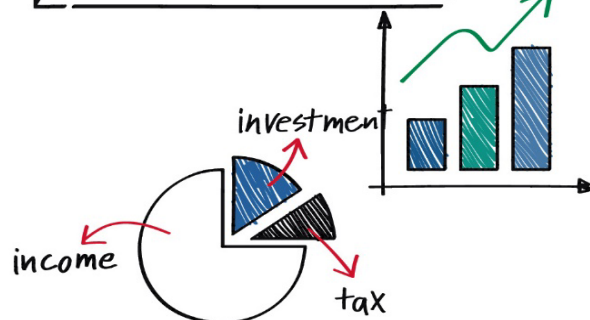
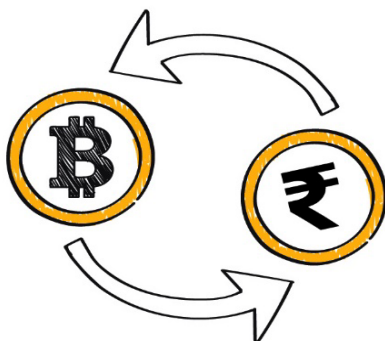
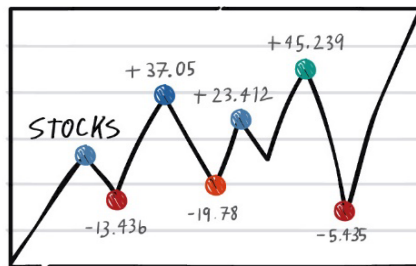




# FINANCIAL LITERACY



Savings



Financial Plan



Gokhale Education Society's

**H.P.T. ARTS & R.Y.K. SCIENCE COLLEGE NASHIK-422005**

**DEPARTMENT OF STATISTICS**



**CERTIFICATE**

This is to certify that the project work entitled “**FINANCIAL LITERACY**” is a bonafide work carried out by,

**02 BHUSHAN NANA BAGUL.**

**03 PRIYANKA PREMANAND BAGUL.**

**11 PRATIKSHA GANESH GHODKE**

**18 SAGAR JAY KANADE**

**26 MAYURI RAVINDRA PATIL**

**32 KHUSHALI SANJAY SHIMPI**

**38 MAYANK SURESH KUMAR VANSI**

WITH PARTIAL FULFILLMENT FOR THE STATISTICS PROJECT OF THE SAVITRIBAI PHULE PUNE UNIVERSITY DURING THE YEAR **2021-2022** .THE PROJECT REPORT HAS BEEN APPROVED AS IT SATISFIES THE ACADEMIC REQUIREMENTS IN RESPECT OF PROJECT WORK PRESCRIBED FOR THE SAID DEGREE.

Signature of the Guide,

Ms. Harshada B. Jadhav.

External Examiner

Head of Department

Prof.Mrs. V. S. Joshi

## ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance from many people and we are extremely privileged to have got this all along the completion of our project. All that we have done is only due to such supervision and we would not forget to thank them. So, it gives us a great pleasure in acknowledging our deep sense of gratitude to all those who helped in completing this project. We select this project to acquire practical knowledge in the subject at the level of UG.

We express our gratitude to Principal **V.N. Suryawanshi** for giving us all support and allowing us to present this project.

We owe our deep gratitude to our Head of the Department Prof. **Mrs. V.S. Joshi ma'am** who took keen interest in our project and guide us all along. We should not forget to remember our project guide Prof. **Mrs. Harshada B. Jadhav ma'am** for their encouragement and guidance till the completion of our project work by providing all the necessary information for developing a good project. We pay our gratitude to **Prof. Yeole Sir** who helped us with Data analytics.

We would like to extend our sincere gratitude to all teaching and non-teaching staff in laboratory for their timely support.

We owe our deep gratitude to economics Department faculty **Prof. Vishwajit Kadam sir** who helped us with all the information about economics and Financial Literacy.

Last but not least we would like to express our gratitude towards the people who gave necessary information required for the project.

And finally, the students of our college, friends and family for their support without which the project could not have been successful one.

Heartily gratitude to all of you!

## INDEX

SR NO.	TOPIC	PAGE NO.
1	Objectives	6
2	Introduction to Financial Literacy	7-11
3	Questionnaire	12-15
4	Coding of Questionnaire	16-20
5	Data	21-22
6	Calculation of financial literacy score	23
7	Information about statistical tools	24-31
	<b>Graphical Representation:</b>	
8	1]Box plot	32
9	2]Rod or Spike Plot	33
10	3]Line chart	34
11	4]Bar graph	35
12	5] Pie Chart	36
13	6] Simple Bar Diagram	37-38
14	Financial Literacy Scale	39
	<b>Chi-Square Test for Independence</b>	

15	1] Dependency of Risk while investing and well Diversified Portfolio:	40
16	2] Dependency of Annual income and availability of Emergency fund	41
	<b>Test for Proportion:</b>	
17	1]To check whether the proportion of financial literacy is equal or different in males and females(18-30 age group)	42
18	2]To check whether the proportion of financial literacy is equal or different in urban and rural area people	44
19	Logistic regression	45-49
20	Naïve Bayes Algorithm	50-51
21	Survey Experience	52
22	Conclusion	53
23	How to increase Financial Literacy	54
24	Limitations	54
25	Reference	55

## OBJECTIVES

- To get an idea about financial literacy among population.
- To check whether the financial literacy depend upon factors like age, gender, region, etc.
- To check whether the proportion of financial literacy is equal or different in males and females.
- To check whether the proportion of financial literacy is equal or different in urban and rural area people.
- To check dependency of Annual income and availability of Emergency fund.
- To check dependency of Risk while investing and well Diversified Portfolio.
- To do graphical analysis for representing the complex data into simpler form.
- To come up with real life implementing solution to increase Financial Literacy among population.

## INTRODUCTION TO FINANCIAL LITERACY

Financial literacy is the ability to understand and effectively use various financial skills, including personal financial management, budgeting, and investing. Financial literacy is the foundation of your relationship with money, and it is a lifelong journey of learning. In simple words we can say that to be financially literate is to know how to manage your money. This means learning how to pay your bills, how to borrow and save money responsibly, and how and why to invest and plan for retirement.

It is the cognitive understanding of financial components and skills such as budgeting, investing, borrowing, taxation, and personal financial management. The absence of such skills is referred to as being financially illiterate. It also includes paying off debt, creating a budget, and understanding the difference between various financial instruments. In sum, financial literacy has a material impact on families as they try to balance their budget, buy a home, fund their children's education, or ensure an income for retirement.

According to the Financial Literacy and Education Commission, there are five key components of financial literacy: earn, spend, save and invest, borrow, and protect. In India April is celebrated as Financial Literacy Month.

*Financial Literacy has become one of the top priorities for most of the countries today as it is directly proportional to the economic growth of the country.*

### **The significance of financial literacy in India are as follow:**

Development of rural areas: Reaching out to rural sections and working on their development can be achieved through financial literacy. This can be achieved by making people more aware about the available resources and right way of utilizing them and boosting our economy.

*According to National Centre for Financial Education survey which says that only 27% of Indians are financially literate. It means that we have a long distance to travel and that puts a lot of responsibility on all of us that how best we take the message of financial literacy across the country. Globally conducted studies by the Organization for Economic Cooperation and Development (OECD) show an alarming rate of financial illiteracy. Consumers fail to answer simple financial questionnaires all over the world. They also show an inadequate understanding of financial principles, risks, and debt.*

- **Understanding the concept of Budgeting:**

A Budget simply maps income and expenses. Most often this is done on a monthly interval as this provides a consistent, yet accurate description of actual income in comparison to expenses.

If you have ever experienced fear about your finances, budgeting will provide a roadmap. Active awareness of what you have coming in, your income and what is going out, your expenses. An individual needs a record keeping system and management of each penny within their system. Budgeting is one component of one's financial picture, saving, cash management, credit management, insurance coverage, buying goods and services, paying for higher education, utilizing employer benefits, saving for retirement, and providing health care. Budgeting is a foundational piece of becoming a wiser consumer and financially literate.

- **What does it mean to invest in the stock market?**

Stocks are an investment that means you own a share in the company that issued the stock. Simply put, stocks are a way to build wealth. This is how ordinary people invest in some of the most successful companies in the world. Investing in stock market requires financial knowledge and proper understanding of things.

- **What is investing in Cryptocurrency?**

Cryptocurrencies are digital assets people use as investments and for online purchases. You exchange real currency, like dollars, to buy “coins” or “tokens” of a certain kind of cryptocurrency.

The rapidly growing popularity and adoption of crypto assets (cryptocurrencies) in India. In a world full of risks and uncertainties, the scope of financial literacy is continuously expanding.

- **What is RD and how it works?**

A Recurring Deposit is a special kind of term-deposit offered by banks. A Recurring Deposit, commonly known as RD, is a unique term-deposit that is offered by Indian Banks. It is an investment tool which allows people to make regular deposits and earn decent returns on the investment. Due to the regular deposit factor and an interest component, it often provides flexibility and ease of investments to individuals.

However, it is essential to know that RDs are different from Fixed Deposits/FDs. RDs are flexible in most aspects. An RD account holder can choose to invest a fixed amount each month while earning decent interest on the amount.



- **What is SIP?**

A Systematic Investment Plan (SIP), more popularly known as SIP, is a facility offered by mutual funds to the investors to invest in a disciplined manner. SIP facility allows an investor to invest a fixed amount of money at pre-defined intervals in the selected mutual fund scheme. The fixed amount of money can be as low as Rs.500, while the pre-defined SIP intervals can be on a weekly/monthly/quarterly/semi-annually or annual basis. By taking the SIP route to investments, the investor invests in a time-bound manner without worrying about the market dynamics and stands to benefit in the long-term due to average costing and power of compounding.

- **What is inflation rate?**

Inflation is the rate of increase in prices over a given period of time. Inflation is typically a broad measure, such as the overall increase in prices or the increase in the cost of living in a country.

**What is WPI and CPI inflation?**

The two most-often used inflation rates in the country are the year-on-year.

- > The wholesale price index (WPI) based inflation rate and.
- > The consumer price index (CPI) based inflation rate.

The former is called the wholesale inflation rate and the latter is called the retail inflation rate.

- **Relation between income and expenditure:**

The difference between income and consumption is used to define the consumption schedule. When income grows, disposable income rises and thus consumers buy more goods. The result is an increase in the consumption of major purchases and non-essential goods. Ideal income and expenditure relation is

**Expenditure < Income**

- **What is passive income?**

A passive income is something that many aspire to achieve. It's a regular form of income that requires little maintenance or effort on your behalf, and most importantly, it's not reliant on you inputting your time on a daily or hourly basis. This type of income will generate on its own, which allows you to focus on other areas of your business and quite literally make money while you sleep.

- **Diversifying your income:**

Diversifying your income means drawing earnings from two or more sources rather than from just one job. A diversified income can protect you against the loss of your sole job. It can help you to save or invest more for your future.

**What Does It Mean to Diversify Income?**

A diversified income derives from more than one source. A diversified portfolio is one that has investments spread across assets such as equities, fixed income, commodities, and cash, and is recommended to limit risk. It also spreads out the holdings within each asset class.

**Benefits of Diversifying Your Income:**

Diversifying your income can help provide stability. It can hedge against income loss due to layoffs, illness, disability, discrimination, and more. Having multiple streams of income can have an incredible impact on your ability to build wealth and fund retirement.

- **Emergency funds**

The term “emergency fund” refers to money stashed away that people can use in times of financial distress. The purpose of an emergency fund is to improve financial security by creating a safety net that can be used to meet unanticipated expenses, such as an illness or major home repairs.

**How to Build an Emergency Fund**

Starting early is the key to setting up an emergency fund, because it helps you build up a comfortable cushion against unexpected emergencies later in life. Getting a start on emergency funds is relatively easy. Here are two simple ways to begin saving for one.

1. Set aside a comfortable amount from your salary each month.
2. Save your tax refund.

## Investment risk

### Q. What Is Risk ?

When you invest, you make choices about what to do with your financial assets. Risk is any uncertainty with respect to your investments that has the potential to negatively affect your financial welfare. For example, your investment value might rise or fall because of market conditions (market risk). Generally speaking, the more financial eggs you have in one basket, say all your money in a single stock, the greater risk you take (concentration risk).

#### ▪ **Market risk**

The risk of investments declining in value because of economic developments or other events that affect the entire market. The main types of market risk are equity risk, interest rate risk and currency risk.

- **Equity risk** – applies to an investment in shares. The market price of shares varies all the time depending on demand and supply. Equity risk is the risk of loss because of a drop in the market price of shares.
- **Interest rate risk** – applies to debt investments such as bonds. It is the risk of losing money because of a change in the interest rate. For example, if the interest rate goes up, the market value of bonds will drop.
- **Inflation risk**

The risk of a loss in your purchasing power because the value of your investments does not keep up with inflation. Inflation erodes the purchasing power of money over time – the same amount of money will buy fewer goods and services.

## QUESTIONNAIRE

**1] Gender:**

- a. Male   b. Female   c. Others

**2] Age group:**

- a. 18-30                      b. 31-40  
c. 41-50                      d. 51 and above

**3] Educational Qualification:**

- a. SSC and below              b. HSC  
c. Graduation                d. Post Graduate and above

**4] Currently you are a:**

- a. Student                      b. Service/Business (man)  
c. Unemployed                d. House-maker  
e. Others

**5] Area you belong to:**

- a. Rural                        b. Urban

**6] Annual income (Parents/Self):**

- a. Below 2 Lakhs              b. 2-5 Lakhs  
c. 5-10 Lakhs                d. Above 10 Lakhs

**7] Do you have any loan (Parents / Self)?**

- a. Yes                        b. No

**8] Do you pay your installments on time?**

- a. Yes                        b. No

**9] How do you make your financial Budget?**

- a. Monthly                      b. Half yearly  
c. Yearly                        d. No Financial Budget

**10] Do you invest in Stocks or Crypto?**

- a. Yes                        b. No

**11] Do you have monthly RD (Recurring deposit) or SIP(Systematic investment plan)?**

- a. Yes                      b. No                      c. Don't Know

**12] Do you know anything about financial literacy (money management)?**

- a. Yes                      b. No

**13] Have you ever learnt how to manage your money and where?**

- a. Course about managing your money  
b. At School as part of subject or any specialized course  
c. Self learnt (Learnt from my parents and through observations)  
d. Unaware  
e. Others

**14] Do you know interest rate of your bank account?**

- a. yes                      b. No

**15] Do you know inflation rate?**

- a. Yes                      b. No                      c. Unaware

**16] Rate yourself for how much you know about the tax you or your parents pay each financial year?**

- a. 0 (No knowledge)  
b. 1(least knowledge about taxation)  
c. 2 (least knowledge about taxation)  
d. 3 (Average knowledge about taxation)  
e. 4 (Average knowledge about taxation)  
f. 5 (Complete knowledge about taxation)

**17] Relation between your income and expenditure habits?**

- a. Expenditure>Income  
b. Expenditure< Income  
c. Expenditure= Income  
d. Don't Know

**18] How much percent of knowledge do you have about 'Insurance' which protects your finances and investment?**

- a. 0-20%                  b. 20-40%  
c. 40-60%                  d. 60-100%

**19] Do you have any passive income?**

- a. Dividend Income (Share Profits)
- b. Rental Income
- c. Gains on investments
- d. No passive Income

**20] Do you think your invested money is well Diversified (Properly distributed)?**

- a. Yes
- b. No

**21] Do you have any emergency fund?**

- a. Yes
- b. No

**22] How much satisfied you are with your spending habits?**

- a. Very Satisfied
- b. Satisfied
- c. Neutral
- d. Dissatisfied

**23] The first thing you will do in financial emergency?**

- a. Take a bank loan
- b. Take a loan from moneylender with high interest rate
- c. Dissolution of RD or SIP or FD
- d. Use emergency fund (savings)
- e. Others

**24] How much are you aware about risks while investing?**

- a. Extremely Aware
- b. Have heard about it
- c. Aware
- d. Not Aware

**25] What are your financial goals?**

- a. Certain amount in saving
- b. Own a luxury item (like Car ,etc)
- c. Own a land
- d. Others

**26] What kind of solution do you think will make people aware about financial literacy?**

- a. It should be taught at school level
- b. There should be a proper training course available
- c. One should apply trial and error method
- d. No suggestions
- e. Others

## CODING OF QUESTIONNAIRE

<b>1] Gender</b>	<b>Code</b>
Male	<b>0</b>
Female	<b>1</b>

<b>2] Age group:</b>	<b>Code</b>
18-30	<b>0</b>
31-40	<b>1</b>
41-50	<b>2</b>
51 and above	<b>3</b>

<b>3] Educational Qualification</b>	<b>Code</b>
SSC and below	<b>0</b>
HSC	<b>1</b>
Graduation	<b>2</b>
Post Graduate and above	<b>3</b>

<b>4] Currently you are a:</b>	<b>Code</b>
Student	<b>0</b>
Service/Business (man)	<b>1</b>
Unemployed	<b>2</b>
House-maker	<b>3</b>
Others	<b>4</b>

<b>5] Area you belong to:</b>	<b>Code</b>
Rural	<b>0</b>
Urban	<b>1</b>

<b>6] Annual income (Parents/Self):</b>	<b>Code</b>
Below 2 Lakhs	<b>0</b>
2-5 Lakhs	<b>1</b>
5-10 Lakhs	<b>2</b>
Above 10 Lakhs	<b>3</b>



<b>7] Do you have any loan (Parents / Self)?</b>	<b>Code</b>
No	0
Yes	1

<b>8] Do you pay your installments on time?</b>	<b>Code</b>
Yes	0
No	1

<b>9] How do you make your financial Budget?</b>	<b>Code</b>
No Budgeting	0
Yearly	1
Half yearly	2
Monthly	3

<b>10] Do you invest in Stocks or Crypto?</b>	<b>Code</b>
No	0
Yes	1

<b>11] Do you have monthly RD (Recurring deposit) or SIP (Systematic investment plan)?</b>	<b>Code</b>
Don't know	0
No	1
Yes	2

<b>12] Do you know anything about financial literacy (money management)?</b>	<b>Code</b>
No	0
Yes	1

<b>13] Have you ever learnt how to manage your money and where?</b>	<b>Code</b>
Unaware	0
At School as part of subject or any specialized course	1
Self-learnt (Learnt from my parents and through observations)	2
Course about managing your money	3
Others	4

<b>14] Do you know interest rate of your bank account?</b>	<b>Code</b>
No	0
Yes	1

<b>15] Do you know inflation rate?</b>	<b>Code</b>
Unaware	0
No	1
Yes	2

<b>16] Rate yourself for how much you know about the tax you or your parents pay each financial year?</b>	<b>Code</b>
0 (Not known)	0
1	1
2	2
3	3
4	4
5(Properly Known)	5

<b>17] Relation between your income and expenditure habits?</b>	<b>Code</b>
Don't Know	0
Expenditure>Income	1
Expenditure=Income	2
Expenditure<Income	3

<b>18] How much percent of knowledge do you have about 'Insurance' which protects your finances and investment?</b>	<b>Code</b>
0-20%	0
20-40%	1
40-60%	2
60-100%	3

<b>19] Do you have any passive income?</b>	<b>Code</b>
No passive Income	0
Rental Income	1
Dividend Income (Share Profits)	2
Gains on investments	3

<b>20] Do you think you have well Diversified Portfolio (investment properly distributed)?</b>	<b>Code</b>
No	0
Yes	1

<b>21] Do you have any emergency fund?</b>	<b>Code</b>
No	0
Yes	1

<b>22] How much satisfied you are with your spending habits?</b>	<b>Code</b>
Dissatisfied	0
Neutral	1
Satisfied	2
Very Satisfied	3

<b>23] The first thing you will do in financial emergency?</b>	<b>Code</b>
Take a loan from moneylender with high interest rate	0
Others	1
Take a bank loan	2
Dissolution of RD or SIP or FD	3
Use emergency fund (savings)	4

<b>24] How much are you aware about risks while investing?</b>	<b>Code</b>
Not Aware	0
Have heard about it	1
Aware	2
Extremely Aware	3

<b>25] What are your financial goals?</b>	<b>Code</b>
Certain amount in saving	0
Own a luxury item (like Car etc.)	1
Own a land	2
Others	3

<b>26] What kind of solution do you think will make people aware about financial literacy?</b>	<b>Code</b>
It should be taught at school level	0
There should be a proper training course available	1
One should apply trial and error method	2
No suggestions	3
Others	4

## DATA

Serial	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Financial Score
1	0	0	2	0	0	0	0	0	3	0	0	1	2	1	2	2	1	1	0	0	1	2	4	0	3	0	15
2	0	1	3	1	1	1	1	1	3	1	2	1	2	1	2	5	3	3	3	1	0	1	4	3	3	0	30
3	0	0	3	2	0	0	0	0	0	0	1	0	0	0	1	0	0	2	1	0	0	0	1	2	1	1	8
4	1	0	2	0	1	2	0	0	3	0	1	1	1	1	2	4	1	1	0	1	0	3	3	0	0	0	18
5	1	0	2	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	2	0	4
6	0	2	3	1	0	1	1	1	0	0	0	0	2	1	0	3	1	0	1	0	1	1	4	0	2	0	12
7	1	0	2	0	0	0	0	1	2	1	0	1	1	1	2	4	1	2	0	1	0	1	2	0	3	0	16
8	0	0	2	1	0	1	0	1	1	0	0	1	2	1	0	1	3	2	0	0	0	3	4	3	0	2	18
9	1	0	3	0	0	0	0	0	0	0	1	1	2	0	0	3	0	0	0	0	0	0	0	0	0	0	5
10	1	0	2	0	1	0	0	0	1	0	0	1	2	1	2	4	0	2	0	1	0	1	4	2	3	0	18
11	0	0	2	0	0	0	0	1	3	0	1	1	2	0	2	3	1	2	0	1	0	0	4	0	0	0	15
12	0	1	3	1	0	3	1	1	3	0	2	1	2	1	1	3	2	2	0	1	0	1	2	0	0	1	16
13	0	0	2	0	0	2	1	1	3	0	0	0	2	0	0	3	1	2	0	1	1	1	4	0	0	1	13
14	1	0	2	0	1	3	1	1	3	0	0	1	2	1	2	4	2	2	0	1	1	2	3	3	0	1	22
15	0	0	1	0	0	0	0	1	0	0	0	1	2	0	2	4	0	2	0	0	0	1	4	2	2	0	16
16	1	0	2	0	0	0	0	1	0	0	1	0	2	0	1	3	1	1	0	1	0	2	3	2	0	3	15
17	1	0	2	0	1	0	0	1	3	0	1	0	2	0	1	2	2	2	0	0	0	2	2	2	2	1	14
18	0	0	0	1	0	0	0	0	3	0	1	0	0	1	0	3	1	1	0	0	1	1	4	0	1	0	13
19	0	0	1	4	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2	0	0	3	3	4
20	1	0	2	0	0	0	0	1	3	0	2	1	2	1	2	3	2	2	0	1	0	2	4	0	3	3	20
21	0	0	2	0	0	1	0	1	3	0	1	0	0	1	2	1	2	0	0	0	0	2	0	1	0	9	
22	1	0	2	0	0	0	1	0	3	1	2	1	2	1	2	5	2	2	2	1	0	2	4	2	0	0	27
23	0	0	3	0	0	0	0	1	3	0	1	1	2	1	2	3	1	2	2	1	1	3	2	3	0	1	23
24	0	3	3	1	0	2	1	1	3	1	2	1	2	1	2	3	3	1	2	1	1	3	4	2	2	2	27
25	1	0	2	0	0	0	1	1	3	0	1	0	2	0	1	1	1	1	0	1	0	0	1	2	0	1	9
26	0	1	3	1	1	2	1	1	3	1	2	0	0	1	2	4	3	1	3	0	0	1	1	3	3	0	22
27	0	2	2	4	0	0	1	0	1	0	1	0	2	1	0	5	2	1	0	1	1	2	1	2	3	2	17
28	1	0	3	1	1	1	1	1	3	1	2	1	2	1	2	3	2	2	3	1	1	1	4	2	3	0	26
29	0	0	2	0	1	2	1	1	3	1	2	1	2	1	2	3	2	3	0	0	1	1	3	2	3	1	22
30	1	0	2	0	1	0	0	1	3	0	1	0	0	1	1	2	1	0	1	1	0	2	2	3	0	0	15
31	0	0	1	0	0	0	0	1	3	1	1	1	2	1	2	3	0	2	0	1	1	3	4	3	1	3	23
32	1	0	3	4	0	0	0	0	3	0	2	0	0	1	1	2	3	2	2	0	1	2	1	0	2	1	17
33	0	0	2	1	0	1	0	1	3	0	2	1	0	1	2	1	3	1	1	1	1	3	0	3	0	0	20
34	1	0	2	0	1	2	0	1	3	0	2	1	2	1	2	5	1	1	0	1	1	2	3	3	0	2	23
35	1	0	2	1	1	0	0	1	3	0	1	1	2	1	1	3	3	1	0	1	1	3	4	0	0	3	20
36	1	0	2	0	0	0	0	0	0	0	1	0	2	1	1	0	0	1	0	1	0	1	2	0	0	0	8
37	1	0	2	0	1	2	1	1	2	0	2	1	2	1	1	4	3	2	0	1	1	2	1	2	0	0	21
38	1	0	2	0	1	0	1	1	3	0	0	0	2	1	1	2	1	2	0	0	0	0	2	0	2	0	9
39	0	0	2	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	3	1	2
40	0	0	3	1	1	1	1	1	3	0	2	1	2	1	2	5	2	2	1	1	1	1	4	2	2	0	25
41	1	0	3	0	1	1	1	1	1	0	2	1	2	1	2	5	2	3	0	1	1	2	4	0	1	0	24
42	0	0	2	0	1	2	0	1	3	0	0	0	2	0	1	3	0	1	0	1	0	1	2	0	0	3	9
43	0	0	2	0	1	2	1	1	0	0	2	0	0	1	2	2	1	1	0	1	1	1	2	0	3	0	14
44	1	0	2	0	0	0	1	1	3	0	1	1	0	0	0	4	1	2	0	0	1	1	3	0	1	2	14
45	0	2	3	1	1	3	1	1	3	1	1	1	2	1	2	5	2	3	1	0	1	0	2	0	2	0	20
46	1	0	2	0	1	0	0	1	2	0	1	0	0	0	0	3	0	0	0	0	0	0	1	0	3	3	5
47	0	0	2	0	1	0	0	3	0	1	1	1	2	0	2	2	3	3	1	0	1	2	4	2	2	0	22
48	0	0	2	1	0	1	1	1	1	0	1	1	2	1	1	3	2	3	0	0	1	2	4	3	2	0	22
49	1	0	2	0	1	0	0	1	1	1	2	1	2	1	2	4	3	1	0	0	1	1	4	3	1	0	24
50	1	3	1	4	0	0	0	0	0	0	2	0	2	0	1	1	0	0	0	0	0	0	1	0	3	3	5
51	1	0	1	0	1	1	0	0	0	1	2	1	2	0	2	5	1	0	2	1	0	1	1	3	3	0	20
52	1	1	0	3	1	1	1	1	3	0	2	1	2	1	1	1	3	2	0	1	1	1	4	3	3	0	21
53	1	0	2	0	1	2	1	1	3	0	0	1	0	0	1	0	0	0	1	0	1	0	1	4	3	2	11
54	0	0	2	0	1	3	1	1	1	0	1	1	2	0	2	2	1	3	0	1	1	1	4	2	1	0	19
55	0	0	2	0	1	2	1	1	2	1	2	1	2	1	2	4	3	3	0	1	1	2	4	3	3	0	28
56	1	0	2	0	1	0	1	1	3	0	0	0	0	1	2	3	1	1	1	0	0	1	3	2	2	0	15
57	1	2	3	1	1	2	1	1	1	1	2	1	2	1	2	5	3	0	0	1	1	2	4	3	1	2	26
58	0	0	2	1	1	2	0	1	3	0	1	0	1	1	1	4	3	2	1	1	1	1	4	2	0	1	22
59	0	0	3	0	1	1	0	1	0	0	1	1	2	1	2	5	3	3	0	1	1	2	4	0	0	0	24
60	0	0	2	0	1	0	0	0	0	0	1	0	0	1	1	4	1	1	1	0	1	1	1	2	3	4	15
61	1	3	3	1	0	0	0	1	1	1	1	1	2	1	2	5	3	1	3	1	0	3	0	3	2	2	25
62	0	0	2	0	0	0	0	0	0	0	2	1	2	1	1	0	3	0	2	1	1	1	0	3	1	0	16
63	0	0	3	1	1	1	1	1	3	1	1	1	2	1	2	3	0	2	3	1	1	3	4	2	0	1	25
64	1	0	3	3	1	1	1	1	3	0	0	1	2	0	0	0	0	0	0	1	1	1	4	2	3	0	10
65	0	0	2	0	1	0	0	0	3	1	2	1	2	1	2	5	3	1	1	1	1	3	1	2	3	0	25
66	1	0	3	0	1	1	1	1	3	0	2	1	2	1	2	4	3	2	0	1	1	2	4	0	2	0	23
67	1	0	2	0	1	2	0	1	3	0	0	1	2	0	2	3	1	1	0	1	0	1	4	0	3	1	14
68	1	0	2	0	0	0	0	1	2	0	1	1	2	1	1	4	2	2	3	1	0	1	2	0	0	0	19
69	0	0	2	0	1	3	1	1	1	0	0	1	4	1	0	1	3	0	0	0	1	2	4	3	2	0	16
70	1	0																									

81	0	0	3	1	0	2	0		3	1	2	0	0	0	1	2	1	3	0	0	1	1	1	4	2	0	0	19
82	1	0	2	0	1	1	1	1	0	0	0	0	2	0	1	2	3	0	0	0	0	1	4	0	3	0	11	
83	1	0	2	0	0	0	0		3	0	0	1	2	1	1	3	0	2	0	1	1	1	4	2	3	0	17	
84	1	0	2	1	0	2	1	1	2	0	1	1	0	1	1	4	3	2	3	1	1	1	2	2	2	0	23	
85	1	1	2	4	1	1	0	1	0	0	2	0	2	1	1	2	0	0	3	1	0	1	4	2	0	1	17	
86	1	0	3	0	0	0	1	0	3	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	3	3	2	
87	1	0	2	0	1	2	0		0	0	0	0	2	0	1	0	0	0	0	1	1	1	4	0	0	3	8	
88	0	0	2	0	1	2	1	1	1	0	2	1	2	1	2	5	3	3	0	0	1	2	3	3	2	0	26	
89	1	0	3	1	0	1	0	1	2	0	0	1	2	1	2	4	3	2	0	0	0	1	2	2	2	1	18	
90	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	3	3	3	
91	0	0	2	0	0	0	0	1	3	1	1	1	2	1	2	4	3	2	0	1	1	2	2	2	3	0	23	
92	1	3	3	4	1	2	1	1	0	0	1	0	2	1	0	3	1	0	1	0	0	0	3	0	0	0	10	
93	1	0	2	1	1	1	0	1	0	0	1	0	2	0	1	5	1	0	0	0	0	0	4	0	1	1	12	
94	1	0	2	0	1	1	0	1	3	1	1	1	2	1	2	3	2	1	0	1	0	1	2	2	1	2	18	
95	0	3	1	2	0	2	0	1	3	1	2	1	2	1	2	3	3	2	3	1	0	1	4	2	0	1	26	
96	0	2	3	1	1	3	0	0	3	0	2	1	1	1	1	5	3	0	2	1	1	2	2	3	2	2	24	
97	0	0	2	4	0	0	0	0	1	0	1	1	0	1	1	2	1	0	0	0	0	1	3	3	3	1	14	
98	0	0	3	1	1	0	0	1	0	1	1	1	2	1	2	1	3	0	3	0	0	0	4	3	1	0	20	
99	0	2	2	0	0	0	1	1	2	1	1	1	1	0	1	4	2	2	0	0	0	2	0	3	2	0	17	
100	0	0	2	0	1	2	1	1	3	0	2	1	2	1	2	4	3	3	1	0	1	2	3	3	0	2	26	
101	1	0	3	4	0	1	0	1	3	0	1	1	2	0	1	3	3	2	2	0	0	1	2	0	3	0	16	
102	1	0	1	0	1	1	0		3	0	0	0	2	0	2	2	2	1	0	0	1	1	3	0	2	1	12	
103	1	0	2	0	1	0	0	1	3	0	0	1	2	0	2	2	2	2	0	1	1	1	4	0	2	0	16	
104	1	0	3	0	0	0	0	0	3	0	0	1	2	0	2	2	2	1	2	0	0	0	1	4	2	0	15	
105	0	1	3	1	1	2	0	0	3	1	2	0	0	1	2	4	3	2	2	0	1	2	4	0	2	0	24	
106	0	0	2	0	1	2	1	1	0	0	1	0	0	0	1	3	3	0	0	1	1	1	4	2	3	3	17	
107	1	0	2	1	1	1	0		3	0	2	0	2	1	0	4	2	1	0	1	1	1	3	0	2	2	16	
108	0	1	2	0	1	0	1	1	3	0	0	0	1	1	1	3	0	0	0	0	0	1	4	0	0	0	10	
109	1	0	2	0	0	3	1	1	3	0	0	1	2	0	2	3	3	3	0	1	1	2	4	2	0	0	22	
110	0	0	3	0	1	1	0		3	1	2	1	2	1	2	5	3	2	3	1	1	2	4	3	3	2	31	
111	1	3	3	1	1	0	1	1	3	0	2	0	2	1	1	0	0	2	1	1	0	3	2	2	0	1	15	
112	0	3	3	1	1	0	0	1	3	0	1	0	0	1	1	0	0	2	1	1	1	3	2	2	0	1	15	
113	0	0	2	0	1	2	1	1	0	1	1	1	2	1	2	3	2	2	0	0	0	0	2	2	3	0	17	
114	1	0	2	0	1	0	0		2	0	2	1	2	1	2	4	2	2	0	1	1	3	3	0	0	0	22	
115	0	0	2	0	1	0	0	0	2	1	2	1	2	1	0	4	3	2	2	1	1	2	3	3	3	0	26	
116	1	0	3	0	1	0	0		3	1	1	1	2	0	1	3	0	1	0	0	1	2	4	3	0	0	18	
117	0	0	2	0	0	0	0		3	1	1	1	2	1	2	5	1	2	0	1	0	1	1	3	3	0	20	
118	1	0	3	1	1	2	0	1	3	1	2	1	1	1	2	5	3	3	3	1	1	2	4	3	3	0	32	
119	1	0	3	1	1	3	1	1	3	0	1	1	2	1	2	5	2	2	0	1	0	0	3	3	2	0	21	
120	0	0	2	0	1	0	1	1	3	0	2	1	2	1	2	2	1	1	2	1	1	1	4	2	2	1	21	
121	1	0	3	1	1	0	1	1	3	0	1	1	0	1	1	4	2	1	0	1	1	1	4	0	2	1	18	
122	0	1	3	1	1	1	0		2	1	0	1	2	0	0	3	3	2	2	1	1	1	2	0	0	0	17	
123	0	1	3	1	1	2	1	1	3	1	2	1	4	1	2	5	1	3	0	0	0	1	3	2	0	0	22	
124	1	0	2	0	1	2	1	1	3	0	0	1	2	1	2	4	3	1	0	1	1	2	4	2	1	0	22	
125	0	0	2	1	0	1	0	1	3	1	2	0	2	1	1	2	2	1	2	0	1	3	4	2	2	0	22	
126	0	0	2	0	1	2	1	1	2	1	2	1	2	1	2	5	0	1	0	0	1	2	4	0	3	0	20	
127	0	0	2	0	1	0	1	1	0	0	1	1	0	1	2	5	3	0	2	1	1	3	2	3	2	1	25	
128	0	1	1	1	1	1	0	1	0	0	1	0	0	0	1	3	2	2	2	0	0	2	0	0	2	3	13	
129	1	0	3	0	0	0	1	0	3	0	2	1	2	1	2	3	0	0	0	1	0	2	4	0	0	2	16	
130	0	0	2	0	1	2	0	1	3	1	1	0	2	0	2	4	1	0	0	0	0	0	2	3	3	0	14	
131	0	0	2	0	1	0	1	1	0	0	2	0	0	1	2	3	0	0	0	1	0	1	1	2	3	1	13	
132	0	0	3	4	1	2	1	1	1	0	2	0	2	1	0	3	3	3	1	1	1	0	1	3	3	0	19	
133	0	0	2	1	1	2	1	1	3	1	2	1	2	1	2	4	2	3	0	1	0	1	4	0	1	0	22	
134	1	0	3	0	1	1	1	1	2	1	2	1	2	1	1	2	2	1	3	0	1	1	4	3	0	1	23	
135	0	0	2	1	1	2	0	0	3	0	2	1	2	1	0	5	2	3	0	0	0	1	3	3	3	3	21	
136	1	0	2	0	1	1	1	1	0	0	1	0	0	1	0	5	2	0	0	0	0	1	2	0	2	3	12	
137	0	0	2	0	1	0	0	0	3	0	1	1	2	1	2	5	1	2	0	0	0	0	3	0	1	1	16	
138	0	1	1	1	1	1	0	0	1	0	1	0	0	0	1	3	2	2	1	0	0	1	2	3	2	1	16	
139	0	0	2	0	1	2	1	1	3	1	2	1	2	1	2	3	0	0	0	0	1	2	4	3	2	3	20	
140	0	0	2	0	1	0	1	1	3	1	1	1	2	1	2	5	0	1	0	1	1	1	3	0	2	2	18	
141	0	0	2	0	1	1	1	1	3	0	1	1	0	1	1	4	2	2	0	0	1	1	4	2	0	0	20	
142	0	0	2	0	0	3	0	1	0	1	2	1	2	1	2	5	3	3	2	1	1	2	4	3	0	0	31	
143	0	0	2	0	1	0	0	1	1	0	1	1	2	1	2	4	3	2	0	1	1	1	3	2	3	1	22	
144	0	0	1	0	1	1	1	1	3	0	2	1	2	1	2	4	2	3	0	0	1	2	1	2	2	0	21	
145	1	0	2	0	1	1	1	1	1	0	1	0	2	0	1	5	0	0	0	0	0	2	4	0	2	3	13	
146	0	0	2	0	1	2	1	1	3	1	2	1	0	1	2	4	3	2	1	1	1	2	4	2	0	0	27	
147	1	0	3	3	1	0	0		3	0	1	1	2	1	2	5	0	1	3	1	1	2	4	3	2	1	25	
148	0	0	2	0	1	0	0		0	0	1	1	2	1	1	2	2	0	1	1	2	3	2	0	1	18		
149	0	0	3	1	1	2	1	1	3	0	2	1	2	1	2	4	3	3	0	0	0	1	4	0	0	2	21	
150	1	0	2	0	1	0	1	1	3	0	0	0	0	0	1	0	2	0	0	1	1	2	2	0	0	0	9	
151	0	0	2	0	1	1																						

## CALCULATION OF FINANCIAL LITERACY SCORE

Financial literacy score is calculated by questions which can denote the financial knowledge of person.

We have included 13 questions whose addition of highest code of per question when calculated comes up to be 33. So, our financial score is out of 33.

For calculation of financial literacy score we took the following questions:

- 1] Do you invest in Stocks or Crypto?
- 2] Do you have monthly RD (Recurring deposit) or SIP (Systematic investment plan)?
- 3] Do you know anything about financial literacy (money management)?
- 4] Do you know interest rate of your bank account?
- 5] Do you know inflation rate?
- 6] Rate yourself for how much you know about the tax you or your parents pay each financial year?
- 7] Relation between your income and expenditure habits?
- 8] How much percent of knowledge do you have about 'Insurance' which protects your finances and investment?
- 9] Do you have any passive income?
- 10] Do you think you have well Diversified Portfolio (investment properly distributed)?
- 11] Do you have any emergency fund?
- 12] How much satisfied you are with your spending habits?
- 13] The first thing you will do in financial emergency?
- 14] How much are you aware about risks while investing?

## STATISTICAL TOOLS

R-Software

Python

MS-Excel

Graphs and charts

Tests

Regression

Data analytics

## INFORMATION ABOUT STATISTICAL TOOLS

After a collection of data second stage is the presentation of data. Classification is the tool of data condensation which makes the analyze easier.

**Frequency:** The number of observations in a class is called as frequency.

**Graphs & Charts:** Graphs are easy to understand and create an effect which lasts for longer time.

Graph represents data and facts in an attractive and impressive manner.

### 1. Simple Bar Diagram:

This is the simplest way of presenting the statistical data classified according to single characteristic. It can be used to present the data like population of different cities, exports of different countries, etc. In general, it can be used for representing any single series but generally it is used to show the categorical series.

### 2. Line Chart:

A line chart is a type of chart that displays information as a series of data points connected by straight line segments. A line chart is a way of visually representing an asset's price history using a single, continuous line.



### 3. Pie Diagram or Pie Chart:

It is a special type of diagram used to represent the whole quantity by a circle and the subdivisions of the whole quantity are shown by the sectors of that circle. This diagram is a two-dimensional diagram. It can be used to represent the subdivision of total budget or total income, etc.

### 4. Rod and Spike Plot:

The graph is used to represent the frequency table or ungrouped frequency distribution of a discrete variable. Spike plot produces a frequency plot for a variable in which the frequencies are depicted as vertical lines from zero.

## CORRELATION:

The extent of linear relationship between the two variables is called as correlation. In order to determine correlation, we require data regarding to two concerned variables. These data are called as bivariate data.

Correlation coefficient lies between -1 to 1.

1. If the correlation is less than 0.3 then there is very poor or negligible correlation.
2. If the correlation lies between 0.3 to 0.8 then there is considerable correlation between the two factors.
3. If the correlation is greater than 0.8 then there is high correlation.

## TESTS:

### 1] Tests for Independence of two attributes:

Suppose that the given data are classified into  $r$  levels of attribute A denoted by  $A_1, A_2, \dots, A_i, \dots, A_r$  and  $s$  levels of attribute B represented by  $B_1, B_2, \dots, B_j, \dots, B_s$ . Then different class frequencies [Cell frequencies] can be represented in the following tabular form.

A \ B	B	B <sub>1</sub> B <sub>2</sub> ..... B <sub>j</sub> ..... B <sub>s</sub>	TOTAL
	A		
A <sub>1</sub>		O <sub>11</sub> O <sub>12</sub> .... O <sub>1j</sub> .... O <sub>1s</sub>	(A <sub>1</sub> )
A <sub>2</sub>		O <sub>21</sub> O <sub>22</sub> .....O <sub>2j</sub> ..... O <sub>2s</sub>	(A <sub>2</sub> )
:		:	:
:		:	:
A <sub>i</sub>		O <sub>i1</sub> O <sub>i2</sub> ..... O <sub>ij</sub> .....O <sub>is</sub>	(A <sub>i</sub> )
:		:	:
:		:	:
A <sub>r</sub>		O <sub>r1</sub> O <sub>r2</sub> ..... O <sub>rj</sub> ..... O <sub>rs</sub>	(A <sub>r</sub> )
TOTAL		(B <sub>1</sub> ) (B <sub>2</sub> ) .... (B <sub>j</sub> ) ..... (B <sub>s</sub> )	N

The above type of table containing r rows and s columns is called as r x s contingency table.

$O_{ij}$  is called observed frequency corresponding to (i,j)<sup>th</sup> cell i=1,2,...r and j=1,2...s.

$$N = \sum_{i=1}^r \sum_{j=1}^s O_{ij} = \text{Total observed frequency.}$$

$$(A_i) = \sum_{j=1}^s O_{ij} = \text{Total of observed frequencies in the } i^{\text{th}} \text{ row. } i=1,2\dots r.$$

$$(B_j) = \sum_{i=1}^r O_{ij} = \text{Total of observed frequencies in } j^{\text{th}} \text{ column. } j=1,2\dots s.$$

We wish to test

Ho: Two attributes A and B are independent

against

H<sub>1</sub>: Two attributes A and B are not independent (associated or disassociated to each other). Under the hypothesis of independence of attributes (i.e., Ho), the expected frequencies corresponding the given observed frequencies are obtained as;

$$e_{ij} = (A_i)(B_j) / N ; i=1,2,\dots,r; j= 1,2,\dots,s$$

(Total of observed frequency in ith row).

$$\text{i.e. } e_{ij} = \frac{\text{(Total of observed frequency in jth column)}}{\text{Grand total of all observed frequencies}}$$

Grand total of all observed frequencies

where  $e_{ij}$  : expected frequency of  $(i,j)^{\text{th}}$  cell. Hence, using the above formula we can find all expected frequencies. Then table of expected frequencies can be prepared as follows:

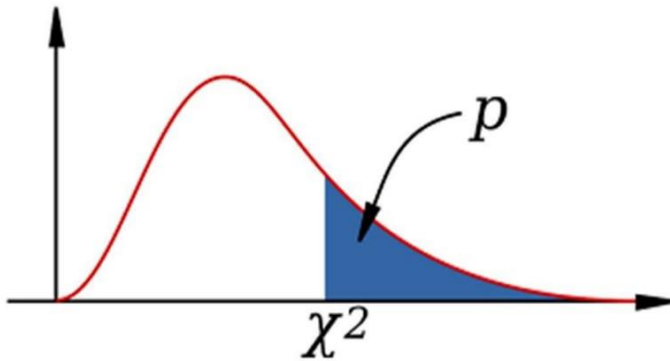
A \ B	B	
	B <sub>1</sub> B <sub>2</sub> .... B <sub>j</sub> ... B <sub>s</sub>	
A <sub>1</sub>	e <sub>11</sub> e <sub>12</sub> ... e <sub>1j</sub> ... e <sub>1s</sub>	
A <sub>2</sub>	e <sub>21</sub> e <sub>22</sub> .... e <sub>2j</sub> .... e <sub>2s</sub>	
:	:	
:	:	
A <sub>i</sub>	e <sub>i1</sub> e <sub>i2</sub> .... e <sub>ij</sub> .... e <sub>is</sub>	
:	:	
:	:	
A <sub>r</sub>	e <sub>r1</sub> e <sub>r2</sub> .... e <sub>rj</sub> .... e <sub>rs</sub>	
	TOTAL	$\sum \sum e_{ij} = N$

Then if  $H_0$  is true, the statistic is

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^r \sum_{j=1}^s \left( \frac{O_{ij}}{E_{ij}} \right)^2 - N$$

follows  $\chi^2$  - distribution with  $(r-1)(s-1)$  d.f. .

The critical region at  $\alpha\%$  level of significance is the shaded region shown below:



(here  $p$  i.e. shaded region is critical region) Thus we reject  $H_0$  at  $\alpha\%$  level of significance if

$\chi^2_{(r-1)(s-1)} \geq \chi^2_{(r-1)(s-1), \alpha}$  and accept  $H_0$  otherwise.

**2] For (2x2) contingency table:**

In particular if we have  $r=2$ ,  $s=2$  i.e., two attributes A and B are at two levels each, then the contingency table will be:

A \ B	B		TOTAL
	B1	B2	
A1	a	b	a + b
A2	c	d	c + d
TOTAL	a + c	b + d	N

Where,

$$N = a + b + c + d$$

Here, test statistic used under  $H_0$  is,

$$\chi^2 = \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

**Decision:** We reject  $H_0$  at  $\alpha\%$  level of significance if

$$\chi^2_{\text{cal}} \geq \chi^2_{\text{table}}$$

Otherwise accept  $H_0$ .

### 3] Testing Equality of two Population Proportions:

Suppose we draw two samples. Suppose these samples give proportions of specific items as  $p_1$  and  $p_2$  respectively. One may be interested in knowing that the population proportions from which these samples are chosen are same. In other words, we want to know whether difference between two sample proportions is negligible and it has arisen merely due to sampling variations.

Let,

$P_1$  = Proportion of specific items in first population.

$P_2$  = Proportion of specific items in second population.

$n_1$  = Size of sample drawn from first sample.

$n_2$  = Size of sample drawn from second sample.

$x_1$  = No. of items of specific type in first sample.

$x_2$  = No. of items of specific type in second sample.

$p_1$  = Proportion of specific items in first sample.

$$= x_1 / n_1$$

$p_2$  = Proportion of specific items in second sample.

$$= x_2 / n_2$$

The hypothesis for such problems will be:

$$\mathbf{H_0 : } P_1 = P_2$$

v/s

$$\mathbf{H_1 : } P_1 \neq P_2$$

Under  $H_0$ ,

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightarrow N(0,1)$$

## Naive Bayes Algorithm:

What is Naive Bayes algorithm?

It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as 'Naive'.

Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods. Bayes theorem provides a way of calculating posterior probability  $P(c/x)$  from  $P(c)$ ,  $P(x)$  and  $P(x/c)$ .

$$\therefore P(c/x) = [P(c) \times P(x/c)] / P(x)$$

## Applications of Naive Bayes Algorithms:

**Real time Prediction:** Naive Bayes is an eager learning classifier and it is sure fast. Thus, it could be used for making predictions in real time.

**Multi class Prediction:** This algorithm is also well known for multi class prediction feature. Here we can predict the probability of multiple classes of target variable.

**Text classification/ Spam Filtering/ Sentiment Analysis:** Naive Bayes classifiers mostly used in text classification (due to better result in multi class problems and independence rule) have higher success rate as compared to other algorithms. As a result, it is widely used in Spam filtering (identify spam e-mail) and Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments).

## BOX PLOT

(Financial Literacy scores)

```
> x=scan("clipboard")
```

Read 170 items

```
> f=fivenum(x)
```

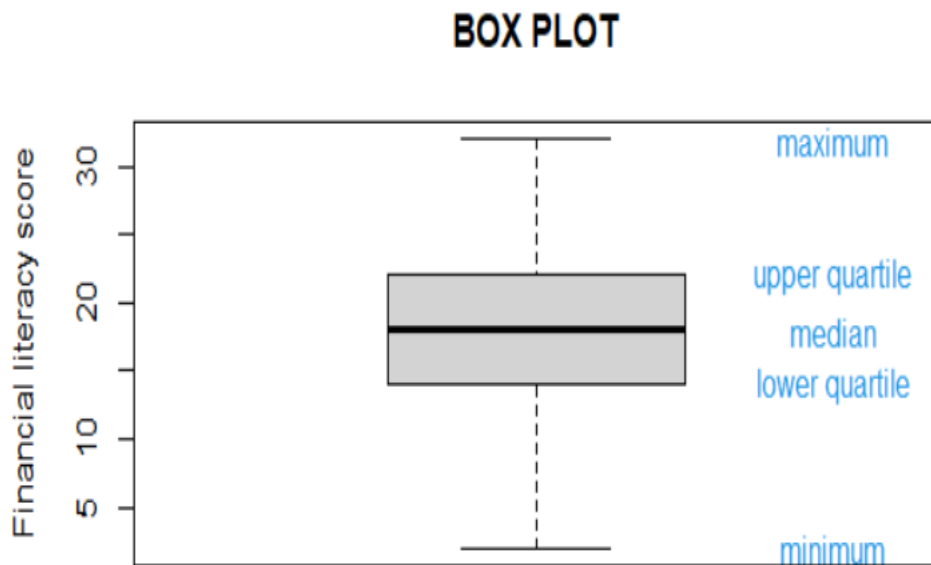
```
> f
```

```
[1] 2 14 18 22 32
```

```
> boxplot(x,main="BOX PLOT",ylab="Financial literacy score")
```

```
> lb=c("minimum","lower quartile","median","upper quartile","maximum")
```

```
> text(rep(1.4,5),f,label=lb)
```



**Interpretation:** From above Box plot we got five number summary which is Minimum score is 2, first quartile=14, Median=18, third quartile=22.5 and Maximum score=32. We can observe that the median is in the middle of the box, and the whiskers are about the same on both sides of the box, then we can say that the data is symmetric.



## ROD AND SPIKE PLOT

(For people rating on knowing about taxation)

**R-commands:**

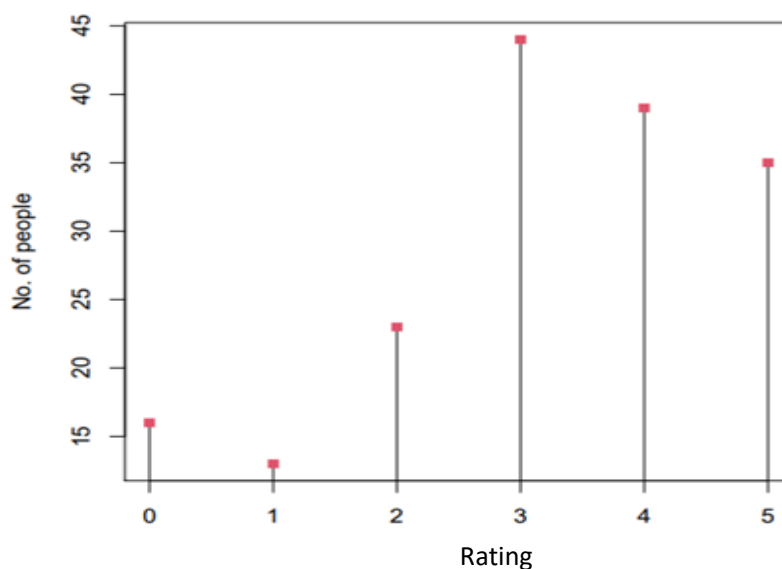
```
> x=0:5
```

```
> y=c(16,13,23,44,39,35)
```

```
> plot(x,y,type="h",xlab="Rate of knowing about TAX",ylab="No. of people")
```

```
> points(x,y,pch=15,col=10)
```

**Rate showing the knowledge about TAX**



- 0** = No knowledge
- 1 & 2** = least knowledge about taxation
- 3 & 4** = Average knowledge about taxation
- 5** = Complete knowledge about taxation

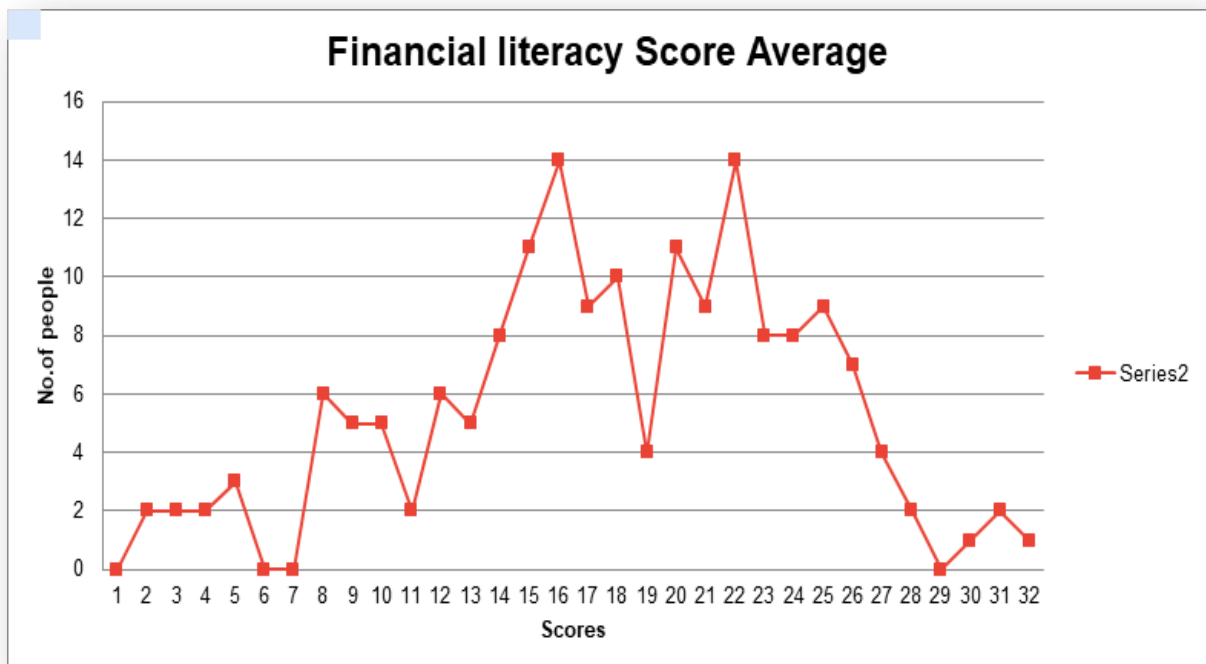
**Interpretation:** The graph shows surprising results as 44 people rate themselves as having average knowledge about taxation, 35 people claims that they have complete knowledge about taxation and only 16 people rate themselves that they are completely unaware about taxation.

## LINE CHART

(Financial Literacy Score V/s Number of people)

Table showing the frequency of people having particular financial score:

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Scores	0	2	2	2	3	0	0	6	5	5	2	6	5	8	11	14
No.	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Scores	9	10	4	11	9	14	8	8	9	7	4	2	0	1	2	1



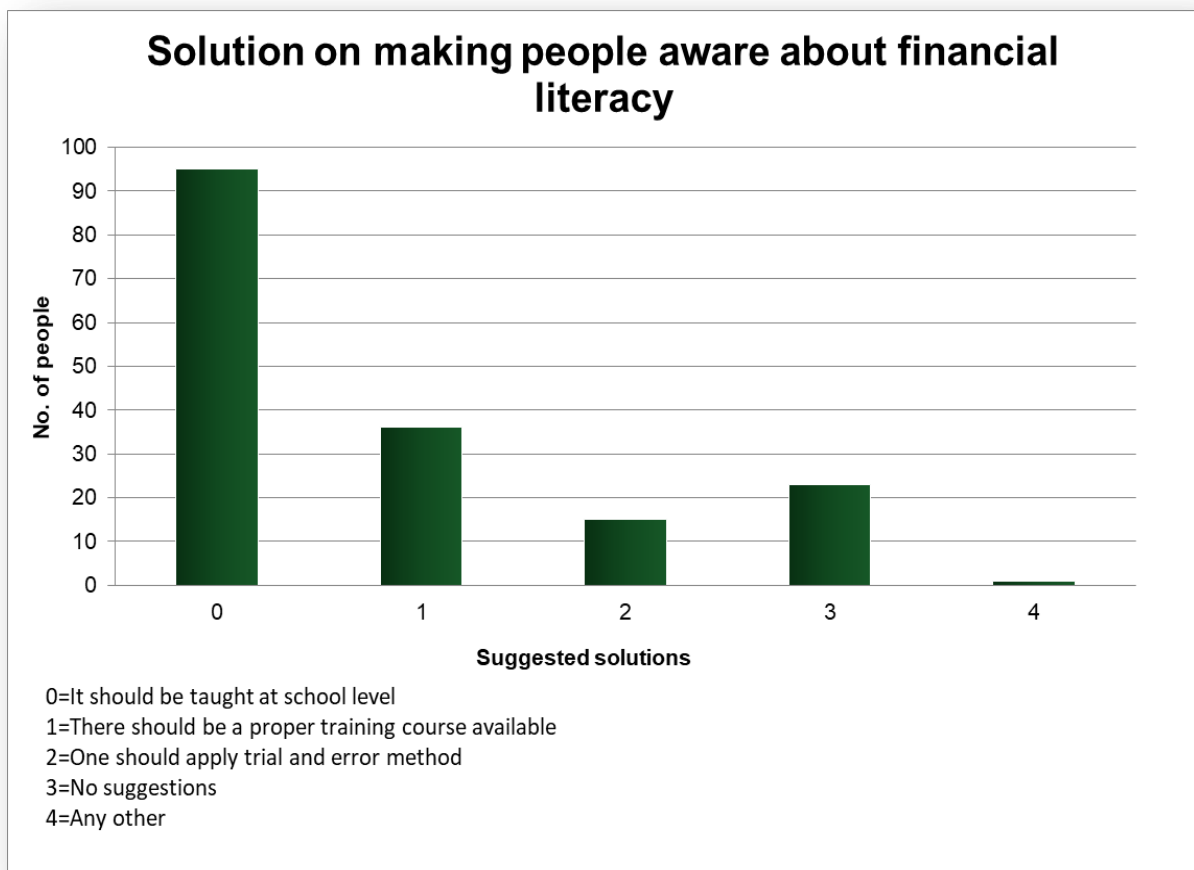
**Interpretation:** It shows that there are 14 people whose financial literacy scores are 16 and 22. Minimum score is observed with 2 people which is 2 and maximum score is observed with 1 person which is 32.

## SIMPLE BAR DIAGRAM

(Solution on making aware about financial literacy)

Pivot table representing the Number of people giving different suggestions:

Row Labels	Count of Q26
0	95
1	36
2	15
3	23
4	1
<b>Grand Total</b>	<b>170</b>



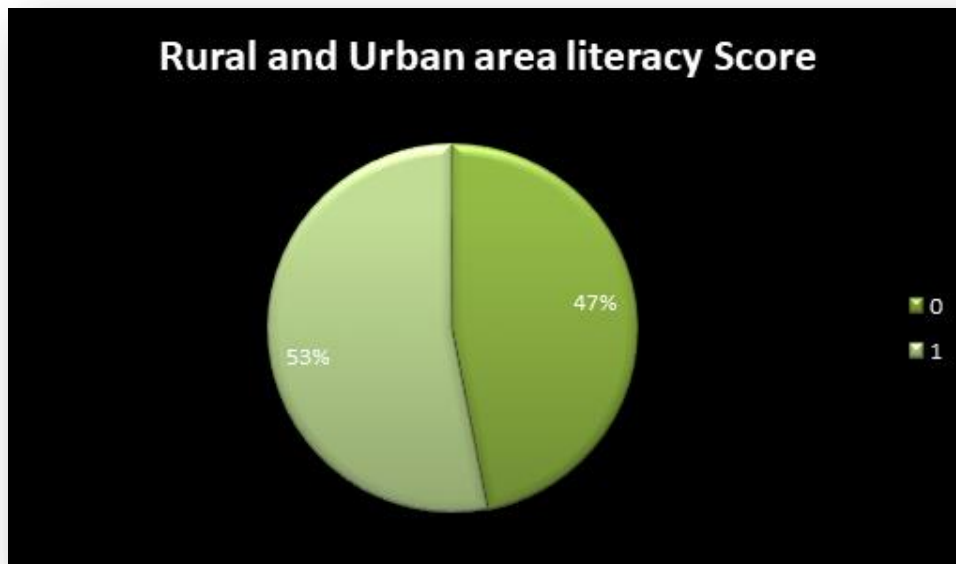
**Interpretation:** It shows that, 95 people suggest that it should be taught at school level about financial literacy and 36 people suggest that there should be a proper training course available.

## PIE CHART

(Financial Literacy score among Urban and Rural)

Pivot Table representing the Sum of Financial Score of Rural and Urban area:

Row Labels	Sum of Financial Score
0 (Rural)	803
1 (Urban)	909
<b>Grand Total</b>	<b>1712</b>



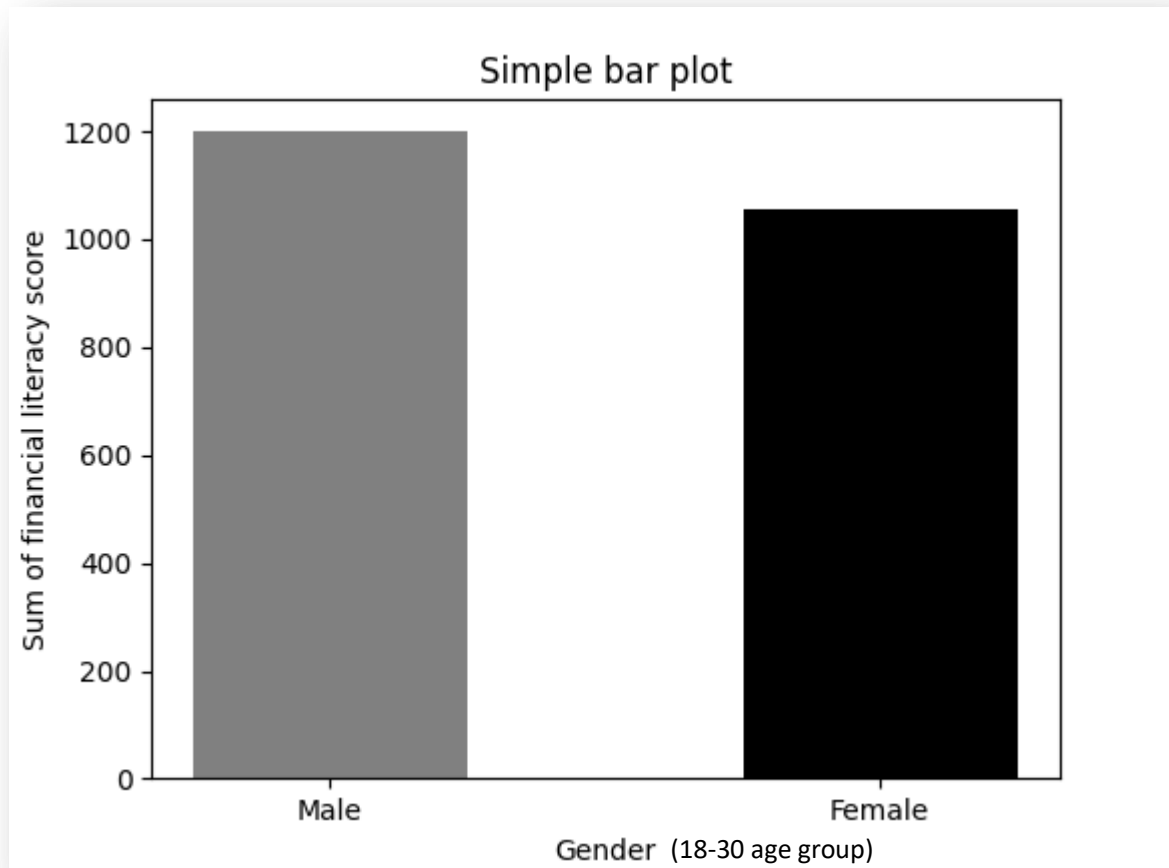
**Interpretation:** Here we can conclude that the financial score of Rural area people is less as compared to Urban area people, in the selected data sample. Therefore, Rural area people are less financial literate than urban area of people.

## SIMPLE BAR DIAGRAM

(Financial literacy in 18-30 age group of Male and Female)

### Python Commands:

```
import matplotlib.pyplot as p
x=["Male","Female"]
x
['Male', 'Female']
y=[1200,1057]
y
[1200, 1057]
p.bar(x,y,width=0.5,color=["grey","black"])
<BarContainer object of 2 artists>
p.xlabel("Gender")
Text(0.5, 0, 'Gender')
p.ylabel("Sum of financial literacy score")
Text(0, 0.5, 'Sum of financial literacy score')
p.title("Simple bar plot")
Text(0.5, 1.0, 'Simple bar plot')
p.show()
```



**Interpretation:** It concludes that the financial literacy among 18-30 age group of Male is more than that of Female in the selected population.

## FINANCIAL LITERACY SCALE

Our Financial Score is out of 33. We have taken the average of total 170 people financial score, which is around 18. Based on the average value we have considered two classes, by which we can give the conclusion about Financial Literacy among the population. Two classes are as below:

### 1] Financial score above 18:

Those individuals who have Financial literacy score above 18 are considered to have high financial literacy.

These individuals understand and effectively use various financial skills, including personal financial management, budgeting, and investing. They have Emergency funds and investment to cope with financial emergency situations.

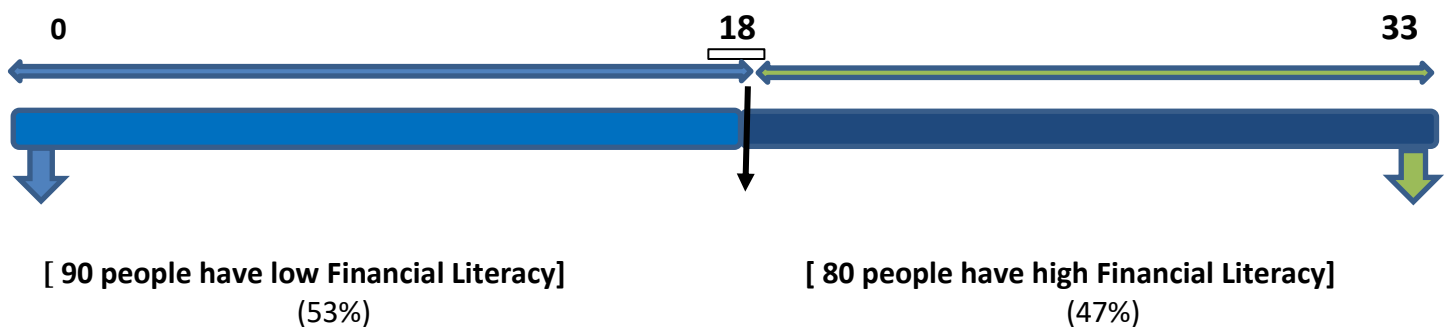
***Out of the 170 population units there are 80 Individuals having score above 18 and are considered to have high financial literacy. Therefore, 47% people have high financial literacy.***

### 2] Financial score equal to or below 18:

Those individuals who have Financial literacy score below 18 are considered to have low financial literacy.

These individuals do not understand and are not able to use various financial skills, including personal financial management, budgeting, and investing. They don't have proper Emergency funds and investments to cope with financial emergency situations.

***Out of the 170 population units there are 90 Individual's having score equal to or below 18 and are considered to have low financially literacy. Therefore, 53% people have low financial literacy.***



## CHI-SQUARE TEST

### 1] DEPENDENCY OF RISK WHILE INVESTING AND WELL DIVERSIFIED PORTFOLIO:

Risk while investing	Well diversified portfolio		Grand Total
	No	Yes	
Aware	15	40	55
Extremely Aware	18	29	47
Have heard about it	18	23	41
Not Aware	23	4	27
<b>Grand Total</b>	<b>74</b>	<b>96</b>	<b>170</b>

Here we want to test

**H0:** Awareness about risk while investing and well diversified portfolio are independent.

v/s

**H1:** Awareness about risk while investing and well diversified portfolio are dependent.

Chi-square test for independency is;

```
> x=c(15,40,18,29,18,23,23,4)
```

```
> m=matrix(x,byrow=T,ncol=2,nrow=4)
```

```
> chisq.test(m,correct=T)
```

Pearson's Chi-squared test

data: m

X-squared = 25.498, df = 3, p-value = 1.215e-05

**Criteria:** Accept H0 if p-value is greater than or equal to level of significance ( $\alpha$ ); otherwise reject it.

Here, p-value = 1.215e-05 < l.o.s=0.05

**Decision:** We may reject H0 at 5% l.o.s. i.e., we may accept H1

**Conclusion:** Awareness about risk while investing and well diversified portfolio are dependent.



## 2] DEPENDENCY OF ANNUAL INCOME AND AVAILABILITY OF EMERGENCY FUND

Annual income (Parents/Self)	Emergency fund		Grand Total
	No	Yes	
2-5 Lakhs	19	25	44
5-10 Lakhs	16	26	42
Above 10 Lakhs	3	8	11
Below 2 Lakhs	39	34	73
<b>Grand Total</b>	<b>77</b>	<b>93</b>	<b>170</b>

Here we want to test

**H<sub>0</sub>:** Annual income and availability of emergency funds are independent.

v/s

**H<sub>1</sub>:** Annual income and availability of emergency funds are dependent.

Chi-square test for independency is;

```
> x=c(19,25,16,26,3,8,39,34)
```

```
> m=matrix(x,byrow=T,ncol=2,nrow=4)
```

```
> chisq.test(m,correct=T)
```

Pearson's Chi-squared test

data: m

X-squared = 4.347, df = 3, p-value = 0.2264

**Criteria:** Accept H<sub>0</sub> if p-value is greater than or equal to level of significance ( $\alpha$ ); otherwise reject it.

Here, p-value = 0.2264 > l.o.s=0.05

**Decision:** We may accept H<sub>0</sub> at 5% l.o.s. i.e., we may reject H<sub>1</sub>

**Conclusion:** Annual income and availability of emergency fund are independent.

## PROPORTION TEST

**1] We have to check whether the proportion of financial literacy is equal or different in males and females (18-30 age group)**

X1=Number of males having high financial literacy

X2=Number of females having high financial literacy

N1=Total number of males

N2=Total number of females

Here we want to test

**H<sub>0</sub>:**  $P_1 = P_2$  v/s

**H<sub>1</sub> :**  $P_1 > P_2$

**H<sub>0</sub>:** Proportion of financial literacy in males and females is same v/s

**H<sub>1</sub>:** Proportion of financial literacy in males is greater than females.

Test for proportion is:

```
> x=c(37,22)
```

```
> x
```

```
[1] 37 22
```

```
> n=c(71,64)
```

```
> n
```

```
[1] 71 64
```

```
> prop.test(x,n,alt="greater")
```

2-sample test for equality of proportions with continuity correction

data: x out of n

X-squared = 3.6135, df = 1, p-value = 0.02866

alternative hypothesis: greater

95 percent confidence interval:

```
0.02451476 1.00000000
```

sample estimates:

```
prop 1 prop 2
```

```
0.5211268 0.3437500
```

**Criteria:** Accept  $H_0$  if p-value is greater than or equal to level of significance ( $\alpha$ ); otherwise reject it.

Here, p-value=0.02866 < l.o.s=0.05

**Decision:** We may reject  $H_0$  at 5% l.o.s i.e., accept  $H_1$

**Conclusion:** Proportion of financial literacy in males is greater than females (18-30 age group).

**2] We have to check whether the proportion of financial literacy is equal or different in urban and rural area people:**

$X_1$ =Number of urban area people having high financial literacy

$X_2$ =Number of rural area people having high financial literacy

$N_1$ =Total number of urban area people

$N_2$ =Total number of rural area people

Here we want to test

**$H_0$  :**  $P_1 = P_2$

v/s

**$H_1$  :**  $P_1 > P_2$

**$H_0$ :** Proportion of financial literacy in urban is equal to rural area people

v/s

**$H_1$ :** Proportion of financial literacy in urban is greater than rural area people.

Test for proportion is:

> x=c(61,19)

> x

[1] 61 19

> n=c(114,56)

> n

[1] 114 56

```
> prop.test(x,n,alt="greater")
```

2-sample test for equality of proportions with continuity correction

data: x out of n

X-squared = 5.0197, df = 1, p-value = 0.01253

alternative hypothesis: greater

95 percent confidence interval:

0.05312587 1.00000000

sample estimates:

prop 1 prop 2

0.5350877 0.3392857

**Criteria:** Accept  $H_0$  if p-value is greater than or equal to level of significance ( $\alpha$ ); otherwise reject it.

Here, p-value= 0.01253 < l.o.s=0.05

**Decision:** We may reject  $H_0$  at 5% l.o.s i.e., accept  $H_1$

**Conclusion:** Proportion of financial literacy in urban is greater than rural area people.

## LOGISTIC REGRESSION

Y=Dependent variable= Well diversified portfolio

X= Independent variable= Awareness about risk while investing

( We have considered 'Not aware' and ' Have heard ' about it as 0 (Not aware about risk while investing) and 'Aware' and 'Extremely aware' as 1 (Aware about risk while investing)

```
> x=scan("clipboard")
```

Read 170 items

```
> x
```

```
> y=scan("clipboard")
```

Read 170 items

```
> y
```

```
> d=data.frame(y,x)
```

```
> d
```

```
> logistic=glm(y~x,data=d,family ="binomial")
```

```
> logistic
```

Call: glm(formula = y ~ x, family = "binomial", data = d)

Coefficients:

(Intercept)	x
-0.4177	1.1553

Degrees of Freedom: 169 Total (i.e. Null); 168 Residual

Null Deviance: 232.8

Residual Deviance: 219.8 AIC: 223.8

```
> summary(logistic)
```

Call:

```
glm(formula = y ~ x, family = "binomial", data = d)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5023	-1.0059	0.8842	0.8842	1.3592

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.4177	0.2478	- 1.685	0.091898 .
x	1.1553	0.3259	3.545	0.000393 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 232.81 on 169 degrees of freedom

Residual deviance: 219.78 on 168 degrees of freedom  
AIC: 223.78

Number of Fisher Scoring iterations: 4

From the output we can observe that,  
 $\beta_0 = -0.4177$   $\beta_1 = 1.1553$

**Logistic regression model:**

$$y = \exp \{-0.4177 + 1.1553 * x\} / [1 + \exp \{-0.4177 + 1.1553 * x\}]$$

$\beta_1 = 1.1553$  implies that  $\psi = \text{odds ratio} = e^{\beta_1} = 3.174 > 1$

When the odds ratio is greater than 1, it describes a positive relationship. The positive relationship means that as awareness about risk while investing "increases," the odds of well diversified portfolio increases. This can be interpreted that being in the (1) group, or being aware about risk while investing, puts you at 3 times greater odds of having well diversified portfolio.

Now, we wish to test significance of regression coefficient( $\beta_1$ )  
i.e., we wish to test

**H<sub>0</sub>:**  $\beta_1 = 0$  (Regressor x is insignificant)

V|s

**H<sub>1</sub>:**  $\beta_1 \neq 0$  (Regressor x is significant)

From the output,

**Null deviance**=Deviance under the null hypothesis  
=Deviance for the model excluding regressor X  
= 232.81

**Residual deviance**= Deviance for the model including regressor X  
= 219.78

Hence,  $G = \text{Null deviance} - \text{Residual deviance}$   
 $G = 13.03$

$$\chi^2_{(1,0.05)} = 3.841$$

**Criteria:** We reject  $H_0$  (null hypothesis) at 5% l.o.s. when  $G \geq \chi^2_{(1,0.05)}$

Here,  $G = 13.03 \geq \chi^2_{(1,0.05)} = 3.841$

**Decision:** We may reject  $H_0$  at 5% l.o.s.

**Conclusion:** We may conclude that regressor x is significant and there is relationship between two variables. (Awareness about risk while investing is significant and there is relationship between awareness about risk while investing and well diversified portfolio)

- **PREDICTION:**

**1] Lets take awareness level as 1(Aware):**

Equation:  $y = \exp \{-0.4177 + 1.1553 \cdot x\} / [1 + \exp\{-0.4177 + 1.1553 \cdot x\}]$

