Statistics 3304: Introduction to Statistical Computing
Fall 2019 Syllabus

Contact Info: Dr. Monnie McGee
Department of Statistical Science
Heroy Hall Room 140
Email: mmcgee@smu.edu

Class Hours: T TH 9:30 – 11:50 p.m. in Harold Simmons Hall 101D

Office Hours: T TH 2:00 – 3:00, Other times by appointment.

Required Course Materials:

- **Computer:** This is a computer intensive course. You need access to your own computer, preferably a laptop that you can bring to class. PC or Mac is acceptable.
- **Software:** Python, R, SAS, and SPSS are required. You can obtain the latest version of R at http://www.r-project.org/. If you wish, you can download RStudio, a software package that facilitates the use of R, from http://www.rstudio.com/. R and RStudio are open source and work for Mac, Linux, and Windows. Python is also open source, available from Python.org. Please download python version 3.
- **SMU has a university site license for SAS and SPSS.** I recommend downloading the full packages. Go to the help desk near Starbucks in Fondren Library if you have trouble downloading SAS and SPSS. There is a video on Canvas that explains how to get SAS and SPSS from apps.smu.edu and via wireless download.

Recommended Materials (I will reference these from time to time):

- **Quick R:** (http://www.statmethods.net/interface/workspace.html). It's free! It's your choice as to whether you purchase the associated textbook at the Manning Publications website.
- **An introduction to R:** http://cran.r-project.org/doc/manuals/R-intro.pdf
- **Alan Elliott and Wayne Woodward. SAS Essentials**
- **SPSS Beginners Tutorials (https://www.spss-tutorials.com/basics/#introduction-to-spss)**

Course Overview

Statistics 3304 focuses on using statistical software to perform basic statistical analysis. In particular, we will use R, SAS, and SPSS to perform the same sorts of statistical analyses that you would have seen in Stat 2331 (regression, t-tests, descriptive statistics). We will typically start with "clean" data and use standard techniques. The point is more to teach the software than it is to teach the statistical methods.

For each of the statistical packages mentioned above, the course will cover the following: language structure, simple programming, descriptive statistics, visualization tools, measures of association, comparison of means and proportions, linear regression, parametric and nonparametric correlation, and smoothing techniques.

Course Objectives:

- To learn basic programming skills.
- To understand of the advantages and disadvantages of various statistical software packages, in particular SPSS, SAS, R, and Python.

- To practice coding and documentation habits that make data reproducible.
- To interpret the output from any statistical software package.
- To communicate the findings of a statistical analysis in a clear, concise, and scientific manner

Prerequisites: Statistics 2331 or equivalent. No calculus is required. No previous statistical software experience is required. Statistics 2331 covers basic graphics, linear regression, probability, confidence intervals, and hypothesis testing. There will be brief reviews of introductory material throughout the semester as we encounter the material in class. Any deficiencies in this material need to be addressed on your own outside of class as early as possible in the semester.

Best Practices for Success in Stat 3304 (and other courses, also)

Attendance. Take responsibility for your commitment. I take attendance at random times, and you can earn points for attendance (see **Grading Policy** below). In most class periods, we do something in class that you will need to turn in before you leave class. Your life will be much easier if you attend class. If you skip multiple class meetings, you risk being dropped at my discretion.

Warning: If you e-mail me to ask me whether there will be attendance taken or a class assignment for a particular class day, I will ignore your e-mail and deduct five points from your grade. There is no acceptable excuse for missing class except for (1) your own illness (2) family emergency – does not include the death of a pet (3) school-sponsored activities. In those three instances only will you be allowed to make up work missed in class. **Note that job interviews, taking roommates to the airport, working on a homecoming float, etc, are NOT excused absences.** This course has ample ways to earn points without your having to make up work should such rare occasions arise.

Citizenship. You need to actively be engaged to succeed in this class. Cell phones, texting, facebooking, tweeting or leisure web browsing are prohibited in class. 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 in class. 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 shy.

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

Hints for Navigating this Course

Before Class: Read any assigned readings and/or watch any videos (posted on Canvas when applicable) for the upcoming class. Sometimes there will be a short assignment to complete before class. Do not panic about your grade on these assignments. They are meant to teach you what you don't know rather than as an assessment of your performance. As such, these short assignments can be done multiple times. Take advantage of this easy A!

During Class: The topic will be interactively discussed and practiced. To successfully participate, you must be familiar with the background material for the day.

After Class: Work on the practice problems at your own pace to hone your basic skills in preparation for the more involved homework assignments. Review material using the online notes provided; revise/clarify any in-class notes you took.

Grading Policy

The grading system is designed to increase mastery of the material and reduce stress due to performance anxiety. I hope you embrace and enjoy this change. Read the below very carefully. **If it sounds ridiculous to you, then you should drop the course now.**

Instead of having a set amount of mandatory activities that you are required to do and then assessing your grade from your performance, I am going to provide a host of opportunities for you to earn points towards your grade, thus allowing you to choose your path according to your interests and skills. Yes, you choose want you want to work on, and keep earning points doing different activities until you achieve the grade you desire. Are you an excellent test-taker? Then take the tests. Not good at taking exams? Then do alternative assignments to earn your points. I would suggest mixing it up and doing a little of everything to cover all your bases and ensure you get enough points.

FAIR WARNING I: You can't wait until the last minute to get the grade you want! The idea that you can do a whole bunch of stuff in the last weeks of class to make up for slacking all semester is a delusion! Almost all of these assignments consist of turning in things every week and/or attending scheduled events all semester long. AND THERE ARE NO 'MAKE-UPS' or 'EXTRA CREDIT' or any other additional opportunities offered at the end of the semester to earn points. **Read my lips – no extra credit!** Plan your semester wisely!!!

FAIR WARNING II: Please note that the point totals shown for each activity are the **maximum** points possible; you will not get all the points simply for participating. For example, you can earn up to 100 points for taking the Final Exam, but if you only get half the questions right (a 50% conventional grade) than you would only receive 50 points towards your final grade (50% of 100 points = 50 points earned). The same is true for other assignments. For some assignments (like weekly online work) you get bonus points for completing all assignments.

FAIR WARNING III: The grading system is designed so that you cannot get an A (or even a B) without attempting multiple types of assignments. So, even if you hate exams, you will probably have to take at least one in order to maximize your grade in the course.

FAIR WARNING IV: If you don’t turn in an assignment the day and time it is due, the penalty for a late assignment is as follows: 1 minute – 12 hours late, 5 percent; 12.1 hours to 24 hours late; 10 percent; more than 24 hours late – not accepted.

Assignment Breakdown.

Assignment	Maximum Possible Points
12 weekly online work (30 points each)	360 + 40 for doing all 12 = 400
Midterm Exam	100
Final Exam	100
Exit and Entrance tickets	25 @ 10 points each (250 total)
Mini-Projects	5 @ 50 points each (250 total)
Coding Challenge	3 @ 50 points each (150 total)
Software discovery mini tutorial	50
Peer Review on Mini Tutorial	10 points each

Statistics in the News	10 points each up to 50 points
Random Attendance (double points on 11/26. 12/3, and 12/5)	?
Helping others	5 points each up to 50
Flash Assignments (on Slack)	?
Total (Maximum)	1400 + ??

Final Grade Breakdown (1400 points + ?):

1200 points = A
1150 points = A-
1100 points = B+
1050 points = B
1000 points = B-
950 points = C+
900 points = C
850 points = C-
750 points = D+
700 points = D (Minimum passing grade for P/F)
less than 700 = LOSE

Explanations and Due Dates for Assignments

Weekly Online Work (WOW - open on Tuesday at 6:00 p.m., due Monday at 11:59 p.m.)

Each week there will be an online assignment that is meant to (a) determine whether students have examined the required material before class and (b) to solidify knowledge of the methods introduced in class. The assignments will be short (10 to 20 questions) and mostly multiple choice. You may need software to help solve some of the problems. The beauty of these quizzes is that you can do them up to five times each and the highest grade will be recorded. There is also a bonus for completing all of the quizzes. This is BY FAR the easiest way to get a head start on the A you want! Not only that, research on teaching and learning has shown that short weekly assignments are good teaching tools. By practicing retrieval of information, you learn that information better. There will be no WOW the week of Thanksgiving.

Midterm Exam (Thursday, October 3)

The mid-term exam is given in class. You will not be expected to write code during the exam, but you will be expected to read and interpret output from code. Should we need to reschedule an exam, at least one week's notice will be given.

Final Exam

The final exam will be given on the day and time scheduled by the registrar's office (see below). It will be given only at that time. If you plan to start your holiday before that date, then you cannot take the final. Period. End of story. Don't even ask. You will be expected to write code during the exam in order to complete a data analysis.

Flash Assignments

On occasion, I will get the urge to give an extra assignment. When I do, I will post questions as a quiz on Canvas and send a message to our course Slack channel to let you know it's there. Flash Assignments have variable point values. I don't know how many there will be during the semester. The quiz will close six hours after the tweet is sent (but there will never be a due date on a weekend). **Once**

a Flash Assignment is closed, I will not reopen it, even if you e-mail me that you intended to do the assignment.

Mini-Projects
In order to improve data analysis skills, you have to do data analysis. Therefore, there will be a question to answer with data assigned every 10 days to 2 weeks. Use of software is required on data analysis assignments. The deliverable is a 5-page (maximum) write-up with introduction, methods, code, results, and conclusion sections.

Coding Challenges
Some of you like to code, and you are good at it. I will ask you to code something, with very little guidance, in a particular language (Python, SAS or R, since SPSS and Excel technically aren't computer languages). Sometimes the coding will involve data wrangling, sometimes it will involve simulation, sometimes it will involve visualization – or all three! The deliverable is an R Markdown or Jupyter document with the code, explanation of the code, and the results embedded. Part of your grade will be determined by whether I can reproduce your results without editing the document.

