A2: Sketching Visualizations

CSE 412: A2 Sketching Visualizations

Jaiden Atterbury, jatter41@uw.edu, Section AA

Part 1

Sketching U.S. Population Growth

The Dataset

US Population Since 1840
(in thousands, rounded to the nearest thousand)

	Boston, MA	St. Louis, MO	Seattle, WA
1840	93	17	-
1850	137	78	-
1860	178	161	0.2
1870	251	311	1
1880	363	351	4
1890	449	452	43
1900	561	575	81
1910	671	687	237
1920	748	773	315
1930	781	822	366
1940	771	816	368
1950	801	857	468
1960	697	750	557
1970	641	622	531
1980	563	453	494
1990	574	396	516
2000	590	347	564
2010	618	319	609
2020	676	302	737

Data sources: Boston (biggestuscities.com, Census), Cleveland (biggestuscities.com, Census), Seattle (biggestuscities.com, Wikipedia)

Brief Process Overview

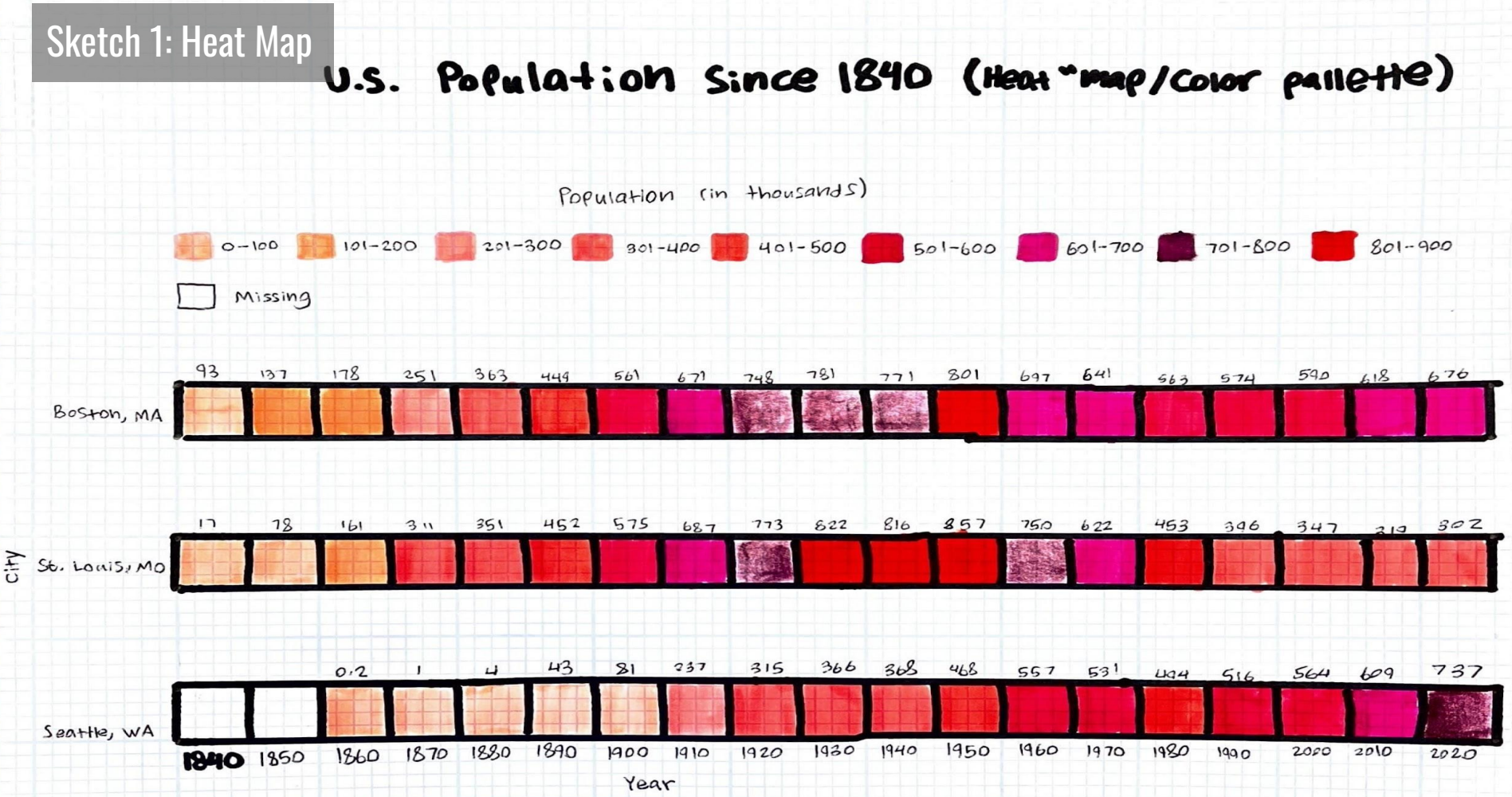
How did you approach the design process?
When starting this part of the assignment, the first thing I did was study the provided dataset. By looking at all of the numbers, as well as missing data points, I had a general sense of what kind of questions I wanted to answer (detailed below) as well as the types of visualizations I needed to make to answer these questions. After this brief peak at the dataset I chose my three visualization types and started sketching. The three visualization types that I chose to sketch were, a heat map/color palette, a time series/line graph, and a vertically stacked bar chart. All three of these visualizations allowed me to answer my primary research question, while also unveiling new parts of the dataset that a single visualization could not do on its own.

What question(s) did you attempt to answer with your visualizations?
My initial intuition, after looking at the population data over the years, was to simply answer, “How has the population in three major U.S. cities changed since 1840?” However, once I noticed that the dataset contained a city from each region of the U.S. (i.e. west coast, central, and east coast), I decided a more relevant and thoughtful question would be, “How has the population in specific regions of the U.S. changed since 1840?”

What did you struggle with?
One of the main things that I struggled with, and never stopped struggling with throughout the entirety of the assignment, was accurately scaling data encodings. For example, in the heat map I made, the color scale I used was not linear and made it very difficult to interpret which value was which, especially when my specific color scale relied on color saturation. Furthermore, in my stacked bar charts, although the total population value for each decade was accurate, the city that took the middle bar was not as accurate as the top or the bottom bar, and this was evident when comparing the city with the middle bar with its corresponding line on the line graph.

What would you do differently next time?
If I were to do this assignment again, I would make sure that I focused less on perfect sketches, and a better amount of time being more creative with my visualization types. By trying to get my sketches visually appealing, I spent an unneeded amount of time on the drawing portion of the assignment, which I could of spent on trying to come up with more creative visualizations. Furthermore, if I would’ve spent less time on the sketching itself, I could’ve spent that wasted time polishing my analysis and rationale portion of the assignment. The last thing that I would do differently when it came to completing part 1, is doing more research on different visualization types, so that I could sketch visualizations that I’ve never used/that unlock certain parts of the data that the other visualizations couldn’t.

Sketch 1: Heat Map



Sketch 1: Design Rationale

Brief description
For my first sketch, the one intended to be more abstract/artistic, I chose to use a heat map in the form of a color palette. In particular, the heat map represents the changes in population since 1840 of each respective U.S. city in the dataset. The heat map portrays these changes through color-coding population and assigning a color to each year for each city.

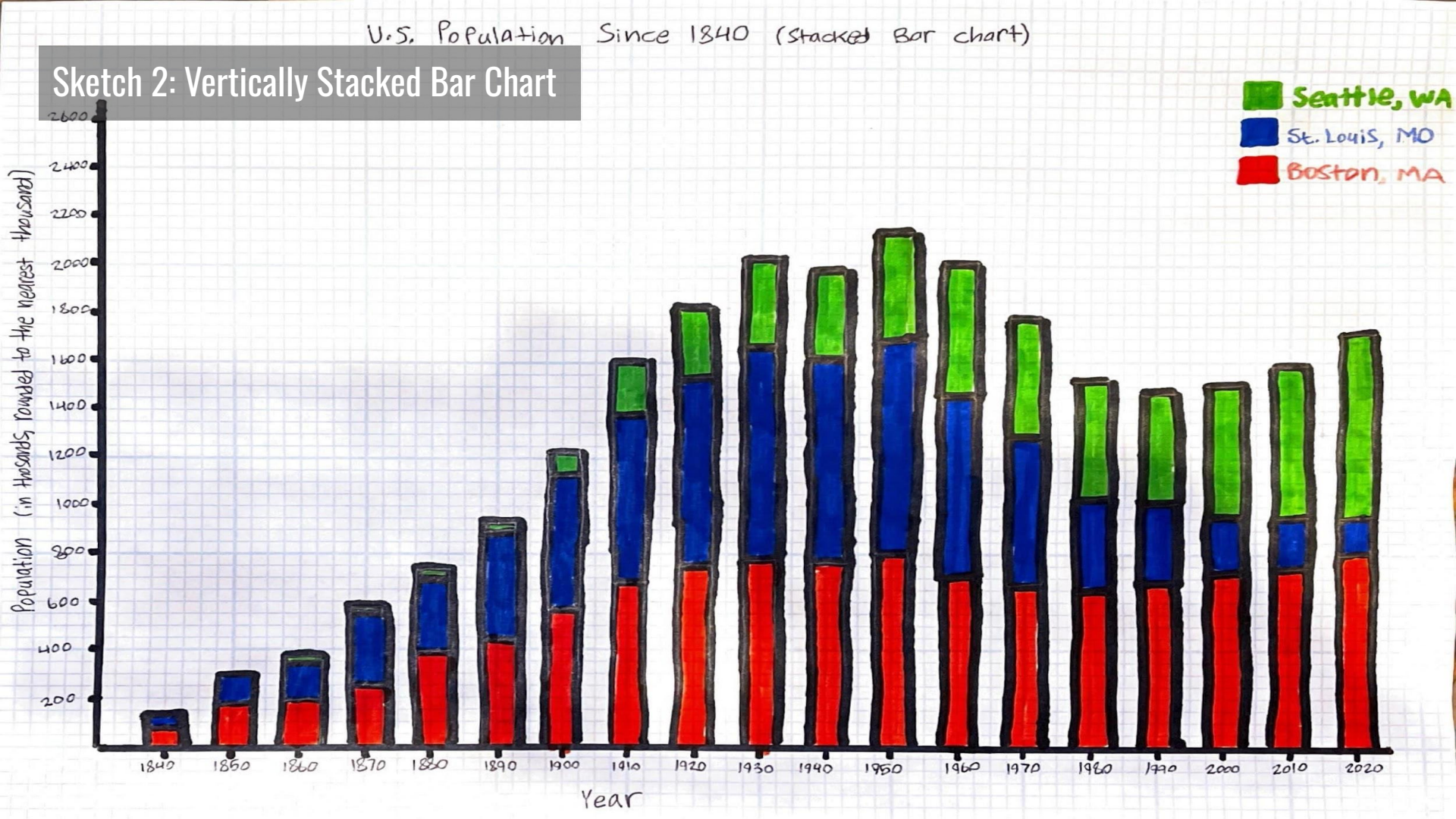
Data representation
The population data provided in the dataset is encoded in the heat map through the use of color corresponding to certain population ranges as denoted by the legend underneath the title. In particular, each population range, in which there are 9 total, are assigned to a different color. Furthermore, in this specific heat map, I decided to try and use similar colors, each in a differing shade (pink, red, purple, etc.) with the darker the color corresponding the higher population numbers, instead of using completely distinct colors. By using similar shades that "blend together," trends in data are easier to visualize without having to constantly consult the legend, as compared to the completely discrete colors case. The y-axis contains the three cities, and the x-axis contains the different decades. The order of the cities on the y-axis, from top to bottom, is Boston, St. Louis, and lastly, Seattle.

Context of use
Given the fact that this heat map is in the form of a color palette, one can see how the population in a given city changes from decade-to-decade, simply by seeing how the shades change between nearby blocks on the palette. Moreover, since the population data is encoded so that darker colors correspond to higher values of population, one can quickly see how the growth of multiple cities compare in a certain time frame, simply by comparing how dark the shades are between the two cities, as well as the direction the shade is going in. In particular, significant changes in shade correspond to cities in which total population is volatile, while on the contrary, infrequent shade changes correspond to cities with no change or steady population changes.

Key strengths/weaknesses
Strengths: The most important strength when it comes to heat maps, is that they allow for the abstraction of data. What I mean by this is that the ranges of population make it so that the exact population in a given year isn't the main focus of the visualization. By encoding data in this way, users can instead focus on general trends in the populations instead of the exact numbers themselves. This becomes beneficial when small changes aren't of great importance. Another strength of heat maps, that is infrequently brought up, is that they are visually appealing. Since the heat map is more visually appealing than the average visualization, users are more likely to be drawn into the story it is telling, in a business setting this could increase click-through rate, profits, etc.

Weaknesses: The biggest weakness of heat maps, in particular heat maps that use progressive/similar shades, is that users can sometimes have a hard time differentiating between shades that are very similar. This is especially relevant in my sketch, as I eyeballed the color scale that was used, and in retrospect, some of the colors are in the wrong place, which led to me having to include the exact population number itself, which I was trying to avoid. Thus, when the coloring is too similar or in the wrong order, all of the benefits and functionalities of using a heat map are no longer present. With that being said, the most fundamental part of heat map is the choice of the color scale. By putting "all of our eggs in one basket," heat maps can be very hit-or-miss in their effectiveness.

Sketch 2: Vertically Stacked Bar Chart



Sketch 2: Design Rationale

Brief description
For my second sketch, one of the sketches intended to be more concrete/informative, I chose to use a vertically stacked bar chart. In particular, each of the bars in the vertically stacked bar chart represents the total population of the three U.S. cities in a specific year from 1840-2020. Furthermore, each bar is broken up into three "sub-bars," each of which represents the total population of the given city in the specific year.

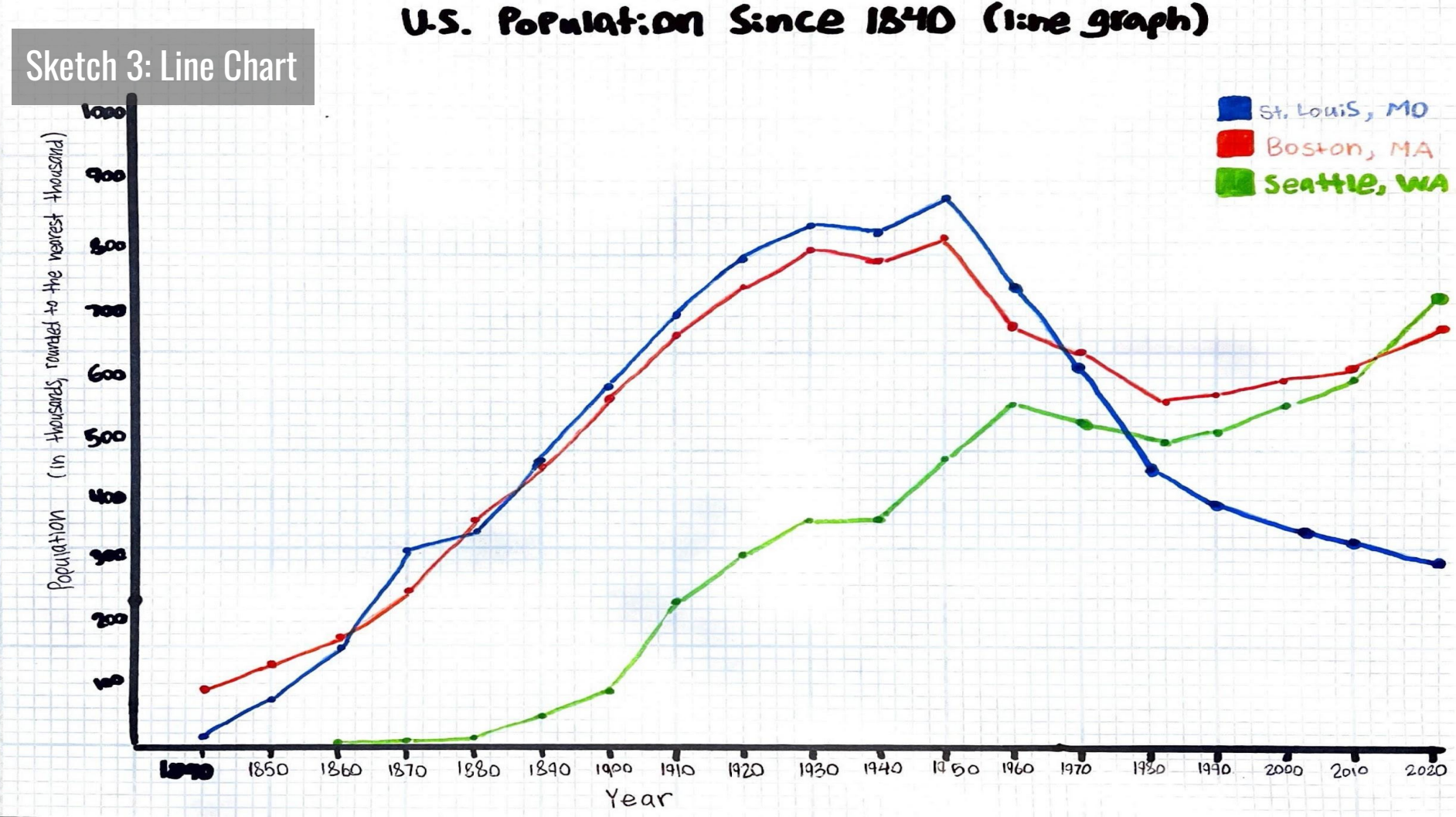
Data representation
As mentioned above, each bar in its entirety represents the total population of all of the three represented cities combined, with the value associated with each bar being represented by the height of the bar itself. The specific value that the bar represents is the corresponding number on the y-axis that is in line with the tallest point of the bar. It is important to note that the x-axis in this case corresponds to the year/decade the population value was taken in. Furthermore, the individual population in each of the three cities is represented as a colored "segment" or "sub-bar" of the entire bar as a whole. The meaning of each color is presented in the legend in the upper-right of the visualization. The population value associated with each sub-bar is also represented by its height, although it is calculated differently than that of the entire bar. This value is calculated by subtracting the y-value corresponding to the bottom of the bar from the y-value corresponding to the top of the bar. As far as the cities are concerned, from top to bottom, Seattle is given green, St. Louis is given blue, and Boston is given red.

Context of use
Unlike the heatmap mentioned in a previous slide, the colors serve no meaning other than representing the different cities. This change in the importance of colors shifts how the users will interact with the data. In particular, users are more focused on the height of the bars in order to gain understanding. The stacked bar chart is good for seeing totals, as well as proportions of the totals based on a certain number of groups. In the lens of population growth, we can see how the total population of these three cities change over time (more precisely than the heat map). At the same time, this visualization also allows us to answer the main research question, as by segmenting the total population by city, we can see how these proportions compare, which will give us a better understanding about how the changes of population in these different regions relate and differ from each other, especially over the decades.

Key strengths/weaknesses
Strengths: One strength of this vertically stacked bar chart, as mentioned above, is that it allows us to see two aspects of population at once. What I mean by this is that we can see how total population has changed, but at the same time we can see how the total population has changed in each of the three featured cities. Another strength of the vertically stacked bar chart, is that proportions of total population are more easily viewed, especially when compared to just looking at raw numbers. Lastly, due to the simplicity of a bar chart, it is easy for anyone to understand, without too much prior knowledge.

Weaknesses: One weakness of this visualization, is that although it is easy to compare population from decade-to-decade, due to how value is encoded into height, it is not as easy to see the exact values of each decade, and it is even harder to find the exact value corresponding to each city in each decade. Unlike a side-by-side bar chart, this stacked bar chart makes it hard to compare change in population between cities over time, as the differing total population heights make small differences hard to detect/compare (especially in the middle bar).

Sketch 3: Line Chart



Sketch 3: Design Rationale

Brief description
For my third and final sketch, one of the sketches intended to be more concrete/informative, I chose to use a line chart/time series to visualize the population dataset. The line chart simply shows how the population in each city has changed over time by plotting points and connecting them with a line (with different colors corresponding to different cities).

Data representation
The way in which the line chart is built is rather simple, for each decade and for each city, the corresponding population value is plotted, with year on the x-axis, and population on the y-axis. After all the points are plotted, points belonging to the same city have lines put between them (pairwise). In order to differentiate between cities, each city is given its own unique color (red, blue, and green). In particular, St. Louis is given blue, Boston is given red, and Seattle is given green. Like the bar chart, and unlike the heat map, color does not encode any value, it is used only to differentiate between the three cities.

Context of use
This visualization type is best used when wanting to compare how populations change over time, especially across certain groups. By looking at the height of data points in a specific decade, we can easily see which city has a higher population. Unlike the stacked bar chart, the height of each point corresponds to the exact value that data point takes on, this is different from the bar chart variation as values for sub-groups need to be calculated using differences of heights. Due to this feature of the line graph, direct comparisons are much easier using a line-chart. Furthermore, due to the connecting lines between consecutive points, one can compare the rate-of-change of each city between two decades, by looking at the slope. These slopes values help us get a general sense of how the population growth in each of these regions have changed over time.

Key strengths/weaknesses
Strengths: One key strength of the line chart, as mentioned above, is that direct comparisons of the exact population values of the three cities are much easier to perform as compared to the previous two visualizations. Another critical feature of the line chart, is that the points and the slope values help us answer two questions at the same time. These questions being, “How does population change over time in these three regions?” as well as, “How *quickly* does population change over time in these three regions?”

Weaknesses: One key weakness of this line chart (and all line charts), that is seldom mentioned, is the fact that population (or any feature being measured) may not change linearly, which the visualization seems to depict. Also, as with most visualizations, the exact population value for each city in each decade can be hard to estimate due to the large ranges on the y-axis, meaning most comparisons will be approximate.

Reflection

Compare your designs. What are their relative strengths and weaknesses?

As can be seen above, the three ways I chose to visualize the data were by using a heat map, a vertically stacked bar chart, and a line chart. Each of these visualizations were built to answer the question, "How has the population in specific regions of the U.S. changed since 1840?" Although each visualization worked to answer this question from a slightly different angle, each highlighting a different part of how population changed in each region of the U.S. since 1840, each of the visualizations had similar strengths and weaknesses. For example, each visualization used color to display an aspect of the data that wouldn't have been possible, or easily done, otherwise. On the other hand, each visualization made it difficult to know the exact population value for each city during each decade. In the heat map, one could only know the population range of each of the cities in a give decade (if they could correctly decipher the proper color), in the stacked bar chart, one could only know the correct population value of each of the cities in a given decade if they subtracted the correct population numbers, and lastly, one could only know the correct population value of each of the cities in a given decade if they correctly eyeballed the y-axis. It is important to note, as highlighted in the line chart rationale section, that these visualizations can answer more questions than the overarching question guiding this part of the assignment.

What was your favorite design and why?

My favorite design out of the three presented above is the heat map/color palette. This visualization is my favorite as it is the most unique out of the three and took me quite a long time to get exactly how I wanted. Even though the color scale isn't what it should be, nor is it how I would've liked it to turn out, I was happy to come up with a creative idea that stretched the definition of what I had originally thought of as a heat map. Lastly, the fact that I was able to morph the heat map into a color palette format was one of my more proud moments in the sketching phase of this course.

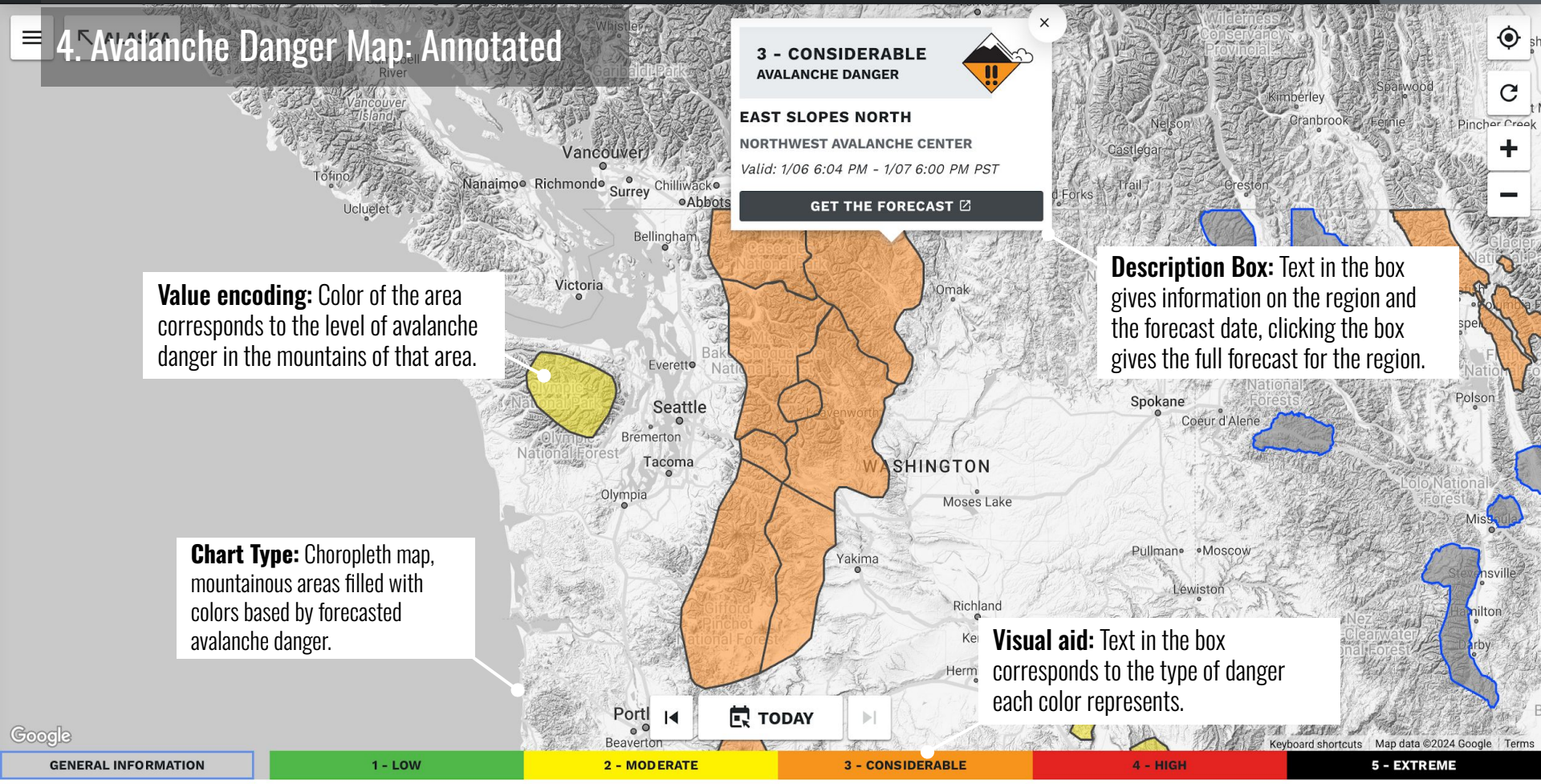
What did you learn from this part of the assignment?

This part of the assignment allowed me to learn how to turn real data into many different sketches, which is something I had never done before (apart from the few lectures we've had so far). Throughout the entirety of my academic journey, I had always used tools to visualize real data that I found on the internet. This assignment, and the aforementioned lectures, have taught me the benefit of sketching as a way to explore the data, as well as different ideas of visualizing said data. With that being said, I'd need to scale back on the quality of the visualizations I created in order to maximize efficiency and emphasize the purpose of sketches in the first place.

Part 2

Sketching Selected A1 Visualizations

4. Avalanche Danger Map: Annotated

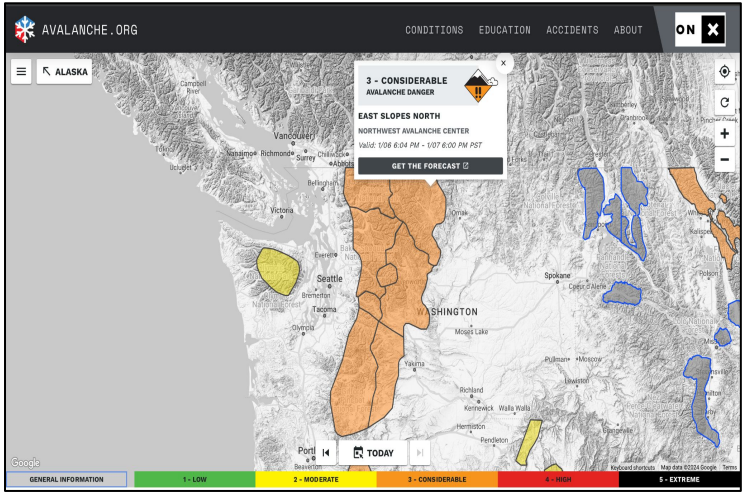


Weather Visualization

Avalanche Danger Map CRITIQUE

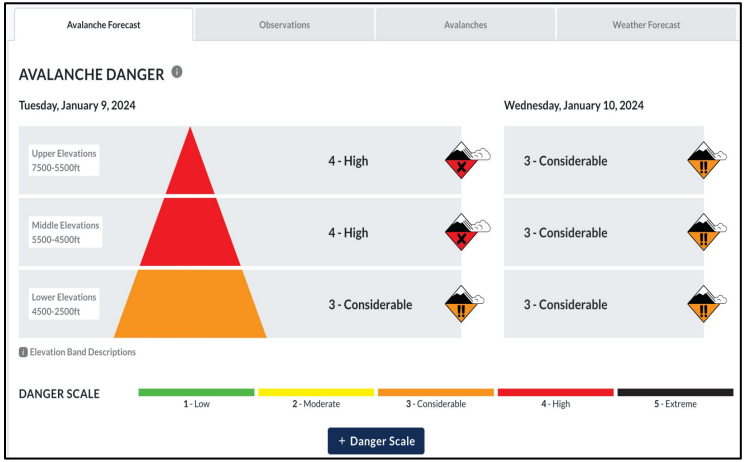
What is the goal of this visualization?

The goal of the avalanche danger map is to inform hikers, climbers, backcountry skiers, and all individuals who intend to travel in mountainous regions, about the current avalanche dangers in these specific mountainous regions. By offering accurate and up-to-date information on a dangerous phenomenon, adventurers are better equipped to make rational and safe decisions when it comes to travelling and entering snow packed areas, especially during the winter months.



How is the data represented—its visual encoding, graphical marks, color, *etc.*?

The main way in which the data is represented is through what is known as a choropleth map, also known as a heat map. A choropleth map is a data visualization tool that encodes certain areas of the map with shades of a color based on the level of a particular value of interest. In the case of the Avalanche Danger Map, the only areas that are shaded in are the mountainous areas of the United States, and these areas are filled in with different colors spanning from green to black based on the perceived avalanche risk in the corresponding area. In general, the darker the color, the greater the risk. Clicking on a specific area gives general information about the region, further clicking will lead to a more in-depth description of current conditions in said region (shown on the bottom right).



Weather Visualization

Avalanche Danger Map CRITIQUE

What are some key strengths of this visualization and why?

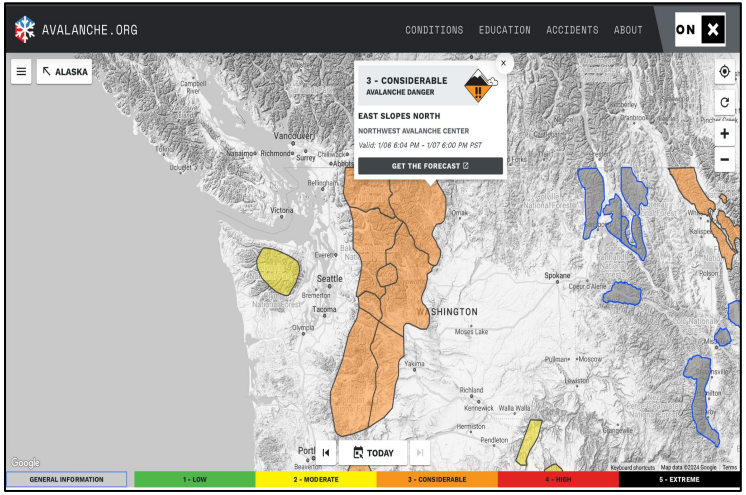
- Since there are only five discrete colors, and all of the chosen colors are unique, it is easy for users to decipher which region corresponds to which specific value, as opposed to a more continuous color encoding.

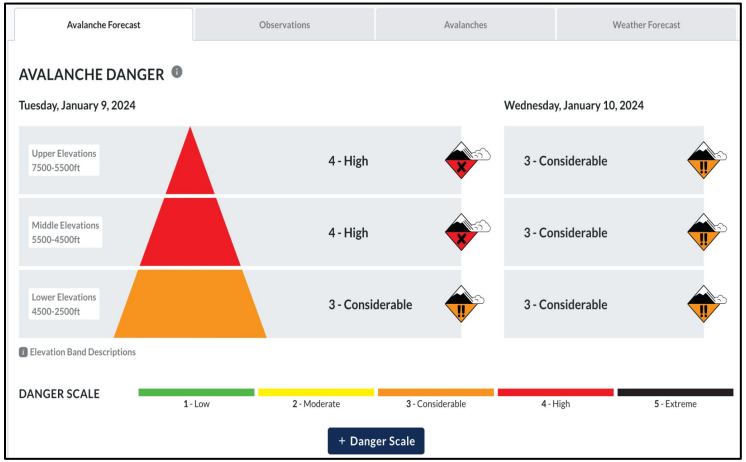
- By creating pop-ups whenever a user hovers over a region, which in turn leads to further and more descriptive forecast after clicking, allows users to do further research after initially assessing conditions with the choropleth map.

What are some key weaknesses of this visualization and why?

- Since a new user won't understand the exact meaning of the warning labels (low, moderate, considerable, etc.), this leads to the visualization being "incomplete" or "less useful" to users who haven't previously been on the website and/or done prior research.

- Due to the small amount of danger categories, some cases/areas getting the same label might not be "equal" in the sense that one could be more dangerous than the other. Users might not know this from the visualization itself, and instead would need to do a lot more further research.





Visualization 1 Sketch Alternative #1

Sketch 1: Design Rationale

Brief description
For my first variation of the avalanche danger map, I decided to think outside of the box and create a “speedometer” to visualize the perceived avalanche danger in the mountainous regions of Washington. For simplicity, only three of the mountainous regions in Washington were considered. Whatever section/area the dial is on is the forecasted avalanche danger level of that region on that date.

Data representation
In each speedometer, there are five separate numbered regions, each representing one of the categories of avalanche danger (low, moderate, considerable, etc.). Each region is color-coded based on its corresponding avalanche danger category. As mentioned above, the dial/arrow points to the forecasted avalanche danger level in that specific region. Below each speedometer, the results are summarized for clearer interpretation. Furthermore, although this feature isn’t implemented in the main visualization, in this sketch, however far to the right the dial is in a specific section, the more perceived danger the specified mountainous area is in as compared to other regions in the same danger level. This means that the dial being farther to the right in the high category equates to more perceived danger than a dial farther to the left in the same category. Different speedometers capture the danger in different areas across the state. In this specific visualization, the three regions that are visualized (for the date 01/10/2024) are West Slopes South, Snoqualmie Pass, and Stevens Pass.

Context of use
With the newly implemented “dial depth” feature explained above, people looking to explore the mountainous regions of Washington can be more accurately educated on the forecasted avalanche dangers of the areas they intend to go to. In that sense, as compared to the purely categorical nature of the data presented in the map, users can see “how much” of a certain danger category a specific region is in, instead of relying on a catch-all category to influence their decision making. For example, if a group of people are only willing to climb a mountain if it is in the “considerable” danger level or below, and they see that the region they are going to climb in is barely in the “high” danger level, they may be more willing to give it a shot as compared to if they had just viewed the map.

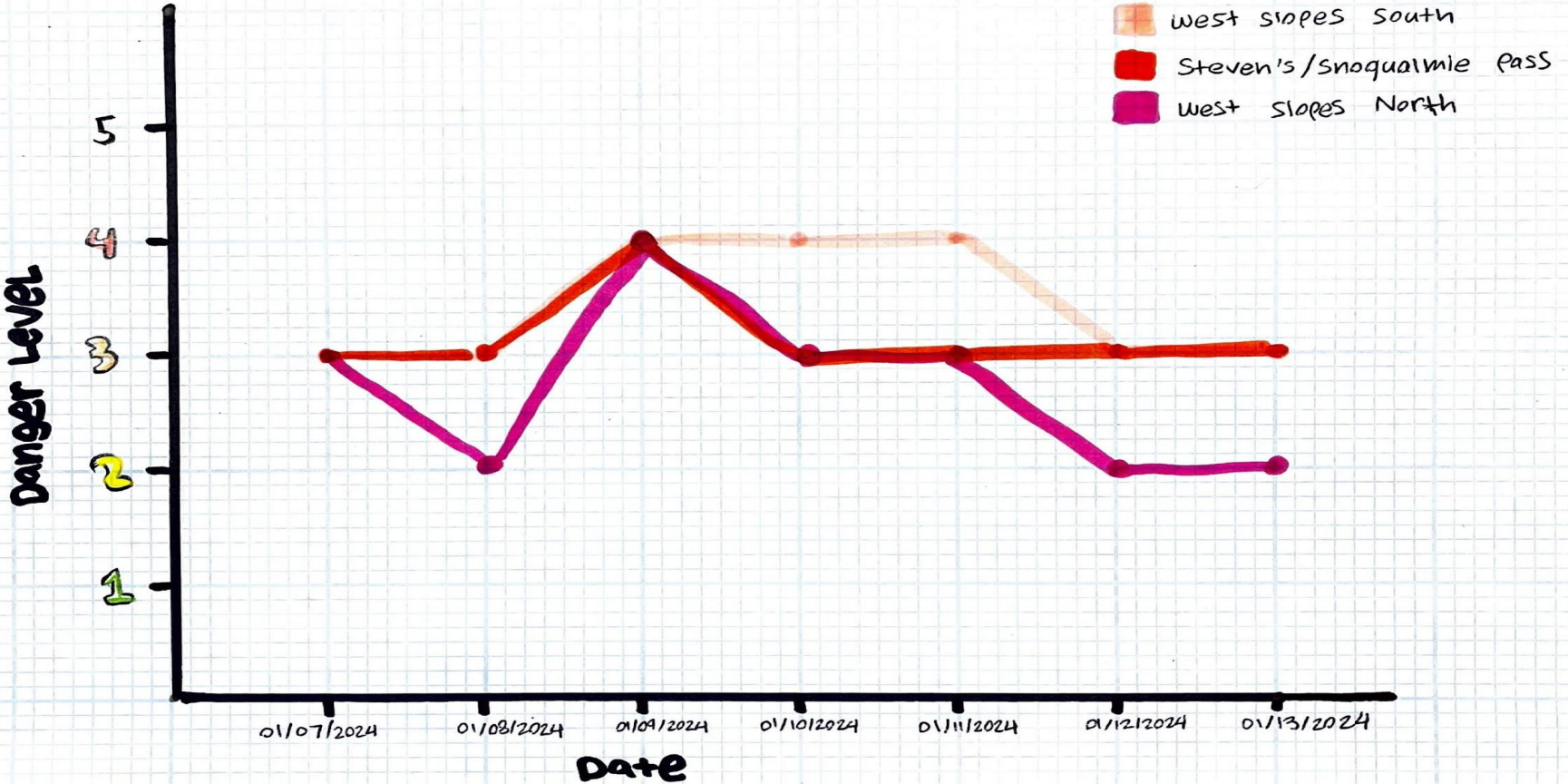
Key strengths/weaknesses
Strengths: Given that the speedometer has all of the avalanche danger information in one place, as compared to the choropleth map that has the information spread out in a legend and in a pop-up, this allows viewers to more easily understand what the forecast is in their region without having to rely on interactivity. Furthermore, by encoding the dial position to correspond to the level of perceived danger within a given category, this visualization type allows for a continuous representation of categorical data.

Weaknesses: Since there must be a new speedometer for each different mountainous region in Washington, this can take up a lot of space and make displaying all of the forecasts on one page cost inefficient. Furthermore, unlike the choropleth map that allows viewers to quickly see the forecast based on geography, the speedometers don’t make it as easy to look up the forecast for a specific region, as you would have to locate the specific region throughout a list of many speedometers. This becomes even more of a problem when the regions don’t have a forecast, which can happen.

Avalanche Danger in Washington (line chart)

Visualization 1 Sketch Alternative #2

Week Ending in 01-13-2024



Sketch 2: Design Rationale

Brief description
For my second variation of the avalanche danger map, I decided to go back to a classic and more concrete visualization type in the line chart to visualize the perceived avalanche danger in the mountainous regions of Washington, over time. For simplicity, only three of the mountainous regions in Washington were considered. Whatever portion of the y-axis a given point is on corresponds to the forecasted avalanche danger in that region on that date.

Data representation
The way in which the line chart is built is very similar to the line chart for the population data, for each date in a given week and for each mountainous region in Washington, the corresponding avalanche danger value is plotted, with date on the x-axis, and perceived avalanche danger on the y-axis. After all the points are plotted, points belonging to the same mountainous region have lines put between them, this is done pairwise. In order to differentiate between the mountainous regions, each region is given its own unique color (pink, red, and violet). In particular, West Slopes South is given pink, Snoqualmie Pass/Stevens Pass is given red, and West Slopes North is given violet. Unlike the speedometer, color does not encode any value, it is merely used only to differentiate between the three mountainous regions. Furthermore, unlike the speedometer, there is no meaning associated to the spaces between the different categories on the y-axis (same for the x-axis).

Context of use
This visualization has many similar uses to the speedometer, such as viewing the forecasted avalanche danger in a specific mountainous region on a given date. However, it has many other uses that change depending on if the plot is viewed in retrospect or as a forecast. If this visualization is viewed at the beginning of the week, users can see how the forecasted avalanche danger in certain mountainous regions around them are supposed to change throughout the given week, this can be used to plan future trips, or to avoid trips to stay away from dangerous situations. On the other hand, if this visualization is viewed in retrospect, viewers of the visualization can see how the danger in certain areas have changed in recent days, months, and even years. This trend information could inform viewers on what to expect during certain times of the year, even when a forecast for the future isn't available.

Key strengths/weaknesses
Strengths: Given that the line chart shows the forecast, or previous forecasts, users have the ability to plan for the future or know what to expect in the future as compared to static/singular day visualizations like the speedometer, which don't give users the foresight to make informed decisions in the long run. The line chart, just as is the case with the choropleth map, stores all the necessary visualization in one screen, which allows for quick lookup.

Weaknesses: Since the line chart, unlike the speedometer, doesn't allow for a continuous encoding of categorical data, decisions at the margins can't be made. Furthermore, since a continuous encoding of categorical data isn't available, there is a lot of overlap present in the line charts. For example, Steven's Pass and Snoqualmie Pass had an identical forecast for the week ending in 01-13-2024. This can become a serious issue when more mountainous regions are put on the chart. The result would be even more unreadable than the typical non-categorical line chart.

Visualization 1 Sketch Alternative #3

Avalanche Danger in Washington (fill-in-dot chart)

For Date ending in 01-10-2024

West Slopes South

4 - HIGH

Snoqualmie Pass

3 - CONSIDERABLE

Stevens Pass

3 - CONSIDERABLE

Sketch 3: Design Rationale

Brief description
For my third and final variation of the avalanche danger map, I decided to go back to being creative and chose what I call a "fill-in-the-dot chart" to visualize the perceived avalanche danger in the mountainous regions of Washington, over time. For simplicity, only three of the mountainous regions in Washington were considered. Whatever dot is filled in corresponds to the forecasted avalanche danger of that region on that date.

Data representation
In each "fill-in-the-dot chart," there are five separate numbered dots, each representing one of the categories of avalanche danger (low, moderate, considerable, etc.). Whichever dot is filled in corresponds to the forecasted avalanche danger in the specific region. The specific dot that is filled in is color-coded based on its corresponding avalanche danger category. Below each "fill-in-the-dot chart," the results are summarized for clearer interpretation. Different "fill-in-the-dot charts" capture the danger in different areas across the state. In this specific visualization, the three regions that are visualized (for the date 01/10/2024) are West Slopes South, Snoqualmie Pass, and Stevens Pass.

Context of use
The "fill-in-the-dot chart" explained above, allows people looking to explore the mountainous regions of Washington state the ability to be more accurately educated on the forecasted avalanche dangers in the areas they intend to go to. By being more aware of the surrounding areas, viewers of this visualization can make more informed decisions on something that directly impacts their safety. This chart would rarely be used in any other contexts, as it has severe limitations and is limited in usage, however it offers an alternative to the choropleth map and can be viewed as simpler by many.

Key strengths/weaknesses
Strengths: Given that the "fill-in-the-dot charts" have all of the avalanche danger information in one place, as compared to the choropleth map that has the information spread out in a legend and in a pop-up, this allows viewers to more easily understand what the forecast is in their region without having to rely on the interactivity of the choropleth map. Furthermore, the simplicity of the chart allows for more users to understand its contents, unlike the multi-layered choropleth map and speedometer that are more nuanced and confusing.

Weaknesses: Unlike the speedometer, no "in-between" values are encoded so all decisions must be made solely based on the specific category given to that region. Furthermore, unlike the line chart, no trends or future forecasts can be analyzed, thus planning must be done on a day-to-day basis. Lastly, unlike the choropleth map that allows viewers to quickly see the forecast based on geography, the "fill-in-the-dot charts" don't make it as easy to look up the forecast, as you would have to locate the specific region throughout a list of "fill-in-the-dot charts." This becomes even more of a problem when the regions don't have a forecast, which is possible.

Visualization 1 Reflection

Compare your designs. What are their relative strengths and weaknesses?

As can be seen above, the three ways I chose to re-visualize the avalanche danger map were by using a speedometer, a line chart, and a "fill-in-the-dot chart." Each of these visualizations were built to visualize the forecasted avalanche danger in all of the mountainous regions in Washington. Although each visualization worked to visualize this data in a slightly different way, each highlighting a different part of the avalanche forecast, each one of the visualizations had similar strengths and weaknesses. For example, each visualization used color to display an aspect of the danger forecast that wouldn't have been possible, or easily done otherwise. In this case, the different categories of avalanche danger were color-coded, as well as the different lines in the line chart. On the other hand, each visualization, when compared to the original choropleth map, was unable to geographically depict the avalanche danger data. Thus, if someone is unaware of the different mountainous regions in the state, they will have a hard time knowing exactly where on the map each region corresponds to.

What was your favorite design and why?

My favorite design out of the three presented above is the speedometer. This visualization, like the heat map/color palette in the first part of the assignment, is my favorite as it is the most unique out of the three and took me quite a long time to get exactly how I desired. I had never created or seen a speedometer to represent data before, so I was very pleased that I was able to come up with the idea on my own during quiz section. I also liked how I was able to incorporate my own feature into the visualization that wasn't present in the original visualization, that being the continuous encoding of the avalanche danger categories.

What did you learn from this part of the assignment?

This part of the assignment allowed me to learn how to turn a choropleth map into different visualizations. Before embarking in this part of the assignment, I didn't think I would be able to turn a choropleth map into three different visualizations. However, after quiz section I was amazed to see just how many different visualizations I was able to make out of something as simple as a choropleth map. In other words, this part of the assignment really allowed me to build confidence in my ability to make multiple versions of the same visualization.

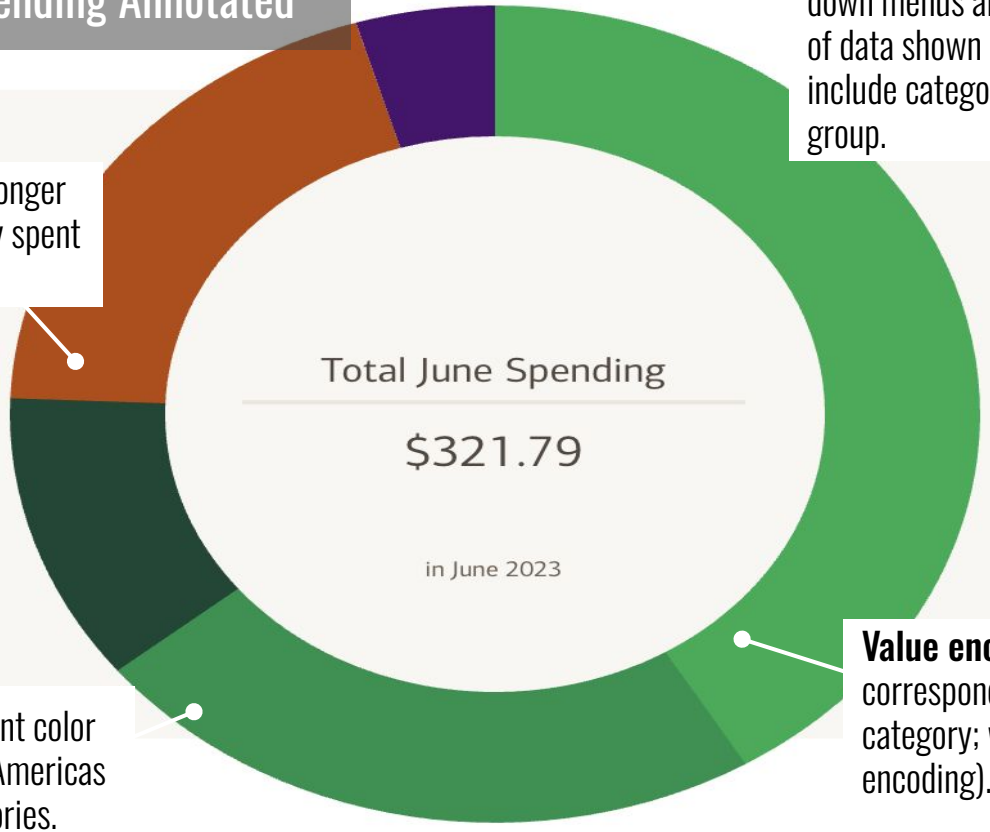
Spending for All Categories in June 2023 in Your Account Group

Lifestyle Visualization

7. Bank of America Spending Annotated

Scroll down menus: Three different scroll down menus allow you to change the type of data shown in the below chart. These include category, month, and account group.

Chart Type: Radial pie chart; longer bars correspond to more money spent in a category.



Value encoding: Length of bar corresponds to money spent in that category; width of bar is static (no value encoding).

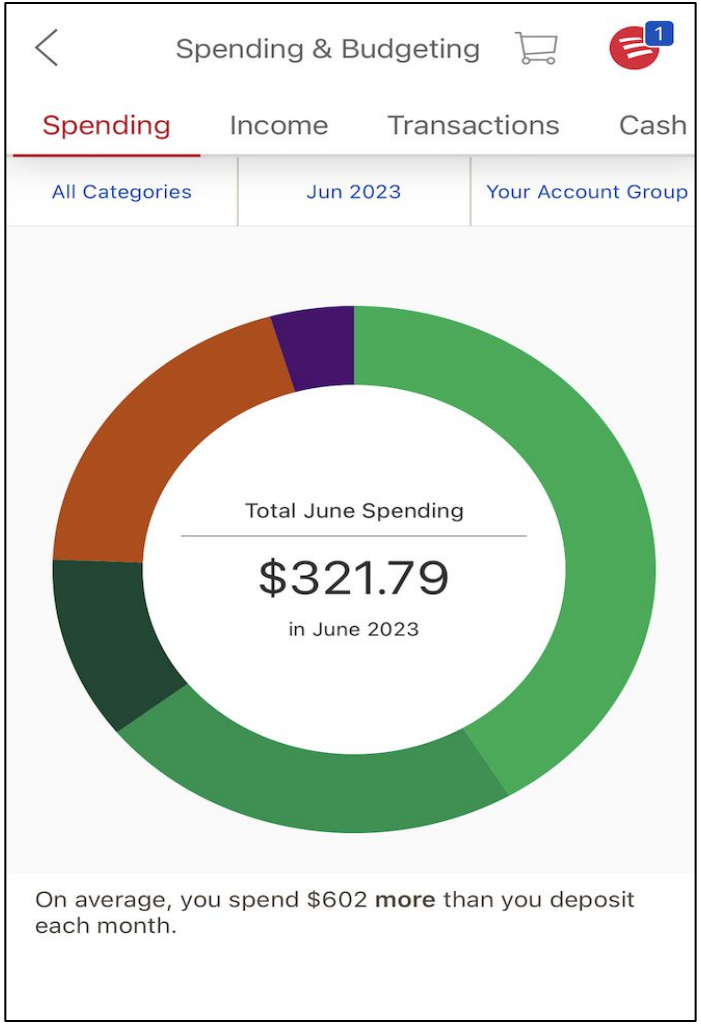
Color encodings: Each different color corresponds to one of Bank of Americas predetermined spending categories.

Lifestyle Visualization

Bank of America Spending CRITIQUE

What is the goal of this visualization?
As mentioned in Assignment 1, Bank of America is one of the most popular banking institutions in the United States. On their online platforms they have a multitude of visualizations that aid users in gaining knowledge of their spending habits, allowing users the chance to become more fiscally responsible. More specifically, this visualization in particular looks to show users their spending breakdown per month.

How is the data represented—its visual encoding, graphical marks, color, *etc.*?
In this "radial pie chart," the main spending data is broken up into different pieces, each with a differing arc length. Each arc is given it's own unique color, with each color representing a different predetermined spending category. Furthermore, the length of each arc corresponds to the amount of money spent in that particular category. In general, the longer the arc, the more money spent in that category. At the top of the visualization, there exists scroll down menus that allow the user to change the data that is presented in the chart, these options include: the categories to include, the month to display, and the account group spending should be taken from.



Lifestyle Visualization

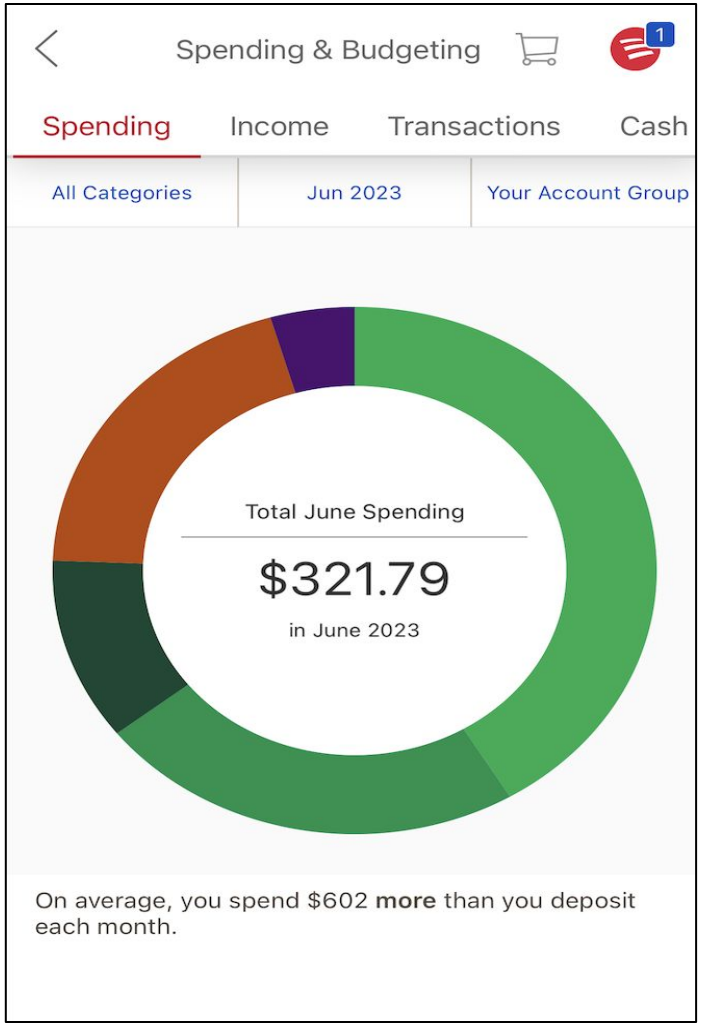
Bank of America Spending CRITIQUE

What are some key strengths of this visualization and why?

- Since the visualization includes the total spending for each month, this allows users the chance to roughly estimate how much spending each arc corresponds to.
- The ability to change certain aspects of the data through the drop down menus, allows users the opportunity to explore different aspects of their spending.
- The overall simplicity of the chart makes it easy for users to understand its meaning and intentions without much prior description.

What are some key weaknesses of this visualization and why?

- Given that the legend is off the screen, and doesn't include all of the categories without further clicking, without prior research or further clocking, users won't know the exact meaning of each category right away.
- Given that the chart is a radial pie chart, instead of a full pie chart, it is harder to view proportions of spending in the radial pie chart then it is using the latter visualization type.

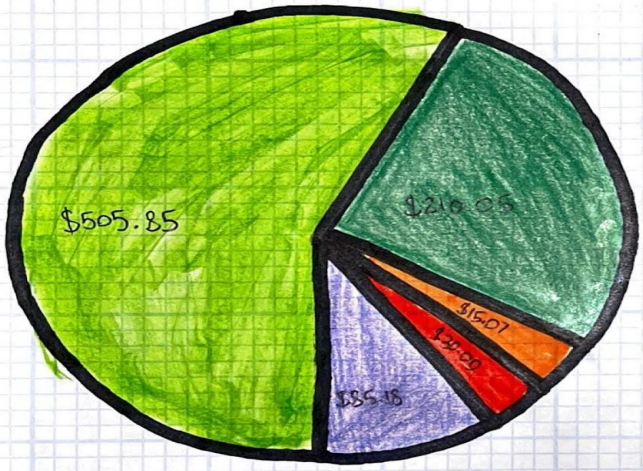


Jaiden's 2023 Spending by Category (Pie Chart)

Visualization 2 Sketch Alternative #1

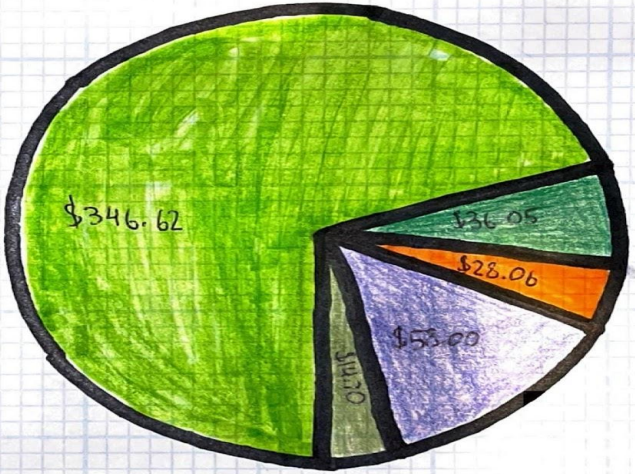
Categories: Transportation, Health, Shopping, Cash/Checks, Finance, Groceries

January 2023



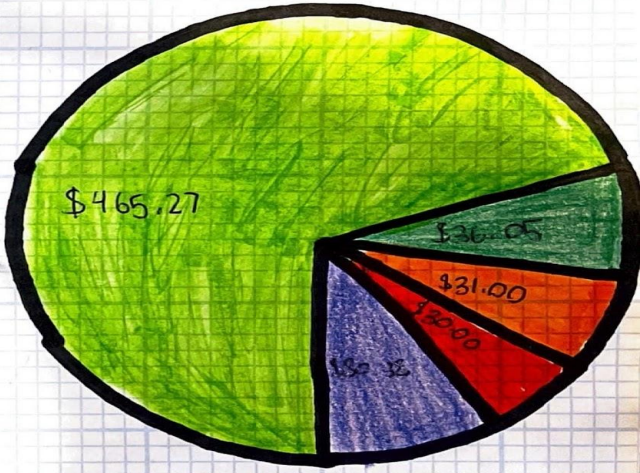
Total spending:
$846.15

February 2023



Total Spending:
$483.43

March 2023



Total Spending:
$642.70

Sketch 1: Design Rationale

Brief description
For my first redesign of the second visualization, I decided to follow from one of the weaknesses of the original visualization and make a pie chart. Due to space limitations, only the first quarter of the year is visualized. In each month, total spending is broken up into categories, which are represented by the slices/areas in the chart.

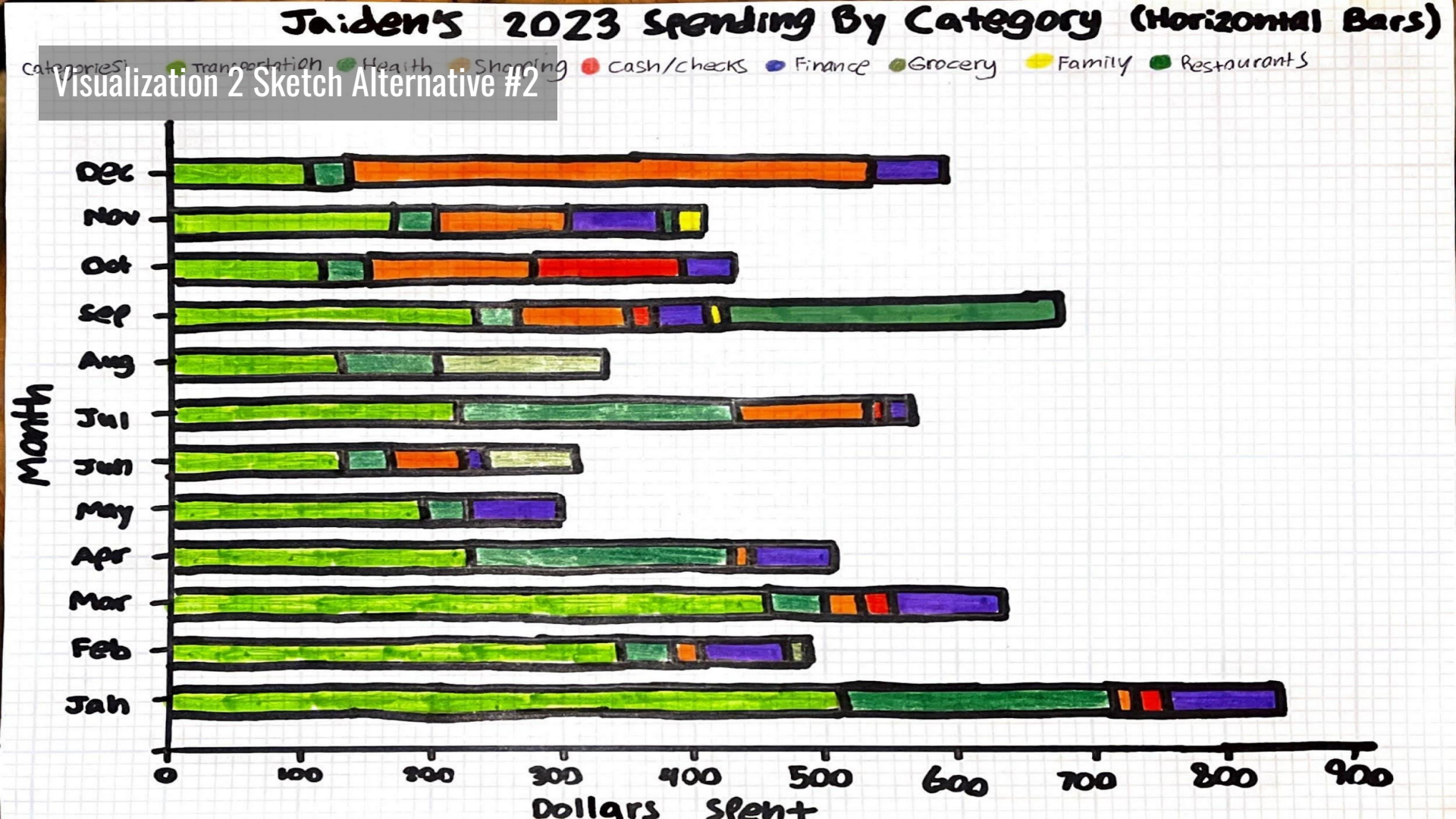
Data representation
In this pie chart, the entire chart represents total spending in the given month. This total spending amount is also put under each pie chart for easier access. Furthermore, each slice of the pie chart corresponds to a different category of spending, and its area corresponds to the percentage of total spending that the specific category contributes to. Furthermore, in each category/area, the exact number of total spending from that category is also displayed. Each category of spending is given its own unique color, which is determined and displayed in the legend below the title. In this pie chart the categories of spending that are represented are: Transportation, Health, Shopping, Cash/Checks, Finance, and Groceries.

Context of use
This visualization is useful in the situation where users want to visualize the proportions of total spending in certain categories. For example, if the user sees that in each month the majority of pie chart is light green (like it is in this case), then the user knows that they often spend the most money on that specific category in each month of the year. This chart allows users to quickly assess which categories are most frequently their biggest spenders, which in turn can inform their decisions on how to spend their money more efficiently.

Key strengths/weaknesses
Strengths: One strength of the pie chart, is that it is easy to decipher which categories are big contributors to total spending, versus which categories are smaller contributors to total spending. This allows users to visualize proportions to the total amount of spending within each month, but also across months.

Weaknesses: One weakness of the pie chart, is that smaller differences are harder for humans to perceive. For example, in the March 2023 pie chart, it is very difficult to decipher between the purple and dark green slice even though the former contributed 12% to total spending, and the latter contributed only 7%. Furthermore, without prior knowledge of the percentages (like I have access to), users still won't know the exact percentage of total spending that each category contributes, instead they only get a rough visual estimate.

Visualization 2 Sketch Alternative #2



Sketch 2: Design Rationale

Brief description
For my second redesign of the Bank of America visualization, I took inspiration from my second visualization of the population dataset, but this time I decided to visualize monthly total spending by category through the use of a horizontally stacked bar chart. In particular, each of the bars in the horizontally stacked bar chart represents the total spending in each month of the given year. Furthermore, each bar is broken up into three "sub-bars," each of which represents the total spending of the given spending category in the specific month of the given year.

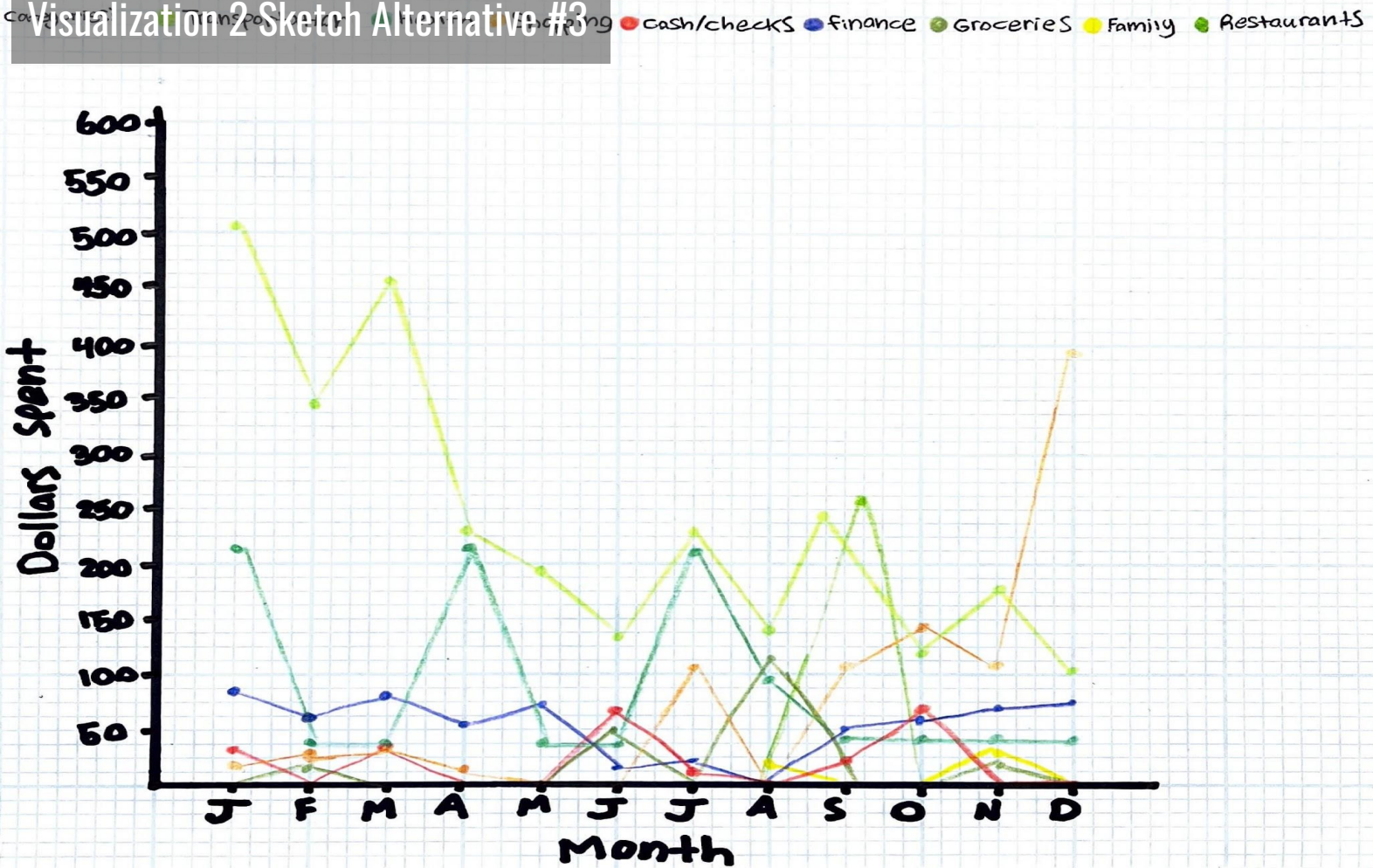
Data representation
As mentioned above, each bar in its entirety represents the total spending of all of the eight represented predetermined spending categories combined, with the value associated with each bar being represented by the length of the bar itself. The specific value that the bar represents is the corresponding number on the x-axis that is in line with the farthest point of the bar. It is also important to mention, that the y-axis in this case, as opposed to the vertically stacked case, represents the months of the year from January to December. Furthermore, the individual spending in each of the eight categories is represented as a colored segment of the entire bar of the whole. The meaning of each color is presented in the legend in the upper-right of the visualization. The total spending value associated with each segment is also represented by its length, although it is calculated differently than the entire bar. This value is calculated by subtracting the x-value corresponding to the leftmost point of the bar from the x-value corresponding to the rightmost point of the bar. The spending categories that are represented in this chart, from left-to-right, are Transportation, Health, Shopping, Cash/Checks, Finance, Grocery, Family, and Restaurants.

Context of use
As was the case for the vertically stacked bar chart, the horizontally stacked bar chart is good for seeing trends in total spending, while at the same time also gaining some insight on the trends in spending for the specific spending categories. However, it is important to note that the stacked bar chart focuses less on categorical trends, and more on categorical proportions. Thus, the horizontally stacked bar chart allows users to see proportions of total spending, like the pie chart, but unlike the pie chart, also allows the users to see how their total spending fluctuated throughout the year, especially in comparison to other months.

Key strengths/weaknesses
Strengths: One key strength of the horizontally stacked bar chart is that it easily allows users to compare total spending across months due to the ease of locating the position on the x-axis. Furthermore, by also breaking down the total spending into the different spending categories, this chart offers more insights than the pie charts on the previous slides.

Weaknesses: One weakness of this visualization, is that although it is easy to compare total spending month-to-month, due to how value is encoded into length, it is not as easy to see the exact total spending value in each month, and it is even harder to find the exact value corresponding to each spending category in each month, due to the varying lengths of other categories. Unlike a side-by-side bar chart, this horizontally stacked bar chart makes it harder to see trends across categories as small changes are hard to detect due to the differing starting position each category has in each month (other than the first category).

Visualization 2 Sketch Alternative #3



Sketch 3: Design Rationale

Brief description
For my third and final variation of the Bank of America spending chart, I decided to go back to my go-to visualization, that being a line chart to visualize my 2023 monthly total spending by category. Whatever portion of the y-axis a given point is on corresponds to the total spending for a given category in a specific month.

Data representation
The way in which the line chart is built is very similar to the line chart for the population data, but is a bit different than the line chart for avalanche danger, as for each month in a given year (in this case for 2023) and for each of Bank of America's predetermined spending categories, the corresponding total spending for that category is plotted, with the month on the x-axis, and total spending in dollars on the y-axis. After all the points are plotted, points belonging to the same spending category have lines put between them, this is done pairwise. In order to differentiate between the different spending categories, each category is given its own unique color. In particular, the spending categories represented in this chart, from left-to-right, are Transportation, Health, Shopping, Cash/Checks, Finance, Grocery, Family, and Restaurants. The color associated with each of the colors is given in the legend above the chart.

Context of use
As with most line charts, users are able to see their trends in total spending per category, more efficiently than if they were to use the stacked bar chart. Furthermore, comparisons between different groups across different months is also easier to do on a line chart due to the fact that it is simpler to find the exact spending per category, per month, as opposed to the tedious calculation that is done with the horizontal bar chart. The results from this chart can be used by a user to better understand their spending habits, which could allow them to make changes and save money in the long run. For example, this line chart efficiently shows me that I spend the most money on shopping in December, this is what we'd expect since that is the holiday season.

Key strengths/weaknesses
Strengths: One strength of this line-chart, in comparison to the previous two redesigns, is that it is easier to see trends in the data, especially within each category. As mentioned above, due to its simplicity, it is easy to compare values between the different groups.

Weaknesses: Due to the fact that there are so many different categories, there is a lot of overlapping between the different lines in the chart, this makes looking at specific categories quite difficult. This crowding happens due to the fact that a lot of the categories have very little spending in them, mainly due to the broadness of the categories, as well as my ability to save money. Lastly, due to the fact that many of the categories have a fair amount of months with zero spending, these specific lines are hard to interpret/understand the trend that is going on, especially with the overcrowding already being present.

Visualization 2 Reflection

Compare your designs. What are their relative strengths and weaknesses?

As can be seen above, the three ways I chose to re-visualize the Bank of America spending chart were by using pie charts, a horizontally stacked bar chart, and the line chart. Each of these visualizations were built to visualize total spending per month, as well as breaking down this total spending into the contributions of each of the predetermined spending categories. Although each visualization sought out to visualize this data from a slightly different lens, each highlighting a different aspect of the total spending, it followed that each one of the visualizations had similar strengths and weaknesses. For example, each visualization used color to represent the same spending categories, which would be difficult or less effective to visualize otherwise. On the contrary, each visualization was pretty basic and only covered the surface level information provided to users by the Bank of America website, this is mainly due to the fact that the information Bank of America gives is in itself pretty surface level.

What was your favorite design and why?

My favorite redesign of the Bank of America spending visualization were the three pie charts. This was my favorite redesign because it allowed me to directly change one of the critiques/weaknesses I had/found in the original visualization. This critique I had was, “given that the chart is a radial pie chart, instead of a full pie chart, it is harder to view proportions of spending in the radial pie chart than it is using the latter visualization type.” Thus, by making these pie charts, I more easily allowed users to view proportions of total spending by category in comparison to the original visualization.

What did you learn from this part of the assignment?

This part of the assignment allowed me to learn how to turn a basic radial pie chart into different visualizations. Before choosing this visualization for this part of the assignment, I didn’t think I would be able to turn such a basic and plain visualization into three unique visualization types. However, after thinking outside of the box, as well as taking inspiration from my previous sketches, I was able to come up with three informative and concrete visualizations.