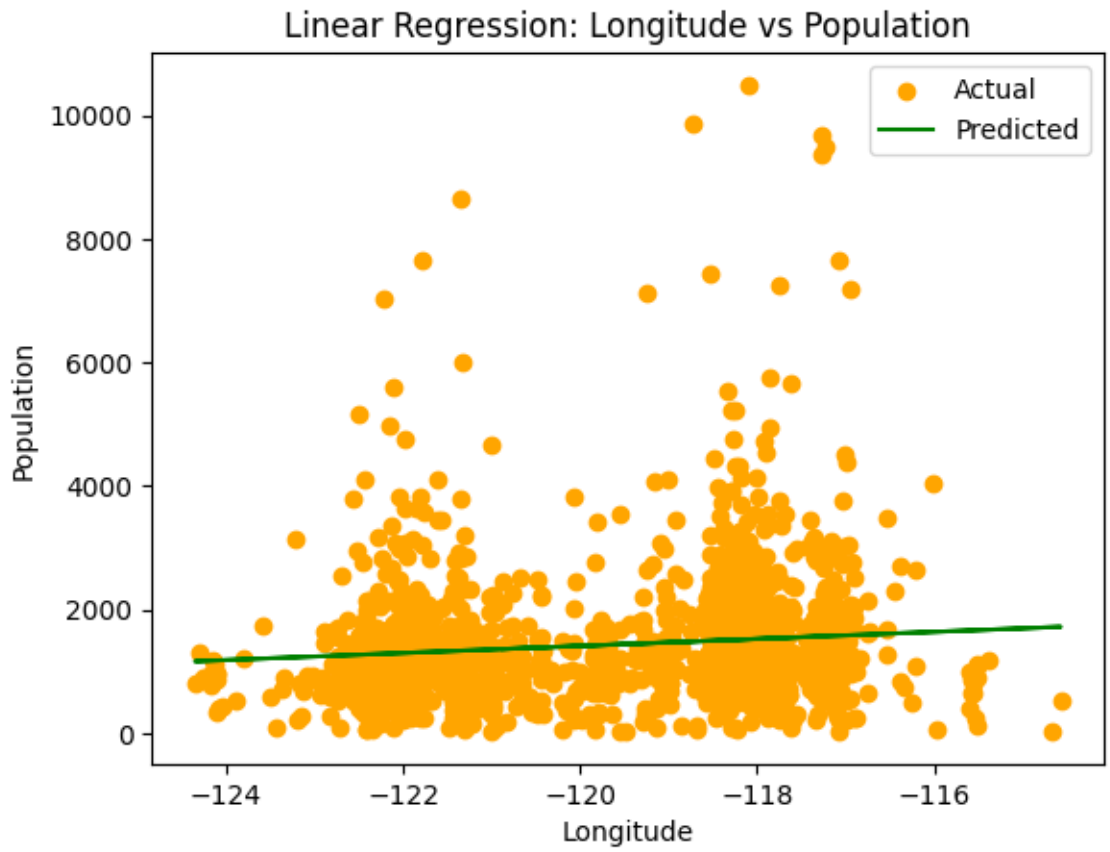
**EXP NO:1 a** **LINEAR REGRESSION**

**DATE:30/07/2024**

**PROGRAM:**

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn import linear_model
df=pd.read_csv('/content/sample_data/california_housing_train.csv')
df.dropna(inplace=True)
xpoints = df["longitude"].values.reshape(-1, 1)
ypoints = df["population"].values
x_train, x_test, y_train, y_test = train_test_split(xpoints, ypoints, test_size=0.1, random_state=42)
reg = linear_model.LinearRegression()
reg.fit(x_train, y_train)
ypoints_pred = reg.predict(x_test)
plt.scatter(x_test, y_test, color="orange", label="Actual")
plt.plot(x_test, ypoints_pred, color="green", label="Predicted")
plt.xlabel("Longitude")
plt.ylabel("Population")
plt.title("Linear Regression: Longitude vs Population")
plt.legend()
plt.show()
```

**OUTPUT:**




**RESULT:**

Thus the program to implement linear regression has been implemented successfully.