

Comparative neighborhood venue analysis between Downtown Toronto and Manhattan

Data-driven recommendations for Downtown Toronto urban development

1. Introduction

The future of cities depends among many other things on their capability to attract people that will help shape their future. The population of a city needs to reflect diversity of people that ensures that the city will strive economically, intellectually and socially.

People with diverse backgrounds, capabilities and ambitions will chose different neighborhoods as preferred place to live. A city that wants to attract people from another city needs to provide neighborhoods with a similar or better quality of life as the neighborhoods of that other city.

Quality of Life is a vast field and different people will consider different aspects as the most relevant. When asking city residents, why they live in their city and not in cheaper and calmer rural areas, one of most frequent answers is 'I love living in this city, because only here I have all the venues, restaurants, museums shops nearby.' In this project we focus on the availability of such venues in urban neighborhoods.

2. Business Problem

In this project we aim to support you, the urban development team of the neighborhoods of Downtown Toronto, in deciding which venues you should attract in order to become attractive to people who consider Manhattan in New York attractive.

We have read that the population of Downtown Toronto is expected to grow from about 250,000 in 2016 to about 475,000 people in 2041. With that growth the geographic population density of Downtown Toronto will become increasingly similar to that of Manhattan. As all basic infrastructure is in place, we consider the quality of life in Downtown Toronto an essential building block in attracting the right mix of people, talent and companies to ensure a sustainable economical, intellectual and social growth for Downtown Toronto.

The project is understood as part of a larger assessment of factors that determine the perceived quality of life in Downtown Toronto. The comparative quantitative analysis of venues in Downtown Toronto with those in Manhattan will provide insights, which complement insights from other projects to inform a comprehensive development strategy for Downtown Toronto.

This quantitative assessment relies on location data retrieved from Foursquare. This data shows the geographic distribution of venues like restaurants, shops, museums and many other.

3. Data sources and acquisition

For this project various three basic data categories will be used:

1. General geospatial data
2. Venue data
3. Mapping data

General geospatial data
New York

The general geospatial data provides us with geographic coordinates for Downtown Toronto and Manhattan. We use the Novatim Python library from geopy.geocoder to get coordinates of the Toronto and Manhattan.

The list of neighborhoods of New York, including Manhattan, has been provided by Coursera and IBM: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json

This json file includes the coordinates of each neighborhood.

Toronto

The list of neighborhoods of Toronto, including Downtown Toronto, is scraped from Wikipedia: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada_M

This list does not include the geographic coordinates. These are retrieved from http://cocl.us/Geospatial_data

The mapping of coordinates to neighborhoods is done through the postal codes included in both data sets.

Venue data

Venue data is retrieved for both cities from FourSquare: https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={} (client_id and client_secret are confidential)

The limit is set by default to 100. The radius is set to 500m. So we retrieve for each neighborhood the first 100 venues within 500m radius from geographic coordinates of the neighborhood. In case that FourSquare knows fewer than 100 venues for a neighborhood, only this smaller number venues will be downloaded. During later normalization it will be ensured that shares of venue categories among all venues in a neighborhood are correctly calculated for all neighborhoods.

Mapping data

For both cities maps are created using the Folium library, which provides us with detailed maps and tools to visualize neighborhoods and venues.

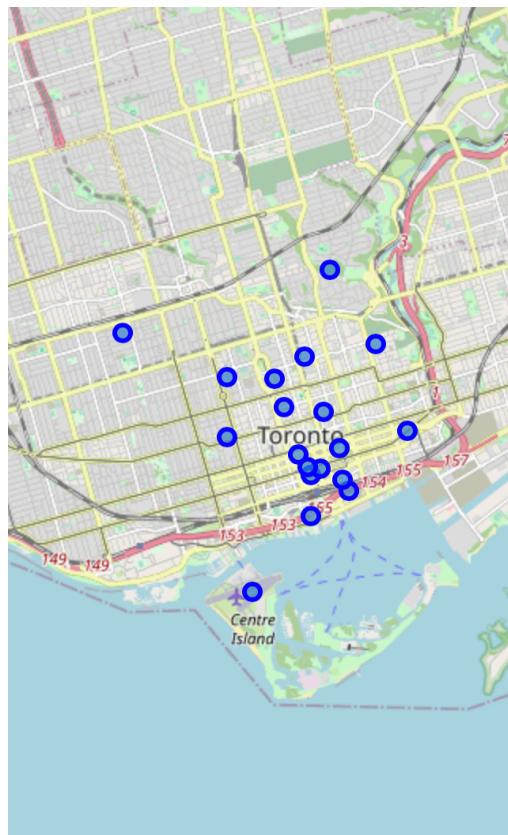
4. Methodology

We will follow the following methodology.

1. Retrieving and cleaning the publicly available geospatial data and venue data from FourSquare
2. Exploring the data, including map visualizations
3. Cluster analysis using a combined dataset of neighborhoods and venues in Downtown Toronto and Manhattan
4. Discussion of the results the cluster analysis and identification of a field of special interest for further statistical analysis
5. Statistical analysis of this field of special interest
6. Discussion of final results
7. Conclusions and recommendations

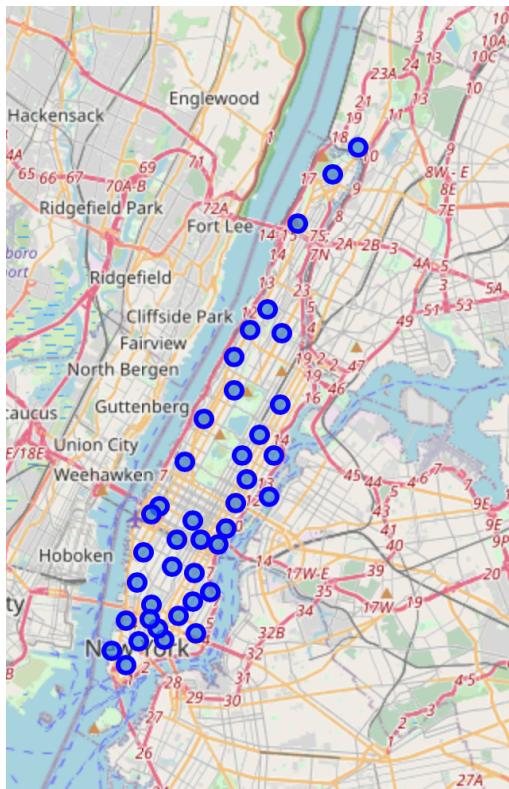
5. Data exploration, analysis and results

5.1 Exploring the neighborhoods



Downtown Toronto consists of 19 neighborhoods.

In total we could retrieve 1253 different venues of 214 different categories in these neighborhoods.



Manhattan consists of 40 neighborhoods.
In total we could retrieve 3219 different venues
of 323 different categories in these
neighborhoods.

5.2 Merging the data for a joint cluster analysis

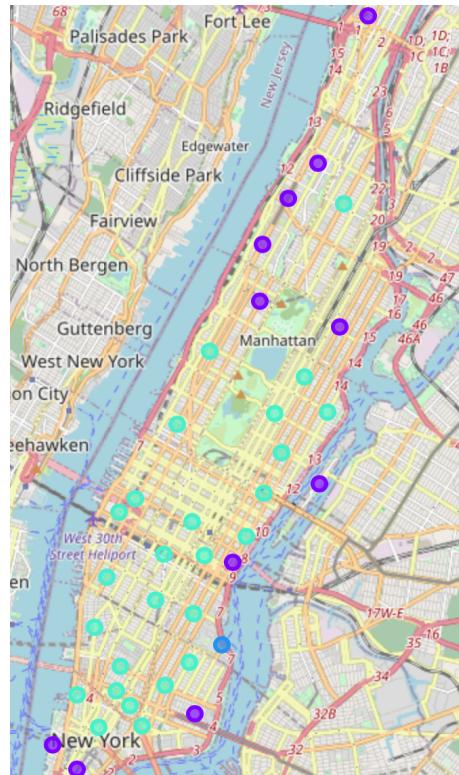
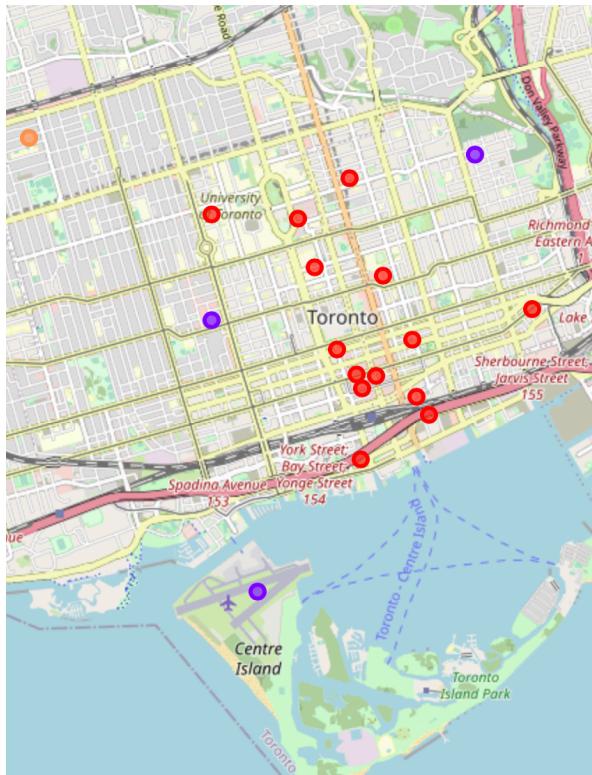
After merging the data frames of Downtown Toronto and Manhattan, we have:

- 59 neighborhoods
 - 360 different venue categories

We have replaced NaN fields with zeros.

5.3 k-means clustering

We have then run k-means clustering with different values of k. Chosing k = 6 has resulted in the clearest clustering pictures:



The two maps reveal that neighborhoods in central areas of Manhattan and Toronto downtown fall into distinct clusters: In Manhattan they fall all into cluster 4 (turkois color). In downtown Toronto they fall all into cluster 1 (red color). This major difference is what we will analyze next in more detail.

These central areas are surrounded in both Manhattan and downtown Toronto by neighborhoods falling the same cluster 2 (purple color). This fact proves that our methodology is correctly clustering by venue category.

The remaining clusters 3, 5 and 6 contain only one neighborhood each. Cluster 3 shows that Stuyvesant Town is distinct from other neighborhoods in Manhattan or downtown Toronto. The Toronto map shows that clusters 5 and 6 are neighborhoods quite far from the the center of downtown and consequently it doesn't come at a surprise to see that these two neighborhoods are not falling into clusters common with other downtown Toronto or Manhattan neighborhoods.

In the following we dive deeper into the differences between the clusters 1 (center of downtown Toronto) and 4 (center of Manhattan). This will help us to identify strategies that could help Toronto attract people and talent from Manhattan.

5.4 Discussion of clustering results of clusters 1 and 4

Let us here have a look at the cluster 1 (the neighborhoods represented in red in above map):

Borough	Neighborhood	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
2	Downtown Toronto	Regent Park, Harbourfront	Coffee Shop	Park	Bakery	Pub	Breakfast Spot	Café	Theater	Shoe Store	Brewery
4	Downtown Toronto	Queen's Park, Ontario Provincial Government	Coffee Shop	Yoga Studio	Bank	Smoothie Shop	Beer Bar	Sandwich Place	Restaurant	Café	Portuguese Restaurant
9	Downtown Toronto	Garden District, Ryerson	Coffee Shop	Clothing Store	Café	Bubble Tea Shop	Cosmetics Shop	Japanese Restaurant	Bookstore	Diner	Pizza Place
15	Downtown Toronto	St. James Town	Coffee Shop	Café	Cocktail Bar	Restaurant	Gastropub	American Restaurant	Beer Bar	Clothing Store	Gym
20	Downtown Toronto	Berczy Park	Coffee Shop	Bakery	Restaurant	Farmers Market	Cocktail Bar	Seafood Restaurant	Beer Bar	Cheese Shop	Park
24	Downtown Toronto	Central Bay Street	Coffee Shop	Café	Sandwich Place	Italian Restaurant	Salad Place	Japanese Restaurant	Burger Joint	Thai Restaurant	Department Store
30	Downtown Toronto	Richmond, Adelaide, King	Coffee Shop	Café	Restaurant	Gym	Hotel	Bar	Clothing Store	Thai Restaurant	Salad Place
36	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	Coffee Shop	Aquarium	Café	Hotel	Fried Chicken Joint	Brewery	Pizza Place	Scenic Lookout	Restaurant
42	Downtown Toronto	Toronto Dominion Centre, Design Exchange	Coffee Shop	Hotel	Café	Restaurant	Japanese Restaurant	American Restaurant	Salad Place	Seafood Restaurant	Bar
48	Downtown Toronto	Commerce Court, Victoria Hotel	Coffee Shop	Restaurant	Hotel	Café	Gym	Deli / Bodega	Japanese Restaurant	Seafood Restaurant	American Restaurant
80	Downtown Toronto	University of Toronto, Harbord	Café	Bookstore	Sandwich Place	Bar	Japanese Restaurant	Bakery	Yoga Studio	Italian Restaurant	Beer Bar
92	Downtown Toronto	Stn A PO Boxes	Coffee Shop	Italian Restaurant	Pub	Restaurant	Café	Beer Bar	Seafood Restaurant	Japanese Restaurant	Hotel
97	Downtown Toronto	First Canadian Place, Underground city	Coffee Shop	Café	Hotel	Restaurant	Gym	Japanese Restaurant	Salad Place	Seafood Restaurant	Asian Restaurant
99	Downtown Toronto	Church and Wellesley	Coffee Shop	Gay Bar	Sushi Restaurant	Japanese Restaurant	Restaurant	Café	Hotel	Mediterranean Restaurant	Yoga Studio
											Men's Store

It is obvious that the Coffee Shops are dominating as 1st most common venue in all these neighborhoods. Note that this cluster was not filtered by 'Downtown Toronto', rather the k-means algorithm provided this list, exclusively made up of neighborhoods in the central part of Downtown Toronto.

Let us now do the same for cluster 4 (the neighborhoods represented in turquoise color in above map):

	Borough	Neighborhood	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
1	Manhattan	Chinatown	Chinese Restaurant	Cocktail Bar	Dessert Shop	Bakery	American Restaurant	Salon / Barbershop	Spa	Noodle House	Optical Shop	Hotpot Restaurant
6	Manhattan	Central Harlem	African Restaurant	Bar	French Restaurant	American Restaurant	Chinese Restaurant	Seafood Restaurant	Cosmetics Shop	Cafeteria	Grocery Store	Library
8	Manhattan	Upper East Side	Exhibit	Italian Restaurant	Coffee Shop	Bakery	Gym / Fitness Center	Yoga Studio	Cosmetics Shop	French Restaurant	Juice Bar	Spa
9	Manhattan	Yorkville	Italian Restaurant	Coffee Shop	Gym	Bar	Deli / Bodega	Sushi Restaurant	Japanese Restaurant	Wine Shop	Diner	Mexican Restaurant
10	Manhattan	Lenox Hill	Italian Restaurant	Sushi Restaurant	Pizza Place	Coffee Shop	Cocktail Bar	Gym	Gym / Fitness Center	Café	Burger Joint	Salon / Barbershop
12	Manhattan	Upper West Side	Italian Restaurant	Café	Bar	Indian Restaurant	Wine Bar	Pizza Place	Bakery	Coffee Shop	Vegetarian / Vegan Restaurant	Thai Restaurant
13	Manhattan	Lincoln Square	Plaza	Café	Concert Hall	Performing Arts Venue	Theater	Italian Restaurant	Gym / Fitness Center	French Restaurant	Indie Movie Theater	Wine Shop
14	Manhattan	Clinton	Theater	American Restaurant	Gym / Fitness Center	Coffee Shop	Cocktail Bar	Sandwich Place	Gym	Hotel	Italian Restaurant	Spa
15	Manhattan	Midtown	Hotel	Coffee Shop	Bakery	Clothing Store	Theater	Steakhouse	Sporting Goods Shop	Sushi Restaurant	Bookstore	Pizza Place
16	Manhattan	Murray Hill	Japanese Restaurant	Coffee Shop	Hotel	Gym / Fitness Center	Sandwich Place	American Restaurant	Bar	Restaurant	Pizza Place	Italian Restaurant
17	Manhattan	Chelsea	Coffee Shop	Art Gallery	Bakery	American Restaurant	Ice Cream Shop	Italian Restaurant	Japanese Restaurant	Park	Cycle Studio	Cupcake Shop
18	Manhattan	Greenwich Village	Italian Restaurant	Sushi Restaurant	Clothing Store	Café	Indian Restaurant	American Restaurant	Gym	Boutique	Bubble Tea Shop	Chinese Restaurant
19	Manhattan	East Village	Bar	Ice Cream Shop	Mexican Restaurant	Pizza Place	Wine Bar	Korean Restaurant	Speakeasy	Italian Restaurant	Vegetarian / Vegan Restaurant	Coffee Shop
21	Manhattan	Tribeca	American Restaurant	Park	Wine Bar	Italian Restaurant	Café	Greek Restaurant	Spa	Coffee Shop	Skate Park	Steakhouse
22	Manhattan	Little Italy	Café	Bakery	Italian Restaurant	Bubble Tea Shop	Chinese Restaurant	Cocktail Bar	Mediterranean Restaurant	Ice Cream Shop	Sandwich Place	Coffee Shop
23	Manhattan	Soho	Clothing Store	Italian Restaurant	Coffee Shop	Boutique	Mediterranean Restaurant	Shoe Store	Bakery	Café	French Restaurant	Pizza Place
24	Manhattan	West Village	Italian Restaurant	American Restaurant	New American Restaurant	Wine Bar	Cocktail Bar	Park	Theater	Jazz Club	Coffee Shop	Sandwich Place
27	Manhattan	Gramercy	Bar	Italian Restaurant	Bagel Shop	Thai Restaurant	American Restaurant	Pizza Place	Coffee Shop	Ice Cream Shop	Diner	Thrift / Vintage Store
30	Manhattan	Carnegie Hill	Coffee Shop	Café	Bookstore	Italian Restaurant	Gym / Fitness Center	Gym	French Restaurant	Yoga Studio	Wine Shop	Vietnamese Restaurant
31	Manhattan	Noho	Italian Restaurant	Hotel	Mexican Restaurant	Yoga Studio	Pizza Place	Coffee Shop	Cocktail Bar	Bookstore	French Restaurant	Art Gallery
32	Manhattan	Civic Center	Coffee Shop	Gym / Fitness Center	Spa	Cocktail Bar	Hotel	French Restaurant	Yoga Studio	American Restaurant	Sushi Restaurant	Sandwich Place
33	Manhattan	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Cosmetics Shop	Burger Joint	Coffee Shop	Dessert Shop	Gym / Fitness Center	Bakery	American Restaurant
34	Manhattan	Sutton Place	Italian Restaurant	Gym / Fitness Center	Furniture / Home Store	Pizza Place	Coffee Shop	Park	Gym	Bar	Thai Restaurant	Beer Bar
35	Manhattan	Turtle Bay	Coffee Shop	Italian Restaurant	Sushi Restaurant	Japanese Restaurant	Park	Seafood Restaurant	Ramen Restaurant	Deli / Bodega	Turkish Restaurant	Karaoke Bar
38	Manhattan	Flatiron	Italian Restaurant	Japanese Restaurant	New American Restaurant	Cycle Studio	Gym	Gym / Fitness Center	Mediterranean Restaurant	Sporting Goods Shop	Spa	American Restaurant
39	Manhattan	Hudson Yards	Hotel	Italian Restaurant	Gym / Fitness Center	American Restaurant	Coffee Shop	Café	Boat or Ferry	Nightclub	Thai Restaurant	Gym

It is obvious that the Restaurants of all different sorts of international cuisines are among the most common venues in all the neighborhoods of this cluster. Note that this cluster was not filtered by ‚Manhattan‘, rather the k-means algorithm provided this list, exclusively made up of neighborhoods in the central part of Manhattan.

5.5 Analysis of the absolute share of restaurants and cafes among all venues in clusters 1 and 4

We suspect that there might be some distortion created by the ‚winner takes all‘ effect of describing clusters by ‚10 most common venue categories‘. Therefore let’s return to absolute numbers, i.e. let’s analyze the absolute numbers of venues per category.

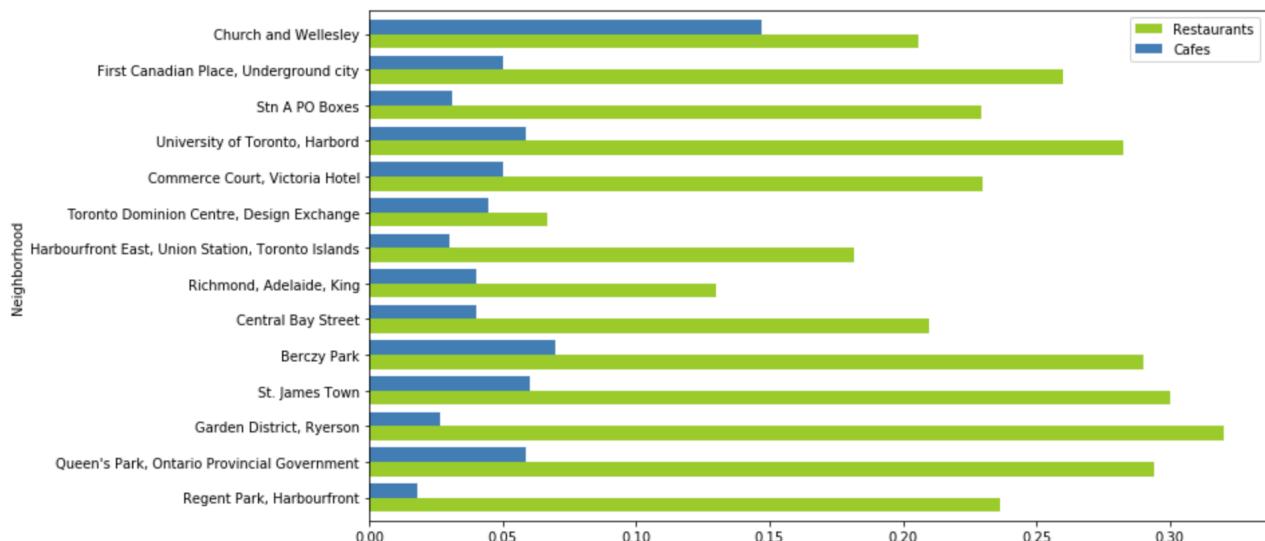
We use one-hot encoding for creating venue summaries per neighborhood.

For comparability of neighborhoods with 100 retrieved venues and neighborhoods with lower venue counts, we normalize the data to obtain the share of restaurants and cafes among all the venues of a neighborhood.

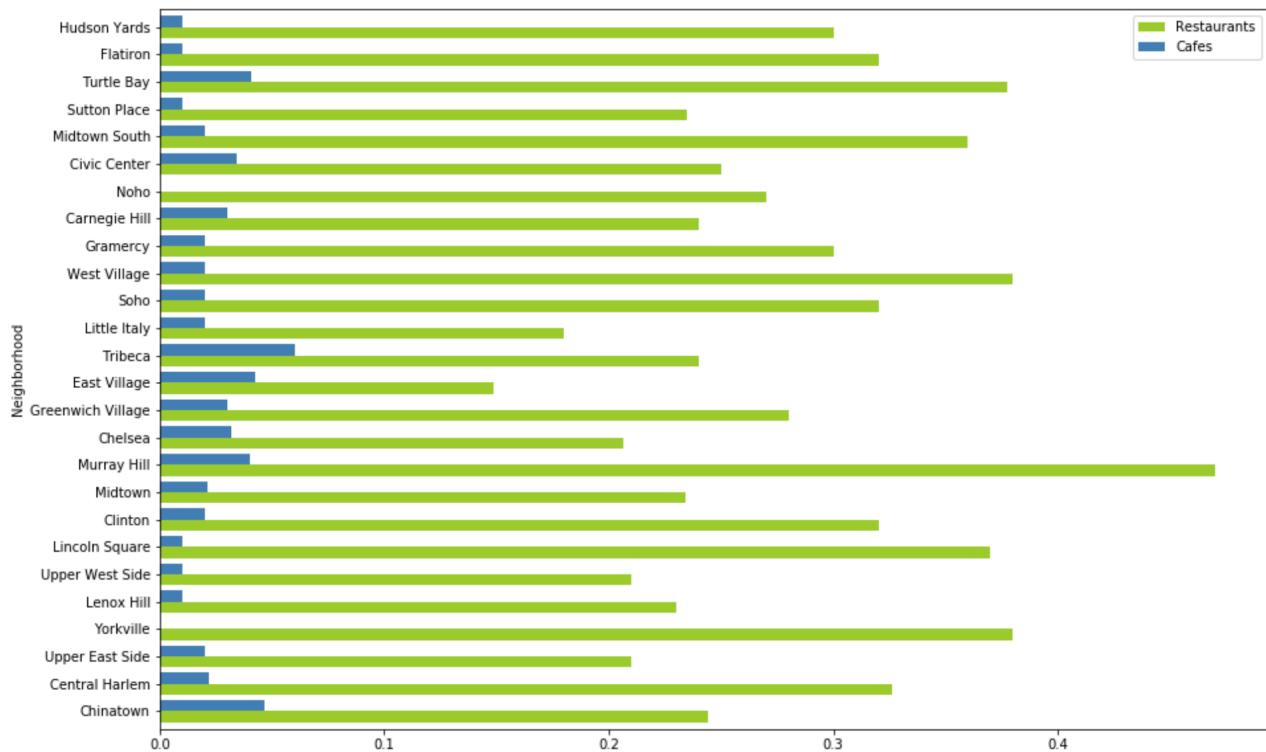
We define ‚restaurant‘ as aggregate count over all venue categories containing ‚restaurant‘ in the category name.

Likewise we define ‚cafe‘ as aggregate count over all venue categories containing ‚coffee shop‘ or ‚café‘ or ‚cafeteria‘.

Here the resulting data for cluster 1:



and for cluster 4:



To quantify the big picture across all the neighborhoods of the two clusters, here the statistical summaries and boxplots. We will discuss these in the final discussion chapter.

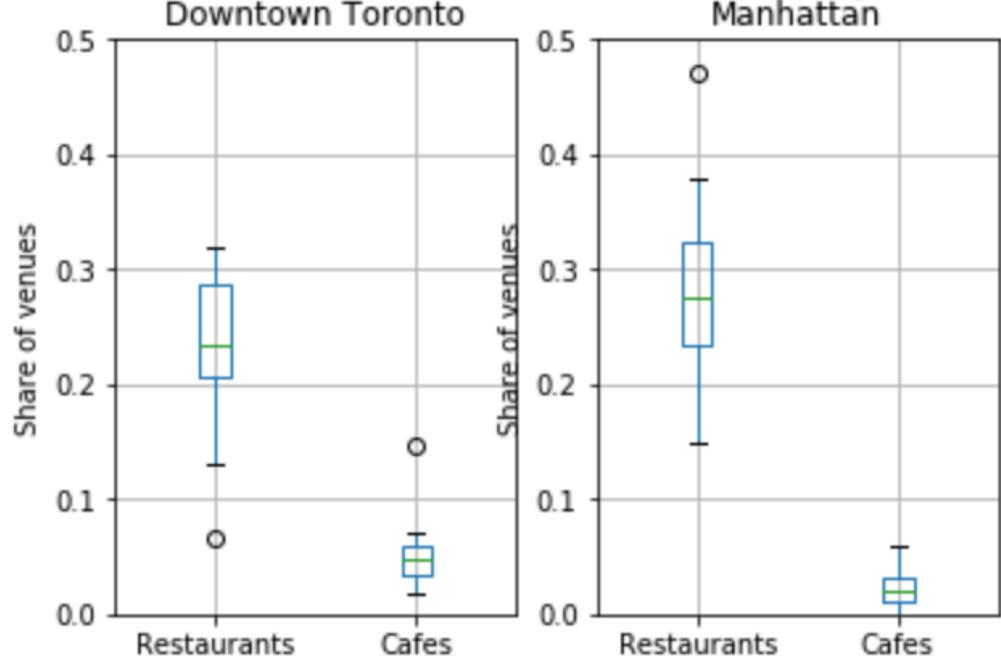
Cluster 1:

Cluster 4:

	Restaurants	Cafes
count	14.000000	14.000000
mean	0.231169	0.051825
std	0.070359	0.031079
min	0.066667	0.018182
25%	0.206912	0.033438
50%	0.233182	0.047222
75%	0.288088	0.058824
max	0.320000	0.147059

	Restaurants	Cafes
count	26.000000	26.000000
mean	0.284686	0.023036
std	0.075318	0.014618
min	0.148936	0.000000
25%	0.234205	0.010051
50%	0.275000	0.020000
75%	0.324565	0.031310
max	0.470000	0.060000

Distribution of restaurants and cafes in central neighborhoods

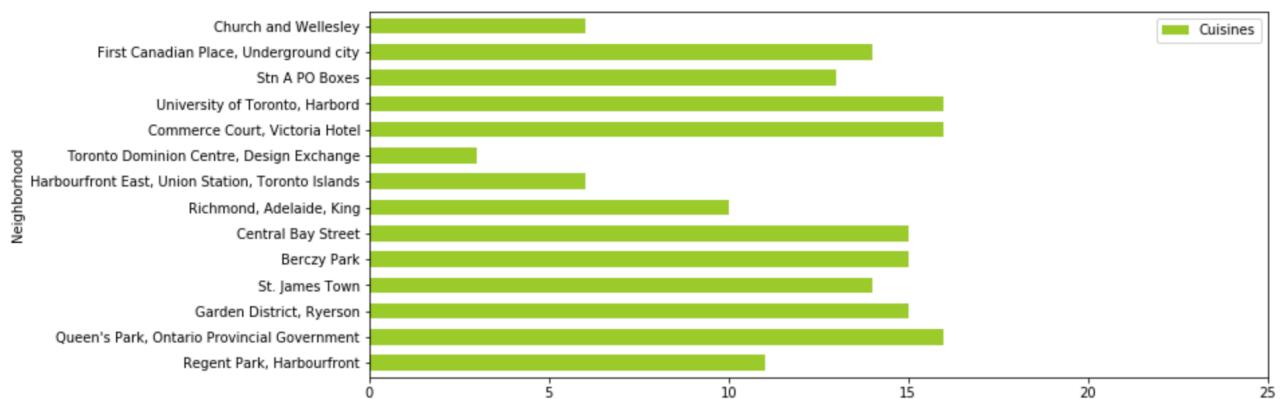


5.6 Analysis of number of international cuisines represented in the restaurants of clusters 1 and 4

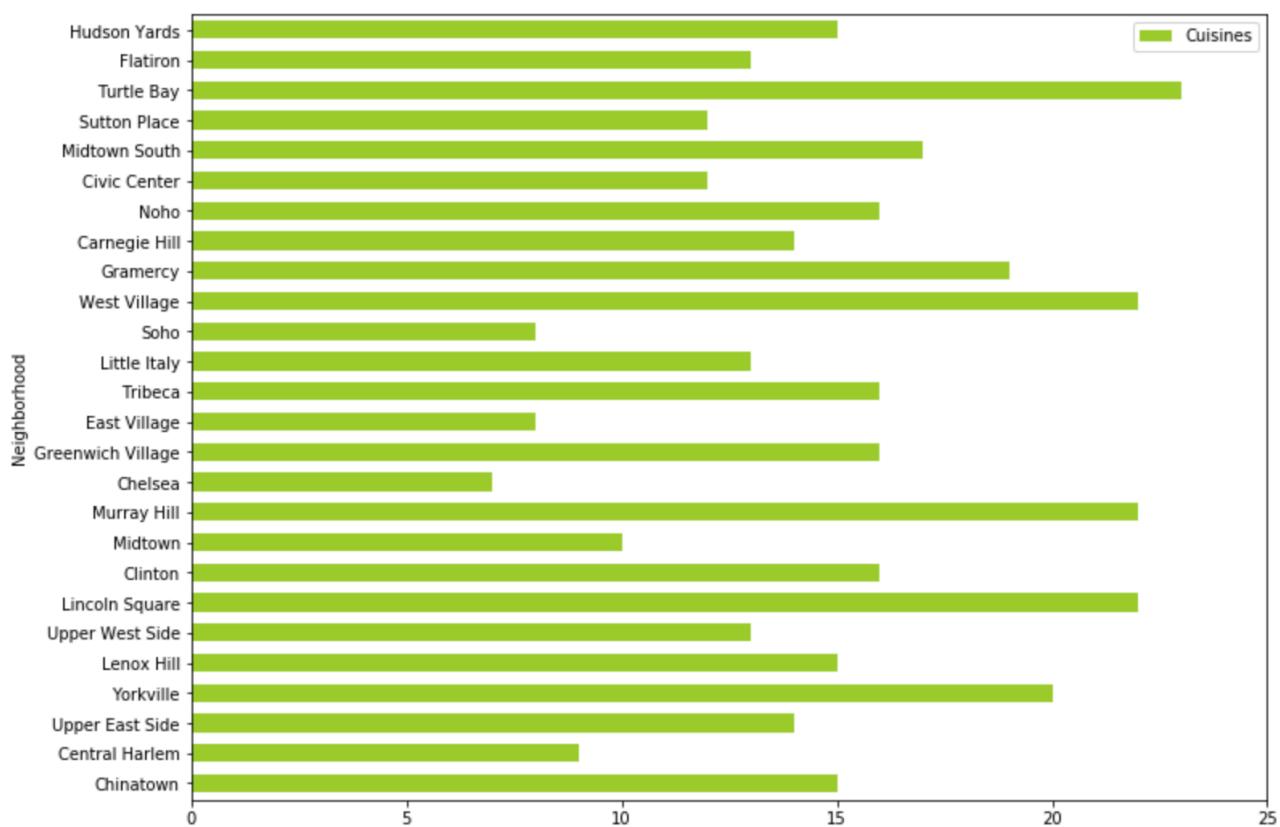
Let us now look at another metric, that might be more relevant to describe the variety of choice of restaurants in the neighborhoods.

For that we sum up the number of different cuisines represented in each of the neighborhoods of the two clusters.

Cluster 1:



Cluster 4:



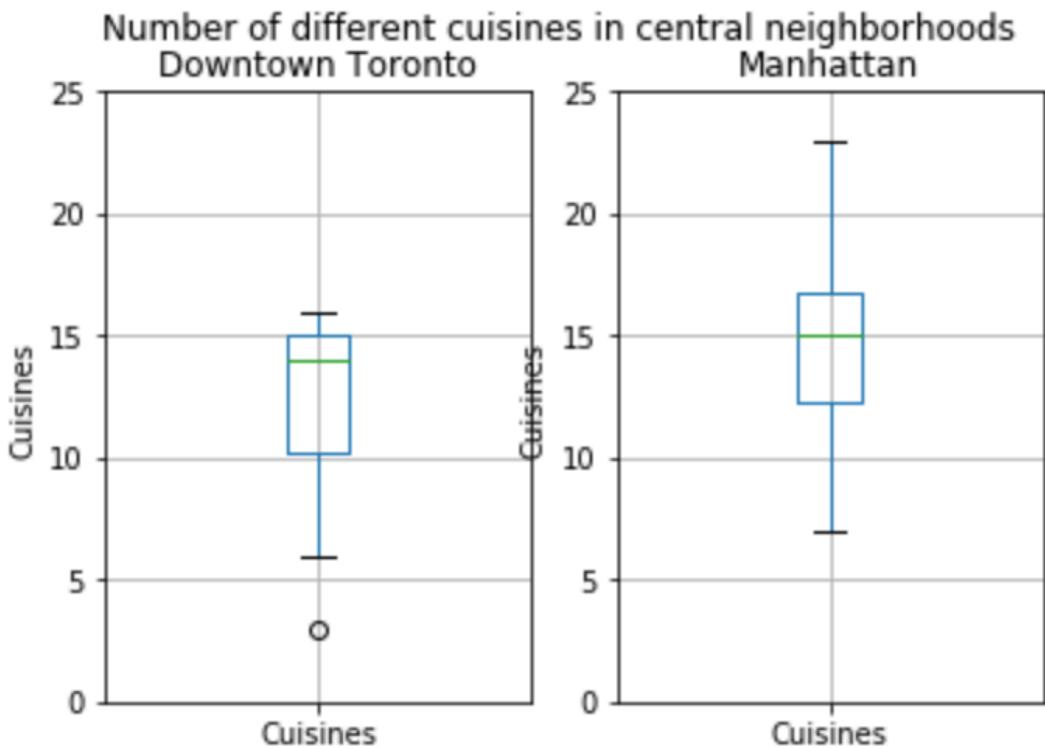
Also here let's summarize across the clusters 1 and 4 with statistical summaries and boxplot charts:

Cluster 1:

Cuisines	
count	14.000000
mean	12.142857
std	4.312007
min	3.000000
25%	10.250000
50%	14.000000
75%	15.000000
max	16.000000

Cluster 4:

Cuisines	
count	26.000000
mean	14.884615
std	4.537197
min	7.000000
25%	12.250000
50%	15.000000
75%	16.750000
max	23.000000



Also here the differences are more at the ‚best‘ and ‚worst‘ neighborhoods rather than in the typical neighborhoods. We‘ll discuss this in the following discussion chapter.

6. Discussion

The k-means clustering displayed a clear separation of neighborhoods in central parts of Downtown Toronto (in cluster 1) and neighborhoods in central parts of Manhattan (in cluster 4).

For simplicity, in the following we use the following short hand:

‚Toronto‘ = neighborhoods in cluster 1

‚Manhattan‘ = neighborhoods in cluster 4

Toronto clearly leads in the share of cafes over all venues in a neighborhood. The mean share of cafes in Toronto is 5.2%, in Manhattan only 2.3%. Also in the neighborhoods with the most and least share of cafes, Toronto leads over Manhattan.

The areas for potential improvement of Toronto are with the restaurants.

The mean share of restaurants over all venues in a neighborhood is 28.5% in Manhattan compared to 23.1% in Toronto.

The difference is even bigger if we focus on the neighborhoods with the highest share of restaurants. This is Murray Hill in Manhattan with 47% and Garden District, Ryerson in Toronto with 32%.

If we focus on the neighborhoods with the lowest share of restaurants, we see a similar picture, with East Village in Manhattan with 14.9% compared to Toronto Dominion Center, Design Exchange with 6.9%.

Also in the variety of choice dimension, which we approximated by counting the number of different international cuisines covered by the restaurants in a neighborhood there is a notable gap.

While the difference in the mean values is relatively modest with 12.1 in Toronto compared to 14.8 in Manhattan, a look at the extremes reveals significant differences.

There are four neighborhoods in Manhattan, in which restaurants represent more than 20 different international cuisines. The Turtle Bay neighborhood offers even 23 different international cuisines. This compares to a maximum of 16 international cuisines in the three best Toronto neighborhoods.

If we focus on the neighborhoods with the lowest number of international cuisines, we see 7 in Chelsea, Manhattan, and just 3 in Toronto Dominion Center, Design Exchange.

7. Conclusion and recommendations

While by no means a comprehensive quality of life study on its own right, the above data analysis of venues in the neighborhoods of Downtown Toronto and Manhattan, provides some insights, which would - if confirmed by complementing studies - suggest to take the following initiatives.

1. Downtown Toronto can build on a very good foundation of venues for eating and drinking. The numbers and variety of restaurants in typical neighborhoods is comparable between Downtown Toronto and Manhattan. Downtown Toronto could even afford to have fewer cafes (or coffee shops) and still be ahead of Manhattan in terms of share of cafes over all venues in a neighborhood.
2. By encouraging additional restaurants, especially of so far underrepresented international cuisines, in the neighborhoods with the highest restaurant share, Downtown Toronto could develop these neighborhoods to reach the benchmark of the best group of neighborhoods in Manhattan.

3. Toronto Dominion Center, Design Exchange is way below the average share of restaurants among all venues as well as below the average number of international cuisines. This might point to more structural problems in this neighborhood, which should be analyzed in a dedicated study.

The author admits that he has recently not been in either city and points out that this study is entirely based on the stated data sources and methodology.