Identifying the Similarities and Dissimilarities between Toronto and New York City (NYC) on Food, Arts & Entertainment and Nightlife

Deshraj Singh Kiran

June 28, 2019

Sample Questions the Project seeks to Answer

- 1. How frequent Japanese Restaurants are in the neighborhoods of Toronto as compared to NYC?
- 2. If Museums play an important role in deciding which place to tour, should you choose Toronto or NYC?
- 3. If Art Gallaries was your favourite place to explore in your free time, would it matter much if you chose to visit NYC rather than Toronto?
- 4. What is the most common Nightlife Spots in Toronto and NYC?
- 5. What are top 3 Foods in Toronto and NYC?
- 6. What Food venues NYC has that Toronto does not have?
- 7. What could I do for entertainment in NYC that I can't do in Toronto?
- 8. Which Borough in NYC and Toronto has the highest venues for Arts & Entertainments?
- 9. Which Food type is most disproportionately found in Toronto as compared to NYC?
- 10. What are the bottom 5 Nightlife Spots in NYC and Toronto?
- 11. What are the top 3 most commonly found Food venues in the neighborhood of Toronto and NYC?

1. New York City (NYC) Neighborhood Location Dataset

- Raw GeoJson data from <u>NYU Spatial Data Repository</u> transformed to a csv file with 'Borough',
 'Neighborhood', 'Latitude' and 'Longitude' features
- A total of 306 neighborhoods found

2. Toronto Neighborhood Location Dataset

- Raw postal code data from <u>Wikipedia</u> web scarapped, cleaned and transformed into a csv file with 'PostalCode', 'Borough' and 'Neighborhood' features
- The csv file then merged with the csv file with geographical coordinates of each postal code from http://cocl.us/Geospatial_data
- The final csv file had 'Borough', 'Neighborhood', 'Latitude' and 'Longitude' features
- A total of 209 unique neighborhoods found.

3. Places API of Foursquare

- The explore endpoint API from <u>Foursquare Developer API</u> used in conjunction with the above mentioned NYC and Toronto datasets
- The *categoryld* of the 'Food', 'Arts & Entertainment' and 'Nightlife Spot' venue categories used to filter the returned data from Foursquare for their data exploratory analysis

Sample New York City venues filtered by Food Venue Category

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Ripe Kitchen & Bar	40.898152	-73.838875	Caribbean Restaurant
1	Wakefield	40.894705	-73.847201	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
2	Wakefield	40.894705	-73.847201	Jackie's West Indian Bakery	40.889283	-73.843310	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Jimbo's	40.891740	-73.858226	Burger Joint
4	Wakefield	40.894705	-73.847201	Cooler Runnings Jamaican Restaurant Inc	40.898276	-73.850381	Caribbean Restaurant

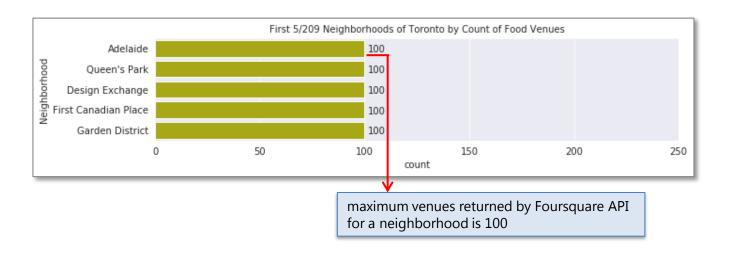
Methodology

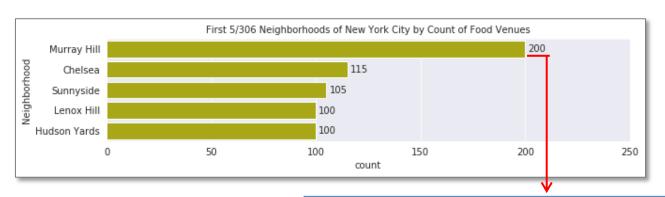
1. Data Preparation

- 2. Exploratory Data Analysis (EDA)
 - Food
 - Arts & Entertainments
 - Nightlife Spot
- 3. K-means Clustering of Neighborhoods

EDA (Food) contd.

Count of Food Venues in Each Neighborhood

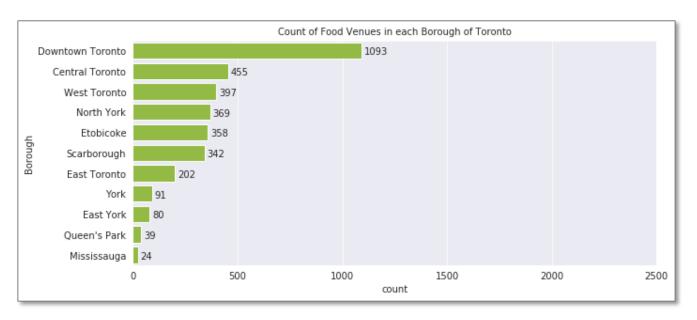


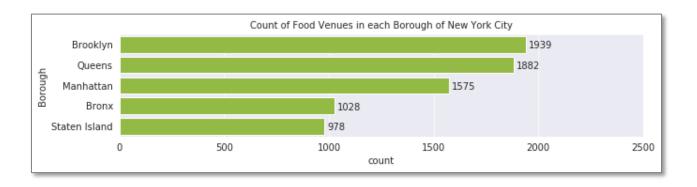


venues from 'Murray Hill' found in both Manhattan and Queens boroughs got added up resulting in more than 100 Food venues for a neighborhood

EDA (Food)

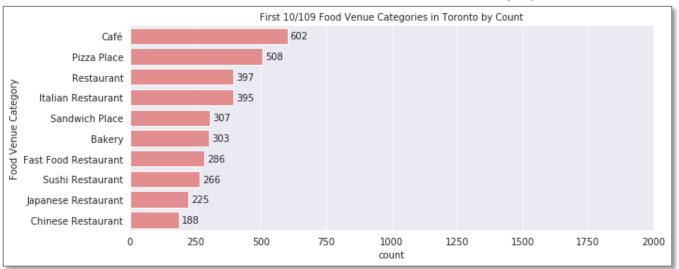
Count of Food Venues in Each Borough

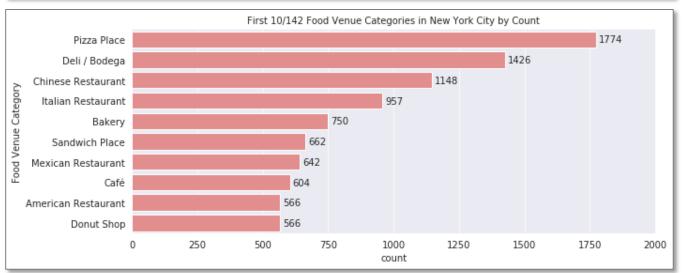




EDA (Food) to be contd.

Count of Food Venues of Each Food Venue Category

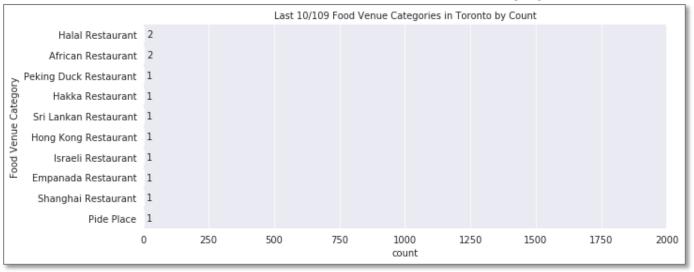


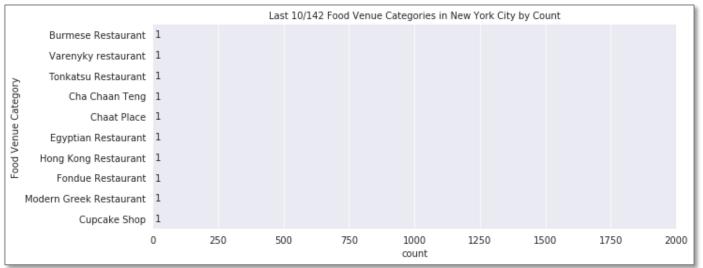


Toronto has 109 unique Food venue categories, while NYC has 142 of them.

EDA (Food) contd.

Count of Food Venues of Each Food Venue Category

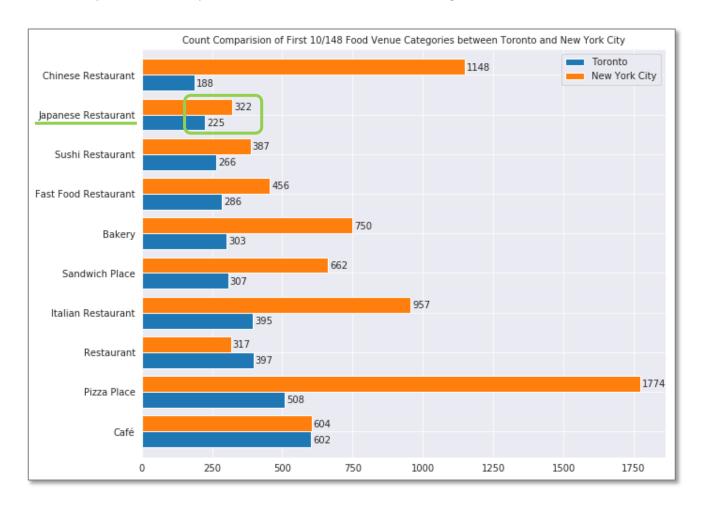




More than Food Venue Categories have a count equal to 1 in both Toronto and NYC. Toronto has just 1 'Sri Lankan Restaurant' while there is just 1 'Burmese Restaurant' in NYC.

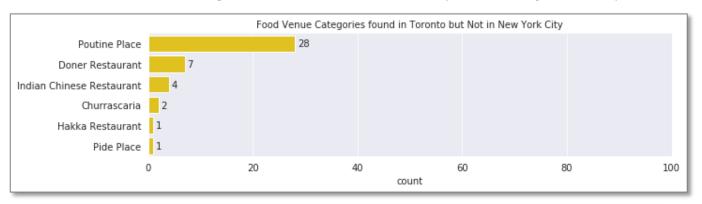
EDA (Food)

Count Comparison of Top 10 out of 148 Food Venue Categories between Toronto and NYC



EDA (Food) to be contd.

Food Venue Categories found in X, but not in Y (Dissimilarity Measure)



A total of 6 Arts & Entertainment venue categories were found in the neighborhoods of Toronto, but not in New York City

EDA (Food) contd.

Food Venue Categories found in X, but not in Y (Dissimilarity Measure)

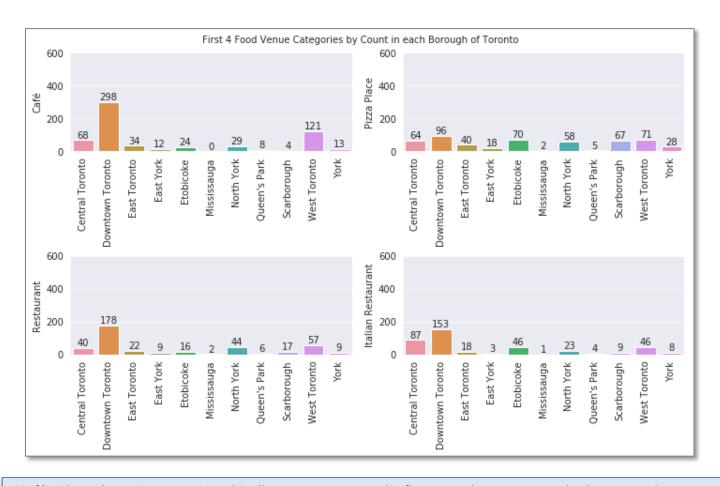
A total of 39 food venue categories were found in the neighborhoods of New York City, but not in Toronto

	Food Venue Category	Count		Food Venue Category	Count
0	Peruvian Restaurant	61	20	Pet Café	4
1	Russian Restaurant	28	21	Swiss Restaurant	4
2	Argentinian Restaurant	26	22	South Indian Restaurant	3
3	Australian Restaurant	16	23	Soba Restaurant	3
4	Szechuan Restaurant	16	24	Molecular Gastronomy Restaurant	2
5	Lebanese Restaurant	12	25	Shabu-Shabu Restaurant	2
6	Japanese Curry Restaurant	11	26	Tex-Mex Restaurant	2
7	Food Stand	11	27	Varenyky restaurant	1
8	Kosher Restaurant	11	28	Romanian Restaurant	1
9	Austrian Restaurant	10	29	Tonkatsu Restaurant	1
10	Paella Restaurant	8	30	Kofte Place	1
11	Himalayan Restaurant	8	31	Egyptian Restaurant	1
12	Caucasian Restaurant	8	32	Cha Chaan Teng	1
13	Kebab Restaurant	8	33	Fondue Restaurant	1
14	Polish Restaurant	7	34	Cupcake Shop	1
15	Scandinavian Restaurant	7	35	Chaat Place	1
16	Czech Restaurant	5	36	Ukrainian Restaurant	1
17	Venezuelan Restaurant	5	37	Burmese Restaurant	1
18	English Restaurant	5	38	Modern Greek Restaurant	1
19	Salvadoran Restaurant	5			

Food Venue Categories found in New York City but Not in Toronto

IBM Applied Data Science Capstone Project

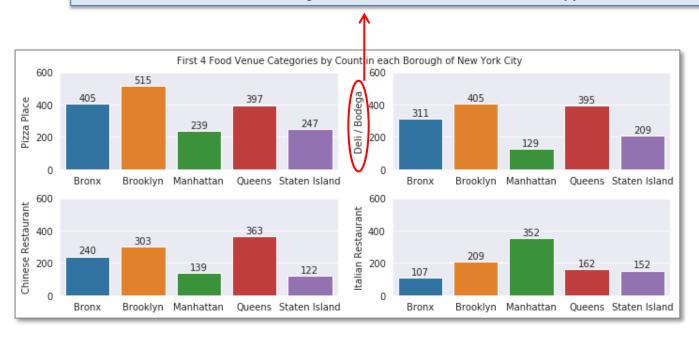
Food Venue Categories by Count in Each Borough



'Café', 'Pizza Place', 'Restaurant' and 'Italian Restaurant' are the first 4 Food venue categories in Toronto by count. 'Downtown Toronto' has the most of them, while 'Mississauga' has the least of them.

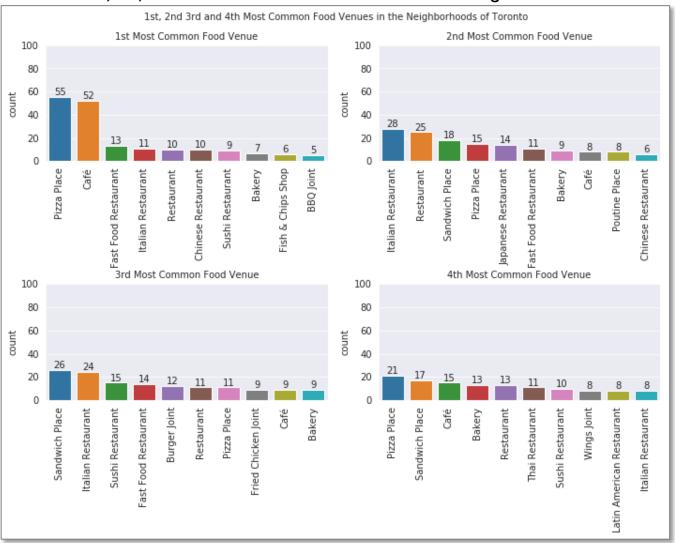
Food Venue Categories by Count in Each Borough

A major dissimilarity or disproportionate Food venue category between NYC and Toronto is 'Deli / Bodega', which has the second highest venue count of 1426 in NYC but only just 128 in Toronto.



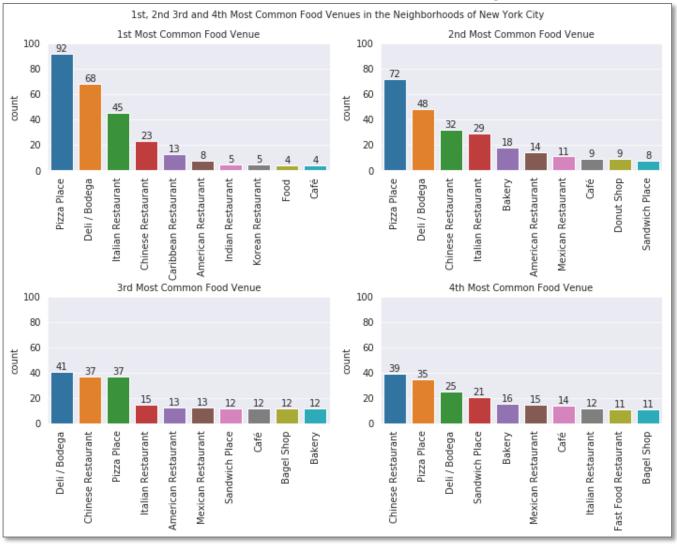
'Pizza Place', 'Deli / Bodega', 'Chinese Restaurant' and 'Italian Restaurant' are the first 4 Food venue categories in NYC by count. 'Brooklyn' has the highest number of 'Pizza Place' Food venues, while 'Manhattan' has the highest count for 'Italian Restaurant'

1st, 2nd, 3rd and 4th Most Common Food Venues in Neighborhoods

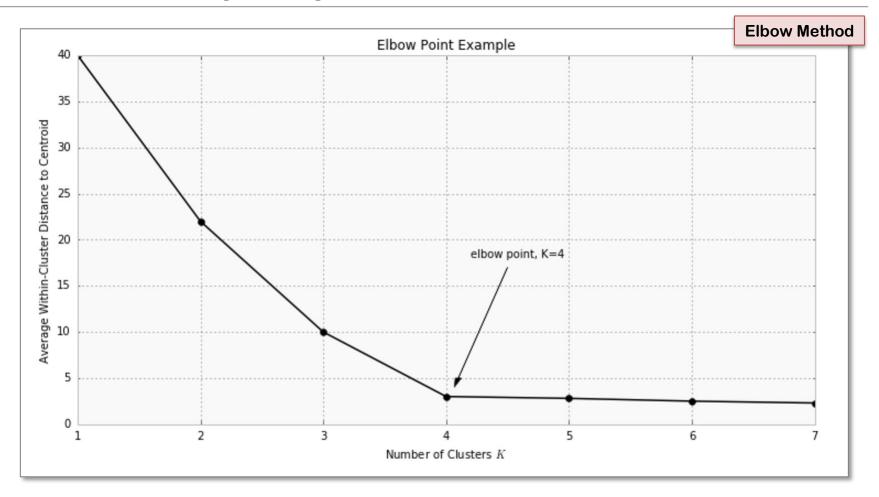


'Pizza Place' beats 'Café' by a very narrow margin to be the 1st most common Food venue in the neighborhoods of Toronto. 'Japanese Restaurant' also find a place among the 2nd most commonly found Food venues in Toronto neighborhoods. EDA (Food) contd.

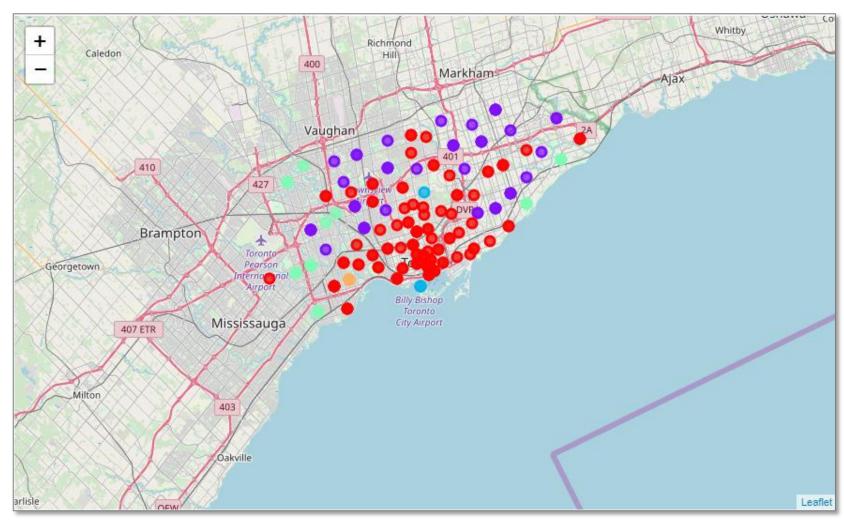
1st, 2nd, 3rd and 4th Most Common Food Venues in Neighborhoods



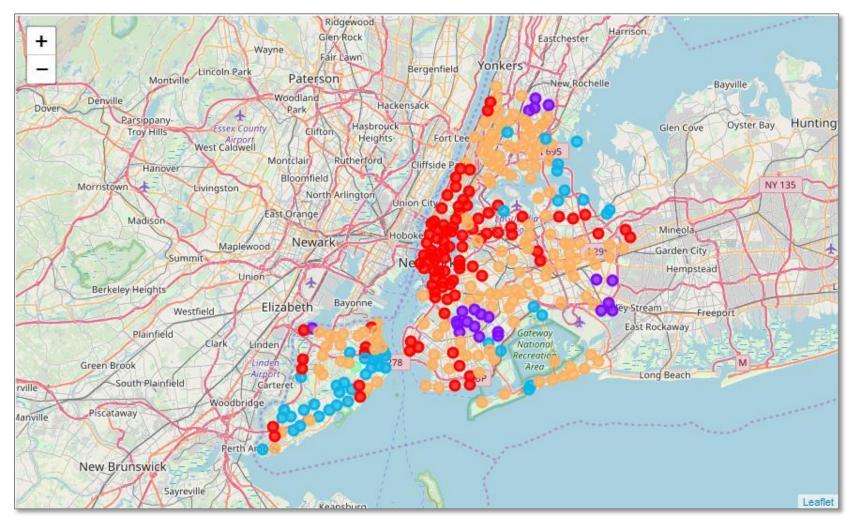
'Pizza Place' to be not just the 1st, but 2nd most common Food venue as well in the neighborhoods of NYC. 'Deli / Bodega' and 'Chinese Restaurant' venues are also very frequent in the NYC neighborhoods.



To find the optimum value of k, we first plotted the **Squared Error (Cost)** for a range of values for **k** and then used the **kneed** project from https://github.com/arvkevi/kneed to automatically find the value of k where the curvature is maximum.



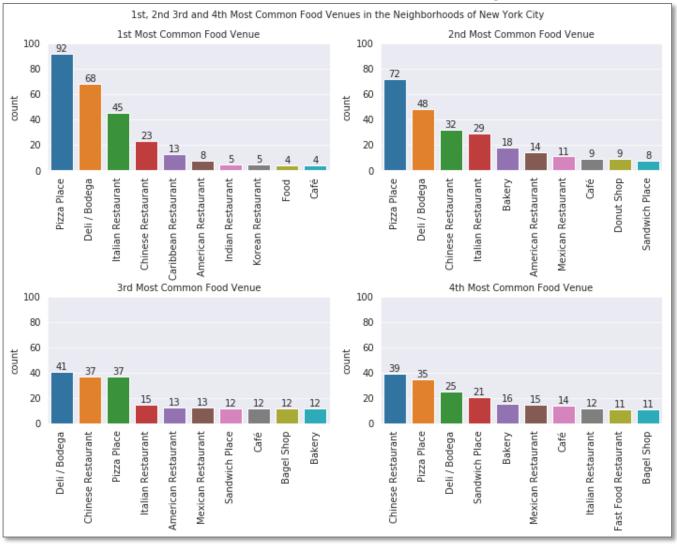
Clustering of neighborhoods in Toronto based on Food Venue Categories (k=5)



Clustering of neighborhoods in NYC based on Food Venue Categories (k=5)

EDA (Food) contd.

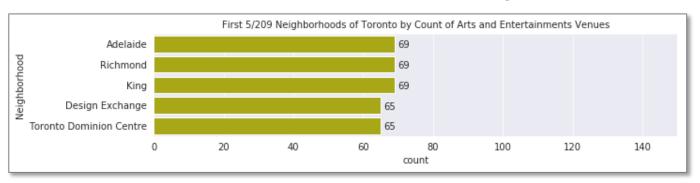
1st, 2nd, 3rd and 4th Most Common Food Venues in Neighborhoods

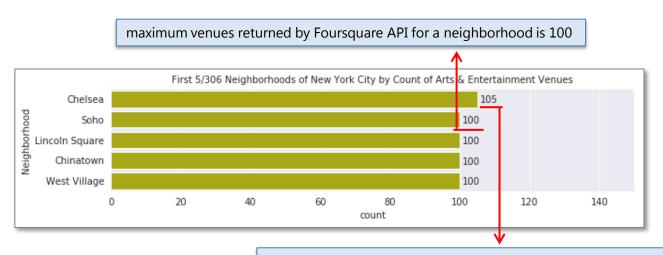


'Pizza Place' to be not just the 1st, but 2nd most common Food venue as well in the neighborhoods of NYC. 'Deli / Bodega' and 'Chinese Restaurant' venues are also very frequent in the NYC neighborhoods.

EDA (Arts & Entertainment)

Count of Arts & Entertainment Venues in Each Neighborhood

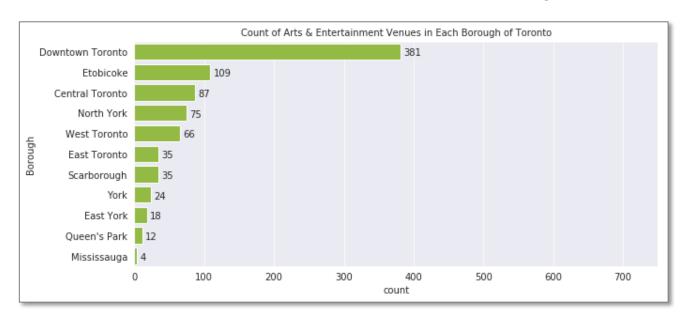


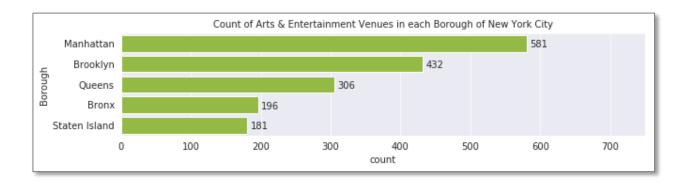


venues from 'Chelsea' found in both Manhattan and Queens boroughs got added up resulting in more than 100 Arts & Entertainment venues for a neighborhood

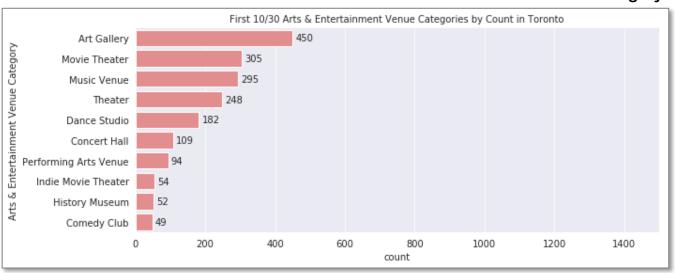
EDA (Arts & Entertainment)

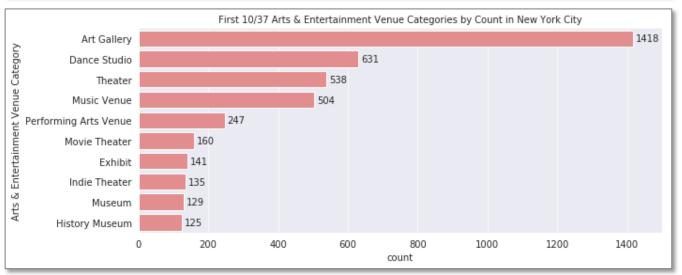
Count of Arts & Entertainment Venues in Each Borough





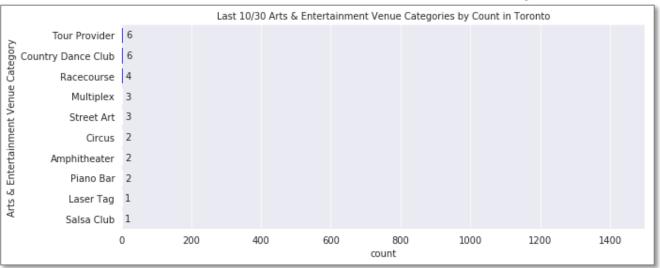
Count of Arts & Entertainment Venues of Each Arts & Entertainment Venue Category

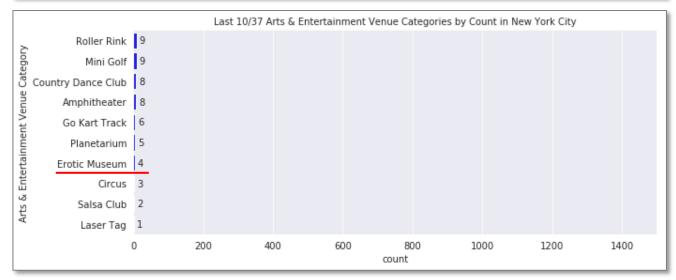




Toronto has 30 unique Arts & Entertainment venue categories, while NYC has 37 of them.

Count of Arts & Entertainment Venues in Each Borough



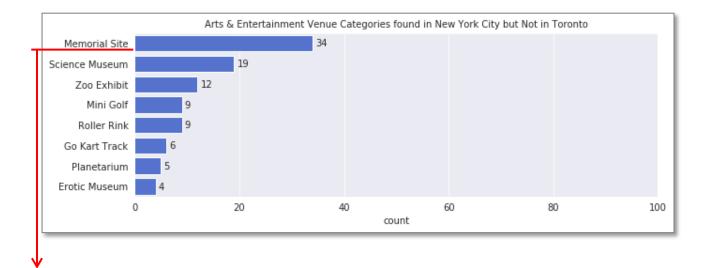


Both Toronto and NYC don't have many 'Salsa Club' venues.

EDA (Arts & Entertainment)

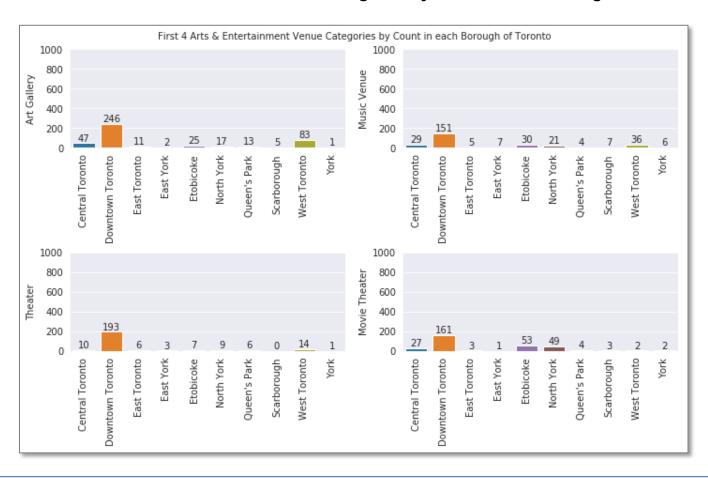
Arts & Entertainment Venue Categories found in X, but not in Y (Dissimilarity Measure)





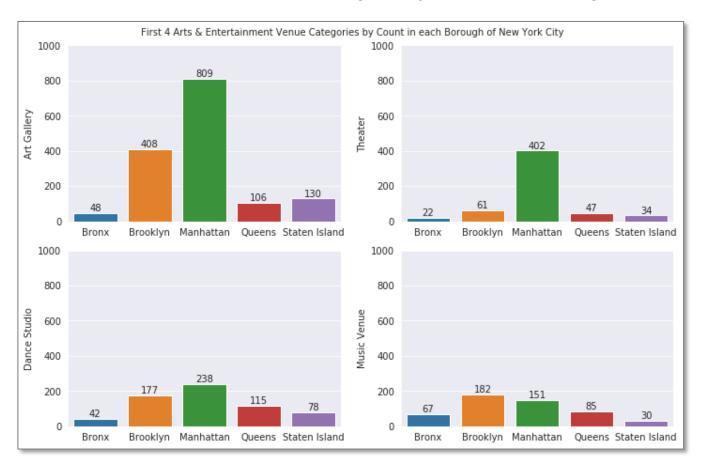
Based on the data gather from Foursquare API Toronto doesn't have any 'Memorial Site' venues, But https://en.wikipedia.org/wiki/Canadian war memorials page from Wikipedia tells us that Toronto has many war 'Mamorial Site' venues.

Arts & Entertainment Venue Categories by Count in Each Borough



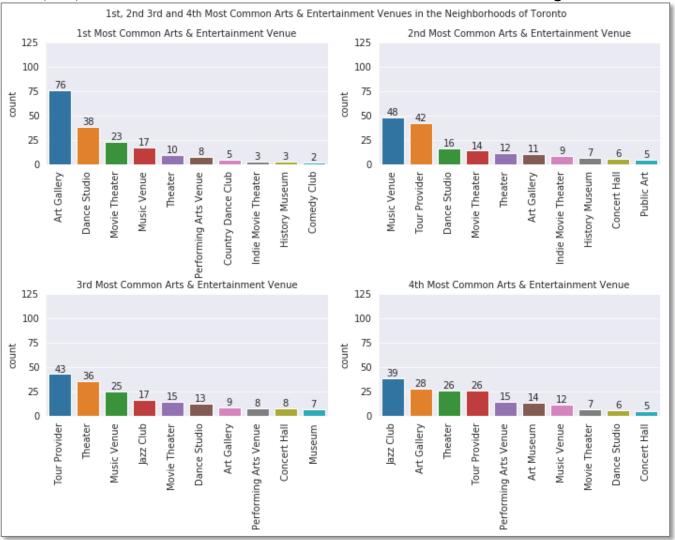
'Art Gallery', 'Movie Theater', 'Music Venue' and 'Theater' were the 4 first 4 Arts & Entertainment venue categories in Toronto by count. 'Downtown Toronto' has the highest number of Arts & Entertainment venues of each of these 4 Food Venue Categories, while others are comparatively quite far behind it

Arts & Entertainment Venue Categories by Count in Each Borough



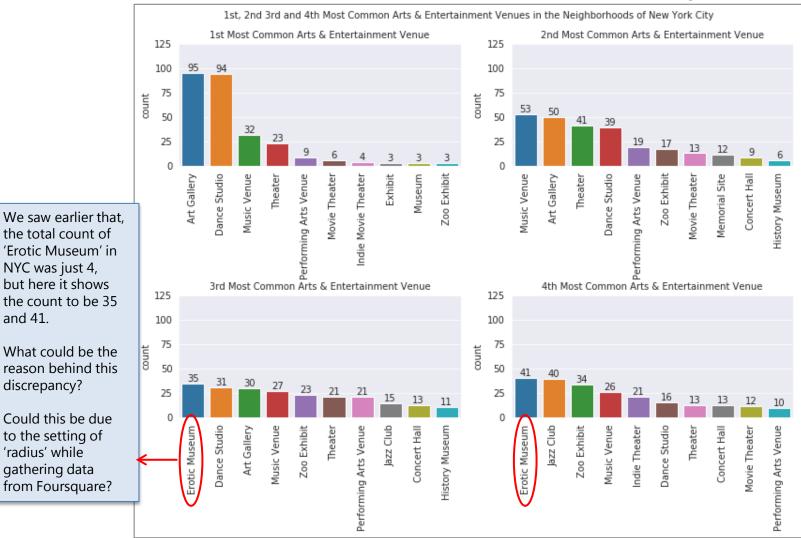
'Art Gallery', 'Dance Studio', 'Theater' and 'Music Venue' were the first 4 Arts & Entertainment venue categories in NYC by count. 'Manhattan' has the highest number of 'Art Gallery', 'Dance Studio', 'Theater' venues, while 'Bronx' has the lowest of these venues'.

1st, 2nd, 3rd and 4th Most Common Arts & Entertainment Venues in Neighborhoods



The 1^{st} most common Arts & Entertainment venue in the neighborhoods of Toronto. 'Music Venue' to be the 2^{nd} most common category. On 3^{rd} we have 'Tour Provider' and on 4^{th} the 'Jazz Club' category

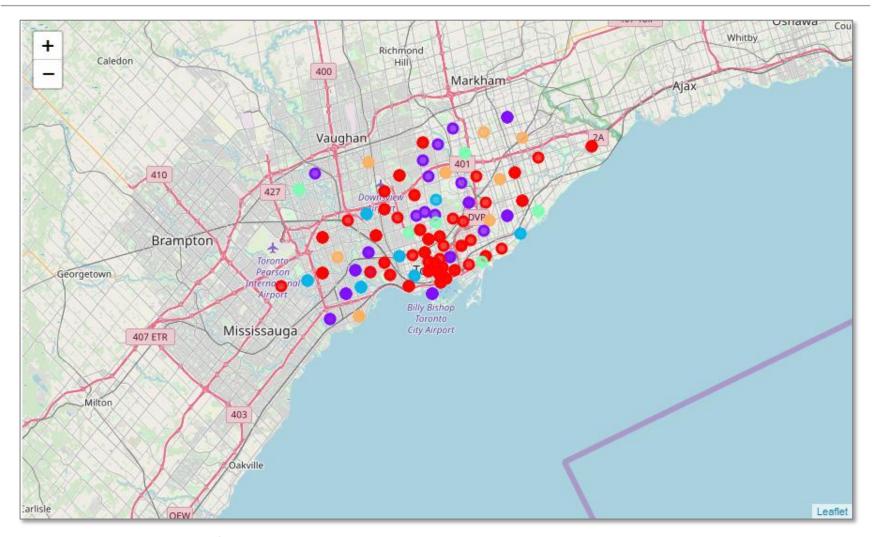
1st, 2nd, 3rd and 4th Most Common Arts & Entertainment Venues in Neighborhoods



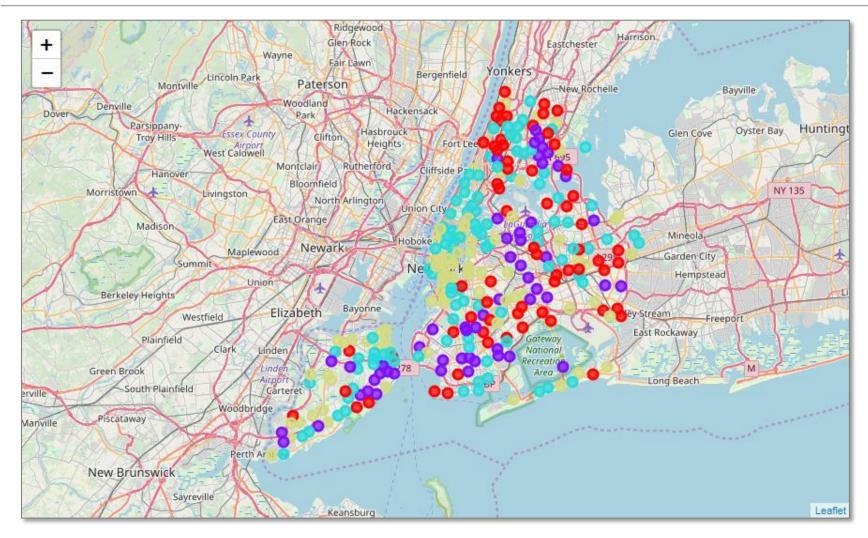
'Art Gallery' and 'Dance Studio' stand neck and neck as the former beats the later just by 1 venue. 'Music Venue' comes 2nd. 'Erotic Museum' seems to be quite well spread out in new York city as it turns out to be the 3rd and 4th most commonly found.

K-means Clustering of Neighborhood (Arts & Entertainment)

to be contd.



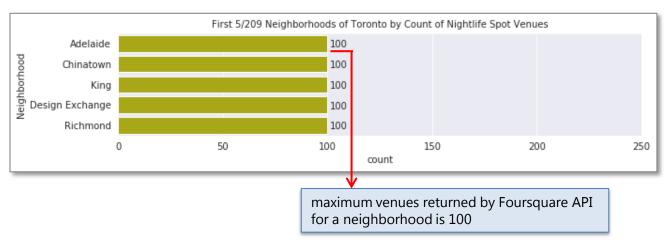
Clustering of neighborhoods in Toronto based on Arts & Entertainment Venue Categories (k=5)

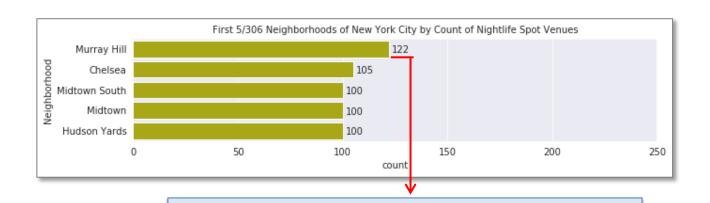


Clustering of neighborhoods in NYC based on Food Venue Categories (k=4)

EDA (Nightlife Spot)

Count of Nightlife Spot Venues in Each Neighborhood





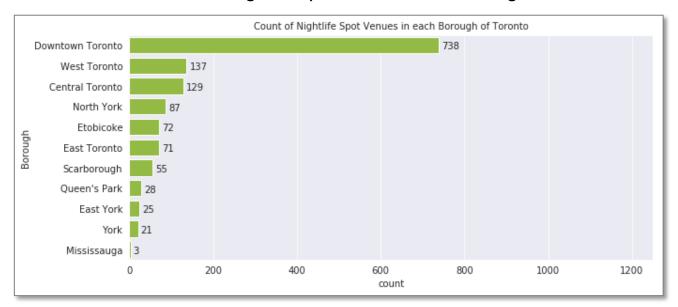
for a neighborhood

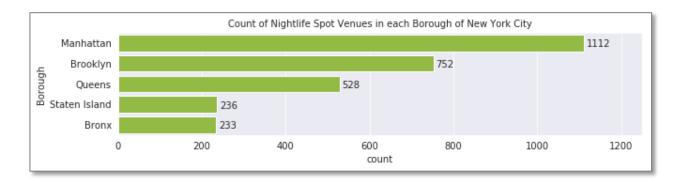
venues from 'Murray Hill' found in both Manhattan and Queens

boroughs got added up resulting in more than 100 Nightlife Spot venues

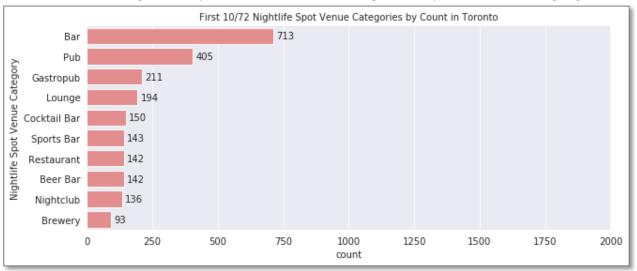
EDA (Nightlife Spot)

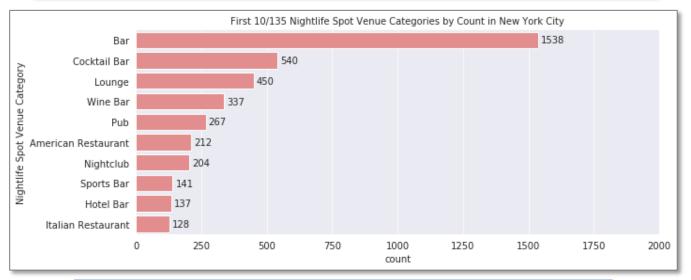
Count of Nightlife Spot Venues in Each Borough





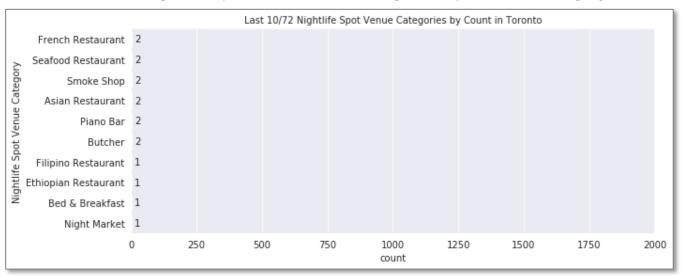
Count of Nightlife Spot Venues of Each Nightlife Spot Venue Category

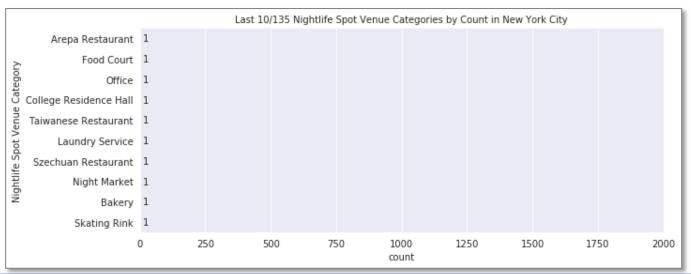




Toronto has 72 unique Nightlife Spot venue categories, while NYC has 135 of them.

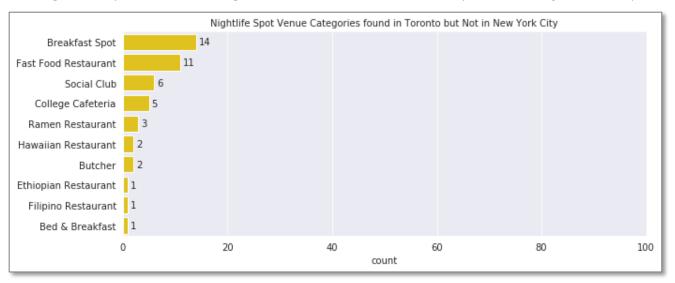
Count of Nightlife Spot Venues of Each Nightlife Spot Venue Category





Toronto had a couple of 'Piano Bar' and 'Night Market' venues, while NYC also has a 'Night Market' and a 'Szechuan Restaurant'

Nightlife Spot Venue Categories found in X, but not in Y (Dissimilarity Measure)



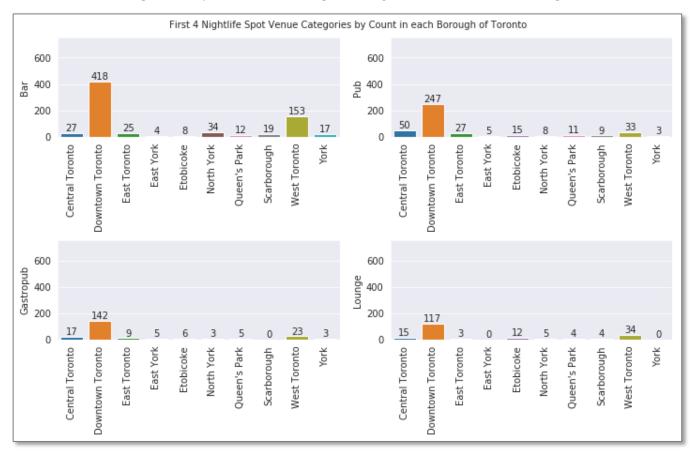
A total of 10 Nightlife Spot venue categories were found in the neighborhoods of Toronto, but not in New York City

Nightlife Spot Venue Categories found in X, but not in Y (Dissimilarity Measure)

ا	Nightlife Spot Venue Category	Count		Nightlife Spot Venue Category	Count		Nightlife Spot Venue Category	Count		Nightlife Spot Venue Category	Count
0	Mexican Restaurant	65	20	Performing Arts Venue	6	40	Food Truck	3	60	Modern European Restaurant	1
1	Caribbean Restaurant	32	21	Tea Room	6	41	Juice Bar	3	61	Bookstore	1
2	Tapas Restaurant	30	22	Fried Chicken Joint	5	42	Ice Cream Shop	3	62	Indie Theater	1
3	Latin American Restaurant	26	23	Cajun / Creole Restaurant	5	43	Nail Salon	2	63	Bakery	1
4	German Restaurant	25	24	Australian Restaurant	5	44	Street Food Gathering	2	64	Pool	1
5	Roof Deck	22	25	Lebanese Restaurant	5	45	Tech Startup	2	65	Market	1
6	Wine Shop	18	26	Turkish Restaurant	5	46	Russian Restaurant	2	66	Scenic Lookout	:
7	Other Nightlife	18	27	Brazilian Restaurant	5	47	Paella Restaurant	2	67	Laundry Service	:
8	Rock Club	17	28	African Restaurant	4	48	Residential Building (Apartment / Condo)	2	68	Building	:
9	Tiki Bar	15	29	Chinese Restaurant	4	49	Middle Eastern Restaurant	2	69	Arepa Restaurant	:
10	Beer Store	11	30	Coworking Space	4	50	Theater	2	70	Professional & Other Places	:
11	Indian Restaurant	11	31	Distillery	3	51	Colombian Restaurant	2	71	Food & Drink Shop	:
12	Southern / Soul Food Restaurant	10	32	Greek Restaurant	3	52	Snack Place	2	72	Taiwanese Restaurant	:
13	Sandwich Place	9	33	Hostel	3	53	Indie Movie Theater	2			
14	Taco Place	8	34	Salon / Barbershop	3	54	Vegetarian / Vegan Restaurant	2			
15	English Restaurant	8	35	Gourmet Shop	3	55	Szechuan Restaurant	1			
16	Vietnamese Restaurant	8	36	Cheese Shop	3	56	Skating Rink	1			
17	South American Restaurant	8	37	Art Gallery	3	57	Food Court	1			
18	Cuban Restaurant	8	38	Harbor / Marina	3	58	Eastern European Restaurant	1			
19	Spanish Restaurant	6	39	General Entertainment	3	59	College Residence Hall	1			

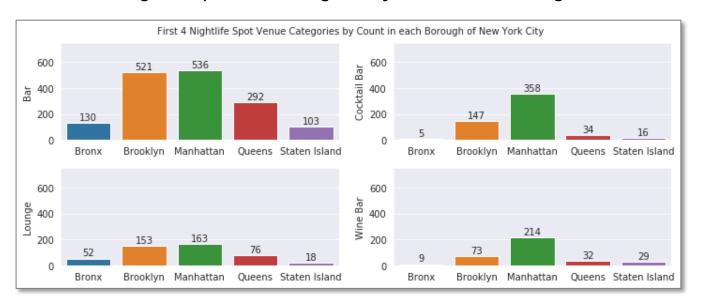
A total of 73 Nightlife Spot venue categories were found in the neighborhoods of NYC, but not in Toronto

Nightlife Spot Venue Categories by Count in Each Borough



'Bar', 'Pub', 'Gastropub' and 'Lounge' were the 4 first 4 Arts & Entertainment venue categories in Toronto by count. 'Downtown Toronto' has the highest number of Nightlife Spot venues of each of these 4 Food Venue Categories, while others are comparatively very far behind it.

Nightlife Spot Venue Categories by Count in Each Borough

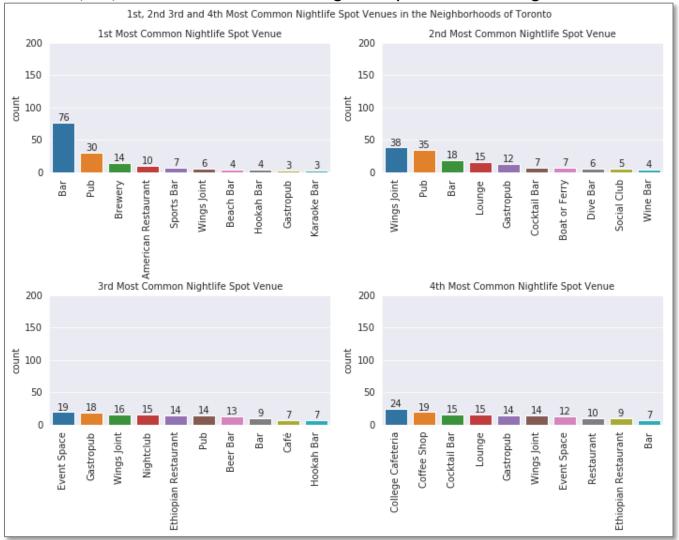


'Bar', 'Cocktail Bar', 'Longue' and 'Wine Bar' were the first 4 Nightlife Spot venue categories in New York City by count.

"Manhattan' has the highest number of all of these 4 venues, while 'Bronx' is pretty close to it when it comes to 'Bar' and 'Lounge'.

'Staten Island' has the minimum number of these 4 Nightlife Spot venue categories among all the 5 boroughs.

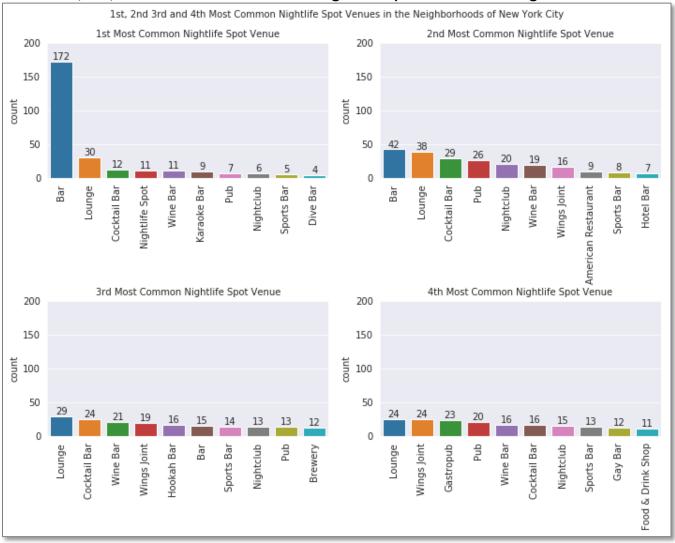
1st, 2nd, 3rd and 4th Most Common Nightlife Spot Venues in Neighborhoods



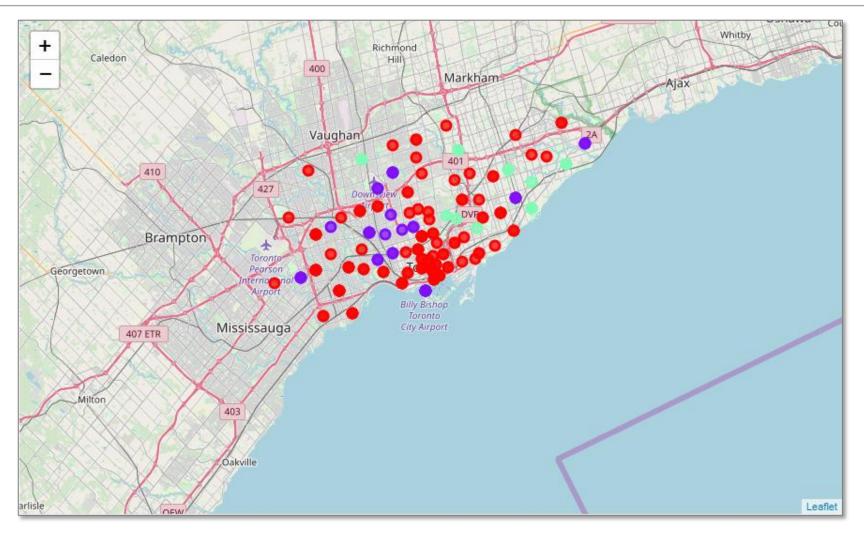
'Bar' is the 1st most common Nightlife Spot venue in the neighborhoods of Toronto. 'Wings Joint' is the 2nd most common category.

On 3rd we have 'Event Space' and 'Gastropub' pretty close to one another and then we have 'College Cafeteria' on 4th.

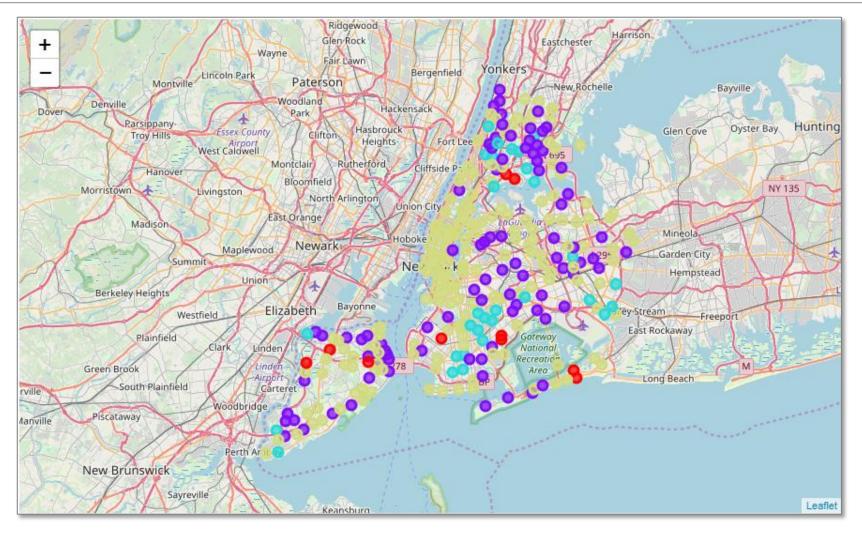
1st, 2nd, 3rd and 4th Most Common Nightlife Spot Venues in Neighborhoods



'Bar' is the 1st most common venue by a long margin than any other Nightlife Spot venue category. 'Lounge' is also quite commonly found in the neighborhoods of New York City.



Clustering of neighborhoods in Toronto based on Nightlife Spot Venue Categories (k=4)



Clustering of neighborhoods in NYC based on Food Venue Categories (k=4)

Food

Dissimilarities

- 1. One of the major dissimilarity between Toronto and New York was that Toronto had 28 of 'Poutine Place' venues while New York City had none of them. Toronto had a total of 6 food venue categories were not found in New York City.
- 2. On the other hand, New York City was found to have 61 of 'Peruvian Restaurant', 28 of 'Russian Restaurant' and 26 of 'Argentinian Restaurant' as the top 3 out of 39 food venue categories that were not found in Toronto.
- 3. Another major dissimilarity between NYC and Toronto is regarding 'Deli / Bodega', which has the second highest venue count in NYC. NYC had 1426 of them while Toronto had just 128 of them.

Similarities

- 1. 5 of the first 10 Food venue categories by count in both Toronto and NYC are common. These include, 'Pizza Place', 'Café', 'Italian Restaurant', 'Bakery' and 'Sandwich Place'. They're all high in numbers. It is a good indication that Based on Food venue categories, Toronto and NYC are more similar than others.
- 2. Out of first 4, 2 of the Food venue categories by count are common in Toronto and NYC. These are, 'Pizza Place' and 'Italian Restaurant'.
- 3. In face the 1st most common Food venue in both Toronto and NYC is 'Pizza Place'. Just like 'Café' is almost the 1st most common Food venue in Toronto, 'Deli / Bodega' is also frequently found in NYC.
- 4. The first 4 Food venue categories by count in Toronto are, 'Café', 'Pizza Place', 'Restaurant' and 'Italian Restaurant'. For NYC, these are, 'Pizza Place', 'Deli / Bodega', 'Chinese Restaurant' and 'Italian Restaurant'.
- 5. When comparing the counts of all the Food venue category between Toronto and New York City, it could be easily seen that Toronto and NYC are more similar to NYC than dissimilar.

Arts & Entertainment

Dissimilarities

- 1. With Arts & Entertainment, Toronto was dissimilar to New York City with just 1 venue category. It was Racecourse. Toronto had 4 of them, while New York City had none.
- 2. On the other hand, the one major dissimilarity between New York City and Toronto was caused by 'Memorial Site' venues. New York City had 34 of them while Toronto had none. Also, New York City also had 4 'Erotic Museum' venues while Toronto had none. Toronto also didn't have any 'Science Museum' and 'Planetarium' was NYC had 19 and 5 of them.
- 3. Although in our data exploration, we got to know that Toronto had no 'Memorial Site' venues, https://en.wikipedia.org/wiki/Canadian_war_memorials page from Wikipedia told us otherwise as it showed us that Toronto in-fact had many war 'Mamorial Site' venues.

Similarities

- 1. Just like Food, when it comes to Arts & Entertainment, both Toronto and NYC has 3 out of 4 venues in common based on their counts. These are 'Art Gallery', 'Theater' and 'Music venue'.
- 2. If this wasn't enough to reflect how similar they are, out of first 10 by count, 7 of venues are common, and if you consider 'Indie Movie Theater' and 'Indie Theater' same, then it make 8 out of first 10 venue categories to be common between the 2 cities. These common venue categories are, 'Art Gallery', 'Movie Theater', 'Music Venue', 'Theater', 'Dance Studio', 'Performing Arts Venue' and 'History Museum.'
- 3. When it comes to first 4 most commonly found venues in Toronto, these include, 'Art Gallery', 'Music Venue', 'Tour Provider' and 'Jazz Club'. For NYC, these are, 'Art Gallery', 'Music Venue', 'Erotic Venue' and even on the 4th position as well we have 'Erotic Venue'.

Nightlife Spot

Dissimilarities

- 1. Nightlife Spot venues get even more interesting when Toronto has 14 of 'Breakfast Spot' venues while NYC has none. Also, there are no 'Ramen Restaurant', 'Hawaiian Restaurant', 'Ethiopian Restaurant' and 'Filipino Restaurant' venues that are open as a Nightlife Spot type venues in NYC. A total of 10 categories of these venues are not found in NYC to make it dissimilar to Toronto.
- 2. On the other hand, NYC has a total of 39 Nightlife Spot venue categories that are not found in Toronto. Almost all of them are restaurants. 65 of 'Mexican Restaurant' to 6 of 'Spanish Restaurant', while there are 16 of 'Rock Club' and 6 of 'Performing Arts Venue' as well.

Similarities

- 1. When it comes to Nightlife Spot both NYC and Toronto are very similar. They're full of 'Bar', 'Pub', 'Lounge', 'Nightclub' and different kind of restaurant venues.
- 2. Out of first 10 by count, 5 of venues are common between the 2 cities. These common venue categories are, 'Bar, 'Pub', 'Lounge', 'Cocktail Bar', 'Sports Bar', and 'Nightclub'.
- 3. Out of first 4 common venues, 'Bar' and 'Pub' are the 1st and 2nd most common Nightlife Spot venues in both Toronto, while in NYC, these are 'Bar' and 'Lounge'.

Discussions

Limitations

- 1. Foursquare returns maximum of 100 venues per neighborhood
- 2. As we keep increasing the radius setting, a neighbourhood might start overlapping with its neighbour neighbourhood resulting in an increase in the total number of venues in a neighbourhood. The default setting by Foursquare is a radius of 250 meters.
- 3. The exact number of venues for a neighbourhood is next to impossible to find as *radius* is going to be fixed when gathering data regardless of the population densities of the neighborhoods in Toronto or NYC.

Data Ambiguity

- 1. We have a 'Restaurant' Venue Category who is among the 2nd, 3rd and 4th most common food venues with a total count of 397 in Toronto. It seems to be a generic restaurant where most of the different kinds of cuisines are served.
- 2. In 'Food' category there is a venue category called 'Food'. It is not clear what type of Food is served here, but NYC has 241 of these while Toronto has 12.

Incorrect Data

1. While exploring data to find the dissimilarities between Toronto and NYC, it turned out that Toronto didn't have any 'Memorial Site' venues. We were quite surprised just with the thought of it. So we did a web search, and it didn't take much time to find out that Toronto had in-fact many war 'Memorial Site' venues, according to https://en.wikipedia.org/wiki/Canadian war memorials Wikipedia page.

Conclusion

- 1. This project and the Notebook it comes with serves as a reference to identify the similarities and dissimilarities between Toronto and New York City on Food, Arts & Entertainment and Nightlife. It can be extended to include data from many other categories provided in the Foursquare Places API.
- 2. As expected, more venue categories are found in NYC than Toronto. This could be attributed to the fact that people from all walks of life from different countries have been coming to US for a better life, so from the culture's perspective it is more diverse than Toronto.
- 3. Data provided by the Foursquare is not perfect and needs to be worked upon. For example, In 'Food' category there is a venue category called 'Food'. It is not clear what type of Food is served here, but NYC has 241 of these while Toronto has 12. Many such venue categories are found in the datasets provided by Foursquare for all the 3 categories explored in this project that need to be made better sense of.

Future Directions

- 1. It would be interesting to gather, explore data and analyse it with a different setting of radius and try to find difference in the outcomes compared this project.
- 2. If findings of this project could be supplemented with the dataset from a Food ordering web portal, https://www.zomato.com for example, it would add additional value to this project and its outcomes.
- 3. Incorporating population along with area of neighbourhoods can aid additional analysis and visualization to complement this project.