Applied Data Science Capstone

1) Introduction/Business Problem

This project is to analyze the neighborhoods in Toronto in order to find/recommend a convenient place to start a restaurant respecting several criteria. It can eventually provide some analysis and comparisons between different towns. It provides every time appropriate tables that represents operations such as merging filtering, selecting data under some conditions, etc. This project will use different tools to obtain data in a very specific geographical area and other files and websites in order to filter and enrich the required data and provide a reliable data quality.

The used criteria about the recommended town are basically related to education average of people, the importance of their incomes as well as average age.

The audience/who would be interested in this project: Any person who desires to *start* a restaurant in Toronto or a person who would *make an investment* in a restaurant but would obtain reliable information about the location choice in the city in order to reduce the risks in such investment.

2) Data

Many data sources will be used in this project, but the most important will be related to the City of Toronto Neighborhood Profiles, and City of Toronto Open Data Catalogue. In addition, the following links, files, or other sources will be also very useful:

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

    Description: this page can be used to scrape and wrangle the data as well as to clean the data and put them into pandas data frames.

* https://cocl.us/Geospatial_data

    Description: This link will be used to get the Geographical coordinates of the neighbourhoods taking into consideration Postal Codes.

* https://foursquare.com/developers/apps

    Description: this is used to a location data platform. It will provide the pertinent data of a specific geographic location that will be used in the analysis and filtering the required queries.

- https://www.toronto.ca/ext/open_data/catalog/data_set_files/2016_neighbourhood_profiles.csv
  Description: it is basically used to get the census information helping to find the best results under the required criteria.
- https://ckan0.cf.opendata.inter.sandbox-toronto.ca/download_resource/1d02b0f0-d735-4469-8f71-ea6d96b319e4?format=geojson&projection=4326
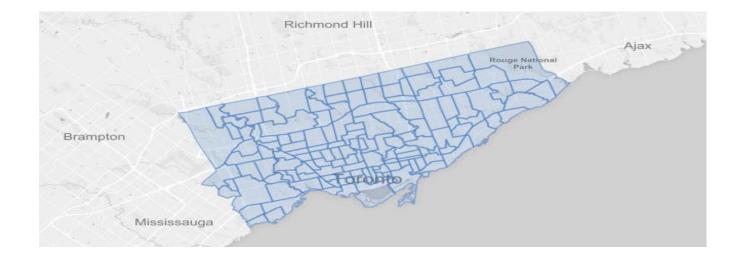 **Description:** this is used as a mapping data

The use of the Data

Similar to what was performed in the previous assignments, these data will be used to go further steps in exploring, segmenting, and clustering the neighborhoods in Toronto.

Most importantly, the data will be largely used to produce the needed statistics, clustering (k-means, etc.) and other methods to refine the final results.

The results that this project provides will be similar to the following:

	PostalCode	Borough	Neighborhood
0	M5G	Downtown Toronto	Central Bay Street
1	M2H	North York	Hillcrest Village
2	M4B	East York	Parkview Hill, Woodbine Gardens
3	M1J	Scarborough	Scarborough Village
4	M4G	East York	Leaside
5	M4M	East Toronto	Studio District
6	M1R	Scarborough	Wexford, Maryvale
7	M9V	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...
8	M9L	North York	Humber Summit
9	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...
10	M1B	Scarborough	Malvern, Rouge
11	M5A	Downtown Toronto	Regent Park, Harbourfront

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
1	M2H	North York	Hillcrest Village	43.803762	-79.363452
2	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
3	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
4	M4G	East York	Leaside	43.709060	-79.363452
5	M4M	East Toronto	Studio District	43.659526	-79.340923
6	M1R	Scarborough	Wexford, Maryvale	43.750071	-79.295849
7	M9V	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...	43.739416	-79.588437
8	M9L	North York	Humber Summit	43.756303	-79.565963
9	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.645711	-79.392732
10	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
11	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636

	neighborhood	neighborhood_latitude	neighborhood_longitude	venue	venue_latitude	venue_longitude	venue_category
0	Agincourt North (129)	43.805441	-79.266712	Saravanaa Bhavan South Indian Restaurant	43.810117	-79.269275	Indian Restaurant
1	Agincourt North (129)	43.805441	-79.266712	Fahmee Bakery & Jamaican Foods	43.810170	-79.280113	Caribbean Restaurant
2	Agincourt North (129)	43.805441	-79.266712	Samosa King - Embassy Restaurant	43.810152	-79.257316	Indian Restaurant
3	Agincourt North (129)	43.805441	-79.266712	Grandeur Palace 華丽宮 (Grandeur Palace 華麗宮)	43.797885	-79.270585	Chinese Restaurant
4	Agincourt North (129)	43.805441	-79.266712	Bestco Food Market 鴻華超級市場	43.796514	-79.270790	Supermarket