**Exp no: 2 Implement programs for visualizing time series data**

**Date: 31/1/25**

**Objectives:**

The primary objective of this analysis is to visualize and analyze the time series data of airline passengers from 1949 to 1960. By plotting the data and performing seasonal decomposition, we aim to identify trends, seasonal patterns, and residuals in the dataset. This will help in understanding the underlying structure of the data and provide insights for further time series forecasting or modeling.

**Background/Scope:**

The Air Passengers dataset is a classic time series dataset that records the number of international airline passengers each month from 1949 to 1960. It is widely used for time series analysis and forecasting due to its clear trend and seasonal patterns. This analysis focuses on visualizing the data, identifying trends, and decomposing the time series into its components (trend, seasonality, and residuals) to gain a deeper understanding of the dataset.

### ****Steps for Time Series Sales Data Preprocessing****:

**Step 1: Load the Dataset**

Load the dataset from a local CSV file and display the first few rows to understand its structure.

import pandas as pd

# Load the dataset

file\_path = "contents/AirPassengers.csv" # Replace with the actual path to your CSV file

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

# Display the first few rows

print(data.head())

**Step 2: Preprocess the Data**

Convert the Month column to a datetime object and set it as the index for time series analysis.

# Convert 'Month' to datetime and set as index

data['Month'] = pd.to\_datetime(data['Month'])

data.set\_index('Month', inplace=True)

# Display the updated dataset

print(data.head())

**Step 3: Visualize the Time Series Data**

Plot the time series data to visualize the trend of airline passengers over time.

import matplotlib.pyplot as plt

# Plot the time series data

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

plt.plot(data.index, data['Passengers'], color='blue', marker='o', linestyle='-')

plt.title('Air Passengers Over Time (1949-1960)')

plt.xlabel('Month')

plt.ylabel('Number of Passengers')

plt.grid(True)

plt.show()

A graph with blue lines

Description automatically generated

**Step 4: Perform Trend Decomposition**

Decompose the time series into its trend using statsmodels.

from statsmodels.tsa.seasonal import seasonal\_decompose

# Perform seasonal decomposition

decomposition = seasonal\_decompose(data['#Passengers'], model='multiplicative')

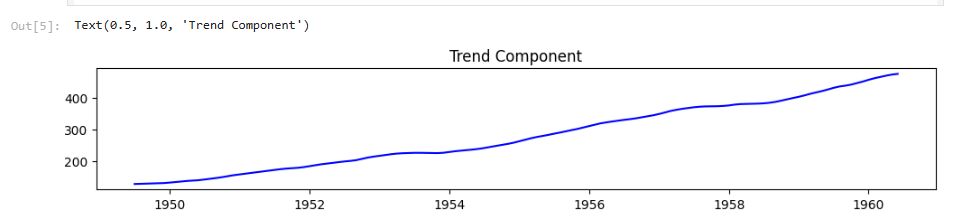
# Plot the decomposition

plt.figure(figsize=(12, 8))

plt.subplot(4, 1, 1)

plt.plot(decomposition.trend, color='blue')

plt.title('Trend Component')



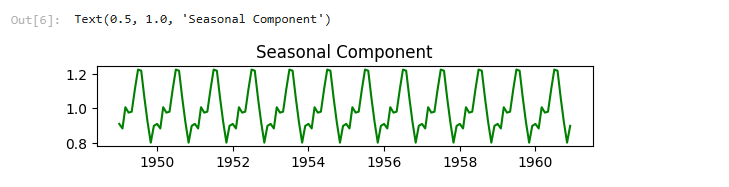
plt.subplot(4, 1, 2)

plt.plot(decomposition.seasonal, color='green')

plt.title('Seasonal Component')

**Step 5: Perform Seasonal Decomposition**

Decompose the time series into its trend using statsmodels.



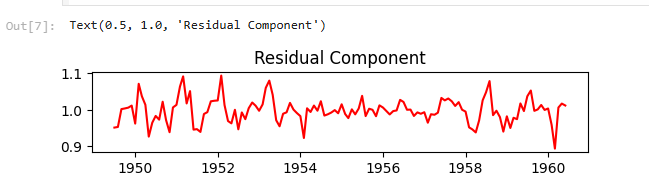
plt.subplot(4, 1, 3)

plt.plot(decomposition.resid, color='red')

plt.title('Residual Component')

**Step 6: Perform Residual Decomposition**

Decompose the time series into its trend using statsmodels.



plt.subplot(4, 1, 4)

plt.plot(data['#Passengers'], color='purple')

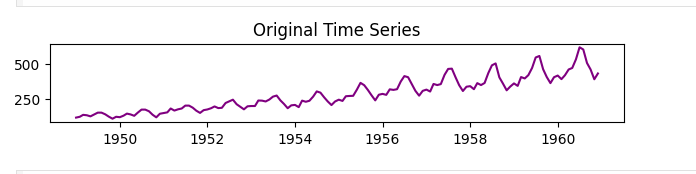
plt.title('Original Time Series')

plt.tight\_layout()

plt.show()

**Step 7: Original Time series Data**

Original Time series Data using statsmodels.



**Step 8: Box plot to show seasonality**

Extract the year and month from the index

import seaborn as sns

data['Year'] = data.index.year

data['Month'] = data.index.month

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

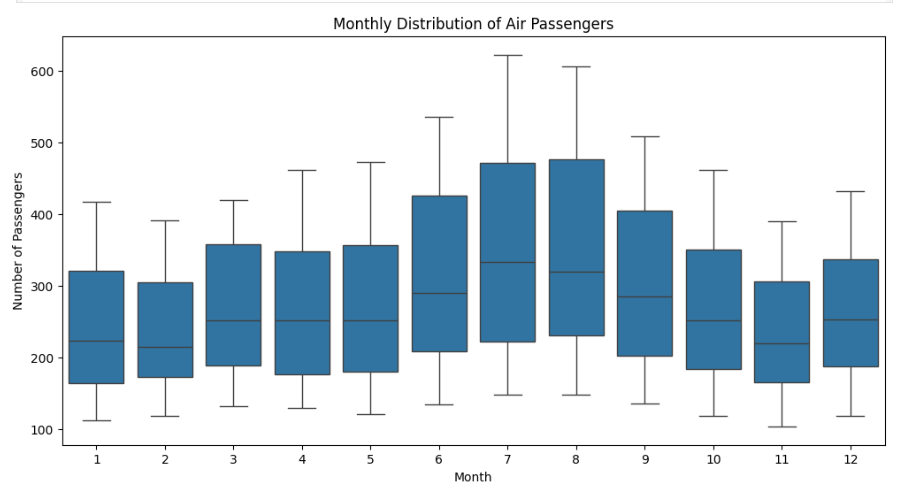
sns.boxplot(x='Month', y='#Passengers', data=data)

plt.title('Monthly Distribution of Air Passengers')

plt.xlabel('Month')

plt.ylabel('Number of Passengers')

plt.show()



**Result:**

Thus the time series dataset is visualized successfully.