

## Tugas Modul 3

Duta Alamin / 123190156

9/26/2021

### R Markdown

Pada dokumen R Markdown ini berisi jawaban-jawaban untuk menjawab Tugas Praktikum Data Science Modul 3 Kelas E. Pertama import dataset “murders” yang ada dalam package dslabs dengan cara:

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

#### 1. Menggunakan fungsi str untuk memeriksa struktur objek “murders”

```
str(murders)

## 'data.frame':    51 obs. of  5 variables:
## $ state      : chr  "Alabama" "Alaska" "Arizona" "Arkansas" ...
## $ abb       : chr  "AL" "AK" "AZ" "AR" ...
## $ region    : Factor w/ 4 levels "Northeast","South",...: 2 4 4 2 4 4 1 2
##              2 2 ...
## $ population: num  4779736 710231 6392017 2915918 37253956 ...
## $ total     : num   135  19  232  93 1257 ...
```

Dari struktur tersebut maka yang menggambarkan karakter dari tiap variabel pada data frame adalah

- Terdiri dari 51 negara
- Data berisi Nama negara bagian, singkatan dari nama negara bagian, wilayah negara bagian, dan populasi negara bagian serta jumlah total pembunuhan pada tahun 2010

#### 2. Nama kolom dapat diketahui dengan menggunakan fungsi name, yang dapat digunakan seperti contoh dibawah ini:

```
names(murders)

## [1] "state"      "abb"        "region"     "population" "total"
```

#### 3. Menggunakan Operator aksesor (\$) untuk mengekstrak informasi singkatan negara dan menyimpannya pada objek “a”

```
murders$abb
```

```
## [1] "AL" "AK" "AZ" "AR" "CA" "CO" "CT" "DE" "DC" "FL" "GA" "HI" "ID" "IL"
      "IN"
## [16] "IA" "KS" "KY" "LA" "ME" "MD" "MA" "MI" "MN" "MS" "MO" "MT" "NE" "NV"
      "NH"
## [31] "NJ" "NM" "NY" "NC" "ND" "OH" "OK" "OR" "PA" "RI" "SC" "SD" "TN" "TX"
      "UT"
## [46] "VT" "VA" "WA" "WV" "WI" "WY"
```

Kemudian disimpan pada objek “a”

```
a = murders$abb
```

Untuk mengetahui jenis kelas bisa menggunakan fungsi `class`

```
class(a)
## [1] "character"
```

#### 4. Menggunakan kurung siku untuk mengekstrak singkatan negara dan menyimpan pada objek “b”

```
b = murders[["abb"]]
```

Membandingkan antara variabel a dan b

- variabel a

```
str(a)
## chr [1:51] "AL" "AK" "AZ" "AR" "CA" "CO" "CT" "DE" "DC" "FL" "GA" "HI"
...

```

- variabel b

```
str(b)
## chr [1:51] "AL" "AK" "AZ" "AR" "CA" "CO" "CT" "DE" "DC" "FL" "GA" "HI"
...

```

Dapat disimpulkan isi variabel a dan b bernilai sama

#### 5. Menggunakan fungsi `level` dan `length` untuk menentukan jumlah region yang memiliki *dataset*

```
length(murders$region)
## [1] 51
levels(murders$region)
## [1] "Northeast" "South" "North Central" "West"
```

Dapat dilihat bahwa banyaknya entri di region yang ada pada dataset murders sebanyak 51 dan ada 4 kategori yang ada

## 6. Menggunakan fungsi table untuk menampilkan tabel baru yang berisi jumlah *state* pada setiap *region*

Sebelum membuat tabel yang berisi data *state* dan *region*, data tersebut dimasukkan ke dalam variabel terlebih dahulu

```
state = murders$state  
region = murders$region
```

Kemudian membuat tabel menggunakan fungsi `table`

```
table(state, region)  
  
##               region  
## state Northeast South North Central West  
## Alabama           0     1           0     0  
## Alaska             0     0           0     1  
## Arizona            0     0           0     1  
## Arkansas           0     1           0     0  
## California         0     0           0     1  
## Colorado           0     0           0     1  
## Connecticut        1     0           0     0  
## Delaware           0     1           0     0  
## District of Columbia 0     1           0     0  
## Florida            0     1           0     0  
## Georgia            0     1           0     0  
## Hawaii             0     0           0     1  
## Idaho              0     0           0     1  
## Illinois           0     0           1     0  
## Indiana            0     0           1     0  
## Iowa               0     0           1     0  
## Kansas             0     0           1     0  
## Kentucky           0     1           0     0  
## Louisiana          0     1           0     0  
## Maine              1     0           0     0  
## Maryland           0     1           0     0  
## Massachusetts      1     0           0     0  
## Michigan           0     0           1     0  
## Minnesota          0     0           1     0  
## Mississippi        0     1           0     0  
## Missouri           0     0           1     0  
## Montana            0     0           0     1  
## Nebraska           0     0           1     0  
## Nevada            0     0           0     1  
## New Hampshire      1     0           0     0  
## New Jersey         1     0           0     0  
## New Mexico         0     0           0     1  
## New York           1     0           0     0  
## North Carolina     0     1           0     0  
## North Dakota       0     0           1     0
```

##	Ohio	0	0	1	0
##	Oklahoma	0	1	0	0
##	Oregon	0	0	0	1
##	Pennsylvania	1	0	0	0
##	Rhode Island	1	0	0	0
##	South Carolina	0	1	0	0
##	South Dakota	0	0	1	0
##	Tennessee	0	1	0	0
##	Texas	0	1	0	0
##	Utah	0	0	0	1
##	Vermont	1	0	0	0
##	Virginia	0	1	0	0
##	Washington	0	0	0	1
##	West Virginia	0	1	0	0
##	Wisconsin	0	0	1	0
##	Wyoming	0	0	0	1

Sehingga kita dapat melihat jumlah *state* pada setiap *region*