

Analysis to locate a Colombian food restaurant in Manhattan: Final Project

Capstone Project - The Battle of Neighborhoods (Week 2)

https://github.com/ElkinSn/Coursera_Capstone/blob/main/FinalProject.ipynb

ELKIN R. SANTAFÉ R.

Jan. 19 / 2021

INTRODUCTION

Colombia is a country that within Latin America enjoys having a unique cultural richness, which is reflected in its gastronomy. The ethnic diversity of Colombia manifests itself in Colombian gastronomy. In Colombian food there's a rich mix of Spanish and Indigenous ingredients with African, Arab and Spanish preparation.

The main problem to be solved is to establish the potential of a Colombian food restaurant in Manhattan. We want the restaurant to be unique, and stand out strongly with the neighboring restaurants through the diversity in the menu. However, New York is known to have a particular Latino influence that is seen not only in the arts but also in food.

The principal ingredients in Colombian gastronomy are: pork, potatoes, beans, corn, chicken, rice and soup.

The menu will design including the most typical Colombian plates that everybody should try:



Ajiaco

Soup made with chicken and three kinds of potatoes using the special Colombian spice guasca, a Colombian specialty.



Hormiga culona

Fried ants, typical of Santander



Lechona

Stuffed pig -- stuffed with rice and meat, typical of Tolima.

Bandeja Paisa

A mix of beans, eggs, meat, rice, plantain, chorizo, fried pig skin, potatoes, and salsa typical of the Paisa region.



Sancochos

Chicken, pork and beef soup



Tamales

Stuffed with rice, chicken, and vegetables



Patacón with everything

Green, fried plantain served with different meats and cheeses.

Arepas and Empanadas

Arepa is the same as the fried plantain. The arepa is a tortilla made of corn flour. Empanadas are small, deep fried pockets filled with rice, meat, vegetables served with spicy sauce.



DATA

Download and Explore Dataset Neighborhood has a total of 5 boroughs (Manhattan, Brooklyn, Queens, The Bronx, and Staten Island) and 306 neighborhoods. What we want to do is to segment the neighborhoods and explore them. For this task it will necessary a dataset that contains the 5 boroughs information and the neighborhoods that exist in each borough. In addition we will need the latitude and longitude coordinates of each neighborhood.



Fig. 1 - Map of the five boroughs of New York (and New Jersey on the left and Long Island on the right). Area of interest is indicated.

The dataset is available on the web using the link: https://geo.nyu.edu/catalog/nyu_2451_34572.

To carry out the information from restaurants, the information available in the FOURSQUARE database will be used. This database has identification codes for different types of restaurants, including those for Latin food and in particular those for Colombian food.

As a first step, the search will focus on New York (Manhattan) and then we will proceed to search all the existing Colombian food restaurants. This will allow us to establish which are the areas that contain the highest density of Colombian restaurants and which ones do not. In this way, we will quickly identify which are the areas with the greatest potential to locate a Colombian food restaurant.

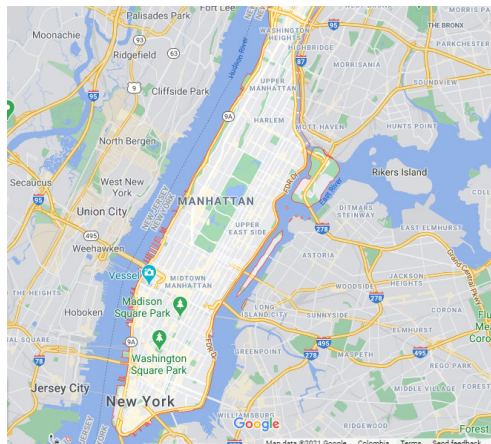


Fig. 2 - Study's area for locating a Colombian Food Restaurant

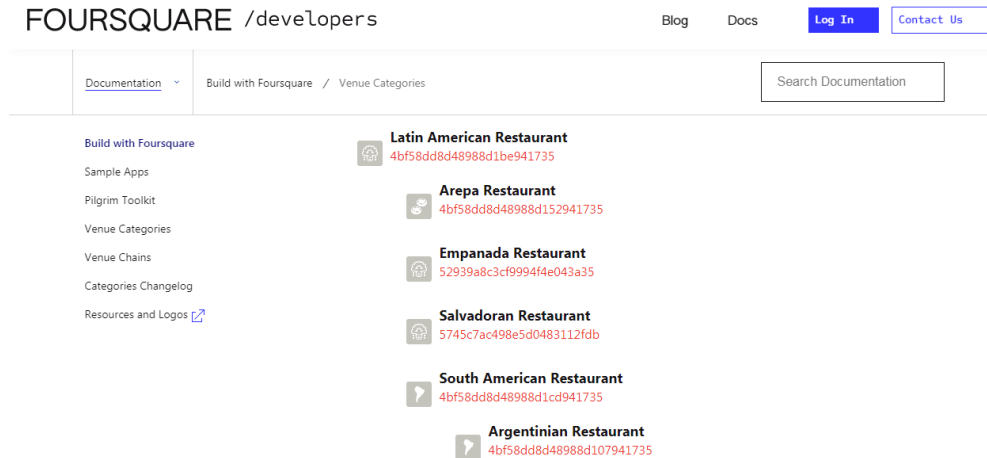


Fig. 3 – ID codes for restaurants from FOURSQUARE

METHODOLOGY

The workflow followed to implement the solution was following:

1. Preparing data base
 - a. Importing Libraries
 - b. Downloading files and placing them on the server
 - c. Transforming the data into a *pandas* data frame
2. Making the base map
 - a. Using *geopy* library to get the latitude and longitude values of New York City.
 - b. Creating a map of New York with neighborhoods
3. Integrating the information of FOURSQUARE venues
 - a. Looking for Colombian restaurants
 - b. Looking for Latino Restaurants
4. Generating competitors map
5. Grouping the information by neighborhoods.
6. Analyzing by neighborhood.
 - a. Raking the venues
7. Clustering the neighborhood
 - a. Clustering process
 - b. Generating new table
 - c. Generating clusters map
 - d. Generating tables by cluster (with the top 10 of venues)
8. Analyzing the information (from tables)

The Machine Learning analysis was implemented in the clustering section where using the *Kmeans* criteria for grouping the neighborhoods. A clustering process generated different tables (ranked on 10) with the top ten of venues related to Latino restaurants. The idea is located the cluster with the less offer of Latino food.

RESULTS

According to the methodology followed, the results by steps are exposed below.

Preparing data base

In this step we import the libraries and we used the link with the information about boroughs and neighborhoods to condense all 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

Fig. 4 – Table result of preparing data base

Making the base map

In this step we import the libraries and we applied the functions over all data for getting the latitude and longitude of each neighborhood. Next, a map was generated for visualize the information.

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Fig. 5 – Table with coordinates of neighborhoods

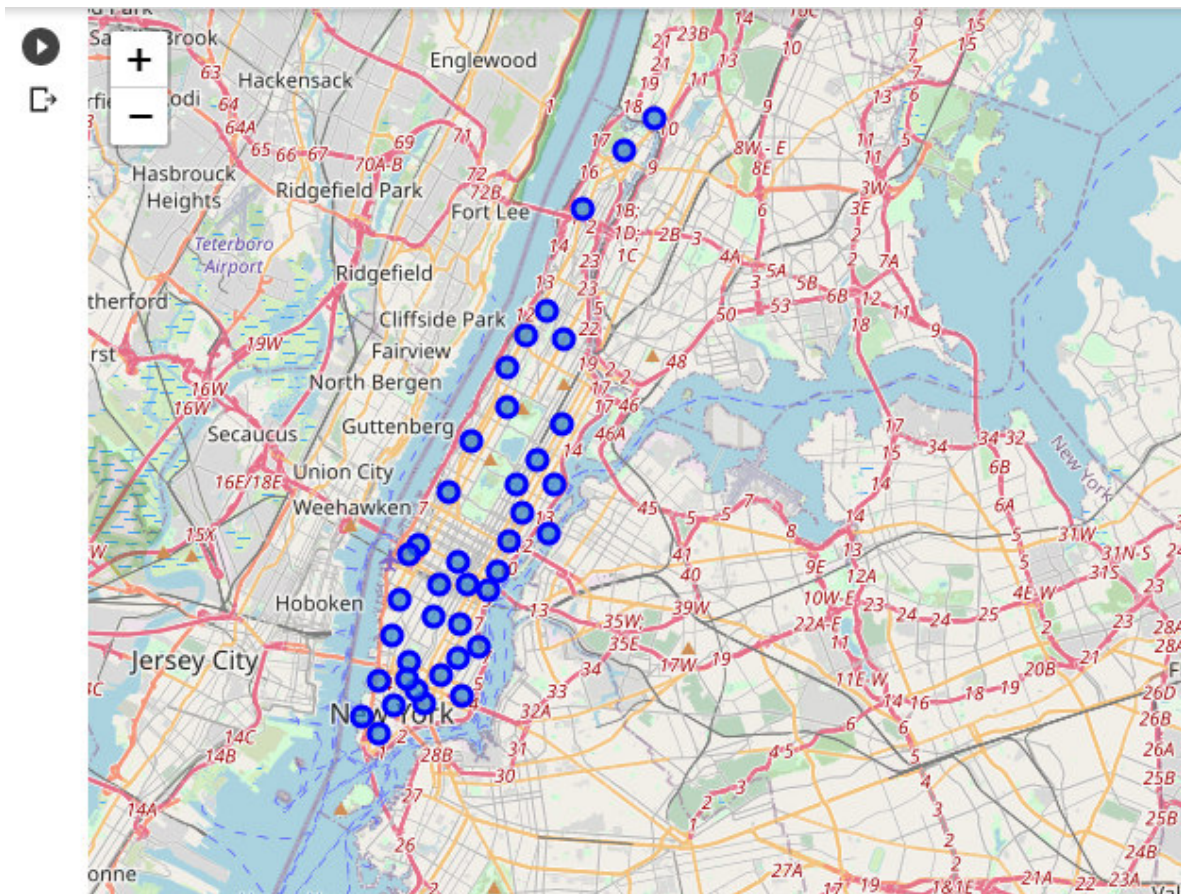


Fig. 6 – Map with neighborhoods

Integrating the information of FOURSQUARE venues

Due the idea is looking for Colombian restaurants (which are the potential competitors), I set the search using the FOURSQUARE ID for Colombian food. The result is the Figure 7.

