**Bull Market or Bear Market: Time Series Price Prediction for Q1 2024:**

John Vincent Deniega, Ravita Kartawinata, Gabi Rivera

Master of Science in Applied Data Science, University of San Diego

**Author Note**

We have no known conflict of interest to disclose.

Correspondence concerning this work should be addressed to the authors above. Email:

jdeniega@sandiego.edu, rkartawinata@sandiego.edu, gabriellarivera@sandiego.edu

***Background Information***

Building forecast models and financial tools to predict major market events such as stock price movement and what factors influence the price direction, is highly lucrative. It’s not only beneficial to the public but more intently the business sectors, government, economist, financial professionals are keen to have predictive insight on making financial and investment decisions. The most recent implications of down-ward stock price are inflation and recession. In 1994 the Board of Governors of Federal Reserve System International Finance Discussion Paper, Ammer mentioned that higher inflation is related to both of lower stock dividend and lower equity return. In addition, as quoted in Kiley 2023, the word “recession” becomes abundantly searched on the internet typically at the beginning of each prolonged downturn in economic activity. There is an apparent interest in a looming downturn to make better financial decisions. As expected, this search trend was captured during the start of the great recession which was influenced by the housing market crash in 2008 (Hudomiet et. al., 2011). This was then followed by the most recent COVID-19 recession which was the second steep decline in economic activity in the last 15 years. This paper was initiated by finding this recent recession and correlating with stock price movement.

Since then, there has been notable literary data derived from 42 recessions in 14 countries using quarterly periods that have high potential for correctly forecasting the next recession (Kroencke, 2022). It is said that stock prices dropped significantly at about 30% at the start of recession while dividends fell on average by 13% (Kroencke, 2022). Prices of stocks seem to fall further compared to real dividends during recession periods with the stock price variance behaving relatively the same as dividend growth variance (Kroencke, 2022). Another likely indicator of recession would be the timing of price-dividend ratio drop which happens in two quarters before recession ensues at around 5.6% (Kroencke, 2022). There’s also recession variance ratio to look at which is the “recession variance over the pre-recession variance” (Kroencke, 2022). For price changes, the recession variance ratio goes up 2.1-fold which is relatively similar to dividend growth at 1.7-fold (Kroencke, 2022). This bank of historic information can be used to help assess forecast models during exploratory data analysis to guide the direction of the model development and interpretation.

Currently, the US stock market is experiencing significant volatility in the last quarter of 2023. This is a perfect time to evaluate the many factors that influence prolonged market downturn such as financial markets, other investments, and current events to determine whether a recession will ensue. Through deeper understanding of the underlying characteristics of the markets and current environment across multiple dimensions, we may be able to leverage our growing knowledge of time series analysis to decompose price movements by price level, price trend, quarterly seasonality, and noise. The goal is to predict with sound data science principles where the US stock market will be by the end of the first quarter of 2024 to ultimately manage risk and build capital.

***Overview of Main Points***

Key components of the stock market include stock exchanges, such as the New York Stock Exchange (NYSE) and NASDAQ, and various financial instruments like equities and exchange-traded funds (ETFs). In this case we are looking at S&P500 ETF (SPY) and Amazon (AMZN) to compare side-by-side the movement of their stock pricing. Fluctuations of the market often occurred on a short-term basis. Therefore, we are looking at short-term time series methodology such as moving average, simple smoothing, exponential smoothing, ARIMA and logistic regression as well as neural network to achieve our goal. Exploratory data analysis will be guided by gathered historical data and empirical data will be used to determine external factor predictors for the forecasting model. Throughout the ADS506 course, other methodology would be included in our analysis.

The stock market is subject to a vast and virtually limitless array of external factors that can impact its performance. Having this large complexity and limited domain knowledge, identifying time series models and forecasts for this module would be a challenge.

***Key Findings***

Dataset is originated from a real-time yahoo finance data. It’s a live data using Python application. The API ticker function retrieves comprehensive daily logs of the stock market movements and metadata. yfinance returns a panda dataframe with multi-level column names from opening prices, high, low, and closing prices as well as volume. Data type of these attributes is class; a one-dimensional table array.

Preliminary EDA findings include the time series of SPDR S&P 500 ETF’s stock price exhibiting non-stationarity over the period of six months from May 2023 to November 2023 using the Augmented Dickey-Fuller test statistic method. Upon applying decomposition on the data, the trend and seasonality needs to be considered before feeding into the model’s like ARIMA that assumes lack of trend and seasonality.

Amazon was also explored at a 5-year span. The original closing price time series fluctuates between $60 to $190. There was a rising trend from 2020 to mid 2021 followed by a downward trend up to 2023’s comeback upward trend. Similar to SPY, AMZN time series displays non-stationarity. The application of LOESS (Locally Estimated Scatterplot Smoothing) decomposition revealed a 5 year downward concave trend with undetectable seasonality when residual is factored. Anomaly detection to account for noise factor will be applied to residuals and evaluated followed by model fitting and forecasting.

***Data Preprocessing***

Data for the stock IDs were pulled from Yahoo Finance Library at 5-year span. No missing data values were observed during initial time series plotting approach on market days. Missing values occurred on weekends and holiday dates. These missing values were then imputed and filled in by propagating the last valid value to the next. Then the time series datasets were converted to a specified frequency in order to be used and fed into time series exploratory tools and models. Data smoothing and differencing were rationalized during exploratory data analysis.

***Exploratory Data Analysis***

Determination of stationarity status was the first step of the platform process. Augmented Dickey-Fuller (ADF) test was employed and based of from the p-values greater than 0.05 significance level results for SPY and AMZN, the time series datasets were both determined not stationary. This is visualized in Figures 1a and 2a from the original time series plot. Figures 1a and 2a also shows the next approach which was to subject the time series to (STL) Seasonal-Trend Decomposition using Locally Estimated Scatterplot Smoothing (LOESS) regression. SPY and AMZN’s trend and seasonal components were clearly parsed in both cases. During transformation to remove the trend and seasonal elements for model exploration, first degree differencing of the series was utilized. ADF tests of p-values lower than 0.05 significance level confirms that the time series were converted to stationary datasets. Figures 3a and 3b presents the autocorrelation (ACF) and partial correlation (PACF) plots of SPY and AMZN. For SPY, ACF suggest lags of [1, 2, 3, 8, 9, 11, 14, 21, 58, 63, 70] can be used to explore AR p parameter value and PACF suggest lags of [1, 2, 3, 8, 9, 11, 14, 21, 58, 63, 70] can be used for AR p parameter value. For AMZN, ACF suggest lags of [1,6,10, 20,31, 32] can be used for MA q parameter value and PACF suggest lags of [1,6,10, 20,31, 32] can be used for AR p parameter value.

As shown in Figures 1b-d and 2b-d, detection of anomaly through the use of STL regression was performed for SPY500 and AMZN. The residual was taken and subjected to ±3 standard deviation threshold (Figures 1c and 2c) to reveal anomalous spike and dips from the original stock curve (Figures 1d and 2d). For SPY500, anomalies were detected during the 2 most recent stock market crashes of 2020 and 2022. Red marks outside of the 99.7% residual’s normal distribution were identified during the times the stock market is the most volatile. For Amazon, it is noticeable that two red marks were detected during the 2020 recession when the overall stock market was crashing. Amazon did well during that recession period compared to majority of the market movers as shown in SPY500 2020 crash. Year 2022 seems to be having increased anomalies up to the most recent stock decline much like SPY500 is at the 2022 mark. Overall, analyzing the time series’ normal distribution provides a good way to detect pending market instability.

Figures 1e and 2e shows additional data exploration on the weekly behavior of the two stocks. SPY500 reveals to have positive price action on Mondays with price avg of $0.96, median of $0.86, and volatility of $2.09 fluctuation. Therefore, Mondays are considered the least volatile day of the week for SPY500. Compared this to Thursday’s, the price action of -$0.07 avg price, $0.024 median, and $3.72 volatility score. Thursdays is then considered the most volatile day of the week for SPY500. For amazon, Tuesdays are the least volatile at -$0.10 avg price, $0.05 median, and $1.58 volatility score. However, this is immediately followed by the most volatile day which is on Wednesdays at -$0.05 avg price, $0.11 median, and $2.22 volatility score.

**Modeling**

***Selecting Modeling Techniques***

Daily stock market logs were mined at the time of API call. Data and associated data types that were retrieved consisted of stock prices (numerical float), volume (integer), and time (Pandas datetime). An additional variable of interest was derived from stock prices to indicate whether or not the daily price for the time period queried resulted in a positive net increase in price (binary). This derived variable serves as the dependent *‘y’* variable for a logistic regression model to be trained to provide a forecast targeting price increase logic.

In order to accomplish this, additional variable predictors were derived as intermediates in the process of fitting the logistic regression model to accomplish this binary outcome. These new predictors are *‘open\_close’* as well as *‘high\_low’* which are used to represent the difference between the two original predictors as a combined item in both cases. The *‘open\_close’* stock price was selected as the most complete predictor of the period that smoothed out intraday price fluctuations. This thereby simplified the numerical difference into a binary outcome field, *‘positive’*, as either *‘0’* or *‘1’* with the latter indicating that the price increased during that discrete period. The binary transformation is used intently to capture the closing stock market price behavior as positive signifying up and negative to down. The time series data – now with additionally constructed features – were then differenced in order to introduce stationarity into the resulting time series for modeling. DataFrames at lag periods one through five were separately constructed for their respective logistic regression model in order to select the best performing lag period and other parameters against the validation set.

In comparison to other methods, the original time series data types pulled from the *yfinance* API were subjected to first-ordered differencing of the closing price predictor before feeding the data into Autoregressive Integrated Moving Average (ARIMA), Simple Exponential Smoothing (SES), and Advanced Exponential Smoothing (AES) methods. The assumption is that the time series exhibits stationarity to remove the influences of both trend and seasonality from the methods. The goal of these aforementioned methods is to predict the closing stock prices at a given target future date solely based off of historical data. Inspection of the SPY500 data’s autocorrelation plot in Figure 3a revealed potentially significant moving average orders at periods 0, 1, 2, 11, 13, 14, 21, 37, and 40. Subsequent inspection of the SPY500 data’s partial autocorrelation plot in Figure 3a revealed potentially significant autoregression orders at periods 1, 2, 3, 8, 9, 11, 14, 21, 58, 63, and 70. This information was key to determining which parameters to iterate through in order to discover the highest performing parameters for lag and moving average orders of ARIMA and AES’s individual models. Same goes with Amazon, periods 1, 6, 10, 20, 31, 32 from ACF and PACF were used to determine the p,d,q parameters of ARIMA.

***Generating a Test Design***

As mentioned earlier, both model-based and data-driven methods (i.e. ARIMA & AES) were used to develop price predictors and logistic regression as well as neural network models to represent stock market movement through price prediction and binary classification. For the price prediction objective, statistical measurements such as root mean square error (RMSE) and mean absolute percentage error (MAPE) were used to evaluate predictive performance between model predictions and truly observed values due to their interpretable values in being in the same unit as the original series as well as a relative proportional measurement as a percentage of the error, respectively.

Additionally, in order to evaluate with consideration for both fitness of the data with complexity of the method used, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were considered when selecting the best method to deploy on unseen data. Of these two criteria, a lower number of parameters and a higher log likelihood of fitting the data aim to mitigate the effects of overfitting a model to the training data.

For binary forecasting, a confusion matrix provided performance metrics for the purpose of evaluating the predictive performance of the classifier against the validation data. Precision, sensitivity, F1 score, and accuracy were included in the confusion matrix to assist in different models matching different investment risk profiles. For more conservative risk profiles, high precision models may be preferred due to their focus on higher probability positive price movements. For more risk-tolerant and higher return risk profiles, high accuracy models may be preferred so as to optimize for higher recall on positive signals, but counter-balanced with the higher specificity on non-positive signals (assuming equally weighted costs on both positive and non-positive periods).

***Building the Models***

SES, AES, ARIMA, logistic regression, and neural network forecasting methods were explored throughout the model building process. SES and ARIMA were omitted out for their insignificant predictive performance. For SPY 500, the data-driven AES method exhibited the highest performance based on RMSE and MAPE among the price predictor models. Among binary forecasters, Cross-sectional MLP (Neural Network) Model exhibited the highest performance based on confusion matrix results. With respect to their respective parameters for SPY500, AES shown in Figure 4c was set to have multiplicative damped trend, additive seasonality, three periods per season, and a heuristic initiation method. Fit was set at 0.1 for both smoothing level and trend parameters. These were determined by running AES through multiple combinations of each parameter such that for both trend and seasonal parameters ‘*add, mul, additive, multiplicative, and none*’ were iterated through to find the combination that yielded the best performing metrics. The same process was performed for parameters in damped trend, seasonal period length, initiation methods, and multiple fits for SPY500. Then, each parameter was subjected to predictive performance criteria sequentially. The validation dataset used for the calculation of the performance criteria is set from the last year period and ending on November 27, 2023. The results are further discussed under assessment and evaluation.

As for SPY 500 logistic regression, the *‘positive’* field was set as the dependent outcome variable. The stock price predictors along with the new predictors derived by lagged periods were differenced at three periods before feeding into the logistic regression model. Then, validation data was set to the next 200 periods after the training model, resulting in the confusion matrix found in Figure 4d. Figure 4e shows the confusion matrix result of the cross sectional MLP (Neural Network) model. Combined iteration of the different MLP parameters were run to determine the best performing hidden layer size, activation, solver, and max iteration number for the model. Grid search was able to produce tanh for activation, 2 as the size of the hidden layer, 2000 for the max iteration, and sgd as the solver method to be used.

For Amazon, auto ARIMA and ARIMA (Figures 5a-b) were also omitted out due to low performance. The AES trend and initialization method was subjected to different iterations of ‘add’ and ‘mul’ as well as None, 'estimated', 'heuristic', 'legacy-heuristic' options, respectively. It was determined that the model performs better at None trend and heuristic method parameters. Seasonal component of AES was defaulted to mul (multiplicative) as add iteration performed poorly visually. The seasonal period was set to 252 for the total number of trading days per year that needed to be forecasted ahead. The model fit smoothing level and trend were both set to 0.5 to correct for the model’s visual location against the training data using the plot as shown in Figure 5c. Amazon’s logistic regression model building approach was the same to SPY500 written above with the confusion matrix result shown in Figure 5d. The last model for AMZN was recurrent neural network (RNN) forecast model as shown in Figure 5e. The timeseries was first normalized using min-max scaling before feeding the dataset into the model. Then the data frame was partitioned for training and test datasets and run into sequential model with simpleRNN and dense method added. After fitting the model and forecast made, the results were transformed back to original scale using inverse transformation on the min-max scaler. The results are then discussed under Assessment and Evaluation.

***Model Assessment and Evaluation***

Figure 4a shows SPY500 SES model forecast with poor predictive curve spanning across the trailing year range. The AIC score was recorded at 3042 while BIC was recorded at 3051, which is to be compared with models throughout this discussion. RMSE was recorded at 31.71 and MAPE was recorded at .06. This may be interpreted as an error of $31.71 USD and a 6% price deviation.

Figure 4b shows SPY500 ARIMA model forecast performing poorly in predicting the validation set similar with SES plot. The forecast curve flattened out throughout the last year prediction range. For performance metrics review, optimal ‘*pdq’* parameters of 14, 1, 1, respectively, yielded an AIC of 10377 and a BIC of 10466. The RMSE was recorded at 418.23 and the MAPE was recorded at 1.0. This may be interpreted as an error of $418.23 USD and a 100% price deviation, suggesting that even with parameter optimization, ARIMA may be unfeasible for price prediction purposes. For AMZN, Figures 5a-b shows the plot for auto-ARIMA and ARIMA with parameters p,d,q manually iterated for best performing outcome. Both models performed poorly as illustrated with a flat forecast.

As shown in Figure 4c, the SPY500 AES method was able to forecast the exponential pattern compared to the validation data. Although, the trend and seasonality patterns were not explicitly reflected to also match the actual validation results it appears that the smoothing sufficiently compensated for increased predictive performance. Among all of the price prediction models, the data-driven AES method performed the best against the validation set with half of the error of SES for only a one-fifth higher information criterion with AIC at 3884 and BIC at 3923. The RMSE was recorded at 15.06 and the MAPE was recorded at .03. This may be interpreted as an impressive minimal error of $15.06 USD and a 3% price deviation as of November 27, 2023’s data retrieval depicted in Figure 10. For AMZN, the AES model performed way better compared to ARIMA. As shown in Figure 5c, the weekly ups and down of the closing prices are capture. However, AES is not able to forecast the steep incline of the validation dataset.

Figure 4d shows SPY500 logistic regression with resulting confusion matrix shown in Figure 9. The overall accuracy is at .62 with precision at .77 and sensitivity at .61. This may be interpreted as a model that is likely best suited for more conservative risk profiled investors due to its higher precision performance. For AMZN, logistic regression did poorly at around 51% accuracy.

Figure 4e shows SPY500 cross-sectional MLP (Neural Network) confusion matrix with accuracy of 93%. Recurrent Neural Network was performed to forecast AZMN. MAPE score was at 1.8% with MSE of 8.2. Looking at Figure 5e, RNN forecast model outperformed all the methods used for AZMN as it was able to predict the validation dataset as closely as possible.

