Capstone Project - The Battle of the Neighborhoods

Diego Albuquerque

1. **Introduction\Bussines problem**

   Background

   The company XZY, have the interest to open a new restaurant Toronto, the speciality of company is regular restaurants always near a Parks, in this bussiness model the restaurants have the operation hours,menu, events and promotions based in the commom use of Parks, it is a specitility of the company.

   Seen this, we have the necessity of use data aproach to base the decision of in what place put a new restaurant.

   Problem

   Based in this context, we hav the question: **Where is the best place to open a new restaurante in Toronto?**, for this answer is necessary to find a place near a park and that not have many restaurants arround, in the vision of stakeholders, how less competition, best!.

2. **Data**

   To find the better place we have many option, drive a car arround the city finding places, use a helicopter or drone to fly above the city or the most easily and cheap is use the foursquare data do see the parks and restaurants arround then.

   In this way, we will map all the parks of Toronto, after all the restaurants and find the parks with less restaurants arround.

3. Libraries Which are Used

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution and Venues

XML: In Webscrap to separate data from presentation and XML stores data in plain text format.

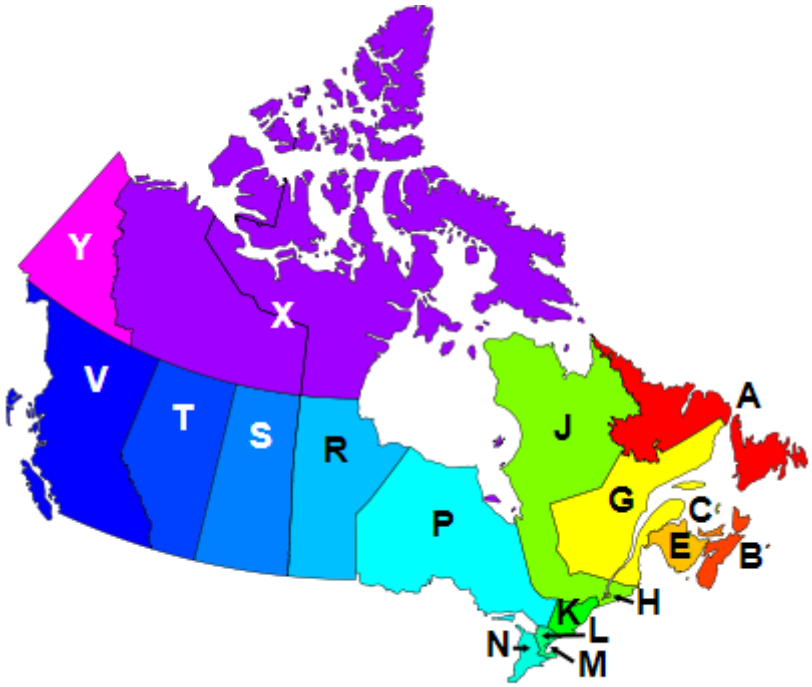
Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module

4. Data acquisition and cleaning

We will start collecting data from wikipedia with the list of postal codes in Toronto, however initially we have to understand how postal codes work in Canada, after a survey it was seen that postal codes follow a pattern of letters as shown below:



Therefore, we saw that Toronto is located in the region served by the letter “M”, knowing this we will see the list of postal codes of Toronto by the page:

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

We now have the following pattern:

Toronto - 103 FSAs

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0. However, the postal code M0R 8T0 is assigned to an Amazon warehouse in Mississauga, and th suggesting that Canada Post may have reserved the M0 FSA for high volume addresses.

M1A *Not assigned*	M2A *Not assigned*	M3A North York (Parkwoods)	M4A North York (Victoria Village)	M5A Downtown Toronto (Regent Park / Harbourfront)	M6A North York (Lawrence Manor / Lawrence Heights)	M7 Qu (Or
M1B Scarborough (Malvern / Rouge)	M2B *Not assigned*	M3B North York (Don Mills) North	M4B East York (Parkview Hill / Woodbine Gardens)	M5B Downtown Toronto (Garden District, Ryerson)	M6B North York (Glencairn)	M7 No
M1C Scarborough (Rouge Hill / Port Union / Highland Creek)	M2C *Not assigned*	M3C North York (Don Mills) South	M4C East York (Woodbine Heights)	M5C Downtown Toronto (St. James Town)	M6C York (Humewood-Cedarvale)	M7 No

With this information in hand, we will perform the “webcraping” to mount pandas on our dataframe, and with the “geocoder” search for latitude and longitude.

5. Exploratory Data Analysis

When we use the “Foursquare API” we have already managed to arrive at some analyzes:

Venues	Amount
Park	50
Restaurants (regular and fast food)	59

As we can see, we have a list of 1.18 restaurants for each Park in Toronto.

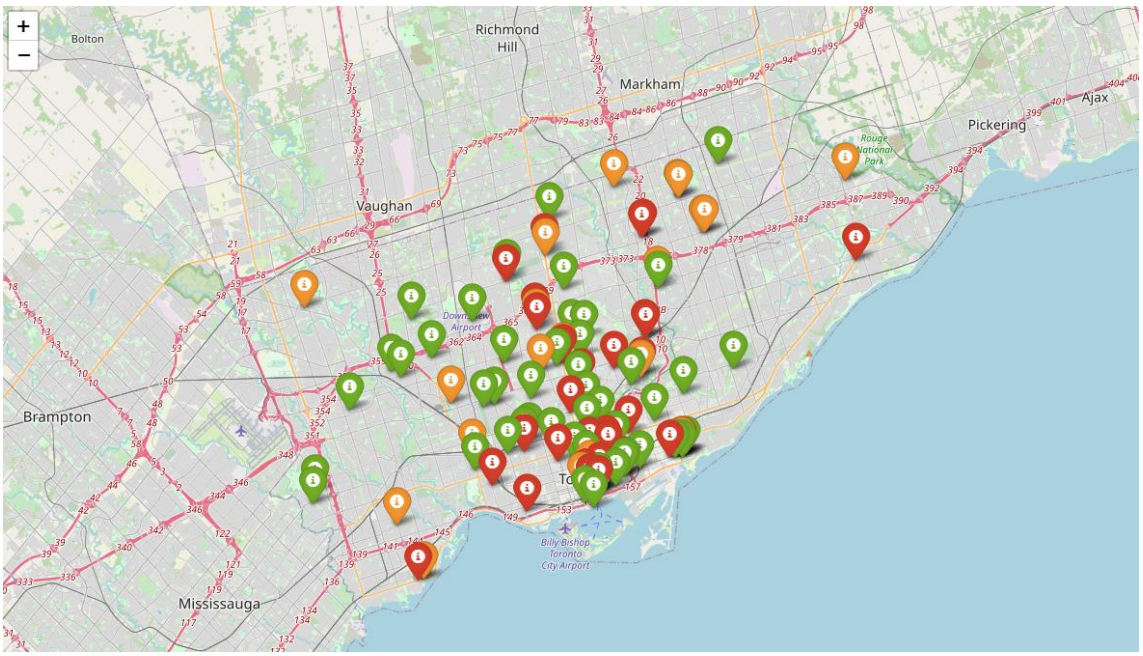
When we look at neighborhoods, we see that Parks are one of the most popular vines in the neighborhood.

Borough	ClusterLabels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
Scarborough	1.0	Playground	Intersection	Park	Escape Room
North York	1.0	Park	Yoga Studio	Deli / Bodega	Escape Room
North York	1.0	Electronics Store	Park	Convenience Store	Yoga Studio
North York	1.0	Fast Food Restaurant	Park	Food & Drink Shop	Yoga Studio
North York	1.0	Airport	Park	Yoga Studio	Deli / Bodega
East York/East Toronto	1.0	Park	Convenience Store	Yoga Studio	Deli / Bodega
Central Toronto	1.0	Park	Swim School	Bus Line	Yoga Studio
Central Toronto	1.0	Park	Tennis Court	Yoga Studio	Deli / Bodega
Downtown Toronto	1.0	Park	Playground	Trail	Dance Studio
York	1.0	Park	Women's Store	Pool	Distribution Center
York	1.0	Park	Convenience Store	Yoga Studio	Deli / Bodega

In this way we will check the map showing the parks in green color

Now let's add the regular restaurants to the map in red

And lastly, let's add the orange fast food restaurants:



In this last chart we can clearly see opportunities for parks with few restaurants around and in return to see the areas where there is greater competition.

6. Conclusion

Purpose of this project was to identify Toronto areas close to parks with low number of restaurants (particularly regular restaurants and fast food

restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new restaurant.

Final decission on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, but the analysis of map and locations showed in this report allow the processing of decision givin more information in in smart way.