EXP No:8 Create an ARIMA model for time series forecasting

Aim:

To develop an ARIMA model for forecasting future electric production based on historical time series data.

Objectives:

- To load and preprocess monthly electric production data.
- To analyze the data for stationarity and seasonality.
- To determine the optimal ARIMA parameters using automated techniques.
- To build, train, and validate an ARIMA forecasting model.
- To forecast future electric production and visualize the results.

Background:

ARIMA (AutoRegressive Integrated Moving Average) is a widely used statistical model for time series forecasting. It captures autocorrelations in data to make predictions. In this project, ARIMA is applied to monthly electric production data to understand trends and forecast future values. This type of forecasting is valuable in energy management, demand planning, and infrastructure development.

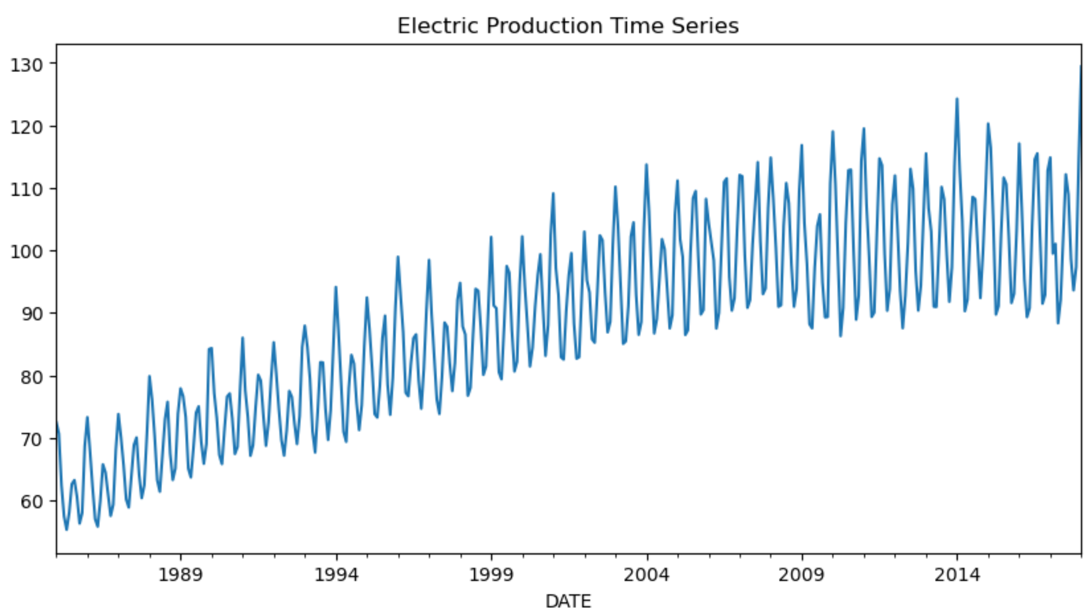
Code:

```
import pandas as pd

import matplotlib.pyplot as plt

from statsmodels.tsa.arima.model import ARIMA

from pmdarima import auto_arima

from datetime import datetime



# Load dataset

df = pd.read_csv(r"C:\Users\Lenovo\Downloads\Electric_Production.csv")

# Convert 'DATE' column to datetime
```

```
df['DATE'] = pd.to_datetime(df['DATE'

# Set 'DATE' as index

df.set_index('DATE', inplace=True)

# Rename the value column

df.rename(columns={'IPG2211A2N': 'Electric_Production'}, inplace=True)

# Plot the original time series

df['Electric_Production'].plot(title='Electric Production Time Series', figsize=(10, 5))

plt.show()

# Use auto_arima to find the best order for ARIMA

stepwise_model = auto_arima(df['Electric_Production'], seasonal=False, trace=True,
                            error_action='ignore', suppress_warnings=True)


# Fit ARIMA model

model = ARIMA(df['Electric_Production'], order=stepwise_model.order)

model_fit = model.fit()


# Summary of the model

print(model_fit.summary())


# Forecast the next 24 months (2 years)

forecast_steps = 24

forecast = model_fit.forecast(steps=forecast_steps)
```

```
# Plot the forecast

plt.figure(figsize=(10, 5))

plt.plot(df['Electric_Production'], label='Historical')

plt.plot(pd.date_range(df.index[-1], periods=forecast_steps+1, freq='MS')[1:], forecast,
label='Forecast', color='red')

plt.title("Electric Production Forecast (ARIMA)")

plt.xlabel("Date")

plt.ylabel("Electric Production")

plt.legend()

plt.tight_layout()

plt.show()
```

