# Capstone Project:
# The Battle of Neighbourhoods
# (Week 1)

## 1. Introduction

### 1.1 Background

Toronto, Canada is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. More than 50 percent of residents belong to a visible minority population group, and over 200 distinct ethnic origins are represented among its inhabitants. In 2010, it was estimated over 100,000 immigrants arrive in the Greater Toronto Area each year. However even within the diversity and the hustle and bustle familiarity and oneness is what one seeks. We, Good Host are an upcoming start-up providing housing solutions that go beyond just the rent.

### 1.2 Business Problem

Given the massive immigrant/ expat population in Toronto, we want to leverage our data analysis expertise and tap into this market and offer a more personalised approach to those seeking housing options. While searching for accommodation avoiding culture shocks to seek familiarity is most often observed. We wish to build an app providing such solutions providing highly customised criteria for selection of preferred locations to live. We are still in the development stage of the app and have released a pilot model using a sample study approach. The sample chosen is from India and is looking for accommodation having nearby access to Indian food, Indian supermarkets, and Indian neighbours.
Given the criteria we are going to consider one parameter initially. The presence of Indian restaurants, which would indicate that it is within a locality wherein its main customers, mostly Indians, live/work nearby. Procurement of ingredients would also be from within the vicinity, indicating India markets nearby.

# 2. Data Collection and Analysis

The input data required to perform the analysis and clustering are as follows:

- List of neighbourhoods in Toronto, Canada

- Latitude and the longitude coordinates of each neighbourhood.

- Venue data for Indian restaurants.

To explore and cluster the neighbourhoods the following will have to be performed:

- Scrapping, wrangling and cleaning of data on Toronto neighbourhoods from the link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- Including the latitude and longitude coordinates of the neighbourhoods using Geocoder Package

- Venue data of neighbourhoods using Foursquare API