THE BATTLE OF NEIGHBORHOODS

IBM DATA SCIENCE PROGRAM – CAPSTONE PROJECT

JAIRO IVAN ORDOÑEZ

DECEMBER 2019

BUSINESS PROBLEM

A Colombian freelancer is looking for a Neighborhood to live in Cartagena or Bogotá, however, he needs these special facilities:

- *Vegetarian or Vegan Restaurant, because He cannot eat animal protein (meat, fish, chicken, etc.)*

- *A gym or Public Park to make physic exercise.*

- *Farmer market to buy the vegetables and fruit for his diet.*

- *Pharmacy to buy medicament to his allergy*

The Colombian ask to help for us to, with cluster models applied in the Data Science Program, we can recommend him the Neighborhoods than accommodate to his requests.

***Note**: We will not evaluate the security or living price in the Neighborhoods, only the closer venues.*

DATA

```
In [119]:    import pandas as pd
             data = pd.read_csv("C_digos_Postales_Nacionales.csv")

In [120]:    data.head()

Out[120]:
```

	noid	codigo_departamento	nombre_departamento	codigo_municipio	nombre_municipio	zona_postal	codigo_postal	limite_norte	limite_sur	limite_e
0	877	15	BOYACA	15621	RONDON	1534	153420	LIMITE URBANO	LIMITE URBANO	LIM URBA
1	1765	25	CUNDINAMARCA	25592	QUEBRADANEGRA	2534	253427	MUNICIPIO UTICA	MUNICIPIO VILLETA	MUNICI NIMAI
2	228	5	ANTIOQUIA	5002	ABEJORRAL	550	55030	LIMITE URBANO	LIMITE URBANO	LIM URBA
3	3540	88	ARCHIPIELAGO DE SAN ANDRES	88001	SAN ANDRES	8800	880007	LIMITE URBANO	SAN LUIS	M CAR
4	1963	41	HUILA	41013	AGRADO	4140	414040	LIMITE URBANO	LIMITE URBANO	K

DATA

BOGOTA

```
In [123]: ▶ bogotaData.head()
```

Out[123]:

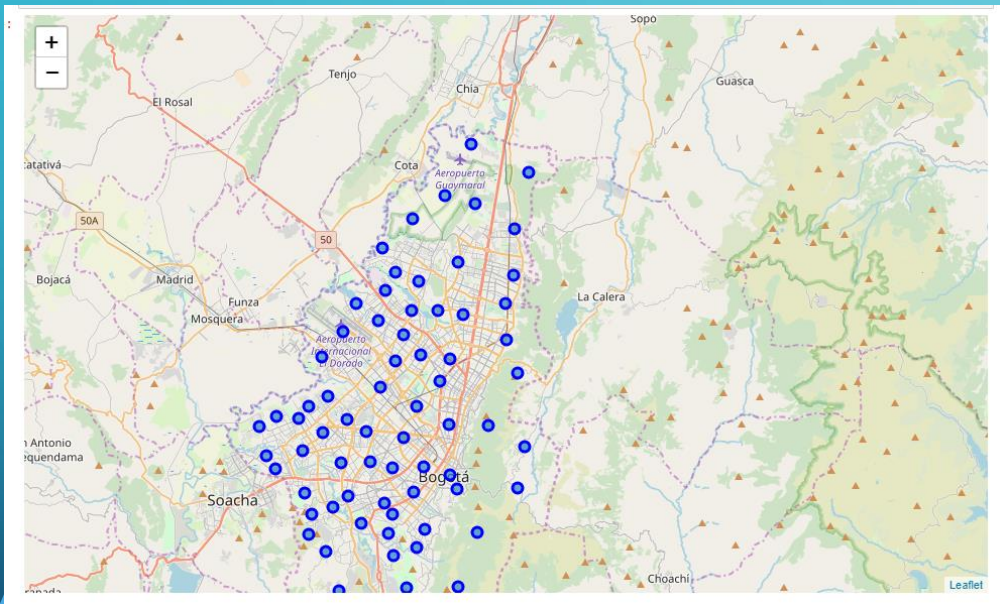|	noid	codigo_departamento	nombre_departamento	codigo_municipio	nombre_municipio	zona_postal	codigo_postal	limite_norte	limite_sur	lir
181	480	11	BOGOTA D.C.	11001	LOCALIDAD USME	1105	110561	RIO TUNJUELITO	VEREDA ARRAYAN	MU CH
239	536	11	BOGOTA D.C.	11001	LOCALIDAD CIUDAD BOLIVAR	1119	111971	CL 97 S	BARRIO MUCHUELO ALTO RURAL	TUN
324	465	11	BOGOTA D.C.	11001	LOCALIDAD USAQUEN	1101	110151	LIMITE URBANO	CL 205	
338	540	11	BOGOTA D.C.	11001	LOCALIDAD SUMAPAZ	1120	112031	VEREDA SAN ANTONIO	QUEBRADA EL TUNAL	LAS
574	541	11	BOGOTA D.C.	11001	LOCALIDAD SUMAPAZ	1120	112041	QUEBRADA EL TUNAL	VEREDA SAN JOSE	MU CU

CARTAGENA

```
In [170]: ▶ mergeDF_CTG = pd.merge(left=postalCodeCTG, right=coordCTG_df, how='left', left_on='codigo_postal', right_on='CodigoPostal')
            mergeDF_CTG.drop(['CodigoPostal'], axis=1, inplace=True)
            mergeDF_CTG['Neighborhood']=mergeDF_CTG.nombre_municipio + "-" + mergeDF_CTG.barrios_contenidos_en_el_codigo_postal + "-" + r
            mergeDF_CTG.head()
```
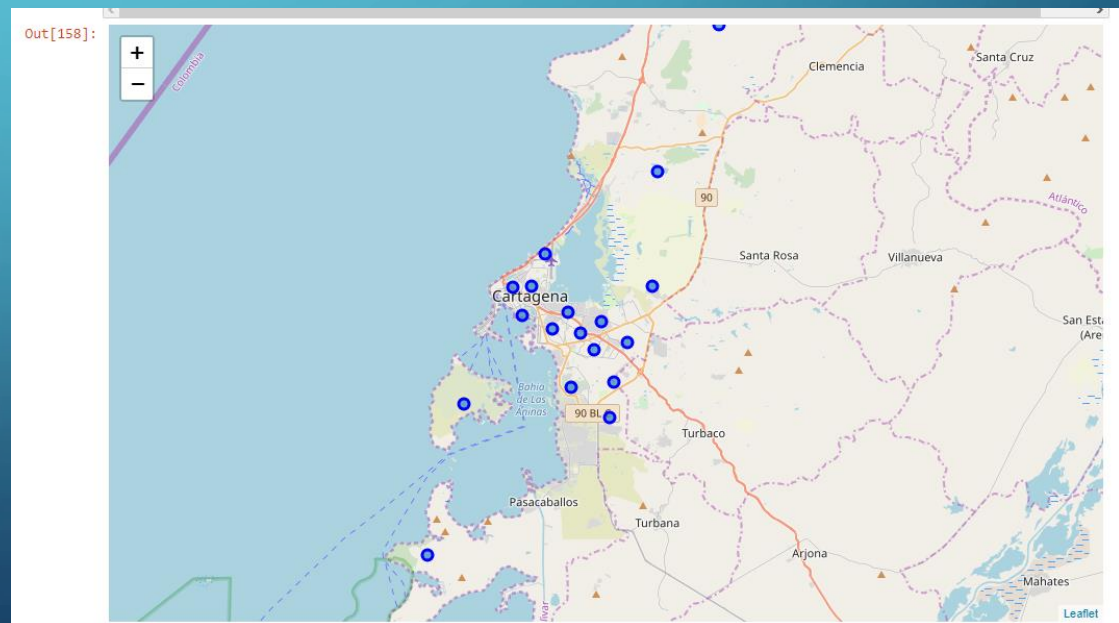
Out[170]:

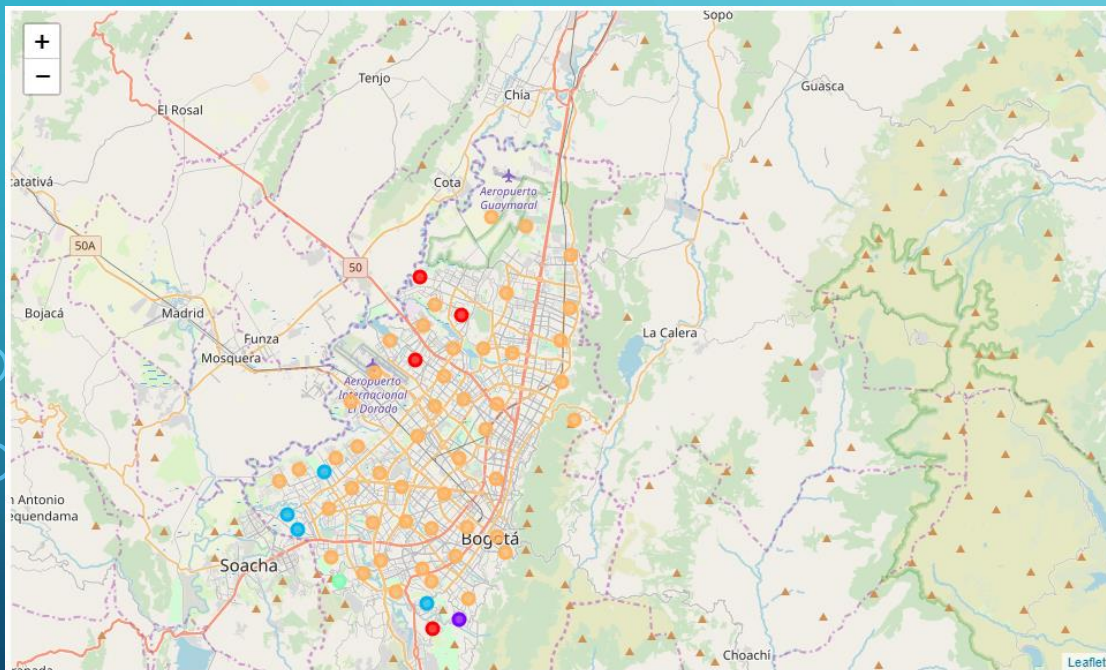|	noid	codigo_departamento	nombre_departamento	codigo_municipio	nombre_municipio	barrios_contenidos_en_el_codigo_postal	codigo_postal	Lati
0	549	13	BOLIVAR	13001	CARTAGENA	La Carolina	130008	10.499
1	560	13	BOLIVAR	13001	CARTAGENA	Sin Informacion de Barrios	130027	10.425
2	558	13	BOLIVAR	13001	CARTAGENA	Sin Informacion de Barrios	130018	10.349
3	559	13	BOLIVAR	13001	CARTAGENA	Sin Informacion de Barrios	130019	10.177
4	550	13	BOLIVAR	13001	CARTAGENA	Sin Informacion de Barrios	130009	10.340

RESULT MAPS
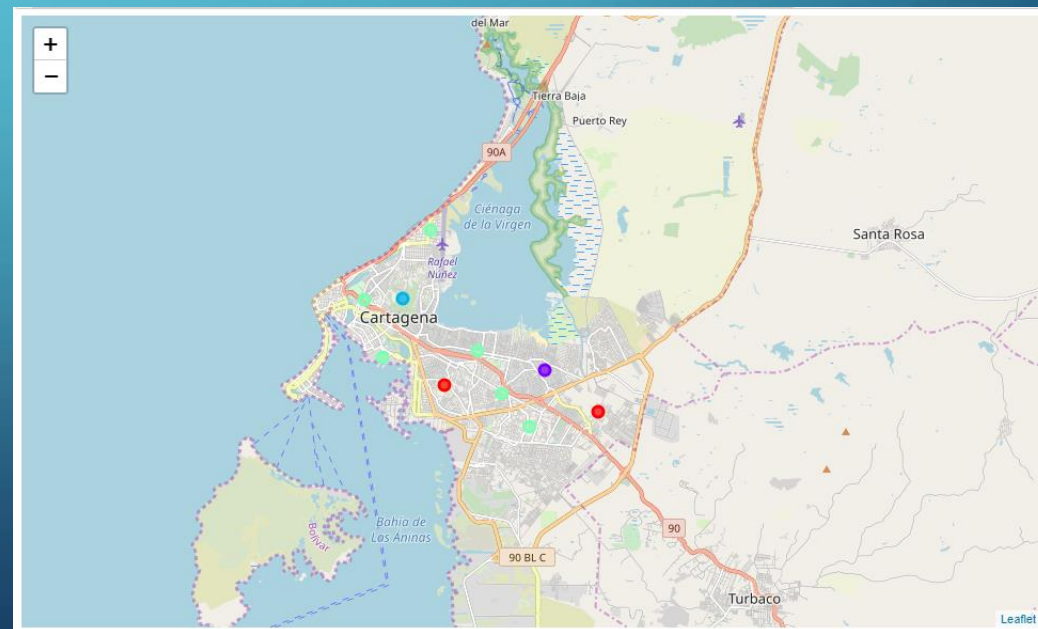
BOGOTA

CARTAGENA

RESULT MAPS - CLUSTERING

BOGOTA

CARTAGENA

CONCLUSION

So, the conclusion to our freelancer Colombian is, if he decides to live in Bogotá:

- *We recommend the "LOCALIDAD SUBA" Neighborhood because its has 3 of the 4 requirements in 2 distinct places.*

- *In "LOCALIDAD BOSA" there are two places with Restaurants Vegetarian or Vegan, but is most common the bars, so if he likes the drink and the dance, we recommend this Neighborhood.*

- *And the most common venue in most of the Neighborhoods in Bogota is the Fast Food Restaurants.*

If he decides to live in Cartagena:

- *In Cartagena there are not Neighborhoods with the special facilities, because it is a coast city and its first product is the fish, so the major of the venues are restaurants but no vegan or vegetarian.*

- *In Cartagena there are not many parks so, if he needs do physic exercise, he would have to swim in the beach.*