

BUAN.6312.003 APPLIED ECONOMETRICS AND TIME SERIES ANALYSIS

STOCK MARKET ANALYSIS AND PREDICTION USING LSTM

Group 5:
Bora Jagannadh
Philip Robbin
Pillai Abhirami Radhakrishnan
Pillai Shubhangi R
Ramaiahgari Bhanu Sree
Sathravada Anusha
Thamatam Vinay
Vaibhav Snehil
Venkatesan Nitisha Sree

CONTENT

3

Introduction

4

Literature Survey

5

Dataset Overview

6

Testing Procedures

7

Results

8

Conclusion



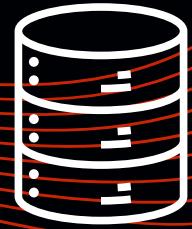
INTRODUCTION

- Predicting the ever-changing stock market is a complex task due to its inherent volatility and dependence on various factors.
- **Introducing LSTM:** This presentation explores Long Short-Term Memory (LSTM) networks, a powerful deep learning technique for analyzing and forecasting time series data like stock prices.
- **Unlocking Insights:** We'll delve into how LSTMs can capture both short-term and long-term dependencies within historical data, allowing for a more comprehensive market understanding.
- **Predictive Power:** Discover how LSTM models can be trained to predict future price movements, potentially aiding informed investment decisions.



LITERATURE SURVEY

- LSTM networks offer a potent solution for stock market analysis, leveraging their ability to capture short-term fluctuations and long-term trends within time series data.
- Through LSTM models, analysts can unlock valuable insights from historical stock price data, facilitating a deeper understanding of market dynamics and aiding in informed decision-making
- A key advantage of LSTM networks lies in their predictive power, enabling the development of models capable of forecasting future price movements with a certain degree of accuracy.
- Despite their promise, challenges such as data scarcity, model interpretability, and market volatility persist, necessitating ongoing research efforts to enhance the applicability of LSTM networks in stock market prediction.



DATASET OVERVIEW

- Our dataset encompasses one year of extensive historical stock market data for leading technology giants, including Apple (AAPL), Google (GOOG), Microsoft (MSFT), and Amazon (AMZN).
- Spanning from 05/2023 to 05/2024, the dataset comprises essential variables such as Open, High, Low, Close, Adjusted Close, and Volume, providing a comprehensive view of each company's performance over time.
- With a total of 3107^7 observations, the dataset offers rich insights into the dynamic nature of the stock market.
- The data was collected using the Yahoo Finance API through the **pandas_datareader** library, ensuring reliability and accuracy in our analysis.



TESTING PROCEDURES

- **Change in Price Over Time:** Plot historical closing prices to visualize price trends.
- **Daily Return of the Stock:** Calculate daily returns to assess daily performance.
- **Moving Average:** Compute moving averages to identify long-term trends.
- **Correlation Between Different Stocks:** Calculate correlations to understand how stocks move relative to each other.
- **Risk Analysis:** Assess risk using metrics like standard deviation or Value at Risk (VaR).
- **Predicting Future Stock Behavior:** Use LSTM to forecast future stock prices by training on historical data.

RESULTS

AAPL.info()

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 253 entries, 2023-05-01 to 2024-05-01
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Open         253 non-null    float64
 1   High        253 non-null    float64
 2   Low          253 non-null    float64
 3   Close        253 non-null    float64
 4   Adj Close   253 non-null    float64
 5   Volume       253 non-null    int64  
 6   company_name 253 non-null    object 
dtypes: float64(5), int64(1), object(1)
memory usage: 15.8+ KB
```

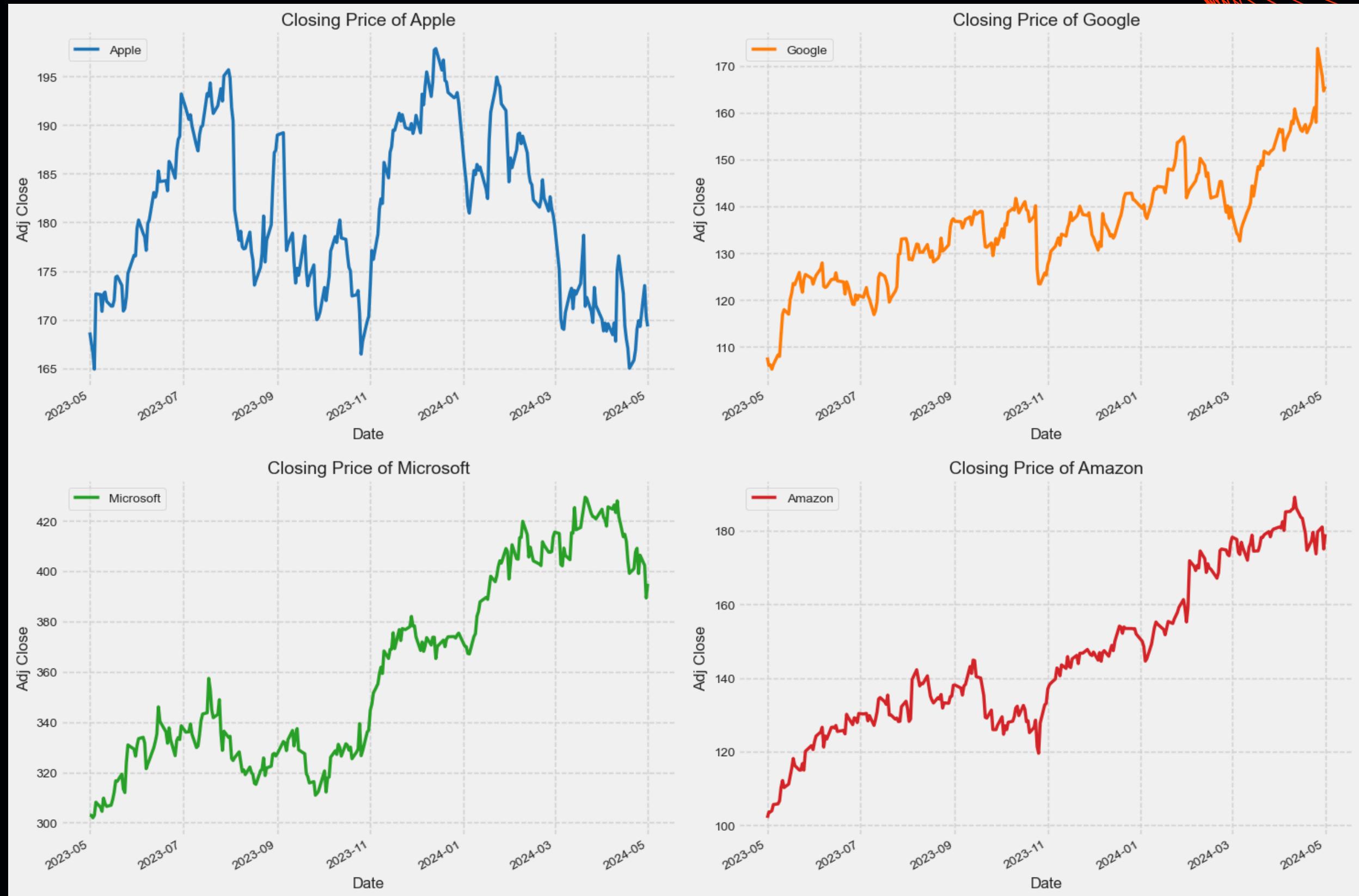
DataFrame including the index dtype and columns, non-null values, and memory usage.

Descriptive statistics summarize central tendency, dispersion, and distribution shape of the dataset, excluding NaN values. It analyzes both numeric and object series, as well as mixed data types within DataFrame column sets.

AAPL.describe()

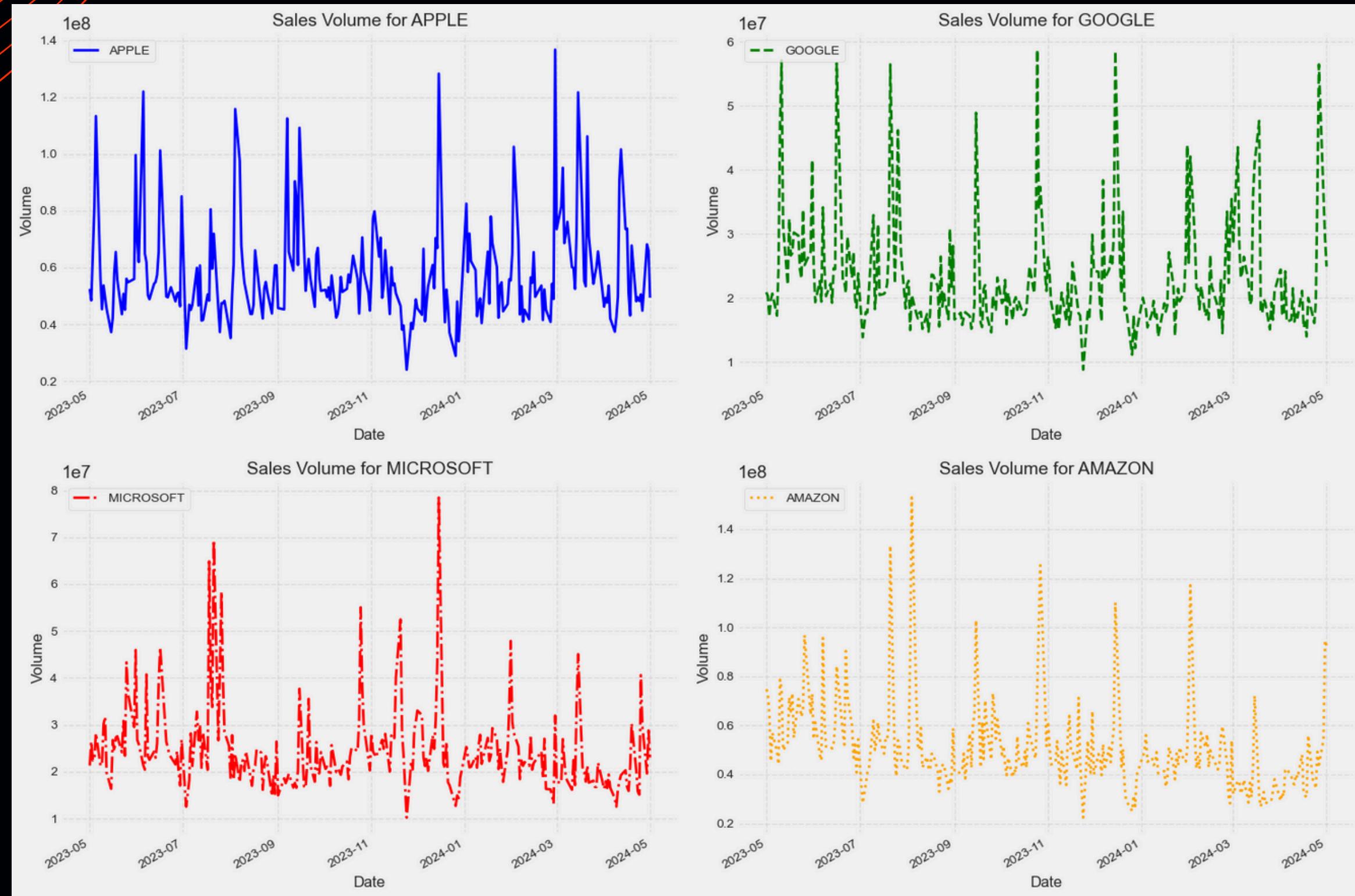
	Open	High	Low	Close	Adj Close	Volume
count	253.000000	253.000000	253.000000	253.000000	253.000000	2.530000e+02
mean	181.112174	182.583953	179.748577	181.207352	180.822665	5.786979e+07
std	8.722094	8.568593	8.693710	8.683122	8.640117	1.761194e+07
min	164.889999	166.399994	164.080002	165.000000	164.909760	2.404830e+07
25%	173.320007	174.990005	172.050003	173.570007	173.229996	4.679290e+07
50%	180.089996	181.929993	178.330002	180.570007	179.859726	5.366560e+07
75%	189.259995	189.990005	187.610001	189.300003	188.850006	6.458890e+07
max	198.020004	199.619995	197.000000	198.110001	197.857529	1.366826e+08

CLOSING PRICE



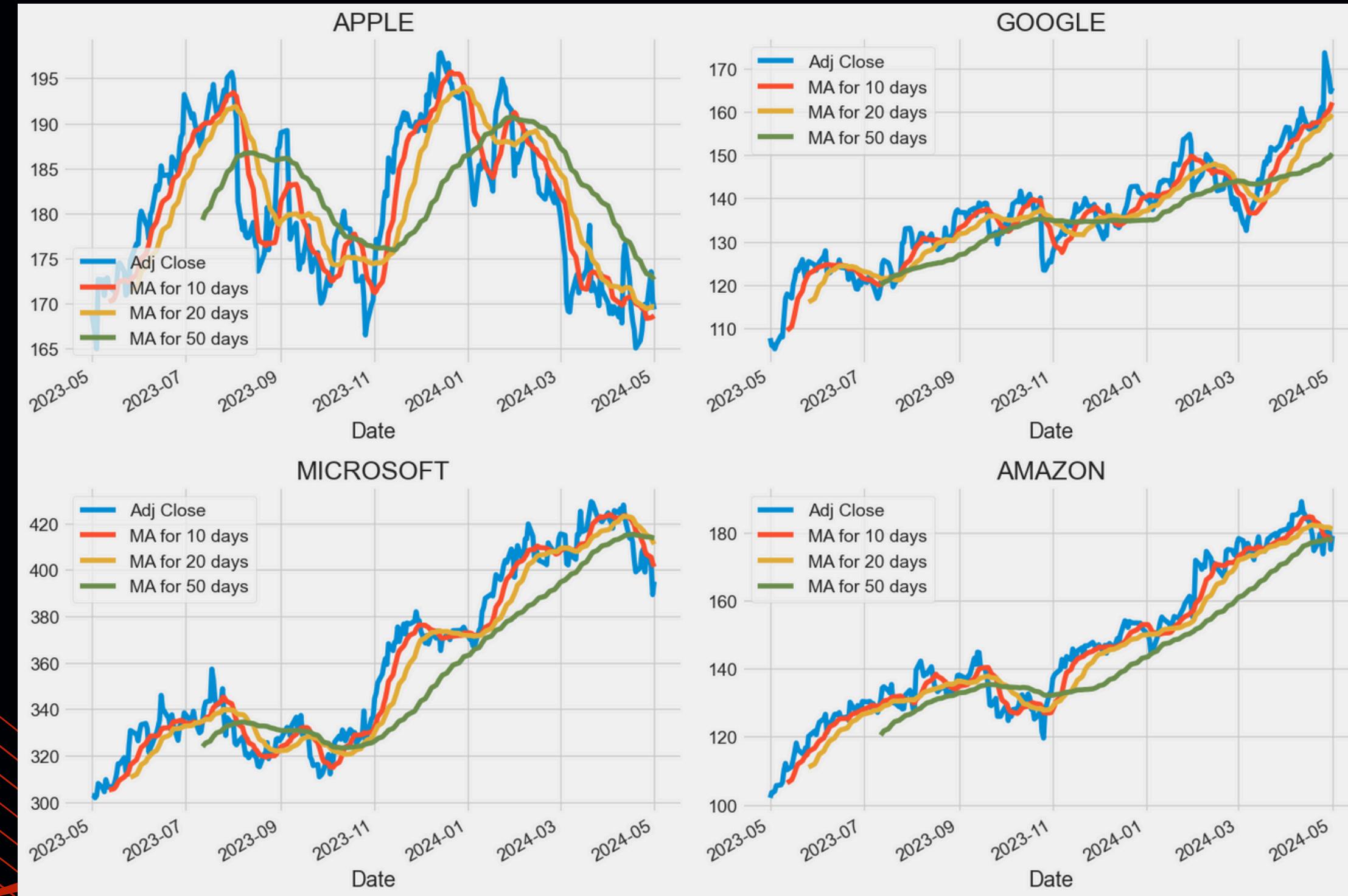
The closing price represents the last traded price of a stock during a regular trading day. It serves as a standard benchmark for investors to monitor the stock's performance over time.

SALES VOLUME



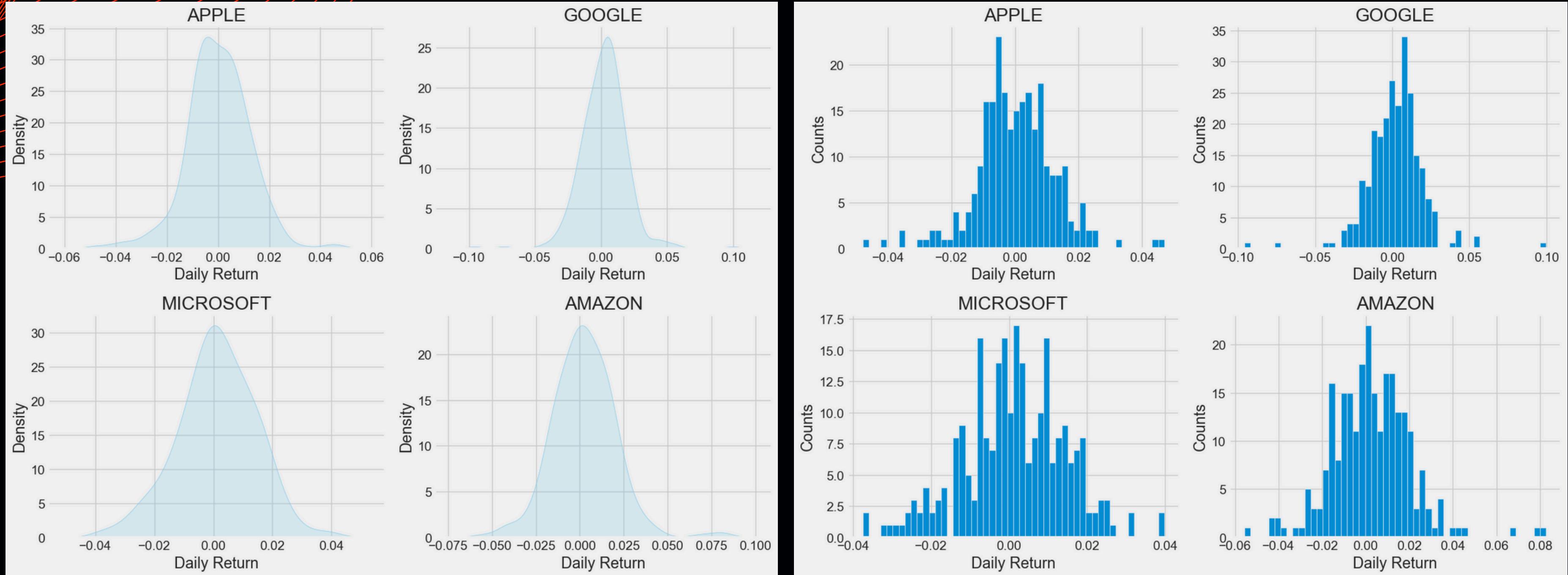
Volume refers to the quantity of an asset or security traded within a specific period, typically a day. It serves as a crucial metric for technical traders, providing insights into market activity and trends.

MOVING AVERAGE OF THE VARIOUS STOCKS



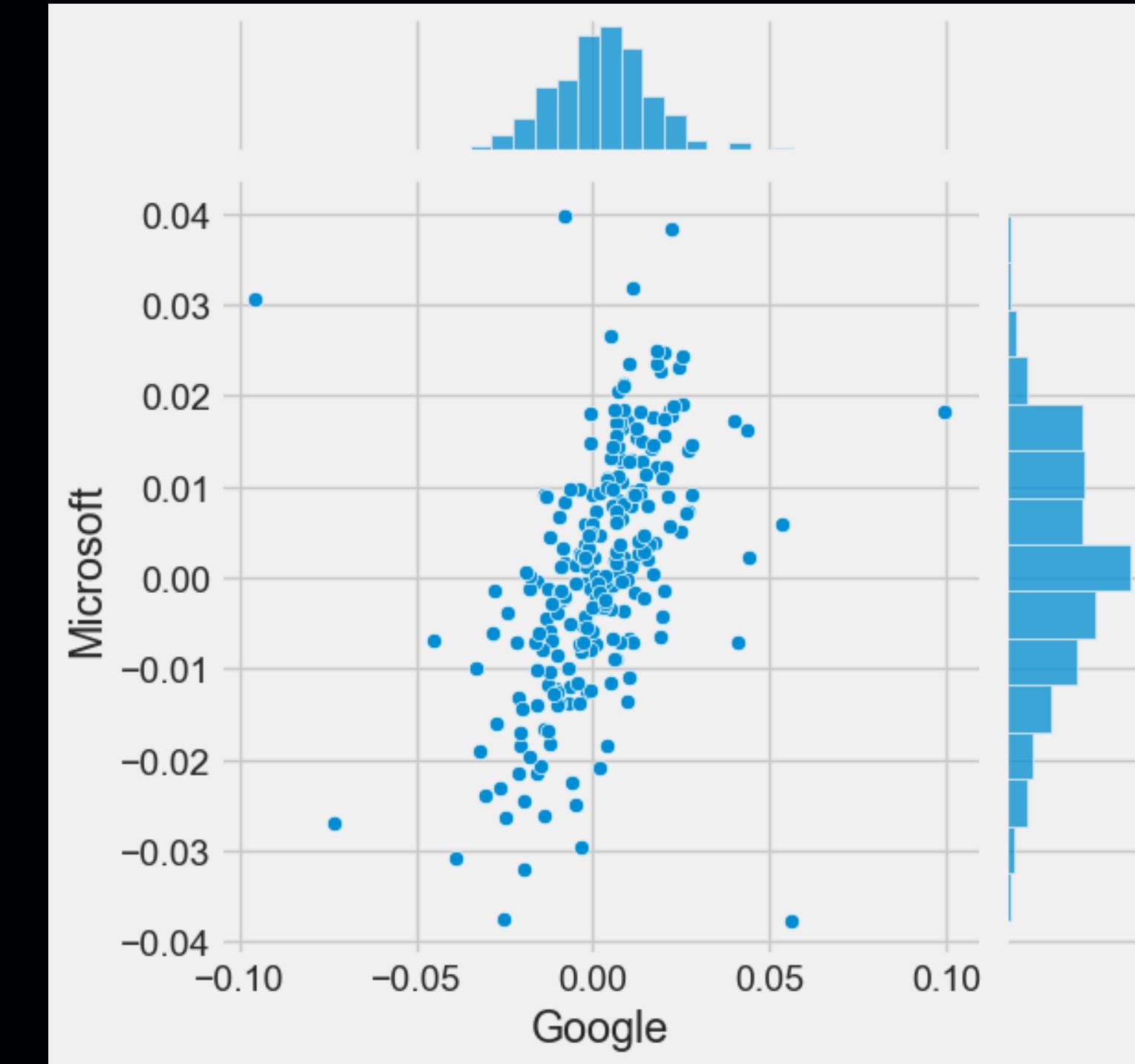
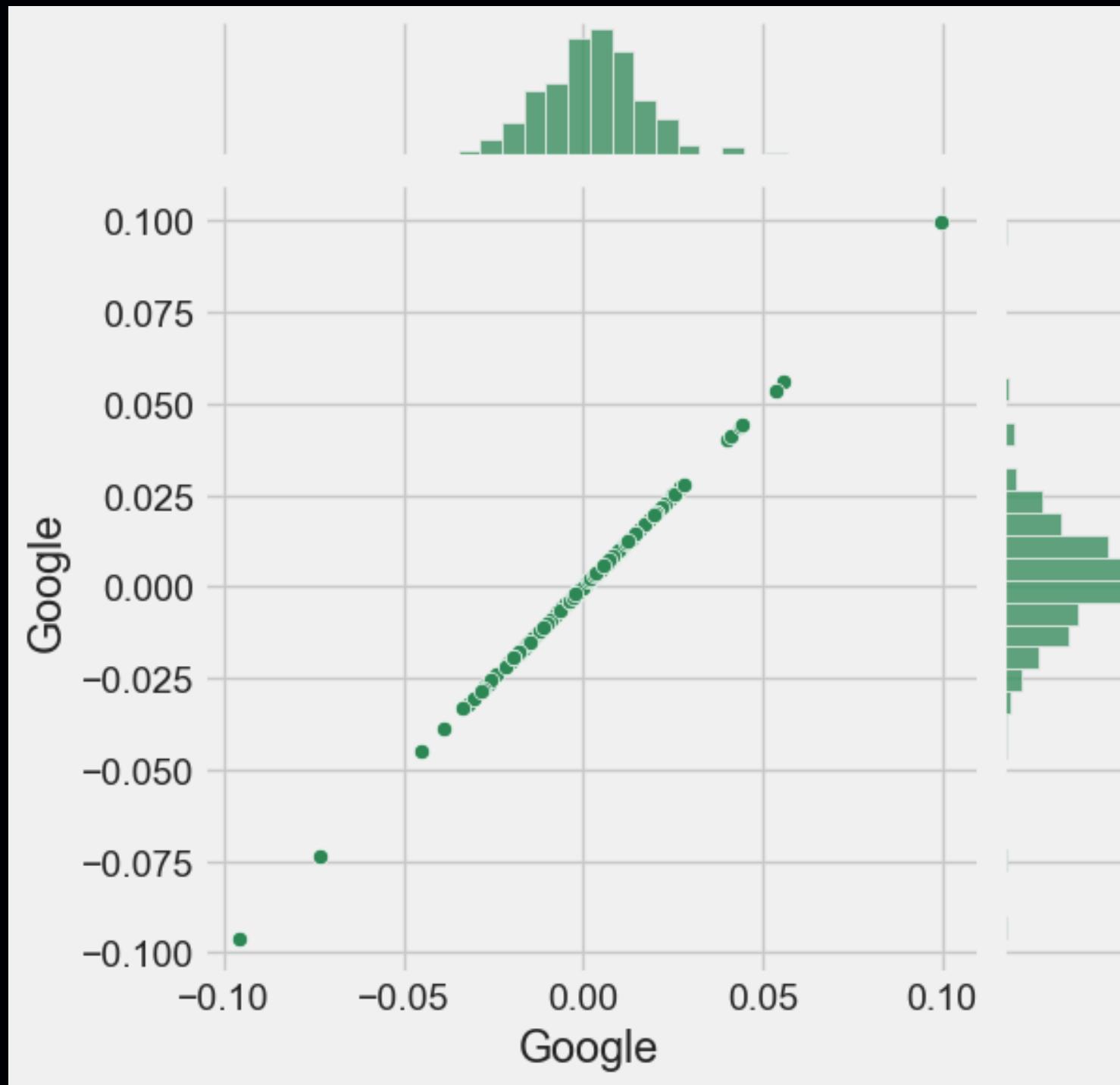
The moving average (MA) is a straightforward technical analysis tool used to smooth out price data by calculating a continually updated average price. This average is computed over a specified time period, such as 10 days, 20 minutes, 30 weeks, or any duration chosen by the trader.

DAILY AVERAGE STOCK RETURN

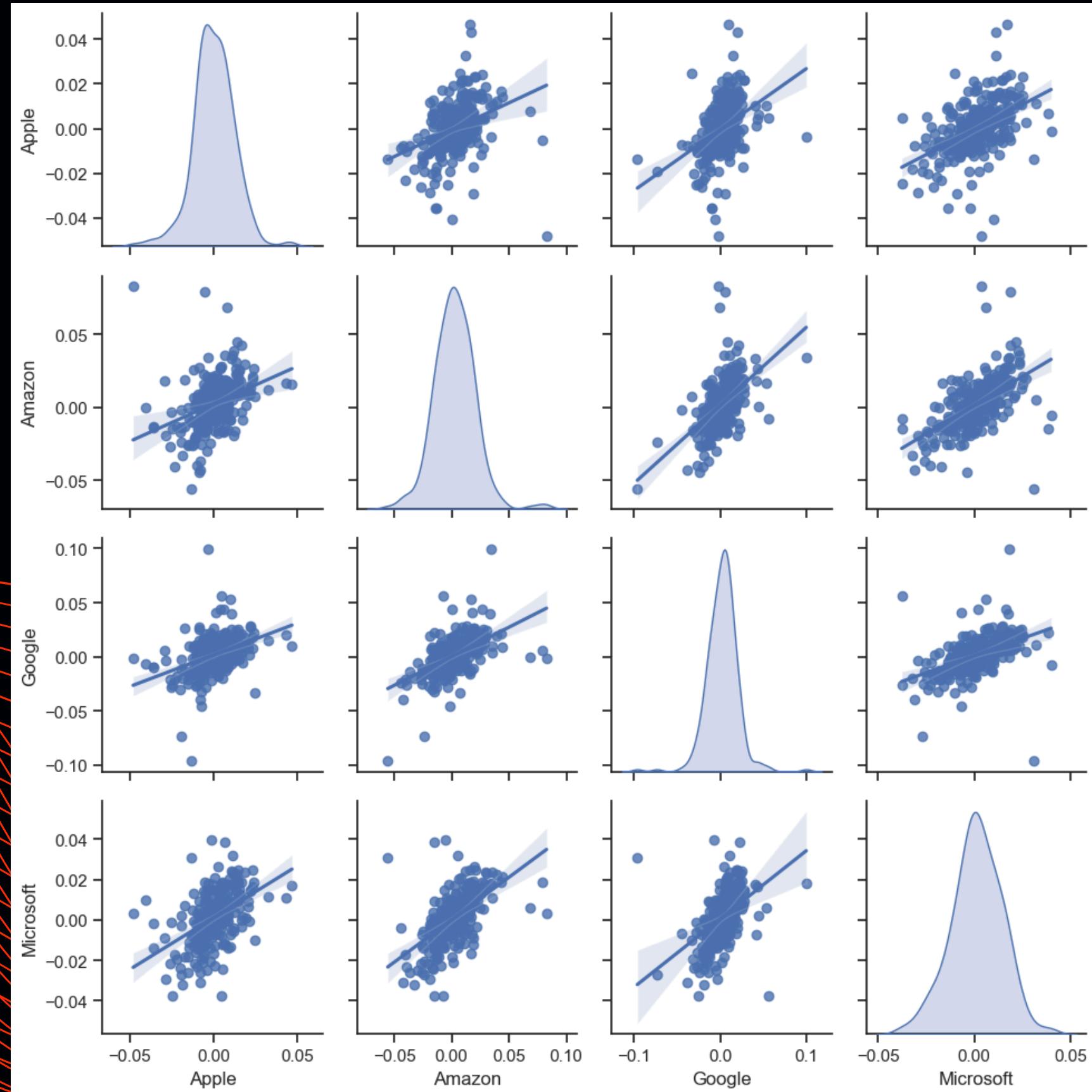


The daily return of a stock on average can be calculated by finding the mean of its daily returns. This is typically done by taking the difference between consecutive days' closing prices, dividing by the previous day's closing price, and then averaging these values.

CORRELATION BETWEEN DIFFERENT STOCKS CLOSING PRICES

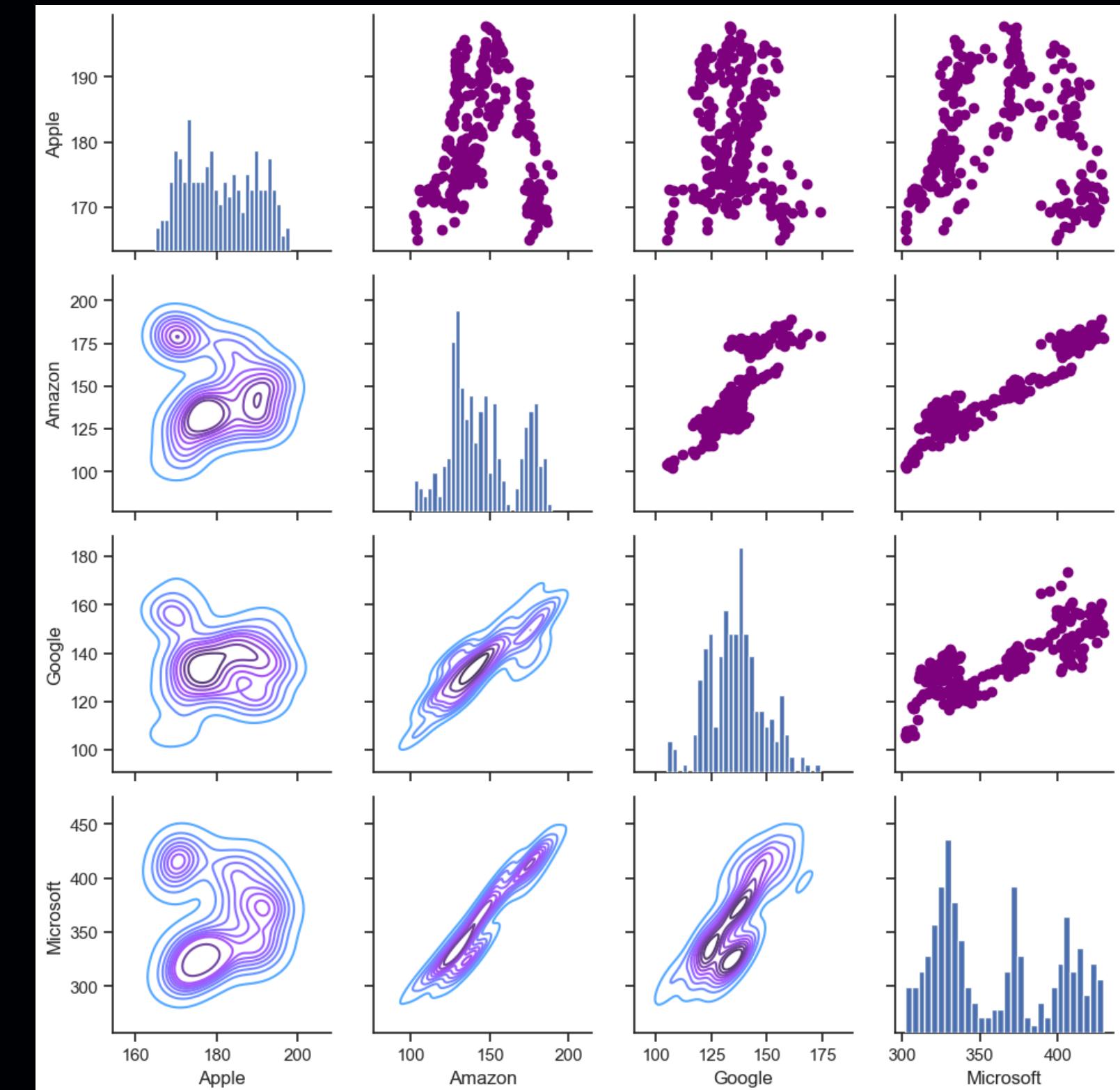
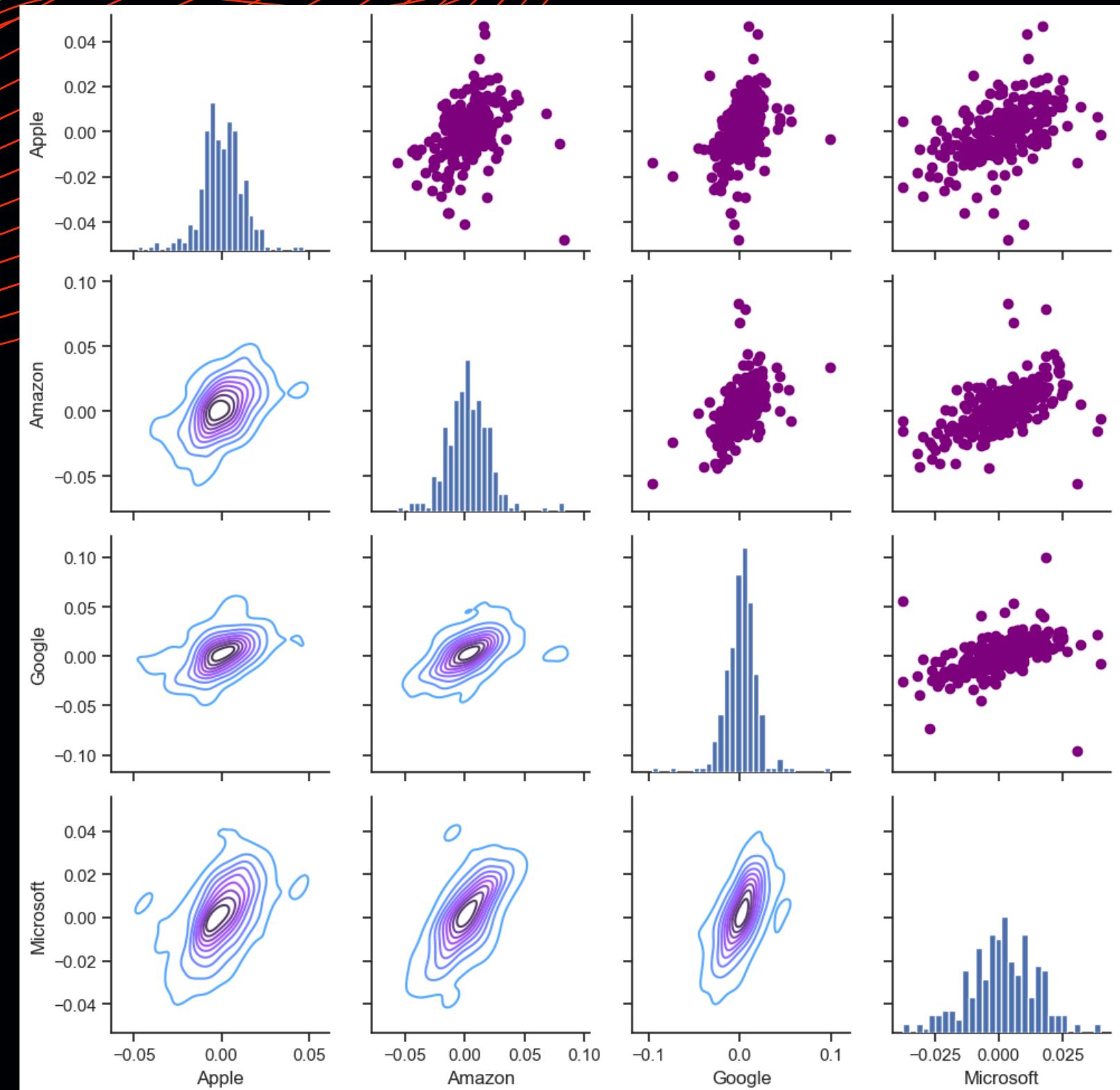


Correlation is a statistical measure indicating how closely two variables move together, with values ranging from -1.0 to +1.0. It quantifies the strength and direction of their linear relationship, but doesn't imply causation or indicate the influence of a third factor.

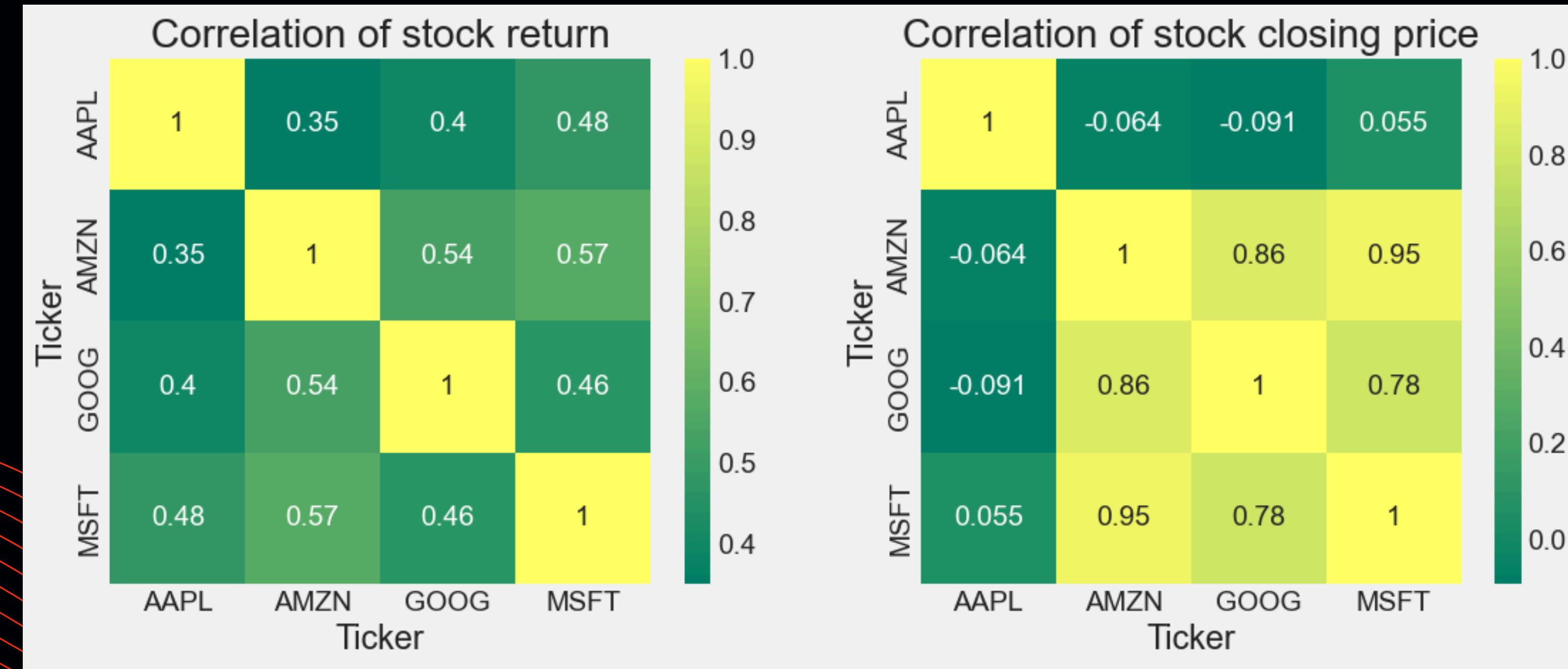


Here, we observe the relationships in daily returns among all stocks. A notable correlation appears between Google and Amazon daily returns, prompting further investigation into this specific comparison.

While `sns.pairplot()` offers simplicity, `sns.PairGrid()` provides complete control over the figure, including the types of plots in the diagonal, upper triangle, and lower triangle. Below is an example demonstrating the versatility of seaborn to achieve this outcome.

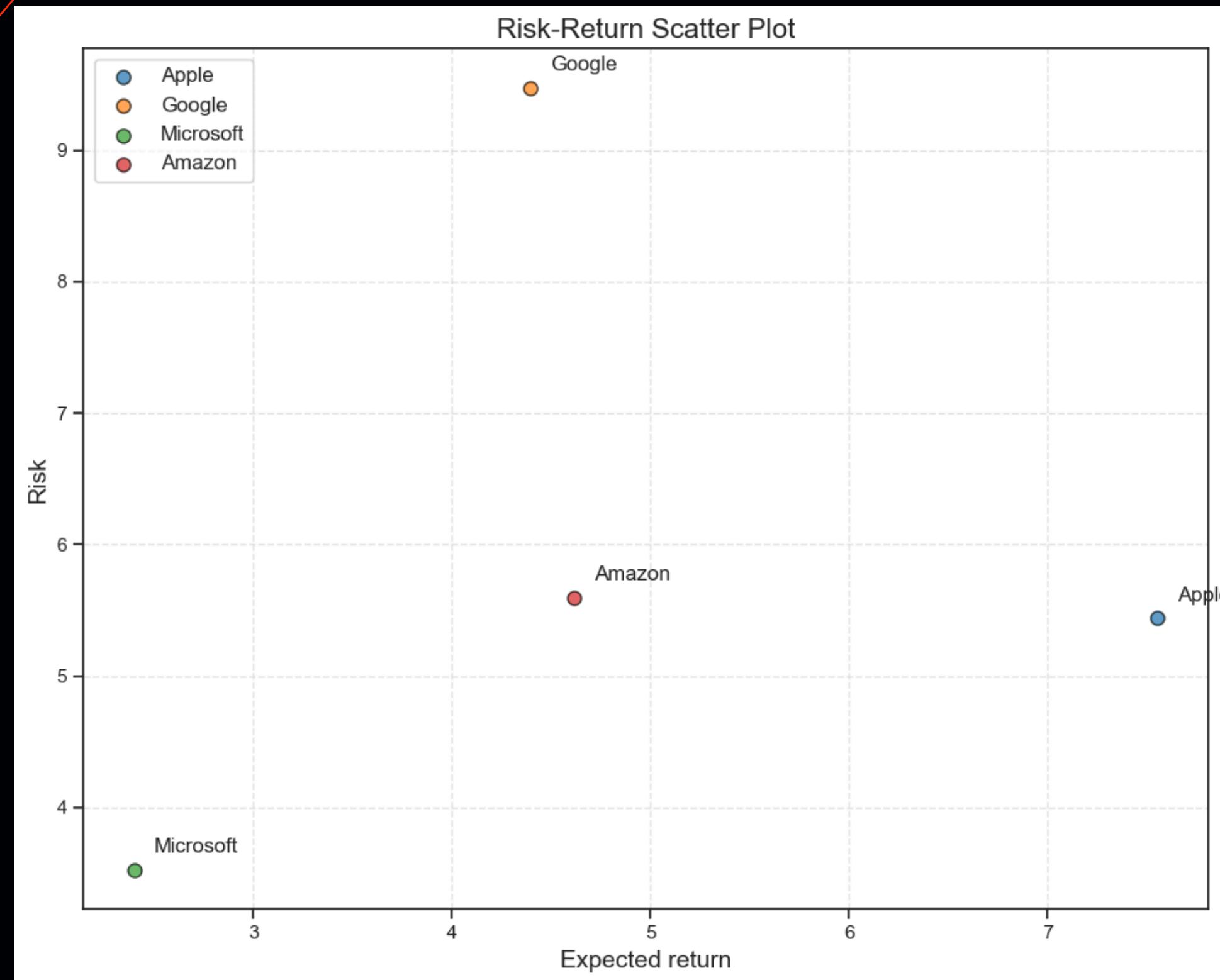


Lastly, we can generate a correlation plot to obtain numerical values for the correlation between the stocks' daily return values. Upon comparing the closing prices, an intriguing relationship emerges between Microsoft and Apple.



As we anticipated from our PairPlot analysis, both numerically and visually, Microsoft and Amazon exhibit the strongest correlation in daily stock returns. Additionally, it's intriguing to note that all technology companies display positive correlations.

RISK BY INVESTING IN A PARTICULAR STOCK

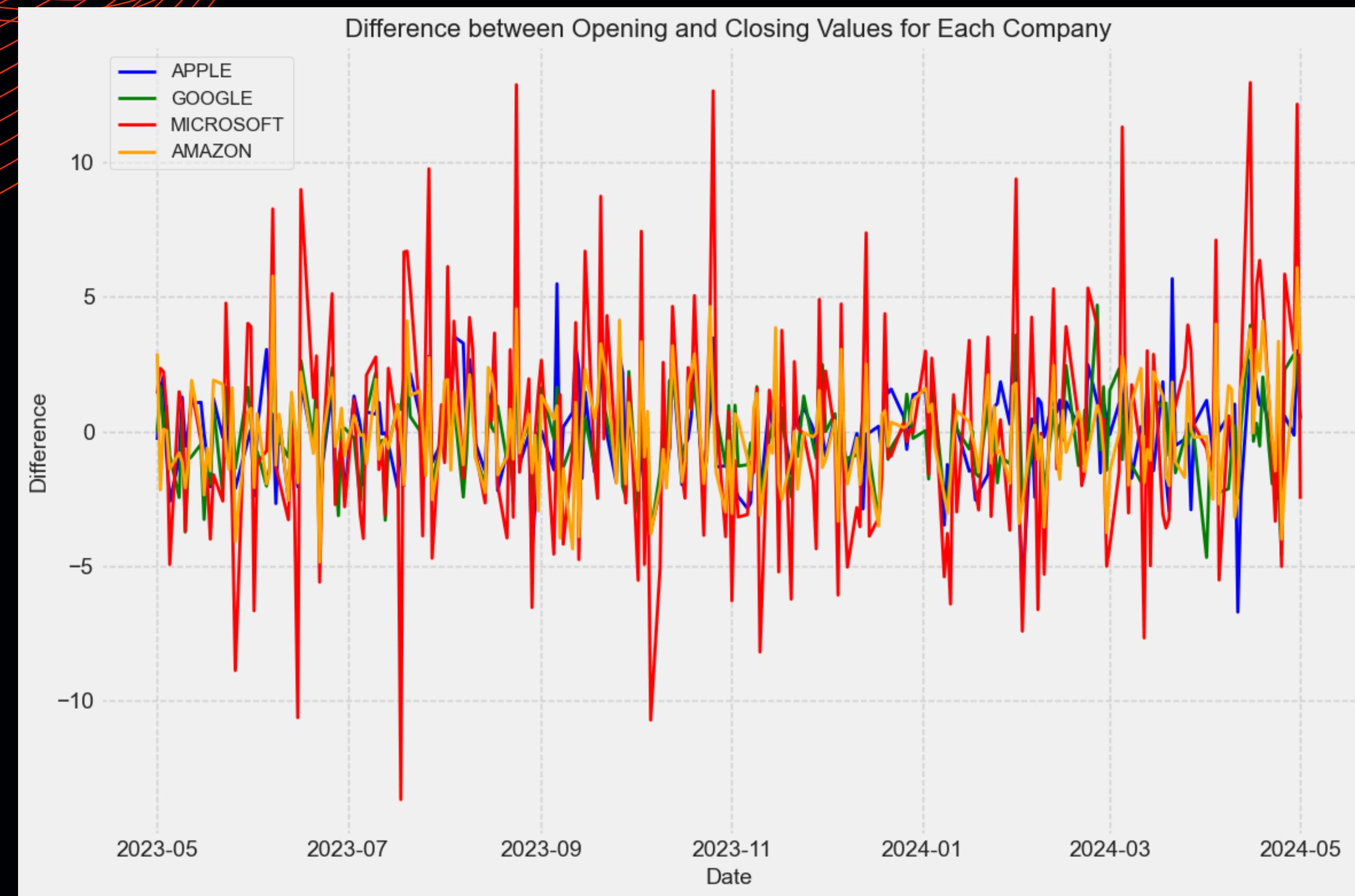


