CAPSTONE PROJECT

By Juan Luis Mejía Villa

Toronto vs New York City

Wich is the best city for the restaurant industry?

- A multinational Company that manufactures home appliances and mechanical/electronic products like refrigerators, microwaves, ovens and stoves wants to diversify its market with new complementary line of products in a specific market niche and in a promising location
- The restaurant industry has offered a consistent growth of 2.1% for the last 20 years in North America (https://aaronallen.com/blog/restaurant-industry-growth).

Stakeholders

- The stake holders of the project are the company owners and managers, they already decide that they want to go all in with the project, they are going to invest what is necessary because their idea is to begin really big, that’s why they want to start in the city with the best profile and the larger number of possible clients for the business, they need an exhaustive analysis to have the best foundation for he decision, and is there where the analytics team is going to solve the business problem.

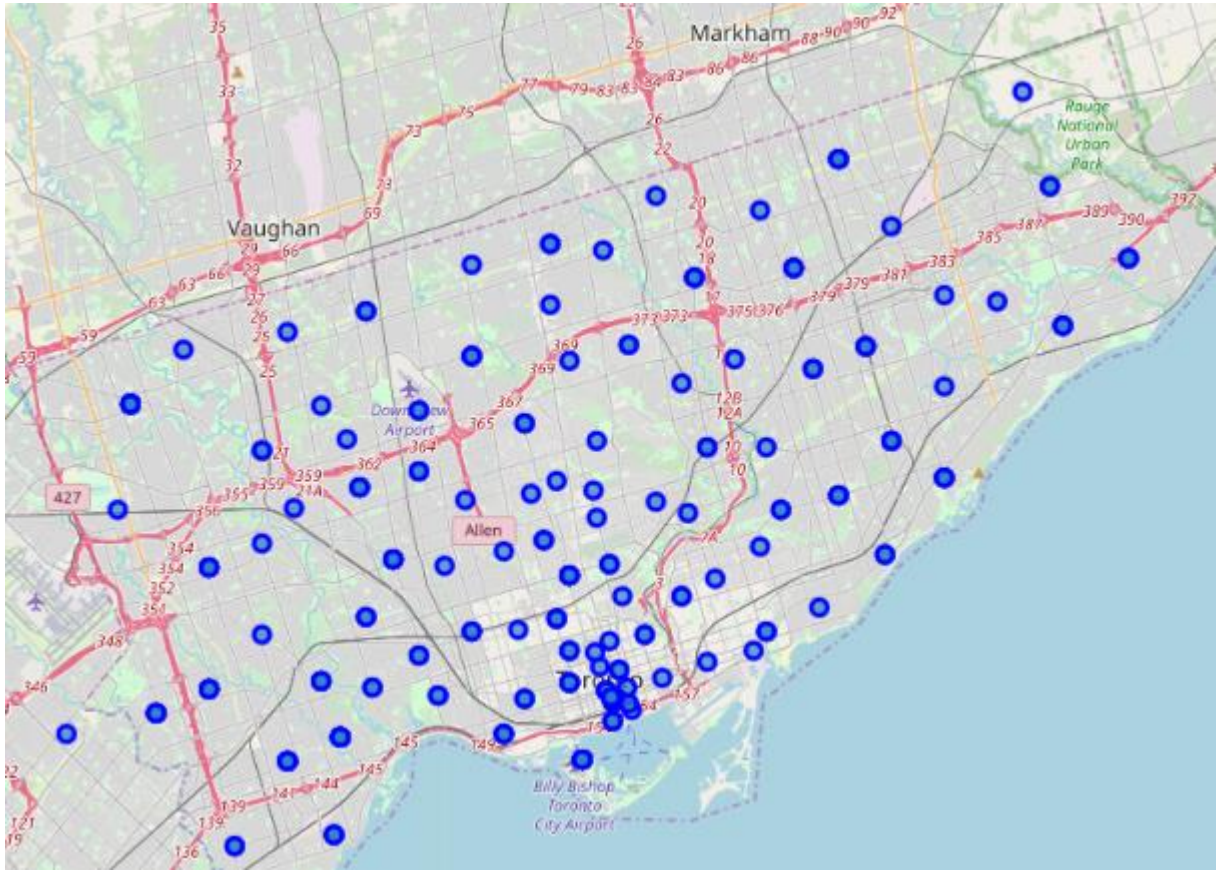
Data

- For these tasks is going to be used Foursquare to retrieve the necessary data trough the API that this page offers.
- We will be using geographical data from both cities; the idea is to have a table for each city that contains the restaurants in the city detailing the restaurant category and locating each restaurant in a neighborhood.
- If needed, the library geocoder will be used to get the latitude and longitude of the cities and neighborhoods.
- Having this data would allow to know the number of restaurants in each city and also in each neighbor and analyzing its categories can be determined in some way the diversity of clients that could be impacted. At the moment that's all the data that is going to be needed for the scope of the project.
- - Toronto data will be taken from wikipedia, the list of Neihborhoods with its Postal Code (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- - Toronto locations (latitud and longitud) are going to be taken from provided csv in the following link: https://cocl.us/Geospatial_data
- - New York City data will be taken from the provided link: https://geo.nyu.edu/catalog/nyu_2451_34572
- - New York City locations (latitud and longitud) are going to be found with geopy

Neighborhood Maps

- New York City

- Toronto



NY venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
2	Wakefield	40.894705	-73.847201	Cooler Runnings Jamaican Restaurant Inc	40.898083	-73.850259	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	SUBWAY	40.890468	-73.849152	Sandwich Place
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

Toronto venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rouge	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
2	Guildwood	43.763573	-79.188711	Swiss Chalet Rotisserie & Grill	43.767697	-79.189914	Pizza Place
3	Guildwood	43.763573	-79.188711	Big Bite Burrito	43.766299	-79.190720	Mexican Restaurant
4	Guildwood	43.763573	-79.188711	Eggsmart	43.767800	-79.190466	Breakfast Spot

Which city?

New York

```
#Number of Restaurants
numb_ny = len(ny_rest["Venue"])

#Divercity

div_ny = len(ny_rest["Venue Category"].unique())

cat_by_venues_ny = div_ny/numb_ny
```

```
print(numb_ny)
print(div_ny)
print(cat_by_venues_ny)
```

```
3895
127
0.03260590500641849
```

Toronto

```
#Number of Restaurants
numb_toronto = len(toronto_rest["Venue"])

#Divercity

div_toronto = len(toronto_rest["Venue Category"].unique())

cat_by_venues_toronto = div_toronto/numb_toronto
```

```
print(numb_toronto)
print(div_toronto)
print(cat_by_venues_toronto)
```

```
1725
77
0.04463768115942029
```

Conclusion

Toronto has 1% more categories per restaurant than NY, but the difference is extremely big in number of restaurant and number of categories, that we defined as divercity.

NEW YORK CITY IS THE CHOSEN ONE