PONDICHERRY UNIVERSITY

DEPARTMENT OF STATISTICS

COMPARISON OF CLASSICAL APPROACH AND DEEP LEARNING APPROACH FOR FINANCIAL STUDIES USING SPSS AND PYTHON.

Presented by

Susanta Dhurua21375062

Introduction: The world's economic is depend on the export and import, in the ancient days we have the policy of give and take. Now a days the policy is same only the medium is changed. Now the country's economic condition is depend on the Share Market. In the world there is two hub of making money as the capacity of digest, one is at BSE (Bombay Stock Exchange) and another is NSE (New York Stock Exchange).

Share Market from the name itself we could having an idea that share of a particular company in the open market where any persons from different region or country can have the privileged to buy the shares.

Shares are available in the form of stocks in the market means for ex if we buy 100 stocks then we have 0.055% of shares on that particular company. Stocks are traded all the time 24*7. In India share market opens from 9:30 am and closes at 3:30. The people having a demmat account and required some necessary documents, they can easily traded in the given time.

Bombay Stock Exchange (BSE) it is situated in the Bombay, India. The place where the BSE is there it is famously known as "Dalal Street" the financial heart of India. There is two things in the share market of everyday life one is "Intraday" and "delivery". The most interesting things about stocks that it will change without any notification.

Intraday trading: it is most interesting and risky market, the simple fundamentals rule of intraday trading is "Buy Today Sell Today" if we sell one stocks while the rate is very high then the persons will get maximum profit or there is another situation is that if anyone buy a stocks while the rate is going down then the person will make profit.

Delivery trading: "buy a stocks today and sell whenever the trader wants". It had a profit of buying a stocks at a lower price and sell at a higher price. In the delivery trading we could buy the stock according to our clear account balance.

There is huge difference in the intraday and delivery trading, the main difference between them is of risk and margins.

There is one of the most famous proverb that get echo in the dalal street is that "Markets are open by immature and closed by professional". This proverb come into the picture when the professional have do the analysis of upcoming days while immature looking into the price of stocks they buy and ignoring the market situations and positions of that particular company.

Abstract:

Adani green energy is one of the most renewable energy sources of the adani groups which is based on India. Adani Enterprises limited had an ownership of 75% in adani green energy.

Adani green energy operates two main business segments solar energy and wind energy. It also has a small presence in other renewable energy technologies such as biomass and small hydro. Adani green energy have successfully completed the project of wind and solar energy in Uttar Pradesh and Tamil Nadu.

Considering the two financial year 2021(from 12th April) -2022 and 2022-2023 (up to 12th April) of adani green energy we are predicting the upcoming day wise closing stocks price, weekly closing stocks price, monthly closing stock price.

Here are some key details about Adani Green Energy:

Ownership: AGEL is majority-owned by Adani Enterprises Limited, which holds a 74.92% stake in the company. The remaining shares are held by public investors.

Business segments: AGEL operates in two main business segments: Solar Energy and Wind Energy. It also has a small presence in other renewable energy technologies such as biomass and small hydro.

Capacity: As of September 2021, AGEL has an installed renewable energy capacity of over 25 GW, making it one of the largest renewable energy companies in the world. Its capacity includes both operational and under-construction projects.

Financials: AGEL reported a consolidated revenue of INR 2,867 crore (US$384 million) in Q1 FY22, up 28% from the same period in the previous year. Its net profit for the same period was INR 219 crore (US$29 million).

Relevant stock variables

Open: The open price in stock refers to the price at which a particular stock begins trading at the beginning of a trading session, such as the opening of the stock market for the day.

Close: The closing price in stock refers to the final price at which a particular stock trades on a particular trading day. It is the last price at which a trade was executed before the market closes for the day.

Adjusted close: The Adjusted Close price in stock refers to the final price at which a particular stock trades on a specific trading day, but with adjustments made for any corporate actions that may have affected the stock's price, such as stock splits, dividends, or mergers.

Volume: Volume in stock refers to the total number of shares of a particular stock that have been traded during a specific period of time, typically a trading day.

Importance of Closing Price.

The **closing price** in **stock** is an important metric that provides **crucial information** for **investors** and **traders** to make decisions about **buying** or **selling** a stock. It helps in price discovery, daily performance tracking, technical analysis, and understanding market sentiment.

Here are some of the reasons **=>**

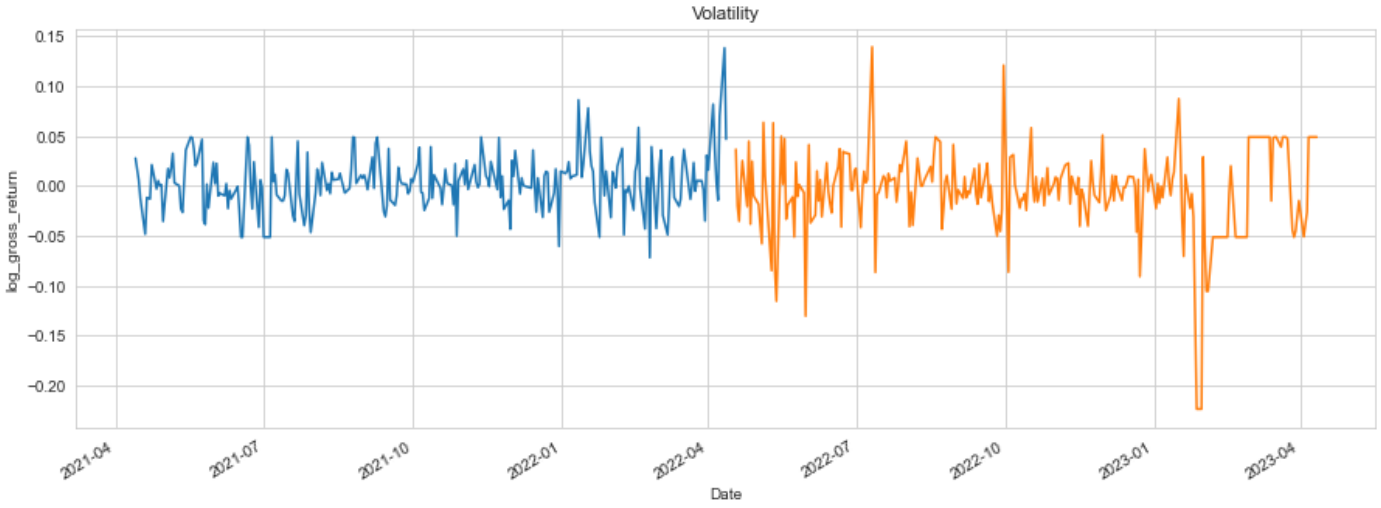
- **Price Discovery**: It helps investors and traders to know the fair value of the stock at the end of the trading day, and to use this information to make informed decisions about buying or selling the stock.
- **Daily Performance**: The closing price is also used to calculate the daily performance of a stock, which is an important metric for investors to track the gains or losses made on their investments.
- **Technical Analysis**: The closing price is an important input for technical analysts who use historical price charts to identify trends and patterns in the stock market.
- **Market Sentiment**: A higher closing price indicates bullish sentiment, while a lower closing price indicates bearish sentiment.

Analysis

Volatility

Considering two different time periods:- **1) 12-04-2021 to 12-04-2022 and 2) 12-04-2022 to 12-04-2023**

to analyse the volatility of the market based on the market log-gross-return. The first time period (1) is taken to see the market fluctuation after Covid-19 pandemic and the second time period (2) is taken to see the effect of market share price after coming out of the report of Hindenburg regarding Adani scam. The volatility report is given below-



The overall picture is telling that market was less volatile at recovery period of the covid-19 pandemic. But the Hindenburg's report had made a huge change in the market and due to that market had become highly volatile at that period after 01/2023. Even after 01/2023, market returns had fallen down rapidly.

Classical modelling

ARIMA (Autoregressive Integrated Moving Average) is a statistical model used to analyze time series data. It combines three elements: auto regression (AR), differencing (I), and moving average (MA). Auto regression refers to the idea that a variable is regressed on its own previous values. Differencing is the process of taking the difference between consecutive observations to remove any trends or seasonality in the data. Finally, moving average refers to using the average of past errors to make predictions for the future.

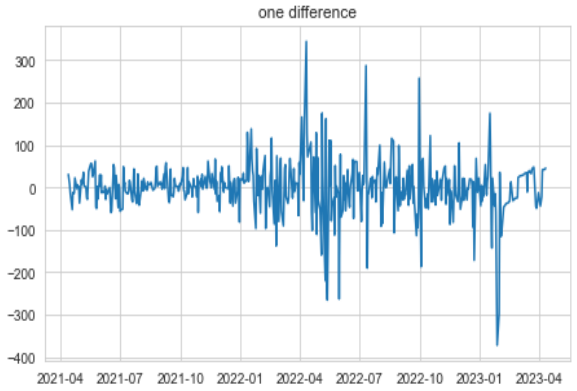
The ARIMA model can be used to forecast future values of a time series based on its past behavior. It is widely used in finance, economics, and other fields where data is collected over time.

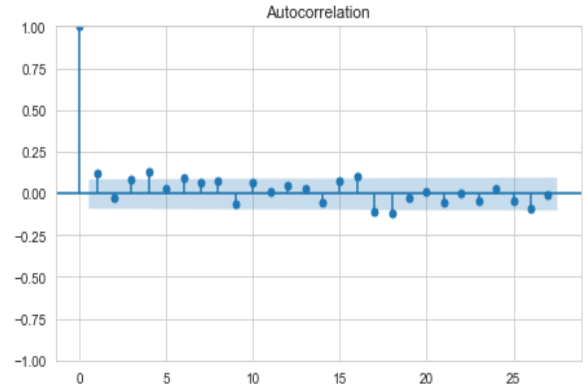
- p determines the number of autoregressive (AR) terms
- d determines the order of differencing
- q determines the number of moving average (MA) terms

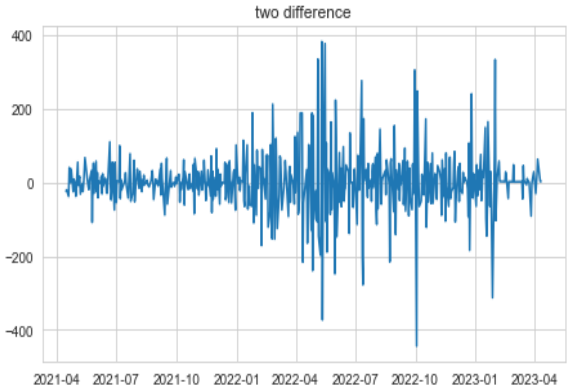
A stationary time series is easier to analyze and model, and its statistical properties can be estimated more accurately. This is because stationary data does not exhibit random or unpredictable changes over time, which makes it possible to use past observations to make reliable predictions about future values.
Stationary in time series refers to a property where the statistical properties of a time series (such as mean, variance, and autocorrelation) remain constant over time. In other words, the data points in a stationary time series exhibit consistent patterns and behavior throughout the entire series, without any significant trend, seasonality, or other time-dependent effects.

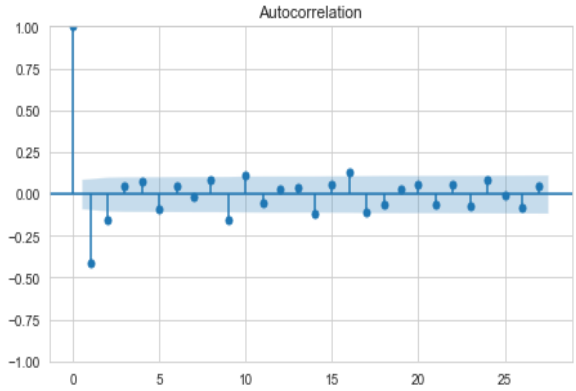
Test of Stationary:
Adfuller test statistic: -1.6771510741751743
P-value: 0.44292170672527176

Estimation of “d” in ARIMA



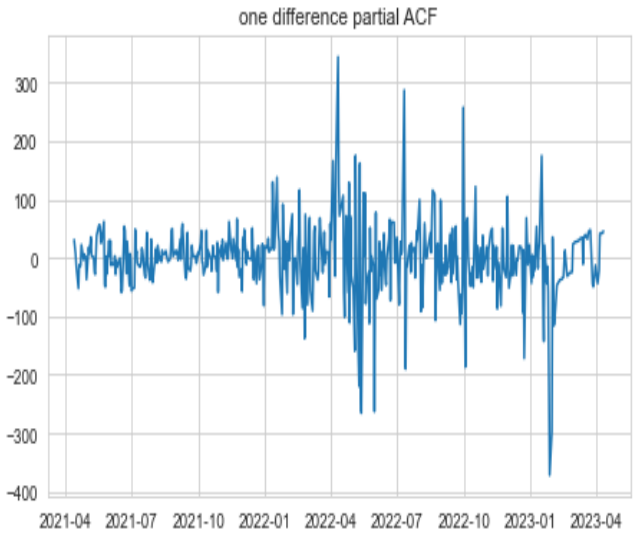


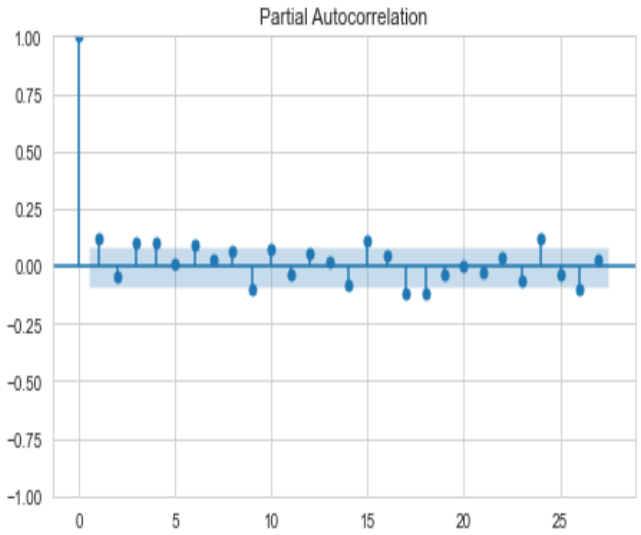




The estimated differencing value is coming out to be 1 based on the auto-correlation plot.

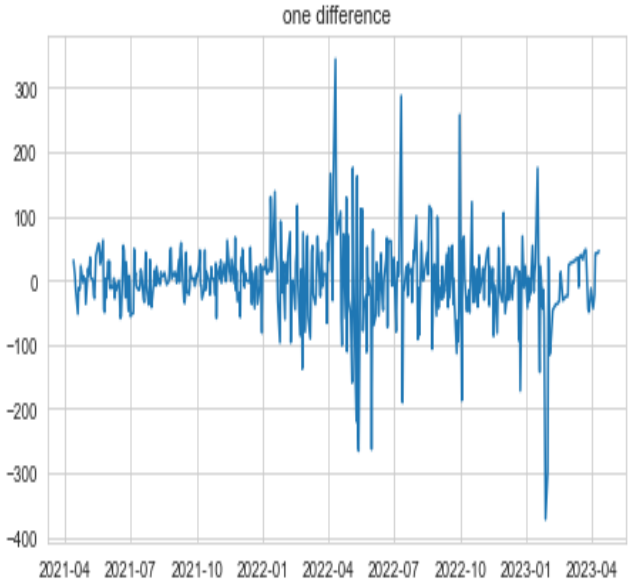
Estimation of “p” in ARIMA taking d=1

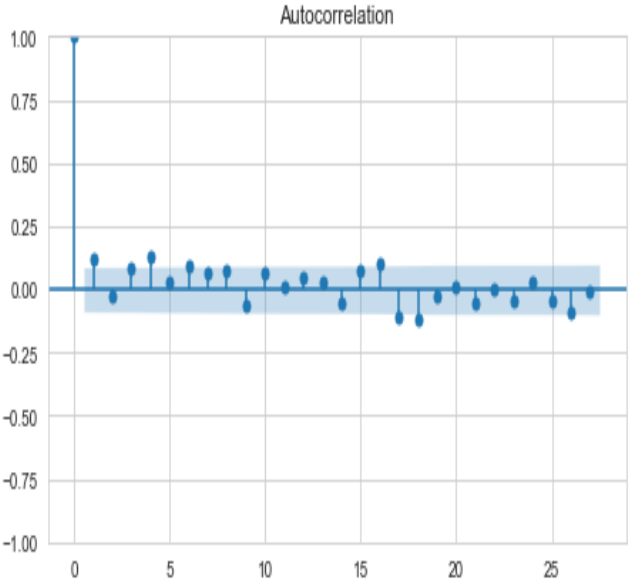




From partial auto-correlation plot, the estimated value of p is coming out to be 3.

Estimation of “q” in ARIMA





The estimated value of q is coming out to be 4 based on the auto-correlation plot.

Now, considering the whole data of closing price as training set for ARIMA (3, 1, 4)

Out[189]:

SARIMAX Results

Dep. Variable:	Close	**No. Observations:**	496
Model:	ARIMA(3, 1, 4)	**Log Likelihood**	-2728.211
Date:	Thu, 13 Apr 2023	**AIC**	5472.421
Time:	00:09:25	**BIC**	5506.058
Sample:	0 - 496	**HQIC**	5485.626

Covariance Type: opg

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.7386	0.086	-8.540	0.000	-0.908	-0.569
ar.L2	0.6229	0.143	4.357	0.000	0.343	0.903
ar.L3	0.6929	0.087	7.924	0.000	0.522	0.864
ma.L1	0.8917	0.090	9.898	0.000	0.715	1.068
ma.L2	-0.5506	0.162	-3.403	0.001	-0.868	-0.233
ma.L3	-0.7133	0.122	-5.826	0.000	-0.953	-0.473
ma.L4	0.0703	0.045	1.569	0.117	-0.018	0.158
sigma2	3574.4049	126.247	28.313	0.000	3326.966	3821.844

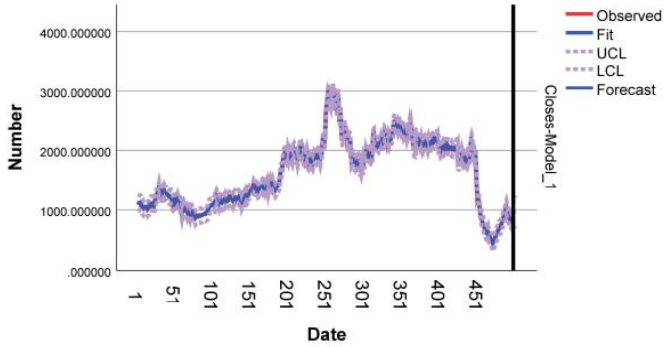
Ljung-Box (L1) (Q):	0.03	**Jarque-Bera (JB):**	978.21
Prob(Q):	0.87	**Prob(JB):**	0.00
Heteroskedasticity (H):	5.23	**Skew:**	-0.32
Prob(H) (two-sided):	0.00	**Kurtosis:**	9.86

Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	Model Fit statistics: R-squared	Model Fit statistics: RMSE	Ljung-Box Q(18): Statistics	Ljung-Box Q(18): DF
Closes-Model_1	0	.080	.989	60.326	30.211	11

Forecast

Model		497	498	499	500
Closes-Model_1	Forecast	959.502839	953.582942	963.722571	966.649299
	UCL	1078.005734	1134.771324	1188.075766	1233.614894
	LCL	840.999944	772.394559	739.369376	699.683704



RMSE= 60.326

Considering 80% of the observation in training set.

Out[172]:

SARIMAX Results

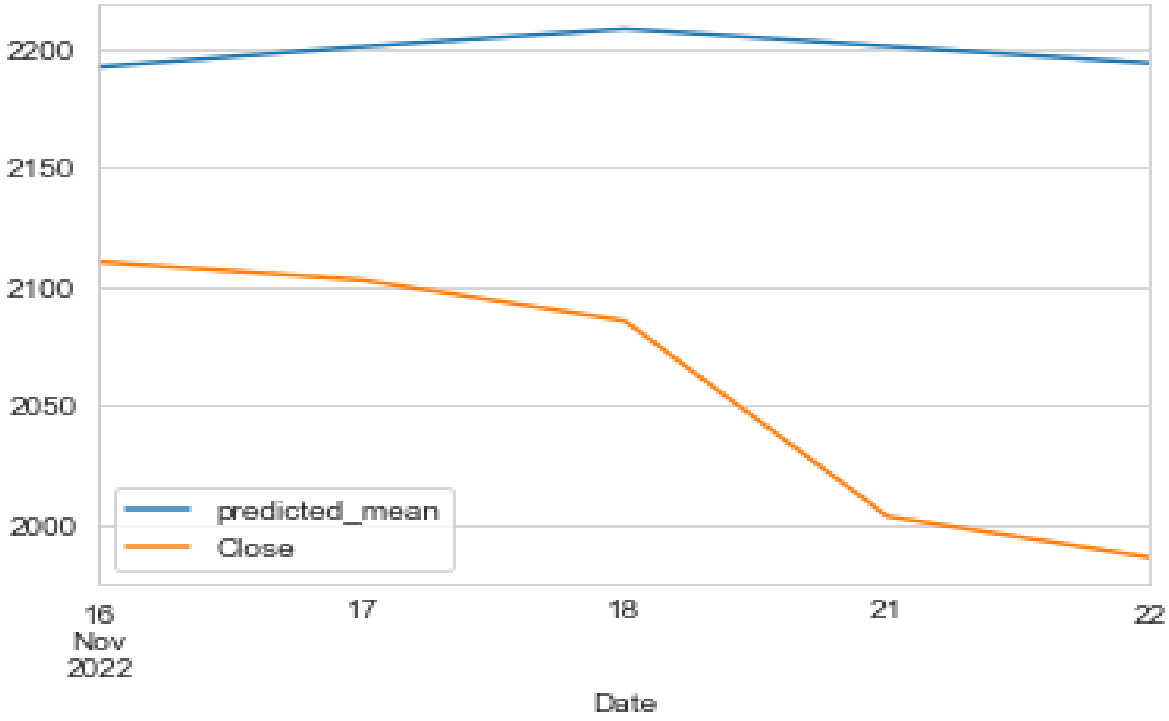
Dep. Variable:	Close	**No. Observations:**	396
Model:	ARIMA(3, 1, 3)	**Log Likelihood**	-2174.435
Date:	Wed, 12 Apr 2023	**AIC**	4362.871
Time:	23:54:04	**BIC**	4390.723
Sample:	0 - 396	**HQIC**	4373.906

Covariance Type: opg

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.8254	0.126	6.540	0.000	0.578	1.073
ar.L2	-0.9290	0.039	-23.923	0.000	-1.005	-0.853
ar.L3	0.8373	0.122	6.873	0.000	0.599	1.076
ma.L1	-0.8107	0.139	-5.853	0.000	-1.082	-0.539
ma.L2	0.9155	0.050	18.358	0.000	0.818	1.013
ma.L3	-0.7571	0.131	-5.798	0.000	-1.013	-0.501
sigma2	3704.6455	162.741	22.764	0.000	3385.678	4023.613

Ljung-Box (L1) (Q):	0.09	**Jarque-Bera (JB):**	547.03
Prob(Q):	0.76	**Prob(JB):**	0.00
Heteroskedasticity (H):	7.53	**Skew:**	0.41
Prob(H) (two-sided):	0.00	**Kurtosis:**	8.71

The RMSE value = 154.38771800357043



Considering weekly close price using Expert Modeler

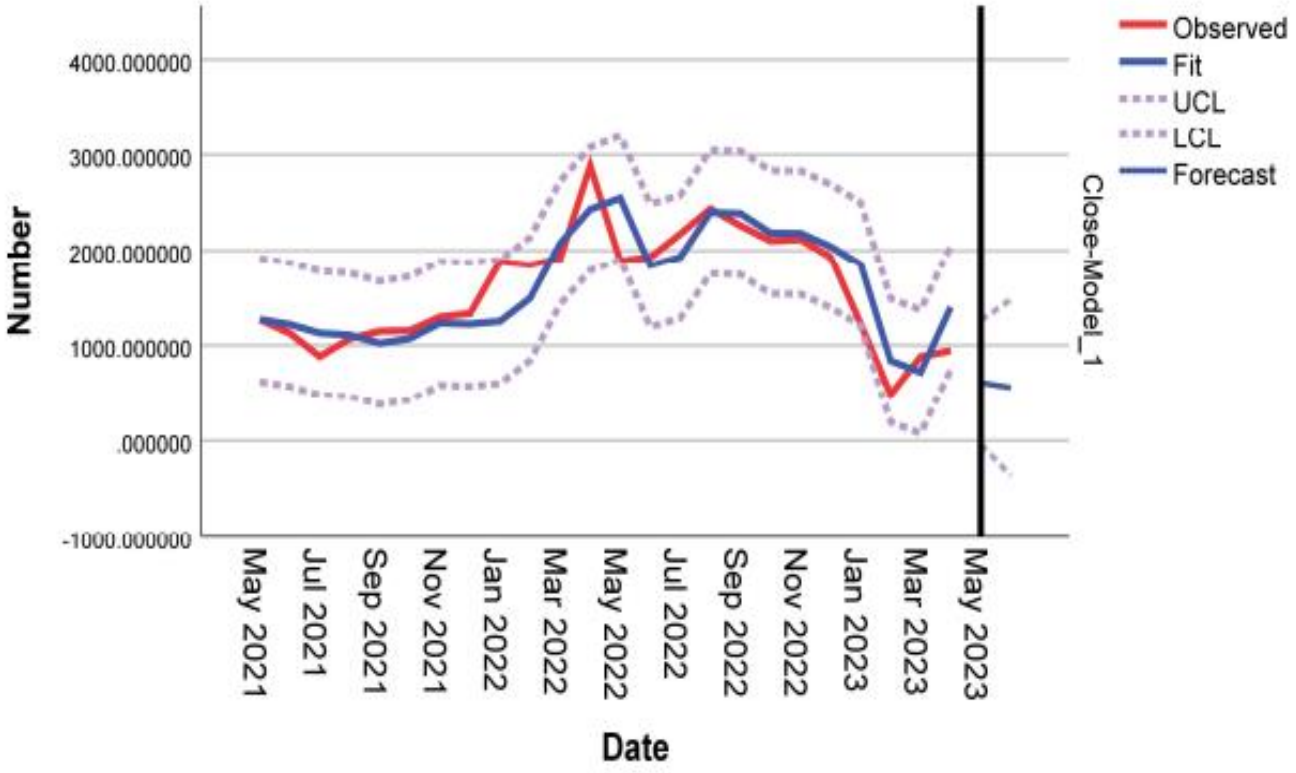
Model Statistics

Model	Number of Predictors	Model Fit statistics			Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	R-squared	RMSE	Statistics	DF	Sig.	
Close-Model_1	0	.105	.972	120.676	20.698	17	.240	0

Forecast

Model		172	173	174	175
Adj_Close-Model_1	Forecast	987.091942	1008.378054	1022.393626	1034.023004
	UCL	1184.106739	1359.096711	1509.709076	1644.962151
	LCL	815.550120	729.837619	662.937861	610.253655

Considering monthly close price using Expert Modeler



Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	Model Fit statistics: R-squared	Model Fit statistics: RMSE	Model Fit statistics: Normalized BIC	Ljung-Box . Statistics
Close-Model_1	0	.697	.738	312.576	11.755	15.639

Forecast

Model		May 2023	Jun 2023
Close-Model_1	Forecast	609.676682	558.352009
	UCL	1257.920038	1474.648328
	LCL	-38.566674	-357.944311

Model	Ljung-Box Q(18) DF	Ljung-Box Q(18) Sig.	Number of Outliers
Close-Model_1	16	.478	0

Deep Learning Approach (LSTM) using day-wise closing price

```
model = keras.Sequential()
model.add(layers.LSTM(100, return_sequences=True, input_shape=(x_train.shape[1], 1)))
model.add(layers.LSTM(100, return_sequences=False))
model.add(layers.Dense(25))
model.add(layers.Dense(1))
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 60, 100)           40800

 lstm_1 (LSTM)               (None, 100)               80400

 dense (Dense)               (None, 25)                2525

