

# Business Proposal

## Introduction:

The purpose of this Project is to help people in exploring better facilities around their neighborhood. It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in **North York, Toronto**.

Lots of people are migrating to various states of Canada and this project is for those people who are looking for better neighborhoods. For ease of accessing to Cafe, School, Super market, medical shops, grocery shops, mall, theatre, hospital, like-minded people, etc.

## Purpose:

The major purpose of this project, is to suggest a better neighborhood in a new city for the person who are shifting there. Social presence in society in terms of like-minded people. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

## Foursquare API:

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

## Work Flow:

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

## Clustering Approach:

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like New York and Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

## Libraries used:

- **Pandas:** For creating and manipulating dataframes.
- **Folium:** Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.
- **Scikit Learn:** For importing k-means clustering.
- **JSON:** Library to handle JSON files.
- **XML:** To separate data from presentation and XML stores data in plain text format.
- **Geocoder:** To retrieve Location Data.
- **Beautiful Soup and Requests:** To scrap and library to handle http requests.
- **Matplotlib:** Python Plotting Module.