

Stat 204: Introduction to Statistical Data Analysis

Fall 2026

Table of contents

1	COURSE INFORMATION	2
1.1	Course Description	3
2	LEARNING OUTCOMES	3
3	REQUIRED MATERIALS, TEXTBOOKS AND TECHNOLOGY	3
3.1	Laptop or Desktop Computer	3
3.2	Technology	4
3.3	Required Textbooks	4
3.4	Other Recommended Books	4
3.5	Prerequisites	4
4	COMMUNICATION	4
5	ASSIGNMENTS & ASSESSMENT	4
5.1	Class participation: 10%	5
5.2	Homework: 30%	5
5.3	Midterm Exam: 25%	6
5.4	Final Project: 35%	6
5.4.1	Final Project Components:	6
5.5	Grading Scale	7
6	COURSE WEEKLY SCHEDULE	7
7	ATTENDANCE AND COURSE DELIVERY POLICY	9
7.1	In-Person Learning Community	9

8	ARTIFICIAL INTELLIGENCE (AI) POLICY	9
8.1	Why This Matters	9
8.2	What Is Prohibited, Permitted, and Required	9
8.3	How Concerns Are Handled	10
8.4	My Own Use of AI	10
8.5	AI Literacy Throughout the Course	11
8.6	Equity and Access	11
8.7	Privacy and Materials	11
8.8	AI for Study Support	11
8.9	Exception: GenAI for the Final Project	12
9	ACCESSIBILITY	12
10	TITLE IX/CARE ADVISORY	12
11	ACADEMIC INTEGRITY	13
12	Academic integrity includes, but is not limited to, the following:	13
13	Academic misconduct includes, but is not limited to, the following:	13
14	STUDENT SUPPORT SERVICES	14
15	RELIGIOUS ACCOMMODATION	14
16	PRINCIPLES OF COMMUNITY	14

1 COURSE INFORMATION

Instructor: Marcela Alfaro Córdoba

Email: macordob@ucsc.edu

Office Hours:

- Tuesday & Thursday, 3:15 pm - 4:00 pm — McHenry Library (*exact location to be updated*)
- Wednesday, 1:00 pm - 2:30 pm — via Zoom
- How to book: book an appointment using [this link](#)

Class Times: Tuesday & Thursday 5:20 pm - 6:55 pm **Location:** McHenry Library 1350
Section (self-guided work): Wednesday 1:20 pm - 2:25 pm, at Soc Sci 2 363. The instructor is available via Zoom for most of this time. **Quarter Dates:** September 24, 2026 - December 4, 2026 **Course Materials:** Canvas and Course Webpage

1.1 Course Description

This course presents tools for exploratory data analysis (EDA) and statistical modeling in R. Topics include: numerical and graphical methods for EDA, linear and logistic regression, ANOVA, PCA, and tools for acquiring and storing large data. No R knowledge is required.

2 LEARNING OUTCOMES

By the end of the course, students should be able to:

- Comfortably load, manipulate, and analyze data within the R environment, building proficiency from no prior R knowledge to advanced statistical computing
- Apply numerical and graphical methods to explore datasets, understand data structures, and create meaningful visualizations using modern R packages
- Fit, diagnose, and interpret common linear models including linear regression, ANOVA, and logistic regression, with emphasis on model assumptions and validation
- Implement principal components analysis (PCA), clustering methods, and regularized regression techniques to handle complex, high-dimensional datasets
- Grasp the mathematical foundations underlying linear regression, ANOVA, logistic regression, and PCA, including key assumptions, theoretical properties, and when each method is appropriate to use
- Integrate exploratory analysis with statistical modeling to draw meaningful conclusions from real-world datasets, culminating in a comprehensive final project
- Use generative AI (genAI) tools responsibly to strengthen the reproducibility of your code and verify that it runs correctly, while relying on your own understanding — not AI — for the written interpretation, argument, and prose of your report and presentation; genAI may be used to polish formatting, not to write content

3 REQUIRED MATERIALS, TEXTBOOKS AND TECHNOLOGY

3.1 Laptop or Desktop Computer

You will need a computer for homework, in-class activities, and projects. You will not be able to complete work on a mobile device. Students who need a laptop can make use of the library's borrow program: <https://library.ucsc.edu/services/computing/borrow-a-laptop>.

