Course Syllabus: MSDS 6371

Live Session Instructors:	Dr. Bivin Sadler and Dr. Jacquie Cheun
Office Hours:	TBA by professor
Course Text:	Ramsey, F. L., and Schafer, D. W. (2013), *The Statistical Sleuth: A Course in Methods of Data Analysis* (3rd ed.), Boston, MA: Brooks/Cole, with associated website www.statisticalsleuth.com.
Prerequisites:	A previous introductory statistics course.
Graders/TAs:	TBA by Professor

The text is available in electronic version from CengageBrain.com and is much less expensive this way!

All elements of the syllabus are subject to change by the instructor.

Before taking this class, you should know

- Statistical methods from an introductory statistics course: appropriate use of the mean and median, interpretation of boxplots and histograms, use of simple linear regression, probability calculations using the Normal distribution, applications of the central limit theorem, interpretation and calculation of confidence intervals for a population mean, interpretation and calculation of a *p*-value for any hypothesis test.

- If you have serious deficiencies in the material mentioned above, you must address these yourself on your own outside of class as soon as possible.

Learning Objectives: The student will

- Gain a better understanding of basic statistical methods covered in introductory statistics courses and problems that arise when these analytic methods are applied to real-life research problems

- Demonstrate understanding of the advantages and disadvantages of a given experimental design, particularly with respect to the type of conclusions that can be made

- Appropriately apply the methods discussed in the course to numerical and categorical data

- Communicate the findings of a statistical analysis in a clear, concise, and scientific manner

Course Expectations

- As a Data Scientist in Training, you need to learn to search out answers to questions for yourself before asking your instructor. 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 one hour and if I can't find anything, I will ask). It's better for you in the long run!!

- Download SAS prior to the first live session meeting. Mac users must use the version at apps.smu.edu. If you have questions on obtaining SAS, please contact help.smu.edu. The live session instructors cannot help you with questions regarding downloading and installation of SAS.

- Play around with SAS. It is a very powerful software program. Asynchronous videos regarding the use of SAS are available in the LMS, but there is no way that they can cover every scenario! Please examine SAS help (available online or within the software itself) before asking your

instructor a question about the use of SAS. You also have the Elliott and Woodward book at your disposal. And don't forget Google!!!

- Watch all asynchronous material for the course. If you do not view required materials for a given week, the instructor has the right to refuse to answer any question over those materials until you have viewed them.

Course Coverage

The course will cover chapters 1–12 of the text. Much of the material may seem like a review to those of you who have had a previous course in statistics. That assessment is somewhat true; however, we will pay more attention to sample size calculation and experimental design than a first course typically does. Furthermore, we will concentrate on understanding WHY a particular technique is appropriate and HOW to interpret the results.

Grading: Midterm Exam (20%), Final Exam (25%), Homework (5%), Project (5%), Live Session Attendance (10%), Asynchronous Video questions {10%), For Live Session Questions (25%)

Homework (5%): Homework will be assigned most weeks and is due on Sunday at 11:59 p.m central time. after that week's live session. No late work will be accepted but we will drop the lowest homework grade. Solutions will be posted likely on the following Tuesday.

All data sets for cases studies and exercises are available (or will be when assigned) on the 2DS website. Datasets for problems assigned out of the book can also be found at the text book website: www.statisticalsleuth.com. Again, your lowest homework grade will be dropped.

Project (5%): In lieu of typical homework for weeks 13 and 14, a group mini-project will be assigned.

Live Session Attendance (10%) Note: **Participation** includes **attending every Live Session from start to finish** (the scheduled hour and a half.) This includes sharing your video at all times to facilitate discussion and the full classroom setting.   For an '80s explanation of the dilemma not being physically present causes, see the following link from the '80s cinema hit *Real Genius*! For those of you who have seen the movie, you will enjoy the whole montage! Those who have never seen it, you are in for a treat (watch the whole movie!) … please pay close attention to 2:45, 3:10 and 4:35. https://www.youtube.com/watch?v=CfL8N_bDo1E

Asynchronous Video questions (10%) The course is structured with both synchronous and asynchronous sessions. In order to participate fully in the synchronous sessions, you must complete ALL of the material for the asynchronous sessions each week. This includes the concept check questions and activities that occur in-between videos and during the asynchronous material.

These are questions asked during the asynchronous videos that require an answer before going on to the next part of the video. Completion **_of all_** of the asynchronous Video Questions is required. In some open ended questions everyone can see each other's answers… this is meant to generate discussion will be very helpful as a tool to help master the material!

For Live Session Questions (25%): The For Live Session Assignment is a set of questions and activates that the student is to work on and provide a response to **_before_** live session.

- The For Live Session Assignment is due 24 hours before live session.

- It is important to note that the goal of the activities is to become familiar with the methods, ideas and implementation involved in that activity so that we can efficiently iron out all the details in live session.

- Analogy: You are building the pieces of puzzle in the For Live Session Assignment and we are putting them together to see the big picture in live session.

- It is **_not_** expected that the student have all the correct answers. The expectation is that each student spend the allotted time (indicated next to the activity) on each activity so that we can discuss the details in live session. In short, the requirement is to give genuine effort. If a student is stumped and works on a problem for the allotted time without much perceived success, that student should have several specific and well written questions so that the professor can address those in live session.

- Again, if you max out the indicated time without finishing the activity and you don't have more time to finish, simply write up what you have learned by that time and **record the questions that were generated** and we will address those in live session!

- We want to develop the questions before live session so that we can use the live session time to effectively answer any questions and master the material! (Build the pieces outside of live session and put them together to see the big picture in live session and in the homework.)

Discussion Boards: Students are not required to post in discussion boards, unless specified by the professor.

Midterm Exam (20%): There will be a midterm exam at week 7 of the course. It will cover concept and hand-calculation questions, as well as a data analysis question. The exam will be held online on **Saturday, February 22, 2020 from 11am to 2pm Central Standard Time**. Please clear your schedule now! We will have a review for Live Session 7.

Final Exam (25%): A final exam will be given the last week of the course. It will be comprehensive, containing concept and hand-calculation questions, as well as a data analysis question. It is a cumulative exam and is similar in format to the midterm. We will have a review for Live Session 15. This test will be from **11am to 2pm Central Standard Time on Saturday, April 18, 2020**.

Rescheduling Exams: Life happens. Should you need to reschedule an exam, please give notice to your live session instructor at least 24 hours prior to the live session in which the exam review is discussed (Unit 7). The notice should be given via e-mail. You and your instructor will discuss the best course of action given your circumstances. Retakes of exams will not be allowed, and a missed exam cannot be made up if notification is received AFTER the exam has taken place.

Flow of the Class. (This will help with the what, when and where!)

1. Complete the Asynch videos and concept check questions for unit X. (The 'X' is for the unit number.)
2. Get the 6371 For Live Session Unit X. pptx slide deck off of our GitHub.
3. This slide deck will have questions on some of the slides, these are the For Live Session questions that make up the For Live Session Assignment. Complete the answers to these questions with at least one slide per question. The idea is that you will present some or all of these in a breakout during live session. Make sure and add at least 4 takeaways on one slide and any questions you have on the last slide.
4. Submit the slide deck to the For Live Session Assignment: Unit X assignment on 2DS 24 hours before your Live Session. This is the absolute latest that they can be turned in without a penalty. It will be counterproductive to work on the slides during or after the live session therefore any slides turned in

after the deadline will receive a maximum score of 70. With that said, please try and turn in the slides 24 hours in advance. If you are able to turn them in the day before live session, you will receive more detailed responses from your professor and I will be able to work your questions into live session ... which is good for the whole class!

5. Come to and participate in live session! Fun!

6. After live session you will have a few problems to work on your own (Homework). Finish these problems and turn them in to the Homework Assignment: Unit X assignment on 2DS (where again the "X" is for the Unit number.) The due date for this is Sunday 11:59pm central time.

7. Your TA will begin to grade these and I will release the solution soon after.

8. Begin the process for the next Unit. (Unit X + 1!)

Honor Code

The Honor Code

When you signed your letter of intent to enroll in the program, you initialed the following statement:

> "I have read and agree to abide by the SMU Honor Code available online at: http://www.smu.edu/StudentAffairs/StudentLife/StudentHandbook/HonorCode"

Please know that the Honor Code is taken very seriously.

Honor Code Violations Include: Academic Sabotage, Cheating, Fabrication, Facilitating academic dishonesty, Plagiarism among others.

Collaborative Work

Data science is a collaborative subject

Most professors encourage collaborative work except when explicitly prohibited (usually on quizzes & exams)

Collaboration means helping one other, not copying answers from one another.

Students who turn in exactly the same answers to the same homework will share the grade assigned (i.e. if two students have the same answers, and the grade on the assignment is a 90, then each student will receive a 45).

Some instructors may impose stricter penalties

Okay	Not Okay
Discussing problems (or even solutions) with other students.	Sending your answers to a student or receiving them from another student.
Telling another student how you did something in SAS (which PROC and how to use it, etc.) or sending "snippets" of code (i.e. an example of how to use a PROC) to help them out.	Sharing anything, including ideas, when expressly forbidden, such as on exams.
	Passing off anyone else's work or explanations as your own.
	Using solutions from prior terms to inform your own work.

Plagiarism

Here is an example of plagiarism:
A regression is a statistical analysis assessing the association between two variables. It is used to find the relationship between two variables.

The following is NOT plagiarism:
"A regression is a statistical analysis assessing the association between two variables. It is used to find the relationship between two variables." (https://www.easycalculation.com/statistics/learn-regression.php).

The difference is in the punctuation and the attribution.
Note that one can self-plagiarize. If you are using something that you wrote (e.g. a blog or a previously published article), please reference yourself.

Consequences

Plagiarism, sabotage, fabrication, and cheating carry high penalties.

Instructors may choose to fail the student on the assignment, give a 0 for the assignment, fail the student for the course, and/or bring the student before the Honor Council at which point expulsion is an option.

The worst penalty from breaking the honor code is the fact that either the person or their peers (or both) miss out on learning the material; they miss out on growth.

The obvious takeaway is:

Have a blast in this course learning the material. We learn through making mistakes and then talking through and correcting those mistakes.

Weekly Breakdown of Course

Unit	Chapter	Topic
1	1	Drawing Statistical Conclusions
2	2	Inference Using t-Distributions
3	3	Data Screening, Assumptions, and Transformations
4	4	Alternatives to the t-Tools for One, Two, and Multiple Samples
5	5	Comparisons Among Several Samples
6	6	Linear Combinations and Multiple Comparison Problem
7	1 - 6	Midterm Exam Week
8	Notes	Correlation and Scatterplots
9	7	Simple Linear Regression: Introduction
10	7	Simple Linear Regression: Prediction
11	8	Regression Diagnostics and Model Refinement
12	9	Multiple Linear Regression
13	10	Inference for Multiple Linear Regression
14	11-12	Model Selection and Validation
15	1 - 12	Review and Final Exam

Best Practices for Success in MSDS 6371 (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 class. These are considered to be disruptive (not to mention rude).

Integrity. A lot of the graded work occurs outside of class, so honesty and integrity are expected in your submissions. Evidence of academic dishonesty will minimally result in zeroes 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 your professor hasn't made a connection clear or has inadvertently left out an important point. Your question gives the professor 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, next by asking questions on the wall, and finally via e-mail to your instructor or other students. **I encourage you and expect you to seek help.** For questions during exams, please e-mail the live session instructor directly.

Collaboration. The formation of study groups and collaboration with your fellow students in tackling the assignments is encouraged. 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 detail.

Submission guidelines for assignments

- Your name must be at the top of the first page and on each successive page.

- Submit solutions in problem order.

- Use an easy-to-read variable-width font (try Ariel, Helvetica, and Geneva fonts—this document is in Helvetica 11 point) with a minimum of 11 point font.

- Relevant SAS/R code and output from the SAS/R console 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 (aka SAS DUMP). Do not include a printout of the data as well. (example, data / datalines statements in SAS.)

- Any graphics from SAS/R 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.

University Policies

- **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 Learning Outcomes:**
Please include in your syllabi all student learning outcomes, both those specific to your course, as well as those that satisfy major and general education requirements.

- **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.

- **Final Exams:**
Final course examinations shall be given in all courses where they are appropriate, and some form of final assessment is essential. Final exams or final assessments must be administered as specified in the official examination schedule, and shall not be administered during the last week of classes or during the Reading Period. Faculty must state clearly in the syllabus the date/time and form of the final exam or

assessment. All exams, tests, and quizzes will be delivered online this fall so that all students, regardless of mode of instruction, have equitable access to testing.

- **Caring Community Connections (CCC) program:**
  This is a resource for anyone in the SMU community to refer students of concern to the Office of the Dean of Students. Faculty play a critical role in identifying students who are experiencing challenges, as you may be the first to notice a change in behavior such as class attendance or performance. The online referral form can be found at smu.edu/deanofstudentsccc. After a referral is submitted, students will be contacted to discuss the concern, strategize options, and be connected to appropriate resources. Additionally, should you have concerns about students and are unclear about what to do, please see the CCC Reference Guide, or contact the Office of the Dean of Students at 214-768-4564.

- **Pregnant and Parenting Students:**
  - Accommodations for 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.

- **Covid-19 Attendance Statement:**

  Students who are experiencing COVID-19 symptoms or who have been notified through contact tracing of potential exposure and need to self-quarantine or isolate must follow the protocols laid out in SMU's Contact Tracing Protocol. To ensure academic continuity, students in these situations will not be penalized and will be provided appropriate modifications to assignments, deadlines, and testing. Please also note that SMUFlex classes might, in rare circumstances, go remote for two week periods to accommodate COVID-related issues. To ensure these necessary accommodations, affected students must:

  - Provide as much advance notification as possible to the instructor about a change in circumstances. Students must notify their instructor about a potential absence as well as plans for a return to class. For cases in which students test positive for COVID-19, they should fill out a CCC form at this link.
  - Communicate promptly with the instructor to establish, as necessary, alternative assignments and/or changes to deadlines and exams. Students are then responsible for meeting the expectations laid out in these alternative arrangements.
  - Continue participation in class via Zoom, as health circumstances permit. Attend class regularly, when not in a situation outlined above, in accordance with safety measures laid out by SMU CAN in the Pledge to Protect (including wearing masks, maintaining social distancing, and cleaning personal space after class). In-person participation in SMUFlex classes is required on students' assigned red/blue rotation days except in cases when students are experiencing illness, are in self-quarantine or in isolation.
  - Students facing multiple or extended COVID-19-related absences or illness can work with the Office of the Dean of Students to consider options such as fully remote learning or medical withdrawal.

This policy, aligned with the SMU Honor Code and the SMU Pledge to Protect, relies on mutual trust and respect between students and faculty to ensure safety, academic integrity, and instructional continuity.

CALENDAR