HW2- CMDA

Submit your script to Dropbox by Monday, September 15, 5pm. Name your script ("YourName_HW2.r").
Use the class notes and scripts, as well as R help and internet resources. Ask questions in the Forum for
your colleagues to answer. Work in teams (the project team is a good start) if you want, but submit your
individual file with your work.

1) Load the phsample.RDATA from the Git link provided in the "PDSR" (Practical Data Science with
   R) textbook on page 31.
2) Describe in a sentence or two (R comment) what the dataset is about.
3) Replicate the textbook Listing 2.12, 2.13, and 2.14 code.
4) Comment for each of the above what the code is doing.
5) Discuss with your team and find a data set for your project. It should have at least 7 variables,
   some numeric, some character, and at least 60 observations.
6) Import your data set, create the R data frame and perform any reshaping operations you see fit
   right now.
7) Provide a short data manual with the explanations: where the data came from; what is the data
   set about; what are the variables; what is the meaning of rows/observations (i.e people, cars
   etc). This should be a pdf document separate from your R script.
8) Save your .RData file, your R script used to create the data, the initial data file (csv etc.) with
   your data and the data manual pdf to your CMDA repository, commit changes and synchronize
   with GitHub.
9) Provide the GitHub link to this repository on the forum in the Project Topic.

Suggested reading (not graded):

"**The Signal and the Noise: Why so Many Predictions Fail but Some Don't**" by Nate Silverman

"**Numbersense: How to Use Big Data to Your Advantage**" by Kaiser Fung