

Kollin Trujillo

COVID-19 Time Series Viz

CS545

Visualization of COVID-19 Time Series Data

d3.js is a Javascript framework used to manipulate SVG images. SVG is a scalable vector graphic that is known for its ability to be interactive and animated. d3.js is used to help simplify the creation of SVG elements and has a multitude of helper functions used to simplify the construction of information rich graphical displays. I want to utilize d3.js to generate a choropleth map to be able to display all the states data and a line graph to be able to focus on trends in specific state data.

Data

I utilized the COVID-19 dataset from the CDC.gov website

(<https://data.cdc.gov/NCHS/Weekly-Provisional-Counts-of-Deaths-by-State-and-S/muzy-jte6>).

This dataset consists of 35 columns which I pruned down to 6: 'Jurisdiction of Occurrence', 'Week Ending Date', 'All Cause', 'Natural Cause', 'COVID-19 (U071, Multiple Cause of Death)', and 'COVID-19 (U071, Underlying Cause of Death)'. Of these 6, I decided to focus on two columns in particular, 'Jurisdiction of Occurrence' (referred to hereon out as 'state') and 'COVID-19 (U071, Underlying Cause of Death)' (referred to hereon out as 'value').

To carry out the visualization of this dataset, I did it in Python. To process the data, I coded up a python script to handle the mundane. From this, I converted the 'Week Ending Date' column using a pandas function to convert the column into a datetime format. I set the

indexing variable to 'Week Ending Date' this allowed me to group the columns by month and year and sum them to get monthly aggregates. From there I reset the index and renamed the columns 'Jurisdiction of Occurrence' to 'state' and 'COVID-19 (U071, Underlying Cause of Death)' to 'value'. Then I constructed a dataframe of only state and value and utilized these in d3.js to construct the visualizations. The only difference in this for the line graphs was the use of aggregating also by state to get data on a per-state basis.

Visualization of Data

To visualize the data, I decided to utilize a Choropleth map and a line graph. I wanted to look at this visualization as a mentality of a "parts of a whole" type thing. We have the parts which consists of a much broader area and consists of the Choropleth map where we can see all the individual state data at once for a given month and year. The whole part of the visualization is the line graph wherein we essentially "zoom in" on a particular state. What I mean by this is that we focus on a particular state and look at all its data, also at a quick glance, but not just in a month isolation like the Choropleth map. From here we can see the whole picture of the data for that state.

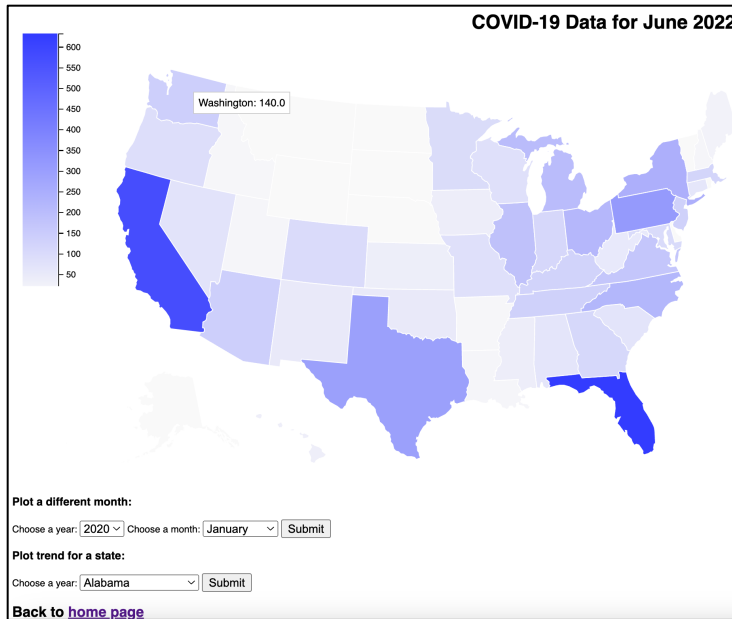


Figure 1: Choropleth visualization of COVID-19 Data for a selected month and year.

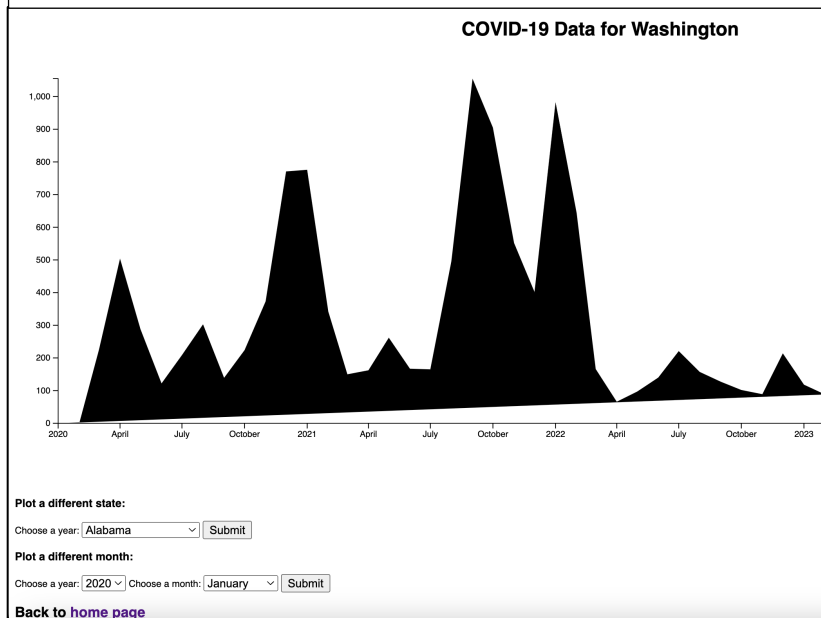


Figure 2: Line graph visualization of COVID-19 Data for a selected state

The two figures above showcase the efficiency of the visualization strategies selected. We can see the parts and the whole of the visualization of the data of the states. The use of the hover effect of the toolbox div tag implemented cannot be understated in its effect for interactivity with the visualization. I believe that it helps to grab a tangible understanding of particular states' data especially when the linearity of the color scale might be difficult to pinpoint exact values.

I had some issues with graphic generation. I am not sure or if it is my code but despite this graphical problem with the construction of the graph, which I am not fully sure if it is a Javascript issue or a SVG limitation, the peaks of the lines were still accurate. Figure 2 highlights this where there is some drift in the baseline as time goes on.

Outcomes and Conclusion

Overall, the outcome was successful. We are able to successfully visualize the data in an easy-to-understand and human-readable format. Along with successful choices of visualization elements, it aids the human ability to comprehend larger amounts of information. I do wish I had more time to optimize some of the UI design but it is usable and gets the information across.

Future Directions

I would really love to expand on this project and integrate other aspects of public health data. I would also seek to, in future versions, utilize the API to visualize more up-to-date current data but for this project it made much more sense to download and digest the dataset locally. I wanted to include animation of time series data and onClick plot the line graph visualization for the state clicked.

Future project design directions would include the utilization of more complex language features and packages such as the use of the Javascript superset Typescript and React to address with complexity of code as the codebase expands and to minimize errors with types that Javascript is known for with larger and more complex codebases.

References

1. <https://data.cdc.gov/NCHS/Weekly-Provisional-Counts-of-Deaths-by-State-and-S/muzy-ite6>
2. https://d3-graph-gallery.com/graph/choropleth_hover_effect.html
3. https://d3-graph-gallery.com/graph/line_basic.html