

TugasModul7

Talitha Fawwaz

2022-10-30

Import dataset “murders”;

```
library(dslabs)
library(tidyverse)

## — Attaching packages — tidyverse
1.3.2 —
## ✓ ggplot2 3.3.6      ✓ purrr 0.3.4
## ✓ tibble 3.1.8       ✓ dplyr 1.0.10
## ✓ tidyr 1.2.0        ✓ stringr 1.4.1
## ✓ readr 2.1.2        ✓ forcats 0.5.2
## — Conflicts —
tidyverse_conflicts() —
## + dplyr::filter() masks stats::filter()
## + dplyr::lag()     masks stats::lag()

data(murders)
```

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

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

## [1] "tbl_df"      "tbl"        "data.frame"
```

2. Gunakan fungsi `group_by` untuk mengkonversi dataset “US murders” menjadi sebuah tibble yang dikelompokkan berdasarkan ‘region’. Jawab:

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

## # A tibble: 51 × 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
```

Gunakan operator pipe sehingga setiap fungsi dapat dipanggil tanpa menambahkan argumen. Gunakan dot operator untuk mengakses populasi. Jawab:

```
murders$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. Jawab:

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

## Warning: `data_frame()` was deprecated in tibble 1.1.0.
## Please use `tibble()` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## # A tibble: 100 × 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
```