Tugas Modul 4

Rama Tri Agung / 123180053/ Praktikum Data Science B

05/11/2020

```
# import library and get data set
library(dslabs)
data(murders)
```

1. Melakukan `sort` pada population dan menampilkan nilai terkecil

```
# menyimpan dan mengurutkan nilai population ke dalam pop
pop <- sort(murders$population)

# menampilkan nilai terkecil
pop[1]
```

```
## [1] 563626
```

2. Menampilkan index data terkecil dari population

```
# mengambil dan mengurutkan nilai population berdasarkan index
pop_2 <- order(murders$population)

# menampilkan index nilai data terkecil dari population
pop_2[1]
```

```
## [1] 51
```

3. Menampilkan index data terkecil dari population menggunakan which.min

```
which.min(murders$population)
```

```
## [1] 51
```

4. Menaampilkan nama negara yang memiliki populasi terkecil

```
# mengambil index data population terkecil
index_min <- which.min(murders$population)

# menampilkan negara dengan populasi terkecil
murders$state[index_min]
```

```
## [1] "Wyoming"
```

5. buat data frame baru yang berisi nama negara bagian dan peringkatnya

```
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. Urutkan data `rank` pada nomor 5

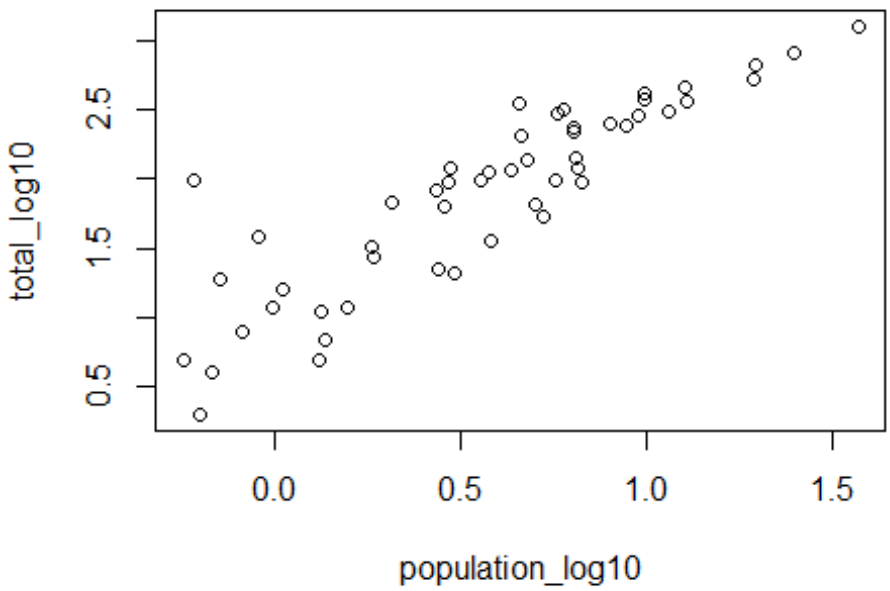
```
ranks <- rank(murders$population)
sort_ranks <- sort(ranks)
index_ranks <- order(ranks)
rank_state <- murders$state[index_ranks]
my_df <- data.frame(state = rank_state, rank = sort_ranks)
my_df
```

```
##                   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
```

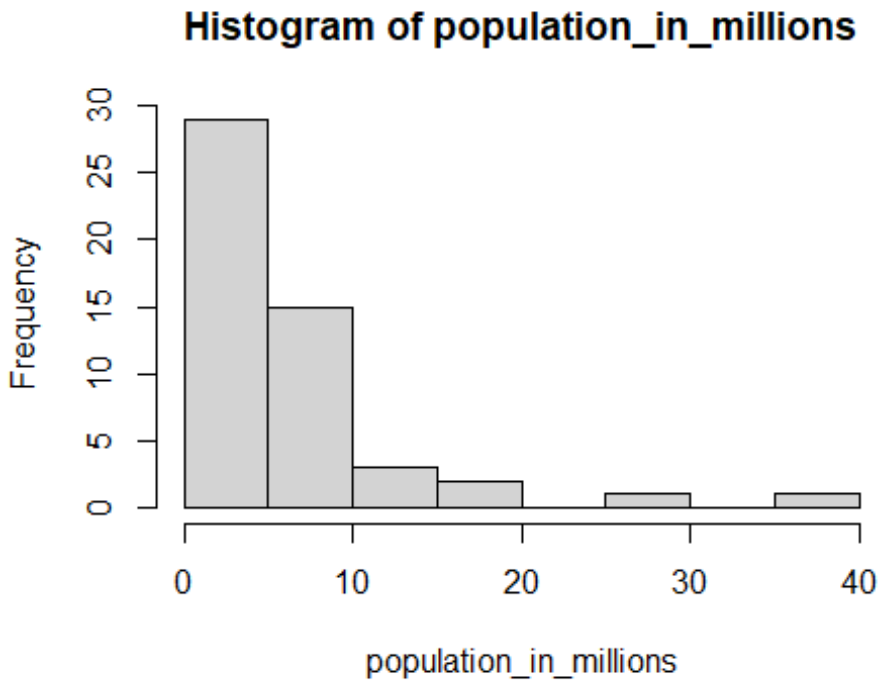
7. Transformasi nilai variabel menggunakan transformasi log10,kemudian tampilkan plot-nya.

```
population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total
population_log10 <- log10(population_in_millions)
total_log10 <- log10(total_gun_murders)
plot(population_log10, total_log10)
```



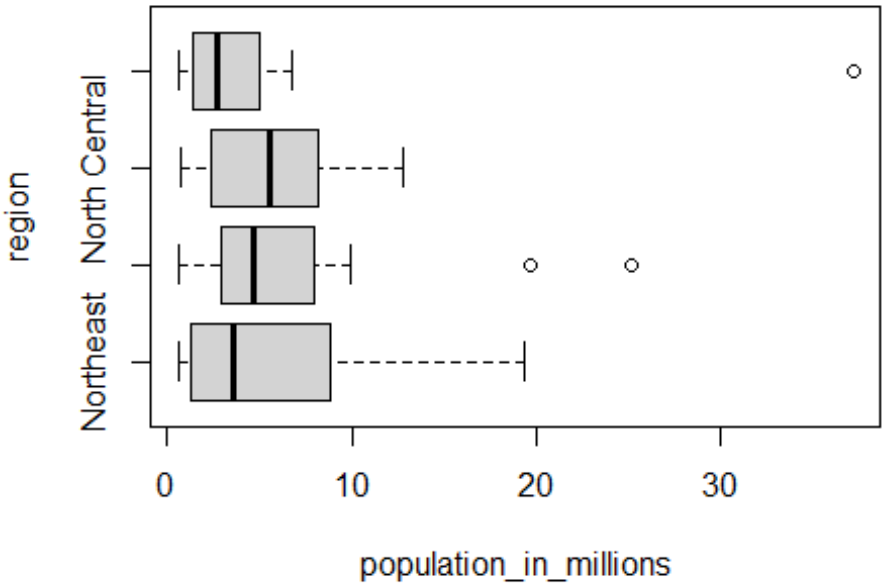
8. Buat histogram dari populasi negara bagian

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



9. Hasilkan boxplot dari populasi negara bagian berdasarkan wilayahnya

```
population_in_millions <- with(murders, population/10^6)
boxplot(population_in_millions~region, data = murders, horizontal = TRUE)
```


region
Northeast
North Central
0
10
20
30
population_in_millions