

Tugas Modul 4

Liek Allyandaru / 123180054 / Praktikum Data Science B

11/6/2020

```
# setup library  
library(dslabs)  
data(murders)
```

1. fungsi sort untuk mengurutkan variabel “pop”

```
# sorting data  
pop <- sort(murders$population)  
  
# menampilkan data terkecil  
pop[1]  
  
## [1] 563626
```

2. indeks dari data yang memiliki nilai populasi terkecil

```
# mengurutkan semua data dan diambil index terkecilnya  
index <- order(murders$population)  
  
# menampilkan index dengan data terkecil  
index[1]  
  
## [1] 51
```

3. fungsi which.min

```
which.min(murders$population)  
  
## [1] 51
```

4. Negara dengan populasi terkecil

```
# mencari pupulasi terkecil  
index_terkecil <- which.min(murders$population)
```

```
# menampilkan nama negara dengan index_terkecil
murders$state[index_terkecil]

## [1] "Wyoming"
```

5. Rank negara berdasarkan populasi

```
ranks <- rank(murders$population)
my_df <- data.frame(state = murders$state, rank = ranks)
my_df
```

```
##           state rank
## 1      Alabama   29
## 2       Alaska    5
## 3      Arizona   36
## 4      Arkansas   20
## 5    California   51
## 6      Colorado   30
## 7    Connecticut   23
## 8      Delaware    7
## 9 District of Columbia    2
## 10     Florida   49
## 11     Georgia   44
## 12     Hawaii   12
## 13     Idaho    13
## 14    Illinois   47
## 15     Indiana   37
## 16      Iowa    22
## 17     Kansas   19
## 18     Kentucky   26
## 19    Louisiana   27
## 20      Maine    11
## 21    Maryland   33
## 22 Massachusetts   38
## 23     Michigan   43
## 24    Minnesota   31
## 25    Mississippi   21
## 26     Missouri   34
## 27     Montana    8
## 28     Nebraska   14
## 29     Nevada    17
## 30 New Hampshire   10
## 31    New Jersey   41
## 32    New Mexico   16
## 33     New York   48
## 34 North Carolina   42
## 35    North Dakota    4
## 36      Ohio    45
## 37    Oklahoma    24
```

```
## 38          Oregon 25
## 39    Pennsylvania 46
## 40      Rhode Island 9
## 41    South Carolina 28
## 42      South Dakota 6
## 43        Tennessee 35
## 44          Texas 50
## 45          Utah 18
## 46        Vermont 3
## 47        Virginia 40
## 48      Washington 39
## 49    West Virginia 15
## 50        Wisconsin 32
## 51          Wyoming 1
```

6. data yang telah diurutkan dari populasi yang paling tidak padat hingga ke yang terpadat.

```
# mengurutkan ranks dari terkecil
ranks <- rank(murders$population)
sorted_ranks <- sort(ranks)
index_ranks <- order(ranks)

# mengurutkan states dari yang terkecil
states <- murders$state[index_ranks]

# membuat data frame
my_df2 <- data.frame(state = states, rank = sorted_ranks)
my_df2
```

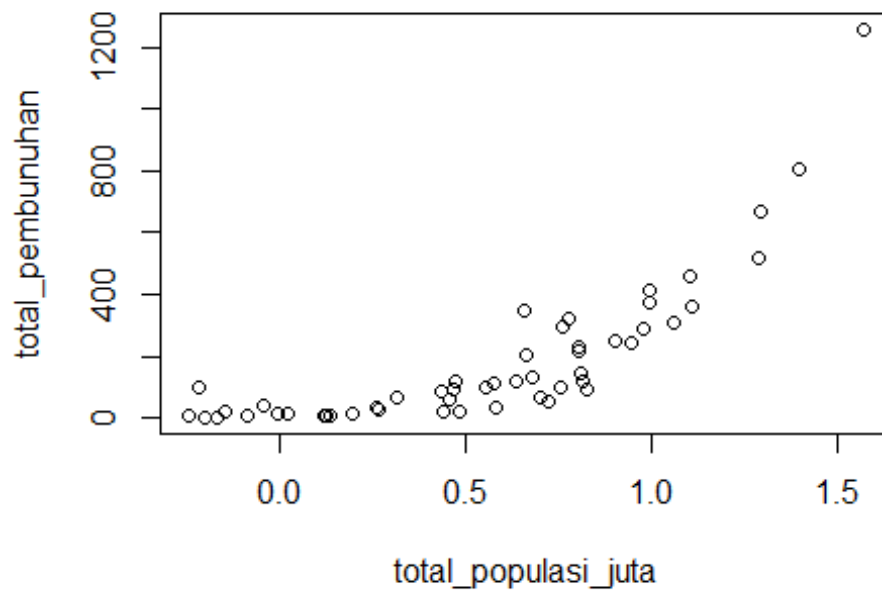
	state	rank
## 1	Wyoming	1
## 2	District of Columbia	2
## 3	Vermont	3
## 4	North Dakota	4
## 5	Alaska	5
## 6	South Dakota	6
## 7	Delaware	7
## 8	Montana	8
## 9	Rhode Island	9
## 10	New Hampshire	10
## 11	Maine	11
## 12	Hawaii	12
## 13	Idaho	13
## 14	Nebraska	14
## 15	West Virginia	15
## 16	New Mexico	16
## 17	Nevada	17

## 18	Utah	18
## 19	Kansas	19
## 20	Arkansas	20
## 21	Mississippi	21
## 22	Iowa	22
## 23	Connecticut	23
## 24	Oklahoma	24
## 25	Oregon	25
## 26	Kentucky	26
## 27	Louisiana	27
## 28	South Carolina	28
## 29	Alabama	29
## 30	Colorado	30
## 31	Minnesota	31
## 32	Wisconsin	32
## 33	Maryland	33
## 34	Missouri	34
## 35	Tennessee	35
## 36	Arizona	36
## 37	Indiana	37
## 38	Massachusetts	38
## 39	Washington	39
## 40	Virginia	40
## 41	New Jersey	41
## 42	North Carolina	42
## 43	Michigan	43
## 44	Georgia	44
## 45	Ohio	45
## 46	Pennsylvania	46
## 47	Illinois	47
## 48	New York	48
## 49	Florida	49
## 50	Texas	50
## 51	California	51

7. plot yang memvisualisasikan total pembunuhan terhadap populasi dan mengidentifikasi hubungan antara keduanya.

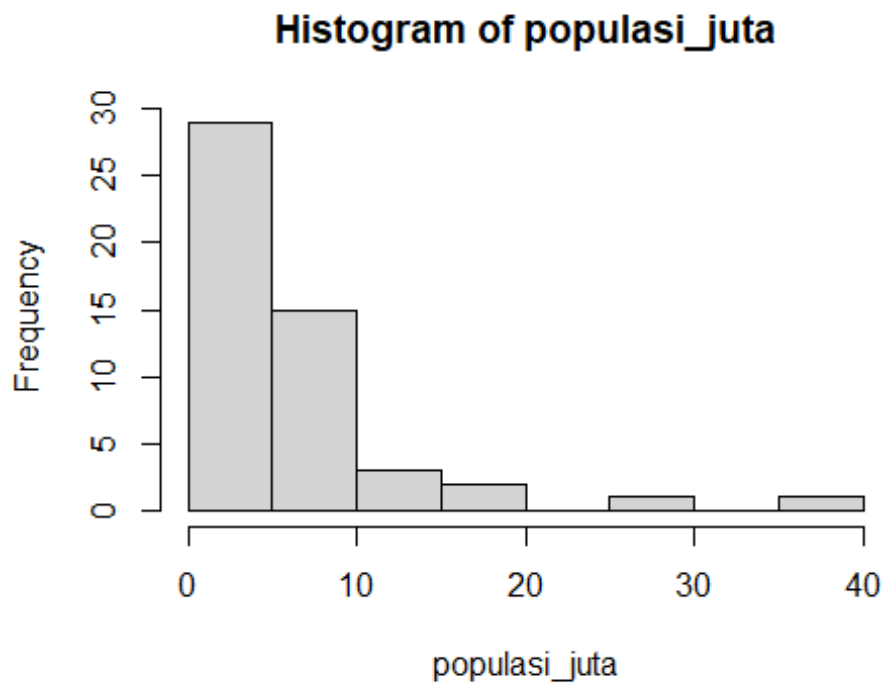
```
total_populasi_juta <- murders$population / 10^6
total_populasi_juta <- log10(total_populasi_juta)
total_pembunuhan <- murders$total

plot(total_populasi_juta, total_pembunuhan)
```



8. Histogram populasi negara bagian

```
populasi_juta <- with(murders, murders$population/10^6)
hist(populasi_juta)
```



9. boxplot dari populasi negara bagian berdasarkan wilayahnya.

```
boxplot((population/10^6)~region, data = murders)
```