3.2 Technology

Personal computer with Canvas, R, and [Positron](#) installed. All these tools should be free for you, please make sure you have access WITHOUT payment. Ask me if you have any problems, and please DO NOT pay for any of them. No assignment in this course requires an AI tool, and no one is advantaged by paying for one.

3.3 Required Textbooks

- **R by Example** (2012) by Jim Albert and Maria Rizzo, Springer Use R! Series A free PDF copy of the textbook is available from the university library: <https://library.ucsc.edu>
- **Veridical Data Science** (2024) by Bin Yu and Rebecca Barter — free online at <https://vdsbook.com/>

3.4 Other Recommended Books

- **R for Data Science** (2017) by Garrett Grolemund and Hadley Wickham: <https://r4ds.had.co.nz/>
- **An Introduction to Statistical Learning** (2013) by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani

3.5 Prerequisites

Familiarity with probability and statistical inference is assumed, along with some experience using matrix algebra operations.

4 COMMUNICATION

All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on this site. I will regularly send course announcements via Ed Discussion. Make sure to check it regularly. If an announcement is sent Monday through Thursday, I will assume that you have read the announcement the next day. If an announcement is sent on a Friday or over the weekend, I will assume that you have read it by Monday.

5 ASSIGNMENTS & ASSESSMENT

Assessment for the course consists of four components: class participation, homework, midterm exam, and final project.

5.1 Class participation: 10%

It is expected that you come to lectures and have discussions with your classmates. Participation matters because the midterm and the project oral exam are both in person, with no AI available: talking through statistical ideas with peers and with me is the practice for them. There will be activities to check participation in every lecture. These activities include weekly checks on the final project work starting on Week 4. If you have to skip lectures at some point, please post it on Ed Discussion (you can use the private mode), that will be my log for grading this part at the end of the quarter.

5.2 Homework: 30%

There will be 5 homework sets throughout the quarter. You may discuss homework assignments with other students; however, homework should be completed and submitted individually, without using any AI tools. Homework is where you build the coding fluency and judgment you need for the midterm. Each assignment lists what, if anything, beyond the class notes and R documentation is permitted; if it says nothing, AI is not permitted. Some homework also includes a short “where AI would fail here” note (for example, hallucinated code or incorrect statistical claims); it is there to build your AI literacy and does not permit AI use on the work.

All homework is due by **11:59 pm (Pacific)** on the dates below, and all of it is due before the midterm:

Homework	Topic	Released	Due
HW1	Quarto tutorial + intro to R	Thu, Sep 24	Thu, Oct 1
HW2	Data import, cleaning, and exploration	Thu, Oct 1	Thu, Oct 8
HW3	Exploratory data analysis and visualization	Thu, Oct 8	Tue, Oct 20
HW4	Statistical modeling: linear and logistic regression	Tue, Oct 20	Thu, Nov 12
HW5	Midterm study guide and practice problems	Thu, Nov 5	Mon, Nov 16

Homework check. Because take-home work is the hardest to verify, for each homework set I will randomly select a sample of students to briefly explain one of their solutions (about 5 minutes) in the next class. Selection is random and is not an accusation. Being unable to explain work you submitted may affect your score on that problem, after a conversation with me (see the AI Policy).

5.3 Midterm Exam: 25%

The midterm exam is tentatively scheduled for **November 19, 2026** (Week 8). More details about the exam format and content will be provided closer to the date.

5.4 Final Project: 35%

The final project should be done in groups of 2 students. These projects should be based on a chosen dataset and should include exploratory data analysis, as well as a full data analysis using the modeling approaches discussed in class. Each project must be kept in a Git repository (or a Quarto project with a commit history) that I can view, and the report must include an AI-use disclosure statement (see the AI Policy).

5.4.1 Final Project Components:

1. **Project Proposal:** A short 1/2 - 1 page proposal that includes the chosen dataset and outlines the goals of the project. Additionally, a 5-minute presentation should be prepared.
2. **Project Presentation:** This course has no written final exam. Instead, project presentations will be held in class during **Week 10 (Dec 1 & 3, 2026)** and are treated as an oral exam: come prepared not only to explain what you did, but also to defend it. These should be about 10 minutes each, and attendance to all presentations is mandatory.
3. **Peer Feedback & Code Review:** During your classmates' presentations, you are responsible for writing feedback, and for reviewing the code of one other project. This peer review is due **Thursday, December 3, 2026**, right after presentations. Consolidated feedback (peer and instructor) will be posted the next day, **Friday, December 4, 2026**, so you can incorporate it into your report before it's due the following Tuesday.
4. **Project Report:** The report should summarize the entire project, and incorporate the feedback given during presentations and the peer code review. This should be no more than 6 pages including figures and tables, but excluding references. The report is due by **5 pm on Tuesday of finals week, December 8, 2026**.

5.5 Grading Scale

Percentage	Grade
90% or above	A
80% to < 90%	B
70% to < 80%	C
60% to < 70%	D
< 60%	F

6 COURSE WEEKLY SCHEDULE

This schedule may change during the quarter. Please consult the class webpage for a more updated version.

Week	Dates	Topic	Homework	Final Project Milestones
1	Sep 24, Sep 29	Introduc- tion and basics of R	HW1 released (Sep 24)	
2	Oct 1, Oct 6	Numerical and graphical sum- maries	HW1 due (Oct 1); HW2 released	
3	Oct 8, Oct 13	Advanced graphics	HW2 due (Oct 8); HW3 released	
4	Oct 15, Oct 20	Linear Re- gression and additional regression topics	HW3 due (Oct 20); HW4 released	
5	Oct 22, Oct 27	Analysis of Variance (ANOVA)		
6	Oct 29, Nov 3	Logistic regression		Proposal due (Oct 29, start of class); proposal presentations (Nov 3)

Week	Dates	Topic	Homework	Final Project Milestones
7	Nov 5, Nov 10	Principal component analysis and clustering	HW5 released (Nov 5)	Remaining proposal presentations (Nov 5)
8	Nov 12, Nov 17, Nov 19	Regularized regression / Midterm Exam	HW4 due (Nov 12); HW5 due (Nov 16)	
9	Nov 24	GenAI for reproducibility & code testing, verifying AI output, and AI bias and environmental cost (Nov 26 - Thanksgiving)		
10	Dec 1, Dec 3	Project presentations (oral exam)		Presentations (Dec 1-3); Peer Feedback & Code Review due (Dec 3); consolidated feedback posted (Dec 4)
Finals Week	Dec 7-11	(<i>no written final exam</i>)		Final Project Report due (Dec 8, 5:00 PM)

In addition to the Tuesday/Thursday sessions above, a self-guided work section meets every Wednesday, 1:20 pm - 2:25 pm, in a separate classroom; the instructor is available via Zoom for most of this time.

This course has no written final exam; project presentations during Week 10 serve as the oral

exam.

7 ATTENDANCE AND COURSE DELIVERY POLICY

7.1 In-Person Learning Community

This course thrives on active participation and collaborative learning. To get the most out of our time together, regular attendance at lectures and discussion sections is essential for your success.

- Exams: All exams will be held in person to ensure fairness and academic integrity. This allows us to provide the best testing environment for all students.
- Course Format: We've designed this as an interactive, in-person experience that builds on daily discussions and activities. If you're unable to attend regularly this quarter, we'd encourage you to consider taking the course when you can fully engage with the learning community.

8 ARTIFICIAL INTELLIGENCE (AI) POLICY

8.1 Why This Matters

On the midterm you will interpret model diagnostics, and in the project presentation you will defend your analysis orally, with no AI available in either. Homework is where you build that judgment. If a tool does the reasoning for you, you skip the practice you will be tested on. The policy below is meant to explain what is allowed and why, not to treat AI use as cheating by default.

8.2 What Is Prohibited, Permitted, and Required

Work	Status	Why
Homework	Prohibited unless the assignment states otherwise	Builds the coding fluency you need for the midterm
Midterm	Prohibited	Measures your own understanding

Work	Status	Why
Project code	Allowed only for reproducibility checks and code testing after the Week 9 session; AI may not write or rewrite your analysis code	Verifying code is a skill; the analysis choices are yours
Project prose, interpretation, and analysis text	Prohibited (grammar correction software is fine; genAI may be used for formatting only)	The interpretation is the point
Studying	Encouraged (see the tips below)	Practice problems and alternative explanations help you learn

Using genAI on graded work outside what this table permits is academic misconduct. The Homework section and the Academic Integrity section follow this same table.

Required disclosure. Wherever AI is permitted on the project (code checks, formatting), your report must include a short “AI use” statement saying which tool you used, what you used it for, and how you verified its output. If you used no AI, say so. Grammar or spell-check software also gets one line in that statement.

8.3 How Concerns Are Handled

I don’t use AI detectors: they are unreliable and produce false positives, especially for non-native English speakers and neurodivergent students. If I have a concern about a piece of work, we will resolve it through conversation, and I may ask you to explain your work. I rely on process evidence instead: the project’s Git or Quarto commit history, the weekly project check-ins starting in Week 4, the random homework checks in the Wednesday section, and a short oral explanation of any submission flagged as concerning. Confirmed misconduct is handled under the Academic Integrity section below.

8.4 My Own Use of AI

I will be open about my own use of AI. I may use AI tools to help draft or polish the format of course materials such as slides and practice problems; I review and take responsibility for everything I post. I do not use AI to decide grades or to write your feedback. AI will not decide any grade in this course.

8.5 AI Literacy Throughout the Course

Responsible AI use is practiced across the quarter, not only in one session. Most homework will include a “where AI would fail here” note, and the Week 9 session and the project cover how to verify AI output: hallucinated code, incorrect statistical claims, and code that runs but answers the wrong question. Week 9 also covers bias in AI systems and their environmental cost, so that you can make informed choices about when using them is worth it.

8.6 Equity and Access

No assignment requires an AI tool, and no one is advantaged by paying for one. Please do not buy any AI subscription for this course.

8.7 Privacy and Materials

Do not upload course materials (textbook excerpts, handouts, assignments, exams) to AI platforms; this creates copyright and author rights issues. Likewise, do not upload sensitive or restricted project datasets (for example, data with personal information or data whose license forbids redistribution) to any AI tool. If you are unsure whether your dataset qualifies, ask me before using it with any tool.

8.8 AI for Study Support

You’re welcome to use AI tools to create practice problems or generate study materials to help you explore concepts from different angles. This can be a great way to test your understanding! Just remember that AI can sometimes provide incorrect information, so always double-check against your course materials and use AI-generated content as a supplement to—not a replacement for—your assigned readings and exercises. Some tips:

- **Smart usage boundaries:** Use AI to understand concepts and check your approach, but always solve practice problems independently first. Remember: AI won’t be available during exams, so build your own problem-solving confidence!
- **Use AI as a teaching assistant:** Ask for step-by-step explanations of concepts, alternative explanations if something doesn’t click, and connections to previous topics you’ve studied.
- **Practice problem generation:** Have AI create variations of homework problems with different numbers or scenarios to test your understanding.
- **Protect class materials:** See “Privacy and Materials” above. Ask general questions about concepts or create your own examples for AI to work with.

When in Doubt, Ask: If you're ever uncertain about whether a particular use of AI would be appropriate for your learning goals, please reach out. We're here to help you succeed while ensuring you're genuinely mastering the material.

8.9 Exception: GenAI for the Final Project

After the Week 9 session on genAI for reproducibility and code testing, you may use genAI tools for your final project to (1) check that your code is reproducible and runs correctly, and (2) improve the formatting of your report and presentation. GenAI may **not** be used to write or rewrite analysis code, prose, interpretation, or analysis text — that must be your own understanding and argument. Every such use must be disclosed in the report's AI-use statement.

9 ACCESSIBILITY

UC Santa Cruz is committed to creating an academic environment that supports its diverse student body. If you are a student with a disability who requires accommodations to achieve equal access to this course, please affiliate with the DRC. I encourage all students who would benefit from learning more about DRC services to contact DRC by phone at 831-459-2089 or by email at drc@ucsc.edu. For students already affiliated, make sure that you have requested Academic Access Letters, where you intend to use accommodations. You can also request to meet privately with me during my office hours or by appointment, as soon as possible. I would like us to discuss how we can implement your accommodations in this course to ensure your access and full engagement in this course.

10 TITLE IX/CARE ADVISORY

The Title IX Office is committed to fostering a campus climate in which members of our community are protected from all forms of sex discrimination, including sexual harassment, sexual violence, and gender-based harassment and discrimination. Title IX is a neutral office committed to safety, fairness, trauma-informed practices, and due process. Title IX prohibits gender discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking. If you have experienced sexual harassment or sexual violence, you can receive confidential support and advocacy at the Campus Advocacy Resources & Education ([CARE](#)) Office by calling (831) 502-2273. In addition, Counseling & Psychological Services ([CAPS](#)) can provide confidential counseling support, (831) 459-2628. You can also report gender discrimination directly to the [University's Title IX Office](#), (831) 459-2462. Reports to law enforcement can be made to UCPD, (831) 459-2231 ext. 1. For emergencies call 911.

11 ACADEMIC INTEGRITY

All members of the UCSC community benefit from an environment of trust, honesty, fairness, respect, and responsibility. You are expected to present your own work and acknowledge the work of others in order to preserve the integrity of scholarship, and the value of the degrees you and your peers are working so hard to earn.

12 Academic integrity includes, but is not limited to, the following:

- Being transparent and forthright in your academic work
- Reading the syllabus
- Asking questions about this academic integrity policy if you don't understand it
- Following exam rules
- Using only permitted materials during an exam
- Viewing exam materials only when permitted by your instructor
- Keeping what you know about an exam to yourself
- Incorporating proper citation of all sources of information
- Disclosing any permitted use of AI (in the project's AI-use statement) and of editing software or grammar tools
- Submitting your own original work

13 Academic misconduct includes, but is not limited to, the following:

- Disclosing exam content during or after you have taken an exam
- Accessing exam materials without permission
- Copying/purchasing any material from another student, or from another source, that is submitted for grading as your own
- Plagiarism, including use of Internet material without proper citation
- Submitting work that was produced in whole or in part by generative artificial intelligence, except for the final project's permitted and disclosed genAI uses (reproducibility/code checking and formatting) described in the AI Policy section above
- Uploading course materials or sensitive/restricted project data to AI tools
- Failing to properly cite external sources when their use is permitted
- Using cell phones or other electronics to obtain outside information during an exam without explicit permission from the instructor
- Submitting your own work in one class that was completed for another class (self-plagiarism) without prior permission from the instructor

Violations of the academic integrity policy can result in a permanent notation on your transcript, a failing grade on an assignment or in the course, and/or dismissal from the university. For the full policy and formal resolution procedures for academic misconduct, please refer to the [Academic Integrity Office](#).

14 STUDENT SUPPORT SERVICES

To learn about the many student services offered at UCSC—such as Learning Support Services (LSS), Resource Centers, or Slug Support—I encourage you to visit the Campus Resources website or click on the “Resources” button (a slug in a heart) at the bottom of the navigation ribbon on the left side of our Canvas page.

15 RELIGIOUS ACCOMMODATION

UC Santa Cruz welcomes diversity of religious beliefs and practices, recognizing the contributions differing experiences and viewpoints can bring to the community. There may be times when an academic requirement conflicts with religious observances and practices. If that happens, students may request reasonable accommodation for religious practices. The instructor will review the situation in an effort to provide a reasonable accommodation without penalty. You should first discuss the conflict and your requested accommodation with your instructor early in the term. You or your instructor may also seek assistance from the [Dean of Students office](#).

16 PRINCIPLES OF COMMUNITY

The University of California, Santa Cruz expressly prohibits students from engaging in conduct constituting unlawful discrimination, harassment or bias. I am committed to providing an atmosphere for learning that respects diversity and supports inclusivity. I ask all members of this class to:

- Be open to and interested in the views of others
- Consider the possibility that your views may change over the course of the term
- Honor the unique life experiences of your colleagues
- Listen to each other’s opinions and communicate respectfully
- Keep confidential discussions of a personal nature
- Ground your comments in course materials and maintain focus on learning together

This syllabus is subject to change. Any changes will be announced in class and posted on Ed Discussion.