Tugas Modul 6

Liek Allyandaru / 123180054 / Praktikum Data Science B

12/3/2020

1. Meambahkan kolom baru dengan nama 'rate'

```
# setup library
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library(dslabs)
data(murders)
murders <- mutate(murders, rate = total / population * 100000)

head(murders)

##        state abb region population total     rate
## 1    Alabama  AL  South    4779736   135 2.824424
## 2     Alaska  AK   West     710231    19 2.675186
## 3    Arizona  AZ   West    6392017   232 3.629527
## 4   Arkansas  AR  South    2915918    93 3.189390
## 5 California  CA   West   37253956  1257 3.374138
## 6   Colorado  CO   West    5029196    65 1.292453
```

2. Fungsi mutate untuk menambahkan kolom baru yang berisi hasil pemeringkatan dari nilai tingkat pembunuhan tertinggi ke terendah

```
# menambahkan kolom rate
murders <- mutate(murders, rank = rank(desc(murders$total)))

head(murders)

##        state abb region population total     rate rank
## 1    Alabama  AL  South    4779736   135 2.824424 20.0
```

```
## 2     Alaska  AK   West     710231    19 2.675186 41.0
## 3    Arizona  AZ   West    6392017   232 3.629527 16.0
## 4   Arkansas  AR  South    2915918    93 3.189390 28.5
## 5 California  CA   West   37253956  1257 3.374138  1.0
## 6   Colorado  CO   West    5029196    65 1.292453 32.0
```

3. Menampilkan nama negara (state) dan singkatan (abb)

```
# menampilkan kolom state dan abb
select(murders, state, abb) %>% head()
```

```
##        state abb
## 1    Alabama  AL
## 2     Alaska  AK
## 3    Arizona  AZ
## 4   Arkansas  AR
## 5 California  CA
## 6   Colorado  CO
```

4. Filter untuk menampilkan 5 negara bagian teratas dengan tingkat pembunuhan tertinggi

```
# filter 5 negara pembunuhan tertinggi
filter(murders, rank <=5)
```

```
##          state abb    region population total     rate rank
## 1   California  CA      West   37253956  1257 3.374138    1
## 2      Florida  FL     South   19687653   669 3.398069    3
## 3     New York  NY Northeast   19378102   517 2.667960    4
## 4 Pennsylvania  PA Northeast   12702379   457 3.597751    5
## 5        Texas  TX     South   25145561   805 3.201360    2
```

5. Filter untuk hanya menampilkan hasil yang terdiri dari: state, rate, dan peringkatnya

```
# mengambil data state, rate, region, dan rank
table1 <- select(murders, state, rate, region, rank)

# mengambil data dengan rate kurang dari 1 dan regiom=n Northeast atau West
table1 <- filter(table1, rate < 1, region == "Northeast" | region == "West")

# menampilkan kolom state, rate, dan rank
select(table1, state, rate, rank) %>% head()
```

```
##          state      rate rank
## 1       Hawaii 0.5145920 47.0
```

```
## 2         Idaho 0.7655102 43.5
## 3         Maine 0.8280881 45.0
## 4 New Hampshire 0.3798036 48.5
## 5        Oregon 0.9396843 36.0
## 6          Utah 0.7959810 39.0
```

Operator Pipe

1. Operator pipe untuk membuat data frame baru dengan nama 'my_states'

```
data(murders)

# menambahkan rate, tingkat(rank), kemudian filter, dan select
my_states <- murders %>%
  mutate(rate = total / population * 100000) %>%
  mutate(rank = rank(desc(total))) %>%
  filter(rate < 1, region == "Northeast" | region == "West") %>%
  select(state, rank, rate)

my_states
```

```
##           state rank      rate
## 1        Hawaii 47.0 0.5145920
## 2         Idaho 43.5 0.7655102
## 3         Maine 45.0 0.8280881
## 4 New Hampshire 48.5 0.3798036
## 5        Oregon 36.0 0.9396843
## 6          Utah 39.0 0.7959810
## 7       Vermont 51.0 0.3196211
## 8       Wyoming 48.5 0.8871131
```