

# STOCK INVESTMENT RECOMMENDATION MODEL

## GROUP MEMBERS:

- 1. HANSEL OMONDI
- 2. SANDRA KIPTUM
- 3. MIRRIAM MUMBUA
- 4. JEFFREY ONGICHO
- 5. CAROLINE KAMUSI



# CONTENTS

What this project covers

---

**01** Project Overview

**02** Business Understanding

**03** Data Understanding & Analysis

**04** Model Development &  
Evaluation

**05** Conclusion

A black and white photograph showing a close-up of a person's hands typing on a laptop keyboard. A pen lies next to the keyboard. The background is dark.

# PROJECT OVERVIEW

This project aims to create a machine learning model that delivers personalized stock investment recommendations based on individual risk appetites.

# PROJECT OVERVIEW



By analyzing historical data from Yahoo Finance, the system will suggest suitable investments across various sectors, empowering investors with data-driven insights to enhance their decision-making and improve investment outcomes.

# BUSINESS UNDERSTANDING

---

## Primary Stakeholders

New Investors

|

## Key Problem Statement

Challenges Investors Face in Stock Selection

- Complexities of financial markets and stock behavior
- Varied risk preferences across investors
- Mismatch between investor risk tolerance and portfolio composition
- Can lead to unintentional financial losses or missed growth opportunities



# BUSINESS UNDERSTANDING

---

## Project Objective

- Bridge the gap between investor risk tolerance and portfolio selection
- Provide tailored stock recommendations based on individual risk preferences

## System Approach

- Automatically filter and categorize stocks by risk level
  - Low-risk, medium-risk, and high-risk portfolios
- Customize recommendations to align with investor risk profiles



# DATA UNDERSTANDING

## OVERVIEW

- Dataset from Yahoo Finance with historical stock price data
  - Daily trading information for analyzing stock performance, volatility, and trends
- 

## DATASET STRUCTURE

- Open: Opening price of the stock on a given day
- High: Highest price during the trading session
- Low: Lowest price during the trading session
- Close: Closing price at the end of the trading day
- Adj Close: Adjusted closing price accounting for corporate actions (e.g., stock splits, dividends)
- Volume: Number of shares traded, indicating market activity and liquidity
- Beta: Measure of stock's volatility relative to the market (e.g., S&P 500)

# DATA UNDERSTANDING

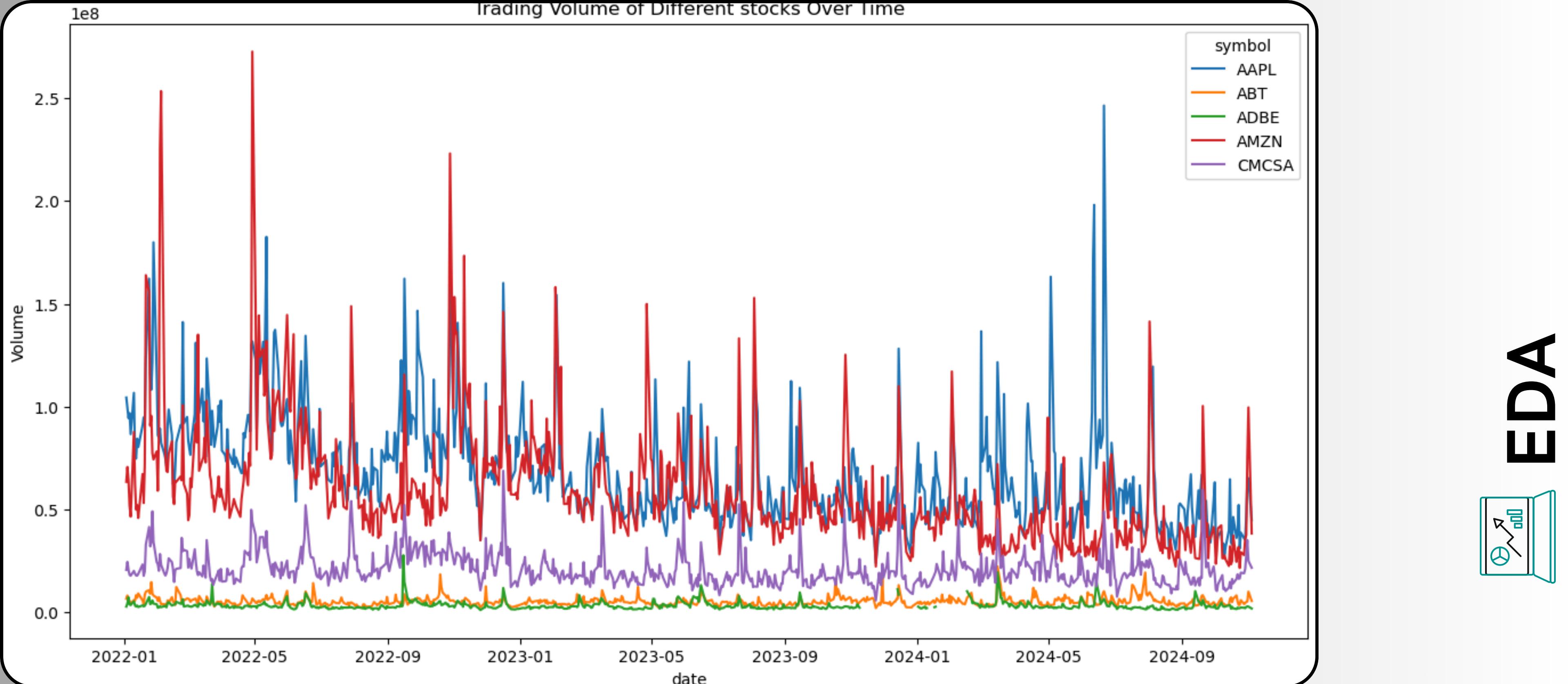
Exploratory Data Analysis



- We conducted a univariate, Bivariate and multivariate analysis performed with further analysis being done inTableau via the link below

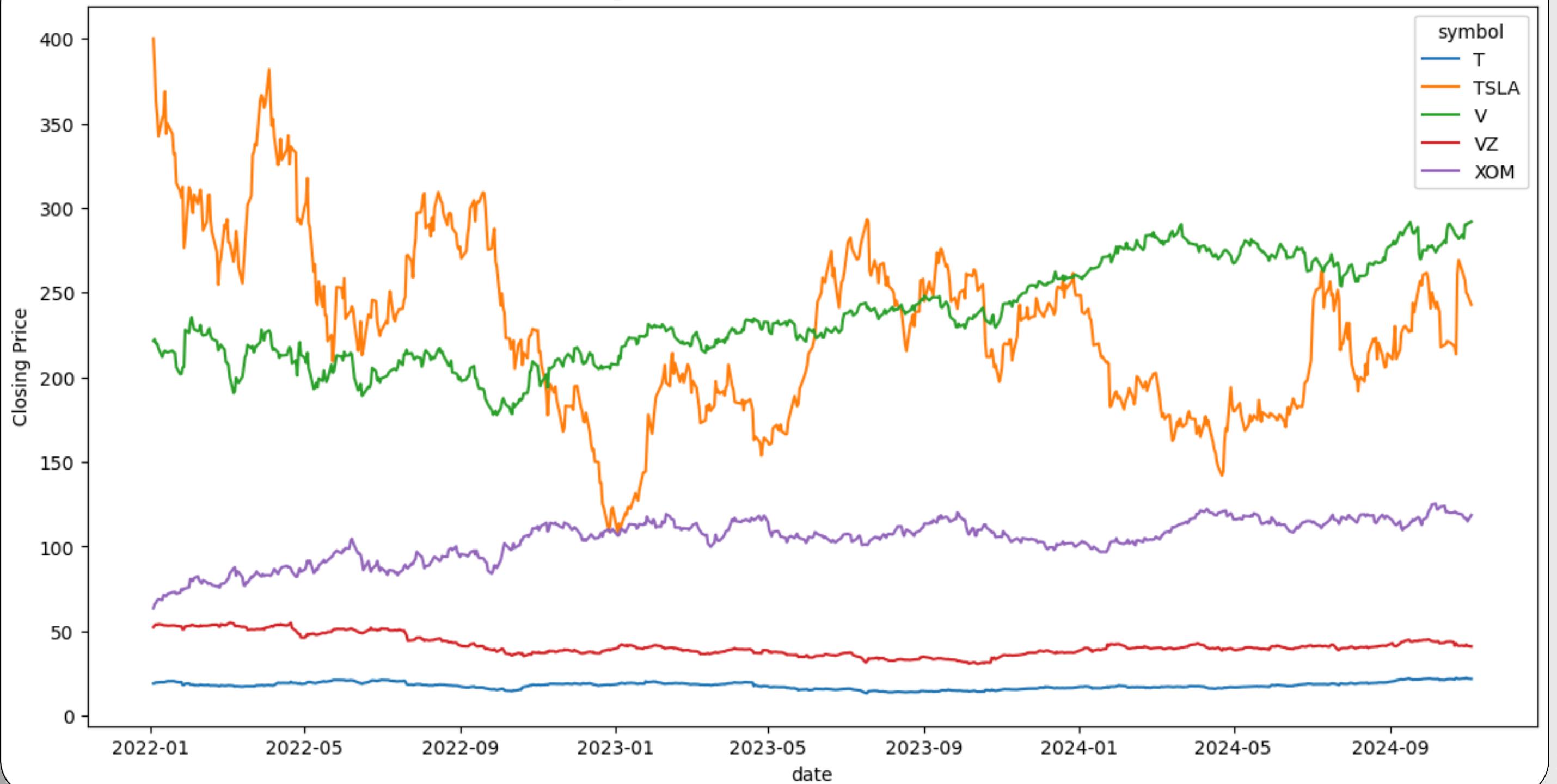
[https://public.tableau.com/views/stocks\\_17301837374410/Dashboard1?:language=en-US&:sid=&:redirect=auth&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/stocks_17301837374410/Dashboard1?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link)





- Stocks like AAPL and AMZN have notably higher trading volumes, indicating greater investor interest and liquidity. The spikes in volume correspond to major company announcements or market events.
- Other stocks, such as ABT and CMCSA, show consistently lower trading volumes, signifying that they are less actively traded.

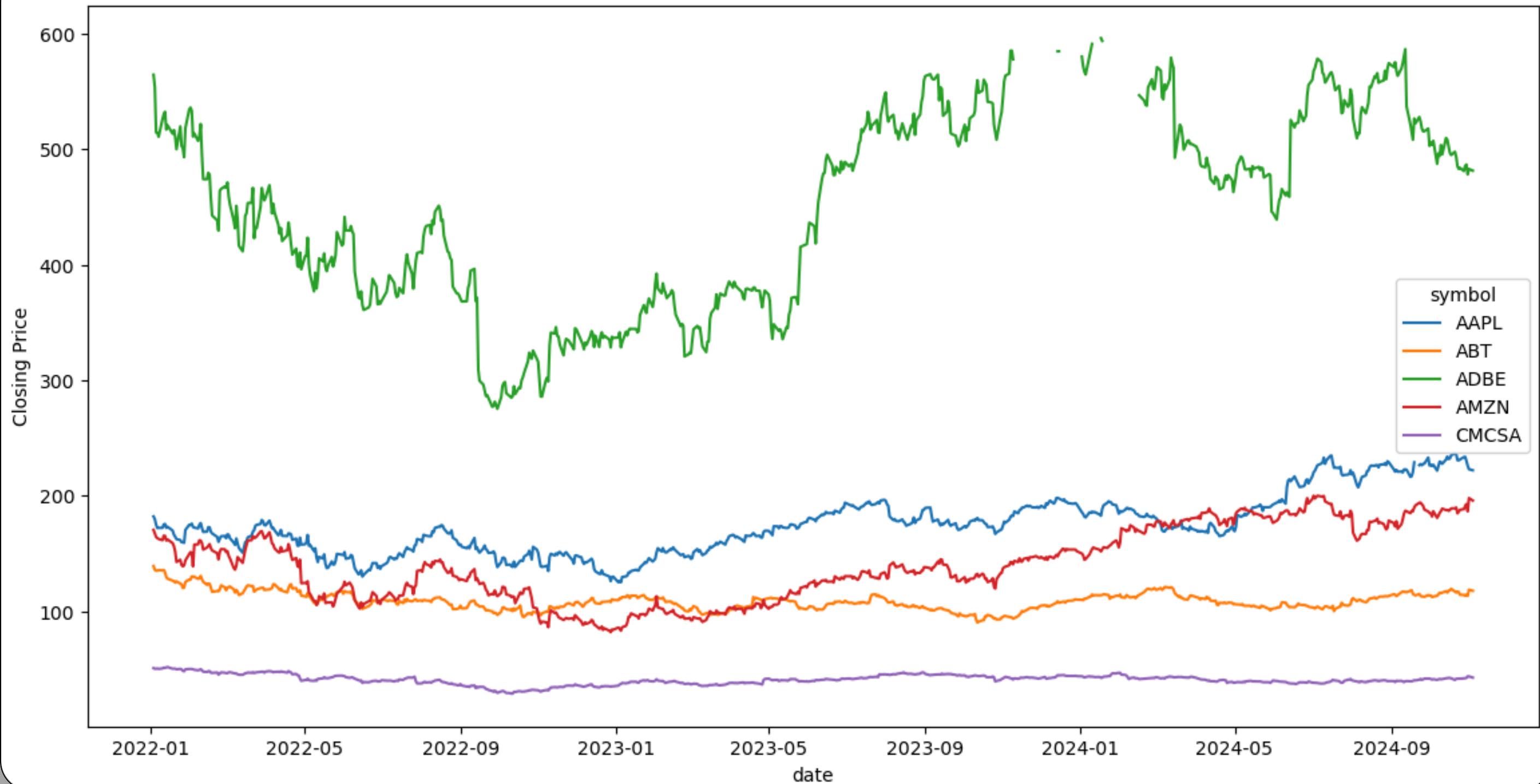
## Closing Prices of The Last Five Stocks Over Time



- TSLA shows significant volatility, with large price fluctuations over time.
- Stocks like T and VZ have relatively stable prices at lower levels

# EDA

Closing Prices of the First Five Stocks Over Time



- ADBE has notably higher volatility and a wider price range
- Stocks like CMCSA, ABT, and AMZN are more stable and remain within lower price ranges.

# FEATURE ENGINEERING AND MODELING



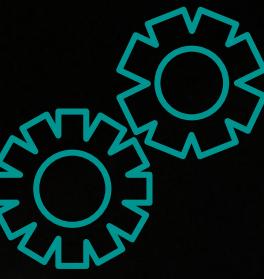
## FEATURE ENGINEERING

We generated the following features:

- **Daily Returns** - to measure stock price change between consecutive days
- **Moving Averages (SMA & EMA)** - SMA provides a smooth average price over a period, while EMA assigns more weight to recent prices, making it more responsive to changes.
- **Volatility**: - Indicates how much the price fluctuates over a period
- **Bollinger Bands**: - A volatility indicator with upper and lower bands. If the price breaks these bands, it may indicate potential reversals.

# FEATURE ENGINEERING AND MODELING

---



## MODELLING

### Agglomerative clustering:

We used Agglomerative clustering but it yielded a lower silhouette score and resolved to proceed with KMeans for clustering.

### KMeans:

- We used KMeans with 3 clusters to categorize stocks into risk groups: Low, Medium, and High risk.

### Cluster Labeling:

- A dictionary (cluster\_mapping) was created to map cluster labels (0, 1, 2) to risk categories: Low Risk, Medium Risk, and High Risk.
- This makes it easier to interpret clusters in terms of risk.

# EVALUATION



## Metric Used: Silhouette Score

- The Silhouette Score is a measure of how similar an object is to its own cluster compared to other clusters.

## Silhouette Score Interpretation

Our Model's Score: 0.595

- A score of 0.595 suggests a moderately well-defined structure in the clusters.
- This value indicates that overall the data points are reasonably close to their assigned clusters.



# CONCLUSION

- We Built a web application where the user is able to key in their risk appetite and receive recommended stocks that they can invest in.





MORINGA  
Discover • Grow • Transform

# Thank you!