**Data:**

We would be using the following publicly available data to achieve this objective:

* List of Neighborhoods (and boroughs) in Toronto
* Geo Coordinates of neighborhoods (Latitudes and Longitudes)
* Venue data of Shopping malls to identify competition and cluster neighborhoods

**How to use the Data:**

The Wiki page: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> gives list of neighborhoods and boroughs in Toronto, Canada. We will scrap this data to extract data from this page using Python and beautifulsoup package. We will then use the Geocode package to fetch the coordinates of the boroughs. After fetching the Latitude and Longitude data of neighborhoods, we will use the Foursquare API to obtain the venue data that is the shopping malls to identify the competition in a neighborhood. Foursquare is one of the biggest location based databases available.

Steps:

* Scraping Toronto neighborhoods data from Wiki page
* Fetch geographical coordinates using Geocoder in Python
* Visualize Toronto’s map
* Use Foursquare to fetch shopping malls data
* Group data by neighborhood and calculate mean of frequency of each category
* Narrow down the data to Shopping malls
* Use k-means clustering to cluster the locations
* Visualize the data using Folium