Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Department of Computer Science and Engineering (Data Science)

Subject: Time Series Analysis

Experiment 5

Jhanvi Parekh

60009210033

D11

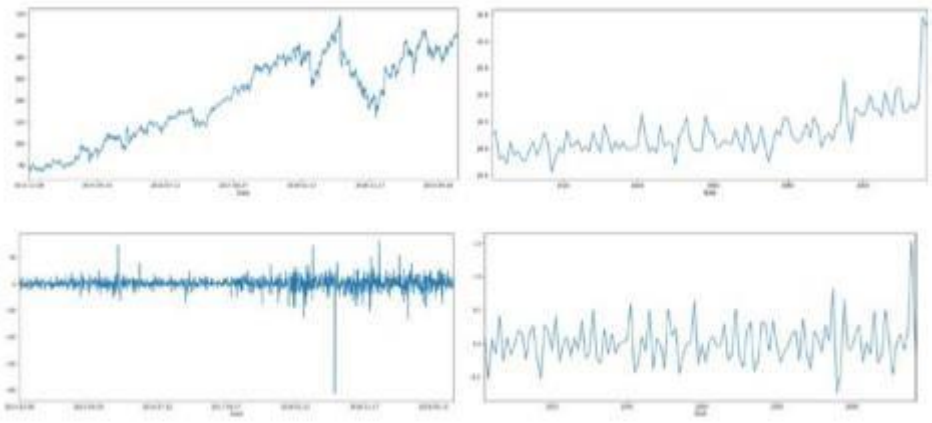
(Stationary – ADF Test)

Aim: Making Data Stationary: Plots
Summary Statistics
Statistics Unit Root Tests
Augmented Dickey – Fuller Test.

Theory:

When time-series data is nonstationary, it means it has trends and seasonality patterns that should be removed. By making data stationary, our data will have a constant mean and variance, and most of the statistical concepts can be applied. To check whether a time series is stationary or not, we can use three types of methods.

- Plots
- Summary statistics
- Statistics tests



Using Plots

The previous graphs show the time-series data in a visual form. This helps us understand whether a time series is linearly increasing or decreasing with time, which is known as a nonstationary time series. A plot that is not increasing or decreasing and shows constant growth over time is known as a stationary time series.

Using Summary Statistics

Summary statistics determine all the essential factors such as mean, standard deviation, variance, min, max, and IQR, which can assist in distinguishing behavior from the data. It helps us to evaluate the seasons or random patterns in the data so we can check for significant differences. Whether the time series is stationary or not, we can split the data in a time series into different groups, where we can compare the difference factors such as mean, variance, and

standard deviation. If the factor that is observed varies and is statistically significant, we say that time series is nonstationary. There are many domains, such as sales, finance, retail, healthcare, etc., where we will find irregular behavior patterns with mean and variance continuously varying because of seasonality and increasing trends.

Using Statistics Unit Root Tests

A statistical test makes a strong assumption about the data. Likewise, we can make some assumptions about time-series data. For this purpose, we have to check that a time-series null hypothesis has been rejected or not. So, we can elucidate the result in order for a particular problem to be meaningful. Statistical unit root tests contain the following five methods, each of which contains a unit root stationary test:

- **Dickey-Fuller (DF) test:** This is based on linear regression. Serial correlation is a big issue of this method.
- **Augmented Dickey-Fuller test:** This solves the serial correlation problem of a DF test and handles big and complex models. This method is widely used.

Lab Assignments to complete:

Perform the following tasks using the datasets mentioned. Download the datasets from the link given:

Link:

https://drive.google.com/drive/folders/1dbqJuZJULas76_Zzkqs-yRd2DbJReJup?usp=sharing

Colab Links:

https://colab.research.google.com/drive/15SfWRxknhLHV3a7Ey1QCY8Nd2Rn_vQG7

Dataset 1: Facebook Stock Market Performance

1. Implement **Augmented Dickey-Fuller test** for unit roots on the dataset and find the following:
    a. Calculate the following
        i. Test Statistic
        ii. p-value
        iii. No Lags Used
        iv. Number of Observations Used
        v. Critical Value 1%, 5%, 10%
    b. Conclude if the time series data contains unit roots or not.

c. Also, infer if the data is stationary or not.
   The dataset fails to reject the null hypothesis because p-value > 0.05 and test statistics value > critical values i.e. the data contains unit roots which makes the data non-stationary.
d. Apply differencing if the data is not stationary.

Link:https://colab.research.google.com/drive/1HhjJ4ipLB9jeb93cLI4fRJ27f-hyorAF?usp=sharing