

# Statistical and Sentiment Analysis On Investment Pattern of Indians

1<sup>st</sup> Shruti Jain  
Dept. of IT  
MPSTME, NMIMS  
Mumbai, India  
shrutijain99917@gmail.com

2<sup>nd</sup> Jash Jain  
Dept. of IT  
MPSTME, NMIMS  
Mumbai, India  
jainjash72@gmail.com

3<sup>rd</sup> Garima Merani  
Dept. of IT  
MPSTME, NMIMS  
Mumbai, India  
garima.merani48@gmail.com

4<sup>th</sup> Jash Patel  
Dept. of IT  
MPSTME, NMIMS  
Mumbai, India  
pateljash2102@gmail.com

**Abstract**—Investments and the knowledge surrounding them have started to pick up the pace in India. As more and more people learn about it and experiment, opinions have begun to formulate around it. In this paper, we conducted a survey to get the idea of people's perspectives with regards to their knowledge on Investment and the different types of Investments people are making. The major purpose of this study is to analyze the survey responses and draw conclusions using data analysis techniques. We performed Sentiment Analysis on the survey participants to identify why people invest, why some do not invest, and what people think about NFT's and Crypto. With this paper, we hope to have a clearer idea on the Investment and Stocks pattern of Indians.

**Index Terms**—Investment Pattern, Statistical Analysis, Sentiment Analysis

## I. INTRODUCTION

“The key is not in spending time, but in investing it.” Investing involves purchasing assets that appreciate in value over time and generate returns in the form of income payments or capital gains. It is a way to set aside money and have that money work itself and grow for you to reap benefits in the future. Investments are basically a great way to generate income and assets. In recent years, India's economy has increased by quite a margin; more and more people are now trying to find ways to secure or grow their money/assets - which has led to an increase in the knowledge of stock trading, real estate investments, mutual funds and many more. Although to invest in many different avenues, one needs to keep a set amount fixed or have a certain amount of savings. This amount depends from person to person and changes with age, profession, income. In this paper, the main objective is to understand the Investment pattern of Indians and their opinions on various investment-related aspects. We have done an analysis and visualization of 200 responses given through a survey form containing 12 questions. Details like name, age group, gender, profession, whether or not they do investments, types of investments made, level of knowledge and expectations from investments, why or why not invest, opinions on NFT's and Crypto were received from the survey. The form was sent to all age groups, gender and people of different professions in

India. This data was then analyzed using graphs, charts, word clouds, sentiment analysis.

## II. RELATED WORKS

Numerous analyses on the various market outlooks have been done in the past. These studies primarily focus on the various capital market instruments, shareholding behavior, new issue market, market efficiency, risk and return, and mutual fund performance. On the other hand, little research has been done on investor sentiment.

[1] The resulting extra desire to buy the association's stock, beyond allure anticipated commercial profits and danger, is specified as evidence of either an individual investor's impact towards a corporation. [2] A work is created to label the tastes and inclinations of a sample of investors selected carelessly by not enough big people. Investment nature of an individual class of family is different from another class of crowd, it may be in the form of risk understanding level, knowledge of miscellaneous properties. [3] The paper evaluates the emotion of the investors because we can study their effect on the sure clock worn on body symbols. It's satisfactory to believe that investors are affected by their desires. Because the peep ideas are so essential, it's important to guarantee that investors aren't under reacting or overreacting to specific timekeeping device characters. [4] Nowadays, mutual funds are the choice that young people like the most. The stock market is flourishing. The Systematic Investment Plan (SIP) is a popular option for young investors to invest in mutual funds.

[5] Market sentiments are a good place to start when looking at the present trend in a company's stock because they can help predict future values and close the gap between expected and actual values. In technical analysis, it is critical to assess the accuracy of various methodologies and determine the suitability of selecting the best algorithm that produces findings that are closest to the actual value for which comparison research is conducted. [6] This study examines investor behavior to determine the best investment opportunities in India. The investment strategy is a plan designed to assist investors in selecting the best suitable investment portfolio that will enable them to attain their financial objectives within a specific time frame. [7] This research is both descriptive and analytical.

120 female respondents in Chennai were chosen using simple random sampling. Using a carefully constructed questionnaire, primary data was gathered. ANOVA, factor analysis, and percentage analysis were the statistical methods employed in the study. [8] Even the non-working women are interested in investment activity. Most of the women have income less than 25000. Most of the women are interested in investing a small part in investment avenues. Most women invest in short term plans. Women are afraid about the security of their investment, therefore most of them are making investments in bank deposits only. [9] The objective of the study was to determine the relationship between the income and investment pattern among employees. The study was conducted on women across both government and private sectors in the city of Coimbatore.

[10] To review the investment patterns of different classes of people. To give major findings in the investment pattern of different classes of people. To give conclusions based on research. [11] The study will be useful in helping to better understand what an investor genuinely expects from a potential investment. In order to do additional research, it will be helpful to comprehend the preferences and behaviors of various class individuals in India. The investors will learn about the numerous financial services the business offers through it. [12] Predicting the stock market is not an easy undertaking. The stock market is a public exchange for exchanging firm shares and derivatives at an agreed price. The trading community employs a variety of methods for making predictions. This essay reviews the literature on artificial neural networks' usage in stock market forecasting and concludes that these networks are particularly effective in forecasting global stock markets.

### III. METHODOLOGY

We had shared a google form with people so that we could get certain inputs regarding investment opinions. 200 people filled this form. We analyzed the data that we received through the form responses. There were 12 questions in the form. We analyzed the important attributes individually and derived conclusions based on the said analysis. We used graphs and charts for the statistical analysis, and word clouds for the sentiment analysis.

#### A. Implementation and Methods

Fig. 1 depicts the system architecture. The implementation was as follows:

**Data Collection:** The data was collected through a survey conducted by sharing a google form questionnaire. 200 people took part in the survey.

**Data Validation:** Most of the answers in the questionnaire were multiple choice based. This helped in avoiding ambiguity of answers. 4 answers were text based. Not all the questions were compulsory. There were 2 sections created, one for people who do invest and one for people who do not.

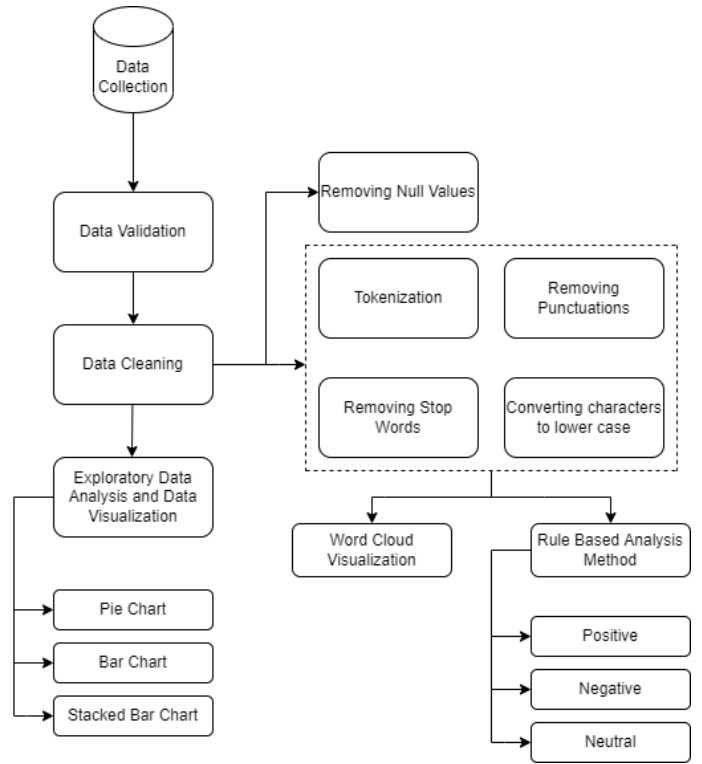


Fig. 1. System Architecture

**Data Cleaning:** For the statistical analysis, we had to clean and preprocess the data in order to remove any null values. One hot encoding was performed on a few columns. All the characters were converted to lowercase. All the unnecessary data was removed. For the sentiment analysis, the data preprocessing was done by tokenization, removing stop words and removing punctuations.

**Exploratory Data Analysis and Data Visualization:** Once all the data was cleaned and preprocessed, we analyzed the statistics through graphs and charts. Pie charts and bar graphs were made using tableau in order to do so. This was done for all the important attributes.

**Sentiment Analysis:** We analyzed the text based answers for the sentiment analysis. We made word clouds based on the responses. We also used rule based analysis method textblob to analyze the sentiment of the answers. This was done by classifying the said answers into 3 categories; positive, negative and neutral.

### B. Attribute Description

TABLE I  
ATTRIBUTE DESCRIPTION TABLE

Attribute	Description
Name	Name of the Respondent (optional)
Age Group	Age of the Respondent
Gender	Gender of the Respondent
Profession	Profession of the Respondent
Do you make investments?	Binary Attribute (Yes or No)
What kind of investments do you make?	Nominal Attribute (Gold, PPF, Mutual Funds, Equity Shares, FDs, Real Estate, EPF, Other)
On a scale of 1-5, where do you consider your knowledge of investments?	Ordinal Attribute (1 to 5)
What are your expectations from the investments you are currently making?	Nominal Attribute (Short Term Return, Long Term Return, Other)
What is your main reason for investing? (one word/one line answer)	Short Answer Text
Why do you not invest in other options?	Short Answer Text
What are your views on the recent boom in Crypto and NFTs?	Long Answer Text
Why do you not invest?	Short Answer Text

### C. Libraries and Tools

The following tools were used for the statistical and sentiment analysis:

Python Programming Language, Google Colaboratory and Google Sheets.

The following libraries and packages were used:

Numpy, Pandas, Seaborn, re, NLTK, Matplotlib, Wordcloud  
and textBlob

## IV. RESULTS

### A. Visualization

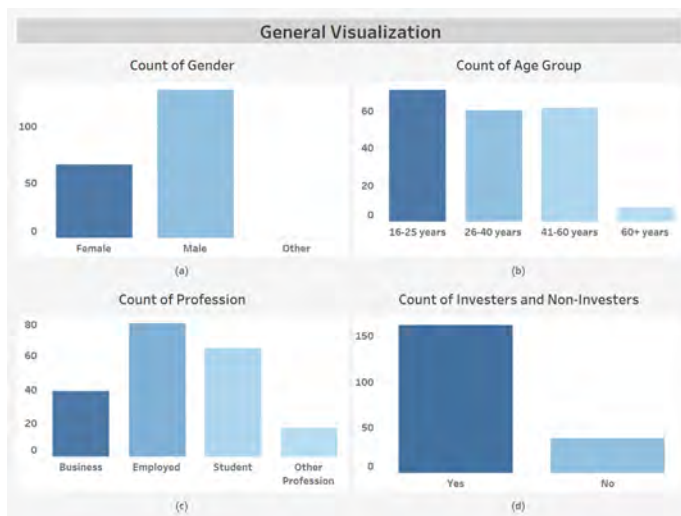


Fig. 2. General Visualization

The responses of the survey are plotted in Fig. 2. The following observations were made as a result of these responses:

Fig. 2(a) It was discovered that the majority of those who participated in the survey were men.

Fig. 2(b) It is observed that most people belong to age age group 16-25 , 41-60 years.

Fig. 2(c) Most people are service class / job oriented or students.

Fig. 2(d) Out of people who filled the form, more than 150 were investors, i.e. they said when they were asked whether they invest or not.

After general visualization, the dataset was further divided into 2 subdatasets, one containing responses of the people who make investments and the other one of those who don't invest. Rows where the people who don't invest and haven't given a reason for their answer were dropped since they did not contribute to our analysis.

One hot encoding was performed on the 'What kind of investments do you make?', 'What are your expectations from the investments you are currently making' and 'Profession' column in order to perform below given analysis

### B. Statistical Analysis

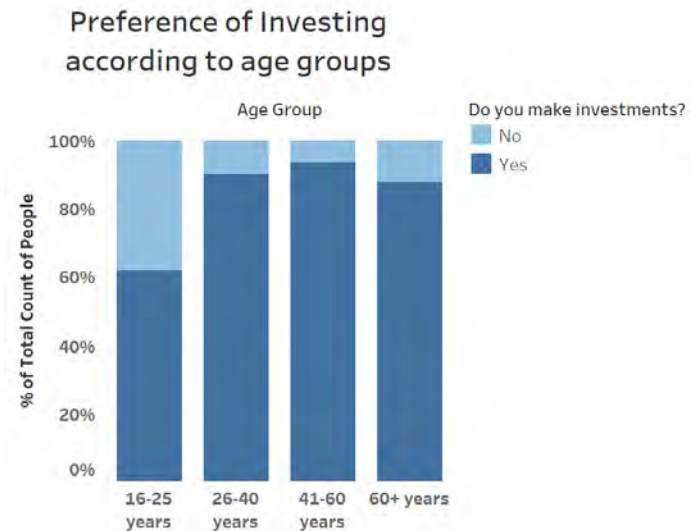


Fig. 3. Investing according to age groups

Fig.3 shows that more than 80% of people of the age group 26-60 years invest and a very small fraction does not invest.

Fig.4 shows that the majority of the investors are males, although there is quite a significant percentage of female investors as well.



Fig. 4. Gender percentage of Investors

Fig.5 shows that out of the people who filled the form and said that they do not invest most of them are males again.

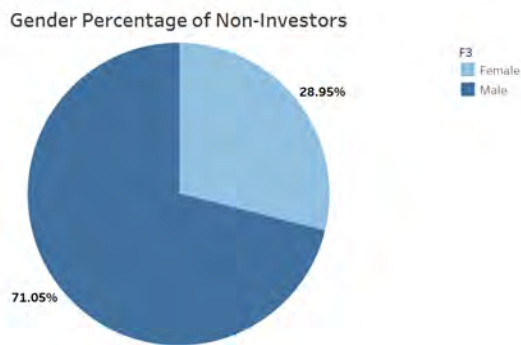


Fig. 5. Gender percentage of Non-Investors

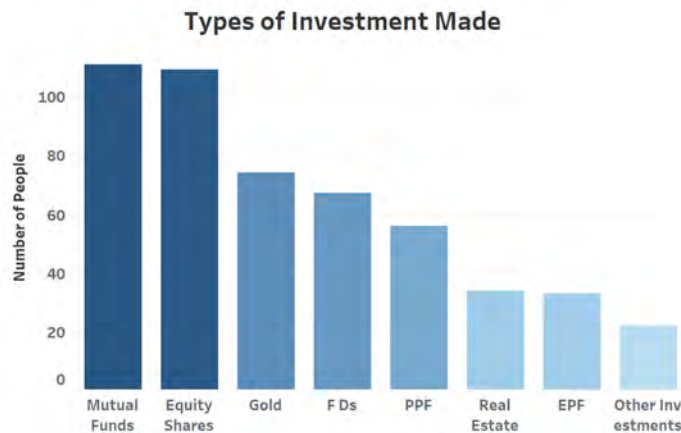


Fig. 6. Different types of Investments

Fig. 6 shows that most of the investors tend to invest in Mutual Fund and Equity Shares. The number of people investing in Real Estate and EPF is significantly less.

Fig. 7 shows the knowledge all investors have on a scale of 1-5. Most of the investors possess an average amount of knowledge regarding investments in general.

