Finding a spot for a Vegan/Vegetarian restaurant in Toronto

A. Introduction

A.1. Description & Disscusion of the Background

Our customer plans to open a chain of Vegan/Vegetarian Restaurants - The Veg-fo-tot (defining a catchy name is not part of this project). Most probable location to start is Toronto. As veganism/vegetarianism is becoming more and more popular there might be places in the city where such venues already exist. There are two findings of this analysis that we want to get:

1. The best locations to set up the first restaurant.

2. Based on the number of similar locations we find to decide if to open additional places or to focus on the first one only.

A.2. Data Description

We will use the following data:

1. Foursquare to find the already existing vegetarian/vegan restaurants.

2. Toronto City's open data Neighborhood profiles to retrieve the income data - Vegan food is not cheap so we need to find out neighborhoods which will be able to afford it. https://portal0.cf.opendata.inter.sandbox-toronto.ca/dataset/neighbourhood-profiles/

2a. We will use the above dataset to additionally retrieve data about population, commuting time, and education of the citizens. The dataset contains hundreds of columns with different data that is why it was necessary to process it deeply.

3. Toronto Police Safety Data Portal to retrieve data about criminality in different neighborhoods - Vegan hipsters might be very uneasy about the surroundings so the atmosphere around our restaurants must be safe. http://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-

4. To retrieve needed geodata and postalcodes for are master file I have used Wikipedia: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada
and again Toronto City's open data service: https://ckan0.cf.opendata.inter.sandbox-toronto.ca/download_resource/1d02b0f0-d735-4469-8f71-ea6d96b319e4?format=geojson&projection=4326

B. Methodology

As the first thing after downloading all datasets do I needed to choose the most interesting categories and clean up the data.

```
[161]: neighprof.shape

[161]: (2383, 146)
```

The main dataset –neighprof, was a massive structure of 2383 rows and 146 columns. This dataset contains a lot of the infomration about Toronto's neighbourhoods.

[162]:

	_id	Category	Topic	Data Source	Characteristic	City of Toronto	Agincourt North	Agincourt South-Malvern West	Alderwood	Annex	Banbury-Don Mills	Bathurst Manor	Bay Street Corridor	Ba
0	1	Neighbourhood Information	Neighbourhood Information	City of Toronto	Neighbourhood Number	NaN	129	128	20	95	42	34	76	
1	2	Neighbourhood Information	Neighbourhood Information	City of Toronto	TSNS2020 Designation	NaN	No Designation	No Designation	No Designation	No Designation	No Designation	No Designation	No Designation	Desig
2	3	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2016	2,731,571	29,113	23,757	12,054	30,526	27,695	15,873	25,797	
3	4	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2011	2,615,060	30,279	21,988	11,904	29,177	26,918	15,434	19,348	
4	5	Population	Population and dwellings	Census Profile 98-316-X2016001	Population Change 2011-2016	4.50%	-3.90%	8.00%	1.30%	4.60%	2.90%	2.80%	33.30%	

After a series of changes to dataset counted 140 rows and 15 columns:

```
[175]: neigh.shape

[175]: (140, 15)

[174]: neigh.head()
```

[174]:

Characteristic	Neighbourhood Number	Population, 2016	Low income (less than $19,999) in %	Low-mid income (less than $39,999) in %	Mid-low income (less than $59,999) in %	Mid-high income(less than $79,999) in %	High income ($80,000 and over) in %	No certificate, diploma or degree in %	Secondary (high) school diploma or equivalency certificate in %	Postsecondary certificate, diploma or degree in %	Less than 15 minutes in %	15 to 29 minutes in %	30 to 44 minutes in %	45 to 59 minutes in %	60 minutes and over in %
Agincourt North	129.0	29113.0	35.791571	16.435956	6.749562	4.276440	24.937313	22.498540	25.624292	37.732285	5.530176	10.905781	10.098581	4.843197	9.171161
Agincourt South-Malvern West	128.0	23757.0	32.474639	16.563539	7.766132	5.093236	19.973061	16.984468	25.634550	43.292503	6.124511	11.512396	10.586353	5.177421	9.386707
Alderwood	20.0	12054.0	26.173884	20.698523	12.900282	11.116642	23.519164	16.633483	24.556164	43.968807	10.038162	15.057242	12.402522	5.226481	7.424921
Annex	95.0	30526.0	22.095918	16.248444	11.662190	21.014873	22.226954	5.192295	13.988076	66.926554	7.632838	20.507109	13.103584	4.619013	2.964686
Banbury-Don Mills	42.0	27695.0	21.628453	17.367756	12.475176	17.367756	21.592345	8.286694	18.595414	57.591623	6.120238	12.547391	11.229464	5.560571	5.650840

From all the categories the most interesting were those related to the population number, income, education and commuting time.

The selected categories were formatted to display their value as a percentage of population in a specific Neighborhood.

An another interesting category (criminality) was not available in the neighprof dataset hence it was necessary to extract it from a different one – crimestats.

This dataset was much smaller and friendly:

[ ]: Starting data clean up for Crime Stats

[180]:
```
crimestats.shape
```

[180]: (140, 56)

[20]:
```
crimestats.head()
```

[20]:

	OBJECTID	Neighbourhood_Crime_Rates_Neigh	Neighbourhood_Crime_Rates_Hood_	Hood_ID	Neighbourhood	Assault_2014	Assault_2015	Assault_2016	Assault_
0	1	Yonge-St.Clair	97	97	Yonge-St.Clair	58	38	51	
1	2	York University Heights	27	27	York University Heights	78	101	111	
2	3	Lansing-Westgate	38	38	Lansing-Westgate	216	203	223	
3	4	Yorkdale-Glen Park	31	31	Yorkdale-Glen Park	121	141	136	
4	5	Stonegate-Queensway	16	16	Stonegate-Queensway	109	140	124	

After data formatting only one category was left – Total Crime per pop representing combined number of crimes in 2018 for specific neighbourhood. In the next steps it was also divided by the number of pop in each neighb.

```
crime.head()
```

:

Hood_ID	Neighbourhood	Total Crime per pop
97	Yonge-St.Clair	175
27	York University Heights	231
38	Lansing-Westgate	318
31	Yorkdale-Glen Park	261
16	Stonegate-Queensway	250

Both datasets were joined together:

]:

Neighbourhood Number	Population, 2016	Low income (less than $19,999) in %	Low-mid income (less than $39,999) in %	Mid-low income (less than $59,999) in %	Mid-high income(less than $79,999) in %	High income ($80,000 and over) in %	No certificate, diploma or degree in %	Secondary (high) school diploma or equivalency certificate in %	Postsecondary certificate, diploma or degree in %	Less than 15 minutes in %	15 to 29 minutes in %	30 to 44 minutes in %	45 to 59 minutes in %	60 minutes and over in %	Neighbourhood	Total Crime per pop
129	29113	35	16	6	4	24	22	25	37	5	10	10	4	9	Agincourt North	0.017621
128	23757	32	16	7	5	19	16	25	43	6	11	10	5	9	Agincourt South-Malvern West	0.006482
20	12054	26	20	12	11	23	16	24	43	10	15	12	5	7	Alderwood	0.018500

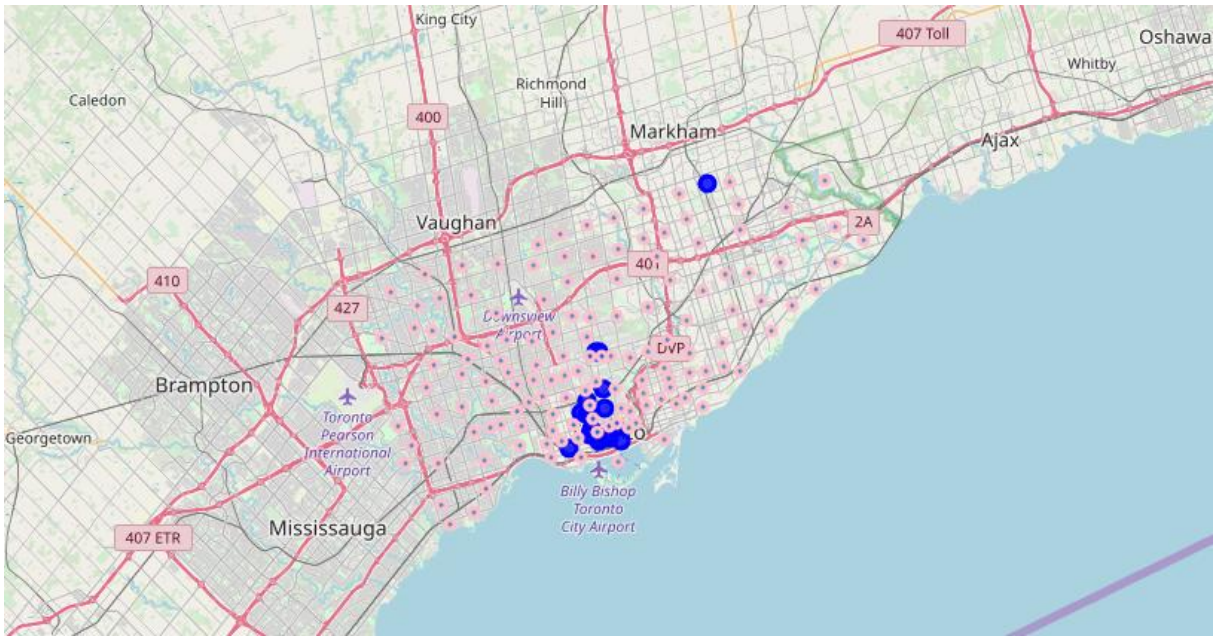
I have used Foursquare to retrieve vegetarian/vegan venues from the city. I did not set any limits and requested the information from the whole area of the city.

