Week 1 - Battle of the Neighbourhoods – Viewing Toronto

The Problem

Toronto is a very varied city that has some great places to live in. The video linked below is looking at 140 neighbourhoods from a magazine 'Toronto Life' and compares various areas to live.

https://youtu.be/7RF1yl29Jms

https://torontolife.com/neighbourhood-rankings/#

The problem now at the start of June 2020 is we have a Corona 19 Virus which is a Pandemic – world wide. So the problem we need to solve is: Where are the rate of infections the lowest? Also as in many places – (if anything like the UK) - are closed so I want to find areas where there are parks to walk and exercise. The problem therefore is not looking at the Toronto Life magazine neighbourhood rankings but looking at low risk areas which have nearby parks.

How to tackle the problem and what data to use:

The first thing to do is to get a list of the Neighbourhoods. I produced a table (data frame) using https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

However only 10 Neighbourhood name match the about list as taken from the Open Data Toronto dataset. https://open.toronto.ca/dataset/neighbourhoods/

So this was used instead. It is difficult to clean so it can be used, Together with

https://drive.google.com/file/d/1euhrML0rkV_hHF1thiA0G5vSSeZCqxHY/view

to get the Neighbour hood rate of infections. I then produce a bubble plot of the infected areas. The next step was to find two houses, for sale, to choose in different neighbourhoods. Both being in low infected areas These where:

8*8 Boulton Ave, Toronto*

and

489 Merton St, Toronto

The geocoder recommended does no work all the time so I used https://www.whatsmygps.com as we only have two locations to find to get the GPS coordinates.

Next I used the Fouresquare API https://foursquare.com/city-guide to find the parks around both locations.

K-means clustering may then be further used to investigate various groups.