CAPSTONE PROJECT - FINAL REPORT

1.Introduction where you discuss the business problem and who would be interested in this project.

-A person, Robert wants to open a new BBQ joint in bogota, colombia.

-Due to bogota's high diversity and very large size. he asked for help in order to find best place to open a new BBQ joint in bogota, colombia.

-Bogota's has 20 different localities (districts) and we aim to find best place.

-we need to choose best locality that has good amount of customers and low amount of competition.

2.Data where you describe the data that will be used to solve the problem and the source of the data.

-The localities of bogota, colombia from wikipedia https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1

-Import the coordinates (latitude, longitude) ot these Localities of Bogotá from Open Street Map APIs

-From Foursquare we will need following venues data:

->The BBQ joint venues of the Localities

->The offices 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 BBQ joint

3.Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.

-For each locality , all office employees, universities and BBQ joint venues data have been collected from foursquare.

-Then for each locality, the sums of the office, university and BBQ joints were computed.

-For each of this 3 categories , a weight has been defined according to what robert's considers the most important.

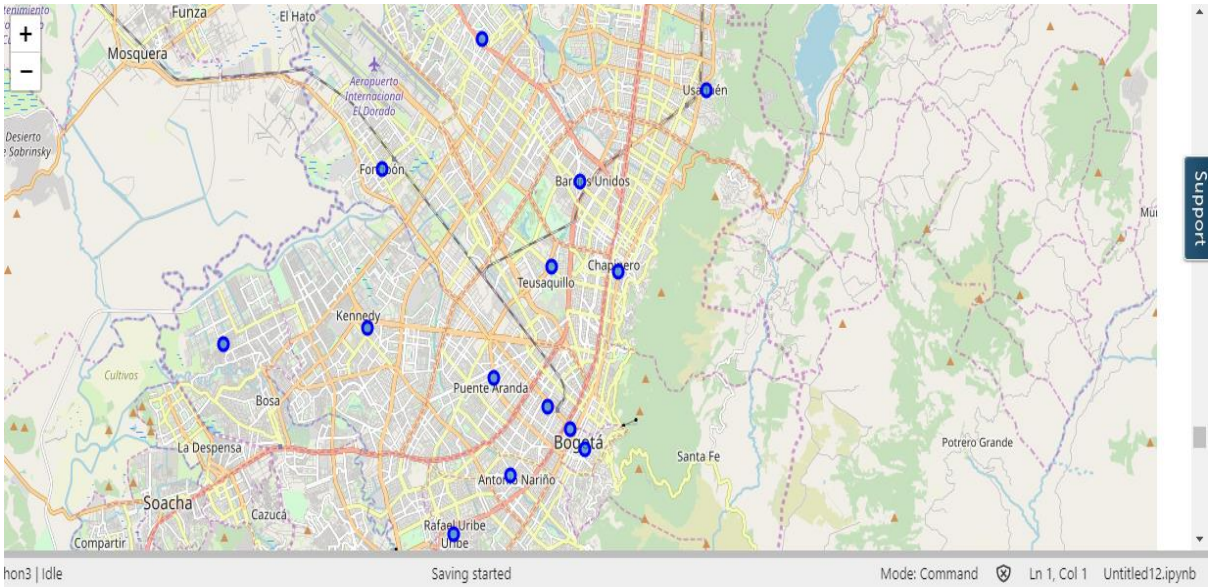
-BBQ joints have been weighted with -1 , since robert wants to avoid concurrence. -universities have been weighted with 1.5, since students are good customers.

-offices have been weighted with 2 , since employees are even better customers.

NOTE-weights are modified according to 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 3 categories

LOCALITIES OF BOGOTA , COLOMBIA:



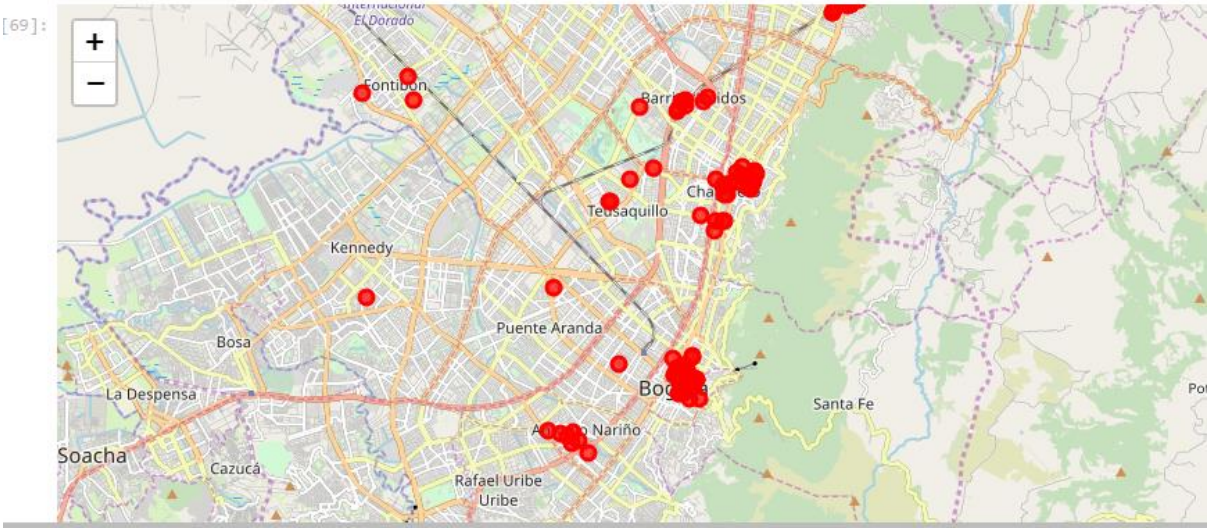
.3]:

	Localidades	Latitude	Longitude
0	Usaquén	4.694969	-74.031093
1	Chapinero	4.645377	-74.061943
2	Santa Fe (Bogotá)	4.602204	-74.078837
3	San Cristóbal (Bogotá)	4.548658	-74.047473
4	Usme	4.411136	-74.129108
5	Tunjuelito	4.561049	-74.127523
6	Bosa (Bogotá)	4.625492	-74.200280
7	Kennedy (Bogotá)	4.629682	-74.149935
8	Fontibón	4.673327	-74.144732
9	Engativá	4.708695	-74.109643
10	Suba	4.761197	-74.082518
11	Barrios Unidos (Bogotá)	4.669679	-74.075483
12	Teusaquillo	4.646410	-74.085441
13	Los Mártires	4.608375	-74.086538
14	Antonio Nariño (Bogotá)	4.589429	-74.099568

BBQ JOINTS IN BOGOTA, LOCALITIES:

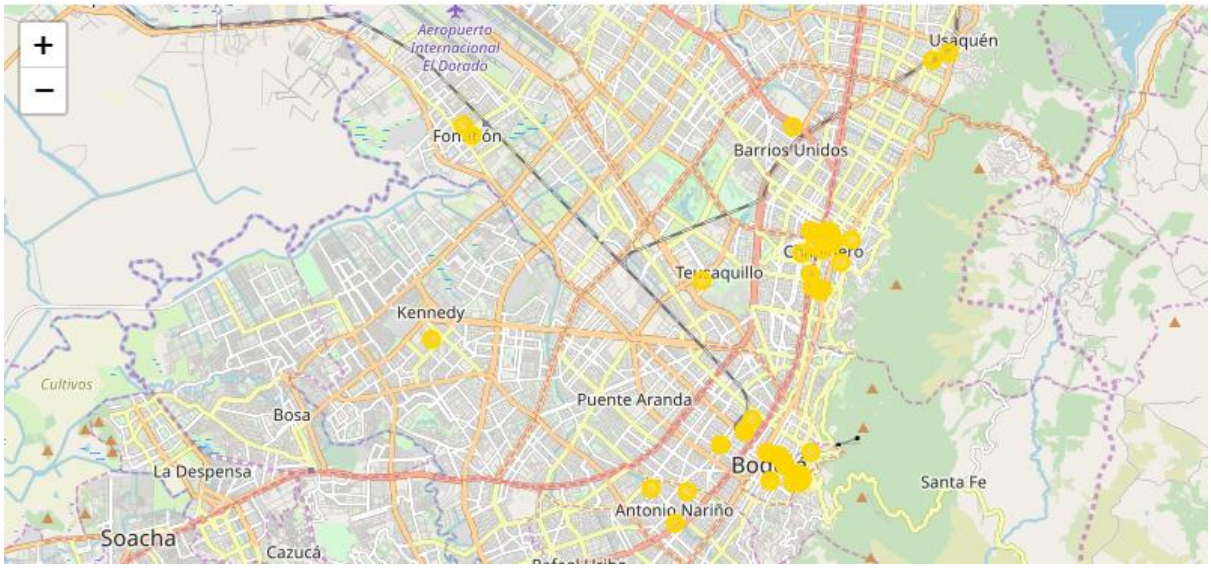
[69]:
```
map_bogota_BBQjoint = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(bogota_venues_BBQjoint, 'red', map_bogota_BBQjoint)
map_bogota_BBQjoint
```

[69]:



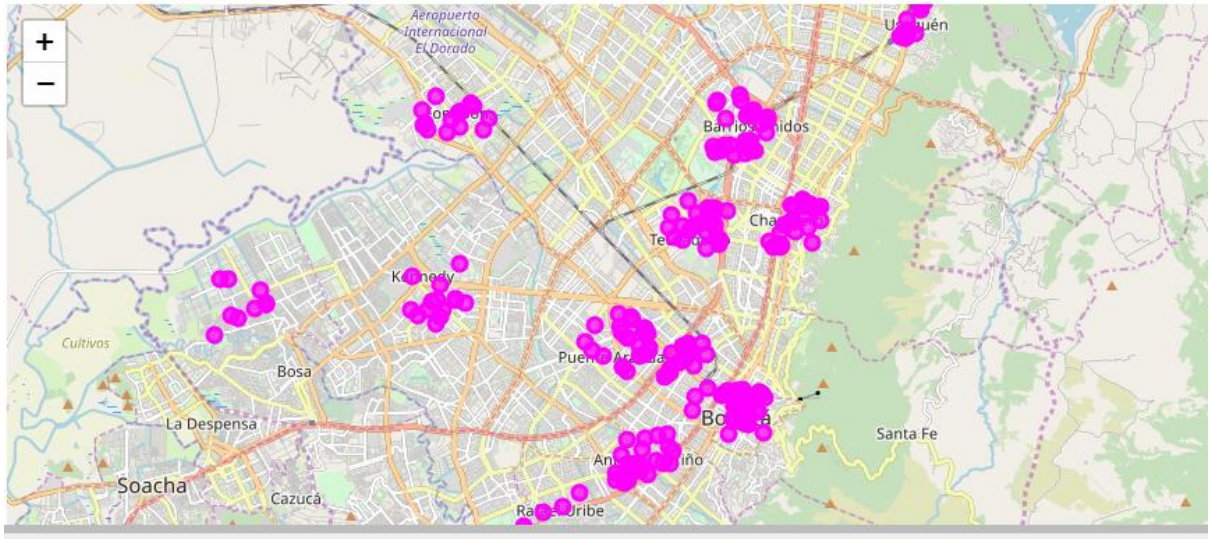
UNIVERSITIES IN BOGOTA, LOCALITIES:

```
map_bogota_university = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(bogota_venues_university, 'gold', map_bogota_university)
map_bogota_university
```



OFFICES IN BOGOTA LOCALITIES:

```
map_bogota_office = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(bogota_venues_office, 'fuchsia', map_bogota_office)
map_bogota_office
```



4.Results section where you discuss the results.

RESULTS:

```
df_weighted['Score'] = df_data['BBQjoint'] * weight_BBQjoint + df_data['Universities'] * weight_university + df_data['Offices'] * weight_offices
df_weighted = df_weighted.sort_values(by=['Score'], ascending=False)
df_weighted
```

	Localidad	Score
16	La Candelaria	56.5
1	Chapinero	50.5
14	Antonio Nariño (Bogotá)	47.5
12	Teusaquillo	47.5
15	Puente Aranda	47.0
11	Barrios Unidos (Bogotá)	43.5
2	Santa Fe (Bogotá)	42.0
13	Los Mártires	39.5
8	Fontibón	30.0
9	Engativá	29.0
7	Kennedy (Bogotá)	28.5
0	Usaquén	28.0

The Locality with the best score is “la candelaria” with 56.5, being the best option

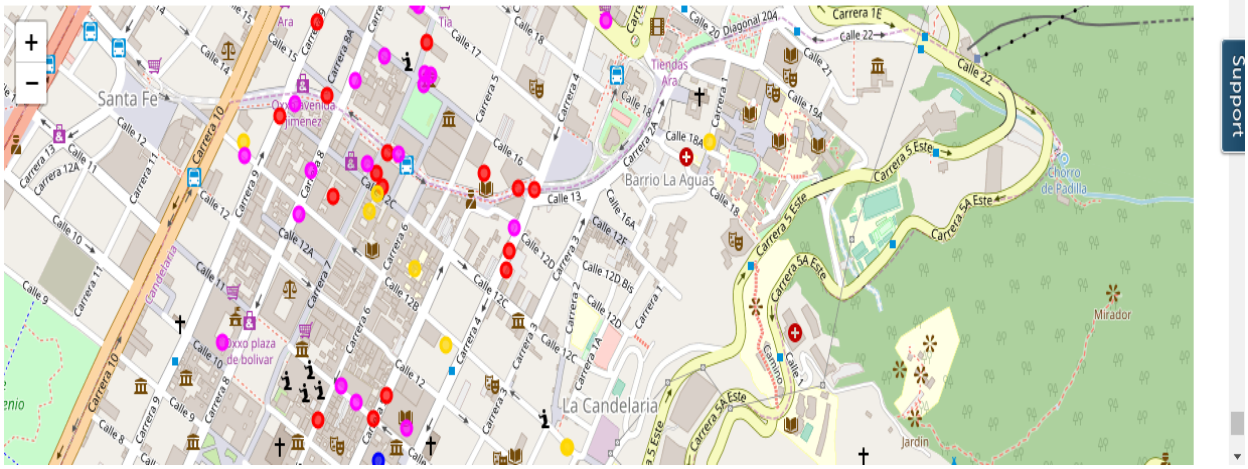
Follows closely “chapinero” with 50.5.

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 BBQ JOINT IS LA CANDELARIA:

```
addToMap(bogota_venues_BBQjoint[bogota_venues_BBQjoint['Localidad'] == 'La Candelaria'], 'red', map_bog_result)
addToMap(bogota_venues_university[bogota_venues_university['Localidad'] == 'La Candelaria'], 'gold', map_bog_result)
addToMap(bogota_venues_office[bogota_venues_office['Localidad'] == 'La Candelaria'], 'fuchsia', map_bog_result)

map_bog_result
```



5.Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

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 BBQ 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.

5.Conclusion section where you conclude the report.

CONCLUSION:

-*La Candelaria Locality is the best option for robert in order to open his BBQ Joint.*