```
Name - Pooja Gandhi
           Task2- Prediction using Unsupervised Machine Learning:K-
           means clustering
           Importing Libraries
In [1]: import numpy as np
           import matplotlib.pyplot as plt
           import pandas as pd
           from sklearn import datasets
           Loading Iris Data
In [2]: iris = datasets.load_iris()
           iris_df = pd.DataFrame(iris.data, columns = iris.feature_names)
           iris_df.head(10)
 Out[2]:
              sepal length (cm) sepal width (cm) petal length (cm) petal width (cm)
                          5.1
                                         3.5
                                                        1.4
                                                                       0.2
           1
                          4.9
                                         3.0
                                                        1.4
                                                                       0.2
           2
                          4.7
                                         3.2
                                                        1.3
                                                                       0.2
           3
                                                        1.5
                                                                       0.2
                          4.6
                                         3.1
                          5.0
                                         3.6
                                                        1.4
                                                                       0.2
           5
                          5.4
                                         3.9
                                                        1.7
                                                                       0.4
                          4.6
                                         3.4
                                                        1.4
                                                                       0.3
           7
                                                                       0.2
                          5.0
                                         3.4
                                                        1.5
                          4.4
                                         2.9
                                                        1.4
                                                                       0.2
                          4.9
                                         3.1
                                                        1.5
                                                                       0.1
In [3]: iris_df.info()
           <class 'pandas.core.frame.DataFrame'>
           RangeIndex: 150 entries, 0 to 149
           Data columns (total 4 columns):
           sepal length (cm) 150 non-null float64
                                   150 non-null float64
           sepal width (cm)
           petal length (cm) 150 non-null float64
           petal width (cm)
                                   150 non-null float64
           dtypes: float64(4)
           memory usage: 4.8 KB
In [4]: iris_df.describe()
 Out[4]:
                  sepal length (cm) sepal width (cm) petal length (cm) petal width (cm)
                       150.000000
                                                                    150.000000
                                      150.000000
                                                     150.000000
           count
                         5.843333
                                        3.057333
                                                       3.758000
            mean
                                                                      1.199333
              std
                         0.828066
                                        0.435866
                                                       1.765298
                                                                      0.762238
                         4.300000
                                        2.000000
                                                       1.000000
                                                                      0.100000
             min
                                                                      0.300000
                                        3.000000
                                                       4.350000
             50%
                         5.800000
                                                                      1.300000
             75%
                         6.400000
                                        3.300000
                                                       5.100000
                                                                      1.800000
                         7.900000
                                        4.400000
                                                       6.900000
                                                                      2.500000
             max
In [5]: iris_df.columns
Out[5]: Index([u'sepal length (cm)', u'sepal width (cm)', u'petal length (cm)',
                   u'petal width (cm)'],
                  dtype='object')
In [6]: iris_df.isnull()
Out[6]:
                sepal length (cm) sepal width (cm) petal length (cm) petal width (cm)
                          False
                                         False
                                                        False
                                                                       False
             1
                          False
                                         False
                                                        False
                                                                       False
             2
                                         False
                                                        False
                          False
                                                                       False
             3
                          False
                                         False
                                                        False
                                                                       False
                          False
                                         False
                                                        False
                                                                       False
             5
                          False
                                         False
                                                        False
                                                                       False
              6
                          False
                                         False
                                                        False
                                                                       False
             7
                          False
                                         False
                                                        False
                                                                       False
                          False
                                         False
                                                        False
                                                                       False
             9
                          False
                                         False
                                                        False
                                                                       False
            10
                                         False
                                                        False
                          False
                                                                       False
            11
                          False
                                         False
                                                        False
                                                                       False
            12
                          False
                                         False
                                                        False
                                                                       False
            13
                          False
                                         False
                                                        False
                                                                       False
            14
                          False
                                         False
                                                        False
                                                                       False
            15
                          False
                                         False
                                                        False
                                                                       False
            16
                          False
                                         False
                                                        False
                                                                       False
            17
                          False
                                         False
                                                        False
                                                                       False
                                                        False
            18
                          False
                                         False
                                                                       False
            19
                          False
                                         False
                                                        False
                                                                       False
            20
                          False
                                         False
                                                        False
                                                                       False
            21
                          False
                                         False
                                                        False
                                                                       False
            22
                          False
                                         False
                                                        False
                                                                       False
            23
                          False
                                         False
                                                        False
                                                                       False
            24
                          False
                                         False
                                                        False
                                                                       False
            25
                          False
                                         False
                                                        False
                                                                       False
            26
                          False
                                         False
                                                        False
                                                                       False
            27
                          False
                                         False
                                                        False
                                                                       False
            28
                          False
                                         False
                                                        False
                                                                       False
            29
                          False
                                         False
                                                        False
                                                                       False
            120
                          False
                                         False
                                                        False
                                                                       False
           121
                          False
                                                        False
                                         False
                                                                       False
            122
                          False
                                         False
                                                        False
                                                                       False
           123
                          False
                                         False
                                                         False
                                                                       False
            124
                          False
                                         False
                                                         False
                                                                       False
           125
                          False
                                         False
                                                         False
                                                                       False
           126
                          False
                                         False
                                                         False
                                                                       False
           127
                          False
                                         False
                                                        False
                                                                       False
           128
                          False
                                         False
                                                        False
                                                                       False
           129
                          False
                                         False
                                                        False
                                                                       False
           130
                          False
                                         False
                                                        False
                                                                       False
           131
                                         False
                                                         False
                          False
                                                                       False
           132
                          False
                                         False
                                                        False
                                                                       False
           133
                          False
                                         False
                                                        False
                                                                       False
           134
                          False
                                         False
                                                        False
                                                                       False
           135
                          False
                                         False
                                                         False
                                                                       False
           136
                          False
                                         False
                                                        False
                                                                       False
           137
                          False
                                         False
                                                        False
                                                                       False
           138
                          False
                                         False
                                                        False
                                                                       False
           139
                          False
                                         False
                                                         False
                                                                       False
           140
                          False
                                         False
                                                        False
                                                                       False
           141
                          False
                                         False
                                                         False
                                                                       False
           142
                          False
                                         False
                                                        False
                                                                       False
           143
                          False
                                         False
                                                        False
                                                                       False
           144
                          False
                                         False
                                                        False
                                                                       False
           145
                                                        False
                          False
                                         False
                                                                       False
           146
                          False
                                         False
                                                        False
                                                                       False
           147
                                         False
                                                         False
                          False
                                                                       False
           148
                          False
                                         False
                                                        False
                                                                       False
           149
                          False
                                         False
                                                         False
                                                                       False
           150 rows × 4 columns
In [7]: iris_df.isnull().sum()
 Out[7]: sepal length (cm)
                                    0
           sepal width (cm)
           petal length (cm)
                                    0
           petal width (cm)
           dtype: int64
          iris_df.corr()
In [8]:
 Out[8]:
                           sepal length (cm) sepal width (cm) petal length (cm) petal width (cm)
            sepal length (cm)
                                  1.000000
                                                 -0.117570
                                                                 0.871754
                                                                               0.817941
                                  -0.117570
                                                 1.000000
                                                                -0.428440
                                                                               -0.366126
            sepal width (cm)
            petal length (cm)
                                  0.871754
                                                 -0.428440
                                                                 1.000000
                                                                                0.962865
                                  0.817941
                                                 -0.366126
                                                                 0.962865
                                                                               1.000000
            petal width (cm)
           Elbow Method for Finding the number of clusters in KMeans Clustering
In [9]: x = iris_df.iloc[:, [0, 1, 2, 3]].values
           from sklearn.cluster import KMeans
           wcss = []
           for i in range(1, 11):
                kmeans = KMeans(n_clusters = i, init = 'k-means++',
                                  max_iter = 300, n_init = 10, random_state = 0)
                kmeans.fit(x)
                wcss.append(kmeans.inertia_)
In [10]: # Plotting the results onto a line graph,
           # `allowing us to observe 'The elbow'
           plt.plot(range(1, 11), wcss)
           plt.title('The elbow method')
           plt.xlabel('Number of clusters')
           plt.ylabel('WCSS') # Within cluster sum of squares
           plt.show()
                                 The elbow method
              700
              600
              500
            SS 400
300
              200
              100
                                  Number of clusters
           Applying KMeans Cluster
```

In [12]: y_kmeans 1, 1, 1, 1, 1, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

y_kmeans = kmeans.fit_predict(x)

```
2, 0, 2, 2, 2, 0, 2, 2, 0, 2, 2, 2, 0, 2, 2, 0])
Visualisation of Clusters
```

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 2, 0, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 0, 0, 2, 2, 2, 0, 2, 0, 2, 0, 2, 2, 0, 0, 2, 2, 2, 2,

In [11]: kmeans = KMeans(n_clusters = 3, init = 'k-means++', max_iter = 300, n_init = 10, random_state

```
In [13]: # Visualising the clusters - On the first two columns
         plt.scatter(x[y_kmeans == 0, 0], x[y_kmeans == 0, 1], s = 100, c = 'red', label = 'Iris-seto
         plt.scatter(x[y_kmeans == 1, 0], x[y_kmeans == 1, 1], s = 100, c = 'blue', label = 'Iris-ver
```

```
sicolour')
plt.scatter(x[y_kmeans == 2, 0], x[y_kmeans == 2, 1], s = 100, c = 'green', label = 'Iris-vir
ginica')
# Plotting the centroids of the clusters
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:,1],
```

s = 100, c = 'yellow', label = 'Centroids')

plt.legend() Out[13]: <matplotlib.legend.Legend at 0xd6ce4a8> 4.0 3.5

3.0 2.5 2.0 Centroids

Thankyou

In []: