Logo

Description automatically generated

**College of Professional Studies**

**Northeastern University San Jose**

**MPS Analytics**

**Course: ALY6015: Introduction to Enterprise Analytics**

**Assignment:**

Module 3 Project-  Forecasting a Time Series

**Submitted on:**

March 12, 2023

**Submitted to:**  **Submitted by:**

Prof: AZADEH MOBASHER NIKSHITA RANGANATHAN

# **ABSTRACT**

Time series forecasting is a statistical method that employs past data to forecast future values for a time series.

Components of time series forecasting:

* The underlying trend of a time series is the general direction in which the data is moving over time.
* Seasonality refers to the recurring patterns or cycles in a time series that are repeated over a fixed period of time, such as weekly, monthly, or yearly.
* The cyclical component of a time series represents fluctuations that occur over long periods of time, such as business cycles or economic cycles.
* The irregular or random component of a time series represents random fluctuations that cannot be explained by the other components, such as unexpected events or external factors.

There are several advantages of time series forecasting, including:

* It helps to improve the accuracy of future predictions by analyzing historical data and identifying trends and patterns. This can lead to better decision making and cost savings for businesses.
* It can be used as an early warning system to identify potential problems or opportunities in advance, enabling companies to take timely action.
* Accurate forecasts enable better planning of resources, such as inventory management, production schedules, and staffing levels.
* It can help businesses anticipate customer demand and provide better customer service by ensuring that products are available when customers need them.
* In order to remain ahead of the competition and react more quickly to market changes, businesses can gain a competitive advantage by using accurate forecasting.
* This technique can be used to forecast future financial metrics, such as revenue, profit, and cash flow, helping businesses to make more informed budgeting and financial planning decisions.

Time series forecasting is commonly used in finance, economics, meteorology, and many other fields where historical data can be used to predict future trends and patterns.

**INTRODUCTION**

The goal of this project is to examine the stock prices of AAPL and HON for a year, which comprises 252 market days. It is divided into three sections: short-term forecasting, long-term forecasting, and regression analysis.

**Part 1: Short-term Forecasting:**

This part involves performing exponential smoothing to forecast both stock prices and determining the most accurate values of the smoothing and trend parameters α and β, respectively.

**Part 2: Long-term Forecasting:**

In part 2, we will use a weighted moving average method and a linear trend method to forecast the stock prices.

**Part 3: Regression:**

We will utilize simple regression to predict the stock prices and perform a residual analysis to evaluate if regression is a suitable method for analyzing the given data.

Table

Description automatically generated

**Figure 1- Table with historical stock prices for AAPL and HON**

**ANALYSIS & INTERPRETATION**

**Part 1: Short-term Forecasting:**

1. **Graphs**

**Chart, line chart

Description automatically generated**

**Figure 2- Time Series Apple**

AAPL's time series indicates a predominantly upward trend in stock prices from November 8, 2019, to October 8, 2020, with some irregular fluctuations. There is a decrease in stock prices in the month of March but it experienced a sharp decline, indicating a sudden change in trend. After that, the stock prices started moving in an upward direction again. The time series is not significant enough for us to interpret any cyclical and seasonal behavior in the data as the data is relatively short.

**Chart, line chart

Description automatically generated**

**Figure 3- Time Series Honeywell**

It seems that the time series of Honeywell International Incorporations (HON) stock prices exhibits a mixed trend with both upward and downward movements. The trend appears to be relatively steady initially. Similar to the previous graph, due to the limited time period of the available data, it is difficult to confirm whether or not the seasonal and cyclic behavior is actually present.

1. **Exponential smoothing**

Exponential smoothing is a forecasting method used in time series analysis to forecast future values based on past data. It is an effective method that can handle trends and seasonality in the data.

In simple exponential smoothing, the forecast for next period is obtained by averaging the previous forecast and the current actual value. The formula is:

Ft = α \* At-1 + (1 - α) \* Ft-1

where:

Ft is the forecast for the next period

At-1 is the actual value for the current period

Ft-1 is the forecast for the current period

α is the smoothing factor, which is a value between 0 and 1 that determines the weight given to the most recent observation. A higher α value gives more weight to the most recent observation, while a lower α value gives more weight to past observations.

MAPD stands for Mean Absolute Percentage Deviation, which is a measure of accuracy for forecasting models. It calculates the percentage difference between the forecasted and actual values but uses the sum of actual values as the denominator.

The formula for MAPD is:

MAPD = Σ(|Actual - Forecast|) / Σ(Actual)) \* 100

Table

Description automatically generated

Table

Description automatically generated

**Figure 4- Exponential smoothing and MAPD values for different α values (Apple)**

**Exponential smoothing values :**

* 114.88 (α = 0.15)
* 115.97 (α = 0.35)
* 117.50 (α = 0.55)
* 118.39 (α = 0.75)

**MAPD Values :**

* 3.91% (α = 0.15)
* 2.48% (α = 0.35)
* 2.08% (α = 0.55)
* 1.96% (α = 0.75)

The results show that the alpha value of 0.75 has yielded the most accurate forecast for Apple stock, with a MAPD value of 1.96%. This means that the forecasted values were, on average, only 1.96% away from the actual values.

**Table

Description automatically generated**

**Table

Description automatically generated**

**Figure 5- Exponential smoothing and MAPD values for different α values (Honeywell)**

**Exponential smoothing values :**

* 175.33 (α = 0.15)
* 179.97 (α = 0.35)
* 182.58 (α = 0.55)
* 183.72 (α = 0.75)

**MAPD Values :**

* 2.84% (α = 0.15)
* 2.19% (α = 0.35)
* 1.90% (α = 0.55)
* 1.77% (α = 0.75)

Based on the results, it can be concluded that the most accurate forecast for Honeywell stock was achieved using an alpha value of 0.75, which resulted in a MAPD value of 1.77%.

It is important to note that as the alpha value increases, there is a decrease in MAPD. This is because a higher alpha value assigns more weight to recent observations, which are closer in time to the actual value, leading to a more accurate forecast.

In my opinion, the alpha value of 0.75 yielded the most accurate forecast for both stocks because it gave the highest weight to the most recent observations, which are likely to be more indicative of the future trend. Additionally, the high weight on recent observations allows the forecast to adjust quickly to any sudden changes or trends in the data.

1. **Adjusted Exponential smoothing**

Adjusted Exponential Smoothing is a forecasting method that uses a combination of level and trend smoothing to make predictions about future values based on historical data. It is an extension of the basic Exponential Smoothing method and takes into account the trends and seasonality in the data.

Tt = β\*(Ft – Ft-1) + (1-β)\*Tt-1

Where:

Ft is the forecasted value for the next period

Ft-1 is the forecast for the current period

Tt-1 is the trend component for the current period

β is the smoothing parameter for the trend

MAPE measures the average absolute percentage difference between the actual values and the forecasted values.

The formula for calculating MAPE is as follows:

MAPE = (1/n) \* Σ(|A - F| / A) \* 100%

**Table

Description automatically generated**

**Table

Description automatically generated**

**Figure 6- Adjusted Exponential smoothing and MAPE values for different** **β values (Apple)**

**Adjusted Exponential smoothing values :**

* 118.04 (β = 0.15)
* 118.54 (β = 0.25)
* 119.24 (β = 0.45)
* 119.22 (β = 0.85)

**MAPE Values :**

* 1.96% (β = 0.15)
* 1.94% (β = 0.25)
* 1.93% (β = 0.45)
* 1.98% (β = 0.85)

For Apple, the value of β that provides the most accurate forecast is 0.45, with a corresponding MAPE of 1.93%. This indicates that past observations played a more important role in predicting future values. This may be due to the fact that Apple's stock price tends to be more stable and less susceptible to sudden market shifts.

**Table

Description automatically generated**

**Table

Description automatically generated**

**Figure 7- Adjusted Exponential smoothing and MAPE values for different** **β values (Honeywell)**

**Adjusted Exponential smoothing values :**

* 184.07 (β = 0.15)
* 184.72 (β = 0.25)
* 185.19 (β = 0.45)
* 184.83 (β = 0.85)

**MAPE Values :**

* 1.97% (β = 0.15)
* 1.93% (β = 0.25)
* 1.88% (β = 0.45)
* 1.84% (β = 0.85)

In the case of Honeywell, the highest accuracy is achieved with β=0.85, indicating that the most recent observations have a higher influence on the forecast. This makes sense for a company like Honeywell, where the demand for its products may change quickly based on market conditions or economic trends.

**Part 2: Long-term Forecasting:**

1. **3-period weighted moving averages and linear trend**

Weighted moving averages is a time series forecasting method that gives more weight to the more recent observations and less weight to the older observations. This method calculates the forecast by taking a weighted average of the previous data points, where the weights assigned to each data point decrease exponentially as the data point becomes further in the past.

In this case, we will be considering 3-period weighted moving average with weights 0.5,0.3, and 0.2.

Once the first 100 periods have been forecasted using the weighted moving average, the observed values from period 101 onwards are used to fit a linear trend. This is done using the TREND function in Excel. The least squares technique is employed by this function to identify the line that best fits the given set of data.

The long-term forecasted values for Apple for periods 253 to 257 were 128.02, 128.42, 128.81, 129.21, and 129.61, respectively.

Honeywell's forecasted values for the long term, for periods 253 to 257, are as follows: 176.05, 176.35, 176.66, 176.96, and 177.27.

**Chart, line chart

Description automatically generated**

**Figure 8- Long term Forecasting(Apple and Honeywell)**

For both Honeywell and Apple, the forecasted values calculated using weighted moving averages closely mirror the observed values and exhibit a similar pattern. On the other hand, the trend-based forecasted values follow a straight line.

**Timeline

Description automatically generated**

**Figure 9- Actual close values(Apple and Honeywell)**

For Apple, the predicted values were consistently higher than the actual close values, indicating that the method of forecasting overestimated the stock price.

On the other hand, the predicted values were very close to the actual close values, indicating that the forecasting method was fairly accurate for this stock.

1. **MAPE**

**Table

Description automatically generated**

**Figure 10- Long term forecasting MAPE (Apple and Honeywell)**

MAPE for Apple and Honeywell using this method is 3.86% and 2.86% respectively.

A lower MAPE value indicates a more accurate forecast, and in comparison with Part 1, we can say that the adjusted exponential smoothing method has yielded more accurate forecasts compared to the other methods for both stocks.

**Part 3: Regression:**

1. **Simple Regression**

The goal of regression analysis is to estimate the effect of the independent variables on the dependent variable and make predictions of the dependent variable based on the values of the independent variables.

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

**Figure 11- Graphs with Predicted Y values, equation and R^2 (Apple and Honeywell)**

Table

Description automatically generated

**Figure 12- Slope, Intercept, R, R^2, Mean, and SD of residual (Apple and Honeywell)**

The table provides statistical information on linear regression analysis for Apple and Honeywell stocks.

For Apple, the slope of the regression line is 0.24, and the intercept of the regression line is 55.80, which represents the estimated stock price at time 0. The correlation coefficient (R) is 0.88, indicating a strong positive correlation between the period and stock price. The coefficient of determination (R2) is 0.77, indicating that 77% of the variation in the stock price can be explained by the time variable.

According to the analysis for Honeywell, there is a weak negative slope (-0.01) indicating a slight downward trend in the stock price over the period. The intercept of the regression line is 159.91. The correlation coefficient (R) is -0.05, showing a weak negative association between time and stock price. The coefficient of determination (R2) is 0.00, which implies that there is no significant linear relationship between time and stock price.

The mean of the residuals is 0.00 for both stocks, indicating that the regression line is a good fit for the data. The standard deviation of the residuals is 9.84 for Apple and 17.97 for Honeywell, indicating that the data points are more spread out for Honeywell than for Apple.

