**Data Science And Engineering (INFO-6105)**

FINAL PROJECT

**DRAFT - 3**

**Team Name** : Maesters

**Team Members**:-

Abhishek Dabas ( 001302629 )

Apurva Mathur ( 001088100 )

Akshata Hosamani ( 001023074 )

**Topic** : Health and Epidemics

Identify the Spread, effect, symptoms of Covid-19 in United States

“Do you Have COVID-19?”

Brief Explanation :

Studies have shown that it takes an average of five years and a half billion dollars to discover and optimize antibodies in a lab. In case of any epidemic, the detection stage starts with the first case and ends with the first intervention activities (e.g. patient isolation, contact tracing, vaccination).

The AIM is to stop the chain of transmission, all health workers in the United States have to do is get the people possibly infected by the sick person into isolation(Quarantine) before these people show signs of Covid-19.

Basically when, R0 drops to zero, the United States is free of viruses.

In this project our team will try to understand the surge in the United States, which is on the top list of cases. As It is an important part of preparedness for a pandemic. We will try to build algorithms implementing Bayesian Statistics, Hypothesis tests, Linear Algebra as well as different Machine Learning Models to accelerate the identification of people affected with Covid-19, based on the travel History, symptoms, spread rate(R0) in the area and history of the disease which can identify the people and help them take effective measures. It could be used in future worst healthcare epidemic situations as well. We will be using our knowledge of Data Science to slow the spread of the infection which is nearly as important as stopping it. It can also help people to be prepared for the precautions of the epidemic and help them take the important measures to stop the spread. Our algorithms will calculate the chances of a person having Positive Covid-19. Everyone could know their chances of being positive for Covid-19, as well as prevent themselves if tested negative with a range defined within which they have to be in to prevent themselves from getting affected, according to the cases in the region. This is ultimately the whole purpose of Quarantine, to stay in a small range as to which there is very less contact with people.

Analyse and Predict :-

1. Spread of the virus/ how contagious is Covid-19 (R-naught)

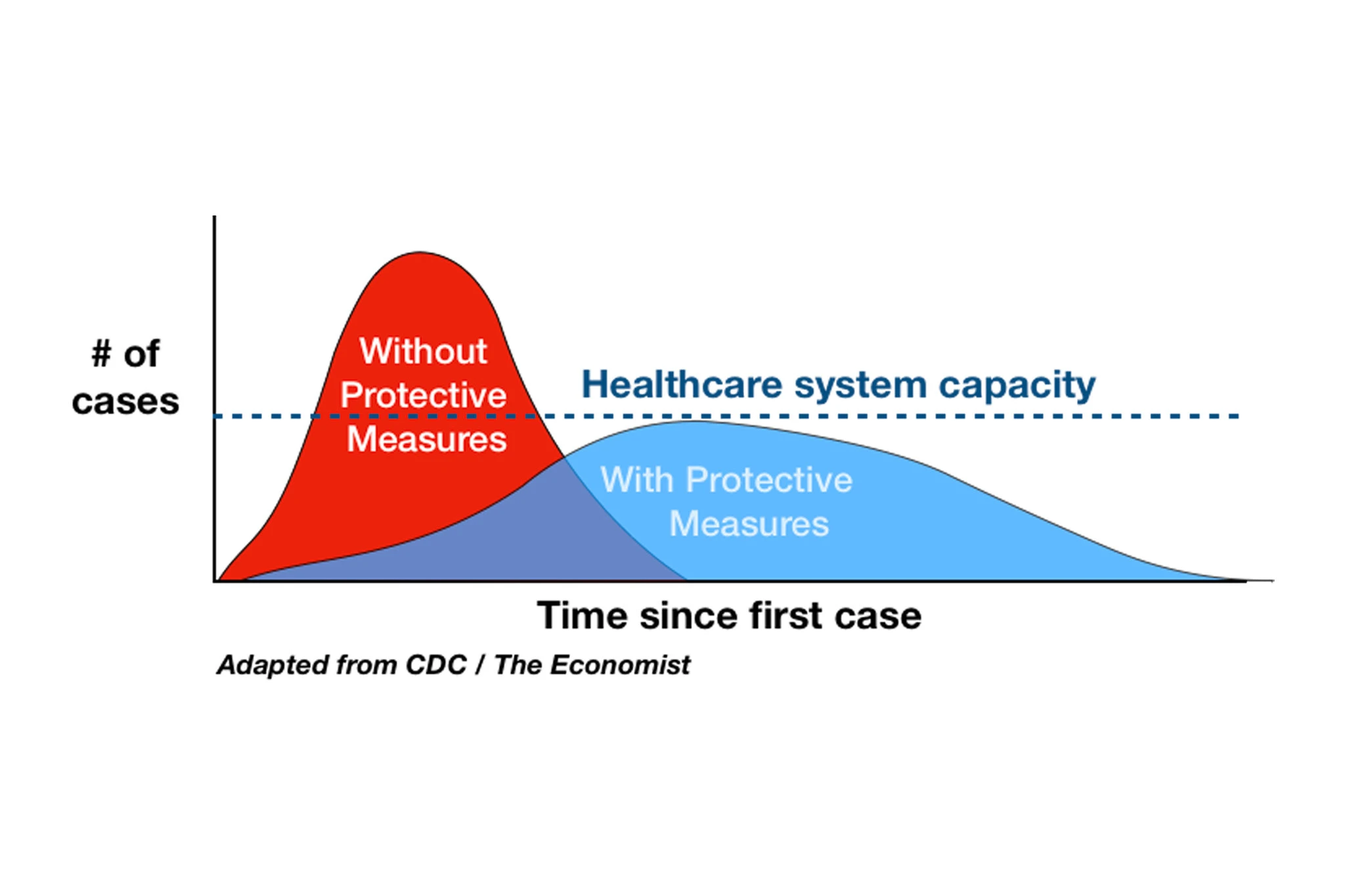
2. Effect of spread on different counties in USA

3. Identify people with CoronaVirus, and specific Group of High Risk Patients ( People with respiratory complications, lung problems, diabetes etc)

“*Flatten the Curve*”-

The goal of this is to reduce the epidemic and halt the spread. Trying to reduce the number of cases that are going on actively at that point of time, which in turn gives the doctors and govt to be prepared and respond to the situation.

The more people reporting with the virus on a given day, the higher the curve; a high curve means the virus is spreading fast. A low curve shows that the virus is spreading slower



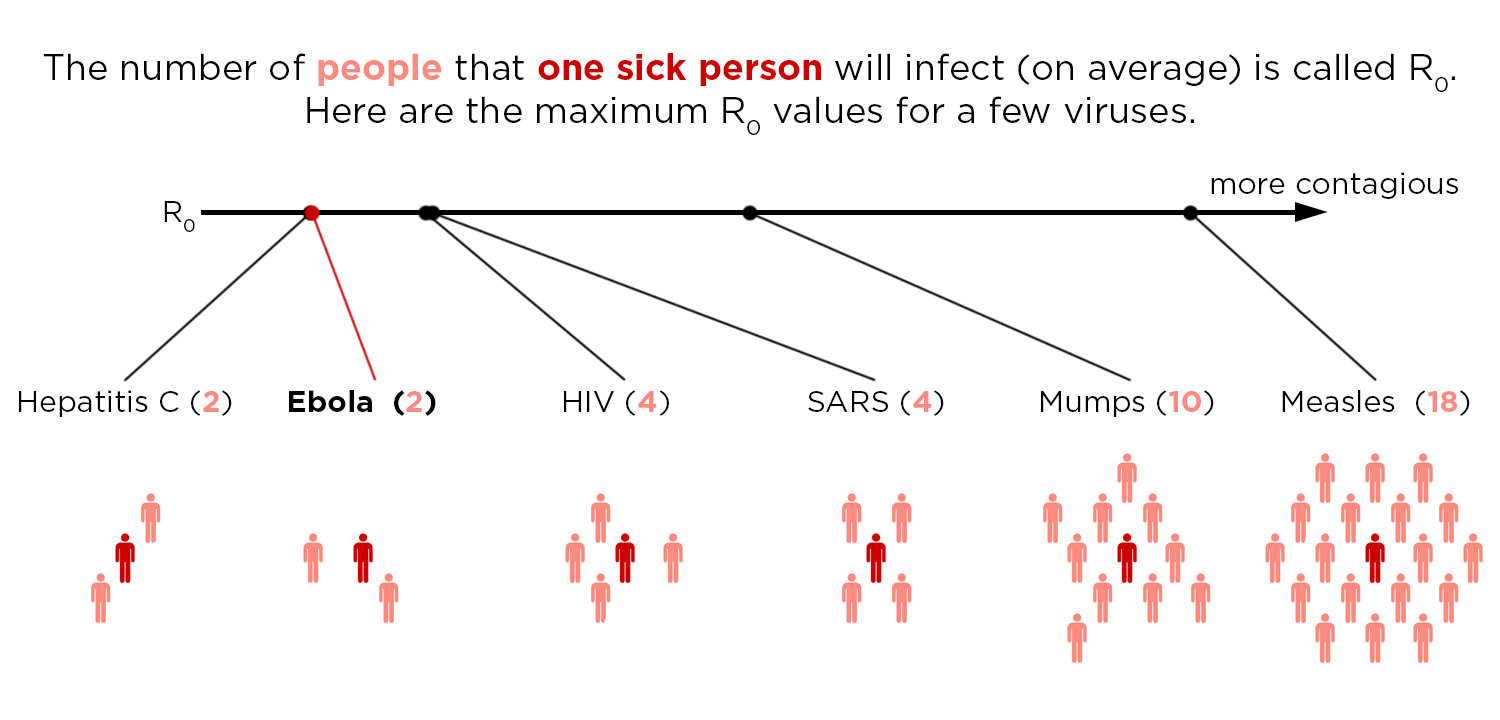
Models used:-

1. SIR Model

<http://mat.uab.cat/matmat/PDFv2013/v2013n03.pdf>

What every researcher wants to know is how far/how vast can the virus spread

R naught:-



The basic reproduction number (R0) is the average number of secondary cases that arise when one primary case is introduced into an uninfected population. These secondary cases arise after a period measured by the serial interval or by the generation time. When R0 is greater than 1, infection may spread in the population, and the rate of spread is higher with increasingly high values of R0. The doubling time (the time required for the incidence to double) was estimated on the basis of the reproduction number and the serial interval.11 After the early phase of exponential growth in case numbers, once infection has become established, the number of people still at risk declines, so the reproduction number falls from its maximum value of R0 to a smaller, net reproduction number, Rt. When Rt falls below 1, infection cannot be sustained. Estimates of R0 and Rt help in evaluating the magnitude of the effort required to control the disease, the way in which transmission rates have fluctuated through time, and the effectiveness of control measures as they are implemented. We estimated Rt over time from the time series of incidence of cases (i.e., a plot of the number of new cases per week over the course of the epidemic) and from our estimate of the serial interval distribution.12 We then estimated R0 for the early stages of the epidemic, when transmission rates were at their highest, on the basis of the date of symptom onset. As described in Supplementary Appendix 1, average estimates of Rt for the period from July 28 to September 7, 2014, which were made on the basis of the date of report to facilitate comparison with future cases, were used to project future cases, allowing for both uncertainty in the estimates of Rt and stochastic variability in the transmission process.

**Take, for example, measles.** The virus is one of the most contagious diseases known to man. It's R0 sits around 18. That means each person with the measles spreads it to 18 people, on average, when nobody is vaccinated. (When everyone is vaccinated, the R0 drops to essentially zero for measles).

**What we plan to do in coming week**:-

1. Read data directly from online( as there are hourly/daily update of cases in the country)
2. Find the R naught for COVID-19
3. Figure Out different parameters required to analyse the range of every person in a county
4. Figure out major symptoms for covid-19 as per data

Reference-

1.Malaria Outbreak Prediction

<http://ijarcet.org/wp-content/uploads/IJARCET-VOL-4-ISSUE-12-4415-4419.pdf>

2.<https://www.npr.org/sections/health-shots/2014/10/02/352983774/no-seriously-how-contagious-is-ebola>

<https://www.nejm.org/doi/full/10.1056/NEJMoa1411100>

3. <https://www.visualcapitalist.com/history-of-pandemics-deadliest/>

4.<https://www.bbc.com/future/article/20200325-covid-19-the-history-of-pandemics>

<https://healthitanalytics.com/news/data-scientists-use-machine-learning-to-discover-covid-19-treatments>

5.<https://www.nytimes.com/article/flatten-curve-coronavirus.html>

6.<https://royalsocietypublishing.org/doi/full/10.1098/rstb.2018.0276#d3e1193>