Data-driven Web Applications
INFO 3300; CS 3300; INFO 5100

Instructor:	Prof. Jeff Rzeszotarski, jeffrz [at] cornell.edu
Time & Location:	M/W/F 11:15-12:05 in Klarman KG70
Course Prerequisites:	CS 2110 & INFO 2300 (or prof. permission)
Course Website:	https://jeffrz.com/info3300/
Course Announcements Q&A:	https://campuswire.com/c/G3730333F
Assignment Submission:	CMS (https://cmsx.cs.cornell.edu)
Graduate TAs:	Maria Antoniak (maa343 [at] cornell.edu) Hajin Lim (hl934 [at] cornell.edu) Sharifa Sultana (ss3634 [at] cornell.edu)
Undergraduate TAs:	See staff directory pinned on Campuswire
Office Hours:	See staff directory pinned on Campuswire
Contacts:	Prof. Rzeszotarski for course administration info3300staff [at] gmail.com for Prof+Grad TAs Post to Campuswire any content questions

Course Description

The web has become an outstanding environment for telling stories with data. This course will cover technologies for representing, modeling and displaying data in the context of interactive web pages. Practical skills for building web pages will be mixed with data mining algorithms and visualization design theory. We will use the D3 Javascript library to develop both static and dynamic visualizations, learn more about programming in Javascript, and explore web scalable vector graphics (SVG). Through design critique and formal study, we will identify the techniques visualization developers employ to create the "right" visualization for a given use case. This course introduces a number of popular data mining models and algorithms which we will incorporate into web visualizations.

Grading

- **Attendance** (10% of final grade)
  Class time will mix lecture, discussion, and hands-on programming exercises, so attendance is important. Attendance will be randomly taken using notecards in class. We will use the following procedure for taking attendance:
  o Before the start of class, put out ~400 notecards, spread in 4 places among the room as we have been doing previously (course enrollment is currently at ~275).
  o Students will answer the questions presented onscreen using the provided notecards during the first 5 minutes of class.
  o Should a think-pair-share activity follow the notecard question, the Professor and TAs will walk the rows, picking up any cards passed to the aisle. Otherwise, students will turn them in at the end of class (which will end 2 minutes early).
  o Students who come in late or otherwise miss turning in their cards have until 2 minutes after class ends to turn in their card to the bag at the front of the hall if they missed the card collection.
  Submissions that are illegible, are turned in later than the 2 minute period following end of class, or are written on materials other than the provided cards will be rejected. The university Code of Academic Integrity will be enforced as necessary. All students have 3 excused absences before grades are penalized, no questions asked. If you go beyond 3 absences due to extreme personal circumstances, reach out to course staff with appropriate supporting materials (see missing work section). Any remaining excused absences at end of the term will be converted to extra credit.

- **Weekly homework** (40% of final grade)
  Homework will be assigned on Mondays, due at 11:59PM ET the following Monday, and returned the following Wednesday. Work will be turned in through CMS. Regrade requests should go in writing to jeffrz [at] cornell.edu or to info3300staff [at] gmail.com. Requests must be made within 1 calendar week of grade release.

- **Two open-ended group projects** (30% of final grade, 15% each)
  The first project will be to design a static (non-interactive) web visualization. The second project will be to design an interactive visualization. Group members will be randomly assigned. Each group will be assigned to a TA advisor. You will send progress reports to your TA listing team member responsibilities and what has been accomplished in the past week, and flagging any problems or questions.

- **Take-home final exam** (for undergraduates in 3300) (20% of final grade)
  OR
  A third group project (for graduate students in 5100) (20% of final grade).
   (a larger interactive visualization project, similar to first two group projects)

- **Extra credit** (up to 5% boost of final grade)
  There will be several opportunities for extra credit this term. The first opportunity for extra credit is hidden right here in the syllabus! There is (at least) one error somewhere in this document. Find it and email it to Professor Rzeszotarski for extra credit. If 75% of the class completes the mid-semester survey, then the class will receive a 1% bonus to their grade. Each leftover excused absence at the end of the term will earn 1/3% extra credit, up to a 1% grade bonus. Another 1% bonus will be offered to students who post a neat visualization example from the web to the "Example Library" tag on Campuswire before Spring Break. No duplicates allowed, so doing this earlier rather than later may be to your benefit. Finally, 2% can be earned via SONA credits (see section in syllabus).

Grade Scale

97.0%-100%: A+	73.0%–76.9%: C
93.0%-96.9%: A	70.0%–72.9%: C-
90.0%–92.9%: A-	67.0%–69.9%: D+
87.0%–89.9%: B+	63.0%–66.9%: D
83.0%–86.9%: B	60.0%–62.9%: D-
80.0%–82.9%: B-	< 60.0%: F
77.0%–79.9%: C+	

Late or Missing Work

[The following rules apply under normal circumstances. If you have experienced a personal crisis or a medical condition contact Prof. Rz as soon as possible. The earlier we hear from you, the more we can help. We can only provide assistance if we are made aware of the problem in a timely fashion.]

If you do not submit work by the deadline, or if you submit a file that does not contain the correct work, we will record a zero. There will be no exceptions. This policy is harsh but necessary for a class of this size. **In order to protect you, your two lowest homework grades will be discarded (excluding projects and exams).** Students sometimes use these "homework passes" during unusually busy weeks, but you are always encouraged to submit as much as you can. Check carefully that you have submitted the correct file in the correct format. We will make an effort to open any file you submit.

In rare circumstances we are able to accept work that was not submitted correctly if you can provide a version of the file that is timestamped in a way that you do not control, such as an email attachment through Cornell's email system to an instructor or TA email address. Your laptop's "last modified" date is not acceptable. In the unlikely event that you are having difficulty with CMS, you may submit work by email before the deadline without penalty.

In-class work

Web programming is a complicated skill, and there's no way to learn without practice. Class will often involve a programming problem that we will work on together. Template HTML files will be made available before class via GitHub.

Laptops

In order to facilitate interactive in-class work, you are allowed to bring a laptop. If you don't have one or don't choose to bring one, work with someone sitting near you. If you have a laptop, you will be expected to use it for relevant work. Laptops are welcome during class lectures. As attentional resources became divided, task performance suffers substantialy. At the same time, participants in studies of attention [rarely estimate the costs of divided attention accurately](). In other words, laptop users are likely to overestimate their ability to multitask. Laptop use has been proven not only to hinder users' classroom performance, but also the [performance of nearby peers]() (whether they also use a computer or not). As a result, if you choose to use a laptop *during lectures* please sit on the left or right portions of the room.

Academic Integrity

We will follow university policies as outlined in the Academic Integrity Handbook. You are encouraged to discuss homework, but each student will complete assignments alone. Using other individuals' code is an important part of programming, but for group projects the code should be substantially the work of the group members except for standard libraries such as D3, lodash, and jQuery. Any code used in projects that was not written by the group members should be placed in separate files and clearly labeled with their source URLs. If you have benefitted from online resources (such as examples or StackOverflow answers), list the URLs in comments in your own code, even if you did not directly copy anything. Recall that some workplaces ban StackOverflow and sites like it outright to avoid diluting their IP. Project work that relates to your other classes or research is encouraged, but you may not recycle assignments. There must be no doubt that the work you turn in for this class was done

for this class and this class alone. Integrity issues will be adjudicated at the discretion of Professor Rzeszotarski and course staff.

Special Accommodations

We will make every effort possible to ensure that the class works for all students. Students who have self-identified to Cornell SDS as needing special accommodations in the classroom should contact the professor at the time of enrollment or the first week of class so that any course materials can be adapted, and other appropriate arrangements made. Failure to do so may prevent us from making the proper accommodations. I will not be recording lectures unless SDS requires it. If there is a specific event that you are concerned about (such as an exam), please inform us at least two weeks in advance so that we have time to make arrangements.

Note on Inclusiveness

This course involves both self-directed assignments as well as in-class exercises. It is possible that in the course of this class you will encounter datasets, visualizations, or arguments that do not match your worldview or perhaps might even be upsetting. I cannot guarantee such events will not occur, and do not want to hamstring the potential projects that teams can complete with a set of content requirements. I expect students to be respectful throughout the critique process and in-class activities, acknowledging that at times discussion can be fraught or argumentative. I will not tolerate intentional displays of disrespect or marginalization during class time and out of class project work. Penalties will range from loss of participation grade to failure of projects or assignments on a case-by-case basis. If during the semester you have concerns about the way the course is going or are having a negative experience, please reach out to me directly and I will work to resolve the situation.

SONA Credits

Many researchers on campus need participants for user studies and other types of experiments. The SONA system allows you to register for studies. You will get 0.5% extra credit for each 30 minute study (or equivalent), up to a maximum of 2.0% for the term. Participating in studies is a great way to find out what real research looks like. To register, go to this URL: https://cornell-comm.sona-systems.com

Course Outline

The following is a tentative course outline, subject to change. The schedule listed on the course web site will take precedence.

Week 1   Introduction
      Visual Channels
Week 2   Javascript + JSON
      Javascript Functions & Events
      Brief history of analytics
Week 3   Intro to SVG
      Intro to d3
      Color theory + models
Week 4   d3 scales
      Data -> d3 scales
      d3 scatterplot (style + scale)
Week 5   d3 line plots, data joins
      d3 histograms & stacking
      Visual perception
Week 6   Regression models
      More on linear regression
Week 7   Demo day
      Multiple regression
      Multivariate displays
Week 8   d3 maps & projection
      d3 choropleth
      Human-Visualization Interaction
Week 9   d3 direct manipulation
      d3 pan+zoom
      Static vs. Interactive Elements
Week 10   d3 emergent layout (bubbleplots)
      d3 network layout
      Network Viz (a.k.a. force direction considered harmful)
Week 11   Scaling up with raster graphics
      Demo day
      Animations & storytelling
Week 12   Dimensionality reduction
      More on dimensionality reduction
      Portraying time
Week 13   d3 text visualizations
      3-D d3 day (a.k.a. 3-don't)
      Designing for novices
Week 14   Usable Uncertainty
      5100 project showcase
Week 15   Future of visualization