

# **Capstone Project**

**Analysing and Clustering  
Neighbourhoods and Venues**

**For  
New York City  
&  
Toronto**

## **Aim of the project:**

- To gain a better understanding of similarities and differences between the two cities
- What are the neighborhoods that are suitable for each type of business, and
- What types of businesses are not very desirable in each city.
- This allows business people to take better and more effective decisions regarding where to open their businesses.

# Data Acquisition and Cleaning:

**Neighbourhood Data Sets:** *(Data supplied by IBM Applied Data Science Course)*

**New York City** = “[https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork\\_data.json](https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json)”

**Toronto** = “[https://en.wikipedia.org/w/index.php?](https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=1011037969)

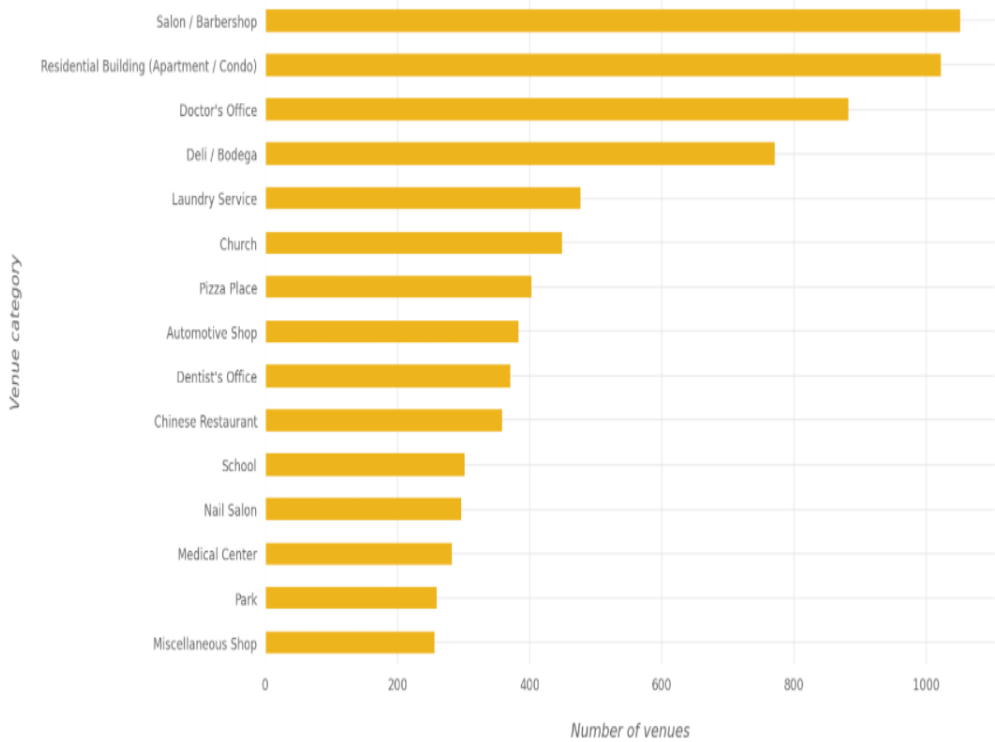
[title=List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M&oldid=1011037969](https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=1011037969)”

Neighborhood Data: datasets that lists the names of the neighborhoods of NYC and Toronto and their latitude and longitude coordinates.

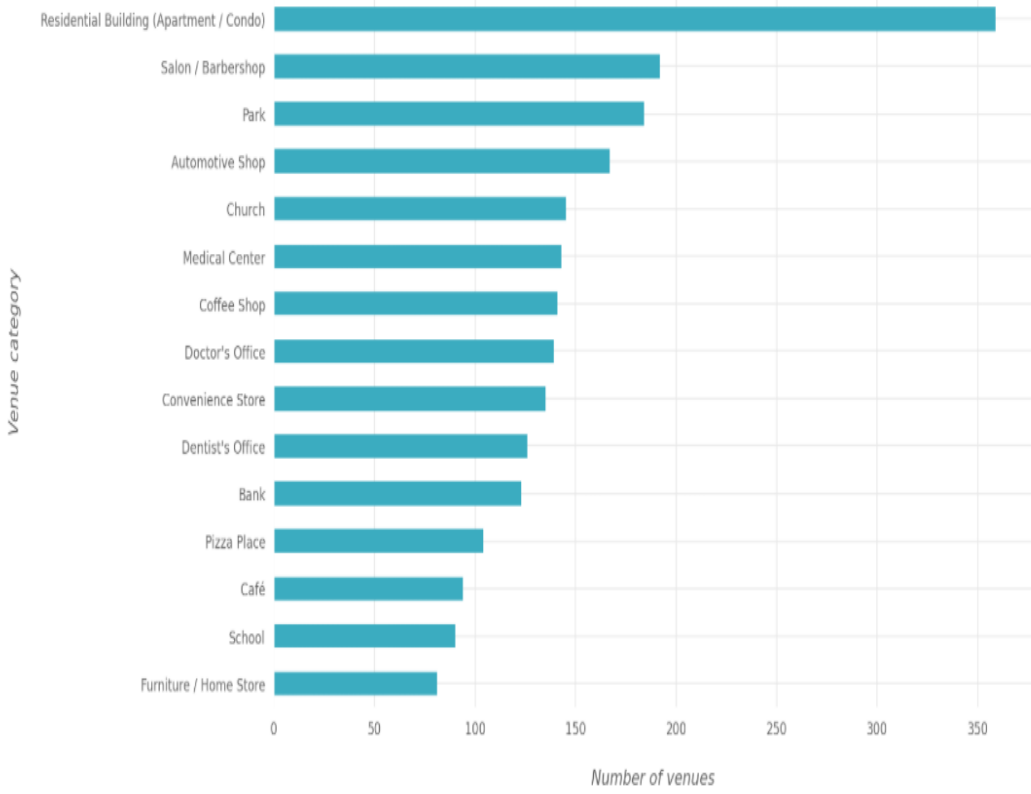
Venues data: data that describes the top 100 venues (restaurants, cafes, parks, museums, etc.) in each neighborhood of the two cities.  
The data should list the venues of each neighborhood with their categories.

This data has been retrieved from Foursquare which is one of the world largest sources of location and venue data.  
Foursquare API will be utilized to get and download the data.

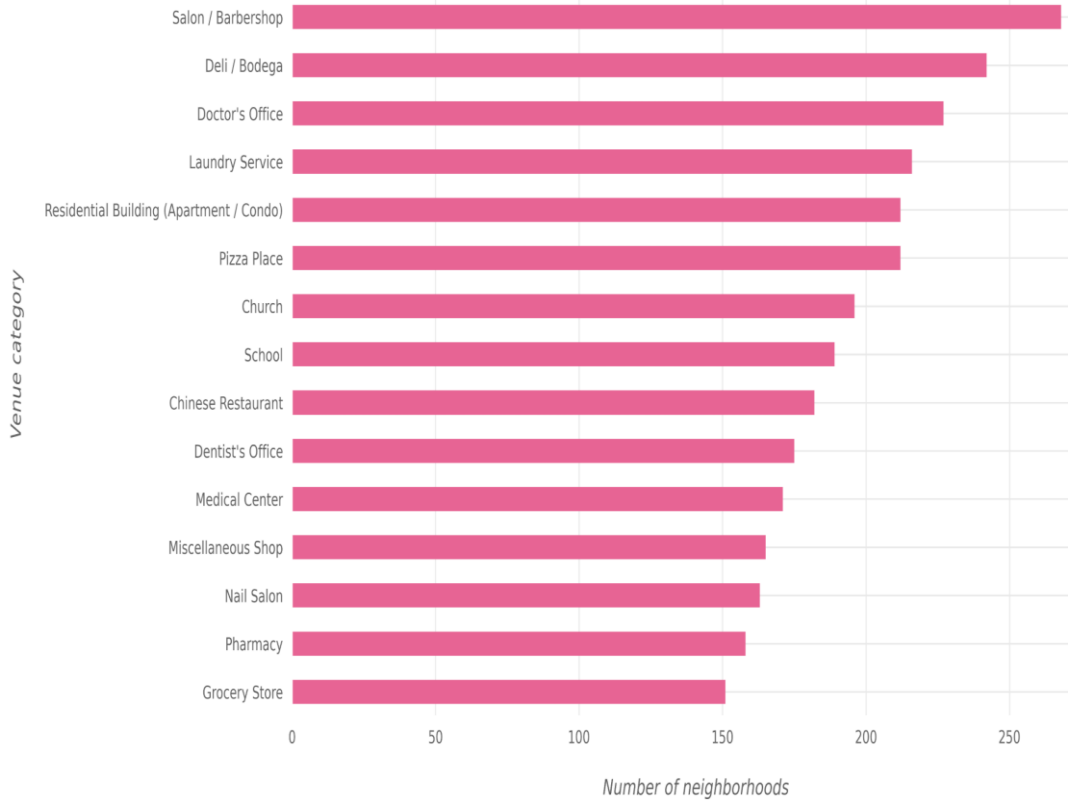
# Most common categories in New York City



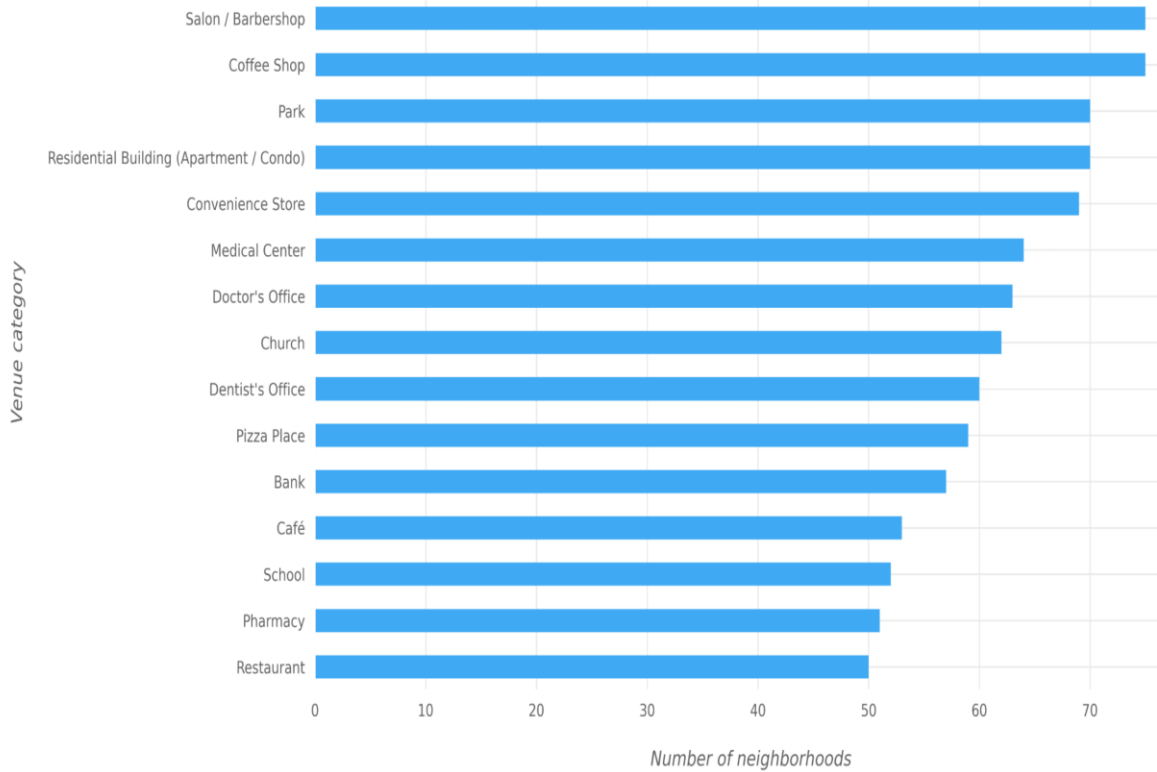
# Most common categories in Toronto



# Most common venues in New York City



# Most common venues in Toronto



## Rare categories in New York City

Venue Category	Count
EV Charging Station	1
Australian Restaurant	1
Israeli Restaurant	1
Tex-Mex Restaurant	1
Well	1
Tree	1
Music Festival	1
Circus	1
Tibetan Restaurant	1
Costume Shop	1
Buddhist Temple	1
Night Market	1
Stadium	1
Buffet	1
Airport	1

## Rare categories in Toronto

Venue Category	Count
Perfume Shop	1
Luggage Store	1
Baggage Claim	1
Corporate Coffee Shop	1
Fish Market	1
Rehab Center	1
Mattress Store	1
Ski Chalet	1
Entertainment Service	1
College Hockey Rink	1
Sake Bar	1
College Basketball Court	1
Argentinian Restaurant	1
Ski Lodge	1
Museum	1

Results from Clustering New York City:

	Neighborhood_	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category
1	Co-op City	School	Residential Building (Apartment / Condo)	Fast Food Restaurant	Dentist's Office	Park	High School	Liquor Store
3	Fieldston	College Academic Building	College Administrative Building	College Residence Hall	Residential Building (Apartment / Condo)	Tech Startup	Historic Site	College Classroom
5	Kingsbridge	Laundry Service	Ice Cream Shop	Park	Bank	Salon / Barbershop	Discount Store	Bagel Shop
10	Baychester	Gas Station	Bank	Automotive Shop	Chinese Restaurant	Donut Shop	Burger Joint	Residential Building (Apartment / Condo)
12	City Island	Harbor / Marina	Italian Restaurant	Antique Shop	Pharmacy	Deli / Bodega	Park	Residential Building (Apartment / Condo)
22	Port Morris	Factory	Bar	Government Building	Financial or Legal Service	Storage Facility	Automotive Shop	Hardware Store
27	Clason Point	Park	Housing Development	Lounge	Salon / Barbershop	Automotive Shop	Intersection	Event Space
28	Throgs Neck	Other Great Outdoors	Deli / Bodega	Chinese Restaurant	Residential Building (Apartment / Condo)	Italian Restaurant	Tattoo Parlor	Japanese Restaurant

Results from Clustering Toronto:

	Borough	Cluster_Labels2	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category
0	Scarborough	0	Automotive Shop	Factory	Coffee Shop	Dentist's Office	Medical Center	Sandwich Place	Convenience Store
1	Scarborough	0	Automotive Shop	Medical Center	Park	General Entertainment	Salon / Barbershop	Bar	Arts & Crafts Store
10	Scarborough	0	Automotive Shop	Furniture / Home Store	Auto Garage	Hardware Store	Storage Facility	Miscellaneous Shop	Factory
12	Scarborough	0	Automotive Shop	Auto Garage	Storage Facility	Doctor's Office	Church	Coffee Shop	Tech Startup
34	North York	0	Automotive Shop	Residential Building (Apartment / Condo)	Auto Dealership	Miscellaneous Shop	School	Bank	Transportation Service
53	Downtown Toronto	0	Automotive Shop	Furniture / Home Store	Italian Restaurant	Light Rail Station	Park	Design Studio	Hardware Store
81	Toronto/York	0	Automotive Shop	Furniture / Home Store	Doctor's Office	Miscellaneous Shop	Gas Station	Event Space	Park
96	North York	0	Automotive Shop	Gas Station	Bank	Furniture / Home Store	Italian Restaurant	Church	Gym
102	Etobicoke	0	Factory	Automotive Shop	Auto Dealership	Hardware Store	Medical Center	Transportation Service	Café

## **Conclusion:**

In this project, the neighborhoods of New York City and Toronto were clustered using K-Means clustering algorithm into multiple groups based on the categories (types) of the venues in these neighborhoods.

The results showed that there are venue categories that are more common in some cluster than the others; the most common venue categories differ from one cluster to the other.