Finding the best place to open an African restaurant in New York City

Thierno Oumar Diallo

June 1, 2020

Table of Contents

1. Introduction

1.1 Background

Finding a good place isn't always easy. For a new business owner who does not have enough money, it is desirable for him to succeed in his project by having a place where he will be able to move forward. What would he gain if he opened his market in a place where there are many products identical to his own? How would he get customers?

Having a good location close to the city centre where there are many attractions is a good first step for a business owner. Because it can help him to increase his turnover quickly.

1.2 Problem

Data that could help determine where a business owner can open a restaurant could include information on the names and coordinates information for neighborhoods, boroughs, and districts in New York City, as well as information on restaurants in New York City. The purpose of this project is to find out which neighborhood will be suitable for a business owner to open a restaurant.

1.3 Interest

It will be very interesting for an entrepreneur to find the best place to invest for the opening of a new African restaurant. This can prevent him from opening in a place where nobody is interested in what he sells. It can therefore help him to save time and money.

2. Data acquisition and cleaning

2.1 Data sources

The New York data can be downloaded here. This is a json file that contains all the information on the neighborhoods. We focus only on boroughs, neighborhoods, latitudes and longitudes from the data.

Then the population data for each neighbourhood can be viewed here. It contains the neighbourhood and the population size. We focus only on 2020 values.

Finally, the population data for each neighbourhood in Brooklyn can be found here. It contains the neighbourhoods and the corresponding population size. The most recent population values we have obtained are for 2010.

2.2 Data cleaning

Data from each of these sources are prepared separately.

For the new york city data, we only select the Borough, Neighborhood, latitude and longitude from the data.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

For population size by borough in New York City, we select only the total size of each borough for the year 2020.

Borough	Population
Brooklyn	2648452
Queens	2330295
Manhattan	1638281
Bronx	1446788
Staten Island	487155

For the population size of each neighbourhood in Brooklyn, we filter the data to select only those neighbourhoods in Brooklyn with the corresponding population size for the year 2010.

	Neighborhood	Population
277	Borough Park	106357
253	Flatbush	105804
261	Crown Heights North	103169
274	East New York	91958
241	Bensonhurst West	88727

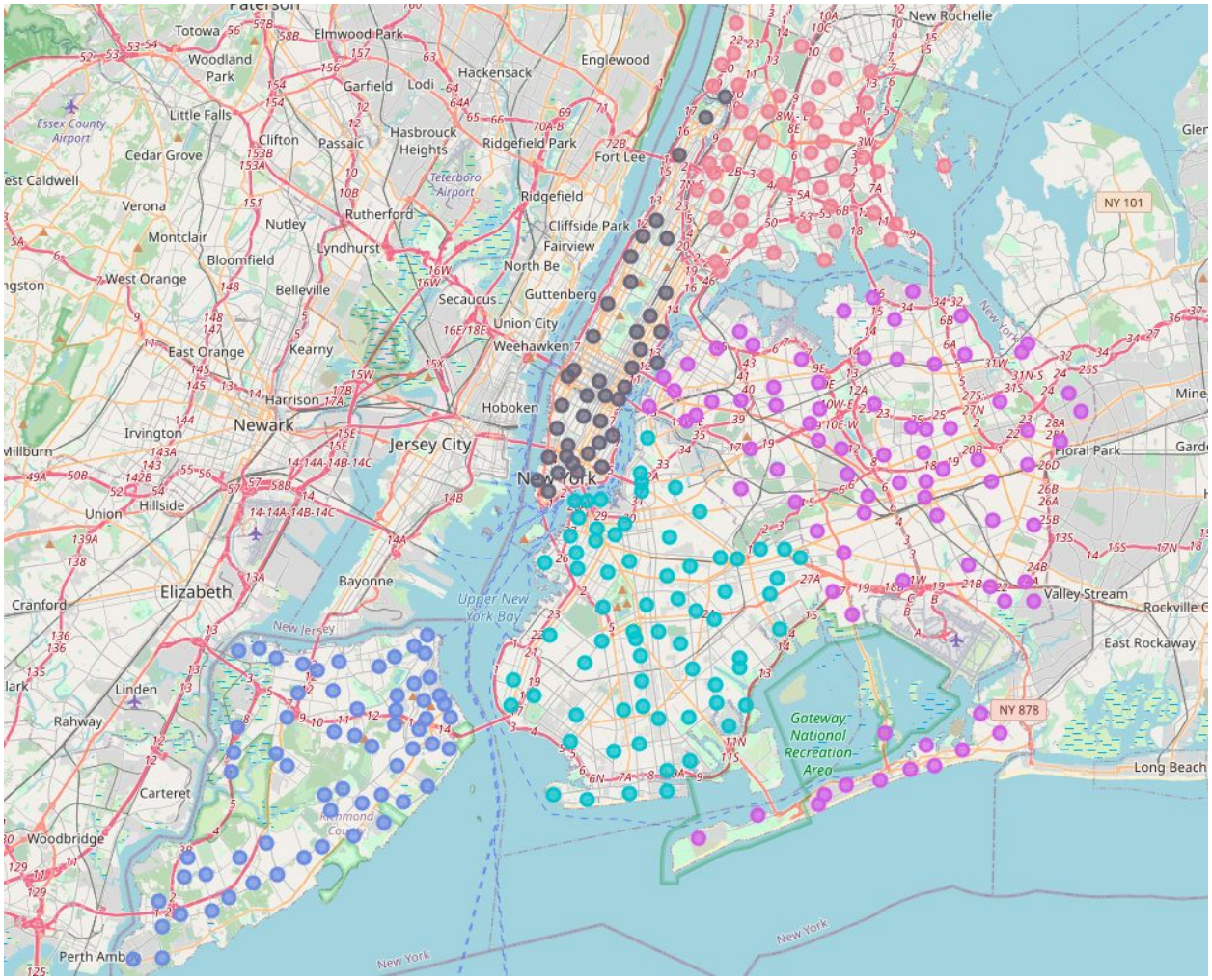
At the end, all the dataframes are merged to obtain more information on each neighbourhood. Then, with Foursquare, we get the information from the restaurants in those locations.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue Category Id	Venue Distance	Is African	Borough	Latitude	Longitude	Population
269	Borough Park	40.633131	-73.990498	Spoons	40.634364	-73.992624	Restaurant	4bf58dd8d48988d1c4941735	226	False	Brooklyn	40.633131	-73.990498	106357
270	Borough Park	40.633131	-73.990498	China Glatt	40.637381	-73.989255	Chinese Restaurant	4bf58dd8d48988d145941735	484	False	Brooklyn	40.633131	-73.990498	106357
271	Borough Park	40.633131	-73.990498	Vostok	40.631181	-73.995791	Eastern European Restaurant	4bf58dd8d48988d109941735	497	False	Brooklyn	40.633131	-73.990498	106357
268	Borough Park	40.633131	-73.990498	Orchidea	40.636390	-73.993899	American Restaurant	4bf58dd8d48988d14e941735	462	False	Brooklyn	40.633131	-73.990498	106357
104	Flatbush	40.636326	-73.958401	El Paso Mexican Grill	40.636081	-73.962079	Mexican Restaurant	4bf58dd8d48988d1c1941735	311	False	Brooklyn	40.636326	-73.958401	105804

3. Exploratory Data Analysis

3.1 New York City Boroughs and Neighborhoods

There are 306 neighborhoods and five boroughs in New York City.
Each borough with a different color

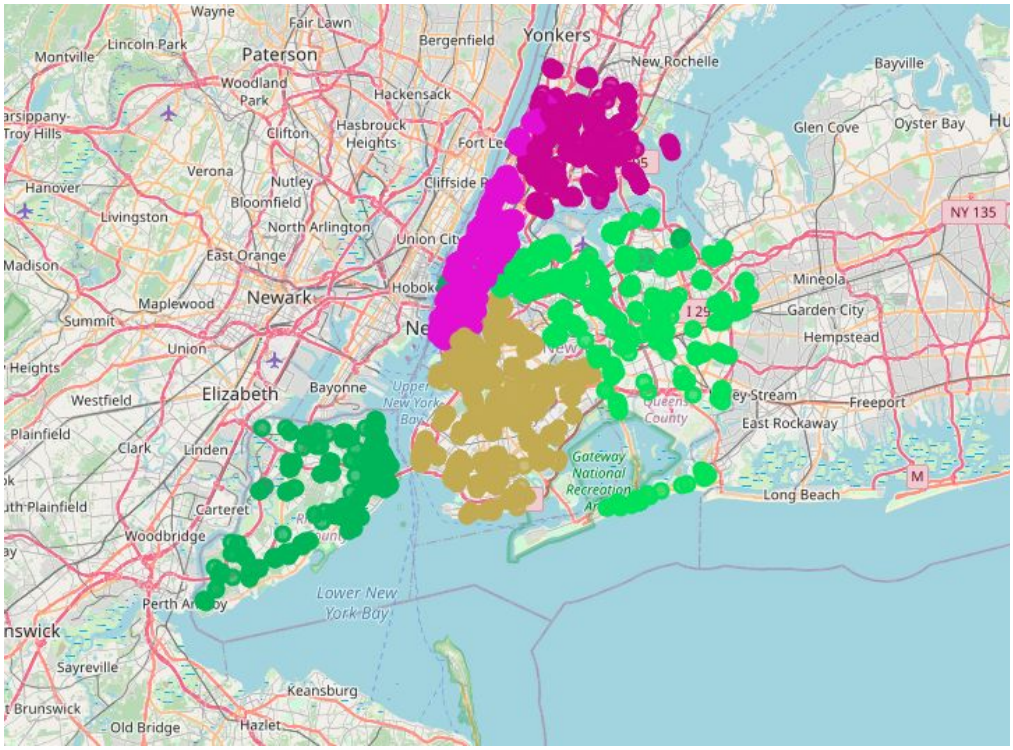


We can see there are a lot of restaurants in New York. The next objectives will be to look for a borough with a lot of people and fewer restaurants.

3.2 Restaurants in New York City

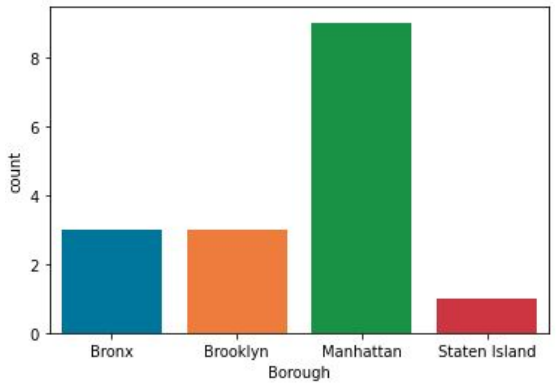
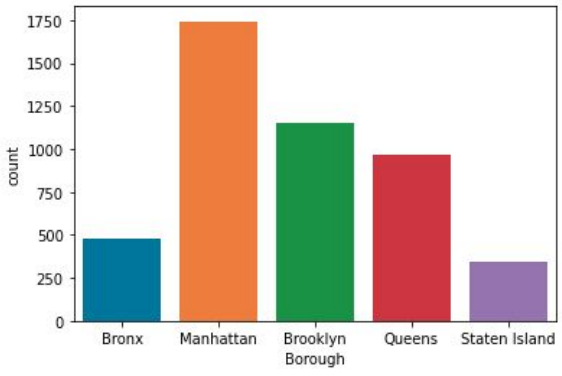
There are 16 African restaurants and `4677` other restaurants in New York City. Notice I'm only focusing on food restaurants. Fast food, coffee shops, pizza shops are not included here.

Each borough with a different color.



To find

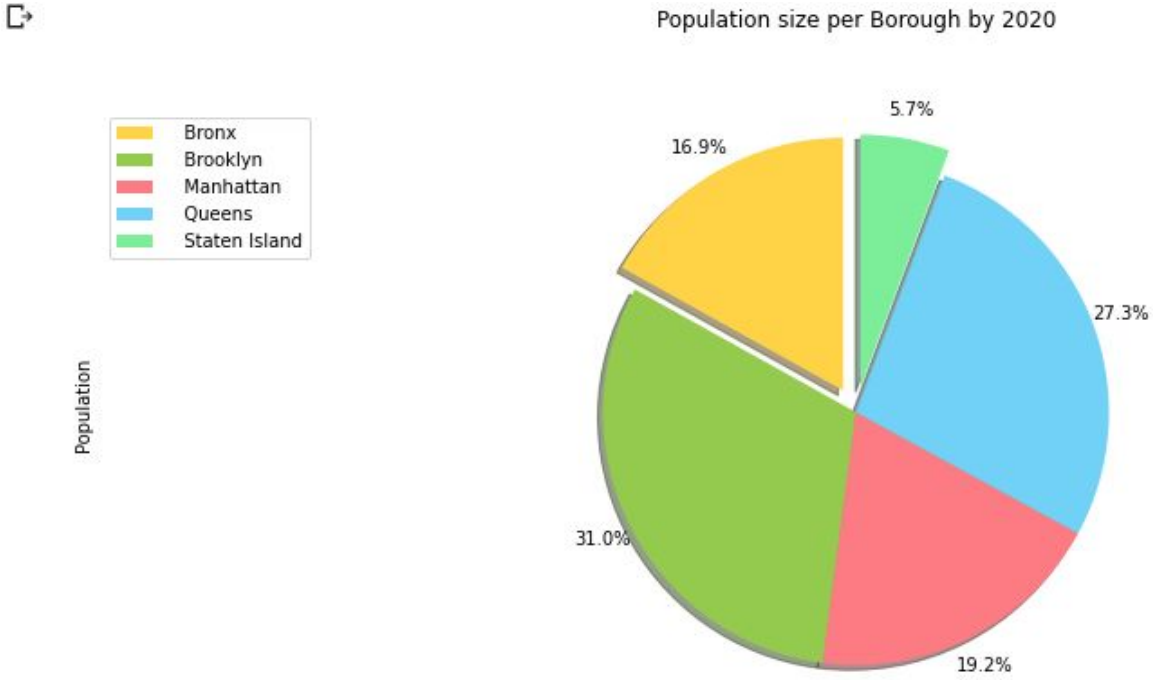
3.3 Number of restaurants per borough

All types of restaurants	African restaurants
	

We can see that Manhattan has many more restaurants than other boroughs. Our goal is to focus on the borough where there are a lot of people and few restaurants.
For this, it is important to know the most populated borough. This will help us to know which one to choose.

3.4 Population size by Borough in 2020

We can see that Brooklyn is the most populous borough in New York City.

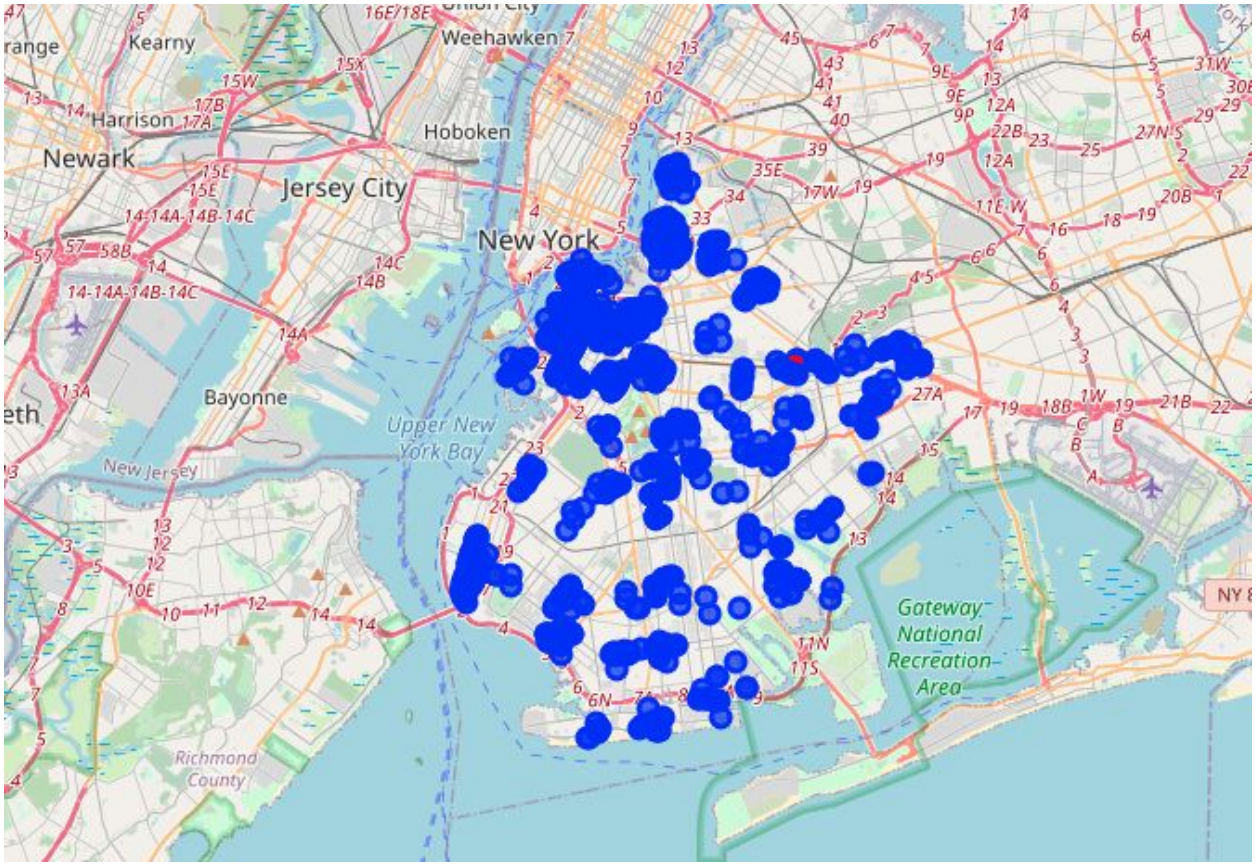


We also saw in section 3.3 that Brooklyn has few African restaurants. Our goal is to focus on the borough which is populated and has few African restaurants. That is why, for the next steps, I
am focusing only on the neighborhoods of Brooklyn.

3.5 Brooklyn restaurants

As I said before, we only focus on food restaurants. Fast food, coffee shops, pizza shops are not included.

There are `1150` restaurants of 70 venues categories.



