1. Introduction: Business Problem

Portugal as become a multicultural country because of mass turism and foreign citizens that decided to move into. This analysis will clear where is the best place to open a restaurant reducing the risk based on population.

There are a lot of restaurants types suchs as: Italian Food; Chinese Food; Brazillian Food; Typical Portuguese; African Food; Indian;

The ideia is that a the better place to open a restaurant is based on: where population has higher income; Portugal Population; Portugal Demographics; Foreign citizens will perfer food from their countries (e.g. Brazillians likes Brazzilian Food, Angola citizen will prefer African Food, ...); Is there competitors in that location?

Target Audience:

This would interest to investors who wants to start a new restaurant in Portugal, with focus to foreign community that lives in Portugal

2. Data

Data Source

I will use the Foursquare API to find venues in Portugal where cateregories are related to Food, inside Food Categorie, i will consider only the ones that has the word restaurant (e.g. Afghan Restaurant, African Restaurant, ...)

I will consider 3 majors datasets:

Portugal Population Distribution:
https://www.pordata.pt/Municipios/Popula%c3%a7%c3%a3o+residente+total+e+por+grandes+grupos+et%c3%a1rios-390

Foreign citizens relation with regions that they live in Portugal:
https://www.pordata.pt/Municipios/Popula%c3%a7%c3%a3o+estrangeira+com+estatuto+legal+de+residente+total+e+por+algumas+nacionalidades-101

It will be possible to cross this data with Forsquare API data to find lack or excess of restaurants related with foreign citizens residents in Portugal.

Income ditribution in Portugal:
https://www.pordata.pt/Municipios/Propor%c3%a7%c3%a3o+de+poder+de+compra-119

Data cleaning

DataSets were in .xls format, so i had to clean some and convert into .csv The main problems were:

encoding; conversions from text to integer; columns labels;

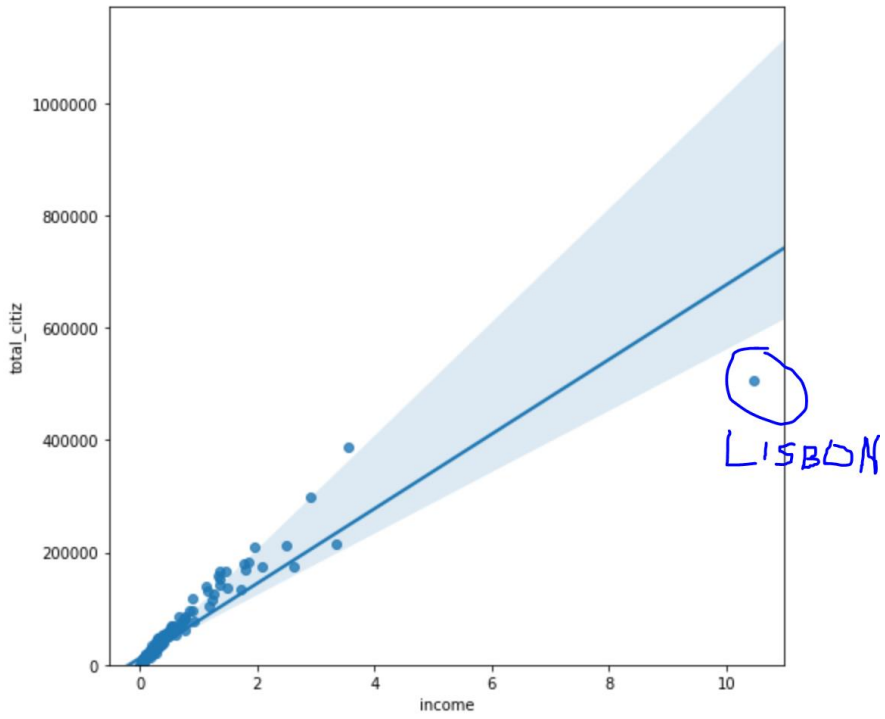
Exploratory Data Analysis

Choose a city

It was easy to choose a city for the study. Lisbon is the capital, where we find most of the Portugal population and is the region where exists more money.

	regiao	total_citiz
213	Lisboa	506654
224	Sintra	387236
47	Vila Nova de Gaia	299879
38	Porto	214936
212	Cascais	212094

	regiao	income
213	Lisboa	10.466
224	Sintra	3.550
38	Porto	3.350
47	Vila Nova de Gaia	2.901
219	Oeiras	2.633



Choose type of restaurant

I notice that African population is most representative of the foreign citizens in Portugal. It is related to the fact that countries (Angola, Cabo Verde…) were colonies of Portugal in the past.

tipo	MunicípioMunicípioMunicípioMunicípioMunicípioM...
regiao	LisboaSintraCascaisAmadoraLouresLouléOdivelasA...
Total	477472
Espanha	14066
Franca	19771
Reino Unido	26445
Ucrania	29197
Romenia	30908
Moldavia	4834
others_ue	77077
Angola	18310
Cabo_Verde	34444
Guine_Bissau	15960
Mocambique	2999
Sao_Tome_Principe	9023
others_african	9035
Brasil	104504
others_americans	13461
China	24856
india	11340
Nepal	11487
others_asian	19258
pt_africans	80736

Viewing the data, I realized that can merge all African countries in 1 column and then I realized that the foreign communities most representative are African (80736) and Brazilian (104504)

The focus of the study will be in Lisbon and restaurants typical African or Brazilian.

3. Methodology

In this report i will detect areas in Lisbon that have low restaurant density, particularly those with low number of African or Brazilian restaurants.

The analysis will be limit 5km area around city center.

I will collect data from Foursquare API to identify restaurants and also African or Brazilian restaurants(according to Foursquare categorization).

Next I will calculate density of all restaurants

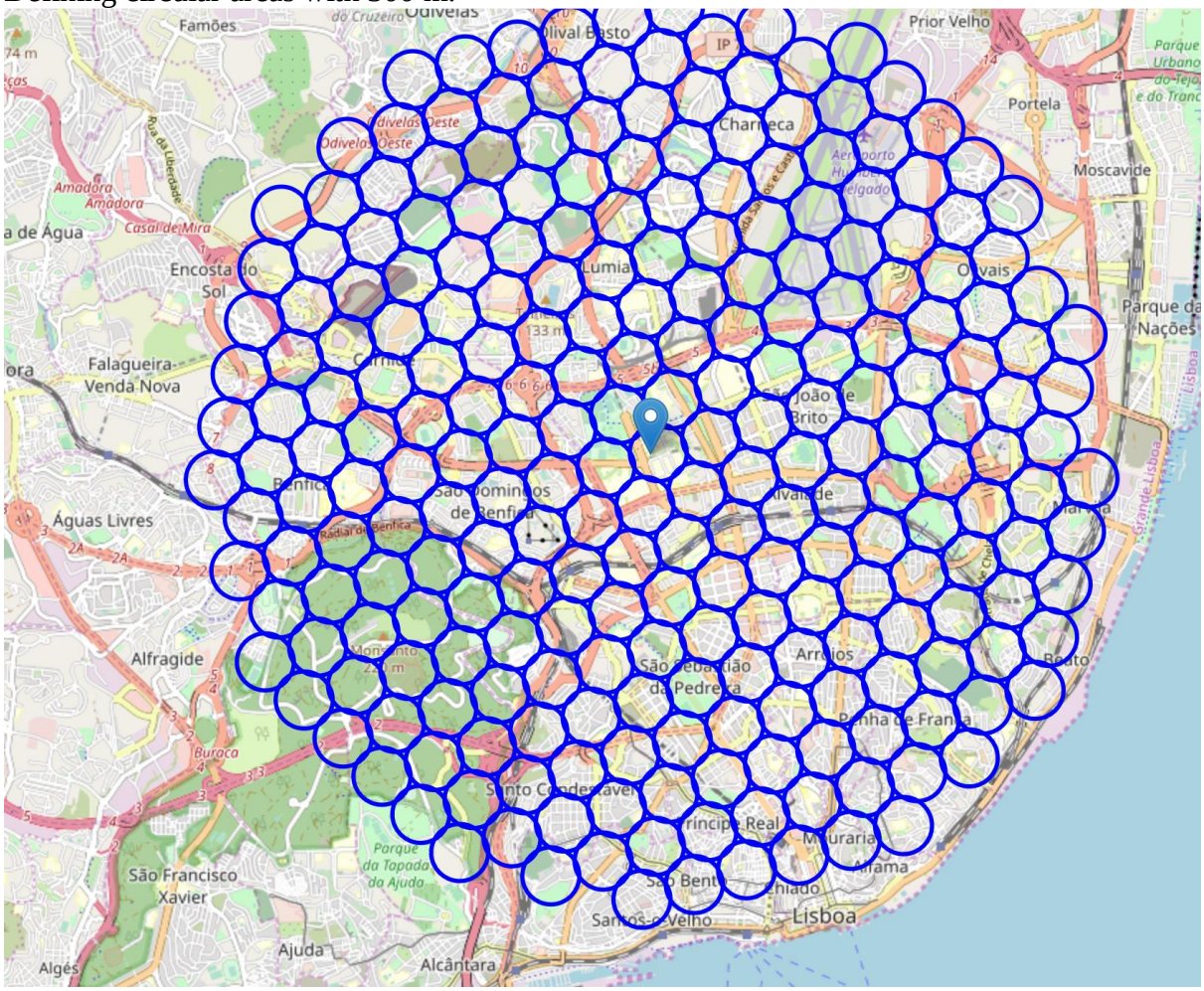
Heatmaps will help to identify some areas with low number of restaurants in general (and no African/Brazilian)

Study one of those areas using k-means.

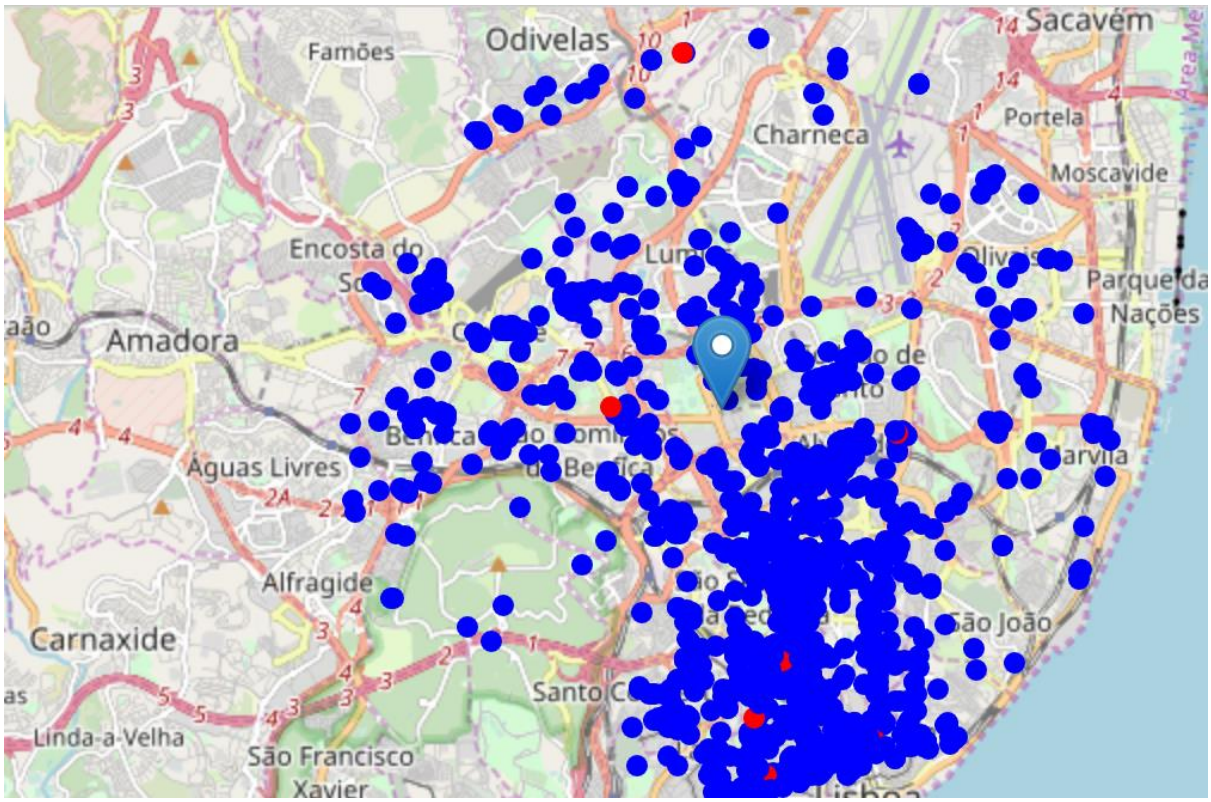
Cluster the results in a map

Analysis

Defining circular areas with 300 m:

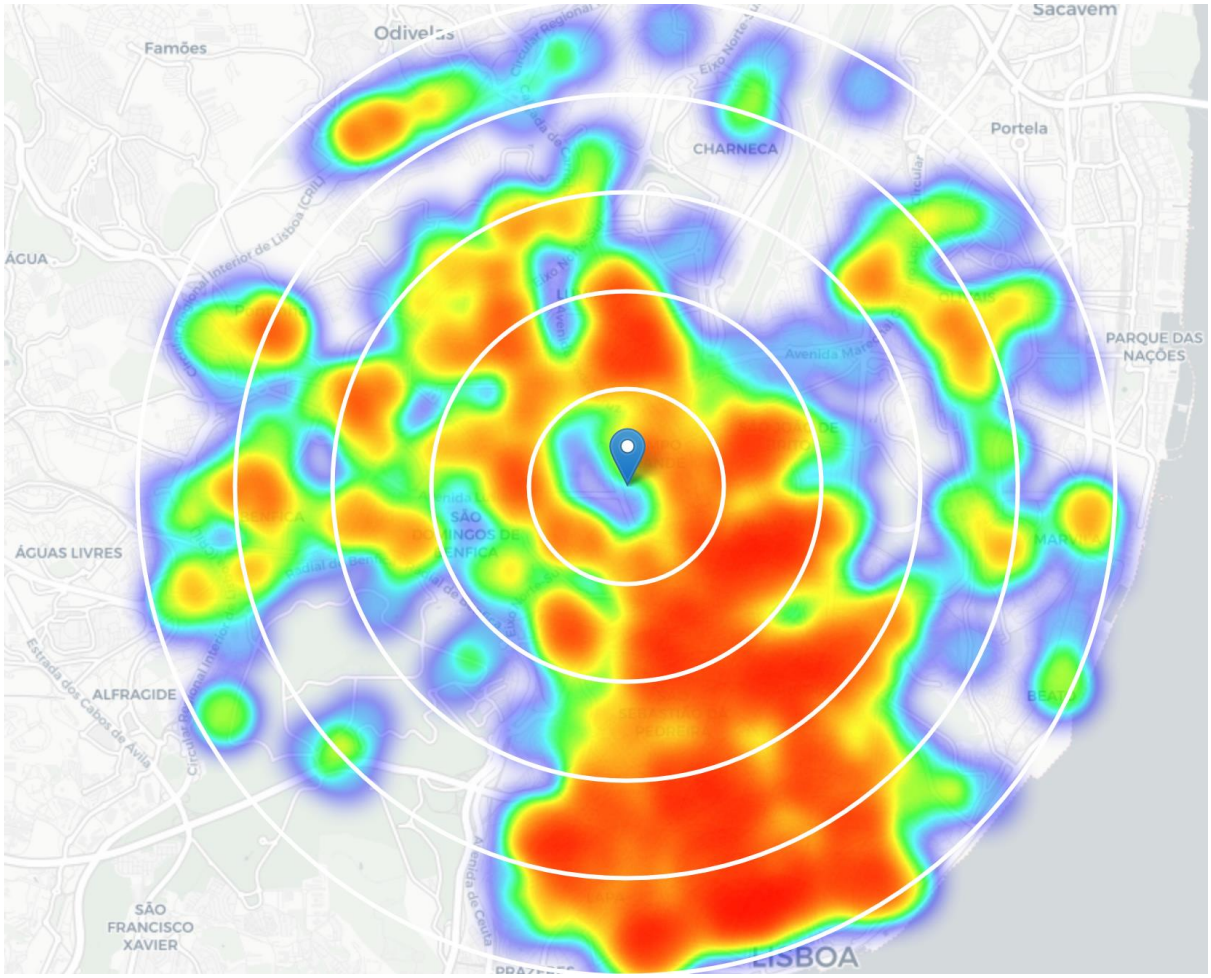


Using the Foursquare API, I collected data from all restaurants, including a column to determinate if the restaurant is African or Brazilian type (according to Foursquare categorization).

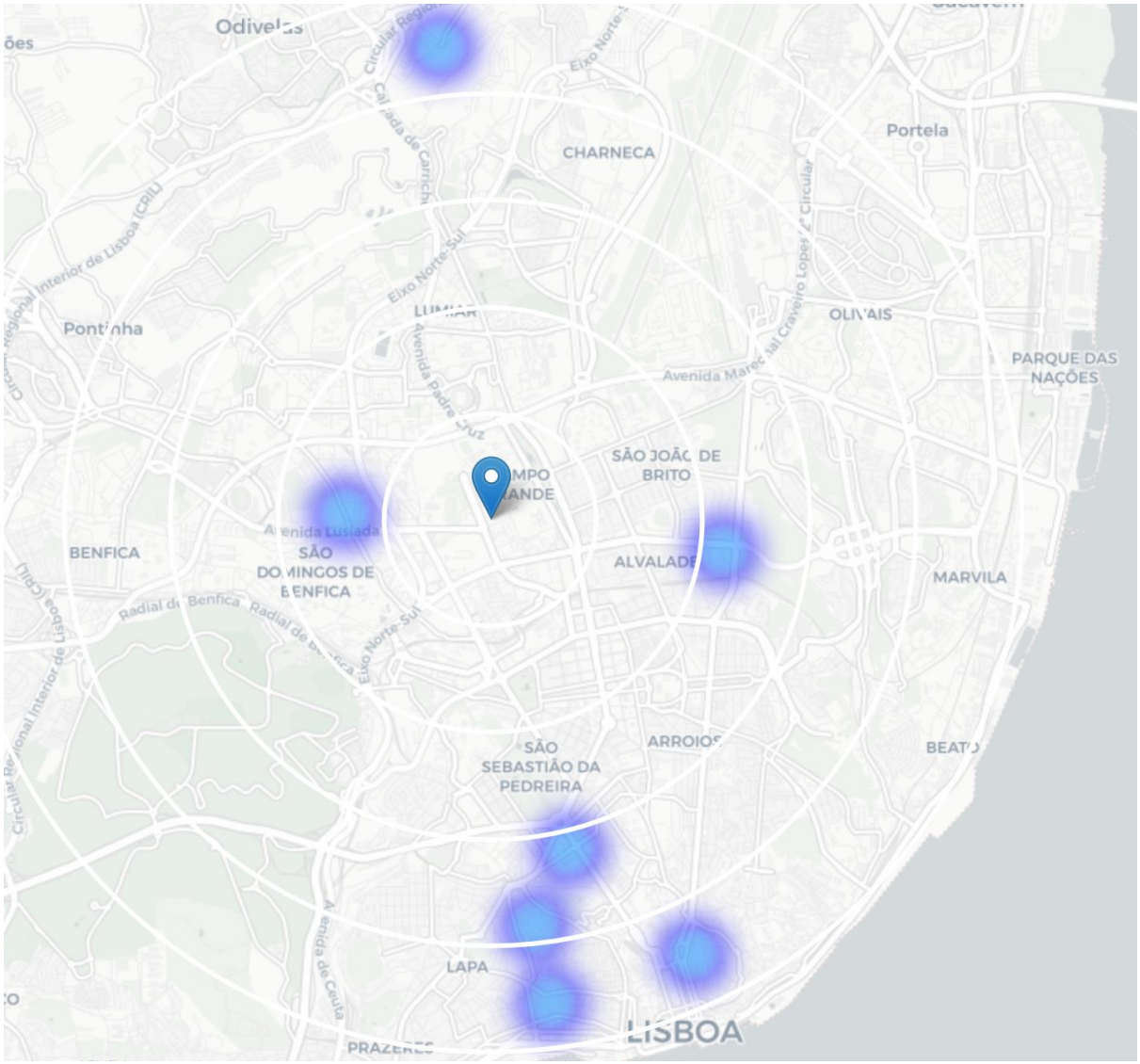


Using heatmaps to determine density of restaurants.

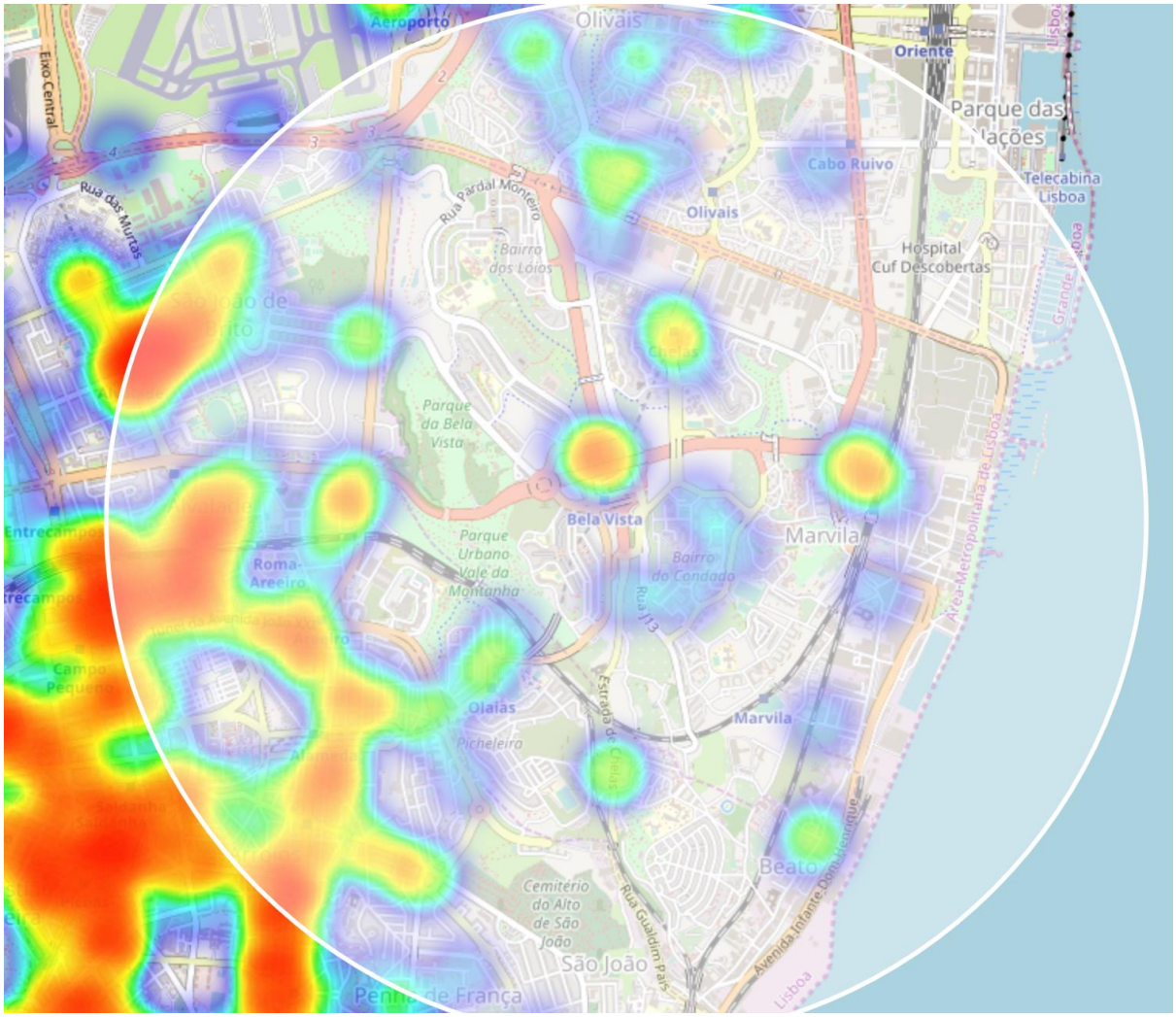
All:



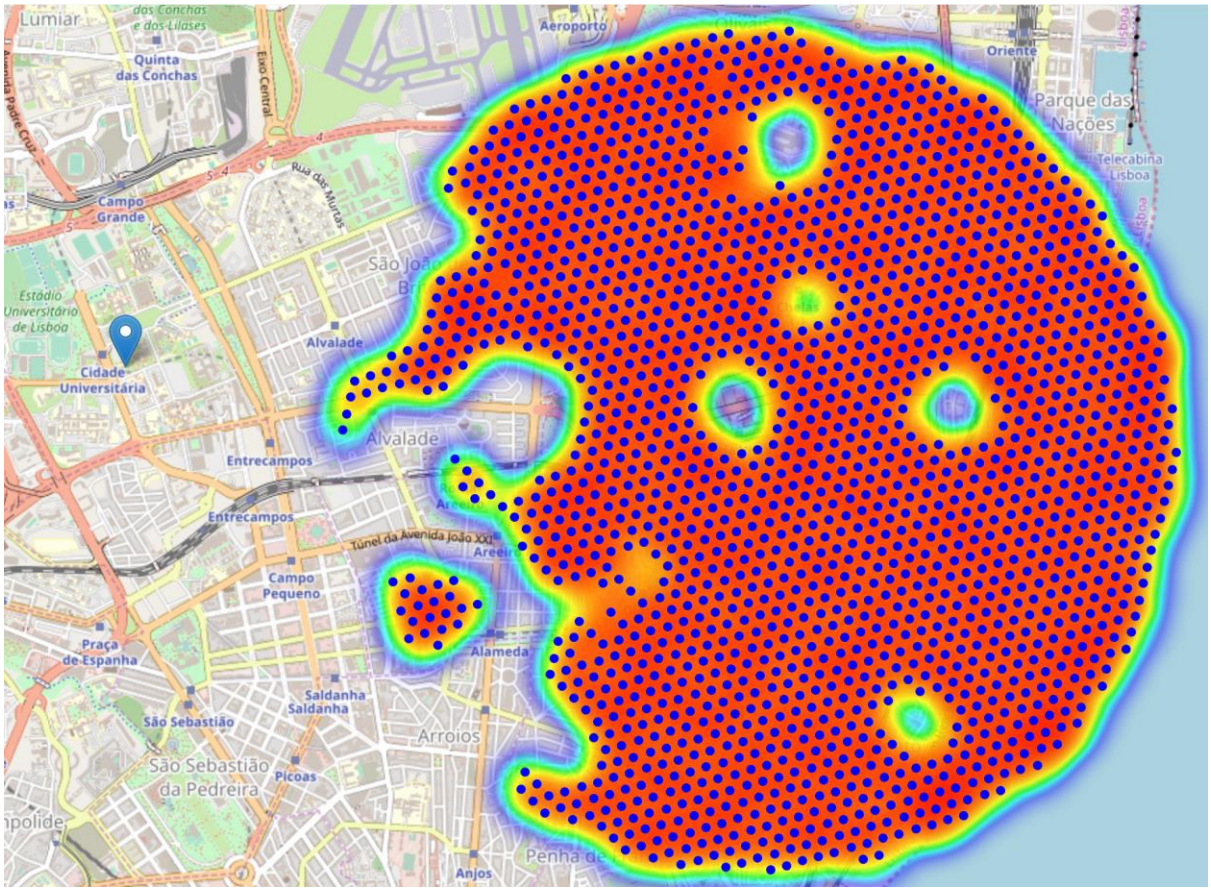
Only African or Brazilian:



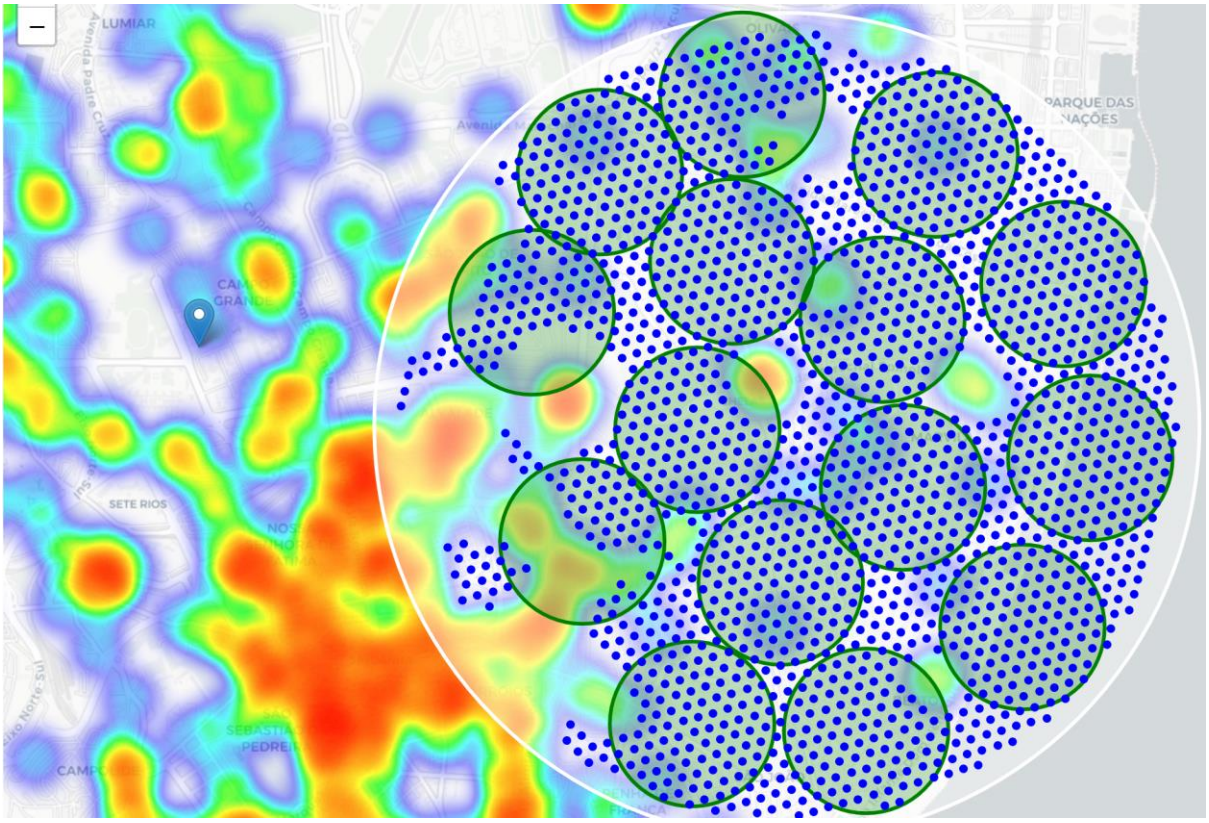
Focus my attention to areas with low density of restaurants and most near city center. East side of Lisbon city (Bela Vista / Marvila)

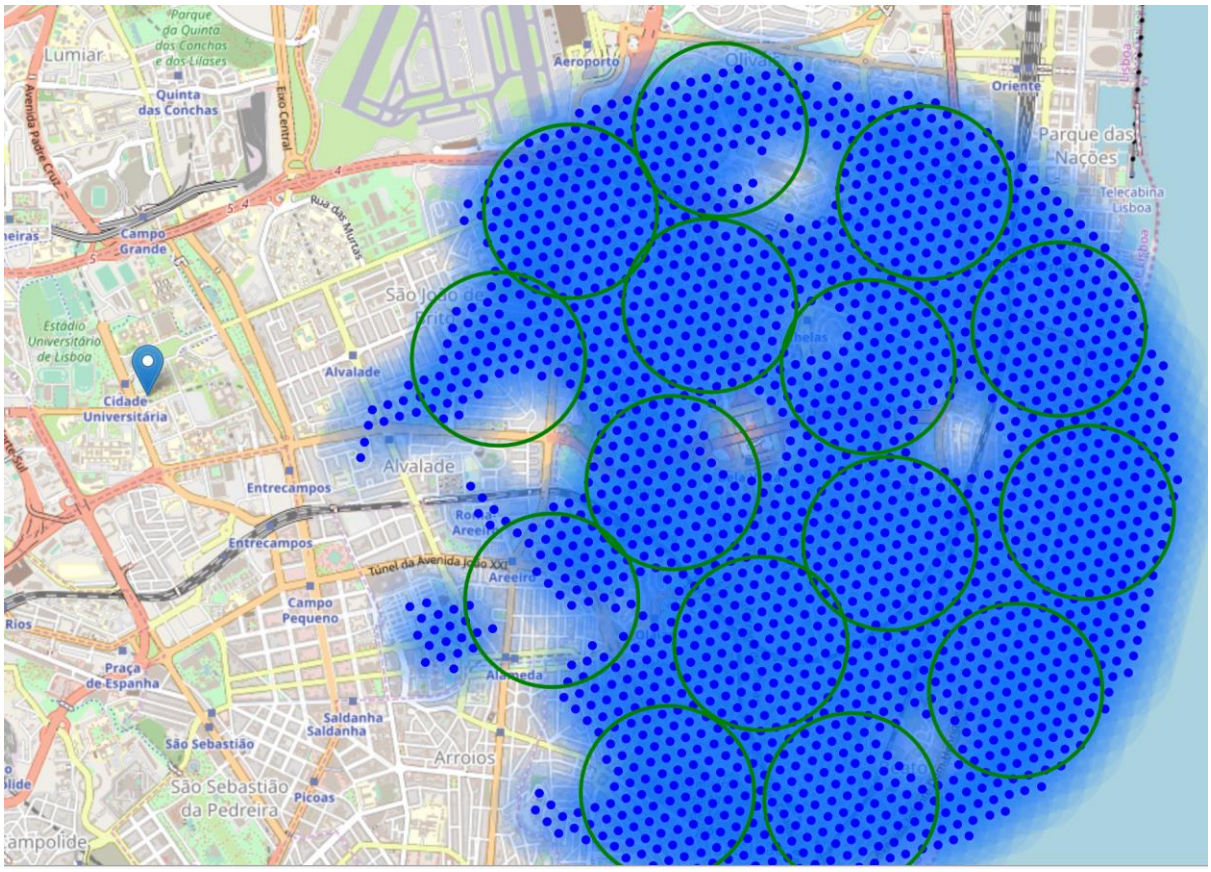


Defining a grid around the candidate area no more than two restaurants in radius of 250 meters and locations without African/Brazilian restaurants in radius of 400 meters.



Create clusters (using **k-means clustering**) of those locations to identify general areas ideal from opening the restaurant


LUMIAR
SETE RIOS


Lumiar
Quinta das Conchas
Campo Grande
Cidade Universitária
Alvalade
Entrecampos
Campo Pequeno
Praça de Espanha
São Sebastião
Saldanha
Arroios
São Sebastião da Pedreira
Picoas
Aeroporto
Oriente
Parque das Nações
Telecabina Lisboa
Rua das Murtas
Túnel da Avenida João XXI

4. Results and Discussion

Lisbon is the most rich and polulated area in Portugal.

Most expressive foreign community in Portugal are Africans and Brazilians

I found 1192 restaurants in Lisbon within 10kms around city center.

Only 8 are African or Brazilian.

Lower concentration of restaurants east of Lisbon

Efforts were made in area of Marvila / Bela Vista (east)

Defined that the restaurant does not have 2 restaurants nearby (250m) and no African/Brazilian in radius of 400m

15 areas wore identified with potential to open an African or Brazilian restaurant

5. Conclusion

Purpose of this project was to identify the best place to open a restaurant in Portugal.

Lisbon city was identified and 15 cluster areas were identified, based on FourSquare API.

Final decission on optimal restaurant location will be made by investors based on specific characteristics of areas identified by this study. I recommend a study abount criminality, proximity of public transportation, turism interess, etc.