

## Course Syllabus: MSDS 6306 Doing Data Science

Course Designer: Monnie McGee, PhD  
Course Lead: Daniel W. Engels, PhD (dwe@smu.edu)  
Course Instructors: Faizan Javed, PhD (faizan.javed@gmail.com)  
Tom Tibbett, PhD (tptibbett@gmail.com)  
Jongwook Woo, PhD (jongwook@alumni.usc.edu)

Required Texts: Gandrud, C. (2015), *Reproducible Research With R and R-Studio*. Boca Raton, FL: CRC Press.  
O'Neil, C., and Schutt, R. (2014). *Doing Data Science: Straight Talk From the Front Line*. San Francisco, CA: O'Reilly Publishers.

Readings: As given in the unit overview for each week.

Prerequisites: None.

### Learning Objectives: The student will

- Be able to form a testable hypothesis from an unstructured problem.
- Apply the principles of reproducible research in testing said hypothesis.
- Use tools such as R, R-Studio, knitr, rmarkdown, and Github to organize and document research so that others can reproduce and/or continue your work.
- Apply basis statistics and graphics to explore data.
- Use the principles of “tidy data” to create clean data sets from messy ones using SAS and R.
- Communicate the findings of a project in a clear, concise, and scientific manner.

### Course Expectations

- Watch all asynchronous videos BEFORE coming to class. Class time is for asking questions and working problems related to the asynchronous material to help solidify what you have watched.
- Read assigned material in the text BEFORE coming to class. Class time is for asking questions and working problems related to the asynchronous material to help you solidify what you have read.
- As a data scientist in training, you need to learn to search out answers to questions for yourself before asking your instructor. The first question I will ask when you come to me with a problem is, “What have you tried to do to answer your own question?” If you have done nothing, then I reserve the right not to answer your question. I expect that everyone in the class follow this policy. **Don't help a classmate unless the classmate can demonstrate that he or she has tried to find a solution.** It's amazing what five minutes on the Internet can turn up! Please, train yourself to try to find an answer for yourself before asking someone else. Give yourself a time limit (e.g., “I will search for 30 minutes, and if I can't find anything, I will ask”). It's better for you in the long run!

**Grading:** Live Session Assignments (50%), Case Studies (30%), Videos and Questions during asynchronous material (BLTs) (15%), Live Session Attendance (5%).

**Live Session Assignments (50%):** Data science is best learned by doing. During the live session, we will have short assignments that will involve solving problems using software. Some live session periods will start with questions on the assigned reading for the week. Any assignments given during live session will be due before the next live session.

**Case Studies (30%):** These are really longer homework assignments. There will be two of these, each equally weighted. Deadline for submission will be determined by the live session instructor.

**Videos and Questions during asynchronous material (BLTs) (15%):** Several modules have questions that are asked during or at the end of a video. These questions are called “Bidirectional Learning Tools” or BLTs. Responses to these questions aid the instructor in determining what is clear and what is not clear to students, and will help the live session instructor structure assignments and activities for the live session. There is a grade for watching each unit videos as well.

**Live Session Attendance (5%):** This ought to be an easy grade! Just show up in the live session! I understand that work/illness/family sometimes makes class attendance impossible. Just let me know if you cannot attend a live session, and make sure you watch the recording for that live session and do any assignments for the session that you missed.

**Rescheduling Assignments:** Life happens. Should you need an extension on an assignment, please let your live session instructor know at least 24 hours prior to the due date. The notice should be given via e-mail. We will discuss the best course of action given your circumstances. **Note: Issues with technology are NOT an excuse for late work.** You have been warned to start your assignments early enough so that you can resolve such issues before they affect your ability to turn in the work.

**SMU Honor Code:** Students are expected to abide by the SMU Honor Code, which can be found online at <http://www.smu.edu/StudentAffairs/StudentLife/StudentHandbook/HonorCode>. The Honor Code prohibits academic sabotage, cheating, fabrication, facilitating academic dishonesty, and plagiarism. Definitions of all of these items can be found on the Honor Code website. The penalty for a first offense is a zero for the assignment. A second offense will result in a failing grade for the course and possible dismissal from the program.

### **Submission guidelines for assignments**

- Submit solutions in problem order.
- Use an easy-to-read variable-width font with a minimum of 11-point font. (These include Arial, Helvetica, and Geneva fonts—this document is in Helvetica 11 point.).
- Relevant code and output must be included in-line at the appropriate point using Courier New (or other fixed width) font, in 10-point size. **Inclusion of irrelevant code or output will be penalized.**

- Any graphics must be electronically cut and pasted in-line at the appropriate point of the write-up. You can use Word to resize the graphics appropriately.
- Any mathematical notation must be provided with appropriate use of subscripts, superscripts, and symbols. Use MS Equation or another equation editor if you submit your work in Word.

## **Best Practices for Success in MSDS 6306 (and other courses also)**

**Attendance.** Take responsibility for your commitment. Attendance means not only being there for synchronous sessions but also participating in asynchronous work.

**Citizenship.** You need to be actively engaged to succeed in this class. Talking on cell phones, texting, "Facebooking," tweeting, or leisure web browsing are prohibited in live session. I consider these to be a disruption (not to mention rude).

**Integrity.** A lot of the graded work occurs outside of class, so I expect honesty and integrity in what you submit for evaluation. Evidence of academic dishonesty will minimally result in zeros for all involved parties, and perhaps University-level disciplinary action. Don't risk your academic career.

**Humility.** Don't get lost! Ask questions during the live session. If something isn't clear to you, it probably isn't clear to others either. Questions may arise because I haven't made a connection clear or have inadvertently left out an important point. Your question gives me a chance to explain more clearly. Don't be proud or shy.

**Organization.** Don't procrastinate! This is a technology-driven course. Count on your computer failing or your wireless connection breaking the night before a due date. Start early and give yourself a chance to succeed.

**Deadlines.** You will generally have a week to complete an assignment. Due dates and times will be clearly indicated. Late submissions will be penalized, but it is much better to turn in work late than not at all (or to turn in incomplete/sloppy work). Work turned in after solutions have been posted to the course website will receive no credit.

**Getting help.** If questions arise while doing assignments/exams, do your best to resolve these questions before the assignment is due, first by taking time to seek answers yourself, and then via e-mail to your instructor or other students. **I encourage you and expect you to seek help.**

**Collaboration.** I encourage the formation of study groups and collaboration with your fellow students in tackling the assignments. Working together in groups on homework is permitted, even encouraged.

**However, every student should write up and complete his or her homework independently.**

**Students who chose to turn in exactly the same work will share the grade assigned.** Talking about problems with other people does help in learning, but just copying the solutions from one another doesn't help!

**Looks do matter!** All assignments must be NEATLY executed and organized. You risk a zero on any assignment submitted in a sloppy manner. See submission guidelines for more details.

## **University Policies (i.e., stuff the instructor must follow and cannot change)**

**Grading Policy: Graduate students must receive a C or better in a course in order to pass the course. If a student must retake a course, then the second grade and the first grade are averaged for the purposes of the overall GPA. Failure to maintain a GPA of 3.0 or better will result in dismissal from the program.**

**Incompletes** will be given only in the case of extraordinary circumstances that prevent you from finishing the semester. You must have completed at least 50% of the course with a passing grade to be eligible for an incomplete.

**Disability Accommodations:** Students needing academic accommodations for a disability must first register with Disability Accommodations & Success Strategies (DASS). Students can call 214-768-1470 or visit <http://www.smu.edu/Provost/ALEC/DASS> to begin the process. Once registered, students should then schedule an appointment with the professor as early in the semester as possible, present a DASS Accommodation Letter, and make appropriate arrangements. Please note that accommodations are not retroactive and require advance notice to implement.

**Religious Observance:** Religiously observant students wishing to be absent on holidays that require missing class should notify their professors in writing at the beginning of the semester, and should discuss with them, in advance, acceptable ways of making up any work missed because of the absence. (See University Policy No. 1.9.)

**Excused Absences for University Extracurricular Activities:** Students participating in an officially sanctioned, scheduled University extracurricular activity should be given the opportunity to make up class assignments or other graded assignments missed as a result of their participation. It is the responsibility of the student to make arrangements with the instructor prior to any missed scheduled examination or other missed assignment for making up the work. (University Undergraduate Catalogue)  
**Work-related travel and meetings count as officially sanctioned activities (McGee's addition).**