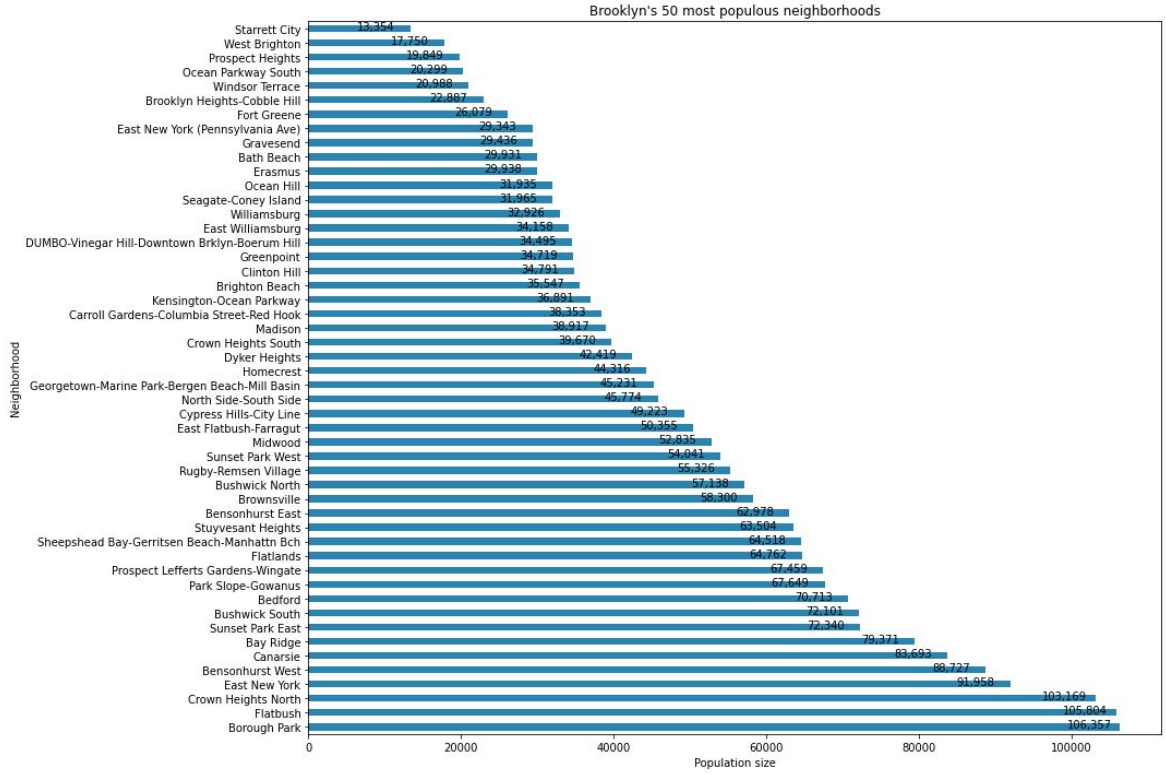
African restaurants are in red. We can see that there are not many of them (just 3). In the next steps, I will focus only on Brooklyn's 50 most populous neighborhoods.

3.6 Brooklyn's 50 most populous neighborhoods

The data collected for population sizes in the various Brooklyn neighbourhoods are from 2010.

	Neighborhood	Population
277	Borough Park	106357
253	Flatbush	105804
261	Crown Heights North	103169
274	East New York	91958
241	Bensonhurst West	88727

Let's take a look at the 50 most populous neighbourhoods

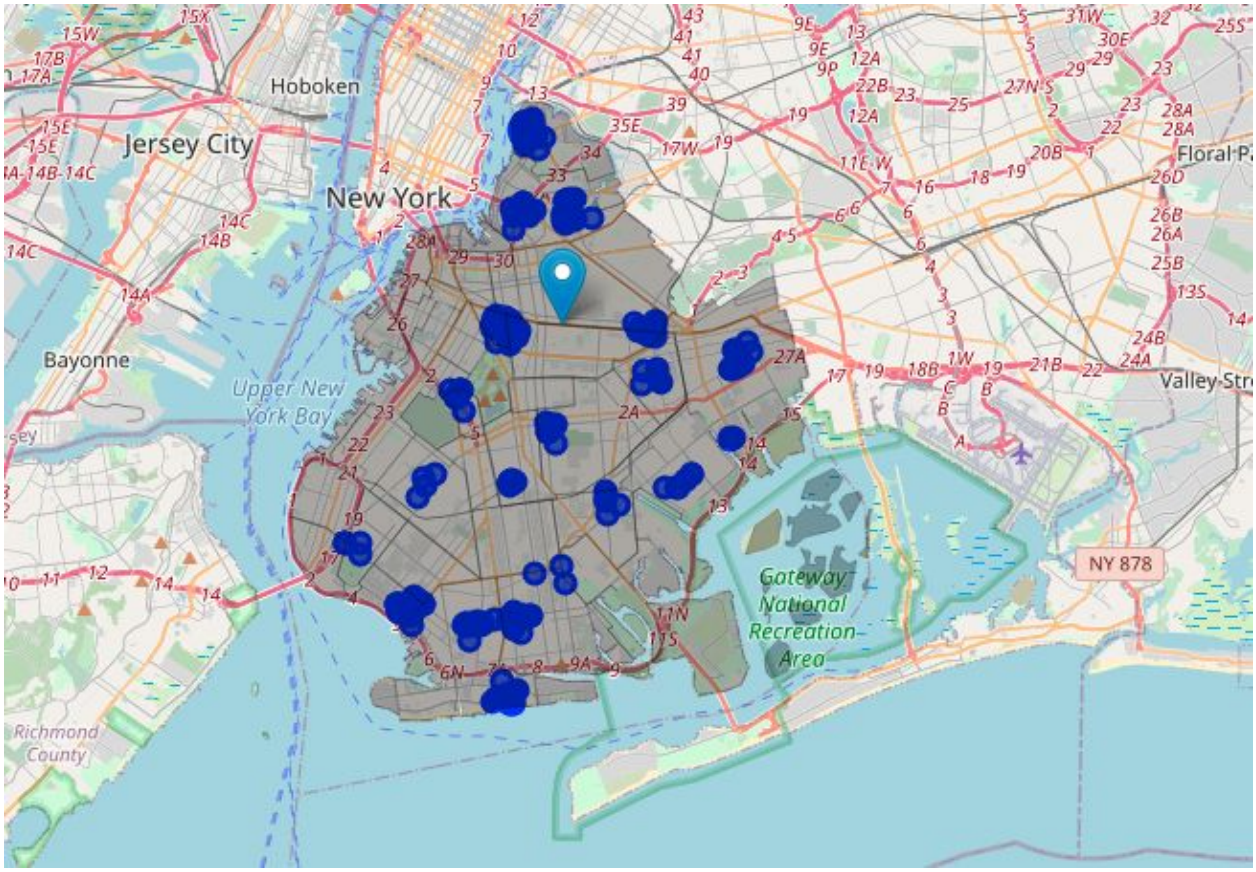


From those 50 most populated neighborhoods, when we merged with the restaurant data, we got 400 lines.

3.6 The 20 populated neighborhoods with the fewest restaurants

We have achieved our first objective, which is to have populated neighborhoods. So the second goal is to select the 20 neighborhoods that have the fewest restaurants.

Let's visualize these **restaurants**



Now that we have the 20 populated neighbourhoods with the least number of restaurants, the objective is to filter them to retain only those restaurants that are no more than 5 km from downtown Brooklyn.

3.7 Distance between each neighborhood and Downtown Brooklyn

I calculated the distance between each neighborhood and downtown Brooklyn. Then I selected the ones that are less than 5 km away.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Category	Venue Latitude	Venue Longitude	Population	Total Numbers	distance km
85	Windsor Terrace	40.656946	-73.980073	Bedawi Cafe	Middle Eastern Restaurant	40.658588	-73.982020	20988	15	4.37109
86	Windsor Terrace	40.656946	-73.980073	Krupa Grocery	American Restaurant	40.659957	-73.980696	20988	15	4.37109
87	Windsor Terrace	40.656946	-73.980073	East Wind Snack Shop	Chinese Restaurant	40.660297	-73.980169	20988	15	4.37109
88	Windsor Terrace	40.656946	-73.980073	Sushi Yama	Sushi Restaurant	40.658542	-73.982054	20988	15	4.37109
89	Windsor Terrace	40.656946	-73.980073	Giovanni's Brooklyn Eats	Italian Restaurant	40.660159	-73.983753	20988	15	4.37109

Finally, we got 113 venues and 5 neighborhoods that are at least 5 km from downtown Brooklyn.

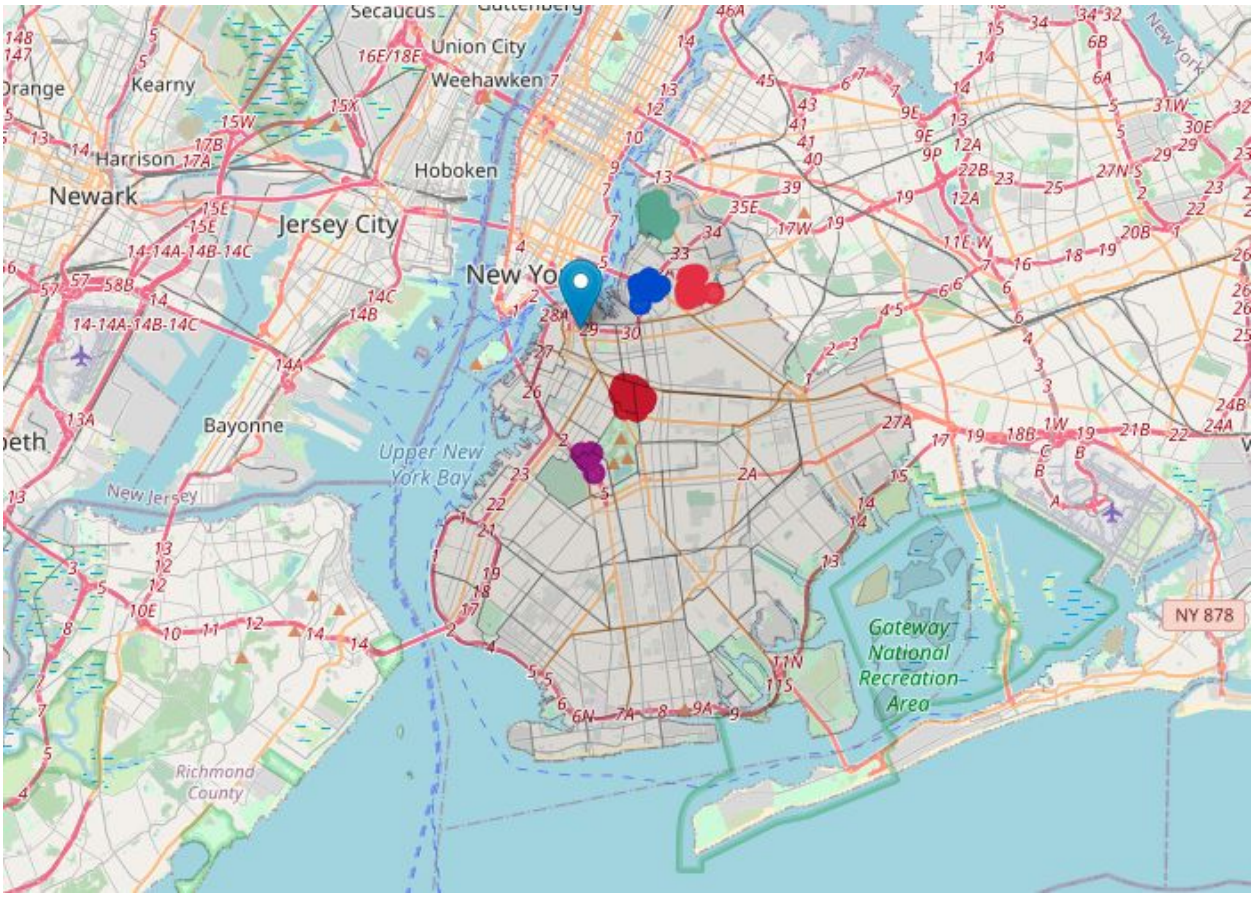
```
The average distance for Windsor Terrace is 4.371090370531356 km
The number of non-african restaurants is 15
Windsor Terrace's population size is 20988


The average distance for Williamsburg is 2.541511044143831 km
The number of non-african restaurants is 15
Williamsburg's population size is 32926


The average distance for East Williamsburg is 4.0881053201884 km
The number of non-african restaurants is 21
East Williamsburg's population size is 34158


The average distance for Greenpoint is 4.571171385078824 km
The number of non-african restaurants is 31
Greenpoint's population size is 34719


The average distance for Prospect Heights is 2.7096356622392506 km
The number of non-african restaurants is 31
Prospect Heights's population size is 19849
```



3.8 Analyze Each Neighborhood

With the 5 neighborhoods obtained, it is now interesting to analyze them to know the most common places. To do this, we will use the category of each venue to know the frequency of its occurrence.

We got 35 categories of the 5 neighborhoods.

Let's see the 5 most common venues for each neighborhood:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	East Williamsburg	Mexican Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	Mediterranean Restaurant
1	Greenpoint	Chinese Restaurant	Mexican Restaurant	Sushi Restaurant	French Restaurant	Restaurant
2	Prospect Heights	Mexican Restaurant	Thai Restaurant	American Restaurant	Diner	Southern / Soul Food Restaurant
3	Williamsburg	Latin American Restaurant	Middle Eastern Restaurant	Caribbean Restaurant	Chinese Restaurant	Diner
4	Windsor Terrace	Sushi Restaurant	Chinese Restaurant	Diner	American Restaurant	French Restaurant

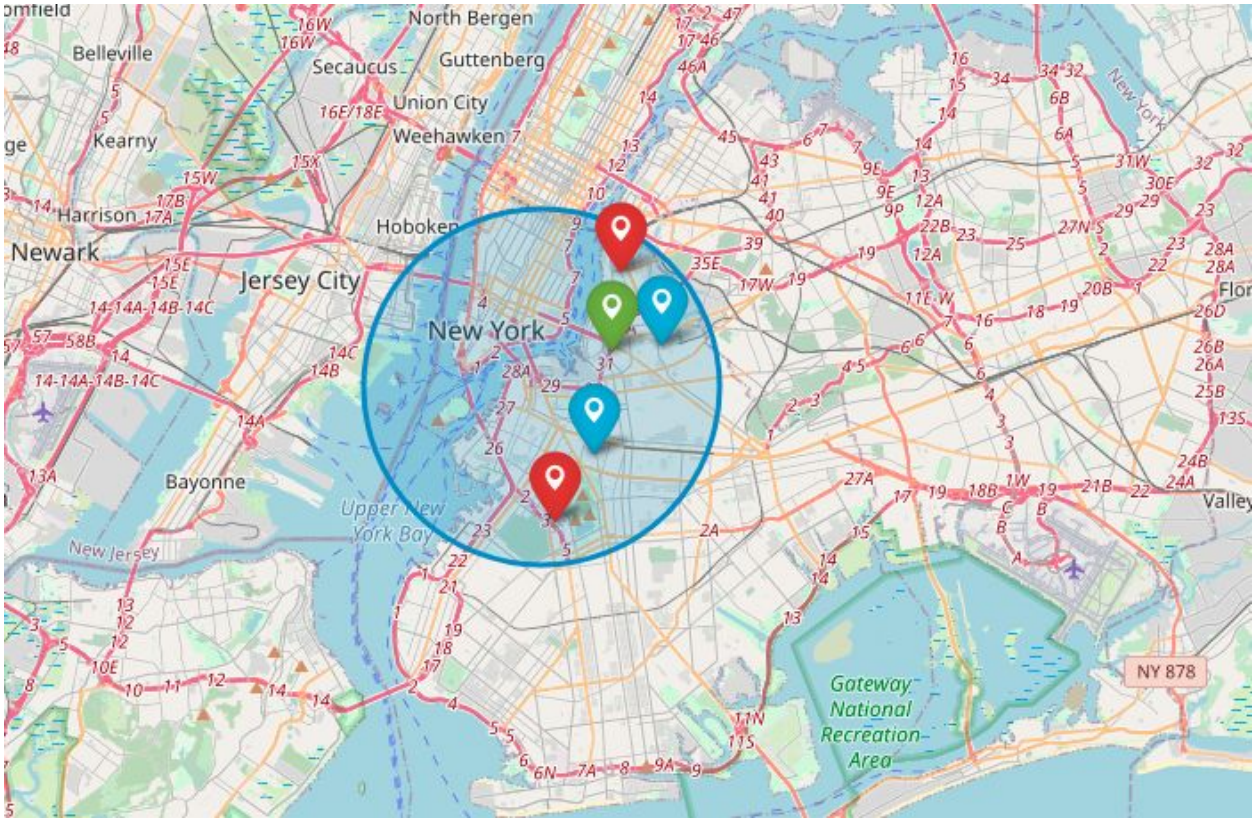
With this result, we see the most common places for each neighborhood. Then we will use the **Kmeans algorithm** to group the neighborhoods together to see which ones have restaurants in common.

4. Modeling

To give the entrepreneur an idea of which neighbourhoods have something in common, we will use the Kmeans algorithm. Since we only have five neighbourhoods, we will group the data into three clusters.

Then we merge the result with our previous data to get more information on each neighbourhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Total Numbers	distance km	Population	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
85	Windsor Terrace	40.656946	-73.980073	15	4.371090	20988	0	Sushi Restaurant	Chinese Restaurant	Diner	American Restaurant	French Restaurant
100	Williamsburg	40.707144	-73.958115	15	2.541511	32926	2	Latin American Restaurant	Middle Eastern Restaurant	Caribbean Restaurant	Chinese Restaurant	Diner
152	East Williamsburg	40.708492	-73.938858	21	4.088105	34158	1	Mexican Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	Mediterranean Restaurant
201	Greenpoint	40.730201	-73.954241	31	4.571171	34719	0	Chinese Restaurant	Mexican Restaurant	Sushi Restaurant	French Restaurant	Restaurant
	Prospect							Mexican			American	Southern / Soul Food



In this map, the circle corresponds to downtown Brooklyn projected at 100 radius.

	Cluster 0	Cluster 1	Cluster 2
Greenpoint	4.57 km 31restaurants 34719 people		
Windsor Terrace	4.37 km 15 restaurants 20988 people		
Prospect Heights		2.70 km 31 restaurants 19849 people	
East Williamsburg		4.08 km 21restaurants	

		34158 people	
Williamsburg			2.54 km 15 restaurants 32926 people

5. Results and Discussion

In this project, the first thing we saw was that there are a lot of restaurants in New York City. That's why we continued the analysis on Brooklyn, which is the most populous borough in New York. Then we focused on the 50 most populous neighborhoods in Brooklyn. This allowed us to select the 20 smallest restaurants in these neighborhoods. Since we would like to have neighborhoods that are not far from the center, we focused only on neighborhoods that are no more than five kilometres from downtown Brooklyn. In the end, we came up with 5 neighborhoods that met our expectations.

Then, With the results of the cluster, the business owner could choose in which neighborhood he would like to open his African restaurant.

If he wants the closest location with the fewest restaurants, he could choose Williamsburg (Cluster 2). In this neighborhood, the most common locations are Latin American and Middle Eastern restaurants.

If he wishes to be in a place where there are a lot of Asian restaurants (Chinese, Japanese,...), he could choose Cluster 0 (Greenpoint or Windsor Terrace).

If he wants to be in a place where there are a lot of Latin American restaurants, he could choose Cluster 1 (Prospect Heights or East Williamsburg).

6. Conclusion

The aim of this project was to help an entrepreneur or business leader to open his African restaurant in Now York. To do this, first of all, it was necessarily a matter of choosing a place where there were not enough restaurants (and no African restaurants). Second, the location must not be far from the city. Thirdly, it must necessarily be a place where there are lots of people.

The aim of this project was to help an entrepreneur or business leader open his African restaurant in Now York. The first step was to choose a location where there were not enough restaurants (and no African restaurants). Then the location had to be close to the city. Third, it had to be a place where there were lots of people.

The entrepreneur or business owner can choose between the proposed neighborhoods. But to improve these results, other factors could be taken into account. For example, to see which neighbourhoods are the safest, the areas most visited by tourists, rental prices, ...