The Predicted Effect of a COVID Vaccine

Prepared by: Henry Luong & Matt Dannheisser

Summary:

Our goal for this project is to predict how the future spread of COVID-19 will be affected by a vaccine given that individuals in certain countries have reported varying willingness to take a vaccine. Our model will provide predictions by country using daily contagion data, IPSOS research reporting a country's likeliness of taking a vaccine, population data, and reported vaccine efficacy reported by the pharmaceutical companies.

Data Sources:

Name:	COVID-19 Spread	Global Attitude on Vaccine	Country Population
Description:	Daily new case of COVID by country. This will confirm total and daily cases all provided by CSSE at John Hopkins University.	Sampled likelihood to take COVID-19 vaccine based on a survey research by the World Economic Forum.	The total population for each country based on the latest United Nations Population Division Estimates
Size:	10,655 Kb	680 Kb	
Location:	Our World in Data	Statista	Population by Country (2020) - Worldometer
Format:	CSV	PDF	HTML
Access Method:	Direct access	Direct access	Web scrape

What initial processing will have to be done on each?

- COVID-19 Spread: The daily spread values will be loaded directly from a CSV into a pandas data frame. Noisy and missing data will be cleaned through Python which creates a clearer model. The daily spread data will then be grouped by each country.
- Global Attitudes on Vaccine: Percentages will have to be extracted from the image as text and converted to a pandas data frame of floats.
- Country Population: The data will be scraped from the website and read directly into a pandas data frame.

How will you combine the datasets, and what will be produced as output?

Initially, we would clean all datasets and group the data according to their countries with the pandas groupby function. Utilizing the pandas library again, the data sets will be combined using the *pd.merge* function using the *Country* column as the key.

What new information will result from combining them?

Combining the COVID-19 Spread data with the Country Population data will enable us to determine the contagion rate and the number of susceptible individuals within each respective country. The Global Attitudes on Vaccine data will be combined with the Country Population to create a figure for the countries' demand for a vaccine. The supply of vaccines will only be realized in real-time, so this will be a variable that will be adjustable on the visual itself. Finally, we will combine the demand and supply data with the contagion rate and a number of susceptible citizens to predict the future spread of the virus.

Description of one resulting visual:

The model will build a curve demonstrating recent COVID cases and will have a division marking when a vaccine is reported which will also be the beginning of forecasted figures. The visual will allow for adjusting the global supply of the vaccine, and it will plot the supply and demand lines for the vaccine. This visual will be faceted by country. We will create an interactive dashboard that can demonstrate these adjustments for each country while also viewing the curve.

Contributions:

	Wk2	Wk3	Wk4	Wk5	Wk6
Matt	Data exploration for determining predictive model	Cont. data manipulation and final decision on predictive model	Create visualizations of historic trend	Integrating visuals into the dashboard	Analyzing the trends and dashboard and summarizing the report
Henry	Data manipulation	Cont. data manipulation and final decision on predictive model	Building the interactive dashboard	Integrating visuals into the dashboard	Analyzing the trends and dashboard and summarizing the report