Capstone Project – Paris to Toronto

Samuel O

1. Introduction

1.1 Problem Background

Toronto is by far one of the greatest cities in the Canada and the world at large. With its multicultural diversity and capitalistic nature, it possesses a booming atmosphere for any business to successfully run. The heart of Canada's finance and banking industries lies in Toronto. It is no surprise that it is the Financial capital of Canada. From Bay Street's high-rises, which house corporations, law firms and the Toronto Stock Exchange to IBM, PepsiCo, etc., The city has shown its financial power and range of opportunities from fashion to technology, legal services, sports, real estate, tourism and finance amongst others. It really is a wonder to see and to live in!

For any young graduate looking to progress in his/her career, Toronto provides the perfect atmosphere of job opportunities, culture, housing, places, dining options ranging from cozy cafes to chic power lunch restaurants, etc. Moving from a different city to Toronto can be quite a change thus it is nice to switch easily. As such, it is very important for any newcomer interested in the business space of Toronto to analyze specifically the similarities between the current residency location and Toronto to ease the change of environment. These analyses would enhance decision making and eventually provide an edge to any person because he/she feels comfortable and is ready to withstand competition and survive in the Toronto scene. These analyses would improve understanding and prove to be valuable.

1.2 Problem Description

As a young graduate looking and pursuing a career in Data Science. I just received an offer of a lifetime to work with great engineers at IBM!

The sad part? I must leave Paris for Toronto, Canada. The IBM Toronto Software Lab is the largest software development laboratory in Canada, and IBM's third largest software lab. It is located in Markham, a community in Toronto, Canada.

Paris is known for its culture, sights, venues, rich diversity and in all a calming environment. Toronto also boasts of the same. For me, moving to Toronto would be quite the challenge and the environment I will have to go to is very important to me. There lies my problem, I must find a location that reminds me of home and has the potential to become my new home!

Several venues remind me the most of home but majorly the variety of food to select from after a very long day. Also, the several amenities ranging from the gyms to the clubs to transportation, etc.

The challenge would be to locate a similar community that benefits my transition from Paris to Toronto.

1.3 Interest

This research will be very interesting for any youth or career person moving from one city to another. It would provide a methodology and approach on how to locate the similarities between the two cities and aid decision making. Others who could be interested in this data are researchers on socio-geographical studies of cities by looking at the social factors as well as the geographical factors.

2. **Data Acquisition and Cleaning**

2.1 **Data Sources**

To solve my problem, I will need to source for data that would show me the top venues and locations at the neighbourhood I reside at in Paris. As such I would require:

1.Data on the Latitude and Longitude of my location (Using Nominatim or Google Map Locator)

2.Venues located in the neighbourhood at a specific radius and at a time version of 7th July, 2020. (Applying the Foursquare API)

With this dataset, I have a basis to begin my research.

I would then need to gather data for Toronto (latitude, longitude, venues, ratings):

These data would be sourced from a Wikipedia page containing the Postal Codes of Ontario and a CSV File containing Geospatial Data of the several Ontario Postal Codes:

1.List of Postal Codes for Ontario: (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

2. Geospatial CSV (http://cocl.us/Geospatial_data)

2.2 **Data Cleaning**

The data on my location in Paris was sourced using Google Locator Maps and the Latitude and Longitude were found thus
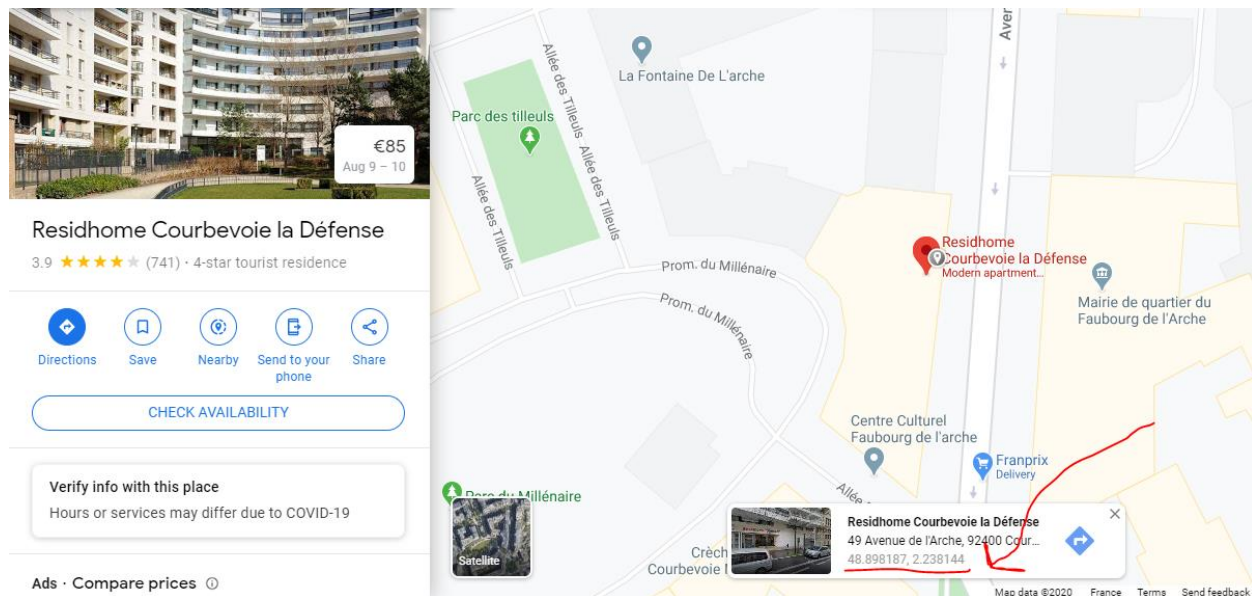
With this and using the Foursquare API, the venues located at a radius of 650m to this residential area in Paris were located and then read into a DataFrame for further analysis later.

	name	categories	lat	lng
0	Parc du Millénaire	Park	48.898194	2.236577
1	Thaïoria	Thai Restaurant	48.900409	2.239551
2	Hôtel Pullman Paris La Défense	Hotel	48.895096	2.239006
3	Starbucks	Coffee Shop	48.896689	2.238164
4	Monoprix	Supermarket	48.896720	2.236800
5	La French Touch	Burger Joint	48.900660	2.232121
6	Grande Arche de la Défense	Monument / Landmark	48.892565	2.235882
7	Place Carpeaux	Plaza	48.893531	2.238529
8	Sushi Fukunoya	Japanese Restaurant	48.900200	2.238700
9	Fnac	Department Store	48.892625	2.239585
10	Villa Min	Korean Restaurant	48.899282	2.245689
11	Decathlon	Sporting Goods Shop	48.892972	2.240244
12	Allée de l'Arche	Plaza	48.894852	2.239296
13	So Thaï	Thai Restaurant	48.896766	2.245842

The data from the Wikipedia page which shows the Postal Codes of Ontario, Canada and the Geospatial data were combined (Will be shown further in Methodology)

This information would be classified by Boroughs and Neighbourhoods and then rated based on the top locations in such areas

In turn the communities would be clustered to aid me by visual and content analysis of the similarities between those clusters and my current location.

All these will in turn aid my Decision Making in picking a similar location and will ease my transition from Paris to Toronto