

Tugas Modul 3

Muhammad Auliya Rasyid

9/28/2021

D. Latihan

```
library(dslabs)
data(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 ...
```

1. Gunakan fungsi `str` untuk memeriksa struktur objek *“murders”*. Manakah dari pernyataan berikut ini yang paling menggambarkan karakter dari tiap variabel pada *data frame*?
 - a. Terdiri dari 51 negara. **(Benar, terlihat dari hasil run structure murders diperlihatkan bahwa dataframe memiliki 51 objek)**
 - b. Data berisi tingkat pembunuhan pada 50 negara bagian dan DC **(Salah)**
 - c. Data berisi Nama negara bagian, singkatan dari nama negara bagian, wilayah negara bagian, dan populasi negara bagian serta jumlah total pembunuhan pada tahun 2010. **(Benar, dilihat dari hasil run fungsi structure bahwa dari dataframe ini terdapat 5 variabel yang didalamnya terdiri dari nama negara bagian, singkatan, wilayahnya, dan populasi dari daerah tersebut. Dimana pada detail dataset pada library dikatakan bahwa data ini bersumber dari US gun murders by state for 2010)**
 - d. str tidak menunjukkan informasi yang relevan. **(benar, karena pada structure, informasi yang ditunjukkan bukanlah informasi mendetail dari isi structure pada sebuah dataframe, namun hanya untuk memberitahu berapa banyak objek dan juga variabel yang terdapat pada sebuah dataframe)**
2. Sebutkan apa saja nama kolom yang digunakan pada *data frame*!

```
names(murders)
```

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

3. Gunakan operator aksesor (\$) untuk mengekstrak informasi singkatan negara dan menyimpannya pada objek *“a”*. Sebutkan jenis `class` dari objek tersebut.

```
a = murders$abb
a
```

```
## [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"
```

```
class(a)
```

```
## [1] "character"
```

4. Gunakan tanda kurung siku untuk mengekstrak singkatan negara dan menyimpannya pada objek “*b*”. Tentukan apakah variabel “*a*” dan “*b*” bernilai sama?

```
b = murders[[2]]
b
```

```
## [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"
```

var a dan b bernilai sama

5. Variabel *region* memiliki tipe data: *factor*. Dengan satu baris kode, gunakan fungsi `level` dan `length` untuk menentukan jumlah *region* yang dimiliki *dataset*.

```
c = levels(murders$region)
length(c)
```

```
## [1] 4
```

Pada dataset jumlah jenis region yang dimiliki hanyalah 4 jenis region saja

```
factor(murders$region)
```

```
## [1] South      West      West      South      West
## [6] West      Northeast South      South      South
## [11] South      West      West      North Central North Central
## [16] North Central North Central South      South      Northeast
## [21] South      Northeast North Central North Central South
## [26] North Central West      North Central West      Northeast
## [31] Northeast West      Northeast South      North Central
## [36] North Central South      West      Northeast Northeast
## [41] South      North Central South      South      West
## [46] Northeast South      West      South      North Central
## [51] West
## Levels: Northeast South North Central West
```

Namun pada tiap daerah bagian (51 negara bagian) memiliki region-nya masing-masing

6. Fungsi `table` dapat digunakan untuk ekstraksi data pada tipe vektor dan menampilkan frekuensi dari setiap elemen. Dengan menerapkan fungsi tersebut, dapat diketahui jumlah *state* pada tiap *region*. Gunakan fungsi `table` dalam satu baris kode untuk menampilkan tabel baru yang berisi jumlah state pada tiap *region*.

```
table(state.region)
```

```
## state.region
## Northeast      South North Central      West
##           9          16          12          13
```