Exercise 1

Charles Carter *

August 9, 2016

1 Load Data

Go to the textbook companion website: `https://studysites.sagepub.com/dancey/main.htm`. The file
"Data for Figure 1_4" is located in Student Resources, SPSS Data Files, Chapter 1. Download that file and
save it in your student directory. Then, open the file and save it to an R object.

```
#delete all objects in your environment (starting over)
rm(list = ls())
#get the name of the file by point and click
data.file <- file.choose()
#print the name of the data file
data.file

## [1] "C:\\Users\\ccc\\stat3127book\\data\\Data for Figure 1_4.sav"

#load the foreign library to read the SPSS file
library(foreign)
#for read.spss() function
data <- read.spss(data.file, to.data.frame = TRUE)

## Warning in read.spss(data.file, to.data.frame = TRUE): C:\Users\ccc\stat3127book\data\Data
## for Figure 1_4.sav:  Unrecognized record type 7, subtype 18 encountered in system file

#prints the data
data

##     Cigarettes Life_expectancy Average_Income Percentage_Overweight
## 1           23              40             72                     41
## 2           25              60             31                     33
## 3           33              27             29                      6
## 4           28              28             27                     18
## 5           26              43             47                     39
## 6           11              75             39                     12
```

---

*carter$_c$harles@columbusstate.edu

```
## 7              19              52              32              11
## 8              26              45              66              48
## 9               5              74              45              29
## 10             29              44              43              24
## 11             27              62              63              46
## 12             16              50              49              23
## 13             17              63              30              12
## 14             21              68              39              40
## 15              4              76              51              37
## 16             50              51              57               5
## 17             16              58              43              19
## 18             17              71              22              14
## 19             30              58              24               6
## 20             23              50              52              22

#prints the names of the variables
names(data)

## [1] "Cigarettes"            "Life_expectancy"       "Average_Income"
## [4] "Percentage_Overweight"
```

2 Descriptive Statistics

Print descriptive statistics.

```
#prints a summary of the data
summary(data)

##    Cigarettes     Life_expectancy Average_Income  Percentage_Overweight
##  Min.   : 4.00   Min.   :27.00   Min.   :22.00   Min.   : 5.00
##  1st Qu.:16.75   1st Qu.:44.75   1st Qu.:30.75   1st Qu.:12.00
##  Median :23.00   Median :55.00   Median :43.00   Median :22.50
##  Mean   :22.30   Mean   :54.75   Mean   :43.05   Mean   :24.25
##  3rd Qu.:27.25   3rd Qu.:64.25   3rd Qu.:51.25   3rd Qu.:37.50
##  Max.   :50.00   Max.   :76.00   Max.   :72.00   Max.   :48.00

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

##                  vars  n  mean    sd median trimmed   mad min max
## Cigarettes          1 20 22.30 10.22   23.0   22.12  8.90   4  50
## Life_expectancy     2 20 54.75 14.44   55.0   55.56 15.57  27  76
## Average_Income      3 20 43.05 14.30   43.0   42.31 17.05  22  72
```
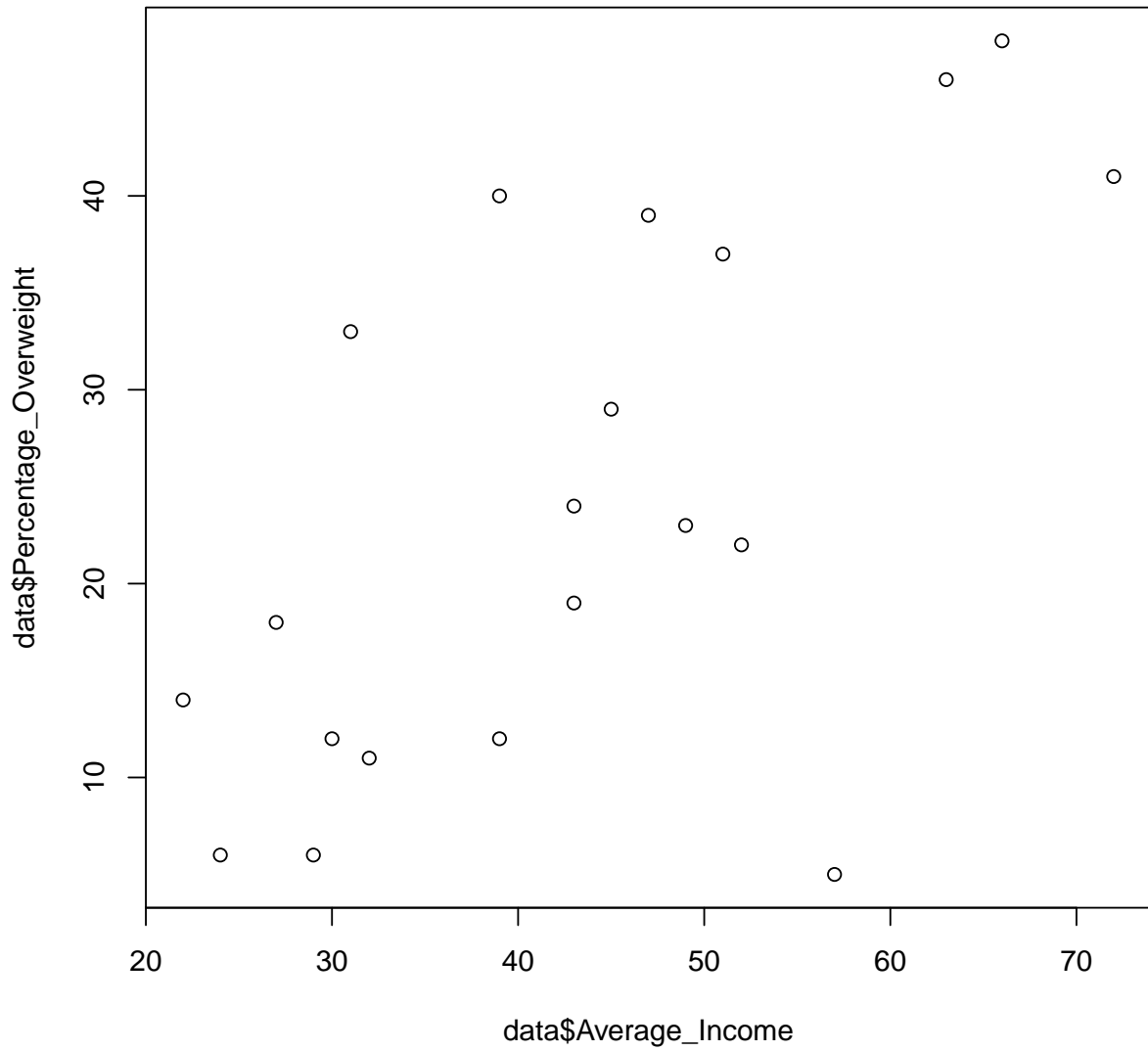
```
## Percentage_Overweight    4 20 24.25 14.02    22.5    23.75 16.31    5   48
##                        range  skew kurtosis    se
## Cigarettes                46  0.47     0.81 2.29
## Life_expectancy           49 -0.25    -0.91 3.23
## Average_Income            50  0.34    -0.98 3.20
## Percentage_Overweight     43  0.21    -1.43 3.14
```

3 Plot the data

```
#plot overweight as a function of income
plot(x=data$Average_Income, y=data$Percentage_Overweight)
```

```
#plot life expectancy as a function of number of cigarettes
plot(x=data$Life_expectancy, y=data$Cigarettes)
```