Exploring Toronto Neighbourhoods for Best Place to Open Ramen Place

Capstone Project - The Battle of the Neighborhoods (Week 2)

Dan Liu
Applied Data Science Capstone

Table of Contents

1.0 Introduction

1.1 Business Problem

This purpose of this project is explore the City of Toronto neighbourhoods for the optimal location to open a Ramen Restaurant.

The following criteria were selected by the stakeholders to determine initial potential locations

1. High population
2. High population density area to allow for highest potential number of initial customers when opening
3. High percentage of population with income greater than $30,000
4. High percentage of population who are living alone
5. High percentage of population who are aged between 15 - 54 years old
6. High population percentage of immigrants originating from East Asian countries
7. Neighbourhood with high residential real estate price
8. Neighbourhood with lower number of existing vendors that are similar to Ramen
9. High Average Total income
10. Low crime rate for patrons
11. Low crime rate for owners and investors

The following weighting will be used to calculate weighted score for each neighbourhood for each criteria.

Criteria	Qualitative Weight	Quantitative Weight
1	Low	0.025
2	High	0.15
3	Medium	0.1
4	High	0.15
5	Medium	0.1
6	Low	0.025
7	Low	0.025
8	Medium	0.1
9	Medium	0.1
10	Medium-High	0.125
11	Medium	0.1

After initial exploration of the available data, it also determined to limit the geographical area to be close to where the owner of the ramen place resides to reduce traveling time:

- South of St. Clair Street
- East of Roncesvalles Avenue
- West of Coxwell Avenue

- Excluding the following areas to avoid existing competitions: Bay Corridor Yonge-Church Corridor.

This reduces the total number of neighbourhoods being examined from 140 to 32. By concentrating on the neighbourhoods close to the downtown core we are also targeting those who may work in downtown are looking to find dinner spots on their commute home to the suburbs.

1.2 Target Audience

The content of this project is intended following group of stakeholders
- Person who is interested in opening their own ramen restaurant in Toronto
- Investors who are interested in partnering to open a ramen restaurant or other similar type of restaurant

2.0 Data

Most of the data to be used to determine the best neighbourhood to open a Ramen Restaurant will be obtained from Toronto Open Data project:
The following particular datasets will be utilized in the project.

2.1 Neighbourhood Profiles
https://open.toronto.ca/dataset/neighbourhood-profiles/

This data set was last refreshed in July 2019 and contains the data for each of City of Toronto's 140 social planning neighbourhoods of Toronto collected during the latest Census of Population in 2016. Data collected contains: age and sex, families and households, language, immigration and internal migration, ethnocultural diversity, Aboriginal peoples, housing, education, income, and labour.
These social planning neighbourhoods were developed by the City of Toronto to help government and community organizations with local planning by providing socio-economic data at a meaningful geographic area. The boundaries of these social planning neighbourhoods are consistent over time, allowing for comparison between Census years
The Census data to provide a portrait of the demographic, social and economic characteristics of the people and households in each City of Toronto neighbourhood, where in this case we are particularly interested in characteristics of people in favour over households.
The data was sourced from a number of Census tables released by Statistics Canada.
Particularly, we are interested in the following data points for each neighbourhoods
- Population and dwellings
    - Population, 2016
    - Population density per square kilometre
- Age characteristics
    - Youth (15-24 years)
    - Working Age (25-54 years)
- Family characteristics of adults
    - Persons living alone (per cent)
- Income of individuals in 2015
    - Total - Employment income groups in 2015 for the population aged 15 years and over in private households - 100% data
        - Under $5,000(including loss), $5,000 to 9,999, $10,000 to $19,999, $20,000 to $29,999, $30,000 to $39,999, $40,000 to $49,999, $50,000 to $59,999, $60,000 to $69,999, $70,000 to $79,999, $80,000 and over
- Total income: Average amount ($)
- Immigrants by selected place of birth
    - Total - Selected places of birth for the immigrant population in private households - 25% sample data
        - Asia
        - China, Hong Kong, Japan, Korea; South, Philippines, Taiwan, Viet Nam
The percentage is calculated by tabulating the total of each subcategory and dividing by the total respondents.

Neighbourhood	Longitude	Latitude	Population, 2016	Population density per square kilometre	Persons living alone (per cent)	Target Populations (15-54 years)	Target Populations (15-54 years) Percent	Target Populations (by income $30,000+)	Target Populations (by income $30,000+) Percent	Average income amount ($)	Target Populations (by immigration from East Asia)	Target Populations (by immigration from East Asia) Percent
Waterfront Communities-The Island (77)	-79.377202	43.633880	65913	8943	35.6	52945	80.325581	40350	65.085894	70600	6735	28.292376
Woburn (137)	-79.228586	43.766740	53485	4345	9.7	29605	55.351968	14785	34.062896	30878	5840	19.350563
Willowdale East (51)	-79.401484	43.770602	50434	10087	16.6	32790	65.015664	19025	42.916761	45326	15360	49.484536
Rouge (131)	-79.186343	43.821201	46496	1260	4.3	25210	54.219718	16770	43.998426	39556	4220	16.497263
L'Amoreaux (117)	-79.314084	43.795716	43993	6144	8.0	22940	52.144659	12590	33.681113	31826	13915	49.117543
Islington-City Centre West (14)	-79.543317	43.633463	43965	2712	17.8	25335	57.625384	19725	52.712453	52787	3915	19.989788
Malvern (132)	-79.222517	43.803658	43794	4948	5.9	24485	55.909485	12365	34.457294	29573	6090	22.635198
Dovercourt-Wallace Emerson-Junction (93)	-79.438541	43.665677	36625	9819	14.6	23715	64.750853	13870	43.582090	39740	2880	20.027816
Downsview-Roding-CFB (26)	-79.490497	43.733292	35052	2337	11.7	19385	55.303549	11375	38.935478	34168	5360	27.459016
Parkwoods-Donalda (45)	-79.330180	43.755033	34805	4691	11.7	19460	55.911507	12605	43.631014	42516	4115	23.910517

Figure 1 Data from Neighbourhood Profile and Neighbourhoods Boundary

2.2 Neighbourhoods

https://open.toronto.ca/dataset/neighbourhoods/

This dataset in .geojson format contains boundaries as well as the geographic coordinates (latitude and longitude) of the 140 City of Toronto neighbourhoods for social planning purposes. The geographic coordinates will be used to call the Foursquare API to determine the upto 10 food vendors within the center of each neighbourhood. The boundaries information will be used to generate choropleth map with Folium on a Leaflet map. The file was converted to .json file for the purpose of this project and easy reading into

	Neighbourhood Number	Neighbourhood	Longitude	Latitude
0	1	West Humber-Clairville (1)	-79.596356	43.716180
1	2	Mount Olive-Silverstone-Jamestown (2)	-79.587259	43.746868
2	3	Thistletown-Beaumond Heights (3)	-79.563491	43.737988
3	4	Rexdale-Kipling (4)	-79.566228	43.723725
4	5	Elms-Old Rexdale (5)	-79.548983	43.721519

Figure 2 Data from Neighbourhood Boundaries

2.3 Foursquare

https://developer.foursquare.com/

Foursquare API is used to collect data on close by venues within a specific radius of a given geographic coordinate. For the purpose of this project we are only looking at food related venues, by pass 'food' to the parameter 'Section' in the Get Venue Recommendations API call. (https://developer.foursquare.com/docs/api/venues/explore) The data returned will be used to explore existing make up for the food venue landscape of each neighbourhood.

20 Most Frequently Occuring Food Venues in Toronto

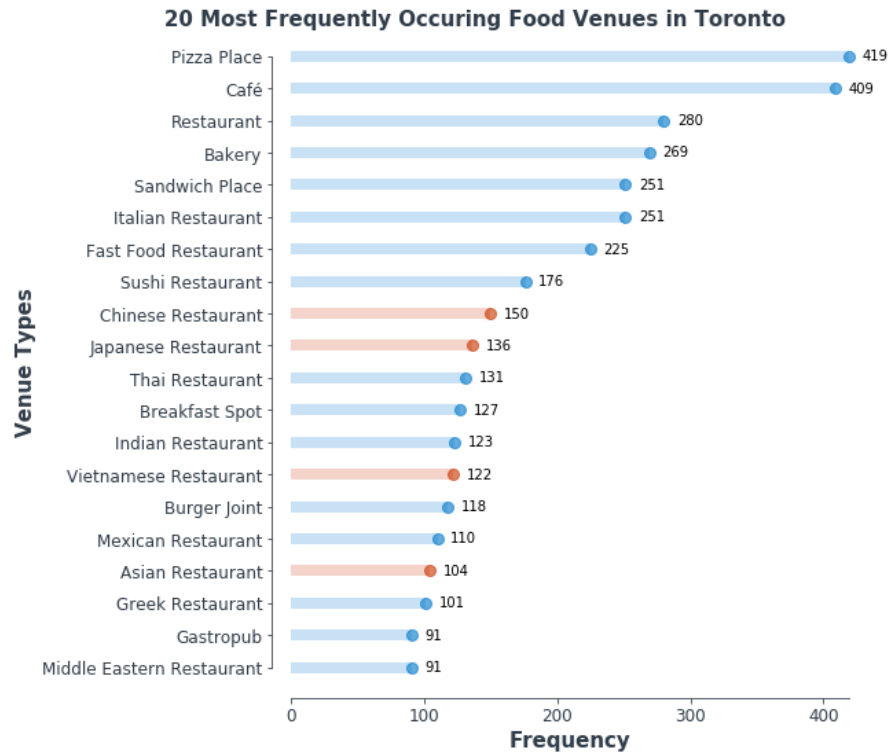| Venue Type	Frequency
Pizza Place	419
Café	409
Restaurant	280
Bakery	269
Sandwich Place	251
Italian Restaurant	251
Fast Food Restaurant	225
Sushi Restaurant	176
Chinese Restaurant	150
Japanese Restaurant	136
Thai Restaurant	131
Breakfast Spot	127
Indian Restaurant	123
Vietnamese Restaurant	122
Burger Joint	118
Mexican Restaurant	110
Asian Restaurant	104
Greek Restaurant	101
Gastropub	91
Middle Eastern Restaurant	91

Figure 3 20 Most Frequently Occurring Food Venues in Toronto

	Neighbourhood	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agincourt North (129)	Chinese Restaurant	Bakery	Vietnamese Restaurant	Pizza Place	Indian Restaurant	Sandwich Place	Fast Food Restaurant	Hong Kong Restaurant	Fried Chicken Joint	Restaurant
1	Agincourt South-Malvern West (128)	Chinese Restaurant	Restaurant	Bakery	Asian Restaurant	Cantonese Restaurant	Sandwich Place	Breakfast Spot	Pizza Place	Hong Kong Restaurant	Vietnamese Restaurant
2	Alderwood (20)	Pizza Place	Donut Shop	Sandwich Place	Moroccan Restaurant	Xinjiang Restaurant	Diner	Dumpling Restaurant	Eastern European Restaurant	Egyptian Restaurant	Empanada Restaurant
3	Annex (95)	Café	Restaurant	Italian Restaurant	French Restaurant	Pizza Place	Vegetarian / Vegan Restaurant	Bakery	Sushi Restaurant	Japanese Restaurant	Indian Restaurant
4	Banbury-Don Mills (42)	Pizza Place	Japanese Restaurant	Café	Restaurant	American Restaurant	Sandwich Place	Caribbean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Burger Joint

Figure 4 Top 10 Food Venue from Foursquare (sample)

Neighbourhood	Ramen Restaurant	Udon Restaurant	Noodle House	Vietnamese Restaurant	Soup Place	Japanese Restaurant	Asian Restaurant	Chinese Restaurant	Korean Restaurant	Taiwanese Restaurant
Agincourt North (129)	0	0	0	2	0	1	0	4	0	0
Agincourt South-Malvern West (128)	0	0	1	2	0	1	3	13	1	0
Alderwood (20)	0	0	0	0	0	0	0	0	0	0
Annex (95)	0	0	0	1	0	4	1	0	3	0
Banbury-Don Mills (42)	0	0	0	0	0	3	0	0	0	0
Bathurst Manor (34)	0	0	0	0	0	1	0	1	0	0
Bay Street Corridor (76)	4	0	0	0	0	6	2	2	0	0
Bayview Village (52)	0	0	0	0	0	1	0	2	0	0

Figure 5 Foursquare Data - Venues Similar to Ramen

2.4 2018 Toronto Detached Prices by Neighbourhood

https://docs.google.com/spreadsheets/d/1GizBkDvGaYm5AAPJ7O0y9eq-n7i81to6WpjejvYFzw0/edit#gid=754169703

This data set contain average detached house prices by 144 TREB (Toronto Real Estate Board) neighbourhoods aggregated by Scott Ingram. As Scott mentioned, detached house prices were used to keep things simple and more apples-to-apples between neighbourhood. (http://www.century21.ca/scott.ingram/blog/Toronto_s_most_and_least_expensive_neighbourhoods_in_2018)

In order for the data to be used, the TREB neighbourhood must be converted to the 140 City of Toronto social planning neighbourhoods. For neighbourhood that does not have any sales, the overall average price for Toronto will be used. This data can be used to determine the state of the real estate price.
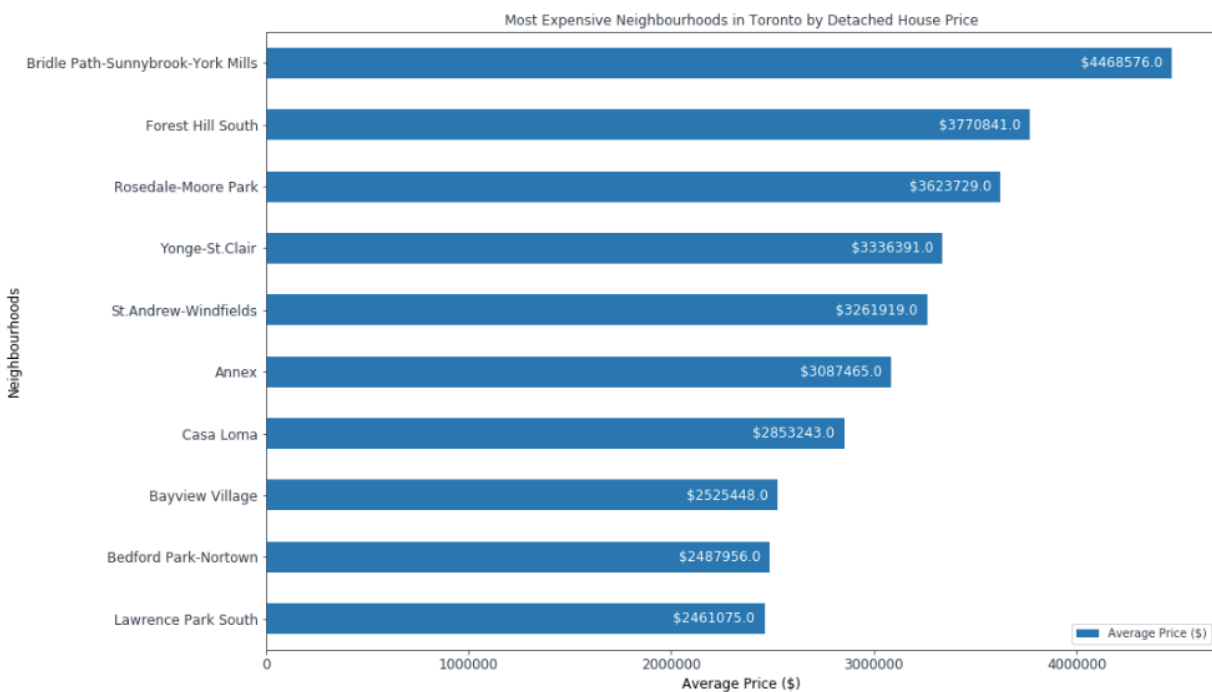

