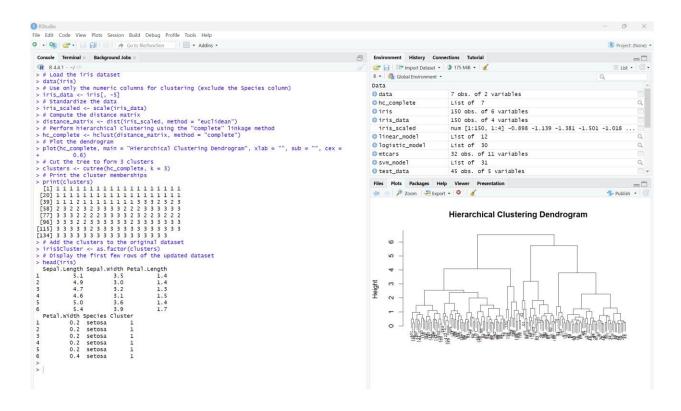
## Exp:9

## Implement clustering techniques – Hierarchical and K-Means

## a) HIERARCHIAL 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)</pre>
# Compute the distance matrix
distance_matrix <- dist(iris_scaled, method = "euclidean")</pre>
# Perform hierarchical clustering using the "complete" linkage method
hc complete <- hclust(distance matrix, method = "complete")</pre>
# 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)</pre>
# Display the first few rows of the updated dataset
head(iris)
```



## **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)</pre>

# 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)</pre>

# 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")

