Latihan 6

Alifia Rahma De Anissa

11/10/2021

1

Gunakan as_tibble untuk mengkonversi tabel dataset “US murders” dalam bentuk tibble dan simpan dalam objek baru bernama ‘murders_tibble’.

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

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
murders_tibble <- as_tibble(murders)
murders_tibble
```

```
## # A tibble: 51 x 5
##    state                abb   region    population total
##    <chr>                <chr> <fct>          <dbl> <dbl>
##  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
## # ... with 41 more rows
```

2

Gunakan fungsi group_by untuk mengkonversi dataset "US murders" menjadi sebuah tibble yang dikelompokkan berdasarkan 'region'.

```
murders %>% group_by(region)
```

```
## # A tibble: 51 x 5
## # Groups:   region [4]
##    state                abb   region    population total
##    <chr>                <chr> <fct>          <dbl> <dbl>
##  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
## # ... with 41 more rows
```

3

Tulis script tidyverse yang menghasilkan output yang sama dengan perintah berikut:

```
exp(mean(log(murders$population)))
```

```
## [1] 3675209
```

```
murders %>% pull(population) %>% log %>% mean %>% exp
```

```
## [1] 3675209
```

4

Gunakan map_df untuk membuat data frame yang terdiri dari tiga kolom: 'n', 's_n', dan 's_n_2'. Kolom pertama harus berisi angka 1 hingga 100. Kolom kedua dan ketiga masing-masing harus berisi penjumlahan 1 hingga n, dimana n menyatakan jumlah baris.

```
library(purrr)

compute_s_n <- function(n){
 x <- 1:n
 tibble(n, s_n = sum(x), s_n_2 = sum(x))
}
n <- 1:100
s_n <- map_df(n, compute_s_n)
s_n
```

```
## # A tibble: 100 x 3
##        n   s_n s_n_2
##    <int> <int> <int>
##  1     1     1     1
##  2     2     3     3
##  3     3     6     6
##  4     4    10    10
##  5     5    15    15
##  6     6    21    21
##  7     7    28    28
##  8     8    36    36
##  9     9    45    45
## 10    10    55    55
## # ... with 90 more rows
```