

## EXERCISE NO. 04

### STATIONARITY OF TIME SERIES DATA

#### AIM:

To check for the stationarity of the time series data.

#### PROCEDURE:

1. Import the necessary libraries.

```
import pandas as pd
from statsmodels.tsa.stattools import adfuller
```

2. Load the time series dataset.

```
df = pd.read_csv('/amazon.csv', encoding='latin1')
```

3. Preprocess the data.

```
month_map = {
    'Janeiro': 'January', 'Fevereiro': 'February', 'Março': 'March',
    'Abril': 'April', 'Maio': 'May', 'Junho': 'June',
    'Julho': 'July', 'Agosto': 'August', 'Setembro': 'September',
    'Outubro': 'October', 'Novembro': 'November', 'Dezembro': 'December'
}

df['month'] = df['month'].map(month_map)
df['date'] = pd.to_datetime(df['month'] + ' ' + df['year'].astype(str), format='%B %Y')
df.set_index('date', inplace=True)

df_monthly = df.resample('ME')['number'].sum()
df_yearly = df.resample('YE')['number'].sum()
```

4. Check for the stationarity of the time series data.

```
adf_test = adfuller(df_monthly.dropna())
print("ADF Statistic:", adf_test[0])
print("p-value:", adf_test[1])
print("Critical Values:")
for key, value in adf_test[4].items():
    print(f"\t{key}: {value}")
```

```
ADF Statistic: -2.6841335058482003
p-value: 0.07682433797724766
Critical Values:
    1%: -3.4594900381360034
    5%: -2.8743581895178485
   10%: -2.573601605503697
```

5. Display if the time series data is stationary or not.

```
if adf_test[1] <= 0.05:
    print("The time series is stationary.")
else:
    print("The time series is not stationary.")
```

```
The time series is not stationary.
```

**RESULT:**

Thus the program to check for stationarity of the time series data has been implemented successfully and verified.