SI 649 Individual Project Report
Haley Johnson
March 19, 2023

For my individual project, I created two data visualizations to accompany the article "Doctors are Failing Patients With Disabilities." When I first read the article, I immediately knew I wanted to create a visualization that showed the disparities in care quality between disabled and able-bodied patients. The Center for Disease Control maintains the Disability and Health Data System, which contains information about the prevalence of disabilities across different demographic categories, health risks and behaviors, and preventative care, and I decided to create a visualization using this dataset.

Interactive Visualization Learning Objective:

- Remember:
    - The viewer can identify that the chart uses population-adjusted rates
    - The viewer can recall when it's necessary to use populated-adjusted rates rather than the "raw" rate
- Understand:
    - The viewer understands that there are disparities in preventive care between disabled and able-bodied people, as well as disparities among different kinds of disabilities
- Analyze:
    - The viewer can analyze the different rates of being up-to-date on preventative care across different ability types

Initially, I struggled to come up with an idea for a static visualization. I wanted to visualize something that was informative and contained information about different types of disabilities (e.g. showing data for individuals with cognitive and visual disabilities separately, rather than just collapsing the data into any disability vs no disability). Lots of my early ideas were either too data dense or not informative enough without interactivity; I realized I was going to have to make decisions about what portions of the data to highlight if I wanted to include.

Ultimately, I decided to visualize how the percentage of disabled people without health insurance has changed since the passage of the Affordable Care Act in 2010. The ACA notably made it illegal for insurance providers to deny coverage on the basis of pre-existing conditions, including disabilities or chronic conditions. Since the passage of the ACA, the percentage of disabled people who are insured has decreased by nearly 50%. I was worried about implying a false correlation by linking this decrease to the ACA, so I did quite a bit of research before starting on this visualization. I found that there is significant research on how the ACA impacted patients with disabilities and that much of this decrease was attributable to policy mechanisms inside the ACA.

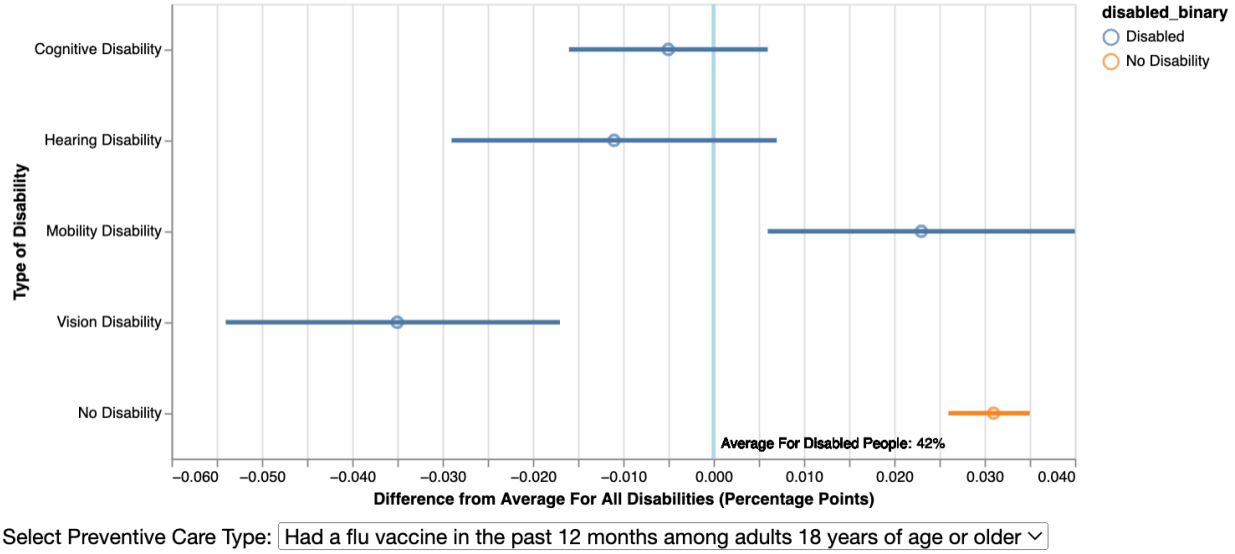
Static Visualization Learning Objectives:

- Remember:
    - The viewer can recognize that there is a general trend, where rates of insurance coverage have been increasing over the period from 2008-2019
- Understand:
    - The viewer will understand that there is a relationship between government policy and healthcare access/outcomes.
    - The viewer can compare rates of insurance coverage between years in the period from 2008-2019, as well as before and after the passage of the Affordable Care Act
- Analyze:
    - The viewer can analyze disparities in insurance coverage between different disabilities and how they've changed from 2008-2019

Interactive Visualization Design Process:

Although my final interactive visualization was fairly close to the initial sketch I submitted, designing it proved to be fairly challenging and involved some difficult decisions.

