**Exp no: 5 Programs for estimating & eliminating trend in time series data-aggregation, smoothing.**

**Date: 1/4/25**

**Aim:**

The aim of this analysis is to estimate and eliminate the trend component in the AirPassengers time series dataset using aggregation and smoothing techniques. By removing the trend, we can isolate the underlying seasonality and noise, which allows for a better understanding of temporal patterns and further analysis, such as forecasting.

**Objectives:**

The primary objectives of this study include aggregating the dataset to highlight long-term trends, applying a moving average to estimate and smooth the trend component, and eliminating this trend to focus on the residuals. Additionally, the project aims to handle missing values caused by the rolling window and save the detrended data for future use in modeling or further analysis.

**Background/Scope:**

Time series data, such as the AirPassengers dataset, often exhibits both long-term trends and seasonal fluctuations. By applying trend estimation and elimination techniques, we can make the data more suitable for analysis, forecasting, and pattern recognition. This study focuses on aggregating and smoothing the data to remove the trend, with the broader scope of improving accuracy in predictive models and gaining insights into cyclical behaviors.

**Steps for Time Series Sales Data Preprocessing:**

**Step 1: Load the Dataset**

The first step is to load the Air Passenger dataset into a pandas DataFrame. The dataset contains monthly international airline passenger data indexed by date.

import pandas as pd

import matplotlib.pyplot as plt

url = 'https://raw.githubusercontent.com/jbrownlee/Datasets/master/airline-passengers.csv'

data = pd.read\_csv(url, parse\_dates=['Month'], index\_col='Month')

data.columns = ['Passengers']

**Step 2: Visualize the Time Series Data**

Plot the time series data to visualize the trend of airline passengers over time.

import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))

plt.plot(df.index, df['#Passengers'], label='Passengers')

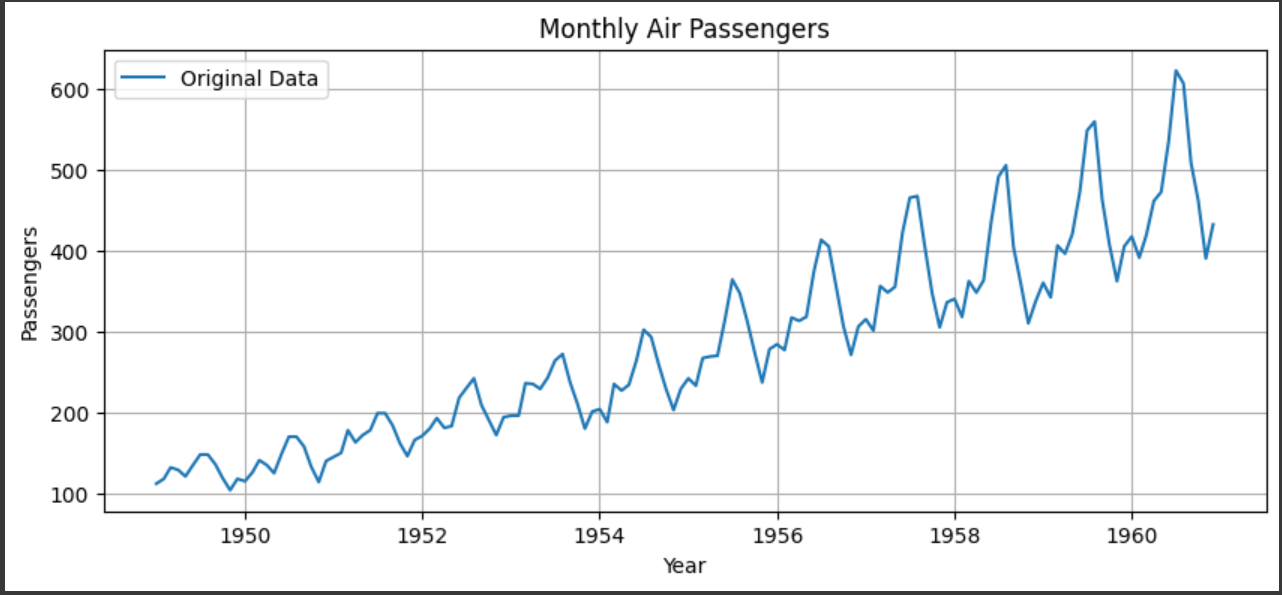
plt.title('Monthly International Airline Passengers (1949-1960)')

plt.xlabel('Date')

plt.ylabel('Number of Passengers')

plt.legend()

plt.show()



**Step 3: Aggregate data (Yearly)**

Yearly aggregation helps smooth short-term noise and makes the long-term trend easier to identify. This is done by computing the average number of passengers each year.

python

yearly\_data = data.resample('A').mean()

# Plot yearly aggregated data

plt.figure(figsize=(10, 4))

plt.plot(yearly\_data, marker='o', label='Yearly Aggregated')

plt.title('Yearly Aggregated Air Passengers')

plt.xlabel('Year')

plt.ylabel('Average Passengers')

plt.legend()

plt.grid(True)

plt.show()

A graph with a line going up

AI-generated content may be incorrect.

**Step 4: Apply moving average smoothing**

A 12-month moving average is used to estimate the trend by smoothing seasonal fluctuations. This helps reveal the overall upward trend in air passenger traffic.

# Step 3: Moving Average (12-month)

data['MA\_12'] = data['Passengers'].rolling(window=12).mean()

# Plot with moving average

plt.figure(figsize=(10, 4))

plt.plot(data['Passengers'], label='Original')

plt.plot(data['MA\_12'], color='red', label='12-Month Moving Average')

plt.title('Trend Estimation using Moving Average')

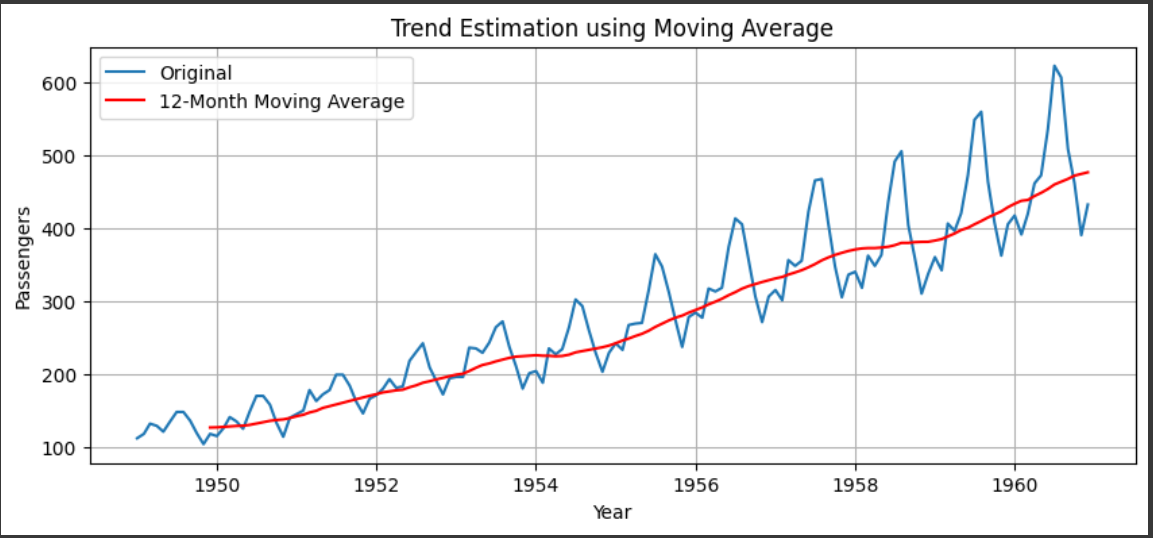
plt.xlabel('Year')

plt.ylabel('Passengers')

plt.legend()

plt.grid(True)

plt.show()



**Step 5: Eliminate trend (Deseasonalize)**

The estimated trend is subtracted from the original series to remove the trend component. The result is a detrended series containing only seasonality and noise.

data['Deseasonalized'] = data['Passengers'] - data['MA\_12']

# Plot deseasonalized data

plt.figure(figsize=(10, 4))

plt.plot(data['Deseasonalized'], label='Deseasonalized Series', color='green')

plt.title('Trend Removed from Time Series')

plt.xlabel('Year')

plt.ylabel('Detrended Passengers')

plt.legend()

plt.grid(True)

plt.show()

A graph showing the growth of a trend

AI-generated content may be incorrect.

**Result:**

Thus, the program to estimate and eliminate the trend in the AirPassengers time series data using aggregation and smoothing techniques has been implemented successfully.