COVID-19 Spread in Ontario, CA Based on Foursquare Location Data

Using k-Means Clustering

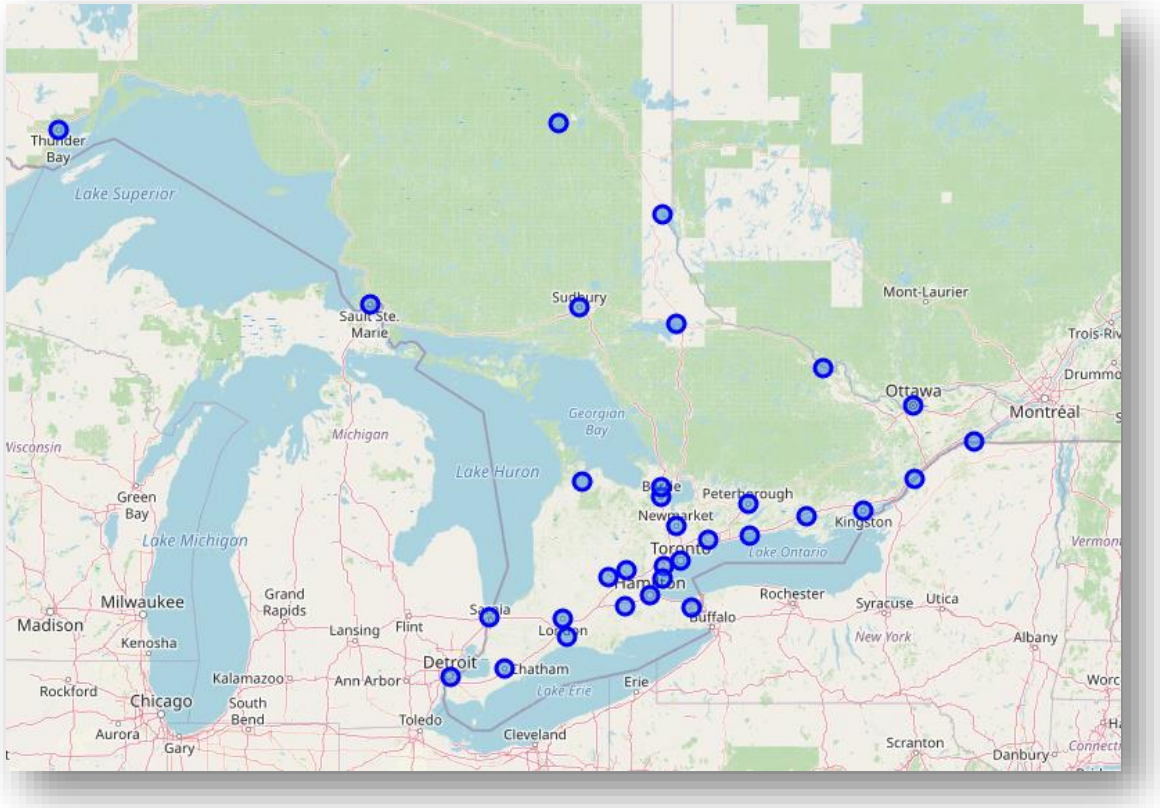
COVID-19 is spreading fast in the world

- Governments around the world are racing with time to find the way to stop the spread
- Human to human transmission is possible
- Human is social by nature
- Does social distancing work in curbing the spread?
- For the purpose of the project, chosen area is Ontario, Canada

Data acquisition

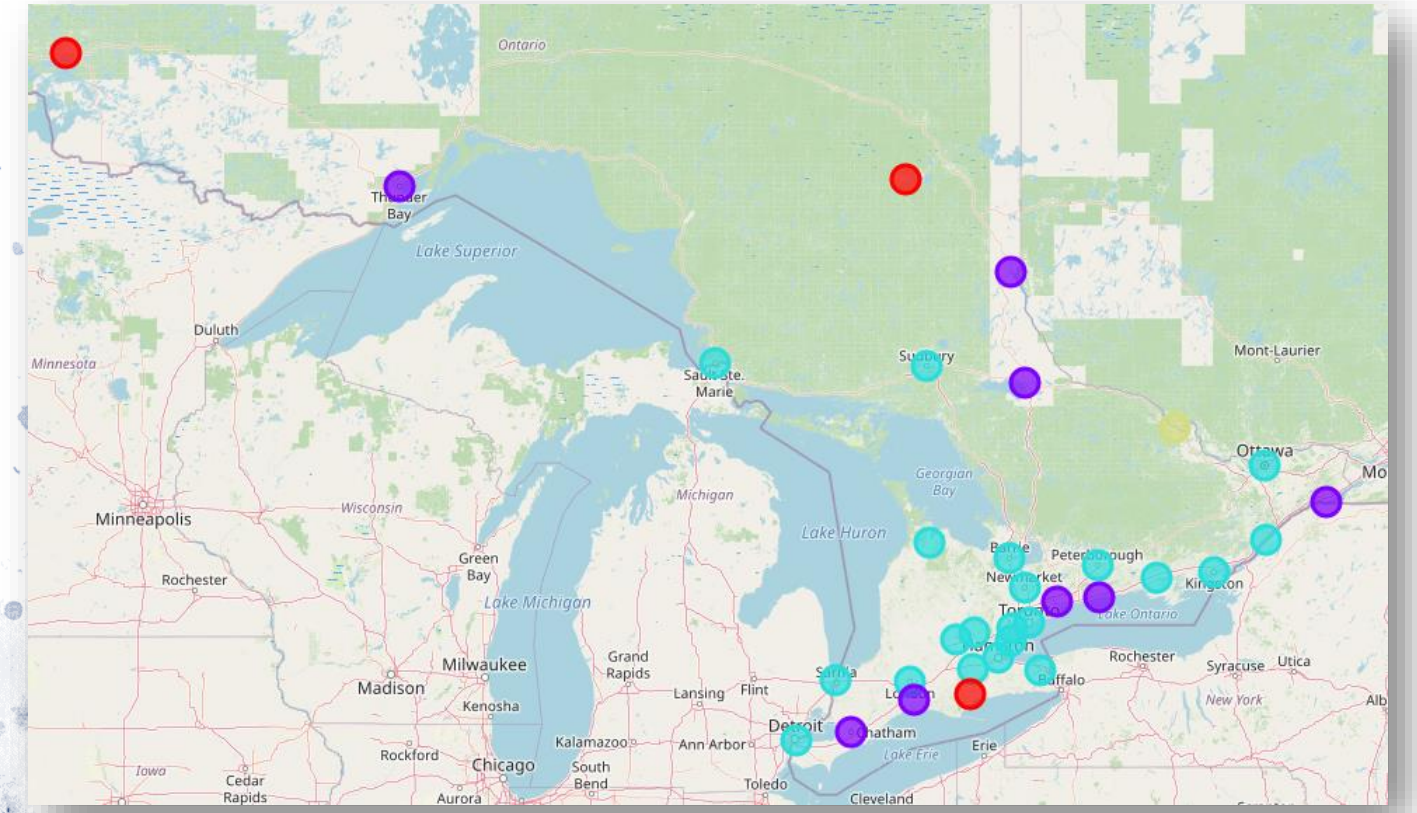
- List of municipality in Ontario, Canada in 2020, data is scraped from https://en.wikipedia.org/wiki/List_of_municipalities_in_Ontario.
- List of COVID-19 positive cases in Ontario, Canada as of 22nd April 2020, data is available publicly in https://data.ontario.ca/dataset/confirmed-positive-cases-of-covid-19-in-ontario.
- Geolocation data of the municipality, acquired from Geopy
- Location data of venues in Ontario, Canada retrieved using Foursquare API

Mapping the Affected Area in Ontario, Canada



The municipality which has confirmed positive cases of COVID-19 is indicated with Blue dots

Clustering Affected Municipality using k-Means Algorithm (k=4)



- There are four clusters
- The municipalities in the same cluster have the same colour
- The features for clustering are from Foursquare API data

Average Positive Cases of COVID-19 in each Cluster

Cluster	Num of City/Municipal	Average Total	Average Contact or Neither
0	3	76.33	58.33
1	8	134.50	111.12
2	22	496.64	226.05
3	1	14.00	4.00

- The highest average is cluster #2, in general each municipality has public gathering place such as pub, coffee house
- Similar case with cluster #1, although cluster #1 has more varied category
- Cluster #0 and #3 have small average number of positive cases

Conclusion and Future

- Cluster of municipalities which has highest average of COVID-19 positive cases happen to have similar features of popular public gathering spaces
- Public gathering space may contribute to human to human transmission of COVID-19 in Ontario, Canada
- This project is limited to Foursquare API data, other variables may contribute to the COVID-19 spread