SI649 / EECS 548 Static Viz Project, Winter 2025

Due date: Sunday February 23rd, 2025 at 11:59pm (Eastern time)

The static viz project for this course is about making communicative visualizations to add detail, context, and nuance to a news article. The article we have selected for this year is about school segregation. The article also makes lots of specific claims and mentions various statistics to help provide context and background for the story, but only includes one data visualization.

Your goal in this project is to design and create **four static visualizations** to accompany the story, to benefit the reader. You will do this individually, rather than as a team. You are free to focus on any aspect of this article you like, including regional, national, individual, or other factors, in order to provide greater detail and context.

Although you have great freedom to choose what data to visualize and how, please keep in mind the following constraints:

1. Each of the four visualizations should plot **different data** (although possibly from the same source), and use **different techniques** (i.e., not all bar charts)
2. You are free to use whatever tools you like to create your visualizations, but you should aim to use a **consistent and distinct style** across the four visualizations. In other words, they should all have a unified style, but you should **make it your own**, rather than just using, for example, the default Tableau style.
3. You should generally aim to **follow Tufte's design principles**. We can summarize these as:
   - Well designed and truthful presentation of interesting data
   - Complex ideas communicated with clarity and precision
   - Gives the viewer greatest number of ideas using the most efficient presentation
   - Nearly always multivariate
4. **Attention to detail** matters a lot! It should be obvious that you have thought carefully about all aspects of each visualization. Is all text readable? Could any non-data ink be removed? Did you choose reasonable colors? Did you use a sensible map projection (if you used a map)? Is it clear what you are trying to communicate, and is this effective?
5. For each figure, also include a **short caption** describing what your visualization shows, as might appear in a news article with visualizations embedded in it, **including a link or reference to the data sources used.**

Article

The story we are basing this project on is called "Choosing a School for My Daughter in a Segregated City", by Nikole Hannah-Jones, published in the *New York Times Magazine* on June 9, 2016. You can find a copy in the files on Canvas, or (if you have subscriber access), you can read it online here: https://www.nytimes.com/2016/06/12/magazine/choosing-a-school-for-my-daughter-in-a-segregated-city.html

Deliverables

1. Final visualizations:
    - For the final submission, create your final, high-quality visualizations to accompany the article. You should **submit these to Canvas as individual PDF figures**. You are free to implement your visualizations in any way you choose.
    - You should also **embed your visualizations within your a copy of the original article** (at the appropriate locations), and then upload it to Canvas as a PDF.
2. Data, code and/or workbook used to create the visualizations:
    - Please also upload a zip file which contains the data, code, and/or workbooks used to create your visualizations.
3. Final report (written as a short, informal blog post):
    - To accompany your final submission, you should write a short informal report, (something like a blog post, probably around 5-6 pages of text in total), that describes your process, due on the same day as the final visualizations. **Please name this file XXX-uniqname-report.pdf**, where XXX is your section number (001, 101, 002, or 102), and uniqname is your UM uniqname.
    - For any **two** of your visualizations, your report should cover:
        - **Tasks**: What comparisons is your visualization designed to support? What will the viewer learn and take away from it? Remember that a good visualization should be designed to effectively support specific domain tasks, and can typically support multiple tasks, not just one.
        - **Data**: Where did you source the data for this visualization from? What questions or concerns do you have about it? Did you have to process or clean up the data in any way? Is this reflected in your final visualization?
        - **Your design process**: What did you try? (feel free to include screenshots). What examples did you look at for inspiration? (again, screenshots are welcome). The expectation is that you tried multiple things and iterated over your design over the weeks of this project. Describe what you liked or didn't

like about your initial design, and how you arrived at your final implementation.
    - **Qualitative self-evaluation**: How would you judge your own design and creation? In what ways is it effective? How could it be improved? Please connect this to principles that you have learned in class.
- Here is an example of a process blog that you may want to refer to as loose inspiration (but please follow the format outlined above, not the format used by this example, and try to be concise): https://www.visualcinnamon.com/2019/04/designing-google-cats-and-dogs/

Evaluation Criteria

For the final submission, we will be looking at:

1. **Insight and contribution**: How much do these visualizations add to the reader's understanding, beyond what is already in the article text? Do they accurately reflect what is in the data? Do they provide nuanced insights?
2. **Clarity and effectiveness**: Is it clear what the visualizations are trying to communicate? How effective are your visualizations for the tasks you have selected?
3. **Design and attention to detail**: Have you used good design principles and considered human perception and cognitive limitations in designing your visualizations? Have you gotten the details correct?
4. **Style**: Do your visualizations use a consistent and intentional style?
5. **Variety**: Do all of your visualizations use different data and different techniques?
6. **Creativity and ambition**: How interesting or ambitious are your visualizations? Do they reflect a considerable amount of effort?
7. **Writing**: Is the report detailed, thoughtful, and well written, explaining your design process?

Aesthetics

Keep in mind that the final results should have both function and form. That means the visualization should *look* good in the end. You are welcome to use any software you like to augment or touch up your static visualization (e.g., Illustrator, Figma, Inkscape, or even Powerpoint), although this is not required. Just make sure to describe your process in your report.

Advice

- Start by reading the article, and thinking about what kinds of questions you have as a reader. Which of those might you be able to find data for?
- Look for data and see what you can find; note that the most interesting or useful data might require some preprocessing.
- Thinking about both the questions you have and the data you can find, come up with tasks that are relevant to the article, that you might be able to effectively support with a visualization.
- Start by sketching ideas for visualizations on paper, and iteratively improve your design.
- Ask yourself, does this visualization actually add something to the article? Each visualization should be thoughtful and strive to support complex reasoning and data.

Analysis vs Exploration

As you approach this project, try to put yourself in the role of a data journalist. That is, you are both an analyst and a communicator. As an analyst, it's up to you to find interesting things in the data (some of which are suggested by the article content), but you should also think critically about the data and analyze it in a rigorous way. Visualization may be an effective tool in exploring and analyzing the data, but these are not likely to be the visualizations you will submit as your final work (except as screenshots of your process in the blog post). In your role as a communicator, you need to put yourself in the mindset of the reader, recognize that they haven't gone through the same work of exploration and analysis that you have, and find a way to effectively communicate the information you want to get across just based on the final visualizations you create. Note: the insights you choose to communicate should be *relevant* to the article, but they do not necessarily need to *agree* with the conclusions of the article's author. Above all, make sure your visualizations *tell the truth*.

Datasets

To get you started, here are some datasets that you might consider working with. You are not obligated to use all (or any!) of these, and you are free to seek out and collect additional data as you see fit. Note that because the article is from several years ago, you have a chance to use data produced *after* the article if you think it would be useful. (Please also see the associated .txt notes file for some data files).

- National-level data on all public and private schools: We will provide a dataset from 2015-16 in .csv format, but you can also search or download additional data (including other years) from: https://nces.ed.gov/datatools/
- Historical New York school test scores (math and language). We will again provide core data in .csv format (extracted from a larger excel file), but you can also find additional data on this site: https://infohub.nyced.org/reports/academics/test-results
- Historical New York graduation rates: We will also provide this in .csv format (again extracted from a larger excel file), with additional data available here: https://infohub.nyced.org/reports/academics/graduation-results
- Additional New York school data (demographic snapshots, absenteeism, etc.) available here: https://infohub.nyced.org/reports/school-quality/information-and-data-overview
- School district boundaries. In case you want to study changes in school zones over time, there is GIS data available here: https://nces.ed.gov/programs/edge/Geographic/DistrictBoundaries
- Generally speaking there is a *lot* of data available on schools! You could also consider comparing against what comparable data would look like for Ann Arbor, Detroit, or other parts of Michigan. Other places to look include:
    - US Census Data https://data.census.gov/cedsci/
    - National Center for Education Statistics: https://nces.ed.gov/
    - NCES Common Core of Data: https://nces.ed.gov/ccd/
    - New York State Education Department: https://data.nysed.gov/
    - New York State GIS data: https://gis.ny.gov/gisdata/
    - NYC Department of Education: https://www.schools.nyc.gov/about-us/reports/doe-data-at-a-glance
    - NYC Open Data: https://data.cityofnewyork.us/
    - National Student Clearinghouse Research Center: https://nscresearchcenter.org/
    - Data provided by individual schools (e.g., https://www.bths.edu/apps/pages/index.jsp?uREC_ID=217764&type=d)