

# Final Project Proposal

Zhouxin Shi

For my final project, I will study the number of cases and deaths of covid-19 in various states in the United States, and visualize the CDC Moderna covid-19-vaccine-distribution-by-state, CDC Pfizer covid-19-vaccine-distribution-by-state and other data. Next, I will study the relationship between vaccination in the United States and the number of covid-19 cases and deaths, In order to find out whether vaccination can significantly reduce the number of covid-19 cases and deaths, as well as the differences in the number of covid-19 cases and deaths caused by vaccines from different manufacturers.

First I would like to intuitively show the number of covid-19 cases and deaths in U.S States. Then, I would show the vaccination situation in U.S States. Most of the data can be accurate to the daily data of U.S States. I would also show a diagram of the relationship between covid-19 cases, deaths and vaccination.

I will decide what data to use and not to use when starting the project. As the project continues, more data may be added, such as the population of each state in the United States, so as to facilitate the study of the proportion of cases, mortality and vaccination rate of covid-19 in the United States, since this is what I interested in study during this period of time.

The data will come from various sources:

- Amazon Web Services (AWS) data lake for analysis of COVID-19 data(<https://datascience.nih.gov/covid-19-open-access-resources>),Root folder(<https://dj2taa9i652rf.cloudfront.net/>)
1. <https://covid19-lake.s3.us-east-2.amazonaws.com/cdc-moderna-vaccine-distribution/csv/cdc-moderna-covid-19-vaccine-distribution-by-state.csv>
  2. <https://covid19-lake.s3.us-east-2.amazonaws.com/cdc-pfizer-vaccine-distribution/csv/cdc-pfizer-covid-19-vaccine-distribution-by-state.csv>
  3. <https://covid19-lake.s3.us-east-2.amazonaws.com/rearc-covid-19-nyt-data-in-usa/csv/us-states/us-states.csv>