Analyzing the Impact of COVID and its Correlation with Vaccines

New Mexico

Supercomputing Challenge

Final Report

April 3rd, 2023

Team 8

Capital High School

Team Members

Julia Almeida

Leslie Gutierrez

Guadalupe Rojo

LaTrisha Padilla

Teacher

Irina Cislaru

Project Mentor

Jon Brown

Table of Contents

1. Executive Summary:

COVID-19 is a contagious respiratory illness that affected the whole world. This project aims to measure the impact of the vaccine and its impact on COVID death rates per state, as well as displaying the death rates in every state as a comparison. For this we needed to determine the effectiveness of the vaccine by calculating the correlation between COVID related fatalities and administered vaccines by state, and doing research on how vaccines work and herd immunity. In order to evaluate the effectiveness we researched, we researched the population of each state and developed a graph using the data we achieved from the CDC's COVID death rates by state dataset. Then, we created a separate graph demonstrating the death to population ratio.

The correlation between COVID related deaths and the number of vaccines that were administered in each state were then calculated and the gathered data was then represented visually with graphs, maps, and a scatter plot. Along with the relationship between COVID death rates and vaccines, we also calculated other information such as average, standard deviation, covariance, and range. Using geopandas, we also represented the correlation with geographical information. We are in the process of completing our maps to compare side-by-side.

2. Introduction:

We know that the COVID-19 pandemic has affected the whole world, and it deeply affected us as a community. Luckily, we were able to come out with a vaccine in very little time compared to other pandemics. Other pandemics took years to come out with a vaccine, while it only took around 9 months to produce a COVID-19 vaccine. Regardless of it being tested and approved by the FDA for people sixteen years and older, many people refused to get the vaccine and be protected from the pandemic. When questioned why, 75% of adults said they were skeptical of COVID or they speculated about the effectiveness of the COVID vaccine ("By The Numbers: Who's Refusing Covid Vaccinations—And Why"). Did the vaccine actually prevent higher death rates in the United States? Should people be less skeptical about how effective this medical intervention is? These are the questions we answered with our project.

3. Solution:

Our project is focused on multiple data sets, which have allowed us to develop a program which measures the impact of COVID-19 and how the vaccine affected the COVID death rates per state. Our project consists of a graph, sectioned in proportions that show the death rates with the total population for each state in order to calculate the impact of COVID-19 and its vaccine. The first step taken with this project was to see the effect of the vaccine to see the correlation between COVID deaths and administered vaccines in each state. Then, we presented this collected data into a scatter plot. By seeing this correlation in a scatter plot, we can also observe the standard deviation and standard range of the given data points.

4. What is Herd Immunity?:

Herd Immunity is when enough people get vaccinated against a contagious disease and it reduces the likelihood of contracting that illness. It happens when a large percentage of a population becomes immune to a contagious disease. It is harder for a disease or virus to spread if sufficient people are protected against it. We have not reached herd Immunity for Covid yet.

5. Procedure:

Our program, written in Python, uses the CDC's COVID Deaths by state dataset. We used this to present the number of COVID related deaths by state in a bar graph (Graph 1). Using the matplotlib library, we made the graph and labeled it with the abbreviation of each state and "COVID-Related Deaths By State". This made it seem like there were an astounding amount of deaths in California, Texas, and Florida. After some experimentation, we researched the population of each state from the United States Census Bureau's website and made the number of deaths a percentage dividing by the population. We also calculated the death rates by multiplying the percentage by one hundred. Then, using the matplotlib library, another graph was made showing the ratio between the COVID related deaths in each state to the population of each state (Graph 2).

(Making of the first graph)

```
#Plot COVID Deaths for each state:
plt.figure(figsize=(11, 5), dpi=100)
plt.bar(x, deaths, width=0.8, color=colors)
plt.xticks(x, abbrStates, rotation='vertical')
plt.xlabel('States')
plt.ylabel('Deaths')
plt.title('COVID-Related Deaths by State')
plt.show()
```

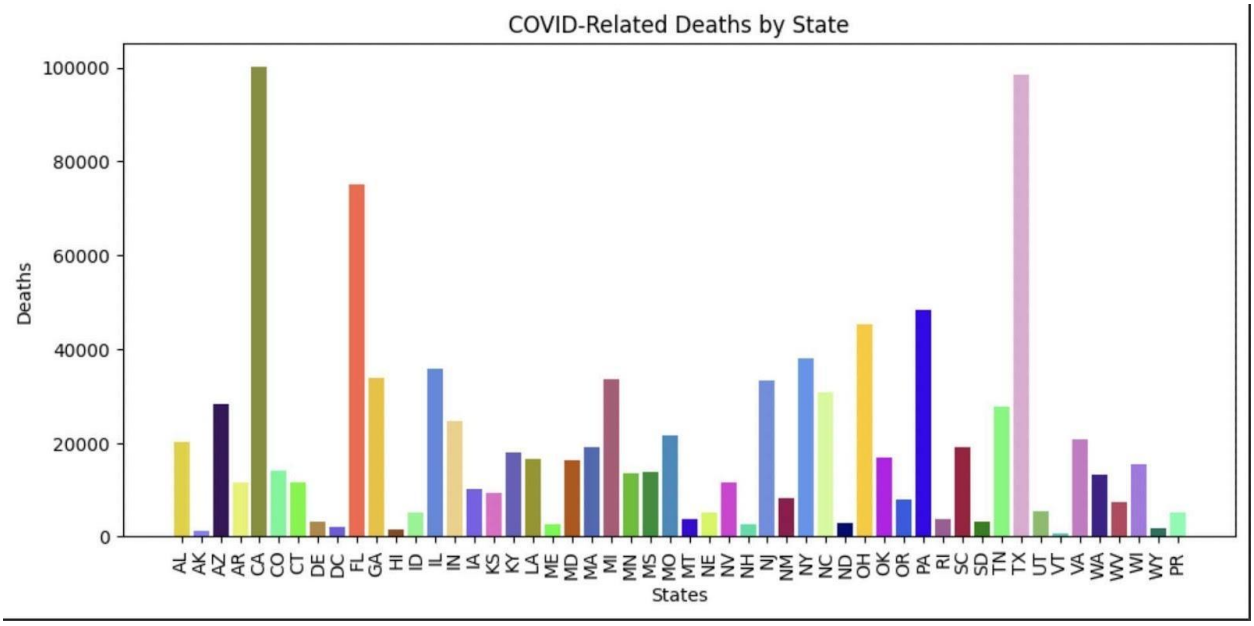
(Making of the second graph)

```
#Plot COVID Death Rates for each state:
plt.figure(figsize=(11, 5), dpi=100)
plt.bar(x, deathPercent, width=0.8, color=colors)
plt.xticks(x, abbrStates, rotation='vertical')
plt.xlabel('States')
plt.ylabel('Death Rates (in %)')
plt.title('COVID Death Rates by State')
plt.show()
```

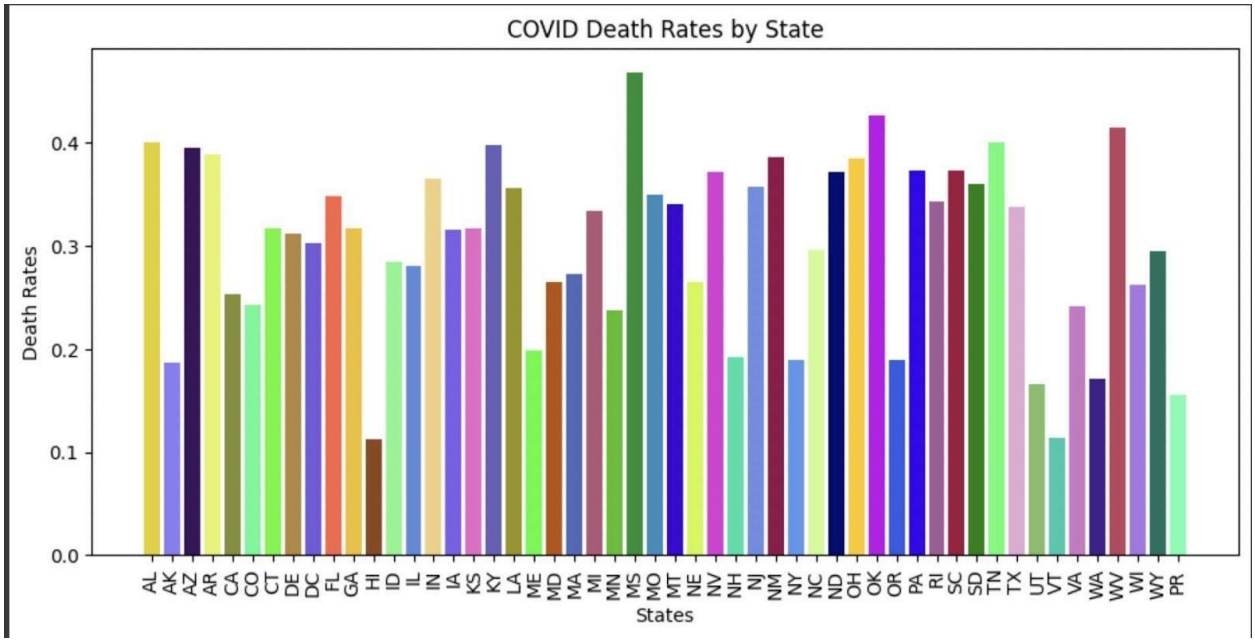
The reason for creating the second graph is because it seemed that some states were not taking strict measures and following certain guidelines set in place to prevent the spread of COVID-19, but it was only because some states have way higher populations than other states. This graph simplified the comparison of COVID death percentages of all the states accurately, comparing the states and their population. After this we began working on our scatter plot (Figure 3)

6. Results:

(Graph 1)

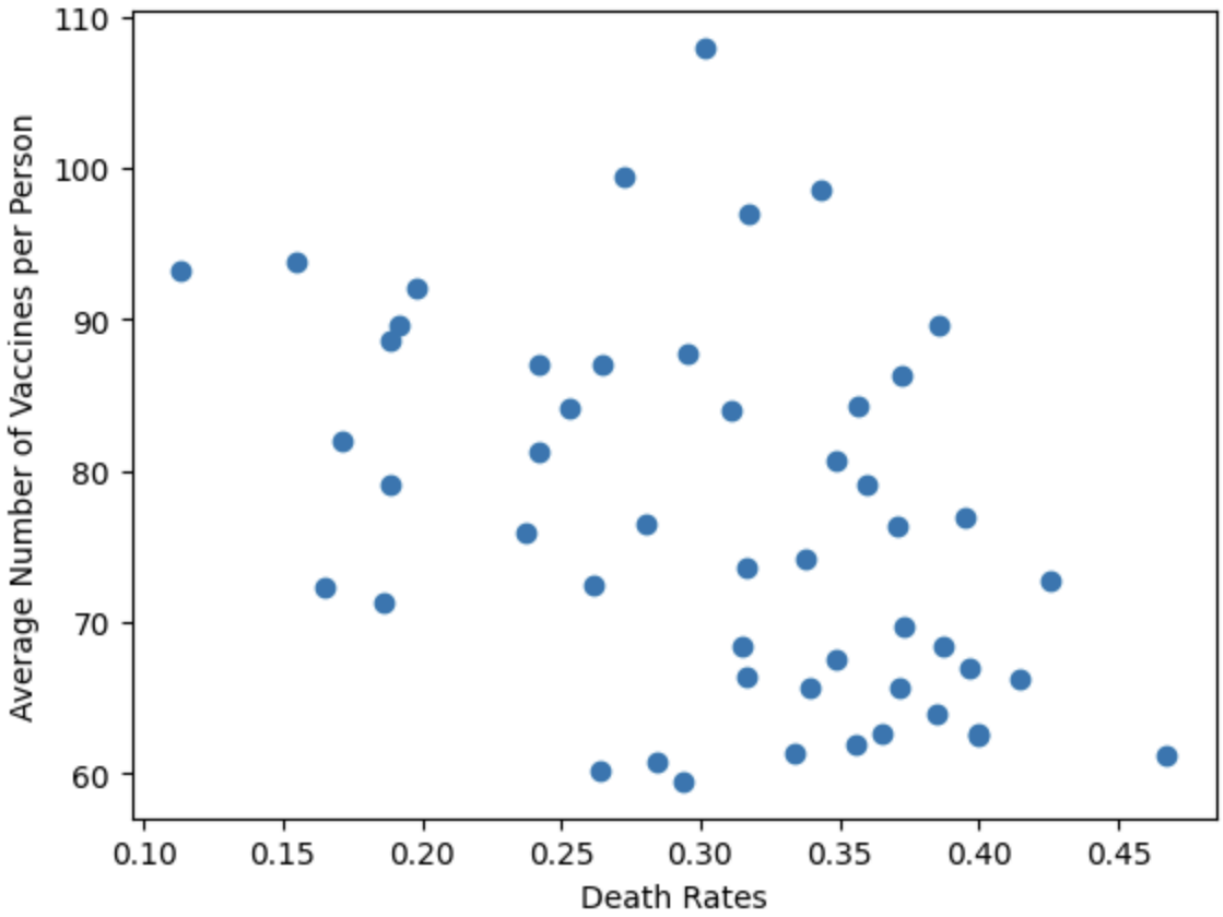


(Graph 2)



In order to evaluate and observe the effectiveness of the COVID vaccine, using the COVID-19 Vaccinations in the United States by County dataset (also from the CDC), we filtered the shown numbers of first dose vaccines administered per county and filtered those numbers to a specific date. Our prediction was that we would be able to notice a negative linear correlation between COVID related deaths and the number of administered vaccines, translating that COVID related deaths decrease as there is an increase of vaccinated people, which is what we saw However, the lack of data from some states made the results inaccurate so we cut them out. This led to the final version of our scatter plot. This scatter plot displays the percent of people with at least one dose by state and death rates.

(Figure 3)



To add more geographic information, we also used the geopandas library to create two different maps of the U.S. These maps are currently being made as a side-by-side comparison of each other, one shows the death rates and the other shows the vaccine percentage.

7. Encountered Issues:

Our project consisted of several datasets. This meant that some information was missing or might have been inaccurate. In the Vaccinations per County dataset there was no data for Hawaii. So we had to cut it out because it interfered with calculating the correlation between the two factors.

```
# cut out HI
x = x[0:11] + x[12:51]
states = states[0:11]+states[12:51]
abbrStates = abbrStates[0:11] + abbrStates[12:51]
deaths = deaths[0:11] + deaths[12:51]
number_ppl_1dose = number_ppl_1dose[0:11] + number_ppl_1dose[12:51]
population = population[0:11] + population[12:51]
```

In the State Population dataset from the United States Census Bureau's website, other areas were included such as New York City, Puerto Rico, and Washington D.C. However, New York City was not included in other sources, so we had to remove it as well. Luckily there was enough data to include Puerto Rico and D.C.

We also encountered an error when making our scatter plot. We received a ValueError, this was because the death rates and amount of vaccines administered were different size arguments. To fix this, we needed to make administered vaccines a percentage as well. We did this by multiplying by 100 and dividing by the population. This made our x and y the same size values.

8. Additional Research:

For more context on the project, we did research on the virus itself and how vaccines are made. There are six different types of vaccines, but the only three used for COVID vaccines are protein subunit vaccines, viral vector vaccines, and mRNA vaccines.

- Moderna and Pfizer-BioNTech COVID-19 vaccines are both messenger RNA vaccines, meaning that they work by giving your cells instructions on how to create proteins to trigger an immune response.
- The Novavax COVID-19 Vaccin is a Protein Subunit meaning it is made with just a small portion of a pathogen
- Johnson & Johnson's Janssen COVID-19 is a viral vector vaccine meaning it is made with a modified version of the pathogen

We also researched the Spanish Flu Pandemic of 1918 and found that an estimated 675,000 Americans died according to the Centers for Disease Control and Prevention over the course of two years. Which compares to the estimated 675,400 deaths from COVID as of September of 2021. Although the Spanish Flu Pandemic caused almost ten times the amount of deaths that Covid did worldwide.

9. Conclusions:

The main part of our project is the scatter plot because it shows the correlation between the death rates and amount of people receiving the first dose. The results of this showed a negative linear correlation of -0.42 which is a significant negative correlation. This means that states where more people were given the vaccine had significantly fewer deaths from Covid. Along with this, the standard deviation was 0.80458 and the covariance was -4.287. Our first map will be colored by the ratio of COVID related deaths in each state and each state's population. The second map will be colored by the effectiveness of the vaccine in each state. These maps help us compare the COVID related deaths before the vaccine and after the vaccine was invented. This can help us see the effectiveness of the vaccine in each state in order to find out whether or not the vaccine has a help to society and if people should be vaccinated against COVID 19.

```
correlation
[[ 1.          -0.42990339]
 [-0.42990339  1.        ]]

covariance
[[ 6.60306167e-03 -4.28747513e-01]
 [-4.28747513e-01  1.50631615e+02]]

standard deviation
0.0804586220976772 3

mean/average
0.3071057907425643 3
```

10. Achievements

Throughout this project we learned how to communicate and work better as a team. We also learned valuable skills and concepts that could help us in the future.

This project helped us to familiarize ourselves better with python. It allowed us to practice outside of the classroom and understand the hard work that goes into this programming environment. We were able to visually represent data in a way that is easy to understand.

11. Furthering Our Project:

Even after this project is complete, we plan to add more information and improve. Although there is now a vaccine for Covid, we haven't reached herd immunity in the United States. As Covid research advances we hope to stay informed and continue to further our research until we do reach herd immunity in the near future.

12. Acknowledgements:

- Irina Cislaru - Thank you for guiding us and encouraging us to push ourselves further into creating such an amazing project. You have pushed us to do our best and have helped us get ready for the real world with your guidance.
- Jon Brown - Thank you for being such an amazing tutor and for being a great help throughout our project. We learned so much from your coding tutorials and we appreciate your patience with us.
- Barbara Teterycz - Thank you for encouraging us to take on this challenge and to develop new skills socially and professionally.
- Mark Galassi - Thank you for all your feedback and suggestions. You have been a great advice provider and it helped us work effectively and professionally.
- The Supercomputing Challenge- for giving us this opportunity to show our coding skills and knowledge of Computer Science.

13. Works Cited:

Hart, R. (2021, September 5). *By the numbers: Who's refusing Covid vaccinations-and why*. Forbes. Retrieved January 7, 2023, from https://www.forbes.com/sites/roberthart/2021/09/05/by-the-numbers-whos-refusing-covid-vaccinations-and-why/?sh=2816fed052ea

World Health Organization. (n.d.). *Coronavirus disease (covid-19)*. World Health Organization. Retrieved November 28, 2022, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Article. Pfizer. (n.d.). Retrieved February 4, 2023, from https://www.pfizer.com/news/articles/understanding_six_types_of_vaccine_technologies

World Health Organization. (n.d.). *Coronavirus disease (covid-19)*. World Health Organization. Retrieved November 28, 2022, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Branswell, H. (2021, September 20). *Covid-19 overtakes 1918 Spanish flu as deadliest disease in American history*. STAT. Retrieved January 2, 2023, from

https://www.statnews.com/2021/09/20/covid-19-set-to-overtake-1918-spanish-flu-as-deadliest-disease-in-american-history/