The Battle of the Neighbourhoods: Toronto vs. New York City

Final Report

11/1/2021

Kolitha Hasitha Wanniarachchi

The Battle of the Neighbourhoods: Toronto vs. New York City

1. INTRODUCTION & BUSINESS PROBLEM

New York and Toronto are two famous cities that are apart from 556km. According to the "BestCities.org", New York, with more than 20 million people, was the 2nd best city in the world in 2020, while Toronto, with almost 7 million population, was 13th best among the thousands of cities around the world. On a more granular level, it is interesting to compare these cities using the following six metrics that are widely used to measure their overall performances.

- **Place**: The perceived quality of a city's natural and built environment, including the sub-categories of weather, safety, neighbourhoods & landmarks, and outdoors.
- **Product**: A city's key institutions, attractions and infrastructure, including the sub-categories of airport connectivity, attractions, museums, university ranking, convention center, and pro sports teams.
- **Programming**: The arts, culture, entertainment and culinary scene in a city, including the sub-categories of shopping, culture, restaurants, and nightlife.
- **People**: The immigration rate and diversity of a city, including the subcategories of foreign-born and educational attainment.
- **Prosperity**: A city's employment and corporate head offices, including the sub-categories of Fortune 500 companies and household income.
- **Promotion**: The number of stories, references and recommendations shared online about a city, including the sub-categories of Google search results, Google trends, Facebook check-ins, Instagram hashtags, and TripAdvisor reviews.

The City of New York has performed exceptionally well on the Programming and Promotions, which indicates the diversity, sustainability and consistency across its boroughs and neighbourhoods to stay persistent and steady. On the other side, Toronto has shown significant improvements in People and Products, which is predictable to a certain extent with its distribution of immigrants and the amount of intelligence they brought to the city.

Given the multi-cultural nature and the range of business opportunities, an investor would like to open a set of new branches of a famous Asian Restaurant chain in both New York and Toronto cities. It would be an exciting investigation to explore the neighbourhoods in both cities concerning the varieties of Asian restaurants and their spread to determine the number of branches and the boroughs to locate them. However, in addition to the socio-economic factors, now the Covid-19 Pandemic has to be included in the equation to optimize decision as both these cities have been severely impacted by multiple waves of this disaster. As requested by the venture capitalist, it is vital to observe both cities to maximize their return on investments while mitigating potential risks.

In the first sprint of this investment project, we will be running a feasibility study on the following objectives.

1. Understand the geographical distribution of restaurants in Toronto and New York cities.
2. Identify the prevalence of the Covid-19 Pandemic in potential neighbourhoods to locate new branches.

2. DESCRIPTION OF DATA

Four different sets of data, as given below, were used to conduct the analysis.

1. Toronto Neighbourhood Data.
2. Toronto Covid-19 Data.
3. New York City Neighbourhood Data.
4. New York City Covid-19 Data.

2.1. TORONTO NEIGHBOURHOOD DATA

Toronto has a total of 10 boroughs and 103 neighbourhoods. To conduct geospatial analysis, it is essential to have latitude and longitude coordinates of each neighbourhood. Wikipedia has a comprehensive list of Toronto neighbourhoods with their postal codes. Coursera capstone project provides a CSV file that contains all the geographical coordinates of Toronto neighbourhoods. Figure 2.1.1 illustrates a cleaned sample dataset.

Figure 2.1.1: Toronto Neighbourhood Data

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

2.2. TORONTO COVID-19 DATA

To examine the prevalence of Covid-19 in Toronto, I used the data available on the City of Toronto website. Toronto Public Health publishes a weekly report in a table format, where we can download as an XML, JSON or a CSV. According to the Toronto Public Health statistics, the total number of confirmed Covid-19 cases in Toronto as of 6th January 2021 was 61556. Figure 2.2.1 illustrates the postal code (FSA) wise total number of case counts in Toronto.

