COVID-19 Vaccination Tracking Website

Chessa Hoekstra*

MIT 6.859, Spring 2021

Figure 1: Viewing Daily Vaccinations per Million People for Alaska. Additional information after hovering over the line gives the count on April 13th.

Abstract

COVID data has become abundant since the pandemic moved to the forefront of our society over a year ago. Along with abundant data comes a need to visualize data, make sense of it, and communicate it to the public. The proposed COVID Vaccination Tracking website provides general users with a single place to view and interact with vaccination data. Specifically, users can select different metrics, focus on certain states, and see more information via tooltips.

Keywords: data visualization, interactive map, coronavirus, vaccination, COVID, hoverable map, line chart

1 Introduction

The spread of coronavirus (or COVID) over the past year has dominated news cycles both nationally and globally. Inherently statistical, case counts posed challenges on how to best make sense of and visualize large datasets. Most tend to agree that data (and corresponding insights) need to be accessible to the general public, but this is no easy task. Accordingly, people have set out to make easily understood source websites and visualizations.

Since the start of 2021, a new (more optimistic) set of numbers has garnered similar need for visualization: vaccination rates. I propose a tool that is straightforward in that it gathers data directly from the CDC, simple in that it avoids advanced metrics, and efficient in that it is delivered all on one web page.

*e-mail: chessa@mit.edu

2 Related Work

Since March of 2020, web pages, journalists, visualization artists have striven to best communicate information about COVID to the public.

The CDC [1] raw data has become the largely-agreed upon gold standard of COVID data, as they are the government-funded party that collects and aggregates it.

If you Google search "covid cases" [2], you will see that Google, in collaboration with the New York Times and other sources, took a straightforward approach, sticking to simple charts (i.e., using only numbers the CDC gives).

Others, such as Coleman for the New York Times [2], opted for more journalistic designs in order to create the popular *memorability* [1] effect for their readers.

In hopes of find insights that are unable to be seen from raw data, Dorling and McLure [3] compute additional statistics to propose more meaning in the data.

The COVID Vaccination Tracker aims to provide one consolidated source that is simple and interactive, able to engage the average user.

[1] cdc.gov
[2] google.com
[3] https://images.theconversation.com/files/326025/original/file-20200407-31007-dits9d.png?ixlib=rb-1.1.0q=45auto=formatw=1000fit=clip

3 METHODS

3.1 Visual Encodings

This website attempts to take advantage of how the public is used to seeing and selecting data. For example, during the pandemic, users have become accustomed to seeing line graphs of cases over time, so line graphs were essential visual encodings to include for vaccinations as well. Additionally, although population data (and therefore case data) does not inherently correspond to state areas [4], selecting states from a map has become an intuitive process in our society. The color encodings for state lines have no inherent value (which is why they are categorical and not shades of the same color).

3.2 Interaction Techniques

I took straightforward approaches on my interactions: radio buttons for single selections, hovering over the line for further tooltip information, and map hovering. Hovering over the map and being able to zoom in for more hover accuracy are both very intuitive actions on a map nowadays. Luckily, Mapbox [5] is a well-documented API that has a natural feel and it was great to get some experience using it. This feature also proves to be very engaging to the user. For all of my interaction techniques, I wanted them to make sense to the user, and to reduce the chance of confusion, I wrote detailed instructions at the top of the website for how to use the interactions.

3.3 Design

The design is meant to be stripped and matter-of-factly. In the peer review session, I had intended for a more playful design (with a green/teal background, fun icons), but redesigned it after I realized I wanted a more professional scope. After all, the CDC (or any other trusted source of health data) would not have unnecessary design unless it was made for kids (a direction I didn't go in).

3.4 Data

I utilized a GitHub repo [6] which loads COVID data daily from the CDC. This ended up being a lifesaver because the CDC only seems to have distinct csv files that you have to download individually.

4 RESULTS

Figures 2 and 3 following are examples of visualizations this system can produce. A user can visit the website before traveling to a new state, gathering info to make informed decision on their activities there. They may also just want to see how their home state compares to other states or other regions in vaccination rates. Tooltip information on proportional values (e.g., 40 out of 100 people are vaccinated in Massachusetts as of this week) is very helpful contextually, whereas a user may not have as much context for raw values (e.g., Massachusetts administered 2,370 doses on April 17th).

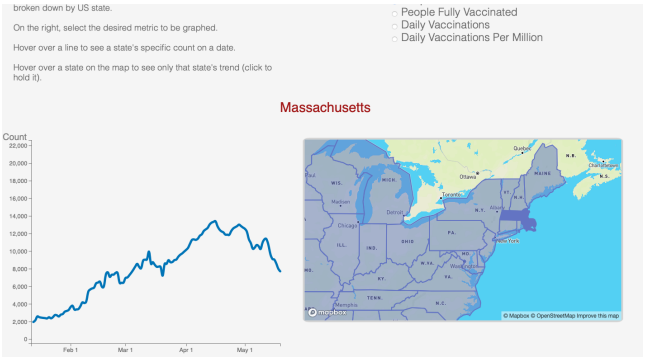

Figure 2: Viewing Daily Vaccinations for Massachusetts

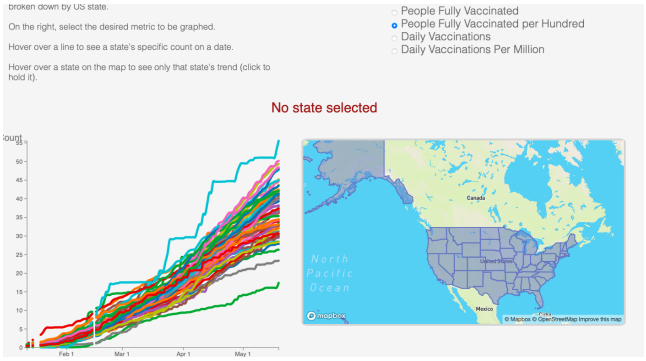

Figure 3: Viewing Total Vaccinated People Per Hundred with no state selected.

5 DISCUSSION

Users are able to learn valuable comparison insights from using this visualization. Additionally, the charts and interactions are intuitive and it is easy (only a few seconds) to pick up on them. Examples of insights that the website has revealed (that would otherwise take a lot of time to compile from raw data) are that vaccination rates are lower in the southeast versus the northeast of the US, some of the the highest daily vaccination peaks have come from US islands such as Hawaii, Guam, and Palau. The interactive Mapbox map implemented on the site has proven very popular among users, especially for discovery of islands or easy investigation of states they wouldn't have otherwise looked at. Though exploratory, this type of interaction still has value in engaging the user.

6 FUTURE WORK

This tool has further potential in adding in more demographics. For example, while the CDC has excellent data based on location and excellent demographics on a national basis, it does not have demographics broken down by location. Future work would include building additional pipelines from

[4] see mapping and cartography lecture http://vis.csail.mit.edu/classes/6.859/lectures/15-Maps.pdf

[5] mapbox.com

[6] https://github.com/owid/covid-19-data/tree/master/public/data/vaccinations

state-level CDCs in order to get this data, though these systems are not all consistent so it would need hard-coding. I am using raw GitHub data [source], but it is still manual retrieval, so in the future maybe CDC has its own direct pipeline as well.

Additionally, if I had more time in the future, it might be nice to turn this set of visualizations into some sort of story, seen through one aspect of COVID; I attempted using a scrollyteller for something of this nature at first, but it resembled a journalistic biased article more than I liked.

REFERENCES

[1] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013. doi: 10.1109/tvcg.2013.234

[2] N. Coleman. On the front page, a wall of grief, Feb 2021.