

# **Finding locations to open a Coffeehouse in Toronto using Machine Learning algorithms.**

## **1. Introduction**

### **1.1 Background**

In this Capstone Project, I am creating a hypothetical scenario for an entrepreneur who wants to open 'Coffeehouse' in the Toronto area. The idea behind this project is that there may be many Coffee shops or Coffeehouses in the Toronto area and it might be difficult for an entrepreneur to find the exact location to open a coffee house. So, by using some techniques of data science we will guide him, on which location he should open a shop to gain maximum profit. Coffee house is an establishment that primarily serves coffee(of various types, e.g. espresso, latte, cappuccino). Some coffeehouses may serve cold drinks such as iced coffee and iced tea; in continental Europe, cafés serve alcoholic drinks. A coffeehouse may also serve food such as light snacks,sandwiches, muffins or pastries. From a cultural standpoint, coffeehouses largely serve as centers of social interaction: the coffeehouse provides patrons with a place to congregate, talk, read, write, entertain one another, or pass the time, whether individually or in small groups. A coffeehouse can serve as an informal club for its regular members. By considering all these points about coffeehouse , finding the best location to open such a coffeehouse is one of the most important decisions for the entrepreneur and I am designing this project to guide him find the most suitable location.

### **1.2 Business Problem**

The objective of this project is to find the most suitable location for the entrepreneur to open a coffeehouse in the Toronto area. So, by using some techniques of data science and machine learning methods such as clustering we will guide him, on which location he should open a shop to gain maximum profit. This project aims to provide solutions to answer the business question: If an entrepreneur wants to open a coffeehouse in toronto,where should he consider to open a coffeehouse?

### **1.3 Target Audience**

The entrepreneur who wants to find the best location to open a coffeehouse.

## 2. Data

To solve this problem we will need data which is associated with the Toronto area. The list of data which will be used is listed below.

1. List of neighborhoods in Toronto, Canada.
2. Latitude and Longitude of these neighborhoods.
3. Venue data related to Coffeehouse. Venue data will help us find the neighborhoods that are most suitable to open a coffeehouse.

## 3. Extracting the data

1. Scrapping of Toronto neighborhoods via Wikipedia.  
([https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M))
2. Getting Latitude and Longitude data of these neighborhoods via a csv file and merging into the Toronto neighborhood data.  
([https://cocl.us/Geospatial\\_data](https://cocl.us/Geospatial_data))
3. Using Foursquare API to get venue data related to these neighborhoods.