```
In  [1]:  import  pandas  as  pd  import
          matplotlib.pyplot as plt
```

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

```
In [3]: df.shape
Out[3]:  (200, 5)
In [7]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
 #   Column                 Non-Null Count  Dtype -
--  ------                 --------------  -----  0
CustomerID             200 non-null    int64
1   Genre                   200 non-null    object
2   Age                     200 non-null    int64
3   Annual Income (k$)      200 non-null    int64  4   Spending Score (1-100)
    200 non-null    int64 dtypes: int64(4), object(1) memory usage: 7.9+ KB
```

```
In [9]:  X=df.iloc[:,[3,4]].values
```

```
In [11]: X
```

```
Out[11]: array([[ 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],
                [ 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],
                [ 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,  53],
                [ 54,  48],
                [ 54,  52],
                [ 54,  42],
                [ 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,  50],
                [ 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],
                [ 72,  71],
                [ 73,   5],
                [ 73,  88],
                [ 73,   7],
                [ 73,  73],
                [ 74,  10],
                [ 74,  72],
                [ 75,   5],
                [ 75,  93],
                [ 76,  40],
                [ 76,  87],
                [ 77,  12],
                [ 77,  97],
                [ 77,  36],
                [ 77,  74],
                [ 78,  22],
                [ 78,  90],
                [ 78,  17],
                [ 78,  88],
                [ 78,  20],
                [ 78,  76],
                [ 78,  16],
                [ 78,  89],
                [ 78,   1],
                [ 78,  78],
                [ 78,   1],
                [ 78,  73],
                [ 79,  35],
                [ 79,  83],
                [ 81,   5],
                [ 81,  93],
                [ 85,  26],
                [ 85,  75],
                [ 86,  20],
                [ 86,  95],
                [ 87,  27],
                [ 87,  63],
                [ 87,  13],
                [ 87,  75],
                [ 87,  10],
                [ 87,  92],
                [ 88,  13],
                [ 88,  86],
                [ 88,  15],
                [ 88,  69],
                [ 93,  14],
                [ 93,  90], [ 97,  32],
                 [ 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]], dtype=int64)
```

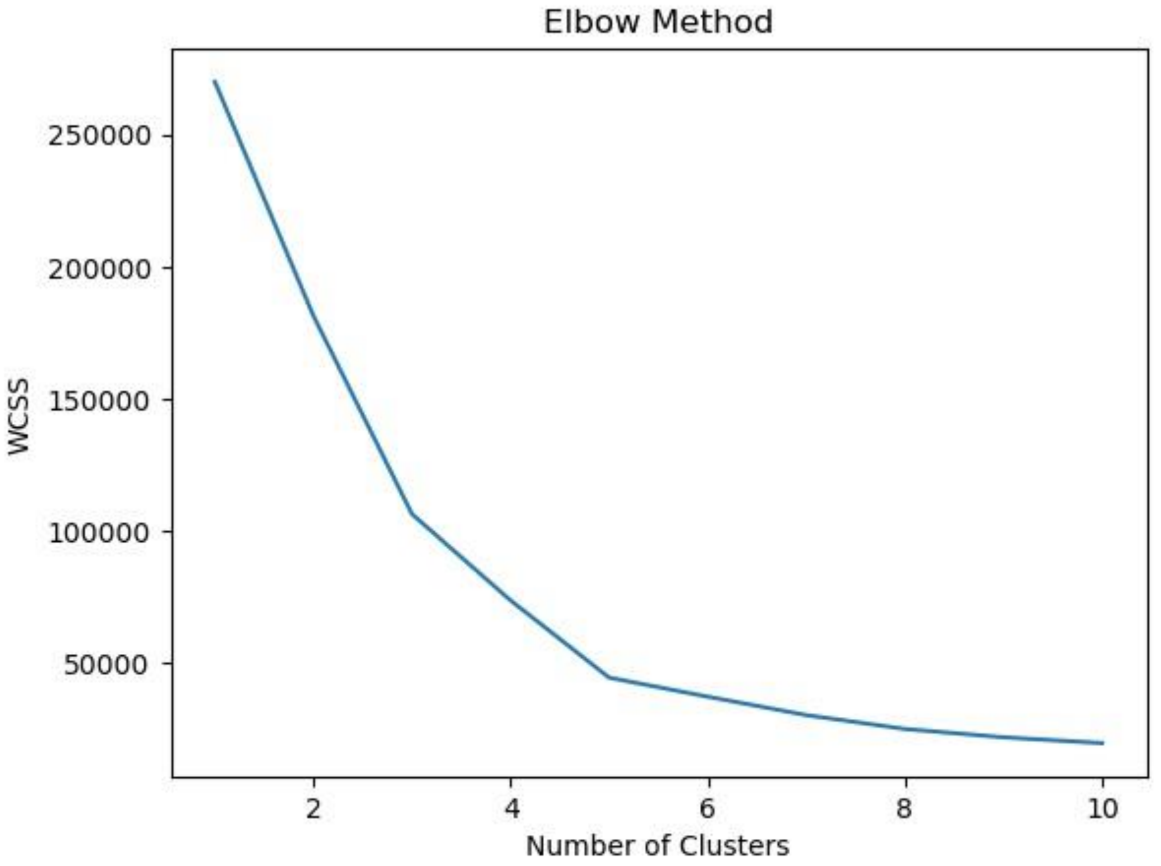
In [13]:

```
from sklearn.cluster import KMeans
# wcss : Within Cluster Sum of Squares : sum of squared difference between insta #
Elbow method for deciding optimum number of clusters, which minimizes value of
wcss=[] # contain wcss value for various number of clusters


for i in range(1,11):
    kmeans=KMeans(n_clusters=i, init='k-means++',max_iter=300,random_state=42)
kmeans.fit(X)
    wcss.append(kmeans.inertia_)

plt.plot(range(1,11),wcss) plt.title("Elbow
Method") plt.xlabel('Number of Clusters')
plt.ylabel('WCSS')
```

```
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar
e less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
```

```
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.   warnings.warn(
```

```
Out[13]: Text(0, 0.5, 'WCSS')
```



In [14]:
```
kmeans=KMeans(n_clusters=5,init='k-means++',max_iter=300,random_state=42)
y_kmeans=kmeans.fit_predict(X)
```

```
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.   warnings.warn(
```

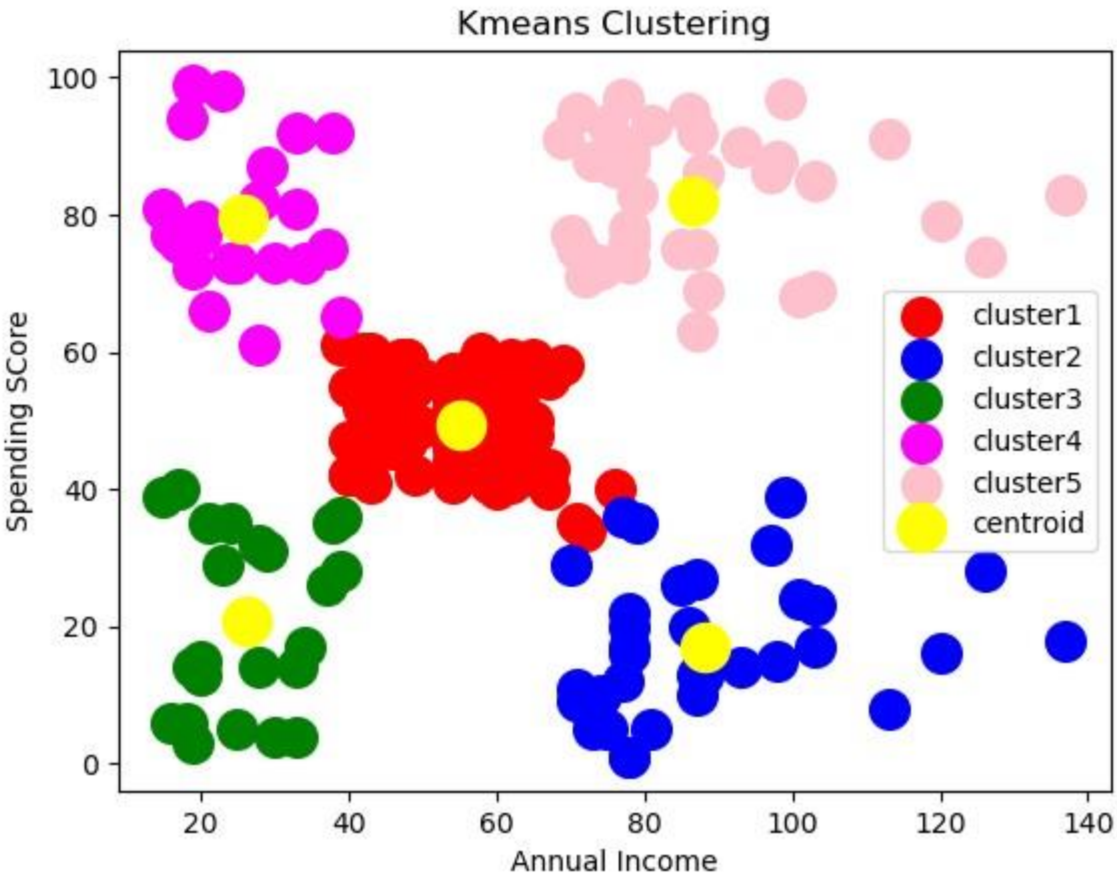
In [15]:
```
y_kmeans
```

Out[15]:
```
array([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, 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])
```

In [16]:
```
y_kmeans.shape
```

Out[16]:
```
(200,)
```

In [17]:
```
plt.scatter(X[y_kmeans==0,0],X[y_kmeans==0,1],s=200,c='red',label='cluster1')
plt.scatter(X[y_kmeans==1,0],X[y_kmeans==1,1],s=200,c='blue',label='cluster2')
plt.scatter(X[y_kmeans==2,0],X[y_kmeans==2,1],s=200,c='green',label='cluster3')
plt.scatter(X[y_kmeans==3,0],X[y_kmeans==3,1],s=200,c='magenta',label='cluster4'
plt.scatter(X[y_kmeans==4,0],X[y_kmeans==4,1],s=200,c='pink',label='cluster5')
plt.scatter(kmeans.cluster_centers_[:,0],kmeans.cluster_centers_[:,1],s=300,c='y
plt.title("Kmeans  Clustering")  plt.xlabel('Annual  Income')  plt.ylabel('Spending
SCore') plt.legend() plt.show()
```



In [18]:

```
import matplotlib.cm as cm from sklearn.metrics import
silhouette_samples, silhouette_score import numpy as np
```

In [19]:

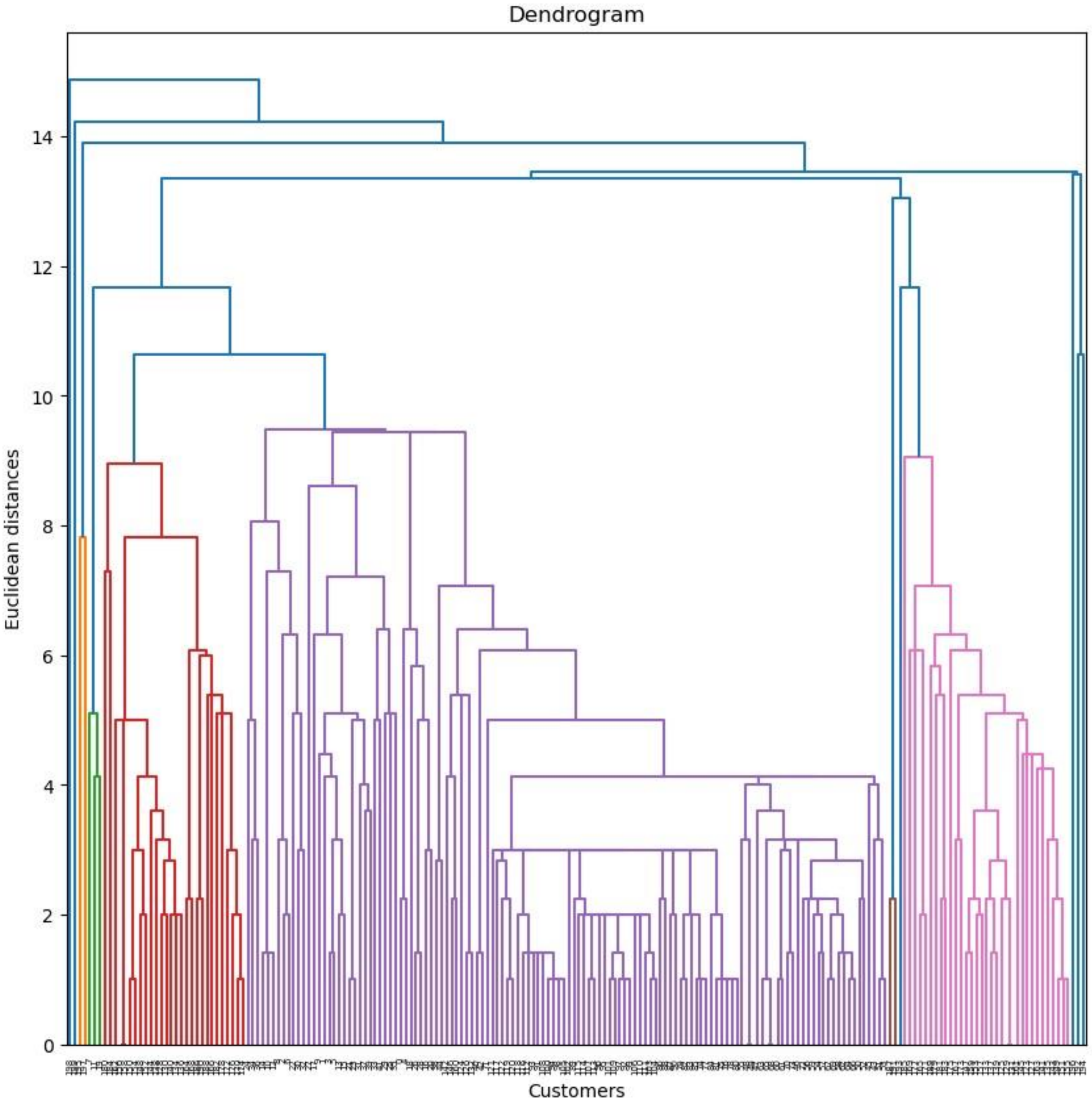
h

```
range_n_clusters = [2, 3, 4, 5, 6]

for n_clusters in range_n_clusters:
     # Initialize the clusterer with n_clusters value and a random generator
    # seed of 10 for reproducibility.     clusterer =
KMeans(n_clusters=n_clusters, random_state=10)
cluster_labels = clusterer.fit_predict(X)

    # The silhouette_score gives the average value for all the samples.
    # This gives a perspective into the density and separation of the formed clu
silhouette_avg = silhouette_score(X, cluster_labels)     print("For n_clusters
=",n_clusters,"The average silhouette_score is :",sil
```

```
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
For n_clusters = 2 The average silhouette_score is : 0.2968969162503008
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870:
Futur eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
For n_clusters = 3 The average silhouette_score is : 0.46761358158775435

C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
For n_clusters = 4 The average silhouette_score is : 0.4931963109249047

C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
For n_clusters = 5 The average silhouette_score is : 0.553931997444648

C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:870: Futur
eWarning: 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(
C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1382: User
Warning: KMeans is known to have a memory leak on Windows with MKL, when there ar e
less chunks than available threads. You can avoid it by setting the environment
variable OMP_NUM_THREADS=1.
  warnings.warn(
For n_clusters = 6 The average silhouette_score is : 0.5376203956398481 In
```

[20]:

```
import scipy.cluster.hierarchy as sch
plt.figure(figsize=(10,10)) dendrogram =
sch.dendrogram(sch.linkage(X, method = 'single'))
plt.title('Dendrogram') plt.xlabel('Customers')
plt.ylabel('Euclidean distances') plt.show()
```



In [21]:

```
from sklearn.cluster import AgglomerativeClustering hc =
AgglomerativeClustering(n_clusters = 5, affinity = 'euclidean', linkage = y_hc =
hc.fit_predict(X)
```

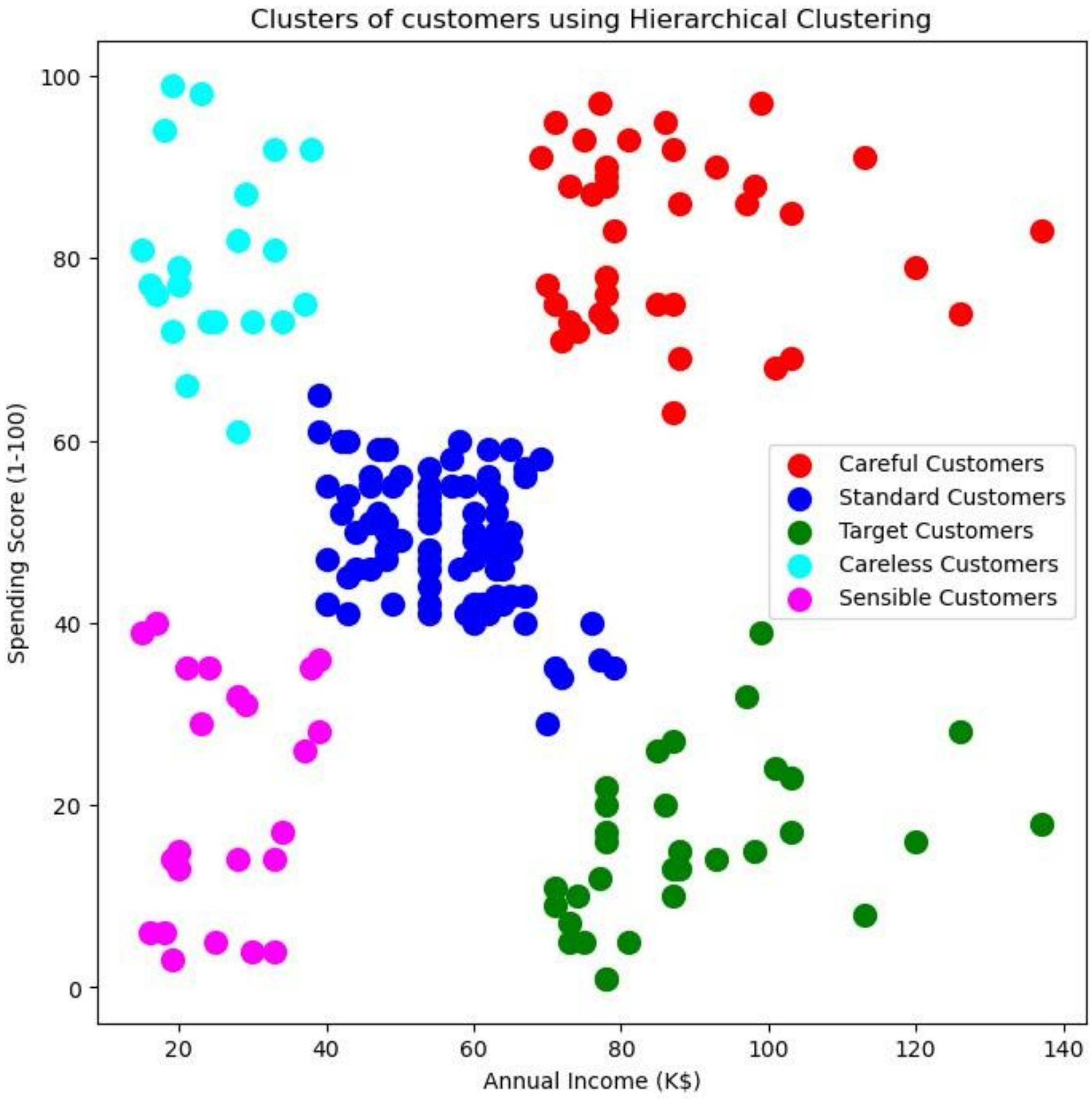
```
    C:\Users\PARTH\anaconda3\Lib\site-packages\sklearn\cluster\_agglomerative.py:98
```

```
3: FutureWarning: Attribute `affinity` was deprecated in version 1.2 and will be
removed in 1.4. Use `metric` instead   warnings.warn(
```

In [22]:

```
# Visualising the clusters

plt.figure(figsize=(8,8)) plt.scatter(X[y_hc == 0, 0], X[y_hc == 0, 1], s = 100,
c = 'red', label = 'Care plt.scatter(X[y_hc == 1, 0], X[y_hc == 1, 1], s = 100, c
= 'blue', label = 'Sta plt.scatter(X[y_hc == 2, 0], X[y_hc == 2, 1], s = 100, c =
'green', label = 'Ta plt.scatter(X[y_hc == 3, 0], X[y_hc == 3, 1], s = 100, c =
'cyan', label = 'Car plt.scatter(X[y_hc == 4, 0], X[y_hc == 4, 1], s = 100, c =
'magenta', label = ' plt.title('Clusters of customers using Hierarchical
Clustering') plt.xlabel('Annual Income (K$)') plt.ylabel('Spending Score (1-
100)') plt.legend() plt.show()
```

```
f
n
r
e
S
```



```
In [ ]: In [ ]:
```