Agglomorative clustering on iris dataset(code)

# Load the data

data("USArrests")

# Standardize the data

df <- scale(USArrests)

# Show the first 6 rows

head(df, nrow = 6)

# Compute the dissimilarity matrix

# df = the standardized data

res.dist <- dist(df, method = "euclidean")

#To see easily the distance information between objects, we reformat the

#results of the function dist() into a matrix using the as.matrix() function.

#In this matrix, value in the cell formed by the row i, the column j, represents

#the distance between object i and object j in the original data set. For instance,

#element 1,1 represents the distance between object 1 and itself (which is zero).

#Element 1,2 represents the distance between object 1 and object 2, and so on.

#Displays the first 6 rows and columns of the distance matrix:

as.matrix(res.dist)[1:6, 1:6]

#creates hierarchial tree

res.hc <- hclust(d = res.dist, method = "ward.D2")

#for creating dendrogram

# cex: label size

library("factoextra")

fviz\_dend(res.hc, cex = 0.5)

# Please refer notes section

# Compute cophentic distance

#The cophenetic distance between two objects is the height of the

#dendrogram where the two branches that include the two objects merge

#into a single branch.

res.coph <- cophenetic(res.hc)

# Correlation between cophenetic distance and

# the original distance

cor(res.dist, res.coph)

# Cut tree into 4 groups

grp <- cutree(res.hc, k = 4)

head(grp, n = 4)

# Number of members in each cluster

table(grp)

# Get the names for the members of cluster 1

rownames(df)[grp == 1]

fviz\_dend(grp, n=4)

#-------------------------

library("cluster")

# Agglomerative Nesting (Hierarchical Clustering)

res.agnes <- agnes(x = USArrests, # data matrix

stand = TRUE, # Standardize the data

metric = "euclidean", # metric for distance matrix

method = "ward" # Linkage method

)

# DIvisive ANAlysis Clustering

res.diana <- diana(x = USArrests, # data matrix

stand = TRUE, # standardize the data

metric = "euclidean" # metric for distance matrix

)

fviz\_dend(res.agnes, cex = 0.6, k = 4)

fviz\_dend(res.diana, cex = 0.6, k = 4)