STA199: Introduction to Data Science and Statistical Thinking
001 — Summer Session I 2022

Instructor Information

Name: Bora Jin
Email: bora.jin@duke.edu
Office Hours & Location: MWTh 11am – 12:15pm, Old Chemistry 003

Teaching Assistant

Name: Camilla Yu
Email: yue.yu652@duke.edu
Office Hours & Location: TuF 12:15pm – 1:15pm, zoom

Class Information

Dates: May 11 – June 23
Lectures: MTuWThF 9:30am – 10:45am
Labs: TuF 11am – 12:15pm
Classroom: Old Chemistry 003
Websites: sakai.duke.edu
https://sta199-summer22.netlify.app
Textbooks: R for Data Science by Wickham & Grolemund (1st Ed. O'Reilly)
Introduction to Modern Statistics by Çetinkaya-Rundel & Hardin (1st Ed. OpenIntro)

Course Objectives

During this course, you will:

1. learn to explore, visualize and analyze data in a reproducible manner
2. gain experience in data wrangling, munging, exploratory data analysis, predictive modeling and data visualization
3. develop your own question(s) about data and use statistical techniques to answer the question(s)
4. practice effective oral and written communication of results

Important Dates:

May 11 (Wed): Classes begin
May 13 (Fri): Drop/add ends
May 30 (Mon): Memorial Day holiday. No classes
Jun 8 (Wed): Last day to withdraw with W
Jun 17 (Fri): Classes end

Jun 22 (Wed): Project report, repository, and slides due
Jun 23 (Thu): Project presentation

To find out more, see the full Duke academic calendar.

Grading

Grading Components

The course grade is determined by the following components:

Components	Percentage
Homework	25%
Labs	15%
Exam 1	17.5%
Exam 2	17.5%
Final Project	20%
Application Exercises	2.5%
Participation	2.5%

1. Homework

In homework, you will apply what you have learned during lecture and lab to complete data analysis tasks. You may discuss homework assignments with other students; however, homework should be completed and submitted individually. Homework must be typed up using R Markdown 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 are learning in the course connects with society more broadly.

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

2. Labs

In labs, you will apply the concepts discussed in lecture to various data analysis scenarios, with a focus on the computation. You will complete lab assignments individually, but can discuss with your classmates. You are expected to use your Git repository on the course's GitHub page as the central platform for work. Commits to this repository will be used as a component of the lab grade. Lab assignments will be completed using R Markdown, correspond to an appropriate GitHub repository, and submitted as a PDF to Gradescope.

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

3. Exams

There will be two, take-home, open-note exams. Through these exams you have the opportunity to

demonstrate what you have learned in the course thus far. Each exam will include small analysis and computational tasks related to the content discussed in lectures, application exercises, homework assignments, and labs. More details about the content and structure of the exams will be discussed during the semester.

4. Final Project

The purpose of the final project is to apply what you have learned throughout the semester to analyze interesting data-driven research questions. The project will be completed in teams, and each team will present their work during lecture on June 23. All team members are expected to contribute equally to the completion of the final project. You are expected to use the team's Git repository on the course's GitHub page as the central platform for collaboration. Commits to this repository may be used as a metric of each team member's relative contribution, and there will be peer evaluation on the team collaboration. **You must complete the final project and be in class to present it in order to pass this course.** More information about the project will be provided during the semester.

5. Application Exercises

Application exercises (AEs) give you an opportunity to practice using the statistical concepts and/or code discussed in lecture. They will typically be started during class and may be assigned to be completed by the next class meeting. Because these AEs are for practice, they will be graded based on completion, i.e., a good-faith effort has been made in attempting all parts. Successful on-time completion of at least 80% of AEs will result in full credit for AEs in the final course grade.

6. Participation

Class attendance and participation in lecture and lab is a firm expectation; frequent absences or tardiness will be considered a legitimate cause for grade reduction.

Final grades will be assigned according to the following scale:

Letter grade	Numeric grade	Letter grade	Numeric grade
A	≥ 93	C	[73, 77)
A–	[90, 93)	C–	[70, 73)
B+	[87, 90)	D+	[67, 70)
B	[83, 87)	D	[63, 67)
B–	[80, 83)	D–	[60, 63)
C+	[77, 80)	F	< 60

Course Policies

Academic Honesty

As a student in this course, you have agreed to uphold the Duke Community Standard as well as

the practices specific to this course. Any violations will automatically result in a grade of 0 on the assignment and will be reported to Office of Student Conduct for further action.

- **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 or explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism.
- On homework or lab assignments, you may not directly share (or copy) code or write up with other students. On the final project, you may not directly share (or copy) code or write up with another team. Unauthorized sharing of the code or write up will be considered a violation for all students involved.

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.

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 talk to me about it.

Attendance Policy Related to COVID Symptoms, Exposure, or Infection

Student health, safety, and well-being are the university's top priorities. To help ensure your well-being and the well-being of those around you, **please do not come to class if you have symptoms related to COVID-19, have had a known exposure to COVID-19, or have tested positive for COVID-19.** If any of these situations apply to you, you must follow university guidance related to the ongoing COVID-19 pandemic and current health and safety protocols. If you are experiencing any COVID-19 symptoms, contact student health at (919) 681-9355.

To keep the university community as safe and healthy as possible, you will be expected to follow these guidelines. Please reach out to me and your academic dean as soon as possible if you need to quarantine or isolate so that we can discuss arrangements for your continued participation in class.

Late Work

- For homework and lab assignments, there will be 10% deduction for each 24-hour period the assignment is late. Note that your lowest homework and lab will be dropped.
- If there are extenuating circumstances that prevent you from completing a homework or lab assignment by the stated due date, please let me know as soon as possible **before the deadline** to waive the late penalty.
- **Late work will not be accepted for exams, AEs, or the final project.**

Excused Absences

Students who miss a class due to a scheduled varsity trip, religious holiday, or short-term illness should fill out an online NOVAP, RHoliday or short-term illness form respectively. Note that these excused absences do not excuse you from assigned homework, it is your responsibility to make alternative arrangements to turn in any assignments in a timely fashion.

If you are faced with a personal or family emergency or a long-range or chronic health condition that interferes with your ability to attend or complete classes, you should contact your academic dean's office. See more information on policies surrounding these conditions at https://trinity.duke.edu/undergraduate/academic-policies/personal-emergencies. Your academic dean can also provide more information.

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. Therefore, you should attend office hours to ask a member of the teaching team about your grading feedback before submitting a regrade request. **No grades will be changed after the final project presentations.**

Logistics

Communication

- All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on the course website, https://sta199-summer22.netlify.app.
- Announcements will be emailed to the class through sakai. Please check your email regularly to ensure you have the latest announcements for the course.
- The teaching team is here to help you be successful in the course. You are encouraged to attend any of the office hours to ask questions as you study the course content and work through assignments. A lot of questions are most effectively answered in-person, so office hours are a valuable resource. Please use them!
- Outside of class and office hours, see the course slack for general questions and discussion.
- Please email me directly if you have any specific questions about grades or personal matters. Please refrain from emailing any course content questions (those should go on slack). **When you email me, please include "STA 199" in the subject line.**

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.

Additional Resources

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! Contant (919) 684-5917, or arc@duke.edu.

Duke Counseling & Psychological Services

Duke Counseling & Pyschological 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 COVID19 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.