Project Report: EDA On Stock Market Dataset

1. Project Overview

The objective of this project is to analyze and compare the performance of four major technology companies—Apple (AAPL), Microsoft (MSFT), Netflix (NFLX), and Google (GOOG)—in the stock market using historical stock price data over the past three months.

The specific goals are to:

Identify trends and patterns in stock price movements.

Calculate moving averages and volatility.

Conduct correlation analysis to examine relationships between different stock prices and attributes.

2. Data Collection and Preparation

The dataset used is a CSV file containing daily stock market data for the four specified tickers. Each row represents a trading day for a specific ticker and includes columns like Date, Open, High, Low, Close, Adj Close, and Volume.

Initial Data Checks:

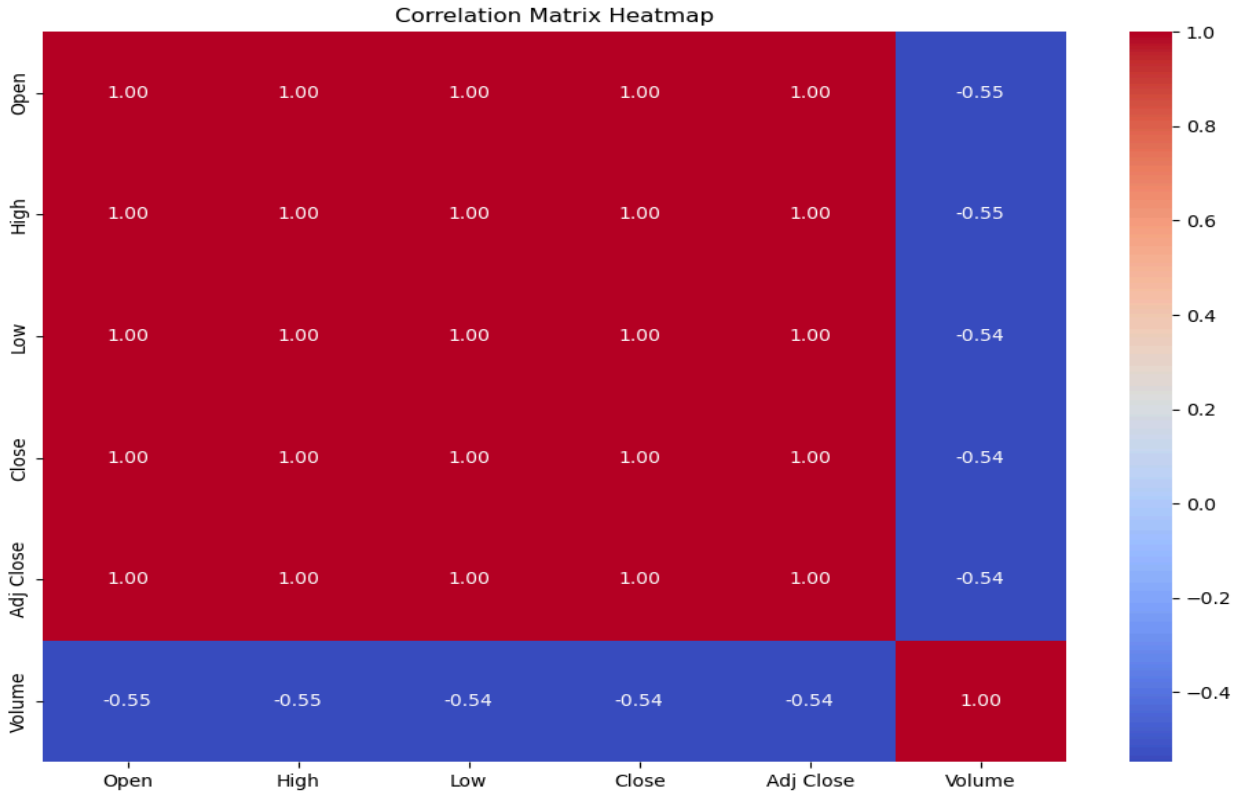
The dataset contains 248 entries and 8 columns.

There are no missing values in any of the columns.

The Date column was converted to a datetime object for time-series analysis.

3. Exploratory Data Analysis (EDA)

A. Correlation Matrix Heatmap (All Features)



This visualization is a Heatmap displaying the Pearson correlation coefficients between all numerical features in the dataset (Open, High, Low, Close, Adj Close, and Volume). The color intensity and the annotated values indicate the strength and direction of the correlation.

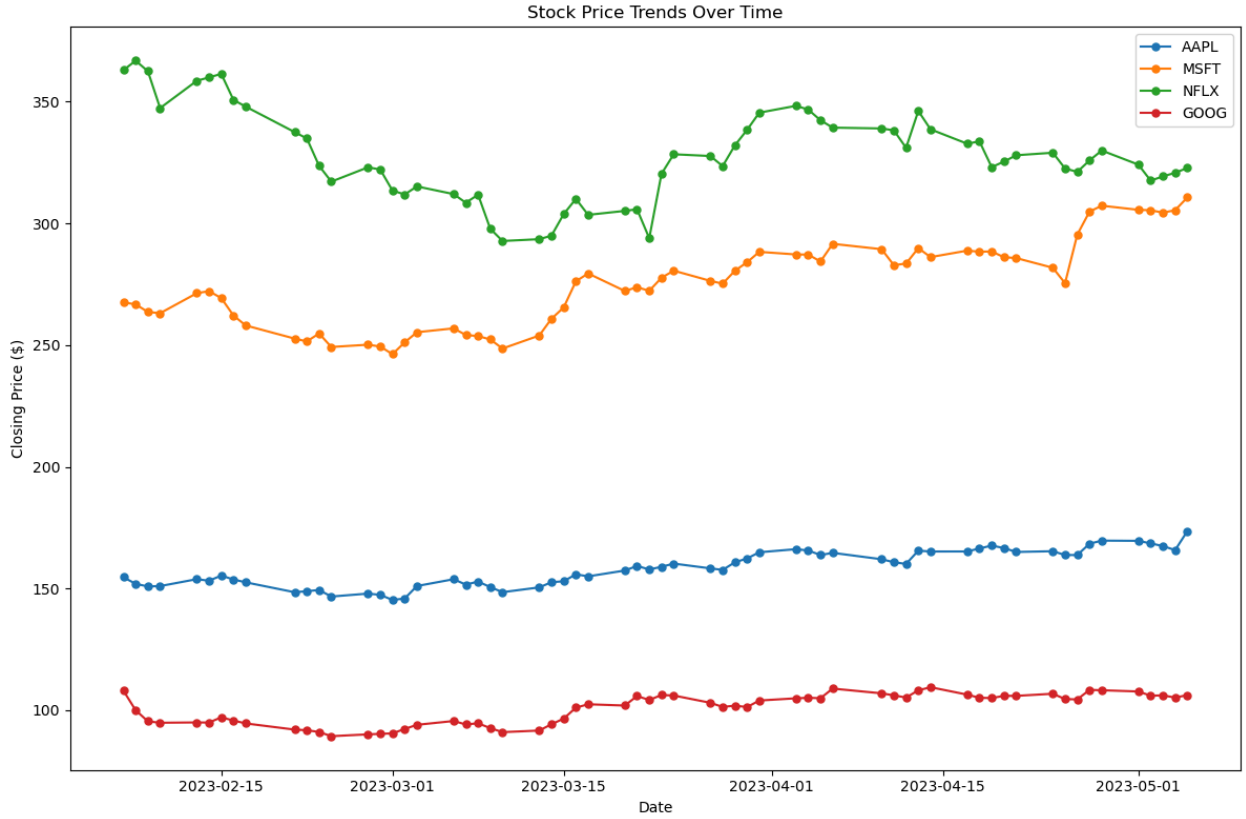
Insights:

Near-Perfect Price Correlation: All the different price metrics (Open, High, Low, Close, Adj Close) are almost perfectly correlated with each other (correlation approx 1.00). This is expected, as they all represent the price of the stock on the same day and tend to move together.

Negative Correlation with Volume: There is a moderate negative correlation (approximately -0.54) between all price metrics and Volume. This suggests that as stock prices tend to increase across the combined dataset, the average trading volume tends to decrease, or vice-versa.

Volume Independence: The Volume is perfectly correlated with itself (1.00) and shows no strong positive correlation with any of the price-related features, confirming its distinct nature as a measure of trading activity rather than a price metric.

B. Stock Price Trend Over Time



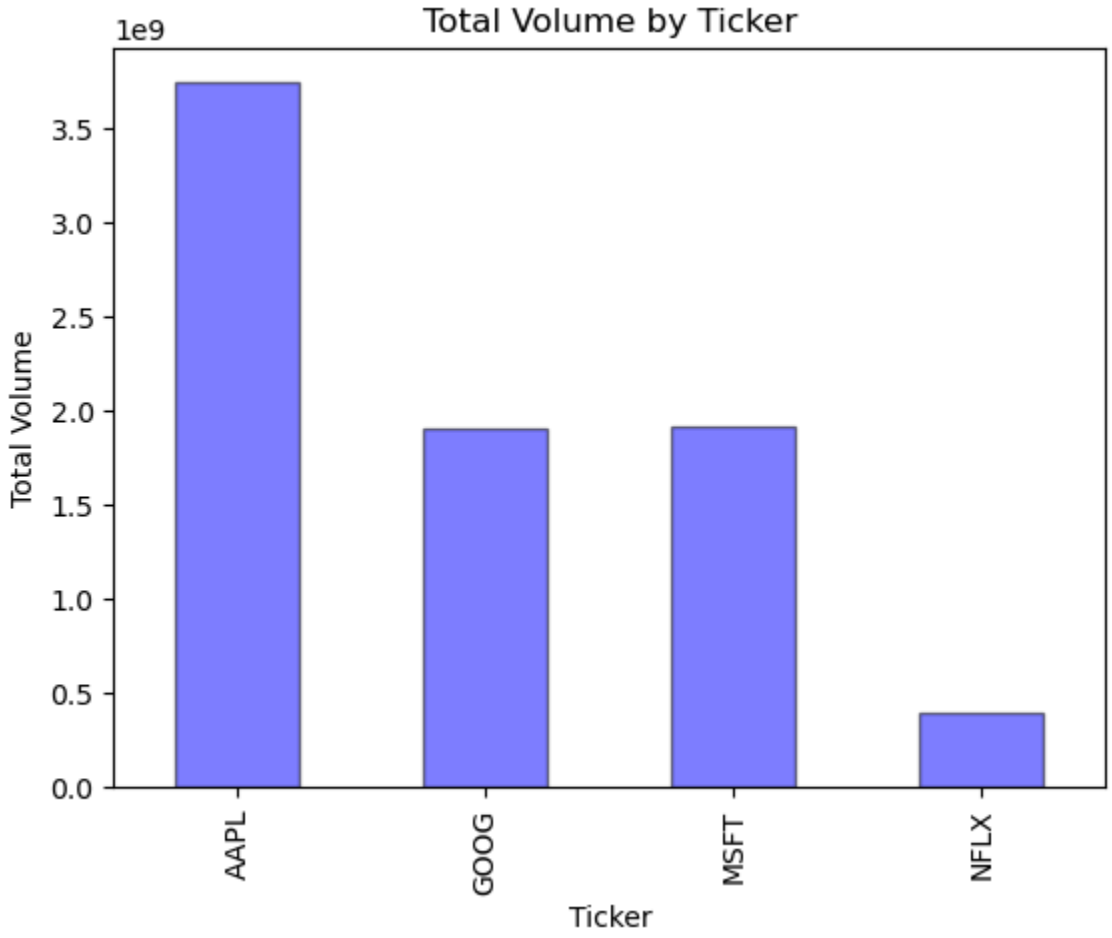
A line plot of the adjusted closing price over time is the most fundamental visualization in finance. It provides a quick, high-level view of performance. By plotting all four stocks on the same axis, we can immediately compare their price action and relative value across the period.

Insights:

The chart quickly reveals the magnitude of the prices, showing that GOOG and MSFT generally trade at higher nominal prices than AAPL and NFLX.

It highlights whether the entire market (represented by these four stocks) was in an overall bullish (upward) or bearish (downward) trend during the three months.

C. Total Volume by Ticker



This visualization is a Bar Chart that shows the cumulative sum of the trading Volume for each of the four tickers (AAPL, GOOG, MSFT, NFLX) over the entire data period. The height of each bar represents the total number of shares traded for that company.

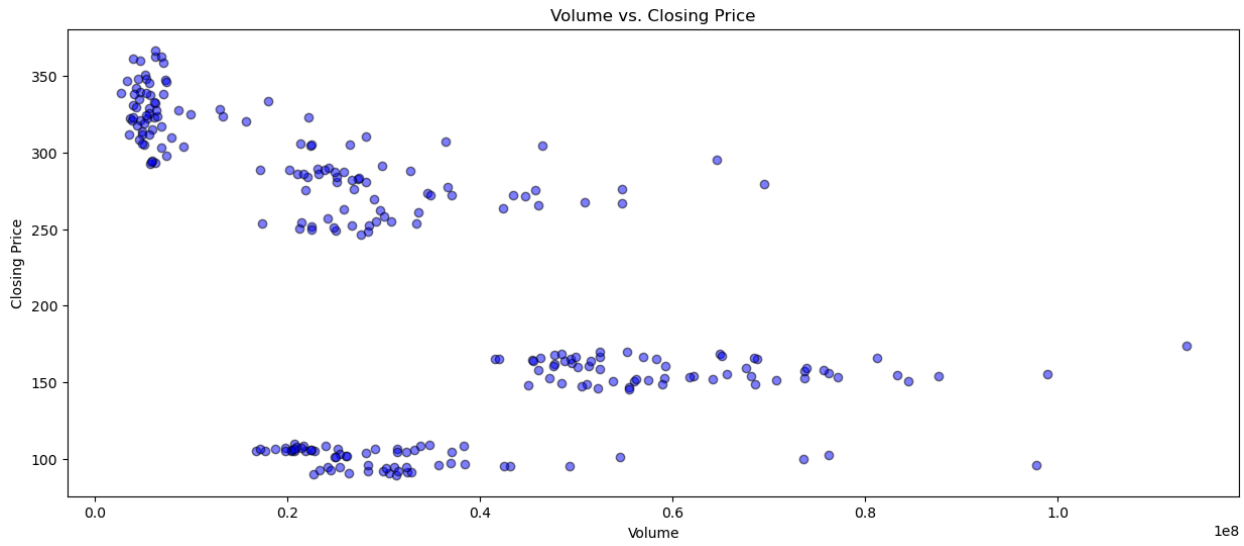
Insights:

AAPL Dominance in Trading Activity: Apple (AAPL) has significantly higher total trading volume than the other three stocks (almost twice that of GOOG and MSFT). This indicates Apple is the most actively traded stock in this specific dataset, suggesting greater market liquidity and investor interest.

GOOG and MSFT Volume Similarity: Google (GOOG) and Microsoft (MSFT) have a very similar total trading volume over the period. This might reflect their comparable market capitalization, investor base, or consistent trading activity within the tech sector.

NFLX Least Traded: Netflix (NFLX) has the lowest total trading volume by a substantial margin. This suggests it is less liquid or less frequently traded compared to the other three mega-cap tech stocks in the dataset.

D. Volume vs. Closing Price (Scatter Plot)



This visualization is a Scatter Plot showing the relationship between Volume (x-axis) and Closing Price (y-axis) for all trading days across all tickers. Each point represents one day's trade data.

Insights:

Price Clustering/Multi-Stock Effect: The data points are grouped into distinct, non-overlapping horizontal clusters. This pattern strongly suggests that the plot combines data from multiple stocks, each trading within a specific, absolute price range (e.g., one cluster is centered around $150, another around $300). Analyzing the data without separating it by ticker can be misleading.

No Clear Volume-Price Correlation: Within each major price cluster (i.e., for individual stocks), the points are scattered widely across the volume axis. This indicates that there is no strong linear correlation between the absolute daily trading volume and the absolute closing price. High trading activity days occur across all prices shown in that range, and vice versa.

Highest Volume at Lower Prices: The data cluster corresponding to the lowest closing prices (around $100 to $170) exhibits the highest absolute trading volumes, reaching up to 1.0 times

10^8. In contrast, the cluster with the highest closing prices (around $300 to $350) generally shows lower volumes. This suggests the stock with the lower price point is the most actively traded (highest liquidity).

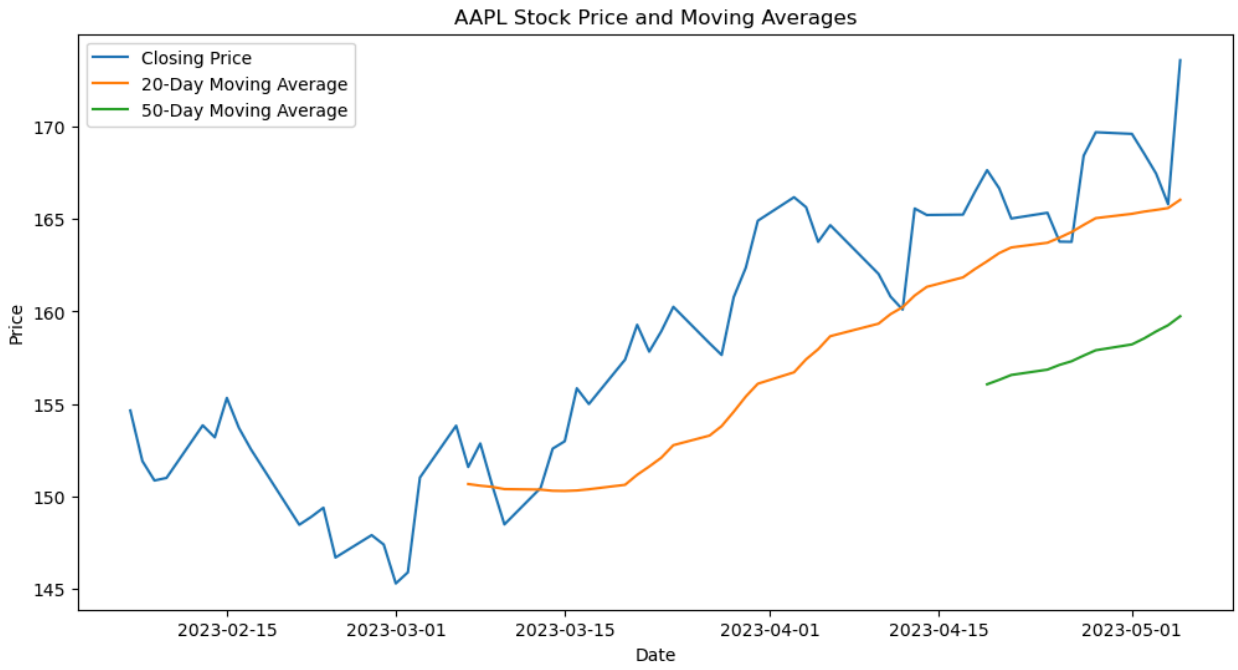
E. Price vs. Moving Averages

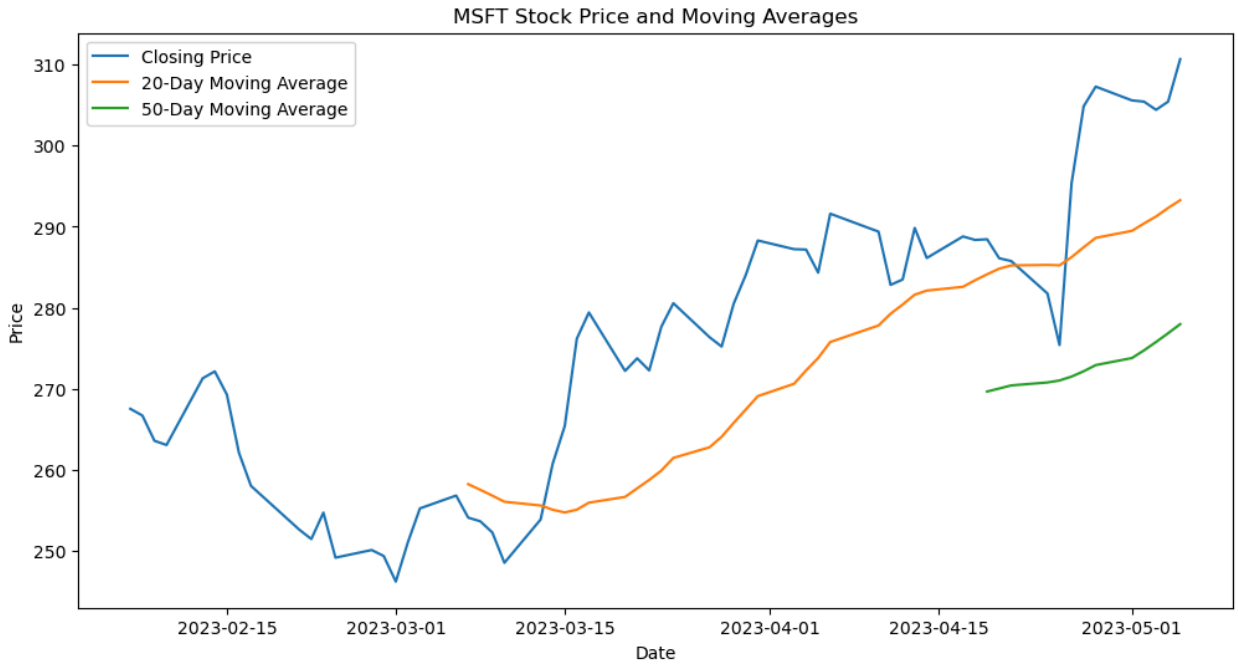
Moving Averages (MA) are essential technical analysis tools used to smooth out short-term fluctuations and define the current trend direction.

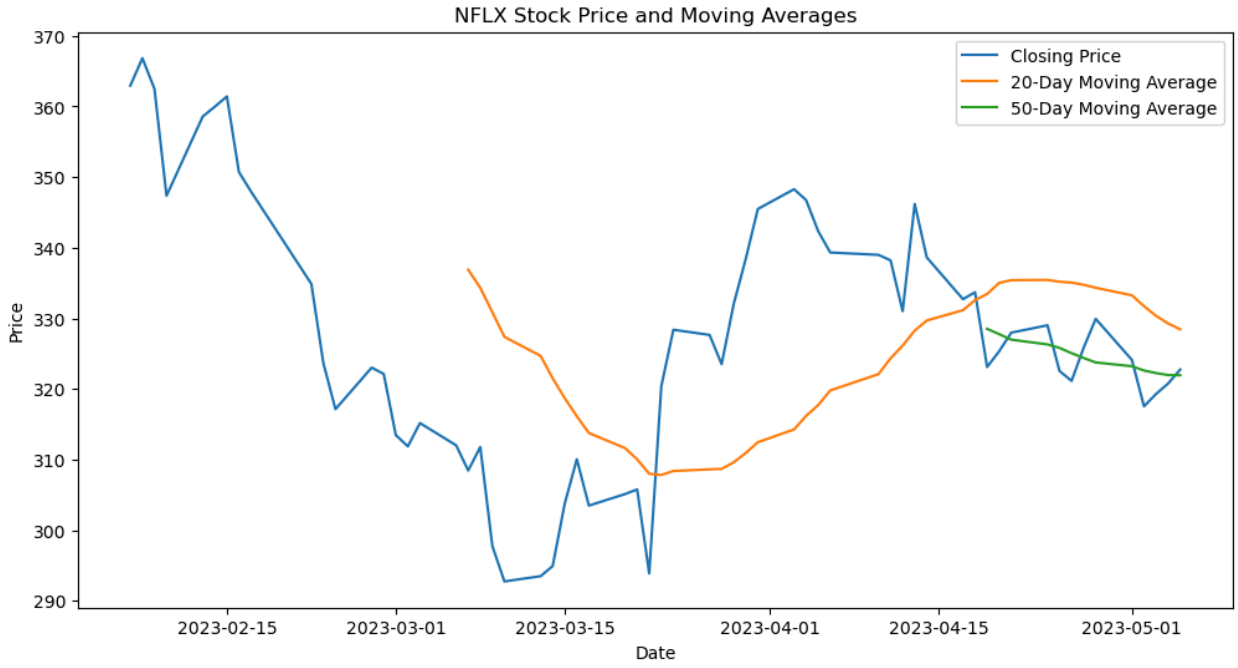
The 20-Day MA represents the short-term trend.

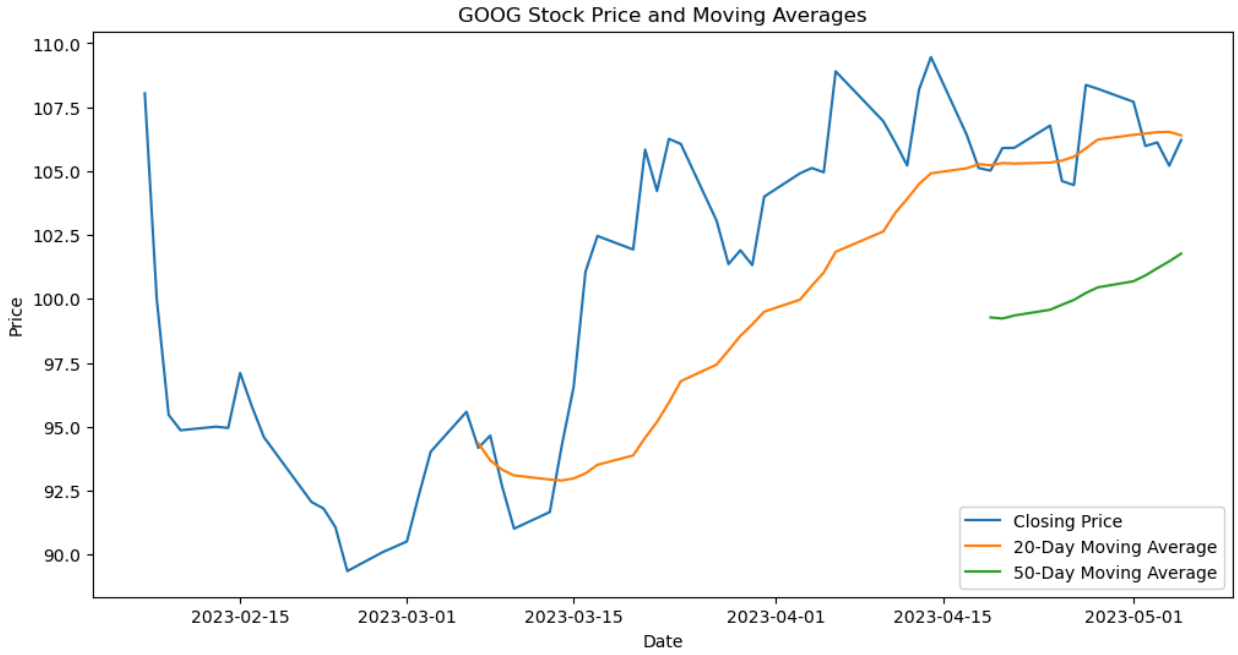
The 50-Day MA represents the mid-term trend.

The visualization plots the raw price alongside these MAs. The key signal is a **crossover**. When the 20-Day MA crosses *above* the 50-Day MA, it is often a *bullish signal* (a Golden Cross). When it crosses *below*, it is a *bearish signal* (a Death Cross).




MSFT Stock Price and Moving Averages
Closing Price
20-Day Moving Average
50-Day Moving Average
Price
Date
310
300
290
280
270
260
250
2023-02-15
2023-03-01
2023-03-15
2023-04-01
2023-04-15
2023-05-01


NFLX Stock Price and Moving Averages
Closing Price
20-Day Moving Average
50-Day Moving Average
Price
Date
370
360
350
340
330
320
310
300
290
2023-02-15
2023-03-01
2023-03-15
2023-04-01
2023-04-15
2023-05-01



Momentum Confirmation: By observing the plots for each company (AAPL, MSFT, NFLX, GOOG), we can determine if the recent price action (short-term) confirms or diverges from the broader trend (mid-term). A stock's price consistently staying above both MAs indicates a strong, well-supported uptrend.

Trend Identification: The slope of the 50-Day MA is crucial. If the 50-Day MA is rising, the mid-term trend is bullish, suggesting a better entry point during price dips that bounce off the MA line (acting as dynamic support).

Support and Resistance: The Moving Averages often act as dynamic support (on the way up) or resistance (on the way down). Observing how many times the stock price tested and successfully bounced off the 50-Day MA can gauge the underlying strength of the trend for each individual stock.

4. Volatility Analysis and Risk Profile

Ticker	Volatility (Standard Deviation of Daily Returns)
AAPL	0.014152
GOOG	0.020710
MSFT	0.017881
NFLX	0.022481

The standard deviation is a measure of risk or volatility. Among these four companies, NFLX (Netflix) exhibits the highest daily volatility (0.022481), while AAPL (Apple) exhibits the lowest (0.014152).

Interpretation of Price Swings: The high volatility of NFLX (0.0225) means that on a typical day, its price return deviates by approximately 2.25% from its average return. This indicates large, frequent price swings compared to Apple's 1.42% daily deviation, which suggests a smoother, more predictable price movement.

Portfolio Management: This differential risk profile is crucial for diversification. An investor seeking stable growth might heavily favor AAPL and MSFT. Conversely, an investor with a higher risk tolerance might allocate more capital to NFLX and GOOG, accepting greater price uncertainty for the potential of higher, short-term returns. The risk profiles clearly show where the largest price fluctuations occurred during the analysis period.

5. Key Insights and Conclusion

Based on the technical analysis over the three-month period, the following key insights were drawn:

Dominant Bullish Trend: Both Microsoft (MSFT) and Apple (AAPL) demonstrated strong, sustained bullish trends, confirmed by rising 50-Day Moving Averages. MSFT showed superior momentum, using its 20-Day MA as dynamic support, indicating very strong, consistent buyer conviction.

Highest Risk Profile: Netflix (NFLX) exhibited the highest daily volatility (0.0225), signifying the largest and most frequent price swings. This points to NFLX as a higher-risk, potentially higher-reward asset compared to the stable growth profile of Apple (AAPL, 0.0142 volatility).

Consolidation: Netflix's frequent price crossings of its MAs suggest it was in a period of consolidation or sideways trading, lacking a clear, defined mid-term trend.

This analysis successfully accomplished all primary objectives, providing a clear, comparative view of the performance, momentum, risk, and interconnectedness of the four technology stocks. The use of Moving Averages effectively assessed momentum, while the standard deviation quantified the differential risk profiles, offering actionable data for portfolio construction and risk management decisions.