

CS/INFO 3300; INFO 5100
Project 1
Due 11:59pm Thursday Mar 1

You are creating a non-interactive data visualization. You will use d3, but the final result will be something that could be published on large-format paper or as a static PNG image on a web page. We will grade your work as it appears when we load the HTML file in a browser. We will not click, hover, or interact with anything. You are encouraged but *not* required to integrate multiple data sources for this project. Each team has been assigned a TA. Your TA is your best resource.

Examples and sources: You are encouraged to get data and inspiration from other sites. Make sure you acknowledge these in comments and in your written description. Any code that you did not write yourself (such as d3) should go in a separate .js file. Unacknowledged code or concept reuse will be handled with standard academic integrity procedures.

Submit a written status report, one per group, to CMS by 11:59pm on the following dates. On-time completion of these reports will be worth **5 points**.

- Friday February 16: Describe five ideas. For each member, list assigned tasks for the following week. Include a filled out and signed work-practices contract.
- Friday February 23: Describe tasks accomplished by each team member. For each member, list assigned tasks for the next week. Include scans or pictures of three to five *hand-drawn* designs or storyboards showing different project designs. Take time with these. There was a strong correlation last year between careful design sketches and high grades.
- Tuesday February 27: Describe tasks accomplished by each team member. For each member, list assigned tasks for the final submission.

About grading: This is an open-ended assignment. With homework we have a specific idea of what we want and we "take off" points when your work deviates from that. The reason project experience is the single most valuable asset you can bring to a new job is that we do *not* have specific ideas about what projects should look like: it's up to you and your teammates. As a result, think of the criteria below as an opportunity to "earn" points, not "lose" points. Our principle with projects is that better work should get better grades. That does not mean that we curve: there's no reason we wouldn't in theory give everyone 100s, but in practice "perfect" grades are extremely rare.

If you have concerns about how your team is working contact your TA individually as soon as possible. In rare circumstances we will differentiate grades with a group, but we are much happier to help a group succeed.

Good practices:

- Start now.

- Talk to each other. Seriously. Listen and value different perspectives.
- If a group member will be unavailable for any period during the project, figure out in advance how you will work around that absence.
- Your ZIP file should contain a directory called `project1`. The web page should be in this directory and should be called `index.html`.
- Use relative paths for data, images, and other resources: do not start URLs with `"/`. Your project will be one directory among many, not the document root.
- Start now.
- Set up a code repository, like Github or Bitbucket. This is always a good idea, but can also provide insurance if something goes wrong with your CMS submission.
- Did I mention that you should start now?

A visualization rubric (Hearst et al, 2016):

- This visualization makes important information visually salient.
- This visualization uses visual components appropriately.
- This visualization successfully presents multiple relevant facts in a single visual pattern.

Your final submission has two parts, a d3-based **static** data visualization (70 pts) and a written description of your visualization (25 pts). Turn in a .zip archive containing:

1. An HTML page called `index.html` containing your visualization. Include any additional script files (such as d3 or jQuery) and any additional data files, preferably in JSON form. We will look for the following elements. You can show your TA a prototype at any time. This section will be graded on the following elements:
 - A. Complexity of the data. Find a dataset that is manageable, but avoid trivial data. There should be more than two variables, for example. An advanced project might combine multiple datasets to provide a unique, novel perspective. Editing is important! Beginning projects often have too little data or too much. Don't overwhelm us with information.
 - B. Technical correctness. The code must actually do what you intend it to do. We also prefer good style in coding: use informative variable names, consistent indenting and whitespace, and informative comments.
 - C. Creativity. Beginning projects often look like online examples or things we've seen before. Advanced projects will make us think "how did they do that?" or use something familiar in an unfamiliar way. Don't be boring.
 - D. Mapping from variables to aesthetic properties. Use scales such as position, shape, color, and text appropriately for variables. Advanced projects give us accurate impressions of the underlying data values, allow us to make comparisons between relevant data points, and balance between focus and context. Beginning projects are often hard to interpret and make comparisons difficult.
 - E. Usability. Someone viewing your work should be able to understand the data values represented in the visualization easily and accurately. Advanced projects make

choices that are clear and intuitive, and may walk us through specific examples. Beginning projects often leave us wondering what we're looking at, or make us read long descriptive paragraphs.

- F. Overall polish. Beginning projects will look like a collection of parts, with default styles. Advanced projects will have a sense of unity, even if they have multiple sections.
 - G. Purpose. What's the point? What are you trying to say? Beginning projects will present information. More advanced projects will have a clear argument and use carefully chosen combinations of aesthetic properties to guide our attention to the evidence that supports that argument.
2. A PDF file containing a written description of your project. There are no specific page or word limits. This document should contain:
- A. A description of the data. Report where you got the data. Describe the variables. If you had to reformat the data or filter it in any way, provide enough details that someone could repeat your results. If you combined multiple datasets, specify how you integrated them. Mention any additional data that you used, such as shape files for maps. **Editing is important! You are not required to use every part of the dataset. Selectively choosing a subset can improve usability. Describe any criteria you used for data selection.** (10 pts)
 - B. A description of the mapping from data to visual elements. Describe the scales you used, such as position, color, or shape. Mention any transformations you performed, such as log scales. (10 pts)
 - C. The story. What does your visualization tell us? What was surprising about it? (5 pts)