Time Series Data Visualization **Roll No: MS2317**

Dataset Used :"stock_data.csv"

Plotting Line for Time Series Data :

The volume column is of continuous data type, we will use line graph to visualize it

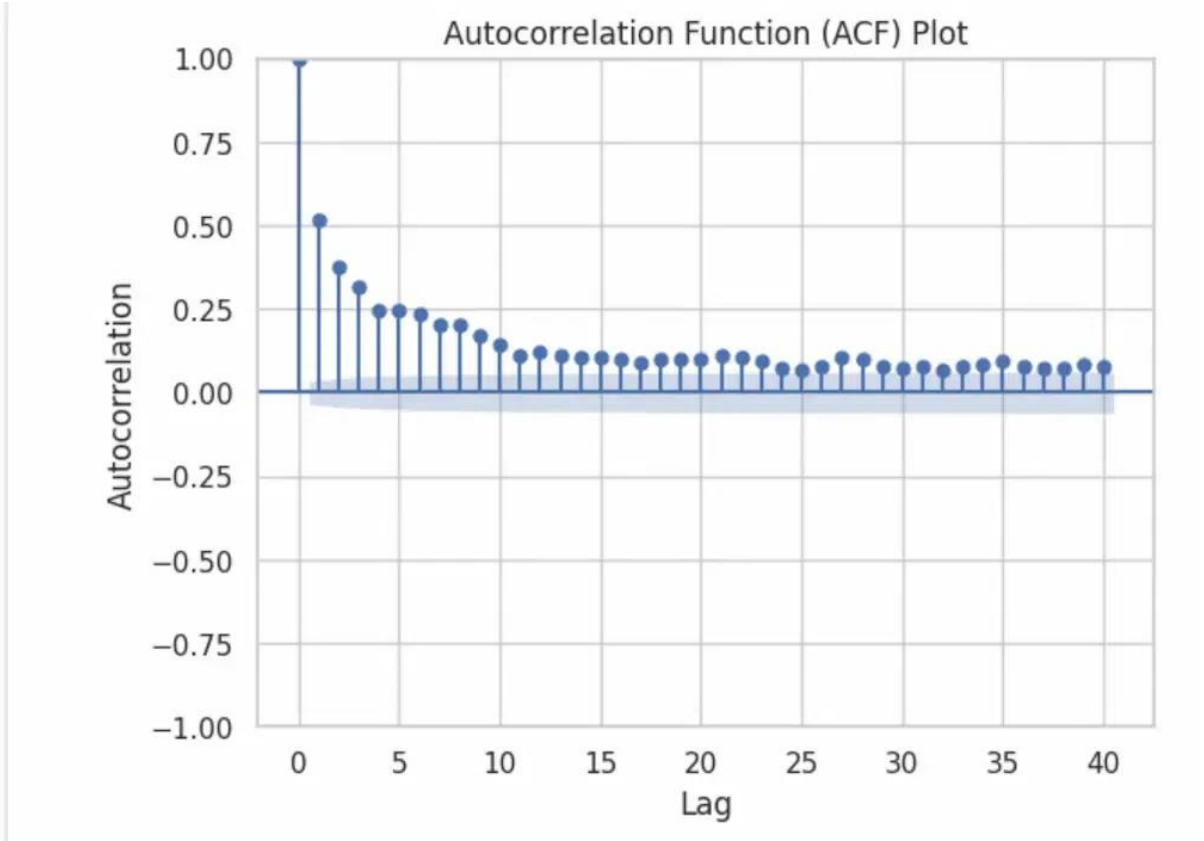
Resampling :

To better understand the trend of the data we will use the resampling method, resampling the data on a monthly basis can provide a clearer view of trends and patterns, especially when we are dealing with daily data.

We have observed an upward trend in the resampled monthly volume data. An upward trend indicates that, over the monthly intervals, the “high” column tends to increase over time.

Detecting Seasonality Using Auto Correlation :

We will detect Seasonality using the autocorrelation function (ACF) plot. Peaks at regular intervals in the ACF plot suggest the presence of seasonality.