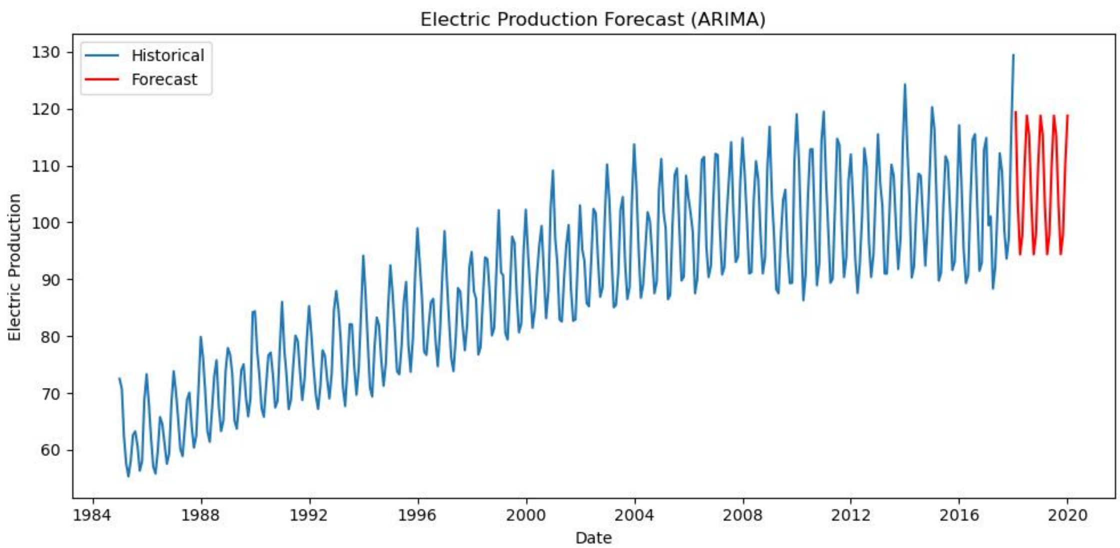
Output:



```
performing stepwise search to minimize aic
ARIMA(2,1,2)(0,0,0)[0] intercept   : AIC=inf, Time=3.39 sec
ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=2750.050, Time=0.03 sec
ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=2691.759, Time=0.04 sec
ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=2624.463, Time=0.05 sec
ARIMA(0,1,0)(0,0,0)[0]             : AIC=2748.185, Time=0.02 sec
ARIMA(1,1,1)(0,0,0)[0] intercept   : AIC=2625.508, Time=0.07 sec
ARIMA(0,1,2)(0,0,0)[0] intercept   : AIC=2525.516, Time=0.11 sec
ARIMA(1,1,2)(0,0,0)[0] intercept   : AIC=2493.118, Time=0.12 sec
ARIMA(1,1,3)(0,0,0)[0] intercept   : AIC=2449.570, Time=0.17 sec
ARIMA(0,1,3)(0,0,0)[0] intercept   : AIC=2449.577, Time=0.14 sec
ARIMA(2,1,3)(0,0,0)[0] intercept   : AIC=inf, Time=0.42 sec
ARIMA(1,1,4)(0,0,0)[0] intercept   : AIC=2431.710, Time=0.26 sec
ARIMA(0,1,4)(0,0,0)[0] intercept   : AIC=2431.461, Time=0.20 sec
ARIMA(0,1,5)(0,0,0)[0] intercept   : AIC=2359.635, Time=0.26 sec
ARIMA(1,1,5)(0,0,0)[0] intercept   : AIC=2349.722, Time=0.39 sec
ARIMA(2,1,5)(0,0,0)[0] intercept   : AIC=2066.370, Time=0.54 sec
ARIMA(2,1,4)(0,0,0)[0] intercept   : AIC=2058.738, Time=0.43 sec
ARIMA(3,1,4)(0,0,0)[0] intercept   : AIC=2081.223, Time=0.52 sec
ARIMA(3,1,3)(0,0,0)[0] intercept   : AIC=inf, Time=0.46 sec
ARIMA(3,1,5)(0,0,0)[0] intercept   : AIC=2060.408, Time=0.58 sec
ARIMA(2,1,4)(0,0,0)[0]             : AIC=2071.225, Time=0.40 sec

Best model:  ARIMA(2,1,4)(0,0,0)[0] intercept
Total fit time: 8.767 seconds
```

```
                               SARIMAX Results
==============================================================================
Dep. Variable:     Electric_Production   No. Observations:                  397
Model:                  ARIMA(2, 1, 4)   Log Likelihood               -1028.613
Date:                 Tue, 15 Apr 2025   AIC                           2071.225
Time:                         11:02:20   BIC                           2099.095
Sample:                     01-01-1985   HQIC                          2082.267
                          - 01-01-2018
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.9986      0.002    522.216      0.000       0.995       1.002
ar.L2         -0.9994      0.001   -859.877      0.000      -1.002      -0.997
ma.L1         -1.3147      0.041    -32.321      0.000      -1.394      -1.235
ma.L2          0.9208      0.074     12.480      0.000       0.776       1.065
ma.L3          0.0158      0.074      0.213      0.831      -0.130       0.161
ma.L4         -0.3899      0.047     -8.328      0.000      -0.482      -0.298
sigma2        10.3486      0.634     16.318      0.000       9.106      11.592
==================================================================================
```

```
Ljung-Box (L1) (Q):                   0.24   Jarque-Bera (JB):               21.25
Prob(Q):                              0.62   Prob(JB):                        0.00
Heteroskedasticity (H):               1.83   Skew:                            0.31
Prob(H) (two-sided):                  0.00   Kurtosis:                        3.95
===================================================================================
```



Result:

Thus the program to create an ARIMA model for time series forecasting is implemented successfully.