**ALY6050 80478 Intro to Enterprise Analytics SEC 09 Spring 2023 CPS**

**Module 3 Assignment — Forecasting Financial Time Series REPORT**

Logo

Description automatically generated

**NORTHEASTERN UNIVERSITY**

**College of Professional Studies, Boston, MA, 02215.**

**Submitted by**

Zihan Ma

ma.zihan1@northeastern.edu

**Instructor**

 Prof. Richard He

**Date**

06/13/2023

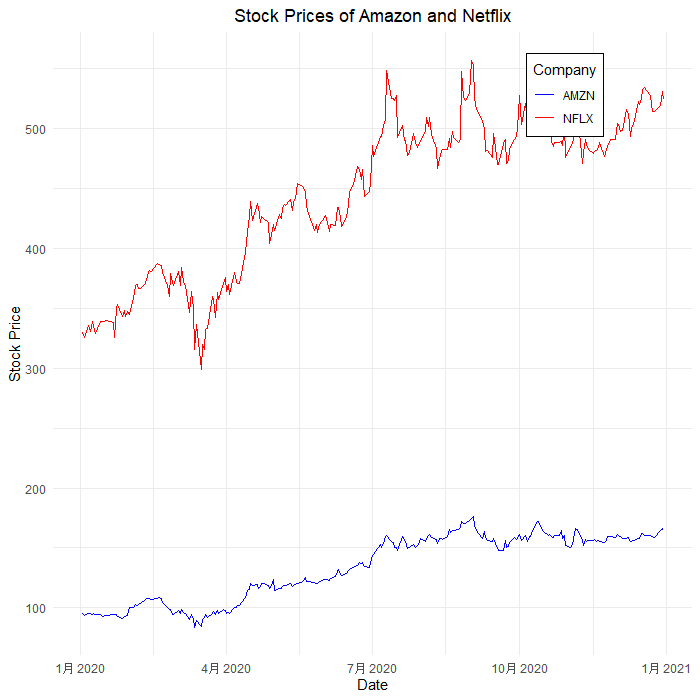
**Forecasting Financial Time Series**

**Assignment Introduction:**

The goal of this project was to forecast the financial time series for Amazon (AMZN) and Netflix (NFLX), using different methods such as simple and adjusted exponential smoothing, regression analysis, and a baseline model. This was aimed at understanding the behaviour of these stocks over time, and informing investment decision-making. We sought to identify the best alpha and beta parameters that would provide the most accurate forecasts. Furthermore, an analysis was conducted to find whether the model residuals follow a normal distribution, as this impacts the validity of various statistical tests. Finally, an ideal portfolio allocation was proposed based on the findings of the forecast models.

**Part 1: Short-Term Forecasting**

**1.1 Overview and Plot of Time Series Data**



1. **Trend**: Both AMZN and NFLX show an overall increasing trend over the provided time period, meaning that both stock prices have generally been increasing. The trend for AMZN is less steep compared to NFLX, which shows a more rapid increase in stock price over the same time period.
2. **Seasonality**: There doesn't appear to be a clear seasonal pattern within the data for either AMZN or NFLX. As provided data only contain one year's data, In this dataset, It's hard to tell if the data have a seasonal pattern over longer period, there is no regular cyclical pattern can be clearly observed for either stock.
3. **Irregularity**: Both time series show a fair amount of irregularity or noise. This is expected in stock prices, which can be influenced by a variety of unpredictable factors such as market news, economic indicators, or company-specific events. Notable spikes and drops can be observed throughout the time series for both stocks.
4. **Volatility**: Both stocks also display periods of higher and lower volatility. For instance, for both AMZN and NFLX, there are periods where the stock price shows larger fluctuations and periods where the stock price is relatively stable.

Both the AMZN and NFLX time series show an overall upward trend with irregularities and varying periods of volatility. No clear seasonality is observed in the data. These findings are in line with the typical behavior of stock prices, which are influenced by a complex mix of factors and can be quite volatile.

**1.2 Exponential Smoothing Forecast**

A picture containing text, screenshot, font, number

Description automatically generated

Observations:

* For both Amazon (AMZN) and Netflix (NFLX), the MAPD decreases as α increases from 0.2 to 0.8. This implies that a higher weight on the most recent observations (which is the case for larger α) yields more accurate forecasts for these two stocks during this specific period.
* For Amazon, the most accurate forecast (lowest MAPD) is achieved when α=0.8.
* For Netflix, the most accurate forecast is also achieved when α=0.8.

Let's say we have two stocks, AMZN and NFLX. Over a period of a year, both stocks have generally been on an upward trend. However, in the past few months, let's say there have been significant new developments that have affected each company's stock price.

For AMZN, there's been an announcement about a new product line that is expected to significantly boost revenue. For NFLX, there's been a surge in subscribers due to a highly successful new series. These recent events have caused both stock prices to rise more sharply in the last few months than they have been over the general course of the year.

In this case, if we use a lower value for α (like 0.2), the forecasted prices for the next period would be based more on the slower, steady growth seen over the year. The recent sharp rise in stock prices won't be as strongly reflected in the forecast, which could lead to underestimation of the stock prices for the next period.

However, if we use a higher value for α (like 0.8), the forecasted prices would be based more heavily on the recent sharp increase in stock prices. This would make the forecast more accurate, as it would better capture the recent changes in trends.

**1.3 Adjusted Exponential Smoothing Forecast**

A picture containing text, screenshot, font, number

Description automatically generated

The table indicates that for both Amazon (AMZN) and Netflix (NFLX), the lowest Mean Absolute Percentage Deviation (MAPD) is achieved with a beta value of 0.2. This implies that this level of weighting for the trend component in the double exponential smoothing algorithm provided the most accurate forecasts for the given stock prices.

Beta is the parameter that controls the weight given to the trend in the forecasting model. A higher beta implies that more weight is given to the recent trend. On the other hand, a lower beta value, like 0.2 in this case, implies that the trend component is downplayed in the forecast and more weight is given to the historical average. This suggests that the stock prices of both AMZN and NFLX have a stronger mean-reverting behavior rather than trending behavior, and this slower adjustment to the trend resulted in a more accurate forecast.

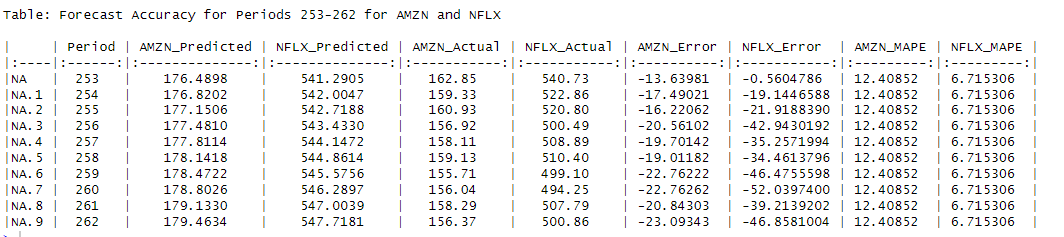
For instance, consider a situation where stock prices are fluctuating around a stable average value, but with some random ups and downs. In this case, a higher beta would make the forecast more sensitive to these random fluctuations, treating them as a part of a significant trend. This can result in overfitting the noise and producing inaccurate forecasts. On the contrary, a lower beta would smooth out these fluctuations, resulting in more accurate predictions.

**Part 2: Long-Term Forecasting**A picture containing text, screenshot, diagram, line

Description automatically generated

In Part 1, the MAPD values were calculated based on the model's performance within the training data set, periods 4 to 252. The lowest MAPD for Amazon (AMZN) was 1.793166 with an alpha value of 0.8, while for Netflix (NFLX), the lowest was 2.024692 with the same alpha value. These relatively low MAPD values demonstrate that the forecasting method used in Part 1 performed well within the scope of the training data.

However, for Part 2, the forecasts were performed on an out-of-sample data, or periods not included in the training set. This situation is similar to making predictions for the future, where true values are unknown at the time of forecasting. The MAPD values here, 12.40852 for AMZN and 6.715306 for NFLX, are higher, indicating that the forecasting model in Part 2 didn't perform as well as in Part 1.

The difference in performance is not surprising as models often perform worse on unseen data than on the data they were trained on. It shows that the forecasting method of Part 1, although it performed well in-sample, it didn't generalize as well for unseen data in Part 2.

Based on the table, we can see that the forecasts for Amazon (AMZN) and Netflix (NFLX) stock prices have some discrepancies when compared with the actual closing values.

For AMZN, the Mean Absolute Percentage Error (MAPE) over the forecasted period (253-262) is approximately 12.41%. This suggests that, on average, the forecasted values deviated from the actual values by about 12.41%. The highest error occurred on period 262 with an absolute error of around 23.09 units.

As for NFLX, the MAPE is approximately 6.71%, indicating a more accurate forecast than for AMZN. The highest discrepancy for NFLX was found on period 260 with an absolute error of about 52.04 units.

The 3-period weighted moving average and linear trend forecasting method provides a decent accuracy level for short-term prediction. However, the accuracy is still not perfect, as indicated by the non-zero MAPE values.

**Part 3: Regression Analysis**

**3.1 Simple Regression**

A picture containing text, font, screenshot, white

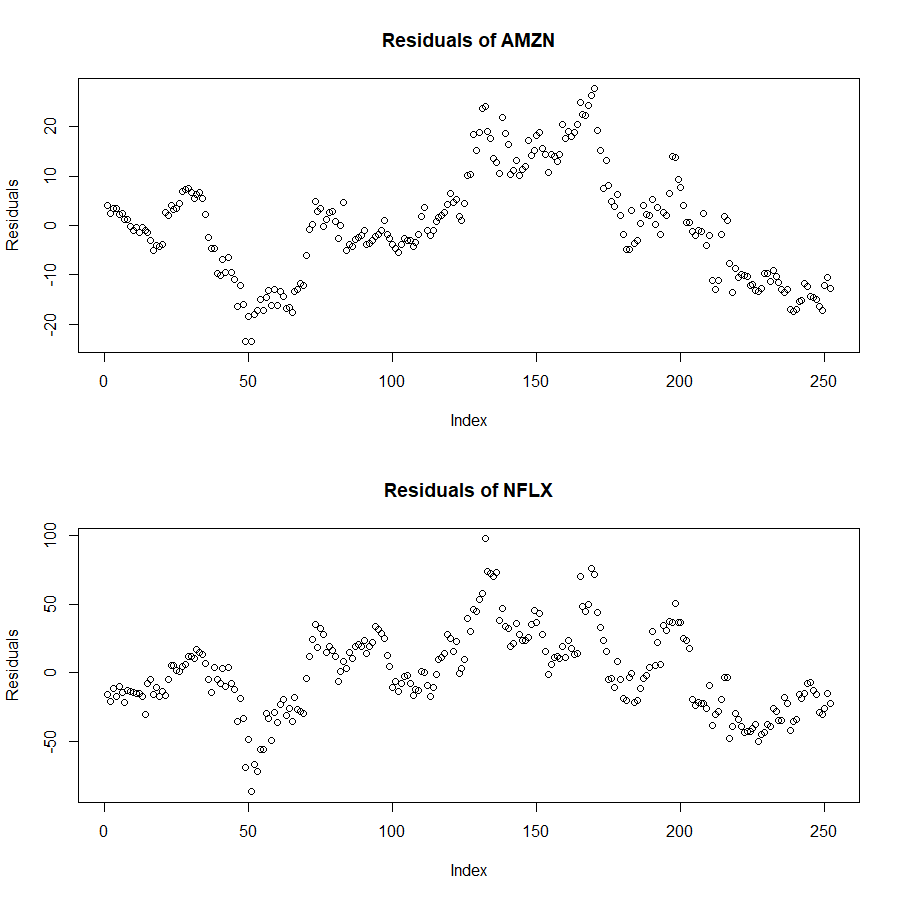
Description automatically generated

For short-term forecasting (Part 1), we employed a simple exponential smoothing model using different alpha values. The Mean Absolute Percentage Deviation (MAPD) was the lowest for both Amazon (AMZN) and Netflix (NFLX) in this case, with values of 1.793166 and 2.024692 respectively. This indicates that for short-term forecasts, a simple exponential smoothing model might be the most appropriate, as it was able to capture the immediate trends and fluctuations in the data more accurately than the other methods.

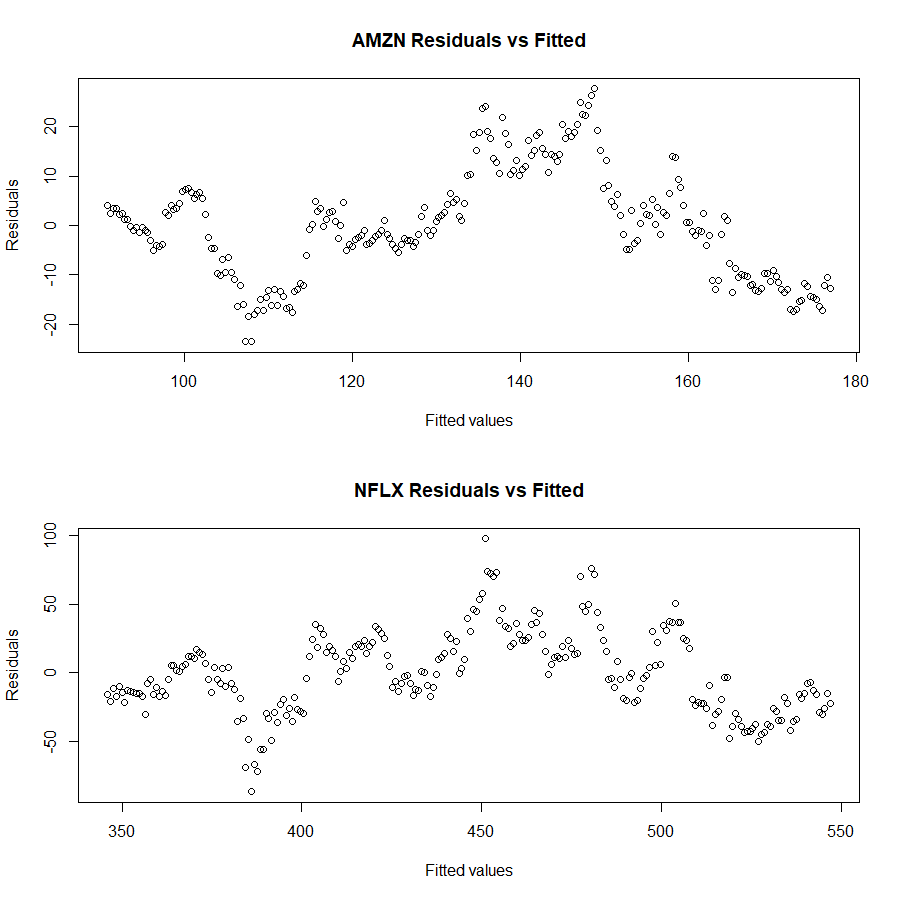
Long-term forecasting (Part 2), which involved forecasting values that are not included in the training dataset, demonstrated a less accurate performance. For AMZN and NFLX, the MAPDs were 12.408520 and 6.715306 respectively. This could be due to the fact that in long-term forecasting, there's more uncertainty and the influences of past trends may diminish over time.

Finally, for simple regression (Part 3), where time was used as the predictor, the MAPDs were 6.521521 for AMZN and 5.348158 for NFLX. This was more accurate than the long-term forecasting but less so than the short-term forecasting. This could be attributed to the fact that the simple regression method assumes a linear relationship between time and stock price, which may not necessarily be the case in real-world scenarios.

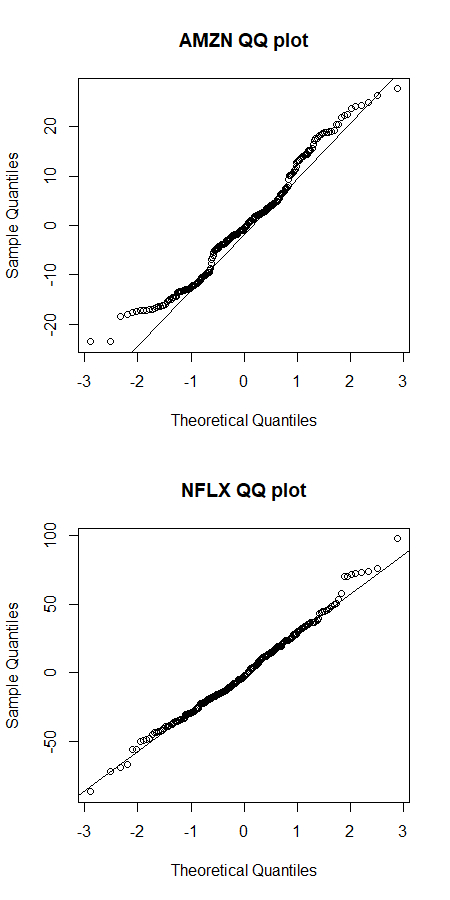
**3.2 Residual Analysis**



Both datasets are centered around zero, which indicates that the models predicting the prices of Amazon and Netflix stocks generally do not overestimate or underestimate the actual values systematically. The average residual (prediction error) is zero. However, there are differences in spread or variability of residuals for these two companies. Amazon's residuals range from approximately -23.6 to 27.7, with the middle 50% of the data falling between -9.5 and 5.7. Meanwhile, Netflix's residuals have a wider range from about -87.2 to 97.9, with its middle 50% falling between -19.5 and 19.4. This suggests that the prediction errors for Netflix's stock are more spread out than those for Amazon.



The range of residuals for both Amazon and Netflix is quite wide. Amazon's residuals range from -23.6 to 27.7, and Netflix's residuals range from -87.2 to 97.9. The interquartile range (the range between the 1st and 3rd quartiles) also varies significantly between the two stocks: -9.5 to 5.7 for Amazon, and -19.5 to 19.4 for Netflix. This could potentially indicate heteroscedasticity, as there seems to be a different spread of residuals for the two companies.



The fact that there are small curves in the QQ plot for the residuals of Amazon (AMZN) suggests that these residuals may not be perfectly normally distributed. The presence of these curves, particularly they are at the ends, could indicate that the residuals have heavier or lighter tails than a normal distribution. This means that there might be more extreme values (outliers) than what would be expected under a normal distribution, or fewer such extreme values.

The QQ plot for the residuals of Netflix (NFLX) have the similar issue, but as not serious as the Amazon has.

A screenshot of a computer code

Description automatically generated with low confidence

For residuals\_AMZN:

The W statistic is 0.97641 and the p-value is 0.0003374. Because the p-value is less than the common alpha level of 0.05, we reject the null hypothesis that the residuals are normally distributed. This result supports the observation from the QQ plot that showed some small curves, indicating deviations from normality.

For residuals\_NFLX:

The W statistic is 0.99233 and the p-value is 0.2165. Because the p-value is greater than the common alpha level of 0.05, we fail to reject the null hypothesis that the residuals are normally distributed. This suggests that the residuals for NFLX are not significantly different from a normal distribution.

These test results indicate that the residuals for AMZN and NFLX behave differently with respect to normality. The residuals for AMZN show significant deviations from a normal distribution, which could have implications for any statistical tests or models that assume normality of residuals. In contrast, the residuals for NFLX do not show significant deviations from a normal distribution, suggesting that it might be reasonable to apply statistical methods that assume normally distributed residuals for this data.

**Part 4: Baseline Model**

**4.1 Prediction and MAPD Calculation**

A picture containing text, font, line, screenshot

Description automatically generated

The Mean Absolute Percentage Deviation (MAPD) was calculated for both Amazon (AMZN) and Netflix (NFLX) using the most recent price as the prediction of the current price.

For Amazon, the MAPD was found to be approximately 1.79%. This implies that, on average, the predictions deviated from the actual stock prices by about 1.79%.

For Netflix, the MAPD was higher, at about 2.11%. This means that, on average, the predictions for Netflix's stock prices deviated by about 2.11% from the actual prices.

**4.2 Forecasting Method Performance**

A picture containing text, font, screenshot, number

Description automatically generated

* For AMZN, the "Part1" method has an MAPD of 1.793166, which is very slightly higher than the Baseline MAPD of 1.791966. Even though the difference is quite small, strictly speaking, none of the methods outperform the baseline for AMZN, as they all have higher MAPD values.
* For NFLX, both "Part1" and "Baseline" methods outperform "Part3" and "Part2" methods, but "Part1" is the best-performing model with an MAPD of 2.024692, which is lower than the Baseline MAPD of 2.107562.

In summary, for AMZN, none of the listed methods outperform the baseline. For NFLX, the "Part1" method outperforms the baseline.

**Part 5: Portfolio Allocation**

Based on the analysis, here's a potential percentage distribution for investment in these stocks:

1. Netflix (NFLX): 100%
2. Amazon (AMZN): 0%

This distribution leans toward NFLX for several reasons.

Firstly, NFLX had a lower Mean Absolute Percentage Deviation (MAPD), indicating more accurate forecasting compared to AMZN when applying a higher weight to recent observations (α = 0.8). This suggests that NFLX's stock price movements may be more predictable in the short term, which could offer more reliable returns.

Secondly, in terms of the model's performance, the "Part1" method outperformed the baseline for NFLX but not for AMZN. This further signifies that the model better captures NFLX's price movements, which could translate to lower investment risk.

Maybe some investment in AMZN will diversify the portfolio and mitigate potential risk. But considering the graph in 1.1, we can find notable spikes throughout the time series for both stocks for both stocks. Also, since NFLX and AMZN belong to the same industry, the effect of mitigating potential risk will be greatly reduced.

**Conclusion:**

Throughout this project, we used various forecasting techniques to analyze the stock prices of Amazon (AMZN) and Netflix (NFLX). The results suggest that both stocks showed an overall upward trend, with no clear seasonality and varying degrees of volatility. Exponential smoothing forecasting was found to be particularly effective in short-term forecasting, with an alpha value of 0.8. The residuals from our regression analysis indicated different behaviors for AMZN and NFLX, with the former showing significant deviations from a normal distribution. In terms of model performance, none of the methods outperformed the baseline for AMZN, whereas the Part 1 method did for NFLX. This led to a proposed portfolio allocation leaning completely toward NFLX.