

# Tugas Praktikum DataScience 1 Modul 3

Muhammad Ikhwan Hanif/123200096

2022-09-14

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

## Soal Nomer 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. b. Data berisi tingkat pembunuhan pada 50 negara bagian dan DC. 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. d. str tidak menunjukkan informasi yang relevan.

```
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 ...
```

jawab = c

## Soal Nomor 2

Sebutkan apa saja nama kolom yang digunakan pada data frame jawab = state, abb, region, population, dan total

```
names(murders)
```

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

## Soal Nomor 3

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

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

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

```
show(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"
```

## Soal Nomor 4

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

```
b <- murders[["abb"]]  
show(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"
```

```
show(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"
```

## Soal Nomor 5

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

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

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

## Soal Nomor 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(murders$region)
```

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