Data Manipulation in R HW

Dean D'sozua

June 29, 2016

#### Install the package HSAUR2.

* After installing the packages, load the package using the library command.
* After the library command, pull up the data set CHFLS using the data command.
* If you're unsure how to do this, try help().

## Installing package:  
# install.packages("HSAUR2")  
   
library(HSAUR2)  
data(CHFLS)  
  
head(CHFLS)

## R\_region R\_age R\_edu R\_income R\_health R\_height  
## 2 Northeast 54 Senior high school 900 Good 165  
## 3 Northeast 46 Senior high school 500 Fair 156  
## 10 Northeast 48 Senior high school 800 Good 163  
## 11 Northeast 46 Junior high school 300 Fair 164  
## 22 Northeast 45 Junior high school 300 Fair 162  
## 23 Northeast 36 Senior high school 500 Excellent 161  
## R\_happy A\_height A\_edu A\_income  
## 2 Somewhat happy 172 Senior high school 500  
## 3 Somewhat happy 170 Senior high school 800  
## 10 Somewhat happy 172 Junior high school 700  
## 11 Somewhat happy 174 Elementary school 700  
## 22 Somewhat happy 172 Junior high school 400  
## 23 Somewhat happy 180 Junior college 900

#### What are the names of the columns of the CHFLS data set?

The names of the colums are obtained as follows:

names(CHFLS)

## [1] "R\_region" "R\_age" "R\_edu" "R\_income" "R\_health" "R\_height"  
## [7] "R\_happy" "A\_height" "A\_edu" "A\_income"

#### What is the class of CHFLS?

The class of CHFLS is obtained as follows:

class(CHFLS)

## [1] "data.frame"

#### What is the average R\_income for each R\_edu ?

For each level of R\_edu we obtain the average as follows:

moredu<-tapply(CHFLS$R\_income, CHFLS$R\_edu, mean)  
moredu

## Never attended school Elementary school Junior high school   
## 209.8889 416.2172 534.6141   
## Senior high school Junior college University   
## 666.7765 1170.4000 1706.8182

#### Create a table of counts of rows by the values in the R\_health column and the R\_happy column. In other words, how many rows are in the data frame for each combination of distinct values in these two columns?

table(CHFLS$R\_health,CHFLS$R\_happy)

##   
## Very unhappy Not too happy Somewhat happy Very happy  
## Poor 2 4 3 1  
## Not good 7 46 77 9  
## Fair 4 67 350 40  
## Good 1 42 459 80  
## Excellent 0 26 166 150

#### What is your working directory?

The working directory can be obtained as follows:

getwd()

## [1] "C:/Users/demon/OneDrive/Documents/GitHub/ANLY\_510-50\_SU2016/Data Manipulation in R"

#### What files exist in your working directory?

The files contained in the working directory are as follows:

list.files()

## [1] "DeanDsouza\_DataManipulationInR\_HW.docx"  
## [2] "DeanDsouza\_DataManipulationInR\_HW.Rmd"   
## [3] "MyChfls.csv"

#### Save the data frame CHFLS as a csv file in your working directory. Show that CHFLS exists as a csv file

# Write the CHFLS data frame as a .csv file in the working directory  
write.csv(CHFLS, file = "MyChfls.csv", fileEncoding = "ISO-8859-1")  
  
# Listing all files to verify that it has been written  
list.files()

## [1] "DeanDsouza\_DataManipulationInR\_HW.docx"  
## [2] "DeanDsouza\_DataManipulationInR\_HW.Rmd"   
## [3] "MyChfls.csv"

As can be seen from the above list of files, **MyChfls.csv** file exists in the working directory showing that it was written successfully.  
 We can also read from this file again as follows:

temp<-read.csv("MyChfls.csv", header = TRUE, sep = ",", quote = "\"", dec = ".", fill = TRUE, comment.char = "")  
head(temp)

## X R\_region R\_age R\_edu R\_income R\_health R\_height  
## 1 2 Northeast 54 Senior high school 900 Good 165  
## 2 3 Northeast 46 Senior high school 500 Fair 156  
## 3 10 Northeast 48 Senior high school 800 Good 163  
## 4 11 Northeast 46 Junior high school 300 Fair 164  
## 5 22 Northeast 45 Junior high school 300 Fair 162  
## 6 23 Northeast 36 Senior high school 500 Excellent 161  
## R\_happy A\_height A\_edu A\_income  
## 1 Somewhat happy 172 Senior high school 500  
## 2 Somewhat happy 170 Senior high school 800  
## 3 Somewhat happy 172 Junior high school 700  
## 4 Somewhat happy 174 Elementary school 700  
## 5 Somewhat happy 172 Junior high school 400  
## 6 Somewhat happy 180 Junior college 900