**The Battle of Neighbourhoods**

**Introduction/Business Problem**

**Problem Statement-**

**Whenever someone looks for opening up a restaurant or any other business venture the main problem faced by them is choosing a location for the set up.**

**But in this era of high knowledge availability we have come up with a solution where using the neighbourhood, location of different points we can optimize the best location problem.**

**For this purpose, I am using Toronto data set to come up with best possible solution for the problem as stated above.**

**Solution for the problem-**

In this project we will try to find an optimal location for an Italian restaurant. Specifically, this report will be targeted to stakeholders interested in opening an Italian restaurant in Toronto, Canada.

Since there are lots of restaurants in Toronto we will try to detect locations on the basis of

1. not already crowded with restaurants.
2. areas with no restaurants in vicinity.
3. close to city centre (downtown Toronto) as possible, assuming that first two conditions are met.

**About Data-**

We will use our data science knowledge to generate a few most promising neighbourhoods based on this criterion. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

For this problem we'll be required to explore, segment, and cluster the neighbourhoods in the city of Toronto. However, the neighbourhood data is not readily available on the internet. For the Toronto neighbourhood data, a Wikipedia page exists [[https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M]](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M%5D) that had all the information I need to explore and cluster the neighbourhoods in Toronto. I scraped the Wikipedia page and wrangled the data, cleaned it, and then read it into a pandas data frame so that it is in a structured format to proceed with my analysis.  
Toronto, the capital of the province of Ontario, is a major Canadian city along Lake Ontario’s north-western shore. It's a dynamic metropolis with a core of soaring skyscrapers, all dwarfed by the iconic, free-standing CN Tower. Toronto also has many green spaces, from the orderly oval of Queen’s Park to 400-acre High Park and its trails, sports facilities and zoo.

Based on definition of our problem, factors that will influence our decision are:

* number of existing restaurants in the neighbourhood (any type of restaurant)
* number of and distance to Italian restaurants in the neighbourhood, if any
* distance of neighbourhood from city centre

We decided to use regularly spaced grid of locations, centred around city centre, to define our neighbourhoods.

Following data sources will be needed to extract/generate the required information:

* For coordinated we'll be using a csv file available with postal code and latitude and longitude information.
* number of restaurants and their type and location in every neighbourhood will be obtained using **Foursquare API**
* coordinate of centre will be obtained using **Google Maps API geocoding**.