Module 3 Project — Forecasting a Time Series

Sourabh D. Khot (ID 002754952)

College of Professional Studies, Northeastern University

ALY 6050: Introduction to Enterprise Analytics CRN 81180

Professor Azadeh Mobasher

Jun 14, 2022

Table of Contents

Introduction

As a portfolio manager in a capital market firm, I handle investments in Apple and Honeywell stocks. For my clients, I have to forecast the future values of these stock prices based on data available for the past year. I have clients of both types – day trading, who are only interested in the next day's closing price for short-term investing, and investors, who invest for the long term. I have to advise them on the future values of Apple and Honeywell stocks and suggest an optimum portfolio from my point of view.

I will build and compare various forecasting methods for accuracy. For short-term forecasting, I will use exponential smoothing to capture seasonality and adjusted exponential smoothing to capture the trend. Weighted moving average until a base point, followed by a linear trend, will be used for long-term forecasting. Finally, I will fit a simple linear regression model to check the linear trend. The final section will suggest a portfolio share based on my reasoning and intuition.

Analysis

Historical stock prices for Apple Inc (AAPL) and Honeywell International Inc (HON) are available for one year (252 market days) from 8-Nov-19 to 6-Nov-20, and short-term forecasting is to be performed to predict prices for 9-Nov-22 (Mon) and long-term forecasting for 9-Nov-22 to 13-Nov-22.

For convenience, the date format used in this project (both report and workbook) is d-MMM-yy (month in abbreviated letters), instead of the original format of dd-MM-yy (month in numericals). Also, the stock prices are available for the market days (Mon-Fri). As the market is closed on weekends, it is assumed that only weekday factors affect the stock prices, and the timeline of consecutive weekdays is considered continuous and linear.

Part 1: Short-term Forecasting

For short-term predictions, we will use exponential smoothing and adjusted exponential smoothing, which can predict only the next immediate unknown value. First, I will analyze the time series visually through line plots.

(i)   Line plots to detect components

Simple line plots of Apple and Honeywell stock prices of historical data are plotted in figures 1 and 2, respectively. This will help us visually analyze the trend, seasonality, and irregularity in the time series.

Figure 1. AAPL Stock Price in $ Figure 2. HON Stock Price in $
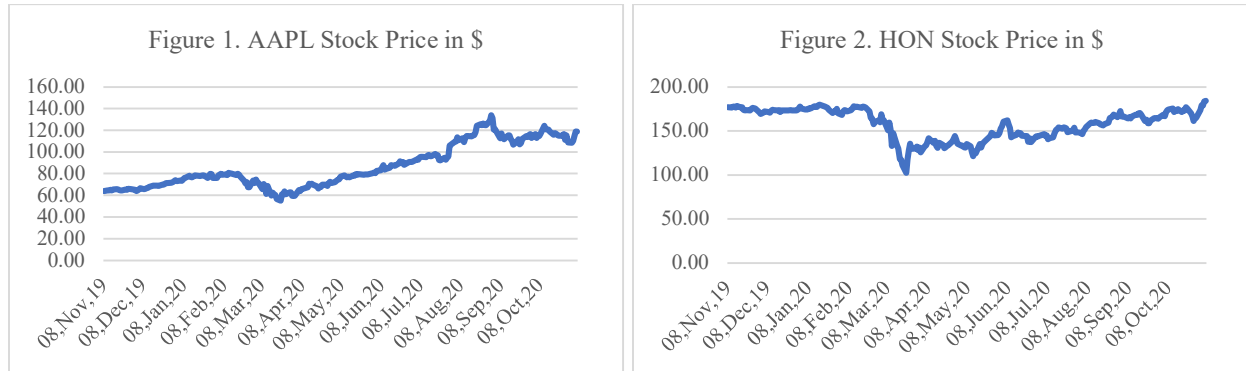
The Apple stock price has had a good trend over most of the year. The trend is bullish (increasing) from Nov-19 to Feb-20, then even more bullish from Apr-20 to Sep-20. Some bearish irregularities during Mar-20 and Oct-20 are random and do not seem to follow any pattern. From the available data of one year, no noticeable seasonal patterns are repeating over fixed intervals of time.

Honeywell stock price seems to follow stationarity from Nov-19 to Feb-20 with little to no dependence on time, post which there is an irregularity with the price unreasonably plummeting in Mar-20. From Apr-20 towards the end (about 70% of the series), there is a steadily increasing trend. There are irrational ups and downs but too random to be called seasonal/cyclical.

(ii)   Exponential Smoothing

I have performed exponential smoothing forecasting to forecast the price for Apple and Honeywell stock for the next immediate period 253 corresponding to 9-Nov-20. I have used various values of the smoothing parameter, α, and also computed the accuracy metric of Mean Absolute Percentage Deviation (MAPD). These results are summarized in Tables 3 and 4.

AAPL / $	$\alpha = 0.15$	$\alpha = 0.35$	$\alpha = 0.55$	$\alpha = 0.75$
MAPD	3.92%	2.49%	2.09%	1.97%
Forecast for period 253 (9-Nov-20)	114.88	115.97	117.50	118.39

Table 3. Exponential Smoothing for Apple Inc Stock Price

For the apple stock, an α of 0.75 has resulted in the lowest MAPD or highest accuracy. Exponential smoothing captures seasonality well, and in this case, higher α is more accurate because the data has no seasonality, i.e., the upcoming value is more dependent on the last actual and less on the last forecast. Accordingly, we will consider the forecast of $118.39 (at $\alpha = 0.75$) as the most accurate forecast of AAPL price for period 253.

HON / $	$\alpha = 0.15$	$\alpha = 0.35$	$\alpha = 0.55$	$\alpha = 0.75$
MAPD	2.86%	2.20%	1.91%	1.78%
Forecast for period 253 (9-Nov-20)	175.33	179.97	182.58	183.72

Table 4. Exponential Smoothing for Honeywell Inc Stock Price

For Honeywell, $\alpha = 0.75$ has the lowest MAPD indicating the comparatively most accurate model. This is because this stock also has no seasonality, with the current value more dependent on the last actual than the last prediction. For period 253, the most accurate forecast of HON price is $183.72 (at $\alpha = 0.75$).

(iii)   Adjusted Exponential Smoothing

To consider the trend parameter, we will further perform adjusted exponential smoothing. The smoothing parameter is fixed at $\alpha = 0.55$, and various trend parameters β are evaluated using the accuracy metric of Mean Absolute Percentage Error (MAPE). The results are summarized in Table 5 and Table 6.

AAPL, $\alpha = 0.55$	$\beta = 0.15$	$\beta = 0.25$	$\beta = 0.45$	$\beta = 0.85$
MAPE	1.97%	1.95%	1.94%	1.98%
Forecast for period 253 (9-Nov-20)	118.04	118.54	119.24	119.22

Table 5. Adjusted Exponential Smoothing for Apple Inc Stock Price

For the apple stock, a value of $\beta = 0.45$ has resulted in the lowest error because, as we checked in Figure 1, the time series has a good trend such that about half of the trend is dependent on the latest incremental forecast and another half on the previous trend. The value of $119.24 (at $\beta = 0.45$) can be considered the most accurate forecast of AAPL for period 253.

HON, $\alpha = 0.55$	$\beta = 0.15$	$\beta = 0.25$	$\beta = 0.45$	$\beta = 0.85$
MAPE	1.98%	1.94%	1.89%	1.85%
Forecast for period 253 (9-Nov-20)	184.07	184.72	185.19	184.83

Table 6. Adjusted Exponential Smoothing for Honeywell Inc Stock Price

$\beta = 0.85$ is the most accurate model for the Honeywell stock since there is a minor trend. The trend component is more dependent on the latest incremental forecast than the previous trend. As per adjusted Exponential Smoothing, the $184.83 ($\beta = 0.85$) is the most accurate forecast.

Part 2: Long-term Forecasting

(i)   Weighted Moving Averages combined with Linear Trend

I have forecasted values for the long term from periods 253 to 257. I will first use the 3-period weighted moving average method to forecast values from periods 4 to 100 with weights of 0.5, 0.3, and 0.2 for most to least recent periods, respectively. Subsequently, using period 101 as the base, I will use the linear trend method to forecast values from periods 101 to 257. The base is chosen as 101 (3-Apr-2020) since from the line graphs of the stocks in figures 1-2, we can see that both stocks follow a consistent bullish trend after crossing Apr-2020.

The forecasted value for the period 253-257 is also compared with actuals from Yahoo.com. The result of long-term forecasting for Apple stock is outlined in figure 7 and table 8, while that for Honeywell stock is given in figure 9 and table 10.


Figure 7. AAPL Long Term Forecast
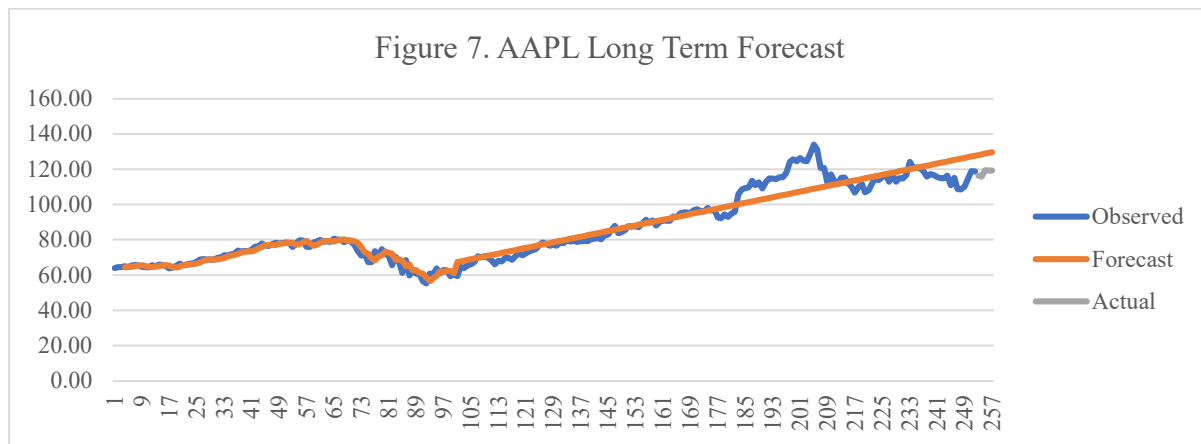
In figure 7, we can see that the overall prediction for Apple stock is close to the observed values. Until period 101, the moving average works well, and from 101 onwards, the linearly forecasted values for the long term follow the overall trend of the graph. There are concerning irregularities between the periods 185 and 209 and after the period 241.

AAPL / $				
Date	**Period**	**Observed**	**Forecast**	**Actual (Yahoo)**
5-Nov-20	251	118.82	127.22	Same as obs.
6-Nov-20	252	118.69	127.62	Same as obs.
9-Nov-20	253	-	128.02	116.32
10-Nov-20	254	-	128.42	115.97
11-Nov-20	255	-	128.81	119.49
12-Nov-20	256	-	129.21	119.21
13-Nov-20	257	-	129.61	119.26

Table 8. Long-term Forecasting for Apple Inc Stock Price

In table 8, comparing the forecasted values of periods 253-257 with the actual values from Yahoo.com, the absolute error is between 9 and 12, which is fine and a fair long-term prediction. Long-term investors may use the model, which may be improved for medium-term investors by taking the base period of linear trend prediction as 217 and reevaluating it.

Figure 9. Honeywell Long Term Forecast
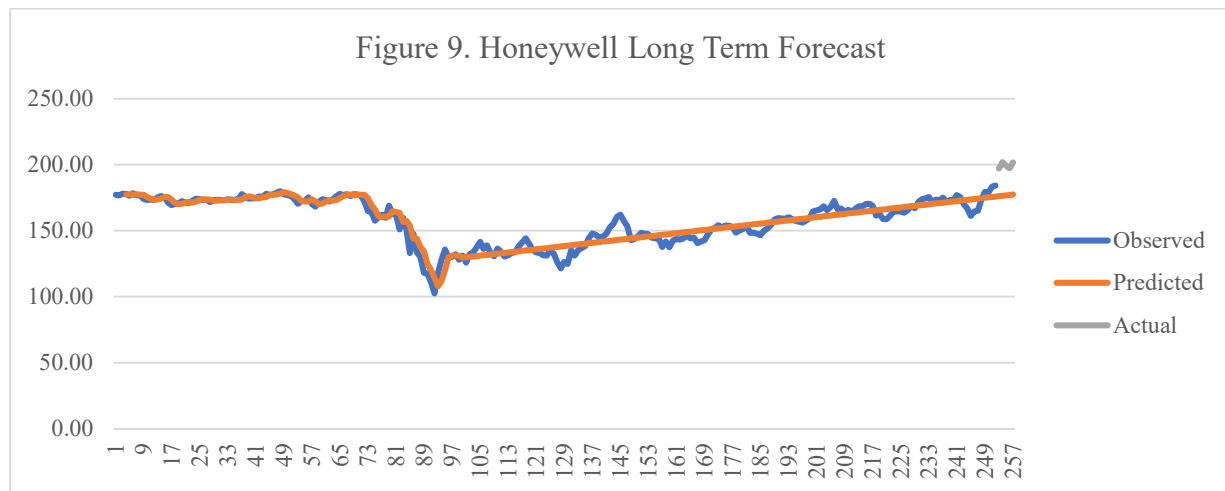
For Honeywell stock, the moving average till 101 period fits well with some lag, post which the linear trend somewhat fits the observed graph. There are noticeable irregularities, with the most prominent ones at 129 and 145.

