

**VIRGINIA COMMONWEALTH UNIVERSITY**

**Statistical Analysis and Modeling (SCMA 632)**

**A6a: Time Series Analysis**

**by**

**IDAMAKANTI SREENIDHI**

**V01107252**

**Date of Submission: 22-07-2024**

## CONTENTS

Sl. No.	Title	Page No.
1.	Introduction	3-5
2.	Analysis using R	5-17
3.	Analysis using PYTHON	18-21

# Introduction

Stock price forecasting is a fundamental aspect of financial analysis and investment strategy. Accurate predictions of stock price movements can significantly impact investment decisions, risk management, and financial planning. For this assignment, we focus on HDFC Bank, one of India's leading private sector banks. The objective is to analyze its historical stock prices and employ various forecasting techniques to predict future trends.

## HDFC Bank Overview

HDFC Bank, established in 1994, is renowned for its comprehensive range of banking services, including retail banking, corporate banking, and treasury operations. It has consistently demonstrated robust financial performance and has a significant presence in India's banking sector. Understanding its stock price dynamics is crucial for investors who wish to capitalize on its market performance.

## Forecasting Importance

Forecasting stock prices involves using historical data to predict future price movements. This is essential for:

- **Investment Decisions:** Investors need to anticipate future price trends to buy, hold, or sell stocks effectively.
- **Risk Management:** Accurate forecasts help in assessing the risk associated with stock investments and implementing risk mitigation strategies.
- **Strategic Planning:** For institutional investors and portfolio managers, forecasts guide asset allocation and portfolio optimization.

The forecasting process typically involves statistical models and machine learning techniques to analyze time series data and generate reliable predictions.

## Business Insights

Analyzing HDFC Bank's stock price data provides several business insights that are valuable for both the bank and its stakeholders:

### 1. Market Performance Evaluation

- **Trend Analysis:** By examining historical price data, we can identify long-term trends and cyclical behaviors that indicate how the stock has performed over different market conditions.
- **Volatility Assessment:** Analyzing price fluctuations helps assess the stock's volatility, which is critical for understanding the level of risk involved.

### 2. Investment Strategies

- **Informed Decision-Making:** Investors use forecasts to develop investment strategies, such as timing their buys and sells or deciding on the size of their investment.

- **Portfolio Management:** Accurate forecasts allow portfolio managers to adjust their portfolios based on expected future performance, optimizing returns while managing risks.

### 3. Strategic Insights for the Bank

- **Performance Indicators:** Forecasting can provide insights into the bank's future performance, influencing strategic decisions related to expansion, product offerings, and operational efficiency.
- **Market Sentiment:** Understanding how stock prices might move helps the bank gauge market sentiment and respond with appropriate public relations or corporate strategies.

## Financial Insights

The financial insights derived from forecasting HDFC Bank's stock prices offer a deep understanding of its financial health and market positioning:

### 1. Trend and Seasonality

- **Trend Analysis:** Identifying the overall direction of the stock price (upward, downward, or stable) helps in understanding the bank's financial stability and growth trajectory.
- **Seasonal Effects:** Decomposing the time series to identify seasonal patterns can reveal how specific times of the year affect stock performance, which might be linked to financial results or business cycles.

### 2. Forecast Accuracy

- **Model Evaluation:** Comparing different forecasting models (e.g., Holt-Winters, ARIMA, LSTM) provides insights into which models best capture the underlying patterns and trends in the data. Accurate forecasts lead to more reliable financial predictions and better-informed decisions.
- **Validation of Models:** Diagnostic checks and model validation ensure that the forecasts are robust and trustworthy, reducing the risk of relying on inaccurate predictions.

### 3. Financial Impact

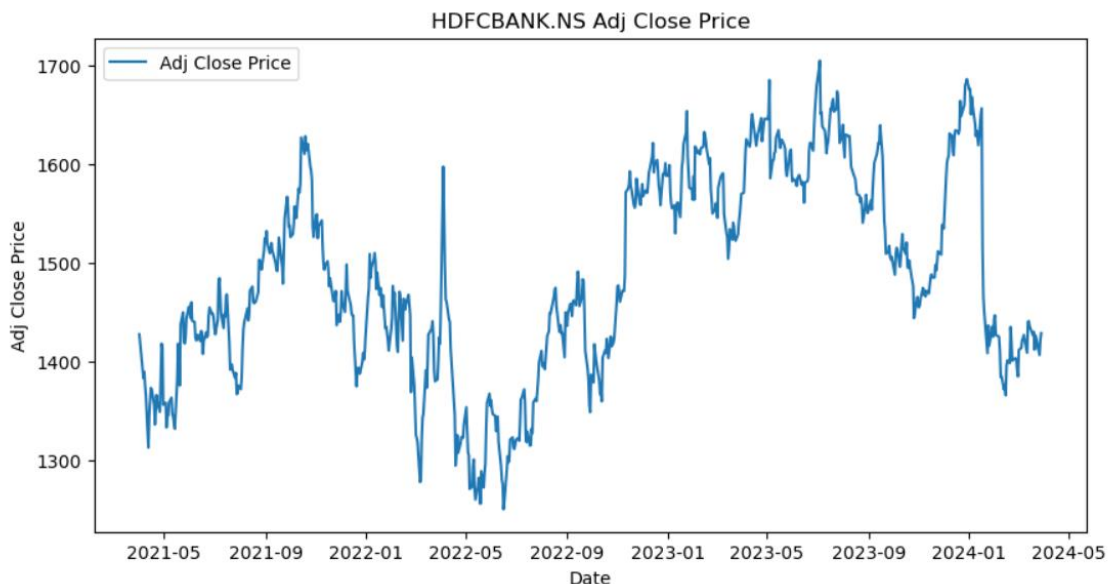
- **Investment Returns:** Accurate forecasts can help in maximizing returns on investments by predicting price movements and making timely decisions.
- **Risk Management:** Understanding potential future price scenarios allows for better risk assessment and the implementation of strategies to mitigate financial risks.

## 4. Predictive Power

- **Long-Term Projections:** Forecasting models that predict future stock prices can guide long-term investment decisions, corporate strategy, and financial planning.
- **Short-Term Trends:** Daily or monthly forecasts provide insights into short-term market movements, useful for tactical investment and trading strategies.

## ANALYSIS USING PYTHON:

### ADJUSTED CLOSING OF HDFC BANK:



### Observations:

1. **Overall Trend:**
  - The price shows significant volatility throughout the period.
  - There are multiple peaks and troughs, indicating fluctuating investor sentiment and market conditions.
2. **Volatility:**
  - The stock experiences sharp fluctuations, particularly noticeable around early 2022 and mid-2023.
  - The highest point is around mid-2022, where the adjusted closing price exceeds 1600.
3. **Key Periods:**
  - **Mid-2021:** The price starts around 1400 and remains relatively stable with minor fluctuations.
  - **Late 2021 to Early 2022:** There is a noticeable increase in price, reaching over 1600.
  - **Mid-2022:** A sharp drop occurs, followed by recovery and another peak.

