AIRLIFT CASE STUDY BY SHAHZEB NAEEM, NUST

Exploring the data and sharing my findings. Do I see any interesting patterns?

My Approach

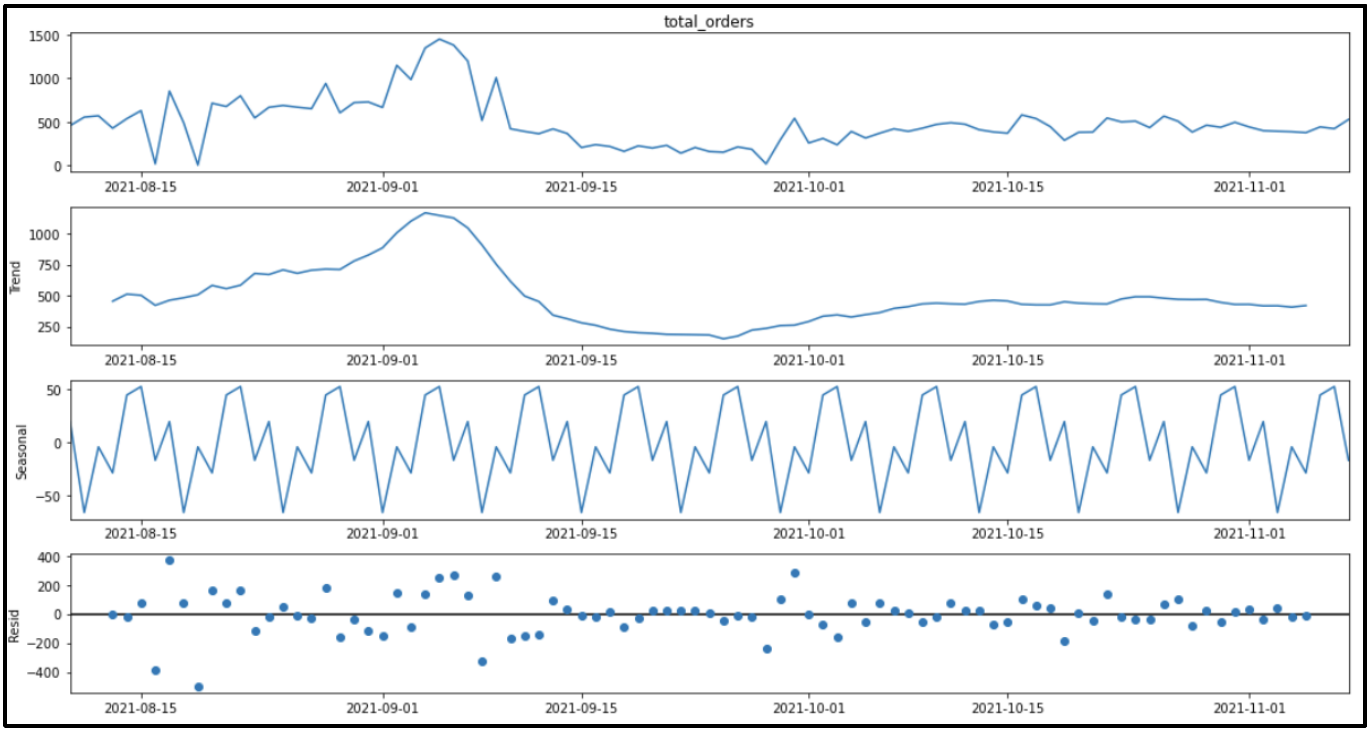
First of all, the data is small in quantity. Plus, it shows the statistics of 2 stores together, with dates that overlap. Henceforth, I wondered if I should treat both stores or together. In fact, I actually started doing so by removing the duplicate dates and summing through excel, only to realize through this question that I must treat them together. Therefore, I simply separated them into two separate excel files, namely, 'Only EW1' and 'Only EW 2'.

The next thing to figure out was how I was going to keep the frequency of the **seasonal_decompose** function to see the patterns of the total orders of the stores more vividly. It was obvious through the scarce amount of data, but since I have to forecast for the daily orders, the frequency should be kept to daily as well.

As for EW1, it had a missing value of total orders on '18/08/2021' and so I utilized functions from pandas to fill it with the mean value. I had the option to fill it with the mode value or the next one, but neither of them turned out to make a difference to the value. I would prefer to use the mode value in such situations since that would be the most repeated total orders in a day.

My Findings

For EW1



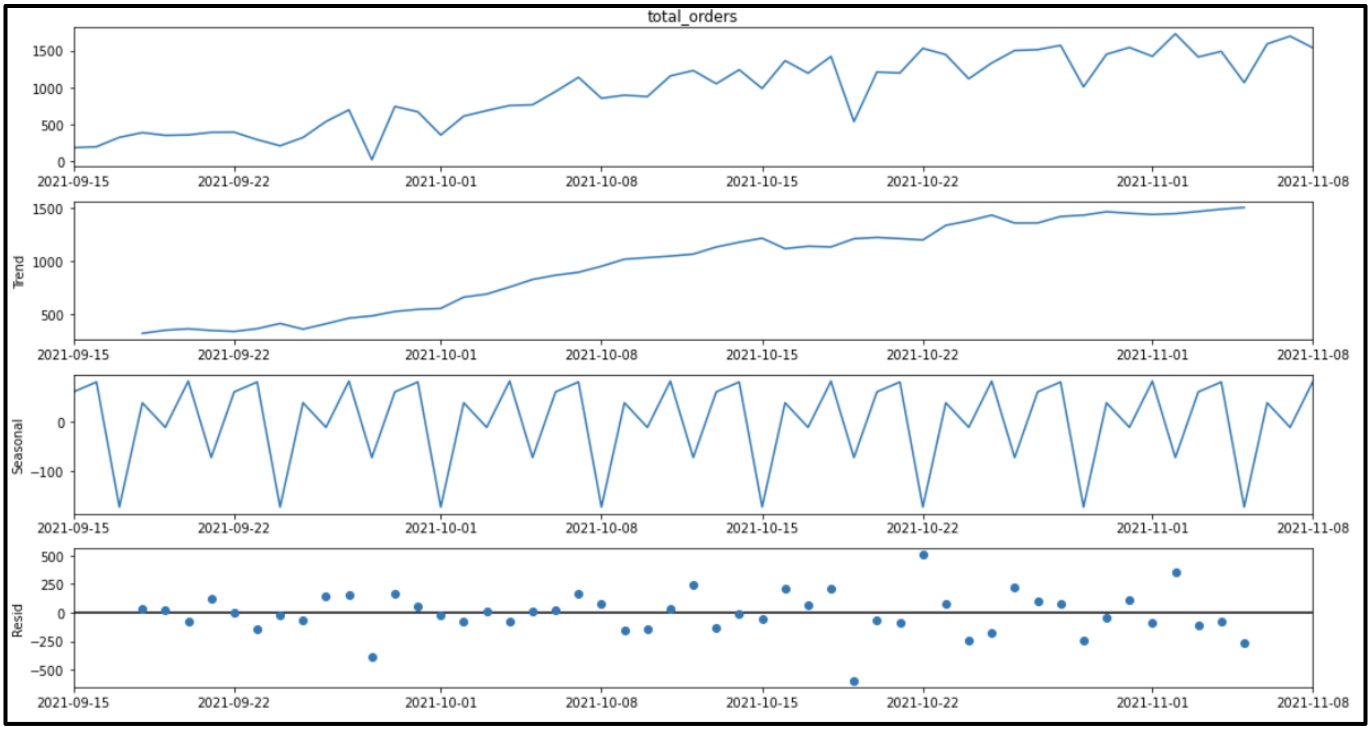
There is not much to take home from here. The *trend* dies down to a rather stagnant and tapered value and stays there. However, the *seasonality* is pretty evident and there are constant drastic dips on Wednesdays with fewer ones on Tuesdays. These are then followed by rises on Fridays and so on. The general trend being almost constant and this repetitive seasonality will definitely be useful in forecasting for the future orders.

The interesting thing here is the sudden rise in total orders for a couple of weeks from 01/09/2021 to about 14/09.2021. Perhaps it could be due to some sort of a shopping season, and something that must be considered when forecasting future orders, even though Artificial Intelligence can be utilized here, but the human element will certainly leave it incomplete and incorrect. Maybe that could be a reason for the mis forecasting as well, the lack of **'Human Intervention'.**

The residual does not really have much to write home about with its variance not being too overly exaggerated in any case.

Now as for EW2, there were no missing days in its data. So, no further functionality was required.

For EW2



Its graphs, however, are certainly much more emphatic. The trend is an increasing one, that too at a pretty quick pace, since the total orders have gone up from around 250 to 1500 in the span of just 50 days or so. Obviously, if the same pattern was being followed by other stores, it would make it difficult for stock to be available everywhere. The seasonality again is clear here, with very dramatic dips on Fridays followed by a dramatic rise on Saturdays and so on.

Now due to the presence of a clear upward trend, and repetitive seasonality, forecasting should not be a problem for this store.

The interesting thing here is the sheer rise in sales suddenly during a much smaller time compared to EW1. Maybe the location of this store is closer to a posher area or has a greater culture of online shopping, but that does not make much sense since the time is too little to make that conclusion. The trend should have been upward since before for that to have been the case. An area cannot just develop within days. So definitely could again be down to some event coming up but even that would be hard to believe since it last for over 30 days and is still rising after that. Looks to point to something that will last permanently, but it is difficult to point at what exactly is causing it. What could certainly be the case is that this data was once all the development in the areas surrounding the store was complete, or that this this store was knew. That could certainly be a legitimate reason for the total orders to rise up so quickly.

The residual again does not really give much away here with even a smaller variance exhibited.

Forecasting daily orders for the next week (9th November, 2021 - 15th November, 2021) for both, EW1 and EW2.

My Approach

The first thing to think of here was that which forecasting technique should I go for. Keeping in mind factors such as longevity and the fact that airlift would certainly have a good amount of data to deal with, **LSTM** (Long Short-Term Memory) made sense. Its error rate is usually lower, and it is more effective, plus it gives more efficient results due to its pattern recognition property lasting over a longer period of time. LSTMs can also perform on non-stationary data which is another plus.

ARIMA (Autoregressive Integrated Moving Average) does give better results for smaller data, and although I was tempted to employ it due to the scarcity of data that I have been provided, I decided it would be better to stick with LSTM due to its practical nature and sense for Airlift. This is because Airlift would certainly have a lot more data, and LSTM outperforms ARIMA in those situations. Not to mention the fact that the parameters are all trained on their own as opposed to the more 'traditional' ML model in ARIMA.

The results that I ended up getting in terms of forecasting for the stated are shown below. I will surely discuss the details related to accuracy of the entire thing in the next question too.

My Findings

For EW1

date	Predictions
2021-11-09	464.073288
2021-11-10	457.851783
2021-11-11	454.965671
2021-11-12	452.266371
2021-11-13	450.658557
2021-11-14	448.063402
2021-11-15	449.353405

For EW2

date	Predictions
2021-11-09	1473.524430
2021-11-10	1482.987226
2021-11-11	1469.844555
2021-11-12	1488.397688
2021-11-13	1431.443779
2021-11-14	1424.106997
2021-11-15	1412.524691

How can you measure forecasting accuracy of a forecasting model?

My Approach

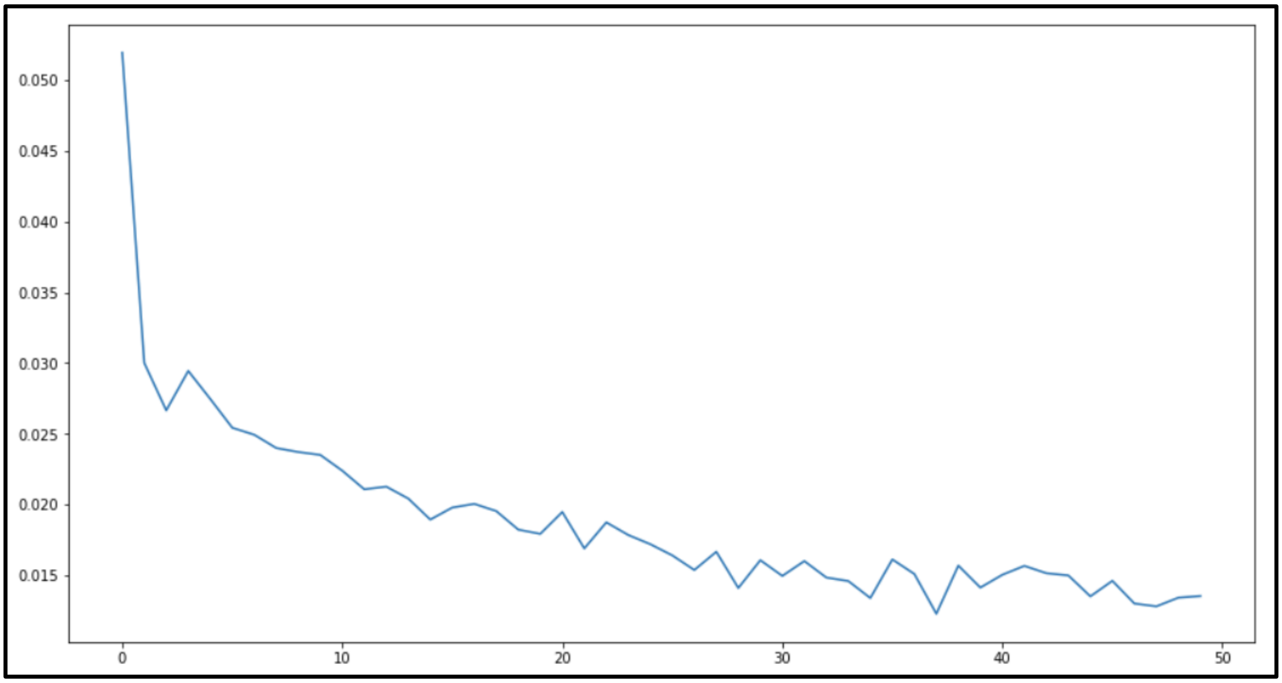
After trying several methods to increase the accuracy of my model, I just came down to the conclusion that I had too less of an amount of data to work with. Albeit the resultant output was not too lacking in its accuracy, over a longer period of time, the inaccuracy would build up so to speak. Then again, I would also have more data to work with.

My Findings

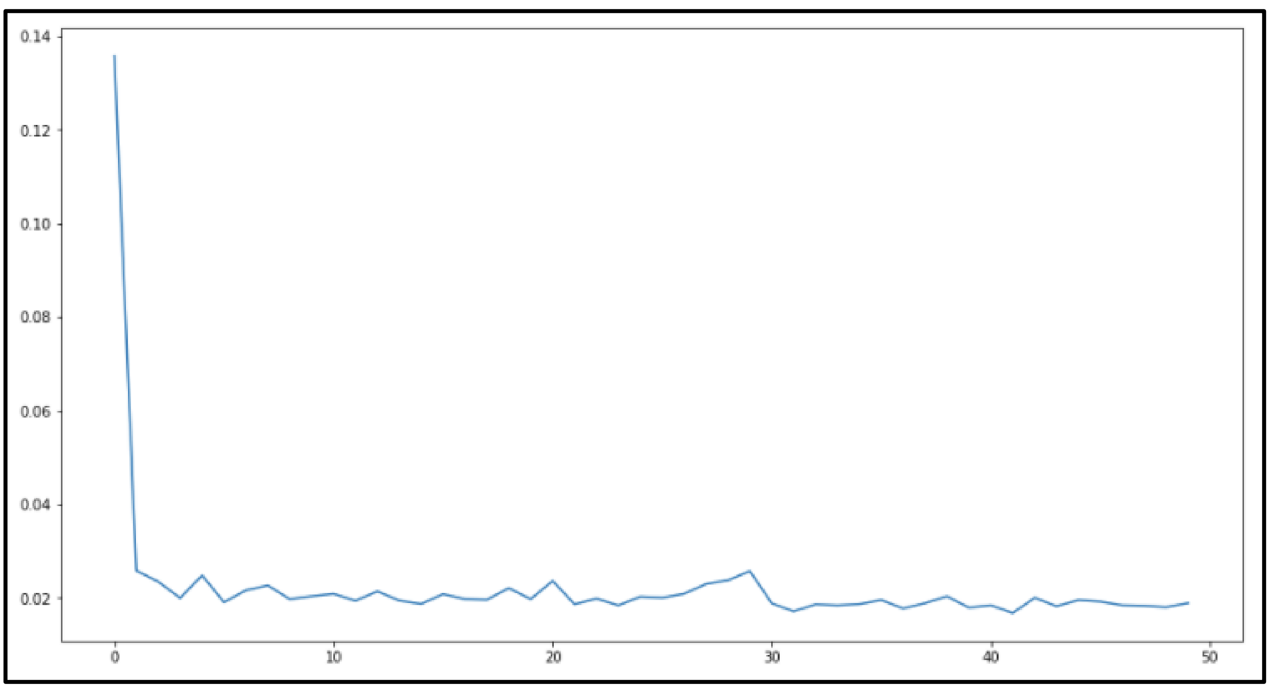
A measure of forecasting accuracy is back testing, or simply put, comparing the output of the model to the output of the original data.

