STAT 155: Introduction to Statistical Modeling

Macalester College
Fall 2022
Section 02: TTh 8:00 am - 9:30 am, Olin-Rice 254
Section 03: TTh 9:40 am - 11:10 am, Olin-Rice 254
Section 04: TTh 1:20 pm - 2:50 pm, Olin-Rice 254

Course Overview

This course provides an introduction to statistical modeling. In other words, it's a crash course on how to turn *data* into *knowledge*. No matter your career path, being able to summarize, interpret, and communicate about data will be crucial, and these are precisely the skills that we'll build in this class. Throughout the semester we'll study the fundamental methods that statisticians use to extract knowledge from data, emphasizing statistical literacy, real data applications, and modern computing over memorizing facts and formulas. In particular, we will:

- **Visualize:** explore, understand, and communicate patterns in data with plots and graphics
- **Model:** select, fit, interpret, and evaluate statistical models to explore and summarize trends in data
- **Infer:** extend observations from sample data to draw conclusions about broader populations
- **Contextualize:** interpret results in context, by considering the methods of data collection, the scientific context, and ethical issues
- **Collaborate:** work productively and effectively in a group setting

Your Instructor

James Normington (he/him)

Office: Olin-Rice 230

Email: jnorming@macalester.edu

Office Hours

- M: 3-4pm
- TTh: 3-4pm
- Th: 12:15pm - 1:15pm
- Location: Olin-Rice 230
- By appointment – please don't hesitate to use this option!
- Scheduled office hours will be done in person. However, given concerns over the COVID-19 pandemic, I'm more than happy meeting by Zoom. Feel free to email me to set this up.

Learning Objectives

By actively participating in this course, you will be able to:

- **Choose**, and **communicate knowledge gained from**, contextually appropriate **statistical summaries** and **models**
  - Identify and describe relevant details of data context
  - Select appropriate summaries (graphical & numerical) and models to answer a scientific question
  - Accurately and effectively interpret data visualizations, numerical summaries, and model coefficients
  - Recognize, quantify, and communicate uncertainty in regression model coefficient estimates
  - Conduct, and correctly interpret, hypothesis tests for regression models
  - Describe potential advantages, limitations, and ethical considerations of a dataset and analysis
- **Use R**, a (free!) programming language and statistical software environment, to **analyze real data**
  - Load and prepare data for analysis
  - Create numerical and graphical summaries of data
  - Fit, perform inference on, and evaluate linear and logistic regression models
  - Write professional, reproducible data analysis reports

Course Materials

The only required material in this course is access to a computer. If you do not have a functional computer, or reliable internet access, don't worry! Just fill out this form; ITS and Financial Aid will work with you to get you what you need.

Every reading will be made freely available online.

Our main text will be:

Stat 155 Notes. Heggeseth, Myint, & Grinde, 2021. [Link]

Additional online resources include:

Modern Dive. Version 1.1.0. Ismay and Kim. 2020. [Link]

OpenIntro Statistics. 4th Edition. Diez, Çetinkaya-Rundel, Barr, 2019. [Link]

RStudio Primers. [Link]

Any others will be accessible, for free, in Moodle.

Topics Schedule

We will tentatively be covering topics as follows:

Week 1: Data principles

Week 2: Data visualization

Weeks 3–6: Linear regression

Weeks 7–9: Logistic regression

Weeks 10–11: Quantifying uncertainty

Weeks 12–13: Hypothesis testing

Week 14: Model evaluation and selection

Week 15: Wrapping up/project work time

Grading Scheme

The grading system in this course is designed to directly ***align your grades with*** our course ***learning*** objectives, and to allow space to ***make and learn from mistakes*** along the way. More specifically, grades will be based on a combination of three components:

(1) Final Project
(2) Quiz: Your performance on three quizzes
(3) Engagement: Demonstrated engagement in course assignments, in-class activities, and final project

See the Grading Rubric for a detailed explanation of the grading calculation.

Major Assessments

Quizzes.

To assess your understanding of course learning objectives, there will be three quizzes spaced throughout the semester. They are (tentatively) scheduled for:

- Quiz 1: Tuesday, October 4
- Quiz 2: Thursday, October 27
- Quiz 3: Tuesday, November 22

Quizzes will be completed during class time, using up to one (8.5 x 11 inch) handwritten page of notes. Each quiz will involve an individual portion (30–40 min), as well as a group portion (15 min). *If you are unable to take a quiz during the scheduled time period, please let me know at least one week in advance* so I can work with you to make alternate arrangements.

Group Data Analysis Project.

You will work on a data analysis project in small groups throughout the semester. This project will ask you to use the skills practiced in class to explore, analyze, and interpret summaries and models fit to a dataset of your choosing. There will be multiple project checkpoints throughout the semester, allowing ample opportunities for instructor feedback. This work will culminate in a **final report** due on the last day of class and a **final presentation** during your final exam period:

- **All Sections Final Report Due: Thursday, December 8**
- **Final Presentations:**
  - Section 02: Friday, December 16, 8am - 10am
  - Section 03: Thursday, December 15, 10:30 am – 12:30 pm
  - Section 04: Thursday, December 15, 1:30pm - 3:30pm

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 (**website**). I encourage you to work with your classmates to discuss material and ideas for your assignments, but in order for you to receive individualized feedback on your own learning, ***all submitted work (including code!) must be written in your own words.***

When you present someone else's work as your own, you are missing a learning opportunity and cheapening the value of your education. If you are tempted to plagiarize someone's work, please consider requesting an extension instead. See below my policy on extensions.

Other Course Policies

Extensions. Life happens! There are circumstances outside of the classroom that make it difficult to focus on schoolwork. These circumstances are often entirely personal; thus, I will grant most extensions, without the need to specify a reason, when requested in advance of the corresponding deadline. My intent is that students can focus on their physical, mental, personal, and interpersonal wellness so there are fewer obstacles to learning. However, I reserve the right to deny an extension if I feel the student is abusing this policy.

Final grades are due Dec. 28th; any unsubmitted work, regardless of circumstances, will count as 0. If your circumstances make it likely that you can't submit a sizeable portion of classwork by then, you should talk with me about alternative arrangements. For deadlines on adding/dropping courses, grading options, and withdrawal, see the Fall 2022 academic calendar here.

Attendance. Attendance is expected, but not mandatory. If you are to miss class, please try to inform me in advance. Your reason for missing class does not need to be included in your message. I reserve the right to adjust the Engagement component of a student's final grade if the student makes a habit of missing class without informing me in advance, or if I suspect a student is abusing this policy.

Late Submissions. Homework submitted late, but within 7 days of the due date, will be assessed a 50% penalty. Homework submitted later than 7 days of the due date will be assessed a 100% penalty.

Course Structure

This class will operate under a "flipped classroom" model. A typical class will proceed as such:

Before Class.

- Complete assigned readings/videos.
- Complete Moodle checkpoint to assess understanding of readings/videos.
  - Unlimited attempts
  - Only highest score recorded

During Class.

- Short lecture (no more than 30 minutes)
- Group activities
  - **Please bring your laptop to class every day.**
  - Google doc to record responses, and to use as a reference for R code

After Class.

- Complete any unfinished parts of in-class activity
- **Most (not all) Fridays**, you will submit a set of practice problems from the week's practice problems and/or applications to your Final Project data
- Prepare for next class session (see "Before Class")
- Work on Final Project; there will be several checkpoints throughout the semester.

My Teaching Philosophy

Active learning. Statistics is an active discipline. The collection, visualization, analysis, and interpretation of data are all active steps in the statistical modeling process. Because of this, a large proportion of class time will be spent actively doing statistics problems with practical, real-world data.

Statistical thinking over statistical procedure. The rote memorization of manual statistical procedures (i) is unnecessary with modern computing, (ii) requires mathematical prerequisites beyond the scope of this course, and (iii) presents statistical techniques as disparate, non-intuitive procedures. Statistical modeling is best learned through intuitively solving real problems, and leaving the tedium of the procedure up to computers (or, for the inclined, preparation for graduate school exams!).

Diversity, equity, and inclusion. Statistics is, in a word, the study of variability. The best statistical models acknowledge and account for all sources of variability, recognizing the diverse and multivariate world we live in. There will be absolutely no discrimination based on race, ethnicity, gender identity, sexual orientation, or any other identities, life experiences, or opinions. If at any point you feel discriminated against, please either reach out to me or find other support at Macalester.

Statistical Software

Computing is an integral part of modern statistics and will thus be a main focus of our course. We will use the R programming language. RStudio, a graphical user interface (GUI) for R, will facilitate our use of R and the creation of R Markdown reports that you will use for submitting assignments. R and RStudio are *free* to use and download.

No computing background is necessary to succeed in this course. When you learn a new computing language, you will stumble at the beginning. That is normal! The key is to learn how to overcome those stumbling blocks (or coding errors). Read the error message. Double check the spelling. Consult Google. Post the code and error message on Slack.

Learning to use R/RStudio will take time and effort, but by the end of this semester you'll be leaving this course with a *very marketable and useful skill* that will serve you well in future courses, jobs, and beyond!

Advice for Success in STAT 155

1. **Actively engage with the course content.** Pay attention in lectures, actively participate during in-class activities, understanding and incorporate my feedback into your Final Project. For Quizzes, do practice problems and understand their solution! If you run out of practice problems, reach out to me.
2. **Ask when you have questions.** Ask during lecture (interrupt me at any time; I'm serious!), come to office hours, visit the preceptors and tutors, ask your classmates, post on Slack. Our best learning comes in the vulnerable moments when we admit we don't understand.

COVID-19 Protocols and Expectations

We all deserve a safe learning environment. Please stay current with the latest COVID-19 guidelines: Shared Community Commitment (Mac Stays Safer 3.0). I will abide with these guidelines very closely.

Some key notes, and my specific expectations:

Vaccination Policy. Macalester requires all employees and students to have received a COVID-19 vaccine and booster, or to have an approved exemption in place. This is not negotiable. If you are eligible for a 2nd booster, I highly recommend it; see more information at the CDC website.

Mask Policy. Masks are required in all public indoor spaces on campus through Wednesday, Sept 14th. After Sept 14th, this requirement may be extended. In the event there is no masking requirement later this semester, I will happily wear a mask in office hours if you would like me to – please don't be shy to ask. During in-class activities and Final Project working sessions, groups should feel free to come up with their own policies (so long as they fit within Macalester's and the State of Minnesota's policies) If there is a disagreement which is harming the group's ability to make progress, please see me. For more information, see the Well-Fitted Mask Policy.

Be Safe, Be Respectful. My primary concern is your well-being. If I or a classmate respectfully asks you to wear a mask, cover your sneeze, wash your hands, etc. – please do so. Your classmates may be immunocompromised, or live/work/relate with someone who is. They may also just not want to get sick! Every student has the right to feel safe attending my lectures in person. We each have varying levels of "COVID cautiousness", so please use both common sense and courtesy for others in my classroom and on campus.

Test Frequently. If you feel symptoms associated with COVID-19, **please** get tested. Please see Shared Community Commitment (Mac Stays Safer 3.0) for information on free, accessible testing.

Stay Home when Sick. If you have COVID-19, or any other contagious disease, **please** don't come to class. Email me and we will work out alternate arrangements.

Health and Wellbeing

Your health and well-being should be prioritized over my course. A student cannot effectively learn statistics if they are under physical, mental, emotional, societal, or financial stress. During class time, eat when you are hungry, drink water when you are thirsty, use the restroom, and step away/out if you are upset or need some air. Please do what is necessary so long as it does not impede your or others' ability to be mentally and emotionally present in the course.

Physical Health. Physical exercise is associated with better grades, better physical & mental health, and an improved quality of life. I highly recommend using the Leonard Center to stay physically active. Participation in intramural sports is another great way to exercise, reduce stress, and develop a sense of community on campus.

Mental Health. You can find a list of mental health resources here.

Financial Health. If you have had an unanticipated financial emergency, please consider using the Emergency Aid Program.

Beyond these resources, if you are having difficulties maintaining your well-being, please contact me. I am invested in your full learning journey, which cannot thrive unless you are OK.

The Environment You Deserve

I am committed to helping you learn and succeed in this course and in your time at Macalester. Enabling this means facilitating environments that support you in the ways that you need.

Respect. There will be *no tolerance for discrimination* based on race, ethnicity, gender identity, religion, sexual orientation, disability, and other identities or life experiences, in and out of the classroom. If you see any discrimination whatsoever, please contact me or find other support at Macalester, and be ensured the matter will be dealt with seriously.

Sensitive Topics. Statistics is used to inform policy, research, and understanding across all aspects of our society. The examples and data we use in class will reflect this. Thus, at times our work might touch on sensitive topics. I will try to give warnings for topics that I perceive may be sensitive. Please let me know if, for any reason, you feel that you are unable to participate in the planned activities due to the topic and we can discuss alternatives.

Accommodations. I am committed to creating an accessible and inclusive classroom for all students, including those with disabilities. If you are in need of any accommodations, please contact Disability Services (visit their website, email disabilityservices@macalester.edu, or call 651-696-6275) to schedule an appointment and discuss your needs. Once you've met with Disability Services, please then set a time to meet with me to discuss your accommodation plan for this course. It is important to *arrange this meeting as early in the semester as possible* (ideally within the first week), in order to ensure that your accommodations can be implemented early on. It is your responsibility to make sure you are registered with Disability Services. If you wait until later in the course, I may not be able to accommodate you retroactively.

Religious Observance. Students may wish to take part in religious observances that occur during this semester. I've done my best to schedule deadlines around major holidays, but if you have a religious observance/practice that conflicts with a scheduled class meeting or assignment deadline, *please contact me at least one week in advance* so we can discuss appropriate accommodations.

Title IX. If you or anyone you know has experienced harassment or discrimination on the basis of sex or gender, know that you are not alone. Macalester provides staff and resources to help and support you. More information is available on the Title IX website.

Please be aware that *all Macalester faculty and preceptors are mandatory reporters*, which means that if we become aware of incidents or allegations of sexual misconduct, we are required to share the matter with the Title IX Coordinator. Although we have to make that notification, you control how your case is handled, including whether or not you wish to pursue a formal complaint. If you would like to speak to someone *confidentially*, contact the Hamre Center (651-696-6275), Chaplain staff (651-696-6298), or other local and national resources listed here.

Preceptors

We have eight preceptors that will be helping out with STAT 155 this semester. They will be providing feedback on your weekly practice problems and holding office hours.

Aidan Appel
he/him
aappel@macalester.edu

Ty Bruckner
he/him
tbruckne@macalester.edu

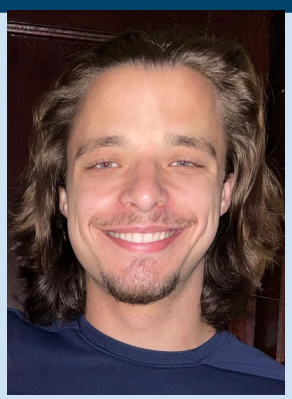
Nicholas Di
he/him/his
ndi@macalester.edu

Ellery Island
she/her
eisland@macalester.edu

Alayna Johnson
she/her
ajohns38@macalester.edu

Valeska Fresquet Kohan
she/hers
vkohan@macalester.edu

Christian Lentz
he/him/his
clentz@macalester.edu



Marshall Roll
he/him
mroll@macalester.edu

MAX Center: Peer Tutoring

The Macalester Academic Excellence (MAX) Center provides tutoring for a variety of math classes, including STAT 155. Check out the schedule for drop in tutoring (no appointment needed).

Q3 General Education Requirement

Successful completion of this course will satisfy the ***Q3 Quantitative Thinking*** general education requirement. You can find details on this requirement here. In particular, in this course you will:

- develop skills in the areas of data collection, data visualization, and data analysis
- practice turning a scientific research question into a statistical one, which can then be investigated using exploratory data analysis and statistical modeling
- think critically about data context
- learn to recognize and acknowledge the limitations of statistical methods and data itself