Capstone project – final presentation

Sally Hany Café Shop Project

Introduction and Business Problem

• A customer, Sally, wants to open a new Café Shop in Bogotá, Colombia.

• Due to Bogotá's high diversity and very large size, she asked me for help in order to find the best spot to place the Café Shop.

• Bogotá has 20 different Localities (Districts) and we aim to find the best one.

• We need to choose a Locality that has good amount of customers and low amount of competition.

Data

• To help Sally in her search we will need to access following data:

• The Localities of Bogotá, Colombia from Wikipedia: https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1

• The coordinates (latitude, longitude) at these Localities of Bogotá from Open Street Map APIs

From Foursquare we will need following venues data:

• The Café Shops venues of the Localities

• The Offices venues of the Localities

• The High schools venues of the Localities

• The Universities venues of the Localities

• We will then leverage the data in order to determine which locality is the most appropriate in order to locate the Café Shop.

Methodology

• For each locality, all office, school, university and Café Shops venues data have been collected from Foursquare.

• Then for each locality, the sums of the office, school, university and Café Shops were computed.

• For each of this 4 categories, a weight (or penalty) has been defined according to what Sally considers the most important.

• Café Shops have been weighted with -1, since Sally wants to avoid concurrence.

• Schools have been weighted with 1, since student are good customers.

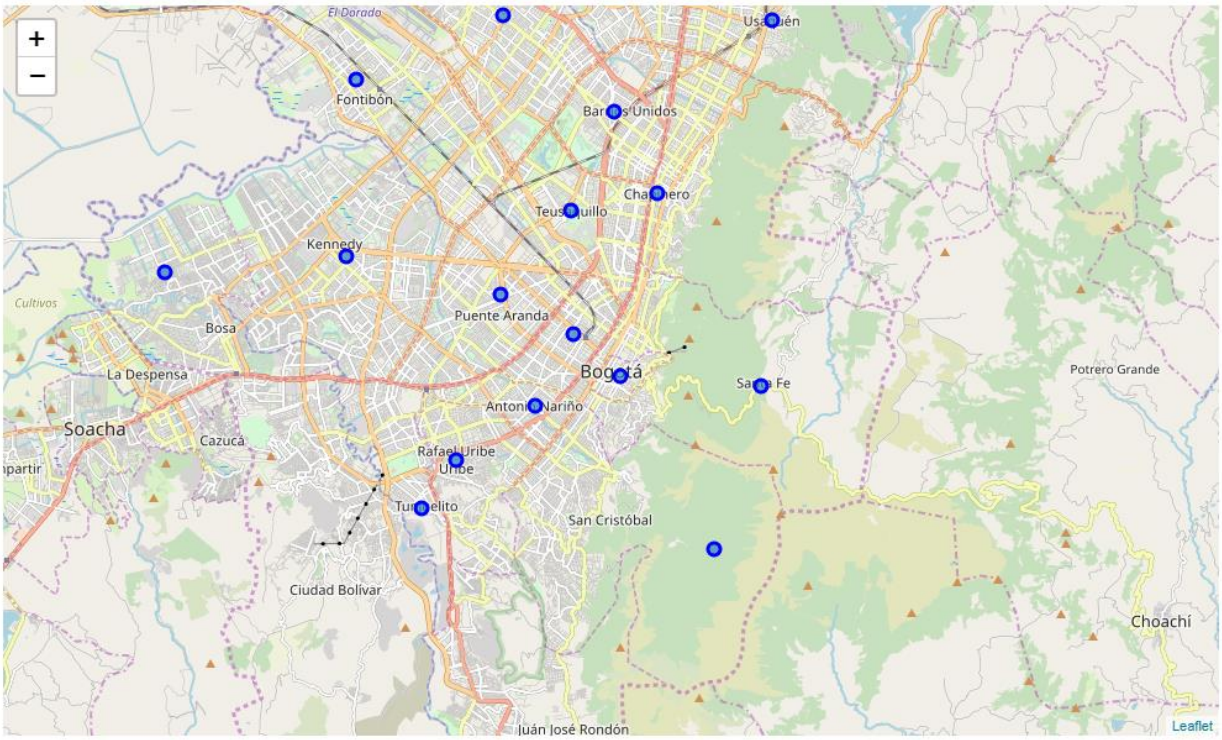
• Universities have been weighted with 1.5, since students are good customers.

• Offices have been weighted with 2, since employees are even better customers.

• Note that the weights can be modified according to the importance of each category.

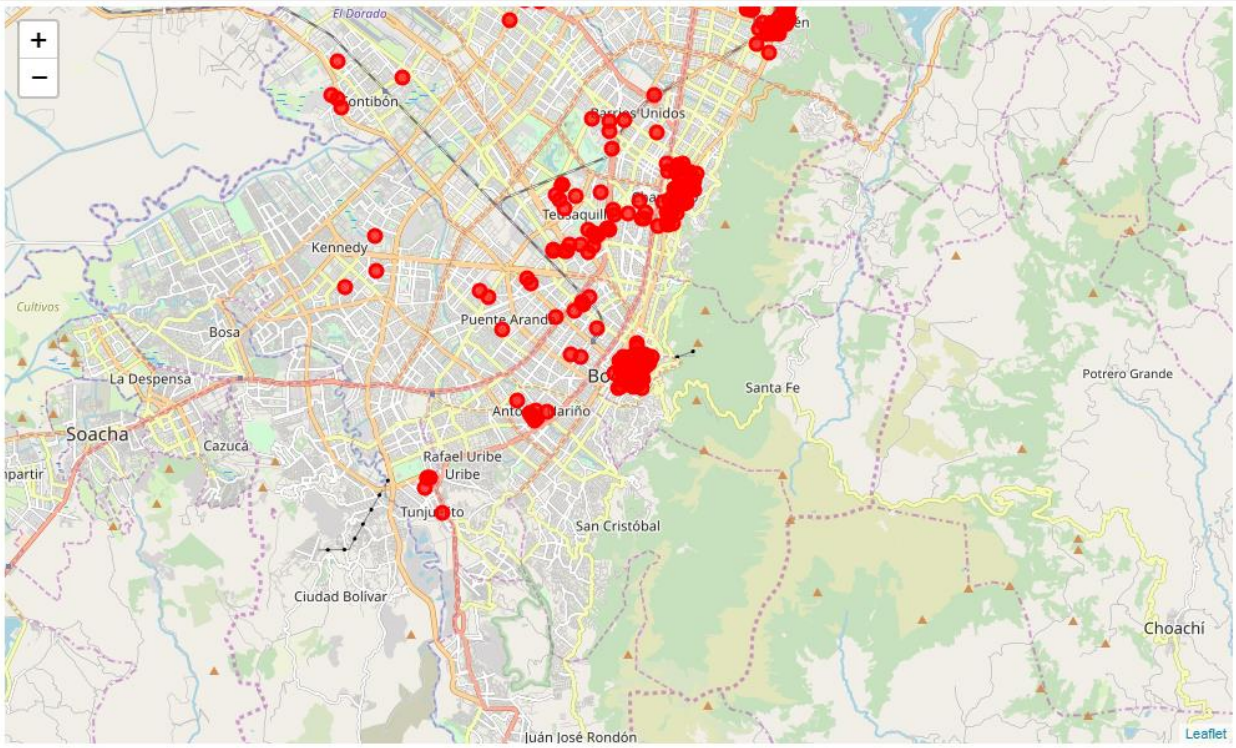
• Lastly, a score was computed for each locality as the weighted sum of the number of venues in each of the 4 categories (school, university, office, Café Shops).

Localities of Bogotá, Colombia

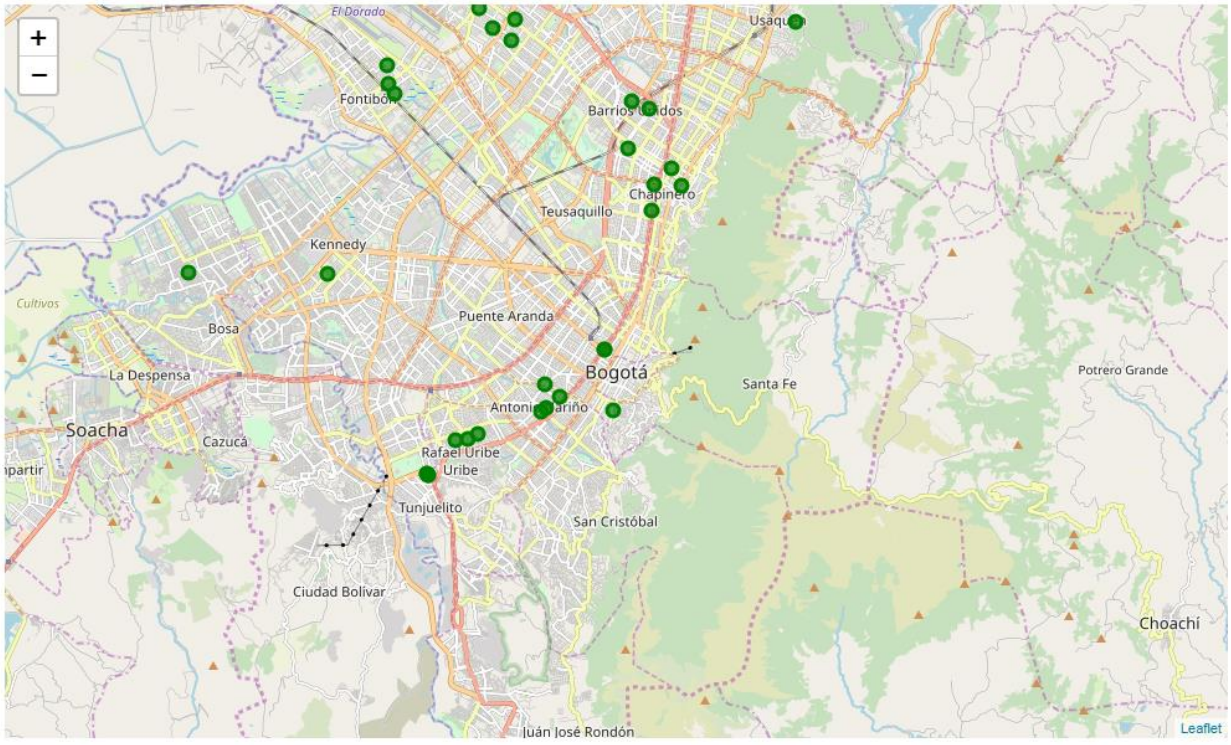


	Localidades	Latitude	Longitude
0	Usaquén	4.695047	-74.031493
1	Chapinero	4.647120	-74.063458
2	Santa Fe (Bogotá)	4.593766	-74.034314
3	San Cristóbal (Bogotá)	4.548658	-74.047473
4	Usme	4.411136	-74.129108
5	Tunjuelito	4.560148	-74.128922
6	Bosa (Bogotá)	4.625492	-74.200280
7	Kennedy (Bogotá)	4.629682	-74.149935
8	Fontibón	4.678737	-74.146988
9	Engativá	4.696628	-74.106120
10	Suba	4.761197	-74.082518
11	Barrios Unidos (Bogotá)	4.669679	-74.075483
12	Teusaquillo	4.642343	-74.087217
13	Los Mártires	4.608375	-74.086538
14	Antonio Nariño (Bogotá)	4.588253	-74.097455
15	Puente Aranda	4.619234	-74.106763
16	La Candelaria	4.596515	-74.073492
17	Rafael Uribe Uribe (Bogotá)	4.573490	-74.119208
18	Ciudad Bolívar (Bogotá)	4.492632	-74.143119
19	Sumapaz (Bogotá)	4.097379	-74.342307