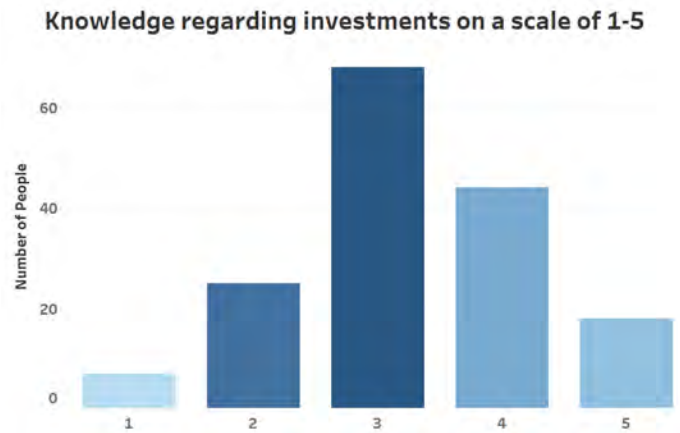


Fig. 7. Knowledge regarding Investments

### Where do people with appropriate knowledge invest

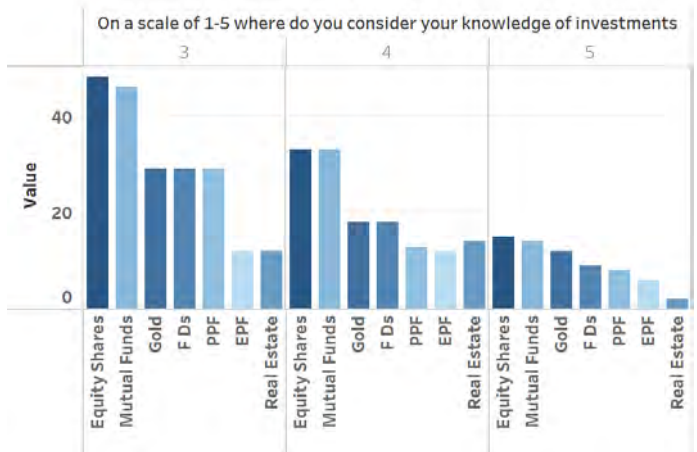


Fig. 8. Major investing areas

Fig. 8 shows that the most popular investments among the people with average and more than average knowledge on investments are equity shares, mutual funds and gold.

Fig. 9 shows most of the investors invest for long term returns. It also shows that most of the investors consist of the employed class of people.

### C. Sentiment Analysis

For building word clouds and analyzing the answers given in the short/long answer questions, nltk and word cloud libraries present in python were used and text was converted to lowercase, punctuation marks, numbers and stop words were eliminated, the frequency of rest of the words was calculated, and word clouds given below were created as a result of this process.

Fig. 10 - Is for the people who regularly make investments. This word cloud was made to understand the main reasons



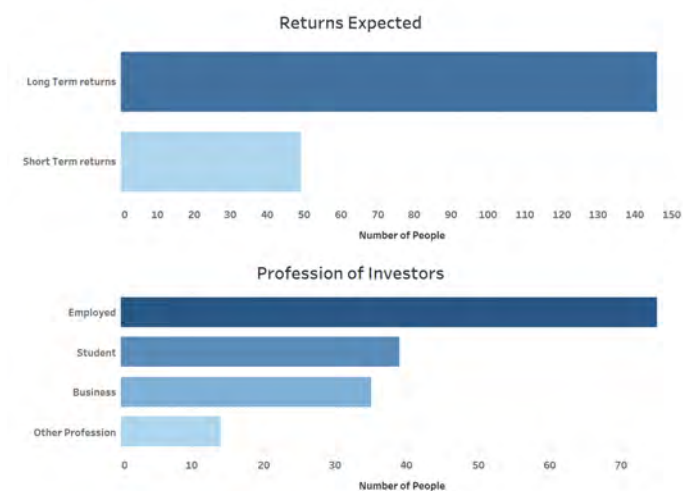


Fig. 9. Investor types and Intentions



Fig. 10. Word Cloud: Main Reasons to Invest

for investing and we found out that wealth creation, income, security, growth were the top reasons mentioned by investors.



