Final Project Proposal

Team members

- Brian Tung, Zhou Mingyue, Angelique Iradukunda

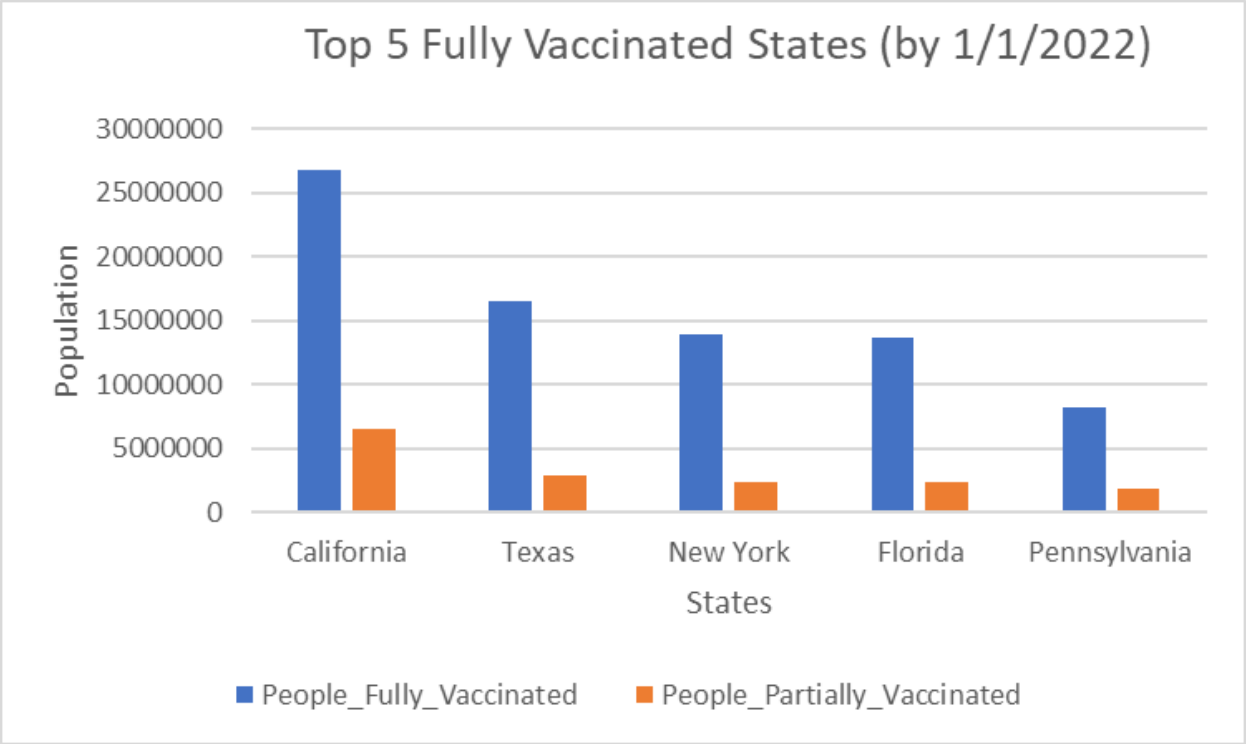
GitHub repository

- Project2_Zhou_Tung_Iradukunda

Primary dataset for analysis

- vaccine_data

Initial plots



Data structure

- Column name, Definition
- FIPS, U.S. State identification code
- Province_State, Name of the State
- Country_Region, Code of the country (US)
- Date, Data collection date
- Vaccine_Type, "Common name of the vaccine provider. Can be either a combination of all vaccine types labeled as 'All', or a specific provider like Moderna or Pfizer"

- Doses_alloc, Cumulative number of doses allocated
- Doses_shipped, Cumulative number of doses that have arrived at the vaccination sites.
- Doses_admin, Cumulative number of doses administered, including booster doses for states where it is reported as part of the total.
- Stage_One_Doses, Cumulative number of first doses administered
- Stage_Two_Doses, Cumulative number of second doses administered
- Combined_Key, "Combination of Province_State, Country_Region"

We anticipate that **FIPS**, **Doses_admin**, **Stage_One_Doses**, and **Stage_Two_Doses** will be useful in analyzing vaccination trends.

Supplemental datasets

- Covid 19 Cases

What we plan to cover in the final report

1. Vaccine effectiveness by manufacturer
    a. Is there a manufacturer that we can recommend for the optimal protection vs COVID?
2. Vaccine allocation and distribution by state
3. Vaccination tendencies in varying geographic locations
    a. Which regions have higher tendencies to get vaccinated?
4. Case and death rates vs number of doses administered
    a. By state, what is the percentage of people who are vaccinated (This we will need the population data per state)