DATA ANALYSIS

04/07/2021

Using R for data analysis

Project for MATH 222 Spring 2021

This is a project for data analysis of the spread and effect of corona virus all over the globe. It was done using "R" as a data analysis language, and the attempt is to provide some relation between countries as well as answering some questions regarding the pandemic.

Joao DeCastro
WCSU
Decastro008@wcsu.edu
Data analysis project

MAT222
Spring 2021
Michael Saccucci

Project Report

Project Name: Corona virus analysis report

03/25/2021 started

Status Code Legend

- Done
- In progress
- Not started
- Problem

DELIVERABLE:

The main point of the project is to use R to analyze data and answer some relevant questions about the pandemic. I will be comparing data specifically from Brazil, where the cases and deaths are increasing a lot. I will be comparing political actions taken in the USA and in Brazil. Also comparing vaccination programs between Israel, Chile and Mexico. Create relations of the effect of those actions and see if there are significant difference between them

Brainstorm on questions, relation to address/answer using data	• Clean data for each question • Decide what type of graph for each • Number of beds in hospital available • PIB of country • How much was invested in public health • Dates of lockdown • How much vaccines were bought? • How many vaccines per day? • Is it working? Lowering cases? • Countries with highest rate of vaccinations • If USA continues in this path when is It going to be done? • When is brazil going to be done with vaccines from this rate • Why brazil is getting so many number of deaths and cases
Get familiar with R Interpret data Clean data	• Clean data • Making data frames, new tables • Make simple plotting • Make some graph
Start to work on answering questions	• Question : Show graphs of investments in public health of countries • Question : Show how many beds available

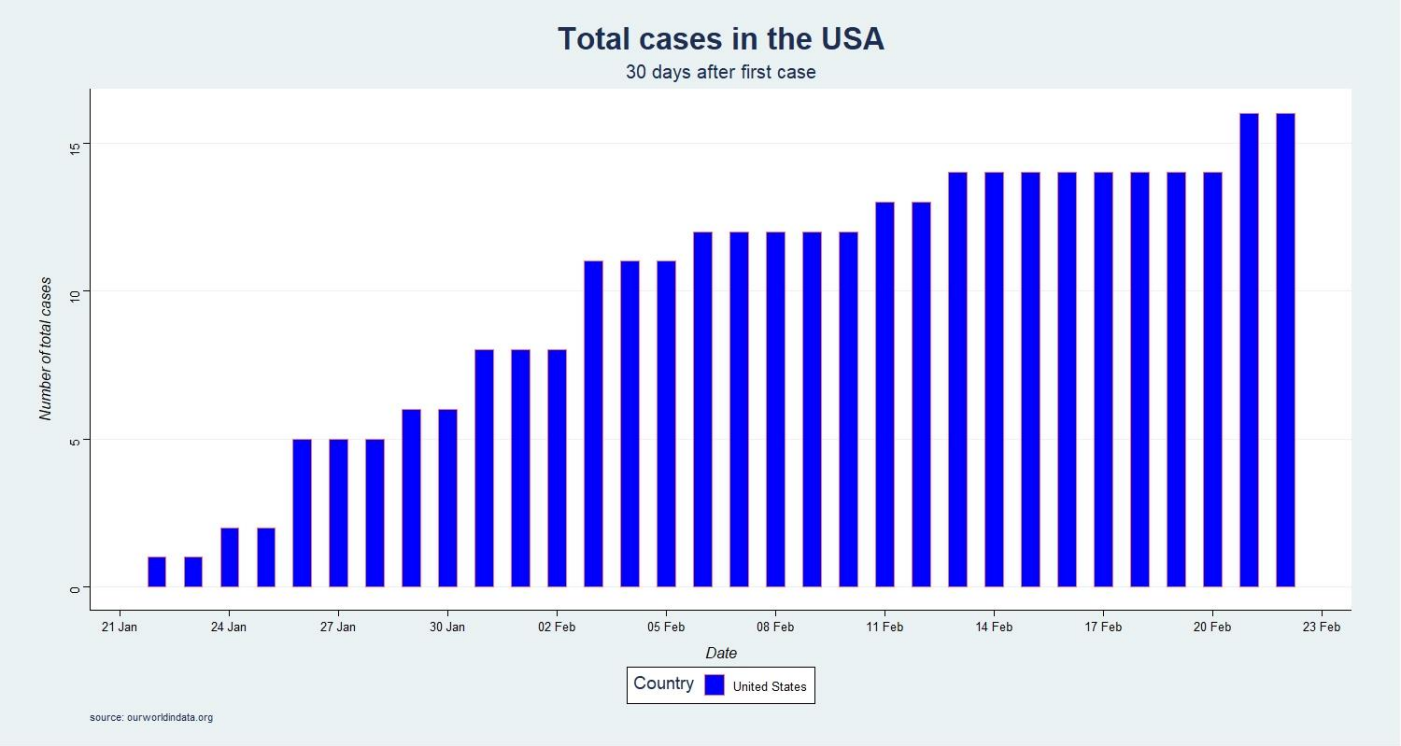
	• Question : Show dates of lockdown and its effects • Question : Events, holyday change the spread of virus? – get dates of holydays etc • Question : Lockdown worked? • Question : Why brazil have so many cases and death now, maybe after carnaval? Cities with lockdown, etc • Question : Countries that bought vaccines when it was available • Question : age of population, diabidies obesity? • Question : Vaccination, per day etc • Question: What percentage vaccine in brazil? • Question : When its done with vaccines •
Learning outcomes	• Clean raw data • Work creating new data frames • Creating specific time, dates, variables • Making simple graphs • Use ggplot2 • Creating better looking graphs • Graph animation
Finishing report	• • •

.

Everything began right after the 2020 holydays, the world was celebrating and wishing new and better things for 2021. We would never imagine that a pandemic was approaching. Most of our generation just read about pandemic on history books, and we were not expecting something like this to happen. Specially nowadays, with so much technology we think that we are safe all the time.

In the USA everything started on January 22nd when the first official corona virus case was confirmed. This is where we can start look at some data and visualize what was happening and what happened a little after the first case was confirmed.

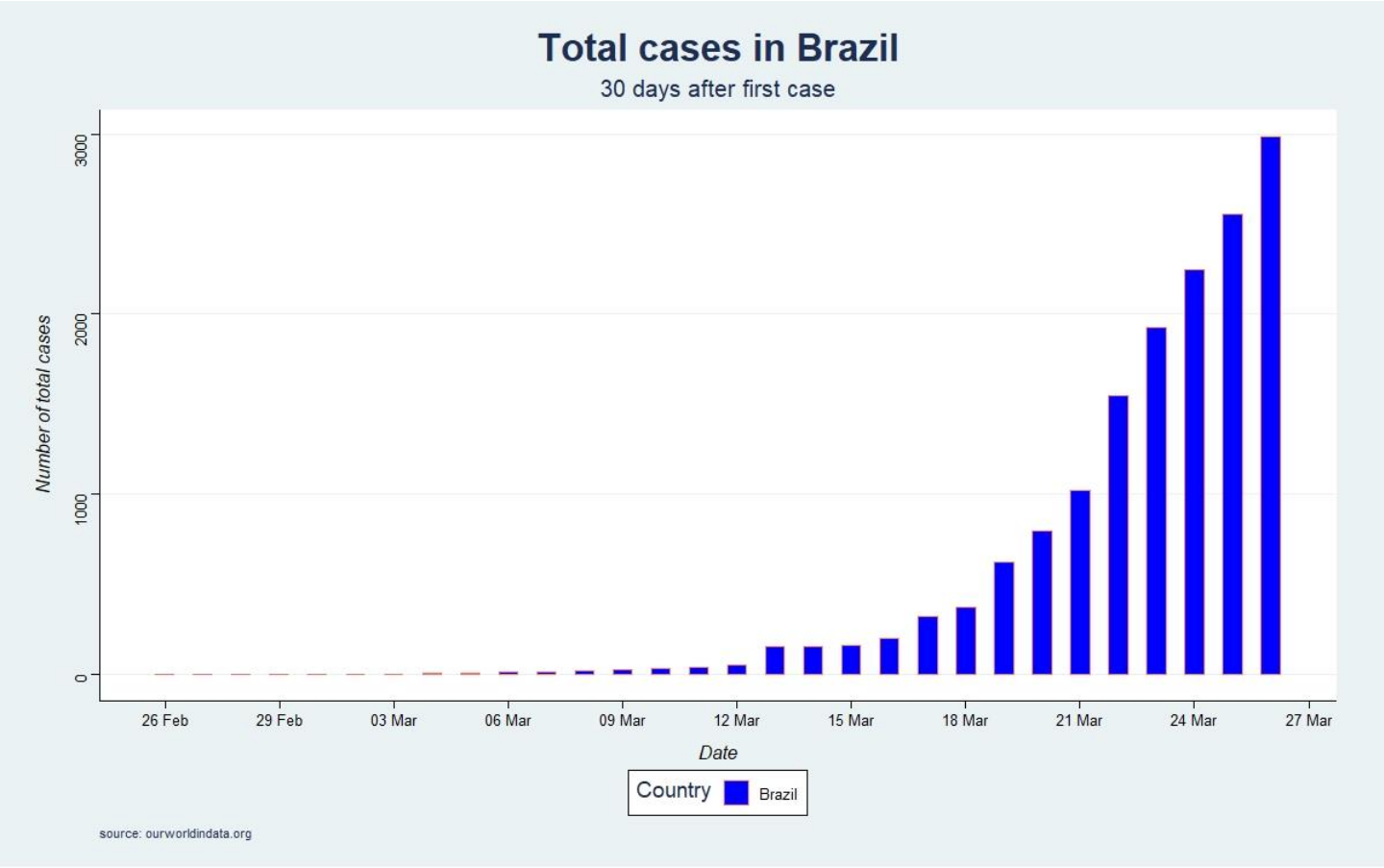
❑ First case in the USA vs 30 days later



- *NOTE: There was not many testing available yet, that was probably why the numbers were low in the beginning. Also, most people that was infected was treating at home*

First case in brazil was in February 26 while in the USA there were only 16 confirmed cases by then, ten days later 237 and ten days later 4679. This is how rapidly the virus can spread.

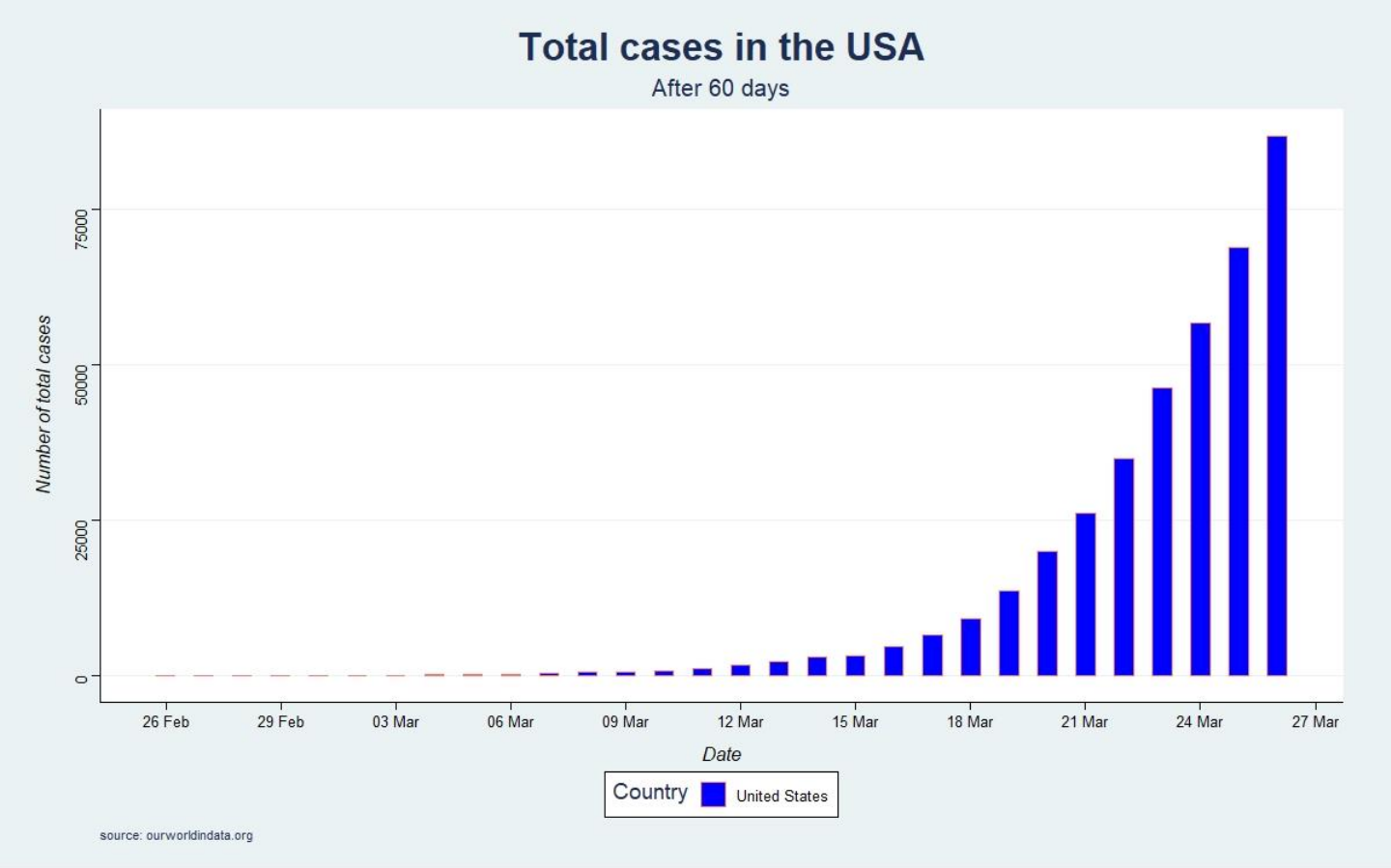
❑ First case in Brazil vs 30 days later



We can clearly see the difference between the two countries. First case in Brazil was almost a month later comparing to the USA, and we can see the difference of the exponential increasing rate of cases in Brazil. In just 10 days cases came from around 500 to 3000 total cases.

Just for comparison, the USA in this same period. Where the virus was around for more then 60 days now

