



STAT220: INTRODUCTION TO DATA SCIENCE

Spring 2026

IMPORTANT INFO

- **Meetings:** MWF 4a, MW 12:30 - 1:40, F 1:10-2:10, CMC 102
- **Professor:** Emily Kurtz, ekurtz@carleton.edu, CMC 225
- **Office Hours:** M 2-3, T 10:30 - 11:30 (by appointment only), W 11:30-12:30, F 12-1, CMC 225
- **Texts**
 - *R for Data Science* (2nd Ed) by Wickham, Çetinkaya-Rundel, Grolemund
 - *Modern Data Science with R* by Baumer, Kaplan, Horton
 - *Fundamentals of Data Visualization* by Wilke
- **Software**
 - Maize RStudio Server
 - R (optional), free
 - RStudio (optional), free

COURSE DESCRIPTION

Stat 220 will cover the computational side of statistics that is not typically taught in an intro or methodology focused course like regression modeling. Most of the data you encounter in introductory stats course are contained in small, tidy .csv files with rows denoting your cases and columns containing your variables. Most of the messiness to these data may have been some missing values (NAs). In this course, we'll learn how to extract information from data in its "natural" state, which is often unstructured, messy and complex. To do this, we will learn methods for manipulating and merging data in standard and non-standard formats, data with date, time, or geolocation variables, text processing and regular expressions, and scraping the web for data. To effectively communicate the information contained in these data, we will cover advanced data visualization methods, including methods for creating interactive graphics. We will primarily use the statistical software R in this course, and cover best practices for reproducible analyses and sharing code.

COURSE OBJECTIVES

After completing this course, you should be able to demonstrate your competency in each of the following areas:

- **Develop** research questions that can be answered by data
- **Acquire** data by importing different file types into R, accessing data through API's, and scraping data from the web
- **Wrangle** common types of data into the form that is needed for analysis
- **Visualize** data to provide insight and uncover relationships and patterns
- **Communicate** your findings in written or oral format
- **Document** your code and **collaborate** across coding projects

COURSE COMPONENTS

MEETINGS

There will be three course meetings per week (Mondays, Wednesdays, and Fridays). Daily attendance and active participation is expected. Course meetings will combine demonstrations/lecture and in-class group exercises. On most days, I'll ask you to complete a reading or watch a short video before class.

ASSIGNMENTS

Homework will be distributed via GitHub. You will submit homework assignments via Gradescope. You will use quarto for all assignments and submit all necessary work for each assignment on GitHub.

PORTFOLIO PROJECTS

Portfolio project require you to integrate several smaller computational tasks and require clear communication of the proposed solution or findings to a broader audience. You will typically work alone or in pairs.

FINAL PROJECT

The final project is a capstone experience synthesizing everything you've learned over the course of the term. This is an opportunity for you to exercise your creativity and create something meaningful. The final project is open-ended and more details will follow. You will work in small groups of about 3.

COMMUNICATION

Assignments and slides will be shared publicly on our course website. Grades will be posted on Gradescope. Any time-sensitive announcements will be sent via Moodle. It is your responsibility to make sure that your notification settings allow time-sensitive announcements to reach you.

GRADING POLICIES

Grades are an imperfect measure of learning, especially when we are all coming to class with different backgrounds and experience levels. This course is designed to reward you for consistently participating, staying on top of the course material, and trying your

hardest. The grading scheme in this class is designed to reflect how we learn “in the real world”: by making mistakes, receiving feedback, and making changes.

HOW EACH ASSIGNMENT IS GRADED

Each assignment will include a short rubric with the specifications of how I will evaluate your work. In general, these are the “marks” that you can receive on each course component:

	How it’s evaluated	How it’s recorded
Home-work Problems	Completeness, Correctness, and Effort	Successful or Not Successful
Portfolio Projects	Completeness, effort, correctness, and communication quality	Excellent, Successful, or Revision
Final Project	Completeness, effort, correctness, and communication quality	Excellent, Successful, or Not Successful

EARNING A COURSE GRADE

Your course grade is assigned using the table below. Each row indicates the minimum percentage of “Successful” results needed to satisfy the requirement of that grade. To earn a grade, complete *all* requirements listed in the row for that grade.

	Homework Problems	Portfolio Projects (4 total)	Final Project
A	85%	2 Excellent, 2 Successful	Excellent
B	75%	4 Successful	Successful
C	65%	3 Successful	Successful
D	50%	2 Successful	Successful

PLUS/MINUS GRADES

“Plus” and “minus” grades will be given if you complete all the requirements for a base letter grade **and** make sufficient progress toward the next grade. Below is an overview:

- If “B” base grade:
 - *and* A in one bin: “B+”
 - *and* A in two bins: “A-”
- If “C” base grade:
 - *and* at least B in one bin: “C+”
 - *and* at least B in two bins: “B-”

- If “D” base grade:
 - *and* at least C in one bins: “D+”
 - *and* at least C in two bins: “C-”
- If “F” base grade:
 - *and* at least D in two bins: “D-”

Example: Mira finishes the course with 91% of homework problems successfully completed, 3/4 portfolio projects marked “excellent”, and a “Successful” final project. Mira satisfies everything in the “B” row, and meets the “A” threshold for two bins (homework and portfolio projects). Mira earns an A-.

ATTENDANCE AND PARTICIPATION POLICY

I expect daily attendance and participation. Missing >5 class meetings or ongoing, consistent issues with being on-task will result in a 1/3 grade deduction.

IMPORTANT POINTS ABOUT THIS GRADING SYSTEM

- Different categories of coursework do not “average together”: you can’t make up for less-than-great work on portfolio projects by doing very well on quizzes, for instance. Each course grade requires *consistent quality across all bins* to earn the grade
- You do not have to do everything. If you want an “A” in the class, for example, you don’t have to do every portfolio assignment perfectly, just do at least two very well and two pretty well.

TOKENS

You will each get 3 tokens for the term. If using a token for an extension, you must submit the token request form **before** the assignment/project/check-in is due. If using a token for a portfolio project revision, you must submit the form within a week of receiving your grade on the project. You will receive a confirmation email of your token request submission as a receipt that one token has been spent.

- Turning in a token provides either (a) a 48-hour extension on a homework assignment, portfolio project, or final project milestone check-in, or (b) a revision on a portfolio project
- Tokens cannot be used for the final project due-date
- You can only revise a portfolio project; you cannot revise homeworks, the final project, or final project milestone check-ins.
- You can revise the same portfolio project multiple times, if needed, but you must spend a token each time.
- A portfolio project must be completed with a good-faith effort to be eligible for revision. If I deem a submission to be “not assessable” due to a lack of effort, then it cannot be revised
- All revisions must be submitted by 5:00pm on the last day of class

ACADEMIC INTEGRITY

You are expected to follow Carleton’s policies regarding academic integrity. I encourage you to *discuss* homework problems with others to figure out tough problems. You should code and write up your solutions on your own. You should collaborate with your

teammates on projects, and should use external resources for background research and debugging, but all work should be original.

Large-language models (e.g. ChatGPT, Gemini, etc.) should only be used for coding or debugging help after you've attempted to code the problem on your own, and you should never type homework problems directly into a prompt. You should never copy and paste any course materials *into* a large-language model, and you should never copy and paste anything *out of* a large-language model into your course materials. Copying, paraphrasing, summarizing, or submitting work generated by anyone but yourself without proper attribution is considered academic dishonesty (this includes output from LLMs).

"AI" tools are new for all of us and it's OK to have questions about what is and isn't appropriate. Please ask if you are unsure of whether or not your actions are complying with the assignment instructions. Always default to acknowledging any help received. Cases of suspected academic dishonesty are handled by the Provost's Office and I am obligated to report any suspected violations of this policy.

COMMITMENT TO AN INCLUSIVE AND COLLABORATIVE ATMOSPHERE

We all come to class with different backgrounds and experiences, and this diversity makes our class environment richer. I value diversity and inclusion, and am committed to a climate of mutual respect and full participation in and out of the classroom. This class strives to be a learning environment that is usable, equitable, inclusive and welcoming, regardless of race, ethnicity, religion, gender and gender identities, sexual orientation, ability, socioeconomic background, and nationality. If you anticipate or experience any barriers to learning, please discuss your concerns with me.

RESOURCES

ACCOMMODATIONS: Carleton College is committed to providing equitable access to learning opportunities for all students. The Office of Accessibility Resources (Henry House, 107 Union Street) is the campus office that collaborates with students who have disabilities to provide and/or arrange reasonable accommodations. If you have, or think you may have, a disability, please contact OAR@carleton.edu to arrange a confidential discussion regarding equitable access and reasonable accommodations. You are also welcome to contact me privately to discuss your academic needs. However, all disability-related accommodations must be arranged, in advance, through OAR.

STATS LAB: The Stats Lab (CMC 304) offers drop-in help R/RStudio help sessions run by friendly and knowledgeable lab assistants on most weekday evenings and some weekend times. Lab assistants may not be familiar with the 220 material, but you are still welcome to drop in and ask basic R questions.

TUTORS: If you find you need more support than office hours and the stats lab can provide, the Academic Support Center offers peer tutoring on the basis of referrals, requests, and availability of tutors. You can request tutoring through a form on their website, or discuss your needs with me and I can submit a referral.

TITLE IX: Please be aware that all faculty are “responsible employees”, which means that if you tell me about a situation involving sexual harassment, sexual assault, dating violence, domestic violence, or stalking, I must share that information with the Title IX Coordinator. Although I have to make this notification, you will control how your case will be handled, including whether or not you wish to meet with the Title IX coordinator or pursue a formal complaint.

And finally....

Take care of yourself. Do your best to maintain a healthy lifestyle this term. Get some fresh air occasionally, especially when coursework has you frustrated. Take some time to rest, but don’t be tempted by excessive short-term dopamine hits that leave you even further behind. Wear a mask when you are sick, and get enough sleep. Your physical and mental health is more important than your grade in this course. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. For more information, see Student Health and Counseling (SHAC), the Office of Health Promotion, or the Office of the Chaplain. If you are experiencing physical or mental health symptoms as a result of coursework, please speak with me so we can address the problem together.