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):

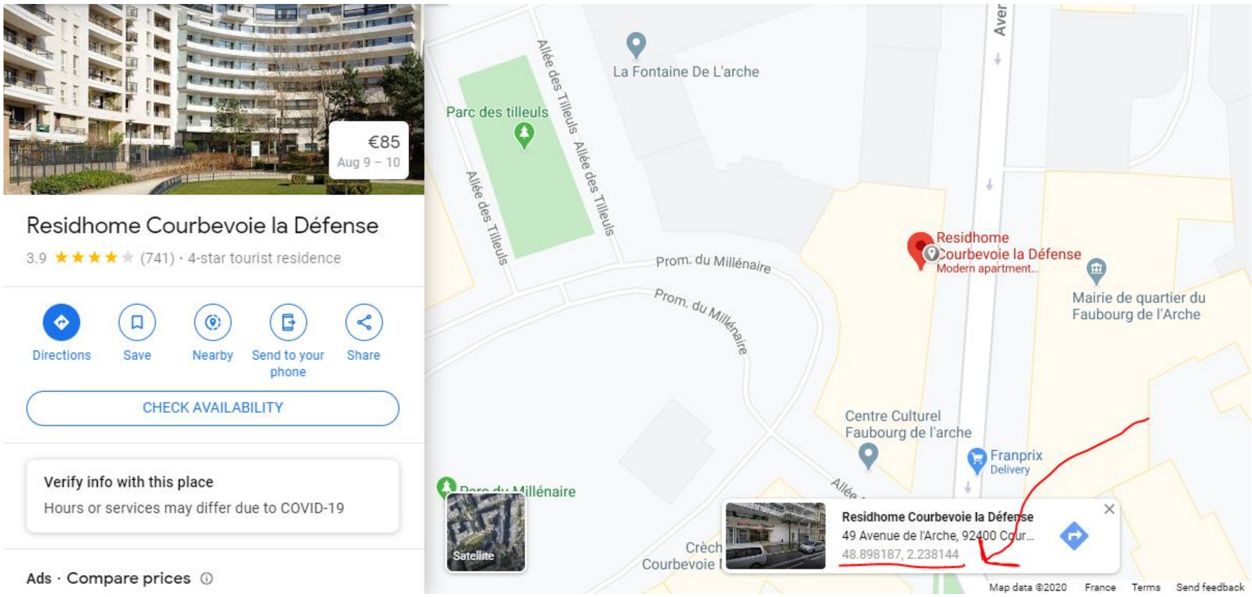
These data would be sourced from a Wikipedia page containing the Postal Codes of Ontario and a CSV File containing Geospatial Data of the several Ontario Postal Codes:

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



Residhome Courbevoie la Défense

3.9 ★★★★☆ (741) · 4-star tourist residence

Verify info with this place

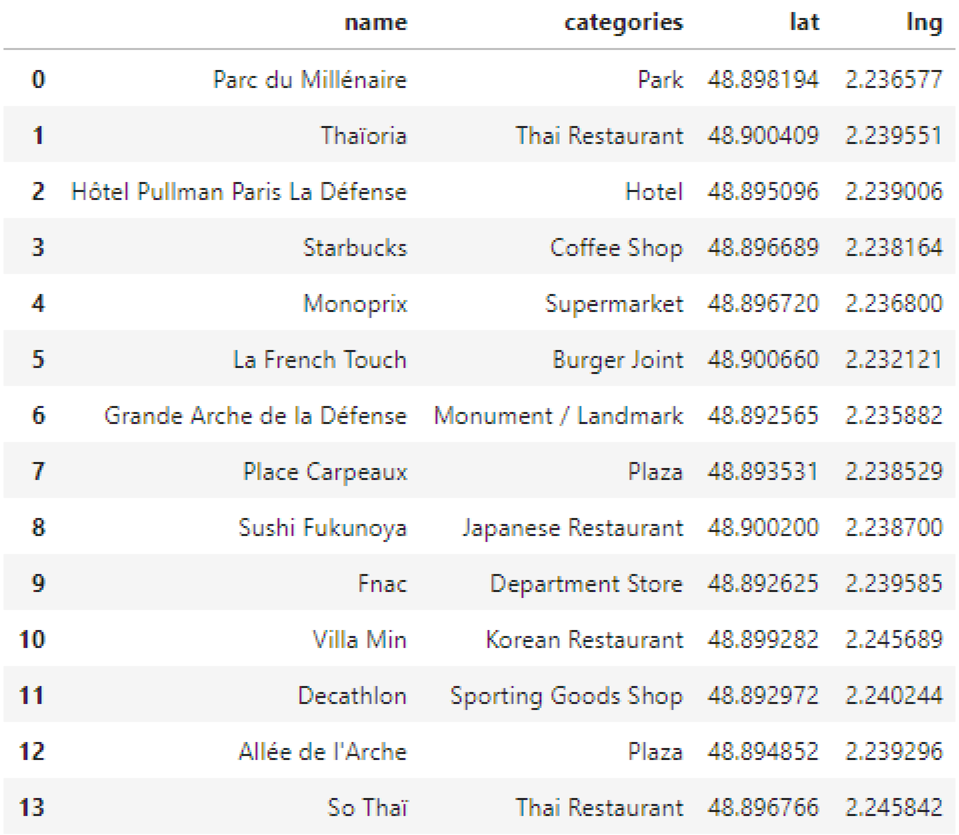
Hours or services may differ due to COVID-19

Residhome Courbevoie la Défense

49 Avenue de l'Arche, 92400 Cour...

48.898187, 2.238144

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	Hótel 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