

**Experiment No.: 9****Aim**

Program to implement k-means clustering technique using any standard dataset available in the public domain

**CO3**

Use different packages and frameworks to implement text classification using SVM and clustering using k-means

**Procedure**

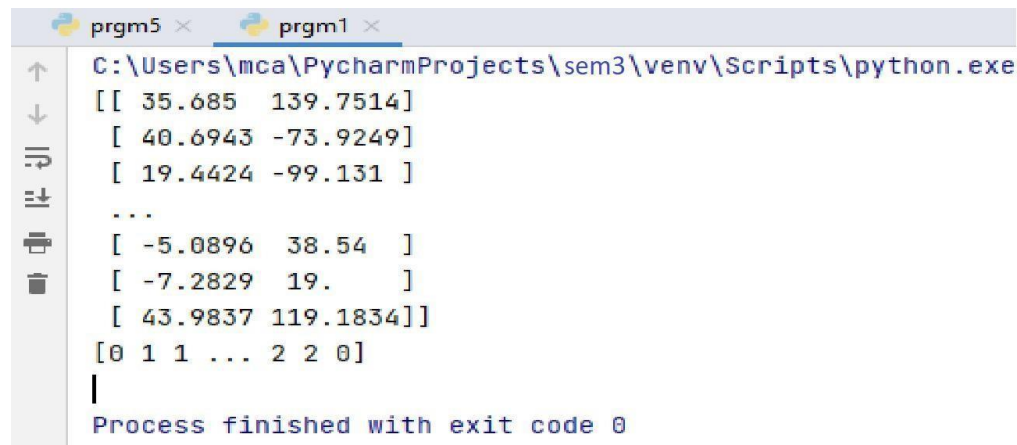
```
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd
dataset = pd.read_csv('data.csv')
x = dataset.iloc[:, [1, 2]].values
print(x)
from sklearn.cluster import KMeans
wcss_list = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(x)
    wcss_list.append(kmeans.inertia_)
mtp.plot(range(1, 11), wcss_list)
mtp.title('The Elbow Method Graph')
mtp.xlabel('Number of clusters(k)')
mtp.ylabel('wcss_list')
mtp.show()

kmeans = KMeans(n_clusters=3, init='k-means++', random_state=42)
y_predict = kmeans.fit_predict(x)
print(y_predict)
```

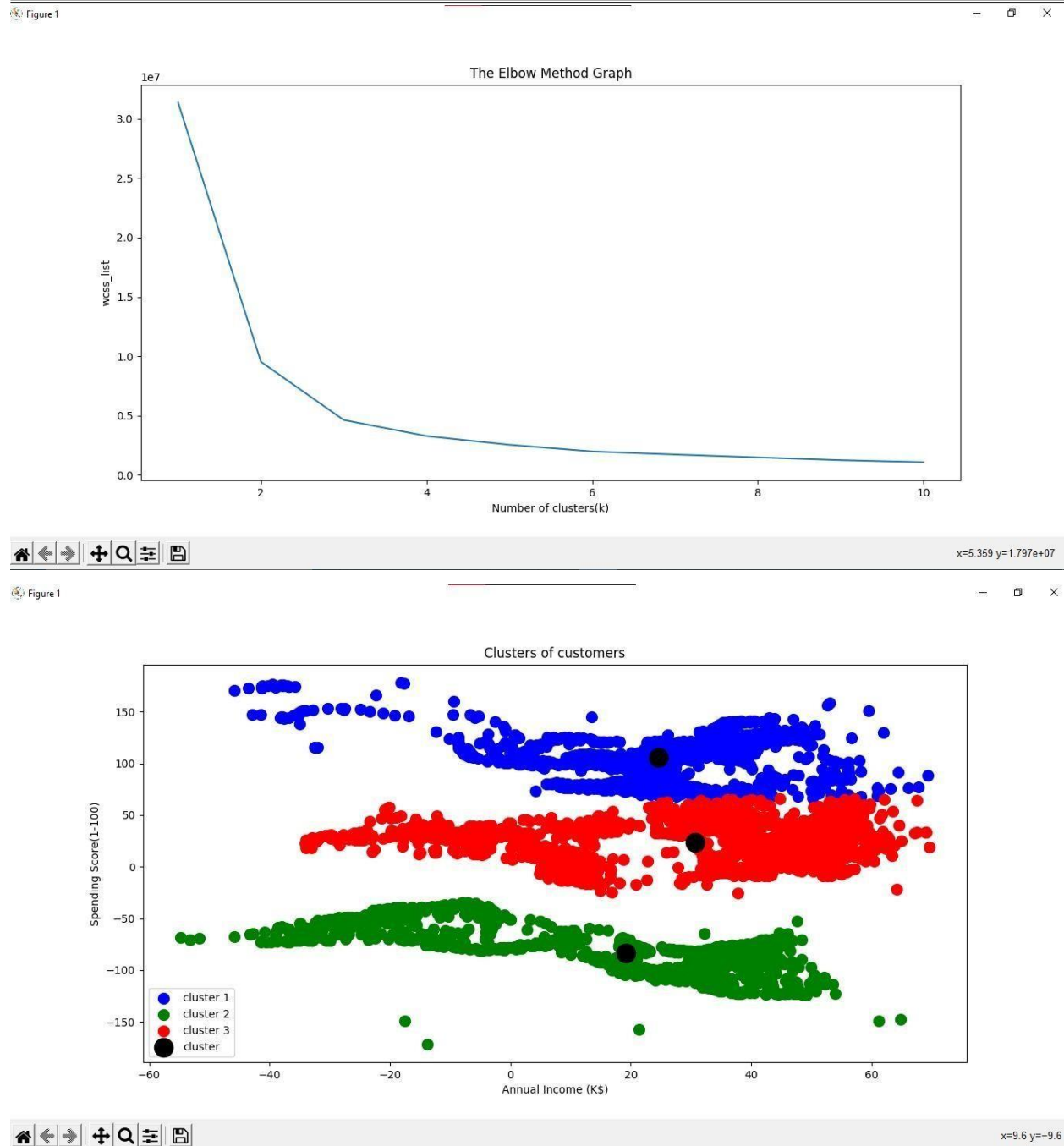
```
mtp.scatter(x[y_predict == 0, 0], x[y_predict == 0, 1], s=100, c='blue', label='cluster 1')
mtp.scatter(x[y_predict == 1, 0], x[y_predict == 1, 1], s=100, c='green', label='cluster 2')
mtp.scatter(x[y_predict == 2, 0], x[y_predict == 2, 1], s=100, c='red', label='cluster 3')
mtp.scatter(kmeans.cluster_centers_[0, 0], kmeans.cluster_centers_[0, 1], s=300, c='black',
label='cluster')

mtp.title('Clusters of customers')
mtp.xlabel('Annual Income (K$)')
mtp.ylabel('Spending Score(1-100)')
mtp.legend()
mtp.show()
```

### **Output Screenshot**



```
prgm5 x prgm1 x
C:\Users\mca\PycharmProjects\sem3\venv\Scripts\python.exe
[[ 35.685 139.7514]
 [ 40.6943 -73.9249]
 [ 19.4424 -99.131 ]
 ...
 [-5.0896 38.54 ]
 [-7.2829 19. ]
 [ 43.9837 119.1834]]
[0 1 1 ... 2 2 0]
|
Process finished with exit code 0
```



## Result

The program was executed and the result was successfully obtained. Thus CO3 was obtained.