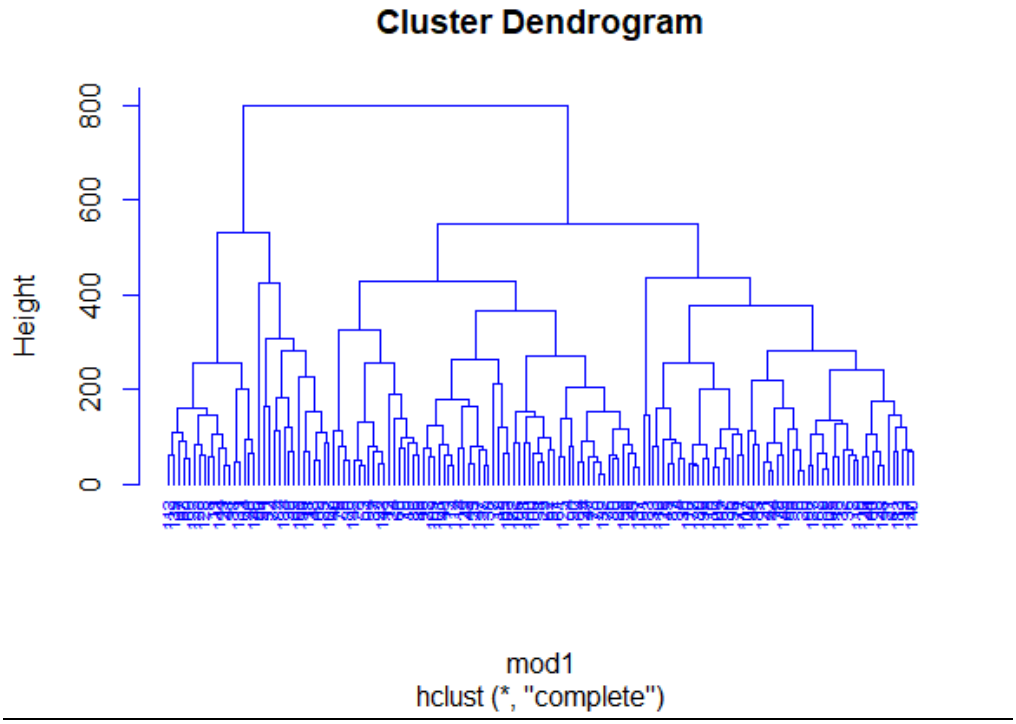
Problem 16.6

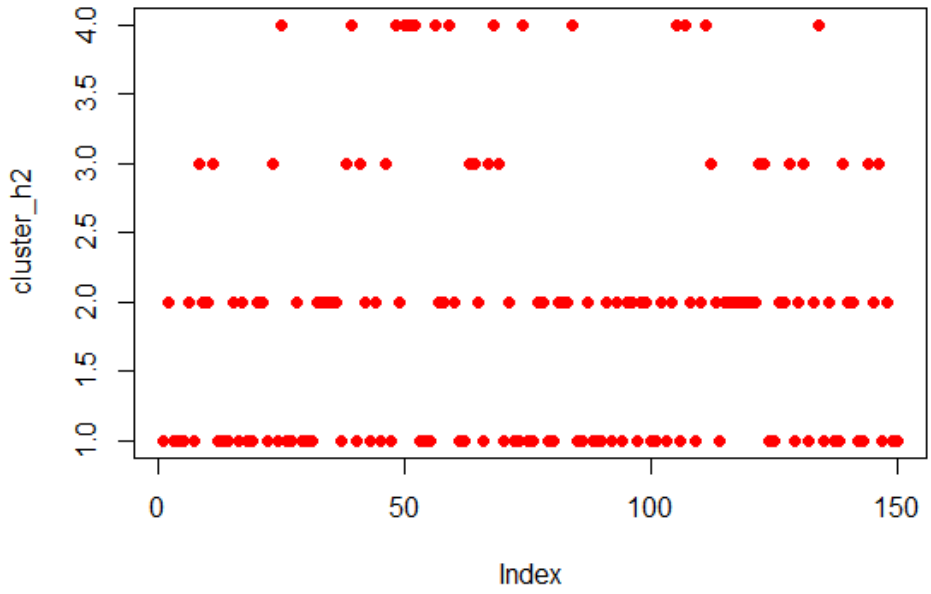
Code:

```
library(cluster)

Lung_d<-data.frame(Lung_Function$fev1father,Lung_Function$FEV1M,Lung_Function$FEV1_OLD,
                   Lung_Function$FFVC,Lung_Function$FVCM,Lung_Function$FVC_OLD)

mod1<-dist(Lung_d,method="euclidean")

cluster_h<-hclust(mod1,method="complete")
plot(cluster_h,hang = -4, cex = 0.6,col="blue")

cluster_h2<-cutree(cluster_h,4)
plot(cluster_h2,col="red",pch=16)

table(cluster_h2,Lung_Function$area)
```

Output:

```
cluster_h2  1  2  3  4
         1 13 19 10 20
         2  8 14  7 26
         3  3  7  0  8
         4  0  9  2  4
```





The 4 Cluster sizes form hierarchical clustering we get are 62, 55, 18 and 15.

Problem 16.8:

Code:

```
library(cluster)

Lung_d<-data.frame(Lung_Function$fev1father,Lung_Function$FEV1M,Lung_Function$FEV1_OLD,
                   Lung_Function$FFVC,Lung_Function$FVCM,Lung_Function$FVC_OLD)

K_means<-kmeans(Lung_d,4)

K_means

table(K_means$cluster, Lung_Function$area)
```

Output:

```
K-means clustering with 4 clusters of sizes 31, 58, 25, 36

Cluster means:
  Lung_Function.fev1father Lung_Function.FEV1M Lung_Function.FEV1_OLD Lung_Function.FFVC Lung_Function.FVCM Lung_Function.FVC_OLD
1                 332.7097            267.3548              262.9032           398.9032           311.2903             299.5484
2                 428.8103            304.4483              162.1207           514.6034           354.8966             185.9310
3                 424.6000            307.8800              441.1200           506.6000           368.2000             518.2000
4                 433.3056            304.2778              297.6667           539.0833           363.7778             350.3611

Clustering vector:
  [1] 1 4 2 1 1 4 2 3 4 2 3 1 2 2 2 2 4 1 2 4 4 2 3 2 3 4 2 2 1 1 1 4 2 2 2 4 2 3 3 1 3 4 2 4 1 1 1 3 4 3 3 3 2 1 2 3 2 4 3 2 2 1 4 1 4 2 4 3 4 1 2
 [72] 1 1 3 1 1 2 4 2 1 4 4 4 3 2 2 2 2 1 1 2 2 4 2 4 2 1 4 2 1 2 4 2 2 3 4 3 4 2 2 3 3 4 2 4 2 2 2 4 2 4 3 3 1 2 2 4 4 2 2 3 1 2 3 1 2 1 2 4 4 2 2
[143] 2 3 4 1 2 2 2 1

Within cluster sum of squares by cluster:
[1]  375419.0 1073140.4  493413.3  412972.1
 (between_SS / total_SS =  63.9 %)

Available components:

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

	1	2	3	4
1	5	13	5	8
2	10	14	8	26
3	3	11	2	9
4	6	11	4	15

The 4 cluster sizes we get after performing K-means clustering are 31, 58, 25, 36 which are uniformly distributed. The hierarchical clustering method gives us 4 clusters that where clusters 1 and 2 are densely populated whereas cluster 3 and 4 are sparsely populated.