Leveraging Location Data to Model Hotspots for Possible COVID-19 outbreaks in Suburban Neighbourhoods

Shaik Abdillah

June 29, 2021

1 Problem Introduction

The COVID-19 pandemic precipitated in the end of 2019, where a then unknown SARS-like virus started spreading through the global population. The pandemic has caused widespread nation-wide lockdowns beginning in early 2020 and its effects are still in play some 2 years later. Much research has been conducted in an attempt to slow the spread of the virus; these included

1. Wearing Face masks
2. Social Distancing
3. Working from Home

It is thus of interest for public health officials to model possible hotspots based on venues that require unmasking and close physical contact. Thus, this project aims to develop a hotspot map of a neighbourhood based on the existing businesses that provide services which violate the above list of recommended safety measures.

2 Data Collection

The neighbourhood in question is ultimately arbitrary; but the location data used will be from the **FourSquare API**. Businessess will be ranked based on:

1. If the service(s) provided requires unmasking
2. The duration of which customers are present for
3. The possiblity of social distancing

Upon this three risk factors, a data set of types of business will be given points, one for each risk factor that they fall under. For example: a barbershop requires the stylist to be in close proximity to the client for an extended time and the client has to be unmasked for an extended period of time. Thus the barbershop will be given 3 points. On the other hand, a Retail shop requires only for the customers to be in close proximity to each other for a short period of time and thus will be given 1 point.

3 Modelling

The neighbourhood will be clustered using K-means clustering to determine centroids of businesses within a particular neighbourhood. Each centroid will then be ranked based on how many risk points all of the businesses in its vicinity and then plotted on a map to provide possible hostpots for local health officials to continually monitor.