**Implement programs for estimating & eliminating trend in time series data – aggregation, smoothing.**

**EX:No.4**

**DATE:**

# AIM:

To Implement programs for estimating & eliminating trend in time series data – aggregation, smoothing..

## OBJECTIVE:

## To estimate and remove trends in time-series air pollution data using aggregation and smoothing techniques.

## BACKGROUND:

* Aggregation(e.g., monthly/yearly averaging) helps identify patterns.
* Smoothing (e.g., moving average, exponential smoothing) removes fluctuations.
* Trend elimination improves forecasting and stationarity.

## SCOPE OF THE PROGRAM:

* Load and clean air pollution data (2012-2021).
* Apply **aggregation** (monthly/yearly averages) to estimate trends.
* Use **moving average smoothing** to reduce noise.
* Apply **exponential smoothing** to highlight trends

**CODE:**

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_csv("Super store.csv", encoding="latin1")

df.columns = df.columns.str.strip()

df['Order Date'] = pd.to\_datetime(df['Order Date'], errors='coerce')

df.set\_index('Order Date', inplace=True)

df['Sales'] = pd.to\_numeric(df['Sales'], errors='coerce')

df\_monthly = df['Sales'].resample('M').sum()

df\_monthly = df\_monthly[df\_monthly < df\_monthly.quantile(0.99)]

df\_monthly = df\_monthly.to\_frame() # Convert series to DataFrame

df\_monthly['Moving\_Avg'] = df\_monthly['Sales'].rolling(window=6).mean()

df\_monthly['Exp\_Smooth'] = df\_monthly['Sales'].ewm(span=6, adjust=False).mean()

plt.figure(figsize=(12, 6))

plt.plot(df\_monthly['Sales'], label="Original Data", color='blue', alpha=0.5, linewidth=1)

plt.plot(df\_monthly['Moving\_Avg'], label="Moving Avg (6-month)", color='red', linewidth=2)

plt.plot(df\_monthly['Exp\_Smooth'], label="Exponential Smoothing", color='green', linewidth=2)

plt.xlabel("Date", fontsize=12)

plt.ylabel("Sales", fontsize=12)

plt.title("Trend Estimation & Elimination using Smoothing", fontsize=14)

plt.legend()

plt.grid(True, linestyle='--', alpha=0.6)

plt.show()

# OUTPUT:

# 

**RESULT:**

Thus, the program using the time series data implementation has been done successfully.