# Karanjot Singh

In [2]:

```
import pandas as pd
from sklearn import datasets
```

In [4]:

```
df=datasets.load_iris()
```

In [5]:

```
x=df.data
y=df.target
```

In [6]:

```
from sklearn.cluster import KMeans
```

In [8]:

```
clust=KMeans(n_clusters=5)
```

In [9]:

```
clust.fit(x)
```

Out[9]:

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=5, n_init=10, n_jobs=None, precompute_distances='auto',
       random_state=None, tol=0.0001, verbose=0)
```

In [11]:

```
pred=clust.predict(x)
pred
```

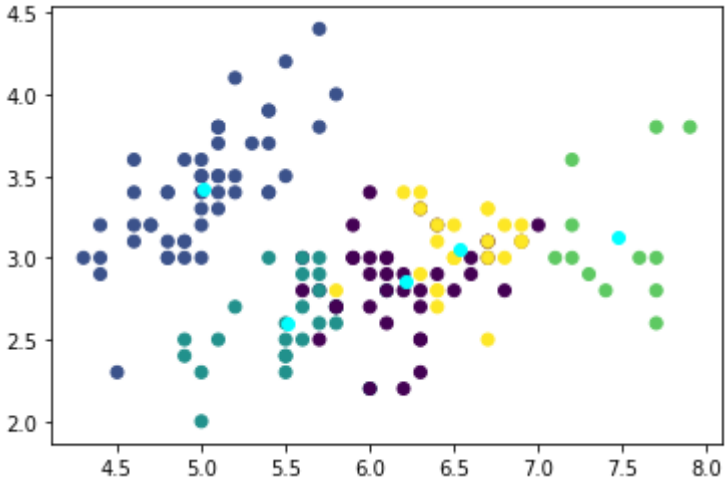
Out[11]:

```
array([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 0, 0, 0, 2, 0, 0, 0, 2, 0, 2, 2, 0, 2, 0, 2, 0,
       0, 2, 0, 2, 0, 2, 0, 0, 0, 0, 0, 0, 0, 2, 2, 2, 2, 0, 2, 0, 0, 0,
       2, 2, 2, 0, 2, 2, 2, 2, 2, 0, 2, 2, 4, 0, 3, 4, 4, 3, 2, 3, 4, 3,
       4, 4, 4, 0, 4, 4, 4, 3, 3, 0, 4, 0, 3, 0, 4, 3, 0, 0, 4, 3, 3, 3,
       4, 0, 0, 3, 4, 4, 0, 4, 4, 4, 0, 4, 4, 4, 0, 4, 4, 0])
```

In [14]:

```
import matplotlib.pyplot as plt
plt.scatter(x[:,0],x[:,1],c=pred)
centers=clust.cluster_centers_
plt.scatter(centers[:,0],centers[:,1],c='cyan')
print(centers)
```

```
[[6.20769231 2.85384615 4.74615385 1.56410256]
 [5.006      3.428      1.462      0.246     ]
 [5.508      2.6        3.908      1.204     ]
 [7.475      3.125      6.3        2.05      ]
 [6.52916667 3.05833333 5.50833333 2.1625    ]]
```




In [15]:

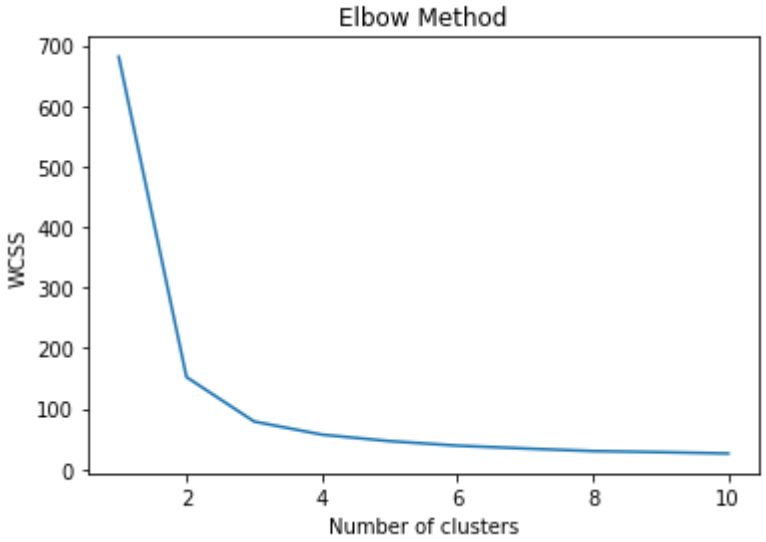
```
from sklearn import metrics
from scipy.spatial.distance import cdist
import seaborn as sns
```

In [16]:

```
wcss = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++',n_init=10, random_state=0)
    kmeans.fit(x)
    wcss.append(kmeans.inertia_)
pl=sns.lineplot(range(1, 11), wcss)
pl.set(title='Elbow Method',xlabel='Number of clusters',ylabel='WCSS')
```

Out[16]:

```
[Text(0, 0.5, 'WCSS'),
 Text(0.5, 0, 'Number of clusters'),
 Text(0.5, 1.0, 'Elbow Method')]
```




In [17]:

```
wcss
```

Out[17]:

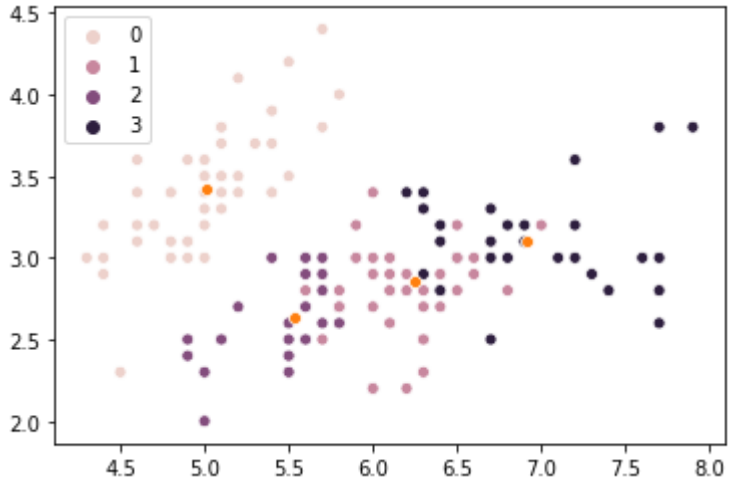
```
[681.3706,
 152.34795176035792,
 78.85144142614601,
 57.228473214285714,
 46.472230158730156,
 39.03998724608725,
 34.299712121212124,
 30.063110617452725,
 28.27172172856384,
 26.094324740540415]
```

In [19]:

```
kmeans = KMeans(n_clusters=4)
prediction = kmeans.fit_predict(x)
sns.scatterplot(x[:,0], x[:,1], hue=prediction)
sns.scatterplot(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1])
```

Out[19]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x270ba151b08>
```




In [20]:

```
prediction
```

Out[20]:

```
array([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, 1, 1, 1, 2, 1, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 1,
       2, 2, 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 1, 2, 1, 1, 1,
       2, 2, 2, 1, 2, 2, 2, 2, 2, 1, 2, 2, 3, 1, 3, 3, 3, 3, 2, 3, 3, 3,
       1, 1, 3, 1, 1, 3, 3, 3, 3, 1, 3, 1, 3, 1, 3, 3, 1, 1, 3, 3, 3, 3,
       3, 1, 1, 3, 3, 3, 1, 3, 3, 3, 1, 3, 3, 3, 1, 1, 3, 1])
```