Math 141 Syllabus

Professor Yan Cui

Fall 2024

Key Information

Lecture and Lab Instructor: Dr. Yan Cui (she/her) – please call me *Yan*

- Email: yanc@reed.edu
- Office: Library 392
- Office Hours: Tuesday 2-3pm; Friday 10:00-11:00am, or by appointment.

Meeting Times

Lecture (MWF)	Lab (Thur): 10:30-11:50am (PHYSIC 122)	1:40-3:00pm (ELIOT 126)	3:10-4:30pm (ELIOT 405)
9:00-9:50am (LIB 204)	F21	F22	F23
12:00-12:50pm (ELIOT 314)	F31	F32	F33

Lab Assistants

- Section 10:30-11:50am: TBD ()
- Section 1:40-3:00pm: TBD ()
- Section 3:10-4:30pm: TBD ()

Lab Assistant Office Hours

All office hours will be held in ETC 105B. XX, XX, and XX are the lab assistants for the sections of Math 141 taught by Dr. Michael Pearce. You are more than welcome to attend any of the office hours listed below, even if they are not for the lab assistant from your lab section!

	Mon	Tues	Wed	Thurs	Fri
4:30-5:30pm	XX	XX	XX	XX	XX
5:30-6:30pm	XX	XX	XX	XX	XX
6:30-7:30pm	XX	XX	XX	XX	XX

Course Information

Course Description

This course introduces foundational concepts in probability and statistics through theory, computation, and practice. We begin by studying foundational concepts in statistical thinking, including appropriate data summary and visualization, probability, and study design. Then, we turn to probability theory and statistics, by studying how to use data to recognize patterns and associations between variables in a linear regression framework, and understand how to quantify our uncertainty in any estimated relationships (a process called *statistical inference*). Inference will be explored through both simulation-based and theoretical lenses. Throughout the semester, we will consider applications to a variety of scientific disciplines, the history of statistics, and ethical considerations in modern statistical practice. The course culminates in a data analysis project that allows students to practice verbal and written communication of statistical results.

Course Learning Outcomes

By the end of the course, you will be able to:

- Translate a research problem from a given dataset into questions that you can answer with appropriate graphics, summary statistics, statistical models, and/or statistical tests.
- Identify a study's structure, and explain how the structure influences potential conclusions.
- Construct graphics and compute summary statistics from data, and represent what they summaries do and do not reveal about the data.
- Perform linear regression and statistical inference, and interpret their results using the language of probability.
- Be aware of ethical issues in statistics, and articulate how the improper use of statistical techniques can misinform and mislead others.
- Communicate the results of a statistical analysis to statistical *and* non-statistical audiences.
- Use `R` and `R Markdown` to develop a reproducible workflow of data management, graphics, and statistical analyses.

Distribution Requirements

This course can be used towards your **Group III, "Natural, Mathematical, and Psychological Science"** requirement. It accomplishes the following learning goals for the group:

1. Use and evaluate quantitative data or modeling, or use logical/mathematical reasoning to evaluate, test or prove statements.
2. Given a problem or question, formulate a hypothesis or conjecture, and design an experiment, collect data or use mathematical reasoning to test or validate it.
3. Collect, analyze, and interpret data.

This course does **not** satisfy the "primary data collection and analysis" requirement.

Textbooks

This class will primarily use three textbooks for readings, all of which are *free and available online*:

- Statistical Inference via Data Science: a ModernDive into R and the Tidyverse, 1st Edition by Ismay and Kim
- Introduction to Modern Statistics, 1st Edition by Çetinkaya-Rundel and Hardin
- OpenIntro Statistics, 4th Edition by Diez, Çetinkaya-Rundel, and Barr (click *"Read Free Sample"*)

Communication and Resources

The following web-based resources will be used for communicating and sharing class information:

- Course Moodle: *announcements, documents, assignments, lecture notes*

- Gradescope: *homework, labs*

Technology

We will make very frequent use of the `R` programming language (and its user-interface, `RStudio`) to manage data, create graphics, run statistical models, and perform statistical tests. All homework and labs will be completed in `RStudio`, which is free to use and can be accessed online through the Reed `RStudio` server: https://rstudio.reed.edu/. Alternatively, you may install `R` and `RStudio` locally on your machine (Please find this link to install R and link to install RStudio).

I would recommend using the online server as we get started. Please let me know ASAP if you do not have access to a personal computer!

Avenues for Help

Office Hours

Attending office hours, either with myself or the lab assistants, is a great way to get help or advice! Office hours are unstructured times in which you may simply "drop-in" and seek assistance with material from lecture, homework, or labs. You may also come to office hours for advice about future statistics courses or research problems. **In the first few weeks of the course, I especially welcome you to come to my office hours to introduce yourself!** I am fully committed to supporting your academic goals and will do everything I can to help you succeed.

A schedule of office hours is on the first page of this syllabus. No appointment is necessary. If my scheduled office hours do not work for your schedule, you may book an appointment using this link: My booking page

Help Outside Office Hours

Outside of office hours, there are many ways to get help:

- Form a regular study group with your peers. Reviewing key concepts from class, discussing strategies for solving homework problems, and looking over code are superb ways to strengthen your knowledge of statistics.
- Post questions to your classmates on the Moodle discussion board.
- Email me directly questions. I will try to respond within 24 hours, although my response may be delayed over the weekend.

In general, *in-person help/assistance is the most productive and efficient.* This is especially true when it comes to coding issues. I encourage you to first seek in-person help, if possible!

Learning Strategies

This course moves at a fairly brisk pace. If you find yourself falling behind, here are some suggestions for getting back on track:

- Carefully review the recommended readings *before* class.
- Actively participate in every class and lab.
- Visit office hours. You can ask questions, or just stop by to review key concepts from class.
- Form a study group. Take turns preparing mini-lectures/reviews for each other.

Course Structure and Assessment

Your grade in the class will be determined by your proficiency in each of the *Course Learning Outcomes*, as demonstrated via the following course components:

1. Attendance and Participation (~10%)

2. Homework Assignments (~25%)
3. Lab Assignments (~25%)
4. Project (~20%)
5. Quizzes and Final Exam (~20%)

Each component is described in greater detail below.

Attendance and Participation

I expect all students to attend lecture and lab regularly. I also expect all students to actively participate (e.g., ask and respond to questions, engage in class/small-group discussion). The topics covered in this course are challenging; attendance and participation will contribute greatly to your learning and the learning of your fellow classmates! To reflect their importance, attendance and participation will constitute about 10% of your course grade.

I understand that attendance is not always possible due to a variety of important reasons. As such, I will not penalize your grade for 2 absences during the semester. Up to 4 absences are permitted for students who have requested accommodations.

Homework Assignments

Homework assignments provide formative opportunities to practice and develop skills in data analysis, written statistical communication, probability theory, and interpretation of statistical results. Your effort and performance in the weekly homework assignments will be the primary method to demonstrate learning of course concepts. Thus, homework assignments constitute about 25% of your course grade.

Homework assignments will be posted on most Fridays to the course Moodle. Your completed homework assignment should be *submitted to Gradescope by the time class starts on the following Friday.* Up to twice throughout the term, you may request a two day extension on your homework assignment by emailing the professor. Except in extraordinary circumstances, **requests must be made prior to an assignment's due date.** I anticipate a total of 9 homework assignments; there will be no homework due the week of a quiz.

Lab Assignments

In addition to weekly homework assignments, lab assignments provide regular and formative opportunities to apply the concepts learned in class to real datasets, as well as practice coding in the R programming language. Due to the high importance of computation in modern statistical practice, lab assignments constitute about 25% of your course grade.

You will work on a lab assignment during each lab section on Thursdays. Your completed lab assignment should be *submitted to Gradescope by the time class starts on the following Wednesday*, though lab time should be enough to complete *most* of the assignment. Up to twice throughout the term, you may request a two day extension on your lab assignment by emailing the professor. Except in extraordinary circumstances, **requests must be made prior to an assignment's due date.** I anticipate a total of 10 lab assignments; there will be no lab assignment due the week after a quiz.

Project

Throughout the term, you will work in groups of 4-6 students on a project that answers a significant research question using real-world data by implementing the techniques covered in our class. If you would like, you may also use more advanced methods from supplementary sources. The project will culminate in a **15-minute in-class presentation in the last week of lecture, and a technical report to be turned in during Finals week**. The project is an excellent opportunity to combine theoretical and practical skills in statistics and to practice the collaborative nature of modern data analysis. Thus, the project will constitute roughly 20% of your course grade.

To allow for regular feedback from the instructor and your peers, you will submit three "in-progress" assignments during the semester. These assignments iteratively develop your final project report. To limit

excessive workload, each project assignment will replace a lab assignment that week! Below is a schedule for the projects.

1) **Wednesday 10/16** — In-Progress Assignment 1: Exploratory Data Analysis (10% of project grade)
    - Propose a research question, evaluate ethics of your data analysis, perform an exploratory data analysis, and fit a preliminary regression model.
    - Each group will select a dataset from an instructor-provided list. If you'd like to study another dataset, you must request to do so in writing no later than Wednesday, 10/2.
2) **Wednesday 11/13** — In-Progress Assignment 2: Revision (10% of project grade)
    - Incorporate the edits I suggest from your first project assignment.
3) **Friday 12/6** — In-Progress Assignment 3: Report Draft and Peer Review (15% of project grade)
    - Complete a first draft of your project report. The report will include all contents of your previous assignment, plus (i) an inferential analysis of your regression model that includes a hypothesis test and a confidence interval, and (ii) a discussion of your results.
    - This project draft will be peer-reviewed in class this same day.
4) **Monday 12/9 and Wednesday 12/11** — In-Class Presentations (25% of project grade)
    - The goal of the presentation is to practice verbal communication of scientific results. More details will be provided at a later date.
5) **Friday 12/13** — Technical Report Due (40% of project grade)
    - The final project report should reflect the requirements of in-progress assignment 3. After peer review, I expect each group will only need to make minimal changes! More details will be provided at a later date.

Quizzes and Final Exam

To encourage holistic learning of the course material, the course includes two in-class quizzes and a final exam. Each quiz/exam will be cumulative, but focus on recent course material. The quizzes will occur on Thursday 10/10 and Thursday 11/7 during your lab section. The quizzes will be 50-minutes in length, allowing students with 1.5 time accommodations the ability to complete the quiz during the lab section. The final exam will occur in-person during finals week (exact time/location to be determined).

I hope to minimize the stress associated with quizzes/exams and provide all students the opportunity to demonstrate their learning. Thus, quizzes and the final exam will together constitute only ~20% of your course grade.

Course Environment and Academic Integrity

I am absolutely committed to fostering a friendly and inclusive classroom environment in which all students have an equal opportunity to learn and succeed. If you experience any barriers to learning, please come to me or a college administrator with your concerns.

- **Code of Conduct** (adapted from the Contributor Covenant, version 2.0.): I expect all members of Math 141 to make participation a harassment-free experience for everyone, regardless of age, body size, visible or invisible disability, ethnicity, sex characteristics, gender identity and expression, level of experience, education, socio-economic status, nationality, personal appearance, race, religion, or sexual identity and orientation.

  I expect everyone to act and interact in ways that contribute to an open, welcoming, diverse, inclusive, anti-racist, and healthy community of learners. Examples of unacceptable behavior include: using sexualized language or imagery, making insulting or derogatory comments, harassing someone publicly or privately, or other unprofessional conduct.

Instead, you can contribute to a positive learning environment by demonstrating empathy and kindness, being respectful of differing viewpoints and experiences, and giving and gracefully accepting constructive feedback.

- **Names and Pronouns**: Everyone deserves to be addressed respectfully and correctly. You are welcome to update your preferred name and gender pronouns at any time. I will also provide an opportunity to share your preferred name and pronouns in-class on the first day!
- **Disability and Accesibility**: If you are a student with a documented disability in need of accommodations, I encourage you to reach out to Reed's Disability and Accessibility Resources (DAR) Office to make the necessary arrangements. If you already have accommodations in place, please feel free to make an appointment with me to discuss your accommodation needs. I am committed to making this course accessible to all students.
- **Academic Integrity**: Unless otherwise noted, students are allowed and encouraged to collaborate on assignments. However, any work that you turn in must be your own. You are welcome to use internet resources to supplement content we cover in this course, with the exceptions of (1) direct solutions to homework or lab problems and (2) generative AI for any purpose except when explicitly permitted.

  These rules are in place to foster your learning! It's one thing to seek out assistance on the internet (practical and modern), and another to copy answers from the internet (no actual knowledge accrued). Breaking this policy will constitute an Honor Principle violation.

Tentative Schedule

Below is a *tentative* course schedule. Up-to-date information can be found on the course Moodle.

Readings for each lecture are in parentheses. Abbreviations below:

- *MD* — Statistical Inference via Data Science: a ModernDive into R and the Tidyverse
- *IMS* — Introduction to Modern Statistics
- *OpenIntro* — OpenIntro Statistics

UNIT I: STATISTICAL THINKING

- Wednesday (9/4): Thinking Like a Statistician
- Thursday (9/5): *Lab 1: Introduction to R and RStudio* (MD 1.1-1.4)
- Friday (9/6): Intro to Data and Data Visualization (MD 2.0-2.1)

- Monday (9/9): The 5 Named Graphs (MD 2.2-2.9)
- Wednesday (9/11): Summary Statistics (IMS 5.2-5.6)
- Thursday (9/12): *Lab 2: Data Visualization and Summary* (MD 3.0-3.1, 3.3-3.4)
- Friday (9/13): Probability I (OpenIntro 3.1-3.2)

- Monday (9/16): Probability II (OpenIntro 3.4-3.5)
- Wednesday (9/18): Sampling Methods (IMS 2.1)
- Thursday (9/19): *Lab 3: Probability*
- Friday (9/20): Study Design and Ethics (IMS 2.2-2.3)

UNIT II: REGRESSION

- Monday (9/23): Linear Models I (IMS 7.0-7.2.2)
- Wednesday (9/25): Linear Models II (IMS 7.2.3-7.2.5)
- Thursday (9/26): *Lab 4: Sampling and Regression in R*
- Friday (9/27): Linear Models III (IMS 7.2.6, 8.1)

- Monday (9/30): Multiple Linear Regression I (IMS 8.2, 24.6.1)
- Wednesday (10/2): *Activity Day:* Multiple Linear Regression II (review readings from Unit II)
- Thursday (10/3): *Lab 5: More Linear Regression in R*
- Friday (10/4): No class (optional project work day)

UNIT III: FOUNDATIONS OF INFERENCE

- Monday (10/7): Sampling Distributions I (MD 7.1-7.2)
- Wednesday (10/9): Sampling Distributions II (MD 7.3-7.5)
- Thursday (10/10): *Quiz I (Units I-II)*
- Friday (10/11): Bootstrapping (MD 8.0-8.2)

- Monday (10/14): *Activity Day:* Bootstrap vs. Sampling Distributions (MD 8.7.1)
- Wednesday (10/16): Confidence Intervals I (MD 8.3), *Project Assignment 1 due*
- Thursday (10/17): *Lab 6: Sampling Distributions*
- Friday (10/18): Confidence Intervals II (MD 8.4-8.5)

- *Monday (10/21) – Friday (10/25): Fall Break*

- Monday (10/28): Hypothesis Testing I (MD 9.0-9.3)
- Wednesday (10/30): Hypothesis Testing II (MD 9.4-9.5, 9.6.3)
- Thursday (10/31): *Lab 7: Statistical Inference in R*
- Friday (11/1): Review of Unit III (review readings from Unit III)

UNIT IV: THEORY-BASED INFERENCE

- Monday (11/4): Random Variables I (OpenIntro 4.1, 4.3)
- Wednesday (11/6): Random Variables II (no reading)
- Thursday (11/7): *Quiz II (Unit III)*
- Friday (11/8): The Central Limit Theorem (IMS 13.1-13.3)

- Monday (11/11): Introduction to Theory-Based Inference (IMS 16.0-16.3)
- Wednesday (11/13): Inference for Proportions (IMS 17.0-17.3), *Project Assignment 2 due ACTIVITY!!*
- Thursday (11/14): *Lab 8: Central Limit Theorem*
- Friday (11/15): Inference for Means I (IMS 19.0-19.3)

- Monday (11/18): *Activity Day:* Inference for Means II (IMS 20.0-20.4)

UNIT V: INFERENCE FOR REGRESSION

- Wednesday (11/20): Inference for Regression I (IMS 24.0-24.5)
- Thursday (11/21): *Lab 9: Inference for Proportions and Means*
- Friday (11/22): Inference for Regression II (IMS 24.6-24.7)

- Monday (11/25): Inference for Regression III (IMS 25.0-25.2)
- Wednesday (11/27): *Activity Day:* Inference for Regression IV (review readings from Unit V)
- *Thursday (11/28): Thanksgiving (no class)*
- *Friday (11/29): Native American Heritage Day (no class)*

- Monday (12/2): Review of Unit V (review readings from Unit V)
- Wednesday (12/4): Project Work
- Thursday (12/5): *Lab 10: Inference for Regression in R*
- Friday (12/6): Project Peer Review, *Project Assignment 3 due*

- Monday (12/9): Project Presentations
- Tuesday (12/10): Final Exam Review
- Wednesday (12/11): Project Presentations
- Friday (12/13): *Project Reports due*

- Finals Week: *In-Person Final Exam (date TBD)*