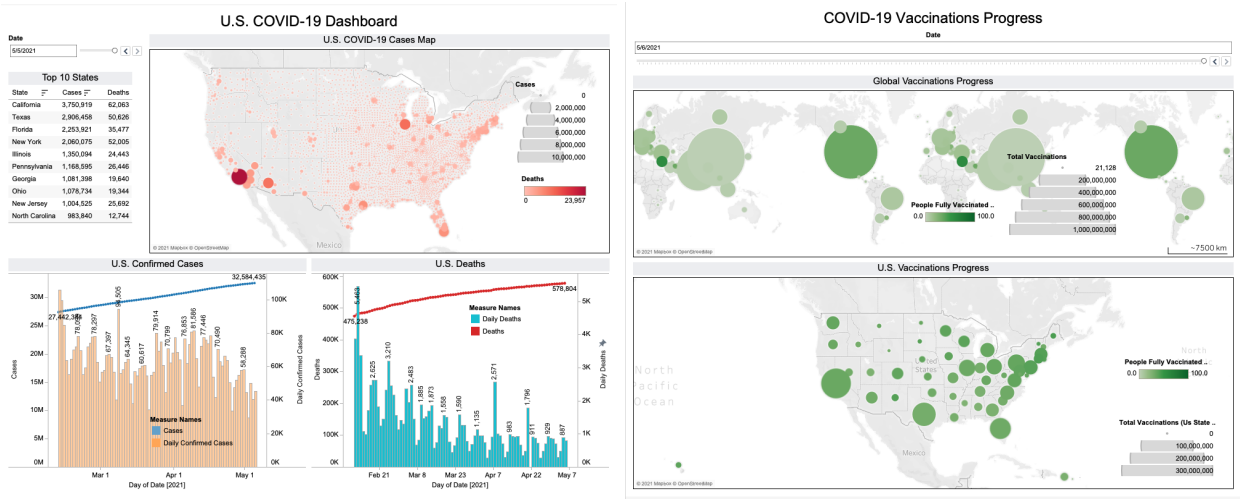
Noboru Hayashi

EN 605.662 Data Visualization

Final Project

Visualization of COVID-19 Spreads and Vaccinations Progress



Left: U.S. COVID-19 Dashboard
Right: COVID-19 Vaccinations Progress

I. Abstract

Currently, we are flooded with all types of COVID-19 related data and information, it is crucial to utilize the visualization softwares and effectively present the massive information to the audience. In the paper, the way to visualize the spreads of COVID-19 and vaccinations with Tableau is demonstrated by utilizing the publicly available datasets from Kaggle.com and integrating the skills and techniques that Data Visualization course by Johns Hopkins University offers.

II. Introduction & Background

COVID-19 has been impacting our society, economy, and daily life a lot since its emergence in late 2019. And today, we are flooded with all sorts of the pandemic related information: the geological spreads, vaccinations over time progress, and some other info such as news reports, tweets and so on. It is a silver lining for us that we have a lot of data to experiment with to demonstrate data visualization techniques and integrate the concepts we learn through this course. In this paper, I will showcase the data visualization skills in Tableau, an interactive data visualization software focused on business intelligence, with some publicly available datasets related to COVID-19, and demonstrate the process of the data exploration and visualization designs.

III. Approach & Results

Several COVID-19 related datasets provided by Kaggle.com are picked for the project:

1. US counties COVID 19 Dataset [1]
2. COVID-19 World Vaccination Progress [2]
3. Covid Vaccinations in United States [3]

Since I am interested in visualizing how COVID-19 and its vaccinations spreads across the world or the U.S counties, I believe my visualizations would mainly be some map graphs with a time range slider or a control to illustrate the numbers of cases, deaths, vaccinations for each part of area over time. So, I have researched some approaches that utilize choropleth maps or scatter map boxes, and there are two feasible solutions for this final project: Plotly Express's scatter map boxes feature and Tableau's map [4][5]. However, for Plotly I was not able to find a

solution to integrate a parameter slider to a scatter map box. So for ease of development, I chose Tableau with which I can visualize the data in a variety of graphs besides maps.

1) U.S. COVID-19 Dashboard

   a) U.S. Top 10 States

Date

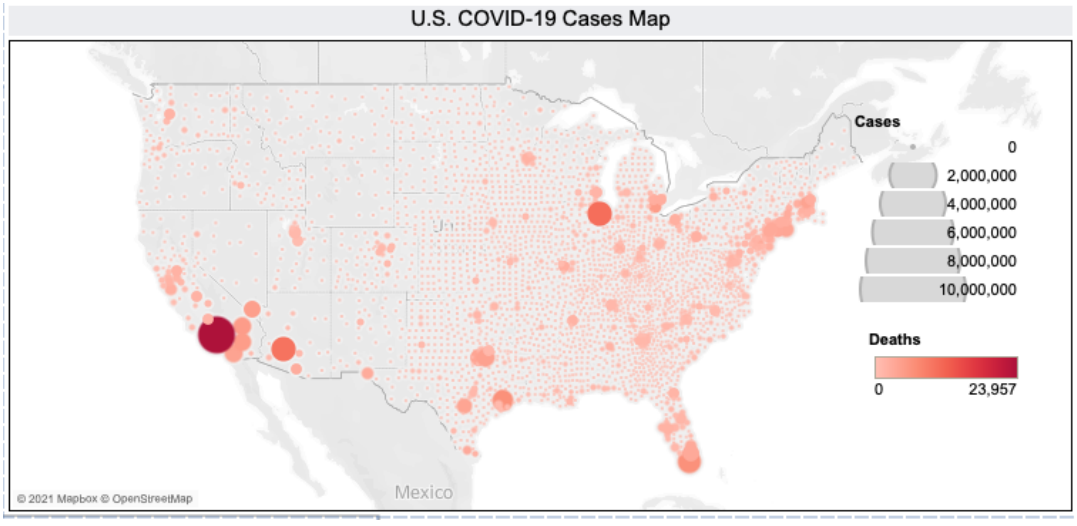
5/5/2021

Top 10 States

State	Cases	Deaths
California	3,750,919	62,063
Texas	2,906,458	50,626
Florida	2,253,921	35,477
New York	2,060,075	52,005
Illinois	1,350,094	24,443
Pennsylvania	1,168,595	26,446
Georgia	1,081,398	19,640
Ohio	1,078,734	19,344
New Jersey	1,004,525	25,692
North Carolina	983,840	12,744

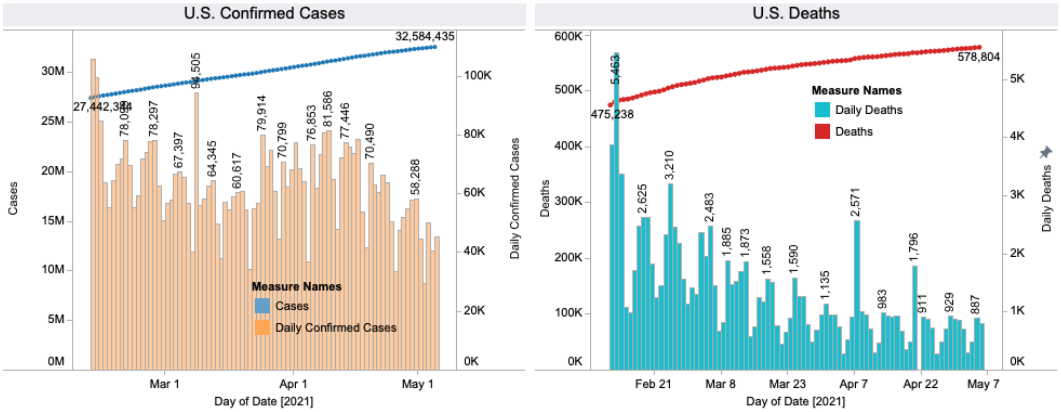
This text table visually illustrates the top 10 states with the most confirmed. COVID-19 cases along with the numbers of death cases over time which is controlled by a ‘Date’ filter.

   b) U.S. COVID-19 Cases Map



This map graph shows the number of cases and deaths for each county in the U.S. with the circle of which size indicates the confirmed cases and color indicates the numbers of deaths. This is also an over-time graph and is controlled by a shared ‘Date’ filter as well.

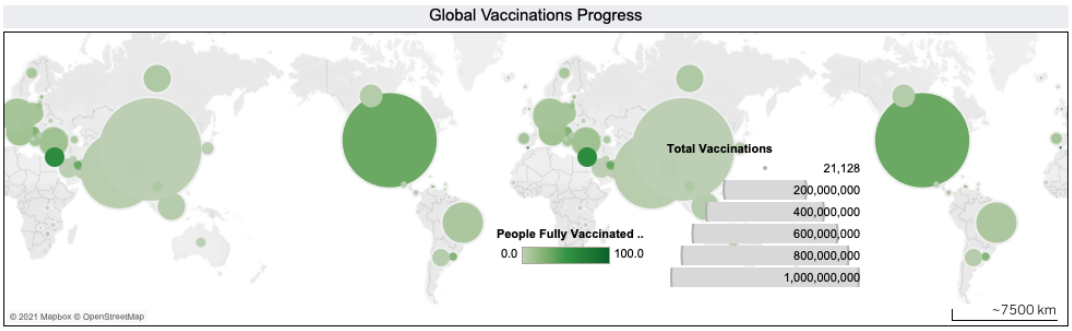
c) U.S. Confirmed Cases & Deaths



These dual combination graphs are having a line chart and a bar chart to show the cumulative and daily numbers for confirmed cases and deaths respectively. The ratio of the cumulative numbers vs. daily numbers is so large that if the axis for two metrics are not synchronized, and for clarity mark labels are added to the bar chart.

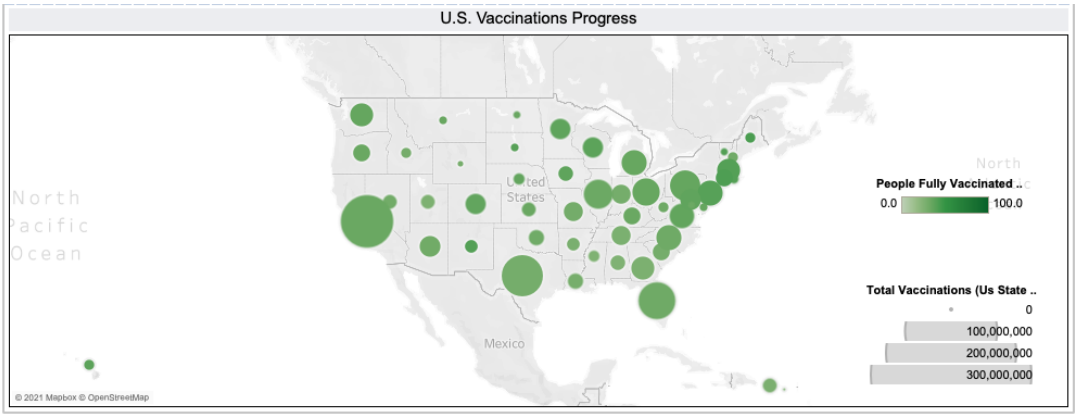
2) Vaccinations Progress Dashboard

a) Global Vaccinations Progress



Dataset “COVID-19 World Vaccination Progress” is used to illustrate how COVID-19’s vaccinations are progressing globally with a choropleth map. In the graph, each circle’s size indicates the total vaccination counts for each country, and its color (white to green) shows the people fully vaccinated per 100 people.

b) U.S. Vaccinations Progress



Dataset “Covid Vaccinations in United States” is used for the visualization. In the choropleth map, each datapoint’s size indicates the number of people who are totally vaccinated for each state, and the color illustrates the ratio of the people fully vaccinated per 100 people.

IV. Conclusion

In the final project, I utilized publicly available COVID-19 related datasets and used Tableau to demonstrate the way to visualize the data and illustrate the world’s situation regarding the pandemic spreads and vaccinations progress. This project was a great opportunity for me to tackle the topic from the real world, and apply the theories and the skills learned through the course.

V. References

[1] US counties COVID 19 Dataset - Kaggle.com

https://www.kaggle.com/fireballbyedimyrnmom/us-counties-covid-19-dataset

[2] COVID-19 World Vaccination Progress

https://www.kaggle.com/gpreda/covid-world-vaccination-progress

[3] Covid Vaccinations in United States

https://www.kaggle.com/gpreda/covid-world-vaccination-progress

[4] Scatter Plots on Mapbox in Python

https://plotly.com/python/scattermapbox/

[5] Build Maps in Tableau

https://help.tableau.com/current/pro/desktop/en-us/maps_build2.htm