Tugas Modul 3

Riris

9/23/2021

## R Markdown

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

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 ...

## Manakah dari pernyataan berikut ini yang paling menggambarkan karakter dari tiap variabel pada data frame?

1. Terdiri dari 51 negara.

length(murders$state)

## [1] 51

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

murders$state

## [1] "Alabama" "Alaska" "Arizona"   
## [4] "Arkansas" "California" "Colorado"   
## [7] "Connecticut" "Delaware" "District of Columbia"  
## [10] "Florida" "Georgia" "Hawaii"   
## [13] "Idaho" "Illinois" "Indiana"   
## [16] "Iowa" "Kansas" "Kentucky"   
## [19] "Louisiana" "Maine" "Maryland"   
## [22] "Massachusetts" "Michigan" "Minnesota"   
## [25] "Mississippi" "Missouri" "Montana"   
## [28] "Nebraska" "Nevada" "New Hampshire"   
## [31] "New Jersey" "New Mexico" "New York"   
## [34] "North Carolina" "North Dakota" "Ohio"   
## [37] "Oklahoma" "Oregon" "Pennsylvania"   
## [40] "Rhode Island" "South Carolina" "South Dakota"   
## [43] "Tennessee" "Texas" "Utah"   
## [46] "Vermont" "Virginia" "Washington"   
## [49] "West Virginia" "Wisconsin" "Wyoming"

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.

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

## 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  
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]]  
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"

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"

*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.

length(levels(murders$region))

## [1] 4

## 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