

# BSDS 100: Intro to Data Science with R

## Assignment 1

by Abbie M. Popa (Adapted from James D. Wilson)

Before we begin, make sure that you have read Chapter 1 of Doing Data Science by Cathy O’Neil and Rachel Schutt entitled ”What is Data Science?” here: <https://www.safaribooksonline.com/library/view/doing-data-science/9781449363871/ch01.html>.

**Directions:** For all questions in this assignment, write complete sentences and fully answer any question that is asked. This assignment is due by canvas submission before class on Thursday (by 2:39 PM, Thursday, August 23). The assignment is worth 30 points. Late assignments will automatically have 10 points deducted.

1. (5 points) Tell me a little about yourself. For instance,
  - (a) What got you interested in taking this course?
  - (b) Have you ever used R before this course? If so, to what extent?
  - (c) Have you used other programming or scripting languages before this course (e.g., C, C++, Matlab, Python)? If so, to what extent?
  - (d) What major are you? What do you plan on doing when you graduate from USF?
  - (e) Tell me at least one thing that you hope to do with R this semester.
  - (f) Feel free to tell me anything else interesting about yourself that you’d like to share.
2. (5 points) From reading the article, what in your opinion is the biggest issue facing the field of “data science,” and what do you think can be done to help fix this issue?
3. (5 points) What, in your opinion, is the biggest difference between traditional statistics and data science? How does technology play a role in this difference?
4. (5 points) Who is responsible for coining the term “data scientist?”
5. (5 points) What, in your opinion, is the biggest difference between data scientists in academia and industry?
6. (5 points) Based on this article, a major obstacle in understanding data science in the first place is understanding what data science actually *is*. From what you’ve read, and what you understand so far, how would you define “data science”?