

EX - 5

06/02/2025

Develop a linear regression model for forecasting time series data

AIM :

To implement a linear regression model for forecasting time series data.

Procedure and Code :

Step 1: Import necessary Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error
```

Step 2 : Read the Dataset

```
file_path = "/content/drive/MyDrive/TimeSereisDatasets/Ex-5/Copy of daily-website-visitors.csv"
df = pd.read_csv(file_path)
```

```
print("First 5 rows of the dataset:")
print(df.head())
```

```
print("\nColumn names in dataset:", df.columns)
```

```
df.columns = df.columns.str.strip()
```

Step 3 : Describe and Process the Data

```
expected_columns = ["Date", "Unique.Visits"]
for col in expected_columns:
    if col not in df.columns:
```

EX - 5

06/02/2025

Develop a linear regression model for forecasting time series data

```
raise ValueError(f"Error: Expected column '{col}' not found. Available columns: {df.columns}")
```

```
df['Date'] = pd.to_datetime(df['Date'])  
df.sort_values('Date', inplace=True)
```

```
df['Unique.Visits'] = df['Unique.Visits'].str.replace(',', '').astype(int)
```

```
df['Days'] = (df['Date'] - df['Date'].min()).dt.days
```

```
X = df[['Days']]  
y = df['Unique.Visits']
```

Step 4: Linear Regression Model

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42,  
shuffle=False)
```

```
model = LinearRegression()  
model.fit(X_train, y_train)
```

```
y_pred = model.predict(X_test)
```

```
mae = mean_absolute_error(y_test, y_pred)  
print(f"\nMean Absolute Error (MAE): {mae:.2f}")
```

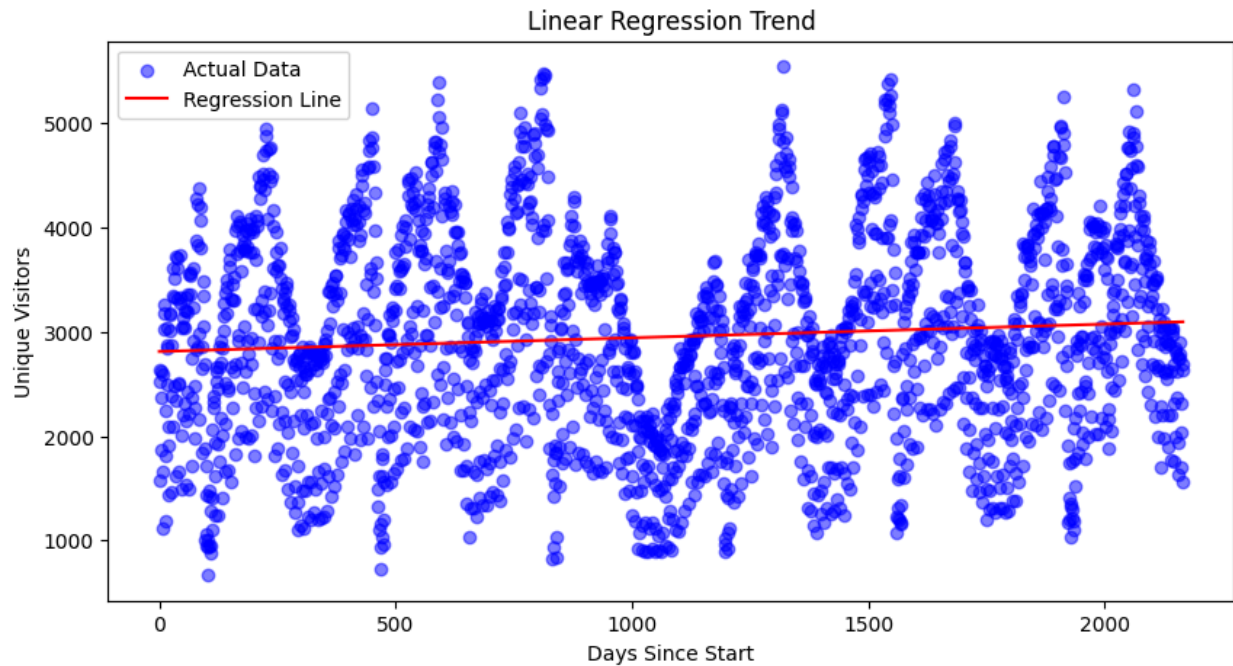
```
plt.figure(figsize=(10, 5))  
plt.scatter(df['Days'], df['Unique.Visits'], color='blue', label="Actual Data", alpha=0.5)  
plt.plot(X, model.predict(X), color='red', label="Regression Line")  
plt.xlabel("Days Since Start")
```

EX - 5

06/02/2025

Develop a linear regression model for forecasting time series data

```
plt.ylabel("Unique Visitors")
plt.title("Linear Regression Trend")
plt.legend()
plt.show()
```



Step 5 : Scattered Plot Visualization

```
print("\nForecasted Visitor Counts for Next 5 Days:")
print(forecast_df.head())
```

```
plt.figure(figsize=(10, 5))
plt.scatter(df['Date'], df['Unique.Visits'], color='blue', label="Actual Data", alpha=0.5)
plt.plot(future_dates, future_predictions, color='green', linestyle='dashed', label="Forecasted Data")
plt.xlabel("Date")
plt.ylabel("Unique Visitors")
plt.title("Time Series Forecasting")
plt.legend()
```

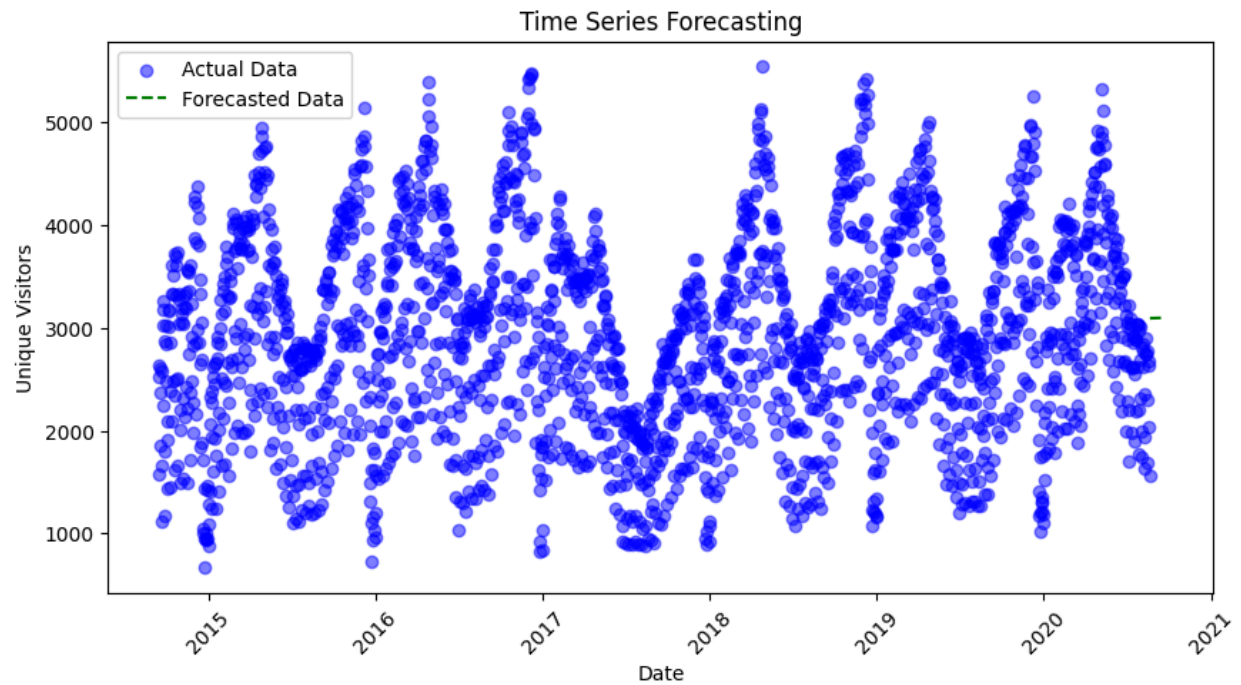
EX - 5

06/02/2025

Develop a linear regression model for forecasting time series data

```
plt.xticks(rotation=45)
```

```
plt.show()
```



Conclusion :

Thus the Linear Regression on Time series data has been executed successfully.