- 60 days after the first case in the USA, comparing to Brazil

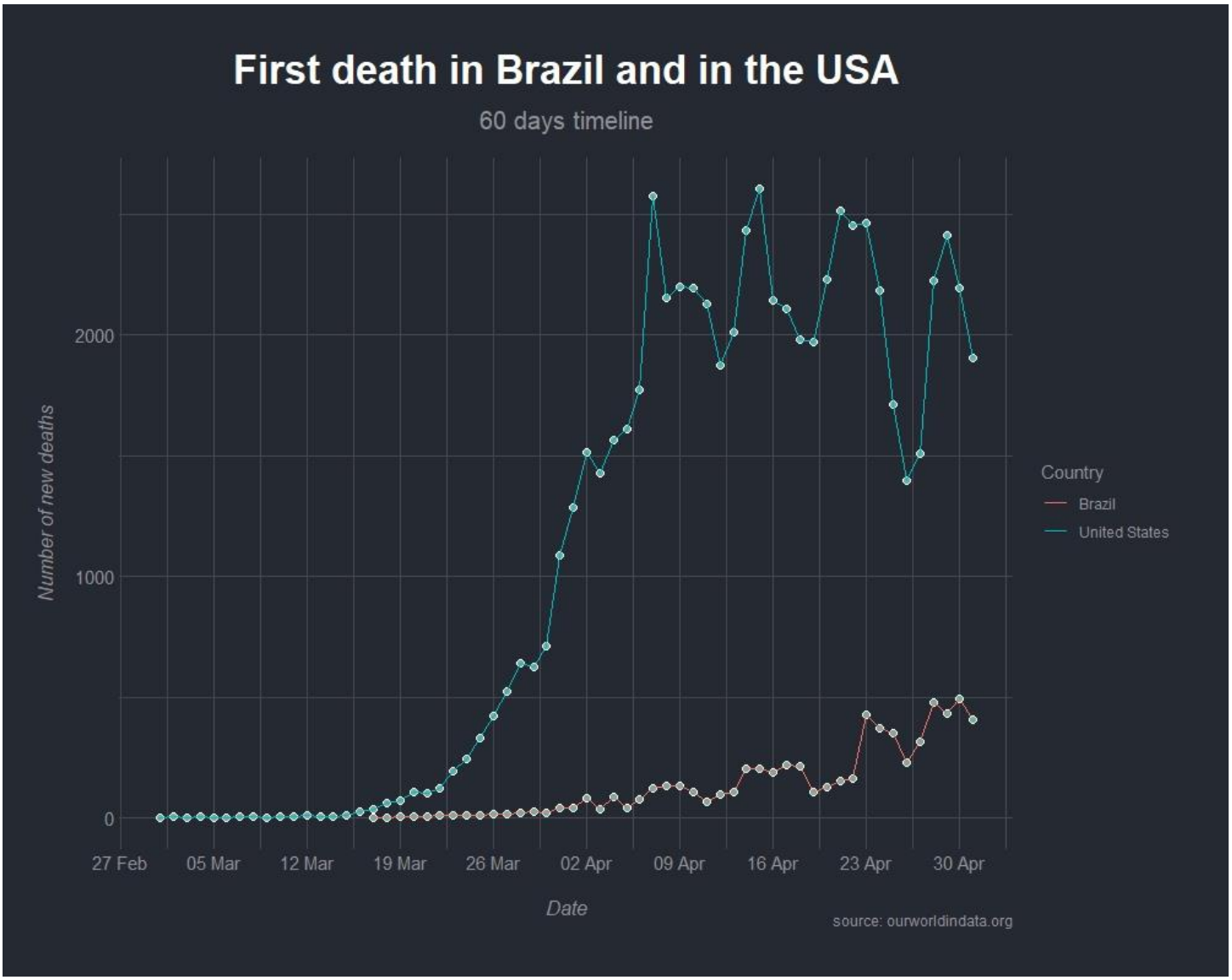


Now it is even clearer. Brazil's first case was approx. 30 days later comparing to the USA, and in this graph, we can see that while brazil was reaching the mark of 3,000 total cases mark, the USA was passing the 80,000. I a week the number came from around 20,000 to almost 80,000.

Just with these 2 graphs any other country in the world should know what was to come. However, as early as August the Brazil's president and his minister passed the chance to buy Pfizer vaccines. Pfizer offered Brazil great number of vaccines to be delivered until December, and president Bolsonaro just refused to buy it, unbelievable.

Now let's see when the first case of confirmed deaths in both countries was, in the USA the first case of death was registered in March 1st and in brazil was in March 17th. With this next graph we can see the timeline of 2 months after the first death in both countries

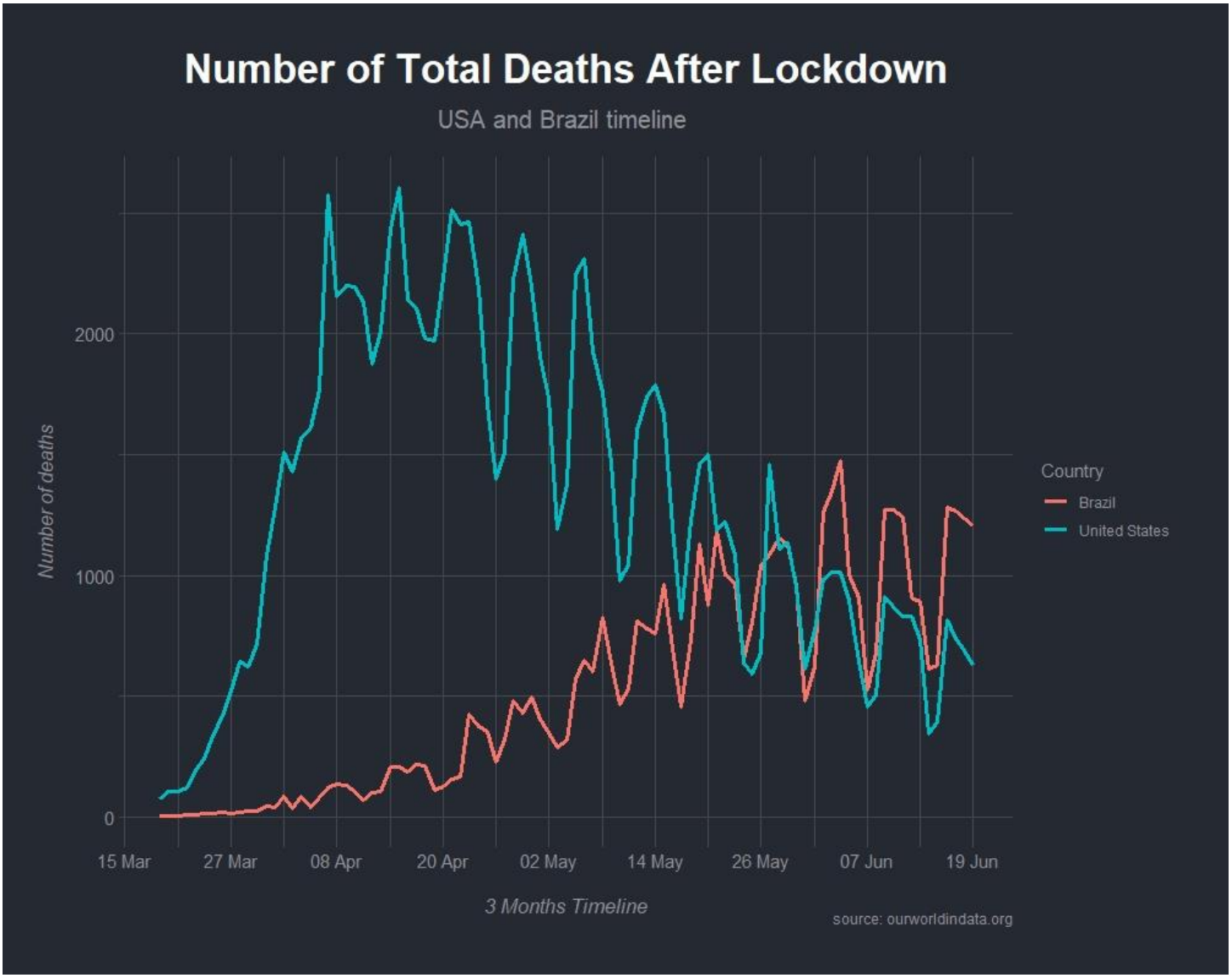
❑ Graph from first deaths and a time lime for 60 days



The number of deaths in the USA increased really fast after the first confirmed death case. After 20 days of the first death, the USA was registering in average 2000 a day. Another thing that we can note analyzing this graph is that the first official lockdown in the USA was in Mar 19th in California and Mar 23rd in Connecticut. We can see that it took some time for the lockdown to start making some difference. During the first 20 days the numbers of deaths was increasing rapidly. (timeanddate, n.d.)

On this next graph I am going to show even further to see if we notice a difference after the lockdown was officialized. Let us make a timeline of 60 days after the lockdown in the USA.

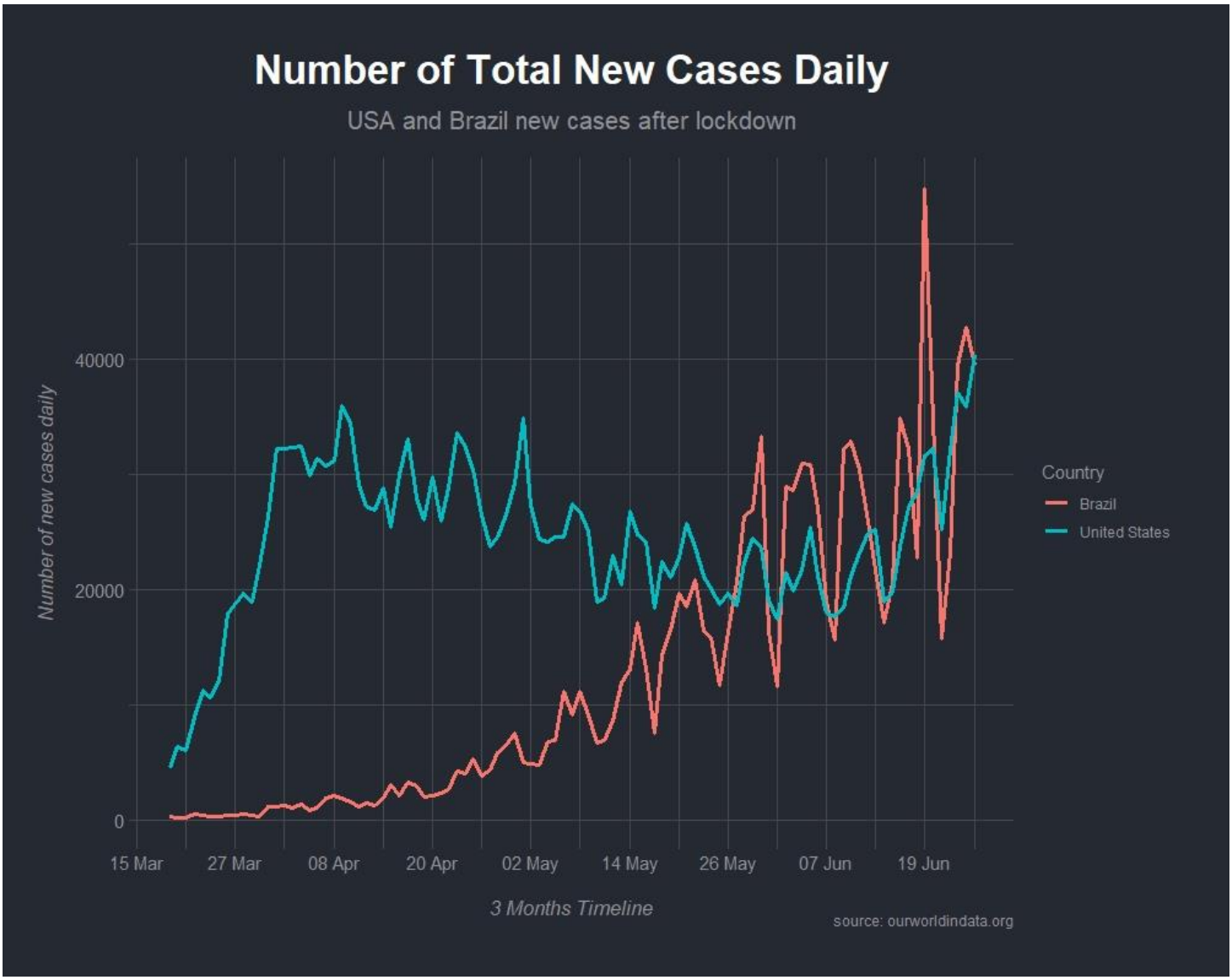
- Graph showing number of deaths daily after the lockdown



