**10. Develop vector auto regression model for multivariate time series data forecasting.**

|  |  |
| --- | --- |
| **EX.N0 : 10** | **Develop vector auto regression model for multivariate time series data forecasting.** |
| **DATE : 12/04/2025** |

**AIM:**

To Develop vector auto regression model for multivariate time series data forecasting.

**PROGRAM:**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from statsmodels.tsa.api import VAR

from statsmodels.tsa.stattools import adfuller

# Step 1: Simulate multivariate population dataset with variability

years = np.arange(1990, 2021)

np.random.seed(42)

population = np.linspace(700000, 1300000, len(years)) + np.random.normal(0, 5000, len(years))

birth\_rate = np.linspace(30, 18, len(years)) + np.random.normal(0, 0.3, len(years))

death\_rate = np.linspace(12, 8, len(years)) + np.random.normal(0, 0.2, len(years))

df = pd.DataFrame({

'Year': years,

'Population': population,

'BirthRate': birth\_rate,

'DeathRate': death\_rate

})

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

# Step 2: Make data stationary (Differencing)

def make\_stationary(data):

df\_diff = data.copy()

for col in df\_diff.columns:

result = adfuller(df\_diff[col])

if result[1] > 0.05: # Not stationary

df\_diff[col] = df\_diff[col].diff()

return df\_diff.dropna()

df\_stationary = make\_stationary(df)

# Step 3: Fit VAR model

model = VAR(df\_stationary)

results = model.fit(maxlags=4, ic='aic')

# Step 4: Forecast next 5 years

forecast\_steps = 5

forecast\_input = df\_stationary.values[-results.k\_ar:]

forecast = results.forecast(y=forecast\_input, steps=forecast\_steps)

# Step 5: Prepare forecast DataFrame

forecast\_index = np.arange(df.index[-1] + 1, df.index[-1] + 1 + forecast\_steps)

forecast\_df = pd.DataFrame(forecast, index=forecast\_index, columns=df.columns)

# Step 6: Revert differencing

last\_values = df.iloc[-1]

forecast\_df\_cumsum = forecast\_df.cumsum()

forecast\_df\_final = forecast\_df\_cumsum.add(last\_values)

# Step 7: Plot historical and forecast data

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

for col in df.columns:

plt.plot(df[col], label=f'Historical {col}')

plt.plot(forecast\_df\_final[col], '--', label=f'Forecast {col}')

plt.title("VAR Forecast for Population Dataset")

plt.xlabel("Year")

plt.ylabel("Values")

plt.legend()

plt.grid(True)

plt.tight\_layout()

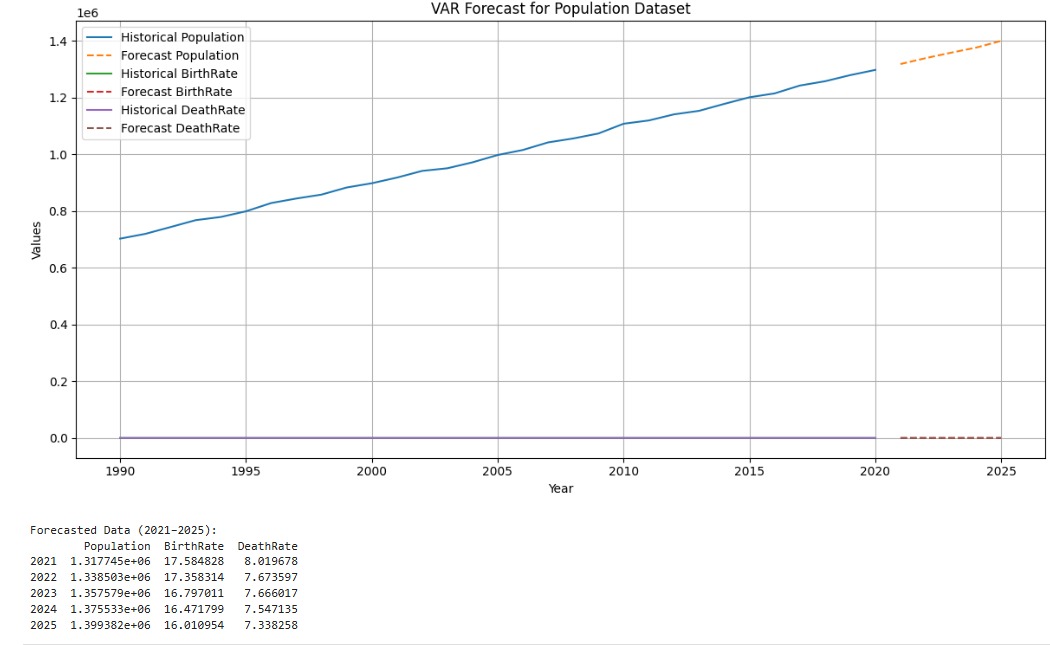
plt.show()

# Step 8: Display forecasted values

print("\nForecasted Data (2021–2025):")

print(forecast\_df\_final)

**OUTPUT:**



**RESULT:**

Thus, Development vector auto regression model for multivariate time series data forecasting.

is executed successfully.