Rishabh Rajesh and Nicholas Gildenhuys

Justin Gould

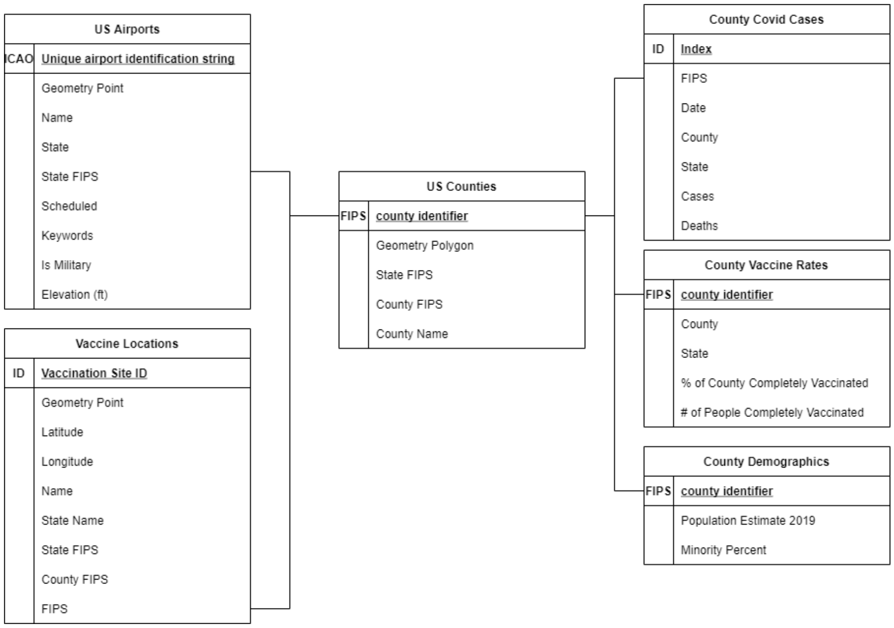
HONR 39900GA

October 19, 2021

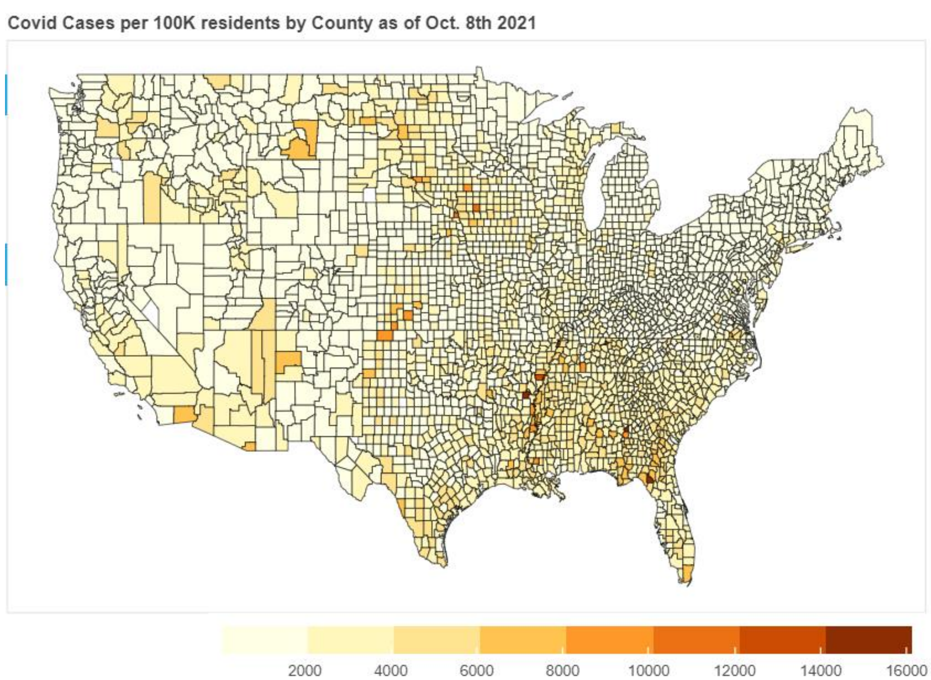
Midterm Report: A COVID Vaccine Distribution Plan

The goal of this project was to create a plan for COVID-19 vaccine rollout and distribution for the United States. With the use of geospatial data about US county demographics, vaccination sites, COVID-19 case and vaccination rates we hope to outline a viable plan for distributing the COVID-19 vaccine equitably. In the process of outlining this task there are some factors that we had to consider. Namely, population, the immunity rate, demographics, and logistically how we plan to distribute such a mass scale of vaccines.

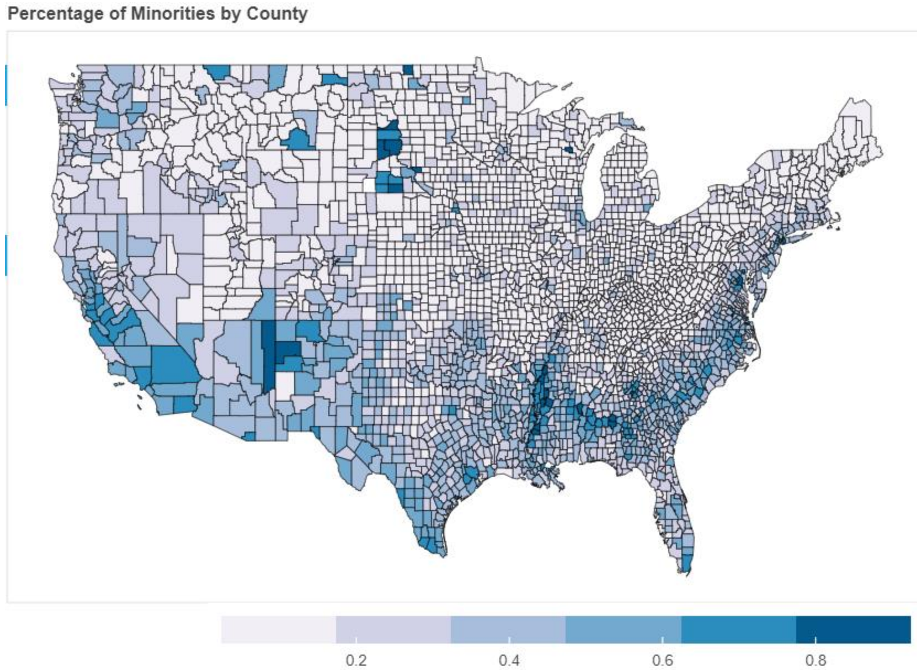
Our approach to this problem was fairly simple. We first gathered data. We pulled most of our data from the us government databases. The CDC for vaccine data and covid case numbers by county. The US Census Bureau for population estimates and demographic data. The given table of CVS locations around the United States. For our consideration into logistics, we pulled from an online database OurAirports for geospatial data on military and civilian airport locations. Once we had the data in its raw form, we used Python coupled with the GeoPandas library to preprocess and clean up the data before committing it to our database. We also made sure to link each table with the appropriate county identifying FIPS code in order to relate it to the other tables.



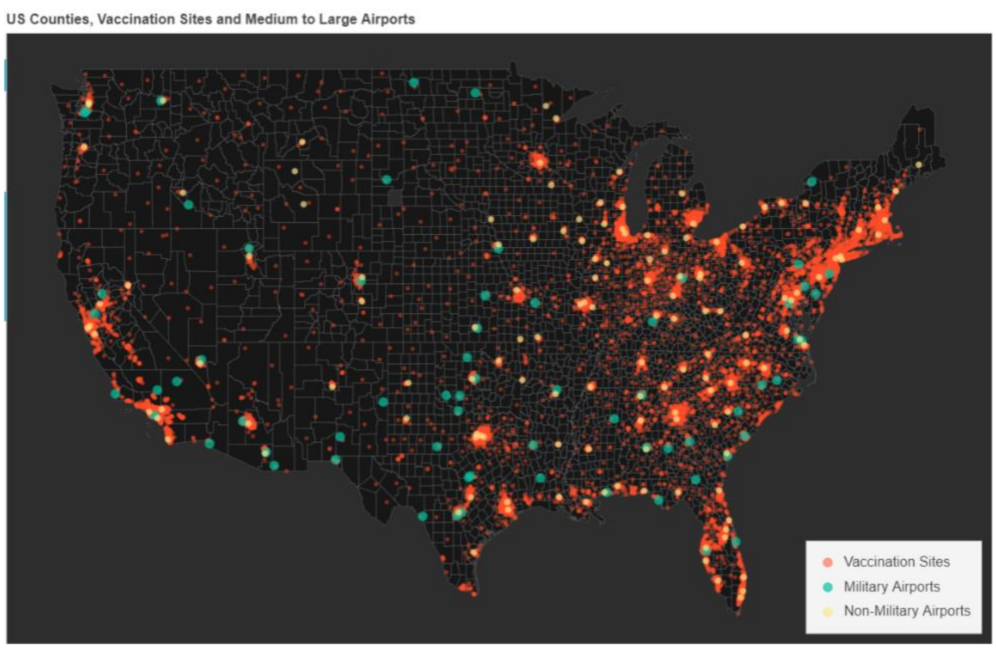
We then utilized the postgres/postgis database for querying the necessary information for our visualizations. Specifically, building a visualization of the COVID-19 case rate by county, as can be seen below:



Along with the distribution of minority groups among US counties.

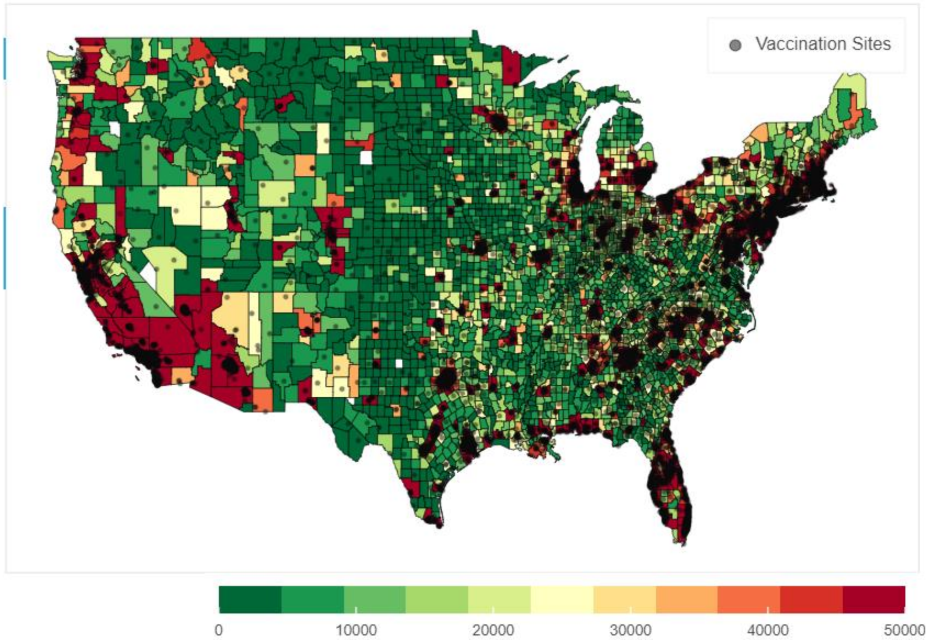


Finally to help visualize the logistics of the distribution effort we mapped out all the viable airport locations for vaccine distribution in each state.



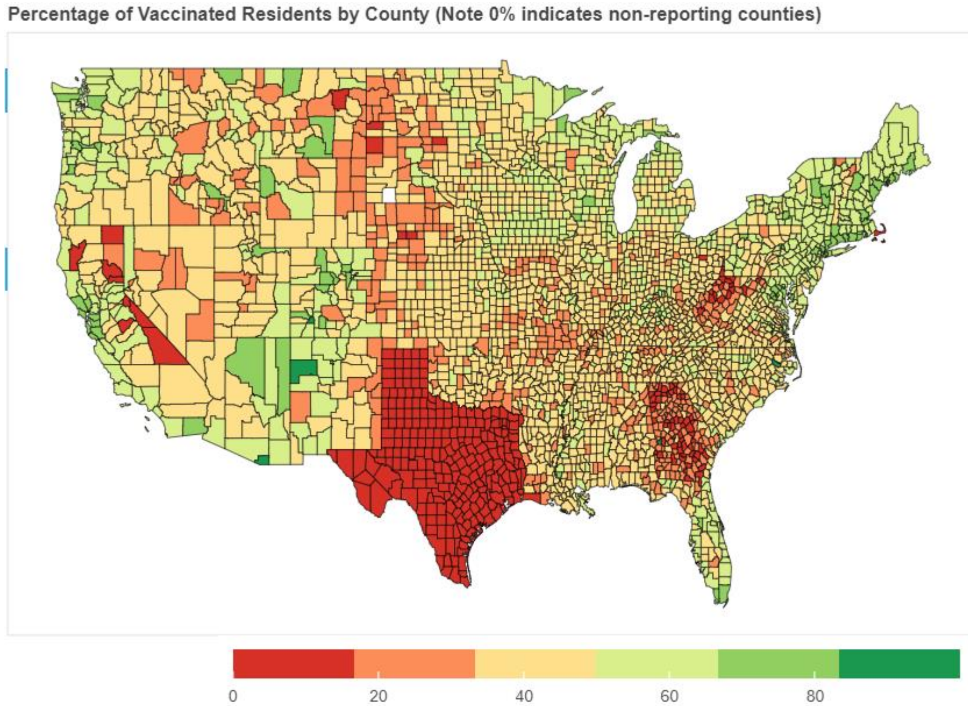
We built these visualizations to outline the current covid status and vaccination needs. Then married that with the data on our distribution network to build a robust display mapping out the number of required vaccine doses for each vaccination location in a state.

Allocated Vaccines Per County



As you can see in our resulting map, the counties of higher density by the mere size of the population will inevitably require more vaccines. This plot is purely by the numbers, in order to apportion vaccines out on the basis of Johnson & Johnson's production would require a bit more tweaking in the data's representation which we would add as an additional feature. However, we do see instances where because how we gauged vaccination rates in certain counties, there are some red counties that seem to require many more vaccine's than their neighbors. This can be attributed to our minority influence metric, and directly shows the disproportionate impact the COVID-19 pandemic has had on the US.

In order to clean up the data and make a cohesive plan we had to make a few assumptions about the data, and some adjustments in cleaning up the data. The first major assumption we made was assuming that the population that got COVID-19 would not require and will not be allocated vaccinations because they have already built immunity. When looking at vaccination hesitancy we found that about 70% of Americans are willing to take the vaccine, and so we allocated 70% of the required vaccines for each county for that county, which by this point in time will be more than enough to cover demand (Ebbs). When looking at vaccination rates, we found that Texas does not report any of their county level data. As can be seen in the map generated below:



To cope with this major gap in information, we took a look at the demographics of each county. As we know, this pandemic has had disparate impacts on minority communities. To cope with this difference, we assigned the vaccination rates in Texan counties based on the majority racial makeup of a county. So, if the county's population majority was a national minority or combination of, then we assigned the county the national vaccination rate for minorities. Else we would assign the county the national vaccination rate. For visualization purposes we left out Alaska and Hawaii from the final maps. Another large assumption we made was that each county had at least one place to distribute vaccines. If a given county had no CVS's we assumed that the town hall would fill in for vaccine distribution, although this is not visualized. Finally with regards to the distribution network and airport mapping, we only mapped out civilian airports that classified as large, and military airports that classified as medium or large. Along with this when calculating the nearest airport for a given vaccination location, we only considered airports within the state in order to cut down on querying time. Although, because of these two constraints we had to add in 2 regional airports in Wyoming because there were no large or medium military or civilian airports in Wyoming in our dataset.

Looking to the future there are many features that we wish we could add to the tool in order to make better decisions when it comes to vaccine distribution. One of the major realizations was just the sheer number of vaccines required to create an immune population. To compensate for this, we saw some other teams make the number of vaccines to be distributed to a given county a proportion of the state's allocation. With our current code base this would not be too difficult to implement and would help model a more realistic distribution pattern as Johnson & Johnson can only make so many vaccines in a given time span. Another improvement that we found would be good to add would be to look not just at the demographic makeup of a county to help inform vaccine distribution, but to look at the median household income of a county. This would in fact probably be a better metric to help model the disparate impact of COVID-19 in the US.

To wrap it all up, the database we created is a cleaned-up relation-based set of tables that, through our visualization, anyone can find the estimated number of vaccines required for a given county, and how to distribute them among the county's vaccination sites. We combined geospatial data regarding population, immunity rates, demographics, and airports to present a cohesive representation of an equitable the vaccine distribution plan. Each of the maps are easily manipulated to find a counties or vaccination site's specific, pertinent data regarding required doses and where to expect them from. More information can be found in the data visualization notebook in the repository cited.

Works Cited

Ebbs, Stephanie. “69% Of Americans Say They've Gotten or Plan to Get COVID-19 Vaccine: Poll.” ABC News, ABC News Network, 5 Mar. 2021, https://abcnews.go.com/Politics/69-americans-plan-covid-19-vaccine-poll/story?id=76281286.

Rajesh and Gildenhuys, https://github.com/rishyraj/covid_vacccine_distribution