

Comparing Toronto and New York

Capstone Project - The Battle of Neighborhoods



Problem & Background

- This final project explores the similarity or dissimilarity in aspects from a tourist point of view regarding food, accommodation, beautiful places, and many more.
- We will explore, segment, and cluster the neighborhoods in New York City and Toronto, as they are famous places in the world. They are diverse in many ways. Both are multicultural as well as the financial hubs of their respective countries.



Data Description

- We have selected two cities to explore their neighborhoods.
- For Downtown Toronto we extracted table of Toronto's Borough from Wikipedia page and cleaned it according to our requirements. Which include eliminating 'Not assignment' values, combine neighborhoods that have some geographical coordinates at each borough and sorted against the concerned borough.
- For data verification and further exploration, we use Foursquare API to get the coordinates of Downtown Toronto and explore its neighborhoods.
- For Manhattan, we used a saved data file which is already explored through foursquare API in which we have extracted all the boroughs of New York and then sorted against the concerned borough.
- Then we explored both neighborhoods as venues and venue categories



Methodology

- For this problem, we'll clean the data, and then read it into a pandas dataframe so that it is in a structured format.
- Once the data is in a structured format, will get the services of Foursquare API to explore the data of two cities, in terms of their neighborhoods.
- The data also include the information about the places around each neighborhood.
- We'll select Manhattan from New York and Downtown Toronto from Toronto.
- Will use k-means clustering algorithm to segment the neighborhoods with similar objects on the basis of each neighborhood data. These objects will be given priority on the basis of foot traffic (activity) in their respective neighborhoods.
- Finally, we'll use the Folium library to visualize the neighborhoods in cities and their emerging clusters. This will help to locate the tourist's areas and hubs..

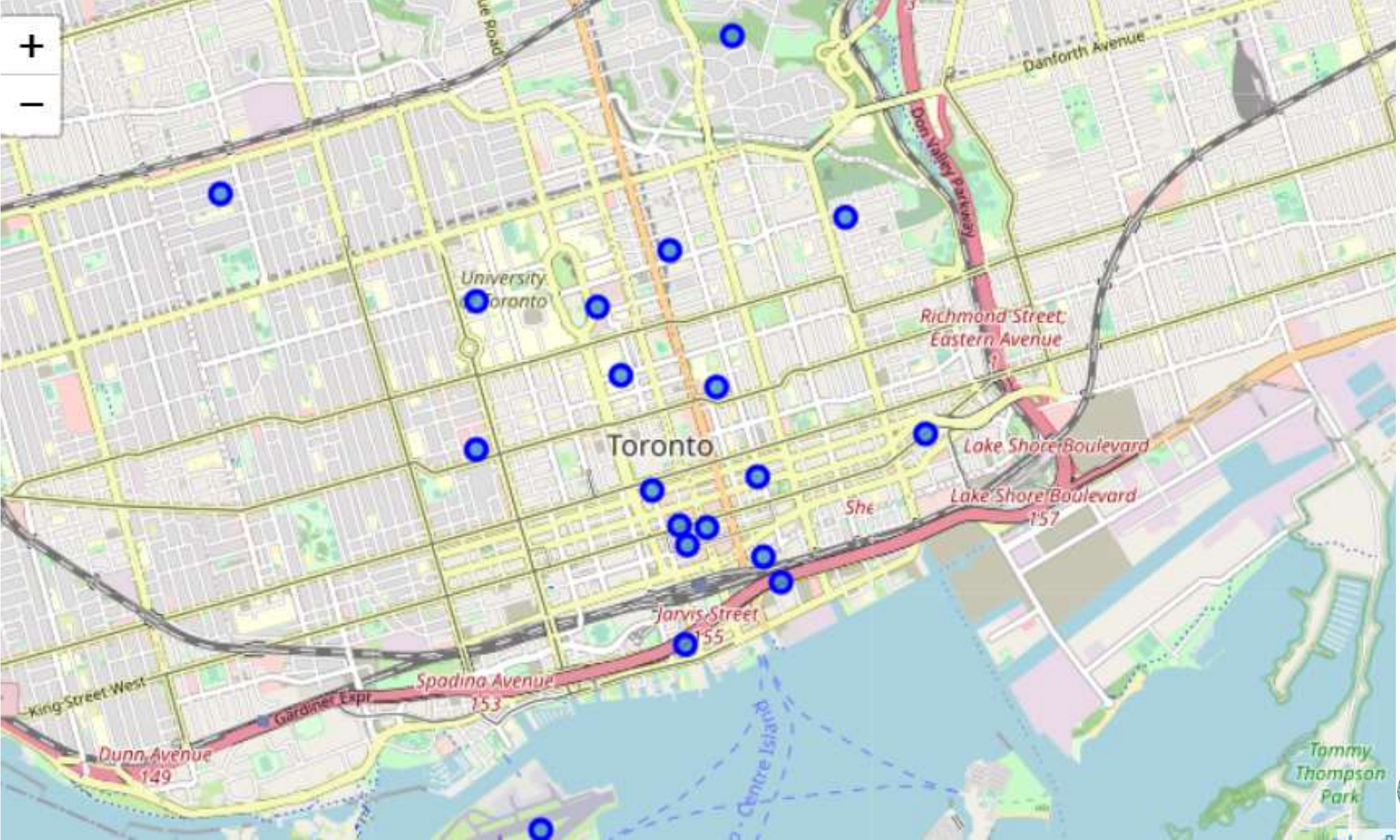


Downtown Toronto

- 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
5	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
6	M6G	Downtown Toronto	Christie	43.669542	-79.422564
7	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568
8	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752
9	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576
10	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
11	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049
12	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049
13	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420





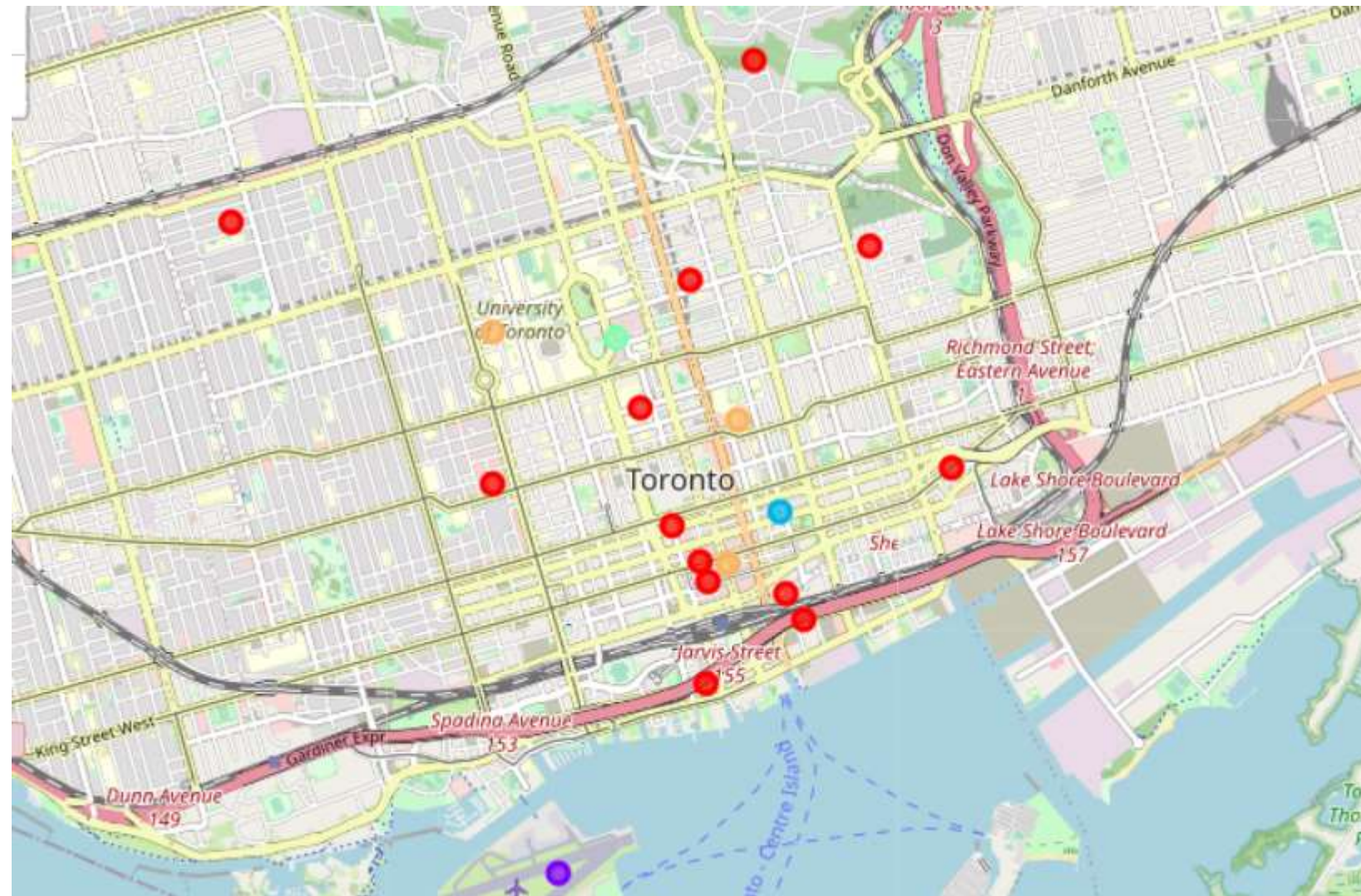
Analysis

- We analyze both boroughs neighborhoods through one hot encoding (giving '1' if a venue category is there, and '0' in case of venue category is not there). On the basis of one hot encoding, we calculate mean of the frequency of occurrence of each category and picked top ten venues on that basis for each neighborhood. It means the top venues are showing the foot traffic or the more visited places.



Clustering Downtown Toronto Neighborhoods

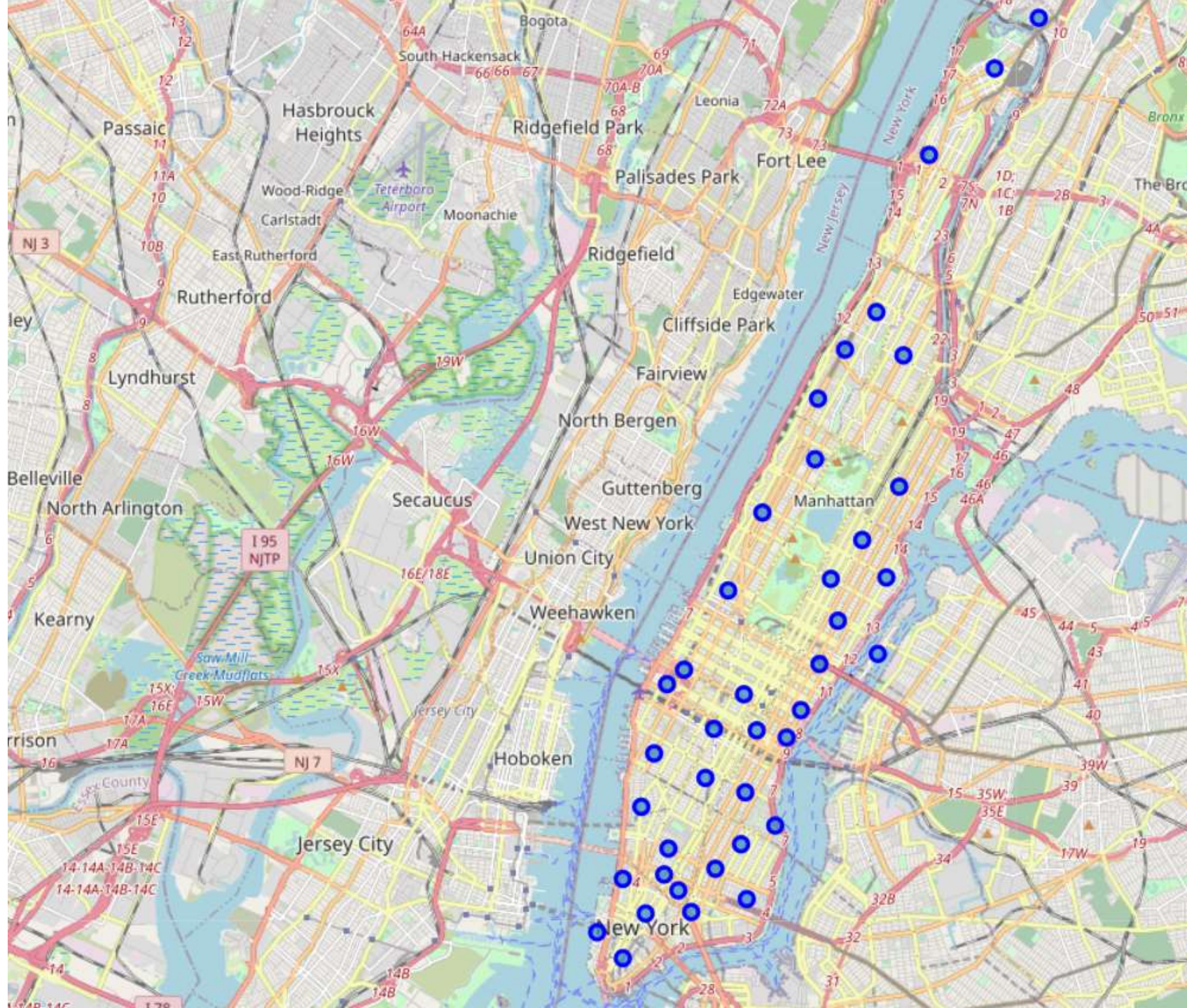
Regent Park, Harbourfront
Queen's Park, Ontario Provincial Government
Garden District, Ryerson
St. James Town
Berczy Park
Central Bay Street
Christie
Richmond, Adelaide, King
Harbourfront East, Union Station, Toronto Islands
Toronto Dominion Centre, Design Exchange
Commerce Court, Victoria Hotel
University of Toronto, Harbord
Kensington Market, Chinatown, Grange Park
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport
Rosedale
Stn A PO Boxes
St. James Town, Cabbagetown
First Canadian Place, Underground city
Church and Wellesley

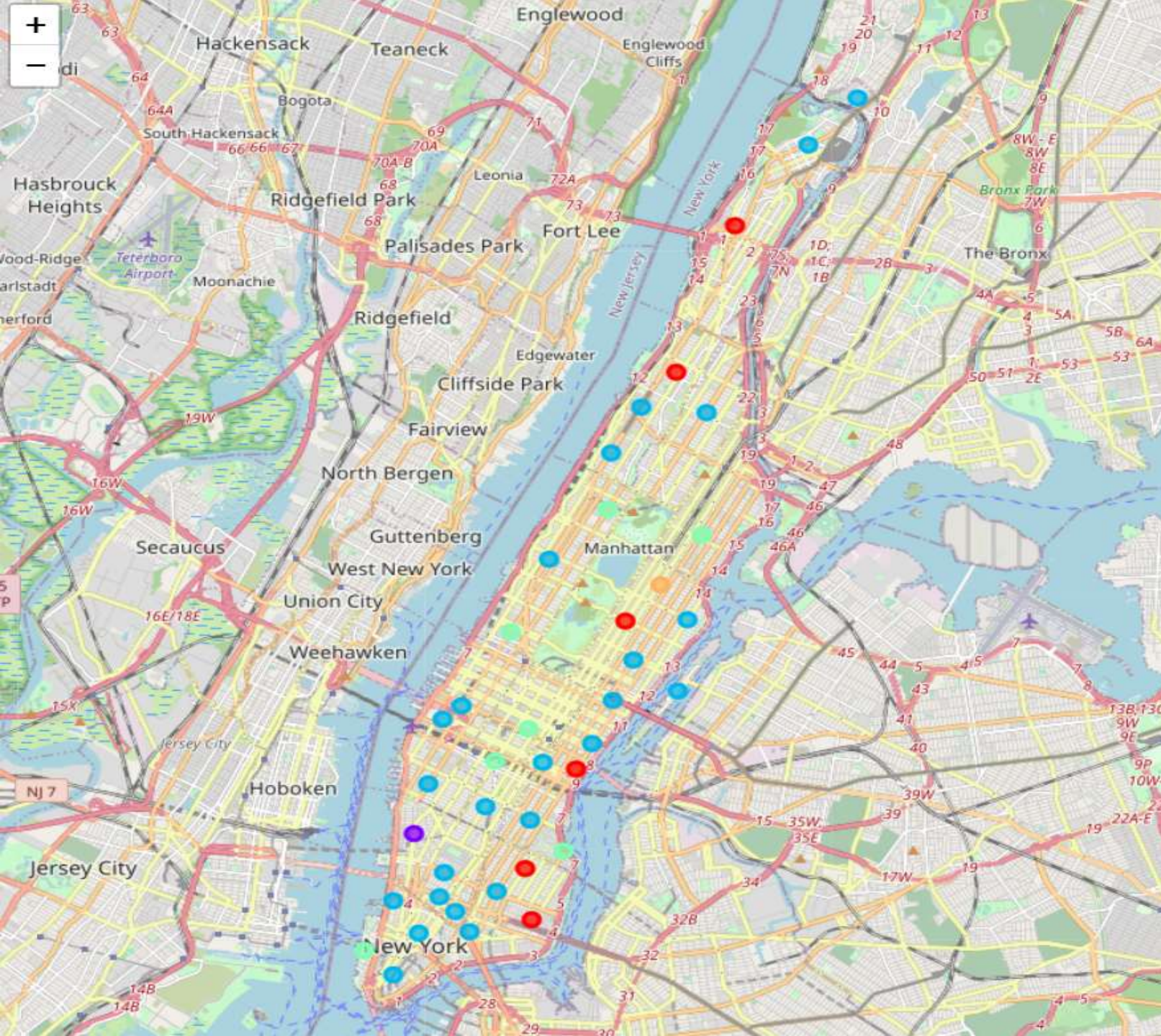


	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688
5	Manhattan	Manhattanville	40.816934	-73.957385
6	Manhattan	Central Harlem	40.815976	-73.943211
7	Manhattan	East Harlem	40.792249	-73.944182
8	Manhattan	Upper East Side	40.775639	-73.960508
9	Manhattan	Yorkville	40.775930	-73.947118
0	Manhattan	Lenox Hill	40.768113	-73.958860
1	Manhattan	Roosevelt Island	40.762160	-73.949168
2	Manhattan	Upper West Side	40.787658	-73.977059
3	Manhattan	Lincoln Square	40.773529	-73.985338
4	Manhattan	Clinton	40.759101	-73.996119
5	Manhattan	Midtown	40.754691	-73.981669
6	Manhattan	Murray Hill	40.748303	-73.978332
7	Manhattan	Chelsea	40.744035	-74.003116
8	Manhattan	Greenwich Village	40.726933	-73.999914
9	Manhattan	East Village	40.727847	-73.982226
0	Manhattan	Lower East Side	40.717807	-73.980890
1	Manhattan	Tribeca	40.721522	-74.010683
2	Manhattan	Little Italy	40.719324	-73.997305
3	Manhattan	Soho	40.722184	-74.000657
4	Manhattan	West Village	40.734134	-74.006180

Manhattan New York







Clustering Manhattan Neighborhoods

- Marble Hill
- Chinatown
- Washington Heights
- Inwood
- Hamilton Heights
- Manhattanville
- Central Harlem
- East Harlem
- Upper East Side
- Yorkville
- Lenox Hill
- Roosevelt Island
- Upper West Side
- Lincoln Square
- Clinton
- Midtown
- Murray Hill
- Chelsea
- Greenwich Village
- East Village
- Lower East Side
- Tribeca
- Little Italy
- Soho
- West Village
- Manhattan Valley
- Morningside Heights
- Gramercy
- Battery Park City
- Financial District
- Carnegie Hill
- Noho
- Civic Center
- Midtown South
- Sutton Place
- Turtle Bay
- Tudor City
- Stuyvesant Town
- Flatiron
- Hudson Yards



Conclusion

- After clustering the data of the respective neighborhoods, both cities (Boroughs) have venues which can be explored and attract the Tourists. The neighborhoods are much similar in features like Theaters, opera houses, food places, clubs, museums, parks etc. As far as concern to dissimilarity, it differs in terms of some unique places like historical places and monuments.
- The downtown Toronto and Manhattan neighborhoods have more like similar venues. As we know that every place is unique in its own way, so that's argument is present in both neighborhoods. The dissimilarity exists in terms of some different venues and facilities but not on a larger extent.

