

**Implement clustering techniques – Hierarchical and K-Means****Aim:**

To implement clustering techniques – Hierarchical and K-Means using RStudio.

**Procedure:****a) HIERARCHICAL CLUSTERING**

# Load the iris dataset

data(iris)

# Use only the numeric columns for clustering (exclude the Species column)

iris\_data <- iris[, -5]

# Standardize the data

iris\_scaled <- scale(iris\_data)

# Compute the distance matrix

distance\_matrix <- dist(iris\_scaled, method = "euclidean")

# Perform hierarchical clustering using the "complete" linkage method

hc\_complete <- hclust(distance\_matrix, method = "complete")

# Plot the dendrogram

plot(hc\_complete, main = "Hierarchical Clustering Dendrogram", xlab = "", sub = "", cex = 0.6)

# Cut the tree to form 3 clusters

clusters <- cutree(hc\_complete, k = 3)

# Print the cluster memberships

print(clusters)

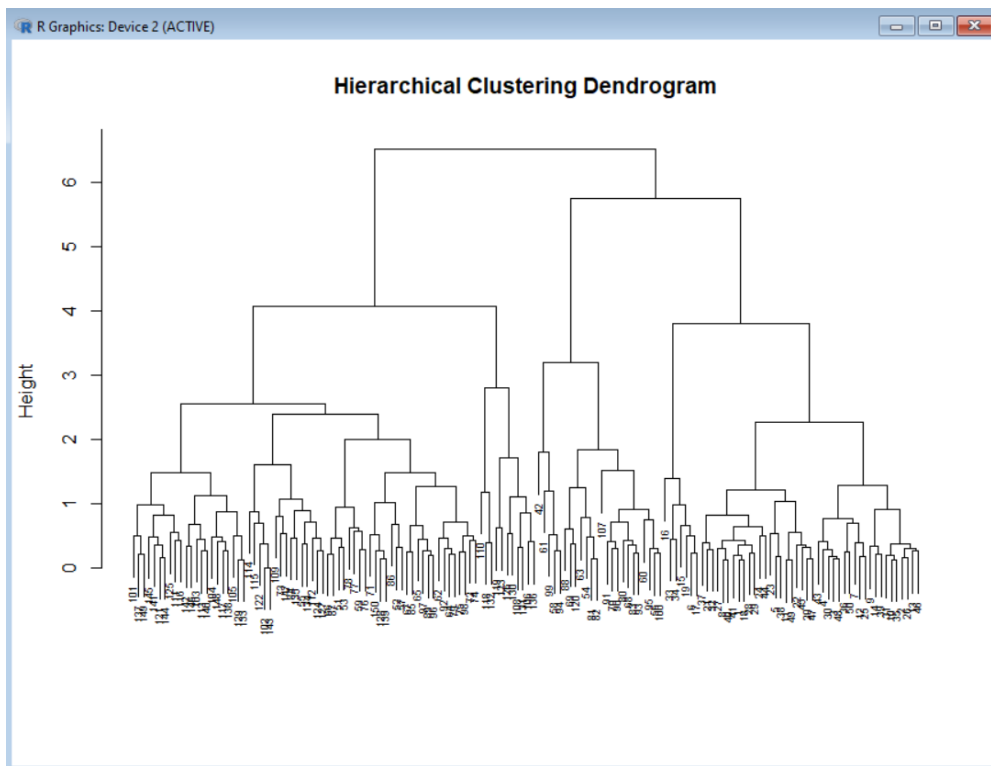
# Add the clusters to the original dataset

iris\$Cluster <- as.factor(clusters)

# Display the first few rows of the updated dataset

head(iris)

**Output:**

[illegible]

### **b) K-MEANS CLUSTERING**

### # Load the iris dataset

data(iris)

# Use only the numeric columns for clustering (exclude the Species column)

```
iris_data <- iris[, -5]  
# Standardize the data  
iris_scaled <- scale(iris_data)  
# Set the number of clusters  
set.seed(123) # For reproducibility  
k <- 3 # Number of clusters  
# Perform K-Means clustering  
kmeans_result <- kmeans(iris_scaled, centers = k, nstart = 25)  
# Print the K-Means result  
print(kmeans_result)  
# Print the cluster centers  
print(kmeans_result$centers)  
# Add the cluster assignments to the original dataset  
iris$Cluster <- as.factor(kmeans_result$cluster)  
# Display the first few rows of the updated dataset  
head(iris)  
# Plot the clusters  
library(ggplot2)  
ggplot(iris, aes(x = Sepal.Length, y = Sepal.Width, color = Cluster)) +  
geom_point(size = 3) +  
labs(title = "K-Means Clustering of Iris Dataset", x = "Sepal Length", y = "Sepal Width")
```

## Output:



## Result:

Thus, clustering techniques – Hierarchical and K-Means using R and RStudio was implemented successfully.