Auto-Regressive Integrated Moving Average (ARIMA) is a model that uses statistical data for time series forecasting.
We use a couple of parameters when utilizing this particular model
We can call them, p, d , and q.
P is the number of lag observations, a lag is a fixed period of time used for the model.
D is the number of times that we draw the difference between raw data, this is also called the degree of difference.
Q is the sliding window we use for the moving average.
We typically start off by visualizing data and plotting points to help identify patterns or trends.
Then we must identify is the data is stationary, or stable over time, so that we can obtain accurate results from the ARIMA model. Stationary data is a must in order to use it correctly. There are many tests we can use to help identify stationarity such as the ADF test which is what we chose to use in our model.
We then plot the data for both correlation and auto correlation by analyzing the time series data. these show both the data we had and the data we have against each other. Correlation directly shows current and lagged observation while auto correlation shows the time series against its own lagged values.
Constructing the model becomes as simple as analyzing our data to select appropriate values for the P, D and Q variables we want to use in our model. This is important because it determines the autoregressive and moving average parts based off of what we pick. This means that we need good values when constructing our model or else we run the risk of severely hindering the correctness of our results.
A few of the problems experienced when creating the model we used for our personal NVIDIA stock predictor were things like documentation websites being down and reference websites being very outdated and using deprecated functions and arguments/parameters that no longer existed.
So when trying to construct our model and plot the results some of the functions we needed to find or arguments to functions that we needed were incredibly difficult to learn about because the documentation website was down and we needed to look for obscure stack overflow questions before we began to find answers or explanations that clarified things.