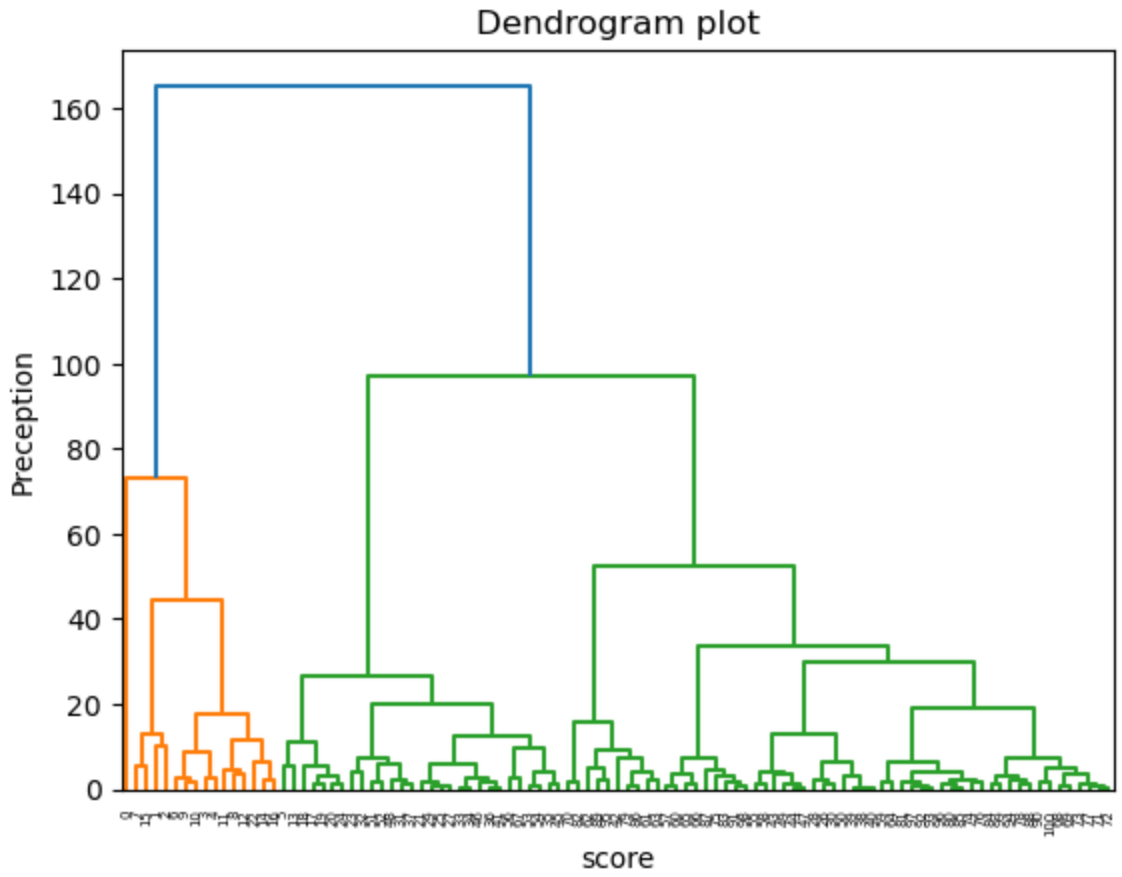
Hierarchical clustering

In [38]:
```
#importing the libraries
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd

#importing the dataset
dataset=pd.read_csv('UniversityRanking_2022.csv')
x=dataset.iloc[:,[4,10]].values

#Finding the optimal number of clusters using the dendrogram
import scipy.cluster.hierarchy as shc
dendro=shc.dendrogram(shc.linkage(x,method="ward"))
mtp.title("Dendrogram plot")
mtp.ylabel("Preception")
mtp.xlabel("score")
mtp.show()
```



In [39]:
```
#training the hierarchical model on dataset
from sklearn.cluster import AgglomerativeClustering
hc=AgglomerativeClustering(n_clusters=5,affinity='euclidean',linkage='ward')
y_pred=hc.fit_predict(x)
```

```
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_agglomerative.py:1005: FutureWarning: Attribute `affinity` was deprecated in version 1.2 and will be removed in 1.4. Use `metric` instead
  warnings.warn(
```

In [37]:
```
mtp.scatter(x[y_pred==0,0],x[y_pred==0,1],s=100,c='blue',label='cluster1')
mtp.scatter(x[y_pred==1,0],x[y_pred==1,1],s=100,c='green',label='cluster2')
mtp.scatter(x[y_pred==2,0],x[y_pred==2,1],s=100,c='red',label='cluster3')
```

```
mtp.scatter(x[y_pred==3,0],x[y_pred==3,1],s=100,c='cyan',label='cluster4')
mtp.scatter(x[y_pred==4,0],x[y_pred==4,1],s=100,c='magenta',label='cluster5')
mtp.title('clusters of customers')
mtp.xlabel('Rank')
mtp.ylabel('Preception')
mtp.legend()
mtp.show()
```

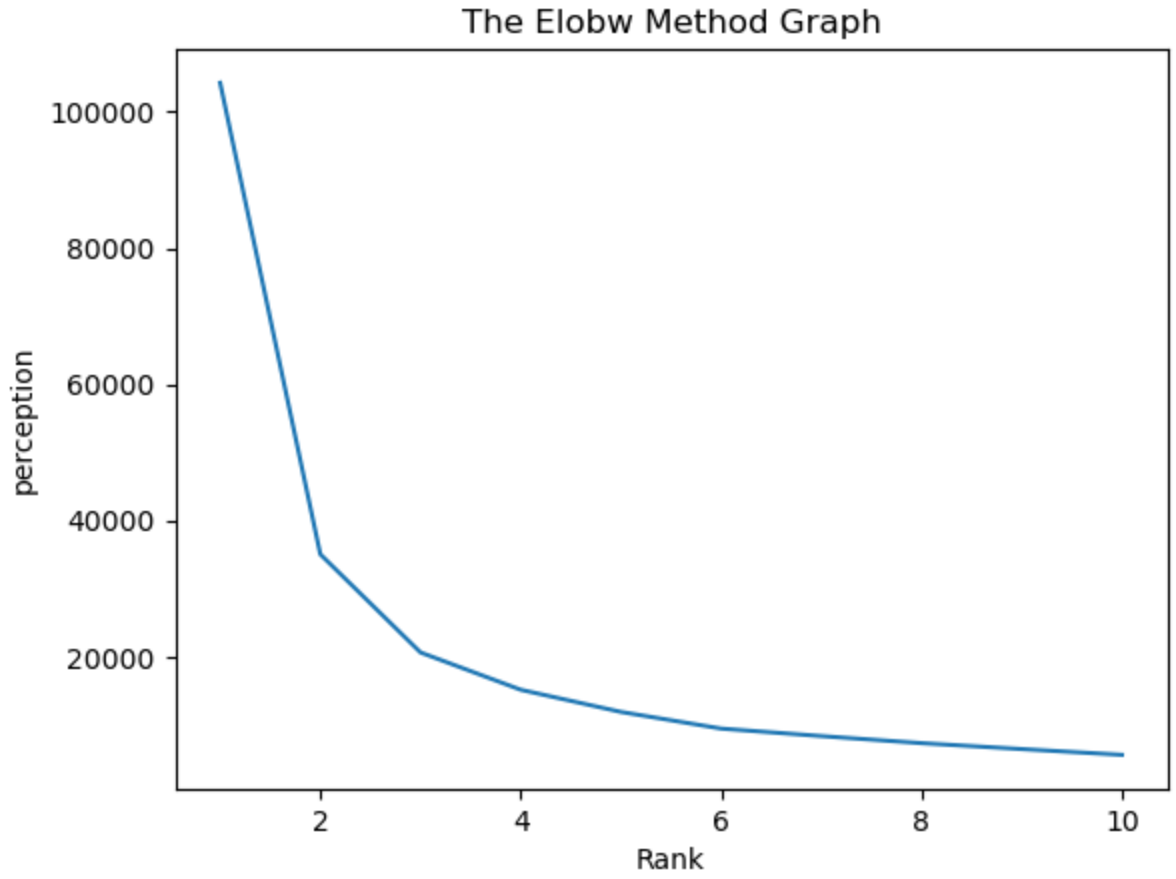
K-Mean Algorithm

In [8]:
```
# importing libraries
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd

# Importing the dataset
dataset = pd.read_csv('UniversityRanking_2022.csv')

x = dataset.iloc[:, [5,10]].values
```

In [9]:
```
#finding optimal number of clusters using the elbow method
from sklearn.cluster import KMeans
wcss_list= []  #Initializing the list for the values of WCSS

#Using for loop for iterations 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 Elobw Method Graph')
mtp.xlabel('Rank')
mtp.ylabel('perception')
```

```
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
```

```
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
```

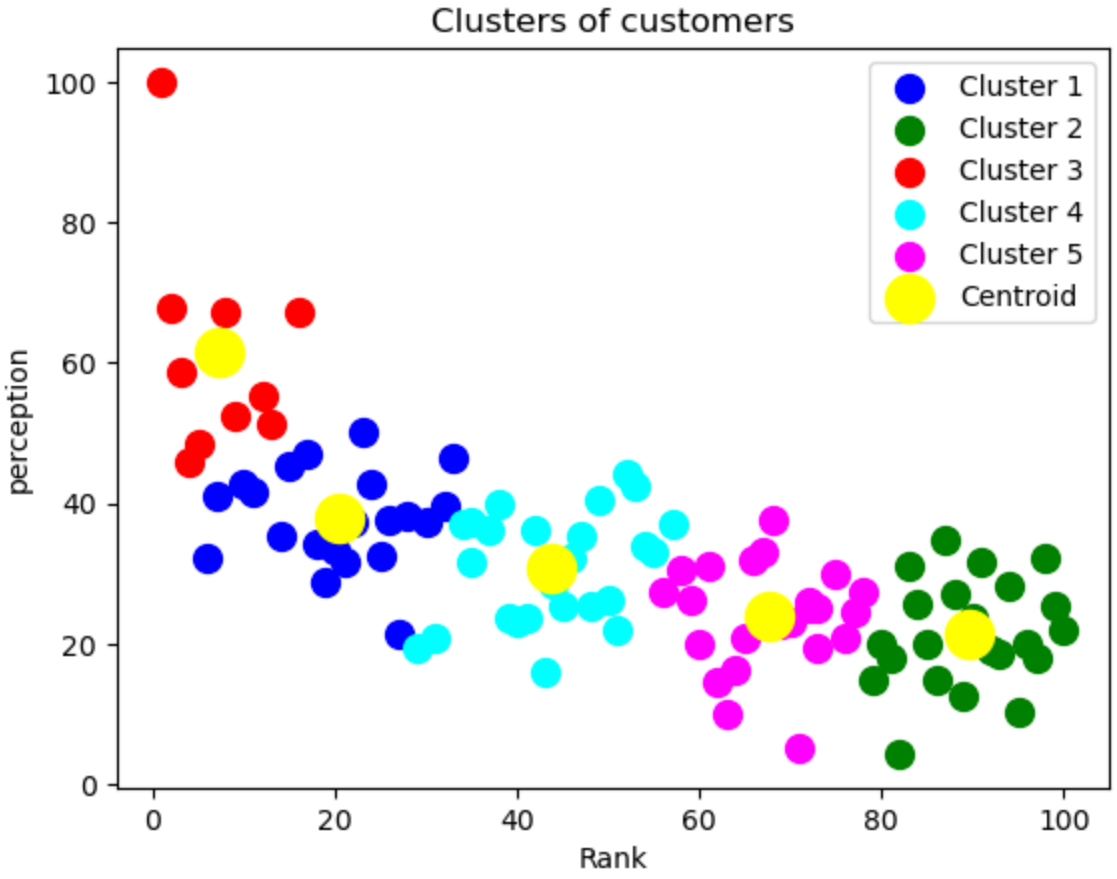


In [10]:
```
#training the K-means model on a dataset
kmeans = KMeans(n_clusters=5, init='k-means++', random_state= 42)
y_predict= kmeans.fit_predict(x)
```

```
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1412: Futu
reWarning: The default value of `n_init` will change from 10 to 'auto' in 1.4. Set the v
alue of `n_init` explicitly to suppress the warning
  super()._check_params_vs_input(X, default_n_init=10)
C:\Users\R.MUNIRANJANI\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1436: User
Warning: 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=1.
  warnings.warn(
```

In [11]:

```
#visulaizing the clusters
mtp.scatter(x[y_predict == 0, 0], x[y_predict == 0, 1], s = 100, c = 'blue', label = 'Cl
mtp.scatter(x[y_predict == 1, 0], x[y_predict == 1, 1], s = 100, c = 'green', label = 'C
mtp.scatter(x[y_predict== 2, 0], x[y_predict == 2, 1], s = 100, c = 'red', label = 'Clus
mtp.scatter(x[y_predict == 3, 0], x[y_predict == 3, 1], s = 100, c = 'cyan', label = 'Cl
mtp.scatter(x[y_predict == 4, 0], x[y_predict == 4, 1], s = 100, c = 'magenta', label =
mtp.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1], s = 300, c = '
mtp.title('Clusters of customers')
mtp.xlabel('Rank')
mtp.ylabel('perception')
mtp.legend()
mtp.show()
```



In [ ]: