Graphical user interface, text

Description automatically generated with medium confidence

Logo

Description automatically generated

**Abhilash Dikshit**

ALY6050: Historical stock prices for TESLA (TSLA) and Amazon (AMZN)

Instructor: Soheil Parsa

Week 3: Forecasting a Time Series

Date: 2023/03/16

**Problem Statement:**

*Use Yahoo Finance to get the historical stock prices for TESLA Inc (TSLA) and Amazon Inc (AMZN) for a total time period of one year, consisting of 252 market days.*

|  |  |
| --- | --- |
|  |  |

We are calculating the start and end dates based on the current date and getting the historical stock prices for TSLA and AMZN for the specified period. Later, we combined the two dataframe into a single dataframe with columns for each stock.

**Part 1. Short Time Forecasting**

**Part 1.a**

*Use a simple line plot of both time series to detect seasonal, irregular, or trend behaviors if any. Write a summary of your observations of both time series in your report.*

|  |  |
| --- | --- |
| In this code, we first calculate the start and end dates as before. We then download the historical stock prices for TSLA and AMZN as before. |  |

Chart, line chart

Description automatically generated

To plot the historical stock prices for TSLA and AMZN, we create a new figure using plt.subplots(), and then plot the adjusted close prices for each stock using ax.plot(). We set the title of the plot using ax.set\_title(), and label the x and y axes using ax.set\_xlabel() and ax.set\_ylabel(), respectively. Finally, we add a legend to the plot using ax.legend(), and display the plot using plt.show().

By examining the plot, we can identify any seasonal, irregular, or trend behaviors in the historical stock prices for TSLA and AMZN. For example, we may observe a seasonal pattern where stock prices tend to increase or decrease during certain times of the year, or a trend where stock prices are increasing or decreasing over time. We may also observe irregular behaviors such as sudden spikes or dips in stock prices.

**Part 1.b**

*Perform exponential smoothing to forecast both prices for period 253. Use successive values of 0.15, 0,45, 0.55, and 0.75 for the smoothing parameter α. Next, calculate the MAPD (Mean Absolute Percentage Deviation) of each forecast; and based on the MAPDs, determine the value of α that has yielded the most accurate forecast for each stock. In your report, describe your results; and explain why in your opinion such values of α have yielded the most accurate forecasts for the two stocks.*

**Text

Description automatically generated**

Output:

TSLA - Best alpha: 0.75, MAPD: 1.21%

AMZN - Best alpha: 0.75, MAPD: 0.73%

The output suggests that the best alpha value for both TSLA and AMZN is 0.75, and the corresponding minimum MAPD for TSLA and AMZN are 1.21%and 0.73%, respectively.

The alpha value determines the weight given to the most recent observation while smoothing the data. A higher alpha value gives more weight to the most recent observation and results in a smoother forecast, while a lower alpha value gives more weight to the past observations and results in a more responsive forecast.

The MAPD is the mean absolute percentage deviation between the forecasted values and the actual values. It measures the accuracy of the forecast and indicates the percentage deviation of the forecast from the actual value.

A lower MAPD indicates a more accurate forecast, and in this case, the MAPD values for both TSLA and AMZN are relatively low, indicating that the forecasts are accurate. The fact that the same alpha value of 0.75 yielded the best forecast for both stocks suggests that this value is likely to be a robust choice for future forecasts.

**Part 1.c**

*Use your exponential smoothing forecast of part (ii) with 𝜶=𝟎.55 and perform an adjusted exponential smoothing to forecast both prices for period 253. Use successive values of 0.15, 0.25, 0.45, and 0.85 for the trend parameters β for both stocks. Next, calculate the MAPEs (Mean Absolute Percentage Error) of your forecasts and determine the values of β that have provided the most accurate forecasts for both stocks. In your report, describe your results and explain why, in your opinion, such values of β have yielded the most accurate forecasts.*

Text

Description automatically generated

Output:

TSLA - Best beta: 0.85, MAPE: 0.40%

AMZN - Best beta: 0.85, MAPE: 0.29%

The output shows the results of an adjusted exponential smoothing forecast for TSLA and AMZN stocks. The forecast is performed using different values of the trend parameter beta (0.15, 0.25, 0.45, and 0.85) along with a fixed smoothing parameter alpha of 0.55, which is used in the exponential smoothing model.

For TSLA, the best beta value that provides the most accurate forecast is 0.85, and it yields an MAPE (Mean Absolute Percentage Error) of 0.40%. This means that the forecasted price is, on average, 0.40% away from the actual price.

For AMZN, the best beta value that provides the most accurate forecast is 0.85, and it yields an MAPE of 0.29%. This means that the forecasted price is, on average, only 0.15% away from the actual price. The smaller the MAPE, the more accurate the forecast.

In summary, the adjusted exponential smoothing forecast with a beta value of 0.45 provides the most accurate forecast for AMZN, while a beta value of 0.29% provides the most accurate forecast for TSLA. It's important to note that the choice of beta values depends on the data and may differ for other stocks or time periods.

**Part 2: Long-term Forecasting**

**Part 2.a**

*For each stock, use a 3-period weighted moving averages to forecast its value during periods 1 through 100. Use the weights 0.5 (for the most recent period), 0.3 (for the period before the most recent), and 0.2 (for two periods ago). Next, use the observed value for period 101 as the base of a linear trend, and use that linear trend to forecast the values of both stocks for periods 101 through 257.Write a summary of your results in your report. Describe how accurate this method of forecasting has been by comparing the forecasted values for periods 253-257 with their actual “Close” values on those specific days (Hint: check the actual values on https://finance.yahoo.com).*

The output presents the results of a chi-squared goodness-of-fit test for two variables labeled as alpha1 and alpha2. The test is used to determine whether the observed frequencies of the variables follow an expected distribution or not. The null hypothesis for this test is that the observed frequencies follow the expected distribution, while the alternative hypothesis is that they do not.

To perform the 3-period weighted moving averages forecast for each stock, we can use the rolling function from pandas and multiply the weights accordingly. We can then use the iloc function to select the last value in each rolling window, which represents our forecasted value for the next period.

For example, for TSLA and AMZN, we can perform the forecast as follows:

Graphical user interface, text, application

Description automatically generated

|  |  |
| --- | --- |
| The code is creating a moving average forecast for the stock prices of Tesla (TSLA) and Amazon (AMZN). It uses a weights array to calculate the moving average and iterates over 100 periods to forecast the values. The historical data for TSLA and AMZN is updated with the predicted values for each iteration. The output is a combined dataframe of the Close prices for both stocks with the head and tail of the dataframe printed to the console. The forecasted values are shown in the last 5 rows of the dataframe. |  |

Text

Description automatically generated

|  |  |
| --- | --- |
| The code provided performs time series forecasting for the TESLA and Amazon stock prices by calculating the slope and forecasting the values for the next 157 periods. The output shows the head and tail of the combined forecasted data for both stocks, where the predicted values for TESLA and Amazon are repeated for all periods. This implies that the forecasted values remain constant for all 157 periods. |  |

**Part 2.b**

*Calculate the MAPEs (Mean Absolute Percentage Error) of your forecasts in question (i) above and compare them with the values obtained for your forecasts in Part 1. For each stock, describe which method has yielded a most accurate forecast.*

*Graphical user interface, text

Description automatically generated*

Output:

Mean Absolute Percentage Error (MAPE) for TESLA Inc (TSLA): 26.19%

Mean Absolute Percentage Error (MAPE) for Amazon Inc (AMZN): 19.48%

The weighted moving averages and linear trend method produced a more accurate forecast for Amazon Inc (AMZN) with a MAPE of 19.48% compared to TESLA Inc (TSLA) with a MAPE of 26.19% based on the Mean Absolute Percentage Error (MAPE). This implies that the linear trend technique outperformed the 3-period weighted moving averages method in predicting the price of AMZN's stock, while neither method performed very well in predicting the price of TSLA's stock. It's crucial to remember that this assessment is based on a brief period of time, and that other elements such as market movements, news, and outside events can have had an impact on the forecasts' accuracy.

**Part 3.**

**Part 3.a**

*For each stock, use simple regression of stock values versus the time periods to predict its values for periods 1 through 257. In your report, describe how the accuracy of this prediction has been compared to the methods used in Parts 1 and 2 of this project.*

**Graphical user interface, text

Description automatically generated**

|  |  |
| --- | --- |
|  |  |

**Part 3.b**

*Perform a residual analysis of your simple regression to verify whether regression is appropriate to use for each of the given data. Determine:*

*● Whether the residuals are independent*

*● Whether the residuals are homoscedastic.*

*● Whether the residuals are normally distributed by plotting a normal probability plot of the residuals*

*● Whether the residuals are normally distributed by performing a Chi-squared test for Normality of the residuals.*

*Chart, bar chart, histogram

Description automatically generated*

**Graphical user interface, chart, line chart

Description automatically generated**

TSLA residuals are not normally distributed (p = 0.0000)

AMZN residuals are not normally distributed (p = 0.0009)

This code first uses the statsmodels library to fit the linear regression models to the TSLA and AMZN data, and then calculates the residuals for each model. We then plot the residuals against the fitted values to check for independence and homoscedasticity.

We also plot normal probability plots of the residuals to check for normality.

Finally, we perform a chi-squared test for normality of the residuals.

**Final Question**

*Suppose that you have decided to form a portfolio Π (Pi) consisting of the above two stock types (denote a share value of AMZN by X and that of TSLA by Y). You are however undecided as to what percentage of your investment should be allocated to the AMZN shares and what percentage should be allocated to TSLA shares. Let these percentages be denoted by P and Q respectively (Obviously, P + Q=100%). In your opinion, what are good values to select for P and Q?*

To determine the appropriate allocation between AMZN and TSLA shares for the portfolio Π, we can use modern portfolio theory. This theory suggests that the optimal allocation depends on the expected returns and risk (measured by the variance) of each asset, as well as their correlation.

To calculate the expected returns and risk, we can use historical data from Yahoo Finance. Let's assume that we have data for the past 3 years, and we want to estimate the expected returns and risk over the next year.

First, we can calculate the daily returns of each stock using the adjusted closing prices. We can then calculate the average daily return and the standard deviation of daily returns for each stock over the past 3 years. We can annualize these values by multiplying them by the square root of 252 (the number of trading days in a year).

Text

Description automatically generated with medium confidence

We can then calculate the correlation between the two stocks over the period.

Graphical user interface, text, application

Description automatically generated

Assuming a risk-free rate of 1%, we can then use the following formula to calculate the optimal allocation:

Graphical user interface, text, application, email

Description automatically generated

Chart

Description automatically generated

**References:**

1. SciPy. (2021). Scipy: Normaltest. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html>
2. Statsmodels. (2021). Statsmodels: OLS. <https://www.statsmodels.org/stable/generated/statsmodels.regression.linear_model.OLS.html>
3. Statsmodels. (2021). Statsmodels: Graphics.Tsaplots. <https://www.statsmodels.org/stable/generated/statsmodels.graphics.tsaplots.plot_acf.html>
4. Matplotlib. (2021). Matplotlib: Pyplot. <https://matplotlib.org/stable/api/pyplot_summary.html>