**About the project:**

In this project, we will focus on the analysis and forecasting of historical stock prices for Honeywell and Apple, spanning 252 market days. This project is divided into two parts, in the first part, we use excel to do the short-term forecasting and in the second part, we will use R to do the long-term forecasting.

In Part 1, we will initiate our exploration with Excel, visually inspecting the time series data through line plots to uncover any discernible trends, seasonality, or irregular patterns. Following this, we will apply exponential smoothing techniques to forecast stock prices for the upcoming 253rd period. Additionally, we employ adjusted exponential smoothing to enhance forecasts by introducing trend parameters β. Exponential smoothing applies a single smoothing factor to all historical data points, while adjusted exponential smoothing adjusts the smoothing factor over time to account for trends in the data through parameter β.

In part 2, we will leverage R for deeper time series scrutiny. Initially, we will download and analyze historical data spanning five years for HON and AAPL. Subsequently, we will fit an AR(1) time-series model to the data and gauge its effectiveness. We will then use Auto.Arima() function in R to find a time series model. This methodology will be extended to forecasting dry wine prices.

Let’s now dive in the project!

**Part 1: Short Term Forecasting**

We have been given stock prices of Apple and Honeywell for 252 days from 2019 and 2020 and we are supposed to predict the stock prices for 253rd day using short term forecasting. Let’s first see the prices of both the stocks given over the time.

Figure. 1 Line chart of Apple and Honeywell Stock Prices Over the years.

**Observations:**

It's noticeable that Honeywell's stock price is higher than Apple's throughout the given period. Additionally, Apple's stock price is generally on the rise, while Honeywell's shows a similar upward trend but with less obvious changes. Both stocks experienced a decline in the middle, but Honeywell's was more noticeable. There doesn't appear to be any seasonal patterns in either stock's prices, but there might be some underlying trends or non-stationarity, which we'll confirm through stationarity tests in the following sections.

Let’s now try to forecast both the stocks using Exponential smoothing.

**Exponential Smoothing:**

Exponential smoothing is a technique used in short-term forecasting where past observations are weighted exponentially to give more importance to recent data, helping to predict future values with reduced noise from older data. It's based on the assumption that recent data points are more relevant for forecasting than older ones. This can be done using the following formula:

|  |
| --- |
| **Fn+1 = α Dn + (1-α) Fn** |

Here alpha is the smoothing factor that determines the weight given to the most recent observation when forecasting future values. Dn is the observed value and Fn is the previous forecasted value.

We will be carrying out exponential smoothing for both the stocks using different alpha values like 0.15, 0.35, 0.55 and 0.75.

We will be measuring the accuracy of the forecasting using MAD and MAPE metrics.

MAD is Mean Absolute Deviation which measures the average difference between actual and forecasted values.

MAPE is Mean Absolute Percentage Error which measures the average percentage difference between actual and forecasted values.

**Exponential Smoothing Forecasting for Apple:**

A white rectangular box with red text

Description automatically generated

A white rectangular table with red text and numbers

Description automatically generated

**Observations:**

The forecasted stock price of Apple on the 253rd day is $114.88 with an alpha value of 0.15, indicating less weight is given to previous observed values. Increasing alpha leads to higher forecasted stock prices, along with decreasing MAD and MAPE, suggesting higher alpha values provide more accurate predictions. Among the given alpha values, using 0.85 yields the most accurate result, as both MAD and MAPE are minimized.IT can be observed that the

**Exponential Smoothing Forecasting for Honeywell:**

A white paper with red text

Description automatically generated

A white rectangular table with red text and numbers

Description automatically generated

**Observations:**

The forecasted value for Honeywell on the 253rd day is $175.33 with an alpha of 0.15, and it increases as the alpha value rises, like what we observed for Apple. It can be observed that MAD and MAPE, both the metrics, are showing similar trends. The lowest MAD and MAPE values occur with an alpha of 0.85, making it the best choice. However, it's important to note that compared to Apple, Honeywell has higher MAD and MAPE values, possibly due to greater stock price variations over time compared to the more stable Apple stock prices.

Let’s now do similar forecasting using adjusted exponential smoothing forecasting method.

**Adjusted Exponential Smoothing Forecasting for Apple:**

Adjusted exponential smoothing is a forecasting method that incorporates trend adjustments alongside exponential smoothing, using both a smoothing parameter (alpha) and a trend parameter (beta). Alpha determines the weight given to the most recent observation, while beta adjusts the smoothing of the trend component. This technique improves forecasts by accounting for both short-term fluctuations and longer-term trends in the data. The formula for the method is given by:

|  |
| --- |
| **Afn = Fn + Tn** |
| **Tn  = β (Fn - Fn-1) + (1-β) Tn-1** |

Here, F denotes forecasted value and T denotes the trend value or adjusting factor.

While following this method, we will fix the alpha value as 0.55 for both the stocks and take different beta values to see the difference between MAPEs.

**Exponential Smoothing Forecasting for Apple:**

A close-up of a number

Description automatically generated

A close-up of a number

Description automatically generated

A screenshot of a computer

Description automatically generated

A close-up of a box

Description automatically generated

**Observations:**

It's clear that using adjusted exponential smoothing leads to smaller MAPE values compared to exponential smoothing, making it a better model. The MAPE value decreases as the beta value increases up to 0.45, but it starts increasing again at 0.85. This suggests that beta values between 0.45 and 0.85 may provide more accurate results.

**Adjusted Exponential Smoothing Forecasting for Honeywell:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A close-up of a number

Description automatically generated

A screenshot of a computer

Description automatically generated

**Observations:**

It can be observed that as we increase the beta value for Honeywell, the MAPE decreases, with the lowest MAPE achieved at a beta value of 0.85. This suggests that for the most precise modified exponential smoothing prediction for Honeywell, a higher trend parameter value is needed.

Overall, we can say that the accuracy of the forecast for Apple and Honeywell using modified exponential smoothing depends significantly on choosing the right beta value or trend parameter. For Apple, a moderate beta value (0.45) gives the most accurate forecast, while for Honeywell, a higher beta value (0.85) works best. These findings suggest that the ideal beta value varies based on each company's traits.

We will now be doing long term forecasting using R.

**Part 2: Time Series using R:**

In this part, we will be building Auto regression models to predict the values of Apple and Honeywell stocks for next 8 months. We will first be using AR(1) model. The AR(1) model, also known as Auto Regressive model of order 1, predicts the next value in a time series based on the most recent observation and a coefficient. In the ARIMA(1,0,0) notation, it means we're fitting an AR(1) model without difference. This model is useful for forecasting when the value of a variable depends mainly on its previous value and can be a good starting point for time series analysis. In AR(1) model we use stationary data which does not show any increasing or decreasing trend over time. Let’s build AR(1) model for both apple and Honeywell:

**AR(1) model for Apple:**

In this part, we will be forecasting the apple stock prices from the given 5 years of data using AR(1) model in R. For this, we will define the ticker symbols for the equities of Apple (AAPL) and Honeywell (HON). Then, we will use the getSymbols function from the quantmod package to download historical stock price data for the previous five years.

A graph with numbers and lines

Description automatically generated

**Observations:** We can see that the AR(1) model predicts values that are nearly stationary for the next 8 months. This might be because the time series data isn't stationary as the AR(1) model expects. In the plot, the blue line shows the actual forecasted values, and the grey area represents the 95% confidence interval for those values. We can say that this is not a good enough model to do long-term forecasting of stock prices.

**AR(1) model for Honeywell prices:**

A similar method as done above will be done for Honeywell stocks.

A graph with numbers and lines

Description automatically generated

**Observations:**

We can see that for Honeywell too, the forecasted values remain fairly steady. The grey area represents the 95% confidence interval, while the blue line shows the forecasted values precisely. The wide prediction interval indicates uncertainty in predicting future values.

We noticed that the AR(1) model didn't work well for the stock prices of both companies. One possible reason could be stationarity. So, we'll now conduct the Augmented Dickey-Fuller (ADF) test to check if the data is stationary.

Stationarity refers to a property of time series data where statistical properties such as mean, variance, and autocovariance remain constant over time. In simpler terms, it means that the data's behavior doesn't change regardless of when it's observed.

**Checking for Stationarity for Apple:**

We will check for stationarity of apple stock prices using ADF test where the hypothesis will be:

**Null Hypothesis:** The stock prices are not stationary over time.

**Alternate Hypothesis:** The stock prices are stationary over time.

A white background with black text

Description automatically generated

**Observations:** It can be observed that the test result shows a Dickey-Fuller statistic of -2.2114 and a p-value of 0.4888. With a p-value higher than the usual significance level of 0.05, we fail to reject the null hypothesis, suggesting that the data is likely not stationary.

**Checking for Stationarity for Honeywell:**

We will follow similar steps to run the ADF test on Honeywell stocks prices.

A black text on a white background

Description automatically generated

**Observations:** For the Honeywell stock prices, the Augmented Dickey-Fuller Test yielded a Dickey-Fuller statistic of -2.4227 and a p-value of 0.3994. With a p-value greater than the typical significance level of 0.05, we fail to reject the null hypothesis. This suggests that the data is likely not stationary.

