Topic: CA Two

- Statistical Analysis of a Dataset: Time Series

Module: B9FT106 Applied Financial Analysis

Faculty: Lynn Monaghan

Student Id:

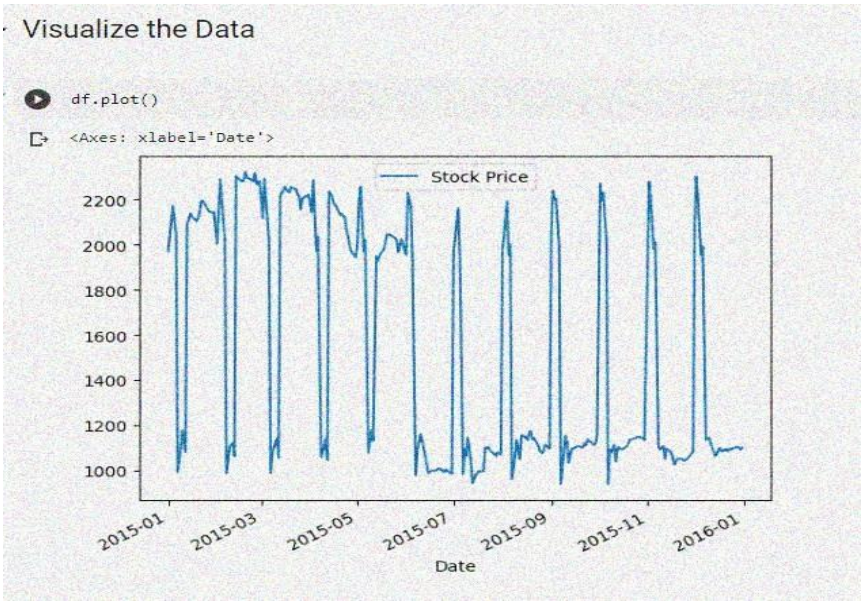
- Siddhi Kelshikar - 10627249

OVERVIEW

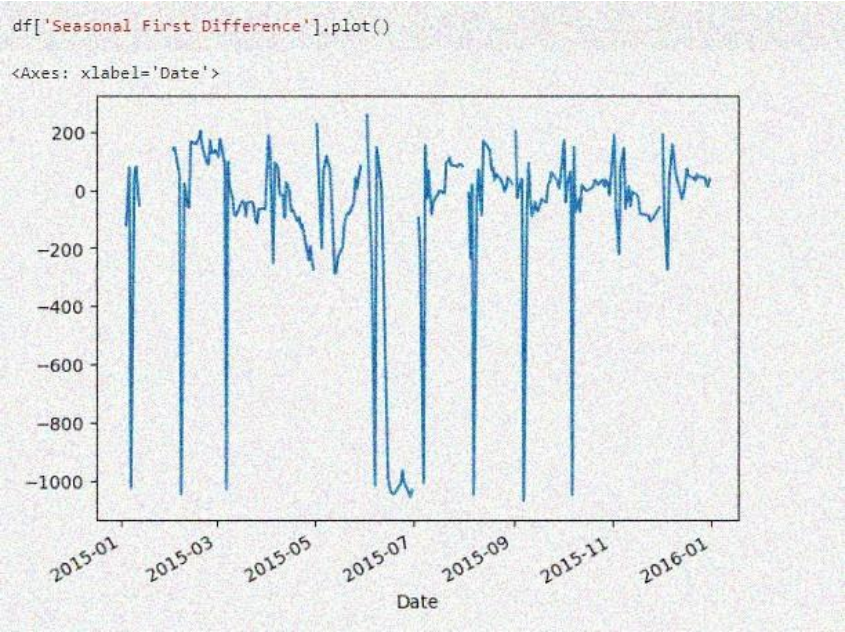
Time series model analysis of stock prices is a well-liked method for forecasting the behaviour of financial assets. In this instance, we used various methods, including data visualisation, data analytics, and machine learning, to analyse the stock price of Infosys, a multinational IT services and consulting firm.

The data, which spans the period from January 1 to December 31, 2015, was obtained from the National Stock Exchange of India Ltd. (NSE), the country's largest stock exchange. We took into account the Volume-Weighted Average Price (VWAP) of the Infosys stock in the dataset, which comprised two columns for date and stock price.

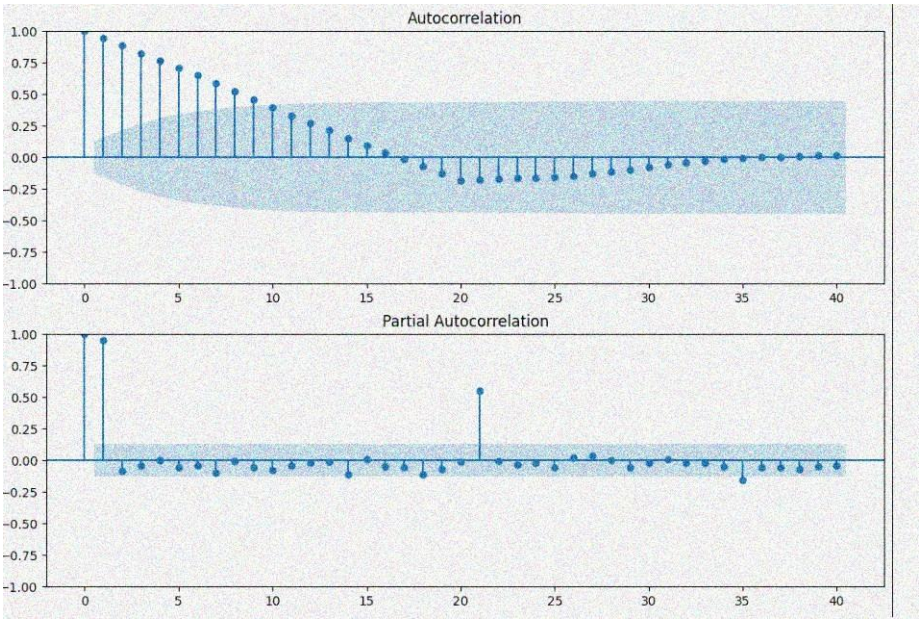
The dataset and all necessary libraries were imported into a Python notebook before we used the info() function to determine the central and variational measures. In order to retrieve the count, mean, standard deviation, min, and max for various parameters, we also transformed the monthly data frame into datetime and utilised the describe() method. We discovered through data analysis that the stock price is seasonal, with monthly variations.



The adfuller test from the statsmodels library was then used to determine whether or not our dataset was stationary. If the data is non-stationary, we use different methods, like differencing, to make it stationary.



The p-value was decreased from 0.759 to 0.109 by applying differencing to determine the Stock Price First Difference, Seasonal First Difference, and Seasonal Second Difference. This shows that the data are stationary, allowing us to continue with the research. We utilised the statsmodels library's plot_acf and plot_pacf functions to further analyse the data. The autocorrelation function (ACF) of a time series, which depicts the relationship between a time series and its lagged values, is plotted using the plot_acf function.



The partial autocorrelation function (PACF) of a time series, which illustrates the correlation between a time series and its lagged values while adjusting for the impact of intermediate lags, is plotted using the plot_pacf function.

ARIMA MODEL

To predict the price of Infosys' shares, we employed the ARIMA model. The model summary's AIC and BIC were 2786.73 and 2797.26, respectively. In order to predict the price of Infosys' stock, we utilised the seasonal order=(1,1,1,20) to represent p, d, q, and shift_value, respectively. In order to forecast the future, we additionally used Date Offset from the pandas.tseries.offsets library.

SARIMAX Results

Dep. Variable:	Stock Price	**No. Observations:**	248
Model:	ARIMA(1, 1, 1)	**Log Likelihood**	-1390.367
Date:	Tue, 02 May 2023	**AIC**	2786.734
Time:	17:40:17	**BIC**	2797.262
Sample:	0	**HQIC**	2790.972
	- 248		
Covariance Type:	opg		

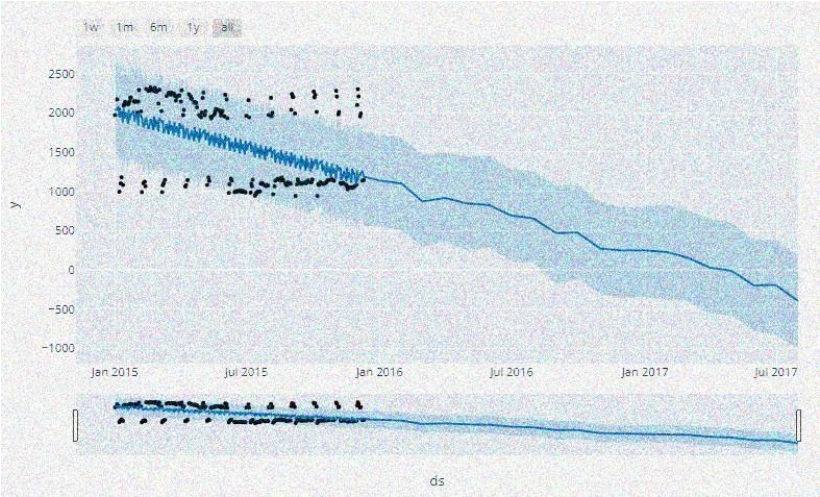
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.1063	21.971	-0.005	0.996	-43.168	42.955
ma.L1	0.1365	21.982	0.006	0.995	-42.948	43.221
sigma2	4566.9149	60.856	75.045	0.000	4447.639	4686.191

Ljung-Box (L1) (Q):	0.00	**Jarque-Bera (JB):**	388451.29
Prob(Q):	0.97	**Prob(JB):**	0.00
Heteroskedasticity (H):	0.24	**Skew:**	-13.17
Prob(H) (two-sided):	0.00	**Kurtosis:**	195.49

The Mean Absolute Error (MAE) is 31.76. This means that the model is moderately accurate in predicting the stock prices as a trend. However, when we look at the Mean Absolute Percentage Error (MAPE) 0f 0.029 or 2.9%, it can be said that the model is pretty accurate in estimating the future trend of stock prices of Infosys. The MAE measures the absolute error in the forecast, while the MAPE measures the relative error in the forecast as a percentage of the actual value.

PROPHET MODEL

We employed the Prophet model in addition to the ARIMA model to predict the price of Infosys' stock. We used the 'ds', 'yhat', 'yhat_lower', and 'yhat_upper' to forecast prices after importing Prophet from the prophet library.



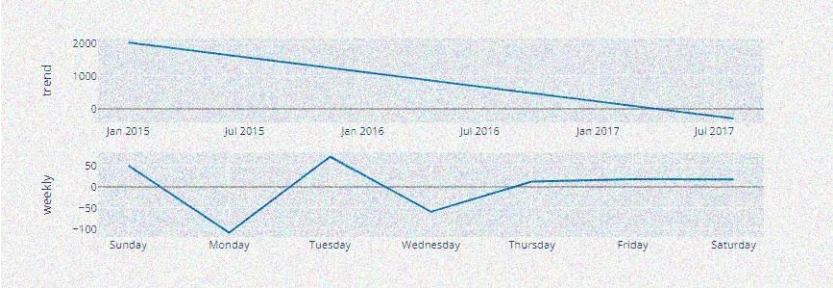
Additionally, we used Prophet's plot_plotly function to create a plotly graph and interpret model parameters like mean_absolute_error, mean_absolute_percentage_error, and mean_squared_error.

We have also plot a plotly graph **from prophet.plot import plot_plotly** and further interpreted the model parameters such as:

mean_absolute_error = 671.1825463472501;

mean_absolute_percentage_error = 0.6155502383143662;

mean_squared_error = 802.7390094074282



From the results of the above parameters, it can be clearly said that the model is not fitted, and is also not an accurate measure to predict the trend of stock prices for Infosys.

CONCLUSION

In conclusion, using time series models to the analysis of stock prices is an efficient technique to foresee the behaviour of financial assets. This is because these models take into account historical data. Utilising a variety of approaches, such as data analytics, machine learning, and data visualisation, enables us to perform a more thorough analysis of the data and locate the information that is most relevant for further examination.

Using the ARIMA and Prophet models, we were able to forecast the stock price of Infosys, and the results of our forecast were explained by looking at the model's input parameters. Because of the analysis, which provided illuminating details about the Infosys stock price, it is now possible to make astute decisions regarding one's investment portfolio.