Exploring for potential new italian restaurant in Bilbao

Data Science Capstone Project

Table of Contents

Introduction

- **BACKGROUND**
- Planning to start a new italian restaurant in the city of Bilbao.
- The real tourist surge though would come much later with the inauguration of the Guggenheim Museum Bilbao in 1997. Thereafter tourist arrivals registered a continued upward trend, reaching over 932,000 visitors in 2018. The trend was exponential considering that in 1995, Bilbao only counted 25,000 tourists.
- It provides a suggestion on what would be the best venue to start a new italian restaurant in a highly visited city with already many good restaurants.
- We will go through the benefits and pitfalls of opening a new italian restaurant in a highly visited city with already many restaurants offering their services. The core of Bilbao is made of eight districts.

Introduction

Introduction

- b. Business Problem
  - This report focusses on the issue of where to open a new italian restaurant in a city like Bilbao, once one has decided to go ahead. Let's imagine an investment company wants to open a new luxury hotel, a first and foremost important decision will be the location for its new italian restaurant.
- c. Interest
  - • On what basis can the company decide its new italian restaurant location?
  - • While selecting the place there are key points to consider like they need to check out like where the most well-visited venues of the city are?
  - • If incase there are already other italian restaurant which have good ratings, will it be risky to open new one near these italian restaurants?

Data

- the factors that will influence the final decision are: -
  - • Number of existing restaurants an coffees in the neighbourhood (any type of restaurant)
  - • Number of and distance to Italian restaurants in the neighbourhood
  - • The tourist places
- The following data sources will be needed to extract/generate the required information:
  - • List of all districts in Bilbao - https://en.wikipedia.org/wiki/Bilbao#Tourism
  - • Coordinates of all districts and venues - GeoPy Nominatim geocoding
  - • Number of restaurants and their type and location in every neighbourhood - Foursquare API - https://developer.foursquare.com

Methodology

- 1º- Collect data on the districts of Bilbao from the internet.
- 2º- Explore each of district and their venues using Foursquare location data. The venues of the districts will be analyzed in detail and patterns will be discovered. This discovery of patterns will be carried out by grouping the districts using k-means clustering. Following this, each cluster will be examined and a decision will be made regarding which cluster fits the shareholder's requirements. The factor that will determine this is the frequency of occurrence of restaurants and other food venues within the cluster.
- 3º -Once a cluster is picked, the districts in that cluster will be investigated with regards to the number of Italian restaurants in its vicinity. The ones that fit the requirements will be further explored and shortlisted based on how small their respective distances to the center of Bilbao are.

Data preparation

a. Scrapping Bilbao Districts Table from Wikipedia

	District	Area km²	Population	Latitude	Longitude
0	Deusto	4.95	51,656	43.271387	-2.949581
1	Uribarri	4.19	38,335	43.266839	-2.919947
2	Otxarkoaga-Txurdinaga	3.90	28,518	43.260369	-2.896813
3	Begoña	1.77	43,030	43.258890	-2.916111
4	Ibaiondo	9.65	61,029	43.256900	-2.923331
5	Abando	2.14	51,718	43.262788	-2.934947
6	Errekalde	6.96	47,787	43.251944	-2.946111
7	Basurto-Zorroza	7.09	33,658	43.268100	-2.965600

Data preparation

b. Bilbao District Map

Data Analysis

Exploring Districts in Bilbao

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Deusto	43.271387	-2.949581	Munich	43.271183	-2.944475	Burger Joint
1	Deusto	43.271387	-2.949581	La Tortilla	43.269937	-2.950183	Spanish Restaurant
2	Deusto	43.271387	-2.949581	La Tapa	43.270046	-2.946396	Café
3	Deusto	43.271387	-2.949581	Restaurante Bordatxo	43.272187	-2.944992	Spanish Restaurant
4	Deusto	43.271387	-2.949581	Deustoarrak	43.271166	-2.945160	Restaurant
5	Deusto	43.271387	-2.949581	Plaza San Pedro de Deusto	43.271936	-2.945661	Plaza
6	Deusto	43.271387	-2.949581	Pulperia Naroa	43.270235	-2.953144	Restaurant
7	Deusto	43.271387	-2.949581	Bertiz	43.271205	-2.946872	Bakery
8	Deusto	43.271387	-2.949581	Telepizza	43.271436	-2.949716	Pizza Place
9	Deusto	43.271387	-2.949581	Moon Hostel Bio	43.272387	-2.946787	Boarding House
10	Deusto	43.271387	-2.949581	Irish Stones	43.271127	-2.945298	Pub

Data Analysis

District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Abando	100	100	100	100	100	100
Basurto-Zorroza	5	5	5	5	5	5
Begoña	25	25	25	25	25	25
Deusto	19	19	19	19	19	19
Errekalde	4	4	4	4	4	4
Ibaiondo	56	56	56	56	56	56
Otxarkoaga-Txurdinaga	2	2	2	2	2	2
Uribarri	7	7	7	7	7	7

Data Analysis

Printing districts along with the top 12 most common venues

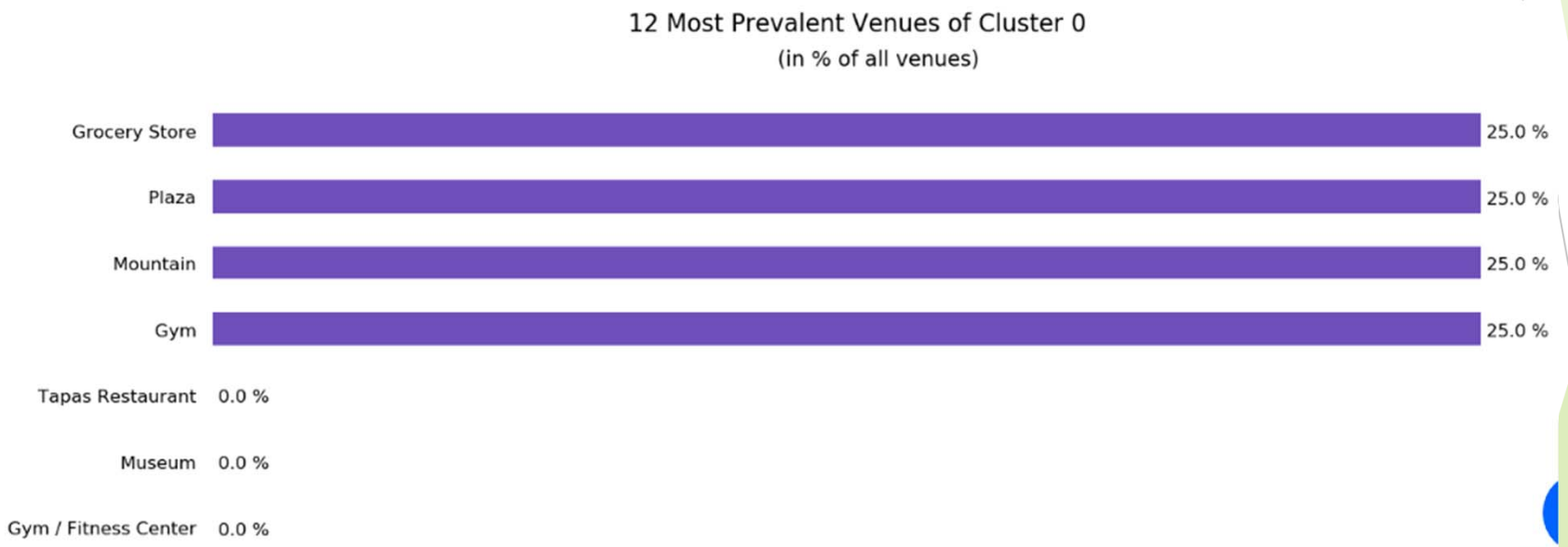
	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Abando	Restaurant	Spanish Restaurant	Tapas Restaurant	Seafood Restaurant	Wine Bar	Cocktail Bar	Plaza	Café	Bar	Bakery
1	Basurto-Zorroza	General Entertainment	Arts & Entertainment	Performing Arts Venue	Flea Market	Playground	Electronics Store	Coffee Shop	Deli / Bodega	Department Store	Dessert Shop
2	Begoña	Bar	Hotel	Coffee Shop	Café	Wine Bar	Boarding House	Gym / Fitness Center	History Museum	Hostel	Chinese Restaurant
3	Deusto	Spanish Restaurant	Restaurant	Food Truck	Plaza	Café	Metro Station	Burger Joint	Hostel	Music Venue	Diner
4	Errekalde	Mountain	Grocery Store	Gym	Plaza	Electronics Store	Cocktail Bar	Coffee Shop	Deli / Bodega	Department Store	Dessert Shop
5	Ibaiondo	Tapas Restaurant	Bar	Restaurant	Plaza	Gay Bar	Hostel	Gastropub	Spanish Restaurant	Café	Farmers Market
6	Otxarkoaga-Txurdinaga	Bar	Wine Bar	Farmers Market	Coffee Shop	Deli / Bodega	Department Store	Dessert Shop	Diner	Electronics Store	Falafel Restaurant
7	Uribarri	Hotel	Gastropub	Metro Station	Tapas Restaurant	Grocery Store	Park	Coffee Shop	Deli / Bodega	Department Store	Dessert Shop

Data Analysis

Clustering Districts

Data Analysis

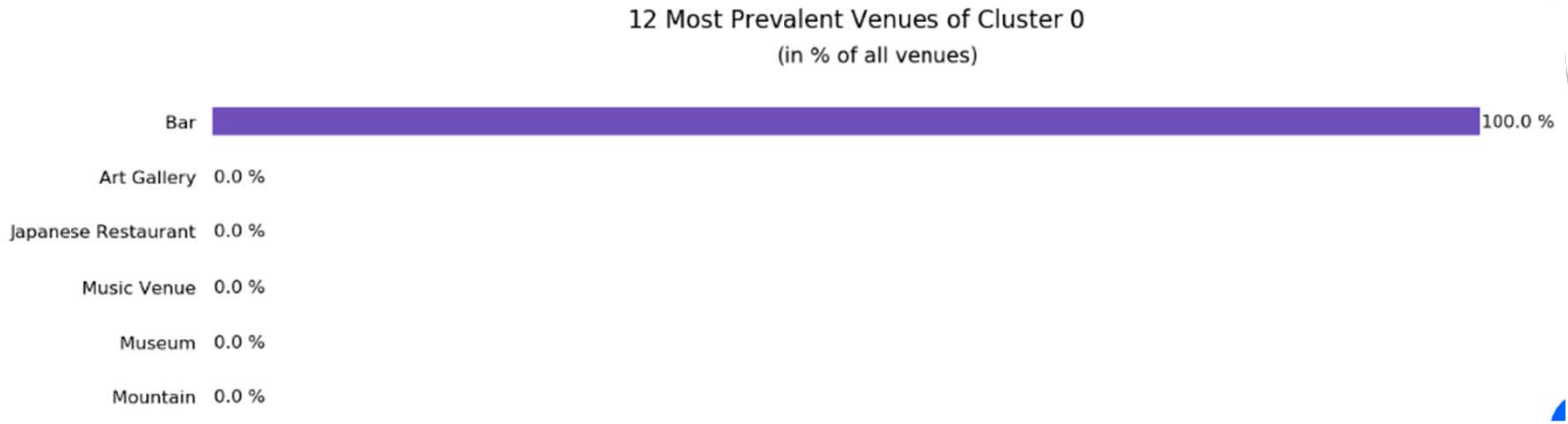
Cluster 0



In this cluster there are no restaurants, coffee shops or bars. These facts indicate that it is not an entertainment area with what would not be the best area to set up a new restaurant.

Data Analysis

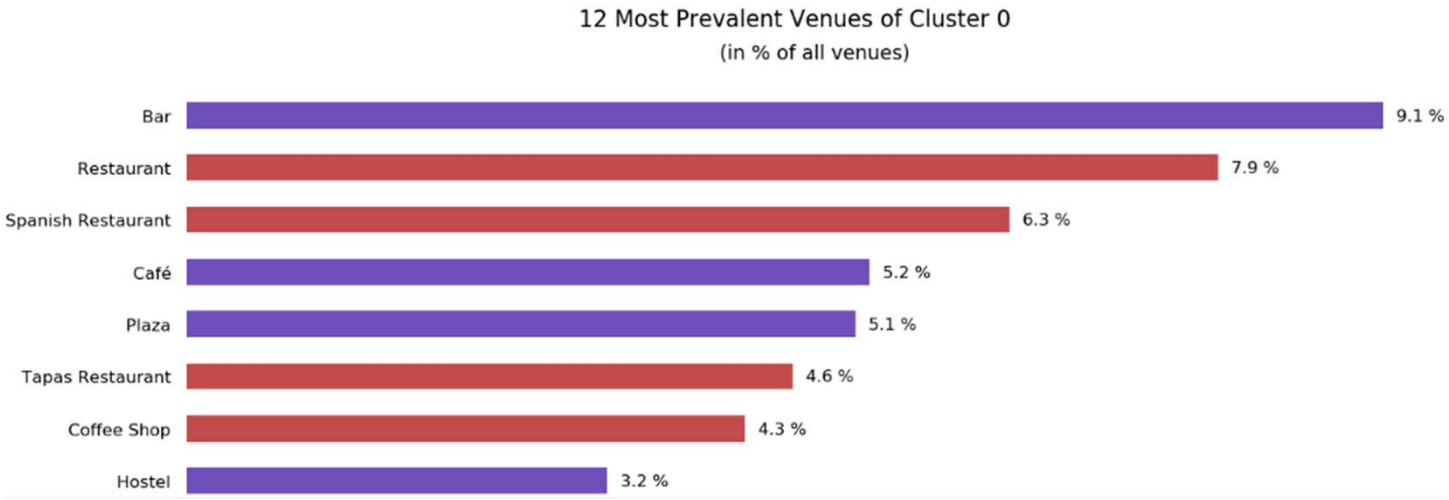
Cluster 1



In this cluster there are no restaurants or coffee shops. These facts indicate that it is not an entertainment area with what would not be the best area to set up a new restaurant. Moreover, it is a mountain area.

Data Analysis

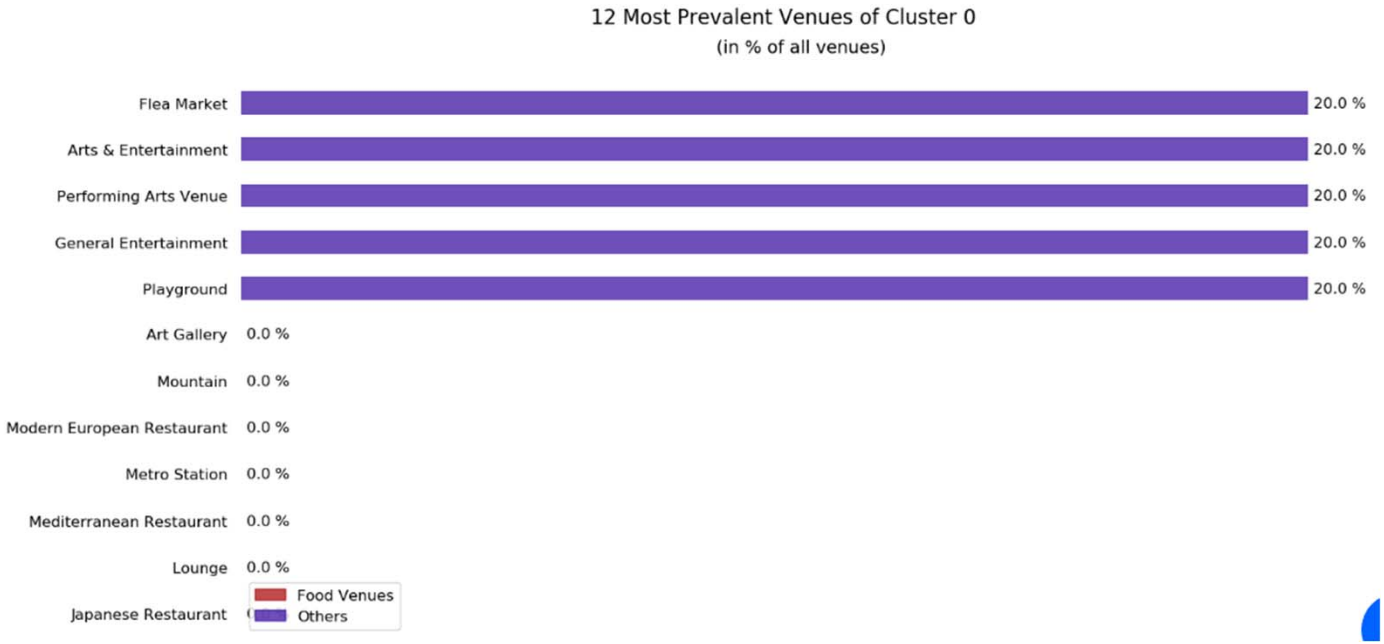
Cluster 2



This cluster is characterized by mostly bars, restaurants and coffee shops. In fact, Deusto, Begoña and Abando are tourist districts of Bilbao. Therefore, any of these three districts could be candidates to set up a new restaurant.

Data Analysis

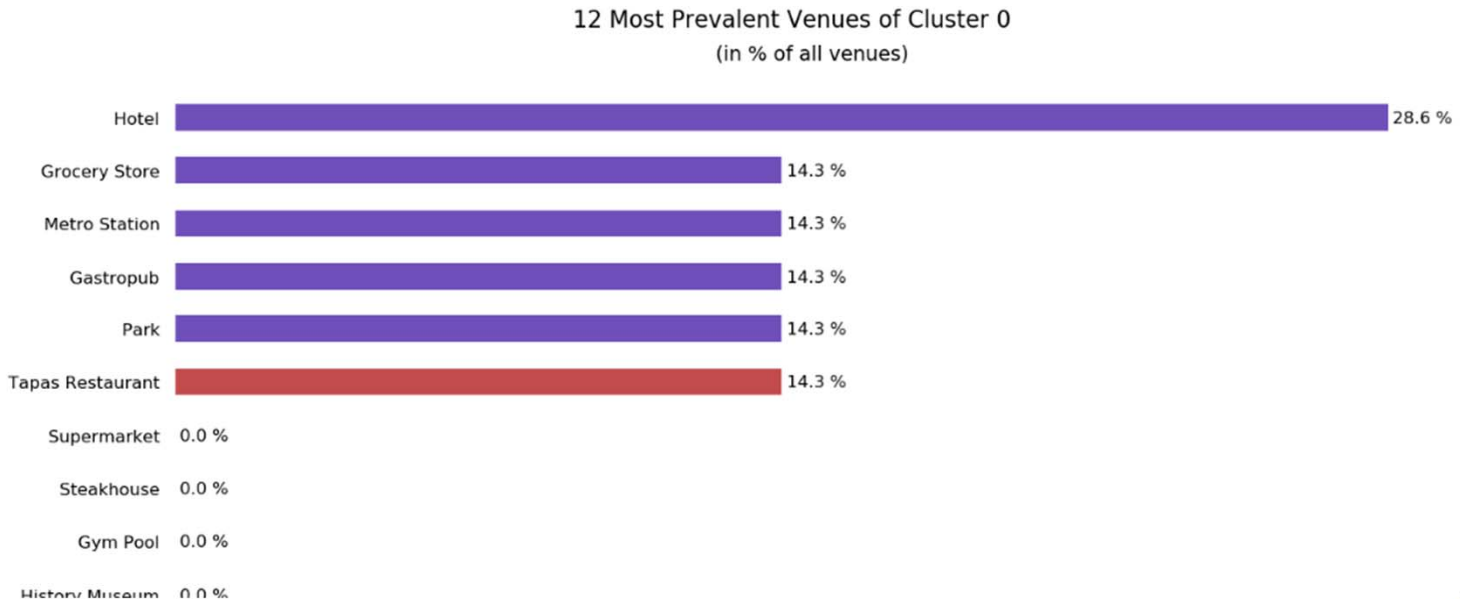
Cluster 4



In this cluster there are no restaurants or coffee shops. These facts indicate that it is not an entertainment area with what would not be the best area to set up a new restaurant.

Data Analysis

Cluster 4



Despite having a hotel, this district has no restaurants, bars or coffees shops that tell us that a new restaurant can succeed.

The chosen cluster is Cluster 2.

Data Analysis

Selecting the district in Cluster 2.

	District	Latitude	Longitude	Cluster Labels
0	Deusto	43.271387	-2.949581	2
1	Otxarkoaga-Txurdinaga	43.260369	-2.896813	2
2	Begoña	43.258890	-2.916111	2
3	Abando	43.262788	-2.934947	2

From the data frames above, it can be observed that Abando has 16 Italian Restaurants within 650 meters from its center, so this district is not a good candidate. Begoña has fewer (2) and Deusto only one. It should be noted that Deusto is a tourist area where University of Deusto and Bilbao Guggenheim Museum are located, while Begoña is not a tourist area. In addition, Deusto has an Italian restaurant at a long distance. For all these reasons, I consider that the best district would be Deusto.

Result and Discussion

- In the beginning of the analysis the data frame of Bilbao Districts was trimmed to include only the ones that had 12 or more venues. This decision was taken as it made sense to set up a restaurant in one of the more popular districts, thereby attracting the attention of a lot more people.
- When clustering the districts, the optimal value of k (k=5) for the dataset was arrived. As a consequence, all districts were grouped into 4 clusters using k-means clustering. In order to examine the deterministic characteristics of each cluster, a data frame for each cluster was created that included their most frequently occurring venues in descending order. A horizontal bar plot was generated showing the top 12 venues for each cluster, highlighting the food venues. This helped in determining the optimal cluster for further analysis. All of the observations pointed in the direction of Cluster 1 being that cluster.
- The following step was to obtain and display the closest Italian restaurants from each neighbourhood in Cluster 1 and their corresponding distances. It was observed that Abando has 16 Italian Restaurants within 700 meters from its center, so this district is not a good candidate.

Result and Discussion

- Begoña has fewer (2) and Deusto only one. It should be noted that Deusto is a tourist area where University of Deusto and Bilbao Guggenheim Museum are located, while Begoña is not a tourist area. In addition, Deusto has an Italian restaurant at a long distance.
- For all these reasons, I consider that the best district would be Deusto.
- In conclusion, this project would have had better results if there were more available data in terms of actual land pricing data within the area, public transportantion access and allowance of more venues exploration with the Foursquare (limited venues for free calls).
- However, based on the available data, my advice to the investment company would be to focus on only Deusto district when investing on a new italian restaurant.