

Finding Similar neighbourhoods Across Cities

Ajay Sharma



Introduction

- The inspiration for this project came from me being completely new in Canada and living in Toronto. And I was looking for some kind of familiarity with my old home in my new home.
- I have lived in Brooklyn, New York, USA for over 6 years, and I loved living in Brooklyn. Now Toronto is my new home.
- Through this project, I have tried to find the neighbourhoods in Toronto that are very similar to Brooklyn.



Problem and Target Audience

- In this project, I would like to do the following
 1. Find the top 5 venues for each neighbourhood/borough.
 2. Find neighbourhoods in the Toronto area that are the closest match to neighbourhoods of Brooklyn in New York in terms of food.
- The target audience for this project is people like me who are trying to find similarities in neighbourhoods across cities.



Data

- To answer the questions, we need data for two cities, Brooklyn and Toronto.
- The new New York city data was downloaded from IBM cloud [here](#).
- The Toronto city data is not available as easily as New York city data. I have used the Toronto neighbourhood data listed on the Wikipedia page [here](#).
- Next, I used Python web scraping libraries to parse the data from the Wikipedia page to compile data in the format I need for further analysis.



Data Processing

- For analysis, I cleaned the data to select attributes that are needed for analysis. The attributes needed for analysis are:

neighbourhood

Borough

Latitude

Longitude

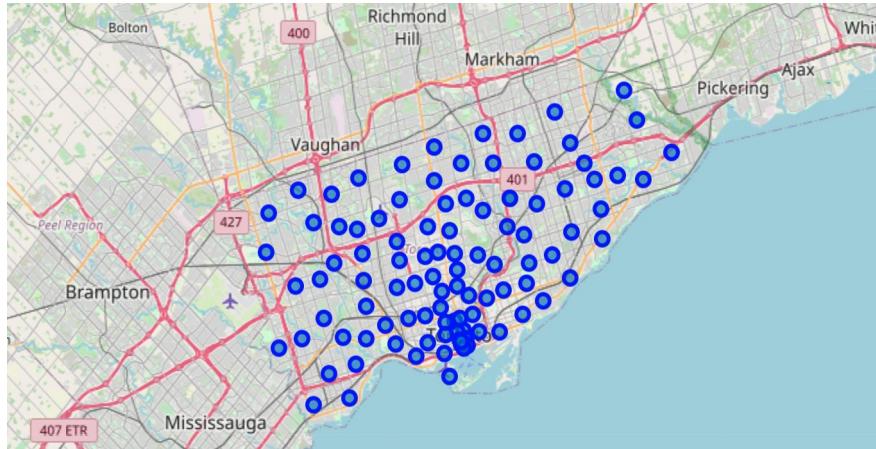
- All these values were present in downloaded data for New York; I filtered the data to select these attributes. And then, I filtered it only to use Brooklyn data for further analysis.
- I used Python Geocoding libraries to add Latitude and Longitude attribute values to Toronto data.
- The Brooklyn and Toronto data were merged for further analysis.

	Borough	Neighborhood	Latitude	Longitude
0	Brooklyn	Bay Ridge	40.625801	-74.030621
1	Brooklyn	Bensonhurst	40.611009	-73.995180
2	Brooklyn	Sunset Park	40.645103	-74.010316
3	Brooklyn	Greenpoint	40.730201	-73.954241
4	Brooklyn	Gravesend	40.595260	-73.973471

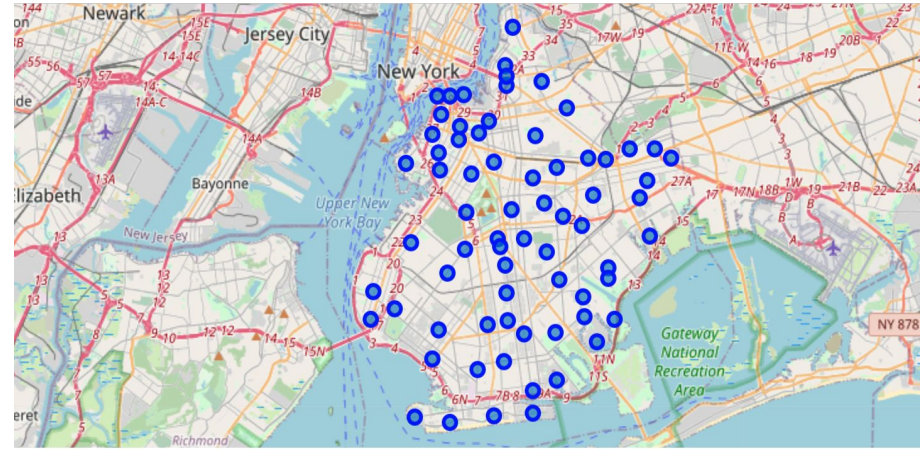
Data Exploration

- Data was plotted on maps to visualize the geographic locations of city neighbourhoods.

Toronto neighbourhoods



Brooklyn neighbourhoods





Analysis

- Foursquare API was used to download and add information for all the venues near the neighbourhoods in our data.
- The data was then filtered only to keep venues that serve food.
- This gave me data for 2,631 venues around neighbourhoods in our data.
- The data were then grouped to find the top 5 food venues neighbourhood and borough separately.



Top 5 venues

- Grouping data by neighbourhood and borough and then sorting by frequency of venue category in descending order gave me a list of the top 5 venues in each neighbourhood or borough. Some of them are listed to show as examples.

By neighbourhood

----Bath Beach----

	venue	freq
0	Donut Shop	0.07
1	Fast Food Restaurant	0.07
2	Italian Restaurant	0.07
3	Chinese Restaurant	0.07
4	Pizza Place	0.07

----Bathurst Manor,Wilson Heights,Downsview North----

	venue	freq
0	Deli / Bodega	0.17
1	Coffee Shop	0.17
2	Mediterranean Restaurant	0.17
3	Middle Eastern Restaurant	0.17
4	Fried Chicken Joint	0.17

By Borough

----Brooklyn----

	venue	freq
0	Pizza Place	0.09
1	Coffee Shop	0.07
2	Bakery	0.05
3	Deli / Bodega	0.05
4	Italian Restaurant	0.05

----Central Toronto----

	venue	freq
0	Coffee Shop	0.13
1	Sandwich Place	0.10
2	Café	0.10
3	Italian Restaurant	0.08
4	Indian Restaurant	0.05



Most common venues

- The data was then analyzed to find the most common venues in the neighbourhood and in the boroughs.

	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
0	Agincourt	Latin American Restaurant	Breakfast Spot	Yemeni Restaurant	Fast Food Restaurant	Deli / Bodega	Dessert Shop	Dim Sum Restaurant	Diner	Distillery	Doner Restaurant
1	Alderwood,Long Branch	Sandwich Place	Coffee Shop	Pizza Place	Yemeni Restaurant	Cuban Restaurant	Deli / Bodega	Dessert Shop	Dim Sum Restaurant	Diner	Distillery
2	Bath Beach	Sushi Restaurant	Pizza Place	Bubble Tea Shop	Italian Restaurant	Fast Food Restaurant	Donut Shop	Chinese Restaurant	Cantonese Restaurant	Sandwich Place	Coffee Shop
3	Bathurst Manor,Wilson Heights,Downsview North	Deli / Bodega	Coffee Shop	Fried Chicken Joint	Pizza Place	Mediterranean Restaurant	Middle Eastern Restaurant	Yemeni Restaurant	Ethiopian Restaurant	Dessert Shop	Dim Sum Restaurant
4	Bay Ridge	Pizza Place	Italian Restaurant	American Restaurant	Greek Restaurant	Bagel Shop	Sandwich Place	Café	Sushi Restaurant	Chinese Restaurant	Snack Place

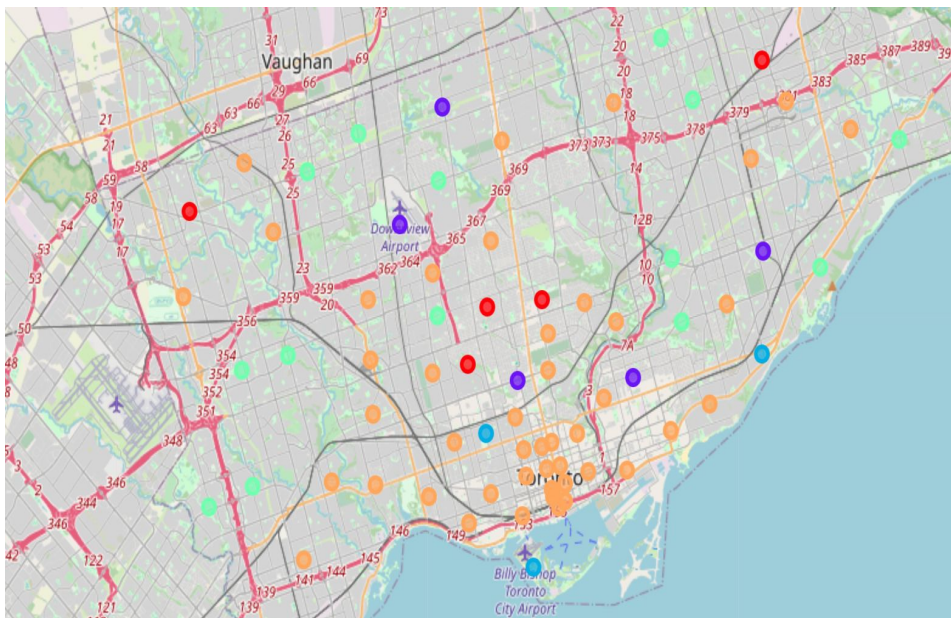
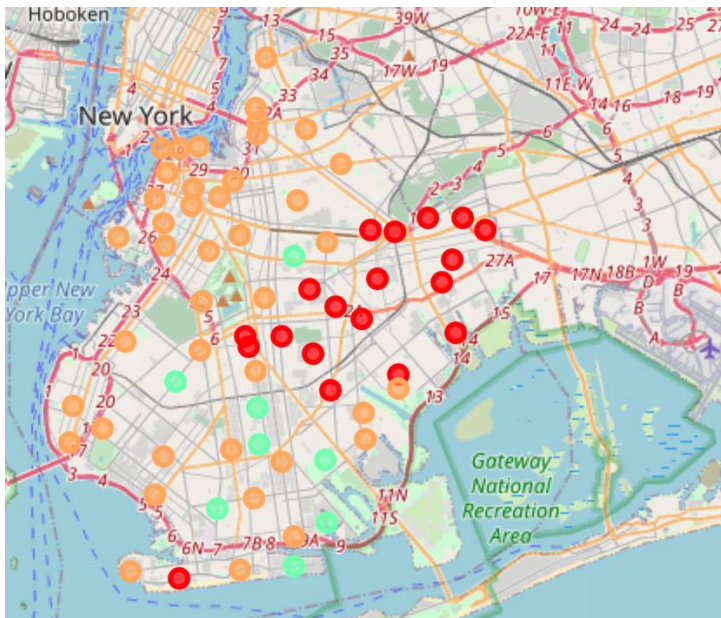
By neighbourhood

By Borough

	Borough	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	Brooklyn	Pizza Place	Coffee Shop	Bakery	Deli / Bodega	Italian Restaurant	Chinese Restaurant	Mexican Restaurant	Ice Cream Shop	Café	Bagel Shop
1	Central Toronto	Coffee Shop	Sandwich Place	Café	Italian Restaurant	Restaurant	Fast Food Restaurant	American Restaurant	Dessert Shop	Indian Restaurant	Thai Restaurant
2	Downtown Toronto	Coffee Shop	Café	Restaurant	Japanese Restaurant	Italian Restaurant	Bakery	Seafood Restaurant	Sushi Restaurant	Pizza Place	American Restaurant
3	Downtown TorontoStn A PO Boxes25 The Esplanade	Coffee Shop	Restaurant	Deli / Bodega	Café	Japanese Restaurant	Italian Restaurant	Fast Food Restaurant	Seafood Restaurant	Breakfast Spot	Fried Chicken Joint
4	East Toronto	Greek Restaurant	Coffee Shop	Italian Restaurant	Ice Cream Shop	Restaurant	Sandwich Place	Bakery	Fast Food Restaurant	Café	Brewery

Clustering

- Finally, I used the K-Means clustering algorithm to separate the neighbourhoods into different clusters based on similarity in food venues around the neighbourhoods.





Results and Discussion

- From the above analysis, as results, I have:
 1. A list of the top 5 venues for each neighbourhood and borough in our data. I can use this list to explore the best cuisines in the Toronto neighbourhood.
 2. I also have a list of Toronto neighbourhoods that are very similar to Brooklyn. I can use this list to explore the places and find a neighbourhood that I would like to settle in the near future.
- When I moved to Canada, I did not know where to start exploring. But, I knew that I wanted to find neighbourhoods similar to my previous home and neighbourhood.
- Food is an essential part of modern culture. And if I had a list of places similar to my previous home in terms of food culture, it will give me an idea about where to start. Diversity and similarity across neighbourhoods can be determined based on the food served in local restaurants. And that is the basis of this project.



Conclusion

Based on the results, I have 47 neighbourhoods similar to my old home, and I cannot wait to start exploring these neighbourhoods as soon as the current pandemic (COVID-19) is over.

A similar analysis could be used to explore/compare any neighbourhood based on some attributes, even if they are different than the ones I have used in the dataset for this project.