COMPREHENSIVE STOCK MARKET ANALYSIS OF LEADING TECH GIANTS

SUBMITTED BY:

AISHARRIYA DASGUPTA (23060242003)

VIDIT VASWANI (23060242098)

CONTENTS

INTRODUCTION

The stock market serves as a vital component of the global economy, providing a platform for companies to raise capital and for investors to buy and sell shares. The stock market plays a pivotal role in determining the financial health and growth potential of companies, particularly in the fast-evolving technology sector. As financial markets continue to evolve, the need for comprehensive analysis of stock performance has become increasingly important. This project focuses on analyzing the stock market data of six influential companies: Apple, Facebook, Google, NVIDIA, Tesla and Twitter.

With the advent of technology and data science, investors and analysts now have access to vast amounts of historical and real-time data. By leveraging programming languages like Python, we can perform sophisticated analysis that uncover trends, patterns, and insights that may not be immediately apparent through traditional analysis methods.These companies, each representing different facets of the technology industry, have demonstrated significant influence on both market trends and global economic patterns.Our goal is to explore various analytical techniques, including technical analysis and time series analysis, to assess the performance of these companies in the stock market.

Through this exploration, we aim to provide valuable insights that can aid investors in making informed decisions. By understanding the historical performance and potential future trends of these stocks, we can contribute to a more nuanced understanding of the financial landscape.
This project not only enhances our programming skills but also deepens our understanding of financial markets and investment strategies. Ultimately, it represents an exciting opportunity to bridge the gap between technology and finance in today's data-driven world.

BACKGROUND

This case study presents an in-depth exploration of the stock market performance of six prominent companies—Apple, Google, Facebook, Nvidia, Tesla, and Twitter. These companies are global leaders in their respective industries, making their stock performance a focal point for both institutional and individual investors. The dataset used for this analysis is sourced from Kaggle and contains a wide range of financial metrics that are instrumental in evaluating the behavior of these companies in the stock market over time. This dataset comprises essential indicators such as opening and closing prices, daily high and low prices, trading volume, and adjusted closing prices, among others.

Such financial data is crucial for understanding the factors influencing the stock market, assessing the financial health of these companies, and predicting future price movements. Analysts, investors, and portfolio managers can leverage this information to make informed decisions, assess risks, and seize opportunities in a highly dynamic market environment.

EXPLORATORY QUESTIONS

1.What do the historical trends in adjusted closing prices and trading volumes of the companies indicate about market activity over time?

2. What do moving averages reveal about short-term versus long-term trends in stock prices?

3. What is the correlation between the returns of different companies, and how can this information guide investment decisions?

4. Are the adjusted closing prices stationary, and what implications does this have for time series forecasting?

5.How accurately can we forecast future stock prices using ARIMA modeling on historical data?

OBJECTIVES OF THE STUDY

1. To visualize and analyze the historical trends in the adjusted closing prices and trading volumes and their fluctuations over time for each company to understand market activity.

2. To assess short-term and long-term trends using moving averages for each company's stock price.

3.To analyze the relationships between the returns of different companies through a correlation matrix and heatmap.

4.To perform stationarity testing and autocorrelation analysis of adjusted closing prices in order to assess their suitability for time series forecasting.

5.To develop a predictive model using ARIMA on a company's adjusted closing price data to forecast future stock prices.

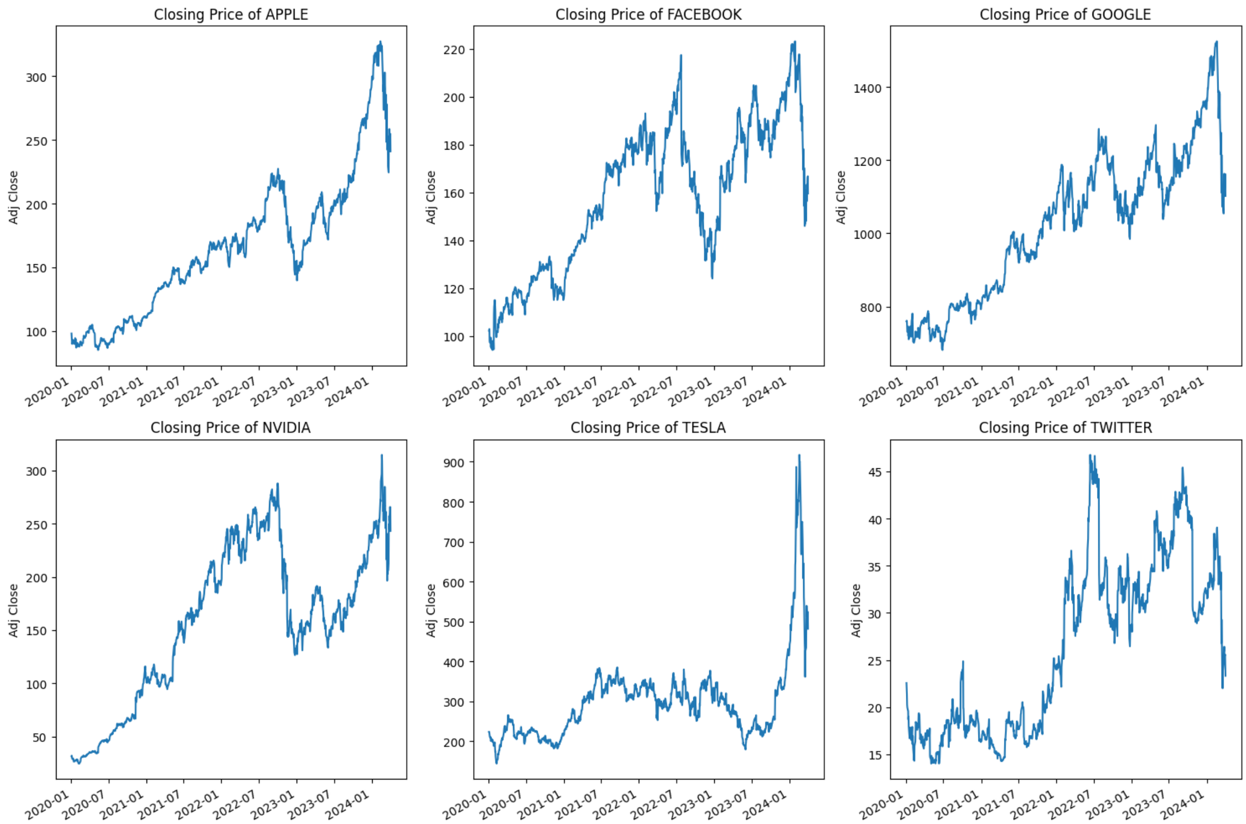
ANALYSIS TECHNIQUES

The techniques used for analysis are:

- **Technical Analysis**: Leveraging historical price and volume data to identify trends, patterns, and indicators that may predict future price movements.
- **Data Visualization**: Developing clear and intuitive visualizations to present stock performance, enabling quick interpretation of complex data and aiding in decision-making.
- **Time Series Analysis**: Examining stock prices over time to detect underlying trends, cycles, and seasonal variations, providing insights into market dynamics.

EXPLORATORY DATA ANALYSIS

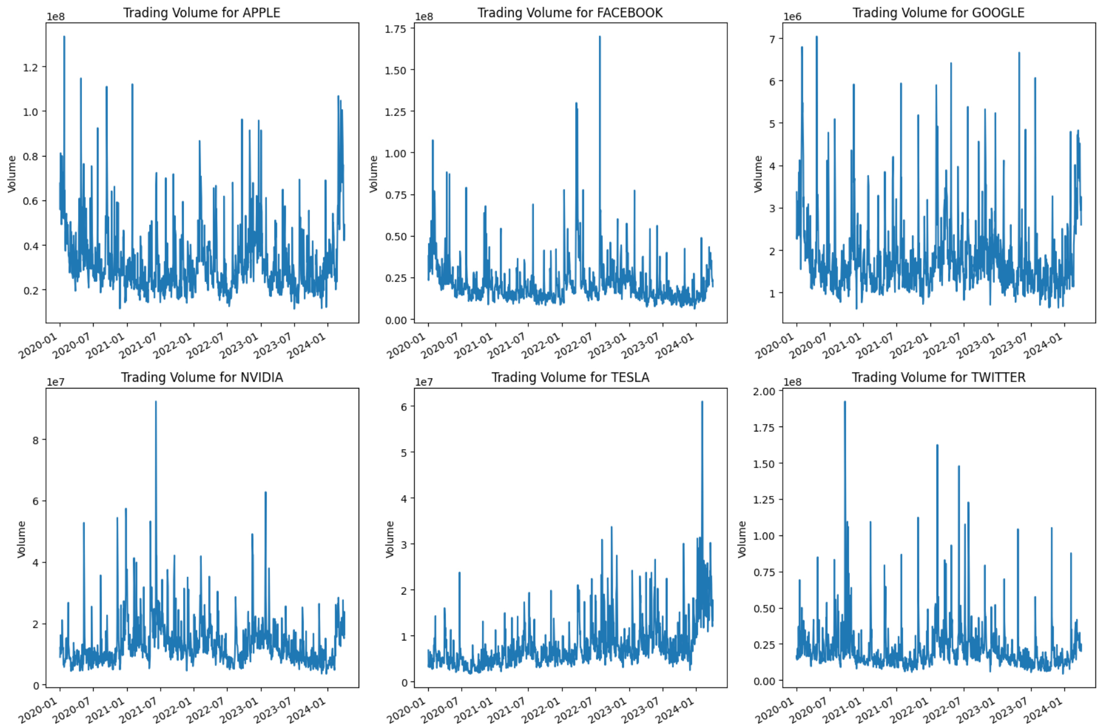
1. Closing Price of All Companies



This graph shows the historical closing prices for the six companies. The closing price represents the final price at which the stock was traded on any given day. Long-term trends in the closing price help us understand how the market perceives the value of each company. For example, companies like Tesla and Nvidia may show significant upward trends, reflecting investor optimism about future growth in electric vehicles and AI. Some stocks, like Tesla, might exhibit high volatility with sharp spikes and drops, indicating that the market views these stocks as riskier. In contrast, a more stable company like Apple might show smoother price movements. From a time series perspective, price data is often non-stationary, meaning that its statistical properties (mean, variance) change over time. This necessitates differencing the data before applying ARIMA models to ensure stationarity. Additionally, the presence of trends, cycles, and seasonality should be carefully considered, as they can inform the choice of model parameters (AR, MA components) to accurately forecast future prices.

