EX:No.1 221501048

25/01/26

**1.Program to implement time series data for import library, load data,Pre-prosessing and visualizing**

**Aim:**

Write a program to implement time series data for import library, load data, Preprocessing and visualising.

**Algorithm :**

Step 1: Install required libraries (if not already installed).

Step 2: Import necessary libraries (pandas, numpy, matplotlib).

Step 3: Load crime data, parse dates, and set 'date' as the index.

Step 4: Remove duplicate timestamps and fill missing values.

Step 5: Select the 'crime’' column.

Step 6: Remove outliers using the IQR method.

Step 7: Ensure daily data frequency.

Step 8: Resample to weekly average (optional, not used in the plot).

Step 9: Create a figure and plot daily pollution levels as a line graph.

Step 10: Set labels, title, and legend for the plot.

Step 11: Show the plot.

**Program Code :**

import pandas as pd

import matplotlib.pyplot as plt

file\_path = r"C:\Users\HDC0422270\Downloads\crime-data-1000.csv" # Adjust the path

data = pd.read\_csv(file\_path)

print("Initial Data:")

print(data.head())

print("\nColumns:", data.columns)

data = data.fillna(0)

date\_column = 'Date' # Replace with the actual name of your date column

if date\_column in data.columns:

data[date\_column] = pd.to\_datetime(data[date\_column], errors='coerce')

data = data.dropna(subset=[date\_column])

data = data.sort\_values(by=date\_column)

else:

print(f"Warning: Column '{date\_column}' not found!")

numeric\_columns = [col for col in data.columns if data[col].dtype in ['float64', 'int64']]

print("\nNumeric Columns for Plotting:", numeric\_columns)

if date\_column in data.columns:

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

if numeric\_columns:

plt.plot(data[date\_column], data[numeric\_columns[0]], label=numeric\_columns[0], marker='o')

plt.title(f"Trend of {numeric\_columns[0]} Over Time")

plt.xlabel("Date")

plt.ylabel(numeric\_columns[0])

plt.legend()

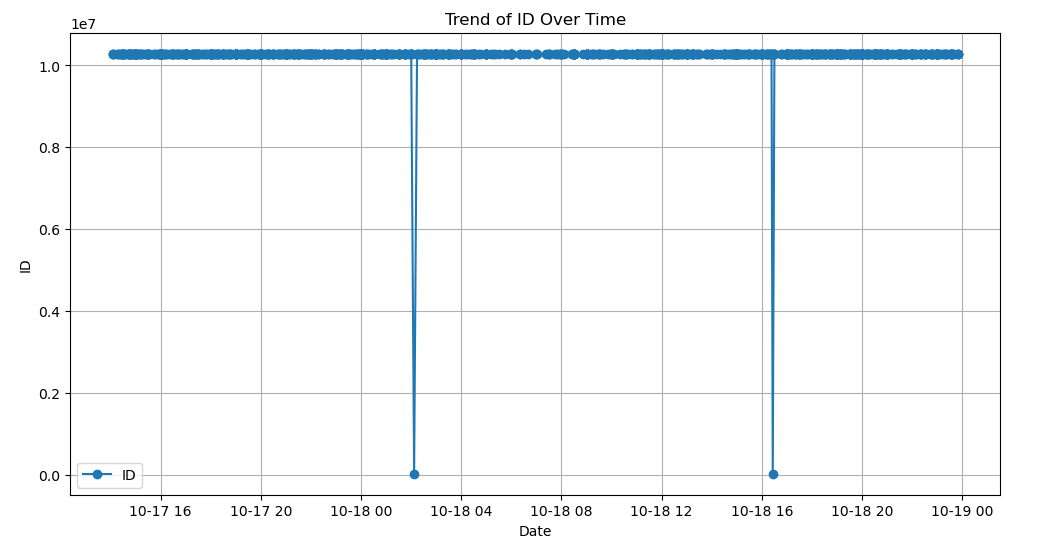
plt.grid()

plt.show()

else:

print("Cannot plot trends without a valid date column!")

**Output :**



**Result:**

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