

Capstone Project - The Battle of Neighborhoods

July 31, 2021

Introduction

In this project we will try to find a location for a new Italian restaurant in Barcelona, Spain.

Barcelona is a city on the coast of northeastern Spain. It is the capital and largest city of the autonomous community of Cataluña, as well as the second most populous municipality of Spain, also it is the fifth most populous urban area in the European Union after Paris, the Ruhr area, Madrid, and Milan and It is one of the largest metropolises on the Mediterranean Sea, located on the coast between the mouths of the rivers Llobregat and Besòs, and bounded to the west by the Serra de Collserola mountain range.

Barcelona is administratively divided into ten districts. The territorial division of the districts responds to historical questions of the city. Most of the districts correspond to old independent municipalities that were annexed to the city during the 19th and 20th centuries, and which still retain their own personality.

The ten districts are: Ciutat Vella, El Ensanche, Sants-Montjuïc, Les Corts, Sarrià-San Gervasio, Gracia, Horta-Guinardó, Nou Barris, San Andrés, San Martín.

We will use the information available from the districts, that is, the neighborhoods in each district and the types of restaurants, specially Italian restaurants, in each.

So, this project is targeted to people that are interested in opening an Italian restaurant in Barcelona.

In order to find the best location to open an Italian restaurant we will follow the following criteria to determine if a location is a good one:

- The location is not crowded with restaurants.
- The location has not Italian restaurants near.
- If it is possible, we want a location close to the city center.

Data

Following the criteria from the Introduction section, we will need the information about the neighborhoods in Barcelona, the restaurants in each neighborhood and also locations where you can eat Italian food, not only Italian restaurants.

First we obtained the information about neighborhoods in Barcelona Open Data(2019): [2019-Superficie](#), where you can find Barcelona's last published area of the neighborhoods of the city of Barcelona.

If we take a look to the information inside we obtain:

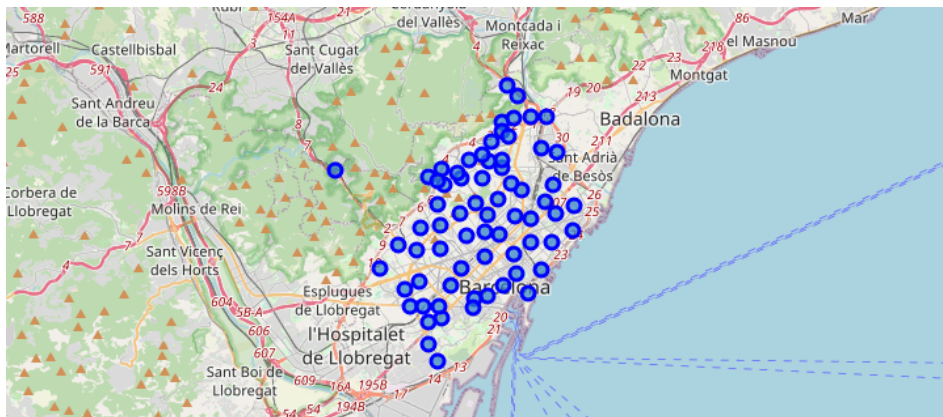
	Codi_Districte	Nom_Districte	Codi_Barri	Nom_Barri
0	1	Ciutat Vella	1	el Raval
1	1	Ciutat Vella	2	el Barri Gòtic
2	1	Ciutat Vella	3	la Barceloneta
3	1	Ciutat Vella	4	Sant Pere, Santa Caterina i la Ribera
4	2	Eixample	5	el Fort Pienc
5	2	Eixample	6	la Sagrada Família
6	2	Eixample	7	la Dreta de l'Eixample
7	2	Eixample	8	l'Antiga Esquerra de l'Eixample
8	2	Eixample	9	la Nova Esquerra de l'Eixample
9	2	Eixample	10	Sant Antoni

That is the first 10 rows of the table, the size is 73x4

The coordinates of each neighborhood is not in the data so we will use Nominatim to Geocode it. So, now after cleaning it, we have the following data set:

	Nom_Districte	Nom_Barri	Latitude	Longitude
0	Ciutat Vella	el Raval	41.379518	2.168368
1	Ciutat Vella	el Barri Gòtic	41.383395	2.176912
2	Ciutat Vella	la Barceloneta	41.380653	2.189927
3	Ciutat Vella	Sant Pere, Santa Caterina i la Ribera	41.388076	2.183398
4	Eixample	el Fort Pienc	41.395925	2.182325
...
68	Sant Martí	Diagonal Mar i el Front Marítim del Poblenou	41.405228	2.213352
69	Sant Martí	el Besòs i el Maresme	41.414979	2.213877
70	Sant Martí	Provençals del Poblenou	41.411948	2.204125
71	Sant Martí	Sant Martí de Provençals	41.416519	2.198968
72	Sant Martí	la Verneda i la Pau	41.423220	2.202940

Now it is time to visualize the neighborhoods around the center of Barcelona.



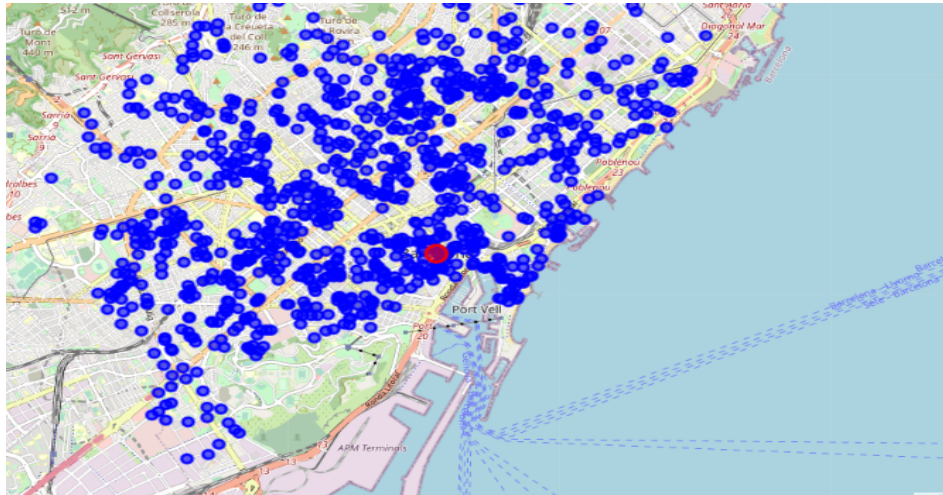
The next step is find the restaurants in each neighborhood. We will obtain the number of restaurants and their type and location in every neighborhood using Foursquare API.

The first search we have carried out using the query word as Restaurant, after that we did another search with the query word as Italian. After that in the last data set we extract all the places that you can eat Italian food.

Because some neighborhoods are very close some locations are repeted so we eliminate the duplicated places. The data set obtenided does not contain the neighborhood of each place, so we used reverse Geocode in order to find the neighborhoods, and finally, we get the data set to work with it.

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formatted
0	ristoante italian fusion	Italian Restaurant	NaN	41.410623	2.213432	[[{"label": "display", "lat": 41.410623, "lng": ...	486	NaN	ES	NaN	NaN	España	[España]
1	Italian District	Italian Restaurant	169 Carrer de Lull	41.400361	2.201006	[[{"label": "display", "lat": 41.400361, "lng": ...	63	08005	ES	Barcelona	Cataluña	España	[169 Carre Lull, 08005 Barcelona Cataluña...
2	Can Italianini	Italian Restaurant	NaN	41.432226	2.177170	[[{"label": "display", "lat": 41.432226, "lng": ...	143	NaN	ES	Barcelona	Cataluña	España	[Barcelona Cataluña, E
4	Ristorante Italiano Da Paolo	Italian Restaurant	Avinguda Madrid, 63	41.378160	2.126805	[[{"label": "display", "lat": 41.378160, "lng": ...	449	08028	ES	Barcelona	Cataluña	España	[Avinguda M 63, 08028 Barcelona Cataluña...
5	L'Angolo Italiano (L'illa Diagonal)	Italian Restaurant	Avinguda Diagonal, 557	41.389323	2.134953	[[{"label": "display", "lat": 41.389323, "lng": ...	486	08029	ES	Barcelona	Cataluña	España	[Avinguda Diagonal, 5 08029 Barc Catal...
7	Bar Cafe La Italiana	Italian Restaurant	Calabria, 181	41.383262	2.150368	[[{"label": "display", "lat": 41.383262, "lng": ...	115	08015	ES	Barcelona	Cataluña	España	[Calabria, 1 08015 Barc Cataluña, E
8	Nonna Sara Ristorante Italiano	Italian Restaurant	225 Carrer de Còrsega	41.391865	2.153005	[[{"label": "display", "lat": 41.391865, "lng": ...	266	08036	ES	Barcelona	Cataluña	España	[225 Carre Còrsega, 0 Barcelona
9	Trattoria Italiana	Italian Restaurant	Calle Roger de Lluria 87	41.395418	2.166061	[[{"label": "display", "lat": 41.395418, "lng": ...	104	NaN	ES	Barcelona	Cataluña	España	[Calle Roge Lluria 87 (C Barcelona
10	Rossini	Italian Restaurant	Plaça Reial, 13	41.379912	2.175864	[[{"label": "display", "lat": 41.379912, "lng": ...	397	08002	ES	Barcelona	Cataluña	España	[Plaça Reial 08002 Barc Cataluña, E

Also we present a view from the locations.



Metodology

The aim of this project is to find the best neighborhood of Barcelona to open a new Italian restaurant.

The total number of neighborhoods in Barcelona are 73 so we need to find a way to cluster them based on their similarities, that are the number and the kind of restaurant.

In first place we check the unique categories in our data set and how many times each category it is repeated.

	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

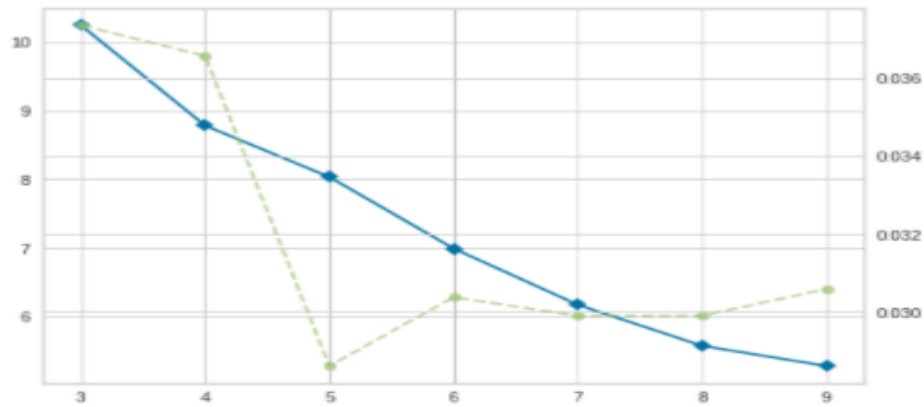
Now we have an idea about what type of restaurants are more common in Barcelona. It is look like there is not too many Italian restaurant so it might be a good decision to open a new one.

We also analyze the frequency of each type of restaurant in all the neighborhoods.

	Neighborhood	African Restaurant	American Restaurant	Arcade	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	BBQ Joint	Bar	Beer Garden	...	Sushi Restaurant	Swiss Restaurant
0	Baró de Viver	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
1	Can Baró	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
2	Canyelles	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
3	Diagonal Mar i el Front Marítim del Poblenou	0.0	0.0	0.0	0.0	0.0	0.038462	0.038462	0.076923	0.038462	...	0.000000	0.0
4	Farró	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
...
65	la Vila de Gràcia	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.090909	0.000000	...	0.000000	0.0
66	les Corts	0.0	0.0	0.0	0.0	0.0	0.045455	0.000000	0.045455	0.000000	...	0.022727	0.0
67	les Planes	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
68	les Roquetes	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0
69	les Tres Torres	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0

In order to clustering the neighborhoods we will use the approach of K-Means algorithm to extract the clusters, produce a map and make an argument on the final result. The first step in this algorithm it is determinate the best " k ", so we will use the analytical approach of the elbow method.

The results obtained with this method is that k is equal to 5.

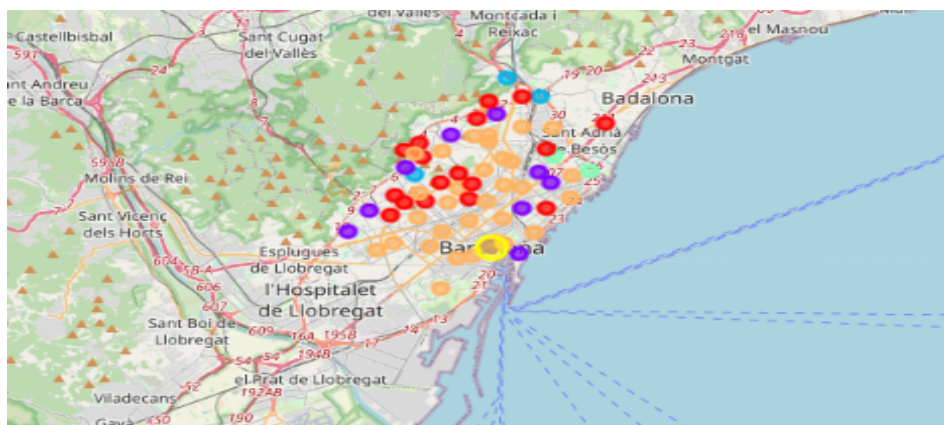


Now we can use the K-Means algorithm to cluster the neighborhood based on the restaurants categories. Once we have the clusters we will analyze with one follow in the optimal way the criteria in the Introduction section. We will do that following the next steps:

- After clustering the neighborhood, we will chose the two clusters with more locations close to the city center.
- Then, in each cluster we will calculate the distance from the neiighborhood to the city center.
- In the next step, we will check the unique categories in our data set and how many times each category it is repeated.
- Finally, 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.

Results

Once we have the five clusters, let's take a look on a folium map:



After clustering the neighborhoods we add to each cluster all the restaurants in all the neighborhoods inside each cluster.

We obtain:

Cluster 1

First we take a look in the first rows in this cluster:

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	form
361	Bar Restaurant Casa Jin	Restaurant	Av. Mare de Deu de Montserrat, 29	41.414135	2.162264	[[{"label": "display", "lat": 41.414135, "lng": ...	218	08024	ES	Barcelona	Cataluña	España	[Av de 080
365	Restaurante Jardí de L'Apat	None	Albert Llamas, 2	41.415155	2.159719	[[{"label": "display", "lat": 41.41515487965063...	247	NaN	ES	Barcelona	Cataluña	España	[All Bai Ca
481	Restaurant Azul	Modern European Restaurant	Padua 108	41.403355	2.143384	[[{"label": "display", "lat": 41.40335464477539...	441	08006	ES	Barcelona	Cataluña	España	[Pa Bai Ca
486	Restaurante Tuttopasta	Italian Restaurant	Santjoanistes, 28	41.402030	2.146542	[[{"label": "display", "lat": 41.40203, "lng": ...	643	08006	ES	Barcelona	Cataluña	España	[Se 080 Ca
487	Restaurante Xewi	Restaurant	C. Sant Elies	41.401724	2.146945	[[{"label": "display", "lat": 41.40172359969239...	688	NaN	ES	Barcelona	Cataluña	España	[C. Bai Ca

Now we want to see the 10 most common categories in this neighborhood. We can see that in this cluster there are 6 Italian food places:

	categories	Count
24	Restaurant	46
28	Spanish Restaurant	15
17	Mediterranean Restaurant	11
3	Breakfast Spot	4
16	Japanese Restaurant	3
23	Pizza Place	3
29	Tapas Restaurant	3
6	Chinese Restaurant	3
8	Coffee Shop	3
15	Italian Restaurant	3

Cluster 2

First we take a look in the first rows in this cluster:

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	form
301	Restaurante 5 Hermanos	Mediterranean Restaurant	Federico García Lorca, 31	41.444106	2.165911	[[{"label": "display", "lat": 41.44410555164136...	565	08042	ES	Barcelona	Cataluña	España	[Fede Lorca Barcel
0	ristoante Italian fusion	Italian Restaurant	NaN	41.410623	2.213432	[[{"label": "display", "lat": 41.410623, "lng": ...	486	NaN	ES	NaN	NaN	España	[Espa
33	Restaurant Siberia	Spanish Restaurant	Selva de Mar 84	41.411146	2.206173	[[{"label": "display", "lat": 41.41114643740437...	192	08019	ES	Barcelona	Cataluña	España	[Selv (Crist Mour
46	Restaurante Dia	Spanish Restaurant	NaN	41.410211	2.218577	[[{"label": "display", "lat": 41.4102114086185, ...	660	NaN	ES	NaN	NaN	España	[Espa
49	Bar Restaurant Angel	Spanish Restaurant	paseo taulat	41.405051	2.213004	[[{"label": "display", "lat": 41.40505146431758...	35	NaN	ES	Barcelona	Cataluña	España	[paseo Barcel Catal

Now we want to see the 10 most common categories in this neighborhood. We can see that in this cluster there are 7 Italian restaurants:

	categories	Count
29	Mediterranean Restaurant	56
39	Spanish Restaurant	42
37	Restaurant	41
41	Tapas Restaurant	8
25	Italian Restaurant	7
1	Asian Restaurant	7
34	Paella Restaurant	7
3	Bar	6
38	Seafood Restaurant	6
26	Japanese Restaurant	4

Cluster 3

First we take a look in the first rows in this cluster:

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	format
275	Restaurante "Gavisé"	Spanish Restaurant	Carrer de Clariania, 33-35	41.447157	2.200350	[{"label": "display", "lat": 41.44715666770935...	317	08030	ES	Barcelona	Cataluña	España	[Carrer Clarian 08030 Ca...
284	Restaurante El Cordero	Spanish Restaurant	Serra de Collserola Nal Park, Carrer del Riude...	41.452670	2.180062	[{"label": "display", "lat": 41.45267027899003...	435	08042	ES	Barcelona	Cataluña	España	[Serra de Collserola Nal Park, Carrer del Riude...
458	restaurante Prince	Spanish Restaurant	NaN	41.412193	2.148900	[{"label": "display", "lat": 41.412193, "lng": ...	424	NaN	ES	NaN	NaN	España	[Españ

This cluster has only 3 restaurants and 2 neighborhoods but it is too far from the city center.

Cluster 4

First we take a look in the first rows in this cluster:

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	forma
280	Restaurant La Font	Food	NaN	41.465404	2.179867	[{"label": "display", "lat": 41.46540356609677...	394	NaN	ES	NaN	NaN	España	[Españ
42	Royal Tandoori Restaurant and cafe	Asian Restaurant	AV.Alfonso XIII 554-56,	41.415585	2.220912	[{"label": "display", "lat": 41.41558522322383...	591	08913	ES	Barcelona	Cataluña	España	[AV.Alf 554-56 (Badal Ba...
21	Restaurante Casa Suso	Food	Carrer de ca n'Oliva, 69	41.424102	2.204519	[{"label": "display", "lat": 41.42410237086280...	164	08020	ES	Barcelona	Cataluña	España	[Carre n'Oliva Barcel
22	Restaurante Chino Ta viang dao	Asian Restaurant	NaN	41.421795	2.205835	[{"label": "display", "lat": 41.42179497479602...	289	NaN	ES	NaN	NaN	España	[Españ

This cluster has only 4 restaurants and 3 neighborhoods but also, it is too far from the city center.

Cluster 5

First we take a look in the first rows in this cluster:

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	form
492	restaurant santalo 113	Restaurant	Santalo 113	41.396570	2.144143	[{"label": "display", "lat": 41.39656965, "lng": ...}	151	NaN	ES	Barcelona	Cataluña	España	[San Barc Cata
494	Restaurant Meating	Restaurant	NaN	41.396829	2.142566	[{"label": "display", "lat": 41.396829, "lng": ...}	128	NaN	ES	NaN	NaN	España	[Esp
495	Restaurant Japonès J8	Japanese Restaurant	Marià Cubí, 179	41.395301	2.145189	[{"label": "display", "lat": 41.39530104454578, ...}	316	08021	ES	Barcelona	Cataluña	España	[Mar (Ami Barc
496	Restaurante Café Madrazo	Mediterranean Restaurant	Madrazo, 50	41.399488	2.147856	[{"label": "display", "lat": 41.39948840282839, ...}	417	08006	ES	Barcelona	Cataluña	España	[Mac 0800 Cata
497	Restaurant Roma	Mediterranean Restaurant	Alfons XII, 41	41.399178	2.146801	[{"label": "display", "lat": 41.39917762103506, ...}	477	08006	ES	Barcelona	Cataluña	España	[Alfo 0800 Cata

Now we want to see the 10 most common categories in this neighborhood. We can see that in this cluster there are 22 Italian food places:

	categories	Count
63	Restaurant	144
69	Spanish Restaurant	93
50	Mediterranean Restaurant	80
20	Chinese Restaurant	51
7	Bar	34
5	Asian Restaurant	32
75	Tapas Restaurant	28
44	Japanese Restaurant	28
43	Italian Restaurant	22
16	Café	18

Determination of good locations

Now if we take a look in to the folium map we start selecting the Cluster 2 and the Cluster 5 in order to continue with the methodology explained in the Methodology section.

For the Cluster 2 the distances are:


```

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 closest neighborhood is el Gòtic:

	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 this neighborhood are 3 Italian restaurants.

So, we will analyze the second neighborhood closer to the center, Santa Caterina:

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

There is 1 Italian restaurant in this neighborhood and it is 345.168 from the center.

For the Cluster 5 the distances are:

```

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
57     2299.342129
58     4328.932901
62     5508.208091
64     1798.463504
65     2818.245370
66     3735.880074

```

The closest neighborhood is la Ribera and the restaurants there are:

	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 this neighborhood and it is 503.878 meters from to the center.

Discussion

So we have two good locations, la Ribera has more restaurants but there 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, so now it depends on the preferences on the criteria.

Finally we have to say that all the data obtained using the Foursquare API is not very useful in Barcelona because it has not the venues with the required accuracy and it is not updated.

Also, some restaurants has not categories assigned, so we loose some information too. Most of them are categorized as Restaurant, so doesn't know what kind of restaurant is.

Conclusion

In this work we used differents tool with the aim of finding an optimal location for a new Italian restaurants in Barcelona, Spain. This is a well known city because it is one of the most visited cities in the world by international visitors. With this in mind we used different data sources to obtain the requiered information. One of them it is the Foursquare API, in this case this in not a good source because it is not updated and it is not very accurate. So, although we found two good locations, this two locations might not be the good ones actually.

If i would like to open a new Italian restaurant i would not use that data source, instead we have the Barcelona Restaurant Open Data updated weekly: [opendatabcn restaurants](#).