



A SKETCH OF A PANDEMIC

PROJECT REPORT

Visualizing The Cases & Mortality of COVID-19 in USA in a Global Context

Raiyan Reza

supervised by
Professor Xianbin Gu

Preface

This author of this report was an undergraduate Computer Science major at New York University Shanghai at the time of conducting this work as part of the course project for Information Visualization taught by Professor Xianbin Gu at NYU Shanghai.

This project can be an excellent illustration of D3 library's capacity in visualizing data and presenting information at an elementary and introductory level. The author of this work may revisit and expand upon this project in the indefinite future.

Acknowledgements

It goes without saying that Professor Xianbin Gu's tutelage and guidance has been indispensable in the version of the work uploaded here on Github. Thus, Professor Gu has my heartfelt and deepest gratitude for the support he provided throughout the duration of this project.

—Raiyan Reza.

Abstract

D3 is a powerful JavaScript library used for creating dynamic and interactive data and information visualization.

This project, A Sketch of A Pandemic, uses the tools D3 offers to illustrate how certain facts of the COVID-19 pandemic in the United States of America and put these facts in a global context.

The information the visualization communicates are the total COVID cases and mortality a, total COVID cases and mortality per a unit of the population, and time evolution of the pandemic cases at the resolution of state-by-state granularity in the United States of America.

Keywords

D3 Library, Action-Target Pair, Tree Map ,Choropleth Map, Line Chart

Contents

1	Introduction	5
2	Data	6
3	Goals and Tasks	6
4	Visualization	7

1 Introduction

A *Sketch of Pandemic* aims to portray the on-going COVID-19 pandemic in two-fold manner: a broad global perspective, and a USA overview. At a global level, the information will manifest in the form of two interconnected views as shown in the picture below.

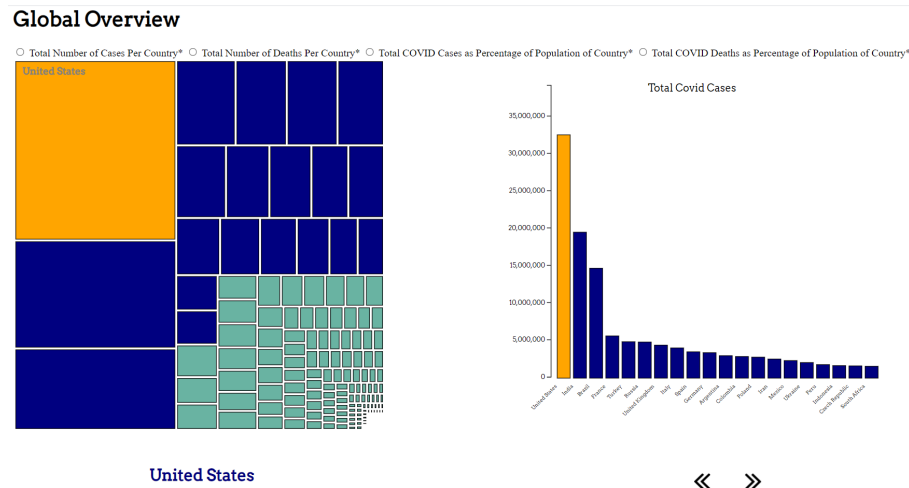


Figure 1: Tree-Map Linked With A Bar Chart

The tree map quickly conveys a bird's eye view of the pandemic and relates the part to a whole. Whereas the bar chart provides more exacting detail on a country-by-country basis. This will permit the audience to situate the USA, which would be the focus of the visualization for the next part, in a global context.

Moving forward, the USA section looks as follows. The next figure below demonstrates two other interlinked views: choropleth map and two separate line charts.

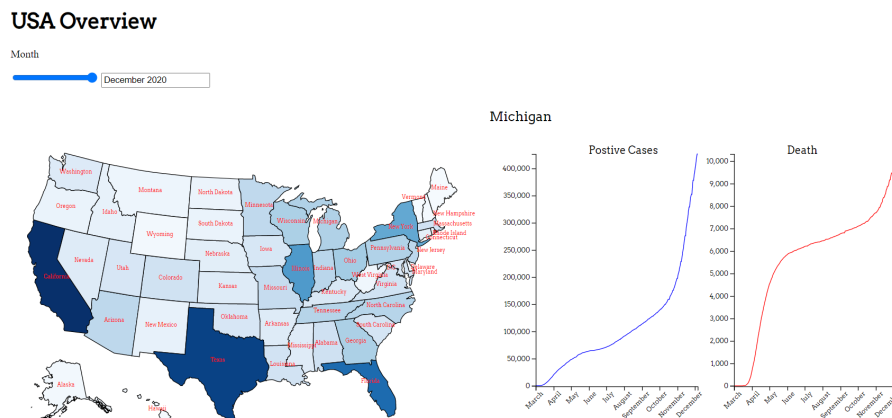


Figure 2: Tree-Map Linked With A Bar Chart

The interactions coded in the visualization will showcase how the pandemic situation evolved

with time. As with the first part, the map will broadcast a bigger narrative across the entire United States of America, whereas the line chart will let the audience survey the COVID-19 situation on a state-to-state basis.

2 Data

The data used for this visualization was drawn primarily from two different sources, both available to the public. The first data set is drawn from the regularly updated Wikipedia csv template [1]. The data presented there is sourced from the works of Ritchie et al. and available on Our World in Data [2]. The data set includes the total number of COVID-19 cases and deaths on country-by-country basis. Something worth keeping in mind is that the data set is very liberal with the criterion of what constitutes a country. Some of the regions listed, therefore, will not qualify as a country by more stringent standards. As the main focus of the visualization in this project is to simply furnish an overview of the COVID-19 situation, no data entry was omitted on the grounds of ambiguity and political reservation with the designation of certain regions as a nation state. The only exceptions were regions that unequivocally failed to meet the status of a nation, such as cruise ships, navy carriers, and such. Those entries were removed. Likewise, entries that had “No Data” rows too were removed. Upon cleaning up, the 240 entries were reduced to 228 entries. An additional column of data was added to include the population of each country. This data was manually added from the Wikipedia entry of each location. Using Excel’s use of formulas this column was used for scaling the total COVID cases and deaths to the population of a country.

For the second part, a data set on Kaggle [3] that summarized the State-by-State daily (inclusive of District-of-Columbia) COVID-situation from 22nd, January 2020 to 6th December, 2020. This produced 15633 entries. The primary processing of the data required was separating State-by-State information by month. For this simply filtering simply using JavaScript sufficed.

The GeoJSON for rendering the map of the USA was taken from a public Github repository [4] .

3 Goals and Tasks

The target audience is the lay person curious about the ongoing pandemic. The task is to show a whole picture and relate the parts accurately to the whole. Therefore, the design idiom the author used to put forward this visualization are of tree map and bar chart for showcasing the

global state of the pandemic. And, for the second part, choropleth map and line charts.

In the first section, the audience would discover distribution (action-target pair). The tree map would showcase how the COVID-19 cases and deaths compare from one country to another. The larger a tile associated with a country appears the more COVID cases or mortality we can associate to it. Since the tiles are placed next to each other in a rectangular layout, the areal proportion a tile takes up in the grid will be indicative of how much more or less COVID has impacted a nation with respect to another nation. The accompanying bar placed therein to address the shortcoming of a tree map, that is, inaccuracy inferring precise information from area. When the need for precision arises, a bar chart may be used to impart the required information. The same information would be presented in four different modes, thereby, letting the user appreciate the how drastically measuring COVID on absolute scale of cases and deaths and cases and deaths scaled to a country's population changes the meaning of the impact of COVID on a nation. In this work, the COVID cases and mortality of US appears less extreme in comparison to other countries when we take into account for the population of the country.

In the second section, the audience would compare trends. The combination of choropleth peth map and trend would assist the user with seeing how COVID-19 spread in USA, and hopefully the picture can be a springboard into further investigating USA's COVID-situation. An astute observer would note that the most impacted regions, unsurprising, were also some of the most populated states.

