Predicting where AwesomeRestaurant&Co should open a restaurant in Toronto

A datascience analysis

Market Analysis

A restaurant company, AwesomeRestaurant&Co, wants to open a business in Toronto.

The company should answer to those question before :

- What are the competition in the districts of Toronto ?
- Is it strong ?
- How many restaurants are they in the central district ?
- What could be a good place to start ?

Data acquisition and cleaning

- We will retrieve from the Foursquare API venues data about restaurant in Toronto
- We will cover the broad center of Toronto for a total of 99 restaurants

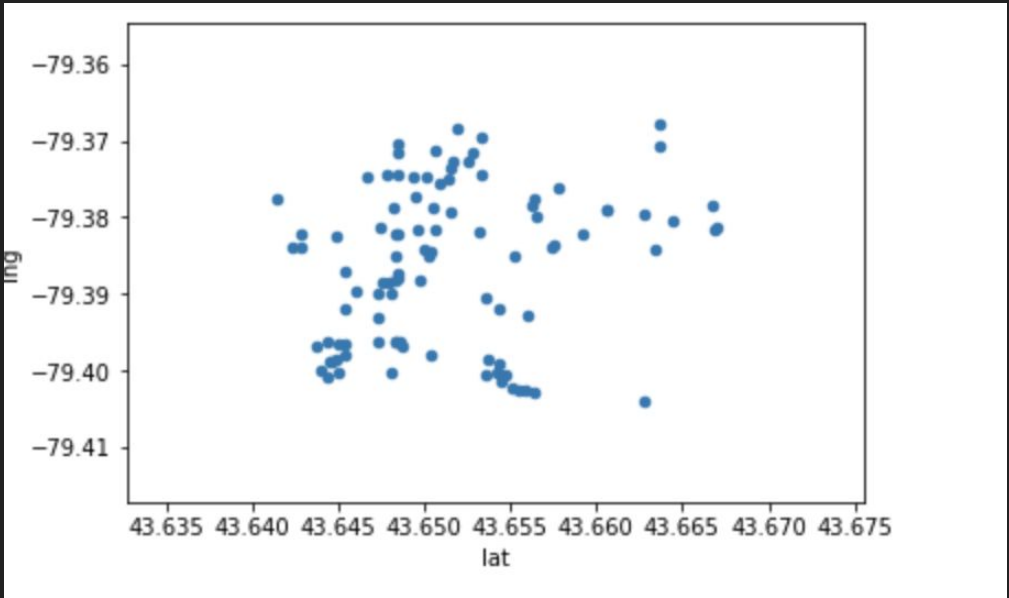
Overview of the dataframe

We will work with a primary dataframe of venue IDs, name, and coordinate of each restaurant in the center of Toronto

Here is an extract of this dataframe :

	venue_ID	name	lat	lng
0	506db1a9e4b0a3f3b31412f0	Richmond Station	43.651569	-79.379266
1	537773d1498e74a75bb75c1e	Eggspectation Bell Trinity Square	43.653144	-79.381980
2	5615b6c4498e3c32c67ad78f	Blaze Pizza	43.656518	-79.380015
3	51755dc7498ece19b7261641	Banh Mi Boys	43.659188	-79.382131
4	4ad69511f964a520e40721e3	The Keg Steakhouse & Bar	43.649937	-79.384196

Visualizing the dataframe



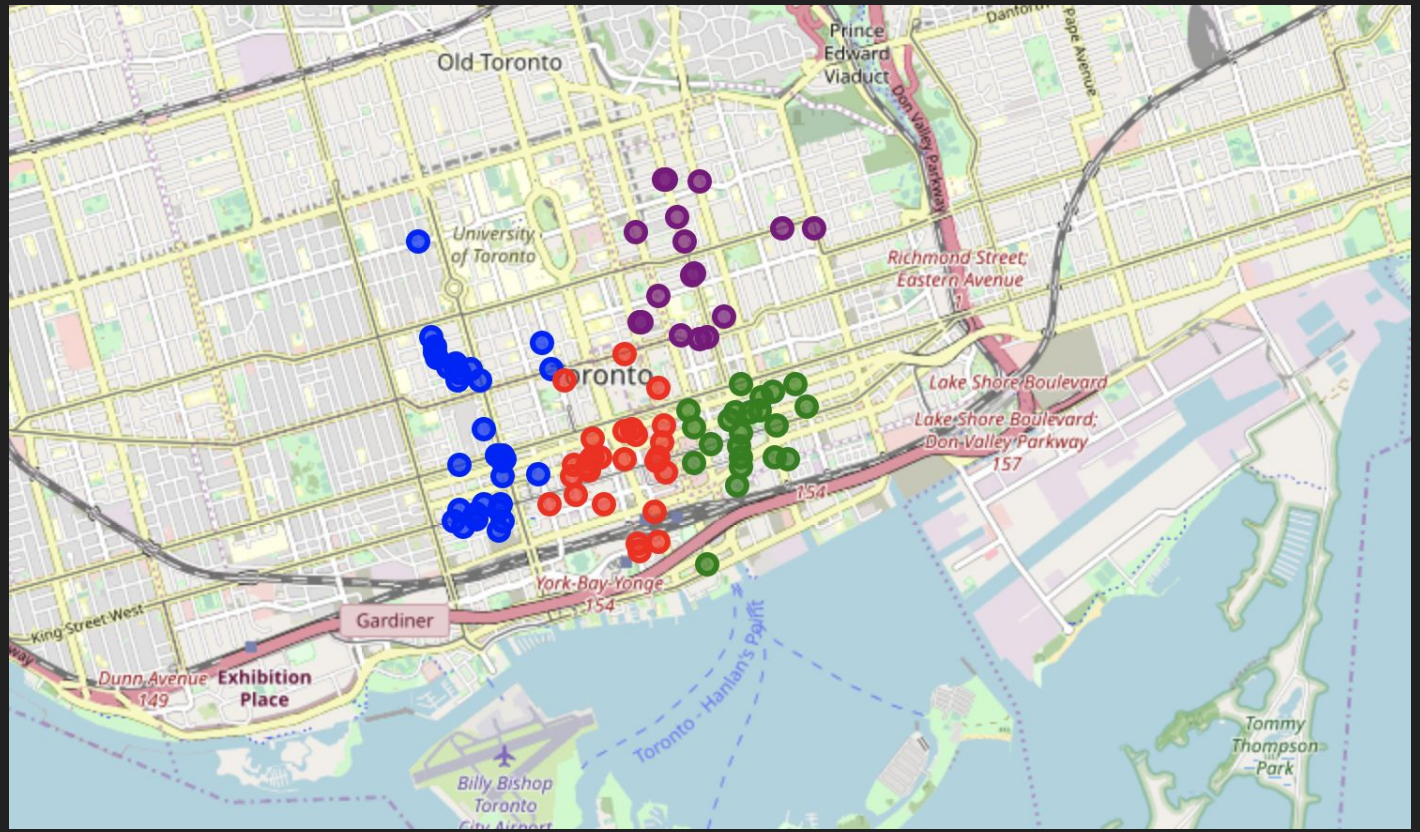
By plotting the restaurants, we can see several clusters (around 3 or 4) immediately.

Clustering of the data

We will run a KMeans algorithm to try to determine the several clusters. The algorithm returns the following :

	venue_ID	name	lat	lng
cluster				
0	33	33	33	33
1	27	27	27	27
2	22	22	22	22
3	17	17	17	17

Clustering of the current restaurants of Toronto



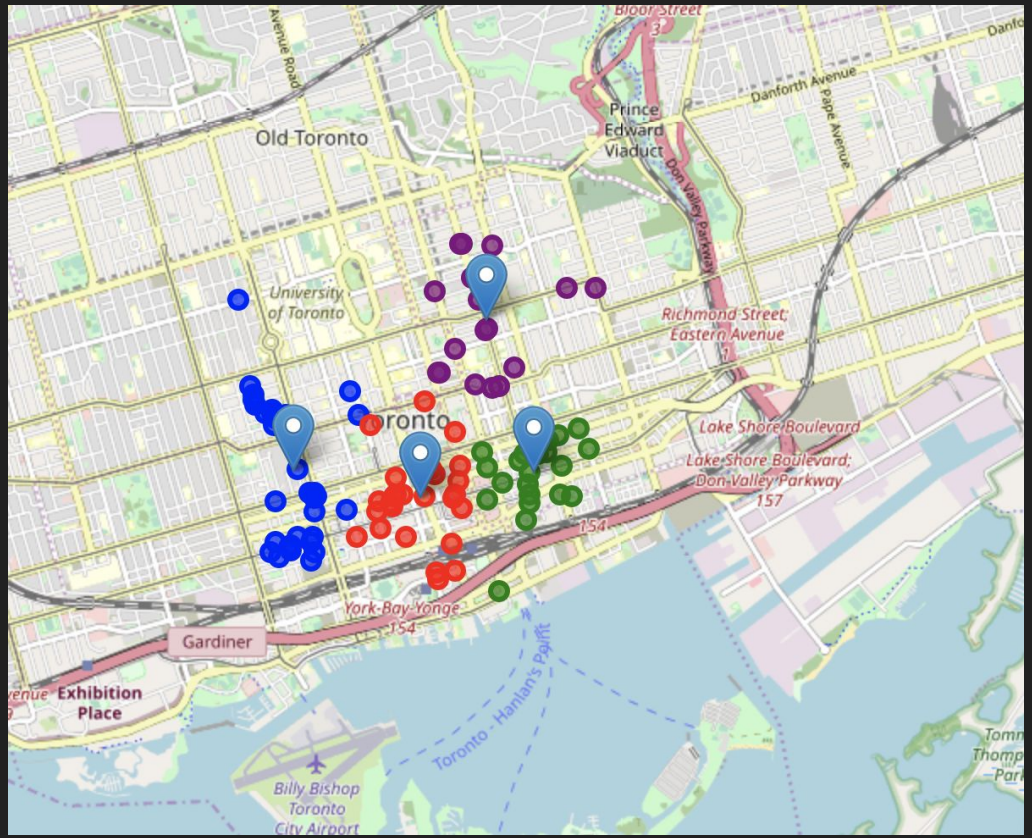
Clustering of the data

We determine there are for clusters of restaurants in the center of Toronto.

We will try to determine the best location for our new restaurants by determining the center of each clusters to optimize foot trafic and take advantages of well known areas

	lat	lng
cluster		
0	43.650090	-79.398597
1	43.648179	-79.385733
2	43.650011	-79.374150
3	43.661219	-79.379013

Recommendations for new restaurants



Conclusion and future directions

We determined 4 different locations to open a new restaurant:

	lat	lng
cluster		
0	43.650090	-79.398597
1	43.648179	-79.385733
2	43.650011	-79.374150
3	43.661219	-79.379013

We could improve the model by taking into account the grades/popularity of the restaurants