28-Tugas Besar Kuliah FSD, Kelas I

Ligar Wisnu Geogiza (19523141) - Wilwatikta Tripawangsa (19523233)

12/18/2020

Case of judicial-expenditures-across-all-50-states

Dari kaggle.com kami mengambil sample data mengenai Pengeluaran pemerintah dan pekerjaan untuk perlindungan polisi, fungsi peradilan dan hukum, dan koreksi di Amerika Serikat pada tahun 2016. Khususnya pengeluaran keadilan per kapita/PC (fiskal 2016) dan pekerjaan keadilan yang setara penuh waktu per 10.000 penduduk (Maret 2016) pemerintah negara bagian dan lokal berdasarkan aktivitas dan negara bagian.

```
library(factoextra)
```

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

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

```
judedata = read.csv("jeee16t08.csv")
head(judedata)
```

```
##        State Population.2016 Total.justice.system.PC Police.Protection.PC
## 1      Total       323071342                  722.70                338.04
## 2    Alabama         4864745                  480.11                257.21
## 3     Alaska          741504                 1297.65                499.27
## 4    Arizona         6945452                  709.77                325.62
## 5   Arkansas         2990410                  503.99                231.09
## 6 California        39209127                 1063.89                448.11
##   Judicial.and.legal.PC Corrections.PC Total.justice.system.Employment
## 1                143.18         241.48                           63.12
## 2                 74.43         148.47                           55.43
## 3                342.55         455.84                           77.18
## 4                141.59         242.56                           66.07
## 5                 73.68         199.21                           68.44
## 6                221.27         394.51                           59.85
##   police.protection.Total.Employment police.protection.Sworn.only.Employment
## 1                              28.51                                    21.73
## 2                              29.08                                    23.00
## 3                              25.68                                    15.58
## 4                              28.30                                    20.43
## 5                              29.50                                    22.11
## 6                              25.65                                    18.27
##   Judicial.and.legal.Employment Corrections.Employment
## 1                         12.83                 21.78
## 2                          9.73                 16.62
```

```
## 3                          20.08                 31.42
## 4                          15.91                 21.85
## 5                          11.52                 27.41
## 6                          11.31                 22.89
```

kita lihat summary dan struktur datanya.

```
summary(judedata)
```

```
##     State           Population.2016     Total.justice.system.PC
##  Length:52          Min.   :   584290   Min.   : 450.1
##  Class :character   1st Qu.:  1793929   1st Qu.: 557.7
##  Mode  :character   Median :  4558222   Median : 667.8
##                     Mean   : 12425821   Mean   : 690.7
##                     3rd Qu.:  7573746   3rd Qu.: 746.6
##                     Max.   :323071342   Max.   :1297.7
##  Police.Protection.PC Judicial.and.legal.PC Corrections.PC
##  Min.   :160.3        Min.   : 72.35        Min.   :141.9
##  1st Qu.:262.2        1st Qu.:112.00        1st Qu.:177.1
##  Median :297.6        Median :134.27        Median :209.2
##  Mean   :322.2        Mean   :141.86        Mean   :226.6
##  3rd Qu.:352.7        3rd Qu.:160.26        3rd Qu.:248.9
##  Max.   :864.2        Max.   :342.55        Max.   :455.8
##  Total.justice.system.Employment police.protection.Total.Employment
##  Min.   : 42.66                  Min.   :19.76
##  1st Qu.: 55.80                  1st Qu.:24.90
##  Median : 61.01                  Median :27.57
##  Mean   : 62.92                  Mean   :28.31
##  3rd Qu.: 66.35                  3rd Qu.:29.65
##  Max.   :112.53                  Max.   :63.36
##  police.protection.Sworn.only.Employment Judicial.and.legal.Employment
##  Min.   :13.68                           Min.   : 7.29
##  1st Qu.:17.82                           1st Qu.:10.79
##  Median :20.95                           Median :12.79
##  Mean   :21.54                           Mean   :13.51
##  3rd Qu.:22.73                           3rd Qu.:15.54
##  Max.   :54.85                           Max.   :29.39
##  Corrections.Employment
##  Min.   :13.60
##  1st Qu.:17.64
##  Median :20.09
##  Mean   :21.09
##  3rd Qu.:24.50
##  Max.   :33.34
```

