Name: Shaikh Inamul Hasan

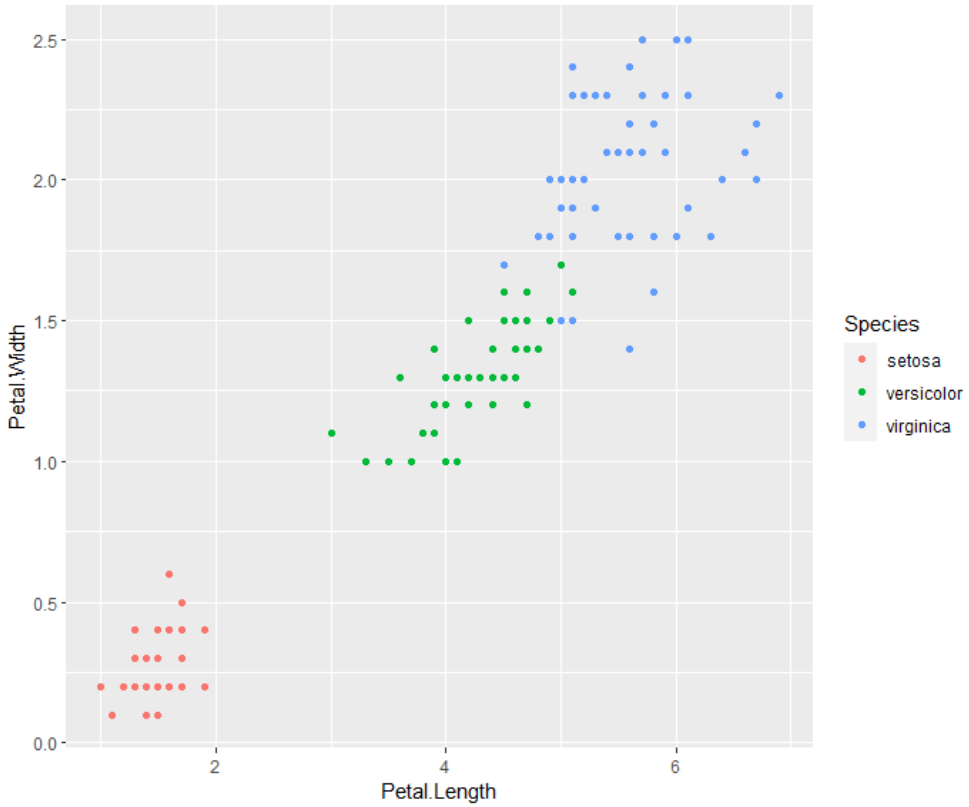
Roll No: 100

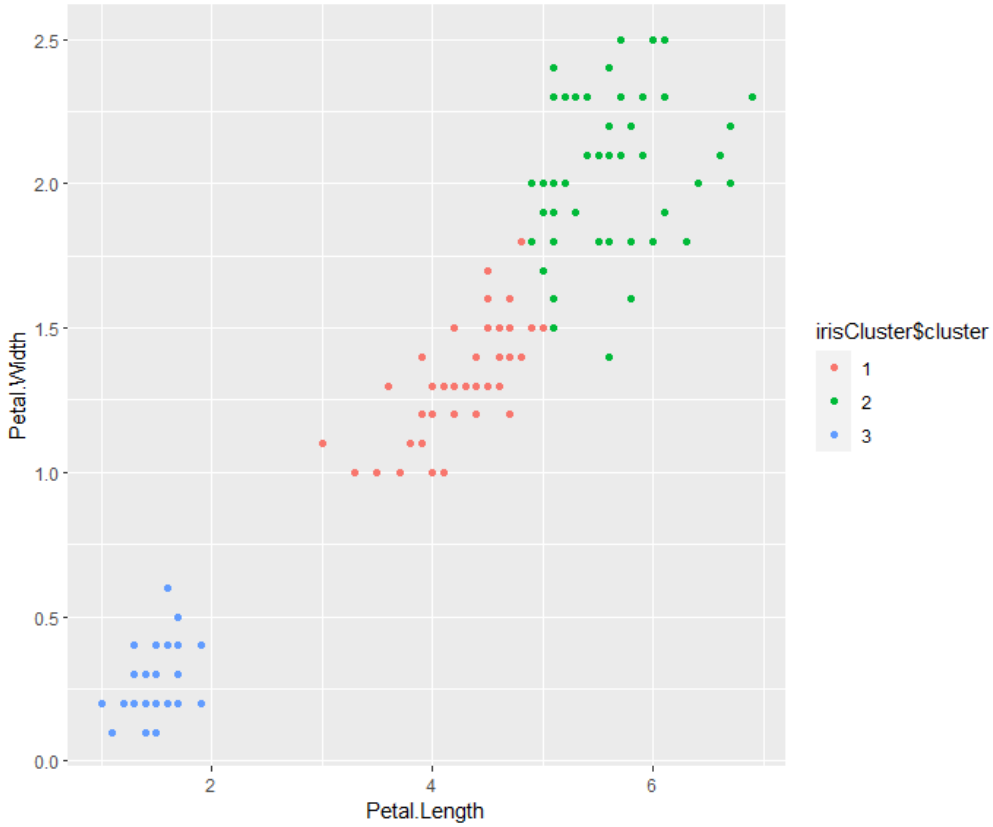
Lab 9: ADMS: practical Assignment: K-Means, Hierarchical clustering.

1. Unsupervised classification: Implementation and analysis of clustering algorithm K-Means.

Code & Output:

```
> Iris<- iris
> library(ggplot2)
> ggplot(iris, aes(Petal.Length, Petal.Width, color = Species)) + geom_point()
> set.seed(20)
> irisCluster <- kmeans(iris[, 3:4], 3, nstart = 20)
> irisCluster
K-means clustering with 3 clusters of sizes 52, 48, 50

Cluster means:
  Petal.Length Petal.Width
1     4.269231    1.342308
2     5.595833    2.037500
3     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 3 3 3 1
 [52] 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 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 2 2
[103] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2

Within cluster sum of squares by cluster:
[1] 13.05769 16.29167  2.02200
 (between_SS / total_SS =  94.3 %)

Available components:

[1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss" "betweenss"
[7] "size"         "iter"         "ifault"
> table(irisCluster$cluster, iris$Species)

    setosa versicolor virginica
  1      0         48         4
  2      0          2        46
  3     50          0         0
> irisCluster$iter
[1] 1
> irisCluster$centers
  Petal.Length Petal.Width
1     4.269231    1.342308
2     5.595833    2.037500
3     1.462000    0.246000
> irisCluster$size
[1] 52 48 50
> irisCluster$ifault
[1] 0
> irisCluster$cluster
  [1] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
 [52] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 2 2
[103] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
> irisCluster$cluster <- as.factor(irisCluster$cluster)
> ggplot(iris, aes(Petal.Length, Petal.Width, color = irisCluster$cluster))+ geom_point()
```

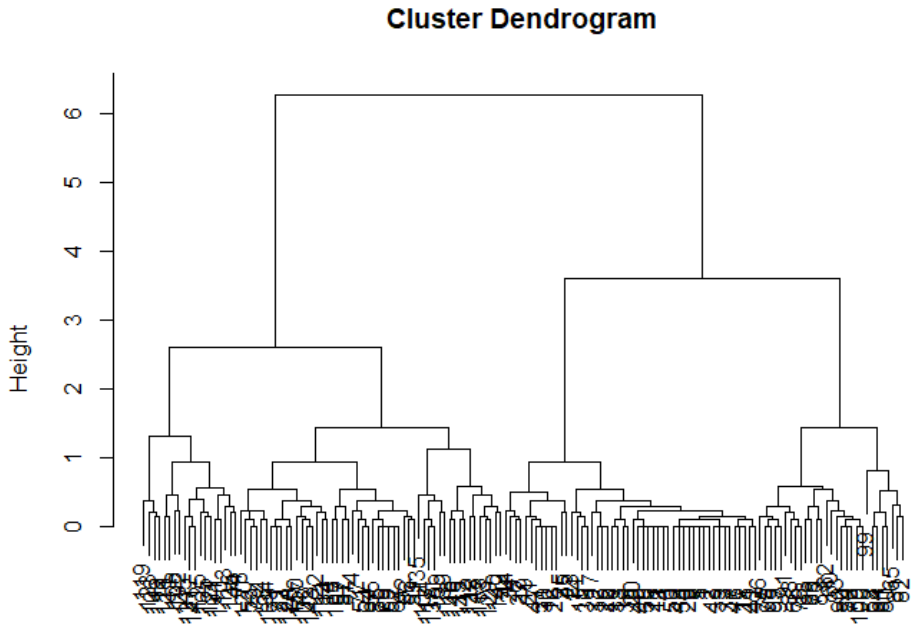

Petal.Width
Petal.Length
0.0
0.5
1.0
1.5
2.0
2.5
2
4
6
Species
setosa
versicolor
virginica

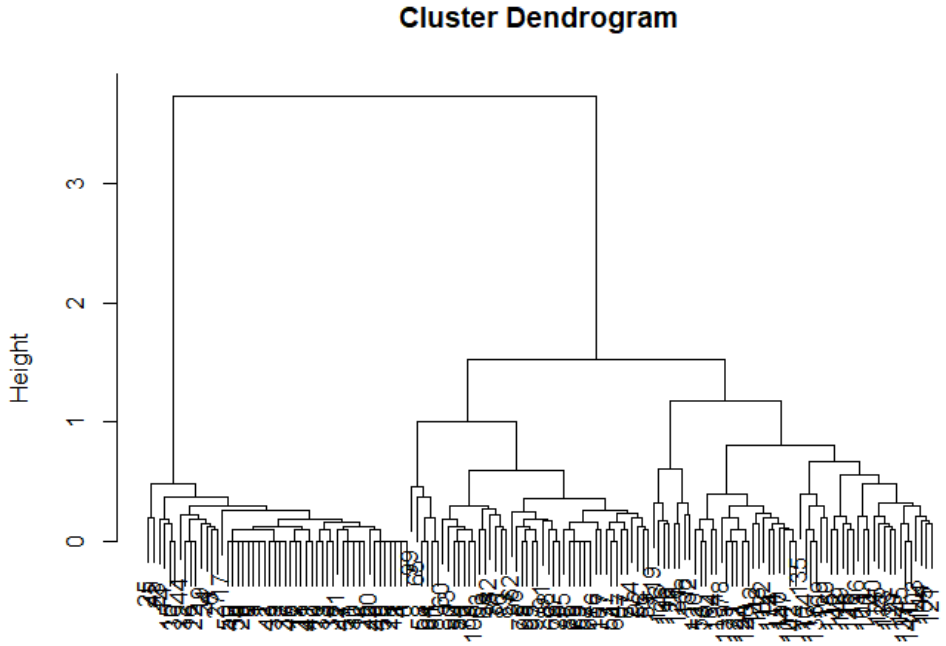

Petal.Width
Petal.Length
0.0
0.5
1.0
1.5
2.0
2.5
2
4
6
irisCluster$cluster
1
2
3

2. Unsupervised classification: Implementation and analysis of clustering algorithm Hierarchical clustering Agglomerative.

Code & Output:

```
> Iris<- iris
> head(iris)
  Sepal.Length Sepal.Width Petal.Length
1          5.1         3.5          1.4
2          4.9         3.0          1.4
3          4.7         3.2          1.3
4          4.6         3.1          1.5
5          5.0         3.6          1.4
6          5.4         3.9          1.7
  Petal.Width Species
1         0.2  setosa
2         0.2  setosa
3         0.2  setosa
4         0.2  setosa
5         0.2  setosa
6         0.4  setosa
> clusters <- hclust(dist(iris[, 3:4]), method="complete")
> clusters

Call:
hclust(d = dist(iris[, 3:4]), method = "complete")

Cluster method   : complete
Distance         : euclidean
Number of objects: 150

> plot(clusters)
> clusterCut <- cutree(clusters, 3)
> clusterCut
  [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [20] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [39] 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 3 2 2 2
 [58] 3 2 3 3 3 3 2 3 3 2 3 2 3 2 3 2 2 3 3
 [77] 2 2 2 3 3 3 3 2 2 2 2 3 3 3 3 2 3 3 3
 [96] 3 3 3 3 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[115] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[134] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
> 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
```


Cluster Dendrogram
Height
0
1
2
3
4
5
6
dist(iris[, 3:4])
hclust (*, "complete")


Cluster Dendrogram
Height
0
1
2
3
dist(iris[, 3:4])
hclust (*, "average")