Opening a new restaurant in Granada

Lisardo Baena Padial

June 03, 2021

1. Introduction: Business Problem.

I am from Granada a little city in the south of Spain and for this project I am going to analyse which location would be the best for opening a new restaurant.

Granada is a city full of students due to the importance of the University one of the oldest in Spain. I would like to open a restaurant for young students, where they can from having some beers to buy the daily menu for taking away.

So at first I am going to look for data of different locations in Granada. I will look for at different webs for obtaining data I need.

Once I have the data I will analyse venues in different locations of Granada. At first I will look for University buildings and select those address where there are students.

After this I will classify the top five locations in two groups with Clustering techniques and I will choose the best address attending to the most common venues and the distance to centre of the city.

2. Obtaining the data

I have obtained data needed for my Analysis through two ways: Searchiing in Google directly and using Foursquare API.

At first I look for Postal codes for Granada on Internet. For doing this I type directly in Google “postal codes Spain Excel” and I found this dataset https://postal.cat/index.es.html.

I upload data in Python and I doing some changes regarding to the names of columns or formats among other to have the data prepared for Analysis. Finally a créate a dataset only for those addresses of the city of Granada capital, excluding towns.

A sample of the dataframe obtained is this:

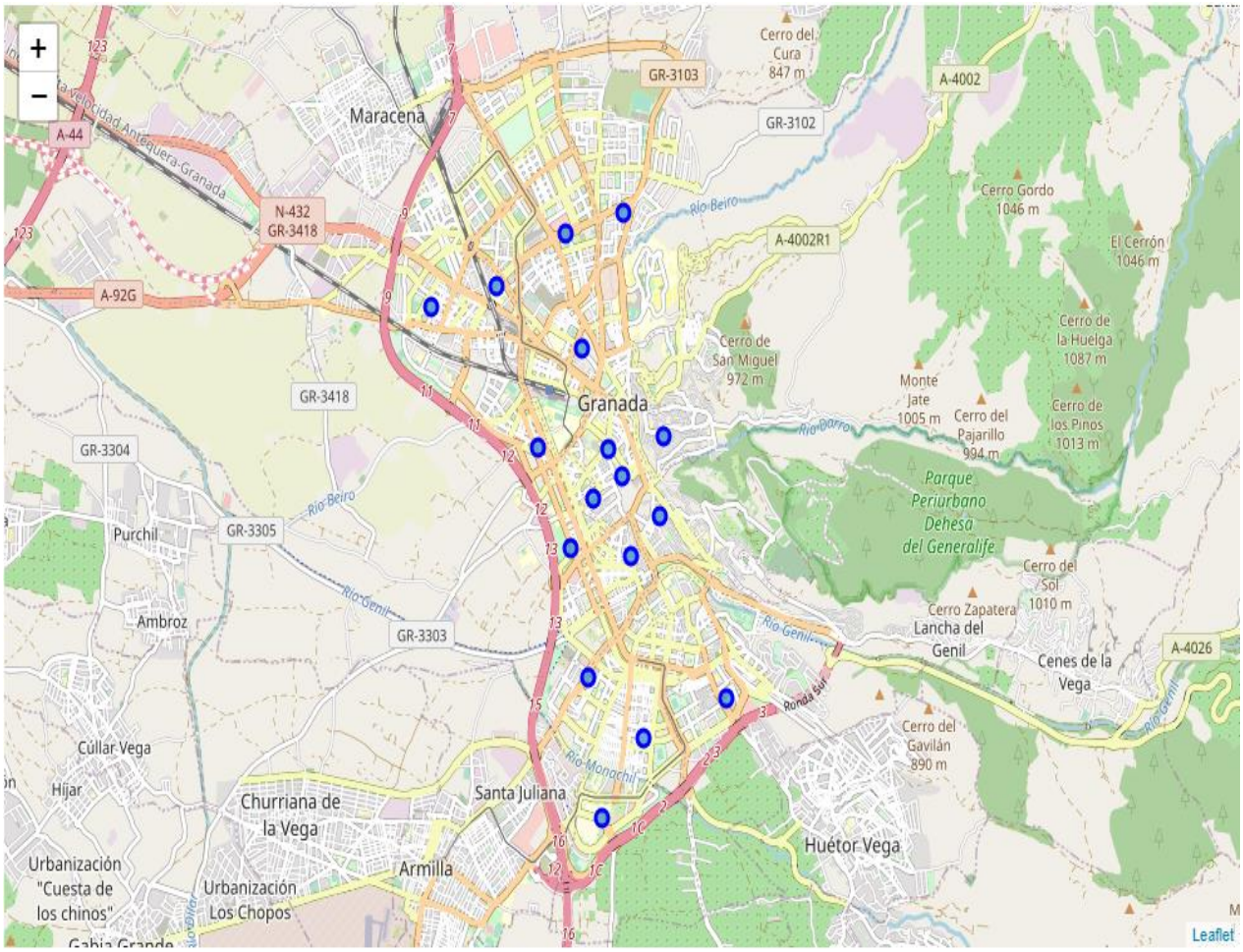
	cp	calle	poblacion	provinciaid	provincia	paisid	pais	direccion
1853	18011	Puerto Lumbreras	Granada	18	Granada	ES	España	Puerto Lumbreras, 18011, Granada
2361	18015	Virgen De La Fuensanta	Granada	18	Granada	ES	España	Virgen De La Fuensanta, 18015, Granada
1344	18004	Maria Moliner	Granada	18	Granada	ES	España	Maria Moliner, 18004, Granada
1236	18015	Leon Felipe	Granada	18	Granada	ES	España	Leon Felipe, 18015, Granada
2185	18009	Sor Cristina Mesa	Granada	18	Granada	ES	España	Sor Cristina Mesa, 18009, Granada
384	18006	Cadiz,avenida (pares Del 2 Al Final)	Granada	18	Granada	ES	España	Cadiz,avenida (pares Del 2 Al Final), 18006, G...
1551	18010	Ortegas,plaza	Granada	18	Granada	ES	España	Ortegas,plaza, 18010, Granada
361	18001	Boqueron,plaza	Granada	18	Granada	ES	España	Boqueron,plaza, 18001, Granada
1732	18007	Pingarron	Granada	18	Granada	ES	España	Pingarron, 18007, Granada
2216	18013	Tete Monteliu	Granada	18	Granada	ES	España	Tete Monteliu, 18013, Granada

After I créate the table with different Addresses in Granada city I grouped data by Postal Code to be more efficient in the analysis. And finally I look for different latitudes and longitudes for Granada Addresses. For this I used geolocator library of Python.

	cp	provincia	pais	direccion	latitude	longitude
0	18001	Granada	España	18001, Granada, España	37.176761	-3.600920
1	18002	Granada	España	18002, Granada, España	37.174797	-3.604873
2	18003	Granada	España	18003, Granada, España	37.179166	-3.612516
3	18004	Granada	España	18004, Granada, España	37.170560	-3.607982
4	18005	Granada	España	18005, Granada, España	37.169961	-3.599788

Finally for having a quick look of Venues in the Postal Code corresponding to the center of Granada city (18001, Granada, España), I use Foursquare and I see that more repeated venues are bars. The problem was that search I was doing didn´t include all the streets for this Postal code, point that I corrected once I begin with Analysis part.

To have more feeling on the data and locations for Granada I created a Folium map representig different Postal Codes in the city.



3. Methodology

For taking the decision on the best place for opening a new student bar in Granada I am going to do the next procedures:

- At first I will explore data for different student locations in Granada (University, Elementary School, Academy, etc.).

- Once I have detected the Addresses with the highest number of student venues I will do an analysis of all the venues in these locations. I will use Foursquare not only at Postal code level, but also at street level. (for trying to obtain enough information).

- At last point for taking a decision I will Cluster these Addresses taking into account the most common venues and I will use the result and the previous analysis for taking a decision.

4. Analysis

4.1 Exploring data for different student locations in Granada.

The aim for this Analysis is to determine which Postal Code location in Granada is the best to open a new restaurant. Taking into account that students and young people will be our customer objective I will be focus in locations with student buildings.

After execute some general searches with Foursquare API I find out that list of student buildings for Granada have to include the next:

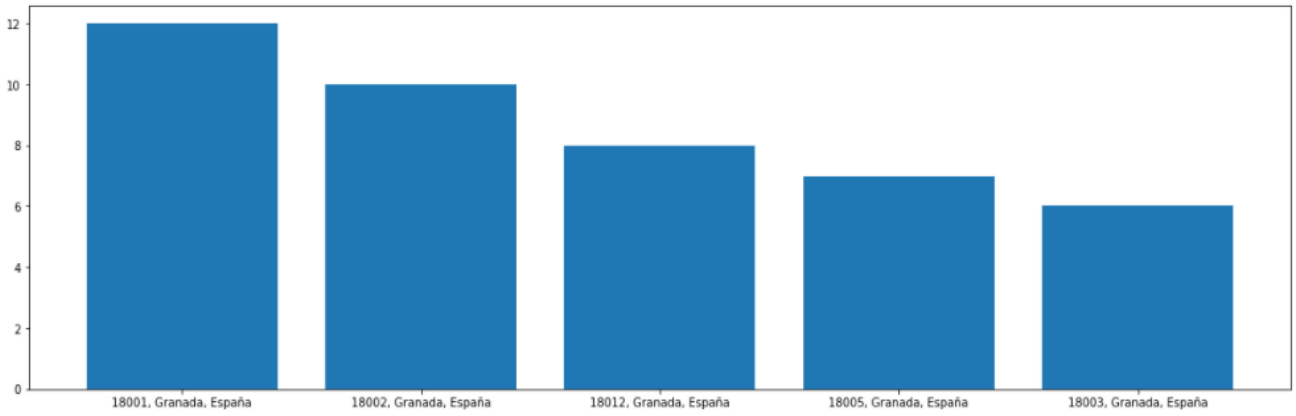
```
]: # I look for venues for this type of locations, in order to identify those places with students (customer objective)

locations = ['University','General College & University', 'Elementary School', 'College Academic Building']

LIMIT= 100
```

Granada is a city really known for the University and a lot of youngsters move from another cities and countries to study there. So for trying to be more effcient in my analysis and be focus in locations a large number students I choose the top5 Addresses according with the highest number of buildings in the list.

	Address	Total
0	18001, Granada, España	12
1	18002, Granada, España	10
2	18012, Granada, España	8
3	18005, Granada, España	7
4	18003, Granada, España	6



4.2 Exploring most common venues in Granada

After identifying and select the Addresses to be analyse, I decided to check the most common venues for each one.

At first I create another dataframe not only with Postal Codes for each Address, but also with all streets in Granda for each Postal code. The objective of this is to extract an enough quantity of data to perform a good analysis. Here I show you a simple of this initial Dataframe:

	cp	calle	poblacion	provinciaid	provincia	paisid	pais	Address	Address2	latitude	longitude
0	18012	Abanilla	Granada	18	Granada	ES	España	Abanilla, 18012, Granada	18012, Granada, España	37.187640	-3.606435
1	18005	Aben Humeya	Granada	18	Granada	ES	España	Aben Humeya, 18005, Granada	18005, Granada, España	37.169961	-3.599788
2	18005	Abu Isac	Granada	18	Granada	ES	España	Abu Isac, 18005, Granada	18005, Granada, España	37.169961	-3.599788
3	18005	Acequia Gorda,callejon	Granada	18	Granada	ES	España	Acequia Gorda,callejon, 18005, Granada	18005, Granada, España	37.169961	-3.599788
4	18001	Acera De Canasteros	Granada	18	Granada	ES	España	Acera De Canasteros, 18001, Granada	18001, Granada, España	37.176761	-3.600920

Once I have this dataframe with all the streets for top5 student addresses in Granada I built a function to extract different venues for each street using Forsquare. The number of venues extracted was of a total of **29.432** venues.

So to be more efficient in the analysis I decided finally analyse the top 10 most common venues grouping by postal code rather than by street. The results obtained show that the most common venues in Granada are all type of restaurants and bars.

	CP	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	18001	Tapas Restaurant	Spanish Restaurant	Plaza	Hotel	Bar	Café	Tea Room	Moroccan Restaurant	Gift Shop	Concert Hall
1	18002	Tapas Restaurant	Spanish Restaurant	Hotel	Plaza	Bar	Japanese Restaurant	Coffee Shop	Gift Shop	Café	Record Shop
2	18003	Tapas Restaurant	Bar	Pub	Hotel	Spanish Restaurant	Gym	Department Store	Restaurant	Food & Drink Shop	Italian Restaurant
3	18005	Tapas Restaurant	Spanish Restaurant	Hotel	Bistro	Restaurant	Plaza	Coffee Shop	Café	Pub	Italian Restaurant
4	18012	Tapas Restaurant	Bar	Café	Plaza	Restaurant	Pizza Place	Hotel	Italian Restaurant	Breakfast Spot	Spanish Restaurant

4.3 Clustering addresses for taking the decision.

Finally I decided to make a Cluster analysis for trying to group different student locations according to the most common venues and find some feature that helps me to choose the best place for my restaurant.

The most common venue for all locations is Tapas Restaurant which give us an evidence of the importance of these type of business in Granada. The rest of the most common venues are more or less the same.

Clustering is difficult with these similar features and I decided to create only 2 Clusters when applying the machine learning algorythm

This is the Cluster 0 >> Addresses in the city center.

```
# Now I examine the different Clusters student locations in Granada - Cluster 0

granada_merged.loc[granada_merged['ClusterLabels'] ==0, granada_merged.columns[[2] + list(range(5, granada_merged.shape[1]))]]
```

	direccion	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	ClusterLabels
0	18001, Granada, España	Tapas Restaurant	Spanish Restaurant	Plaza	Hotel	Bar	Café	Tea Room	Moroccan Restaurant	Gift Shop	Concert Hall	0
1	18002, Granada, España	Tapas Restaurant	Spanish Restaurant	Hotel	Plaza	Bar	Japanese Restaurant	Coffee Shop	Gift Shop	Café	Record Shop	0
3	18005, Granada, España	Tapas Restaurant	Spanish Restaurant	Hotel	Bistro	Restaurant	Plaza	Coffee Shop	Café	Pub	Italian Restaurant	0

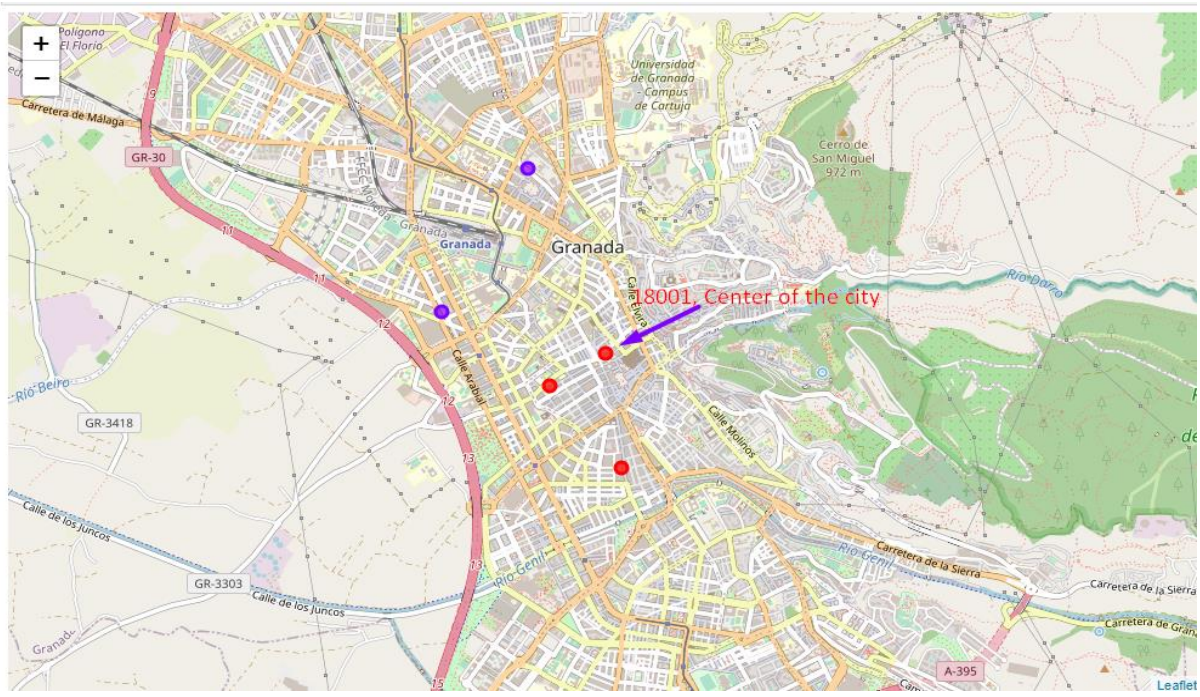
This is the Cluster 1 >> Addresses not in the city center.

```
# Now I examine the different Clusters student locations in Granada - Cluster 1

granada_merged.loc[granada_merged['ClusterLabels'] ==1, granada_merged.columns[[2] + list(range(5, granada_merged.shape[1]))]]
```

	direccion	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	ClusterLabels
2	18003, Granada, España	Tapas Restaurant	Bar	Pub	Hotel	Spanish Restaurant	Gym	Department Store	Restaurant	Food & Drink Shop	Italian Restaurant	1
4	18012, Granada, España	Tapas Restaurant	Bar	Café	Plaza	Restaurant	Pizza Place	Hotel	Italian Restaurant	Breakfast Spot	Spanish Restaurant	1

The results obtained differentiate these locations in two groups, but continue being difficult to find differences and as consequence choose the best place for opening our restaurant. I analyse and the prices to rent a local in Granada are higher in the city center. I decide to create a map Folium for checkin Clusters created and check wihic loccation would be the cheapest.



5. Conclusion

It is really difficult to determine which place could be the best location in Granada for opening a restaurant due to there are a lot of type of similar venues.

However to conclude with this project and after analysing different features described in this report, I thinl the best place for opening a new restaurant is **“18012, Granada, España”.**

The reasons to choose this location like the best for opening my restaurant are the next:

- This is the third location with more student places >> 8 student builidings.

- This is included in the Cluster of "Not in the center of the city" >> Probably will be the place with lowest rental prices.