**Introduction to Enterprise Analytics**

**Module 3 Project: Forecasting Financial Time Series**

**ALY 6050**

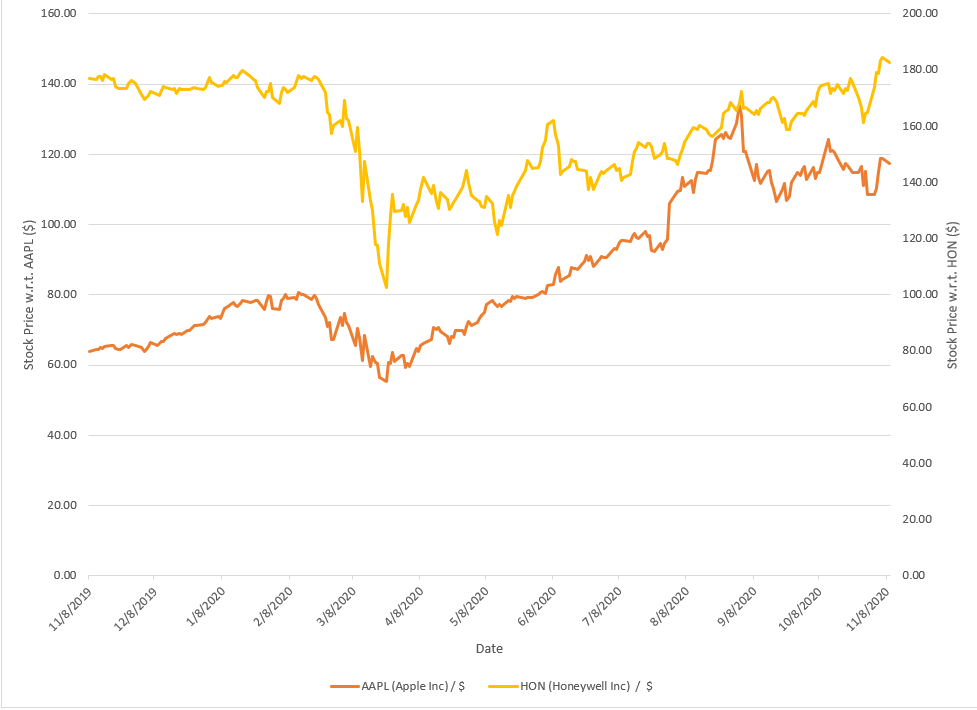
**Anukriti Baijal**

**Baijal.a@northeastern.edu**

**Introduction:**

The historical stock prices for Apple Inc. (AAPL) and Honeywell International Inc. (HON) for a period of one year, comprising 252 market days, are provided.

**PART 1:**



From the above trend chart, we observe the following:

For AAPL:

* There is steady increase in the stock prices from November 2019 to February 2020. However, there is sudden decline between the months March and April 2020.
* There is a steady increase from April to August.
* A sudden increase is seen for the month of September 2020. It might be due to the release of new Apple products.

For HON:

* The stocks are steady for a period of few months, from November 2019 to February 2020.
* A sudden decline in stocks is seen between March and April 2020. This might be due to the lockdown implementation worldwide.
* The stocks stay quite volatile for the rest of the duration. However, there is a slight increase in the stocks overall.

Now we perform exponential smoothening using different values for the smoothening parameter:

|  |  |
| --- | --- |
| **AAPL** | |
| **α** | **MAPD** |
| 0.15 | 3.91 |
| 0.35 | 2.48 |
| 0.55 | 2.08 |
| 0.75 | 1.96 |

|  |  |
| --- | --- |
| **HON** | |
| **α** | **MAPD** |
| 0.15 | 2.84 |
| 0.35 | 2.19 |
| 0.55 | 1.90 |
| 0.75 | 1.77 |

For both the stocks, the most accurate results are obtained when the value of the exponential smoothening parameter is 0.75. We know this as the corresponding MAPD value is the least when compared. The range of possible alpha values is 0 to 1, inclusive. Because they average out changes over time, lower values give greater weight to historical observations and generate smoother fitted lines. Larger values limit the degree of averaging by the earlier data, which results in a more jagged line since they place a higher emphasis on the current data. In this case, both the stocks show erratic behavior for the most part. Therefore, higher value of the parameter produced better yield.

Next, we perform adjusted exponential smoothening using the α=0.55 and forecast values for the period 253 for both the stocks:

|  |  |
| --- | --- |
| AAPL | 117.50 |
| HON | 182.58 |

Further, we use different values for the trend parameter, β, for both the stocks:

|  |  |
| --- | --- |
| **AAPL** | |
| **β** | **MAPE** |
| 0.15 | 0.081 |
| 0.25 | 0.080 |
| 0.45 | 0.079 |
| 0.85 | 0.084 |

For AAPL stocks, β=0.45 yield the best results with Mean Absolute Percentage Error being 0.079%. Since it updates the trend component for each observation, this method mimics dynamic gradients. Like alpha, beta can range from 0 to 1, inclusive. Larger values provide more importance to recent observations, enabling the trend component to respond to changes in the trend more quickly.

|  |  |
| --- | --- |
| **HON** | |
| **β** | **MAPE** |
| 0.15 | 0.090 |
| 0.25 | 0.086 |
| 0.45 | 0.081 |
| 0.85 | 0.077 |

For HON stocks, β=0.85 yield the best results with Mean Absolute Percentage Error being 0.077%. This might be because the HON is more volatile than AAPL as seen in the graph above. Therefore, higher value of β yields more accurate result when compared to the β for AAPL.

**PART 2:**

Now, we calculate the weighted moving average for the period 1 to 100. We use 0.5 for the most recent stock, 0.3 for the previous period and 0.2 for stock price two periods ago as the weights for the calculations.

Then, we use the observed value for the period 101 as the base for the linear trend forecasting to forecast stock prices for periods 102 through 257. Below are the actual and predicted stock prices during those periods.

**AAPL:**

**Actual Values:**

Graphical user interface

Description automatically generated

**Predicted:**

|  |  |  |
| --- | --- | --- |
| **Date** | **Period** | **Predicted Value** |
| 11/13/2020 | 257 | |  | | --- | | 124.30 | |
| 11/12/2020 | 256 | 124.17 |
| 11/11/2020 | 255 | 124.07 |
| 11/10/2020 | 254 | 123.94 |
| 11/09/2020 | 253 | 123.78 |

As we can see, the predicted values are higher than the actual values. The actual values range from 115.97-119.49. However, the predicted values range from 123.78-124.30. This means that the prediction method is not accurate enough.

**HON:**

**Actual Values:**

Graphical user interface, text, application, email

Description automatically generated

**Predicted:**

|  |  |  |
| --- | --- | --- |
| **Date** | **Period** | **Predicted Value** |
| 11/13/2020 | 257 | |  | | --- | | 179.68 | |
| 11/12/2020 | 256 | 179.29 |
| 11/11/2020 | 255 | 178.88 |
| 11/10/2020 | 254 | 178.41 |
| 11/09/2020 | 253 | 177.94 |

As we can see, the predicted values are lower than the actual values. The actual values range from 196.99-201.98. However, the predicted values range from 177.94-179.68. This means that the prediction method is not accurate enough.

Next, we compare the Mean Absolute Percentage Error for both the stocks:

|  |  |  |
| --- | --- | --- |
| **Stock** | **MAPE (Part 1)** | **MAPE (Part 2)** |
| AAPL | 0.079 | 5.50 |
| HON | 0.077 | 3.65 |

Comparing the MAPEs for both the stocks, the most accurate method of forecasting was the adjusted exponential smoothening with the MAPE being 0.079% for AAPL stocks and 0.077% for HON stocks. Whereas, when we look at the linear trend forecasting, the MAPEs are much higher at 5.50% for AAPL and 3.65% for HON.

**PART 3:**

Now, we perform simple regression of stock prices versus the time period for both the stocks:

Graphical user interface, table

Description automatically generated with medium confidence

Therefore, the regression equation is as follows:

AAPL:

HON:

Using the above equations, we forecast values for periods 1 to 257. Now, we compare these with the predicted values from part 1 and 2:

AAPL:

|  |  |  |  |
| --- | --- | --- | --- |
| Actual Values | Part 1 | Part 2 | Part 3 |
| 119.26 | 119.24 | |  | | --- | | 124.30 | | 118.44 |
| 119.21 | 118.01 | 124.17 | 118.19 |
| 119.49 | 113.42 | 124.07 | 117.95 |
| 115.97 | 109.35 | 123.94 | 117.71 |
| 116.32 | 108.30 | 123.78 | 117.46 |

As we can see from the above table, Part 3, that is the simple regression, has the most accurate predictions. This means that the difference between the predicted values and the actual value is the least for simple regression performed in part 3.

HON:

|  |  |  |  |
| --- | --- | --- | --- |
| Actual Values | Part 1 | Part 2 | Part 3 |
| 201.54 | 184.83 | |  | | --- | | 179.68 | | 153.58 |
| 197.24 | 183.78 | 179.29 | 153.61 |
| 199.29 | 179.74 | 178.88 | 153.65 |
| 201.98 | 179.99 | 178.41 | 153.68 |
| 196.99 | 173.76 | 177.94 | 153.71 |

From the above table, the adjusted exponential smoothening predicted the most accurate result out of all the predictions. This means that the difference between the predicted values and the actual value is the least for the method performed in part 1.

Next, we perform residual analysis for the simple regression:

AAPL:

Graphical user interface, chart, application, line chart

Description automatically generated

From the above graph for Period Residual Plot, we do not notice random scattering of points. However, the points are not clustered in one location, rather spread out in a line. Therefore, the residuals are not independent. Also, from the graph Period Line Fit, we can see that the line is straight. There is not pattern that is observed. This means that the variance is quite constant, so we do not fail homoscedasticity for this data. Also, the in the Normal Probability Plot, the line is straight. This is the behavior that would be expected if the residuals were truly normally distributed.

Table

Description automatically generated with medium confidence

Since, the p-value=1, we can state with 95% confidence that the data is normally distributed.

HON:

Chart, line chart

Description automatically generated

The Period Residual Plot graph shown above does not show a random distribution of points. The points, however, are dispersed over a line rather than grouped together in one spot. The residuals are therefore not independent. Also, we can observe that the line is straight from the Period Line Fit graph. There is no pattern. As a result, we do not violate homoscedasticity for this data because the variance is rather constant. Moreover, the line is straight in the normal probability plot. If the residuals were normally distributed, then this behavior would be anticipated.

Table

Description automatically generated with medium confidence

Since, the p-value=1, we can state with 95% confidence that the data is normally distributed.

**QUESTION:**

While both the stocks AAPL and HON have a good value, however the distribution of investment suitable, in my opinion, would be 70-75% in AAPL and the remaining 25-30% in HON. This is because the fluctuations in HON is more as compared to AAPL. Also, there much more prominent growth in the stocks of AAPL when compared to HON. Further, when we look at the standard deviations of the residuals, it is much greater for HON than for AAPL. Therefore, more investment in AAPL would be beneficial.

**References:**

[1] Mark (February 3, 2021). *How to Run a Normality Test in Excel: Chi-Square Goodness-of-Fit.* Retrieved on March 12, 2023, from <https://turbofuture.com/computers/excels-easiest-and-most-robust-normality-test-the-chi-square-goodness-of-fit-test>

[2] Excel Easy. *Forecast*. Retrieved on March 11, 2023, from <https://www.excel-easy.com/examples/forecast.html>

[3] Holland, G. *Easy Forecasting in Excel*. Retrieved on March 11, 2023, from <https://absentdata.com/excel/excel-forecasting/>

[4] Zach (March 6, 2020). *How to Find Weighted Moving Averages in Excel.* Retrieved on March 10, 2023, from <https://www.statology.org/weighted-moving-average-excel/>

[5] Zach (February 27, 2020). *How to Calculate Mean Absolute Percentage Error (MAPE) in Excel*. Retrieved on March 10, 2023, from <https://www.statology.org/mape-excel/>