```
> p=exp(-0.4177+1.1553*1)
```

```
> p
```

```
[1] 2.090911
```

```
> p1=p/(1+p)
```

```
> p1
```

```
[1] 0.6764708
```

**Interpretation:** If the individual is Aware (with code 1) then we can say that there is 67% chance that he has well Diversified portfolio.

**2] Lets take awareness level as 0 (Not Aware):**

Equation:  $y = \exp \{-0.4177 + 1.1553 \cdot x\} / [1 + \exp\{-0.4177 + 1.1553 \cdot x\}]$

```
> p=exp(-0.4177+1.1553*0)
```

```
> p
```

```
[1] 0.6585598
```

```
> p1=p/(1+p)
```

```
> p1
```

```
[1] 0.3970673
```

**Interpretation:** If the individual is Unaware (with code 0) then we can say that there is only 39% chance that he has well Diversified portfolio.

- **CONFUSION MATRIX:**

		<u>Actual Results</u>	
		Positive	Negative
<u>Model Predictions</u>	Positive	<u><b>True Positive</b></u> The number of observations the model predicted were positive that were actually positive	<u><b>False Positive</b></u> The number of observations the model predicted were positive that were actually negative
	Negative	<u><b>False Negative</b></u> The number of observations the model predicted were negative that were actually positive	<u><b>True Negative</b></u> The number of observations the model predicted were negative that were actually negative

```
> pmodel=predict(logistic)
> pmodel
> t=table(pmodel>0.5,d$y)
> t
```

```
      0   1
```

```
FALSE 41 27
```

```
TRUE  33 69
```

```
> TP=69
```

```
> TP
```

```
[1] 69
```

```
> FP=33
```

```
> FP
```

```
[1] 33
```

```
> FN=27
```

```
> FN
```

```
[1] 27
```



> TN=41

> TN

[1] 41

> accuracy=(TP+TN)/(TP+TN+FP+FN)

> accuracy

[1] 0.6470588

### Interpretation:

True Positives (TP) : The model predicted positive and the actual label is positive =69

True Negative (TN) : The model predicted negative and the actual label is negative =41

False Positive (FP) : The model predicted positive and the actual label was negative =33

False Negative (FN): The model predicted negative and the actual label was positive =27

When we talk of accuracy, we are referring to how close the measured value (what we are predicting) is to the known values. To calculate the accuracy of a model from our confusion matrix we would sum the correct answers (TP + TN) and divide it by the total number of instances (TP + TN + FP + FN).

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{False Positives} + \text{True Negatives} + \text{False Negatives}}$$

The Accuracy of our model is 64%

## NAIVE BAYES ALGORITHM

Financial literacy scores above average (18) are considered as high financially literate people (coded as 'Yes')

Financial literacy scores equal or below average (18) are considered as low financially literate people (coded as 'No')

X1, X2, X3 and X4 are not different variables but levels of one single variable that is

### Budgeting

X1=Half Yearly

X2=Monthly

X3=No Financial budget

X4= Yearly

### • BUDGETING AND FINANCIAL LITERACY:

Count of Financial literacy	Financial literacy			
	No	Yes	Grand Total	
<b>How do you make your financial Budget?</b>				
Half yearly	6	10	16	<b>X1</b>
Monthly	41	53	94	<b>X2</b>
No Financial Budget	33	6	39	<b>X3</b>
Yearly	10	11	21	<b>X4</b>
<b>Grand Total</b>	<b>90</b>	<b>80</b>	<b>170</b>	

P(Yes)	0.470588235
P(No)	0.529411765

P(X1/Yes)	0.625
P(X2/Yes)	0.563829787
P(X3/Yes)	0.153846154
P(X4/Yes)	0.523809524
P(X1/No)	0.6
P(X2/No)	0.773584906
P(X3/No)	5.5
P(X4/No)	0.909090909

P(Yes/X1)	0.480769231
P(Yes/X2)	0.393155814
P(Yes/X3)	0.024260804
P(Yes/X4)	0.33869839
P(No/X1)	0.519230769
P(No/X2)	0.606844186
P(No/X3)	0.975739196
P(No/X4)	0.66130161

**Interpretation:** The people who do half yearly budgeting (X1) have P(Yes/X1) more than others, so we can conclude that those who do half yearly budgeting are more financially literate. While, people who don't do budgeting (X3) are less financially literate as they have more P(No/X3) than other variables.

## **SURVEY EXPERIENCE**

Collecting the data for any type of project is the main part of that project. It may be a time-consuming task to collect the data by primary data collection method.

For this T.Y.BSc Statistics project, we adopt online data collection technique to collect the data.

Firstly, we created our questionnaire by using Google Forms.

Then we circulated it to respective respondents by social media so that they can be able to fill up their valuable responses.

In rural area we took efforts by individually explaining the questions and then marking down the responses (Village name: Malvadi,Pimpalgoan)

Some of the people were also interested in knowing the results of the project and encouraged us to perform this kind of projects.

Overall, our experience was good and satisfactory.

## CONCLUSION

From overall project we can conclude the following things:

- Out of the 170 population units there are 80 Individuals having score above 18 and are considered to have high financial literacy. Therefore, 47% people have high financial literacy. 90 Individual's having score equal to or below 18 and are considered to have low financial literacy. Therefore, 53% people have low financial literacy. This concludes that there is high time to take initiative to spread awareness about financial literacy as more than half of the population has very low financial literacy.
- According to statistical analysis we found out Males are more financially literate than Females. Which put our society in inequality unrest so we need to put light on providing same opportunities to all genders.
- According to statistical analysis we found out urban area people are more financially literate than rural area people. So, spreading awareness in rural area is important to grow our economy as whole.
- Financial Literacy is dependent on factors like Annual income and educational qualification.
- There is strong relationship between Awareness about risk while investing and well diversified portfolio. By confusion matrix we found out that accuracy of fitted logistic regression model is 64% .So, we as a society can take initiatives to spread awareness about risk while investing in various financial instruments.
- By Naïve baye's algorithm we found out that people who do half yearly budgeting have more probability of having high financial literacy.
- From the project we can conclude that Financial Literacy plays an important role in everyone's life to cope up with financial emergency. It is important thing as it can protect you from poverty trap and can provide you stable life with proper finances.

## HOW TO INCREASE FINANCIAL LITERACY

- 1] It should be taught at school level
- 2] Arranging various training courses
- 3] Increase familiarity with your own finances
- 4] Use Financial Management tools
- 5] Take financial advice from experts
- 6] Books and magazines remain a great way to learn about financial topics.

## LIMITATIONS

1. This data and overall survey is limited for only 170 observations that we have taken into account. Hence, whatever results and conclusion we obtained are only for these respective responds.
2. The results may differ if the numbers of observations either increases or decreases.
3. The sample size is not too large. Practically, this type of survey should contain large number of samples.
4. The results may vary from sample to sample.
5. As we got more data for age group 18-30 and very less for other age groups hence, we were not able to give proper conclusion for the whole population.
6. The results and conclusion of this project are limited only for this project.

## REFERENCES

**Links** of the online sites used in project:

<https://www.financialeducatorsCouncil.org/>

[https://en.wikipedia.org/wiki/Financial\\_literacy](https://en.wikipedia.org/wiki/Financial_literacy)

<http://www.thepennyhoarder.com/>

<https://www.champlain.edu/centers-of-excellence/>

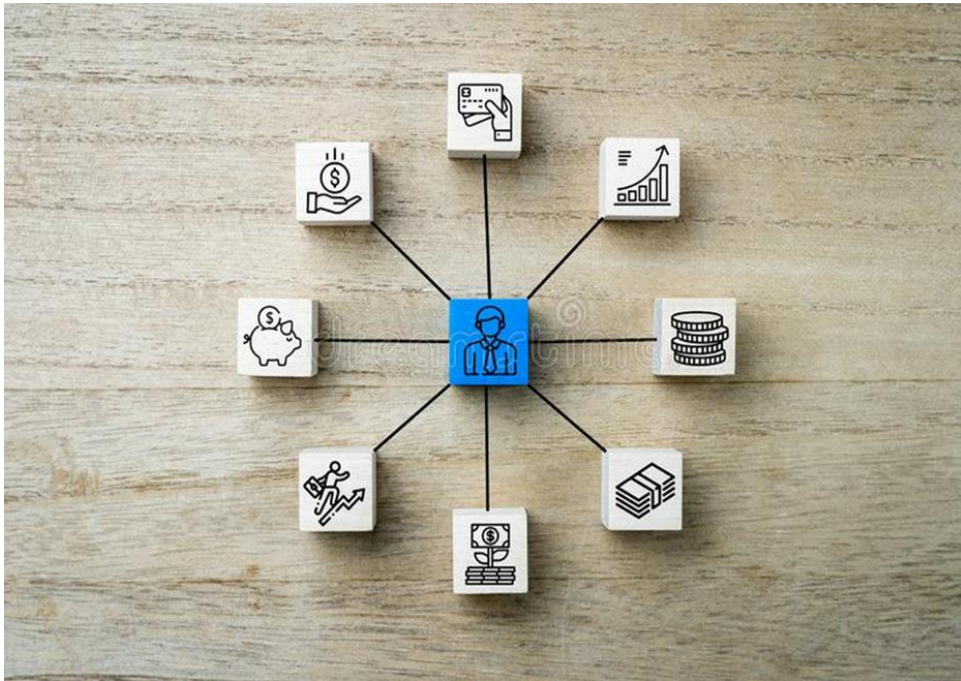
### **Bibliography:**

- 1) Descriptive Statistics
- 2) Sampling Distributions and Inference
- 3) Statistical Methods and Use of R-Software
- 4) (NISM) (X1 &X2) National Institutes of Securities Markets
- 5) Introduction to regression analysis

### • **Questionnaire Link:**

[https://docs.google.com/forms/d/1-WVeyfRsi9HKRXEDz9HC1hrEif2lp7\\_XUuyeCPtgKbs/edit](https://docs.google.com/forms/d/1-WVeyfRsi9HKRXEDz9HC1hrEif2lp7_XUuyeCPtgKbs/edit)

# THANK YOU!



**“Never spend your money before you have earned it.”**