

# PREDICTING THE STOCK MARKET



# US STOCKS AND THE MAG 7

- A combined market capitalization of about \$14 trillion
- Exposure to high-growth technologies such as high-end software and hardware, cloud computing, and Artificial Intelligence
- Each of the seven stocks has outperformed the S&P 500 by a huge margin in the past decade



# THE BIG IDEA

- Predict the long-term performance of any large-cap stock over an year based on –
  - Fundamental Factors – EPS and P/E Ratio
  - Technical Factors – GDP Growth, Inflation, Unemployment Rate, Retail Sales and Industrial Production
  - News and Market Sentiment



# THE DATA SCIENCE APPROACH

- Gather last 10 years' data from multiple sources and APIs
- Clean and Pre-process data
- Preliminary EDA
- Machine Learning – Train a regression model on a subset of the data
- Test the model on another subset of data
- Predict the performance of the Mag-7 stocks over the next year



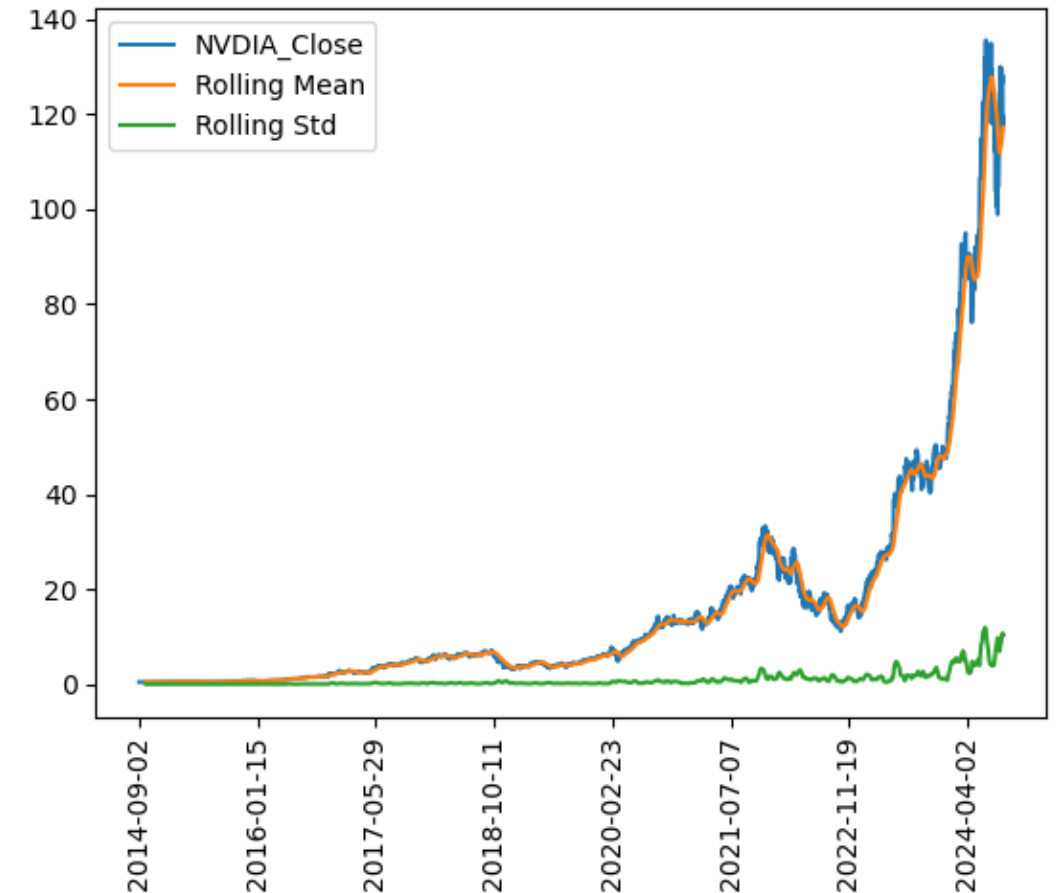
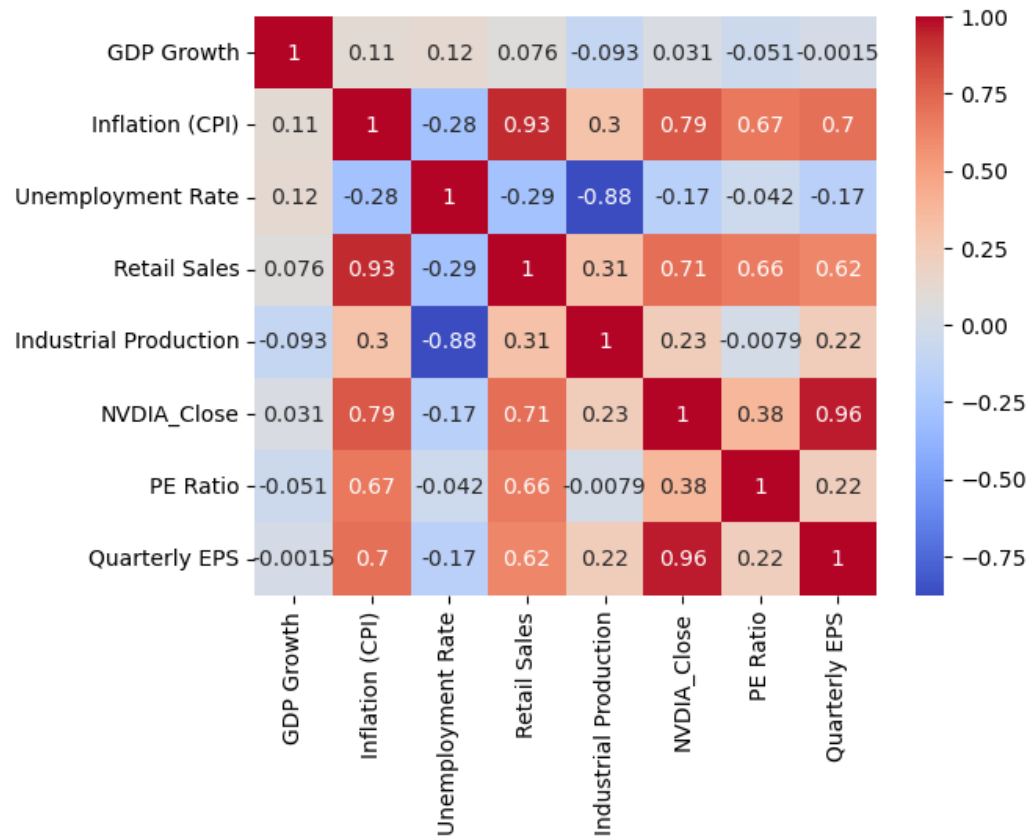
# DATA COLLECTION AND PRELIMINARY EDA