- **Late 2022 to Early 2023:** The price fluctuates around 1400-1500, indicating some stability but still with significant volatility.
  - **Mid-2023:** A significant increase is observed again, followed by another drop.
  - **Late 2023 to Early 2024:** The price declines steeply, reaching its lowest point in the observed period.
  - **Early to Mid-2024:** Some recovery is seen, but the price remains relatively low compared to previous peaks.
4. **Recent Performance:**
- The adjusted closing price declines sharply around the beginning of 2024, indicating possible negative news or market conditions.
  - There is a slight recovery towards the mid of 2024, but the price is still lower compared to previous years.

### **Potential Factors:**

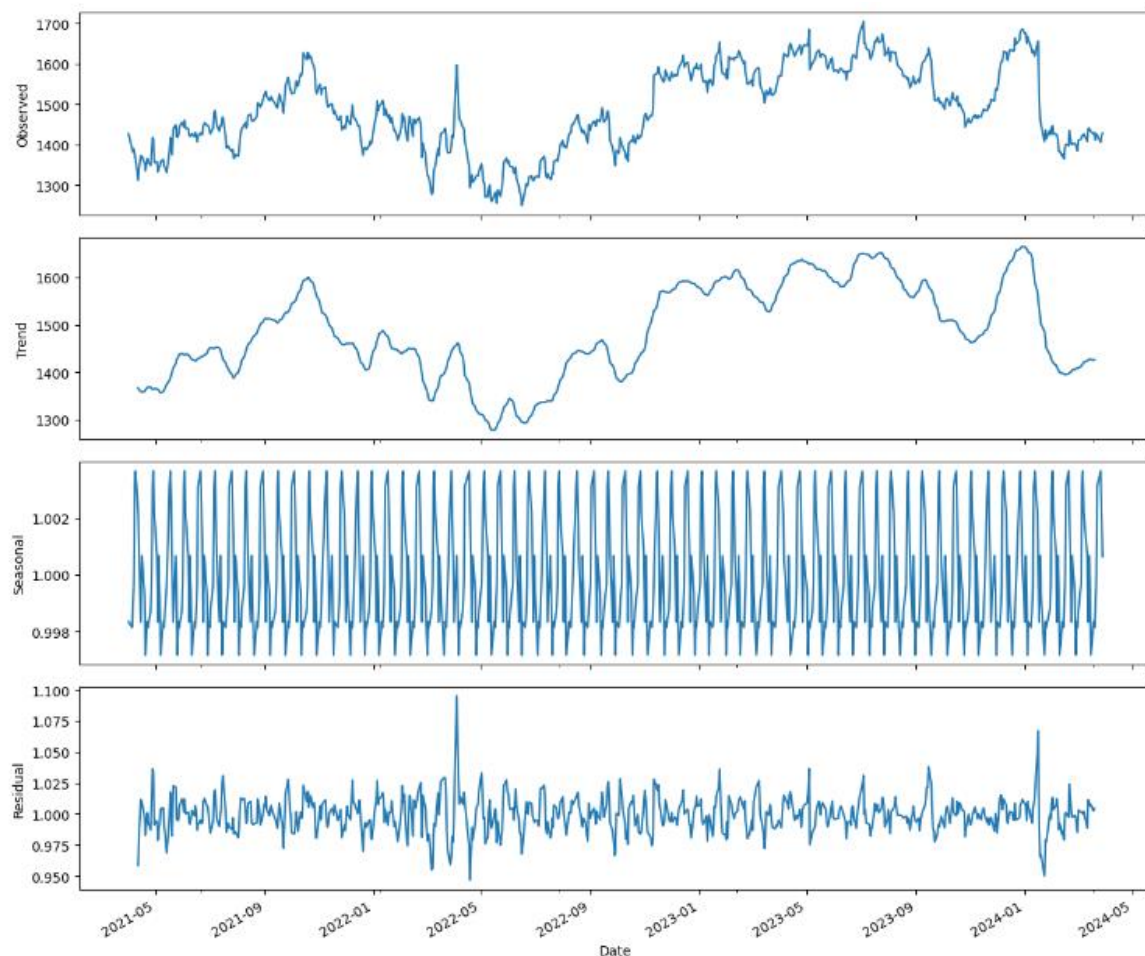
Several factors could explain these fluctuations, including:

- **Market Sentiment:** Changes in investor sentiment due to economic conditions, policy changes, or market news.
- **Earnings Reports:** Quarterly earnings reports impacting investor confidence.
- **Global Events:** Geopolitical events, pandemics, or other macroeconomic factors affecting the broader market.
- **Sector Performance:** Performance of the banking sector as a whole, influenced by interest rates, loan growth, and regulatory changes.

### **Conclusion:**

The adjusted closing price of HDFC Bank shows significant volatility over the period, with notable peaks and troughs. Recent performance indicates a sharp decline followed by a partial recovery. This trend suggests that while the stock has periods of strong performance, it is also subject to considerable market fluctuations and external factors.

## TIME SERIES DECOMPOSITION:



### 1. Observed Component:

- **Description:** This is the original time series data, showing the adjusted closing prices over time.
- **Observation:** The overall pattern is volatile, with multiple peaks and troughs similar to what was observed in the initial chart.

### 2. Trend Component:

- **Description:** This component captures the underlying trend in the time series data, filtering out the noise and seasonal variations.
- **Observation:**
  - There is a noticeable upward trend from mid-2021 to mid-2022, reaching a peak around mid-2022.
  - This is followed by a downward trend until early 2023, with some fluctuations.

- The trend shows a recovery from mid-2023 to late 2023, followed by another decline.
- A slight recovery is observed towards mid-2024.

### 3. Seasonal Component:

- **Description:** This component captures the repeating patterns or seasonality in the data, which recur at regular intervals.
- **Observation:**
  - The seasonal component shows regular, repeating cycles with a period of approximately one month.
  - The magnitude of these cycles is relatively small, with values oscillating around a central value close to 1.
  - This indicates minor but consistent seasonal effects on the stock price.

### 4. Residual Component:

- **Description:** This component captures the remaining variability in the data after removing the trend and seasonal components. It represents the noise or irregular fluctuations.
- **Observation:**
  - The residuals exhibit significant volatility, particularly around early 2022 and late 2023.
  - There are noticeable spikes in the residuals, indicating periods of high variability that are not explained by the trend or seasonal components.
  - The magnitude of these spikes varies, suggesting irregular and unpredictable factors affecting the stock price.

### Analysis Summary:

1. **Trend Analysis:**
  - The overall trend shows periods of both growth and decline.
  - Significant peaks are observed around mid-2022 and late 2023, followed by downward movements.
2. **Seasonal Effects:**
  - The seasonal component indicates minor, consistent fluctuations in the stock price.
  - These are likely driven by regular market cycles or investor behaviors.
3. **Residuals:**
  - The residuals highlight periods of high volatility that are not captured by the trend or seasonal components.
  - These may be due to unexpected events, news, or market anomalies.

