The Battle of Neighborhoods

Coursera Capstone Project

Explore restaurants in Lisbon's Famous Viewpoints

A work by
Beatriz Barros

1. Introduction

This final project explores the famous Viewpoints that exit in Lisbon, regarding the type of restaurants. The goal is to understand the number and type of restaurants present at this specific locations – an information that can be helpful to business personnel who wants to invest or open a restaurant or for a turist looking for the best places to visit.

Lisbon is the stunning capital city of Portugal, and is one of the most charismatic and vibrant cities of Europe, with an estimated population of 505,526 within its administrative limits in an area of 100.05 km. Lisbon's superb natural setting, spread across seven hills facing the Tagus River, offers a network of terraces from which to contemplate the beauty of the city, that are perfect spots to enjoy the sunset. For that reason, rooftop bars and restaurants are very common and popular at this locations with diverse types of food offer since Portuguese food, Mediterranean, Tapas, among others.

1.1. Business problem

The aim of this project is to understand the spatial distribution of the official Viewpoints of Lisbon, and the number and type of restaurants that exist at each one of them in order to understand which locations would be better to open a restaurant or to plan a trip, and what type of food is popular.

2. Target Audience

The target of audience for this problem includes:

1. Business entrepreneur who wants to invest or open a restaurant.

2. Tourist who wants to understand the best Viewpoints to visit.

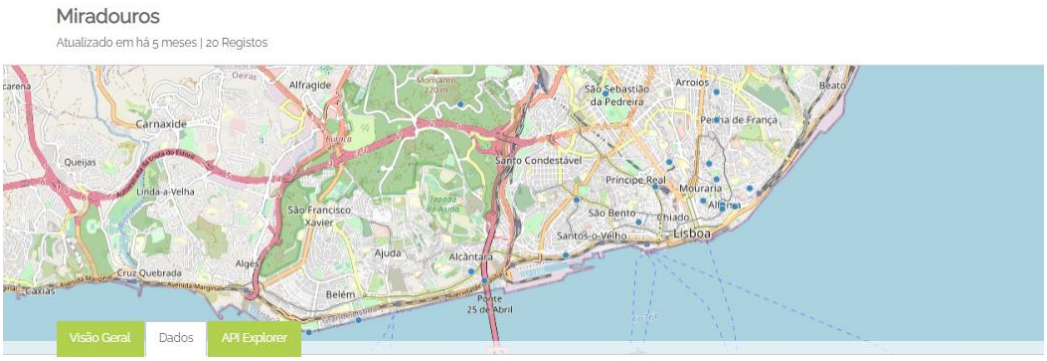
3. A local person who wants to know the best locations to enjoy a meal at the end of the day or schedule an event.

4. A Business Analyst or Data scientist who want to implement exploratory Data Analysis techniques to obtain and analyze data to solve a problem.

3. Data section

To solve this problem, the first data required is a dataset of the Viewpoints of Lisbon, regarding its Name and Location. Since there are numerous Viewpoints known, but not all of them are considered official, the data source must be chosen carefully.

Lisbon's City Council website (https://www.lisboa.pt) is the first source that should be consulted. Luckily, it contains the information that is required. This is the page dedicated to Viewpoints:



30/09/2020 Camada de Elementos Personalizar Licença

Transferência APIs

A exibir 1 a 10 de 20

Sugestão: filtre as colunas usando

OBJECTID	COD_SIG	IDTIPO	CAT_NOME	CAT_ID	INF_ID	INF_NOME	INF_MORADA
1	0805531053001	2	Espaços Verdes e Lazer - Miradou..	211819	12514	Miradouro do Panorâmico de Mon..	Estrada Bela Vista
2	0202512431001	2	Espaços Verdes e Lazer - Miradou..	211819	12522	Experiência Pilar 7	Avenida da Índia, Por
3	16	7	Espaços Verdes e Lazer - Miradou..	211819	3232	Miradouro de Santo Estêvão	Largo de Santo Estêv
4	41	7	Espaços Verdes e Lazer - Miradou..	211819	1948	Miradouro do Castelo de São Jorge	Castelo de São Jorge
5	20	7	Espaços Verdes e Lazer - Miradou..	211819	1958	Miradouro Sophia de Mello Breyn..	Largo da Graça
6	23	7	Espaços Verdes e Lazer - Miradou..	211819	1960	Miradouro de Santa Clara	Campo de Santa Clar
7	32	7	Espaços Verdes e Lazer - Miradou..	211819	3233	Miradouro da Basílica da Estrela	Praça da Estrela 120c
8	4	7	Espaços Verdes e Lazer - Miradou..	211819	1965	Miradouro do Torel	Rua Júlio de Andrade
9	21	7	Espaços Verdes e Lazer - Miradou..	211819	1959	Miradouro das Portas do Sol	Largo das Portas do S
10	11	7	Espaços Verdes e Lazer - Miradou..	211819	1953	Miradouro do Padrão dos Descobr..	Avenida de Brasília

https://geodados-cml.hub.arcgis.com/datasets/miradouros/data?geometry=-9.366%2C38.688%2C-8.975%2C38.735

The table contains the following information for each Viewpoint:

- ID
- Name
- Address
- Phone Number
- Fax
- Email
- Website
- Description
- Municipal Information
- Coordinates

Since the essential information for this project is Name and Location, this dataset is appropried. So, the next step is to scrape this dataset. After using BeautifulSoup to scrape information from web pages, create a Dataframe and populate it with the content of the website, a complete dataframe is obtained.

Scraping the dataset from the Web Page of Lisbon's City Council

In [3]:
```
list_url = "https://geodados-cml.hub.arcgis.com/datasets/miradouros/data"
source = requests.get(list_url).text
soup = BeautifulSoup(source, 'xml')
table=soup.find('table')
```

In [4]:
```
#Create dataframe
column_names = ['OBJECTID','COD_SIG','IDTIPO','CAT_NOME','CAT_ID','INF_ID','INF_NOME','INF_MORADA','INF_TELEFONE','INF_FAX','INF_EMAIL','INF_SITE','INF_DESCRICAO','INF_AUTOR_DESCRICAO','INF_FONTE','INF_OBS','INF_ACTIVO','INF_MUNICIPAL','Globa
df = pd.DataFrame(columns = column_names)
```

In [5]:
```
for tr_cell in table.find_all('tr'):
    row_data=[]
    for td_cell in tr_cell.find_all('td'):
        row_data.append(td_cell.text.strip())
    if len(row_data)==3:
        df.loc[len(df)] = row_data

df
```

	OBJECTID	COD_SIG	IDTIPO	CAT_NOME	CAT_ID	INF_ID	INF_NOME	INF_MORADA	INF_TELEFONE	INF_FAX	INF_EMAIL	INF_SITE	INF_DESCRICAO	INF_AUTOR_DESCRICAO	INF_FONTE	INF_OBS	INF_ACTIVO	INF_MUNICIPAL
0	1	0805531053001	2	Espaços Verdes e Lazer - Miradouros	211819	12514	Miradouro do Panorâmico de Monsanto	Estrada Bela Vista	None	None	None	None	Localizado no Parque Florestal de Monsanto, na...	None	DMEVAE	None	1	0
1	2	0202512431001	2	Espaços Verdes e Lazer - Miradouros	211819	12522	Experiência Pilar 7	Avenida da Índia, Ponte 25 de Abril - Pilar 7	+351 211117880	None	info@lismarketing.pt	www.visitlisboa.com	None	None	www.visitlisboa.com	None	1	0
2	3	16	7	Espaços Verdes e Lazer - Miradouros	211819	3232	Miradouro de Santo Estêvão	Largo de Santo Estêvão	+351 213 912 600	None	daev@cm-lisboa.pt	None	Situado no coração de Alfama, do adro da Igrej...	None	-	SG/DMC	1	0
3	4	41	7	Espaços Verdes e Lazer - Miradouros	211819	1948	Miradouro do Castelo de São Jorge	Castelo de São Jorge	+351 218 800 620	None	info@castelodesaojorge.pt	www.castelodesaojorge.pt	O Castelo de São Jorge oferece uma vista privi...	Isabel Santos	-	None	1	0
4	5	20	7	Espaços Verdes e Lazer - Miradouros	211819	1958	Miradouro Sophia de Mello Breyner Andresen (Mi...	Largo da Graça	(351) 213 912 600	None	daev@cm-lisboa.pt	None	Localizado junto à Igreja e antigo Convento da...	Oásis Alfacinhas: Guia Ambiental de Lisboa - I...	-	None	1	1
5	6	23	7	Espaços Verdes e Lazer - Miradouros	211819	1960	Miradouro de Santa Clara	Campo de Santa Clara	+351 218 533 270	None	daev@cm-lisboa.pt	None	É um ponto alto da cidade com vista panorâmica...	None	-	SG/DMC	1	0
6	7	32	7	Espaços Verdes e Lazer - Miradouros	211819	3233	Miradouro da Basílica da Estrela	Praça da Estrela\n1200-667 LISBOA	+351 213 960 915	None	None	None	Do zimbório desfruta-se de bela vista: a oeste...	None	-	SG-DMC	1	0
7	8	4	7	Espaços Verdes e Lazer - Miradouros	211819	1965	Miradouro do Torel	Rua Júlio de Andrade	None	None	None	None	Deste miradouro disfruta-se de uma vista desaf...	None	-	SG/DMC	1	0
8	9	21	7	Espaços Verdes e Lazer - Miradouros	211819	1959	Miradouro das Portas do Sol	Largo das Portas do Sol\n	+351 213 912 600	None	daev@cm-lisboa.pt	None	Localizado no Largo das Portas do Sol, daí o s...	http://revelartx.cm-lisboa.pt, www.guiadacidad...	-	None	1	1
9	10	11	7	Espaços Verdes e Lazer - Miradouros	211819	1953	Miradouro do Padrão dos Descobrimentos	Avenida de Brasília\n	+351 213 031 950	None	info@padraodosdescobrimentos.pt	www.padraodosdescobrimentos.pt	No terraço existe um espetacular miradouro, do...	None	-	None	1	1

In [11]:
```
df.shape
```

Out[11]: (20, 19)

Since the website also provides a shape file (GeoJson), the table can also be extracted just by using the link:

```
Lisbon_viewpoints_geodf = gpd.read_file('https://opendata.arcgis.com/datasets/02763aa9ec6b4d719841f713e682b12a_1.geojson')
df = Lisbon_viewpoints_geodf
```

Let's then use this table since it additionally contains the coordinates of each Viewpoint:

NICIPAL	GlobalID	geometry
0	6eed37fd-2b1c-4d74-bf46-651b8b94737d	POINT (-9.18461 38.72862)
0	783577af-9715-49a8-a797-b6cef6a60d69	POINT (-9.17955 38.70001)
0	51166c6a-a6bf-4d6a-9fdf-b2490dc737b6	POINT (-9.12773 38.71207)
0	8ced2062-5a41-479a-927f-d17cadd6352a	POINT (-9.13421 38.71299)

Now the complete dataset contains many different information. However, for this problem, the only information needed is Name, Address and Coordinates, so we can create a second dataframe with only this information. Let's also change the columns name for it to be easily understood, and create two new columns for Latitude and Longitude, extracting the values from the "geometry" column:

```
from shapely.wkt import loads as load_wkt

centroid_list = []

for polygon in df_lisbonviews["Coordinates"]:
    box_str = str(polygon)
    p1 = load_wkt(box_str)
    point = p1.centroid
#     print(type(p1.centroid.x))
#     print(p1.centroid.y)
    centroid_list.append((p1.centroid.y, p1.centroid.x))

lat_centr, lon_centr = zip(*centroid_list)

df_lisbonviews['Latitude'] = lat_centr
df_lisbonviews['Longitude'] = lon_centr

df_lisbonviews.head(10)
```

The final dataset:

	Viewpoint	Adress	Latitude	Longitude
0	Miradouro do Panorâmico de Monsanto	Estrada Bela Vista	38.728618	-9.184607
1	Experiência Pilar 7	Avenida da Índia, Ponte 25 de Abril - Pilar 7	38.700008	-9.179553
2	Miradouro de Santo Estêvão	Largo de Santo Estêvão	38.712070	-9.127729
3	Miradouro do Castelo de São Jorge	Castelo de São Jorge	38.712991	-9.134210
4	Miradouro Sophia de Mello Breyner Andresen (Mi...	Largo da Graça	38.716427	-9.131584
5	Miradouro de Santa Clara	Campo de Santa Clara	38.714970	-9.124694
6	Miradouro da Basílica da Estrela	Praça da Estrela\n1200-667 LISBOA	38.713097	-9.160581
7	Miradouro do Torel	Rua Júlio de Andradel	38.719323	-9.141286
8	Miradouro das Portas do Sol	Largo das Portas do Sol\n	38.712346	-9.130366
9	Miradouro do Padrão dos Descobrimentos	Avenida de Brasília\n	38.693614	-9.205697
10	Miradouro da Rocha de Conde de Óbidos	Rua Presidente Arriaga;	38.704181	-9.162765
11	Miradouro da Penha de França	Largo da Penha de França\n1170-298 LISBOA	38.730633	-9.131550
12	Miradouro de Santa Luzia	Largo de Santa Luzia\n	38.711646	-9.130360
13	Miradouro de Santo Amaro	Calçada de Santo Amaro	38.701464	-9.182371
14	Miradouro do Parque Eduardo VII	Alameda Cardeal Cerejeira\n	38.730358	-9.154480
15	Miradouro de São Pedro de Alcântara	Rua de São Pedro de Alcântara	38.715299	-9.144277
16	Miradouro do Monte Agudo	Rua Heliodoro Salgado	38.726145	-9.131528
17	Miradouro da Torre de Belém	Avenida de Brasília\n1400-038	38.691578	-9.215928
18	Miradouro de Santa Catarina	Rua de Santa Catarina\n1200-403 LISBOA	38.709515	-9.147672
19	Miradouro da Senhora do Monte	Largo do Monte\n	38.718975	-9.132994

3. Approach

With this clear and simple dataset, is possible to visually analyze the locations of each viewpoint and explore, through the Foursquare API, the venues near this locations. Therefore, the following approach will be realized:

- Visualize the distribution of Viewpoints using Folium
- Use Foursquare API to get all venues for each Viewpoint
- Map the distribution of venues around Viewpoints (Folium)
- Analysis to find the most common types of restaurants
- Analyzing using Clustering (K-Means)
- Compare the different clusters in terms of Location and Number/Type of restaurants
- Inference From these Results and related Conclusions