In [1]: import numpy as np
 import pandas as pd
 import matplotlib.pyplot as plt
 import seaborn as sb

In [2]: df=pd.read_csv('Customers.csv')

In [3]: df.head(15)

Out[3]:

CustomerID	Gender	Age	Annual Income (\$)	Spending Score (1-100)	Profession	Work Experience	Family Size
1	Male	19	15000	39	Healthcare	1	4
2	Male	21	35000	81	Engineer	3	3
3	Female	20	86000	6	Engineer	1	1
4	Female	23	59000	77	Lawyer	0	2
5	Female	31	38000	40	Entertainment	2	6
6	Female	22	58000	76	Artist	0	2
7	Female	35	31000	6	Healthcare	1	3
8	Female	23	84000	94	Healthcare	1	3
9	Male	64	97000	3	Engineer	0	3
10	Female	30	98000	72	Artist	1	4
11	Male	67	7000	14	Engineer	1	3
12	Female	35	93000	99	Healthcare	4	4
13	Female	58	80000	15	Executive	0	5
14	Female	24	91000	77	Lawyer	1	1
15	Male	37	19000	13	Doctor	0	1
	1 2 3 4 5 6 7 8 9 10 11 12 13	1 Male 2 Male 3 Female 4 Female 5 Female 6 Female 7 Female 8 Female 9 Male 10 Female 11 Male 12 Female 13 Female 14 Female	1 Male 19 2 Male 21 3 Female 20 4 Female 23 5 Female 31 6 Female 22 7 Female 35 8 Female 23 9 Male 64 10 Female 30 11 Male 67 12 Female 35 13 Female 58 14 Female 24	CustomerID Gender Age Income (\$) 1 Male 19 15000 2 Male 21 35000 3 Female 20 86000 4 Female 23 59000 5 Female 31 38000 6 Female 22 58000 7 Female 35 31000 8 Female 23 84000 9 Male 64 97000 10 Female 30 98000 11 Male 67 7000 12 Female 35 93000 13 Female 58 80000 14 Female 24 91000	Customerib Gender Age Income (\$) Score (1-100) 1 Male 19 15000 39 2 Male 21 35000 81 3 Female 20 86000 6 4 Female 23 59000 77 5 Female 31 38000 40 6 Female 22 58000 76 7 Female 35 31000 6 8 Female 23 84000 94 9 Male 64 97000 3 10 Female 30 98000 72 11 Male 67 7000 14 12 Female 35 93000 99 13 Female 58 80000 15 14 Female 24 91000 77	Customerio Gender Age Income (\$) Score (1-100) Profession 1 Male 19 15000 39 Healthcare 2 Male 21 35000 81 Engineer 3 Female 20 86000 6 Engineer 4 Female 23 59000 77 Lawyer 5 Female 31 38000 40 Entertainment 6 Female 22 58000 76 Artist 7 Female 35 31000 6 Healthcare 8 Female 23 84000 94 Healthcare 9 Male 64 97000 3 Engineer 10 Female 30 98000 72 Artist 11 Male 67 7000 14 Engineer 12 Female 35 93000 99 Healthcare 13 Female 58	Customeria Age Income (\$) Score (1-100) Profession Experience 1 Male 19 15000 39 Healthcare 1 2 Male 21 35000 81 Engineer 3 3 Female 20 86000 6 Engineer 1 4 Female 23 59000 77 Lawyer 0 5 Female 31 38000 40 Entertainment 2 6 Female 22 58000 76 Artist 0 7 Female 35 31000 6 Healthcare 1 8 Female 23 84000 94 Healthcare 1 9 Male 64 97000 3 Engineer 0 10 Female 30 98000 72 Artist 1 11 Male 67 7000 14 Engineer 1 </th

In [4]: df.shape

Out[4]: (2000, 8)

In [5]: df.info

Out[5]:		d method Data Spending Scor			CustomerID	Gender	Age	Annual	Income
	0	1	Male	19	15000				39
	1	2	Male	21	35000				81
	2		Female	20	86000				6
	3		Female	23	59000				77
	4		Female	31	38000				40
	-	,							
	1995	1996	Female	71	 184387				 40
	1996		Female	91	73158				32
	1997	1998	Male	87	90961				14
	1998	1999	Male	77	182109				4
	1999	2000	Male	90	110610				4 52
	1999	2000	Maic	90	110010				32
		Professio	n Work	Experience	Family Size				
	0	Healthcar		1	4				
	1	Enginee		3	3				
	2	Enginee		1	1				
	3	Lawye		9	2				
	4	Entertainmer		2	- 6				
	• • •			• • •	• • •				
	1995	Artis	it.	8	7				
	1996	Docto		7	7				
	1997	Healthcar		9	2				
	1998	Executiv		7	2				
	1999	Entertainmer		5	2				
		cc. cacr		3	-				

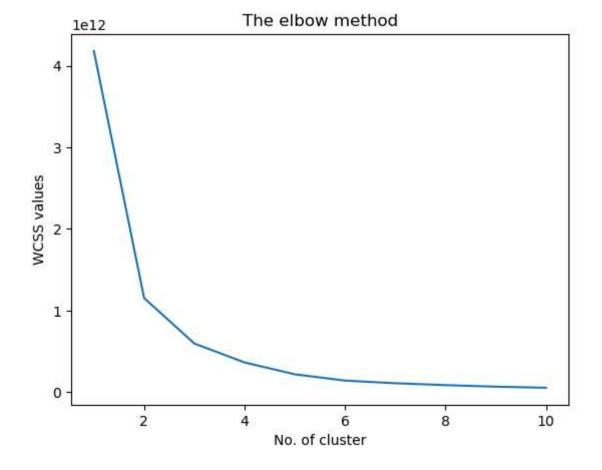
[2000 rows x 8 columns]>

```
In [6]: df.isnull().sum
 Out[6]: <bound method NDFrame._add_numeric_operations.<locals>.sum of
                                                                                   CustomerI
             Gender
                        Age Annual Income ($) Spending Score (1-100) \
          0
                      False
                              False False
                                                          False
                                                                                    False
          1
                      False
                              False False
                                                          False
                                                                                    False
          2
                              False False
                                                          False
                                                                                    False
                      False
          3
                      False
                              False False
                                                          False
                                                                                    False
          4
                              False False
                                                          False
                                                                                    False
                      False
                        . . .
                                . . .
                                        . . .
                                                             . . .
                                                                                      . . .
          1995
                      False
                              False
                                      False
                                                          False
                                                                                    False
                              False False
                                                                                    False
          1996
                      False
                                                          False
          1997
                      False
                              False
                                      False
                                                          False
                                                                                    False
          1998
                      False
                              False False
                                                          False
                                                                                    False
          1999
                              False
                                      False
                                                          False
                                                                                    False
                      False
                Profession
                             Work Experience
                                               Family Size
          0
                      False
                                        False
                                                      False
          1
                      False
                                        False
                                                      False
          2
                      False
                                        False
                                                      False
          3
                      False
                                        False
                                                      False
          4
                      False
                                        False
                                                      False
                        . . .
                                          . . .
                                                        . . .
          . . .
                      False
                                        False
                                                      False
          1995
          1996
                      False
                                        False
                                                      False
          1997
                      False
                                        False
                                                      False
          1998
                      False
                                        False
                                                      False
          1999
                      False
                                        False
                                                      False
          [2000 rows x 8 columns]>
         X= df.iloc[:,[3,4]].values
 In [7]:
 In [8]:
         print(X)
          [[ 15000
                        39]
           35000
                        81]
           [ 86000
                         6]
                        14]
           [ 90961
           [182109
                         4]
           [110610
                        52]]
         from sklearn.cluster import KMeans
In [26]:
          wcss =[]
```

```
In [27]: for i in range(1, 11):
    kmeans = skc.KMeans(n_clusters=i, init='k-means++', random_state=0)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)
```

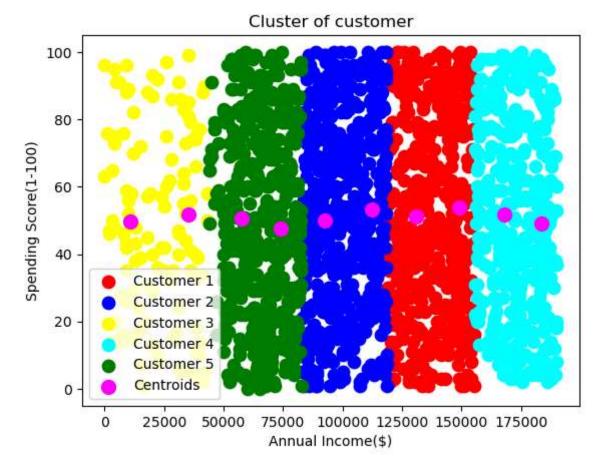
C:\Users\Sridip\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1036:
UserWarning: KMeans is known to have a memory leak on Windows with MKL, when
there are less chunks than available threads. You can avoid it by setting the
environment variable OMP_NUM_THREADS=8.
 warnings.warn(

```
In [28]: plt.plot(range(1,11), wcss)
    plt.title('The elbow method')
    plt.xlabel('No. of cluster')
    plt.ylabel('WCSS values')
    plt.show()
```



```
In [30]: kmeansmodel = KMeans(n_clusters=5, init='k-means++', random_state=0)
```

```
In [31]: y_kmeans = kmeansmodel.fit_predict(X)
```



In []: