# Simulating the Spread of COVID-19

Aleksey Panas, Anand Singh, Jannate Rahim, Rohit Shetty

*5 November 2021*

## 1 Project Description

COVID-19 is a disease caused by a virus called SARS-CoV-2. This disease emerged around December 2019 causing millions of deaths all across the world. The virus spreads from an infected person to others through respiratory droplets and aerosols. The practice of social distancing and vaccines are factors that helped reduce the spread of the virus. Much data has been collected to analyze the exponential spread of Covid-19, and assist with slowing it. The data sets are crucial in scientific research as they allow us to test or check future insights and will be quite necessary for future long-term studies.

The goal of our project is to present a simplified model that will simulate the spread of an epidemic in a given population. With this model, we will try to simulate some real-world conditions and interpret how these conditions affect the spread of an epidemic(Covid-19) in this case.

A 'curve' in our context is a graphical representation of the onset of illness among cases associated with an outbreak. The term "Flatten the Curve" means to spread out the rate of infection to prevent an overwhelmed health care system and infrastructure. During our research, we will explore many factors that affect the curve. We will also analyze how each of the preventative measures (social distancing, quarantine, vaccinations, etc) and socio-economic factors (population density, access to healthcare, etc) affect this curve.

The question we want to ask is, **given the nature of a particular urban center or town, can we accurately predict the outcome of a Covid outbreak if it were to happen there** and identify causing factors that lead to an undesirable "curve" outcome?

With the gradual improvement of the simulation through feedback from real-world data and scientific research, we might get a usable tool for predicting the risk an epidemic poses on regions with particular characteristics. This would allow testing policies in advance, and foreseeing the leading causes of the spread of infection.

## 2 Relevant Dataset

A relevant dataset that we have found could be useful to our project comes from the Government of Canada. This dataset holds information about various Toronto neighborhoods and their CoVID data. The dataset is available in many formats, including CSV and JSON, and contains statistics on the population of each neighborhood, such as the percentage of elderly people, density of the population, percent young people, and other similar numbers. Alongside this, the dataset shows CoVID case progression, testing practices, and outbreaks. There are a total of 61 variable columns, some of which do not necessarily pertain to CoVID since the dataset also contains air pollution information. Below is a trimmed sample of the first 20 observations within this data showing only 15 of the 61 columns that might be most crucial to our analysis.

**Link to Dataset:**

https://open.canada.ca/data/en/dataset/2d86f026-10b 4-44ac-a68b-80a9dd5dd390/resource/ab558292-2e62-4b71-944b- aa6c19cc5d41?inner_span=True

**Trimmed Sample:**

| id | Neighbourhood name | population | population density | neighbourhood area | % unemployed | % young | % crowded housing | % over 65 | % over 85 | % public transit | F cases | M cases | outbreak cases | tests per 1k |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | West Humber-Clairville | 33312 | 1117 | 30 | 9.6 | 21.2 | 17.4 | 14.9 | 1.8 | 28.1 | 216 | 250 | 200 | 63.7 |
| 2 | Mount Olive-Silverstone-Jamestown | 32954 | 7291 | 5 | 12.1 | 30.7 | 30.9 | 10.8 | 0.9 | 32.6 | 270 | 313 | 68 | 50.6 |
| 3 | Thistletown-Beaumond Heights | 10360 | 3130 | 3 | 10.4 | 24.7 | 15.9 | 18.1 | 3.4 | 23.7 | 68 | 67 | 27 | 68.9 |
| 4 | Rexdale-Kipling | 10529 | 4229 | 2 | 10.9 | 22.5 | 14.0 | 16.4 | 2.8 | 28.4 | 48 | 55 | 77 | 67.9 |
| 5 | Elms-Old Rexdale | 9456 | 3306 | 3 | 10 | 24.5 | 17.7 | 13.5 | 1.5 | 32.9 | 68 | 57 | 12 | 63.7 |
| 6 | Kingsview Village-The Westway | 22000 | 4356 | 5 | 9.9 | 23.5 | 20.6 | 16.3 | 2.6 | 30.2 | 178 | 215 | 39 | 63.4 |
| 7 | Willowridge-Martingrove-Richview | 22156 | 4007 | 6 | 8.5 | 18.5 | 11.1 | 22.1 | 4.0 | 25.4 | 108 | 95 | 27 | 76.3 |
| 8 | Humber Heights-Westmount | 10948 | 3981 | 3 | 7.4 | 18.0 | 8.7 | 27.8 | 8.7 | 27.6 | 28 | 29 | 153 | 89.0 |
| 9 | Edenbridge-Humber Valley | 15535 | 2840 | 5 | 6.1 | 13.2 | 8.2 | 21.2 | 4.3 | 28.3 | 35 | 32 | 80 | 104.9 |
| 10 | Princess-Rosethorn | 11051 | 2138 | 5 | 6 | 10.3 | 2.6 | 18.3 | 2.9 | 19.8 | 17 | 19 | 7 | 98.8 |
| 11 | Eringate-Centennial-West Deane | 18588 | 2171 | 9 | 7.4 | 15.3 | 6.6 | 20.3 | 3.0 | 22.4 | 56 | 45 | 17 | 78.6 |
| 12 | Markland Wood | 10554 | 3614 | 3 | 6.2 | 11.7 | 5.6 | 25.6 | 5.1 | 23.3 | 19 | 16 | 9 | 88.4 |
| 13 | Etobicoke West Mall | 11848 | 6582 | 2 | 7.9 | 17.7 | 14.6 | 17.4 | 4.3 | 31.5 | 32 | 31 | 87 | 75.2 |
| 14 | Islington-City Centre West | 43965 | 2712 | 16 | 7.3 | 11.9 | 10.7 | 16.8 | 3.4 | 37.2 | 125 | 118 | 248 | 84.6 |
| 15 | Kingsway South | 9271 | 3593 | 3 | 7.5 | 7.2 | 1.5 | 21.5 | 3.5 | 31.7 | 8 | 9 | 32 | 126.8 |
| 16 | Stonegate-Queensway | 25051 | 3199 | 8 | 6.7 | 13.1 | 7.4 | 16.6 | 3.0 | 29.8 | 33 | 40 | 21 | 95.0 |
| 17 | Mimico (includes Humber Bay Shores) | 33964 | 4915 | 7 | 6.2 | 12.0 | 8.0 | 15.2 | 1.7 | 28.0 | 73 | 88 | 64 | 86.7 |
| 18 | New Toronto | 11463 | 3342 | 3 | 8.7 | 16.6 | 9.1 | 13.3 | 1.2 | 32.9 | 29 | 29 | 57 | 99.2 |
| 19 | Long Branch | 10084 | 4584 | 2 | 7.1 | 15.1 | 9.4 | 13.9 | 1.4 | 30.0 | 17 | 19 | 10 | 87.5 |
| 20 | Alderwood | 12054 | 2435 | 5 | 6.1 | 19.5 | 7.0 | 16.7 | 2.7 | 21.3 | 23 | 34 | 13 | 94.4 |

## 3 Computational Plan

Since we are building a simulation model, we will not manipulate any input data. Rather, it will generate simulated data and use real-world data to show comparisons. With that, here is a brief summary of the flow our model will follow.

The model will use Pygame as a GUI engine (which we know it isn't perfect for, but it will do!) to allow the user to initiate new and keep track of existing simulations. Simulations will not be a black box, but rather, will display a real-time visual of people in the form of dots (or perhaps something prettier if computational power permits) and will simulate their basic day-to-day behaviors. Prior to the start of the simulation, the characteristics of the location being simulated will be specifiable, allowing the user to indicate the density of the city, the population's inclination or reluctance to wear masks, adhere to social distancing, or vaccinate, and other such variables. Each "person" created for the simulation will then have their personality traits randomized based on those initial settings.

As a technicality, the graphics display of the simulation will be decoupled from the computation via threading, allowing the user to speed up or slow down the simulation at will, and even potentially introduce factors midway. As the simulation runs, simulated data will be collected and saved to a database or CSV file depending on user preference. At any time, the user can save and quit a simulation for later and view the simulated data side-by-side with imported real-world data for comparison. These comparisons can be made fairly feature-rich if enough time is dedicated, allowing for side-by-side plots of real-world vs simulated to be made intractable, allowing to overlay, subtract, or add them, and maybe other operations that seem necessary based on the data being compared.

## 4 Works Cited


"The University of Alabama at Birmingham." *UAB*,
http://www.uab.edu/.

*Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention,
http://www.cdc.gov/.

"What Is Coronavirus?" *Johns Hopkins Medicine*
https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus.

"These Simulations Show How to Flatten the Coronavirus Growth Curve."
*The Washington Post*, WP Company, 14 Mar. 2020, https://www.washingtonpost.com/graphics/2020/world/corona-simulator/.

"Toronto Air Pollution and Covid-19 Data by Neighbourhood - Toronto Air Pollution and COVID-19 Data by Neighbourhood." *Open Government Portal* ,
https://open.canada.ca/data/en/dataset/2d86f026-10b4-44ac-a68b-80a9dd5dd390/resource/ab558292-2e62-4b71-944b-aa6c19cc5d41?innerspan=True.