

Name:Aditya khuman

UID: 2021300061

Experiment no.: 06

Aim: Experiment Design for Creating Visualizations using D3.js on Finance Dataset

Dataset 1:

Netflix stock data from 2014 to 2023:

date: The date on which the stock data was recorded.

open: The stock's opening price on the given day.

high: The highest price the stock reached during the trading day.

low: The lowest price the stock reached during the trading day.

close: The stock's closing price on the given day.

volume: The number of shares traded on that day.

rsi_7: The 7-day Relative Strength Index (RSI), which indicates the momentum of the stock. Values range from 0 to 100, with higher values indicating an overbought condition and lower values indicating oversold conditions.

rsi_14: The 14-day Relative Strength Index, similar to the 7-day RSI but over a longer period.

cci_7: The 7-day Commodity Channel Index (CCI), used to identify cyclical trends in the stock. High values suggest overbought conditions, and low values suggest oversold conditions.

cci_14: The 14-day Commodity Channel Index, similar to the 7-day CCI but over a longer period.

sma_50: The 50-day Simple Moving Average, an average of the stock's closing prices over the past 50 days.

ema_50: The 50-day Exponential Moving Average, a weighted moving average that gives more importance to recent data points.

sma_100: The 100-day Simple Moving Average.

ema_100: The 100-day Exponential Moving Average.

macd: The Moving Average Convergence Divergence, a momentum indicator that shows the relationship between two moving averages (usually the 26-day and 12-day EMAs). Positive values suggest bullish momentum, while negative values suggest bearish momentum.

bollinger: The value of the upper Bollinger Band, a volatility indicator that uses standard deviation to create an envelope around the stock's moving average.

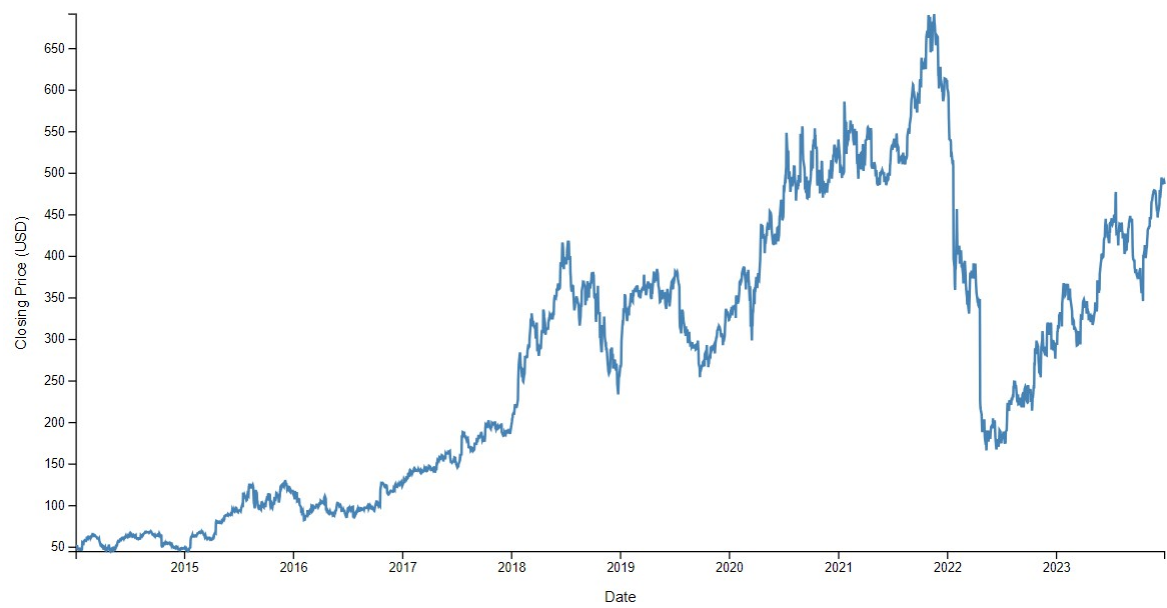
TrueRange: The true range, a measure of volatility that takes into account the most significant price changes within a day (including gaps between the previous day's close and the current day's open).

atr_7: The 7-day Average True Range (ATR), an indicator of the stock's volatility over the past 7 days.

atr_14: The 14-day Average True Range, similar to atr_7 but over a longer period.

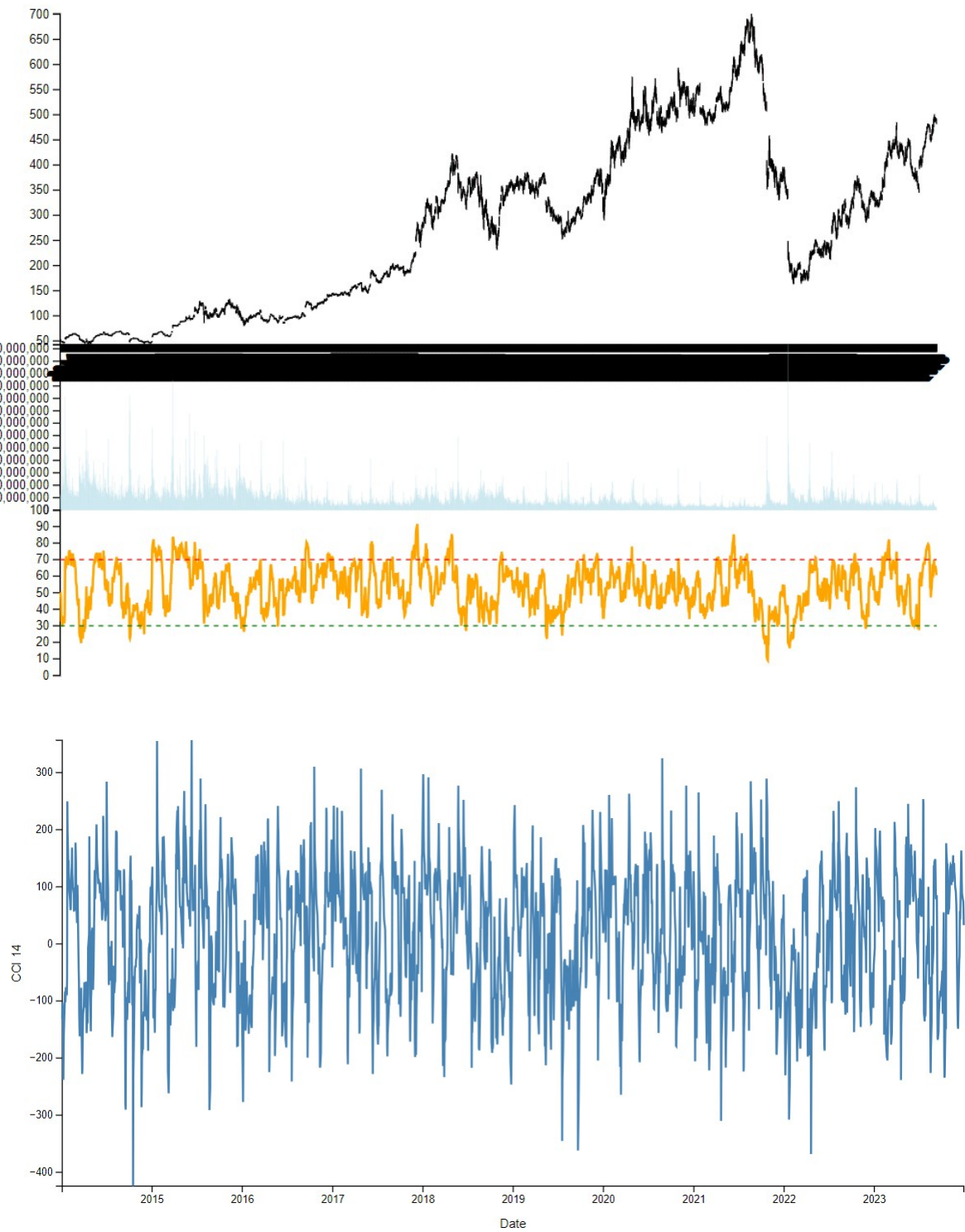
Visualizations:

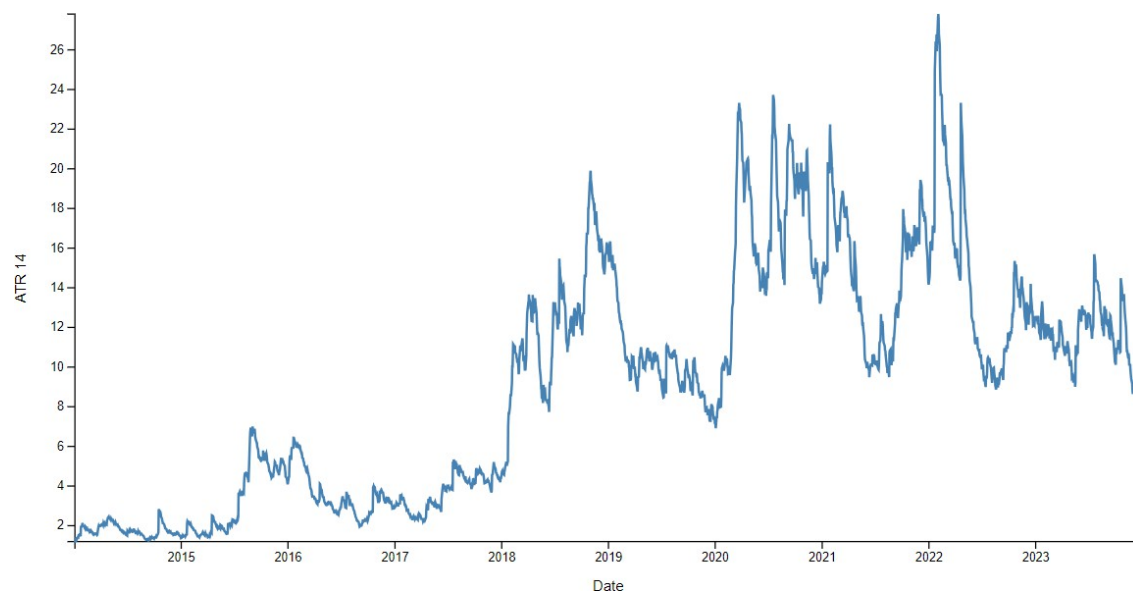
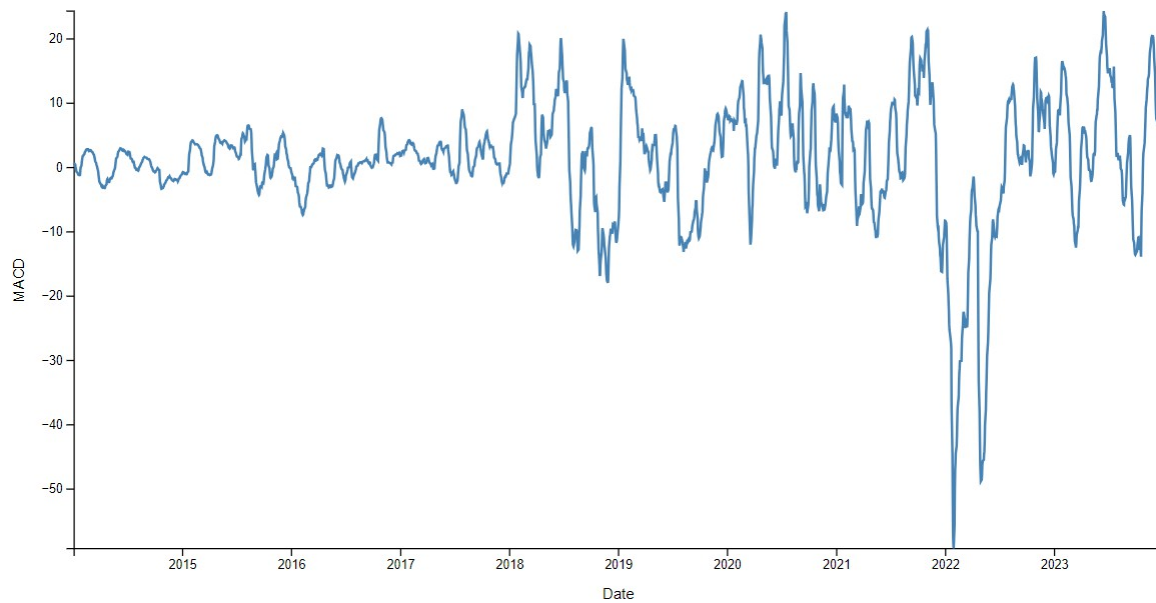
Netflix Stock Data Visualizations (2014-2023)



Stock price - time line chart

Candlestick Chart with Volume and RSI





Candel stick chart with volume and RSI.

Timeline Chart (Closing Price over Time):

- **Represents:** This chart shows the stock's closing price over time.
- **X-axis:** Date (from 2014 to 2023).
- **Y-axis:** Closing price in USD.
- **Purpose:** To track the trend of Netflix's stock price over the years.

Volume Bar Chart (Trading Volume over Time):

- **Represents:** This chart displays the volume of Netflix stocks traded over time.
- **X-axis:** Date (showing years to reduce clutter).
- **Y-axis:** Trading volume (number of shares).
- **Purpose:** To show how much Netflix stock was traded daily, providing insight into market activity.

RSI (14-day Relative Strength Index over Time):

- **Represents:** This chart tracks the 14-day RSI (Relative Strength Index), which measures momentum by comparing gains and losses over a specific time period.
- **X-axis:** Date.
- **Y-axis:** RSI value (typically ranging between 0 and 100).
- **Purpose:** To help identify whether Netflix stock is overbought or oversold at various times.
- **CCI 7 Chart (7-day Commodity Channel Index over Time):**
 - **Represents:** The 7-day Commodity Channel Index (CCI) measures price deviation from the statistical mean. It helps in identifying overbought and oversold levels.
 - **X-axis:** Date.
 - **Y-axis:** CCI value (typically ranging from -200 to +200).
 - **Purpose:** CCI helps detect cyclical trends in the stock price.
- **CCI 14 Chart (14-day Commodity Channel Index over Time):**
 - **Represents:** Similar to CCI 7, but over a longer period (14 days).
 - **X-axis:** Date.
 - **Y-axis:** CCI value (-200 to +200).
 - **Purpose:** Provides a longer-term view of price deviation and cycles.
- **MACD Chart (Moving Average Convergence Divergence):**
 - **Represents:** MACD is a trend-following indicator that shows the relationship between two moving averages of a stock's price (e.g., 12-day and 26-day EMA).
 - **X-axis:** Date.
 - **Y-axis:** MACD value.
 - **Purpose:** Helps identify bullish or bearish momentum in the stock's price.
- **SMA 50 Chart (50-day Simple Moving Average over Time):**
 - **Represents:** The 50-day simple moving average, which smooths out price fluctuations over the last 50 days.
 - **X-axis:** Date.
 - **Y-axis:** SMA value.
 - **Purpose:** SMA is a basic trend indicator, showing the average closing price over a given period.
- **EMA 50 Chart (50-day Exponential Moving Average over Time):**
 - **Represents:** The 50-day exponential moving average, similar to the SMA but giving more weight to recent prices.
 - **X-axis:** Date.
 - **Y-axis:** EMA value.
 - **Purpose:** EMA reacts more quickly to recent price changes compared to SMA.

- **SMA 100 Chart (100-day Simple Moving Average over Time):**
 - **Represents:** The 100-day simple moving average.
 - **X-axis:** Date.
 - **Y-axis:** SMA value.
 - **Purpose:** A longer-term trend indicator.
- **EMA 100 Chart (100-day Exponential Moving Average over Time):**
 - **Represents:** The 100-day exponential moving average.
 - **X-axis:** Date.
 - **Y-axis:** EMA value.
 - **Purpose:** A longer-term EMA, giving insight into long-term trends.
- **Bollinger Bands Chart:**
 - **Represents:** Bollinger Bands consist of a moving average and two standard deviations plotted above and below the moving average. This shows price volatility.
 - **X-axis:** Date.
 - **Y-axis:** Stock price with upper and lower bands.
 - **Purpose:** To show potential high and low points relative to past price movements.
- **ATR 7 (Average True Range over 7 days):**
 - **Represents:** The 7-day ATR, which measures market volatility.
 - **X-axis:** Date.
 - **Y-axis:** ATR value.
 - **Purpose:** ATR helps gauge how much stock price moves on average during a day.
- **ATR 14 (Average True Range over 14 days):**
 - **Represents:** The 14-day ATR.
 - **X-axis:** Date.
 - **Y-axis:** ATR value.
 - **Purpose:** Provides a longer-term view of price volatility.

Observations:

- **Volume chart:** Shows stock trading volume over time.
- **RSI (7 & 14):** Indicates momentum and overbought/oversold conditions.
- **CCI (7 & 14):** Measures price deviation from a mean and detects cycles.
- **MACD:** Tracks momentum and trend direction.
- **SMA/EMA (50 & 100):** Identifies trends and smooths price data over time.
- **Bollinger Bands:** Measures volatility with potential buy/sell signals.
- **ATR (7 & 14):** Gauges market volatility.

Dataset 2:

```
finance.csv > data
1 CustomerID, Age, Income, LoanAmount, CreditScore, LoanTerm, InterestRate, AccountType, LoanStatus, Region, YearsWithBank, ProductCategory
2 1, 25, 55000, 20000, 700, 60, 5.5, Savings, Approved, North, 3, Mortgage
3 2, 42, 120000, 50000, 750, 36, 3.2, Checking, Approved, West, 8, Car Loan
4 3, 36, 90000, 30000, 620, 48, 6.5, Savings, Denied, South, 5, Personal Loan
5 4, 29, 70000, 15000, 680, 72, 4.1, Savings, Approved, East, 2, Car Loan
6 5, 53, 140000, 75000, 800, 120, 2.8, Checking, Approved, North, 10, Mortgage
7 6, 41, 100000, 40000, 720, 84, 3.9, Savings, Approved, West, 7, Personal Loan
8 7, 34, 85000, 22000, 690, 36, 5.7, Checking, Denied, South, 6, Car Loan
9 8, 46, 130000, 60000, 780, 60, 3.1, Savings, Approved, East, 9, Mortgage
10 9, 31, 62000, 18000, 650, 24, 6.8, Checking, Approved, North, 3, Personal Loan
11 10, 58, 150000, 85000, 820, 180, 2.5, Savings, Approved, West, 15, Mortgage
12 11, 26, 57000, 21000, 710, 60, 5.4, Savings, Approved, South, 3, Car Loan
13 12, 39, 95000, 45000, 730, 48, 3.7, Checking, Approved, East, 7, Personal Loan
14 13, 45, 125000, 55000, 760, 72, 3.3, Savings, Approved, North, 8, Mortgage
15 14, 37, 87000, 32000, 640, 36, 6.3, Savings, Denied, West, 6, Personal Loan
16 15, 28, 68000, 17000, 670, 60, 5.9, Checking, Approved, South, 4, Car Loan
17 16, 52, 135000, 70000, 790, 96, 2.9, Savings, Approved, East, 12, Mortgage
18 17, 43, 110000, 48000, 740, 84, 3.5, Checking, Approved, North, 9, Personal Loan
19 18, 33, 82000, 24000, 660, 48, 6.0, Savings, Denied, West, 5, Car Loan
20 19, 48, 128000, 62000, 770, 60, 3.0, Savings, Approved, South, 11, Mortgage
21 20, 30, 60000, 16000, 690, 24, 5.6, Checking, Approved, East, 3, Personal Loan
22 21, 55, 145000, 78000, 810, 144, 2.6, Savings, Approved, North, 13, Mortgage
23 22, 40, 98000, 41000, 720, 60, 4.0, Checking, Approved, West, 7, Personal Loan
24 23, 29, 72000, 19000, 680, 36, 5.8, Savings, Denied, South, 4, Car Loan
25 24, 47, 132000, 63000, 780, 96, 3.1, Savings, Approved, East, 10, Mortgage
26 25, 35, 90000, 35000, 700, 60, 5.5, Checking, Approved, North, 6, Personal Loan
```

Dataset Description

- CustomerID:**
 - Type:** Integer
 - Description:** A unique identifier for each customer. This helps in tracking and referencing individual records in the dataset.
- Age:**
 - Type:** Integer
 - Description:** The age of the customer in years. This can be useful for analyzing loan eligibility or trends based on age demographics.
- Income:**
 - Type:** Integer
 - Description:** The annual income of the customer in monetary units (e.g., dollars). Higher income may correlate with better loan approval rates and larger loan amounts.
- LoanAmount:**
 - Type:** Integer
 - Description:** The total amount of the loan requested by the customer. This is important for assessing loan risk and repayment capacity.
- CreditScore:**
 - Type:** Integer
 - Description:** A numerical representation of the customer's creditworthiness, typically ranging from 300 to 850. Higher scores indicate lower risk for lenders.
- LoanTerm:**
 - Type:** Integer

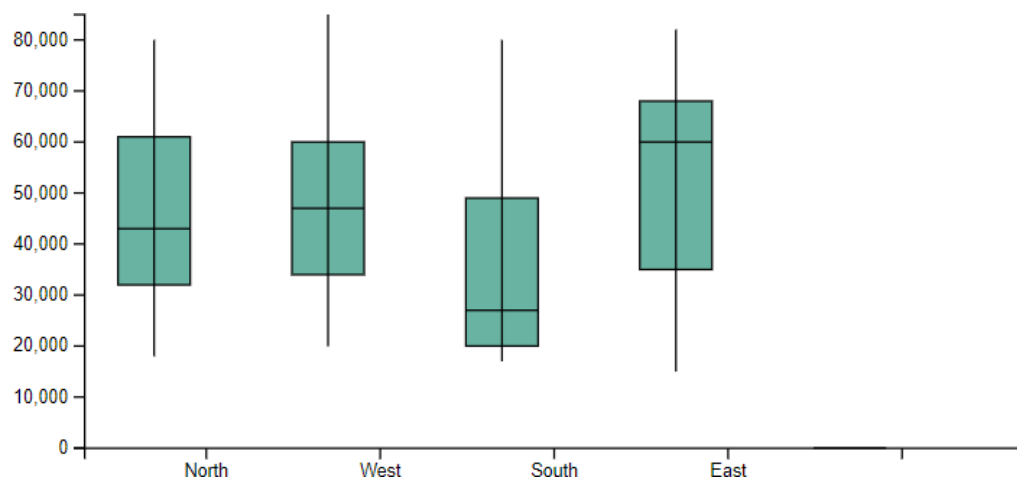
- **Description:** The duration of the loan in months. This can affect the monthly payment amount and overall interest paid over the life of the loan.
- 7. **InterestRate:**
 - **Type:** Float
 - **Description:** The annual interest rate charged on the loan, expressed as a percentage. It is a crucial factor affecting the total cost of the loan.
- 8. **AccountType:**
 - **Type:** Categorical (String)
 - **Description:** The type of bank account the customer holds (e.g., Savings, Checking). This may relate to customer loyalty and eligibility for loans.
- 9. **LoanStatus:**
 - **Type:** Categorical (String)
 - **Description:** The outcome of the loan application (e.g., Approved, Denied). This is critical for evaluating the success rate of loan applications.
- 10. **Region:**
 - **Type:** Categorical (String)
 - **Description:** The geographic region where the customer resides (e.g., North, South, East, West). This can provide insights into regional lending trends.
- 11. **YearsWithBank:**
 - **Type:** Integer
 - **Description:** The number of years the customer has been with the bank. Longer relationships may lead to better loan terms or approval chances.
- 12. **ProductCategory:**
 - **Type:** Categorical (String)
 - **Description:** The type of loan product being applied for (e.g., Mortgage, Car Loan, Personal Loan). This is useful for analyzing the distribution of loan types among customers.

Visualization:

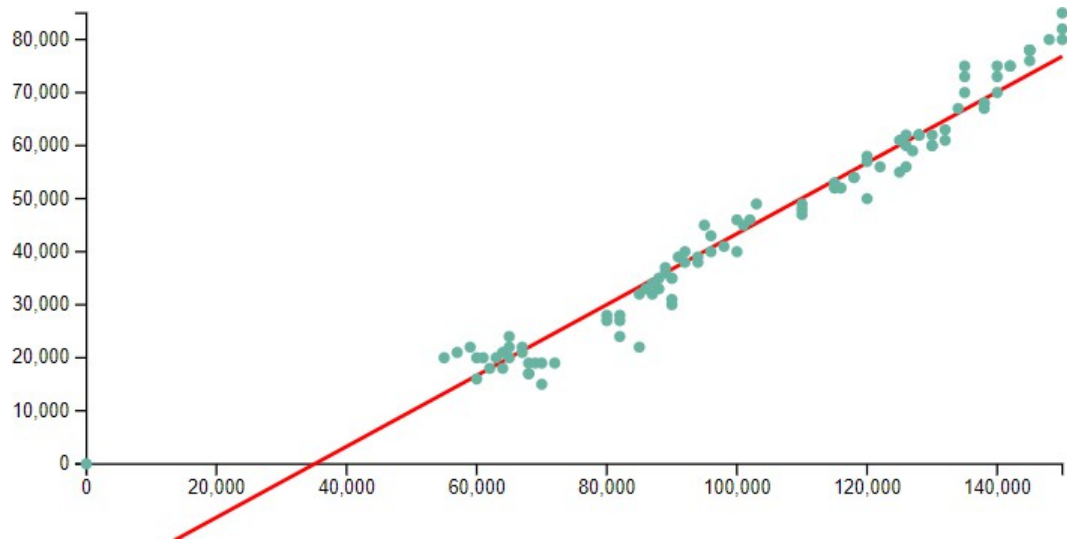
Word plot:

Car Loan Mortgage
Mortgage Car Loan
Car Loan Mortgage
Personal Loan Car Loan

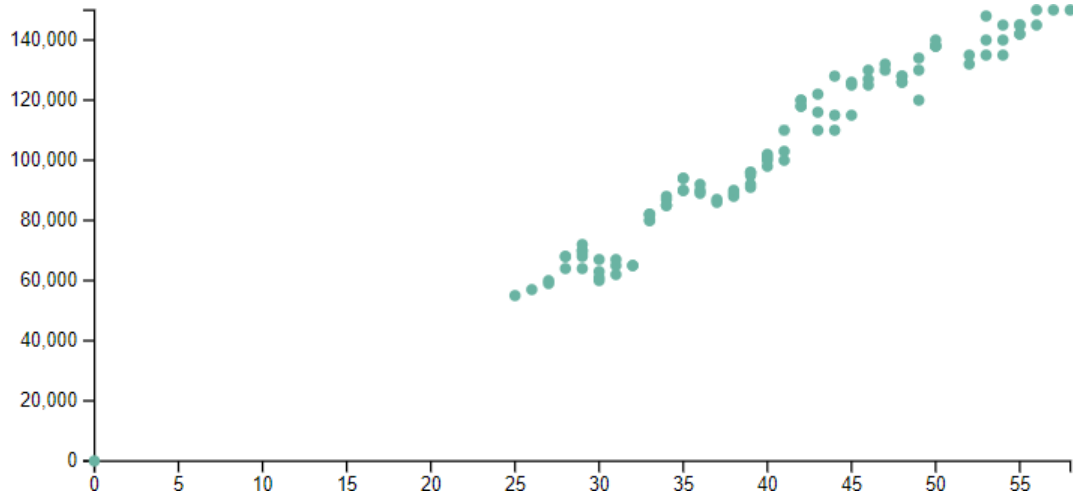
Box plot:



Regression plot:



Jitter plot:



Word Cloud (Product Categories by Income)

- **Purpose:** Visualize different `ProductCategory` items (like Loan, Credit Card, Insurance, etc.) and their relative importance by scaling the size of each word based on `Income`.
- **X-Axis / Y-Axis:** No axes here as it's a word cloud. The size of each word is determined by the `Income` value for each product category.
- **What it Shows:** The bigger the word, the higher the income associated with that product category. It helps identify which financial products are tied to higher incomes.

2. Box Plot (Loan Amount by Region)

- **Purpose:** Display the distribution of `LoanAmount` across different `Region` values using a box plot.
- **X-Axis:** Represents different `Region` categories (e.g., North, South, East, West).
- **Y-Axis:** Represents the `LoanAmount`.
- **What it Shows:** For each region, it displays the quartiles (Q1, median, Q3) of loan amounts, the interquartile range (IQR), and the minimum and maximum loan amounts. It helps compare loan amounts across regions and see the spread of the data.

3. Violin Plot (Loan Amount by Credit Score)

- **Purpose:** Show the distribution of `LoanAmount` across different `CreditScore` values using a violin plot.
- **X-Axis:** Represents the `CreditScore`.
- **Y-Axis:** Represents the `LoanAmount`.
- **What it Shows:** The plot shows the density of the data (how many loans are taken at different credit scores) and the variation in loan amounts. Thicker parts of the violin indicate more frequent occurrences.

4. Regression Plot (Income vs. Loan Amount)

- **Purpose:** Display the relationship between `Income` and `LoanAmount` and fit a regression line to show the trend.
- **X-Axis:** Represents the `Income`.
- **Y-Axis:** Represents the `LoanAmount`.
- **What it Shows:** Scatter plot points show the data distribution between income and loan amounts, and the red regression line shows the linear relationship. It helps to observe whether higher income is generally associated with larger loan amounts.

5. Jitter Plot (Income by Age)

- **Purpose:** Show the distribution of `Income` by `Age` using jittered points to reduce overlap and make the distribution clearer.

- **X-Axis:** Represents the *Age*.
- **Y-Axis:** Represents the *Income*.
- **What it Shows:** Each point represents an individual's income for a given age, spread slightly to prevent points from overlapping. This plot helps visualize how income is distributed across different ages, showing any patterns or clusters.

Conclusion:

In this practical, i learnt about various visualizations using D3 js and used the finance data.