

Rental Unit – Rating Engine

Introduction:

One of the biggest challenges while looking for a new home is to find neighborhood information of a rental listing. Important information like proximity to schools, restaurants, playgrounds etc. could be very useful while looking at a listing in a rental website. This additional information would give a potential tenant to decide the worthiness of a house based on his/her preferences. For example, families with young children would be interested in schools and other child activities while young couples would be interested in dining and entertainment options nearby.

For this project, I will focus on the city of Toronto. The solution would provide scores to a rental listing based on 7 metrics

1. Arts & Entertainment (Museums, Movie theatres, etc.)
2. Food options (Restaurants, Cafes, Deli's etc.)
3. Nightlife (Bars, Night clubs, etc.)
4. Medical centers (Hospitals, Clinics etc.)
5. Education (Schools and other places of learning)
6. Spiritual (Places of worship, Meditation centers, etc.)
7. Shops (Convenience, Grocery stores, etc.).

Ratings would be calculated based on comparing the number of venues nearby with the average number of nearby venues Toronto.

The goal of this project is to give easily interpretable scores that would highlight the location advantage for a rental listing, compared to other listings, and aid the customer in decision making.

Data:

Neighborhood data is obtained from Toronto city website with standard neighborhood names (<https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/neighbourhood-profiles/>). Latitude and Longitude for each neighborhood identifier is obtained using Nominatim geocoder. Foursquare API venue search is then used to get the number of venues in a neighborhood for the 7 categories we are interested in. Top level categories are used to capture all sub-categories as well.

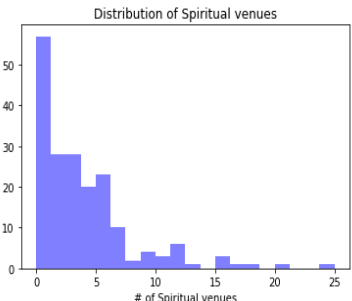
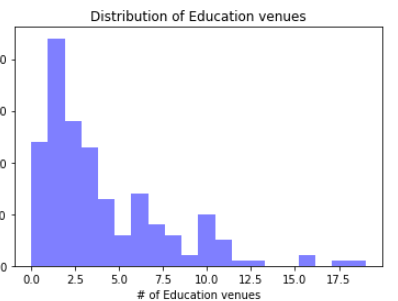
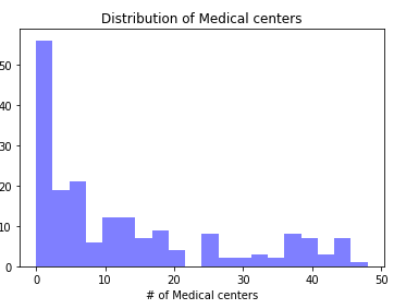
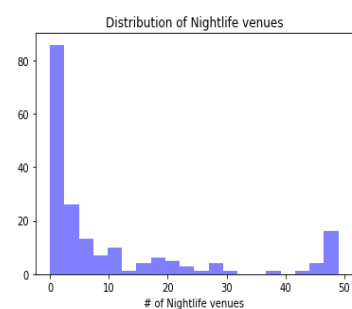
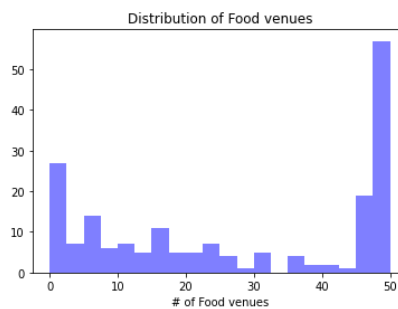
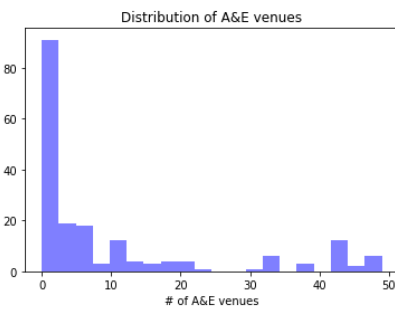
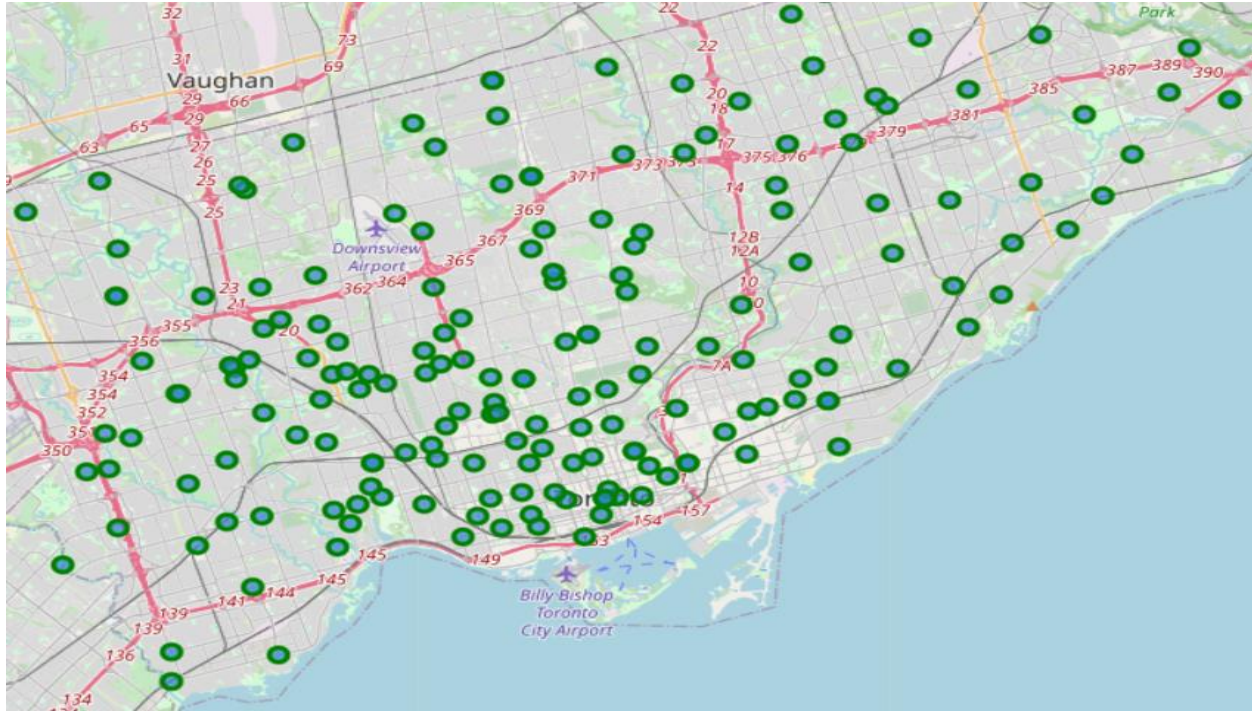
	Neighborhood	Latitude	Longitude	A&E	Food Options	Nightlife	Medical centers	Education	Spiritual	Shops
0	Agincourt North	43.808038	-79.266439	2	32	2	5	6	0	35
1	Agincourt South	43.785353	-79.278549	2	45	2	17	3	5	29
2	Alderwood	43.601717	-79.545232	3	21	4	9	1	1	31
3	Amesbury	43.706162	-79.483492	1	11	0	2	1	2	15
4	Annex	43.670338	-79.407117	33	48	23	33	11	15	44

Rental unit addresses for this project are provided as a csv file, while in a real-time application all available rental units in an area requested by the user will be provided as input.

Rental Unit – Rating Engine

Methodology:

There are about 200 neighborhoods in Toronto, with the downtown having more venues than some of the suburban neighborhoods, although the suburbs are fast catching up. For our rating requirement, we assume the 200 neighborhoods to be normally distributed samples of Toronto. The metrics we obtain from these samples could be used to calculate mean and standard deviation of the corresponding metrics for Toronto.



Rental Unit – Rating Engine

Due to the non-negative data and the max venue limit 50 from foursquare APIs, we expect the distribution to be skewed. From the graphs, we can infer that for A&E, Nightlife, Medical centers, Education and Spiritual are positively skewed and Food and shops negatively skewed.

```
df_Toro.describe()
```

	Latitude	Longitude	A&E	Food Options	Nightlife	Medical centers	Education	Spiritual	Shops
count	189.000000	189.000000	189.000000	189.000000	189.000000	189.000000	189.000000	189.000000	189.000000
mean	43.705238	-79.405799	10.169312	27.455026	10.391534	13.100529	3.735450	3.910053	27.105820
std	0.049391	0.097077	14.759523	19.557997	14.937285	13.875796	3.747888	4.118892	15.546068
min	43.592005	-79.599116	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	43.665519	-79.475998	1.000000	7.000000	1.000000	2.000000	1.000000	1.000000	13.000000
50%	43.695689	-79.410915	3.000000	25.000000	3.000000	7.000000	2.000000	3.000000	32.000000
75%	43.743176	-79.352594	12.000000	48.000000	12.000000	19.000000	6.000000	5.000000	40.000000
max	43.823174	-79.150768	49.000000	50.000000	49.000000	48.000000	19.000000	25.000000	50.000000

Ratings are calculated as follows. The range of results are limited from “mean-std” to “mean+std”. “mean-std” and below are rated as “1”, “mean” is rated 3 and “mean+std” and above are rated “5”, while any intermediate value is interpolated linearly.

```
def get_rating(value,mean,std):
    rating = 3
    if value>=(mean+std):
        rating = 5
    elif value<=(mean-std):
        rating = 1
    elif value>mean:
        rating = 3+(value-mean)/std*2
    else:
        rating = 3-(mean-value)/std*2
    return round(rating,1)
```

Using the same venue search Foursquare API, we find the number of venues around each rental unit in our list. To balance the venues around the commercial center of neighborhood vs a residential location, the venue search on the rental unit is performed with a radius of 1km vs the 500m neighborhood search.

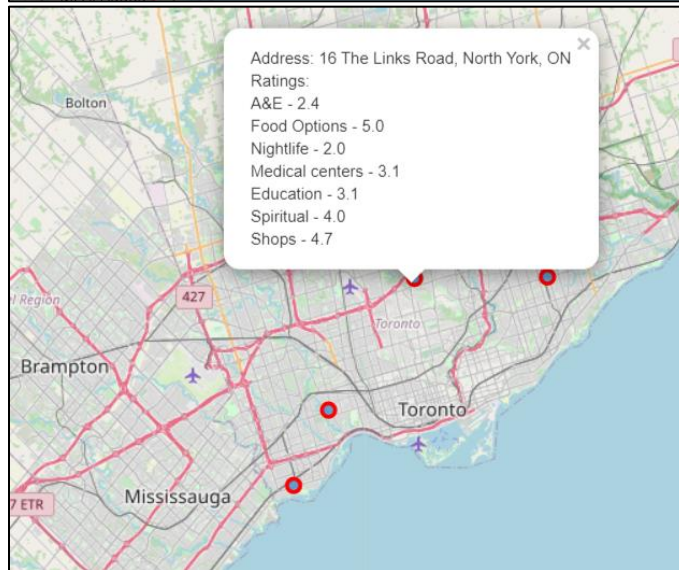
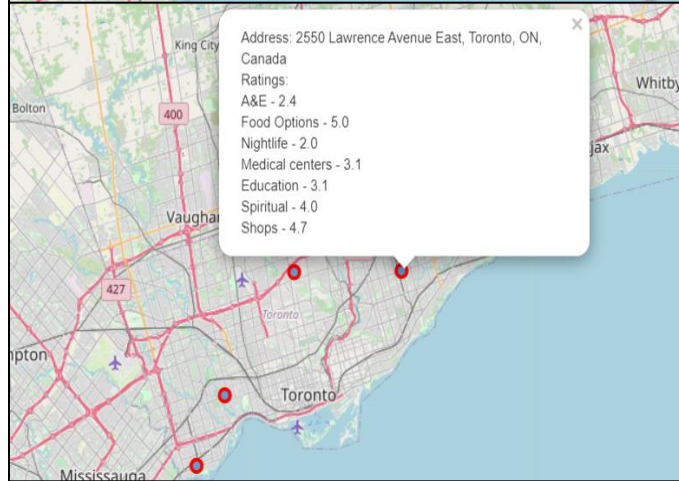
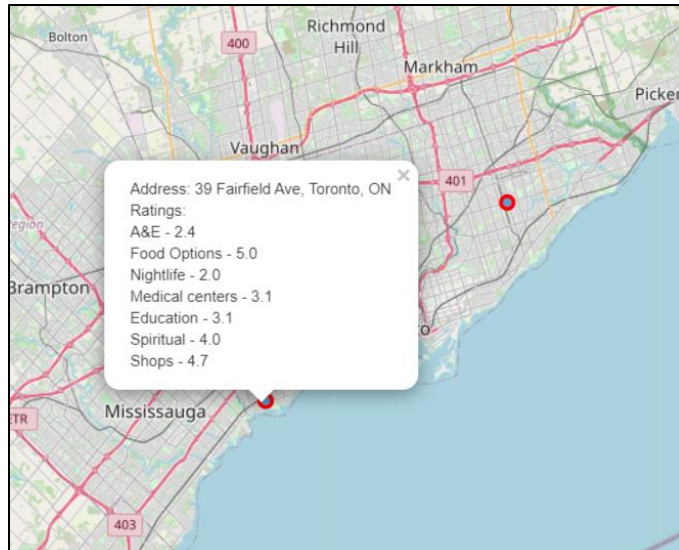
Results:

Ratings were calculated for the rental units we are interested in and plotted in a map, with a popup showing the ratings for that unit. As expected, units closer to commercial streets had higher ratings vs units in deep residential areas. Toronto being a populous city with well spread commercial areas, food and grocery options are available throughout and all rental units ranked high in these categories. Nightlife and Medical centers are the two venue categories that differentiated rental units. Again, this is expected based on the geography and population distribution.

Rental Unit – Rating Engine

The geographical coordinate of "39 Fairfield Ave, Toronto, ON" is 43.59705006153846, -79.52569320769231.			
Index	#nearby venues	rating	
0 A&E	21	4.5	
1 Food Options	45	4.8	
2 Nightlife	10	2.9	
3 Medical centers	12	2.8	
4 Education	7	4.7	
5 Spiritual	7	4.5	
6 Shops	40	4.7	
The geographical coordinate of "2 Old Mill Dr, Toronto, ON" is 43.6526324, -79.4900624.			
Index	#nearby venues	rating	
0 A&E	19	4.2	
1 Food Options	49	5.0	
2 Nightlife	18	4.0	
3 Medical centers	36	5.0	
4 Education	8	5.0	
5 Spiritual	11	5.0	
6 Shops	42	4.9	
The geographical coordinate of "60 Horsham ave, Toronto, ON" is 43.77408805102328, -79.41745152303463.			
Index	#nearby venues	rating	
0 A&E	28	5.0	
1 Food Options	50	5.0	
2 Nightlife	41	5.0	
3 Medical centers	38	5.0	
4 Education	16	5.0	
5 Spiritual	16	5.0	
6 Shops	39	4.5	
The geographical coordinate of "16 The Links Road, North York, ON" is 43.750688999999994, -79.4025758174603.			
Index	#nearby venues	rating	
0 A&E	10	3.0	
1 Food Options	47	5.0	
2 Nightlife	10	2.9	
3 Medical centers	31	5.0	
4 Education	5	3.7	
5 Spiritual	10	5.0	
6 Shops	26	2.9	
The geographical coordinate of "2550 Lawrence Avenue East, Toronto, ON, Canada" is 43.751345, -79.2668813.			
Index	#nearby venues	rating	
0 A&E	6	2.4	
1 Food Options	49	5.0	
2 Nightlife	3	2.0	
3 Medical centers	14	3.1	
4 Education	4	3.1	
5 Spiritual	6	4.0	
6 Shops	40	4.7	

Rental Unit – Rating Engine



Rental Unit – Rating Engine

Discussion:

The ratings obtained by the algorithm matched my experiences staying in the addresses used in the program. We can use these ratings as a good starting point when comparing rental listings. There could be further improvements to the listing information by suggesting top trending venue for each category and the distance/time to those venues from the rental unit. Proximity data would further induce a potential renter to choose one rental unit over another.

Our current rating is based on established neighborhood commercial centers and calculating mean and standard deviation based on that data. If we were to use true random samples (randomly picking latitude, longitude within Toronto), ratings might change if the mean differs significantly from our current calculation. Further analysis is required to see if true random samples have any significant effect on venue mean and standard deviation.

Conclusion:

To assist renters in Toronto who are looking at rental property listings, this project provides them with proximity information on 7 categories of venues such as education, health, food options, etc. Instead of showing number of options, this project provides them with a rating that would be much more useful as a comparative tool.

This project calculated the ratings for 7 categories that are identified as relevant to most groups of renters. Ratings are calculated based on the average availability of venues belonging to the 7 categories in Toronto in a given radius and the availability near the rental. Improvements to rating calculation and to the user interaction (input/output format) were discussed and could improve the usability of the project.