Toronto and NYC comparison study

Zeyad Al Mothafar

# Introduction

This project aims to compare the restaurants in both Toronto and NYC (Manhattan to be specific).

As a food enthusiast I had the curiosity to dig deeper and know more about the restaurants in NYC and Toronto, since I lived in NYC for almost 2 years and tried a variety of restaurants. On the other hand, I have an interest in Toronto and its rigid well known variety of restaurants of different cultural backgrounds.

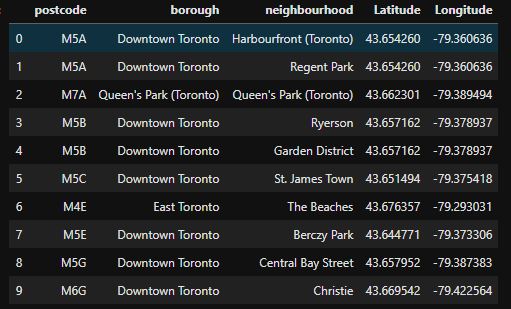
One of the things to consider, is the variance and uniqueness when it comes to the categories of restaurants in both cities. As a result, food enthusiasts can know which city has a wider range of options for restaurants on a multicultural perspective.

Categorizing the restaurants into different clusters based on their price categories, ratings, likes count, longitude and latitude is one of the aspects that’s shown in this study.

Another interesting thing to take a look at is the counts of the top restaurants' categories in both cities so that if a contractor wants to open a restaurant, knows which city doesn’t have that many restaurants from a specific category so that it might have a bigger customer base as well as profits.

# The Data

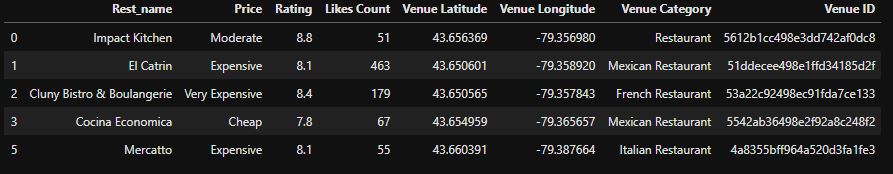
## Toronto Dataset:



Then the data for the restaurants is retrieved using the Foursquare API on the first Toronto dataset by retrieving the nearby venues for each neighborhood and then filtering out the restaurants from the venues.



Then for a subset of 200 restaurants the API gets called using the restaurants ID’s to retrieve each restaurant details.



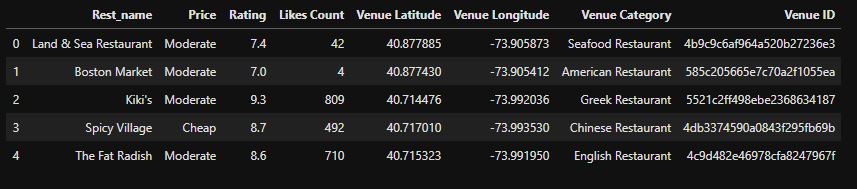
## New York City Dataset:



Then the data for the restaurants is retrieved using the Foursquare API on the first NYC dataset by retrieving the nearby venues for each neighborhood and then filtering out the restaurants from the venues.



Then for a subset of 200 restaurants the API gets called using the restaurants ID’s to retrieve each restaurant details.



# Methodology:

Datasets for both cities retrieved and formatted from the resources provided at the resources section using the “Requests” and “JSON” Libraries in Python then the datasets get converted into Pandas DataFrames.

Nearby venues for each neighborhood in the cities retrieved using the Foursquare API in order to get the restaurants in all of the neighborhoods. After that is filtering the restaurants from the venues and retrieving the details for the restaurants including price, likes count and rating.

Next is Clustering a subset of 200 restaurants in both cities (due to API limitations) based on price, likes count, rating, latitude and longitude.

In order to start with clustering the restaurants, the price category needs to be “One-hot encoded” then the whole data of the restaurants gets normalized using the “Standardscaler “library in “Sklearn” followed by using K-Means clustering with k=4 and analyzing the clusters attributes then visualizing the clusters on the map.

# Results

Example of the restaurant categories in both Toronto and NYC:

Toronto:

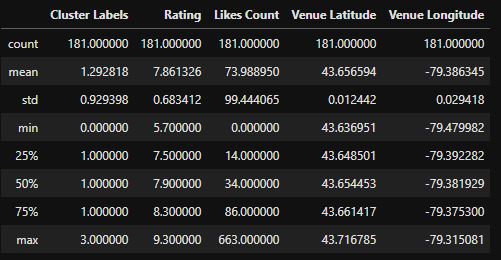
* Italian Restaurants
* Vegetarian / Vegan Restaurants
* American Restaurants
* Japanese Restaurants
* Thai Restaurants

NYC:

* Sushi Restaurants
* Chinese Restaurants
* French Restaurants
* Seafood Restaurants
* Indian Restaurants

Knowing that, there exists about fifty restaurant categories in Toronto vs. seventy-six existing in NYC.

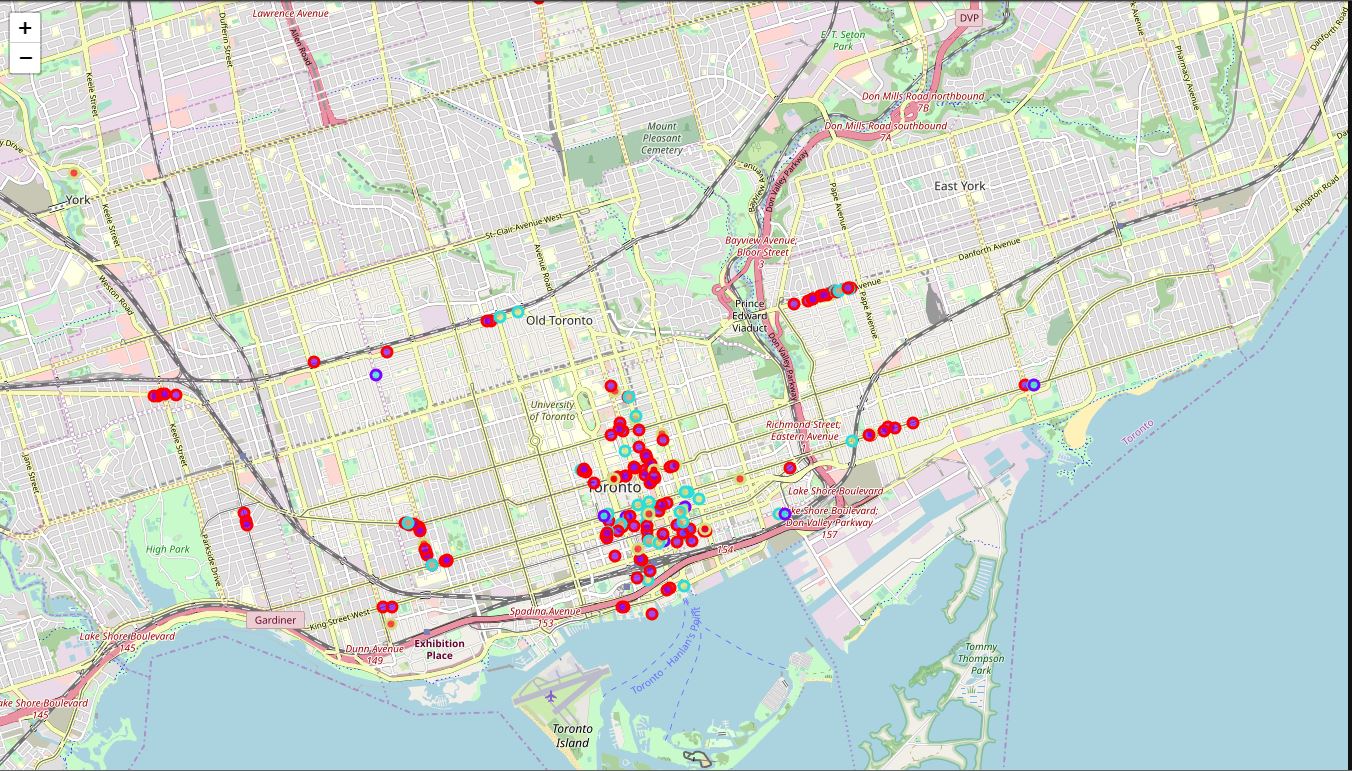
Some stats about the cluster labels, restaurants ratings and likes count in Toronto:



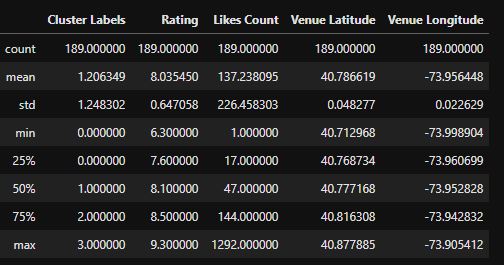
Count of each price category in Toronto:

|  |  |
| --- | --- |
| Price Category | Count |
| Cheap | 25 |
| Moderate | 114 |
| Expensive | 35 |
| Very Expensive | 7 |

Visualizing of the restaurants 4 clusters:



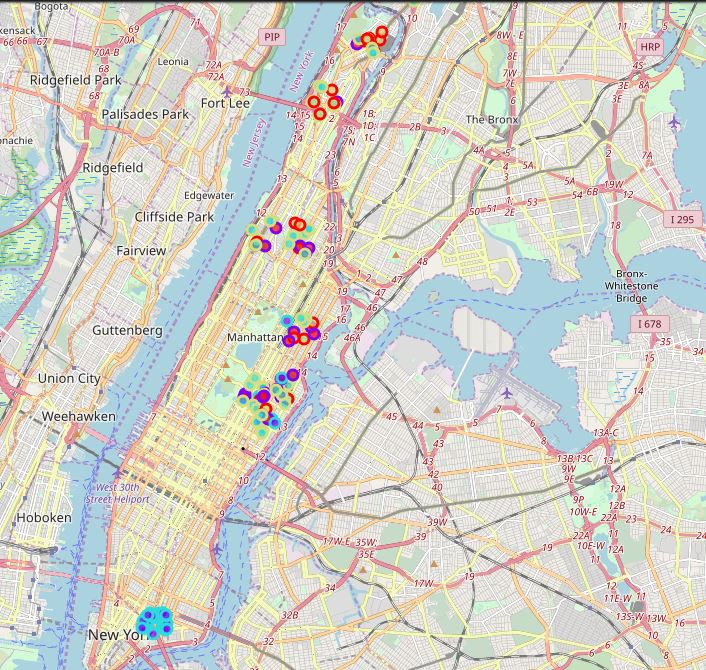
Some stats about the cluster labels, restaurants ratings and likes count in NYC:



Count for each price category in NYC:

|  |  |
| --- | --- |
| Price Category | Count |
| Cheap | 42 |
| Moderate | 105 |
| Expensive | 32 |
| Very Expensive | 10 |

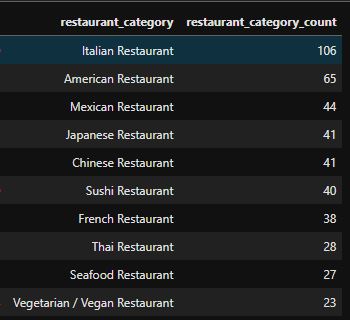
Visualizing the restaurants 4 clusters:



Top restaurants categories and their counts in Toronto:



Top restaurants categories and their counts in NYC:



## Discussion:

Based on the results we can say that:

* If a tourist wants the cheaper city NYC would win with 42 cheap restaurants out of the 200 are and 25 cheap restaurants in Toronto.
* If a tourist wants to look at likes and rating of restaurants NYC wins with an average of 8 for restaurant rating and average of 137 likes count.
* If a tourist wants to try Greek food Toronto wins with 29 Greek restaurants while NYC has less or equal than 23.
* If a tourist wants to try a French restaurant NYC wins with 38 restaurants while Toronto has less or equal to 28 French restaurants.
* For more variety choice of restaurants NYC is the winner with 70 unique restaurant categories while Toronto have 50 unique restaurants categories.
* For contractors that want to build a new restaurant and wants to decide in which city they can look at the Top 10 restaurant categories and their counts, for example if the restaurant is Vegan restaurant NYC might have better future potentials for the contractor because there won’t be as many vegan restaurants as in Toronto.
* We can notice that restaurants in general are more clumped up near each other in NYC and more spread in Toronto.

# Conclusion

In the end, NYC looks like a winner with restaurants and multicultural choices for tourists but that big difference in unique categories of restaurants give the contractors an opportunity to open new restaurants in Toronto that might not even have a competitor in the same category.

# References

1. <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>
2. <http://cocl.us/Geospatial_data>
3. <https://cocl.us/new_york_dataset>