Vishal Saravanan

Information Visualization Section 2

Flinsch

15 December, 2024

Data Source

It was odd when I opened the file because I had to open the Excel file as a text file (implemented as a CSV). There were many more columns than any other project we have done before. Usually, there are fewer categorical columns than numerical columns, but there were a lot in this case. Using at least 500 values in each graph/chart was relatively easy because there was a data value for almost every date from 2020 to 2022. I noticed that some columns had too many null values for me to use with certain countries. For example, "Hospital Patients" for Afghanistan mainly had/all null values compared to a country with more data like the United States. Other columns that I felt did not have enough data values were "Handwashing Facilities," "Life Expectancy," and "Hospital Beds." I noticed that even though COVID originated in Wuhan, China. Before using them in my graphs, I used "owid-covid-codebook" to understand what each variable represented. I changed the dataset by converting the"location" column to a geographical category. The "location" column had some values like "High Income" and "North America," for which I had to filter out nulls. I did this for most of the sheets where I removed the null values instead of setting them to 0 (which would lead to inaccuracy). The sources I used for research revolved around CDC and my own experiences with COVID.

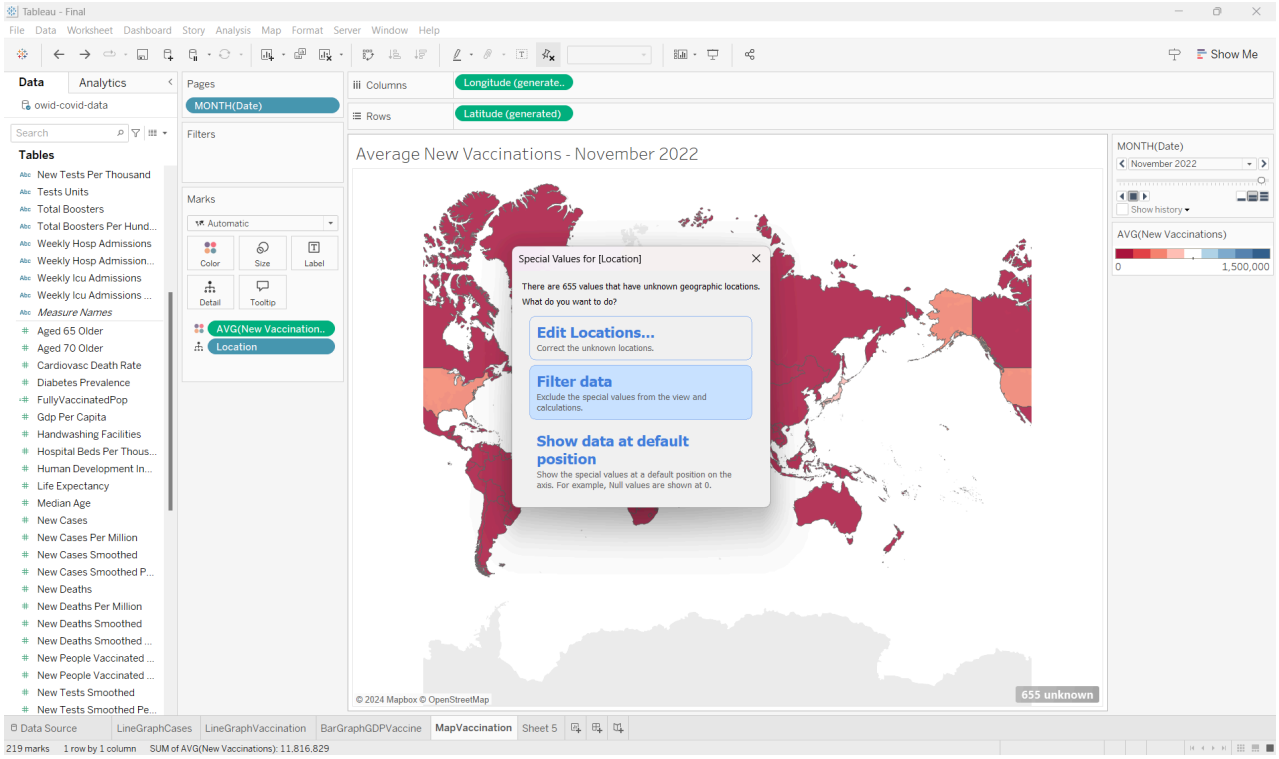


Tableau - Final
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Show Me
Data
Analytics
owid-covid-data
Search
Tables
New Tests Per Thousand
Tests Units
Total Boosters
Total Boosters Per Hund...
Weekly Hosp Admissions
Weekly Hosp Admission...
Weekly Icu Admissions
Weekly Icu Admissions ...
Measure Names
Aged 65 Older
Aged 70 Older
Cardiovasc Death Rate
Diabetes Prevalence
FullyVaccinatedPop
Gdp Per Capita
Handwashing Facilities
Hospital Beds Per Thous...
Human Development In...
Life Expectancy
Median Age
New Cases
New Cases Per Million
New Cases Smoothed
New Cases Smoothed P...
New Deaths
New Deaths Per Million
New Deaths Smoothed
New Deaths Smoothed ...
New People Vaccinated ...
New People Vaccinated ...
New Tests Smoothed
New Tests Smoothed Pe...
Pages
MONTH(Date)
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
AVG(New Vaccination..
Location
Columns
Longitude (generate..
Rows
Latitude (generated)
Average New Vaccinations - November 2022
Special Values for [Location]
There are 655 values that have unknown geographic locations.
What do you want to do?
Edit Locations...
Correct the unknown locations.
Filter data
Exclude the special values from the view and calculations.
Show data at default position
Show the special values at a default position on the axis. For example, Null values are shown at 0.
© 2024 Mapbox © OpenStreetMap
655 unknown
MONTH(Date)
November 2022
Show history
AVG(New Vaccinations)
0
1,500,000
Data Source
LineGraphCases
LineGraphVaccination
BarGraphGDPVaccine
MapVaccination
Sheet 5
219 marks
1 row by 1 column
SUM of AVG(New Vaccinations): 11,816,829

