</b>
1	2	3	4	5	6

		<b>Over all-Life</b>	<b>Colleg e-Life</b>
--	--	--------------------------	--------------------------

	I feel confident analysing a long-term problem to find a solution concerning my ...		
	I feel confident in representing my ideas concerning my ...		
	I feel confident contributing to discussions about strategies on my ...		
	I feel confident setting targets/goals on my...		
	I feel confident contacting people to discuss problems concerning my ...		
	I feel confident sharing information with a group of students about my...		
	If I should find myself in a jam about my ..., I could think of many ways to get out of the jam.		
	At the present time, I am energetically pursuing my ... goals.		
	There are lots of ways around any problem concerning my...		
0	Right now, I see myself as being pretty successful concerning my...		
1	I can think of many ways to reach my current goals regarding ...		
2	At this time, I am meeting the goals that I have set for myself concerning...		
3	When I have a setback with ..., I have trouble recovering from it, moving on.		
4	I usually manage difficulties one way or another concerning my ...		
5	I can be "on my own" so to speak, if I have to regarding my...		
6	I usually take stressful things in stride with regard to my...		

7	I can get through difficult times at school because I've experienced difficulty before concerning  my...		
8	I feel I can handle many things at a time with my...		
9	When things are uncertain for me with regards to ..., I usually expect the best.		
0	If something can go wrong for me with my ..., it will.		
1	I always look on the bright side of things regarding my...		
2	I'm optimistic about what will happen to me in the future as it pertains to my...		
3	With regards to my ..., things never work out the way I want them to.		
4	I approach my ... as if "every cloud has a silver lining."		

**PROJECT REPORT**  
**A STUDY ON THE IMPACT OF CELEBRITY ENDORSEMENT**  
**ON CONSUMERS' BUYING BEHAVIOR**

Submitted by:

EMMA BARBERA

Register No:

SB19PSY014

Under the guidance of

MS. JISHA SEKHAR

In partial fulfillment 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 IMPACT OF CELEBRITY ENDORSEMENT ON CONSUMERS' BUYING BEHAVIOR”, is a bonafide record submitted by MS. EMMA BARBERA, Reg.no. SB19PSY014, in partial fulfillment 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



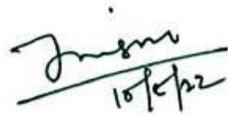
Ms. Jisha Sekhar

Assistant Professor

Department of Psychology

St. Teresa's College, Ernakulam

External Examiner:



## **DECLARATION**

I, Emma Barbera, hereby declare that the study presented in the dissertation entitled, “The Impact of Celebrity Endorsement on Consumers' Buying Behavior ”, 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. Jisha Sekhar, 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

Emma Barbera

Date: 07/05/2022

## **ACKNOWLEDGEMENT**

It is not possible to prepare any project without the guidance and support of other people. This one is certainly no exception. I would like to express my gratitude to the Department of Psychology, St. Teresa's College, Ernakulam for providing me with the opportunity to undertake this research and for their kind help and encouragement which helped me in completion of this project.

I am highly indebted and acknowledge my utmost sense of gratitude to my research guide, Ms. Jisha Sekhar, Assistant Professor, Psychology, for encouraging and guiding me throughout all the phases of my research as well as for the constant supervision and in providing the necessary information regarding the project.

I extend my sincere thanks to my parents and all who have supported me throughout this time. My thanks and appreciations also go to my friends in developing the project and people who have willingly helped me out with their abilities.

Above all, I thank God Almighty for blessing me in all stages of this project and for its successful completion.

Thanking You

Emma Barbera

## TABLE OF CONTENTS

<b>SL.NO.</b>	<b>TOPICS</b>	<b>PAGE NO.</b>
<b>1</b>	<b>INTRODUCTION</b>	
1.1	Background of study	2
1.2	Problem Statement	3
1.3	Need and Significance of the study	3
1.4	Scope of the study	3
1.5	Objectives of the study	4
1.6	Limitations of the study	4
<b>2</b>	<b>REVIEW OF LITERATURE</b>	
2.2	Literature Review	6
<b>3</b>	<b>THEORETICAL FRAMEWORK</b>	<b>10</b>
<b>4</b>	<b>RESEARCH AND METHODOLOGY</b>	
4.1	Objectives	22
4.2	Hypothesis	22

4.3	Research Design	22
4.4	Sources of Data	23
4.5	Sample Design	23
4.6	Sample Size	23
4.7	Sampling Method	23
4.8	Method of Data Collection	23
4.9	Drafting a Questionnaire	24
4.10	Data Analysis Techniques	24
<b>5</b>	<b>DATA ANALYSIS</b>	<b>27</b>
<b>6</b>	<b>FINDINGS</b>	<b>33</b>
<b>7</b>	<b>RECOMMENDATIONS</b>	<b>36</b>
<b>8</b>	<b>CONCLUSIONS</b>	<b>38</b>
	<b>REFERENCES</b>	<b>39</b>
	<b>APPENDIX</b>	<b>40</b>

**CHAPTER 1**  
**INTRODUCTION**

## INTRODUCTION

### 1.1 BACKGROUND

Watching popular personalities convert themselves to excellent salesmen is something that we witness quite often. Celebrities vouching for or promoting brands and their products is one of the quickest ways for brands to develop an association and a bond in the mind of its consumers. Hence celebrity endorsements are naturally recognized as a viable option for brands to increase their popularity, awareness and credibility. In a world where celebrities are treated as role models, people are changing their life to replicate that of their favourite celebrity. This influence endorsed by the celebrities creates a great impact on the buying behavior of common people. Which attracts the customers and ultimately increases the company productivity.

Celebrities are not always creating any kind of effect on an individual's mind in terms of buying. But mostly it gives a great impact on the perception of choosing any product. We are always thinking that if our favourite celebrity is using them, then we should use that to be like them. These celebrities need not be a superstar, but someone who the target audience can relate to. It has been recognized as a single psychological effect that purchasing a product that has been promoted by a celebrity an individual admires, will allow that particular individual to emulate the celebrity's desired characteristics or attract similar people into their lives.

From the past 150 years advertising is changing in different phases from taking classical forms to modern. Now a days it's the best strategy used by marketers to influence customers by showing celebrities with their products, which includes different appeals like, exciting, absurdity, sexual etc. Belch, G. and Belch, M. acknowledged that the main aim of formulating such specific strategies is to get high brand revelation, longing, concentration and curiosity. For this, marketers attach famous personalities with their products.

McCracken (1989) stated that these famous personalities' had great influence on the consumer's buying behavior that's why it comes off as the most attractive tool of advertising today. The major aim to do advertising and adopt this strategy is to influence customers towards the products. Marketing values have transformed throughout the years. Today marketers make use of celebrities to endorse their products so that they can achieve an edge over their competitors. But this requires great understanding of the concept in order for the celebrity

endorsement to be successful. Hence several factors have to be weighed in order for the celebrity endorsement to be successful.

## **1.2 PROBLEM STATEMENT**

Does celebrity endorsement encourage consumers to take some kind of action towards buying the endorsed product(s)? The study aims to find whether the use of celebrities while endorsing cosmetics products have any impact on the purchase intention among young adult females in Ernakulam district.

## **1.3 NEED AND SIGNIFICANCE OF THE STUDY**

This research will help to achieve major insights into the impact celebrity endorsements have on consumers' buying behavior. It will assist the marketers to get a better understanding on the different attributes that are important and matters the most to the consumers. Celebrity endorsement influences the behaviours of consumers to buy a specific product through cognition. These perceptions are perceived by consumers through their senses, discernment, attention, recall, reasoning, language, et cetera. Hence this research would provide major help to increase productivity in various businesses.

## **1.4 SCOPE OF THE STUDY**

This is a study done to check whether celebrity endorsements affect buying behavior of the people and to check whether there is a shift in their behavior because of celebrity endorsements. Major emphasis of this research is to determine how marketers select particular celebrities to influence consumers' buying patterns. It also emphasizes how the marketers create a desired image for themselves by selecting strong publicly desirable celebrities. This research also aims at identifying the key factors that may influence consumers buying behavior through celebrity endorsement.

## **1.5 OBJECTIVES OF THE STUDY**

The main aim of studying the impact and influence of celebrity endorsement on consumers buying intention is to find the answers to the following:

- To assess the impact celebrities have on the consumers' buying intention with regards to the credibility of the chosen celebrity.
- To assess the impact celebrities have on the consumers' buying intention with regards to the attractiveness of the chosen celebrity.
- To assess the impact that celebrity endorsement have on its consumers in terms of product fit match between the product/brand and the celebrity.
- To see how celebrities transfer meaning to the products they endorse and to what extent does it convince and shape consumers' buying intention.

## **1.6 LIMITATIONS OF THE STUDY**

The limitations of this study are as follows:

- As the research will be conducted within Ernakulam district, the perspective of consumers outside Ernakulam district will be left unnoticed.
- Since the respondents will be of young female individuals (between the ages of 18 and 30), individuals above the age of 30 would not be taken into consideration whose opinion regarding celebrity endorsement would likely be very different from the selected sample.
- Respondents belonging to the urban sector, who have an understanding of this issue, are selected while individuals of rural areas will be ignored. Hence, the opinions of people residing in rural areas are not taken into consideration in this study.

**CHAPTER 2**  
**REVIEW OF LITERATURE**

## REVIEW OF LITERATURE

### 2.1 LITERATURE REVIEW

Advertisers regularly practice the strategies intended to attract customer's interest to their message and to differentiate their offerings from rival products with the anticipation of influencing buying behavior of the customer. In today's competitive world, a quality is positioned on an approach, which can accomplish these objectives. One challenge at such a plan includes the use of a celebrity representative. According to Atkin and Block (1983), there are numerous bases as to why a famous endorser may be dominant. First, such a representative attracts consideration toward the commercials into the messy flow of communication. In addition, celebrities are conventionally observed as being greatly active individuals with eye-catching and likable traits.

The use of famous persons in promotions is traced back to the nineteenth century and these general promotional practices have revealed a large quantity of intellectual as well as realistic considerations. Mainly academic analysis of celebrity support encompasses the sphere of spokesperson credibility and charismatic representatives, and recommends that famous persons exercise their impact on customers through apparent traits. A number of research studies reveal that use of eye-catching celebrities serve as a foundation to enhance feelings towards the ads. This mind-set to the commercials is identified as psychological circumstances that are exercised by persons to systematize the manner, how to recognize the surroundings as well as organize the manner a person reacts towards it.

Today the mass media are flooded through descriptions along with information concerning superstars, and because of this, celebrities have high reputations, distinctive traits, and fascinating descriptions according to the community's opinion. Celebrities regularly emerge in promotions in connection among customer goods or services. By means of skill to pierce the hectic mess of publicity, portray customer consideration, produce high memory rates, generate as well as distinguish brand descriptions thus, create trade and income, superstar endorsement have demonstrated to be a helpful approach. No doubt dealers spend huge amounts of capital in utilizing superstars to sponsor their brands.

McCracken (1998) proposes that a superstar is considered like a memorial, entertainer or representative of the business organization. Investigation has established that spokesperson endorsement influences consumers' mind-set in common and it may change the feelings of customers towards the commercial and products as well. This may perhaps effect the improvement of the acquisition plan and as an outcome in the increase of trade. Researchers have intended for extensive concentration to consumer's mind-set to the commercial as a sentimental creation and intervening influence on brand attitudes and acquiring intentions.

This study's main purpose was to study the impact of the celebrity endorsement on effective brand management and evaluate associated factors that contribute to the success or failure of the endorsement. Celebrity endorsement effects are moderated by brand symbolism, such that brands that communicate something about the user yield stronger effects than brands that do not. So, it explains that not just the celebrity but existing brand value plays a very crucial and important role in changing the perception of the target market. This study portrays the light on different purchase patterns of a commodity or a service, when a celebrity is associated with them. The methodology adopted by the researchers was collection of data from libraries and first hand data to conclude the study. The result was that the people would purchase more of the same goods or services due to celebrity endorsements as the consumer tends to neglect the negative effects of the use of that product and also becomes more loyal to the brand.

The researchers of the study focus on the decision making process of consumers after and before the association of a celebrity with a product. The study also showcases the adverse effects of celebrity endorsements. The endorsement is only successful when the right celebrity is associated with the right product. Hence, celebrity endorsement can't be treated as a 'mantra' for success.

## **2.2 RESEARCH GAP**

This research about the "impact of celebrity endorsement on consumers buying behavior" will be useful in many ways :

- There is a lack of research in the field of cosmetic segments in terms of celebrity endorsement. So, the results of this research will highlight how effective the technique of celebrity endorsement is in the cosmetics industry.
- The research about celebrities appearing in cosmetics advertisements in Ernakulam district is rarely done. Hence, the conclusions that would be derived with this research will help in understanding the buying behavior of the females in Ernakulam district.
- This research will help to understand the reaction of women towards celebrity endorsed cosmetic products and will help to judge how much trust women place in celebrities endorsing cosmetics products.

**CHAPTER 3**  
**THEORETICAL FRAMEWORK**

## **THEORETICAL FRAMEWORK**

### **Celebrities**

Erdogan (1999) concludes that celebrities are those people who are well known by the large number of people. They have special uniqueness and features like magnetism, unusual standard of living or special skills that are not commonly experiential in common people. That it can be said that in society they are different from the common people. Among the model forms of celebrities, actors (e.g. Saif Ali Khan, Shah, Salman Khan, Amitabh Bachan etc), models (e.g. Parineeti Chopra, Ali Zafar, Bipasha Basu, Kareena Kapoor etc), Sports-men (e.g. Yousuf Pathan, Shahid Afridi, Sachin Tendulkar, etc.) are significant.

### **Celebrity Endorsement**

Khatri (2006) studied that celebrity endorsement is the Promotion strategy to attract the customers. By analyzing the current market, now it's become the need of the marketers to use the different famous personalities to relate with their brands to create unique identity of the brand and to do famous his company's brand or product, which results high expenditure for the company to use that strategy, however nowadays it is used to be a powerful strategic tool to get maximum profit. It also shows that this can carry risk, because there is no sureness that the celebrity can come up with the sales generation of the firm. But it creates a buzz and gives popularity to the company and the brand. This can increase the expectation of the customers in terms of real stars by delivering the company promise. There are certain perspectives where the real person can work better than the celebrities' endorsement, but not always.

### **Factors Affecting Celebrity Endorsement and Relevant Models**

There are different factors of celebrities influencing the consumer buying behavior when such celebrities endorse any brand. Marketers have to consider some factors while selecting celebrities to endorse brands to get the desired result to use this strategy of celebrity endorsement in terms of brand image, consumer buying behavior and attainment of desired market share. Past researchers have focused on a different mix of factors related to celebrity endorsement

influencing consumer buying behavior, this research is focused on a relatively different and most appropriate mix of factors that a marketer should consider while selecting celebrities to endorse any brand to get a desired effect on consumer buying behavior.

### **a) Credibility**

It refers to the level of believability a beneficiary has towards a message given by the source and expertise of the source or deliverer. In the past source credibility was just considered as an endorser credibility in a commercial. It is proved as a significant basis affecting customer buying behavior and their attitude towards marketing communications. Past publications described three facets of it which are trustworthiness, expertise & attractiveness.

#### **❖ Source Credibility Model.**

Celebrity's integrity relies on his or her physical attractiveness, trustworthiness and expertise that's why researchers who want to study the effectiveness of celebrity endorsement have to consider this model (Ohanian, 1990).

### **i) Trustworthiness**

Trustworthiness is the degree of believability or trust listeners have for communicators (Hovland, 1953). In the case of celebrities it refers to the buyer's trust on celebrities' message for the brand (Ohanian, 1991). In other words if customers have a trust on celebrities being endorsed then the message of such celebrities will be more influential and will help a lot in changing the minds of customers towards the desired action that organizations want for their brand. (Miller and Baseheart, 1969).

#### **→ Impacts of Trustworthiness**

Previously it was researched that more the customers have a trust on communicator the more they get persuaded by the message of that communicator and more the chances of getting customers mind change for the purchase of that brand (Miller & Baseheart, 1969)

### **ii) Expertise**

Expertise is the measure of a communicator's aptitude and attitude which influences customers' decision towards purchase of a brand. Researchers believe that celebrities with

relative and high expertise prove to be more successful in persuading the customers than those who just have physical attractiveness not the combination of attractiveness and expertise (Till and Busler 1998).

### → Impacts of Expertise

Expertise is the ability of a communicator that proves as a helpful factor in taking the decision by a shopper. They generally get inspired by the sayings of learned and expert communicators. Proficient learning skills and expertise of endorsers has a significant reliability impact on customers' belief towards the communication of such endorsers (David H. Silvera, Benedikte Austad, 2004).

### b) Physical Attractiveness

Attractiveness has the power that can easily grab the attention of viewers but it doesn't just relate with the body but the entire physical traits comes under the measure of attractiveness which includes, hair color, facial features, height, weight, complexion, etc and normally public figures have to be attractive to have an influence on their audience and viewers (Temple, 2009).

### ❖ Source Attractiveness Model

Attractive celebrities are more commonly used by organizations for brands' promotions since the inception of this marketing strategy (Erdogan, 1999), once the customers get inspired with the attractive personality of celebrity then they get more involved in the message such celebrities give and then the chances of customers conviction towards the communication of celebrities become more fruitful (Baker & Churchill, 1977; Chaiken, 1979; Debevec & Keman, 1984) such fruitful results probability is normally low with less attractive celebrities.

### i) Similarity

The human brain recognizes celebrities similarly to how it recognizes people we actually know. The effect is that, if consumers happen to be fans, they place a higher value on products that celebrities are endorsing – it is as if they are receiving advice from a valued friend. Consumers feel more sympathetic towards a brand, if their products are promoted by a celebrity they admire or relate to. It's a simple psychological effect: Subconsciously people believe that purchasing a product that's promoted by a celebrity they admire, will allow them to emulate the

celebrity's desired traits or attract similar people into their lives. They will associate the celebrities' success, beauty, athletic skill etc. with a particular product. Shimp (2007) asserted that celebrity similarity involves the degree to which celebrity endorser of a brand matches the target viewers of commercials in terms of some features such as age, gender, ethnicity, et cetera. It is perceived that consumers tend to prefer and rely on messages from celebrity endorsers of a brand that shares these common characteristics. When celebrity endorsers and consumers share similar qualities, such as needs, goals, interest, lifestyle, et cetera the celebrity endorser is better valued in the eyes of the consumers.

## **ii) Familiarity**

Familiarity is a form of remembering in which a situation, event, place, person, or the like provokes a subjective feeling of recognition and is therefore believed to be in memory, although it is not specifically recalled. Familiarity has been shown to be positive and reassuring for most people, while perceived similarity results in the assumption that people have more in common, facilitating warmer, more comfortable interactions (Schneider et al., 2012), all of which facilitate attraction. People are more attracted to that which is familiar. Hence, familiarity with celebrity endorsers can play a role in consumer buying behavior.

## **iii) Likeability**

Likability pulls people toward you. When we think someone likes us, we tend to like them as well. Hence, celebrities who are liked by consumers are better consumers may buy products only best they care aft the person who doong. Likability virtually always helps boost their reputation among their peers. Likeability can increase conformity among people.

## **c) Emotional Involvement**

Emotional Involvement means the purchase decision of customers is directly proportional to the use and likeness by celebrity (endorser) for the brand that celebrity is endorsing (Cronley et al., 1999; Silvera and Austad, 2004). Adding to this, emotional attachment of viewers with their favorite celebrities becomes more fruitful in terms of their believability for the message given by that celebrity.

## **❖ Emotional Involvement Model**

Endorser likes and uses the item he supported. (Cronley et al., 1999; Silvera & Austad, 2004). When an endorser is accepted to like or utilize the item being supported, the purchaser states of mind towards the brand and promotion enhancement.

#### **d) Meaning Transfer**

(McCracken 1986), he says that in simple terminology we can say that endorsers carry their individual meaning to the product. Every celebrity has their own distinctive set of meaning and lifestyle. It clarifies the adequacy of big name spokespersons in terms of what purchasers connect with the endorser and in the end exchange to the brand.

#### **❖ Meaning Transfer Model**

Process of transferring a concept of a product through an endorser to a customer (McCracken, 1986). The appropriate fit between brand features and advocator's personality brings higher chances of likelihood of consumers observation and purchase intention. There are three stages to it. First stage covers the message transferred by celebrity to the brand, in the second stage a message from the brand gets transferred to the customer and in the third stage that message leaves the impact on the customer's psyche and ultimately on customers' decision to buy that brand.

#### **e) Product Match-Up**

Product match-up is construed as ensuring a similarity between the spokesperson's characteristics and the product attributes so as to enhance the advertisement effectiveness. Product specific associations include the associations that vary in their importance depending on the type of product category. The findings have significant implications for academics, brand managers and celebrity management companies.

#### **❖ The Product Match-Up Hypothesis**

The above model explains that there has to be an almost ideal match between the celebrity characteristics of the personality and the features of the brand. (Erdogan 1999) says that a successful matchup can be determined by the extent of the fitness between the celebrity and the brand endorsed by that celebrity. This same concept is further stated by Michael (1989) who believes the same. Another study done by Ohanian in 1991 reveals that simply such celebrities

should be endorsing who is compatible and the consumers perceive them to possess expertise too.

### **Consumer Buying behavior**

Buying behavior is a process by which a person searches for the product/services they need or want, make a decision to buy the required and most suitable one from different alternatives, use and then dispose of it. For making a marketing decision, the buying process model is playing a very important role for any one. It makes marketers think about each step of this process rather than just purchase decisions because if marketers just consider the purchase decision, it may be too late for a business to influence the choice of customers. According to this model the customer passes through all stages for purchasing every goods or services. However, in more regular purchases, customers often skip some stages (Kirmani & Shiv1998).

### **Stages in Consumer Choice Making**

Process starts by acknowledging a need or problem called the problem recognition stage. Then a person starts searching the information regarding the solution of a problem this is known as the information search stage. Then as a result of the information collected individuals become able to evaluate the alternatives they have to resolve their problem. Then the time comes to make a purchase decision for the most appropriate alternative they have evaluated among all they have in their choice. Then on using the selected alternative, buyers become able to assess the performance of the brand whether it fulfilled the desired expectations or not or whether to buy it again or not.

### **Factors affecting Buying Behavior**

Brewster, Sparrow and Vernon (2007) explain about factors that affect buying behavior and vary from person to person, age to age, and area to area. Every society follows its own norms, culture and values. At different stages of life our preferences change because of our age, needs,

lifestyle, earning and psychological factors. These factors can be Internal (memory and way of thinking) or External (media, word of mouth, publicity and feedback).

There are several factors, which pressure the buying behavior :

**a) Cultural influences**

It has the broadest and the deepest influence on buying behavior. Brewster, Sparrow and Vernon (2007) define culture as a shaping process, ' for a culture to exist, member of a group or society having different values and norms, which vary from time to time. Cultural values can change and have to be watched by marketers. Ignoring this deepest and widest factor can be very costly for a company in terms of image and profit.

**b) Social influences**

Social influences are those influences that clearly mold buying behavior, it affects through reference group, family members and social class (Ahmed & Saeed 2014).

**c) Family influences**

Family life cycle and family decision making has the most influence on one's buying behavior.

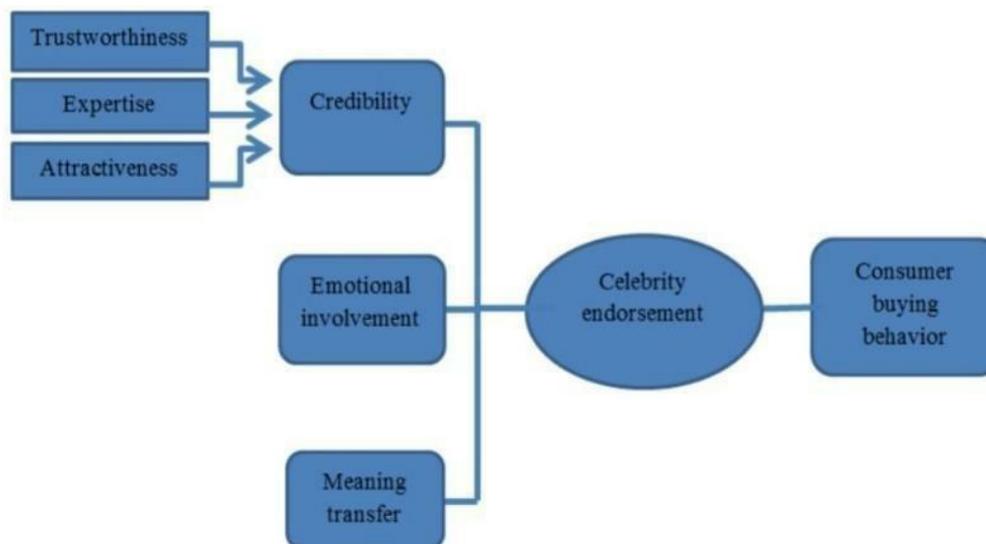
**d) Psychological influences**

These influences are related to our perception, learning, memory and motivation. It changes buying behavior through making the perceived picture of the product in the customer's mind. Customer buying behavior can be influenced by different factors like: perception, beliefs, society, personality, information choices, preferences and communication.

### **Impact of Endorsement on Buying Behavior**

Ranjbarian, Shekarchizade & Momeni (2010) agreed that advertisement is the action that persuades individuals of any particular market to buy services and products or services. Through different ways the advertisement message can be spread like TV ads, radio publicity, print promotion, online advertising, billboard marketing, in-store advertisement, WOM advertising, and endorsement. Now the question arises which category of promotion is best? The best nature

of advertisement depends on the type of industry or firm and its necessities and desires. McCracken (1989) found that celebrities' endorses characterized an effective way of transferring meaning to brands. The common conviction of the marketers is that there is a significant and huge impact of those advertisements, which are endorsed by the famous celebrities comparable with the non-endorsed celebrities.



### Positive Effects of Celebrity Endorsement

(R. Croft) cited that as competition is increasing between the firms to attract more consumers towards their brands, celebrities are increasingly used by the marketers to endorse their products. (Kulkarni and Gaulkar 2005) mentions that featuring a famous personality helps marketers in solving the problem of over communication. Celebrities because of their well knownness can assist advertisements to become more popular and stand out from the clutter and increase communicative ability. Celebrity endorsement helps in improving the brand's image and also polishes the company's image. Celebrity endorsement can also prove to be a powerful tool in entering foreign markets, it helps the company to overcome numerous issues. For instance, L'Oreal uses Ashwariya Rai and Sonam Kapoor in their White Perfect range, Kareena Kapoor in Lakme Eye conic kajal and Kajol in Olay aging cream are all considered popular in India. It is

also seen that those products which are endorsed by celebrities' stand out from other competing products because of their high level of recall and celebrities also create an impact on consumers' minds by making the advertisement memorable so it can be easily retrieved at the time of shopping. Lastly, celebrities who appear in any advertisement for endorsing a product, they are actually capable of breathing life in a falling brand and can help that falling brand to rise, in other words, they create new interest and excitement in consumers. All these arguments point to one conclusion that celebrity endorsement creates a positive impact on consumers' buying behavior (Goldsmith, Lafferty and Newell 2000).

### **Negative Effects of Celebrity Endorsement**

However celebrity endorsement has countless benefits but there are also certain risk factors that are associated with celebrity endorsement as a marketing strategy. Negative publicity regarding the celebrity is a major risk with endorsing a celebrity and there are other factors that can lead to serious consequences like :

- Popularity starts decreasing
- Moral issues
- Over endorsing can lead losing credibility
- Overshadowing of endorsed products.

Negative publicity about a celebrity endorsing a brand can change the perception of the consumer about that celebrity and it may also damage the brand reputation resulting in marketers to pay a big price over the celebrity's misdeeds and face serious humiliation and embarrassment. (Kumar 2010) mentions that there are times when consumers actually only notice the celebrity appearing rather than the product, hence this idea to promote that brand fails miserably. (Cooper 1984) says that "the product, not the celebrity, must be the star." This overshadowing is also called "vampire effect" because there is lack of clarity for the consumers because they are found to take more notice and interest in the celebrity rather than the interest. Another issue that arises is overexposure when marketers employ well recognized celebrities to endorse their brands and

it ultimately confuses the consumers and they are unable to correctly recall about that brand which celebrity stands for.

### **Consumer's buying behavior in India**

India is home to a host of languages, climates, and traditions. Defining the target audience is key to understanding consumer behavior in India. While city startups attract major investment, rural villages hail electricity and sanitation. Marketing in this diverse region can be daunting. But the potential is huge. India is set to overtake the US economy by 2030, second only to China. It already has the second-largest internet population – and only 41% of citizens are online. Big brands are already investing in Indian expansion.

India is known as a country of savers, though spending is on the rise. Disposable income is increasing. By 2030, 80% of households will be middle-income, compared to 50% today. This growing middle class is expected to drive consumer spending – buying more and buying better.

Access to credit also plays a key role. Indian millennials still value traditional priorities, like income and home ownership. But they have other goals, too. This shows in their spending priorities. Value for money is an important factor in determining consumer behavior in India. Indian shoppers are well-informed and want to get a good deal – even for luxury products. Moreover, eCommerce is new to many Indians, particularly outside the big cities. Programs like Amazon Easy are connecting traditional stores to the eCommerce sector. Innovative strategies like this help bridge the gap between customers and new technology. It's not surprising personal communication and trust are so highly valued.

Bargaining is a national pastime. Building trust also means ears to the ground. Research your audience carefully and consider expanding one city at a time. Then follow up with great service and a clear returns policy. The Indian family is changing. Traditional, multi-generational households are in decline. This can affect buying behavior. Marriage is still central to most people's lives. It's not unusual for people to spend 20% of their lifetime's earnings on their wedding. This means weddings are a major focus for the apparel, beauty, events, and luxury sectors. Themes of nurture, care, and affection are prominent in advertising – and successful. Hence, consumer behavior in India is complex and fast-changing.

**Celebrities Appearing in Cosmetics Advertisements**

BRANDS	PRODUCTS	NAME OF CELEBRITIES
L'Oreal	White Perfect	Ashwaria Rai
Oriflame	Foundation and Lip Color	Sonali Bendre
Lakme	Eyeconic	Kareena Kapoor
Olay	Aging Cream	Kajol

**CHAPTER 4**  
**RESEARCH AND METHODOLOGY**

## **RESEARCH AND METHODOLOGY**

### **4.1 RESEARCH OBJECTIVES**

The objectives of this research study are to get information about:

- The impact celebrities have on consumers' buying intention in regards of credibility.
- The impact attractiveness of celebrities have on the extent to which consumers convinced.
- The impact celebrity endorsement creates on consumers in terms of product fit match between the product and the celebrity.
- The extent of transfer of meaning to the products endorsed by celebrities in shaping buying intention.

### **4.2 HYPOTHESIS**

Following are the hypothesis of this research:

H1: Celebrity endorsed advertisement is considered to be effective in terms of buying intention in the cosmetic industry, when the celebrity used is credible.

H2: Celebrities who are attractive, create a positive impact on consumer's buying intention.

H3: The celebrity/product match-up positively influence consumers to purchase the celebrity endorsed product.

H4: Celebrity endorsed products transfer meanings which positively affect consumers buying intention.

### **4.3 RESEARCH DESIGN**

The research design of this study is also cross-sectional as it takes place at a single point in time. The participants are selected based on particular variables of interest. It considers numerous characteristics at once and can provide information about the current population.

#### **4.4 SOURCES OF DATA**

**Primary Data:** Primary data will be gathered through the distribution of questionnaires to the respondents and their answers will be recorded, which will be the primary data.

**Secondary Data:** Secondary data would be collected through information given in reports, newspapers, magazines, articles and textbooks.

**Software Used:** SPSS, Excel and spreadsheets would be used.

#### **4.5 SAMPLE DESIGN**

The population involves the desired sample size of 303 females aged between 18 to 30 years from Ernakulam district. The method of selecting the sample was done by using the snowball sampling. The respondents were required to be regular purchasers of cosmetic products.

#### **4.6 SAMPLE SIZE**

The sample size is 303 respondents from the adult female population aged between 18 to 30 years.

#### **4.7 SAMPLING METHOD**

The sampling method of snowball sampling has been used in the study. It is a form of non-probability sampling. The participants selected were all females within the age group of 18-30. In the present study, the selected participants recommended potential participants, who

themselves were observed and asked to nominate others and so on until a sufficient number of participants were obtained.

#### **4.8 METHOD OF DATA COLLECTION**

The data in the present study has been collected from the population by giving out an online questionnaire through Google Forms. A consent form was provided at the beginning of the questionnaire to make sure that the confidentiality of the participant's data will be maintained. This was followed by a few questions that collected the demographic details of the participants. The participants took about 10 to 15 minutes to complete the questionnaire and it was ensured that they filled every item of the questionnaire. The questionnaire involves a number of questionnaire items that asks a question and provides a set of response options for participants to choose from. From this, the highest responses given in each item is considered for further analysis and the final results are to be obtained using the SPSS software version 28.0.

#### **4.9 DRAFTING QUESTIONNAIRE**

The questionnaire used in this study is used to assess the impact of celebrity endorsements on consumer buying behavior. The first three questions were scored based on a four point, two point and three point likert scale respectively. The fifth question was scored based on a choice scale of preference. The remaining questions were scored based on rating scales.

#### **4.10 DATA ANALYSIS TECHNIQUE**

##### **CORRELATION**

The data analysis technique of correlation is to be used in the present study. A correlation is a statistical measurement of the relationship between two variables. Correlation means association - more precisely it is a measure of the extent to which two variables are related. There are three possible results of a correlational study: a positive correlation, a negative correlation, and no

correlation. Possible correlations range from +1 to -1. A zero correlation indicates that there is no relationship between the variables.

A positive correlation is a relationship between two variables in which both variables move in the same direction. Therefore, when one variable increases as the other variable increases, or one variable decreases while the other decreases. A negative correlation (inverse correlation) is a relationship between two variables in which an increase in one variable is associated with a decrease in the other. A zero correlation exists when there is no relationship between two variables.

There are four types of correlations in statistics;

- a) Pearson correlation
- b) Kendall rank correlation
- c) Spearman correlation
- d) Point-Biserial correlation.

→ **Spearman correlation**

Spearman's rank correlation coefficient or Spearman's  $\rho$ , named after Charles Spearman and often denoted by the Greek letter rho, 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.

## **REGRESSION**

Regression is a statistical technique that is used to measure and describe the strength and shape of the relationship between two or more variables. It is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables. It can be utilized to assess the strength of the relationship between variables and for modeling the future relationship between them. Regression analysis includes several variations, such as linear, multiple linear, and nonlinear. The most common models are simple linear and multiple linear. Nonlinear regression analysis is commonly used for more complicated data sets in which the dependent and independent variables show a nonlinear relationship. Regression analysis offers numerous applications in various disciplines, including finance.

**CHAPTER 5**  
**DATA ANALYSIS**

## DATA ANALYSIS

The aim of the present study was to investigate the impact of celebrity endorsements in consumer buying behavior. Spearman's correlation and linear regression was used to find the degree of relation as well as the prediction level.

The normality test between the independent and dependent variables didn't follow normal distribution. Hence the test used in the analysis of the data is non-parametric

## CORRELATION

**TABLE 4.1:** Correlation coefficient between Celebrity Endorsement and Consumer Buying Behavior (Spearman's rho)

Independent Variable N	Dependent Variable	r	p (2 tailed)
Celebrity Endorsement 303	Consumer Buying Behavior	.266**	.001

\*\* Correlation is significance at 0.01 level (2 tailed)

Spearman's correlation coefficient was computed to assess the relationship between celebrity endorsements and consumer buying behavior. From Table 4.1, the p-value is less than 0.001. This indicates that there is a statistically highly significant low positive relationship between celebrity endorsements and consumer buying behavior ( $p = .001$ ,  $r = 0.266$ ). In other words, we can infer that there is a weak relationship between the two variables.

## REGRESSION

**TABLE 4.2:** Regression Analysis Model Summary table between Celebrity Endorsement and Consumer Buying Behavior

<b>Model</b>	<b>R</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Std. Error of the Estimate</b>
1	.217	0.047	0.024	2.18651

a. Predictors: (Constant), Celebrity Endorsement

In table 4.2, R value represents the simple correlation between the variables. It points out the linear relationship between two variables (celebrity endorsement and consumer buying behavior) which is .217. This indicates a low degree of correlation. R Square is the coefficient of determination. The R square value indicates how much of the total variation in the dependent variable (consumer buying behavior), can be explained by the independent variable (celebrity endorsement). In this case, 4.7% of variance in consumer buying behavior is accounted for by celebrity endorsement. It shows that there is a positive weak relationship between the two variables

**TABLE 4.3:** Regression Analysis ANOVA table.

<b>Model</b>		<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig</b>
1	Regression	69.642	7	9.949	2.081	.045
	Residual	1410.345	295	4.781		
	Total	1479.987	302			

a. Dependent Variable: Consumer Buying Behavior

## b. Predictors: (Constant), Celebrity Endorsement

The ANOVA table predicts how well the regression equation fits the data (ie. predicts the dependent variable). It shows whether the overall model has been accepted or not, which is determined by the significance value. In table 4.3, the significance level is 0.45 which is greater than the p value. This indicates that there is stronger evidence in favor of the alternative hypothesis. Also, p value is less than 0.001 which is less than 0.05. This indicates that the regression model statistically significantly predicts the outcome variable (ie., it is a good fit for data).

**TABLE 4.4:** Regression Analysis Coefficients table

<b>Model</b>		<b>Unstandardized B</b>	<b>Standardized Coefficients Beta</b>	<b>t</b>	<b>Sig.</b>
1	<b>Constant</b>	9.950		24.591	<.001
	<b>Expertise of the celebrity</b>	0.164	0.105	1.143	0.254
	<b>Trustworthiness of the celebrity</b>	0.219	0.131	1.318	0.188
	<b>Similarity between the endorser and consumer</b>	0.060	0.031	0.398	0.188

0.280	<b>Familiarity of the celebrity</b>	0.182	0.100	1.083
0.944	<b>Likeability of the celebrity</b>	-0.011	-0.006	-0.070
0.910	<b>Product-celebrity matchup</b>	-0.019	-0.010	-0.113
	<b>Meaning transferred to the product</b>	0.317	0.188	2.117 0.035

a. Dependent Variable: Consumer Buying Behavior

Table 4.4 shows the regression analysis coefficients obtained for each dimension of celebrity endorsement and consumer buying behavior. A linear regression was calculated to predict consumer buying behavior based on celebrity endorsement. Results of the linear regression indicated that there was a collective significant effect between celebrity endorsement and consumer buying behavior.

From table 4.4, it was found that the impact the expertise and trustworthiness of the celebrities have on consumer buying behavior had a significant level which was more than 0.001. These two factors are the subsequent properties of one of the dimensions measured in celebrity endorsement, which is credibility. This indicates that the credibility of the celebrity has no effect on consumer buying intention. Hence, hypothesis 1 is rejected.

Similarly, the impact the similarity between the endorser and consumer, familiarity of the celebrity and likeability of the celebrity have on consumer buying behaviour also had a significant level which was more than 0.001. Since these three factors are the subsequent properties of another of the dimensions measured in celebrity endorsement, which is attractiveness, it can be interpreted that the attractiveness of the celebrity has no effect on consumer buying intention. Hence, hypothesis 2 is rejected.

It was also found that the factor of product-celebrity matchup also had a significance level of more than 0.001. Which can again be interpreted that product-celebrity matchup had no significant effect on consumer buying intention. Hence, hypothesis 3 is rejected.

However, the aspect of meaning transfer to the product of celebrity endorsement on consumer buying behavior was found to have a significance level less than 0.001. This indicates that the property of meaning transfer in celebrity endorsement does have a positive effect on consumer buying behavior. Therefore, hypothesis 4 is accepted. That is celebrity endorsed products do transfer meanings which positively affect consumers buying intention.

## **CHAPTER 6**

### **FINDINGS**

## FINDINGS

The findings show that celebrity endorsement is effective and a relationship is one between four elements of celebrity endorsement with consumer's purchase intention.

The first attribute of celebrity endorsement is "credibility" which comprises two factors namely expertise and trustworthiness. As the research suggests, consumers, in the field of cosmetics, frame their purchase intention which is not necessarily in favor of those cosmetics which are endorsed by such celebrities who have some expertise in that field and are considered to be trustworthy. This might be an indication of the consequence in the recent loss of trust in celebrities due to a lack of perceived genuinity and expertise in them as perceived by consumers, especially young female consumers. Young females in the current generation are becoming less impressionable and seem to value other aspects of an endorser when it comes to their evaluation process in decision making and their ultimate choice of decision.

The second attribute of celebrity endorsement is "attractiveness", which is further divided into three factors namely similarity, familiarity and likeability. Since the domain of cosmetics is very fragile therefore attractiveness of a celebrity here doesn't play a very vital role because liking a celebrity wouldn't mean that a cosmetic will actually produce the desired result to the consumer or their skin as it did to the celebrity endorsing it. The research may also imply that young females, while making a decision regarding the purchasing of cosmetics, are very selective since the purchasing of cosmetics is highly subjective in its use and benefits. That is, the superficial qualities of the celebrity doesn't play a significant role in consumers' buying intention. Hence, the study suggests that attractiveness of a celebrity has no relationship with purchase intention.

The third attribute of celebrity endorsement is "product celebrity match-up", which means that the celebrity endorsing a product has an image that goes with the product he/she is endorsing. This attribute implies that the product should match with the celebrities in different ways including personality, passion and career. In other words, the products endorsed by the celebrity must show some level of relevance with them. The study suggests it has no relationship with the purchase intention. This might indicate that even if the product does match-up with the celebrity, female consumers of cosmetics still, for some reasons, find it a good enough reason to

purchase cosmetic items. This can again come down to the perceived genuinity of the celebrity and also likeness and need for the endorsed products.

The fourth and the last attribute of celebrity endorsement is “meaning that is transferred to the product by the celebrity”. The study found it has a positive relationship with the purchase intention because often consumers think that using such celebrity endorsed cosmetics will make them look more glamorous and classy. In other words, female consumers at the end of the day, buy cosmetics goods for the perceived appeal of not the celebrity, but what it brings to them. It also depends on how much the products resonate and can meet their needs and wants. It can also depend on the effectiveness in the transference of meaning of the product to the consumers by the endorser. It can also be what the seeming product value and brand name brings for the consumer.

Therefore “meaning transferred” positively impacts the purchase intention of female consumers in the field of cosmetics. That is, even if the celebrity endorser is credible, attractive and has a relevant matchup with the product endorsed, it ultimately depends upon the impact the product or the celebrity endorser has on the consumers' psyche which ultimately comes down to the consumer's decision why or why not to buy a particular brand or product.

**CHAPTER 7**  
**RECOMMENDATIONS**

## RECOMMENDATIONS

According to the study, marketers while using celebrities to endorse their cosmetics products, and subsequently their brand, should make sure that they fulfill the criteria of the following:

- Celebrities, while transferring meaning to the product they are endorsing, should depict reality. Unrealistic beauty standards and goals do not sit well with the current generation, especially in females. Moreover, inclusivity is a huge factor in the industry of cosmetics, especially nowadays. Marketers while using celebrities in endorsements must aim at setting a realistic example without coming off as fake or biased. The consumers must in turn find the endorsement genuine, relevant and meaningful.
- Celebrities with any kind of negative publicity should not endorse products. This holds true for any business, not just in cosmetics. One of the most important things a brand/business can do is to show their hands on commitment towards customer satisfaction and commitment. If consumers find the endorsement meaningless or even disregarding or disrespectful in any way through the use of an infamous celebrity, it can come off as extremely ignorant and distasteful. It can even lead to the defamation of the brand indefinitely.
- Marketers must always concentrate on what their brand stands for when thinking of potential endorsers. It will not generate sales and revenue to have a celebrity conflicting with brand identity just to capitalize on a current “it” person. If marketers are unsure of their brand’s identity with consumers, conduct initial research to see what consumers associate with your brand. In this age of social media, having a celebrity endorser who is ‘off brand’ could potentially hurt a company’s identity. In other words, celebrity endorsement should not come off as a "PR stunt".
- Celebrities, who possess a good name and profile in the field of cosmetics and in general, should endorse cosmetics as consumers get influenced by such appeal and more importantly, their reputation. With the ongoing "jumping on the bandwagon" trend, consumers might as well buy products just because it is popular to do so. Hence, marketers must be highly selective and adequate in choosing and employing celebrity endorsers. A consumer may buy the product because it might give them some sort of status or recognition that is tied with the celebrity endorsing the product.

**CHAPTER 8**  
**CONCLUSION**

## **CONCLUSION**

The study at hand was primarily conducted to examine the relationship between and the impact of celebrity endorsements on consumer buying behavior. It has been concluded that there is a statistically highly significant low positive relationship between celebrity endorsements and consumer buying behavior. The attribute "meaning transfer to the product" of celebrity endorsement is found to have a significant impact on behavior and intention of buying cosmetics among females in the young adult population of Ernakulam district. That is, female consumers may be interested to buy cosmetics goods for reasons other than the celebrity endorsement. These reasons may be the price, availability, quality and quantity of the product. Moreover, celebrity endorsement and consumer buying behavior was found to have a weak correlation.

## REFERENCES

- Baker, M. J., & Churchill, G. A. (1977). The impact of physically attractive models on advertising evaluations. *Journal of Marketing Research*, 14(4), 538–555.  
<https://doi.org/10.1177/002224377701400411>
- Chaiken, S. (1979). Communicator physical attractiveness and persuasion. *Journal of Personality and Social Psychology*, 37(8), 1387–1397.  
<https://doi.org/10.1037/0022-3514.37.8.1387>
- Friedman, H. H., Termini, S., & Washington, R. (1976). The effectiveness of advertisements utilizing four types of endorsers. *Journal of Advertising*, 5(3), 22–24.  
<https://doi.org/10.1080/00913367.1976.10672647>
- Goldsmith, R. E., Lafferty, B. A., & Newell, S. J. (2000). The impact of corporate credibility and celebrity credibility on consumer reaction to advertisements and brands. *Journal of Advertising*, 29(3), 43–54. <https://doi.org/10.1080/00913367.2000.10673616>
- McCracken, G. (1989). Who is the celebrity endorser? Cultural Foundations of the endorsement process. *Journal of Consumer Research*, 16(3), 310. <https://doi.org/10.1086/209217>
- Misra, S., & Beatty, S. E. (1990). Celebrity spokesperson and brand congruence. *Journal of Business Research*, 21(2), 159–173. [https://doi.org/10.1016/0148-2963\(90\)90050-n](https://doi.org/10.1016/0148-2963(90)90050-n)
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3), 39–52. <https://doi.org/10.1080/00913367.1990.10673191>
- Roy, S., Jain, V., & Rana, P. (2013). The moderating role of consumer personality and source credibility in celebrity endorsements. *Asia-Pacific Journal of Business Administration*, 5(1), 72–88. <https://doi.org/10.1108/17574321311304549>
- Silvera, D. H., & Austad, B. (2004). Factors predicting the effectiveness of celebrity endorsement advertisements. *European Journal of Marketing*, 38(11/12), 1509–1526.  
<https://doi.org/10.1108/03090560410560218>
- Till, B. D., & Busler, M. (1998). Matching products with endorsers: Attractiveness Versus Expertise. *Journal of Consumer Marketing*, 15(6), 576–586.  
<https://doi.org/10.1108/07363769810241445>

## APPENDIX

### Impact of Celebrity Endorsement on Consumers' Buying Behavior

1) How frequently do you encounter celebrity endorsed advertisements?

- Very often
- Occasionally
- Hardly
- Never

2) Which advertisement catches your attention the most?

- Celebrity endorsed advertisement
- Non celebrity endorsed advertisement

3) Do you trust celebrity endorsed advertisements?

- Yes always
- No, not at all
- Sometimes only

4A) Do you purchase cosmetics based on celebrity endorsed advertisements?

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

4B) Will your purchase intention be negative to purchase cosmetics which has side effects

but are endorsed by a celebrity?

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

## Impact of Celebrity Endorsement on Consumer Buying Behaviour

5) Which of the factor listed below influence you to purchase celebrity endorsed cosmetics product?

- Credibility of the celebrity
- Attractiveness of the celebrity
- Product celebrity match up
- Meaning that are transferred to the product

9) Rate these factors based on their importance and effectiveness that celebrities in the cosmetics field should possess (1 being the lowest and 5 being the highest)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Celebrities in ads help me recall products more promptly					
I perceive celebrity endorsement as very effective					
Celebrity endorsements are very influential					
Expertise of the celebrity influence my purchase decision					
Trustworthiness of celebrity influence mu purchase decision					
Familiarity of the celebrity influence my purchase decision					

## Impact of Celebrity Endorsement on Consumer Buying Behaviour

Likeability of the celebrity influence my purchase decision					
Product and celebrity match influence my purchase decision					
Celebrities can provide or transfer meaning to the product and this influences my purchase decision					
Negative publicity of the celebrity also impact my purchase decision					
I find celebrity endorsed product more classy, desirable and a symbol of status					
Using celebrity endorsed product makes me feel more glamorous and I feel more confident about myself					
I also recommend my friends to use celebrity endorsed products					

Project Report

On

**AN ENSEMBLE MODEL FOR  
ADVERTISEMENT CLICK THROUGH RATE  
PREDICTION**

*Submitted*

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

**MASTER OF SCIENCE**

*in*

**APPLIED STATISTICS AND DATA ANALYTICS**

*by*

**AKSHAYA K.U**

**(Register No. SM20AS001)**

**(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, **AN ENSEMBLE MODEL FOR ADVERTISEMENT CLICK THROUGH RATE PREDICTION** is a bonafide record of the work done by Ms. **AKSHAYA K.U** 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:

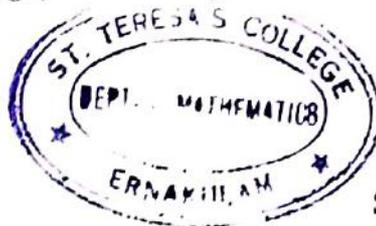
Place: Ernakulam

*Sreelakshmi*

Ms **SREELAKSHMI M S**

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

*Vahud*



*Shanty B.P*

Ms.shanty. B.P  
Assistant Professor HOD ,  
Department of statistics,  
St. Teresa's College(Autonomous),  
Ernakulam.

External Examiners

1: *Kumar Kan*

*26/5/2022*

2: *m.j.l.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 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: 9-5-22



**AKSHAYA K.U**

**SM20AS001**

# ACKNOWLEDGEMENTS

Primarily I would like to thank the supreme power the almighty god for being able to complete this project with success.it is my privilage to express my profound gratitude to my project guide Ms SREELAKSHMI M S ,Assistant professor ,Department of Mathematics and Statistics, St.Teresa's college Ernakulam, for the valuable support and guidance throughtout my project.

I am externaly thankful to the teaching and non-teaching staff pf the Department of Mathemtics and Statistics.

I also very thankful to HoD for their valuable suggestions, critical examination of work during the progress .

I would like to express my gratitude to each and everyone who helped and encourage me throughtout this project and made this a successful one

Ernakulam.

Date: 9-5-22



AKSHAYA K.U

SM20AS001

# ABSTRACT

Globally, the Internet has become the most prominent and accessible means of disseminating information about an event or pitching, advertising, and selling a product. Any advertising campaign's success hinges on reaching the right audience .class of target audience and, in the future, transform them into potential customers Engines of sear Google, Yahoo, and Bing are just a few of the most popular search engines utilised by businesses to sell their products. Aside from that, other websites, such as [www.alibaba.com](http://www.alibaba.com), which receives a lot of traffic, also provide B2B services. customers to create their own ad campaign The appearance of the advertisement, the daily maximum bill The audience's age and gender, the position's bid price, and the size of the advertisement are all factors. Businesses are primarily charged based on the amount of clicks on their advertisements, while certain websites additionally charge them. with a per-billing-cycle fixed charge This necessitates the need for advertising platforms to monitor and report on their data. Investigate these influencing elements in order to get the most out of your adverts. Additionally, it is critical for organisations to properly personalise these aspects in order to obtain the desired results. maximum number of clicks This study proposes a click-through-rate prediction algorithm that considers a number of factors. With advancements over time, the criteria described above can be used to forecast whether or not an advertisement would generate a click. the present systems in terms of data collection, features exploited, and strategies employed to increase the precision. We utilised an ensemble model with a weighted scheme and got an accuracy of 0.91 on a unit scale when predicting the likelihood of a user clicking on an advertisement

# 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>
1.1 IMPORTANCE . . . . .	2
1.2 MOTIVATION . . . . .	2
1.3 PROBLEM STATEMENT . . . . .	3
1.4 RESEARCH OBJECTIVES . . . . .	5
<b>2 LITERATURE REVIEW</b>	<b>6</b>
2.1 MACHINE LEARNING BASED APPROACH . . . . .	6
2.2 DEEP LEARNING BASED APPROACH . . . . .	9
<b>3 DATASET</b>	<b>13</b>
3.1 Description of Dataset . . . . .	13
<b>4 EDA AND FEATURE ENGINEERING</b>	<b>17</b>
4.1 Data preparation . . . . .	17
4.2 Exploratory Analysis . . . . .	18
4.2.1 Click Variations by granular time . . . . .	18
4.2.2 Relationship between the number of clicks and non -clicks received by Avazu on any single day	19
4.2.3 Relationship between the anonymous attribute ‘C1’ and the count of the clicks . . . . .	21

<b>5</b>	<b>MODEL BUILDING</b>	<b>22</b>
5.1	Logistic Regression . . . . .	22
5.2	Random Forest Algorithm . . . . .	23
5.3	XGBoost Algorithm . . . . .	24
5.4	Decision Tree . . . . .	25
5.5	Ensemble Learning Technique . . . . .	26
<b>6</b>	<b>BENCH MARKING TECHNIQUES</b>	<b>29</b>
6.1	The ROC AUC Curve . . . . .	29
6.2	Curve Of Sensitivity And Specificity . . . . .	30
6.3	Confusion matrix . . . . .	30
<b>7</b>	<b>EXPERIMENT AND OBERVATION</b>	<b>32</b>
7.1	Machine learning experiment . . . . .	32
7.1.1	Approach 1 . . . . .	32
7.2	Approach 2 . . . . .	36
7.3	Approach 3 . . . . .	37
7.4	Comparison of the Approaches . . . . .	38
7.5	Results . . . . .	39
7.5.1	Result 1 . . . . .	39
7.5.2	Result 2 . . . . .	39
7.5.3	Result 3 . . . . .	39
7.5.4	Result 4 . . . . .	39
<b>8</b>	<b>CONCLUSION AND FUTURE WORK</b>	<b>41</b>
	<i>REFERENCES</i> . . . . .	42

# Chapter 1

## INTRODUCTION

---

Much of the business has been made available to people as digital content over the internet in recent years. Most firms now have no choice but to market their items online through promotional adverts that can help them scale their business numerous times over. Advertisements on Google's Search Engine Results Page, Facebook pages, and Amazon's website are all examples of this. In today's world, users are exposed to a wide range of adverts. Text-based ads, video-based ads, and other interactive ads are only a few examples.

The click through rate (CTR) is a percentage that compares the number of individuals who see an advertisement to the number of people who actually click on it. The effectiveness of an advertisement is directly related to the click through rate. A firm can use this metric to see how well the keywords they've picked are performing in their advertising campaign. It's recurring in the sense that if businesses can predict the expected CTR before signing up for an ad campaign, they may choose the keywords and other features for the ad accordingly. By altering the keywords, bids, banner placements, and display size, a company could create various adverts (ads). Companies are eager to advertise their products aggressively utilising search engines and social media in a highly competitive and volatile market, and they pay fees based on the number of hits their ads generate [2]. From the standpoint of the search engines, this is a reasonable method of invoicing clients, and the click through rate serves as a good rationale.

## 1.1 IMPORTANCE

Given that advertising has become the primary source of revenue for many websites and mobile applications, it is unquestionably worthwhile to investigate the efficiency of advertising campaigns. As a result, predicting the click through rate has become crucial, which is why this study was conducted. This research will aid business owners in targeting their potential clients with the appropriate ad structure on a high-quality website. This study, on the other hand, will be valuable for search engines and website owners in determining which categories of advertisements are appropriate, as well as assisting online marketers with ad bidding rates, space, and orientation based on the traffic and rank of their website. As a result, this is an intriguing topic of research because it will provide intuitive insights to both advertisement platform owners and company suppliers in selecting advertisements that are beneficial in terms of monetary conversion.

Because the data was sensitive and not made public, a lot of study in this field has been stalled. The majority of search engines gathered the information and utilised it to determine the cost per click of adverts. There are now only a few data sets available for researchers from Avazu Labs and Alibaba, an ecommerce business to business (B2B) platform.

Any advancements in predicting whether or not an advertisement will be clicked would allow many organisations who wish to try a new advertising platform to properly budget for internet marketing and, as a result, examine the effectiveness of their business in the eyes of customers as a whole.

## 1.2 MOTIVATION

Even if the improvement in the prediction of the click through rate is minor, it could have a considerable impact on the stakeholder's finances [2][3]. The entire process of online advertising is dependent on the likelihood of a consumer seeing an ad and either clicking it or not. While it is important to note that a user clicking on an advertisement does not

guarantee that the user will become a converted customer, it is one of the greatest ways to make an educated guess that this person may be a future customer. When using the pay-per-click payment system, the ad publisher must precisely anticipate the likelihood of a user clicking on an advertisement in order to keep a large number of advertisers under their control. as this is the only key factor in determining the final revenue, it becomes a good area of research for the machine learning enthusiasts to work on.

Nonetheless, the issue may be understood from the perspective of the publishers who run the advertising platforms, the advertisers who subscribe to these platforms, and the consumers who see the adverts alongside their search engine results. Accurate advertising is yet another significant field of research that most publishers are now focusing on to find ways to enhance, as this will help them display the most relevant adverts to consumers. It is clearly insufficient for publishers to display ads solely based on the keywords that advertisers select for their campaigns, as there are numerous other highly weighted factors that are critical in this dilemma. For instance, it might be worthy depending heavily on the geo-location in addition to the search term from the consumer. Hence, considering this as a bigger problem, prediction of the click through rate becomes the sub problem in aiding the publishers to charge the advertisers and also for the advertisers to choose the right mode of online advertisements.

### 1.3 PROBLEM STATEMENT

The fundamental topic addressed in this study is predicting the advertisement's click-through rate, which is directly proportional to its effectiveness. When determining the success of an advertisement, several crucial factors are taken into account. The click through rate (CTR), buy list, and follow-on search behaviour, among other domain-specific metrics, are among them. Because the adverts are sold to businesses on a pay-per-click basis, it is critical for the publishers to maximise clicks, which is a direct advantage to the business in terms of marketing. Based

on the qualities of the advertisement, we will use prior statistics from an advertisement publisher to estimate whether or not a person will click on it.

The click through rate (CTR) is the ratio of the number of clicks received by an advertisement to the number of impressions received [2]. The total number of times a user sees an advertisement is referred to as an impression. This is an estimate of how likely a person will click on the recommended advertisement. Throughout this research, we assume that an advertisement will be effective if it has a higher ratio of clicks to views.

Another issue revolves around the different aspects that influence an advertisement's CTR. There are certain obvious factors, such as the user's profile, the user's search history, and the advertiser's profile, among others. However, there could be a slew of hidden linked factors that influence the CTR. When constructing models that estimate probability, it is critical to analyse and account these feature relationships. Because the features, such as the typical beer and diaper correlation, are implicit and nonlinear [1,] this study focuses on feature selection and analysis using various machine learning models.

The dataset used in this study is primarily made up of advertisement details registered by publishers and consumer behaviour logs. Because many of the data's properties are anonymous for security and privacy reasons, and the majority of them are categorical, the dataset becomes extremely dimensional [2]. There have been a variety of machine learning algorithms used to handle the challenge of forecasting click through rate. This scenario allows us a lot of flexibility in terms of analysing and ignoring features that have a minor impact on the anticipated likelihood.

In summary, the main challenges are to analyse and select the appropriate set of minimal features from the high dimensionality of the data, as well as to predict the click through rate using a combination of machine learning algorithms such as Random Forest and Logistic Regression, as well as Decision Trees, all of which are suitable for this

binary classification.

## 1.4 RESEARCH OBJECTIVES

The specific objectives of this research are listed below.

- Exploratory analysis on the data set to understand the features on the distribution of the data based the features.
- Preprocess the data to remove the null values and meaningless information.
- Feature selection on the dataset to identify and remove those features that do not contribute significantly to the prediction of the probability of a user clicking any given advertisement.
- Once these rich set of features are decided, one hot encode the features to create vectors and build machine learning models to find the probability.
- Bench mark the predictions of the various models using the confusion matrix, ROC AUC (Area Under the Receiver Operating Characteristics) curve, the sensitivity and specificity curves
- Also, create an ensemble of the models by varying the feature set and bench mark the same.

## Chapter 2

# LITERATURE REVIEW

---

Advertisements have a significant impact on recruiting targeted clients, and the way advertisements are presented has an impact on real sales. Furthermore, mobile commercials, in particular, are vital in an environment where time is of the essence: who will be the first to publish an advertisement and who will profit? The history click information is used to predict the Ads' CTR for a normal, ongoing advertisement, however this strategy does not work in the case of a new advertisement CTR prediction because there is not enough historical data for these new advertisements

### 2.1 MACHINE LEARNING BASED APPROACH

Fang et al. provide a method for predicting the Click-Through Rate of new advertisements. The Bayesian Network was used to create a model to predict the CTR of the new advertising. The initial stage was to create the Keyword Similarity model by building a Bayesian network of terms that characterise adverts in specific domains, also known as the keyword Bayesian Network (KBN). The second stage was finding keywords from advertisements with known CTRs that were comparable to the new ad keywords in order to approximate KBN conclusions. As a result, these commonalities can be utilised to locate advertising that are comparable to the new ad. Similar ads that are described by similar keywords will play a part in pre-

dicting the new ad CTR using the known ads' CTRs, based on the KBN inference's results. The tests were carried out using a test dataset of 27512 keywords. The dataset was created by combining the keywords from all of the adverts. The following were discovered as a result of the research. I In the event of a small number of keywords and a specified hardware setup, the KBN creation method is practical. (ii) The KBN inferences were accurate to a degree, with a reasonable error average. As a result, on average, the CTR forecast strategy for new advertising was accurate. Future work should be done to improve the proposed method in this study, which includes improving the KBN constructing method by adding data-intensive computing 1682 CMC, 2021, vol.66, no.2 techniques and improving the accuracy of the CTR prediction method in general. Similarly to the previous study, Kondakindi et al. [4] used the same dataset to train and evaluate the model with certain changes, such as the addition of new features and the removal of irrelevant fields. The study used a logistic regression technique with sufficient data pre-processing and received a high ranking in the Kaggle competition with a logloss score of 0.3938.

Furthermore, Dembczynski et al. employed decision rules to forecast the CTR of new adverts in another investigation. They also make suggestions to improve the quality of the adverts. The study makes use of Microsoft's Beyond Search dataset, which has 386,857,679 entries that explain the quality attributes of adverts. The research is divided into two sections. First, the maximum likelihood estimation was used to estimate the CTR for existing advertising (MLE). As a result, the system was able to learn all model parameters from real historical data. They compared the outcomes of the suggested model to the nave Bayes model after producing the mathematical function. The second step was to estimate the CTR for first-time viewers. They employed an algorithm for learning an ensemble of decision rules to forecast the CTR. The dataset was separated into three sets: training, validation, and testing, in proportions

of 25Percentage, 25 percentage, and 50 percentage, respectively. However, the study has a flaw in that the dataset utilised in the study only comprises the quality characteristic, leaving the contents of the advertisement out, rendering the suggestions useless.

Using numerous machine learning approaches such as Logistic Regression, Random Forest, and Gradient Boosting, Shi et al. created a model to predict the CTR and average cost per click (CPC) of a keyword. Contextual features and historical features were separated from the dataset's features. Historical features are more useful than contextual features for CTR prediction. The study's findings show that linear regression and Random Forest worked well for CTR prediction, while Gradient Boosting generated the worst results for both CTR and average CPC prediction.

Gai et al. introduced the Pies-wise Linear Model (LS-PLM), which is thought to be a solution for the CTR prediction domain's non-linear, large-scale, and sparse data problem. The model (LS-PLM) focuses on large-scale CTR prediction applications. It employs a divide-and-conquer strategy, partitioning the feature space into local sectors and fitting the linear classification model to each one. As a result of this procedure, weighted linear prediction combinations were created. The experimental results were discovered by running the algorithm on seven datasets obtained on different days from a system of mobile display advertising items. The proposed model was compared to the Logistic Regression (LR) Model, which was also utilised with the same dataset. The trials were designed to assess the performance of the LS-PLM, and the results showed that the Pies-wise Linear Model outperformed Logistic Regression and may be used in CTR prediction systems.

Using a logistic regression model, Xiong et al. [8] suggested a technique to predict advertising CTR based on user actions. To eliminate sparsity and redundancy, the features were first transformed into meaningful numerical parameters. Second, the problem of class imbalance was solved using a down sampling-based technique.

The features were then classified using heuristic thinking, and the inductive features were described using gradient trees. The dataset used in this study was a Tencent SOSO advertising log file. With an average R2 of 0.05 percent and an RMSE of 50.5 percent, the suggested model outperformed the baseline techniques. Saraswathi et al. also employed CTR prediction to determine how interested website visitors are in a given advertisement. The Naive Bayes Classifier, Logistic Regression, SVM, and Decision Tree are all examples of machine learning approaches that combine features. The data scientist compiled the dataset, which included numerous features. Human characteristics such as Frequent Time Spent on the Website, Lifetime, Field Revenue, Frequent Internet Usage, and Gender were incorporated into the model. The accuracy rate was 96 percent, which was far greater than previous models.

## 2.2 DEEP LEARNING BASED APPROACH

Deep Learning is also commonly used to forecast Click Through Rate. For the presence or absence of clicks as labels on binary classification, Edizel et al. [10] developed a Deep Character-Level Click-Through Rate Prediction (LCTR) model. The research is divided into two layers: word level and character level. With the query as the input into the system, the study creates a deep learning model to forecast the query and the advertisement CTR. The prediction of the CTR is the system's output. In comparison to the baseline model, both the word and character level techniques achieved considerable results. The feature selection and word2vec-based method have a big impact on the model's performance. The study demonstrated the importance of deep learning in CTR prediction.

Guo et al. [11] proposed the DeepFM model to combine the strengths of factorization machines for recommendation with deep learning for feature learning in a new neural network design. The studies were carried out on two real-world datasets: the Criteo dataset,

which employed 90 percentage of the data for training and 10 percentage for testing, and 7 consecutive days of user records from a commercial App for training and a single day record for testing. DeepFM combined the training of a deep and an FM component. It improved performance due to various advantages, including: (1) no pre-training is required; (2) the system learns both high and low-order feature interactions; and (3) the system learns both high and low-order feature interactions. (3) To avoid feature engineering, the system implemented a feature embedding sharing mechanism.

Wang et al. proposed a feature learning method for advertising data based on deep learning. K-means was used to aggregate similar objects into the same cluster. The study introduced a hybrid model Attention Stacked Autoencoder (ASAE) for advertising CTR estimation. The ASAE model trains a deep component and Factorization machines (FM) component together. Extensive experiments were conducted on two datasets (Frappe and SIGKDD) and then they compared the results with 5 baseline Models. The ASAE model showed a better performance and that due to the following reasons: (1) It does not need any pretraining; (2) It learns both high- and low-order feature interactions; 3) It uses the attention mechanism in neural network modelling to improve FM and make feature interactions contribute effectively to prediction.

Similarly, Zhou et al. suggested a method for predicting advertising CTR using a deep neural network and fuzzy logic theory. The following algorithms were utilised as a baseline for performance comparison in this study: LR, FM, GBDT+FM, and DBNLR. The FDNN approach has the best performance, whereas the LR method has the poorest. When compared to other machine learning algorithms, LR performs admirably [6,12]. Jiang et al. conducted another investigation that combined fuzzy logic and deep neural networks. The study surpassed other deep neural network tests and is more noise resistant.

Cacheda et al. [16] conducted research using three models: (1) A

hashing function was employed to reduce the amount of features in a logistic regression (LR) machine learning model; (2) Factorization Machines (FMs) using only two features; (3) Field-weighted Factorization Machines (FwFMs). These models were created using two data sets: Criteo CTR and Oath CTR. The results reveal that the FwFMs model outperforms the other two models despite having much less features. Surprisingly, the model that employed LR beat the model that used Factorized machine method for weighting fields. In addition, Chen et al. employed a Kaggle–Avazu dataset of time-series mobile advertisements in another investigation. Deep Belief Nets (DBN), a fusion model assessment, was proposed in the study; DBN was powerful yet simple to use. The DBN approach was then compared to Logistic Regression (LR) and Support Vector Regression (SVR). Finally, the DBN model beat the other two techniques, according to the findings.

Another study used an attentive deep interest (ADI) based model to capture the user’s interest in the relevant advertisement for the unique users. In addition, the different models were trained (Frappe and MovieLens) and tested (Books and Electronics) using four real-world datasets. In addition, seven approaches were evaluated to see if ADI produced the greatest outcomes. In comparison to the other groups, the study found that ADI had the best results. Because it compares seven different deep learning algorithms, the study is quite informative.

In addition, the FiBiNET model was proposed, which combined a shallow model with a deep neural network into a deep model to focus on the value of features. The FiBiNET is made up of six layers. To compare the different models, two public datasets [Criteo and Avazu] were used. Finally, because it emphasises the relevance of features, the FiBiNET model outperforms typical machine learning techniques. In comparison to earlier research, this one used the most complex algorithm of them all, with six layers of deep learning algorithms.

All prior research has demonstrated the importance of CTR and the combination of machine learning and deep learning for CTR prediction. Various firms consider CTR to be an important part of their ad efforts. Because it is still a hardcore issue in the realm of business, Machine Learning communities are still investigating new possibilities. With that in mind, we set out to complete a study that yielded a better result (AUC ROC) with a smaller number of features. We attempted to test a variety of machine learning algorithms in order to obtain the best accurate findings for the benefit of those who were interested.

# Chapter 3

## DATASET

---

This section discusses the dataset that this study has used for experiments and the bench marking. The details of the data are as below.

### 3.1 Description of Dataset

The dataset used in our study is the Click-Through Rate Prediction Competition Dataset from the Kaggle data science community [24]. The original data was collected from Avazu. The Avazu is a leading advertising platform that has provided the researchers with the dataset. The total training dataset contains information about 40 million mobile advertisements in 10 days. This can be used for the training purpose. Additionally, the test data also covers the same set of features as the training data and has features of the ads that were displayed in a day.

The dataset is seen to contain 21 different attributes of which the target attribute is the click details which is binary typed, 0 meaning an advertisement was not clicked and 1 meaning the click of an ad by the user. The dataset contains about 8 of the 21 features as the anonymous features. These are categorical variables and could be the detailed information about the users' profile and the advertisers' profile and are hashed to a unique value for the researchers to frame the vectors [4][9]. Some of the key qualities that qualify

this data set for this research are its enormity and coverage about the user, the advertisement and also about the advertiser. This helps us with rich useful information in categorizing the data. Although there are anonymous attributes, they are hashed to unique value and hence becomes seamless when we need to reduce the dimensional space of the dataset in hand

	id	click	hour	C1	banner_pos	site_id	site_domain	site_category	app_id	app_domain	...	device_type	device_conn_type	C14	C15
0	601394868	0	2014-10-21	1005	0	030440fe	08ba7db9	76b2941d	ecad2386	7801e8d9	...	1	0	18993	320
1	-59070594	0	2014-10-21	1005	1	0eb72673	d2f72222	f028772b	ecad2386	7801e8d9	...	1	0	16208	320
2	-1859646727	0	2014-10-21	1005	0	6c5b482c	7687a86e	3e814130	ecad2386	7801e8d9	...	1	0	17654	300
3	497487217	0	2014-10-21	1005	0	85f751fd	c4e18dd6	50e219e0	feb1138	82e27996	...	1	0	21611	320
4	-1852466777	0	2014-10-21	1005	0	1fbe01fe	f3845767	28905ebd	ecad2386	7801e8d9	...	1	0	15706	320

Figure 3.1: Dataset

The various attributes and their definitions are tabulated below.

- id : The unique identifier for all details that corresponds to one occurrence of an advertisement. This is a continuous variable.
- click : The target variable, 0 means an advertisement was not clicked and 1 means the ad was clicked. This is a categorical variable, binary typed.
- hour: The hour, in YYMMDDHH format. We could break this down and add additional features during the cleaning process. This is a continuous variable
- banner position: The position in the screen where the advertisement was displayed. This shows the prominent place for an advertisement to get the attention of the user. This is a categorical integer
- site id: The identifier to unique identify a site in which the advertisement was displayed. This is a hashed value.
- site domain: The domain information of the website in which the advertisement was displayed.

- site category: This is a categorical variable representing the field to which the website belongs to. This can be used to understand if any site category has more visitor attraction during any particular time.
- app id: The identifier to unique identify a mobile application in which the advertisement was displayed. This is a hashed value
- app domain: The domain information of the application in which the advertisement was displayed.
- app category: This is a categorical variable representing the field to which the application belongs to. This can be used to understand if any app category has more visitor attraction during any particular time. This is similar to the site category and can be compared relatively to check if app has more clicks over the website.
- device id: The unique identifier that marks the device from which the click was captured. This is a hashed continuous variable and can be repeated in the data set.
- device ip: The ipv4 address of the device from which the click was received. Hashed to a different value for privacy reasons to avoid trace back to the device.
- device model: The model of the device. We choose not to use this value.
- device type: The type of the device, is a categorical variable and has around 7 categories.
- device connection type: This is a hashed value about the connection type. We do not use this value for forming the vector.
- C1: An anonymous variable. It has influence over the prediction.
- C14-C21: Anonymous categorical variables that might have information about the advertisers' profile and the users' profile like the age, gender, etc.

We could classify these attributes in the data set into four sets and one target variable ‘click’

- \* Target feature : click
- \* site features : site id, site domain, site category
- \* device feature: device id, device ip, device model, device type, device conn type
- \* anonymized categorical features: C14-C21

The site and app specific attributes give information about the advertiser who are running the advertisement campaign. The anonymous attributes could contain the user specific and revenue specific details which are hidden for privacy reasons. The train data is approximately 6 Gigabytes and test data is approximately 1.5 Gigabytes in size. We intend to randomly sample the data to avoid over fitting during bench marking.

## Chapter 4

# EDA AND FEATURE ENGINEERING

---

### 4.1 Data preparation

Feature engineering [4] is a method of filtering and selecting high-quality characteristics for machine learning algorithms to learn from and reliably anticipate the target value [1]. These are the major elements in the training dataset that are directly related to the goal variable 'click.' Quality elements should be independent of one another, according to the literature [3][8]. If two characteristics have virtually identical meanings, it is not worth including in the feature set because it will diminish the model's accuracy. Features in our dataset, such as device id and device ip, will not help us forecast whether or not an advertising was clicked.

We also discovered that the dataset's 'hour' property is an integer in the format YYMMDDHH. This feature is beneficial in terms of comprehending statistics. For example, a company might expect a high volume of inquiries during peak hours but not at other times. However, from the standpoint of prediction, training the model with this feature reduced the accuracy, thus this characteristic has been broken down into 'hour, day, year, and month.' To add these additional attributes to the dataset,

we employ data frames.

There were a few null values in the dataset, but they were insignificant. Site-based advertisements lacked value in comparison to app-based advertisements. The IP address values were not formatted correctly. The missing values are usually attributed to the most commonly occurring attributes discovered during the data exploration process.

The dataset’s monstrosity was one of its major flaws [2]. The data was so large that even Google Collab’s TPU infrastructure couldn’t handle it. For each of our machine learning models to be trained, we took a sample of the data. Furthermore, many of the characteristics were categorical and anonymous. When training the model, we employed one hot encoding to convert the features to vectors [9]

## 4.2 Exploratory Analysis

The next step as in any machine learning project is to understand the data [1]. We use the standard exploration techniques and plot the results to get an idea about the various values and implicit meanings contained in the features of our dataset [14]. Out of all the explorations computed, the most relevant and useful insights are included here.

### 4.2.1 Click Variations by granular time

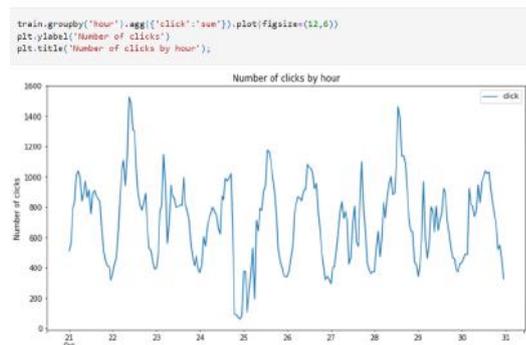


Figure 4.1: Number of clicks by hour

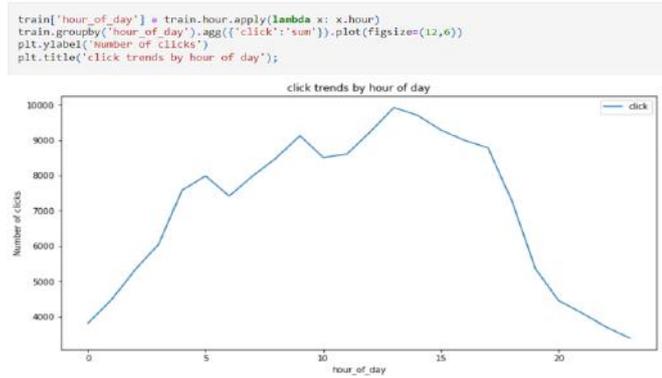


Figure 4.2: Number of clicks by hourofday

### Let's take impressions into consideration

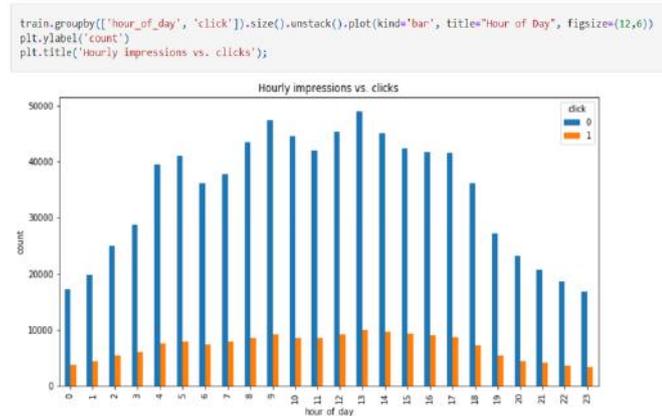


Figure 4.3: hourly impression vs clicks

#### 4.2.2 Relationship between the number of clicks and non - clicks received by Avazu on any single day

This is a key insight that helps us understand the relationship between the number of ads that were shown to the user to the number of ads that were clicked by them. These two are the essential factors that help us in determining the click through rate of an advertisement.

We have based this plot on the banner position additionally and understand that the values range form 0-7 for this categorical variable and most of the ads in this system were displayed at positions 0 and 7. We also use the device type to understand the relationship.

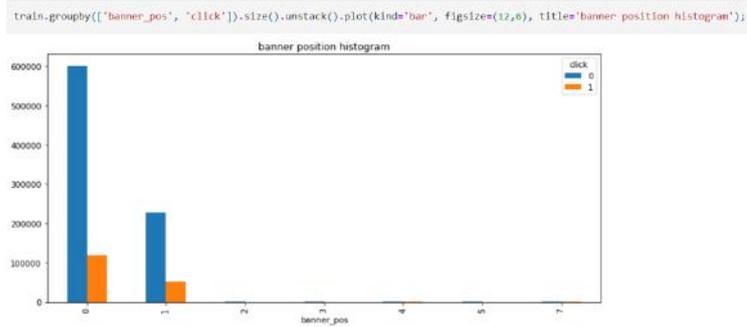


Figure 4.4: No. of clicks to non- clicks based on banner

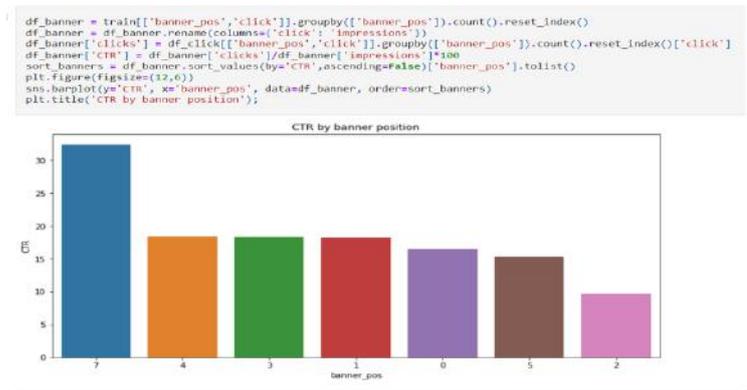


Figure 4.5: No. of impression to non-click based on banner

Although banner position 0 and 1 have the highest number of impressions and clicks, banner 7 enjoys the highest click through rate. Increasing the number of ads placed on banner position 7 seems to be a good idea.

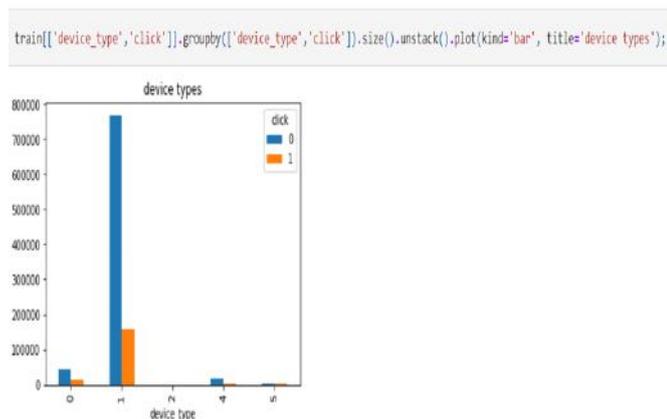


Figure 4.6: No. of clicks to non- clicks based on device type

### 4.2.3 Relationship between the anonymous attribute ‘C1’ and the count of the clicks

The attribute C1 is a categorical variable that is anonymized. We see this as an important feature for the training as most of the advertisement that received a click or not depends on this feature. We could see that the attribute value, 1005 has the maximum number of clicks and impressions.

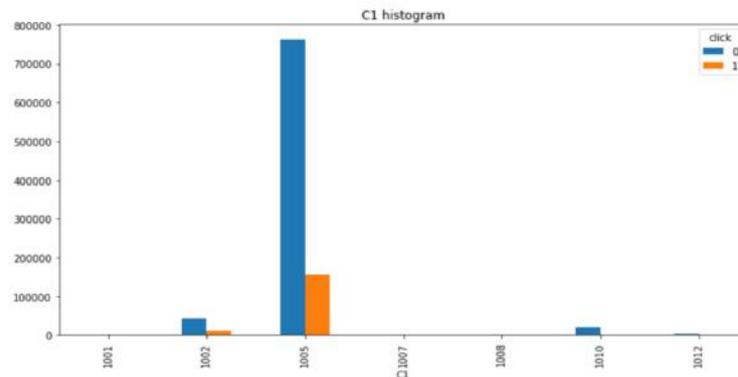


Figure 4.7: anonymous attribute ‘C1’ vs the count of the clicks

```
df_c1 = train[['c1', 'click']].groupby(['c1']).count().reset_index()
df_c1 = df_c1.rename(columns={'click': 'impressions'})
df_c1['clicks'] = df_click[['c1', 'click']].groupby(['c1']).count().reset_index()['click']
df_c1['CTR'] = df_c1['clicks']/df_c1['impressions']*100

plt.figure(figsize=(12,6))
sns.barplot(y='CTR', x='c1', data=df_c1)
plt.title('CTR by c1');
```

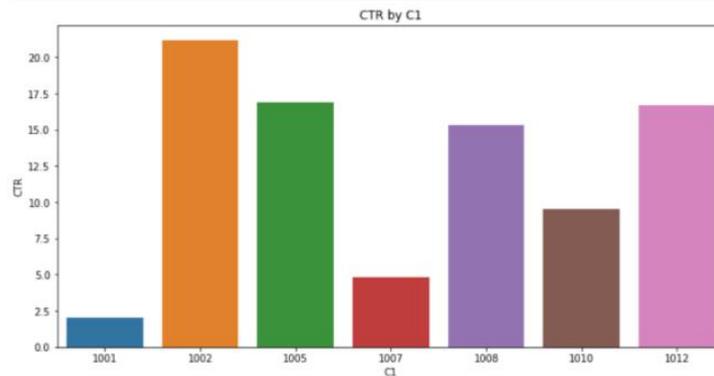


Figure 4.8: CTR by c1

## Chapter 5

# MODEL BUILDING

---

Based on the literature that studied recommendation of e-commerce products to the users and the relevant literatures that worked on predicting the click through rate of any advertisement dataset, we have used the following machine learning algorithms to predict the probability that a user will click the ad.

- \* Logistic Regression (Online Method)
- \* Random Forest Algorithm
- \* XGBoost algorithm
- \* Decision Tree model

We have predominantly used the Scikit learn libraries to implement these models.

### 5.1 Logistic Regression

Logistic regression, despite its name, is a classification model rather than regression model. Logistic regression is a simple and more efficient method for binary and linear classification problems. It is a classification model, which is very easy to realize and achieves very good performance with linearly separable classes. It is an extensively employed algorithm for classification in industry. The logistic regression model, like the Adaline and perceptron, is a statistical method for binary classification

that can be generalized to multiclass classification. Scikit-learn has a highly optimized version of logistic regression implementation, which supports multiclass classification task

For most classification problems, logistic regression [10] is a widely utilised projection analytic approach. It can be used to classify binary, multicategorical, and ordinal targets. When there are more than two classification categories, we can use softmax strategies in the regression learning process. We train using a set of occurrences of the binary classification issue, and we compute the goal probability for any occurrence of the same classification problem during testing. The relationship between the predicted value and all of the other features examined during the training phase is revealed by the final predicted probability. We leverage the logistic regression algorithm's online learning feature to deliver predictions in real time [1], and we design the data pipeline using Apache Kafka streams. When employing this technique, it's vital to remember to delete correlated features to avoid overfitting or underfitting the trained model [4]. To compute the probability for each data point in the test set, we utilise the scikit learn library's `predict_proba()` method

## 5.2 Random Forest Algorithm

As the name implies, a Random Forest is made up of a vast number of trees. Individual decision trees that collaborate to get a result. The class prediction is specified by each tree in a random forest, and the result is the most predicted class among the decision of trees. The Because trees defend each other from individual faults, Random The effects of the forest are incredible. Despite the fact that some trees may If you guess incorrectly, numerous other trees will fix the ultimate result. The trees will be able to move in the appropriate direction as a group as a result of this prediction. By mixing numerous weak

learners, Random Forests avoid overfitting. They are underfit since only a subset of all training samples are used. Random Forests have the ability to handle big datasets. Random Forests can handle a high number of variables in a data set with ease. During the forest creation process, they also make an unbiased assessment of the generalisation error. They can also precisely estimate the data that has been lost. Due to the random nature of the forest generation process, the fundamental disadvantage of Random Forests is their lack of reproducibility. Furthermore, the final model and subsequent outcomes are difficult to understand because they incorporate a large number of independent decision trees.

The random forest learning algorithm is a combination of different decision tree learning algorithms [12]. The algorithm chooses a different set of features, and if we look at the Random Forest implementation closely, we can see that it can intuitively figure out which of the features that are used as training parameters are good candidates for predicting the target variables. The random forest method can learn the relative value of any feature this way [13]. We could also fine-tune the model's hyper parameters and maximum features.

### 5.3 XGBoost Algorithm

The XGBoost is a well-known boosting approach that can be used to increase a model's efficacy while reducing training time. On several different classification problems, the XGBoost [4][14] algorithm has been benchmarked and compared to the random forest technique, with considerable improvements 38 in running time. The XGBoost, like the random forest, is an ensemble method that use both bagging and boosting techniques. The set of features is bundled together, and the target variable is determined based on that set. The error residue is determined using the boosting technique, and the decision tree evolves in

the next iteration to limit the occurrence of errors. In contrast to the bagging technique, the tree does not grow as much, as boosting will use a smaller number of splits in the iterations. These techniques make this model a natural choice for classification problems.

XG Boost is a fine-tuned and customised version of Gradient Boosting that boosts performance and speed. The scalability of XG Boost in all settings is the most crucial aspect in its success. The XG Boost is more than ten times quicker than typical methods on a single machine and scales to billions of samples in distributed or memory-limited environments. The scalability of XG Boost is due to a number of major algorithmic improvements. A theoretically justified weighted quantile sketch approach for handling instance weights in approximate tree learning is one of these improvements, as is an unique tree learning algorithm for handling sparse data. XG- Boost, on the other hand, makes use of out-of-core computation, allowing data scientists to process hundreds of millions of instances on a single workstation. Finally, integrating these tactics to construct an end-to-end system that can handle even larger data sets with the fewest clusters possible resources fascinates me even more.

## 5.4 Decision Tree

A well-known Machine learning technique is the decision tree classifier. A decision tree is a tree structure that resembles a flowchart, with each leaf node representing the conclusion, an inside node indicating a feature or attribute, and a branch representing a decision rule. A decision tree's root node is the node at the very top. It learns to split data based on the attribute's value. It is known as recursive partitioning because it partitions the tree recursively. This feature allows the tree classifier to deal with a wider variety of data sets, both numer-

ical and categorical. Furthermore, decision trees are ideal for dealing with nonlinear attributes-to-class correlations. The Gini Variety Index is employed as a known measure for overall performance, and an impurity function is determined on a regular basis to check the quality of the division for each node. The decision tree is adaptive in practice, allowing it to model nonlinear or unusual relationships. It is capable of deciphering the interaction of predictors. It can also be easily comprehended due to its binary form. The decision tree, on the other hand, has a variety of flaws that cause it to abuse data. Furthermore, it is difficult to update a decision tree with new samples.

The features provided to the decision tree model during the training phase are used to build a tree. To manage the model fitting, the correlated parameters must be deleted [7]. The rules will be constructed by the decision trees, and these rules will be matched on each new occurrence of the problem in the test set to return the associated target value [7]. The ensemble of these models is what we employ. The data is sampled at random and fed into these models. We aggregate the results of the forecast to return the probability once the prediction is made.

## 5.5 Ensemble Learning Technique

Ensemble methods are techniques that create multiple models and then combine them to produce improved results. ensemble methods usually produces more accurate solutions than a single model would

- \* Bagging is mainly applied in classification and regression. it increases the accuracy of models through the use of decision trees, which reduce variance to a large extent. the reduction of variance increases accuracy, hence eliminating over fitting, which is a challenge to many predictive models
- \* Boosting is an ensemble technique that learns from previous predictor mistakes to make better prediction in the future.

the technique combines several weak base learners to form one strong learner. Boosting takes many forms, which includes gradient boosting, AdaBoost etc.

- \* AdaBoost makes use of weak learners that are the form of the decision trees which most included one split that is popularly known as decision stumps. AdaBoost's main decision stump comprises observation carrying similar weights
- \* Gradient boosting adds predictors sequentially to the ensemble, where preceding predictors correct their successors, thereby increasing the accuracy of the model. New predictors are fit to counter the effects of errors in the previous predictors. The gradient of descent helps the gradient booster in identifying problems in learners' predictions and countering them accordingly.
- \* Voting ensemble involves summing the predictions made by classification models

We create an ensemble of all the above discussed algorithms [15]. We use the weighted classification method to create an ensemble model. The weights for each of these models [11] were decided based on the performance of the models and are discussed in the implementation section later. The training data is randomly sampled ( $S_1 - S_m$ ) for 25000 data points and are fed to the decision tree, online logistic regression, the stochastic gradient and the random forest machine learning algorithms to develop classifiers  $C_1$  to  $C_m$ . The classifiers individually predict the probability and the outputs are combined to infer The ensemble result

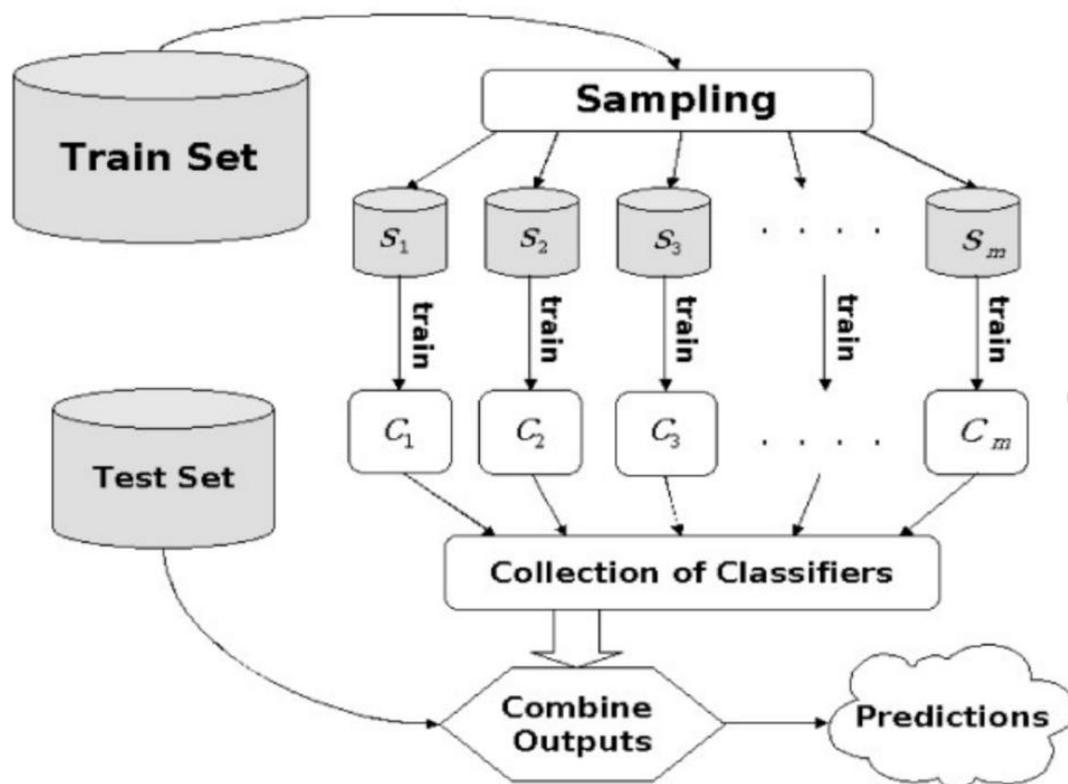


Figure 5.1: sample ensemble model

## Chapter 6

# BENCH MARKING TECHNIQUES

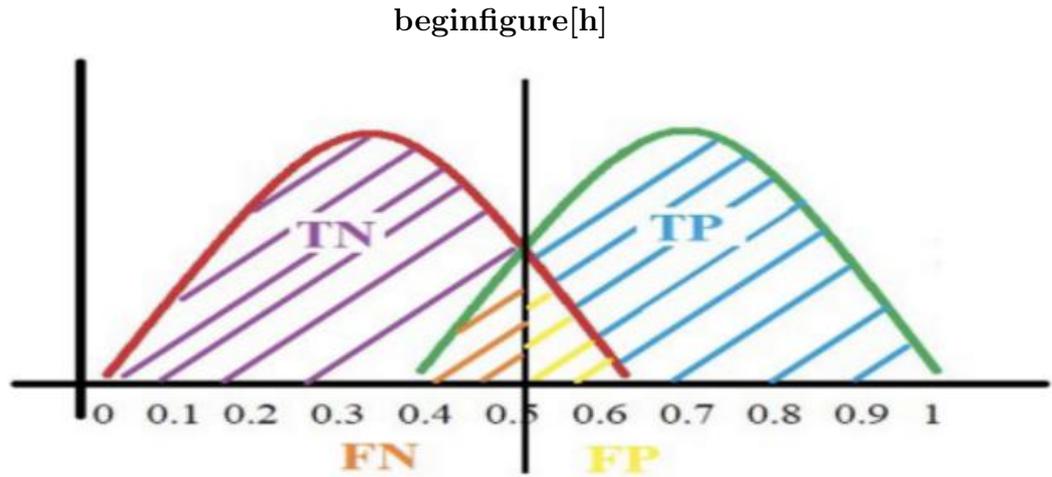
---

In This Study We Use The Following Evaluation Techniques

- \* The Roc Auc Curve
- \* The Sensitivity Specificity Curve
- \* Confusion matrix Confusion matrix

### 6.1 The ROC AUC Curve

The Receiver Operating Characteristics [20] can be expanded from the ROC. The Area Under the ROC is referred to as the AUC. This metric can be used to compare the categorization model's outcomes. We could see how well the model predicts True Positives (TP) and True Negatives (TN) by visualising it (TN). This refers to the model's ability to discriminate between two classes of the target variable under consideration [14]. The ROC AUC curve selects all potential threshold intervals between 0 and 1, and all occurrences of positive predictions crossing the threshold are true positives (shown in blue), whereas all occurrences of negative predictions crossing the threshold are false positives (shown in yellow). True negatives (shown in purple) and False negatives (shown in orange) have similar signals.



ROC AUC curve

## 6.2 Curve Of Sensitivity And Specificity

This is an extra benchmarking procedure that we use to determine the prediction's accuracy [16][17]. When there is an imbalance in the incident points in the training data, this measure is critical. The recall, or sensitivity, is the ratio of the number of incidents predicted as true to the overall forecasts. Similarly, specificity refers to the proportion of erroneous predictions compared to total predictions. The sensitivity and specificity of a model indicate its suitability for predicting a class of the target variable (true/false).

## 6.3 Confusion matrix

The confusion matrix [23] is a more common accuracy measure in the machine learning projects. It gives the relationship between the actual class versus the predicted class of the target value. We can generate the absolute matrix or the normalized confusion matrix [18].

	Predicted class		
Actual Class		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

confusion matrix

## Chapter 7

# EXPERIMENT AND OBERVATION

---

We Wrangle The Dataset And Use The Data Preparation Techniques Mentioned In The Section 4. Post These Steps, We Build The Decision Tree, The Logistic Regression, And The Random Forest Machine Learning Models Discussed Using The Libraries Mentioned Below, To Name A Few

- NumPy
- Csv
- Pickle
- Scikitplot
- Matplotlib
- Sklearn
- Seaborn

### 7.1 Machine learning experiment

#### 7.1.1 Approach 1

Due to the massive quantity of data available, we sample 25000 data points out of the 100000 data points available for training. This method is safe to use because the distribution of clicks against non-clicks is consistent across days [8]. We were able to solve the computing challenges thanks to this sampling strategy. Instead of the 'hour' element in the

dataset, we introduce three attributes: the day, the hour of the day, and the date. During the data preparation process, all device-specific features are also eliminated. To transform categorical data to a fixed length vector, we use one hot encoding [8]. Following that, we run the above-mentioned machine learning algorithms, adding and removing anonymous categorical variables as part of the parameter tuning process to improve accuracy. The experiments, as well as their outcomes, are summarised here. We use the ROC AUC curve [22], the accuracy and recall curves, as well as the Confusion matrix [24] to benchmark. All of the algorithms are benchmarked using the same method. Google Colab was used, which may be found at <https://colab.research.google.com>. For rapid execution during the training phase, we employ the Tensor Processing Unit, which is the default K80 Core in the online Google Collab environment.

#### Results of Decision Tree

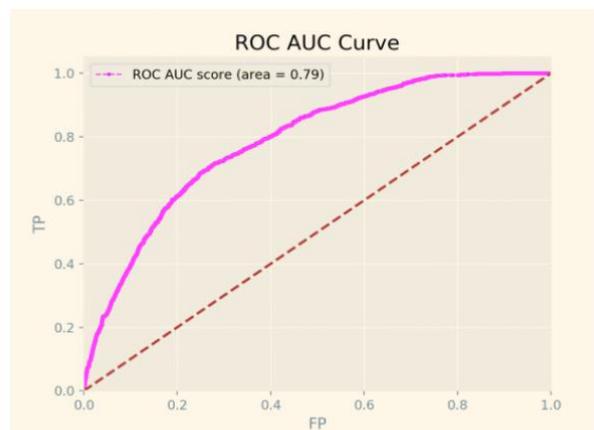


Figure 7.1: ROC AUC Curve – Decision tree model

The graph above shows that the decision tree model can correctly categorise positive and negative data points, click versus no click in this case, in 79 percent of the cases. Because of the large imbalance between the classes of data points in the CTR prediction dataset, the ROC AUC is an appropriate bench marking tool. During the exploration phase, we discovered that 80 percent of the advertisements in the dataset had no clicks, while only 20 percentage of the data samples had a click.

The below confusion matrix demonstrates that the model was able

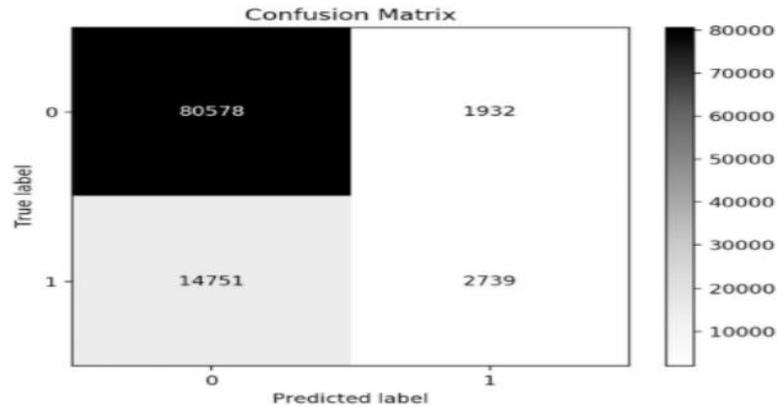


Figure 7.2: Confusion Matrix Decision tree model

to accurately classify data points that did not receive a click as 0 in the vast majority of cases. However, we can see that the model frequently fails to correctly detect genuine positives. We see this as a result of the dataset’s imbalance. We solve this problem by employing the sample techniques outlined in approach 3.

#### Result of online learning

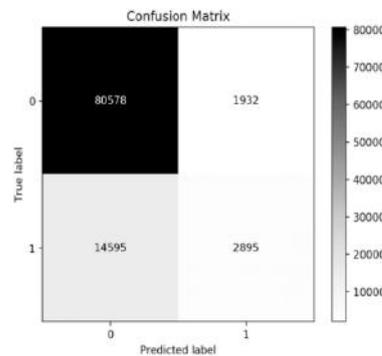


Figure 7.3: confusion matrix online learning

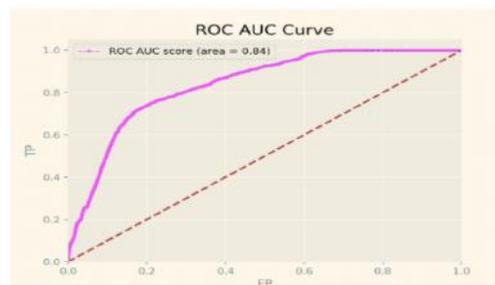


Figure 7.4: Roc Auc online learning

## Results of Random Forest

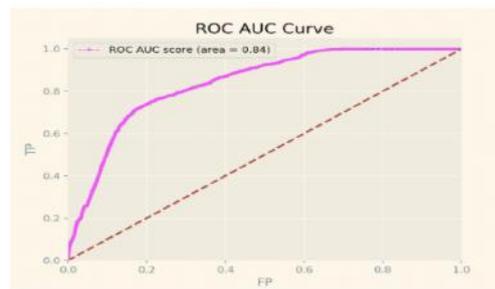


Figure 7.5: ROC AUC- Random forest

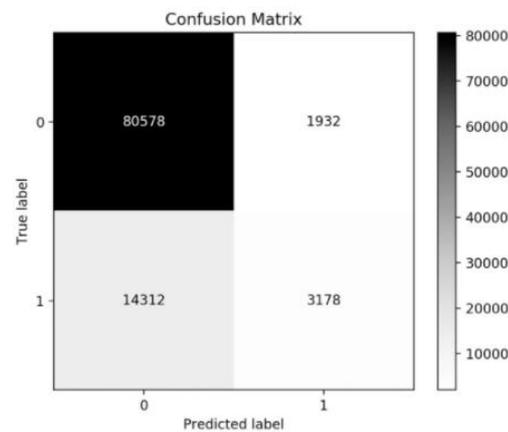


Figure 7.6: Convex matrix Random forest

Due to the imbalance in the data, we notice that all of the classifiers predict non-clicks better than clicks when employing each of the above discussed classifiers for prediction. Furthermore, data points that are incorrectly predicted by one model are correctly predicted by other models. This necessitates the employment of approach 2, which is the ensemble model of all of the above-mentioned classifiers

### Result of XGBoost

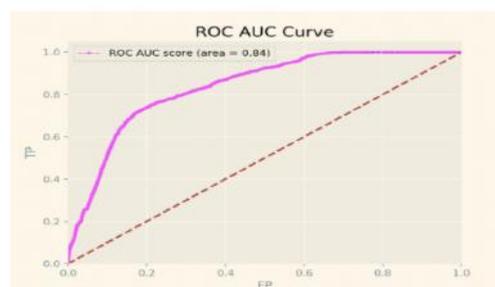


Figure 7.7: ROC AUC- XGboost

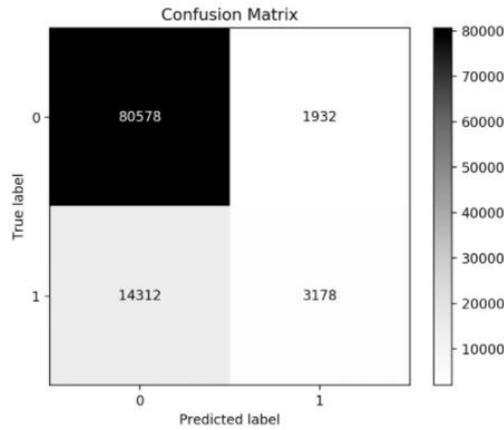


Figure 7.8: Confusion Matrix- xGboost

The results of the XGBoost technique was also very close to the Random Forest technique.

## 7.2 Approach 2

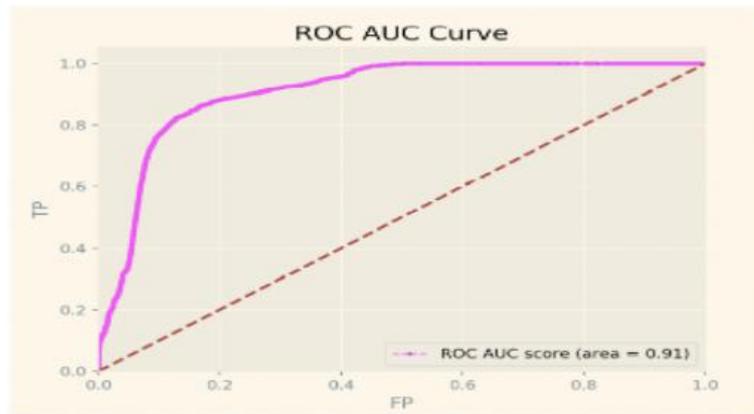


Figure 7.9: ROC AUC Curves Ensemble Model

Using the weighted strategy and the voting mechanism, we produced an ensemble of all of the above models. With the dataset’s imbalance, the weighted approach proved to be effective. the XGBoost and Random forest model is given a higher weight [18] than all the other models, , and all the other models are evaluated equally. One such experiment’s outcome is depicted below. The ROC AUC value is 0.91, which is the best in the group.

### 7.3 Approach 3

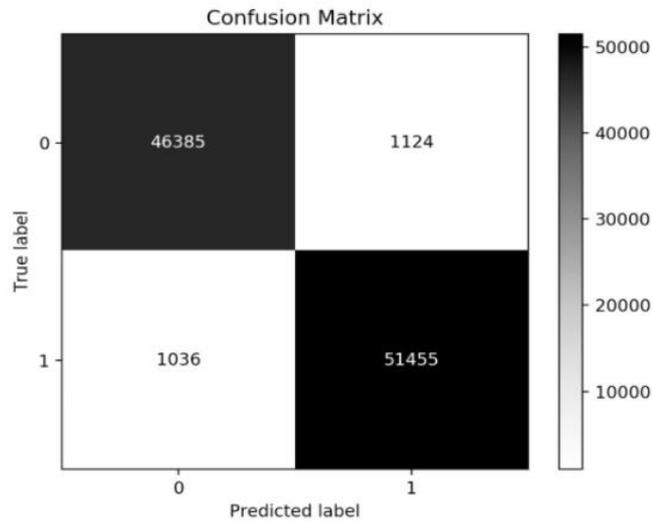


Figure 7.10: Confusion Matrix Over Sampled data

With all of the previous approaches, the model accurately classifies a non-clicked advertisement as 0 (True Negative), but also correctly forecasts the majority of clicked advertisements as 0, indicating that the model has a high precision value but needs to improve on false negatives. This is almost definitely due to a data set imbalance between the two classes of target variables. As a result, the target variable 1 is oversampled, bringing both target categories into balance. The outcomes of this experiment have been recorded. The confusion matrix shows that the model is currently doing well across the board [11][17]

## 7.4 Comparison of the Approaches

The accuracy of the machine learning models indicated in the confusion matrix and the ROC-AUC plots provided in sections 7.2 to 7.3 is represented by the numbers in the table.

Model	Accuracy score (+/-0.03)	ROC AUC	Confusion matrix
Decision Trees	0.833	0.79	Works well for True negatives, but misclassification with true positives.
Logistic regression (Online)	0.835	0.81	Works well for True negatives, but misclassification with true positives.
Random Forest	0.838	0.84	Works well for True negatives, but misclassification with true positives.
<u>XGBoost</u>	0.8383	0.84	Works well for True negatives, but misclassification with true positives.
Ensemble	0.845	0.91	Works well for True negatives, but misclassification with true positives.
Over sampling – Ensemble	0.85	0.91	Reduced error rate

Figure 7.11: comparison of the approaches

## 7.5 Results

The following observations are made in this study.

### 7.5.1 Result 1

The data in the dataset is evenly distributed throughout all days, with an imbalance in the target category. When the outcomes of the above approaches were compared, it was discovered that the IXGBoost and the random forest model had higher AUC values, thus these models were given larger weights in the ensemble experimentation. The overall AUC is 0.91, which is higher than previous studies [4][14]. However, the model is heavily reliant on categorical anonymous variables and does not cope well with data imbalance [15]. In terms of precision and recall, white papers from advertising platforms such as Google and Alibaba report superior outcomes. It's possible that they have full access to the data because there are no anonymous variables.

### 7.5.2 Result 2

When using the over sampled data to train the ensemble model, the overall best result was obtained. There are only 1036 false negatives out of 50000 positive data samples, and only 1124 false positives out of 50000 negative data samples. As a result, the rate of misclassification error is kept to a bare minimum.

### 7.5.3 Result 3

When compared to the weighted scheme, the ensemble model's performance with voting deteriorates. This is because when we round off the probability, a lot of misclassification occurs. Instead, we went with the weighted ensemble method, which produced the best results.

### 7.5.4 Result 4

For the ensemble, we randomly sampled the data for each of the classifier and results were observed to be consistent. Hence, the ensemble

**technique with over sampling to get balanced data is the overall winning approach in this study.**

## Chapter 8

# CONCLUSION AND FUTURE WORK

---

The rapid growth of the social media as the online advertising platform has made it compelling for the researches to focus on improving the misclassification rate in the prediction of the click through rate of an advertisement. This will help the advertiser choose the right set of attributes to target the potential audience for their business. The accuracy can also help the advertising platforms to decide on the cost of the advertisement. In this research, the steps from preparation of the dataset to the building of the classification model were discussed in detail. The results of each step were also discussed and bench marked against the base lines. A series of experiments were also performed and the results were captured and compared. The initial study was hard due to the enormous amount of data at hand. Nevertheless, the feature engineering techniques and sampling strategies were some of the key parameter tuning steps that helped improve the overall performance of the research. As a next step, this study can be improved in terms of the efficiency making it more usable. Given access to all the anonymized variables, noise in the data can be perfectly removed. We could tune the parameters wisely which could help us solve the more complex problem of choosing the right advertisement at the right time for a user.

## REFERENCES

- [1] P. Chan, X. Hu, L. Zhao, D. Yeung, D. Liu and L. Xiao, "Convolutional Neural Networks based ClickThrough Rate Prediction with Multiple Feature Sequences", Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018.
- [2] P. Barford, I. Canadi, D. Krushevskaja, Q. Ma and S. Muthukrishnan, "Adscape", Proceedings of the 23rd international conference on World wide web - WWW '14, 2014.
- [3] G. Zhou, K. Gai, X. Zhu, C. Song, Y. Fan, H. Zhu, X. Ma, Y. Yan, J. Jin and H. Li, "Deep Interest Network for Click-Through Rate Prediction", Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining - KDD '18, 2018.
- [4] C. Jie-Hao, L. Xue-Yi, Z. Zi-Qian, S. Ji-Yun and Z. Qiu-Hong, "A CTR prediction method based on feature engineering and online learning", 2017 17th International Symposium on Communications and Information Technologies (ISCIT), 2017.
- [5] S. Zhang, Z. Liu and W. Xiao, "A Hierarchical Extreme Learning Machine Algorithm for Advertisement Click-Through Rate Prediction", IEEE Access, vol. 6, pp. 50641-50647, 2018.
- [6] H. Guo, R. Tang, Y. Ye, Z. Li and X. He, "Deep FM: A Factorization-Machine based Neural Network for CTR Prediction", JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015 11.
- [7] M. Freser, D. Kosir, "Analysis of the open advertising data set", 6th Jožef Stefan International Postgraduate School Students' Conference, Pg. 111-120, Volume: 1, May 2014.
- [8] Xiangyang She, Shaopeng Wang, "Research on Advertising Click-Through Rate Prediction Based on CNN-FM Hybrid Model", 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics, Year: 2018 , Volume: 02, Page s: 56 – 59.
- [9] Jing Ma, Xian Chen, Yueming Lu, Kuo Zhang" A click-through rate prediction model and its applications to sponsored search advertising", International Conference on Cyberspace Technology, Year: 2013, Page s: 500 – 503.

- [10] Antriksh Agarwal, Avishkar Gupta, Tanvir Ahmad “A comparative study of linear learning methods in click-through rate prediction”, 2015 International Conference on Soft Computing Techniques and Implementations, Year 2015, Pages 97-102.
- [11] Feng Zhou, Hua Yin, Lizhang Zhan, Huafei Li, Yeliang Fan, Liu Jiang “A Novel Ensemble Strategy Combining Gradient Boosted Decision Trees and Factorization Machine Based Neural Network for Clicks Prediction”, 2018 International Conference on Big Data and Artificial Intelligence, Year 2018, Pages 29- 33.
- [12] Qianqian Wang, Fang'ai Liu, Shuning Xing, Xiaohui Zhao, Tianlai Li “Research on CTR Prediction Based on Deep Learning”, IEEE, Year: 2018, Page 1-1.
- [13] Xiaokang Qiu, Yuan Zuo, Guannan Liu, “ETCF: An Ensemble Model for CTR Prediction”, 2018 15th International Conference on Service Systems and Service Management, year 2018, pages 1-5.
- [14] Yuyu Zhang, Hanjun Dai, Chang Xu, Jun Feng, Taifeng Wang, Jiang Bian, Bin Wang, and Tie- Yan Liu. “Sequential click prediction for sponsored search with recurrent neural networks.”, AAAI, Conference on Artificial Intelligence, Year 2014.
- [15] Weinan Zhang, Tianming Du, and Jun Wang. “Deep learning over multi-field categorical data - - A case study on user response prediction”, Year 2016, European Conference on Information Retrieval.
- [16] Jing Ma, Xian Chen, Yueming Lu and Kuo Zhang, ”A click-through rate prediction model and its applications to sponsored search advertising,” International Conference on Cyberspace Technology (CCT 2013), Beijing, China, 2013, pp. 500-503.
- [17] R. Kumar, S. M. Naik, V. D. Naik, S. Shiralli, Sunil V.G and M. Husain, ”Predicting clicks: CTR estimation of advertisements using Logistic Regression classifier,” 2015 IEEE International Advance Computing Conference (IACC), Bangalore, 2015, pp. 1134-1138.
- [18] B. Tan and R. Srikant, ”Online Advertisement, Optimization and Stochastic Networks,” in IEEE Transactions on Automatic Control, vol. 57, no. 11, pp. 2854-2868, Nov. 2012

[19] Q. Cheng, P. K. Varshney and M. K. Arora, "Logistic Regression for Feature Selection and Soft Classification of Remote Sensing Data," in *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 4, pp. 491-494, Oct. 2006.

