**Implement program for visualizing time series data**

**AIM:**

To implement program for visualizing time series data using jupyter notebook.

**ALGORITHM:**

1. Load the taxi dataset then clean and load the values

2. Preprocess the values such as variables in the dataset

3. Plot the graph and visualize the values

**CODE:**

**import** pandas **as** pd

**import** seaborn **as** sns

**import** matplotlib.pyplot **as** plt

**import** numpy **as** np

**from** statsmodels.graphics.tsaplots **import** plot\_acf

**from** statsmodels.tsa.seasonal **import** seasonal\_decompose

file\_path **=** "Downloads/train.csv/taxi dataset.csv"

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

print(df**.**head())

*# Convert to datetime*

df['datetime'] **=** pd**.**to\_datetime(df[['year', 'month', 'day']]) **+** pd**.**to\_timedelta(df['hour\_of\_day'], unit**=**'h')

*# Set index*

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

*# Resample to daily total fares*

df\_daily **=** df[['total\_amount']]**.**resample('D')**.**sum()

*## 1. Heatmap (Fare Trends by Hour and Day)*

df['day\_of\_week'] **=** df**.**index**.**dayofweek *# Monday=0, Sunday=6*

df['hour'] **=** df**.**index**.**hour

pivot\_table **=** df**.**pivot\_table(values**=**'total\_amount', index**=**'day\_of\_week', columns**=**'hour', aggfunc**=**np**.**sum)

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

sns**.**heatmap(pivot\_table, cmap**=**"coolwarm", annot**=False**)

plt**.**xlabel("Hour of Day")

plt**.**ylabel("Day of Week")

plt**.**title("Heatmap: Total Fare Amount by Hour and Day")

plt**.**show()

*## 2. Histogram (Fare Distribution)*

plt**.**figure(figsize**=**(10, 6))

sns**.**histplot(df['total\_amount'], bins**=**50, kde**=True**, color**=**'blue')

plt**.**xlabel("Total Fare Amount")

plt**.**ylabel("Frequency")

plt**.**title("Histogram of Total Fare Amounts")

plt**.**grid()

plt**.**show()

*# 3. Hourly Fare Distribution (Boxplot)*

plt**.**figure(figsize**=**(10, 6))

sns**.**boxplot(x**=**df['hour\_of\_day'], y**=**df['total\_amount'])

plt**.**xlabel("Hour of Day")

plt**.**ylabel("Total Fare Amount")

plt**.**title("Fare Amount Distribution by Hour of Day")

plt**.**grid()

plt**.**show()

*## 4. Autocorrelation Plot (ACF)*

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

plot\_acf(df\_daily['total\_amount']**.**dropna(), lags**=**30)

plt**.**title("Autocorrelation Plot of Total Fare Amount")

plt**.**show()

*## 5. Scatter Plot (Fare vs. Hour)*

plt**.**figure(figsize**=**(10, 6))

sns**.**scatterplot(x**=**df['hour\_of\_day'], y**=**df['total\_amount'], alpha**=**0.5)

plt**.**xlabel("Hour of Day")

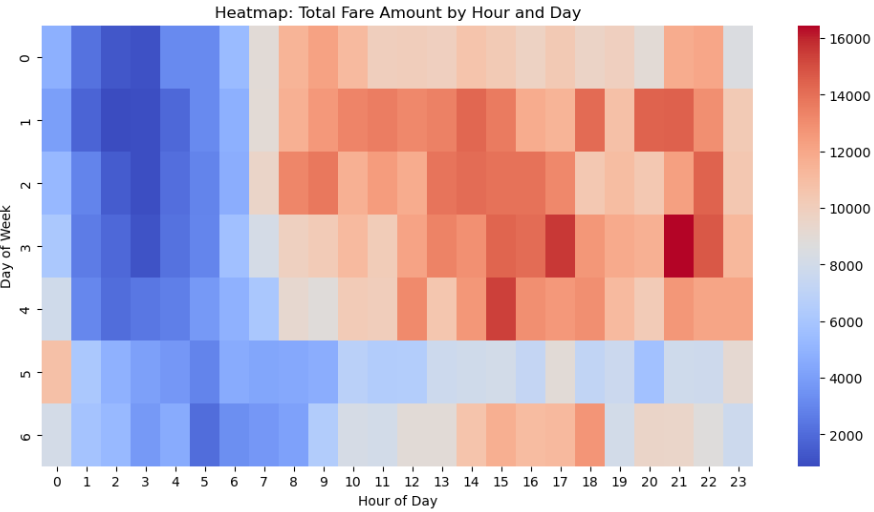
plt**.**ylabel("Total Fare Amount")

plt**.**title("Scatter Plot: Total Fare Amount vs. Hour of Day")

