

SC321: Statistical Modeling

Spring 2024

Professor:	Jerzy Wieczorek
Office:	Lovejoy 110
Telephone:	(207) 859-5841
Email:	jawieczo@colby.edu
Office hours:	See our Moodle page
Lectures:	Mon / Weds / Fri, 12:00–12:50pm, Davis 308
Assignments:	Homeworks are due on Friday at 4:00pm, unless otherwise instructed.
Exams:	There will be two midterm exams, each consisting of two parts, tentatively on Weds + Fri, March 20 + 22 and on Weds + Fri, May 8 + 10 .
Data Analysis Reports:	There will be 3 or 4 reports (individually and in groups). In lieu of a final exam, the last report will be due TBD during finals week, May 15-20 .
Prerequisites:	SC212 or consent of the instructor.
Textbook:	“STAT2: Modeling with Regression and ANOVA” by Cannon, Cobb, et al. (Second Edition —has a rainbow on cover)

Course Description

Statistics is at the center of advances in science in the 21st century. Almost every area of modern science depends on statistics in some form. It serves not only the sciences, but the field’s influence has grown to affect all aspects of our lives from health to business to public policy to entertainment. Thanks to advances in technology, we now live in a world inundated with data. How can we make use of all this data to yield useful information? Statistical modeling provides a way of making sense of data—a way to gain insight about processes and populations. Extracting useful information requires more than learning a laundry list of methods; it requires careful consideration of context. What is the question of interest and what theory motivates the investigation? How was the study designed? How were the data collected? What can ultimately be learned from the analysis—that is, what is the scope of inference?

In this course, we will explore several flexible and powerful approaches to statistical modeling, including simple linear regression, multiple regression, model diagnostics, analysis of variance (both simple and multifactor), and logistic regression for binary response variables and binomial counts.

Course Objectives

By the end of the semester:

- Students will be able to extend linear regression concepts learned in an introductory statistics course to a setting where there is more than one predictor variable.

- Students will learn how to formulate research questions using statistical models.
- Students will understand the implications of bias, variation, confounding, and interaction in the context of statistical inference.
- Students will learn how to check regression model assumptions and assess the fit of different types of models.
- Students will be able to model some types of data that violate the standard linear regression assumptions (e.g., data with binary outcomes).
- Students will be able to identify proper regression modeling strategies depending on the type of data encountered.
- Students will be comfortable using software to fit statistical models and use models to answer research questions.
- Students will demonstrate the ability to communicate statistical results effectively.

Grading Policies

- All numeric grades are on a scale from 0 to 100.
- Final grades will be computed according to the following weights (subject to change at the instructor's discretion):

Homework Score	15%
Exam #1	20%
Exam #2	20%
Data Analysis Reports	40%
Class Engagement	5%

- Final letter grades will be determined according to the following rules (subject to change at the instructor's discretion):

A	≥ 90
B	$[80, 90)$
C	$[70, 80)$
D	$[60, 70)$
R	< 60

- Instead of accepting late work, I will calculate your final homework grades out of a point total equivalent to dropping one homework.

If there are $NrHWs$ homework assignments, and the max possible point total across all HWs is $MaxTotal$, then your final score will be scored out of $MaxTotal * (NrHWs - 1) / NrHWs$, up to a max of 100%:

$FinalScore = YourTotal / (MaxTotal * (NrHWs - 1) / NrHWs)$, capped at 100%.

For example, if there are 11 homeworks and each is worth 100 points, the max possible total is 1100 points—but your final HW grade would be scored out of only 1000 points:

$FinalScore = YourTotal / 1000$. Even if you miss a few points here and there, you could earn a full-credit final homework score by submitting all assignments.

Course Components

1. **Lectures.** Regular attendance is essential and in-class participation is expected. The main topics of the course will be covered during the lecture. You are also responsible for any additional material covered in the assigned readings and homework.

If you miss a lecture, you are responsible for the material covered during the lecture you have not attended. Students are expected to take notes and follow along with example problems in class. Some (but not all) course notes and example code will be posted on the course website.

This class brings together students with varying backgrounds in science, math, and statistics, so it is especially vital that you raise questions in class when you don't understand a particular topic, or if you have a perspective on a topic that you feel would contribute to class discussion.

Please bring your textbook with you to class; we will occasionally work through problems in small groups, often using R on our classroom computers. During some lectures, we may assign short small-group activities resulting in a sketch or short writeup to be handed in at the end of class. These informal activities are designed to help you learn and will not be graded for correctness, only monitored for participation. Please always write your name on the sheets you hand in.

2. **Classroom computers.** Our class will meet in a computer lab. These computers are to be used for educational purposes only. Students who use computers during class time for purposes beyond SC321 (email, internet surfing, paper writing, etc.) are obviously not participating in class, and this will be reflected in your final grade. Your hands should be off of the keyboard unless you're using R or our Moodle page.

When working on small-group activities, please do discuss the activity with other students, ask other students for help, and help other students, as long as the talking is not disruptive. Talking is not allowed during exams, however.

3. **Homework.** Homework problems provide you with the opportunity to learn, practice, and test your knowledge and understanding of the material. All material found in the homework may show up in later homeworks and/or the exams.

Homeworks are due on Fridays at 4:00pm, submitted through the course website. Students should submit a single .Rmd file and its knitted .html output file, unless otherwise specified.

Although you are encouraged to discuss problems with each other, I expect each person to submit their own work. You may choose to work with a partner on data analyses and thus produce the same computer output—that's okay, but I want written interpretations and responses to be in your own words. And remember, an answer to a statistics question is almost never just a number—I am looking for thoughtful explanations and conclusions.

The assignments are designed to thoroughly test your knowledge of topics we discuss in class and are definitely not meant to be one-night jobs. We will give you adequate time to work on the problems, and the graders will work hard to return your homework in a timely manner. Unfortunately, this means that **late homework will not be accepted**. Instead, the grading policy allows for the equivalent of dropping 1 homework (see above).

4. **Code.** All code should be written in R and RMarkdown. Students should follow one of Hadley Wickham's popular style guidelines: the one-page [style guide from Advanced R](#), or its expanded version, the [tidyverse Style Guide](#).

5. **Exams.** There will be two exams during the semester, each consisting of two in-class parts. These exams will focus on your ability to use the statistical software, to interpret results, to express an understanding of statistical concepts, and to engage in statistical thinking on open-ended questions. Specific details about the content and format of the exams will be available closer to the exam dates. The exams are closed-book and closed-notes unless otherwise specified.

Tentative dates: Weds+Fri March 20+22, and Weds+Fri May 8+10.

You must meet with me within the first two weeks of classes if you need to reschedule an exam date or request accommodations.

6. **Data Analysis Reports.** Throughout the semester you will conduct statistical analyses on selected datasets and write up your results in formal papers. You'll work both individually and collaboratively on these. I anticipate that 3 or 4 reports will be assigned over the course of the semester. The last of these will be due during our regular final exam period (in lieu of a final exam).

Computing

- All projects and other course assignments must be written in R and RMarkdown unless otherwise specified.
- Students with laptops and personal computers are encouraged to download the latest versions of R and RStudio. Instructions to do this will be given during the first week of classes.
- During lab time, students are welcome to use the campus computers or Colby's RStudio server (<https://rstudio.colby.edu/>)—at least in the first few weeks. It is possible that students may eventually need to use their own computers, since the lab computers may not support some recent R packages that we will use later in the course.
- If you do not have a laptop or personal computer that can run RStudio, and you will need to use lab computers all semester or to access Colby's RStudio server instead, please **immediately** notify the instructor.

Administrative Procedures and Logistics

- **Masks.** According to guidance from the Dean of Faculty, "Regardless of campus-wide policies, faculty and teaching staff may require masks in classrooms and other teaching spaces. More generally, we appreciate everyone respecting the needs of others and wearing masks as requested, and we ask that you have a mask with you when you are on campus." Depending on conditions, I may expect the SC321 classroom to mask in order to protect those of us with compromised immune systems, those of us with high-risk family members, and those of us who simply want to avoid transmitting the virus to others or being re/infected themselves.
- **Lectures.** Use common courtesies: arrive on time; do not leave early; no cell-phone use allowed; do not be disruptive in class; participate in class when the instructor asks questions; etc. The use of laptops/tablets/etc is allowed only for course-related purposes; students are encouraged to follow along and run code in class.
- **Course Materials and Announcements:** The syllabus, lab assignments, homework assignments, solutions, assigned readings, any supplementary material, and grades for this

course can be found on the course web page on Moodle: <https://moodle.colby.edu/>. **Please check Moodle regularly.**

Moodle will also be used to send out course announcements. **Please check your Colby email account regularly.**

- **Communication.** If you have any questions related to the class material, homework problems and exams, please ask the instructor during class or office hours.

Please use email only to address administrative and logistic issues. Questions about homework submitted by email may not be answered quickly or at all. You should not expect a reply within 24 hours.

- **Homework Format.** Homeworks should have the student's name at the very top/beginning. Questions should be answered in order. All answers should be clearly marked and labeled. Answers should be written in the context of the problem when applicable. Proper spelling and grammar should always be used – this means using complete sentences, proper punctuation, etc. Deviating from this format may result in your assignment not being graded.

You are encouraged to discuss homework problems and collaborate with classmates. However, the work you submit must be **your own**. This means, in particular, that each student must independently write up each problem, including all code and written responses. Instances of identical, nearly identical, or copied homework will be considered cheating and plagiarism.

- **Deadline extensions and missed classes.** Colby College supports the religious practices of students, faculty, and staff. Students are expected to notify their instructors of their intent to fulfill the obligations of their religious tradition well in advance of these days. For this class I ask that you notify me by email at least 14 days in advance of the date in question.