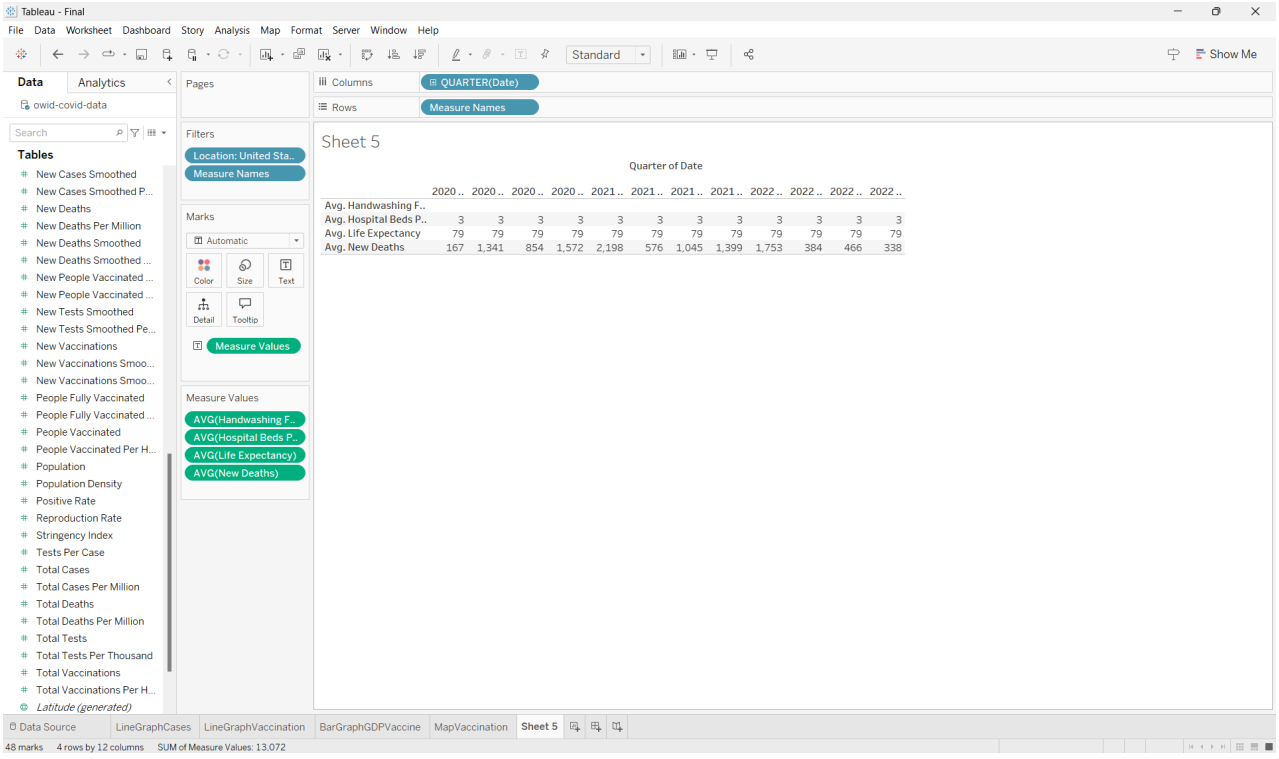


Tableau - Final
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
owid-covid-data
Search
Tables
New Cases Smoothed
New Cases Smoothed P...
New Deaths
New Deaths Per Million
New Deaths Smoothed
New Deaths Smoothed ...
New People Vaccinated ...
New People Vaccinated ...
New Tests Smoothed
New Tests Smoothed Pe...
New Vaccinations
New Vaccinations Smoo...
New Vaccinations Smoo...
People Fully Vaccinated
People Fully Vaccinated ...
People Vaccinated
People Vaccinated Per H...
Population
Population Density
Positive Rate
Reproduction Rate
Stringency Index
Tests Per Case
Total Cases
Total Cases Per Million
Total Deaths
Total Deaths Per Million
Total Tests
Total Tests Per Thousand
Total Vaccinations
Total Vaccinations Per H...
Latitude (generated)
Pages
Filters
Location: United Sta..
Measure Names
Marks
Automatic
Color
Size
Text
Detail
Tooltip
Measure Values
Measure Values
AVG(Handwashing F...
AVG(Hospital Beds P...
AVG(Life Expectancy)
AVG(New Deaths)
Columns
QUARTER(Date)
Rows
Measure Names
Sheet 5
Quarter of Date
2020 .. 2020 .. 2020 .. 2020 .. 2021 .. 2021 .. 2021 .. 2021 .. 2022 .. 2022 .. 2022 .. 2022 ..
Avg. Handwashing F..
Avg. Hospital Beds P.. 3 3 3 3 3 3 3 3 3 3 3 3
Avg. Life Expectancy 79 79 79 79 79 79 79 79 79 79 79 79
Avg. New Deaths 167 1,341 854 1,572 2,198 576 1,045 1,399 1,753 384 466 338
Data Source
LineGraphCases
LineGraphVaccination
BarGraphGDPVaccine
MapVaccination
Sheet 5
48 marks
4 rows by 12 columns
SUM of Measure Values: 13,072

New Cases of COVID-19 in Different Locations

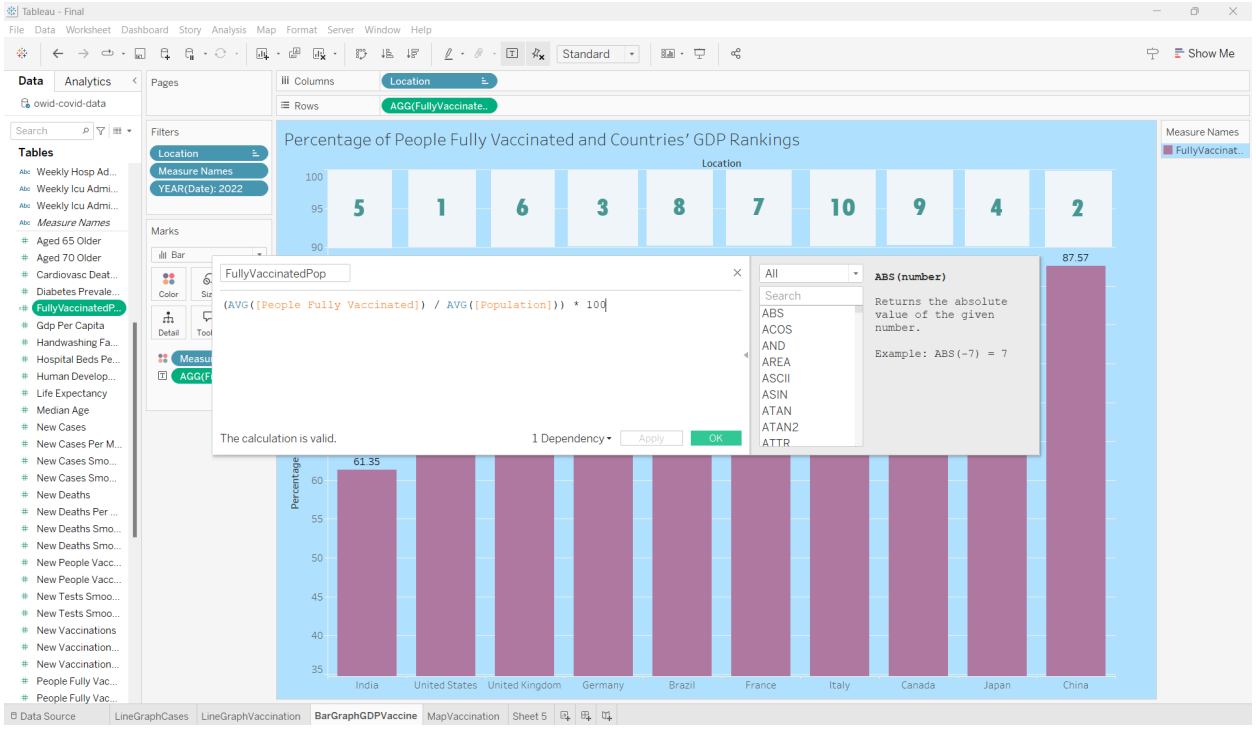
I chose a main graph that could be referred to in my story when looking at other graphs. The graph shows the number of New Cases of COVID-19 across the countries with the most cases. I chose my countries by referring to the CDC's database of the most affected countries(and the virus's origin, China). I included numerous annotations to refer to specific variants that arose and led to an increased number of new cases. I also highlighted trends and abnormalities, such as a spike in new cases in India in May 2021. I used a line graph to show how many new cases there were throughout the pandemic and made adjustments to have multiple countries. Each country has a different color to contrast with the light background. I chose to divide the date column into columns. The date column was interesting because I could divide it into years, quarters, months, etc., which was my first time seeing this. I experimented with it for some time and chose quarters because the text fit on my graph, unlike months. I also ensured that I chose average new cases (which would find the average across each quarter) and used the smoothed graph, which looked much cleaner than the normal one.

New Vaccinations of COVID-19 in different locations

I began this graph by duplicating the previous graph and deleting the "New Cases Smoothed" column before adding the "New Vaccinations Smoothed" column. Once more, I analyzed trends and added annotations based on the graph. The x-axis starts at the end of 2020 Q4 rather than 2019 Q4 in the previous graph because vaccinations were not created and released to the public until a year after the virus started spreading. I planned to have this graph and the previous graphs be the more general ones that can be referred to in my story when making the other four graphs.

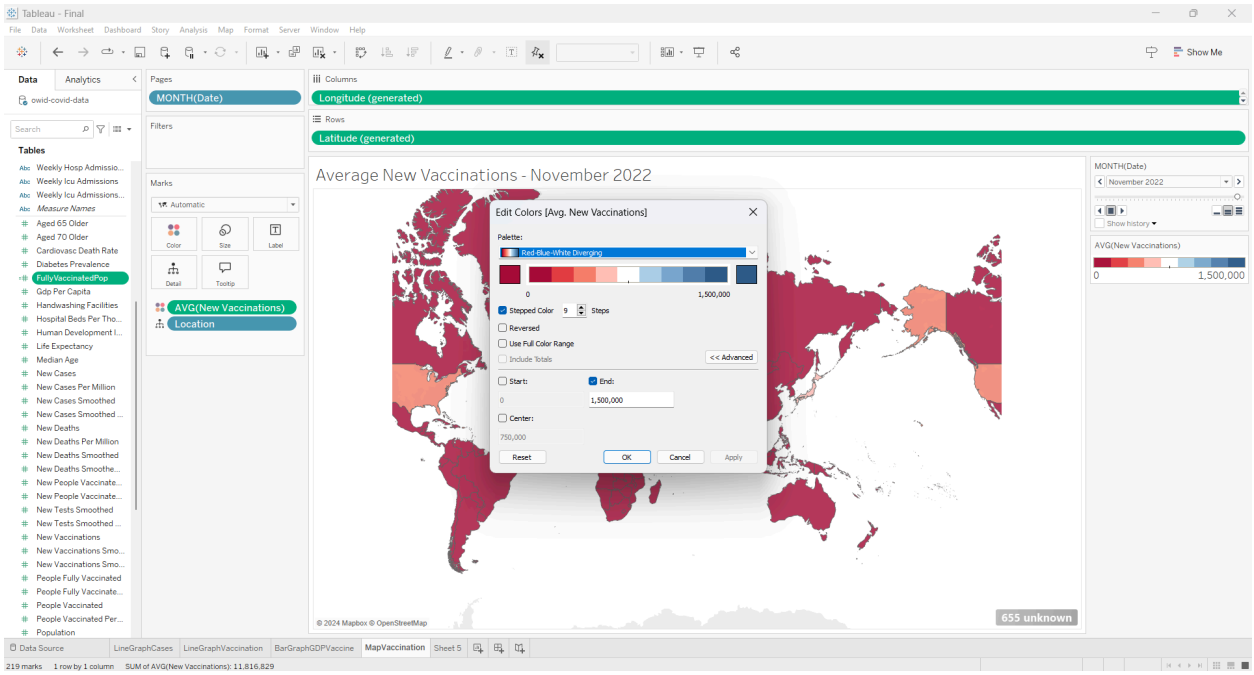
Percentage of People Fully Vaccinated and Countries' GDP Rankings

I chose the countries with the ten highest overall GDPs to calculate the percentage of fully vaccinated people in those countries. My initial assumption was that countries with a higher GDP would have more money to vaccinate people. To measure the percentage of people fully vaccinated in a country, I divided the average number of people vaccinated by the average population. I used annotations and a bold font to rank each GDP above each country and the exact vaccination percentage as text above each bar. I chose colors that looked appealing together (blue/green/purple) and ensured that these measurements were done in 2022 (the period that makes the most sense).



Average New Vaccinations

I had already converted location to a geographical variable and included the Date column on the pages. I divided based on months (which would give me 36 pages to play through). I edited the colors such that it would go from red to blue and have nine steps, starting from 0 and going to 1.5 million new vaccinations. I played around with the "End" parameter to ensure it would make the most sense.



New Deaths and Total Deaths

I created two line charts representing the new and total deaths over time. I did not put them on the same chart because they had a different y-axis scale. The annotations I added explained spikes in new deaths and trends. I filtered the location as the entire world chose a period that made sense (months) and implemented a contrasting color scheme. I also figured out how to adjust the lines and line ends for the annotations.

Bar Chart GDP Per Capita

I made this chart similar to the other bar chart but focused on countries with the top five highest GDP per capita and added pages for the user to scroll through. I used the same calculated field from earlier. I checked the data, and the population values for Qatar were less than the number of people vaccinated, so I have to assume there is at least a 10% margin of error. I ensured that the y-axis scale was from 0-100, even if Qatar exceeded that mark.

Global Impact of COVID-19: Trends, Vaccinations, and Economic Insights

Creating the story was very easy. I added my title as text and adjusted the size/font before placing it on the top left corner of each sheet. Then, I adjusted the layout to dots and made the size of each story automatic so that everything would fit properly. I added a title for my story to represent what I was trying to show.

Works Cited

Centers for Disease Control and Prevention. (n.d.). Centers for Disease Control and Prevention. https://www.cdc.gov/

Worldometer. (2024a). GDP by country. Worldometer. https://www.worldometers.info/gdp/gdp-by-country/

Worldometer. (2024). GDP per capita. Worldometer. https://www.worldometers.info/gdp/gdp-per-capita/