Clustering With DBScans

Nama : Mahesa Cadi Rajasa (19523122)

```
library(factoextra)
```

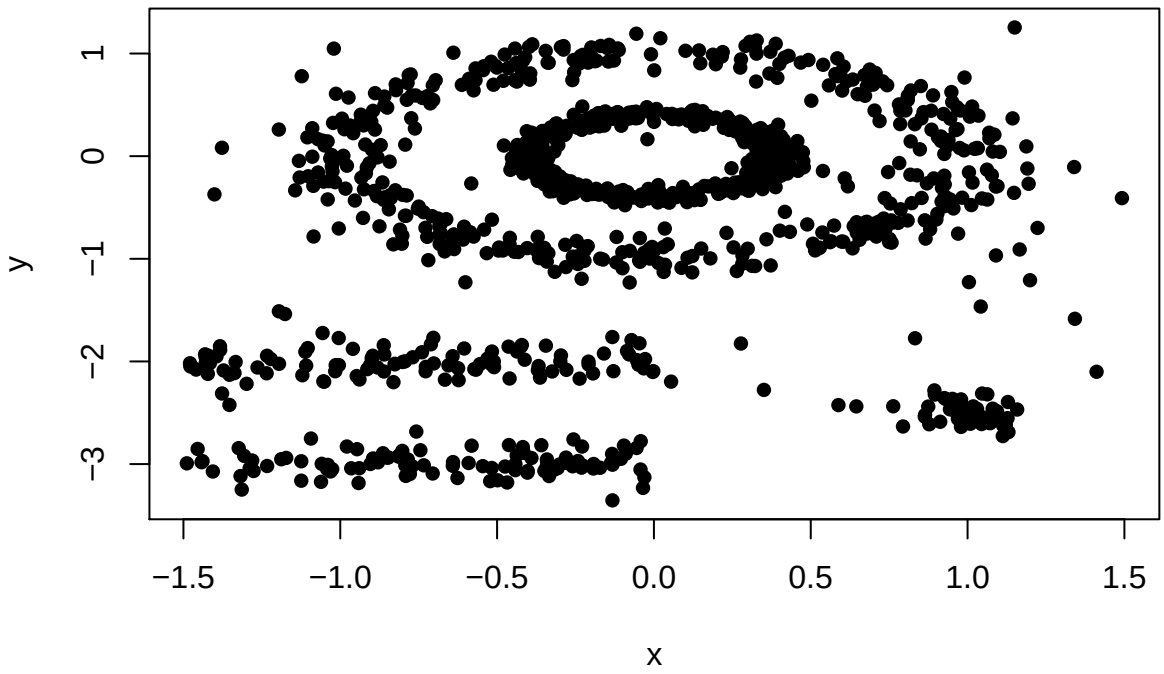
```
## Warning: package 'factoextra' was built under R version 4.0.3
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.0.3
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

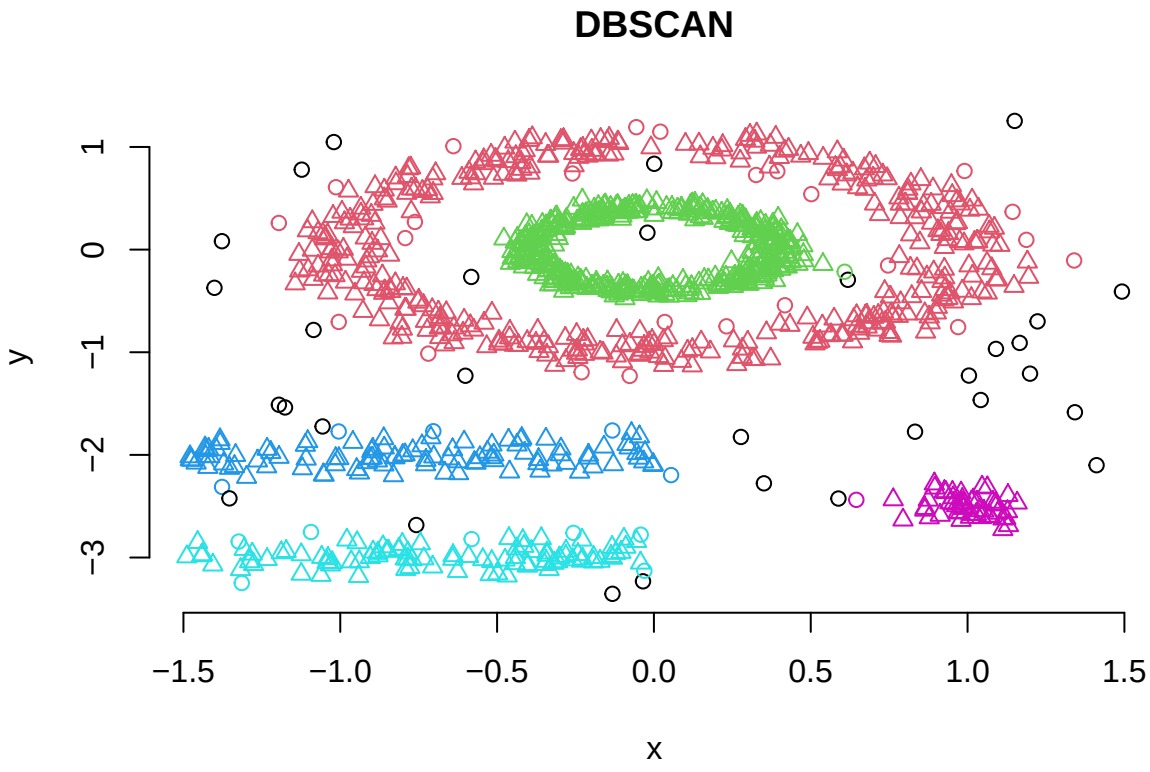
```
data("multishapes")
df <- multishapes[, 1:2]
plot(df, pch=16)
```



```
library(fpc)
```

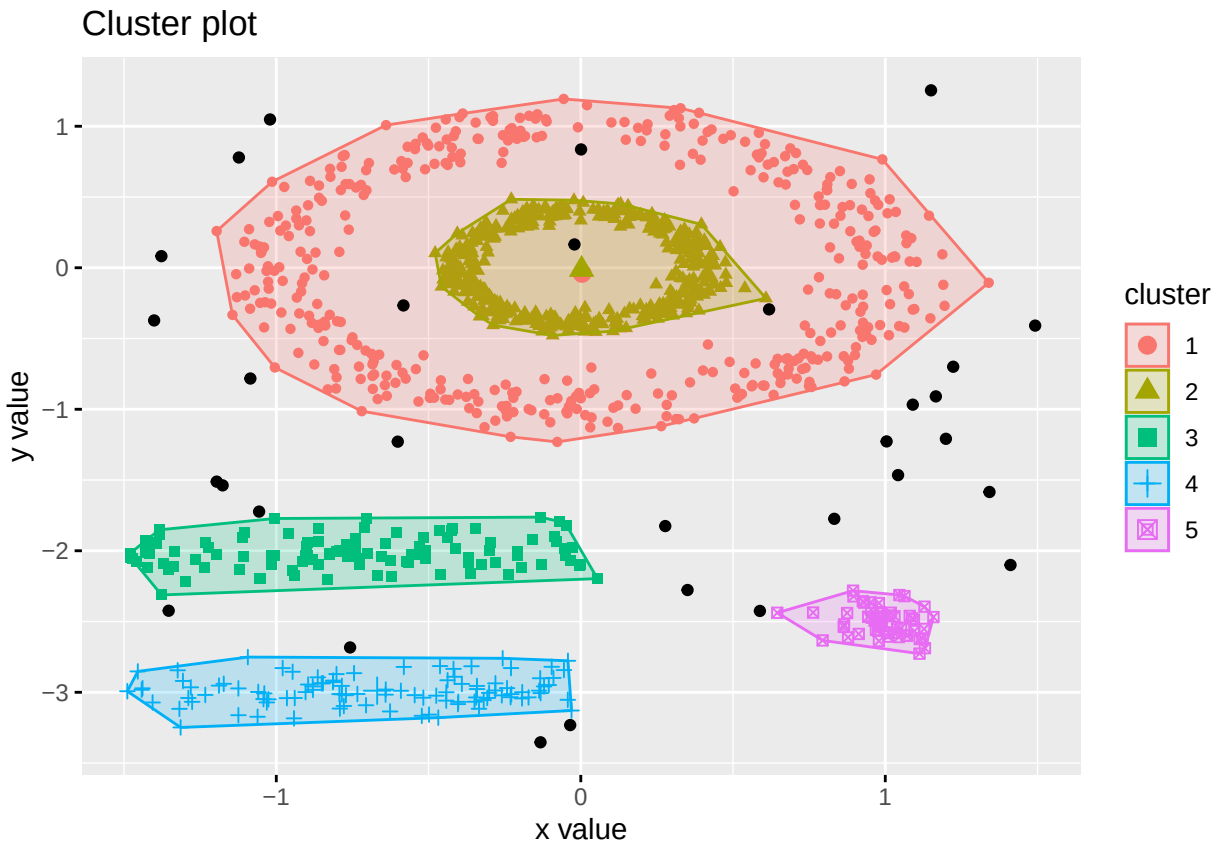
```
## Warning: package 'fpc' was built under R version 4.0.3
```

```
set.seed(123)
db <- dbscan(df, eps = 0.15, MinPts = 5)
plot(db, df, main = "DBSCAN", frame = FALSE)
```




```
library(factoextra)
fviz_cluster(db, df, stand = FALSE, frame = FALSE, geom = "point")
```

```
## Warning: argument frame is deprecated; please use ellipse instead.
```




```
print(db)
```

```
## dbscan Pts=1100 MinPts=5 eps=0.15
##         0   1   2   3  4  5
## border 31  24   1   5  7  1
## seed    0 386 404  99 92 50
## total  31 410 405 104 99 51
```

```
db$cluster
```

```
##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1
##  [75] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [112] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [149] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [186] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [223] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [260] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [297] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [334] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [371] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2
## [408] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [445] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [482] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [519] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
##  [556] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [593] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [630] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [667] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [704] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [741] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [778] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3
##  [815] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
##  [852] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
##  [889] 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 0 4 4 4 4 4 4 4 4 4 0
##  [926] 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
##  [963] 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
## [1000] 4 1 5 1 2 0 0 0 0 0 1 0 2 0 2 0 1 1 4 0 0 3 0 0 2 0 0 1 0 0 0 0 1 2 0 0 0
## [1037] 1 1 1 1 0 1 3 0 3 0 0 0 0 3 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
## [1074] 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
```

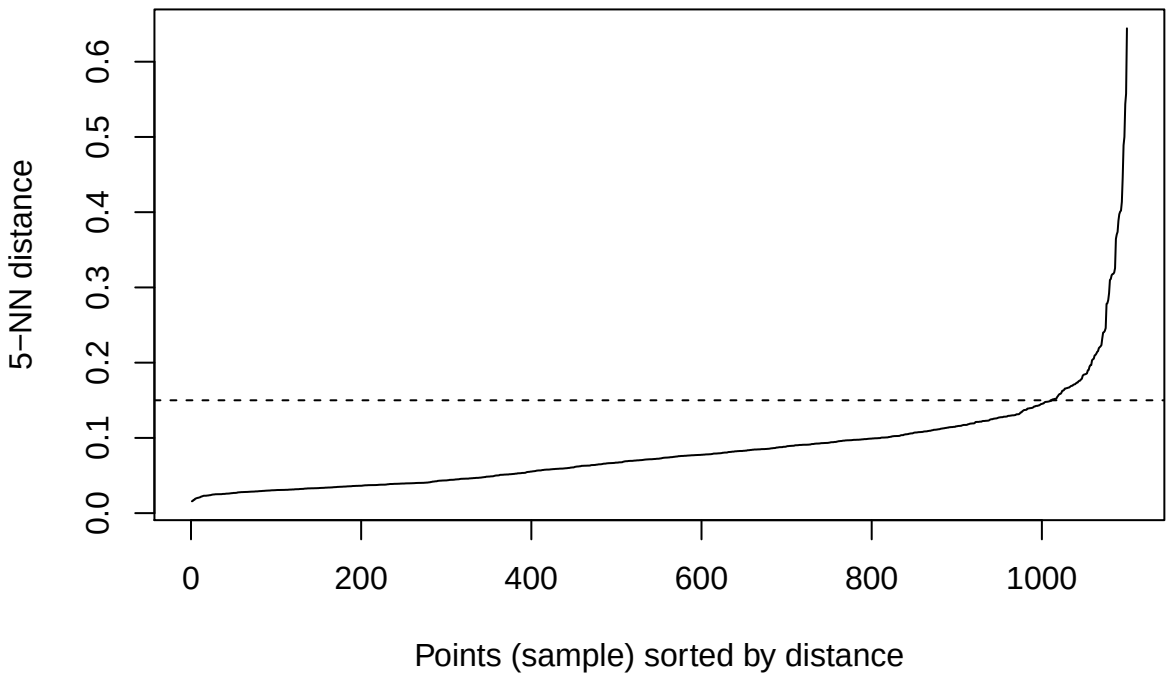
```
library(dbscan)
```

```
## Warning: package 'dbscan' was built under R version 4.0.3
```

```
##
## Attaching package: 'dbscan'
```

```
## The following object is masked from 'package:fpc':
##
##     dbscan
```

```
kNNdistplot(df, k = 5)
abline(h = 0.15, lty = 2)
```



Latihan

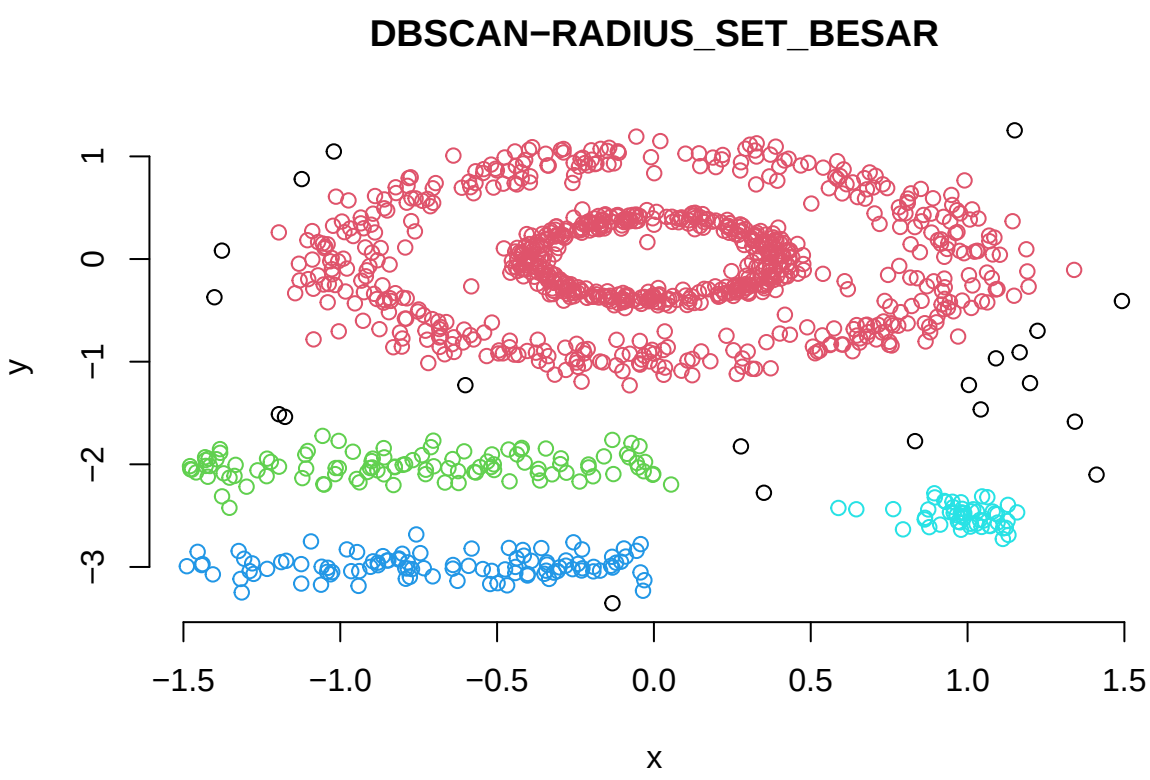
1. Jika radius di ubah dan minPts = 5 tetap

```
library(fpc)
library(factoextra)
set.seed(123)

##Radius Semakin Besar
db <- dbscan(df, eps = 0.20, MinPts = 5)
```

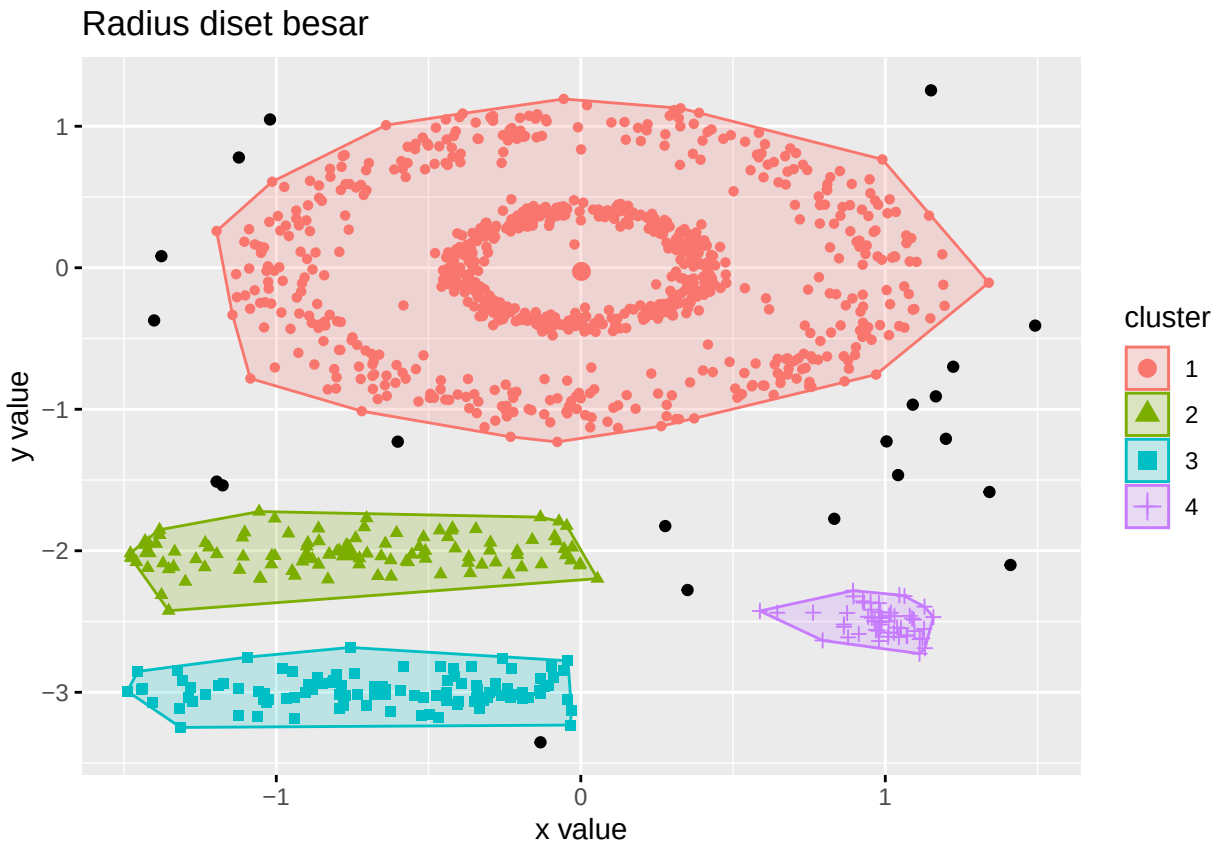
```
## Warning in dbscan(df, eps = 0.2, MinPts = 5): converting argument MinPts (fpc)
## to minPts (dbscan)!
```

```
plot(db, df, main = "DBSCAN-RADIUS_SET_BESAR", frame = FALSE)
```




```
fviz_cluster(db, df, main = "Radius diset besar", stand = FALSE, frame = FALSE, geom = "point")
```

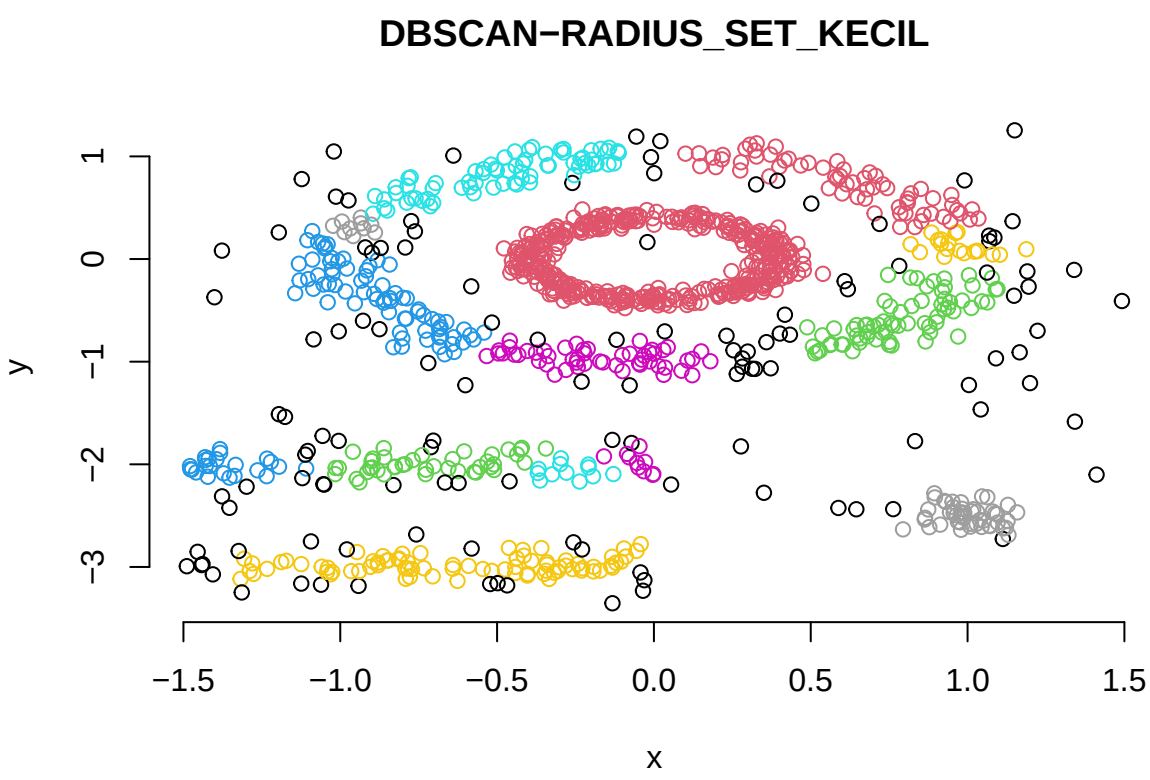
```
## Warning: argument frame is deprecated; please use ellipse instead.
```




```
##Radius Semakin Kecil
db <- dbscan(df, eps = 0.10, MinPts = 5)
```

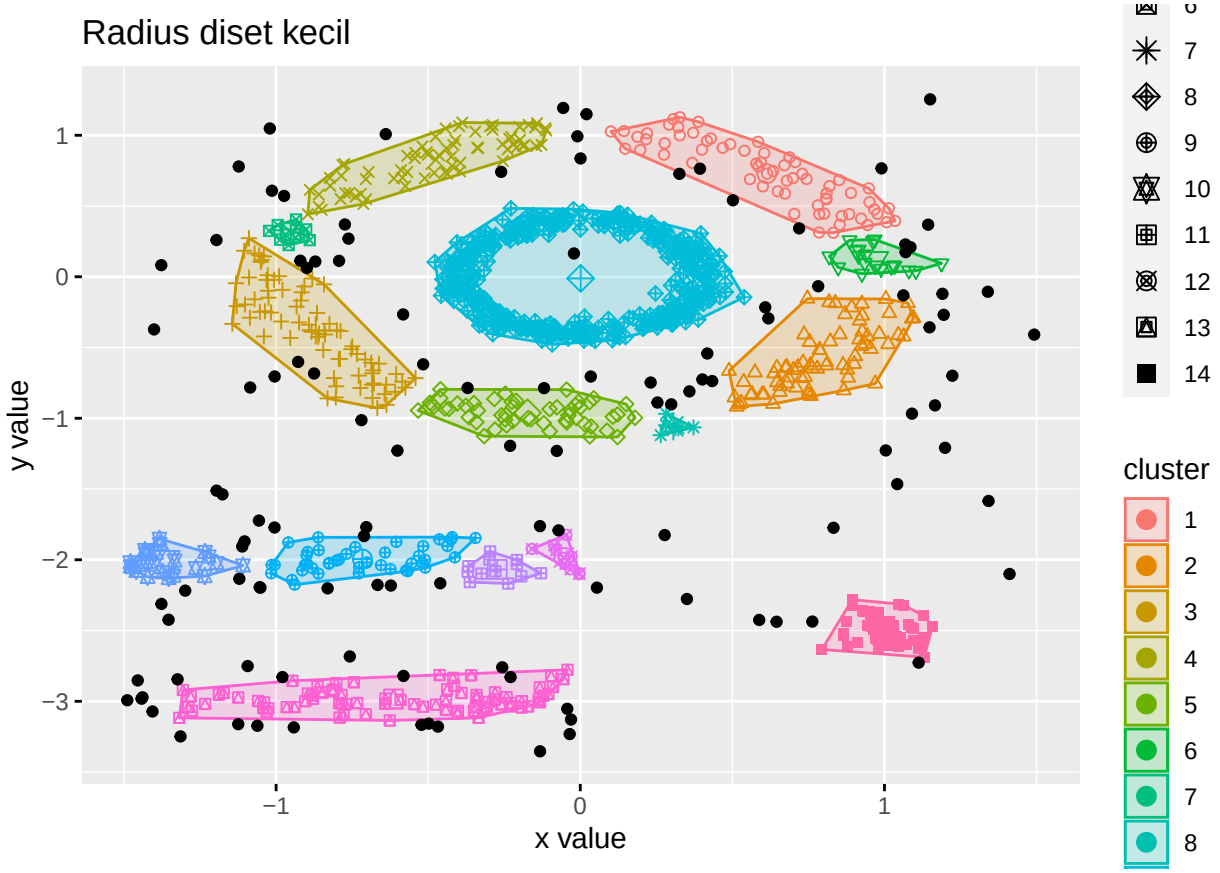
```
## Warning in dbscan(df, eps = 0.1, MinPts = 5): converting argument MinPts (fpc)
## to minPts (dbscan)!
```

```
plot(db, df, main = "DBSCAN-RADIUS_SET_KECIL", frame = FALSE)
```



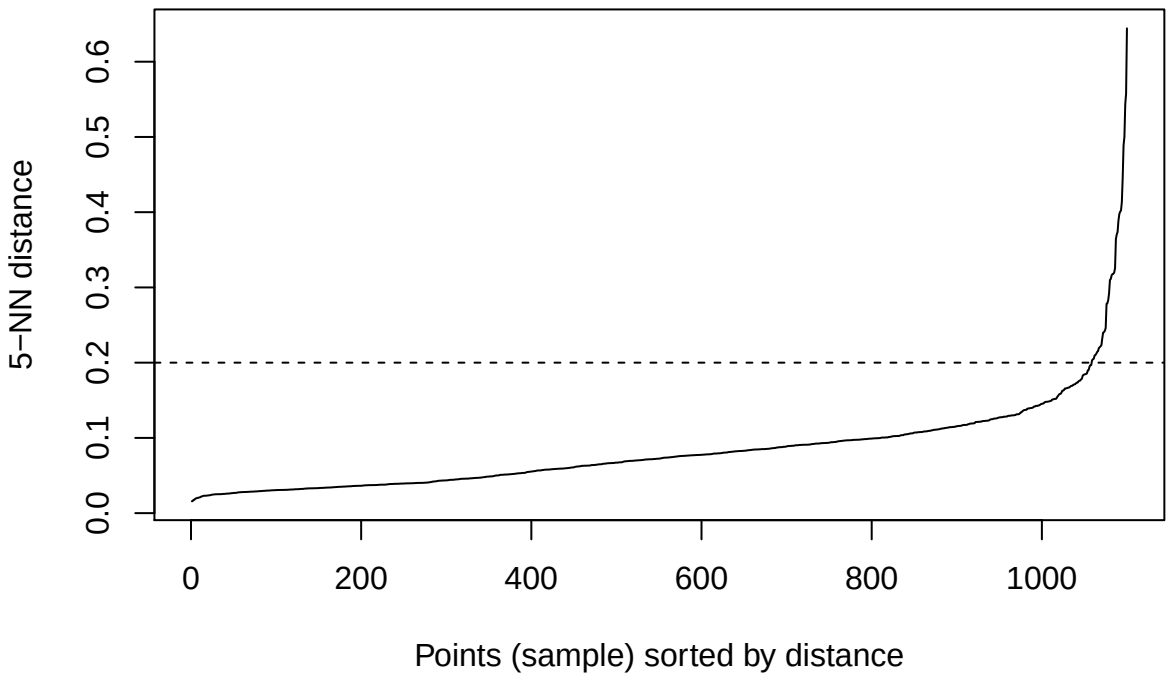

```
fviz_cluster(db, df, main = "Radius diset kecil", stand = FALSE, frame = FALSE, geom = "point")
```

```
## Warning: argument frame is deprecated; please use ellipse instead.
```

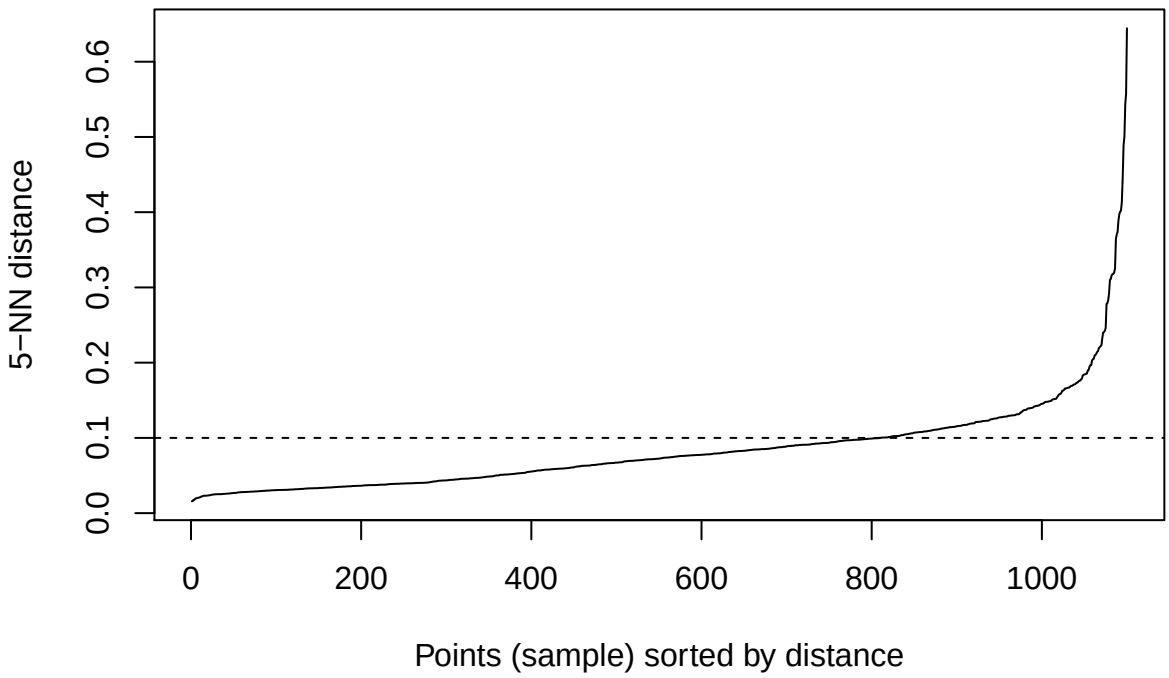



```
library(dbscan)

##Radius Semakin Besar
kNNdistplot(df, k = 5)
abline(h = 0.20, lty = 2)
```




```
##Radius Semakin Kecil
kNNdistplot(df, k = 5)
abline(h = 0.10, lty = 2)
```



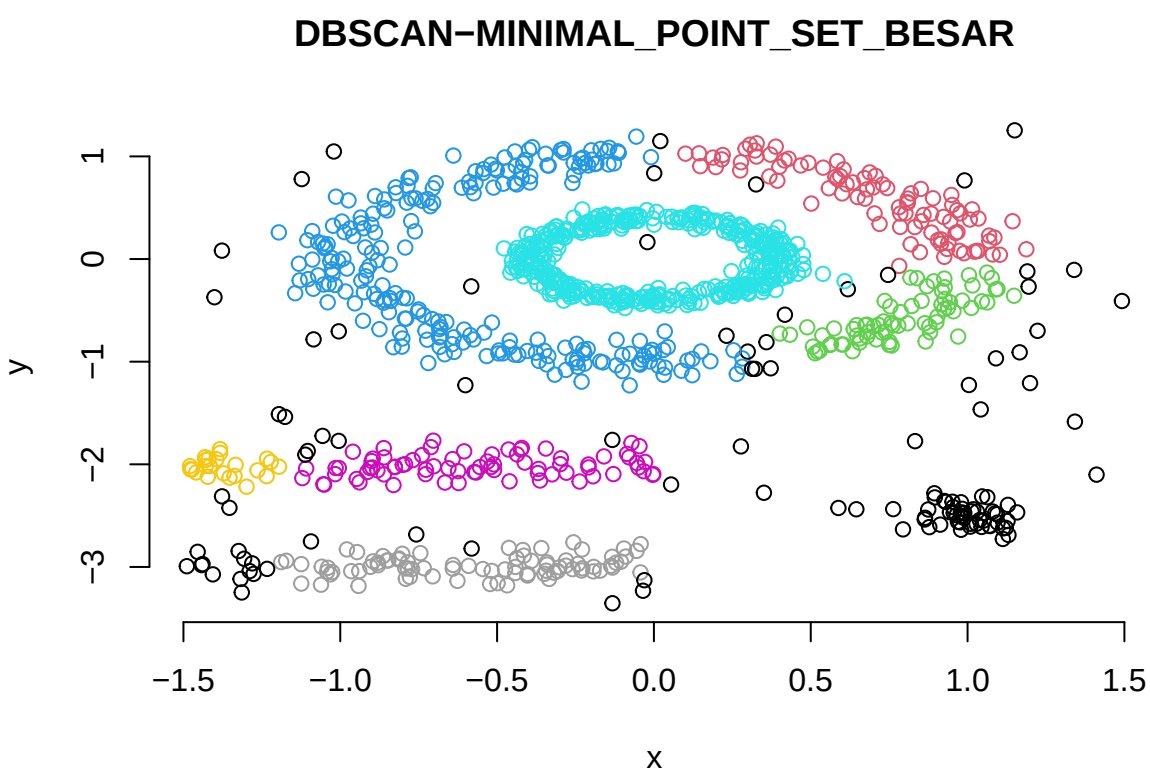
2. Jika minPts di ubah dan radius = 0.15 tetap

```
library(fpc)
library(factoextra)
set.seed(123)

##Minimal Point Semakin Besar
db <- dbscan(df, eps = 0.15, MinPts = 10)
```

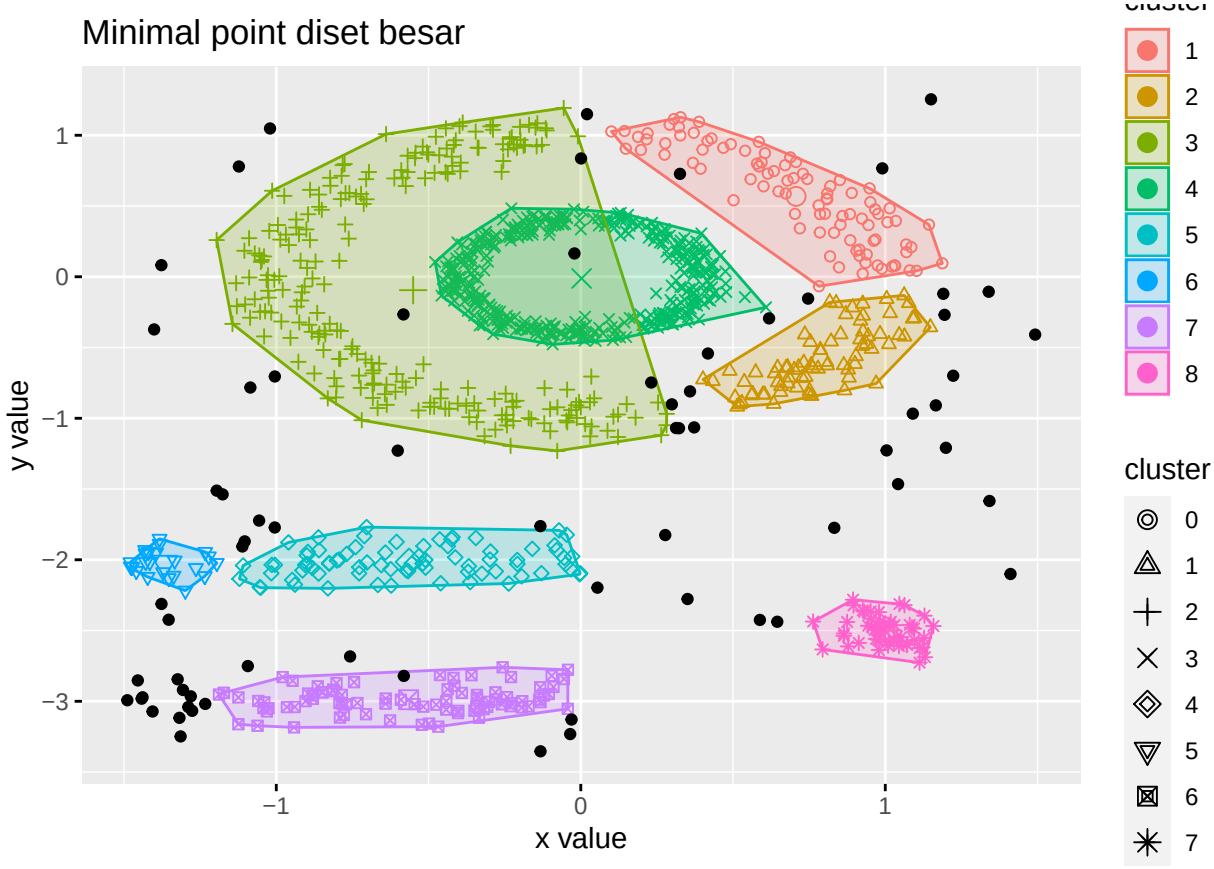
```
## Warning in dbscan(df, eps = 0.15, MinPts = 10): converting argument MinPts (fpc)
## to minPts (dbscan)!
```

```
plot(db, df, main = "DBSCAN-MINIMAL_POINT_SET_BESAR", frame = FALSE)
```




```
fviz_cluster(db, df, main = "Minimal point diset besar", stand = FALSE, frame = FALSE, geom = "point")
```

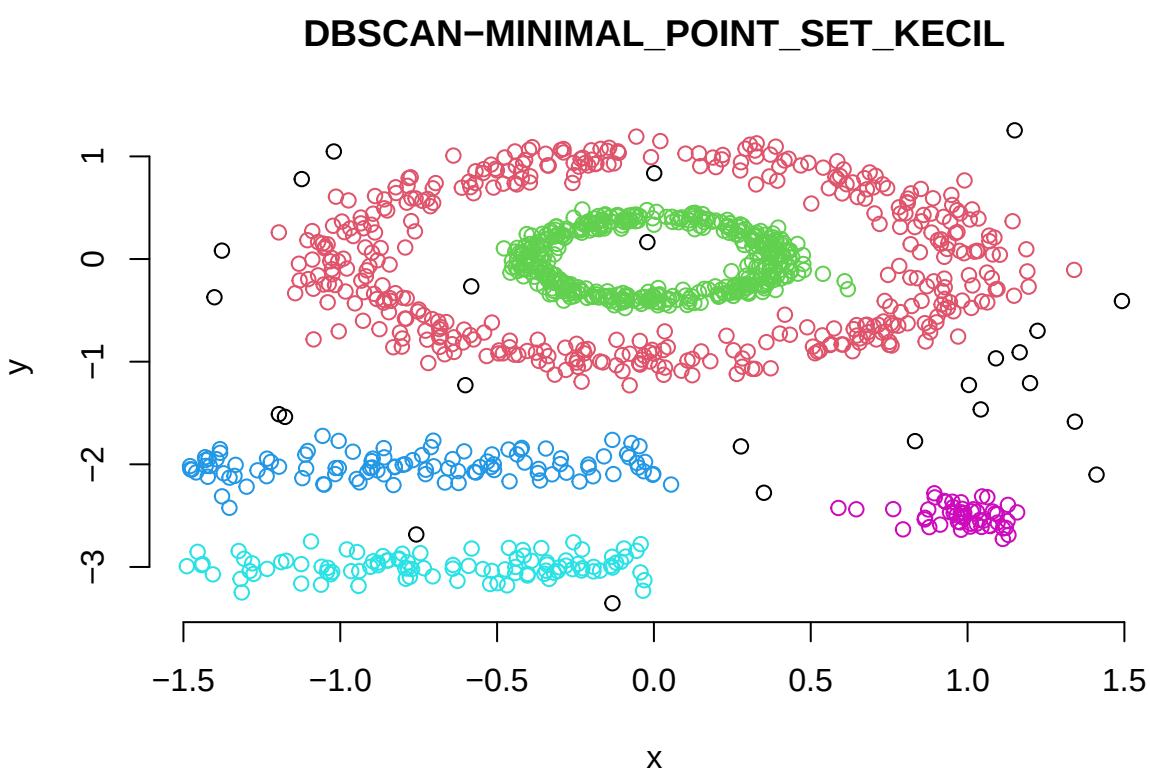
```
## Warning: argument frame is deprecated; please use ellipse instead.
```




```
##Minimal Point Semakin Kecil
db <- dbscan(df, eps = 0.15, MinPts = 3)
```

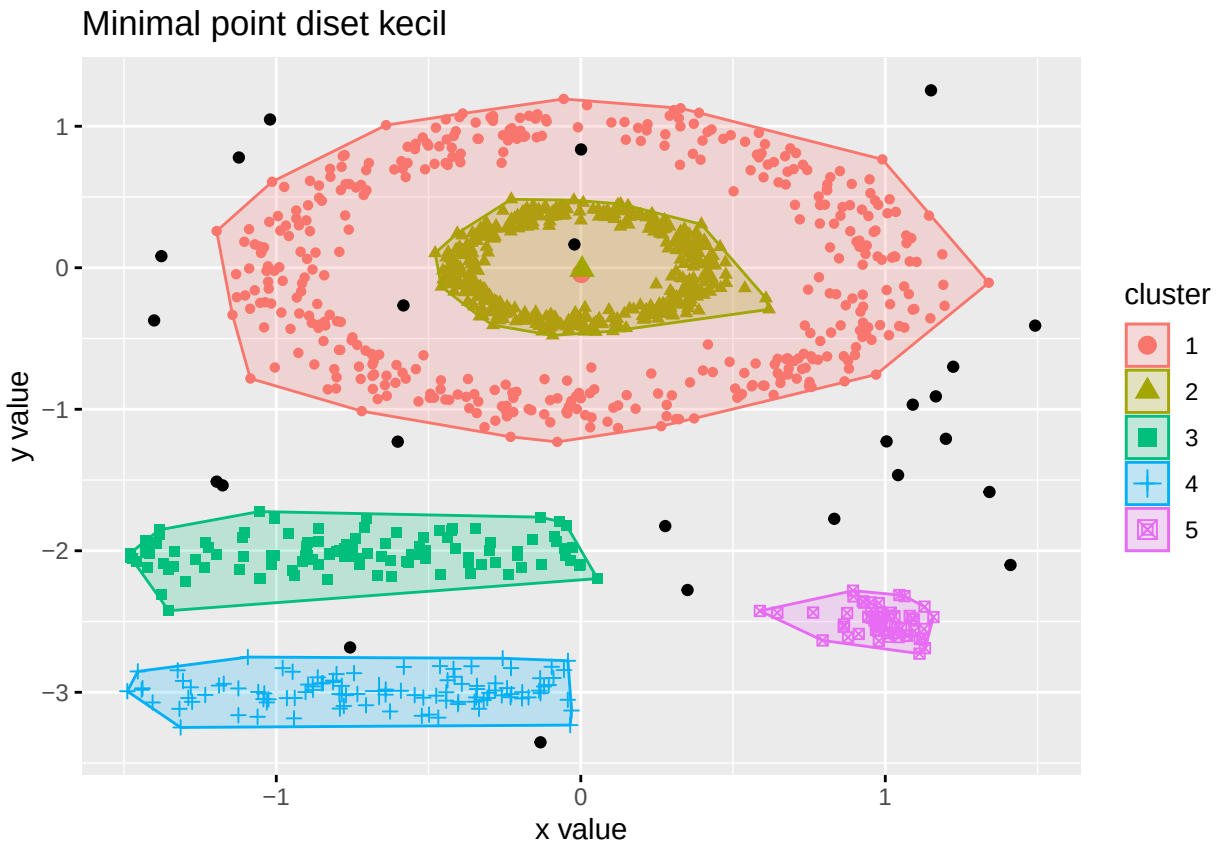
```
## Warning in dbscan(df, eps = 0.15, MinPts = 3): converting argument MinPts (fpc)
## to minPts (dbscan)!
```

```
plot(db, df, main = "DBSCAN-MINIMAL_POINT_SET_KECIL", frame = FALSE)
```

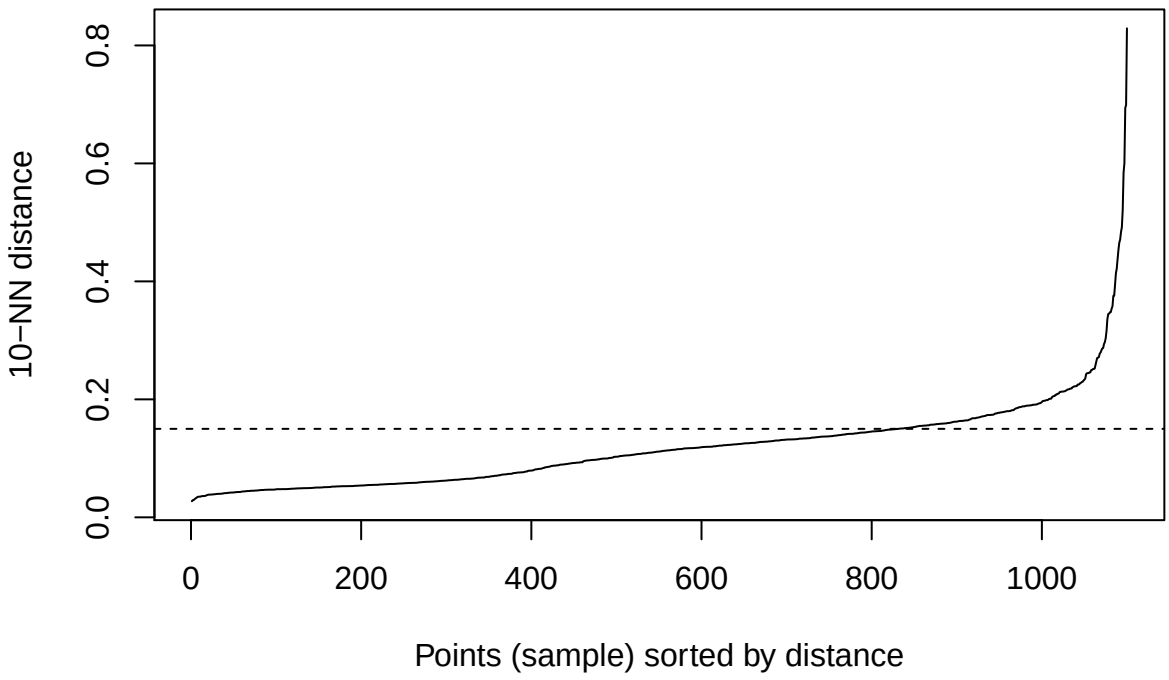



```
fviz_cluster(db, df, main = "Minimal point diset kecil", stand = FALSE, frame = FALSE, geom = "point")
```

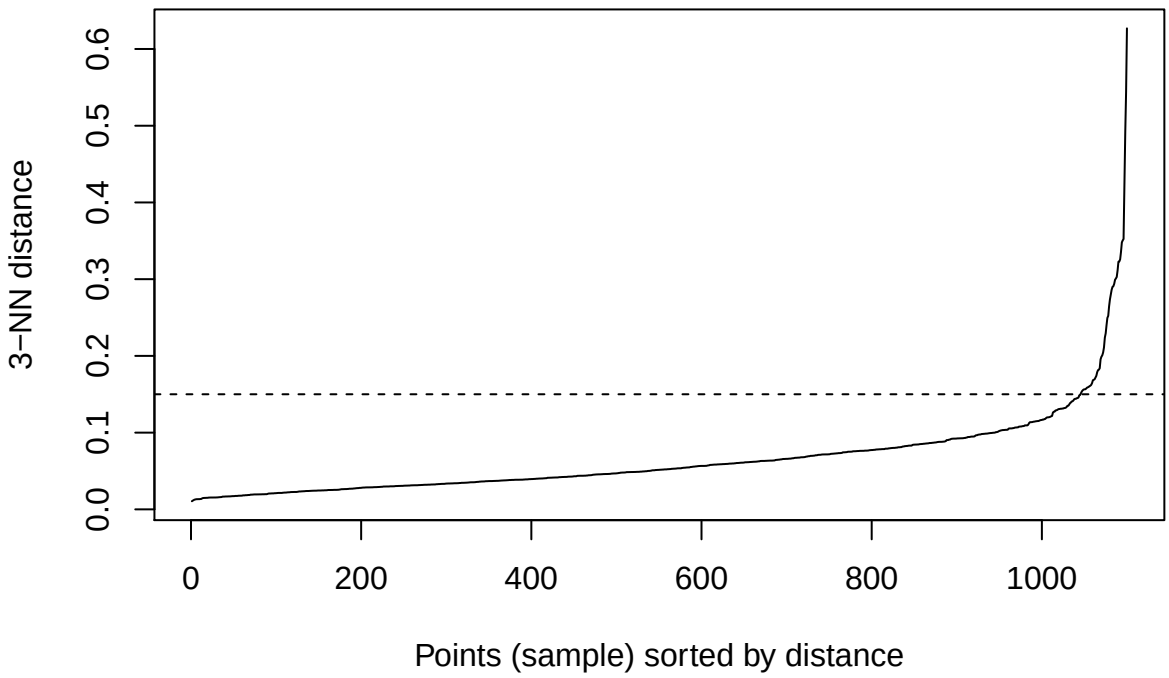
```
## Warning: argument frame is deprecated; please use ellipse instead.
```




```
library(dbscan)
##Minimal Point Semakin Besar
kNNdistplot(df, k = 10)
abline(h = 0.15, lty = 2)
```




```
##Minimal Point Semakin Kecil
kNNdistplot(df, k = 3)
abline(h = 0.15, lty = 2)
```



Kesimpulan

Jika radius di set semakin besar maka beberapa noise akan ikut menjadi bagian dari claster sedangkan jika radius di set semakin kecil maka noise akan bertambah begitupun dengan banyaknya claster yang terbentuk.

Jika minimal point di set semakin besar maka akan terbentuk beberapa claster baru dan noise nya bertambah sedangkan jika minimal point di set semakin kecil maka claster dan noise akan berkurang.