Statistics 6336:Statistical Methods and Data Analysis
Fall 2021 Syllabus

Instructor:	Dr. Monnie McGee
	Department of Statistical Science
	Heroy Hall Room 142
	Email: mmcgee@smu.edu
Class Hours:	T TH 2:00 – 3:20 p.m. Hyer 106
Office Hours:	T TH 3:30 – 4:30 in Heroy 142, by appointment on Fridays. Mondays and Wednesdays are for research!

Mask Policy: In consideration of students who have vulnerable family members at home, masks are required for all students. This masking policy is subject to change during the semester, and any changes will be posted clearly in Canvas announcements.

Course Overview

The first semester of Statistical Analysis is an introductory graduate-level course that emphasizes statistical methods for univariate data. Students will become familiar with the concepts of sampling distributions and statistical inference (confidence intervals and tests of statistical hypotheses) for averages and standard deviations obtained from simple random samples, and for two-sample inference, with either independent samples or matched-pairs samples. In addition, students will learn methods for handling count data. Students will learn about resampling methods (jackknife and bootstrap), multiple testing, nonparametric methods, and testing of the equality of many means (analysis of variance).

The course is a blend of applications and the theory necessary to fully understand the appropriateness of the methods used to analyze the data. They will receive instruction in the use of SAS statistical software, which enables you to concentrate most of your time on interpretation and decision-making. Both oral and written communication of statistical ideas will be emphasized.

Course Objectives:
- Recognize research designs that will allow inference of causation and those that will not.
- Distinguish when a sample is representative of a population, and determine how sample selection affects inferences to a population.
- Describe the importance of randomization in data collection and treatment assignment.
- Interpret output from statistical software, in particular, SAS.
- Carry out hypothesis tests and obtain confidence intervals for discrete and continuous data, and interpret the results in context.
- State the assumptions required for validity of t-tools, and describe their robustness to violation of these assumptions.

Course Materials:
- **Textbook:** Ramsey, FL and Schafer, DW (2013). *The Statistical Sleuth: A Course in Methods of Data Analysis*, 3rd Edition. New York: Brooks/Cole.
- **Software:** We will use SAS and R in this course. You can obtain the latest version of R at http://www.r-project.org/. **SMU has a university site license for SAS.** If you have a PC, I recommend downloading the full package. I will post a video on how to get SAS downloaded on your computer. If you have a MAC, you will have to use the Citrix server version of SAS.

- **Recommended:** Elliott, AC and Woodward, WA (2015). *SAS Essentials, 2nd Edition*. New York: Wiley.

Canvas LMS: Supplemental course materials will be available on the course-management system Canvas. The Calendar on the course website gives assigned readings associated with each lecture. You will also be assigned readings from time to time from statistical journals. Please read these materials before class. All lecture notes will be made available on Canvas. Remember that the notes do not replace the textbook, nor cover all the material in it. The lecture notes are cross-referenced with the textbook in the course schedule to facilitate the integration of the textbook and the lecture notes.

Course Model: This course will be taught using "Just in Time Teaching", which is a version of the flipped classroom model. That means that some lecture material will be given outside of class via readings from the textbook or other items provided by the instructor. You are expected to examine these materials prior to coming to class. You may also be asked to bring a short assignment to class. During class, we will work problems from the textbook and discuss strategies for analyzing data.

Deadlines: You will generally have a week to complete an assignment after it is assigned. 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 weigh the risk of turning in sloppy/incomplete work on time versus turning in complete work late. Please do not email me and ask for my permission/advice on turning in late work. You are adults and I trust your decision on this.** Late penalties are as follows:
- 1 minute to 12 hours late: 5% of grade
- 12 hours 1 minute late to 24 hours late: 10% of grade
- 24 hours 1 minute to 48 hours: 25% of grade (see exception below!)
- After 48 hours – not accepted.

Exception: On the rare occasion when I grade an assignment and post the solutions within 48 hours of the published due date, no late work will be accepted after grades and solutions are posted.

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.**

The discussion page will be enabled for each data analysis assignment. You are free to discuss problems with one another and give hints. However, do not post correctly running, complete code or complete answers to problems.

Collaboration: Working together in groups on homework is permitted. **However, every student should write-up and complete his or her assignments independently. Students who chose to turn in exactly the same work will both receive a 0 for the assignment regardless of who copied from whom.** Talking about problems with other people does help in learning, but just copying the solutions from one another doesn't.

Study Groups: Please form groups for studying. It is best if your groups meet once per week to go over concepts learned in class and to do problems in the text, but not necessarily assigned as homework.

Grading: Class Summaries (5% of grade), Weekly Online Work (5% of grade), In-Class Work (10%), Writing & Journaling (10% of Grade), Mini-Projects (15%), Mid-term (15%), Course Project (20%), Final (20%). Specific Details on each type of assignment follow.

Class Summaries (5% of grade):

At the end of each class, I will ask one of you to prepare a summary the material from that class day to be given at the beginning of the next class. It is best if you prepare for this as soon after the previous class period as possible. The purpose of this is three-fold (1) solidify the material in your head (2) give your class mates a summary of the material and (3) practice oral presentation skills. Some sort of slides is required.

Weekly Online Work (5% of grade):
Each week there will be assigned reading material. The purpose of the weekly online work is to solidify key aspects and concepts of the reading material. These assignments will be mostly multiple choice and automatically graded in Canvas.

In-Class Work (10% of Grade):
There will always be something to do in class. Most of the time it will be a data analysis question that will require coding and analysis. I will ask you to finish these in class and turn them in at the end of class.

Writing & Journaling (10% of grade):
Eventually, each one of you will write two very important papers in this program. The first will come in the second year after you pass your basic exams (which you will all do on the first try, I'm sure!). It is the PhD Qualifying exam (lovingly called the "Supertest"), where you will summarize and synthesize three journal articles on similar topics. The second one is your dissertation. This year I am experimenting with some writing exercises to prepare you for writing these documents and others as you continue your career at SMU.

The first type of assignment is a journal, where you will keep solutions to homework that is not assigned in class (see study groups), solutions to previous basic methods exams, and notes from the text that you take as you read. I expect that you will take notes on the weekly Friday seminars, also. Journals can be either handwritten or electronic. They are due on the last day of class and I will give them back to you in plenty of time to use them to study for the final.

There will also be formal writing assignments throughout the semester. Information on these assignments will come later in class. Participation in class discussions in Canvas also counts in this category.

Mini-Projects (20% of grade)
In order to improve data analysis skills, you have to do data analysis. Therefore, there will be a data analysis problem assigned every 2 - 4 weeks. Use of software is required on data analysis assignments. Part of the grade for each assignment is whether or not the code was reproducible. Assignments will be started in class and you will have some time each class period to work on them and ask questions. Even so, you will probably need time outside of class to complete the work. Parts of each assignment will be done in class.

Submission guidelines for Mini Projects

- 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 SAS or R must be electronically cut-and-pasted in-line at the appropriate point of the write-up. I will show you how to download graphics from SAS and R 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.

Group Course Project due December 3 in class (20 %): One of the issues facing the discipline of statistics is the lack of reproducibility of p-values from various studies. In 2016, the ASA published a statement on the use of p-values. This semester's project is to examine that statement and other associated papers. The first part of the project is to read the ASA statement (found on Canvas) and the articles found in the special issue of The American Statistician on p-values. After that, there are several directions that the project can go.

1. Education: How should p-values be taught to students in introductory statistics courses? Can you come up with an activity or a demonstration that would help students understand the issues scientific reproducibility where the p-value is concerned? Demonstrate your method in class.

2. Research: Some researchers have suggested better ways to determine statistical significance. Are they really better? You will have to define "better" and test your results with theory or simulation.

3. Literature: Is anyone using these methods other than statisticians? Is the word getting out? This will involve a literature search of disciplines other than statistics.

4. Data Analysis: Take a paper from another discipline and analyze the data using one of the "new" techniques for determining the magnitude of differences between treatments.

5. Experimental Design: Design an experiment to determine whether a particular method for teaching substitutes for p-values is better than another one – or maybe whether students retain the "new" p-value method better than the one typically taught.

6. History: Where did the notion of a p-value come from? Why $p < 0.05$? Is that reasonable? How long did it take before someone questioned the idea?

7. Big Data: How can we measure differences between groups with extremely large numbers of observations? The traditional p-value will always be small in this case.

8. Another idea that I haven't thought of yet. Please talk to me about your idea before you get too far into it.

This is a group project. Groups will be chosen by your interest in the topic. Please do not talk to your friends and decide what you will do together. Each of you needs to think about this on your own, then tell me what you want to do. I will ask you to give me a first and second choice. If your choice is 8, then you need to write down a brief description of what you want to do. Maybe someone else will have a similar idea, and I can group you together. Groups will form after the mid-term. Please read the articles by that time.

Mid-Term Exam (20% of grade)
The mid-term exam will be given in class. You may be expected to write code (SAS or R) during exams, and to read and interpret output from code. Should we need to reschedule the exam, at least one week's notice will be given. The midterm will be given sometime in mid-October, likely the week after Fall Break.

Final Exam (25% of grade) – Monday, December 14, 11:30 a.m. to 2:30 p.m.
I do not make the final exam schedule and I cannot change the date or time. Please make your winter break plans accordingly. If there is interest, I will schedule a final exam review for a reading day at a time that fits everyone in the class.

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 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. (https://www.smu.edu/StudentAffairs/Chaplain/ReligiousHolidays).

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. (See 2020-2021 SMU Undergraduate Catalog under "Enrollment and Academic Records/Excused Absences.")

Student Academic Success Programs: Students needing assistance with writing assignments for SMU courses may schedule an appointment with the Writing Center through Canvas. Students wishing support with subject-specific tutoring or success strategies should contact SASP, Loyd All Sports Center, Suite 202; 214-768-3648; https://www.smu.edu/sasp.

Pregnant and Parenting Students: Under Title IX students who are pregnant or parenting may request academic adjustments by contacting Elsie Johnson (elsiej@smu.edu) in the Office of the Dean of Students, or by calling 214-768-4564. Students seeking assistance must schedule an appointment with their professors as early as possible, present a letter from the Office of the Dean of Students, and make appropriate arrangements. Please note that academic adjustments are not retroactive and, when feasible, require advance notice to implement.