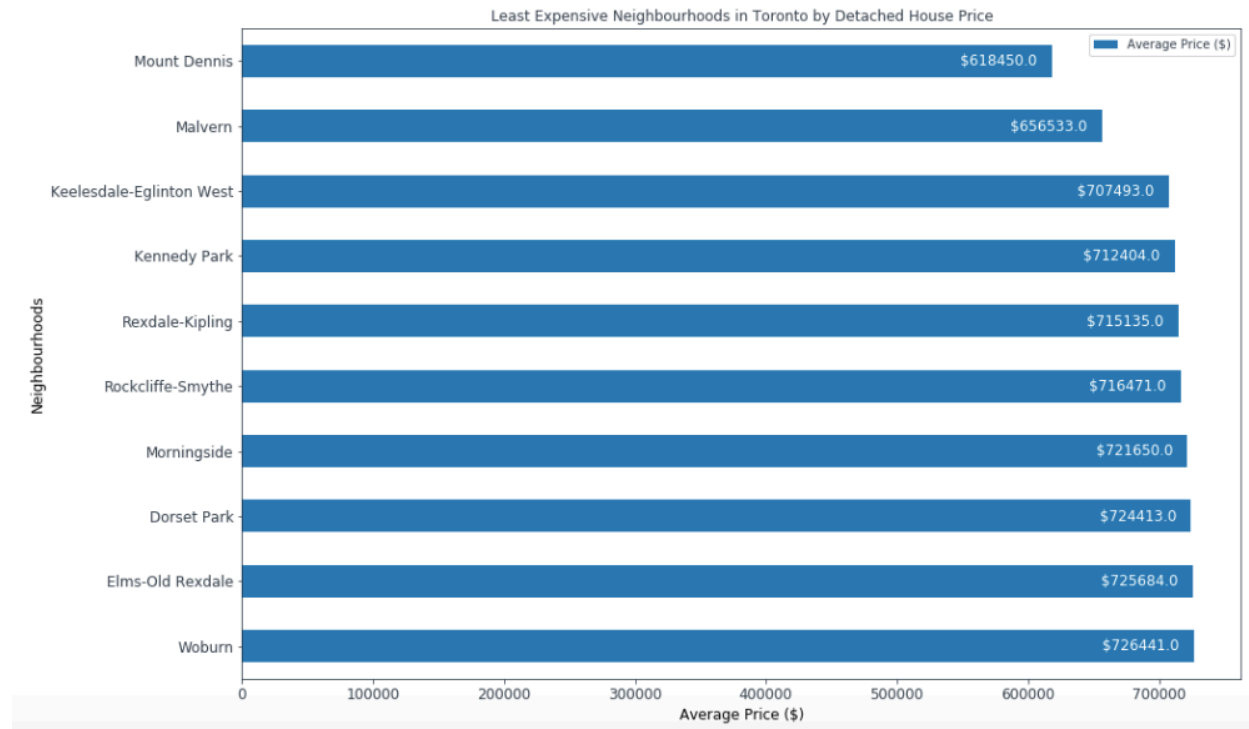*Figure 6 Most Expensive Real Estate Neighbourhoods in Toronto*

Figure 7 Least Expensive Real Estate Neighbourhoods in Toronto

2.5 Neighbourhood Crime Rates

http://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-

This dataset contains Crime Data by Neighbourhood. the 2014-2018 period. Counts are available for Assault, Auto Theft, Break and Enter, Robbery, Theft Over and Homicide. Data also includes four year averages and crime rates per 100,000 people by neighbourhood based on 2016 Census Population.

The crime rates will be used to calculate the safety rate for the potential patrons and the owners and investors. For patrons we will consider Auto Theft, Robbery, and Homocide. For owners and investors we will consider Theft Over.

Definition of Crimes

Theft Over: The act of stealing property in excess of $5,000 (excluding auto theft).

Robbery: The act of taking property from another person or business by the use of force or intimidation in the presence of the victim.

Auto Theft: The act of taking or another person's vehicle (not including attempts). Auto Theft figures represent the number of vehicles stolen.

Break and Enter: The act of entering a place with the intent to commit an indictable offence therein.

