**CAPSTONE PROJECT**

**BATTLE OF THE NEIGHBOURHOOD**

**Introduction and problem description**

The surroundings of Lawrence park neighbourhood appears to have a lot of schools and universities. This implies that there are a lot of young people residing or moving in and out of this neighbourhood and there might be a demand for business that are aimed towards the younger generation. The problem description here is what kind of business can a businessman undertake in this neighbourhood.

**Data**

The data required is available on the Wikipedia page and foursquare API can be used.

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

**Methodology**

* Using the list of neighborhoods, we connect to the Foursquare API to gather information about venues inside near of Central Toronto borough. We take the radius to be 1000 meter. It means that we have asked Foursquare to find venues that are at most 1000 meter far from Central Toronto borough.
* Processing the Data and Creating a Data Frame for All the Venues inside the Central Toronto borough. This actions are working with language Python.
* When the data is completely ordered, we will perform processing on that raw data to find our desirable features for each venue. Our main feature is the category of that venue. After this stage, the column "Venue's Category" will be One-hot encoded and different venues will have different feature-columns. After On-hot encoding we will integrate all venues columns to one column and joint columns to "Total Joints" column. After applying k-Means Clustering of Machine Learning Techniques.

