

Coca-Cola Stock Forecasting using ARMA Model

Zach Joachim, Taylor Jones, Trevor Wai, Carson Watkin

March 30, 2024

Abstract

The goal of this paper is to build a model to forecast stock prices and profits of the Coca-Cola Company. Using an Autoregressive Integrated Moving Average (ARIMA) model to create this forecast. Stock prices tend to follow trends during certain periods of time. We attempt to use this data to demonstrate that this data can be used to create reasonable predictions of the behavior of stock prices...

1 Problem Statement and Motivation

Stock forecasting is vital for a company that wants to be successful and profitable. It helps companies with financial planning, allowing companies to allocate resources efficiently and make informed decisions about investments and expenditures. It allows companies to identify and mitigate risks associated with a volatile market. Finally, along with many other things, forecasting provides a benchmark for the company's actual performance to be measured against. By comparing the forecasted stock prices with the actual prices, companies can evaluate the effectiveness of their strategies and operations.

Clearly the ability to forecast stock prices and other important quantities is of great use and interest to companies everywhere.

2 Data

The Coca-Cola stock price history data came from a dataset on Kaggle's website. The dataset contains 7 quantities indexed by the day of those quantities. These values are the stock's opening price, highest price point in the day, lowest price point in the day, the closing price of the stock, the volume, dividends, and stock splits of the stock for the day. Finding data like this was difficult to find for quantities like profits, margins, or quantities for specific drinks that fall under the Coca-Cola umbrella. However, the stock price for a company is a good indicator as to what the values are and where they are heading.

The stock price data required a little cleaning. For about 50% the data, the index or date for the stock info comes in the form of YYYY-MM-DD, while the other 50% has the time next to the date, in the same format, YYYY-MM-DD HH:MM:SS-UTC. For the purposes of our analysis, we only needed the date, and the format it was in was acceptable, so what we needed to do was delete part of the index that gave the time. To do this, we stripped the index into its list of strings and only kept the part that gave the date, we made this into a new column called `df['date']`, and deleted the original column.

3 Results

3.1 ARIMA Forecast

We wanted to see if an ARIMA model could accurately predict the 100 days into the future. Because stock prices fluctuate through the day we decided to look at opening prices, closing prices, the high for that day and the low for that day. Because we wanted to look at each of these different prices individually we were able to use the ARIMA model from statsmodels.

Insert Graphs here

When trying to predict the last 100 days of 2016 we found that the ARIMA model was not accurately able to capture the fluctuating prices of each day. Our model failed to capture the peaks and valleys and rather gave us a very linear prediction. Looking at all four different types of prices yielded very similar results neglecting the ups and downs that come from each day.

3.2 Random Forests

3.3 Time Series Decomposition of Closing Prices

Because prices of stocks fluxuate so much we decided to use a seasonal of prices to break down the data. It revealed the trend line as well as the seasonal periodic graph as well as the residual noise. We found this to be a good method to break down the data because stock prices are slightly periodic through the day and are definetly not perfectly linear and so this allowed us to notice trends in our data and see how it changes over time.

3.4 Taylor's Stuff

4 Analysis

4.1 ARIMA

4.2 Random Forests

5 Ethical Implications

The primary ethical concern of our project is the application of our analysis. Individuals wanting to manipulate the stock market would be able to use our model as their evidence to artificially manipulate supply and demand. This interferes with the free and fair operation of the stock market and is illegal in the United States. An example of how manipulators can use our analysis to affect the market is through stock bashing. Our analysis could predict the fall of a stock price and then perpetrators can use this as evidence for stock bashing or convincing the public that a stock price is worth less. Therefore allowing them to purchase the stock at a cheaper price to try and later turn a profit.

Therefore we should add a disclaimer in our project that all of the analysis and predictions created are simply models and not prophecies of events to come. Therefore people understand that our analysis is simply trying to predict how the stock prices are going to fluxuate and this can simply be a factor in their buying and selling of shares.

6 Conclusion