Experiment 8

Student Name: Kanishk Soni
UID: 20BCS9398
Branch: BE-CSE
Section/Group: 20BCS-DM_708B
Semester: 6th
Subject Name: Data Mining Lab
Subject Code: 20CSP-376

1. Aim:

To perform the hierarchical clustering using R programming.

2. Objective:

- hclust in the stats package and agnes in the cluster package for agglomerative hierarchical clustering.
- diana in the cluster package for divisive hierarchical clustering.

3. Code and Output:

PROGRAM

```
library(datasets)

library(cluster)

library(factoextra)

library(purrr)


df = iris[, 1:4]

df = na.omit(df)

df = scale(df)

d = dist(df, method = "euclidean")
```

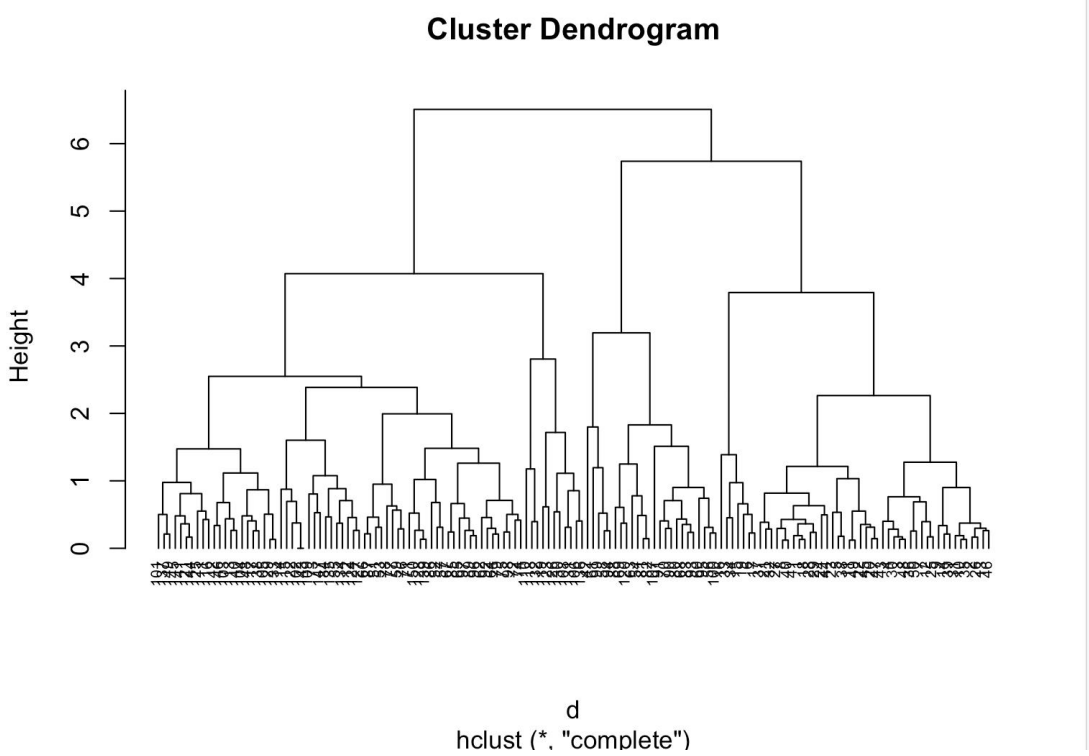
```
hc1 = hclust(d, method = "complete")

plot(hc1, cex = 0.6, hang = –1)


sub_groups = cutree(hc1, k = 3)

fviz_cluster(list(data = df, cluster = sub_groups))


plot(hc1, cex = 0.6, hang = –1)

rect.hclust(hc1, k = 3, border = 2:4)


hc3 = diana(df)

hc3$dc

pltree(hc3, cex = 0.6, hang = –1, main = "Diana's Dendogram")
```

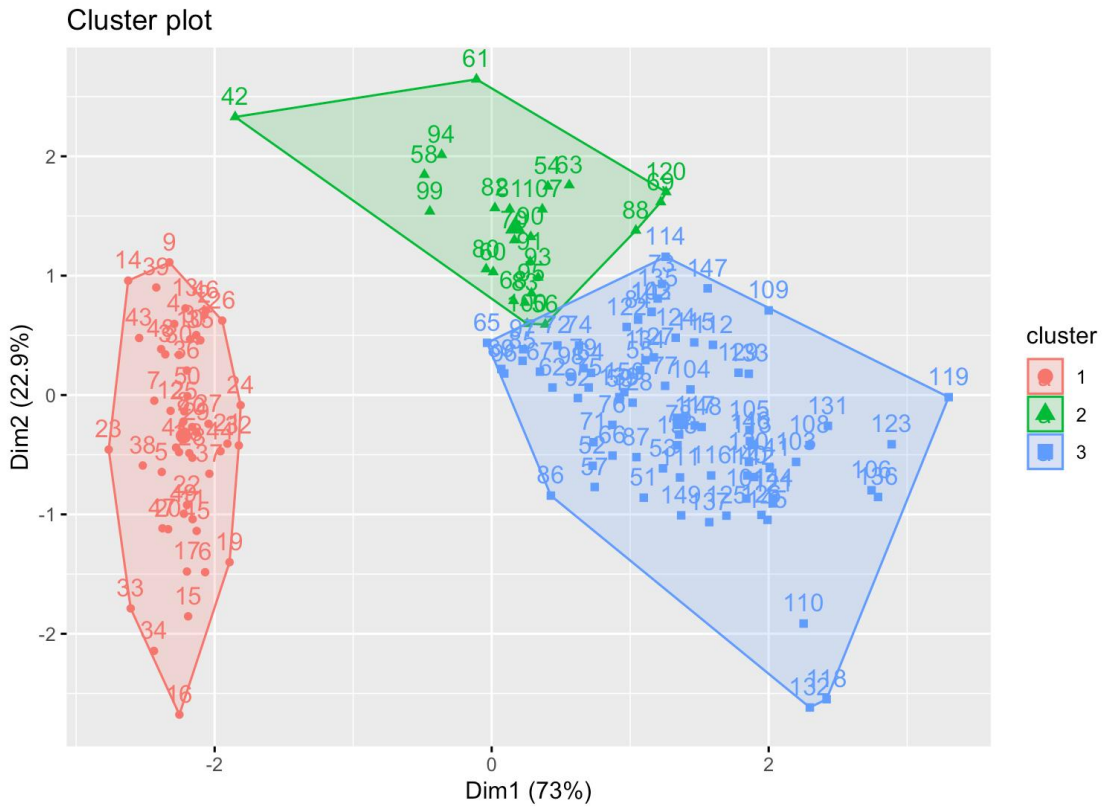
CONSOLE

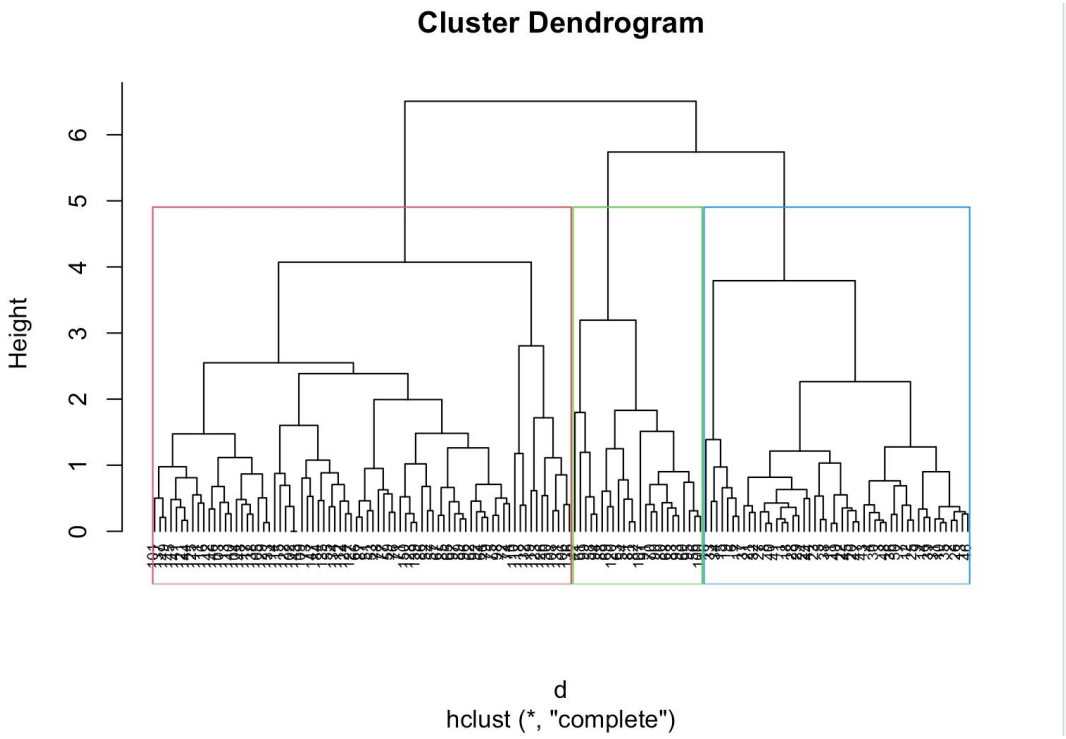
```
> library(datasets)
> library(cluster)
> library(factoextra)
Loading required package: ggplot2
Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3W
Ba
> library(purrr)
> df = iris[, 1:4]
> df = na.omit(df)
> df = scale(df)
> d = dist(df, method="euclidean")
> hc1 = hclust(d, method = "complete")
```

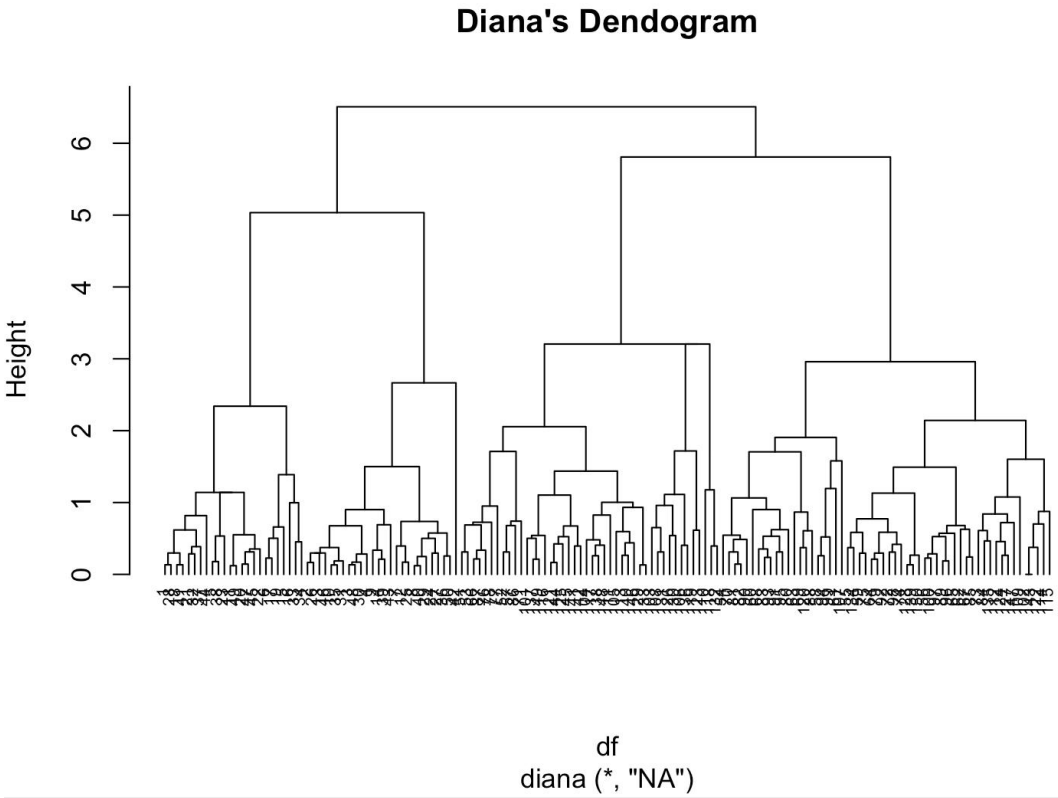
```
> plot(hc1, cex = 0.6, hang = –1)
> sub_groups = cutree(hc1, k = 3)
> fviz_cluster(list(data = df, cluster = sub_groups))
> plot(hc1, cex = 0.6, hang = –1)
> rect.hclust(hc1, k = 3, border = 2:4)
> hc3 = diana(df)
> hc3$dc
[1] 0.9397208
> pltree(hc3, cex = 0.6, hang = –1, main = "Diana's Dendogram")
>
```

4. Output:




Cluster plot
Dim2 (22.9%)
Dim1 (73%)
cluster
1
2
3


Cluster Dendrogram
Height
d
hclust (*, "complete")



Learning Outcomes:

- Learnt how to do Hierarchical cluster analysis.
- In which we learnt about Agglomerative hierarchical clustering and Divisive hierarchical clustering.
- We also learnt how to make Dendogram for the analysis.