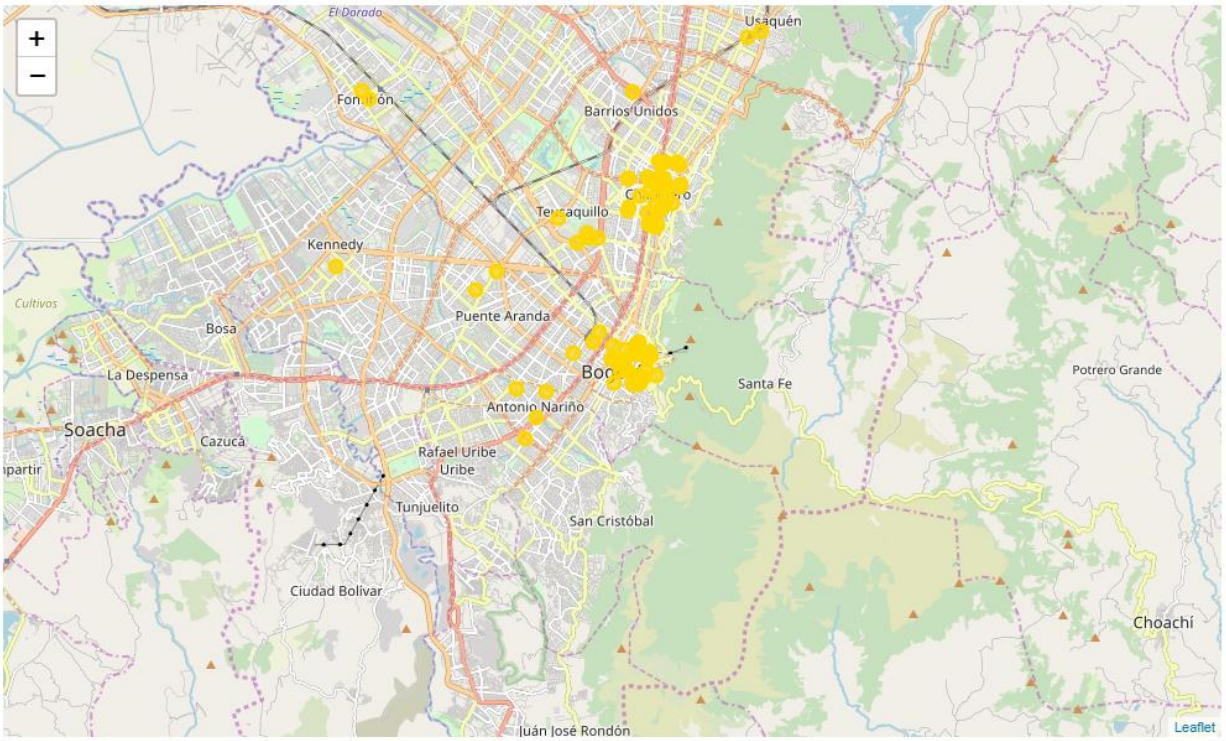
Café Shops in Bogotá Localities



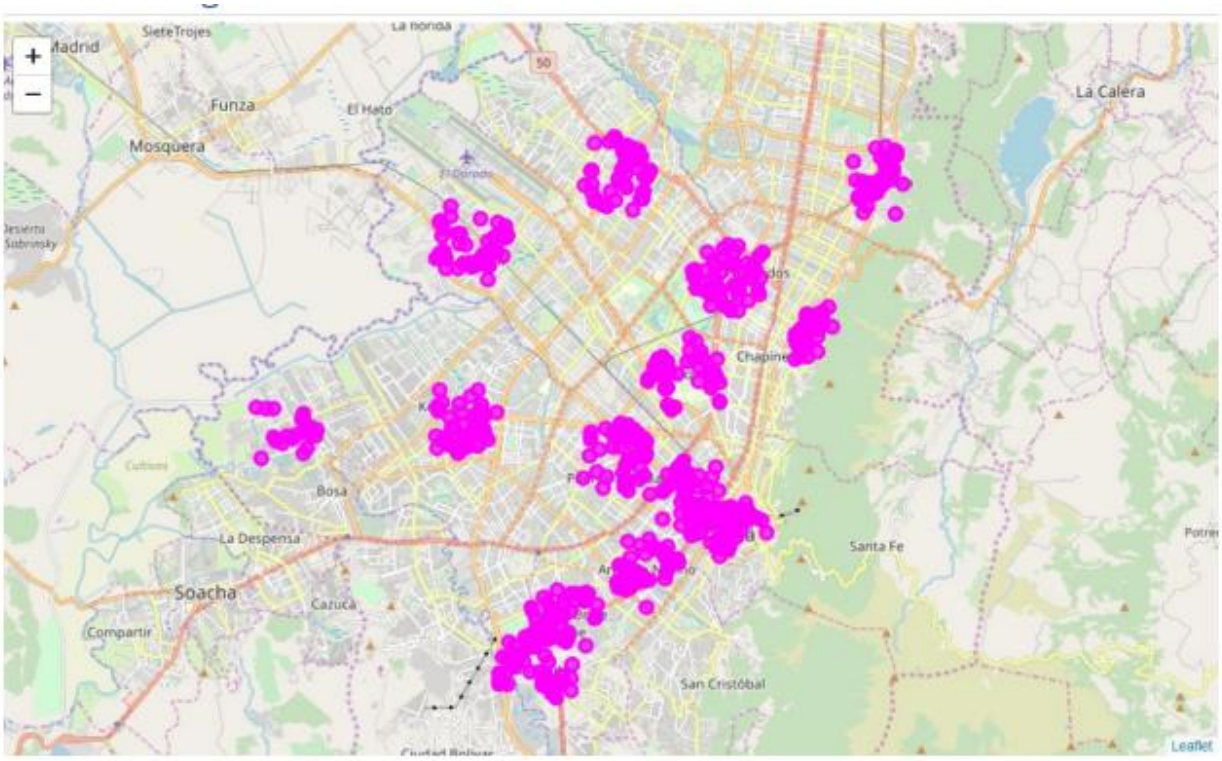
High Schools in Bogotá Localities



Universities in Bogotá Localities



Offices in Bogotá Localities



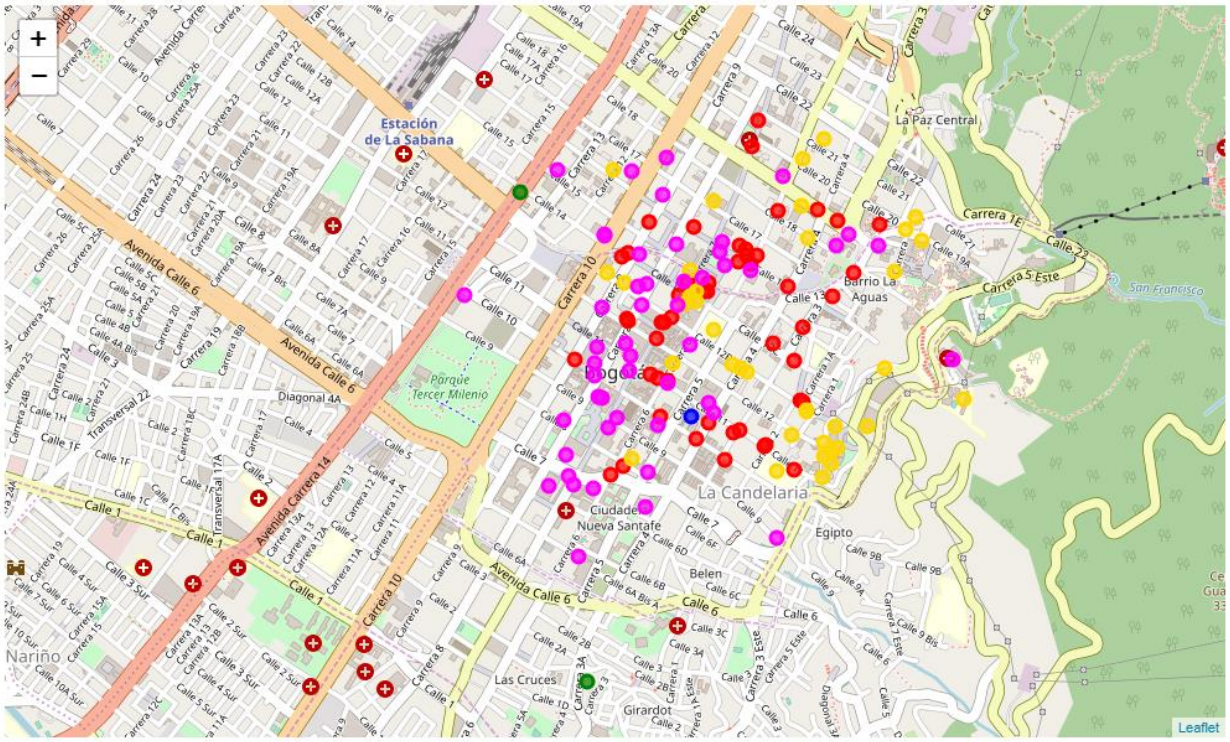
Results

	Localidad	Score
1	Chapinero	110.0
16	La Candelaria	107.5
9	Engativá	98.5
14	Antonio Nariño (Bogotá)	98.0
11	Barrios Unidos (Bogotá)	95.5
17	Rafael Uribe Uribe (Bogotá)	95.0
15	Puente Aranda	94.0
8	Fontibón	93.0
7	Kennedy (Bogotá)	91.5
13	Los Mártires	90.5
5	Tunjuelito	86.0
12	Teusaquillo	81.5
0	Usaquén	57.5
10	Suba	54.0
6	Bosa (Bogotá)	45.0
4	Usme	0.0
3	San Cristóbal (Bogotá)	0.0
2	Santa Fe (Bogotá)	0.0
18	Ciudad Bolívar (Bogotá)	0.0
19	Sumapaz (Bogotá)	0.0

- The Locality with the best score is “Chapinero” with 110.0, being the best option.
- Follows closely “La Candelaria” with 107.5.

These options maximize the number of potential customers from offices and universities and at the same time have not too large competence.

Best Place for the Café Shop in Bogotá is "Chapinero"



Recommendation The following analysis can be improved with following extensions:

• Consider more categories. For example like "Night life" which is also a good source for customers. But also like "Restaurants", which even if not Café Shops may be some concurrence if too many.

• In the Locality itself, it can also be computed the distance between all the venues in order to find a place with the most number of potential customers.

• Using smaller geographical areas like Neighborhoods could improve the accuracy for the scores.