**Create an ARIMA Model for time series forecasting**

**EX:No.8**

**DATE: 08/03/25**

# AIM:

To implement a program for creating an ARIMA Model for time series forecasting.

## ALGORITHM:

* Import required libraries (pandas, matplotlib, statsmodels)
* Load the cleaned office supply sales dataset
* Convert the date column to datetime and set it as the index
* Resample the daily sales data to get monthly average sales
* Drop missing values (if any)
* Split the data into training and testing sets (80% train, 20% test)
* Fit an ARIMA model on the training data with specified order
* Forecast sales for the test period using the fitted model
* Plot and compare the actual vs forecasted sales data

**CODE:**

import pandas as pd

import matplotlib.pyplot as plt

from statsmodels.tsa.arima.model import ARIMA

from pandas.plotting import register\_matplotlib\_converters

register\_matplotlib\_converters()

# Load the cleaned sales dataset

df = pd.read\_csv("/content/cleaned\_sales\_data.csv")

# Convert 'Order Date' to datetime and set as index

df['Order Date'] = pd.to\_datetime(df['Order Date'], errors='coerce')

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

# Aggregate daily sales, then resample to monthly average

df\_daily = df['Sales'].resample('D').sum()

df\_monthly = df\_daily.resample('M').mean()

# Drop NA values (if any)

df\_monthly.dropna(inplace=True)

# Split into training and test sets (80% training, 20% testing)

split\_idx = int(len(df\_monthly) \* 0.8)

train = df\_monthly[:split\_idx]

test = df\_monthly[split\_idx:]

# Fit the ARIMA model (simple order to start with)

model = ARIMA(train, order=(1, 1, 1))

model\_fit = model.fit()

# Forecast for the length of the test set

forecast = model\_fit.forecast(steps=len(test))

# Plot training, actual test, and forecasted values

plt.figure(figsize=(14, 6))

plt.plot(train.index, train, label='Training Data')

plt.plot(test.index, test, label='Actual Sales')

plt.plot(test.index, forecast, label='Forecasted Sales', linestyle='--', color='orange')

plt.title("Monthly Office Supply Sales Forecast using ARIMA")

plt.xlabel("Date")

plt.ylabel("Sales")

plt.legend()

plt.grid(True)

plt.tight\_layout()

plt.show())

# OUTPUT:

# **Screenshot 2025-04-30 055410**

**RESULT:**

Thus, the program is implemented and verified successfully.