Program No:11

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

Program

```
import numpy as np
import matplotlib.pyplot as mtp
import pandas as pd
dataset=pd.read_csv('Mall_Customers.csv')
x=dataset.iloc[:,[3,4]].values
print(x)
from sklearn.cluster import KMeans
wcss_list = []
for i in range(1, 11):
      kmeans = KMeans(n_clusters=i, init='k-means++')
      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=5,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='Cluster0')
mtp.scatter(x[y_predict == 1,0], x[y_predict ==1,1], s=100, c='green', label= 'Cluster1')
mtp.scatter(x[y_predict == 2,0], x[y_predict ==2,1], s=100, c='red', label= 'Cluster2')
mtp.scatter(x[y_predict == 3,0], x[y_predict ==3,1], s=100, c='cyan', label= 'Cluster3')
mtp.scatter(x[y_predict == 4,0], x[y_predict ==4,1], s=100, c='magenta', label= 'Cluster4')
mtp.scatter(kmeans.cluster_centers_[:,0],kmeans.cluster_centers_[:,1], s = 300,)
mtp.title('clusters of customers')
mtp.xlabel('annual incmome (k$')
mtp.ylabel('spending score (1-100)')
mtp.legend()
mtp.show()
```

OUTPUT

```
C:\Users\ajcemca\AppData\Local\Programs\Python\Python39\python.exe C:/Users/ajcemca
[[ 15  39]
 [ 15  81]
 [ 16   6]
 [ 16  77]
 [ 17  40]
 [ 17  76]
 [ 18   6]
 [ 18  94]
 [ 19   3]
 [ 19  72]
 [ 19  14]
 [ 19  99]
 [ 20  15]
 [ 20  77]
 [ 20  13]
 [ 20  79]
 [ 21  35]
 [ 21  66]
 [ 23  29]
 [ 23  98]
 [ 24  35]
 [ 24  73]
 [ 25   5]
 [ 25  73]
 [ 28  14]
```

```
[ 25   73]
[ 28   14]
[ 28   82]
[ 28   32]
[ 28   61]
[ 29   31]
[ 29   87]
[ 30    4]
[ 30   73]
[ 33    4]
[ 33   92]
[ 33   14]
[ 33   81]
[ 34   17]
[ 34   73]
[ 37   26]
[ 37   75]
[ 38   35]
[ 38   92]
[ 39   36]
[ 39   61]
[ 39   28]
[ 39   65]
[ 40   55]
[ 40   47]
[ 40   42]
```

```
[ 42   52]
[ 42   60]
[ 43   54]
[ 43   60]
[ 43   45]
[ 43   41]
[ 44   50]
[ 44   46]
[ 46   51]
[ 46   46]
[ 46   56]
[ 46   55]
[ 47   52]
[ 47   59]
[ 48   51]
[ 48   59]
[ 48   50]
[ 48   48]
[ 48   59]
[ 48   47]
[ 49   55]
[ 49   42]
[ 50   49]
[ 50   56]
[ 54   47]
[ 54   54]
```

```
[ 54  51]
[ 54  55]
[ 54  41]
[ 54  44]
[ 54  57]
[ 54  46]
[ 57  58]
[ 57  55]
[ 58  60]
[ 58  46]
[ 59  55]
[ 59  41]
[ 60  49]
[ 60  40]
[ 60  42]
[ 60  52]
[ 60  47]
[ 60  50]
[ 61  42]
[ 61  49]
[ 62  41]
[ 62  48]
[ 62  59]
[ 62  55]
[ 62  56]
[ 62  42]
```

```
[ 63   46]
[ 63   43]
[ 63   48]
[ 63   52]
[ 63   54]
[ 64   42]
[ 64   46]
[ 65   48]
[ 65   50]
[ 65   43]
[ 65   59]
[ 67   43]
[ 67   57]
[ 67   56]
[ 67   40]
[ 69   58]
[ 69   91]
[ 70   29]
[ 70   77]
[ 71   35]
[ 71   95]
[ 71   11]
[ 71   75]
[ 71    9]
[ 71   75]
[ 72   34]
```

```
 [ 97  86]
 [ 98  15]
 [ 98  88]
 [ 99  39]
 [ 99  97]
 [101  24]
 [101  68]
 [103  17]
 [103  85]
 [103  23]
 [103  69]
 [113   8]
 [113  91]
 [120  16]
 [120  79]
 [126  28]
 [126  74]
 [137  18]
 [137  83]]
[2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2
 3 2 3 2 3 2 0 2 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 4 1 4 0 4 1 4 1 4 0 4 1 4 1 4 1 4 1 4 0 4 1 4 1 4
 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1
 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4]
```

Figure 1 — □ ✕

The elbow method Graph

Number of clusters (k)

clusters of customers