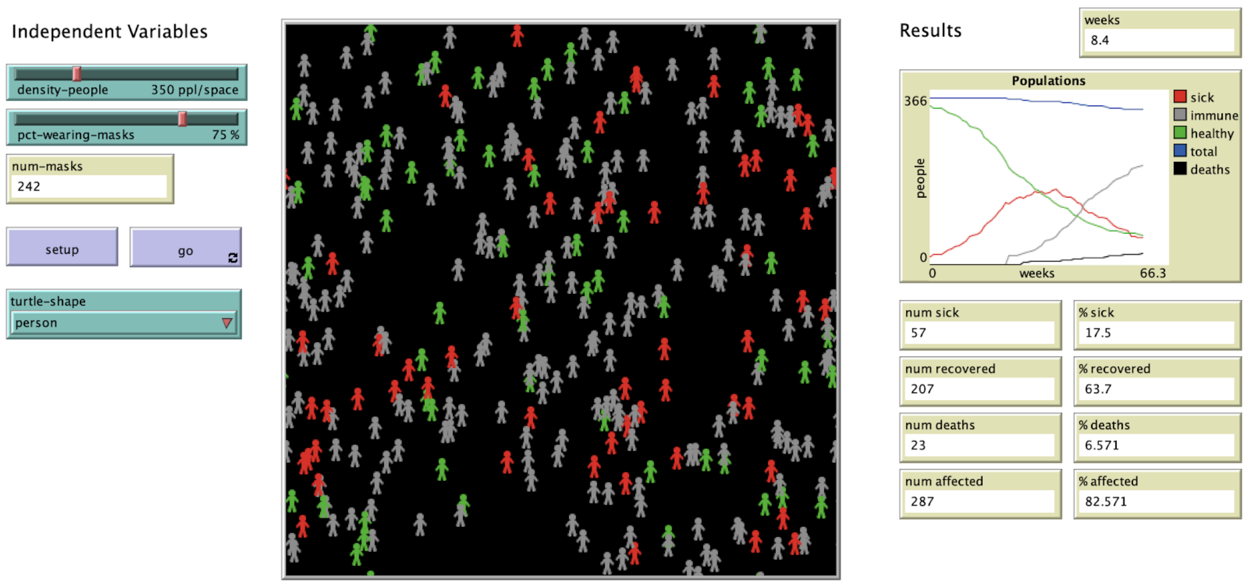
Impacts of Safety Precautions on COVID-19

OIDD 325: Case 4

Grace Jiang, Dana Cohen

I. Experimental Overview

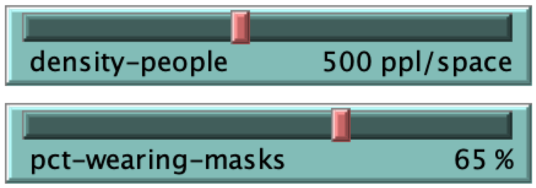
We wanted to conduct an experiment to see how safety guidelines such as social distancing and wearing a mask can affect the spread of viruses such as COVID-19.



II. Experiment Setup

A. Independent Variables

We modified 2 **independent variables** in this experiment:



- **Degree of Social Distancing** (`density-people`): We quantified the degree of social distancing by introducing a new variable, density-people. Higher densities represent a lower degree of social distancing, and vise versa.
- **% Population Wearing Mask** (`pct-wearing-mask`): Percentage of population wearing a mask. This percentage directly affected the infectiousness of the disease (`infectiousness`) in the following way:
    - **Mask Off**: When two turtles both did not wear a mask, we used an infection rate of 99%.
    - **Mask On**: When at least one of the turtles wore a mask, we used an infection rate of 20%, taken from a study in Beijing.

B. Control Variables

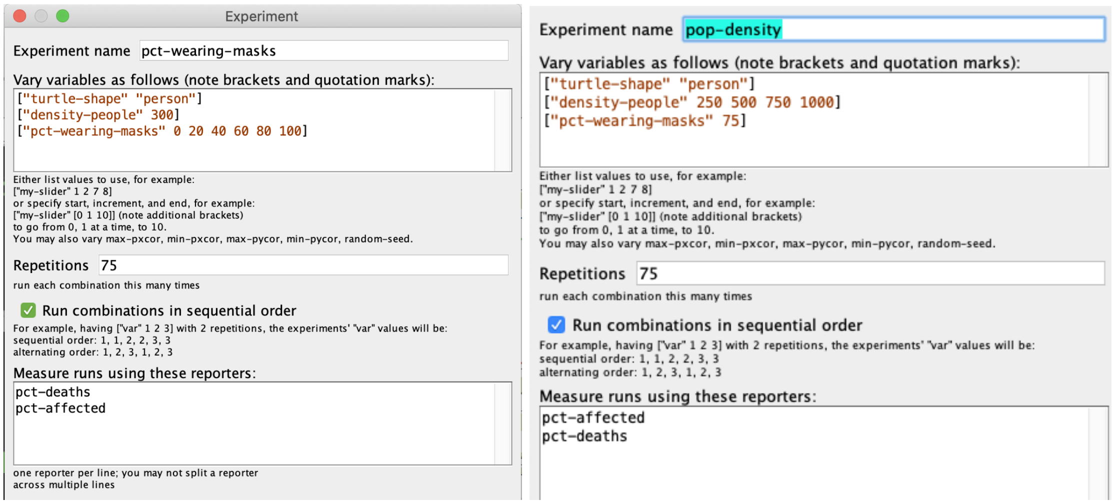
While testing each of the independent variables, we treated all the other independent variables as **control variables**, each with their own default value as listed below:

- **Degree of Social Distancing** (`density-people`): 300 people/space
- **% Population Wearing Mask** (`pct-wearing-mask`): 75%

Some other **control variables** that we kept constant were:

- **Death Rate** (`chance-recover`): 90%. A turtle has a 90% chance of recovering from the virus and becoming immune, and a 10% chance of dying from the virus. We chose this percentage as a high-end estimate to the severity of COVID, especially for older age groups.

- **Virus Duration** (`duration`): 21 days. The virus has a duration period of 21 days before the patient either recovers or dies.



Behavior Spaces we Ran

C. Dependent Variables

Finally, we looked to several **dependent variables** to measure the impacts of the independent variables on the spread and impact of the virus. We kept track of several metrics, such as:

- **Percent of Deaths in Population** (`pct-deaths`): This reporter keeps track of the percentage of deaths from the virus happen from the total population.
- **Percent of Population Affected** (`pct-affected`): This reporter keeps track of the percentage of people affected from the virus from the total population. We defined “affected” as anyone who has been sick from the virus, whether or not they recovered, died, or are still sick.

III. Analysis

A. Varying Degrees of Social Distancing

To measure how much varying social distancing affected the infectiousness and impact of the virus, we looked at two main reporters: **% of population affected** by the virus and **% of population died** by the virus.

Here were our results:

```
#pivot table of population density vs deaths at the end of the run
pd.pivot_table(df1, index=['density-people'],
               values = ['pct-deaths'], aggfunc = [np.mean])
```

	mean
	pct-affected
density-people	
250	65.839301
500	83.545705
750	88.746230
1000	91.415213

	mean
	pct-deaths
density-people	
250	82.970667
500	89.837333
750	89.889778
1000	90.092000

From a quick glance, there seems to be a positive correlation between the density of people (lower social distancing) and the percentage of people affected and dead from the virus. To confirm this observation, we ran an ANOVA-test on the values to see if there was a significant difference between the 4 baselines of social distancing.

H_0: Varying densities of people has no correlation with percentage of people affected and percentage of people dead
H_1: Varying densities of people has some correlation with percentage of people affected and percentage of people dead

Here were our results:

```
import scipy.stats as stats

#one-way anova for densities vs pct-affected
affected250 = df[df['density-people'] == 250]['pct-affected']
affected500 = df[df['density-people'] == 500]['pct-affected']
affected750 = df[df['density-people'] == 750]['pct-affected']
affected1000 = df[df['density-people'] == 1000]['pct-affected']
stats.f_oneway(affected250, affected500, affected750, affected1000)
```

	pct-affected	**pct-deaths**
t-statistic	779.3134097	226.102168
p-value	0.0	3.0542718e-76

Because both p-values were less than α = 0.01, we were able to reject the null hypothesis. In other words, our ANOVA tests confirmed that there was a significant correlation between people density, percentage of people affected, and percentage of people dead.

