

# Exercise 2

Charles Carter \*

August 9, 2016

## 1 Between Groups

Create a data frame with two integer columns, Age and Sex, with the supplied data. Age = 24, 18, 19, 32, 19, 25, 22. Sex = M, F, M, F, F, F, M.

```
###delete all objects in your environment (starting over)
# rm(list = ls())
##create abn empty data frame
# data1 <- data.frame(Age = integer(), Sex = character())
##enter the data using an interactive grid
# data1 <- edit(data1)
#make sure Sex is a factor
is.factor(data1$Sex)

## [1] TRUE

#print data
data1

##   Age Sex
## 1  24  M
## 2  18  F
## 3  19  M
## 4  32  F
## 5  19  F
## 6  25  F
## 7  22  M

#prints the names of the variables
names(data1)

## [1] "Age" "Sex"
```

---

\*cartercharles@columbusstate.edu

```

#prints a summary of the data
summary(data1)

##      Age      Sex
## Min.   :18.00  M:3
## 1st Qu.:19.00  F:4
## Median :22.00
## Mean   :22.71
## 3rd Qu.:24.50
## Max.   :32.00

#for describe() function
library(psych)
#prints various descriptive statistics
describe(data1)

##      vars n  mean   sd median trimmed  mad min max range  skew kurtosis
## Age      1 7 22.71 4.89      22   22.71 4.45  18  32    14  0.74    -0.92
## Sex*     2 7  1.57 0.53        2    1.57 0.00   1   2     1 -0.23    -2.20
##          se
## Age  1.85
## Sex* 0.20

```

## 2 Within Groups

Enter the following data in a data frame with three fields, Participant, Before, and After. Participant = 1, 2, 3, 4, 5, 6, 7. Before = 166, 182, 194, 321, 190, 258, 124. After = 160, 142, 167, 207, 192, 198, 100.

```

#create three integer vectors
Participant <- as.integer(c(1, 2, 3, 4, 5, 6, 7))
Before <- as.integer(c(166, 182, 194, 321, 190, 258, 124))
After <- as.integer(c(160, 142, 167, 207, 192, 198, 100))
#create data frame named data2
data2 <- data.frame(cbind(Participant, Before, After))
#print data2
data2

##   Participant Before After
## 1           1    166   160
## 2           2    182   142
## 3           3    194   167
## 4           4    321   207
## 5           5    190   192
## 6           6    258   198
## 7           7    124   100

```

```

#prints a summary of the data
summary(data2)

##      Participant      Before      After
## Min.      :1.0    Min.      :124    Min.      :100.0
## 1st Qu.:2.5    1st Qu.:174    1st Qu.:151.0
## Median :4.0    Median :190    Median :167.0
## Mean   :4.0    Mean   :205    Mean   :166.6
## 3rd Qu.:5.5    3rd Qu.:226    3rd Qu.:195.0
## Max.   :7.0    Max.   :321    Max.   :207.0

#for describe() function
library(psych)
#prints various descriptive statistics
describe(data2)

##           vars n   mean    sd median trimmed   mad min max range  skew
## Participant   1  7  4.00  2.16      4    4.00  2.97   1   7     6  0.00
## Before        2  7 205.00 64.84    190  205.00 35.58 124 321   197  0.58
## After         3  7 166.57 37.30    167  166.57 37.06 100 207   107 -0.54
##           kurtosis    se
## Participant   -1.71  0.82
## Before        -1.10 24.51
## After         -1.22 14.10

```

### 3 Within Groups, Part 2

Create a plain text data file using a text editor like Notepad, and save the file in the same directory as your R. Name the file “data3.csv.” Make sure that the file has the extension “csv,” Notepad will use the “txt” extension by default, and you must change it.

```

#import your data file into R
data3 <- read.csv("data3.csv", header = TRUE, stringsAsFactors = TRUE)
#print the data
data3

##      Occupation Stress
## 1      Nurse      32
## 2      Nurse      28
## 3      Nurse      22
## 4      Nurse      35
## 5      Nurse      29
## 6      Nurse      27
## 7      Nurse      26
## 8 Junior Doctor    31

```

```
## 9 Junior Doctor 23
## 10 Junior Doctor 29
## 11 Junior Doctor 34
## 12 Junior Doctor 26
## 13 Junior Doctor 24
## 14 Consultant 19
## 15 Consultant 16
## 16 Consultant 11
## 17 Consultant 22
## 18 Consultant 20
## 19 Consultant 17

#generate descriptive statistics
summary <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = summary)
summary

##          Group.1 x.Min. x.1st Qu. x.Median x.Mean x.3rd Qu. x.Max.
## 1 Consultant 11.00 16.25 18.00 17.50 19.75 22.00
## 2 Junior Doctor 23.00 24.50 27.50 27.83 30.50 34.00
## 3 Nurse 22.00 26.50 28.00 28.43 30.50 35.00

length <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = length)
length

##          Group.1 x
## 1 Consultant 6
## 2 Junior Doctor 6
## 3 Nurse 7

mean <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = mean)
mean

##          Group.1 x
## 1 Consultant 17.50000
## 2 Junior Doctor 27.83333
## 3 Nurse 28.42857

median <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = median)
median

##          Group.1 x
## 1 Consultant 18.0
## 2 Junior Doctor 27.5
## 3 Nurse 28.0

sd <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = sd)
sd
```

```
##          Group.1      x
## 1    Consultant 3.834058
## 2 Junior Doctor 4.262237
## 3         Nurse 4.197505

var <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = var)
var

##          Group.1      x
## 1    Consultant 14.70000
## 2 Junior Doctor 18.16667
## 3         Nurse 17.61905

range <- aggregate(data3$Stress, by = list(data3$Occupation), FUN = range)
range

##          Group.1 x.1 x.2
## 1    Consultant  11  22
## 2 Junior Doctor  23  34
## 3         Nurse  22  35
```