Table

Description automatically generated

**Figure 13- Regression MAPE(Apple and Honeywell)**

MAPE values obtained from the regression analysis are higher than those from other methods such as exponential smoothing and 3-period weighted moving averages and linear trend, which could suggest that the regression model is not the most appropriate method for predicting future values of the dependent variable.

1. **Checking for properties**

* **Whether the residuals are independent**

**Chart, scatter chart

Description automatically generated**

**Figure 14- Plots for checking independence (Apple and Honeywell)**

The Residuals vs Period plots indicate that the residuals for both Apple and Honeywell are not independent and display a trend, indicating a relationship between the residuals and the variables. This violates the independence assumption required for regression analysis

* **Whether the residuals are homoscedastic**

**Chart, scatter chart

Description automatically generated**

**Figure 15- Plots for checking Homoscedasticity (Apple and Honeywell)**

The above plots for both the stocks indicate that the residuals are not homoscedastic in nature, as they are distributed unevenly and their variance does not appear to be constant. This suggests that the assumption of constant variance required for regression analysis may not be met.

* **Whether the residuals are normally distributed**

**Chart, line chart

Description automatically generated**

**Figure 16- Plots for checking Normal distribution (Apple and Honeywell)**

These plots suggest that the residuals do not follow a normal distribution since they are not closely aligned with the observed values.

* **Whether the residuals are normally distributed – Chi-Square test**

**Diagram

Description automatically generated with low confidenceTable

Description automatically generated**

**Figure 17- Chi-Square tests (Apple and Honeywell)**

**Null hypothesis:** The observed data follows a normal distribution.

**Alternative hypothesis:** The observed data does not follow a normal distribution.

**The alpha value is 0.05**

The P-value for both stocks **– Apple and Honeywell**, is less than the significance level of 0.05. Therefore, we reject the null hypothesis at the 95% confidence level.

Therefore, we can conclude that there is sufficient evidence to suggest that the observed data significantly deviates from a normal distribution. In other words, **the data does not follow normal distribution**.

**Portfolio 𝜫 (Pi)**

Determining the appropriate allocation of investment between two stocks (AAPL and HON) in a portfolio requires careful consideration of various factors such as risk tolerance, investment goals, and market conditions.

Based on the analysis of the time series data for both stocks, it is evident that Apple stock is more stable and has exhibited a consistent uptrend in the past. On the other hand, Honeywell's stock price has been more volatile and has shown a mixed trend.

Given these observations, I would suggest an allocation of 60% of the portfolio to Apple stock and 40% to Honeywell stock. This allocation would allow for a good balance between stability and risk. The higher allocation to Apple stock would ensure that the portfolio has a steady and consistent return and the remaining portion of the portfolio can be used to take advantage of the potentially higher returns associated with Honeywell stock. AAPL is a technology-based stock and has more potential for growth compared to HON, which is a more established company.

Additionally, I would suggest that the portfolio should be regularly monitored and altered as needed to ensure that the desired risk-return balance is maintained.

**CONCLUSION**

In conclusion, it is evident that predicting stock prices is a complex task, and there are many factors that need to be taken into consideration when making decisions about investments.

It was found that the Exponential Smoothing method provided better and more accurate forecasts compared to other methods

The residuals were found to be non-independent, non-homoscedastic, and non-normally distributed. The Chi-Square test also confirmed that the residuals significantly deviate from a normal distribution. Therefore, the regression models may not be the most appropriate method for predicting future values of the dependent variable.

Investing 60% of the portfolio into Apple shares and 40% into Honeywell shares would be a good choice. This range of investment would offer a balanced portfolio that reduces risk, taking into account the company's historical data, stock prices, and volatility.

**REFERENCES**

Bluman, A. G. (2018). Elementary Statistics, 10th ed. McGraw Hill.

Kabacoff, R. I. (2011). R in action: Data analysis and graphics with R. Manning Publications Co.

Home - RDocumentation. (n.d.). <https://www.rdocumentation.org/>