

# Battle of the Neighborhoods : Moroccan restaurant

Marwane Lotfi

August 29, 2020

## 1. Introduction

### 1.1 Description and Background¶

Canada is a historical land of immigration, and this cultural mesh created a very wide and rich display of different cultures, colors and tastes. At each corner people can find a variety of ethnical cuisine restaurants and shops, some more widely popular than others.

As many type of cuisines already are famous (italian, american) or many that benefits from a booming interest (asian, japanese) ther are yet continents that are timidly trying to make their food get some keen interest by the general public.

African food is still not as popular in the public panel and is yet to show it's full potential, bringing unexplored flavours to the city of Toronto.

Morocco is a historical melting pot that combines influences from plethora of cultures and ethnicities : berbers, arabs, sub saharians, phenicians, jews, romans...

And this richness can be felt into its very traditional but exquisite cuisine but unfortunately not as well known.

The goal of this project would be to open a **restaurant chain of moroccan themed food with some african twist**. The simultaneous presence of the stores would have a great impact for advertising moroccan food across the city.

## 2. Data

### 2.1 Description of the Data

The data to use would consist in boroughs and neighborhoods of the city of **Toronto** famous for their restaurant and ethnic food concentration. Spots that are well known by foodies and that already holds a customer basis.

The data should also pinpoint areas with good potential that is free from fierce concurrence.

The source of data would be a combination of **Wikipedia** boroughs list with **Foursquare API** data of Toronto Venues.

The classification should be made regarding the ethnical type of food served and be analyzed regarding similar type of foods.

## 2.2 Data Extraction : Toronto Boroughs from Wikipedia

We are going to extract all the **Toronto Boroughs** name and confront them with their respective coordinate data to generate a Dataframe basis

Scraping the data from Wikipedia and Confronting them to the Foursquare API Database of Toronto Extracting the Different Venues out of the main the Boroughs of Toronto.

Out[11]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
3	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa
4	Regent Park, Harbourfront	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant
5	Regent Park, Harbourfront	43.65426	-79.360636	Corktown Common	43.655618	-79.356211	Park
6	Regent Park, Harbourfront	43.65426	-79.360636	The Distillery Historic District	43.650244	-79.359323	Historic Site
7	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot
8	Regent Park, Harbourfront	43.65426	-79.360636	The Extension Room	43.653313	-79.359725	Gym / Fitness Center
9	Regent Park, Harbourfront	43.65426	-79.360636	Dominion Pub and Kitchen	43.656919	-79.358967	Pub

*Toronto Venues from Wikipedia and Foursquare API with geospatial Data*

## 3. Methodology

### 3.1 Isolating the different restaurant types

Hence the fact that we are focusing on opening a restaurant, we do isolate all the different restaurant venues.

Regroup the data into a pandas dataframe

Entrée [18]:

```

toronto_resto = pd.DataFrame(restaurant_type_list)
toronto_resto.reset_index(drop=True,inplace=True)
print("Toronto Resto shape : ",toronto_resto.shape)
toronto_resto.head()

```

Toronto Resto shape : (383, 7)

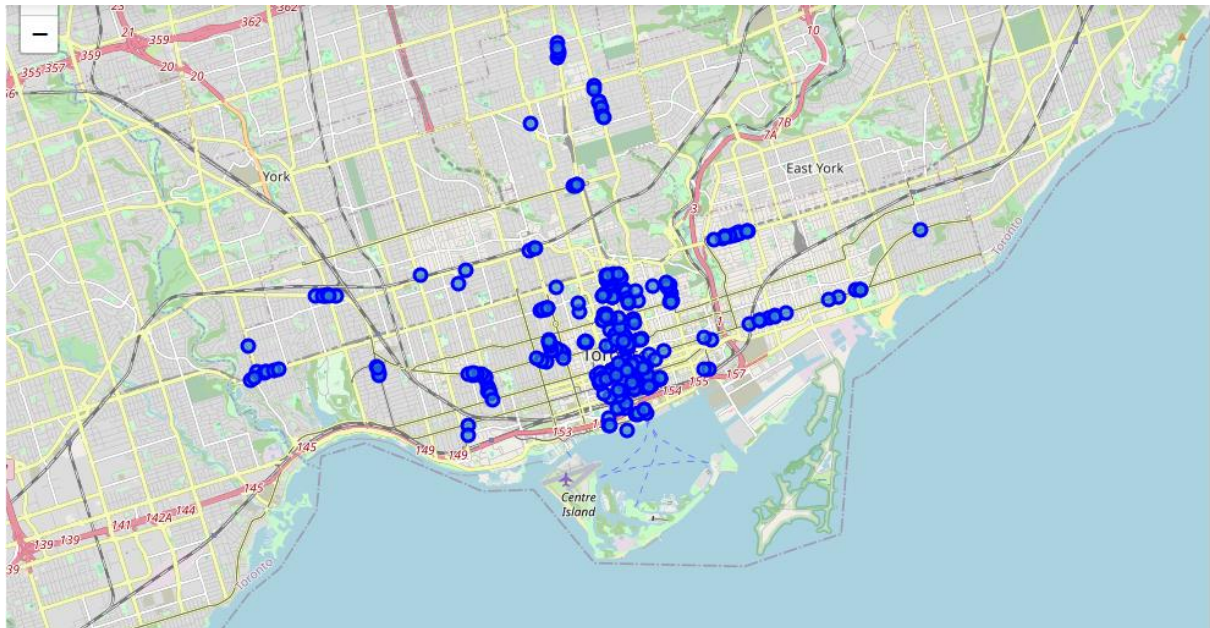
Out[18]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park, Harbourfront	43.654260	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant
1	Regent Park, Harbourfront	43.654260	-79.360636	Cluny Bistro & Boulangerie	43.650565	-79.357843	French Restaurant
2	Regent Park, Harbourfront	43.654260	-79.360636	El Catrin	43.650601	-79.358920	Mexican Restaurant
3	Regent Park, Harbourfront	43.654260	-79.360636	Flame Shack	43.656844	-79.358917	Restaurant
4	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Nando's	43.661728	-79.386391	Portuguese Restaurant

*Pandas Dataframe of Restaurant venues*

## 3.2 Displaying the different restaurants on a map

Displaying all these restaurants on the map to see the distribution



Map of the restaurant distribution in Toronto

## 3.3 Analysing each Restaurant Type

We start by applying a one hot encoding to prepare the data for frequency analysis

```
Entrée [24]: toronto_onehot = pd.get_dummies(toronto_resto[['Venue Category']],prefix="",prefix_sep="")
toronto_onehot.head()
```

Out[24]:

	Afghan Restaurant	American Restaurant	Asian Restaurant	Belgian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Colombian Restaurant	Comfort Food Restaurant	...	Portuguese Restaurant	Ramen Restaurant	Restau
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	1	0	0

Applying one hot encoding

Grouping the frequency of each venue type by Neighborhood

	Neighborhood	Afghan Restaurant	American Restaurant	Asian Restaurant	Belgian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Colombian Restaurant	...	Portuguese Restaurant	Ramen Restaurant	Rest
0	Berczy Park	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.000000	0.1
1	Brockton, Parkdale Village, Exhibition Place	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.000000	0.5
2	Business reply mail Processing Centre, South C...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.000000	0.5
3	Central Bay Street	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.052632	0.052632	0.0
4	Christie	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.000000	0.5

5 rows × 48 columns

*Restaurant type frequency per neighborhood*

Then we regroup the data into a Dataframe showing up the descendant order by Most common restaurant type per Neighborhood

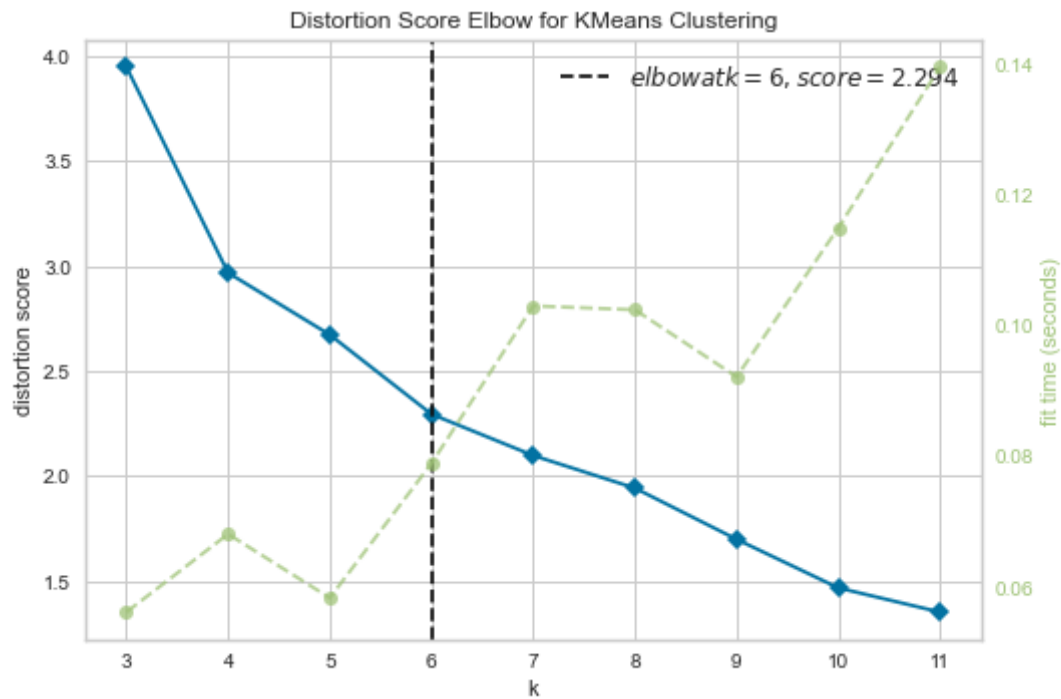
	Neighborhood	1st Most Common Restaurant Type	2nd Most Common Restaurant Type	3rd Most Common Restaurant Type	4th Most Common Restaurant Type	5th Most Common Restaurant Type	6th Most Common Restaurant Type	7th Most Common Restaurant Type	8th Most Common Restaurant Type	9th Most Common Restaurant Type	10th Most Common Restaurant Type
0	Berczy Park	Seafood Restaurant	Restaurant	Comfort Food Restaurant	Eastern European Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	Japanese Restaurant	French Restaurant	Thai Restaurant
1	Brockton, Parkdale Village, Exhibition Place	Italian Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant
2	Business reply mail Processing Centre, South C...	Fast Food Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant
3	Central Bay Street	Italian Restaurant	Japanese Restaurant	Indian Restaurant	Portuguese Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Korean Restaurant	Modern European Restaurant	New American Restaurant	Middle Eastern Restaurant
4	Christie	Italian Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant
5	Church and Wellesley	Japanese Restaurant	Sushi Restaurant	Restaurant	Mediterranean Restaurant	Indian Restaurant	Persian Restaurant	American Restaurant	Caribbean Restaurant	Fast Food Restaurant	Mexican Restaurant

*Dataframe of the most common restaurant per neighborhood*

## 3.4 Clustering the Neighborhood

In this part we gonna apply **Machine Learning** tools on our data. Hence the fact that we want to find the best areas to open our joint, one of the best feature to apply is the clustering method.

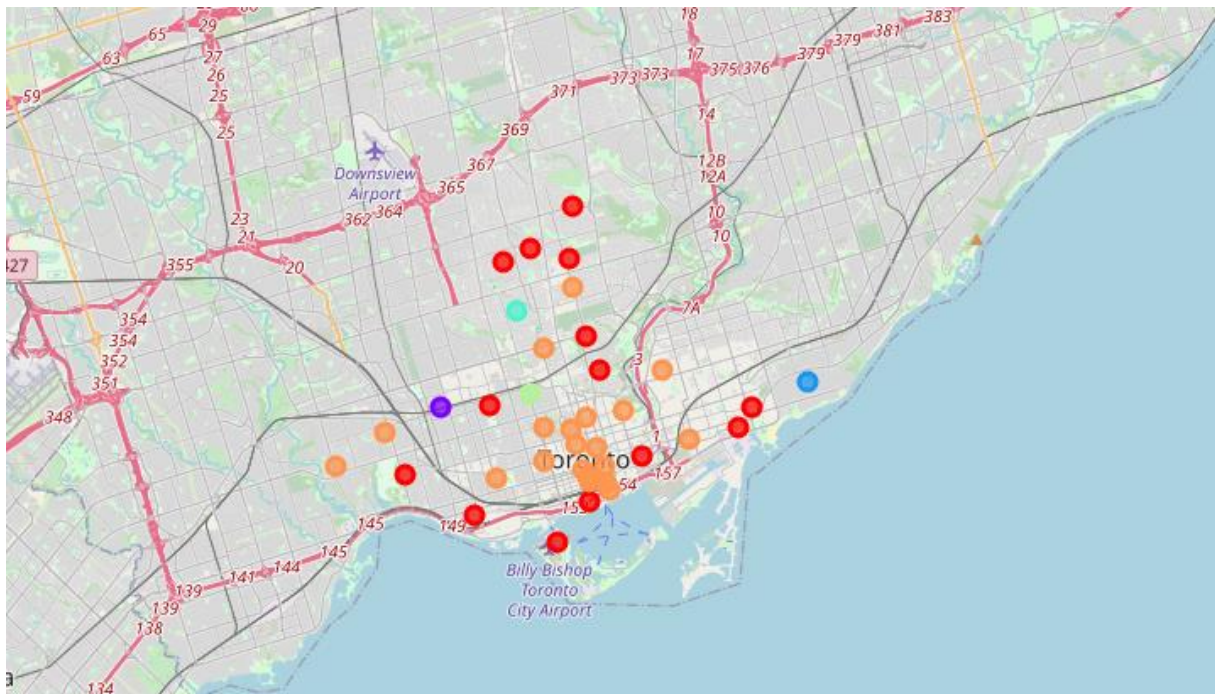
To identify exactly how many clusters (**k**) we should aim for to get a more accurate clustering we gonna use a testing method to determine so commonly known as **Elbow Method**.



application of k elbow

We can clearly notice that following the method the best fit for clusters is :  $k = 6$

We map the **Neighborhoods** on a **folium** map to better see the distribution of the neighborhoods into clusters



map of the neighborhood clusters

## 4. Results

By analyzing the data we can clearly identify a pattern that draws itself.

We can see that the clusters **1,2,3,4** and **6** concentrate their restaurants into specific areas of Toronto (West, East, Central) areas.

While the richest variety of restaurants considered is located within the **Downtown area** (Cluster 6), which can seem logical at first glance regarding the historical aspect of the site.

As we did see previously, the sole moroccan venues are located in the same area **St. James Town** in **Downtown Toronto** :

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
50	St. James Town	43.651494	-79.375418	Berber SOCIAL	43.648114	-79.374046	Moroccan Restaurant
51	St. James Town	43.651494	-79.375418	The Sultan's Tent and Cafe Moroc	43.647928	-79.374341	Moroccan Restaurant

*location of the moroccan restaurants in Toronto*

## 5. Discussion

Ultimately, and due to very few competition, the best option to open our Moroccan themed restaurant Brand would be in various areas in Downtown excepting **St. James Town**.

Even though opening in spots surrounding this area would be an interesting fit as it already has a customer basis looking for north african flavors.

## 6. Conclusion

In conclusion, we can clearly notice how powerful a tool machine learning and Data Science are. In a few lines of codes, and some logic we were able to make an efficient Data analysis from online available sources that mixed together can offer the premices of a solid Business plan or Marketing campaign.

As simple of an application it can be made, this information processing can offer an invaluable amount of hints for business opportunities.