Where to open a coffee shop in Toronto?

Felipe de la Mora

April 08th, 2019

1. Introduction

I've been living in Mexico all my life, and I love my country, but something happened when I had the chance to visit Canada, since then I have always dreamed of moving into the beautiful Canada and Toronto is one of the two options I have considered, with Vancouver too. Why do I want to move into Canada? Well, Canada's gross domestic product was 1.76 trillion dollars on 2017 and the GDP per capita was $48,100. Amazing numbers, right? Canada is roughly 3 times the size of Mexico, has more fresh water than any country with millions of lakes and what to say of the famous Canada's hospitality?

Let's imagine that my dream comes true and I am able to move into Toronto, so the next question pops up: If I am opening a coffee shop, where should I do it? What neighborhood will have less coffee shops?

So who will find this analysis helpful? Well, for any person willing to invest in a business (in this case a coffee shop) in the lovely Toronto.

I will look for neighborhoods where Coffee shop is not a common venue, so my (future) coffee shop can stand a chance and will be able to attract new customers easily.

To find out where to open my coffee shop I will do a deep analysis on Toronto's neighborhoods, compare them by using Foursquare data to find what venues are near each location, making it easier to decide where we would like to open a business.

I will follow the course idea of clustering based on venues and frequency of venues, then show them on a map.

2. Data

To perform this analysis, I will be using the list of Postal Codes of Toronto, which is a free resource on table format on Wikipedia allowing me to get all Neighborhoods of Toronto. And it looks something like this:

Toronto - FSAs [edit]

Note: There are no rural FSAs in Toronto, hence no postal codes start with M0.

Postcode	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront
M5A	Downtown Toronto	Regent Park
M6A	North York	Lawrence Heights
M6A	North York	Lawrence Manor
M7A	Queen's Park	Not assigned

By performing web scrapping techniques, I downloaded the table into a data frame that was able to manipulate and merge with the other section of data.

The other part of the data is coming from Foursquare API, where I was able to extract the following fields: Venue Name, Venue Category, Venue Location. This will be key to understand and visualize Toronto's business situation.

Foursquare data looks like this:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	KFC	43.754387	-79.333021	Fast Food Restaurant
2	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
3	Parkwoods	43.753259	-79.329656	TTC stop - 44 Valley Woods	43.755402	-79.333741	Bus Stop
4	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena

3. Methodology

In first step we have collected the required data: location and type (category) of every venue for each neighborhood.

Second step in our analysis will be calculation and exploration of venues frequencies across the neighborhoods - we will create labels for those categories that contain the word coffee, and then we will use heatmaps to identify with low number coffee shops.

In third and final step we will focus on clusters of locations that the most common venues are not coffee shops. I will present map of all such locations but also using k-means clustering of those locations to create clusters and identify neighborhoods that are potential places to put a coffee shop.

4. Results and Discussion

As you can see, each of the clusters fall into a category thanks to K-Means algorithm, for example, we could say that cluster 0 is where you could easily find cafeteria and bars, in case you would like to share a nice afternoon with a friend. On the contrary, cluster 2 is where you find tons of outside activities and places like parks and rivers.

And although, there are a lot of coffee shops in Toronto, there are some neighborhoods that a coffee shop is not one of the most common venues, to prove this, K-Means algorithm helped us to identify 4 clusters where the most common venues are not coffee shops. We could select one of these clusters, like cluster 2where we can easily identify that most of the venues are parks and playgrounds and maybe a coffee shop near would be a good opportunity to invest in or take for example cluster 3, it looks like a financial cluster since most of the venues are banks and we can see some dessert shops and women's store, sounds like a good combination to open a little and cozy coffee shop, right?

This is not a definitive guide, but just an exploratory process to better understand Toronto's neighborhoods and venues, there could be many other factors that could change which neighborhood is suitable to open a coffee shop.

5. Conclusion

The purpose of this project was to use Toronto's neighborhoods data and Foursquare data to find recommended zones where someone could open a Coffee Shop. Clustering with K-Means of locations helped us achieve a better overall look into Toronto's venues distribution and spotted some interesting options, this is providing a picture of how Toronto looks like in terms of businesses.

Decision will be made by adding more factors like people density, pricing, and time series, proximity to people dense places, among others.