HON / $ Date	Period	Observed	Forecast	Actual (Yahoo)
5-Nov-20	251	183.28	175.44	Same as obs.
6-Nov-20	252	184.27	175.75	Same as obs.
9-Nov-20	253	-	176.05	196.99
10-Nov-20	254	-	176.35	201.98
11-Nov-20	255	-	176.66	199.29
12-Nov-20	256	-	176.96	197.24
13-Nov-20	257	-	177.27	201.54

Table 10. Long-term Forecasting for Honeywell Inc Stock Price

In table 10, the actual values are given for the long-term forecast for the period 253-257. The absolute error is between 20 to 26, which is concerning. In fact, figure 9 shows that the highest error is during this long-term forecast. This error seems to be because of the irregularity component, which cannot be predicted, i.e. the stock price is inherently volatile. So this may be the best possible model and can be used only by the long-term investors.

(ii)   MAPE of long-term forecasting and comparison with short-term forecasting

I have computed MAPE of the previous model (part 2 (i)) where I have forecasted using the 3-period weighted average till 100 and then used linear trend with 101 as the base. Only periods 4-252 are used for calculating MAPE since before 4, there are no forecasts available, and after 252, there are no observed values. The actual Yahoo.com values from 253 are available, however, excluded for a fair comparison with the part 1 models, which considered values only till 252. For comparison, in the exponential smoothing of Part 1 (i), MAPD was computed. In contrast, the adjusted exponential smoothing forecast of Part 1 (ii) was built on this model with MAPE computed. Hence, we will compare the latest long-term model only with the adjusted exponential smoothing forecast with $\alpha = 0.55$ and a value of β that has the lowest MAPE. The results are compared in Table 11.

	Model from Part 2 (i)	MAPE	Model from Part 1 (ii)	MAPE
AAPL / $	Wt Mov Avg + linear trend from base of 101	3.76%	Adj Exp Smoothing with $\alpha = 0.55, \beta = 0.45$	1.94%
HON / $	Wt Mov Avg + linear trend from base of 101	2.69%	Adj Exp Smoothing with $\alpha = 0.55, \beta = 0.85$	1.85%

Table 11. MAPE of long-term forecasting & comparison with short-term

For the Apple stock, the part 2 model had a MAPE of 3.76%, which in absolute terms is not very large, but when compared to the adjusted exponential smoothing MAPE (1.96%), it is almost double. However, the latter forecasting method can only predict the next immediate value. Thus, the adjusted exponential smoothing will be preferred for short-term forecasting, while for long-term prediction, the part 2 method may be used and could be improved with a different base of 217 or so.

The Honeywell stock has part 2 MAPE of 2.69%, which is good although slightly higher than the adjusted exponential smoothing method of 1.85%. Part 1 model is preferred when short-term forecasting. However, the part 2 model may be used for long-term forecasting since it may be the best model as errors are due to irregularities that cannot be captured.

Part 3: Regression

(i)   Simple Regression & its Accuracy

Simple regression was performed on the Apple and Honeywell stock prices to predict the linear trend. The following regression equations were obtained (with period = 1 for 8-Nov-2019 and so on for every market day). The regression line is plotted in figures 12 and 13.

