Practical No: 8

Clustering algorithm

AIM: Write an application to implement clustering algorithm.

Code: Unsupervised Learning

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import scipy.cluster.hierarchy as shc
from sklearn.cluster import AgglomerativeClustering

# Read the customer data from a CSV file
customer_data = pd.read_csv('/content/mall_customers.csv')

# Display the shape and the first few rows of the data
print(customer_data.shape)
customer_data.head()

# Extract the relevant columns from the data
data = customer_data.iloc[:, 3:5].values

# Create a dendrogram plot
plt.figure(figsize=(10, 7))
plt.title("Customer Dendrograms")
dend = shc.dendrogram(shc.linkage(data, method='ward'))

# Perform hierarchical clustering
cluster = AgglomerativeClustering(n_clusters=5, affinity='euclidean', linkage='ward')
cluster_labels = cluster.fit_predict(data)

# Create a scatter plot to visualize the clusters
plt.figure(figsize=(10, 7))
plt.scatter(data[:, 0], data[:, 1], c=cluster_labels, cmap='rainbow')
plt.show()
```
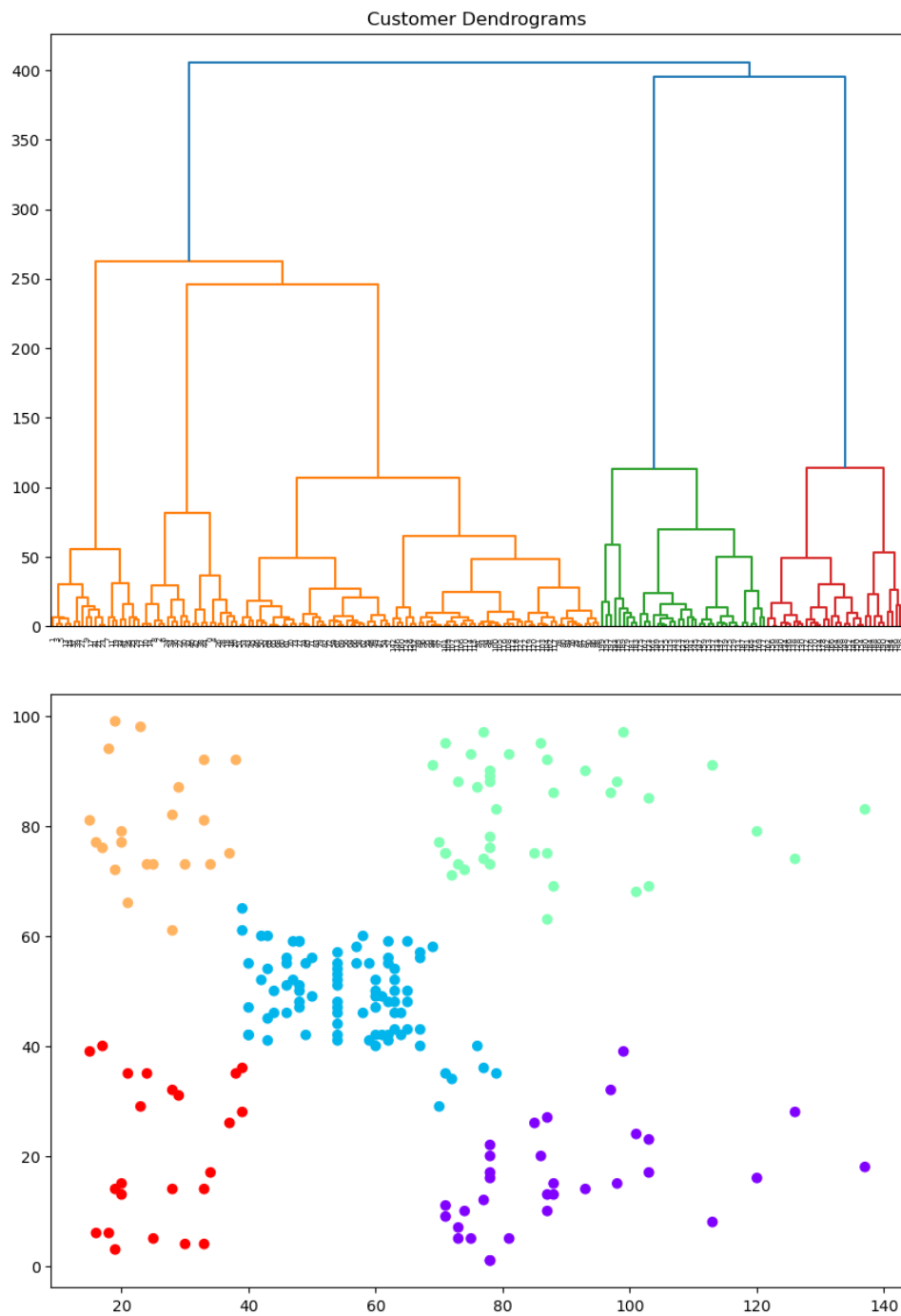
OUTPUT:

Out[5]:

	customer_id	gender	age	annual_income	spending_score
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40

Customer Dendrograms

b) Clustering

Code:

```
from numpy import where
from sklearn.datasets import make_classification
from matplotlib import pyplot

x,y =
make_classification(n_samples=1000,n_features=2,n_informative=2,n_redundant=0,n_cluste
rs_per_class=1,random_state=4)
 for class_value in range(2):
   row_ix=where(y==class_value)
   pyplot.scatter(x[row_ix,0],x[row_ix,1])
pyplot.show()
```

OUTPUT: