**5.Program to check stationarity of a time series data**

**Aim:**

To implement a program in Python to check the stationarity of a time series dataset using the Augmented Dickey-Fuller (ADF) Test.

**Procedure:**

**Step 1: Import Libraries**

* Import pandas for data manipulation.
* Import adfuller from statsmodels.tsa.stattools to perform the ADF test.

**Step 2: Load and Inspect the Dataset**

* Load the dataset using pandas.read\_csv().
* Strip any leading or trailing spaces in column names.
* Print column names to understand the structure.

**Step 3: Preprocess the Date Column**

* Convert the Date column to datetime format.
* Set the Date column as the DataFrame index.
* Sort the DataFrame by date to maintain proper chronological order.

**Step 4: Aggregate Global Daily Data**

* Group the data by the index (Date) and sum all numeric columns to get global totals.
* Calculate daily new confirmed cases by applying .diff() on the Confirmed column.

**Step 5: Prepare the Time Series**

* Drop any NaN values caused by the diff() operation.
* Store the cleaned daily new cases as a time series variable (ts).

**Step 6: Perform the ADF Test**

* Use adfuller() to test for stationarity of the daily\_cases time series.
* Capture the ADF test results including the statistic, p-value, used lags, and critical values.

**Step 7: Display and Interpret Results**

* Print the ADF statistic and the corresponding p-value.
* Compare the p-value with 0.05 to determine if the series is stationary:
  + If p-value < 0.05: the series is **stationary** (no unit root).
  + If p-value >= 0.05: the series is **not stationary** (unit root present).

**Code:**

import pandas as pd

from statsmodels.tsa.stattools import adfuller

df = pd.read\_csv("/content/time-series-19-covid-combined.csv")

df.columns = df.columns.str.strip()

print("📌 Column names:", df.columns.tolist())

df["Date"] = pd.to\_datetime(df["Date"], format="%Y-%m-%d")

df.set\_index("Date", inplace=True)

df.sort\_index(inplace=True)

daily\_df = df.groupby(df.index).sum()

daily\_df["daily\_cases"] = daily\_df["Confirmed"].diff()

ts = daily\_df["daily\_cases"].dropna()

adf\_result = adfuller(ts)

print("\n📊 ADF Test Results:")

print(f"ADF Statistic: {adf\_result[0]:.4f}")

print(f"p-value: {adf\_result[1]:.4f}")

print(f"Used lags: {adf\_result[2]}")

print("Critical Values:")

for key, value in adf\_result[4].items():

print(f" {key}: {value:.4f}")

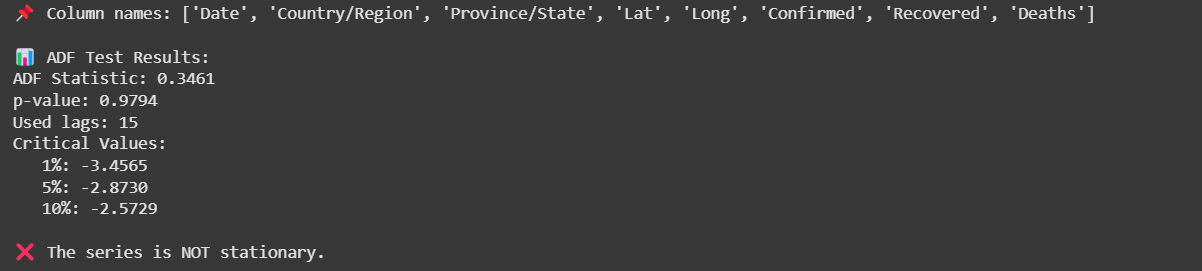
if adf\_result[1] < 0.05:

print("\n✅ The series is stationary.")

else:

print("\n❌ The series is NOT stationary.")

**Output**:



**Result:**

The program to check stationarity of time series data(COVID 19) is successfully implemented and verified