

# CAPSTONE - APPLIED DATA SCIENCE FINAL SUBMISSION

## Week 4

### A description of the problem and a discussion of the background

One of my friends is looking for entering restaurant business in Toronto. He observed that restaurant sector is growing due to busyness of people, less tendency to cook at home and more appetite to spend less time during meals. He couldn't decide what kind of restaurant to open and in which location. He is not willing to open a big restaurant due to his budget constraints. He asked me to recommend which restaurant and where to open.

### A description of the data and how it will be used to solve the problem

The data which I thought to use

- Toronto Demographic (Wikipedia): In order to determine the restaurant type (in terms of cuisine), I decided to use population changes in Toronto in terms of ethnic groups. I will use ethnic groups because we can predict a type of foods for our restaurant by ethnic groups. As there are data for year 2016 and 2011, I will calculate the increase in population changes in the ethnic groups. At the same time this data will not be enough for us to determine because we also need to take the percentage of this ethnic group in overall population into consideration in order to eliminate noises and outliers.
- Toronto borough, neighborhoods lists (Wikipedia, locations): We will use postal codes in order to have a clean list of boroughs with neighborhood values.
- Venues in that locations (from Foursquare): We will use Foursquare to find out how many venues are there in each neighborhood, density of venues in that location, what type of venues there are.