

Restaurants in New York and Toronto

Marina Hunt

Table of contest:

Introduction

Data

Methodology

Result

Discussion

Conclusion

References

Introduction

Opening a restaurant is one of the most important financial decisions an individual will make in their lifetime. It can be a daunting experience for some.

Throughout project, I explore the in-depth process of opening a restaurant in New York and Toronto. Where the processes diverge between the two cities, I explain the differences. It’s our goal to educate potential restaurant owner about the differences and make it easier to understand.

New York and Toronto are two of biggest cities in North East. New York is famously diverse city. New York City welcomed a record 65.2 million visitors, comprising 51.6 million domestic and 13.5 million international visitors, the ninth consecutive year of tourism growth

Long recognized as one of the most livable cities in the world, Toronto has only recently started receiving the attention it deserves as a tourist destination, and there’s never been a better time to visit. Urban renewal projects have transformed industrial zones, and an energetic dining scene churns out a steady stream of innovative restaurants.

Data

For this assignment, I used following data:

Wikipedia page of Toronto

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

https://cocl.us/Geospatial\_data

New York City data

https://cocl.us/new\_york\_dataset

Foursquare API to pull the following location data on restaurants in Toronto and New York

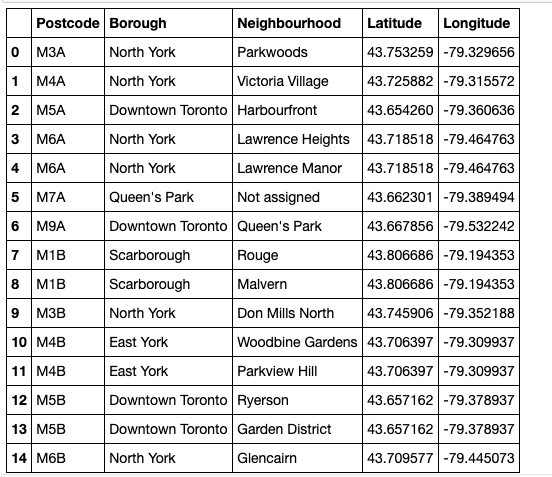
* Venue Name
* Venue ID
* Venue Location
* Venue Category
* Count of Likes

I have used Wiki page of Toronto and https://cocl.us/Geospatial\_data

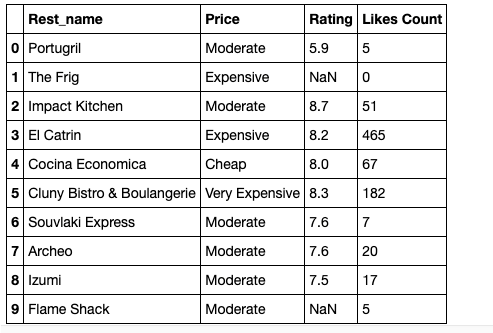
to scab a table to dataframes.







I have cluster similar restaurants together. Based on the output from foursquare, user can easily determine what type of restaurants are best to eat at based on feedback.



I used <https://cocl.us/new_york_dataset> to explore various neiborhoods New York.





Check the quality of the rating with Foursquare



Methodology

From this JSON the following attributes are extracted and added to the Dataframe:

* Restaurant ID
* Restaurant Category Name
* Restaurant Category ID
* Restaurant Postalcode
* Restaurant City
* Restaurant Latitude
* Restaurant Longitude
* Venue Name
* Venue Latitude
* Venue Longitude

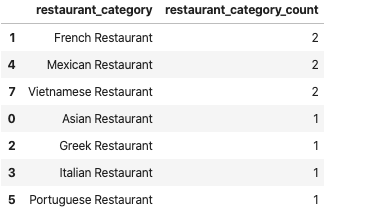
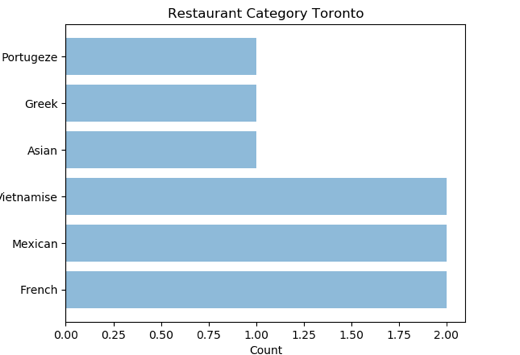
From Foursquare data, we got 326 unique categories of restaurants. Based on the output from Foursquare data, I used filtering the restaurants venues in order to easy determine what type of restaurant include best price, rating and likes count.

To get better understading of the data we I will visualize it. In order to do so, I am performing K-Means clustering to visualize the groups of the best rated category. Using K-Means algorithm from Scikit-learn library to obtain clusters.

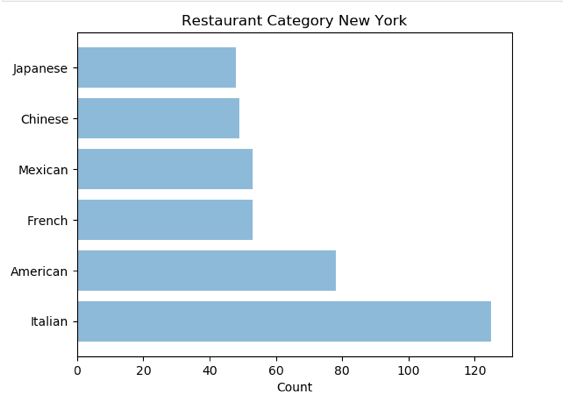
Result

What is very interesting is these results by category.

Restaurant Category in Toronto

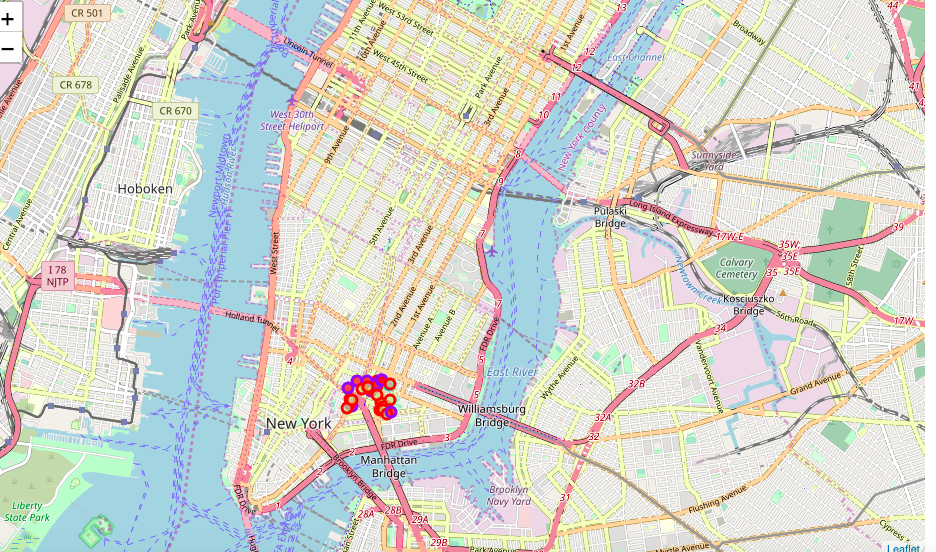


Restaurant Category in New York

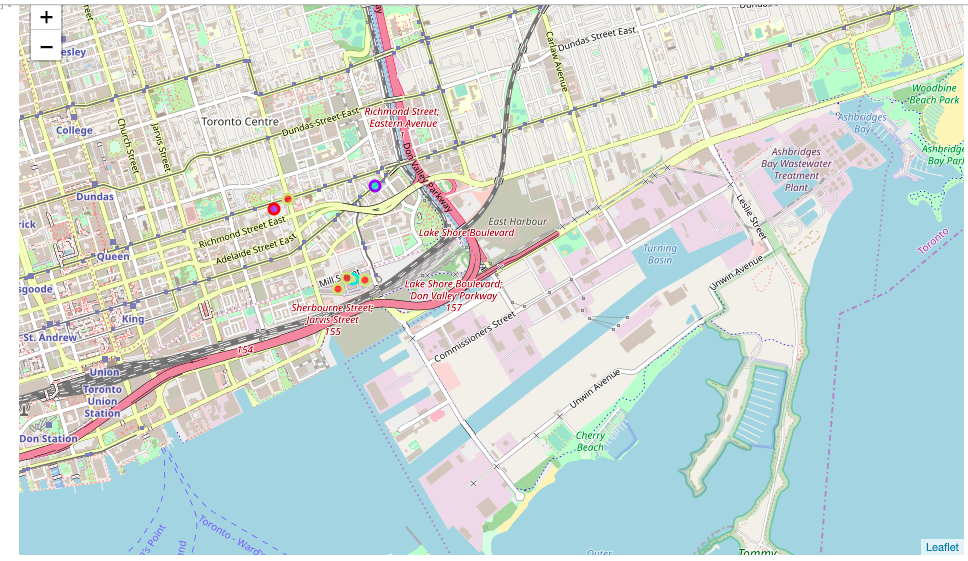


Visualization

Map New York



Map Toronto



Discussion

The idea for the Capstone Project is to show that when driven by venue and location data from FourSquare, that it is possible to present the future possible location to open new restaurant.

According to this analysis, the person who want to open restaurant can decide on a city location.

Conclusion

Toronto area will provide the least competition for potential restaurant owner compare to New York area.