STAT 155: Introduction to Statistical Modeling

General Education Designation: Quantitative Thinking (Q3)

Professor: Taylor Okonek (she/her)

Email: tokonek@macalester.edu

Office: Olin Rice 126

Section	Meeting Time	Room
01	MWF 3:30 - 4:30 pm	OLRI 241

Office Hours:

- Mondays 11:00 am - 12:00 pm
- Tuesdays 4:00 - 5:00 pm
- Thursdays 2:00 - 3:00 pm

Note: Office hours are typically first-come, first-served. However, I will do my best to include everyone present in the conversation. If the office hour times listed above don't work for you, I am **available by appointment**; email me to set up a time to meet. More expectations on communication are in the Course Communication section.

Preceptors

We have many preceptors for STAT 155 this semester! You may attend office hours for any of the STAT 155 preceptors, not just those grading/attending our sections (Ryan). The course Moodle site contains detailed information for preceptor office hour times/locations.

The role of an MSCS preceptor is to help students with content questions, assist in the navigation of available resources, advise on studying approaches for classes, and assist with concepts, tools, and skills needed for problem sets. Students are accountable for their own learning; as such, preceptors are not allowed to share answers to assignments (unless specifically directed by the instructor), are not expected to immediately know the right approach, or provide assistance outside of office hours.

Additional guidelines and expectations on how to interact with preceptors can be here.

Textbook

There is no required textbook for this course. Throughout the course, readings may be assigned from these notes, or other sources. All links and materials needed can be found in the Schedule tab of this website.

The best resources for this course are course content videos, attending and participating in class / assignments, and office hours. I am happy to talk about possible additional materials / strategies for effective learning at any time throughout the semester!

Learning Objectives

Upon completion of this course, students should be able to...

- Use (and interpret) graphical and numerical summaries of data.
- Identify the appropriate statistical model to answer a research question, fit the model in R, and interpret the relevant resulting output from these models in context of the scientific question at hand.
- Quantify and communicate uncertainty in regression model coefficient estimates.
- Conduct, and correctly interpret, hypothesis tests for regression models.
- Evaluate regression models using fit metrics such as correlation, R^2, and adjusted R^2.
- Articulate the role that probability and conditional probability play in statistical inference.

- Describe potential advantages, limitations, and ethical considerations of a data set and statistical analysis.
- Identify common pitfalls in statistics (e.g., spurious correlation vs. causal relationships, extrapolation, multicollinearity, multiple testing, practical vs. statistical significance).

Tentative Schedule

A detailed, day-to-day schedule will be filled in throughout the semester in the Schedule tab of this website. A rough outline by week number is as follows:

1. Collecting and Summarizing Data
2. Univariate Visualization, Simple Linear Regression
3. Simple Linear Regression: Model Evaluation & Transformations
4. Simple Linear Regression: Categorical Predictors
5. Multiple Linear Regression
6. Confounding, Interaction Terms, **MSCS Capstone Days**
7. Multiple Linear Regression: Model Building
8. Probability/Odds, Simple Logistic Regression

** Spring Break**

9. Multiple Logistic Regression
10. Tools for Inference
11. Confidence Intervals, Hypothesis Testing
12. Hypothesis Testing: F-Tests, P-values
13. Review / Project work time
14. Review / Project work time
15. **Last Day of Class**

Course Communication

Course Website: All course materials will be posted on the course website. The Schedule tab of this website will contain information about due dates, in-class activities, and general day-to-day going-ons.

Moodle: Moodle will be used to post grades, assignment feedback, announcements, and more. Please check the course Moodle page every day before class, to make sure you aren't missing any assignments!

Email: Outside of office hours, email is the best way to discuss topics of a personal nature (e.g., grades, missing class) or to ask other questions you don't feel comfortable asking in a group setting. **Note that I will not send same-day responses to messages sent after 6:00pm.** You can email me about an extension day after 6:00pm, I will respond the next day.

Please communicate respectfully with your classmates and myself, and let me know if there are ways classroom communication can be made more accessible to you.

Classroom Environment

I am committed to helping you learn and succeed. Part of this means facilitating environments that support you in the ways that you need. Everyone comes from a different path through life, and it is our duty to listen to each other without judgment, and to respect one another. There will be **no tolerance** for discrimination based on race, ethnicity, gender, religion, sexual orientation, disability, or other identities.

Statistics is used to understand the world around us as it pertains to many relevant and timely issues. As such, we may sometimes address topics that are sensitive for you. This is likely to be truer than ever given what has transpired in our world recently. If you feel that you may not be able to complete a certain activity, do a reading, or partake in a discussion due to the nature of the topic at hand, please reach out to me so that you can help me understand your situation, and so that we can discuss the possibility of replacing the item with an alternative. I am more than happy to work with each of you to ensure that you feel comfortable in your learning.

Grades

Evaluations / Assessments

Quizzes: There will be three quizzes throughout the course of the semester.

- Quiz 1: Friday 2/14. 1 hour in class.
- Quiz 2: Friday 3/28. 1 hour in class. Will cover material from the first quiz to some extent because of the way that material in this course builds on earlier ideas.
- Quiz 3: Saturday 5/10, 1:30 - 3:30 pm (during the final exam period). The exam will be written to take ~ 1.25 hours, but you will have the full two hour period to complete it. Will cover content from the whole course.

Quiz format:

- Fully pen/pencil and paper. You will **not** need to write code or use a calculator, but you **will** need to be able to read and interpret output from R code.
- You are allowed to bring a 3x5 index card with notes written on both sides. Typing your notes and pasting them on the card is also fine.

Showing growth: You can earn up to 50% of missed points back on the first two quizzes if you complete a quiz correction and reflection. You must:

- Write a reflection of how you prepared for the quiz and where you felt strongest and more uncertain in your understanding before taking the quiz.
- Submit your quiz corrections along with your reflection to the instructor, no later than one week after quizzes have been handed back.

You cannot do this for the third quiz, due to time constraints at the end of the semester.

Practice Problems: There will be 8 practice problems throughout the course, building off of topics learned in videos/our in-class activities. Practice problems will be due on Wednesdays at 5pm. You are encouraged to work together on practice problems, but **the work that you submit, including code, should be your own**.

You will receive qualitative feedback on all questions on these practice problem sets. Each *question* (or occasionally, parts of a question) will receive an overall score on the following scale:

- High pass: 2 points
- Pass: 1 point
- Needs improvement: 0 points

Thus, the number of overall points on a given Practice Problem may vary depending on the length of the assignment. **Your lower practice problem score will be dropped from your final grade, so long as you have submitted all 8 practice problems and demonstrated effort on each one.**

Checkpoints: Some class periods will have course videos and readings assigned ahead of time. For each class period where this is the case, a checkpoint quiz (on Moodle) must be completed *30 minutes before class starts*. Checkpoints may be attempted as many times as you'd like, but to earn completion credit for a given checkpoint you *must* score 100% by your final attempt.

These short quizzes are designed to ensure that you stay on top of course material, since much of the content in this course builds on itself.

Group Project: Information on the group project can be found on the Project tab.

Late Policy

Throughout the quarter, you may use up to **three**, three-day extensions. These three extensions can be used on Problem Sets *only*, not quizzes. The purpose of deadlines (and extensions) are to keep you accountable for your own learning, to keep you on track with the pace of the course (which builds upon itself throughout the semester), and to provide preceptors and myself with the ability to provide you with timely feedback on assignments. Since the Problem Sets are due every two weeks, you *must* begin working on them early if you want to succeed.

Extensions can be used automatically, without letting me know in advance. The Moodle dropboxes for assignments will close exactly 3 days after the original deadline, and I will not accept work submitted after that point unless there are extenuating circumstances that you have communicated with me about *ahead of the original deadline*. If you email me a completed assignment after a 3-day extension is up, I may have the preceptors provide you with feedback, but you will **not receive credit** for the assignment (equivalent to "Needs Improvement" on every question of the relevant assignment).

I expect you to keep track of how many extensions you've used. I will email you a reminder if you have used all three of your extensions and have none remaining.

If you have run out of extensions and/or an extenuating circumstance occurs that impacts your ability to submit assignments on time, please email me to discuss the situation. I am happy to be flexible as long as you communicate!

Academic Integrity

As a globally- and community-oriented institution, Macalester College expects respectful exchange of ideas. Students are expected to be familiar with the college standards on academic integrity (link). You are encouraged to work with your classmates on assignments and activities, but in order to receive individualized feedback on your own learning, **all submitted work (including code) must be written in your own words**. Issues of academic integrity will be taken very seriously, and any concerns about this policy being violated will be followed up on individually.

Course Grading System

Your overall course grade will consist of three, evenly-weighted components (Quizzes, Practice Problems, and the Project), **modified by** your Checkpoint Quiz completion.

- 33% Practice Problems
- 33% Quizzes
- 33% Project

Overall percentages will correspond to the following letter grades:

Letter Grade	Overall Course Percentage
A	>90%
A-	87-90%
B+	84-87%
B	81-84%
B-	78-81%
C+	75-78%
C	72-75%
C-	69-72%
D	<69%

Note that the cutoffs for letter grades above may be altered at the end of the semester to be *lower*, but will not be moved higher than what is listed above (i.e. they would only potentially move in your favor).

Checkpoint quizzes are graded for completion. If you complete $\geq 80\%$ of the checkpoint quizzes before the time they are due, your overall course grade will **not** be negatively affected. If you complete *less* than 80% of the checkpoint quizzes before the time they are due, your overall course grade will be lowered by 1/3 of a letter grade (i.e. B –> B-, A- –> B+, etc.).

Statistical Software

We will use the free, R programming language extensively throughout this course. RStudio (an interface for R) will facilitate our use of R. You may use RStudio in one of two ways:

1. Online: Go to https://rstudio.macalester.edu, and log in with your full Mac email address and your usual Mac password to get access
2. Desktop version: Download for Windows or Mac at https://posit.co/downloads/. **Note**: You first need to download and install R on your computer in order to use the desktop version of RStudio

More detailed instructions on downloading, installing, and getting started with R, RStudio, and an important package called Quarto is available on the Tech Setup tab.

Accommodations

I am committed to ensuring access to course content for all students, including those with disabilities. If you have a disability, please meet with me early in the semester to discuss your accommodation plan. If you have not yet obtained a plan or are unsure if you have a disability that requires accommodation, please contact Disability Services: disabilityservices@macalester.edu, or call 651-696-6275.

Religious Observance

Students may wish to take part in religious observances that occur during the semester. If you have a religious observance/practice that conflicts with your participation in the course, please contact me before the end of the second week of the semester to discuss appropriate accommodations.

In an effort to respect religious diversity, I request that students who plan to observe a religious holiday during scheduled class meetings/ class requirements talk to me about reasonable consideration by the end of the second week of the course.

Title IX

Macalester is committed to providing a safe and open learning and living environment for all students, staff, and faculty. Any community member experiencing sexual harassment, sexual violence, relationship violence, or stalking, is encouraged to seek help and support.

Please be aware that as a faculty member, it is my responsibility to report disclosure about sexual harassment, sexual misconduct, relationship violence, and stalking to the Title IX Office. The purpose of this report is to ensure that anyone experiencing harm receives the resources and support they need. I will keep this information private, and it will not be shared beyond this required report.

You may also contact Macalester's Title IX Coordinator directly (phone: 651-696-6258; e-mail: titleixcordinator@macalester.edu); she will provide you with supportive measures, resources, and referrals. Additional information about how to file a report (including anonymously) is available on theTitle IX website.

If you prefer to speak with someone confidentially, or need 24/7 support, there are resources available on- and off-campus to assist you:

- Counseling Services at the Hamre Center – on-campus counseling resource for students **Free, Urgent, Phone Counseling (Press 2) is available to Macalester students anywhere in the world, 24/7/365.** Speak to a licensed mental health counselor 24 hours a day by calling Hamre Center at 651-696-6275, then press or say option 2 when prompted.
- Center for Religious and Spiritual Life – chaplains may be reached via 651-696-6298 or email religiouslife@macalester.edu
- SOS Sexual Violence Services of Ramsey County 651-266-1000 (24-Hour Crisis Hotline staffed by trained advocates)
- Rape, Assault, and Incest National Network (RAINN) 1-800-656-4673 (24-Hour Hotline and Live Chat)
- Additional local and national resources can be found athttps://www.macalester.edu/title-ix/find-support/

Land Acknowledgement

I would like to honor the fact that we are on Dakota land. This is the ancestral homeland of the Dakota people who were forcibly exiled from the land because of aggressive and persistent settler colonialism. I make this acknowledgment to honor the Dakota people, ancestors, and descendants, as well as the land itself.

About Your Professor

Hi! Please call me Taylor or Professor Okonek (never Dr. Okonek, it sounds weird to me). My pronouns are she/her/hers. I was born on Madeline Island, Wisconsin, and grew up partially there / partially in White Bear Lake, Minnesota. I went to St. Olaf College for undergrad, and double majored in Mathematics and Religion, with a concentration in Statistics. I am a first-generation college student from a low-income background and was a part of the TRIO McNair program at St. Olaf. I completed my Ph.D. in Biostatistics at the University of Washington, Seattle in 2023. My research interests are in demography, mortality estimation, spatial and survey statistics, and survival analysis (and their intersections!).

When I'm not thinking about / doing Statistics, I like to rock climb, hike, cook, read, play video games (mostly Path of Exile currently, and Celeste is my favorite game of all time), and hang out with my two cats, Alice and Ellie. I'm always happy to talk about research if you have any questions, and happy to chat about your interests/goals as well, be they statistics-adjacent or not!

Note

This syllabus is subject to change at any time! Announcements of changes will be made in class or via email.