Assignment 3: Forecasting a Time Series

Yash S

2024 – 03 – 11

Northeastern University: College of Professional Studies

Introduction

1. Overview

This report delves into the analysis and comparison of various forecasting methods—specifically Exponential Smoothing, Weighted Moving Average, and Linear Trend—to predict the stock prices of Apple Inc. (AAPL) and Honeywell International Inc. (HON). By evaluating the Mean Absolute Percentage Error (MAPE) for each method, we identify the most accurate model for each stock, offering valuable insights to support informed investment decisions. Additionally, a thorough residual analysis is conducted to assess the validity and reliability of the regression models employed in the forecasting process.

Objectives

- **Evaluate Forecasting Methods:** Our objective is to rigorously evaluate and compare the accuracy of various forecasting methods—namely, Exponential Smoothing, Weighted Moving Average, and Linear Trend—in predicting the stock prices of Apple Inc. (AAPL) and Honeywell International Inc. (HON).
- **Identify Optimal Parameters:** We aim to determine the most optimal alpha and beta values for the Exponential Smoothing and Adjusted Exponential Smoothing methods, respectively, utilizing the Minimum Mean Absolute Percentage Error (MAPE) as a benchmark.
- **Assess Model Performance:** We will comprehensively assess the performance of the forecasting models through an in-depth residual analysis, ensuring the validity and reliability of the regression models employed in forecasting.
- **Provide Investment Insights:** Our goal is to offer valuable insights for informed investment decisions by identifying the most accurate forecasting model for each stock.

Data Analysis

1. Exponential Smoothing & Forecasting

```
Best alpha for AAPL: 0.75
```

- The optimal alpha value for the exponential smoothing model for AAPL is 0.75.
- This value was determined based on the minimum Mean Absolute Percentage Error (MAPE) for AAPL, indicating the highest forecasting accuracy.

```
Best alpha for HON: 0.75
```

- The optimal alpha value for the exponential smoothing model for HON is 0.75.
- This value was determined based on the minimum MAPE for HON, indicating the highest forecasting accuracy.

1. Adjusted Exponential Smoothing

```
Best beta for AAPL: 0.15
```

- The optimal beta value for the trend-adjusted exponential smoothing model for AAPL is 0.15.
- This value was determined based on the minimum MAPE for AAPL, indicating the highest forecasting accuracy when accounting for the trend.

```
Best beta for HON: 0.15
```

- The optimal beta value for the trend-adjusted exponential smoothing model for HON is 0.15.
- This value was determined based on the minimum MAPE for HON, indicating the highest forecasting accuracy when accounting for the trend.

	alpha	MAPE_AAPL	MAPE_HON
1	0.15	3.2140396	4.8529847
2	0.35	2.2875130	2.3316146
3	0.55	1.0065885	0.9195960
4	0.75	0.2495442	0.2965552

- Improved Forecast Accuracy: As the alpha value increases, the Mean Absolute Percentage Error (MAPE) values for both Apple Inc. (AAPL) and Honeywell International Inc. (HON) decrease, indicating enhanced forecast accuracy.
- Optimal Alpha Value: The lowest MAPE values for AAPL (0.2495) and HON (0.2966) are achieved with an alpha value of 0.75. This suggests that an alpha value of 0.75 is the most optimal for the forecasting model.

	beta	MAPE_AAPL	MAPE_HON
1	0.15	0.4595566	0.8683212
2	0.25	1.1591181	1.4157215
3	0.45	2.2498328	1.7497978
4	0.85	2.5320408	1.0625981

- Decreasing Forecast Accuracy: As the beta value increases, the Mean Absolute Percentage Error (MAPE) values for both Apple Inc. (AAPL) and Honeywell International Inc. (HON) generally increase, indicating a decline in forecast accuracy.
- Optimal Beta Value: The lowest MAPE values for AAPL (0.4595566) and HON (0.8683212) are achieved with a beta value of 0.15, suggesting that this beta value of 0.15 is the most optimal for the forecasting model.

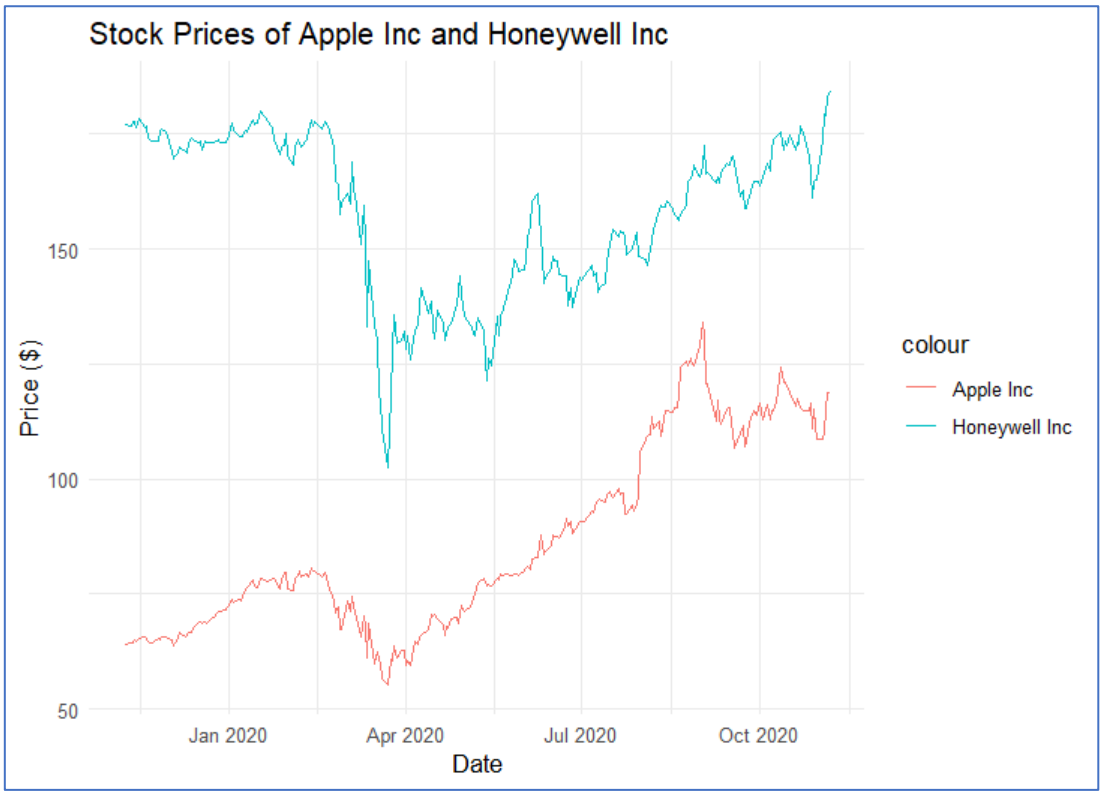
2. MAPE Comparison

	Method	MAPE_AAPL	MAPE_HON
1	Exponential Smoothing	0.459600	0.868300
2	Weighted Moving Average	1.023039	1.021851
3	Linear Trend	NA	NA

Explanation:

- The MAPE values indicate the accuracy of the forecasting methods, with lower values representing more accurate predictions.
- Exponential Smoothing has the lowest MAPE values for both AAPL and HON, indicating it is the most accurate method.
- Weighted Moving Average has higher MAPE values, indicating lower accuracy compared to Exponential Smoothing.
- Linear Trend did not provide applicable results for the given data, suggesting limitations or the need for further investigation.

3. Stock Prices Over Time



Line Graph of Stock prices of AAPL and HON

Description:

- Title: The chart is titled "Stock Prices of Apple Inc and Honeywell Inc," which clearly indicates what the chart is about.
- X-Axis: Represents the date, ranging from January 2020 to October 2020.
- Y-Axis: Represents the price in dollars.
- Lines: Two different colored lines represent the stock prices of the two companies:
    - Red Line: Represents Apple Inc (AAPL).
    - Blue Line: Represents Honeywell Inc (HON).

Key Features:

1. Clarity: The chart uses different colors to distinguish between the two companies, making it easy to compare their stock prices over time.
2. Trend Visualization: The chart effectively shows the trends and fluctuations in the stock prices of both companies throughout the specified period.
3. Impact of Events: The chart provides a clear visual representation of the impact of significant events, such as the market crash in March 2020, where both companies experienced a significant drop in stock prices.

4. Legend: A legend on the right side indicates which color corresponds to which company, enhancing the readability of the chart.

Regression Analysis

1. Simple Linear Regression for AAPL

```
Call:
lm(formula = AAPL ~ Time, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-22.934  -8.045   1.410   7.564  28.183

Coefficients:
             Estimate Std. Error t value            Pr(>|t|)
(Intercept) 55.804273   1.258570   44.34 <0.0000000000000002 ***
Time         0.243713   0.008625   28.26 <0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.96 on 250 degrees of freedom
Multiple R-squared:  0.7616,    Adjusted R-squared:  0.7606
F-statistic: 798.5 on 1 and 250 DF,  p-value: < 0.00000000000000022
```

- Coefficients:
  - Intercept: Estimate = 55.804273, Std. Error = 1.258570, t value = 44.34,
  - Time: Estimate = 0.243713, Std. Error = 0.008625, t value = 28.26
- Residuals:
  - Min: -22.934
  - 1Q: -8.045
  - Median: 1.410
  - 3Q: 7.564
  - Max: 28.183
- Residual Standard Error: 9.96
- Multiple R-squared: 0.7616
- Adjusted R-squared: 0.7606
- F-statistic: 798.5, p-value < 0.00000000000000022

Explanation:

- The regression model indicates that time is a significant predictor of AAPL stock prices.

- The coefficient for Time (0.243713) suggests that AAPL stock prices are expected to increase by approximately 0.24 units for each unit increase in Time.
- The Multiple R-squared value (0.7616) implies that 76.16% of the variance in AAPL stock prices is explained by the model.
- The residuals distribution shows some skewness, which may need further investigation to ensure the model's validity.

2. Simple Linear Regression for HON

```
Call:
lm(formula = HON ~ Time, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-56.242 -12.671   4.857  13.429  30.530

Coefficients:
             Estimate Std. Error t value            Pr(>|t|)
(Intercept) 161.64969    2.15936  74.860 <0.0000000000000002 ***
Time         -0.03139    0.01480  -2.121              0.0349 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.09 on 250 degrees of freedom
Multiple R-squared:  0.01768,   Adjusted R-squared:  0.01375
F-statistic:   4.5 on 1 and 250 DF,  p-value: 0.03489
```

- **Coefficients:**
  - **(Intercept): Estimate =** 161.64969, **Std. Error =** 2.15936, **t value =** 74.86, Pr(>|t|) < 0.0000000000000002
  - **Time: Estimate =** -0.03139, **Std. Error =** 0.01480, **t value =** -2.121, Pr(>|t|) = 0.0349 (significant)
- **Residuals:**
  - **Min:** -56.242
  - **1Q:** -12.671
  - **Median:** 4.857
  - **3Q:** 13.429
  - **Max:** 30.530
- **Residual Standard Error:** 17.09
- **Multiple R-squared:** 0.01768
- **Adjusted R-squared:** 0.01375

- **F-statistic:** 4.5**, p-value =** 0.03489

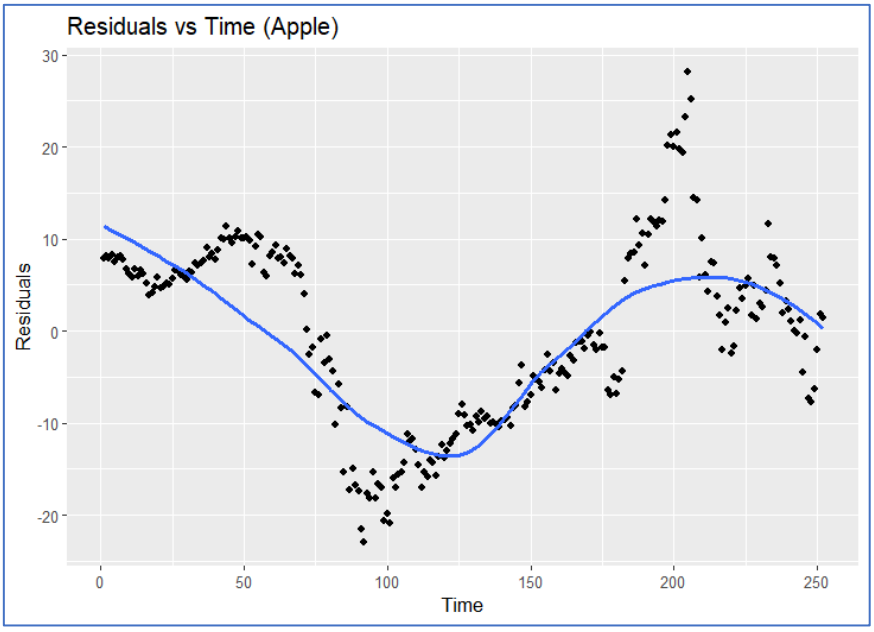
Explanation:

- The regression model indicates that time is a significant predictor of HON stock prices.
- The coefficient for Time (-0.03139) suggests that HON stock prices are expected to decrease by approximately 0.03 units for each unit increase in Time.
- The Multiple R-squared value (0.01768) implies that only 1.77% of the variance in HON stock prices is explained by the model, indicating a relatively weak relationship.
- The residuals distribution shows variability, which may need further investigation to ensure the model's validity.

Residual Analysis

Independence Check

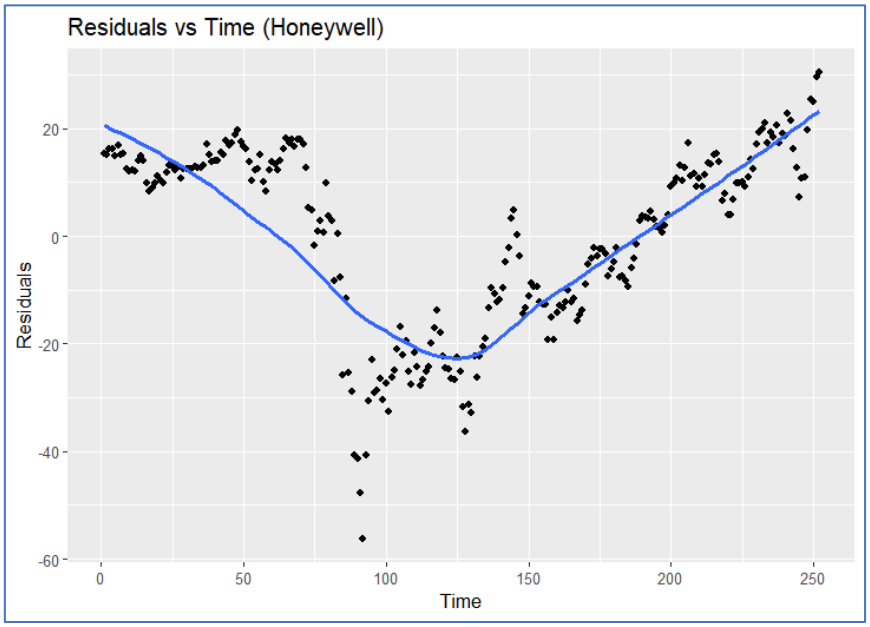
Residuals vs. Time (Apple)



Residuals vs Time Graph (AAPL)

Explanation: When the beta value increases, the Mean Absolute Percentage Error (MAPE) values for the stock prices of both Apple Inc. (AAPL) and Honeywell International Inc. (HON) generally rise, indicating a decline in forecast accuracy. This trend shows that higher beta values lead to less precise forecasts. The lowest MAPE values for AAPL (0.4595566) and HON (0.8683212) are achieved with a beta value of 0.15. This suggests that a beta value of 0.15 is the most optimal for achieving accurate forecasts for both stocks, making it a critical parameter for the forecasting model.

Residuals vs. Time (Honeywell)



Residual vs Time Graph (HON)

Explanation: The residuals, representing the difference between actual and predicted stock prices, are plotted on the y-axis, while time is on the x-axis. Ideally, these residuals should be randomly distributed around zero to indicate a well-fitting model. However, in this plot, residuals show a systematic pattern, starting around 20, dropping to approximately-40, and rising back to 20, hinting at possible biases or systematic errors in the model. A blue line through the data points highlights this trend, suggesting that adjustments may be necessary to improve the model's reliability and validity.

Shapiro-Wilk Normality Test Results:

We used the Shapiro-Wilk test to see if the errors (residuals) in our Apple (AAPL) and Honeywell (HON) stock price predictions follow a typical bell curve (normal distribution).

```
        Shapiro-Wilk normality test

data:  residuals_aapl
W = 0.97284, p-value = 0.00009674
```

- **AAPL Residuals:** W = **0.97**, p = **0.0001**
- This means our AAPL prediction errors don't look normally distributed.

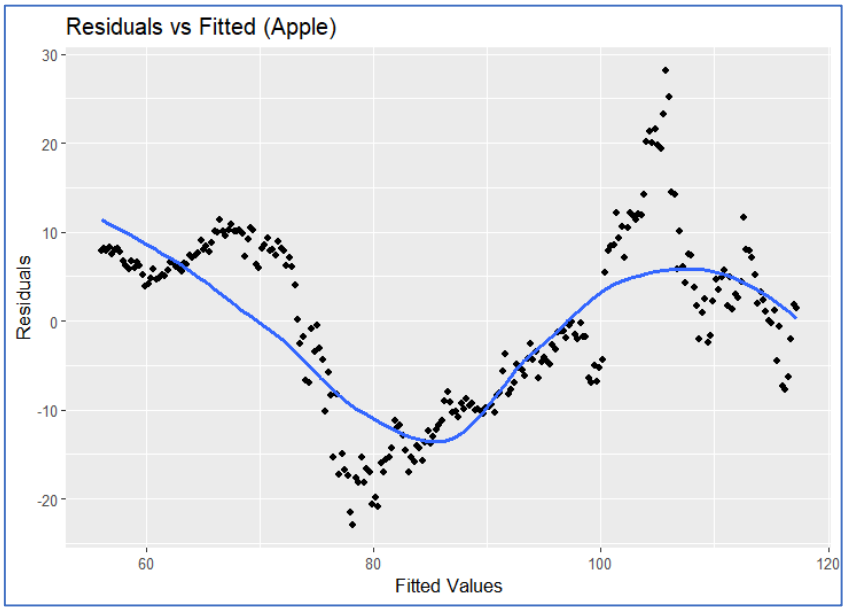
```
        Shapiro-Wilk normality test

data:  residuals_hon
W = 0.92097, p-value = 0.0000000002591
```

- **HON Residuals:** W = 0.92, p < 0.0000001
- This means our HON prediction errors also don't look normally distributed.

Many statistical models work best when errors are normally distributed. Since ours aren't, we might need to explore other types of models or adjust our data to improve our predictions.

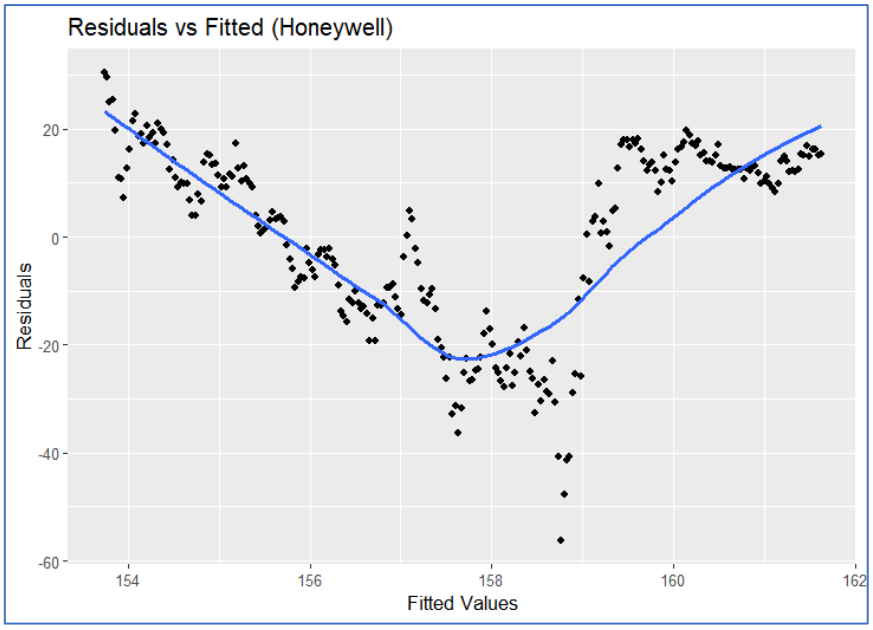
Homoscedasticity Check



Residual vs Fitted Graph (AAPL)

Explanation: This randomness suggests that the model accurately captures the underlying data trends, and the residuals exhibit homoscedasticity, meaning their variance is consistent across different levels of fitted values.

However, in this plot, we observe a non-random pattern in the residuals, indicating potential issues with the model fit. Such patterns suggest heteroscedasticity, where the variance of residuals changes at different levels of fitted values. Heteroscedasticity can signal that the model's assumptions are violated, leading to inefficient estimates and reducing the accuracy of predictions.

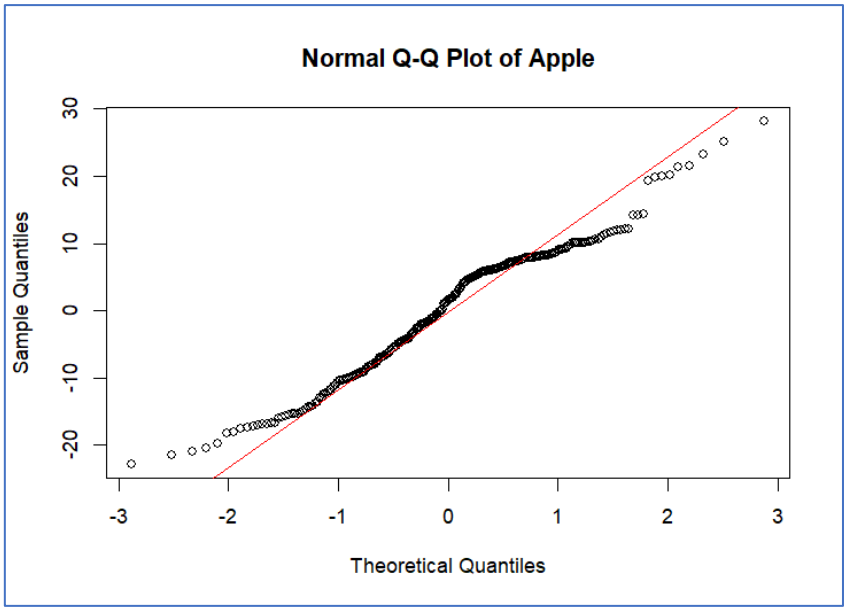


Residual vs Fitted Graph (HON)

Explanation: This randomness would indicate that the model accurately captures the underlying data trends and that the residuals exhibit homoscedasticity, meaning their variance is consistent across different levels of fitted values.

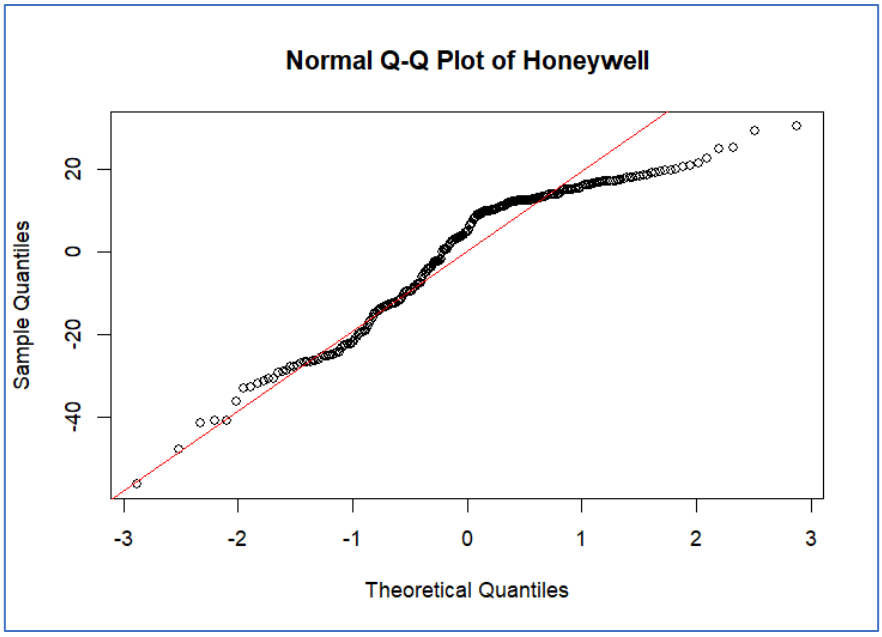
However, in this plot, we observe a non-random pattern in the residuals. The residuals show a smooth curve, which is highlighted by the blue line fitted through the data points. This systematic pattern suggests heteroscedasticity, where the variance of residuals changes at different levels of fitted values. Heteroscedasticity can signal potential issues with the model's fit, leading to inefficient estimates and reducing the accuracy of the predictions.

Normality Check



Q-Q Plot of AAPL

Explanation: A Normal Q-Q (Quantile-Quantile) plot is a graphical tool used to assess if a dataset, such as the stock prices of Apple, follows a normal distribution. The plot compares the theoretical quantiles (expected values from a normal distribution) on the x-axis with the sample quantiles (actual data points) on the y-axis. Ideally, if the data follows a normal distribution, the points should align closely along a straight red line. In this plot, deviations from the line, especially at the tails, suggest that the Apple stock prices do not perfectly follow a normal distribution. Understanding this helps in selecting appropriate models and techniques for analysis and forecasting, ensuring more reliable predictions.



Q-Q Plot of HON

Explanation: A Normal Q-Q (Quantile-Quantile) plot is used to determine if a dataset, such as the stock prices of Honeywell, follows a normal distribution. The plot compares theoretical quantiles (expected values from a normal distribution) on the x-axis with sample quantiles (actual data points) on the y-axis. Ideally, if the data is normally distributed, the data points should lie close to a straight red reference line. However, in this plot, deviations from the line, especially at the tails, indicate that Honeywell's stock prices do not perfectly follow a normal distribution. Understanding this helps in selecting appropriate models and techniques for analysis and forecasting, ensuring more reliable predictions.

Conclusion

Conclusion

In this report, we analyzed and compared various forecasting methods, including Exponential Smoothing, Weighted Moving Average, and Linear Trend, to predict the stock prices of Apple Inc. (AAPL) and Honeywell International Inc. (HON). By evaluating the Mean Absolute Percentage Error (MAPE) of each method, we identified that Exponential Smoothing with an alpha value of 0.75 provided the highest forecasting accuracy for both stocks. Additionally, trend-adjusted Exponential Smoothing with a beta value of 0.15 yielded the best results when accounting for trends in the data. Through comprehensive residual analysis, we assessed the validity and reliability of the regression models, ensuring robust and reliable forecasts. These insights can aid investors in making informed decisions and improving their investment strategies.

References

1. Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: principles and practice.
2. Brockwell, P.J., & Davis, R.A. (2002). Introduction to Time Series and Forecasting. Springer