

Creation and Use of a Coronavirus Simulation for Spread Analysis in a Closed Environment

By: Hyun Min Kim

Abstract

This project aimed to simulate the spread of the coronavirus by using a spread simulator to analyze how different preventive measures impacted the rate of spread within a community. By creating a simulator as a web app, human movement was visualized, and datasets with different combinations of preventive measures were collected. Multiple calculations were performed on the collected datasets to find the optimal values to represent each subset of data. The study found that preventing travel and contact with other individuals resulted in the slowest rate of infection while washing hands had no significant effect.

Introduction

The understanding and analysis regarding the rate of the spread of a virus is a crucial step towards the containment of a worldwide pandemic. Around late 2019, a newly introduced strand of the coronavirus started spreading around the world at a rapid rate. Although suspected to have originated in China's Hubei province, researchers around the world are not certain of its origin at the time of writing of this paper. When the virus enters the human body, it becomes classified as a disease that is called "Covid-19". Because there are no known cures and vaccines for this disease, leaders and government officials around the world have taken various approaches to deal with its spread. These leaders around the world took the national interest into consideration and created measures based on it. In the early months of the spread, many governments put preventive measures aimed to control the spread of the virus. The measures caused many changes around the world and devastated many economies in the world because many businesses and stores were shut down, leading to fewer profits and unemployment for those involved in jobs that require physical human contact. Around mid-2020, some countries like the US and Belarus strayed from this strict approach and changed to a lax one, trusting their people to make decisions regarding their health themselves. On the other hand, countries like South Korea and New Zealand continued their enforcement of strict measures to control the spread of the virus, aiming to lessen the rate of transmission and to find the origins of the spread.

Some regulations put in place to control the spread are the mandatory quarantine of people who enter the country, the requirement of wearing masks, and much more. The effectiveness of these measures in containing and lessening the transmission of the virus has been debated between groups of people. Some claim that the enforcement of such regulation is a waste of resources because it has no impact on the spread while others believe that such measures are necessary for slowing down the spread and containing the virus. As studies about the possible methods of coronavirus transmission were conducted, it became evident that the use of certain methods such as masks would lessen the spread. The World Health Organization (WHO) stated that saliva and various secretions from the body can travel through the air and infect other individuals ¹¹. Because the WHO and some government agencies are not willing to implement these measures as law, the organization strongly encourages physical distancing among individuals and wearing masks through its Q/A page about the coronavirus. Researchers are currently creating modals and visualizations to study the spread of the virus in multiple regions of the world and to help the public understand the current state of the pandemic.

This project aimed to build on the various studies currently being done regarding the coronavirus spread in an interesting way. The purpose of this study is to simulate the spread of the coronavirus in a closed community, providing visualizations of the simulated spread and data for analysis on the rate of spread. A web app of the simulation was developed for the ease of access by the public. Various functions and spread logic were implemented within the code to replicate human movement and interaction.

Methods/Procedure

Tools Used:

- Vanilla Javascript
- CSS
- HTML
- Node.js
- Express.js

- astar.js
- Chart.js
- Bootstrap 4
- Bootstrap-Modal

*The full code is located in <https://github.com/gomgom03/Corona-Project>

*The project is deployed online at coronaspreadsim.me

Programming Overview

The project website was deployed on <website> and the backend routing was done with Express.js. All HTML elements were styled and animated with Bootstrap 4. The user cannot access the variables and functions on the Javascript console because all of them are stored within the IIFE (Immediately Invoked Function Expression) scope. The full code is available in the Github link on the right corner of the website.

Creating the Spread Simulation

Instantiation of Variables

All variables and functions in the Javascript file for the simulation page were stored within the IIFE scope. All the configurable variables can be seen by clicking the configuration options button on the simulation page. However, all other variables that aren't configurable but can be changed when adding to this project like the **pixelSize** variable are defined in the scope as individual variables. The Javascript object that holds most data is the **world** object which contains the <canvas> element, the positions of all human objects, and all buildings. Most interactive HTML elements such as buttons were saved as **const** element objects and their respective event listeners were attached to them. All imported functions and variables from Chart.js, Bootstrap.js, and astar.js are stored globally and available for usage within the IIFE scope. The **param** object is a crucial part of the simulation, providing the basis on various factors within the simulation such as the probability of coronavirus transmission and whether the simulation will be run on Map 1 or Map 2. The various probabilities within the **param** object were found through the Center for Disease Control (CDC), World Health Organization (WHO), and Healthline ^{2,3,5,10,12}.

Starting the Simulation - Setup

When the user clicks the blue buttons to start the simulation, variables stored previously are reset and all the configuration parameters are stored in the **param** object as properties. The user can choose whether or not the scenes of the simulation will be saved, in which case a significant amount of the client's RAM would be used. If the user hasn't inputted a value in a field, the preset default value will stay. Then, the **world.start** function is called to create the <canvas> element which displays the individual scenes in the simulation. To get the sample map for the simulation, the **AJAX "GET"** request is sent to the server which in turn sends a pre-made .txt file filled with numbers. This information is used to draw the sample map on the <canvas> . After the map data is received, gather data about the different features of the map such as the locations and borders of houses. The map information is also used to create a **walkableTiles** property of the **world** object which stores whether a human can walk on a certain tile. Humans are also created during the setup: the **humansPerHousehold** property within **param** is used to put this number of people into all the houses. Each human is given the properties **curCoords**, which is the x and y coordinates of the person object's current location, and **originCoords**, the coordinates of the person's house.

Starting the Simulation - Before First Iteration

Before the first iteration of the simulation, each **human** object is given the **coronaState** property which indicates if the person has the disease, is carrying the virus, or is asymptomatic. At the start of each day in the simulation, each **human** is given tasks that the person has to do by the end of the day such as visiting a park, going to the groceries, or going to the restaurant. These tasks are randomly assigned for each **human** object and are ordered randomly by the Fisher-Yates sorting algorithm. Every simulation starts with one infected human who is chosen by random among the **human** objects. This person is asymptomatic/not going to the hospital because he/she is the origin of the spread - if that person goes to the hospital, then the simulation would likely show no further spread of the virus. Finally, the simulation is instantiated where an asynchronous function **queueScene** calls **makeScene**. The **makeScene** function is called and put

in a task queue as a Promise and after the task is done, handled by **.then** function. The **.then** function sets a timeout to call the same function again from within that function, effectively creating a loop. The simulating was done this way because this enables the user to interact with the page while the simulation is happening in the background. If a regular function is used to continuously or recursively call the **makeScene** function, the user experience would be horrible because Javascript restricts user interaction until all the task queued is completed - that is, when nothing else is on the Javascript call stack.

Each Iteration in the Simulation

Each iteration of the simulation represents 10 seconds of a day. It is established that at the start of each day, all **human** objects are given tasks to do by the end of the day. Throughout the day, a human randomly decides whether to stay at the current location, in which case the person “wanders” around (as shown in Figure 2, moving around within boundaries of a park or a house) or fulfill a task. When a **human** object decides to complete a task and go to a location like the groceries, the nearest location is found by applying the distance formula between the **human**’s **curCoords** property and all the coordinates of the location. After the closest location is found, a path to that location is created using the **astar.js** tool using the A* pathfinding algorithm ⁹ which takes in the **walkableTiles** object, start coordinates, and the end coordinates. This path is stored in the **curPath** object and the human follows this path every iteration until the human reaches the final destination. A human object following the path in the **curPath** object is seen below in Figure 3. While the **moveHuman** function is responsible for moving humans as described above, the **infect** function transmits the virus between humans. The transmission of the virus occurs in various steps. First, a **human** object can physically receive the virus by chance if there is another infected **human** object nearby. When this happens, the human’s **coronaState** property changes to the “carrying” state where the human physically has the virus on them, specifically on their hands. During each iteration when the person is “carrying” the virus, that person has a certain percentage probability of getting infected by the virus. In the case where the probability test for infection returns true, then another test is called to determine whether the human is symptomatic or asymptomatic. If the human becomes symptomatic, the **human** object

is directed to the hospital and after reaching the destination, is removed from the array containing all humans because that human remains in the hospital.

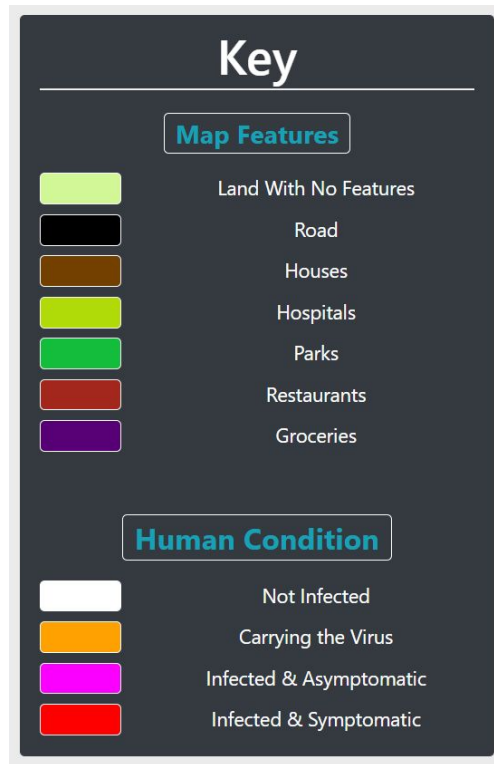


Figure 1. Key

Figure 1 Note: The key displays how the colors on the maps and its matching features or human condition.



Figure 2. Visual of Human Movement and Infection

Figure 2 Note: Three people are in a house initially but after the one carrying the virus becomes infected and asymptomatic, leaves the house to the hospital.

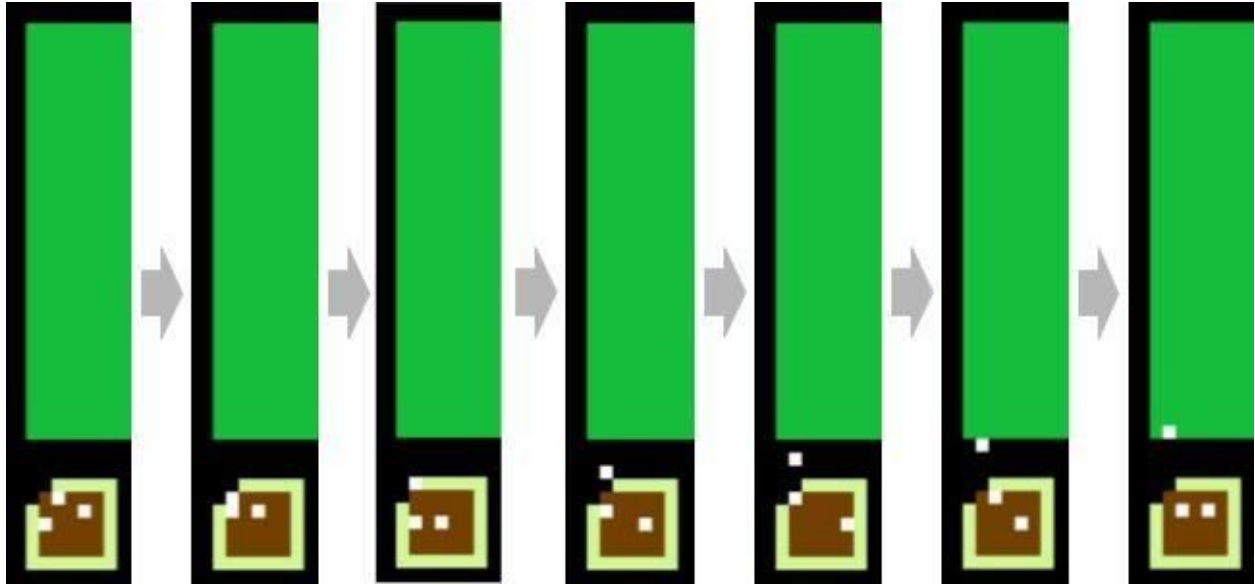


Figure 3. Visual of Human Changing States

Figure 3 Notes: A visualization of a human changing from a “wander” state to traveling somewhere. In this case, the human object selects the “parks” task and travels to the park.

Ending the Simulation

After 2 weeks pass or after 100 percent of the population is infected (whichever comes first), the simulation stops, and the loop involving the async function gets terminated. All data is presented on the graph created by **charts.js** and saved as an object for later use. The data presented on the **chart.js** canvas is collected every 20 minutes in the simulation by default, and it stores the percentage of the population infected at that given time. A copiable format of the data, created for Google Sheets or Excel, is copied to clipboard after the copy button is clicked. If there are more iterations of the simulation requested, most variables are reset to their original values and a new dataset is created and the simulation is started again.

Testing/Data Collection

Testing the impact of various preventative measures for the coronavirus, all combinations of the measures included in the simulation were tested: wearing masks, avoiding nonessential travel (parks and restaurants), and washing hands after returning home. Each combination of the measures was tested by running 10 iterations of the simulation with these measures. There were

three people per household in the simulations because the average people per household in the 2019 census by the government was rounded to three ³. After all the iterations for each simulation, the collected data (the same data that is presented on the web app) is put into an excel sheet for further analysis. The data collection described here was done on both Map 1, which represents a randomly generated town with 100 houses, and on Map 2, which represents an organized community of 576 houses. The two maps are shown in Figure 4 and Figure 5 below.

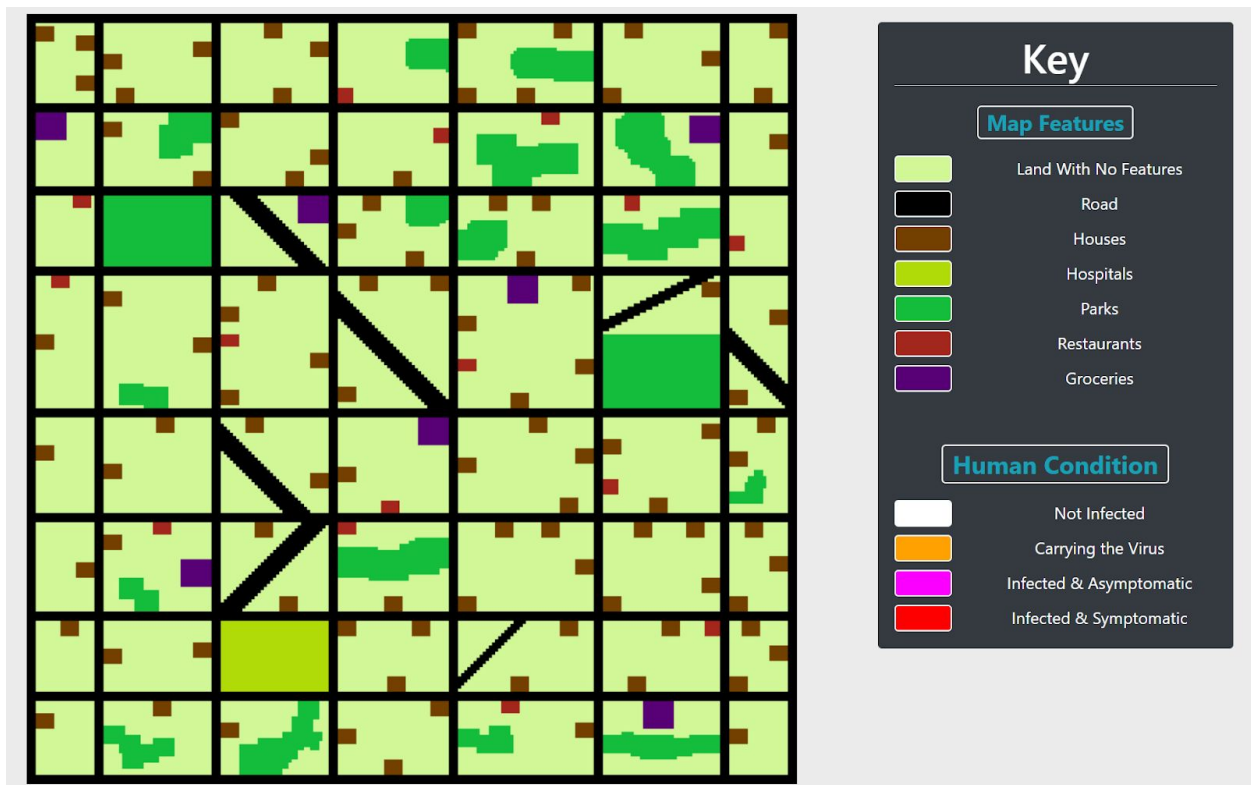


Figure 4. Image of Map 1

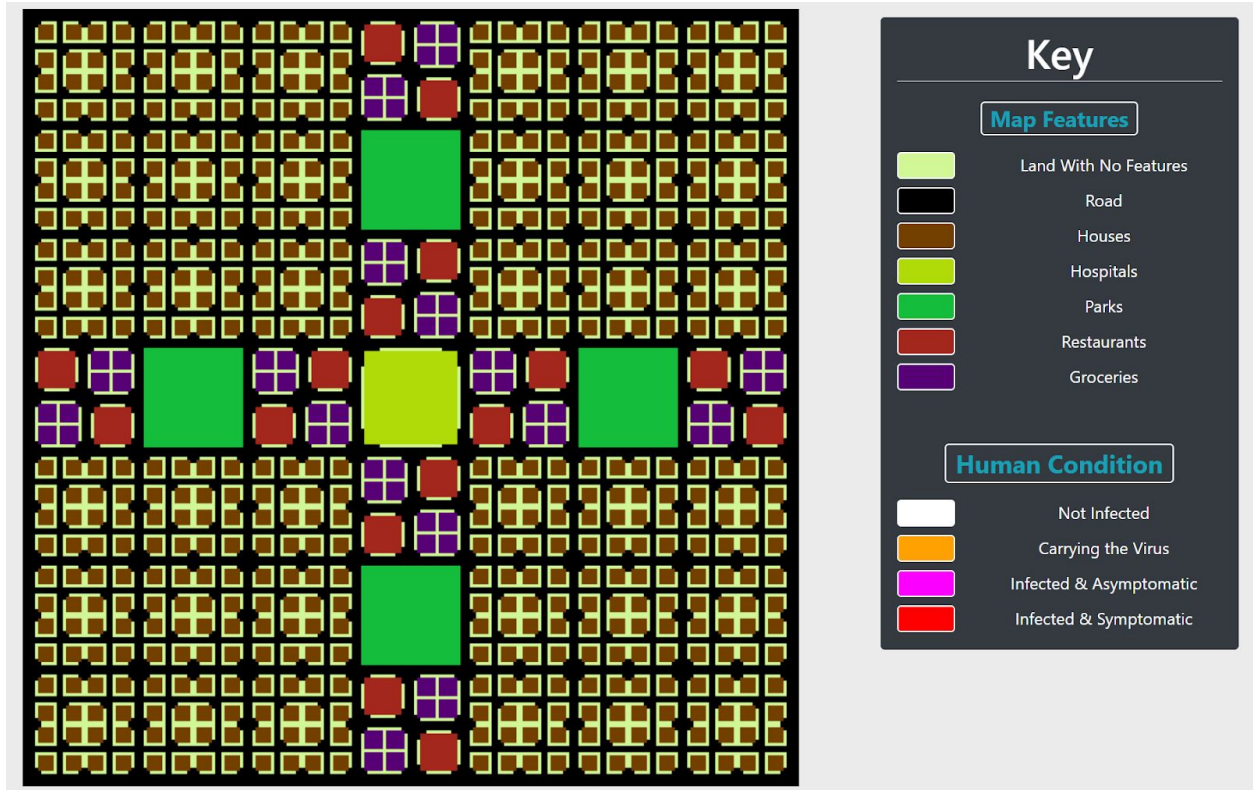


Figure 5. Image of Map 2

Data

Ten runs of the simulation for each combination of preventable measures were done and stored in a spreadsheet. These runs with different combinations of measures were saved in separate sheets for individual analysis. Within the sheets, six values were calculated for each time when the data of the percentage of the population infected was recorded: mean, median, minimum, maximum, average of the minimum and maximum, and the mean of the six middle values. For example, for when $t=0$ minutes, the 10 percentage values collected in each dataset were used to calculate these values. The mean of the six middle values is a “trimmed mean” calculation where the two highest values and the two lowest values are omitted and the mean of the remaining six values are found. This method for finding a basis for analysis was used to compare the values for all the different tests with various combinations of measures. All the data calculated for Map 1 and Map 2 were plotted in graphs below. Two tables were also created for

Map 1 and Map 2, containing the maximum percentage for each simulation with different combinations of measures. This table also calculated the time each simulation took to reach 95% of this maximum value and the rate of infection to 95%.

Dataset with Default Settings: No Preventive Measures (Map 1)

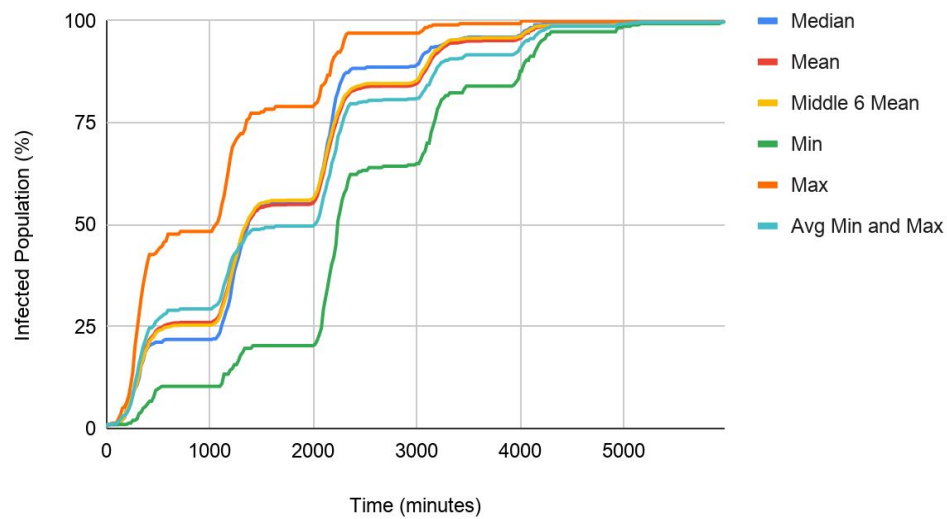


Figure 6. Data collected by simulating an area with no measures on Map 1

Dataset with Default Settings: No Preventive Measures (Map 2)

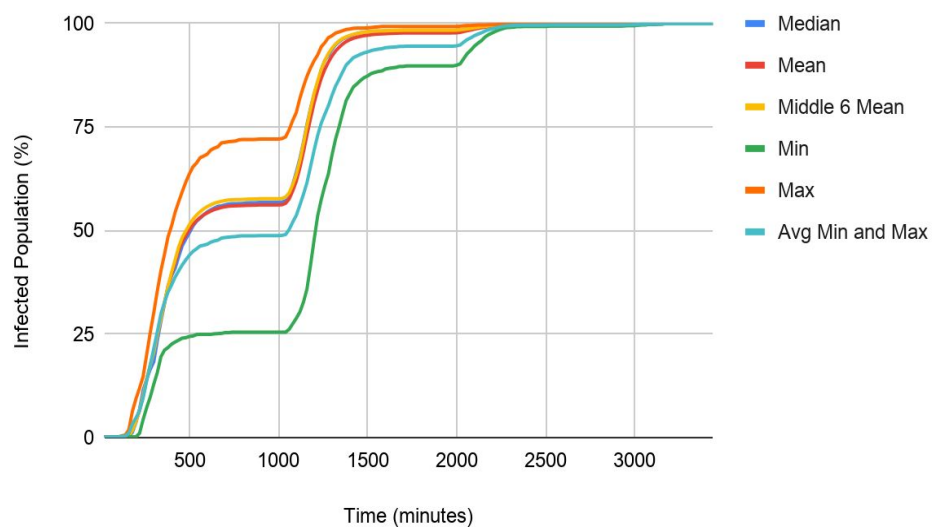


Figure 7. Data collected by simulating an area with no measures on Map 2

Dataset with All 3 Preventive Measures (Map 1)

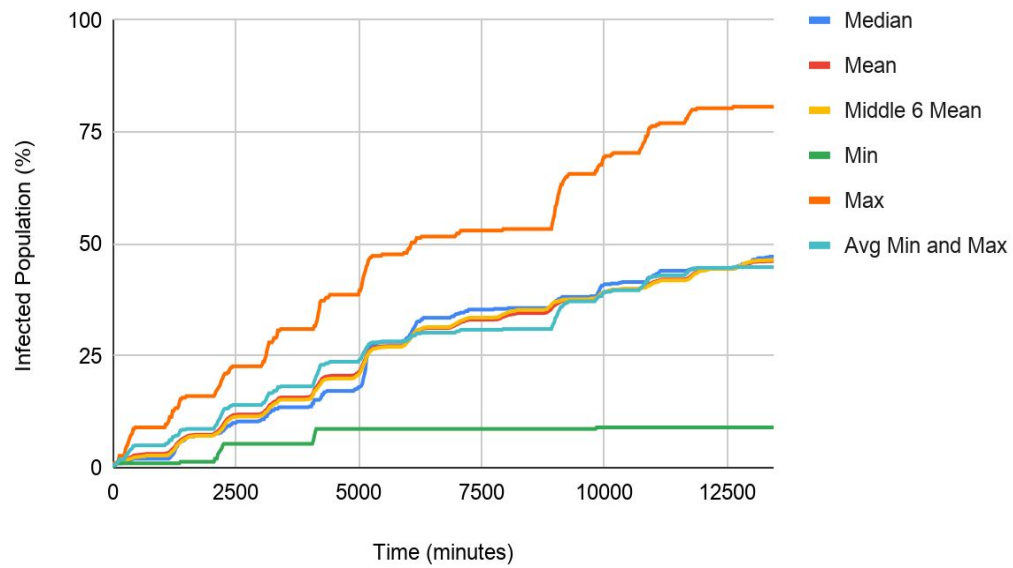


Figure 8. Data collected by simulating an area with all three measures on Map 1

Dataset with All 3 Preventive Measures (Map 2)

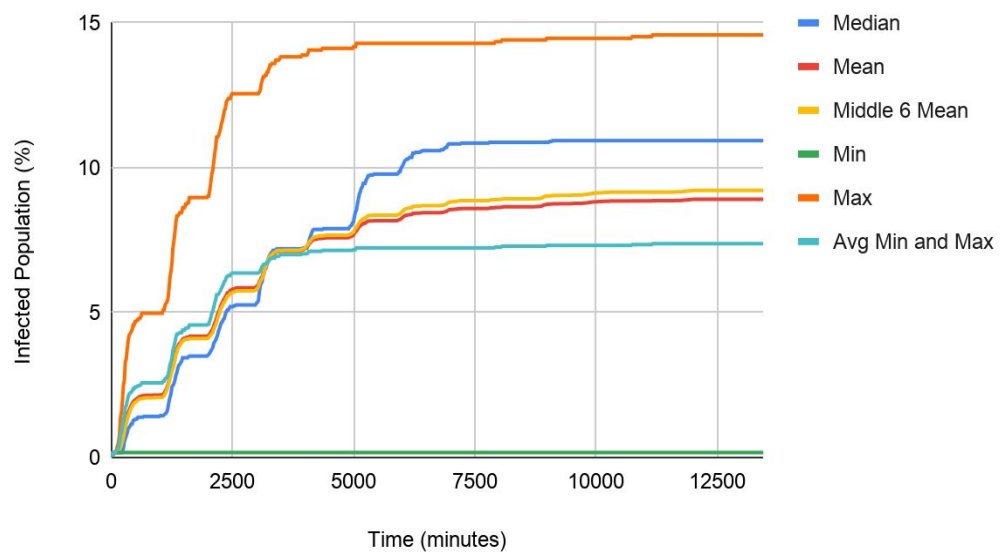


Figure 9. Data collected by simulating an area with all three measures on Map 2

Middle 6 Mean For All Test Cases (Map 1)

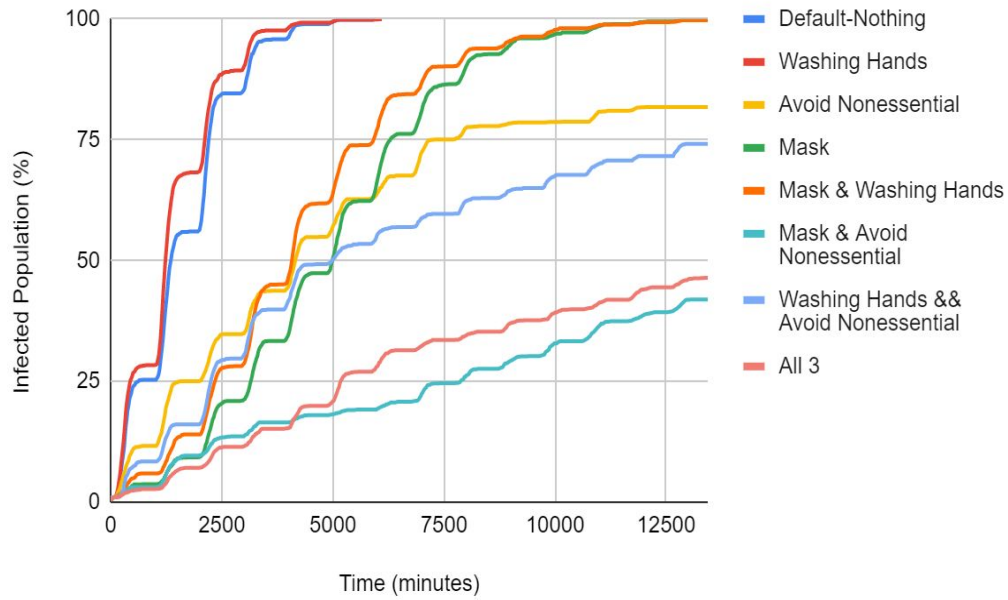


Figure 10. Comparing the different test cases/measures put in place using the mean value of the six middle values for each iteration (Map 1)

Middle 6 Mean For All Test Cases (Map 2)

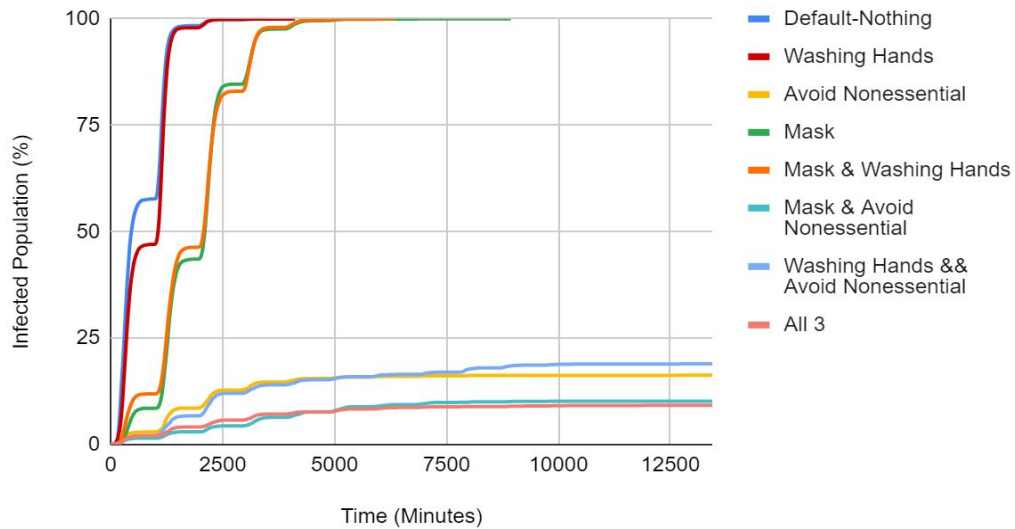


Figure 11. Comparing the different test cases/measures put in place using the mean value of the six middle values for each iteration (Map 2)

	Default (Nothing)	Washing Hands	Mask	Avoid Nonessenti als	Mask & Washing Hands	Mask & Avoid Nonessenti al	Washing Hands & Avoid Nonessenti al	All 3
Max Percentage Hit (%)	100	100	100	81.8	100	41.9	74.2	46.4
Time to Reach 95% of max (min)	3300	3140	9040	8120	8900	12700	11100	12040
Rate of Infection 95% (% per min)	0.028787	0.030254	0.010508	0.009570	0.010674	0.003134	0.006350	0.003661
Rate of Infection Compared to Default (% dec)	0	-5.10	63.50	66.76	62.92	89.11	77.94	87.28

Figure 12. Analysis of Maximum Population Infected for Map 1

Figure 12 Notes: Figure 12 summarizes the time it took to reach 95% of the max percentage value, and the rate of infection from start to 95% for each of the test cases (Map 1)

	Default (Nothing)	Washing Hands	Mask	Avoid Nonessenti al	Mask & Washing Hands	Mask & Avoid Nonessen tial	Washing Hands & Avoid Nonessenti al	All 3
Max Percentage Hit (%)	100	100	100	16.3	100	10.2	19	9.2
Time to Reach 95% of max (min)	1300	1360	3240	4380	3240	7060	8740	6920
Rate of Infection 95% (% per min)	0.073076	0.069852	0.029320	0.003535	0.029320	0.001372	0.002065	0.001263
Rate of	0	4.41	59.88	95.16	59.88	98.12	97.17	98.27

Infection Compared to Default (% dec)								
---------------------------------------	--	--	--	--	--	--	--	--

Figure 13. Analysis of maximum population infected for Map 2

Figure 13 Notes: Figure 13 summarizes the time it took to reach 95% of the max percentage value, and the rate of infection from start to 95% for each of the test cases (Map 2)

Analysis

The data collected above shows interesting results regarding the spread of the coronavirus. Figures 7 to 10 are examples of the analysis done on the individual test cases, four examples out of 16 total - eight different combinations of measures for the two maps. With the results for each test case, the 10 datasets, the mean, median, middle six mean, maximum, minimum, and the average of the max and min were found on one graph as shown in the Figures 7-10 analysis examples. The 16-simulation data and the graphs showed that getting the trimmed mean of the six middle values resulted in the most consistent representation of the percentage infected at most data points. The trimmed mean produced the lowest standard deviation values and was effective in getting rid of the outliers. Therefore, this analysis method was used for further analysis in Figures 12 and 13. The median value was not used because it failed to consider multiple values. Because of this, the median is prone to bias in instances where the middle value or the middle two values are skewed. The trends of the results were logical because with more measures to control the spread of the coronavirus, the rate of infection was lower.

With no measures in the controlled environment, the percentage of people infected after a 2 week period for the two maps were both 100% while 46.4% of people in Map 1 and 9.2% of the people in Map 2 were infected with all three measures in place: washing hands upon returning home, wearing masks, and avoiding nonessential travel. Although the two communities represented in Map 1 and Map 2 had stark differences in the population size and the degree of structuredness of the community, they presented similar trends. Among the three preventative

measures tested in the various simulations, avoiding nonessential travel was most impactful in lessening the rate of spread at 0.009570% per minute for Map 1 and 0.003535% per minute for Map 2. Because each person only travels to grocery stores, the number of people each human contact throughout the day is lessened. This test case was particularly interesting because the results for Map 1 indicated around a 67% reduction in the rate of the spread compared to the default simulation while that of Map 2 showed around 95%. Thus, avoiding nonessential travel had a more significant impact on densely populated communities than in less populated ones. The measure of wearing masks proved to be the next effective measure in slowing and containing the spread. These results are reasonable and similar for both maps with the rate of infection per minute decreasing about 60%. This is because wearing masks significantly reduces the probability of coronavirus transmission between individuals while the number of people contacted stays unimpacted. The washing hands measure, on the other hand, didn't have a significant effect on the rate of spread, showing only a 4.41% decrease in the rate of spread compared to the default test for Map 2 and even showing an increase in the rate of spread for Map 1.

For the test cases with more than one measure put in place, the trend was similar. The "washing hands" measure had little to no impact when paired with wearing masks. When paired with the "avoid nonessential" measure, it decreased the rate of spread by a noticeable amount compared to that of the single "avoid nonessential measure" in both Map 1 and Map 2. Although this was the case, the max percentage hit for the two test cases were similar for both trials on Map 1 and Map 2. Even though the "washing hands" showed some impact in the previous case, the test case with all 3 measures produced the similar results to the test case combining "avoid nonessential" and "mask", showing little to no impact of the "washing hands" measure once again. Both Map 1 and Map 2 data for these two test cases produced a difference of less than 3% regarding the percent decrease in infection rate and reached a similar maximum percentage of population infected.

Conclusion

The coronavirus spread simulation was created and allowed the analysis of spread within a closed community. The simulation was created to emulate a real-world environment with human movement patterns and logical methods of virus spread. By simulating the spread of the virus with various test cases with different preventive measures, the effectiveness of the measures was found. After testing with different combinations of these measures, the most effective test cases were the “Mask & Avoid Nonessential” and “All 3” cases. Various tests indicated that the “Avoid Nonessential” measure was most effective in lessening the rate of spread followed by the “Mask” and “Washing Hands” measures. These results align with the CDC’s advice of “Social Distancing” and the emphasis put on limiting contact: “Only visit stores selling household essentials in person when you absolutely need to”.³ It is thereby evident that the prevention of contact with other individuals by only travelling to essential places is more effective than the attempt at reducing these risks after or during human contact by washing hands or wearing masks.

Discussion

The study on the spread of the coronavirus is something that is being done by various individuals and groups. For example, the Wolfram team is doing data analysis on the current coronavirus data based on their curated computable data⁶ while individual researchers like Renyi Zhang are involved in identifying airborne transmission of the coronavirus¹⁴. Among these researchers, some researchers such as Thomas Riisgard⁸ and Benjamin Fredrich⁴ have used similar simulations to collect data as well. Each has configurable factors that impact the spread of the virus. This project’s take on analyzing the rate of spread was different from others on the web because it simulates realistic human movement and interaction. The emphasis of this project was different from other simulations on the web: the focus was on creating closed and realistic communities with humans moving randomly while also fulfilling the daily tasks such as going to a park. This method of simulating opens the door to the analysis of real communities around the world. For example, this simulation can be furthered to perform case studies of the

spread in large cities such as Chicago and New York by importing drawn maps of these cities to the simulation. Also, more configurable options such as preventive measures can be added to the simulation.

Along with its advantages, this simulation also has its limitations. First off, the map input and logic were created to support 2D spaces rather than 3D. Not being able to model human movement in a 3D environment restricts the accurate representation of spread in spaces such as apartments and multi-story buildings. Another major limitation of this simulation is the amount of memory and time it uses to compute and store the computed data. When running on a computer with less RAM with the “save scenes” configuration on, the browser runs out of memory and crashes. Regarding computation, since this simulation also deems important the users’ access to a visual representation of spread through human movement, various calculations must be done to move, infect, and set a path for a human object. This causes the data collection process to take a significant amount of time. Lastly, a part of this simulation that needs to be addressed for future use is the difficulty of creating various maps for testing. Each of the maps created are imported from a .txt file that has 250 x 250, 62500 different numbers that creates the map. For the testing of various maps, the creation of a tool for map creation may be beneficial.

A limitation within the method of data collection was the length of the simulation period and the lack of a quarantine measure for infected individuals. The data was collected through a two-week period for each trial, taking recovery from the Covid-19 out of consideration. Also, symptomatic and infected individuals were immediately directed to the hospital after infection, excluding spread through quarantining individuals. For long term spread with further measures such as quarantining, further analysis is needed. Simulating spread for larger areas for a longer period will make analysis by logistic regression a viable option.

Another part of the simulation that must be taken into consideration is ability for authorities to enforce the laws presented. The hardships of enforcement can be due to public pushback, cost, or many other factors. For example, requiring people to wear masks is a

relatively hard task because an abundant supply of masks and financial support needs to be available for people to have access to it. Also, people have raised concerns about the inconvenience of masks and claimed that it is their right to practice their free will by refusing to wear masks. Although hard and costly to enforce, many countries have opted to supply their citizens with masks and take on the financial burden. Requiring the people in a community to only travel to essential places such as the groceries is extremely difficult to properly enforce because it is based on accountability. Also, this would hurt many businesses such as restaurants and those in the service industry, negatively impacting the community's economy. Having people wash their hands is also a rather easy measure to properly enforce with little to no pushback. Since washing hands is part of a daily routine for many individuals with little to no cost, it is the easiest measure for authorities to recommend.

The discussion of the three preventive measures in the simulation presents an interesting trend: the harder the measure is to implement in society, the greater impact it will have on lowering the rate of spread. Finding the optimal preventive measure seems to be one of the greatest challenges researchers around the world have in this era of the coronavirus. Leaders and authorities around the world have considered various measures and have enforced some measure in the hopes of containing the spread. Although methods such as quarantining and wearing masks have shown some levels of success, more research must be done to find the most optimal preventive measures. This remains as the greatest obstacle before the introduction of commercially available Covid-19 vaccines.

Works Cited

1. Black, Paul E. *Fisher-Yates Shuffle*, Dictionary of Algorithms and Data Structures, 6 May 2019, 10:22:33, xlinux.nist.gov/dads/HTML/fisherYatesShuffle.html.
2. Brosseau, Lisa, and Roland Berry Ann. "N95 Respirators and Surgical Masks." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 14 Oct. 2009, blogs.cdc.gov/niosh-science-blog/2009/10/14/n95/#:~:text=Collection%20efficiency%20of%20surgical%20mask,their%20performance%20in%20these%20studies.
3. Bureau, US Census. "Historical Households Tables." *The United States Census Bureau*, The United States Census Bureau, 10 Oct. 2019, www.census.gov/data/tables/time-series/demo/families/households.html.
4. Fredrich, Benjamin. *Coronavirus Simulation*, Katapult-Magazin, 2020, corona.katapult-magazin.de/.
5. "Coronavirus." *World Health Organization*, World Health Organization, 2020, www.who.int/health-topics/coronavirus#tab=tab_1.
6. "COVID-19 Coronavirus Data & Resources." *COVID-19 Data & Resources*, Wolfram, 2020, www.wolfram.com/covid-19-resources/.
7. "COVID-19 Pandemic Planning Scenarios." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 10 July 2020, www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html.

8. Hansen, Thomas Riisgaard. "Coronavirus Simulator." *Corona*, 2020, corona-land.org/start.
9. *Introduction to A**, Stanford University, theory.stanford.edu/~amitp/GameProgramming/AStarComparison.html.
10. Plater, Roz. "50 Percent of People with COVID-19 Aren't Aware They Have Virus." *Healthline*, Healthline Media, 28 May 2020, www.healthline.com/health-news/50-percent-of-people-with-covid19-not-aware-have-virus.
11. "Q&A: How Is COVID-19 Transmitted?" *World Health Organization*, World Health Organization, 9 July 2020, www.who.int/news-room/q-a-detail/q-a-how-is-covid-19-transmitted.
12. Schumaker, Eric. "Asymptomatic and Presymptomatic People Transmit Most COVID-19 Infections: Study." *ABC News*, ABC News Network, 8 July 2020, abcnews.go.com/Health/asymptomatic-presymptomatic-people-transmit-covid-19-infections-study/story?id=71647268.
13. "Social Distancing, Quarantine, and Isolation." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 15 July 2020, www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html.
14. Zhang, Renyi, et al. "Identifying Airborne Transmission as the Dominant Route for the Spread of COVID-19." *PNAS*, National Academy of Sciences, 30 June 2020, www.pnas.org/content/117/26/14857.