Our final observation was that the percentage of people affected goes up with decreasing returns for proportional increases in density. In other words, the density of people after a certain point begins to matter less as infections happen to almost everyone regardless.

Significance: Based on our results, we can confirm that the density of people in a space affects the infectiousness and overall death rates of the virus. This is in-line with what doctors and epidemiologists say: not gathering in big spaces (especially enclosed ones) with many people is one way to keep down infections and deaths.

B. Varying Percentages of Mask Wearing

To measure how much varying percentages of mask-wearing affected the infectiousness and impact of the virus, we looked at two main reporters: **% of population affected** by the virus and **% of population died** by the virus.

Here were our results:

```
#pivot table of mask wearing vs percent affected over full time period
pd.pivot_table(df, index=['pct-wearing-masks'],
               values = ['pct-affected'], aggfunc = [np.mean])
```

	mean
	pct-affected
pct-wearing-masks	
0	79.171640
20	75.796601
40	74.638007
60	63.949679
80	45.319220
100	9.954145

	mean
	pct-deaths
pct-wearing-masks	
0	9.666667
20	8.666667
40	8.333333
60	5.733333
80	4.200000
100	1.866667

From a quick glance, there seems to be a negative correlation between the percentage of the population wearing a mask and the percentage of people affected and dead from the virus. In other words, the higher percentage of the population that wears masks, the less people are affected and dead from the virus. To confirm this observation, we ran an ANOVA-test on the values to see if there was a significant difference between the different percentages of mask-wearing.

H_0: Varying percentages of wearing a mask has no correlation with percentage of people affected and percentage of people dead

H$_1$: Varying percentages of wearing a mask has some correlation with percentage of people affected and percentage of people dead

```
import scipy.stats as stats

#one-way anova for densities vs pct-affected
affected0 = df[df['pct-wearing-masks'] == 0]['pct-affected']
affected20 = df[df['pct-wearing-masks'] == 20]['pct-affected']
affected40 = df[df['pct-wearing-masks'] == 40]['pct-affected']
affected60 = df[df['pct-wearing-masks'] == 60]['pct-affected']
affected80 = df[df['pct-wearing-masks'] == 80]['pct-affected']
affected100 = df[df['pct-wearing-masks'] == 100]['pct-affected']
stats.f_oneway(affected0, affected20, affected40, affected60, affected80, affected100)
```

	pct-affected	**pct-deaths**
t-statistic	76676.14384	57638.166414
p-value	0.0	0.0

Because both p-values were less than $\alpha = 0.01$, we were able to reject the null hypothesis. In other words, our ANOVA tests confirmed that there was a significant correlation between percentage of people wearing masks, percentage of people affected, and percentage of people dead.

Significance: Our results confirm that wearing a mask is statistically significant in reducing the percentage of people infected, and consequently the percentage of people dying. Therefore, they are in line with the current nationwide requirement for people going into stores and public spaces to wear masks in order to reduce infection rates.

V. Conclusion

Our model demonstrates how COVID can spread in a population and how two key safety guidelines often discussed currently in the news affects transmission of infection and mortality.

Future modifications to our experiment could include considering factors such as changing infectiousness rate based off if the receiver is wearing a mask (right now, our model only affects infectiousness if the sick person is wearing a mask), varying death rates for the different age groups, and introducing factors such as superspreaders, asymptomatic people, etc.

Our model validates the state-mandated practices of social distancing and wearing masks as being statistically significant to lowering percentage of population infected as well as percentage of population dead as a result of covid as a result. Despite some of the simplifications and assumptions in our model, it should still be regarded as informative that the actions we are taking are effective and correct.