EXERCISE NO 05

ESTIMATING AND ELIMINATING TREND - AGGREGATION SMOOTHING

AIM:
To estimate and eliminate trends in time series dataset by aggregation and smoothing.

PROCEDURE:

1. Import the necessary libraries:

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

2. Load the time series data:

```
# Load dataset
file_path = "/content/amazon.csv"
df = pd.read_csv(file_path, encoding='latin1')
```

3. Pre-process the data:

```
# Map Portuguese month names to English
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.set_index('date', inplace=True)

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

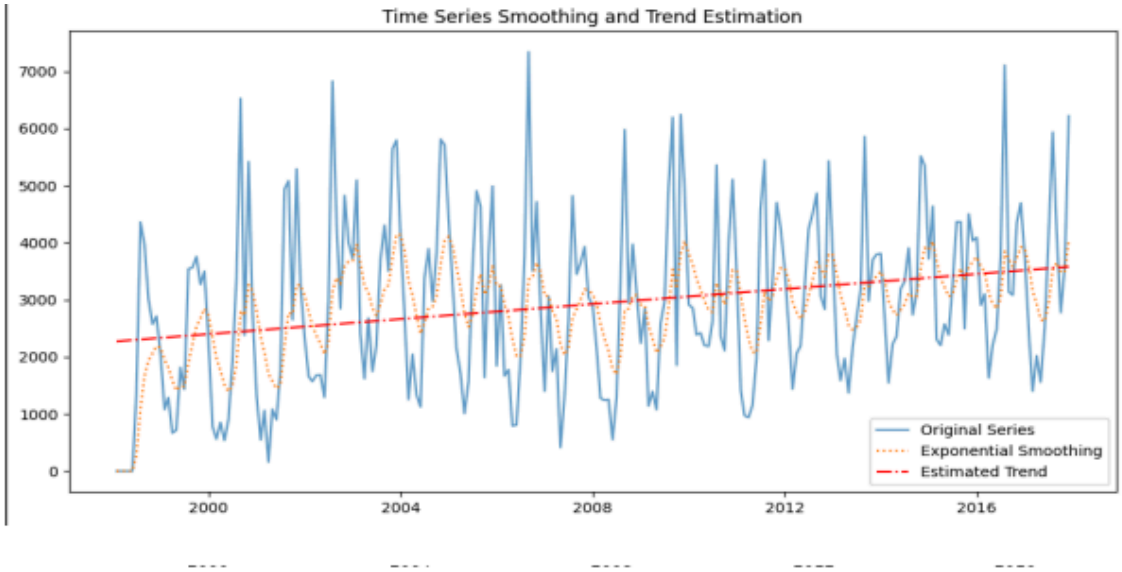
4. Smoothing and Detrending:

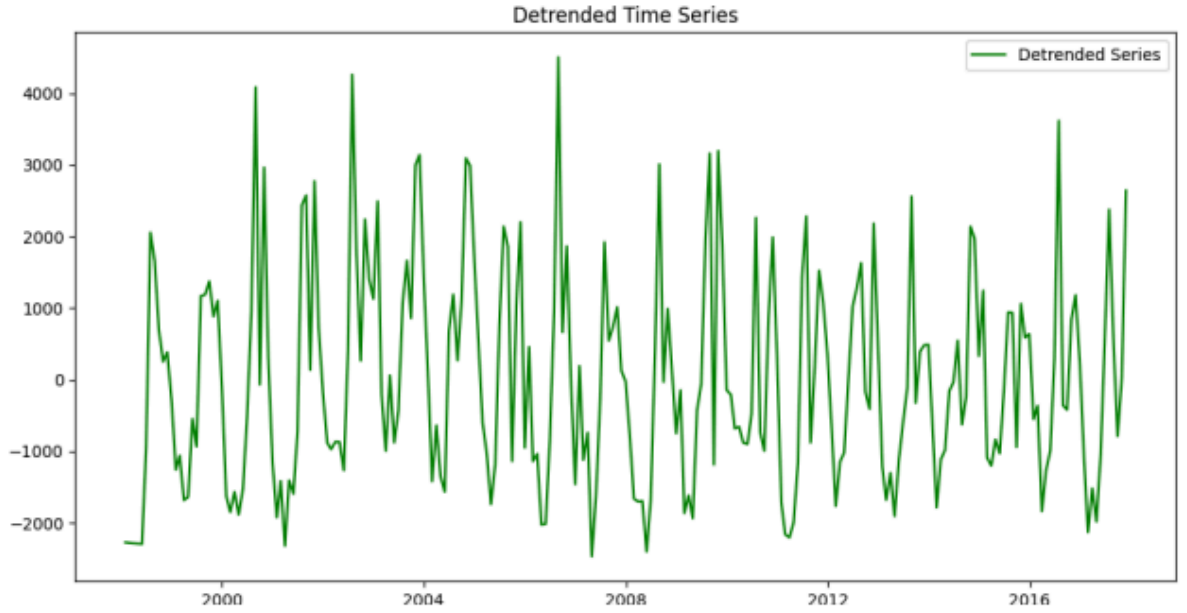
```
# Exponential Smoothing
alpha = 0.2
df_exp = df_monthly.ewm(alpha=alpha, adjust=False).mean()

# Detrending using Linear Regression
X = np.arange(len(df_monthly)).reshape(-1, 1)
y = df_monthly.values.reshape(-1, 1)

model = LinearRegression()
model.fit(X, y)
trend = model.predict(X).flatten()
df_detrended = df_monthly - trend
```

5. Visualization:

```
# Plot results
plt.figure(figsize=(12, 6))
plt.plot(df_monthly, label='Original Series', alpha=0.7)
plt.plot(df_exp, label='Exponential Smoothing', linestyle='dotted')
plt.plot(df_monthly.index, trend, label='Estimated Trend', linestyle='dashdot', color='red')
plt.legend()
plt.title('Time Series Smoothing and Trend Estimation')
plt.show()

plt.figure(figsize=(12, 6))
plt.plot(df_monthly.index, df_detrended, label='Detrended Series', color='green')
plt.legend()
plt.title('Detrended Time Series')
plt.show()
```

Output:


Time Series Smoothing and Trend Estimation
7000
6000
5000
4000
3000
2000
1000
0
2000
2004
2008
2012
2016
Original Series
Exponential Smoothing
Estimated Trend


Detrended Time Series
4000
3000
2000
1000
0
−1000
−2000
2000
2004
2008
2012
2016
Detrended Series