



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

Experiment-3.1

Student Name: Nabha Varshney

UID: 20BCS4995

Branch: CSE

Section/Group: 20BCS-DM-704 (A)

Semester: 6th

Date of Performance: 28th Apr 2023

Subject Name: Data Mining

Subject Code: 20CSP- 351

Aim – To perform the hierarchical clustering using R programming.

Objective-

- ♦ Represent the reading of file using R studio
- ♦ Displaying the graph using cluster, purrr and datasets.
- ♦ Demonstration of Cluster analysis by hierarchical clustering.

Script and Output-

```
# Load required packages

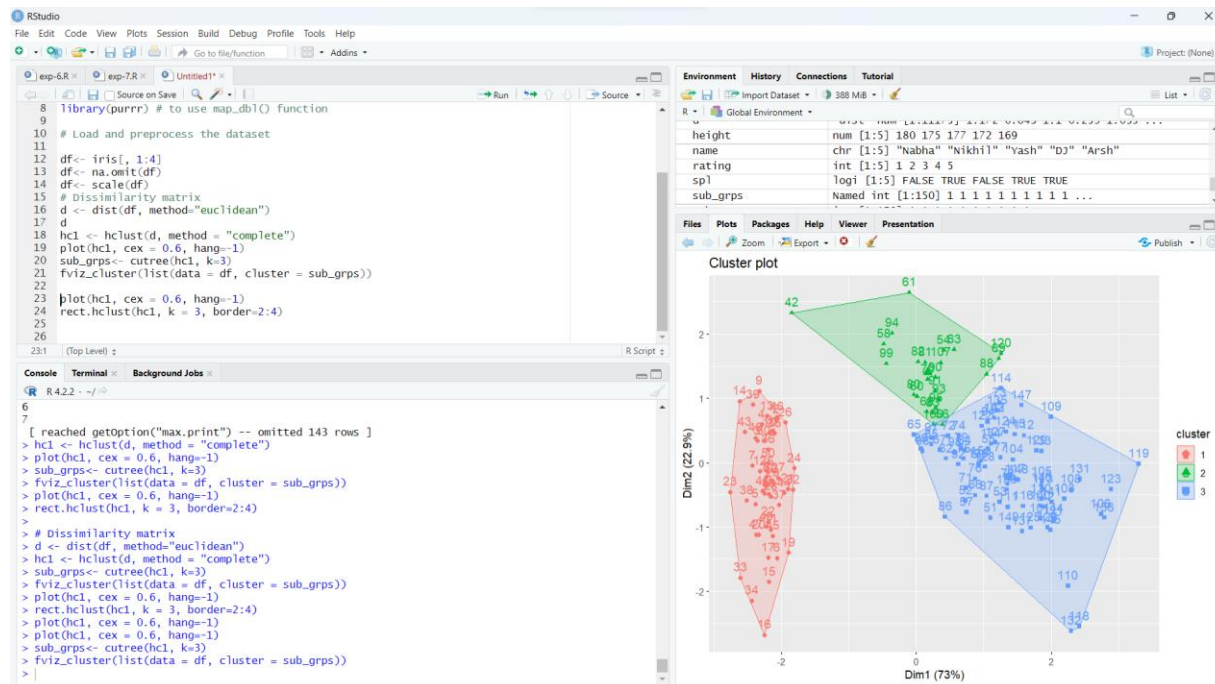
library(datasets) # contains iris dataset
library(cluster) # clustering algorithms
install.packages("factoextra")
library(factoextra) # visualization
install.packages("purrr")
library(purrr) # to use map_dbl() function

# Load and preprocess the dataset

df<- iris[, 1:4]
df<- na.omit(df)
df<- scale(df)
# Dissimilarity matrix
d <- dist(df, method="euclidean")
d
hc1 <- hclust(d, method = "complete")
plot(hc1, cex = 0.6, hang=-1)
sub_grps<- cutree(hc1, k=3)
fviz_cluster(list(data = df, cluster = sub_grps))

plot(hc1, cex = 0.6, hang=-1)
rect.hclust(hc1, k = 3, border=2:4)
```

Output-





DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