Another method is Mean Squared Error (**MSE**) whilst training, which turned out to be quite small, and actually did quite well at face value if one were to look at it its graph for both EW1 and EW2.

For EW1



For EW2

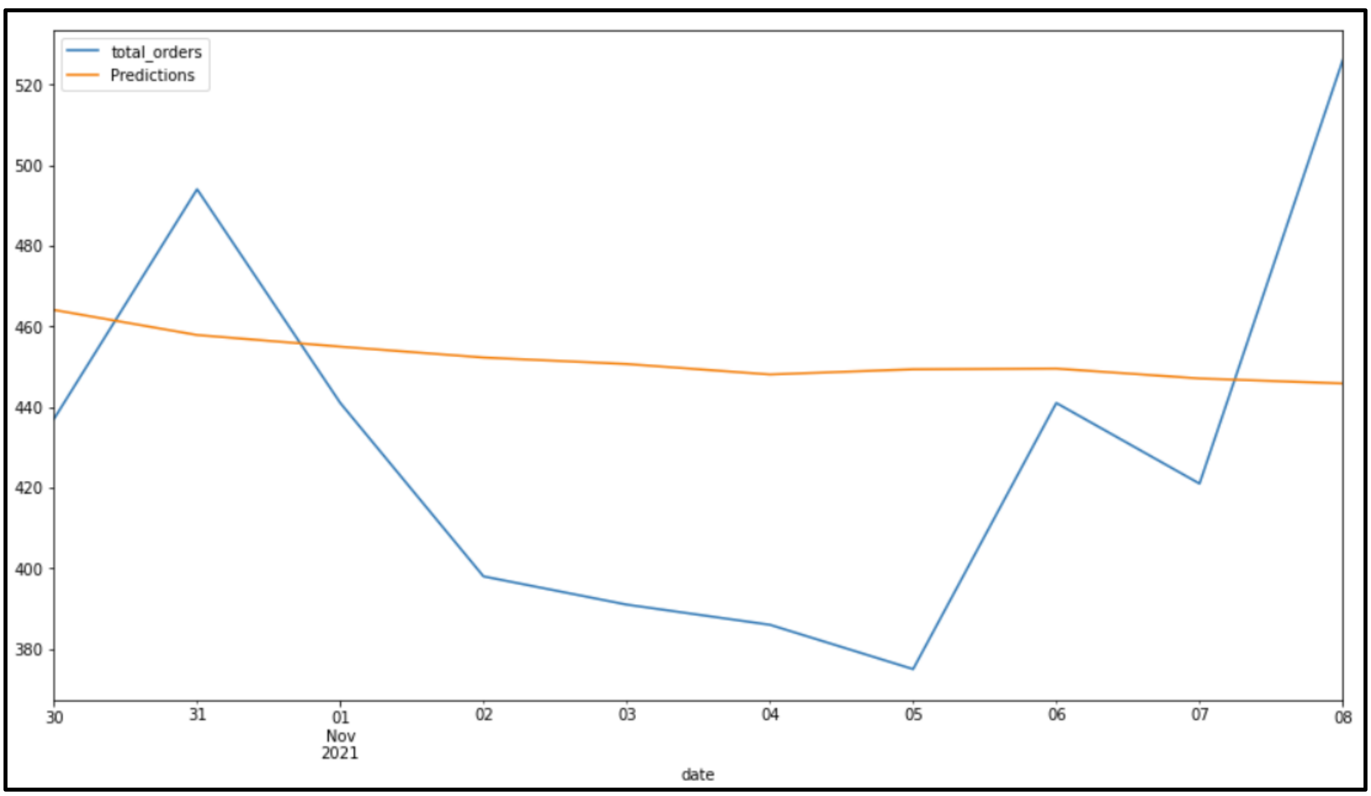


Another method also comes into play which is the **MAE**, Maximum Absolute Error, and that did not do too badly here either. This was again done only on the train data.

Yet another method is the Root Mean Squared Error, which after going through a research paper and many websites turned out be something incorporated pretty consistently with LSTM usage. This was done on the test data. It turned out to be 50.3 after a lot of manipulation and tries for EW1. For EW2, it turned out to be 207.8 which actually makes sense.

