STA 221 Syllabus

Fall 2024

Course info

Class meetings

Lecture	Tue & Thu 10:05 - 11:20am	Old Chemistry 116
Lab 01	Fri 8:30 - 9:45am	Perkins Link #5
Lab 02	Fri 10:05 - 11:20am	Perkins Link #5

Teaching team

Name	Role
Prof. Maria Tackett	Instructor
Kat Husar	Head TA
	Lab 01L leader
Kelly Huang	Classroom TA
Janice Kim	Classroom TA
Cathy Lee	Lab 02L leader
Alan Wang	Lab 01 helper

Office hours times and locations on Canvas.

Course description

In STA 221, students will learn how linear and logistic regression models are used to explore multivariable relationships, apply these methods to answer relevant and engaging questions using a data-driven approach, and learn the mathematical underpinnings of the models. Students will develop computing skills to implement a reproducible data analysis workflow and

gain experience communicating statistical results. Throughout the semester, students will work on a team project where they will develop a research question, answer it using methods learned in the course, and share results through a written report and presentation.

Topics include applications of linear and logistic regression, analysis of variance, model diagnostics, and model selection. Regression parameter estimation via maximum likelihood least squares will also be discussed. Students will gain experience using the computing tools R and GitHub to analyze real-world data from a variety of fields.

Prerequisites

Either any STA 100-level course or STA 230, 231, or 240L and MATH 216, 218, or 221. The recommended co-requisite is STA 230, 231, or 240L. Interested students with different backgrounds should seek instructor consent.

Course learning objectives

By the end of the semester, you will be able to…

- analyze data to explore real-world multivariable relationships.
- fit, interpret, and draw conclusions from linear and logistic regression models.
- implement a reproducible analysis workflow using R for analysis, Quarto to write reports and GitHub for version control and collaboration.
- explain the mathematical foundations of linear and logistic regression.
- effectively communicate statistical results to a general audience.
- assess the ethical considerations and implications of analysis decisions.

Course materials

While there is no official textbook for the course; readings will be made available as they are assigned. We will use the statistical software R. Students will be able to access R through Docker containers provided by Duke Office of Information Technology. See the computing page for more information.

Course community

Inclusive community

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength, and benefit. It

is my intent to present materials and activities that are respectful of diversity and in alignment with Duke's Commitment to Diversity and Inclusion. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities. To help accomplish this:

- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. If you prefer to speak with someone outside of the course, your academic dean is an excellent resource.
- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please let me or a member of the teaching team know.

Pronouns

Pronouns are meaningful tools to communicate identities and experiences, and using pronouns supports a campus environment where all community members can thrive. Please update your gender pronouns in Duke Hub. You can find instructions to do so here. You can learn more at the Center for Sexual and Gender Diversity's website.

Accessibility

If there is any portion of the course that is not accessible to you due to challenges with technology or the course format, please let me know so we can make appropriate accommodations.

The Student Disability Access Office (SDAO) is available to ensure that students are able to engage with their courses and related assignments. Students should be in touch with the Student Disability Access Office to request or update accommodations under these circumstances.

Communication

All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on the course website, sta221-sp25.netlify.app.

Links to Zoom meetings may be found in Canvas. Periodic announcements will be sent via email and will also be available through Ed Discussion and Canvas Announcements. Please check your email regularly to ensure you have the latest announcements for the course.

Email

If you have questions about assignment extensions, accommodations, or any other matter not appropriate for the class discussion forum, please email me directly at maria.tackett@duke.edu. **If you email me, please include "STA 221" in the subject line.** Barring extenuating circumstances, I will respond to STA 221 emails within 48 hours Monday - Friday. Response time may be slower for emails sent Friday evening - Sunday.

Five tips for success

Your success on this course depends very much on you and the effort you put into it. Your TAs and I will help you be providing you with materials and answering questions and setting a pace, but for this to work you must do the following:

1. Complete all the preparation work before class.
2. Ask questions. As often as you can. In class, out of class. Ask me, ask the TAs, ask your friends, ask the person sitting next to you. This will help you more than anything else. If you get a question wrong on an assessment, ask us why. If you're not sure about the homework, ask. If you hear something on the news that sounds related to what we discussed, ask. If the reading is confusing, ask.
3. Do the readings and other preparation work.
4. Do the homework and lab.The earlier you start, the better. It's not enough to just mechanically plow through the exercises. You should ask yourself how these exercises relate to earlier material, and imagine how they might be changed (to make questions for an exam, for example.)
5. Don't procrastinate. The content builds upon what was taught in previous weeks, so if something is confusing to you in Week 2, Week 3 will become more confusing, Week 4 even worse, etc. Don't let the week end with unanswered questions. But if you find yourself falling behind and not knowing where to begin asking, come to office hours and work with a member of the teaching team to help you identify a good (re)starting point.

Getting help in the course

- If you have a question during lecture or lab, feel free to ask it! There are likely other students with the same question, so by asking you will create a learning opportunity for everyone.

- The teaching team is here to help you be successful in the course. You are encouraged to attend *office hours*[1] to ask questions about the course content and assignments. Many questions are most effectively answered as you discuss them with others, so office hours are a valuable resource. You are encouraged to use them!
- Outside of class and office hours, any general questions about course content or assignments should be posted on the class discussion forum Ed Discussion. There is a chance another student has already asked a similar question, so please check the other posts in Ed Discussion before adding a new question. If you know the answer to a question posted in the discussion forum, you are encouraged to respond!

Check out the Support page for more resources.

What to expect in the course

Lectures and labs

Lectures and labs are designed to be interactive, so you gain experience applying new concepts and learning from each other. My role as instructor is to introduce you to new methods, tools, and techniques, but it is up to you to take them and make use of them. A lot of what you do in this course will involve writing code, and coding is a skill that is best learned by doing. Therefore, as much as possible, you will be working on a variety of tasks and activities during the lectures and labs. You are expected to prepare for class by completing assigned readings, attend all lecture and lab sessions, and meaningfully contribute to in-class exercises and discussion. Additionally, some lectures will feature application exercises that will be graded based on completing what we do in class.

You are expected to bring a laptop, tablet, or any device with internet and a keyboard to each class so that you can participate in the in-class exercises. Please make sure your device is fully charged before you come to class, as the number of outlets in the classroom will not be sufficient to accommodate everyone.

Teams

You will be assigned to a team at the beginning of the semester. You are encouraged to sit with your teammates in lecture and you will also work with them in the lab sessions. All team members are expected to contribute equally to the completion of the group activities, labs and the final project. You will be asked to complete teamwork evaluations and self-reflections throughout the semester. Failure to adequately contribute to an assignment can result in a penalty to your score relative to the team's overall mark.

[1] Office hours are times the teaching team set aside each week to meet with students. Click here to learn more about how to effectively use office hours.

You are expected to make use of the provided GitHub repository as the central collaborative platform. Commits to this repository will be used as one of several metrics of each team member's relative contribution for each project.

Activities & Assessment

You will be assessed based on six components: application exercises, homework, labs, exams, project, and teamwork.

Labs

In labs, you will apply the concepts discussed in lecture to various data analysis scenarios, with a focus on the computation and communication. Most lab assignments will be completed in teams, and all team members are expected to contribute equally to the completion of each assignment. You are expected to use the team's Git repository in the course's GitHub organization as the central platform for collaboration. Commits to this repository will be used as a metric of each team member's relative contribution for each lab, and there will be periodic peer evaluation on the team collaboration. Lab assignments will be completed using Quarto, correspond to an appropriate GitHub repository, and submitted for grading in Gradescope.

The lowest lab grade will be dropped at the end of the semester.

Homework

In homework, you will apply what you've learned during lecture and lab to complete data analysis tasks and explain the underlying mathematics. You may discuss homework assignments with other students; however, homework should be completed and submitted individually. Similar to lab assignments, homework must be typed up using Quarto and GitHub and submitted as a PDF in Gradescope.

One homework assignment will be dedicated to a ***statistics experience***. The statistics experience is an opportunity to engage with statistics and data science outside of the classroom through podcasts, books, seminars, data analysis competitions, and other activities. As you complete these experiences, the goal is to consider how the material you're learning in the course connects with society more broadly.

The lowest homework grade will be dropped at the end of the semester.

Exams

There will be two exams in this course. Each exam will include a closed-notes in-class component and an open-note take-home component. Through these exams you have the opportunity to demonstrate what you've learned in the course thus far. The exams will focus on both conceptual understanding of the applied and mathematical content and application through analysis and computational tasks. The exams will be based on content in reading assignments, lectures, application exercises, homework, and lab assignments. More detail about the exams will be given during the semester.

Project

The purpose of the final project is to apply what you've learned to analyze an interesting data-driven research question. The project will be completed with your lab teams, and each team will present their work through a written report and presentation. More information about the project will be provided during the semester. You can learn more on the final project page.

Participation (Application exercises + teamwork)

Application exercises

You will get the most out of the course if you actively participate in class and when working with your team. Parts of some lectures will be dedicated to working on Application Exercises (AEs). AEs are submitted by pushing your work to the relevant GitHub repo.

AEs will be graded based on making a good-faith effort to attempt all questions covered in class. You are welcome to, but not required, to work on AEs beyond lecture.

Successful effort on at least 80% of AEs will result in full credit for AEs in the final course grade.

Teamwork

Given the collaborative nature of statistics and data science work, teamwork will be a key part of this course. You will work in teams for in-class activities, lab assignments, and the final course project. There will be periodic peer and self-evaluations to reflect on the team's collaboration. These evaluations will be counted as part of the participation grade.

Grading

The final course grade will be calculated as follows:

Category	Percentage
Homework	30%
Final project	15%
Labs	10%
Exams (2 Midterms)	40%
Participation (AEs + Teamwork)	5%

The final letter grade will be determined based on the following thresholds:

Letter Grade	Final Course Grade
A	>= 93
A-	90 - 92.99
B+	87 - 89.99
B	83 - 86.99
B-	80 - 82.99
C+	77 - 79.99
C	73 - 76.99
C-	70 - 72.99
D+	67 - 69.99
D	63 - 66.99
D-	60 - 62.99
F	< 60

Course policies

Duke Community Standard

All students must adhere to the Duke Community Standard(DCS): Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, and accountability. Citizens of this community commit to reflect upon these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity.

To uphold the Duke Community Standard, students agree:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself honorably in all my endeavors;and
- I will act if the Standard is compromised.

Academic honesty

TL;DR: Don't cheat!

- The homework assignments must be completed individually and you are welcomed to discuss the assignment with classmates at a high level (e.g., discuss what's the best way for approaching a problem, what functions are useful for accomplishing a particular task, etc.). However you may not directly share answers to homework questions (including any code) with anyone other than myself and the teaching assistants.
- You may not discuss or otherwise work with others on the exams. Unauthorized collaboration or using unauthorized materials will be considered a violation for all students involved. More details will be given closer to the exam date.
- For the projects and team labs, collaboration within teams is not only allowed, but expected. Communication between teams at a high level is also allowed however you may not share code or components of the project or team labs across teams.
- **Reusing code**: Unless explicitly stated otherwise, you may make use of online resources (e.g. StackOverflow) for coding examples on assignments. If you directly use code from an outside source (or use it as inspiration), you must explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism.
- **Use of artificial intelligence (AI)**: You should treat AI tools, such as ChatGPT, the same as other online resources. There are two guiding principles that govern how you can use AI in this course:[2] (1) *Cognitive dimension:* Working with AI should not reduce your ability to think clearly. We will practice using AI to facilitate—rather than hinder—learning. (2) *Ethical dimension*: Students using AI should be transparent about their use and make sure it aligns with academic integrity.
  - **AI tools for code:** You may make use of the technology for coding examples on assignments; if you do so, you must explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism. You may use these guidelines for citing AI-generated content.
  - **No AI tools for narrative:** Unless instructed otherwise, AI is **not** permitted for writing narrative on assignments. In general, you may use AI as a resource as you complete assignments but not to answer the exercises for you. You are ultimately responsible for the work you turn in; it should reflect your understanding of the course content.

[2] These guiding principles are based on *Course Policies related to ChatGPT and other AI Tools* developed by Joel Gladd, Ph.D.

If you are unsure if the use of a particular resource complies with the academic honesty policy, please ask a member of the teaching team.

Regardless of course delivery format, it is the responsibility of all students to understand and follow all Duke policies, including academic integrity (e.g., completing one's own work, following proper citation of sources, adhering to guidance around group work projects,and more).Ignoring these requirements is a violation of the Duke Community Standard. Any questions and/or concerns regarding academic integrity can be directed to the Office of Student Conduct and Community Standards at conduct@duke.edu.

Late work policy

The due dates for assignments are there to help you keep up with the course material and to ensure the teaching team can provide feedback in a timely manner. We understand that things come up periodically that could make it difficult to submit an assignment by the deadline. *Note that the lowest homework and lab assignment will be dropped to accommodate such circumstances.*

- Homework and labs may be submitted up to 2 days late. There will be a 5% deduction for each 24-hour period the assignment is late.
- The late work policy for exams will be provided with the exam instructions.
- The late work policy for the project will be provided with the project instructions.

