MSDS 460 Final Project: Agent-Based Disease Simulation

Group Members

- Albert Lee
- Alberto Olea
- Maddie Stapulionis
- Migus Wong
- Thomas Young

1. Abstract

Our group chose to simulate the spread of disease in a community using an agent-based model to capture variability in individual behavior. The goal of this simulation is to get a better understanding of how certain variables - such as vaccination rate, isolation rate, and disease transmission likelihood - influence the outcome of a pandemic. These findings could be used to inform future policies and best practice guides for disease responses and to better guide how to strike a balance between restrictiveness and public safety.

Epidemiology is the field of study regarding the distribution, patterns and determinants of health which affect the interaction between a disease and the surrounding population. This field has been extensively studied for centuries due to its government, business, and population implications.

Rather than treating all individuals as identical, our approach introduces randomness and variability through an agent-based model. This method allows individual decisions to have a lasting impact on the overall outcome of the simulation, similar to how superspreader events can critically impact the spread rate of a disease.

Given that epidemiology is a well-established field with models far more complex than the time constraints of this term allow, the model developed here serves as the introductory step towards disease modeling. It provides a firm foundation that should be sufficient to generate reasonably accurate results based on the assumptions made.

2. Data Modeling

Agent based modeling allows us to capture a great number of interactions and actions of individuals. (Hunter, 2025) In our simulation, we model a closed community consisting of 500 agents randomly positioned in a two-dimensional area measuring 500 by 500 units. Each agent has specific characteristics that influence the simulation's dynamic. This includes spatial coordinates (x, y), infection status indicating current disease state, and immunity status representing temporary protection following recovery. Agents have a numerical resistance level that is initially set as 0.3. This determines the probability of resisting infection upon exposure to an infected agent.

Agents also maintain counters to track their disease progression and immune duration. The infection duration counter is initially set to 50 timesteps thus determining the length of an infection. Further, the immunity duration counter indicates the period of temporary immunity after recovery (initialized to 50 timesteps).

To start the disease spread, 10 agents are randomly infected at the beginning. All agent’s position and health information are updated every timestep according to predetermined interaction rules, thereby simulating a closed community.

While this was the initial baseline scenario that was simulated and analyzed, the following simulation variables were also modified to test how a population would be affected under different parameters:

2.1 Number of Agents

This represents the total starting number of agents being simulated. As the total number of agents increases, the likelihood of agents making contact with each other increases. As a result, this variable can be directly related to the population density of the environment.

2.2 Agents Infected

This is the number of agents that are infected from the onset of the simulation.

2.3 Agent Resistance

This parameter controls how contagious the disease is. A higher resistance level means healthy agents are less likely to contract the disease if they come into contact with an infected agent.

2.4 Proximity Threshold

This defines the maximum distance within which agents are considered to have made contact. A lower proximity threshold reduces the likelihood of interactions, while a higher threshold increases potential exposure.

2.5 Step Size and Number of Timesteps

Determines how far agents move in each iteration of the simulation and the total duration of the simulation. Larger step sizes lead to faster movement and potentially more interactions between agents, while a greater number of timesteps allows for longer observation of disease spread and population dynamics.

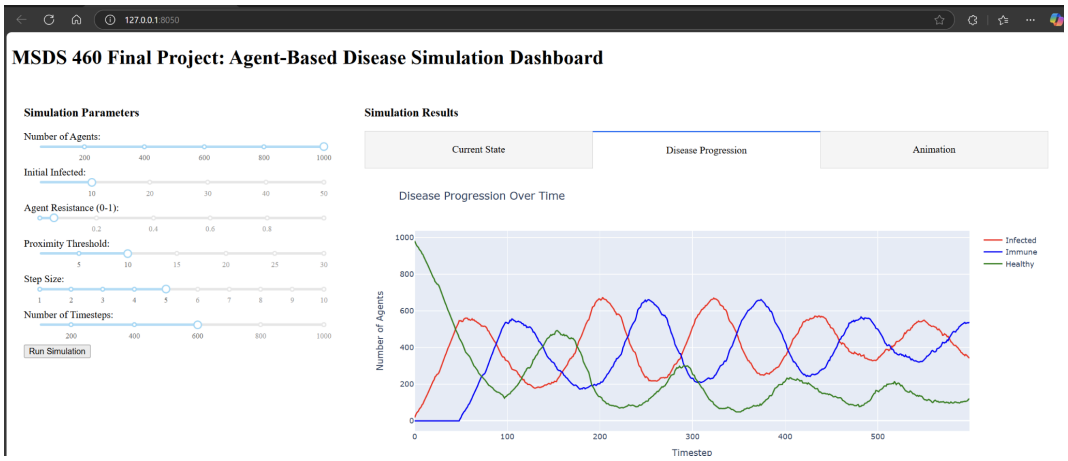
3. Algorithm and Approach

Our disease spread simulation begins by randomly distributing 500 agents across a 500x500 unit space. At the beginning, 10 agents are randomly selected to be infected, which initiates the spread of the disease within the population. Each simulation step consists of agents moving randomly within their boundaries and simulating realistic movement/behavior. In addition to following movement, distances between all agents are calculated and agents within a proximity limit of 10 units have the potential to transmit the disease. The spread of infection is probabilistically determined based on each agent's resistance level. If susceptible agents are near infected individuals they are checked for possible infection. If infection occurs then agent's

statuses are infected and duration counters are updated. When infected agents recover they acquire temporary immunity lasting 50 timestamps. This process is repeated over 500 timesteps, which allows us to observe and analyze any evolving dynamics of disease transmission. At each timestep, the position and infection status of each agent is evaluated and tracked.

4. Implementation

The described simulation parameters were built in Python using standard packages such as numpy and scipy. To add a second level of interactivity to the simulations, Plotly's Dash framework was utilized to create a front-end application that allows users to adjust many of the parameters seen above and view the transmission of the disease over time.



4.1 Assumptions and Limitations

Although this simulation is intended to mimic how a disease interacts with a human population, there were several assumptions made that do not necessarily reflect reality. Firstly, the simulation assumes that human movement is 100% random (unless near the edges of our defined space) meaning there is an equal probability of agent interactions; this does not reflect that in reality, agents often travel away from their home and interact with other agents. Ventilation and agent specific spread probabilities based on masking would also benefit the efficacy of this model. (Griffith et. al, 2025) Furthermore, family units are not represented in this simulation and would impact how a disease spreads given the proximity of space shared. In future studies, it would make sense to develop a network of connections between agents since the interactions between agents should not only be proximity-based, but also relationship based.

There are other demographic features of the population that were implemented here but do play a pivotal role in epidemiology such as age distribution. Given that elderly populations tend to be more susceptible towards negative outcomes, we would expect an older population in a community to be significantly less resilient towards recovery and may warrant the need to implement another infection status, death.

While this does tend to occur randomly in the simulations, it would also be interesting to simulate the effect of superspreader events. Large social gatherings have, as seen during the COVID-19 pandemic, have a significant impact and the spread of a disease not only because these gatherings typically produce a high density of agents in a concentrated area, but also because after these events, agents disperse back throughout the community thus increasing the infected area and supercharging the spread rate of a disease in terms of area covered.

5. Results and Conclusions

While our simulation provided valuable insights into the dynamics of disease spread within a closed community, it is also clear that a simulation such as this does not perfectly match peer to peer interactions. The assumption of random interactions and random movement is more than likely incorrect in reality so thus far, the simulation is not ready to provide actionable recommendations on disease response recommendations.

However, our simulation provided valuable insights into the dynamics of disease spread within a closed community of agents and forms the necessary building blocks of how infections may spread. The framework here can be built upon to create more complex social structures and event flows to better mimic human behavior and potential responses as “agent-based models are able to capture complex interactions between factors and emergent results based on agents’ decisions within the model that other types of models cannot” (Hunter, 2018). The development of an interactive dashboard also gives the necessary user interface for anyone to better understand and develop several frameworks to test. Future enhancements could improve its accuracy and reduce the number of assumptions needed to make the simulation a useful tool in disease management and regulation.

Our simulation model provided valuable insights but we believe for it to truly be a real world model, we would need to factor in additional real-world scenarios such as transportation dynamics, population identity (age, gender, race, etc.), preventative/mitigation strategies, and mutation scenarios.

6. Citations

Blanco, Rafael, Gustavo Patow, and Nuria Pelechano. "Simulating Real-Life Scenarios to Better Understand the Spread of Diseases Under Different Contexts." *Scientific Reports* 14, no. 1 (2024): 2694. https://doi.org/10.1038/s41598-024-52903-w..

Hazel Griffith, Cristina Ruiz-Martin, Gabriel Wainer, "A Discrete-event modeling method to study human behavior for spread of diseases on university campuses." *Computers & Industrial Engineering,* volume 200, (2025):. https://doi.org/10.1016/j.cie.2024.110732.

Hunter, Elizabeth, Brian Mac Namee, and John Kelleher. "An Open-Data-Driven Agent-Based Model to Simulate Infectious Disease Outbreaks." *PLOS ONE* 13, no. 12 (2018): e0208775. https://doi.org/10.1371/journal.pone.0208775.

Technical Explainer: Infectious disease transmission models. (2025, January 6). CFA: Modeling and Forecasting. https://www.cdc.gov/cfa-modeling-and-forecasting/about/explainer-transmission-models.html#:~:text=Agent%2Dbased%20Models,-Compartmental%20models%20can&text=ABMs%20simulate%20disease%20spread%20at,disease%20status%20(Figure%204).