PRACTICAL 8

Perform the data clustering using clustering algorithm.

k-means clustering using R

#apply K means to iris and store result

newiris <- iris

newiris$Species <- NULL

(kc <- kmeans(newiris,3))

```
K-means clustering with 3 clusters of sizes 21, 96, 33

Cluster means:
  Sepal.Length Sepal.Width Petal.Length Petal.Width
1     4.738095    2.904762     1.790476    0.3523810
2     6.314583    2.895833     4.973958    1.7031250
3     5.175758    3.624242     1.472727    0.2727273

Clustering vector:
   [1] 3 1 1 1 3 3 3 3 1 1 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 1 1 3 3 3 1 3 3 3 1
  [40] 3 3 1 1 3 3 1 3 1 3 3 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
  [79] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 [118] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

Within cluster sum of squares by cluster:
[1]  17.669524 118.651875   6.432121
 (between_SS / total_SS =  79.0 %)

Available components:

[1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
[6] "betweenss"    "size"         "iter"         "ifault"
```

#Compare the Species label with the clustering result

table(iris$Species,kc$cluster)

```
              1  2  3
  setosa     17  0 33
  versicolor  4 46  0
  virginica   0 50  0
```

#Plot the clusters and their centers

plot(newiris[c("Sepal.Length","Sepal.Width")],col=kc$cluster)

points(kc$centers[,c("Sepal.Length","Sepal.Width")],col=1:3,pch=8,cex=2)

dev.off()

#Plot the clusters and their centre

plot(newiris[c("Sepal.Length","Sepal.Width")],col=kc$cluster)