Experiment 3.1

Student Name: Aayush Gurung
UID: 20BCS5323
Branch: CSE
Section/Group: 20BCS_DM_607/A
Semester: 6
Subject Name: Data Mining Lab
Subject Code: 20CSP- 376

Aim: To perform hierarchical clustering using R programming.

Objective: To Learn about hierarchical Clustering.

Theory: Hierarchical cluster analysis (also known as hierarchical clustering) is a clustering technique where clusters have a hierarchy or a predetermined order. Hierarchical clustering can be represented by a tree- like structure called a Dendrogram.

There are two types of hierarchical clustering:

- Agglomerative hierarchical clustering: This is a bottom-up approach where each data point starts in its own cluster and as one moves up the hierarchy, similar pairs of clusters are merged.
- Divisive hierarchical clustering: This is a top-down approach where all data points start in one cluster and as one moves down the hierarchy, clusters are split recursively.

Script/Code/Steps:

```
install.packages("dplyr")
library(dplyr)
distance_mat <- dist(mtcars, method = 'euclidean')
distance_mat
set.seed(240)
Hierar_cl <- hclust(distance_mat, method = "average")
Hierar_cl
plot(Hierar_cl)
abline(h = 10, col = "blue")
fit <- cutree(Hierar_cl, k = 3 )
fit
table(fit)
      rect.hclust(Hierar_cl, k = 3, border =
     "blue")
```

```r
install.packages("dplyr")
library(dplyr)
distance_mat <- dist(mtcars, method = 'euclidean')
distance_mat
set.seed(240)
Hierar_cl <- hclust(distance_mat, method = "average")
Hierar_cl
plot(Hierar_cl)
abline(h = 10, col = "green")
fit <- cutree(Hierar_cl, k = 3 )
fit
table(fit)
rect.hclust(Hierar_cl, k = 3, border = "green")
```

Output Screenshots:

```
> library(dplyr)
> distance_mat <- dist(mtcars, method = 'euclidean')
> distance_mat
                     Mazda RX4 Mazda RX4 Wag  Datsun 710 Hornet 4 Drive Hornet Sportabout
Mazda RX4 Wag         0.6153251
Datsun 710           54.9086059    54.8915169
Hornet 4 Drive       98.1125212    98.0958939 150.9935191
Hornet Sportabout   210.3374396   210.3358546 265.0831615    121.0297564
Valiant              65.4717710    65.4392224 117.7547018     33.5508692       152.1241352
Duster 360          241.4076490   241.4088680 294.4790230    169.4299647        70.1767262
Merc 240D            50.1532711    50.1146059  49.6584796    121.2739722       241.5069657
Merc 230             25.4683117    25.3284509  33.1803843    118.2433145       233.4924012
Merc 280             15.3641921    15.2956865  66.9363534     91.4224033       199.3344960
Merc 280C            15.6724727    15.5837744  67.0261397     91.4612914       199.3406564
Merc 450SE          135.4307018   135.4254826 189.1954941     72.4964325        84.3888482
Merc 450SL          135.4014424   135.3960351 189.1631745     72.4313532        84.3683999
Merc 450SLC         135.4794674   135.4723157 189.2345426     72.5718466        84.4332423
Cadillac Fleetwood  326.3395903   326.3355070 381.0926242    234.4403876       116.2804201
Lincoln Continental 318.0469808   318.0429333 372.8012090    227.9726091       108.0624299
Chrysler Imperial   304.7203408   304.7169175 359.3014906    218.1548299        97.2049146
Fiat 128             93.2679950    93.2530993  40.9933763    184.9689734       302.0377212
Honda Civic         102.8307567   102.8238713  52.7704607    191.5518700       310.0324645
Toyota Corolla      100.6040368   100.5887588  47.6535017    192.6714187       309.5581776
Toyota Corona        42.3075233    42.2659224  12.9654743    138.5304725       252.3331988
Dodge Challenger    163.1150750   163.1134210 217.7795805     72.4403915        48.9838851
AMC Javelin         149.6047203   149.6014522 204.3188913     61.3601899        61.4274240
Camaro Z28          233.2228758   233.2248748 286.0049209    163.6632641        70.9665308
Pontiac Firebird    248.6780270   248.6762035 303.3583889    156.2240346        40.0052475
Fiat X1-9            92.5048389    92.4940020  39.8815148    184.4471198       301.5669483
Porsche 914-2        44.4033659    44.4073589  13.1357109    139.1579524       254.1452553
Lotus Europa         65.7328377    65.7362635  25.0948550    163.2367437       272.3582423
Ford Pantera L      245.4247064   245.4293785 297.2940489    180.1140339        89.5934049
Ferrari Dino         66.7661029    66.7764167  90.2415509    130.5523007       215.0673853
Maserati Bora       265.6454248   265.6491465 309.7718171    229.3419352       170.7094473
Volvo 142E           39.1894029    39.1626037  20.6939436    137.0363299       248.0063378
                        Valiant  Duster 360   Merc 240D    Merc 230    Merc 280   Merc 280C
Mazda RX4 Wag
Datsun 710
Hornet 4 Drive
Hornet Sportabout
Valiant
```

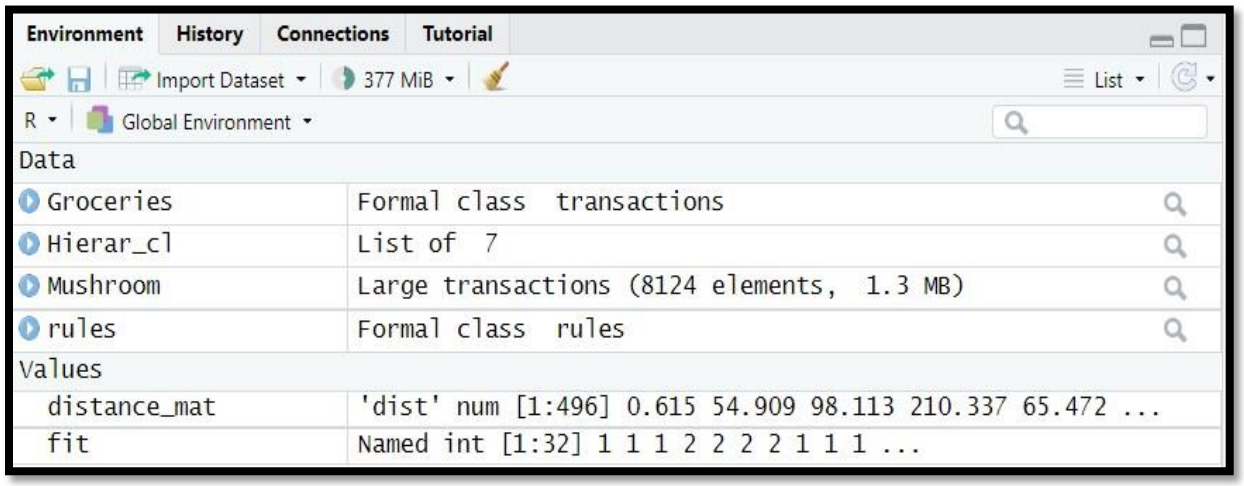
```
Console   Terminal   Background Jobs
R 4.2.2 · ~/
Camaro Z28                100.3046106 105.6062618
Pontiac Firebird           85.8075196  99.2836114  86.2665759
Fiat X1-9                 253.6624046 240.5266823 325.1490914       339.1396182
Porsche 914-2             206.6452569 193.3080584 276.8924414       292.1646488  48.3775209
Lotus Europa              226.5004836 212.7568765 287.6179004       311.3862342  49.8406880
Ford Pantera L            118.7516779 123.3832044  19.3589023       101.7389686 336.7018783
Ferrari Dino              174.9280395 161.1060307 216.7489910       255.0570519 127.8210813
Maserati Bora             185.9059273 185.1553411 102.5946154       188.3240020 349.1199576
Volvo 142E                201.3682522 187.6978440 266.5277736       286.7497823  60.4120429
                     Porsche 914-2 Lotus Europa Ford Pantera L Ferrari Dino Maserati Bora
Mazda RX4 Wag
Datsun 710
Hornet 4 Drive
Hornet Sportabout
Valiant
Duster 360
Merc 240D
Merc 230
Merc 280
Merc 280C
Merc 450SE
Merc 450SL
Merc 450SLC
Cadillac Fleetwood
Lincoln Continental
Chrysler Imperial
Fiat 128
Honda Civic
Toyota Corolla
Toyota Corona
Dodge Challenger
AMC Javelin
Camaro Z28
Pontiac Firebird
Fiat X1-9
Porsche 914-2
Lotus Europa            33.7678653
Ford Pantera L         288.5852993  297.5376920
Ferrari Dino            87.9105966   80.4553451    224.4587490
Maserati Bora          303.9222549  303.2796468     86.9383253  223.5342175
Volvo 142E              18.7555858   27.8104457    277.4803312   70.4751034   289.1157363
```

```
> set.seed(240)
> Hierar_cl <- hclust(distance_mat, method = "average")
> Hierar_cl

Call:
hclust(d = distance_mat, method = "average")

Cluster method   : average
Distance         : euclidean
Number of objects: 32

> plot(Hierar_cl)
> abline(h = 10, col = "green")
> fit <- cutree(Hierar_cl, k = 3 )
> fit
          Mazda RX4       Mazda RX4 Wag          Datsun 710      Hornet 4 Drive
                  1                   1                   1                   2
  Hornet Sportabout             Valiant          Duster 360           Merc 240D
                  2                   2                   2                   1
           Merc 230            Merc 280           Merc 280C          Merc 450SE
                  1                   1                   1                   2
        Merc 450SL         Merc 450SLC  Cadillac Fleetwood Lincoln Continental
                  2                   2                   2                   2
  Chrysler Imperial            Fiat 128         Honda Civic      Toyota Corolla
                  2                   1                   1                   1
     Toyota Corona    Dodge Challenger         AMC Javelin          Camaro Z28
                  1                   2                   2                   2
   Pontiac Firebird           Fiat X1-9       Porsche 914-2        Lotus Europa
                  2                   1                   1                   1
    Ford Pantera L        Ferrari Dino       Maserati Bora          Volvo 142E
                  2                   1                   3                   1
> table(fit)
fit
 1  2  3
16 15  1
> rect.hclust(Hierar_cl, k = 3, border = "green")
```



Output of Cluster Dendrogram: