

Agglomerative Hierarchical Clustering

```
rm(list = ls()) # Free up memory
```

```
gc()          # Garbage Collection
```

```
#install.packages("cluster")
```

```
library(cluster)
```

```
data1 <- read.csv("seeds_dataset1.csv", header = TRUE)
```

```
#dd <- read.csv("seeds_dataset1.csv", header = TRUE)
```

```
View(data1)
```

```
plot(data1)
```

```
# To calculate the distance provide 'method' parameter to the dist function.
```

```
# Available distance measures are Euclidean, Manhattan, Maximum, Minkowski, etc
```

```
# To use cosine similarity measure, use cosine() function in 'lsa' package.
```

```
# dist() function computes and returns the distance matrix computed by using the
```

```
#specified distance measure to compute the distances between the rows of a data matrix.
```

```
#Default method is "euclidean" but options include "maximum", "manhattan",
```

```
#"canberra", "binary" or "minkowski".
```

```
distMat <- dist(data1,method = "euclidean")
```

```
distMat
```

```
# Use 'method' option of hclust() to provide merging criteria like, single,
```

```
# complete, average, etc.
```

```
Clust1 <- hclust(distMat,method="single")
```

Clust1

```
plot(Clust1) #plot the clusters
```

```
dend <- as.dendrogram(Clust1)# Convert hclust into a dendrogram and plot
```

```
# Plot the dendrogram
```

```
plot(dend)
```

```
*****
```