Capstone Project - NICE RESTAURANT

Applied Data Science Capstone by IBM/Coursera

Table of contents

1. Introduction: Business Problem

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 **Nice**, France.

Since there are lots of restaurants in Nice we will try to detect **locations that are not already crowded with restaurants**. We are also particularly interested in **areas with no Italian 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.

2. Data

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 Italian restaurants in the neighborhood, if any
- distance of neighborhood from city center

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

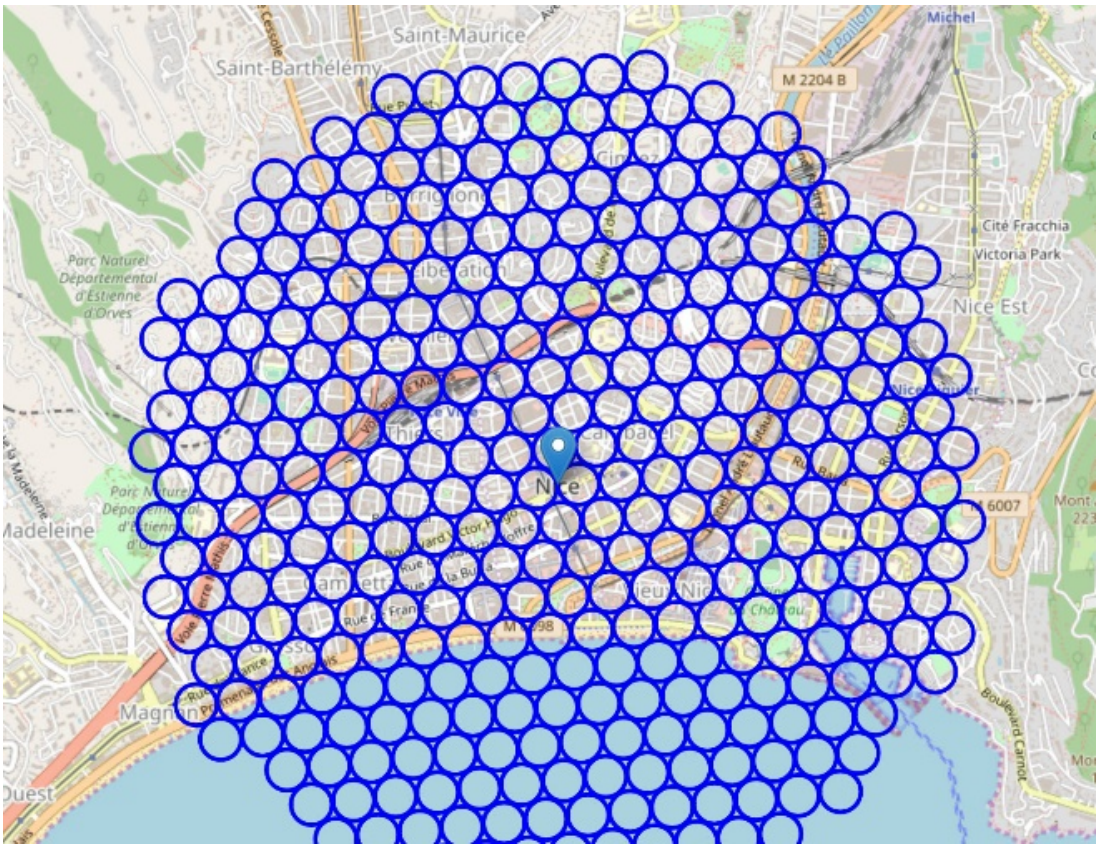
Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **OpenStreetMap reverse geocoding**
- number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
- coordinate of Nice center will be obtained using **GeoPy API geocoding** of well-known Nice location (Nice Etoile Jean Medecin)

2.1. Area Candidates

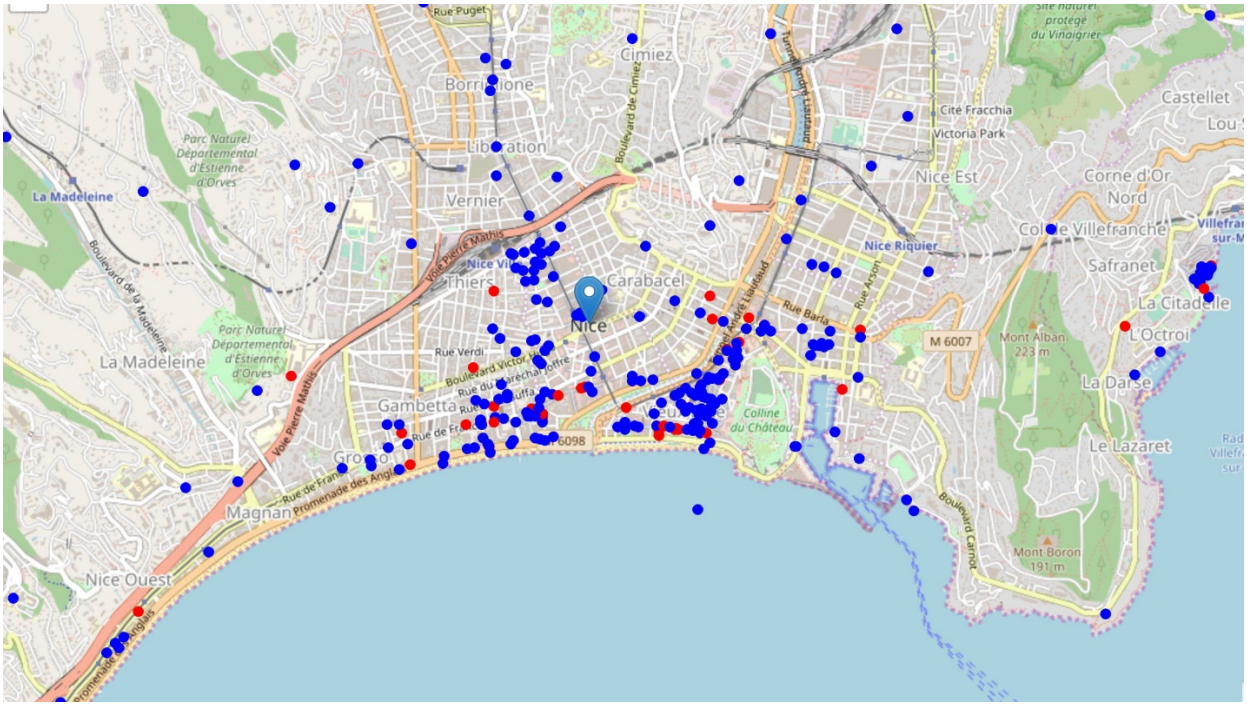
We created a grid of area candidates, equally spaced, centered around city center and within ~3km from Nice Center. Our neighborhoods will be defined as circular areas with a radius of 150 meters, so our neighborhood centers will be 300 meters apart.

To accurately calculate distances, we need to create our grid of locations in Cartesian 2D coordinate system which allows us to calculate distances in meters (not in latitude/longitude degrees). Then we'll project those coordinates back to latitude/longitude degrees to be shown on Folium map.



2.2. Area Restaurants

The Total number of restaurants is 355 (3 km around Nice Center).
Total number of Italian restaurants is 45 (3 km around Nice Center).
Percentage of Italian restaurants is 12.68%
Average number of restaurants in neighborhood: 0.87 (400 m around).
This is the distribution of restaurants. Italian restaurants are in red and others are in blue.



3. Methodology

In this project we will direct our efforts on detecting areas of Nice that have low restaurant density, particularly those with low number of Italian restaurants. We will limit our analysis to area ~3km around city center.

In first step we have collected the required **data: location and type (category) of every restaurant within 3km from Nice center**. 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 Nice - we will use **heatmaps** to identify a few promising areas close to center 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 150 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 / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

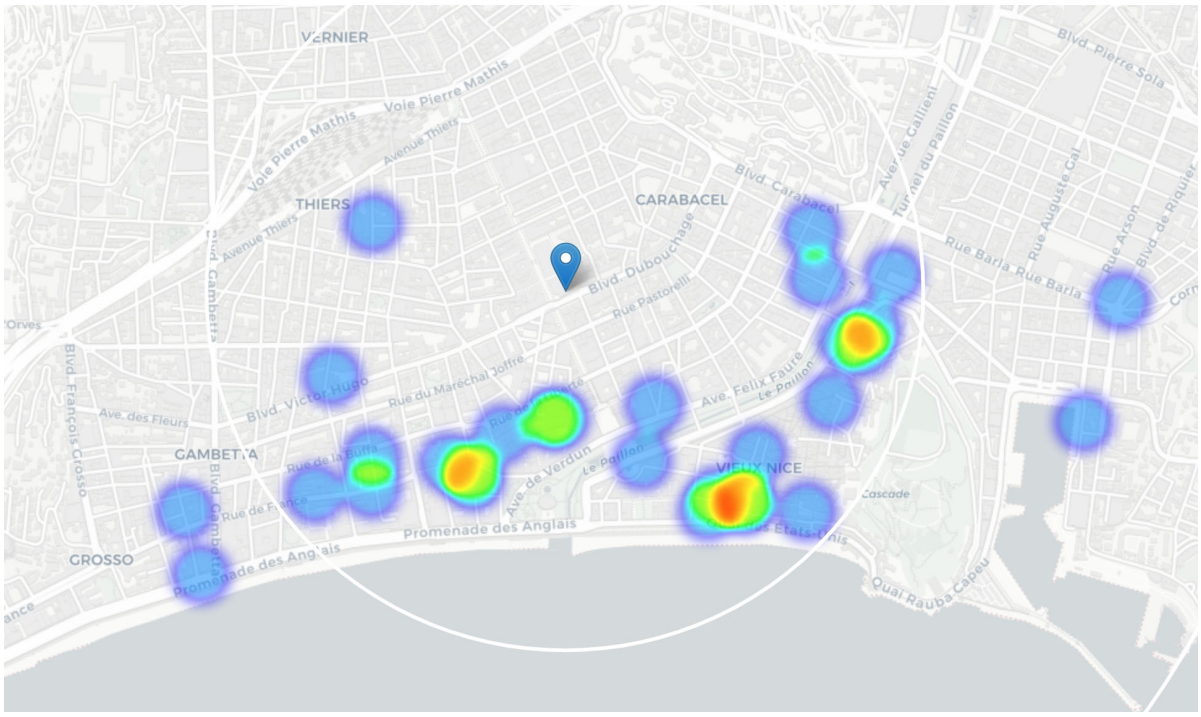
4. Analysis

The Average distance to closest Italian restaurant from each area center: 578 meters.

We create a map showing heatmap / density of restaurants and try to extract some meaningful info from that.

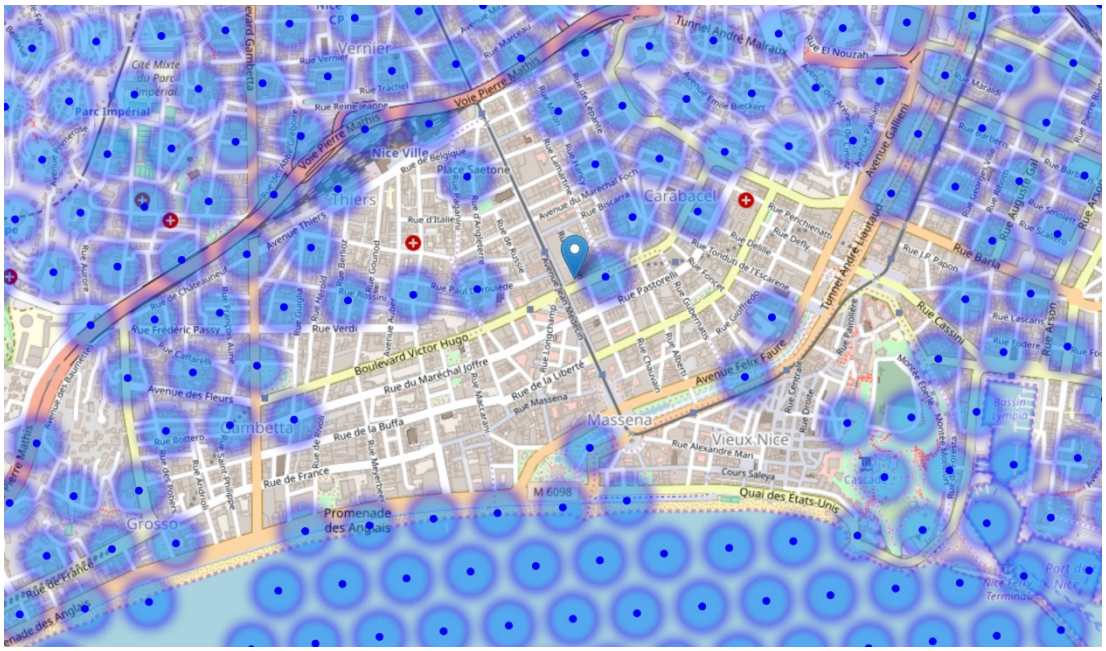
We notice that restaurants are frequent in Promenade des Anglais, Vieux Nice and Carré d'OR.

We create another heatmap map showing **heatmap/density of Italian restaurants** only.

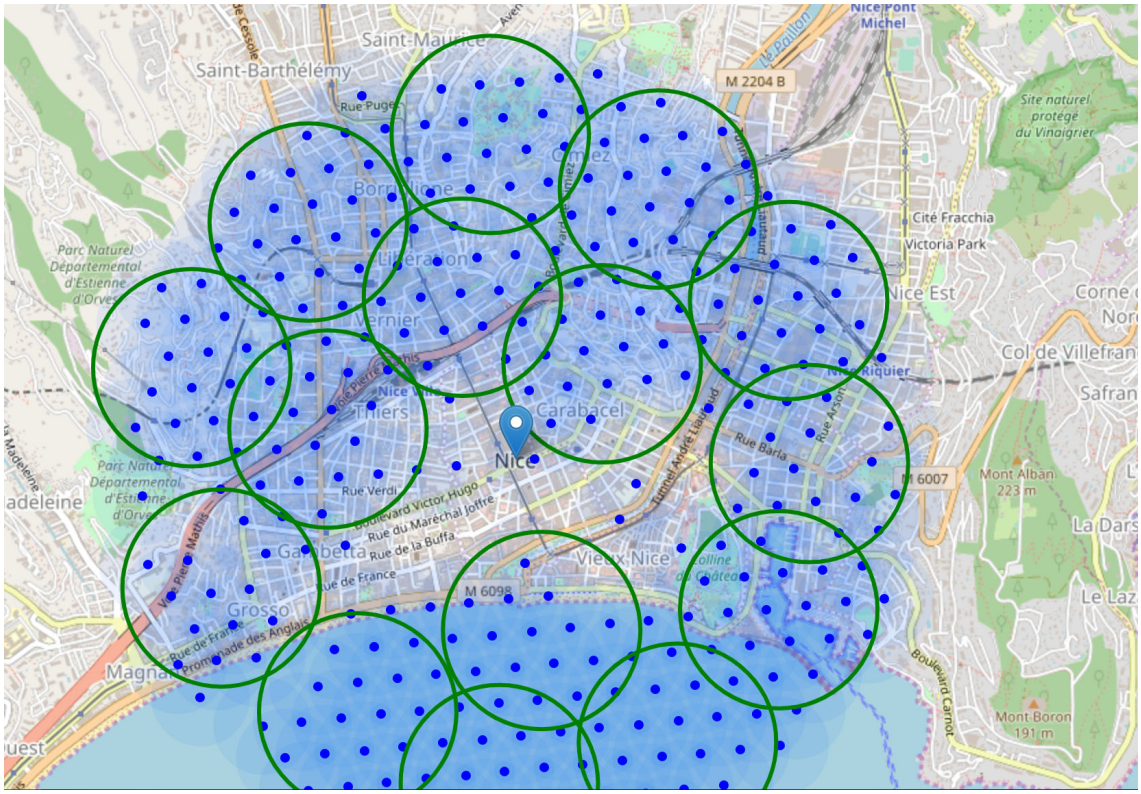


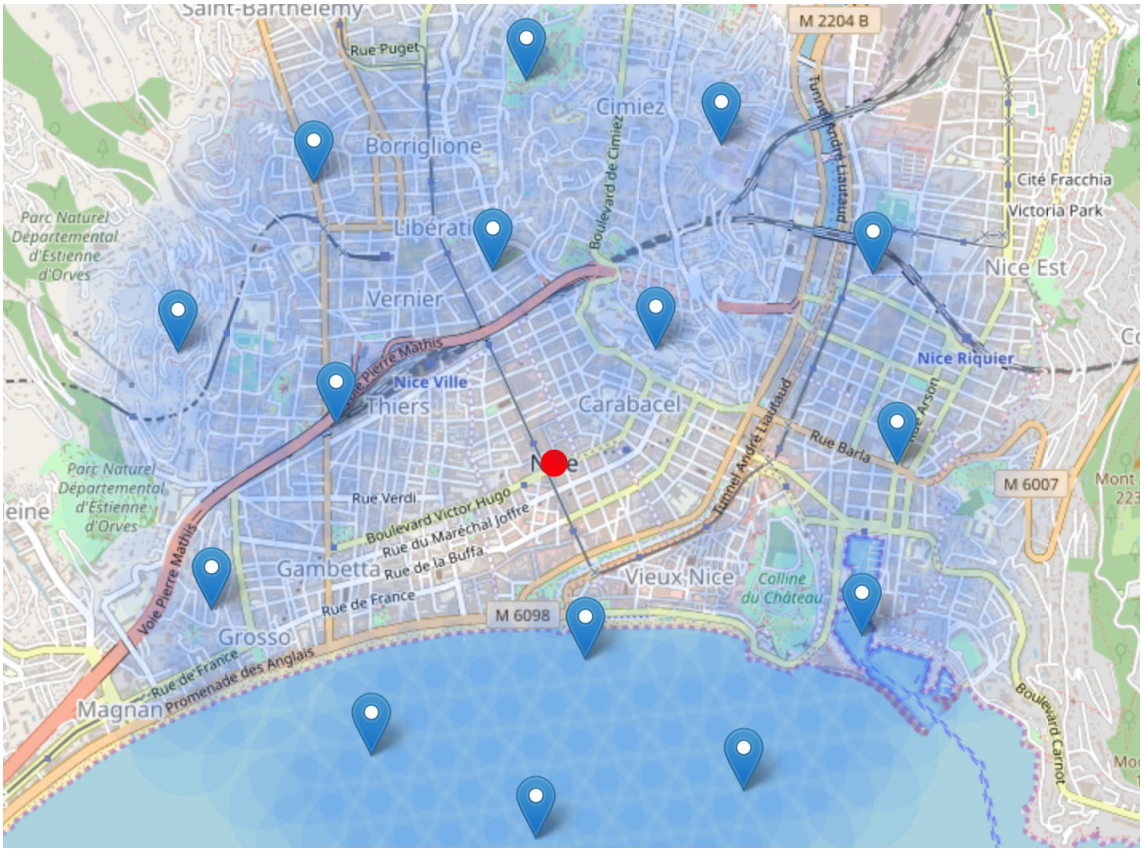
This map is not so 'hot' (Italian restaurants represent a subset of ~12% of all restaurants in Nice). There is almost no Italian restaurants in the center, the north and in the west.

We filtered the locations that do not have more than two restaurants in radius of 400m, and no Italian restaurant closer than 150m. Any of those locations is a potential candidate for a new Italian restaurant, at least based on nearby competition.



We cluster the locations to create centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis.





Addresses of centers of areas recommended for further analysis :

1) Villa Saint-George, Avenue Mirabeau, Libération, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06000, France => 0.9km from Nice Center
2) Raubà capeu, Promenade des Anglais, Vieux Nice, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06300, France => 1.6km from Nice Center
3) Massena, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06046, France => 1.6km from Nice Center
4) Avenue des Diables Bleus, Riquier, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06300, France => 1.6km from Nice Center
5) Escalier, Avenue Dauphiné, L'Archet, La Madeleine, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06100, France => 1.7km from Nice Center
6) L'Archet, Grosso, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06046, France => 1.6km from Nice Center
7) 15, Boulevard de Cessole, Saint-Barthélémy, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 0610, France => 1.6km from Nice Center
8) 58, Corniche Sainte-Rosalie, Cimiez, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06106, France => 1.6km from Nice Center
9) Bassin des Amiraux, Digue, Le Port, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06300, France => 1.5km from Nice Center
10) Opera Plage, Promenade des Anglais, Vieux Nice, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06046, France => 0.9km from Nice Center
11) Parc Valrose - Université de Nice - Faculté des Sciences, Avenue Valrose, Saint-Maurice, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06106, France => 1.7km from Nice Center
12) Hi Beach, Promenade des Anglais, Vieux Nice, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06046, France => 1.5km from Nice Center
13) Montée Desambrois, Carabacel, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06000, France => 0.7km from Nice Center
14) 19, Rue Arson, Riquier, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06300, France => 1.5km from Nice Center
15) Voie Pierre Mathis, Quartier Jean-Médecin, Nice, Alpes-Maritimes, Provence-Alpes-Côte d'Azur, France métropolitaine, 06000, France => 1.0km from Nice Center

This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants and no Italian restaurants nearby, all zones being fairly close to city center.

5. Results and Discussion

Our analysis shows that are not many restaurants in Nice and even fewer Italian restaurants. Highest concentration of restaurants was detected in Promenade des Anglais, Vieux Nice and Carré d'OR. These neighborhoods may be more attractive for tourists.

We first created a dense grid of location candidates (spaced 100m appart); those locations were then filtered so that those with more than two restaurants in radius of 150m and those with an Italian restaurant closer than 400m were removed. Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers 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 15 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 Nice center 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 taken into account and all other relevant conditions met.

6. Conclusion

Purpose of this project was to identify Nice areas close to center 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 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 centers 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 neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood