Week5_Capstone Project -
The Battle of Neighborhoods-Presentation

Friday, February 5, 2021 11:49 PM

Introduction

In this project we will do some researches from Toronto city's data to find a new restaurant in Toronto.
First, we will try to find the best location which someone wants to open a new restaurant in the city, and which types of income group will be attracted most to it based on the population of the neighborhood.

Business Problem

Background

Toronto is the provincial capital of Ontario and the most populous city in Canada and the fourth most populous city in North America, which is with a population of 2,731,571 in 2016. The Toronto census metropolitan area (CMA), of which the majority is within the Greater Toronto Area (GTA), held a population of 5,928,040, making it Canada's most populous CMA. Toronto is the fastest growing city in North America. and is the anchor of an urban agglomeration, known as the Golden Horseshoe in Southern Ontario, located on the northwestern shore of Lake Ontario.

Toronto encompasses a geographical area formerly administered by many separate municipalities. Former municipalities include East York, Etobicoke, Forest Hill, Mimico, North York, Parkdale, Scarborough, Swansea, Weston and York. Throughout the city there exist hundreds of small neighbourhoods and some larger neighbourhoods covering a few square kilometres.

There is vast population and big geographical area, there also exists big competition between businesses. Therefore it became very challenging for stake holder or new business to decide which area they should start their business to get higher revenue with lowest possible competition.

Purpose

Target Audience:
New Business who wants to open a restaurant in Toronto.

The objective of this report is to determine the best possible location to open a Restaurant in Toronto, based on the different localities of the city, already established restaurants in varios geographical location and ease of accessibility by maximum number of people so that the revenue from the latest venture can be maximized.

Therefore I focus on that borough during my analysis. We define potential neighborhood based on the number of restaurant which are operating right in each neighborhood. Toronto has full potential but also is a very challenging district to open a business because of high competition. New restaurant should be open in an area that inadequate neighborhood in this way the restaurant can attract more customers. Therefore, this analysis necessary to ensure that we have enough

customers and that we are not so close to other restaurants places，make sure it will be a steady and sustainable business.

Data

In this assignment we will do some researches from Toronto city's data.

First, we will try to find the best location which someone wants to open a new restaurant in the city, and which types of income group will be attracted most to it based on the population of the neighbourhood.

For my analysis I choose some data resources:\
1、 City of Toronto Neighbourhood Profiles use this Census data to provide a portrait of the demographic, social and economic characteristics of the people and households in each City of Toronto neighbourhood. \
The profiles present selected highlights from the data, but these accompanying data files provide the full data set assembled for each neighbourhood.\

Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"

2、 Load Toronto geospatial coordinates and merge to Toronto Postal Code Data

Link: http://cocl.us/Geospatial_data \
Then, using geospatial mege to toronto pastal code data

3、 Toronto neighborhoods populations broken down by postal code

Link：https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Tables/File.cfm?T=1201&SR=1&RPP=9999&PR=0&CMA=0&CSD=0&S=22&O=A&Lang=Eng&OFT=CSV

4、 Toronto neighborhoods average after tax income broken down by postal code \
joined income data to the toronto data,then saved a copy of the data set for checking and explaination

Canadian families and unattached individuals had a median after-tax income of $57,000 in 2016.

Link： https://www150.statcan.gc.ca/n1/daily-quotidien/180313/dq180313a-eng.htm

5、Using foursquare data to get information about restaurants in Toronto

Link: https://foursquare.com/explore?mode=url&ne=44.418088%2C-78.362732&q=Restaurant&sw=42.742978%2C-80.554504

6、Onehot encode and count restaurants

Methodology

K-Means Clustering.

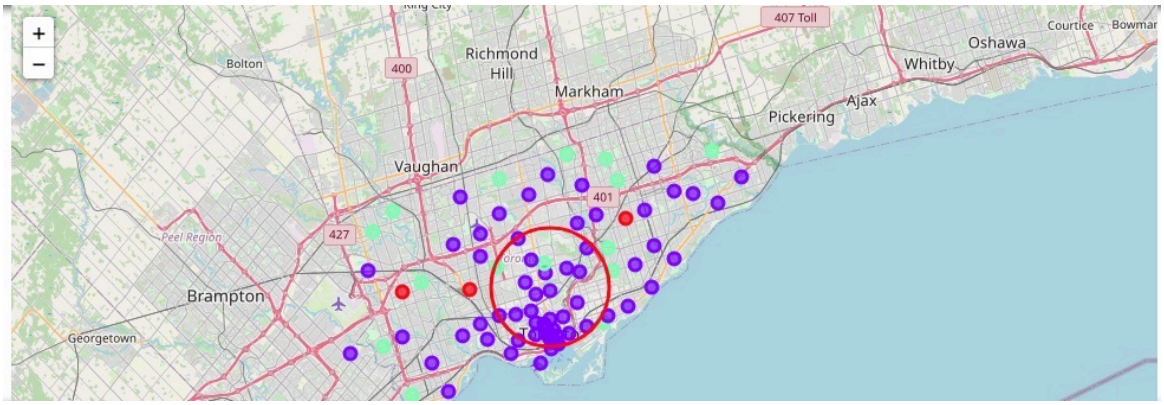
I chose K-Means clustering as the algorithm:
K-means clustering is an iterative clustering algorithm where the number of clusters K is predetermined and the algorithm iteratively assigns each data point to one of the K clusters based on the feature similarity. \
Which is refered from https://towardsdatascience.com/clustering-algorithms-for-customer-segmentation-af637c6830ac

Approach

1、Collect the Toronto city data from Wikipedia.
2、Using FourSquare API we will find all venues for each neighborhood.
3、Pointing out how much poputlation each neighborhoods is.
4、Matching The average after tax income for all of these neighborhoods.
3、Filter out all venues that are nearby by locality.
4、Using aggregative rating for each resturant to find the best places.
5、Visualize the Ranking of neighborhoods using folium library(python)

Result

The exact Address to locate would be: 268 Balliol Street, ON M4S 1C2, Canada or lat: 43.6991598, lng: -79.3878871



Discussion

We got a glimpse of the Restaurants in Toronto and were able to find out some interesting insights which might be useful to people with business interests. Let's summarize our findings:

As we built our list of neighborhoods with Restaurant venues exclusively we discovered most neighborhoods were similar and the greatest concentration of restaurants was in Central Toronto and downtown Toronto. This might seem obvious but it would also appear that these are some of the most affluent neighborhoods in Toronto so there appears to be correlation. By Locating in the general vicinity of the Exact location my friend could be geographically centered in this cluster and poised to service his restaurant customer base with the greatest efficiency.

When we built our our K-Means dataset we used Silhouette analysis to tell us there was a lot of similarity between neighborhoods and the most common restaurants contained with in. Really there was only 2 types of cluster or neighborhoods in greater Toronto. The vast majority of those were in 1 cluster. So Toronto restaurants might be many but they are very homogeneously located near the center of Toronto.

Of the 103 Toronto Neighborhoods gathered only 55.3% or 57 Neighborhoods are above the median after-tax income. 37.8% or 39 Neighborhoods are below he median after-tax income. 6.7% or 7 neighborhoods did not register as it appears their populations are too low. It appears that the greatest concentration of affluence is near

central Toronto. We decided to keep all neighborhoods in the dataset regardless of income of population as the majority were close enough.

Conclusion

The recommendation is backed up with demonstrated data analysis, purpose was to serch the optimal location for a new restaurant. While nothing can ever be 100% certain ， it will certainly be better informed than the stakeholder was prior to asking for my help.

Final decission on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (easy for car parking), proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.