

# Capstone Project

Isha Singla

## Introduction and Business Problem:

- A customer, Tom, wants to open a new burger joint in Bogotá, Colombia.
- Due to Bogotá's high diversity and very large size, he asked me for help in order to find the best spot to place the burger joint.
- 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:

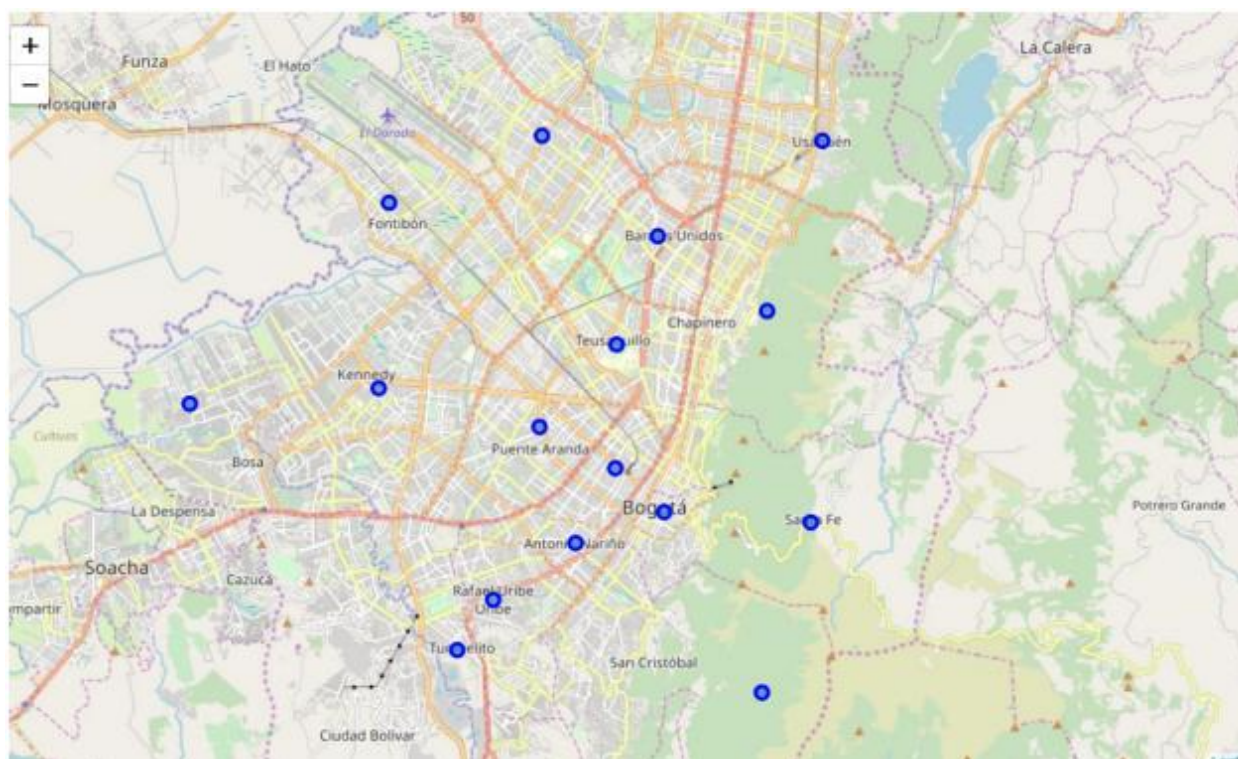
- To help Tom in his search we will need to access following data:
- The Localities of Bogotá, Colombia from Wikipedia:  
[https://es.wikipedia.org/wiki/Anexo:Localidades\\_de\\_Bogot%C3%A1](https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1)
- The coordinates (latitude, longitude) of these Localities of Bogotá from Open Street Map APIs
- From Foursquare we will need following venues data:
- The burger joint 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 burger joint.

## Methodology:

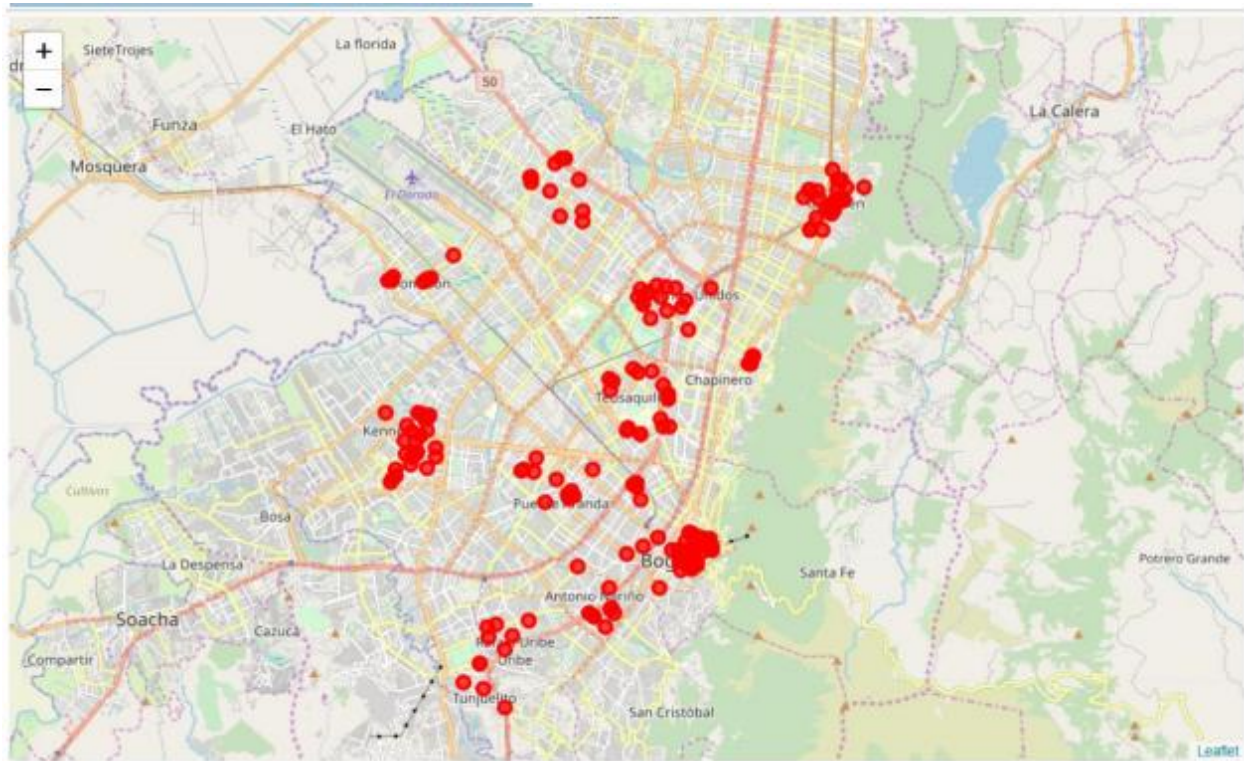
- For each locality, all office, school, university and burger joints venues data have been collected from Foursquare.
- Then for each locality, the sums of the office, school, university and burger joints were computed.
- For each of this 4 categories, a weight (or penalty) has been defined according to what Tom considers the most important.
- Burger Joints have been weighted with -1, since Paolo 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, burger joints).

## Localities of Bogotá, Colombia:



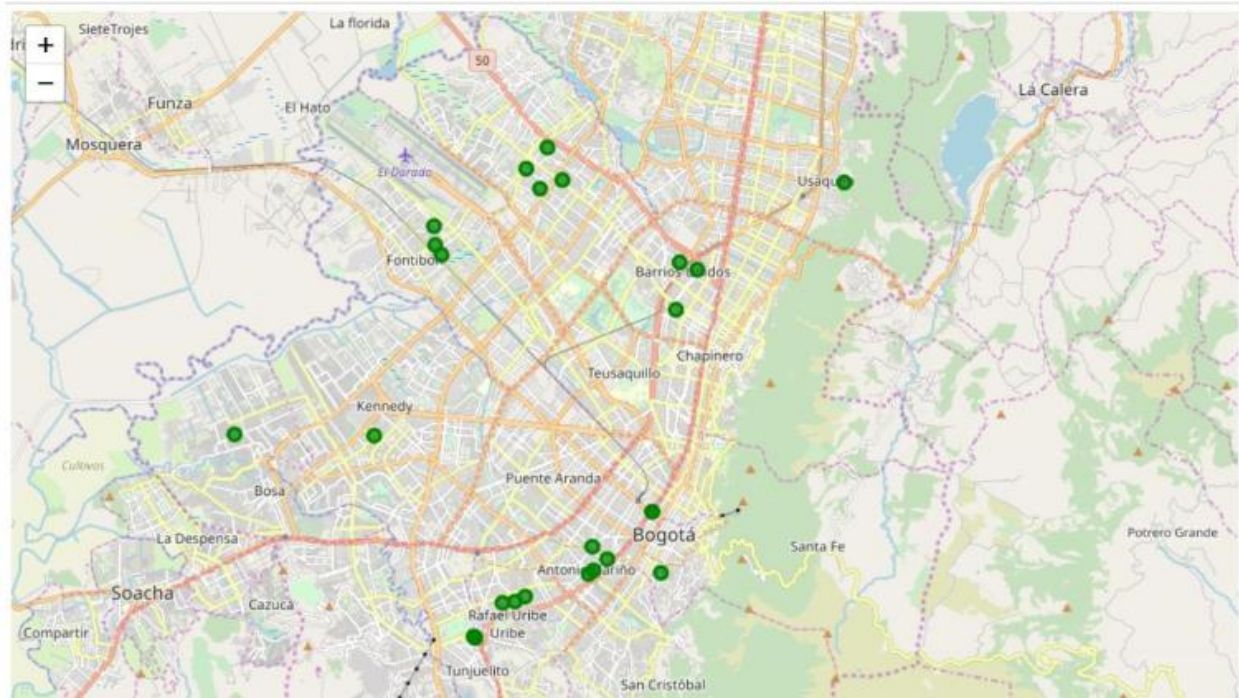
	Localidades	Latitude	Longitud
0	Usaquén	4.695047	-74.031493
1	Chapinero	4.649917	-74.046269
2	Santa Fe (Bogotá)	4.593766	-74.034314
3	San Cristóbal (Bogotá)	4.548658	-74.047473
4	Usme	4.411136	-74.129108
5	Tunjuelito	4.560148	-74.128922
6	Bosa (Bogotá)	4.625515	-74.200328
7	Kennedy (Bogotá)	4.629451	-74.149927
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	4.641244	-74.086336
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)	4.098800	-74.341530

## Burger Joints in Bogotá Localities

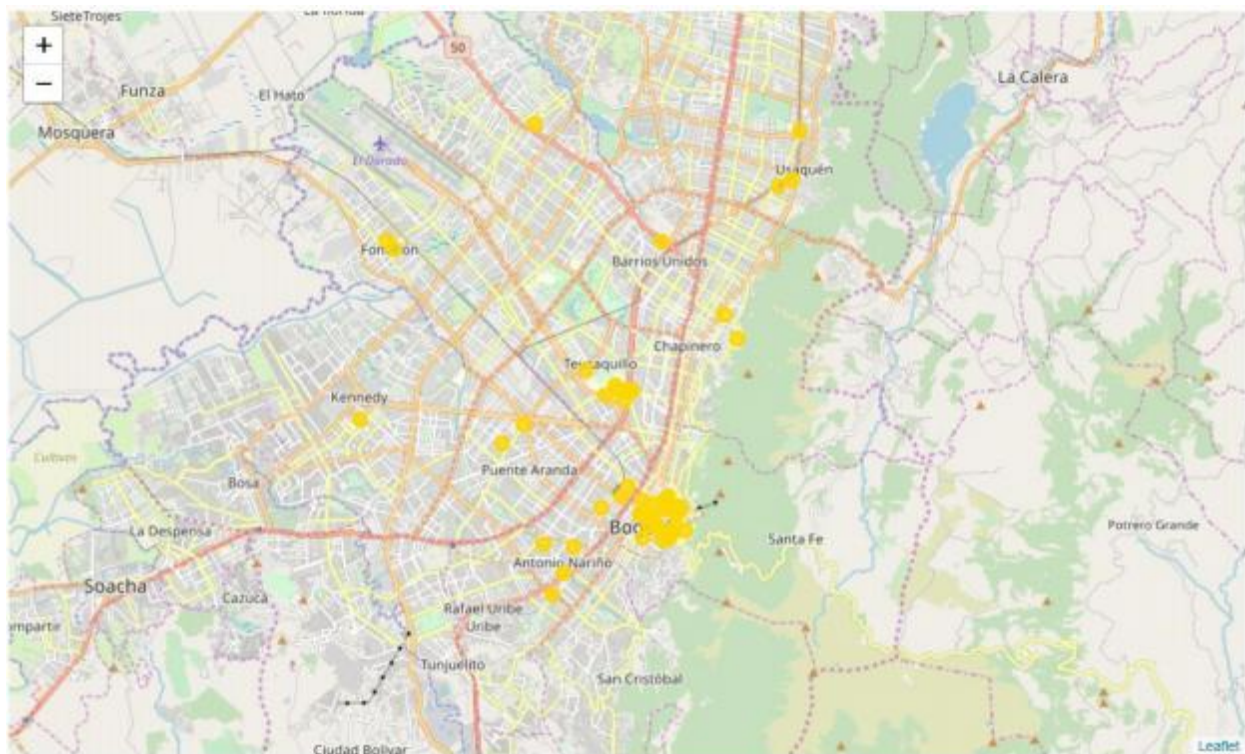


## High School in Bogota Localities:

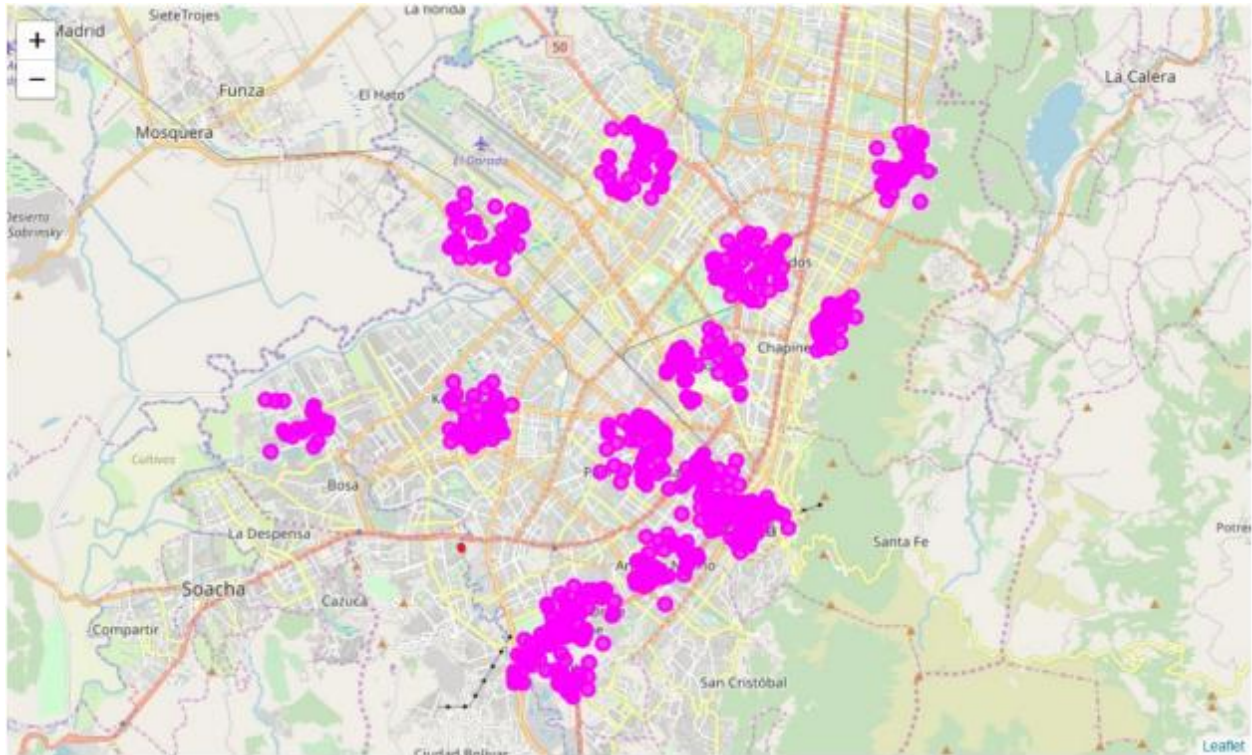




## Universities in Bogotá Localities:



## Offices in Bogotá Localities:



Results:

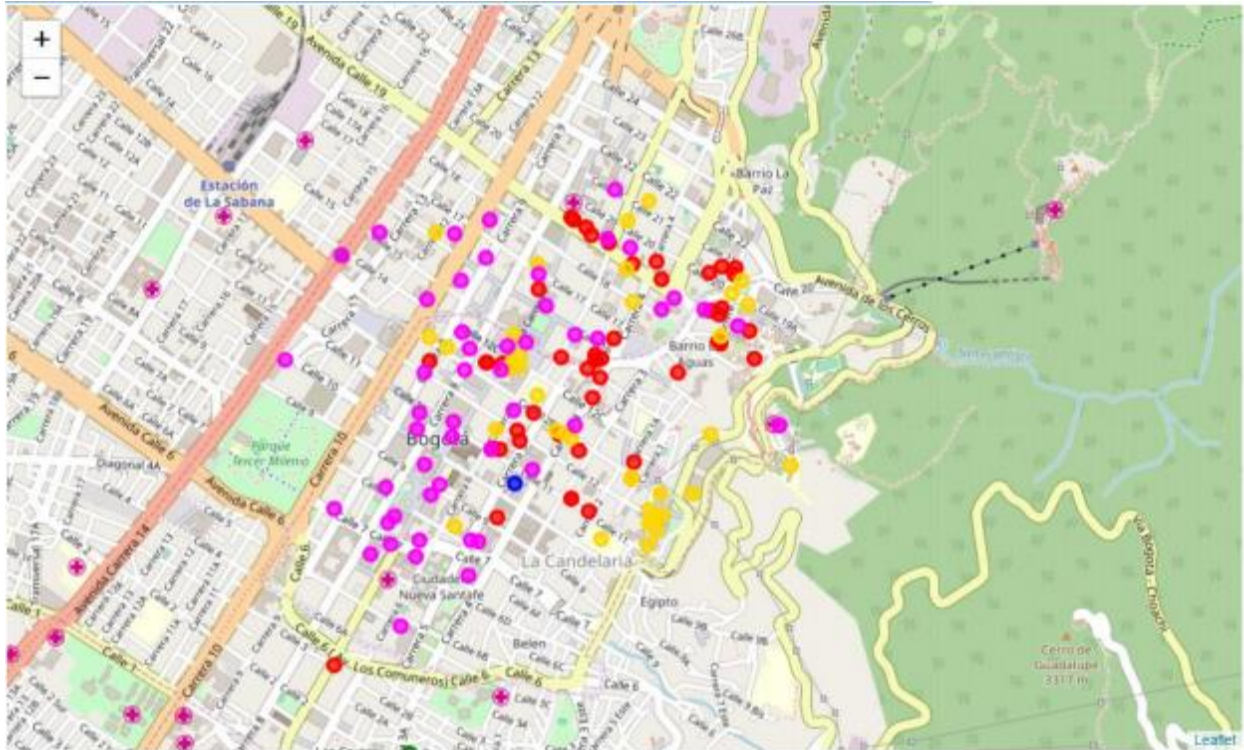
	Localidad	Score
16	La Candelaria	109.5
14	Antonio Nariño (Bogotá)	101.0
8	Fontibón	99.0
17	Rafael Uribe Uribe (Bogotá)	98.0
5	Tunjuelito	98.0
1	Chapinero	97.0
9	Engativá	94.5
12	Teusaquillo	94.0
13	Los Mártires	93.5
15	Puente Aranda	92.0
11	Barrios Unidos (Bogotá)	87.5
7	Kennedy (Bogotá)	75.5
0	Usaquén	74.5
10	Suba	54.0
6	Bosa (Bogotá)	49.0
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 “La Candelaria” with 109.5, being the best option.
- Follows closely “Antonio Nariño” with 101.

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 Burger Joint in Bogotá is “La Candelaria”:**





## 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 burger joints 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.