

Name : Pranali Baravkar

RBT23CS004

TY A-div(A1 Batch)

```
import pandas as pd
import numpy as np

data = {
    'User ID': np.random.randint(10000000, 99999999, 100),
    'Gender': np.random.choice(['Male', 'Female'], 100),
    'Age': np.random.randint(18, 60, 100),
    'EstimatedSalary': np.random.randint(20000, 150000, 100),
    'Purchased': np.random.choice([0, 1], 100)}

df = pd.DataFrame(data)

df.to_csv('USERS.csv', index=False)

print("Sample 'USERS.csv' file created.\n")

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

data_set = pd.read_csv('user_data.csv')

x = data_set.iloc[:, [2, 3]].values

y = data_set.iloc[:, 4].values

x_train, x_test, y_train, y_test = train_test_split(
    x, y, test_size=0.25, random_state=0)

st_x = StandardScaler()

x_train = st_x.fit_transform(x_train)

x_test = st_x.transform(x_test)

print("✅ Dataset processed and ready for modeling.")
```

Output :

Sample 'USERS.csv' file created.

✅ Dataset processed and ready for modeling.