Software Discovery Mini Tutorial (November 26 and December 3)
SPSS, R, SAS and Python can do many things that I will not have time to do in this class. Some of those things I know about, some I don't. For these assignments, you are to find something that one of our software packages does that we haven't talked about in class. For R or Python, you might find an interesting package that we haven't mentioned. For SAS, you might describe a new procedure. There are many possibilities. The end result is a presentation to the class that (1) describes the novel method and (2) gives an example using data. You must prepare PowerPoint or KeyNote slides. Bonus points if you tell me something I don't know! Note: I have reserved two class periods for these tutorials. The length of the presentation will depend on the number of students who sign up for this option. The sign up sheet will be available in early November.

Peer Review of Software Tutorial (when tutorials are given)
One of the objectives of this course is to practice methods that result in reproducible data analysis. In order to assess this objective, students will be asked to review the work of classmates. If a classmate can't run your code on the data you provide and get the same results that you did, then the data aren't reproducible. For each assignment, students will be asked to run the code of other students and check the results. Completing each peer review ON TIME is worth 10 points for the reviewer and is 10% of the software tutorial grade.

Statistics in the News (due at any time during class or office hours)
This is not technically related to software, although it could be. Your job is to find an article from a reputable news source that discusses the results of a scientific study. Send a link of your article to Slack and write a short paragraph mentioning the explanatory and response variables involved, the individuals studied, the sample, the population, the result, your assessment of the result and an explanation for your assessment.

Helping Others (at any time)
Some of you already have a lot of experience with statistical software. Some do not. Those of you with experience should help those that do not have experience. In order for the helper to get points, the one who is helped must write post a Slack message giving me the name of the person who helped them, what they did, and the date and time it was done. This is a concrete way that the person who was helped can say "Thank You". Help can be given inside or outside of class.

Random Attendance (on randomly selected class days)
I will take attendance on random days during the semester. I don't even know what days or how many. I'm not even sure how many points I will give (it could be 5 points, or maybe 25 points – depends on my mood). I will show you how I determine the random attendance days when we study simulation.

Submission guidelines for data analysis assignments

- **The answers and code you submit must be your own.** Copying verbatim from another student's assignment or a previously posted solution (I know about Chegg and Course Hero) will result in a 0 for all students involved, regardless of who copied from whom.
- Your name must be at the top of the first page and on each successive page.
- Submit solutions in problem order.
- Both the text of the problem and the solution must be provided for every problem assigned.
- If you use MS Word to type homework, use an easy to read variable-width font (I like Ariel, Helvetica, and Geneva fonts) with a minimum of 12 point font.
- Relevant computer code and output from the code 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 from R must be electronically cut-and-pasted in-line at the appropriate point of the write-up. Resize graphics window in R to optimize the readability. I will show you how to do this in class.
- 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.
- **Looks do matter!** All assignments must be NEATLY executed and organized. You risk a zero on any assignment submitted in a sloppy manner.

Other Policies

Deadlines: Due dates and times will be clearly indicated. Assignments will be submitted via Canvas. Late assignments will be penalized. It is up to you to weight the risk of turning in sloppy/incomplete work on time versus turning in complete work late. Any assignment turned in after the grades for that assignment have been posted on Canvas will not be graded and will receive a grade of 0.

The definition of "late" is after 11:59 p.m. on the published due date of the assignment, unless otherwise specified. If you miss class on the day homework is due, and then turn in the homework after class, you will be penalized five percent. Don't skip class to complete a homework assignment!

Exception: WOWs and Flash Assignments have hard deadlines. They cannot be turned in late.

Getting Help: If questions arise while doing assignments/exams, you must resolve these questions before the assignment is due by asking questions in class, collaborating other students, or by discussing the problem with me personally in office hours or by appointment. **I encourage you and expect you to seek my 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 assignments independently. Students who chose to turn in exactly the same work will receive zeroes, regardless of who**

copied whom. Talking about problems with other people does help in learning, but just copying the solutions from one another doesn't.

Important Dates (may be changed with one week's notice):

- **Midterm Exam: Thursday, October 3 during Class**
- **FINAL EXAM: POST THE DATE OF THE FINAL EXAM ON SLACK FOR YOUR FIRST FLASH QUIZ! WORTH 5 POINTS. DUE TUESDAY, AUGUST 27 AT 11:59 P.M.**

From the Provost (i.e. University Policies over which I have little control)

Incompletes will only be given 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/SASP/DASS to begin the process. Once approved and registered, students will submit a DASS Accommodation Letter to faculty through the electronic portal DASS Link and then communicate directly with each instructor to 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 me in writing (e-mail is OK) by Wednesday, September 11, and should discuss with me, 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 will 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)