Stat 331/531: Statistical Computing with R
Winter 2023

Contact

Dr. Emily Robinson

- Email: erobin17@calpoly.edu
- Office: Building 25 Office 103 (by Statistics Department Office)

Course Forum:

For questions of general interest, such as course clarifications or conceptual questions, please use the class discussion forum. I encourage you to give your post an informative title, so that other people can find it. For example, *"How do I color bars in a barplot with ggplot?"* is a better discussion title than *"help with plotting"*.

While your posts are not anonymous, in this case there is no such thing as a bad question!

Course Info

Times:

Mondays/Wednesdays

- Section 70: 10:10am - 12:00pm
- Section 71: 12:10pm - 2:00pm

Room: 38-0123 (Math & Science)

Office Hours

Day	Time
Tuesdays	10:10am - 11:00am, in-person (25-103)
Thursdays	10:10am - 11:00pm, in-person (25-103)
Fridays	1:15pm - 2:15pm, by appt (https://calendly.com/calpoly-emily-robinson/15min)

Office hours by appointment are required to be scheduled at least 1-hour prior to the meeting. You can schedule either Zoom or in-person office hour appointments through Calendly.

Course Description

Introduction to programming for statistical analysis. Covers basic programming concepts necessary for statistics, good computing practice, and use of built-in functions to complete basic statistical analyses.

Prerequisites

Entrance to STAT 331/531 requires successful completion of a Stat II qualifying course and an introductory computing course.

Learning Objectives

This course will teach you the foundations of statistical computing principles in the language of R.

After taking this course, you will be able to:

- Work with the RStudio IDE and quarto documents.
- Import, manage, and clean data from a wide variety of data sources.
- Visualize and summarize data for informative exploratory data analysis and presentations.
- Write efficient, well-documented, and tidy R code.
- Program random experiments and simulations from probability models
- Extend your R skills independently through documentation and online resources.

Additionally, it is my hope that you will learn to:

- Extend your R skills independently through documentation and online resources.
- Be thoughtful, deliberate, and ethical in your use of R and similar tools.
- Use R to be playful, creative, and fun!
- Contribute to and participate in the R Open Source Community.

Course Resources

Textbook

The primary textbook for this course modified from material by Susan VanderPlas and other resources listed below. It is available for free at https://earobinson95.github.io/stat331-calpoly/. It is under construction/a work in progress, so it may be hard to work more than a couple of weeks ahead in this class using the primary textbook.

In addition, you may find it useful to reference some of the following resources that I have consulted while assembling the textbook. Most are available online for free, though some require an institutional email address.

- R for Data Science (2nd edition)
- Modern Dive
- Posit Primers
- Advanced R

Equipment

Although you may always work on the Studio computers, I strongly recommend that you use your own personal laptop for this course if you have one.

Chromebooks and ipads will **not** be sufficient to use R.

If this requirement is limiting for you, please contact me ASAP.

Class Schedule & Topic Outline

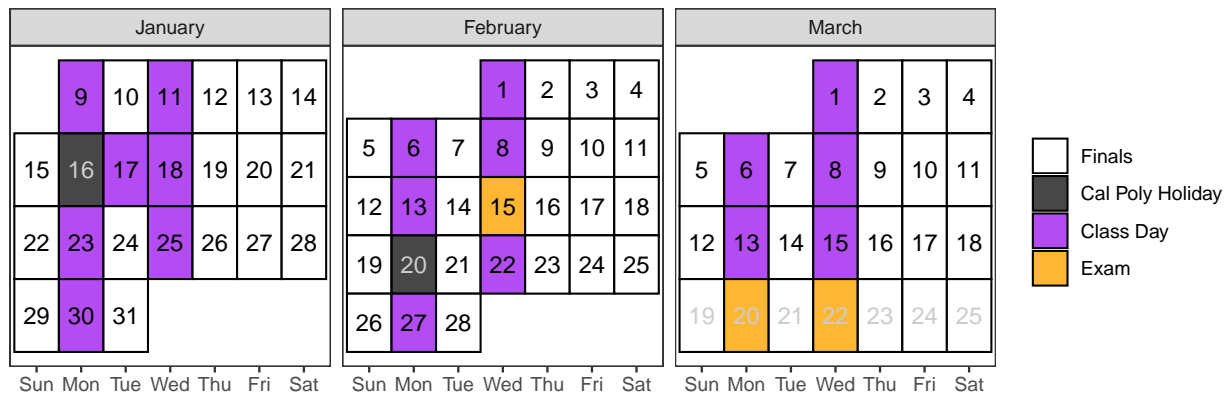This schedule is tentative and subject to change.

Figure 1: Tuesday, January 17th will follow a Monday class schedule.

Table 2: Tentative schedule of class topics and important due dates

Date	Topic
Jan 9, Jan 11	Getting Started
Jan 17, Jan 18	Basic Variable Types & Control Structures
Jan 23, Jan 25	Data Structures & Programming with Data
Jan 30, Feb 1	Data Cleaning and Manipulation (dplyr)
Feb 6, Feb 8	Data Transformations (tidyr)
Feb 13	Data Visualization (ggplot2)
Feb 15	Midterm Exam
Feb 22	Strings-Factors-Dates
Feb 27, Mar 1	Debugging & Version Control
Mar 6, Mar 8	Simulation & Modeling
Mar 13, Mar 15	Simulation & Modeling
Mar 20	Final Exam
Mar 22	Final Exam

Course Policies

Assessment/Grading

Your grade in STAT 331/531 will contain the following components:

Assignments	Weight
Flex Points	5%
Practice Activities	10%
Homework Assignments	15%
Project I	15%

Assignments	Weight
Midterm	15%
Project II	15%
Final Exam	25%

Lower bounds for grade cutoffs are shown in the following table. I will not "round up" grades at the end of the quarter beyond strict mathematical rules of rounding.

Letter grade	X +	X	X -
A	.	93	90
B	87	83	80
C	77	73	70
D	67	63	60
F	<60		

Interpretation of this table:

- A grade of 85 will receive a B.
- A grade of 77 will receive a C+.
- A grade of 70 will receive a C-.
- Anything below a 60 will receive an F.

General Evaluation Criteria

In every assignment, discussion, and written component of this class, you are expected to demonstrate that you are intellectually engaging with the material. I will evaluate you based on this engagement, which means that technically correct but low effort answers which do not demonstrate engagement or understanding will receive no credit.

When you answer questions in this class, your goal is to show that you either understand the material or are actively engaging with it. If you did not achieve this goal, then your answer is incomplete, regardless of whether or not it is technically correct. This is not to encourage you to add unnecessary complexity to your answer - simple, elegant solutions are always preferable to unwieldly, complex solutions that accomplish the same task.

While this is not an English class, grammar and spelling are important, as is your ability to communicate technical information in writing; both of these criteria will be used in addition to assignment-specific rubrics to evaluate your work.

Course Expectations

You will get out of this course what you put in. The following excerpt was taken from Rob Jenkins' article "Defining the Relationship" which was published in The Chronicle of Higher Education (August 8, 2016). This accurately summarizes what I expect of you in my classroom (and also what you should expect of me).

"I'd like to be your partner. More than anything, I'd like for us to form a mutually beneficial alliance in this endeavor we call education.

I pledge to do my part. I will:

- *Stay abreast of the latest ideas in my field.*
- *Teach you what I believe you need to know; with all the enthusiasm I possess.*
- *Invite your comments and questions and respond constructively.*
- *Make myself available to you outside of class (within reason).*
- *Evaluate your work carefully and return it promptly with feedback.*
- *Be as fair, respectful, and understanding as I can humanly be.*
- *If you need help beyond the scope of this course, I will do my best to provide it or see that you get it.*

In return, I expect you to:

- *Show up for class each day or let me know (preferably in advance) if you have some good reason to be absent.*
- *Do your reading and other assignments outside of class and be prepared for each class meeting.*
- *Focus during class on the work we're doing and not on extraneous matters (like whoever or whatever is on your phone at the moment).*
- *Participate in class discussions.*
- *Be respectful of your fellow students and their points of view.*
- *In short, I expect you to devote as much effort to learning as I devote to teaching.*

What you get out of this relationship is that you'll be better equipped to succeed in this and other college courses, work-related assignments, and life in general. What I get is a great deal of professional and personal satisfaction. Because I do really like you [all] and want the best for you."

Make Mistakes!

Programming is the process of making a series of silly or stupid mistakes, and then slowly fixing each mistake (while adding a few more). The only way to know how to fix these mistakes (and avoid them in the future) is to make them. (Sometimes, you have to make the same mistake a few dozen times before you can avoid it in the future). At some point during the class, you will find that you've spent 30 minutes staring at an error caused by a typo, a space, a parenthesis in the wrong place. You may ask for help debugging this weird error, only to have someone immediately point out the problem... it is always easier to see these things in someone else's code. This is part of programming, it is normal, and you shouldn't feel embarrassed or sorry (unless you put no effort into troubleshooting the problem before you asked for help)

If you manage to produce an error I haven't seen before, then congratulations. You have achieved something special, and that achievement should be celebrated. Each new and bizarre error is an opportunity to learn a bit more about the programming language, the operating system, or the interaction between the two.

Other Policies

See academicprograms.calpoly.edu/content/academicpolicies.

Participation and Attendance

I do not take formal attendance in this class. However, it is my expectation that you remain in class and on task until you have finished all your activities and assignments. Consistent, repeated failure to attend class or actively participate in the online portions of the course will affect the participation portion of your grade.

If you are feeling ill, please **do not come to class**. Instead, email me, review the material and work on the participation activity and weekly lab assignment, and then schedule an appointment with me to meet virtually.

Late Policy

Your typical homework assignments will be weekly labs. These labs are due on Fridays at midnight.

There are no extensions on lab assignments. There is a grace period through Sunday at midnight, during which your late lab recieves a 10% grade deduction.

Labs turned in later than Sunday will not be reviewed, and instead will be quickly checked for completeness and recieve an automatic grade of 50%.

Diversity and Inclusion

It is my goal for everyone to feel safe and comfortable in my classroom. If there is any way I can make the course more welcoming for you, please do not hesitate to ask.

In particular, if you have a disability, I will gladly work with you to make this class accessible.

I encourage you to also contact the Disability Resource Center (Building 124, Room 119 or at 805-756-1395), who can help you register for extra accomodations such as extended exam time.

Academic Integrity and Class Conduct

Simply put, I will not tolerate cheating or plagiarism.

Any incident of dishonesty, copying, exam cheating, or plagiarism will be reported to the Office of Student Rights and Responsibilities.

Cheating will earn you a grade of 0 on the assignment and an overall grade penalty of at least 10%. In circumstances of flagrant cheating, you may be given a grade of F in the course.

Paraphrasing or quoting another's work without citing the source is a form of academic misconduct. This included the `R` code produced by someone else! Writing code is like writing a paper, it is obvious if you copied-and-pasted a sentence from someone else into your paper because the way each person writes is different.

Even inadvertent or unintentional misuse or appropriation of another's work (such as relying heavily on source material that is not expressly acknowledged) is considered plagiarism. If you are struggling with writing the `R` code for an assignment, please reach out to me. I would prefer that I get to help you rather than you spending hours Googling things and get nowhere!

If you have any questions about using and citing sources, you are expected to ask for clarification.

For more information about what constitutes cheating and plagiarism, please see academicpro-grams.calpoly.edu/content/academicpolicies/Cheating.