

EXERCISE NO 06

MOVING AVERAGE SMOOTHING FOR DATA PREPARATION AND TIME SERIES FORECASTING

AIM:

To prepare data by moving average smoothing and time series forecasting.

ALGORITHM:

1. Load & Preprocess Data: Read dataset, convert dates, and aggregate monthly.
2. Apply moving average: weighted moving average.
3. Visualize Results: Plot original, moving average series.

PROCEDURE:

1. Import the necessary libraries:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

2. Load the time series data:

```
df = pd.read_csv("C:/Users/Lenovo/Downloads/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',
errors='coerce')
df = df.dropna(subset=['date'])
df.set_index('date', inplace=True)

# Aggregate data
df_monthly = df.resample('M')['number'].sum()
df_yearly = df.resample('Y')['number'].sum()
```

4. Moving average implementation:

```
window_size = 5
df_monthly_smooth = df_monthly.rolling(window=window_size, center=True).mean()

# Forecasting future values using Moving Average
forecast_period = 12 # Forecast for the next 12 months
last_values = df_monthly.tail(window_size) # Take last 'window_size' values
moving_avg_forecast = np.mean(last_values) # Average of last 'window_size' values

forecast_dates = pd.date_range(start=df_monthly.index[-1], periods=forecast_period+1,
                                freq='M')[1:]
forecast_values = np.full(forecast_period, moving_avg_forecast)
```

5. Visualization:

```
plt.figure(figsize=(12, 6))
plt.plot(df_monthly.index, df_monthly, label='Original', alpha=0.6)
plt.plot(df_monthly.index, df_monthly_smooth, label=f'Smoothed (Moving Avg, {window_size}-months)',
         linestyle='dashed')
plt.plot(forecast_dates, forecast_values, label=f'Forecast (Next {forecast_period} Months)',
         linestyle='dotted', color='red')
plt.legend()
plt.title('Moving Average Smoothing & Forecasting')
plt.show()
```

6. Output :



