Rental Unit – Rating Engine

Rating availability of important venues near a listing

Currently limited information to compare rental units in Toronto listings

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.

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.

The project would provide scores to a rental listing based on 7 metrics listed below, based on various Toronto neighborhood statistics.

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.).

We will use Toronto city public data and Foursquare API to get venue information

- Neighborhood data is obtained from Toronto city website with standard neighborhood names

- Latitude and Longitude for each neighborhood is obtained using Nominatim geocoder.

- Foursquare API venue search is 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

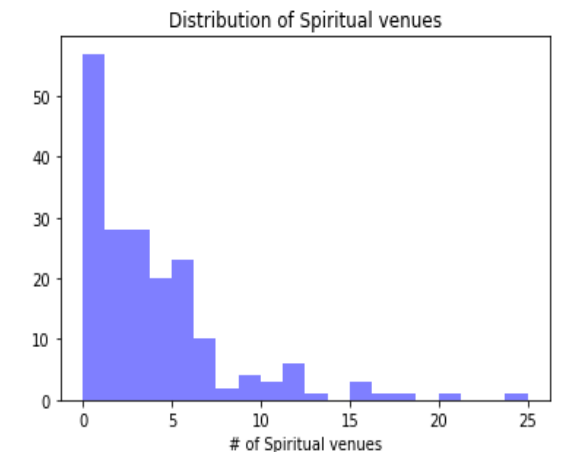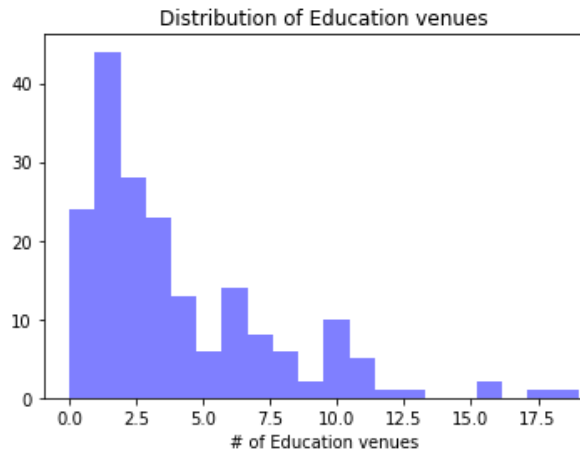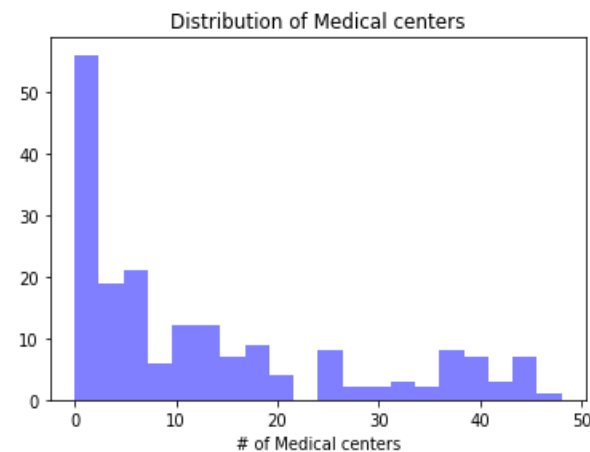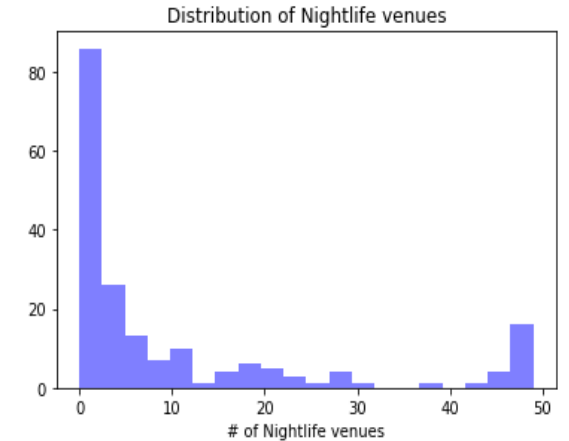# Neighborhood Venue count within a given radius is used to calculate Toronto average
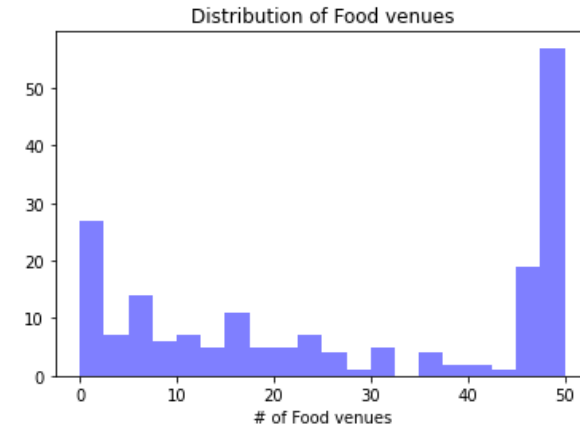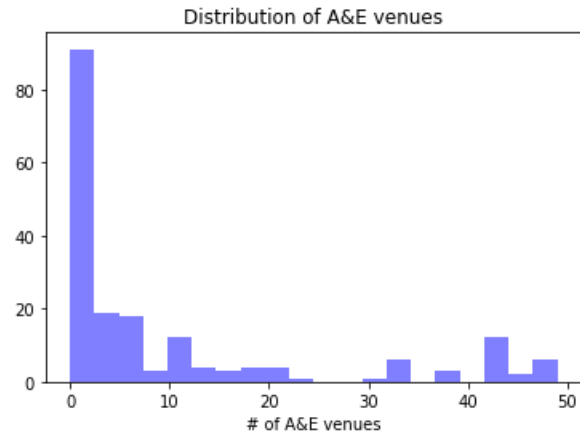
- For our rating requirement, we assume the 200 neighborhoods to be normally distributed samples of Toronto. The metrics we obtain from these samples could be used to calculate mean and standard deviation of the corresponding metrics for Toronto

On plotting distribution of number of venues, we infer skewed distribution

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.


Distribution of A&E venues

Distribution of Food venues

Distribution of Nightlife venues

Distribution of Medical centers

Distribution of Education venues

Distribution of Spiritual venues

Ratings are calculated on a linear scale from "Mean-SD" to "Mean+SD"

Ratings are calculated as follows:

- Range of results are limited from "Mean-SD" to "Mean+SD".
- "Mean-SD" and below are rated as "1"
- "Mean" is rated 3
- "Mean+SD" and above are rated "5"
- Other Intermediate values are interpolated linearly

```python
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)
```

Ratings are generated and provided on a map at the corresponding listing location
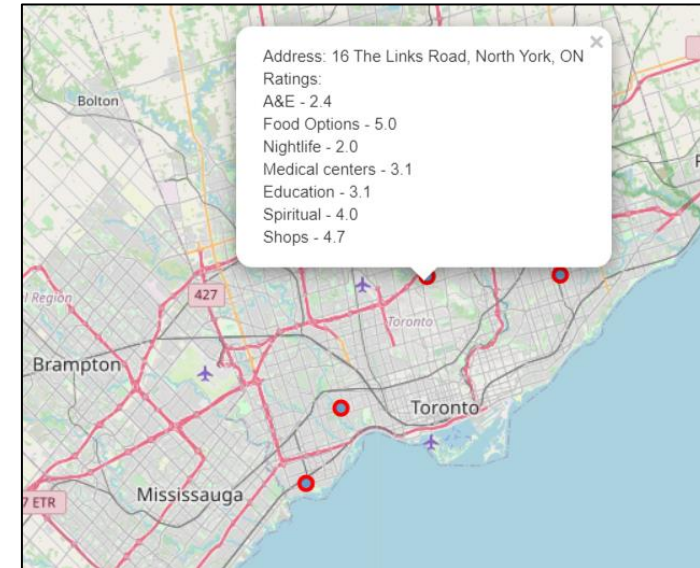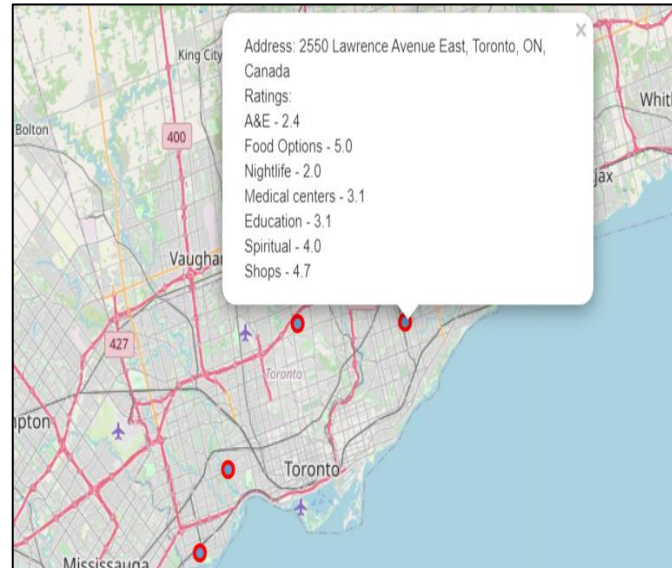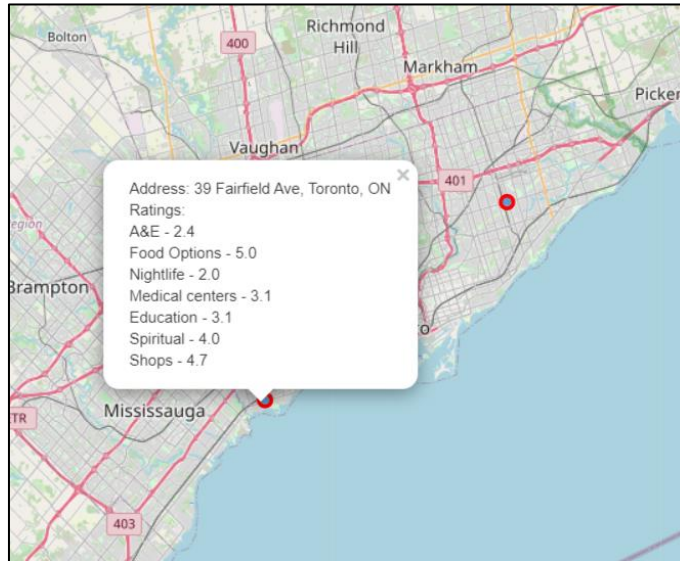
Ratings were calculated for the units we are interested and plotted in a single map, with popups showing ratings and address information.

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.

Ratings gave a baseline comparison of the venue proximity for the rental units

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.

Thank you