

PROGRAM NO : 12

AIM : Program to implement K-Means Clustering technique using any standard dataset available in the public domain

PROGRAM

```
import matplotlib.pyplot as mtp
import pandas as pd

dataset=pd.read_csv('Mall_Customers.csv')
x=dataset.iloc[:,[3,4]].values
print(x)

#find elbow
from sklearn.cluster import KMeans

wcss_list=[] #initializing the list for the values of WCSS (sum of squared distance b/w each value)

#using the loop for iteration from 1 to 10
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()

#training the K-Means model on a dataset

kmeans=KMeans(n_clusters=5,init='k-means++',random_state=42)
y_predict=kmeans.fit_predict(x)
print(y_predict)

#Visualizing the Clusters
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(x[y_predict == 3,0],x[y_predict == 3,1],s=100,c='cyan',label='Cluster 4')
mtp.scatter(x[y_predict == 4,0],x[y_predict == 4,1],s=100,c='magenta',label='Cluster 5')

mtp.scatter(kmeans.cluster_centers_[0,0],kmeans.cluster_centers_[0,1],s=300,c='black')

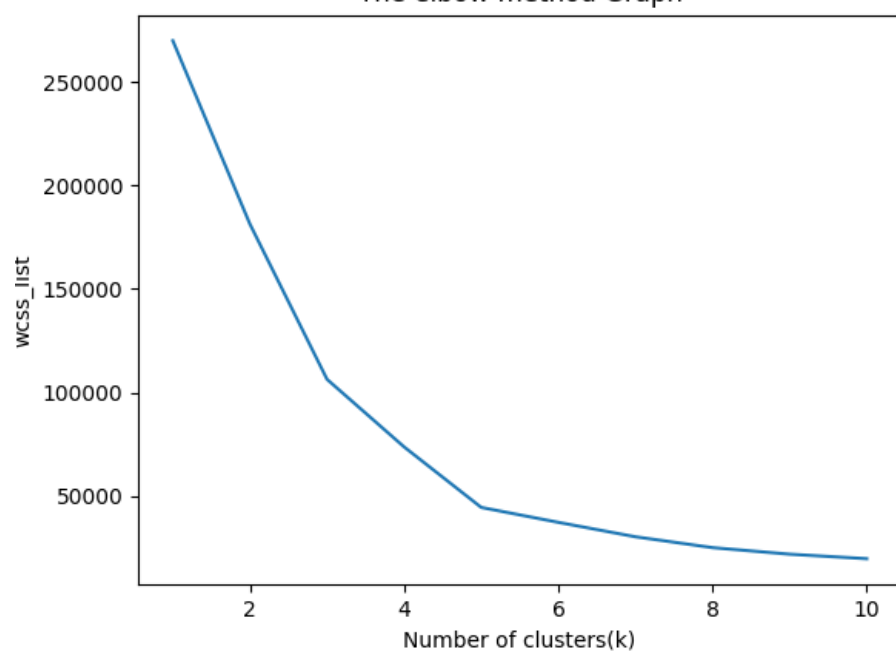
mtp.title("Clusters of Customers")
```

```
mtp.xlabel("Annual Income (K$)")
mtp.ylabel("Spending Score (1-100)")
mtp.legend()
mtp.show()
```

OUTPUT

[illegible]

The elbow method Graph



Clusters of Customers