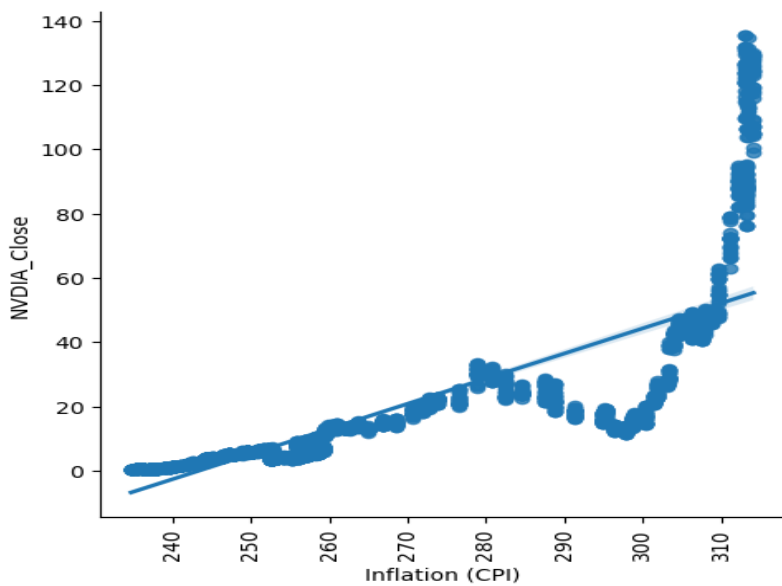
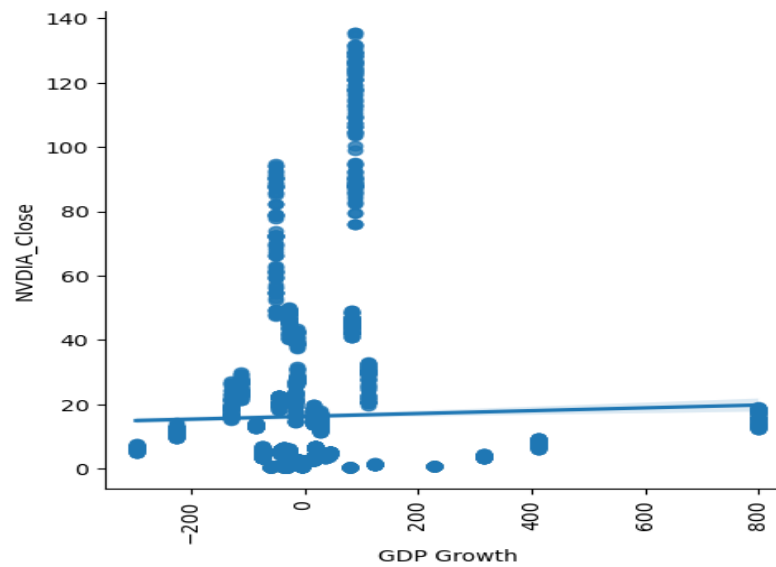
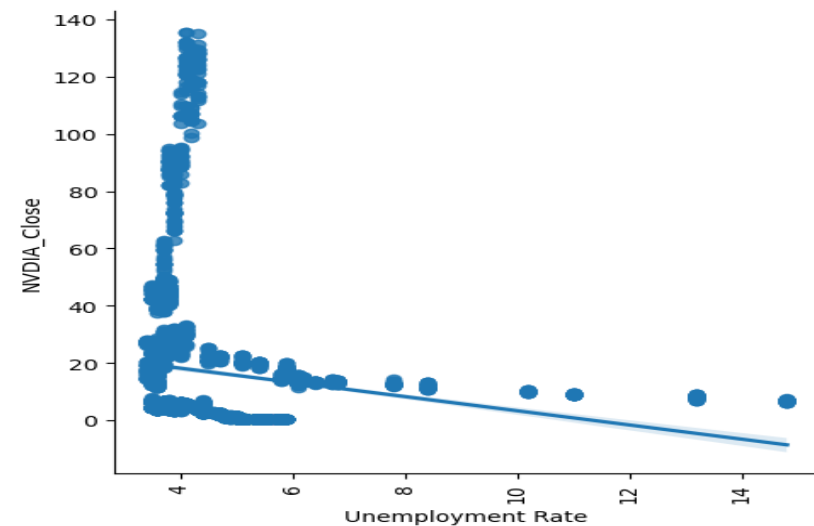
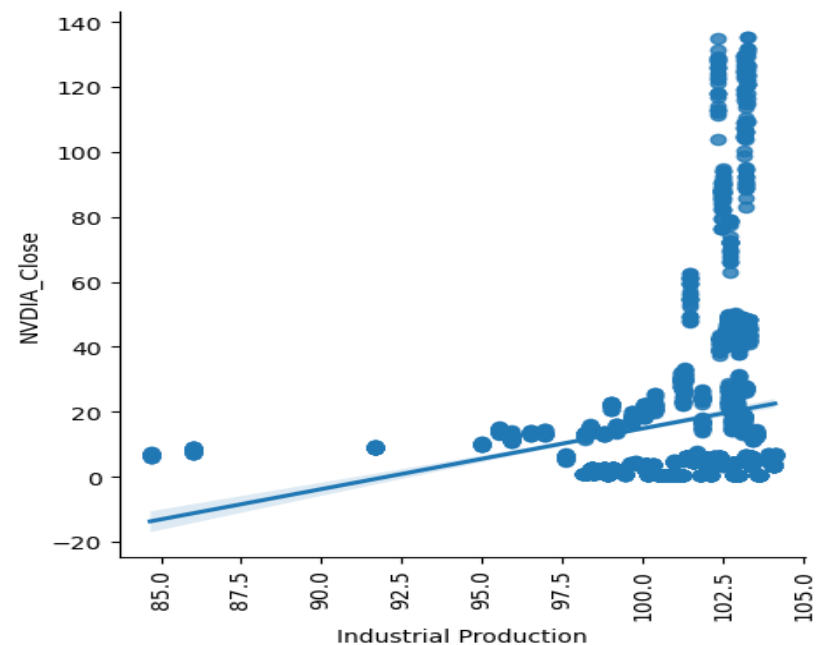
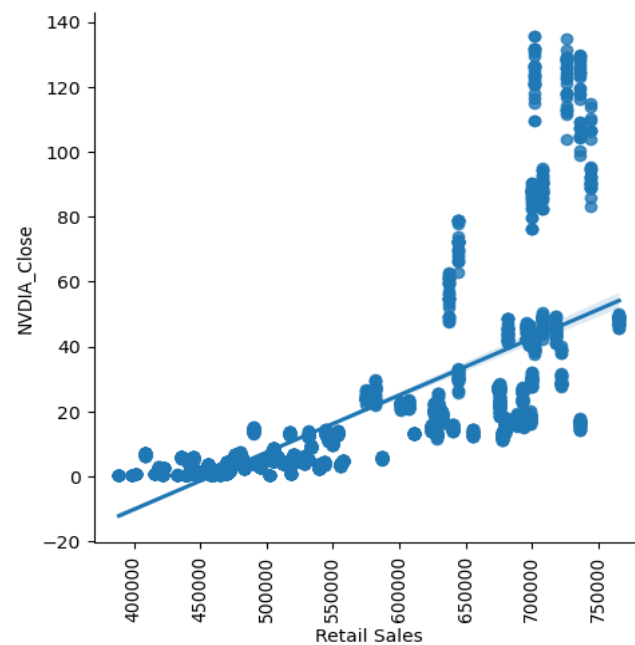
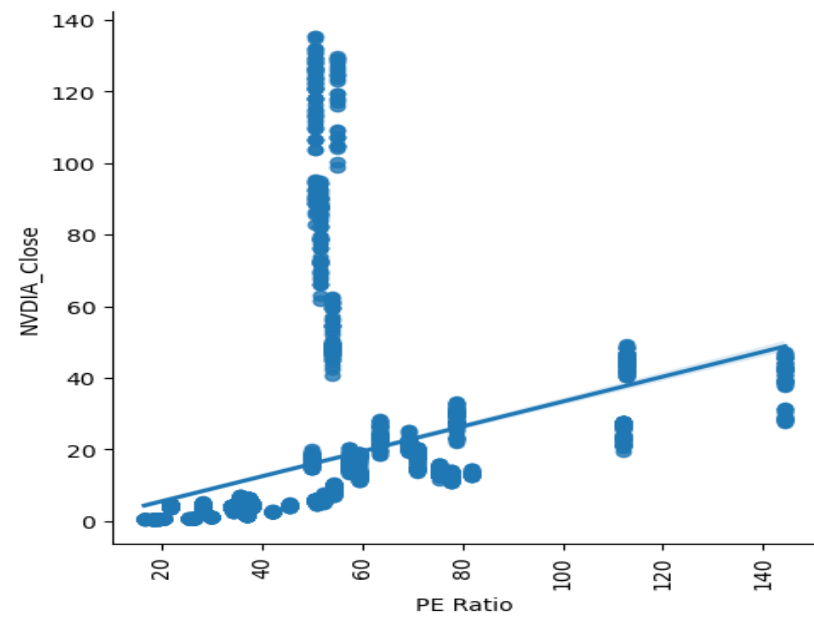
- Stock Price data for the Mag-7 stocks collected from yfinance API using python.
- Macroeconomic indicators data collected from the FRED API
- EPS and PE Ratios collected from “Macrotrends” website by saving the webpage and extracting data through pandas
- Forward filling of quarterly data across multiple dates
- Joined data from different sources to enable processing
- Created initial multivariate and univariate visualizations





# DATA COLLECTION AND PRELIMINARY EDA





By - Rohit Grover



# FUTURE WORK

- Sentiment Analysis of News and Social Media data related to the stocks and feed them into the model
- Feature engineering and training the model





WALL ST  
55

**THANK YOU**

