WRI363

Feb 29/24

Victoria Chui
victoria.chui@mail.utoronto.ca

Victoria (Me)

- PhD student in the Faculty of Information, STG
- I study human-centered algorithm design and AI tools for decision-making in healthcare systems
  - HCAD (theoretical, speculative, participatory) approaches
  - To design/code machine learning algorithms
  - That administrators and clinicians use in everyday healthcare systems
  - To make ethical decisions or assist workflow in unbiased manners
- I have a background in data science (MI) and biochemistry (HBSc)

Agenda

01

Dear Data Activity

02

Critiquing Activity

03

Visualization Ethics

04

Conclusion

01

Dear Data
Activity

Dear Data

Dear Data was a year-long project started by Giorgia and Stefanie, where weekly postcards representing a week's worth of data were sent back and forth from London to New York.

- Each week Giorgia/Stefanie created a data visualization on a postcard
- The visualizations represented a week's worth of _ _ _ _
- The postcards had the visualization on the front and a legend on the back
- The week's worth of _ _ _ _ was noted each week and translated into a hand-drawn visualization
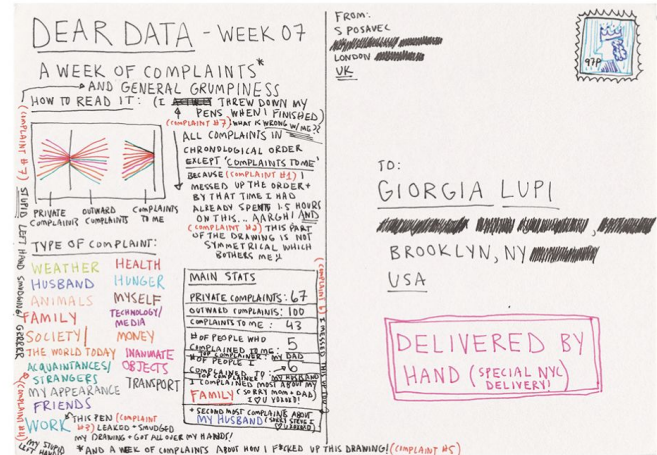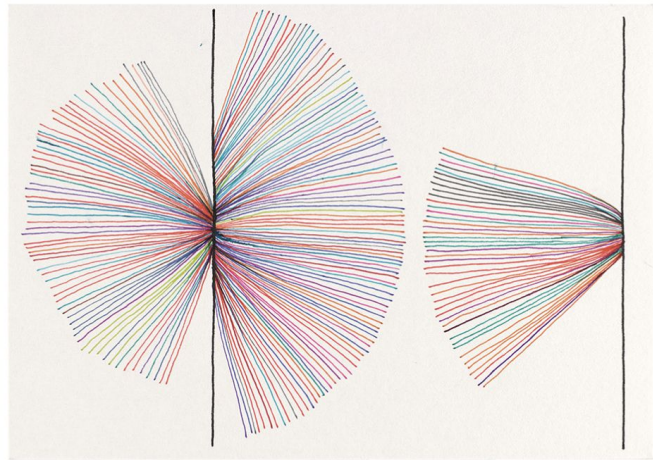
Note the hand-drawn stamps: these postcards were delivered in person in New York!
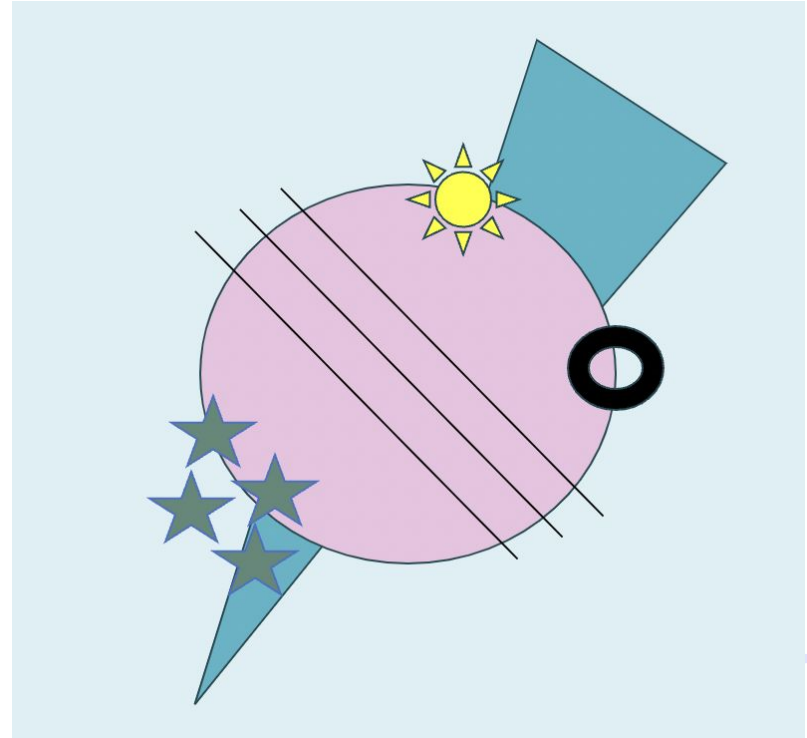
Dear Data Activity (10-15 min)

Our Goal

- To create a Dear Data visualization
- To explore how data can be visualized in a variety of ways
- To understand the many "mundane" activities that can become data

To Get Started…

- Go to the PowerPoint link on Quercus
- Find the legend and example
- Create your own! Add your own information and representations

Dear Data Activity

What Did We Learn?

- ___

02

Critiquing Activity

Terrible Maps

Bad/Misleading Data Visualization Pitfalls

1. Omitting data/cherry picking

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts without a common axis

5. Choosing a nonoptimal chart type

   a. Continuous vs discrete

6. Poor colour choices

7. Providing no context

Terrible Maps

Data Visualization Pitfalls

1. Omitting data/cherry picking

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts without a common axis

5. Choosing a nonoptimal chart type

    a. Continuous vs discrete
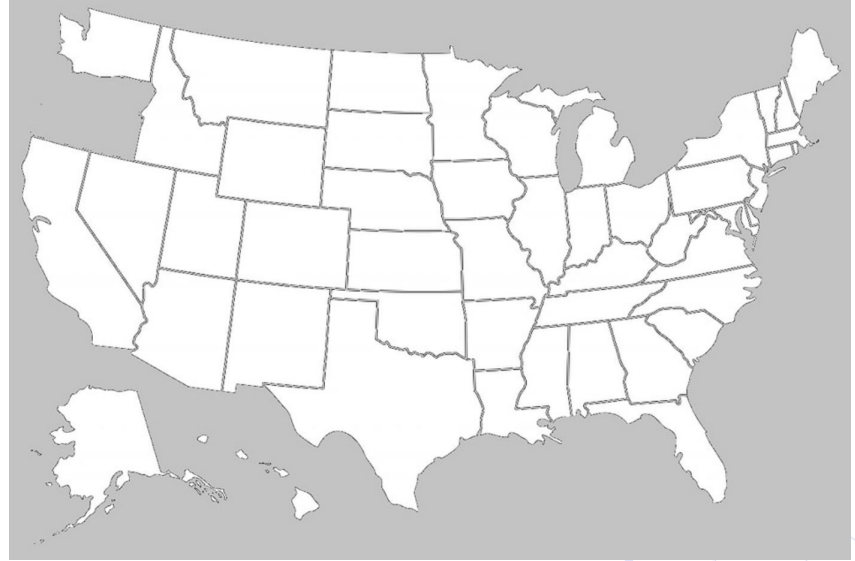
6. Poor colour choices

7. Providing no context

On Reddit, Instagram people post "terrible maps":

- What makes them "terrible"?
- What makes them funny?
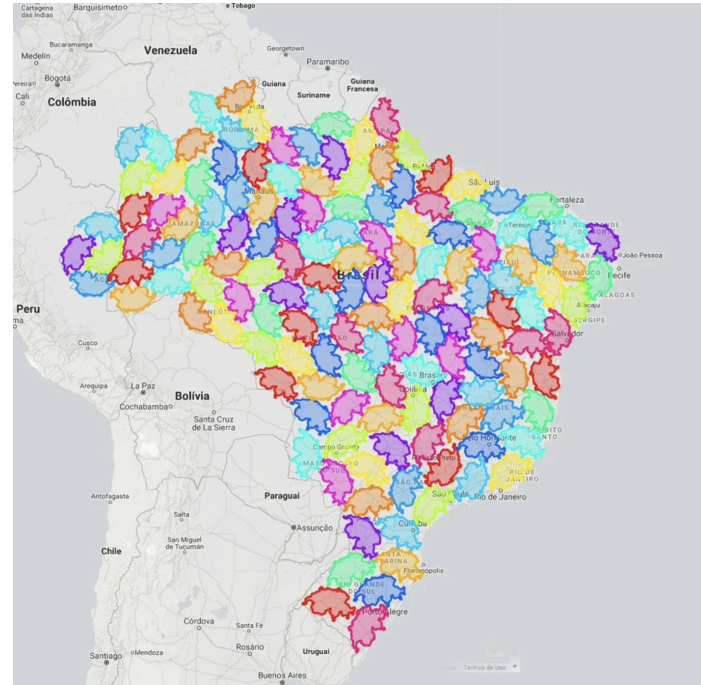

"If the US were an Oregon-donor"

Terrible Maps

Data Visualization Pitfalls

1. Omitting data/cherry picking

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts without a common axis

5. Choosing a nonoptimal chart type

    a. Continuous vs discrete

6. Poor colour choices

7. Providing no context

How many Switzerland's fit in Brazil?

Terrible Maps

Data Visualization Pitfalls

1. **Omitting data/cherry picking**

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts without a common axis

5. **Choosing a nonoptimal chart type**

    a. Continuous vs discrete

6. Poor colour choices

7. **Providing no context**

How many Switzerland's fit in Brazil?

Terrible Maps

Data Visualization Pitfalls

1. Omitting data/cherry picking

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts without a common axis

5. Choosing a nonoptimal chart type

   a. Continuous vs discrete
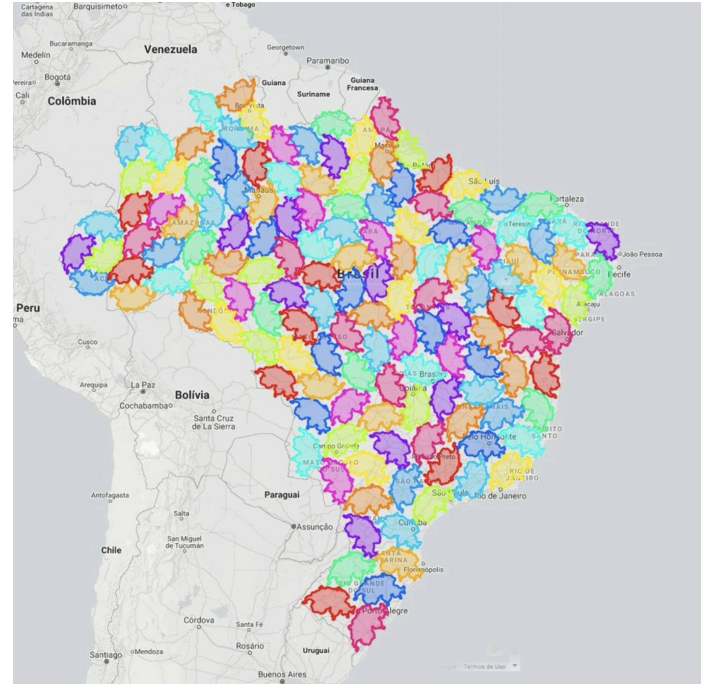
6. Poor colour choices

7. Providing no context

"Europe if it were birds"

Terrible Maps

Data Visualization Pitfalls

1. Omitting data/cherry picking

2. Truncating Y (or X) axis

3. Implying correlation = causation

4. Dual-axis charts **without a common axis**

5. Choosing a **nonoptimal chart type**

   a. Continuous vs discrete

6. **Poor colour choices**

7. **Providing no context**

"Europe if it were birds"

03

Data Viz Ethics

Ethics - Skill

Data Visualization Experts Bear Responsibility

- Not every member of your team can accessibly/impactfully translate data → knowledge.

- Data is subjective, political, and incomplete.

- Building expertise in the knowledge you are representing is important for impactful communication.

- Fine tuning your ability to share information creates a stronger narrative.

- "we are not all equal in our relationship to big data"

Teaching Students How (Not) to Lie, Manipulate, and Mislead with Information Visualization
- Mahmud, A. et al. (2017)

Ethics - Unethical Representations

Common Unethical Design Pitfalls

- Using inappropriate visualization techniques.

- Telling a story from/for only a singular perspective (yours).

- Manipulating users of a different data/visual literacy.

- Being unreceptive to feedback.

- Creating misinformation.

Teaching Students How (Not) to Lie, Manipulate, and Mislead with Information Visualization
- Mahmud, A. et al. (2017)

Ethics - What Do We Do?

Exercises and Strategies

- Rely on common metrics (maps).
- Follow a list of reflexive prompts for data collection, analysis, visualization, and self-reflection.
- What is the origin or source of your data? Who or what collected the data? At what point in the data collection process are you interjecting your analysis?
- What's kind of narrative does the visualization create? Who is it aimed at? Is it a widely accessible narrative? What are the potential social and political implications of the visualization?
- Who is your design and research team comprised of? What assumptions and biases are embedded in the data? How might these impact the data's use? Who or what is privileged by the data?

Teaching Students How (Not) to Lie, Manipulate, and Mislead with Information Visualization

- Mahmud, A. et al. (2017)

04

Conclusion

Takeaways

- Data visualization *experts* are few and desired
- Both technical and theoretical understandings are important for clear information communication through data
- Data visualization can be FUN!
  - And it can be FUN to critique data visualizations
- Anytime you convey/distort/share information, there are ethical considerations
  - Be accountable to yourselves and your stakeholders
- You (can) share data in visual formats everyday!