CSC/SDS 109: Communicating with Data

HW 01: Data & Data-Visual Mapping

This is an individual or pair assignment-- you pick (collaboration is highly encouraged)!

Goals:

- Explore data documentation in real-world visualizations
- Evaluate the data visual mappings in real-world visualizations

Instructions

Choose Visualizations

Choose a visualization from the galleries here: https://www.vislies.org/2024/

Choose a visualization from a news reporting outlet.

Choose one additional visualization from any third source.

Deliverables

Create a document that contains:

- A screenshot and link to each visualization you chose
- For each visualization--
  - Look for the raw data source
    - **If you cannot find it**:
      - Document the steps you took to search for it (your search should take at least 20 minutes before calling it)
      - From the visualization, can you infer who collected the data?
      - From the visualization, can you infer when was the data collected?
      - From the visualization, can you infer what each observation (row) in the data represents?
      - From the visualization, can you infer what variables (columns) are included in the data?
      - From the visualization, can you infer if the dataset representative of the population it seeks to capture? (ex. If the dataset claims to be US Colleges,

are college across the entire country included? Are public and private institutions included? Etc.)

- Pick one of the questions you answered above. What implications does your answer have for biases or ethical issues present in the data?

- **If you can find it**:
  - Provide a link to it
  - Who collected the data?
  - When was the data collected?
  - What does each observation (row) represent?
  - What variables (columns) are included?
  - Is the dataset representative of the population it seeks to capture? (ex. If the dataset claims to be US Colleges, are college across the entire country included? Are public and private institutions included? Etc.)
  - Pick one of the questions you answered above. What implications does your answer have for biases or ethical issues present in the data?

- Critically read the visualization
  - What mark(s) does the visualization use?
  - What does one mark represent?
  - What variables are shown in the visualization?
  - What visual channel(s) are used to represent each variable (i.e. what is the data-visual mapping)?
  - For each variable-visual channel pair, does the visual channel shown work for the type of data it represents?
  - Do you notice any inaccuracies in the visualization?
  - Do you think the visualization is well designed? Why or why not?

Submission

Submit your deliverable(s) as a PDF on Gradescope. If you worked with a partner, submit as a group (https://guides.gradescope.com/hc/en-us/articles/21863861823373-Adding-Group-Members-to-a-Submission).

Rubric

The following matches the rubric you will see on Gradescope.

	Points	Criteria
For each visualization:	0.25	Screenshot
	0.25	Link
	1	Raw data link or search procedure
	0.5	Data collector
	0.5	Data collection time
	0.5	ID observation
	0.5	ID variables
	1	Representativeness of data
	1	Data biases
	0.5	ID marks
	0.5	Mark explanation
	0.5	ID variables
	1	Data-visual mapping
	1	Data-visual mapping evaluation
	0.5	Inaccuracies
	0.5	General evaluation
TOTAL	30	