

INFO6205 Project: Viral Diseases Simulator

Manik Kumar, Pallak Singh, Yatish Pitta

{kumar.man, singh.pal, pitta.y}@northeastern.edu

December 11, 2020

Abstract

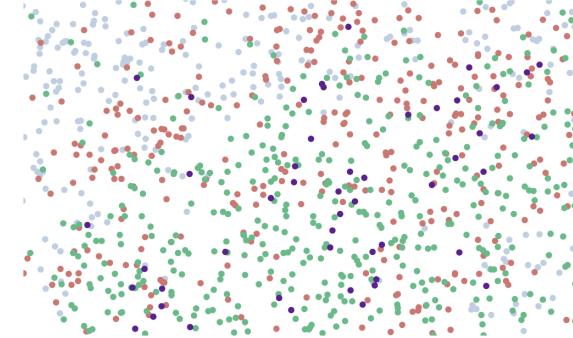
Infectious diseases continue to remain one of the major causes of morbidity in humans and animals alike all across the world. The frequency of outbreaks that have been reported over the past few decades has increased by a concerning amount. More recently, the newly identified SARS-Cov-2 has led to a large number of deaths along with over 50 million cases worldwide. As of December 8 2020, there haven't been any publicly administered vaccines released to control the impact of COVID-19 thereby, there has been extensive research not just in therapy for the infected but for controlling the spread of the disease itself. This requires cooperation between the government, the scientists and the general public in order to have an effective response. With the current pandemic, one can see that there has been resistance within the general public when it comes to following social distancing advisories and mask mandates which can be attributed to the lack of awareness when it comes to how viruses work and how they can be controlled. This project worked towards building an application capable of simulating a typical spread of viral diseases and how the spread and the impact of the disease varies depending on factors like K and R values, social distancing, mortality rates, mask mandates and even the specific effectiveness of a particular of mask. We also provided a helpful UI for better accessibility and information regarding each of the factors taken into consideration.

1 Introduction

With the spread of the virus SARS-Cov-2, which has now been named COVID-19 by World Health Organization, the world has found itself struggling on all fronts. From the dwindling economy to the overwhelmed healthcare capacity, this pandemic has been unpredictable and devastating. With the cases still surging and showing different behaviors in different areas, researchers are turning to mathematical models to predict how the disease is spreading. These models take into account different factors related to the virus along with specific population factors and simulate how the virus will behave. Simulation models can be useful not only to researchers, but also to the general public in order to understand a virus without the scientific jargon and visualize the impact of the measures being taken to control its spread. This encouraged us to build a tool that is accessible to a non-scientist and allows them to tweak different factors to see how it impacts the spread and effects of a virus.

Our application is a Viral Disease Simulator capable of visualizing the spread of a viral disease, providing real-time graphs and statistics to better understand the simulation. As shown in Figure 1, one can easily visualize how social distancing, mask mandates, hospital capacity affect the statistics.

Simulation



Statistics

Frame:	297	Population:	1000
Currently Infected:	309	Masks Wearing:	At frame 320
Healthy People:	238 or 23.8%	Social Distancing:	Active
Immune People:	404 or 40.4%	People Distancing:	500 or 50.0%
Total Deaths:	49 or 4.9%	Total Infected:	762
Healthcare Conditions:	Extreme		

Figure 1: Simulation on the Viral Diseases Simulator

We also provide a helpful UI which allows the user to modify configuration in order to simulate different scenarios. Our application comes with data loaded for the Influenza and COVID-19 disease. Users can customize virus related configurations to create their own scenarios.

2 Aim

The aim of this project is to model and simulate the spread of viral diseases such as COVID-19 and Influenza, and understand how factors like the K value, R value, population density, infection range, mortality rate and different intervention policies like social distancing, mask mandates, and travel restrictions impact the outcome of the viral outbreak. Our goal is also to provide a helpful GUI to interact with the application easily and provide an option to the users to tweak various factors and simulate how the virus behaves if that factor is changed. For complete understanding of the model, our aim is to be as comprehensible as possible through the means of charts and graphs.

3 Project Details

Viral Diseases Simulator is a stochastic simulation model that can project the spread of different viral diseases. The model works by taking into account the R and K factors of the disease, the mortality rate, recovery time and projecting the number of infections and state of the outbreak based on the healthcare capacity, intervention, and protection policies.

The program is written in Python, with the help of the numpy library to aid in the mathematical modelling of the diseases. The computer graphical simulation is done using the Matplotlib library and the graphical user interface is written in Python, with the help of the Tkinter package and ttk module for styling.

Visualization style used in this project was inspired by Paul Val Gent's Covid visualization[1]. And certain elements in designing the algorithm were inspired by Grant Sanderson's YouTube videos discussing the current pandemic[2][3].

Details for the implementation and the final deliverable can be found in section 4 and section 6. The executable binaries of the application is available for macOS, Windows and Linux on the project repository on GitHub.

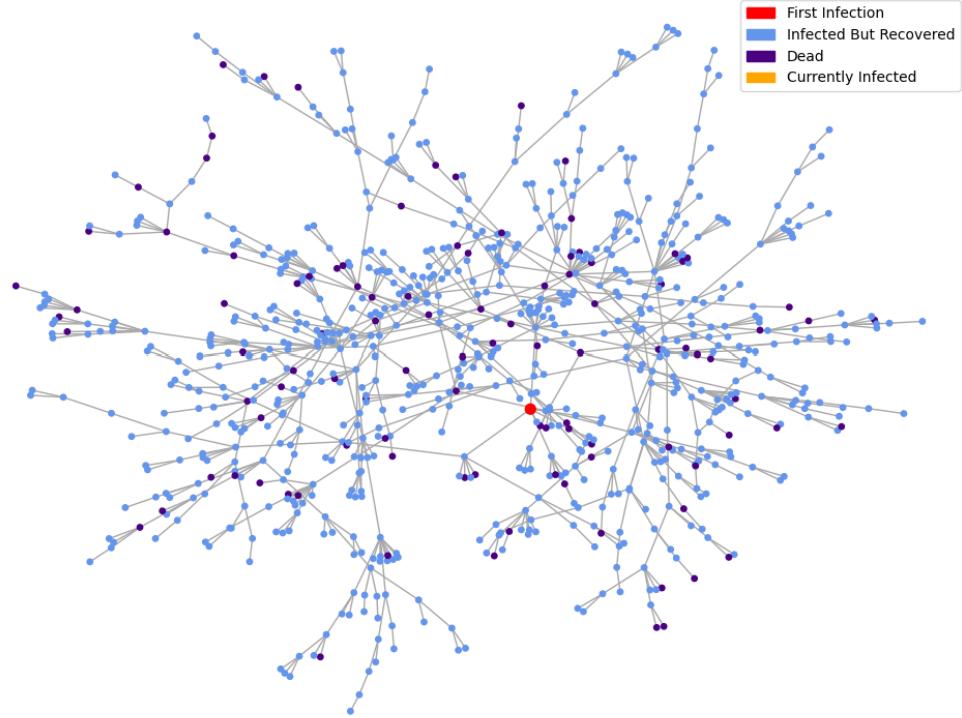


Figure 2: Network of spread from the first infection.

3.1 Features

The program provides the functionality to simulate any viral disease by allowing the user to configure different parameters specific to the disease. All parameters available to the user and their respective functions are provided in section 6. User can also view a network graph showing the spread of the virus from the first infected person once they end the simulation. All details regarding the algorithm can be found in section 5.

The application comes in built with configuration data for two diseases –

1. **COVID-19** gathered from [4],[5]
2. **Influenza** gathered from [6],[5]

The simulation can be observed in two modes:

1. **Render Mode** Renders the simulation into a .mp4 video file on the disk. User can select the render mode from the radio buttons present on the right of the UI screen.
2. **Live Simulation Mode** Displays the simulation live on a matplotlib window.

User also has the option to run the application in two modes:

1. **GUI (Graphical User Interface) mode:** This mode allows the user to configure different parameters of the simulation model through an intuitive UI and observe the simulation as they make the changes. The UI can be seen in 3.

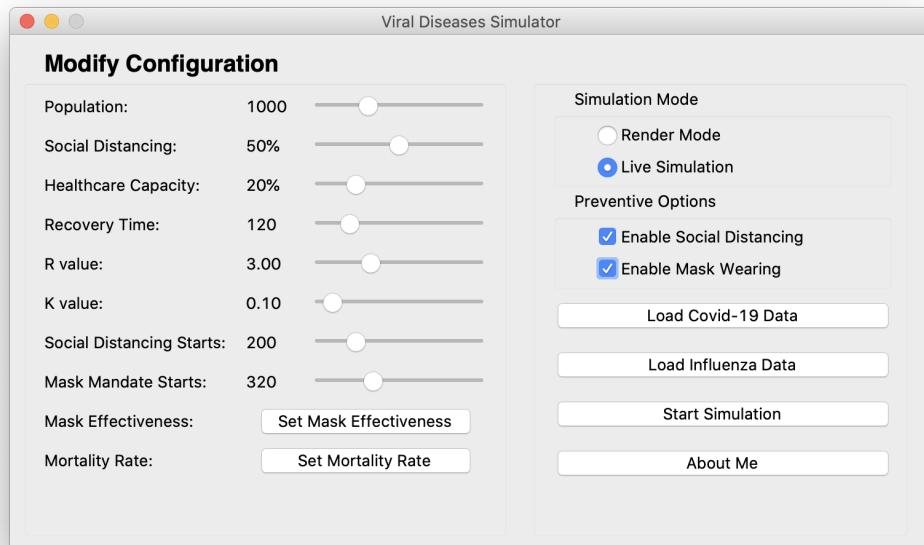


Figure 3: GUI screen user is welcomed with

2. **CLI mode:** This mode allows the user to run the simulation without the need to interact with the GUI. The simulation can be configured through the config/config.ini file. This will default to covid.stats portion of the config when it comes to variables that define the characteristics of a virus(stuff like K or R value). Change the values in the covid.stats to change virus characteristics. The user can run the application in the CLI mode by entering the following command on their terminal:

```
python main.py --disable-UI
```

3.2 Structure of the Application

The core algorithm is located in the /src folder and contains all the modules necessary for the simulation.

The /gui folder contains the code for the graphical user interface built using Tkinter.

The /test folder contains all the unit tests for the model.

The following are the core modules of the model:

1. **config_util.py:** Contains the ConfigUtil class, which has the functionality to parse configuration files and return configuration values.
2. **population.py:** Contains the Population class, which holds the main data container containing the information about each person. It also has methods to initialize the data container, and getters and setters for every data type.
3. **population_util.py:** Contains the PopulationUtil class, which is used to instantiate the Population class by passing in the appropriate configuration values from the configuration file. It also contains the move method, which is called at every frame interval and does the heavy lifting of calling the update methods to update the positions of each individual and calling the infect method.
4. **virus_util.py:** Contains the Virus class, which contains the main utility methods to infect individuals based on different factors and update the state of each individual.
5. **movements.py:** Contains the Movement class, which contains utility functions to update the position of individuals on the X-Y plane.
6. **visualization.py:** Contains the Visualization class, which makes use of the matplotlib library to plot the population and simulate the model.

The UML diagram representing the core modules and how our application is designed to include those modules is provided in Figure 4.

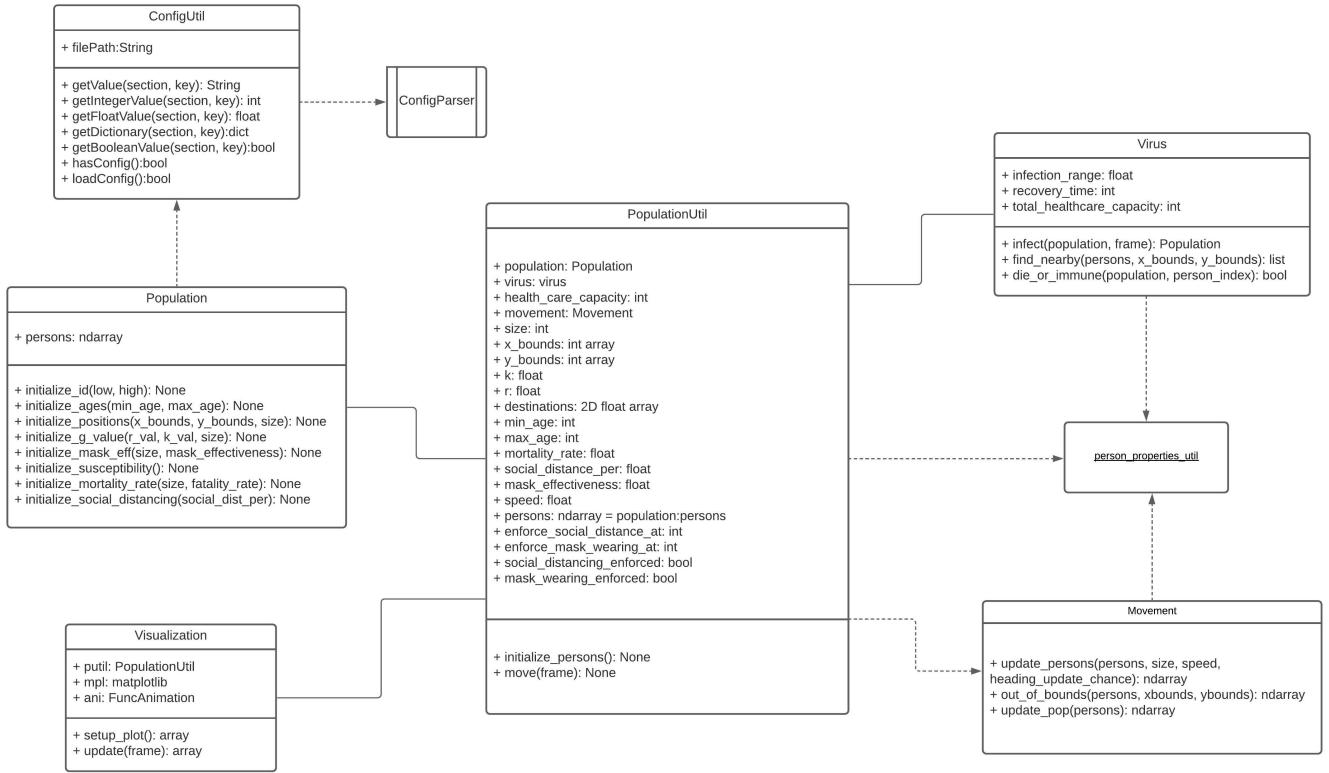


Figure 4: UML diagram representing the core modules of the model

4 Implementation

4.1 Data Used for Model

To create a sufficiently competent model, we start with loading in various statistics about the virus we're trying to model. Using those values and other user provided data, a set of random individuals are initialized which will contain various different variables using which the model algorithm will work.

The following virus statistics are being used.

1. **R Value:** R value or the reproduction number is defined as the average number of individuals an infected person will infect over the course of the spread [7].
2. **K Value:** K value or the dispersion value is defined as the variance in the number of individuals an infected person will infect [5].
3. **Infection Range:** The physical distance under which it's plausible for an infected individual to spread the virus to someone else.

4. **Mask Effectiveness:** This defines how effective masks are in protecting a healthy individual when they come within the infection range of an infected person. Mask effectiveness are defined for the following different masks.
 - (a) Cloth Masks
 - (b) Surgical Masks
 - (c) N95 Masks
5. **Mortality Rate:** This defines mortality rate of the virus for the following age ranges.
 - (a) 0 to 19 years old.
 - (b) 29 to 49 years old.
 - (c) 50 to 69 years old.
 - (d) 70 years old and above.

6. **Recovery Time:** This defines how long on an average it takes for someone to recover from the virus.

The user provides following data-points using which different kinds of simulations can be visualized.

1. The size of the population.
2. What the healthcare capacity is.
3. If there would be any social distancing measures
4. How many individuals will respect the social distancing measures.
5. When would the social distancing measures start.
6. If there would be mask mandates.
7. When would the mask mandates start

To create a model that was sufficiently fast enough to update in real time and produce a simulation we decided to use matrices to store individual variables. The matrix storing all the data has 18 data columns,

1. **ID:** Every individual is assigned a unique ID.
2. **Age:** Stores the Age of the individual.
3. **Starting Position X-axis:** X-axis of the position this individual will start from.
4. **Starting Position Y-axis:** Y-axis of the position this individual will start from.
5. **X-axis Heading:** X-axis of the direction this individual will be heading towards.

6. **Y-axis Heading:** Y-axis of the direction this individual will be heading towards.
7. **Wander Range X:** X-axis range of where an individual can wander in.
8. **Wander Range Y:** Y-axis range of where an individual can wander in.
9. **Speed:** Speed of movement of an individual.
10. **Current State:** Current health status of the individual, e.g: Healthy, Immune, or Deceased.
11. **Individual Infection Value:** Number of individuals this individual has the potential to infect if infected. This value is determined through a normal distribution taking in the R and K values of the virus.
12. **Mask Effectiveness:** If this individual would wear a mask, if yes what kind of mask(e.g: cloth mask, N95, or surgical mask) and how effectively would it protect against the virus.
13. **Social Distancing:** If this individual will participate in social distancing or not.
14. **Susceptibility:** How susceptible is this individual for catching the virus, this field depends upon the infection range of the virus, what kind of mask the individual is wearing and if they're social distancing.
15. **Mortality:** Percentage that determines an individual's mortality chance. This value depends upon the individual's age, how deadly the virus is, and if healthcare capacities are full. If healthcare capacity conditions are extreme, mortality rates rise.
16. **Infected At:** Stores when this individual was infected at for algorithmic purposes.
17. **Hospitalization Status:** Stores whether or not this person was hospitalized.
18. **Infected By:** Stores who this individual was infected by, this is used to create the virus spread network graph.

4.2 How R and K values are being incorporated in the model

As explained above, R is the reproduction number which is defined as the average number of individuals an infected individual will spread the virus to[7]. And the K is the dispersion value which is defined as the variance in number of individuals an infected individual will infect. A smaller value of K means more variance and a bigger value of K means less variance.

As an example, for a virus with a R value 4.0 and a low K value like 0.1, it implies on an average each infected individual will infect 4 other individuals but on the independent level, there is going to exist a big variance, such that there will exist people who infect barely anyone and there will exist people who will infect 10 people or more.

Now, if the same virus was given a larger K value like 25.0, it would imply that on the individual level most people will infect only 4 people. Hence, a small variance for the number of individuals an infected person will infect.

From the explanations above, we can clearly notice that the variance in R is inversely proportional to the value of K.

$$Var[R] \propto \frac{1}{K} \quad (1)$$

To incorporate the above information in the model, we created a data field which would store number of individuals each person can spread the virus to, called "Individual Infection Value". Now to generate individual infection values that vary proportionally to the inverse of K while keeping the average equal to R we used a normal distribution[8].

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (2)$$

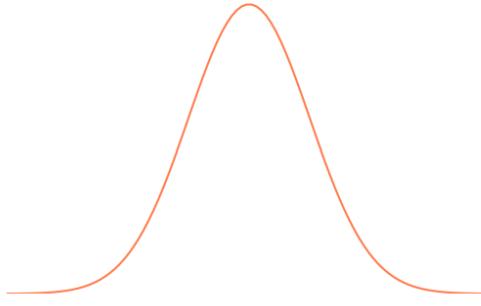


Figure 5: What a normal distribution looks like.

The above equation represents a normal distribution, where μ is the mean of the distribution and σ is the standard deviation.

Using R and K in the normal distribution, using equation 1 and 2, we get,

$$F(x) = \frac{1}{\frac{1}{K}\sqrt{2\pi}} e^{-(x-R)^2/\frac{2}{K}} \quad (3)$$

When initializing the values for the individual infection value, equation 3 is utilized to generate a random normal distribution, while keeping the mean of all the values equal to R while the variance is kept inversely proportional to K. Now this data is used in the algorithm as a metric of how many individual an infected individual can spread the virus to.

5 Algorithm

The following depicts and explains the algorithm developed which models the virus spread and simulates it.

Initialization Phase of the algorithm:

1) Step 1: Load the following virus statistics and user specified data mentioned as in section 4.1.

- (a) R Value
- (b) K Value
- (c) Infection Range
- (d) Mask Effectiveness
- (e) Mortality Range
- (f) Recovery Time
- (g) Population
- (h) Healthcare Capacity
- (i) Social Distancing Start Time(If Enabled)
- (j) Percentage of Individuals Social Distancing(If Enabled)
- (k) Mask Mandate Start Time(If Enabled)

2) Step 2: Initialize Data

- (a) Create a matrix where number of rows would be equal to the population with a unique ID for everyone and number of columns is 17 as explained in section 4.1.
- (b) Generate a random spread of ages using a continuous uniform distribution[9] in a specified range so that they represent a more lifelike age spread.
- (c) Generate random starting (x, y) co-ordinates using a continuous uniform distribution for every individual.
- (d) Generate random directional (x, y) co-ordinates using a continuous uniform distribution and random wander regions $(x1, x2)$ and $(y1, y2)$ for every individual.
- (e) Initialize Individual Infection Value by generating it using a normal distribution which uses R and K values as explained in section 4.2.
- (f) Set the speed of movement using a continuous uniform distribution in a specified range for every individual.
- (g) Initialize individuals who would be social distancing. Using the percentage specified by the user, random individuals are chosen to participate in social distancing. Social Distancing kicks in after the specified time.
- (h) Initialize mask effectiveness for people, if someone would wear a mask and what kind of mask they wear is again decided using a random function. This value would be 0 for individuals not wearing a mask.

- (i) Calculate and initialize susceptibility rate for each individual. Susceptibility is calculated by factoring in what kind of mask the individual is wearing and if they're social distancing.
- (j) Calculate and initialize mortality chance for each individual. Mortality chances depend upon the individuals age and the viruses mortality rate for that age range.
- (k) Choose a random individual, and infect them.
- (l) Initialize a integer, $frame = 0$, which will keep track of every single frame/turn in the simulation loop.

Now entering the looping phase of the algorithm.

1) Step 1:

If $frame ==$ Time at which social distancing starts if enabled:

Start social distancing.

If $frame ==$ Multiple of specified time:

Randomly switch up people following social distancing.

2) Step 2:

If $frame ==$ Time at which mask mandates starts if enabled:

Start mask mandates.

3) Step 3: Move Individuals:

For Individuals who are not deceased and aren't social distancing:

$[X \text{ co-ordinate value}] = [X \text{ co-ordinate value}] + [\text{Speed}] * [X \text{ direction value}]$

$[Y \text{ co-ordinate value}] = [Y \text{ co-ordinate value}] + [\text{Speed}] * [Y \text{ direction value}]$

4) Step 4: Infect People:

- (a) Find individuals in the infection range for other infected people range whose individual infection value is not 0:
- (b) Infect them based on randomly generated chances using a continuous uniform distribution and their susceptibility.
- (c) Decrement individual infection value for individuals who infect others.

5) Step 5: Decide outcome:

- (a) Find infected individuals who have been infected for over the specified recovery time.
- (b) Using randomly generated chances using a continuous uniform distribution and their mortality chance decide either they survive for not.
- (c) Mark them as immune if they survive, or deceased if they did not.
- (d) If the number of currently infected people is extremely above the healthcare capacity at the current moment the chances of an individual dying increases 3 folds.

6) Step 5: Update Values for Random Individuals:

Update the directions and speed of 20% of the population, using a uniform distribution again.

7) Step 6: Check Out of Bounds:

Check if an individual is headed outside of the defined simulation canvas, and chance their directions such that they do not exit the plane.

8) Step 7: $frame = frame + 1$ and jump to **step 1**.

6 Output

Our GUI looks like the following:

From the configuration screen, we can set following things:

1. Population size
2. Number of people social distancing as a function of the population size
3. The capacity of the hospital as a function of the population size
4. The R value for the virus
5. The K value for the virus
6. The time social distancing starts (if at all)
7. The time mask mandate starts (if at all)
8. Mask effectiveness of 3 types of masks:
 - (a) Cloth masks
 - (b) Surgical masks
 - (c) N95 masks
9. Mortality rates of 4 age groups:
 - (a) 0-19
 - (b) 20-49
 - (c) 50-69
 - (d) 70+

Following are a series of screenshots which depict how the GUI looks, and how a user can go around operating the simulation system.

In the figure below, user can click on "Set Mask Effectiveness" button to set how effective different kinds of masks are, or click on "Set Mortality Rate" to set mortality rates for different age groups. User can also enable/disable mask mandates and social distancing using the radio buttons.

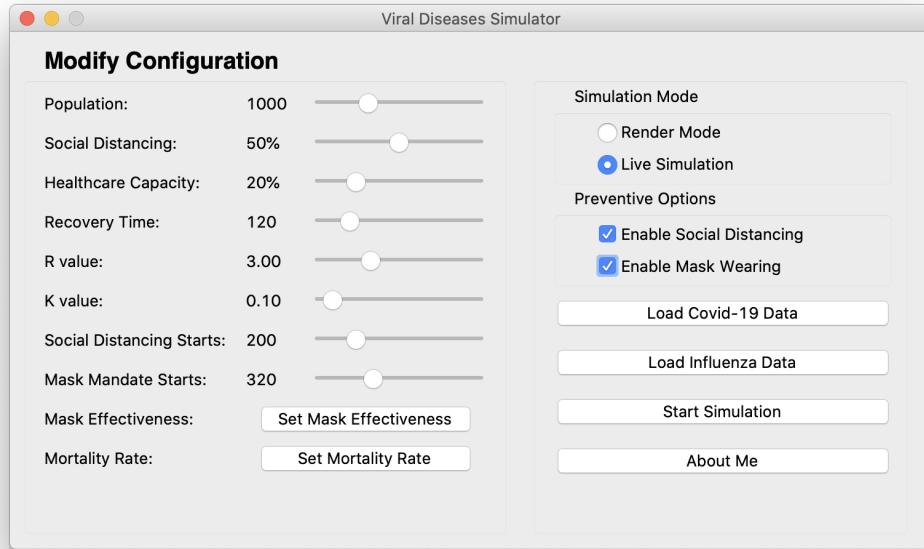
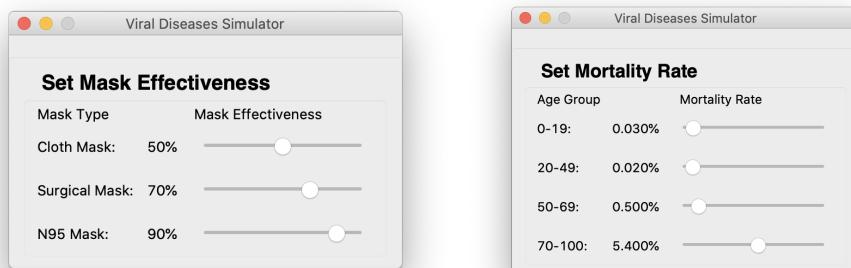


Figure 6: UI the user is greeted with.

Using the "Load Covid-19 Data" and "Load Influenza Data" button can load the virus statistics to the the program.



(a) Dialog to edit mask effectiveness.

(b) Dialog to edit mortality rate.

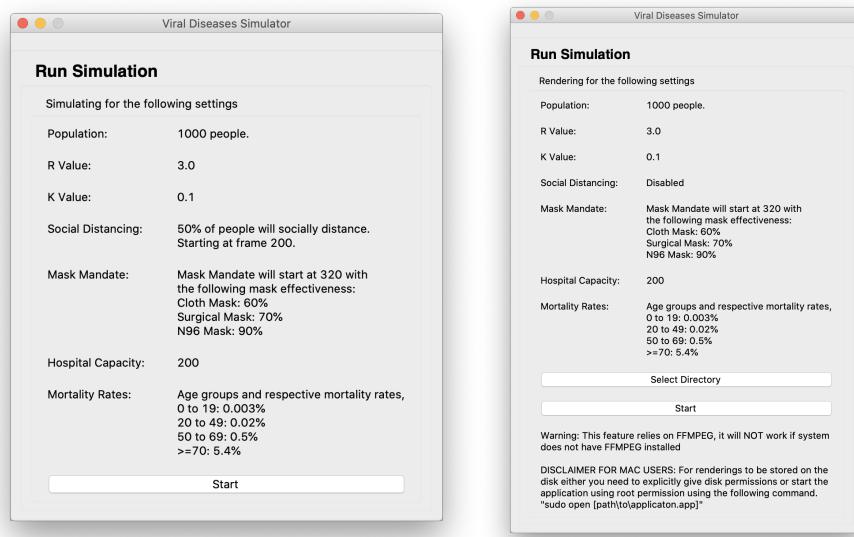
Figure 7: Dialog boxes in the GUI to edit mortality rates and mask effectiveness

After the user has decided the preferences they want to use, they can decide whether they want to

render the simulation to a file or view it live, and then use "Start Simulation" button to initialize the simulation. In live simulation mode, when the user closes the simulation, a network graph of how the virus spread is generated and displayed.

After clicking on the start button, the user is treated to the following dialog boxes, where the set settings are shown to the user for confirmation before simulation execution. In case the user was in render mode, an additional option for deciding the path to store the render and a disclaimer about additional required libraries and elevated access are shown.

Clicking on start will start the live simulation or the render.



(a) Start dialog in simulation mode.

(b) Start dialog in render mode.

Figure 8: Dialog boxes when starting simulation.

The figure below represents how the simulation either viewed live or rendered looks. The top part of the simulation is a scatter of dots which represent people, with light grey dots representing individuals who are healthy/haven't contracted the virus, orange dots represent individuals who are currently infected, green dots represent individuals who have recovered and are immune, and indigo dots represent individuals who have been deceased.

The middle part of the simulation is a live line graph which depict current trends and data for total number of infected, currently infected, immune and deceased people against time. Vertical lines in the graph depict the time when social distancing or mask mandates start, with yellow representing social distancing and pink representing mask mandates.

The third and the last part of the simulation show raw statistics of the current simulation, with various data points like population, number and percent of immune and deceased people, when and if preventative measures will start, and how the healthcare conditions are at a given time.

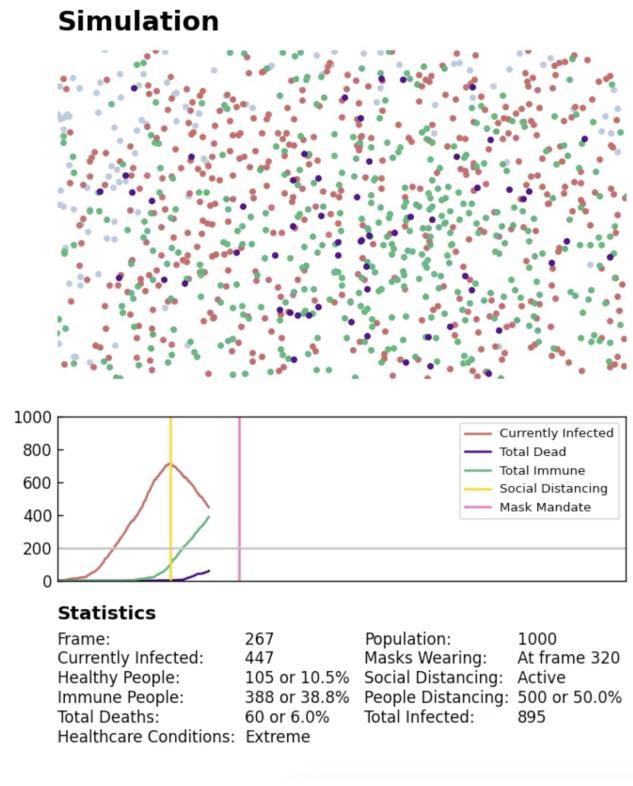


Figure 9: How the simulation looks.

6.1 Unit Testing

Unit tests have 90% coverage of the simulation software.



(a) Screenshot depicting different unit tests.

Name	Stmts	Miss	Cover
<hr/>			
src\config_util.py	52	11	79%
src\movements.py	31	0	100%
src\person_properties_util.py	19	0	100%
src\population.py	85	17	80%
src\population_util.py	64	7	89%
src\virus_util.py	43	6	86%
test\test_config_util.py	30	0	100%
test\test_movements.py	44	0	100%
test\test_population.py	60	10	83%
test\test_population_util.py	43	5	88%
test\test_virus_util.py	74	0	100%
<hr/>			
TOTAL	545	56	90%

(b) Screenshot depicting code coverage.

Figure 10: Screenshots providing insight in unit testing.

7 Mathematical Analysis & Evidence

In this section, we explore different factors involved in the spread and impacts of a virus. We vary the factors differently and provide results for each case. We analyze the results are we obtained from our experiment simulations.

7.1 Exponential Growth

The exponential growth rate of an outbreak indicates the severeness of the spread of the disease. The growth rate and the reproduction number (R value) are ultimately correlated because with R value greater than 1, the number of infections increase exponentially as each person goes on infect at least 1 person. Once the R value goes below 1, the outbreak slowly starts to contract. This may happen due to two reasons: the entire population gets infected as indicated in or due to control measures like social distancing and mask mandates taken to reduce the number of people each person goes on to infect. As indicated in Figure 11 and Figure 12.

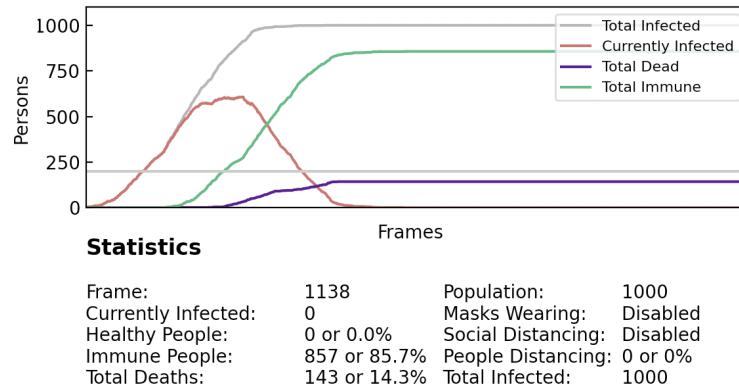


Figure 11: Virus Spread with Time for COVID-19 with no Control Measures

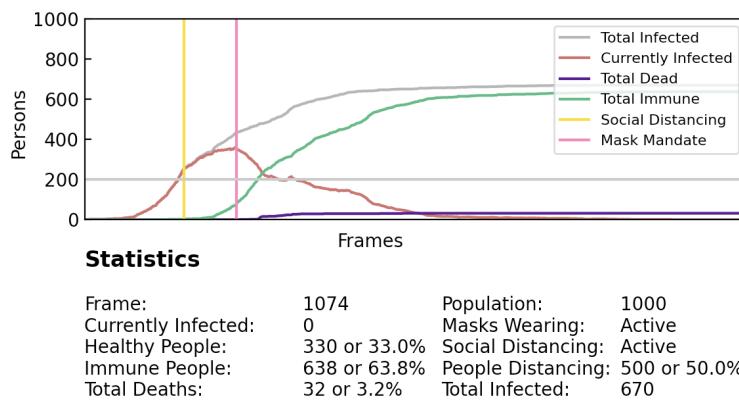


Figure 12: Virus Spread with Time for COVID-19 with Social Distancing and Mask Mandates

In Figure 11, one can see that the graph of the total number of total infected grew exponentially and reached a saturation once the entire population had been infected. Meanwhile, in Figure 12, with social distancing and mask mandate being followed by a percent of the population, the infection grew exponentially, but reached a saturation point at a much lower number with 330 people not catching the disease at all.

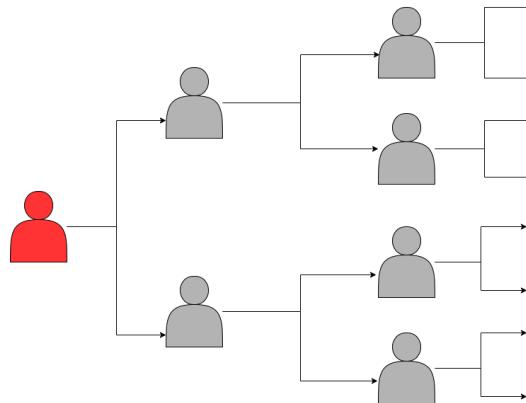


Figure 13: Assuming a virus with R value of 2, the virus grows exponentially

A visual representation of the effects of reproduction number (R value) on how the disease spreads is given in Figure 13. Assuming a virus with R value of only 2, the disease grows exponentially if control measures are not taken.

7.2 R value

As explained in Section 4, the R value indicates the average number of people an infected person goes on to infect. By running a simulation for 3 scenarios: R value equals 1, R value equals to 3 (COVID-19) and R value equals to 8 and keeping all other factors constant one can notice how fast the virus spreads with increasing R value. The simulations for all three cases are shown in Figure 14, 15, 16.

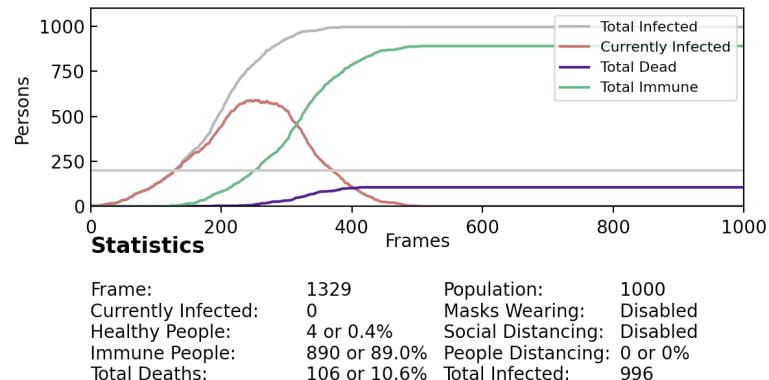


Figure 14: Virus Spread with Time when R value = 1 and K value = 1

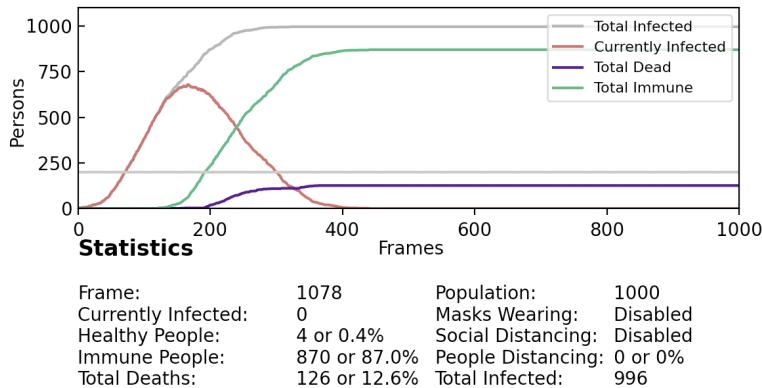


Figure 15: Virus Spread with Time when R value = 3 and K value = 1

From the shown graphs, one can see that for $R = 1$, the virus was spread much slower than for higher R values with the peak current infection rate being observed at roughly 250th frame. The curve for currently infected is also much more spread out meaning that the healthcare facility capacity is not overwhelmed all that much, thereby causing lower number of deaths. The number of deaths in these scenario was only 106 (10.6%) people. This is also known as flattening the curve, a common term being used these days to promote control measures amongst the population so as to not let healthcare facility capacity reach its limit.

For $R = 3$, the curve is spread out although not as much as for $R = 1$ and the peak is reached around 200th frame. The number of deaths in this scenario is 126 (12.6%) people.

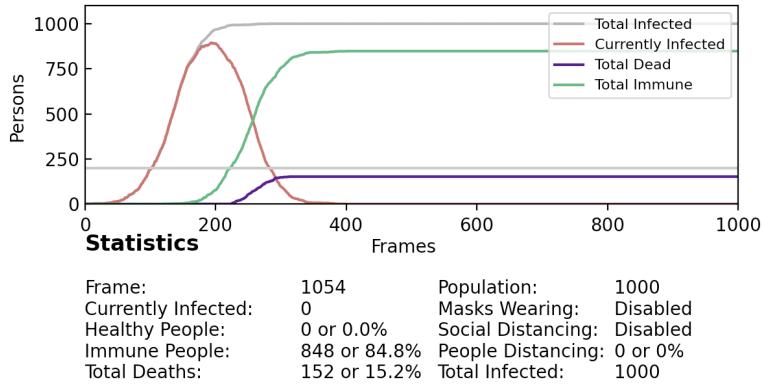


Figure 16: Virus Spread with Time when R value = 8 and K value = 0.1

For $R = 8$, the peak is observed way earlier and there comes a point where barely anyone has recovered but the entire population has already contracted the disease. The peak for the currently infected is higher, steeper and reached before the 200th frame. This is an extremely devastating scenario for the healthcare facilities as they are much more likely to run out of ICUs and hospital beds causing a higher number of deaths. The number of deaths observed in this scenario is the highest i.e 152 people (15.2%) which is to be expected. We can conclude that faster the infection spreads, the more likely it is to cause a large number of spreads as hospitals are not able to

accommodate every person who may require healthcare facilities.

When viewing the network graph for $R = 1$ and $R = 8$, one can see that for $R = 8$ where the virus spreads rapidly, the network is more dense as 1 person goes on to infect a lot more people and as opposed to the virus spreading from 1 person to another. This can be seen in Figure 14, 16.

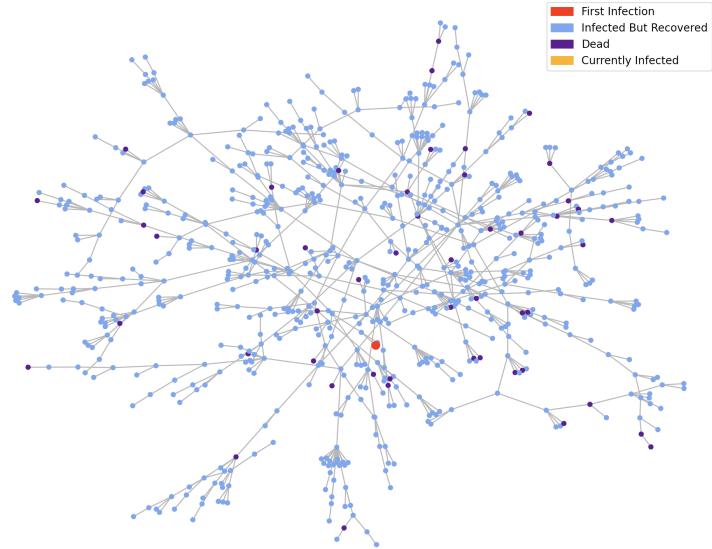


Figure 17: Virus Spread Network Visualization for R value = 1

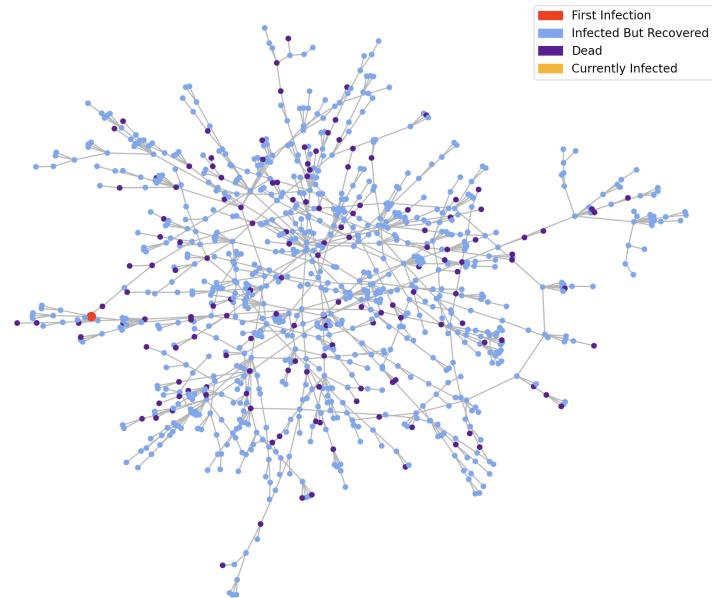


Figure 18: Virus Spread Network Visualization for R value = 8

7.3 K value

As also explained in Section 4, K value is the dispersion of individual reproduction value of the virus for a person. For lower K values, the variance is much higher, thereby a person is more likely to either not spread the disease at all or cause superspreader events. Hence, for lower K values, one would expect more superspreader events and faster exponential growth as compared to higher K values for which the the infection curve would be more spread out. Our experiment simulation results can be seen in Figure 19 and 20.

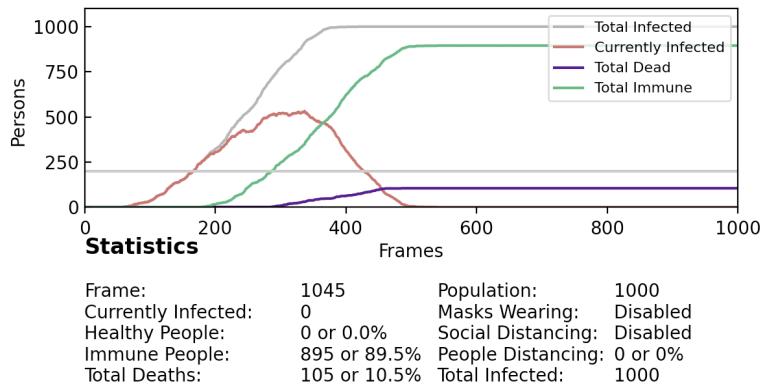


Figure 19: Virus Spread with Time when K value = 3 and R value = 3

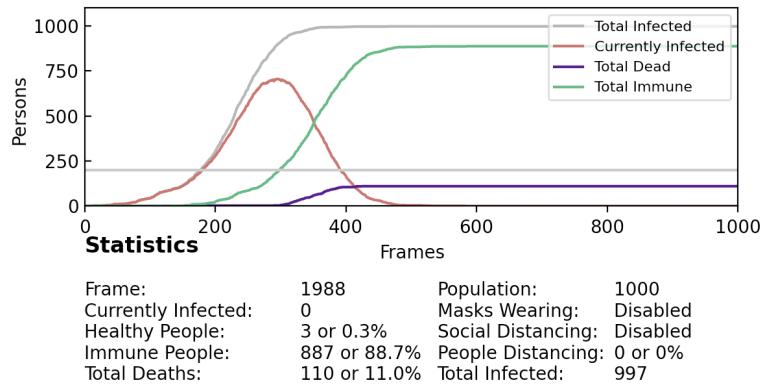


Figure 20: Virus Spread with Time when K value = 0.03 and R value = 3

One can observe that for K value = 0.03, the curve for current infections was steeper while for K = 3, the curve for current infections was much more spread out. This indicates that when K value is lower, there are more chances of superspreader events causing faster exponential growth.

7.4 Social Distancing

With the spread of COVID-19, each country, even each town had different responses to the pandemic. From the data emerging from different areas showing the virus demonstrating different growths one cannot stress enough the importance that social distancing has played in maintaining and reducing the virus spread. However, there are two main factors that come into play when one talks about how useful social distancing has been. These factors are mainly how soon the advisory was put into place and the percentage of the population that is social distancing[10]. We discuss the impact of social distancing in two different cases:

1. **How early social distancing was put into place:** We run 3 different experiments for social distancing during COVID-19: No social distancing put into place, social distancing implemented at an earlier stage of the outbreak, social distancing implemented at a later stage of the outbreak. In the latter two cases, 50% of the population is social distancing. Our results our reported in Figure 21, 22, 23.

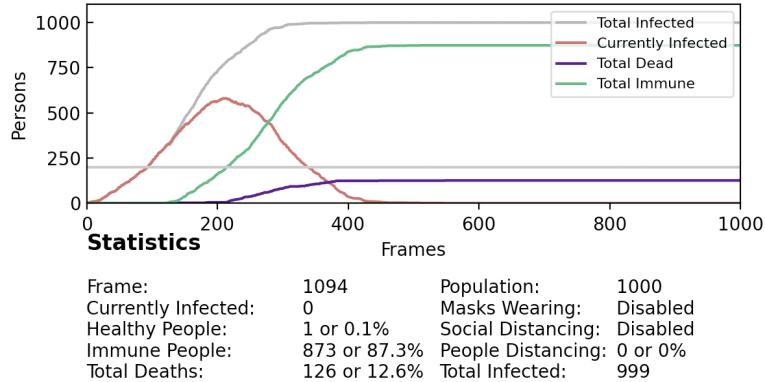


Figure 21: Virus Spread with Time for COVID-19 with No Social Distancing

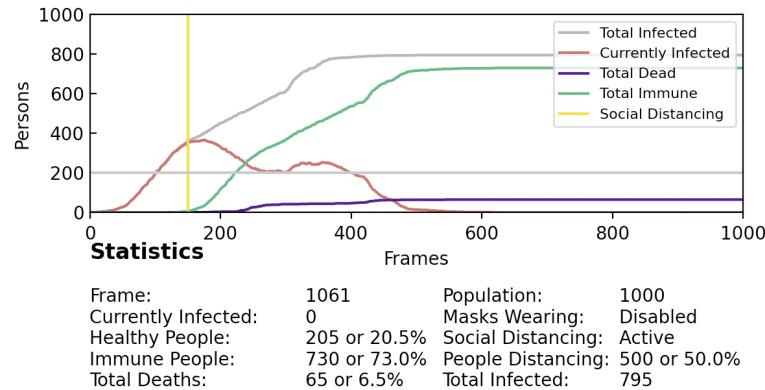


Figure 22: Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing starting at Frame 150

In Figure 21, where no one socially distances at all, we observe that the virus grows expo-

nentially until all the population has contracted the disease. 126 (12.6%) people died in this case and the entire population got the virus.

In Figure 22, where social distancing starts at an early stage (150th frame), one can see that at one point, at most only 400 (40.0%) people were infected which greatly impacted the death rates as only 65 (6.5%) people died in this case. This happened because the healthcare facilities were not overwhelmed by the surge of cases at once and thereby could admit the people who required hospitalizations, thereby reducing the morbidity of the virus. Total number of infections in this case was only 795 (79.5%) people.

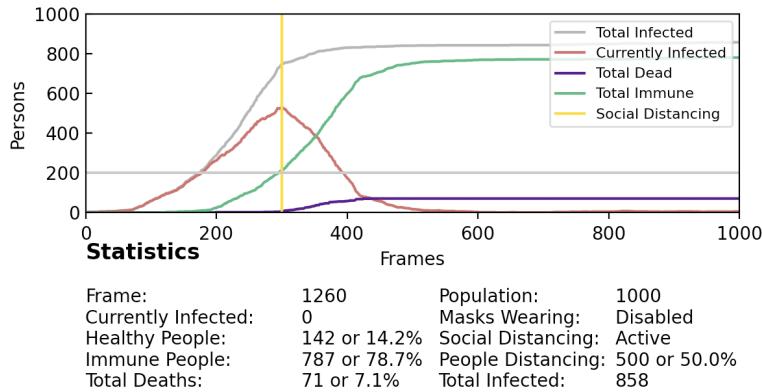


Figure 23: Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing starting at Frame 300

In Figure 23, one can see that once social distancing advisory was put into place even at a later stage (300th frame), the number of current infections started declining almost immediately. With 858 (85.8%) total infections and 71 (7.1%) deaths even with the advisory implemented at a later stage, it is easy to tell how important social distancing is when it comes to controlling the outbreak.

2. **Percentage of people social distancing:** We run 3 different experiments for social distancing during COVID-19: 20% of the population is social distancing, 50% of the population is social distancing, 80% of the population is social distancing. In all these experiments, social distancing starts at Frame 200. Our results are reported in 24, 25, 26

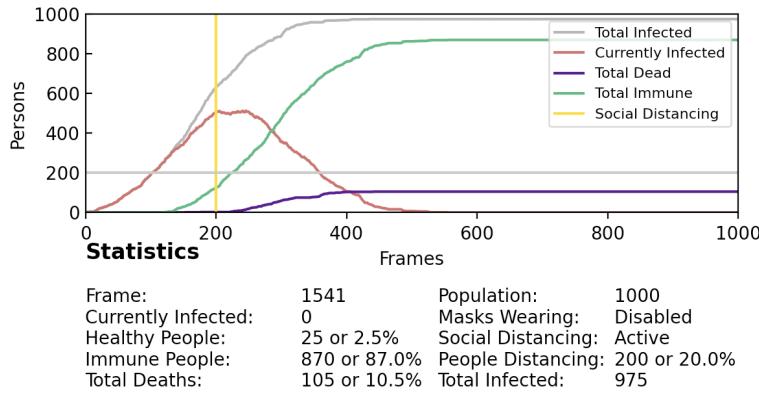


Figure 24: Virus Spread with Time for COVID-19 with 20% of the Population Social Distancing at Frame 200

In Figure 24, we can see that since only 20% of the population was social distancing, the infection curve was spread out but that did not stop the infection from spreading to the entire population. This implies that 20% of the population social distancing is not enough to control the outbreak. This scenario demonstrates the highest number of deaths i.e 105 (10.5%).

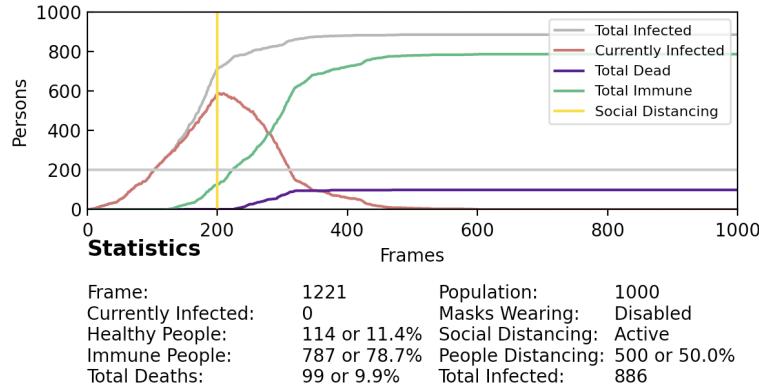


Figure 25: Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing at Frame 200

In Figure 25, there is slight improvement from the above case, where a larger section of the population never contracted the virus at all i.e 114 people (11.4%). This can be attributed to 50% of the population social distancing. The total deaths in this case were also lower in comparison i.e 99 (9.9%).

In Figure 26, we see the best results, where a big section of the population never contracted the virus at all i.e 587 people (58.7%). This can be attributed to 80% of the population social distancing. The total deaths in this case were the lowest i.e 68 (6.8%). This goes on to show that outbreak control is possible if a larger than half of the population social distances.

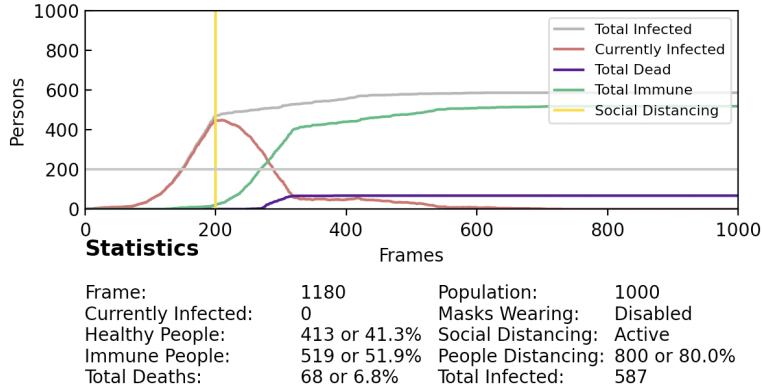


Figure 26: Virus Spread with Time for COVID-19 with 80% of the Population Social Distancing at Frame 200

7.5 Population Density

Population density plays a role in how fast the virus spreads. When people are sparsely located, they are less likely to come in contact with an infected individual[11]. When they are located in a more dense area, the virus spreads much faster as the probability of a healthy individual coming in contact with an infected person increases. There have been contradicting studies specific to COVID-19 that suggest that population density may not be linked to higher density if the activity in the population is low [12][13]. These studies are based on real urban settings with some sort of social distancing and mask mandates already in place. However, in our experiments, since there are no control measures in place to reduce activity/movement, the virus spreads faster. To show this, we run three simulation experiments for COVID-19: a population consisting of only 500 people, 1000 people, and 3000 people while keeping the size of the area and all the other factors the same. Here we see how fast the virus spreads for all 3 scenarios. The results are reported in Figure 27, 28, 29.

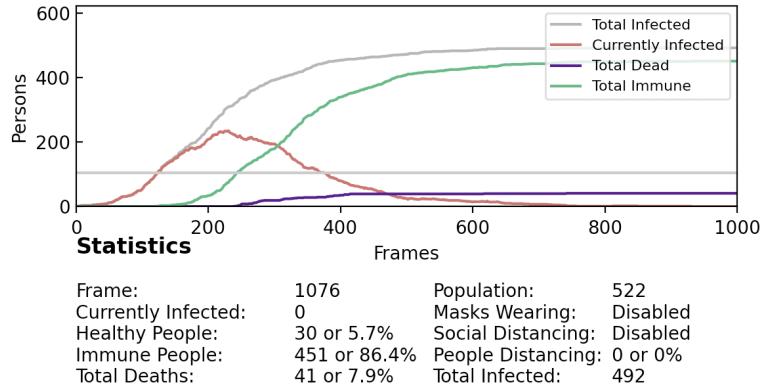


Figure 27: Virus Spread with Time for COVID-19 with Population Size = 522

In Figure 27, we can see that the virus spreads slowly as people are distributed on the map in a

more sparse manner. With a population size of only 522 distributed evenly on a map, the likelihood of a healthy person coming in contact with an infected person decreases. A small section of the people (30 people/5.7%) end up not contracting the virus at all.

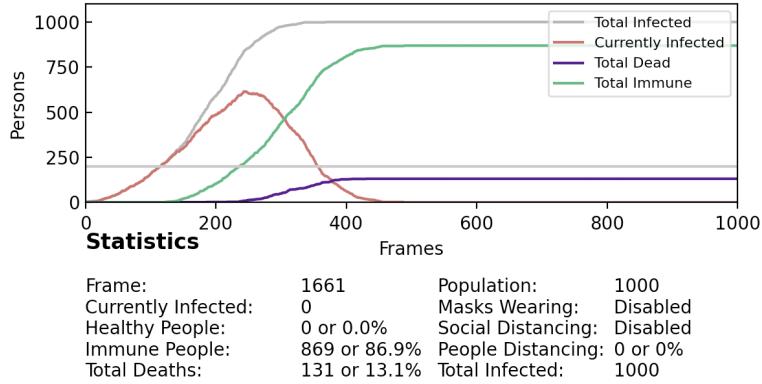


Figure 28: Virus Spread with Time for COVID-19 with Population Size = 1000

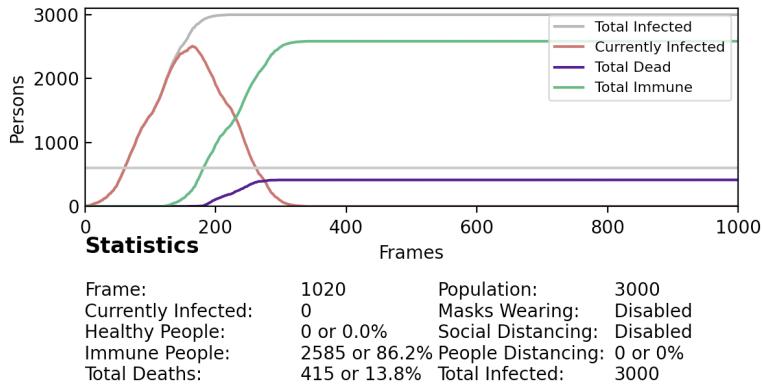


Figure 29: Virus Spread with Time for COVID-19 with Population Size = 3000

In Figure 28, we can see that the virus spreads comparatively faster with the entire population contracting the disease at around 250th frame.

In Figure 29, we can see that the virus spreads the fastest with the entire population contracting the virus even before the 200th frame. This happens because the people are so densely populated in the same area that they are very likely to come in contact with a diseased person when there are no social distancing measures and high activity at all times. We observe that death rates between the population do not differ by all that much because as the population increases, the healthcare capacity scales up accordingly. This is consistent with our findings from our research [12].

7.6 Mask Mandates

For mask mandates, we see that masks alone don't affect the number of cases drastically. This is because they do not protect an individual a 100% from contracting the virus if they were to come within the infection range of an infected individual. Mask mandates do help flatten the curve so that the hospital capacities[14] are not overwhelmed. We distribute 3 types of masks and have 4 cases configured for the population evenly:

1. No Masks: Efficiency is 0%
2. Cloth Masks
3. Surgical Masks
4. N95 Masks

We run 2 simulation experiments for COVID-19: no mask mandates, mask mandates starting at frame 70. Our results are reported in Figure 30, 31.

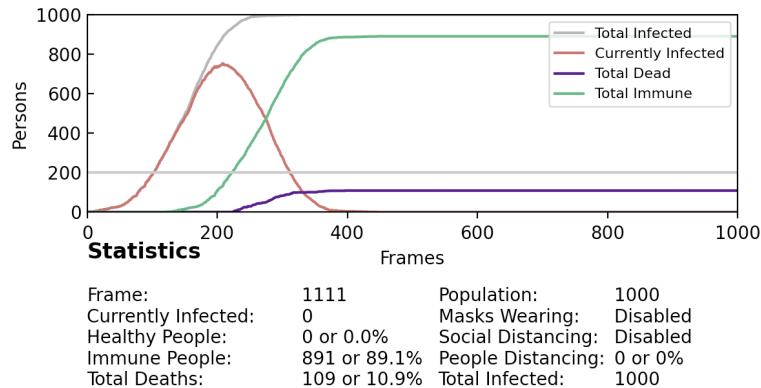


Figure 30: Virus Spread with Time for COVID-19 with No Mask Mandate

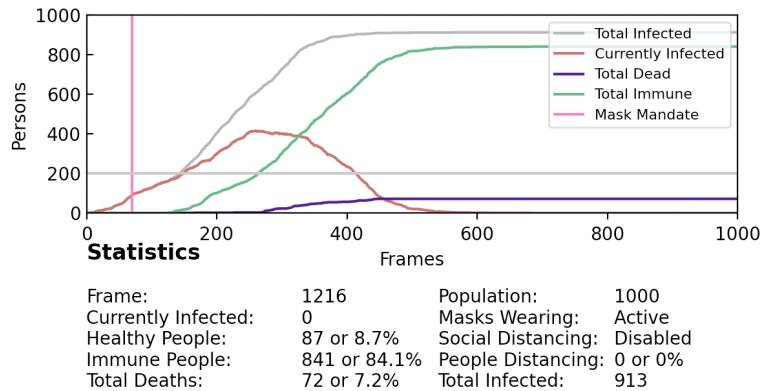


Figure 31: Virus Spread with Time for COVID-19 with Mask Mandate starting at Frame 70

As one can see, mask mandates don't change the total number of infections by much, however they do affect the death rates. By wearing masks, people flatten the curve of current infections, thereby, allowing hospitals to function more or less under their full capacity and having space for people to be hospitalized if they need it. In Figure 31, the number of deaths are 72 (7.2%) and the curve is more evenly spread out through time. At one point no more than 40% of the population is infected. However, in Figure 30, the curve for current infection is steeper, reaches the peak much earlier thereby overwhelming the hospital capacity and causing more deaths. In this scenario, 109 (10.9%) people died which is much higher than for a mask mandate.

7.7 Healthcare Capacity

The capacity of the healthcare facilities can affect the death rates drastically. If the virus grows faster and no control measures are taken, there are a large portion of people infected at the same time and the healthcare capacities are overwhelmed. This leads to a lot of people who may need medical intervention not get the help they need promptly or at all, thereby, causing their demise[15]. We found that hospital capacities are generally set as a percent of the total population. We run 3 simulation experiments: healthcare capacity is 20% of the population, healthcare capacity is 50% of the population, healthcare capacity is 80% of the population. The results are reported in Figure 32, 34, 34.

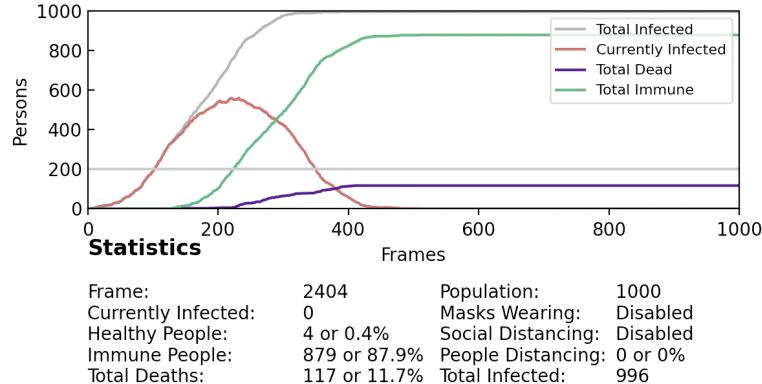


Figure 32: Virus Spread with Time for COVID-19 with Hospital Capacity = 20%

In Figure 32, we can observe that with the hospital capacity being 20% of the total population, the number of deaths were quite high i.e 117 people (11.7%). This implies that a lot of people couldn't get the medical intervention that they needed.

In Figure 33, with the hospital capacity being 50% of the total population, the number of deaths decreased as the hospital could hold a larger amount of people. The number of deaths in this scenario were 76 people (7.6%).

In Figure 34, we have hospital capacity to be 80% of the total population. This is an unlikely scenario because in a normal case, no one would expect more than even 40% of the people to be hospitalized at the same time. However with the current COVID-19 pandemic, the surges in the

infection affecting each individual differently, the hospitals are running at capacity and there are deaths which could've been prevented otherwise. In the Figure

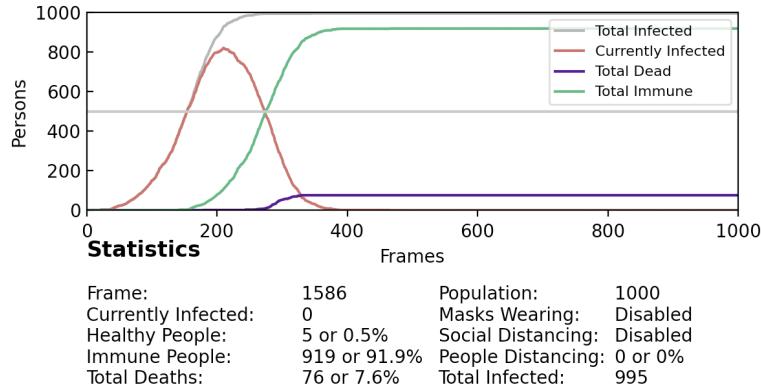


Figure 33: Virus Spread with Time for COVID-19 with Hospital Capacity = 50%

34, one can see that only 35 people or 3.5% of the population died as everyone got the medical intervention they needed because the hospital capacity was 80% of the total population.

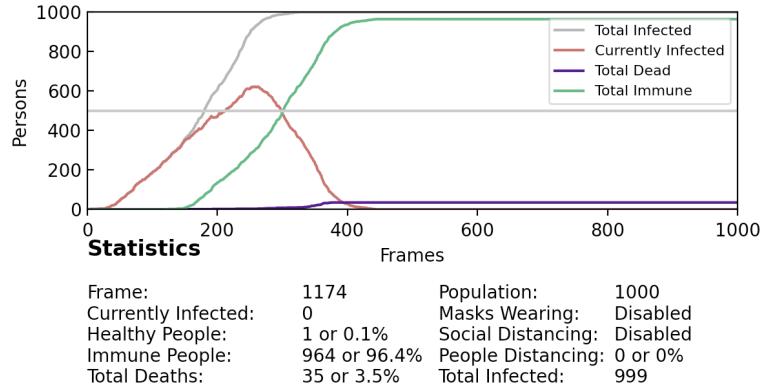


Figure 34: Virus Spread with Time for COVID-19 with Hospital Capacity = 80%

7.8 Outbreak Control

After seeing our analysis above, one may ask: how can we control the outbreak as the vaccine is still not widely available to us? We here following the WHO guidelines[16] try to show that if social distancing and mask mandates are implemented early and followed by a large portion of the population, it is very much possible to be safe from the most devastating of outbreaks. For this experiment, we create a virus with a higher R value = 8, a lower K value (0.1) which means that there are more chances of superspreader events and 1 person can go on infect 8 people on an average. However, we make more than 50% of the population practice social distancing and wear masks from an early stage. Our results are reported in Figure 35.

The results in Figure 35 is quite an optimistic but hopeful way of showing that even with a disease

with a reproduction value (R value) = 8 and K value = 0.1, if more than 50% of the people social distance and wear masks from an early stage of the outbreak, the number of cases can be controlled. One can also see that the current cases always remained under the hospital capacity and only 7 (0.7%) people died. In our section for the R values, where no one was social distancing or wearing masks and the R value was 8 (please see Figure 16), the entire population got the infection before frame 200 and 152 people died (15.2%). Here, 60% of the population started social distancing from Frame 150, and started wearing masks at Frame 70. This immensely helped control the spread.

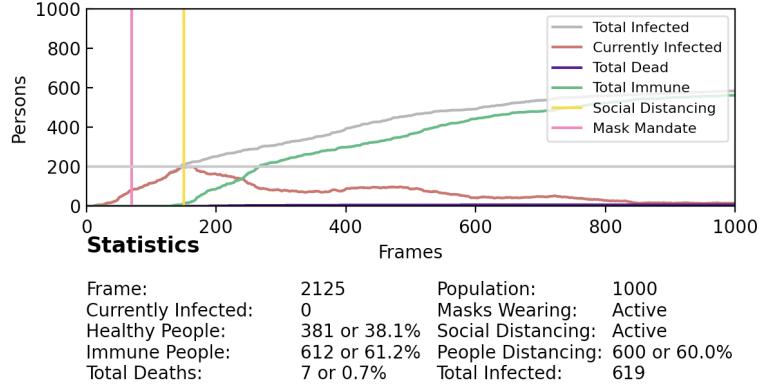


Figure 35: Virus Spread with Time for a Virus with R value = 8, K value = 0.1, 60% of the Population Social Distancing Starting at Frame 150 and Mask Mandate starting at Frame 70

7.9 Influenza vs COVID-19

In our application, we have pre-loaded data for Influenza[6][5] and COVID-19[4][5]. These two viral diseases have very different characteristics. The tables are given in Table 1, 2.

	R value	K value
COVID-19	3.00	0.1
Influenza	1.8	0.9

Table 1: R and K values of COVID-19 and Influenza

	Mortality Rate: 0-19	Mortality Rate: 20-49	Mortality Rate: 50-69	Mortality Rate: 70+
COVID-19	0.003%	0.02%	0.5%	5.4%
Influenza	0.02%	0.05%	0.06%	0.86%

Table 2: Mortality Rate for 4 age groups for COVID-19 and Influenza

We provide an analysis for both by running the simulation. Our results are reported in Figure 36, 37.

The results of our experiment simulations were as expected. As one can see, in COVID-19 virus (SARS-CoV-2), with a high R value and a low K value, there are more superspreader events and

the virus grows rapidly. The entire population gets infected and the death toll 13.6% of the total population or 136 people. In case of Influenza, with a R value of 1.8 and K value 0.9, the spread of the disease is much slower. Influenza having lower mortality rates brings the death toll to a comparatively very low number of only 0.4% or 4 people.

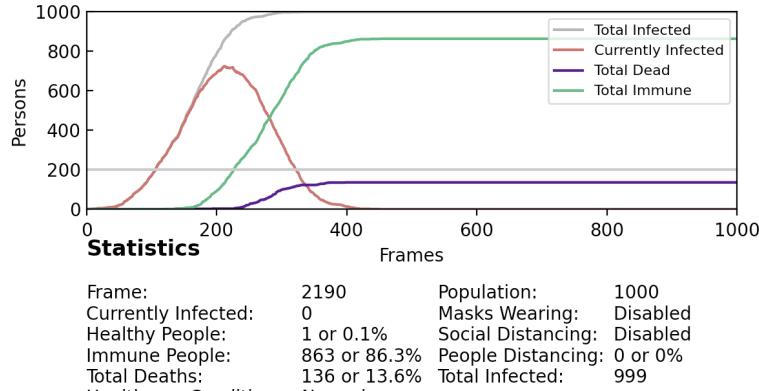


Figure 36: Virus Spread with Time with COVID-19 Virus

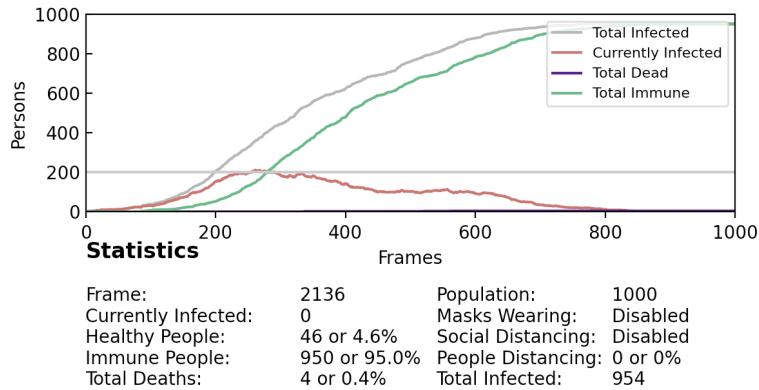


Figure 37: Virus Spread with Time for Influenza Virus

8 Conclusion

We developed an application capable of simulating a virus spread that took into account various factors such as the virus specific characteristics (R value, K value, recovery time, mortality rates, mask effectiveness) and population specific characteristics (population size, social distancing, mask mandates) to demonstrate an outbreak and its impact. From our analysis above, we found that the simulation model demonstrated results which were on par with our expectations, observations in real world, and our research. Some of the conclusions that can be made using our simulation models are:

1. R and K value of a virus together can have a huge effect on how fast the virus spreads and

its exponential growth. When R value is less than 1, that implies that the virus outbreak is contracting and will disappear by itself soon. K value affects the individual R value of a person; how many people the infected individual will spread the virus to. With a lower K value, the individual R value shows more variance and a person may thereby spread the virus to a lot of people (such as the case in superspreader events) or not spread the virus at all (being a deadend in terms of the spread). The most devastating impact of a virus occurs when a virus has a very high R value and a very low K value. We analyzed the individual impacts of R and K value in section 7.2 and 7.3.

2. Population density also plays a role in how fast a virus spreads. If a large amount of people are situated in small area, the virus will spread much faster as compared to a more sparsely populated region. This is consistent with our research [11]. Of course, the impact of population density can be reduced by control measures such as social distancing and mask mandates which can also be seen in our analysis. Urban areas are in fact linked with lower infection rates for COVID-19 because they have higher healthcare capacity, more people social distancing, wearing masks and more awareness [13]. Population density is also not linked with high death rates because the healthcare capacity is a function of the population size and hence, scales up accordingly allowing people to get the medical intervention that they may need[12]. The effects of high hospital capacity and control measures have been analyzed in section 7.7, 7.6 and 7.4 respectively.
3. Flattening the curve can help reduce the death rates by allowing healthcare capacity to function normally. This is consistent with our research [17] that the longer the infection can be prevented, the more equipped an area is to deal with it. We analyze this in section 7.4, and 7.6. We demonstrate the effect of healthcare capacity in section 7.7
4. With appropriate control measures in place, even virus with an R value as high as 8 with a potential to have devastating consequences can be controlled early. This requires social distancing and mask wearing to be implemented as early as possible. We analyzed outbreak control in section 7.8 and analyzed the individual effects of social distancing and mask wearing in section 7.4 and 7.6 respectively. It is imperative that a larger portion of the population follow these control measures in order to prevent higher surges in infections. It is the responsibility of the area officials to put appropriate control measures for the public as early as possible.

List of Figures

1	Simulation on the Viral Diseases Simulator	2
2	Network of spread from the first infection.	3
3	GUI screen user is welcomed with	4
4	UML diagram representing the core modules of the model	6
5	What a normal distribution looks like.	9
6	UI the user is greeted with.	13
7	Dialog boxes in the GUI to edit mortality rates and mask effectiveness	13
8	Dialog boxes when starting simulation.	14
9	How the simulation looks.	15
10	Screenshots providing insight in unittesting.	15
11	Virus Spread with Time for COVID-19 with no Control Measures	16
12	Virus Spread with Time for COVID-19 with Social Distancing and Mask Mandates	16
13	Assuming a virus with R value of 2, the virus grows exponentially	17
14	Virus Spread with Time when R value = 1 and K value = 1	17
15	Virus Spread with Time when R value = 3 and K value = 1	18
16	Virus Spread with Time when R value = 8 and K value = 0.1	18
17	Virus Spread Network Visualization for R value = 1	19
18	Virus Spread Network Visualization for R value = 8	19
19	Virus Spread with Time when K value = 3 and R value = 3	20
20	Virus Spread with Time when K value = 0.03 and R value = 3	20
21	Virus Spread with Time for COVID-19 with No Social Distancing	21
22	Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing starting at Frame 150	21
23	Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing starting at Frame 300	22
24	Virus Spread with Time for COVID-19 with 20% of the Population Social Distancing at Frame 200	23
25	Virus Spread with Time for COVID-19 with 50% of the Population Social Distancing at Frame 200	23

26	Virus Spread with Time for COVID-19 with 80% of the Population Social Distancing at Frame 200	24
27	Virus Spread with Time for COVID-19 with Population Size = 522	24
28	Virus Spread with Time for COVID-19 with Population Size = 1000	25
29	Virus Spread with Time for COVID-19 with Population Size = 3000	25
30	Virus Spread with Time for COVID-19 with No Mask Mandate	26
31	Virus Spread with Time for COVID-19 with Mask Mandate starting at Frame 70	26
32	Virus Spread with Time for COVID-19 with Hospital Capacity = 20%	27
33	Virus Spread with Time for COVID-19 with Hospital Capacity = 50%	28
34	Virus Spread with Time for COVID-19 with Hospital Capacity = 80%	28
35	Virus Spread with Time for a Virus with R value = 8, K value = 0.1, 60% of the Population Social Distancing Starting at Frame 150 and Mask Mandate starting at Frame 70	29
36	Virus Spread with Time with COVID-19 Virus	30
37	Virus Spread with Time for Influenza Virus	30

List of Tables

1	R and K values of COVID-19 and Influenza	29
2	Mortality Rate for 4 age groups for COVID-19 and Influenza	29

References

- [1] Paul van Gent. Python covid-19 ('corona virus') simulation, 2020.
- [2] Grant Sanderson. Continuous uniform distribution, March 27, 2020.
- [3] Grant Sanderson. Exponential growth and epidemics, March 8, 2020.
- [4] Chris T Bauch. Estimating the covid-19 r number: a bargain with the devil? 2020.
- [5] Adam Kleczkowski. Is the k number the new r number? what you need to know, 2020.
- [6] Carrie Reed Manoj Gambhir Matthew Biggerstaff, Simon Cauchemez and Lyn Finelli. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. 2014.

- [7] Timothy F. Leslie Y. Tony Yang Paul L. Delamater, Erica J. Street and Kathryn H. Jacobsen. Complexity of the basic reproduction number (r_0). 2019.
- [8] Wikipedia. Normal distribution, 2020.
- [9] Wikipedia. Continuous uniform distribution, 2020.
- [10] Laura Matrajt and Tiffany Leung. Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease. 2020.
- [11] Patrick M. Tarwater. The effects of population density on the spread of disease. 1999.
- [12] Sadegh Sabouri Shima Hamidi and Reid Ewing. Does density aggravate the covid-19 pandemic? 2020.
- [13] Matthew Keegan. Why cities are not as bad for you as you think. 2020.
- [14] Wei Lyu and George L. Wehby. Community use of face masks and covid-19: Evidence from a natural experiment of state mandates in the use. 2020.
- [15] Md. Mazharul Islam Jahidur Rahman Khan, Nabil Awan and * Olav Muurlink6. Healthcare capacity, health expenditure, and civil society as predictors of covid-19 case fatalities: A global analysis. 2020.
- [16] WHO Interim Guidelines. Infection prevention and control of epidemic- and pandemic-prone acute respiratory diseases in health care. 2017.
- [17] Sarah Zhang. The simple rule that could keep covid-19 deaths down. 2020.