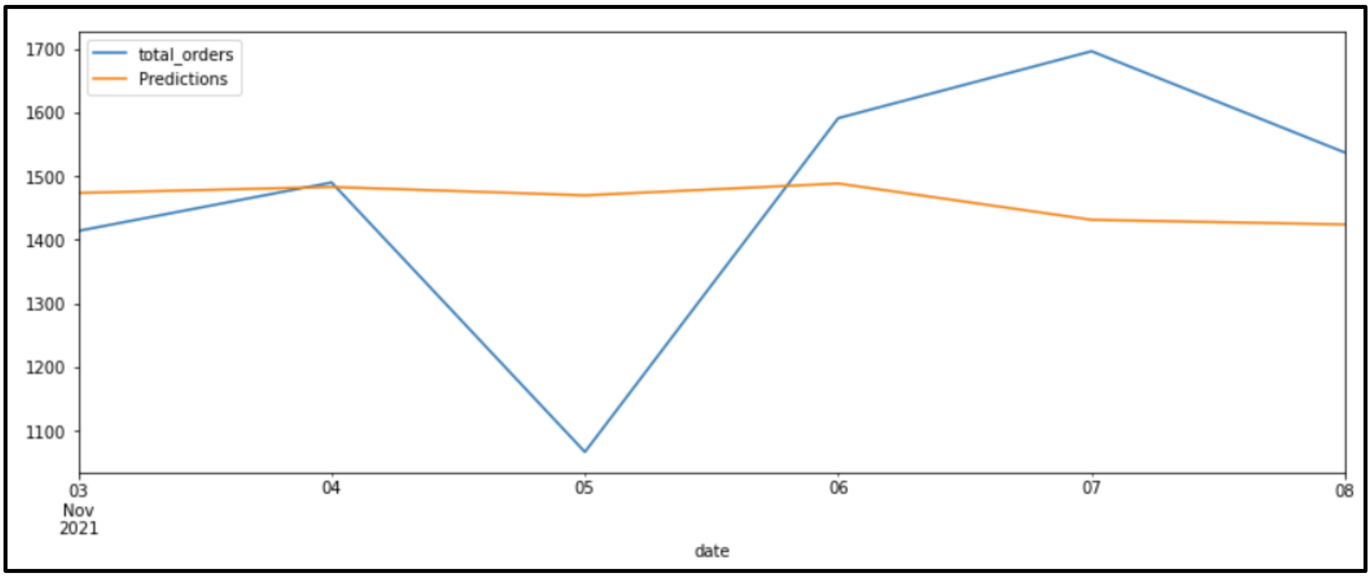
One would think that the MSE graph looks so good for EW2 and better than EW1. Henceforth, how could EW2 have a worse RMSE than EW1? The answer to this is evident from their test results vs their prediction results.

For EW1



We can see here that the difference between the test values and predicted values is small in terms of scale. The reason for this is that the seasonal variation and trend variation was not as much if you were to have a look at those graphs as well for EW1 that is. It is a different story for EW2 though.

For EW2



We can see that although its test data was smaller, yet its RMSE turned out to be greater than EW1. The reason is evident from its trend and seasonality graphs as well, both spanning over a much greater y axis (total orders).

One could still argue that the training set itself was smaller for EW2 since its dataset only had 50 odd values compared to the 90 odd ones of EW1 and that is the cause of this discrepancy and greater error for EW2. We frankly cannot be sure of either since we just have such little data to work with.

I thought of employing GANs and creating more data, but due to not only the lack of time, also the testing status of **GANs** (Generative Adversarial Networks) when it comes to working with anything other than images, I decided not to go ahead with it. Plus, even GANs or **Autoencoding** would have difficulty deciphering the lack of patterns in this data to create a newer one.

Really hope you enjoyed and liked my work. I have tried to do whatever I could with the time and ability that I was given. See you in the next round, hopefully :)

Thankyou for enhancing my learning and opening my mind up to so much more!

Ref:

https://www.researchgate.net/publication/355429916_A_Comparative_Analysis_of_the_ARIMA_and_LSTM_Predictive_Models_and_Their_Effectiveness_for_Predicting_Wind_Speed

https://www.youtube.com/watch?v=S8tpSG6Q2H0

Many others to resolve the coding errors and employ various strategies to optimize the models and so forth.