[20] "ROC AUC Curve", 2019, Available online

<https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>

[21] "ROC Curve for accuracy" 2019, Available online at <http://fastml.com/what-you-wanted-to-know-about-auc/>

[22] "Scikit Learn for classification models", 2019, Available online at <http://benalexkeen.com/scoring-classifier-models-using-scikit-learn/>

[23] "The Scikit Learn Documentation", 2019, Available online at <https://scikitplot.readthedocs.io/en/stable/metrics.html>

[24] "Seaborn Plots", 2019, Available online at

<https://seaborn.pydata.org/tutorial/relational.html#relating-variables-with-scatter-plots>

[25] "The Matplotlib documentation", 2019, Available online at <https://matplotlib.org/>

[26] Zhu J., Guan Q., Zhao X. (2015) Multi-class JPEG Image Steganalysis by Ensemble Linear SVM Classifier. In: Shi YQ., Kim H., Pérez-González F., Yang CN. (eds) *Digital-Forensics and Watermarking. IWDW 2014. Lecture Notes in Computer Science*, vol 9023. Springer, Cham

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

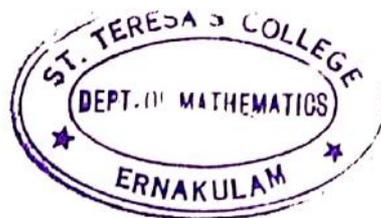


CERTIFICATE

This is to certify that the dissertation entitled, **STUDY ON ONLINE CLOTH SHOPPING BEFORE AND DURING COVID 19** is a bonafide record of the work done by Ms. **EVANA T D** under my guidance as partial fulfilment 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.

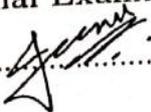
Date: 04/03/2022  
Place: Ernakulam

  
**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:  .....

2: .....

Project Report

On

**STUDY ON ONLINE CLOTH SHOPPING  
BEFORE AND DURING COVID**

*Submitted*

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

**BACHELOR OF SCIENCE**

*in*

**MATHEMATICS**

*by*

**EVANA T D**

(Register No. AB19AMAT015)

*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, **STUDY ON ONLINE CLOTH SHOPPING BEFORE AND DURING COVID 19** is a bonafide record of the work done by Ms. **EVANA T D** under my guidance as partial fulfilment 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: 15/02/2022

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:.....

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.

**EVANA T D**

Date: 15/02/2022

**AB19AMAT015**

# ACKNOWLEDGEMENT

I wish to express my gratitude to our project guide, Dr Elizabeth Reshma M T of the Department of Mathematics and Statistics of St Teresa's College(Autonomous), Ernakulam. The completion of this project could not have been possible without the assistance of Smt. Reshmi S and Smt. Rosmin Raju.

I thank all the teachers who in one way or another shared their support. Also, I would like to thank everyone who helped me in finishing my project.

Ernakulam.

Date: 15/02/2022

EVANA T D

**AB19AMAT015**

# 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>1.1 SIGNIFICANCE OF STUDY</b> . . . . .	2
<b>1.2 OBJECTIVES OF THE STUDY</b> . . . . .	3
<b>1.3 LIMITATIONS OF THE STUDY</b> . . . . .	3
<b>2 DATA DESCRIPTION</b>	<b>4</b>
<b>2.1 DATA SOURCE</b> . . . . .	4
<b>2.2 DATA DESCRIPTION</b> . . . . .	4
<b>3 METHODOLOGY</b>	<b>6</b>
<b>3.1 STATISTICAL SURVEY</b> . . . . .	6
<b>3.2 TEST USED</b> . . . . .	7
<b>4 DATA ANALYSIS</b>	<b>9</b>
<b>4.1 GRAPHICAL ANALYSIS</b> . . . . .	9
<b>4.2 STATISTICAL ANALYSIS</b> . . . . .	24
<b>4.2.1 EFFECT OF FACTORS ON THE PREFERENCE             OF SHOPPING</b> . . . . .	24
<b>4.2.2 RELATION BETWEEN MONTHLY INCOME             AND MONTHLY PURCHASE RATE</b> . . . . .	27
<b>4.2.3 EFFECT OF FACTORS ON THE ROLE OF DIS-             COUNT</b> . . . . .	29

5	CONCLUSION	33
	<i>REFERENCES</i> . . . . .	36
	<i>ANNEXURE</i> . . . . .	37

# Chapter 1

## INTRODUCTION

---

For the last decade, the evolution in technology is remarkable. We are in the era of globalization where online shopping is one of the greatest achievements. For the last couple of years, the increase in online shopping is very evident due to the widespread of Covid-19. Our whole country went to a complete lockdown by the end of March 2020. Every sector was negatively affected by it, directly or indirectly. The economy of India, like many other countries, has decreased considerably. During these dark times, one of the possible means to buy things was online shopping.

Before this study was conducted, many people had already done studies over online shopping in this pandemic period. Among these studies, we referred some which happened to be helpful.

'Trust in online shopping during COVID 19 - A case study from Kosovo (Faculty of Economics, University of Prishtina "Hasan Prishtina", Kosovo)' states that online shopping is the easiest way in this busy world especially during COVID 19. They studied the convenience of online shopping. The increase in e-shopping with the development of information technology has an impact on the world economy. E-shopping enables us to be more global and also reduces trade barriers. They concluded that the shopping behaviour changes due to some factors.

'A study on the influence of COVID 19 pandemic on customers' buying behaviour (by Bharati Aggarwal and Deepa Kapoor)' says that online shopping is an e-business that allows purchasing goods by using the in-

ternet with the help of a web browser. They concluded that the internet plays an important role in e-business that provides various facilities (by Rani and Charumathi).The study says that the increase in the usage of the internet has changed the trend in shopping. They concluded that the online shopping has tremendously increased during COVID 19.

'Online shopping behaviour during COVID 19 pandemic in India (by Amit Ranjan, Madhavendra Misra and Jitendra Yadav)' affirms that e-commerce is a platform where customers can enjoy their shopping through different websites and channels. They say that the customers need something beyond their expectations and gratification. They just don't need high-quality products but also require multi-channel services.

'Impact of COVID 19 pandemic on online consumer purchasing behaviour (by Shengyu Gu, Beata S'lusarczyk, Sevba Hajizada, Irina Kvalyova and Amina Sakhbieva)' asserts that by 2020 the number of e-commerce publications has increased more compared to 2000. They concluded that seven clusters were introduced through the analysis of e-commerce publications.

## 1.1 SIGNIFICANCE OF STUDY

The study tries to examine user behaviour and experiences of consumers participating in online shopping. Online shopping has created a space for small scale businesses to become successful. So by examining consumer behaviours, business owners can make optimized marketing plans and strategies to improve their commerce. The survey also investigates the factors responsible for consumers' choices and preferences in online shopping. This paper also mentions various challenges people face in online shopping. The results could also be used for future references.

## 1.2 OBJECTIVES OF THE STUDY

- Identify various factors affecting preference of shopping and role of discount
- Relation between monthly income and monthly purchase rate

## 1.3 LIMITATIONS OF THE STUDY

The study is limited to the Ernakulam district of Kerala. The study is only based on consumer behaviour in shopping online. The data collected for research purposes is based exclusively on the primary data provided by the respondents. There is a risk of personal bias. So the accuracy is not right. Due to time constraints and other restrictions, the survey was limited to only 261 participants.

## Chapter 2

# DATA DESCRIPTION

---

### 2.1 DATA SOURCE

Here we have used the primary method. Due to the pandemic, the survey was done by means of google form. Moreover, an online questionnaire is cost-effective, especially in this era of social distancing. We have collected 332 samples out of which we selected 261 samples which are from Ernakulam district. For getting more effective results samples are only collected from Ernakulam district. The age group which we are concentrating on is 15-65, which includes students, employees, unemployed, housewives, etc.

### 2.2 DATA DESCRIPTION

The variables are

- Age
- Gender (Female/Male)
- Place of residence (Urban/Rural)
- Monthly income (Below 10,000/10,000-50,000/50,000-1,00,000/Above 1,00,000)
- Status of online purchase - Yes(have purchased online) or No(have not purchased online)

- **Role of discount - Yes(it has a role) or No(it does not have a role)**
- **Monthly purchase rate (Below 1,000/-, Below 2,000/-, Above 2,000/-)**

## Chapter 3

# METHODOLOGY

---

### 3.1 STATISTICAL SURVEY

It is a scientific process of collection and analysis of numerical data. The facts and figures related to our area of interest are called data, which may be qualitative or quantitative. Planning is the most essential step in the statistical survey. A questionnaire is a list of questions used for the collection of information in an investigation. The success of the study depends to a very great extent on the questionnaire. In the absence of adequate and accurate data, the application of statistical techniques cannot fetch good and reliable results. Then the classification is done based on the characteristic under consideration. After the collection and classification of data, statistical analysis and interpretation of the results are done. Finally, a report that contains a complete description of all the stages of enquiry is published.

Statistical surveys allow us to reach thousands of participants. Surveys provide a high level of general capability in representing a large population. It is of low cost and can be administered to the participants in a variety of ways. Because of the high representativeness brought about by the survey method, it is often easier to find statistically significant results than other data gathering methods.

Statistical surveys have disadvantages too. Questions that bear controversies may not be precisely answered by the participants. The general questions may not be as appropriate for all the participants as they should be.

## 3.2 TEST USED

### CHI-SQUARE TEST

It is a statistical hypothesis test that is used to compare the observed results with expected results. It is used to find whether there is a relationship between two variables. The two commonly used Chi-square tests are the Chi-square goodness of fit test and the Chi-square test of independence. Chi-square goodness of fit is used when you have a single measurement variable and Chi-square test of independence is used when you have two measurement variables. Here, we use the chi-square test of independence.

We take the null hypothesis,  $H_0$ : There is no relation and the alternate hypothesis as  $H_1$ : There is a relation. Degrees of freedom =  $(d_1 - 1)(d_2 - 1)$ ; where  $d_1$  is the category for the first variable and  $d_2$  is the category for the second variable. To get a more reliable result, we take  $\alpha = 0.05$  (5%). The P-value of 0.05 or greater is considered critical, anything less means the deviations are significant and the null hypothesis being tested must be rejected. The  $\chi^2$  test is a one-sided test because we never have negative values of  $\chi^2$ .

For  $\chi^2$ , the sum of the difference of observed and expected squared is divided by the expected, thus chi-square is always a positive number or it may be close to zero. The theoretical value depends on both the alpha value and the degrees of freedom of our data.

Let  $E_i$  be the expected frequency of the  $i$ th class and  $O_i$  be the observed frequency of the  $i$ th class, then

$$\chi^2 = \sum_{i=1}^k (O_i - E_i)^2 / E_i$$

follows the  $\chi^2$  distribution with  $k-r-1$  degrees of freedom where  $k$  is the number of classes and  $r$  is the number of independent constraints to be satisfied by the frequencies. Chi-square test often refers to tests for which the distribution of the test statistic approaches the  $\chi^2$  distribution asymptotically, meaning that the sampling distribution of the test statistic approximates a chi-square distribution more and more closely as sample size increases.

## Chapter 4

# DATA ANALYSIS

---

### 4.1 GRAPHICAL ANALYSIS

We have collected a total of 332 samples out of which 261 is from Ernakulam district. As we are only concentrating on Ernakulam district, our sample size is 261.

In 261, 54 were males and 207 were females. About 90% of the respondents belong to the age group 18-25. From urban 162 responded and from rural 99 responded.

Out of the 261 samples collected, 231(89%) have purchased clothes online at least once and the remaining 30(11%) have never purchased clothes online.

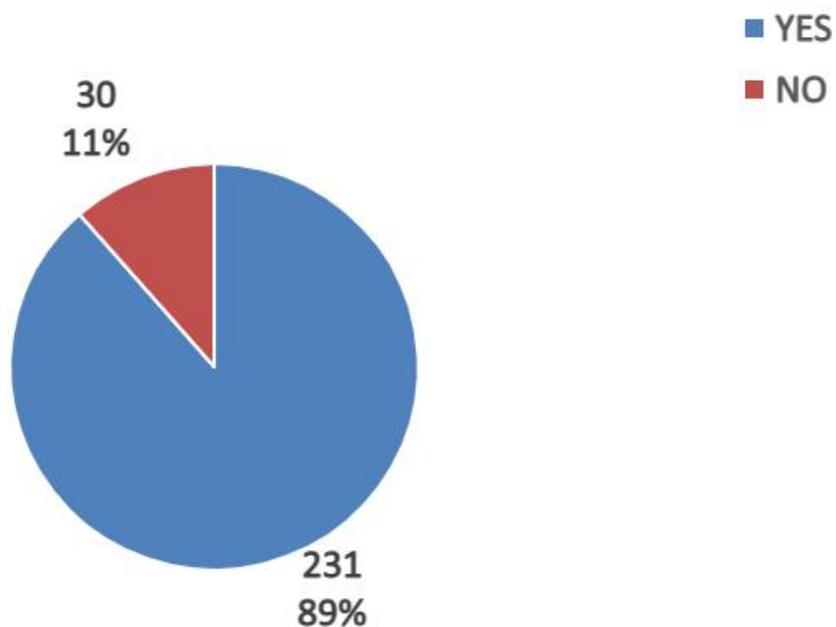


Figure 4.1: Status of online purchase

One of the biggest concerns among people who have never purchased online is that the expected product is not received. Some other concerns include online transactions are riskier, risk of identity theft and lack of knowledge about online shopping.

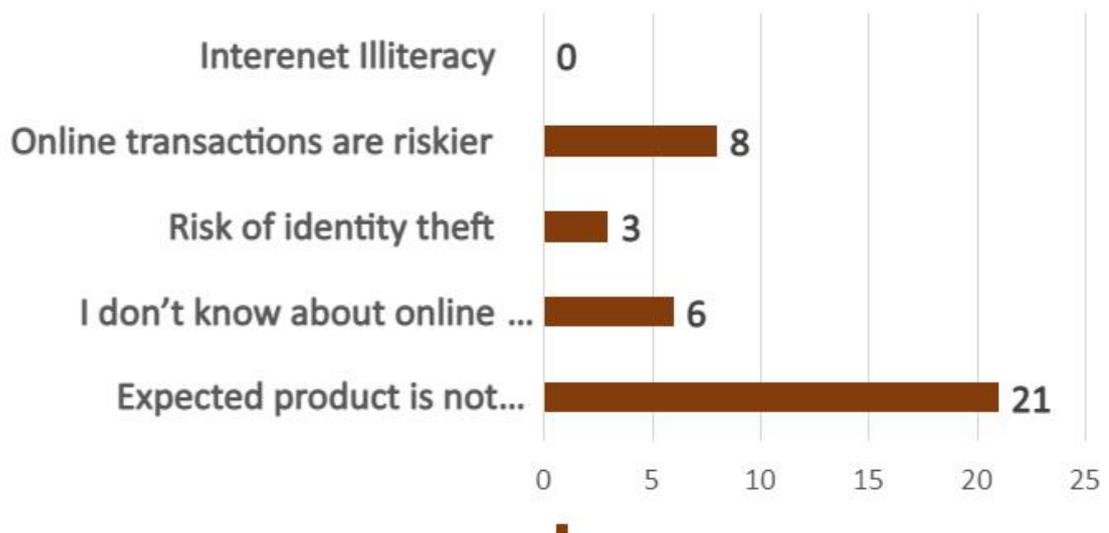


Figure 4.2: Reasons for not participating in online shopping

From the 231 responses on the frequency of online shopping, 163 (71%) people shop online once in a few months, 28 (12%) shop online once in a month, 25 (11%) twice or thrice in a month and 15 (6%), more than thrice in a month.

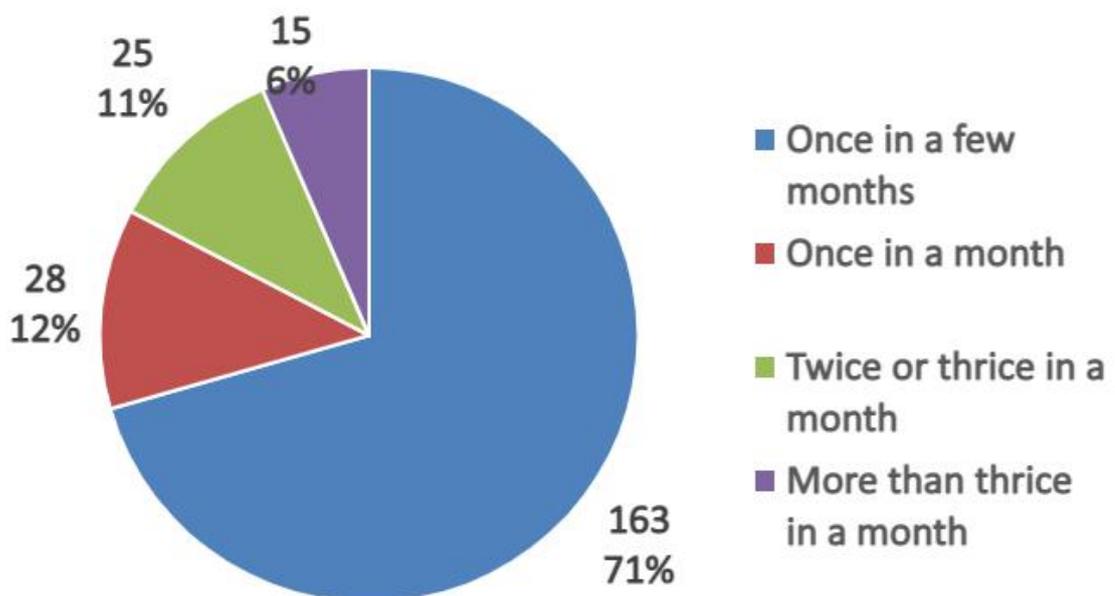


Figure 4.3: Frequency of online shopping

From the 231 responses on the frequency of people purchasing in person, 151 (65%) people shop once in a few months, 45 (19%) shop once in a month, 20 (9%) twice or thrice in a month and 15 (7%), more than thrice in a month.

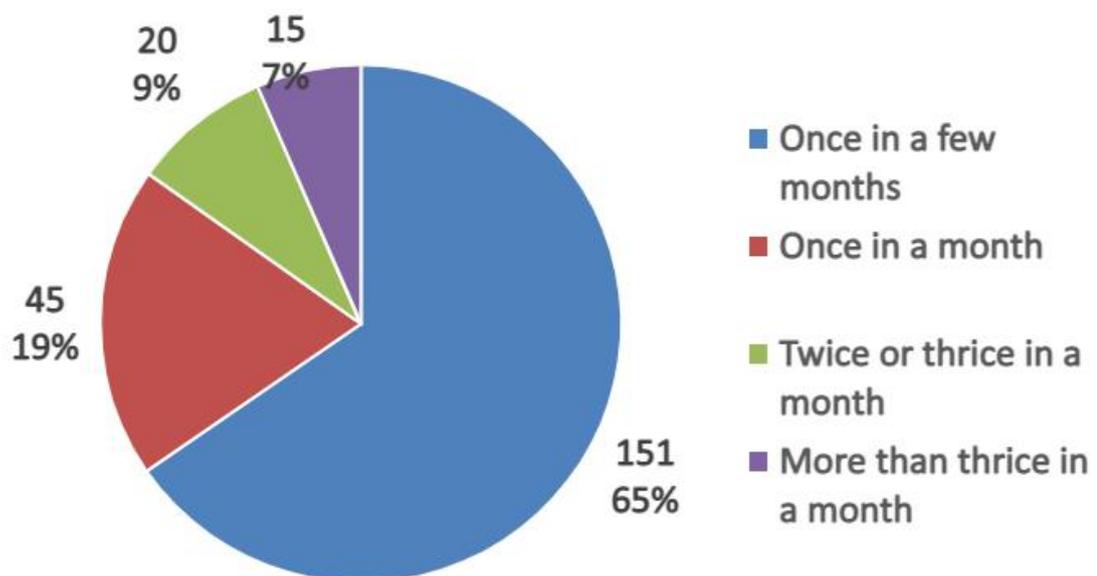


Figure 4.4: Frequency of in-person shopping

153 people out of 231 would rate their overall online shopping experience as average, 72 rate their experience as excellent and the remaining 6, as poor.

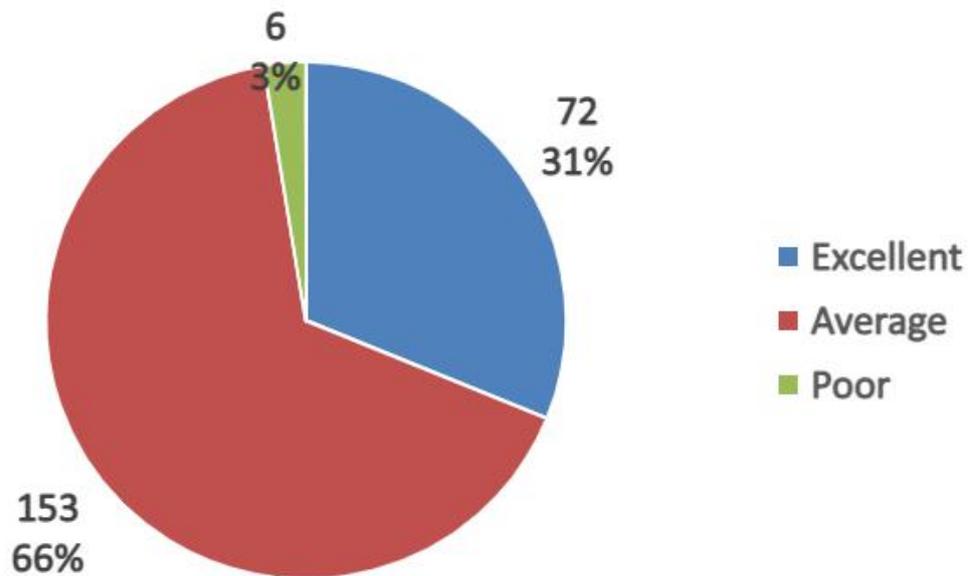


Figure 4.5: Rating - Online shopping

Some of the main reasons for people to prefer online are a wide range of choices (29%), time-saving (24%), Convenience and flexibility(14%), easy returns(14%), discounts/ offers(9%), and no contact home delivery(9%). There are other(1%) reasons as well.

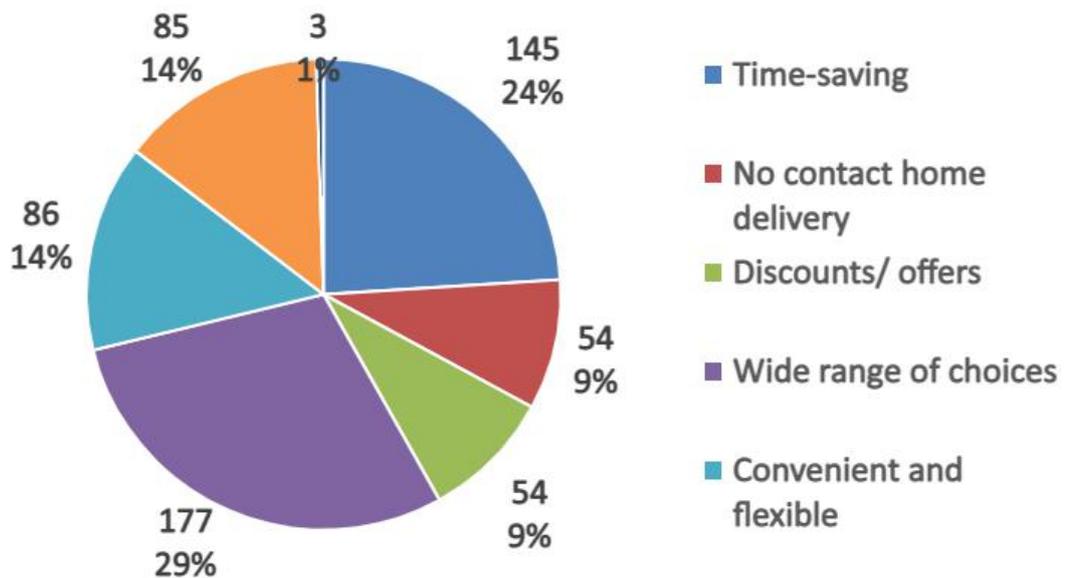


Figure 4.6: Advantages of online shopping

Among the concerns people have in online shopping, the prominent one is poor quality (50%). Other concerns include breach of personal information/ payment details (39%), poor internet connection (8%) and others (3%).

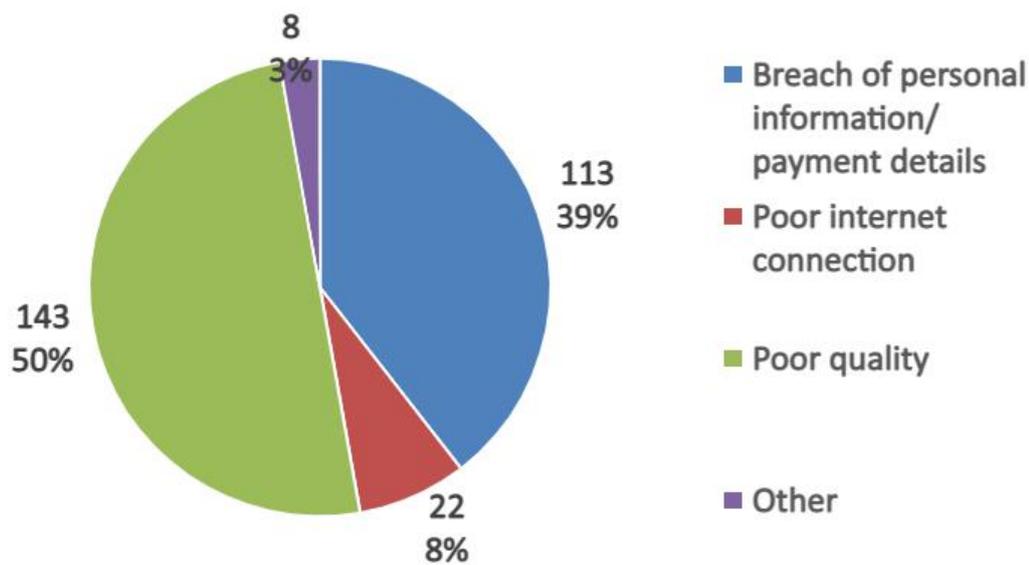


Figure 4.7: Concerns in online shopping

Out of the 231 people who responded to the impact of Covid 19 on online shopping, 188 (81%) answered that their online shopping increased during Covid 19 and the remaining 43 (19%) did not find any increase.

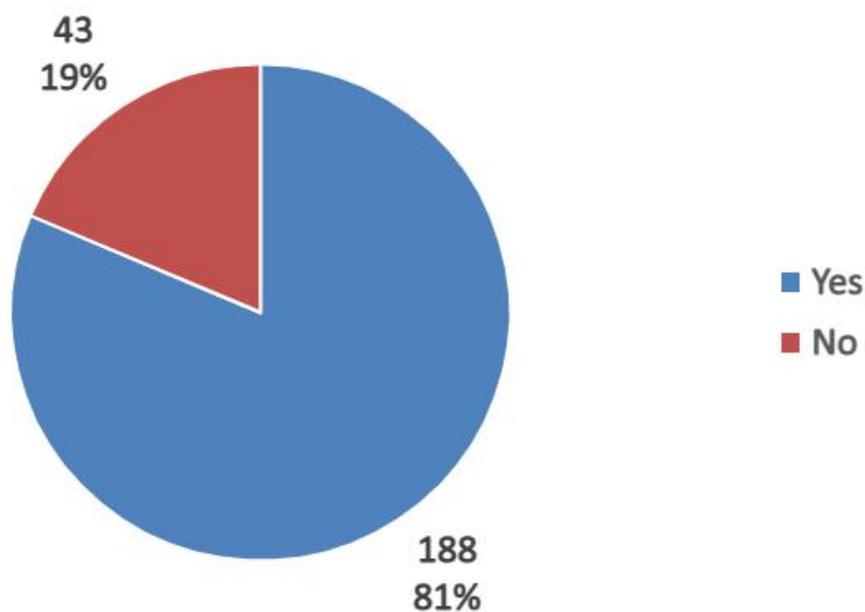


Figure 4.8: Covid 19 and Online shopping

Of the 188 people who responded that their online shopping increased during Covid, only 12 admitted to have done their entire shopping on-line, 90 says more than half of their purchases was done online and 86 says that only limited purchases were increased online.

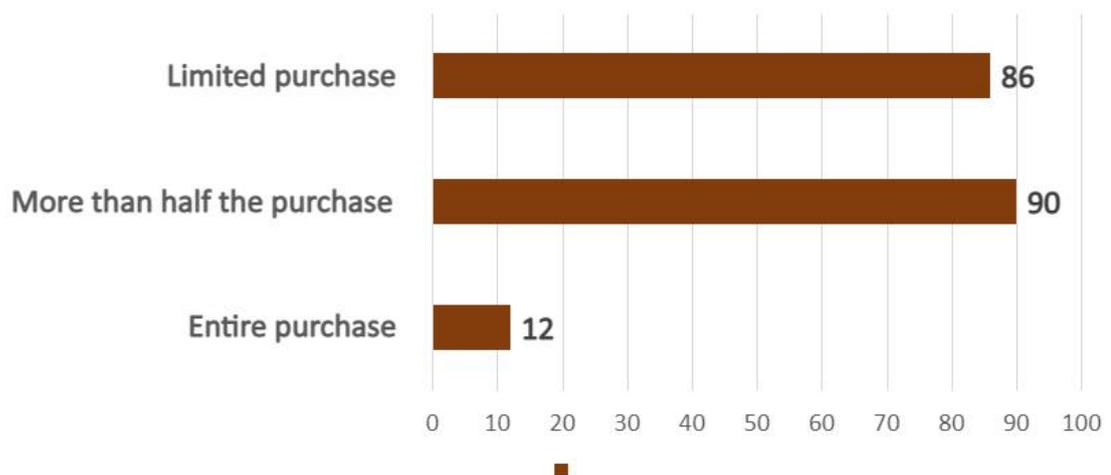


Figure 4.9: Impact of Covid 19 on online shopping

Regarding the amount spent on online shopping every month, 133 spend money below 1,000/-, 68 spend below 2,000/- and the remaining 30 spend more than 2,000/- on online shopping on a monthly basis.

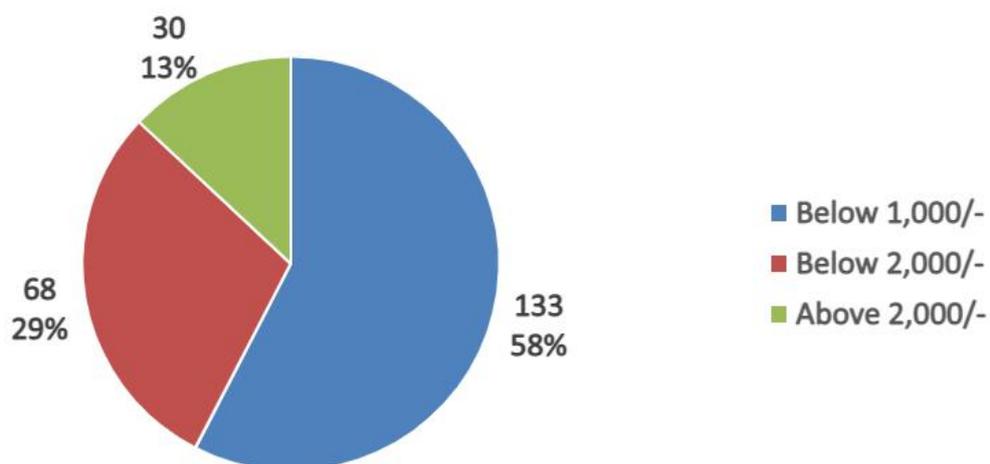


Figure 4.10: Monthly purchase rate

To the query about the discounts/offers in online shopping, almost 88% of the people responded that discounts play a role in their online shopping and 12% responded otherwise.

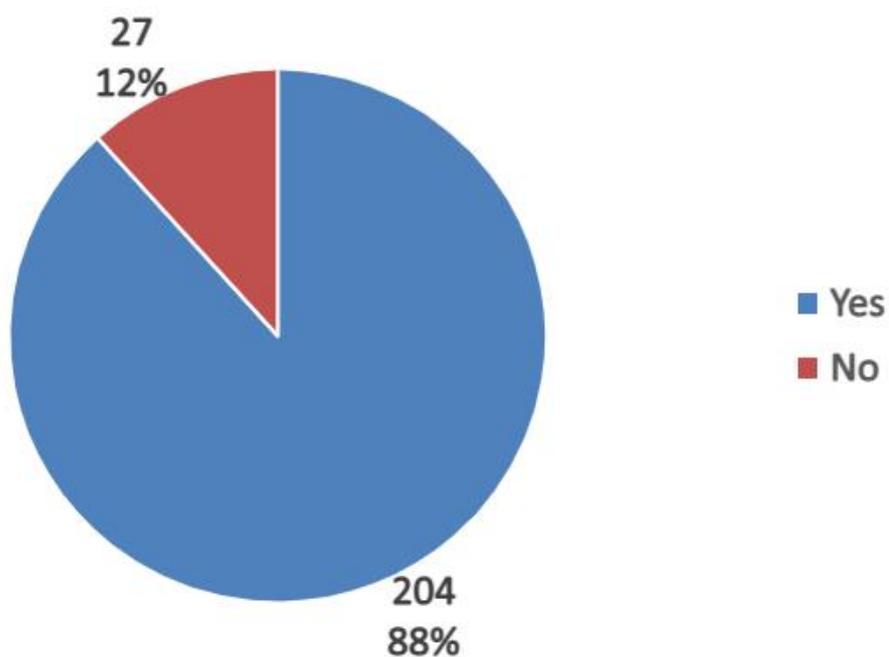


Figure 4.11: Discounts and Online shopping

Out of the 204 people who responded that discounts play a role in their online shopping, 116 look for discounts only if they want to buy something and 88 buy clothes whenever there is a discount.

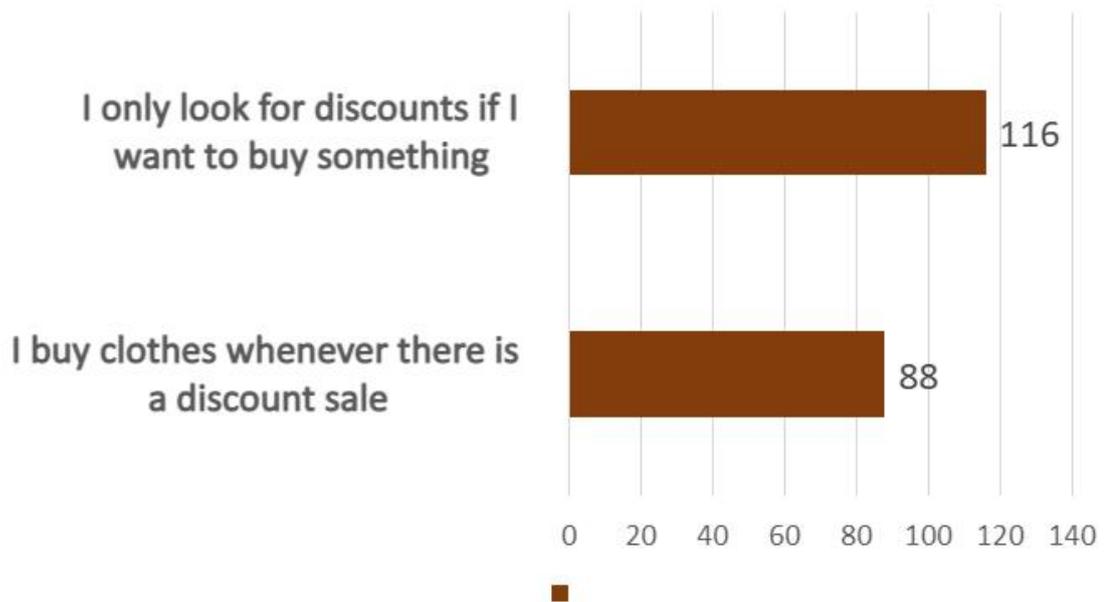


Figure 4.12: Role of discounts in online shopping

About the ease in finding products online, 40% of people responded that it was easy for them to find what they were looking for, 13% as very easy, 44% responded to it as neutral, and 3% as difficult.

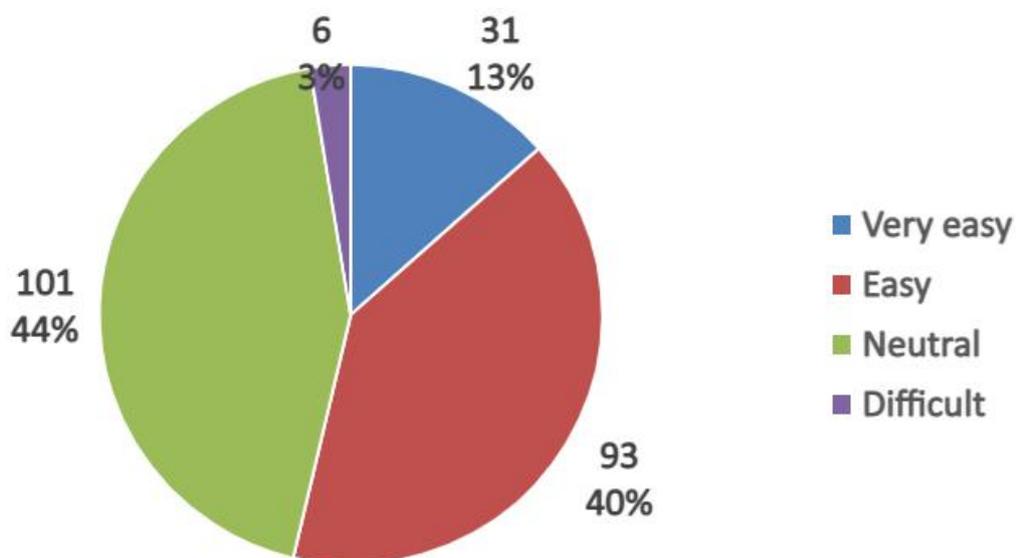


Figure 4.13: Finding products online

Among the responses on how likely they are to recommend online shopping to their friend (1 being the least and 5 being the most), 48% are more likely to recommend, only a few responded that they are least likely to recommend.

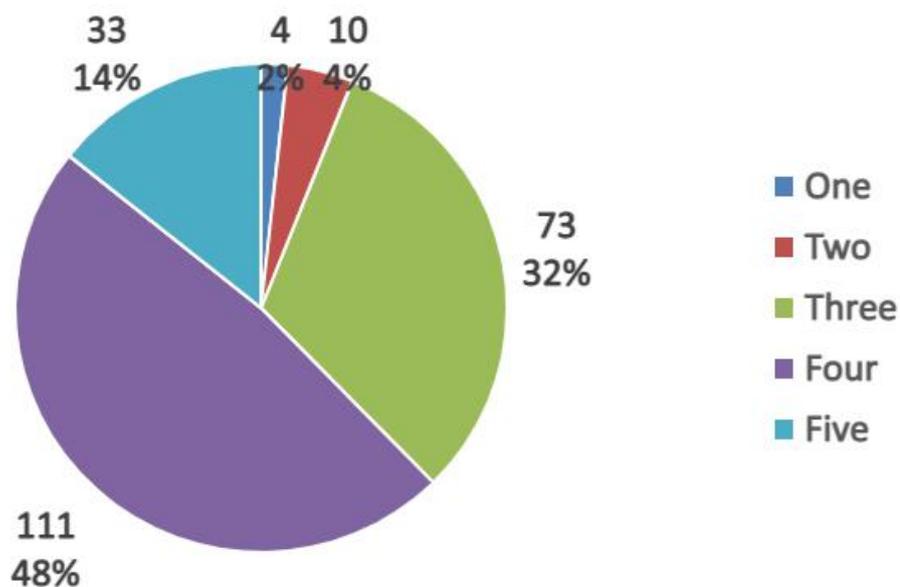


Figure 4.14: Recommending online shopping

## 4.2 STATISTICAL ANALYSIS

### 4.2.1 EFFECT OF FACTORS ON THE PREFERENCE OF SHOPPING

As we are doing the chi-square test of independence, two measurement variables are needed. So we are considering demographic variables like age, gender, place of residence and financial variables like monthly income as one variable and response to the question 'Have you ever done online shopping?' as the other.

#### COMPARISON BETWEEN PREFERENCE AND GENDER

$H_0$ : There is no relation between preference and gender.

$H_1$ : There is relation between preference and gender.

PREFERENCE	MALE	FEMALE	TOTAL
YES	47	184	231
NO	7	23	30
TOTAL	54	207	261

Figure 1: Preference vs Gender - Observed frequency

PREFERENCE	MALE	FEMALE	TOTAL
YES	47.79310345	183.206896	231
NO	6.206896552	23.7931034	30
TOTAL	54	207	261

Figure 2: Preference vs Gender - Expected frequency

$$p\text{-value} = 0.703972477$$

Since the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between preference and gender. In this 21st century, everyone has a smartphone and access to an internet connection irrespective of their gender. Hence, the preference for shopping is independent of gender.

## COMPARISON BETWEEN PREFERENCE AND RESIDENCE

$H_0$ : There is no relation between preference and residence.

$H_1$ : There is relation between preference and residence.

PREFERENCE	URBAN	RURAL	TOTAL
YES	148	83	231
NO	14	16	30
TOTAL	162	99	261

Figure 3: Preference vs Residence - Observed frequency

PREFERENCE	URBAN	RURAL	TOTAL
YES	143.3793103	87.62068966	231
NO	18.62068966	11.37931034	30
TOTAL	162	99	261

Figure 4: Preference vs Residence - Expected frequency

$$p\text{-value} = 0.0646$$

Since the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between preference and residence. Due to the COVID crisis, people living in both rural and urban areas had no other choice but to shop online.

## COMPARISON BETWEEN PREFERENCE AND AGE

$H_0$ : There is no relation between preference and age.

$H_1$ : There is relation between preference and age.

AGE GROUP	YES	NO	TOTAL
15-19	46	6	52
20-25	163	20	183
26-65	22	4	26
TOTAL	231	30	261

Figure 5: Preference vs Age - Observed frequency

AGE GROUP	YES	NO	TOTAL
15-19	46.02298851	5.977011494	52
20-25	161.9655172	21.03448276	183
26-65	23.01149425	2.988505747	26
TOTAL	231	30	261

Figure 6: Preference vs Age - Expected frequency

$$p\text{-value} = 0.800756564$$

Here, the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between preference and age. People of all age groups started shopping online as they were forced to stay home due to the lockdown.

**COMPARISON BETWEEN PREFERENCE AND INCOME**

$H_0$ : There is no relation between preference and income.

$H_1$ : There is relation between preference and income.

INCOME	YES	NO	TOTAL
BELOW 10,000	88	15	103
10,000-50,000	96	9	105
50,000-1,00,000	32	3	35
ABOVE 1,00,000	15	3	18
<b>TOTAL</b>	<b>231</b>	<b>30</b>	<b>261</b>

Figure 7: Preference vs Income - Observed frequency

INCOME	YES	NO	TOTAL
BELOW 10,000	91.16091954	11.83908046	103
10,000-50,000	92.93103448	12.06896552	105
50,000-1,00,000	30.97701149	4.022988506	35
ABOVE 1,00,000	15.93103448	2.068965517	18
<b>TOTAL</b>	<b>231</b>	<b>30</b>	<b>261</b>

Figure 8: Preference vs Income - Expected frequency

$$p\text{-value} = 0.45703882$$

As we can see that the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between the preference and income of people. People belonging to all income groups have increased their purchases online because of Covid 19. So there is no relation between a particular income group and online shopping.

#### 4.2.2 RELATION BETWEEN MONTHLY INCOME AND MONTHLY PURCHASE RATE

Here the variables we use are financial variables - monthly income and response to the question 'How much do you spend on online shopping every month?'.

### COMPARISON BETWEEN MONTHLY INCOME AND MONTHLY PURCHASE RATE

$H_0$ : There is no relation between monthly income and monthly purchase rate.

$H_1$ : There is relation between monthly income and monthly purchase rate.

INCOME	BELOW 1,000	BELOW 2,000	ABOVE 2,000	TOTAL
BELOW 10,000	55	25	8	88
10,000 - 50,000	59	26	11	96
50,000 - 1,00,000	15	10	7	32
ABOVE 1,00,000	4	7	4	15
<b>TOTAL</b>	<b>133</b>	<b>68</b>	<b>30</b>	<b>231</b>

Figure 9: Purchase Rate vs Income - Observed frequency

INCOME	BELOW 1,000	BELOW 2,000	ABOVE 2,000	TOTAL
BELOW 10,000	50.66666667	25.9047619	11.42857143	88
10,000 - 50,000	55.27272727	28.25974026	12.46753247	96
50,000 - 1,00,000	18.42424242	9.41991342	4.155844156	32
ABOVE 1,00,000	8.636363636	4.415584416	1.948051948	15
<b>TOTAL</b>	<b>133</b>	<b>68</b>	<b>30</b>	<b>231</b>

Figure 10: Purchase Rate vs Income - Expected frequency

$$p\text{-value} = 0.094193177$$

As the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between monthly income and monthly purchase rate. Even people with high income do not prefer to shop for clothes above 2000/- online, maybe due to poor quality and other concerns faced in online shopping.

### 4.2.3 EFFECT OF FACTORS ON THE ROLE OF DISCOUNT

Here we consider the demographic variables like age, gender, place of residence and financial variable - monthly income as well as the response to the question 'Do discounts/offers play a role in your online shopping?'

#### COMPARISON BETWEEN DISCOUNT AND RESIDENCE

$H_0$ : There is no relation between discount and residence.

$H_1$ : There is relation between discount and residence.

DISCOUNT	URBAN	RURAL	TOTAL
YES	131	73	204
NO	17	10	27
TOTAL	148	83	231

Figure 11: Discount vs Residence - Observed frequency

DISCOUNT	URBAN	RURAL	TOTAL
YES	130.7012987	73.2987013	204
NO	17.2987013	9.701298701	27
TOTAL	148	83	231

Figure 12: Discount vs Residence - Expected frequency

$$p\text{-value} = 0.898549779$$

Since the p-value is greater than 0.05, we accept  $H_0$  and hence there is no relation between the role of discount and place of residence. Due to the sudden upsurge of COVID 19, everybody looks for discounts. Hence there is no particular relation between them.

**COMPARISON BETWEEN DISCOUNT AND INCOME**

$H_0$ : There is no relation between discount and income.

$H_1$ : There is relation between discount and income.

INCOME	YES	NO	TOTAL
BELOW 10,000	79	9	88
10,000-50,000	88	8	96
50,000-1,00,000	25	7	32
ABOVE 1,00,000	12	3	15
<b>TOTAL</b>	<b>204</b>	<b>27</b>	<b>231</b>

Figure 13: Discount vs Income - Observed frequency

INCOME	YES	NO	TOTAL
BELOW 10,000	77.71428571	10.28571429	88
10,000-50,000	84.77922078	11.22077922	96
50,000-1,00,000	28.25974026	3.74025974	32
ABOVE 1,00,000	13.24675325	1.753246753	15
<b>TOTAL</b>	<b>204</b>	<b>27</b>	<b>231</b>

Figure 14: Discount vs Income - Expected frequency

$$p\text{-value} = 0.14167761$$

Since the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between the role of discount and monthly income. In order to save money, people look for discounts, especially in online shopping.

**COMPARISON BETWEEN DISCOUNT AND GENDER**

$H_0$ : There is no relation between discount and gender.

$H_1$ : There is relation between discount and gender.

DISCOUNT	MALE	FEMALE	TOTAL
YES	39	165	204
NO	8	19	27
TOTAL	47	184	231

Figure 15: Discount vs Gender - Observed frequency

DISCOUNT	MALE	FEMALE	TOTAL
YES	41.50649351	162.4935065	204
NO	5.493506494	21.50649351	27
TOTAL	47	184	231

Figure 16: Discount vs Gender - Expected frequency

$$p\text{-value} = 0.202288914$$

Here the p-value is greater than 0.05, we accept  $H_0$  and there is no relation between the role of discount and Gender. Both males and females prefer discounts while doing online shopping. So we could clearly say that there is no particular relation between discount and gender.

**COMPARISON BETWEEN DISCOUNT AND AGE**

$H_0$ : There is no relation between discount and age.

$H_1$ : There is relation between discount and age.

AGE GROUP	YES	NO	TOTAL
15-19	40	6	46
20-25	147	16	163
26-65	17	5	22
<b>TOTAL</b>	<b>204</b>	<b>27</b>	<b>231</b>

Figure 17: Discount vs Age - Observed frequency

AGE GROUP	YES	NO	TOTAL
15-19	40.62337662	5.376623377	46
20-25	143.9480519	19.05194805	163
26-65	19.42857143	2.571428571	22
<b>TOTAL</b>	<b>204</b>	<b>27</b>	<b>231</b>

Figure 18: Discount vs Age - Expected frequency

$$p\text{-value} = 0.198625881$$

As we can see the p-value is greater than 0.05, there is no relation between the role of discount and age. Most people of all age groups look for discounts in order to save money and to purchase more clothes.

## Chapter 5

# CONCLUSION

---

Online shopping offers many advantages to the consumers like a wide range of products, customer feedback and ratings, shopping in the comfort of our home etc. Those were the main reasons for an increase in online shopping. But in the current scenario, social distancing plays a major role to avoid the virus, so online shopping is the only safest option one can opt for. It also has some other added benefits like discounts and offers.

Through this study we couldn't find any variables particularly responsible for the behaviour pattern of the consumers. Nevertheless, the online shopping of 81% respondents from Ernakulam district has increased during this pandemic. Also 88% of the respondents say that discount is an important factor for them to shop online.

Age group doesn't play a role in people's choice to shop online. People tend to shop online if their technical knowledge is above average. In this era many of the older generation as well as the middle aged people are as good as the younger generation in terms of technology. So it is not the age group but the technical knowledge of people that make them shop online. Moreover COVID created a situation where online shopping is the safe way to buy things especially for the older generation.

Nowadays everybody wants to dress up and look stylish no matter their gender. Everyone is on their own journey to create new styles. Following the trends and fashion is one of the ways to achieve it. For those people, online stores offer a variety of clothes at a comparatively affordable price. Also one of the biggest advantages of online shopping over in-person shopping is that it has all the sizes from xs to xxxl.

Another factor that doesn't influence people's decisions to go for online shopping is the monthly income. Online we can buy clothes of all price ranges. So people from all income groups can participate in online shopping evenly.

In this peak time, people from all places of residence would go for online shopping as it is more convenient. People's knowledge on technology and access to internet connection does not depend on where they live (urban or rural).

One of the biggest concerns people have on online shopping is the quality of the product received. Even if it is your 50th time doing online shopping, the quality of the clothes you buy might not be predictable. So the majority of the respondents from all income groups buy clothes under 2,000 rupees so that a large amount of money won't be at risk. This would lead that the monthly income and monthly purchase rate have no particular relation.

Discount is something we all look for irrespective of our age, gender, place of residence and financial condition. One of the major motives behind it is to save money. But in most cases people don't need a reason to look for discounts. The thought that we bought something for a price lower than its MRP would give us satisfaction. This satisfaction is a universal feeling

So in general, though it has some drawbacks, online shopping can be considered as a reliable and efficient way to buy clothes even when COVID 19 and such pandemics are wiped out completely.

## REFERENCES

---

- [1] Ramakrishna Pillai's **TEXTBOOK OF COMPLEMENTARY STATISTICS** for Semester 1 and 2 revised by Lalitha R Pillai
- [2] [https://www.researchgate.net/publication/353057164\\_Online\\_Shopping\\_Behavior\\_during\\_COVID-19\\_Pandemic\\_An\\_Indian\\_Perspective](https://www.researchgate.net/publication/353057164_Online_Shopping_Behavior_during_COVID-19_Pandemic_An_Indian_Perspective)
- [3] <https://www.mdpi.com/0718-1876/16/6/125/htm>
- [4] [https://www.jmp.com/en\\_us/statistics-knowledge-portal/chi-square-test.html](https://www.jmp.com/en_us/statistics-knowledge-portal/chi-square-test.html)
- [5] [https://drive.google.com/file/d/18wl\\_gFuhj\\_2PFefhofVRXctKedU8\\_eza/view?usp=drivesdk](https://drive.google.com/file/d/18wl_gFuhj_2PFefhofVRXctKedU8_eza/view?usp=drivesdk)

# ANNEXURE

---

## Questionnaire

1. Age:
2. Gender:
3. Place of residence
  - Urban
  - Rural
4. Monthly income:
  - Below 10,000/-
  - 10,000/- or above till 50,000/-
  - 50,000/- or above till 1,00,000/-
  - Above 1,00,000/-
5. Have you ever purchased clothes online?
  - Yes
  - No
6. If no, what are the reasons?
  - I don't know about online shopping
  - Online transactions are riskier
  - Internet illiteracy
  - Expected product is not received
  - Risk of identity theft
7. How often do you shop online?
  - Once in a few months
  - Once in a month
  - Twice or thrice in a month
  - More than thrice in a month

8. How often do you purchase in-person?

- Once in a few months
- Once in a month
- Twice or thrice in a month
- More than thrice in a month

9. How would you rate your overall online shopping experience?

- Poor
- Average
- Excellent

10. Why do you prefer online shopping?

- Time-saving
- No contact home delivery
- Discounts/ offers
- Wide range of choices (collection, rate, size, colour, etc.)
- Convenient and flexible
- Easy returns
- Other

11. What are your biggest concerns regarding online shopping?

- Breach of personal information/ payment details
- Poor internet connection
- Poor quality
- Other

12. Have your online shopping increased during Covid 19?

- Yes
- No

13. If yes, to what extent?

- Entire purchase
- More than half the purchase
- Limited purchase

14. How much do you spend on online shopping every month?
- Entire purchase
  - More than half the purchase
  - Limited purchase
15. Do discounts/offers play a role in your online shopping?
- Yes
  - No
16. If yes, by how?
- I buy clothes whenever there is a discount sale
  - I only look for discounts if I want to buy something
17. How easy was it for you to find what you were looking for online?
- Very easy
  - Easy
  - Neutral
  - Difficult
18. On a scale of 1 to 5, how likely are you to recommend online shopping to your friend? (1 being the least and 5 being the most)
- One
  - Two
  - Three
  - Four
  - Five

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM



CERTIFICATE

This is to certify that the dissertation entitled, **STUDY ON ONLINE CLOTH SHOPPING BEFORE AND DURING COVID 19** is a bonafide record of the work done by Ms. **EVANA T D** under my guidance as partial fulfilment 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

**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:

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 Reshna 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: 04/03/2022



EVANA T D

AB19AMAT015

14. How much do you spend on online shopping every month?
- Entire purchase
  - More than half the purchase
  - Limited purchase
15. Do discounts/offers play a role in your online shopping?
- Yes
  - No
16. If yes, by how?
- I buy clothes whenever there is a discount sale
  - I only look for discounts if I want to buy something
17. How easy was it for you to find what you were looking for online?
- Very easy
  - Easy
  - Neutral
  - Difficult
18. On a scale of 1 to 5, how likely are you to recommend online shopping to your friend? (one being the least and five being the most)
- One
  - Two
  - Three
  - Four
  - Five



**ST. TERESA'S COLLEGE (AUTONOMOUS)  
AFFILIATED TO MAHATMA GANDHI UNIVERSITY**



**PROJECT REPORT  
ON  
TRAVELORE**

In partial fulfillment of the requirements for the award of the degree of

**BACHELOR OF SCIENCE IN  
COMPUTER APPLICATIONS  
[TRIPLE MAIN]**

**Submitted By**

**Evania Joseph**

**III B.Sc. Computer Applications [Triple Main]**

**Register No: SB19CA011**

**Under the guidance of**

**Mrs. Elizabeth Paul**

**DEPARTMENT OF COMPUTER APPLICATIONS**

**2019-2022**

**ST. TERESA'S COLLEGE (AUTONOMOUS)  
AFFILIATED TO MAHATMA GANDHI UNIVERSITY**



**PROJECT REPORT  
ON  
TRAVELORE**

In partial fulfillment of the requirements for the award of the degree of

**BACHELOR OF SCIENCE IN  
COMPUTER APPLICATIONS  
[TRIPLE MAIN]**

**Submitted By**

**Evania Joseph**

**III B.Sc. Computer Applications [Triple Main]**

**Register No: SB19CA011**

**Under the guidance of**

**Mrs. Elizabeth Paul**

**DEPARTMENT OF COMPUTER APPLICATIONS**

**2019-2022**



## **CERTIFICATE**

This is to certify that **Ms. EVANIA JOSEPH. (Reg. No: SB19CA011)** 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 "**TRAVELORE**" 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**

A handwritten signature in green ink, appearing to read "Deborah M.", is written over a horizontal line.

Chief Executive Officer





## CERTIFICATE

This is to certify that the project report entitled "TRAVELORE" a bona fide record of the work done by Evania Joseph during the year 2021-22 and submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Applications (Triple main) under Mahatma Gandhi University.

*Roh*

Head of the Department

*George*

Internal Examiner



*J G/A/22*

External Examiner

## DECLARATION

I, Evania Joseph, B.Sc. Computer Applications ( Triple Main ) final year student of St.Teresa's college (Autonomous) Ernakulam, Register No. SB19CA011, 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:

Place: Ernakulam

Evania Joseph

## ACKNOWLEDGEMENT

First and foremost we thank God Almighty for his blessings. We take this opportunity to express our gratitude to all those who helped me in completing this seminar successfully. We wish to express my sincere gratitude to the Manager Rev. Dr. Sr. Vinitha CSST, Director Rev. Sr Emeline CSST and the Principal Dr. Lizzy Mathew for providing all the facilities. We express my sincere gratitude towards the Head of the department Mrs. Raji S Pillai and the Course coordinator Mrs. Sheeba Emmanuel for the support. We deeply express my sincere thanks to my guide Mrs.Elizabeth for her proper guidance and support throughout the seminar work. We have indebted to my beloved teachers whose cooperation and suggestion throughout the project helped us a lot. We thank all my friends and classmates for their support. We convey my hearty thanks to my parents for the moral support, suggestion and encouragement.

Evania Joseph

## SYNOPSIS

Travelore Project intends to embrace the travel and tourism culture of Kerala by giving opportunities to various hotels and travel agencies to help in the running and progression of their business progression as well as making use of this to provide various exploring facilities and packages to our users. Named as one of the ten paradises of the world by National Geographic Traveler, Kerala is famous especially for its ecotourism initiatives and beautiful backwaters. Its unique culture and traditions, coupled with its varied demography, have made Kerala one of the most popular tourist destinations in the world. So users are continuously searching for a perfect package for their travel based on their interests and needs as well as many small/large scale hotels and travel agencies are searching for a perfect medium to connect to the customers. Hence this project aims to provide support such various hotels and travel agencies in their business as well as provide their various travelling and exploring facilities to the users based n their specified need and interest.

