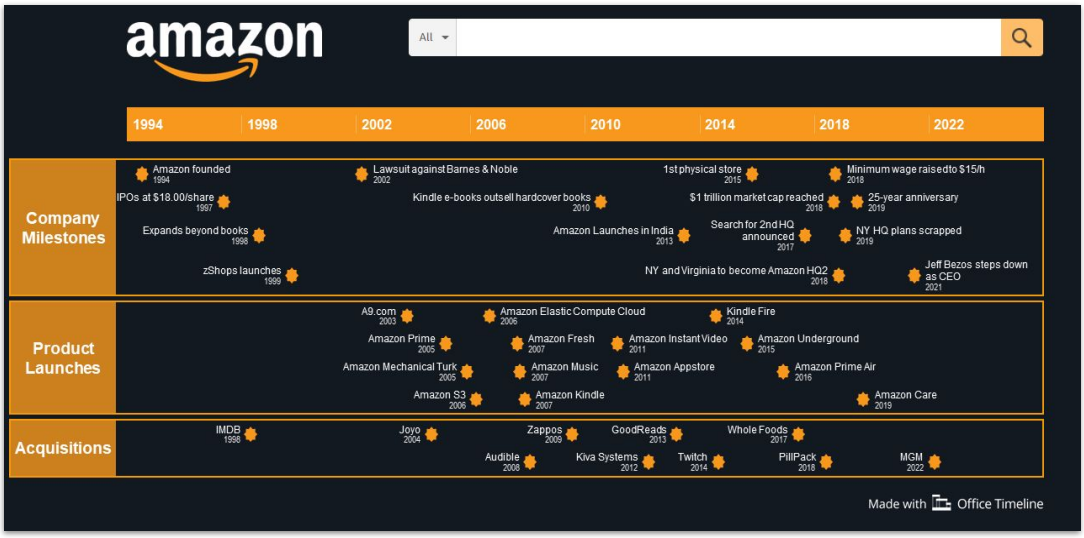
Amazon Stocks Machine Learning Project

Tech Talent & Strategy, Data Science | February-May '23 Cohort | Instructor: Preston Menke

Erik Barraza Cordova | James Brewer | Anabelle Capois Espinal | Christopher Endriga | Kalel Lopez | Layla Wynn

Company Background



Hypothesis

- Amazon's stock price will decrease into 2019
- Late into 2019 it will increase as demand would rise when stock price lowers
- Price increase leads to demand decrease in 2020
- The model prediction will not represent how the stock actually performed
    - COVID19 Pandemic = positive impact

About the Data

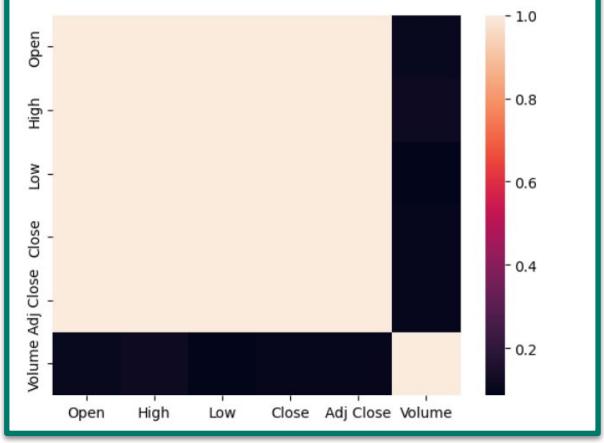
- Gathered from Yahoo FInance import
- Dates Used: 01/2016 - 01/2020
- Dates Predicted: 2019, 2020
- Robust = AMZN stock is consistently profitable regardless of market direction
- Some competitor data will be visualized to get a glimpse of market performance

	Open	High	Low	Close	Adj Close	Volume
count	1006.000000	1006.000000	1006.000000	1006.000000	1006.000000	1.006000e+03
mean	63.757561	64.318716	63.076373	63.729558	63.729558	8.576559e+07
std	23.675035	23.887567	23.394806	23.645184	23.645184	4.358994e+07
min	23.900499	24.674999	23.700001	24.103500	24.103500	1.762600e+07
25%	41.033751	41.115748	40.731499	40.925875	40.925875	5.633200e+07
50%	59.936501	60.270750	59.420500	59.735750	59.735750	7.314300e+07
75%	87.474253	88.173378	86.568748	87.498249	87.498249	1.015140e+08
max	101.905502	102.525002	100.650002	101.975502	101.975502	3.313000e+08

Data Correlation

- Volume of stocks sold has no linear correlation with the stock price
- The different stock prices utilized have a perfectly positive linear correlation between each other
  - They all represent the same price at different points of the day
- We will utilize the closing price

	Open	High	Low	Close	Adj Close	Volume
Open	1.000000	0.999656	0.999441	0.999112	0.999112	0.104011
High	0.999656	1.000000	0.999300	0.999548	0.999548	0.114650
Low	0.999441	0.999300	1.000000	0.999595	0.999595	0.086079
Close	0.999112	0.999548	0.999595	1.000000	1.000000	0.099484
Adj Close	0.999112	0.999548	0.999595	1.000000	1.000000	0.099484
Volume	0.104011	0.114650	0.086079	0.099484	0.099484	1.000000



Stock Market Terms

Naked Shorting

- The process of selling stocks that do not exist. For example, if AMazon only has 1 million shares as part of their market cap, but an investor wants to purchase 200 thousand shares but only 100 thousand shares are available, but the the order is still fulfilled, that is an example of naked shorting.

Short Selling

- Usually occurs when market makers/hedge funds believe a stock will drop in price, so they sell stocks before actually owning them to make profit. If their plan works, the pay back the shares that were borrowed/loaned out and keep the profits.

Bullish Market

- Is the time of when the market or certain stocks within the market are on an upward trajectory.

Bearish Market

- Is the time of when markets are on a downward trend as a whole.

What Influences the Stock Price?

Market trends

Most Common Halts For Common Traders

- LUDP Or LULD - This type of halt is triggered by high volatility trading. Good examples would be meme stocks like GME (Gamestop) and/or AMC. These stocks were considered volatile as it was encouraged by mostly online communities to buy the stocks to save them from going bankrupt. In return, these stocks soared in price which triggered halts. These types of halts last anywhere from 5-15 minutes once the price of the stock jumps to over 10% within a short timeframe.

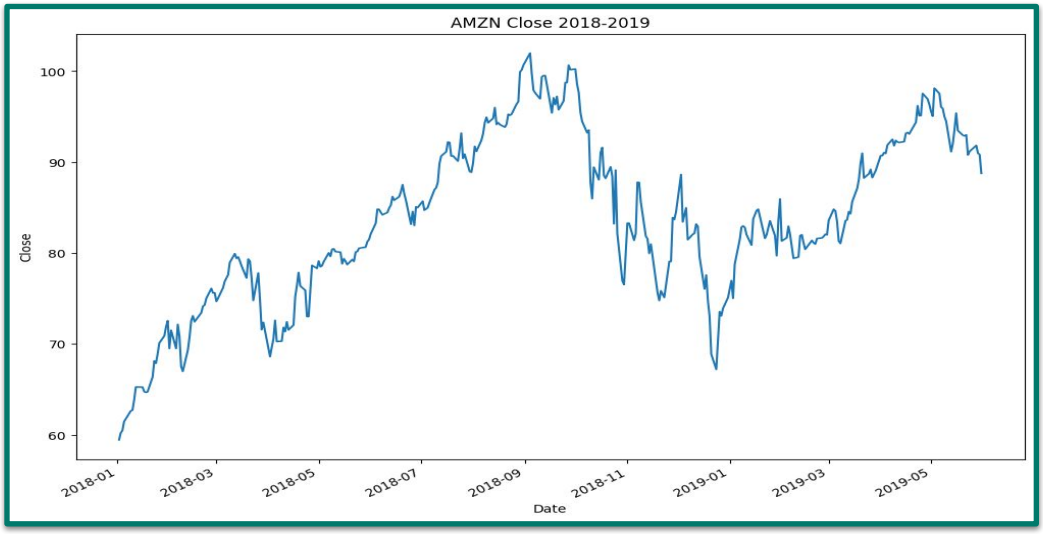
Presidencies or General Elections

- Incoming Presidents make changes/amendments as they see fit - Current rules/regulations could be more relaxed while new rules/regulations would be much more strict.

Natural Events

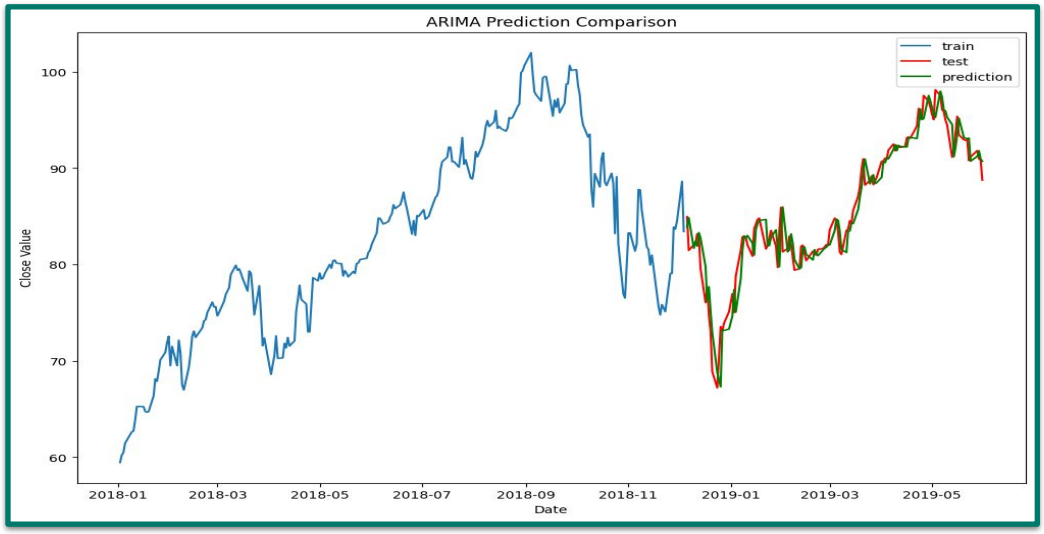
- COVID-19 is the most recent natural event that caused a shift of the stock market. Many jobs were lost which resulted into the rise of day-traders who solely focused of purchasing certain stocks and selling them within the same day, repeating this process until satisfaction.

2018 Training Data



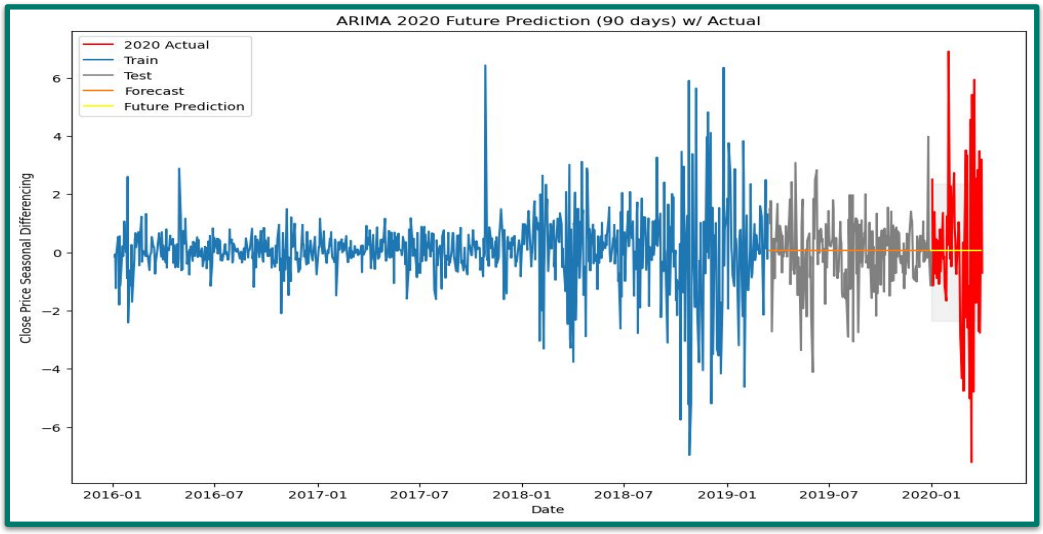
The stock performance for Amazon from January 2018 to May/June 2019 can be observed, this will be used as the training data for the machine learning model

ARIMA 2019 Prediction



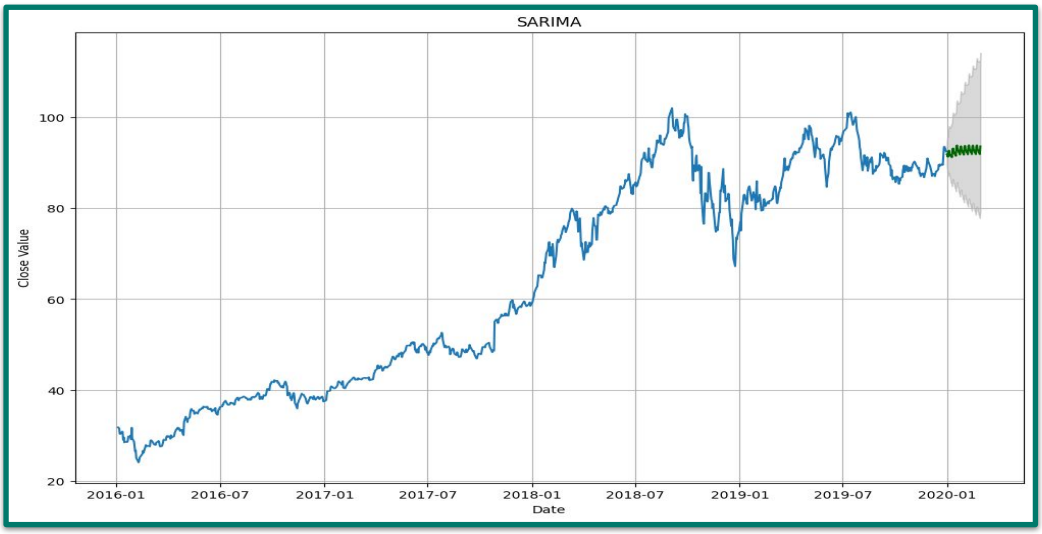
After using the ARIMA model we can see that the prediction of the model appears to be similar to what was actually observed for the stock from December 2018 50 June 2019

Stationary Arima 2020 Prediction



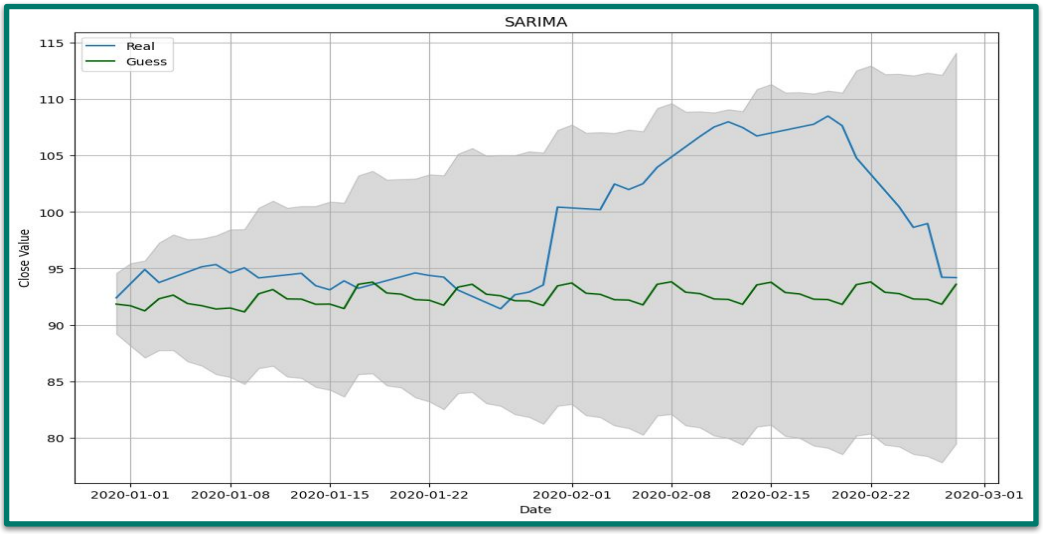
If we use the ARIMA model after converting the data to stationary data, we can predict the closing price's seasonal differencing values. In this model, data from 2016 to 2019 was utilized for the training data. 20% of the tail of the data was tested and we then predicted 90 days into 2020. As can be observed, the actual 2020 data was well beyond the confidence area, which could perhaps be explained in large part to the positive impacts the 2020 pandemic had on Amazon.

Visual 3 - SARIMA 2020 Prediction



Using the SARIMA model instead we can try to predict even further into the future. We used this model to look at what the stock price may be some days into 2020. The green line provides the mean, while the gray area represents the direction the price can go.

2020 Actual Stock Performance vs Prediction

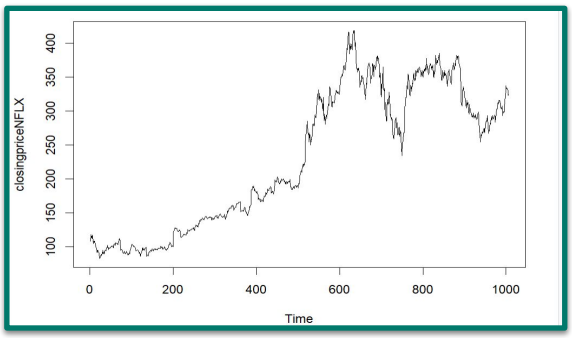


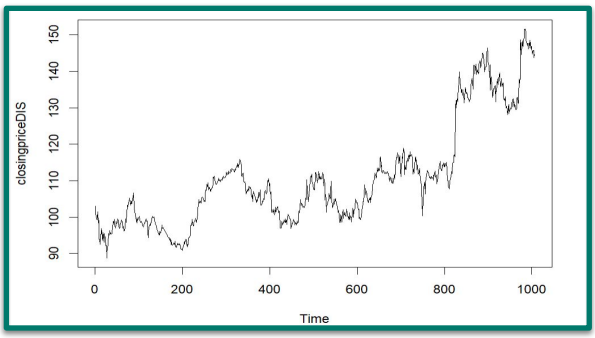
Here we can take a close look at the prediction area and compare it to what was actually observed in 2020. Early in 2020 we can see that the model is almost on par with the actual, however, as we go further into 2020 there is more of a discrepancy. This gap can more than likely be explained by the 2020 Global Pandemic, which had several household more dependant than ever on Amazon deliveries due to lock down. Surprisingly, the actual data is still within the bounds of the prediction model.

Delivery Services Competitor(s)

Here we can observe the the stock performance of one of Amazon's closest competitors in the retail/delivery service industry. They experienced a more steady upward trend than Amazon, but overall we can observe some similar patterns, such as a significant dip in January 2019. We can perhaps expect a similar prediction model result and actual performance as that observed for Amazon stock.

