**EX:No.4 Stationarity check 221501010 25/03/25**

**AIM :** To implement linear regression on a time series data .

**PROCEDURE:**

1. Read the time-series data from the CSV file.

2. Convert the date column to datetime format and set it as the index.

3. Visualize the data using a line graph.

4. Apply the Augmented Dickey-Fuller (ADF) test to check stationarity.

5. Print the ADF statistic, p-value, and critical values.

6. Determine if the series is stationary based on the p-value.

**IMPLEMENTATION :**

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression

import matplotlib.pyplot as plt

import pandas as pd

import matplotlib.pyplot as plt

from statsmodels.tsa.stattools import adfuller

# Load the dataset

file\_path = "/content/ch3\_airline\_passengers.csv"

df = pd.read\_csv(file\_path)

# Convert 'Month' to datetime format and set it as the index

df["Month"] = pd.to\_datetime(df["Month"])

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

# Define ADF test function

def adf\_test(series):

    result = adfuller(series)

    print("ADF Statistic:", result[0])

    print("p-value:", result[1])

    print("Critical Values:")

    for key, value in result[4].items():

        print(f"   {key}: {value}")

    if result[1] <= 0.05:

        print("Conclusion: The series is stationary.")

    else:

        print("Conclusion: The series is non-stationary.")

    return result

# Plot the time series data

plt.figure(figsize=(10, 5))

plt.plot(df["AirPassengers"], label="Airline Passengers", color="blue")

plt.title("Airline Passengers Over Time")

plt.xlabel("Year")

plt.ylabel("Number of Passengers")

plt.legend()

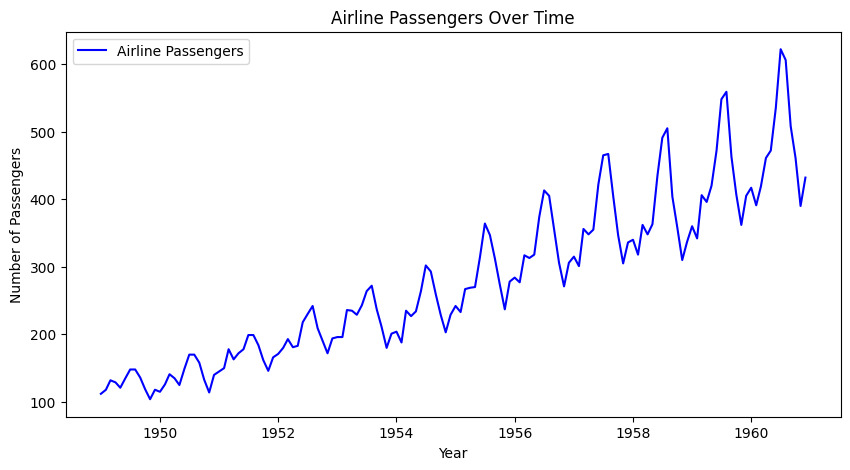
plt.show()

# Perform ADF Test

print("ADF Test on Original Data:")

adf\_test(df["AirPassengers"])

**OUTPUT:**



**RESULT :** Thus linear regression has been implemented on a time series data.