Finding optimal locations to open restaurant/grocery businesses.

Introduction

An international grocery and restaurant chain looking forward opening their business locations in the city of Toronto. They wanted to identify optimum locations having maximum businesses potential and required to generate business intelligence to form a strategy in establishing their new business locations.

In the week 3 assignment we note that the Toronto city has 140 postal zip codes assigned to 103 different boroughs. This project will conduct analyzing population demographics, financial and household data in those neighborhoods and cluster them based on their similarity. It will also find the existing venues creating competition (e.g. Restaurants, Grocery stores) and other venues in the proximity which adds new businesses opportunities.

Data

Data Sources

Source #1: City of Toronto's Open Data Catalogue

URL: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/

Contains information licensed under the Open Government License – Toronto.

Name	Description	File Format
Population Demographics	Population data by neighborhood including population by age categories, culture and ethnicity.	csv
Safety	Crime incidents by neighborhood	csv
Economy	Income, credit rating	csv
Hosing	Households, House price	csv

The data from Open Data Catalogue will be used to cluster neighborhoods based on their similarity characteristics. This will help the business to group neighborhoods when forming custom business strategies to their targeted neighborhoods. This data will also be used in finding the optimum business locations.

Source #2: Datasets used in the Week3 Assignment, Neighborhood Segmentation and Clustering

Name	Description	File Format
Postal Codes Dataset	List of postal codes of Toronto, Canada web scrapped from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M	csv
Geospatial Coordinates of Postal Codes	Geographical coordinates of each postal code: http://cocl.us/Geospatial_data provided in the assignment.	csv

Source #3: Foursquare APIs location data

URL: https://developer.foursquare.com

Name	Description	File Format
Venues	Location, Venue Name, Venue Type, Ratings (if any)	API

The foursquare dataset will be used to identify competitive business locations in each neighborhood (e.g. Grocery stores, restaurants) as well as venues which adds new businesses opportunities (e.g. Schools, Offices, Attractions, Shopping Malls, etc.).