
DATA SCIENCE

Reliance Stock Price

Time Series Forecasting

CONTENTS:

- INTRODUCTION
- EDA
- VISUALIZATION
- MODEL BUILDING AND FORECASTING
- COMPARISON TABLES
- DEPLOYMENT RESULTS

INTRODUCTION:

FORECASTING PROJECTS ON PREDICTING THE STOCK PRICES OF RELAINCE INDUSTRIES :

Reliance Industries Limited is a Fortune 500® company and the largest private sector corporation in India. It is an Indian conglomerate holding company headquartered in Mumbai, Maharashtra, India.
Reliance owns businesses across India engaged

- ENERGY
- PETROCHEMICALS
- RETAIL
- DIGITAL SERVICES - JIO
- NEW ENERGY
- MEDIA & ENTERTAINMENT

The number of shares of RIL are approx.3.1 billion. The promoter group, Ambani family, holds approx. 46.32% of the total shares whereas the remaining 53.68% shares are held by public shareholders, including FII and corporate bodies. Life Insurance Corporation of India is the largest non-promoter investor in the company, with 7.98% shareholding.

FEATURES :

DATE : STOCK PRICE ON THE SPECIFIC DATE STARTING FROM 1996

OPEN : THE OPENING PRICE OF THE STOCK FOR THE GIVEN DATE

HIGH : THE HIGHEST PRICE OF THE STOCK FOR THE GIVEN DATE

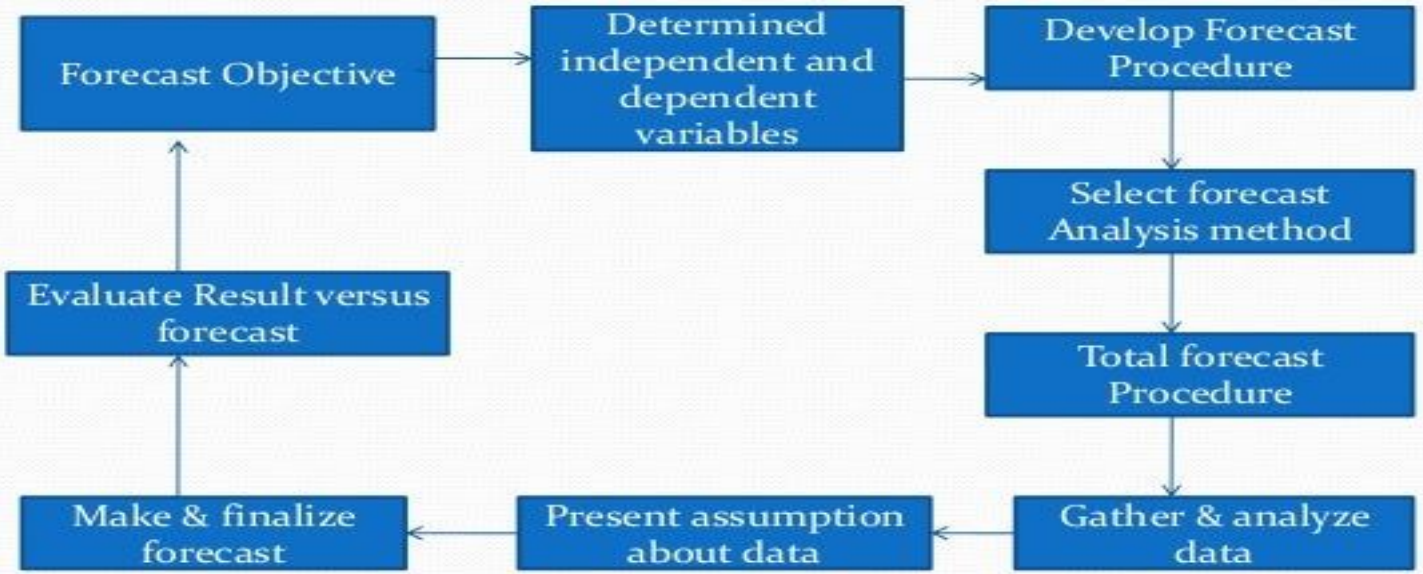
LOW :THE LOWEST PRICE OF THE STOCK FOR THE GIVEN DATE

CLOSE : THE CLOSING PRICE OF THE STOCK FOR THE GIVEN DATE

ADJ CLOSE :ADJUSTED CLOSE PRICE ADJUSTED FOR SPLITS AND DIVIDEND AND/OR CAPITAL GAIN DISTRIBUTIONS.

VOLUME : THE VOLUME OF THE STOCK ON THE GIVEN DATE

Forecasting Process



Exploratory data analysis:

- Importing of all important libraries.
- Installation of python packages.
- Importing of reliance csv data into the program.
- Understanding data types of the columns.
- Checking the shape of the data.
- Checking missing values.
- Filling missing values.
- Converting the date column into date time format and setting it as index
  - For time series forecasting.
- Checking the statistical values of the data through describe() function.

EDA (Exploratory Data Analysis):

Descriptive Statistics:

The shape of the dataset is (7045,7).

Dataset Information:

```
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Date       7035 non-null   object
 1   Open       7035 non-null   float64
 2   High       7035 non-null   float64
 3   Low        7035 non-null   float64
 4   Close      7035 non-null   float64
 5   Adj Close  7035 non-null   float64
 6   Volume     7035 non-null   float64
dtypes: float64(6), object(1)
memory usage: 439.7+ KB
```

Checking for Missing / Nan Values:

```
Date          0
Open         10
High         10
Low          10
Close        10
Adj Close    10
Volume       10
dtype: int64
```

The rows have any missing values are:

	Date	Open	High	Low	Close	Volume
1895	2003-04-14	NaN	NaN	NaN	NaN	NaN
2165	2004-04-26	NaN	NaN	NaN	NaN	NaN
2287	2004-10-13	NaN	NaN	NaN	NaN	NaN
3603	2010-02-06	NaN	NaN	NaN	NaN	NaN
4081	2012-01-07	NaN	NaN	NaN	NaN	NaN
4120	2012-03-03	NaN	NaN	NaN	NaN	NaN
4250	2012-09-08	NaN	NaN	NaN	NaN	NaN
4292	2012-11-11	NaN	NaN	NaN	NaN	NaN
4629	2014-03-22	NaN	NaN	NaN	NaN	NaN
4858	2015-02-28	NaN	NaN	NaN	NaN	NaN

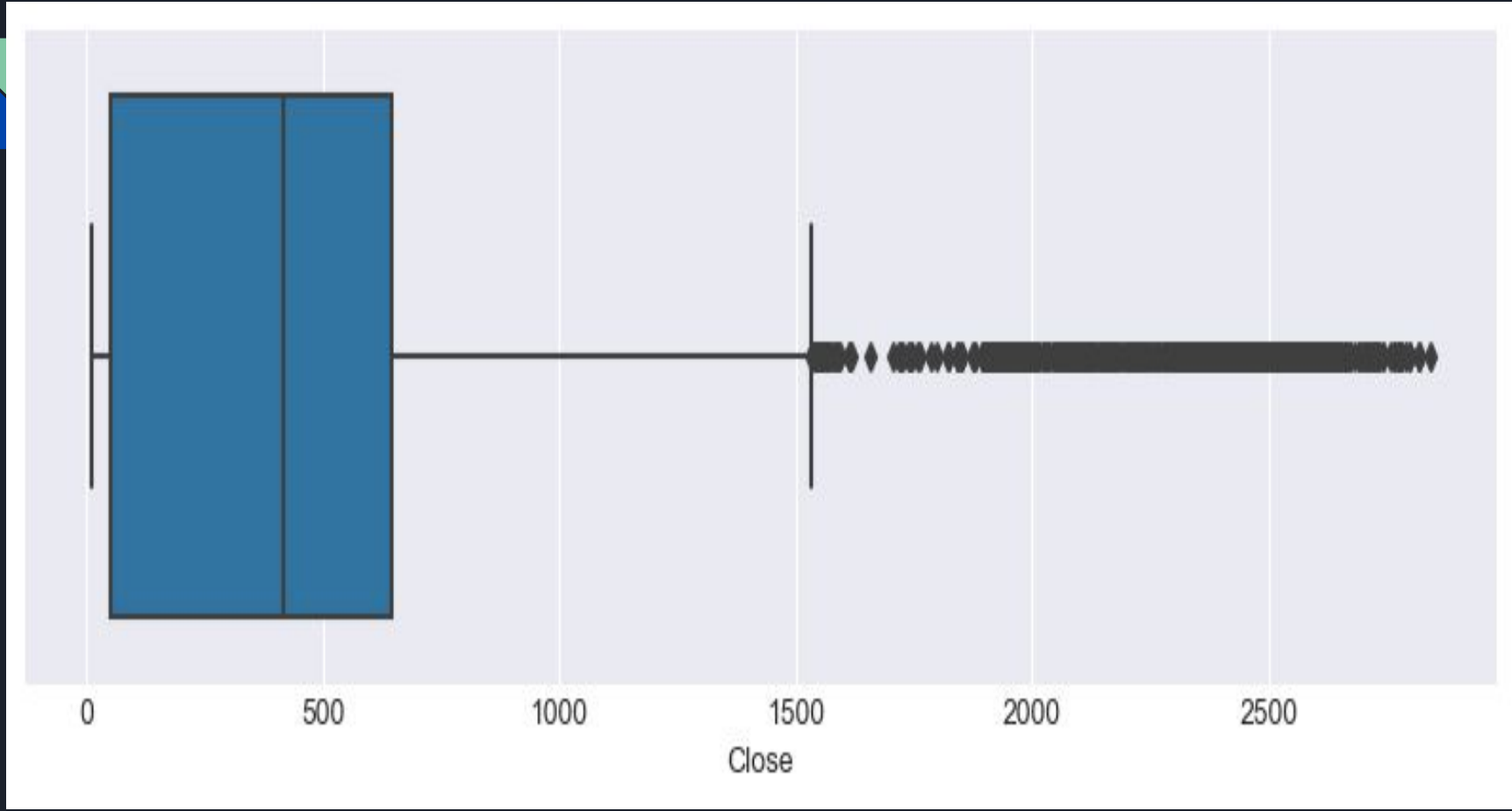
The percentage of missing values is 0.141. So we dropped the data.

Data Stats

	Close
count	7035.000000
mean	610.886871
std	739.898136
min	11.890704
25%	52.633965
50%	417.740234
75%	644.868530
max	2841.850098

Outlier Detection:

```
INTER QUARTILE RANGE IS  592.2345654999999
FIRST QUARTILE RANGE IS  644.86853
THIRD QUARTILE RANGE IS  52.633964500000005
          UPPER TAIL IS  940.9858127499999
          LOWER TAIL IS  -243.4833182499999
```

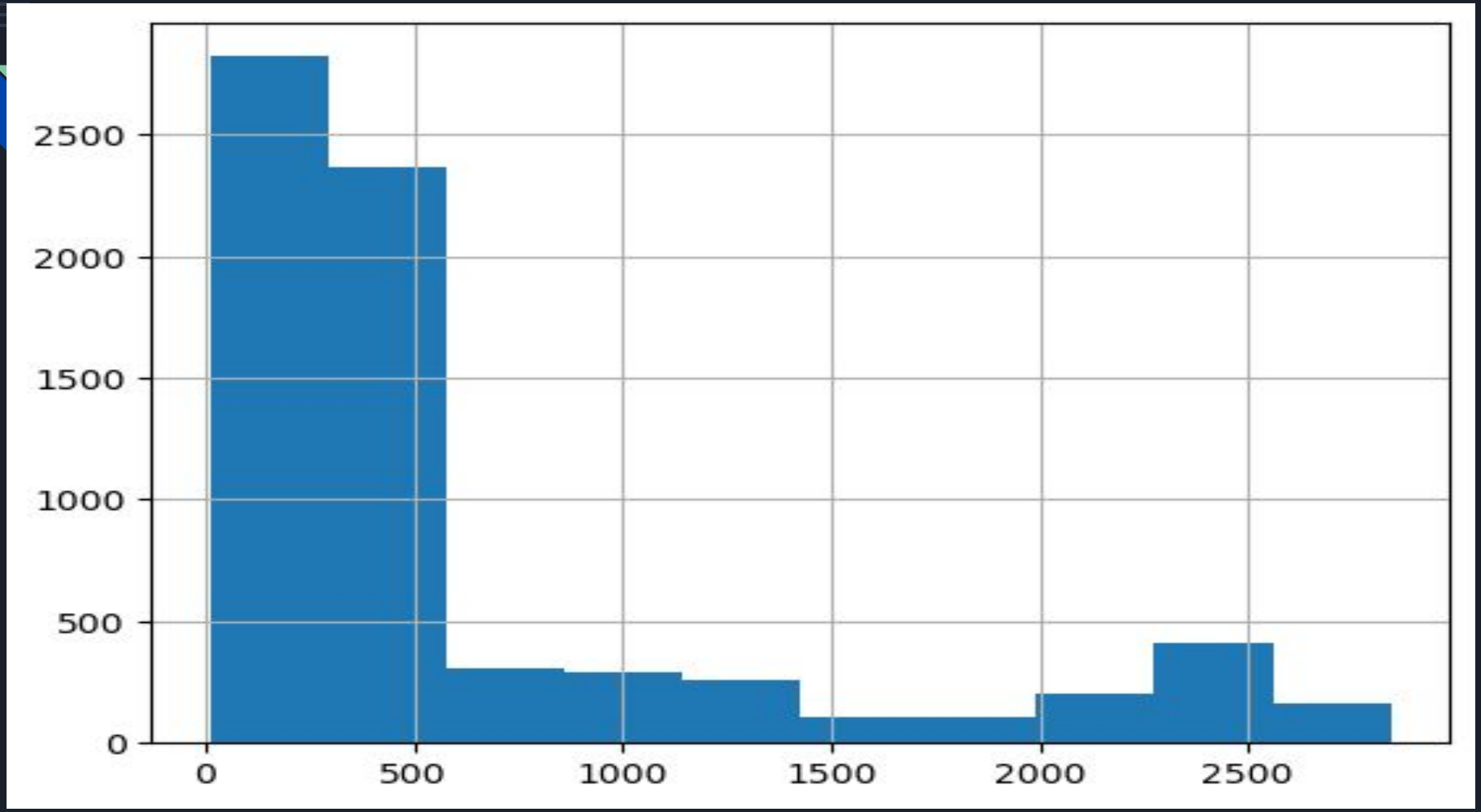

0
500
1000
1500
2000
2500
Close

Skewness of the Data:

```
Close            1.566294
Month           -0.009149
Year             0.028751
Daily Returns    9.447249
dtype: float64
```

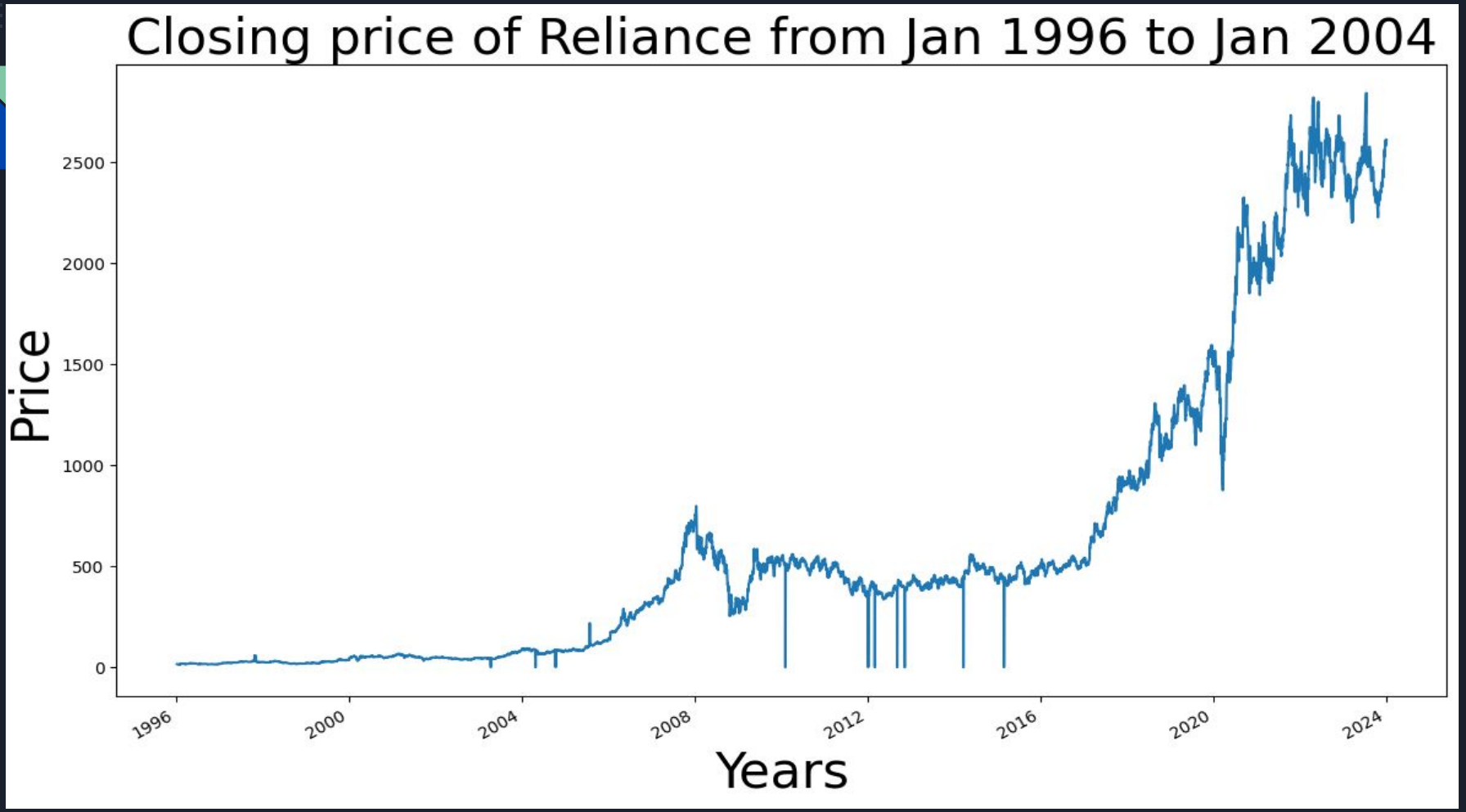
Inference : The skewness of a dataset measures the asymmetry of its distribution.

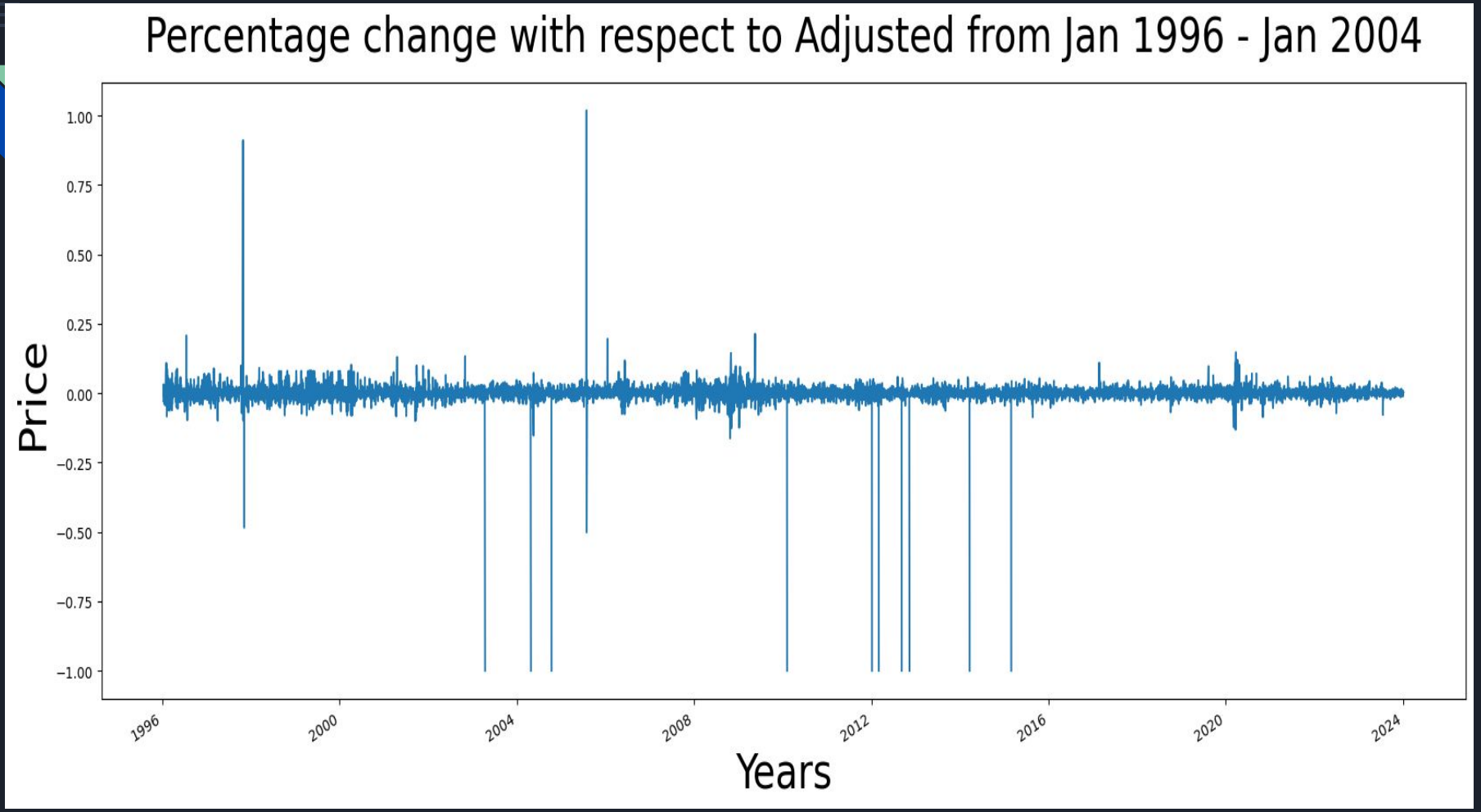
SKEWNESS 1.566294 : A positive skewness indicates that the data has a long tail to the right, meaning that there are more extreme values on the right side of the distribution. In the context of stock prices, a positive skewness might suggest that there are occasional large increases in the stock price, leading to a right-skewed distribution.


2500
2000
1500
1000
500
0
0
500
1000
1500
2000
2500

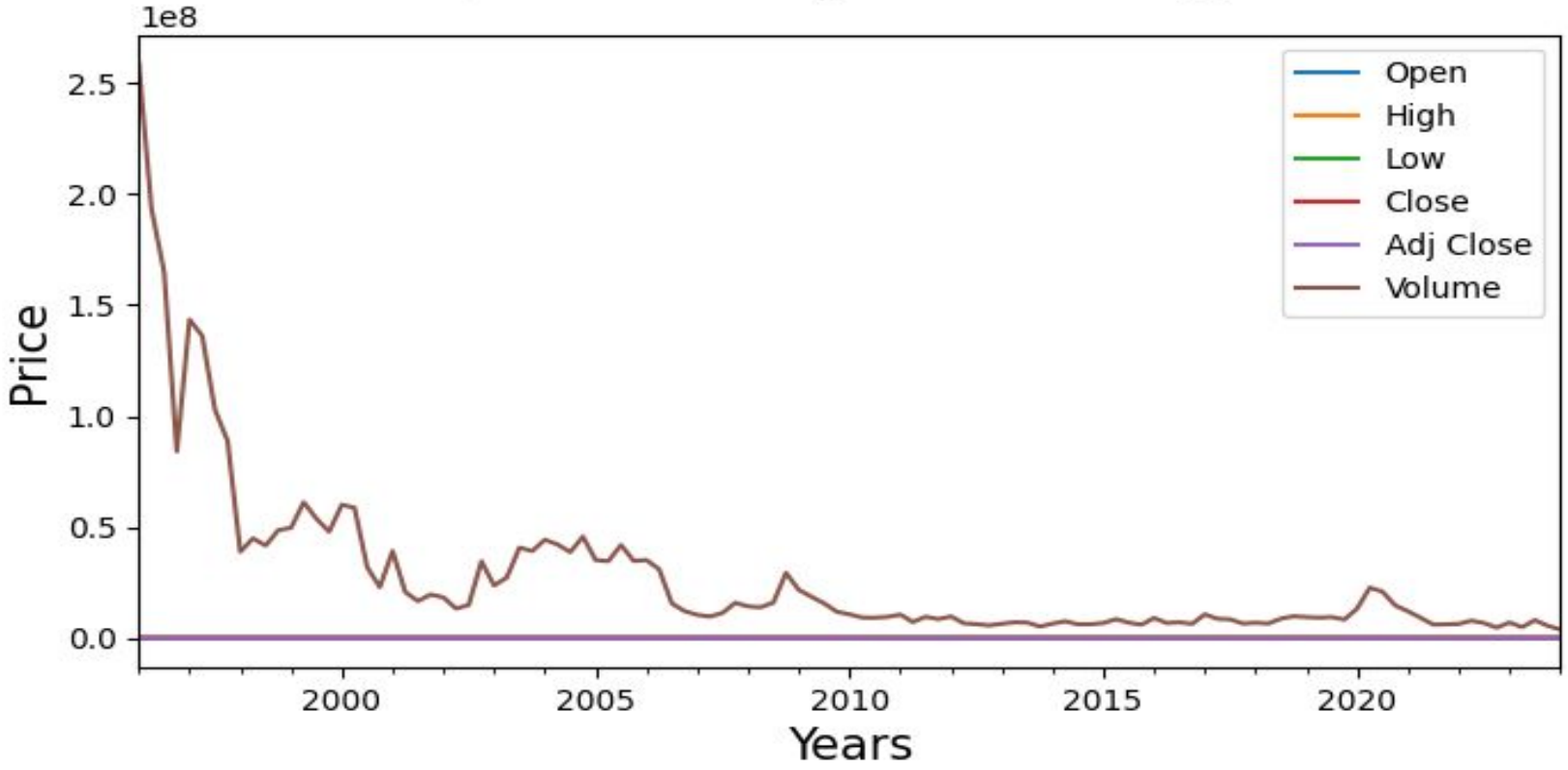
Exploratory Data Analysis Visualizations:

- Closing stock price from Jan 1996 to Jan 2004.
- Percentage change with respect to Adjusted close from Jan 1996 to Jan 2004.
- Quarterly Average line Plot.
- Comparing the Quarterly mean of both High and Low(Line plot).
- Comparing the Quarterly mean of both High and Low(Bar plot).
- Plotting the Close price through a histogram.
- Violinplot depicting Closing stock price.
- Boxplot for Outlier detection.
- Monthly visualization using box plot.

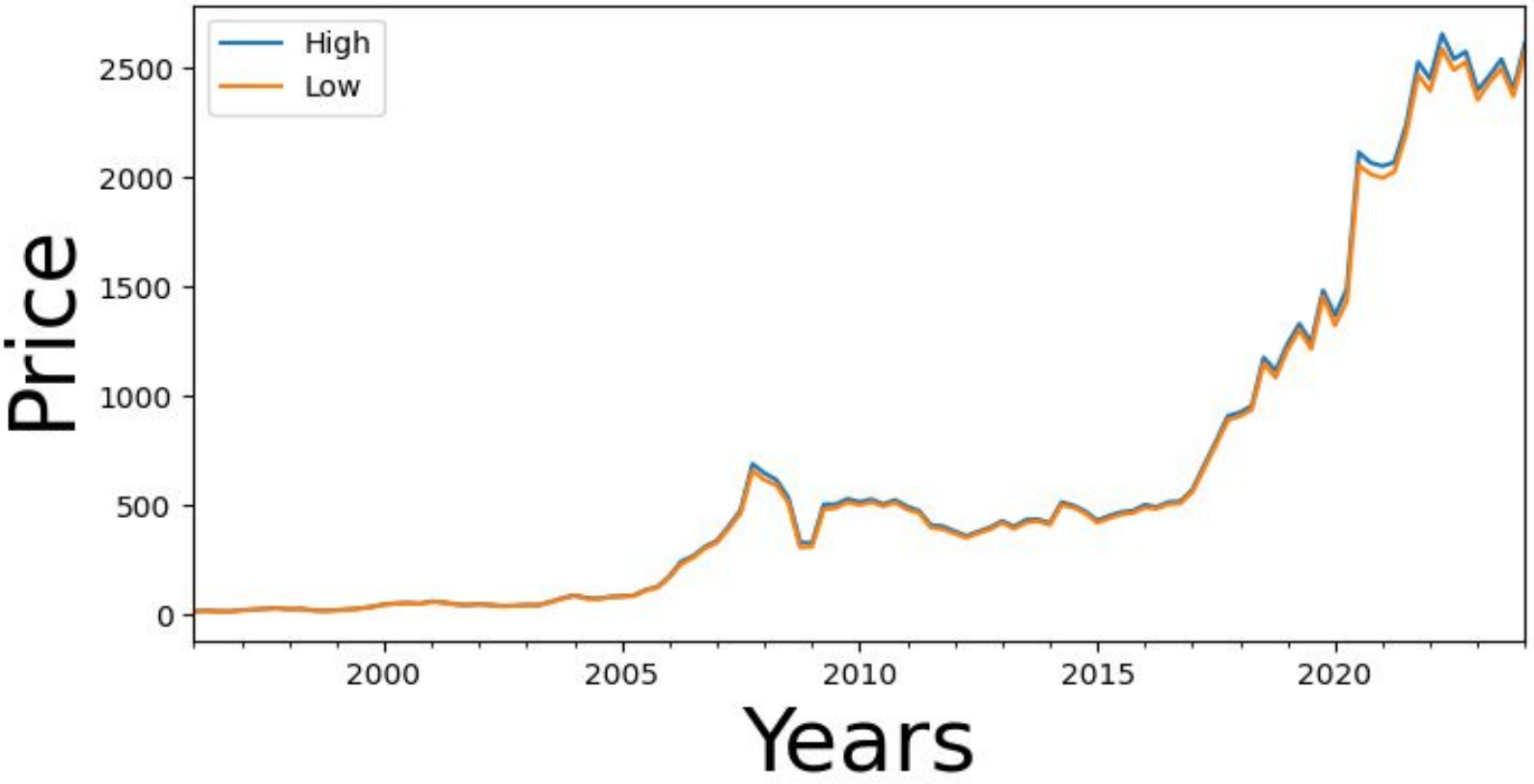

Closing price of Reliance from Jan 1996 to Jan 2004
Price
2500
2000
1500
1000
500
0
1996
2000
2004
2008
2012
2016
2020
2024
Years


Percentage change with respect to Adjusted from Jan 1996 - Jan 2004
Price
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
1996
2000
2004
2008
2012
2016
2020
2024
Years

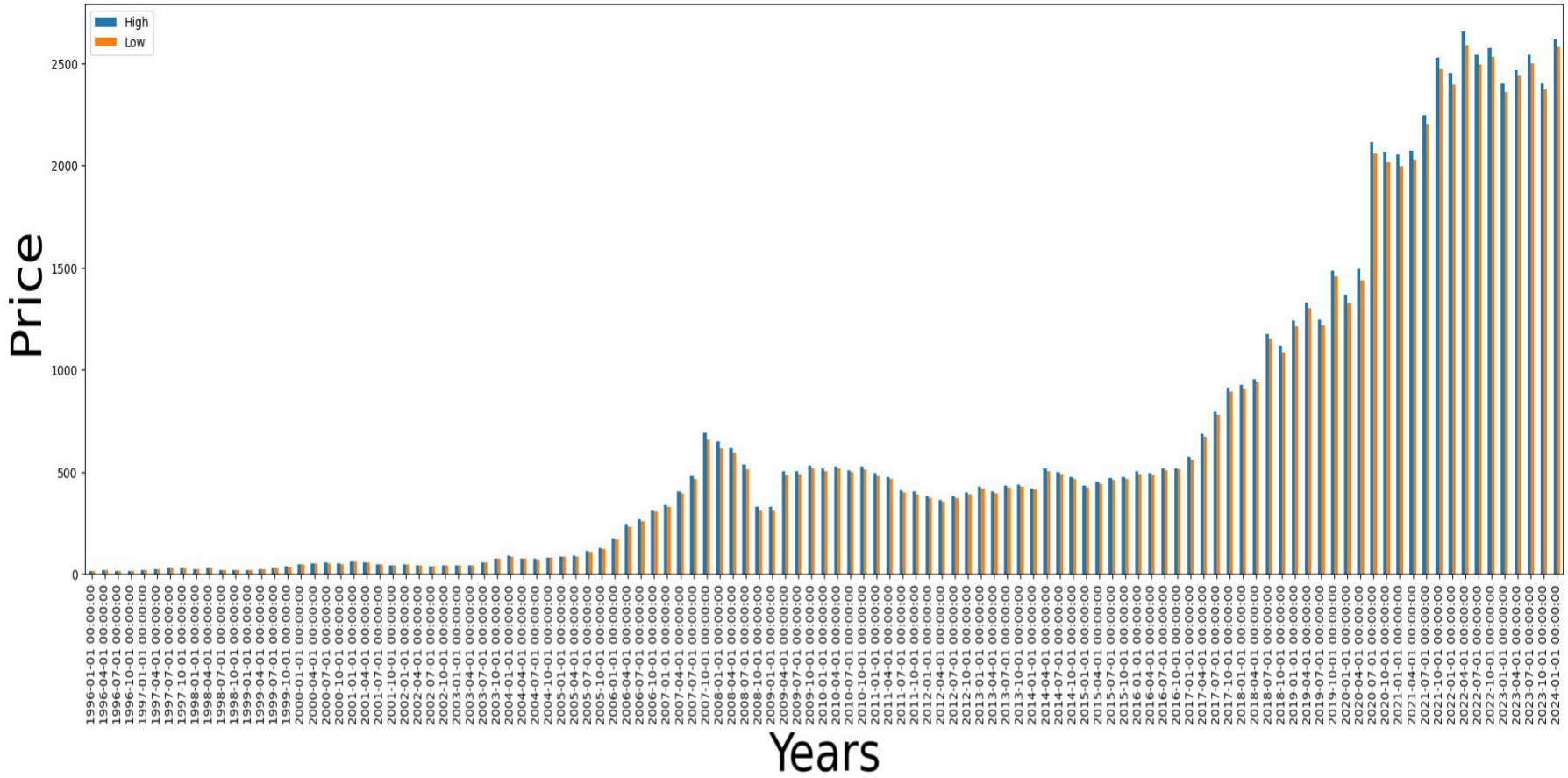
Quarterly Average



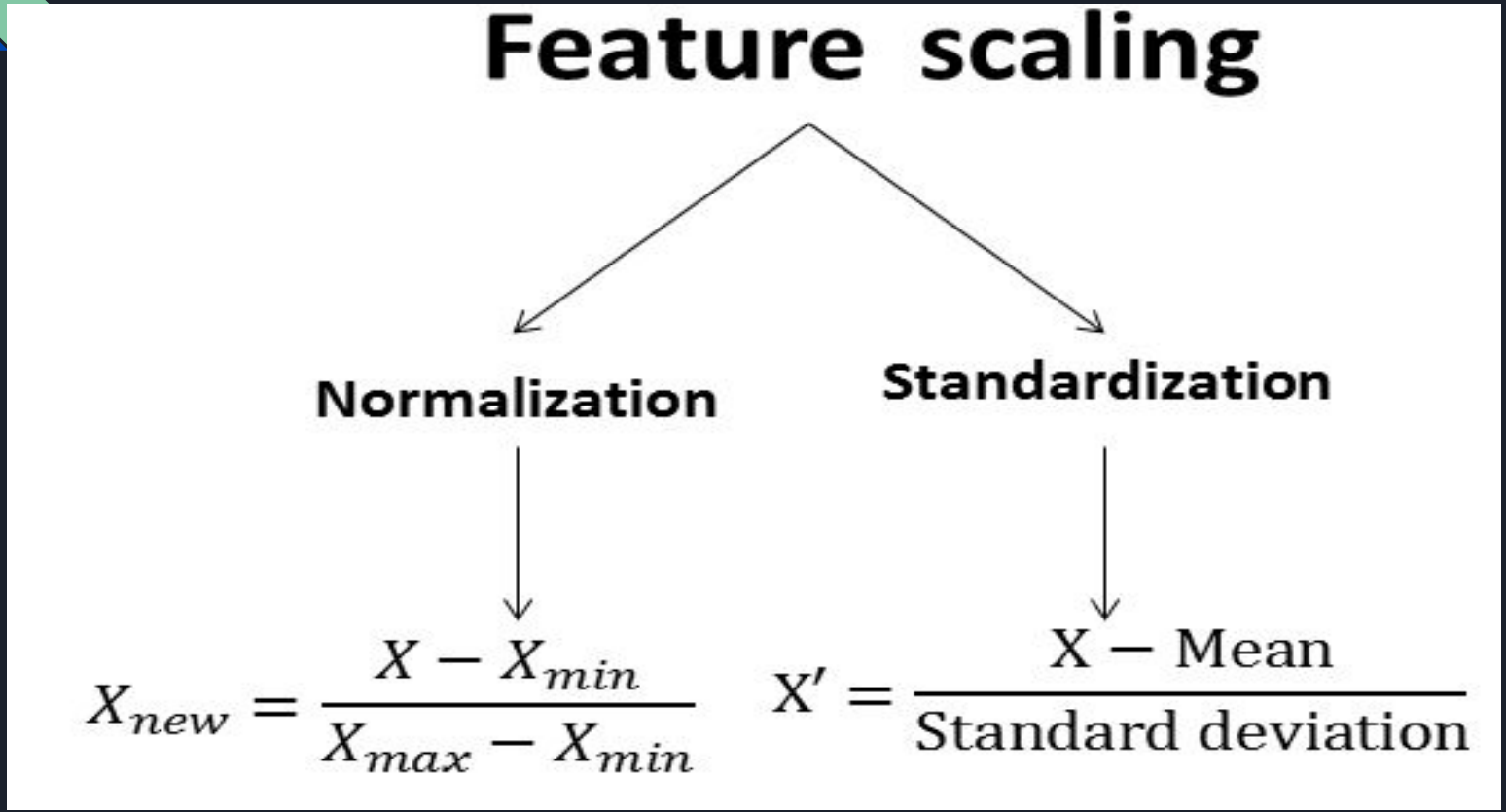
Quarterly mean of both high and low



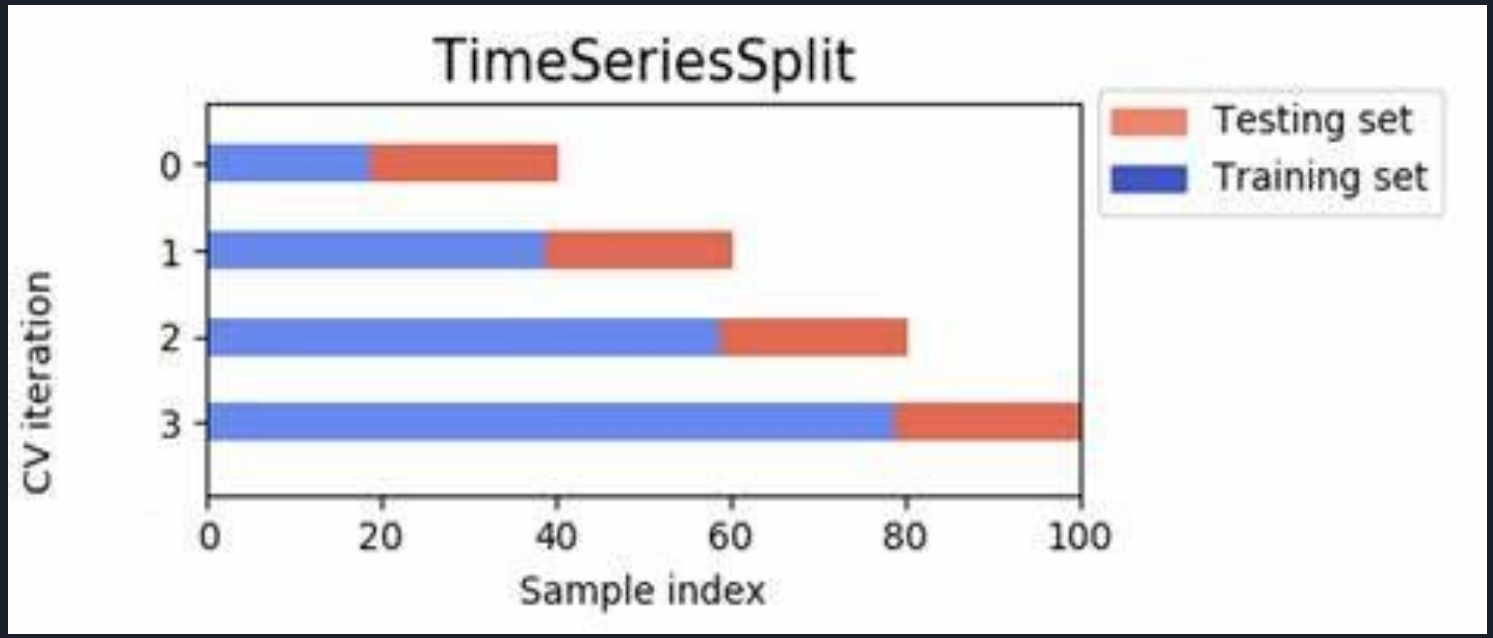
Quarterly mean bar of both high and low



Scaling of Data:



Splitting of Data:

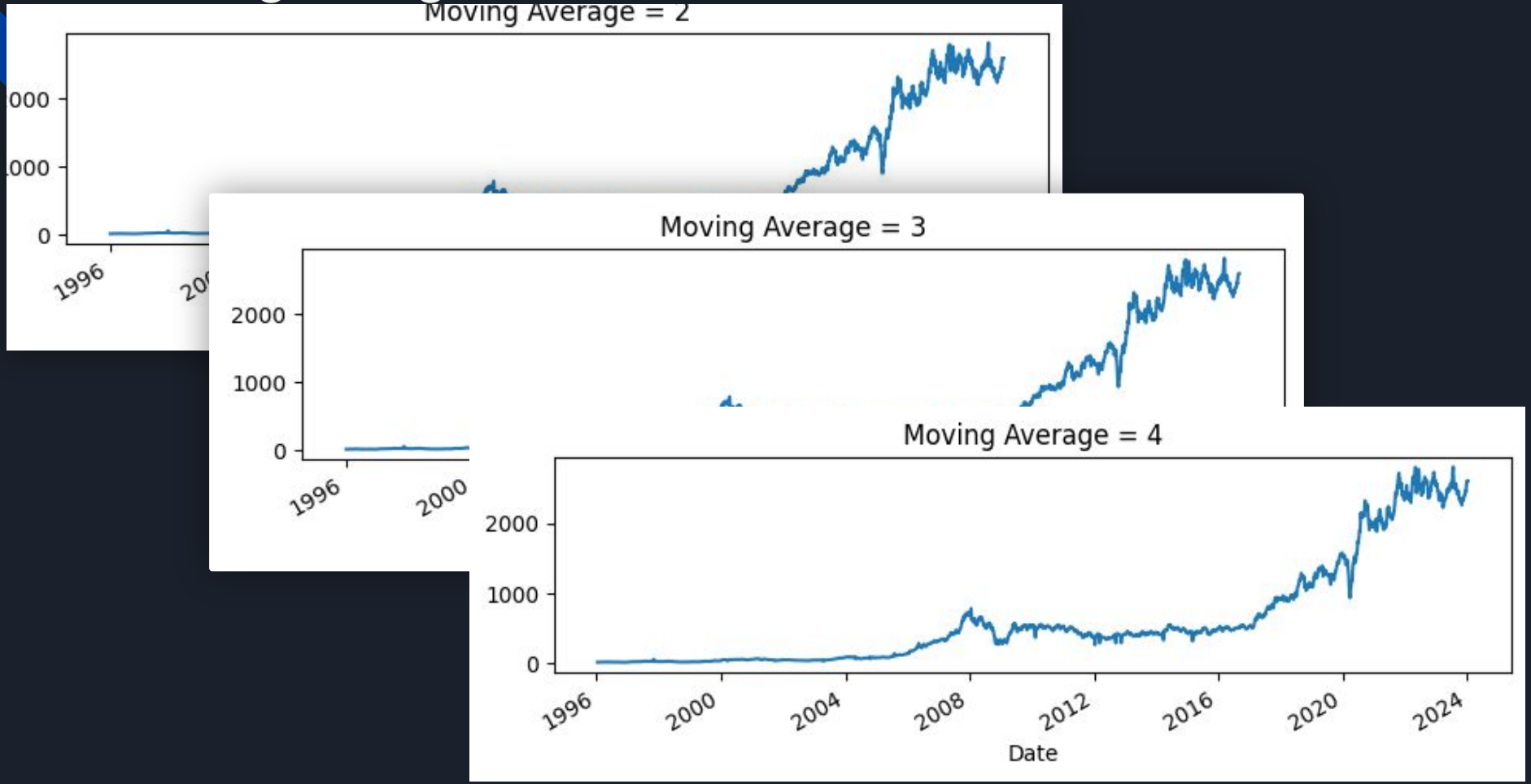


Model Building in Time Series Forecasting

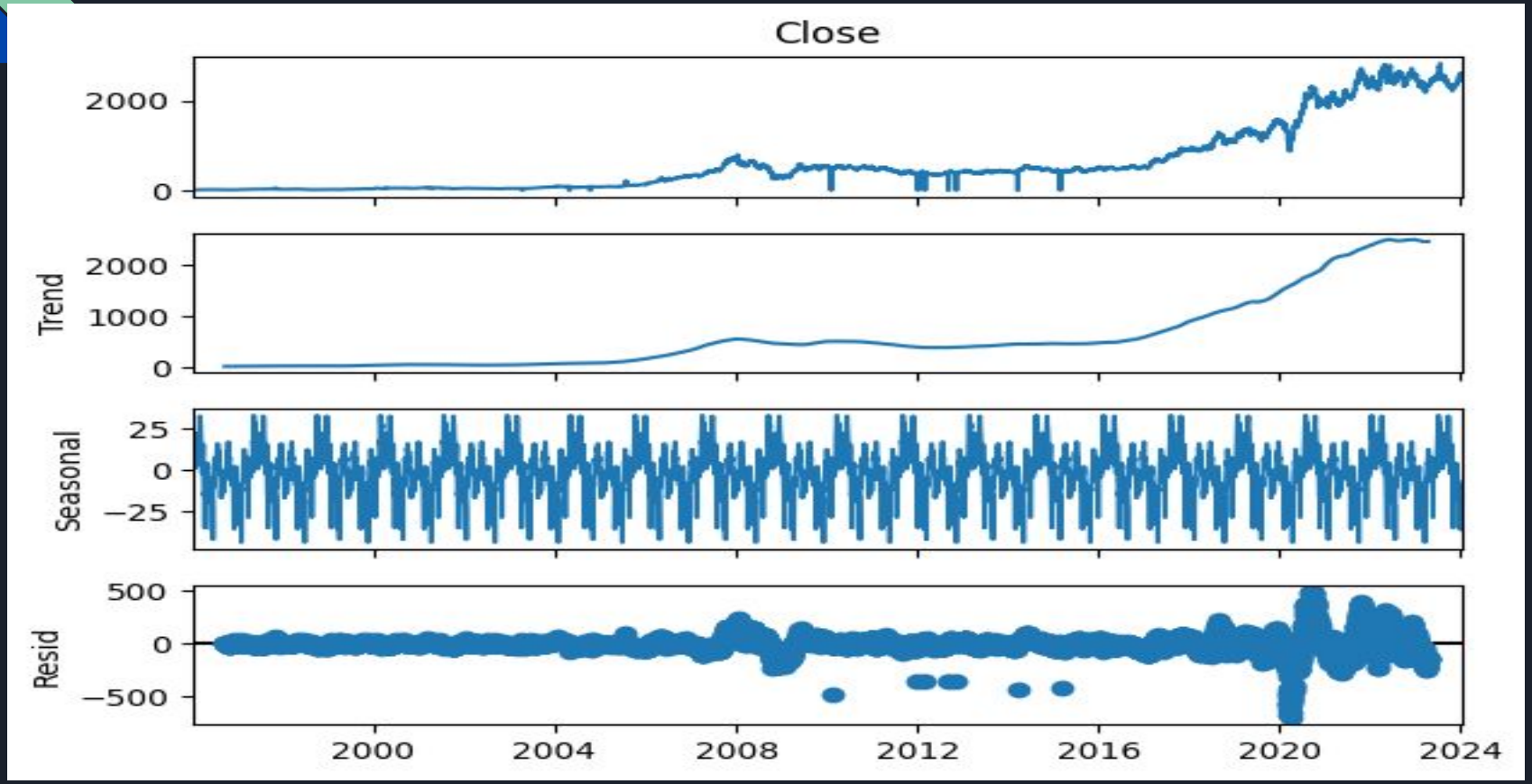
Top 10 Time Series Forecasting Methods in Data Science

Method	Description
ARIMA	Autoregressive Integrated Moving Average
Exponential Smoothing	Includes Holt-Winters and other variants
Prophet	True positives over all actual positives.
SARIMA	Harmonic mean of precision and recall.
LSTM	Model's ability to distinguish between classes.
GRU	Average absolute difference between predicted and true.
VAR	Average squared difference between predicted and true.
Theta Method	Square root of MSE.
TBATS	Loss based on probability estimates.
Facebook Prophet	Variance explained by the model.

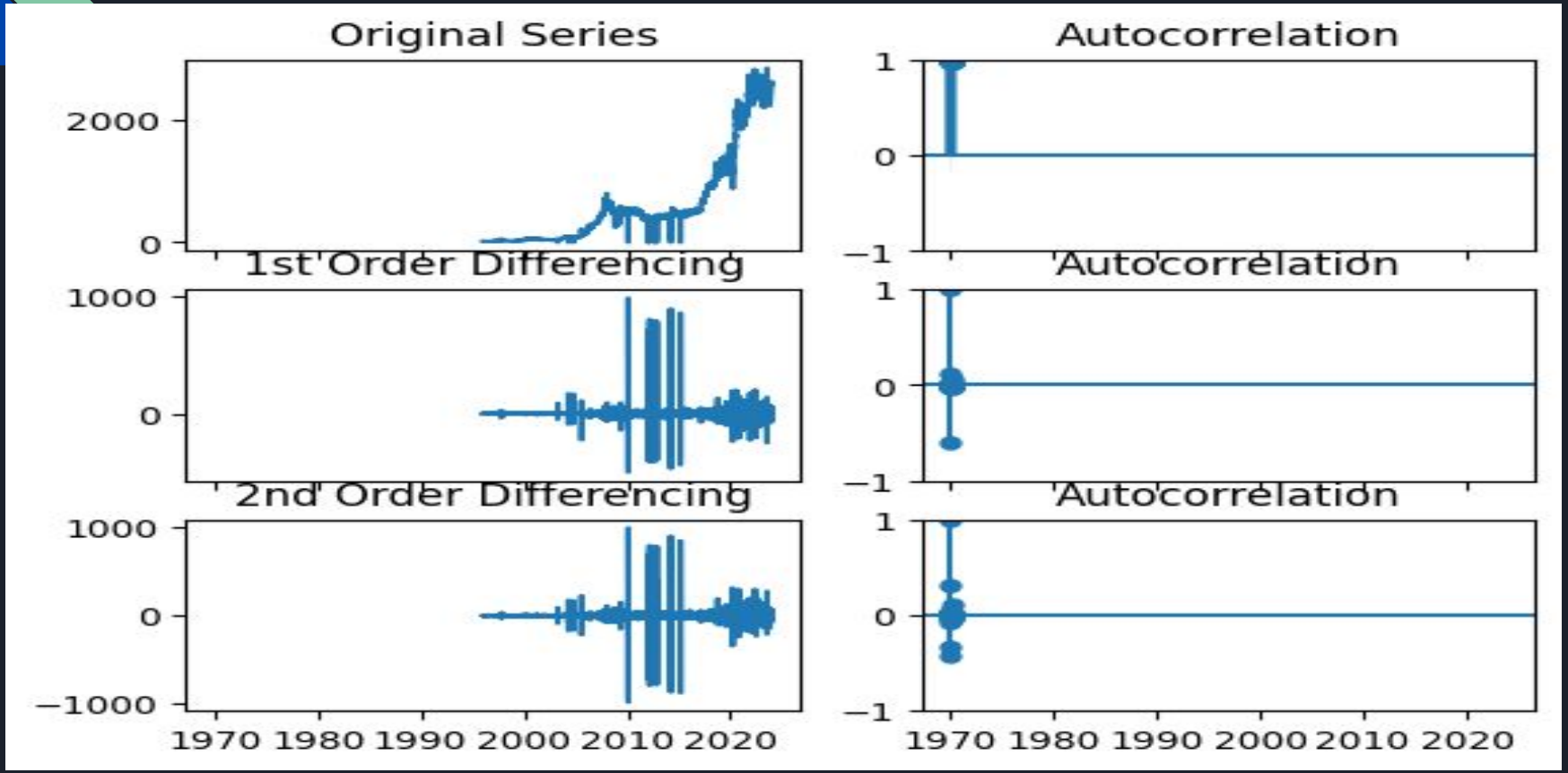
Moving Average:



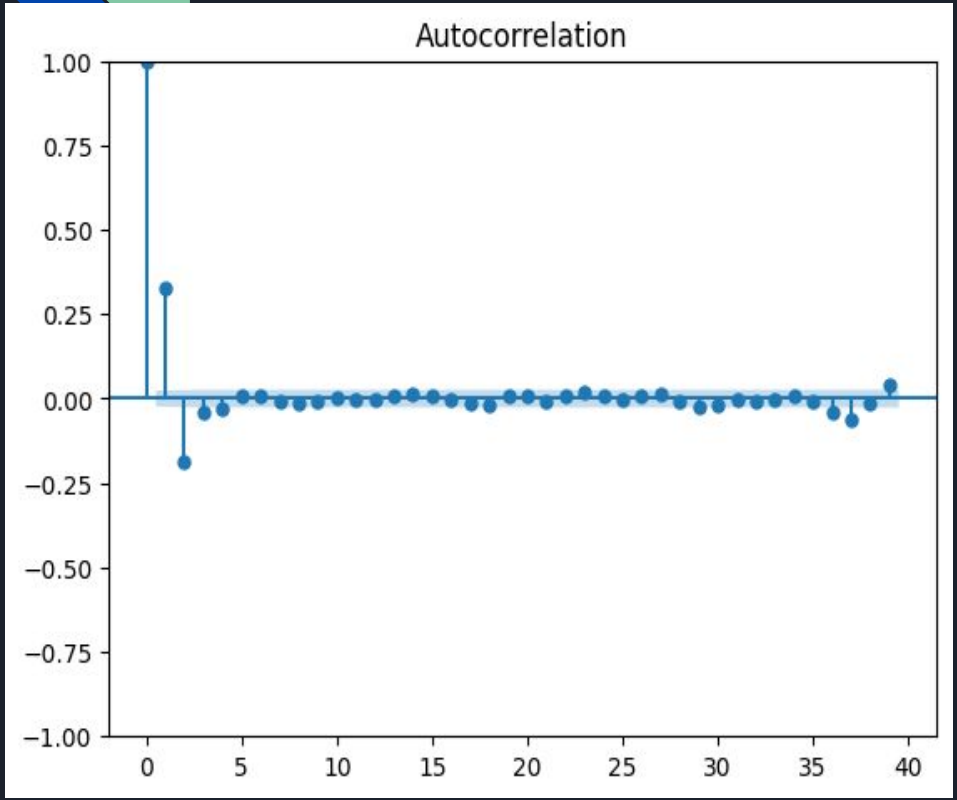
Decomposition of Data:

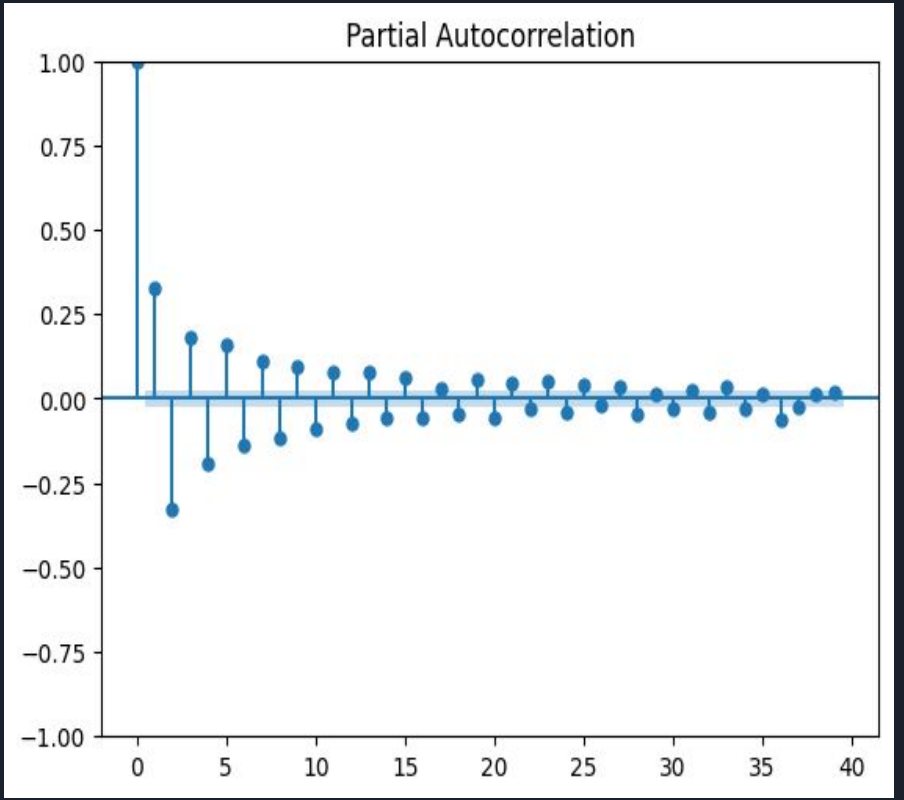


ACF plot after differencing:



ACF and PACF plot:





ARIMA model:

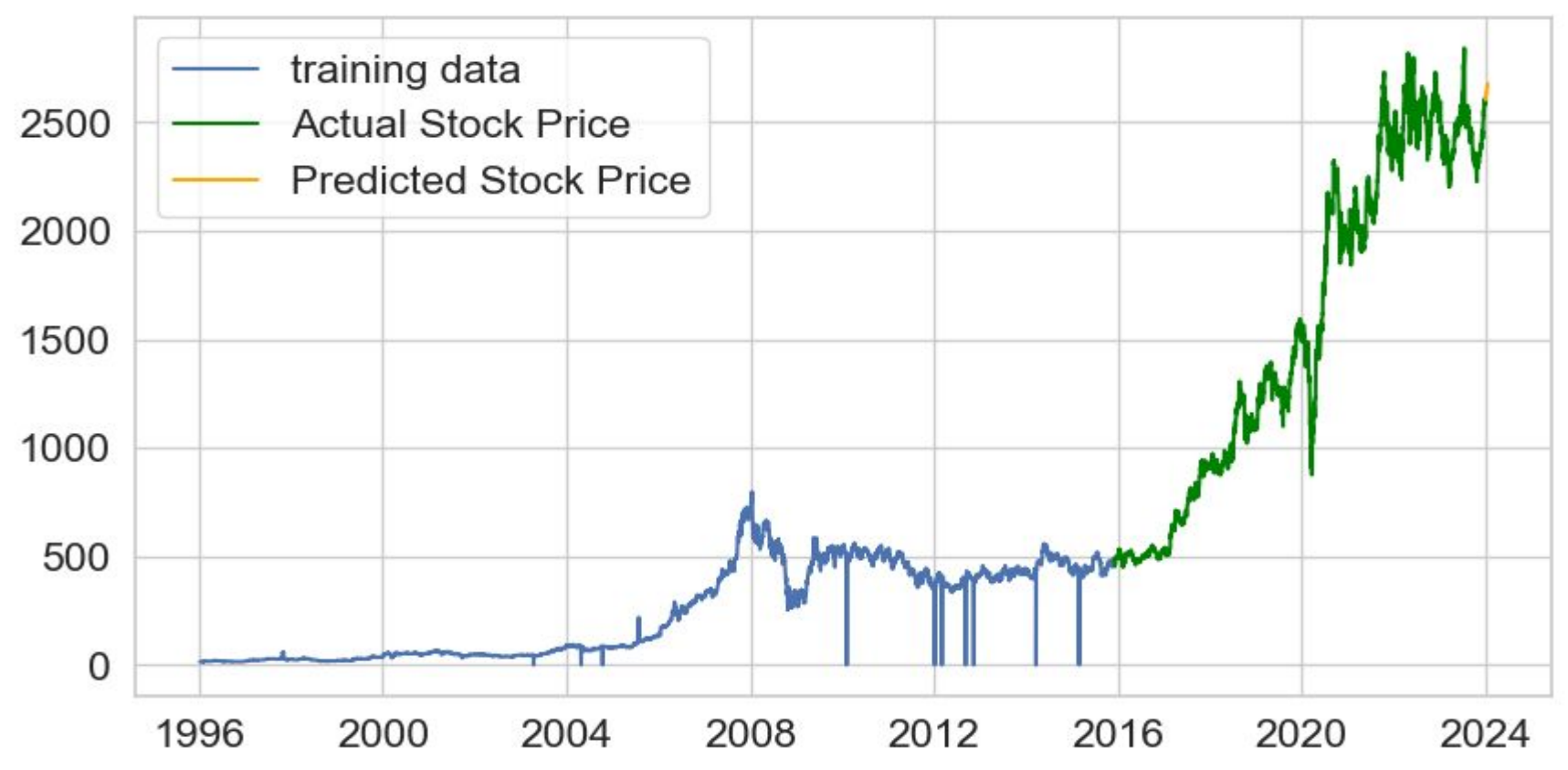
- ARIMA model applied through stats model, the best value for p,d,q has been foundto be (2,2,1) fetching the following metric values.
- MSE: 267.840
- MAE: 14.3465
- RMSE: 16.3658
- MAPE: 0..005531
- 15 days of forecasted Data has been drawn and plotted along with the train and test data.

INFERENCE:

- The data shows a upward trend from 2016 onwards.
- There has been fluctuations in the data.

The forecasted values of 15 days shows a upward trend in the data.

Forecast for 15 days:



SARIMA Model:

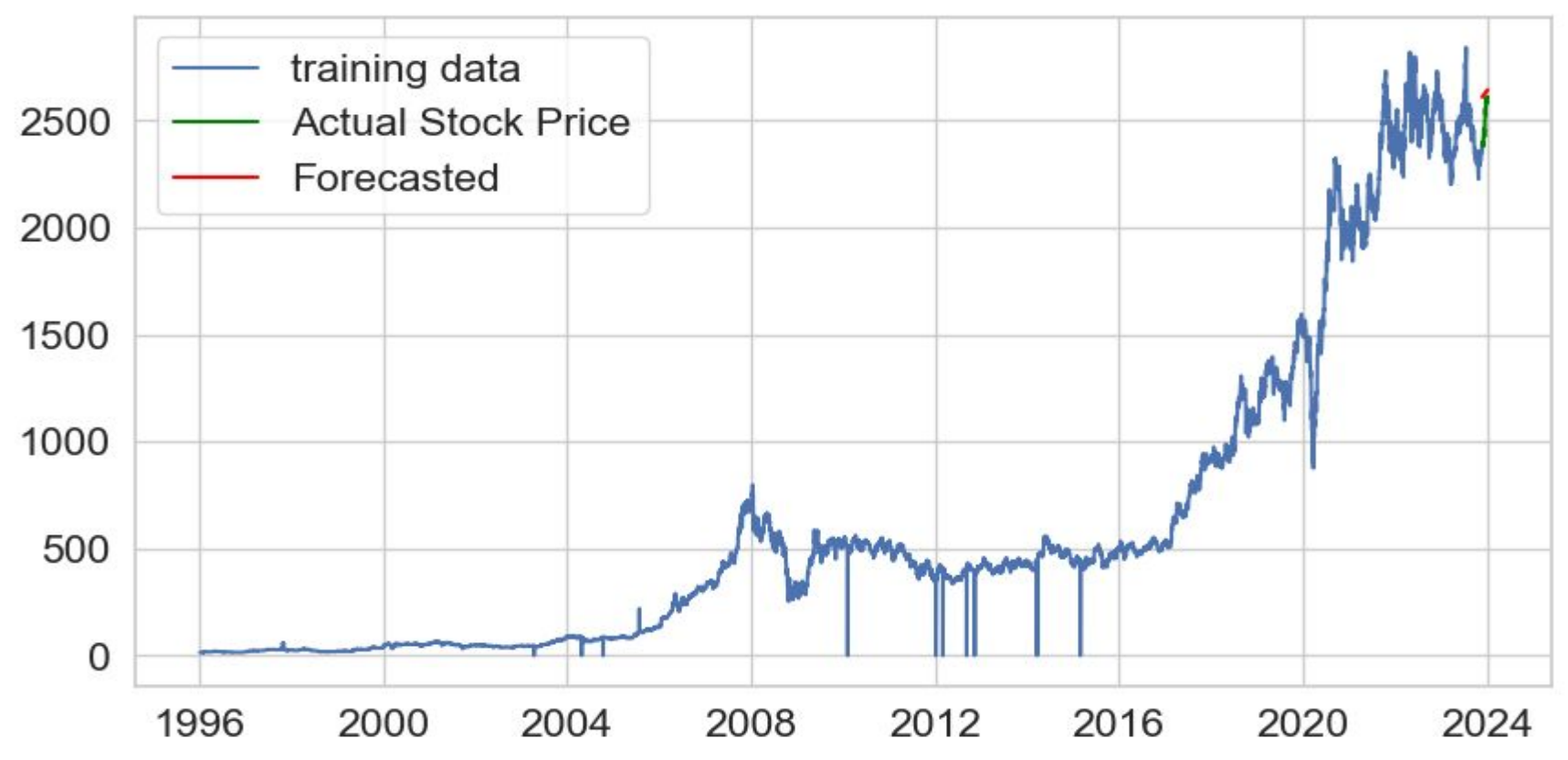
- SARIMA model has been applied with max of 0 to 3 for p q and d.
- Generated a ARIMA model with seasonality that is a SARIMAX model with (1,1,1) and seasonality (2,1,0,12) values for p, d, and q.
- Metrics obtained for the above Data is as follows:
- MSE: 226.1910
- MAE: 12.6365
- RMSE: 15.0396
- MAPE: 0.00487877
- 15 days forecast has been drawn and plotted along with the train and test data.

INFERENCE:

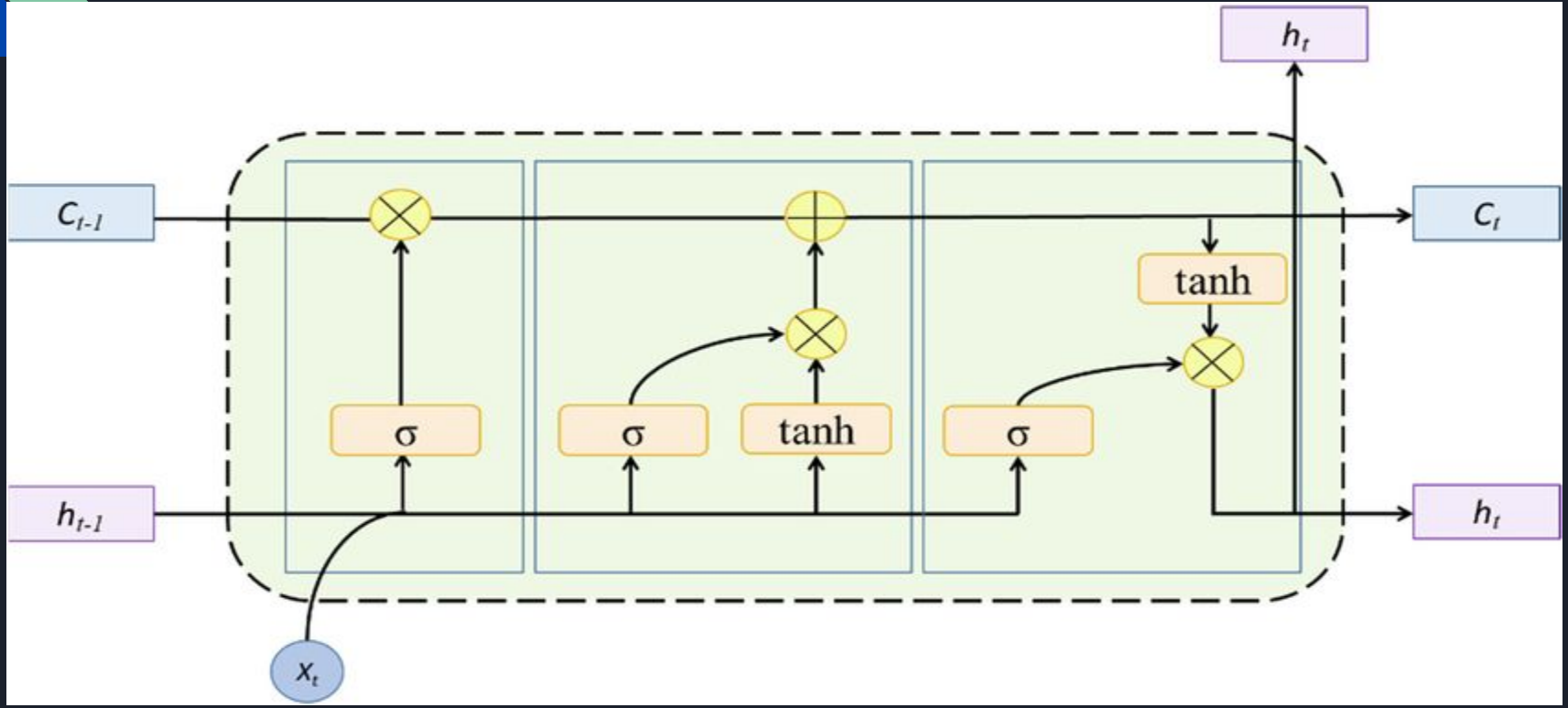
- The data shows a upward trend from 2016 onwards.
- There has been fluctuations in the data.

The forecasted values of 15 days shows a upward trend in the data.

Forecasting for 30 days:



LSTM model(working):



LSTM model:

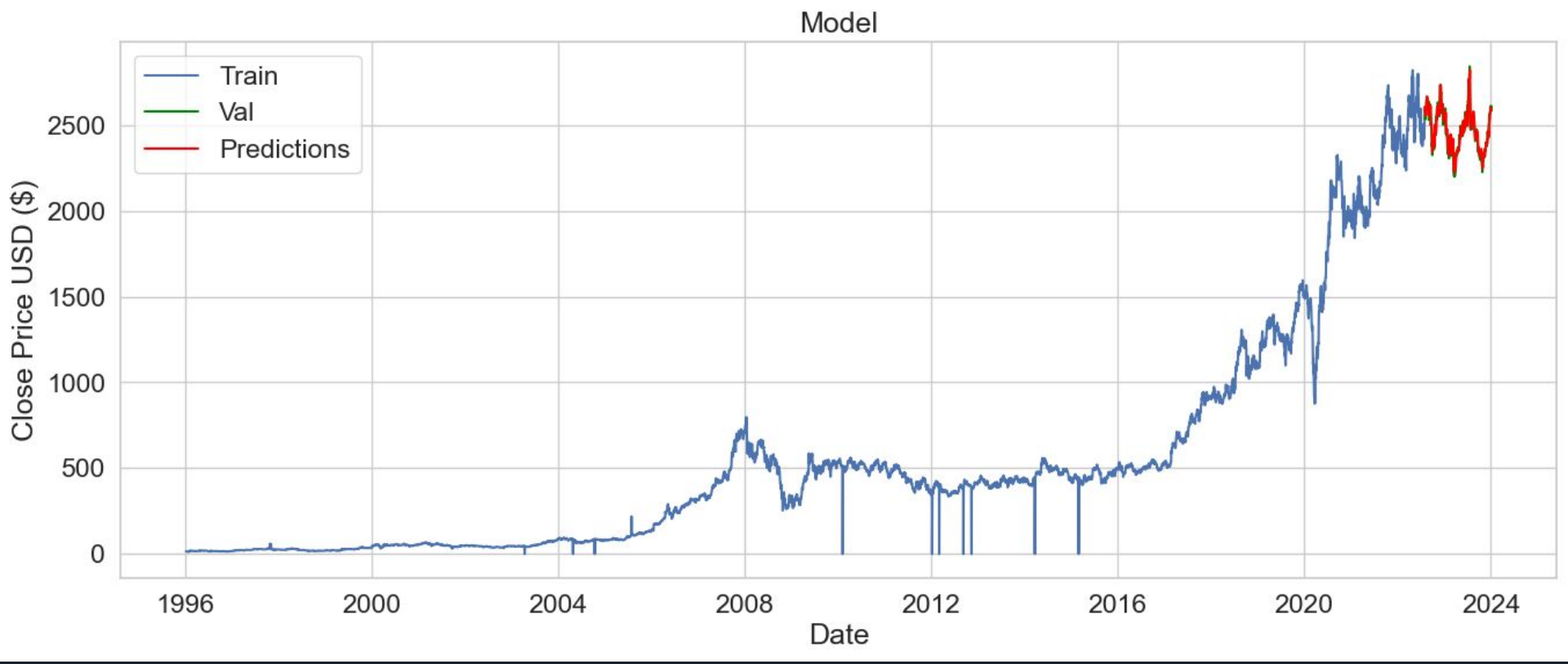
- LSTM (Long- short term memory) model applied .
- Three hidden layers created with one output layer.
- Since LSTM is a part of neural network, it follows the gradient descent and the optimizer used is Adam, and the loss function is taken as mean squared error as its a regression problem.
- fetching the following metric values.
- MSE: 548.597
- MAE: 14.8856
- RMSE: 23.42216
- MAPE: 0.0535
- 15 days of forecasted Data has been drawn and plotted along with the train and test data.

INFERENCE:

- The data shows a upward trend from 2016 onwards.
- There has been fluctuations in the data.

The forecasted values of 15 days shows a upward trend in the data.

LSTM Model:


Model
Train
Val
Predictions
Close Price USD ($)
2500
2000
1500
1000
500
0
1996
2000
2004
2008
2012
2016
2020
2024
Date

Prophet Model:

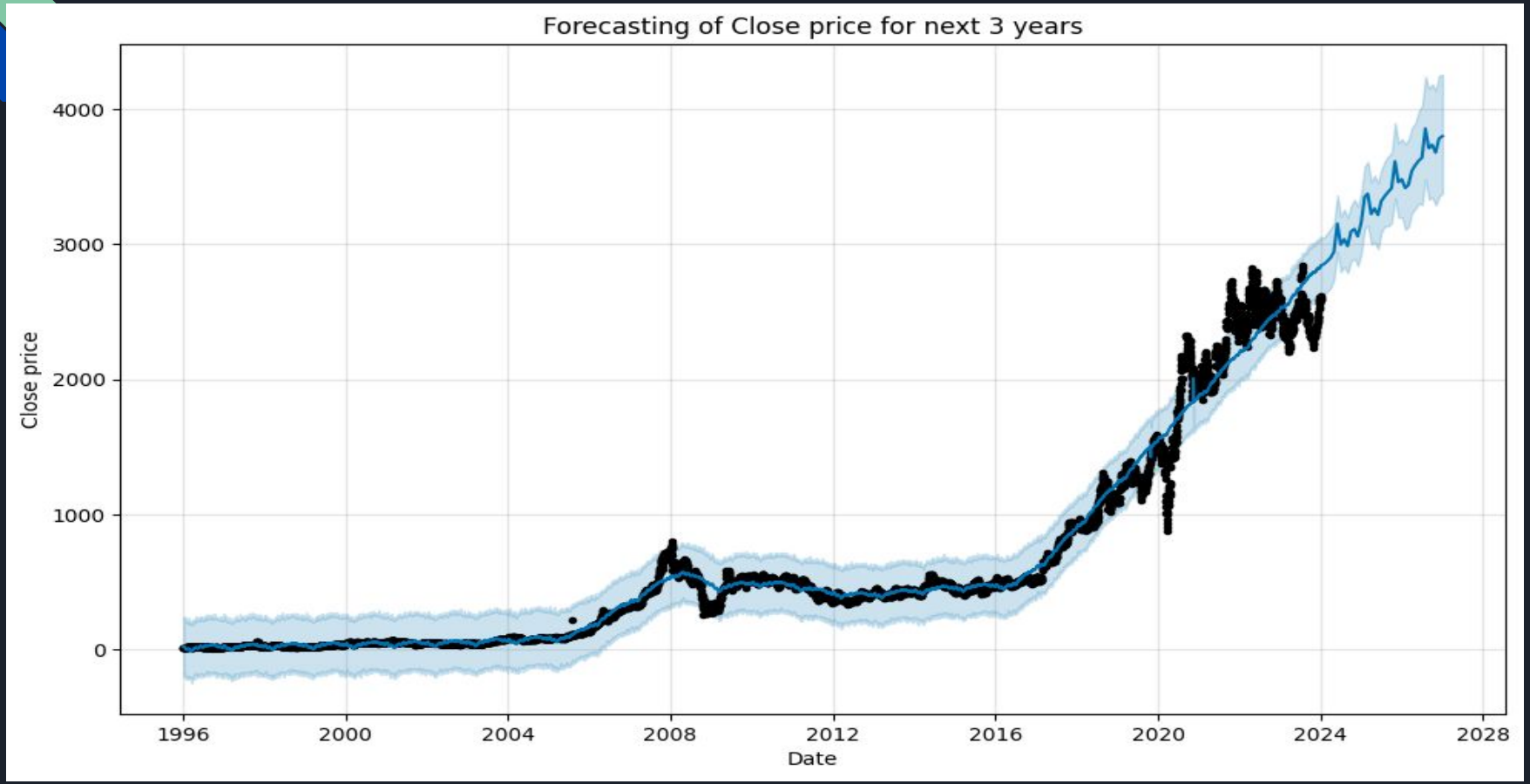
- The Prophet model is a method for forecasting time series data that is based on an additive model.
- It can handle data with trends, seasonality, and holidays, and it can automatically adjust for changes in the data. .
- fetching the following metric values.
- MSE: 11534.503
- MAE: 56.8109
- RMSE: 107.3988
- MAPE: 0.18806
- 3 years of forecasted Data has been drawn and plotted along with the train and test data.

INFERENCE:

- The data shows a upward trend from 2016 onwards.
- There has been fluctuations in the data.

The forecasted values of 3 years shows a upward trend in the data.

Forecast for the next 3 years:



Comparison of different models with different scores:

- **MSE(Mean Square error):**

Model Name	Score
ARIMA	267.840
SARIMA	226.191
LSTM	548.597
PROPHET	11534.503

- **MAE(Mean Absolute error):**

Model Name	Score
ARIMA	14.3465
SARIMA	12.6365
LSTM	14.8856
PROPHET	56.8109

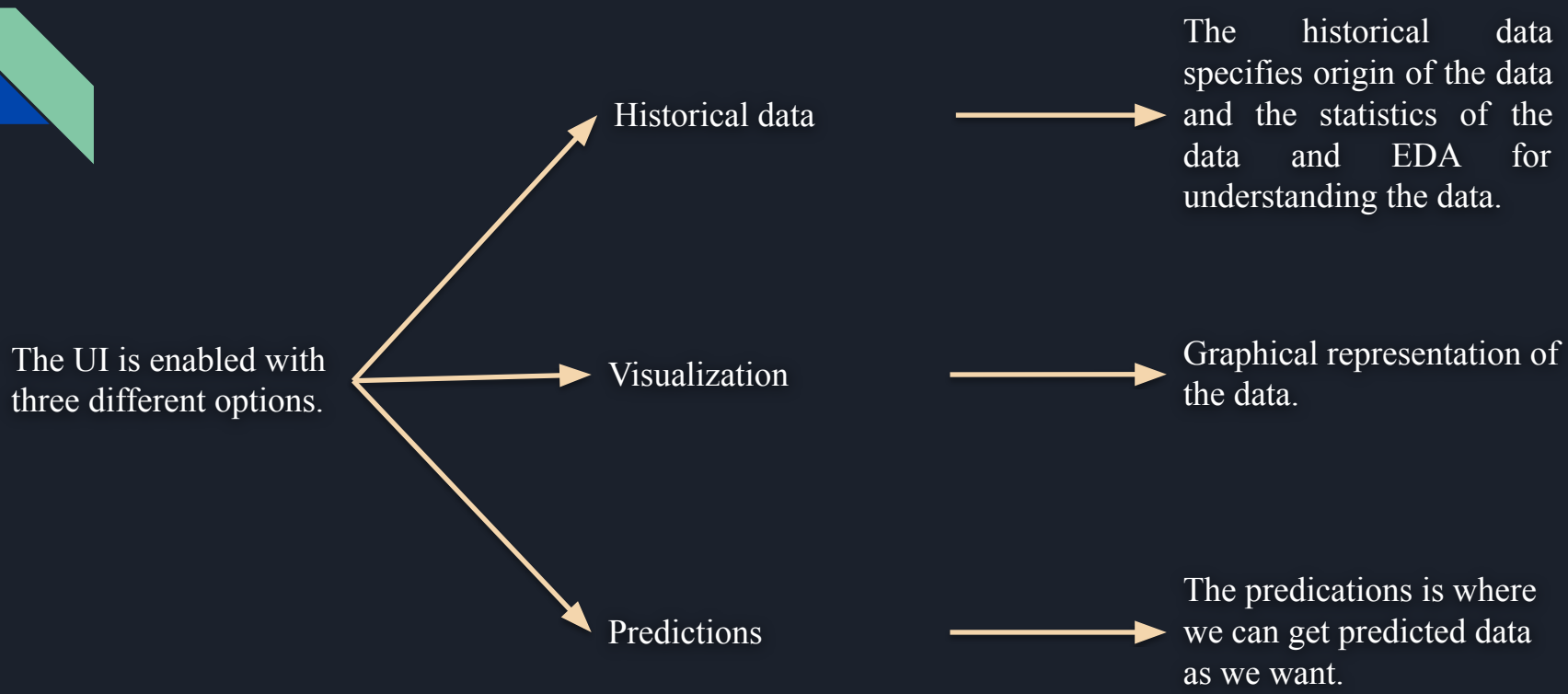
- **RMSE(Root Mean Square error):**

Model Name	Score
ARIMA	16.3658
SARIMA	15.0396
LSTM	23.4221
PROPHET	107.3880

- **MAPE(Mean Absolute Percentage error):**

Model Name	Score
ARIMA	0.005531
SARIMA	0.004878
LSTM	0.05355
PROPHET	0.18806

Deployment:


The UI is enabled with three different options.
Historical data
Visualization
Predictions
The historical data specifies origin of the data and the statistics of the data and EDA for understanding the data.
Graphical representation of the data.
The predications is where we can get predicted data as we want.

Historical data:

Options

- Historic data
- Visualization
- Predictions

Deploy

Stock market Prediction of Reliance Industries

Historic data from 1996 to 2024 Jan

Stock Market

Show Table

	Open	High	Low	Close	Adj Close	Volume
1996-01-08 00:00:00	15.6273	15.6389	14.8769	15.0316	9.8601	86,288,584.0000
1996-01-09 00:00:00	14.5868	14.8537	14.0994	14.5984	9.5759	179,415,702.0000
1996-01-10 00:00:00	14.3122	14.6564	14.1574	14.2232	9.3298	127,653,926.0000
1996-01-11 00:00:00	14.0801	14.7377	14.0607	14.6680	9.6216	189,051,436.0000
1996-01-12 00:00:00	14.7763	14.8885	14.4166	14.5133	9.5201	172,918,416.0000
1996-01-15 00:00:00	14.4708	14.5443	14.2812	14.3625	9.4211	92,291,448.0000
1996-01-16						

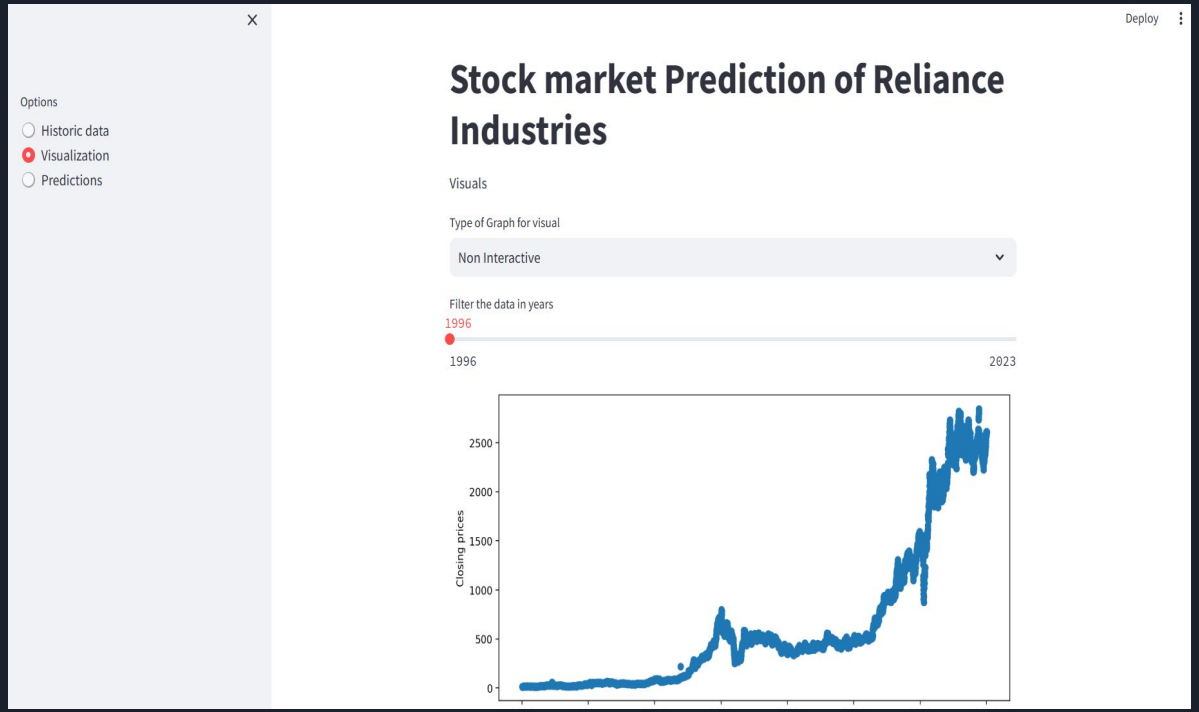
Visualization:

Here in the visualization window we have two modes:

- ❑ Interactive mode
- ❑ Non interactive mode

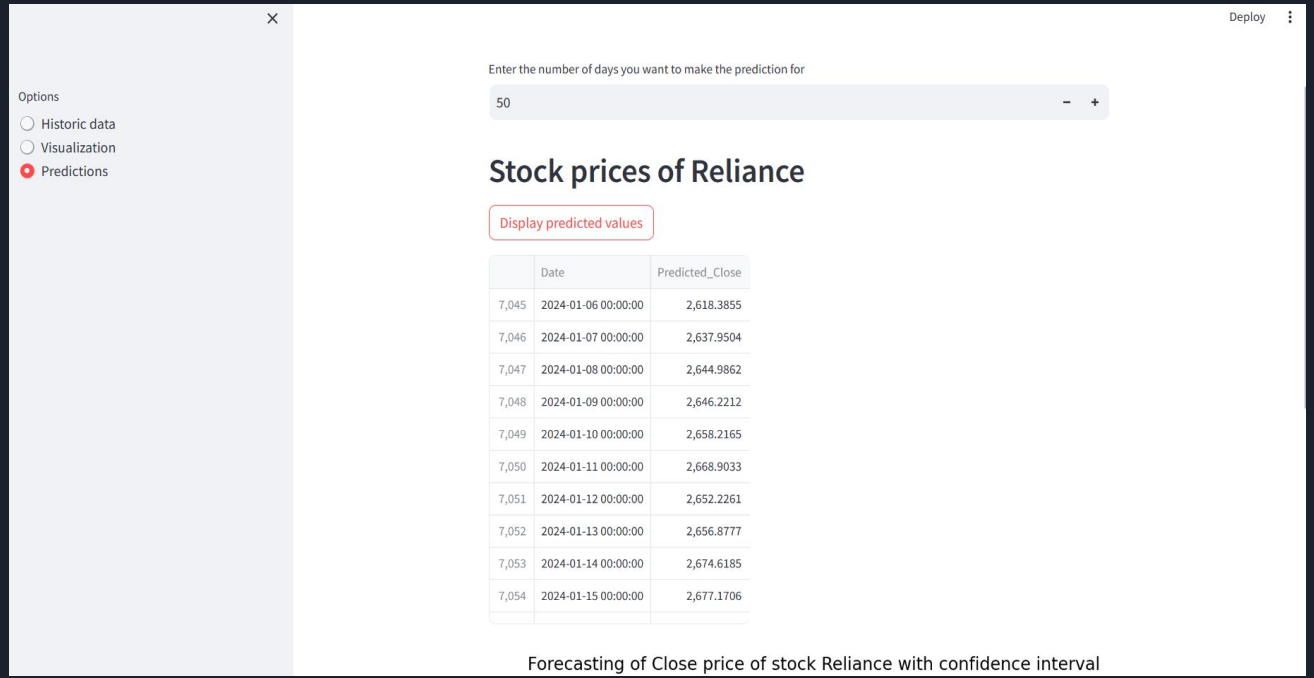
Interactive mode: we can observe whole data graphical representation at a time.

Non-Interactive mode: Here we can customize or filter how many years we require.



Predictions:

- The final page here we can get predicted data as much we require and we can download the data as for reference.



THANK YOU