Simulating the spread of COVID-19 in enclosed environments, a supermarket setup

Oscar Jackson

April 17, 2022

Abstract

In this paper I demonstrate a modular random walk COVID-19 simulation that can be used to assess the spread of COVID-19 within a set environment. The example scenario I use is a supermarket, in which I have studied the impacts of mask usage and room layout in a post COVID mandates time. I have found that if someone is positive for COVID-19 the best option is to wear the most protecting mask you own to stop the transfer to other people in social hubs, even if social distancing is no longer practiced it can reduce the number of infected by 59.1% $\pm$ 11.1%. It was discovered that the layout of a room has an impact on the infection rate, providing evidence that a rooms layout can be optimised to limit the spread of air born diseases.

1 Introduction

1.1 The Pandemic

The COVID-19 pandemic has drastically changed our understanding of pathogens, and the science of simulations has struggle to keep up with it in the times needed to enact real change in practice. The people in charge have been given the difficult job of balancing the economy and the well being of their people, some have chosen a zero COVID approach and others have looked towards herd immunity. Consistently they have said that they have been following the science which itself was making strides and bounds in understanding, meaning the science of one week could be outdated by the next, this caused confusion throughout the population as the media presented contradicting information. A great example of this was the World Health Organisation who stated "There is no specific evidence to suggest that the wearing of masks by the mass population has any potential benefit. In fact, there's some evidence to suggest the opposite in the misuse of wearing a mask properly or fitting it properly," Dr. Mike Ryan an executive director of the WHO health emergencies program [How20], this has later been proved wrong and they have changed their stance encouraging mask usage [Ell20] as one of the best tools to limit the spread [BB21].

Early on in the pandemic there where many reports of different simulations looking at a range of spreading events and the impacts of a bad or misguided simulations/science was really seen, which fed in to the already confusing information space of COVID. The importance of quick clear science, has been vital to the response of countries around the world, the transparency of this science and of the models assumptions and drawbacks is key to stopping the flow of misinformation.

This project aims to create a modular simulation package which can be easily adapted for different environments and situations. This software will allow the analysis of high traffic social environments (airports, supermarkets, libraries, malls, nightclubs, parties) and make the best choices to minimise the spread of COVID, such as limiting the number of people allowed in or changing the room layout.

1.2 Simulations

From the start of the pandemic many groups have used computer simulations to model the spread of COVID in many ways, there is almost a simulation for each part of COVID, from world population simulations using differential equations to aerosol modelling for COVID-19 droplet analysis and prediction.

2 Method

My simulation will focus on the spread of COVID within enclosed environments such as rooms or large open plan buildings, to see how human movement results in COVID spreading to other people in the simulation. In this short time scale the rules are different to other simulations and intern different conclusions can be made, the results of such are dependent on the programmed human behaviour. It is a much smaller scale than many other simulations, but it is important to know the behaviour in such environments as they are some of the major COVID-19 spreading locations as they function as social meeting points. Finding the probability of getting infected for spending a certain amount of time in a store can give people knowledge on the safest practice when shopping or looking at the effects of layout can lead store owners to limit the spread by only moving the shopping aisles around.

I have designed a random-walk simulation with a fully customizable environment, I shall be using it to simulate a medium sized supermarket and model the volume of contact events to determine the best practice to limit the spread of COVID in a local meeting place with high traffic. Now that most COVID-19 guidelines are scrapped I will focus on simulating the system with none of these enforced (no 100% mask usage and no social distancing), the current variant, and most likely the variants to come [VHMea21], is extremely infectious having very high transmissibility. On this note immunity does not tend to prevent against infection only severity when having the disease, thus I will assume that a contact event will spread COVID with 100% certainty.

A Random walk simulation works well to simulate this scenario as it simplifies human movement and allows for many simulations to be run in a feasible time-frame, making it a type of Monty Carlo model.

The assumptions made are incredibility important as they underpin the final results.

2.1 Assumptions

Face masks . . .

- Assume only 3 types of masks, 2layer cloth, disposable surgical and n95,
- Use data of national mask usage [fNS22] 68% from the 22 of February 2022,
- Assume not 100% fitted corrected, due to nature of face mask fitting assume 80% infection chance when standing right next to contagious person [fDCP21]

COVID-19 . . .

- Consider the distance travelled by aerosol droplets, only from breathing.
- All people in the simulation will have no symptoms as this changes the droplet distance drastically (sneezing for example).
- COVID droplets do not linger in the air after the person has moved away, A fully ventilated room is a fair assumption for a supermarket

Humans . . .

- Movement based on Human stride size and field of view, optional stationary periods.
- Behaviour simplified to movement towards closest partner.

2.2 Parameters

Test Environment . . .

- Supermarket size 20000 square foot [VN16] 200ft by 100ft (around 60m by 32m),
- Supermarket layout else separation of 2m.
- As said above assume well ventilated, large space with sufficient air conditioning.

Number of people . . .

- 50 people, estimation from trip to local supermarket
- Value that represents real shop usage but does not take a long time to simulate.

Time in store to simulate ...

- Average time spent in supermarket [Lak20] is 47 minutes
- Use as time to run each simulation as the contagious person would then leave, so no more spreading. Means that each person is spending the average time in store.

2.3 The Program

The simulation has been made in python and uses MATPLOTLIB, NUMPY, PANDAS, and CVS. Full code and instructions for modification can be found on my GitHub [Jac22]. In this paper they have been run on Spyder through the anaconda portal on a PC with an AMD Ryzen 5 5600X 6-Core Processor (3.70 GHz) and 32.0 GB of ram. 1343 steps with 50 people took around 20 minutes to perform 100 simulations.

The simulation can be imported as an library, it is a class based modular system, which allows control over most parameters. The main object is a Human class which holds the properties of each person in the simulation.

2.3.1 The human class

Attributes:

- **Azimuth:** Direction the person is looking (float, from 0 to 2pi)
- **Position:** x and y cords (numpy array of floats)
- **Move Mask:** Mask of illegal square meters (2D numpy array of Boolean's)
- **Index:** Unique ID for each human (string)
- **Face Mask Type:** Does person have mask on? if so what type? (int)
- **Infected:** Is this person infected with COVID-19? (Boolean)
- **Contagious:** Will this Infected person Spread COVID-19? (Boolean)
- **Immunity:** What is the chance this person will become infected in a contact event (float)
- **Droplet Distance:** Distance in which COVID-19 spreads in environment (float)
- **Mask Inefficiency:** The base inefficiency of the mask the person is using (float, 0 to 1)
- **Infect Factor B:** Helps define the overall Mask Inefficiency that depends on distance (float)
- **Time Infected:** Time (in step number) when this person became infected (int)

2.3.2 Moving the person

To move the person one step, the person will change the direction they are looking, randomly sampled from a set normal distribution spanning the average field of view, then a step amount is made and the person is moved in that direction. This simulates human movement much better than just totally random movement see Fig.1. This process is repeated a set number of times for each person creating the full simulation. To find the number of steps to simulate the average time in a supermarket must be taken into account.

The azimuth angle is defined from the positive y direction, this will identify the direction the person is looking, the azimuth angle is randomised at the start through a uniform distribution and every time before they take a step they will turn to move another direction, the distribution of this change of direction is quite important to simulate human movement. I have assumed that a person will only move in a direction within their field of view, this is around 220 deg or 3.8 rads. Therefore I will sample the change in looking direction from the distribution which spans this FOV, it is a normal distribution created using numpy normal with a scale 0.75 see Fig.2. Looking at natural human movement there is a higher chance a person will continue on the way they are looking or

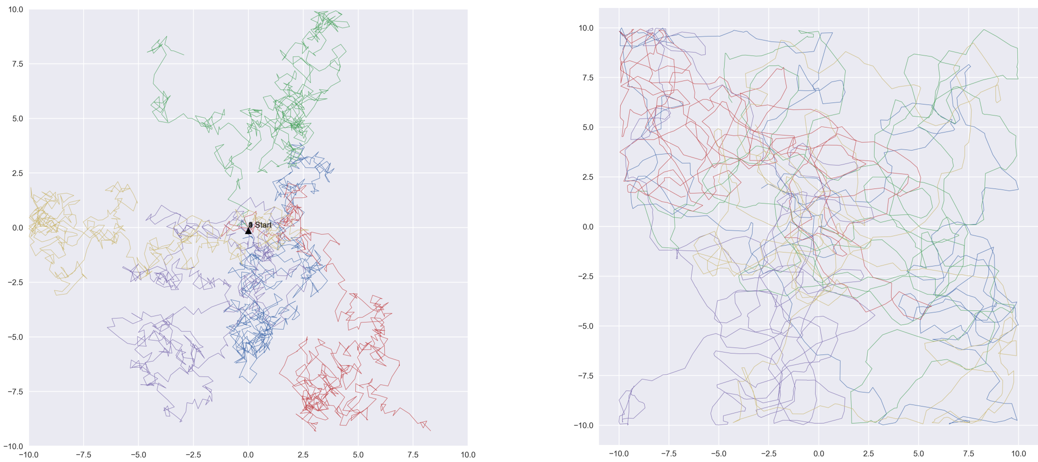

Figure 1: Comparing totally random movement (left) to a more realistic walking pattern (right).

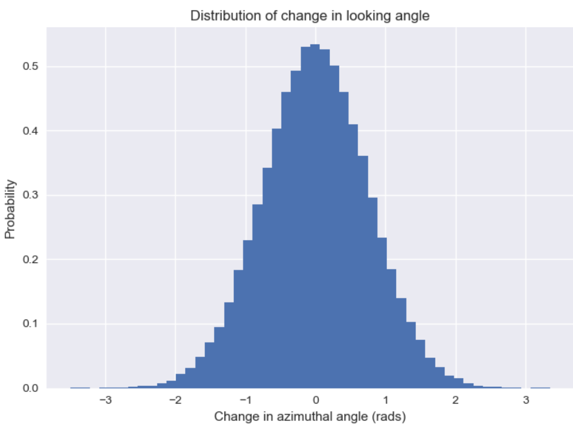

Figure 2: The distribution that the change in looking angle is sampled from.

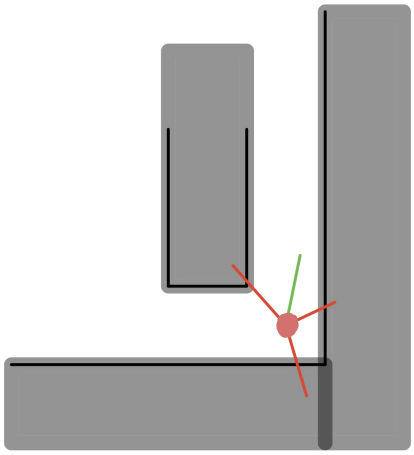

Figure 3: The person will look for a legal route (green).

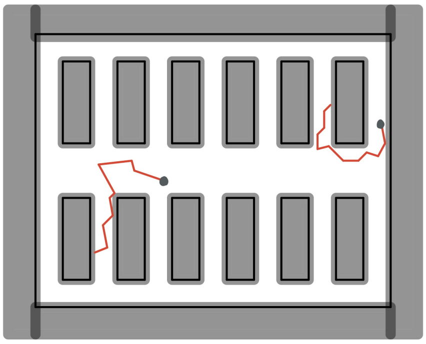

Figure 4: The route people will move in the set environment.

adjust this a small amount, thus the distribution peaks at 0. This distribution can be changed if this behaviour is determined to be unrealistic.

Human stride size is 1.5m [Hau18], in a supermarket people will stop or move slower if they are getting something of the shelves so I will sample the stride side from a uniform distribution from 0 to 1.5 m. There is also the option to sample from 0.5 to 1.5m but also have a chance of the person not moving, this produces slightly different movement but would require an increase of the step number as the average speed of the person is lowered, see results for full analysis.

2.3.3 Defining the environment

There is a separate py script that contains the functions needed to create the legal positions mask, it contains functions for each room tile seen that makes up the room, it reads the layout from a CSV file. The person will look for moves that can be taken, when it can then it can only move see Fig.3, the environment is spit into 1 x 1 m sections which are empty (can move into) or full (can not move into), this will simulate a person walking around these barriers through the room see Fig4.

The room layout I have chosen is based on several medium sized Tesco supermarkets I have visited, at around 20000 square foot. This was chosen as it is big enough to observe interesting results but not to big so that it is hard to fill it with people, this size on average would have 50 people visiting it excluding the staff who I will assume are wearing full protection and are testing negative. As this layout is customise-able it is possible to analyse results from different layouts.

2.3.4 Introducing COVID

Each person can be infected and can be contagious, contagious people are infected and can spread COVID to other people. Spread of COVID-19 occurs via airborne aerosols and droplets.

People who are contagious with COVID can release particles and droplets of respiratory fluids that contain the SARS CoV-2 virus into the air when they breath, speak, cough and sneeze)

Mask Type:	None	Two-layer cloth	Surgical	N95
Inefficacy:	100%	3.6%	0.67%	0%
Droplet Distance, when breathing:	1.25m	0.61m	0.15m	0m

Figure 5: The Properties of each face mask used in this simulation [AGS+20].

Knowing this each person has an 'infection distance' which is only dependant on mask usage, it is the distance that aerosol COVID-19 droplets travel in the air. See 5 .

My assumption is that the closer you get to another person the higher the chance you get infected as the number of droplets is increased and if the mask is not fitted correctly there will be a cloud of aerosol around the person in question increasing the risk of an infection. To replicate this, I have created a basic probability density for each mask type, i have chosen a quadratic distribution with 0.8 y intercept and droplet distance x intercept, excluding N95 which is assumed to be fitted correctly.

To find out if a person has now become infected the distance to all other people must be calculated and then compared to the droplet distances, looking for one of these people to be contagious so the virus passes on. To do this a function finds the distance from one person to all other people in the room the distances are given as difference vectors and there corresponding magnitudes, these are used to find the closest person and to know if they are within infection distance.

If they are within droplet distance then it must be determined if the droplets pass through the mask of the uninfected individual, the mask inefficacy is combined with a value (dependant on the distance to the contagious person) from the probability density of that mask simulating a loose-fitting mask. Then using this chance, it is determined through a uniform distribution if there has been a contact event and the person is now assumed to be infected. Mask usage and type is sampled from a uniform distribution with set boundaries e.g. 0.32, 0.82, 0.96 means 32%, 50%, 14% and 4% for type 1, 2, 3 and 4 face masks.

2.3.5 Other human behaviour

People being social creatures implies they moving towards each other, in this scenario it is not that important. Assuming a small amount of human interactions in the shop it still is a valid parameter, in other cases such as a party this value can be higher to simulate people talking and gathering together. To do this a person will be moved towards people surrounding them, the amount is defined by an attractiveness factor. Currently they are only moved towards one other person (the closest) this is to save computing power, but the code can be adjusted to move towards every person or the 5 closest people depending on the scenario.

2.3.6 Measuring the spread

The simulation will return several results each time it is run, a number of contact events and a list of times when a person was effected thus also giving the number of people who were infected over the run time of the simulation. Using this data many results can be determined when the simulation is run a large number of times. Some examples are the spread rate vs time and the range of possible outcomes (number infected in the end) see results section for more examples.

2.3.7 Improvements

This simulation makes many assumptions but these could be ignored and the behaviour simulated, the design of the code makes it simple to introduce new features. Coughing and sneezing could be introduced as it changes droplet distance [HHL+21] making a more realistic model, an improved distance infection probability relation or having lingering droplets in air means a well ventilated room removes more assumptions. But all these improvements will increase computing cost so optimisation would be needed to run code for larger numbers of people, possibly multiprocessing. In a supermarket setup the ability to Introduce different tile types for one-way lanes for example or better resolution in the room layout will allow more detailed layout analysis.

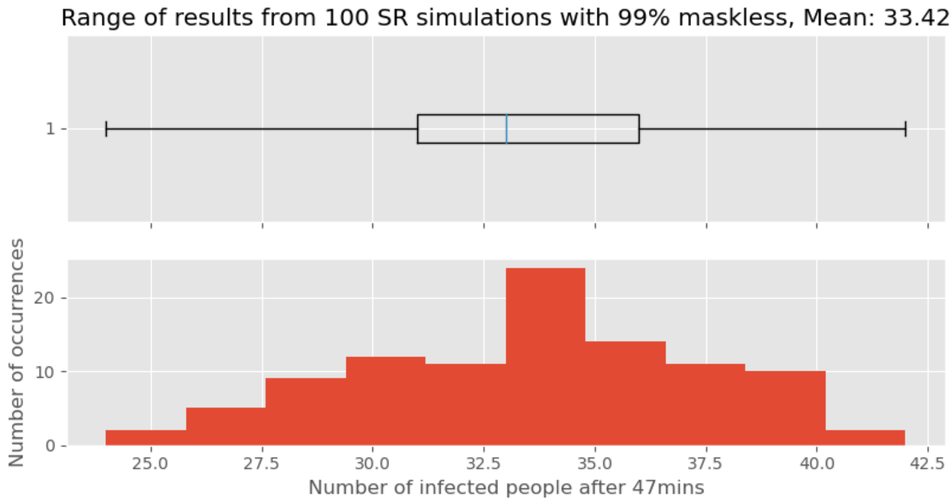

Figure 6: The range of final infected numbers from 100 simulations with 50 people and 99% of the population not wearing masks, zero immunity and simulating in the standard layout.

3 Results

Using the results from multiple runs of simulations many results can be found, each simulation will return a list containing times (in step number) of infection for each person (who has been infected) meaning the length of list is number of people infected over the whole simulation. When doing a large number of simulations it is easy to find the average number of infected at a curtain step time, a timeline of the number of new infected people can be made for each simulation by just looping over each step number and counting the number of instances and having a cumulative total. This average infection rate can be fitted and the impacts of parameters can be modelled. The number of contact events is also interesting and averages can also be taken over multiple runs. The range of possible outcomes (number of infected at the end of the simulation) is useful as it could show the unpredictable nature of such a setup.

3.1 Range of Outcomes

It is important to see if the distribution of final outcomes has a clear distribution and is not uniform or the simulations would not have clear results, in Fig6 a histogram shows the distribution of results and the corresponding mean and range for 99% people not having face masks, there is a clear mean and range. For the current mask usage a similar but less defined result is seen, Fig7, here this difference in outcome can also be seen, the average drops from 33 to 28.

The standard deviation can be used as an error of the results.

3.2 Time vs percentage exposed

Converting the steps to time is done by a simple scale factor. An average and standard deviation is found over 100 runs of the simulation, looking at 99% people not having face masks results in a clear positive relation, applying a fitting function to the curve a relation was found see Fig8.

And percentage exposed is found easily by dividing by the number of people in the simulation, see Eq.1. This is the infection rate equation, ni is the fraction infected, NT is the total number of people in the room, t is the time in minutes and alpha and beta are found through fitting for each set of starting parameters.

$$n_i = \frac{\ln(t\alpha + 1)}{N_T}\beta \tag{1}$$

3.3 Comparing room types

It is important to test if the room layout actually effects the results, therefore i have made 4 layouts to test, one is the standard room layout seen in section 2.3.3, one is a more restricted layout as I want to see if forcing people closer together with increase the spread then there is an empty room

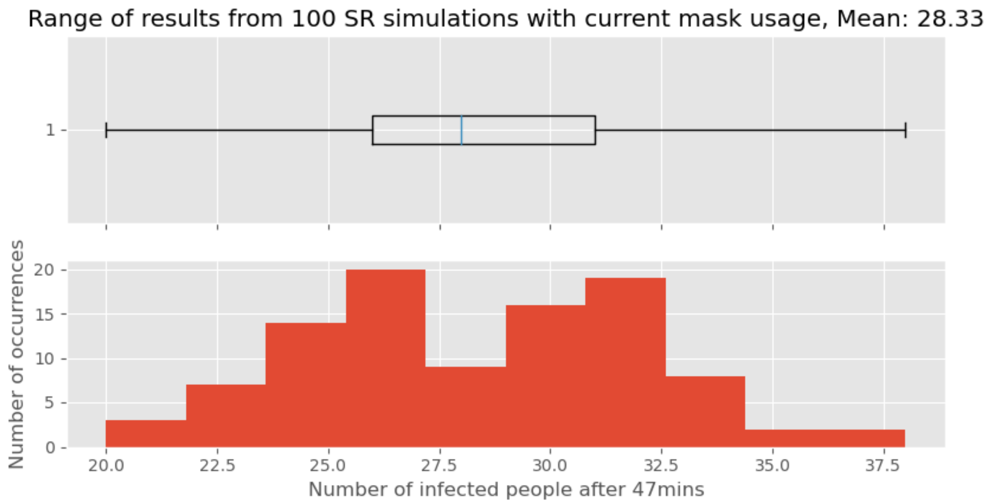

Figure 7: The range of final infected numbers from 100 simulations with 50 people with current mask usage, zero immunity and simulating in the standard layout.

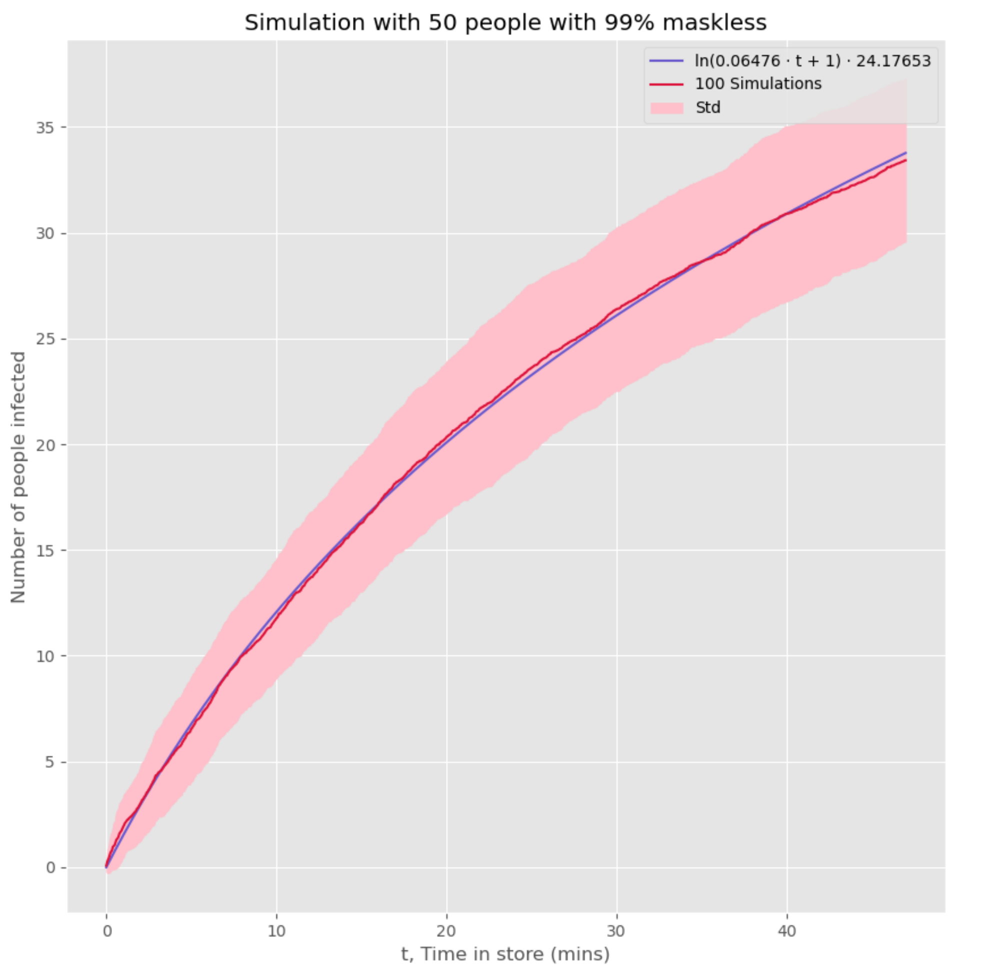

Figure 8: 100 simulations with 50 people and 99% of the population not wearing masks, zero immunity and simulating in the standard layout, shaded area shows the standard deviation.

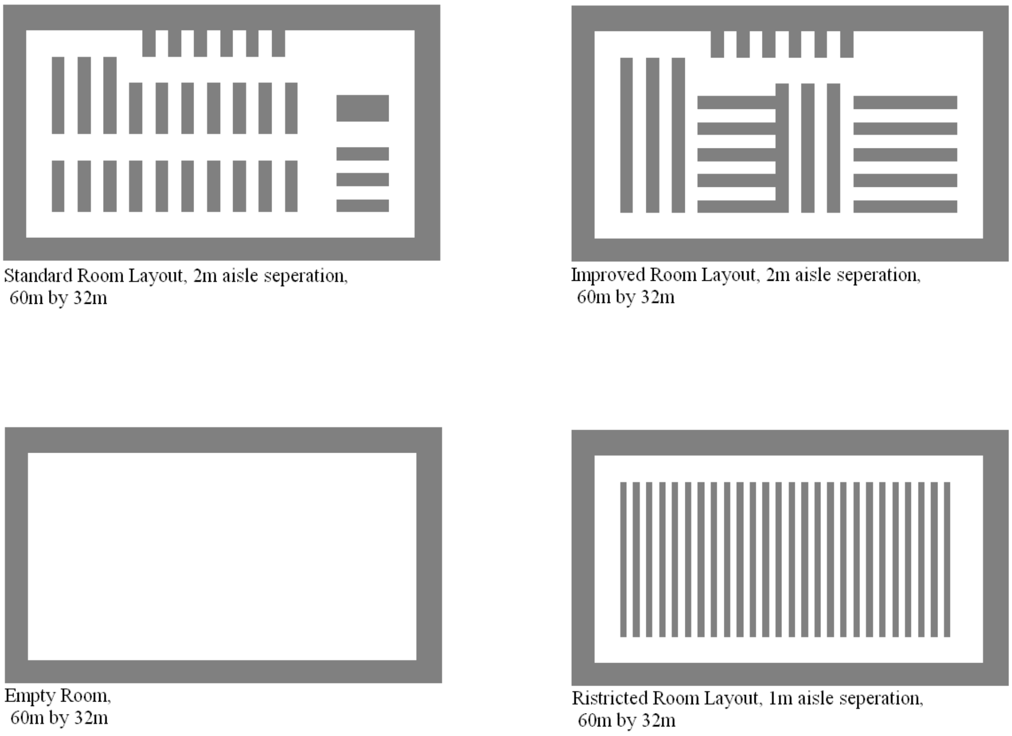

Figure 9: All room layouts being tested.

as the control. Also making up the list is my concept of an improved layout, I consider how too break the freedom of one person moving around the store and how to separate out people limiting the spread. See all of the layouts in Fig9.

When running the simulation 100 times for each layout, and with 99% people not having face masks (to only look at room type effects) clear differences are seen in the infection rate, Fig.10. When preforming the simulations for the restricted room originally the infection rate looked similar to the empty room one, this is due to the step size being bigger than the depth of the aisles meaning the people where stepping over them. To fix this I made one step into two half size ones, this may have effected the results so it is not a completely fair test Overall a difference is seen for each room layout, the Empty room had 74.6% ± 3.3% people infected after 47 minutes compared to 66.8% ± 6% with the standard room layout. It is also possible to reduce the infections by changing the room layout, seen by the 59.6% ± 9.3% result from the improved layout a 7% difference, with more thought and testing this could be reduced even more.

3.4 Effects of mask usage

Looking at the effects of the contiguous person wearing different or no masks shows great evidence towards the effectiveness of masks, even in a fairly constricted layout and no enforced social distancing. The more protective the mask of the starting contagious person is the less COVID spreads in the room, see Fig.11, and assuming a relationship similar to section 3.2 the same fitting can be done to forecast infection levels. Just by the contagious person wearing a two layer mask the percentage of people infected over 47 minutes drops from 66.8% ± 7.7% to 46.2% ± 7.3% and lower with a surgical mask to only 7.7% ± 3.4%.

3.5 Number of contact events

The number of contact events, separative from the number of infected, is an interesting parameter, there is a chance during one that COVID is not transmitted but the increase in the total number will definitely increase the chances of it spreading. Comparing the number of contact events over the full 47 minutes against the number of people in the standard room layout shows a clear positive

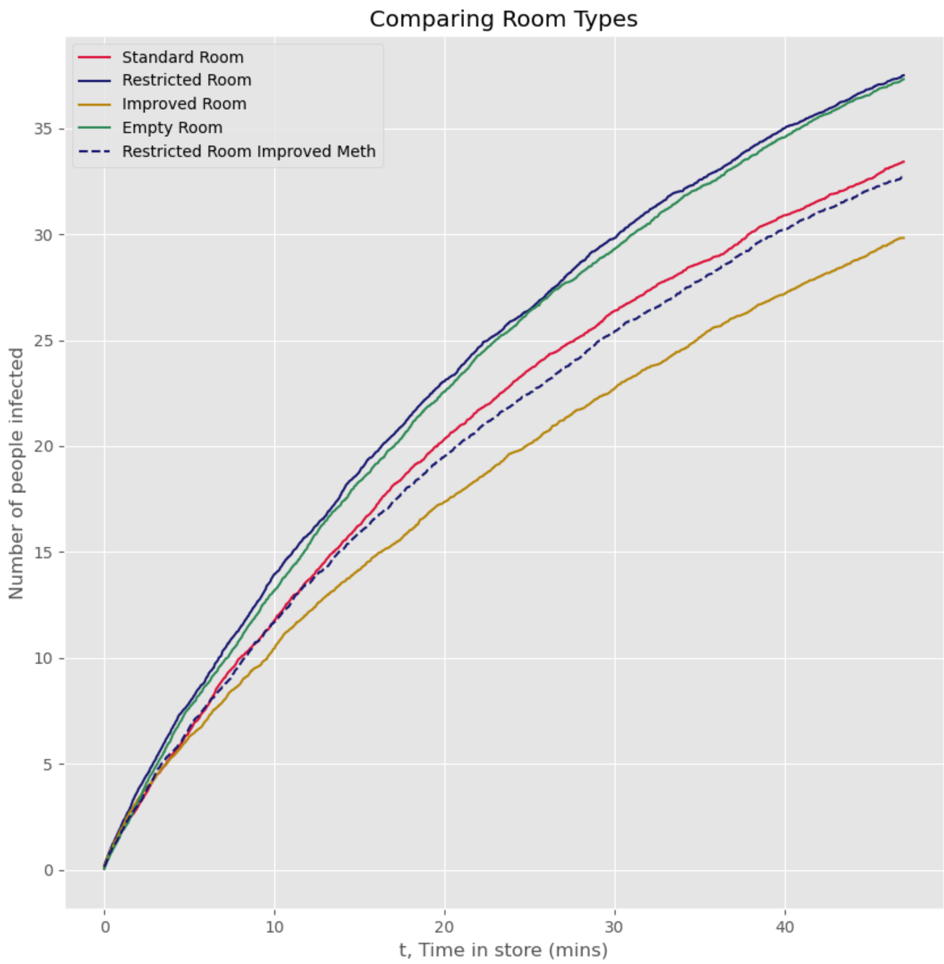

Figure 10: Comparing different room layouts see Fig#.

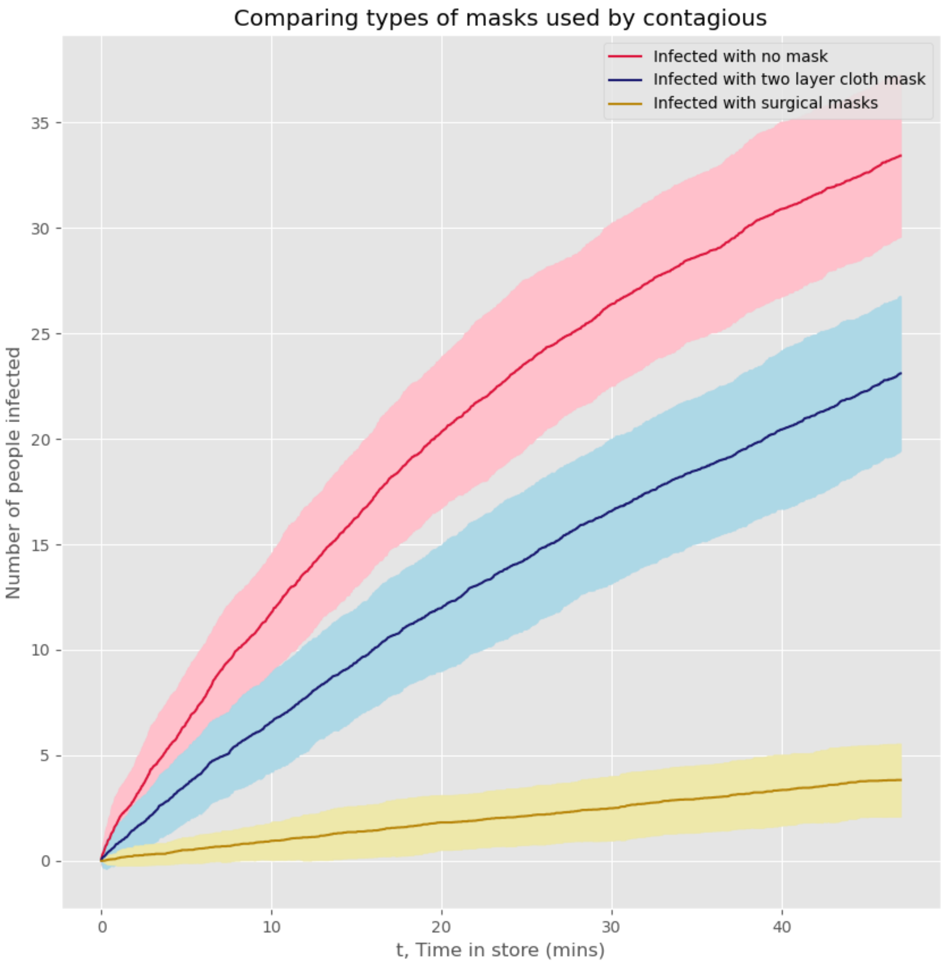

Figure 11: The infection rates when the starting contagious person is wearing different types of masks, shaded areas show the standard deviation.

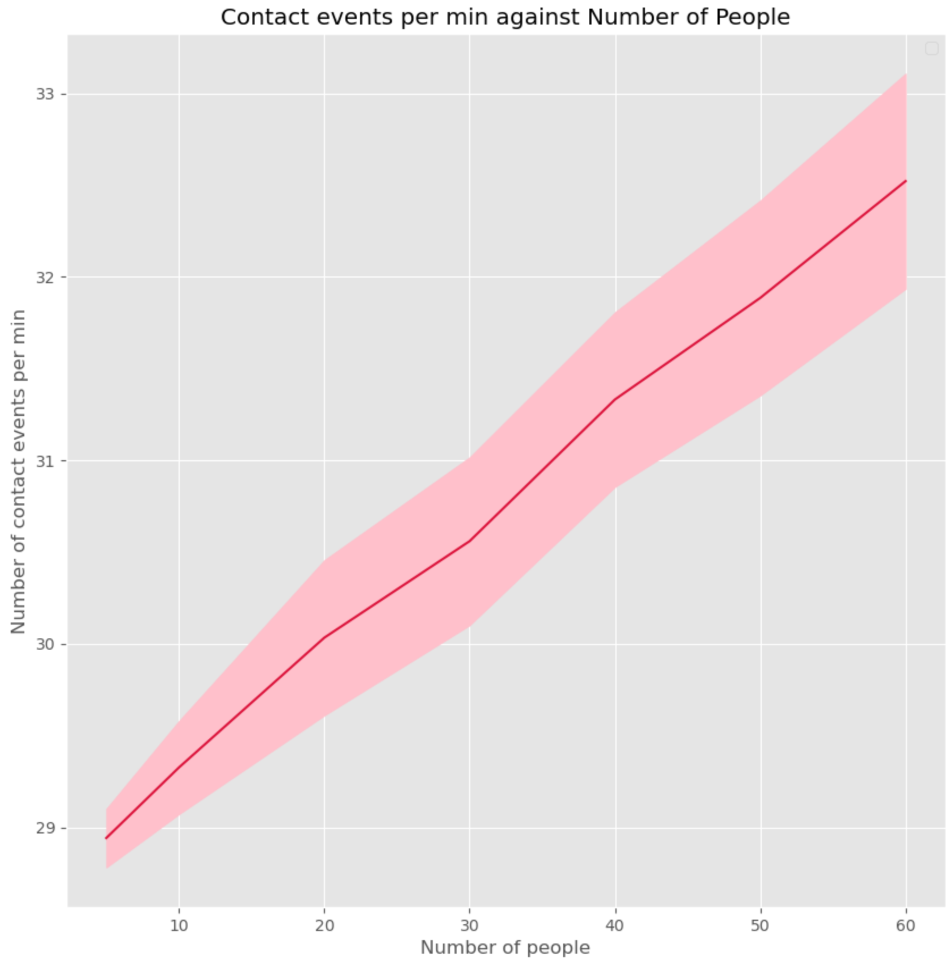

Figure 12: Comparing the average number of contact events per minute to the number of people in the room.

linear relationship, see Fig.12, the slope of this line can be compared to other layouts but more optimisation must be done to view the relationship at higher population numbers than 60.

4 Discussion

The random walk system has worked to correctly simulate COVID-19 spread through the population, the effects of mask usage are clearly seen when comparing the infection rates and the percentage of people infected in an average shopping time. More interesting I have proved that the layout of the store does have an impact in the spread of COVID at this macro scale, and this software can be used to design room layouts that minimise the spread.

The validity of the simulation results depends on the properties of masks and the behaviour of COVID-19 aerosol droplets, if any further advancements are made on this the code has the ability to adapt to the new science and produce more accurate results. I have learnt the importance of assumptions and methods to creating usable numerical simulations on complected scenarios, I have only scratched the surface of the effects of layout and more work should be done on this to find a more clear relationship.

5 Conclusion

The biggest factor in the transmitting of COVID-19 was the type of or lack of mask used by the starting contagious person, now that you are no longer required by law to isolate when testing positive I have shown that the best option now is to wear the most protecting mask you own to stop the transfer to other people in social hubs, even if social distancing is no longer practiced.

The layout of a room can have an impact on the spread of COVID, this software and future iterations will be able to guide people into designing rooms that can passively limit the spread of COVID.

6 Appendix

6.1 Random Walk

The random walk is the process describing the motion of an object in space, it is taken in random steps defined by a probability density and produces a path that can be used in many types of simulations. Brownian motion, animal movement, quantum theories and financial economics are some uses.

6.1.1 Examples

It is used in polymer physics to describe an ideal chain and other module properties. One paper [BB87] uses a closed random walk to represent the properties of an entangled polymer molecule.

The random walk hypothesis in economics is the hypothesis that prices on the stock market and other markets (such as crypto) evolve as a random walk. As its only a hypnosis there are many opions on this, one paper [LM87] outright rejects it for the period of (1962-1985) but others[Lee92] find some evidence in the weekly return series (the rises of stocks over a week, a common parameter used to decide investment)

6.2 scipy.optimize.curvefit

During this project I used the curve fit function from scipy, this uses non-linear least squares to fit a function defined by the user to data. It returns the optimal values for the parameters while minimizing the sum of the squared residuals.

References

[AGS+20] Javed Akhtar, Abner Luna Garcia, Leonardo Saenz, Sarada Kuravi, Fangjun () Shu, and Krishna Kota. Can face masks offer protection from airborne sneeze and cough droplets in close-up, face-to-face human interactions?—a quantitative study. *Physics of Fluids*, 32(12):127112, 2020.

[BB87] M G Brereton and C Butler. A topological problem in polymer physics: configurational and mechanical properties of a random walk enclosing a constant are. *Journal of Physics A: Mathematical and General*, 20(12):3955–3968, aug 1987.

[BB21] John T. Brooks and Jay C. Butler. Effectiveness of Mask Wearing to Control Community Spread of SARS-CoV-2. *JAMA*, 325(10):998–999, 03 2021.

[Ell20] Ralph Ellis. Who changes stance, says public should wear masks. *webMD*, 2020.

[fDCP21] Centers for Disease Control and Prevention. Improve the fit and filtration of your mask to reduce the spread of covid-19. *Archive*, 2021.

[fNS22] The Office for National Statistics. Coronavirus and the social impacts on great britain, 1 april 2022 dataset. *Health and well-being*, 2022.

[Hau18] Nina Hauptman. The average walking stride length. *CHRON*, 2018.

[HHL+21] Jeremy Howard, Austin Huang, Zhiyuan Li, Zeynep Tufekci, Vladimir Zdimal, Helene-Mari van der Westhuizen, Arne von Delft, Amy Price, Lex Fridman, Lei-Han Tang, Viola Tang, Gregory L. Watson, Christina E. Bax, Reshama Shaikh, Frederik Questier, Danny Hernandez, Larry F. Chu, Christina M. Ramirez, and Anne W. Rimoin. An evidence review of face masks against covid-19. *Proceedings of the National Academy of Sciences*, 118(4):e2014564118, 2021.

[How20] Jacqueline Howard. Who stands by recommendation to not wear masks if you are not sick or not caring for someone who is sick. *CNN*, 2020.

[Jac22] Oscar Jackson. Random walk covid-19 simulation. *Github: https://github.com/orangeduice*, 2022.

[Lak20] Rebecca Lake. Grocery shopping statistics: 23 fun size facts to know. *CreditDonkey*, 2020.

[Lee92] Unro Lee. Do stock prices follow random walk?:: Some international evidence. *International Review of Economics Finance*, 1(4):315–327, 1992.

[LM87] Andrew W Lo and A. Craig MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. Working Paper 2168, National Bureau of Economic Research, February 1987.

[VHMea21] Erik Volz, Verity Hill, John T. McCrone, and Anna Price et al. Evaluating the effects of sars-cov-2 spike mutation d614g on transmissibility and pathogenicity. *Cell*, 184(1):64–75.e11, 2021.

[VN16] Julie Vasquez-Nicholson. Uk supermarket chain profiles 2016. *GAIN Report*, 2016.