

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

Final Report

1/10/2021

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.

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

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-Donauda	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

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

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.