Daily percentage returns are typically calculated by comparing the expected return with the standard deviation of the daily returns. This comparison provides insight into the volatility or risk associated with the stock's daily performance.

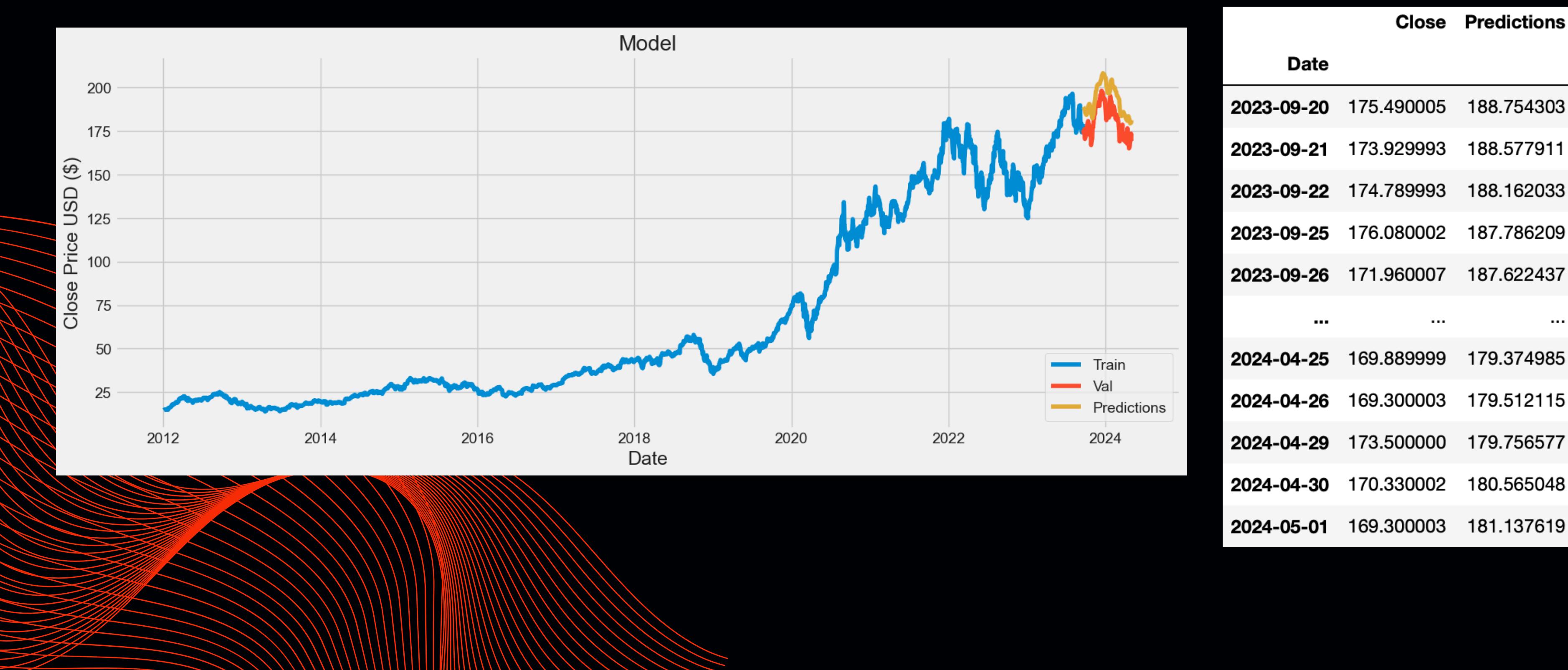
TRAINING THE CLOSING PRICE STOCK PRICE OF APPLE INC

	Open	High	Low	Close	Adj Close	Volume
Date						
2012-01-03	14.621429	14.732143	14.607143	14.686786	12.433824	302220800
2012-01-04	14.642857	14.810000	14.617143	14.765714	12.500648	260022000
2012-01-05	14.819643	14.948214	14.738214	14.929643	12.639426	271269600
2012-01-06	14.991786	15.098214	14.972143	15.085714	12.771558	318292800
2012-01-09	15.196429	15.276786	15.048214	15.061786	12.751300	394024400
...
2024-04-25	169.529999	170.610001	168.149994	169.889999	169.889999	50558300
2024-04-26	169.880005	171.339996	169.179993	169.300003	169.300003	44838400
2024-04-29	173.369995	176.029999	173.100006	173.500000	173.500000	68169400
2024-04-30	173.330002	174.990005	170.000000	170.330002	170.330002	65934800
2024-05-01	169.800003	172.690002	169.110001	169.300003	169.300003	49377085





PREDICTING THE CLOSING PRICE STOCK PRICE OF APPLE INC





Conclusion

- Comprehensive analysis conducted on stock market data of tech giants: **Apple, Google, Microsoft, and Amazon.**
- Examination of **one-year** historical data to uncover **trends, correlations, and stock price volatility.**
- Advanced data visualization techniques used to present insights clearly and effectively.
- Statistical analysis employed to identify significant patterns and relationships.
- Findings offer valuable insights for investors navigating the dynamic technology sector.
- Actionable recommendations provided to assist investors in making informed decisions.

THANK YOU