

CAPSTONE PROJECT - THE BATTLE OF LISBON NEIGHBORHOODS

DANIEL LORAS LEITE

November 2019

CONTENTS

| | | |
|-----|---|----|
| 1 | Introduction | 3 |
| 1.1 | Background | 3 |
| 1.2 | Business Problem description | 3 |
| 2 | Data Section | 4 |
| 2.1 | Data of Current Situation | 4 |
| 2.2 | Data Required to resolve the problem | 4 |
| 2.3 | Mapping of Data | 7 |
| 3 | Methodology section | 8 |
| 3.1 | Strategy and Exploratory analysis | 8 |
| 3.2 | K-Cluster analysis | 12 |
| 4 | Results section | 13 |
| 5 | Discussion section and Conclusions | 15 |
| 6 | Limitations and recommendation for future study | 16 |

LIST OF FIGURES

| | | |
|-----------|---|----|
| Figure 1 | Touristic places of interest in Lisbon using Foursquare API | 5 |
| Figure 2 | Restaurants dataframe | 6 |
| Figure 3 | Final dataframe regarding the places of interest in Lisbon. | 6 |
| Figure 4 | Borough Table | 7 |
| Figure 5 | Lisbon Map with places of interest | 7 |
| Figure 6 | Number of Citizens Living in each Lisbon borough | 8 |
| Figure 7 | Number of Restaurants/Bars in each Lisbon borough | 9 |
| Figure 8 | Lisbon Map with places of interest | 9 |
| Figure 9 | Lisbon Map with places of interest | 10 |
| Figure 10 | Cluster Map of restaurants. | 10 |
| Figure 11 | General Cluster Map of Lisbon Restaurants | 11 |
| Figure 12 | Cluster Map of restaurants red zone. | 11 |
| Figure 13 | Cluster initial dataframe. | 12 |
| Figure 14 | Cluster 1. | 13 |

¹ Department of Electrical and Computer Engineering, Northeastern University, Boston, United States

| | | |
|-----------|--------------------|----|
| Figure 15 | Cluster 2. | 13 |
| Figure 16 | Cluster 3. | 13 |
| Figure 17 | Cluster 4. | 13 |
| Figure 18 | Cluster 5. | 14 |

1 INTRODUCTION

1.1 Background

Lisbon is Portugal's amazing capital and one of Europe's most charismatic and vibrant cities. It is a city that mixes traditional heritage effortlessly with striking modernism and progressive thinking. Lisbon offers a rich and varied history as a holiday destination, a buzzing nightlife and a glorious year-round climate. This guide will provide an introduction to Lisbon by answering some of the common vacation questions and providing links to more detailed information. The post-crisis evolution of Lisbon and the corresponding tourism boom demonstrate the dramatic impact of overtourism and its lasting ramifications. Discover Lisbon's culture and history as you retrace the famous journey to India by Vasco de Gama. Take a trip to Belem where you'll be able to witness the historic monuments of the city that mark the Portuguese people's fascinating accomplishments, like Mosteiro Dos Jeronimos. Visit Alfama (the oldest district in Lisbon) and try one of the many local craft shops. Take a break to go shopping for souvenirs and taste some genuine Portuguese wine. Admire the scenic river of Lisbon by coach as you pass through the historic and architectural treasures of the city. Before being taken back to your starting point, enjoy live commentary and soak up the famous sights of the city.

1.2 Business Problem description

Legislation encouraging foreign investment, the sharing economy, and a strong start-up scene combined with the city being a comparatively cheap destination has resulted in tons of travelers heading to Lisbon. It is also relevant that Portugal also has lovely, tiled buildings, beautiful beaches, and a vibrant food scene.

Taking this into account is undoubtedly an asset to invest in Portugal, and especially in Lisbon. In this case, implementing machine learning technology is necessary to help entrepreneurs make wise and successful decisions in Lisbon. As a result, the market issue that we are currently facing is:

- What type of business would be most beneficial in being invested in Lisbon?
- What is the best neighborhood in Lisbon to open this business?
- What does tourist hope to find in Lisbon?
- What is most attractive in the city?

One of the strengths of Portuguese culture is the wonderful cuisine. Having this, an idea that could be implemented in Lisbon would be the case of an entrepreneur that desires to be opening a new restaurant/bar. To solve this business problem, it is firstly needed to have a good database of both Lisbon boroughs /neighborhoods as well as the already established restaurants/bars or any food/drink places. This process will be described in the next section.

2 DATA SECTION

2.1 Data of Current Situation

Churches, cathedrals, and a magnificent castle are just a few of the sightseeing splendors of the city center to tick off. There are also world-class museums such as the impressive Museu Calouste Gulbenkian and Museu Nacional de Arte Antiga where can be seen artistic drawings and historical relics. In Sintra, where royal palaces sit in royal splendor surrounded by lush, ancient woodland, more architectural wonders are unveiled. To be able to visit them all as tourist attractions, a city of Lisbon offers an easy and organized uma rede of public transport. According to many tourism guides Lisbon is known mainly for five unique features and characteristics:

- Fado
- Neighbourhood
- Viewpoints
- Festivals
- Food

To tourists dining in Lisbon, one of the big surprises is how cheap the food is. The prato do dia dish is one of Europe's cheapest options! Fresh fish and seafood choices are popular in many restaurants as it is suitable for a seafaring nation: bacalhau (cod) is a delicious staple. The traditional tasks tucked away in the backstreets of the city are looking for hearty provincial fare. The gourmet hotspots of Lisbon deliver a truly international flavor for more sophisticated palates, with Mediterranean gastronomy especially well represented. Exotic Brazilian, Mozambique and other ex-colonies cuisine add spice to an ethnic menu that is already vibrant. There are plenty of Indian restaurants, and some of the sushi bars in the city have gained celebrity status.

2.2 Data Required to resolve the problem

Since gastronomy is a considerable point between Lisbon and Portugal, data collection for analysis should be carefully and organized. In order to make the analysis with the best possible results, we started by creating a dataframe with the boroughs of Lisbon as well as their geographical data and the number of population in each. To accomplish this first task it was used the data from a Wikipedia table [1] where all the data was reorganized so that there was no redundancy. Having this it was used Microsoft Excel in order to create a CSV containing the Latitude and Longitude of each Lisbon borough, necessary to make the dataframe complete as it is shown in fig.5.

| | Borough | Population | Area(km ²) | Latitude | Longitude |
|----|-------------------------|------------|------------------------|-----------|-----------|
| 0 | Ajuda | 15 617 | 288 | 38.707500 | -9.198333 |
| 1 | Alcântara | 13 943 | 5,07 | 38.706389 | -9.174167 |
| 2 | Alvalade | 31 813 | 534 | 38.746944 | -9.136111 |
| 3 | Areeiro | 20 131 | 174 | 38.740278 | -9.128056 |
| 4 | Arroios | 31 653 | 213 | 38.728889 | -9.138889 |
| 5 | Avenidas Novas | 21 625 | 299 | 38.738889 | -9.145833 |
| 6 | Beato | 12 737 | 2,46 | 38.734722 | -9.105833 |
| 7 | Belém | 16 528 | 10,43 | 38.700000 | -9.200000 |
| 8 | Benfica | 36 985 | 803 | 38.751111 | -9.202222 |
| 9 | Campo de Ourique | 22 120 | 165 | 38.715278 | -9.166944 |
| 10 | Campolide | 15 460 | 277 | 38.726389 | -9.163333 |
| 11 | Carnide | 19 218 | 369 | 38.760833 | -9.183611 |
| 12 | Estrela | 20 128 | 4,60 | 38.713333 | -9.160000 |
| 13 | Lumiar | 45 605 | 657 | 38.765278 | -9.158611 |
| 14 | Marvila | 37 793 | 7,12 | 38.745278 | -9.104167 |
| 15 | Misericórdia | 13 044 | 2,19 | 38.711389 | -9.148056 |
| 16 | Olivaís | 33 788 | 809 | 38.773611 | -9.117500 |
| 17 | Parque das Nações | 21 025 | 544 | 38.768056 | -9.093889 |
| 18 | Penha de França | 27 967 | 2,71 | 38.730000 | -9.131667 |
| 19 | Santa Clara | 22 480 | 336 | 38.785278 | -9.145000 |
| 20 | Santa Maria Maior | 12 822 | 3,01 | 38.712778 | -9.135556 |
| 21 | Santo António | 11 836 | 149 | 38.724167 | -9.145000 |
| 22 | São Domingos de Benfica | 33 043 | 429 | 38.743611 | -9.170000 |
| 23 | São Vicente | 15 339 | 1,99 | 38.719444 | -9.126389 |

Figure 1: Touristic places of interest in Lisbon using Foursquare API

All the code necessary for obtaining the first dataframe can be found in the file *1_Lisbon_Borough.ipynb*, this file and every file mentioned in this report can be found in the GitHub repository [2]. Having all data from the 24 Lisbon boroughs ready, we started by creating a dataframe dedicated to the restaurant data. The code regarding the dataframe arrangement of restaurants can be seen in the file *2_Lisboa_Restaurant*. Firstly it was needed to have access to the CSV file containing the Lisbon borough data.

Using the data of the boroughs of Lisbon, namely the latitude and longitude variables, a for loop was created to access the Foursquare API and save the data for all the boroughs with places of tourist interest, including hotels, bars, restaurants, historic centers, among others.

| | name | id | categories | lat | lng |
|------|---------------------------|--------------------------|-----------------------|-----------|-----------|
| 0 | Palácio Nacional da Ajuda | 4b0588a3f964a5207bd122e3 | Historic Site | 38.707653 | -9.197758 |
| 1 | Restaurante Andorinhas | 4d9885d59079b1f7a0182d0a | Restaurant | 38.704911 | -9.199349 |
| 2 | Páteo Alfacinha | 4c532ced72cf0f47267c71d2 | Restaurant | 38.706537 | -9.194202 |
| 3 | Jardim Botânico da Ajuda | 4c8b582be51e6dcb8e7671de | Botanical Garden | 38.706430 | -9.201222 |
| 4 | Churrasqueira do Marquês | 4c48033e76d72d7fa2043f4d | BBQ Joint | 38.703996 | -9.199402 |
| ... | ... | ... | ... | ... | ... |
| 1001 | Mercado de Santa Clara | 4e886bc5be7b88449a912b01 | Event Space | 38.715564 | -9.125582 |
| 1002 | Jardim Botto Machado | 4c962b6e82b56dcbd0f9deaa | Garden | 38.715877 | -9.123740 |
| 1003 | Feira da Ladra | 4b0588a8f964a520cfd222e3 | Flea Market | 38.715368 | -9.125244 |
| 1004 | Cafe De Calçada | 54d74df1498ec1066a06efaf | Bistro | 38.718287 | -9.131190 |
| 1005 | Pastelaria Náná | 50952535e4b024b7d7b7e600 | Portuguese Restaurant | 38.718530 | -9.121602 |

Figure 2: Restaunts dataframe

After Having the dataframe with all touristic places, it was needed to filter only the restaurants/bars or any kind of food/drink places. This was a needed step because the focus of this project investment was to open a new restaurant/bar in Lisbon. According to this objective there was no need to analyze the historic places/hotels and other places that do not represent a treat to the opening of the restaurant.

| | Places | ID | Categories | Latitude | Longitude |
|-----|------------------------------|--------------------------|-------------------------------|-----------|-----------|
| 0 | Restaurante Andorinhas | 4d9885d59079b1f7a0182d0a | Restaurant | 38.704911 | -9.199349 |
| 1 | Páteo Alfacinha | 4c532ced72cf0f47267c71d2 | Restaurant | 38.706537 | -9.194202 |
| 2 | Estufa Real | 4b0588a4f964a520ced122e3 | Restaurant | 38.706840 | -9.201975 |
| 3 | Alcântara 50 | 50899fb2e4b0167a9c2eddf4 | Portuguese Restaurant | 38.705462 | -9.173533 |
| 4 | O Palácio | 4c5c82867735c9b6507f8c72 | Seafood Restaurant | 38.706357 | -9.173442 |
| ... | ... | ... | ... | ... | ... |
| 433 | Penalva da Graça | 4f89e174e4b00a6262549ad1 | Seafood Restaurant | 38.720722 | -9.130070 |
| 434 | Taproom Oitava Colina | 5b4928789f8a9f002c28cc08 | Beer Bar | 38.718390 | -9.131880 |
| 435 | O Cardoso do Estrela de Ouro | 4c892b94a0ffb60c7f4228c5 | Portuguese Restaurant | 38.720650 | -9.130091 |
| 436 | Tazza In Giro | 5a09ef5d2619ee11bd25fffc | Vegetarian / Vegan Restaurant | 38.715800 | -9.125121 |
| 437 | Pastelaria Náná | 50952535e4b024b7d7b7e600 | Portuguese Restaurant | 38.718530 | -9.121602 |

Figure 3: Final dataframe regarding the places of interest in Lisbon.

The thought process behind this is that likes are a proxy for quality. The more likes there are, the better the restaurant is. This might be incorrect but API call issues (how many I can use for free) holds me back from getting price / rating data. I will then bin this data into a quality categorical variables so we can cluster appropriately. Having this into account to accomplish this project, this solutions seemed the most logical for having a comparative method.

| | Places | ID | Categories | Latitude | Longitude | Likes |
|-----|------------------------------|--------------------------|-------------------------------|-----------|-----------|-------|
| 0 | Restaurante Andorinhas | 4d9885d59079b1f7a0182d0a | Restaurant | 38.704911 | -9.199349 | 23 |
| 1 | Páteo Alfacinha | 4c532ced72cf0f47267c71d2 | Restaurant | 38.706537 | -9.194202 | 44 |
| 2 | Estufa Real | 4b0588a4f964a520ced122e3 | Restaurant | 38.706840 | -9.201975 | 25 |
| 3 | Alcântara 50 | 50899fb2e4b0167a9c2eddf4 | Portuguese Restaurant | 38.705462 | -9.173533 | 27 |
| 4 | O Palácio | 4c5c82867735c9b6507f8c72 | Seafood Restaurant | 38.706357 | -9.173442 | 86 |
| ... | ... | ... | ... | ... | ... | ... |
| 433 | Penalva da Graça | 4f89e174e4b00a6262549ad1 | Seafood Restaurant | 38.720722 | -9.130070 | 12 |
| 434 | Taproom Oitava Colina | 5b4928789f8a9f002c28cc08 | Beer Bar | 38.718390 | -9.131880 | 12 |
| 435 | O Cardoso do Estrela de Ouro | 4c892b94a0ffb60c7f4228c5 | Portuguese Restaurant | 38.720650 | -9.130091 | 9 |
| 436 | Tazza In Giro | 5a09ef5d2619ee11bd25fffc | Vegetarian / Vegan Restaurant | 38.715800 | -9.125121 | 6 |
| 437 | Pastelaria Náná | 50952535e4b024b7d7b7e600 | Portuguese Restaurant | 38.718530 | -9.121602 | 3 |

438 rows × 6 columns

Figure 4: Borough Table

2.3 Mapping of Data

The data will be used as follows: Use Foursquare and geopy data to map the 24 boroughs for all Lisbon and use foursquare and geopy data to map the location of restaurants and places of interest for the business. After this cluster all data in a map in order to be able to identify the neighbourhoods each location separately. This will allow to have a simplified version of all the data. Addresses from restaurant locations will be converted to geodata (lat, long) and will be used in the *folium.Map* in order to the examination and analysis easier.

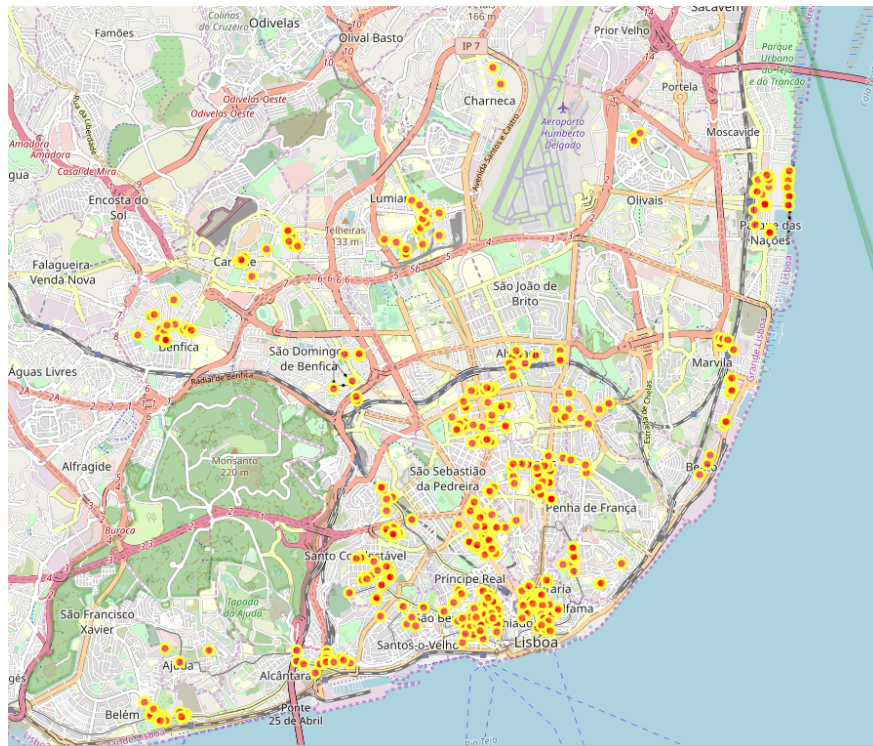


Figure 5: Lisbon Map with places of interest

3 METHODOLOGY SECTION

3.1 Strategy and Exploratory analysis

After getting the data sets "clean", it was created some data visualizations in order to better interpret and study the results. Initially it was analysed the number of citizens per borough was carried out in order to be aware of the influence that this factor may have on local trade. The graph bar in fig.6 shows the population for each of the boroughs in the dataframe.

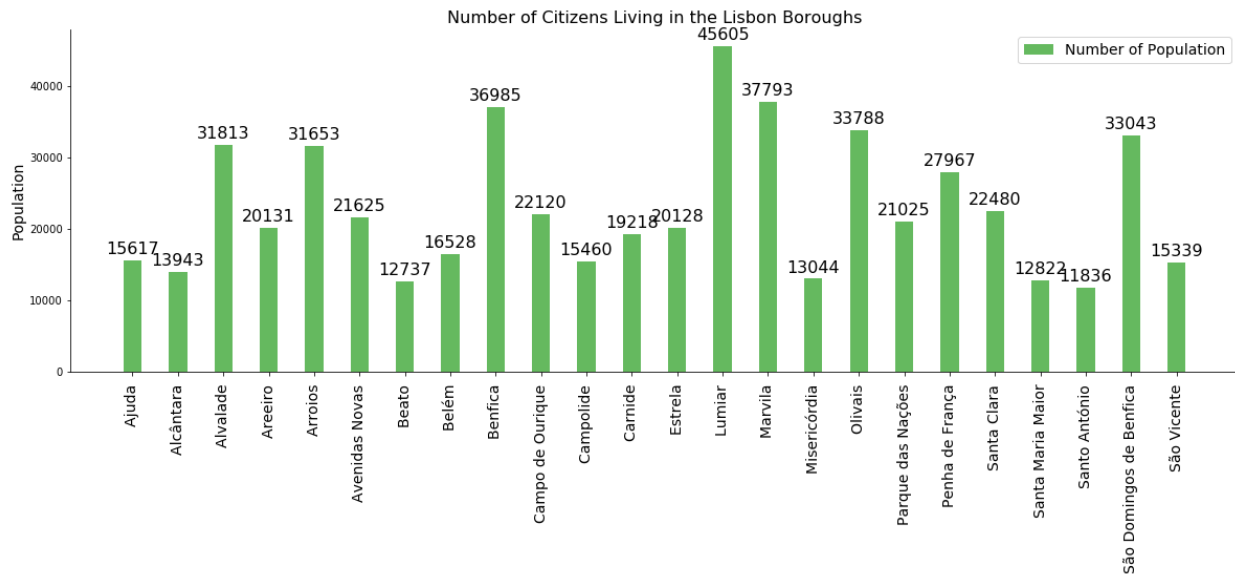


Figure 6: Number of Citizens Living in each Lisbon borough

After knowing the population of each borough, it is important to relate this number to the number of restaurants and bars in each borough. For this purpose, the bar graph shown in figure 7 was created. It is possible to verify that there is no reliable relationship between the number of inhabitants and the number of restaurants. At most the correlation that can be made is that the areas with the largest housing tend to have a number of medium / low restaurants.

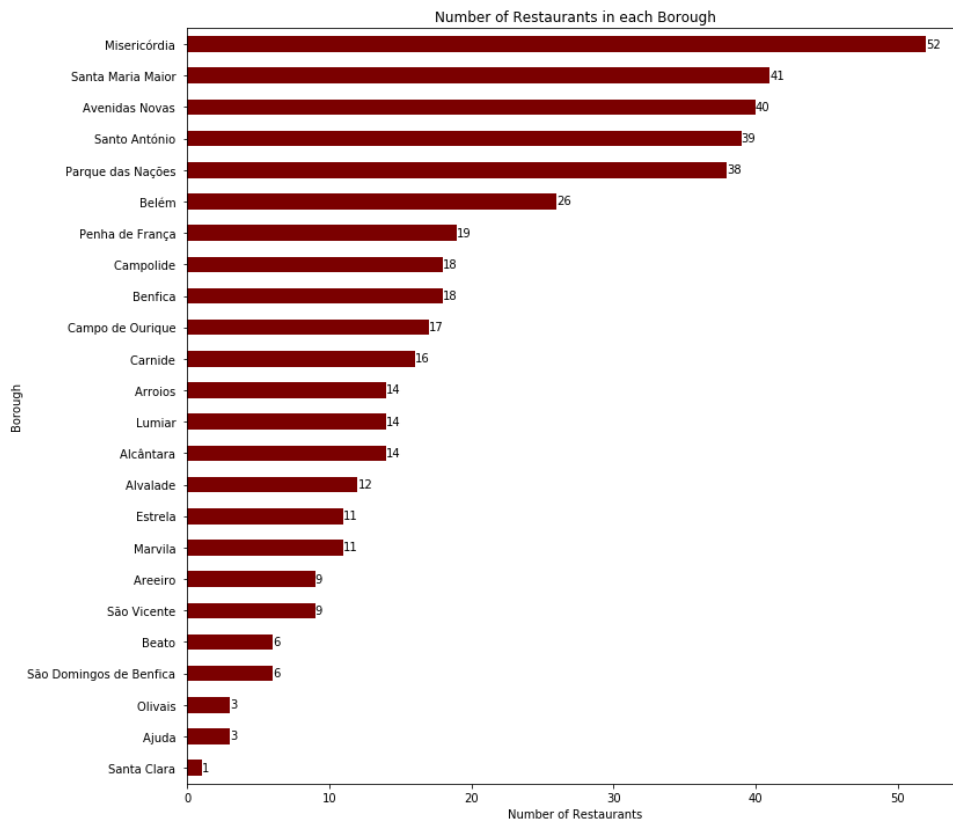


Figure 7: Number of Restaurants/Bars in each Lisbon borough

Taking into account an overview of all of Lisbon, the 15 categories with the most established restaurants were filtered. Through figure 8 it can be seen that clearly stand out the "Restaurants" which are mainly those traditional restaurants with a high variety, and in the first place the Restaurants specialized in Portuguese food.

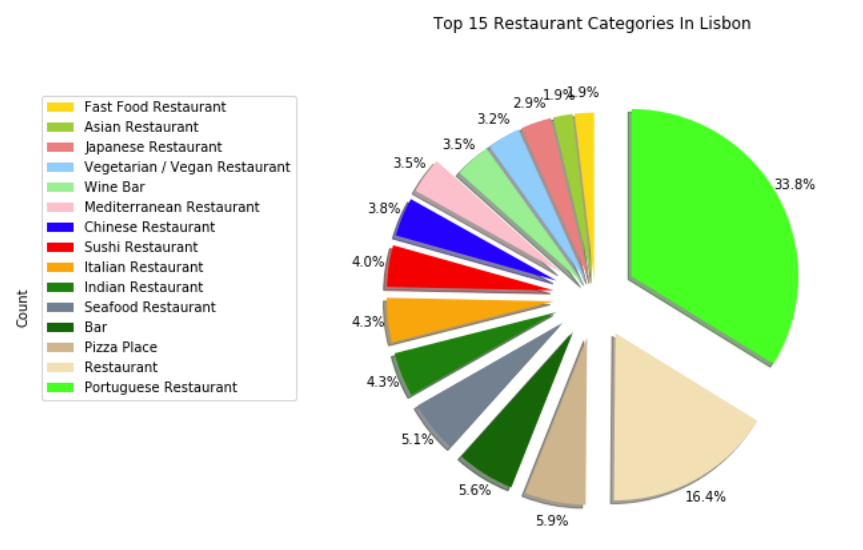


Figure 8: Lisbon Map with places of interest

It is possible to verify through figure 10 that the red data point has the largest cluster of restaurants, with 133 elements. As such, it has been analyzed in greater detail in Figure 18 that there is enormous competition especially in Campo Pequeno, Arroios and the Marquês de Pombal areas.

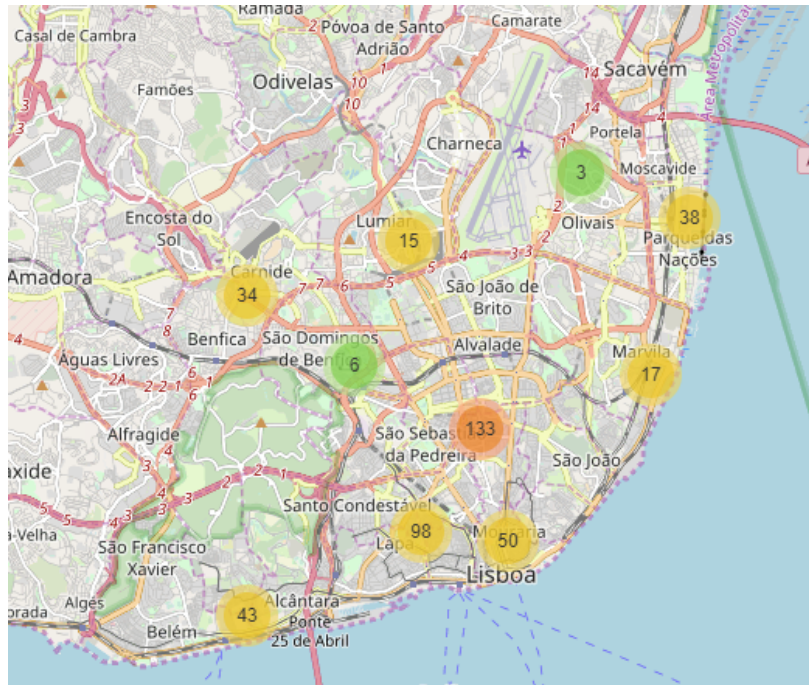


Figure 11: General Cluster Map of Lisbon Restaurants

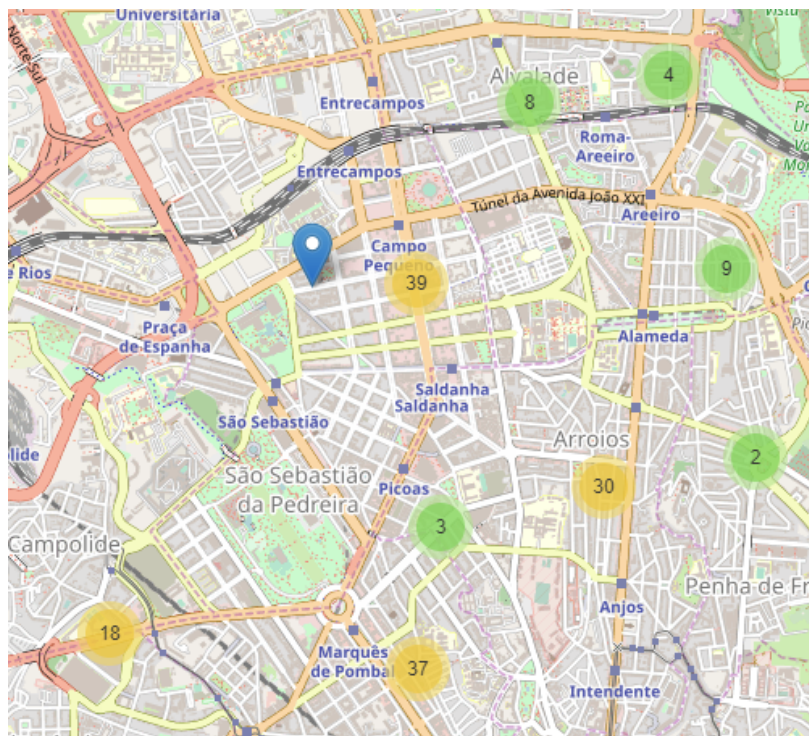


Figure 12: Cluster Map of restaurants red zone.

3.2 K-Cluster analysis

Subsequently, K-Means Clustering was performed the Lisbon boroughs according to which restaurant category used from Foursquare information to experience each borough's opportunity. A hot encoding was performed as the first step of cluster analysis to give binary values to each restaurant category. After that, borough names clustered the data to find out how many category of each class occur within each of the boroughs. We could obtain a list of the most common venue categories in each borough based on the frequency as follows.

| | Places | Borough | Likes | Latitude | Longitude | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|---|------------------------|-----------|-------|-----------|-----------|----------------|--------------------------|--------------------------|-----------------------------|-----------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| 0 | Restaurante Andorinhas | Ajuda | 23 | 38.704911 | -9.199349 | 1 | Restaurant | Wine Bar | Eastern European Restaurant | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food | Fast Food Restaurant |
| 1 | Páteo Alfacinha | Ajuda | 44 | 38.706537 | -9.194202 | 1 | Restaurant | Wine Bar | Eastern European Restaurant | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food | Fast Food Restaurant |
| 2 | Estufa Real | Ajuda | 25 | 38.706840 | -9.201975 | 1 | Restaurant | Wine Bar | Eastern European Restaurant | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food | Fast Food Restaurant |
| 3 | Alcântara 50 | Alcântara | 27 | 38.705462 | -9.173533 | 2 | Portuguese Restaurant | Mediterranean Restaurant | Seafood Restaurant | Eastern European Restaurant | Pizza Place | Restaurant | Indian Restaurant | Beer Bar | Sushi Restaurant | Beer Garden |
| 4 | O Palácio | Alcântara | 86 | 38.706357 | -9.173442 | 2 | Portuguese Restaurant | Mediterranean Restaurant | Seafood Restaurant | Eastern European Restaurant | Pizza Place | Restaurant | Indian Restaurant | Beer Bar | Sushi Restaurant | Beer Garden |

Figure 13: Cluster initial dataframe.

4 RESULTS SECTION

Regarding the K-Means clustering methodology used the results can be seen below:

| | Borough | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|-----|-------------------------|----------------|--------------------------|--------------------------|--------------------------|-------------------------------|--------------------------|-------------------------------|--------------------------|-----------------------------|--------------------------|-----------------------------|
| 52 | Avenidas Novas | 0 | Portuguese Restaurant | Restaurant | Italian Restaurant | Vegetarian / Vegan Restaurant | Pizza Place | Asian Restaurant | Fast Food Restaurant | Sushi Restaurant | Brazilian Restaurant | Japanese Restaurant |
| 124 | Benfica | 0 | Portuguese Restaurant | Seafood Restaurant | Restaurant | Asian Restaurant | Sushi Restaurant | Food | Chinese Restaurant | Pizza Place | Dim Sum Restaurant | French Restaurant |
| 142 | Campo de Ourique | 0 | Portuguese Restaurant | Bar | Restaurant | Seafood Restaurant | Italian Restaurant | Japanese Restaurant | Pizza Place | Indian Restaurant | Cantonese Restaurant | Brazilian Restaurant |
| 159 | Campolide | 0 | Restaurant | Portuguese Restaurant | Seafood Restaurant | French Restaurant | Japanese Restaurant | Fast Food Restaurant | Falafel Restaurant | Dim Sum Restaurant | Pizza Place | Health Food Store |
| 177 | Carnide | 0 | Portuguese Restaurant | Sushi Restaurant | Restaurant | Dim Sum Restaurant | Tapas Restaurant | Asian Restaurant | Food | Mediterranean Restaurant | Pizza Place | Food Truck |
| 204 | Lumiar | 0 | Pizza Place | Japanese Restaurant | Fast Food Restaurant | Chinese Restaurant | Restaurant | Vegetarian / Vegan Restaurant | Italian Restaurant | Sushi Restaurant | Wine Bar | Eastern European Restaurant |
| 218 | Marvila | 0 | Restaurant | Portuguese Restaurant | Pizza Place | Argentinian Restaurant | Mediterranean Restaurant | Wine Bar | French Restaurant | Food Truck | Food Service | Food |
| 284 | Parque das Nações | 0 | Portuguese Restaurant | Restaurant | Sushi Restaurant | Chinese Restaurant | Seafood Restaurant | Bar | Hotel Bar | Italian Restaurant | Pizza Place | Falafel Restaurant |
| 422 | São Domingos de Benfica | 0 | Bar | Japanese Restaurant | Food Truck | Seafood Restaurant | Fast Food Restaurant | Portuguese Restaurant | Wine Bar | Eastern European Restaurant | Gay Bar | French Restaurant |

Figure 14: Cluster 1.

| | Borough | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|-----|---------|----------------|--------------------------|--------------------------|-----------------------------|-----------------------------|-----------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| 0 | Ajuda | 1 | Restaurant | Wine Bar | Eastern European Restaurant | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food | Fast Food Restaurant |
| 92 | Beato | 1 | Restaurant | Tapas Restaurant | Cantonese Restaurant | Wine Bar | Eastern European Restaurant | Gay Bar | French Restaurant | Food Truck | Food Service | Food |
| 281 | Olivais | 1 | Restaurant | Chinese Restaurant | Wine Bar | Eastern European Restaurant | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food |

Figure 15: Cluster 2.

| | Borough | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|-----|-------------------|----------------|--------------------------|--------------------------|-------------------------------|-------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|-------------------------------|
| 3 | Alcântara | 2 | Portuguese Restaurant | Mediterranean Restaurant | Seafood Restaurant | Eastern European Restaurant | Pizza Place | Restaurant | Indian Restaurant | Beer Bar | Sushi Restaurant | Beer Garden |
| 38 | Arroios | 2 | Portuguese Restaurant | Indian Restaurant | Vegetarian / Vegan Restaurant | Italian Restaurant | Brazilian Restaurant | Chinese Restaurant | Restaurant | Mediterranean Restaurant | Argentinian Restaurant | Falafel Restaurant |
| 229 | Misericórdia | 2 | Portuguese Restaurant | Bar | Wine Bar | Cocktail Bar | Restaurant | Italian Restaurant | Juice Bar | Pizza Place | Brazilian Restaurant | French Restaurant |
| 342 | Santa Maria Maior | 2 | Portuguese Restaurant | Wine Bar | Restaurant | Indian Restaurant | Bar | Mediterranean Restaurant | African Restaurant | Ramen Restaurant | Food Service | Vegetarian / Vegan Restaurant |
| 428 | São Vicente | 2 | Mediterranean Restaurant | Portuguese Restaurant | Indian Restaurant | Vegetarian / Vegan Restaurant | Bar | Beer Bar | Seafood Restaurant | Empanada Restaurant | French Restaurant | Food Truck |

Figure 16: Cluster 3.

| | Borough | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|-----|-------------|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| 341 | Santa Clara | 3 | Portuguese Restaurant | Wine Bar | Hotel Bar | Health Food Store | Gay Bar | French Restaurant | Food Truck | Food Service | Food | Fast Food Restaurant |

Figure 17: Cluster 4.

| | Borough | Cluster Labels | 1st Most Common Category | 2nd Most Common Category | 3rd Most Common Category | 4th Most Common Category | 5th Most Common Category | 6th Most Common Category | 7th Most Common Category | 8th Most Common Category | 9th Most Common Category | 10th Most Common Category |
|-----|-----------------|----------------|--------------------------|--------------------------|-------------------------------|--------------------------|---------------------------|--------------------------|--------------------------|-----------------------------|-----------------------------|---------------------------|
| 17 | Alvalade | 4 | Portuguese Restaurant | Bar | Persian Restaurant | Pizza Place | Snack Place | Indian Restaurant | Thai Restaurant | Beer Bar | Eastern European Restaurant | French Restaurant |
| 29 | Areeiro | 4 | Portuguese Restaurant | Asian Restaurant | Italian Restaurant | Chinese Restaurant | Restaurant | Pizza Place | Wine Bar | Eastern European Restaurant | French Restaurant | Food Truck |
| 98 | Belém | 4 | Portuguese Restaurant | Restaurant | Pizza Place | Chinese Restaurant | Fast Food Restaurant | Italian Restaurant | Japanese Restaurant | Juice Bar | Mediterranean Restaurant | French Restaurant |
| 193 | Estrela | 4 | Portuguese Restaurant | Pizza Place | Vegetarian / Vegan Restaurant | Bar | Japanese Restaurant | Restaurant | French Restaurant | Food Truck | Food Service | Food |
| 322 | Penha de França | 4 | Portuguese Restaurant | Snack Place | Chinese Restaurant | Indian Restaurant | Asian Restaurant | Italian Restaurant | Restaurant | Pizza Place | Middle Eastern Restaurant | Bar |
| 383 | Santo António | 4 | Portuguese Restaurant | Restaurant | Pizza Place | Russian Restaurant | Middle Eastern Restaurant | Italian Restaurant | Japanese Restaurant | Brazilian Restaurant | Himalayan Restaurant | Mediterranean Restaurant |

Figure 18: Cluster 5.

Based on our initial parameters, we were able to discover the best neighborhood after various analyzes. Now we're going to review all the analyzes completed in this project before and finally come to the conclusions about starting up a business as an entrepreneur.

5 DISCUSSION SECTION AND CONCLUSIONS

We find from this study that the five boroughs below are the best places to build a student building, based on the borough's opportunity of business. One option would be to open a restaurant dedicated to Portuguese food in the Olivais region of cluster 2, because despite the tourist trend being "Portuguese Restaurant", this borough has no competition at this level which makes it stand out as a Portuguese restaurant. On the other hand, it is an area with a considered population of 33788, which allows it not to depend exclusively on tourists, so this region stands out from the rest of cluster 2. Another opportunity would be to open, for example, a pizzeria in Santa Clara, cluster 4. Because it has no such venture it is one of the most prominent tourist categories. In addition to this I would like to highlight, the reward with the efforts, time and money spent. I believe this course with all the topics covered is well worthy of appreciation. This project showed me a practical application to solve the real situation with the use of Data Science tools that have an impact on personal and financial impact. Folium mapping is a very powerful technique for consolidating information and thoroughly and confidently making the analysis and decision. I'd recommend using it in similar situations.

6 LIMITATIONS AND RECOMMENDATION FOR FUTURE STUDY

We only took into consideration which sort of restaurant categories are common and how many likes that restaurant and category has, among various factors that determine a good borough. A higher number of influencing factors might give a more precise algorithm as well as a more perfectionist analysis. Apart from that, it is also needed to mentioned that all project was done only using the default/free access to Foursquare what might be a limiting factor in terms of data access.

REFERENCES

- [1] Lista de freguesias de lisboa. URL https://pt.wikipedia.org/wiki/Lista_de_freguesias_de_Lisboa.
- [2] Capstone repository. URL https://github.com/Danielorasleite/Coursera_Capstone.