

The recommended Neighborhoods to open a catering services in Casanearshore

1. Introduction

1. Background

From a small port at the beginning of the 20th century, to a big city, Casablanca is the largest city in Morocco stretched on the Atlantic coast and located 90 km south of the capital Rabat, this city is considered today as the economic and commercial capital of the Morocco Kingdom.

Casablanca is the first most populated city in Morocco with approximately 4.75 million inhabitants. It is the third largest city in Africa after Lagos and Cairo.

Industrial activity is predominant and is a powerful catalyst for the growth and extension of the urban fabric. It dominates the national industrial space by the weight of its production equipment, the diversity of manufactured goods and its place at the top of the hierarchy of industrial centers of the country. Indeed, Casablanca employs 39% of the workforce, accounts for 35% of the country's electricity consumption and absorbs 1,231 million tons of cement. The ports of Casablanca and Mohammedia account for 55% of foreign trade. Its airport serves 51% of passengers.

Casanearshore park is a service area located in Morocco, dedicated to companies specializing in software development, infrastructure management, back office banking and insurance and customer relationship management

- built in a wooded park of 53 hectares
- 20 000 jobs on site
- billion dirhams contribution to GDP by 2015
- multinational companies already installed in the park

The park is characterized by : More than 300 000 m² of offices and services with Considerable pool of human resources and Special incentive framework and competitive operation cost. Moreover, the park provides its customers with a Food Court with a capacity of 900 seats with selected Food Franchise offering various foods, as well as inter-companies restaurants.

Opening a restaurant is all about location. However, not every restaurant is suitable for every location, and vice versa. It comes down to a combination of restaurant concept and ideal customer. If you can define your restaurant type and identify your target demographic and its most populated areas, you'll be well on your way to choosing a restaurant location that sets your business up for success.

Understand the customer environment is mandatory before deciding where to open a restaurant. Environment that surrounds your potential restaurant location is all about the neighborhood characteristics: population, schools, industry, market, shopping centers and so on. The neighborhood industries can give an idea about customer income level, age and if there is a large employee pool that can work within your restaurant concept. Is there a college in the area or entertainment options? This can help to determine customer age interval and can show you whether the neighborhoods you're considering will benefit from future growth.

Another way to identify where you'll find your ideal customer type is to look at competing businesses; restaurants often choose to be located next to their fiercest competitors. Existing restaurant types in the neighborhood gives an idea about the customer types also.

Customer environment understanding is the best way to understand your customer needs and to determine which style your restaurant falls under: Fast food, Bistro, casual dining, Fine dining

2. Problem

In this project we try to find the optimal location for a restaurant with the concept of collective catering near to casablanca casanearshore. The catering services can be dedicated, primarily, to casanearshore employees and students. This work uses location data of Casablanca neighborhoods to explore and analyses venues categories in casanearshore neighborhood and thereafter proposes a clustering of this park (casanearshore) which recommend where a Catering services can be implemented.

Select the appropriate features among venues categories to perform the clustering is crucial. Features may include:

- distance to casanearshore park
- restaurant concepts: restaurant and coffee related words,
- Customer type: education and industry related words
- Population : market and shopping related words
- distance to other restaurants that provide the same services
- distance to market and shopping centers

If we can get demographics data related to each venue in this neighborhood, it will be interesting to check if there is a correlation with our data and include it in the model.

Limitation : latitude and longitude data must be correct and complete otherwise the neighborhood will be covered partially.

2. Data acquisition and cleaning

2.1 Data sources

This work addresses three data sources. Casablanca neighborhood data is extracted from the post referential (http://codepostal.ma/search_mot.aspx?keyword=casablanca) which includes all postal codes and their neighborhoods of the Moroccan cities. To capture geolocation data, Google geolocation api is used. This last provides the geographic coordinates: latitude and longitude of each neighborhood by concatenating the city name “Casablanca” and the neighborhood “name”. The last data source is Foursquare referential to get all venues of each neighborhood particularly, the venues categories and locations.

2.2 Data cleaning

Data are downloaded or scraped from each source then combined into one dataframe. Multiple neighborhood entries exist for the same postal code. Some missing values for latitude and longitude or wrong location of the neighborhood (outside Casablanca) were identified and dropped, to keep just the neighborhoods that contain the correct coordinates.

2.2.1 Locate Casanearshore neighborhoods

Based on latitude and longitude of Casanearshore center (Table 1), we have created a hexagonal grid of cells equally spaced covering our area of interest that represents approximately 2* 2 kilometers around Casanearshore center. Each cells area has a radius of 300 meters. To accurately calculate distances, the Cartesian 2D coordinate system which allows us to calculate distances in meters (not in latitude/longitude degrees) has been defined and used. Then Cartesian coordinates will be projected back to latitude/longitude degrees to be shown on Folium map as shown in Figure 1.

This approach has generated 38 candidate neighborhood centers.

Table 1 : Casanearshore coordinates

Position	latitude and longitude Coordinates	Cartesian coordinates
Casanearshore center	Longitude =-7.64010629999999 latitude =33.5252545	X=-1623198.1413343325 Y=3948747.710696002



Figure 1: Casanearshore map and grid cells using Folium

For each cells in the grid, Google Maps Api is used to get for each location latitude and longitude that were calculated automatically using the Cartesian coordinates, the approximate google address of the position. 38 addresses were retrieved and the distances from Casanearshore center to those addresses and the distance to closest catering restaurant were calculated as shown in the Table 2. The average distance to the closest Catering restaurant from each area center is 604.60 meters

Table 2 : list of some addresses with their coordinates

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Catering restaurant
0	Rue 100	33.510421	-7.644396	-1.623998e+06	3.947150e+06	1787.133900	0	1434.609731
1	Unnamed Road	33.511571	-7.638435	-1.623398e+06	3.947150e+06	1610.542634	0	1012.563982
2	10 Route de Bouskoura, Casablanca 20000	33.512721	-7.632473	-1.622798e+06	3.947150e+06	1647.375967	1	730.497149
3	P3020	33.513871	-7.626512	-1.622198e+06	3.947150e+06	1885.165133	0	872.906840
4	Res.dandoun	33.514169	-7.648566	-1.624298e+06	3.947669e+06	1540.479815	0	942.827361

2.2.1 Locate Casanearshore venues

Using Foursquare Api, the venues of each neighborhood of Casanearshore are retrieved and analyzed. In this study, two types of venues have been checked and retrieved to get knowledge about (1) the population characteristics and (2) the restaurant typology for an area:

For population characteristics type, we are interested to check if there are enough professional and student people in each selected area using the Foursquare categories below:

- Professional & Other Places : 4d4b7105d754a06375d81259
- College & University category : 4d4b7105d754a06372d81259
- Business Center : 56aa371be4b08b9a8d573517
- Business Service : 5453de49498eade8af355881
- Office : 4bf58dd8d48988d124941735

The second type that concerns restaurant typology is related to the following Foursquare categories:

- For generic restaurant, we selected all food category : 4d4b7105d754a06374d81259
- For Catering competitors, we selected :
 - o Restaurant : 4bf58dd8d48988d1c4941735
 - o Buffet : 52e81612bc57f1066b79f4

- Comfort Food Restaurant : 52e81612bc57f1066b7a00
- Cafeteria : 4bf58dd8d48988d128941735
- Food court : 4bf58dd8d48988d120951735
- Theme restaurant : 56aa371be4b08b9a8d573538

The total of retrieved venues using Foursquare is 114 which are distributed as mentioned in the Table 3. The percentage of catering restaurant among the total number of restaurant is: 10,61% and the average number of restaurants in each area of radius of 300 meters is 4.34.

Table 3 : distribution of venues among the selected types

Type	Number of venues
Number of restaurant	198
Number of Catering restaurants	21
Number of business (population characteristics)	179

2.3 Exploratory Data Analysis

This section explores the '**restaurant density**' across different areas of Casanearshore. Using the heatmaps, the study identifies a few areas close to center with low number of restaurants in general (*and* no Catering restaurants) and focuses the attention on those promising areas.

2.3.1 location of restaurants and business/college in the Area

Firstly, it is important to have a visual insight about the collected data of our area of interest using a map. Different colors are used to differentiate between business (yellow), restaurant (red) and catering restaurant (bleu).

The map in Figure 2 illustrates in yellow how much businesses around our area of interest with two concentration points : in the center and the north of Casanearshore marker. The bleu points which present the restaurants are distributed in the west (from north to south) of the casanearshore center. Moreover, the catering restaurants with red color are few but presents in our area, they are distributed in the west like the other restaurants except for 2 or 3 ones.

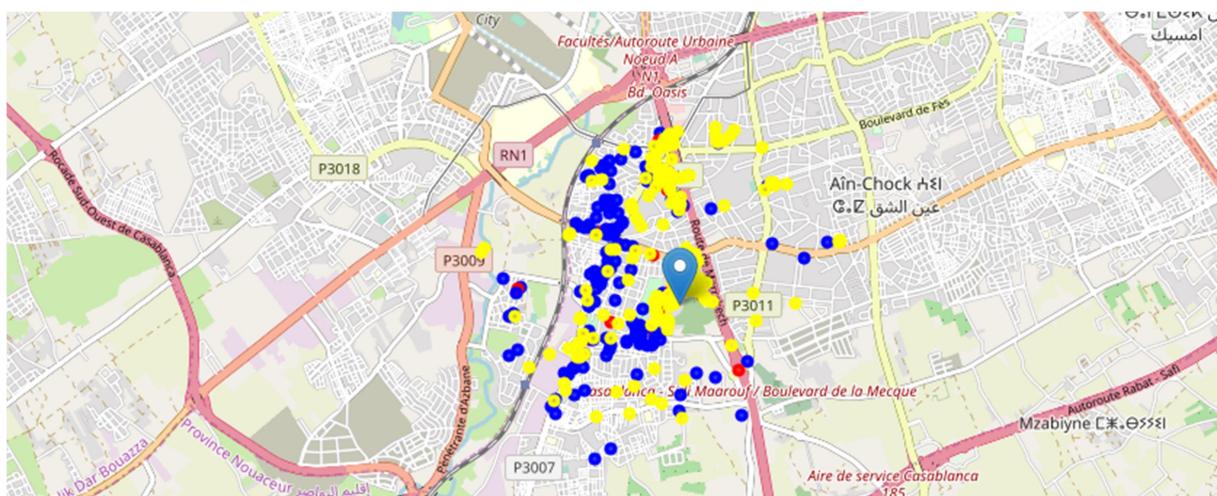


Figure 2 : collected venues of Casanearshore in a map

2.3.2 Relationship between number of restaurant and distance from center

This section analyses the number of restaurants and the distance from center, the distance to catering restaurant and their relationship.

This study found that some neighborhoods in casanearshore area have no restaurants and for others we can have 26 restaurants with 4 restaurants as mean which is still little to give more choice to customers.

The Figure 3 shows the distribution of distance from center of the restaurants, 50% of restaurants are located between 1024 meters (25% quartile) and 1664 meters (75% quartile) with a mean of 1290 meters. However, the distance to a catering restaurant varies between 77 meters (min) and 1617 (max), 50% are below 850 meters. This last metrics illustrate the sparsity that exists between neighborhoods in Casanearshore area in terms of distance to catering restaurants. Some employees/students from Casanearshore area have a quick access to catering services and others not. We note also that more restaurants are installed from 500 meters and mores



Figure 3 : the distributions of distance from center and number of restaurants metrics

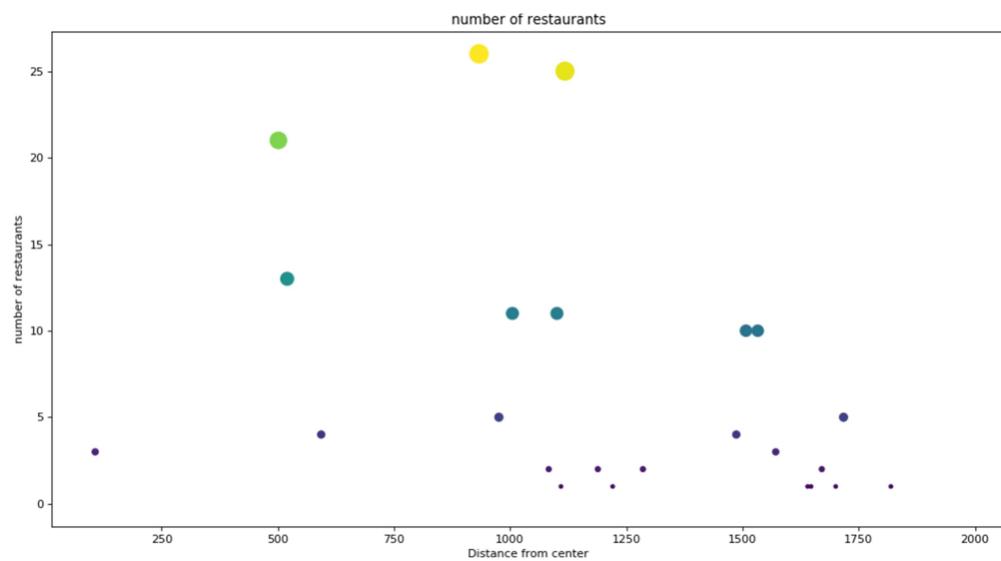


Figure 4 : number of restaurants and distance from center relationship

2.3.1 Density exploration of restaurants and catering in the Area

This section goes through the restaurant density in our area of interest and explores characteristics of this density to identity non covered zones.

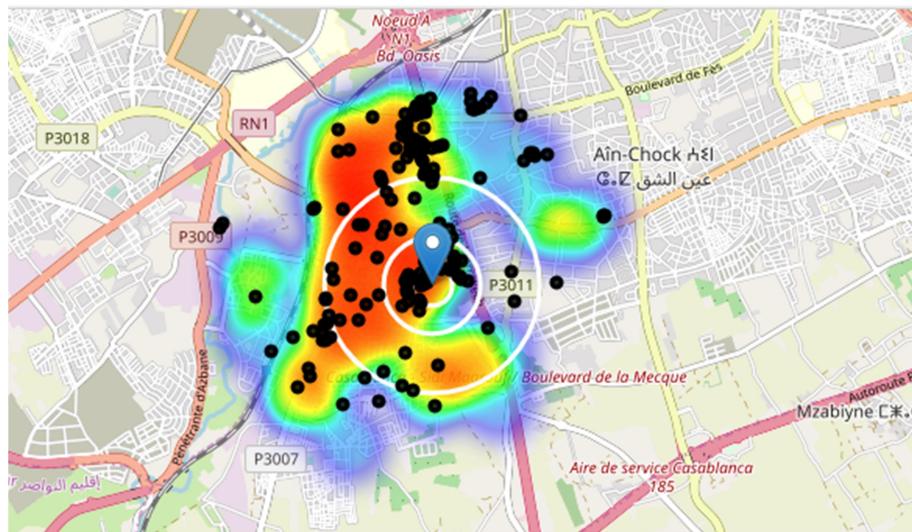


Figure 5 : heatmap of restaurants density

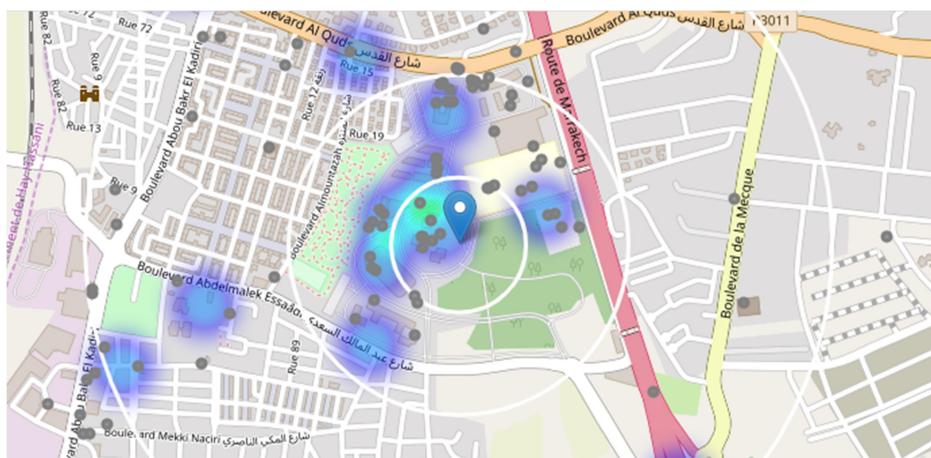


Figure 6 :heatmap of catering density(zoomed)

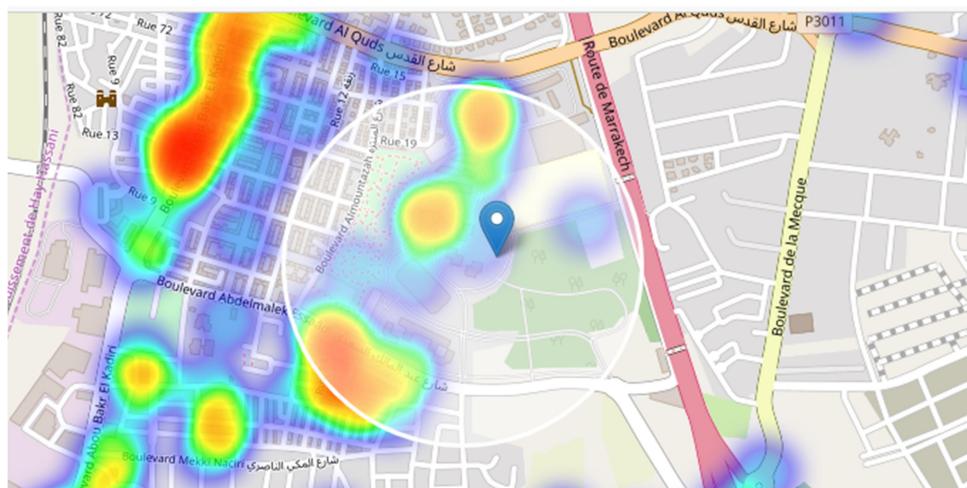


Figure 7 : map of the area of interest

Figure 5 shows a heatmap of density of restaurants to extract some meaningful information. Also, it shows three borders of casanearshore boroughs which indicate the distance of 200 m, 500 m and 1100 m from Casanearshore center. Looks like a few pockets of low restaurant density closest to the center can be found in the east from casanearshore and some in the North east and south-east.

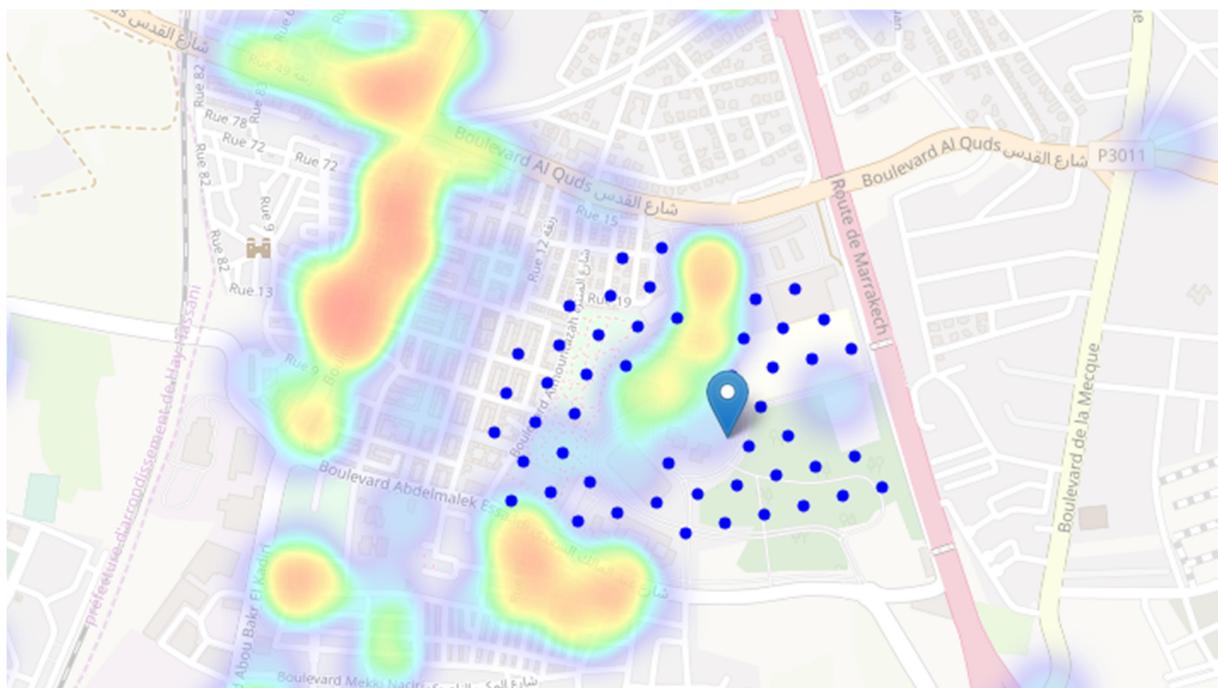
The Catering restaurants density is different from restaurant density. This map(cf Figure 6) is not so 'hot' (Catering restaurants represent a subset of ~4.34% of all restaurants in casa nearshore) but it also indicates higher density of existing catering restaurants directly south-west and west from casa nearshore site, with closest pockets of low catering restaurant density positioned east, south-east and North from center. There is also a low density of catering restaurant in the south with approximately 200 m of distance from the center.

2.3.2 Definition of area of interest

Based on last section, this study focus the analysis on areas: east, North and south-east from Casanearshore center, we will keep the center of our area of interest and reduce its size to have a radius of **500 m**(cf Figure 7). This places our location candidates mostly (The white border):

- near to boulevard "Abdelmalek Essaadi"(near to route de Marrakech),
- and in the green park in the north of this boulevard,
- and near to boulevard "almountazah" (garden)
- another potentially interesting zone is the north east of boulevard "almountazah" with large low restaurant density north-east from city center,

The east of casanearshore center cannot be among the candidates, since it's a part of Casanearshore park that includes only business centers.



2.4 Clustering Modeling

In this section, the application of clustering algorithm Kmeans is illustrated using geospatial data and map presentation. The present study uses the popular Kmeans clustering algorithm for cluster analysis. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells.

Three key features of k-means that make it efficient are often regarded as its biggest drawbacks:

- Euclidean distance is used as a metric and variance is used as a measure of cluster scatter.
- The number of clusters k is an input parameter: an inappropriate choice of k may yield poor results. That is why, when performing k-means, it is important to run diagnostic checks for determining the number of clusters in the data set.
- Convergence to a local minimum may produce counterintuitive ("wrong") results.

In this study, we set the k number of cluster to 10 clusters, since our zone of interest is small with 250 m of radius. Our clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates.

Addresses of those cluster centers will be a good starting point for exploring the neighborhoods to find the best possible location of the catering restaurant based on neighborhood specifics.

Figure 8 illustrates those zones on a city map, using purple circle to indicate our 10 clusters. An additional heatmap is displayed to show the density of restaurant in those zones.

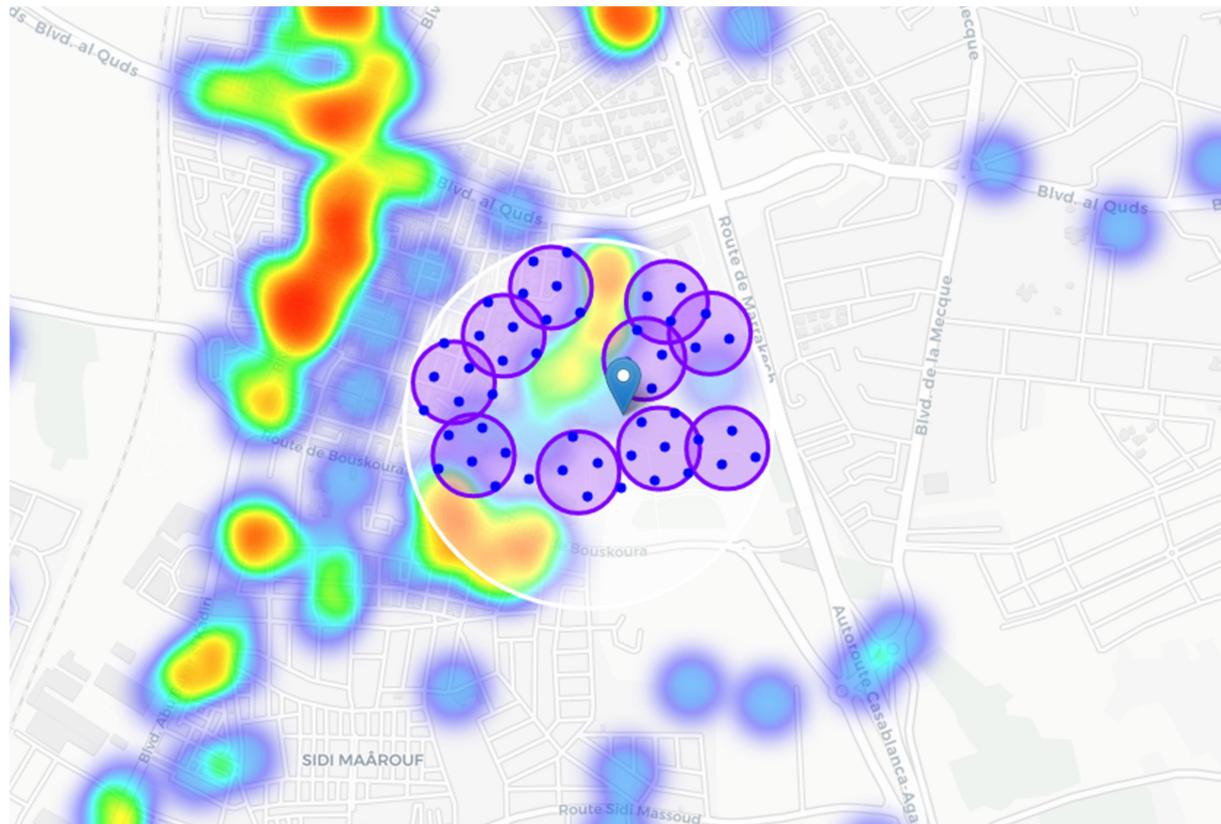


Figure 8 : the 10 generated clusters of our area of interest

Addresses of centers of each cluster areas are retrieved using reverse location for further analysis as shown bellow:

Unnamed Road, Casablanca	=> 0.5km from casanearshore
shore 12-B Casanearshore, Casablanca	=> 0.2km from casanearshore
Casablanca Nearshore, Casablanca	=> 0.4km from casanearshore
Km 9,500 Route de Nouasseur BP. 8114 - Casablanca	Oasis, Casablanca => 0.3km from casanearshore
346 Boulevard 60, Casablanca	=> 0.4km from casanearshore
N11, Casablanca	=> 0.3km from casanearshore
Km 9,500 Route de Nouasseur BP. 8114 - Casablanca	Oasis, Casablanca => 0.2km from casanearshore
Unnamed Road, Casablanca	=> 0.4km from casanearshore
Unnamed Road, Casablanca	=> 0.4km from casanearshore
Unnamed Road, Casablanca	=> 0.1km from casanearshore

This concludes our analysis. We have created 10 addresses representing centers of zones containing locations with low number of restaurants and no Catering restaurants nearby, all zones being fairly close to Casanearshore center (all less than 500m from casanearshore,). Although zones are shown on map with a radius of ~100 meters (purple circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most of the recommended zones are located in the west of the casanearshore center, which we have identified as interesting due to being accessible for business/student people and also possible to build services in, and well connected by public transport.

2.1 Results and Discussion

Our analysis shows that although there is a great number of restaurants in Casanearshore (~2000 in our initial area of interest which was 12x12km around casanearshore center), there are pockets of low restaurant density fairly close to west. Highest concentration of restaurants was detected north-south and west south from casanearshore center, so we focused our attention to areas near West, south-east and east, corresponding to boroughs near to boulevard "Abdelmalek Essaadi"(near to route de Marrakech), in the green park in the north of this boulevard, and near to boulevard "almountazah" (garden)

After directing our study to this more narrow area of interest (covering approx. 250x250m west from center) 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 100m and those with an catering restaurant closer than 100m 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/startng points for more detailed local analysis based on other factors.

Result of all this is 10 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Catering restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new

catering service! Purpose of this analysis was to only provide info on areas close to Casanearshore center but not crowded with existing restaurants (particularly Catering) - 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.

Références:

https://en.wikipedia.org/wiki/K-means_clustering

<https://fitsmallbusiness.com/choose-a-restaurant-location/#additional-5>