Syllabus

Course description

This course introduces students to the discipline of statistics as a science of understanding and analyzing data. Themes include data collection, exploratory analysis, inference, and modeling. Focus on principles underlying quantitative research in social sciences, humanities, and public policy. Research projects teach the process of scientific discovery and synthesis and critical evaluation of research and statistical arguments. Readings give perspective on why in 1950, S. Wilks said, "Statistical thinking will one day be as necessary a qualification for efficient citizenship as the ability to read and write."

In this course, you will learn how to effectively make use of data in the face of uncertainty: how to collect data, how to analyze data, and how to use data to make inferences and conclusions about real world phenomena. Critiquing data-based claims and evaluating data-based decisions is at the core of this course. Throughout the course students acquire a conceptual understanding and mastery of statistical and quantitative reasoning tools in order to be able to make such critiques and evaluations.

In addition, students are presented with novel data sets and application examples on a daily basis, and they use these data to model outcomes and make inferences about unknown population characteristics. Students learn that the first step of any analysis is identifying the assumptions and conditions necessary to apply the statistical technique(s) required to answer the research question at hand. Students not only learn the mechanics of the quantitative analysis, but also how to interpret conclusions based on quantitative evidence in context of the data and the research questions as well as identifying limitations due to data collection and study design.

For the lab component of this course students prepare weekly lab reports presenting statistical analysis of real data. In addition, students complete two independent data analysis projects where they answer significant research questions via the analysis of real data using statistical inference and modeling tools.

Learning objectives

The course learning objectives are as follows:

1. Recognize the importance of data collection, identify limitations in data collection methods, and determine how they affect the scope of inference.

2. Use statistical software to summarize data numerically and visually, and to perform data analysis.

3. Have a conceptual understanding of the unified nature of statistical inference.

4. Apply estimation and testing methods to analyze single variables or the relationship between two variables in order to understand natural phenomena and make data-based decisions.

5. Model numerical response variables using a single or multiple explanatory variables.

6. Interpret results correctly, effectively, and in context without relying on statistical jargon.

7. Critique data-based claims and evaluate data-based decisions.

8. Complete research projects demonstrating mastery of statistical data analysis from exploratory analysis to inference to modeling.

Prerequisites

This course has no pre-requisites.

Workload

You are expected to put in ~6 hours of work / week outside of class. Some of you will do well with less time than this, and some of you will need more.

Tips for success

- Complete the reading before a new unit begins, and then review again after the unit is over.

- Be an active participant during lectures and labs.

- Ask questions - during class or office hours, or by email. Ask me, your TAs, and your classmates.

- Do the problem sets - start early and make sure you attempt and understand all questions.

- Start your project early and and allow adequate time to complete it.

- Give yourself plenty of time time to prepare a good cheat sheet for exams. This requires going through the material and taking the time to review the concepts that you're not comfortable with.

- Do not procrastinate - don't let a unit go by with unanswered questions as it will just make the following unit's material even more difficult to follow.

Textbook

Readings for the course will come from the following textbook which is **freely available online**. You do not need to purchase a physical copy of the book to succeed in this class.

`[ims]`: Mine Çetinkaya-Rundel and Jo Hardin. Introduction to Modern Statistics. 2nd edition. OpenIntro, 2024.

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 have a name that differs from those that appear in your official Duke records, please let me know! You'll be able to note this in the Getting to know you survey.
- 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 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: sta101-f23.github.io.

I will regularly send course announcements via email and Canvas, make sure to check one or the other of these regularly. If an announcement is sent Monday through Thursday, I will assume that you have read the announcement by 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.

Where to get help

- 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 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. Please use them!
- Outside of class and office hours, any general questions about course content or assignments should be posted on the course Slack. There is a chance another student has already asked a similar question, so please check the other posts on Slack before adding a new question. If you know the answer to a question posted on Slack, I encourage you to respond!

Check out the Support page for more resources.

I want to make sure that you learn everything you were hoping to learn from this class. If this requires flexibility, please don't hesitate to ask.

- You *never* owe me personal information about your health (mental or physical) but you're always welcome to talk to me. If I can't help, I likely know someone who can.

- I want you to learn lots of things from this class, but I primarily want you to stay healthy, balanced, and grounded during this crisis.