$$AAPL_{predicted} = 55.80 + 0.24 \times period, \quad R^2 = 0.7616$$

$$HON_{predicted} = 161.65 \pm 0.13 \times period, \quad R^2 = 0.0177$$

Figure 12. Regression on AAPL Stock Price

Figure 13. Regression on HON Stock Price
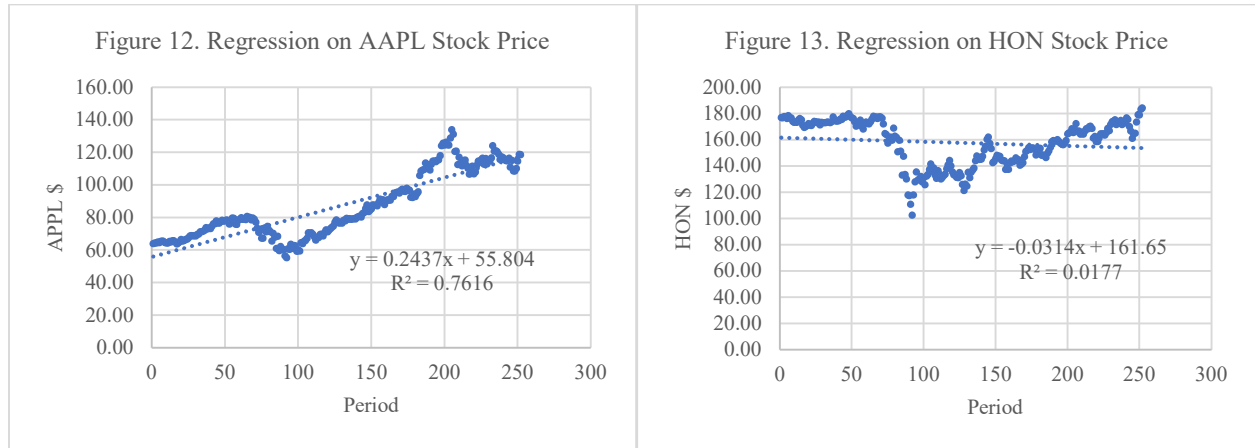
About 76% of the variation in Apple stock price is explained by time, which is essentially a trend, while 24% may be due to seasonal or irregularity factors. Similarly, only 1.7% of the variation in Honeywell's stock price is explained by time. So overall, there is a negligible trend for Honeywell, and almost all the variation is due to other factors (however, this is not true if we split the series before period 101 and after 101). MAPE for the regression model is computed and compared against previous models in Table 14.

	Model from Part 3 (i)	MAPE	Model from Part 2 (i)	MAPE	Model from Part 1 (ii)	MAPE
AAPL / $	Regression	10.56%	Wt Mov Avg + linear trend	3.76%	Adj Exp Smoothing with α = 0.55, β = 0.45	1.94%
HON / $	Regression	9.76%	Wt Mov Avg + linear trend	2.69%	Adj Exp Smoothing with α = 0.55, β = 0.85	1.85%

Table 14. MAPE of regression & comparison with previous parts

From table 14, we see that regression is the worst performing model in terms of MAPE in the case of both Apple and Honeywell stock prices. Generally, a stock price is volatile and has every component of trend, seasonality, and irregularity in the long term. A simple regression generalizes too much, assuming a linear price relationship with time, and hence may not work with stock prices. However, linear regression with more variables might perform better than simple linear regression with only time.

(ii)   Regression Residual Analysis

Independence of Residuals

Figure 15. AAPL: Residual vs. Period


Figure 16. HON: Residual vs. Period
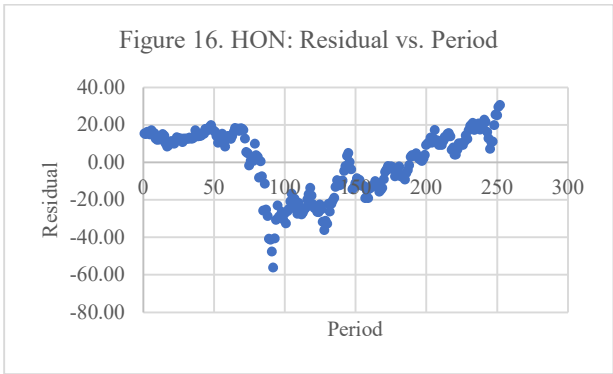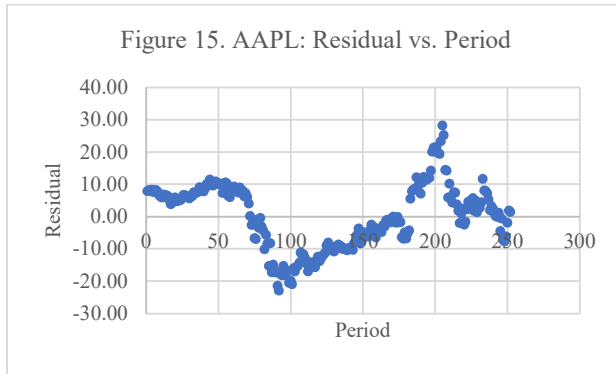
In the case of both Apple and Honeywell, the period vs. residual plots do not follow a single-piece pattern. However, they do not seem completely random also. There might be some dependence on time, and this assumption may void the validity of regression in both stocks.

Homoscedasticity of Residuals

Figure 17. AAPL: Residual vs. Predicted

Figure 18. HON: Residual vs. Predicted
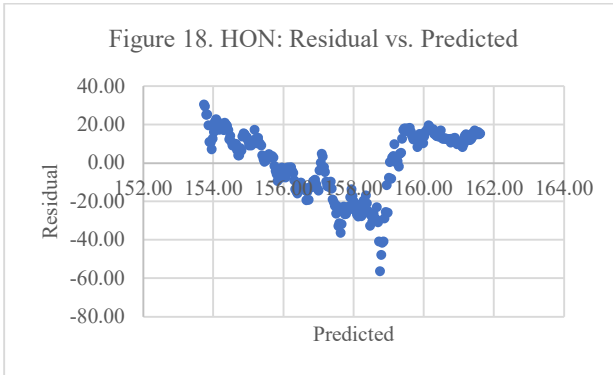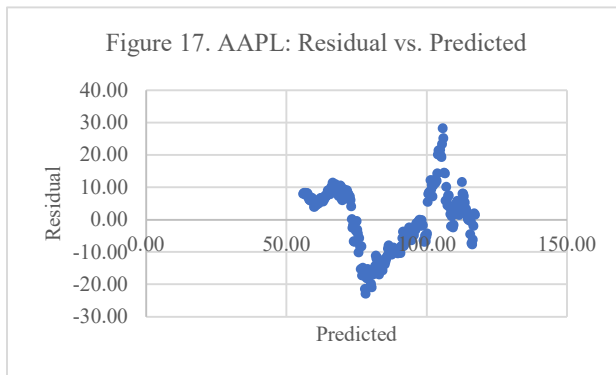
The predicted vs. residual plot has no direct pattern or an increase or decreasing trend, and the plots seem fairly random. Visually, we may say there is no significant unequal variance in the residuals over the predicted values, and this regression assumption of homoscedasticity may be considered valid.

Residuals normally distributed: Normal probability plot

Figure 19. AAPL: Normal Prob Plot of Residuals

Figure 20. HON: Normal Prob Plot of Residuals
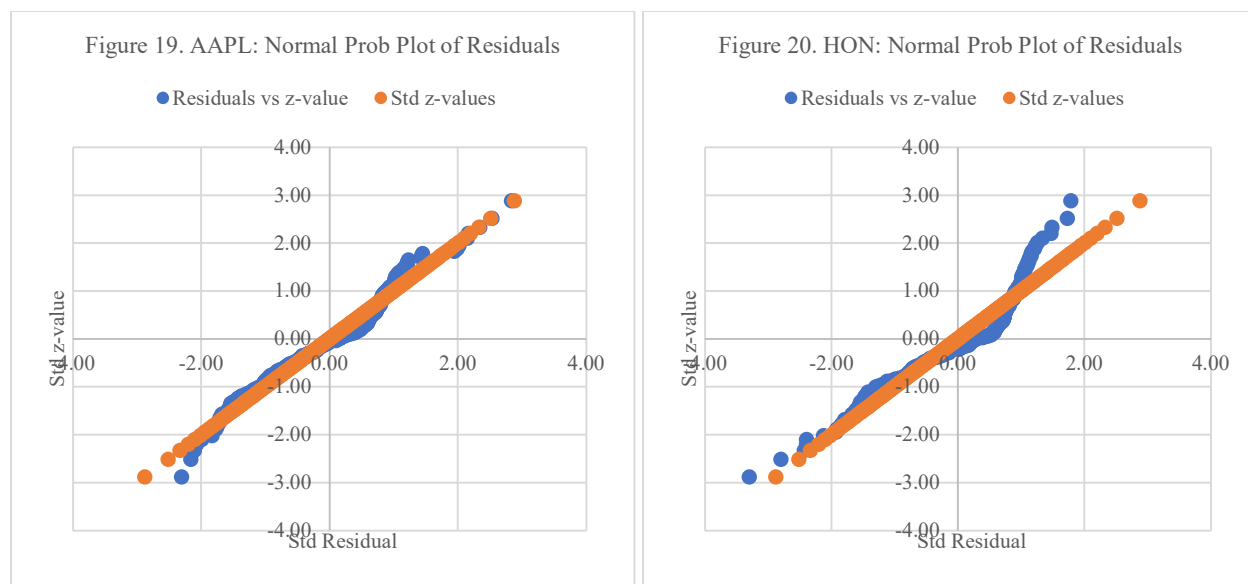
Analyzing the normal probability plot (or QQ plot) of residuals, we can say that both the stock residuals do not seem to follow a normal distribution. Especially the Honeywell stock has a significant deviation from the normal line at 45° towards almost the entire range. The Apple stock also has a lesser, but considerable deviation from the normal line. Hence, both the residuals do not seem to follow a normal distribution.

Residuals normally distributed: Chi-squared test

A chi-square goodness-of-fit test of residuals of both stocks was performed against normal distribution. The visual fit is shown in figures 21-22.

Figure 21. AAPL: Obs Residual vs Theo Normal

Figure 22. HON: Obs Residual vs Theo Normal
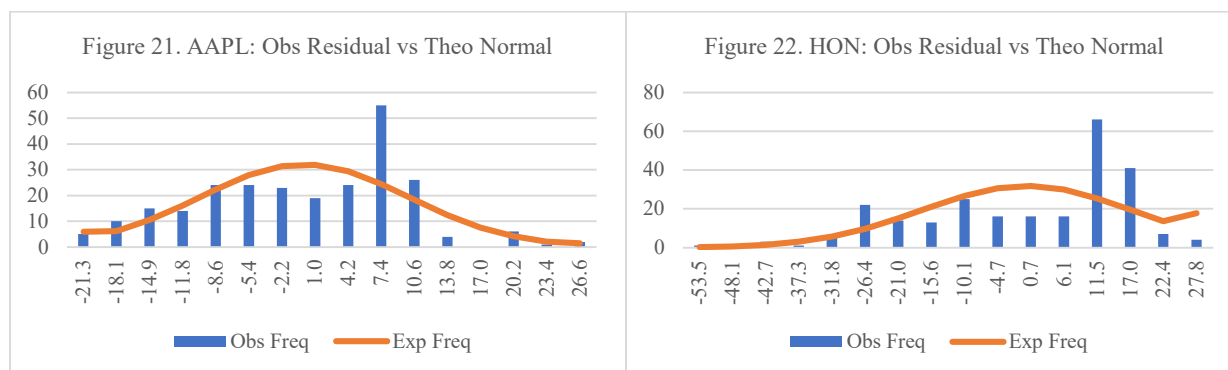
In figure 22 of Honeywell stock, the right end of the theoretical normal distribution line graph goes higher toward the last class. This is not an error but exists because the right tail of normal distribution extends beyond the range of the observed distribution, and the last bin has cumulatively included the longer truncated tail of the distribution, which is more than the second last bin.

For the hypothesis test, the theoretical normal distribution's two parameters (mean and variance) are estimated from the observed distribution. Hence, the degrees of freedom are 13. The hypothesis is defined below, and the test results are summarized in table 23.

H_0: A normal distribution is a good fit for the residuals

H_1: A normal distribution is not a good fit for the residuals

	AAPL (Apple Inc)	HON (Honeywell Inc)
Level of Significance	0.05	0.05
d.f. *(bins - para est - 1)*	13	13
Chi-Squared Test Statistic	250	250
Chi-Squared p-value	0.0000	0.0000
Interpretation	Sufficient evidence to reject H_0	Sufficient evidence to reject H_0

Table 23. Chi-Square Test for Normality of Residuals

In figures 21-22, the normal distribution does not seem to be a better fit, which is confirmed by the test as the p-value against both the distribution tends to zero. Hence, there is sufficient evidence to reject the claim that residuals do NOT follow a normal distribution.

The assumption of normality of residuals is clearly not met. Also, they do not seem independent. Hence linear regression is not valid for predicting the stock prices of Apple and Honeywell.

Question

(i)   Portfolio Allocation

Technically, to decide what percentage of my portfolio should be allocated between Apple and Honeywell, I need to know the horizon of investment (short or long term) and the investor's risk appetite. I also need to know the probability distribution of the two stock prices after Nov-20 onwards. Consecutively, I can perform a Monte Carlo simulation on various portfolio scenarios with different proportions of Apple and Honeywell shares and arrive at the optimum portfolio with best profits. However, the probability distribution of the stock prices cannot be known in advance. Another method could be to understand the trend vs. variance of a stock. A risk-taking person will allocate a higher proportion to stocks with a high trend and high irregularity. Similarly, a risk-averse person may allocate a higher proportion to low but consistent positive trend and low irregularity.

Considering my case, I am a slightly risk-averse person investing for the medium term (six months). Visually analyzing figures 1 and 2, replicated below for convenience, Honeywell was flat before May-20, while Apple gave more consistent returns except in Mar-20 and Apr-20. Also, post-Apr-2020, Apple has a slightly more positive trend, and lesser irregularity. Both the stocks show an upward trend, and post-Sep-20, Honeywell seems to outpace the growth of Apple. Hence I will allocate 70% to Apple and 30% to Honeywell.

Figure 1. AAPL Stock Price in $ Figure 2. HON Stock Price in $
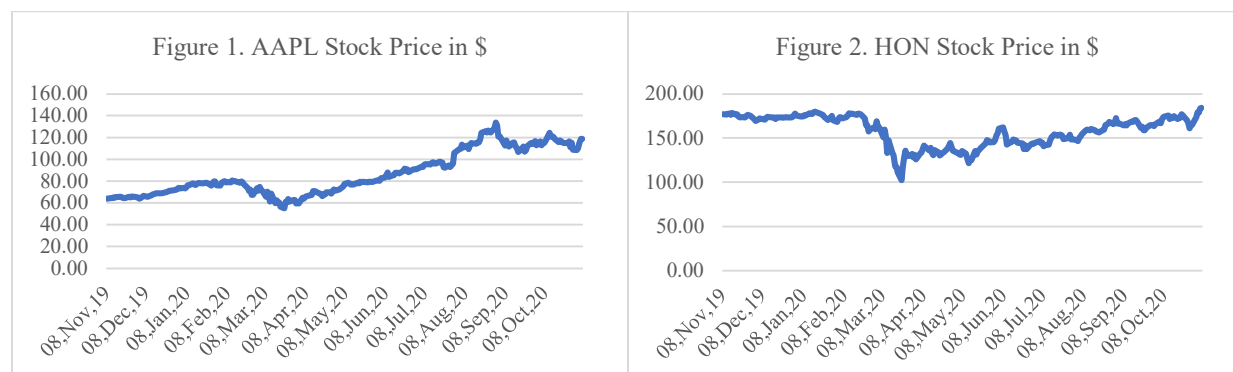
Conclusion

Visually, Apple stock prices have a trend and less irregularity while Honeywell has a trend after Apr-20 with slightly more irregularity. Both stocks do not have noticeable seasonality in the one year and in the long term, exhibit optimism.

For short-term forecasting, exponential smoothing is optimum (least errors) at a high α value of 0.75, indicating least seasonality for both the stock prices. For adjusted smoothing, with α fixed at 0.55, the best value of β is 0.45 for Apple and β for 0.85 for Honeywell stocks highlighting that Apple has a good trend, while Honeywell overall does not have a prominent trend as such.

Long-term forecasting using weighted average and linear trend from the base period of 101 was performed; however, the errors increased from the adjusted exponential smoothing, with errors almost doubling for Apple and increasing somewhat for Honeywell. For Apple, the long-term forecasting may improve with the base changed to 217 or more,

Simple regression was performed on both the time series; however, errors rocketed at least five times higher compared to the best of the short-term forecasting methods, with the Honeywell regression model not able to explain more than 1.7% of the variation in stock price. Instead, the regression method can be dismissed as the assumptions of residual normality and independence were not met.

For a slightly risk-averse person for the medium term (about six months), I would suggest allocating 70% of the portfolio to Apple and 30% to Honeywell since Apple has a consistent trend with less irregularity in the last eight months, and Honeywell also seems to pick up growth in the last two months.

References

(n.d.). Retrieved from Yahoo Finance: https://finance.yahoo.com/

APA Style Table: APA.org. (n.d.). Retrieved from https://apastyle.apa.org/style-grammar-guidelines/tables-figures/tables

Canvas Module 3 Project: Forecasting a Time Series. (n.d.). Retrieved from https://northeastern.instructure.com/courses/110059/assignments/1346054

Evans, J. R. (2013). *Statistics, data analysis, and decision modeling.* New York: Pearson.