

kmeans

August 28, 2023

###CH.Mahesh #####21X05A6712 #####CSE-DS

##**Project Title:-** #####Analysis and prediction of “Mall_Customers” of american mall market called PHONIX Mall.To find out how many customers are visited to a particular shop.On the basics of prediction of Annual Income vs Spending Score

##DISCLAIMER ##### In this paricular dataset we assume Annual income as a centroid and Spending Score from the range 1-100 called as the data nodes of the statement

###Problem Statement #####The American Finance market as per the GDP of 2011 “PHONIX_TRILLUMS” mall and the first range out of 5.The owner of the mall wants to be exact which particular shop all the products such in different kind of clusters in entire mall. #####As a data scientist Engineer predict the futuristic financial marketfor upcoming GDP rate based on no of clusters #####The client want atleast top 5 clusters(shops).

```
[26]: #import the numpy, matlot, pandas libery's
import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
```

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

```
[33]: ## <THE ELBOW METHOD>
#from sklearn used "sklearn.cluster" attribute and import KMeans
#Take a distance from from centroid to cluster point with WrapsColumnExpression.
# 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()

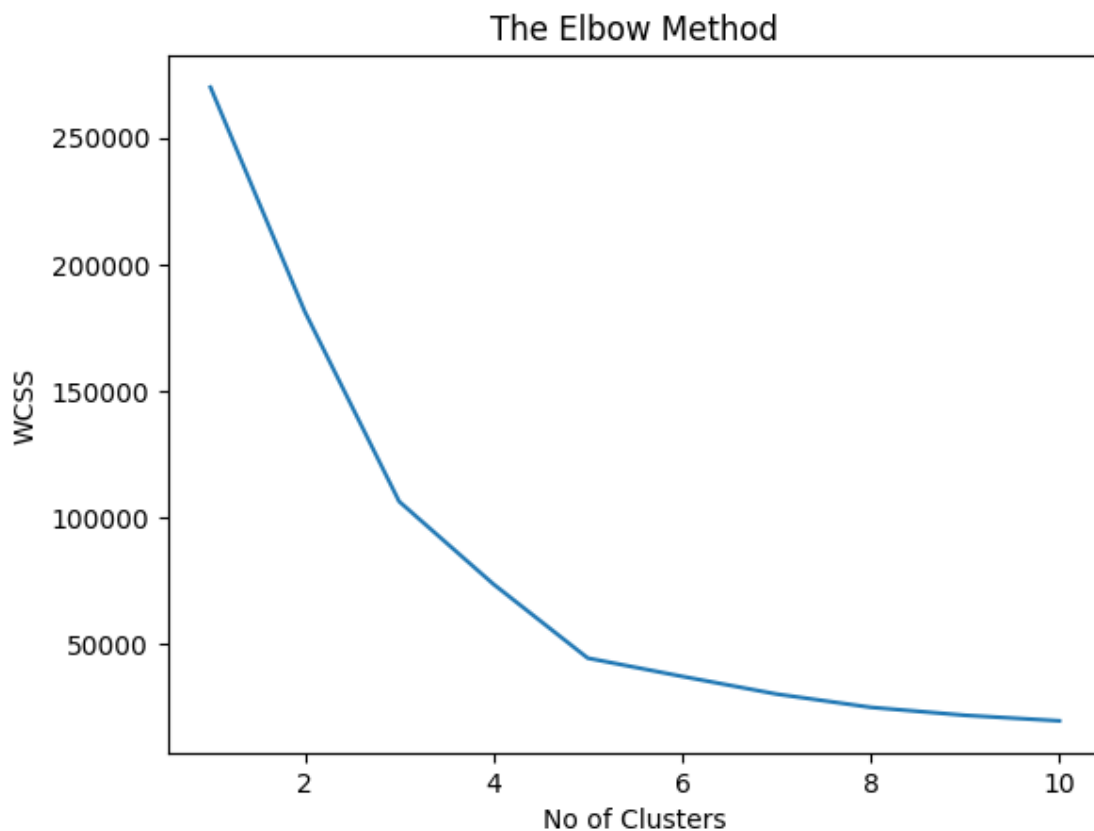
```

```

/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(

```

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



```
[36]: for i in range(1,11):
        kmeans=KMeans(n_clusters = 3,init="k-means++",random_state=42)
        y_kmeans=kmeans.fit_predict(X)
        # 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 = 'green',
            label = 'Cluster 3')
plt.scatter(X[y_kmeans == 3, 0], X[y_kmeans == 3, 1], s = 100, c = 'yellow',
            label = 'Cluster 4')
plt.scatter(X[y_kmeans == 4, 0], X[y_kmeans == 4, 1], s = 100, c = 'cyan',
            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()

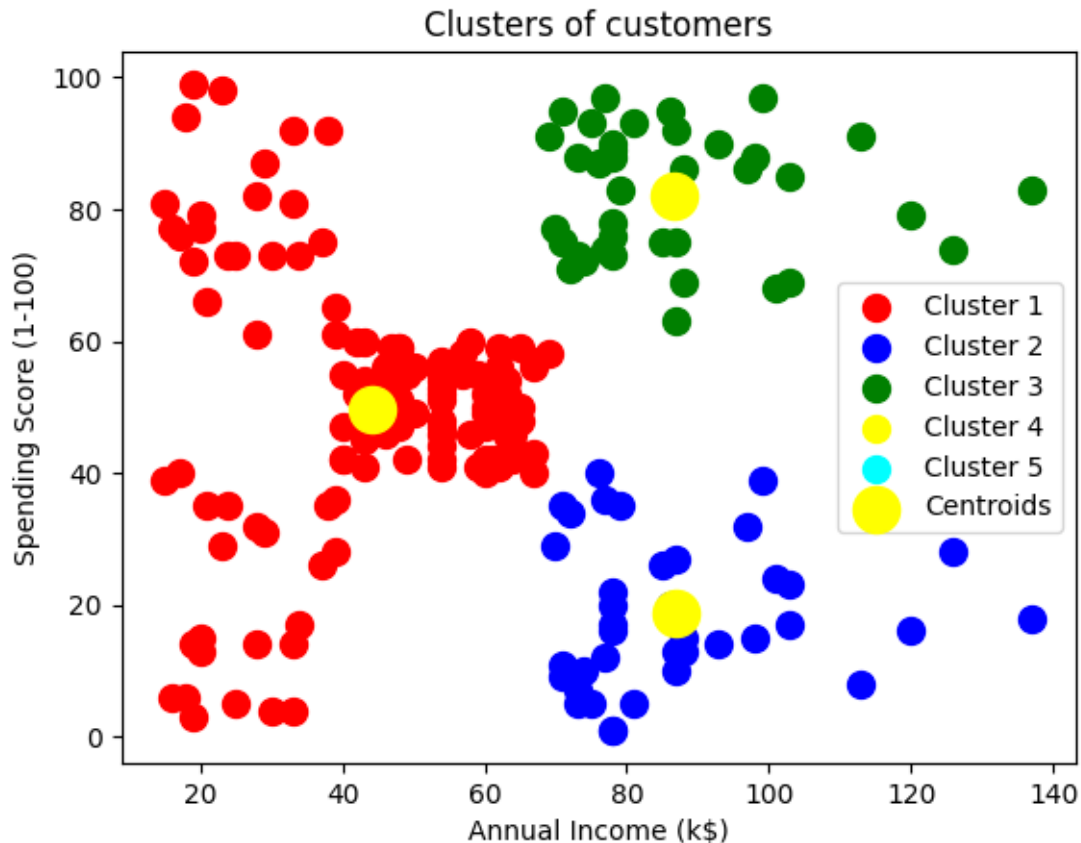
```

```

/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(

```

```
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(
```



[]:

###CONCLUSION:- #####According using Machine Learning Algorithm K means clustering we count that cluster 1 which consists RED colour is a highest cluster which attach more 50 data nodes.

###REFERENCES:- #####The model building algorithm develop for all kinds of clusteration values.The yellow spot represents the “CENTROID” which is max of 3.

[]: