nrcm-kmeans2

August 28, 2023

##**NAME**:K.swetha ##**BRANCH**:Data Science ##**ROLLNO**:21X05A6730 ##**COLLEGE**:Narsmiha Reddy Engineeering College

###**PROJECT TITLE**: Analysis and predection of"Mallcustomers.csv" of american mall markets called as phonix Mall to find out how many customers are visited to a particular a shop on the basis of these predection of anual income vas spending score

###**DISCLIMER**:In this particular dataset we assume annual income as a centroide and spending score from the range 1 to 100 calld as datanodes of the clusters

###**PROBLEM STATEMENT**:The American finance market as per the GDPof 2011 "phonex_trillums" Mall as in the first range out of five.The owner of the Mall wants to be exact which particular shop or product search in different kinds of clusters in entire Mall ### As a Data Science engineer predict the futuristic finanical market for upcoming GDP rate based on number of clusters ###The client want atleast five top clusters(shops).

##** PROBLEM STATEMEN**T:The American finance market as per the GDPof 2011 "phonex_trillums" Mall as in the first range out of five.The owner of the Mall wants to be exact which particular shop or product search in different kinds of clusters in entire Mall As a DaPRSTATEMENTOBLEMta Science engineer predict the futuristic finanical market for upcoming GDP rate based on number of clusters The client want atleast five top clusters(shops).

[10]:
```
## Import pandas, numpy, matplotlib
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

[11]:
```
#Read the dataset take variable name called "dataset" only.
dataset = pd.read_csv("/content/Mall_Customers.csv")

# without printing this data add in separet variable as input variable Caqpital␣
 ↪X only. loc index by select the all row ,
#and give the required colum index like[3,4].for this particular dataset.
x = dataset.iloc[:,[3,4]].values
```

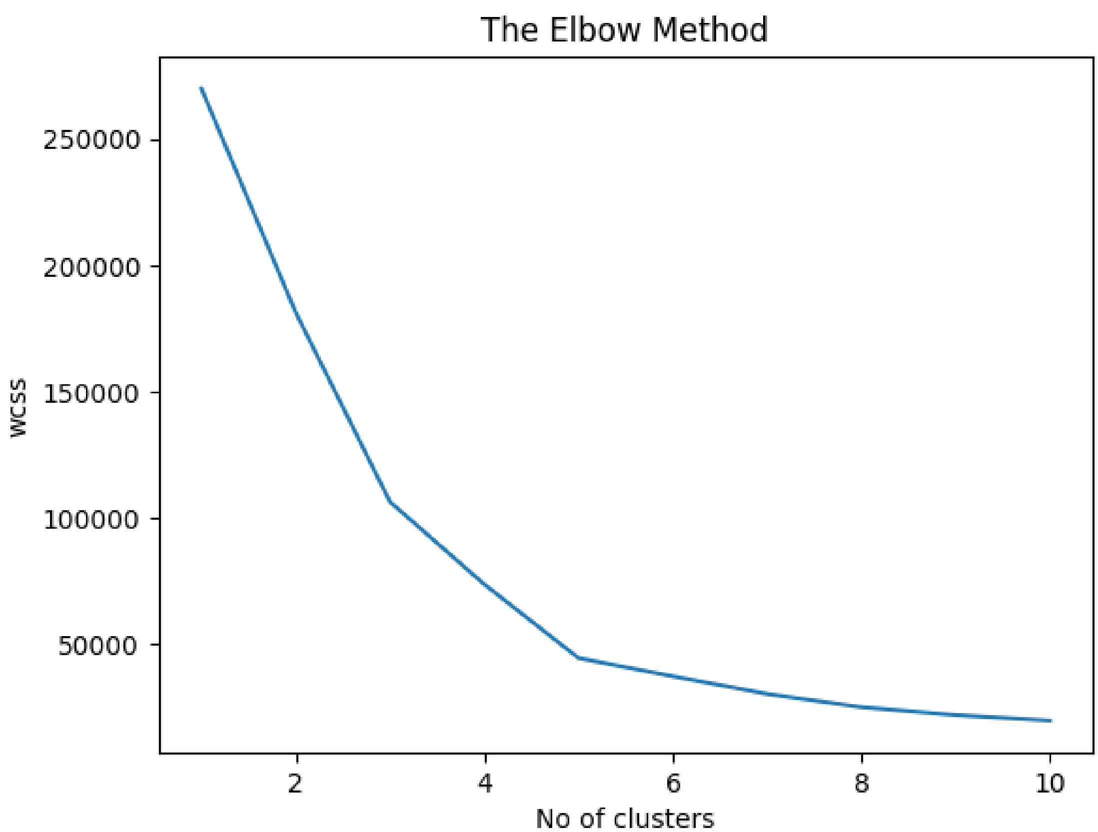
[21]:
```
from pandas.core.common import random_state
## <THE ELBOW METHOD>
#from sklearn used "sklearn.cluster" attribute and import KMeans
#Take a distance from from centroid to cluster point with WrapsColumnExpression.
```

```python
# Assume you have 10 cluster and iterate the for up to range 10 with iterater␣
↪kmeans++.
# Fit the model if value comes too samlla in range.
#For clustering in wcss ,inertia is adding / appending is required.(kmeans.
↪inertia_)#defalut usecase.
#Plot the poarticular graph along with the wcss and your range which you taken␣
↪as input variable.
#Add title "The Elbow Method".
#Lable x variable as "No of Customers".
#Lable y variable as "WCSS".
#Plot the graph using plt.show().
from sklearn.cluster import KMeans
wcss = []
for i in range(1,11):
  kmeans = KMeans(n_clusters = i, init="k-means++", random_state = 42)
  kmeans.fit(x)
  wcss.append(kmeans.inertia_)
plt.plot(range(1,11),wcss)

plt.title("The Elbow Method")
plt.xlabel("No of clusters")
plt.ylabel("wcss")
plt.show()
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
```

```
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
```

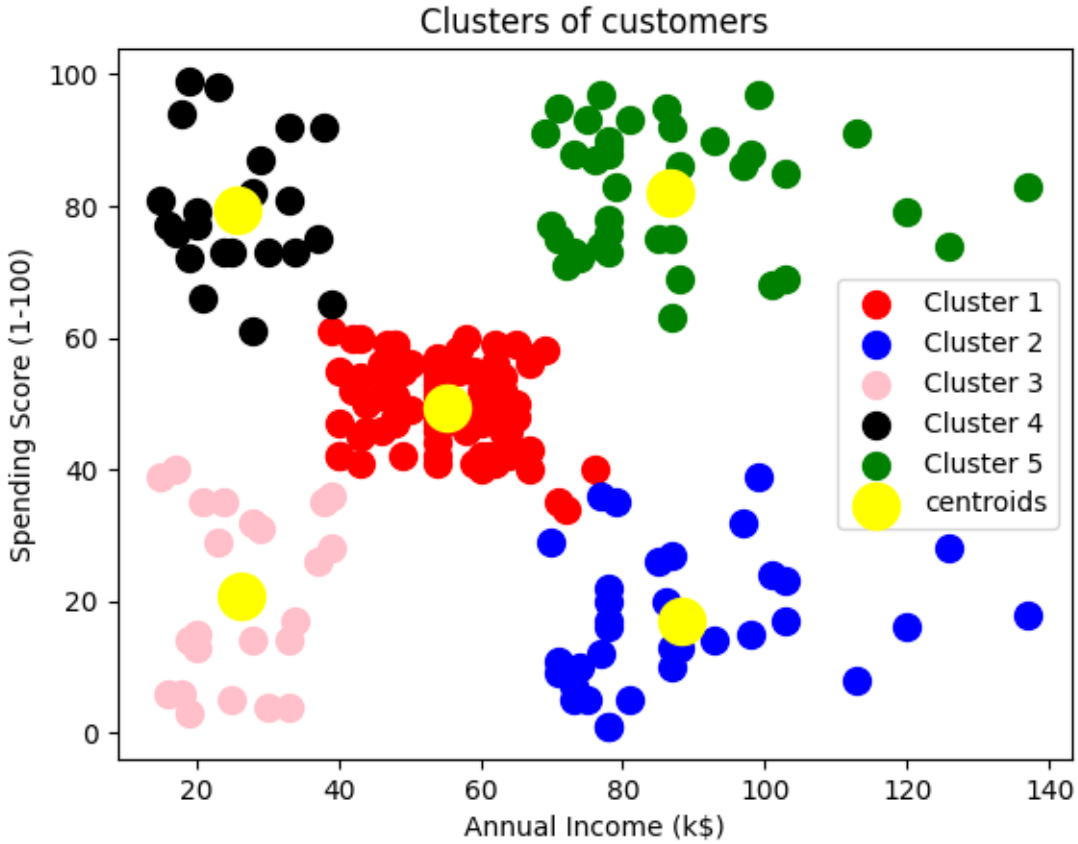


[35]:
```
for i in range(1,11):
  kmeans = KMeans(n_clusters = 5 , init="k-means++", random_state = 42)
  y_kmeans=kmeans.fit_predict(x)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
```

[36]:
```python
# Take any no of cluster and run you take 5.
plt.scatter(x[y_kmeans == 0, 0], x[y_kmeans == 0, 1], s = 100, c = 'red', label␣
 ↪= 'Cluster 1')
plt.scatter(x[y_kmeans == 1, 0], x[y_kmeans == 1, 1], s = 100, c = 'blue',␣
 ↪label = 'Cluster 2')
```

```
plt.scatter(x[y_kmeans == 2, 0], x[y_kmeans == 2, 1], s = 100, c = 'pink',␣
 ↪label = 'Cluster 3')
plt.scatter(x[y_kmeans == 3, 0], x[y_kmeans == 3, 1], s = 100, c = 'black',␣
 ↪label = 'Cluster 4')
plt.scatter(x[y_kmeans == 4, 0], x[y_kmeans == 4, 1], s = 100, c = 'green',␣
 ↪label = 'Cluster 5')
#Write Code for rest.SS
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1], s =␣
 ↪300, c = 'yellow', label = 'centroids')
plt.title('Clusters of customers')
plt.xlabel('Annual Income (k$)')
plt.ylabel('Spending Score (1-100)')
plt.legend()
plt.show()
```



###**CONCLUSION**: According to the model basics predections using machine learning alogorithm kmeans clusytering we found that clusters were which consist red color is a highest cluster which attach more than 50 datanodes.

[ ]:

###**REFERENCES**:The model buliding algorithm develop for all kinds of clusteration values.The yellow spots represents centroids which is max TO max 3