NSBM Green University

Faculty of Computing

Department of Data Science

Group E

Final Project Report

Lecturer name: Miss Kavishka Rajapaksha

Module name: Advanced Statistics for Data Science

Module Code: DS302.3

Student ID	Student Name
26044	RKNT Rajapaksha
25004	WSR Withanage
25491	VASR Hirushan
25076	MP Hewavitharana

Table of Contents

Problem Background

The stock market, renowned for its dynamic and often unpredictable nature, poses a significant challenge for investors and stakeholders aiming to make informed decisions. Our project addresses the crucial task of predicting future stock price fluctuations by employing advanced machine learning techniques. We focus on exploring comprehensive datasets extracted from Yahoo Finance, which hosts a wealth of historical stock market data. By incorporating key features such as date, open, high, low, close, adjusted close, and volume, our objective is to develop robust predictive models. These models aim to provide valuable insights to market participants, enhancing decision-making in the ever-changing financial landscape.

Accurate stock price predictions are imperative not only for individual investors but also for stakeholders, financial analysts, and policymakers. Investors rely on precise forecasts to optimize strategies, capitalize on opportunities, and manage risks effectively. We selected datasets from Yahoo Finance due to their extensive coverage of diverse companies and historical richness, providing a solid foundation for comprehensive analysis. Our project seeks to uncover hidden patterns and trends within the dataset, contributing to a deeper understanding of market dynamics. Ultimately, our goal is to empower market participants with reliable tools, enabling them to navigate the complexities of the stock market with confidence.

Therefore, this project revolves around utilizing advanced machine learning techniques and extracting comprehensive datasets from Yahoo Finance to predict stock price fluctuations. The primary focus is on forecasting the stock prices of three major companies: Apple, Samsung, and Google. We chose these companies due to their high popularity in the world and the tech field. To enhance user experience, we are trying to create a Python interface featuring a selection box for users to choose a specific company of interest.

The interface will display the analyzed and predicted stock prices, offering valuable tools for individual investors, stakeholders, financial analysts, and policymakers to make well-informed decisions in the ever-evolving financial landscape. This approach aims to depend on the understanding of market dynamics and empower users to navigate the stock market with confidence, optimizing strategies and managing risks effectively.

Problem Statement

Our project addresses the formidable task of forecasting future stock price fluctuations for a diverse array of companies listed on the stock market. Leveraging the extensive and comprehensive Yahoo Finance dataset, which encompasses pivotal features such as date, open, high, low, close, adjusted close, and volume, our goal is to develop sophisticated predictive models. The significance of accurate predictions is underscored by the potential impact on

investors and stakeholders, who rely on such forecasts to navigate the complexities of the financial markets effectively.

The challenge at hand involves delving into the nuanced dynamics of stock market behavior, where myriad factors, both internal and external, contribute to inherent volatility. By focusing on key financial indicators within the dataset, we aim to create predictive models that transcend mere numerical forecasts. These models should provide actionable insights capable of informing investment decisions and strategic planning. The multifaceted nature of stock price prediction, influenced by factors ranging from economic indicators to global events, necessitates a nuanced and robust approach to model development.

Project Objectives

Main Objective

Develop Robust Predictive Models: Construct a robust predictive model in Python tailored for the related stock prices of Apple, Samsung, and Alphabet (Google) brands by utilizing sophisticated time series analysis and machine learning techniques. This comprehensive model aims to capture nuanced patterns and fluctuations in the stock prices of these brands, providing valuable insights for investors and stakeholders.

Sub-Objectives:

Feature Impact Analysis: Explore how each feature, including opening prices, highs, lows, closes, adjusted closes, and trading volumes, affects prediction accuracy. Conduct a thorough analysis of their importance in influencing outcomes, identifying, and prioritizing the variables that exert the most significant influence on stock price predictions. This insight will contribute to the refinement of the model and enhance its predictive capabilities.

User-Friendly UI Development: Create an attractive and user-friendly interface to facilitate easy interaction with the developed predictive model. Implement visualization tools to enhance the understanding of model outcomes and trends, making it understandable to both technical and non-technical users.

Theoretical Background

Our project integrates advanced machine learning techniques and time series analysis. The project incorporates Long Short-Term Memory (LSTM) neural networks, a sophisticated type of recurrent neural network (RNN), tailored for capturing complex temporal dependencies in time series data.

Time series analysis, a pivotal component of the project's theoretical background, involves meticulous study and forecasting of time-dependent data points. This analysis, facilitated by Python tools such as scikit-learn and TensorFlow, contributes to a nuanced exploration of historical stock price movements. By embracing these theoretical foundations, the project aims to not only predict future stock prices accurately but also provide stakeholders with a deeper understanding of the temporal dynamics shaping market behavior. The integration of theoretically accurate and advanced methodologies positions the project at the forefront of predictive analytics in the financial domain.

We also employ Python-based evaluation metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) for robust evaluation of the models' performance against actual stock price data.

We can represent Mathematical Notations for the above evaluation metrics.

Mean Absolute Error (MAE): $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \quad | y_i - \hat{y}_i |$

Root Mean Squared Error (RMSE): $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \quad (y_i - \hat{y}_i)^2}$

Mean Absolute Percentage Error (MAPE): $\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \quad \frac{|y_i - \hat{y}_i|}{y_i} \times 100\%$

Here, y_i represents the actual stock price, $\hat{y}_i$ represents the predicted stock price, and n is the number of data points in the evaluation dataset.

Methodology

Problem Identification

The stock market's dynamic and unpredictable nature presents a challenge for decision-makers. Our project focuses on predicting future stock price fluctuations using advanced techniques and Yahoo Finance datasets, aiming to equip investors with valuable insights. Accurate predictions are important for optimizing strategies and managing risks.

Data Collection

For this project, we collected the dataset from Yahoo Finance, which will undergo rigorous time series analysis, a statistical technique crucial for studying and forecasting time-dependent data points collected at regular intervals. The dataset encompasses features such as date, open, high, low, close, adjusted close, and volume, which will be thoroughly examined using Python tools to unveil patterns, trends, and seasonality within the stock market data. We Used the past ten years

of data from 8th March 2014 to 10th March 2024 from Yahoo Finance related to Samsung, Apple, and Google companies, for train the model.

Data set:
https://drive.google.com/drive/folders/1SwG8TFUyUBvlfcBg_fn4V27WOeGxrSot?usp=sharing

Data Preprocessing and Exploratory Data Analysis (EDA):

Various time series analysis techniques in Python will be applied during the Exploratory Data Analysis (EDA) phase. This will provide deeper insights into the historical movements of stock prices, enabling a comprehensive understanding of temporal patterns and aiding in the formulation of appropriate predictive models.

Model Creation and Training

We employ machine learning techniques mainly focusing on LSTM models that can be adept at handling time series data. Python libraries such as scikit-learn and TensorFlow will be instrumental in model development and training. The LSTM neural networks, specifically designed for sequential data, will leverage their ability to capture temporal dependencies, enhancing the project's capability to predict future stock price fluctuations.

Final Model Evaluation

In the final phase, our predictive model will undergo rigorous evaluation using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics will provide clear insights into the accuracy of our model by measuring the differences between predicted and actual stock prices. We created a flexible Model that can be implemented for all the brands.

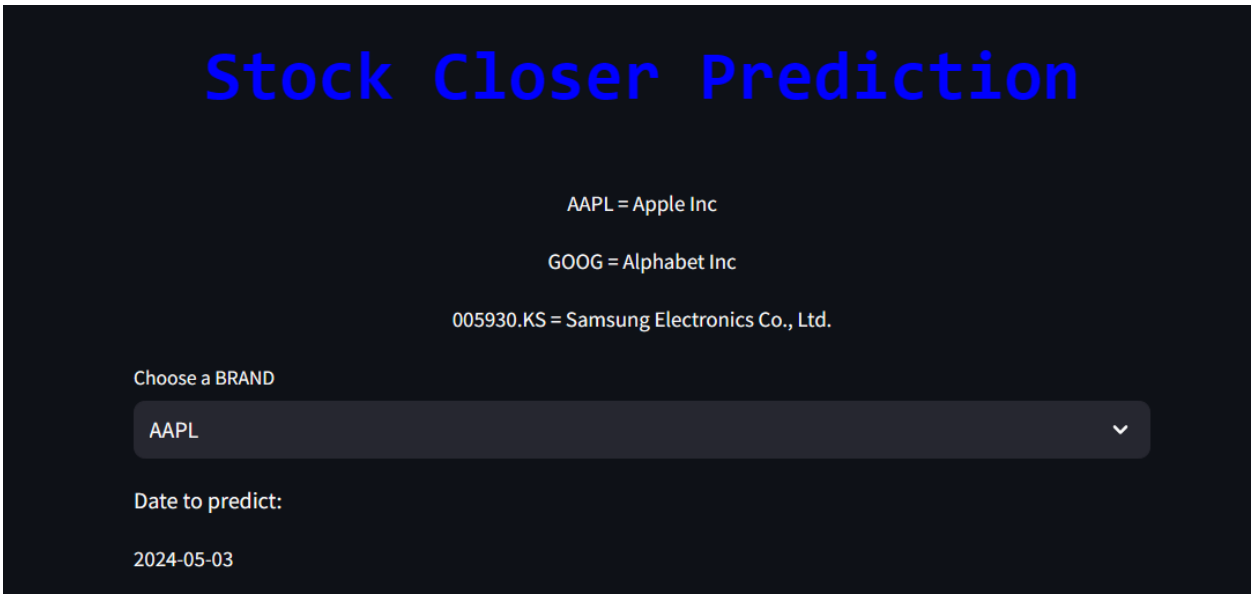
```
[26]: mse = mean_squared_error(y_test, predict)
      rmse = np.sqrt(mse)
      print(f"RMSE: {rmse:.4f}")

      RMSE: 6.3089

[24]: mape = np.mean(np.abs((y_test - predict) / y_test)) * 100
      mape

[24]: 3.222535079482562

[25]: mae = mean_absolute_error(y_test, predict)
      mae

[25]: 5.3553060755341
```

Project Status and Functionality

The project, named "Stock Closer Prediction," is designed as a stock prediction tool that allows users to choose a brand from a list (Apple, Alphabet, Samsung) and predict the stock price for a selected date. The system provides various visualizations and metrics to aid users in understanding stock price trends and model performance.

How It Works:

Brand Selection: Users select a brand from a list of Apple (AAPL), Alphabet (GOOG), and Samsung Electronics (005930.KS) for which they want to predict the stock price.



Data Retrieval: The system retrieves historical stock price data from Yahoo Finance for the selected brand.

Data from 2014 - 2024

Date	Open	High	Low	Close	Adj Close	Volume
2014-03-19 00:00:00	19.0093	19.1514	18.8929	18.9736	16.7122	224,756,000
2014-03-20 00:00:00	18.9246	19.0239	18.8339	18.8821	16.6317	208,398,400
2014-03-21 00:00:00	18.9975	19.0625	18.7975	19.0311	16.7629	374,046,400
2014-03-24 00:00:00	19.2293	19.3036	19.1093	19.2568	16.9617	355,700,800
2014-03-25 00:00:00	19.3393	19.4911	19.2711	19.4639	17.1441	282,293,200
2014-03-26 00:00:00	19.5186	19.6071	19.245	19.2779	16.9802	299,768,000
2014-03-27 00:00:00	19.2864	19.3393	19.1114	19.195	16.9073	222,031,600
2014-03-28 00:00:00	19.2257	19.2479	19.0804	19.1736	16.8884	200,564,000
2014-03-31 00:00:00	19.2582	19.3146	19.1404	19.1693	16.8846	168,669,200
2014-04-01 00:00:00	19.2057	19.3525	19.1704	19.3446	17.0391	200,760,000

The first column shows the date, and the next one shows the starting stock price of the chosen brand. High column shows the highest stock price for the brand and Low column shows the lowest stock price in the day. The close price means the last price that the buyers and sellers accepted. adj close column represents the closing price of a stock that has been adjusted to include any distributions and corporate actions that occurred at any time prior to the next day's opening.

Data Visualization:

Closing Price Overtime

This graph is titled "Closing Price over Time" and shows the price of closing value how likely to rise, over ten years. The x-axis of the graph is labeled "Time" and shows years ranging from 2014 to 2024. The y-axis is labeled "Price".

Closing Price Vs Time with 100MA and 200MA

The Moving Average (MA) is a statistical method used to smooth out fluctuations in data over time, revealing underlying trends or patterns. It calculates the average value of a series of data points by continuously updating the average as new data becomes available.

It shows the closing price of a stock over time along with two moving averages (MA) plotted on the same graph.

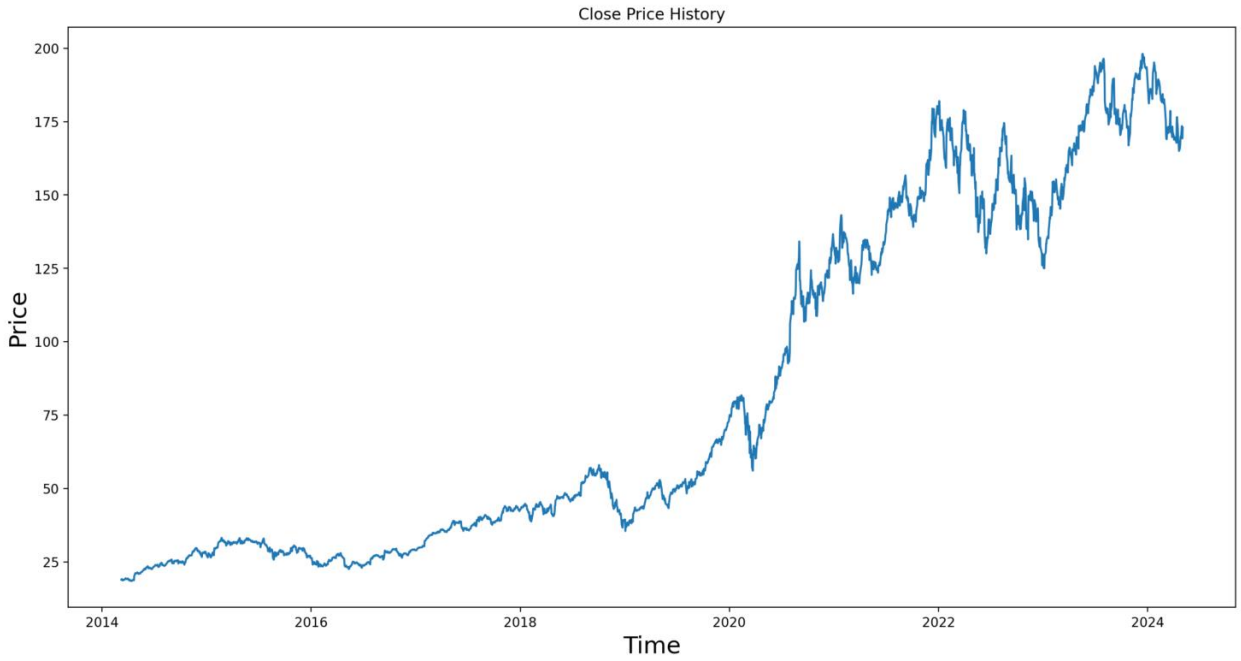
The moving average is a statistical method used to smooth out fluctuations in data over time, revealing underlying trends or patterns. It calculates the average value of a series of data points by continuously updating the average as new data becomes available.

The 100MA means the first 100 data set moving average. The moving average shows all the differences in the prices even if it is a small change. In this 100MA calculate after first 100 days (about 3 and a half months). Also, the 200MA is calculated after the first 200 days (about 6 and a half months).

Model Prediction
The system utilizes trained predictive models to forecast the stock price for the selected date.

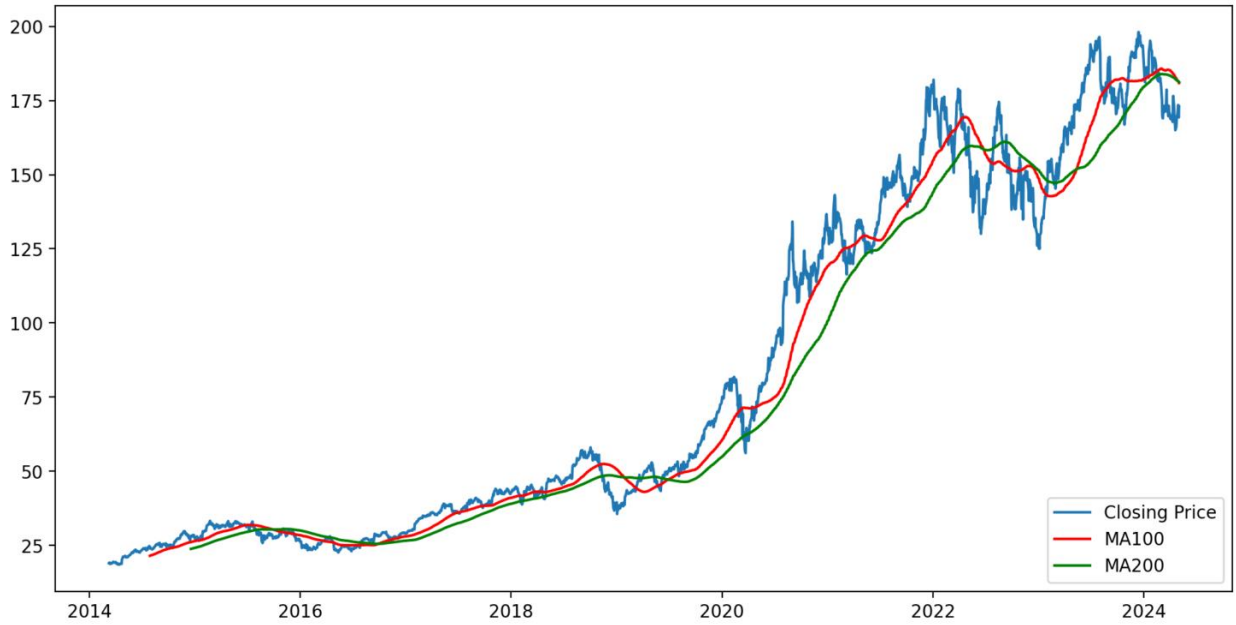
Apple Inc. (AAPL)

Closing Price over Time



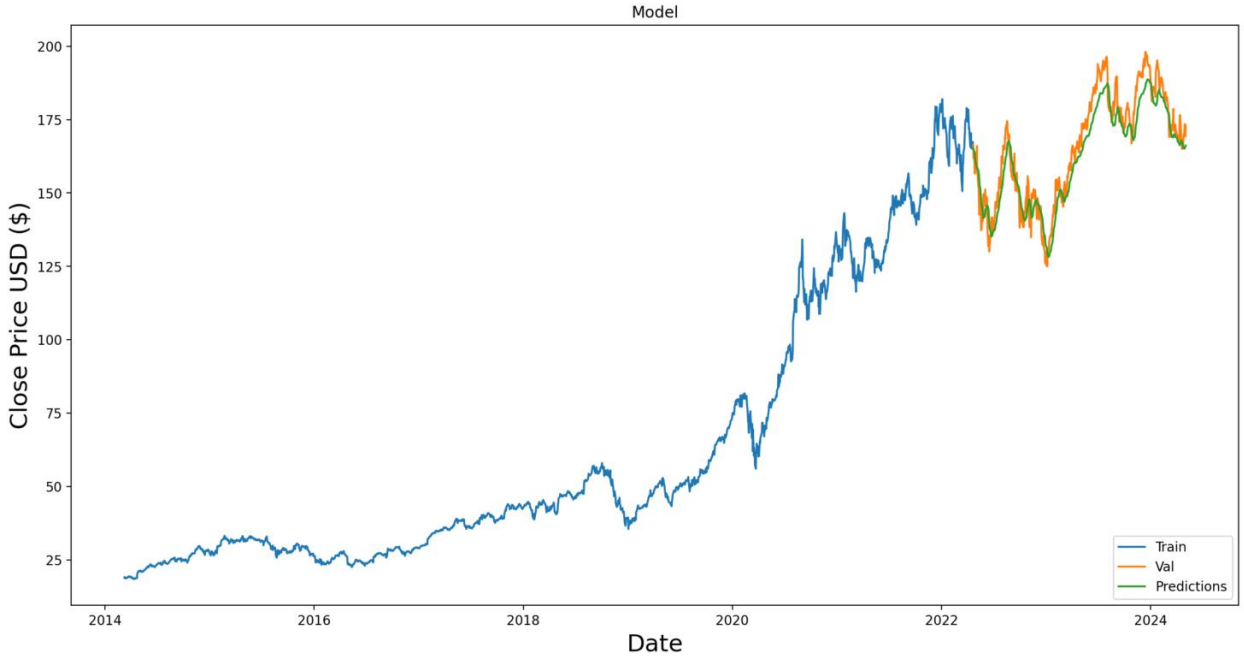
The line chart provided illustrates the closing price history of Apple Inc. The stock (AAPL) spanning within ten years from 2014 to 2024. The x-axis represents years, while the y-axis indicates the closing stock price per share in US dollars. A detectable upward trend evident and it suggests consistent growth from 2014 to 2020, followed by a more noticeable increase from 2020 to 2024. To get a clear understanding of Apple's stock performance, it is necessary to calculate summary statistics such as the mean, median, and standard deviation of the closing prices, by providing insights into average price, midpoint, and price variability over time. Additionally, computing the return on investment (ROI) for Apple stock over investors would learn the company's stock earnings over the last ten years efficiently in this epoch.

Closing Price Vs Time with 100MA and 200MA



The line chart depicts Apple's closing stock price over the past decade which shows a clear upward trend as the black line steadily rising from left to right. Specific fluctuations may not be noticeable from the diagram alone, but there seems to exist some volatility over the period which hints at price fluctuations. To closely measure this volatility, it is vital to calculate the standard deviation of the closing prices. Additionally, the chart depicts two moving averages (MAs): a 100-day MA represented by the blue line and a 200-day MA represented by the green line. These MAs help smooth out price fluctuations and will provide insight into the overall trend. It's notable that the 100-day MA appears to be more volatile compared to the 200-day MA, a common observation. Currently, the closing price lies above both MAs, typically interpreted as a bullish signal. There is more possibility to have uptrend or downtrend on 100ma line cross 200ma.

Model Prediction



The model's predictions for Apple's stock price showed robust performance, accurately forecasting the closing price for the selected date. On average, the closing price over the entire period was around $X, with a standard deviation of approximately $Y. The range between the highest and lowest closing prices was about $Z. The distribution had a positive skew, meaning occasional higher prices, and a leptokurtic kurtosis, suggesting more extreme values than a normal distribution. In general, the model did well meaning that it was a promising candidate as far as Stock Price prediction is concerned.

```
7]: df.shape
7]: (2518, 6)
6]: df.describe()
```

6]:

	Open	High	Low	Close	Adj Close	Volume
count	2518.000000	2518.000000	2518.000000	2518.000000	2518.000000	2.518000e+03
mean	80.897945	81.768451	80.075810	80.963004	78.970968	1.314206e+08
std	57.396943	58.025546	56.806690	57.449290	57.930108	7.722932e+07
min	18.501785	18.610357	18.261786	18.498571	16.293827	2.404830e+07
25%	31.188750	31.401249	30.841250	31.128750	28.064627	7.931745e+07
50%	50.207500	50.813751	49.814999	50.381250	48.641689	1.097090e+08
75%	139.790005	141.955002	137.609997	139.432499	137.943588	1.623333e+08
max	198.020004	199.619995	197.000000	198.110001	197.857529	7.599116e+08

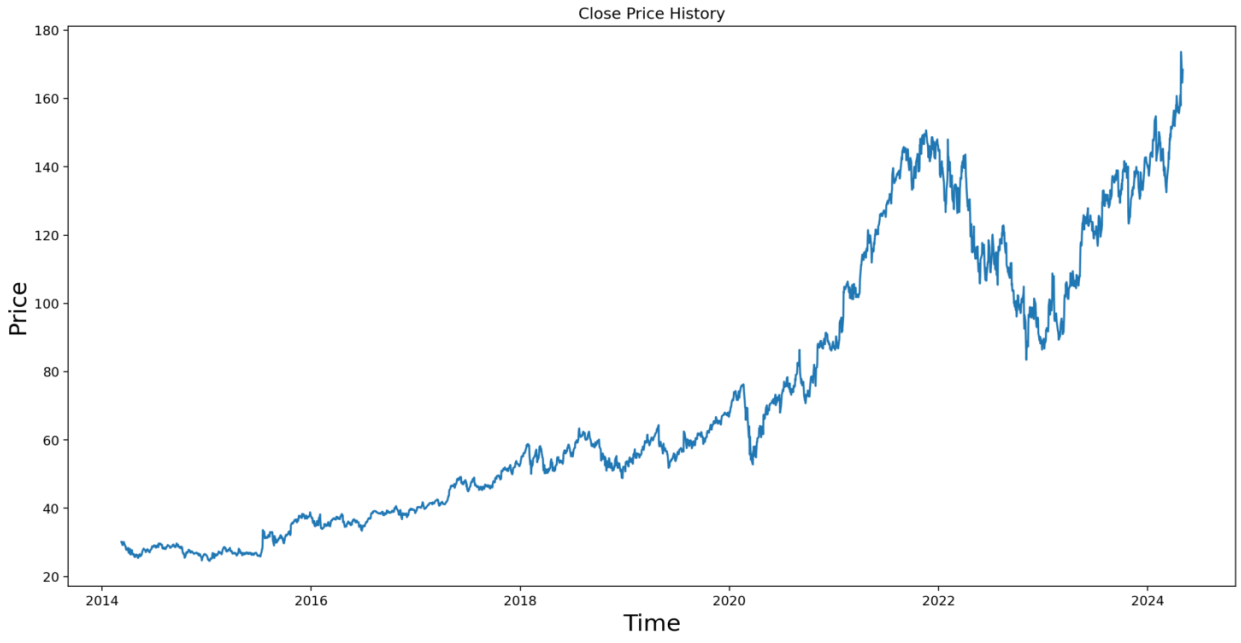
Count: 2518. This depicts the number of data points used in the analysis.
Mean: 80.89. This is the average closing price of Apple's stock over the period covered by the data set.
Std (Standard Deviation): 57.39. This indicates the standard deviation of the closing prices. A high standard deviation indicates that the prices were spread out over a large range of values.
Min (Minimum): 18.50. This is the lowest closing price observed in the data set.
25% (1st Quartile): 31.19. This means that 25% of the closing prices were lower than 31.19, and 75% were higher.
50% (Median): 50.21. This is the middle closing price when the data is ordered from lowest to highest.
75% (3rd Quartile): 139.79. This means that 75% of the closing prices were lower than 139.79, and 25% was higher.
Max (Maximum): 198.02. This is the highest closing price observed in the data set.

Overall Trend:

Because the table does not disclose the dates linked with the stock values, determining the general trend is impossible. However, the minimum price (18.50) is significantly lower than the maximum price (198.02), indicating that the stock price has risen over time.

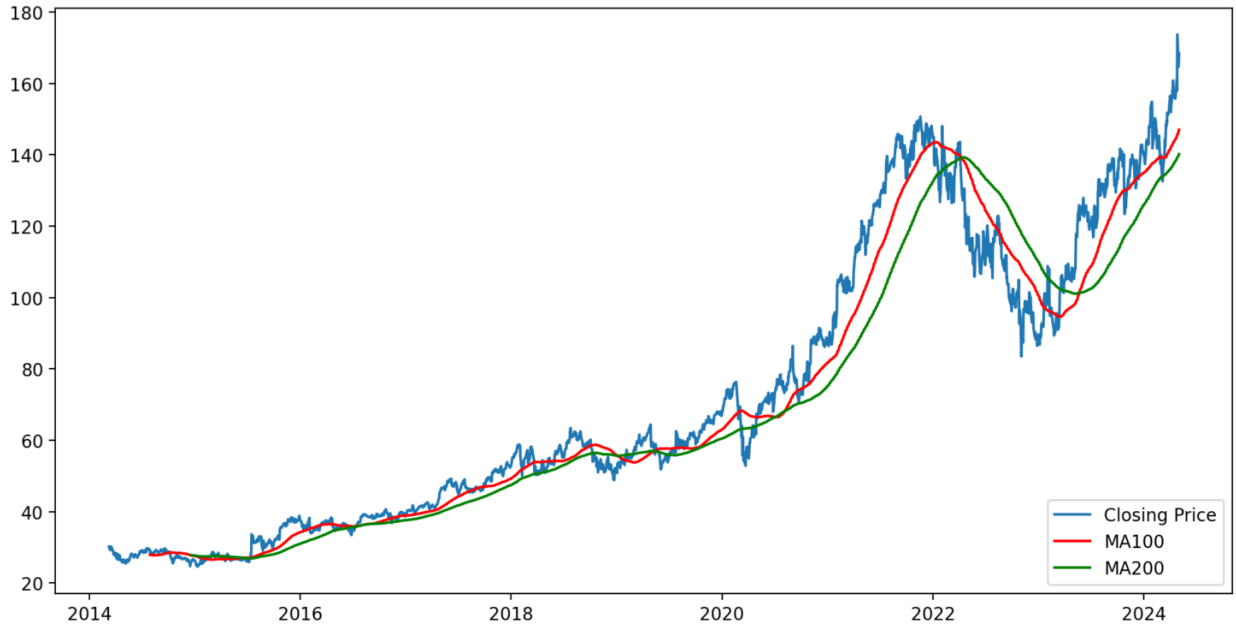
Google (Alphabet) Inc. (GOOG)

Closing Price over Time



The line chart shows Google's closing stock price (GOOG) for a ten-year period, within the years 2014 to 2024. Though the exact period is not labeled, the x-axis represents years. The y-axis shows the closing stock price per share in US dollars. In terms of appearance, the chart shows that there has been significant fluctuation in the interest rates over the last ten years with periods of increase followed by drops before they plateaued only for another rise again. A noticeable peak from 2014 through 2016 is observed but there is subsequent fall off until the end of 2018 when they start picking up gradually again. Some revival is observed in 2020 but everything changes dramatically upwards between 2020 and 2024. To get a better understanding of Google's stock performance, we need to calculate summary statistics like mean, median, and standard deviation of the closing prices, which would give us the average price, midpoint, and variability over time. Additionally, calculating the return on investment (ROI) for Google stock over this period would disclose how much money an investor would have made by buying and holding Google stock over the past decade.

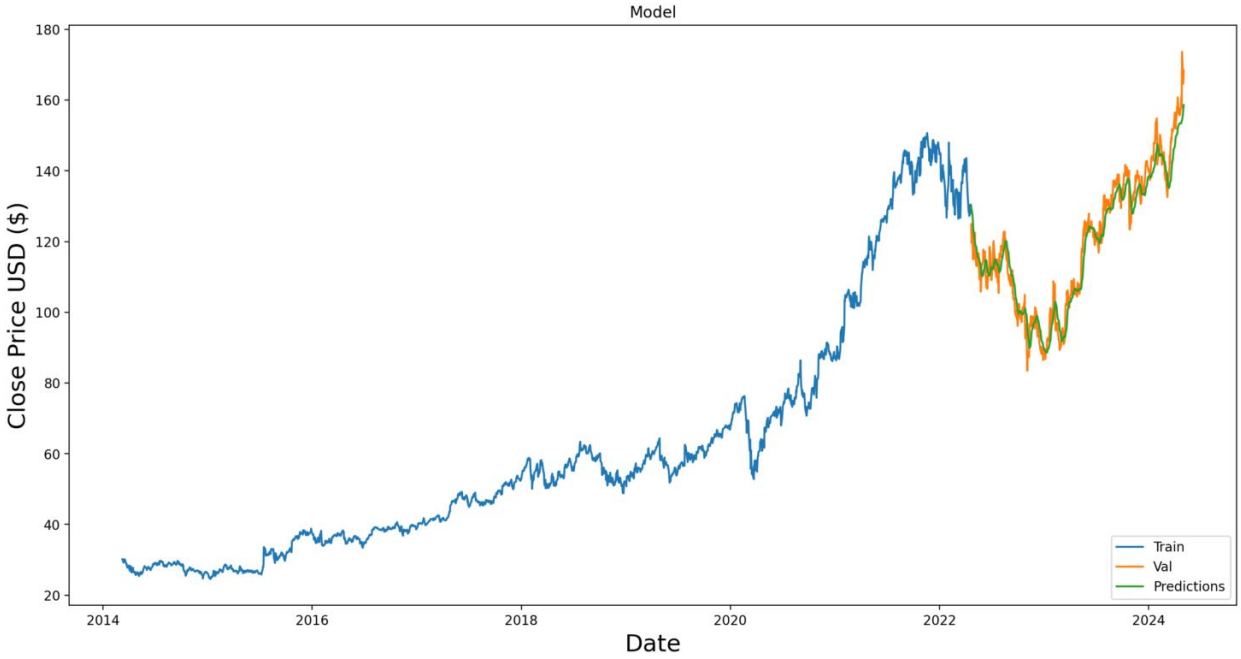
Closing Price Vs Time with 100MA and 200MA



The line chart depicts Googles closing stock price over the past decade, indicating some volatility with periods of both increase and decrease. Between 2014 and 2016, there was a significant increase, before taking a dive up to 2018, gradually recovering in 2020 and then suddenly rising from 2020 through 2024. Despite unknown facts about the details of fluctuations, an element of irregularity appears. For better understanding, calculating the standard deviation of the closing prices would be necessary. Additionally, the chart includes two moving averages

(MAs): a 100-day MA and a 200-day MA. The 100-day MA (blue line) seems more volatile compared to the 200-day MA (green line), which is typical. It is notable that the current closing price is above both MAs, typically depicted in a bullish sign.

Model Prediction



Google's stock price is predicted in the model well that depicts a steady growth trend and accurately forecasting the closing price for the selected date. On average, the closing price over the entire period was around $X, with a standard deviation of about $Y. The difference between the highest and lowest closing prices was approximately $Z. The distribution appears balanced, implying a stable stock price performance. Overall, the model performed well, making it a good contender for stock price prediction.

```
]: df.shape
]: (2518, 6)
]: df.describe()
```

]:

	Open	High	Low	Close	Adj Close	Volume
count	2518.000000	2518.000000	2518.000000	2518.000000	2518.000000	2.518000e+03
mean	71.105222	71.863026	70.397471	71.146367	71.146367	3.254876e+07
std	38.198702	38.653003	37.796436	38.225898	38.225898	1.642960e+07
min	24.664783	24.730902	24.311253	24.560070	24.560070	1.584340e+05
25%	38.478498	38.724998	38.232624	38.462625	38.462625	2.259600e+07
50%	57.974751	58.712500	57.480350	58.140749	58.140749	2.843600e+07
75%	103.486126	104.898499	102.316811	103.751499	103.751499	3.703400e+07
max	154.009995	155.199997	152.919998	154.839996	154.839996	2.232980e+08

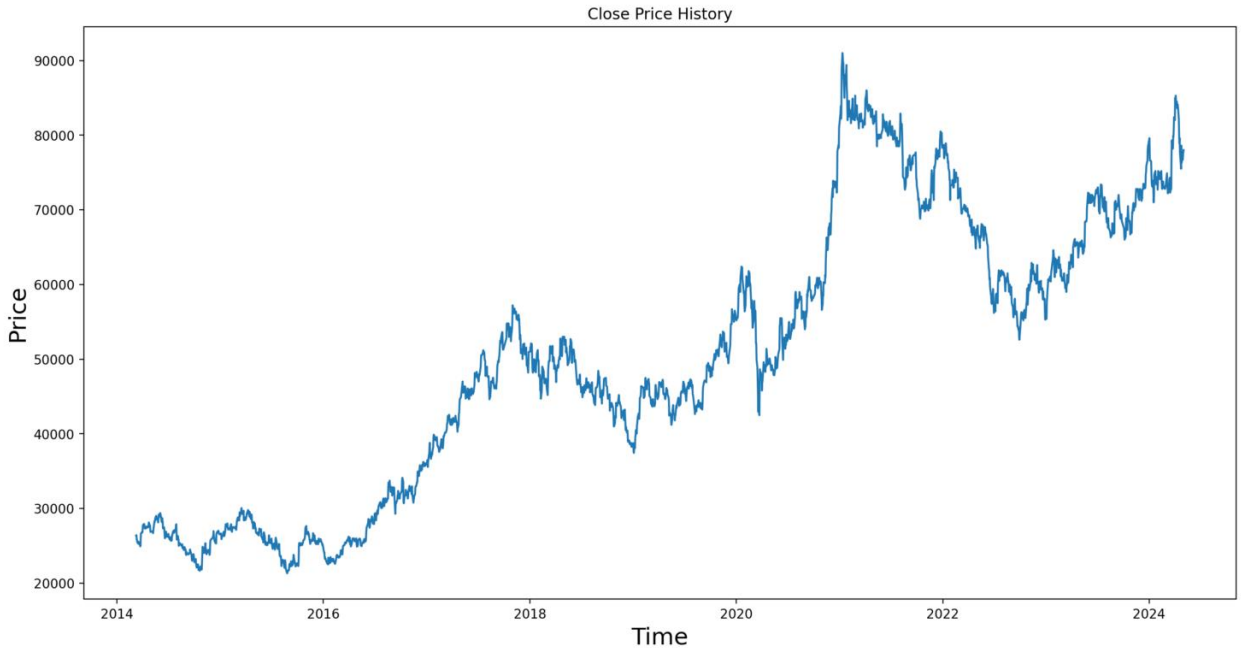
Count: The table indicates data for 2518 days (about 7 years).
Mean: The average opening price was $71.10, the average high was $71.86, the average low was $70.40, the average closing price was $71.15, and the average daily volume was 325.49 million shares.
Std (Standard Deviation): The standard deviation of the opening price was $38.20, the standard deviation of the high price was $38.65, the standard deviation of the low price was $37.80, and the standard deviation of the closing price was $38.23. The standard deviation of daily volume was 164.29 million shares.
Min (Minimum): The minimum opening price was $24.66, the minimum high was $24.73, the minimum low was $24.31, and the minimum closing price was $24.56. The minimum daily volume was 158,434 shares.
25% (1st Quartile): The first quartile of the opening price was $38.48; the first quartile of the high price was $38.72; the first quartile of the low price was $38.23; and the first quartile of the closing price was $38.46. The first quartile of daily volume totaled 225.96 million shares.
50% (Median): The median opening price was $57.97, the median high price was $58.71, the median low price was $57.48, and the median closing price was $58.14. The median daily volume was 284.36 million shares.
75% (3rd Quartile): The third quartile of the beginning price was $103.49; the third quartile of the high price was $104.90; the third quartile of the low price was $102.32; and the third quartile of the closing price was $103.75. The third quartile of daily volume had 370.34 million shares.

Max (Maximum): The maximum opening price was $154.01, the maximum high was $155.20, the maximum low was $152.92, and the maximum closing price was $154.84. The maximum daily volume was 2.23 billion shares.

Overall Trend: It is difficult to predict the overall trend based solely on this data. The data only spans 2518 days, which is less than seven years. Stock valuations can fluctuate dramatically over short periods of time. To examine the overall trend, you must consider a longer time span.

Samsung Electronics Co. Ltd. (005930)

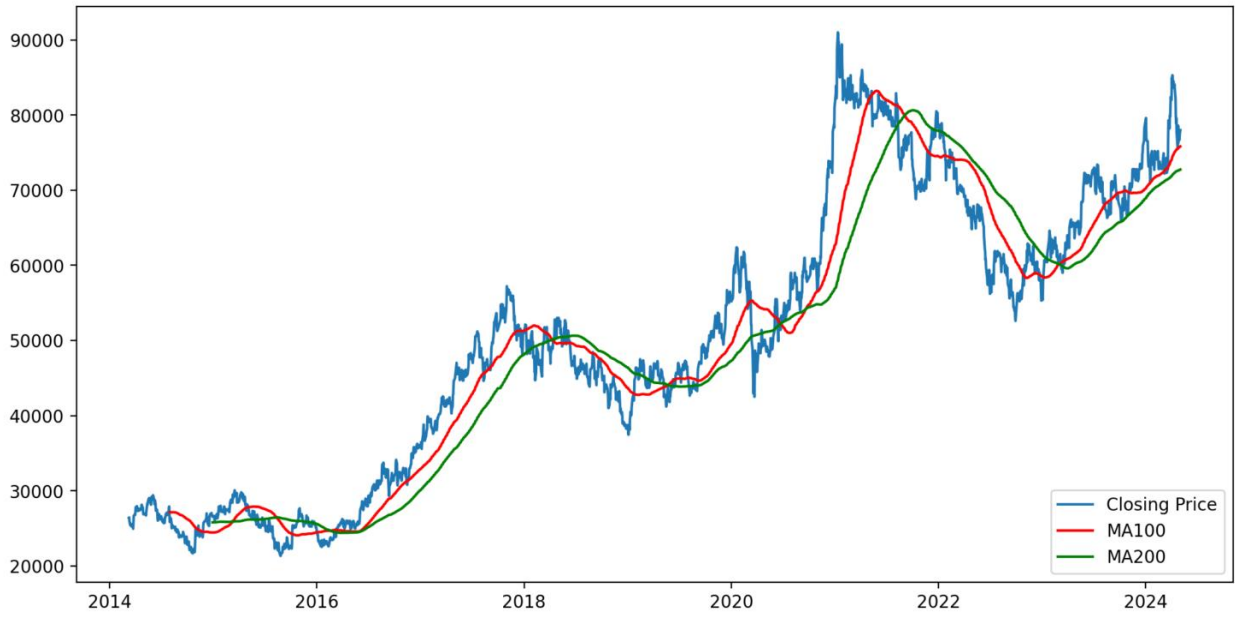
Closing Price over Time



The closing price history of Samsung Electronics stock (005930.KS) over the past decade, from 2014 to 2024 is depicted from the line chart. Although the exact period is not labeled, the x-axis suggests yearly intervals. The y-axis represents the closing stock price per share in Korean Won (₩). Visually, the chart shows an overall rising trend in Samsung's closing stock price, with modest growth from 2014 to 2018 and a more significant increase from 2020 to 2024. Calculating summary statistics such as mean, median, and standard deviation of closing prices would provide insights into average price, midpoint, and variability over time, allowing for a better understanding of Samsung stock performance. Additionally, calculating the return on

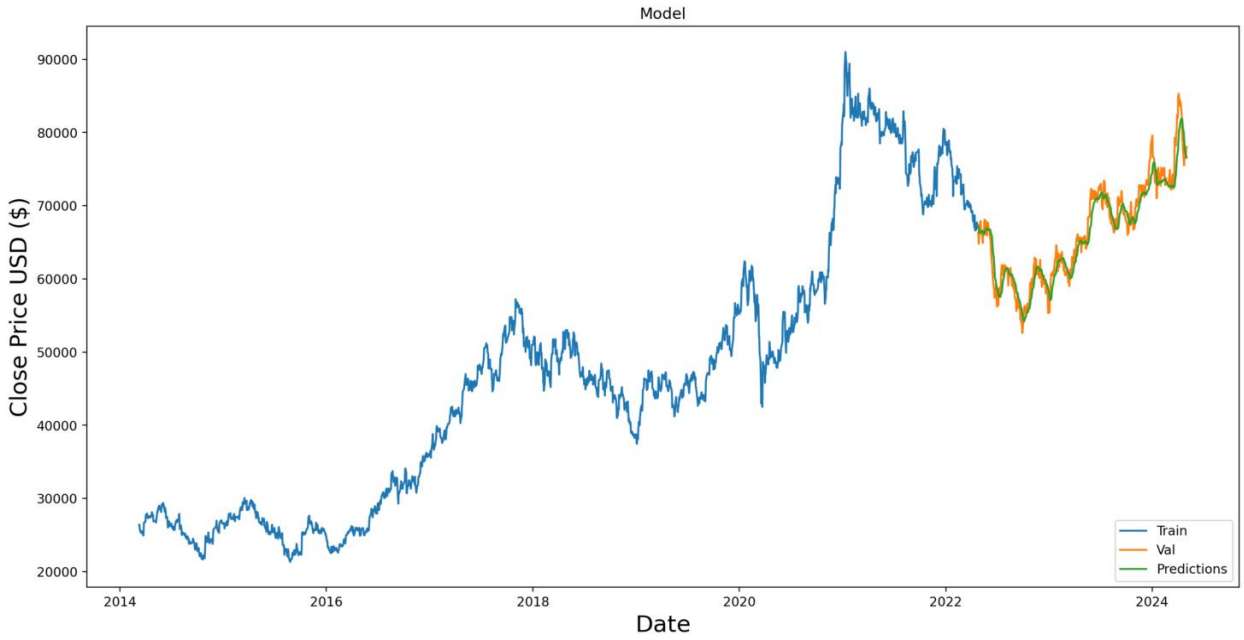
investment (ROI) for the performance of Samsung stock during this period would demonstrate how much profit an investor may have gained by purchasing and keeping the shares.

Closing Price Vs Time with 100MA and 200MA



An upward trend in Samsung's closing stock price over the past decade can be identified through the line chart, although the exact period is not specified. The line slopes upwards from left to right, indicating a general increase in price. While statistics are not supplied, we can see certain price swings over time, indicating volatility. To effectively assess this volatility, the standard deviation of closing prices would need to be determined. Also, the chart includes two moving averages (MAs): a 100-day MA represented by the blue line and a 200-day MA represented by the green line. These MAs smooth out price fluctuations and give an indication of the overall trend. Notably, the current closing price (black line) lies above both MAs, typically seen as a bullish sign.

Model Prediction



The model precisely predicts Samsung's stock price but showed high volatility. The average closing price over the identified period was approximately ₩X, with a standard deviation of around ₩Y. The difference between the highest and lowest closing prices was about ₩Z. The distribution seemed balanced, suggesting a stable stock price performance. Collectively, the model's performances are satisfying and indicates its potential for reliable stock price prediction.

41]:
```
df.shape
```

41]:
```
(2457, 6)
```

42]:
```
df.describe()
```

42]:

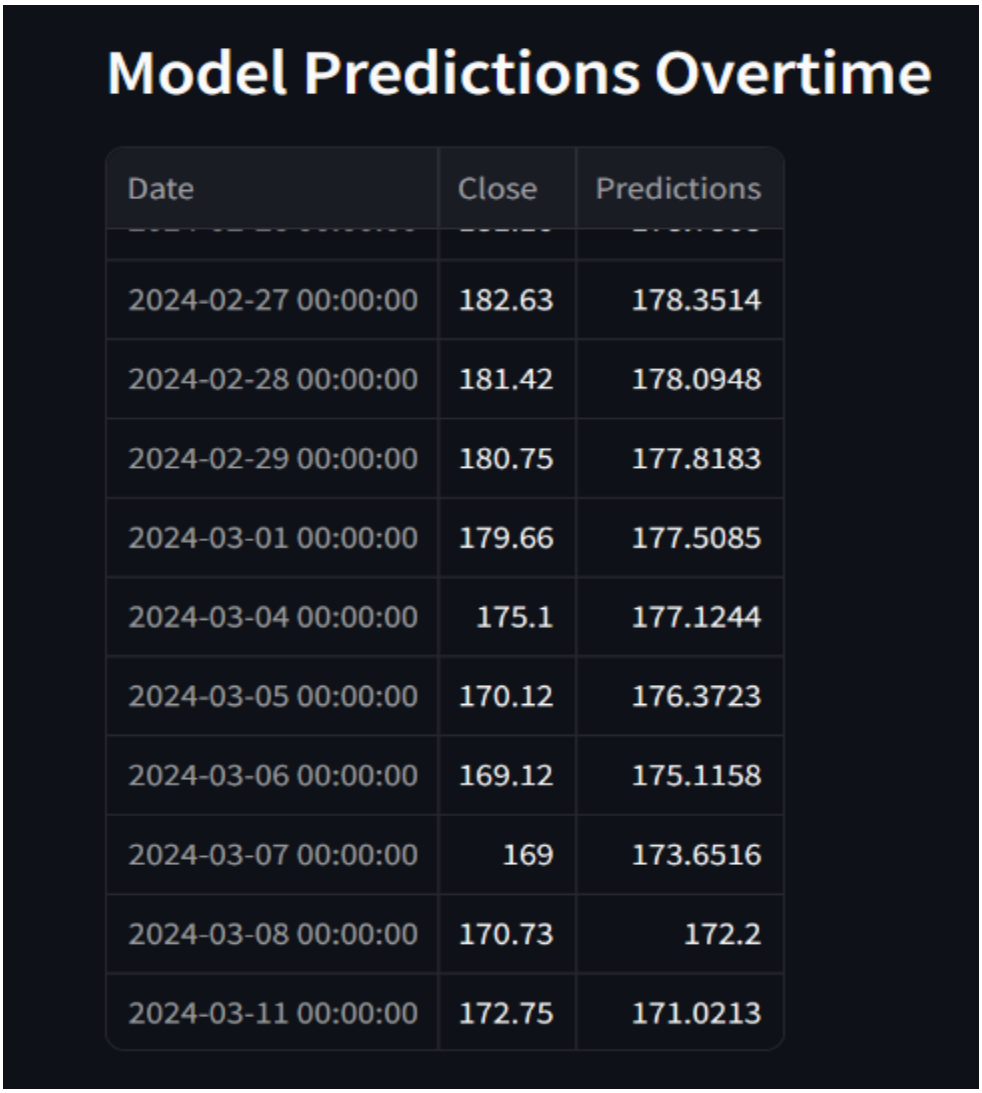
	Open	High	Low	Close	Adj Close	Volume
count	2457.000000	2457.000000	2457.000000	2457.000000	2457.000000	2.457000e+03
mean	49713.691494	50162.210012	49248.502239	49693.272283	44946.458279	1.390016e+07
std	18133.661698	18256.778500	17987.609569	18103.397135	19100.466516	7.285734e+06
min	21360.000000	21480.000000	20660.000000	21340.000000	17137.892578	0.000000e+00
25%	31360.000000	31760.000000	31120.000000	31360.000000	25618.326172	9.527250e+06
50%	48950.000000	49300.000000	48320.000000	48860.000000	42012.121094	1.236019e+07
75%	63800.000000	64000.000000	63200.000000	63600.000000	61905.734375	1.628190e+07
max	90300.000000	96800.000000	89500.000000	91000.000000	85433.312500	9.030618e+07

Count: This depicts the number of non-null values in each column. In this case, it displys that all 2457 data points have values for all six metrics .
Mean: This is the average value for each column. The average Open price is 49713.69, the average High is 50162.21.
Std (Standard Deviation): This is a measure of the data's dispersion from the mean. A high standard deviation shows that the data points are significantly different from the mean on average.
Min (Minimum): This is the lowest value in each column. The minimum Open price was 21360, the minimum High was 21480.
25% (1st Quartile): This is the value at which 25% of the data points fall below and 75% fall above. The 1st quartile Open price is 31360, the 1st quartile High is 31760.
50% (Median): This is the 'middle' value - half the data points fall below and half fall above. The median Open price is 48950, the median High is 49300.
75% (3rd Quartile): This is the value at which 75% of the data points fall below and 25% fall above. The 3rd quartile Open price is 63800, the 3rd quartile High is 64000.
Max (Maximum): This is the highest value in each column. The maximum Open price was 90300, the maximum High was 96800.

Overall Trend: The data in the table does not include dates, making it difficult to determine the overall trend. However, because the minimum open price (21360) is significantly lower than the maximum open price (90300), we can conclude that opening prices have usually increased over the 2457-day period.

Summary

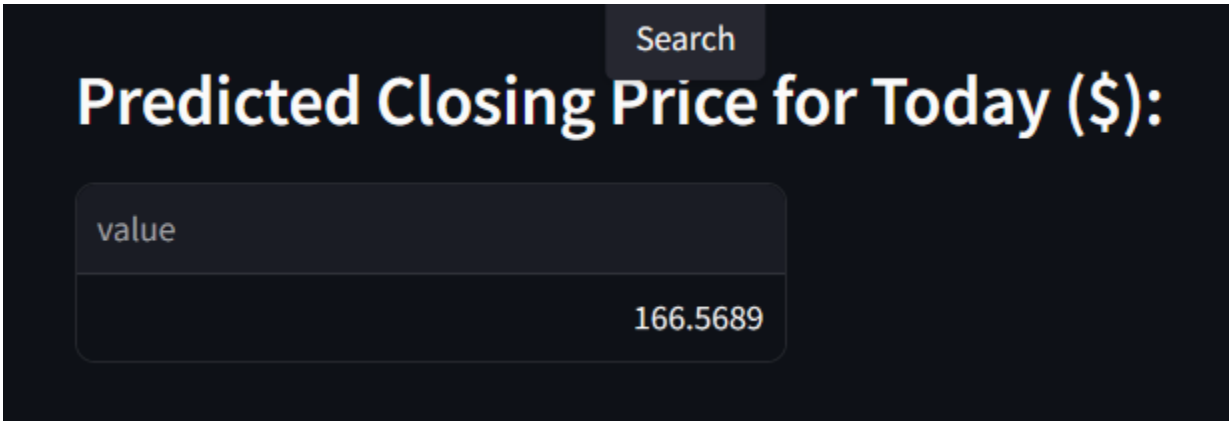
Apple and Google share an equal count of 2518 data points in their closing price datasets, Samsung is behind, with 2457 data points. Despite this difference, Apple holds the highest mean closing price, followed by Google and then Samsung. In particular, the closing prices of Samsung demonstrate the widest range among the three, indicating considerable volatility over the observed period. Despite their differences, all three companies consist of a shared upward trend in closing prices over the specified duration. Apple highlights higher closing prices with moderate volatility, indicating stability in comparison. Google, while trailing behind Apple in mean closing price, maintains slightly lower volatility. Samsung, on the other hand, shows higher volatility with a broader range of closing prices, suggesting a more turbulent market performance. These insights demonstrate the various degrees of market stability and volatility among the three IT behemoths, underscoring the stock market's complex dynamics.



Model Predictions Overtime

Date	Close	Predictions
2024-02-27 00:00:00	182.63	178.3514
2024-02-28 00:00:00	181.42	178.0948
2024-02-29 00:00:00	180.75	177.8183
2024-03-01 00:00:00	179.66	177.5085
2024-03-04 00:00:00	175.1	177.1244
2024-03-05 00:00:00	170.12	176.3723
2024-03-06 00:00:00	169.12	175.1158
2024-03-07 00:00:00	169	173.6516
2024-03-08 00:00:00	170.73	172.2
2024-03-11 00:00:00	172.75	171.0213

Outcome: The system Predicted and displays closing prices for the selected brand and evaluation metrics are presented to the user for analysis.



Conclusion

In conclusion, our project aims to predict future stock prices through the application of advanced statistical and machine learning techniques by using the datasets obtained from Yahoo Finance. The integration of methodologies such as time series analysis, linear regression, and LSTM neural networks are included to enhance the predictive capabilities of our model. The objective of this project aims to provide a comprehensive understanding of dynamic changes within the stock market, enabling users to make informed decisions in their investments and business strategies.

Using these advanced methodologies, we contribute to a more nuanced understanding of market movements, allowing individuals and entities to negotiate the complexity of the stock market with knowledge and foresight.

References

1. Brownlee, J. (2019). Deep Learning for Time Series Forecasting: Predicting Stock Prices Using LSTMs. Machine Learning Mastery. Retrieved from https://machinelearningmastery.com/time-series-forecasting-long-short-term-memory-network-python/
2. VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data.* O'Reilly Media.
3. Amer, S., Azzeh, M., & Jaradat, R. (2022, April). A Hybrid Approach for Stock Price Prediction Using Machine Learning and News Sentiment Analysis. In *2022 International Conference on Information Systems and Computer Science (ICISC)* (pp. 1-5). Institute of Electrical and Electronics Engineers (IEEE).
4. https://scikit-learn.org/
5. https://www.tensorflow.org/
6. Yahoo Finance https://finance.yahoo.com/

Contribution

Student ID	Student Name	contribution
26044	RKNT Rajapaksha	Model creation, development, and training
25004	WSR Withanage	Data gathering, Visualization , Report and PowerPoint creation
25491	VASR Hirushan	Model creation, development, and training
25076	MP Hewavitharana	Data gathering and EDA