The presence of seasonality is typically indicated by peaks or spikes at regular intervals, as there are none there is no seasonality in our data.

Detecting Stationarity :

We will perform the ADF test to formally test for stationarity.

The test is based on the:

- Null hypothesis that a unit root is present in the time series, indicating that the series is non-stationary.
- The alternative hypothesis is that the series is stationary after differencing (i.e., it has no unit root).
- The ADF test employs an augmented regression model that includes lagged differences of the series to determine the presence of a unit root.

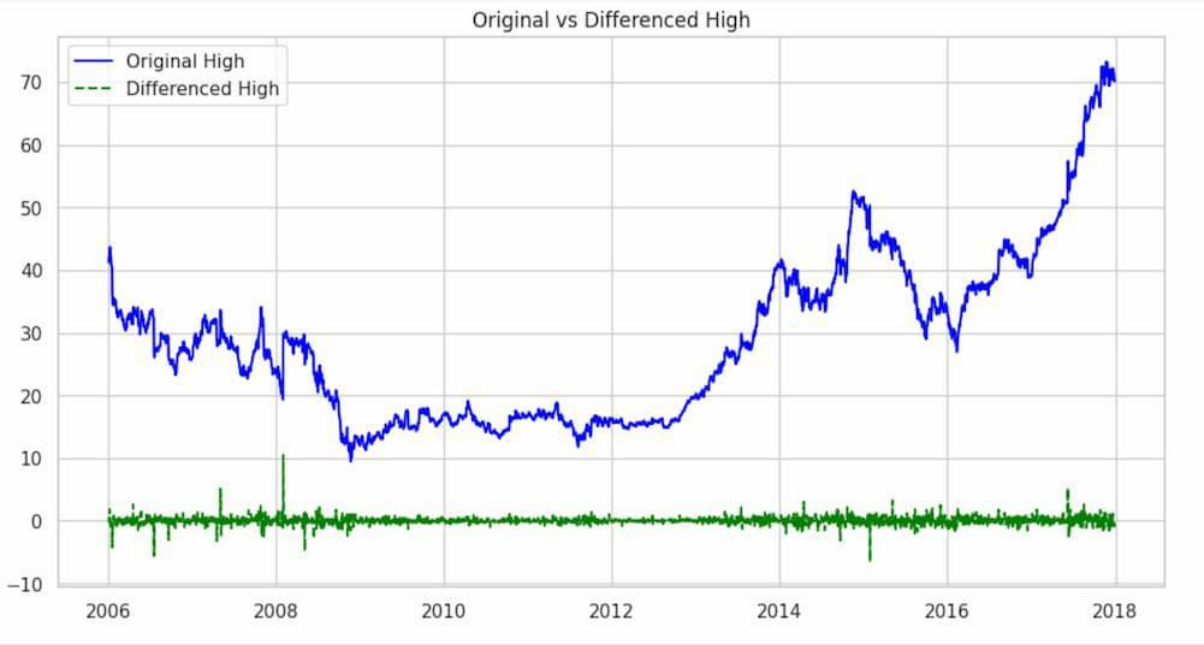
```
ADF Statistic: 0.7671404880535936
p-value: 0.9910868050318213
```

```
Critical Values: {'1%': -3.4325316347197403, '5%': -
2.862503905260741, '10%': -2.5672831121111113}
```

- Based on the ADF Statistici.e > all Critical Values, So, we accept the null hypothesis and conclude that the data does not appear to be stationary according to the Augmented Dickey-Fuller test.
- This suggests that differencing or other transformations may be needed to achieve stationarity before applying certain time series models.

Smoothing the date using Differencing and Moving Average:

Differencing involves subtracting the previous observation from the current observation to remove trends or seasonality.



The df['High'].diff() part calculates the difference between consecutive values in the 'High' column. This differencing operation is commonly used to transform a time series into a new series that represents the changes between consecutive observations.

This calculates the moving average of the ‘High’ column with a window size of 120(A quarter) , creating a smoother curve in the ‘high_smoothed’ series. The plot compares the original ‘High’ values with the smoothed version.