**Places for setting up new business in Toronto**

**1. Introduction**

**1.1 Description & Discussion of the Background**

Toronto is the provincial capital of Ontario and the most popular city in Canada. It has a population of 2,731,571. Toronto is an international center of business, finance, arts and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. A business owner decides to open a restaurant where there is a lower real estate cost. At the same time, he may also want to choose the district according to population density. However, it is difficult to obtain information that will guide investors in the direction.

**1.2 Problem**

Data that might contribute to determining proper location to start business include the location of other restaurant, where the real estate index is placed on Toronto and each district is clustered according to the venue density. This project aims to predict where should be perfect place for opening a restaurant considering a large number of factors.

**2. Data acquisition and cleaning**

**2.1 Data sources**

To consider the problem we can list the data as below:

* Neighbourhood, postal code and borough data can be found in Wikipedia page, [https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M,](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)  Data is cleaned and only Toronto area is selected.
* The most common venues of given Borough of Toronto are found in Foursquare API.
* The center coordinates of each Borough are found from <https://cocl.us/Geospatial_data>

**2.2 Data cleaning**

Data downloaded or scraped from multiple sources were combined into one table.

There are several problems with the datasets. The features “Borough” and “Neighbourhood’ have some missing values “Not assigned”. The missing values of “Borough” are dropped, and the missing values of “Neighbourhood” are replaced by the corresponding “Borough” name. In addition, neighbourhoods under same postal code are combined together. In addition, coordinates (latitude, longitude) are added into data according to its borough.

Since the business owner wants to start his business in Toronto, so only “East Toronto”, “West Toronto”, “Downtown Toronto”, “Central Toronto” are left, other areas are removed from data.