Lectures and lab

The goal of both the lectures and the labs is for them to be as interactive as possible. My role as instructor is to introduce you new 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 throughout each lecture and lab. Attendance will not be taken during class but you are expected to attend all lecture and lab sessions and meaningfully contribute to in-class exercises and discussion.

If you have a laptop or table, you may find it helpful to bring it to class so that you can take part in the in-class exercises. However, you will typically be working in groups, so as long as someone at your table has a laptop, that will be sufficient.

Assessments and grading

Attendance and participation

I expect you to be present at class meetings and to actively participate in discussions and activities. While your attendance and participation do not factor directly into your course grade, they do factor indirectly in that your learning is a direct result of engaging with the material. If you do need to miss class for any reason, you can check the class Moodle page to see what we covered. I will also post all slides.

Interactive tutorials

You will be assigned a number of interactive tutorials each week from the textbook. You will be asked to submit these on a weekly basis and graded on a check/no check basis.

Make sure to add your name and your **Net ID** before generating the hash. Submit these in Canvas.

If you've completed at least 80% of the tutorials, you'll get all available points for this component.

Labs

In labs, you will apply the concepts discussed in class to various data analysis scenarios. Labs will focus on both computation and conceptualization. Lab assignments will be completed using Quarto and submitted as PDF for feedback. While you may collaborate with others on lab assignments, you will need to submit your own final lab report.

Quizzes

We will have occasional short in-class quizzes. The purpose of these is to check-in where you are in your understanding of various concepts throughout the semester.

Projects

There will be a mid-semester prediction project and a final project. The prediction project will introduce you to conducting independent analyses and writing a formal report using a pre-specified data set. The final project allows you to explore a question and data set of your own. More details about the projects will be provided during the semester. Projects will be completed in teams.

You will be assigned to a different team for each of your two projects. 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 each project and you will be asked to evaluate your team members after each assignment is due. Failure to adequately contribute to an assignment will result in a penalty to your mark relative to the team's overall mark.

See Section for dates and times of project deadlines. Project deadlines cannot be changed. If you can't be in class for the final project presentation, you should drop this class.

Course policies

Collaboration

- While you are welcome to discuss lab assignments with classmate, your lab report should b

- For the projects, collaboration within teams is expected. Communication between teams at

- **Online resources:** I am well aware that a huge volume of R code is available on the web to solve any number of problems. Unless I explicitly tell you not to use something, the course's policy is that you may make use of any online resources (e.g., StackOverflow) but you must explicitly cite where you obtained any code you directly use (or use as inspiration). This is not a course in learning R, but rather interpreting the results of analysis.

- **Use of generative artificial intelligence (AI)**: You should treat generative AI, such as ChatGPT, the same as other online resources. There are two guiding principles that govern how you shoulud use AI in this course:[1] (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.

  - **AI tools for narrative:** Unless instructed otherwise, you may not use generative AI to write narrative on assignments. In general, you may use generative 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.

Regardless of course delivery format, it is the responsibility of all students to understand and follow all University of Portland 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).

Late work & extensions

Due dates for assignments are there to give you structure and to help you keep up with the course material. They also help me provide you with feedback in a timely manner. That said, I understand that things come up periodically that can make it difficult to complete an assignment by the deadline. Life happens!

If something comes up that prevents you from completing an assignment by the stated due date, just send me email and let me know when I can expect your work. You do not need to provide an explanation. In general, I will expect to receive your work within **one week**.

---

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

There are, however, two **hard deadlines** to be aware of: Oct. 11th (Friday before Fall break) and Dec. 6th (last day of classes). Except in unusual circumstances, I will not accept work after those dates.

Important dates

- **Tuesday, August 27:** Classes begin
- **Friday, August 30:** Last day to add/drop
- **Friday, October 13:** First hard deadline, Project 1 due
- **Monday, October 15**: Fall Break begins
- **Friday, November 17:** Exam 2 - Take home due
- **Monday, November 25:** Last day to Withdraw
- **Thursday, November 28:** Thanksgiving
- **Friday, December 6**: Last day of classes; 2nd hard deadline
- **Monday, December 9**: Section H: Project 2 presentations
- **Wednesday, December 11**: Section I: Project 2 presentations