

Finding the best location to open an Italian restaurant in Barcelona city

Rafael Diaz / June 28, 2020

1. Introduction

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an Italian restaurant in Barcelona, Spain.

Since there are lots of restaurants in Barcelona we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Polish restaurants in vicinity. We would also prefer locations as close to city center as possible, assuming that first two conditions are met.

We will use our data science powers to generate a few most promising neighborhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Obviously, entrepreneurs' teams would be very interested in accurate prediction of player improvement, for competitive advantage and business values. Others who are interested in NBA such as fans and fantasy basketball players may also be interested.

2. Data acquisition and cleaning

2.1 Data sources

Based on definition of our problem, factors that will influence our decision are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to another popular place in the neighborhood, if any
- distance of neighborhood from city center

I decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract and generate the required information:

- District centres of candidate areas will be obtained using Barcelona Open Data source from Local authorities.
- Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
- Coordinate of Barcelona center will be obtained using Barcelona Open Data source of city centre location (Plaza Catalunya).
- Barcelona's main Districts coordinates will be obtained via GoogleMaps and Foursquare.

2.2 Data cleaning and selection

Geolocation data have been fetched out of Foursquare and GoogleMaps for the potential neighborhoods where to place the restaurant.

First, we have fetched the coordinates of Plaza Catalunya which is the touristic center of Barcelona city. Then we have created hexagonal grid of cells equally distant from each other as the zones to explore.

Applying this calculation on python, there are 364 candidate neighborhood centers generated. Latitude and longitude coordinates can be easily obtained out of GoogleMaps together with the corresponding X and Y coordinates over the UTM map zone. Those coordinates have been used to explore the different candidate centers.

Let's visualize the data we have so far: city center location and candidate neighborhood centers using Folium

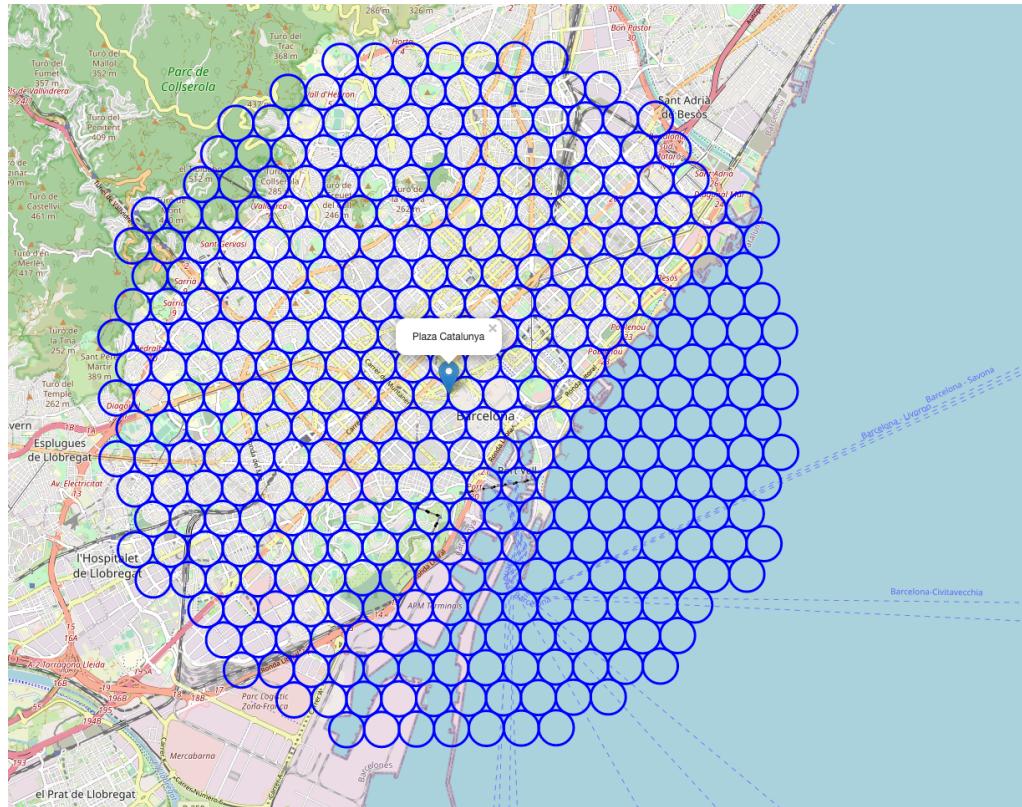


Figure 1. Map of Barcelona with projected candidate neighborhood centers from Plaza Catalunya

We then extracted the coordinates of centers of districts and areas to be evaluated from Plaza Catalunya as per our locations grid using GoogleMaps API.

Once I had the locations, I used Foursquare API to get info on restaurants in each neighbourhood centres. We're interested in venues in 'food' category, but only those that are proper restaurants - coffee shops, pizza places, bakeries etc. are not direct competitors so we don't care about those.

I have included in the list only venues that have 'restaurant' in category name, and we'll make sure to detect and include all the subcategories of specific 'Italian restaurant' category, as we need info on Italian restaurants in the neighbourhood.

- Total number of restaurants: **2,573**
- Total number of Italian restaurants: **163**
- Percentage of Italian restaurants: **6.34%**
- Average number of restaurants in neighborhood: **13.84722222222221**

So now we have all the restaurants in area within few kilometres from Barcelona centre, and we know which ones are Italian restaurants! We also know which restaurants exactly are in vicinity of every neighbourhood candidate centre.

This concludes the data gathering phase - we're now ready to use this data for analysis to produce the report on optimal locations for a new Italian restaurant!

3. Methodology and Data Analysis

3.2 Methodology

In this project we will direct our efforts on detecting areas of Barcelona that have low restaurant density, particularly those with low number of Italian restaurants. We will limit our analysis to area approximately 6km around the city centre.

In first step we have collected the required **data: location and type (category) of every restaurant within 6km from Barcelona centre** (Plaza Catalunya). We have also **identified Italian restaurants** (according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of '**restaurant density**' across different areas of Barcelona - we will use **heatmaps** to identify a few promising areas close to centre with low number of restaurants in general (*and* no Italian restaurants in vicinity) and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create **clusters of locations that meet some basic requirements** established in discussion with stakeholders: we will take into consideration locations with **no more than two restaurants in radius of 250 meters**, and we want locations **without Italian restaurants in radius of 400 meters**. We will present map of all such locations but also create clusters (using **k-means clustering**) of those locations to identify general zones / neighbourhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

3.2 Data Analysis

Let's perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the **number of restaurants in every area candidate**:

calculate the **distance to nearest Italian restaurant from every area candidate centre** (not only those within 300m - we want distance to closest one, regardless of how distant it is).

Average distance to closest Italian restaurant from each area centre: 1465.343661241832

OK, so **on average Italian restaurant can be found within ~1,500m** from every area centre candidate. That's not fairly close, so we need to filter our areas further!

Let's create a map showing **heatmap / density of restaurants** and try to extract some meaningful info from that. Also, let's show **borders of Barcelona boroughs** on our map and a few circles indicating distance of 1km, 2km and 3km from Plaza Catalunya.

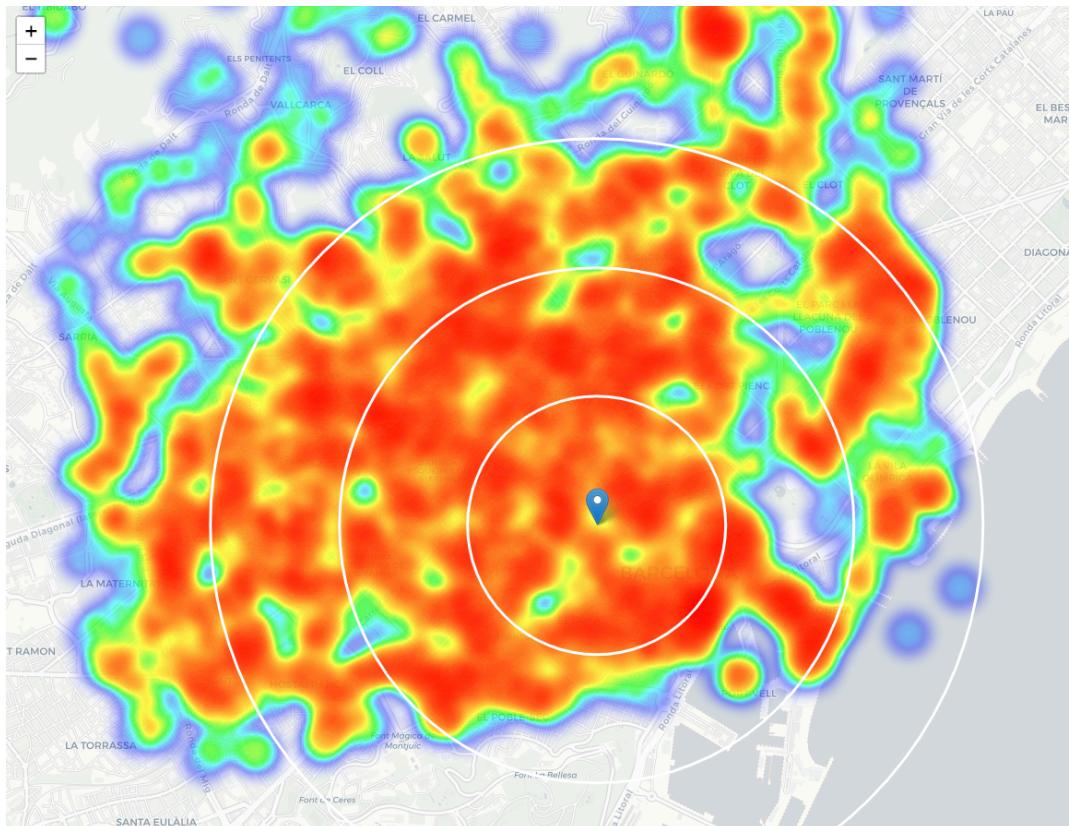


Figure 2. Heat-Map of restaurants in Barcelona from Plaza Catalunya

Looks like a few pockets of low restaurant density closest to city centre can be found **north-east and south-west from Plaza Catalunya**.

Let's create another heatmap map showing **heatmap/density of Italian restaurants only**.

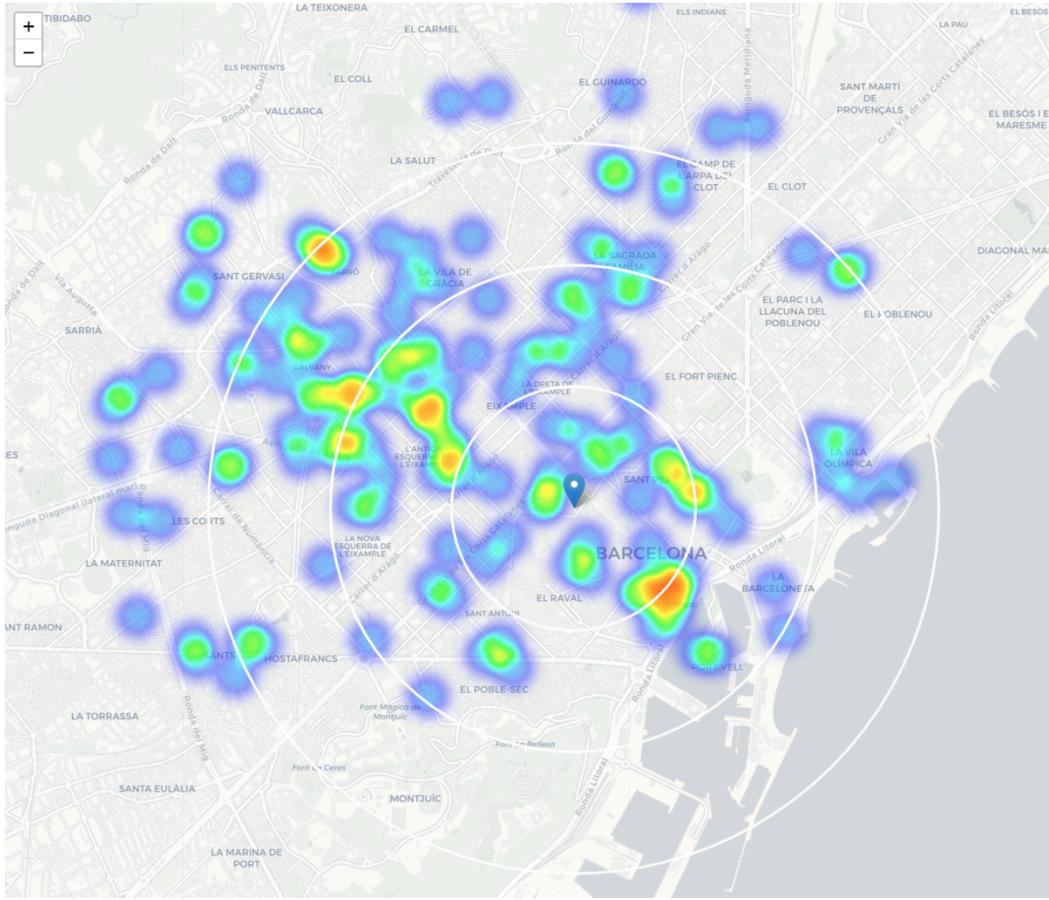


Figure 3. Heat-Map of Italian restaurants in Barcelona from Plaza Catalunya

This map is not so 'hot' (Italian restaurants represent a subset of ~13.5% of all restaurants in Barcelona) but it also indicates higher density of existing Italian restaurants directly north-west from Plaza Catalunya, with closest pockets of **low Italian restaurant density positioned north-east and south-west from city centre**.

Based on this we will now focus our analysis on areas *north-east and south-west from the city centre* - we will move the centre of our area of interest and reduce its size to have a radius of **3km**. This places our location candidates mostly in boroughs **Hostafrancs and Poblenou** which we will check whether they have popular places where locals and tourist come by often.

I have explored these locations in terms of popular places using Foursquare as shown in the following map (Poblenou in yellow dots and Hostafrancs in orange dots) in contrast to the restaurants in the area (blue circles) and Italian restaurants (red circles).

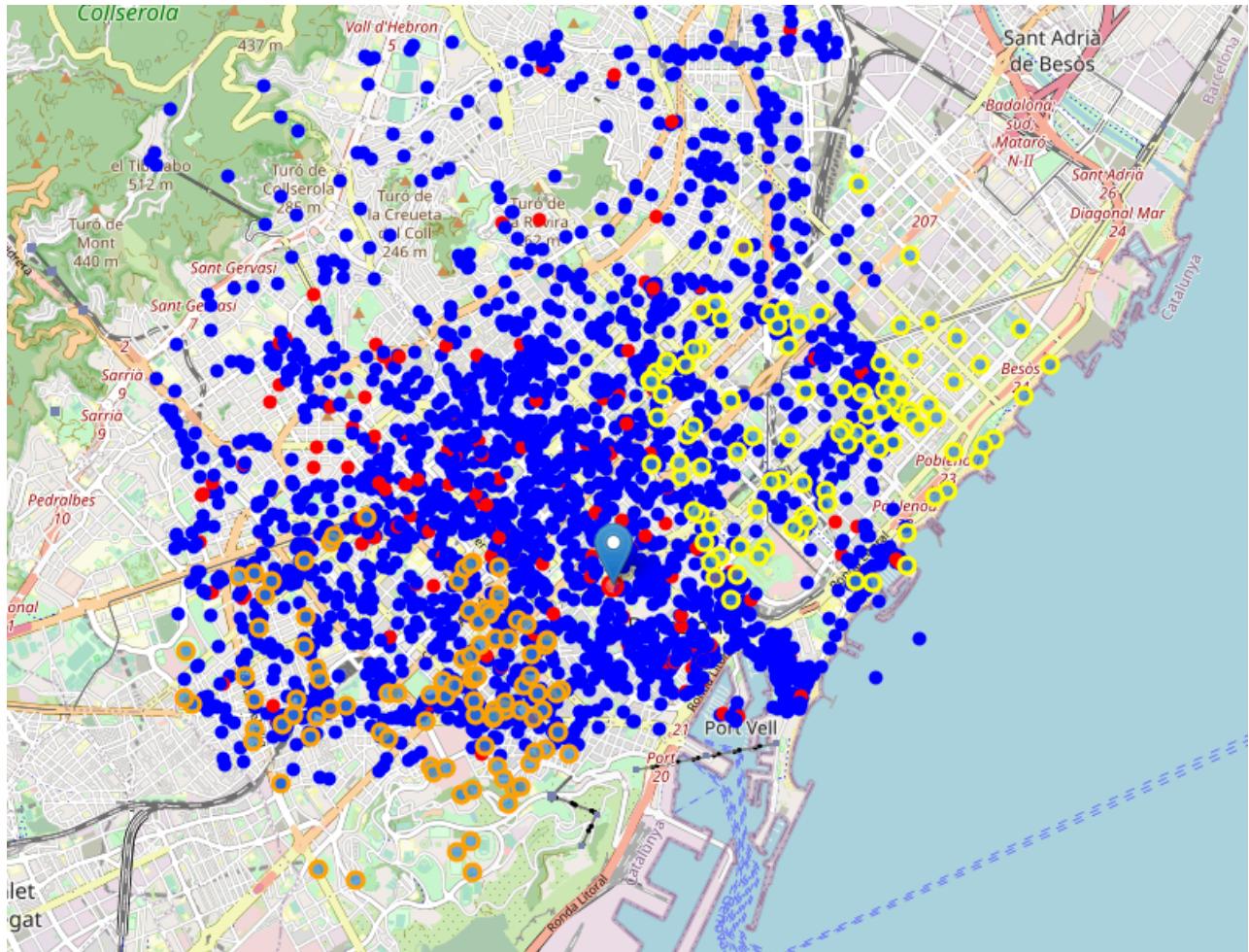


Figure 4. Restaurant, Italian restaurants and popular places in Barcelona from Plaza Catalunya

Let's define new, narrower regions of interest, which will include low-restaurant-count parts of Hostafrancs and Poblenou closest to Plaza Catalunya.

Now let's calculate two most important things for each location candidate: **number of restaurants in vicinity** (we'll use radius of **250 meters**) and **distance to closest Italian restaurant** to Hostafrancs and Poblenou.

We might just want to explore only **Hostafrancs** given the higher number of popular places in the area also considering the number of **Italian restaurants** nearby.

	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Italian restaurant
0	41.375088	2.142933	428329.220632	4.580751e+06	12	357.265635
1	41.400527	2.201729	433271.981651	4.583528e+06	4	527.250781

We need to **filter** those locations we're interested only: **locations with no more than two restaurants in radius of 250 meters**, and **no Italian restaurants in radius of 250 meters**.

=====
Addresses of centres of areas recommended for further analysis
=====

Carrer de la Creu Coberta, 99, 08014 Barcelona => 2.6km from Plaza Catalunya
Carrer de Llull, 186, 08005 Barcelona => 3.1km from Plaza Catalunya

This concludes our analysis.

We have created 2 addresses representing locations with low number of restaurants and no Italian restaurants nearby, all zones being fairly close to city centre (both around 3km from Plaza Catalunya). These locations should be considered only as a starting point for exploring area neighbourhoods in search for potential restaurant locations. The zones located in Hostafrancs and Poblenou boroughs, which we have identified as interesting due to being popular with tourists, are fairly close to city centre and well connected by public transport.

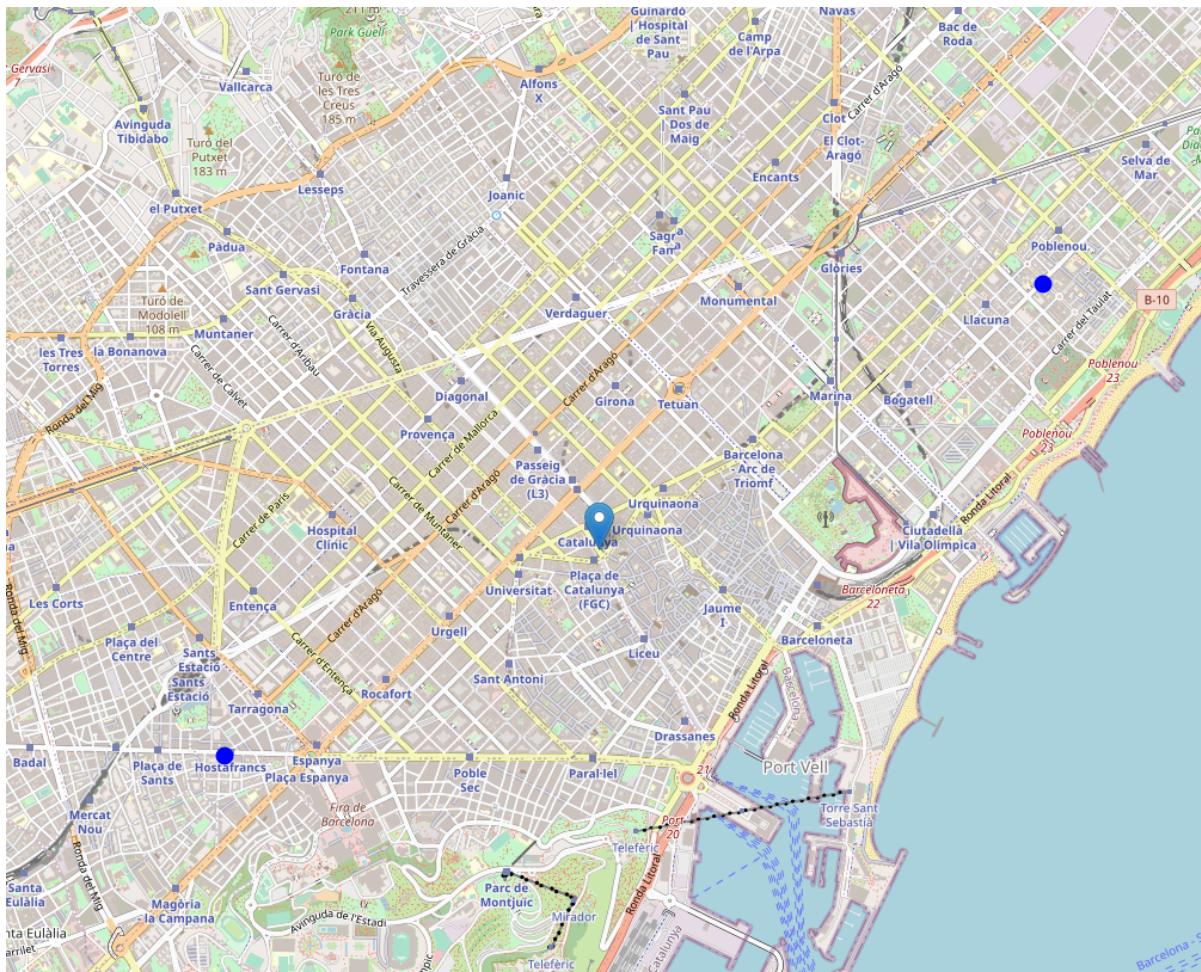


Figure 5. The recommended places for establishing an Italian restaurant in Hostafrancs and Poblenou

4. Results and Discussion

Our analysis shows that although there is a great number of restaurants in Barcelona (2500+ in our initial area of interest which was 6x6km around Plaza Catalunya), there are pockets of low restaurant density fairly close to city centre. Highest concentration of restaurants was detected north-west and south from Plaza Catalunya, so we focused our attention to areas south-west and north-east, corresponding to boroughs Hostafrancs and Poblenou, which offer a combination of popularity among tourists, closeness to city centre, strong socio-economic dynamics *and* a number of pockets of low restaurant density.

After directing our attention to these narrower areas of interest, we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Italian restaurant closer than 250m were removed.

Those location candidates were zones of interest which contain a number of location candidates. Addresses of centres of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 2 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Italian restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Barcelona centre but not crowded with existing restaurants (particularly Italian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors considered and all other relevant conditions met.

5. Conclusion

Purpose of this project was to identify Barcelona areas close to centre with low number of restaurants (particularly Italian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Italian restaurant. By calculating restaurant density distribution from Foursquare data, we have first identified general boroughs that justify further analysis (Hostafrancs and Poblenou), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centres were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise and proximity to major roads, real estate availability, prices, social and economic dynamics of every neighbourhood etc.