Ex.No -9 Roll No - 210701161

Implement clustering techniques – Hierarchical and K-Means

AIM:

To Implement clustering techniques – Hierarchical and K-Means using R.

PROCEDURE:

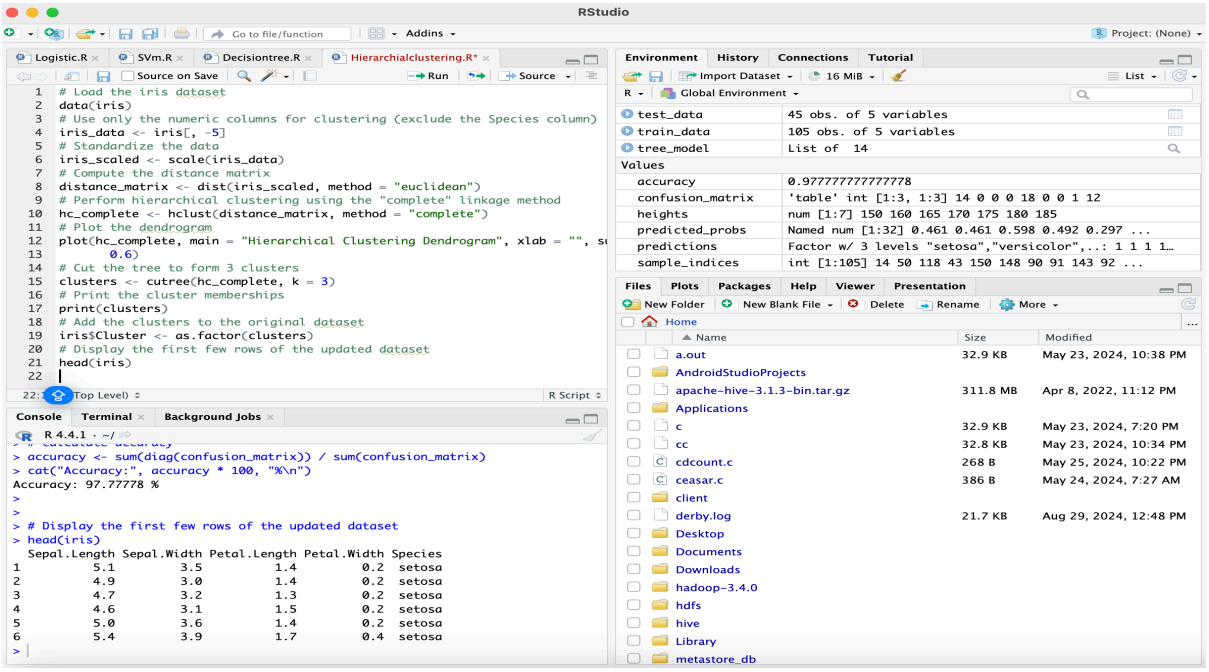
1. Collect and load the dataset from sources like CSV files or databases.
2. Clean and preprocess the data, including handling missing values and scaling features.
3. Determine the number of clusters (K) for K-Means, or decide on the stopping criterion for Hierarchical Clustering.
4. Choose the appropriate clustering algorithm: K-Means for partitioning, Hierarchical for nested clustering.
5. Apply the K-Means algorithm using fit_predict to assign data points to clusters.
6. Apply the Hierarchical Clustering algorithm using AgglomerativeClustering for hierarchical clusters.
7. Visualize the clusters with scatter plots for K-Means, and dendrograms for Hierarchical Clustering.
8. Evaluate clustering performance using metrics like silhouette score or inertia (for K-Means).
9. Fine-tune the clustering by adjusting the number of clusters or linkage criteria.
10. Interpret the results to understand the structure and relationships within the data.

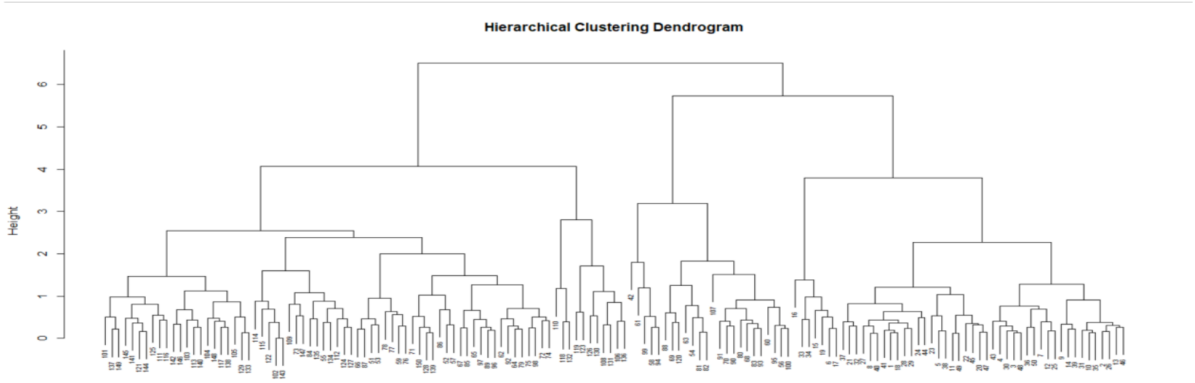
CODE:

```
# 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 datasethead(iris)
```

OUTPUT:





RESULT:

Thus, to Implement clustering techniques – Hierarchical and K-Means using R has successfully executed.