Final Project Proposal

Yiru Fei

**Project Question**

I plan to explore the distribution of new COVID-19 cases in special cities with high population density, and try to find the relationship between the number of patients and other conditions, such as, age, sex.

**Data Description**

I will use the data from HARVED Data verse, COIVD-19 Data Hub and time series data from JHU. And the data includes

* ZIP code
* Total positive cases
* Total tested number
* Number of new cases each day
* Number of hospitalizations each day
* Number of deaths each day
* Number of recovered each day
* Rates by age
* Rates by sex
* Population
* Date of government mandated social distance

**Analysis methods**

First, according to the daily counting number of new cases, deaths, and hospitalization, I plan to explore the distribution by using the MCMC method and try to figure out the maximum number of new cases in one day.

Then I want to compare the result by using time series and try to figure out whether age, sex, and population affect the number of infections and the probability of being infected.