#### 1. Introduction

### 1.1 Background

Fast food restaurants most of the time are used by people who want to have a quick lunch break between their daily activities. Given this background they are ideal for high transit zones with students/office workers that don't have a lot of free time to wait for their food and want to spend their resting time eating outdoors.

### 1.2 Investors interest - Requests

The investors want to set up a new fast food restaurant branch in the city of Toronto. They want to know what would be a good location to set up their business. The conditions asked for this project is that the restaurant needs to be near two universities, preferably near the city center and student campus with high people transit. The branch is big enough to attract customers so competition in the area is not a problem for them.

# 2. Data acquisition and methodology

#### 2.1 Data sources

All the datasets used are public. The ones that I will be using to solve the investors request are the following:

- Toronto Adult Education <a href="https://open.toronto.ca/dataset/wellbeing-youth-alternative-education">https://open.toronto.ca/dataset/wellbeing-youth-alternative-education</a>
- Toronto Postal Codes for geolocation –
  <a href="https://en.wikipedia.org/w/index.php?title=list\_of\_postal\_codes\_of\_Canada:\_M&oldid=1">https://en.wikipedia.org/w/index.php?title=list\_of\_postal\_codes\_of\_Canada:\_M&oldid=1</a>
  011037969
- FourSquare geolocation Data for venues in Toronto

## 2.2 Methodology

I'll first be using K-cluster nearest neighbor to pinpoint the clusters with a high concentration of venues. The data will be manipulated to find places with high concentration of normal restaurants/coffee shops and find were the investors fast food restaurant will face less competition.

In addition to that I'll map the universities and campuses in the city in order to define high transit areas with students. In order to do this I'll map them and after that superimpose the locations to the venues clusters.

## 2.3 Data Analysis

# 2.3.1 Methodology

In order to effectively process the data I'll sort the venues by neighborhood using FourSquare API resulting in charts like these:

	Nei ghborho od	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	- 79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	- 79.2928 69	Neighborhood
4	The Danforth West, Riverdale	43.679557	-79.352188	Men Essentials	43.677820	-79.351265	Cosmetics Shop

I'll also use superimposed maps between universities locations and venue data to get the best location for the restaurant.