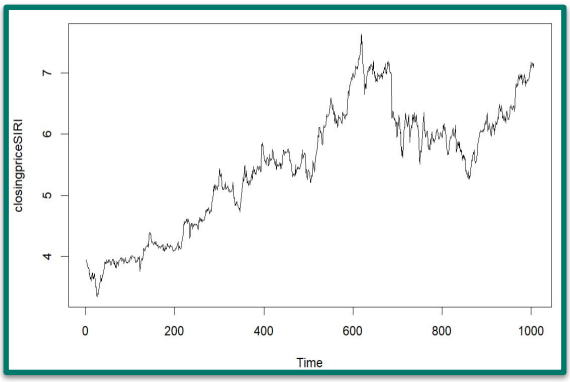
Video Streaming Competitor(s) 2016-2019

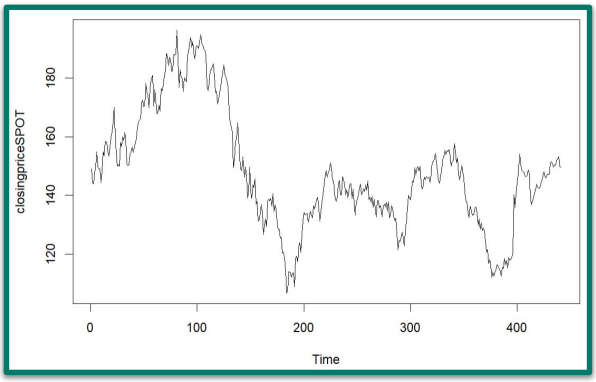




This visual is comparing closing price action of Netflix and Disney streaming services. The X axis represents the timespan in days. So from January 2016 to January 2020 is a total of 1000 days or 4 years. The Y axis represents the price at the time.

Music Streaming Competitor(s) 2016-2019

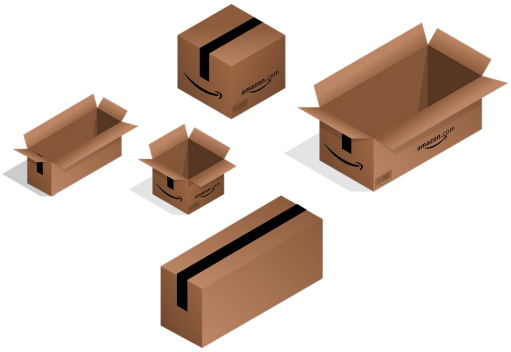




This visual is comparing closing price action of Spotify and Sirius music streaming services. The X axis represents the timespan in days. So from January 2016 to January 2020 is a total of 1000 days or 4 years.The Y axis represents the price at the time.

Takeaways

- There are several factors that influence the performance of a stock
- Machine Learning models are a useful tool for predicting stock prices with historical data of stock performance
- It is not always possible to have an accurate prediction due to events not before experienced, such as the COVID19 Pandemic
- A more complex model is needed which takes into consideration similar historical events and the performance of companies providing similar services during those events
- Comparing AMZN stock with competitors is tricky, as Amazon serves diverse industries, but the stock is represents company as a whole



Sources

- Machine Learning with Python - https://www.geeksforgeeks.org/machine-learning-with-python/
- Stock Price Prediction using Machine Learning in Python - https://www.geeksforgeeks.org/stock-price-prediction-using-machine-learning-in-python/
- Machine Learning to Predict Stock Prices - https://towardsdatascience.com/predicting-stock-prices-using-a-keras-lstm-model-4225457f0233
- Understanding ARIMA Models for Machine Learning - https://www.capitalone.com/tech/machine-learning/understanding-arima-models/
- Creating an ARIMA Model for Time Series Forecasting - https://towardsdatascience.com/creating-an-arima-model-for-time-series-forecasting-ff3b619b848d
- Robust Data - https://www.forbes.com/sites/ellevate/2021/06/08/the-importance-of-using-robust-analysis-to-understand-change/?sh=55ddad943dab https://www.investopedia.com/terms/r/robust.asp
- 2018 SEC Filing Amazon - https://www.sec.gov/Archives/edgar/data/1018724/000101872419000004/amzn-20181231x10k.htm
- 2018 AMZN Year End Price Action Why Amazon Stock Is Up 29% in 2018 | The Motley Fool
- Searchable Stock Halts Trade Halts - Current (nasdaqtrader.com)
- Stock Price Prediction with Machine Learning
- Stock Price Prediction using Machine Learning