2. Trading Volume of All Companies

This graph represents the number of shares traded for each of the six companies over a defined period.



Trading volume is a measure of how actively a stock is being bought or sold. High trading volumes typically suggest significant market activity, often in response to company-specific news (e.g., earnings reports, new product launches) or macroeconomic events (e.g., interest rate changes). Peaks in trading volume can be linked to investor sentiment. For instance, higher volumes during periods of economic downturns might indicate that investors are reacting to fear or uncertainty, whereas during booms, increased volume could signal optimism. Trading volume can often be correlated with price volatility. In time series models, this relationship can be quantified by analyzing the autocorrelation of trading volumes, which often exhibit a pattern of persistence (i.e., high volumes tend to follow high volumes). This can be modeled using ARIMA (Autoregressive Integrated Moving Average) or similar approaches to understand the lagged effects of trading volume on price movements.

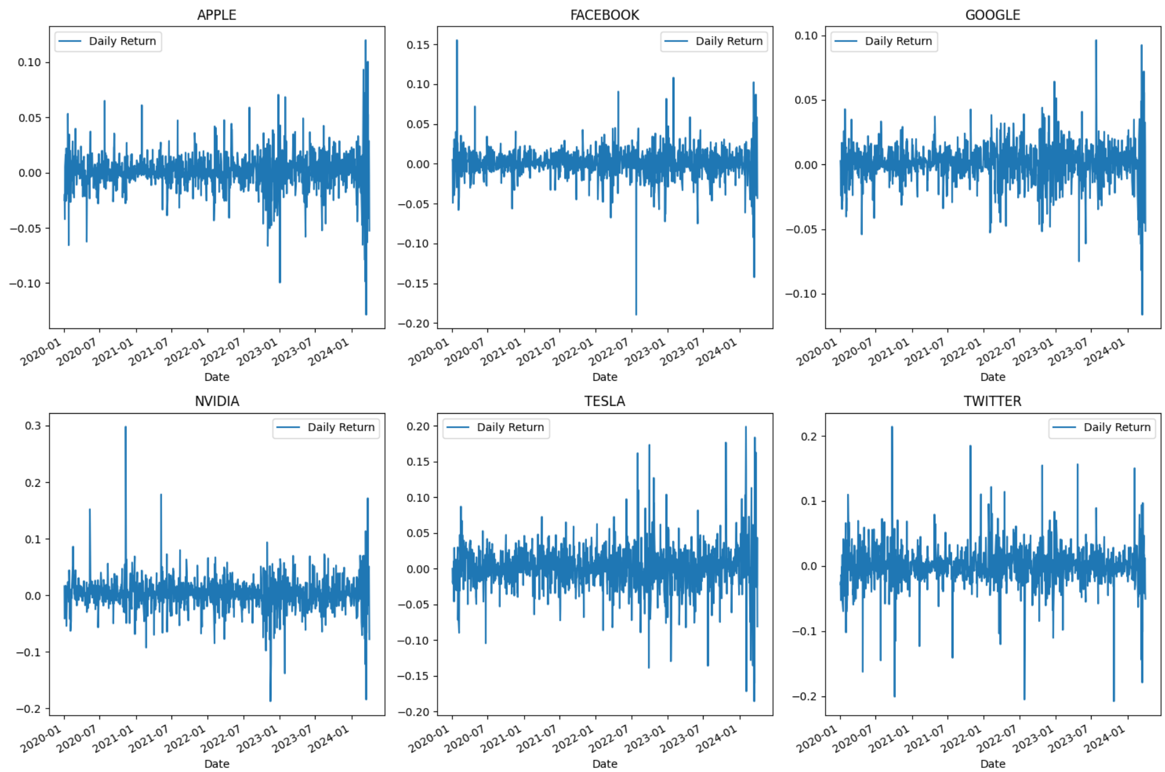
3. Adjusted Closing Price and Moving Averages (10, 20, 50 days)

The adjusted closing price accounts for corporate actions like dividends or stock splits, while the moving averages (10-day, 20-day, 50-day) smooth out short-term fluctuations in stock prices.

Adjusted Closing Price provides a more accurate representation of a stock's value over time because it factors in events that could distort the true price (e.g., stock splits). This is essential for long-term price comparisons. The moving averages help smooth out volatility in the stock prices. A 10-day moving average will show short-term trends, reacting quickly to price changes, while the 50-day moving average highlights long-term trends. If a stock's adjusted closing price stays consistently above its moving average, it indicates an uptrend, which might signal bullish market sentiment. Conversely, if it drops below, it could indicate a downtrend. Moving averages are particularly useful in identifying trends and smoothing out noisy data. When applying ARIMA models, identifying whether the series is trending upwards or downwards can determine whether differencing is needed. The differences between the moving averages can also suggest potential turning points (crossovers) where trends may reverse.

4. Daily Returns for All Companies

This graph shows the daily percentage change in stock prices for each company, which gives a view of the volatility and market risk associated with the stocks.

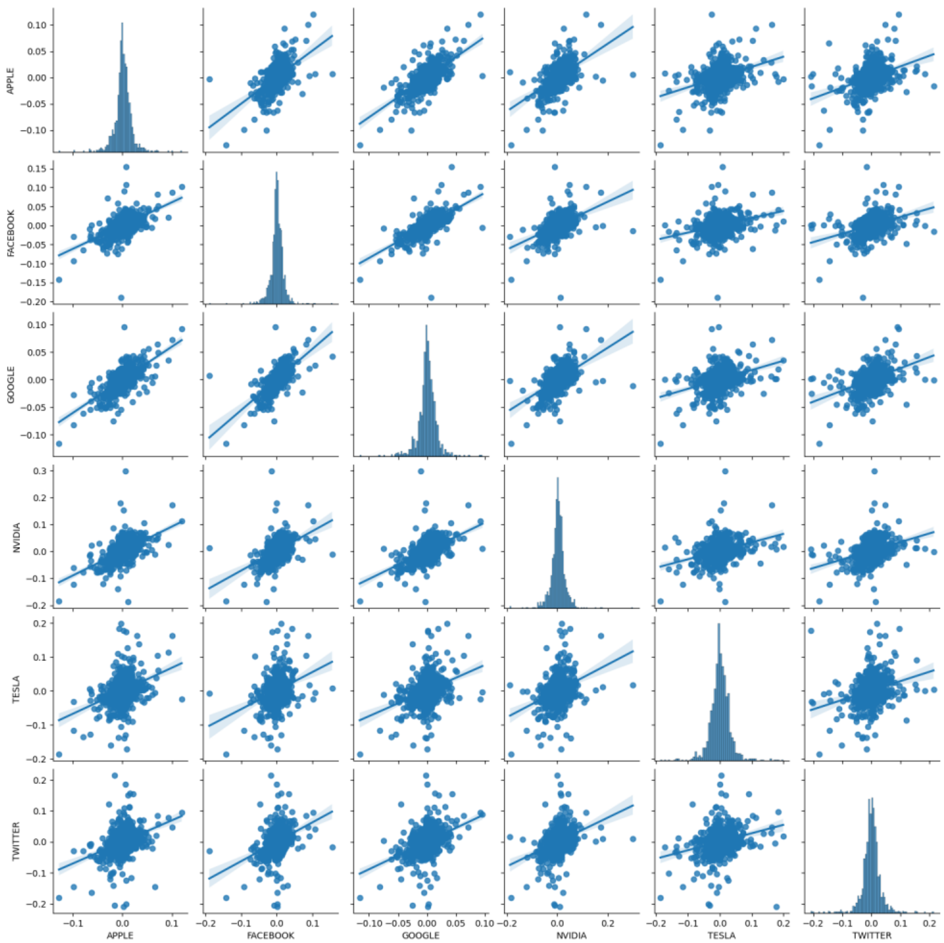


Daily returns are a more precise measure of how risky a stock is. For instance, Tesla might have more extreme spikes in daily returns compared to Apple, reflecting higher volatility. These large swings could stem from news, earnings surprises, or industry shifts. Daily returns reflect short-term market sentiment. Higher volatility means higher risk, which might not be acceptable for all investors. Understanding these fluctuations is crucial in portfolio management and risk assessment. Daily returns are generally stationary, making them suitable for direct ARIMA modeling without the need for differencing. However, the variance in returns (i.e., volatility) often changes over time, which suggests the need for models like GARCH (Generalized Autoregressive Conditional Heteroskedasticity) to capture periods of high and low volatility. Modeling these dynamics is key for accurate financial forecasting, as volatility clustering (periods of high volatility followed by more high volatility) is a common phenomenon in financial markets.

STATISTICAL ANALYSIS

1. Pairplot of Percentage Change of Returns

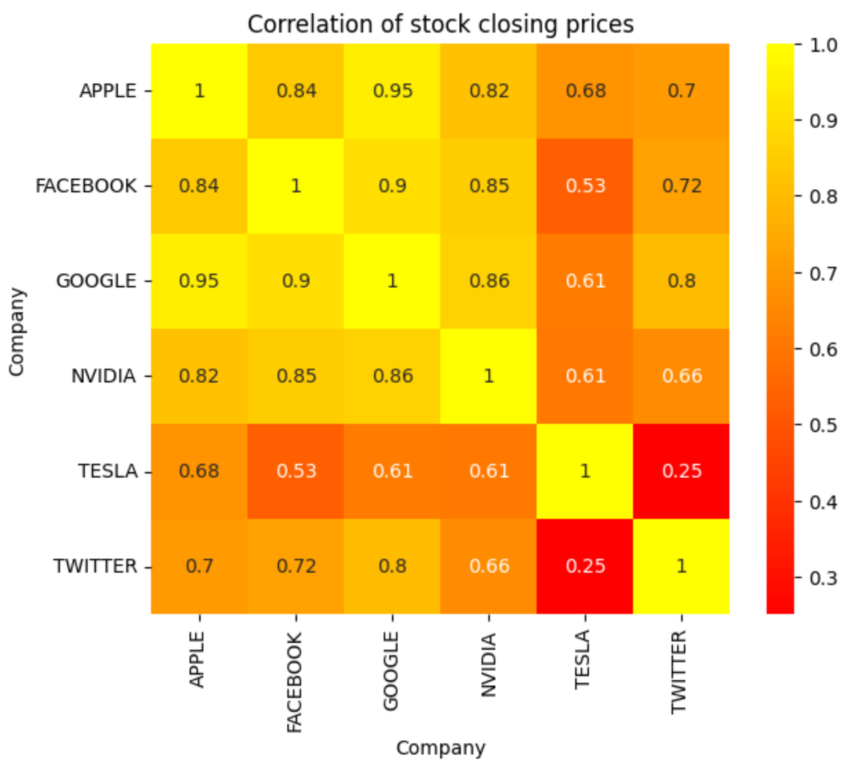
This is a scatterplot matrix showing the pairwise relationships between the percentage changes (returns) of the six companies.



If two stocks have a strong positive relationship in their percentage change, they are likely reacting to the same economic or industry factors. For instance, tech giants like Apple and Google may show similar reactions to tech sector trends, regulation, or macroeconomic conditions like interest rates. For investors, identifying pairs of stocks that are weakly correlated or negatively correlated can provide diversification benefits. For example, if Apple and Nvidia show weak correlation in returns, holding both in a portfolio could reduce risk. Correlations between returns can guide multivariate time series models like VAR (Vector Autoregressive) models. A high correlation suggests that stocks are influenced by common factors, while low correlation implies more independent movements. In an ARIMA context, these relationships can be useful for understanding how the performance of one stock may help predict another.

2. Correlation of Stock Closing Prices

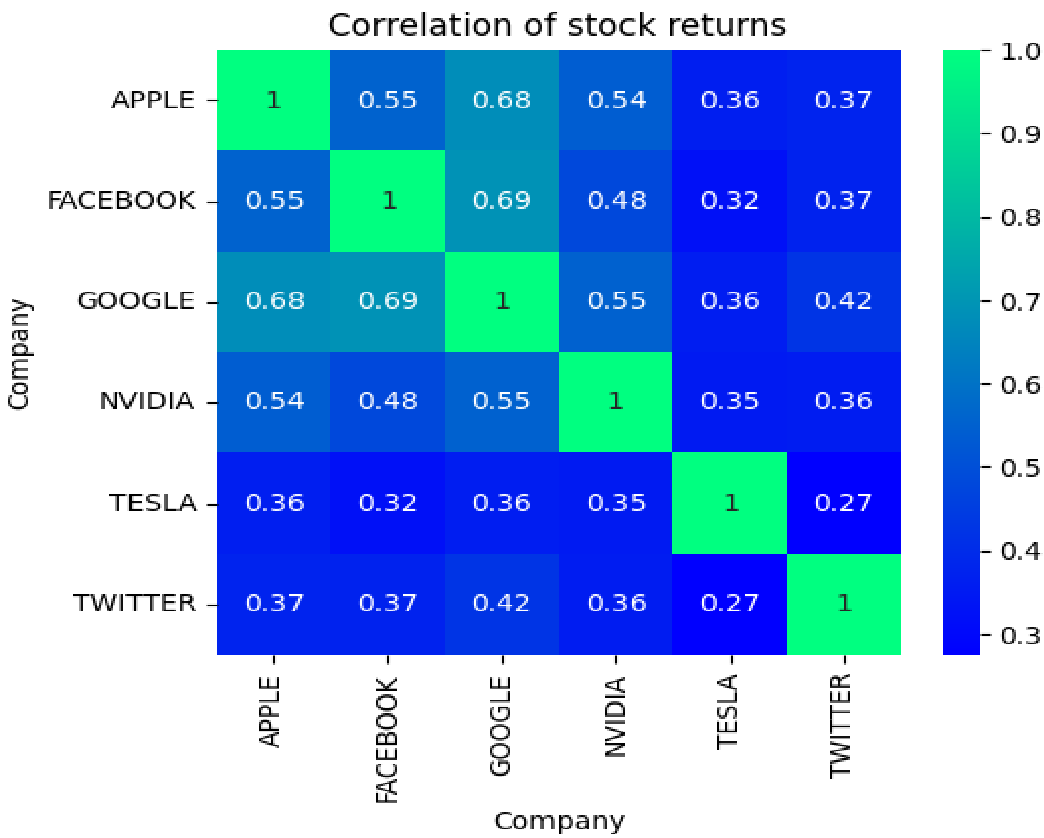
This heatmap shows the correlation coefficients between the closing prices of the six companies.



A high correlation (near 1) indicates that two companies' stock prices tend to move in the same direction. For example, Apple and Google, both major technology companies, may exhibit strong positive correlation due to their shared exposure to tech industry dynamics. A low or negative correlation means that the stocks tend to move in opposite directions or independently of each other. This is useful for portfolio diversification, as stocks that are not correlated can reduce overall portfolio risk. Correlations in price data can inform the econometric modeling process. For instance, if stocks show high correlation, it may suggest that external macroeconomic variables (like interest rates or inflation) are influencing both stocks similarly. Modeling these correlations helps in understanding the co-movements and may be useful in constructing a co-integrated time series model if the stocks are related in the long run.

3. Correlation of Stock Returns

This heatmap shows the correlation between the daily returns of the six companies.

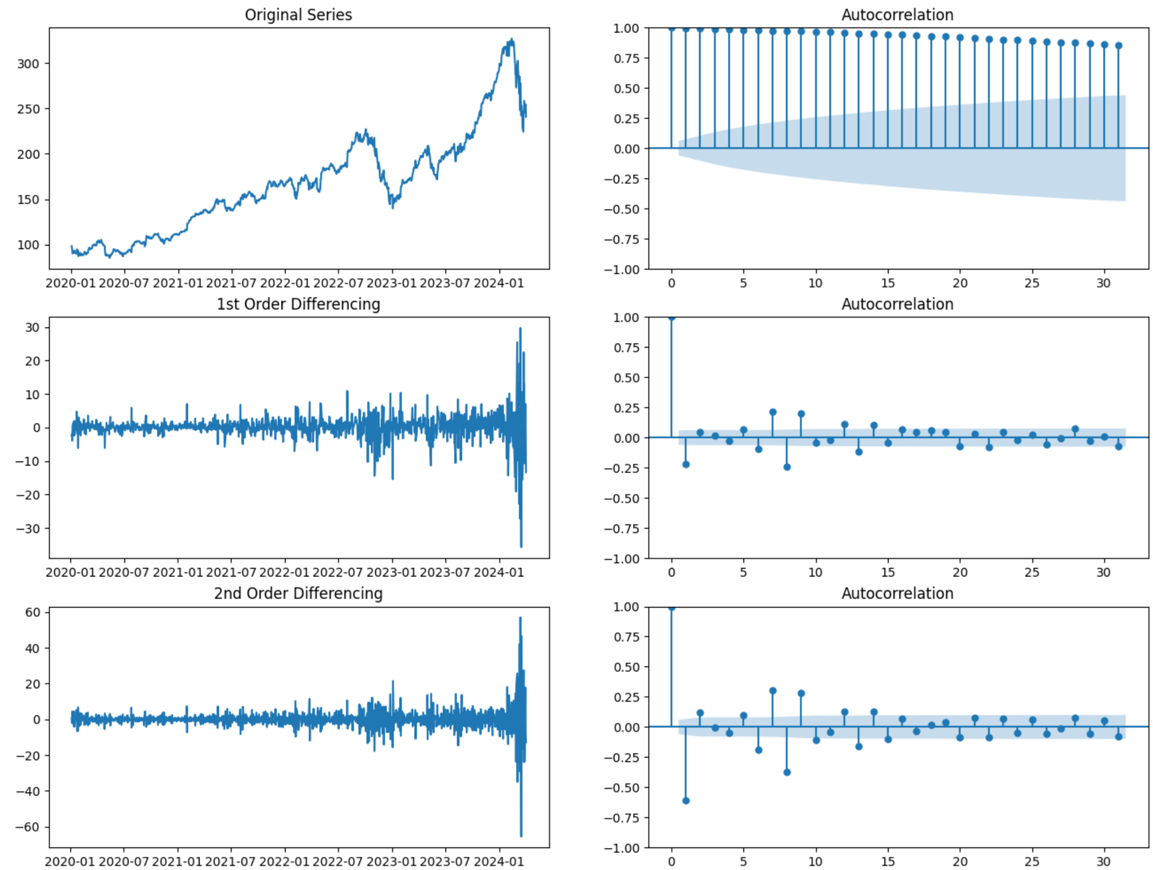


If two companies' returns are highly correlated, they are likely influenced by similar market events. For example, companies in the tech sector like Google and Nvidia might exhibit a high correlation because they are impacted similarly by sector-wide changes such as new technological advancements or regulation. Lower correlation implies that the stocks are responding to different factors, which can provide diversification benefits for investors. A well-diversified portfolio would ideally include stocks with low correlation to balance risk. When modeling returns using ARIMA or GARCH, correlation helps in identifying co-movements. A highly correlated set of returns suggests common volatility patterns, and models like multivariate GARCH (MGARCH) can be used to predict the time-varying volatility of returns across multiple stocks.

MACHINE LEARNING MODEL

1. Apple Stock Original and Differenced Series with ACF and PACF

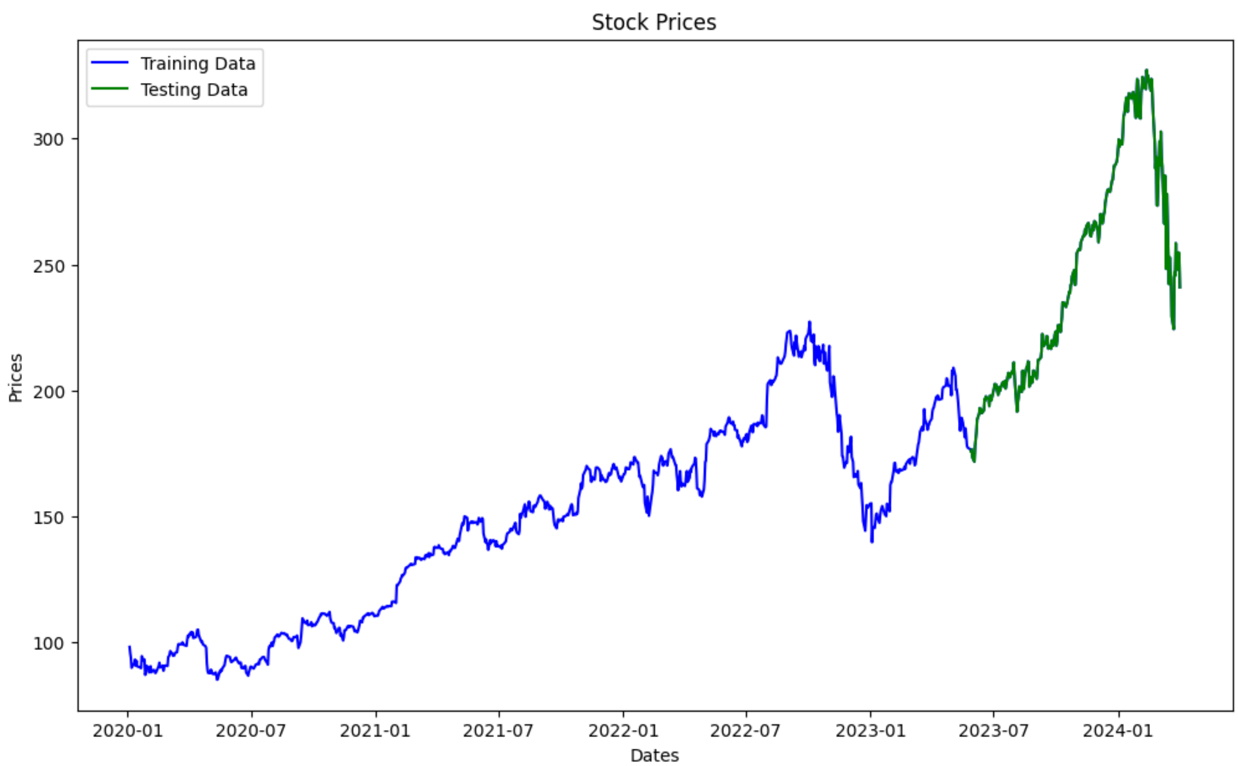
This set of graphs shows the original time series for Apple's stock prices, the differenced series, and the corresponding Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots.



The original series likely exhibits trends or non-stationarity, which violates the assumptions of ARIMA models. Differencing the series (by subtracting consecutive values) removes these trends, making the series stationary and suitable for ARIMA modeling. The ACF plot shows how the current value of the time series is related to its past values (lagged observations). If the ACF decays slowly, it suggests that the time series is non-stationary, while a quick drop-off suggests stationarity. The PACF plot helps identify the number of autoregressive (AR) terms in the ARIMA model. A significant spike at the first lag in the PACF plot suggests an AR(1) process, whereas additional significant spikes suggest higher-order autoregressive terms. ACF and PACF are essential diagnostic tools for identifying the appropriate ARIMA model. In an economic context, using these tools helps determine whether a stock's price is influenced more by its own historical values (AR terms) or by past shocks (MA terms). These insights guide the choice of ARIMA parameters (p, d, q), ensuring the model adequately captures the time series dynamics.

2. Apple Stock Train and Test Data Plot

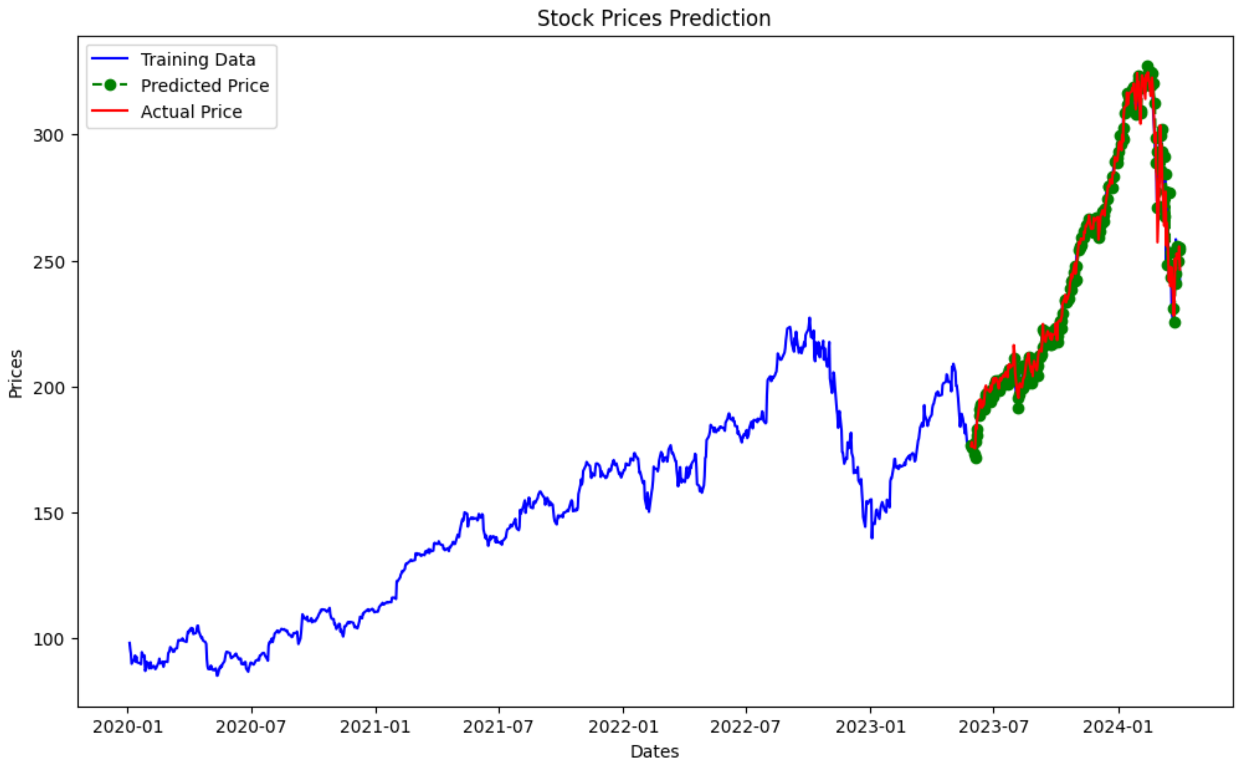
This plot splits the Apple stock data into training and test sets, with the training set used to fit the ARIMA model and the test set used to validate the model's predictions.



The purpose of separating data into training and test sets is to evaluate the forecasting accuracy of the ARIMA model. The model is fitted on the training data, and its forecasts are compared against the actual data in the test set. If the predicted stock prices in the test set closely follow the actual stock prices, the model is performing well. Large deviations would suggest that the model is either underfitting (failing to capture key dynamics) or overfitting (capturing noise instead of real trends). The split between training and test sets is a critical step in evaluating a model's out-of-sample forecasting ability. Econometric models are often evaluated based on their ability to generalize beyond the data used to train them, making this step essential for assessing the real-world applicability of the ARIMA model.

3. Apple Stock Prices Prediction Plot

This graph shows the actual Apple stock prices and the predicted prices generated by the ARIMA model.



The degree to which the predicted prices match the actual prices indicates the accuracy of the ARIMA model. A close match suggests the model has successfully captured the underlying trends and patterns in the data. If the model continues to track the actual prices well into the future, it indicates robustness. However, if the predictions start to deviate significantly from actual prices, it may suggest the need for more complex models, such as adding exogenous variables (macroeconomic factors) or using more advanced techniques like ARIMAX (ARIMA with exogenous variables). A successful prediction model should be able to provide confidence intervals for the forecasts. The econometric emphasis here would be on the predictive power of the ARIMA model. The residuals (differences between actual and predicted values) should ideally be white noise, meaning no discernible patterns are left unmodeled. If residuals show patterns, the model may need further refinement.

FUTURE SCOPE

The stock market analysis employs ARIMA modeling to forecast the adjusted closing prices of six companies by leveraging key time series concepts like autocorrelation and stationarity. ARIMA's strength lies in transforming non-stationary stock prices into a stationary series, allowing for reliable short-term forecasting. Additionally, the analysis of daily returns highlights volatility, especially for stocks like Tesla and Nvidia, suggesting the potential for further refinement with GARCH models to capture time-varying volatility. The correlation of returns and prices across stocks provides insights into market co-movements and diversification opportunities, making the analysis useful for portfolio management. By splitting data into training and test sets, your model's forecasting accuracy is evaluated, and residual analysis ensures that the model captures all significant patterns. Although ARIMA is effective for time series forecasting, incorporating exogenous factors like interest rates or macroeconomic indicators could enhance predictive power. Ultimately, the analysis offers practical forecasting tools for decision-making, risk management, and economic forecasting, though further refinements such as GARCH or multivariate models (VAR) could improve the model's robustness in highly volatile market conditions.

Appendix

The links to the codes and dataset have been attached below:

1. **Google Colab Notebook**
   https://colab.research.google.com/drive/1gCGsqVW-GfIirbMyNfhCKUKhlxDkfWOr?usp=sharing

2. **Dataset**
   https://docs.google.com/spreadsheets/d/1Y4oo41fkgXjevP3JPBOgnN_aWebul5pZqgr-pERoOxs/edit?usp=sharing