

# MSDS 6306: Doing Data Science

## Live session Unit 02 assignment

**Due: 1 hour before your 3<sup>rd</sup> live session (September 13)**

### Submission

**ALL (non-swirl) MATERIAL MUST BE KNITTED INTO A SINGLE, LEGIBLE, AND DOCUMENTED HTML DOCUMENT.** Formatting can be basic, but it should be easily human-readable. Unless otherwise stated, please enable `{r, echo=TRUE}` so your code is visible.

### Questions

1. (20 points) Put the following questions into RMarkdown as headers (or subheaders if you're making question numbers headings). Be sure to answer the questions underneath each header.
  - What is a basic workflow for reproducible research?
  - What are five practical tips for making research reproducible?
  - Give an example of how you might implement each tip.
  - Which one of these do you think will be the most difficult?
2. (20 points) Download and complete “air\_hist.R” code from the Files Tab on 2DS. You will build scatter plots using the plot function.
  - a. As described in the “TODO Assignment 2: Question 2a” comment, complete the plot function regarding monthly temperature.
  - b. As described in the “TODO Assignment 2: Question 2b” comment, complete the plot function involving ozone.
  - c. Translate these to RMarkdown and put them in your overall homework RMarkdown file.
3. (20 points) Download and complete “knit\_cars.Rmd” in the Files tab on 2DS.
  - a. As described in the “TODO Assignment 2: Question 3a” comment, complete a plot function for Temperature and Pressure
  - b. As described in the “TODO Assignment 2: Question 3b” comment, complete a similar plot function that reverses the two axes.
  - c. This is written in RMarkdown, so just transfer it to your RMarkdown script.
4. (40 points) Complete Modules 8 to 11 in the R Programming course of Swirl. **Copy your code/output to a separate .txt file. It does not need to be included in your**

**RMarkdown file. The grader has requested at minimum to show the 90%-100% progress bar for each Module and what output you had for it.**

- a. Complete “8: Logic”
- b. Complete “9: Functions”
- c. Complete “10: lapply and sapply”
- d. Complete “11: vapply and tapply”

**Reminder**

To complete this assignment, please submit **one** RMarkdown and matching HTML file that includes questions 1-3, and a .txt file containing solely your swirl output (Question 4) at least one hour before your live session on September 13, 2017. Please submit all files at the same time; only one submission is granted.

Good luck!