Coursera Capstone Final Project

Teri Nguyen_Shank

<div align="center">**Battle of the Neighborhoods of Toronto**</div>

Introduction

Toronto is the capital city of the Canadian province of Ontario. With a recorded population of 6,197,000 in 2020, it is the most populous city in Canada. It is an international center of business, finance, arts, and culture. It takes a spot in the top ten list of the "World's Most Economically Powerful Cities" according to Forbes Magazine. Toronto attracts investment due to the size of their economy, their projected future prosperity, and cost of living and quality of life. It also is a lively place when it comes to entertainment, culture, and food. The cities are filled with friendly people. You can find world-class theater, shopping, and restaurants here.

Where to live in Toronto? Depends on your preferences and lifestyle, finding a good place to live can be one of the most daunting tasks when moving to Toronto. Suburbs such as Scarborough, North York, and a few others are favored by families and are among the best places to live in Toronto. While many visitors shirk Scarborough as too suburban, it tends to be a popular location for new immigrants to settle because of its highly affordable part of town with some of the cheapest rents and dining options. North York it is one of the six administrative districts of Toronto. It encompasses most of uptown Toronto with its largely residentials and many micro-neighborhoods of its own.

This project offers a brief analysis of the two different Toronto's datasets, aimed at giving you an idea about the boroughs, and the neighborhoods. It will then explore two most dense boroughs - North York/Downtown Toronto and Scarborough and their trending venues.

Data Overview

The data that will be used for this project are collected from multiple sources as Wikipedia, GeoSpatial Data set that has been prepared as s CSV file, and Foursquare API.

1. List of postal codes of Canada: M from Wikipedia: This is a list of postal codes in Canada that is organized in a form of a table where the first letter is M.
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

2. Geographical coordinates of each postal code from a csv file are ready to be extracted. GeoSpatial_Coordinate.CSV
3. Foursquare API: Foursquare location data

https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_token={}&v={}&query={}&radius={}&limit

4. Toronto information
https://en.wikipedia.org/wiki/Toronto

Methodology

1. Explore and segment the boroughs in the city of Toronto based on the postal code and borough information.
- Download and Explore Dataset

- Web scrapping: scraping the Wikipedia page using BeautifulSoup package and wrangle the data, clean it, and then read it into a structural format pandas data frame with 3 columns as bellow.

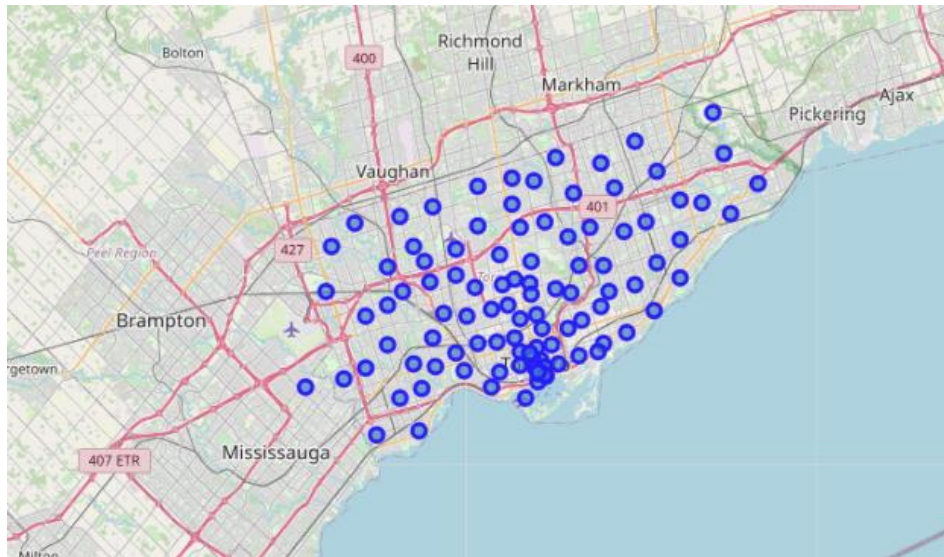|	PostalCode	Borough	Neighborhood
0	M1B	Scarborough	Malvern / Rouge
1	M1C	Scarborough	Rouge Hill / Port Union / Highland Creek
2	M1E	Scarborough	Guildwood / Morningside / West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

- Download csv file using wget command and access data to create the following 3 columns data frame.

	PostalCode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

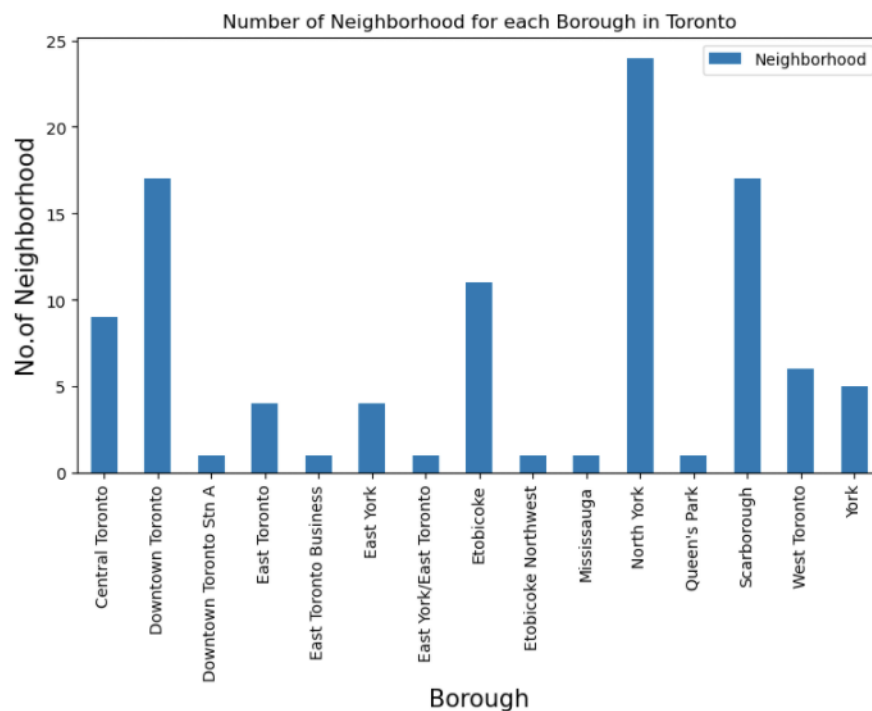- Merge the two above data frames together into one with 5 columns.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern / Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill / Port Union / Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood / Morningside / West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	Kennedy Park / Ionview / East Birchmount Park	43.727929	-79.262029
7	M1L	Scarborough	Golden Mile / Clairlea / Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffside / Cliffcrest / Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff / Cliffside West	43.692657	-79.264848

- Once the data is in a structured format, we analyze it.
- Toronto has a total of 15 boroughs and 103 neighborhoods. To segment the neighborhoods and explore them, we create a dataset that contains the 15 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. Then we use the visualization libraries like Folium and Matplot to visualize the results.
- Using geopy library to get the latitude and longitude values of Toronto and then create map of Toronto with Folium using those values.
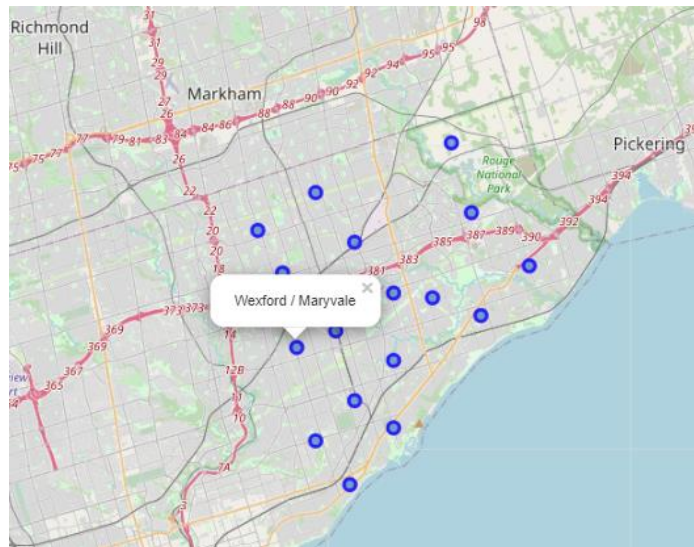
Blue dots represent boroughs and neighborhoods.

- Using matplot library to plot the boroughs with its neighborhoods on a bar graph.

- According to the graph above the analysis is then narrowed down to two 2 boroughs - Scarborough and North York - due to the highest count of neighborhoods.
2. Define Foursquare API criteria and create get venues functions.
   The idea is to leverage the Foursquare location data to explore different neighborhoods within two boroughs to come up with different sets of venues within a radius of 500 meters.
- Define limit and radius
  LIMIT = 100 # limit of number of venues returned by Foursquare API
  radius = 500 # define radius
- Define Foursquare Credentials and Version

- Create 2 functions to repeat the same process to all the boroughs and neighborhoods in Toronto.
  get_category_type(): extracts the category of the venues.
  getNearbyVenues(): extracts the nearby venues.
3. Explore Scarborough.
- Create a new data frame of the Scarborough data.
- Get the latitude and longitude coordinates of Scarborough using geopy library.
- Map Scarborough with 17 neighborhoods

Blue dots represent neighborhoods.

- Define the corresponding URL, define limit and radius, and get request to Foursquare for venues categories.
  LIMIT = 100 # limit of number of venues returned by Foursquare API
  radius = 500 # define radius
- Using get_category_type() to get venues and their categories in Scarborough and then clean and filter venues into a data frame with 4 columns: Postal Code, Boroughs, Neighborhoods, Latitudes, and Longitudes

	name	categories	lat	lng
0	Hot Topic	Clothing Store	43.775450	-79.257929
1	Tommy Hilfiger	Clothing Store	43.776015	-79.257369
2	New York Fries - Scarborough TC	Restaurant	43.775135	-79.256858
3	lululemon athletica	Clothing Store	43.775517	-79.258423
4	CANBE Foods Inc	Indian Restaurant	43.773546	-79.246082
5	Subway	Sandwich Place	43.771477	-79.252805
6	Subway	Sandwich Place	43.775102	-79.257115
7	Petro-Canada	Gas Station	43.769724	-79.261677
8	St. Andrews Fish & Chips	Fish & Chips Shop	43.771865	-79.252645

Foursquare returned 39 venues

- Using getNearbyVenues() to extract number of venues in Scarborough then use groupby function to filter and group venues into a list of unique venues returned by Foursquare.

89 venues were returned resulting in 56 unique categories.

- Visualize Scarborough's venues on the map using Folium.

- Analyze these neighborhoods to get the top 10 common venues using One hot coding algorithm.

```
----Agincourt)----
                       venue  freq
0   Latin American Restaurant  0.25
1                      Lounge  0.25
2              Breakfast Spot  0.25
3              Clothing Store  0.25
4         American Restaurant  0.00
5                   Pet Store  0.00
6         Korean BBQ Restaurant  0.00
7              Medical Center  0.00
8                Metro Station  0.00
9          Mexican Restaurant  0.00
```

Top 10 trending venues in Agincourt

- Explore the first neighborhood in Scarborough and its venues – Malvern/Rouge.

	name	categories	lat	lng
0	Kim Kim restaurant	Chinese Restaurant	43.753833	-79.276611
1	Kairali	Indian Restaurant	43.754915	-79.276945
2	Karaikudi Chettinad South Indian Restaurant	Indian Restaurant	43.756042	-79.276276
3	Pho Vietnam	Vietnamese Restaurant	43.757770	-79.278572
4	Big Al's Pet Supercentre	Pet Store	43.759279	-79.278325

- Send a request to the API to search for a particular venue – Vietnamese restaurant - within 500 meters radius of Scarborough.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
49	Dorset Park / Wexford Heights / Scarborough To...	43.75741	-79.273304	Pho Vietnam	4bc75a3c6501c9b630433e29	43.75777	-79.278572	Vietnamese Restaurant

- Get the venue's overall rating. Make a call to the Foursquare API with venue ID listed above to get the ratings. Pho Vietnam has a rating of 5.9.
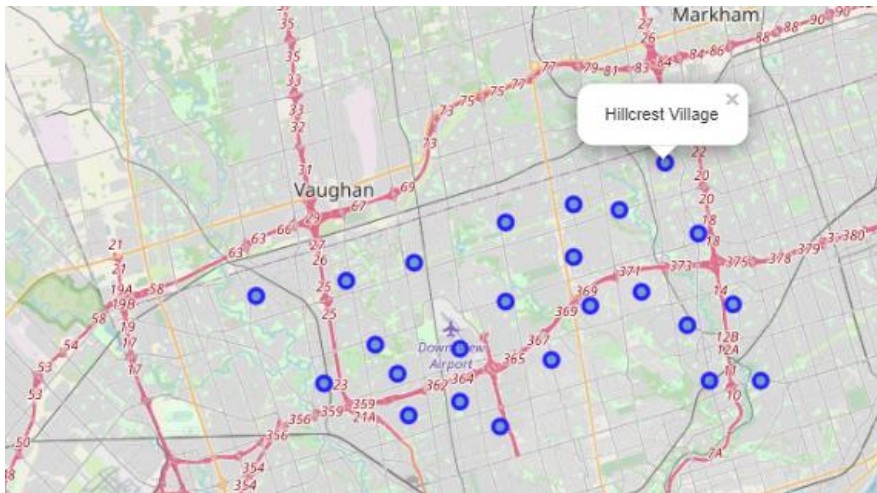
4. Explore North York
   We do the same process and data analysis as the above borough.
- Data frame of North York along with Postal code, Neighborhood, Latitude, Longitude

- Map North York with 24 neighborhoods according to its latitude and longitude
- Define the corresponding URL and get request.
- Repeat the process of calling functions get_category_type() and getNearbyVenues(): to extract categories and venues then filter and group venues into a list of unique venues returned by Foursquare.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M2H	North York	Hillcrest Village	43.803762	-79.363452
1	M2J	North York	Fairview / Henry Farm / Oriole	43.778517	-79.346556
2	M2K	North York	Bayview Village	43.786947	-79.385975
3	M2L	North York	York Mills / Silver Hills	43.757490	-79.374714
4	M2M	North York	Willowdale / Newtonbrook	43.789053	-79.408493
5	M2N	North York	Willowdale)Sout	43.770120	-79.408493
6	M2P	North York	York Mills West	43.752758	-79.400049
7	M2R	North York	Willowdale)Wes	43.782736	-79.442259
8	M3A	North York	Parkwoods	43.753259	-79.329656
9	M3B	North York	Don Mills)Nort	43.745906	-79.352188

Blue dots represent neighborhoods.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
0	Hillcrest Village	43.803762	-79.363452	New York Fries	4d8771f6651041bd5e499b30	43.803664	-79.363905	Fast Food Restaurant
1	Hillcrest Village	43.803762	-79.363452	Eagle's Nest Golf Club	4ad9dce6f964a520651b21e3	43.805455	-79.364186	Golf Course
2	Hillcrest Village	43.803762	-79.363452	AY Jackson Pool	4e1fad69fa761d67106a9989	43.804515	-79.366138	Pool
3	Hillcrest Village	43.803762	-79.363452	Villa Madina	5331a799498e3bfbae4a5a04	43.801685	-79.363938	Mediterranean Restaurant
4	Hillcrest Village	43.803762	-79.363452	Duncan Creek Park	4c64ca1b772bef3bfadbc2aa	43.805539	-79.360695	Dog Run
...	...	...	...	...	...	...	...	...
239	North Park / Maple Leaf Park / Upwood Park	43.713756	-79.490074	Queen's GreenBelt Park	4f8ba717121d29fd94277ce8	43.712355	-79.484754	Trail
240	North Park / Maple Leaf Park / Upwood Park	43.713756	-79.490074	Mika's Trim	590bdb6bbe70781ade1dab8a	43.714068	-79.496113	Construction & Landscaping
241	Humber Summit	43.756303	-79.565963	HNS ARARAT	54fb21be498e513a0a853128	43.757519	-79.563653	Furniture / Home Store
242	Humber Summit	43.756303	-79.565963	Paul Davis Canada of York Region	57b20e5e498e2f95667c4fd6	43.755852	-79.560256	Home Service
243	Humberlea / Emery	43.724766	-79.532242	Strathburn Park	4e38862f1f6ecc2de159fb4d	43.721765	-79.532854	Baseball Field

Foursquare returned 244 venues with 101 unique categories

```
----Bathurst Manor / Wilson Heights / Downsview North--
                       venue  freq
0                       Bank  0.09
1                Coffee Shop  0.09
2  Middle Eastern Restaurant  0.05
3                Supermarket  0.05
4                Pizza Place  0.05
5                   Pharmacy  0.05
6                 Restaurant  0.05
7                      Diner  0.05
8             Sandwich Place  0.05
9               Deli / Bodega  0.05
```
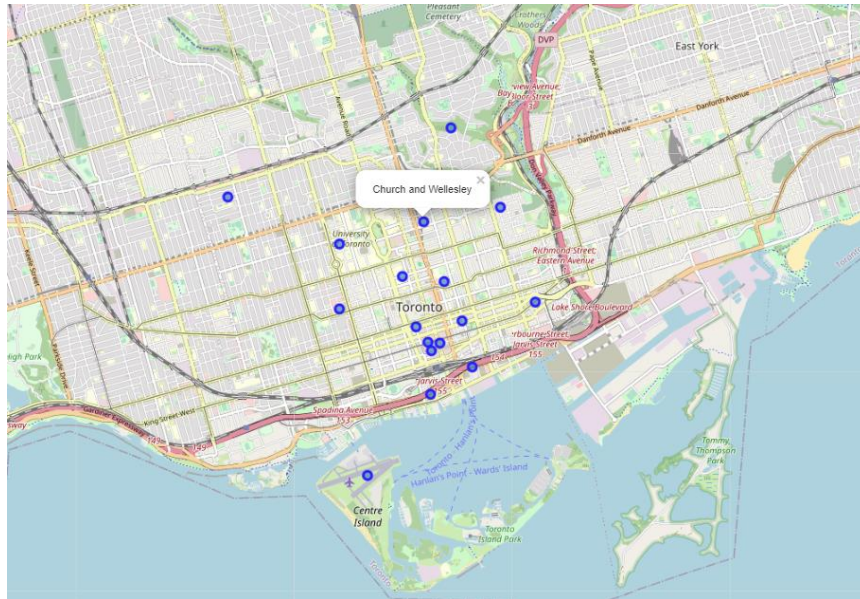
> Analyze the neighborhoods to get the top common 10 venues using a machine learning algorithm called One hot coding. This is the top 10 trending venues in this neighborhood.

- Explore the first neighborhood in North York and its venues – Fairview / Henry Farm / Oriole
- Send a request to the API to search for a specific type of venues – Vietnamese restaurant - in North York within 500 meters radius.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
89	Willowdale)Sout	43.770120	-79.408493	Pho 88 Vietnamese Cuisine	4b4779f0f964a520543326e3	43.770456	-79.413064	Vietnamese Restaurant
150	Northwood Park / York University	43.767980	-79.487262	Phở Liên Hương	51030632e4b0840c0152c004	43.764437	-79.489775	Vietnamese Restaurant
199	Lawrence Manor / Lawrence Heights	43.718518	-79.464763	Lac Vien Vietnamese Restaurant	4ccc5aebee23a14370591ea8	43.721259	-79.468472	Vietnamese Restaurant

- Get the venue's overall rating. Make a call to the Foursquare API with venue ID listed above to get their rating. Pho 88 has a rating of 6.1 and Pho Lac Vien has a rating of 7.6 while Pho Lien Huong has none.
5. Explore Downtown Toronto
   We do the same data analysis process as the above boroughs.
- Data frame of Downtown Toronto along with Postal code, Neighborhood, Latitude, and Longitude
- Map Downtown Toronto with its according latitude and longitude
- Define the corresponding URL and get request.
- Repeat the process of calling functions get_category_type() and getNearbyVenues(): to extract categories and venues then filter and group venues into a list of unique venues returned by Foursquare.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
1	M4X	Downtown Toronto	St. James Town / Cabbagetown	43.667967	-79.367675
2	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
3	M5A	Downtown Toronto	Regent Park / Harbourfront	43.654260	-79.360636
4	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
5	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
6	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
7	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
8	M5H	Downtown Toronto	Richmond / Adelaide / King	43.650571	-79.384568
9	M5J	Downtown Toronto	Harbourfront East / Union Station / Toronto Is...	43.640816	-79.381752

Blue dots represent the neighborhoods.

	name	categories	lat	lng
0	Blaze Pizza	Pizza Place	43.656518	-79.380015
1	Silver Snail Comics	Comic Shop	43.657031	-79.381403
2	Burrito Boyz	Burrito Place	43.656265	-79.378343
3	The Elm Tree Restaurant	Modern European Restaurant	43.657397	-79.383761
4	UNIQLO ユニクロ	Clothing Store	43.655910	-79.380641
...	...	...	...	...
91	Sol'exotica Tanning Spa	Tanning Salon	43.654562	-79.376663
92	Starbucks	Coffee Shop	43.654841	-79.375815
93	Somethin' 2 Talk About	Middle Eastern Restaurant	43.658395	-79.385338
94	Booster Juice	Smoothie Shop	43.652658	-79.378459
95	Bubble Bath & Spa	Spa	43.659050	-79.385344

96 rows × 4 columns

- Foursquare finally returned 202 unique categories out of 1102 venues

```
----Berczy Park----
                venue  freq
0          Coffee Shop  0.10
1         Cocktail Bar  0.05
2               Bakery  0.05
3    Seafood Restaurant  0.03
4          Cheese Shop  0.03
5           Restaurant  0.03
6             Beer Bar  0.03
7       Farmers Market  0.03
8                Diner  0.02
9                 Café  0.02
```

Analyze the neighborhoods to get the top common 10 venues using a machine learning algorithm called One hot coding. This is the top 10 venues in this neighborhood.

- Send a request to the API to search for a specific type of venues – Vietnamese restaurant - in Downtown Toronto within 500 meters radius. Banh Mi Nguyen Hung has the top rating of 8.9 followed by Asian Bowl: 6.8 and Pho Tien: 6.1

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
219	Garden District, Ryerson	43.657162	-79.378937	Fine Asian Bowl	4aea2b14f964a520ebb921e3	43.655387	-79.380326	Vietnamese Restaurant
905	Kensington Market / Chinatown / Grange Park	43.653206	-79.400049	Banh Mi Nguyen Huong	4af45a8af964a52097f121e3	43.653628	-79.398376	Vietnamese Restaurant
960	Kensington Market / Chinatown / Grange Park	43.653206	-79.400049	Saigon Lotus Restaurant	58a7c100076be13f60d1dff5	43.654311	-79.399225	Vietnamese Restaurant
961	Kensington Market / Chinatown / Grange Park	43.653206	-79.400049	Pho Tien	50e6282de4b026ac78e0ee70	43.652832	-79.406242	Vietnamese Restaurant

Results

	Scarborough	North York	Downtown Toronto
Neighborhoods	17	24	17
Total Venues	89	244	1102
Unique Categories	56	101	202
Vietnamese restaurants	1	3	4
Rating	5.9	7.6, 6.1, and 0	8.9, 6.8, and 6.1

According to my data analysis and the above table, Downtown Toronto has the most venues and options for all types of cuisine, but since North York surrounds most of uptown Toronto, it is the best borough between the 2.

Discussion/Conclusion

Toronto is one of the world's most multicultural cities and famous for a diverse choice of food. Smaller suburbs surround the large cities like Scarborough's neighborhoods offering options for people who prefer to step back from big city life while uptown neighborhoods in North York offers a wider selection of necessities. The two boroughs are diverse and culturally unique, varying from a bohemian hangout to a foodie paradise. The objective of this analysis is to help newcomers compare the 2 boroughs, within the boroughs segmented into each neighborhood to discover the top trending venues, and finally narrow it down to one cuisine.