Capstone: The Battle of Neighborhoods

Data Description

1. Requirements

As we need to explore, segment, and analyze the neighborhoods in the city of Toronto, the Toronto neighborhoods data is key for this project. We need the names of neighborhood, their populations, hospital admission scenario. Here, we also need the exact geolocations of each neighborhoods.

We need the data how hospitals in scattered all over Toronto. For this, their geolocation is must. We need the data how pharmacy stores are distributed all over the city. The data should contain the coordinates for each of the Pharmacies in Toronto that will help us to further obtaining more information critical for this project. We will also like to obtain the key information; such as average income for each neighborhood which might be one of the key factors for the neighborhood of choice in the final recommendation report.

2. Collections

From '*http://www.torontohealthprofiles.ca/*', we obtain hospital admission rate in different neighborhoods in Toronto, Canada. It includes: neighborhoods name, their population, admission scenario. We use geopy for integrating these observations with geolocation. It doesn't have average income of person in each neighborhood. We scrape and clean the income data of each neighborhood from '*https://www.cmhc-schl.gc.ca/en/professionals/housing-markets-data-and-research*' and left join with above neighborhood data.

We use Foursquare API for getting Venue's data using the geolocation of Neighborhoods. We Filtered this data for getting the Pharmacies Data in Toronto, Ontario

3. Datasets

a. Neighborhoods Data

It has the all the neighborhoods name, their respective populations, hospital admission scenario. Along with these, it also comprises geolocation, and income.

	Neighbourhood	Population	Male Admission	Female Admission	Both Admission	Latitudes	Longitudes	Income
0	West Humber-Clairville	34080	2453	3151	5604	43.723370	-79.597457	71699.0
1	Rexdale-Kipling	10455	938	1090	2028	43.722114	-79.572292	64800.0
2	Elms-Old Rexdale	9970	674	978	1652	43.721770	-79.552173	63156.0
3	Humber Heights-Westmount	10600	950	1335	2285	43.695785	-79.520832	76669.0
4	Edenbridge-Humber Valley	14895	1029	1269	2298	43.670672	-79.518855	144418.0
5	Markland Wood	10395	804	944	1748	43.631239	-79.585434	NaN
6	Etobicoke West Mall	11005	821	1058	1879	43.643549	-79.565325	80527.0
7	Islington-City Centre West	38130	2829	3966	6795	43.662577	-79.531789	NaN

b. Toronto Venues Data

It has Venues names with their respective geolocation, its category, and the neighborhood where it lies.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	West Humber-Clairville	43.72337	-79.597457	Fortinos	43.721438	-79.596291	Grocery Store
1	West Humber-Clairville	43.72337	-79.597457	Mandarin Buffet	43.720360	-79.594387	Chinese Restaurant
2	West Humber-Clairville	43.72337	-79.597457	Humber Arboretum	43.724653	-79.608006	Garden
3	West Humber-Clairville	43.72337	-79.597457	Bulk Barn	43.720290	-79.600887	Food & Drink Shop
4	West Humber-Clairville	43.72337	-79.597457	New York Fries - Woodbine Centre	43.720568	-79.599609	Restaurant
5	West Humber-Clairville	43.72337	-79.597457	TD Canada Trust	43.719630	-79.599896	Bank
6	West Humber-Clairville	43.72337	-79.597457	Winners	43.719819	-79.594923	Department Store
7	West Humber-Clairville	43.72337	-79.597457	Woodbine Racetrack	43.716673	-79.604962	Racecourse

c. Toronto Pharmacy Data

It stores the Pharmacies, their respective names, and the neighborhood where it lies.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Humber Heights-Westmount	43.695785	-79.520832	Shoppers Drug Mart	43.701264	-79.519705	Pharmacy
1	Edenbridge-Humber Valley	43.670672	-79.518855	Shoppers Drug Mart	43.661721	-79.518881	Pharmacy
2	Etobicoke West Mall	43.643549	-79.565325	Shoppers Drug Mart	43.641312	-79.576924	Pharmacy
3	Islington-City Centre West	43.662577	-79.531789	Shoppers Drug Mart	43.663067	-79.531753	Pharmacy
4	Islington-City Centre West	43.662577	-79.531789	Thorncrest Drug Store	43.662988	-79.531817	Pharmacy
5	Kingsway South	43.647381	-79.511333	Shoppers Drug Mart	43.647623	-79.510860	Pharmacy
6	New Toronto	43.600763	-79.505264	Shoppers Drug Mart	43.601677	-79.502239	Pharmacy
7	Long Branch	43.592005	-79.545365	Rexall	43.593111	-79.538327	Pharmacy
8	Alderwood	43.601717	-79.545232	Sav-On Drug Mart	43.609891	-79.549618	Pharmacy
9	Glenfield-Jane Heights	43.756799	-79.518050	Shoppers Drug Mart	43.756147	-79.515843	Pharmacy

d. Neighborhoods without Pharmacies Data

It stores the information of neighborhoods without any retail Pharmacies.

	Neighbourhood	Population	Male Admission	Female Admission	Both Admission	Latitudes	Longitudes	Income
0	West Humber-Clairville	34080	2453	3151	5604	43.723370	-79.597457	71699.0
1	Rexdale-Kipling	10455	938	1090	2028	43.722114	-79.572292	64800.0
2	Elms-Old Rexdale	9970	674	978	1652	43.721770	-79.552173	63156.0
5	Markland Wood	10395	804	944	1748	43.631239	-79.585434	NaN
9	Mimico	26635	1758	2080	3838	43.616677	-79.496805	71772.0

It stores the information of neighborhoods with retail Pharmacies.

	Neighbourhood	Population	Male Admission	Female Admission	Both Admission	Latitudes	Longitudes	Income
3	Humber Heights-Westmount	10600	950	1335	2285	43.695785	-79.520832	76669.0
4	Edenbridge-Humber Valley	14895	1029	1269	2298	43.670672	-79.518855	144418.0
6	Etobicoke West Mall	11005	821	1058	1879	43.643549	-79.565325	80527.0
7	Islington-City Centre West	38130	2829	3966	6795	43.662577	-79.531789	NaN
8	Kingsway South	9170	607	745	1352	43.647381	-79.511333	198998.0

4. Conclusion

We will analyze these data frames visually and try to identify the location for setting up the Pharmacy Store for Orange Drugs.

With all these features, techniques and data, we will then be able to come up with a best recommendation to the management of Orange Drugs to their problem which is where is the best neighborhood for them to first start off to offer their services. For an example, we will not want to enter a neighborhood whereby there is already a high concentration of Pharmacy available or there is a high trending of such stores upcoming in the neighborhood. We will like to recommend a neighborhood where we know that there will be a higher demand of such delivery service due to the lack of supply in that area.