2. Data

Opening a new business such as a restaurant requires an analysis of several factors. If done meticulously, number of factors could be well above a hundred. However, for the purpose of this report we will focus on the following – potential customers, competition and safety in several neighborhoods in Toronto. Description of factors and data sources is below.

2.1. Data sources

There are several data sources in the project:

1. Foursquare location data from a developer's account
https://foursquare.com/
2. Safety data extracted from "Toronto Police Service Public Safety Data Portal"
http://data.torontopolice.on.ca/pages/open-data
3. Population data across neighborhoods was extracted from the 2016 Census, Toronto
https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm
4. Demographic data from the Toronto Open Data Portal
https://www.toronto.ca/city-government/data-research-maps/open-data/
5. Neighborhood data and postal codes for Toronto
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
6. Langitude and latitude coordinates of neighborhoods
CSV file "Geospatial coordinates.csv"

2.2. Data description

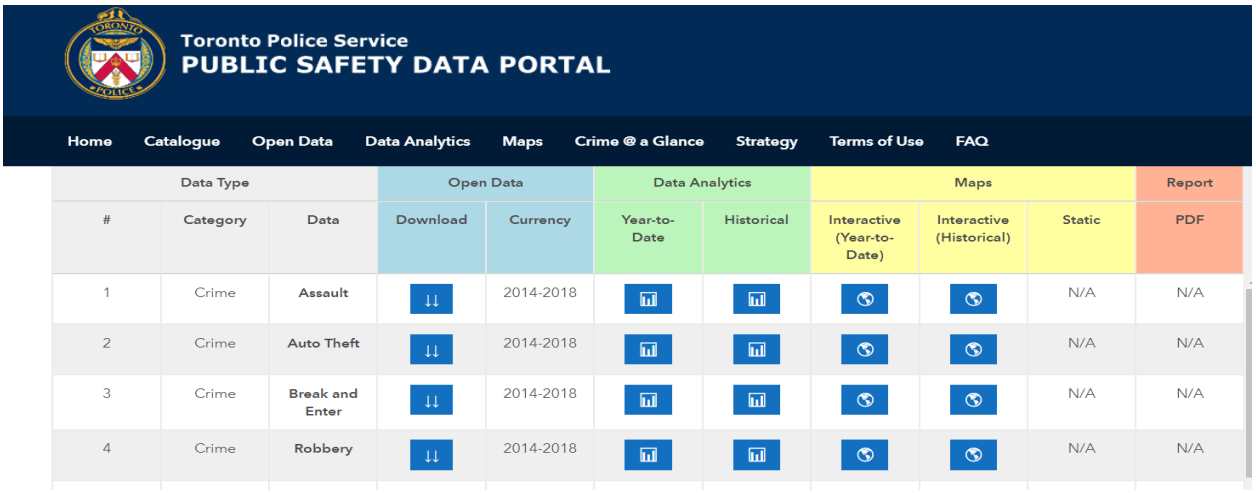
Data collected from the mentioned sources above will allow us to analyze neighborhoods and figure out which are more suitable to opening new restaurants. Using foursquare location data, we will primarily analyze competition, for example a number of restaurants operating in the neighborhood (if data allows, maybe even the types of restaurants).

[34]:

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	neighborhood	postalCode	state
0	Korin	Furniture / Home Store	57 Warren St	US	New York	United States	Church St	73	[57 Warren St (Church St), New York, NY 10007, United States]	[{'label': 'display', 'lat': 40.71482437714839, 'lng': -74.00940425461492}]	40.714824	-74.009404	Tribeca	10007	NY
1	Juice Press	Vegetarian / Vegan Restaurant	83 Murray St	US	New York	United States	btwn Greenwich St & W Broadway	202	[83 Murray St (btwn Greenwich St & W Broadway), New York, NY 10007, United States]	[{'label': 'display', 'lat': 40.71478769908051, 'lng': -74.0111317502157}]	40.714788	-74.011132	NaN	10007	NY
2	Chambers Street Wines	Wine Shop	148 Chambers St	US	New York	United States	btwn West Broadway & Hudson St	88	[148 Chambers St (btwn West Broadway & Hudson St), New York, NY 10007, United States]	[{'label': 'display', 'lat': 40.715773063928374, 'lng': -74.00971823312332}]	40.715773	-74.009718	NaN	10007	NY
3	Takahachi Bakery	Bakery	25 Murray St	US	New York	United States	at Church St	187	[25 Murray St (at Church St), New York, NY 10007, United States]	[{'label': 'display', 'lat': 40.713652845301894, 'lng': -74.0088038953017}]	40.713653	-74.008804	NaN	10007	NY
4	Takahachi	Sushi Restaurant	145 Duane St	US	New York	United States	btwn W Broadway & Church St	146	[145 Duane St (btwn W Broadway & Church St), New York, NY 10013,	[{'label': 'display', 'lat': 40.71652647412374, 'lng': -74.00810108466207}]	40.716526	-74.008101	NaN	10013	NY

Image 1. An example of venue information obtained by a request to a Foursquare API

The safety data from Toronto Police will give us data on the crime rates of certain areas as we don't want to open a restaurant in an area with high levels of crime. It is available in CSV format and will be transformed into Pandas data frames for cleaning.



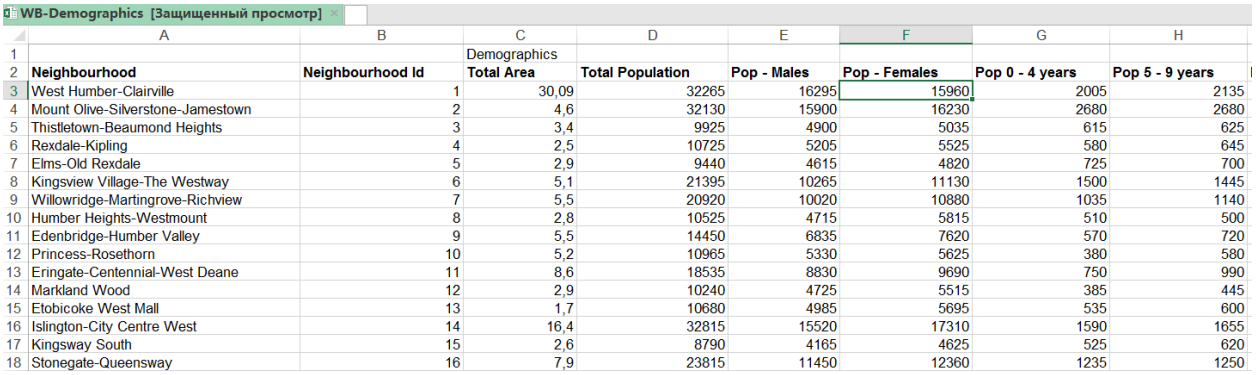
Toronto Police Service
PUBLIC SAFETY DATA PORTAL

Home | Catalogue | Open Data | Data Analytics | Maps | Crime @ a Glance | Strategy | Terms of Use | FAQ

Data Type			Open Data		Data Analytics		Maps			Report
#	Category	Data	Download	Currency	Year-to-Date	Historical	Interactive (Year-to-Date)	Interactive (Historical)	Static	PDF
1	Crime	Assault	↓↓	2014-2018					N/A	N/A
2	Crime	Auto Theft	↓↓	2014-2018					N/A	N/A
3	Crime	Break and Enter	↓↓	2014-2018					N/A	N/A
4	Crime	Robbery	↓↓	2014-2018					N/A	N/A

Image 2. Data on Toronto crime from the Police Service.

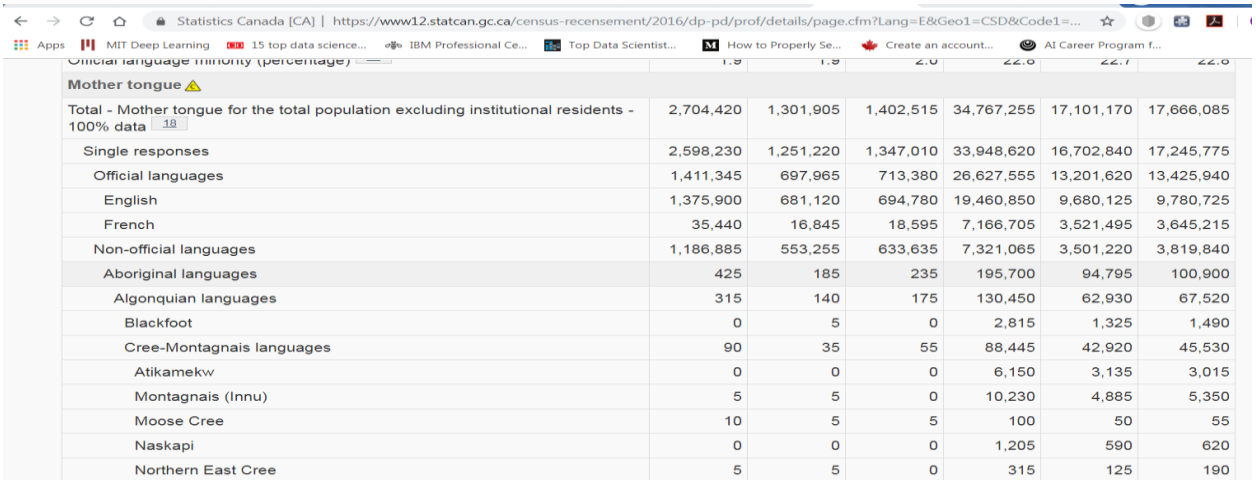
The demographic and population data will let us analyze potential demand and maybe give us a clue in the type of restaurant that should be opened (for example, the predominance of a certain language in the area might mean that type of cuisine will be popular). It is available in CSV format and will be read into a Pandas data frame for cleaning and analysis.



WB-Demographics [Защищенный просмотр]

	A	B	C	D	E	F	G	H
1			Demographics					
2	Neighbourhood	Neighbourhood Id	Total Area	Total Population	Pop - Males	Pop - Females	Pop 0 - 4 years	Pop 5 - 9 years
3	West Humber-Clairville	1	30,09	32265	16295	15960	2005	2135
4	Mount Olive-Silverstone-Jamestown	2	4,6	32130	15900	16230	2680	2680
5	Thistletown-Beaumond Heights	3	3,4	9925	4900	5035	615	625
6	Rexdale-Kipling	4	2,5	10725	5205	5525	580	645
7	Elms-Old Rexdale	5	2,9	9440	4615	4820	725	700
8	Kingsview Village-The Westway	6	5,1	21395	10265	11130	1500	1445
9	Willowridge-Martingrove-Richview	7	5,5	20920	10020	10880	1035	1140
10	Humber Heights-Westmount	8	2,8	10525	4715	5815	510	500
11	Edenbridge-Humber Valley	9	5,5	14450	6835	7620	570	720
12	Princess-Rosethorn	10	5,2	10965	5330	5625	380	580
13	Eringate-Centennial-West Deane	11	8,6	18535	8830	9690	750	990
14	Markland Wood	12	2,9	10240	4725	5515	385	445
15	Etobicoke West Mall	13	1,7	10680	4985	5695	535	600
16	Islington-City Centre West	14	16,4	32815	15520	17310	1590	1655
17	Kingsway South	15	2,6	8790	4165	4625	525	620
18	Stonegate-Queensway	16	7,9	23815	11450	12360	1235	1250

Image 3. A view of the demographic data for Toronto neighborhoods in xls format.



Statistics Canada [CA] | https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=...

Apps | MIT Deep Learning | 15 top data science... | IBM Professional Ce... | Top Data Scientist... | How to Properly Se... | Create an account... | AI Career Program f...

Mother tongue						
Total - Mother tongue for the total population excluding institutional residents - 100% data	2,704,420	1,301,905	1,402,515	34,767,255	17,101,170	17,666,085
Single responses	2,598,230	1,251,220	1,347,010	33,948,620	16,702,840	17,245,775
Official languages	1,411,345	697,965	713,380	26,627,555	13,201,620	13,425,940
English	1,375,900	681,120	694,780	19,460,850	9,680,125	9,780,725
French	35,440	16,845	18,595	7,166,705	3,521,495	3,645,215
Non-official languages	1,186,885	553,255	633,635	7,321,065	3,501,220	3,819,840
Aboriginal languages	425	185	235	195,700	94,795	100,900
Algonquian languages	315	140	175	130,450	62,930	67,520
Blackfoot	0	5	0	2,815	1,325	1,490
Cree-Montagnais languages	90	35	55	88,445	42,920	45,530
Atikamekw	0	0	0	6,150	3,135	3,015
Montagnais (Innu)	5	5	0	10,230	4,885	5,350
Moose Cree	10	5	5	100	50	55
Naskapi	0	0	0	1,205	590	620
Northern East Cree	5	5	0	315	125	190

Image 4. Data on Toronto languages

The neighborhood data and postal codes with longitude and latitude are needed for tools to make the visualizations and make the analysis easier to understand.

	Postcode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

Image 5. Transformed data from postal code and neighborhood data with geographical coordinates