Homicide/Murder: The homicide category includes the offences of First Degree Murder, Second Degree Murder, and Manslaughter. A homicide occurs when a person directly or indirectly, by any means, causes the death of another human being. Deaths caused by criminal negligence, suicide, or accidental or justifiable homicide (i.e self-defence) are not included. Homicide data is compiled based on the Homicide Squad Case List Log. Count is based on offence (i.e each deceased victim)
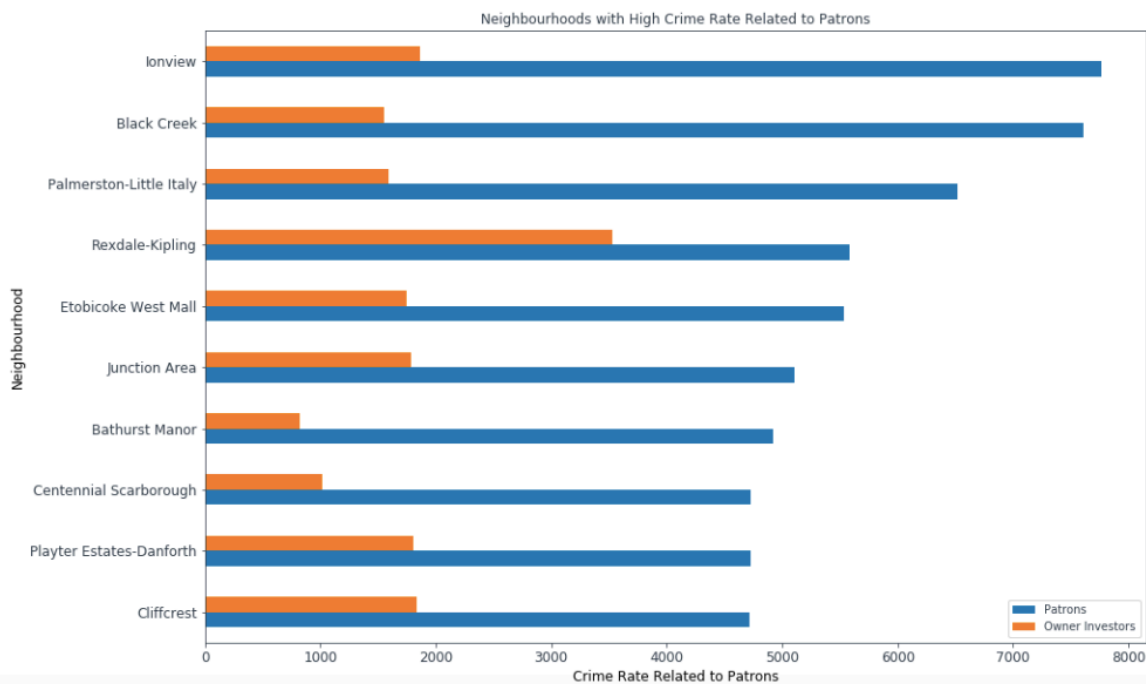

Figure 8 High Crime Rate Neighbourhood in Toronto

3.0 Methodology

To Narrow down the scope of this project, we will limit the neighbourhoods examined to the following geographical areas. This reduces the total number of neighbourhoods being examined from 140 to 32.

- South of St. Clair Street
- East of Roncesvalles Avenue
- West of Coxwell Avenue
- Excluding the following areas to avoid existing competitions: Bay Corridor Yonge-Church Corridor.

By concentrating on the neighbourhoods close to the downtown core we are also targeting those who may work in downtown are looking to find dinner spots on their commute home to the suburbs.

A decision matrix will be used to determine potential locations with the following criteria:

1. High population
2. High population density area to allow for highest potential number of initial customers when opening
3. High percentage of population with income greater than $30,000
4. High percentage of population who are living alone
5. High percentage of population who are aged between 15 - 54 years old
6. High population percentage of immigrants originating from East Asian countries
7. Neighbourhood with high residential real estate price
8. Neighbourhood with lower number of existing vendors that are similar to Ramen
9. High Average Total income
10. Low crime rate for patrons
11. Low crime rate for owners and investors

A standardized score for each criteria will be calculated for each neighbourhood by dividing the value from the neighbourhood profiles by the average of the neighbourhoods examined.

The following weighting will be applied to calculate weighted score for each neighbourhood for each criteria

Criteria	Qualitative Weight	Quantitative Weight
1	Low	0.025
2	High	0.15
3	Medium	0.1
4	High	0.15
5	Medium	0.1
6	Low	0.025
7	Low	0.025
8	Medium	0.1
9	Medium	0.1
10	Medium-High	0.125
11	Medium	0.1

The neighbourhoods with top combined weighted score will be further examined and recommended.

4.0 Analysis

A decision-making matrix was calculated using the criteria and dividing the value from the neighbourhood profiles by the average of the neighbourhoods examined.

Neighbourhood Number	Neighbourhood Number	Neighbourhood	Population Score	Population Density Score	Income $30,000+ Score	Population Living Alone Score	15-54 Years Old Score	Immigration from East Asia Score	Real Estate Score	Ramen Score	Average Income Score	Patron Related Crime Rate Score	Owner Related Crime Rate Scoe	Total Score
76	77	Waterfront Communities-The Island (77)	0.091	0.135	0.128	0.239	0.128	0.029	0.020	0.949	0.109	0.983	0.964	3.777
97	98	Rosedale-Moore Park (98)	0.029	0.068	0.129	0.156	0.077	0.016	0.053	0.962	0.322	0.969	0.965	3.745
72	73	Moss Park (73)	0.028	0.223	0.098	0.271	0.114	0.023	0.021	0.898	0.091	0.964	0.951	3.682
94	95	Annex (95)	0.042	0.164	0.112	0.202	0.098	0.021	0.045	0.886	0.175	0.962	0.953	3.659
73	74	North St.James Town (74)	0.026	0.669	0.070	0.190	0.108	0.033	0.020	0.784	0.051	0.786	0.879	3.616
84	85	South Parkdale (85)	0.030	0.145	0.076	0.211	0.108	0.027	0.021	0.962	0.055	0.967	0.985	3.586
81	82	Niagara (82)	0.043	0.153	0.133	0.239	0.132	0.024	0.026	0.886	0.109	0.863	0.972	3.579
95	96	Casa Loma (96)	0.015	0.086	0.124	0.159	0.082	0.012	0.042	0.975	0.256	0.911	0.881	3.542
69	70	South Riverdale (70)	0.038	0.047	0.100	0.126	0.100	0.055	0.018	1.000	0.083	0.987	0.973	3.529
70	71	Cabbagetown-South St.James Town (71)	0.016	0.126	0.107	0.219	0.094	0.027	0.023	0.860	0.098	0.974	0.966	3.510
86	87	High Park-Swansea (87)	0.033	0.074	0.119	0.143	0.092	0.014	0.027	1.000	0.110	0.938	0.954	3.503
65	66	Danforth (66)	0.013	0.129	0.097	0.101	0.088	0.019	0.017	0.962	0.086	0.981	0.989	3.483
88	89	Runnymede-Bloor West Village (89)	0.014	0.096	0.120	0.071	0.088	0.013	0.019	0.975	0.111	0.970	0.965	3.442
92	93	Dovercourt-Wallace Emerson-Junction (93)	0.051	0.148	0.086	0.098	0.104	0.021	0.017	0.924	0.062	0.955	0.973	3.438
64	65	Greenwood-Coxwell (65)	0.020	0.130	0.094	0.110	0.095	0.037	0.015	0.937	0.072	0.963	0.964	3.437
82	83	Dufferin Grove (83)	0.016	0.129	0.090	0.130	0.105	0.019	0.026	0.898	0.063	0.970	0.938	3.385

5.0 Results

Top 10 Neighbourhood to open Ramen Restaurant in Toronto

	Neighbourhood Number	Neighbourhood	Total Score
76	77	Waterfront Communities-The Island (77)	3.777
97	98	Rosedale-Moore Park (98)	3.745
72	73	Moss Park (73)	3.682
94	95	Annex (95)	3.659
73	74	North St.James Town (74)	3.616
84	85	South Parkdale (85)	3.586
81	82	Niagara (82)	3.579
95	96	Casa Loma (96)	3.542
69	70	South Riverdale (70)	3.529
70	71	Cabbagetown-South St.James Town (71)	3.510

Choropleth Map of the Total Scores of the selected neighbourhoods
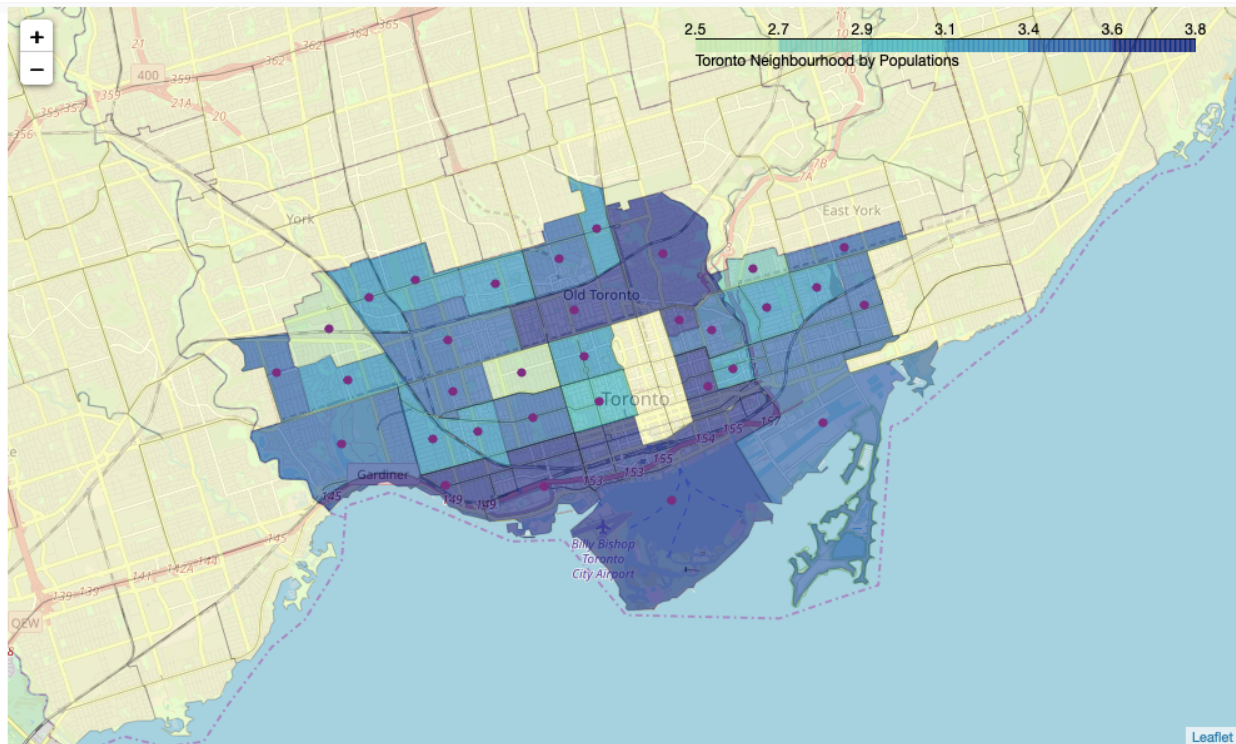
6. 0 Discussion & Recommendations

Based on the results calculated we can make the following observations and recommendations:

- Based on the weighted criteria selected the most optimal neighborhood to open a ramen restaurant would be Waterfront Communities — The Island. With a total score of 3.777. This area has low number of existing food establishments that are similar to ramen, low crime rates, high population, and individuals with income more than $30,000.
- Close in second place for recommendation is Rosedale-Moore Park. This area has high average income, low crime rate, and low number of existing food establishments that are similar to ramen.
- In a surprise outcome, Moss Park came in third as recommended place to open a ramen restaurant, even though the neighbourhood itself has a bad reputation within the city of Toronto. The neighbourhood has relative low crime rate, high percentage of people living alone as well as high percentage of 15-54 years old populations.
- Annex is the fourth recommended neighbourhood
- The following neighbourhoods should be avoided based on the calculated scores:
  - Weston-Pellam Park
  - North Riverdale
  - Kensington-Chinatown
  - Playter Estates-Danforth
  - Junction Area
  - Palmerston-Little Italy

This is due to higher crime rate and higher number of existing similar to ramen food establishments.

7.0 Conclusion

Purpose of this project was to help stakeholders identify optimal neighbourhoods in Toronto to open a ramen restaurant. Four main data sources were used: Foursquare, Toronto Neighbourhood Profiles, Toronto Average Detached House Sale Price 2018, and Toronto Crime Rate by Neighbourhood.

A total of 11 criteria were selected: population, population density, population living alone, population aged between 15-54, population with greater than $30,000 income, average income, existing food establishment similar to ramen, detached house prices and crime rate affecting patrons and owner.

Each selection criteria were assigned a weight. The neighbourhood data were standardized by dividing it by the mean of each category. The weights were then multiplied to calculated the weighted score. The neighbourhood with highest total weighted score is recommended for the best location to open the ramen restaurant.

The final decision on restaurant location will be made by stakeholders based on other factors that could not be considered for this specific project such as availability of commercial real estate space and price, parking availability, traffic, and etc..

Due to the limited scope of this of the analysis, only a small section of the neighbourhoods were selected and with a small set of features. With further features selected the 140 neighbourhoods could be clustered into groups using machine learning algorithm to determine groups of neighbourhood with similar characteristics. However, with the current the model, it is easily replicated for future census data as well as updated Foursquare data.