Time Series Forecasting of Google's Stock Prices

Statistical Learning
Professor: Gianluca MASTRANTONIO

Author: Ghassane BEN EL AATTAR

Politecnico di Torino

Contents

1 Introduction

This report presents a comprehensive analysis of the closing stock prices of Google (GOOG) over a specific time period, from 2013 to 2017.

The stock market is a complex system influenced by a myriad of factors, both internal corporate events and external global events. Analyzing stock prices can provide insights into the company's performance, market trends, and potential future movements. The primary objective of this analysis is to understand the historical behavior of Google's stock prices and to forecast its future movements.

Forecasting stock prices is a pivotal task in finance, aiding investors in making informed decisions. Specifically, closing prices are of significant interest as they represent the consensus of value for a stock at the end of a trading day, encapsulating the day's volatility and investor sentiment.

By delving deep into the patterns, seasonality, and other characteristics of Google's closing prices, this report aims to provide a robust foundation for predictive modeling and strategic investment planning.

2 Time Series

A time series is a sequence of data points, typically consisting of successive measurements made over a time interval.

Mathematically, a time series $\{X_t\}$ can be represented as:

$$X_t = f(t, S_t, \epsilon_t) \tag{1}$$

where t is the time index, S_t is the systematic structure of the series, and ϵ_t is the random error term. The systematic structure can further be decomposed into trend, seasonality, and cyclic components.

The essence of time series analysis lies in decomposing the series to understand its underlying patterns and then modeling these patterns for forecasting.

In the context of our task, the time series represents the closing stock prices of Google. Analyzing this time series can unveil patterns in stock price movements, the influence of external events, and the overall trajectory of the stock's value.

By modeling these patterns, we aim to forecast future stock prices, providing valuable insights for investment strategies and decision-making processes.

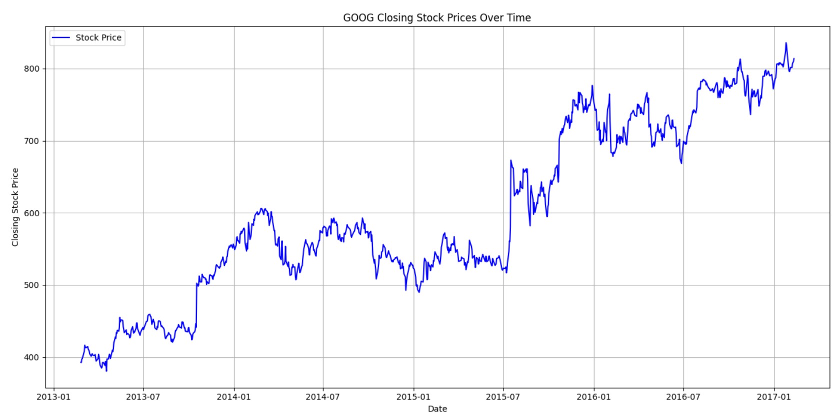

Figure 1: Google Stock Price Over Time

2.1 Time Series Decomposition

Time series decomposition is a foundational method in time series analysis, aiming to deconstruct a time series into its constituent components. Most time series data, including stock prices, can be described using three systematic components:

- **Trend**: Represents the underlying trajectory of the data over a long period. In the context of Google's closing stock prices, the trend might capture the company's overall growth trajectory, influenced by factors such as increasing market share, product innovations, or global expansion.

- **Seasonality**: Captures regular and predictable changes that recur at fixed intervals. For stock prices, seasonality might manifest in patterns related to quarterly earnings reports, annual sales events, or other regular corporate activities.

- **Cycles**: Unlike seasonality, cycles represent medium-term changes caused by circumstances or events that don't have a fixed repetition period. In stock prices, cyclical components might capture broader economic cycles, industry-wide shifts, or market sentiments that don't adhere to a fixed calendar schedule.

In addition to these systematic components, there's a non-systematic component called **noise**. This captures random fluctuations in the data that can't be attributed to the trend, seasonality, or cycles. Noise might result from unforeseen events, such as sudden market shocks, unexpected corporate news, or other unpredictable influences.

In the task of analyzing Google's closing stock prices, decomposition allows us to isolate these components, providing a clearer understanding of the underlying patterns and behaviors in the data. By understanding each component separately, we can more effectively model and forecast future stock prices. For instance, recognizing a consistent upward trend can inform long-term investment strategies, while understanding seasonality can guide short-term trading decisions. Similarly, identifying cyclical downturns might prompt more cautious investment approaches, even if the overall trend is positive.

$$Y_t = T_t + S_t + C_t + R_t \tag{2}$$

where Y_t is the observed series, T_t is the trend component, S_t is the seasonal component, C_t is the cyclical component, and R_t is the residual or noise component.

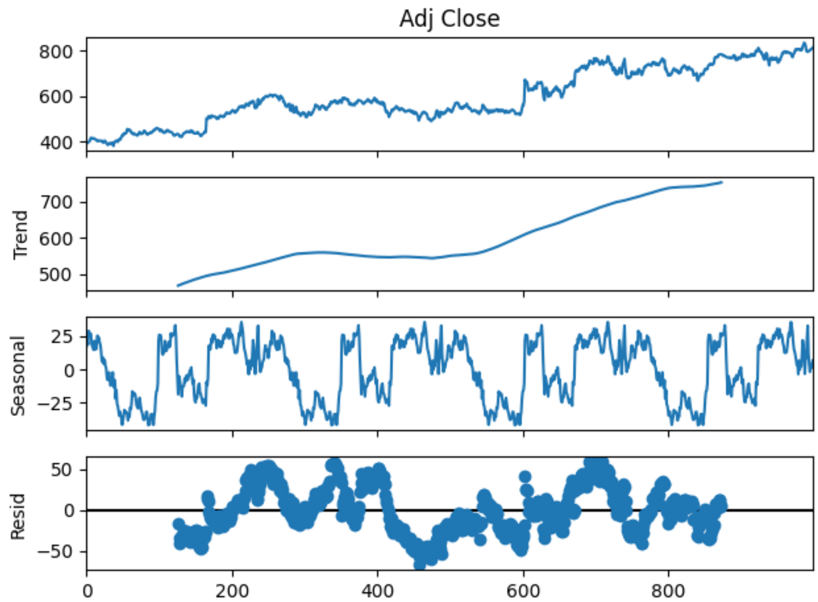

Figure 2: Time Series Decomposition of Google Closing Stock Prices.

3 Trend and Seasonality Analysis

3.1 Trend Analysis of Google's Stock Prices

Upon analyzing the trend plot for Google's closing stock prices, several key observations emerge:

1. There is a **distinct upward trend** over the years, signifying a consistent increase in the value of Google's stock. This is indicative of the company's robust performance and positive market reception.

2. Some intervals showcase a **flattened trend**, suggesting periods of stagnation or minimal growth. Notably, these phases are transient and are succeeded by renewed upward trajectories.

3. Occasional **dips in the trend line** are evident, pointing to temporary declines in stock prices. Crucially, these downturns are short-lived and are invariably followed by a return to the prevailing upward trend.

In essence, the trend plot underscores a predominant theme: Google's stock prices have largely witnessed an upward trajectory over the years, punctuated by brief spells of slower growth or minor declines.

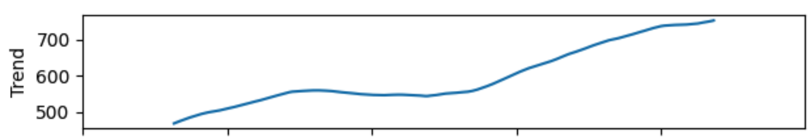

Figure 3: Trend of Google's Stock Prices

3.2 Seasonality Analysis

As described before, seasonality is a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year. These fluctuations are often tied to the time of year, such as increased sales of winter clothing during winter months.

Let's now analyse seasonality for Google's stock prices, from the plots below.

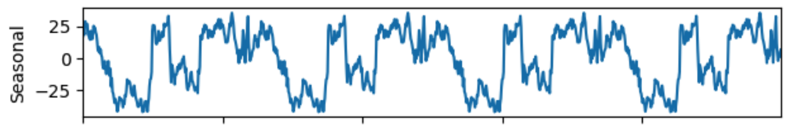

Figure 4: Seasonality Component of Google's Stock Prices

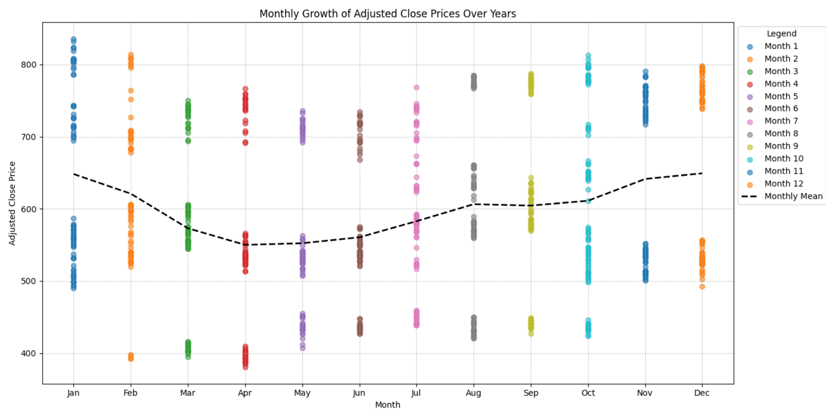

Figure 5: Monthly Growth of Google's Stock Prices of the Years

Analyzing the seasonality plot for Google's closing stock prices yields several noteworthy insights:

- A **recurring pattern** in the data is discernible, hinting at some degree of seasonality. Nevertheless, the amplitude of these seasonal oscillations is modest when juxtaposed with the overarching trend. This observation underscores the trend as the primary driving force in this time series.

- The seasonality lacks a **uniform rhythm**. This inconsistency can be attributed to the intricate nature of stock prices, which are swayed by a plethora of factors, many transcending seasonal influences.

- Certain intervals exhibit a **heightened seasonal component**. Such pronounced seasonality could align with specific events or annual periods that wield a marked influence on Google's stock valuation, such as product launches, annual sales events, or industry-wide phenomena.

To encapsulate, the seasonality plot intimates the presence of seasonal patterns in Google's stock prices. However, these patterns exhibit variability and their magnitude is overshadowed by the dominant trend. This infers that, while seasonality might contribute to stock price fluctuations, a myriad of other determinants exert a more pronounced influence.

Let's now further confirm our assumptions by analysing the polar seasonal plot generated from the dataset, which is a way of visualizing the average monthly closing stock prices for each year. It can potentially help identify any seasonal patterns in the data.

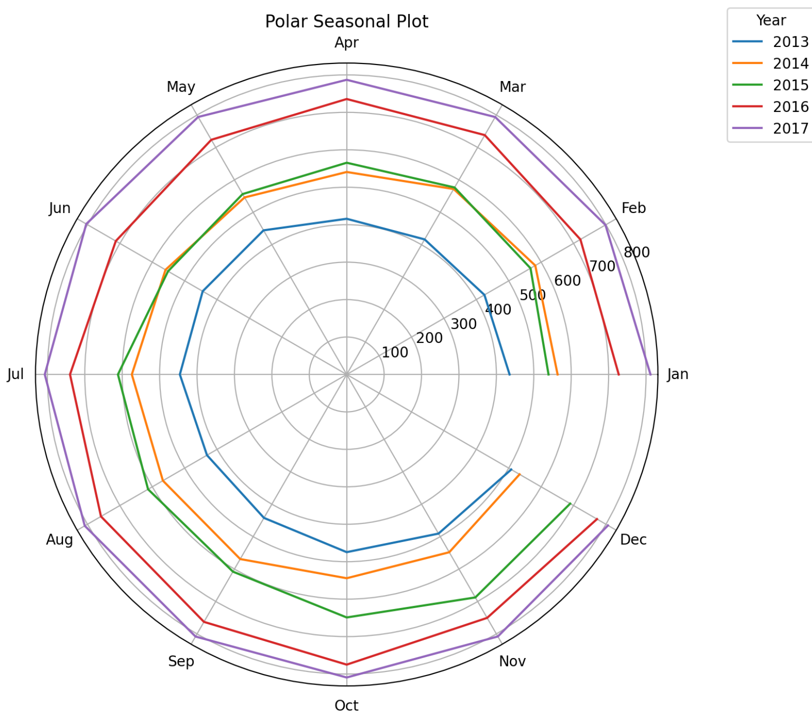

Figure 6: Polar Seasonal Plot of Google's Stock Prices Over the Years

As it's possible to see, the polar seasonal plot offers a unique perspective on the cyclical behavior of Google's closing stock prices:

- Each trajectory in the plot corresponds to a distinct year, with individual points on the trajectory denoting the average closing stock price for each month within that year.

- The circular configuration of the plot symbolizes the cyclical progression of a year. This design ensures that the commencement and culmination of the year (i.e., January and December) are contiguous, illustrating the seamless transition from one year to its successor.

- The radial magnitude from the circle's epicenter to a specific point on a trajectory is indicative of the stock price. Consequently, an augmented radial distance signifies an elevated stock price.

- Despite the plot's illustrative nature, discerning definitive patterns is challenging due to the superimposition of trajectories. A preliminary observation suggests an absence of consistent peaks or troughs during identical months across different years. This insinuates that Google's stock price might not be profoundly influenced by a robust seasonal component on a monthly scale.

In essence, while the polar seasonal plot provides a cyclical overview of Google's stock prices, it intimates a potential lack of strong monthly seasonality. Although the seasonality appears to be weaker than the trend factor, stationary models require the removal of the latter and thus making seasonality, although minimal, problematic for forecasting.

4 Stationarity

Stationarity is a pivotal concept in time series analysis. A time series is said to be stationary if its statistical properties, such as mean, variance, and autocorrelation, remain constant over time. Mathematically, for a stationary time series, the joint probability distribution of $X_t, X_{t+1}, \ldots, X_{t+n}$ remains unchanged for any lag n. In simpler terms, a stationary series does not exhibit trends, seasonality, or cycles that might change over time.

- **Relevance to Forecasting**: Stationarity is crucial because many forecasting methods, including ARIMA and its variants, assume that the time series is stationary or can be transformed to become stationary. Stationary time series are inherently more predictable as their future statistical properties are not different from their past.

- **Application to Google's Stock Prices**: In the context of our task with Google's closing stock prices, ensuring stationarity is essential for accurate forecasting. The original series exhibited non-stationarity, necessitating a transformation. We applied differencing to the series, stabilizing its statistical properties before modeling. This transformation was critical in enhancing the reliability of our forecasts.

4.1 Autocorrelation Function (ACF) Plot Analysis

Autocorrelation, in time series analysis, refers to the correlation of a series with its own past values. It measures the linear predictability of the series at time t using its value at time $t - k$, where k is the lag. The Autocorrelation Function (ACF) provides a systematic way to visualize and quantify this correlation across various lags.

Upon differencing the series to achieve stationarity, we analyze the ACF plot. This plot offers a graphical representation of the time series' autocorrelation across a range of lags.

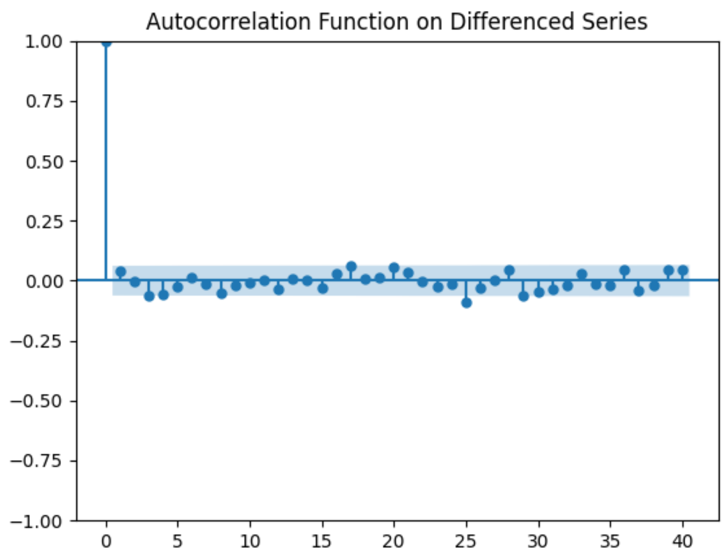

Figure 7: ACF Plot for Differenced Series

The ACF plot for our differenced series shows a sharp drop after the first lag and oscillates around zero thereafter. This rapid decline towards zero suggests that the differencing has been effective in removing autocorrelation from the data, which is a characteristic of a stationary time series. Subsequent lags beyond the first show no significant correlation, confirming the lack of systematic pattern like trend or seasonality in the differenced series.

The ACF plot indicates that the transformed data is suitable for modeling with ARIMA-type statistical methods, which assume stationarity as a prerequisite for forecasting.

5 Time Series Forecasting Models

5.1 ARIMA (AutoRegressive Integrated Moving Average)

ARIMA, an acronym for AutoRegressive Integrated Moving Average, is a prominent time series forecasting method that blends the Autoregressive (AR) model, Moving Average (MA) model, and differencing to induce stationarity (Integrated).

5.1.1 Components of ARIMA

- **AR (AutoRegressive)**: This component captures the influence of preceding values on the current value in a time series. It posits that current observations are linearly dependent on their previous values with a degree of correlation. The AR part of the ARIMA model is specified by the equation:

  $$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$$

  Here, $\phi_1, \phi_2, \ldots, \phi_p$ are the coefficients that measure the impact of past lags p on the current value y_t, and ε_t is the error term, often assumed to be white noise.

- **I (Integrated)**: This denotes the differencing process, which is the subtraction of the current and previous observations to achieve stationarity. The 'Integrated' aspect of ARIMA reflects the number of differencing operations required to stabilize the mean of the time series, denoted by d. A non-stationary series can often be made stationary through differencing, thereby eliminating trends and seasonality.

- **MA (Moving Average)**: This component models the relationship between the current observation and the residual errors from a moving average model applied to past observations. It provides a way to incorporate the 'shock' effects from previous forecast errors into the current prediction. The MA part is formulated as:

  $$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

In this equation, $\theta_1, \theta_2, \ldots, \theta_q$ are the coefficients of the model, and q represents the number of lagged forecast errors in the model.

5.2 ETS (Error, Trend, Seasonality)

The ETS model stands for Error, Trend, and Seasonality, and it is an alternative to ARIMA for forecasting time series data. This approach decomposes a time series into three interpretable components, each of which can be modeled either additively or multiplicatively, providing a flexible framework to capture a variety of time series behaviors.

5.2.1 Mathematical Formulation of ETS

- **Error (E)**: The error component represents the random fluctuations in the time series that cannot be explained by the model's trend or seasonal components. Mathematically, it can be represented as:

  $$E_t = Y_t - (T_{t-1} + S_{t-1})$$

  where E_t is the error at time t, Y_t is the actual observation at time t, and T_{t-1} and S_{t-1} are the trend and seasonal components from the previous period, respectively. The error term can follow an additive model where errors are roughly constant over time or a multiplicative model where errors grow in proportion to the level of the series.

- **Trend (T)**: The trend component captures the long-term progression of the series, indicating whether the data is increasing, decreasing, or stable over time. The trend can be none (indicating a flat trend), additive (linear), or multiplicative (exponential), and is given by:

  $$T_t = T_{t-1} + \beta E_{t-1}$$

  for an additive model, where β is the trend smoothing parameter, and by:

  $$T_t = T_{t-1} \times (1 + \beta E_{t-1})$$

  for a multiplicative model.

- **Seasonality (S)**: This component reflects the seasonal fluctuations in the series—patterns that repeat at regular intervals, such as daily, monthly, or quarterly. The seasonality can also be modeled additively or multiplicatively:

$$S_t = S_{t-L} + \gamma E_{t-1}$$

for an additive model, where L is the length of the seasonal period and γ is the seasonal smoothing parameter, and by:

$$S_t = S_{t-L} \times (1 + \gamma E_{t-1})$$

for a multiplicative model.

The ETS model is often represented as $ETS(Error, Trend, Seasonality)$, with each component specified by its type—additive (A) or multiplicative (M). For instance, an $ETS(A, A, M)$ model would have an additive error component, an additive trend, and a multiplicative seasonality.

In the context of Google's stock prices forecasting, the ETS model provides a framework to capture patterns and make forecasts considering the error, trend, and potential seasonal patterns. Understanding these components separately allows for a nuanced interpretation of the time series and more informed decision-making. For example, a clear upward trend in the data can signal a long-term investment opportunity, while a strong seasonal pattern may suggest optimal timing for buying or selling.

5.3 SARIMA Model

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a sophisticated time series forecasting method used to capture and model complex seasonality patterns within data. It is particularly valuable when dealing with time series data that exhibits regular, repeating patterns at fixed intervals, such as daily, monthly, or yearly seasonality.

The SARIMA model combines three fundamental components: Seasonal, Autoregressive (AR), and Moving Average (MA). Each of these components plays a crucial role in modeling the underlying patterns within the data.

5.3.1 Seasonal Component

The seasonal component of the SARIMA model introduces seasonality into the time series analysis. It takes into account the periodic patterns that occur at regular intervals. The notation SARIMA(p, d, q)(P, D, Q, s) specifies the seasonal component, where:

- p represents the number of non-seasonal autoregressive (AR) terms.
- d denotes the order of non-seasonal differencing necessary to make the time series stationary.
- q signifies the number of non-seasonal moving average (MA) terms.
- P indicates the number of seasonal AR terms.
- D represents the order of seasonal differencing required to remove seasonal trends.
- Q stands for the number of seasonal MA terms.
- s represents the number of time steps in each seasonal period (e.g., 12 for monthly data with yearly seasonality).

5.3.2 Mathematical Formulation

The mathematical formulation of the SARIMA model is expressed as follows:

$$(1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p)(1-\Phi_1 L^s-\Phi_2 L^{2s}-\ldots-\Phi_P L^{Ps})(1-L^s)^D(1-L)^d y_t =$$

$$(1+\theta_1 L+\theta_2 L^2+\ldots+\theta_q L^q)(1+\Theta_1 L^s+\Theta_2 L^{2s}+\ldots+\Theta_Q L^{Qs})\epsilon_t$$

In this equation:

- y_t represents the observed time series at time t.
- L denotes the lag operator, where $L^s y_t$ refers to y_{t-s}.
- $\phi_1, \phi_2, \ldots, \phi_p$ are the non-seasonal AR coefficients.
- $\Phi_1, \Phi_2, \ldots, \Phi_P$ are the seasonal AR coefficients.
- d and D are the non-seasonal and seasonal differencing orders, respectively.
- $\theta_1, \theta_2, \ldots, \theta_q$ represent the non-seasonal MA coefficients.
- $\Theta_1, \Theta_2, \ldots, \Theta_Q$ are the seasonal MA coefficients.
- ϵ_t represents white noise.

The SARIMA model aims to capture both short-term and seasonal dependencies within the time series data. It does so by considering the historical values, seasonal patterns, and differencing operations. By utilizing these components, SARIMA provides a powerful framework for forecasting and analyzing time series data with intricate seasonal trends.

6 Forecasting

Since seasonality could not be observed in the data, we'll focus the rest of the analysis on ARIMA, SARIMA and ETS, comparing their behavior and performances in terms of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

6.1 ARIMA Model Selection

The ARIMA(3, 1, 0) model was carefully chosen after a comprehensive examination of the data's statistical properties and the application of model selection criteria. The ARIMA model, abbreviated as AutoRegressive Integrated Moving Average, is specified by three critical parameters: p, d, and q. These parameters define the autoregressive, differencing, and moving average components of the model.

For our dataset, the initial step was to ensure stationarity, which required differencing the data once ($d = 1$). This stationarity was confirmed by a significant reduction in autocorrelation, as observed in the Autocorrelation Function (ACF) plot. Subsequently, we utilized the auto ARIMA function from Python's pmdarima library to automate the ARIMA modeling process. This function systematically explores different combinations of p and q and selects the optimal model based on chosen information criteria, typically the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).

In our case, the ARIMA(3, 1, 0) model was chosen due to its lower AIC value compared to other models, indicating a better fit to the data while penalizing for model complexity. The inclusion of three autoregressive terms ($p = 3$) suggests a correlation between the current value and its three immediate past values. The absence of moving average terms ($q = 0$) indicates that the forecast error is not explicitly influenced by past forecast errors. This choice is well-suited for capturing the specific data dynamics.

6.2 SARIMA Model Selection

For the SARIMA model, the parameters (1, 1, 0)(2, 1, 0)[12] were chosen using an iterative process and the Python library, similar to the ARIMA model selection. The SARIMA model, short for Seasonal AutoRegressive Integrated Moving Average, extends ARIMA to account for seasonality patterns in time series data. The parameters (1, 1, 0) represent the non-seasonal components, while (2, 1, 0)[12] represents the seasonal components. Specifically:

- $p = 1$ is the number of non-seasonal autoregressive (AR) terms.
- $d = 1$ is the order of non-seasonal differencing required to make the time series stationary.
- $q = 0$ is the number of non-seasonal moving average (MA) terms.
- $P = 2$ is the number of seasonal AR terms.
- $D = 1$ is the order of seasonal differencing required to remove seasonal trends.
- $Q = 0$ is the number of seasonal MA terms.
- $s = 12$ is the number of time steps in each seasonal period, indicating monthly data with yearly seasonality.

Both the ARIMA and SARIMA models were chosen based on their ability to capture the specific data patterns and dynamics effectively, as indicated by lower information criterion values.

6.3 Forecasting Insights

The following plot shows the results generated by applying these three models to out data.

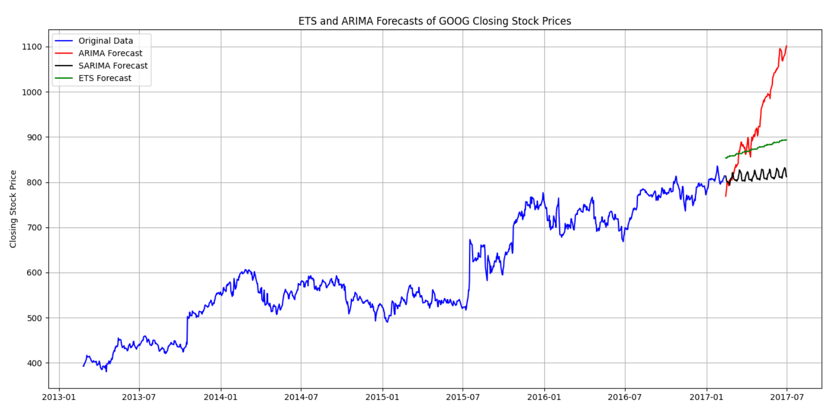

Figure 8: Forecasts of Google's Stock Prices

The ARIMA and SARIMA models, which rely on differencing to render the data stationary, suggest a continuation of the existing trend, as indicated by the red and black forecast lines respectively. The ARIMA forecast shows a more optimistic projection, while the SARIMA forecast implies a slightly more varied future trend, acknowledging potential seasonal effects.

The ETS forecast, depicted in green, adapts to the non-stationarity within the data and forecasts a trend that resembles the original data, although the seasonality does not appear as apparent as the SARIMA forecast.

This divergence may reflect the ETS model's capacity to capture inherent level, trend, and seasonal patterns more effectively.

The contrast in the projections of the ARIMA, SARIMA, and ETS models can be ascribed to their unique assumptions and structural approaches to time series forecasting. While ARIMA focuses on recent past values and requires stationarity, SARIMA extends this by incorporating seasonal differencing to address seasonality. The ETS model, on the other hand, is designed to directly model the data's non-stationary components, potentially providing a more holistic view of the underlying patterns in the stock prices.

6.4 Analysis of Forecasting Errors

To assess the accuracy and reliability of our forecasting models, we computed two standard error metrics: the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). These metrics offer insights into the average magnitude of errors between predicted and observed values.

6.5 Confidence Intervals in Time Series Forecasting

Confidence intervals in time series forecasting estimate the degree of uncertainty associated with forecasted values. They are crucial for understanding the reliability of predictions made by statistical models like ARIMA and SARIMA.

A confidence interval (CI) is a type of interval estimate of a population parameter and is used to indicate the reliability of an estimate. Mathematically, a confidence interval for a parameter is defined by two values: the lower bound (LB) and the upper bound (UB). For a given dataset and a chosen confidence level (commonly 95%), the CI is calculated to predict the range within which the true value of the parameter lies with that level of confidence.

The general form of a confidence interval is given by:

$$CI = \left(\hat{\theta} - Z \times SE, \hat{\theta} + Z \times SE\right)$$

where:

- $\hat{\theta}$ is the point estimate of the parameter (e.g., the mean).
- Z is the Z-score from the standard normal distribution corresponding to the desired confidence level.
- SE is the standard error of the estimate, which measures the statistical accuracy of the estimate.

For a 95% confidence level, the Z-score is typically 1.96, assuming the underlying distribution is normal.

In the ARIMA model, confidence intervals are computed for each forecasted value. For instance, an interval of [-16.22, 22.63] suggests that, with 95% confidence, the actual value will fall within this range. Narrow intervals at the start indicate high certainty in the immediate forecast, while wider intervals as the forecast extends imply increasing uncertainty.

The SARIMA model, which includes seasonal components, shows varying widths in its confidence intervals. Early wide intervals (e.g., [793.68, 832.88]) reflect the model's adjustment to the seasonal patterns, resulting in higher initial uncertainty. Fluctuating widths throughout the forecast period represent the model's dynamic response to seasonal trends.

The width of the confidence intervals in both ARIMA and SARIMA models reflects the level of uncertainty or variability in the forecasted stock prices. It is important to interpret these intervals within the context of inherent stock market volatility and the limitations of the forecasting model. They provide a range of likely future values but cannot guarantee absolute accuracy, especially given the influence of unforeseen market factors.

6.6 Error Metrics Defined

- **Mean Absolute Error (MAE)**: This represents the average of the absolute differences between the forecasted and actual values. It provides a direct measure of the average magnitude of errors without considering their direction. The mathematical formulation is as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

where y_i is the actual value, $\hat{y}_i$ is the forecasted value, and n is the number of data points.

- **Root Mean Squared Error (RMSE)**: This measures the square root of the average of squared differences between forecasted and actual values. It assigns more weight to larger errors than smaller ones, making it sensitive to outliers. The formula is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

6.7 Error Metrics and Interpretation

The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are essential metrics for evaluating the accuracy of our forecasting models. They measure the average magnitude of the errors between the forecasts and the actual observed values.

- **ARIMA Model**:

$$\begin{aligned} \text{MAE} &: 146.990 \\ \text{RMSE} &: 169.479 \end{aligned}$$

- **ETS Model**:

$$\begin{aligned} \text{MAE} &: 87.330 \\ \text{RMSE} &: 88.868 \end{aligned}$$

- **SARIMA Model**:

$$\begin{aligned} \text{MAE} &: 27.303 \\ \text{RMSE} &: 31.687 \end{aligned}$$

The MAE represents a direct measure of the average forecast error magnitude, while the RMSE, by squaring the errors before averaging, gives more weight to larger errors. In this context, the ARIMA model exhibits a relatively higher RMSE compared to its MAE, indicating the presence of some larger forecast errors. Conversely, the ETS model shows a similar MAE and RMSE, suggesting a more consistent error size.

However, the SARIMA model stands out with the lowest MAE and RMSE among the models, indicating superior forecast accuracy. The choice between these models should also consider factors beyond these metrics, such as the nature of the data, the forecasting horizon, and the specific costs associated with forecast errors in the application domain.

In summary, the SARIMA model outperforms both the ARIMA and ETS models in terms of forecast accuracy, as evidenced by the lower MAE and RMSE values.

7 Conclusion

In our comprehensive analysis of Google's closing stock prices, we addressed the non-stationarity of the time series through differencing, paving the way for the application of ARIMA, SARIMA, and ETS forecasting models. The ARIMA(3, 1, 0) model was chosen for its simplicity and strong performance according to the Akaike Information Criterion (AIC), while the SARIMA(1, 1, 0)(2, 1, 0)[12] model extended this by accounting for any potential seasonal effects not greatly evident in the data. The ETS model complemented these approaches by directly modeling the data's non-stationary aspects, offering a different perspective that does not rely on differencing.

The forecasts from the models indicated a continuation of the existing trend, with the SARIMA model clearly hinting at possible seasonal variations. The ETS model forecast showed a less visible seasonality than SARIMA. ARIMA forecasted a clear and optimistic upwards trajectory of the stock price. Our error analysis revealed that the SARIMA model outperformed the ARIMA and ETS models in accuracy, exhibiting the lowest Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

These findings underscore the importance of utilizing multiple models to capture the full spectrum of patterns present in time series data, especially when preparing for the unpredictability of stock price movements. While our models provide structured forecasts based on historical data, they do not account for unforeseen events that could significantly impact the market. Thus, any predictions should be approached with caution, as the volatile nature of the stock market often defies the most sophisticated statistical models.