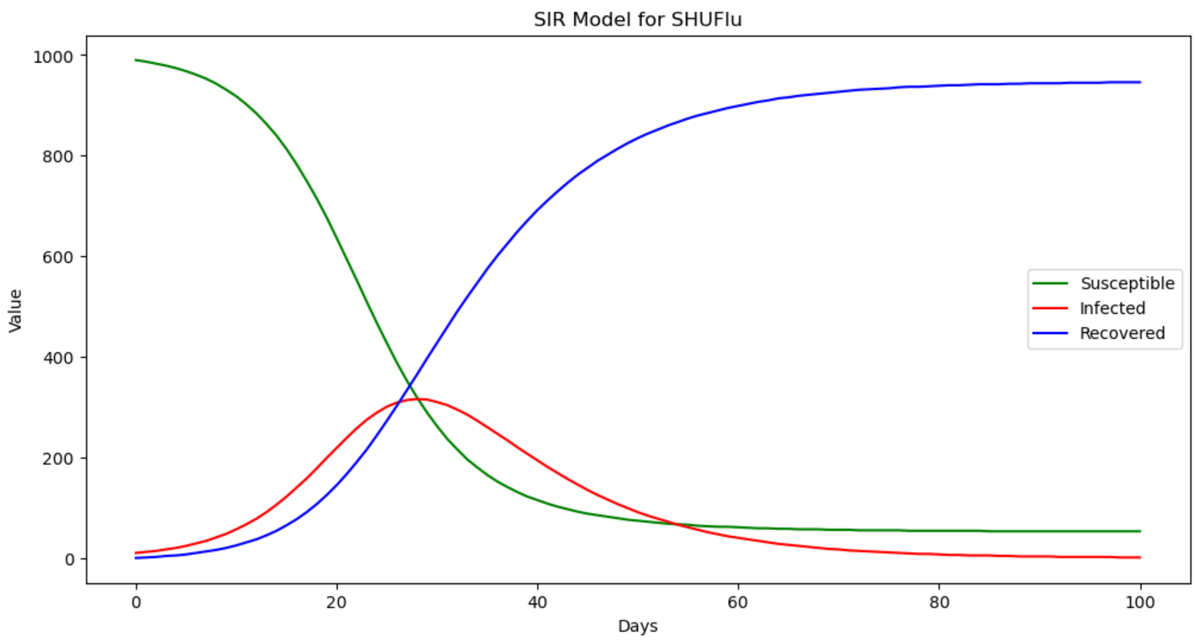
Final Project for Mathematical Modeling Course: Analysis of Epidemic Spread Using Mathematical and Computational Models

Benjamin Nicholson
Seton Hill University



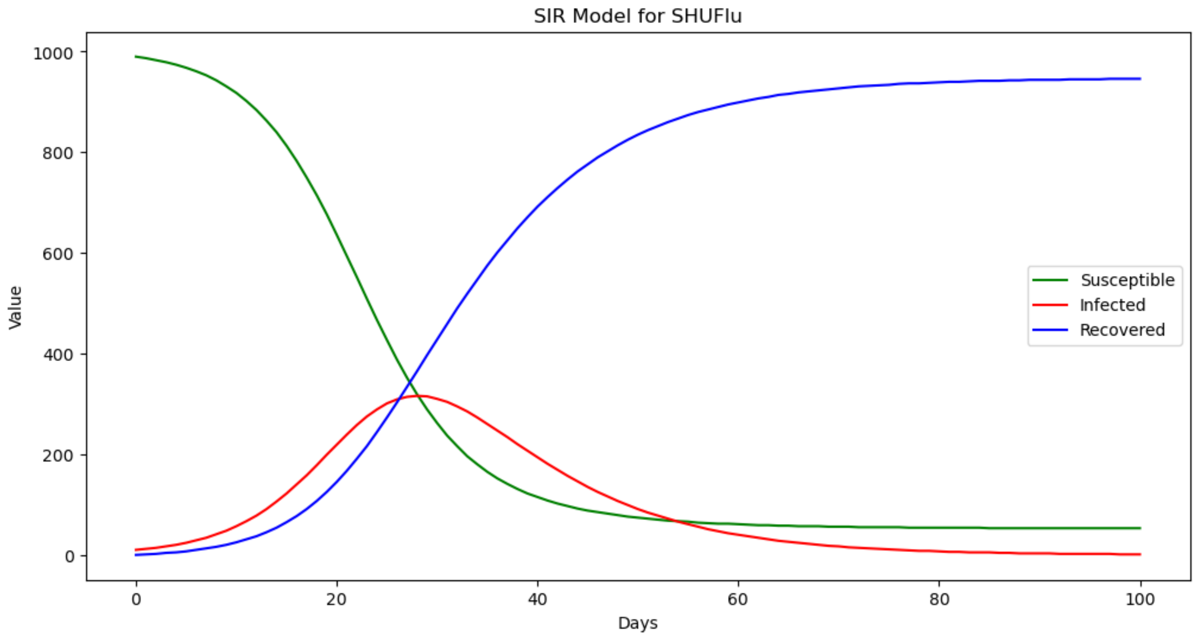
GitHub Link: https://github.com/bennicholson2/SIR-Model

Introduction

Infectious diseases such as SHU Flu have spread throughout populations for millennia however in 2024 due to the vastly growing understanding of their spread they can be understood on a deeper level. The method of understanding that will be discussed in this report is the SIR model which incorporates susceptible, infectious and recovered individuals into a system of relationships. Through understanding SIR models, an understanding of how the disease spreads is understood which can help reduce the possibly devasting impacts such as the recent Covid-19 Pandemic which begun in 2020. The goal of this project is to understand how the spread of SHU Flu has impacted the given population with the goal to predict disease spread, assess intervention strategies, and understand the dynamics of the epidemic through different scenarios.

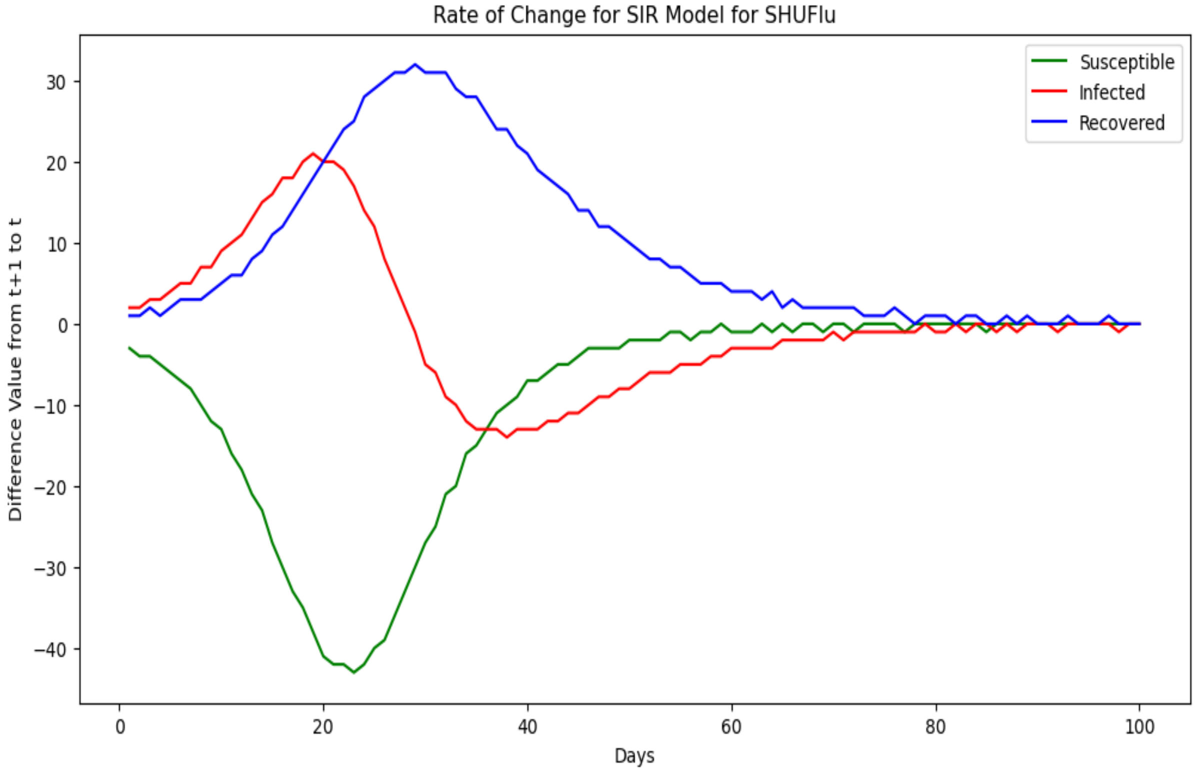
Exploratory Data Analysis

The data files that have been provided consist of susceptible, infectious and recovered populations over 100 days. As well as a file which has the contacts between 1000 people. The majority of the analysis will be utilising the SIR model data. The first step is to model the historical data to understand how it has spread through the population and to understand its impacts.

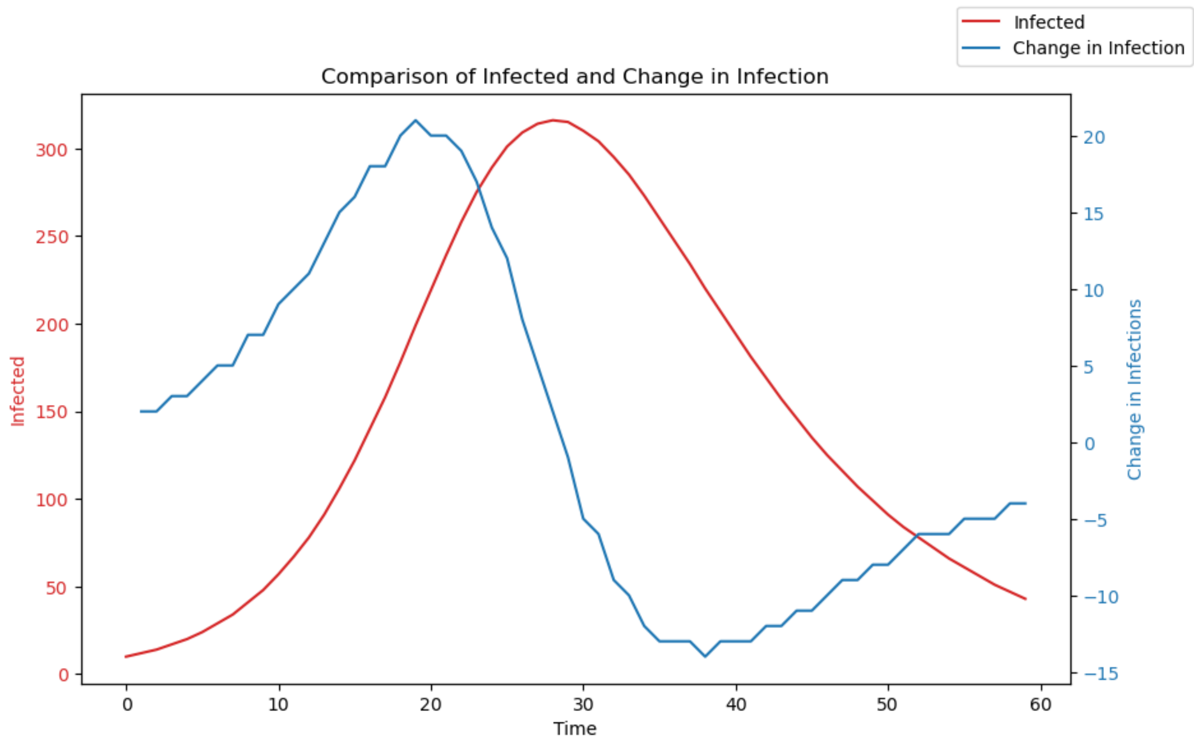


Susceptible rates start high and are always decreasing which means that there is no reinfections or new people coming into the population. Essentially there are 1000 people and once you have been infected you have immunity. There begins with 10 infectious people with a maximum number of infections taking place at day 29 where there is 316 infectious people. The recovered rate should be the safe as the infected rate as there is no death so populations will have the disease but eventually recover.

There is only one wave as there is an assumption of no new populations being accessed to the disease as well as there being equal spread of infection across the population. To understand the wave of infections it is important to observe how the change in infectious people takes place from day to day, which is displayed below.

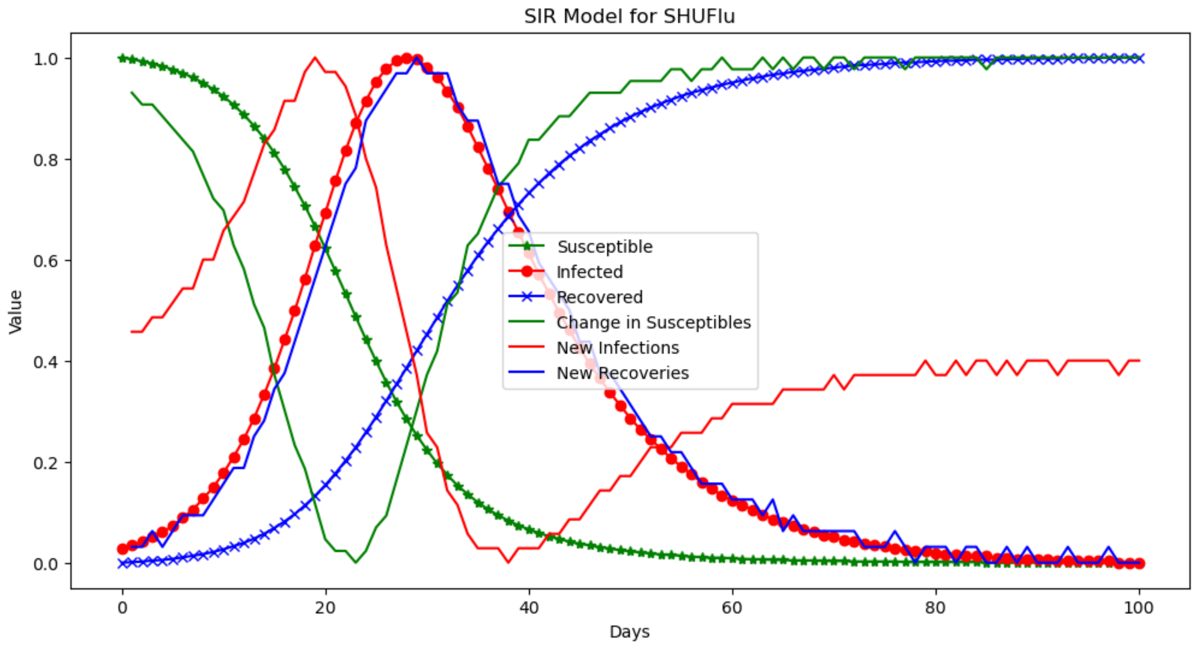


Infectious rates changing from day to day peak at around day 20 (8 days prior to the peak of infections) with a rate of 21 new infections which is about 2% of the population per day at the peak. The negative wave peak is the maximum rate of infections decreasing. Where as the recovered and susceptible peaks only occur once.



There is an 8 day delay in the peak of the change of infections with the current infections this highlights the importance of public health interventions when the rates of infection get higher rather than the active infections. This is because the growth can take place at an exponential rate. The reason for the lack of exponential growth in the case of the SHU Flu is because there is not a high enough population for the virus to sweep through. The highest rate of infection was 31.6% with 316 people being infected with the disease at one point. The reason this was the peak was because the number of susceptible people was not high enough for the infection to continue. However in the case of Covid-19 where billions of people are exposed it makes sense as to why places such as the USA had millions of infections over small periods of time. There was only 53 people not infected by the SHU Flu in the population of 1000 people, once again highlighting the importance of health interventions which is going to be discussed in 'Simulation and Analysis' of this report.

By looking at the SIR data and the rates of change you can observe how these factors are related to one another. Taking the normalised data allows for a direct comparison of how maximum and minimum of waves take place.



New recoveries and infected follow almost exact trends which means that the rate of new recoveries will follow a similar pattern to new infections. This means there is not too long of a delay in recoveries. If there was a longer delay you would expect recovery rates to have a bigger lag. Susceptible and recovered have inverse relationships to one another. New infections and Susceptible have opposite relationships to one another again.

Model the Spread of the Disease

The rate of change of the SIR model can be described by the differential equations.

$$\frac{dS}{dt} = -\frac{\beta I S}{N},$$
$$\frac{dI}{dt} = \frac{\beta I S}{N} - \gamma I,$$
$$\frac{dR}{dt} = \gamma I.$$

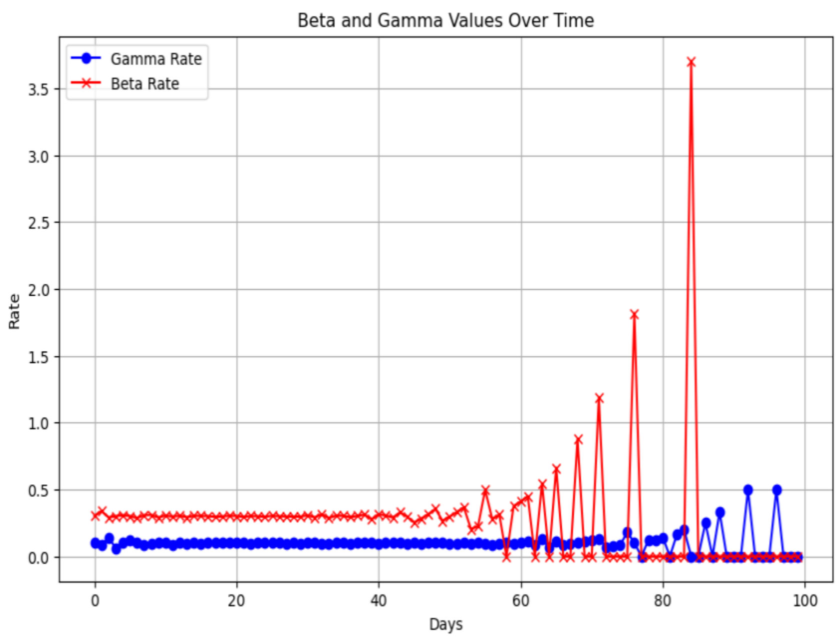
These equations can be used to model how rates of susceptible, infectious and recovered numbers change from one interval to the next. These includes the parameters Beta and Gamma. The Beta rate is the transmission rate of the disease which is the probability of a susceptible individual becoming infected upon contact. This is also used to find the reproduction rate which is the number of people that are infected for each infection. If you have the average duration of infection you can find the reproduction rate. Meanwhile the gamma rate is the recovery rate of individuals are infected with a disease. Essentially is represents how quickly individuals recover from the infection and become immune.

Using the historical data from the data file, you can solve for the Beta and Gamma rates for each day by transposing the listed differential equations:

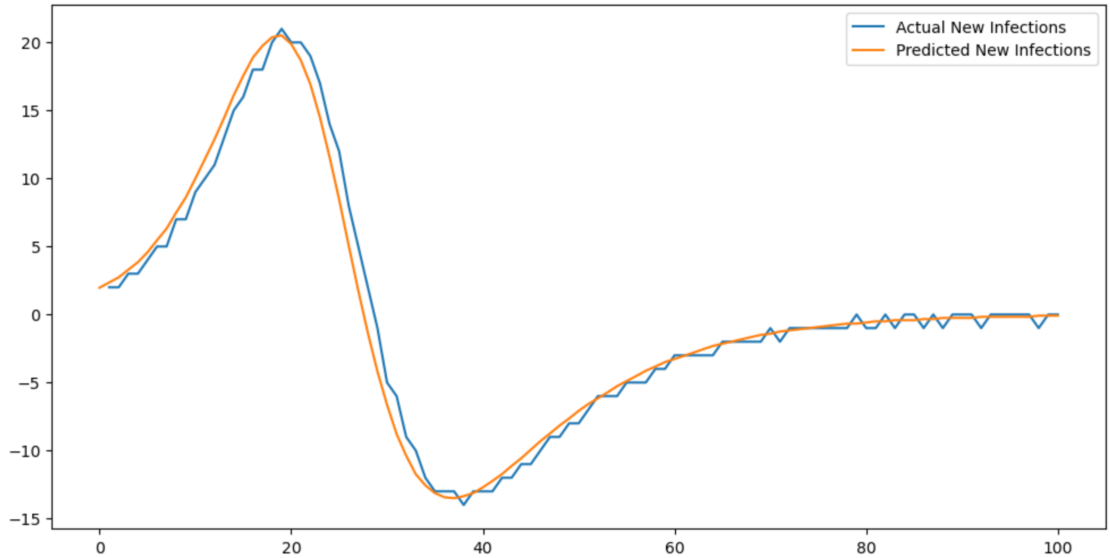
$$\beta = -\frac{N}{SI} \cdot \frac{dS}{dt}$$

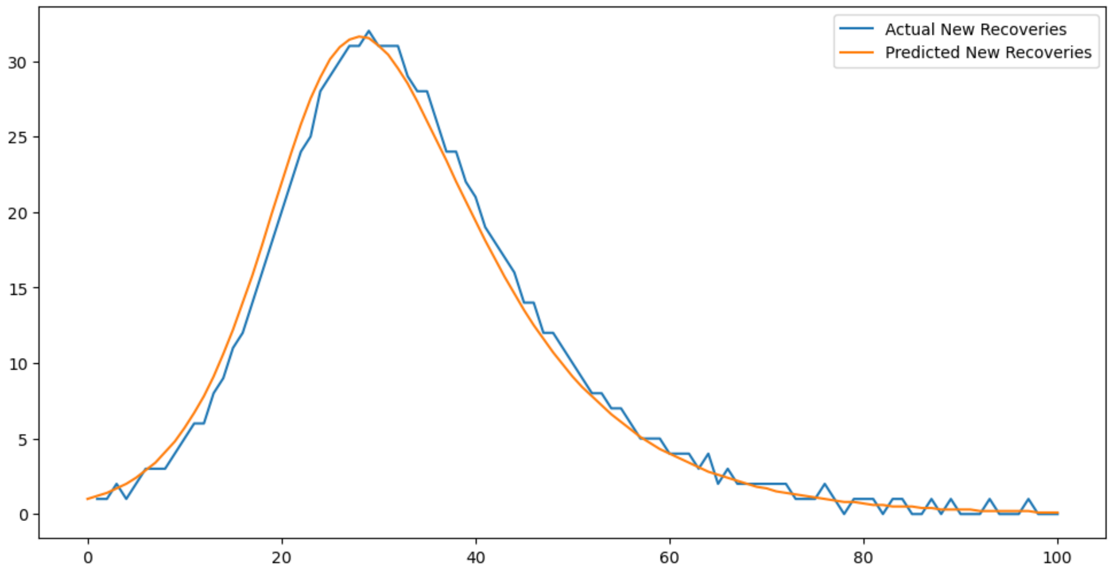
$$\gamma = \frac{1}{I}\frac{dR}{dt}$$

The Beta rate and



Gamma rate can be observed on the above graph as being about 0.3 and 0.1 respectively. Deriving these values allows for modeling of the SIR model over time without needing the historical data. Allowing you to apply the model to different populations and different changes in certain rates over time to change how the model changes. When comparing the derived model with the actual result there is an MAE of less than 1.

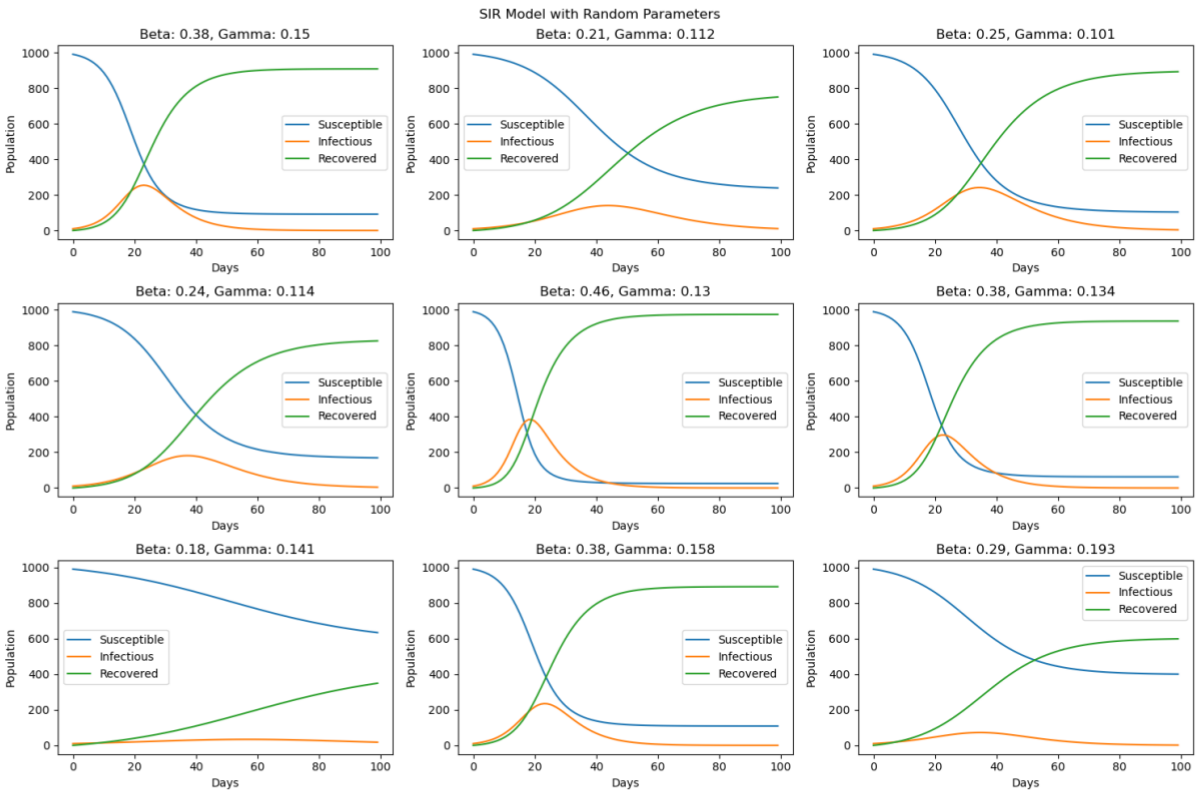




Stochastic Modeling

Beta and gamma rates heavily impact how the disease spreads through a population. A higher beta rate with a lower gamma rate will see higher infection rates and vice versa. To build an understanding of how the SIR model moves with changes in beta and gamma rates these parameters.

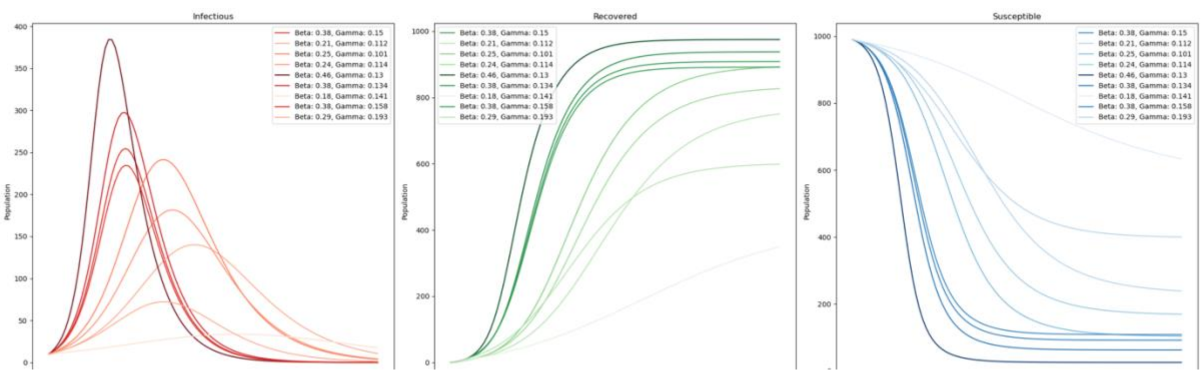
- Beta mean: 0.3 with a standard deviation of 0.1
- Gamma mean: 0.1 with a standard deviation of 0.05



The grid allows you to explore a variety of different relationships that take place.

- Higher beta rates with lower gamma rates (2,2)
- Similar beta and gamma rates (3,1)
- Lower beta rates with higher gamma rates (1,2)

These rates can be put on the same plot. The darker the colour the higher difference in beta vs gamma rate which can further prove how different rates show different models in the SIR.



Predictive Neural Network

Neural Networks act as a black box which makes it very difficult to understand how the patterns that are found amongst layers in neural networks operate. Understanding that a neural network is for results and less for modeling using specific parameters allows for a proper use of its powers. You cannot expect to understand the intricate patterns but it will give very accurate results and this is observed using 90% of the historical data as training data while testing the model on the last 10%.

	Infected	Neural Network	Fold	Day	Susceptible	Recovered	Difference from NN	Percentage Difference from NN
2	1	2.44965	1	99	53	946	1.44965	144.965000
3	12	11.57500	1	74	55	933	-0.42500	-3.541667
4	4	4.67785	1	86	53	942	0.67785	16.946250
5	6	6.06552	1	82	54	940	0.06552	1.092000
6	219	213.07600	1	20	636	145	-5.92400	-2.705023
7	40	39.07520	1	60	61	899	-0.92480	-2.312000
8	15	14.58230	1	71	56	929	-0.41770	-2.784667
9	106	96.18190	1	14	840	54	-9.81810	-9.262358
10	2	3.16961	1	92	53	944	1.16961	58.480500
11	84	84.93050	1	51	72	844	0.93050	1.107738

You can see how close the values from the neural network are from the real data. This displays the power of the neural network, however an understanding of how its predictive nature can be used is not understandable due to it being a black box model.

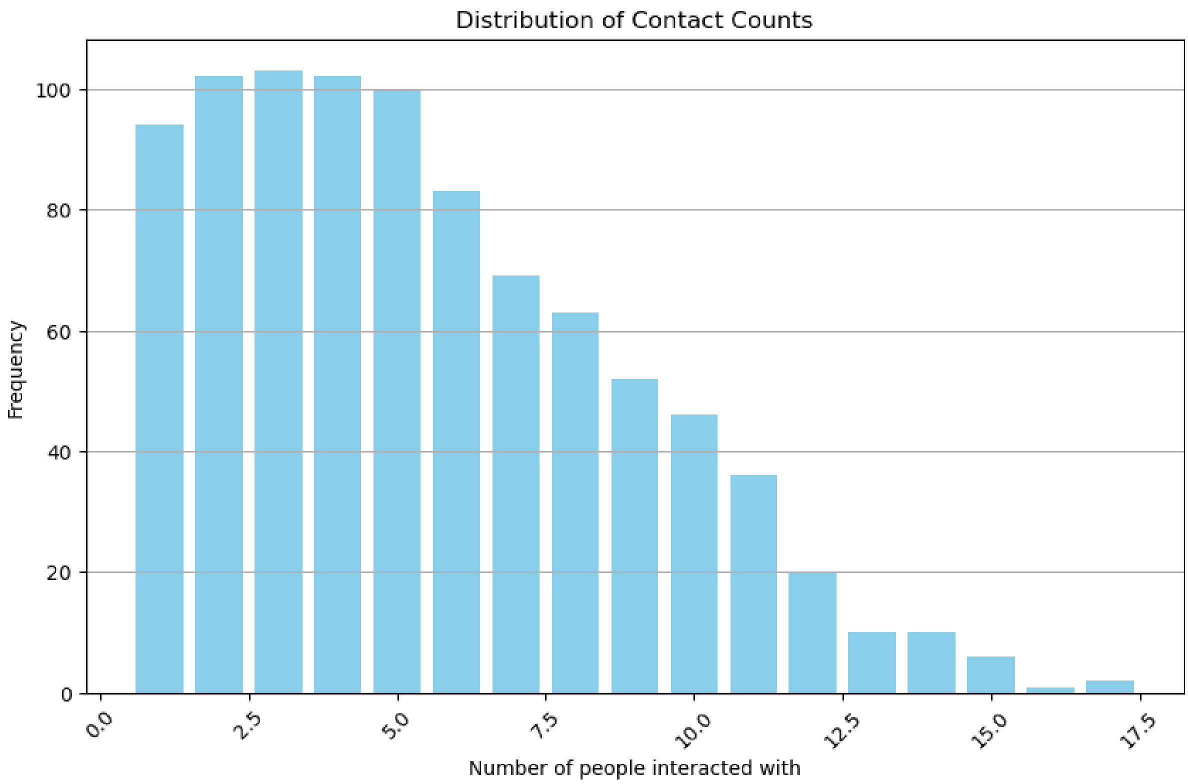
Graph Theoretical Analysis

Exploring the contact tracing data, a number of people super spreaders have been identified, where they infected interact with a much higher number of people than the rest of the population. They are listed to the right.

	Number of People Infected
Person1	
52	17
137	17
47	16
20	15
139	15
283	15
73	15
66	15
16	15
115	14

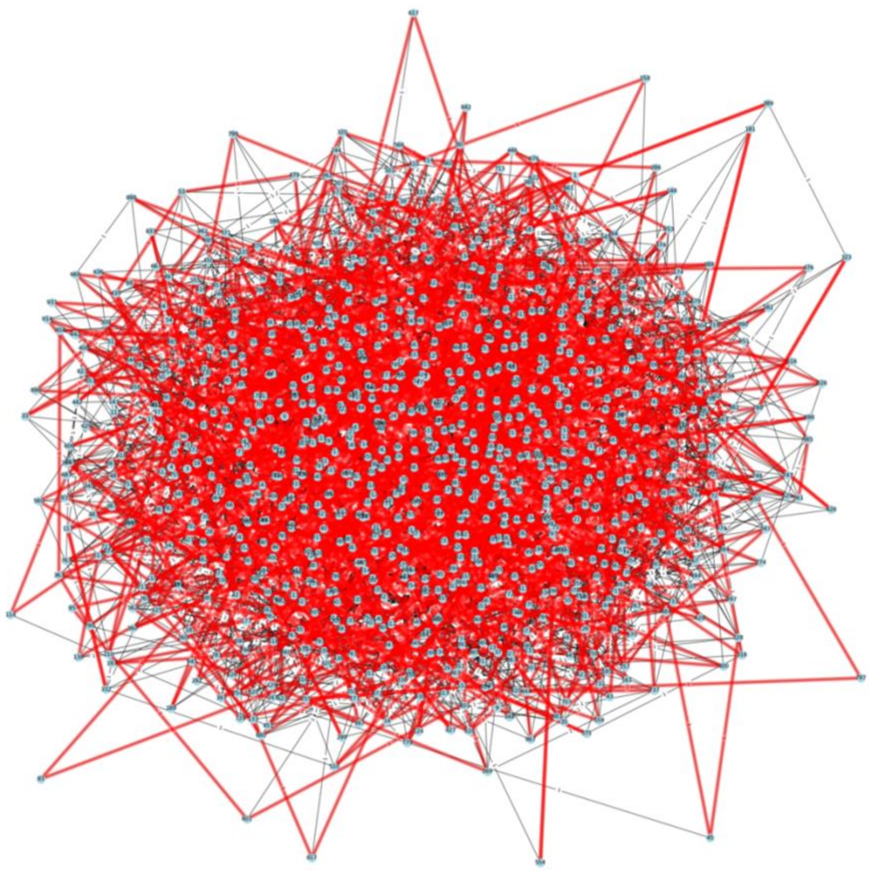
Summing the number of people who share the number of people that they are in contact in can help understand the spread of the distribution. The graph below illustrates the spread of the data.

	Number of People with Number of Contacts
Number of Contacts	
1	94
2	102
3	103
4	102
5	100
6	83
7	69
8	63
9	52
10	46
11	36
12	20
13	10
14	10
15	6
16	1
17	2



The distribution reflects a positively skewed data where there is a shift in the distribution to the left. This is because there are a few people such as the super spreaders that have a much higher than average number of contacts which is why the mean is shifted to the right where the median is further to the left.

Mapping how the contacts interact with one another when there is 1000 people proves to be quite complicated.



Simulation and Analysis

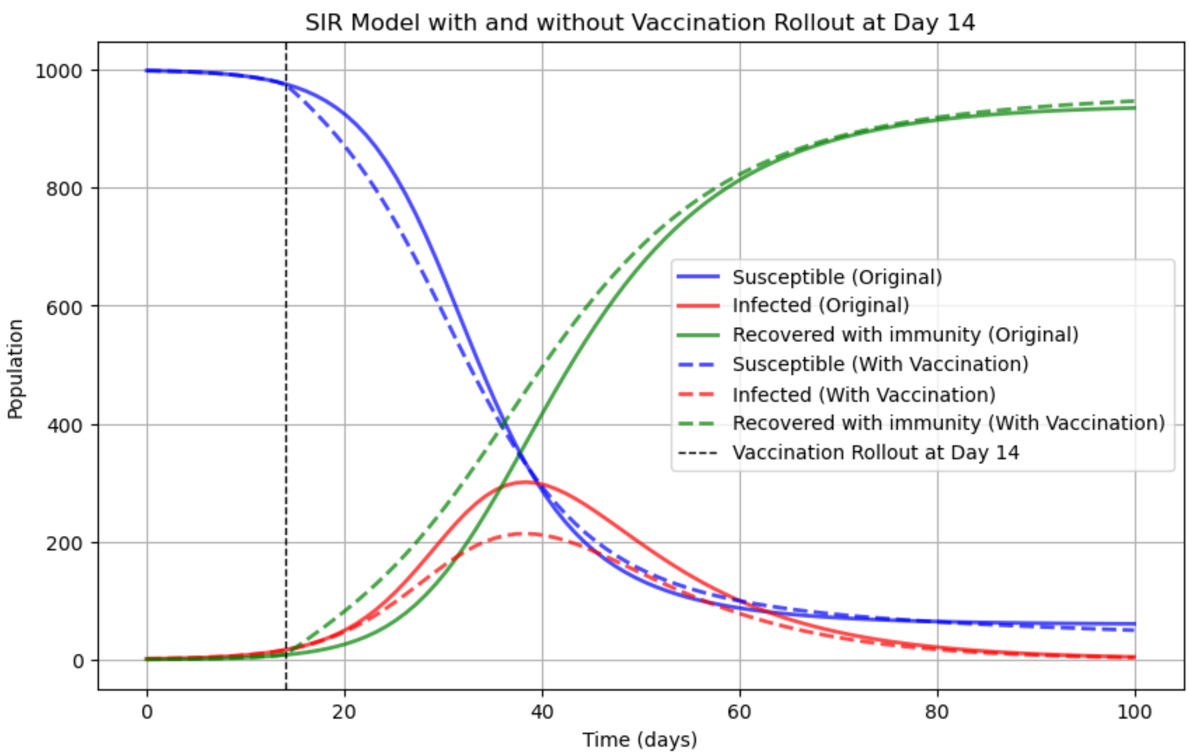
This section is going to look at the following public health interventions to help 'flatten the curve' which was a common feature of the Covid-19 Pandemic.

1. Vaccination
2. Social Distancing
3. Lockdowns

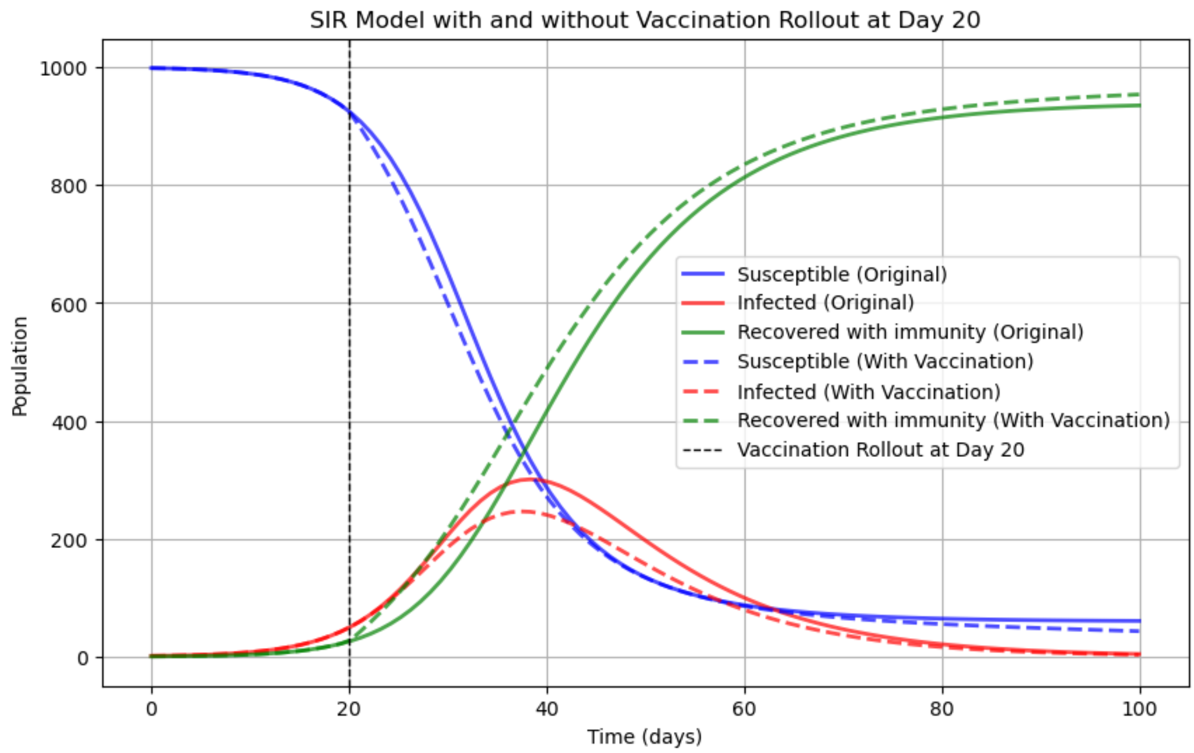
Vaccination rollouts are a common public health intervention to reduce the number of people who can be infected by a particular disease. Exposure to the virus in a safe level will help create natural antibodies effectively you are given the disease in a small enough dose to not get sick but will know how to fight off the disease. This can be measured in the SIR model by adding another equation to the system. This is going to be a vaccination rate * the number of susceptible people which is taken away from the susceptible number of people. This happens because those with the vaccine are essentially removed from the system.

To visualise the importance of an earlier intervention from a health point of view the report is going to look at different days of the time interval and observe how that would change the number infections and relationships of the SIR model.

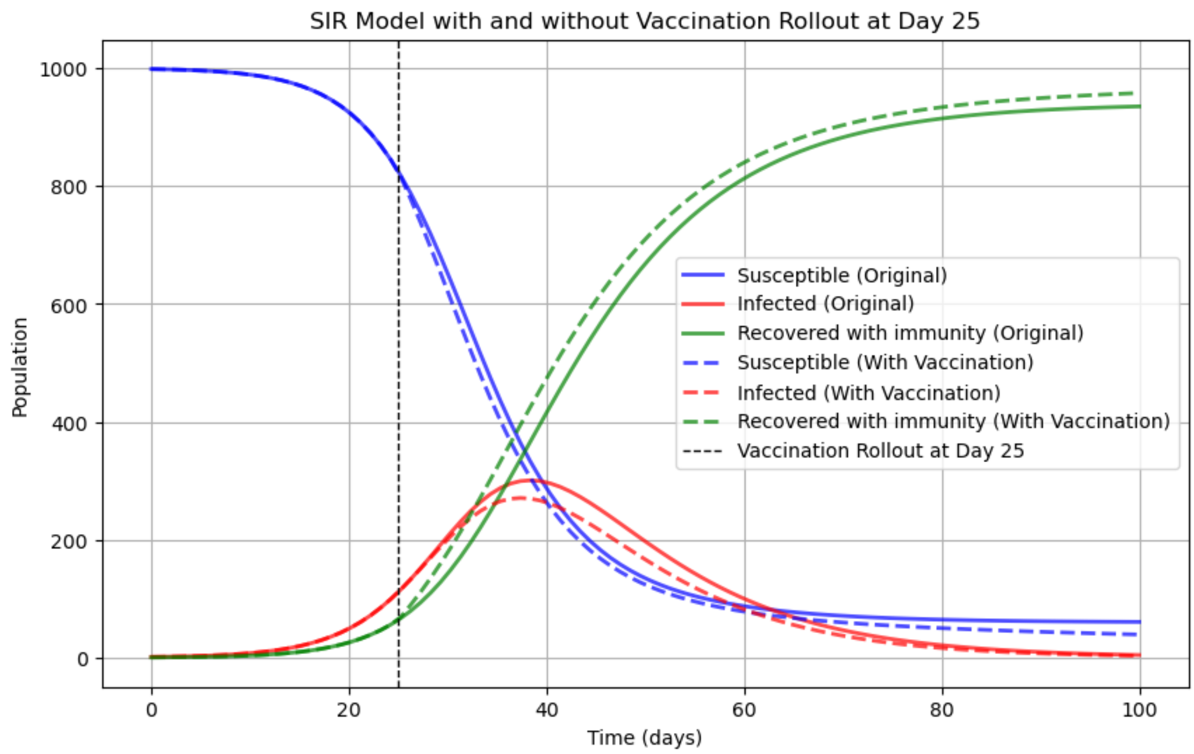
10% of population has disease at day 14



20% of population has disease at day 20



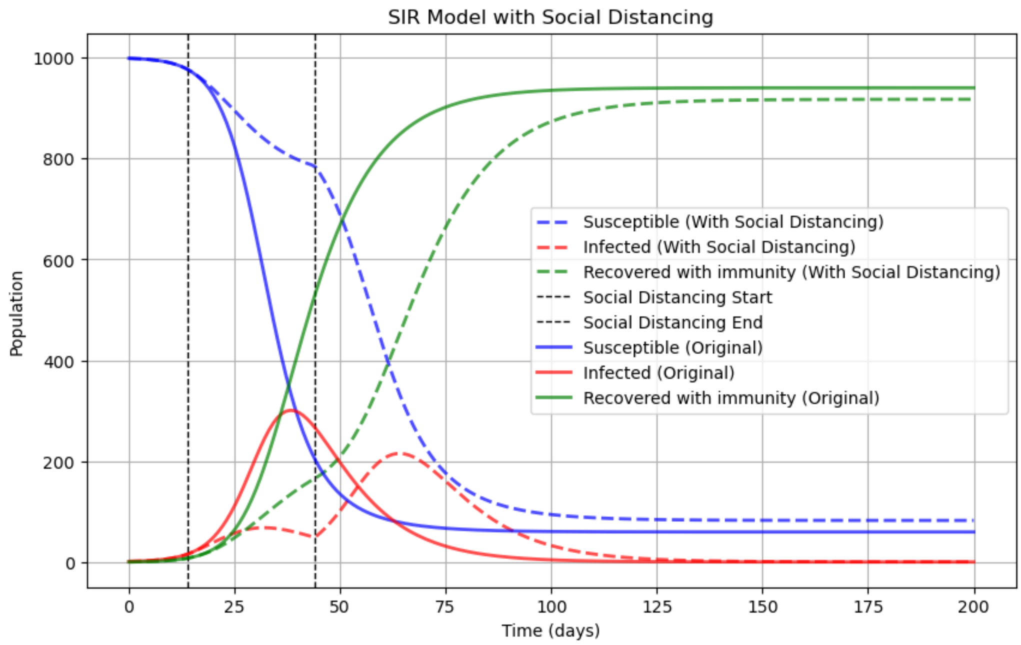
30% of population has disease at day 25



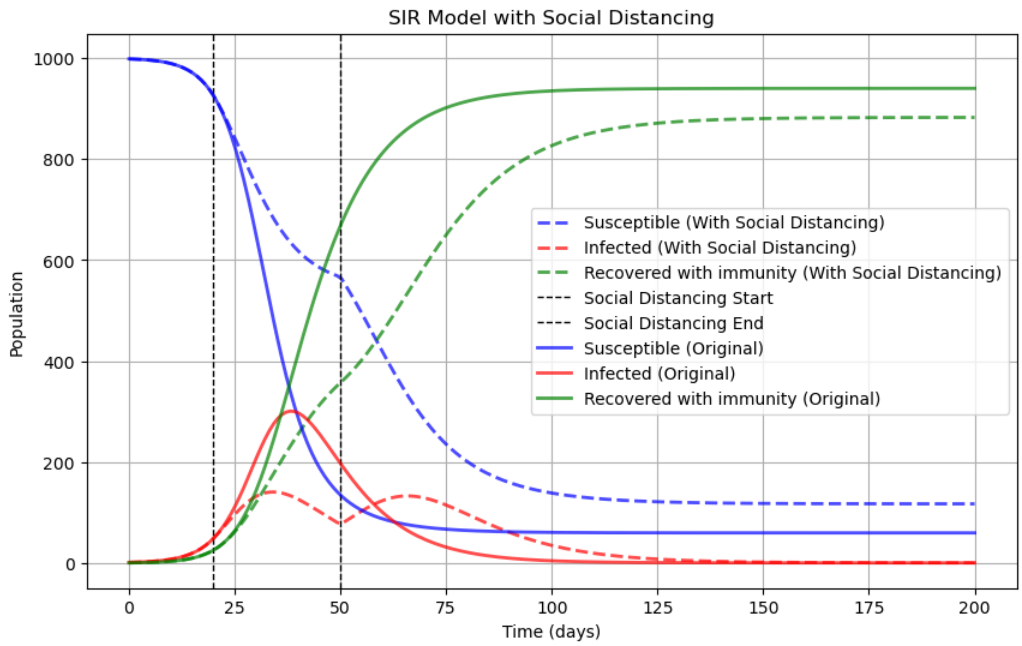
Notice that the biggest difference was when the vaccination took place earlier. The curve was flattened to where there was only 20% of the population at one point in time with the disease compared to the 31.6% of the population which could be the difference between access to ICU's and not. This proves its helpfulness.

Social Distancing is the next public health intervention that can take. This is the encouragement for people to avoid coming within 2 metres of one another which is where the transmission rate is the highest. As explored before this reflects the beta rate, so when this is dropped the number of infections also drops. To get a better reflection of how social distancing might work in a population I am going to incrementally drop the transmission rate as it becomes a trend to socially distance which increases its effectiveness compared to at the start where people usually forget. The social distancing rule will stay in place for 30 days.

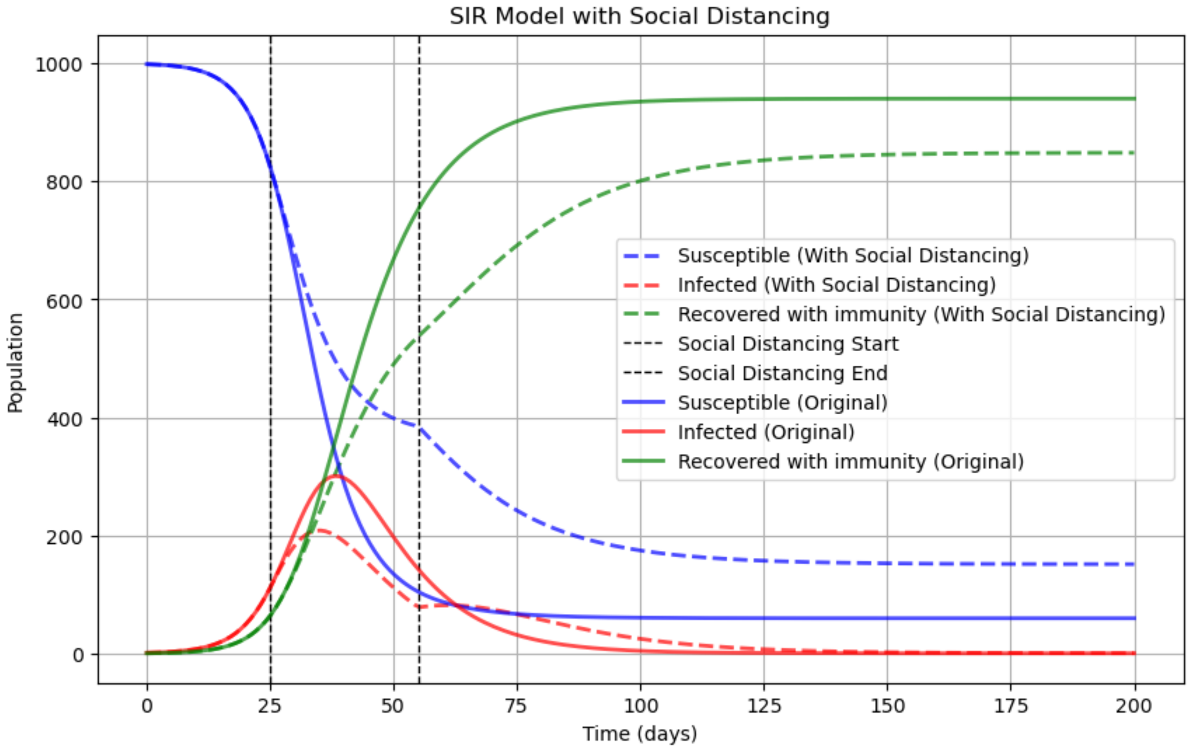
Social distance from day 14-44 when 10% of population has disease



Social distance from day 20-50 when 20% of population has disease



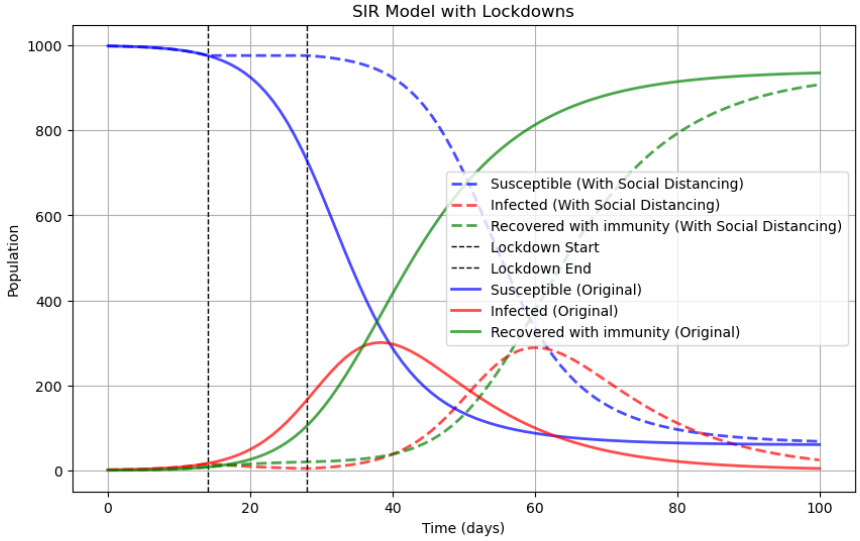
Social distance from days 25-55 when 30% of population has disease



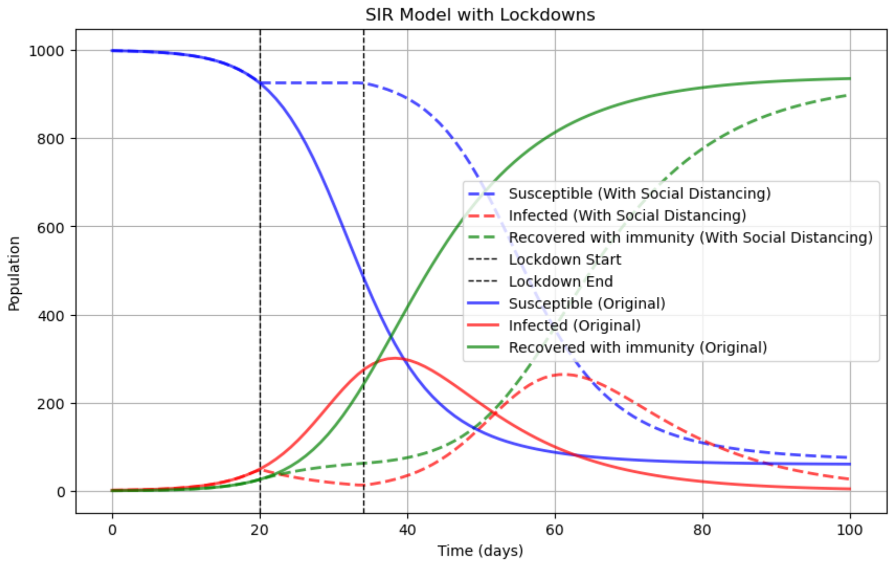
Observe that social distancing simply spaces out the infection, having too high of social distancing to then releasing back into public is not going to help as they are still susceptible to the disease. It appears that it is best at around 20% of the population to implement this public health intervention.

Lockdowns are the strictest public health intervention and was used heavily during the Covid-19 pandemic. It was proven to be a very effective tool from a health point of view but from a social and economic point of view it was extremely detrimental to a lot of global economies and populations, likely to last generations. This can be simulated by instantly making the transmission rate to 0 as the number of people who are interacting with one another is far lower. The lockdown will start at those same intervals but will only last for a total of 14 days.

Social distance from days 14-28 when 10% of population has disease



Social distance from days 20-34 when 20% of population has disease



Social distance from days 25-49 when 30% of population has disease