# Prep Exercise (PE01) Getting Started

### General Instructions

For this PrepEx you will upload this single WORD file into blackboard. You can find this by going to Weekly Content, Week 1, and then clicking on the PE01 DropBox near the lower 1/3 of your blackboard screen. Each week, Prep Exercises are completed using the Word document, Homework is completed via an R-script (or file).

A few highlights to pay close attention to as you create your R code:

* We expect that you will benefit from the code presented in Introduction to Data Science, as well as examples provided by the instructors, the Internet, and other sources. But you must always, always, always give credit to your sources. For example, if you find a line of code someone else wrote that helps you with developing a solution, by all means borrow that line of code, but make sure to use a comment to document who wrote the code. The same is true from helpful videos or tutorials from internet and YouTube.
* Building on the previous point, when submitting code as your solution to a homework, the comments matter at least as much, if not more than the code itself. A good rule of thumb is that every line of code should have a comment, and every meaningful block of code should be preceded by a comment block that is just about as long as the code itself. Note that the comments go before (not after) the code. As appropriate, comments can also be to the right of a piece of code. Also, white space between logical blocks of code can make your code easier to read. As noted above, you can use comments to identify your submission as your own and you can use comments to give proper credit to your sources. See below for an example of good commenting:
* Sample:

################################################

# IST 687, Standard Homework Heading

#

# Student name: Thadhani Hitesh Chandrakumar

# Homework number: PE01

# Date due: Wed 28th Aug 2019 11:59PM

#

# Attribution statement: (choose the statements that are true)

# 1. I did this work by myself, with help from the book and the professor

# 2. I did this homework with help from the book and the professor and these Internet sources: <provide the urls>

# 3. I did this homework with coaching from <Name of another student> but did not cut and paste any code

# Run these three functions to get a clean test of homework code

dev.off() # Clear the graph window

cat('\014') # Clear the console

rm(list=ls()) # Clear all user objects from the environment!!!

# Set working directory

# Change to the folder containing your homework data files

setwd("C:/Syracuse Fall 2019/Introduction to Data Science IST687 Saltz/ISTX87/Homework")

### Prep Exercise (note some questions require you to do something on your machine but not on this form. This week you’ll see that you need to provide your work to questions 5 and 6. Note: Q6 asks you to include a Screen Shot of your R-code. Be sure to include the comments at the top that show your name.)

1. **Install R (see pg 24 and visit** [**https://cran.r-project.org/**](https://cran.r-project.org/) **for more info)**
2. **Install R Studio Desktop – Open Source License (**[**https://www.rstudio.com/products/rstudio/**](https://www.rstudio.com/products/rstudio/)**)**
3. **While you are there take some time to Take a Tour of the RStudio IDE – pay particular attention to the descriptions of the 4 panes (windows) within RStudio.**
4. **The following should be coded using RStudio. Use the source code pane, and then execute the lines of code, to see the results in the console. Remember to use the required comment block at the top and to comment your lines of code. Be sure you know where you are saving your file. You’ll need to use this file for your homework this week (save the file as a .R file).**
   1. Define the following vectors, which represent the weight and height of people on a particular team (in inches and pounds):

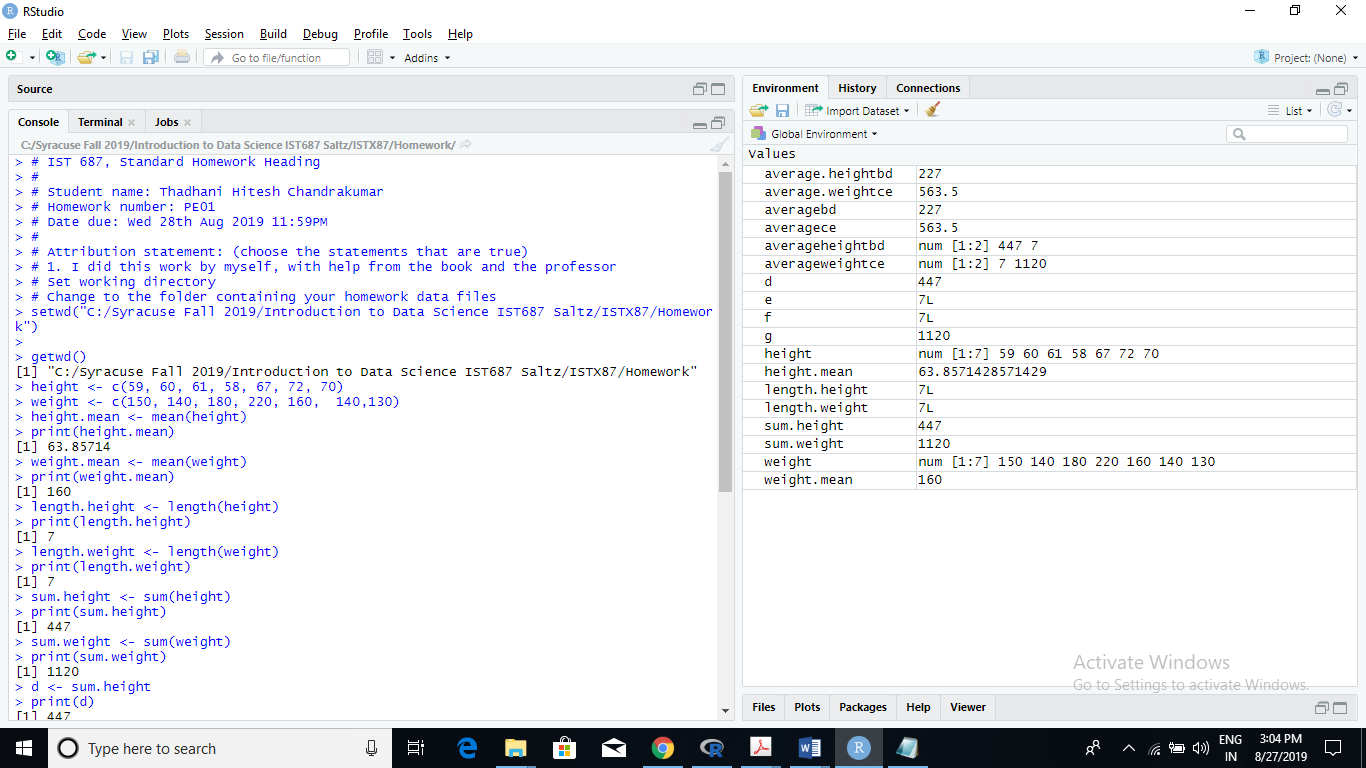
**height <- c(59, 60, 61, 58, 67, 72, 70)**

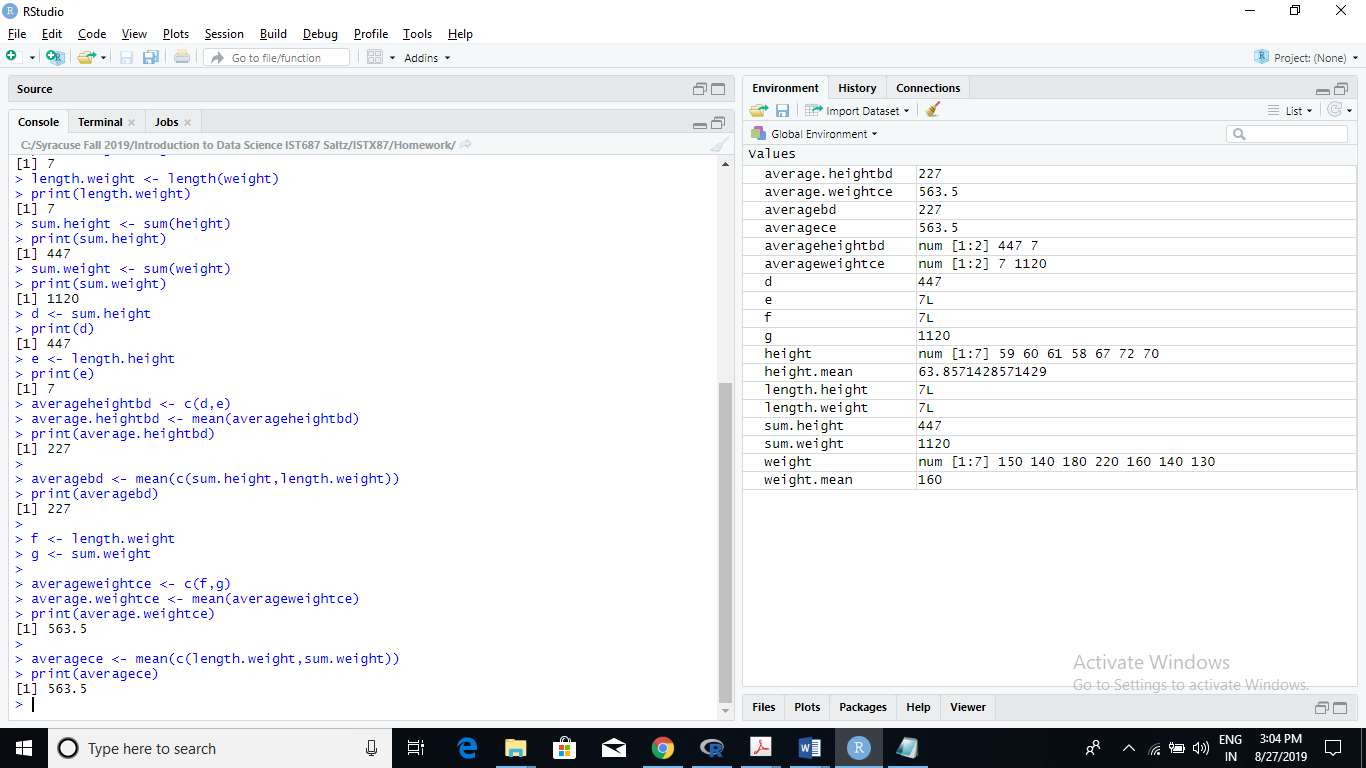
**weight <- c(150, 140, 180, 220, 160, 140,130**)

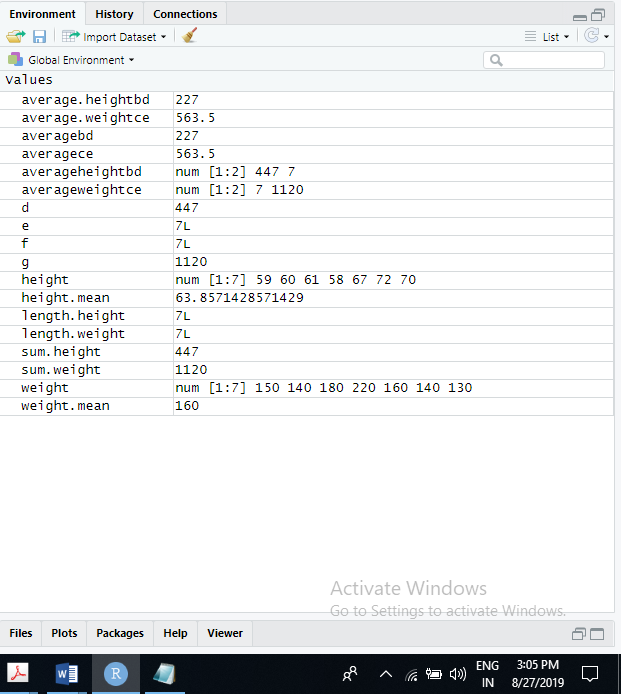
* 1. Define a variable:

**a <- 150**

1. **Calculating statistics using R and fill in the answers where asked to do so:**
   1. Compute the average height and average weight with the **mean()** function
      1. Average height: 63.85714
      2. Average weight: 160
   2. Show the number of observations in height with the **length()** function - 7
   3. Show the number of observations in weight with the **length()** function - 7
   4. Calculate the sum of the heights with the **sum()** function - 447
   5. Calculate the sum of the weights with the **sum()** function
      1. Sum of all weights: 1120
   6. Compute the average height by combining results from B and D
      1. Average height: 227
   7. Recompute the average weight by combining results from C and E
      1. Average weight: 563.5
2. Paste screen captures of your source code pane and your global environment pane below.







1. **Be sure to save your work as the file will become the starting point for this week’s homework. For homework each week, you will turn in the R code. For Prep Exercises like this one you will upload THIS WORD DOCUMENT with your answers.**

***You must submit all Prep Exercises in Blackboard the evening before the lecture.*** Late PrepXs or work submitted in the wrong area of blackboard will not be graded or accepted for credit.