

Tugas Modul 3

Rama Tri Agung / 123180053 / Praktikum Data Science B

02/11/2020

```
# import library and get data set
library(dslabs)
data(murders)
```

1. Fungsi str

a. Terdiri dari 51 negara.

```
length(murders$state)
```

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

b. Data berisi tingkat pembunuhan pada 50 negara bagian dan DC.

```
murders
```

| ## | state | abb | region | population | total |
|-------|----------------------|-----|---------------|------------|-------|
| ## 1 | Alabama | AL | South | 4779736 | 135 |
| ## 2 | Alaska | AK | West | 710231 | 19 |
| ## 3 | Arizona | AZ | West | 6392017 | 232 |
| ## 4 | Arkansas | AR | South | 2915918 | 93 |
| ## 5 | California | CA | West | 37253956 | 1257 |
| ## 6 | Colorado | CO | West | 5029196 | 65 |
| ## 7 | Connecticut | CT | Northeast | 3574097 | 97 |
| ## 8 | Delaware | DE | South | 897934 | 38 |
| ## 9 | District of Columbia | DC | South | 601723 | 99 |
| ## 10 | Florida | FL | South | 19687653 | 669 |
| ## 11 | Georgia | GA | South | 9920000 | 376 |
| ## 12 | Hawaii | HI | West | 1360301 | 7 |
| ## 13 | Idaho | ID | West | 1567582 | 12 |
| ## 14 | Illinois | IL | North Central | 12830632 | 364 |
| ## 15 | Indiana | IN | North Central | 6483802 | 142 |
| ## 16 | Iowa | IA | North Central | 3046355 | 21 |
| ## 17 | Kansas | KS | North Central | 2853118 | 63 |
| ## 18 | Kentucky | KY | South | 4339367 | 116 |
| ## 19 | Louisiana | LA | South | 4533372 | 351 |
| ## 20 | Maine | ME | Northeast | 1328361 | 11 |
| ## 21 | Maryland | MD | South | 5773552 | 293 |
| ## 22 | Massachusetts | MA | Northeast | 6547629 | 118 |
| ## 23 | Michigan | MI | North Central | 9883640 | 413 |
| ## 24 | Minnesota | MN | North Central | 5303925 | 53 |

| | | | | | |
|-------|----------------|----|---------------|----------|-----|
| ## 25 | Mississippi | MS | South | 2967297 | 120 |
| ## 26 | Missouri | MO | North Central | 5988927 | 321 |
| ## 27 | Montana | MT | West | 989415 | 12 |
| ## 28 | Nebraska | NE | North Central | 1826341 | 32 |
| ## 29 | Nevada | NV | West | 2700551 | 84 |
| ## 30 | New Hampshire | NH | Northeast | 1316470 | 5 |
| ## 31 | New Jersey | NJ | Northeast | 8791894 | 246 |
| ## 32 | New Mexico | NM | West | 2059179 | 67 |
| ## 33 | New York | NY | Northeast | 19378102 | 517 |
| ## 34 | North Carolina | NC | South | 9535483 | 286 |
| ## 35 | North Dakota | ND | North Central | 672591 | 4 |
| ## 36 | Ohio | OH | North Central | 11536504 | 310 |
| ## 37 | Oklahoma | OK | South | 3751351 | 111 |
| ## 38 | Oregon | OR | West | 3831074 | 36 |
| ## 39 | Pennsylvania | PA | Northeast | 12702379 | 457 |
| ## 40 | Rhode Island | RI | Northeast | 1052567 | 16 |
| ## 41 | South Carolina | SC | South | 4625364 | 207 |
| ## 42 | South Dakota | SD | North Central | 814180 | 8 |
| ## 43 | Tennessee | TN | South | 6346105 | 219 |
| ## 44 | Texas | TX | South | 25145561 | 805 |
| ## 45 | Utah | UT | West | 2763885 | 22 |
| ## 46 | Vermont | VT | Northeast | 625741 | 2 |
| ## 47 | Virginia | VA | South | 8001024 | 250 |
| ## 48 | Washington | WA | West | 6724540 | 93 |
| ## 49 | West Virginia | WV | South | 1852994 | 27 |
| ## 50 | Wisconsin | WI | North Central | 5686986 | 97 |
| ## 51 | Wyoming | WY | West | 563626 | 5 |

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.

```
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
## $ population: num  4779736 710231 6392017 2915918 37253956 ...
## $ total     : num  135 19 232 93 1257 ...
```

d. str tidak menunjukkan informasi yang relevan.

ternyata `str` menunjukkkan informasi yang relevan

`str` menunjukkan struktur data

```
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
## $ population: num  4779736 710231 6392017 2915918 37253956 ...
## $ total     : num  135 19 232 93 1257 ...
```

```
## $ population: num 4779736 710231 6392017 2915918 37253956 ...
## $ total      : num 135 19 232 93 1257 ...
```

2. Nama kolom yang digunakan pada data frame

```
# state, abb, region, population, dan total
names(murders)
```

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

3. Sebutkan jenis class dari abb

```
# ekstrak informasi singkatan negara dan menyimpannya pada objek "a"
menggunakan operator aksesor
```

```
a <- murders$abb
class(a)
```

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

4. Tentukan apakah variabel "a" dan "b" bernilai sama

```
# ekstrak singkatan negara dan menyimpannya pada objek "b" menggunakan
tanda kurung siku
```

```
b <- murders[[2]]
class(b)
```

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

```
# cek kesamaan a dan b
```

```
class(a) == class(b)
```

```
## [1] TRUE
```

5. Menentukan jumlah region yang dimiliki dataset

```
level_region <- levels(murders[[3]])
length(level_region)
```

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

6. Menampilkan tabel baru yang berisi jumlah state pada tiap region

```
table(murders[[3]])
```

```
##
##      Northeast      South North Central      West
##           9          17          12          13
```