EX:No.4 DATE: 22/03/25	Implement programs to check stationary of a time series data

AIM:

To clean, preprocess, and visualize house data, to check the data stationary or non stationary.

ALGORITHM:

Step 1: Read CSV, convert date column, set index, and sort by time.

Step 2:Check if the time series is stationary using the Augmented Dickey-Fuller test.

Step 3:If $p > 0.05$, the series is non-stationary; if $p \leq 0.05$, it's stationary.

Step 4: Compute series.diff().dropna() to remove trends and achieve stationarity.

Step 5:Plot the original non-stationary series and stationary differenced series for comparison.

CODE AND DESCRIPTION:

```
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller

# Load your dataset
df = pd.read_csv("ma_lga_12345.csv")
```

```
# Convert 'saledate' to datetime format
df['saledate'] = pd.to_datetime(df['saledate'], format="%d/%m/%Y")
```

```
# Sort dataset by date
df = df.sort_values('saledate')
```

```
# Aggregate by date (if multiple values exist for a date, take the mean)
df = df.groupby('saledate', as_index=False)['MA'].mean()
```

```
# Set the date as index
df.set_index('saledate', inplace=True)
```

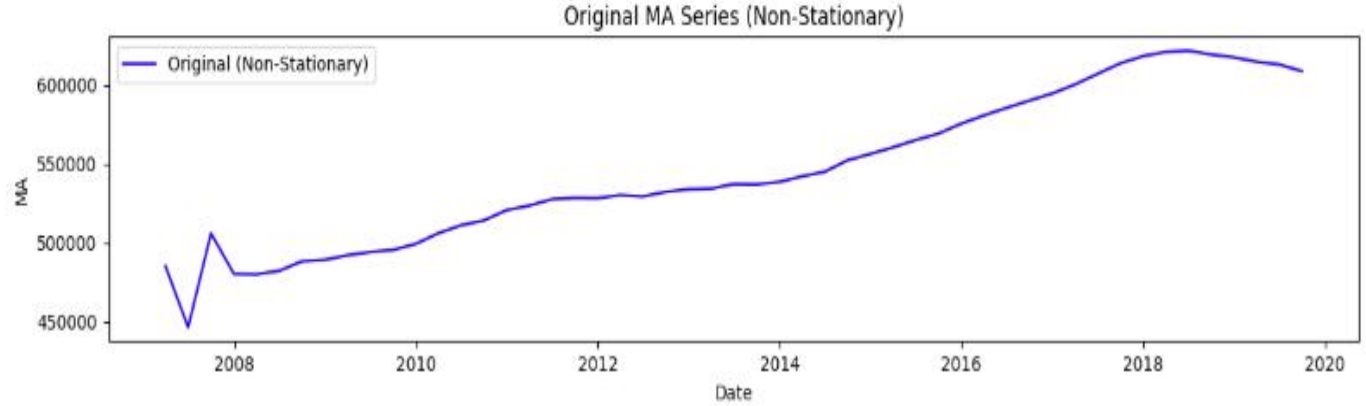
```
# Extract the 'MA' time series
series = df['MA']
```

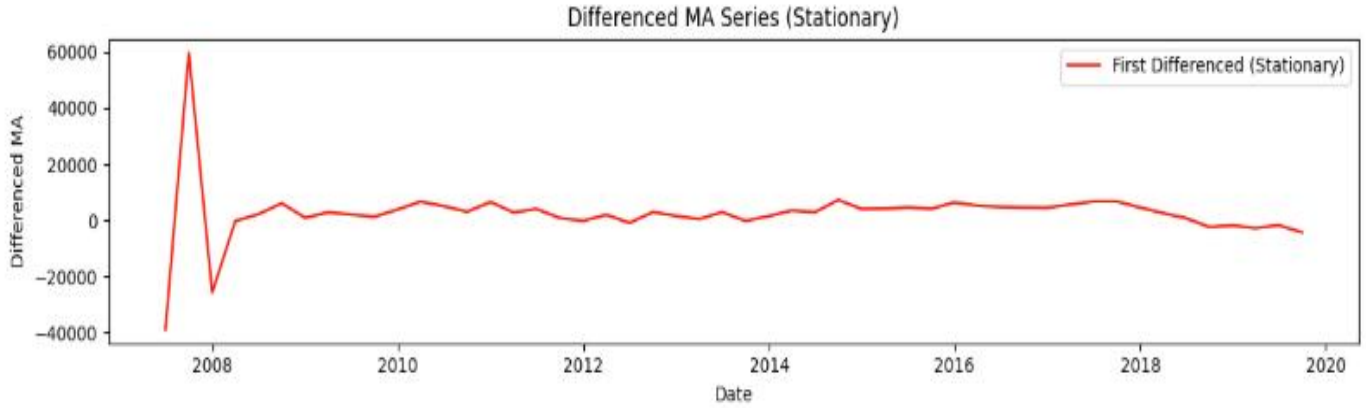
```
# Function to perform ADF test
def adf_test(series, title):
```

```
    """Performs the Augmented Dickey-Fuller (ADF) test and prints results."""
    result = adfuller(series)
    print(f"\n  ADF Test on {title}")
    print(f"ADF Statistic: {result[0]}")
    print(f"p-value: {result[1]}")
    print(f"Critical Values: {result[4]}")

    if result[1] > 0.05:
        print("❌ The series is NON-STATIONARY (Fail to reject H0).")
    else:
        print("✅ The series is STATIONARY (Reject H0).")

# ADF test on original data
adf_test(series, "Original MA Series")

# First differencing to make the data stationary
series_diff = series.diff().dropna()

# ADF test after differencing
adf_test(series_diff, "Differenced MA Series (Stationary)")

# Plot both graphs
plt.figure(figsize=(12, 6))

# Plot Original Series (Non-Stationary)
plt.subplot(2, 1, 1)
plt.plot(series, color='blue', label="Original (Non-Stationary)")
plt.title("Original MA Series (Non-Stationary)")
plt.xlabel("Date")
plt.ylabel("MA")
plt.legend()

# Plot Differenced Series (Stationary)
plt.subplot(2, 1, 2)
plt.plot(series_diff, color='red', label="First Differenced (Stationary)")
plt.title("Differenced MA Series (Stationary)")
plt.xlabel("Date")
plt.ylabel("Differenced MA")
plt.legend()

plt.tight_layout()
plt.show()
```

OUTPUT

```
📌 ADF Test on Differenced MA Series (Stationary)
ADF Statistic: -2.601940841322446
p-value: 0.09259679812708749
Critical Values: {'1%': np.float64(-3.6155091011809297), '5%': np.float64(-2.941262357486514), '10%': np.float64(-2.6091995013850418)}
❌ The series is NON-STATIONARY (Fail to reject H0).
```





RESULT:

Thus, the program has been completed and verified successfully.