

Tugas 3 - Praktikum Data Science

Muhammad Rizal / 123170036

13/2/2020

No 1

A.

```
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"
```

```
### false.
```

```
### bukan negara, melainkan negara bagian. kurang tepat.
```

B

```
### false
```

```
### di data.frame tersebut tidak ada variabel yang berisi tingkat pembunuhan
```

C

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
```

```
### true  
### pilihan jawaban ini paling menggambarkan dari objek data.frame murders
```

D

```
?str  
### false  
### str digunakan untuk menampilkan struktur dari objek r tersebut.
```

Jawab:

Jadi yang paling menggambarkan karakter dari data.frame adalah **C** karena dari pilihan c terdapat semua karakter yang menggambarkan data.frame tersebut.

No 2

```
names(murders)
```

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

Jawab

state, abbb, region, population, total

No 3

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

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

Jawab

class dari objek tersebut adalah chr atau character

No 4

```
a <- murders['abb']
b <- murders$abb
isTrue <- a == b
print(isTrue)
```

```
##      abb
## [1,] TRUE
## [2,] TRUE
## [3,] TRUE
## [4,] TRUE
## [5,] TRUE
## [6,] TRUE
## [7,] TRUE
## [8,] TRUE
## [9,] TRUE
## [10,] TRUE
## [11,] TRUE
## [12,] TRUE
## [13,] TRUE
## [14,] TRUE
## [15,] TRUE
## [16,] TRUE
## [17,] TRUE
## [18,] TRUE
## [19,] TRUE
## [20,] TRUE
## [21,] TRUE
## [22,] TRUE
## [23,] TRUE
## [24,] TRUE
## [25,] TRUE
## [26,] TRUE
## [27,] TRUE
## [28,] TRUE
## [29,] TRUE
## [30,] TRUE
## [31,] TRUE
## [32,] TRUE
## [33,] TRUE
## [34,] TRUE
## [35,] TRUE
## [36,] TRUE
## [37,] TRUE
## [38,] TRUE
## [39,] TRUE
## [40,] TRUE
## [41,] TRUE
## [42,] TRUE
## [43,] TRUE
## [44,] TRUE
## [45,] TRUE
## [46,] TRUE
## [47,] TRUE
```

```
## [48,] TRUE
## [49,] TRUE
## [50,] TRUE
## [51,] TRUE
```

Jawab

Kedua objek tersebut bernilai sama **TRUE** karena bertipe data sama serta memiliki konten yang sama.

No 5

```
# library(plyr)
# count(murders$region) # menghasilkan frekuensi dari setiap elemen

class(murders$region)
```

```
## [1] "factor"
```

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

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

Jawab

Untuk menghitung jumlah total dari **levels** yang ada di data.frame **murders** pada variabel factor **region** bisa menggunakan **levels** kemudian dibungkus dengan **length**, **levels** akan menampilkan elemen - elemen yang ada di tipe data factor secara unik, jadi jika elemen - elemen tersebut berjumlah lebih dari satu elemen dari elemen yang sama maka yang diambil hanya elemen yang **berbeda** saja.

No 6

```
table(murders$region, murders$state) # menghasilkan frekuensi dari setiap elemen
```

```
##
##      Alabama Alaska Arizona Arkansas California Colorado Connecticut
## Northeast      0      0      0      0      0      0      1
## South          1      0      0      1      0      0      0
## North Central  0      0      0      0      0      0      0
## West           0      1      1      0      1      1      0
##
##      Delaware District of Columbia Florida Georgia Hawaii Idaho
## Northeast      0      0      0      0      0      0
## South          1      1      1      1      0      0
## North Central  0      0      0      0      0      0
## West           0      0      0      0      1      1
```

```
##
##      Illinois Indiana Iowa Kansas Kentucky Louisiana Maine Maryland
## Northeast      0      0      0      0      0      0      1      0
## South          0      0      0      0      1      1      0      1
## North Central  1      1      1      1      0      0      0      0
## West           0      0      0      0      0      0      0      0
##
##      Massachusetts Michigan Minnesota Mississippi Missouri Montana
## Northeast      1      0      0      0      0      0      0
## South          0      0      0      1      0      0      0
## North Central  0      1      1      0      1      1      0
## West           0      0      0      0      0      0      1
##
##      Nebraska Nevada New Hampshire New Jersey New Mexico New York
## Northeast      0      0      1      1      0      1
## South          0      0      0      0      0      0
## North Central  1      0      0      0      0      0
## West           0      1      0      0      1      0
##
##      North Carolina North Dakota Ohio Oklahoma Oregon Pennsylvania
## Northeast      0      0      0      0      0      1
## South          1      0      0      1      0      0
## North Central  0      1      1      0      0      0
## West           0      0      0      0      1      0
##
##      Rhode Island South Carolina South Dakota Tennessee Texas Utah
## Northeast      1      0      0      0      0      0
## South          0      1      0      1      1      0
## North Central  0      0      1      0      0      0
## West           0      0      0      0      0      1
##
##      Vermont Virginia Washington West Virginia Wisconsin Wyoming
## Northeast      1      0      0      0      0      0
## South          0      1      0      1      0      0
## North Central  0      0      0      0      1      0
## West           0      0      1      0      0      1
```

Jawab

Dengan menggunakan table kita bisa menampilkan frekuensi dari setiap elemen dari tipe data factor.