

# Tugas Fungsi Sorting dan Visualisasi Data

Elisia Dwi Rahayu

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
print("Nilai dari objek pop:")
```

```
## [1] "Nilai dari objek pop:"
```

```
pop
```

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

```
opo <- sort(pop)
print("Hasil pengurutan:")
```

```
## [1] "Hasil pengurutan:"
```

```
opo
```

```
## [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
```

```
print("Nilai terkecil:")
```

```
## [1] "Nilai terkecil:"
```

```
opo[1]
```

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

## Soal Nomor 2

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

```
x <- murders$population
print("Populasi:")
```

```
## [1] "Populasi:"
```

```
x
```

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

```
index <- order(x)
print("Hasil pengurutan populasi:")
```

```
## [1] "Hasil pengurutan populasi:"
```

```
x[index]
```

```
## [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
```

```
i <- order(x)
print("Indeks dari hasil pengurutan populasi:")
```

```
## [1] "Indeks dari hasil pengurutan populasi:"
```

```
i

## [1] 51  9 46 35  2 42  8 27 40 30 20 12 13 28 49 32 29 45 17  4 25 16  7 37 38
## [26] 18 19 41  1  6 24 50 21 26 43  3 15 22 48 47 31 34 23 11 36 39 14 33 10 44
## [51]  5
```

```
print("Indeks dengan populasi terkecil:")
```

```
## [1] "Indeks dengan populasi terkecil:"
```

```
i[1]
```

```
## [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.

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

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

### Soal Nomor 5

Untuk membuat data frame baru, contoh script yang dapat digunakan adalah sebagai berikut:

```
temp <- c(35, 88, 42, 84, 81, 30)
city <- c("Beijing", "Lagos", "Paris", "Rio de Janeiro",
"San Juan", "Toronto")
city_temps <- data.frame(name = city, temperature = temp)
```

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)
ranks
```

```
## [1] 29 5 36 20 51 30 23 7 2 49 44 12 13 47 37 22 19 26 27 11 33 38 43 31 21
## [26] 34 8 14 17 10 41 16 48 42 4 45 24 25 46 9 28 6 35 50 18 3 40 39 15 32
## [51] 1
```

```
rk <- c(ranks)
st <- c(murders$state)
my_df <- data.frame(states = st, ranking = rk)
my_df
```

```
##           states ranking
## 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
```

## 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

```
x <- murders$population
y <- murders$state
ind <- order(x)
rk <- c(rank(x)[ind])
st <- c(y[ind])
my_df <- data.frame(states = st, ranking = rk)
my_df
```

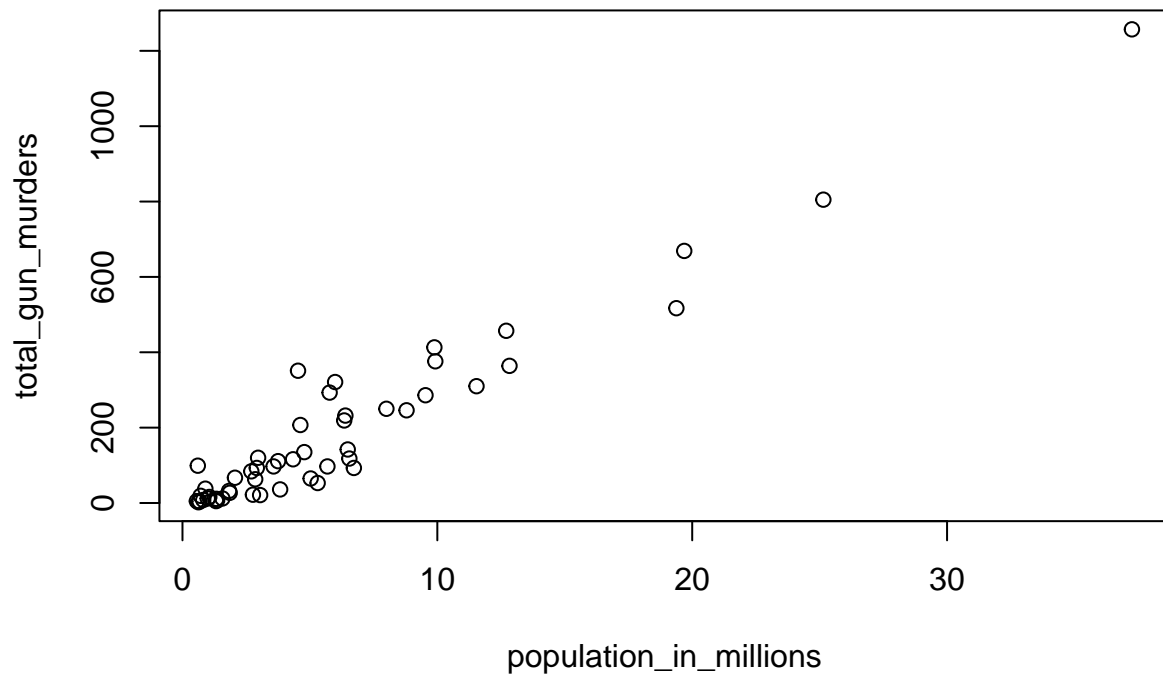
```
##          states ranking
## 1      Wyoming      1
## 2 District of Columbia  2
## 3      Vermont      3
## 4      North Dakota    4
## 5      Alaska          5
## 6      South Dakota    6
## 7      Delaware       7
## 8      Montana        8
## 9      Rhode Island    9
## 10     New Hampshire   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

## Soal Nomor 7

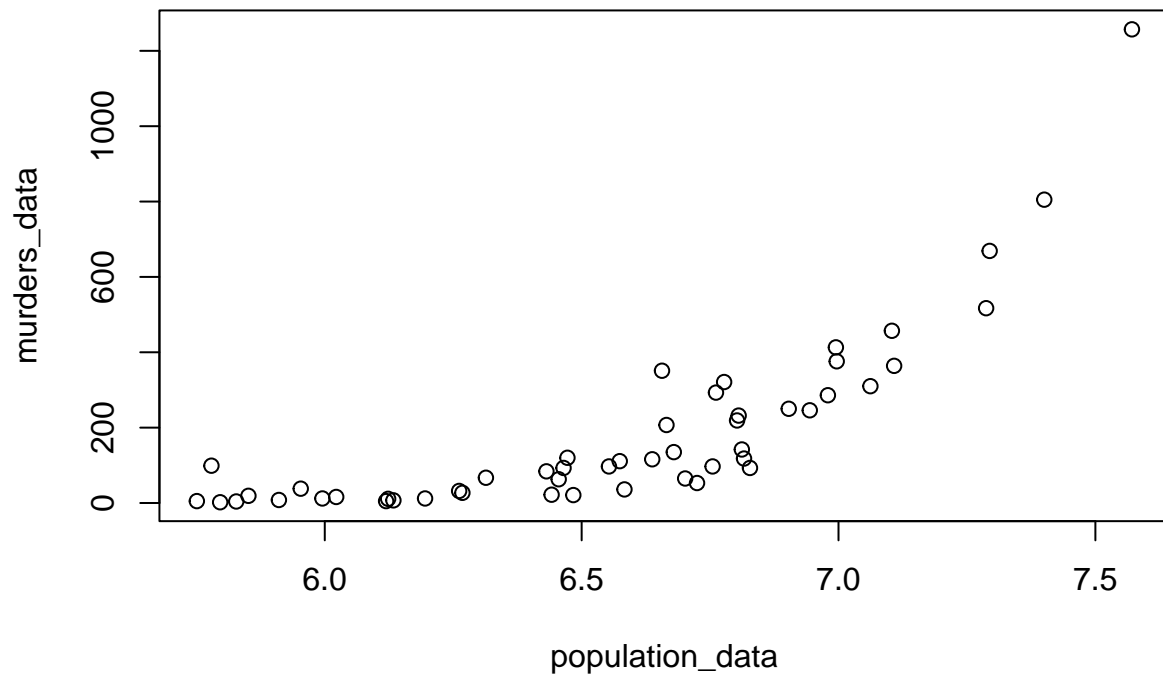
Untuk keperluan analisis data, akan dibuat plot yang memvisualisasikan total pembunuhan terhadap populasi dan mengidentifikasi hubungan antara keduanya. Script yang digunakan:

```
population_in_millions <- murders$population/106
total_gun_murders <- murders$total
plot(population_in_millions, total_gun_murders)
```



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_data <- log10(murders$population)
murders_data <- murders$total
plot <- plot(population_data, murders_data)
```

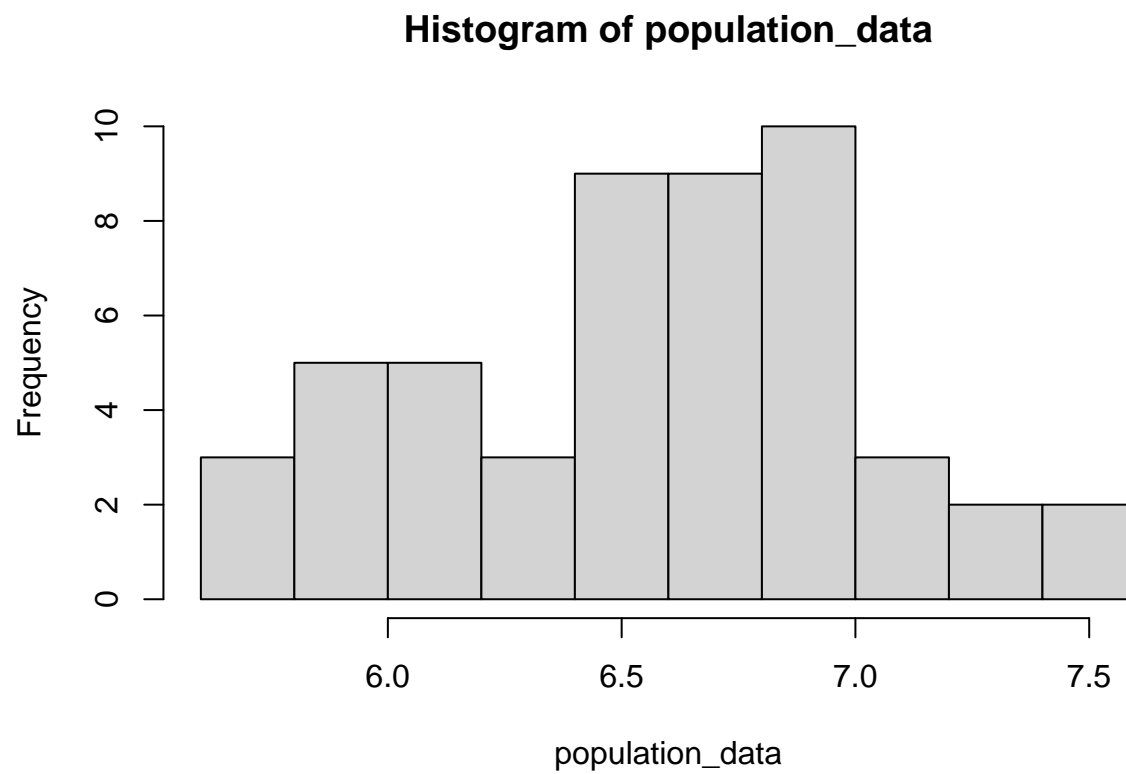


### Soal Nomor 8

Buat histogram dari populasi negara bagian.

```
population_data <- log10(murders$population)
hist(population_data)
```





### Soal Nomor 9

Hasilkan boxplot dari populasi negara bagian berdasarkan wilayahnya.

```
population_data <- log10(murders$population)
boxplot(population_data~region, data = murders)
```

