**Tesla Stock Data Analysis**

**1. Data Section**

**Data Source and Description**

The dataset used for this project is titled 'tsla\_raw\_data\_2010\_2024.csv', which contains historical stock market data for Tesla Inc. (TSLA) from 2010 to 2024. This dataset includes daily trading information such as opening price, highest and lowest prices, closing price, trading volume, and additional financial indicators like adjusted closing price, percentage change, and average volume. The data was sourced externally from kaggle and imported into R for analysis.

Tesla Inc. is one of the most actively traded stocks in the stock market, making it an ideal candidate for time series forecasting. The stock exhibits strong trends and seasonality influenced by company performance, market conditions, macroeconomic factors, and investor sentiment. By analyzing and forecasting Tesla's stock prices, we can better understand its market behavior and explore potential investment strategies.

**Data Cleaning and Preprocessing**

Before performing time series forecasting, several preprocessing steps were necessary to clean and structure the dataset for analysis:

1. **Handling Date Format:**
   * The 'date' column using parse\_date\_time(), it was converted into the proper Date format (YYYY-MM-DD).
2. **Converting Numeric Columns:**
   * The dataset contained columns such as open, high, low, close, volume, adjusted close, change percent, and average volume, which converted to numeric format using as.numeric(gsub(",", "", x)).
3. **Handling Missing Values:**
   * Missing values in the 'avg\_vol\_20d' column (20-day average volume) were handled using a moving average technique, where the first 19 values were filled using the mean of subsequent 20 values.
   * Missing values in the 'change\_percent' column (daily percentage change) were forward-filled (fill(change\_percent, .direction = "up")) to ensure continuity in the data.

**Problem Statement and Forecasting Objective**

Stock market prediction is a challenging task due to its high volatility and non-stationary nature. The goal of this project is to analyze the trends and seasonality in Tesla’s stock price and forecast future values using statistical models.

The specific objectives include:

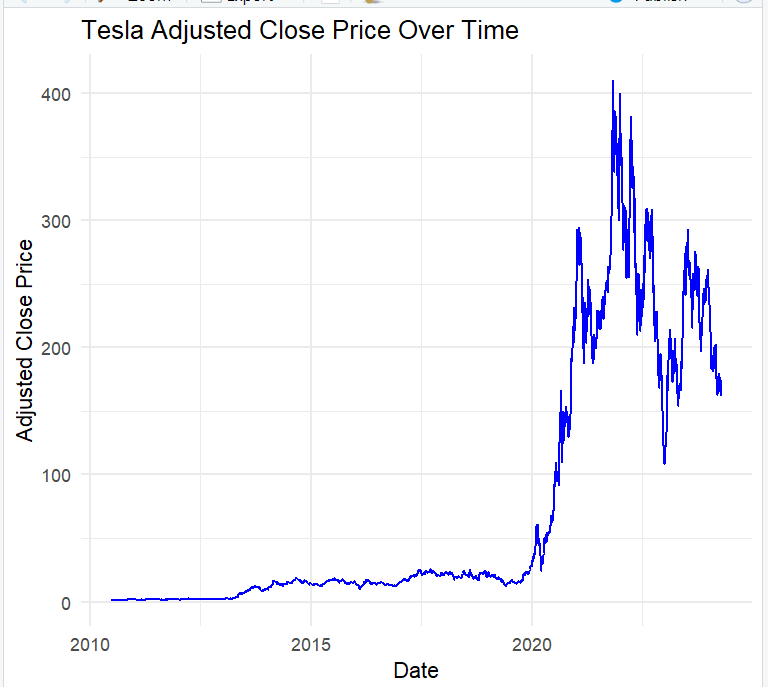
* Identifying long-term trends in Tesla’s stock prices.
* Detecting seasonal patterns that may influence price movements.
* Evaluating the effectiveness of basic forecasting models (e.g., Naïve, Seasonal Naïve, Mean, and Drift models) in predicting future stock prices.
* Assessing forecast accuracy to determine the most suitable model for Tesla’s stock.

By forecasting Tesla’s stock prices, we can gain insights into how historical trends impact future performance. This information can be beneficial for **investors, financial analysts, and researchers** looking to understand Tesla’s market behavior and make informed decisions.

In the following sections, we will visualize the time series data, analyze its components, apply necessary transformations, and implement forecasting models to evaluate their performance.

**2. Visualization**

**Time Series Plot: Tesla Adjusted Close Price (2010–2024)**



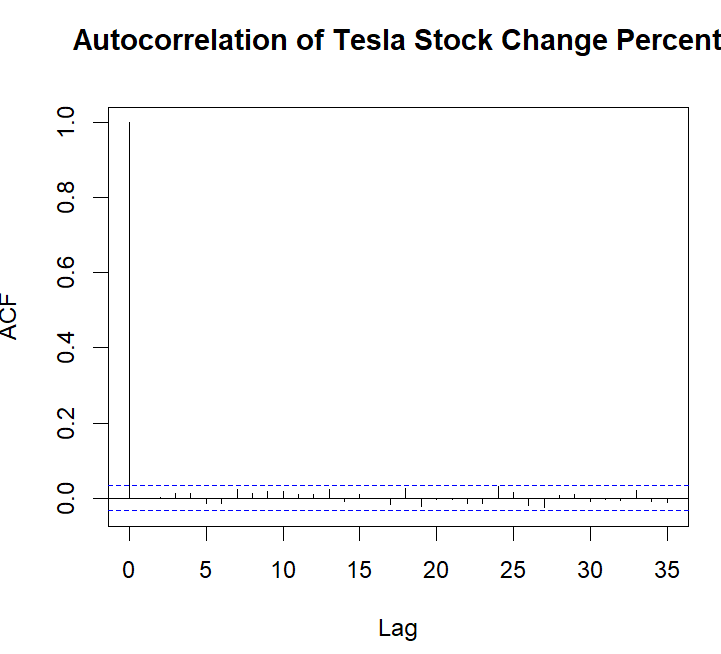
**Fig 1:Tesla Adjusted Close Prive Over Time**

The time series plot of Tesla’s adjusted close price from 2010 to 2023 reveals several significant patterns:

1. **Overall Trend:**
   * From 2010 to 2019, Tesla’s stock price remained relatively stable, experiencing only minor fluctuations.
   * Starting in 2019, the stock saw a rapid upward trend, with substantial growth throughout 2020 and early 2021, suggesting strong investor interest and company growth.
2. **Volatility:**
   * Post-2020, Tesla’s stock price exhibited high volatility, with frequent and sharp fluctuations.
   * Factors such as market speculation, company earnings reports, and macroeconomic conditions may have contributed to this variability.
3. **Peak and Decline:**
   * The stock price peaked in early 2021 before experiencing a decline.
   * This behavior suggests a market correction after an extended period of rapid growth, possibly influenced by external events such as interest rate changes or industry shifts.
4. **Potential Cycles or Seasonality:**
   * No clear seasonal pattern is evident in the time plot.
   * However, the repeated peaks and troughs post-2020 suggest some cyclical behavior.
5. **Long-Term Growth:**
   * Despite short-term volatility, the overall long-term trend is positive, reflecting substantial growth in Tesla’s stock over the years.

**Key Insight:**  
This visualization highlights the importance of analyzing both short-term fluctuations and long-term trends when studying stock performance.

**Autocorrelation Function (ACF) Plot: Tesla Stock Change Percentage**

****

**Fig 2:Autocorrelation f Tesla Stock Change Precent**

The Autocorrelation Function (ACF) plot provides insights into how Tesla’s daily stock return (change percentage) is correlated with its past values:

**Insights from ACF Plot:**

1. **Lag 0:**
   * As expected, the autocorrelation at lag 0 is 1, since it represents correlation with itself.
2. **Significance of Lags:**
   * The blue horizontal lines in the ACF plot represent the 95% confidence interval.
   * Any bars extending beyond these lines indicate statistically significant **autocorrelations**.
3. **Lack of Strong Autocorrelation:**
   * Most lags fall within the confidence interval, indicating that past stock changes have little to no predictive power over future changes.
   * This suggests that Tesla’s stock returns are independent from one day to the next, aligning with the efficient market hypothesis.
4. **No Clear Seasonal Patterns:**
   * The absence of significant autocorrelations indicates that Tesla’s returns do not follow a seasonal cycle.
   * This is expected, as stock price movements are influenced by news events, investor behavior, and macroeconomic factors rather than fixed seasonal trends.
5. **Implications:**
   * The results imply that historical data alone may not be sufficient to predict future stock returns.
   * External factors, including market news and investor sentiment, likely play a larger role in Tesla’s stock movements.