From this graph its clear that the lockdown worked and saved several lives, we can see the total number of deaths decreasing. Also, we can see that the number of deaths in Brazil was still increasing because there were places without an official lockdown and, in big cities like Sao Paulo social distancing was almost impossible.

Let us look at the number of new cases after lockdown for both countries

❑ Number of new cases after lockdown

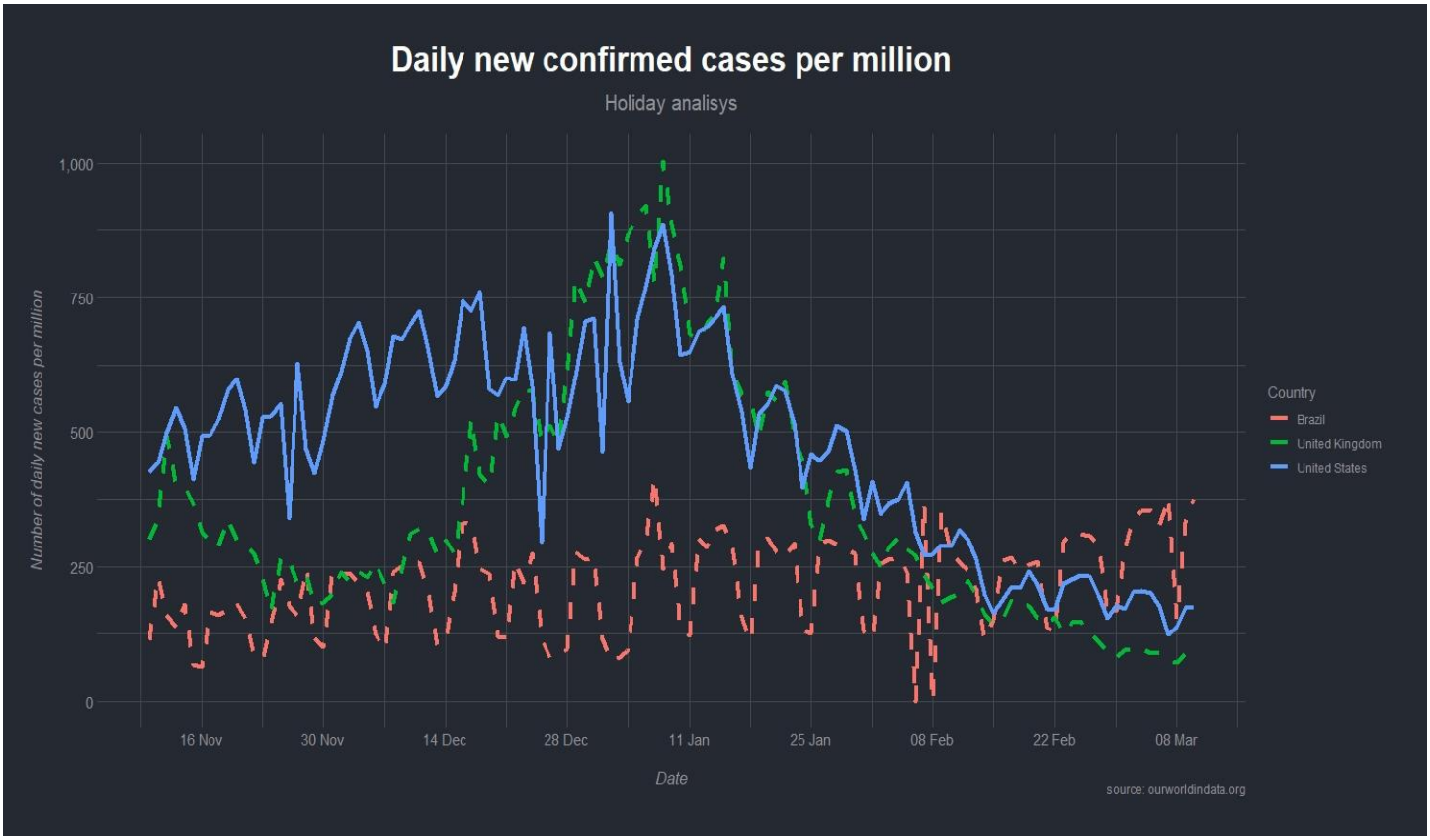


We see that the daily cases in the USA was somewhat stable for some time, comparing to Brazil that was increasing the number of daily new cases. Now, it is well known how fast this virus can spread and when we look far in time huge numbers are accounted. I am going to move forward in time to analyze the data on and after important holidays.

To be able to analyze it, we are going to look at period of time such as Thanksgiving, Christmas and Carnaval in brazil.

2021

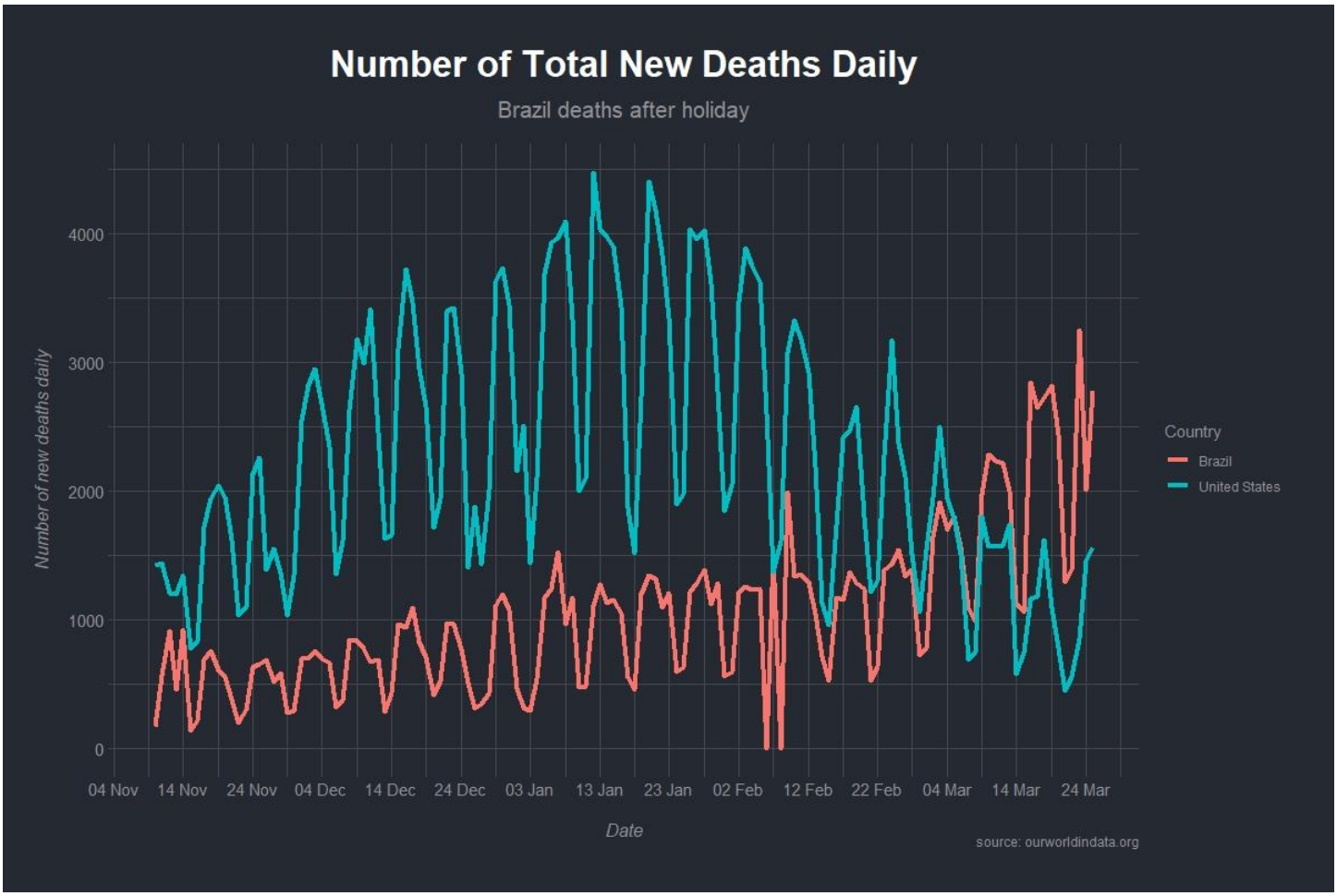
❑ Graph showing daily cases after thanksgiving



I added United Kingdom to compare with the USA because in this period Brazil was not able to keep with the demand of tests. That is why it shows low number of new cases. But it is clear that after the holidays the number of cases increased a lot, specially in the United Kingdom. The curve right after the holidays is very clear, and the number of cases decreasing after he holyday.

Another way that we can compare the situation in Brazil is looking at the number of confirmed deaths, since there were not enough tests. The number of deaths can be used to interpret the number of cases, according to the Imperial College it was estimated around 2% deaths of a sample population infected with covid.

