The Battle of Neighborhoods

Capstone Project:

Coffee Shop in Toronto

_____

June 17, 2020

_____

Applied Data Science Capstone
IBM Data Science Professional Certificate on Coursera

1. Introduction

1.1. Background

Toronto is the provincial capital of Ontario. With a recorded population of 2,731,571 in 2016, it is the most populous city in Canada and the fourth most populous city in North America. The Greater Toronto Area (GTA) had a 2016 population of 6,417,516. The city covers an area of 630.20 square kilometers (243.32 sq mi) and comprises six districts – East York, Etobicoke, North York, Old Toronto, Scarborough and York – which, were amalgamated to form Toronto's present boundaries in 1998. The city is the anchor of the Golden Horseshoe, an urban agglomeration of 9,245,438 people (as of 2016) surrounding the western end of Lake Ontario. Toronto is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. More than 50 percent of residents belong to a visible minority population group, and over 200 distinct ethnic origins are represented among its inhabitants. While the majority of Torontonians speak English as their primary language, over 160 languages are spoken in the city.

Toronto is a prominent center for music, theatre, motion picture production, and television production, and is home to the headquarters of Canada's major national broadcast networks and media outlets. Its varied cultural institutions, which include numerous museums and galleries, festivals and public events, entertainment districts, national historic sites, and sports activities, attract over 43 million tourists each year. Toronto is known for its many skyscrapers and high-rise buildings, in particular the tallest free-standing structure in the Western Hemisphere, the CN Tower.

The city is home to the Toronto Stock Exchange, the headquarters of Canada's five largest banks, and the headquarters of many large Canadian and multinational corporations. Its economy is highly diversified with strengths in technology, design, financial services, life sciences, education, arts, fashion, aerospace, environmental innovation, food services, and tourism.

1.2. Problem

The objective of this project is to use Foursquare location data and regional clustering of venue information to determine what might be the 'best' neighborhood in Toronto to open a Coffee Shop. This city is home to many coffee shops where coffee addicts find their favorite drink. Whether you're on your way to work looking for a quick cup or a tourist fueling up while visiting the best attractions and parks or you're simply looking for a spot to sit down and map out the best things to do during your local adventures, Toronto is one of the best options to drink a very good coffee.
But...
will we find the most suitable location for an entrepreneur to open a new Coffee Shop in Toronto?

1.3. Interest

This project is aimed towards entrepreneurs or business owners who want to open a Coffee Shop in Toronto.

The success criteria of the project will be a good recommendation of neighborhood choice to open a new Coffee shop based on the lack of this type of establishment in that location, near of neuralgic centers and with little to no competition.

2. Data acquisition and cleaning

2.1. Data sources

For the Toronto neighborhood data, a Wikipedia page exists that has all the information you need to explore and cluster the neighborhoods in Toronto. You will be required to scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe so that it is in a structured format.

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

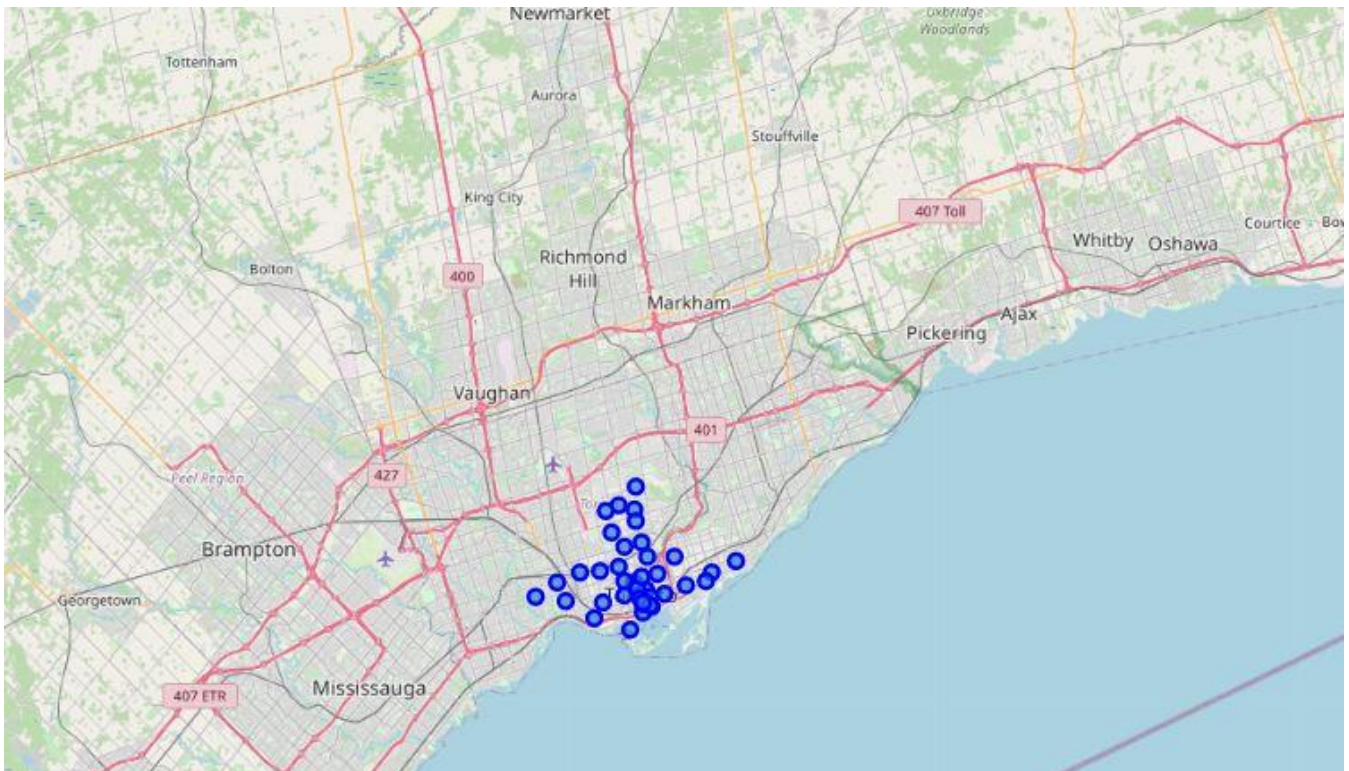|	Postal Code	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills

Now that you have built a dataframe of the postal code of each neighborhood along with the borough name and neighborhood name, in order to utilize the Foursquare location data, you need to get the latitude and the longitude coordinates of each neighborhood. You will use the Geocoder Python package.

URL: https://geocoder.readthedocs.io/index.html

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

Toronto city geographical coordinates data Will be utilized as input for the Foursquare API that Will be leveraged to provision venues information for each neighborhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot
3	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
4	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa

3. Methodology

4. Results

5. Discussion

6. Conclusion