Initially, I wanted to highlight disparities across different kinds of disabilities. I wanted to center the x-axis at 0 and for each type of disability, I'd show how their rate of being up-to-date compared to the average for all disabilities. So if for a certain type of care, on average 50% of disabled people are up to date, but only 45% of people with hearing values are, the value would be -5%.



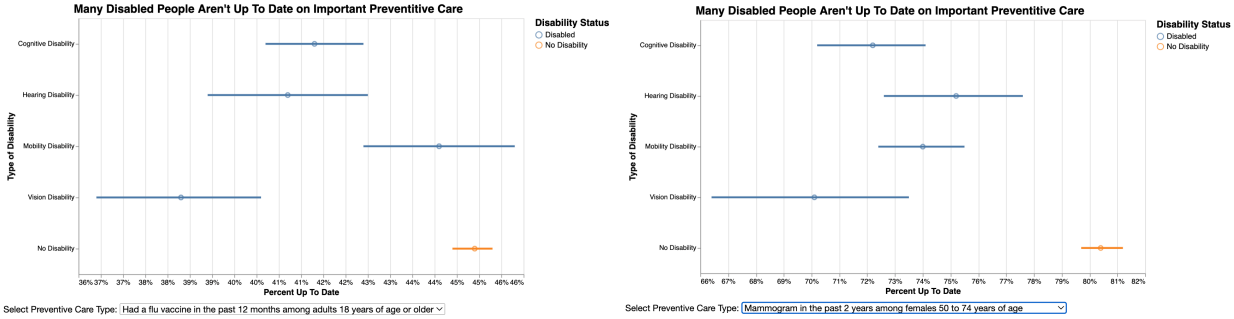
Previous version of my interactive visualization

While I liked the idea of doing something different than a standard line plot, I felt like this design wasn't really showing the underlying data. Some types of preventative care have much higher

up-to-date rates for both disabled and able-bodied people. For example, only around 40% of people get their flu vaccines but 70-80% of people are up to date on cervical cancer screenings. This is completely lost when I just show the data as deviations from the average.

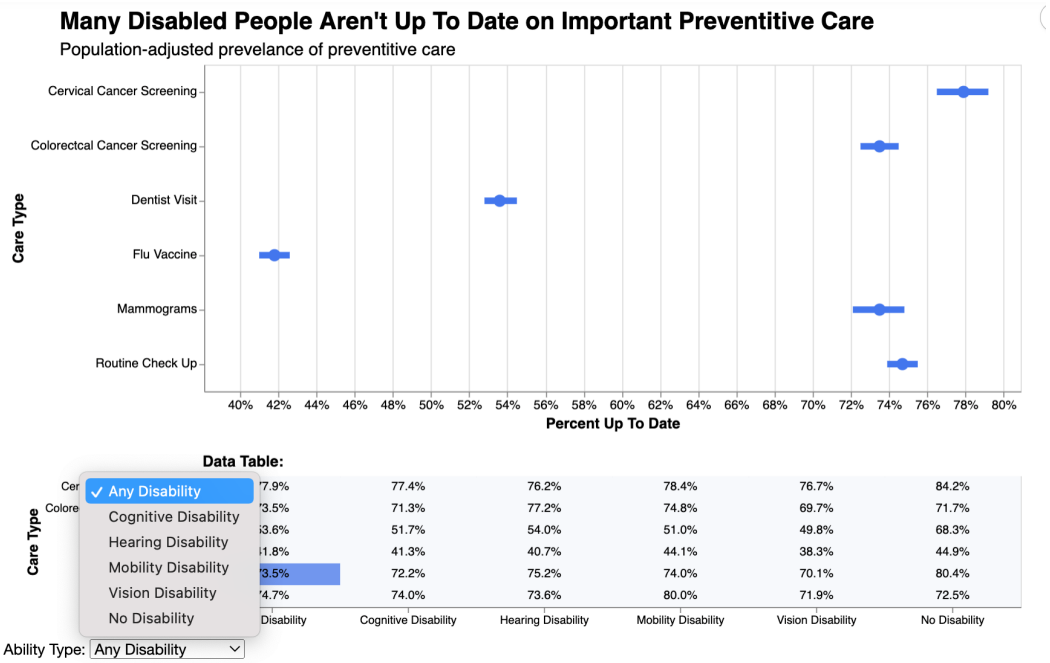
Likewise, I'd argue that missing a cancer screening is much more serious than missing the flu shot one year. To me, this is reflected in the fact that being up-to-date on cancer screenings has a much higher base rate in the population. Showing a 1% deviation from the mean as the same for all preventative care types, irrespective of the base rate, felt a bit reductive. I ultimately decided to abandon the idea of visualizing deviations from the average and just show the raw data instead.

This was when I began to run into issues with scales. For each "view" on the dropdown menu, Altair automatically generates a new scale for the x-axis:



Note the different scales on the x-axis for the different types of preventative care

I also ran into some technical constraints using Altair. I wanted to move the location of the dropdown menu to the top of the page, or between the chart and the table, to make it more clear what information was being shown.

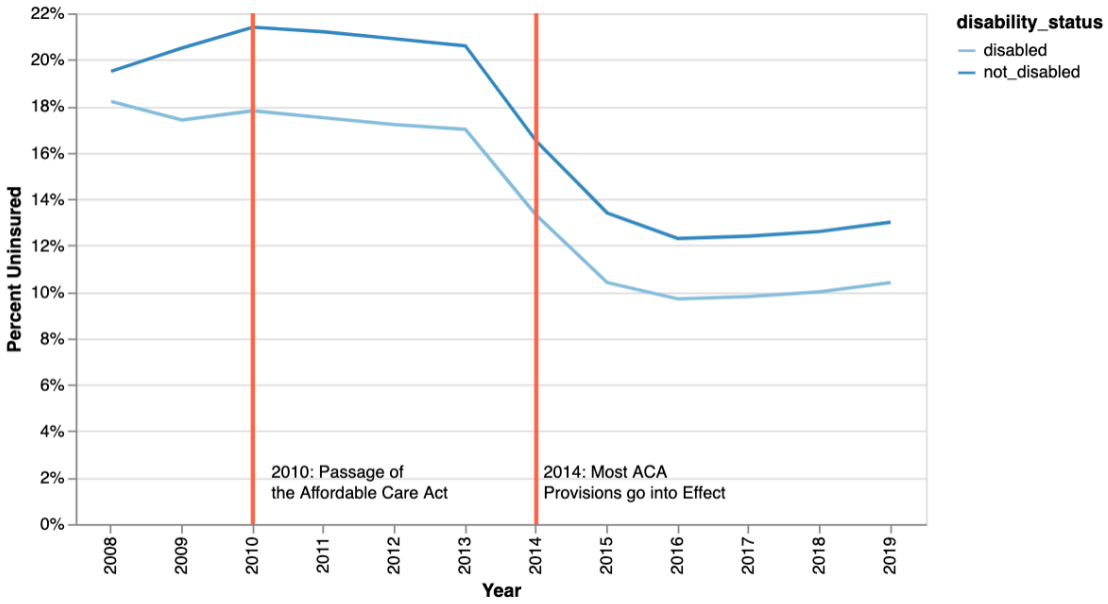


In Altair, the default position for widgets is in the bottom left-hand corner

Unfortunately, this is not possible in Altair. There are some workarounds using HTML if you're displaying a chart in Jupyter Notebook, but they didn't seem to work in Streamlit (or rather, I don't have the right knowledge of HTML to adapt those solutions to Streamlit). I set the default view to "any disability," so hopefully the title makes it clear what the default view is showing. However, I think putting the dropdown box in a more discoverable location would have improved the effectiveness of visualization.

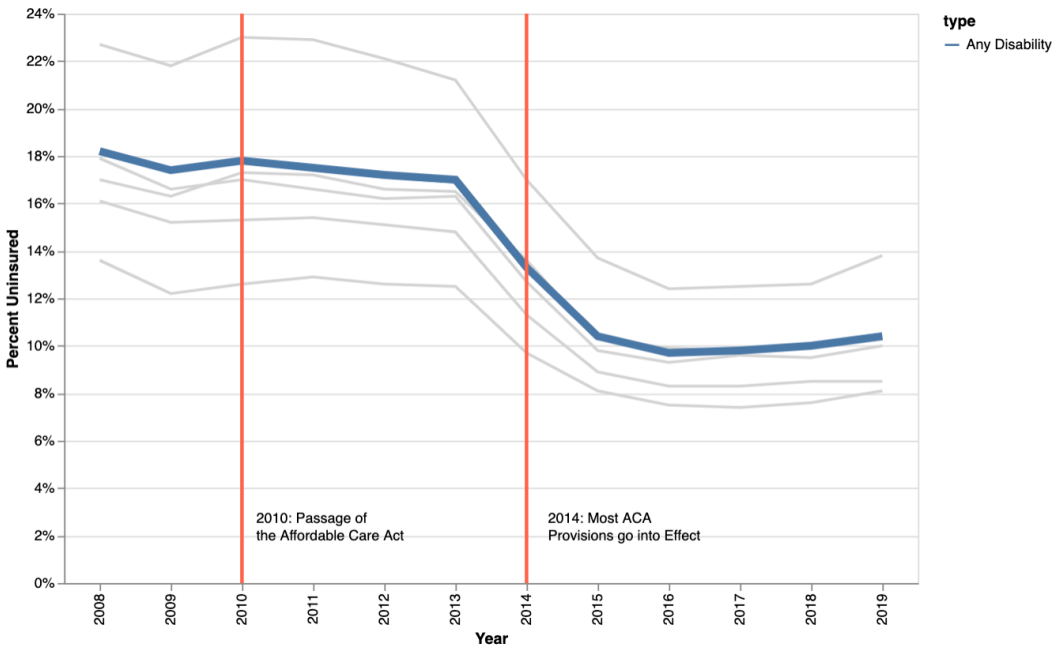
Static Visualization Design Process:

In an earlier version of my static visualization, I was just showing rates of insurance coverage for disabled vs able-bodied Americans. I liked that this highlighted that disabled Americans are less likely to be insured and that this gap has persisted even after the passage of the ACA. However, I felt like this visualization wasn't "data-dense" enough and decided to explore other approaches.



Previous version of my static visualization

Next, I tried adding lines for different types of disabilities. My main design challenge was balancing between highlighting the overall trend (the percentage of disabled people who are uninsured has fallen overall) vs disparities between different types of disabilities (e.g. individuals with vision disabilities continue to be less likely to have health insurance).



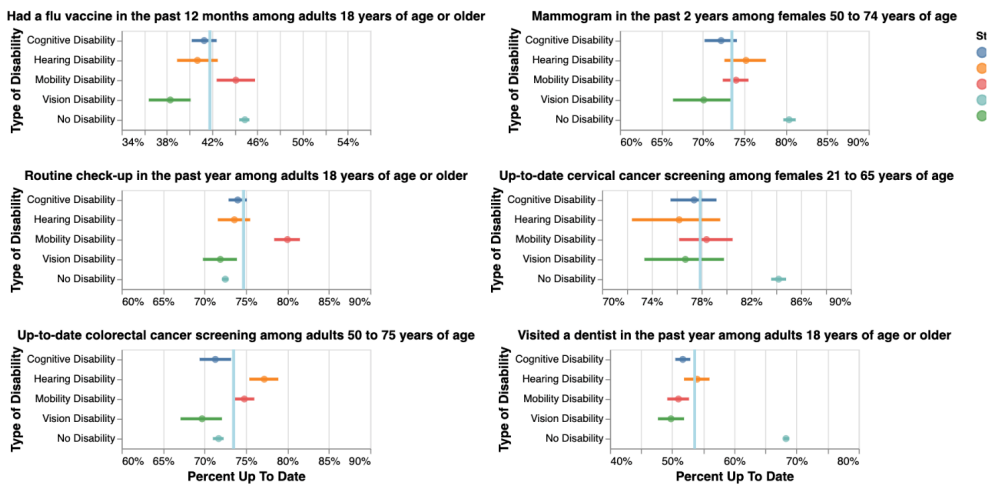
Previous version of my static visualization

I played around with making each line a different color to distinguish between different kinds of disabilities but felt like this took the emphasis off the overall trend, which was the main piece of information I wanted to highlight. After playing around with a few kinds of encodings, I decided using a combination of shape and color was the most effective approach. There's some overlap between the lines, which makes certain data points hard to see. For instance, the line for people with hearing disabilities is mostly covered by the line "any disability" between 2013-2019. My intention was more so to highlight the overall trend than to make it easy to pick out the insurance coverage rate in any given year for any disability, so I was okay with this tradeoff.

The last step in creating my static visualization was to create the "information box" below the line chart. I wanted this to be a "sidebar" visualization and thought it was necessary to provide additional context on the relationship between the ACA, insurance coverage, and disability since this wasn't discussed in the original article. I considered using Adobe Photoshop to add a text box but discovered that this was possible in Altair using a combination of mark_rect() and mark_text().

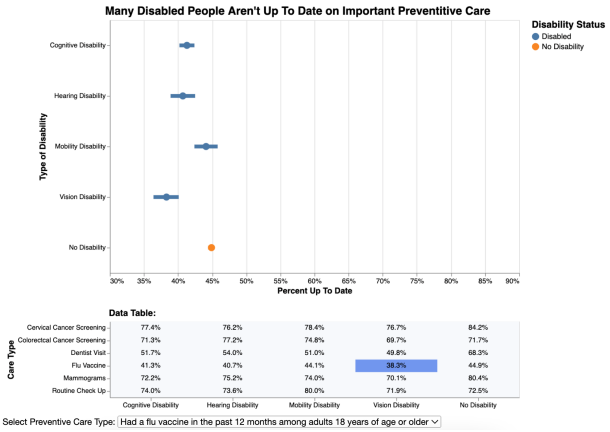
Qualitative Evaluation:

My interactive visualization proved to be a "wicked problem," like we discussed in lecture 8. I made around 4 versions of this visualization that I think I could've submitted for this project but really struggled to decide which one was the right one or even what criteria I should use to determine if one visualization was better than another. In particular, I had trouble deciding if I should use small multiples or a drop-down menu. I liked that small multiples enabled more direct comparisons between subsets of the data, but decided that it was ultimately too much information to show at once.



Alternate version of my interactive chart using small multiples

Once I decided to use a dropdown menu, I had to decide if ability type or preventative care type should be the variable in the drop-down menu.



I never really resolved this question — I created polished versions of both options on Streamlit before deciding which one to submit. I think the relative effectiveness of having care types in the dropdown menu vs ability types depends on the type of comparisons I wanted to enable. The final version I created has the ability type variable in the drop-down menu and preventative care type on the y-axis. I thought this arrangement had a more logical default view — when the page loads, the viewer sees the average percentage up-to-date among all disabled people. Having a default view that was easy to interpret was especially important to me because of the technical constraints I faced trying to move the dropdown menu in Altair.

I initially disliked Tufte and his focus on minimalism when we discussed him in class, but I found his principles helpful while working on my visualizations. His principle that designers should “above all else show the data” was very useful for me when I was creating my interactive visualization and decided to go from showing deviations from the average to raw values. Likewise, his idea of data-to-ink ratios helped me articulate why I was dissatisfied with earlier versions of my static visualization — I wasn’t encoding enough data relative to the rest of the “ink” on my chart.

My personal style isn’t quite as minimalist as Tufte’s, but this experience has made me appreciate his principles a lot more. Striking the right balance between aesthetic appeal and data density can be difficult sometimes; keeping Tufte’s ideas in mind throughout the design process helped me prioritize clarity.

Overall, I think my static visualization is very effective. It would complement the article by providing a case study of how policy tools can be used to improve healthcare outcomes for disabled people. The article discusses a number of tools to address disparities — such as requiring hospitals to create plans to address access issues or helping medical offices understand what ADA compliance entails. The static visualization is meant to suggest that government mandates can be very effective at addressing medical discrimination.

My interactive visualization, on the other hand, was designed to be exploratory rather than persuasive. I think it adds less to the article than my static visualization because it already extensively discusses how disabled people are less likely to receive preventative care. As I discussed above, I’m also not sure what the best way to show this information is.

As a designer, I think I sometimes reach a point where I've worked on something for so long it's difficult for me to tell if it's "good" anymore. I definitely hit that point with my interactive visualization and don't feel completely satisfied with the finished product. I "tested" the visualization on my roommates, who are both political science majors, and they were able to complete some basic tasks using the visualization, so at least I'm confident that it's interpretable and enabled basic exploratory analysis.