Toronto and NYC comparison study

Zeyad Al Mothafar

# Introduction

This project aims to study some differences of Toronto and NYC (Manhattan specifically) which are two big cities in the world. The first thing to take a look at, is different unique categories of restaurants in both cities, so that if someone wants to know which city has wider range of multicultural choices for restaurants to try out.

The second study will be on clustering the restaurants of both cities based on their price categories, ratings and likes count, getting the counts for each pricing category in each city and then visualizing the clusters on the map.

The last thing to take a look at is the counts of the top restaurants categories in both cities so that if a contractor from a specific nationality wants to open a restaurant knows which city doesn’t have that many restaurants from their city so that it might have bigger customer base and profits.

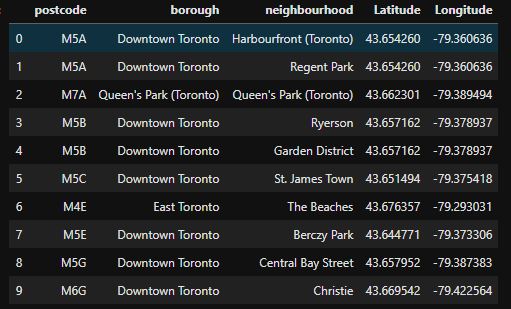
# The Data

## Toronto Dataset:

The Toronto dataset source:

<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> combined with

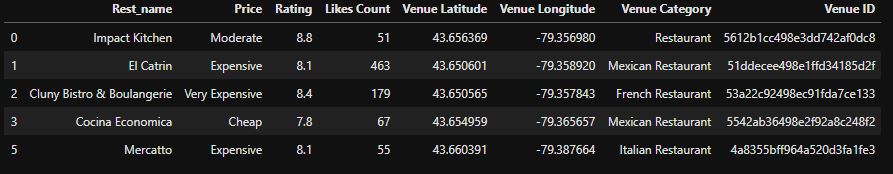
<http://cocl.us/Geospatial_data>



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:

The NYC dataset source:

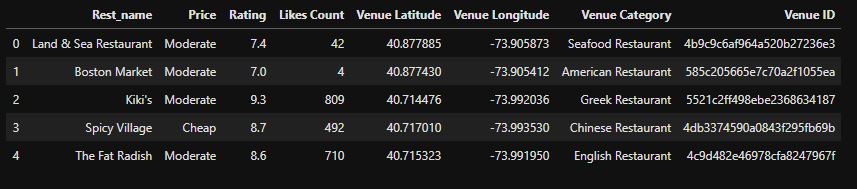
<https://cocl.us/new_york_dataset>



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.



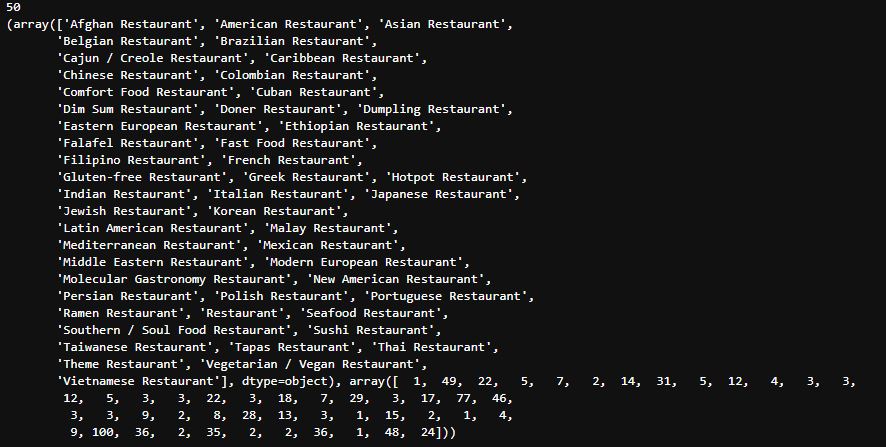
# Methodology:

1. Datasets for both cities retrieved using the “Requests” and “JSON” Libraries in Python.
2. Datasets converted into Pandas data frames.
3. Nearby venues for each neighborhood in the cities retrieved using the Foursquare API.
4. Filtering out the restaurants from the venues.
5. Retrieving the details for the restaurants.
6. Getting the count of unique restaurants categories.
7. Clustering a subset of 200 restaurants in both cities (due to API limitations) based on price, likes count, rating, latitude, longitude using K-Means clustering with k=4.
8. Studying and analyzing the clusters attributes.
9. Visualizing the clusters of the restaurants on the map.

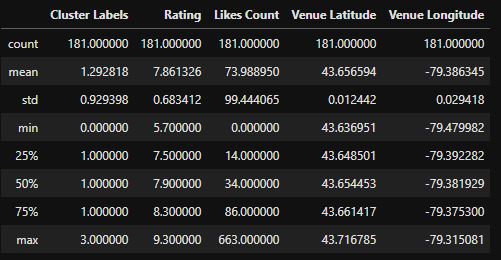
# Results

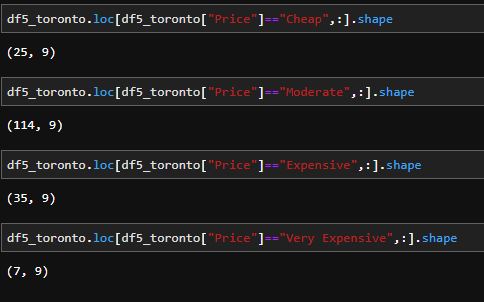
## Toronto results:

Fifty unique restaurants categories in Toronto with their counts:

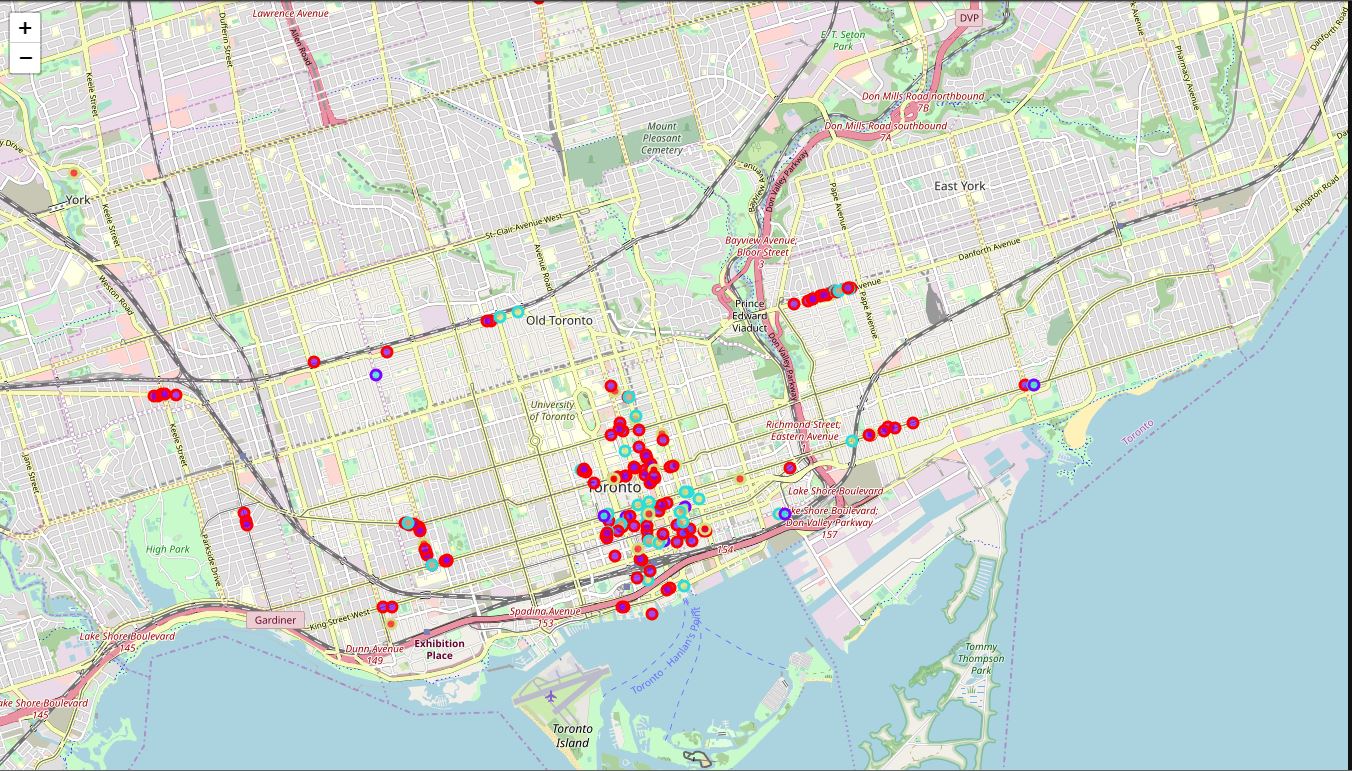


Some stats about the cluster labels, restaurants ratings and likes count:



Count of each price category in Toronto:

Visualizing of the restaurants 4 clusters:

