DSE I2700 - Visual Analytics

(Rahul Chandani, Nikhita Kannam, Jose Salas)

Project 1: COVID-19 Vaccine Efficacy Dashboard.

Objective

The objective of this project is to find a closer estimation of the true protection percentage of a country considering the number of vaccines, its efficacy, the vaccine type composition of a given country and the percentage of its vaccinated population.

Introduction

Despite different COVID-19 Vaccine being available for over 2 years, hesitancy about the vaccine's ability to protect our immune system against serious disease still lingers. Due to the lack of understanding from the general population on what vaccines are and how they work, it is often misunderstood that vaccines guarantee an 100% protection from any given disease. While it is true that no vaccine offers a 100% protection, there have been vaccines that are highly effective such as the smallpox vaccine or polio vaccine, which needs three doses to offer a 99% to 100% effectiveness. Additionally, some major news outlets tend to report only on vaccine numbers and equate them linearly proportional to its efficacy, even if unintended. It's often assumed that if 95% of the population is vaccinated, then the protection rate is of the same magnitude; however, this is not the case. The average conversation of covid vaccine tends to overlook the efficacy of the vaccine and the vaccine type used in each country. Countries with different vaccines and different vaccine composition will end up having different efficacy overall.

Assumptions:

- Vaccine efficacy levels reported taken at face value.
- Vaccine provided replicates the same efficacy despite people's specific health conditions.
- Level of protection based on the number of vaccine doses administrated.

Unknowns:

- Number of vaccines given to each person.

DSE I2700 - Visual Analytics

(Rahul Chandani, Nikhita Kannam, Jose Salas)

Data Processing and Cleaning

The original data frame used for the analysis of the countries true efficacy includes the country, date, vaccine manufacturer and the cumulative total vaccinations. In addition to this data frame, a second data frame that includes the efficacy of different vaccine types with respect to different variants in terms of severe disease and infection was used to complement the original data frame. These two data sets were merged in order to have the percentage protection for each vaccine type in each country.

Theoretically, if there was only one vaccine being administrated in a country, the efficacy of this vaccine would be considered the protection percentage against this disease for the country administrating this vaccine. However, since there are different vaccine types being administrated in different countries, the percentage protection of a country is a combination of the efficacy of all the vaccines being administrated with respect to the proportion of this vaccine in such country. For this reason, it was important to the percentage distribution of the vaccines being used in the studied countries, and new column labeled 'Vaccine % in each country' to store this value was created. This vaccine percentage distribution was calculated by dividing the total vaccinations corresponding to each vaccine type in each country by the total number of vaccines in such given country. Once this number was calculated, it was used to calculate the vaccine efficacy the countries. For example, given the vaccine type Moderna, the protection provided by the efficacy of the Moderna vaccine against Omicron Infections was calculated by multiplying the Moderna efficacy against Omicron infections times the percentage that Moderna represents of vaccine administrated in a given country. This same procedure was followed to calculate the protection provided by the efficacy of all the vaccines against Omicron Infections used a country, as well as for different cases, severe disease, for different variants like the Alpha, and Delta variants. Once the true protection against infections was obtained, it was possible to calculated the possible breakthroughs, 1 – country's true protection, and new columns were added to store these values.

After the country's true efficacy or protection was calculated, it was important to know the number of people vaccinated in each country to have a better understanding of the number of people truly protected in a given country. The efficacy/protection calculated in the previous section only account for the people that was vaccinated; however, each country has different percentages of their population vaccinated. Thus, in order to have a more accurate account of the protection

against the disease of a given country, it was important to bring unvaccinated people into the picture. Since the percentage protection of each country was already calculated, the next step was to multiply this by the number of people vaccinated; and then divide it by the country's population. For this reason, the 'Country_Vaccinations1.csv' dataset that contain the total people vaccinated for each country was used, as well as the 'worlds_population.csv' dataset to get the total population. These two columns were combined with our main data frame. The number of people protected against a certain variant was calculated by multiplying the efficacy per vaccine with the total number of people vaccinated. For example, in order to get the number of people protected against infection for the omicron variant we multiply the omicron efficacy per vaccine for infection with the people vaccinated. To get the percentage of people protected, we divided the number of people protected with the total population; by doing this, it was possible to obtain the percentage protection of a country that includes vaccinated and unvaccinated people and depict a more accurate situation for each country in the continuous fight against COVID-19.

Conclusion.

The level of protection doesn't only depend on the number of vaccines distributed but also the proportion of the vaccine in a country multiplied by its efficacy. Because this is what contributes to the countries level of protection. Additionally, a part of already known factors like age, sex and underlying health conditions, the proportion of the vaccines distributed in each country and the vaccines' efficacy plays an important roles when discussing the level of protection against Covid-19.