**Develop A Linear Regression Model For The Time Series Data**

**EX.No:1 DATE: 25/01/2**

**AIM:**

To build a linear regression model for electricity production data, analyze trends over time, and visualize the results.

**ALGORITHM:**

1. Load the electricity production data from the CSV file.
2. Parse the DATE column and convert it to datetime format.
3. Create a numerical time index representing the number of days since the start.
4. Handle missing values by dropping or imputing them if necessary.
5. Fit a linear regression model using the time index as the independent variable and production values as the dependent variable.
6. Generate predictions from the trained model.
7. Plot the actual data points and the fitted linear regression trend to visualize therelationship.

**CODE:**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.linear\_model import LinearRegression

file\_path = "C:\\Users\\Lenovo\\Downloads\\Electric\_Production.csv"

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

print("Initial Data Info:")

print(df.info())

df[df.columns[0]] = pd.to\_datetime(df[df.columns[0]], errors='coerce')

df = df.dropna().drop\_duplicates()

df = df.sort\_values(by=df.columns[0]).reset\_index(drop=True)

df["TimeIndex"] = (df[df.columns[0]] - df[df.columns[0]].min()).dt.days

X = df[["TimeIndex"]]

y = df[df.columns[1]]

model = LinearRegression()

model.fit(X, y)

df["Predicted"] = model.predict(X)

print("\nData Info After Processing:")

print(df.info())

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

plt.scatter(df[df.columns[0]], y, label="Actual Data", color="blue", s=10)

plt.plot(df[df.columns[0]], df["Predicted"], label="Linear Regression Trend", color="red")

plt.xlabel("Year")

plt.ylabel("Electric Production")

plt.title("Linear Regression Model for Time Series Data")

plt.legend()

plt.grid(True)

plt.show()

**OUTPUT:**

**A graph showing a line graph

AI-generated content may be incorrect.**

**RESULT:**

Thus the program has been completed and verified successfully.