

COVID-19 Related Visualization

Andy Fan

dufa3828@colorado.edu

Mengchen Gong

mego6698@colorado.edu

Qinglu Sun

qisu4484@colorado.edu

Yilin Gong

yigo7755@colorado.edu

ABSTRACT

In the previous approximate 400 days, we've been through the 2020 elections, COVID-19 Pandemic and we are still experiencing COVID-19 spreading in current days. It's been more than 1 year since the outbreak of COVID-19. From the outbreak to vaccines, it's been a long path, based on the data retrieved related to tweets hashtags, the highest ranking hashtags in tweets during a pandemic are a more negative category like: fear and sadness. We want to visualize the COVID-19 infection/hospitalization/death both worldwide and in the US to give us a clear view of the trend in the past year. We see that the US has a really high death rate among countries in the world. And for the death live data, infection live data, we can see that India in the past a couple weeks has a really high rate of infection and death. With the data about vaccinations started in 2021, we plotted the state-wise correlations between the election voting results to the vaccination receive result, there is a positive correlation between percentage of democrats in each state to the percentage of people who get at least one dose of vaccines. With the increasing vaccinated percentage worldwide, the majority of the countries' hospitalized rate has been decreasing.

1. INTRODUCTION

In the beginning state of the COVID19 epidemic, due to the lack of timely control and the high flow of people, a blowout outbreak began globally. It has had a serious and universal impact on the mental health of people around the world, with increased anxiety, anger, fear, depression and other stress reactions. The rise of infected persons and shortage of masks and medical resources in the news has even increased people's fear. For the second half of 2020, people begin to pay more attention to personal hygiene, wearing masks and anti-virus tissues when going out, and washing hands frequently at home, reducing the spread of the virus. At present, the vaccines are being distributed and vaccines are divided into inactivated vaccines, nucleic acid vaccines, recombinant protein vaccines and etc. In this project, we want to analyze 2020 election results, infection/confirmed/hospitalized/recovered cases results, and vaccination rates in all states. We want to compare the above 3 analysis and see how election results relate to the infection rate and vaccination rate. On top of that we would like to analyze the emotion of people worldwide, the difference between vaccines and their symptoms, impact of policies with a timeline with positive/death/vaccination rate.

2. METHODS

We found multiple datasets online, which some are static and some have been updated in real time. So some of the plots we have is based on the data we have right now. In general, we

mainly used Python, and it's libraries like Plotly, Altair, matplotlib, Pandas, numpy etc. In this project. We used different plots to visualize the trend or the correlations between different aspects like vaccine, hospitalized rate, sex, and 2020 election results. The main types we used in this study are choropleth maps (countries, states, counties), scatter plots, line plots, bar charts.

We mainly used 3 types of choropleth maps to show summaries of geographic characteristics within each area. First type is simple choropleth maps for countries, states or counties. We used worldwide maps to show country-wise death/infection trends. The US map is used for 1. State-wise when we need to compare it to election results by states; 2. County-wise for a more in-depth look at the death rate in each county due to potential factors like higher ratio of senior group, less access to any ventilator or hospital in general. The Second type is a combination of choropleth map and timeline slider, users can drag the bar to see the death, infection results and how the trend changes by date. The third type is a combination of choropleth maps and line charts showing that there are hospitalized trends for each state one the map. Line plots are used to represent trends for vaccine and hospitalized rate in the past a couple months as well. We also used bar charts to see the distributions of vaccine manufacturing worldwide, age groups and its vaccine type received, sex groups distributions of getting vaccine doses, and the distribution of side-effects after receiving the vaccine.

3. RELATED WORK

Since the outbreak of COVID-19 pandemic, people from all sorts of backgrounds have put lots of effort trying to track the severity of the pandemic happening around the world. In the US, from the earliest COVID-19 tracking visualization built by John Hopkins University[10], to the most recent one built by Centers for Disease Control and Prevention [2] and all other medias such as New York Times[16], CNN[9], and The Washington Post[17]. All of these provide detailed visualization about new COVID-19 cases, death, and hospitalization within the US. For visualization around the global, Our World in Data[14] and Worldmeater[18] provide detailed views country by country. These visualizations became one of the most critical one that people all around the world follow closely every day.

For works that combine both the COVID-19 cases and the US presidential election result, Brookings[6] made a few county-by-county maps to compare the spread of COVID-19 and its relation with the 2016 presidential election.

4. RESULTS

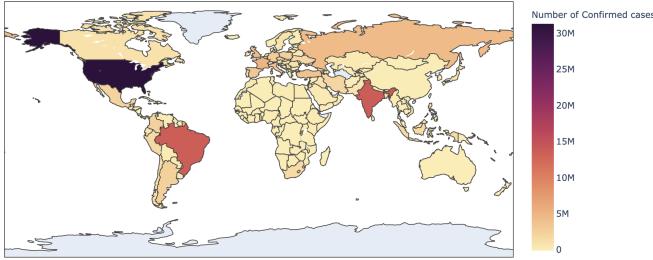
In this section, we will talk about the findings in COVID-19 infection/hospitalization/death rate results, vaccine-related results

with election & vaccine correlations, and emotions analysis worldwide.

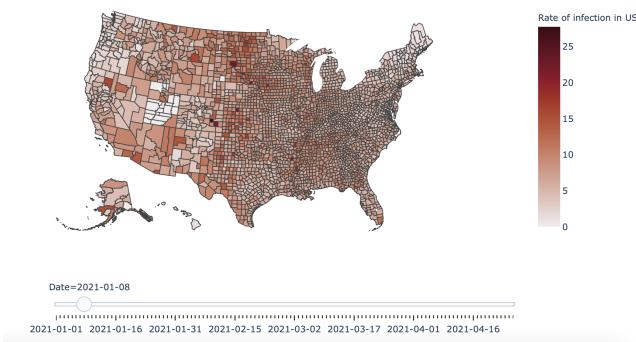
4.1. COVID-19 Infection/Confirmed cases

We visualized the number of confirmed cases for COVID-19 both worldwide and in the US with a slider for the timeline for 2021[4]. Worldwide confirmed cases provided a big picture of the COVID-19 infection view, and then an in-detailed US map which showed the trend of the infection rate in 2021.

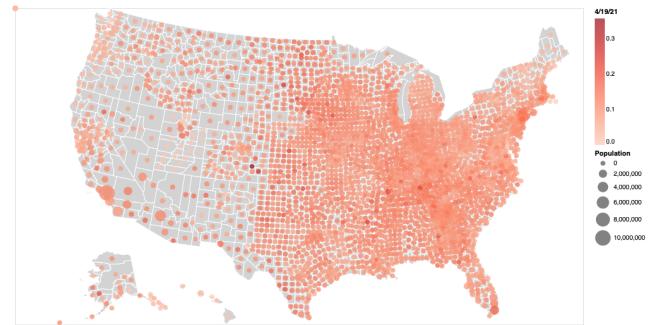
For the world trend, we can see that the deepest colored country, which means the highest total number of confirmed cases is the US. At the same time, India and Brazil also had a deeper color scheme shown in the map. For country India, it matches the recent news about the rapid increasing cases and lack of validator, oxygen availability.



Below is the US choropleth map with a slide bar for infection rates started in early January 2021 [5]. We drew this because the confirmed cases have overwhelmed largely in 2020, and it's been the most cases in the world. We visualized them by the rate of each county because we believed that only showing the number of cases is not helpful to show the risk level without any bias since some counties might have higher cases than others because of large populations. After gathering the number of confirmed cases in each county, we manually divided the cases number by population to get the infection rate in each county. From the interactive graph by adjusting the timeline, we can see that the infection rate decreased compared to early 2021.



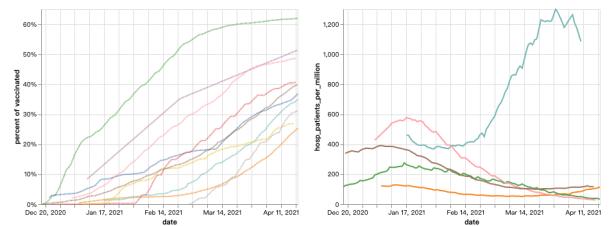
We noticed that the infection rate seems higher on the eastern side than the western side, so we visualized the points' geography by the latest days, trying to find the patterns between population and infection rate. .



On the visualization above, we could clearly see that the population didn't have a strong relationship with the infection rate, it might seem to have a relationship because of more crowded or dense counties on the eastern side. Although the U.S has been the most dangerous country in the world last year, from the trends, it might become the dream country again.

4.2. COVID-19 Hospitalizing

For the hospitalizing rate, we plotted a side by side trend line graph for vaccination and hospitalization. We also plot a choropleth map with state-wise hospitalizing trends for the US [3][8].



From the first visualization, we picked top 10 vaccinated countries across the globe. Based on these trend lines, we can clearly see that as the vaccination percentage rises, the hospitalization rate decreases (except for one outlier).



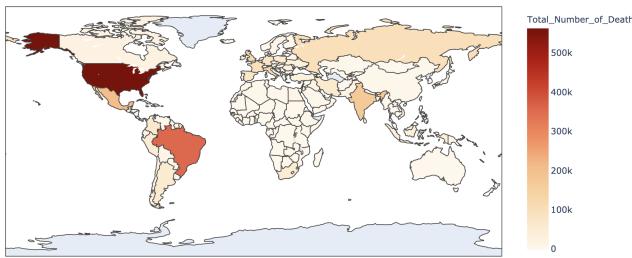
Above is a state-wise hospitalizing trend in the US. In order to show the trend of the rate of hospitalization, as well as preserving the geographical feature, we decided to use multiple trend lines to represent both the information. From this visualization, we can notice the speed of hospitalization slows down within the recent months.

Based on the above two visualizations, we can further confirm that taking COVID vaccine really does help in reducing the severity of the syndrome of COVID-19.

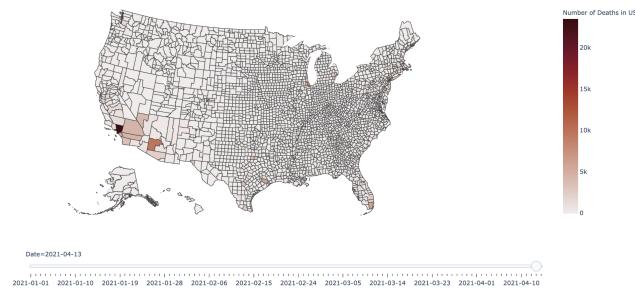
4.3. COVID-19 Death & Recover

We visualized the number of deaths, number of confirmed cases and number of recovered cases for COVID-19 both worldwide and in the US along with sliders for the timeline [4][5].

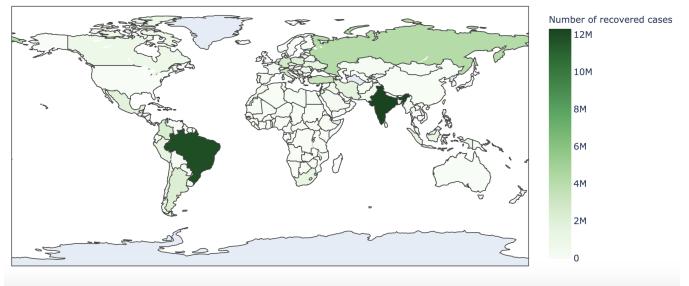
For the world-total death view, we see that the US, Brazil and India have high total numbers of deaths compared to other countries, which correspond to the world confirmed cases map in 4.1. Because there are in general a bigger number of confirmed cases, the number of deaths is also high.



For the visualizations in the US, we plotted the choropleth maps based on counties. The reason that we plot our visualizations in counties instead of states is because the number of confirmed cases of COVID-19 in almost all states in the United States has risen sharply, which has led to the possibility of the collapse of the US medical system. Although high prices have prevented many people from going to the hospital, the sharp increase in the number of patients will still raise the problem of medical resources shortage, especially in those counties at the higher risk of being overwhelmed with the majority of the residents being at high risk where being high risk is defined by health conditions and behaviors such as age(65 or older), diabetes, asthma, smoke cigarettes and etc. Because the computation time for visualizing the data from 04/2020-present is high, we also established functions where the input is the date, and output is the visualization of the number confirmed/recovered cases in US/worldwide.



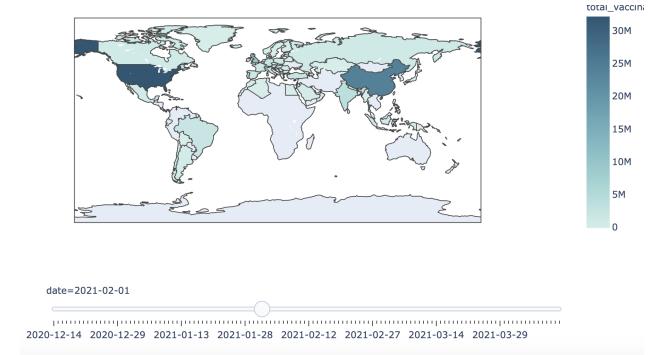
Below is the total number of recovering cases worldwide. Similar to the previous visualizations of the total number of confirmed/death cases, we see that Brazil and India have much more deeper color compared to other countries. The only difference is that we wasn't able to find data about recovered cases in the US due to the privacy issues.



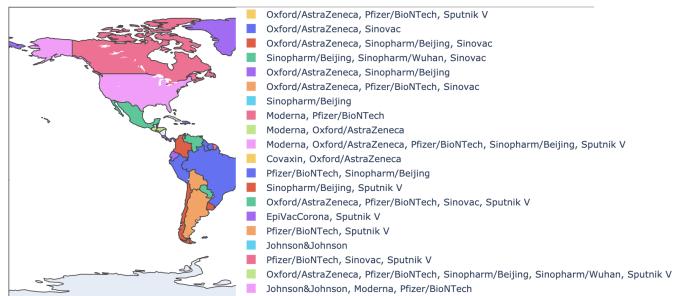
4.4. COVID-19 Vaccine & Election Result

We plotted the total number of vaccinations around the world, and we visualized what the different types of vaccines had been distributed in different countries.. In addition, we plotted histograms of vaccinations distribution by different age-groups; the total number of people vaccinated by sex; and side-effects distributions after vaccination. For the data in the US, we plotted the total number of people received at least one dose and vaccine type among all the doses for each state. Last but not the least, we did analysis about the relations between election and vaccine.

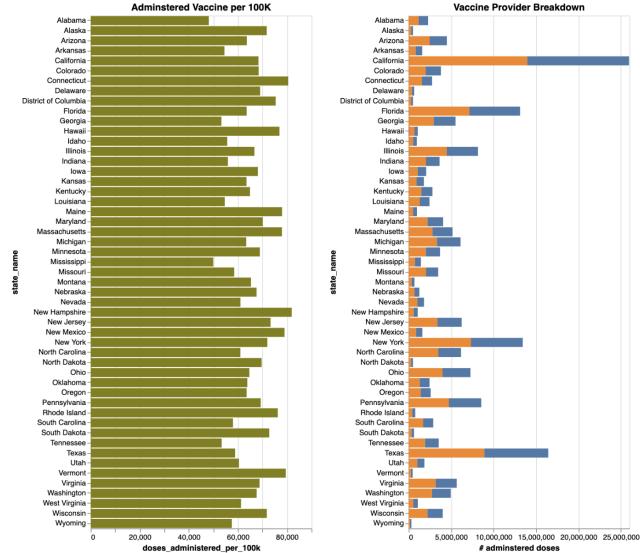
Below is the total number of countries' vaccinations around the world [8]. It is very obvious that the US and then China so far are the countries with the highest number of vaccinations among all the countries. However, because we don't have data for majority countries in Africa, middle-east and Australia, we don't know how many vaccines have been distributed.



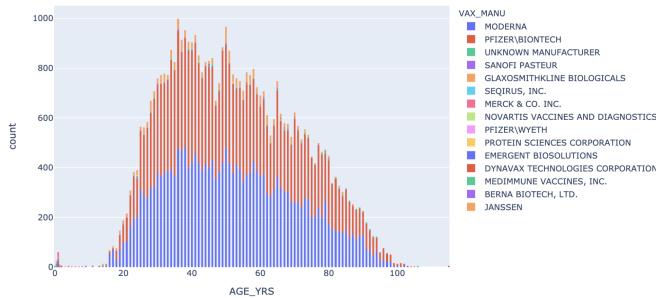
Below is the plot of vaccines manufactured used in different countries in America. We can see that Moderna, Pfizer and Johnson & Johnson are the most used vaccination types here in the US [13].



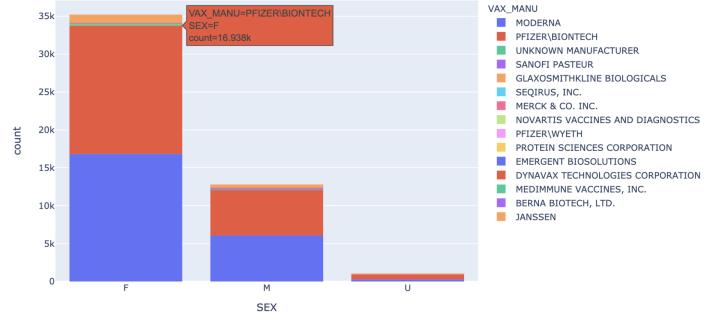
To take a closer look at the US vaccine distribution. We broke it down into states. Using side to side barcharts, the left one showed how percentage vaccines have been distributed per 100k, and by clicking on the bars on the left, a clear breakdown of the total number of vaccine manufacturers will be displayed on the right panel.



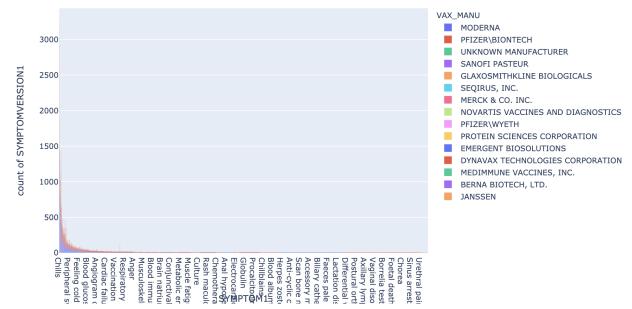
The histogram showed the distribution of types of vaccines among different age groups in the US. For vaccine, vaccine type distribution for all age groups, we can see that approximately age 30-60 has the highest total counts. We can see that for all the groups there is about a 50-50 vaccine type split between Modena and Pfizer\Biotech. The remaining vaccine types only contain a very small percentage for all age groups.



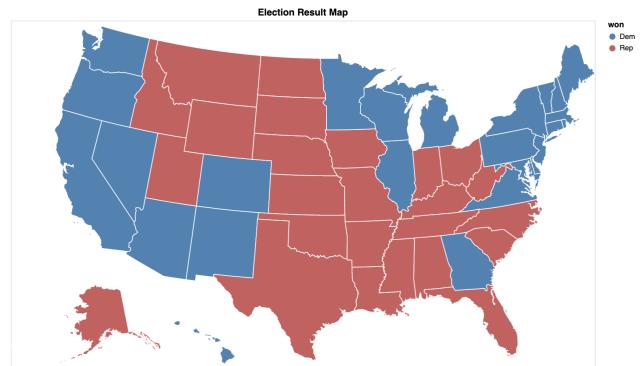
As for total number of vaccines received and sex groups are shown below. Here “U” stands for unknown, because it only contains a smaller amount of cases, we will focus on the female/male group analysis. We can see that more than half of the populations who received vaccines are Females. The total number of males who received the vaccine is less than half of the total number for females.



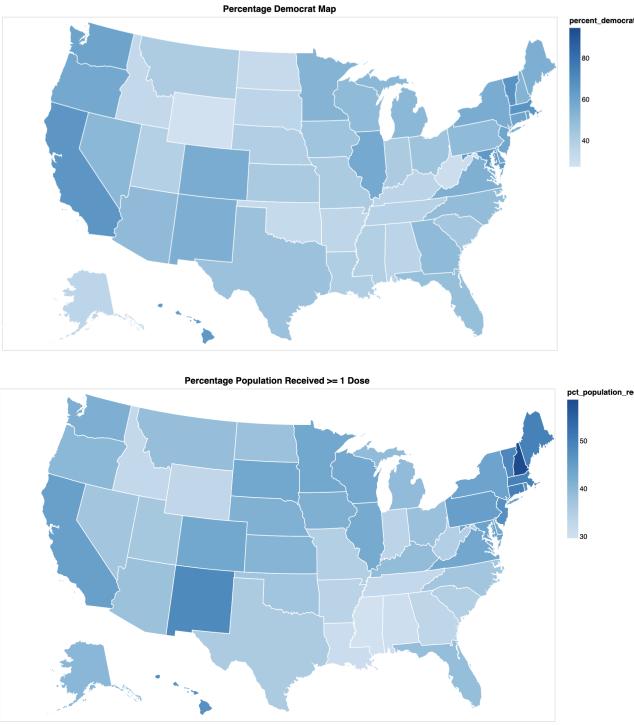
We plotted the count of different symptoms vaccine receivers reported in a decreasing order[7]. We can see that the decreasing trend is really rapid, and the majority of reported symptoms are light like chills, feeling abnormal, headache, tight/sore throat, arthralgia and so on.



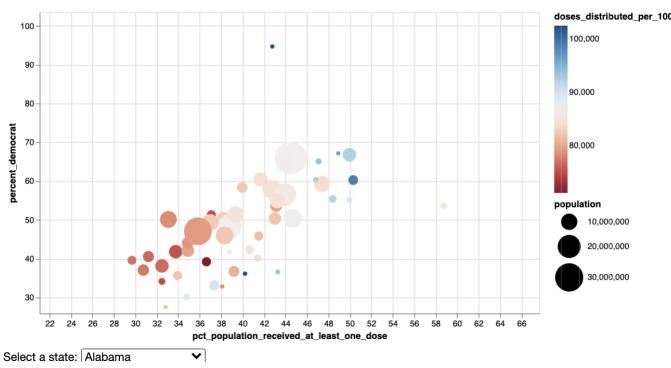
Last part of the vaccinations analysis, we did election-vaccine correlation visualizations. We want to see how elections impact the vaccination received/distributed in different states. For this section, all the US choropleth maps are colored based on states, because when we are comparing the 2020 election results, it is easier to see the correlations, color depth based on states. Here below is an overview of the 2020 election results[11].



To better show the choropleth map of 2020 election with vaccine distributions, we plotted percentage of democrats in each state, and also plotted percentage of population that receive at least 1 dose in each state[15]. Below, the top one is the percentage of democrats for the 2020 election and the bottom one is the percentage of the population that received at least 1 dose. Based on the plot, we can see that 2 plots are really similar to each other. Majority of the darker colors’ state in election results (darker = higher democrat%) have dark colors in the vaccine results (darker = higher vaccine dose(≥ 1) receiver%).



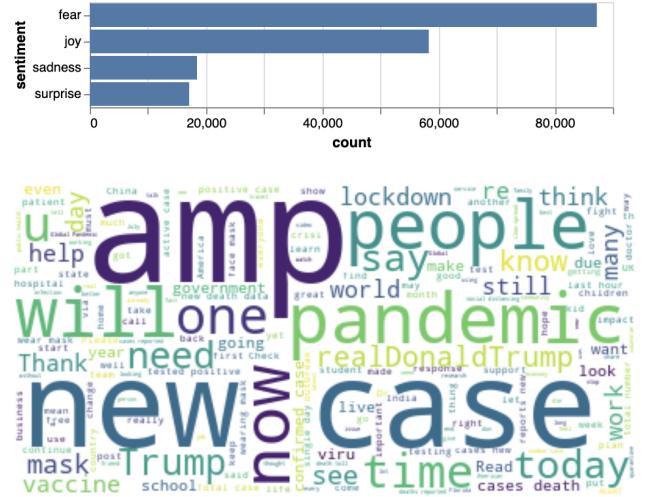
To see the correlation between election and vaccine better, we did a scatter plot. X-axis is the number of people who get a minimum of 1 dose, y-axis is percentage of democrat, color is number of doses distributed per 100k, and the size is each state's population. From the result, we can clearly see a positive correlation between the percentage of people who get ≥ 1 vaccine and percentage of democrat for each state. What we can also see from the graph is that if the distribution of doses per 100k is high (blue), it's more likely that the corresponding state has a higher percentage of people get at least 1 dose, at the same time, it's more likely those states have a higher percentage of democrat votes.



4.5. COVID-19 Emotions worldwide

We used COVID-19 tweet data with hashtags for the emotion visualization [12]. By using the tweets that have covid19 hashtag, we use spark and spark nlp for sentiment analysis using pretrained models of emotion classifier and negative/positive/neutral

classifier to analyze how people feel about this epidemic, visualizing the geolocation of people who are tweeting and generate the word cloud visualization for visually highlight the keywords that appear frequently in the tweets related to COVID-19. We can see that the highest count among the tweets related to COVID-19 is fear, followed by joy, sadness, and surprise. As for word cloud, we exclude word "COVID-19" itself, for the remaining words, we can see that people, new, pandemic, help, lockdown, mask, Trump were the words that have been commonly used.



5. DISCUSSION

World-wide trend of COVID-19 Infection/death/recovered have the similar characteristics that the US has the highest number of infection/death/vaccination cases. In other words, because the US has the highest number of confirmed cases, the counts of death are also up because of it.

Confirmed cases are still increasing nowadays all around the world. From the world map compared to a week ago, India's color has been deeper and deeper. Even though the color of it is not as deep as the US and Brazil, India's confirmed cases are increasing rapidly, which is corresponding to the recent news about the COVID-19 spreading situation in India. As for the US infection rate, according to the graph on 4.1, we observed that it didn't have a clear pattern of the changes in counties, but the growing rate has decreased in overview. In early 2021, the growing rate of confirmed cases became the zenith and rapid decreasing after that. The reason might be due to the vaccine released and enhanced the awareness of prevention.

From the vaccination trend and hospitalization trend comparison, we can see that the use of COVID-19 vaccines truly reduced the number of hospitalized cases, which proved the importance of receiving vaccines. The importance of vaccines is not preventing receivers from getting the virus, but to effectively reduce the likelihood of hospitalization and death. The state-by-state visualization further confirms the fact that vaccination does help people around the US get less significant syndrome.

For the recovered/death worldwide data result, it is related to the total number of confirmed cases. It means that if a country has a bigger number of people getting infected, with the same rate of

recovery/death, there will be more people recovering/passing away among the population that got infected. Because we only have the US death data in 2021, when dragging the slide bar there is an upsurge of death count in January. What likely caused the rapid growth is that back in January, it was still the first stage of vaccination, so not many people are vaccinated back at that time, also because of the holiday season, combined with the spring break for schools, there were a lot of travel happened, which increase the spreading trend and indirectly increased the death count after that. Above that after the announcement of the election result, there are a lot of protests going on in many states without proper mask wearing, which is another reason why the confirmed cases count and death counts have increased by a huge margin.

As the graph showed in 4.3, county-wise US choropleth map looks really unevenly distributed, we see that there are a small number of counties with a super high number of death count. There are 2 main potential causes of the uneven distribution based on the research [1]. First, those counties with a high death count have a higher ratio of senior citizens. Since the senior group is the most innocent group when it comes to fighting COVID-19, those counties have a super high death count. Second, uneven distribution of hospital and medical resources. Some of the counties, states only have a few hospitals with equipment like vadelators. When people living in the nearby counties need any COVID-19 related treatment, they need to go to the near-by equipped hospitals, which increases the counts in those counties with the medical services.

As for vaccinations, the US has been good at vaccine distributions and right now, vaccinations are available for all the public who are older than 16. There have been more vaccine availabilities around, and more and more people are getting their vaccines. Based on the results of the vaccine distribution and age groups vaccination counts for gender distributions about vaccine counts, what is interesting to see is that the majority of the people who received vaccines are females, which means that there is a higher percentage of females who will experience various side-effects compared to males.

Compared 2020 Elections results to vaccination data in different states in the US. There is a positive correlation between state's democrats percentage and state's percentage of population receiving ($>=1$) vaccines/ total number of distributed vaccines per 100k. The correlation makes sense. It is likely that for the democrat states, there are more policies to get a higher vaccine distribution and encouraging people to get the vaccines. By saying that it doesn't mean the governments of states with higher republican votes don't get as much as distribution of vaccines. It is more because residents of those states who might have lied and misled about the need/usage of vaccinations from the previous presidency, which still impact their choice of getting COVID-19 vaccines or not till now. Among democrats, there is more social sentiment to get the vaccination, and there is a higher acceptance rate of the COVID-19 vaccine.

REFERENCES

- [1] Butchireddygari, L. (2020, April 22). *How One High-Risk Community In Rural South Carolina Is Bracing For COVID-19*. FiveThirtyEight. <https://fivethirtyeight.com/features/how-one-high-risk-community-in-rural-south-carolina-is-bracing-for-covid-19/>
- [2] Centers for Disease Control and Prevention. (n.d.). *CDC COVID Data Tracker*. Centers for Disease Control and Prevention. Retrieved April 29, 2021 from <https://covid.cdc.gov/covid-data-tracker/#datatracker-home>
- [3] Centers for Disease Control and Prevention. (n.d.). *COVID-NET Laboratory-confirmed COVID-19 hospitalizations*. Centers for Disease Control and Prevention. Retrieved April 29, 2021 from <https://covid.cdc.gov/covid-data-tracker/#covidnet-hospitalization-network>
- [4] Centers for Disease Control and Prevention. (n.d.). *Trends in Number of COVID-19 Cases and Deaths in the US Reported to CDC, by State/Territory*. Centers for Disease Control and Prevention. Retrieved April 29, 2021 from https://covid.cdc.gov/covid-data-tracker/#trends_dailytrends_cases
- [5] Dong, E., Du, H., & Gardner, L. (2020). *An interactive web-based dashboard to track COVID-19 in real time*. The Lancet. Infectious diseases, 20(5), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- [6] Frey, W. H. (2020, June 3). *Mapping COVID-19's spread from blue to red America*. Brookings. <https://www.brookings.edu/blog/the-avenue/2020/05/29/mapping-covid-19s-spread-from-blue-to-red-america/>
- [7] Grag, A. (2021, March). *COVID-19 World Vaccine Adverse Reactions, version 4*. Retrieved April 29, 2021 from <https://www.kaggle.com/ayushggarg/covid19-vaccine-adverse-reactions>
- [8] Hasell, J., Mathieu, E., Beltekian, D. et al. *A cross-country database of COVID-19 testing*. Sci Data 7, 345 (2020). <https://doi.org/10.1038/s41597-020-00688-8>
- [9] Hernandez, S., Manley, B., O'Key, S., & Pettersson, H. (2021). *Tracking Covid-19 cases in the US*. CNN. <https://edition.cnn.com/interactive/2020/health/coronavirus-us-maps-and-cases/>
- [10] Johns Hopkins University & Medicine. (n.d.) Coronavirus Resource Center. <https://coronavirus.jhu.edu/>
- [11] Mooney, P. (2020, October). *2020 USA Election: Vote Percentages by State*. Kaggle. <https://www.kaggle.com/paultimothymooney/2020-usa-election-vote-percentages-by-state>
- [12] Preda, G. (2020, August). *COVID19 Tweets, version 24*. Retrieved April 29, 2021 from <https://www.kaggle.com/gpreda/covid19-tweets>
- [13] Preda, G. (2021, April). *COVID-19 World Vaccination Progress, version 110*. Retrieved April 29, 2021 from

- https://www.kaggle.com/gpreda/covid-world-vaccination-progress?select=country_vaccinations_by_manufacturer.csv
- [14] Roser, M., Ritchie, H., Ortiz-Ospina, E., & Hasell, J. (2020) *Coronavirus Pandemic (COVID-19)*. OurWorldInData.org.
<https://ourworldindata.org/coronavirus>
- [15] The Associated Press. (2021, February). *CDC COVID-19 Vaccine Tracker*. Retrieved April 29, 2021 from
<https://data.world/associatedpress/cdc-covid-19-vaccine-tracker>
- [16] The New York Times. (2021). *Coronavirus in the U.S.: Latest Map and Case Count*. The New York Times.
<https://www.nytimes.com/interactive/2021/us/covid-cases.html>
- [17] The Washington Post (2021). *U.S. coronavirus cases and state maps: Tracking cases, deaths*. The Washington Post.
<https://www.washingtonpost.com/graphics/2020/national/coronavirus-us-cases-deaths/>
- [18] Worldometer. (n.d.). *COVID-19 CORONAVIRUS PANDEMIC*. Worldmeter. Retrieved April 29, 2021 from
<https://www.worldometers.info/coronavirus/#countries>