EXP NO: 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)
# 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)</pre>
# Display the first few rows of the updated dataset head(iris)
head(iris)
```

OUTPUT:

```
    Histogram.R ×    Bar Chart.R ×    Box Plot.R ×    Decision Tree.R ×    Decision Tree.R ×    SYM.R ×
        Source on Save Q / - | #Load the iris dataset
                                                                                                                                                              Run Source - =
        data(iris)
       #Use only the numeric columns for clustering
iris_data <- iris[,-5]</pre>
       #Standardize the data
iris_scaled <- scale(iris_data)</pre>
   10 #Compute the distance matrix
   11 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 dendogram
plot(hc_complete, main = "Hierarchical Clustering dendogram", xlab="", sub="", cex=0.6)
       #Cut the tree to form 3 clusters| clusters <-cutree(hc_complete, k=3)
        #Print the cluster memberships
       #Add the clusters to the original dataset
   26 iris$Cluster <- as.factor(clusters)
   28 #Dusplay the first few rows of the updated dataset
29 head(iris)
Console Terminal × Background Jobs ×
R 4.4.1 · C:/Personal/Kavin/DA/Exp9/
> #Load the iris dataset
> data(iris)
> #Use only the numeric columns for clustering
> iris_data <- iris[,-5]</pre>
 > #Standardize the data
> iris_scaled <- scale(iris_data)
  #Compute the distance matrix
 distance_matrix <- dist(iris_scaled, method = "euclidean")</pre>
>
> #Perform hierarchical clustering using the "complete" linkage method
> hc_complete <- hclust(distance_matrix, method="complete")
 /
>#Plot the dendogram
> plot(hc_complete, main = "Hierarchical Clustering dendogram", xlab="", sub="", cex=0.6)
> #Cut the tree to form 3 clusters
> clusters <-cutree(hc_complete, k=3)
 > #Print the cluster memberships
> #Add the clusters to the original dataset
> iris$Cluster <- as.factor(clusters)</pre>
> #Dusplay the first few rows of the updated dataset
> #Dusplay the first few rows of the updated dataset
> head(iris)
   Sepal.Length Sepal.Width Petal.Length Petal.Width Species Cluster
             ength Sepai. Width Petai. Length Petai. W

5.1 3.5 1.4

4.9 3.0 1.4

4.7 3.2 1.3

4.6 3.1 1.5

5.0 3.6 1.4

5.4 3.9 1.7
                                                             0.2 setosa
0.2 setosa
0.2 setosa
0.2 setosa
0.2 setosa
0.2 setosa
0.4 setosa
```

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$center)
# Add the cluster assignments to the original dataset
iris$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")
```

OUTPUT:

