The Battle of Neighborhoods

ABSTRACT

An analysis of the city of Toronto, CA with a comparison to the city of Paris, FR.

Ljiljana Nikolajev

IBM Capstone Project

1 Introduction

1.1 Background

For this project I am a real estate agent recruited by a family of four, about to relocate from Paris, France to Toronto, Canada. The client is looking to relocate in a neighborhood that is more secure than their current city. Ideally, the client wishes to live in a family friendly average neighborhood, where schools (elementary and high schools) are at walking distance from their homes. As they are big foodies and love discovering new types of food, they would like to have great diversity in restaurant options close to their home. Both parents will be working at the Michael Garron hospital and wished for their commute to be no longer than 20 minutes by car.

1.2 Problematic

Moving from one country to another is not an easy task, particularly when one does not have a point of reference. This project will help to determine the best borough in Toronto for this family to relocate and ensure that their transition into this new life is as smooth as possible

2. Data Acquisition and Cleansing

2.1 Data Source

- To assess the level of criminality in Toronto, we will scrap and analyze data from the open data portal from the city of Toronto.
- To compare these data, criminality in Paris was extracted from this article, sourced from the Ministre de L'interieur
- To help define the boroughs and neighborhood of Toronto, the open data portal for the city of Toronto was accessed again.

2.2 Data Cleaning

Toronto crime dataset first had to be cleaned before it could be used. First, this data set contained all major crime incidents in the city of Toronto from 2014-2019. Since we are only interested in comparing the data from 2019, all previous years had to be dropped.

	X	Y	Index_	event_unique_id	occurrencedate	reporteddate	premisetype	ucr_code	ucr_ext	offence	...	occurrencedayofyear	occurrencedayofweek	occurrencehour	MCI	Division	Hood_ID	Neighbourhood	Long	Lat	ObjectId
0	-8.816401e+06	5.434587e+06	701	GO-20141756319	2014/03/24 00:00:00+00	2014/03/24 00:00:00+00	Commercial	1430	100	Assault	...	83.0	Monday	1	Assault	D42	132	Malvern (132)	-79.199081	43.800281	1
1	-8.837252e+06	5.413357e+06	901	GO-20143006885	2014/09/27 00:00:00+00	2014/09/29 00:00:00+00	Other	2120	200	B&E	...	270.0	Saturday	16	Break and Enter	D52	76	Bay Street Corridor (76)	-79.386383	43.662472	2
2	-8.862433e+06	5.422276e+06	702	GO-20141756802	2014/03/24 00:00:00+00	2014/03/24 00:00:00+00	Commercial	2120	200	B&E	...	83.0	Monday	6	Break and Enter	D23	1	West Humber-Clairville (1)	-79.612595	43.720406	3
3	-8.833104e+06	5.431887e+06	703	GO-20141760570	2014/03/24 00:00:00+00	2014/03/24 00:00:00+00	Apartment	2120	200	B&E	...	83.0	Monday	15	Break and Enter	D33	47	Don Valley Village (47)	-79.349121	43.782772	4
4	-8.845311e+06	5.413667e+06	902	GO-20142004859	2014/05/03 00:00:00+00	2014/05/03 00:00:00+00	Commercial	1610	210	Robbery - Business	...	123.0	Saturday	2	Robbery	D11	90	Junction Area (90)	-79.458778	43.664490	5

5 rows × 29 columns

First, all non-relevant columns were dropped. Second, because there were null values in the dataset, all row that contained a non-value for the "occurenceyear" column were dropped. Once this was done, it was possible to isolate all incidents from year 2019 (*crimes3_df*).

	MCI	occurrenceyear	Lat	Long	Neighbourhood
0	Assault	2019.0	43.810932	-79.227135	Malvern (132)
1	Assault	2019.0	43.663906	-79.384155	Church-Yonge Corridor (75)
2	Assault	2019.0	43.655777	-79.380676	Church-Yonge Corridor (75)
3	Assault	2019.0	43.723015	-79.415932	Bedford Park-Nortown (39)
4	Break and Enter	2019.0	43.648773	-79.528748	Islington-City Centre West (14)

In order to pass the neighborhood into a choropleth map, a new dataframe was created with all lat, lonf and total crimes value for each neighborhood (*crimes5_df*)

Out[4]:

	Neighbourhood	Count
0	Agincourt North (129)	210
1	Agincourt South-Malvern West (128)	309
2	Alderwood (20)	85
3	Annex (95)	598
4	Banbury-Don Mills (42)	214

Evolution des crimes et délits à Paris Le Parisien

Comparaison janvier à septembre 2018 et 2019

Catégories	Janv. à sept. 2019	Evolution 2018/2019
• Vols - dont cambriolages	144 641 13 743	+ 12,4 % + 7,9 %
• Délits économiques et financiers (escroqueries)	29 909	+ 4,6 %
• Violences	17 784	+ 7,1 %
• Destructions et dégradations	12 757	+ 15,4 %
• Stups	9 157	- 10,6 %
• Outrage et violence sur personnes dépositaires de l'autorité	4 307	+ 20,4 %
• Violences sexuelles	2 690	+ 4,9 %
• Port ou détention d'armes prohibées	2 662	+ 15 %
• Crimes et délits contre mineurs - dont viols	1 518 837	+ 7,4 % + 11,5 %
• Autres	8 786	+ 11,2 %
TOTAL	234 211	+ 10,5 %

SOURCE : MINISTÈRE DE L'INTÉRIEUR. LP/INFOGRAPHIE.

Unfortunately, there are no open source for crime data for the Paris metro area. Data were extracted from the aforementioned article, for which the data were provided by the French government. To standardize the categories, all assaults were grouped together. The same was for robberies and thefts. Data are from Jan - Sept inclusively. Since the data were in an image and not in a table, the values were added to a panda dictionnary that was further transfored in a dataframe (*paris_df*).

3. Methodology

Techniques that will be used through this project are as such:

Data Analysis will help us narrow down the sagest neighborhood

1) we will create a new dataframe that will contain all crimes by types.
    a. Pie charts will then be used to compare Paris and Toronto criminality (example shown)
2) We will use the panda library again to isolate the 10 more dangerous and the 10 safest neighborhoods of Toronto. Graph bars will then be use to illustrate the difference.

3) Using folium and choropleth, all crimes will be transposed onto a map of Toronto to provide an additional visual.
4) All neighborhoods within 10 km distance from the hospital will be selected for further analysis

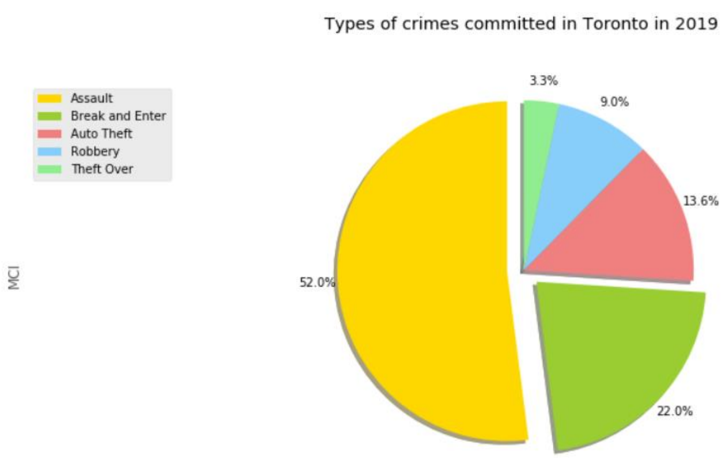
Foursquare API will help us find the venues in each of these neighborhoods and help us narrow down the list even more

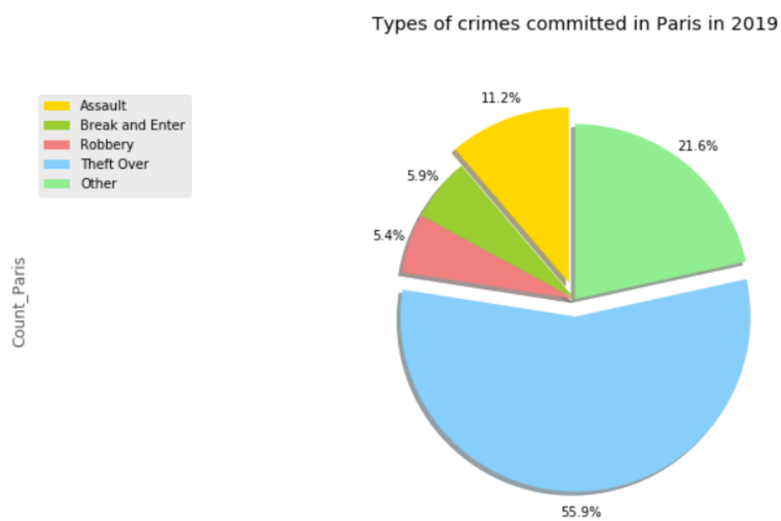
5) Foursquare API will be used to first, determine if each neighborhood contained both a elementary and high schools that are close to each other.
6) Foursquare API will be used to retrieve all restaurants venues in the selected neighborhoods.
7) For the selected neighborhoods, the one with the overall best profile, as determined by school presence and by greatest variety of restaurants, will be selected.

4. Results

4.1 Data Analysis

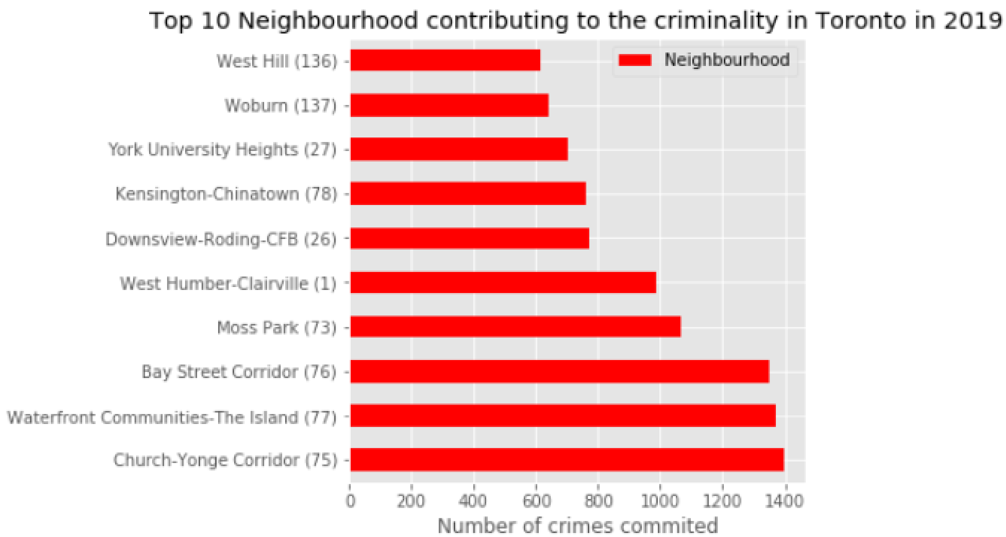
- A dataframe was created to determine the type of crimes committed in Toronto (*type_df*). Pie charts were generated to compare the types of crimes incidence in Paris and Toronto

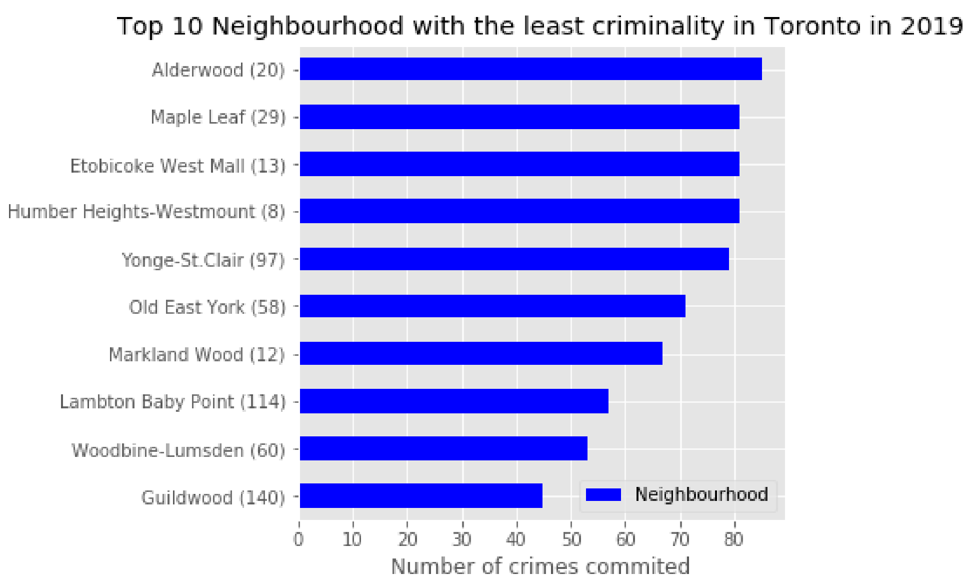




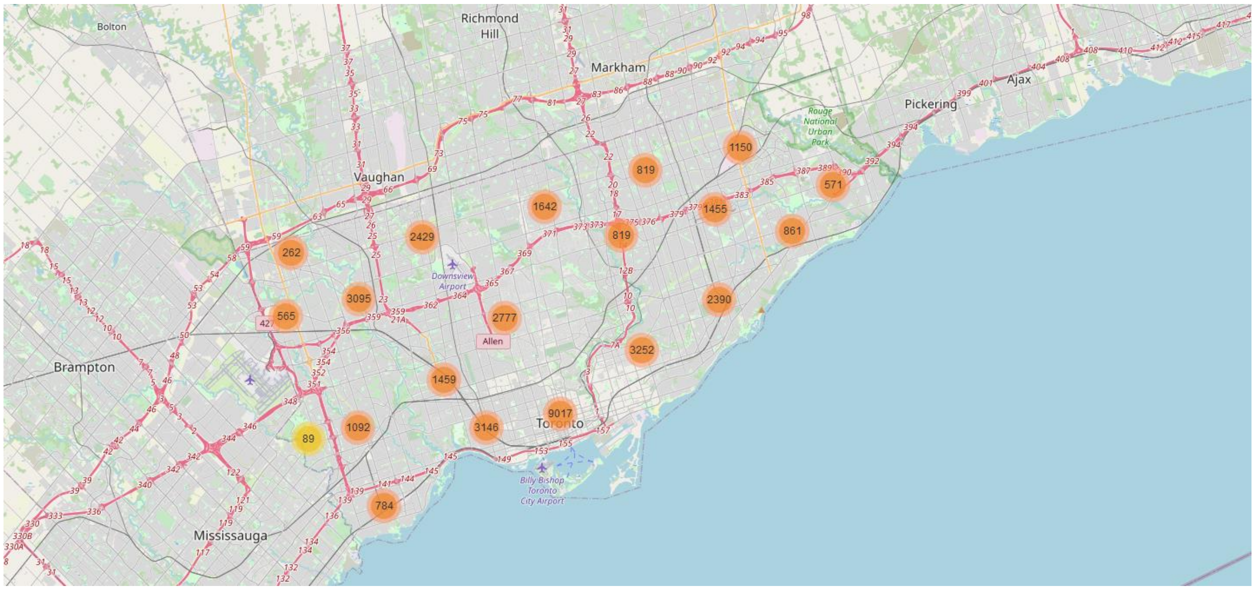
From these data, we see that 37,674 crimes were committed in Toronto in 2019, while 234,211 crimes were committed in Paris during the months on January to September 2019. While Toronto might appear safer than Paris, population density is not the same.

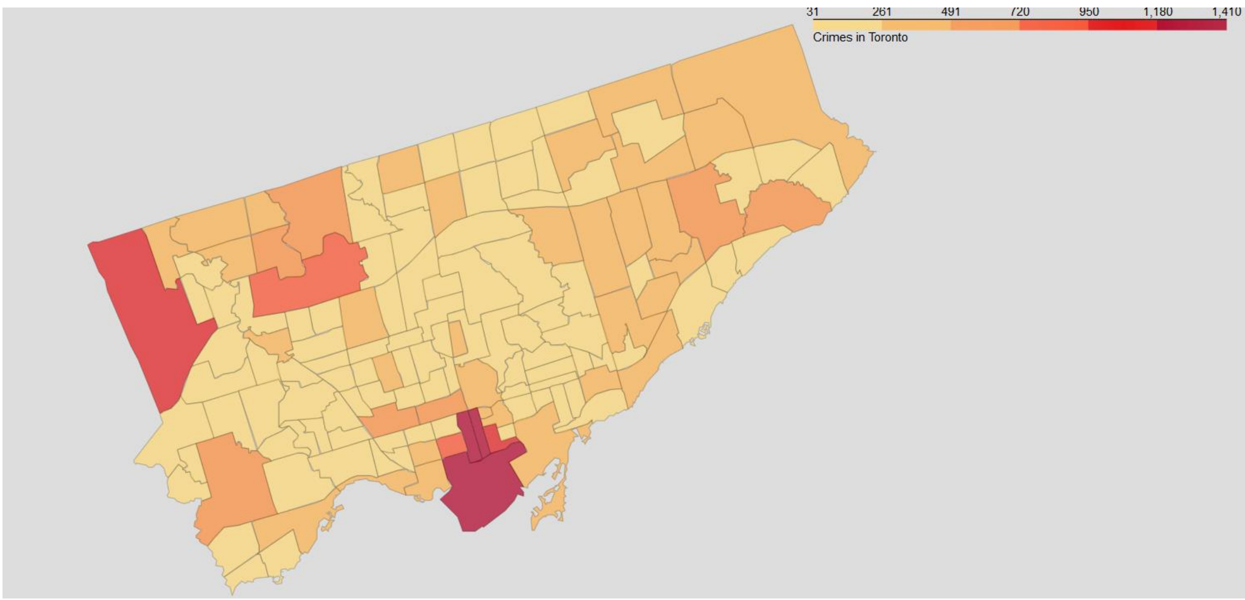
With 2.7M habitants, Toronto (+GTA) has a crime rate of 13.9 crimes per million of habitant. Paris (+ Metro area) has a population of 10.2M habitant and a crime rate of 23.0 crimes per million habitants. While Toronto is slightly safer than Paris on general, the worst and the safest neighborhoods in Toronto were isolated as shown in the bar graphs below:



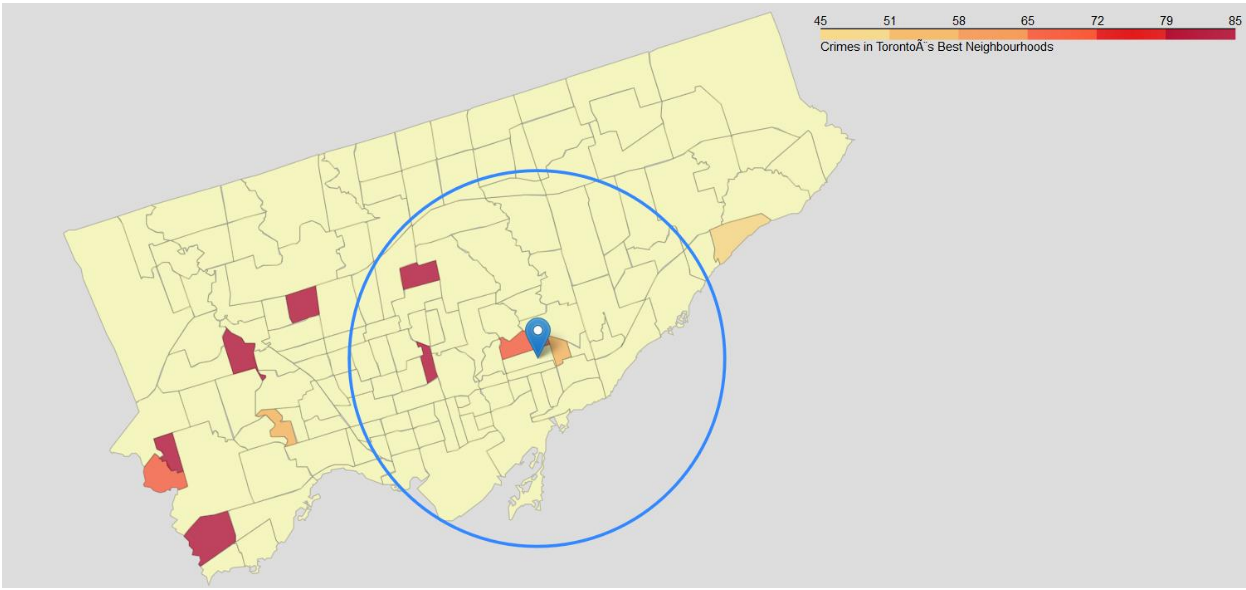


For visual effect, all crimes were represented on a folium map and using a choropleth map. This will also help us visualize in which part of Toronto criminal activity takes place.





Using these data, we decide to focus on the 10 safest neighborhoods to continue our analysis. Since our client wishes to live within a 10km distance from their future work location, we use folium map, choropleth and folium.Circle to narrow down all adequate neighborhoods. As seen in the map below, we end up with four neighborhood that fits this criterion, and for which the criminality is minimal: **Yonge-St-Clair, Old East York, Woodbine-Lumsden, and Lawrence Parc North**. All four neighborhoods will be considered for further evaluation.



4.2 Foursquare API

For each of the four selected neighborhoods and using a function returning all venues for each, we determine the amount of venue per sector. Using the groupby function, we can see in the table below, that with a radius of 3km, Lawrence Park North only has 10 venues. This is not enough to satisfy our client, who wishes to have a diversified selection of venues in their future residential area. This neighborhood is therefore excluded from the analysis.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Lawrence Park North (105)	10	10	10	10	10	10
Old East York (58)	30	30	30	30	30	30
Woodbine-Lumsden (60)	30	30	30	30	30	30
Yonge-St.Clair (97)	30	30	30	30	30	30

For the three-remaining neighborhood, we call the Foursquare API again to return the school venues:

1. **Yonge- St-Clair** has:
    a. 5 elementary and high school venues
    b. 27 restaurants within 2km radius, with limited variety, with ½ of the choice being Asian restaurants (see index)
2. **Old East York** has:
    a. 4 elementary and high school venues
    b. 29 restaurants within 2km radius with great food type variety (see index)
3. **Woodbine-Lumsden** has
    a. 4 elementary and high school venues
    b. 28 restaurants within 2km radius with great food type variety

5. Discussion

In this project, we explore the various neighborhoods of Toronto, in an effort to find the best place for an immigrant family to leave. Initial criteria include

1. 10 km proximity to workplace (Michael Garron Hospital)
2. Lowest criminality possible
3. Proximity to elementary and high school
4. Diverse culinary scene

Starting with a dataset of crimes committed in Toronto, we quickly notice that the city was generally safer than the city of Paris, from which our clients are coming from. Using only the data from 2019, we realized that the crime rate per habitant in Toronto was of 13.9 crimes per million of habitants, while it was of 23.0 crimes per million habitants for Paris. Since the family want to leave in the safest environment, only the 10 safest neighborhoods were selected for further consideration.

It is important to denote that the dataset from Paris was not as thorough as the one from Toronto and that the months of October to December were not included. We do not think this impact our reasoning, since Toronto is anyway already safer than Paris.

We then mapped our 10 neighborhoods on a choropleth map to visualize their relative distance from the Michael Garrison Hospital. Only the neighborhood that were within a 10 km radius were selected.

Following this, data were extracted from the foursquare API to retrieve school and restaurant venues. Only three neighborhoods were good enough to pass the round.

While Foursquare API is a powerful tool, we recognize that it is hard to pinpoint a particular location from a neighborhood, without a precise address. It is therefore possible that the obtained results vary slightly although non significantly, once a house in a particular neighborhood will be selected by the client. To take that into account, we selected a radium relatively large in regards to the surface areas of the neighborhoods. As we can see with the restaurant venues, from Old east York and Woodbine-Lumsden, there are some overlap in the restaurants' selection.

6. Conclusion

In conclusion, it is our belief that all 3 remaining neighborhoods, that is Yonge-St-Clair, Old east York and Woodbine-Lumsden are good fit for the family. Given that **Old east York and Woodbine-Lumsden** are geographically close and have better (and more) restaurant venues (although here is some overlap between the list), we think the family should focus their research in this region of Toronto.

7. Annex

	categories
Chinese Restaurant	4
Indian Restaurant	3
Restaurant	3
Japanese Restaurant	2
Italian Restaurant	2
Cantonese Restaurant	1
Eastern European Restaurant	1
Thai Restaurant	1
Sushi Restaurant	1
Diner	1
Music Venue	1
French Restaurant	1
General Entertainment	1
New American Restaurant	1
Steakhouse	1
Bistro	1
Event Space	1
Café	1

Figure 1: Yonge-St-Clair Restaurant List

	categories
Chinese Restaurant	5
Ethiopian Restaurant	4
Middle Eastern Restaurant	2
Indian Restaurant	2
Restaurant	2
Burger Joint	2
Caribbean Restaurant	2
Hungarian Restaurant	1
Breakfast Spot	1
Asian Restaurant	1
African Restaurant	1
Thai Restaurant	1
American Restaurant	1
Turkish Restaurant	1
Miscellaneous Shop	1
Mexican Restaurant	1

Figure 2: Woodbine-Lumdsen Restaurants

	categories
Chinese Restaurant	5
Ethiopian Restaurant	4
Caribbean Restaurant	3
Greek Restaurant	3
Restaurant	3
Thai Restaurant	2
Middle Eastern Restaurant	2
American Restaurant	1
African Restaurant	1
Indian Restaurant	1
Breakfast Spot	1
Miscellaneous Shop	1
Turkish Restaurant	1
Lounge	1

Figure 3: Old East York Restaurants