Capstone Project - The Battle of Neighborhoods

Belén León Pérez

July 31, 2021

Section 1 | Introduction

Finding a location for a new Italian restaurant in Barcelona

Barcelona is the fifth most populous urban area in the European Union after Paris, the Ruhr area, Madrid, and Milan.

Barcelona is one of the largest metropolises on the Mediterranean Sea.

Also, we choose Barcelona because it is one of the most visited cities in the world by international visitors.

Italian food is really popular around the world, so this project is targeted to people who want to open a new Italian restaurant in Barcelona.

Finding a location for a new Italian restaurant in Barcelona

Criteria

The location is not crowed with restaurants.

The location has not Italian restaurants near.

If it is possible, we want a location close to the city center.

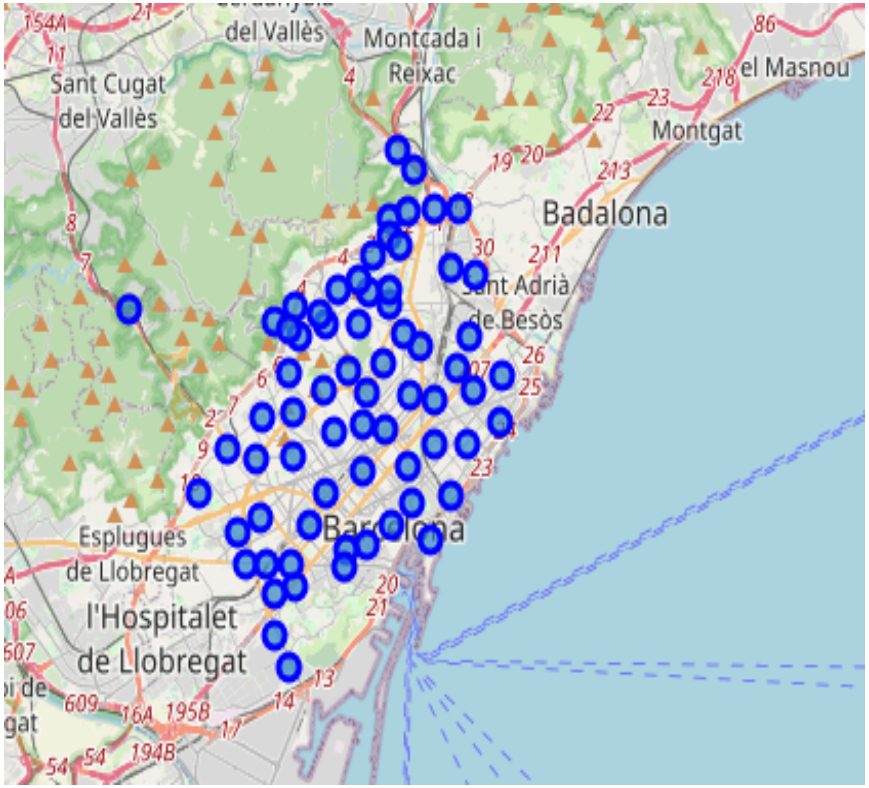
Section 2 | Data

Data used and sources of the data

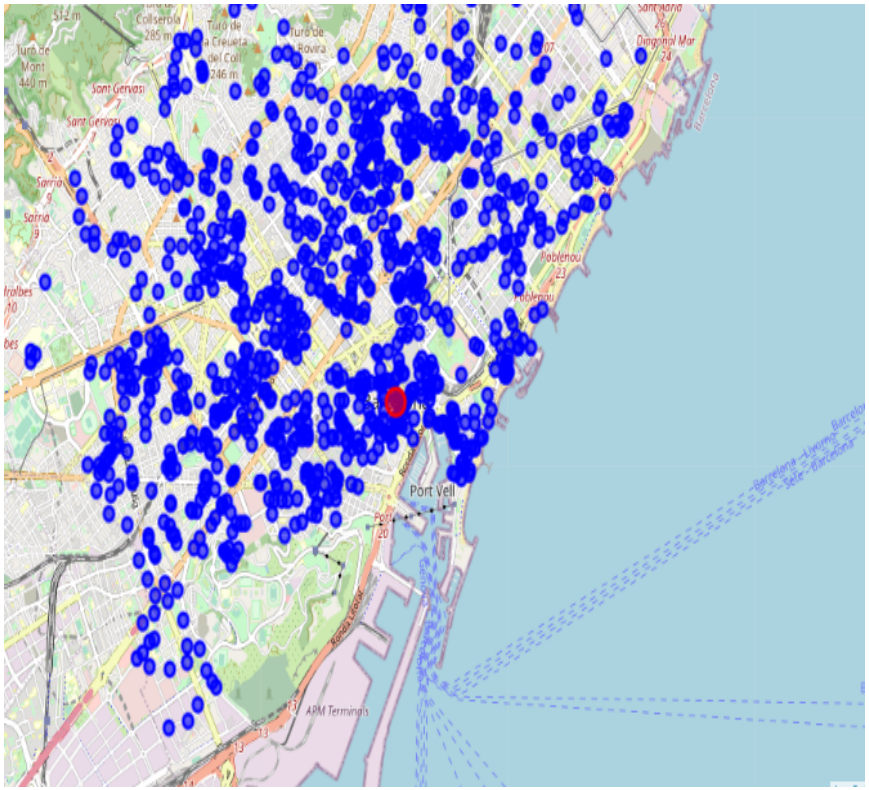
We will need the information about the neighborhoods in Barcelona, the restaurants in each neighborhood and also locations where you can eat Italian food.

We obteined the information about neighborhoods in Barcelona Open Data: (2019): opendata-ajuntament.barcelona.cat

We obtained the number of restaurants and their type and location in every neighborhood using Foursquare API



Barcelona is divided into 73 neighborhoods.



There are 1070 restaurants and Italian food venues in Barcelona.

Ten most common restaurants in Barcelona

	categories	Count
67	Restaurant	228
73	Spanish Restaurant	150
53	Mediterranean Restaurant	147
21	Chinese Restaurant	57
5	Asian Restaurant	44
7	Bar	41
79	Tapas Restaurant	39
47	Japanese Restaurant	34
46	Italian Restaurant	31
27	Diner	22

There are 228 restaurants without an specific category.

Most common restaurant is Spanish Restaurant.

Italian restaurants are in eigth place.

Section 3 | Methodology

Analytic approach

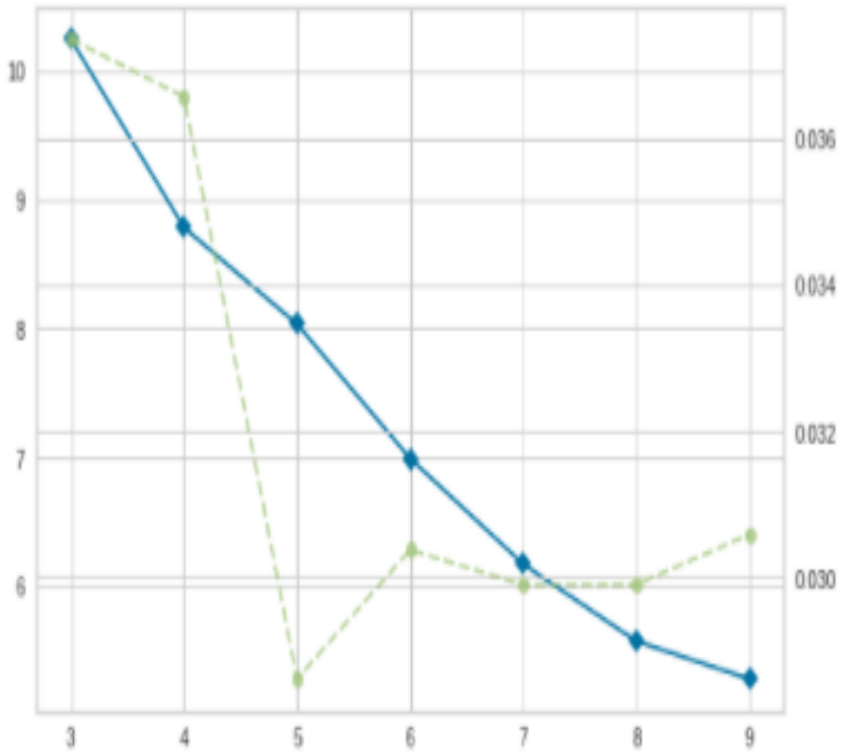
We want to cluster the neighborhoods based on their similarities, that are the number and the kind of restaurant.

We will use the approach of K-Means algorithm to extract the clusters.

We will produce a map and make some arguments on the final result.

The first step in this algorithm it is determinate the best "k".

Analytical approach of the elbow method



After this we choosed $k = 5$.

Now we can use the K-Means algorithm to cluster the neighborhood based on the restaurants categories.

Last step is define a method to start analyzing the clusters.

Defining a method to start analyzing the clusters

Method

We will chose the two clusters with more locations close to the city center.

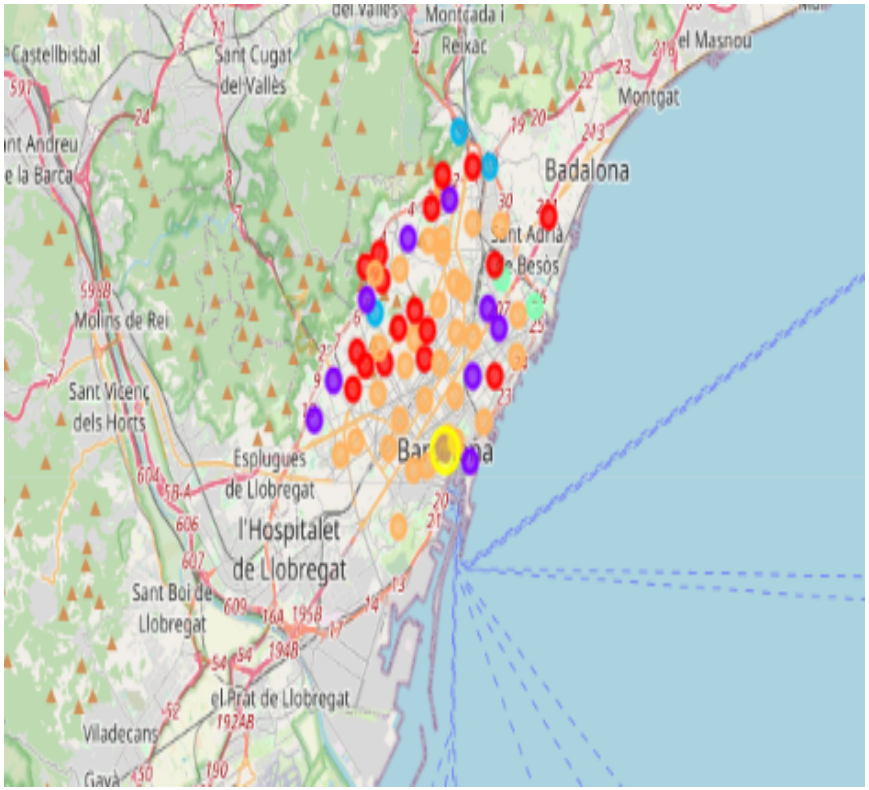
In each cluster we will calculate the distance from the neihgborhood to the city center.

We will check the unique categories in our data set and how many times each category it is repeated.

If we can not choose a location in the two clusters we will repeat the steps with the next two clusters closer to the center.

Section 4 | Results

Results obtained



We have the five clusters in a folium map

The yellow circle is the center of the city.

Following the method we choosed the Cluster 2 (purple) and the Cluster 5 (orange) to start looking for a good location.

Cluster 2

2	39584.277831
3	3894.471740
7	5772.621671
10	5511.820557
12	3922.353529
18	4144.572656
20	345.168480
22	5020.546333
39	70.598895
40	2296.787147
41	2821.179667
46	5148.935413
48	1074.042898
55	6601.346096

The closer neighborhood is 70.6 meters from the city center.

There is another one close to the city center, 345.17 meters.

The closer neihborhood is el Gòtic, the second one is Santa Catarina.

Cluster 2: closest neighborhoods

	categories	Count
8	Mediterranean Restaurant	11
14	Spanish Restaurant	10
13	Restaurant	6
7	Italian Restaurant	3
10	Paella Restaurant	2
16	Tapas Restaurant	2
6	Halal Restaurant	1
5	Greek Restaurant	1
1	Café	1
9	Moroccan Restaurant	1

In el Gòtic, there are 3 Italian restaurants and in total 38 restaurants . Distance: 70.6 metres.

	categories	Count
3	Restaurant	2
4	Spanish Restaurant	2
0	Bar	1
1	Italian Restaurant	1
2	Mediterranean Restaurant	1

In Santa Catarina there is 1 Italian restaurants and in total 7 restaurants.
Distance: 345.17 metres.

Cluster 5

5	3181.637985
6	3195.485961
9	62528.370793
11	5585.090401
13	6087.545847
14	1401.222932
19	25059.617926
21	2795.772779
26	6790.030261
27	5233.204112
29	4687.322844
30	6335.772689
32	3125.719987
33	5100.397430
34	3205.957561
35	34957.592773
37	1503.427678
38	3881.048304
42	3879.467563
43	845.590253
44	5537.032607
45	4453.637587
47	2034.804991
49	1596.424780
51	4344.861552
53	2378.069748
56	503.878118

The closer neighborhood is 503.9 meters from the city center.

The next one is too far from the city center, 845.6 meters.

The closer neihborhood is la Ribera.

Cluster 5: closest neighborhood

	categories	Count
6	Mediterranean Restaurant	4
13	Tapas Restaurant	3
7	Mexican Restaurant	2
10	Restaurant	2
11	Seafood Restaurant	2
12	Spanish Restaurant	2
0	Argentinian Restaurant	1
1	Breakfast Spot	1
2	Café	1
3	Cantonese Restaurant	1
4	Comfort Food Restaurant	1
5	Food Court	1
8	Molecular Gastronomy Restaurant	1
9	Paella Restaurant	1

There is not Italian restaurant in la Ribera and in total there are 23 restaurants.

Distance: 503.878 meters.

In this neighborhood there are more restaurants than in Santa Catarina.

Section 5 | Discussion

We have two good locations: Ia Ribera and Santa Catarina.

La Ribera has more restaurants but tere is not an Italian restaurant.

In Santa Caterina there is one Italian restaurant but there are less restaurant and it is closer to the center.

Problems found

The data obtained using the Foursquare API has not the venues with the required accuracy.

The data obtained using the Foursquare API is not updated.

Some restaurants has not categories assigned, so we lost some information too.

Section 6 | Conclusion

Conclusion

We used differents tool with the aim of finding an optimal location for a new Italian restaurants in Barcelona.

We used different data sources to obtain the requiered information

The Foursquare API, in this case this in not a good source.

Then, although we found two good locations, this two locations might not be the good ones actually.

Instead of using Foursquare API, we have the Barcelona Restaurant Open Data updated weekely: opendatabcn restaurants.