MEDELLÍN CHICHARRÓN LOCATION

By: Marco Cruz Dawkins
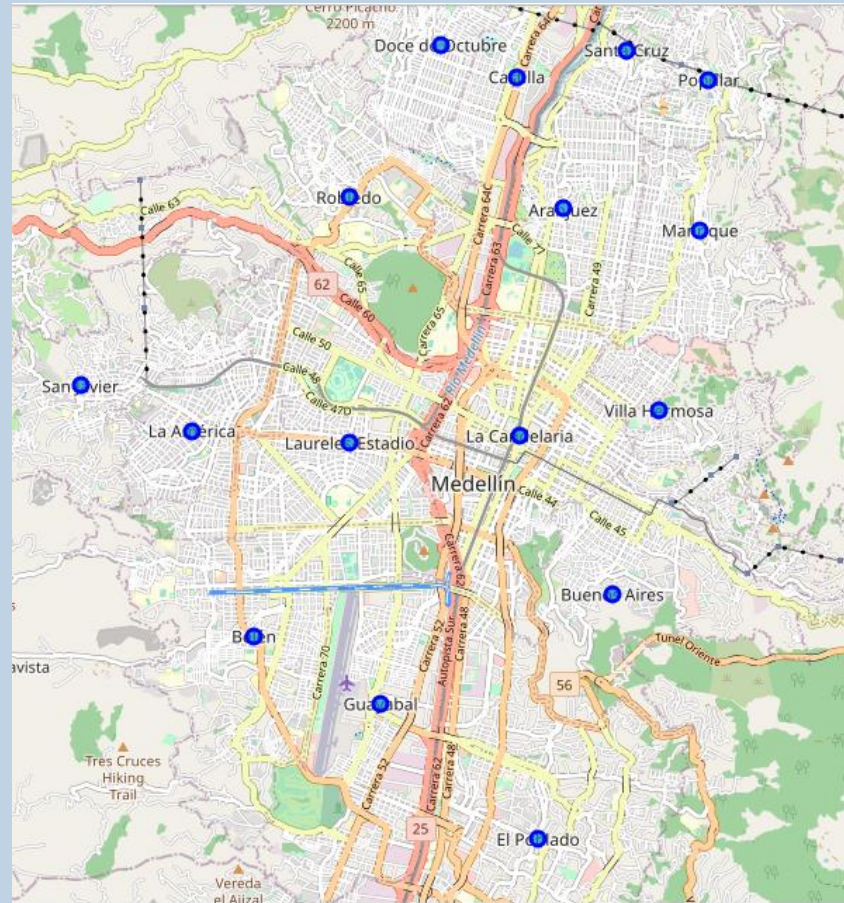
Description of the Problem

◦ Andrea thinks there is no specialized chicharron restaurant in Medellin knowing that this city culturally loves this dish because it is the best part of the Bandeja Paisa. But she does not know where would be the best spot for this restaurant having in mind that is a new category of the restaurant in the city.

◦ Andrea family have a unique recipe for this dish that they are willing to share and she also is taking a trip to get trained in Barranquilla, Colombia that has more than 10 specialized chicharron restaurant that works 24/7 so her venture has the best chicharron once it's launched.

◦ For these reasons she wanted to use Data Science to pick the best location for her restaurant.
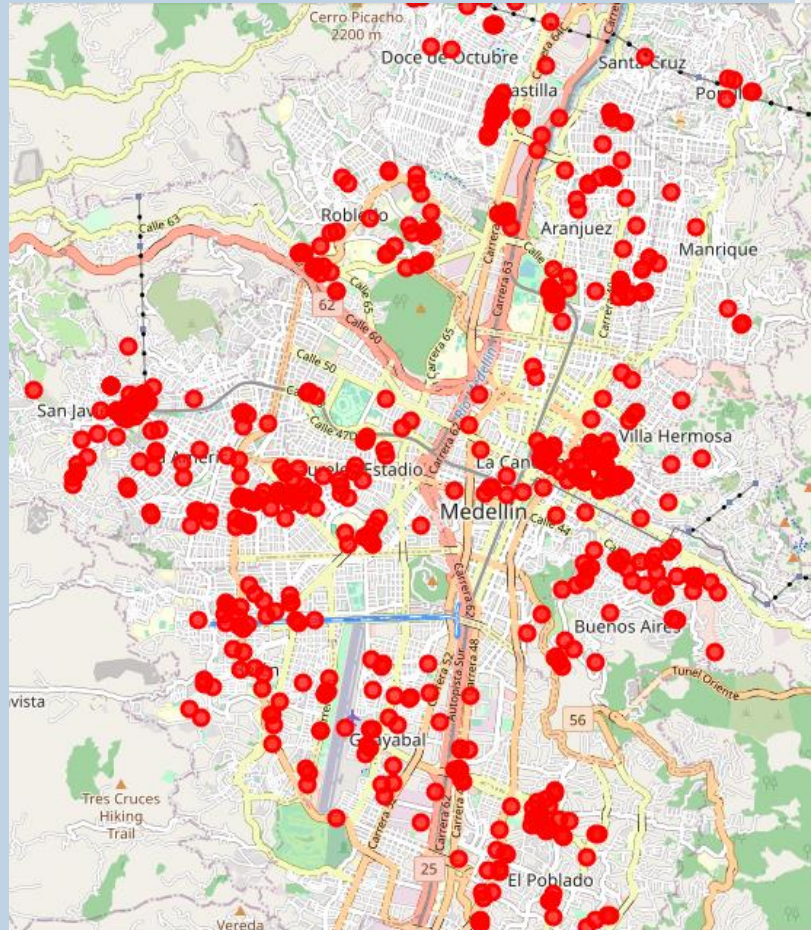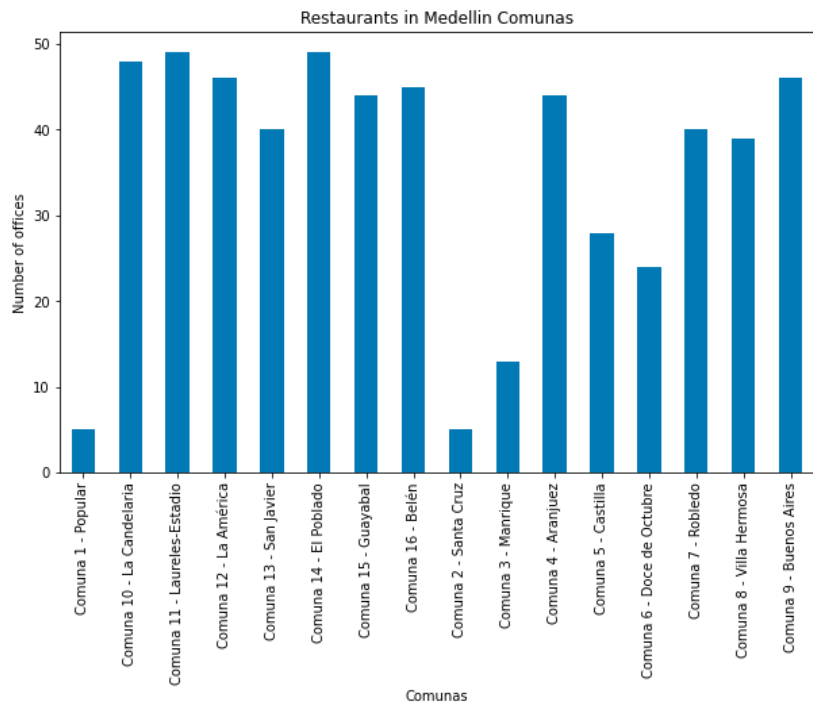
Data acquisition and cleaning

◦ Medellin in divided in 16 comunas, data scraped form
https://es.wikipedia.org/wiki/Anexo:Barrios_de_Medell%C3%ADn

◦ Latitude and longitude for these comunas was acquired using
https://geopy.readthedocs.io/en/stable/

◦ Venues in 4 categories (restaurants, universities, bars and hotels) in the 16 Comunas, data scraped https://developer.foursquare.com/

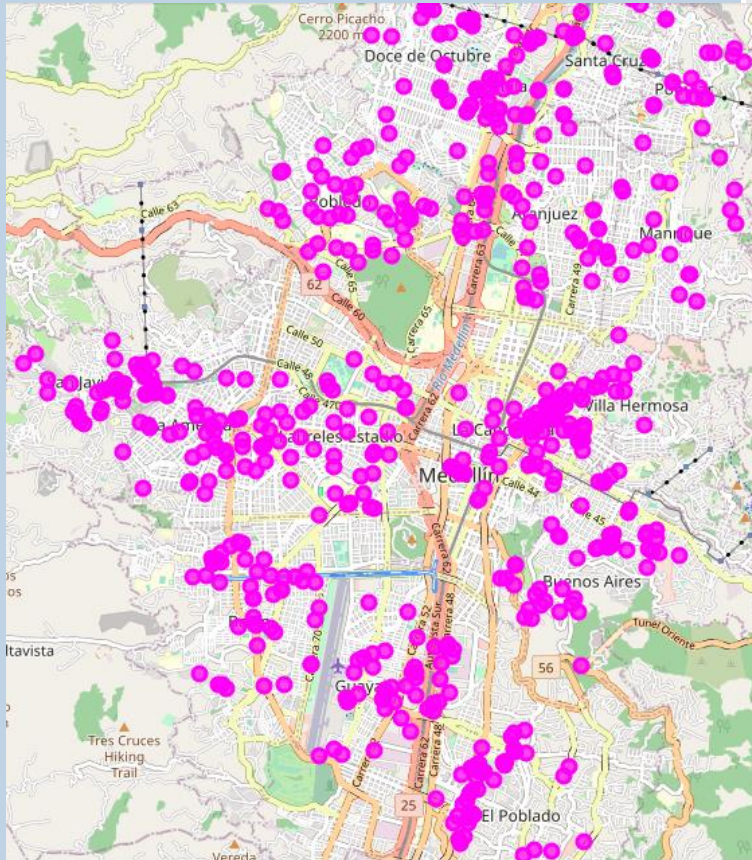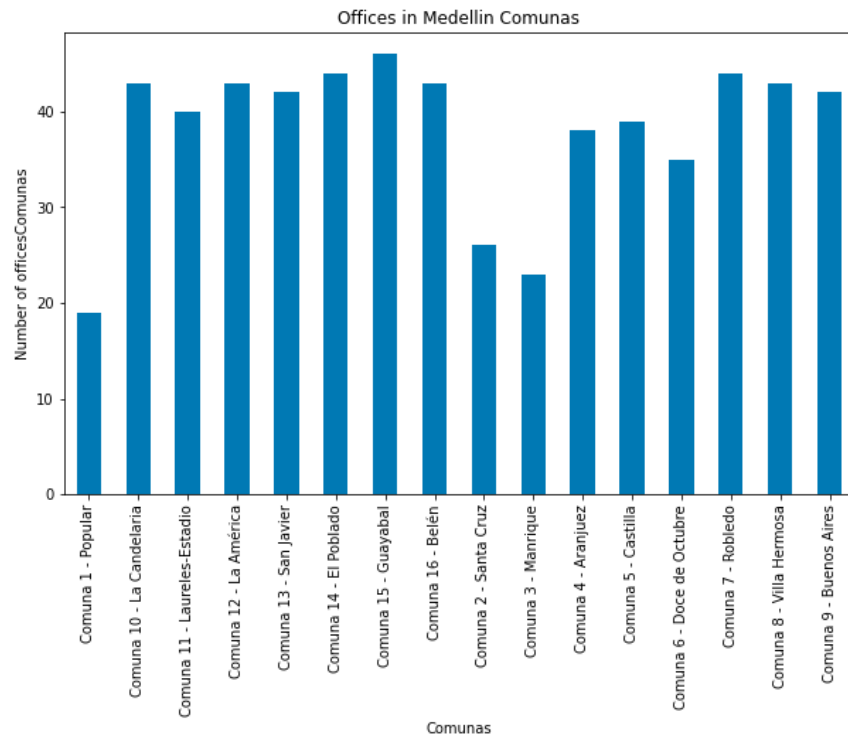|	Comunas	Latitude	Longitude
0	Comuna 1 - Popular	6.293344	-75.544825
1	Comuna 2 - Santa Cruz	6.296956	-75.554852
2	Comuna 3 - Manrique	6.275061	-75.545892
3	Comuna 4 - Aranjuez	6.277723	-75.562599
4	Comuna 5 - Castilla	6.293699	-75.568264
5	Comuna 6 - Doce de Octubre	6.297565	-75.577556
6	Comuna 7 - Robledo	6.279074	-75.588767
7	Comuna 8 - Villa Hermosa	6.253104	-75.550897
8	Comuna 9 - Buenos Aires	6.230779	-75.556594
9	Comuna 10 - La Candelaria	6.250102	-75.567915
10	Comuna 11 - Laureles-Estadio	6.249238	-75.588654
11	Comuna 12 - La América	6.250701	-75.607904
12	Comuna 13 - San Javier	6.256274	-75.621543
13	Comuna 14 - El Poblado	6.201092	-75.565665
14	Comuna 15 - Guayabal	6.217549	-75.584891
15	Comuna 16 - Belén	6.225745	-75.600348

Data acquisition and cleaning

Using Pandas, Beautiful Soup and Requests libraries the following data frame.

Data acquisition and cleaning

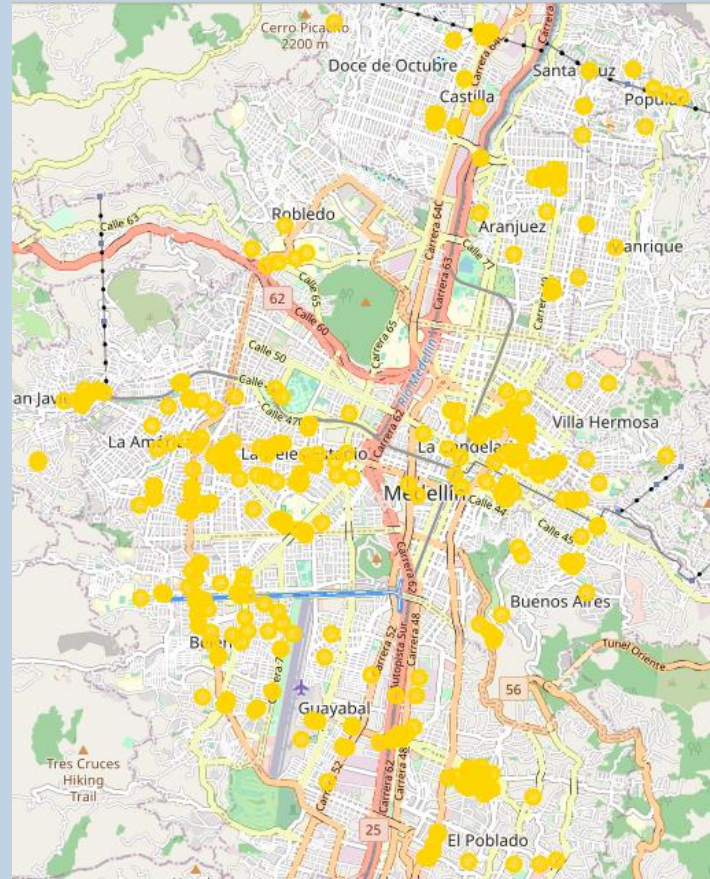Using Foursquare we acquire the information about the four venues categories.
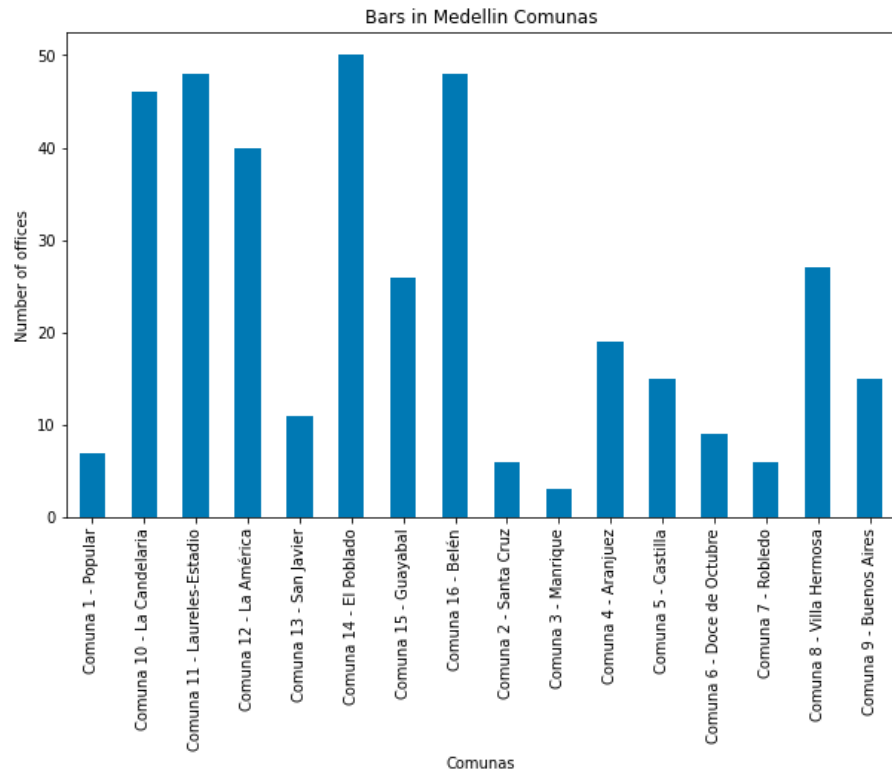
- Food (Restaurant).

Data acquisition and cleaning

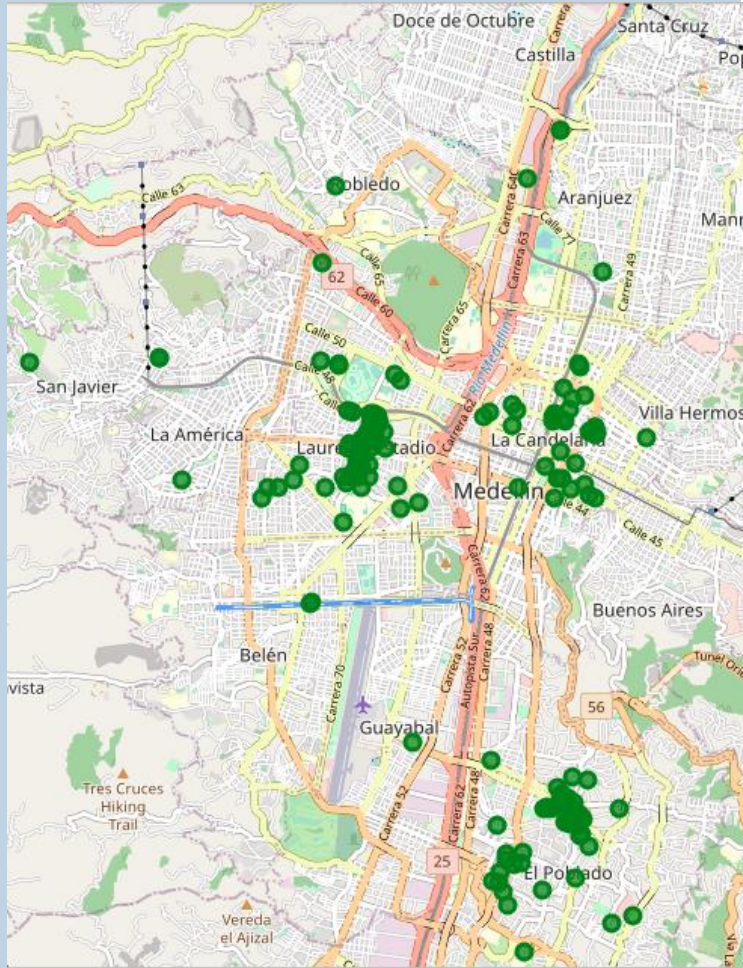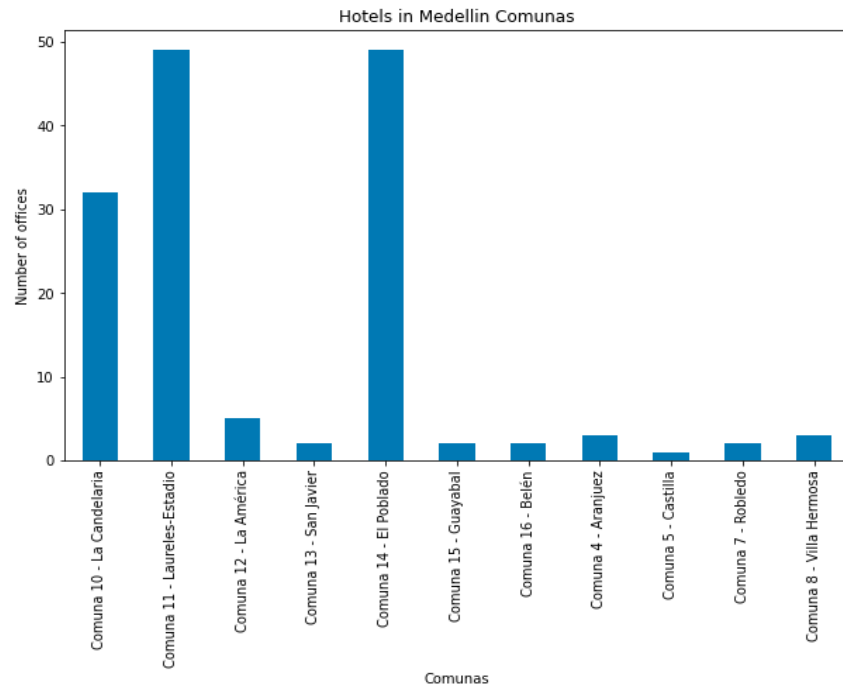Using Foursquare we acquire the information about the four venues categories.

- Offices.

Data acquisition and cleaning

Using Foursquare we acquire the information about the four venues categories.

- Night clubs.

Data acquisition and cleaning

Using Foursquare we acquire the information about the four venues categories.

- Hotels.

	Comuna	Latitude	Longitude	Food	Offices	Nightlife Spot	Hotels
0	Comuna 1 - Popular	6.293344	-75.544825	5.0	19.0	7.0	0.0
1	Comuna 2 - Santa Cruz	6.296956	-75.554852	5.0	26.0	6.0	0.0
2	Comuna 3 - Manrique	6.275061	-75.545892	13.0	23.0	3.0	0.0
3	Comuna 4 - Aranjuez	6.277723	-75.562599	44.0	38.0	19.0	3.0
4	Comuna 5 - Castilla	6.293699	-75.568264	28.0	39.0	15.0	1.0
5	Comuna 6 - Doce de Octubre	6.297565	-75.577556	24.0	35.0	9.0	0.0
6	Comuna 7 - Robledo	6.279074	-75.588767	40.0	44.0	6.0	2.0
7	Comuna 8 - Villa Hermosa	6.253104	-75.550897	39.0	43.0	27.0	3.0
8	Comuna 9 - Buenos Aires	6.230779	-75.556594	46.0	42.0	15.0	0.0
9	Comuna 10 - La Candelaria	6.250102	-75.567915	48.0	43.0	46.0	32.0
10	Comuna 11 - Laureles-Estadio	6.249238	-75.588654	49.0	40.0	48.0	49.0
11	Comuna 12 - La América	6.250701	-75.607904	46.0	43.0	40.0	5.0
12	Comuna 13 - San Javier	6.256274	-75.621543	40.0	42.0	11.0	2.0
13	Comuna 14 - El Poblado	6.201092	-75.565665	49.0	44.0	50.0	49.0
14	Comuna 15 - Guayabal	6.217549	-75.584891	44.0	46.0	26.0	2.0
15	Comuna 16 - Belén	6.225745	-75.600348	45.0	43.0	48.0	2.0

Data processing

For each Comuna we counted the total of the venues in each category that we selected.

Data processing

	Comuna	Score
13	Comuna 14 - El Poblado	53.05
10	Comuna 11 - Laureles-Estadio	51.25
9	Comuna 10 - La Candelaria	45.15
15	Comuna 16 - Belén	35.15
11	Comuna 12 - La América	33.10
14	Comuna 15 - Guayabal	27.00
7	Comuna 8 - Villa Hermosa	26.50
3	Comuna 4 - Aranjuez	22.55
8	Comuna 9 - Buenos Aires	21.10
12	Comuna 13 - San Javier	19.60
4	Comuna 5 - Castilla	18.90
6	Comuna 7 - Robledo	18.10
5	Comuna 6 - Doce de Octubre	14.75
1	Comuna 2 - Santa Cruz	9.40
2	Comuna 3 - Manrique	8.25
0	Comuna 1 - Popular	8.05

Having the total of each venue, their were weighed as follows:

Positive weight, because Andrea wants to open a Chicharron joint and thus wants as much traffic as possible

Weight food = 0.1

Positive weight because employees are even better customers
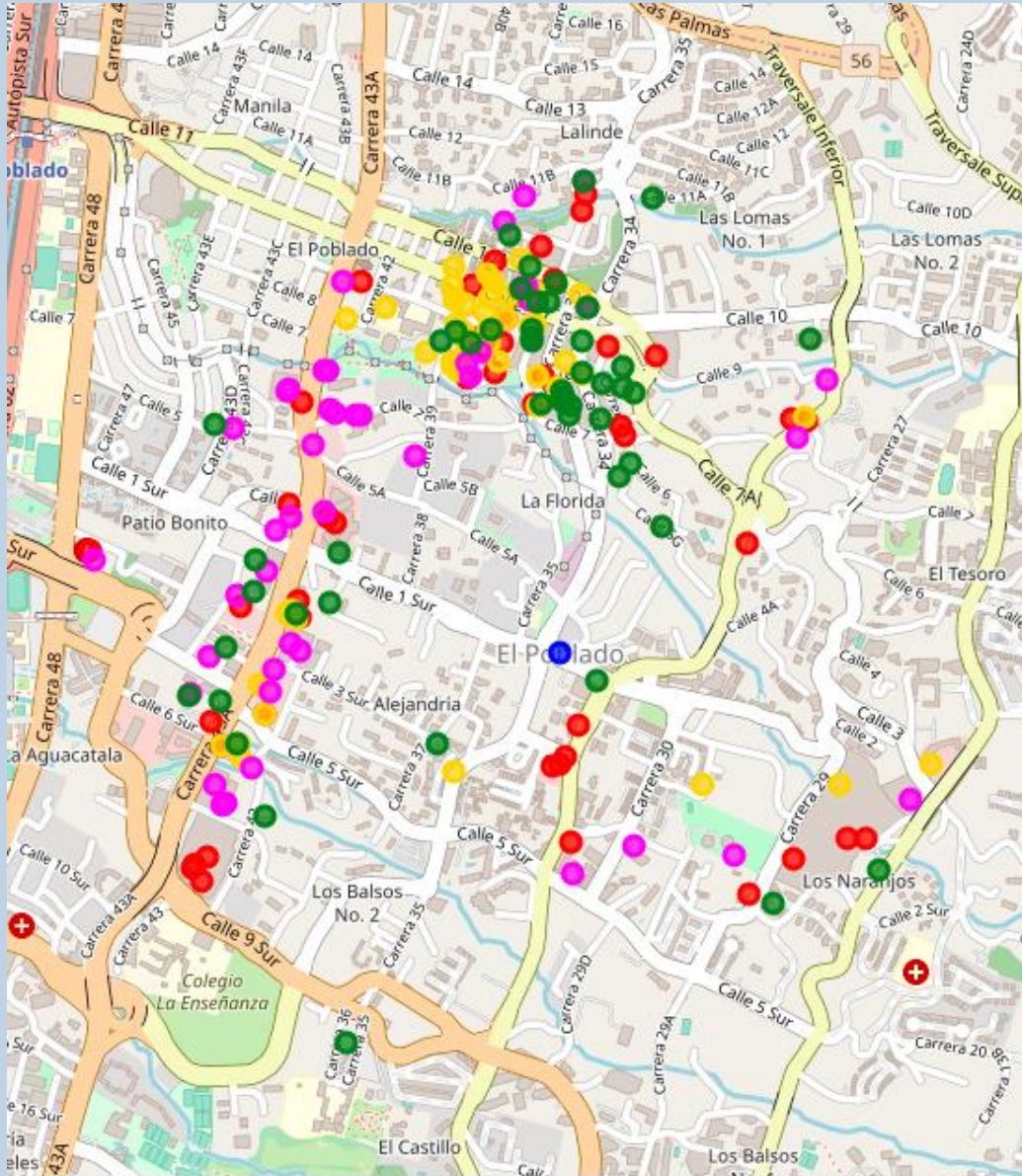
Weight offices = 0.25

Positive weight because drunk people are the best customers for greasy food

Weight bars = 0.4

Positive weight because tourist always want to taste a cultural dish and a lot like to party
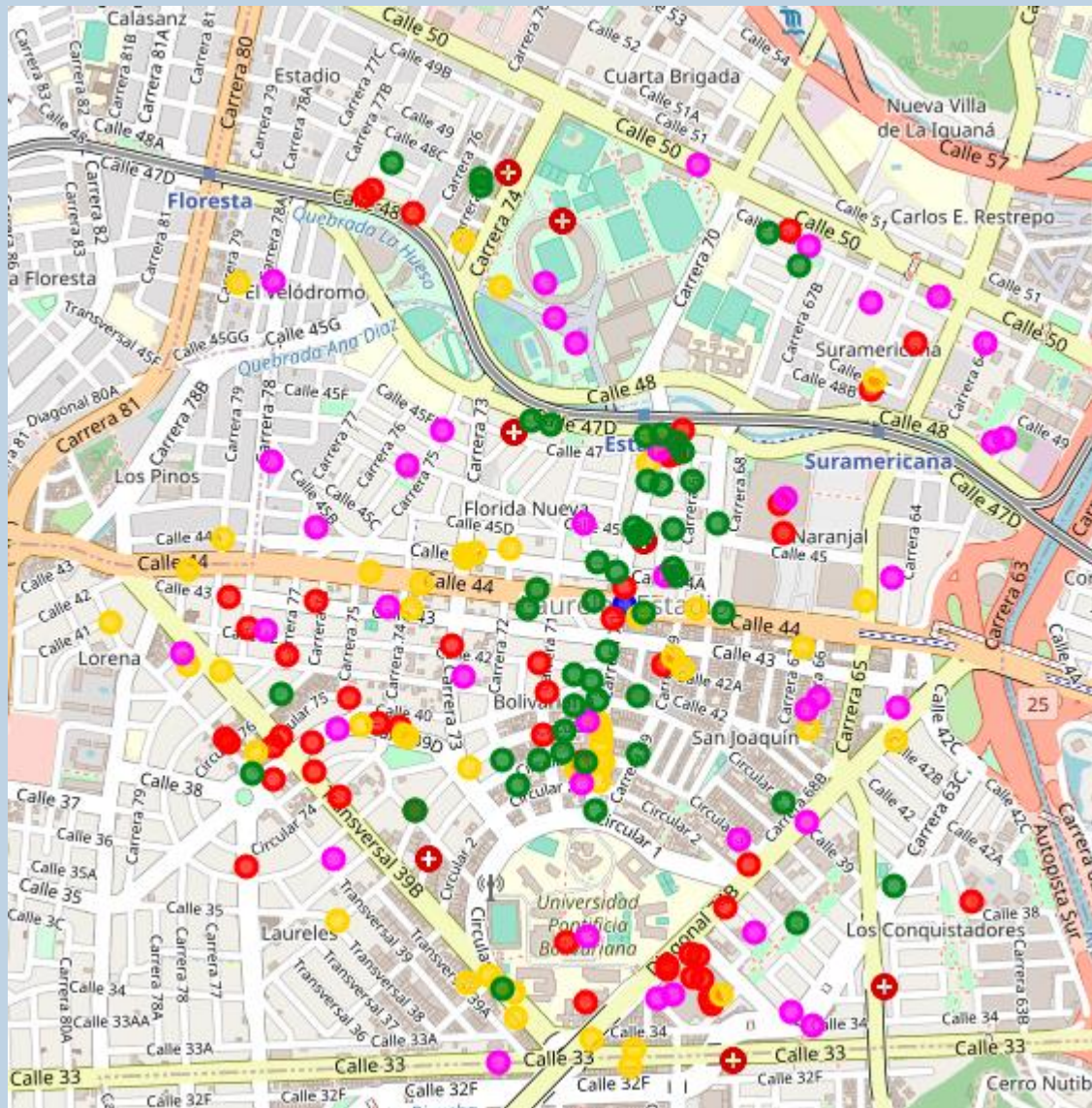
Weight hotels = 0.35

With this a score was calculated.

Data processing

The Comuna with the highest score was El Poblado, with a score of 53.05.

Data processing

With the second highest score, Comuna 11, Laureles-Estadio with a score of 51,25.

Conclusion

◦ Assigning a weight to each Comuna according to the amount of venues that are of interest, two Comunas gave a very high score, El Poblado with 53 points and Laureles-Estadio with 51 points, looking at the distribution of the venues of each Comuna, El Poblado have the venues concentrated in one area, this would be very helpful for people traffic and for the business to get exposed, for this reason the best ubication for Andrea venture is El Poblado and should have Laureles-Estadio on stand-by for future expansions.