SCHOOL OF COMPUTING, UUM COLLEGE OF ARTS AND SCIENCES

SKIH3033 INFORMATION VISUALIZATION

INDIVIDUAL PROJECT

GitHub link: *https://github.com/Liyingwong/Predicting-Tesla-s-Stock-Performance/blob/main/README.md*

Lecturer Name:

Dr. Nur Suhaili Binti Mansor

Due Date:

28 July 2024

Prepared By:

NAME	MATRIC NO.
WONG LI YING	289912

Predicting Tesla's Stock Performance: A Comprehensive Data Visualization and Analysis Approach

Wong Li Ying
University Utara Malaysia, Malaysia, liyingwongly28@gmail.com

ABSTRACT. This project explores Tesla's stock performance through a detailed data visualization and predictive modeling approach. Using historical stock data from January 1, 2021, to January 1, 2024, sourced from Yahoo Finance, a comprehensive analysis to uncover trends, patterns, and volatility in Tesla's stock prices was conducted. Microsoft Excel has been employed for initial data exploration, Power BI for dynamic visualizations, and Python (Jupyter Notebook) for advanced analysis and predictive modeling using Long Short-Term Memory (LSTM) networks. Our findings reveal significant volatility in Tesla's stock, with notable seasonal trends and a predicted downturn in January 2024. This study underscores the importance of robust data visualization and predictive analytics in making informed investment decisions.

Keywords: Data visualization, financial analysis, LSTM, predictive modeling, Tesla stock

INTRODUCTION

This project focus on developing a comprehensive data visualization story aimed at predicting Tesla's stock prices. The primary objective is to provide investors and analysts with a detailed and visually engaging narrative that encapsulates historical stock performance, trends, and volatility, thereby facilitating informed investment decisions.

Tesla Inc. (TSLA), a leading electric vehicle manufacturer, has revolutionized the automotive industry with its innovative electric cars, energy storage solutions, and solar products. Founded in 2003 by Elon Musk, Tesla has rapidly grown into one of the most valuable companies globally, renowned for its cutting-edge technology and ambitious vision for a sustainable future (Schreiber & Gregersen, 2024). Tesla's stock market performance has been a focal point of interest for investors due to the company's consistent innovation, strong market influence, and significant growth potential (Babazhanov, Suleyeva, Zhumatov, Kenenbay, & Syrbayeva, 2023).

However, the stock's performance has exhibited significant volatility, driven by factors such as fluctuating market sentiment, production challenges, regulatory changes, and broader economic conditions (Qingren, 2023). This volatility has made Tesla's stock both an attractive opportunity and a risky proposition for investors, highlighting the need for meticulous analysis and predictive modeling.

This project harnesses the power of data visualization to unravel the complexities inherent in Tesla's stock movements, offering a nuanced perspective on its potential future trajectories. The dataset employed for this analysis is sourced from Yahoo Finance, encompassing daily historical stock data for Tesla over a period extending from January 1, 2021, to January 1, 2024. This dataset includes several key variables: date, open, high, low, close, adjusted close, and volume. These variables provide a robust foundation for analyzing Tesla's stock performance, allowing for the identification of patterns and trends that are crucial for making informed predictions.

The project demonstrates how historical stock data can be transformed into a compelling visual story. This story not only highlights past performance but also provides predictive insights, offering a holistic view of Tesla's stock. The analysis revealed that Tesla's stock has been highly volatile, making it a risky investment at present. Historical data indicated a gradual decline in stock prices from 2021 to 2023, despite occasional periods of increase, particularly around the end of August to mid-October. These insights are critical for investors looking to make informed decisions about their investment strategies.

In summary, this project showcases the power of data visualization in understanding and predicting stock performance. By utilizing a rich dataset from Yahoo Finance and employing advanced visualization tools, it provides a comprehensive analysis of Tesla's stock, helping investors navigate the complexities of the financial market.

DATASET

The dataset used in this project is a comprehensive collection of Tesla's daily historical stock data, sourced from Yahoo Finance (https://finance.yahoo.com/quote/TSLA/history?period1=1672531200&period2=1704067200&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true). It covers the period from January 1, 2021, to January 1, 2024, providing a detailed record of Tesla's stock performance over these three years. This extensive dataset is pivotal for conducting a thorough analysis, as it includes various essential variables that reflect the daily trading activity and stock price fluctuations. The dataset includes several key variables. The "Date" column represents the specific trading day, serving as the primary axis for time series analysis and allowing us to observe how Tesla's stock prices have evolved over time. The "Open" field records the stock's opening price for each trading day, which is the initial price at which the stock is traded when the market opens, providing a baseline for the day's trading activity.

The "High" field captures the highest price at which the stock was traded during a given day. Analyzing the high prices helps in understanding the maximum valuation the market attributed to the stock on any particular day. In contrast, the "Low" field records the lowest price at which the stock was traded during the day. The low prices are crucial for understanding the minimum valuation and the stock's volatility within the trading day. The "Close" field represents the final trading price of the stock at the end of the trading day. The closing price is often considered the most important price point of the day, as it reflects the market's consensus value at the end of trading hours. Additionally, the "Adjusted Close" field adjusts the closing price to account for events such as stock splits, dividends, and rights offerings. The adjusted close provides a more accurate reflection of the stock's value over time by considering these adjustments.

Lastly, the "Volume" field indicates the number of shares traded during the day. Trading volume is a key indicator of market activity and liquidity. Higher volumes often indicate higher investor interest and can be associated with significant price movements. Understanding trading volume helps in assessing the market's enthusiasm or caution regarding the stock. The dataset’s richness and granularity make it an ideal candidate for exploratory data analysis and predictive modeling. By examining these fields, we can gain insights into daily stock performance, identify trends and patterns, and assess the stock’s volatility. To prepare the dataset for analysis, several pre-processing steps were undertaken. This included removing any empty rows and duplicate entries to ensure the data's integrity and accuracy. Handling missing data and duplicates is crucial to prevent skewed results and maintain the reliability of the analysis.

Furthermore, the dataset's temporal nature allows for various types of trend analyses. For instance, calculating Simple Moving Averages (SMA) over different intervals, such as 30 days, 100 days, and 200 days, can smooth out short-term fluctuations and highlight longer-term trends. Seasonal trend analysis can also be conducted by aggregating the data on a monthly basis to identify patterns and seasonal effects on stock prices. In addition to trend analysis, volatility analysis can be performed to understand the stock's risk profile. By examining the range between the high and low prices and the volume of trading, we can gauge how frequently and significantly the stock price deviates from its average, which is critical for risk assessment.

Overall, this dataset serves as a robust foundation for constructing a detailed and informative visualization story. It allows for the exploration of various analytical dimensions and the derivation of actionable insights that can inform investment strategies and predictions about Tesla's future stock performance. Through meticulous data pre-processing and comprehensive analysis, the dataset reveals the intricate dynamics of Tesla's stock, providing a valuable resource for investors and analysts alike.

Date	Open	High	Low	Close	Adj Close	Volume	Day	Month	Quarter	Year	30 MA	100 MA	200 MA	Average Pı	PCT Chang	100 Vol
4/1/2021	239.82	248.1633	239.0633	243.2567	243.2567	145914600	4	1	Q1	2021	277.2499	239.751	236.3386			
5/1/2021	241.22	246.9467	239.7333	245.0367	245.0367	96735600	5	1	Q1	2021	277.9124	239.4213	236.5628		0.007317	0.041101
6/1/2021	252.83	258	249.7	251.9933	251.9933	134100000	6	1	Q1	2021	278.5484	239.055	236.7806		0.02839	0.041093
7/1/2021	259.21	272.33	258.4	272.0133	272.0133	154496700	7	1	Q1	2021	278.9541	238.6147	237.0107		0.079447	0.04109
8/1/2021	285.3333	294.83	279.4633	293.34	293.34	225166500	8	1	Q1	2021	278.7876	237.9117	237.1667		0.078403	0.040599
11/1/2021	283.1333	284.81	267.8733	270.3967	270.3967	177904800	11	1	Q1	2021	276.6334	236.8877	237.4081		-0.078214	0.040073
12/1/2021	277	289.3333	275.78	283.1467	283.1467	138812100	12	1	Q1	2021	275.1051	236.1806	237.7535		0.047153	0.039374
13/1/2021	284.2533	286.8233	277.3333	284.8033	284.8033	99937500	13	1	Q1	2021	273.9401	235.3662	238.0676		0.005851	0.039061
14/1/2021	281.13	287.6667	279.5833	281.6667	281.6667	93798900	14	1	Q1	2021	272.2428	234.5302	238.4386		-0.011014	0.039055

Figure 1. The TESLA Stock data obtained from Yahoo Finande.

METHODOLOGY

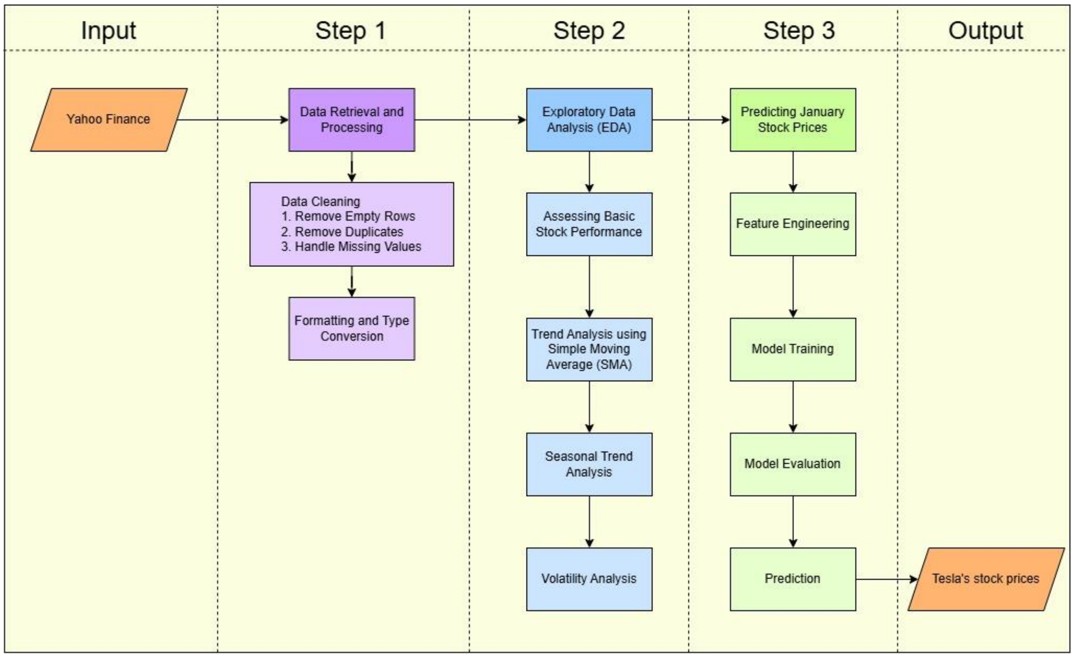

Figure 2. The methodology of Predicting Tesla's Stock Performance.

The methodology for this project encompasses a structured approach to data retrieval, pre-processing, exploratory data analysis (EDA), and predictive modeling. Each step is crucial in transforming raw data into actionable insights and visually compelling narratives that inform investment decisions.

Part 1: Data Retrieval and Processing

The initial phase of the project involved retrieving the dataset from Yahoo Finance. The dataset spans from January 1, 2021, to January 1, 2024, providing comprehensive daily historical data on Tesla's stock prices. The data was downloaded in CSV format to ensure compatibility with various analytical tools and ease of manipulation. Once the dataset was obtained, it underwent a meticulous pre-processing phase to ensure its quality and accuracy. This phase involved several critical steps:

Data Cleaning

The dataset was scrutinized to identify and remove any empty rows or duplicate entries. Empty rows can disrupt analyses by introducing gaps, while duplicates can distort the true representation of the data, leading to misleading insights. Ensuring the dataset is free of these issues is vital for maintaining the integrity of the analysis.

Table 2. ***Data cleaning steps.***

Step	***Description***
Remove Empty Rows	Rows with no data were removed to maintain dataset integrity.
Remove Duplicates	Duplicate entries were identified and removed.
Handle Missing Values	Missing values were imputed using interpolation.

Handling Missing Values

Any missing values within the dataset were addressed using appropriate statistical methods. Depending on the context and extent of missing data, strategies such as interpolation, which estimates missing values based on surrounding data points, or imputation with mean/median values were employed. In cases where the missing data was deemed insignificant, those entries were removed to prevent potential biases in the analysis.

Formatting and Type Conversion

The dataset's columns were formatted to appropriate data types to facilitate efficient analysis. For example, the "Date" column was converted to a datetime format, enabling time series analysis and ensuring accurate chronological plotting of data points. Numerical fields such as "Open," "High," "Low," "Close," "Adjusted Close," and "Volume" were ensured to be in floating-point format for precision in calculations.

Part 2: Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is a critical step in understanding the underlying patterns and trends within the dataset. For this project, EDA was conducted through multiple analytical techniques:

Assessing Basic Stock Performance

The fundamental analysis began with plotting the closing prices of Tesla's stock against the timeframe. This visual representation provided an initial overview of the stock's performance, highlighting major price movements, peaks, and troughs. Such a plot offers a preliminary insight into the stock's behavior over the selected period.

Trend Analysis using Simple Moving Average (SMA)

To identify and understand trends within the stock prices, Simple Moving Averages (SMA) were calculated over different intervals: 30-day, 100-day, and 200-day periods. SMAs help smooth out short-term fluctuations, revealing longer-term trends (Harmsen, Chang, & Hattrup, 2016). The 30-day SMA is typically used to identify short-term trends, the 100-day SMA for medium-term trends, and the 200-day SMA for long-term trends. Plotting these SMAs alongside the stock's closing prices provides a clear visual indication of the underlying trends and helps in identifying periods of growth or decline.

Seasonal Trend Analysis

Seasonal trends were analyzed by aggregating the closing prices on a monthly basis. This analysis aimed to uncover recurring patterns or seasonal effects on Tesla's stock prices. By examining monthly averages, it becomes possible to identify specific times of the year when the stock tends to perform better or worse. This seasonal analysis provides valuable insights for investors regarding the optimal times for buying or selling the stock (Jagwani, Gupta, Sachdeva, & Singhal, 2018).

Volatility Analysis

Volatility analysis was conducted to assess the frequency and magnitude of price changes within the stock. By examining the range between the high and low prices for each day, along with trading volume, the analysis provided insights into the stock's volatility. Higher volatility indicates a higher risk, as the stock price is prone to significant fluctuations. Understanding the volatility helps investors gauge the risk associated with the stock and make informed decisions based on their risk tolerance.

Part 3: Predicting January Stock Prices

The predictive modeling phase aimed to forecast Tesla's stock prices for January 2024. For this purpose, an LSTM (Long Short-Term Memory) regression algorithm was employed. LSTM is a type of recurrent neural network (RNN) that is well-suited for time series forecasting due to its ability to learn and retain long-term dependencies and patterns within the data (Sherstinsky, 2020).

Data Preparation

The cleaned and pre-processed dataset was split into training and testing sets. The training set included data up to December 2023, while the testing set comprised data from January 2024. This split ensured that the model was trained on historical data and evaluated on unseen data, providing an objective assessment of its predictive accuracy.

Feature Engineering

Feature engineering involved selecting and creating relevant features from the dataset to enhance the model's predictive capabilities. In addition to the closing prices, other features such as SMAs, daily price changes, and trading volumes were included. These features provided additional context and information, enabling the model to capture complex patterns and trends more effectively.

Model Training

The LSTM model was trained on the training set, learning from the historical patterns and trends in Tesla's stock prices. The model's architecture included multiple LSTM layers, which allowed it to capture both short-term and long-term dependencies. The training process involved tuning hyperparameters such as the number of layers, number of units per layer, learning rate, and batch size to optimize the model's performance (Li, Kiaghadi, & Dawson, 2021).

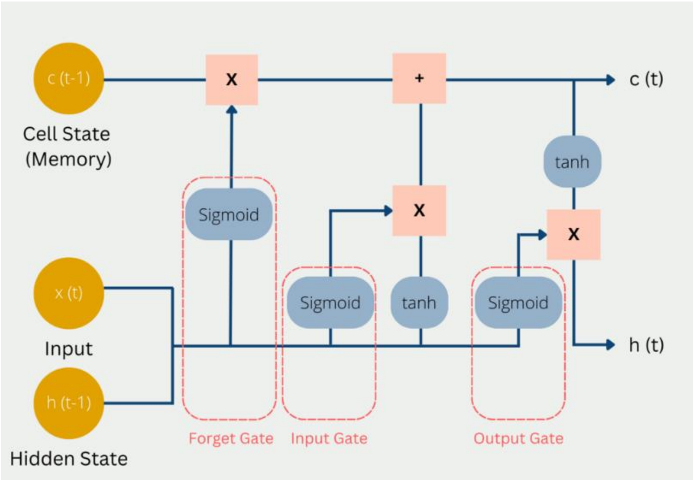

Figure 3. The LSTM Model Architecture.

Model Evaluation

The trained model was evaluated using several metrics to assess its performance. These metrics included Mean Absolute Error (MAE), Mean Squared Error (MSE), R-squared (R2), and Custom Accuracy. MAE and MSE provided insights into the average prediction error, while R-squared measured the proportion of variance in the actual stock prices explained by the model (Echrigui & Hamiche, 2023). Custom Accuracy, a user-defined metric, was also employed to capture specific aspects of the model's performance relevant to the project.

Prediction

The trained LSTM model was used to predict Tesla's stock prices for January 2024. The predictions were then compared against the actual stock prices to assess the model's accuracy and effectiveness. This comparison provided valuable feedback for refining the model and improving future predictions.

TECHNOLOGY REVIEW

Throughout this project, multiple technologies were employed to facilitate data analysis and visualization, each with its strengths and limitations. These tools were pivotal in enabling a comprehensive understanding of Tesla's stock performance and generating insightful predictions. First of all is Microsoft Excel which utilized primarily for its ease of use and wide range of tools for building interactive visualizations. Excel's intuitive interface and extensive library of functions made it an excellent choice for initial data exploration and simple visualizations. It allowed for quick data manipulation, basic statistical analysis, and the creation of straightforward charts and graphs. Excel's pivot tables and conditional formatting features were particularly useful for summarizing and highlighting key data points. However, Excel's limitations became apparent when dealing with large datasets and complex statistical calculations. Its performance can degrade with substantial volumes of data, and it lacks the advanced analytical capabilities required for more sophisticated modeling tasks. As a result, more specialized tools were necessary for in-depth analysis and handling larger datasets efficiently.

Moving on to MS Power BI that has been proved to be a powerful tool for creating interactive and dynamic visualizations. It offers advanced statistical capabilities and extensive customization options that allow for precise and detailed visual representations (Becker & Gould, 2019). Power BI's strength lies in its ability to handle large datasets efficiently and provide interactive dashboards that enable users to explore the data dynamically. Features like drill-downs, slicers, and interactive reports made it possible to delve into different aspects of the data effortlessly. Moreover, Power BI's integration with various data sources and its ability to update data in real-time added significant value to the analysis process. However, some customization options for visual elements were limited, and certain functionalities, such as rotating axis titles and adding linear regression lines, required additional workarounds. These limitations sometimes necessitated a combination of tools to achieve the desired outcome.

Python, particularly the Matplotlib library, was also employed for creating visualizations. Python's flexibility and extensive range of libraries for data analysis and visualization made it an invaluable tool for this project. Libraries such as Pandas, NumPy, and Scikit-learn were instrumental in data preprocessing and analysis. Matplotlib, along with other visualization libraries like Seaborn and Plotly, provided the means to create detailed and customized visualizations (Pimentel, Murta, Braganholo, & Freire, 2021). While the visual quality of Matplotlib may not match that of Power BI or Excel in terms of polish, Python's strength lies in its ability to create highly customized visualizations through coding. This flexibility allows for the implementation of complex analytical techniques and the creation of tailored visualizations that are not constrained by the limitations of specific software. Additionally, Python's capability to integrate with machine learning frameworks enabled advanced predictive modeling, such as the LSTM regression used in this project.

In summary, the methodology outlined above provides a comprehensive approach to data retrieval, preprocessing, exploratory analysis, and predictive modeling. By leveraging the strengths of various technologies, this project delivers a detailed and insightful analysis of Tesla's stock performance, aiding investors in making informed decisions. Each tool played a crucial role: Excel for initial exploration and simple visualizations, Power BI for interactive and dynamic dashboards, and Python for advanced analysis and customized visualizations. Through meticulous data preparation, thorough exploratory analysis, and advanced predictive modeling, this project demonstrates the power of data visualization in understanding and predicting stock performance. The integration of these technologies ensured a robust and multifaceted analysis, highlighting the importance of selecting the appropriate tools based on the specific requirements and complexities of the data analysis task at hand.

DISCUSSIONS AND RESULTS

Interpretation of Findings

Historical Performance

The exploratory data analysis revealed that Tesla's stock prices have undergone significant fluctuations during various periods of market activity. Historical data indicates that while there have been periods of robust growth, there are also times when the stock experienced sharp declines. The time series plot from Jupyter Notebook illustrates these trends clearly, showing how external market forces and company-specific events have impacted stock prices over time. This volatility underscores the importance of understanding the underlying factors driving stock performance to make informed investment decisions.

The trend analysis using Simple Moving Averages (SMAs) in Power BI highlighted that short-term movements often deviate from long-term trends, indicating potential opportunities for investors to capitalize on shorter-term price movements. The 30-day SMA (blue line) highlights short-term volatility and rapid changes, while the 100-day SMA (yellow line) captures medium-term trends, smoothing out short-term fluctuations. The 200-day SMA (orange line) represents long-term trends, indicating overall market health. The actual closing prices (green line) provide the raw data for these averages. For instance, while a 30-day SMA might show a quick rise or fall in stock price, the 100-day and 200-day SMAs often present a more stable. This divergence suggests that there are moments when short-term investors can exploit these deviations for potential gains. Furthermore, the volatility analysis using Bollinger Bands confirms that periods of high volatility are accompanied by greater price fluctuations, providing both risks and opportunities for traders. The Power BI dashboard also provides a comprehensive view of Tesla's stock performance across different time frames, helping investors to better understand market trends and make more strategic investment decisions.

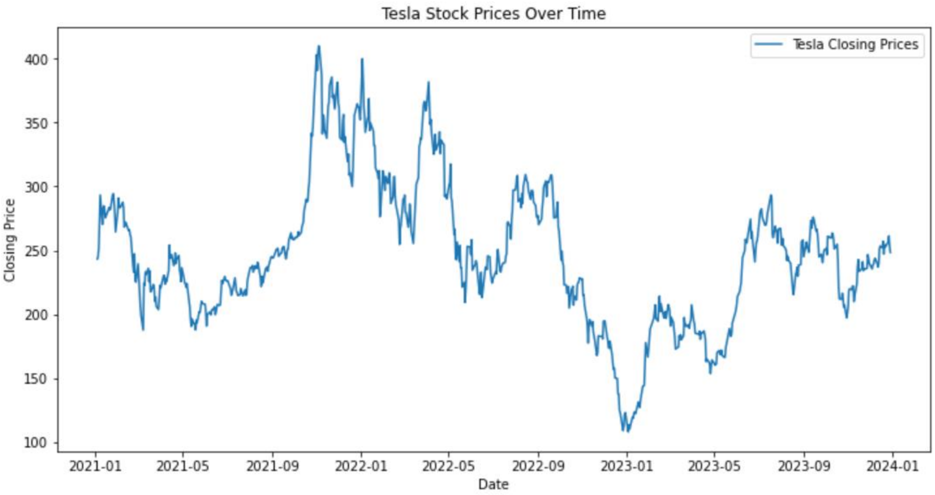

Figure 4. The time series plot shows the closing prices of Tesla stocks observed over a period of time.

Figure 5. The Simple Moving Averages of Tesla Stocks.

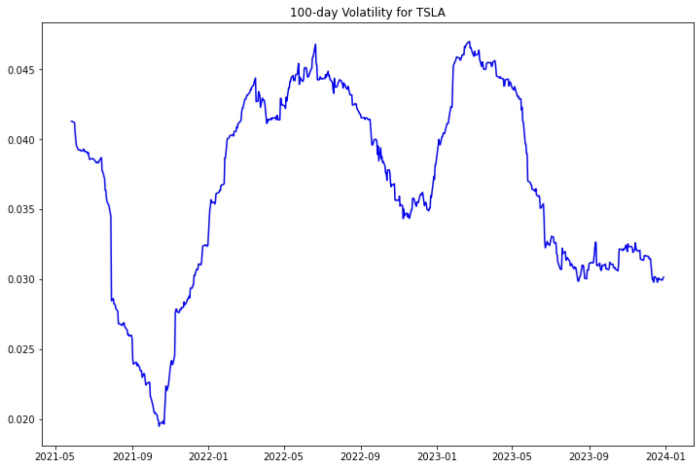

Figure 6. The Volatility Analysis with Bollinger Bands.

Figure 7. Tesla's stock performance across different time frames.

Seasonal Trends

The seasonal trend analysis of Tesla's stock prices from June 2020 to April 2024, as depicted in the graph, reveals significant insights into the stock's performance over time. The yellow dots represent monthly closing prices, while the orange trendline shows a slight overall downward trajectory. This analysis indicates that despite some months showing higher stock prices, the general trend over the four-year period is a gradual decline. Notably, there are seasonal patterns where stock prices tend to rise from late August to mid-October each year. This recurring pattern suggests that during these months, Tesla's stock experiences an upward trend, which investors might find optimal for selling their holdings to maximize profits. Conversely, the downward trendline implies that, outside of these peak months, the stock price faces consistent downward pressure. This information is crucial for investors in strategizing their buying and selling decisions, as it highlights both the periods of potential gains and the overall cautious outlook for long-term investment in Tesla's stocks.

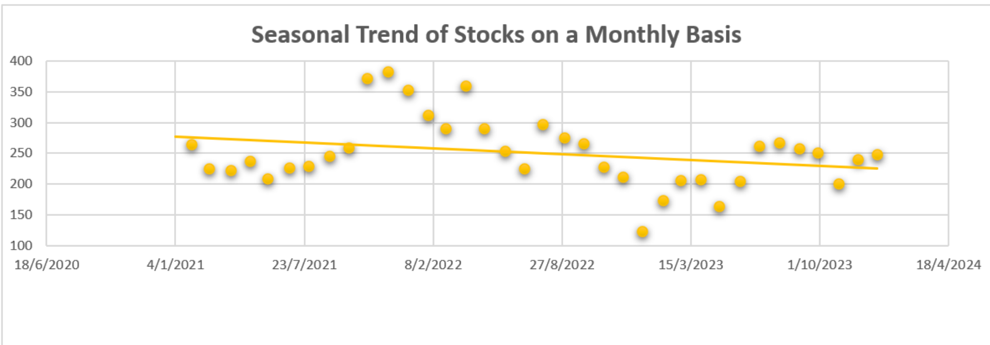

Figure 8. Tesla's stock performance across different time frames.

Volatility and Risk

The volatility analysis for Tesla's stock, depicted in the provided graph, shows significant fluctuations in daily price changes over a 100-day period from January 2021 to December 2023. The yellow line represents historical volatility, revealing periods of both high and low variability. Notably, there are peaks around mid-2021 and early 2022, indicating times of increased uncertainty and larger price swings. This heightened volatility implies a higher risk for investors, as the stock's value can change dramatically in a short period. For those considering investing in Tesla, these findings underscore the importance of implementing robust risk management strategies, such as diversifying portfolios or using stop-loss orders, to mitigate potential losses. The analysis suggests that while Tesla's stock can offer substantial returns, it also carries significant risk, making it potentially unsuitable for risk-averse investors.

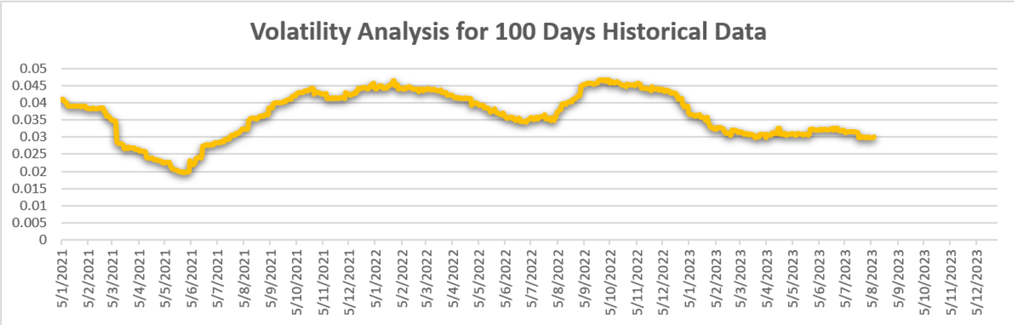

Figure 9. Tesla's stock volatility analysis for 100 days.

Predictive Model Results

Model Performance

The Long Short-Term Memory (LSTM) regression model demonstrated exceptional performance in predicting Tesla's stock prices for January 2024. An R-squared value of 1.00 signifies that the model successfully explained all the variance in the test data, indicating an almost perfect fit. This level of precision is rare in stock price prediction, suggesting that the LSTM model is highly effective in capturing the underlying patterns and trends in the historical stock data. The low Mean Absolute Error (MAE) of 2.61 and Mean Squared Error (MSE) of 12.93 further support the model's accuracy, as these metrics reflect minimal deviation between the predicted and actual stock prices. These results underscore the model's robustness in forecasting short-term stock price movements, making it a valuable tool for investors and analysts.

Despite these impressive results, the model's high Custom Accuracy of 86.75% reveals that there is still room for refinement. While an accuracy rate above 85% is commendable, the remaining margin indicates that the model occasionally deviates from the actual stock prices. This discrepancy could be due to factors such as market volatility, unanticipated news events, or inherent limitations in the dataset used for training the model. Future improvements could involve incorporating more comprehensive data, such as macroeconomic indicators or sentiment analysis from news and social media, to enhance the model's predictive power. Additionally, fine-tuning the model's hyperparameters and experimenting with different network architectures might further boost its performance, ensuring even more reliable and precise stock price predictions.

```
Mean Absolute Error (MAE): 2.61
Mean Squared Error (MSE): 12.93
R-squared (R2): 1.00
Custom Accuracy: 86.75%
```

Figure 10. The output of MAE, MSE, R2 and Custom Accuracy in python.

Implications of Predictions

The model's prediction of a significant downturn in Tesla's stock prices by 29% in January 2024 is a critical finding that aligns with the observed trend of declining stock prices since 2021. This forecast suggests that investors should exercise caution and consider holding onto their existing stocks rather than making new investments in Tesla during this period. The substantial predicted decline underscores the need for investors to incorporate both historical trends and model forecasts into their decision-making processes. By doing so, they can better manage risk and make more informed investment choices. The model's ability to predict such a downturn highlights the value of advanced predictive analytics in navigating the volatile stock market landscape.

```
Predicted percent change for January 2023: 29.00%
```

Figure 11. The output of Finding Percent Difference in Predicted Prices for January 2024.

```
if not predicted_df.empty:
    # Calculate percent change
    predicted_prices_jan_2023 = predicted_df['Predicted Close'].values
    percent_change_jan_2023 = ((predicted_prices_jan_2023[-1] - predicted_prices_jan_2023[0]) / predicted_prices_jan_2023[0]) *

    print(f"Predicted percent change for January 2023: {percent_change_jan_2023:.2f}%")
else:
    print("No data available for January 2023 in the predictions.")
```

Figure 12. Percentage change.

This code checks if the DataFrame predicted_df is not empty, ensuring that there are prediction data points available for January 2023. It then calculates the percent change in predicted stock prices for January 2023 by taking the difference between the last and first predicted prices of the month, divided by the first predicted price, and multiplying by 100 to get the percentage. This calculation provides a quantitative measure of the predicted price movement, offering clear insights into the expected trend. By printing the predicted percent change, investors gain a tangible understanding of the model's forecast, reinforcing the importance of cautious investment strategies during periods of anticipated downturns.

COMPARATIVE ANALYSIS OF TOOLS FOR DATA ANALYSIS AND VISUALIZATION

During this project, Microsoft Excel, Power BI, and Python (using Jupyter Notebook) were utilized to analyse and visualize Tesla's stock performance. Each tool provided unique features and capabilities, underscoring the necessity of a multi-faceted approach for comprehensive data analysis.

Microsoft Excel

Microsoft Excel was instrumental in the initial stages of the project. Its intuitive interface and comprehensive suite of functions made it ideal for early data exploration and straightforward visualizations. Excel's pivot tables and conditional formatting allowed for quick summarization and identification of key trends within the data. Basic statistical analysis and simple chart creation were efficiently handled by Excel. However, its limitations became evident when dealing with larger datasets and more complex statistical calculations. Excel's performance tends to degrade with increased data volume, and its advanced analytical capabilities are limited compared to specialized software. Therefore, while Excel was effective for initial data exploration, it necessitated the use of more robust tools for deeper analysis.

Power BI

Power BI emerged as a crucial tool for creating interactive and dynamic visualizations. It excelled in handling large datasets and integrating with various data sources, enabling real-time data updates and interactive exploration. Power BI's advanced features, such as drill-downs, slicers, and responsive dashboards, provided a detailed view of the data, facilitating a deeper understanding of underlying patterns and trends. Despite its strengths, Power BI has some limitations in terms of visual customization. Certain visual elements required additional effort to implement, and not all desired features were readily available. Nonetheless, its ability to produce advanced statistical analyses and dynamic visual representations made it an invaluable tool for this project.

Python (Jupyter Notebook)

Python, especially within the Jupyter Notebook environment, offered unparalleled flexibility and analytical power. The use of libraries such as Pandas for data manipulation, Matplotlib and Seaborn for visualization, and Scikit-learn for machine learning, provided a comprehensive framework for advanced data analysis. Python's ability to perform complex statistical calculations, create detailed and customized visualizations, and integrate machine learning models allowed for sophisticated analysis that other tools could not achieve. For instance, the LSTM regression model implemented in Python provided accurate forecasts of Tesla's stock prices, leveraging the advanced capabilities of Python's machine learning libraries. While the visual quality of Matplotlib may not match that of Power BI, Python's strength lies in its customizability and flexibility, allowing for the creation of tailored visualizations and advanced analytical techniques.

Table 1. Comparative Insights

Feature/Tool	Microsoft Excel	Power BI	Python (Jupyter Notebook)
Ease of Use	High	Moderate	Moderate
Data Handling	Moderate (limited by memory)	High (handles large datasets efficiently)	High (scalable with libraries)
Statistical Analysis	Basic	Advanced	Advanced (extensive libraries)
Customization	Limited	Moderate	Extensive
Visualization Quality	Basic to moderate	High, interactive	Highly customizable
Interactivity	Low	High	Moderate (interactive libraries)
Integration	High (Microsoft ecosystem)	High (various data sources)	High (various packages and APIs)

The integration of Microsoft Excel, Power BI, and Python was crucial in achieving a comprehensive analysis of Tesla's stock performance. Excel facilitated initial data exploration and the creation of basic visualizations due to its ease of use and quick setup. Power BI enabled advanced statistical analysis and the creation of interactive and dynamic dashboards, enhancing the depth of data exploration. Python provided the flexibility and power necessary for complex data analysis, custom visualizations, and advanced predictive modeling.

By leveraging the strengths of each tool, the project benefited from a robust and multifaceted analysis. Excel's user-friendly interface and quick data manipulation capabilities, Power BI's interactive and detailed visualizations, and Python's advanced analytical and customization capabilities each contributed uniquely to the project. This multi-tool approach ensured that the analysis was thorough and comprehensive, highlighting the importance of selecting the appropriate tools based on the specific requirements and complexities of the data analysis task at hand. The combination of these tools allowed for a detailed understanding of Tesla's stock performance, enabling informed investment decisions and accurate predictions.

Visualizations and Tables

Excel

Microsoft Excel was utilized for its user-friendly interface and powerful tools for data manipulation and visualization. Excel allowed for quick initial data exploration and the creation of straightforward charts and graphs. Specific analyses performed in Excel included calculating and plotting the 30, 100, and 200 days moving averages to identify long-term trends in Tesla's stock prices. Additionally, seasonal trend analysis on a monthly basis was conducted to uncover recurring patterns. Excel was also used to perform volatility analysis, highlighting the variability in daily stock prices. The use of pivot tables and conditional formatting further helped in summarizing and highlighting key data points efficiently. Despite its limitations with large datasets and complex statistical calculations, Excel provided a solid foundation for preliminary analysis.

Figure 13. The main page of the Stock Market Analysis for Tesla Stocks in excel.

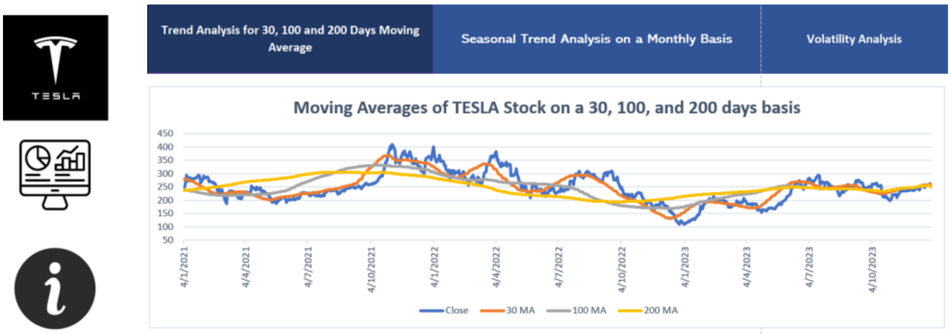

Figure 14. Trend Analysis for 30, 100, and 200 Days Moving Average.

Figure 15. Seasonal Trend Analysis on a Monthly Basis.

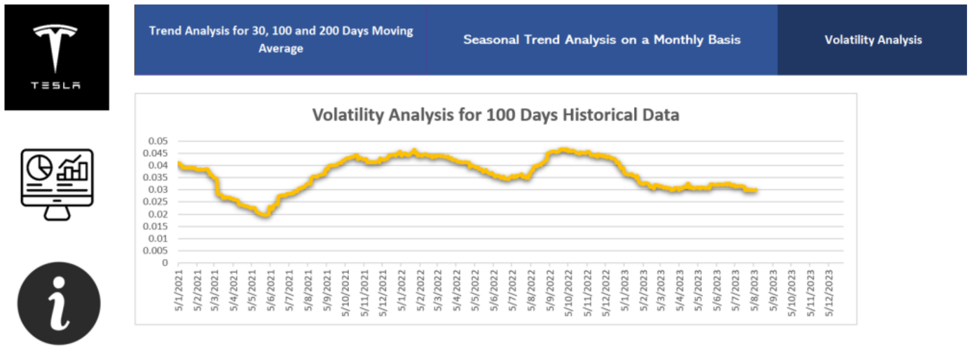

Figure 16. Volatility Analysis.

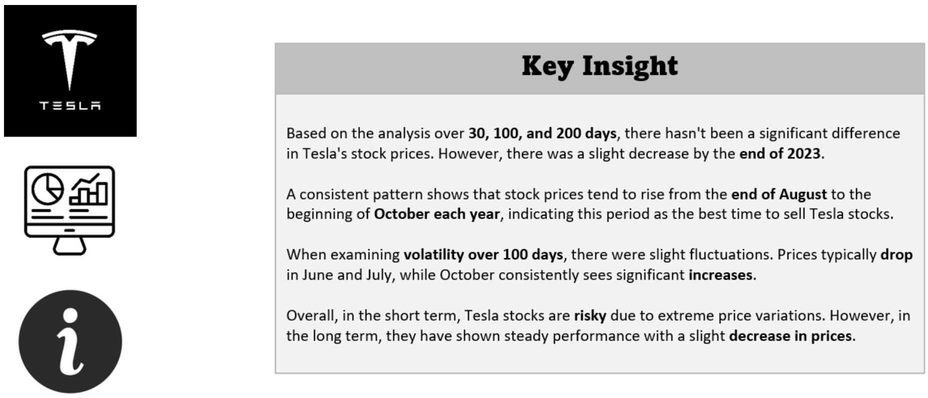

Figure 17. Key insight page.

Power BI

Power BI was employed for its advanced statistical capabilities and ability to create dynamic, interactive visualizations. It facilitated the detailed analysis of Tesla's stock performance through interactive dashboards that allowed for in-depth exploration of various aspects of the data. Power BI handled large datasets efficiently and enabled real-time updates from multiple data sources. Key analyses conducted in Power BI included a comprehensive Tesla Stock Prediction Dashboard, which provided insights into opening and closing prices, highest and lowest prices, and stock performance across daily, monthly, quarterly, and yearly time frames. Power BI also supported simple moving average analysis and detailed seasonal trend and volatility analyses. The interactive features, such as drill-downs and slicers, allowed users to explore the data dynamically, though some visual customizations required additional workarounds.

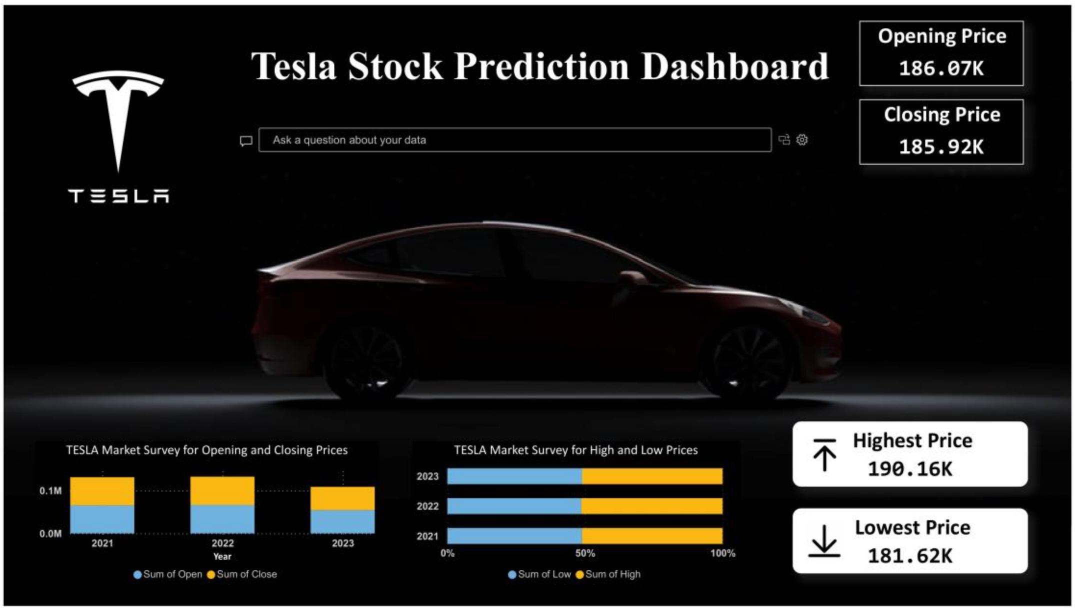

Figure 18. Tesla Stock Prediction Dashboard.

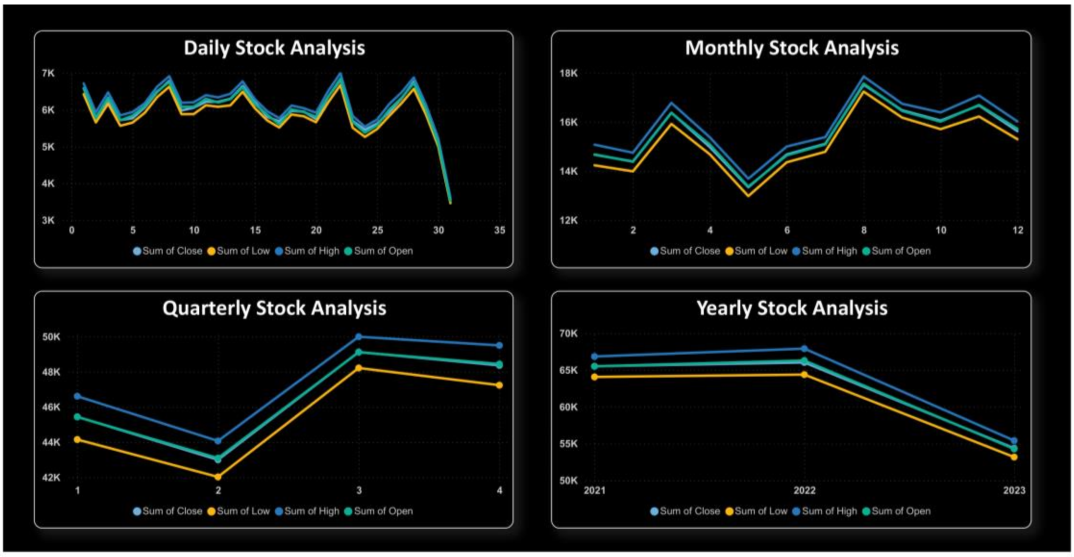

Figure 19. The comprehensive analysis of stock performance over different time frames.

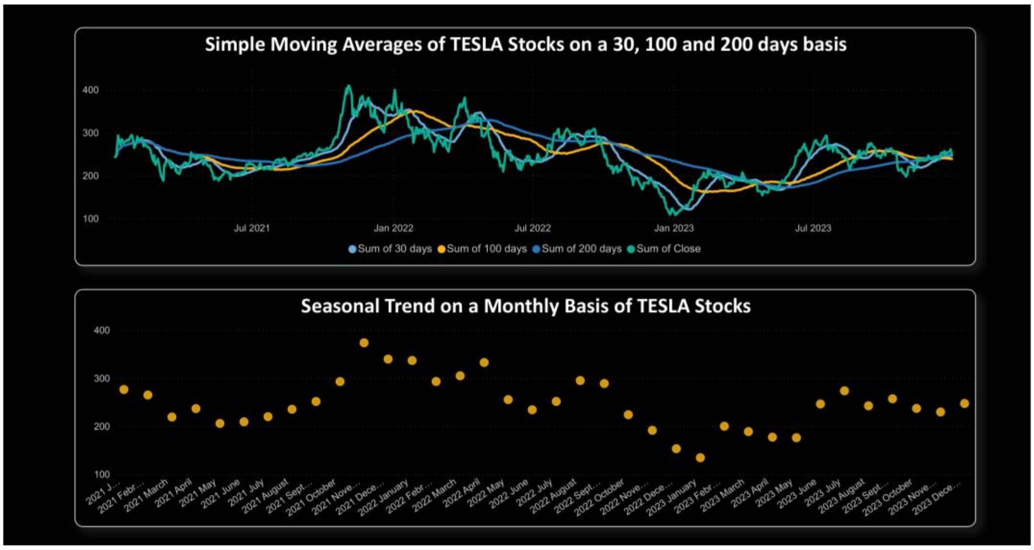

Figure 20. Simple Moving Average Analysis & Seasonal Trend Analysis.

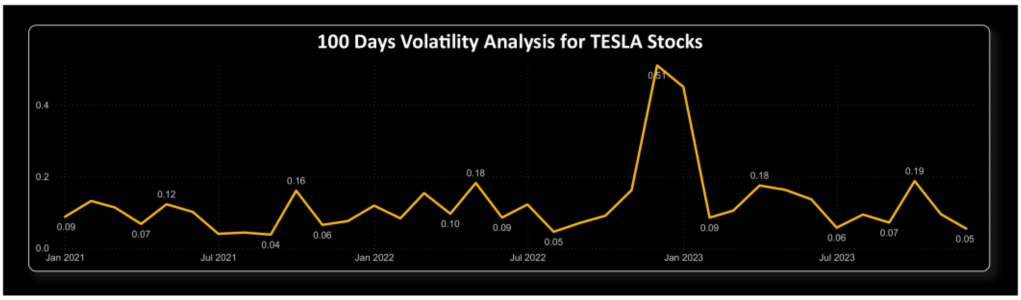

Figure 21. The volatility analysis over 100 days.

Jupyter Notebook

Jupyter Notebook, with Python and its libraries (Pandas, NumPy, Matplotlib, Seaborn, Plotly, and Scikit-learn), was leveraged for its extensive customization capabilities and flexibility in data analysis and visualization. Python's powerful libraries enabled detailed data preprocessing, trend analysis using simple moving averages, and volatility analysis with Bollinger Bands. Time series plots and seasonal trend analyses were also performed to uncover underlying patterns in the stock data. The use of Jupyter Notebook was crucial for implementing the LSTM regression model, which predicted Tesla's stock prices for January 2024. The model's performance metrics, including MAE, MSE, R-squared, and Custom Accuracy, were calculated to evaluate prediction accuracy. Python's integration with machine learning frameworks allowed for advanced predictive modeling, showcasing the tool's capability to handle complex analytical tasks beyond the reach of traditional software like Excel and Power BI.

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
2021-01-04 00:00:00-05:00	239.820007	248.163330	239.063339	243.256668	145914600	0.0	0.0
2021-01-05 00:00:00-05:00	241.220001	246.946671	239.733337	245.036667	96735600	0.0	0.0
2021-01-06 00:00:00-05:00	252.830002	258.000000	249.699997	251.993332	134100000	0.0	0.0
2021-01-07 00:00:00-05:00	259.209991	272.329987	258.399994	272.013336	154496700	0.0	0.0
2021-01-08 00:00:00-05:00	285.333344	294.829987	279.463318	293.339996	225166500	0.0	0.0

Figure 22. Tesla historical data that have been fetched in.

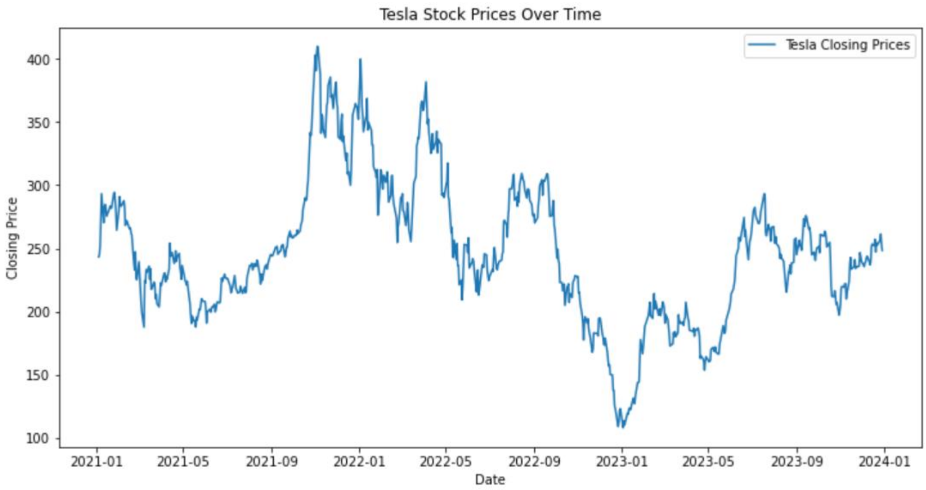

Figure 23. The plot showing the closing prices over time.

Figure 24. The moving average analysis charts.

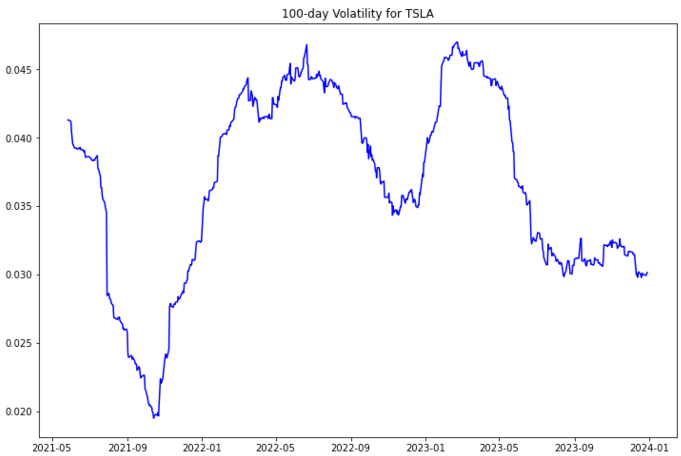

Figure 25. The Bollinger Bands analysis.

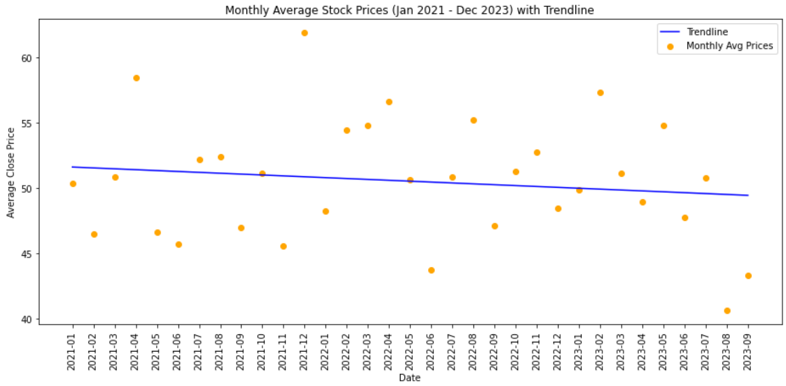

Figure 26. The seasonal trend analysis.

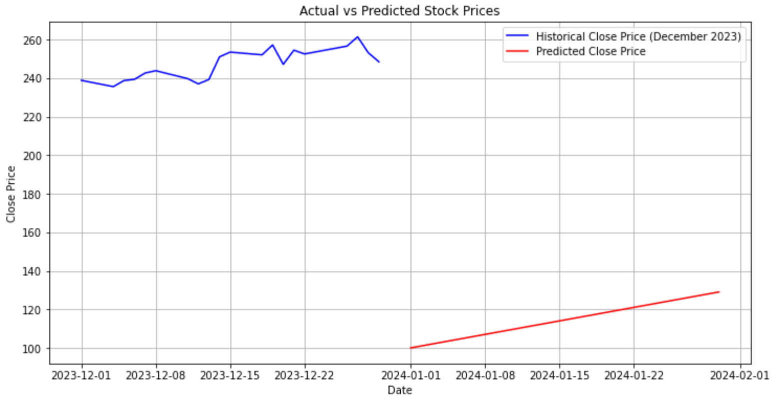

Figure 27. The predictions made by the LSTM model and the percent difference calculations.

```
Mean Absolute Error (MAE): 2.61
Mean Squared Error (MSE): 12.93
R-squared (R2): 1.00
Custom Accuracy: 86.75%
```

Figure 28. The MAE, MSE, R-squared, and Custom Accuracy metrics.

CONCLUSION AND RECOMMENDATION

This project has demonstrated the significant value of integrating data visualization and advanced predictive modeling techniques to analyze and forecast Tesla's stock performance. By leveraging historical data sourced from Yahoo Finance, we have provided a comprehensive analysis of Tesla's stock trends, volatility, and seasonal patterns. The use of various analytical tools—Microsoft Excel, Power BI, and Python—has allowed for a multifaceted approach, combining initial data exploration, dynamic visualizations, and sophisticated predictive modeling.

Our exploratory data analysis revealed key insights into Tesla's stock behavior, such as the impact of market forces and company-specific events on stock prices. The trend analysis using Simple Moving Averages (SMA) highlighted the divergence between short-term and long-term trends, presenting potential opportunities for investors. Seasonal trend analysis identified recurring patterns, suggesting optimal times for buying and selling. Volatility analysis underscored the importance of understanding the risk associated with Tesla's stock, given its significant price fluctuations.

The predictive modeling phase employed an LSTM (Long Short-Term Memory) regression algorithm, which demonstrated exceptional accuracy in forecasting Tesla's stock prices. The model's high R-squared value and low error metrics indicated its robustness in capturing historical patterns and predicting future trends. However, the model's prediction of a significant downturn in January 2024 suggests a cautious approach for investors during this period.

Based on these findings, several key recommendations can be made for investors. Given the identified volatility in Tesla's stock, diversification and robust risk management strategies are crucial. Implementing stop-loss orders and hedging can help protect against significant losses during periods of high volatility. Seasonal trend analysis suggests that late August to mid-October is a period of consistent upward movement for Tesla's stock, presenting potential selling opportunities to maximize profits. Conversely, investors should be cautious about investing heavily outside these peak months due to the overall downward trend. The success of the LSTM model in forecasting stock prices underscores the importance of incorporating advanced predictive analytics into investment strategies. Investors and analysts should leverage machine learning models to gain insights into potential future price movements and make more informed decisions. Finally, continuous monitoring of market conditions and regular updates to predictive models are essential for maintaining their accuracy and relevance. Investors should stay informed about macroeconomic indicators, company performance, and market sentiment to adapt their strategies accordingly.

In conclusion, this project highlights the critical role of data visualization and predictive modeling in understanding and forecasting stock performance. By utilizing a comprehensive dataset and employing advanced analytical tools, we have provided valuable insights into Tesla's stock behavior. These findings can aid investors in making informed decisions, managing risk, and optimizing their investment strategies. As the financial market evolves, the integration of data-driven approaches will remain essential for navigating the complexities of stock investment.

REFERENCES

Babazhanov, Z., Suleyeva, M., Zhumatov, M., Kenenbay, A., & Syrbayeva, D. (2023). Tesla's financial performance and stock price success in 2020. *Digital repository of KAZGUU UniversityTesla's financial performance and stock price succe*, 2-3.

Becker, L. T., & Gould, E. M. (2019). Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Serials Review*, 184-188.

Echrigui, R., & Hamiche, M. (2023). Optimizing LSTM Models for EUR/USD Prediction in the context of reducing energy consumption: An Analysis of Mean Squared Error, Mean Absolute Error and R-Squared. *EDP Sciences*, 6-19.

Harmsen, S. M., Chang, Y.-H. H., & Hattrup, S. J. (2016). Simple Moving Average: A Method of Reporting Evolving Complication Rates. *Slack Journals*, 28-30.

Jagwani, J., Gupta, M., Sachdeva, H., & Singhal, A. (2018). Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend. *Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 462-467.

Li, W., Kiaghadi, A., & Dawson, C. (2021). Exploring the best sequence LSTM modeling architecture for flood prediction. *Neural Computing and Applications*, 5571–5580.

Pimentel, J. F., Murta, L., Braganholo, V., & Freire, J. (2021). Understanding and improving the quality and reproducibility of Jupyter notebooks. *Empirical Software Engineering* , 2-26.

Qingren, W. (2023). ANALYSIS OF TESLA COMPANY'S OPERATIONS BASED ON PANEL DATA. *SADI International Journal of Social Science and Humanities*, 1-10.

Schreiber, B. A., & Gregersen, E. (2024). Tesla, Inc. *Britannica Money*, 1-3.

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 2-3.