

## CMDA 3654 – Fall 2015

### Homework 1

**Due:** Friday 9/4/15 at noon

Submit a single R script and a companion PDF file containing your comments and the outputs to the scholar dropbox. To name your files, please refer to the course syllabus.

1. Fill out the survey below:

[https://virginiatech.qualtrics.com/SE/?SID=SV\\_5oiz3Vhmf204XBj](https://virginiatech.qualtrics.com/SE/?SID=SV_5oiz3Vhmf204XBj)

2. Install R and RStudio.

3. Breeze through Part one of the “Practical Data Science with R” textbook.

3. Go over the “basics.R” script file posted on the course website.

4. Create a free account on GitHub website. Download and Install the GitHub desktop application.

5. Read the following guides from the GitHub website:

- Hello World
- Getting your project on GitHub
- Forking projects

6. Consider the dataset on <http://data.princeton.edu/wws509/datasets/effort.dat>

- a) Print out only the observations with 0 effort.
- b) Print the names of the variables.
- c) Print the names of the rows using one function.
- d) Print out the first and the last 10 observations in the dataset.
- e) Export your data frame as both a text and a .csv files.

7. Load the phsample.RDATA from the Git link provided in the “PDSR” (Practical Data Science with R) textbook on page 31.

- a) Describe in a sentence or two (R comment) what the dataset is about.
- b) Replicate the textbook Listing 2.12, 2.13, and 2.14 code.
- c) Comment for each of the above what the code is doing.