**Capstone Project – The Battle of Neighborhoods**

**Data description:**

For the initial data set, I will start by extracting the data from Open Data Soft (ODS) to get a list of the neighborhoods in Chicago. ODS has thousands of datasets available publicly that can be filtered according to state, keywords, etc. For Toronto, the data will be extracted from the Wiki link that was used in the previous exercise. Both links are provided below:

Chicago dataset:

<https://public.opendatasoft.com/>



Toronto dataset:

<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>



Once these datasets are extracted and read into a pandas data frame, I will first filter for the respective cities and clean the data. This will ensure there are no duplicates or null values and that the data is consistent. To gather the latitude and longitude values, we will be using the **geocoder** package and create a function that will get the coordinates for each neighborhood. We will also need data about the different venues in different neighborhoods for both cities. In order to get this information, we will be using the **Foursquare API** to make calls. Foursquare provides data on location which includes multiple venues and areas of interest in a given radius. This can also include venue names, locations, and photos. After finding the list of neighborhoods and creating the respective data frames, we will connect with the Foursquare API to get information about the location within a 500-meter radius.

The data retrieved from the Foursquare API will contain information such as:

* Name of neighborhood
* Neighborhood latitude and longitude
* Name of venue
* Venue latitude and longitude
* Type of venue

We will be using a clustering approach to explore and segment the neighborhoods. In the end, we can look at the top venues that were visited in each cluster and determine answers to the following questions:

* Comparing Chicago and Toronto, which city has more restaurants as the 1st most visited venue?
* In particular, are their certain neighborhoods that do not have restaurants in their “top 10 visited” list?
* Is there a specific type of restaurant that we can open that isn’t already there?

Doing so can help our potential businessman in choosing the best city and location for opening his/her restaurant. In addition, this could also help future businessmen in determining the different types of business that are already largely available and what type of business they can invest in or open in these locations.