Re-Creating COVID-19 Analysis (Project 23)

P Omkar Ashrit(CS1914)
Amit Kumar Jaiswar(CS1932)

July 2020

Contents

1 Introduction

COVID-19 pandemic is a global health emergency caused by Severe Acute Respiratory Syndrome Corona virus 2 (SARS-CoV-2). In this study we have put together the timeline for different aspects of the pandemic growth in Indian perspective.This study also aims to evaluate whether simulation models might be proficiently used to anticipate a possible increase of SARS-Cov-2 spread.

2 Analysis

2.1 Data Collection

"COVID-19 India Org Data Operations Group" has open sourced their database for research purpose. We have used their API to collect the raw data for our interest. We have managed to perform some analysis which gives us the overview of spread of the pandemic in India.

2.2 COVID Spread in India

We have plotted the timeline for daily cases and cumulative cases in India.For all the states of India, we have the total number of infected, recovered, deceases cases till date. We have the timeline for each state and their daily and cumulative cases. We also have number of tests performed in India and in each of it's states for each day.

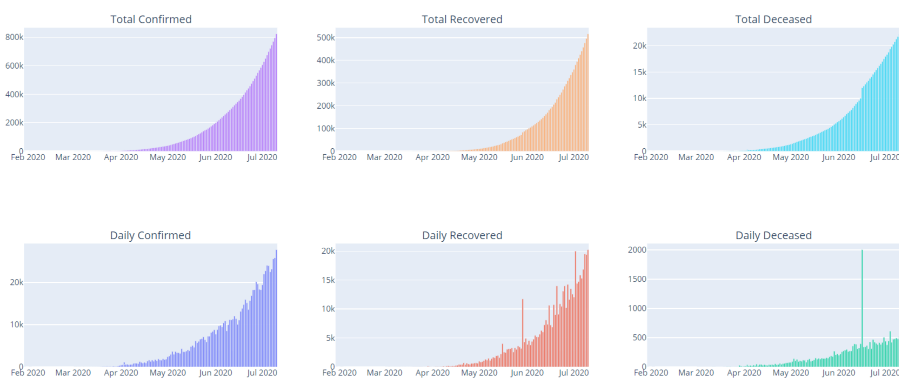

Figure 1: India Timeline

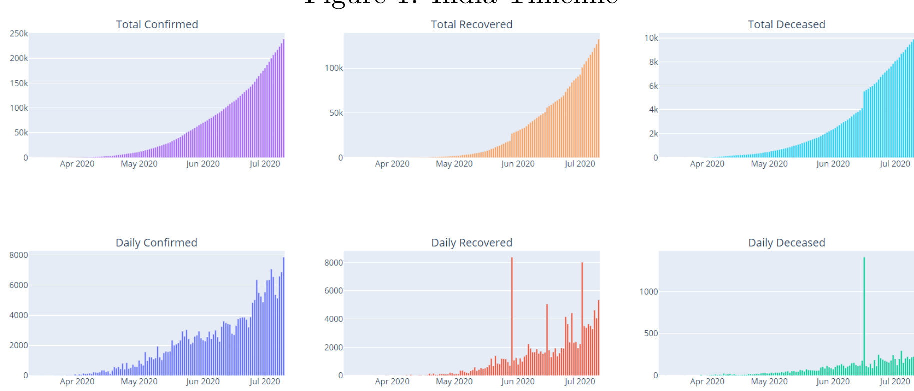

Figure 2: State Timeline(Maharashtra)

2.3 India vs Other Countries

When we see daily news reports on COVID-19, it's really difficult to build a coherent picture of what's actually going on, because the numbers are changing so quickly (which is exactly what you expect with exponential growth) that they're almost immediately out of date. We know that epidemics tend to grow exponentially at first, and also that exponential growth is really really hard for our human brains to understand because of just how crazily fast it is. So we have used the following ideas for this plot.

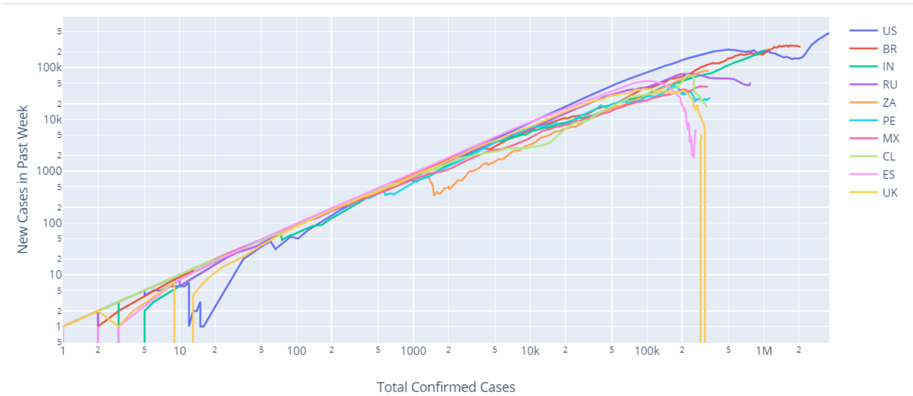

Figure 3: Top 10 Countries

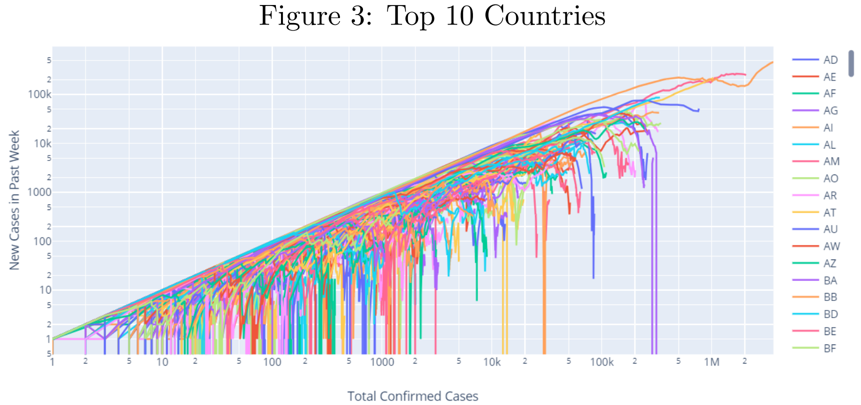

Figure 4: Top 50 Countries

- The first is to plot on a logarithmic scale, since that's the natural scale for exponential growth. This scales up small numbers and scales down large numbers, making the growth equally apparent on all scales, and lets us compare the growth in countries with very different numbers of cases.
- The second idea is to catch changes early, by looking at change itself. For example, if we look at the growth of cases in South Korea, we can see that at first they're exponential, and later, the growth slows down. But when we're halfway up this curve, it's hard to tell by eye that it's slowing down as it still looks exponential. When the number of new cases each week flattens out or goes down, we have escaped the exponential growth zone.
- The third idea is not to plot against time. Usually, when we see exponential growth, the

number of cases is plotted versus time. But the spread of the disease does not care if it's March or April; it only cares about two things: how many cases there are, and how many new cases there will be - that is, the growth rate. The defining feature of exponential growth is that the number of new cases is proportional to the number of existing cases, which means that if you plot new cases vs total cases, exponential growth appears as a straight line.

So these are what we plotted on our graph: the number of new cases (aka the growth rate) is on the y axis, and the cumulative number of cases is on the x axis, both on logarithmic scales.

2.3.1 Observation

This gives us a beautiful-horrible graph that shows where all countries are in their COVID-19 journeys. It makes it obvious that the disease is spreading in the same manner everywhere. We're all headed on the same trajectory, just shifted in time and it makes it obvious where public health measures like testing, isolation, social distancing, and contact tracing have started to beat back the disease, and where they either are not working or haven't had time to show up in the data. In nearly every country (*so far), the number of cases grows at a roughly similar rate, until it does not.