Colby rules state that student athletes are permitted to miss a class for a contest but only at the discretion of the professor of that class. Students should communicate with the professor directly and as soon as possible (ideally in the first week of classes) about any conflicts between their athletic schedule and our class schedule.

In general, extensions will not be granted for students because they are behind on work, had a busy week, etc. Extensions for **reasonable academic purposes** (e.g. job interview) or **extreme circumstances** (e.g. hospitalization) will generally be granted. If you believe you have a reasonable request for an extension, please request this at least 48 hours before an assignment is due, along with the reason for requesting an extension. At the top of the assignment, please clearly write that you received an extension on the assignment.

- **Academic accommodations.** I am available to discuss academic accommodations that any student with a documented disability may require. Please note that you'll need to provide a letter from the Dean of Studies Office documenting your approved accommodations. **Please meet with me within two weeks of the start of the semester to make a request for accommodations** so that we can work together with the College to make the appropriate arrangements for you. Colby's Office for Student Access and Disability Services is the primary contact for accommodations and any questions related to educational testing and documentation: <https://life.colby.edu/get-support/access-disability-services/>
- **Sexual misconduct/Title IX statement.** Colby College prohibits and will not tolerate sexual misconduct or gender-based discrimination of any kind. Colby is legally obligated

to investigate sexual misconduct (including, but not limited to sexual assault and sexual harassment).

If you wish to speak confidentially about an incident of sexual misconduct, please contact Colby Counseling Services (207-859-4490) or the Director of the Gender and Sexual Diversity Program, Emily Schusterbauer (207-859-4093).

Students should be aware that faculty members are considered responsible employees; as such, if you disclose an incident of sexual misconduct to a faculty member, they have an obligation to report it to Colby's Title IX Coordinator. "Disclosure" may include communication in-person, via email/phone/text, or through class assignments.

To learn more about sexual misconduct or report an incident, visit <https://life.colby.edu/your-safety/sexual-violence-title-ix/>

- **Academic honesty and consequences for academic dishonesty.** Always ask if you are unsure whether your actions comply with the assignment instructions. Always acknowledge any help received on assignments: list the names of the people you worked with, and cite any external sources you used (including generative AI). You are encouraged to discuss assignments with your classmates, but the work you submit must be your own.

Honesty, integrity, and personal responsibility are cornerstones of a Colby education and provide the foundation for scholarly inquiry, intellectual discourse, and an open and welcoming campus community. These values are articulated in the Colby Affirmation and are central to this course. Students are expected to demonstrate academic honesty in all aspects of this course.

Academic dishonesty includes, but is not limited to:

- plagiarism (including quoting sources without “ ” around the borrowed words and a citation);
- presenting another's work as one's own (this includes generative AI such as ChatGPT);
- buying or attempting to buy papers or projects for a course;
- fabricating (or using generative AI to fabricate) information or citations;
- knowingly assisting others in acts of academic dishonesty;
- violating clearly stated rules for taking an exam (e.g., keeping your phone on your person when told to place it in a bag) or completing homework;
- misrepresentations to faculty within the context of a course; and
- submitting the same work, including an essay that you wrote, in more than one course without the permission of instructors.

Academic dishonesty is a serious offense against the college. Sanctions for academic dishonesty are assigned by an academic review board and may include failure on the assignment, failure in the course, or suspension or expulsion from the College. All students found responsible for violating academic integrity standards will have a disciplinary letter placed in their file for 6 years after they leave Colby. This letter can be released to requesting parties (e.g., medical/law school, employers). Thus, the consequences of dishonesty far exceed the benefits (not to mention the fact that cheating is a disservice to you and your learning process).

For more on recognizing and avoiding plagiarism, see:
<https://libguides.colby.edu/avoidingplagiarism>

For resources and information on academic integrity, see:

<https://www.colby.edu/academics/academic-integrity/>

- **A special note on AI and ChatGPT.** Under no circumstances should you be using any type of AI in preparing reports or exams. This includes, but is not necessarily limited to, using any kind of online AI chat (e.g. ChatGPT). All writing that is conducted in this class must be original and be your own. If you have any questions about this, ask me.
- **The Colby Affirmation.** Colby College is a community dedicated to learning and committed to the growth and well-being of all its members. As a community devoted to intellectual growth, we value academic integrity. We agree to take ownership of our academic work, to submit only work that is our own, to fully acknowledge the research and ideas of others in our work, and to abide by the instructions and regulations governing academic work established by the faculty. As a community built on respect for ourselves, each other, and our physical environment, we recognize the diversity of people who have gathered here and that genuine inclusivity requires active, honest, and compassionate engagement with one another. We agree to respect each other, to honor community expectations, and to comply with College policies. As a member of this community, I pledge to hold myself and others accountable to these values.
<https://www.colby.edu/academics/academic-integrity/the-colby-affirmation/>