**Practical No: 10**

**Aim:**- Implementation and analysis of clustering algorithms like K-Means , Agglomerative.

**Theory:**

K-Means Clustering is an [Unsupervised Learning algorithm](https://www.javatpoint.com/unsupervised-machine-learning), which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

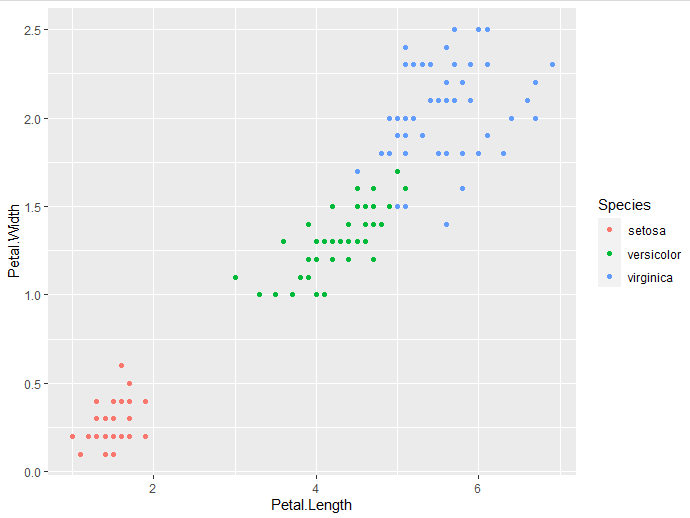
**R Commands:-**

head(iris)

install.packages("ggplot2",dependencies = TRUE)

library(ggplot2)

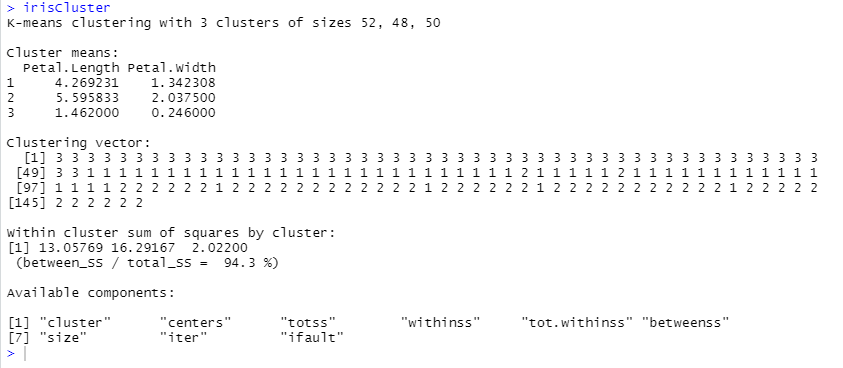
ggplot(iris, aes(Petal.Length, Petal.Width, color = Species)) + geom\_point()

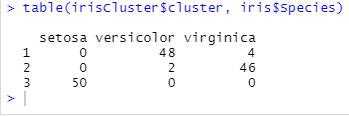


set.seed(20)

irisCluster <- kmeans(iris[, 3:4], 3, nstart = 20)

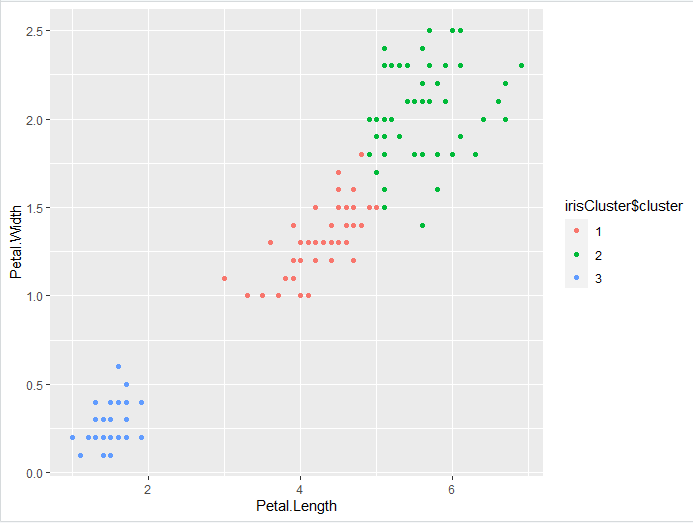
irisCluster





irisCluster$cluster <- as.factor(irisCluster$cluster)

ggplot(iris, aes(Petal.Length, Petal.Width, color =irisCluster$cluster)) + geom\_point()



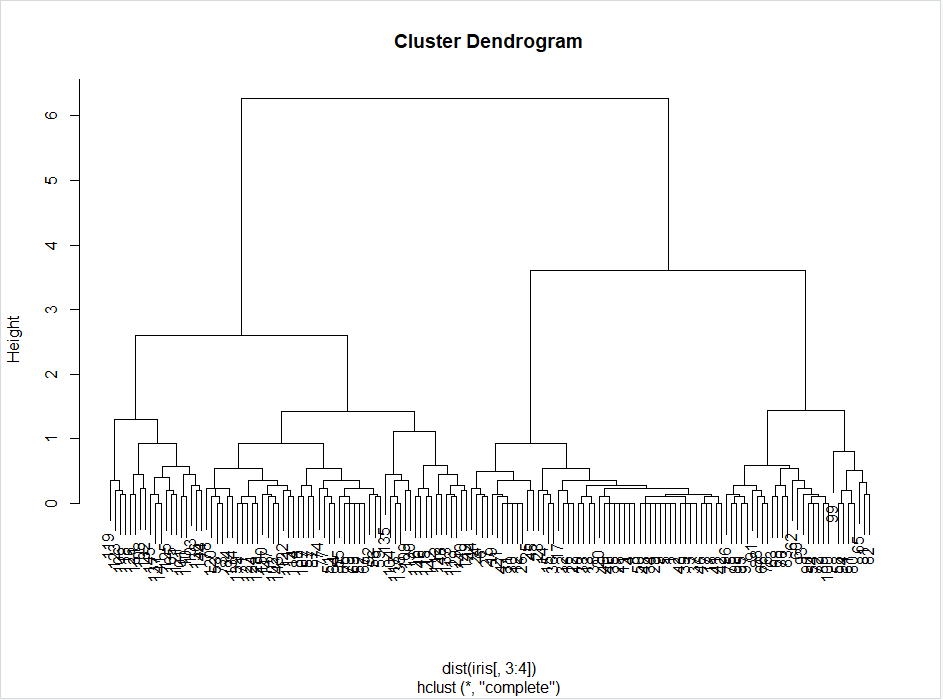
#Agglomerative Clustering

**Agglomerative Clustering is**Also known as bottom-up approach or hierarchical agglomerative clustering . A structure that is more informative than the unstructured set of clusters returned by flat clustering. This clustering algorithm does not require us to prespecify the number of clusters

head(iris)

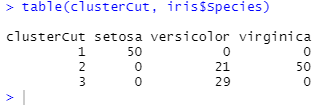
clusters <- hclust(dist(iris[, 3:4]))

plot(clusters)\



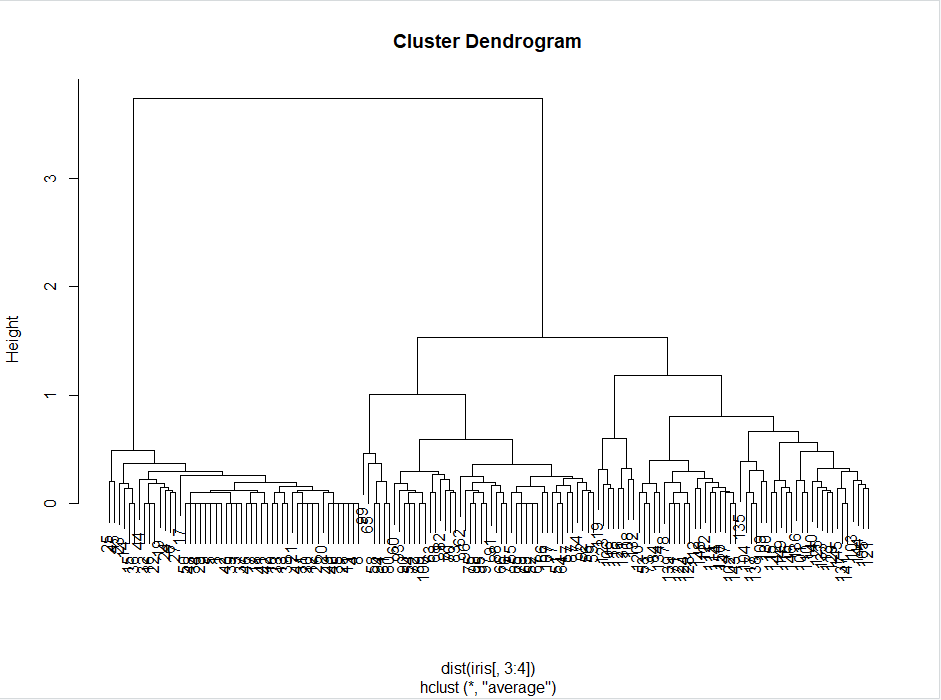
clusterCut<- cutree(clusters, 3)

table(clusterCut, iris$Species)



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

plot(clusters)



clusterCut <-cutree(clusters, 3)

table(clusterCut, iris$Species)

ggplot(iris, aes(Petal.Length, Petal.Width, color = iris$Species)) +geom\_point(alpha = 0.4, size = 3.5) + geom\_point(col = clusterCut) + scale\_color\_manual(values = c('black','red', 'green'))

