Time Series Forecasting of Apple Inc's Quarterly Net Sales

College of Information Technology │ Western Governor's University │ Graduate Capstone

Written Report By: Nicole Burns

Date: April 9th, 2024

D214 — Data Analytics Graduate Capstone

Problem Statement

Since its inception in 1976, Apple has grown to become the first company with a 3 trillion-dollar market value (Vlastelica, 2023). Because of this, Apple's financial performance reflected by its quarterly net sales, is of keen interest to stakeholders such as investors, analysts, and competitors. These stakeholders rely heavily on an accurate understanding of Apple's financial health and its growth trajectory to make informed decisions. The ability to forecast future sales with a high level of accuracy is therefore essential, influencing a range of strategic decisions from investment strategies to marketing and resource allocation. Moreover, a proficient forecast of Apple's quarterly net sales provides insights beyond simple projections. It can reveal the underlying patterns, such as seasonality (e.g. higher sales during the holidays or product launches) and cyclical trends. Given its capability to account for seasonality, the SARIMA model is the best technique to use for this study (Patra, 2023).

Hypothesis

Vaswani and Prasad (2023) performed a study utilizing a SARIMA model that resulted in the ability to forecast sales of wholesale trade to retail trade with 90% accuracy. This study aims to build upon that foundation by evaluating whether a SARIMA model can effectively forecast Apple's quarterly revenue with more than 80% accuracy. The null hypothesis suggests that it cannot, and the alternate hypothesis suggests that it can.

The Data Analysis Process

The analysis process consisted of downloading a CSV of Apple's quarterly income statement via Yahoo Finance, spanning from September 1985 to December 2023. The data is comprised of 51 rows detailing the financial data names, and 156 columns representing the dates of each quarter's end. The analysis was performed using Python as the preferred programming language, within the Jupyter Notebook environment.

Following the importation of the CSV, the dataset was transposed, switching the financial names to the column headings, and the quarterly dates to the row labels. Numerical values were converted from objects to floats, commas were removed, and the values were reduced to improve readability. All columns not pertinent to the analysis were removed. A graph was generated as well as an Augmented Dickey-Fuller (ADF) test, for inspection of stationarity.
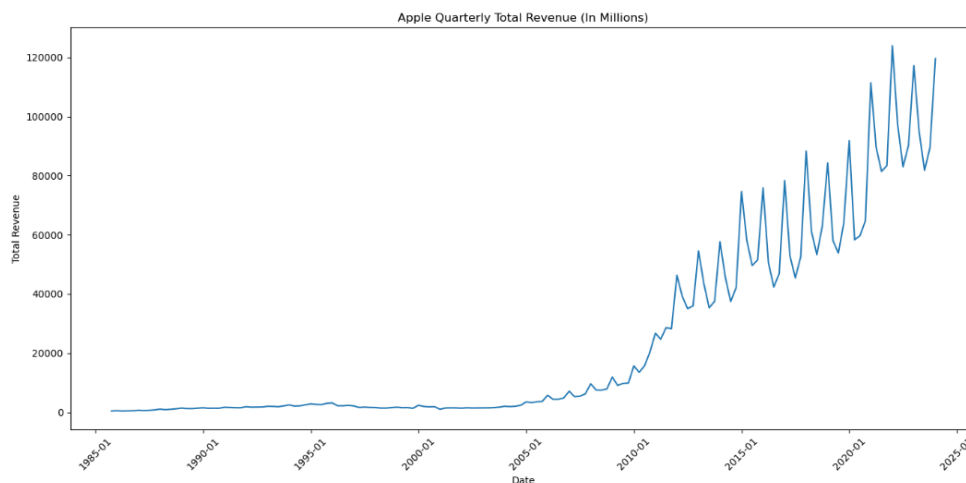
The graph above displays a distinct upward trend and the Augmented Dickey-Fuller (ADF) test returned a p-value of ~0.99, indicating non-stationarity within the dataset. Delving deeper into the dataset's characteristics, seasonal decomposition was employed. In the graph

