DSCI 320: Assignment #1 – Self-Evaluation

By Selena Shew 68692490

Sept. 30, 2023

	Mastery	Proficient	Satisfactory	Explanation
DESCRIPTION				
What data is represented inthis visualization?	Thorough and precise description of the data represented. Specific bydescribing the attribute types of each feature.	Accurate description of thedata but does not include details for all of the data attributes shown	Limited or incorrect description of the data.	I give a clear and in-depth explanation of the data represented in the provided table.
What is the intent of the viz	Describes the goal of the viz,what message it is trying to get across, discusses the audience and possible affective considerations	Describes a goal but fails to discuss affect, audience in a meaningful way	Describes a goal by just quoting the title	I clearly discuss the intent of but do not go in-depth of the intended audience.
How is data encoded in thevisualization and channel- attribute mapping	Comprehensive explanation ofdata encoding, marks and channels. Talks about the attributes, does not confuse explain for critique	Good explanations of data encodings, marks and channels. Sticks to the description, but may omit 1 mark or channel.	Leaves out some marks or channels and has errors. Failsto use proper viz vocab.	I give a clear and in-depth explanation of the data represented in the provided table.
CRITIQUE				
Expressiveness & Effectiveness	Critiques design choices based on the principles, highlights at least 2 examples	Critiques design choices basedon principles but does not ground them in examples from the visualization	Is not included.	I clearly explain the expressiveness and effectiveness of the visualization based on the principles, and explained several specific and clear examples.

Critique the visualization by expanding on the characteristics of the channelsand how they are suited for the given attribute?	Comprehensive critique of channel characteristics, providing detailed examplesand analysis.	Adequate critique of channel characteristics but lacks depthor examples.	Sticks to the mundane.	I give clear and detailed examples and analysis of the data channels.
Address what questions is thevis effective in answering.	Addresses questions effectively, covering allrelevant aspects.	Partially addresses questionsbut misses some key aspects.	Highlights 1 question of interest.	I list several questions that the visualization attempts to answer.

Redesign				**Explanation**
Redesign is legible.	Redesign is highly legible, allowing users to easily interpret each datum. Makes great use of space, color, and it detailed.	Redesign is legible and allows users to easily interpret each datum, makes a great use of color and demonstrates effortto convey meaning.	Redesign is somewhat legible but could be clearer.	The redesign is legible and easy to understand, but I apologize for my lack of artistic skill.
Addressing limitations of the original viz.	Redesign creatively and effectively addresses at least one limitation of the original visualization.	Redesign addresses one limitation but lacks depth or creativity.	Redesign address one limitation using a fairly standard approach.	Hopefully by taking this class I will become inspired and gain a greater sense of creativity and artistic skill.
Issues from the original viz that it addressed.	Provides a detailed and clear explanation of how the redesign addresses issues from the original visualization.	Adequate explanation of issues addressed but lacksdepth.	Explanation of issues addressed is minimal	My explanation is clear and addresses exactly how the redesign addresses issues from the original chart.
Limitations of their re-design.	Clearly identifies and thoroughly explains the limitations of the redesign.	Identifies some limitationswith a basic explanation.	Identifies limitations but lacksexplanation or depth.	My explanation is clear but could be more thoroughly explained.

DSCI 320: Assignment #1 – Into the Critique of It

By Selena Shew 68692490

Sept. 30, 2023

The Approved Visualization

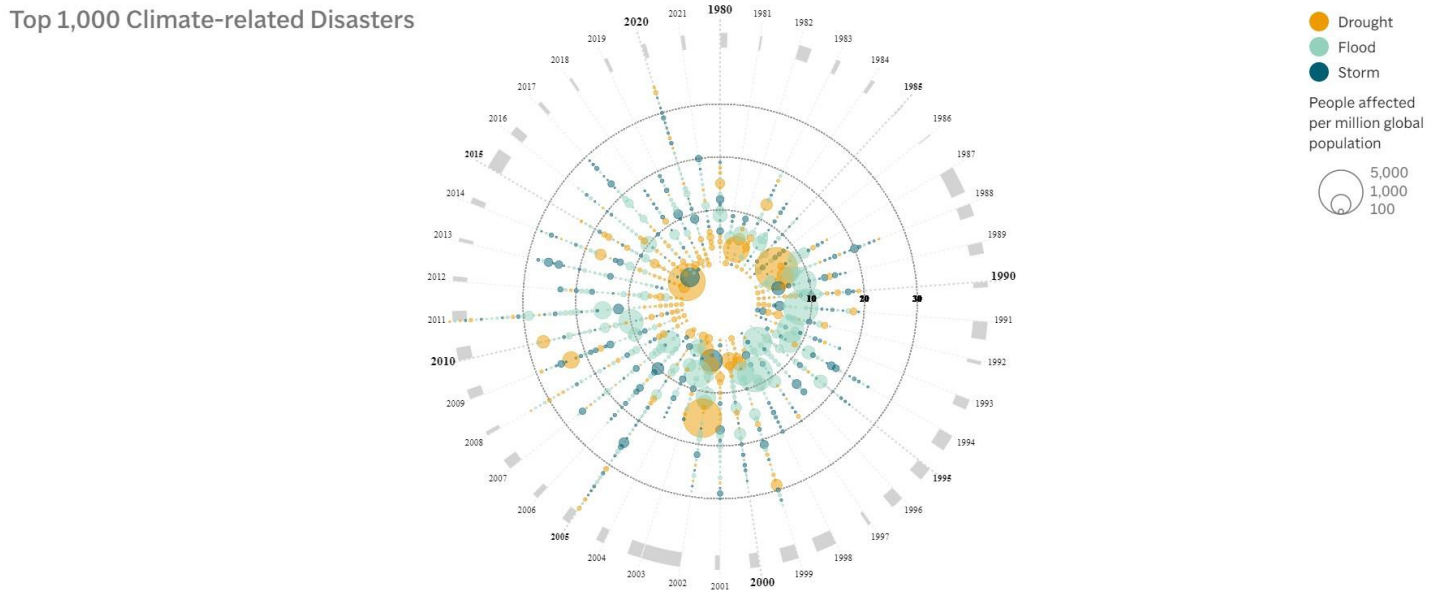


Image Reference:

https://projects.apnews.com/features/2023/climate-change-disease-animals-warming-earth/index.html

Raw Data Source Used to Make the Visualization:

https://public.emdat.be/

Description

Data Semantics	Data Attribute Types	Marks	Channels	Notes
The type of climate-related disaster	Nominal	Circle	Colour: - Orange for drought - Green for flood - Teal for storm	
The number of people affected by the climate-related disaster (per million global population)	Quantitative	Circle	Size/Area: - Smallest area represents 100 people per million affected - Medium area represents 1,000 people per million affected - Largest area represents	

			5,000 people per million affected	
The number of climate-related disasters occurring per year	Quantitative	Circle (the dotted rings)	Size/Area: - The inner ring marked with 10 represents the boundary of 10 disasters The middle ring marked with 20 represents the boundary of 20 disasters - The outer ring marked with 30 represents the boundary of 30 disasters	*This number is out of the filtered top 1000 climate disasters that occurred between 1980-2021
Year	Temporal	Line	Line Width/Thickness: - Thicker lines represent every 5 years	
?? (Missing Data Semantic)	??	Rectangle	Size/Area	*located underneath each year, the data semantic is never explained anywhere in the article or legend

Artifacts of this visualization include the horizontal axis position and vertical axis position, as the data is intended to be read in a clockwise direction instead.

The intent of this visualization is to depict the top 1,000 climate-related disasters that have occurred between 1980 and 2021, whether each of those disasters was a drought, flood, or storm, and the number of people that were affected by each disaster. The accompanying text alongside this image notes that the authors wanted to illustrate this in order to bring light to the fact that erratic weather patterns caused by climate change create conditions for diseases to easily spread, thus harming humans.

Critique

To begin, the expressiveness of a visualization refers to whether the chosen data channels match the data attribute types. In regards to this visualization, the expressiveness is very low for several reasons. From looking at the raw dataset, it is clear that the data semantic of 'the number of people affected by the climate-related disaster (per million global population)' is quantitative, and should be thus reflected by an increasing numeric scale on the vertical axis. However, the authors of this article have decided not to utilize a vertical axis, and instead rounded the numbers to the nearest groupings of 100, 1,000, or 5,000, which are then represented on the visualization in terms of area. This is especially problematic as

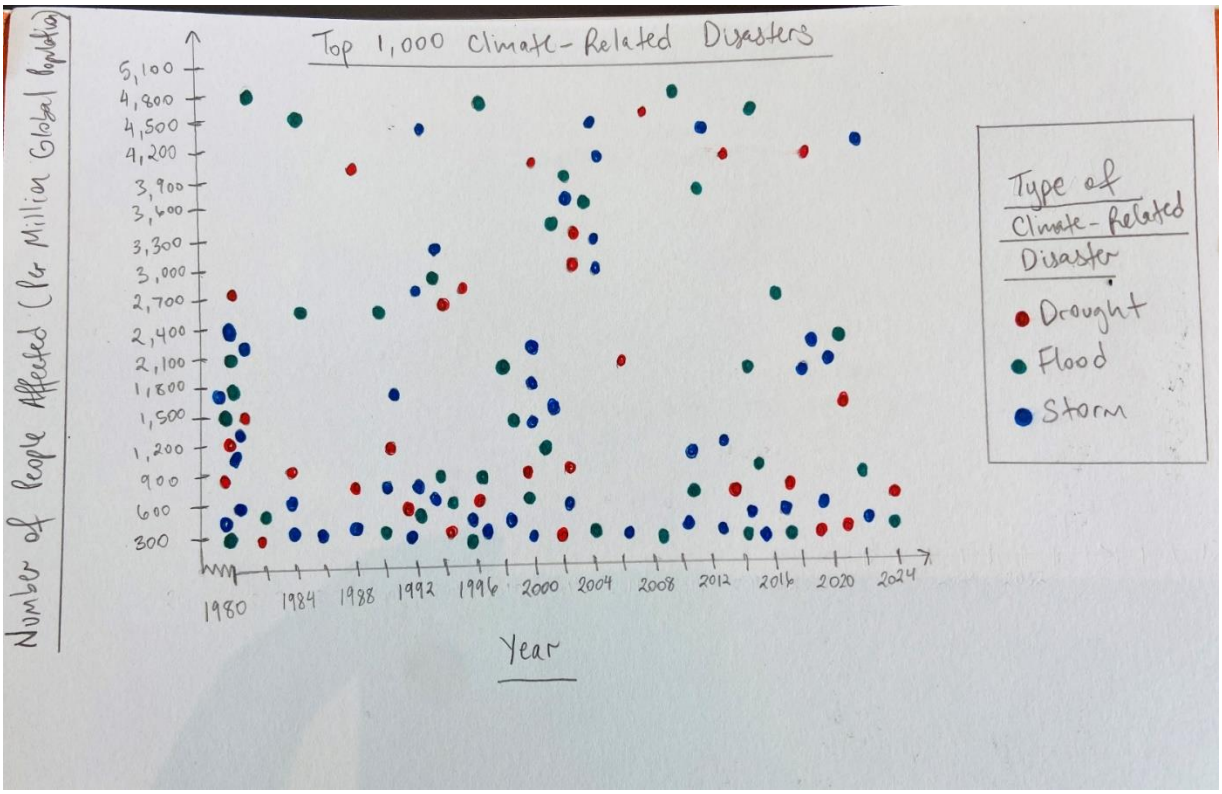
humans inherently are not great at accurately being able to discern area, making it difficult to distinguish between the 100 and 1,000 values (which are depicted quite similarly in size) on the chart. Another quantitative attribute, the 'number of climate-related disasters occurring per year,' is also not well represented. While the use of circles of increasing area marked with the numbers 10, 20, and 30 generally denote approximately how many disasters occurred per year, the use of these ordinal groupings does not accurately reflect the exact number of disasters that occurred per year, and makes it extremely difficult to determine the exact number. Additionally, the temporal attribute of the years from 1980-2021 are strangely depicted in a clockwise pattern instead of on the horizontal axis. This then leads to the weird interpretation of the year of 2021 time travelling back into 1980 at the top of the visualization. Most peculiarly, there are grey rectangles (the data mark) of varying areas (the data channel) located underneath each year that are never explained anywhere in the legend or article text or raw data, and I have no idea what they represent at all.

Next, the effectiveness of a visualization refers to whether the importance of each data attribute matches the data channel rankings. Unfortunately, this does not appear to have been accomplished any better than the expressiveness of the visualization. Firstly, the nominal data of the type of climate-related disaster (whether it was a flood, drought, or storm) are not well-distinguished from each other by the chosen colours: orange, green, and teal respectively (the teal and green are too similar and interfere with each other). The orange appears to popout as if to signify that droughts have a greater sense of importance compared to the other disaster types, which was not the intention of the authors, nor an accurate reflection of the raw data. The chosen areas to represent the groups of number of people affected by each disaster are weirdly scaled to imply that the values of 100 and 1,000 are a lot closer to each other in terms of significance than they should be, with only the 5,000-value group being easily discernable on the chart. There is also extreme interference from all of the different data attributes overlapping on top of each other (circles on top of circles within a big circle).

Finally, the questions that the visualization are effective at answering include which years had higher or lower numbers of major disasters occurring, how many people in general were affected by each disaster, and what the disaster type was.

Redesign Explanation

Redesign Guidelines I: Effectiveness



In my redesign, my main goal was to make the visualization simpler and easier to understand. I decided to encode my data in a scatterplot format (with each circle representing a disaster that occurred), with 'Year' on the horizontal axis and 'Number of People Affected (Per Million Global Population) on the vertical axis. I kept the same name of the plot ('Top 1,000 Climate-Related Disasters') to denote that the plot only shows the top 1,000 global climate-related disasters. Next, I added a legend to the right-hand side to denote whether each disaster was a drought, flood, or storm by the colour.

By redesigning the visualization in this manner, I addressed the issues of the data semantics being represented by the wrong data channels (such as quantitative data being represented by ordinal groupings), eliminated the accidental popout as all three colours of red, blue, and green are now equally distinct, made the timescale more easily legible along the horizontal axis to show the increase in time, and made the quantitative values of the number of people affected more apparent along the vertical axis. Therefore, I believe that my visualization now better addresses the original questions of the type of disasters that occurred alongside the number of people that were affected bye each disaster. However, some limitations of my design include the fact it is still hard to count the exact number of disasters out of the top 1,000 that have occurred each year, and that I faced physical limitations in how much of the data that I could reasonably draw by hand.

Redesign Guidelines II: Aesthetic

For this redesign, I have replaced the data marks for each disaster type with symbols of a flame, waterdrop, and lightning bolt. I also made the visualization "interactive" by showing that by hovering over each data point or by hovering over each year, a data box pops up with more information.

Video Reflection

Address the three prompts.

Watch the video https://www.youtube.com/watch?v=fp-WNqaQG0s Reflect on it by addressing the three questions below:

1. Prior to watching the video, what was your understanding of who created vis and how they were created?

   Prior to watching the video, I thought that scientists and researchers were the ones who had created visualizations in order to better showcase their research results. Before computers became commonplace, I assumed that they were all hand-drawn and therefore only showcased a few data points deemed the most important, due to the physical limitations of the visualizer.

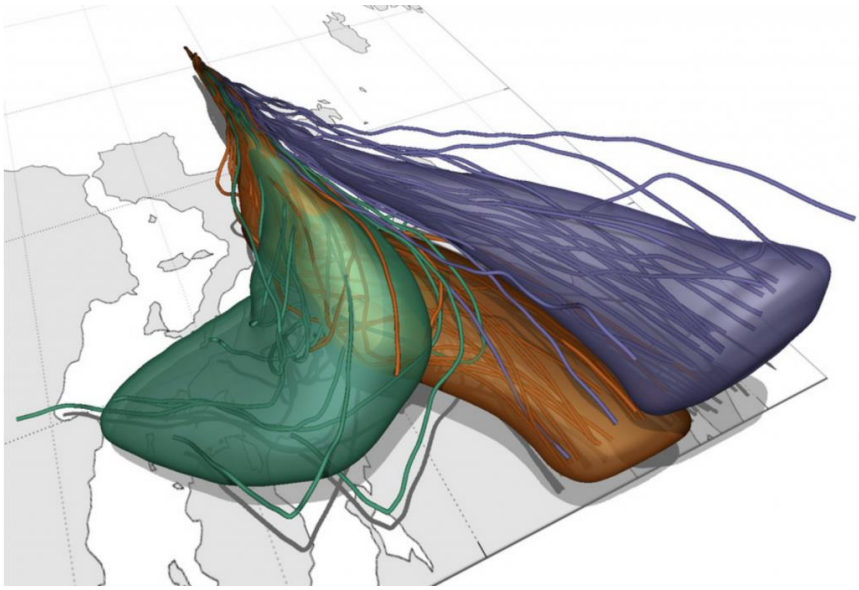
2. What was the most interesting thing you learned?

   The most interesting thing was learning about xenographphobia, which is the fear of unknown or weird charts. It's shocking that this is a real fear! I believe that this only further highlights how important it is to ensure that visualizations are effective and expressive in order to ensure that any viewer can easily and quickly understand it.

3. What was the most surprising thing you learned?

   I didn't know before that William Playfair, a Scottish engineer, was the one who invented the line, pie, and bar chart. Having learned that fact, I'm surprised that Scotland has not widely advertised this or leaned into becoming the hub for data science and data visualization.

Xeno Graphics



Reference: https://xeno.graphics/streamline-variability-plot/

My favourite plot from the Xenographics website is the streamline variability plot. This plot aims to show how an ensemble of streamlines flows through a selected location within Euclidean space (AKA 2D or 3D).