**Figure 1a**

*STL Decomposision of SPY500 Time Series Using LOESS Technique.*

A graph of different types of data

Description automatically generated with medium confidence

**Figure 1b**

*SPY500 Closing Price (USD) Time Series in 5-year Span.* A graph showing the growth of stock prices

Description automatically generated

*Note.* Orignal and Estimated (STL’s Trend & Seasonal) projection.

**Figure 1c**

*SPY500: STL decomposision Residual at Thresholds of ± 3 Std Dev of Normal Distribution.*A graph showing a wave of sound

Description automatically generated with medium confidence

**Figure 1d**

*SPY500 5-yr Time Series with Marked Anomalies Outside of 99.7% Normal Distribution.*

A graph with numbers and lines

Description automatically generated

**Figure 1e**

*SPY500 Price Change Vs. Days of the Week Boxplots.*

A graph of different colored squares

Description automatically generated

**Figure 2a**

*STL Decomposision of Amazon Time Series Using LOESS Technique*

*A graph of different types of sales

Description automatically generated with medium confidence*

**Figure 1b**

*Amazon Closing Price (USD) Time Series in 5-year Span.*

*A graph showing the stock market

Description automatically generated*

*Note.* Orignal and Estimated (STL’s Trend & Seasonal) projection.

**Figure 1c**

*Amazon: STL decomposision Residual at Thresholds of ± 3 Std Dev of Normal Distribution.*

*A graph showing a sound wave

Description automatically generated*

**Figure 1d**

*Amazon 5-yr Time Series with Marked Anomalies Outside of 99.7% Normal Distribution.*

*A graph with lines and numbers

Description automatically generated*

**Figure 1e**

*Amazon Price Change Vs. Days of the Week Boxplots.*

*A graph of different colored squares

Description automatically generated*

**Figure 3a**

*SPY500 ACF and PACF Plots.*

A graph with blue dots

Description automatically generated A graph with blue dots

Description automatically generated

**Figure 3b**

*Amazon ACF and PACF Plots.*

A graph with blue dots

Description automatically generated A graph with blue dots and numbers

Description automatically generated

**Figure 4a**

*SPY500: SES Forecast Model Plot.*

A graph with blue lines

Description automatically generated

**Figure 4b**

*SPY500: ARIMA Forecast Model Plot.*

A graph with orange and blue lines

Description automatically generated

**Figure 4c**

*SPY500: AES Forecast Model Plot.*

*A graph with orange and blue lines

Description automatically generated*

**Figure 4d**

*SPY500: Logistic Regression Confusion Matrix.*

A screenshot of a graph

Description automatically generated

**Figure 4e**

*SPY500: Cross-sectional MLP (Neural Network) Model Confusion Matrix.*

A screenshot of a graph

Description automatically generated

**Figure 5a**

*Amazon: Auto-ARIMA Forecast Model Plot.*

A graph showing the growth of the stock market

Description automatically generated

**Figure 5b**

*Amazon: ARIMA Forecast Model Plot.*

A graph showing the growth of the stock market

Description automatically generated

**Figure 5c**

*Amazon: AES Forecast Model Plot.*

**A graph showing a line graph

Description automatically generated with medium confidence**

**Figure 5d**

*Amazon: Logistic Regression Confusion Matrix.*

**A screenshot of a graph

Description automatically generated**

**Figure 5e**

*Amazon: Recurrent Neural Network (RNN) Forecast Model Plot.*

**A graph of a graph

Description automatically generated**

**References**

Chatgpt. ChatGPT. (2023, November 10). http://www.openai.com/product/chatgpt

Industrial Business Machines Corporation (2021). *Introduction to CRISP-DM*. Industrial

Business Machines Corporation. https://www.ibm.com/docs/en/spss-modeler/saas

?topic=guide-introduction-crisp-dm

Hudomiet, P., Kézdi, G., & Willis, R. J. (2011). *Stock Market Crash And Expectations Of American Households*. Journal of applied economics, *26*(3), 393–415. <https://doi.org/10.1002/jae.1226>

Kiley, M. T. (2023, January 16). *Recession Signals and Business Cycle Dynamics: Tying the Pieces Together*. Finance and Economics Discussion Series 2023-008. <https://doi.org/10.17016/FEDS.2023.008>

Kroencke, T. A. (2022, July 7). *Recessions and the stock market*. Journal of Monetary Economics. <https://www.sciencedirect.com/science/article/pii/S0304393222000976>

Ammer, J. (1994, April). Inflation, inflation risk, and stock returns - Federal Reserve Board. <https://www.federalreserve.gov/pubs/ifdp/1994/464/ifdp464.pdf>