EDP 380C.28: Statistical Modeling & Simulation in R

Fall 2022

Unique: 12104

Instructor:	Brian T. Keller, Ph.D.	**Email:**	bk@utexas.edu
Time:	M 9:00 a.m. – 12:00 p.m.	**Room:**	SZB 4.302 (In-Person)
Office Hours:	M 1:00 p.m. – 1:50 p.m.	(SZB 5.038L)	
	Th 11:00 a.m. – 11:50 a.m.	(SZB 5.038L)	

---

https://utexas.instructure.com/courses/1346790

Prerequisites

– EDP 380C.26 Data Exploration & Visualization in R

– EDP 380C.12 Survey of Multivariate Methods

Description

Over the last few decades, computers have revolutionized statistics and our lives. In today's modern age, having a firm grasp of computer programming or "coding" will serve you well throughout the rest of your life. This course aims to provide quantitative methods students with the necessary skills needed to succeed in this modern age. This course will introduce you to the fundamentals of simulation, statistical computing, and programming in R.

In addition to these specific skills, we will discuss broader concepts, such as abstraction or functional and object-oriented programming, which will serve you outside the realm of statistical computing. You will be introduced to version control software (i.e., Git), code debugging, and code optimization. Ultimately, you will be **challenged** to think critically about coding.

Objectives

By the end of this course, you should...

– Understand how to simulate data and analyze Monte Carlo simulation studies.

– Be able to translate equations and statistical algorithms into runnable code.

– Be able to perform Array (matrices) programming in R.

– Have an understanding of fundamental programming concepts such as abstraction, functional programming, and object-oriented programing

– Be better coders!

Format

This course meets once a week for two hours and forty-five minutes in person; thus, attendance is critical. Most classes will consist of a small lecture (approximately 30 minutes), followed by an introduction to the week's lab. You are expected to stay and work in class in the lab, and I will be there to help and provide feedback.

Evaluation

Your course grade will be based on completing the course labs, a project proposal, a project outline, and the final project itself. **Note, you must obtain a satisfactory on all course labs to receive a grade. Otherwise, you will be given an <u>incomplete</u> in the class**. The breakdown of the final grade as follows.

Grade Break Down

	Percentage
Labs	45%
Project Proposal	5%
Project Outline	10%
Final Project	40%
Total	100%

Your final grade will be assigned letter grades based on the percentage of total points obtained in the course. Grades will be determined based on the following rubric (expressed in the percentage of total points).

		A	$\geq$	90
90	>	B+	$\geq$	87
87	>	B	$\geq$	84
84	>	B−	$\geq$	80
80	>	C+	$\geq$	77
77	>	C	$\geq$	74
74	>	C−	$\geq$	70
70	>	D+	$\geq$	67
67	>	D	$\geq$	64
64	>	D−	$\geq$	60
60	>	F		

Please note: To ensure fairness, all grades are absolute, and will not be rounded up or down at any stage. Thus a B− will be inclusive of all scores of greater than and equal to 80 and less than 84. The University does not recognize the grade of A+; thus in fairness to you, I will not give an A−.

Textbook

There is no required textbook for this class; however, below is a useful resource book.

Wickham, H. (2019). *Advanced R* (Second ed.). CRC press.

This book is freely available on the web and via the UT Library.

Teaching Assistant

There is no teaching assistant assigned to this course.

License

Copyright Brian T. Keller (2022) – All Rights Reserved. This course material is for your use only and cannot be distributed without explicit written consent by Brian T. Keller. Unless explicitly stated in documents or provided code, no other expressed or implied license is given.

Tentative Course Outline

Labs

Each lab will consist of a document that provides you with the requirements necessary to complete a specific topic—for example, simulating data from a linear regression. You are allowed to collaborate with one other partner on the lab, but you must provide your own completed lab. Each lab is graded on an insufficient/sufficient scale for the following four items:

1. **Completeness of the requirements and Organization:**

   – Did you satisfy the requirements given to you in the lab?
   – Does the code run without error?
   – Does the code obtain the correct results?

2. **Design and Organization:**

   – Is the code flow decomposed into manageable, logical chunks?
   – Is common code unified (i.e., not copy/pasted same chunks multiple times).

3. **Cleanliness/Readability:**

   – Did you follow the Style Guide?
   – Is the code free of clutter or unused/dead code?
   – Are you consistent with the use of whitespace, indention, and capitalization?
   – Are you breaking sections of the code with additional white space?
   – Are variable and function names consistent and descriptive?

4. **Commenting:**

   – Do you provide informative comments throughout your code?
   – This includes per function level, per logical section, and per module.
   – Are inline comments used when needed to decipher complex expressions?

You must obtain sufficient on all four categories to obtain an overall "sufficient" for the lab. Upon submitting your lab, I will grade it and provide you with feedback. If you do not receive a "sufficient," you must make the appropriate changes based on the feedback and resubmit it until you receive a "sufficient."

Project

You have two options for the final project:

1. **A small-scale Monte Carlo simulation study:**

   For the small-scale simulation, you must design, run, and analyze a small-scale Monte Carlo simulation that compares statistical methods across various conditions. This includes performing all the following in R: Implementing the data generation procedure, analyzing the data, and analyzing the simulation. The organization of all code and workspace must adhere to the guidelines discussed throughout the course labs (see points above). In addition, you must provide a written document that discusses the simulation conditions, the data generating model you used, how parameters were determined for the simulation, and a brief discussion of the results (tables and graphs are allowed). The number of conditions and replications highly depends on the type of statistical model. As a good rule of thumb, the simulation run time should be approximately something you can run during a weekend on your computer.

2. **A small documented R package:**

   For a small R package, you must design, implement, and test all functions and classes to achieve a specific purpose. In addition, you must provide complete help documentation for all exported R functions/classes and a brief vignette demonstrating how to use your package (see examples in R packages on CRAN). All code implementation, code documentation, and organization of the package must adhere to the guidelines discussed throughout the course labs (see points above). The number of specific functions and objects that must be exported and documented for the package will vary depending on the complexity of the task you choose. As a good rule of thumb, you must create one S4 class object, export five different methods, and five other functions.

Note that this project must be something new that you have not done before (i.e., if you are currently going through QP or dissertating, select something different). Regardless of the option, you must first provide an initial proposal and a project outline.

Proposal

The proposal is approximately half a page (no more than one page) and informs me of two things. First, which option you are selecting to do. Second, a statement of purpose for your project. This statement of purpose gives me a broad overview of the project and why you want to do it. Please do not give me specifics (i.e., which conditions you will investigate in the simulation). Said differently, think of this as an abstract, a summary of what you envision. Once I review your initial proposal, I will give you the okay to move forward.

Outline

Once I've approved your project proposal, you can begin on the project outline. The substance of the project outline depends on which option you choose, and a separate document offers more details, but generally, it provides a much more detailed structure of your project. For a simulation, this includes (but is not limited to) the data generating model, a discussion of conditions that you will be manipulating, the number of replications, how you will estimate the resulting statistical models (I will allow the use of other packages and software), and what measures you will use to

evaluate the simulation (e.g., bias, coverage, mean-square error). For an R package, you will need to outline the purpose of the package and provide a detailed skeleton of the published functions, classes, and methods you will include without the actual implementations.

Submission and Due Dates

All submissions will be made under the Assignments section in the course canvas. The specific directions to submit are on each assignment page, and the due dates and times are included as part of the assignment. Each assignment closes at the end of the date, and you will no longer be able to submit the assignment. If you have trouble completing the assignment in time, please contact me, and we can discuss options. Below are the due dates for your reference.

Assignment	Due Date
Proposal	September 12, 2022
Outline	October 14, 2022
Project	December 2, 2022

General Policies

COVID-19

The university offers the following website for up to date information about COVID 19 policies.

<p style="text-align:center">https://protect.utexas.edu</p>

Furthermore, the Family Educational Rights and Privacy Act (FERPA), prohibits me from divulging information about a student testing positive in the class if I am made aware of it. Therefore, please do not ask me any questions related to such a topic. Instead, the appropriate information will be provided to you via communications from the University Health Services (UHS). If a student does test positive and has been on campus in a class with other students, UHS states that those students will receive two levels of communication from UHS—a general notice that one of their classmates tested positive, and a directed notice if they were identified as a possible close contact. The information in these notices ought to be quite specific on next steps. Finally, if I am required to quarantine at any time during the semester, instruction will continue to be taught remotely via Zoom until I am allowed to return to campus.

Statement on Learning Success

Your success in this class is important to me. We will all need accommodations because we all learn differently. If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we'll develop strategies to meet both your needs and the requirements of the course. I also encourage you to reach out to the student resources available through UT. Many are listed on this syllabus, but I am happy to connect you with a person or Center if you would like.

Student Rights & Responsibilities

– You have a right to a learning environment that supports mental and physical wellness.

– You have a right to respect.

– You have a right to be assessed and graded fairly.

– You have a right to freedom of opinion and expression.

– You have a right to privacy and confidentiality.

– You have a right to meaningful and equal participation, to self-organize groups to improve your learning environment.

– You have a right to learn in an environment that is welcoming to all people. No student shall be isolated, excluded or diminished in any way.

With these rights come responsibilities:

– You are responsible for taking care of yourself, managing your time, and communicating with the teaching team and with others if things start to feel out of control or overwhelming.

– You are responsible for acting in a way that is worthy of respect and always respectful of others.

- – Your experience with this course is directly related to the quality of the energy that you bring to it, and your energy shapes the quality of your peers' experiences.

- – You are responsible for creating an inclusive environment and for speaking up when someone is excluded.

- – You are responsible for holding yourself accountable to these standards, holding each other to these standards, and holding the teaching team accountable as well.

Personal Pronoun Use (She / He / They / Ze / Etc)

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. Class rosters are provided to the instructor with the student's legal name, unless they have added a "preferred name" with the Gender and Sexuality Center. I will gladly honor your request to address you by a name that is different from what appears on the official roster, and by the gender pronouns you use (she/he/they/ze, etc). Please advise me of any changes early in the semester so that I may make appropriate updates to my records.

Services for Students with Disabilities

This class respects and welcomes students of all backgrounds, identities, and abilities. If there are circumstances that make our learning environment and activities difficult, if you have medical information that you need to share with me, or if you need specific arrangements in case the building needs to be evacuated, please let me know. I am committed to creating an effective learning environment for all students, but I can only do so if you discuss your needs with me as early as possible. I promise to maintain the confidentiality of these discussions. Any student with a documented disability who requires academic accommodations should contact Services for Students with Disabilities at 471-6259 (voice) or 512-410-6644 (Video Phone) as soon as possible to request an official letter outlining authorized accommodations. For more information, visit

<div align="center">

http://ddce.utexas.edu/disability/about/.

</div>

Counseling and Mental Health Center

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support.

<div align="center">

http://www.cmhc.utexas.edu/individualcounseling.html

</div>

Title IX Reporting

Title IX is a federal law that protects against sex and gender-based discrimination, sexual harassment, sexual assault, sexual misconduct, dating/domestic violence and stalking at federally funded

educational institutions. UT Austin is committed to fostering a learning and working environment free from discrimination in all its forms. When sexual misconduct occurs in our community, the university can:

1. Intervene to prevent harmful behavior from continuing or escalating.

2. Provide support and remedies to students and employees who have experienced harm or have become involved in a Title IX investigation.

3. Investigate and discipline violations of the university's relevant policies.

Beginning January 1, 2020, Texas Senate Bill 212 requires all employees of Texas universities, including faculty, report any information to the Title IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations) must be reported. I am a Responsible Employee and must report any Title IX related incidents that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee. If you would like to speak with someone who can provide support or remedies without making an official report to the university, please email advocate@austin.utexas.edu. For more information about reporting options and resources, visit http://www.titleix.utexas.edu/, contact the Title IX Office via email at titleix@austin.utexas.edu, or call 512-471-0419.

Although graduate teaching and research assistants are not subject to Texas Senate Bill 212, they are still mandatory reporters under Federal Title IX laws and are required to report a wide range of behaviors we refer to as sexual misconduct, including the types of sexual misconduct covered under Texas Senate Bill 212. The Title IX office has developed supportive ways to respond to a survivor and compiled campus resources to support survivors.

Academic Integrity

Each student in the course is expected to abide by the University of Texas Honor Code:

> *As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity.*

Plagiarism is taken very seriously at UT. Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Otherwise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT's Academic Honesty and the University Honor Code which can be found at the following web address:

https://deanofstudents.utexas.edu/conduct/standardsofconduct.php