Final Project Proposal

Yiru Fei

**Project Question**

I plan to explore the distribution of new COVID-19 cases in New York and Massachusetts, and try to try to calculate the number of secondary infections produced by a single infection.

**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
* Number of new cases each day
* Number of hospitalizations each day
* Number of deaths each day
* Number of recovered each day
* 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, and try to predict the number of new cases in the next few days.

Secondary, based on daily new cases, I will calculate the number of secondary infections produced by a single infection by using Bayes’ Theorem.

Finally, concluded the differences and similarities between New York and Massachusetts.