Francisco Morillo
Applied Data Science Capstone Project
Opening a Distribution Center for Restaurants in Barcelona
August 19, 2018

DISTRIBUTION CENTER FOR RESTAURANTS IN BARCELONA

CAPSTONE PROJECT

This Capstone Project consists in scraping a Wikipedia Page to get the necessary information of the names of the neighborhoods in Barcelona; and using the dataset provided by its own government to have information on the territorial distribution of income by neighborhood. Using Google API to get Longitudes and Latitudes of the different Neighborhoods in Barcelona and using Foursquare API to get the nearby venues of each of the neighborhoods. Then using unsupervised machine learning K-means algorithm, cluster the neighborhoods, to finally, analyze each cluster in order to identify the cluster that we need. Identifying the best

neighborhood from this cluster and finally finding the centroid among the venues to use as the location for the Distribution Center.

1. Business Problem

Barcelona, located in north-eastern Spain and on the shores of the Mediterranean, is one of the principal European cities.

Given it's location, Barcelona is rich in food and cuisine culture, with the Mediterranean Diet as a key driver. Barcelona offers a number of different elements that make it an attractive place to do business. Today the city is a highly interesting location for new economic activities.

For that reason we are evaluating which could be the perfect location to open a Distribution Center for Restaurants.

The distribution Center will be in charge of distributing the necessary raw materials for the restaurants.

In order to decide, where It would be the best place to set up a Distribution Center, we will be using the Territorial Distribution Income Index, to understand how each of the neighborhoods are economically.

This business problem is intended for people interested in examples on how to use the foursquare API to cluster neighborhoods and leverage a business proposal

2. Data

For the realization of this project, the data and libraries that were used are:

- The neighborhoods name in Barcelona, found in: https://es.wikipedia.org/wiki/Distritos_de_Barcelona
- The 2016 Territorial Income Distribution for the neighborhoods in Barcelona, found in:http://opendata-ajuntament.barcelona.cat/data/dataset/79bdf758-dae1-485b-800c-be9f8cfa9360/resource/1d9ff171-6f23-45c1-b02f-203b0589f08a/download/2016_distribucio_territorial_renda_familiar.csv
- Google API Geocode
- Foursquare API

- Sklearn K Means Clustering

- Pandas

- Folium

3. Methodology

For this project we used K-means clustering as the algorithm to classify each of the neighborhoods. Besides of using the Foursquare API to get the nearby venues of each of the neighborhoods, we used amount of population and the territorial income distribution index (RFD) as variables to consider.

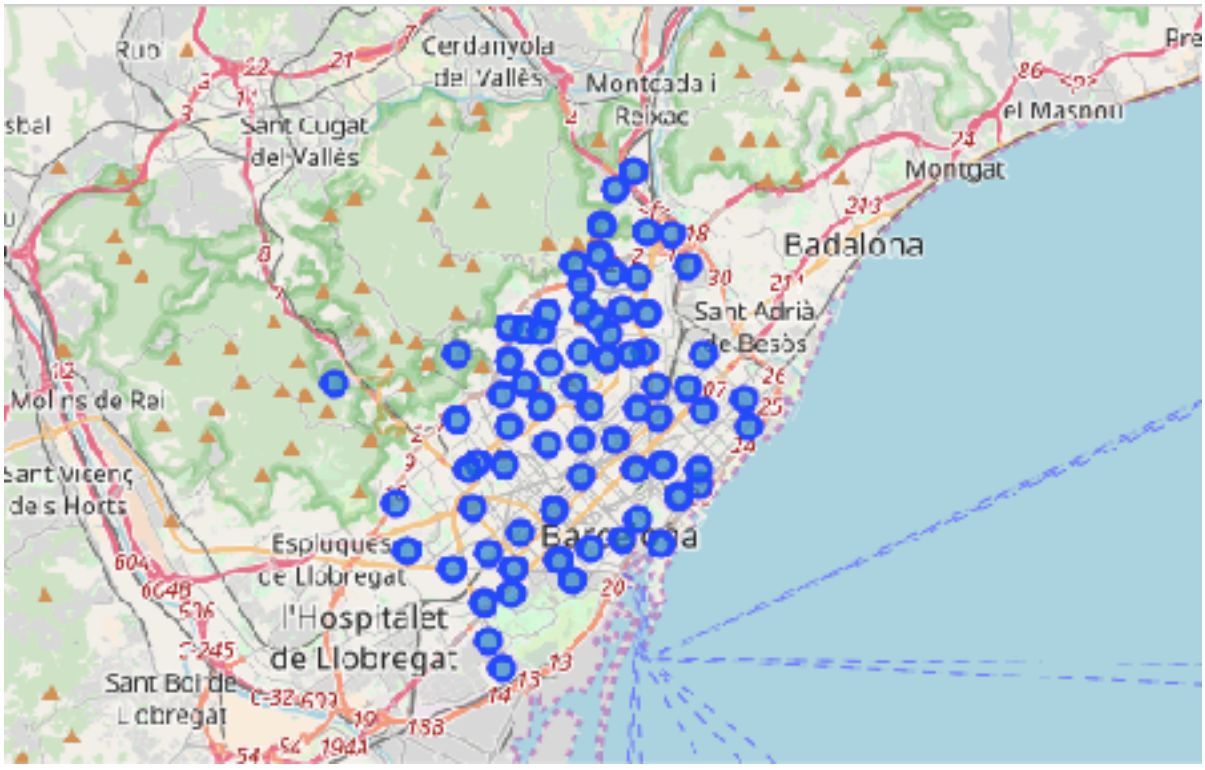
Seeing that we were handling categorical variables, we needed to use one hot encoder to get dummy variables.

Also as the population and RFD index were not in range as the rest of the variables, we needed to standardize the data frame.

The RFD Index tells us the economic capacity of the residents in Barcelona, in which the media is 100, so if a neighborhood is under 100, it does not have a high economic capacity for spending compared to the others. If the RFD is above 100, it means that is a high class neighborhood, compared to the others.

4. Results

First we visualize each of the 73 neighborhoods in Barcelona using Folium



Then we add the RFD index

	Neighborhood	Latitude	Longitude	Population	RFD Index
0	Baix Guinardó	41.412131	2.169357	26583	92.0
1	Barrio Gótico de Barcelona	41.381900	2.179186	15729	110.5
2	Barrio de Les Corts	41.389019	2.132749	45975	119.9
3	Barrio de Sant Antoni (Barcelona)	41.377266	2.159331	38182	104.0
4	Barrio de la Sagrada Familia (Barcelona)	41.404484	2.175726	61349	95.9
5	Buró de Vivar	41.444099	2.197847	2511	72.6
6	Camp d'en Grassot i Gràcia Nova	41.404068	2.155900	34275	109.1
7	Can Baró	41.417228	2.153444	8959	95.0

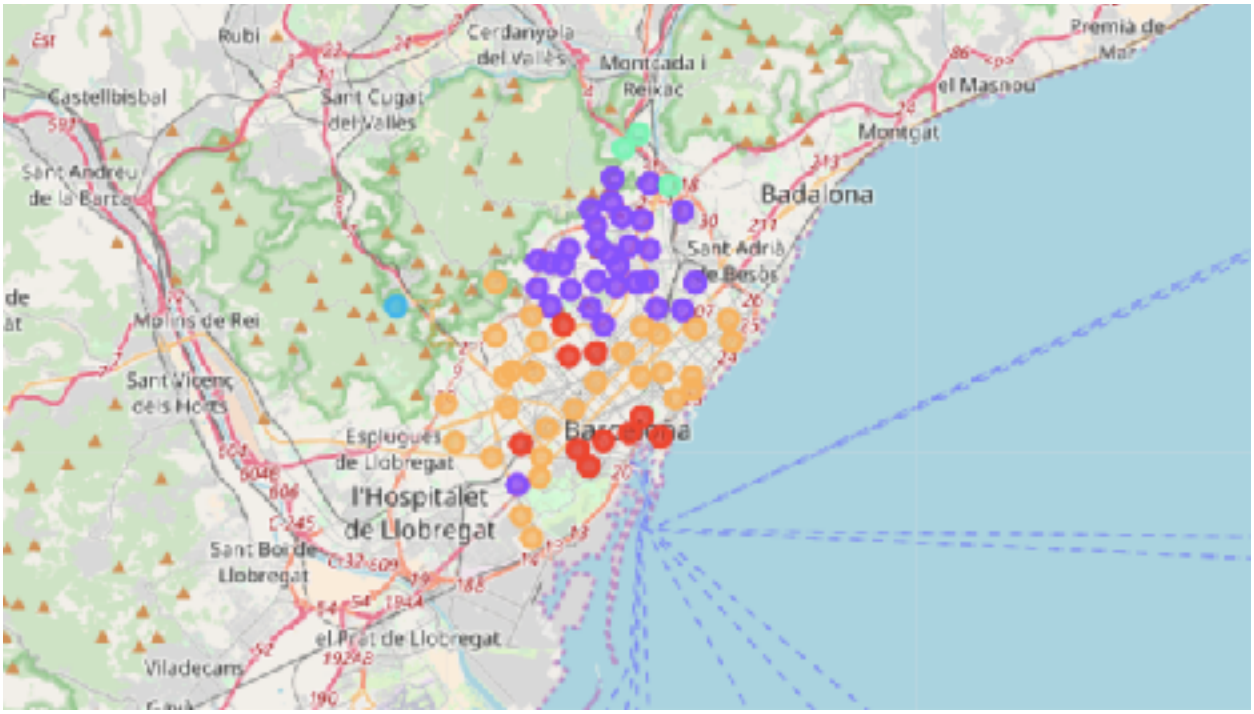
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Baix Guinardó	41.412131	2.169357	Basilicata	41.410857	2.166199	Latin American Restaurant
1	Baix Guinardó	41.412131	2.169357	Xinquito Agua	41.409808	2.166879	Beer Garden
2	Baix Guinardó	41.412131	2.169357	CIR Castillejos	41.412432	2.170713	Gym
3	Baix Guinardó	41.412131	2.169357	Phenomena	41.409693	2.171712	Indie Movie Theater
4	Baix Guinardó	41.412131	2.169357	Bella e Buon	41.411936	2.173393	Italian Restaurant

and population for the neighborhoods to the Data-frame

	Neighborhood	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
0	Baix Guinardó	Bar	Grocery Store	Japanese Restaurant	Bakery	Café	Mediterranean Restaurant	Gym	Hotel
1	Barrio Gótico de Barcelona	Italian Restaurant	Hotel	Bar	Cocktail Bar	Coffee Shop	Ice Cream Shop	Dessert Shop	Grocery Store
2	Barrio de Les Corts	Café	Hotel	Coffee Shop	Cocktail Bar	Japanese Restaurant	Italian Restaurant	Asian Restaurant	Burger Joint
3	Barrio de Sant Antoni (Barcelona)	Café	Mediterranean Restaurant	Bar	Cocktail Bar	Hotel	Coffee Shop	Argentinian Restaurant	Japanese Restaurant

Foursquare API lets get information about nearby venues on each of the neighborhoods.

We now know the name of the venues, where they are located and it's category



We then sorted the data frame to know which where the most common venues in each of the neighborhoods in Barcelona

Using K-means algorithm for clustering and using as variables the frequency of venues in neighborhoods and it's population and RFD Index we classified them into 5 categories

	Neighborhood	Population	RFD Index	Cluster Labels	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
1	Barrio Gótico de Barcelona	15329	1105	0	Italian Restaurant	Hotel	Bar	Cocktail Bar	Coffee Shop	Ice Cream Shop	Dessert Shop	Candy Store
3	Barrio de Sant Antoni (Barcelona)	38182	1040	0	Café	Mediterranean Restaurant	Bar	Cocktail Bar	Hotel	Coffee Shop	Argentinian Restaurant	Japanese Restaurant
6	Camp d'en Grassot i Gràcia Nova	34375	1081	0	Mediterranean Restaurant	Bar	Hostel	Farmers Market	Café	Brewery	Indie Movie Theater	Japanese Restaurant
22	El Raval	47274	746	0	Cocktail Bar	Bar	Mediterranean Restaurant	Hotel	Bookstore	Japanese Restaurant	Beer Bar	Coffee Shop
28	La Barceloneta	15068	848	0	Mediterranean Restaurant	Ice Cream Shop	Burger Joint	Bar	Cocktail Bar	Breakfast Spot	Beach	Beer Bar
40	La Salud (Barcelona)	13123	1164	0	Bar	Hotel	Mediterranean Restaurant	Café	Grocery Store	ArtGallery	Chinese Restaurant	Italian Restaurant
54	Pueblo Seco	40104	761	0	Mediterranean Restaurant	Bar	Cocktail Bar	Café	Brewery	Italian Restaurant	Beer Bar	Hotel
61	Sant Pere, Santa Caterina i la Ribera	22380	978	0	Bar	Cocktail Bar	Italian Restaurant	Coffee Shop	Hotel	Dessert Shop	Mediterranean Restaurant	Art Museum
62	Sants	40631	923	0	Bar	Mediterranean Restaurant	Hotel	Gym / Fitness Center	Gastropub	Hostel	Farmers Market	Bakery
72	Vila de Gracia	50670	1015	0	Bar	Café	Mediterranean Restaurant	Ice Cream Shop	Arts & Crafts Store	Asian Restaurant	Bookstore	Cocktail Bar

4.1 Cluster 0

4.2 Cluster 1

	Neighborhood	Population	RFD Index	Cluster Labels	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
[illegible]	Baix Guinardó	25669	90.0	1	Bar	Grocery Store	Japanese Restaurant	Bakery	Café	Mediterranean Restaurant	Gym	Hotel	Italian Restaurant	Breakfast Spot
[illegible]	Baró de Viver	2611	75.6	1	Clothing Store	Electronics Store	Burger Joint	Fast Food Restaurant	Cosmetics Shop	Coffee Shop	Italian Restaurant	Japanese Restaurant	Ice Cream Shop	Grocery Store
[illegible]	[illegible]	[illegible]	[illegible]	1	Gym	[illegible]	[illegible]	Bar	Bakery	[illegible]	Latin American Restaurant	[illegible]	[illegible]	Italian Restaurant
[illegible]	Can Peguera	2216	51.5	1	Grocery Store	Café	Food & Drink Shop	Bar	Burger Joint	Bakery	Breakfast Spot	Hot Dog Joint	Italian Restaurant	Metro Station
5	[illegible]	5025	54.4	1	Café	Bar	Market	Pool	Food & Drink Shop	Food Court	Mediterranean Restaurant	Grocery Store	Music Venue	Climbing Gym
11	Congrés i els Indians	14021	71.7	1	Grocery Store	Bar	Bakery	Farmers Market	Italian Restaurant	Falafel Restaurant	Cocktail Bar	Chinese Restaurant	Gym	Café
13	El Carmel	32037	64.6	1	Mediterranean Restaurant	Historic Site	Chinese Restaurant	Italian Restaurant	Coffee Shop	Bar	Bakery	Athletics & Sports	Metro Station	Food & Drink Shop
16	El Coll	7301	84.4	1	Bar	Museum	Café	Mountain	Grocery Store	Hotel	Burger Joint	Botanical Garden	College Gym	Italian Restaurant
[illegible]	[illegible]	[illegible]	[illegible]	1	Bar	[illegible]	[illegible]	Gym	Italian Restaurant	Breakfast Spot	Coffee Shop	[illegible]	Chinese Restaurant	Gym / Fitness Center
22	El Valle de Hebron	5887	81.5	1	Metro Station	Café	Gym	Grocery Store	Chinese Restaurant	Coffee Shop	Basketball Court	Hostel	Breakfast Spot	Hotel
[illegible]	Horta (Barcelona)	25667	80.6	1	Café	Bakery	Mediterranean Restaurant	Burger Joint	Hostel	Food & Drink Shop	Metro Station	Chinese Restaurant	Coffee Shop	Grocery Store
[illegible]	La Bordeta	13010	70.6	1	Café	Grocery Store	Bakery	Bar	Mediterranean Restaurant	Hotel	Japanese Restaurant	Farmers Market	Burger Joint	Italian Restaurant
[illegible]	La Clota	560	91.4	1	Café	Mediterranean Restaurant	Gym	Food & Drink Shop	Grocery Store	Chinese Restaurant	Bakery	Hostel	Coffee Shop	Deli / Bodega
31	La Font d'en Fargues	9423	84.6	1	Café	Bar	Grocery Store	Food & Drink Shop	Italian Restaurant	Burger Joint	Hot Dog Joint	Mediterranean Restaurant	Historic Site	German Restaurant

4.3 Cluster 2

	Neighborhood	Population	RFD Index	Cluster Labels	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
69	Vallvidrera, Tibidabo i les Planes	4641	130.6	2	Hostel	BBQ Joint	Concert Hall	Comic Shop	Comfort Food Restaurant	Comedy Club	College Gym	College Cafeteria	Coffee Shop

4.4 Cluster 3

	Neighborhood	Population	RFD Index	Cluster Labels	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
10	Ciutat Meridiana	10056	34.3	3	Metro Station	Grocery Store	Café	Mediterranean Restaurant	Music Venue	Comedy Club	College Gym	College Cafeteria	Coffee Shop
43	La Trinitat Vella	10048	45.9	3	Metro Station	Music Venue	Circus	Comfort Food Restaurant	Comedy Club	College Gym	College Cafeteria	Coffee Shop	Cocktail Bar
67	Vallbona (barrio)	1354	51.9	3	Metro Station	Grocery Store	BBQ Joint	Furniture / Home Store	Cupcake Shop	Music Venue	Comedy Club	College Gym	College Cafeteria

4.5 Cluster 4

	Neighborhood	Population	RFD Index	Cluster Labels	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
2	Barrio de Les Corts	46976	119.8	4	Café	Hotel	Coffee Shop	Cocktail Bar	Japanese Restaurant	Italian Restaurant	Asian Restaurant
4	Barrio de la Sagrada Familia (Barcelona)	51348	95.8	4	Italian Restaurant	Historic Site	Bank	Burger Joint	Gym	Hotel	Bar
12	Diagonal Mar i Front Marítim del Poblenou	13455	152.6	4	Mediterranean Restaurant	Hotel	Bank	Italian Restaurant	Clothing Store	Coffee Shop	Burger Joint
13	Dreta de l'Eixample	48716	170.0	4	Hotel	Mediterranean Restaurant	Hostel	Boutique	Japanese Restaurant	Cocktail Bar	Bakery
14	El Besòs y el Maresme	22508	54.7	4	Hotel	Café	Italian Restaurant	Mediterranean Restaurant	Clothing Store	Ice Cream Shop	Burger Joint
15	El Eixample (Barcelona)	12584	57.0	4	Hotel	Mediterranean Restaurant	Beach Bar	Café	Bar	Bakery	Beach
16	El Camp de Arpa del Clot	38252	82.7	4	Hotel	Bar	Bakery	Italian Restaurant	Café	Grocery Store	Korean Restaurant
18	El Clot	26528	58.7	4	Hotel	Mediterranean Restaurant	Gym / Fitness Center	Bakery	Grocery Store	Cocktail Bar	Japanese Restaurant

5. Discussion

After obtaining the clusters of the neighborhoods, we begin to analyze each of the clusters to really understand what it is telling us.

In the cluster 0 , we can tell that the average of the RFD is about 100 and the most common venues are mostly bars. Seeing that we are searching for an optimal location for our distribution center for restaurants we will continue analyzing the other clusters.

Cluster 1 seems to group the average of the neighborhoods and the ones that have most population. Its most common venues are grocery stores and Metro Stations, and it's RFD index is pretty low.

Cluster 2 only contains Tibidabo, which makes sense considering that it is outside of the city and is mostly the amusement park located at the mountain.

Cluster 3 has a very low RFD index and low population, which seems to tell us that these neighborhoods are not ideal for our distribution center, noticing as well that its most common venue is the metro station

Finally cluster 4, gathers the neighborhoods that have mostly a high RFD index and its most common venues are hotels and restaurants. Cluster 4 is ideal for us to find the perfect neighborhood for our distribution center.

From these neighborhoods we are going to select the one that has the highest RFD index, making the one that has a better distribution of income from the neighborhoods in Barcelona

Taking this into consideration, the neighborhood that we are selection for the opening of a new distribution Center for restaurants is Pedralbes with a RFD index of 242.4, way above the average

6. Conclusion

Now that we know in which neighborhood we are going to set up the distribution center. We are going to select the exacts coordinates of the Distribution Center. By finding the centroid of the coordinates of the restaurants in a 500 meter radius from Pedralbes, we decide to set up our distribution Center in (41.3894233187, 2.11117622151). In the image we can visualize the Distribution Center as the red circle.