

Tugas 3 Modul 4

Nathan

10/2/2021

R Markdown Modul 4

```
library(dslabs)
data("murders")
```

Including Plots

1. Gunakan operator aksesori (\$) untuk mengakses variabel populasi dan menyimpannya pada objek baru "pop". Kemudian gunakan fungsi sort untuk mengurutkan variabel "pop". Pada langkah terakhir, gunakan operator () untuk menampilkan nilai populasi terkecil.

```
x<-sort(murders$population)
x[1]
```

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

2. Tampilkan indeks dari data yang memiliki nilai populasi terkecil. Petunjuk: gunakan fungsi order.

```
x<-order(murders$population)
x[1]
```

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

3. Dengan fungsi which.min, Tulis satu baris kode yang dapat menampilkan hasil yang sama dengan langkah diatas.

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

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

4. Tampilkan nama negara yang memiliki populasi terkecil.

```
murders$state[which.min(x)]
```

```
## [1] "Nevada"
```

5. Untuk membuat data frame baru, contoh script yang dapat digunakan adalah sebagai berikut: Gunakan fungsi rank untuk menentukan peringkat populasi dari tiap negara bagian, dimulai dari nilai terkecil hingga terbesar. Simpan hasil pemeringkatan di objek baru “ranks”, lalu buat data frame baru yang berisi nama negara bagian dan peringkatnya dengan nama “my_df”.

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

```
##      rank      state
## 1      51    Alabama
## 2       9    Alaska
## 3      46    Arizona
## 4      35    Arkansas
## 5       2    California
## 6      42    Colorado
## 7       8    Connecticut
## 8      27    Delaware
## 9      40 District of Columbia
## 10     30    Florida
## 11     20    Georgia
## 12     12    Hawaii
## 13     13    Idaho
## 14     28    Illinois
## 15     49    Indiana
## 16     32    Iowa
## 17     29    Kansas
## 18     45    Kentucky
## 19     17    Louisiana
## 20      4    Maine
## 21     25    Maryland
## 22     16    Massachusetts
## 23      7    Michigan
## 24     37    Minnesota
## 25     38    Mississippi
## 26     18    Missouri
## 27     19    Montana
## 28     41    Nebraska
## 29      1    Nevada
## 30      6    New Hampshire
## 31     24    New Jersey
## 32     50    New Mexico
## 33     21    New York
## 34     26    North Carolina
## 35     43    North Dakota
## 36      3    Ohio
## 37     15    Oklahoma
## 38     22    Oregon
## 39     48    Pennsylvania
## 40     47    Rhode Island
## 41     31    South Carolina
## 42     34    South Dakota
## 43     23    Tennessee
## 44     11    Texas
```

```
## 45 36          Utah
## 46 39          Vermont
## 47 14          Virginia
## 48 33          Washington
## 49 10          West Virginia
## 50 44          Wisconsin
## 51 5           Wyoming
```

6. Ulangi langkah sebelumnya, namun kali ini urutkan `my_df` dengan fungsi `order` agar data yang ditampilkan merupakan data yang telah diurutkan dari populasi yang paling tidak padat hingga ke yang terpadat. Petunjuk: buat objek “ind” yang akan menyimpan indeks yang diperlukan dalam mengurutkan data populasi

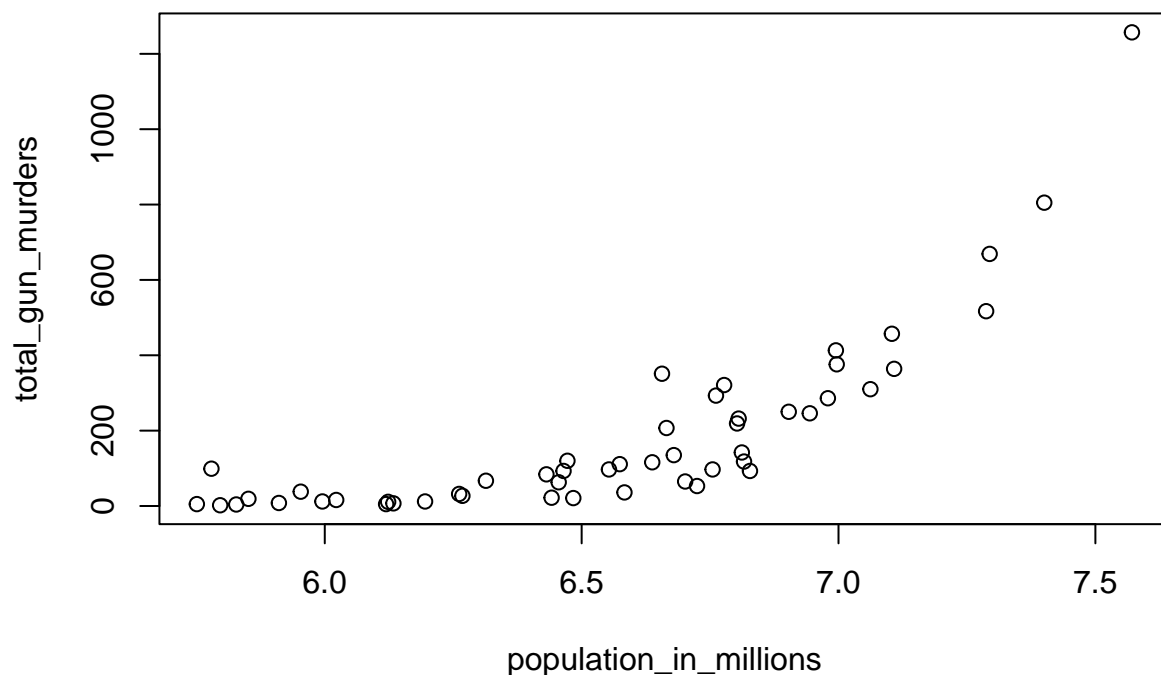
```
ind <- order(x)
my_df <- data.frame(rank = ranks[ind], state = murders$state[ind], index = ind)
my_df
```

```
##      rank      state index
## 1      1      Nevada   29
## 2      2    California    5
## 3      3        Ohio   36
## 4      4        Maine   20
## 5      5        Wyoming  51
## 6      6    New Hampshire  30
## 7      7      Michigan   23
## 8      8    Connecticut    7
## 9      9        Alaska    2
## 10     10    West Virginia  49
## 11     11        Texas   44
## 12     12        Hawaii   12
## 13     13        Idaho   13
## 14     14        Virginia  47
## 15     15      Oklahoma   37
## 16     16    Massachusetts  22
## 17     17      Louisiana   19
## 18     18      Missouri   26
## 19     19      Montana   27
## 20     20      Georgia   11
## 21     21      New York   33
## 22     22      Oregon   38
## 23     23      Tennessee  43
## 24     24    New Jersey   31
## 25     25      Maryland   21
## 26     26    North Carolina  34
## 27     27      Delaware    8
## 28     28      Illinois   14
## 29     29      Kansas   17
## 30     30      Florida   10
## 31     31    South Carolina  41
## 32     32        Iowa   16
## 33     33      Washington  48
## 34     34    South Dakota  42
## 35     35      Arkansas    4
```

```
## 36 36 Utah 45
## 37 37 Minnesota 24
## 38 38 Mississippi 25
## 39 39 Vermont 46
## 40 40 District of Columbia 9
## 41 41 Nebraska 28
## 42 42 Colorado 6
## 43 43 North Dakota 35
## 44 44 Wisconsin 50
## 45 45 Kentucky 18
## 46 46 Arizona 3
## 47 47 Rhode Island 40
## 48 48 Pennsylvania 39
## 49 49 Indiana 15
## 50 50 New Mexico 32
## 51 51 Alabama 1
```

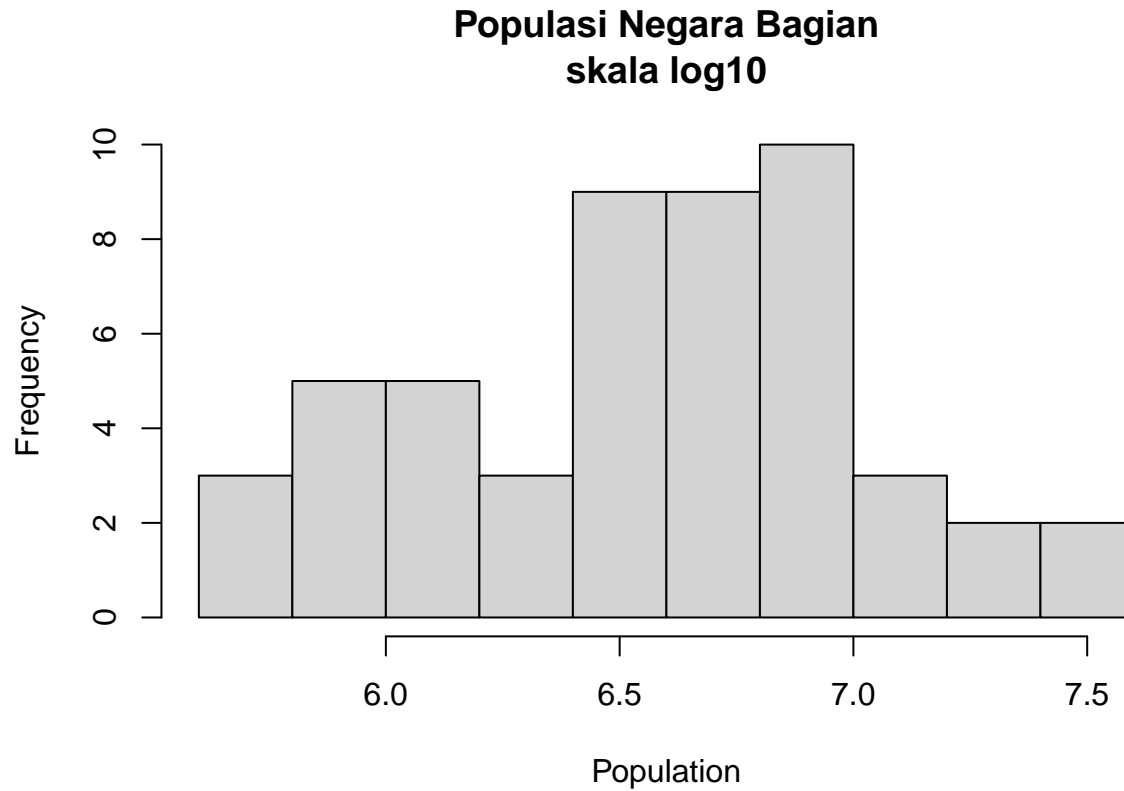
7. Untuk keperluan analisis data, akan dibuat plot yang memvisualisasikan total pembunuhan terhadap populasi dan mengidentifikasi hubungan antara keduanya. Script yang digunakan: Perlu diingat bahwa beberapa negara bagian memiliki populasi di bawah 5 juta, sehingga untuk mempermudah analisis, buat plot dalam skala log. Transformasi nilai variabel menggunakan transformasi log10, kemudian tampilkan plot-nya.

```
population_in_millions <- log10(murders$population)
total_gun_murders <- murders$total
plot(population_in_millions, total_gun_murders)
```



8. Buat histogram dari populasi negara bagian.

```
hist(population_in_millions, main = "Populasi Negara Bagian\nskala log10", xlab = "Population")
```



9. Hasilkan boxplot dari populasi negara bagian berdasarkan wilayahnya.

```
boxplot(population_in_millions~region, data = murders)
```