## Table of Contents

1.INTRODUCTION.....	1
1.1 About Project.....	2
1.2 About Organization.....	2
1.3 Objectives of project and organization.....	2
2.SYSTEM ANALYSIS.....	3
2.1 Introduction.....	4
2.2 Existing System.....	4
2.3 Proposed System.....	4
2.4 System Specification.....	4
2.5 Operating System.....	4
2.6 Languages and Software Packages.....	4
2.7 Hardware and Software Specifications.....	5
3.SYSTEM DESIGN.....	6
3.1 Introduction.....	7
3.2 Data Flow Diagram.....	7
3.3 Data Dictionary.....	10
3.4 Database Design.....	11
4. SYSTEM DEVELOPMENT.....	13
4.1 Introduction.....	14
4.2Process Description.....	14
4.3 Code Design.....	15
5. SYSTEM TESTING AND IMPLEMENTATION.....	25
5.1Introduction.....	26
5.2Implementation.....	26
5.3Debugging.....	26
5.4System Security.....	27
5.5Scope for Future Enhancement.....	27
6.CONCLUSION.....	28
7. APPENDIX.....	29
8. BIBLIOGRAPHY.....	39

# **1.INTRODUCTION**

## **1.1 About Project**

Travelore is a Kerala based travel and tourism app where hotels and travel agencies can register and provide their packages to the users and users can select and book them based on their needs and interests.

## **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 project and organization**

The main objective of the project is to embrace the culture and explore the scope of travel and tourism in Kerala by acting as a medium between various small/large scale hotels and travel agencies and users thereby giving the opportunity to the hotels and travel agencies in their business development as well as providing travelling and exploring facilities to the users based n their specified need and interest.

The main objective of the organization 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**

## 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 analyzing 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

To find a suitable tour package, often people have to search on many different sites. Even if they get to access all facilities under a single system it may not be trustworthy or having packages based on their budget and other specified needs. Also they may not provide emergency support to them. In case of direct booking there is no guarantee in getting a room or travel package suitable based on their needs and interests.

## 2.3 Proposed System

Travelore is an all in one application where users can search for a tour package among various small/large scale hotels and travel agencies which are trustworthy and safe since they are been verified and approved by the admin when registering. Since there are many options the users can more likely find a suitable package based on their budget and other needs. Also users need not to go and directly book the services rather than that, they can book hotels and travel agencies with an advance payment which ensures guaranty and clarity for their journey. Since it is an all in one application, users need not to search various sites for hotels and agencies.

## 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 system software that manages computer hardware, software resources, and provides common services for computer programs.

## 2.6 Languages and Software Packages

### Java

Java is a class-based, object-oriented programming language that is designed to have as few implementation dependencies as possible. It produces software for multiple platforms.

### Android

Android is a mobile operating system developed by Google, based on the Linux kernel and designed primarily for touch screen mobile devices such as smart phones 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 .

## **Python**

Python is a widely used high-level programming language for general-purpose programming. An interpreted language, Python has a design philosophy that emphasizes code 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 interpreters are available for many operating systems, allowing Python code to run on a wide variety of systems. Python is a multi-paradigm programming language: object-oriented programming and structured programming are fully supported.

## **SQL Database**

Microsoft SQL Server is a relational database management system developed by Microsoft. As a database server, it is a software product with the primary function of storing and retrieving data as requested by other software applications.

The scope of SQL includes data insert, query, update and delete, schema creation and modification, and data access control. SQL became a standard of the American National Standards Institute (ANSI), and of the International Organization for Standardization (ISO). Most SQL code is not completely portable among different database systems without adjustments.

## **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 behavior of a system to be developed and may include a set of use cases that describe interactions the users will have with the software.

Operating System: WINDOWS 8 or above for better performance

Front end: Python (For web application), Android (Mobile Application)

Back end: MYSQL

Software: SubLime Text, WAMP, Android Studio

Web Browser: Internet Explorer/Google Chrome/Firefox

Web Server: Apache

### **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.

Hard Disc: 320GB.

Display Type: PC Display.

Keyboard: PC/AT Enhanced PS/2Keyboard (110/10Key).

Mouse: First/Pilot Mouse Serial (c48).

Input Device: Mouse, keyboard

Output Device: Monitor, Mobile Display

## **3.SYSTEM DESIGN**

# Travelore

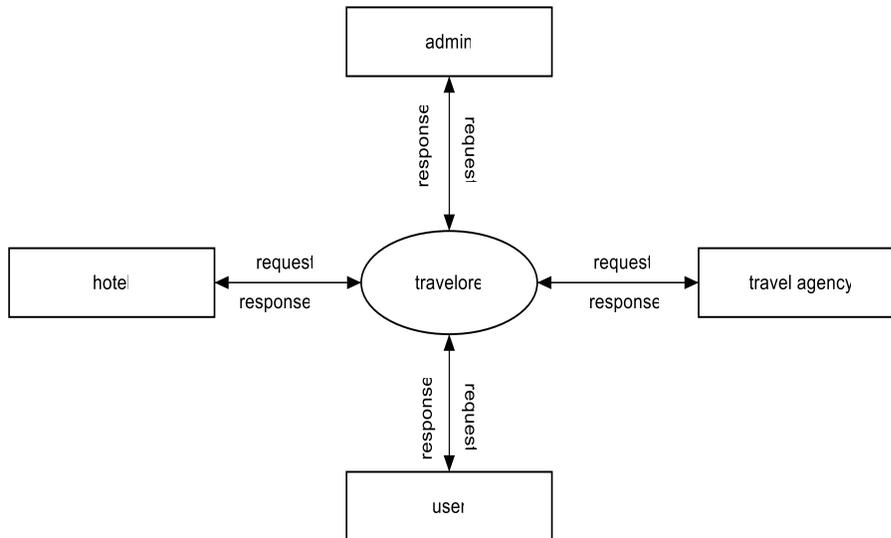
## 3.1 Introduction

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.

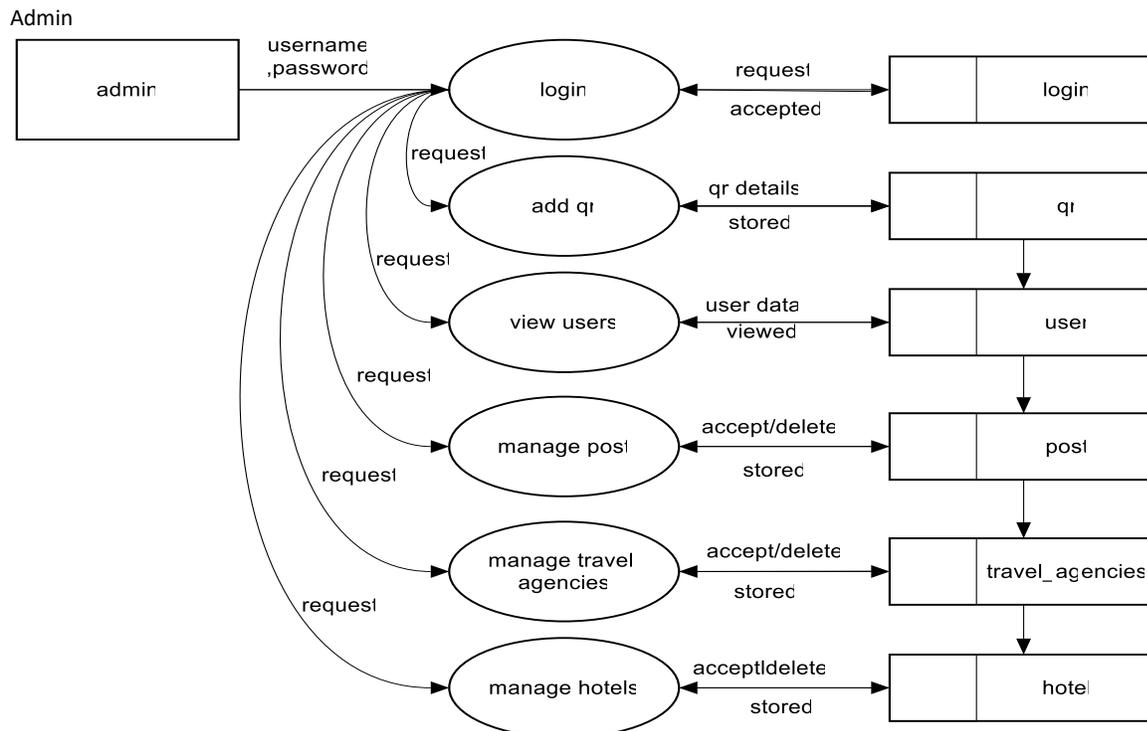
## 3.2 Data Flow Diagram

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

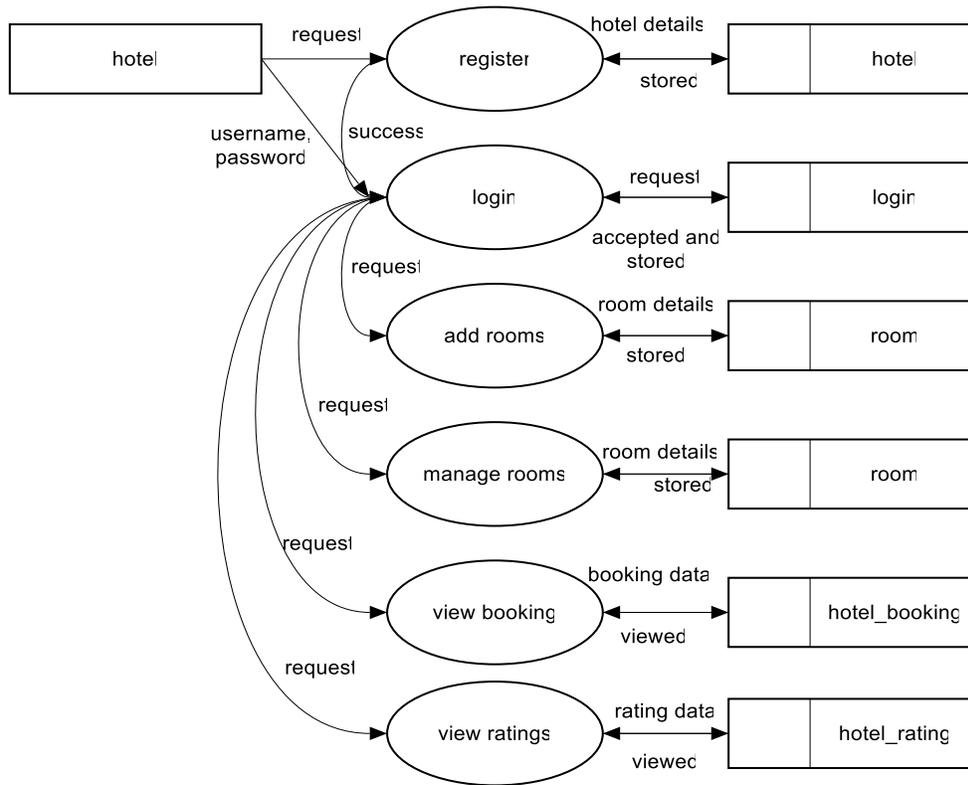


### DFD Level One

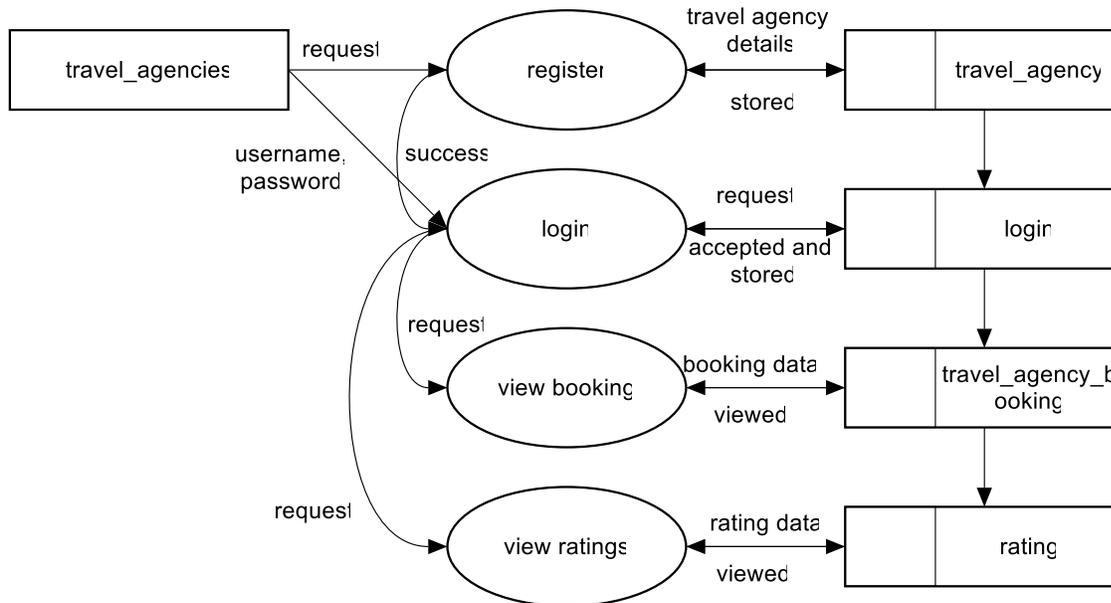


# Travelore

## Hotel

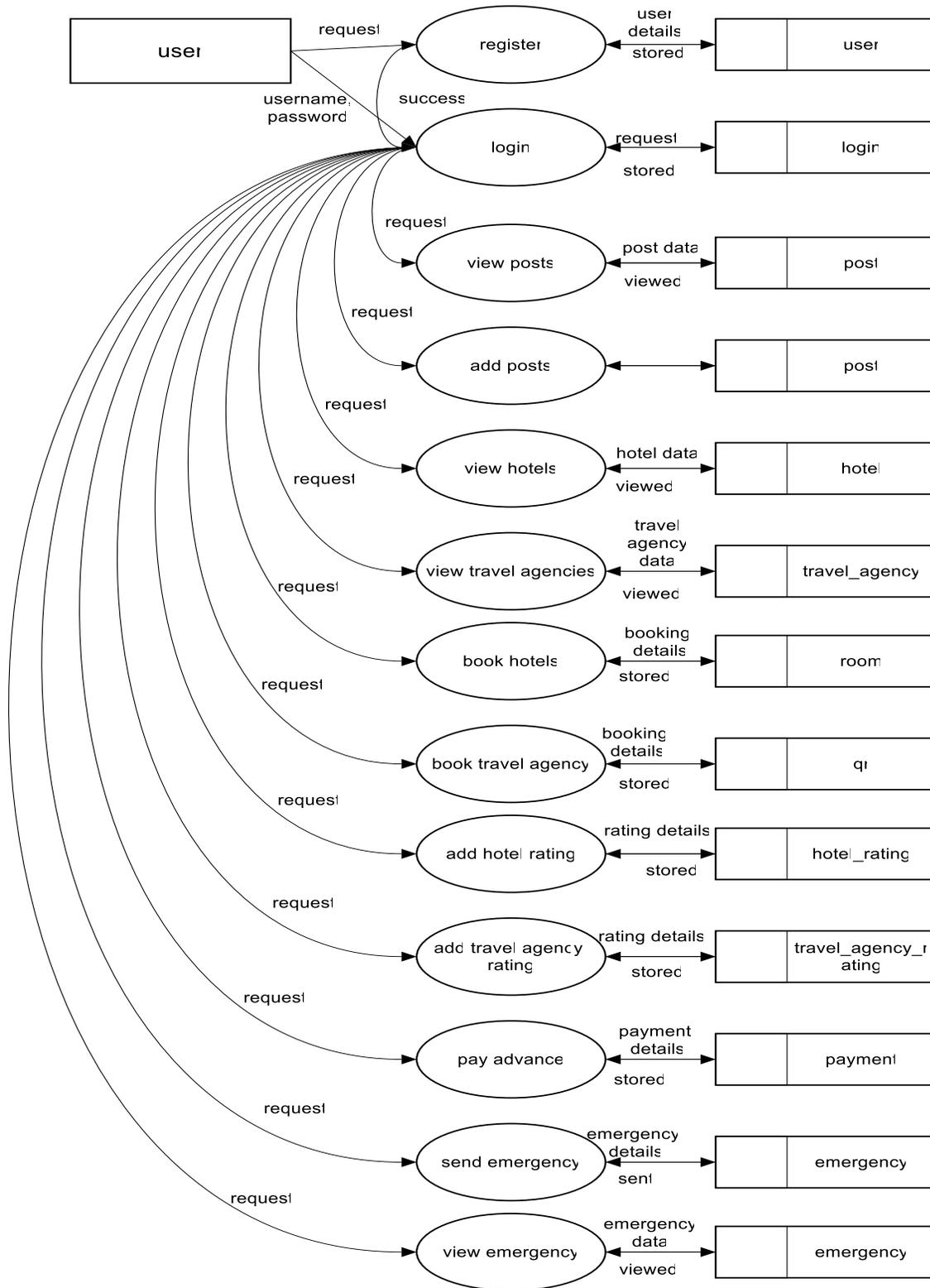


## Travel Agency



# Travelore

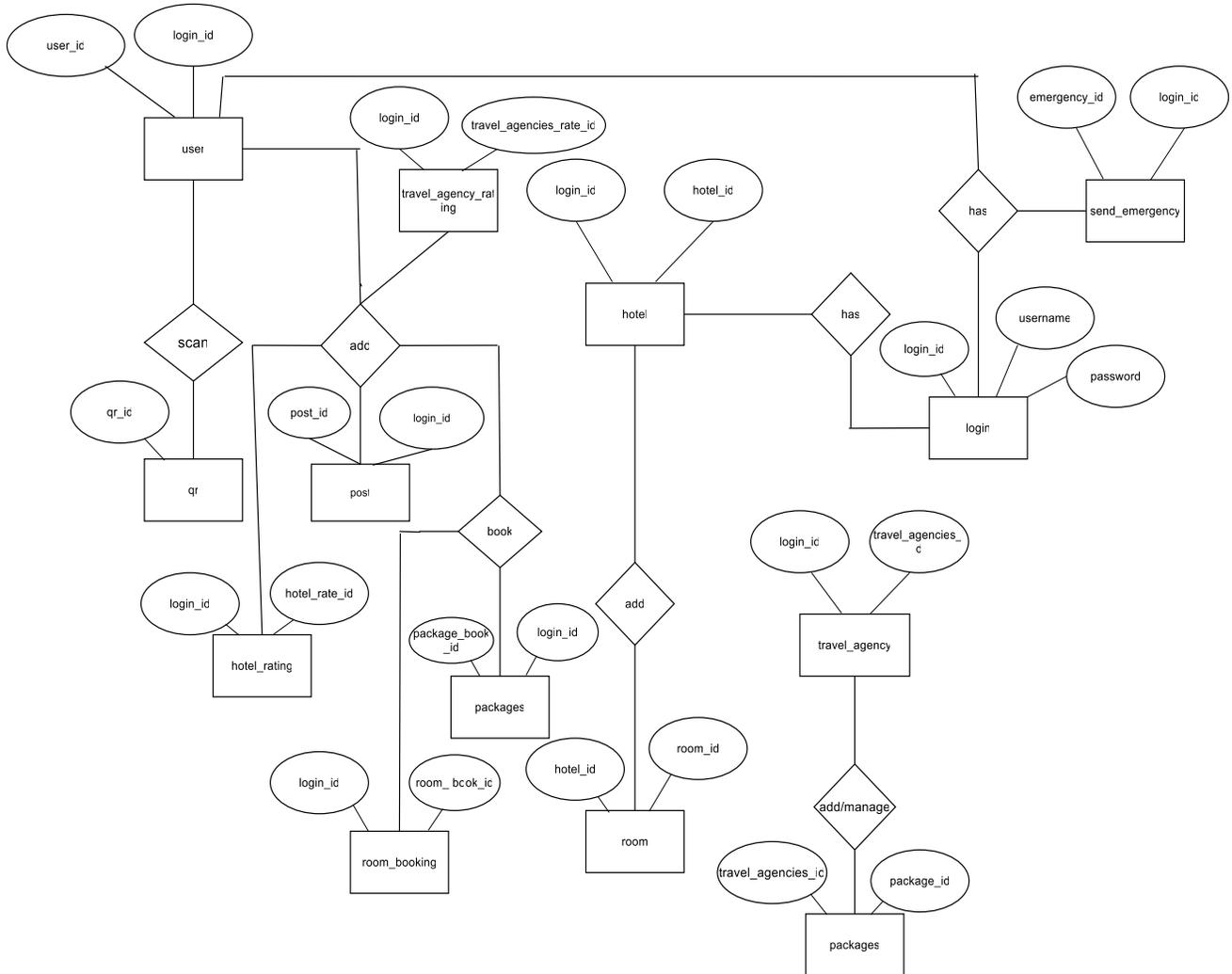
User



# Travelore

## E-R Diagram

Entity relationship diagrams provide a visual starting point for database design that can also be used to help determine information system requirements throughout an organization.



### 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.

# Travelore

## 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(100)	yes	
password	varchar(100)	yes	
user_type	varchar(100)	yes	

- hotel

Column	Type	Null	Default
hotel_id	int(11)	no	
login_id	int(11)	yes	
name	varchar(100)	yes	
place	varchar(100)	yes	
phone_no	varchar(100)	yes	
email	varchar(100)	yes	
latitude	varchar(100)	yes	
longitude	varchar(100)	yes	

- room\_booking

Column	Type	Null	Default
room_book_id	int(11)	no	
login_id	int(11)	yes	
room_id	int(11)	yes	
date	varchar(100)	yes	
booking_date	varchar(100)	yes	
no_of_days	varchar(100)	yes	
status	varchar(100)	yes	

- travelagency

Column	Type	Null	Default
travel_agencies_id	int(11)	no	
login_id	int(11)	yes	
name	varchar(100)	yes	
place	varchar(100)	yes	
phone_no	varchar(100)	yes	
email	varchar(100)	yes	
latitude	varchar(100)	yes	
longitude	varchar(100)	yes	

- package\_booking

Column	Type	Null	Default
package_book_id	int(11)	no	
login_id	int(11)	yes	
package_id	varchar(100)	yes	
date	varchar(100)	yes	
booking_date	varchar(100)	yes	
no_of_person	varchar(100)	yes	
status	varchar(100)	yes	

# Travelore

---

- user

Column	Type	Null	Default
user_id	int(11)	no	
login_id	int(11)	yes	
first_name	varchar(100)	yes	
last_name	varchar(100)	yes	
age	varchar(100)	yes	
gender	varchar(100)	yes	
Phone_no	varchar(100)	yes	
email	varchar(100)	yes	

- payment

Column	Type	Null	Default
payment_id	int(11)	no	
booking_id	int(11)	yes	
amount	varchar(100)	yes	
date	varchar(100)	yes	

- qr

Column	Type	Null	Default
qr_id	int(11)	no	
heading	varchar(100)	yes	
description	varchar(100)	yes	
image	varchar(500)	yes	

- post

Column	Type	Null	Default
post_id	int(11)	no	
user_id	int(11)	yes	
post_details	varchar(100)	yes	
image	varchar(500)	yes	
latitude	varchar(100)	yes	
longitude	varchar(100)	yes	
date	varchar(100)	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

Different processes of each module are as given below:

### i) Admin

#### 1. Admin Login

The admin must enter the web application using the existing username and password to log into their account in the login fragment.

#### 2. Add and remove QR

The admin can add or remove qr and its details in the web application.

#### 3. View Users

The admin can view registered users of the application.

#### 4. View and manage Posts

The admin can view and manage the posts posted by users.

#### 5. Manage hotels

The admin can accept or reject the hotels registered in the web application any time. The hotel can access the application only after when admin accept their registration.

#### 6. Manage travel agencies

The admin can accept or reject the travel agencies registered in the web application any time. The hotel can access the application only after when admin accept their registration.

### ii) Hotels

#### 1. Hotel Registration

Hotels can register in the application by filling the registration form with their required details. But they will get the access to the application only after when admin accepts their registration.

#### 2. Hotel Login

The admin must enter the web application using the existing username and password to log into their account in the login fragment.

#### 3. View and manage rooms

The hotels can view, add and manage their room details in the web application.

#### 4. View Booking

The hotels can view their bookings done by the users.

#### 5. View ratings

The hotels can view their ratings given by the users.

### iii) Travel Agencies

#### 1. Travel agency Registration

Hotels can register in the application by filling the registration form with their required details. But they will get the access to the application only after when admin accepts their registration.

## 2. Travel agency Login

The admin must enter the web application using the existing username and password to log into their account in the login fragment.

## 3. Add and manage Packages

Travel Agencies can add and manage their travel packages.

## 4. View Bookings

Travel Agencies can view their bookings done by the users.

## 5. View ratings

Travel agencies can view their ratings given by the users.

## iv) Users

### 1) Add post

Users can add posts publicly in the application.

### 2) View post

Users can view their as well as other users posts.

### 3) View and book hotels

Users can view the list of hotels and can book them by paying their respective advance amount.

### 4) View and book travel agencies

Users can view the list of travel agencies and their packages and can book them by paying their respective advance amount.

### 5) Scan QR

Users can scan qr codes and can access its description.

### 6) Send emergency messages

Users can send emergency messages in the application.

### 7) View emergency details

Users can view a summary of the emergency messages sent in the application.

## 4.3 Code Design

### Room Booking

```
package com.example.travelore;
```

```
import androidx.appcompat.app.AlertDialog;  
import androidx.appcompat.app.AppCompatActivity;
```

```
import android.app.DatePickerDialog;  
import android.content.DialogInterface;  
import android.content.Intent;  
import android.content.SharedPreferences;  
import android.os.Bundle;  
import android.preference.PreferenceManager;  
import android.text.InputType;  
import android.util.Log;  
import android.view.View;  
import android.widget.AdapterView;  
import android.widget.Button;  
import android.widget.DatePicker;
```

# Travelore

---

```
import android.widget.EditText;
import android.widget.Toast;

import org.json.JSONObject;

import java.text.SimpleDateFormat;
import java.util.Calendar;
import java.util.Locale;

public class user_book_room extends AppCompatActivity implements JsonResponse{
    EditText e1, e2;
    Button b1,b2;
    String booking_date, no_of_days;
    SharedPreferences sh;
    private DatePickerDialog fromDatePickerDialog;

    private SimpleDateFormat dateFormatter;

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.activity_user_book_room);
        sh = PreferenceManager.getDefaultSharedPreferences(getApplicationContext());
        e1 = (EditText) findViewById(R.id.et1);
        e2 = (EditText) findViewById(R.id.ed);
        b1 = (Button) findViewById(R.id.btn1);
        b2 = (Button) findViewById(R.id.btn2);
        e2.setInputType(InputType.TYPE_NULL);
        e2.requestFocus();
        setDateTimeField();
        dateFormatter = new SimpleDateFormat("dd-MM-yyyy", Locale.US);

        b1.setOnClickListener(new View.OnClickListener() {
            @Override
            public void onClick(View v) {
                no_of_days = e1.getText().toString();
                booking_date = e2.getText().toString();
                JsonRequest JR = new JsonRequest();
                JR.json_response = (JsonResponse) user_book_room.this;
                String q = "/user_book_room?lid=" + sh.getString("log_id", "") + "&rid=" + user_view_room.rids + "&booking_date=" +
                booking_date + "&no_of_days=" + no_of_days;
                q = q.replace(" ", "%20");
                JR.execute(q);
            }
        });
        b2.setOnClickListener(new View.OnClickListener() {
            @Override
            public void onClick(View view) {
                startActivity(new Intent(getApplicationContext(), user_pay_advance.class));
            }
        });
    }

    private void setDateTimeField() {
        e2.setOnClickListener(new View.OnClickListener() {
```

```
@Override
public void onClick(View arg0) {
    // TODO Auto-generated method stub
    fromDatePickerDialog.show();
}
});
Calendar newCalendar = Calendar.getInstance();
fromDatePickerDialog = new DatePickerDialog(this, new DatePickerDialog.OnDateSetListener() {

    public void onDateSet(DatePicker view, int year, int monthOfYear, int dayOfMonth) {
        // TODO Auto-generated method stub
        Calendar newDate = Calendar.getInstance();
        newDate.set(year, monthOfYear, dayOfMonth);
        e2.setText(dateFormatter.format(newDate.getTime()));
//        bdate=e2.getText().toString();
    }
}, newCalendar.get(Calendar.YEAR), newCalendar.get(Calendar.MONTH), newCalendar.get(Calendar.DAY_OF_MONTH));

}

@Override
public void response(JSONObject jo) {
    try {
        String status = jo.getString("status");
        Log.d("pearl", status);

        if (status.equalsIgnoreCase("success")) {
            Toast.makeText(getApplicationContext(), "Updated SUCCESS", Toast.LENGTH_LONG).show();
            startActivity(new Intent(getApplicationContext(), user_view_room.class));
        } else {
            startActivity(new Intent(getApplicationContext(), user_book_room.class));

            Toast.makeText(getApplicationContext(), " failed.TRY AGAIN!!", Toast.LENGTH_LONG).show();
        }
    } catch (Exception e) {
        // TODO: handle exception
        e.printStackTrace();
        Toast.makeText(getApplicationContext(), e.toString(), Toast.LENGTH_LONG).show();
    }
}

}
```

## QR

```
package com.example.travelore;

import androidx.appcompat.app.AppCompatActivity;

import android.content.SharedPreferences;
import android.os.Bundle;
import android.preference.PreferenceManager;
import android.util.Log;
import android.widget.AdapterView;
import android.widget.EditText;
import android.widget.ImageButton;
import android.widget.ImageView;
import android.widget.Toast;

import com.squareup.picasso.Picasso;

import org.json.JSONArray;
import org.json.JSONObject;

public class qr_code extends AppCompatActivity implements JsonResponse{
    EditText e1,e2;
    ImageView i1;
    SharedPreferences sh;
    String[] tile,description,value;
    String s;

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.activity_qr_code);
        sh = PreferenceManager.getDefaultSharedPreferences(getApplicationContext());
        e1=(EditText)findViewById(R.id.et1);
        e2=(EditText)findViewById(R.id.et2);
        i1=(ImageView)findViewById(R.id.ib);
        JsonRequest JR=new JsonRequest();
        JR.json_response=(JsonResponse)qr_code.this;
        String q="/qr?id="+AndroidBarcodeQrExample.vals;
        q=q.replace(" ","%20");
        JR.execute(q);
    }

    @Override
    public void response(JSONObject jo) {
        try {

            String status = jo.getString("status");
            Log.d("pearl", status);

            if (status.equalsIgnoreCase("success")) {
                JSONArray ja1 = (JSONArray) jo.getJSONArray("data");

                tile = new String[ja1.length()];
                description=new String[ja1.length()];
```



```
import android.widget.ImageButton;
import android.widget.Toast;

import org.json.JSONObject;

import java.io.ByteArrayOutputStream;
import java.io.File;
import java.io.FileInputStream;
import java.io.InputStream;
import java.text.SimpleDateFormat;
import java.util.Calendar;
import java.util.HashMap;
import java.util.Locale;
import java.util.Map;

public class user_add_post extends AppCompatActivity implements JsonResponse{
    EditText e1,e2;
    ImageButton i1;
    Button b1;
    String post_details,date,latitude,longitude;
    SharedPreferences sh;
    final int CAMERA_PIC_REQUEST = 0, GALLERY_CODE = 201;
    public static String encodedImage = "", path = "";
    private Uri mImageCaptureUri;
    byte[] byteArray = null;
    private DatePickerDialog fromDatePickerDialog;

    private SimpleDateFormat dateFormatter;

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.activity_user_add_post);
        sh = PreferenceManager.getDefaultSharedPreferences(getApplicationContext());
        e1 = (EditText) findViewById(R.id.et1);
        e2=(EditText)findViewById(R.id.ed);
        i1 = (ImageButton) findViewById(R.id.ib);
        b1 = (Button) findViewById(R.id.btn1);
        e2.setInputType(InputType.TYPE_NULL);
        e2.requestFocus();
        setDateTimeField();
        dateFormatter = new SimpleDateFormat("dd-MM-yyyy", Locale.US);

        i1.setOnClickListener(new View.OnClickListener() {
            @Override
            public void onClick(View v) {
                selectImageOption();
            }
        });

        b1.setOnClickListener(new View.OnClickListener() {
            @Override
            public void onClick(View v) {
                post_details = e1.getText().toString();
                date=e2.getText().toString();
                sendAttach();
            }
        })
    }
}
```

```
    });
}

private void sendAttach() {
    try {
        SharedPreferences sh = PreferenceManager.getDefaultSharedPreferences(getApplicationContext());
//        String uid = sh.getString("uid", "");

        String q = "http://" + ipsettings.text + "/api/user_upload_file";
//        String q = "http://" + IpSetting.ip + "/api/user_upload_file";

        Toast.makeText(getApplicationContext(), "Byte Array:" + byteArray.length, Toast.LENGTH_LONG).show();

        Map<String, byte[]> aa = new HashMap<>();

        aa.put("image", byteArray);
        aa.put("latitude", LocationService.lati.getBytes());
        aa.put("longitude", LocationService.logi.getBytes());

        aa.put("lid", sh.getString("log_id", "").getBytes());
        aa.put("post_details", post_details.getBytes());
        aa.put("date", date.getBytes());
//        aa.put("lname", lname.getBytes());
//        aa.put("house", house.getBytes());

        FileUploadAsync fua = new FileUploadAsync(q);
        fua.json_response = (JsonResponse) user_add_post.this;
        fua.execute(aa);
    } catch (Exception e) {
        Toast.makeText(getApplicationContext(), "Exception upload : " + e, Toast.LENGTH_SHORT).show();
    }
}

private void selectImageOption() {

/*Android 10+ gallery code
    android:requestLegacyExternalStorage="true"*/

    final CharSequence[] items = {"Capture Photo", "Choose from Gallery", "Cancel"};

    AlertDialog.Builder builder = new AlertDialog.Builder(user_add_post.this);
    builder.setTitle("Take Photo!");
    builder.setItems(items, new DialogInterface.OnClickListener() {
        @Override
        public void onClick(DialogInterface dialog, int item) {

            if (items[item].equals("Capture Photo")) {
                Intent intent = new Intent(MediaStore.ACTION_IMAGE_CAPTURE);
                //intent.putExtra(MediaStore.EXTRA_OUTPUT, imageUri);

                startActivityForResult(intent, CAMERA_PIC_REQUEST);

            } else if (items[item].equals("Choose from Gallery")) {
```

```
        Intent i = new Intent(Intent.ACTION_PICK, android.provider.MediaStore.Images.Media.EXTERNAL_CONTENT_URI);
        startActivityForResult(i, GALLERY_CODE);

        } else if (items[item].equals("Cancel")) {
            dialog.dismiss();
        }
    }
});
builder.show();
}
@Override
protected void onActivityResult(int requestCode, int resultCode, Intent data) {

    super.onActivityResult(requestCode, resultCode, data);

    if (requestCode == GALLERY_CODE && resultCode == RESULT_OK && null != data) {

        mImageCaptureUri = data.getData();
        System.out.println("Gallery Image URI : " + mImageCaptureUri);
        // CroppingIMG();

        Uri uri = data.getData();
        Log.d("File Uri", "File Uri: " + uri.toString());
        // Get the path
        //String path = null;
        try {
            path = FileUtils.getPath(this, uri);
            Toast.makeText(getApplicationContext(), "path : " + path, Toast.LENGTH_LONG).show();

            File fl = new File(path);
            int ln = (int) fl.length();

            InputStream inputStream = new FileInputStream(fl);
            ByteArrayOutputStream bos = new ByteArrayOutputStream();
            byte[] b = new byte[ln];
            int bytesRead = 0;

            while ((bytesRead = inputStream.read(b)) != -1) {
                bos.write(b, 0, bytesRead);
            }
            inputStream.close();
            byteArray = bos.toByteArray();

            Bitmap bit = BitmapFactory.decodeByteArray(byteArray, 0, byteArray.length);
            i1.setImageBitmap(bit);

            String str = Base64.encodeToString(byteArray, Base64.DEFAULT);
            encodedImage = str;
            // sendAttach1();
        } catch (Exception e) {
            Toast.makeText(getApplicationContext(), e.toString(), Toast.LENGTH_LONG).show();
        }
    } else if (requestCode == CAMERA_PIC_REQUEST && resultCode == Activity.RESULT_OK) {

        try {
            Bitmap thumbnail = (Bitmap) data.getExtras().get("data");
            ByteArrayOutputStream baos = new ByteArrayOutputStream();
            thumbnail.compress(Bitmap.CompressFormat.JPEG, 100, baos);
```

```
        i1.setImageBitmap(thumbnail);
        byteArray = baos.toByteArray();

        String str = Base64.encodeToString(byteArray, Base64.DEFAULT);
        encodedImage = str;
//        sendAttach1();
    } catch (Exception e) {
        e.printStackTrace();
    }
}

}

private void setDateTimeField() {

    e2.setOnClickListener(new View.OnClickListener() {

        @Override
        public void onClick(View arg0) {
            // TODO Auto-generated method stub
            fromDatePickerDialog.show();
        }
    });
    Calendar newCalendar = Calendar.getInstance();
    fromDatePickerDialog = new DatePickerDialog(this, new DatePickerDialog.OnDateSetListener() {

        public void onDateSet(DatePicker view, int year, int monthOfYear, int dayOfMonth) {
            // TODO Auto-generated method stub
            Calendar newDate = Calendar.getInstance();
            newDate.set(year, monthOfYear, dayOfMonth);
            e2.setText(dateFormatter.format(newDate.getTime()));
//            bdate=e2.getText().toString();

        }
    }, newCalendar.get(Calendar.YEAR), newCalendar.get(Calendar.MONTH), newCalendar.get(Calendar.DAY_OF_MONTH));
}

}

@Override
public void response(JSONObject jo) {
    try {
        String status = jo.getString("status");
        Log.d("pearl", status);

        if (status.equalsIgnoreCase("success")) {
            Toast.makeText(getApplicationContext(), "Uploaded SUCCESS", Toast.LENGTH_LONG).show();
            startActivity(new Intent(getApplicationContext(), user_home.class));

        } else {
            startActivity(new Intent(getApplicationContext(), user_add_post.class));

            Toast.makeText(getApplicationContext(), " failed.TRY AGAIN!!", Toast.LENGTH_LONG).show();
        }

    } catch (Exception e) {
        // TODO: handle exception
        e.printStackTrace();
    }
}
```

# Travelore

---

```
        Toast.makeText(getApplicationContext(), e.toString(), Toast.LENGTH_LONG).show();
    }
}
```

## **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 phase is the phase, which involves the process of converting a new system design into an operational one. It is the key stage in achieving a successful new system. Implementation is the stage if the project, where the theoretical design is turned into a working system.

The implementation plan consists of the following steps:

- Testing the developed system with the sample data.
- Detection and correction of errors.
- Making necessary changes in the system.
- Training and involvement of user personnel.
- Installation of software utilities.

## 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

The application was divided into smaller components and tested individually. Each code was executed separately to ensure accuracy.

### • Integration Testing

Each small component was integrated or combined into a module to ensure that each module works properly when put together. This was done to check connectivity between modules.

### • System Testing

The system as a whole was tested by combining every module. This was to ensure that each process have a particular order. This was to ensure that the system does not crash while using.

### • Output Testing

After performing the validation testing, the next step is the output testing of the enhanced system. No system could be useful if it does not produce the required output in the required format. The outputs generated or the displayed by the system are tested by asking the users about the format required by them.

### • User acceptance Testing

User acceptance testing comprises a completed and successful end-to-end system test with review of the results by one or more users with specific knowledge.

### • Validation Testing

Validation testing is confirmation that a given information meets its intended use and the needs.

## **5.4 System Security**

Password encryption is used to protect each hotel, travel agency and user's details.

## **5.5 Scope for Future Enhancement**

The current system is flexible and can be modified in the future

## **6.CONCLUSION**

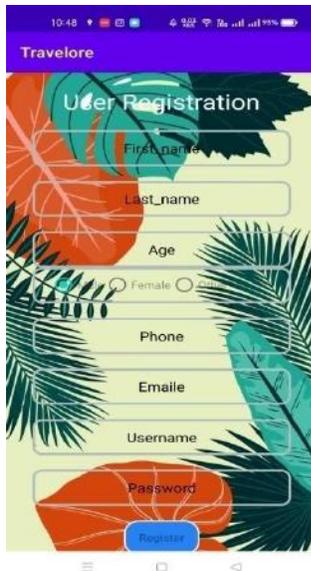
The software developed has fulfilled the necessary requirements as required by the user. It is ensured that all the programs are working properly in the “Travelore” application. The system is used to operate in a user-friendly manner. Proper documentation done from different areas provides smooth running of all the operations without difficulty. The system has achieved the objective to embrace the culture and explore the scope of travel and tourism in Kerala by acting as a medium between various small/large scale hotels and travel agencies and users thereby giving the opportunity to the hotels and travel agencies in their business development as well as providing travelling and exploring facilities to the users based on their specified need and interest. The project has been implemented and tested.

## **7. APPENDIX**

## INTERFACE

### 1. User(Android)

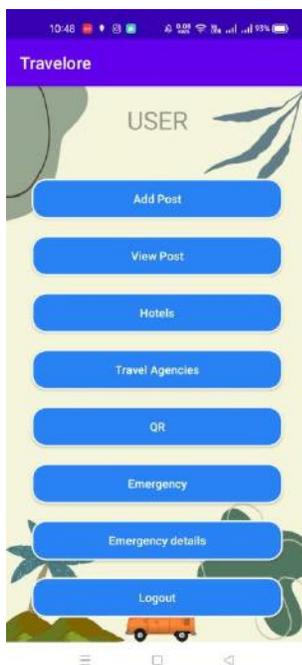
#### Registration



#### Login

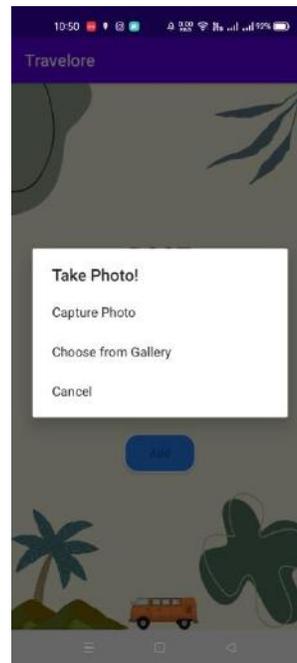


#### User Login



# Travelore

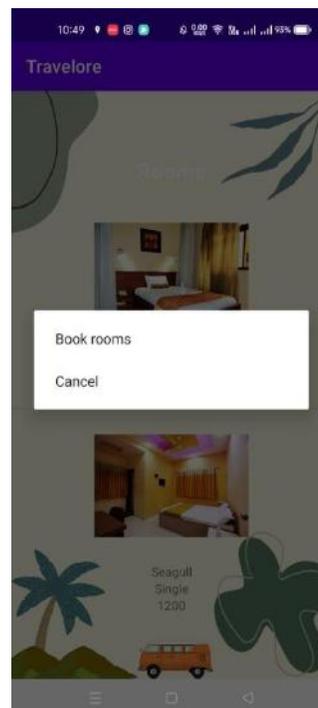
## User adds post



## User view hotels



## User book rooms

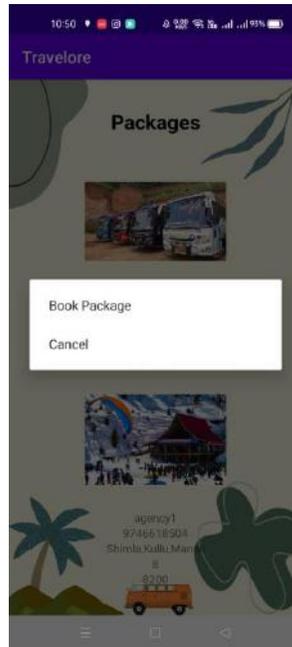


# Travelore

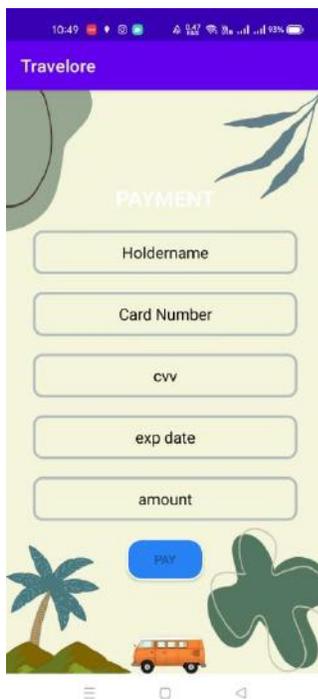
User view travel agencies



User book travel agencies



User pays advance to book hotels and travel agencies



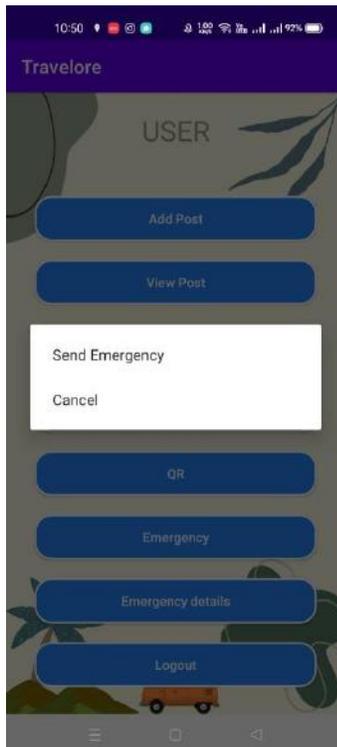
User scan QR



SCAN QR CODE



User send emergency alert

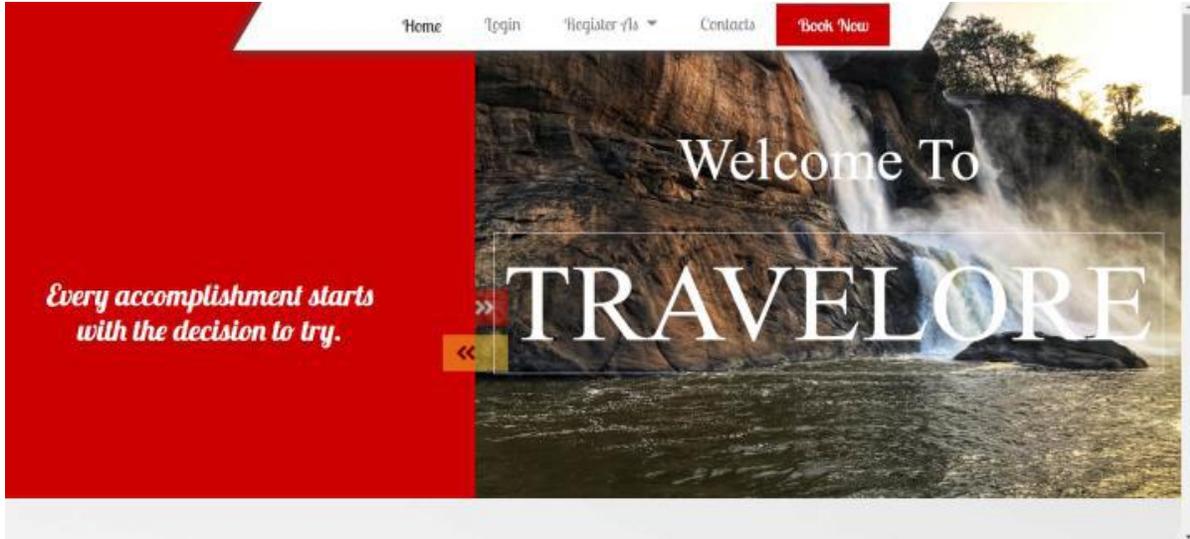


# Travelore

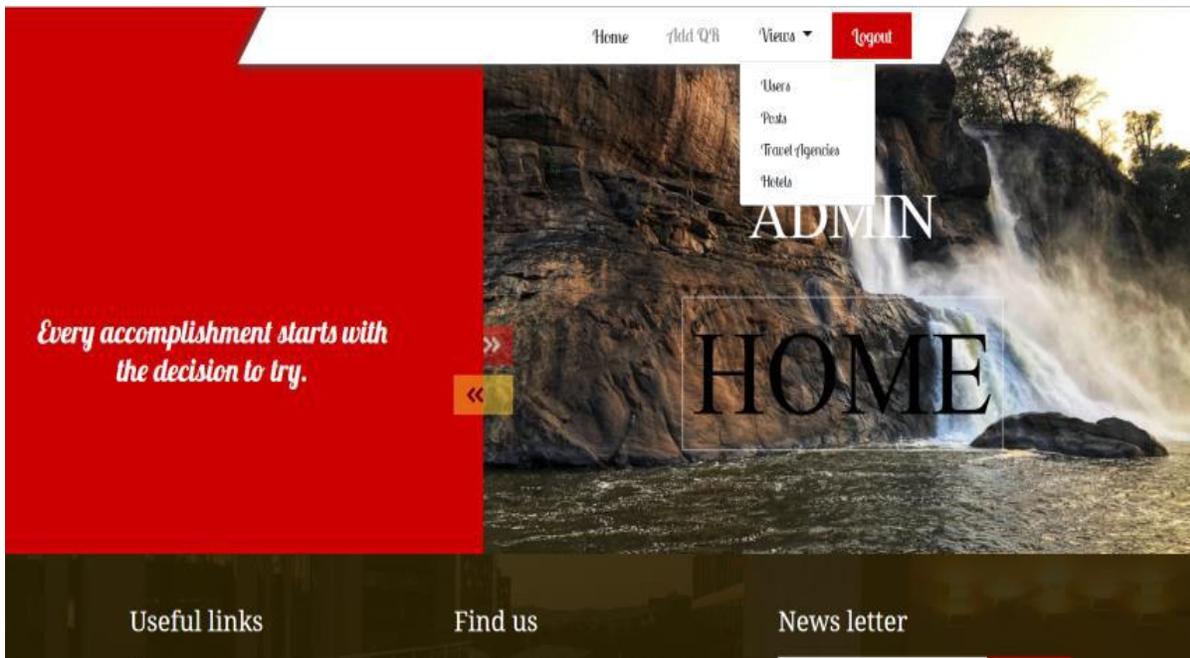
---

## 2. WEB

### Login



### Admin home



## Hotel Registration

Home Login Register (dropdown) Contacts Book Now Logout

### Hotel Registration

Name

Place

Phone

Email

Latitude  View Map

Longitude

Username

Password

Register

## Travel agency registration

Home Login Register (dropdown) Contacts Book Now Logout

### Travel Agency Registration

Name

Place

Phone

Email

Latitude  View Map

Longitude

Username

Password

Register

## Adding QR

QR

Heading

Description

Image  No file chosen

## Admin managing hotels

HOTELS

Hotel Name	Place	Phone Number	Email	Latitude	Longitude	
Seagull	Fort Kochi	4672349502	seagull@gmail.com	9.976684570645943	76.2796058147821	<input type="button" value="Reject"/>
Ginger House	Fort Kochi	3456789987	gingerhouse@gmail.com	9.975164737631026	76.27997774601927	<input type="button" value="Reject"/>

Useful links  
[Home](#)  
[About](#)

Find us  
Healing Center, 176 W Street name,  
New York, NY 10014, US

News letter  
Email

## Admin managing travel agencies

The screenshot displays the 'Travel Agencies' management interface. At the top, there is a navigation bar with 'Home', 'Add Q&A', 'View', and 'Logout' buttons. Below this is a table listing travel agencies with columns for Agency Name, Place, Phone, Email, Latitude, and Longitude. Each row includes a 'Reject' button, except for the last row which has 'Accept' and 'Reject' buttons. The footer contains three sections: 'Useful links' (Home, About, Rooms, Blog, Contacts), 'Find us' (address, phone numbers, and email), and 'News letter' (Email input field and Submit button). Social media icons for Facebook, Twitter, and Instagram are also present.

Agency Name	Place	Phone	Email	Latitude	Longitude	
agency1	ekm	2563201478	a@gmail.com	9.980938863268337	76.2766170501709	Reject
sdf	chalakudi	7451236985	as@gmail.com	9.980743879720944	76.27677912271129	Reject
dddddd	ekm	2563201478	a@gmail.com	9.965527812628109	76.29649428689596	Reject
pavzham	Fort Kochi	9654456789	pavzham@gmail.com	9.978037092405636	76.28578562435241	Accept Reject

**Useful links**

- Home
- About
- Rooms
- Blog
- Contacts

**Find us**

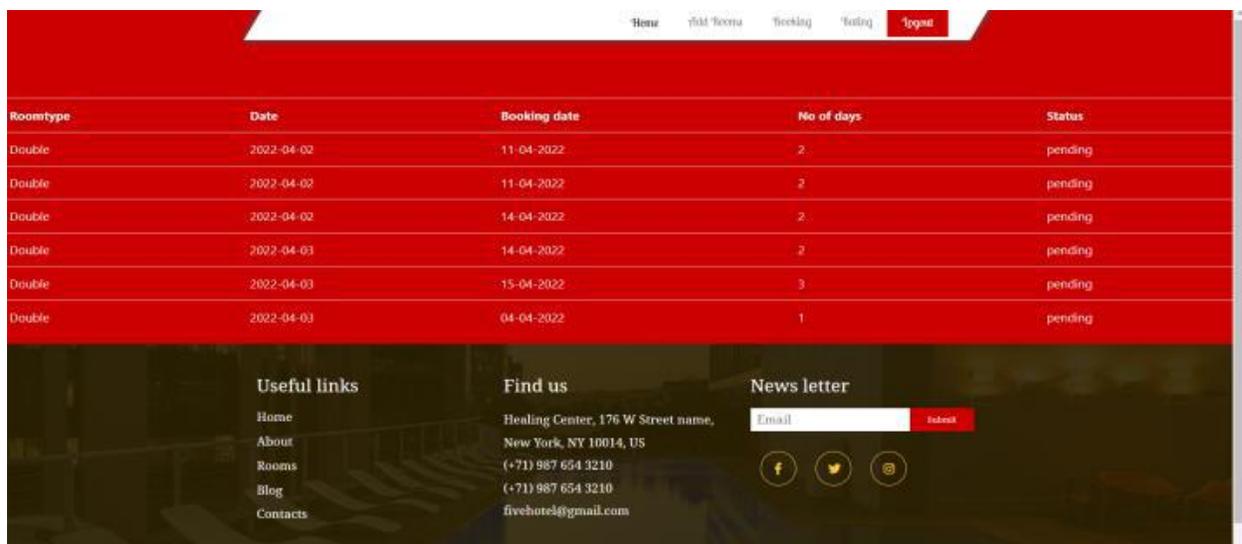
Healing Center, 176 W Street name,  
New York, NY 10014, US  
(+71) 987 654 3210  
(+71) 987 654 3210  
fivehotel@gmail.com

**News letter**

Email:

[f](#) [t](#) [i](#)

## Travel Agency views booking



The screenshot shows a web application interface with a red header and a table of bookings. The table has five columns: Roomtype, Date, Booking date, No of days, and Status. Below the table is a footer with three sections: Useful links, Find us, and News letter.

Roomtype	Date	Booking date	No of days	Status
Double	2022-04-02	11-04-2022	2	pending
Double	2022-04-02	11-04-2022	2	pending
Double	2022-04-02	14-04-2022	2	pending
Double	2022-04-03	14-04-2022	2	pending
Double	2022-04-03	15-04-2022	3	pending
Double	2022-04-03	04-04-2022	1	pending

**Useful links**

- Home
- About
- Rooms
- Blog
- Contacts

**Find us**

Healing Center, 176 W Street name,  
New York, NY 10014, US  
(+71) 987 654 3210  
(+71) 987 654 3210  
fivehotel@gmail.com

**News letter**

Email

[f](#) [t](#) [i](#)

## **8. BIBLIOGRAPHY**

## Reference Website

- <https://stackoverflow.com/>
- <https://www.tutorialspoint.com/>
- <https://www.javatpoint.com/>
- <https://www.mysqltutorial.org/>

सी इ पी सी आई प्रयोगशाला  
एवं अनुसंधान संस्थान



CEPCI LABORATORY &  
RESEARCH INSTITUTE

Cashew Bhavan, Mundakkal, Kollam-691 001, Kerala, India, E-mail:cepcilab@cashewindia.org  
Tel:+91 474 2761003, 2742704, 2742504 Web: www.cepcilab.org.in

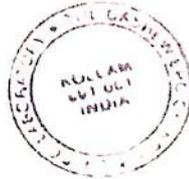
No.QUL/M.Sc./2022/ 284

CERTIFICATE

This is to certify that this dissertation entitled "Extraction of pectin from banana and mango peels and its applications" is a bonafide record of independent work carried out by Ms. Evelin Pradeep M.Voc Food processing technology, St.Terasas College Ernakulam during three month project work under the guidance and supervision of Dr. Prabhakumari.C of this Institute. It is further certified that she has independently reviewed the literature, performed all the tests, analyzed the results and critically discussed the findings in presented data.

This is in partial fulfillment of the requirements for the M.Voc degree in Food processing technology M.G University.

12-07.2022  
Kollam



  
(Dr. Prabhakumari C.)  
Dy. Principal Scientist

**Dr. C. PRABHAKUMARI**  
Deputy Principal Scientist

Established by:



Accreditations & Approvals

- NABL as per ISO/IEC 17025 : 2017 in the fields of Chemical and Biological Testing (TC-6088)
- Bureau of Indian Standards (BIS) for Packaged Drinking Water Testing ● FSSAI notified Laboratory

CIN: U74999KL1955NPL000300

**“INCORPORATION OF HIBISCUS ROSA-SINESIS  
EXTRACT IN THE BARLEY BASED NOODLES”**

*Dissertation submitted to Mahatma Gandhi University in partial  
fulfillment of the requirements for the award of degree of*

*Bachelor of Vocational Studies*

*Submitted by*

**ANNIE PHILIP**

**(REGISTER NO.: VB19FPT007)**

**EVLIN PETERS**

**(REGISTER NO.: VB19FPT013)**

**NAMITHA SUNIL**

**(REGISTER NO.: VB19FPT023)**

*Under the guidance of*

**Ms. Sandhra Santhosh**

Assistant Professor

**B.VOC. FOOD PROCESSING TECHNOLOGY**



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

COLLEGE WITH POTENTIAL FOR EXCELLENCE

Nationally Re-Accredited at 'A++' Level (4<sup>th</sup> Cycle)

Affiliated with Mahatma Gandhi University

Kottayam- 686560

March 2022