

Tugas Modul 4

Ahmad Abdullah Azzam

10/6/2021

D.Latihan

Pada latihan ini seluruh soal akan menggunakan *dataset: AS murders*.

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

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.

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

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

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

```
print(order(murders$population)[1])
```

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

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

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

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

4. Tampilkan nama negara yang memiliki populasi terkecil.

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

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

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(city_temps)
my_df<-data.frame(name=city, rank=ranks)
print(my_df)
```

```
##           name rank
## 1      Beijing    7
## 2        Lagos    8
## 3         Paris    9
## 4  Rio de Janeiro 10
## 5       San Juan 11
## 6       Toronto 12
## 7      Beijing    2
## 8        Lagos    6
## 9         Paris    3
## 10  Rio de Janeiro    5
## 11      San Juan    4
## 12      Toronto    1
```

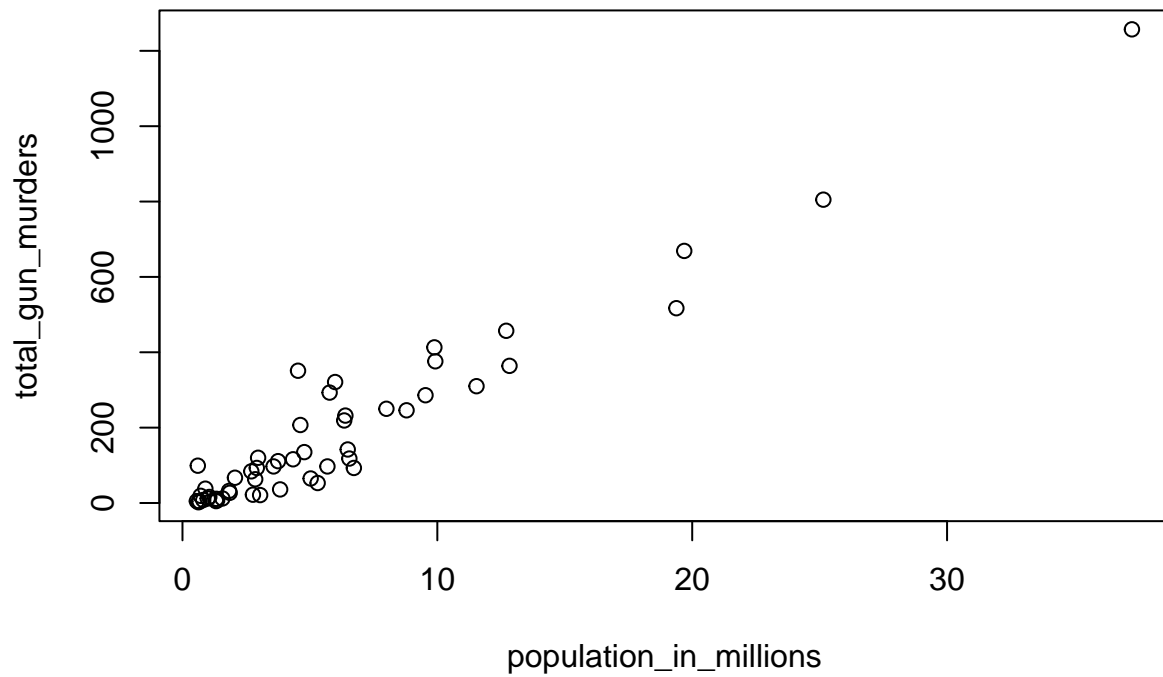
6. Ulangi langkah sebelumnya, namun kali ini urutkan `my_df` dengan fungsi `order` agar data yang ditampilkan merupakan data yang telah diurutkan dari populasi yang paling tidak padat hingga ke yang terpadat. Petunjuk: buat objek “*ind*” yang akan menyimpan indeks yang diperlukan dalam mengurutkan data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

```
ind<-order(city_temps$temperature)
my_df<-data.frame(name=city_temps$name[ind], temperature=city_temps$temperature[ind])
print(my_df)
```

```
##           name temperature
## 1      Toronto          30
## 2      Beijing          35
## 3         Paris          42
## 4       San Juan          81
## 5  Rio de Janeiro          84
## 6        Lagos          88
```

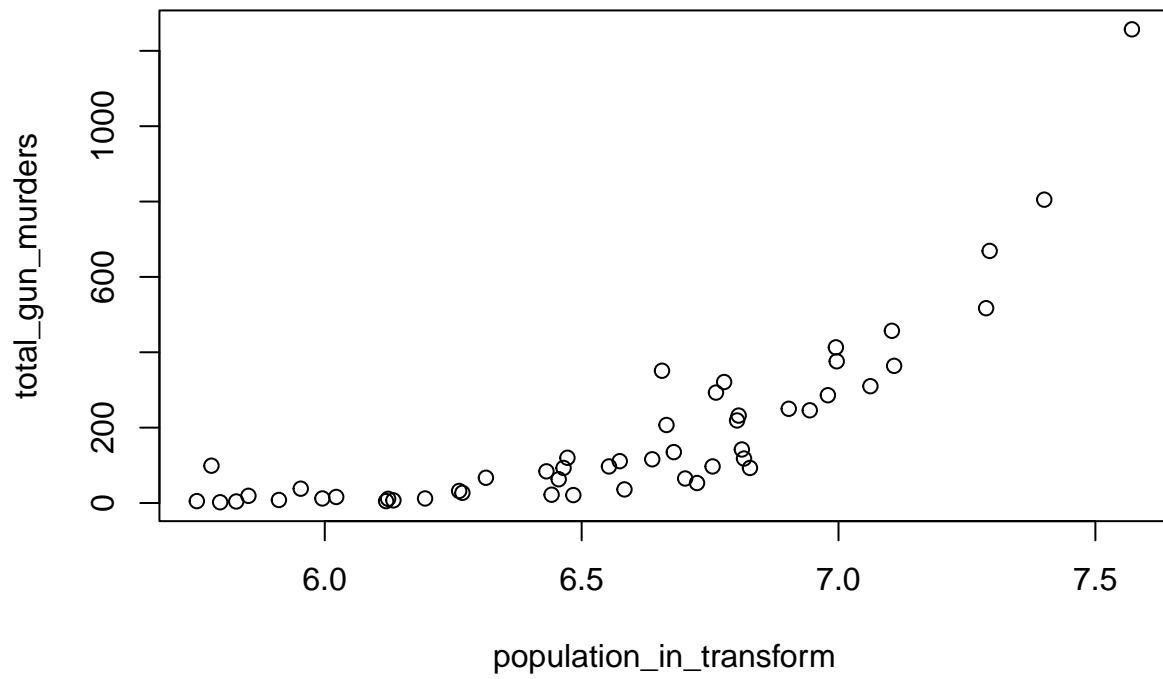
7. Untuk keperluan analisis data, akan dibuat plot yang memvisualisasikan total pembunuhan terhadap populasi dan mengidentifikasi hubungan antara keduanya. *Script* yang digunakan:

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



Perlu diingat bahwa beberapa negara bagian memiliki populasi dibawah 5 juta, sehingga untuk mempermudah analisis, buat plot dalam skala log. Transformasi nilai variabel menggunakan transformasi `log10`, kemudian tampilkan plot-nya.

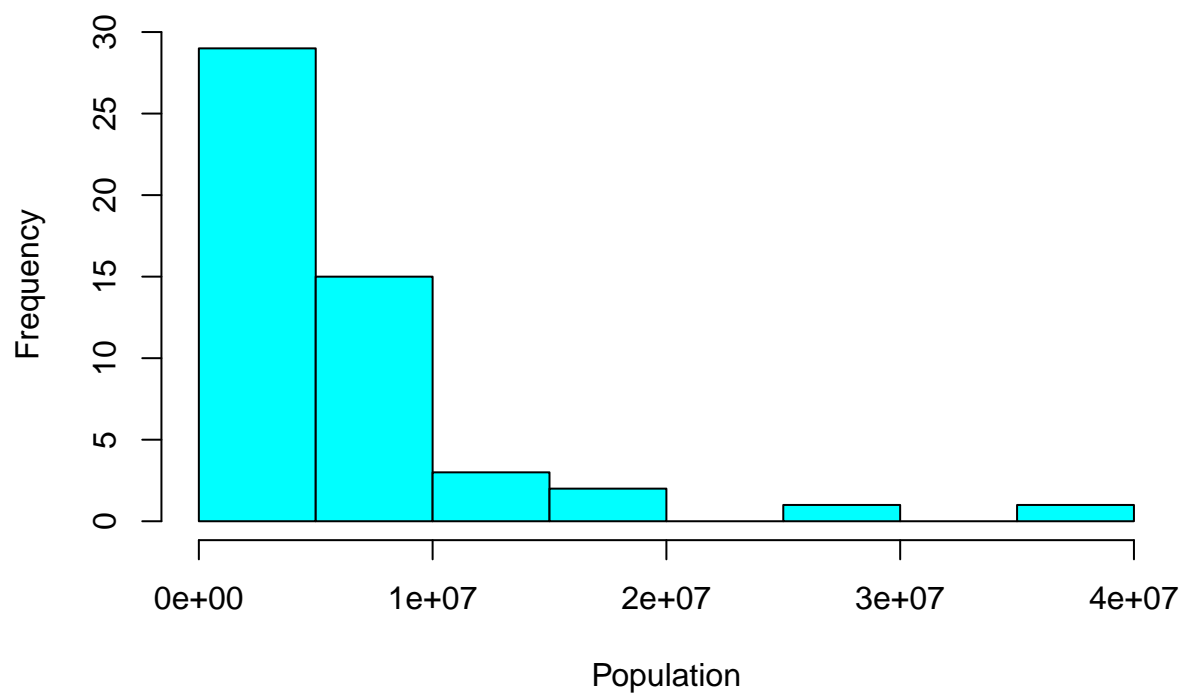
```
population_in_transform<-log10(murders$population)
total_gun_murders<-murders$total
plot(population_in_transform,total_gun_murders)
```



8. Buat histogram dari populasi negara bagian.

```
hist(murders$population, col="cyan", main="Histogram Populasi Negara Bagian", xlab="Population")
```

Histogram Populasi Negara Bagian



9. Hasilkan *boxplot* dari populasi negara bagian berdasarkan wilayahnya.

```
boxplot(population~region, data=murders, col=c("cyan"), cex.main=1.2, main="BOXPLOT Populasi Negara ber
```

BOXPLOT Populasi Negara berdasarkan wilayah

