> head(iris)

Sepal.Length Sepal.Width Petal.Length Petal.Width Species

1 5.1 3.5 1.4 0.2 setosa

2 4.9 3.0 1.4 0.2 setosa

3 4.7 3.2 1.3 0.2 setosa

4 4.6 3.1 1.5 0.2 setosa

5 5.0 3.6 1.4 0.2 setosa

6 5.4 3.9 1.7 0.4 setosa

> kc<-kmeans(iris[,-5],3)

> kc

K-means clustering with 3 clusters of sizes 62, 38, 50

Cluster means:

Sepal.Length Sepal.Width Petal.Length Petal.Width

1 5.901613 2.748387 4.393548 1.433871

2 6.850000 3.073684 5.742105 2.071053

3 5.006000 3.428000 1.462000 0.246000

Clustering vector:

[1] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

[41] 3 3 3 3 3 3 3 3 3 3 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1

[81] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 2 2 2 1 2 2 2 2 2 2 1 1 2 2 2 2 1

[121] 2 1 2 1 2 2 1 1 2 2 2 2 2 1 2 2 2 2 1 2 2 2 1 2 2 2 1 2 2 1

Within cluster sum of squares by cluster:

[1] 39.82097 23.87947 15.15100

(between\_SS / total\_SS = 88.4 %)

Available components:

[1] "cluster" "centers" "totss" "withinss" "tot.withinss"

[6] "betweenss" "size" "iter" "ifault"

> table(iris$Species,kc$cluster)

1 2 3

setosa 0 0 50

versicolor 48 2 0

virginica 14 36 0

> x=iris[,-5]

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

> points(kc$centers[,c("Sepal.Length","Sepal.Width")], col=1:3,pch=23,cex=3)

> clusters <- hclust(dist(iris[, 3:4]))

> plot(clusters)

> clusterCut <- cutree(clusters, 3)

> table(clusterCut, iris$Species)

clusterCut setosa versicolor virginica

1 50 0 0

2 0 21 50

3 0 29 0

> clusters <- hclust(dist(iris[, 3:4]), method = 'average')

> plot(clusters)

> clusterCut <- cutree(clusters, 3)

> table(clusterCut, iris$Species)

clusterCut setosa versicolor virginica

1 50 0 0

2 0 45 1

3 0 5 49