 dense_1 (Dense)             (None, 1)                 26

=================================================================
Total params: 123,751
Trainable params: 123,751
Non-trainable params: 0
_________________________________________________________________
```

The RMSE value = 32.395354288667946

Models

Time	Model	RMSE
Daily (considering whole data)	ARIMA (3, 1, 4)	60.326
Daily (considering 80% data)	ARIMA (3, 1, 3)	154.387
Daily (considering 95% data)	ARIMA (0, 2, 2)	148.055
Weekly(considering whole data)	ARIMA (1, 1, 0)	120.676
Monthly (whole data)	Simple Seasonal	312.576
LSTM (for day wise data)	LSTM	32.395

Conclusions:

Comparison of Classical and Deep learning approach through this analysis

- **LSTM (Long Short-Term Memory) and ARIMA (Autoregressive Integrated Moving Average) are both popular models used in time series forecasting.**
- ARIMA is a statistical model that can be used to predict future values of a time series based on its past values. It models the data as a linear combination of its past values and the past errors. ARIMA models are well-suited for modeling stationary time series data with linear dependencies.
- LSTM, on the other hand, is a type of recurrent neural network that can learn long-term dependencies in time series data. Unlike ARIMA, LSTM can handle non-linear dependencies and can capture complex patterns in the data. **LSTM models are particularly useful when the data has non-linear trends or seasonality.**
- One advantage of ARIMA is that it is easy to interpret and understand the underlying parameters. It also works well with relatively small amounts of data. However**, it requires that the data be stationary, and it may not perform well when the data has complex patterns.**
- **LSTM typically requires a large amount of data for training, particularly for capturing long-term dependencies**.
- **LSTM models have shown to perform well on complex time series data and can outperform ARIMA in certain cases.**

References:

1. https://finance.yahoo.com/
2. https://365datascience.com/
3. https://economictimes.indiatimes.com/markets