

R_Module4

Joe Vargovich

9/2/2020

```
library(tidyverse, quietly=TRUE)
options(dplyr.summarise.inform=FALSE)
```

Exercise 1

```
data("ChickWeight")
write.csv(ChickWeight, "chickWeight.csv")
head(ChickWeight)
```

Load the dataset chickweight

```
##   weight Time Chick Diet
## 1     42    0     1     1
## 2     51    2     1     1
## 3     59    4     1     1
## 4     64    6     1     1
## 5     76    8     1     1
## 6     93   10     1     1
```

```
day10_20 <- ChickWeight %>% filter(Time == 10 | Time == 20)
head(day10_20)
```

Apply the filter function to get measurements from day 10 OR day 20.

```
##   weight Time Chick Diet
## 1     93   10     1     1
## 2    199   20     1     1
## 3    103   10     2     1
## 4    209   20     2     1
## 5     99   10     3     1
## 6    198   20     3     1
```

#Get the mean and sd of each group and display them.

```
day10_20 %>%
  group_by(Diet) %>%
  summarize(n = n(),
            mean.weight = mean(weight),
            sd.weight = sd(weight))
```

```
## # A tibble: 4 x 4
##   Diet      n mean.weight sd.weight
##   <fct> <int>      <dbl>      <dbl>
```

```
## 1 1      36      130.      56.6
## 2 2      20      157.      71.4
## 3 3      20      188      86.6
## 4 4      19      177.      61.3
```

Exercise 2

The OpenIntro textbook on statistics includes a data set on body dimensions. Instead of creating an R chunk for each step of this problem, create a single R pipeline that performs each of the following tasks.

```
#a. Load the file
body <- read.csv('http://www.openintro.org/stat/data/bdims.csv')

#b. The column sex is coded as a 1 if the individual is male and 0 if female. This is a non-intuitive l
body <- body %>%
  mutate(sex.MF = if_else(sex == 1, 'Male', 'Female'))

#c. The columns wgt and hgt measure weight and height in kilograms and centimeters (respectively). Use
body <- body %>%
  mutate(bmi = (wgt)/(hgt/100)^2)

#d My calculation was not initially correct, I was using the wrong units.

#e. The function cut takes a vector of continuous numerical data and creates a factor based on your giv
body <- body %>%
  mutate(Age.Grp = cut(age,
                        breaks=c(10,20,30,40,50,60,70),
                        right=FALSE))

#f. Find the average BMI for each Sex by Age.Grp combination.
body %>%
  group_by(sex.MF, Age.Grp) %>%
  summarize(n = n(),
            mean.bmi = mean(bmi))
```

```
## # A tibble: 12 x 4
## # Groups:   sex.MF [2]
##   sex.MF Age.Grp      n mean.bmi
##   <chr> <fct>    <int>    <dbl>
## 1 Female [10,20)     23     21.8
## 2 Female [20,30)    139     21.8
## 3 Female [30,40)     61     22.5
## 4 Female [40,50)     31     24.3
## 5 Female [50,60)      4     22.7
## 6 Female [60,70)      2     23.7
## 7 Male   [10,20)      3     25.5
## 8 Male   [20,30)    132     24.2
## 9 Male   [30,40)     57     24.9
## 10 Male  [40,50)     38     26.4
## 11 Male  [50,60)     12     24.8
## 12 Male  [60,70)      5     23.9
```