Tugas Pertemuan 11: Manipulasi Data

Sahda Huwaidah Estiningtyas

30 Mei 2024

**Soal 7-20** menggunakan data msleep.csv.

> dat <- read.csv(file="C:/data/msleep.csv")  
> dat <- as\_tibble(dat)  
> glimpse(dat)  
Observations: 83  
Variables: 11  
$ name <fct> Cheetah, Owl monkey, Mountain beaver, Greate...  
$ genus <fct> Acinonyx, Aotus, Aplodontia, Blarina, Bos, B...  
$ vore <fct> carni, omni, herbi, omni, herbi, herbi, carn...  
$ order <fct> Carnivora, Primates, Rodentia, Soricomorpha,...  
$ conservation <fct> lc, NA, nt, lc, domesticated, NA, vu, NA, do...  
$ sleep\_total <dbl> 12.1, 17.0, 14.4, 14.9, 4.0, 14.4, 8.7, 7.0,...  
$ sleep\_rem <dbl> NA, 1.8, 2.4, 2.3, 0.7, 2.2, 1.4, NA, 2.9, N...  
$ sleep\_cycle <dbl> NA, NA, NA, 0.1333333, 0.6666667, 0.7666667,...  
$ awake <dbl> 11.9, 7.0, 9.6, 9.1, 20.0, 9.6, 15.3, 17.0, ...  
$ brainwt <dbl> NA, 0.01550, NA, 0.00029, 0.42300, NA, NA, N...  
$ bodywt <dbl> 50.000, 0.480, 1.350, 0.019, 600.000, 3.850,...

#### 11. Susun data yang terdiri dari name, order, sleep\_total dalam urutan naik berdasarkan kolom order dan sleep\_total dan filter untuk sleep\_total >= 10. Lalu tampilkan 3 pengamatan pertama.

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

# Memanggil data  
sleep <- read.csv("C:/Users/LENOVO/OneDrive/Documents/UNY/Semester 2/msleep.csv")  
head(sleep)

## name genus vore order conservation  
## 1 Cheetah Acinonyx carni Carnivora lc  
## 2 Owl monkey Aotus omni Primates <NA>  
## 3 Mountain beaver Aplodontia herbi Rodentia nt  
## 4 Greater short-tailed shrew Blarina omni Soricomorpha lc  
## 5 Cow Bos herbi Artiodactyla domesticated  
## 6 Three-toed sloth Bradypus herbi Pilosa <NA>  
## sleep\_total sleep\_rem sleep\_cycle awake brainwt bodywt  
## 1 12.1 NA NA 11.9 NA 50.000  
## 2 17.0 1.8 NA 7.0 0.01550 0.480  
## 3 14.4 2.4 NA 9.6 NA 1.350  
## 4 14.9 2.3 0.1333333 9.1 0.00029 0.019  
## 5 4.0 0.7 0.6666667 20.0 0.42300 600.000  
## 6 14.4 2.2 0.7666667 9.6 NA 3.850

# Menyusun data  
result <- sleep %>%   
 select(name, order, sleep\_total) %>%   
 arrange(order, sleep\_total) %>%   
 filter(sleep\_total >= 10)  
  
head(result, n = 3)

## name order sleep\_total  
## 1 Tenrec Afrosoricida 15.6  
## 2 Dog Carnivora 10.1  
## 3 Jaguar Carnivora 10.4

#### 12. Buat variabel baru rem\_proportion yang didefinisikan sebagai sleep\_rem dibagi dengan sleep\_total lalu tambahkan pada data. Lalu tampilkan 3 pengamatan pertama.

result <- sleep %>%   
 mutate(rem\_proportion = sleep\_rem / sleep\_total)  
  
head(result, n = 3)

## name genus vore order conservation sleep\_total sleep\_rem  
## 1 Cheetah Acinonyx carni Carnivora lc 12.1 NA  
## 2 Owl monkey Aotus omni Primates <NA> 17.0 1.8  
## 3 Mountain beaver Aplodontia herbi Rodentia nt 14.4 2.4  
## sleep\_cycle awake brainwt bodywt rem\_proportion  
## 1 NA 11.9 NA 50.00 NA  
## 2 NA 7.0 0.0155 0.48 0.1058824  
## 3 NA 9.6 NA 1.35 0.1666667

#### 13. Buat variabel baru rem\_proportion yang didefinisikan sebagai sleep\_rem dibagi dengan sleep\_total dan bodywt\_grams yang didefinisikan sebagai bodywt \* 1000 lalu tambahkan pada data. Lalu tampilkan 3 pengamatan pertama.

result <- sleep %>%   
 mutate(rem\_proportion = sleep\_rem / sleep\_total,  
 bodywt\_grams = bodywt \* 1000)  
  
head(result, n = 3)

## name genus vore order conservation sleep\_total sleep\_rem  
## 1 Cheetah Acinonyx carni Carnivora lc 12.1 NA  
## 2 Owl monkey Aotus omni Primates <NA> 17.0 1.8  
## 3 Mountain beaver Aplodontia herbi Rodentia nt 14.4 2.4  
## sleep\_cycle awake brainwt bodywt rem\_proportion bodywt\_grams  
## 1 NA 11.9 NA 50.00 NA 50000  
## 2 NA 7.0 0.0155 0.48 0.1058824 480  
## 3 NA 9.6 NA 1.35 0.1666667 1350

#### 14. Buat ringkasan statistik rata-rata, simpangan baku, minimum, maksimum dan total dari sleep\_total berdasarkan grup order.

summary\_stats <- sleep %>%   
 group\_by(order) %>%   
 summarise(  
 avg = mean(sleep\_total, na.rm = TRUE),   
 min = min(sleep\_total, na.rm = TRUE),   
 max = max(sleep\_total, na.rm = TRUE),   
 sd = sd(sleep\_total, na.rm = TRUE),   
 total = n()  
 )  
  
print(summary\_stats)

## # A tibble: 19 × 6  
## order avg min max sd total  
## <chr> <dbl> <dbl> <dbl> <dbl> <int>  
## 1 Afrosoricida 15.6 15.6 15.6 NA 1  
## 2 Artiodactyla 4.52 1.9 9.1 2.51 6  
## 3 Carnivora 10.1 3.5 15.8 3.50 12  
## 4 Cetacea 4.5 2.7 5.6 1.57 3  
## 5 Chiroptera 19.8 19.7 19.9 0.141 2  
## 6 Cingulata 17.8 17.4 18.1 0.495 2  
## 7 Didelphimorphia 18.7 18 19.4 0.990 2  
## 8 Diprotodontia 12.4 11.1 13.7 1.84 2  
## 9 Erinaceomorpha 10.2 10.1 10.3 0.141 2  
## 10 Hyracoidea 5.67 5.3 6.3 0.551 3  
## 11 Lagomorpha 8.4 8.4 8.4 NA 1  
## 12 Monotremata 8.6 8.6 8.6 NA 1  
## 13 Perissodactyla 3.47 2.9 4.4 0.814 3  
## 14 Pilosa 14.4 14.4 14.4 NA 1  
## 15 Primates 10.5 8 17 2.21 12  
## 16 Proboscidea 3.6 3.3 3.9 0.424 2  
## 17 Rodentia 12.5 7 16.6 2.81 22  
## 18 Scandentia 8.9 8.9 8.9 NA 1  
## 19 Soricomorpha 11.1 8.4 14.9 2.70 5