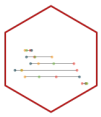


INFO 3312/5312 - Data Communication

Syllabus

Overview

Instructor

- Dr. Benjamin Soltoff
- Office: 284 CIS Building
- Email: soltoffbc@cornell.edu
- Office hours: Virtually, by appointment

Course logistics

- Meets online asynchronously from June 1 - June 18
- 3 credits, offered for a letter grade
- Prerequisites: INFO 2950/2951 or INFO 5001. Prior experience with R and Git/GitHub is required.

> ! All times are in Eastern Time (ET)
>
> Regardless of your physical location, all deadlines and times for this course are in Eastern Time (ET). Please make sure to convert to your local time zone as needed.

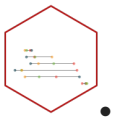
Course description

Data scientists often present information to disseminate their findings. This course introduces theories and applications of communicating with data, with an emphasis on visualizations. To support this approach, we will focus on the what, why, and how of data visualization. “What” focuses on specific types of visualizations for a particular purpose, as well as tools for constructing these plots. In “how” we will focus on the process of generating a data visualization from pre-processing the raw data, mapping attributes of the data to plot aesthetics, strategically determining how to define the visual encoding of the data for maximal accessibility, and finalizing the visualization to consider the importance of visual appeal. In “why” we discuss the theory tying together the “how” and the “what”, and consider empirical evidence of best-practices in data communication.

Course learning objectives

By the end of the semester, you will...

- Implement principles of designing and creating effective data visualizations.



- Evaluate, critique, and improve upon one's own and others' data visualizations based on how good a job the visualization does for communicating a message clearly and correctly.
- Post-process and refine plots for effective communication.
- Master using R and a variety of modern data visualization packages to reproducibly create data visualizations.
- Work reproducibly individually and collaboratively using Git and GitHub.

Office hours

Office hours are held virtually on Zoom by appointment.

Textbooks

All books are **freely available online**.

ggplot2: Elegant Graphics for Data Analysis	Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen	Springer, 3rd edition (in progress)	Hard copy only available of 2nd edition
Fundamentals of Data Visualization	Claus O. Wilke	O'Rielly Media, 2019	Hard copy available

Course community

We want you to feel like you belong in this class and are respected. Cornell University (as an institution) and we (as human beings and the instructor of this course) are committed to full inclusion in education for all persons. If for any reason you feel that we have failed these goals, please either let us know or report it, and we will address the issue.

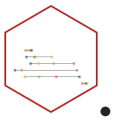
Services and reasonable accommodations are available to persons with temporary and permanent disabilities, to students with DACA or undocumented status, to students facing mental health or other personal challenges, and to students with other kinds of learning challenges. Please feel free to let me know if there are circumstances affecting your ability to participate in class. Some resources that might be of use include:

- Office of Student Disability Services: https://sds.cornell.edu
- Cornell Health CAPS (Counseling & Psychological Services): https://health.cornell.edu/services/counseling-psychiatry
- Undocumented/DACA Student support: https://dos.cornell.edu/undocumented-daca-support/undergraduate-admissions-financial-aid

Academic accommodations

We want all students to have the opportunity to be successful in this course. Accommodations can help provide some flexibility and equitable classroom access.

Per university policy, this course provides the following accommodations:



- Disability Accommodations
- Religious-Observance Accommodations
- Title IX Accommodations
- Varsity Athlete Accommodations
- Medical Accommodations
- Military Service
- Other Accommodations

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.

Student Disability Services is available to ensure that students are able to engage with their courses and related assignments. Students should be in touch with Student Disability Services to request or update accommodations under these circumstances.

If you have an approved SDS accommodation, please send a copy of this letter to the instructors at soltoffbc@cornell.edu so we can ensure your accommodations are implemented in this course.

Communication

All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on the course website: info3312.infosci.cornell.edu.

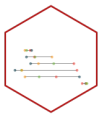
Announcements will be posted through Canvas Announcements periodically. Please check Canvas (or ensure Canvas announcements are forwarded to your email) to ensure you have the latest announcements for the course.

Email

If there is a question that's not appropriate for the public forum, please email us at soltoffbc@cornell.edu. Barring extenuating circumstances, we will respond to INFO 3312/5312 emails within 24 hours Monday - Friday. Response time may be slower for emails sent Friday evening - Sunday.

Where to get help

- If you have a question, **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.
- CayugaViz is a course-specific AI chatbot that you can ask questions about the course content, assignments, and logistics. It is trained on the course materials and is a great resource for getting help with the course.
- Any general questions about course content or assignments should be posted on the course discussion forum. There is a chance another student has already asked a similar question, so please check the other posts on GitHub Discussions before adding a new



question. If you know the answer to a question posted on the discussion board, I encourage you to respond!

- If you have a question that's not appropriate for the public forum, please email us at soltoffbc@cornell.edu.
- I do not have set office hours for this course, but I am happy to meet with you virtually to discuss any questions you have about the course or troubleshoot assignments. Please email me to set up an appointment. I can usually meet within 24 hours of your request.

Activities & Assessment

The activities and assessments in this course are designed to help you successfully achieve the course learning objectives. They are designed to follow the **Prepare, Practice, Perform** format.

- **Prepare**: Includes reading assignments to introduce new concepts and ensure a basic comprehension of the material.
- **Practice**: Includes application exercises where you will begin to apply the concepts and methods introduced in the prepare assignment.
- **Perform**: Includes homeworks and projects. These assignments build upon the prepare and practice assignments and are the opportunity for you to demonstrate your understanding of the course material and how it is applied to communicate effectively using real-world data.

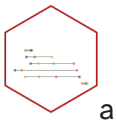
! Your responsibilities for your education

My role as the instructor is to provide you with the resources and support you need to succeed in this course. However, your success ultimately depends on your active participation and engagement in the course. You are expected to:

- Complete the assigned readings and lectures on time.
- Actively participate in the course.
- Complete all assignments on time and to the best of your ability.
- Seek help when needed, whether from the course staff, classmates, or other resources.
- **Take responsibility for your own learning and development throughout the course.**

Readings (Prepare)

Each day Monday-Friday you are expected to complete a set of assigned readings. These will primarily be lecture notes derived from the in-person format of this course, supplemented with external readings. The readings are designed to help you prepare for the active components of the class (e.g. application exercises, homeworks, projects). I don't assign



an absurd volume of readings, nor do I expect you to fully and completely understand every single thing you've read. But you should have a familiarity with the concepts and techniques.

Application exercises (Practice)

Every day you will have some set of **Application Exercises** (AEs). These exercises will give you an opportunity to apply the concepts and techniques introduced in the prepare assignment. Sometimes they are baked into the lecture notes, and sometimes they are separate documents.

Due to the nature of the online format and because these AEs are for practice, they are not graded. However completing them will be very advantageous for you to be successful in the course. They will help you prepare for the homeworks and projects, which are graded.

Homework (Perform)

In homework, you will apply what you've learned to complete visualization and communication tasks. You may discuss homework assignments with other students; however, homework should be completed and submitted individually. Homework must be typed up using Quarto and submitted as a PDF in Gradescope.

Homework assignments are due 11:59 pm on the indicated due date.

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

Project (Perform)

The purpose of the projects is to apply what you've learned throughout the semester to solve some sort of real-world problem.

- Project 1: Students will create a data visualization using an assigned dataset. Completed in the middle of the course.
- Project 2: Students will choose a dataset and create a series of data visualizations to communicate a story about the data. Completed at the end of the course.

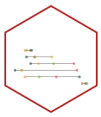
Both projects will include a written report and an oral assessment.

Grading

The final course grade will be calculated as follows:

Category	Percentage
Project 1	30%
Project 2	40%
Homework	30%

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



Letter Grade	Final Course Grade
A+	>= 98
A	93 - 97.99
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

Academic honesty

TL;DR: Don't cheat!

Please abide by the following as you work on assignments in this course.

Discussing assignments

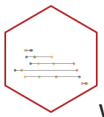
You may discuss individual homework assignments with other students; however, you may not directly share (or copy) code or write up with other students. Unauthorized sharing (or copying) of the code or write up will be considered a violation for all students involved.

Reusing code

Unless explicitly stated otherwise, you may make use of online resources (e.g. StackOverflow) for coding examples on assignments. You may not directly copy and paste from these sources, but instead you need to adapt the code to fit your specific task. You must explicitly cite where you obtained the code using a code comment `#` immediately near the appearance of the reused code in the file. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism.

Use of generative artificial intelligence (GAI)

Cornell's report on Generative Artificial Intelligence for Education and Pedagogy outlines many of the potential benefits and drawbacks to using GAI in the classroom. In this course,



we see the value of coding assistants such as GitHub Copilot and ChatGPT to generate code from text. However as an introductory course, we need to ensure that GAI is not used as a substitute or replacement for student learning. GAI should not be used by students to replace your ability to think clearly. Students who use GAI should use it to **facilitate**, rather than **hinder**, learning.

- ✅ **GAI tools for reference purposes:** You may make use of the technology as a reference tool, similar to looking up the documentation for a function or Googling your problem. For example, I hate writing regular expressions. Absolutely loathe it. Say I have a dataset where I need to clean a character column to remove all words that are within double asterisk symbols. I might ask ChatGPT

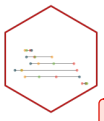
  > How do I make a scatterplot using **ggplot2** in R?

- ✅ **GAI tools for writing my code:** You may use GAI tools to assist in writing code in this class. You are expected to understand how any/all submitted code works. Any assignment for which you use GAI as more than a reference tool will require a written self-reflection to consider how you used GAI tools, what skills you acquired through the assignment, and how you believe you demonstrate mastery of the learning objectives for the course.

  You may not make use of the technology as a substitute for critical thinking. For example, you may not upload your data file to a GAI platform and ask it to create charts and statistical models for you. You are taking this course, not a GAI tool. I reserve the right to orally assess any student on their submissions to verify they meet the learning objectives for the assignment; **students who fail to satisfactorily demonstrate they have met the learning objectives may receive a grade penalty of up to 100% on the assignment.**

- ❌ **GAI tools for narrative:** unless instructed otherwise, you may not use GAI 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.



⚠ Potential consequences of academic dishonesty

Any violations in academic honesty standards as outlined in the Cornell University Code of Academic Integrity and those specific to this course will result in a 0 for the assignment (or possibly more) and will be reported to the College of Engineering Academic Integrity Hearing Board.

This course is participating in **Accepting Responsibility** (AR), which is a pilot supplement to the Cornell Code of Academic Integrity (AI). For details about the AR process and how it supplements the AI Code, see the AR website.

Extra credit

Students can earn up to a maximum of 1 percentage point towards their final grade through the extra credit assignment. This is the only opportunity for extra credit in the course.

Late work & extensions

The due dates for assignments are there to help you keep up with the course material and to ensure the course staff can provide feedback within a timely manner. **This is especially important in a compressed three-week format.** However we understand that things come up periodically that could make it difficult to submit an assignment by the deadline.

Late work

- **Homework assignments**: A **slip day** allows you to submit an assignment 24 hours after the deadline and still receive credit without a late penalty. You are provided with a total of **4 slip days** for the entire semester. Slip days may be used on **homework assignments**. You can use up to 1 slip day for a given homework assignment. Note that the lowest homework assignment will be dropped at the end of the semester.

  To use your slip days, just submit your assignment late. No need to email telling us you are submitting using your slip days. Check Canvas to see how many of your slip days you have used before submitting an assignment late.

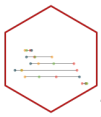
  If you use a slip day, **do not submit anything to Gradescope before the submission deadline**. We may begin grading before the slip day deadline and we will grade whatever submission we see in Gradescope.

  If you run out of slip days or fail to submit your assignment prior to the slip day deadline without prior permission then your assignment will not be accepted.

- **Projects**: Late work is not accepted.

Waiver for extenuating circumstances

If you need a bit of extra time, **please use your slip days**. Slip days are specifically intended for legitimate reasons for needing an extension like disability, religious observance, Title IX, student athletics, medical problems, and military service.



If using your slip days for accommodations is not working for you or if you have an SDS accommodation which includes deadline flexibility, you may request a deadline extension in-advance of the deadline. We will work with you to develop reasonable accommodations that align with your individual situation.

To request a deadline extension:

1. Commit and push the work you have completed up to this point on the assignment.
2. Email soltoffbc@cornell.edu. In your email clearly state
    a. The assignment
    b. What you have already completed on the assignment.
    c. What you have left to complete.
    d. Your proposed deadline extension (e.g. *Monday, February 8th at 11:59pm.*)

Regrade requests

Regrade requests must be submitted on Gradescope within 48 hours 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.