

Name : Cheruvathoor Abin Anto  
Roll No : TIT2425008

## PRACTICAL 7

The file Iris.csv contains 50 samples from each of 3 species of Iris (Iris setosa, Iris virginica, Iris versicolor). Build DBScan clustering Model and plot it.

```
iris.df <- read.csv("Iris.csv")
```

```
#libraries
```

```
install.packages("fpc")
```

```
library(fpc) #For dbscan function
```

```
#removing labels from the data set
```

```
iris_1 <- iris[-5]
```

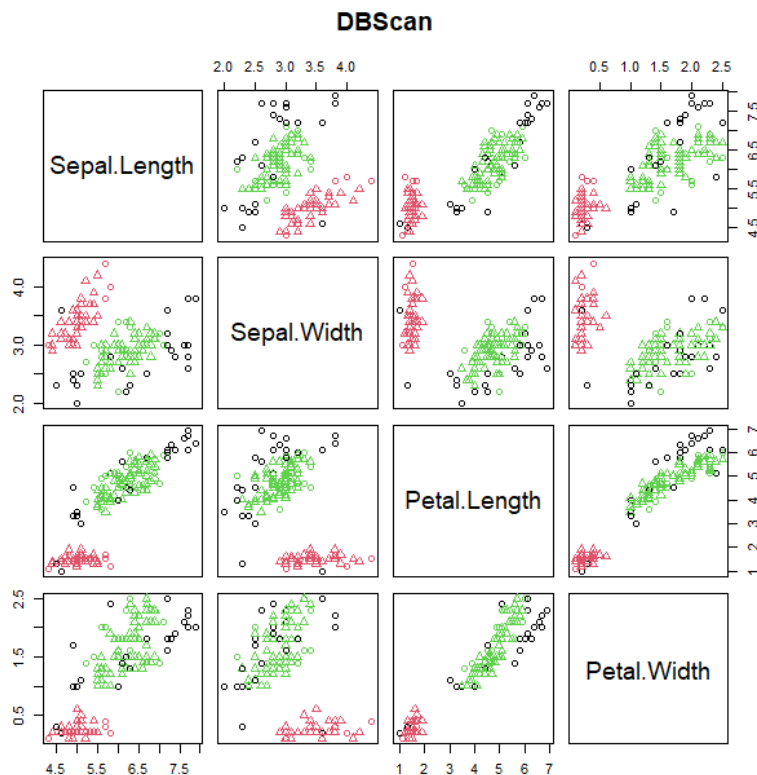
```
View(iris_1)
```

```
downloaded 824 kB  
package 'DEoptimR' successfully unpacked and MD5 sums checked  
package 'mclust' successfully unpacked and MD5 sums checked  
package 'flexmix' successfully unpacked and MD5 sums checked  
package 'prabclus' successfully unpacked and MD5 sums checked  
package 'diptest' successfully unpacked and MD5 sums checked  
package 'robustbase' successfully unpacked and MD5 sums checked  
package 'kernlab' successfully unpacked and MD5 sums checked  
package 'fpc' successfully unpacked and MD5 sums checked  
  
The downloaded binary packages are in  
  C:\Users\admin\AppData\Local\Temp\RtmpUR9dR1\downloaded_packages  
> library(fpc)
```

```
#Fitting DBScan Clustering Model to Training Dataset
```

```
Dbscan_cl <- dbscan(iris_1, eps = 0.45, MinPts = 5)
```

```
Dbscan_cl
```



## PRACTICAL 8 : Hierarchical Clustering

`mtcars`(motor trend car road test) comprises fuel consumption, performance and 10 aspects of automobile design for 32 automobiles. Perform hierarchical clustering and plot the dendrogram and cut the tree by the no. of clusters.

### DESCRIPTION OF THE DATASET

A data frame with 32 observations on 11 (numeric) variables.

1. mpg - Miles/(US) gallon
2. cyl – Number of cylinders
3. disp – Displacement (cubic inches)
4. hp – Gross horsepower
5. drat – rear axle ratio
6. wt – Weight (1000 lbs)
7. qsec – ¼ mile time
8. vs - Engine(0 = V-shaped, 1 = straight)
9. am – Transmission (0 = automatic, 1 = manual)
10. gear – Number of forward gears
11. carb – Number of carburetors

```
mtcars.df <- read.csv("mtcars.csv")
```

```
#libraries
```

```
install.packages("dplyr")
```

```
library(dplyr) #hclust function
```

```
library(ggplot2)
```

```
package 'dplyr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
  C:\Users\admin\AppData\Local\Temp\RtmpUR9dR1\downloaded_packages
> library(dplyr) #hclust function
Attaching package: 'dplyr'
```

```
#Finding distance matrix
```

```
distance_mat <- dist(mtcars, method = 'euclidean')
```

```
distance_mat
```

```

> distance_mat <- dist(mtcars, method = 'euclidean')
> distance_mat

```

	Mazda RX4	Mazda RX4 Wag	Datsun 710	Hornet 4 Drive
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
Duster 360	241.4076490	241.4088680	294.4790230	169.4299647
Merc 240D	50.1532711	50.1146059	49.6584796	121.2739722
Merc 230	25.4683117	25.3284509	33.1803843	118.2433145
Merc 280	15.3641921	15.2956865	66.9363534	91.4224033
Merc 280C	15.6724727	15.5837744	67.0261397	91.4612914
...	...	...	...	...
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				
Ford Pantera L				
Ferrari Dino	224.4587490			
Maserati Bora	86.9383253	223.5342175		
Volvo 142E	277.4803312	70.4751034	289.1157363	

#Finding Hierarchical clustering model to training dataset

```
Hierar_cl <- hclust(distance_mat, method = "average")
```

```
Hierar_cl
```

```

> Hierar_cl <- hclust(distance_mat, method = "average")
> Hierar_cl

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

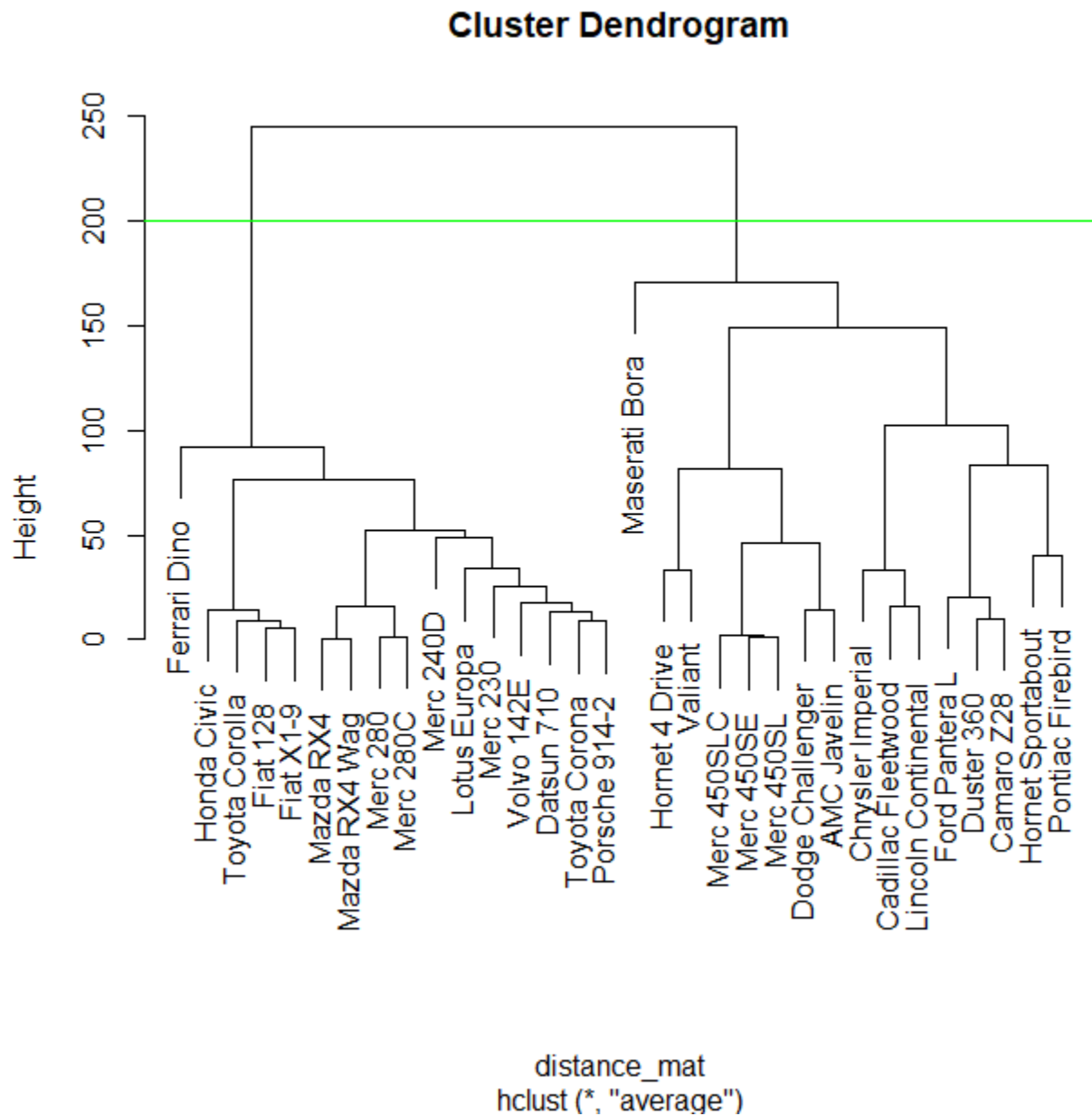
Cluster method   : average
Distance          : euclidean
Number of objects: 32

> |

```

```
#Plotting dendrogram
plot(Hierar_cl)
```

File History Resize



```
#choosing number of clusters
```

```
#cutting tree by height
```

```
abline(h = 200, col = "green")
```

```
fit=cutree(Hierar_cl,k=3)
```

fit

```
> 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)

rect.hclust(Hierar\_cl,k=3,border="red")

```
> #choosing number of clusters
> #cutting tree by height
> abline(h = 200, col = "green")
> fit=cutree(Hierar_cl,k=3)
> table(fit)
fit
 1  2  3
16 15  1
> rect.hclust(Hierar_cl,k=3,border="red")
> |
```

---

Cluster Dendrogram

