

Question 9

Niraj Raj Kharel (16)

2024-05-31

Question 9

Question 9

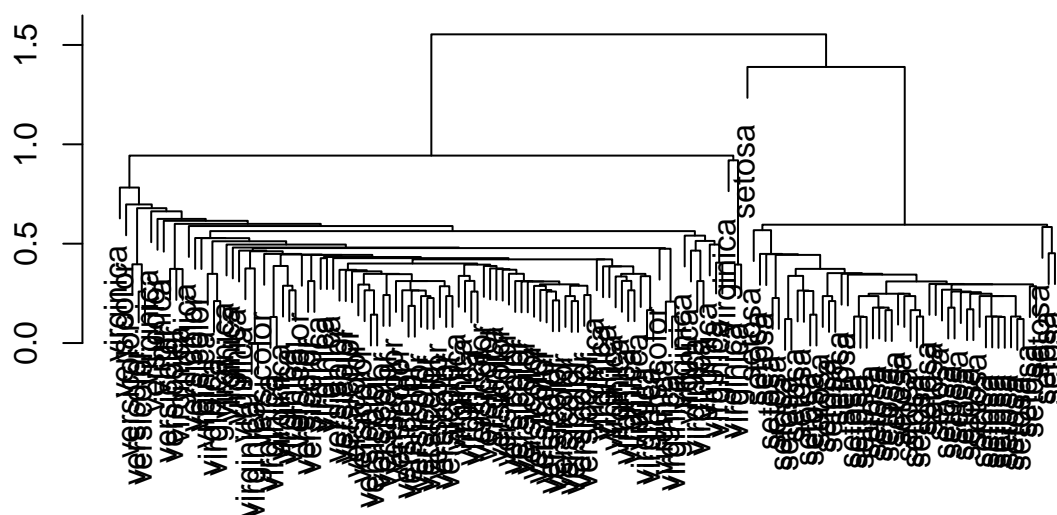
```
data(iris)
ir_data <- iris
ir_label <- iris$Species
ir_data <- ir_data[, -5]
head(ir_data)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1          5.1          3.5          1.4          0.2
## 2          4.9          3.0          1.4          0.2
## 3          4.7          3.2          1.3          0.2
## 4          4.6          3.1          1.5          0.2
## 5          5.0          3.6          1.4          0.2
## 6          5.4          3.9          1.7          0.4
```

#a

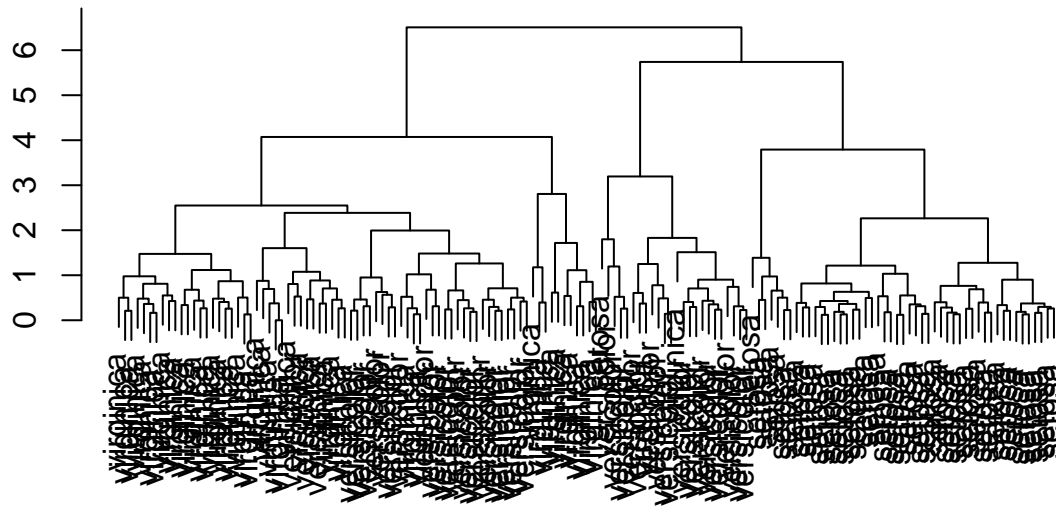
```
sd.data <- scale(ir_data)
data.dist <- dist(sd.data)
plot(hclust(data.dist, method = "single"), xlab = "", sub = "", ylab = "",
     labels = ir_label, main = "Single Linkage")
```

Single Linkage



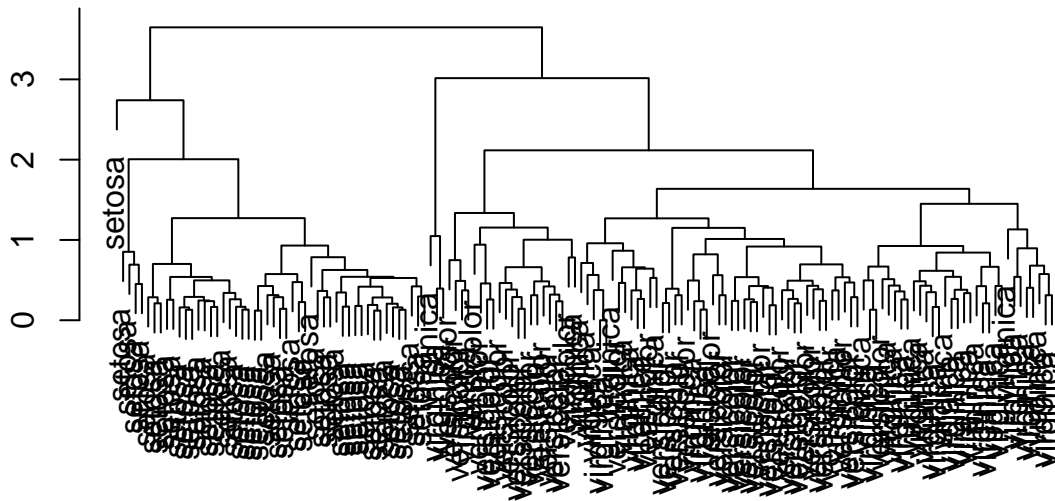
```
#b
plot(hclust(data.dist), xlab = "", sub = "", ylab = "",
     labels = ir_label, main = "Complete Linkage")
```

Complete Linkage



```
#c
plot(hclust(data.dist, method = "average"), xlab = "", sub = "", ylab = "",
      labels = ir_label, main = "Average Linkage")
```

Average Linkage

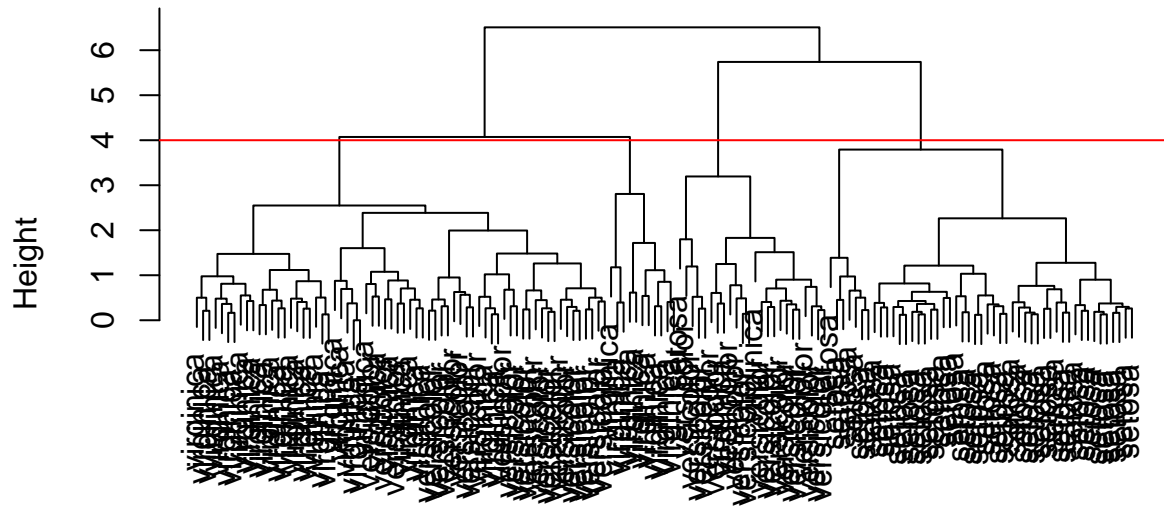


```
#d
# Best hierarchical model
hc.out <- hclust(dist(sd.data))
hc.clusters <- cutree(hc.out, 4)
table(hc.clusters, ir_label)

##          ir_label
## hc.clusters setosa versicolor virginica
##          1      49           0           0
##          2       1          21           2
##          3       0          29          37
##          4       0           0          11

par(mfrow = c(1, 1))
plot(hc.out, labels = ir_label)
abline(h = 4, col = "red")
```

Cluster Dendrogram



```
dist(sd.data)  
hclust (*, "complete")
```