below, there is a very clear consistency of growth shown in the trend and repeated intervals of up-and-down movements shown in the seasonality. These patterns likely correspond to periodic events (holidays) or sales cycles, which are crucial for understanding the timing and magnitude of Apple's sales dynamics.
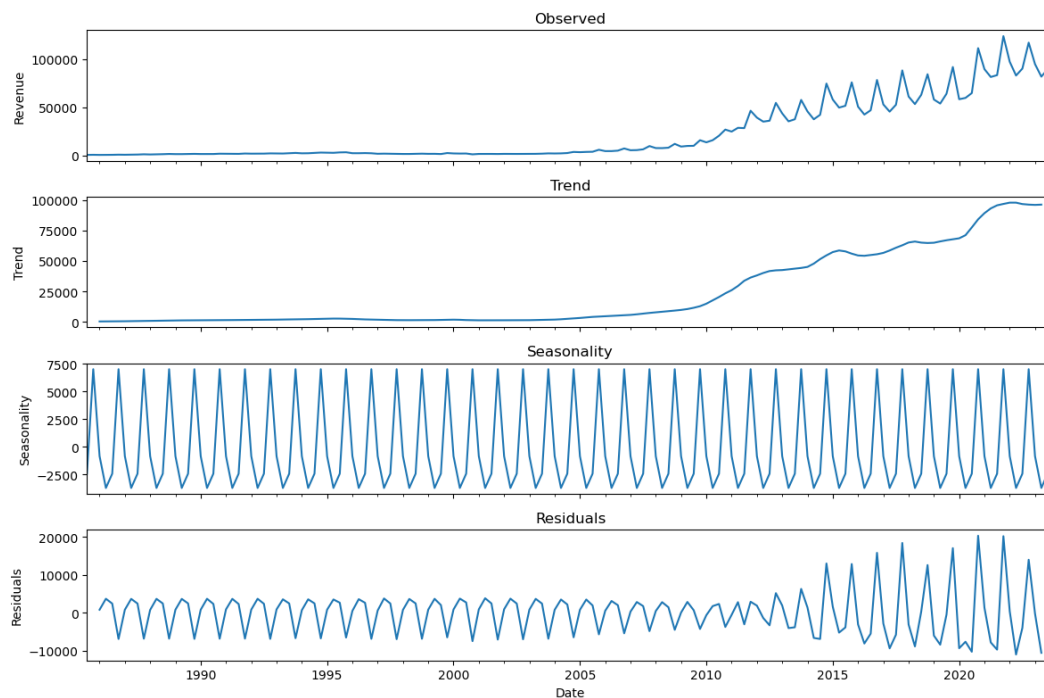

The residuals were relatively small, suggesting that the trend and seasonal components accounted for a significant portion of the data's variation. Upon splitting the data into test/validation sets, the 'auto_arima' function from the 'pmdarima' package in Python was employed. This function automates the selection of the best parameters for the SARIMA model, simplifying the model-building process. After model fitting, additional diagnostic plots were generated to assess the SARIMA model's fit to the data. In the final stage of the analysis, the model was used to generate forecasted values, a graph was created to visualize both the historical data and the projections, and error metrics such as Mean Absolute Error (MAE), Root Mean

Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) were calculated to quantify the forecast accuracy.
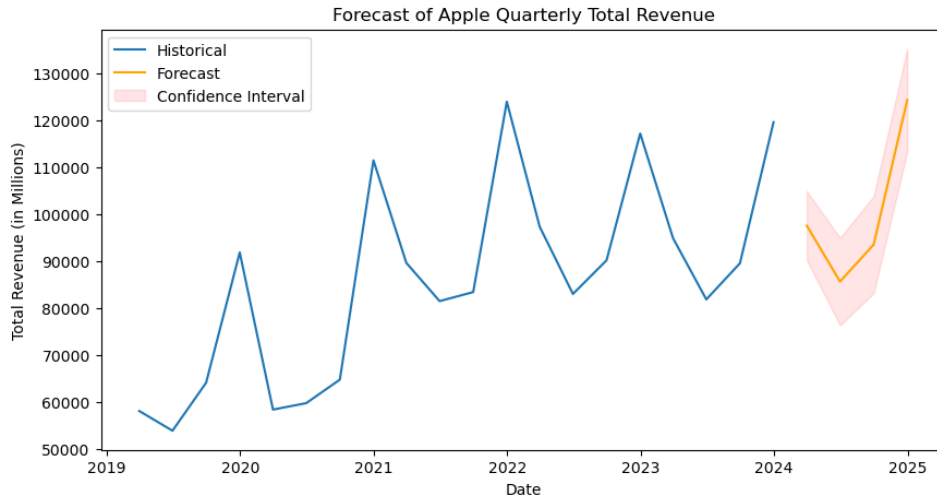
The Results