>> Looking for Colombian Restaurants

```
[ ] #https://developer.foursquare.com/docs/resources/categories
#Colombian Restaurant: 58daa1558bbb0b01f18ec1f4
neighborhoods = neighborhoods[neighborhoods['Borough'] == 'Manhattan'].reset_index(d
newyork_venues_colombian = getNearbyVenues(names=neighborhoods['Neighborhood'], lati
newyork_venues_colombian.head()
```

https://api.foursquare.com/v2/venues/search?client_id=QSQC42XYF50MXN2PNVTIVQ5ASUGUY5

```
{'meta': {'code': 200, 'requestId': '6004a6ecba07166a6ac3254e'}, 'response': {'venue
[]
Empty DataFrame
Columns: []
Index: []
```

Fig. 7 – Results of searching Colombian food.

It can be seen from the result that there are no restaurants specialized in Colombian food in the area. This is good news. However, as is often the case in some countries, the Colombian food menu can be integrated into Latin food menu cards. So what we will do now is to look for Latin food restaurants within the area and see how they are distributed in order to identify the opportunities.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Hola España	40.879160	-73.904553	Latin American Restaurant
1	Marble Hill	40.876551	-73.91066	La Caridad	40.869219	-73.903219	Latin American Restaurant
2	Marble Hill	40.876551	-73.91066	Leche y Miel	40.883742	-73.901857	Latin American Restaurant
3	Marble Hill	40.876551	-73.91066	La Tia	40.878808	-73.905187	Latin American Restaurant
4	Marble Hill	40.876551	-73.91066	Empanadas Monumental	40.883476	-73.901964	Empanada Restaurant

Fig. 8 – Results of searching Latino food restaurants.

Generating competitors map

The information obtained from the previous step is dense. The map shown in the figure 9 expose a high amount of Latino food restaurants. The competition is high. So what will be done next is to group the information by neighborhoods and at the same time rank the top 10 positions for each group.

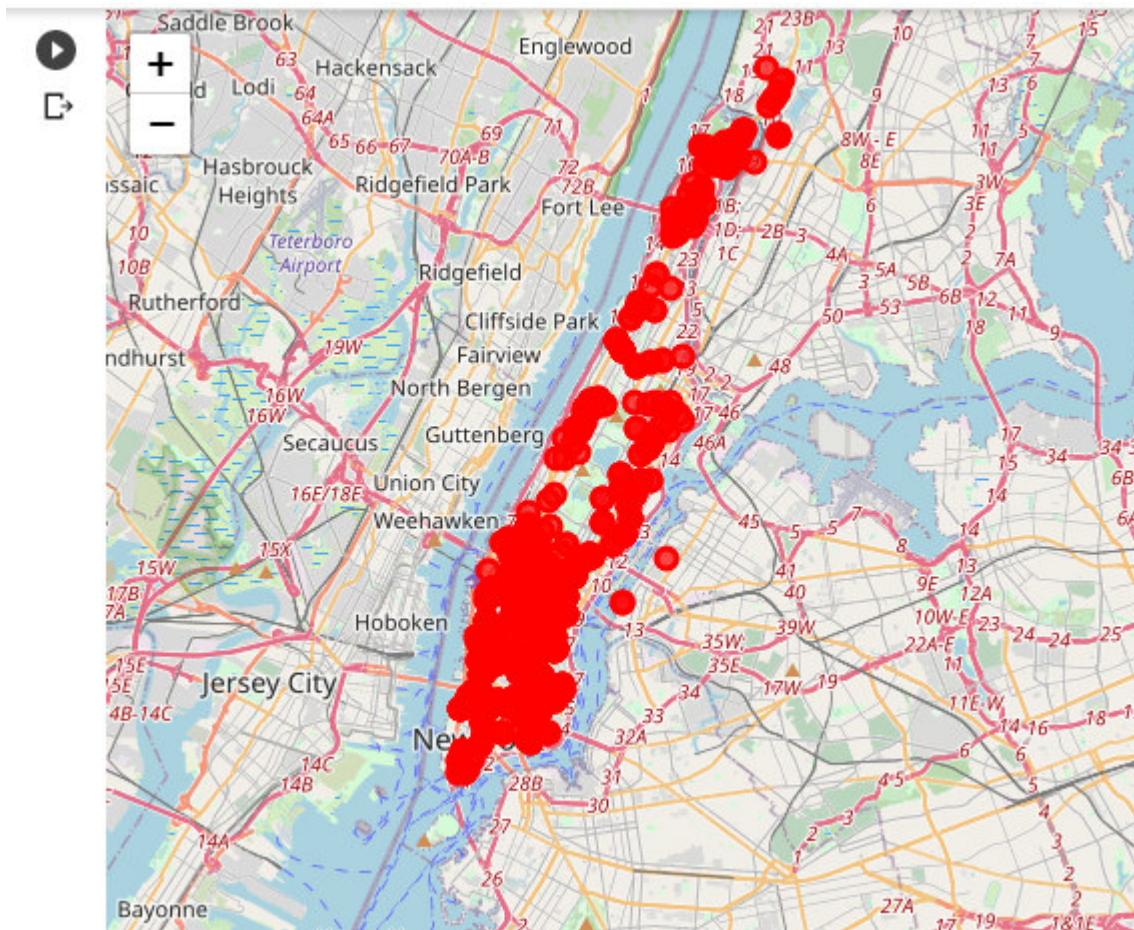


Fig. 9 – Results of searching Latino food restaurants.

Grouping the information by neighborhoods.

What has been done now is to generate a new table that groups the information by neighborhoods.

East Village	40	40	40	40	40	40
Financial District	8	8	8	8	8	8
Flatiron	32	32	32	32	32	32
Gramercy	25	25	25	25	25	25
Greenwich Village	35	35	35	35	35	35
Hamilton Heights	11	11	11	11	11	11
Hudson Yards	34	34	34	34	34	34
Inwood	36	36	36	36	36	36
Lenox Hill	12	12	12	12	12	12
Lincoln Square	20	20	20	20	20	20
Little Italy	41	41	41	41	41	41
Lower East Side	39	39	39	39	39	39
Manhattan Valley	16	16	16	16	16	16
Manhattanville	10	10	10	10	10	10
Marble Hill	14	14	14	14	14	14

Fig. 10 – Results of searching Latino food restaurants.

Analyzing by neighborhood.

What has been done now is to generate tables with the categories associated to each neighborhood for applying a ranking.

	Neighborhood	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	Bakery	Bar	Brazilian Restaurant	Café	Caribbean Restaurant	Churrascaria	Cuban Restaurant	Deli / Bodega	Empanada Restaurant	Food Truck
0	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	1	0

	Neighborhood	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	Bakery	Bar	Brazilian Restaurant	Café	Caribbean Restaurant	Churrascaria	Cuban Restaurant	Deli / Bodega	Empanada Restaurant	Food Truck
0	Battery Park City	0.111111	0.111111	0.000000	0.000000	0.000000	0.111111	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.1
1	Carnegie Hill	0.090909	0.090909	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
2	Central Harlem	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.1
3	Chelsea	0.086957	0.130435	0.000000	0.043478	0.000000	0.043478	0.000000	0.043478	0.000000	0.043478	0.000000	0.000000	0.0
4	Chinatown	0.076923	0.102564	0.000000	0.000000	0.000000	0.051282	0.000000	0.000000	0.000000	0.000000	0.025641	0.000000	0.0
5	Civic Center	0.222222	0.166667	0.000000	0.000000	0.000000	0.055556	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
6	Clinton	0.042553	0.085106	0.000000	0.000000	0.000000	0.170213	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.1
7	East Harlem	0.058824	0.058824	0.000000	0.000000	0.058824	0.000000	0.058824	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
8	East Village	0.050000	0.075000	0.000000	0.000000	0.000000	0.075000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
9	Financial District	0.000000	0.125000	0.000000	0.000000	0.000000	0.125000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.1
10	Flatiron	0.125000	0.093750	0.000000	0.031250	0.000000	0.062500	0.000000	0.031250	0.000000	0.031250	0.031250	0.031250	0.0

Fig. 11 – Results of searching Latino food restaurants.

Raking the venues

The idea of ranking the places is to be able to define the real competition to which the future restaurant could be subjected.

C*	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Battery Park City	Latin American Restaurant	Arepa Restaurant	Brazilian Restaurant	Empanada Restaurant	Middle Eastern Restaurant	Peruvian Restaurant	South American Restaurant	Argentinian Restaurant	Bar	Café
1	Carnegie Hill	Latin American Restaurant	Peruvian Restaurant	Arepa Restaurant	Argentinian Restaurant	Food Truck	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Caribbean Restaurant
2	Central Harlem	Latin American Restaurant	Food Truck	Spanish Restaurant	Empanada Restaurant	Wine Bar	Caribbean Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Café
3	Chelsea	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Arepa Restaurant	Bakery	Brazilian Restaurant	French Restaurant	Caribbean Restaurant	Cuban Restaurant	Mexican Restaurant
4	Chinatown	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Arepa Restaurant	South American Restaurant	Brazilian Restaurant	Mexican Restaurant	Tapas Restaurant	Venezuelan Restaurant	Deli / Bodega

Fig. 12 – Raking the best venues by neighborhood.

Clustering the neighborhoods.

The process of clustering neighborhoods allows for better organization of information. 10 clusters have been generated within which the list of the best 10 sites has been placed (Figure 13). Additionally, a map with color codes has been created indicating the location of these clusters (Figure 14) in order to see where the best opportunity area for a Colombian food restaurant is located. Additionally, the tables of each cluster (Figures 15 to 24) have been generated to be able to see which neighborhoods are within each cluster and which are the places that are there providing Latin food service. This will later serve as a filter to choose the cluster that groups the fewest neighborhoods and within this to identify which could be the competitors of a future restaurant.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	2	Latin American Restaurant	Tapas Restaurant	Lounge	Empanada Restaurant	Wine Bar	Churrascaria	Food Truck	Deli / Bodega
1	Manhattan	Chinatown	40.715618	-73.994279	1	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Arepa Restaurant	South American Restaurant	Brazilian Restaurant	Mexican Restaurant	Tapas Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	8	Latin American Restaurant	Empanada Restaurant	South American Restaurant	Seafood Restaurant	Caribbean Restaurant	Arepa Restaurant	Italian Restaurant	Spanish Restaurant
3	Manhattan	Inwood	40.867684	-73.921210	2	Latin American Restaurant	Venezuelan Restaurant	Spanish Restaurant	Lounge	Empanada Restaurant	Wine Bar	Churrascaria	Deli / Bodega
4	Manhattan	Hamilton Heights	40.823604	-73.949688	2	Latin American Restaurant	Empanada Restaurant	Wine Bar	Churrascaria	French Restaurant	Food Truck	Deli / Bodega	Cuban Restaurant

Fig. 13 – Raking the best venues by neighborhood.

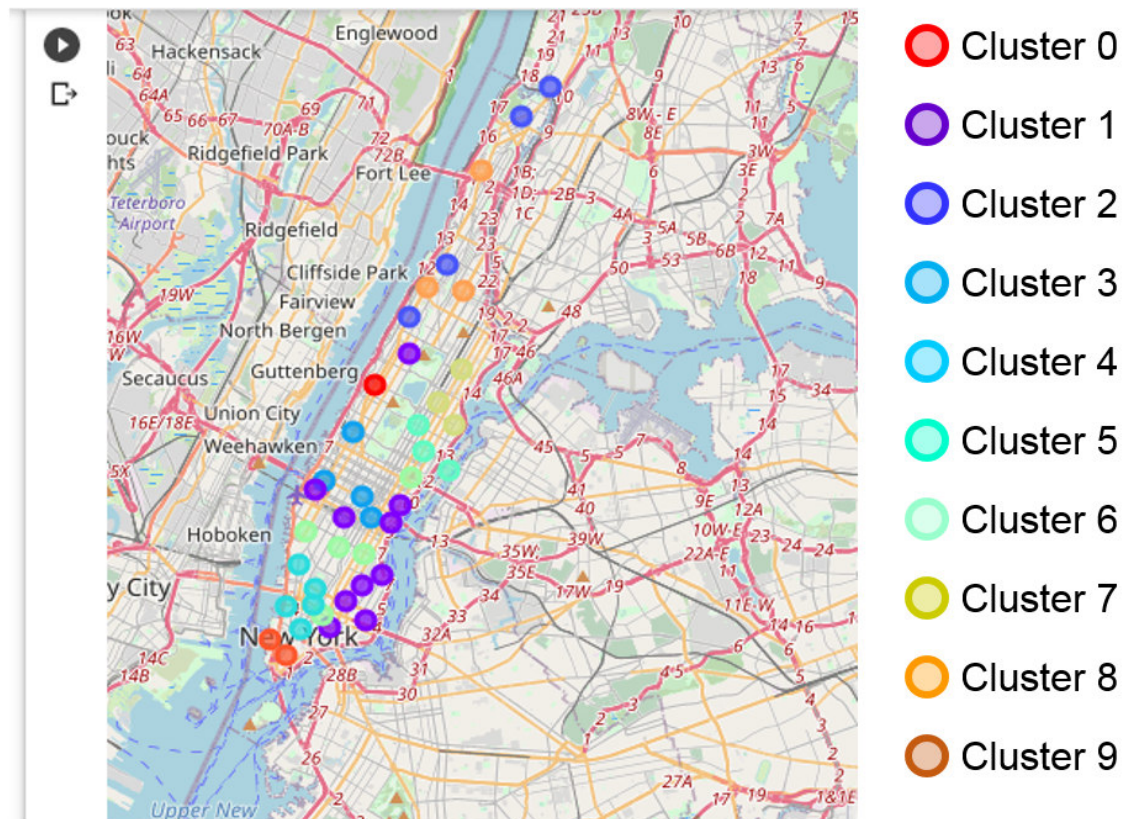


Fig. 14 – Raking the best venues by neighborhood.

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
12 Upper West Side	South American Restaurant	Peruvian Restaurant	Latin American Restaurant	Brazilian Restaurant	Spanish Restaurant	Caribbean Restaurant	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria

Fig. 15 – Neighborhoods and restaurants in cluster 0.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Chinatown	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Arepa Restaurant	South American Restaurant	Brazilian Restaurant	Mexican Restaurant	Tapas Restaurant	Venezuelan Restaurant	Deli / Bodega
19	East Village	Latin American Restaurant	Argentinian Restaurant	Brazilian Restaurant	Arepa Restaurant	Mexican Restaurant	Tapas Restaurant	South American Restaurant	Empanada Restaurant	Peruvian Restaurant	Food Truck
20	Lower East Side	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Empanada Restaurant	South American Restaurant	Brazilian Restaurant	Tapas Restaurant	Mexican Restaurant	Venezuelan Restaurant	Deli / Bodega
25	Manhattan Valley	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Argentinian Restaurant	Brazilian Restaurant	Café	Churrascaria	Food Truck	Empanada Restaurant	Deli / Bodega
31	Noho	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Empanada Restaurant	Arepa Restaurant	Tapas Restaurant	Brazilian Restaurant	South American Restaurant	Mexican Restaurant	Venezuelan Restaurant
33	Midtown South	Latin American Restaurant	Brazilian Restaurant	South American Restaurant	Arepa Restaurant	Empanada Restaurant	Peruvian Restaurant	Argentinian Restaurant	Food Truck	Office	Caribbean Restaurant
35	Turtle Bay	Latin American Restaurant	Empanada Restaurant	Peruvian Restaurant	Seafood Restaurant	Food Truck	Brazilian Restaurant	Jazz Club	Argentinian Restaurant	Bar	Café
36	Tudor City	Latin American Restaurant	Empanada Restaurant	Argentinian Restaurant	Food Truck	Brazilian Restaurant	Peruvian Restaurant	Wine Bar	Churrascaria	Deli / Bodega	Cuban Restaurant
37	Stuyvesant Town	Latin American Restaurant	Argentinian Restaurant	Wine Bar	Brazilian Restaurant	Food Truck	Deli / Bodega	Peruvian Restaurant	Tapas Restaurant	Bar	Café
39	Hudson Yards	Latin American Restaurant	Empanada Restaurant	Peruvian Restaurant	Brazilian Restaurant	Argentinian Restaurant	Office	French Restaurant	Arepa Restaurant	Steakhouse	Café

Fig. 16 – Neighborhoods and restaurants in cluster 1.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Marble Hill	Latin American Restaurant	Tapas Restaurant	Lounge	Empanada Restaurant	Wine Bar	Churrascaria	Food Truck	Deli / Bodega	Cuban Restaurant	Caribbean Restaurant
3	Inwood	Latin American Restaurant	Venezuelan Restaurant	Spanish Restaurant	Lounge	Empanada Restaurant	Wine Bar	Churrascaria	Deli / Bodega	Cuban Restaurant	Caribbean Restaurant
4	Hamilton Heights	Latin American Restaurant	Empanada Restaurant	Wine Bar	Churrascaria	French Restaurant	Food Truck	Deli / Bodega	Cuban Restaurant	Caribbean Restaurant	Jazz Club
26	Morningside Heights	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Wine Bar	Food Truck	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Caribbean Restaurant

Fig. 17 – Neighborhoods and restaurants in cluster 2.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
13	Lincoln Square	Brazilian Restaurant	Peruvian Restaurant	Latin American Restaurant	Argentinian Restaurant	Empanada Restaurant	Mexican Restaurant	Pizza Place	Arepa Restaurant	South American Restaurant	Bar
14	Clinton	Latin American Restaurant	Brazilian Restaurant	Peruvian Restaurant	Empanada Restaurant	Argentinian Restaurant	South American Restaurant	Arepa Restaurant	Steakhouse	Food Truck	Office
15	Midtown	Latin American Restaurant	Brazilian Restaurant	Empanada Restaurant	South American Restaurant	Arepa Restaurant	Food Truck	Argentinian Restaurant	Peruvian Restaurant	Steakhouse	Office
16	Murray Hill	Latin American Restaurant	Brazilian Restaurant	Empanada Restaurant	Food Truck	South American Restaurant	Argentinian Restaurant	Arepa Restaurant	Deli / Bodega	Peruvian Restaurant	Italian Restaurant

Fig. 18 – Neighborhoods and restaurants in cluster 3.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
18	Greenwich Village	Latin American Restaurant	Peruvian Restaurant	Arepa Restaurant	Brazilian Restaurant	Argentinian Restaurant	Mexican Restaurant	Empanada Restaurant	Spanish Restaurant	South American Restaurant	Nightclub
21	Tribeca	Peruvian Restaurant	Latin American Restaurant	Arepa Restaurant	Brazilian Restaurant	Argentinian Restaurant	Spanish Restaurant	Asian Restaurant	Bakery	Bar	Café
23	Soho	Latin American Restaurant	Peruvian Restaurant	Arepa Restaurant	Argentinian Restaurant	Brazilian Restaurant	Spanish Restaurant	Mexican Restaurant	Tapas Restaurant	Empanada Restaurant	Asian Restaurant
24	West Village	Peruvian Restaurant	Brazilian Restaurant	Latin American Restaurant	Arepa Restaurant	Spanish Restaurant	Bakery	Nightclub	Caribbean Restaurant	Mexican Restaurant	Tapas Restaurant
32	Civic Center	Latin American Restaurant	Arepa Restaurant	Peruvian Restaurant	Argentinian Restaurant	Brazilian Restaurant	Middle Eastern Restaurant	Asian Restaurant	Bakery	Bar	Café

Fig. 19 – Neighborhoods and restaurants in cluster 4.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
8	Upper East Side	Latin American Restaurant	Juice Bar	Argentinian Restaurant	Arepa Restaurant	Spanish Restaurant	French Restaurant	Tapas Restaurant	Asian Restaurant	Bakery	Bar
10	Lenox Hill	Latin American Restaurant	Juice Bar	Empanada Restaurant	Arepa Restaurant	Spanish Restaurant	French Restaurant	Argentinian Restaurant	Asian Restaurant	Bakery	Bar
11	Roosevelt Island	Latin American Restaurant	Arepa Restaurant	Churrascaria	French Restaurant	Food Truck	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Caribbean Restaurant	Jazz Club

Fig. 20 – Neighborhoods and restaurants in cluster 5.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
17	Chelsea	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Arepa Restaurant	Bakery	Brazilian Restaurant	French Restaurant	Caribbean Restaurant	Cuban Restaurant	Mexican Restaurant
22	Little Italy	Latin American Restaurant	Arepa Restaurant	Argentinian Restaurant	Peruvian Restaurant	South American Restaurant	Brazilian Restaurant	Mexican Restaurant	Tapas Restaurant	Venezuelan Restaurant	Deli / Bodega
27	Gramercy	Latin American Restaurant	Argentinian Restaurant	Brazilian Restaurant	Mexican Restaurant	Arepa Restaurant	Nightclub	Empanada Restaurant	Deli / Bodega	Peruvian Restaurant	Food Truck
34	Sutton Place	Latin American Restaurant	Empanada Restaurant	Juice Bar	Argentinian Restaurant	Food Truck	Jazz Club	Arepa Restaurant	Seafood Restaurant	Puerto Rican Restaurant	Steakhouse
38	Flatiron	Latin American Restaurant	Arepa Restaurant	Argentinian Restaurant	Brazilian Restaurant	Mexican Restaurant	Caribbean Restaurant	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Nightclub

Fig. 21 – Neighborhoods and restaurants in cluster 6.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	East Harlem	Latin American Restaurant	Arepa Restaurant	Puerto Rican Restaurant	Mexican Restaurant	Café	Peruvian Restaurant	Bar	Argentinian Restaurant	Brazilian Restaurant	Bakery
9	Yorkville	Latin American Restaurant	Arepa Restaurant	Peruvian Restaurant	Argentinian Restaurant	Food Truck	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Caribbean Restaurant
30	Carnegie Hill	Latin American Restaurant	Peruvian Restaurant	Arepa Restaurant	Argentinian Restaurant	Food Truck	Empanada Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Caribbean Restaurant

Fig. 22 – Neighborhoods and restaurants in cluster 7.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Washington Heights	Latin American Restaurant	Empanada Restaurant	South American Restaurant	Seafood Restaurant	Caribbean Restaurant	Arepa Restaurant	Italian Restaurant	Spanish Restaurant	Asian Restaurant	Bakery
5	Manhattanville	Latin American Restaurant	Food Truck	Spanish Restaurant	Empanada Restaurant	Wine Bar	Caribbean Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Café
6	Central Harlem	Latin American Restaurant	Food Truck	Spanish Restaurant	Empanada Restaurant	Wine Bar	Caribbean Restaurant	Deli / Bodega	Cuban Restaurant	Churrascaria	Café

Fig. 23 – Neighborhoods and restaurants in cluster 8.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
28	Battery Park City	Latin American Restaurant	Arepa Restaurant	Brazilian Restaurant	Empanada Restaurant	Middle Eastern Restaurant	Peruvian Restaurant	South American Restaurant	Argentinian Restaurant	Bar	Café
29	Financial District	Latin American Restaurant	Peruvian Restaurant	Argentinian Restaurant	Empanada Restaurant	South American Restaurant	Brazilian Restaurant	Middle Eastern Restaurant	Churrascaria	Deli / Bodega	Cuban Restaurant

Fig. 24 – Neighborhoods and restaurants in cluster 9.

DISCUSION

When analyzing the clusters, what can be observed is that the area with the greatest potential for development chosen is the one with the least amount of neighborhoods. In this case, we have chosen cluster 0. Although there is only one neighborhood, this is positive given that the level of competition is lower. In this way, Colombian food can be introduced on a smaller scale but benefiting from voice to voice. Within cluster 0, what stands out the most is that in the top 10 of the restaurants there is a Latin American restaurant and a Caribbean restaurant, which could include Colombian food in their menu, but could not go into many details given the wide range which may be the gastronomy of these areas. It is also striking that there is an *Empanadas* restaurant, which suggests direct competition with one of the products included in the menu, which should be removed from the food menu.

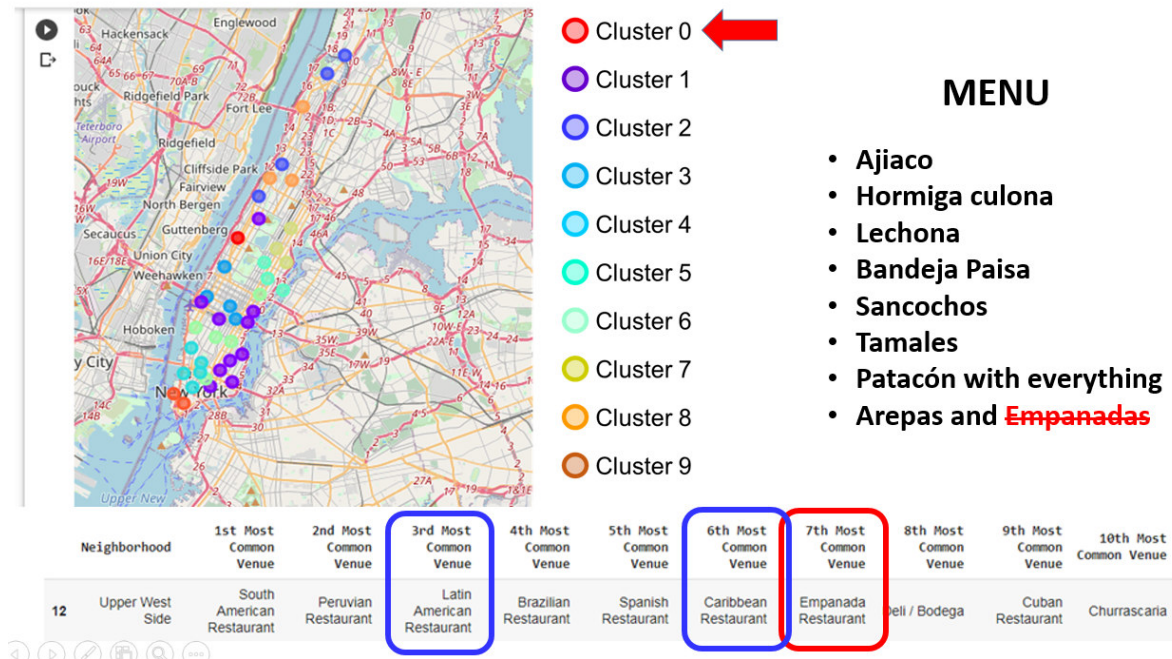


Fig. 25 – Analysis performed on Cluster 0.

CONCLUSIONS

- For the final IBM Capstone project it was chosen to work a small part of the marketing study to open a Colombian food restaurant in Manhattan.
- Within the implemented program, the unsupervised learning algorithm *Kmeans* was used to perform clustering support. Folio was also used to develop the maps and information from the FOURSQUARE database to generate categories and rank data.
- As a result of the study, it was possible to identify an area of Manhattan (Upper West Side) as a potential to open this restaurant.
- In the identified area, 2 indirect competitors were detected (one for Latin American food and another for Caribbean food) which could include Colombian food in their menus, but it is based on the fact that because the offer is very wide, very possibly they have discarded many of the dishes that they intend to offer.
- In the identified area, 1 direct competitor was detected (empanada restaurant) which is specialized in a single product that is also found in the menu that is being proposed. This suggests that such a product should exit the menu, at least initially.

References.

- <https://www.colombia.com/gastronomia/recetas-colombianas/ajiaco-r3>
- <https://www.vanguardia.com/area-metropolitana/bucaramanga/convocan-a-artistas-a-pintar-hormigas-culonas-como-simbolo-regional-LYVL311684>
- [https://github.com/Arunk-is-here/IBMcapstone/blob/b29d0389e4869bdbcc5580569706df12b991adcb/Capstone%20Project%20-%20The%20Battle%20of%20Neighborhoods%20\(Week%201\).ipynb](https://github.com/Arunk-is-here/IBMcapstone/blob/b29d0389e4869bdbcc5580569706df12b991adcb/Capstone%20Project%20-%20The%20Battle%20of%20Neighborhoods%20(Week%201).ipynb)
- <https://www.mcomidacolombiana.site/platos-fuertes-colombianos/lechona/>
- <https://aulanews.uao.es/2020/04/02/la-bandeja-paisa-de-colombia-al-paladar/>
- <https://antojandoando.com/recetas/arroces/tamales-colombianos/>