**Comparing Boroughs in Toronto, Ontario, Canada to Boroughs in New York City, New York, USA**

Capstone Project - The Battle of the Neighborhoods (Week 2)

Applied Data Science Capstone by IBM/Coursera

**Table of contents**

Introduction: Business Problem

* Data
* Methodology
* Analysis
* Results and Discussion
* Conclusion

**Introduction: Business Problem**

In this project we will try to find which neighborhoods in Ney York City, New York most closely resemble which neighborhoods in Toronto, Ontario, Canada. Specifically, this report will be targeted to stakeholders interested in moving from Toronto, Ontario, Canada to New York City, New York. There are numerous factors that make up a neighborhood, but in this study to focus on comparing the number of restaurants and proximity to major airports in Toronto and New York.

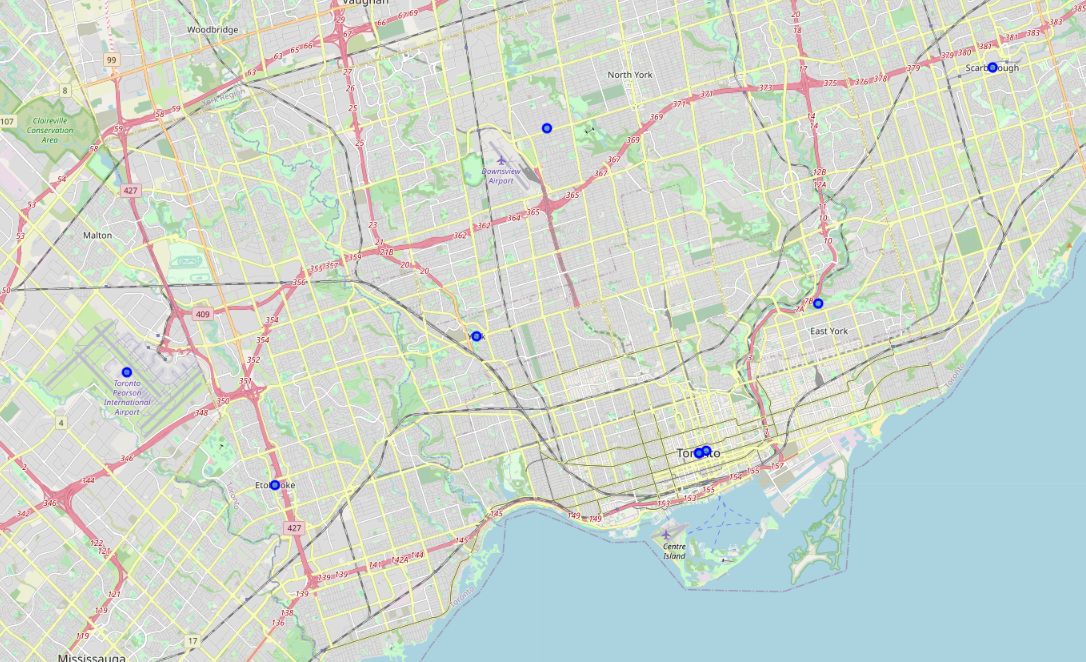
**Data**

Based on definition of our problem, factors that will influence our decision are:

* number of existing restaurants in the neighborhood (any type of restaurant)
* number of and distance to major airports

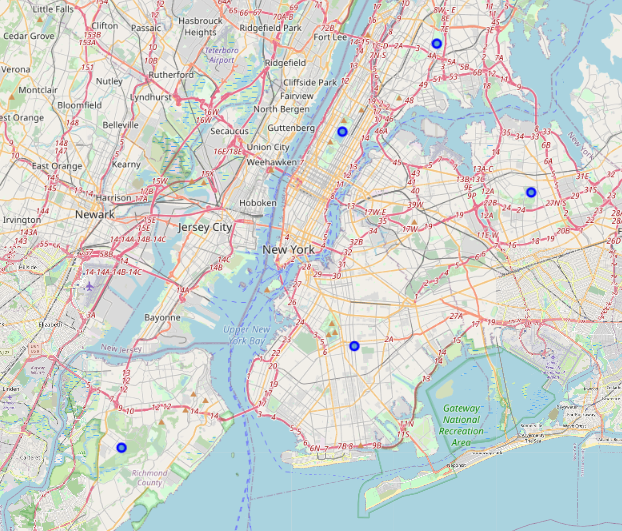
We decided to use established boroughs within each city (10 in Toronto, 5 in New York) to define our neighborhoods. The Toronto boroughs are:

* Central Toronto
* Downtown Toronto
* East Toronto
* East York
* Etobicoke
* Mississauga
* North York
* Scarborough
* West Toronto
* York



The established boroughs for New York are:

* Bronx
* Brooklyn
* Manhattan
* Queens
* Staten Island



The following data sources will be needed to extract/generate the required information:

* Coordinates for the centroid areas or each borough will be requested and used using geolocator library in Python.
* The number of restaurants in every neighborhood will be obtained using Foursquare API
* Coordinates of each major airport will be requested and used using geolocator library in Python.

**Methodology**

For this project, we will be looking at the number of each type of restaurant in the boroughs of both Toronto and New York, and then putting each borough in Toronto into one of five categories: each category is a borough in New York. Since we would theoretically be moving from Toronto to New York, we will use the boroughs of New York as the category headers and the boroughs of Toronto into each category based on how closely they match each borough of New York.

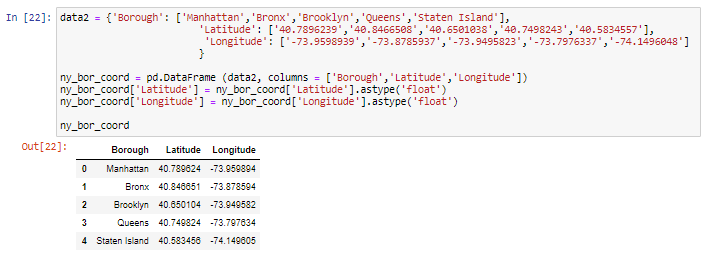
|  |  |
| --- | --- |
| New York Borough | Corresponding Toronto Borough |
| Bronx | ??? |
| Brooklyn | ??? |
| Manhattan | ??? |
| Queens | ??? |
| Staten Island | ??? |

We will then do the same analysis for the boroughs of Toronto and New York based on the distance from the center of each borough to the single largest airport in the area. Again, each borough in Toronto will be placed in a category defined by a borough in New York to show which boroughs in Toronto most closely resemble which borough in New York.

The first step will always be to import the appropriate libraries into Python so we can begin our analysis.



We will need exact latitude and longitude coordinates in order to use the FourSquare API in Python. After using the libraries we imported earlier, we can build two dataframes that we will use to power the FourSquare API.

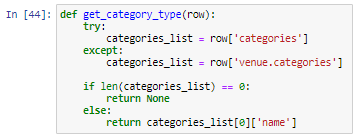




The FourSquare API allows us to pull in information about venues based on coordinates. To set this up, we will use the following code to pull in venue names, categories, and latitude/longitude coordinates.



After submitting the request to the API, we will need to set up a usable dataframe and clean the data. Once we have the columns named and cleaned, we can submit a request for each set of coordinates in both New York and Toronto.

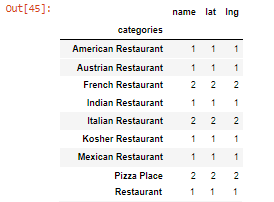




The resulting dataframes for each request are quite large, mostly due to the density of both cities.



We can filter the results as needed, and in this case, we find that there are 12 total restaurants in the Bronx.

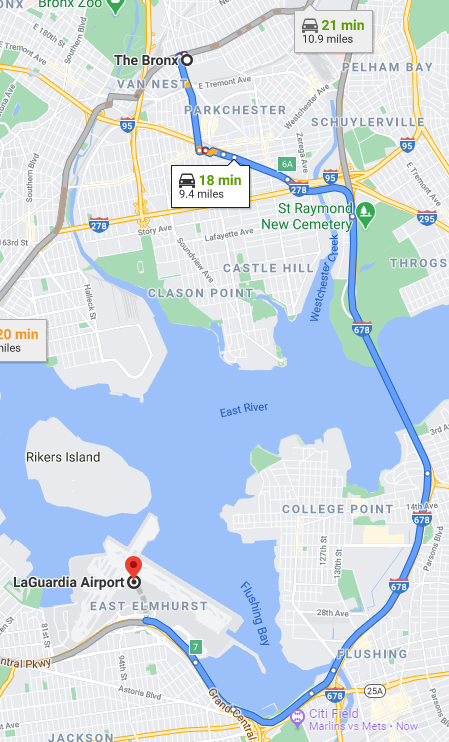


Next, using the same method for each borough, let's count the number of restaurants.





The second part of the analysis will require us to map the driving distance from the center point of each borough to the airport. For the sake of this analysis, we will use LaGuardia International Airport in New York and Toronto Pearson International Airport in Toronto. We will use the Google Maps driving distance as the standard for each measurement.



After mapping the distance between each borough and the airport, the results are as follows:

|  |  |
| --- | --- |
| New York Borough | Dist. To Airport |
| Bronx | 9.4 miles |
| Brooklyn | 12.6 miles |
| Manhattan | 7.9 miles |
| Queens | 6.3 miles |
| Staten Island | 25.9 miles |

|  |  |
| --- | --- |
| Toronto Boroughs | Dist. To Airport |
| East York | 22.4 miles |
| Scarborough | 21.1 miles |
| East Toronto | 17.6 miles |
| Downtown Toronto | 16.4 miles |
| Central Toronto | 16.2 miles |
| North York | 12.7 miles |
| York | 11.8 miles |
| Mississauga | 9.7 miles |
| Etobicoke | 9.3 miles |
| West Toronto | 8.4 miles |

All results are noted in miles instead of miles and kilometers so that we have a true comparison between the distances.

**Analysis**

Now that we have our restaurant counts and distances to the airport, we can start to place the Toronto boroughs into the appropriate categories.

For the restaurant counts, the corresponding boroughs are listed below:

|  |  |
| --- | --- |
| New York Borough | Corresponding Toronto Borough |
| Bronx | Downtown Toronto, Scarborough, West Toronto |
| Brooklyn | Central Toronto, Mississauga |
| Manhattan | East Toronto |
| Queens | North York |
| Staten Island | East York, Etobicoke, York |

As you can see, the Bronx and Staten Island had the most matches with Brooklyn a close second. It is a good sign to see that every Toronto borough can be very closely matched with a corresponding Toronto borough in this instance.

Next we looked at driving distance from the center of each borough to the major airport. These results were more concentrated than the restaurant number results we saw. There were two ways to divide the Toronto boroughs into each category, and both methods are shown here.

The first method was to stratify the Toronto boroughs into pairs after the boroughs were sorted from furthest to shortest distance to the airport. This yielded two boroughs in each category, regardless of the actual distance similarity.

|  |  |
| --- | --- |
| New York Borough | Corresponding Toronto Borough |
| Staten Island | East York, Scarborough |
| Brooklyn | East Toronto, Downtown Toronto |
| Bronx | Central Toronto, North York |
| Manhattan | York, Mississauga |
| Queens | Etobicoke, West Toronto |

After completing this specific methodology, we needed a more accurate way to categorize the Toronto boroughs based on distance. Thus, we categorized the Toronto boroughs based on the closest match to the New York boroughs. This yielded more accurate results that better fit the spirit of the analysis.

|  |  |
| --- | --- |
| New York Borough | Corresponding Toronto Borough |
| Staten Island | East York, Scarborough |
| Brooklyn | East Toronto, Downtown Toronto, Central Toronto, North York |
| Bronx | Mississauga, Etobicoke |
| Manhattan | West Toronto |
| Queens | West Toronto |

**Conclusion**

Purpose of this project was to identify which boroughs in Toronto would most closely resemble the 5 boroughs in New York. This would allow any stakeholder moving from Toronto to New York to have a much better understanding of where to look for housing based on restaurant amenities and proximity to the international airport. By calculating restaurant density using FourSquare data we have identified and categorized the Toronto boroughs into the appropriate categories. This will allow stakeholders to make more informed decisions when moving, taking a lot of the stress of that activity away.