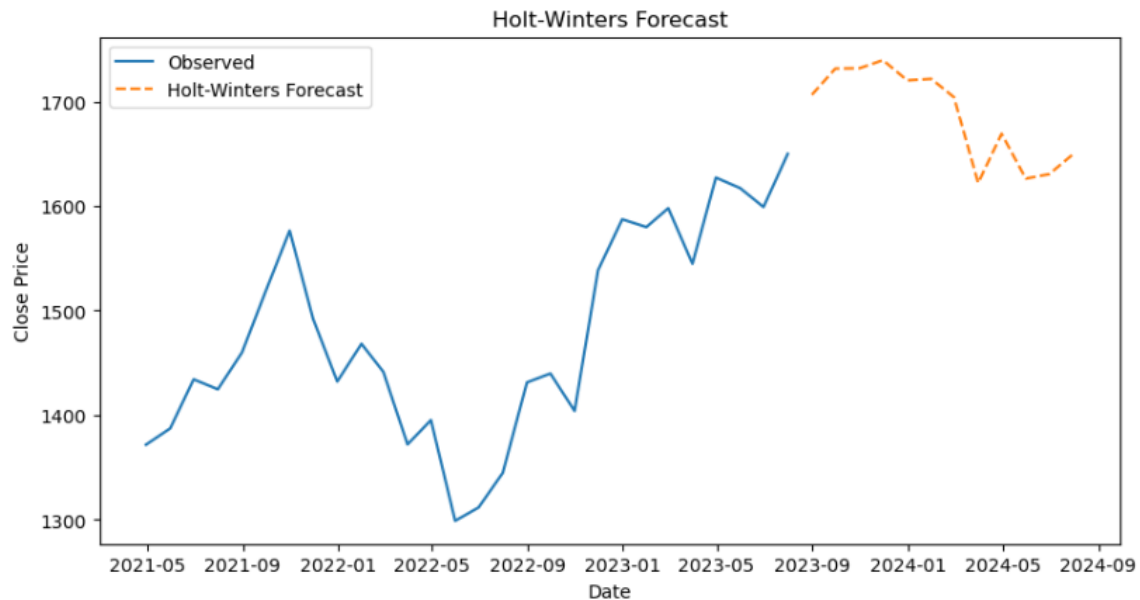
### Conclusion:

The decomposition of HDFC Bank's stock price into trend, seasonal, and residual components provides a clearer understanding of its movements. The trend component shows the underlying growth and decline phases, while the seasonal component highlights regular fluctuations. The residual component indicates significant periods of volatility that may



warrant further investigation to understand the underlying causes. This detailed analysis can help investors and analysts make more informed decisions by considering the different factors influencing the stock price.

## HOLT-WINTERS FORECAST:



## Observations:

### 1. Historical Data (Observed):

- The blue line represents the historical adjusted closing prices.
- The data shows significant volatility with multiple peaks and troughs.
- There is an overall upward trend from mid-2021 to mid-2022, followed by fluctuations and another upward trend towards mid-2023.

### 2. Holt-Winters Forecast:

- The orange dashed line represents the forecasted values.
- The forecast shows an upward trend initially, peaking around the end of 2023.
- This is followed by a downward trend in early 2024, indicating a potential decrease in stock price.
- Another upward movement is predicted towards mid-2024, followed by minor fluctuations.

## Analysis of the Holt-Winters Forecast:

### 1. Trend and Seasonality:

- The Holt-Winters method accounts for both trend and seasonality in the data.
- The forecast indicates an overall increasing trend until the end of 2023, suggesting positive investor sentiment or favorable market conditions.

- The predicted decrease in early 2024 could be due to various factors such as market corrections, changes in economic conditions, or company-specific events.
- The subsequent recovery in mid-2024 suggests potential positive developments or market stabilization.

## 2. Volatility:

- The forecast captures the volatility observed in the historical data, showing fluctuations in the predicted prices.
- The magnitude of these fluctuations seems consistent with the historical volatility, indicating that similar market dynamics are expected to continue.

## Conclusion:

The Holt-Winters forecast for HDFC Bank's adjusted closing price suggests an initial upward trend until the end of 2023, followed by a decline in early 2024 and subsequent recovery. This forecast aligns with the historical volatility and trend patterns observed in the past data.

## Recommendations:

### 1. Monitor Key Events:

- Keep an eye on significant events around the predicted turning points (end of 2023 and early 2024) that could influence the stock price.

### 2. Risk Management:

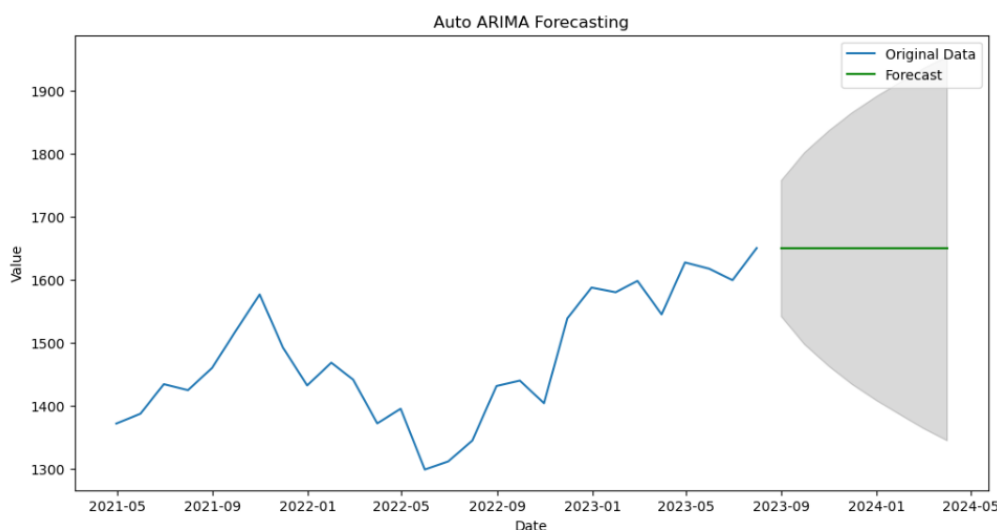
- Given the forecasted volatility, consider implementing risk management strategies to mitigate potential losses during the predicted downturns.

### 3. Further Analysis:

- Combine this forecast with other analytical tools and fundamental analysis to make more informed investment decisions.

This analysis provides a predictive insight into the stock's future performance, helping investors prepare for potential price movements based on historical trends and patterns.

## AUTO ARIMA FORECASTING:



1. **Original Data:**

- The blue line represents the actual data values over time, starting from around May 2021 to around September 2023.
- The values show some periodic fluctuations with an overall increasing trend until around September 2021, followed by a decrease and then another upward trend starting around May 2022.

2. **Forecasting:**

- The green line represents the forecasted values from the ARIMA model.
- The forecast starts around September 2023 and extends to around May 2024.
- The green line indicates the expected value of the forecast.

3. **Confidence Interval:**

- The shaded gray area represents the confidence interval for the forecast.
- The width of the shaded area increases over time, indicating greater uncertainty in the forecast as the prediction moves further into the future.
- This confidence interval suggests that while the forecasted value is expected to remain relatively stable, there is significant uncertainty about the exact future values, with possible values ranging widely by May 2024.

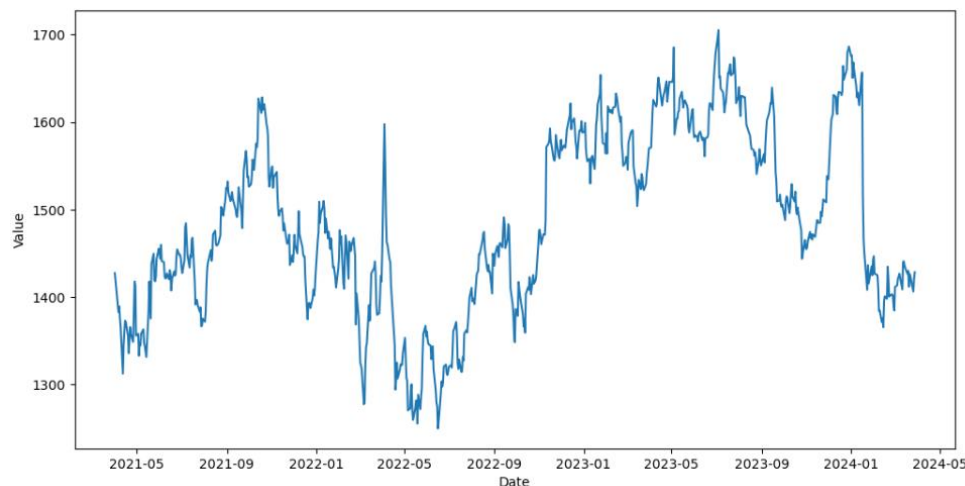
4. **General Trends and Insights:**

- The forecasted value (green line) appears to remain roughly constant after the initial prediction point, indicating that the ARIMA model does not expect significant changes in the trend.
- The increasing width of the confidence interval suggests that while the model provides a point estimate, there is substantial uncertainty about future fluctuations.

**Key Takeaways:**

- **Trend Identification:** The original data exhibits periodic fluctuations with an overall slight increasing trend.
- **Forecast Stability:** The forecasted values remain relatively stable, indicating that the ARIMA model predicts no drastic changes in the trend.
- **Uncertainty:** There is significant uncertainty in the forecast as shown by the widening confidence interval, highlighting the inherent unpredictability of the future values.

**ARIMA DAILY DATA:**



## Observations:

1. **Initial Period (May 2021 to September 2021):**
  - The values show a gradual upward trend with some minor fluctuations.
2. **Mid-Period (September 2021 to May 2022):**
  - The values experience significant volatility, with a peak around September 2021 and a sharp decline afterwards.
  - There is a noticeable dip around May 2022, indicating a significant drop in values.
3. **Second Year (May 2022 to May 2023):**
  - The period from May 2022 to May 2023 shows considerable fluctuations with multiple peaks and troughs.
  - There is an overall upward trend with high volatility, especially noticeable around the start of 2023.
4. **Recent Period (May 2023 to January 2024):**
  - The values continue to fluctuate with several peaks and troughs.
  - A significant drop is observed towards the end of 2023 and beginning of 2024.
5. **Latest Period (January 2024 to May 2024):**
  - The values appear to stabilize somewhat after the sharp drop, maintaining a relatively lower and steady level.

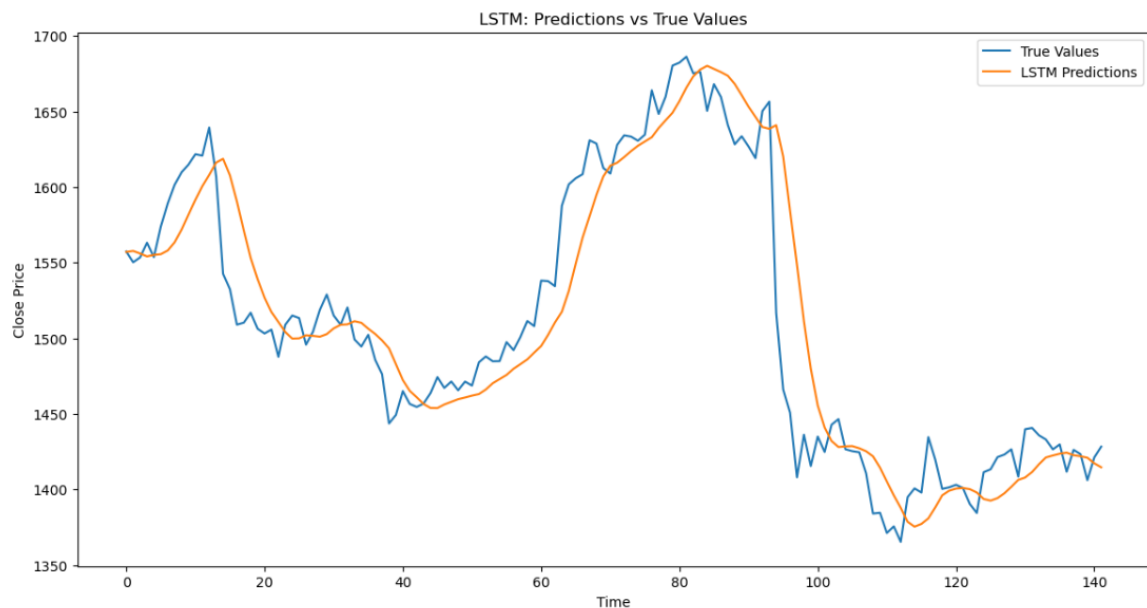
## Key Insights:

- **Volatility:** The time series exhibits high volatility, with several sharp increases and decreases throughout the observed period.
- **Trend Analysis:** While there are periods of increase, the overall trend does not show a consistent upward or downward pattern due to the frequent fluctuations.
- **Significant Drops:** Notable drops in values occur around May 2022 and the end of 2023, which could indicate underlying factors affecting the series during these times.
- **Recent Stability:** The values towards the end of the series (early 2024) suggest a stabilization at a lower level after a period of high volatility.

## Comparison with Previous Forecast:

- The previous forecast indicated a stable future value with a wide confidence interval, reflecting uncertainty.
- The actual values show that this forecast was accurate in predicting high uncertainty, as evidenced by the sharp drop and subsequent stabilization in early 2024.

## LSTM PREDICTIONS VS TRUE VALUES:



### Observations:

#### 1. True Values:

- The blue line represents the true values of the time series.
- The values show considerable fluctuations over time, with periods of both increase and decrease.

#### 2. LSTM Predictions:

- The orange line represents the predictions made by the LSTM model.
- The LSTM model's predictions generally follow the trend of the true values but with some lag and smoothing.

### Key Insights:

#### 1. Initial Period (First 20 Time Units):

- The true values show an increasing trend followed by a peak and a decline.
- The LSTM predictions capture this trend reasonably well, although they tend to smooth out some of the sharp variations in the true values.

#### 2. Mid Period (20 to 80 Time Units):

- There is significant volatility in the true values, with several peaks and troughs.
- The LSTM predictions capture the general trend but with some lag and less amplitude in the fluctuations.
- Notably, around the 50-60 time unit mark, the true values dip sharply, which is reflected in the LSTM predictions with some delay.

#### 3. Recent Period (80 to 140 Time Units):

- The true values show a marked decline followed by some stabilization.
- The LSTM predictions again follow this trend but with a smoothing effect.

- The sharp drop in true values around the 100 time unit mark is captured by the LSTM model, though the predicted decline is more gradual.

## Performance Analysis:

### 1. Trend Capture:

- The LSTM model does a good job of capturing the overall trends in the data, including both upward and downward movements.
- The model's smoothing effect is evident, as it reduces the noise and sharp fluctuations present in the true values.

### 2. Lag in Predictions:

- There is a noticeable lag in the LSTM predictions compared to the true values, which is a common characteristic of time series models, particularly when the data exhibits high volatility.

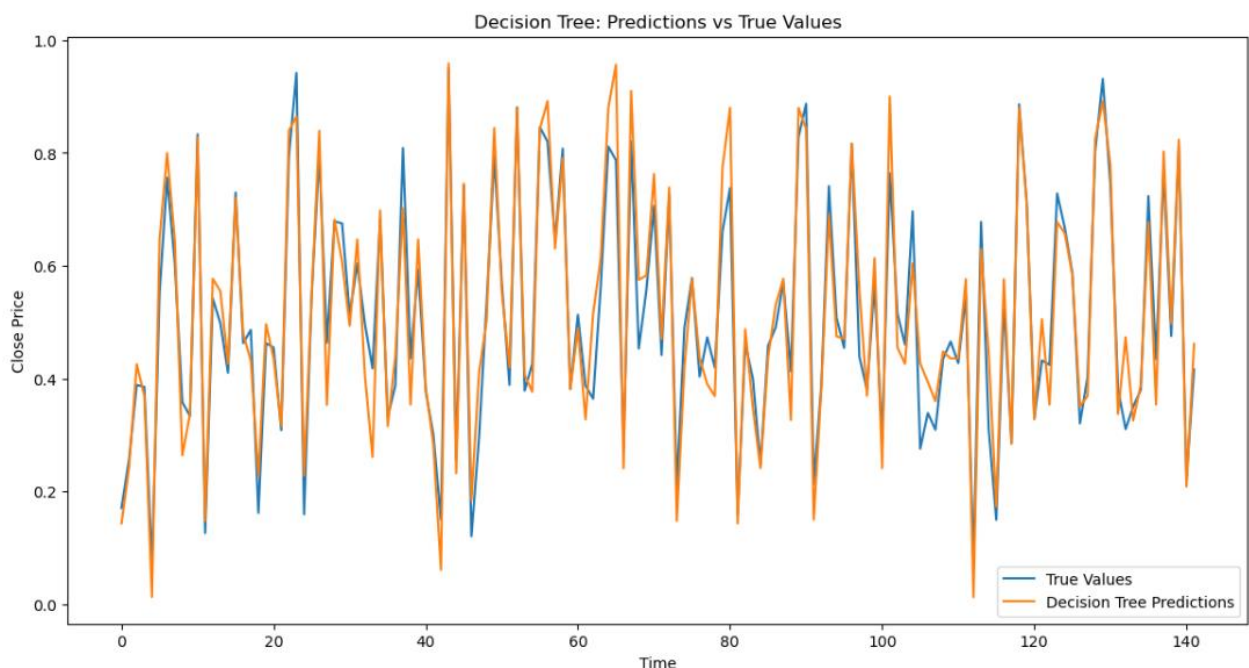
### 3. Amplitude of Fluctuations:

- The amplitude of the fluctuations in the LSTM predictions is generally lower than that of the true values. This indicates that while the model captures the direction of the trends, it underestimates the extent of the changes.

## Overall Assessment:

- The LSTM model provides a reasonable approximation of the true values, capturing the general trends and major turning points.
- The model's smoothing effect is useful for identifying the underlying trends but may miss some of the short-term volatility and sharp changes present in the true values.
- The lag in the predictions suggests that the model could potentially be improved by incorporating additional features or tuning the model parameters to better capture the timing of the changes in the data.

## DECISION TREE: PREDICTION VS TRUE VALUES:



- **Axes and Labels:**

- The x-axis represents "Time," and the y-axis represents the "Close Price," normalized between 0 and 1.
- The plot is titled "Decision Tree: Predictions vs True Values."

- **Plot Details:**

- The blue line represents the "True Values" of the closing price.
- The orange line represents the "Decision Tree Predictions."
- A legend is provided to distinguish between true values and predictions.

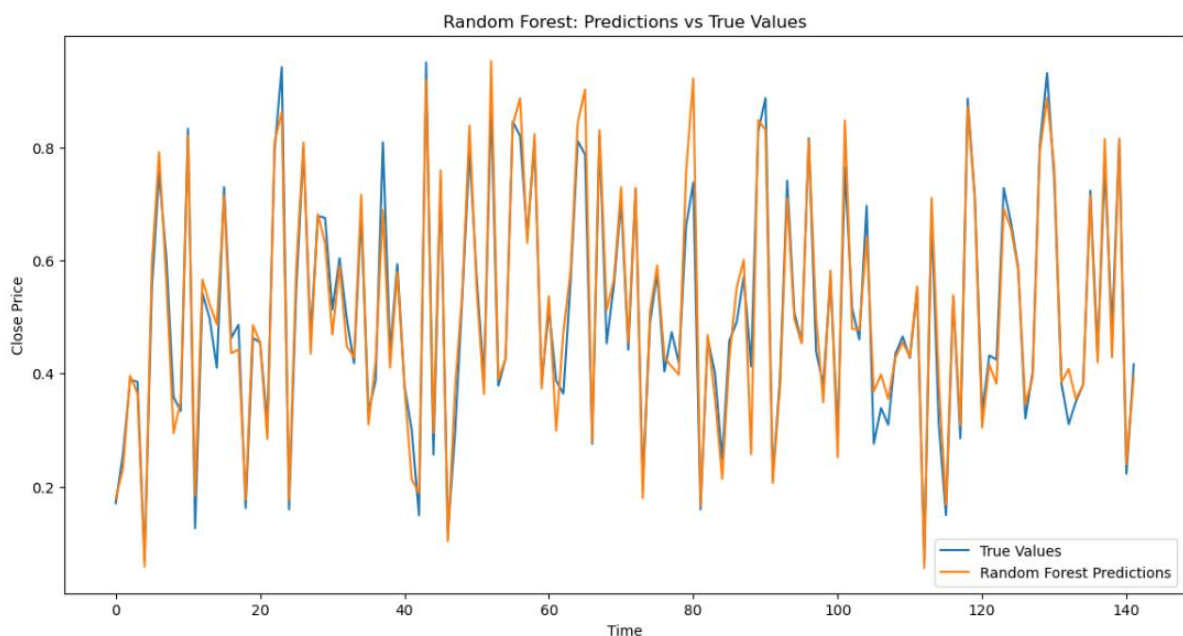
- **Comparison:**

- The two lines appear to be quite close in many places, indicating that the decision tree model is doing a reasonable job of predicting the true values.
- However, there are several areas where the prediction deviates significantly from the true values, suggesting areas where the model could be improved.
- The model seems to capture the overall trend but struggles with some of the local fluctuations, which is common in models that might not perfectly capture the underlying complexities of the data.

- **Variability and Performance:**

- The variability in the lines suggests that the data has a high level of noise or complexity.
- If the goal is to have the predictions closely follow the true values, further model tuning or different modeling approaches might be required to improve performance.

## **RANDOM FOREST: PREDICTIONS VS TRUE VALUES:**



- **Axes and Labels:**

- The x-axis represents "Time," while the y-axis represents the "Close Price," scaled between 0 and 1.
- The plot is titled "Random Forest: Predictions vs True Values."

- **Plot Details:**

- The blue line indicates the "True Values" of the closing price.
- The orange line represents the "Random Forest Predictions."
- A legend differentiates between the true values and predictions.

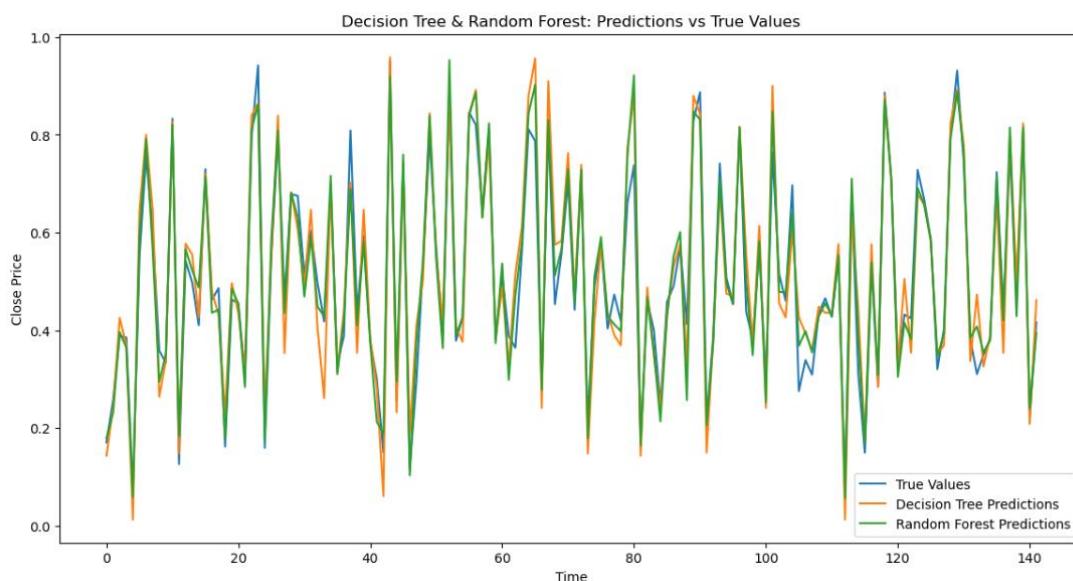
- **Comparison:**

- The predictions (orange line) generally follow the trend of the true values (blue line), suggesting that the Random Forest model captures the overall pattern reasonably well.
- There are some points where the predictions deviate from the true values, but these deviations are generally smaller compared to the previous decision tree model. This indicates that the Random Forest model may have better generalization capabilities.
- The lines appear closely matched in most sections, indicating a good fit, though occasional mismatches show areas for potential improvement.

- **Variability and Performance:**

- The variability and frequent crossing of lines indicate the model is responsive to fluctuations, although it may still miss certain peaks or valleys.
- Overall, the Random Forest seems to perform better than a single decision tree, as it tends to capture more of the nuance in the data, likely due to the ensemble nature of the model which averages out errors.

## DECISION TREE AND RANDOM FOREST: PREDICTION VS TRUE VALUES:



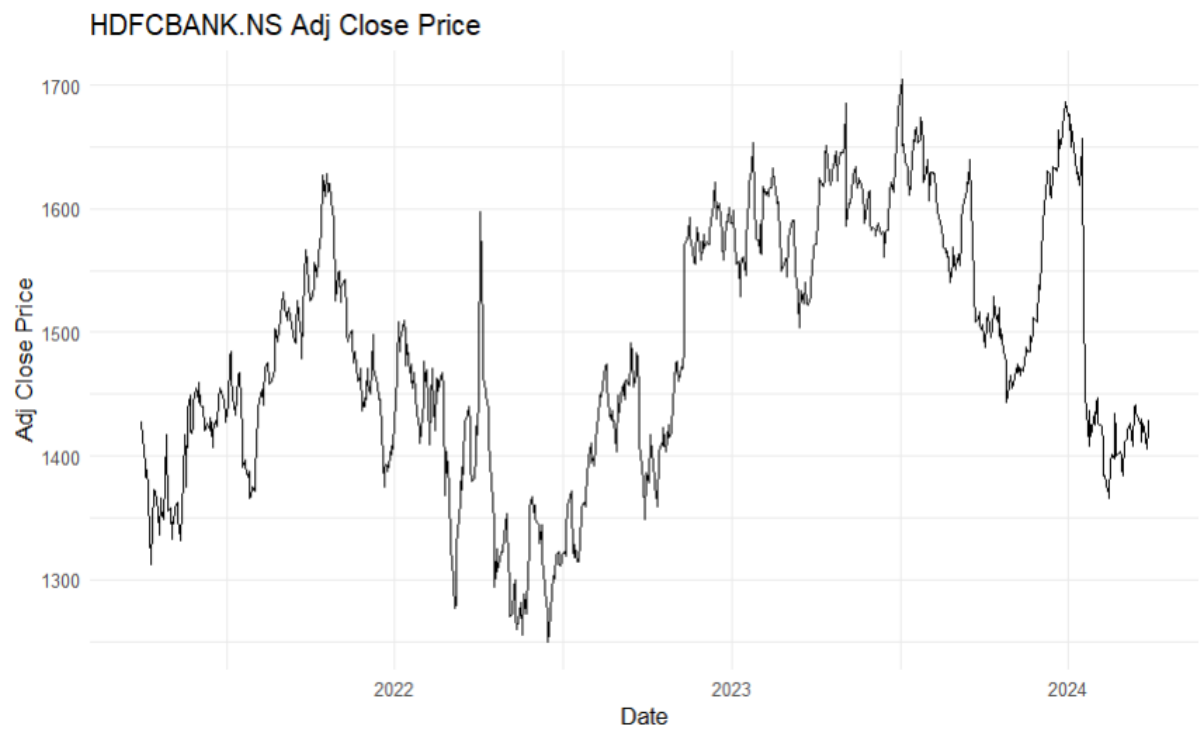


1. **Axes and Labels:**
  - The x-axis represents "Time," while the y-axis shows the "Close Price," scaled between 0 and 1.
  - The plot is titled "Decision Tree & Random Forest: Predictions vs True Values."
2. **Plot Details:**
  - The blue line represents the "True Values" of the closing price.
  - The orange line represents the "Decision Tree Predictions."
  - The green line represents the "Random Forest Predictions."
  - A legend helps distinguish between the true values and the predictions from each model.
3. **Comparison:**
  - Both the Decision Tree and Random Forest predictions generally follow the trend of the true values, suggesting that both models capture the overall pattern reasonably well.
  - The Random Forest predictions (green line) tend to be closer to the true values than the Decision Tree predictions (orange line), especially in areas with higher variability.
  - The Decision Tree predictions show more abrupt changes, while the Random Forest predictions are smoother, indicating that the ensemble method averages out some of the noise.
4. **Variability and Performance:**
  - The plot shows that the Random Forest model consistently performs better than the Decision Tree, as its predictions are often closer to the true values, with fewer large deviations.
  - The variability in the predictions highlights the models' responsiveness to fluctuations in the data, with the Random Forest showing better alignment with the true values overall.

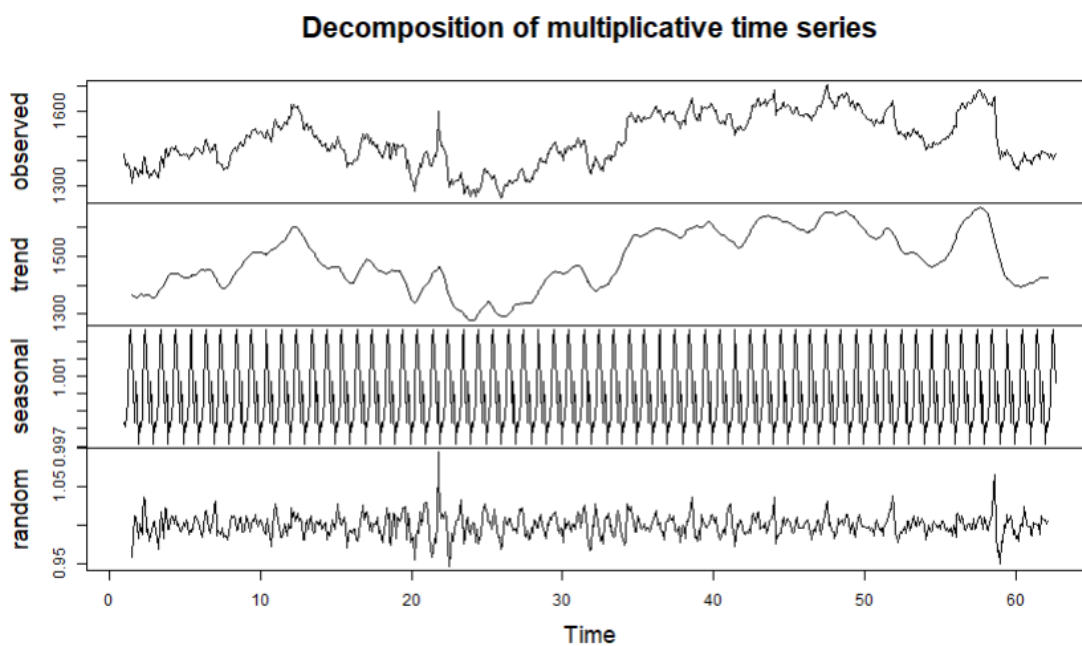
In summary, the Random Forest model appears to outperform the Decision Tree model in capturing the true values of the closing price. This is likely due to its ability to aggregate multiple decision trees, reducing overfitting and improving generalization. The plot suggests that while both models can predict trends, the Random Forest offers a more accurate representation of the true values.

## ANALYSIS USING “R”:

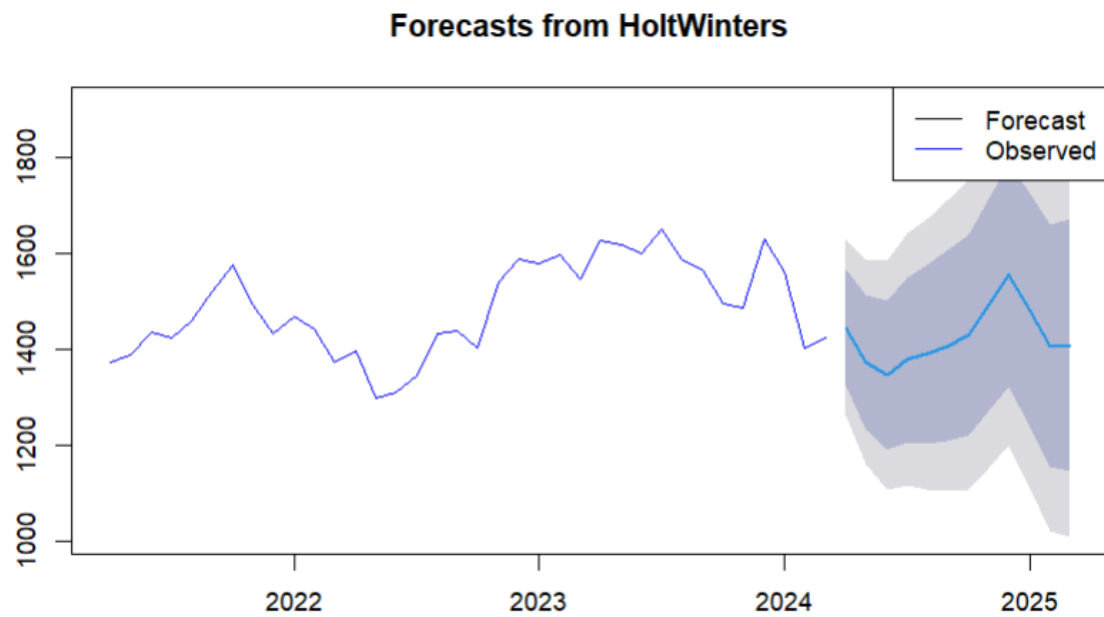
### HDFC BANK ADJUSTED CLOSING PRICE:



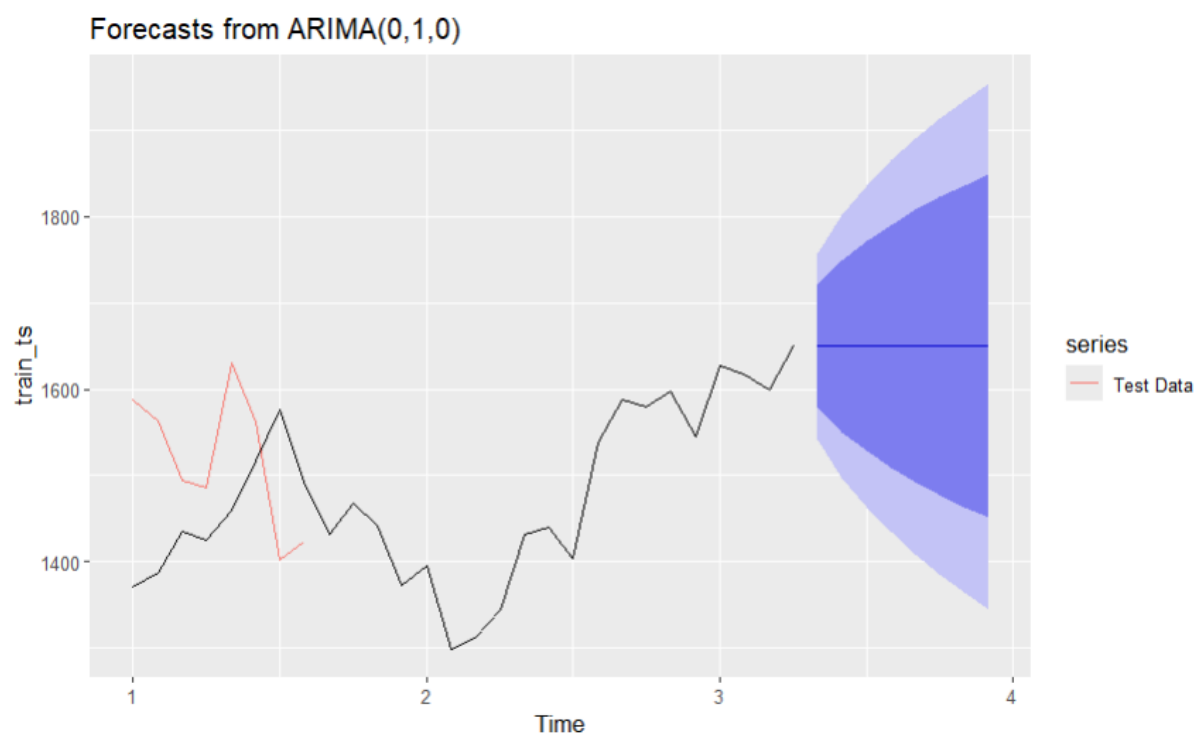
### DECOMPOSITION OF MULTIPLICATIVE TIME SERIES:



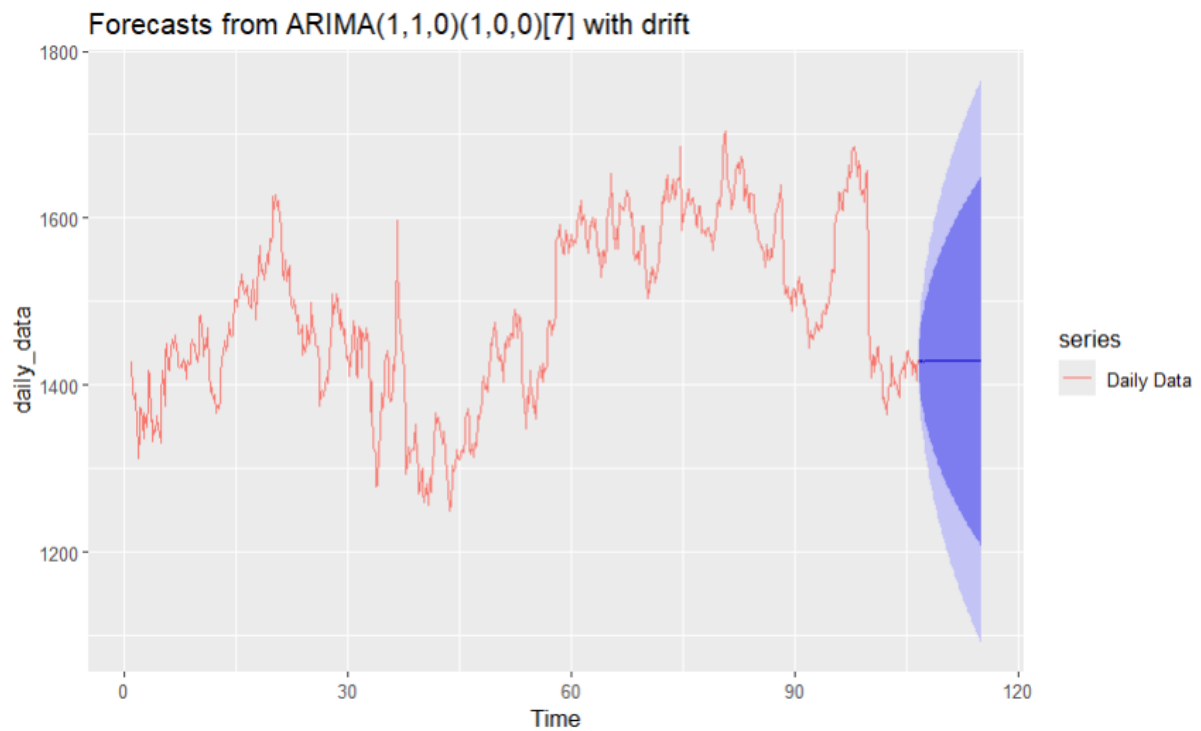
## HOLT-WINTERS FORECAST:



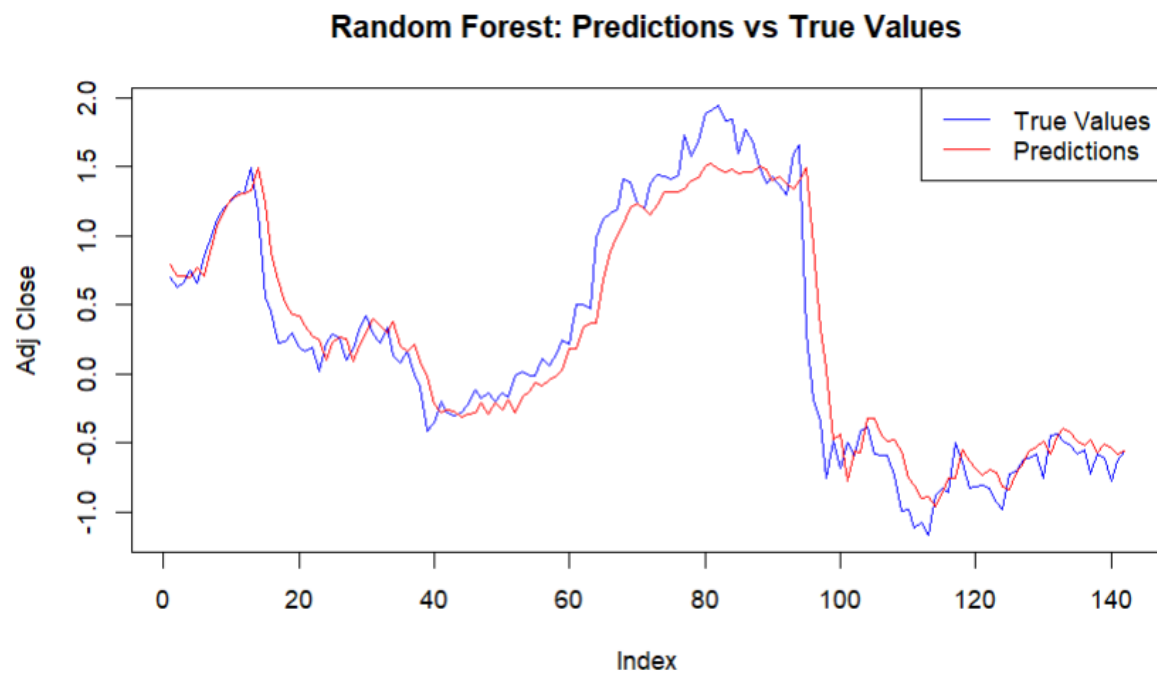
## ARIMA DAILY DATA:



## ARIMA DAILY FORECAST:



## RANDOM FOREST: PREDICTION VS TRUE VALUES:



## DECISION TREE & RANDOM FOREST: PREDICTIONS VS TRUE VALUES:

