Alessandro Hammond and Paige Defiori

# Group Project Proposal

## -An introduction of your research question

Our research question pertains to the current COVID-19 pandemic and its spread throughout the globe. Specifically, we are seeking to identify the correlation between the socioeconomic status of each country, the rate of infectection, and the rate of vaccination. What relationship exists between an individual's country of residence and their ability to receive the vaccine? What factors are considered when determining which countries will receive the vaccine? What social inequities are resulting from the distribution of the COVID-19 vaccine?

## -An explanation of why it is important to you, why it matters to others, and what is at stake

When COVID-19 first began to spread, most assumed that it would only last about two weeks before we were able to overcome the disease. However, over a year later, we are still seeing the devastating impact of the disease's widespread effects. Each of us in the group has witnessed the trauma of COVID-19 firsthand as we have seen friends and family diagnosed with the virus. Never before have we seen a virus so severely impact our economy, health, governments, and more, as we will likely see fallout from this year for the rest of our lives. While the medical community is focused on finding cures, vaccines, and treatments, other members of the community must seek to analyze the social and political issues that will arise as a result. If our society fails to probe these issues more deeply now while they are still relatively minor, we may see continued exacerbation of the problems for years to come.

## -A description of the spatial scope (e.g. Boyle Heights or Hong Kong), and why space and/or time matters for your project

As we are looking to track the correlation between the rate of infection and the rate of vaccination in specific countries based on their socioeconomic status, we plan on tackling a space scope of the entire world broken down by country and continent, which will therefore allow us to analyze the underlying social inequities of global medical care. Furthermore, since the spread and vaccine of the COVID-19 is a relatively new phenomenon in human history, we will be focusing on the entire timeline of the disease as recorded in the below data sets. This will specifically begin at the first recorded case from the first data set, which is January 20th, 2020, and run until today, which will allow us to thoroughly analyze the entire development of COVID-19.

## -A preliminary but definitive description of data sources (at least two) that you will use Include datasource with links

https://docs.google.com/spreadsheets/d/1jS24DjSPVWa4iuxuD4OAXrE3QeI8c9BC1hSlqr-NMiU/edit#gid=1187587451
This data set demonstrates the progression of COVID-19 beginning in January 2020 throughout the world. The first sheet specifically details information about many of the early cases, including the individual's country of residence, age, gender, symptoms, hospitalization status, and how they believe to have been exposed to COVID-19. In addition, this data set specifically emphasizes the impact of coronavirus in the Chinese provinces, detailing the daily and total cases in each individual province.

https://github.com/owid/covid-19-data/blob/master/public/data/vaccinations/locations.csv
This data set focuses on tracking worldwide vaccination trends by country, specifically identifying which vaccine each country is receiving and from what source, beginning in the middle of January 2020. However, this data set only includes countries that have already received the vaccine, making it incomplete of underserved countries.

https://ourworldindata.org/covid-vaccinations
This data set tracks the rate of vaccination in all countries that have received it already, beginning in December 2020 and tracking the total number of people who have been vaccinated. This data is available in multiple formats, including a line chart, a heat map, and an alphabetized table that also tracks daily absolute change, making the information incredibly simple to consume and track. In addition, because of the map, it is easy to see which countries have not yet received any vaccinations, making it very complimentary to the above data sets.

COVID Vaccinations - Our World in Data
The datasource above is COVID-19 cumulative vaccines distributed (recorded by administration per 100 people). Notably, this may not equal the total number of people vaccinated, depending on the specific dose regime (i.e. if people receive multiple doses).
Worldometer Live COVID Updates
Above is a datasource on 221 countries COVID-19 cases, including: total cases, new cases, tests, total deaths, new deaths, population number infected and more. This data with the vaccination data from the first source can be contrasted (cases vs. vaccine distribution).
Additional sources:
DXY.cn COVID records
Vaccine Locations
All off the data sources above on the COVID-19 pandemic are useful for our research on the pandemic and vaccines, and they are meta-analyses formed from data given from each country.

**-A scope that explains the intended analysis and resulting visualizations for your project**

By comparing the initial spread of the disease as well as the individual rate of vaccination in each country, we will be able to analyze the policies and social inequities involved in COVID-19 politics. In order to best visualize this data, we are going to produce two complimentary global heat maps in order to thoroughly compare and contrast the pace of these two events. The first

heat map, drawn from the first data set, will allow us to analyze the speed of infection in different parts of the world. The second heat map, which will use a combination of data from the second and third data sets, to compare the speed of global vaccination. By layering these two maps, we hope to successfully analyze the discrepancies between countries that were most infected by the virus with those receiving the most support from the global medical community, therefore highlighting the social inequities that result from the pandemic. With the data from the sources above, we intend to compare the rates of covid occurring globally and contrast those numbers with the distribution of the vaccines. For visualizations, a choropleth map will show the general view of covid-19 cases throughout the world. For a more specific view of our data, a line or bar graph of will show a contrast between the COVID-19 cases by country and their respective vaccine distribution. The data sets can be compared on the same graph or two separate graphs, side-by-side, depending on depth/details analyzed from the data.

**-A concluding paragraph of what insights you expect to gain from your research**

From our basic understanding of the COVID-19 pandemic, it is already apparent that the countries who were most severely impacted by the coronavirus were those with little medical resources, and therefore will likely be the last to receive vaccinations and treatments. As a result, these countries will likely have the most challenging recovery from the coronavirus socially, economically, politically, and medically, as they are being denied basic healthcare from simply based on their socioeconomic status. By creating our two heat maps, we hope to prove this theory, which will consequently highlight the need of many underserved countries as they still struggle through this pandemic. As a result, this research will empower the medical community to better prepare for the inequitable fallout if there is another major pandemic, allowing us to learn from our mistakes to create a better future.