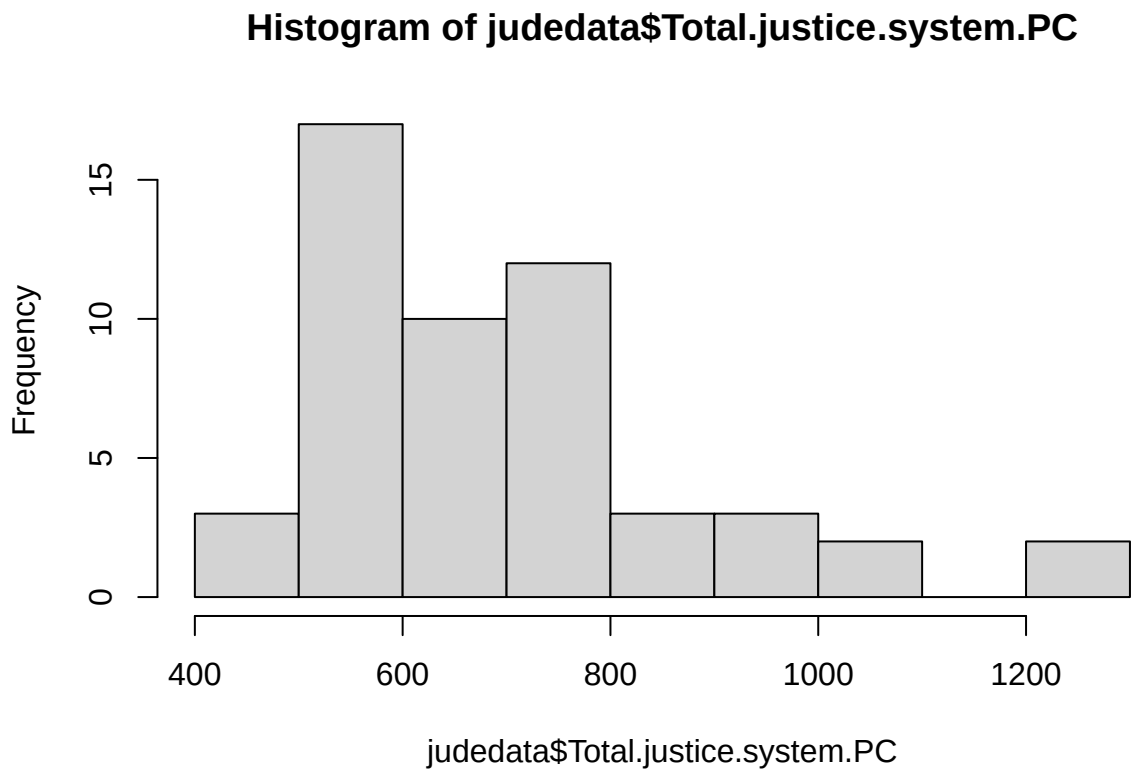
```
str(judedata)
```

```
## 'data.frame':    52 obs. of  11 variables:
##  $ State                                : chr  "Total" "Alabama" "Alaska" "Arizona" ...
##  $ Population.2016                      : int  323071342 4864745 741504 6945452 2990410 39209127 5
##  $ Total.justice.system.PC              : num  723 480 1298 710 504 ...
##  $ Police.Protection.PC                 : num  338 257 499 326 231 ...
```

```
##  $ Judicial.and.legal.PC                 : num  143.2 74.4 342.6 141.6 73.7 ...
##  $ Corrections.PC                        : num  241 148 456 243 199 ...
##  $ Total.justice.system.Employment       : num  63.1 55.4 77.2 66.1 68.4 ...
##  $ police.protection.Total.Employment    : num  28.5 29.1 25.7 28.3 29.5 ...
##  $ police.protection.Sworn.only.Employment: num  21.7 23 15.6 20.4 22.1 ...
##  $ Judicial.and.legal.Employment         : num  12.83 9.73 20.08 15.91 11.52 ...
##  $ Corrections.Employment                : num  21.8 16.6 31.4 21.9 27.4 ...
```

untuk melihat total keadilan sistem perkapita kita sampling dengan histogram

```
hist(judedata$Total.justice.system.PC)
```



##Pre-processing

```
library(caTools)

set.seed(99)
split = sample.split(judedata$Population.2016, SplitRatio = 1.25)

training_set = subset(judedata, split == TRUE)
test_set = subset(judedata, split == FALSE)
```

```
training_set
```

```
##       State Population.2016 Total.justice.system.PC Police.Protection.PC
```

```
## 29 Nebraska         1905924                  600.27               240.71
##    Judicial.and.legal.PC Corrections.PC Total.justice.system.Employment
## 29                 97.48         262.08                           60.62
##    police.protection.Total.Employment police.protection.Sworn.only.Employment
## 29                              26.11                                   19.32
##    Judicial.and.legal.Employment Corrections.Employment
## 29                         11.06                  23.45
```

```
test_set
```

```
##                   State Population.2016 Total.justice.system.PC
## 1                 Total       323071342                  722.70
## 2               Alabama         4864745                  480.11
## 3                Alaska          741504                 1297.65
## 4               Arizona         6945452                  709.77
## 5              Arkansas         2990410                  503.99
## 6            California        39209127                 1063.89
## 7              Colorado         5540921                  711.18
## 8           Connecticut         3578674                  767.90
## 9              Delaware          949216                  910.60
## 10 District of Columbia          686575                 1268.29
## 11              Florida        20629982                  701.08
## 12              Georgia        10304763                  568.02
## 13               Hawaii         1428105                  679.89
## 14                Idaho         1682930                  593.95
## 15             Illinois        12826895                  671.84
## 16              Indiana         6633344                  450.08
## 17                 Iowa         3131785                  504.91
## 18               Kansas         2911263                  564.97
## 19             Kentucky         4438229                  505.74
## 20            Louisiana         4678215                  741.85
## 21                Maine         1331370                  504.40
## 22             Maryland         6004692                  912.84
## 23        Massachusetts         6826022                  689.91
## 24             Michigan         9951890                  620.48
## 25            Minnesota         5523409                  663.84
## 26          Mississippi         2988298                  507.07
## 27             Missouri         6087203                  536.46
## 28              Montana         1040863                  729.41
## 30               Nevada         2919772                  812.32
## 31        New Hampshire         1342373                  546.03
## 32           New Jersey         8874516                  791.11
## 33           New Mexico         2092789                  824.87
## 34             New York        19641589                 1048.95
## 35       North Carolina        10156679                  576.85
## 36         North Dakota          754353                  675.36
## 37                 Ohio        11635003                  656.22
## 38             Oklahoma         3926769                  558.29
## 39               Oregon         4091404                  808.33
## 40         Pennsylvania        12783538                  714.15
## 41         Rhode Island         1057063                  760.94
## 42       South Carolina         4958235                  465.48
## 43         South Dakota          862890                  547.21
## 44            Tennessee         6645011                  571.23
```

```
## 45                Texas        27937492                  592.19
## 46                 Utah         3042613                  503.63
## 47              Vermont          623644                  650.75
## 48             Virginia         8410946                  705.52
## 49           Washington         7294680                  660.33
## 50        West Virginia         1830929                  555.79
## 51            Wisconsin         5772958                  708.58
## 52              Wyoming          584290                  997.13
##    Police.Protection.PC Judicial.and.legal.PC Corrections.PC
## 1                338.04                143.18         241.48
## 2                257.21                 74.43         148.47
## 3                499.27                342.55         455.84
## 4                325.62                141.59         242.56
## 5                231.09                 73.68         199.21
## 6                448.11                221.27         394.51
## 7                338.09                136.11         236.98
## 8                345.66                231.06         191.18
## 9                366.65                219.12         324.84
## 10               864.21                193.01         211.07
## 11               380.44                114.70         205.94
## 12               263.81                117.93         186.28
## 13               321.41                206.21         152.27
## 14               271.63                118.67         203.64
## 15               406.65                120.78         144.41
## 16               189.20                 86.20         174.68
## 17               248.60                114.46         141.86
## 18               274.42                113.44         177.11
## 19               160.32                141.84         203.58
## 20               345.07                153.43         243.34
## 21               221.53                 84.49         198.39
## 22               441.17                155.37         316.30
## 23               366.53                159.38         164.00
## 24               252.10                115.04         253.35
## 25               348.04                132.43         183.37
## 26               248.46                 92.00         166.61
## 27               286.44                 92.24         157.78
## 28               278.90                203.06         247.45
## 30               405.63                162.91         243.78
## 31               293.32                104.86         147.84
## 32               388.21                172.68         230.22
## 33               334.82                157.37         332.68
## 34               505.16                211.65         332.14
## 35               322.10                 75.23         179.52
## 36               275.78                167.42         232.15
## 37               333.68                150.09         172.46
## 38               271.19                 95.56         191.54
## 39               321.55                189.50         297.27
## 40               290.79                148.86         274.49
## 41               412.11                138.70         210.13
## 42               240.77                 72.35         152.37
## 43               233.25                105.77         208.19
## 44               279.13                115.18         176.92
## 45               272.58                107.66         211.95
## 46               215.29                115.28         173.06
```

```
## 47              301.79              129.11         219.84
## 48              277.95              123.12         304.45
## 49              267.91              155.59         236.83
## 50              211.49              151.51         192.79
## 51              324.53              114.83         269.22
## 52              413.82              216.09         367.22
##    Total.justice.system.Employment police.protection.Total.Employment
## 1                            63.12                                28.51
## 2                            55.43                                29.08
## 3                            77.18                                25.68
## 4                            66.07                                28.30
## 5                            68.44                                29.50
## 6                            59.85                                25.65
## 7                            62.81                                28.12
## 8                            61.58                                28.14
## 9                            78.60                                27.45
## 10                          112.53                                63.36
## 11                           64.42                                29.69
## 12                           66.31                                26.65
## 13                           68.62                                28.54
## 14                           61.64                                24.05
## 15                           60.34                                31.95
## 16                           53.15                                23.61
## 17                           47.08                                22.95
## 18                           66.26                                31.24
## 19                           55.93                                21.95
## 20                           79.39                                37.75
## 21                           42.66                                21.34
## 22                           71.70                                31.73
## 23                           61.35                                29.64
## 24                           49.58                                20.81
## 25                           51.89                                22.52
## 26                           57.69                                32.37
## 27                           67.90                                30.41
## 28                           59.66                                25.39
## 30                           60.66                                26.60
## 31                           50.31                                28.67
## 32                           76.64                                38.70
## 33                           76.91                                29.69
## 34                           86.98                                44.03
## 35                           60.49                                27.51
## 36                           58.50                                25.09
## 37                           62.78                                26.98
## 38                           57.15                                29.57
## 39                           54.55                                21.79
## 40                           63.48                                25.13
## 41                           54.64                                28.92
## 42                           60.16                                28.45
## 43                           53.08                                24.34
## 44                           61.78                                31.61
## 45                           61.91                                27.25
## 46                           47.71                                20.83
## 47                           52.69                                23.94
## 48                           66.47                                26.06
```

```
## 49                            49.34                              19.76
## 50                            58.25                              22.48
## 51                            60.33                              27.63
## 52                            85.03                              34.78
##    police.protection.Sworn.only.Employment Judicial.and.legal.Employment
## 1                                    21.73                           12.83
## 2                                    23.00                            9.73
## 3                                    15.58                           20.08
## 4                                    20.43                           15.91
## 5                                    22.11                           11.52
## 6                                    18.27                           11.31
## 7                                    19.81                           14.09
## 8                                    22.43                           17.23
## 9                                    21.58                           20.28
## 10                                   54.85                           29.39
## 11                                   19.78                           15.12
## 12                                   21.33                           13.89
## 13                                   22.12                           23.58
## 14                                   16.55                           13.14
## 15                                   25.49                           10.68
## 16                                   17.99                           11.39
## 17                                   17.32                           10.15
## 18                                   22.14                           13.23
## 19                                   17.20                           16.05
## 20                                   28.80                           15.42
## 21                                   17.34                            7.29
## 22                                   23.08                           14.21
## 23                                   24.10                           13.99
## 24                                   16.68                           10.77
## 25                                   16.53                           11.43
## 26                                   24.30                            8.92
## 27                                   22.74                           12.74
## 28                                   17.39                           15.93
## 30                                   17.18                           14.73
## 31                                   21.96                            8.05
## 32                                   29.95                           21.54
## 33                                   22.62                           17.54
## 34                                   38.35                           15.90
## 35                                   22.46                            7.43
## 36                                   20.95                           12.99
## 37                                   20.28                           17.03
## 38                                   20.95                           10.80
## 39                                   14.58                           11.51
## 40                                   20.88                           14.17
## 41                                   22.89                           12.02
## 42                                   22.72                            9.68
## 43                                   19.46                           11.47
## 44                                   25.46                           11.32
## 45                                   19.87                           10.00
## 46                                   14.38                            9.83
## 47                                   16.02                           11.30
## 48                                   20.93                           10.71
## 49                                   13.68                           11.54
## 50                                   17.97                           15.30
```

```
## 51                                21.22                          9.40
## 52                                25.28                         16.91
##    Corrections.Employment
## 1                   21.78
## 2                   16.62
## 3                   31.42
## 4                   21.85
## 5                   27.41
## 6                   22.89
## 7                   20.60
## 8                   16.20
## 9                   30.87
## 10                  19.78
## 11                  19.61
## 12                  25.77
## 13                  16.50
## 14                  24.44
## 15                  17.71
## 16                  18.14
## 17                  13.98
## 18                  21.79
## 19                  17.92
## 20                  26.22
## 21                  14.03
## 22                  25.76
## 23                  17.72
## 24                  18.00
## 25                  17.95
## 26                  16.40
## 27                  24.75
## 28                  18.34
## 30                  19.32
## 31                  13.60
## 32                  16.40
## 33                  29.68
## 34                  27.05
## 35                  25.55
## 36                  20.41
## 37                  18.77
## 38                  16.78
## 39                  21.25
## 40                  24.18
## 41                  13.70
## 42                  22.03
## 43                  17.27
## 44                  18.85
## 45                  24.67
## 46                  17.04
## 47                  17.45
## 48                  29.70
## 49                  18.04
## 50                  20.47
## 51                  23.29
## 52                  33.34
```

##Cross-validation

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
## v purrr   0.3.4

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(caret)
```

```
## Loading required package: lattice

##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
##     lift
```

```
set.seed(99)
```

```
sample <- createDataPartition(judedata$Population.2016, p = 0.75, list = FALSE)
```

```
training_judedata <- judedata[sample, ]
```

```
testing_judedata <- judedata[-sample, ]
```

```
model <- lm(Population.2016 ~ Total.justice.system.PC, data = training_judedata)
```

```
predictions <- predict(model, testing_judedata)
```

```
data.frame( R2 = R2(predictions, testing_judedata $ Population.2016),
            RMSE = RMSE(predictions, testing_judedata $ Population.2016),
            MAE = MAE(predictions, testing_judedata $ Population.2016))
```

```
##          R2     RMSE     MAE
## 1 0.3951158 11266342 9862963
```

Pemodelan

Data ini memiliki atribut sebanyak 11 multivariate berkarakteristik integer, real, dan string(pada state).

Model pembelajaran mesin yang digunakan adalah DBSCAN Density-Based Spatial Clustering of Application with Noise (DBSCAN) merupakan sebuah metode clustering yang membangun area berdasarkan kepadatan yang terkoneksi (density- connected). Setiap objek dari sebuah radius area (cluster) yang mengandung setidaknya sejumlah minimum data.

Evaluasi, akurasi dan apakah masuk akal ? Penggunaan algoritma DBSCAN untuk pengolahan data clustering, regresi dan klasifikasi. Dengan kerapatan titik-titiknya.

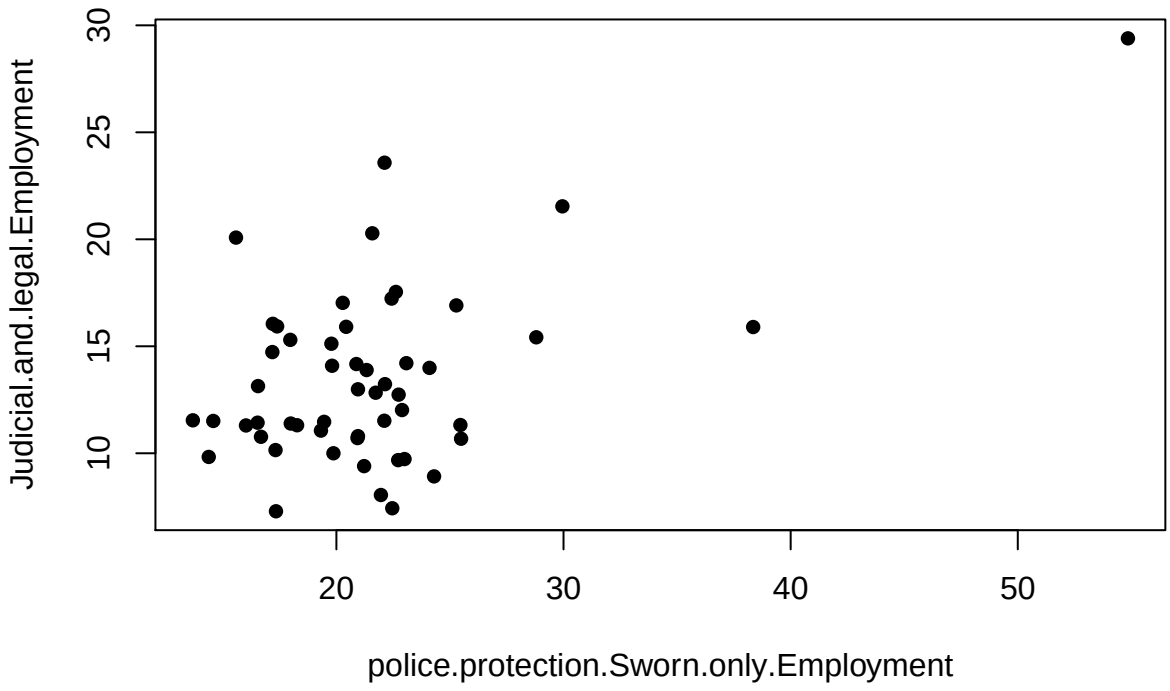
Implementasi DBSCAN

```
df <-judedata[, 9:10 ]
df
```

```
##    police.protection.Sworn.only.Employment Judicial.and.legal.Employment
## 1                                    21.73                          12.83
## 2                                    23.00                           9.73
## 3                                    15.58                          20.08
## 4                                    20.43                          15.91
## 5                                    22.11                          11.52
## 6                                    18.27                          11.31
## 7                                    19.81                          14.09
## 8                                    22.43                          17.23
## 9                                    21.58                          20.28
## 10                                   54.85                          29.39
## 11                                   19.78                          15.12
## 12                                   21.33                          13.89
## 13                                   22.12                          23.58
## 14                                   16.55                          13.14
## 15                                   25.49                          10.68
## 16                                   17.99                          11.39
## 17                                   17.32                          10.15
## 18                                   22.14                          13.23
## 19                                   17.20                          16.05
## 20                                   28.80                          15.42
## 21                                   17.34                           7.29
## 22                                   23.08                          14.21
## 23                                   24.10                          13.99
## 24                                   16.68                          10.77
## 25                                   16.53                          11.43
## 26                                   24.30                           8.92
## 27                                   22.74                          12.74
## 28                                   17.39                          15.93
## 29                                   19.32                          11.06
## 30                                   17.18                          14.73
## 31                                   21.96                           8.05
## 32                                   29.95                          21.54
## 33                                   22.62                          17.54
## 34                                   38.35                          15.90
## 35                                   22.46                           7.43
## 36                                   20.95                          12.99
## 37                                   20.28                          17.03
## 38                                   20.95                          10.80
## 39                                   14.58                          11.51
## 40                                   20.88                          14.17
## 41                                   22.89                          12.02
```

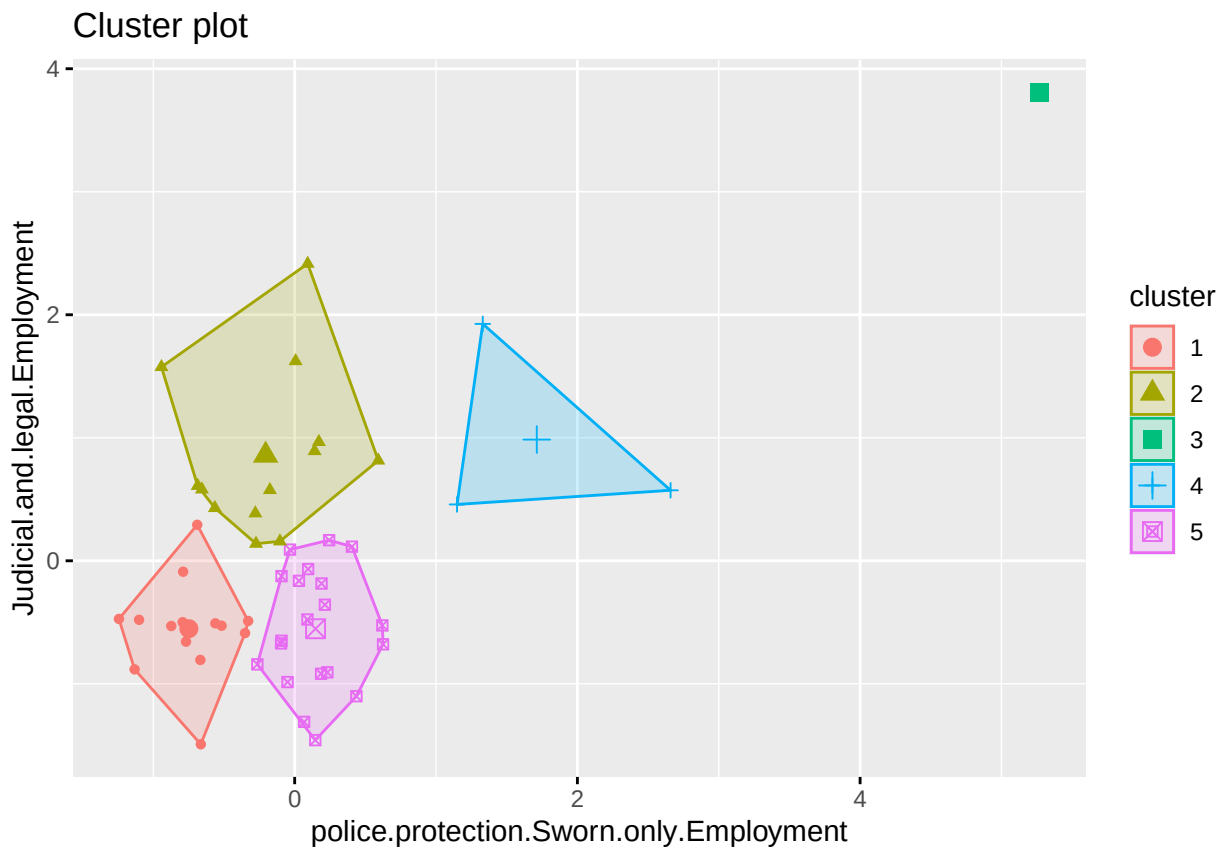
```
## 42                                22.72                                9.68
## 43                                19.46                               11.47
## 44                                25.46                               11.32
## 45                                19.87                               10.00
## 46                                14.38                                9.83
## 47                                16.02                               11.30
## 48                                20.93                               10.71
## 49                                13.68                               11.54
## 50                                17.97                               15.30
## 51                                21.22                                9.40
## 52                                25.28                               16.91
```

```
plot(df, pch=16)
```

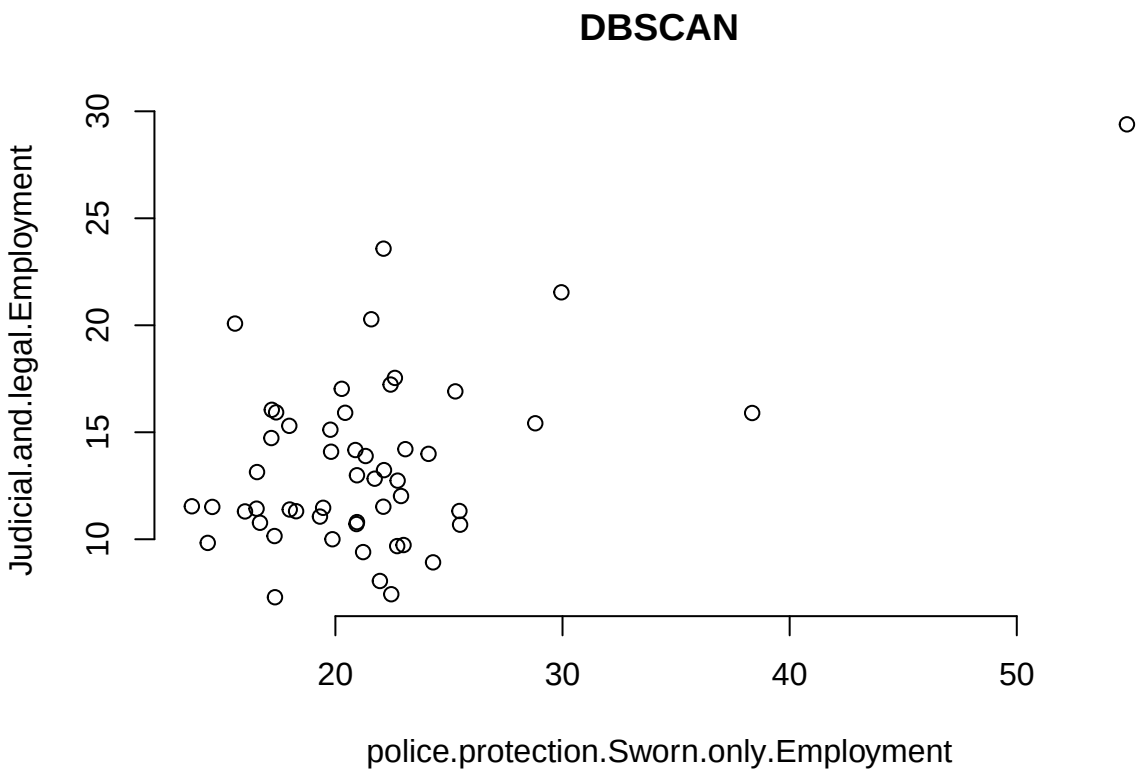



```
set.seed(99)
km <- kmeans(df, 5, nstart = 25)
fviz_cluster(km, df, frame = FALSE, geom = "point")
```

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

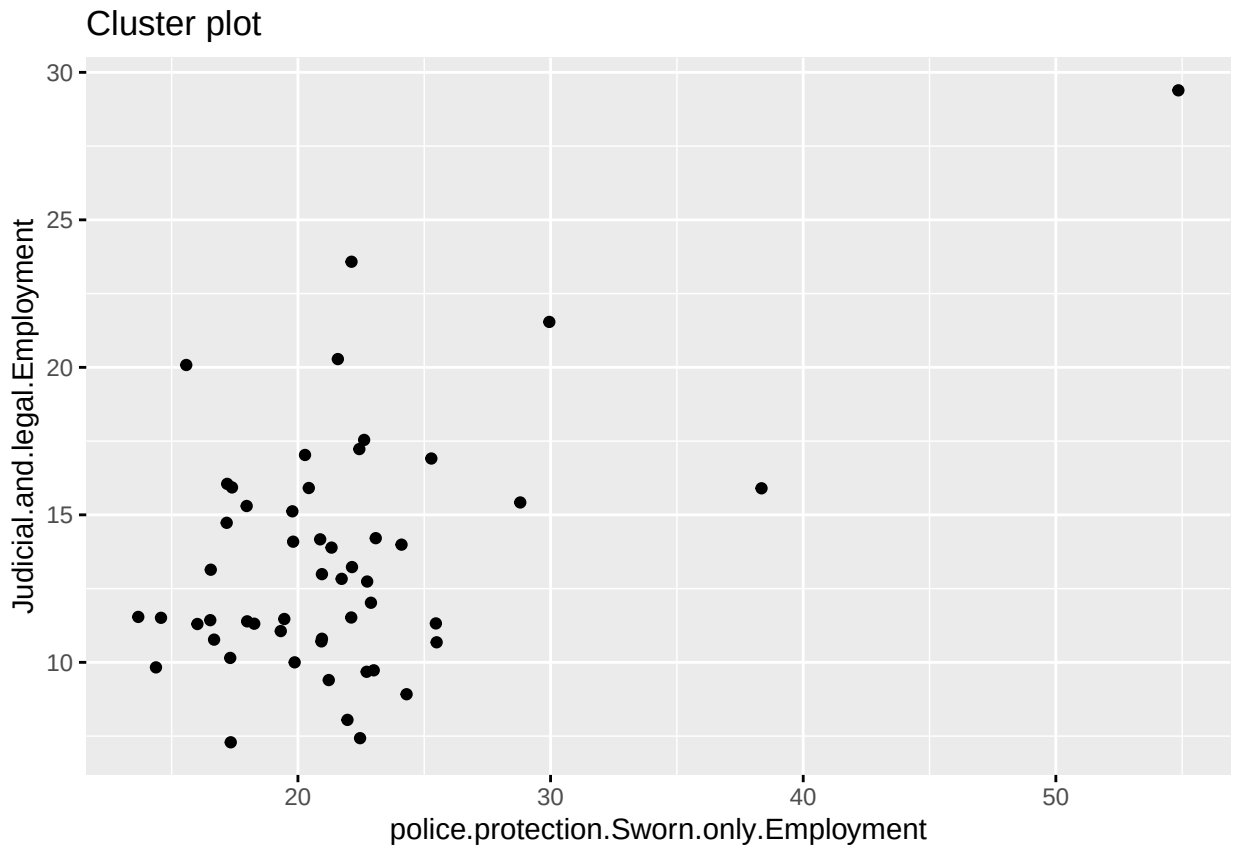



```
library("fpc")
set.seed(99)
db <- dbscan(df, eps = 0.15, MinPts = 5)
plot(db, df, main = "DBSCAN", frame = FALSE)
```




```
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=52 MinPts=5 eps=0.15
##
##  0
## 52
```

```
db$cluster
```

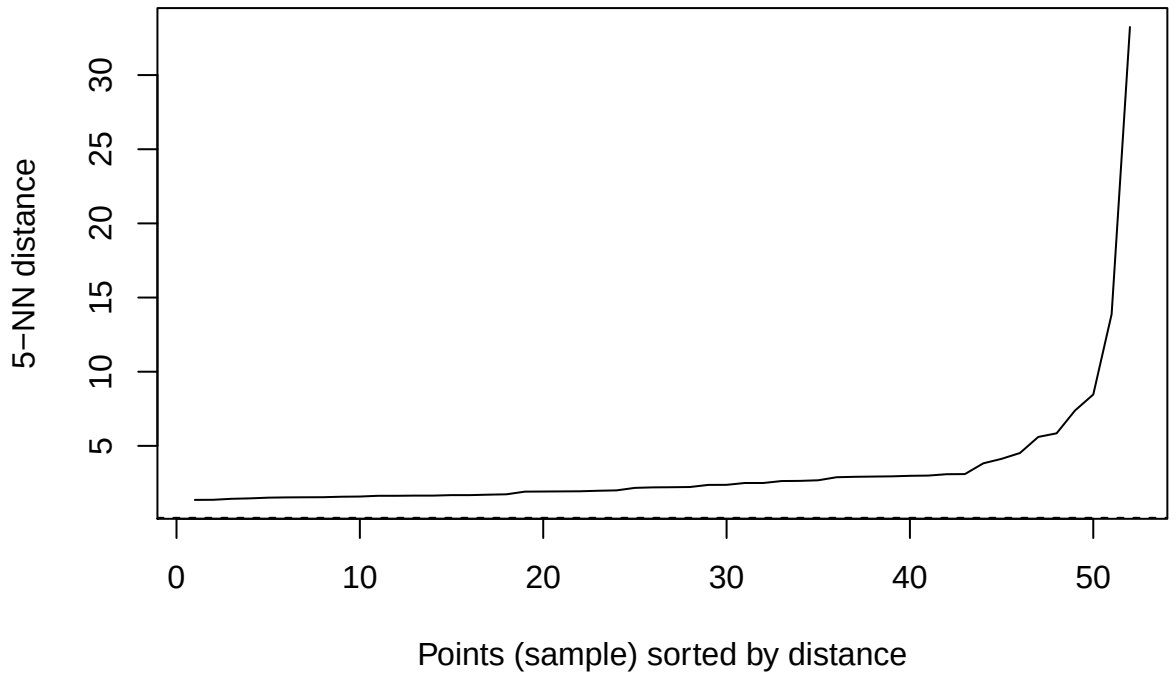
```
##  [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [39] 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

```
library(dbscan)
```

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

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

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



Kesimpulan

Kluster plot memperlihatkan bahwa lembah pada kurva adalah ketika y >= 15. Hasil tersebut yang dinamakan sebagai nilai eps. Alternatif lain untuk menentukan minPts adalah dengan menghitung mean dari nilai k-distance dari semua data point.