❑ Comparing number of deaths after holiday



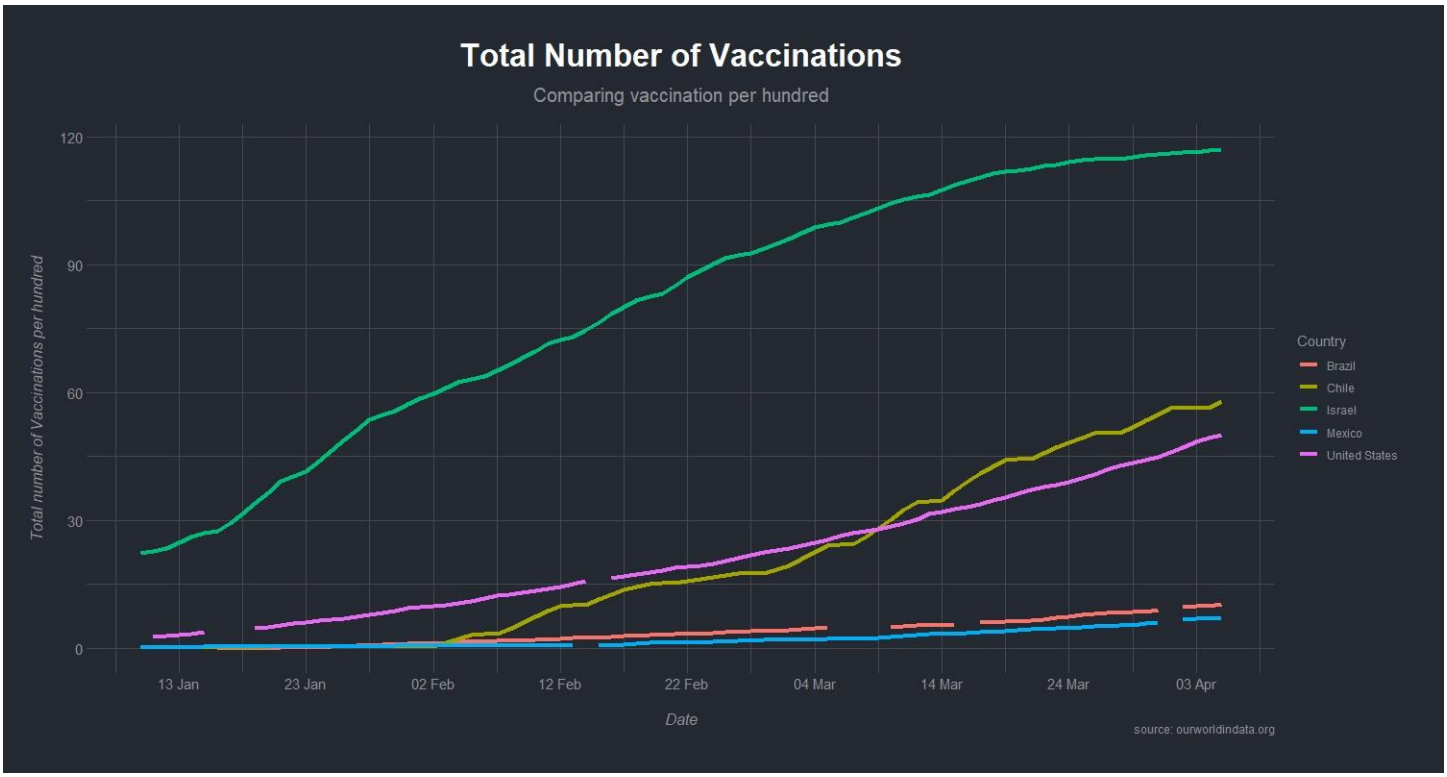
Now we can have a better look comparing the two countries. Brazil increased a lot the number of deaths after February (after carnaval), and it was around 3,000 deaths per day which we can interpret as 150,000 cases per day. This number would make more sense comparing to the high number of daily cases in the USA from the previous graph.

The number of daily deaths in Brazil is making people really worried, these past few days the number passed 4,000 deaths in one day. This is a record, while in the USA the numbers are decreasing in a slow rate. Unfortunately, people are not getting vaccines either. They are behind vaccinating 60 years old people just now, and again this is an effect of political decisions made back in 2020 declining the opportunity to buy vaccines when it was available. According to the website medicalxpress, "Brazil has ordered 100 million doses of the Pfizer-BioNTech COVID-19 vaccine and another 38 million of Johnson & Johnson's". However, it is a little to late now, companies are going to

be able to provide the vaccines mid-October. People are sad because Brazil has one of the best immunization programs on the globe, but without vaccines there is nothing to do.

Let us look at the immunization program in some countries so we can have a better idea.

❑ Comparing total number of vaccinations per hundred



Israel is one of the first countries to reach almost the total population vaccinated, and chile is the first country in the South America. The USA has been doing a good job with almost the half of population vaccinated. Brazil and Mexico appear very behind on the immunization race, both are about 10 per hundred people vaccinated. In this rate Brazil and Mexico would finish the immunization by 2022.

Is it lowering the cases?

CONCLUSION

Works Cited

(n.d.). Retrieved from https://ourworldindata.org/

https://www.timeanddate.com/holidays/us/lockdown-day-1. (n.d.).