2.3.2 Limitations of the Plot

- The main goal of the plot is to emphasize deviations from exponential growth - that is, to amplify the light at the end of the tunnel, so it may be less informative for other purposes.
- Logarithmic scales distort : 10,000 looks really close to 1,000 on a log scale; this kind of distortion might allow people to take COVID-19 less seriously.
- Another important limitation is that this graph (basically every other COVID-19 graph that we have seen) is not actually showing the true number of cases, just the number of detected cases.

2.4 Insights

With all the data available, we have calculated several insights viz. Confirmed cases per million population, Active rate, Recovery rate, Mortality rate, Avg. growth rate, Number of tests per million, Percentage of positive cases in total tests performed, Total cases vs new cases.

3 Simulations

We have discussed 3 simulations to model how an pandemic spreads. In each of the simulation the total population has been divided into 3 categories.

- Infected : Red colour in the simulation
- Recovered : Green color in the simulation. We take green as either recovered or decease or quarantined. In either case the green particles are unable to spread the disease.
- Susceptible : Black color in the simulation

In these simulations, for every unit of time a susceptible person crosses path with someone with the disease, they'll have some probability of contracting it themselves. So we're using physical collision as a stand-in for things like shaking hands, touching the same surfaces, sneezing on each other. Then for each infectious person, after a given period of time, they'll recover and no longer be able to spread the disease.

We are not epidemiologists, so we would be very hesitant to generalize any of the observations from the simulations without deeper consideration.

The link to the simulations website : `https://omkar1610.github.io/COVID-19/`

3.1 Simulation 1

In the first simulation, everyone just moves around the city, and the infection follows the rules we have laid out. The parameters of the simulation are described below.

- Radius : This parameter changes the size of each particle.
- Speed : This parameter controls the speed at which particles move randomly in the city.
- Number of Samples : This is the total no of particles in the simulation.
- Initial Infection : This is the percentage of people infected initially before starting the simulation.
- Day/Sec : This represents 1 second in real time is equivalent to how many days in the simulation.
- Infection Days : This represents number of days each infected particle takes for recovery(or decease).

3.2 Simulation 2

In most towns, people don't actually spend their days randomly wandering through the city like the previous simulation. Often there's a common destination, like a central market or a school, that people need to go to. In this simulation, we introduce a central spot that people regularly visit then return from. The parameters of the simulation are described below.

- Radius : This parameter changes the size of each particle.
- Speed : This parameter controls the speed at which particles move randomly in the city.

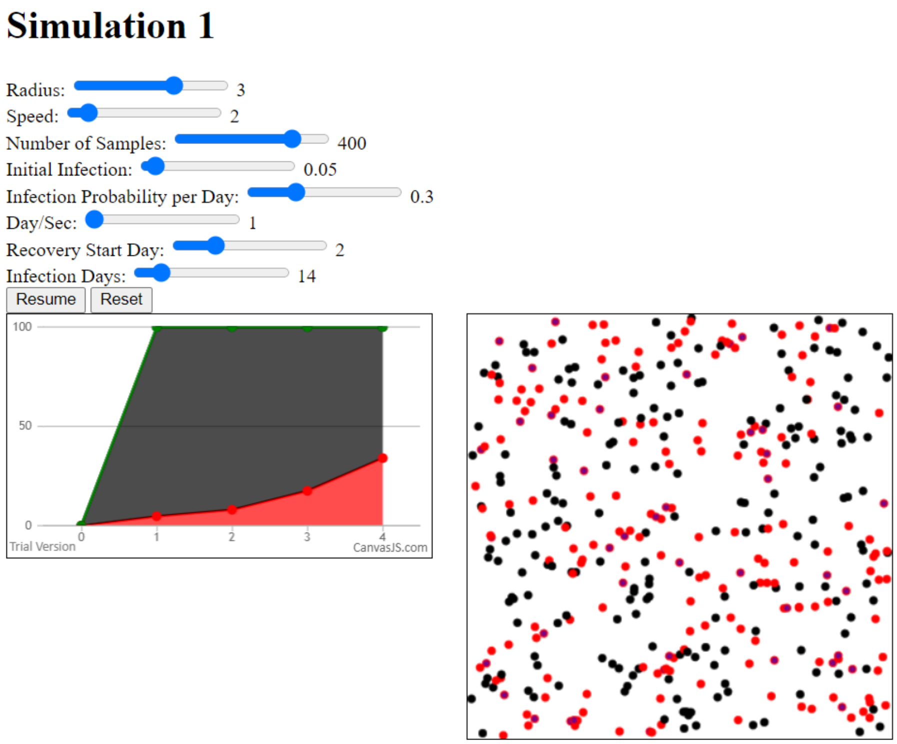

Figure 5: Simulation 1 : Starting

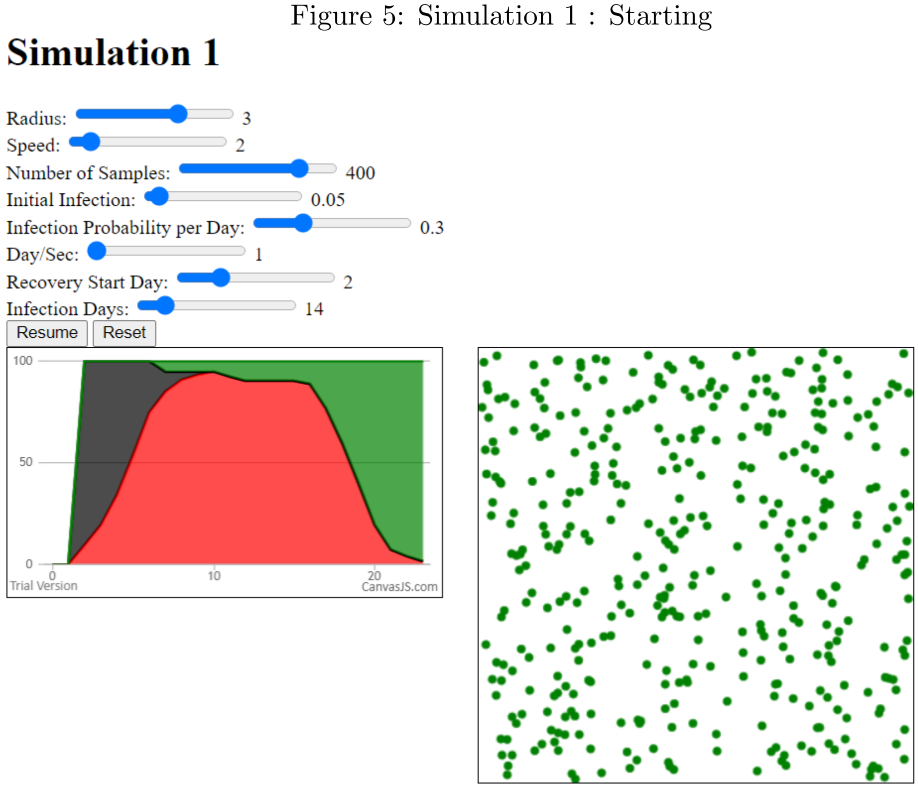

Figure 6: Simulation 1 : Finished

- Number of Samples : This is the total no of particles in the simulation.
- Initial Infection : This is the percentage of people infected initially before starting the simulation.
- Day/Sec : This represents 1 second in real time is equivalent to how many days in the simulation.
- Infection Days : This represents number of days each infected particle takes for recovery(or decease).
- Visit Probability : This is the probability of people going to the central location.

3.3 Simulation 3

In this simulation, we have a few separate communities, with travel between them. Every day, each person will have some probability of traveling to another community, where they arrive at a random location. That's our basic setup for this simulation.

- Radius : This parameter changes the size of each particle.
- Speed : This parameter controls the speed at which particles move randomly in the city.
- Number of Samples per Community : This is the total no of particles in the simulation.
- Number of Initially infected Community : This signifies how many communities are infected initially before starting the simulation.
- Initial Infection per Community : This signifies what percentage of particles in each initially infected community is infected before starting the simulation.
- Probability of moving to different Community : This signifies what are the chances of any individual moving to a different community.
- Initial Infection : This is the percentage of people infected initially before starting the simulation.
- Day/Sec : This represents 1 second in real time is equivalent to how many days in the simulation.
- Infection Days : This represents number of days each infected particle takes for recovery(or decease).

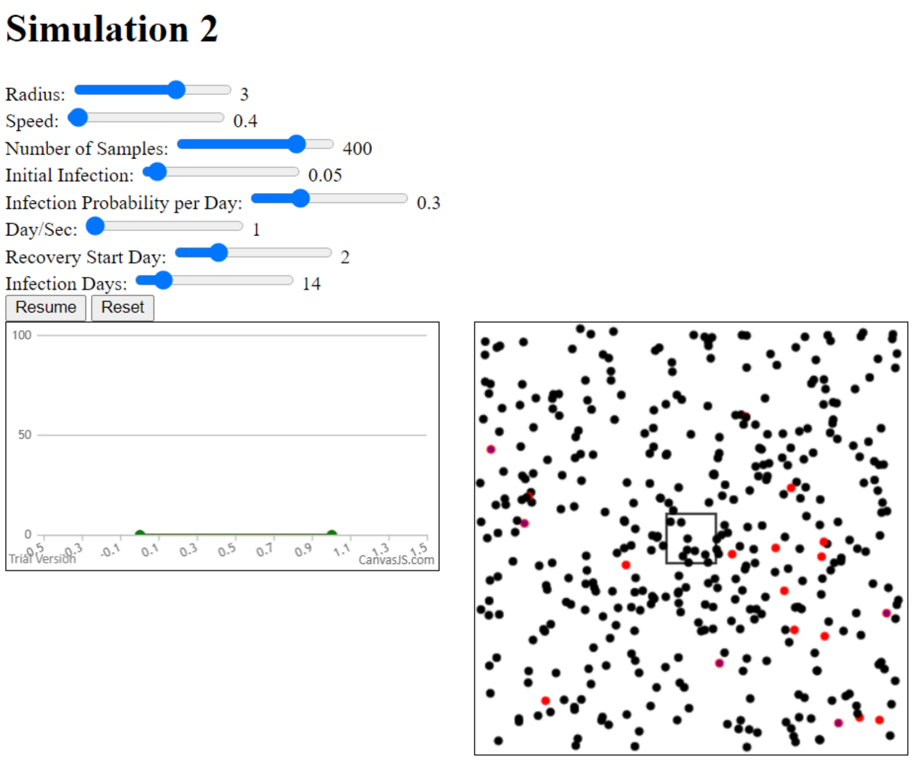

Figure 7: Simulation 2 : Starting

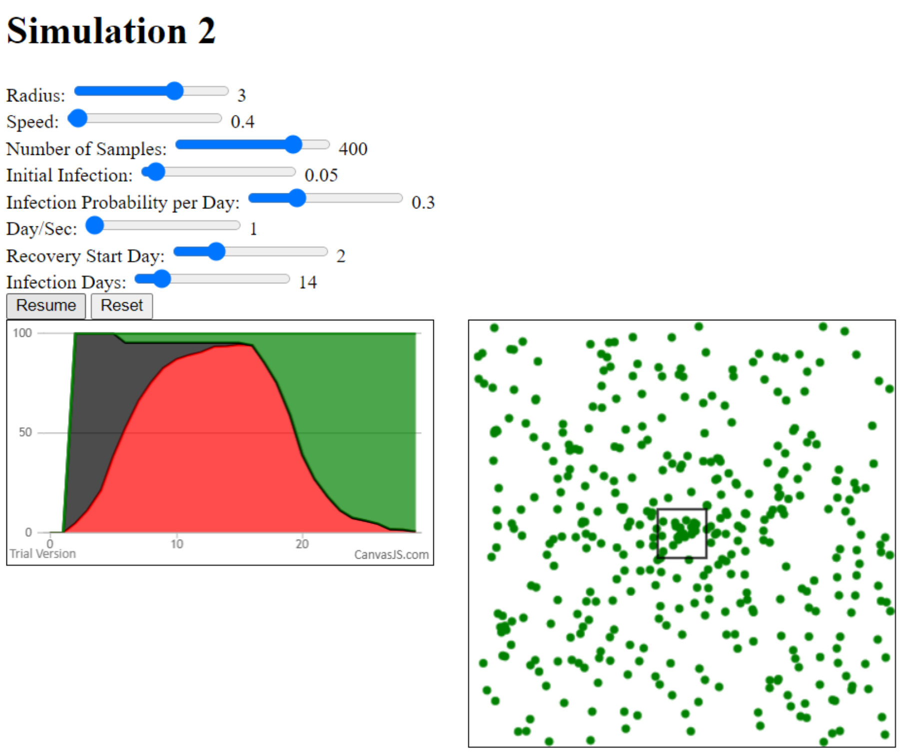

Figure 8: Simulation 2 : Finished

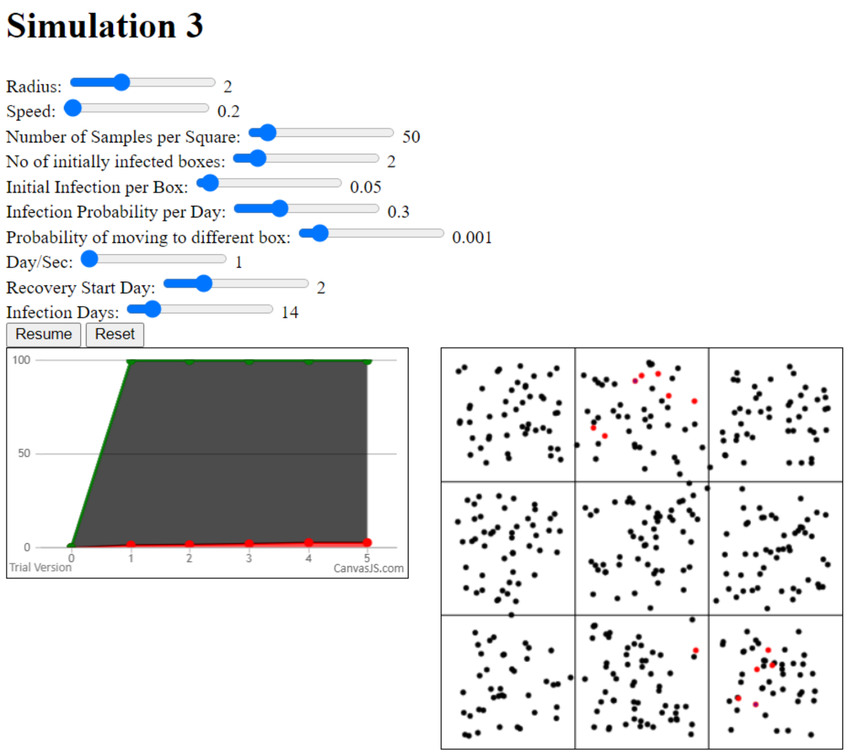

Figure 9: Simulation 3 : Starting

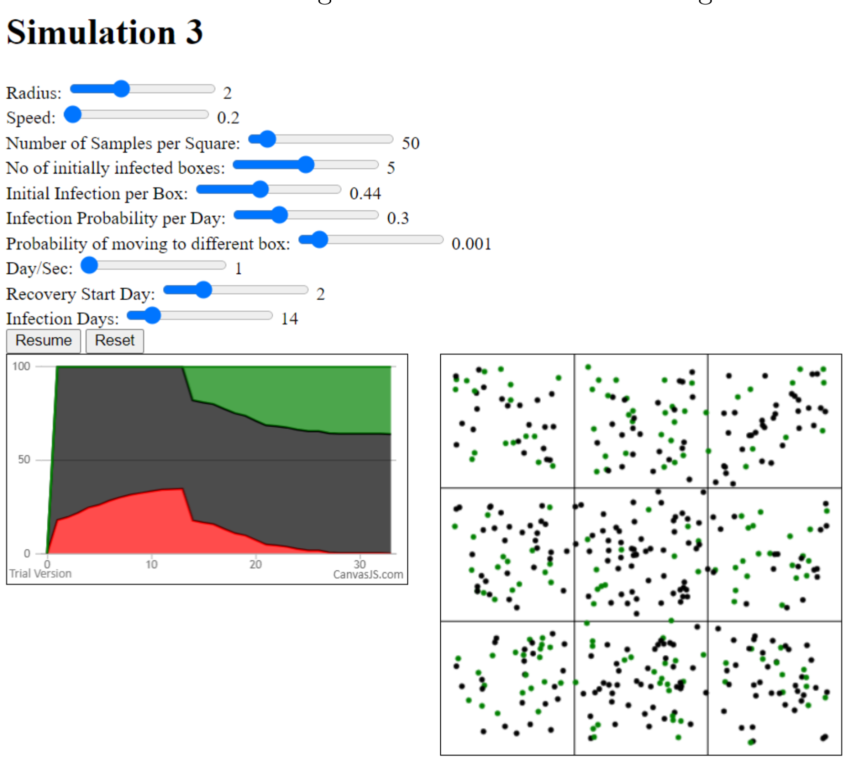

Figure 10: Simulation 3 : Finished

4 Conclusion

In this report we have tried to analyse how a pandemic spread is affected by different aspects of the society. With simulations in hand, we have explored lots of different combinations and the curve was indeed flattened with measures like less contact and early detection of the cases. We have also explored real data on India and found out some insights which give us a realization on how the pandemic is growing in our own country.

5 Future Improvements

We are working on adding few more features to the simulations. The features are listed below.

- Social Distancing : We can add a layer of social distancing, which will let people reduce contact with other people.
- Infection Radius : It's not required for people to collide physically to contract the disease. The disease can spread from an infected person with some infection radius.
- Quarantine : We can take out infected people to a separate location after their disease has been detected.
- Asymptotic Case : There can be some infected cases which are not showing any symptoms. Hence remain undetected.

Other than these features, we are also working on putting everything in the form of a single website.

References

[1] http://www.healthdata.org/data-visualization/covid-19-us-state-state-projections

[2] https://www.washingtonpost.com/2020/world/corona-simulator/?arc404=true

[3] https://covid19.healthdata.org/united-states-of-america

[4] https://www.genomedetective.com/app/typingtool/cov

[5] https://www.washingtonpost.com/graphics/2020/world/corona-simulator

[6] https://prajwalsouza.github.io/Experiments/Epidemic-Simulation.html

[7] https://www.go-fair.org/implementation-networks/overview/vodan

[8] https://github.com/nytimes/covid-19-data

[9] https://github.com/ieee8023/covid-chestxray-dataset

[10] https://healthweather.us/?mode=Atypical

[11] https://www.who.int/india/emergencies/india-situation-report

[12] https://api.covid19india.org/

[13] https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

[14] https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?