Pengolahan Data Time Series

Data time series adalah rangkaian observasi yang diukur pada waktu tertentu dan diurutkan berdasarkan urutan waktu. Data jenis ini sering ditemui dalam berbagai bidang, seperti keuangan (harga saham), meteorologi (suhu harian), kesehatan (detak jantung), dan banyak lagi.

# Karakteristik Data Time Series

Data time series memiliki beberapa karakteristik unik yang perlu diperhatikan:

* Ketergantungan Waktu: Nilai-nilai dalam data series paling bergantung satu sama lain berdasarkan urutan waktu.
* Tren (Trend): Pola jangka panjang dalam data, seperti peningkatan atau penurunan nilai rata-rata seiring waktu.
* Seasonalitas (Seasonality): Pola berulang yang terjadi dalam interval waktu tertentu, seperti penjualan yang meningkat setiap akhir tahun.
* Siklus (Cycles): Fluktuasi jangka panjang yang tidak memiliki pola musiman yang tetap.

# Analisis Time Series

Analisis time series bertujuan untuk memahami pola-pola dalam data time series dan menggunakan untuk membuat prediksi atau mengambil keputusan.

## Langkah-langkah Analisis Time Series:

1. Visualisasi Data: Membuat plot time series untuk melihat pola-pola dalam data, seperti tren, seasonality, dan outlier.
2. Decomposisi Time Series: Memecah time series menjadi komponen-komponen: tren, seasonality, dan residu (noise).
3. Peramalan (Forecasting): Memprediksi nilai-nilai masa depan dari time series berdasarkan pola-pola yang teridentifikasi.

# Visualisasi Data Time Series

Visualisasi adalah langkah awal yang penting dalam analisis time series. Beberapa jenis plot yang sering digunakan untuk memvisualisasikan data time series antara lain:

* **Line Plot:** Plot garis sederhana untuk menunjukkan perubahan nilai terhadap waktu.
* **Seasonal Plot:** Plot yang menunjukkan nilai-nilai untuk setiap musim dalam satu tahun.
* **Seasonal Subseries Plot:** Plot yang menunjukkan nilai-nilai untuk setiap musim secara terpisah.
* **Autocorrelation Plot (ACF):** Plot yang menunjukkan korelasi antara suatu titik waktu dengan titik-titik waktu sebelumnya.

## Tugas Hari 13

Anda akan bekerja dengan dataset “AirPassengers” yang tersedia di R. Dataset ini berisi data bulanan jumlah penumpang maskapai penerbangan internasional dari tahun 1949 hingga 1960.

**Tugas:**

1. **Membuat dan Visualisasi Data:**

* Membuat plot time series dari data tersebut. Apakah terdapat adanya tren atau seasonality?
* Explorasi (opsional)

**Tugas:**

1. **Visualisasi:**

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_csv('AirPassengers.csv') # Membaca File

df['Month'] = pd.to\_datetime(df['Month']) # Ubah Kolom 'Month' menjadi tipe data datetime

df.set\_index('Month', inplace=True) # Set kolom 'Month' sebagai index

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

plt.plot(df['#Passengers'])

plt.title('Jumlah Penumpang Maskapai Penerbangan Internasional (1949-1960)')

plt.xlabel('Bulan')

plt.ylabel('Jumlah Penumpang')

plt.show()

1. Tambahan:

from statsmodels.tsa.seasonal import seasonal\_decompose

# Decomposisi time series

result = seasonal\_decompose(df['#Passengers'], model='multiplicative')

# Plot komponen-komponen

result.plot()

plt.show()