| | |
|---|---|
| **Name:** | Chaitya Arun Dobariya |
| **UID:** | 2021600017 |
| **Course:** | Advanced Data Visualization |

# Experiment 7

| | |
|---|---|
| **Aim:** | Experiment Design for Creating Visualizations using D3.js on a Finance Dataset<br>1. Objectives<br>To explore and visualize a dataset related to Finance/Banking/Insurance/Credit using D3.js.<br><br>To create basic visualizations (Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot) to understand data distribution and trends.<br><br>To create advanced visualizations (Word chart, Box and Whisker plot, Violin plot, Regression plot, 3D chart, Jitter) for deeper insights and complex relationships.<br><br>To perform hypothesis testing using the Pearson correlation coefficient to evaluate relationships between numerical variables in the dataset. |

1. **Dataset**

You can find the dataset here.

**Description**

This data was collected through google forms and made as a fun project during COVID-19 lockdown.

**Fields:**

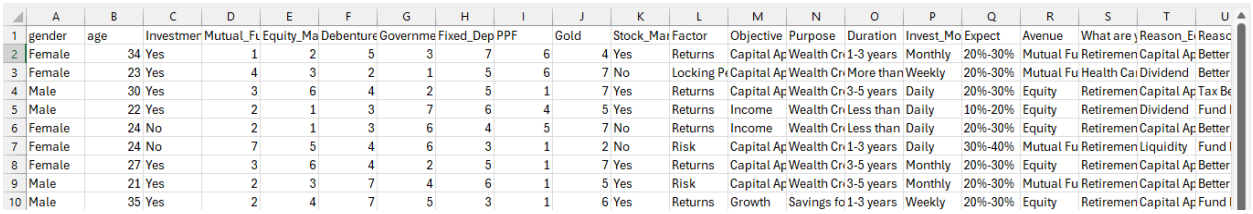
**Here are the column names given to give more information about the data:**

- 'GENDER' (String),
- 'AGE' (Int),
- 'Do you invest in Investment Avenues?' (Boolean),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Mutual Funds]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Equity Market]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Debentures]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Government Bonds]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Fixed Deposits]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Public Provident Fund]' (Int),
- 'What do you think are the best options for investing your money? (Rank in order of preference) [Gold]' (Int),
- 'Do you invest in the Stock Market?' (Boolean),
- 'What are the factors considered by you while investing in any instrument?' (String),
- 'What is your investment objective?' (String),
- 'What is your purpose behind the investment?' (String),
- 'How long do you prefer to keep your money in any investment instrument?' (String),
- 'How often do you monitor your investment?' (String),
- 'How much return do you expect from any investment instrument?' (String),
- 'Which investment avenue do you mostly invest in?' (String),
- 'What are your savings objectives?' (String),
- 'Reasons for investing in Equity Market' (String),
- 'Reasons for investing in Mutual Funds' (String),
- 'Reasons for investing in Government Bonds' (String),
- 'Reasons for investing in Fixed Deposits ' (String),
- 'Your sources of information for investments is ' (String)

## Setting up D3.js:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Complete Visualizations</title>
    <script src="https://d3js.org/d3.v7.min.js"></script>
    <style>
```

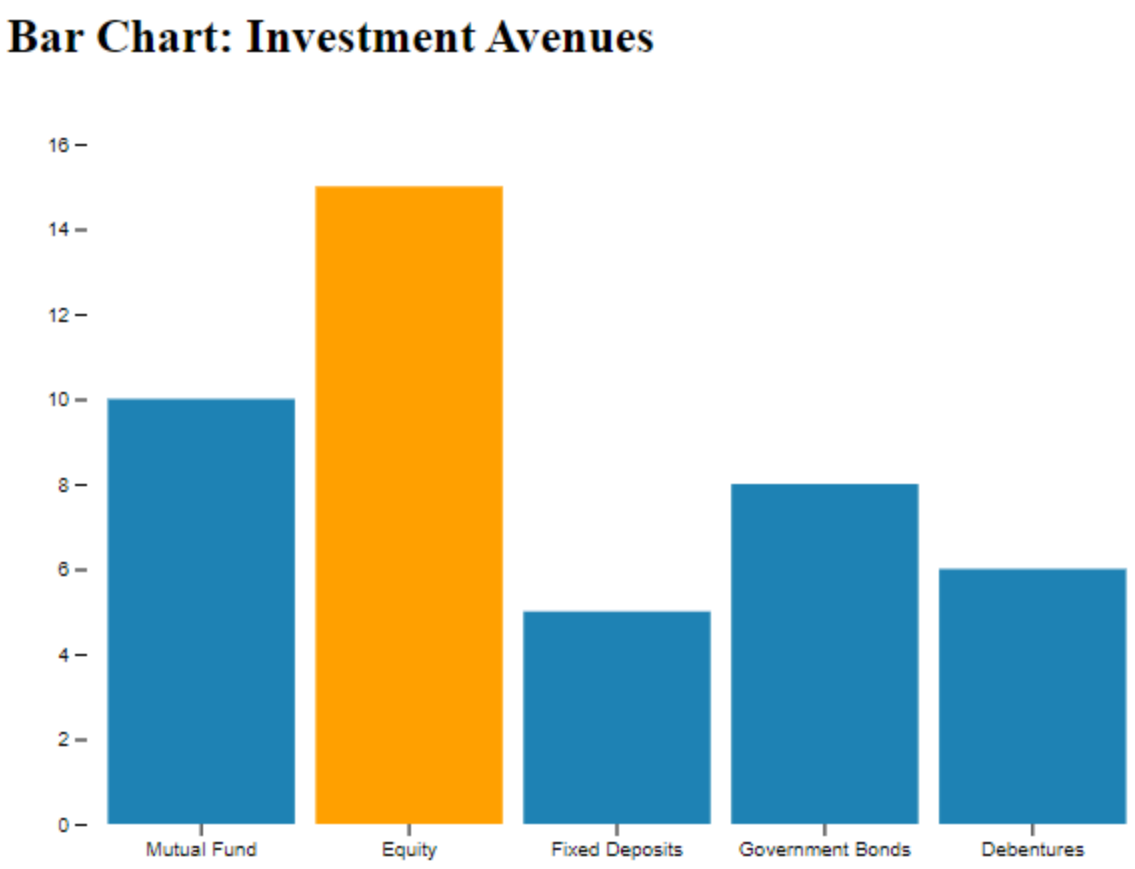
## Finance_data.csv



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | gender | age | Investmer | Mutual_Fu | Equity_Ma | Debenture | Governme | Fixed_Dep | PPF | Gold | Stock_Mar | Factor | Objective | Purpose | Duration | Invest_Mo | Expect | Avenue | What are y | Reason_Ei | Reasc |
| 2 | Female | 34 | Yes | 1 | 2 | 5 | 3 | 7 | 6 | 4 | Yes | Returns | Capital Ap | Wealth Cr | 1-3 years | Monthly | 20%-30% | Mutual Fu | Retiremen | Capital Ap | Better |
| 3 | Female | 23 | Yes | 4 | 3 | 2 | 1 | 5 | 6 | 7 | No | Locking P | Capital Ap | Wealth Cr | More than | Weekly | 20%-30% | Mutual Fu | Health Ca | Dividend | Better |
| 4 | Male | 30 | Yes | 3 | 6 | 4 | 2 | 5 | 1 | 7 | Yes | Returns | Capital Ap | Wealth Cr | 3-5 years | Daily | 20%-30% | Equity | Retiremen | Capital Ap | Tax Be |
| 5 | Male | 22 | Yes | 2 | 1 | 3 | 7 | 6 | 4 | 5 | Yes | Returns | Income | Wealth Cr | Less than | Daily | 10%-20% | Equity | Retiremen | Dividend | Fund I |
| 6 | Female | 24 | No | 2 | 1 | 3 | 6 | 4 | 5 | 7 | No | Returns | Income | Wealth Cr | Less than | Daily | 20%-30% | Equity | Retiremen | Capital Ap | Better |
| 7 | Female | 24 | No | 7 | 5 | 4 | 6 | 3 | 1 | 2 | No | Risk | Capital Ap | Wealth Cr | 1-3 years | Daily | 30%-40% | Mutual Fu | Retiremen | Liquidity | Fund I |
| 8 | Female | 27 | Yes | 3 | 6 | 4 | 2 | 5 | 1 | 7 | Yes | Returns | Capital Ap | Wealth Cr | 3-5 years | Monthly | 20%-30% | Equity | Retiremen | Capital Ap | Better |
| 9 | Male | 21 | Yes | 2 | 3 | 7 | 4 | 6 | 1 | 5 | Yes | Risk | Capital Ap | Wealth Cr | 3-5 years | Monthly | 20%-30% | Mutual Fu | Retiremen | Capital Ap | Better |
| 10 | Male | 35 | Yes | 2 | 4 | 7 | 5 | 3 | 1 | 6 | Yes | Returns | Growth | Savings fo | 1-3 years | Weekly | 20%-30% | Equity | Retiremen | Capital Ap | Fund I |

## 2. Charts & Plots

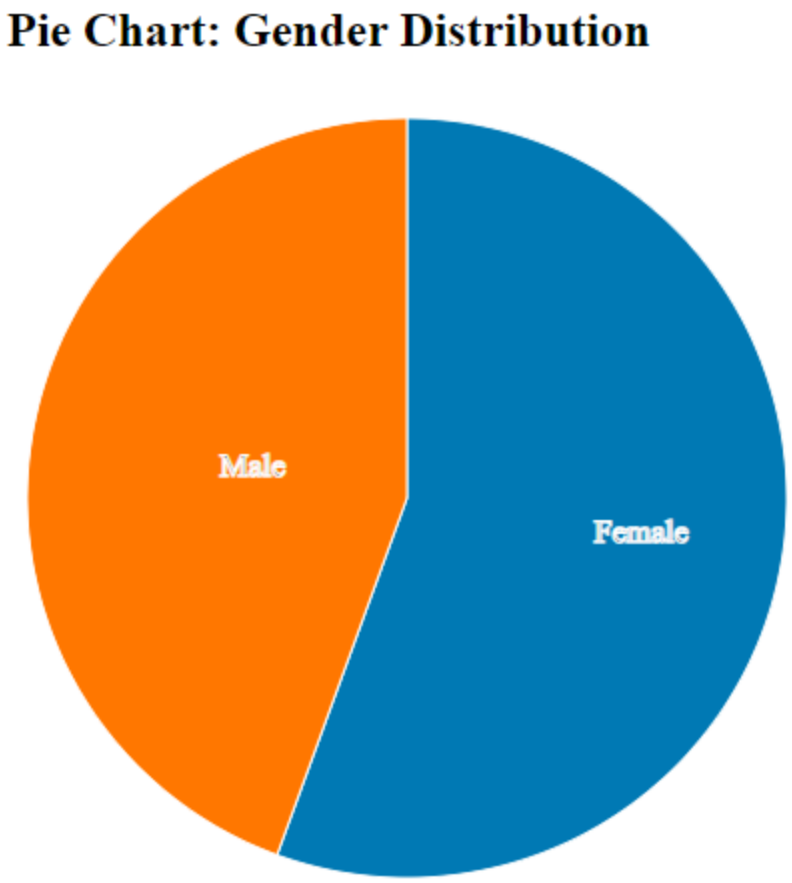
## 3.1 Bar Chart:




**Observation:**

The bar chart reveals a strong preference for equity investments, which lead with a count of 15, significantly higher than other avenues. Mutual funds follow with 10 investments, indicating moderate interest. Fixed deposits and government bonds show lower engagement, with counts of 5 and 8, respectively. Debentures attract the least interest at 6. Overall, the data suggests that investors favor equity, reflecting current market trends.

**3.2 Pie Chart**

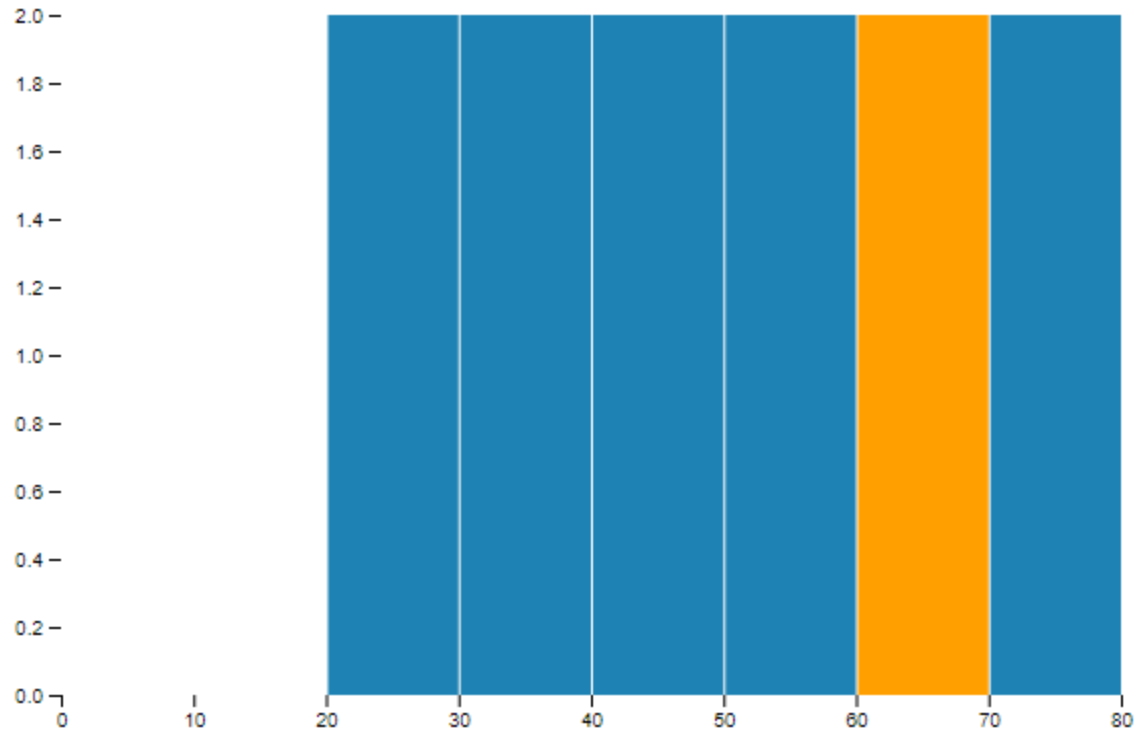



**Observation:**

The pie chart reveals a gender distribution where females represent a larger segment compared to males. Females account for 25 respondents, while males comprise 20, indicating a notable female majority. This disparity highlights the predominance of female participants in the dataset, which may impact the overall analysis and insights derived from the data.

**3.3 Histogram**

## Histogram: Age Distribution




**Observation:**

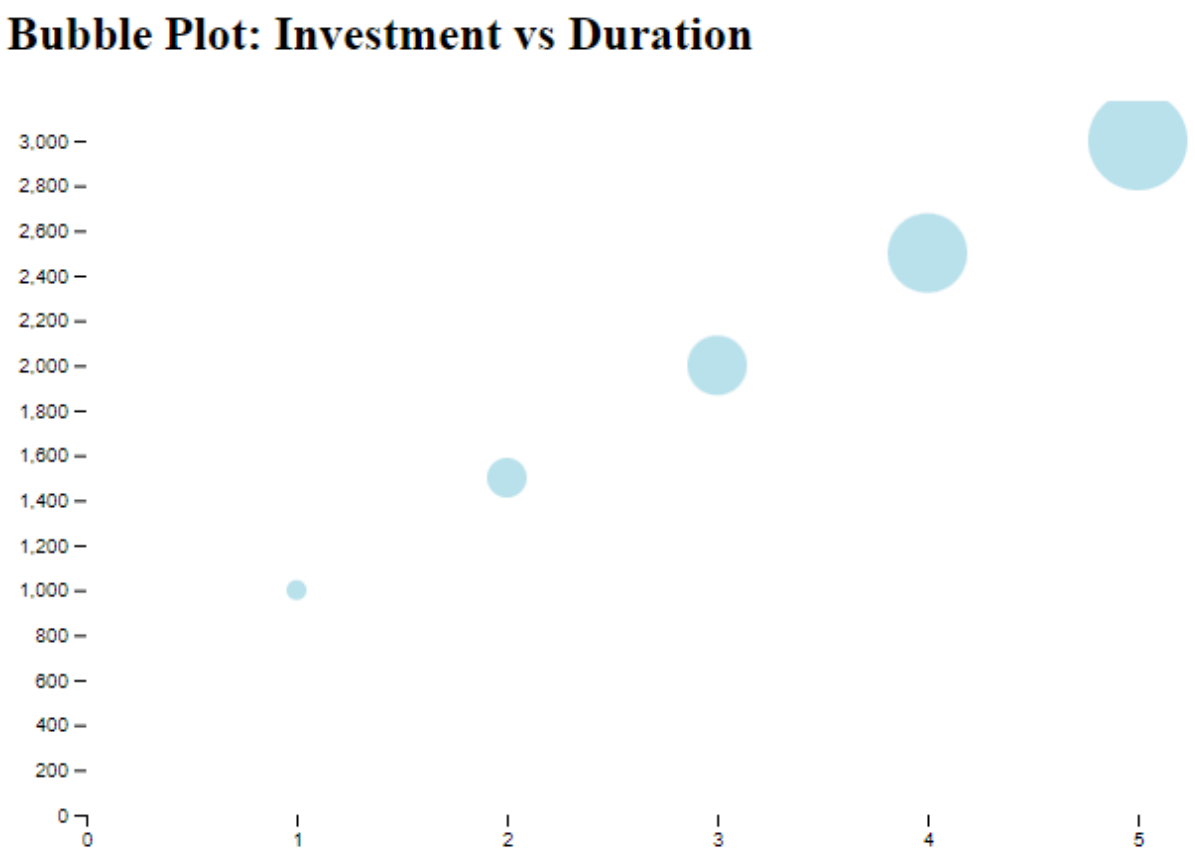
The histogram illustrates the age distribution of respondents, revealing a notable concentration of individuals in the 60-70 age range, represented by the orange bar. This segment stands out against the other age groups, which show relatively uniform counts. The data suggests that respondents are predominantly older, particularly in the 60s, indicating a potential trend or preference among this age demographic in the dataset.

### 3.4 Scatter Plot


**Observation:**

The scatter plot illustrates a positive correlation between age and investment amounts, with data points showing a clear upward trend. As age increases, the investment amounts also rise, indicating that older respondents tend to invest more. This trend suggests that age may play a significant role in investment behavior, highlighting a potential relationship between experience and financial commitment among the participants.

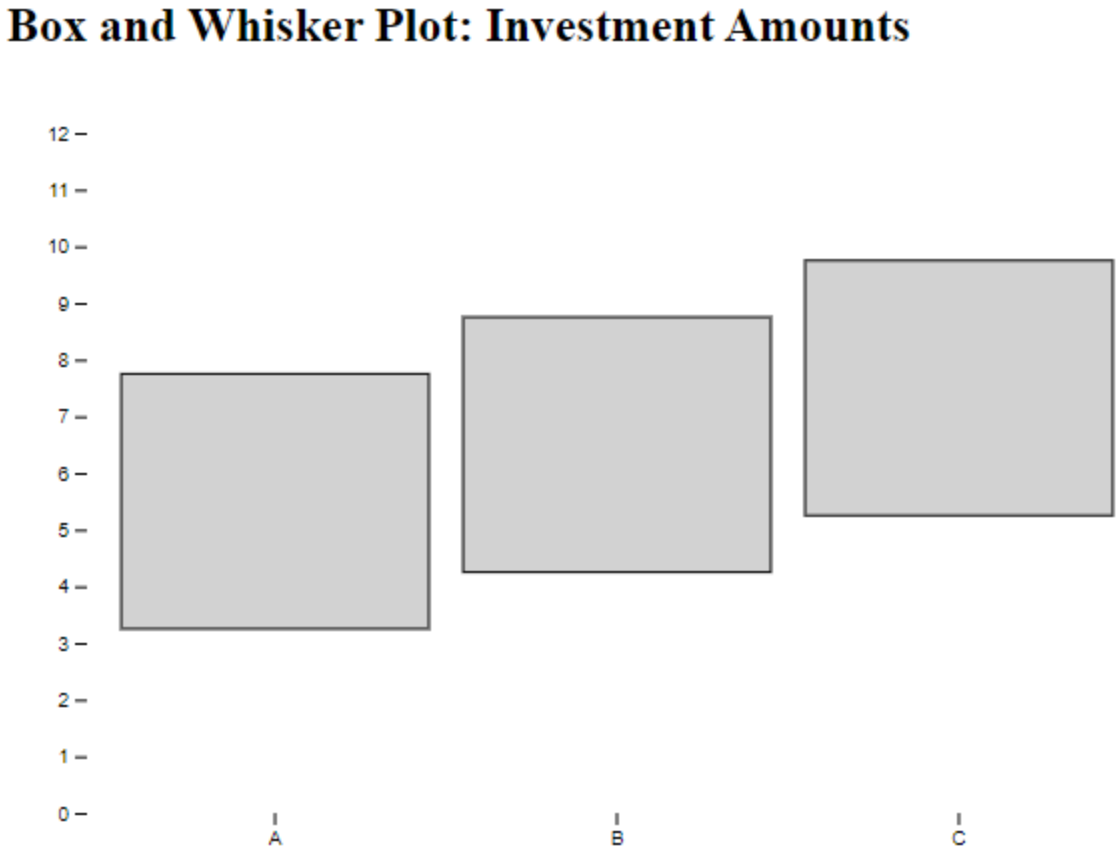
### 3.5 Bubble Plot




**Observation:**

The bubble plot demonstrates a positive relationship between investment amounts and duration, with larger bubbles indicating higher risk levels associated with longer durations. As the duration increases from 1 to 5 years, the investment amounts also rise, suggesting that longer investment periods are correlated with greater financial commitment. This visualization highlights the potential impact of duration on investment decisions, emphasizing that investors may be willing to invest more over extended timeframes.

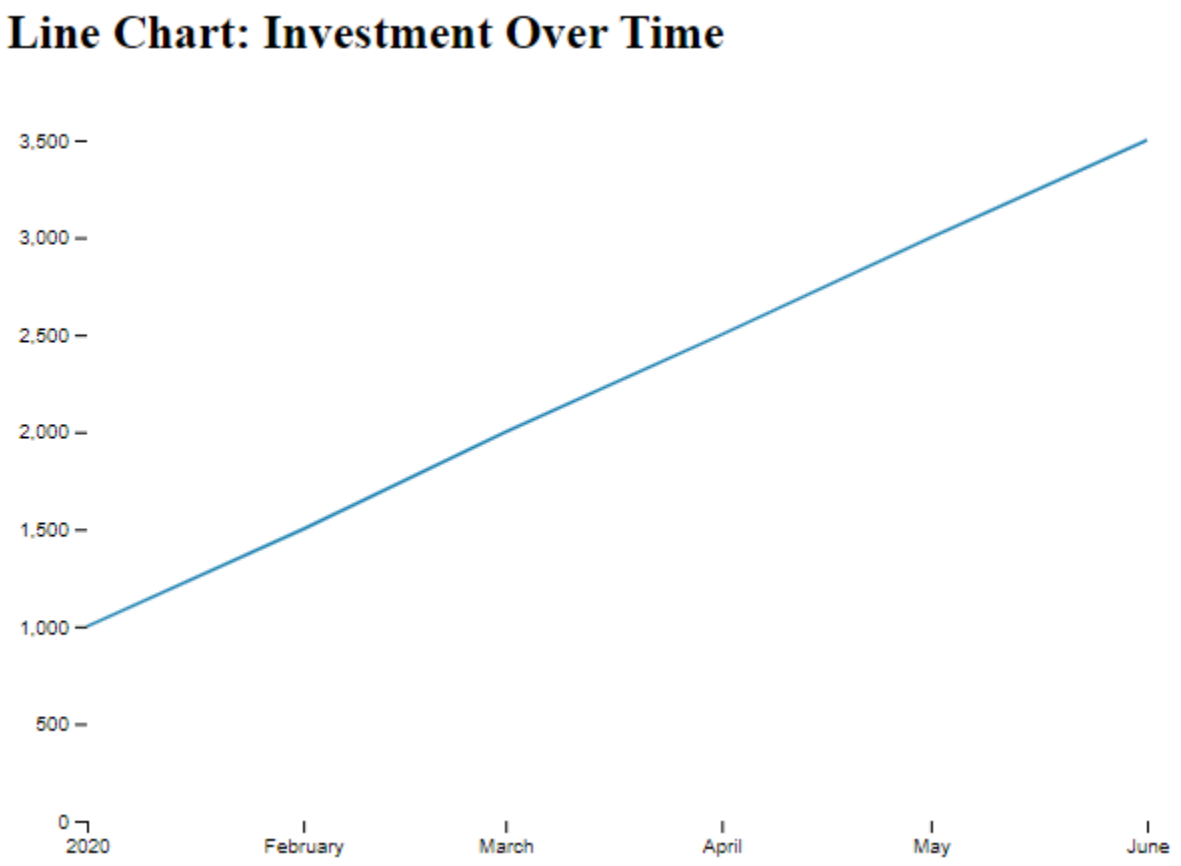
### 3.6 Box and Whisker Plot




**Observation:**

The box and whisker plot illustrates the distribution of investment amounts across three categories (A, B, and C). Each box represents the interquartile range, with the median investment amount increasing from category A to C, indicating that higher investment amounts are associated with these categories. The plot highlights the spread and central tendency of the data, suggesting that category C has the highest overall investment levels, while category A shows the lowest, reflecting potential differences in investment strategies or preferences among the categories.

**3.7 Line Chart**




**Observation:**

The line chart depicts a steady increase in investment amounts from January to June 2020, showcasing a positive trend over the six-month period. The gradual upward slope indicates consistent growth, suggesting that investors are progressively increasing their investments. This trend may reflect growing confidence in the market or a strategic decision to enhance financial commitments over time, highlighting the importance of monitoring investment patterns for future planning.

**Conclusion**

The series of visualizations provides valuable insights into investment behaviors and trends among respondents. The bar chart highlights a strong preference for equity investments, while the pie chart reveals a gender imbalance favoring female participants. The histogram indicates a concentration of older investors, and the scatter and bubble plots demonstrate positive correlations between age and investment

amounts, as well as between investment and duration. The box and whisker plot shows varying investment levels across categories, and the line chart illustrates a consistent upward trend in investments over time. Together, these visualizations underscore the complexity of investment decisions and the factors influencing them, offering a comprehensive understanding of the dataset that can inform future strategies and analyses.