

```

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns

# df = pd.read_csv('/content/User_Data (1).csv')
df = pd.read_csv('User_Data.csv')

df.isnull().sum()

    User ID      0
    Gender      0
    Age         0
    EstimatedSalary  0
    Purchased    0
    dtype: int64

X = df.iloc[:, [2,3]].values
Y = df.iloc[:, 4].values

from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)

from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()

X_train = sc_X.fit_transform(X_train)

X_test = sc_X.transform(X_test)

from sklearn.linear_model import LogisticRegression
classifier = LogisticRegression(random_state = 0)
classifier.fit(X_train, Y_train)

Y_Pred = classifier.predict(X_train)
Y_Pred

from sklearn.metrics import confusion_matrix
cm = confusion_matrix(Y_train, Y_Pred)
cm

    array([[190, 15],
          [ 36, 79]])

sns.heatmap(pd.DataFrame(cm), annot=True)

from sklearn.metrics import accuracy_score
print(Y_test.shape)
print(Y_Pred.shape)

len(Y_test) == len(Y_Pred)

print(f"Length of Y_test: {len(Y_test)}")
print(f"Length of Y_Pred: {len(Y_Pred)}")

    Length of Y_test: 80
    Length of Y_Pred: 320

```

```
if len(Y_test) < len(Y_Pred):  
    Y_test = np.pad(Y_test, (0, len(Y_Pred) - len(Y_test)))  
elif len(Y_Pred) < len(Y_test):  
    Y_Pred = np.pad(Y_Pred, (0, len(Y_test) - len(Y_Pred)))  
  
accuracy = accuracy_score(Y_test, Y_Pred)  
accuracy
```