Figure 2.2.1: Toronto Covid-19 Data

	_id	Assigned_ID	Outbreak Associated	Age Group	Neighbourhood Name	FSA	Source of Infection	Classification	Episode Date	Reported Date	Client Gender	Outcome	Currently Hospitalized
0	736701	1	Sporadic	50 to 59 Years	Willowdale East	M2N	Travel	CONFIRMED	22/01/2020	23/01/2020	FEMALE	RESOLVED	No
1	736702	2	Sporadic	50 to 59 Years	Willowdale East	M2N	Travel	CONFIRMED	21/01/2020	23/01/2020	MALE	RESOLVED	No
2	736703	3	Sporadic	20 to 29 Years	Parkwoods-Donalda	M3A	Travel	CONFIRMED	05/02/2020	21/02/2020	FEMALE	RESOLVED	No
3	736704	4	Sporadic	60 to 69 Years	Church-Yonge Corridor	M4W	Travel	CONFIRMED	16/02/2020	25/02/2020	FEMALE	RESOLVED	No
4	736705	5	Sporadic	60 to 69 Years	Church-Yonge Corridor	M4W	Travel	CONFIRMED	20/02/2020	26/02/2020	MALE	RESOLVED	No

2.3. NEW YORK CITY NEIGHBOURHOOD DATA

New York City has a total of 5 boroughs and 306 neighbourhoods. JSON file '**newyork_data.json**' provided by Coursera contains the latitude and longitude coordinates of each neighbourhood. Figure 2.3.1 illustrates the first five records of the dataset.

Figure 2.3.1: New York City Neighbourhood Data

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

2.4. NEW YORK CITY COVID-19 DATA

New York government publishes daily Covid-19 statistics in their "**NYC Health**" GitHub repository, and it contains a plethora of NYC covid-19 data in various verticals. As per the records, the cumulative number of confirmed Covid-19 cases in New York was 327721. Figure 2.4.1 illustrates the sample of data.

Figure 2.4.1: New York City Covid-19 Data

	MODIFIED_ZCTA	NEIGHBORHOOD_NAME	BOROUGH_GROUP	COVID_CASE_COUNT	COVID_CASE_RATE	POP_DENOMINATOR	COVID_DEATH_COUNT
0	10001	Chelsea/NoMad/West Chelsea	Manhattan	786	2846.48	27613.09	24
1	10002	Chinatown/Lower East Side	Manhattan	2929	3888.60	75322.71	172
2	10003	East Village/Gramercy/Greenwich Village	Manhattan	1354	2508.44	53977.81	35
3	10004	Financial District	Manhattan	131	4407.63	2972.12	1
4	10005	Financial District	Manhattan	220	2512.21	8757.23	0

In the data analysis, neighbourhood data will be used to segment the venues and Covid-19 data will be merged with neighbourhood data to figure out the prevalence of the pandemic in the clustered neighbourhoods.

3. METHODOLOGY

3.1. AGGREGATE DATA

The first step of the analysis was merging neighbourhood data and Covid-19 data to generate master data tables for two cities, which are shown below in Figures 3.1.1 and 3.1.2.

Figure 3.1.1: Merged Toronto Neighbourhood and Covid-19 Data

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Case Count
0	M3A	North York	Parkwoods	43.753259	-79.329656	511.0
1	M4A	North York	Victoria Village	43.725882	-79.315572	374.0
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	838.0
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	788.0
4	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242	634.0

Figure 3.1.1: Merged New York City Neighbourhood and Covid-19 Data

	Borough	Neighborhood	Latitude	Longitude	Case Count
0	Bronx	Wakefield	40.894705	-73.847201	2364.0
1	Bronx	Co-op City	40.874294	-73.829939	1222.0
2	Bronx	Fieldston	40.895437	-73.905643	433.0
3	Bronx	Riverdale	40.890834	-73.912585	1365.0
4	Bronx	Kingsbridge	40.881687	-73.902818	4000.0

3.2. EXPLORE NEIGHBOURHOODS

Then, using the Foursquare APIs, explored the different venues in each neighbourhood for both cities. To analyse the distribution, grouped all the venues by their category and populated a frequency table. Next, selected the top 10 venues for each neighbourhood, sorting by the frequencies in descending order. Figures 3.2.1 and 3.2.2 illustrate the top 10 venues based on the neighbourhood in both cities.

Figure 3.2.1: Top 10 Venues in Toronto Neighbourhoods

	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	Lounge	Breakfast Spot	Latin American Restaurant	Skating Rink	Yoga Studio	Electronics Store	Doner Restaurant	Donut Shop	Drugstore	Eastern European Restaurant
1	Alderwood, Long Branch	Pizza Place	Pharmacy	Coffee Shop	Sandwich Place	Pub	Dance Studio	Gym	Concert Hall	Dim Sum Restaurant	Ethiopian Restaurant
2	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Pharmacy	Fried Chicken Joint	Shopping Mall	Sandwich Place	Diner	Supermarket	Middle Eastern Restaurant	Sushi Restaurant
3	Bayview Village	Japanese Restaurant	Bank	Chinese Restaurant	Café	Eastern European Restaurant	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Yoga Studio
4	Bedford Park, Lawrence Manor East	Sandwich Place	Italian Restaurant	Coffee Shop	Grocery Store	Toy / Game Store	Liquor Store	Juice Bar	Butcher	Café	Restaurant

Figure 3.2.2: Top 10 Venues in New York City Neighbourhoods

	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	Allerton	Pizza Place	Chinese Restaurant	Deli / Bodega	Supermarket	Fried Chicken Joint	Breakfast Spot	Martial Arts School	Fast Food Restaurant	Gas Station	Donut Shop
1	Annadale	Pizza Place	Liquor Store	Train Station	Restaurant	Diner	Park	Filipino Restaurant	Ethiopian Restaurant	Event Service	Event Space
2	Arrochar	Pizza Place	Deli / Bodega	Italian Restaurant	Bus Stop	Bagel Shop	Sandwich Place	Pharmacy	Athletics & Sports	Outdoors & Recreation	Mediterranean Restaurant
3	Arverne	Surf Spot	Metro Station	Sandwich Place	Bus Stop	Board Shop	BBQ Joint	Beach	Café	Donut Shop	Pizza Place
4	Auburndale	Italian Restaurant	Sushi Restaurant	Toy / Game Store	Train	Miscellaneous Shop	Pet Store	Pharmacy	Fast Food Restaurant	Athletics & Sports	Mobile Phone Shop

3.3. K-Means Clustering

Finally, using the k-means clustering method segmented the extracted venue categories in both cities to see the geographical distribution and similarities. Clustered neighbourhoods in Toronto and New York City are illustrated in Figures 3.3.1 and 3.3.2.

Figure 3.3.1: Clustered Toronto Neighbourhoods

	Cluster Labels	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	3	Agincourt	Lounge	Latin American Restaurant	Breakfast Spot	Chinese Restaurant	Drugstore	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Yoga Studio
1	3	Alderwood, Long Branch	Pizza Place	Gym	Sandwich Place	Coffee Shop	Skating Rink	Pub	Dog Run	Diner	Discount Store	Distribution Center
2	3	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Frozen Yogurt Shop	Bridal Shop	Sandwich Place	Diner	Restaurant	Intersection	Supermarket	Sushi Restaurant
3	3	Bayview Village	Café	Bank	Japanese Restaurant	Chinese Restaurant	Donut Shop	Distribution Center	Dog Run	Doner Restaurant	Drugstore	Diner
4	3	Bedford Park, Lawrence Manor East	Coffee Shop	Italian Restaurant	Sandwich Place	Thai Restaurant	Grocery Store	Indian Restaurant	Juice Bar	Liquor Store	Comfort Food Restaurant	Pharmacy

Figure 3.3.2: Clustered New York City Neighbourhoods

|	Cluster Labels	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	4	Allerton	Deli / Bodega	Pizza Place	Supermarket	Fried Chicken Joint	Pharmacy	Breakfast Spot	Spa	Gas Station	Fast Food Restaurant	Martial Arts School
1	1	Annadale	Park	Food	Pharmacy	Diner	Sushi Restaurant	Restaurant	Bakery	Pizza Place	American Restaurant	Dance Studio
2	1	Arrochar	Bus Stop	Italian Restaurant	Deli / Bodega	Pizza Place	Food Truck	Middle Eastern Restaurant	Outdoors & Recreation	Athletics & Sports	Mediterranean Restaurant	Supermarket
3	1	Arverne	Surf Spot	Sandwich Place	Metro Station	Board Shop	Bus Stop	Beach	BBQ Joint	Coffee Shop	Donut Shop	Café
4	1	Auburndale	Italian Restaurant	Gymnastics Gym	Mobile Phone Shop	Mattress Store	Fast Food Restaurant	Toy / Game Store	Sushi Restaurant	Pet Store	Pharmacy	Noodle House

4. RESULTS

4.1. Toronto Clusters

Toronto was segmented into 5 clusters as shown in the figure 4.1.1.

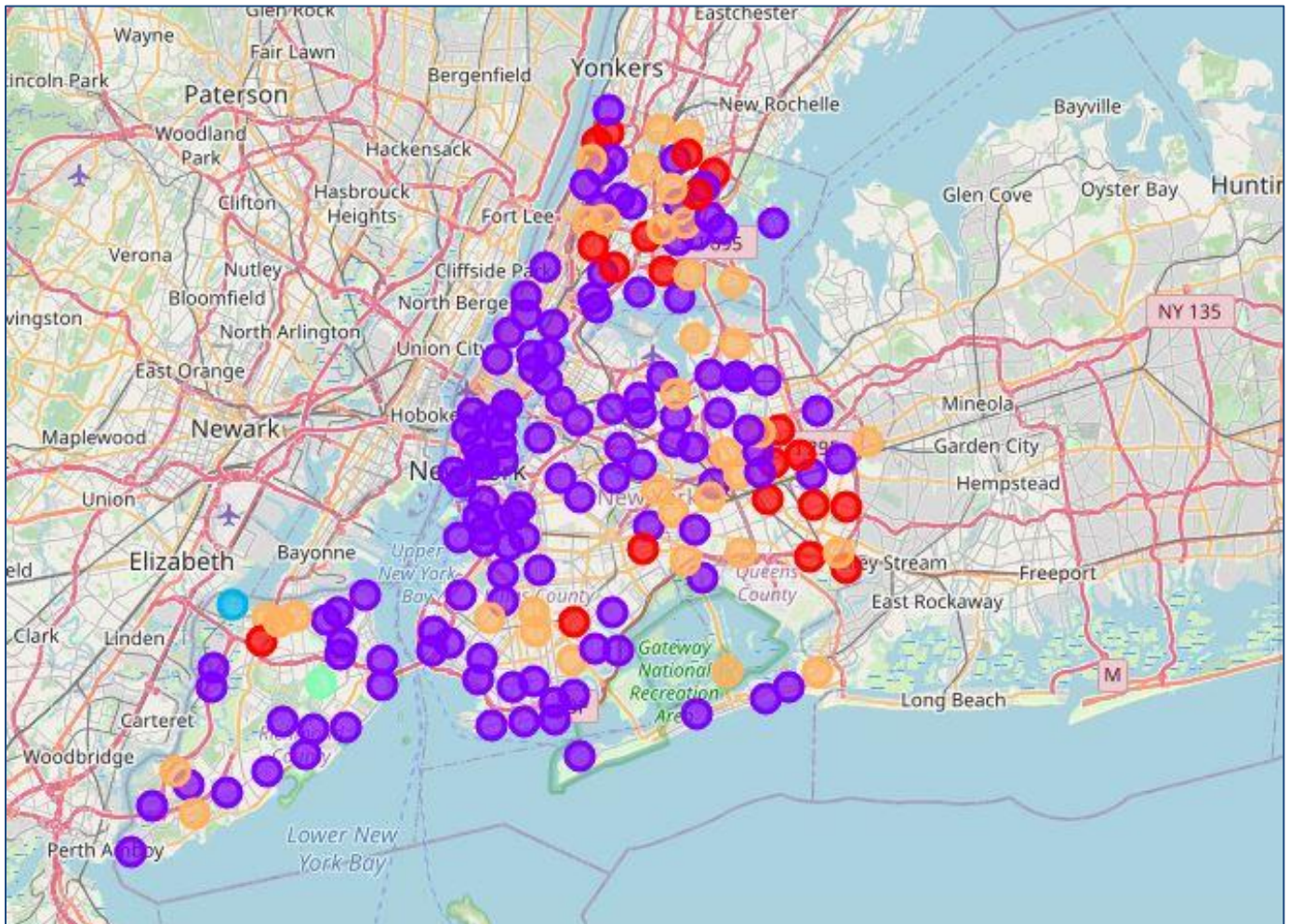**Figure 4.1.1: Clustered Neighbourhoods in Toronto**
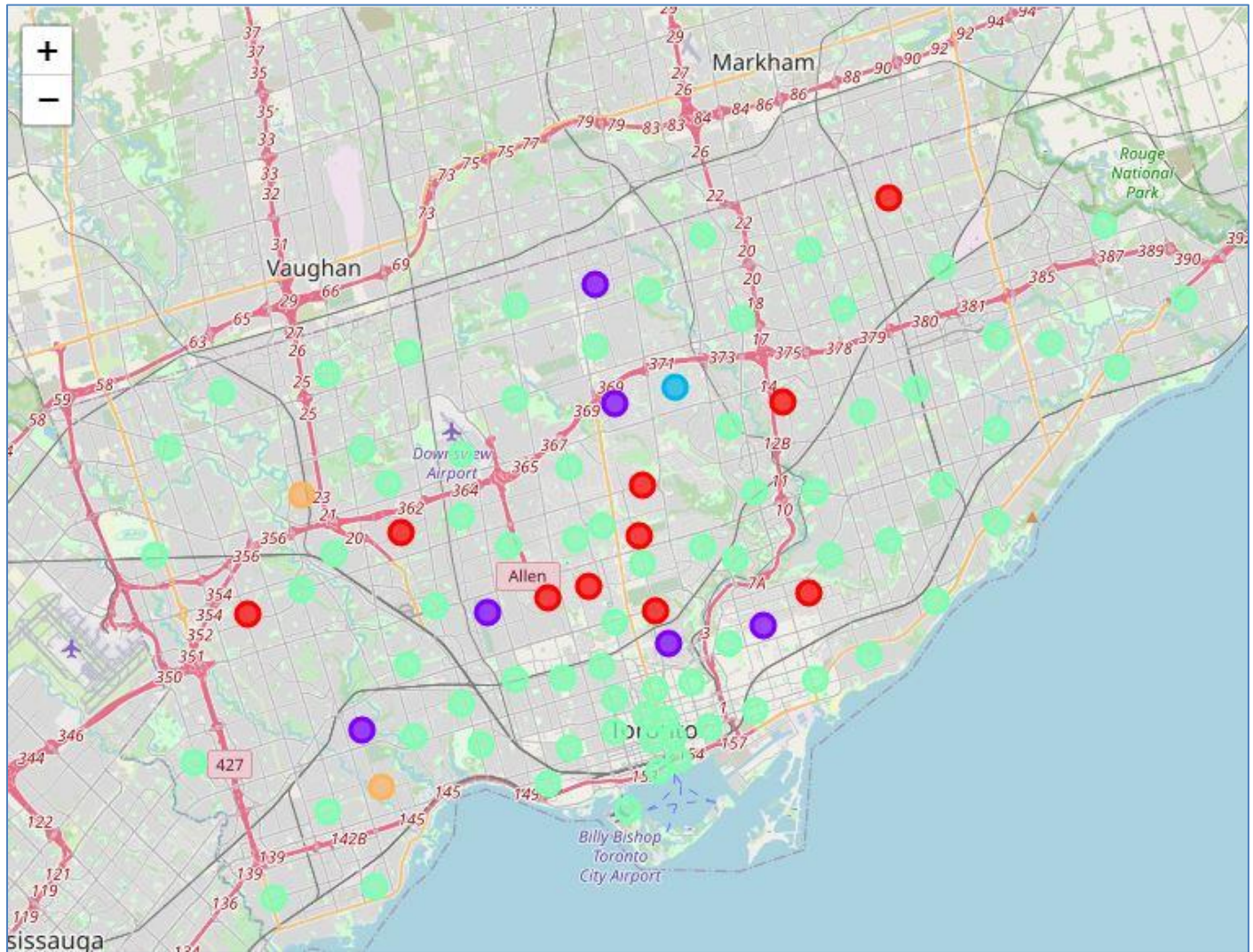
CLUSTER	NATURE	NUMBER OF NEIGHBOURHOODS	AVG COVID-19 CASES PER NEIGBURHOOD
Cluster 1 ●	Recreational cluster with parks, trails, swimming pools, hockey arenas, playgrounds with *a few Asian Restaurants*.	10	505
Cluster 2 ●	Fitness cluster with yoga studios, parks, playgrounds, pools with *no Asian Restaurants*	6	374
Cluster 3 ●	A mixed cluster with ample recreation centres, restaurants and other stores.	1	93
Cluster 4 ●	The largest cluster with many banks, cafes, supermarkets and Japanese, Italian, Thai, Korean, Mexican restaurants.	74	687
Cluster 5 ●	Services cantered cluster with a few Eastern European Restaurants and drug stores.	2	664

4.2 New York City Clusters

Figure 4.2.1 illustrated the 5 clusters of New York City based on its venue categories.

Figure 4.2.1: Clustered Neighbourhoods in New York City

CLUSTER	NATURE	NUMBER OF NEIGHBOURHOODS	AVG COVID-19 CASES PER NEIGBURHOOD
Cluster 1 🔴	Food cluster with a range of restaurants, food trucks, bakeries, cafes, pizza shops and ice cream shops	20	1930
Cluster 2 🟣	The largest and the competitive cluster with different types of restaurants, recreation centres and service centres	124	1685
Cluster 3 🔵	The entertainment cluster with a bar, yoga studio, event service, exhibit and an Ethiopian Restaurant.	1	570
Cluster 4 🟢	Fit & Food cluster with a park, trail and different set of restaurants	1	1495
Cluster 5 🟡	Pizza cluster with a large number of pizza shops and a range of recreational activity centres	35	1895

5. DISCUSSION

As per results shown in the previous chapter, it is crystal clear that all the neighbourhoods have been severely impacted by Covid-19 pandemic, which should be a major concern in decision making when it comes to pick a location to open a new Asian Restaurant.

When we look at the five clusters in Toronto, neighbourhoods in cluster 2 indicate a strong potential in opening a new restaurant as they contain no Asian restaurants, but a set of fitness centres around. That implies, having healthy food options in the menu would definitely enhance the customer base at the beginning and then it can be generalised into the community gradually. On the other side, cluster 2 has the second lowest average Covid-19 cases per neighbourhood. Neighbourhoods in recreational cluster, namely cluster 1 has a few Asian restaurants with 3rd lowest average Covid-19 cases. It would be an interesting to locate a branch there to evaluate the competition. Neighbourhoods in other clusters have a wide range of restaurants and also significant number of Covid-19 cases.

As expected, all the neighbourhoods in New York City have enormous number of Covid-19 cases compared to Toronto. However, neighbourhoods in cluster 5 consist a plenty of pizza shops, but only few other types of restaurants, while cluster 3 has only an Ethiopian restaurant with many entertainment services. All other clusters in New York City encompass board spectrum of restaurants including Japanese, Italian, Thai, Mexican, European and Korean.

6. CONCLUSION

It was an enchanting analysis to figure out the potential locations to position a new Asian Restaurant. The significance of this study is an unavoidable Covid-19 pandemic impact had to consider over the competitor analysis, given the circumstances. With respect to the objectives outlined for the first sprint of the investment project, we were able to identify potential neighbourhoods in Toronto and New York cities with a priority. Below table depicts the priority clusters in both cities.

CITY	*PRIORATISED CLUSTERS*	*DEPRIORATISED CLUSTERS*
Toronto	Cluster 2 & Cluster 1	Clusters 3, 4 & 5
New York City	Cluster 5 & Cluster 3	Clusters 1, 2 & 4

Since we have filtered a set of prospective neighbourhoods, further studies can be conducted to determine the number of branches and specific locations by analysing the following factors.

- Age & Gender wise Menu Preferences
- Ethnicity Diversity
- Competitive Landscape
- Pricing Models
- Transportation

7. REFERENCES

https://www.toronto.ca/home/covid-19/covid-19-latest-city-of-toronto-news/covid-19-status-of-cases-in-toronto

https://open.toronto.ca/catalogue/?search=Covid%2019&sort=score%20desc

https://open.toronto.ca/dataset/covid-19-cases-in-toronto/

https://docs.ckan.org/en/latest/api/index.html

https://www.nytimes.com/interactive/2020/nyregion/new-york-city-coronavirus-cases.html#zipcode

https://www1.nyc.gov/home/search/index.page?search-terms=COvid+CAses

https://www1.nyc.gov/site/doh/covid/covid-19-data-totals.page#zip

https://github.com/nychealth

https://github.com/nychealth/coronavirus-data/tree/master/totals

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

https://www.coursera.org/learn/data-analysis-with-python/home/welcome.

https://www.coursera.org/learn/python-for-applied-data-science-ai/home/welcome

https://www.coursera.org/learn/python-for-data-visualization/home/welcome

https://www.coursera.org/learn/applied-data-science-capstone/home/welcome