CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS (WEEK 2)

Best Neighborhoods to settle in, according to some criteria

Table of contents

A - Introduction: Business Problem

◦ It appears generally difficult for people to choose the neighborhood where they want to settle in. Mostly when they are not natives of the city in which they intend to stay.

◦ Indeed, Mr. Chris has just got his Canadian's PR card. He is coming from Cameroon and wants to settle in Toronto with his wife and their 3 years old child. Mr. Chris needs us to make some recommendations in order to help him choose the best neighborhood.

◦ Mr. Chris is a junior Data Scientist looking to graduate from one of the universities in Toronto. He would like mostly to find a neighborhood not far of garden school for his child.

◦ Nowadays, there are some great tools that we can use to effectively overcome such a problem. In our case, we are going to use a machine learning tool known as "Segmentation and Clustering", combined with the "Foursquare location data".

B - Data Section: Data understanding

Data Sources

◦ We are going to use **Toronto dataset** scraped from the web. Here is the link: ('https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'). This table is consisted of 3 columns: postcodes, boroughs and neighborhoods.

◦ We will also get **Toronto's Postcode location data** by downloading from the web. Here is the link: (http://download.geonames.org/export/zip/). This geospatial table is consisted of 4 columns: postcodes, boroughs, latitudes and longitudes.

B - Data Section: Data understanding

Data Collection

◦ By using the **BeautifulSoup package for web scraping**, we will pull out information of Postal code, boroughs and neighborhoods, from the table scraped on the web.

◦ By using the **Foursquare API**, we will obtain location data and explore venues around.

◦ The borough and neighborhood table joined with location data, present the latitude and the longitude of each neighborhood.

◦ The exploration of venues presents venues near of the neighborhood of Toronto. We will use it to filter gardens, bus station, and medical centers as well. Now we are ready to prepare the data for the modelling section. Before that let's look at the methodology that we are going to use.

C - Methodology Section

Data wrangling

Exploration of neighborhoods in Toronto

```
(210, 3)
```

	Postcode	Borough	Neighbourhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront
5	M6A	North York	Lawrence Heights
6	M6A	North York	Lawrence Manor

C - Methodology Section

Data wrangling

Geopatial data of neighborhoods in Toronto

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.7545	-79.3300
1	M4A	North York	Victoria Village	43.7276	-79.3148
2	M5A	Downtown Toronto	. Harbourfront	43.6555	-79.3626
3	M6A	North York	Lawrence Heights	43.7223	-79.4504
4	M6A	North York	Lawrence Manor	43.7223	-79.4504

C - Methodology Section

Data wrangling

Explore Neighborhoods in Toronto

```
(4573, 7)
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.7545	-79.3300	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.7545	-79.3300	GreenWin pool	43.756232	-79.333842	Pool
2	Parkwoods	43.7545	-79.3300	Variety Store	43.751974	-79.333114	Food & Drink Shop
3	Victoria Village	43.7276	-79.3148	Victoria Village Arena	43.723481	-79.315635	Hockey Arena
4	Victoria Village	43.7276	-79.3148	Tim Hortons	43.725517	-79.313103	Coffee Shop

C - Methodology Section

Data wrangling

Explore Neighborhoods in Toronto: filter bus station, medical center, pharmacy, garden, metro, and order

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
0	Agincourt North	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
1	Albion Gardens	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
2	Alderwood	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
3	Beaumond Heights	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
4	Bedford Park	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
5	Bloordale Gardens	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
6	Cabbagetown	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
7	Church and Wellesley	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
8	Clairlea	Bus Line	Metro Station	Bus Station	Pharmacy	Medical Center	Garden Center	Garden

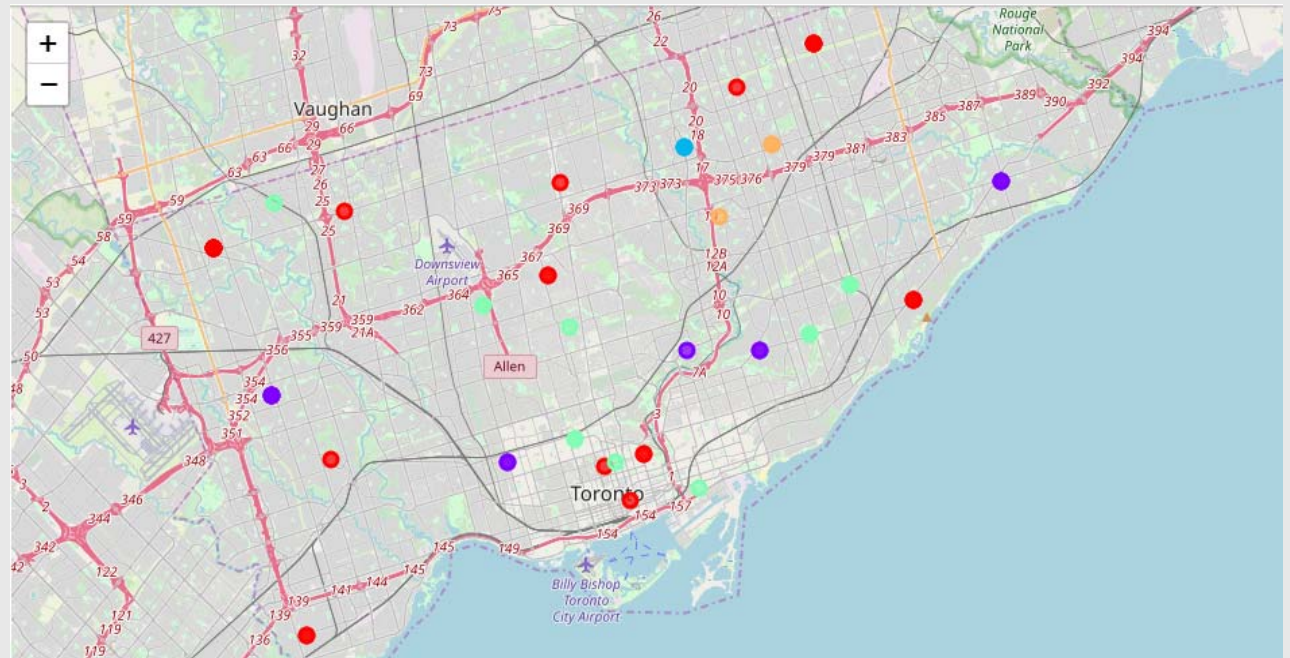# C - Methodology Section

Clustering

Set number of cluster to 5

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
3	M6A	North York	Lawrence Heights	43.7223	-79.4504	4	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
4	M6A	North York	Lawrence Manor	43.7223	-79.4504	4	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
5	M7A	Downtown Toronto	Queen's Park	43.6641	-79.3889	0	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
6	M9A	Queen's Park	Queen's Park	43.6662	-79.5282	0	Pharmacy	Metro Station	Medical Center	Garden Center	Garden	Bus Station	Bus Line
10	M4B	East York	Woodbine Gardens	43.7063	-79.3094	2	Pharmacy	Bus Line	Metro Station	Medical Center	Garden Center	Garden	Bus Station

D - Results section

- **Cluster N°0 (red):** Pharmacy, **Metro Station**, Medical Center, Garden Center

- **Cluster N°1 (violet):** Pharmacy / **Garden Center**, Metro Station, Medical Center

- **Cluster N°2 (blue):** **Bus line** / Pharmacy, Medical Center / Metro Station

- **Cluster N°3 (green):** Pharmacy, **Metro Station**, Medical Center, Garden Center

- **Cluster N°4 (orange):** Pharmacy, **Metro Station**, Medical Center, Garden Center

E - Discussion section

Assuming that he would like mostly to be nearest of a garden center in order to easily lead his child to school, we recommend all the neighborhoods in:

Cluster N°1: East and Central Toronto, around Yorkville and Studio District

In another hand, if he would like to be nearest of a Bus line in order to make easiest his movement:

Cluster N°2: Mostly in Scarborough

For cases more generals, he can choose:

Cluster N°3 or 4: Others neighborhoods

F - Conclusion section

According to all above, we can recommend to M. Chris to settle down in cluster N° 1, around **Yorkville** and **Studio District**. Due to the proximity in first with a garden center, then with metro station.