



# **IBM Applied Data Science Capstone Project: Finding the best areas to set up hospitals in Toronto, Canada**

Manasa Sriram



# Introduction

## Business Problem:

To find locations where there are no hospitals or very few hospitals so that they can cover a larger area of people to provide healthcare facilities to a wider group of people



# Data

## Data required and sources

- List of neighborhoods in Toronto, Canada: Scraping of Toronto Neighborhoods from the wikipedia page for the list of neighborhoods:  
[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- Latitude and longitude of the neighborhoods: Getting the latitude and longitude of these neighborhoods from the Geocoder Package
- Data on the venues present in the neighborhood: Using Foursquare API to get venue data related to these neighborhoods.



# Methodology

- The data on the neighborhoods in Toronto, Canada was extracted from the above given wikipedia URL using BeautifulSoup
- Using [Toronto location data by IBM](#), the latitude and longitude for each postal code was found.
- Using the folium and geocoders libraries, we visualise the dataframe in the form of an interactive map.
- The machine learning concept used was KMeans clustering



## Results

From the data it was found out that there was one hospital, Toronto Western Hospital located in Kensington Market / Chinatown / Grange Park.

From this it can be recommended to set up new hospitals in neighborhoods away from the above mentioned one as this would be the most efficient solution



## Discussion

As seen from the data source, we only take the postal codes that start from 'M'.

We could get a more accurate solution to our problem if we have the data of all the postal codes but that would make the program more complex