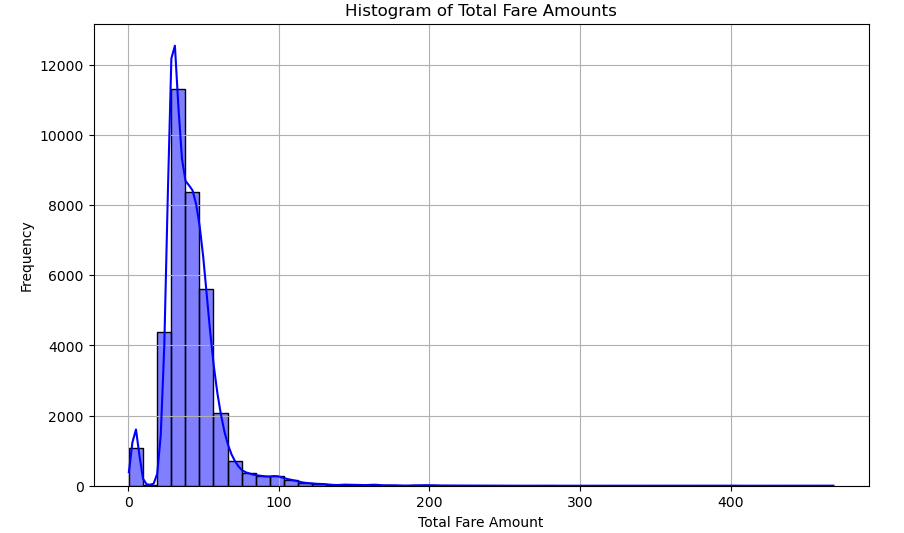
plt**.**grid()

plt**.**show()

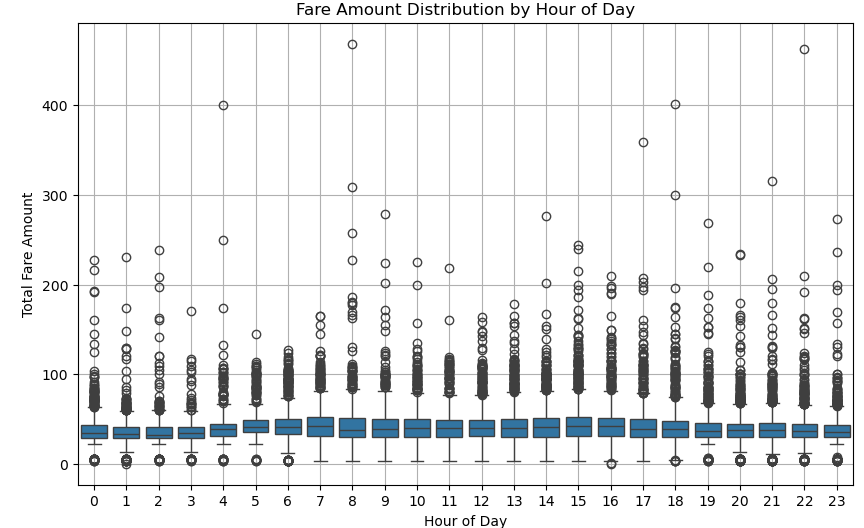
**OUTPUT:**

*## 1. Heatmap (Fare Trends by Hour and Day)* ****

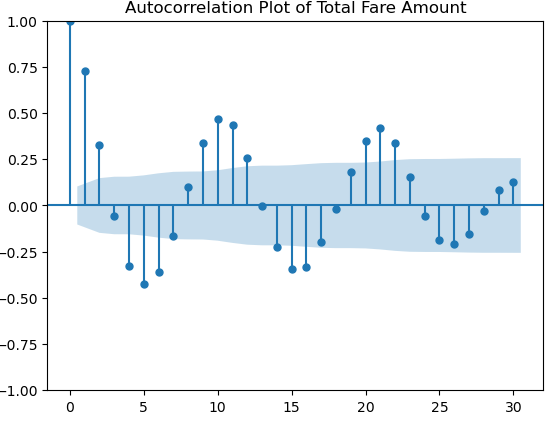
*## 2. Histogram (Fare Distribution)*

****

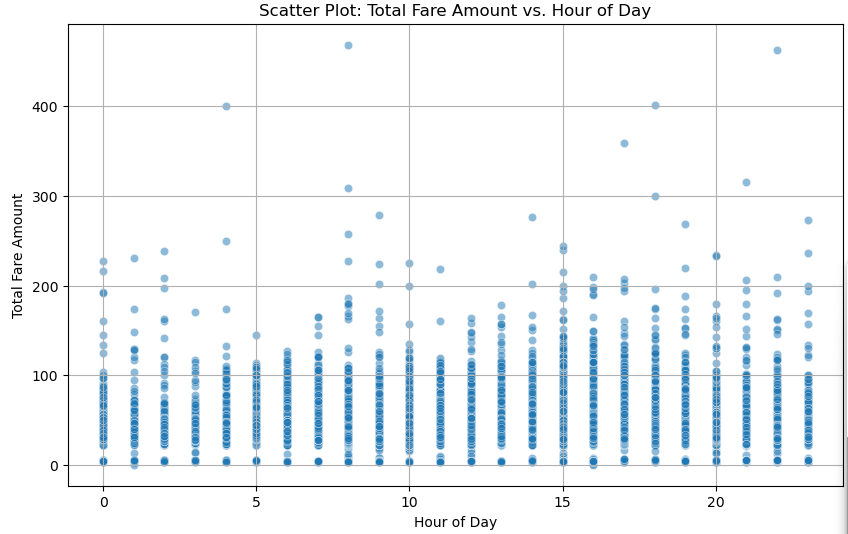
*# 3. Hourly Fare Distribution (Boxplot)*

****

*## 4. Autocorrelation Plot (ACF)*

****

*## 5. Scatter Plot (Fare vs. Hour)*



**RESULT:**

The program to execute visualizing time series data completed successfully and the output is verified.