Fig. 11. Word Cloud: Opinions on Cryptocurrency and NFTs

Fig 11 - This word cloud shows the opinions of people on the recent boom in Cryptocurrency and NFTs. Most of the people think it is volatile, don't have enough knowledge or feel it is risky to invest due to its volatile nature. Although they do see future scope in this area.



Fig. 12. Word Cloud: Main Reasons to not Invest

Fig. 12 - This word cloud contains the reasons of the people who aren't willing to invest. Main reasons are lack of knowledge and understanding, risk and less returns.

Sentiment analysis was performed on the following columns:

- 1) What is your main reason for investing ?. For this column we found out that there were 133 neutral , 25 positive and 10 negative sentiment responses
- 2) Why do you not invest?. For which we received 26 neutral, 9 positive and 3 negative sentiment responses.
- 3) What are your views on the recent boom in Crypto and NFTs?. For this question we received 78 neutral, 35 positive and 20 negative sentiment responses.

## V. CONCLUSION

The economy is booming, the job market has been doing so well and there has been an upsurge in the graph for salaries. The new generation of youth in India will have money in their pockets and ample opportunity to put it to good use, if they can shift from Bank Account Savings to the Capital Market. They need lucrative options to put their money in for long term returns, wealth creation but are scared or confused due to lack of understanding, knowledge and the risk involved. It has been observed that due to such a variety of options people want more diversification based on the risk reward ratio for a smoother PNL curve which is also one of the reasons people opt out from investing in Cryptocurrency, or have a less concentrated Crypto Portfolio. Thus people with half decent knowledge are regularly investing in Equities and Mutual Funds for a return that beats the FD. Anand Mahindra recently said in a speech “Stock market trading should be made compulsory in middle school education.” Based on our analysis and results, we strongly believe that this should be enacted.

## ACKNOWLEDGMENT

We would like to express our gratitude for all of the individuals who took out the time to fill the survey which was shared. We would also like to thank our professor, Dr. Ashwini Rao for her patient guidance, encouragement, and useful critiques of this research work. This research paper has

been prepared by Shruti Jain, Jash Jain, Garima Merani and Jash Patel, all in Fourth Year, BTech Information Technology, MPSTME, NMIMS University.

#### REFERENCES

- [1] A. Mehraj, "Impact of behavior on investment decisions: An empirical study of Jammu and Kashmir," International journal of multidisciplinary educational research, Srinagar, 2021.
- [2] J. Aspara, "Individuals' Affect-Based Motivations to Invest in Stocks: Beyond Expected Financial Returns and Risks," Journal of Behavioral Finance, Helsinki, 2009.
- [3] IA. Chatterjee, "Investor classification and sentiment analysis," IEEE International Conference, San Francisco, 2016.
- [4] C. Y. PATEL, "A study of investment perspective of salaried people," Asia Pacific Journal of Marketing & Management Review, Mumbai, 2012.
- [5] R. Srivastava, "Capital market forecasting by using sentimental analysis," 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, 2016.
- [6] N. Bhaskaran, "Investors behaviour in various investment avenues a study," IJMT, Tiruchirappalli, 2012.
- [7] D. Harini, "A study on the investment behavior of working women with reference to Chennai city," IJMRASC, Chennai, 2021.
- [8] D. Anilkumar, "A study on the behaviour and pattern of women investment," Calicut, 2021.
- [9] V. Jisha, "A Study on the perception of investment pattern among urban working women with reference to Coimbatore City," IJESC, Coimbatore, 2017.
- [10] D. Y. Kumar, "Analysis of investment pattern of different class of people- A review," IJRAR, 2019.
- [11] S. Sabarinathan, "A review research on investment behaviour of different class people in India," Dogo Rangsang Research Journal, Erode, 2020.
- [12] D. R.K., "Application of Artificial Neural Network for stock market predictions: A review of literature," International Journal of Machine Intelligence, Nanded, 2010.