28.Question: K-Means Clustering for Customer Segmentation

You are working for an e-commerce company and want to segment your customers into distinct

groups based on their purchasing behavior. You have collected a dataset of customer data with

various shopping-related features.

Write a Python program that allows the user to input the shopping-related features of a new

customer. The program should use K-Means clustering from scikit-learn to assign the new customer

to one of the existing segments based on the input features.

Code;

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

**from** sklearn.cluster **import** KMeans

**from** sklearn.preprocessing **import** StandardScaler

data\_path **=** r"C:\Users\vara prasad\Downloads\customer\_data.csv" *# Update this path to where you saved the CSV file*

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

**if** 'CustomerID' **not** **in** df.columns:

df['CustomerID'] **=** range(1, len(df) **+** 1)

X **=** df[['Annual\_Spend', 'Visit\_Frequency', 'Time\_on\_Site']]

scaler **=** StandardScaler()

X\_scaled **=** scaler.fit\_transform(X)

kmeans **=** KMeans(n\_clusters**=**3, random\_state**=**42)

kmeans.fit(X\_scaled)

df['Cluster'] **=** kmeans.labels\_

print("Clustered Customer Data:")

print(df)

**def** classify\_new\_customer(annual\_spend, visit\_freq, time\_on\_site):

new\_customer **=** pd.DataFrame([[annual\_spend, visit\_freq, time\_on\_site]], columns**=**['Annual\_Spend', 'Visit\_Frequency', 'Time\_on\_Site'])

new\_customer\_scaled **=** scaler.transform(new\_customer)

cluster **=** kmeans.predict(new\_customer\_scaled)

**return** cluster[0]

print("\nEnter details for the new customer:")

annual\_spend **=** float(input("Annual Spend: "))

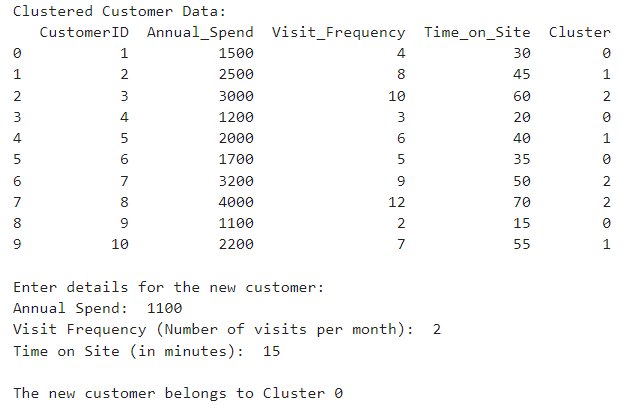
visit\_freq **=** int(input("Visit Frequency (Number of visits per month): "))

time\_on\_site **=** float(input("Time on Site (in minutes): "))

new\_customer\_cluster **=** classify\_new\_customer(annual\_spend, visit\_freq, time\_on\_site)

print(f"\nThe new customer belongs to Cluster {new\_customer\_cluster}")

output:



Dataset:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CustomerID | Annual\_Spend | Visit\_Frequency | Time\_on\_Site | |
| 1 | 1500 | 4 | 30 |  |
| 2 | 2500 | 8 | 45 |  |
| 3 | 3000 | 10 | 60 |  |
| 4 | 1200 | 3 | 20 |  |
| 5 | 2000 | 6 | 40 |  |
| 6 | 1700 | 5 | 35 |  |
| 7 | 3200 | 9 | 50 |  |
| 8 | 4000 | 12 | 70 |  |
| 9 | 1100 | 2 | 15 |  |
| 10 | 2200 | 7 | 55 |  |