<u>Final Project – Project Idea</u>

Team Members:
Ruoyi Gao
Yi Sun
Yixiao Zhou
Ziyue Xia

Title: Covid-19 Global Vaccine Tracker

Abstract: We are interested in finding out how countries are rolling out vaccination programs. Making use of the dataset from Our World in Data project at the University of Oxford together with NLP analysis of twitter data, we expect to explore the following question:

- Worldwide-level: How does vaccination rate differ from country to country? We would like to explore more detailed questions such as percentage of each country's population who received at least one dose and who have been fully vaccinated.
- Worldwide-level: How fast are countries vaccinating? Which country administered most doses in the past 7 days? Which country administered most doses throughout the Covid-19 timeline?
- Worldwide-level: How does vaccination doses relate to country variations in health care infrastructure, say, it is expected that countries with higher GDP per capita will have more people getting vaccinated (in per 100K people). We would also look at indicators like region and mortality rate.
- Country-level: How does Covid-19 affect each country? We can also ask more detailed questions such as infection rate, death rate, vaccines offered and vaccination phase.
- Country-level: Does the public's attitudes and experiences with Covid-19 vaccinations causes low uptake in some countries? For example, it is expected that spread of safety concerns of vaccine amongst Twitter users (e.g., #covidsideeffect) would increase public vaccine hesitancy. We determined three possible barriers: confidence (i.e., safety of the vaccine), complacency (i.e., the perceived need for the vaccine), and convenience (i.e., how accessible the vaccine is to people).[1] What would be specific barrier to vaccination in each country?

Techniques: ggplot2, ggmap, plotly, shiny, ThreeJS, NLP text mining

Data Description:
- Vaccination Dataset: https://github.com/owid/covid-19-data/tree/master/public/data/vaccinations#country-metadata
    - Country-level time series data on COVID-19 vaccinations

---

[1] "Covid-19 Vaccine Hesitancy: Analysing Twitter to Identify Barriers to Vaccination in a Low Uptake Region of the UK." Frontiers in Digital Health 3 (2022). https://doi.org/10.3389/fdgth.2021.804855.

- o Country-level time series data on percent of population given at least one dose of vaccine
- o Country-level time series data on percent of population fully vaccinated
- Country Dataset: https://data.world/
  - o Country-level GDP per capita
  - o Country-level region
- Twitter API:
  - o We plan to analyze sentiments towards vaccination from twitter data, whether it is mentioned alongside positive or negative words.

Visualizations: