Syllabus

INFO 3312/5312 - Data Communication

Instructor

- Dr. Benjamin Soltoff
- Office: Gates Hall 216
- Email: soltoffbc@cornell.edu
- Office hours: Wed 9-11am

Course logistics

- Meets TuTh 1:00 - 2:15 pm for 28 sessions
- Lab sections meet on Fridays at varying times for 15 sessions
- 4 credits, offered for a letter grade
- Prerequisites: INFO 2950. Prior experience with R and Git/GitHub is preferred.

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

Click here for the instructor and TA office hours and locations.

You are welcome to attend the office hours for **any** INFO 3312/5312 TA, regardless of section.

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 instructors 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.

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 course staff 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 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!

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 48 hours Monday - Friday. Response time may be slower for emails sent Friday evening - Sunday.

Assessment

Assessments in this course are designed to help you successfully achieve the course learning objectives. They are comprised of three components: attendance and participation, homework assignments, and projects.

- **Attendance and participation** is expected throughout the semester.
- **Homework assignments** (6), due every other week (roughly), completed individually and submitted via Gradescope. Each homework assignment is worth 10% of the course grade. Lowest homework assignment score is dropped.

- **Projects** (2), mid-semester and end of semester, completed in teams.
  - Project 1: Teams will be given a dataset to visualize. Project 1 is worth 20% of the course grade.
  - Project 2: Teams will pick a dataset of interest to them and/or build an R package that implements a new type of data visualization in R. Project 2 is worth 30% of the course grade.

  The deliverables for each project will include a data visualization, a write up of the process and findings, and a presentation. For the second project, you will be encouraged to think beyond a traditional two-dimensional data visualization (e.g. interactive web apps/dashboards, data art, generative art, physical/tangible visualizations, **ggplot2** extensions, etc.).

  Each project will have a peer review component to provide at least one round of feedback during the process of development. Teams will provide confidential peer feedback regarding their teammates upon completion of the project. The scores from the peer evaluations, along with individual contributions tracked by commits on GitHub, will be used to ensure that each student has contributed to the teamwork.

  All team members must take part in the presentation. Presentations can be given in person in class.

Grading

The final course grade will be calculated as follows:

Category	Percentage
Homework	50%
Project 1	20%
Project 2	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

Letter Grade	Final Course Grade
C	73 - 76.99
C-	70 - 72.99
D+	67 - 69.99
D	63 - 66.99
D-	60 - 62.99
F	< 60

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.

You are expected to make use of the provided GitHub repository as their central collaborative platform. Commits to this repository will be used as a metric (one of several) of each team member's relative contribution for each project.

Graduate requirements for INFO 5312

Students in INFO 5312 have additional expectations in the course:

- INFO 5312 homework will at times be graded against a more stringent rubric
- INFO 5312 students will be grouped together for all projects
- Project rubric line items will have more stringent criteria for INFO 5312 students
- INFO 5312 students will complete an additional individual project during the semester

The final course grade for students registered in INFO 5312 will be calculated as follows:

Category	Percentage
Homework	40%
Project 1	15%
Project 2	25%
Individual project	20%

The final letter grade will be determined using the same thresholds as for INFO 3312.

Course policies

Academic honesty

TL;DR: Don't cheat!

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

- 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.
- **Automated help**: you may use AI code helpers such as GitHub CoPilot and ChatGPT as reference tools only. We recognize that these tools are becoming quite powerful, but relying on them prevents you from meeting the core learning objectives of this course. All code you submit in this class must be written by a human being – you.

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.

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. 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 **3 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 run out of slip days or submit after the slip day deadline, you may submit your homework assignment for up to 70% credit.

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

Regrade requests

- **Homework assignments**: 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.
- **Projects**: Copy the template below into an email. Send that email to soltoffbc@cornell.edu.

  Subject:

  > **INFO 3312/5312 (REQUEST) Regrade**

  Email Message *Template*:

  > NetID: *TODO* Team Name: *TODO* Assignment: *TODO*
  >
  > **Directly state the mistake(s)** in the grading of your assignment. Be **specific** *and* **specify the total points** that you believed should be returned for **each mistake**. (1-3 *brief and concise* bullets):
  >
  > *TODO*
  >
  > *(optional)* If necessary, **briefly** explain why your approach to this assignment is a **good choice** (1-3 *brief and concise* bullets):
  >
  > *TODO*

  Tips:

  - When writing, please be respectful, thoughtful, and professional.
  - **Be brief and concise.** Bullet points are encouraged. Please do not write a lengthy explanation.
  - Form and ground your argument based on **ideas and principles presented in this course**. This is the primary criteria we use to evaluate your regrade request.

– **Assume that we made a mistake**; avoid accusing us being unfair or punishing you.