

Tugas Modul 4_Praktikum Data Science

Taufik Sahid Fadhil

2022-09-20

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

Soal Nomor 1

Gunakan operator aksesor (\$) 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.

```
pop<-murders$population
sort(pop)
```

```
## [1] 563626 601723 625741 672591 710231 814180 897934 989415
## [9] 1052567 1316470 1328361 1360301 1567582 1826341 1852994 2059179
## [17] 2700551 2763885 2853118 2915918 2967297 3046355 3574097 3751351
## [25] 3831074 4339367 4533372 4625364 4779736 5029196 5303925 5686986
## [33] 5773552 5988927 6346105 6392017 6483802 6547629 6724540 8001024
## [41] 8791894 9535483 9883640 9920000 11536504 12702379 12830632 19378102
## [49] 19687653 25145561 37253956
```

```
pop[51]
```

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

Soal Nomor 2

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

```
max(order(murders$population))
```

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

Soal Nomor 3

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

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

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

Soal Nomor 4

Tampilkan nama negara yang memiliki populasi terkecil

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

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

Soal Nomor 5

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

Soal Nomor 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(my_df$rank)
my_df[ind,]
```

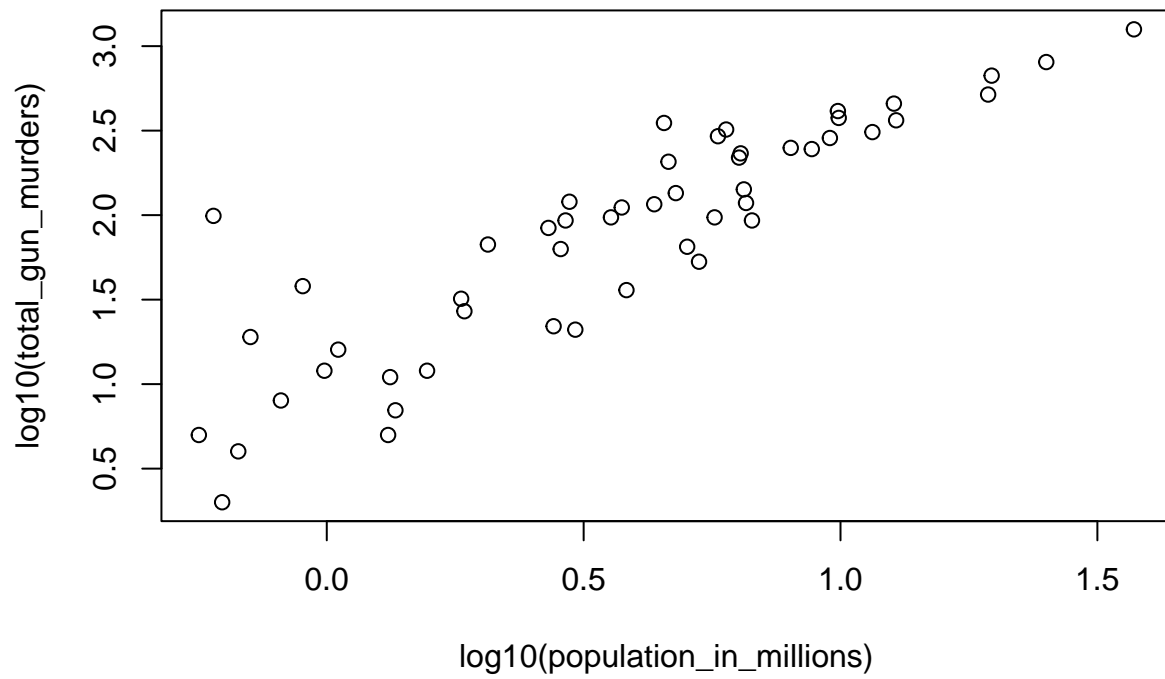
```
##           state rank
## 51      Wyoming    1
## 9 District of Columbia 2
## 46      Vermont    3
## 35    North Dakota    4
## 2       Alaska     5
## 42    South Dakota    6
## 8       Delaware    7
## 27      Montana     8
## 40    Rhode Island    9
## 30    New Hampshire   10
## 20      Maine        11
## 12      Hawaii       12
## 13      Idaho        13
## 28      Nebraska     14
## 49    West Virginia   15
## 32      New Mexico    16
## 29      Nevada       17
```

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

Soal Nomor 7

Untuk keperluan analisis data, akan dibuat plot yang memvisualisasikan total pembunuhan terhadap populasi dan mengidentifikasi hubungan antara keduanya. 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.

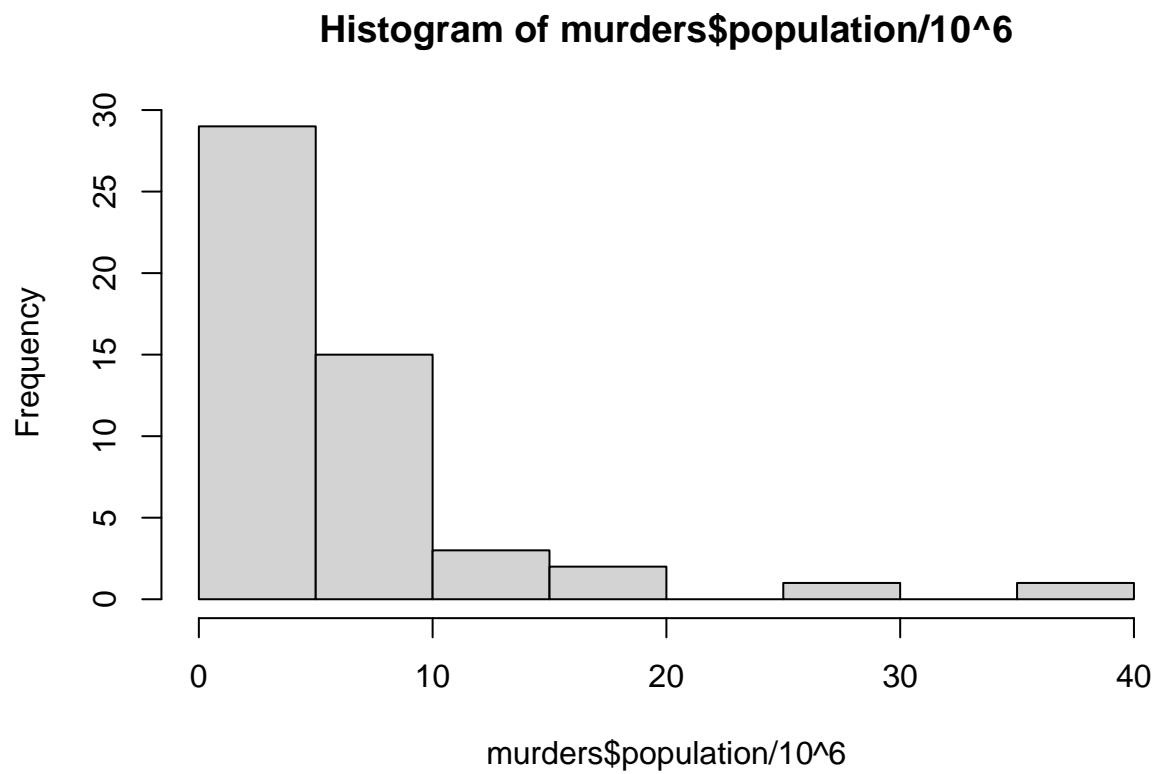
```
population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total
plot(log10(population_in_millions), log10(total_gun_murders))
```



Soal Nomor 8

Buat histogram dari populasi negara bagian.

```
hist(murders$population/ $10^6$ )
```



Soal Nomor 9

Hasilkan boxplot dari populasi negara bagian berdasarkan wilayahnya.

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