We observed that both the stocks are likely not stationary. Therefore, we using AR(1) model for forecasting is not a good option. We can use Auto.Arima() function in R to forecast the values.

ARIMA is a forecasting model in R used for time series data, particularly for non-stationary data where statistical properties like mean and variance change over time. The auto.arima function in R automatically identifies the best ARIMA model, considering factors like trend, seasonality, and noise, making it convenient for time series forecasting tasks.

**Auto.Arima Model for Apple:**

A computer screen shot of a number

Description automatically generated

A graph with numbers and lines

Description automatically generated

**Observation:** The auto.arima model for Apple stock prices suggests an ARIMA(0,1,0) model with a drift term. The drift coefficient is estimated to be 0.1009 with a standard error of 0.0686. The model's log likelihood is -2898, and the information criteria (AIC, AICc, BIC) are 5800.01, 5800.02, and 5810.28 respectively. It can be seen that ARIMA(0,1,0) model I giving better forecasts which are not simply straight lines. This could be because ARIMA is effective for non-stationary data as it applies differencing techniques to transform the data into a stationary form. By removing trends and seasonality, it helps in capturing the underlying patterns in the data, giving better forecasts.

**Auto.Arima Model for Honeywell:**

A black text on a white background

Description automatically generated

A graph with numbers and lines

Description automatically generated

**Observations:** The ARIMA model for Honeywell stock prices suggests a simple ARIMA(0,1,0) model without any additional terms. The model's performance metrics, such as log likelihood, AIC, AICc, and BIC, indicate that this basic ARIMA model may not fully capture the underlying patterns in the data, potentially limiting its forecasting accuracy. However, it is not a simple straight line as obtained in AR(1) model so, it is overall a better choice than the first AR(1) model.

**Auto.Arima for Dry Wine Data:**

We will now be forecasting long term values for Dry-Wine data using Auto-Arima model in R. We will first import the data and transform it to a time series object and then we will use a similar method as above to fit ARIMA model.

A screenshot of a computer

Description automatically generated

A graph of a wine chart

Description automatically generated with medium confidence

**Observations:** It can be seen that the resulting ARIMA(1,1,0)(0,1,1)[12] model suggests one autoregressive term, one seasonal moving average term, and a seasonal differencing term. The model's coefficients indicate negative relationships, with ar1 and sma1 values of -0.4652 and -0.6416, respectively. The ARIMA model selected indicates a negative relationship between the current observation and its lagged value, as well as a negative impact of the previous seasonal error on the current observation. The model's performance metrics, including log likelihood and information criteria, suggest a reasonable fit to the data. The forecasted dry wine prices, as depicted in the plot, show expected fluctuations over the forecast horizon. Therefore, we can say that The chosen ARIMA model captures both the autoregressive and seasonal dynamics present in the dry wine data, allowing for more accurate forecasts.

**Conclusion:**

In this project, we explored the analysis and forecasting of historical stock prices for Honeywell and Apple over a 252-day period. The project was divided into two parts: short-term forecasting using Excel and long-term forecasting using R.

In the Excel portion, we visually inspected the time series data for trends, seasonality, and irregular patterns. We applied exponential smoothing techniques and adjusted exponential smoothing to forecast stock prices, evaluating accuracy using metrics like MAD and MAPE

In the R portion, we utilized AR(1) models and the auto.arima function to fit time series models to the data. Despite initial attempts with AR(1) models, we found that the data was likely not stationary, leading us to explore more advanced forecasting methods like ARIMA. The auto.arima function in R helped identify the best ARIMA model for forecasting stock prices.

Observations from the analysis revealed that adjusted exponential smoothing generally provided more accurate forecasts compared to simple exponential smoothing. For both Apple and Honeywell, the choice of trend parameter (beta) significantly impacted forecast accuracy.

Additionally, ARIMA models proved effective in capturing the underlying patterns in the data, providing better long-term forecasts compared to AR(1) models. The ARIMA models accounted for non-stationarity by applying differencing techniques, resulting in improved forecast accuracy.

In conclusion, the accuracy of forecasting stock prices depends on various factors such as data stationarity, choice of forecasting method, and parameter selection. By leveraging advanced forecasting techniques like ARIMA and adjusting exponential smoothing, we can enhance the accuracy of financial time series forecasts, enabling better decision-making in enterprise analytics.

**References:**

Hayes, A. (2024, February 23). Autoregressive Integrated Moving Average (ARIMA) Prediction Model. Investopedia. <https://www.investopedia.com/terms/a/arima.asp>

<https://aws.amazon.com/what-is/autoregressive-models/>

<https://online.stat.psu.edu/stat510/lesson/1/1.2>