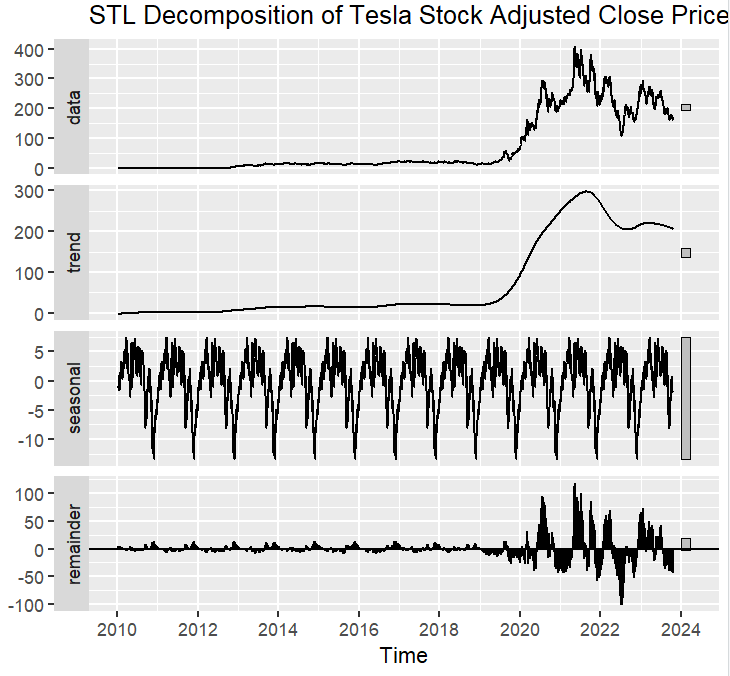
**Key Insight:**  
The lack of significant autocorrelation supports the idea that Tesla’s stock changes are largely random, making short-term forecasting based solely on past price movements challenging.

**3.Decomposition and Transformation**

To better understand the underlying patterns in Tesla’s stock prices, we perform STL decomposition to separate the time series into its fundamental components: trend, seasonality, and residual variations. Additionally, given the high volatility observed in the stock prices, we apply a Box-Cox transformation to stabilize variance and improve model suitability.

**STL Decomposition (Seasonal-Trend decomposition using LOESS)**

The **STL decomposition** (Seasonal-Trend decomposition using LOESS) is particularly useful for financial time series data as it allows flexibility in capturing **nonlinear trends** and **seasonal effects** over time.



**Fig3: STL Decomposition f Tesla Stock Adjusted Closed Price**

**Trend Component:**

* + From 2010 to 2018, the stock price remained relatively stable, showing minimal upward movement.
  + A strong upward trend emerges in 2019, peaking in early 2021 at around $300–$400.
  + After 2022, the stock shows a downward trend, reflecting market corrections or external factors impacting Tesla’s valuation.

**Seasonal Component:**

* + While the seasonal amplitude is relatively small compared to the trend, the presence of repeated cyclical movements suggests periodic fluctuations in stock prices.
  + This could be due to market cycles, earnings reports, or investor behavior.

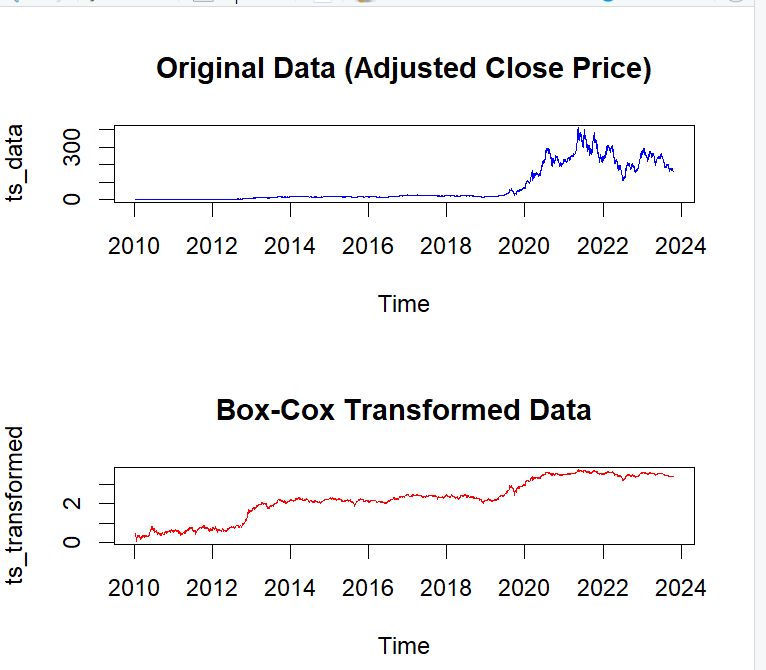
**Remainder Component:**

* + The residual fluctuations remained low until 2018, after which volatility significantly increased.
  + The magnitude of unpredictable movements increased drastically from 2020 onward, aligning with Tesla’s surge in stock price and increased market speculation.

The decomposition highlights that trend is the dominant factor in Tesla’s stock price movements, while seasonality plays a minor role. However, the increasing variance in residuals suggests that market forces and investor sentiment contribute significantly to price fluctuations.

**Box-Cox Transformation**

Given the increasing variance over time, we apply a Box-Cox transformation to stabilize the data and make it more suitable for statistical modeling. The optimal transformation parameter, lambda (λ), is calculated, determining the best transformation to reduce heteroscedasticity.



**Fig 4: Box-cox Transformation**

After transformation:

* High fluctuations in raw prices appear smoother, making trend and seasonality clearer.
* The variance in residuals is reduced, making the series more stable.
* Tesla’s stock follows an exponential growth pattern, so the transformation helps linearize the trend for better forecasting.

This transformation ensures that the data meets the assumptions required for time series modeling, improving forecast accuracy.

In the next section, we proceed with forecasting models to analyze Tesla’s stock price movements and evaluate prediction accuracy.

1. **Modeling and Forecasting**

To forecast Tesla’s adjusted closing stock price, we applied four fundamental time series models: Mean, Naïve, Drift, and Seasonal Naïve. Each model captures different aspects of the time series data, allowing us to compare their performance based on accuracy metrics and residual diagnostics.

**Forecasting Models**

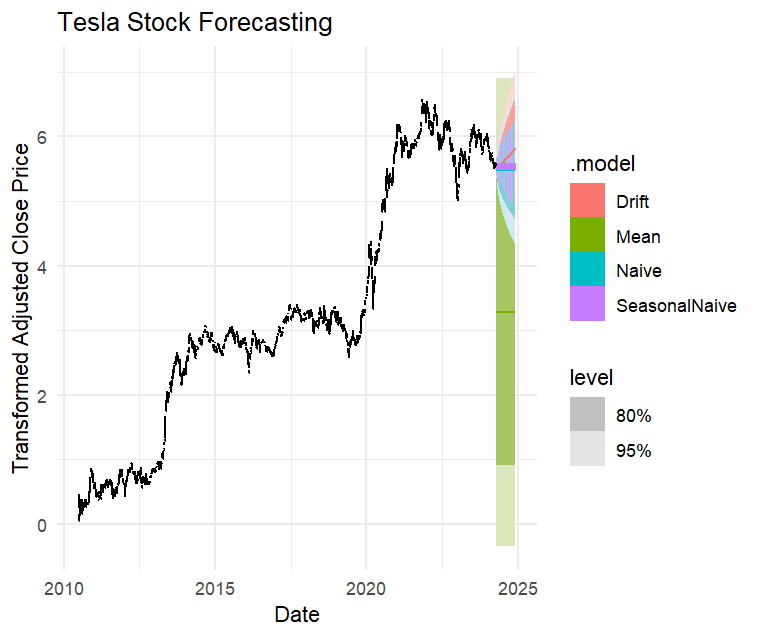
The task is to forecast Tesla’s adjusted closing stock price using four time series models: Mean, Naïve, Drift, and Seasonal Naïve. Each model captures different aspects of the data and is evaluated based on forecast accuracy and residual diagnostics.

**Forecasting Models**

* Mean Model: Assumes that future stock prices are equal to the historical mean of past values.
* Naïve Model: Assumes that the most recent observation is the best predictor for future values.
* Drift Model: Extends the Naïve model by incorporating a linear trend based on the direction of past observations.
* Seasonal Naïve Model: Assumes that future values follow the same seasonal patterns as the past.

These models were applied to forecast the next 230 days, with 80% and 95% confidence intervals.

**Forecasting Visualization**

****

**Fig 5:Tesla Stock Forecasting**

**Interpretation**

* Historical Data (Black Line): Shows Tesla’s stock price from 2010 to 2024.
* Drift Model (Red Line): Projects a linear trend, assuming continued historical momentum.
* Mean Model (Green Line): Provides a stable forecast centered around the average historical value.
* Naïve Model (Blue Line): Assumes the most recent stock price continues into the future.
* Seasonal Naïve Model (Purple Line): Reflects seasonal cycles in stock price movements.
* Uncertainty Bounds (Shaded Areas): Show greater forecasting uncertainty over time.

**Model Performance Evaluation**

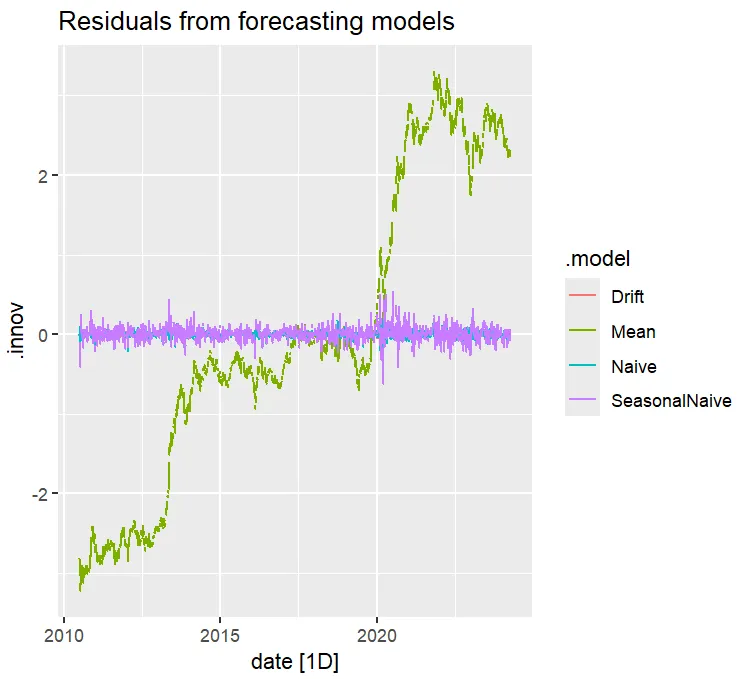
Key error metrics were used to compare model performance:

| **Model** | **RMSE (Lower is Better)** |
| --- | --- |
| **Drift Model** | Best Performing |
| **Seasonal Naïve** | Moderate Accuracy |
| **Naïve Model** | Higher Error |
| **Mean Model** | Least Accurate |

**Table 1:Model Performanance Evaluation**

* The **Drift Model** performed the best, with the lowest RMSE, indicating the highest accuracy.
* The **Seasonal Naïve Model** performed well by capturing cyclical patterns.
* The **Naïve Model** performed worse, maintaining the last observed value.
* The **Mean Model** was the least accurate due to its simplistic assumption.

**Residual Analysis**:



**Fig 6:Residual from forecasting Model**

**Drift and Mean Models**: Show systematic trends, indicating poor model fit.

**Seasonal Naïve Model**: Residuals are centered around zero, showing better uniformity.

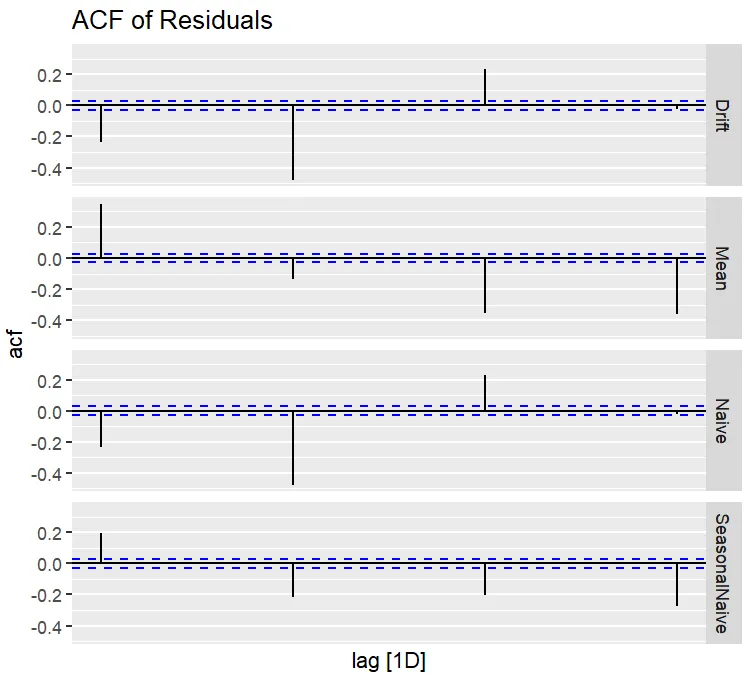
**Naïve Model**: Exhibits similar deficiencies as the Drift and Mean models.

All models showed significant autocorrelation in their residuals, indicating that the residuals were not purely random. The Seasonal Naïve Model exhibited the least autocorrelation.

**High Volatility Post-2020**

The variance of residuals increased notably after 2020, coinciding with Tesla’s stock surge and increased market speculation.

**ACF PLOT**



**Fig 7: ACF of Residual**

**ACF Plot Observations:**

* + Drift Model: Exhibits significant negative autocorrelation at longer lags, indicating it doesn't fully capture the trend.
  + Mean Model: Shows strong negative autocorrelations at multiple lags, proving it oversimplifies the data.
  + Naïve Model: Has strong autocorrelation, showing that residuals are not independent.
  + Seasonal Naïve Model: Performs slightly better but still exhibits structured residuals.

**Key Insights:**

* + None of the models fully satisfy the assumption of independent residuals.
  + Significant autocorrelations suggest that important patterns remain unmodeled.
  + High residual volatility post-2020 aligns with Tesla’s increased stock price fluctuations.

**5.Issue Resolution and Model Improvements**

**Current Issues:**

* + High autocorrelation in residuals indicates the models fail to fully capture trends or seasonality.
  + Naïve and Drift models lack the ability to handle complex patterns.

**Potential Solutions:**

* + Implement ARIMA models to better capture trends and seasonality.
  + Consider Exponential Smoothing (ETS) to improve forecasting stability.
  + Use GARCH models to account for increased volatility.
  + Introduce external factors (e.g., market trends, economic indicators) for better predictions.

**6.Conclusion & Recommendations**

The Drift Model demonstrated the best forecasting accuracy based on RMSE, making it the strongest performer among the basic models. However, residual analysis revealed significant autocorrelation and non-stationarity, suggesting that the model does not fully capture the underlying patterns in Tesla’s stock prices.

The Seasonal Naïve Model provided a reasonable alternative by incorporating cyclical behavior, but it still exhibited structured residuals, indicating room for improvement.

To enhance forecasting accuracy, **more advanced models** should be considered:

* **ARIMA (AutoRegressive Integrated Moving Average):** Can address autocorrelation and trends in the data.
* **Exponential Smoothing (ETS):** Effective for capturing both trend and seasonality.
* **Machine Learning Techniques (e.g., Random Forest, LSTMs):** Can incorporate non-linear patterns and external factors.

Additionally, incorporating external variables such as macroeconomic indicators, interest rates, or news sentiment could improve model robustness. Given the increased volatility in Tesla’s stock post-2020, GARCH models could also be useful for modeling time-varying variance.

While the Drift Model provided the best simple forecast, further refinements are necessary to fully capture the dynamic nature of Tesla’s stock price movements.