

The Battle of the Neighborhoods

AFRICAN RESTAURANT

Introduction

- The Uhuro group is a prestigious group of African restaurants, having already won several awards for the quality of services and dishes that bring the delight and representation of African culture to any taste;
- The group's remarkable success within the African continent has attracted the attention of important investors, who intend to expand the restaurant franchise to major financial cities in the world such as New York, Dubai, Shanghai, London etc;
- In this process, the first city to receive the expansion is New York City. With this in mind, the Uhuro group decided to hire data science services to assess in which neighborhoods it would be best to install the restaurant;

Problem

- New York City is known for its diversity in international cuisine, making it possible to find the best of each continent or country in every corner of the city. This factor presents a great risk for new investments in the food industry;
- My client wants to open an African restaurant in New York City. The quality of service and field of cultures of African dishes the Uhuro group already has, the problem faced is the deficient knowledge that stakeholders have about the place where they intend to open the establishment;
- This lack of knowledge constitutes a disadvantage and risks, as the choice of the wrong place will affect the success of the venture and may alienate investors from the expansion project, so it is in the group's interest to obtain enough data for them to make decisions.

Data acquisition and cleaning

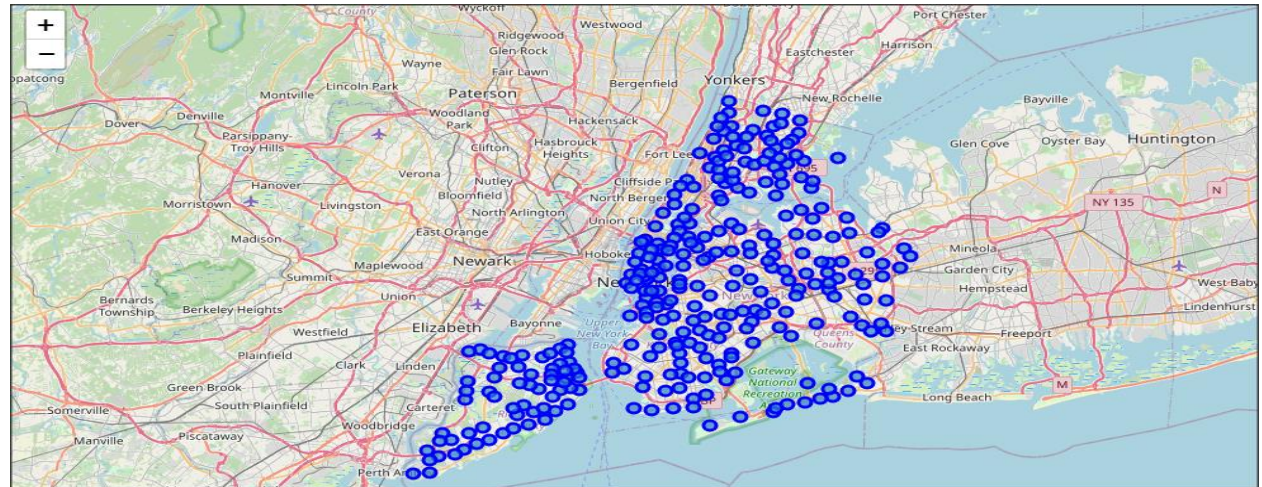
- In order to carry out the necessary analysis so that the group can make concise decisions about which neighborhood to install the restaurant in, it is necessary to obtain the list of New York City neighborhoods and have for each neighborhoods the list of nearby venues;
- The list of New York City neighborhoods was found on the following website: https://geo.nyu.edu/catalog/nyu_2451_34572, this list found in geojson format;
- To obtain the list of venues for each neighborhood we will use the Foursquare tool, where for each neighborhood a request was made for its API with 100 venues within a 500m radius

Data acquisition and cleaning

- Reading the geojson of New York City we come across the dataset and its representation on the map below:

Out[10]:

	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



Data acquisition and cleaning

- Obtaining the data from the neighborhoods, the next step is to obtain the data from the venues, to obtain this data we use the Foursquare platform, where we make requests for its API to obtain the venues, in this request we intend to obtain the 100 venues near the neighborhoods in one 500m radius, and at the end for each of the neighborhoods we list their venues. With this junction we obtain a new dataset where we have the list of each venue for each neighborhood, as shown in the figure below:

	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	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
3	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

Analysis

- Transport information about venues categories to dataset columns and convert information into numbers;
- Filter to get an African restaurant dataset;
- Group neighborhoods and calculate means;

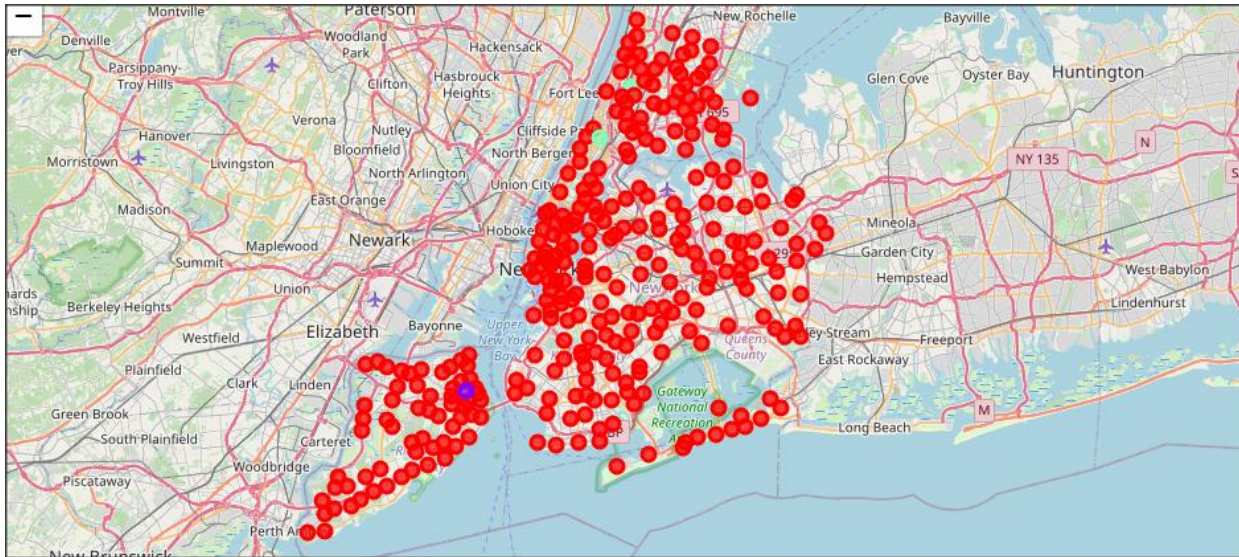
	Neighborhood	African Restaurant
0	Allerton	0.000000
1	Annadale	0.000000
2	Arden Heights	0.000000
3	Arlington	0.000000
4	Arrochar	0.000000
5	Arverne	0.000000
6	Astoria	0.000000
7	Astoria Heights	0.000000

Analysis

- Kmeans was used to perform the clustering;
- We use as desired number of clusters 3;
- The division of clusters :
 - Neighborhoods with a low rate of African restaurants;
 - Neighborhoods with an average African restaurant index;
 - Neighborhoods with a high rate of restaurants with African cuisine;

Results

- We can visualize the results obtained with the application of the analysis



Red: low index;

Purple: average index

Green: high index

Results

- As we can see in the figure above, we have more neighborhoods with low African restaurants in New York City;
- It is easier to point out which neighborhoods are not recommended to install the establishment due to competition;

Conclusion

- Few neighborhoods in New York City have restaurants with African cuisine;
- This fact not only encourages the installation of the restaurant in the city, but also represents a space for growth within the city itself, that is, being able to expand the restaurant; other neighborhoods in New York City;
- With the great quality of services of the Uhuro group and the great variety of Neighborhoods in the city of New York without African results, it is safe to say that this bet is likely to have great gains.

Thanks!