4 Visualization

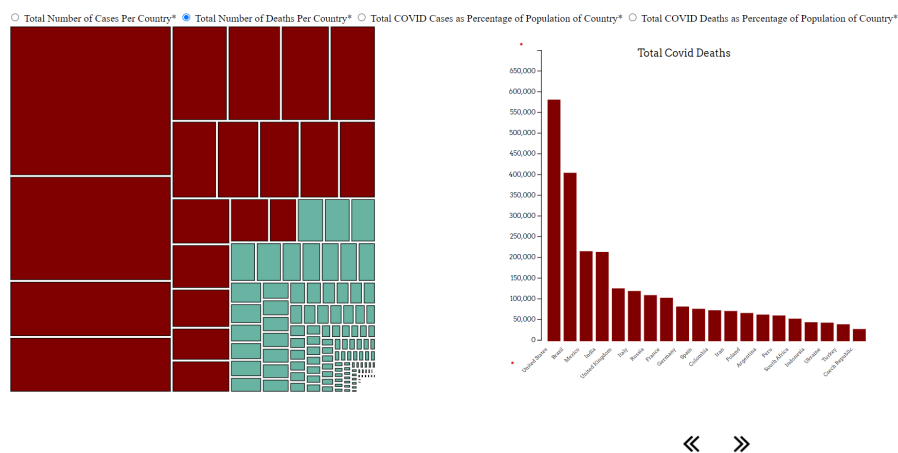


Figure 3: View Generated When Selecting The Total Number of Deaths Per Country

Figure 3 shows the view when the audience selects the “Total Number of Deaths Per Country*”. Figure 4 depicts the radio buttons up close.

☒ Total Number of Cases Per Country* ☐ Total Number of Deaths Per Country* ☐ Total COVID Cases as Percentage of Population of Country* ☐ Total COVID Deaths as Percentage of Population of Country*

Figure 4: Close Up Of The Radio Buttons

The bar chart on the right of the tree map are clearly color coordinated. The tiles that are visualized by the bar chart share the same color on the tree map, which in this case is maroon.

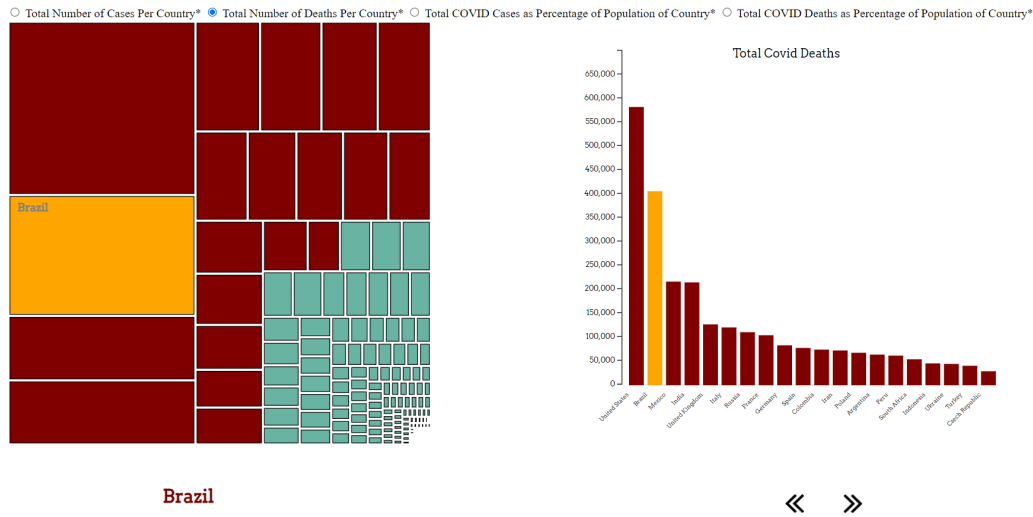


Figure 5: Interaction Between The Tree Map And Bar Chart

Figure 5 shows the interaction that occurs when a user hovers the mouse over a tile. If the bar chart is displaying the country, both the bar representing that country on the bar chart and the title representing the country on the tree map gets highlighted. The color channel therefore encodes the category here: country. A label appears both on the tile and underneath the tree map to convey this information. The area of the tile encodes here the total number of deaths per country. In different viewing modes the map encodes the relevant category of information. The height of each bar represents the same information but it is more precise so that it prevents information loss that occurs in use of area channels. At the same time the tree map lets the user see how the country compares with other countries with regards to COVID-19 statistics.

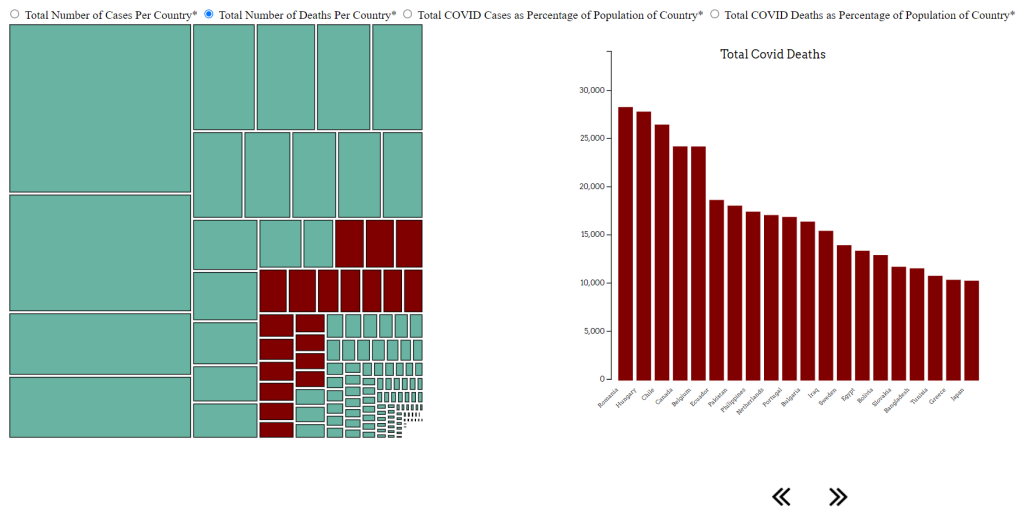


Figure 6: Scrolling Along The Bar Chart

Figure 6 shows scrolling. The buttons shown are fully functional. And, as clearly illustrated, scrolling along the bar chart selects the relevant tiles on the tree map. Both the tree map and bar chart are sorted in descending order, so scrolling right means going from higher COVID cases/mortalities to lower.

USA Overview

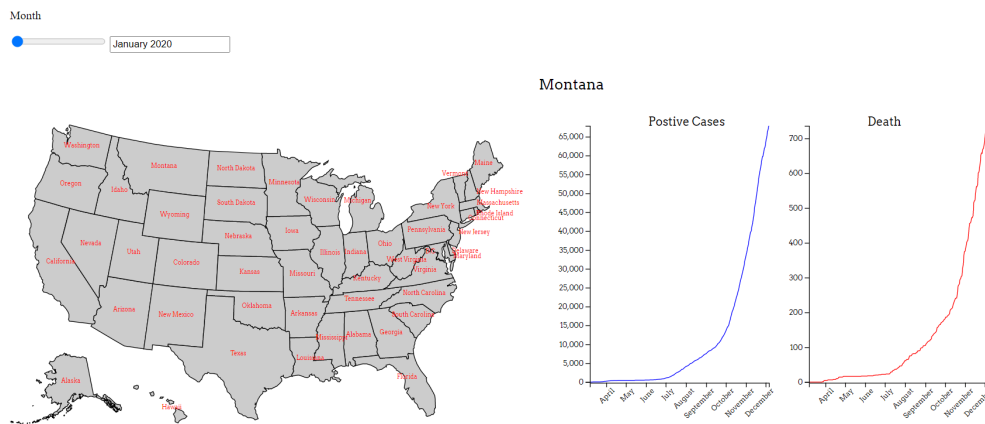


Figure 7: Choropleth Map And Line Chart

Figure 7 shows the other two views generated. A choropleth map and line chart. Using the slider, the map shows information about COVID-19 outbreak during the month of January. The gray signifies there were no positive cases in any of the states.

Figure 8 shows what happens when a user utilizes the slider. The image on the left is from June 2020. And, the shades of blue represent the COVID-19 outbreak, the dark the blue the more cases there are. And, the image on the right shows the COVID-19 situation in December. Clearly, with the passage of time the outbreak grew worse. Another thing the picture quickly establishes

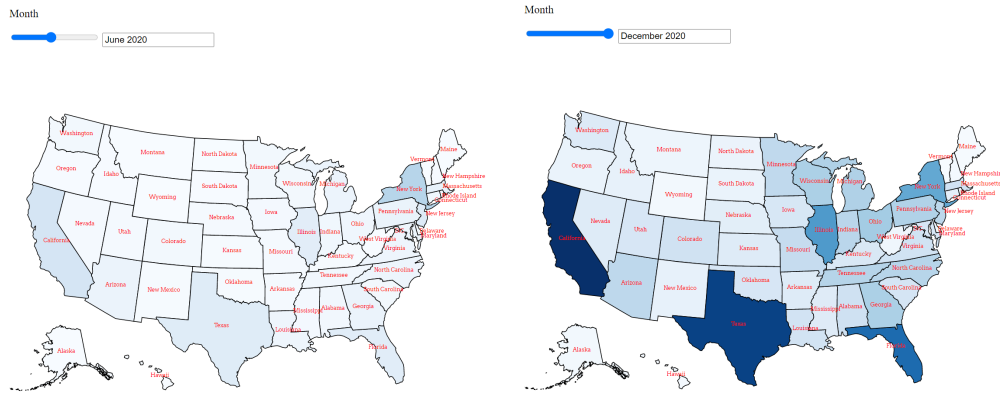


Figure 8: Choropleth Map of USA in June vs December of 2020

is there's considerable state-by-state variation. Heavily populated regions like California, Texas, Florida, and New York are much darker shades of blue, whereas Alaska takes on a very light hue.

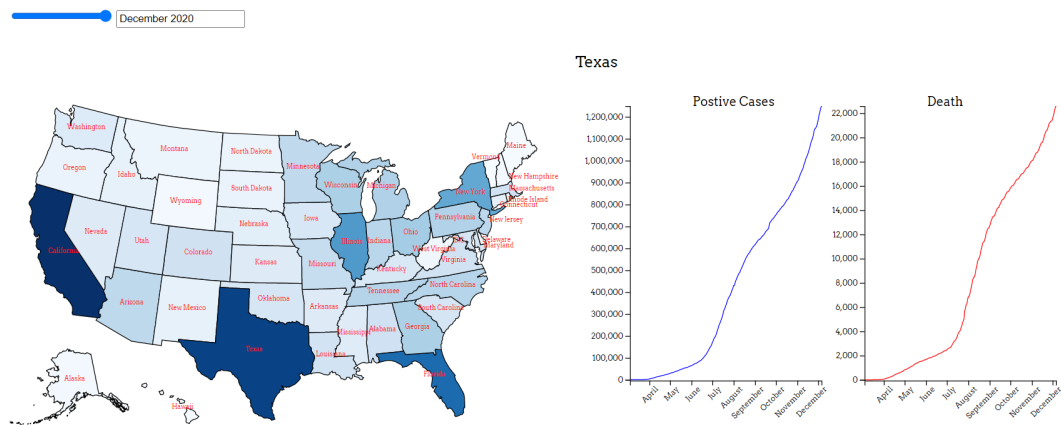


Figure 9: Choropleth Map And Line Chart Working In Coordination

If a user hovers the mouse over a State, they can see a summary of the state on the right hand side. The two line-chart shows the number of COVID-19 cases and deaths over the months respectively. Doing this, the user would note that the shape of the line is indeed different from one state to another.

This maybe indicative of factors influencing the extent and spread of COVID-19 in each state, such as but not limited to, COVID-19 measures, population density, and availability of hospital beds in a given state.

References

- [1] NA, “Covid-19 pandemic data,” https://en.wikipedia.org/wiki/Template:COVID-19_pandemic_data, 2020.
- [2] E. Mathieu, H. Ritchie, L. Rod  s-Guirao, C. Appel, C. Giattino, J. Hasell, B. Macdonald, S. Dattani, D. Beltekian, E. Ortiz-Ospina, and M. Roser, “Coronavirus pandemic (covid-19),” *Our World in Data*, 2020, <https://ourworldindata.org/coronavirus>.
- [3] S. Rajkumar, “Covid-19 in usa,” <https://www.kaggle.com/datasets/sudalairajkumar/covid19-in-usa>, 2020.
- [4] M. Chandra, “Basic us state map,” <https://gist.github.com/michellechandra/0b2ce4923dc9b5809922>, 2016.