



Travel Recommendation

CAPSTONE PRESENTATION

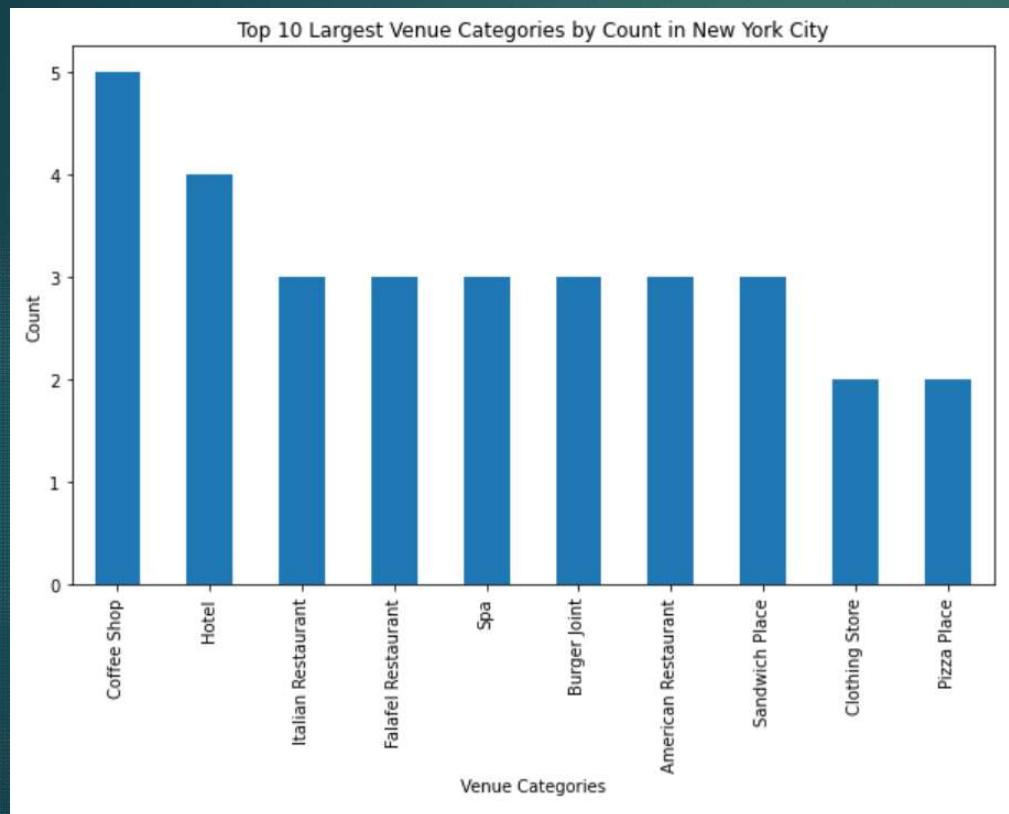
Recommendation services are key to BeGone International

- ▶ Increase reader/traveler engagement from its outlet across its new and existing user base
- ▶ Increase recommendation services from BeGone International
- ▶ Marketing of local and destination venues
- ▶ Demonstration of data analytic capabilities

Data

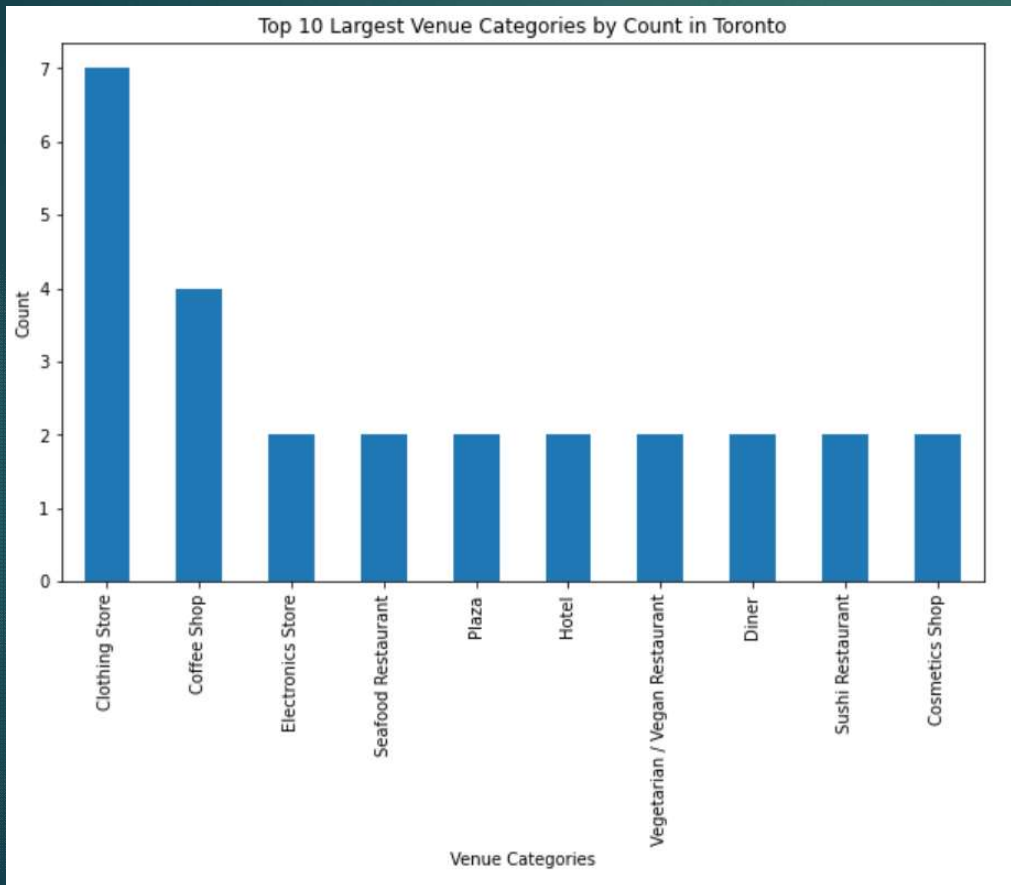
- ▶ For this initiative, The research resources are using location using Python libraries and venue data from publically accessible form from FourSquare. From this data source, data on New York City and Toronto venues was acquired.
- ▶ Venue information within the vicinity of the neighborhoods was retrieved using the Foursquare location API. Maps and graphs were generated using Python libraries to visualize the location and venues.
- ▶ Anomalies in data were excluded, i.e. venues without valid postal codes.
- ▶ Mail codes removed from Toronto postal codes, as they did not inform or add value to the initiative .

Statistical Analysis on Source Data



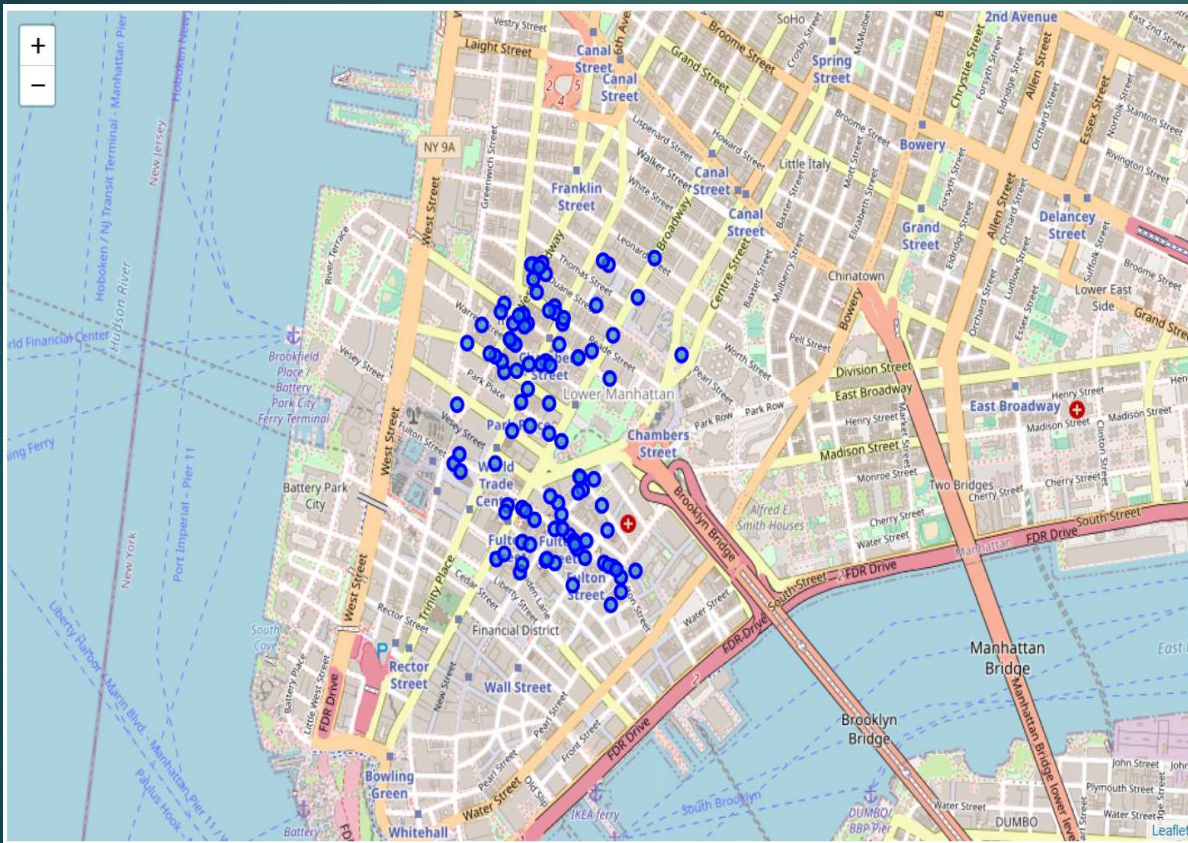
Coffee Shop	5
Hotel	4
Italian Restaurant	3
Falafel Restaurant	3
Spa	3

Statistical Analysis on Source Data



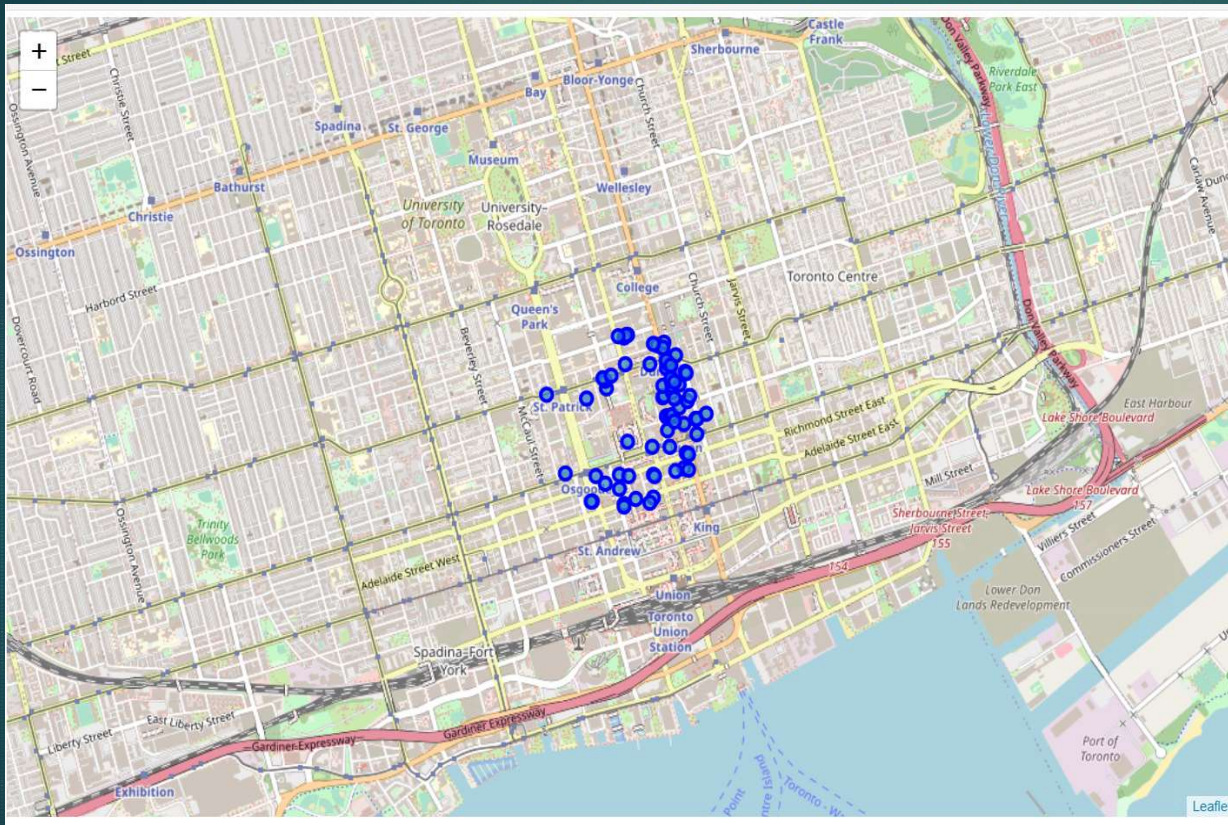
Clothing Store	7
Coffee Shop	4
Electronics Store	2
Seafood Restaurant	2
Plaza	2

Map of New York City, NY



- 100 venues were selected with 500 meters of New York City
- 100 venues were assessed
- 4 zip codes were captured

Map of Toronto, Canada



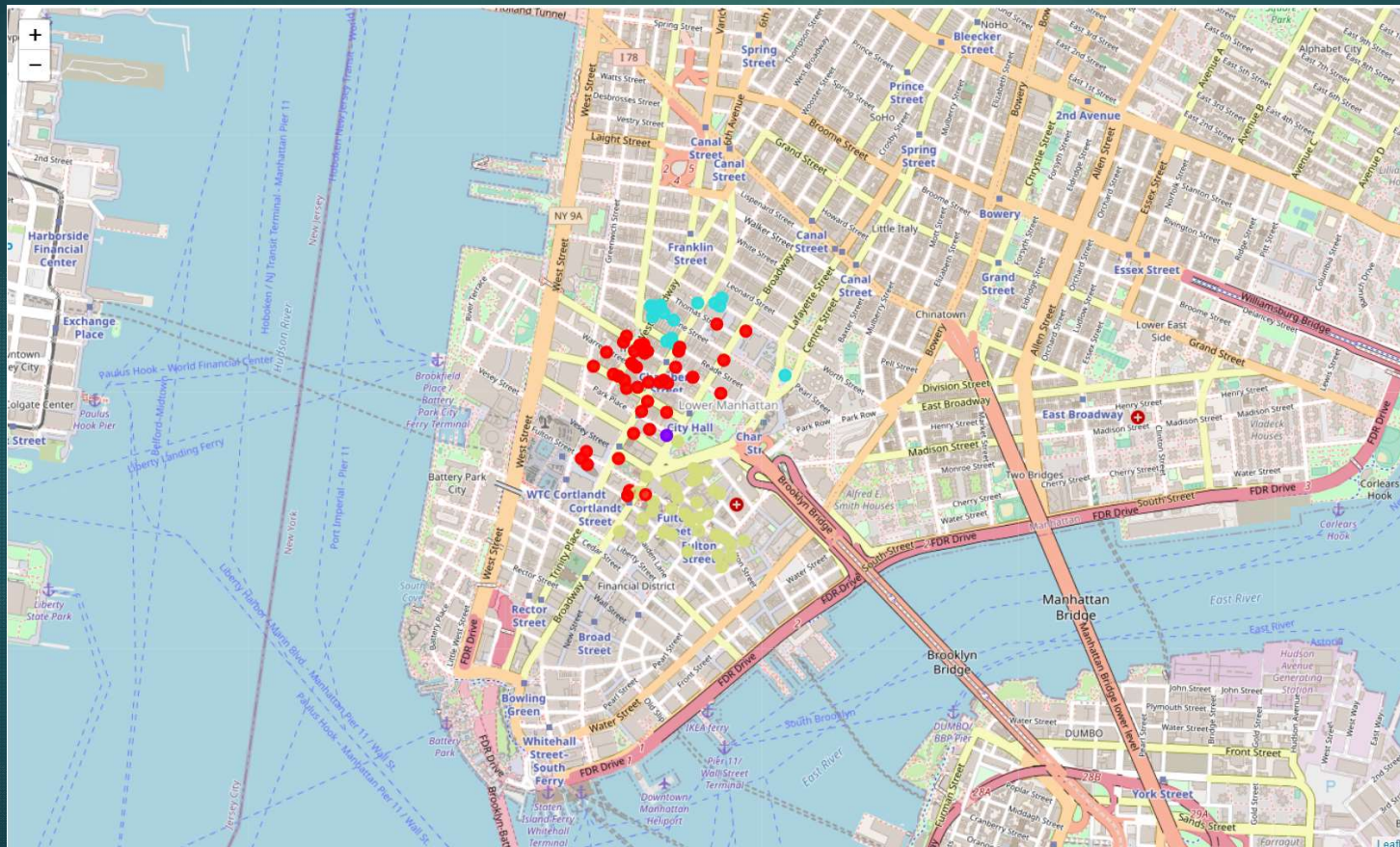
- 100 venues were selected with 500 meters of New York City
- 78 venues were assessed
- 7 zip codes were captured

Data

The use of the FourSquare data on venues within New York and Toronto required the following maintenance, in order to execute this initiative:

- ▶ Removal of Toronto venues with Null postal codes, from the any further processing or analysis
- ▶ Re-formatting of the Toronto postal codes, to remove trailing mail delivery codes

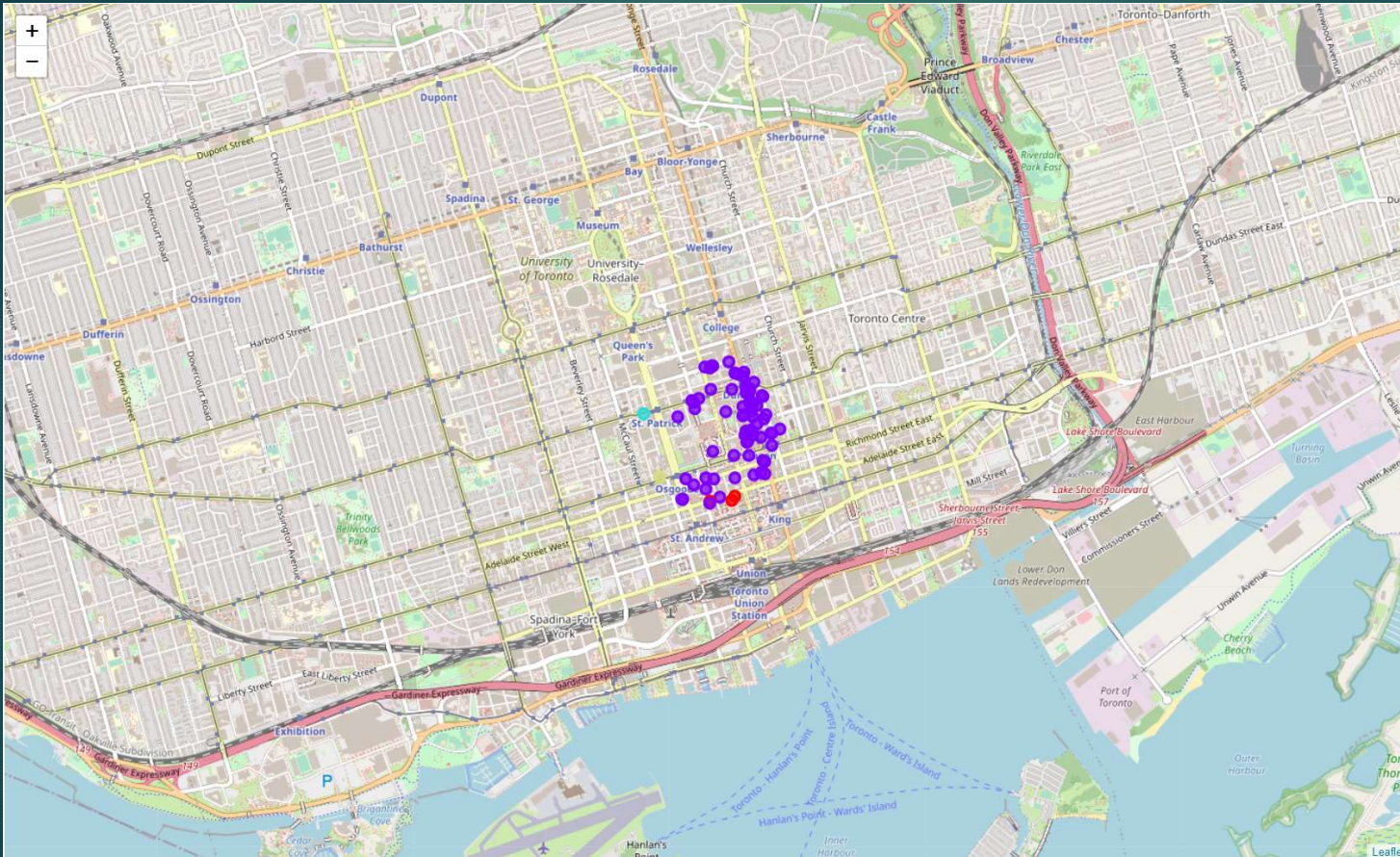
Clustering of New York City venues



Using the 1st most common venue category for each cluster:

- Cluster 0 – Coffee Shop
- Cluster 1 – Building
- Cluster 2 – Sushi Restaurant
- Cluster 3 – Sandwich Place

Clustering of Toronto venues



Using the 1st most common venue category for each cluster:

- Cluster 0 – American Restaurant
- Cluster 1 – Clothing Store
- Cluster 2 – Cafe
- Cluster 3 – Coffee Shop

Results

- ▶ Using categorical attributes, specifically the venue type, the data set on venues had to be transformed. Those attributes were converted into frequency of occurrence of a venue category across a postal code for both New York City and Toronto.
- ▶ Using k-means (N =number of postal codes per location), those postal codes were associated or clustered based on the highest frequency of venue categories in that location.
- ▶ The cluster of zip codes was generated based most common venue category (with the highest frequency).
- ▶ Visualizations were created to demonstrate venue's location within those postal code and venue clusters in both New York City and Toronto.
- ▶ Refer to the resulting data frames.

New York City

Cluster Labels	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
0	Indian Restaurant	Coffee Shop	Wine Shop	Hotel Bar	Hotel	Women's Store	Gym	Furniture / Home Store	French Restaurant	Falafel Restaurant	42
1	Plaza	Building	Yoga Studio	Discount Store	Comic Shop	Coworking Space	Cuban Restaurant	Dance Studio	Donut Shop	Cocktail Bar	2
2	Sushi Restaurant	Spa	Gym / Fitness Center	Restaurant	Asian Restaurant	Bakery	Cocktail Bar	Falafel Restaurant	Medical Center	Molecular Gastronomy Restaurant	15
3	Hotel	Coffee Shop	Italian Restaurant	Burger Joint	Pizza Place	Sandwich Place	Yoga Studio	Fast Food Restaurant	Hotel Bar	Gym	41

Toronto

Cluster Labels	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
0	Café	Vietnamese Restaurant	Comic Shop	Hotel	Gym / Fitness Center	Gastropub	Furniture / Home Store	Fast Food Restaurant	Electronics Store	Diner	1
1	Plaza	Breakfast Spot	New American Restaurant	Noodle House	Office	Colombian Restaurant	Concert Hall	Restaurant	Seafood Restaurant	Monument / Landmark	16
2	Coffee Shop	Vegetarian / Vegan Restaurant	Colombian Restaurant	Gym / Fitness Center	Gastropub	Furniture / Home Store	Fast Food Restaurant	Electronics Store	Diner	Dessert Shop	2
3	Latin American Restaurant	Smoothie Shop	Department Store	Vietnamese Restaurant	Colombian Restaurant	Gym / Fitness Center	Gastropub	Furniture / Home Store	Fast Food Restaurant	Electronics Store	3
4	American Restaurant	Gym / Fitness Center	Salad Place	Colombian Restaurant	Gastropub	Furniture / Home Store	Fast Food Restaurant	Electronics Store	Diner	Dessert Shop	3
5	Clothing Store	Theater	Diner	Cosmetics Shop	Coffee Shop	Electronics Store	Vietnamese Restaurant	Plaza	Furniture / Home Store	Video Game Store	36
6	Japanese Restaurant	Breakfast Spot	Hotel	Italian Restaurant	Dessert Shop	Modern European Restaurant	New American Restaurant	Coffee Shop	Poke Place	Bubble Tea Shop	17

Observations

Based on the research findings,

- ▶ Coffee Shops was the most common venue out of the top 10 categories within the locations Toronto and New York City.
- ▶ 4 zip codes were assessed in New York City and 42 zip codes were assessed in Toronto.
- ▶ 100 venue were assessed in New York City and 68 venues were assessed from Toronto, showing that data set from Toronto had discrepancies.
- ▶ These categories are the first most common venue category. The analysis is supported based on the count of venue categories captured in the provided bar graphs.

Conclusion

- ▶ The use of clustering for generating recommendation services should be proposed for further investment and/or evaluation given the overlap in identified venue types.
- ▶ Using a larger data set or data with missing values. This impacts number of venues and frequency of venue categories available to be evaluated.
- ▶ The use of additional venue categories, in addition to those in this analysis, may provide specific destinations of interests to their clients based on destination cities or even specific zip codes.