Waiver for extenuating circumstances

If there are circumstances that prevent you from completing a lab or homework assignment by the stated due date, you may email me at maria.tackett@duke.edu before the deadline to waive the late penalty. In your email, you only need to request the waiver; you do not need to provide explanation. *This waiver may only be used once in the semester, so only use it for a truly extenuating circumstance.*

If there are circumstances that are having a longer-term impact on your academic performance, please let your academic dean know, as they can be a resource. Please let me know if you need help contacting your academic dean.

Regrade Requests

Regrade requests must be submitted on Gradescope within a week of when an assignment is returned. Regrade requests will be considered if there was an error in the grade calculation or if you feel a correct answer was mistakenly marked as incorrect. Requests to dispute the number of points deducted for an incorrect response will not be considered. Note that by

submitting a regrade request, the entire question will be graded which could potentially result in losing points.

No grades will be changed after the final project presentations.

Attendance policy

Every student is expected to attend and participate in lecture and labs. There may be times, however, when you cannot attend class. Lecture recordings are available upon request for students who have an excused absence. See the Lecture recording request policy for more detail. If you miss a lecture, make sure to review the material and complete the application exercise, if applicable, before the next lecture. Labs dedicated to completing the lab assignment and collaborating with your lab team. If you miss a lab session, make sure to communicate with your lab TA and teammates about how you can make up your contribution. If you know you're going to miss a lab session and you're feeling well enough to do so, notify your lab TA and teammates ahead of time.

More details on Trinity attendance policies are available here.

Lecture recording request

Lectures will be recorded on Panopto and will be made available to students with an excused absence upon request. Videos shared with such students will be available for a week after the lecture date. To request a particular lecture's video, please fill out the form at the link below. Please submit the form within 24 hours of missing lecture to ensure you have sufficient time to watch the recording. Please also make sure that any official documentation, such as STINFs, Dean's excuses, NOVAPs, and quarantine/removal from class notices from student health are also uploaded to the form.

https://forms.office.com/r/rvsgNcumRy

About one week before each exam, the class recordings will be available to all students. These recordings will be available until the start of the exam.

Accommodations

Academic accommodations

If you need accommodations for this class, you will need to register with the Student Disability Access Office (SDAO) and provide them with documentation related to your needs. SDAO will work with you to determine what accommodations are appropriate for your situation. Please note that accommodations are not retroactive and disability accommodations cannot

be provided until a Faculty Accommodation Letter has been given to me. Please contact SDAO for more information: sdao@duke.edu or access.duke.edu.

Religious accommodations

Students are permitted by university policy to be absent from class to observe a religious holiday. Accordingly, Trinity College of Arts & Sciences and the Pratt School of Engineering have established procedures to be followed by students for notifying their instructors of an absence necessitated by the observance of a religious holiday. Please submit requests for religious accommodations at the beginning of the semester so that we can work to make suitable arrangements well ahead of time. You can find the policy and relevant notification form here: trinity.duke.edu/undergraduate/academic-policies/religious-holidays

Academic and wellness support

Academic Resource Center

There are times may need help with the class that is beyond what can be provided by the teaching team. In those instances, I encourage you to visit the Academic Resource Center. The Academic Resource Center (ARC) offers free services to all students during their undergraduate careers at Duke. Services include Learning Consultations, Peer Tutoring and Study Groups, ADHD/LD Coaching, Outreach Workshops, and more. Because learning is a process unique to every individual, they work with each student to discover and develop their own academic strategy for success at Duke. Contact the ARC to schedule an appointment. Undergraduates in any year, studying any discipline can benefit! Contact ARC@duke.edu, 919-684-5917.

CAPS

Duke Counseling & Psychological Services (CAPS) helps Duke Students enhance strengths and develop abilities to successfully live, grow and learn in their personal and academic lives. CAPS recognizes that we are living in unprecedented times and that the changes, challenges and stressors brought on by the COVID-19 pandemic have impacted everyone, often in ways that are tax our well-being. CAPS offers many services to Duke undergraduate students, including brief individual and group counseling, couples counseling and more. CAPS staff also provides outreach to student groups, particularly programs supportive of at-risk populations, on a wide range of issues impacting them in various aspects of campus life. CAPS provides services to students via Telehealth. To initiate services, you can contact their front desk at 919-660-1000.

Important dates

- **January 8**: Classes begin
- **January 20**: Martin Luther King Jr. Day holiday.
- **January 22**: Drop/Add ends
- **March 10 - 14**: Spring break
- **March 26**: Last day to withdraw with "W"
- **November 27 - 29**: Thanksgiving recess
- **April 23**: Classes end
- **April 24 - 27**: Reading period
- **April 28 - May 3**: Final exam period

Click here for the full Duke academic calendar.