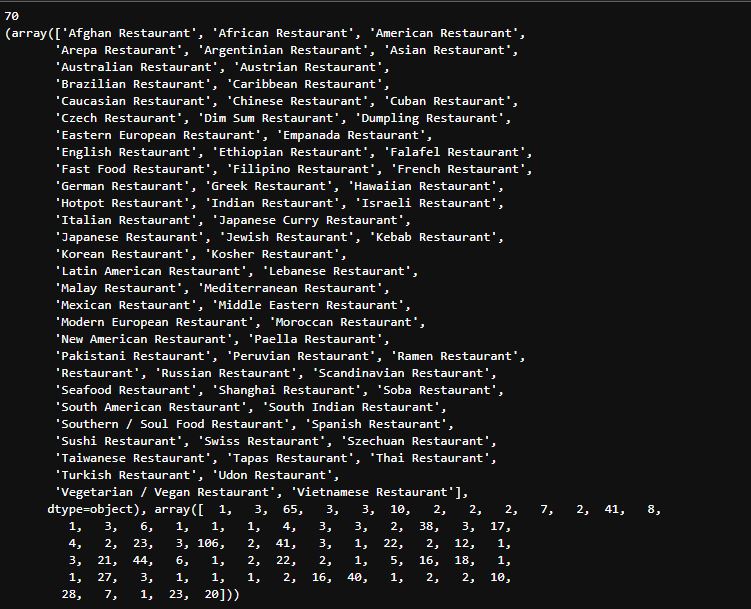


Top restaurants categories and their counts:

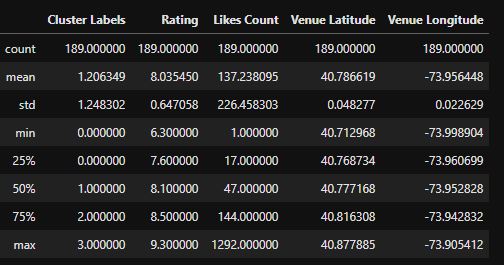


## Manhattan results:

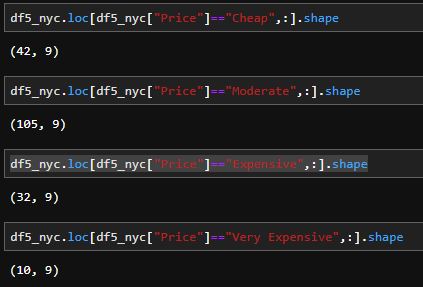
Seventy unique restaurants categories in Toronto with their counts:



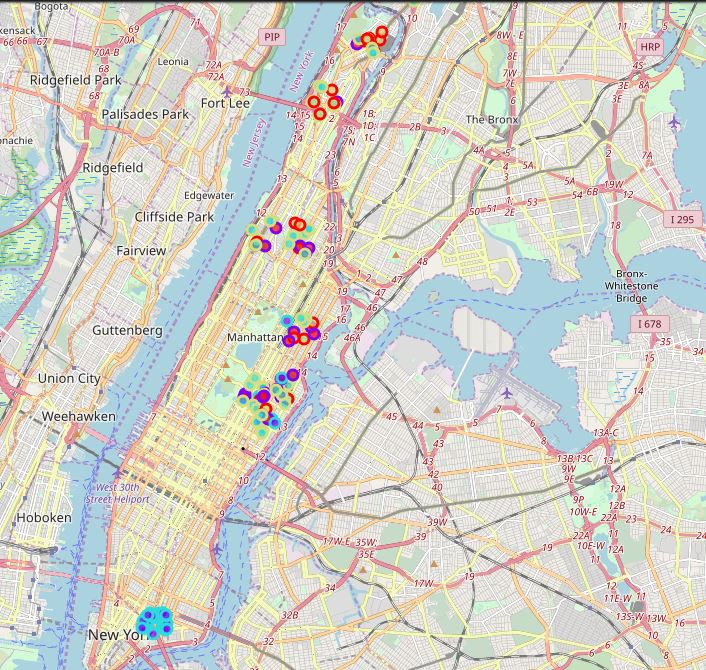
Some stats about the cluster labels, restaurants ratings and likes count:



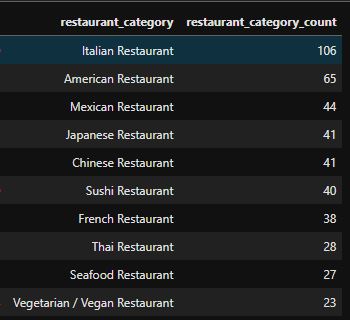
Count for each price category in NYC:



Visualizing the restaurants 4 clusters:



Top restaurants categories and their counts:



## 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.