As previously mentioned, following the model's creation were additional plots to assess the model's fit to the data. It was observed that the model had managed to account for a significant amount of the data's autocorrelation, which is a positive sign. However, the distribution and extreme values in the residuals suggest that the model could be improved. The spikes in residuals over time, the skewed histogram, and the deviations from normality of the tails in the Q-Q plot indicate that the model may not be fully capturing all the data's underlying patterns.
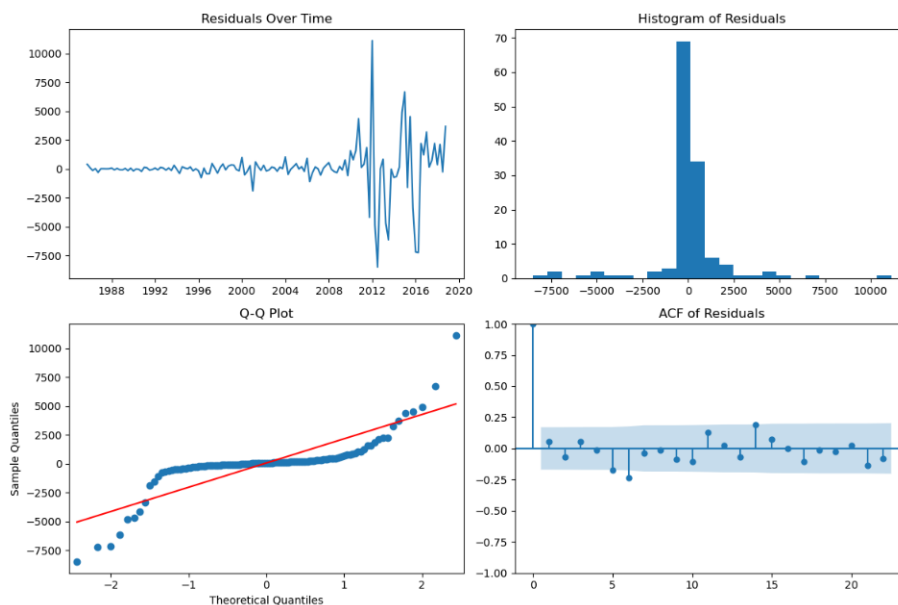
The graph below visualizes Apple's historical quarterly revenue along with projections spanning 21 quarters into the future. The wide confidence intervals suggest that there is uncertainty in the forecast as time increases.
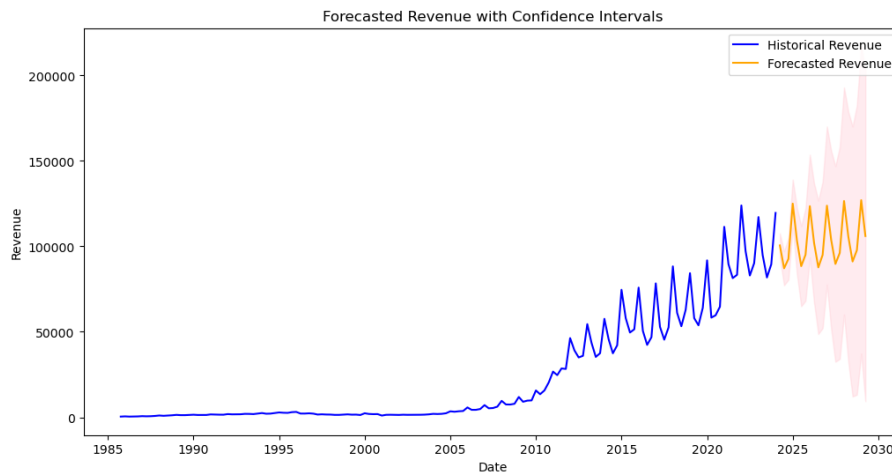
MAPE (Mean Absolute Percentage Error), a calculation that determines how accurate forecasts are when comparing them to the actual values (Azaria, 2024), resulted in an error percentage of 33.0%. This suggests that on average, the forecasted values deviate from the actual values by 33%. It can be inferred that the model has a moderate ability to make predictions with 67% accuracy. On the other hand, when experimenting with a more recent historical timeframe and a reduced projection period (5 years historical/1-year forecast), the MAPE significantly improved, reducing to 3.97%. This improvement in the MAPE suggests that the model's predictions are about 96% accurate for this shorter forecasting period.

Forecast of Apple Quarterly Total Revenue

Limitations

The results above indicates that the SARIMA model is much more effective at capturing and forecasting the patterns in the data over the short term compared to the long term. However, there are some important things to keep in mind. Apple only records revenue on an annual and quarterly basis, limiting the amount of data available for the model to be trained on. The more data for training, the more reliable the prediction accuracy. It is also best to consider the relevance of the time span of the dataset. Years spanning as far back as 1985 for example, are considered irrelevant to the analysis in comparison to recent times, due to the significant changes in the economy or regarding Apple's current consumer sentiment. This is why forecasting too far into the future creates uncertainty. It's tough to account for unexpected market shifts when forecasting.

Proposed Actions

It's advised to employ the SARIMA model primarily for short-term revenue forecasting, treating it as an informative tool rather than a prediction. It should be used in conjunction with business insights and other market knowledge. Additionally, future research should include external factors like consumer sentiment, competitive actions, economic indicators, and the impact of new product launches to build a model that better accounts for external influences and trends. Pairing SARIMA with other forecasting methods or machine learning models may also leverage various strengths and increase prediction accuracy.

Expected Benefits

The expected benefit of this analysis was to generate a SARIMA model capable of utilizing Apple's historical quarterly revenue to effectively forecast future quarterly revenue with > 80% accuracy. The analysis revealed that while the model exceeded expectations in short-term forecasting using recent sales data, its ability to achieve the desired level of accuracy over longer periods is constrained. This insight enables stakeholders interested in Apple's future performance to use this information as the foundation for their strategic decisions. It's advisable for such decisions to also take into account external factors and to utilize a combination of forecasting methods to validate the results obtained in this study.

Sources

Title Page Photo by Unknown Author is licensed under CC BY-NC and provided by Bing

Patra, T. D.-D. (2023, September 24). *Sarima vs Arima for Timeseries Analysis Model*. Medium.
 Retrieved April 3, 2024, from https://dhirajpatra.medium.com/sarima-vs-arima-for-
 timeseries-analysis-model-a600ab544b1f

Vaswani, A., Prasad , Dr. P. C., & Padhi, Dr. P. K. (2023). Journal of University of Shanghai for
 Science and Technology. *Time Series Analysis: An Application of SARIMA Model in
 General Trade to Forecast Sales*, *25*(3), Retrieved April 3, 2024, from
 https://jusst.org/wp-content/uploads/2023/03/Time-Series-Analysis-An-Application-of-
 SARIMA.pdf. https://doi.org/10.51201/6735

Vlastelica, R., & Bloomberg. (2023, June 30). *Apple just made history by becoming the first
 company with a $3 trillion market value-'all powered by a device people look at 4 hours
 a day'*. Fortune. Retrieved April 3, 2024, from https://fortune.com/2023/06/30/apple-
 history-3-trillion-market-value/

Yahoo! (2024, April 1). *Apple Inc. (AAPL) income statement*. Yahoo! Finance. Retrieved April 1,
 2024, from https://finance.yahoo.com/quote/AAPL/financials