Initially the result was 20+ venues, but after additional filtering and matching their postcodes with neighborhood's names the final list contained 17 spots:

[41]:

	name	categories	address	cc	city	country	crossStreet	formattedAddress	labeledLatLngs	lat	lng	state	id	PostalCode	Neighbourhood
0	Lotus Pond Vegetarian Restaurant 蓮花素食	Vegetarian / Vegan Restaurant	3838 Midland Ave.	CA	Scarborough	Canada	btwn Passmore & McNicoll Ave.	[3838 Midland Ave. (btwn Passmore & McNicoll Ave.), Scarborough ON M1K 5V5, Canada]	[{'label': 'display', 'lat': 43.819420734434544, 'lng': -79.294681915014}]	43.819421	-79.294682	ON	4b78354df964a52070bc2ee3	M1K	East Birchmount Park
1	Mad Radish	Vegetarian / Vegan Restaurant	2293 Yonge Street	CA	Toronto	Canada	NaN	[2293 Yonge Street, Toronto ON M4P, Canada]	[{'label': 'display', 'lat': 43.707672, 'lng': -79.398424}]	43.707672	-79.398424	ON	5bd1fc9bfdb9a7002c331d1c	M4P	Davisville North
2	Fresh	Vegetarian / Vegan Restaurant	90 Eglinton Avenue East	CA	Toronto	Canada	Yonge & Eglinton	[90 Eglinton Avenue East (Yonge & Eglinton), Toronto ON M4P 1A6, Canada]	[{'label': 'display', 'lat': 43.707324410453595, 'lng': -79.39564918411965}]	43.707324	-79.395649	ON	521e0c6c04939a8ad55d93d3	M4P	Davisville North
3	nutbar	Vegetarian / Vegan Restaurant	1240 Yonge St	CA	Toronto	Canada	NaN	[1240 Yonge St, Toronto ON M4T 1W5, Canada]	[{'label': 'display', 'lat': 43.682810954698965, 'lng': -79.39194466929969}]	43.682811	-79.391945	ON	5890f70e65be581749d7522e	M4T	Moore Park
4	Vegan Bear	Vegetarian / Vegan Restaurant	NaN	CA	Toronto	Canada	NaN	[Toronto ON M5B 1R7, Canada]	[{'label': 'display', 'lat': 43.657124, 'lng': -79.380948}]	43.657124	-79.380948	ON	5cf43e873ba767002c95ce10	M5B	Ryerson

I have visualized them together with all the neighborhoods using folium library on the map of Toronto:



Later I have combined the number of venues for each neighborhood with the main table creating two new columns:

45 to 59 ninutes in %	60 minutes and over in %	Neighbourhood	Total Crime per pop	Number of veg venu	Has veg venue?
4	9	Agincourt North	0.017621	0.0	0.0
5	9	Agincourt South-Malvern West	0.006482	0.0	0.0
5	7	Alderwood	0.018500	0.0	0.0
4	2	Annex	0.005045	3.0	1.0
5	5	Banbury-Don Mills	0.002925	0.0	0.0

With this data I have used the K-means algorithm to cluster similar neighborhoods in 3 groups: 0-Bad Spot, 1-OKSpot, 2-GoodSpot:

[70]:

Labels	Population, 2016	Low income (less than $19,999) in %	Low-mid income (less than $39,999) in %	Mid-low income (less than $59,999) in %	I inc $7
0	17012.770833	21.770833	14.979167	10.437500	1
1	20264.607595	30.569620	17.670886	7.936709	
2	24158.000000	24.384615	17.461538	12.538462	1

Below version with the naming adopted:

Neighbourhood	Number of veg venu	Has veg venue?	Labels
Agincourt North	0.0	0.0	OKspot
Agincourt South-Malvern West	0.0	0.0	OKspot
Alderwood	0.0	0.0	OKspot
Annex	3.0	1.0	GoodSpot
Banbury-Don Mills	0.0	0.0	BadSpot
Bathurst Manor	0.0	0.0	OKspot
Bay Street Corridor	1.0	1.0	GoodSpot

Later on I have used the choropleth map to visualize the results.

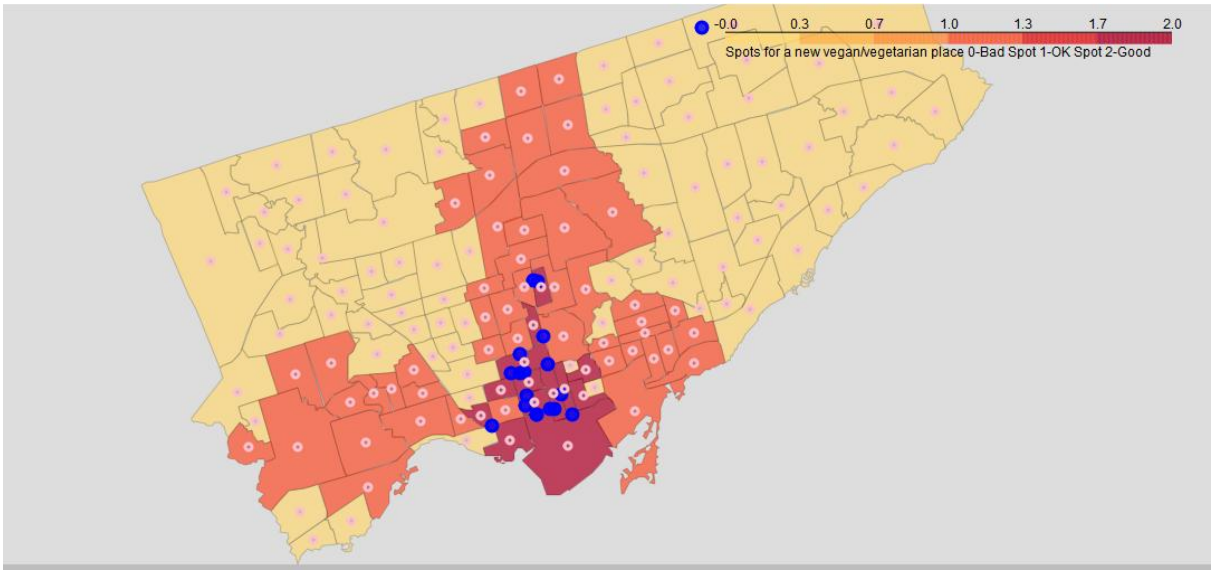
C. Results

After grouping the results we have found that 61 out of 139 neighborhoods would be worth to be treated as a potential place for a new vegetarian/vegan restaurant with 13 GoodSpot and 48 OKspot results.

[72]:

Labels	Neighbourhood
BadSpot	79
GoodSpot	13
OKspot	48

Here is how it looks on the map of Toronto.



D. Discussion

As already presented in the data above, Toronto has a very limited number of vegetarian/vegan venues numbering around 20 spots with most of them located more or less in the center of the city. Some of the neighborhoods in the top category already have vegetarian/vegan venues. Question which may arise is – should these neighborhoods be there?

The answer is – yes. As the choice of this kind of venues in the city is limited it should not affect the decision of treating the neighborhood as attractive. These places one of those in which our potential customers would look either for a vegetarian/vegan restaurant or a place to eat in general.

F. Conclusion

Toronto is still a place where competition on the vegetarian/vegan restaurants market is not high so with a good location and an idea for a good menu it could be a promising place for starting this kind of business.