A Way to Find The Best Place

Clustering the Neighbourhoods of Bogota

Richard Alejandro Mora Perilla

06 February 2021

Introduction

Normally when people travel they make a plan or a routine to be able to know certain places of the site to which they are going to travel, however, prior to this, people usually look for those relevant places to which they would like to go according to the opinions or comments of the people who are near there.

Business Problem

The problem to be solved is that of which sites could be known according to the good opinions of people from places close to that site, in order to make a decision before making your trip also this data collection can also help to recognize a good place to live in the city.

Data Description

We require geolocation data for both Bogota. The city's zip codes serve as a starting point. Using the zip codes we use we can find the most popular towns, districts, places and their categories of places.

Bogota

To derive our solution, We scrape our data from https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1

This wikipedia page has information about all the localities.

1. *town* : Name of the localitie.
2. *post_code* : Postal codes.

Foursquare API Data

We will need data about different venues. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information

about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

Neighborhood Neighborhood Latitude Neighborhood Longitude Venue Name of the venue e.g. the name of a store or restaurant Venue Latitude Venue Longitude Venue Category

Methodology

We will be creating our model with the help of Python so we start off by importing all the required packages.

In [3]:
```
import pandas as pd
import requests
import numpy as np
import matplotlib.cm as cm
import matplotlib.colors as colors
import folium
from sklearn.cluster import KMeans
```

Localities

We begin to start collecting and refining the data needed for the business solution.

Data Collection

To get the localities in Bogota, we start by scraping the list of areas of Bogota wiki page.

In [2]:
```
1 url = "https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1"
2 wiki = requests.get(url)
3 wiki
```

Out[2]: <Response [200]>

Response 200 means that we are able to make the connection

	Nº	Localidad	Códigos Postales	Superficie km²[2]	Población[3]	Densidad hab/km²
0	1	Usaquén	110111-110151	65.31	501 999	7 686.4
1	2	Chapinero	110211-110231	38.15	139 701	3 661.88
2	3	Santa Fe	110311-110321	45.17	110 048	2 436.3
3	4	San Cristóbal	110411-110441	49.09	404 697	8 243.98
4	5	Usme	110511-110571	215.06	457 302	2 126.39
5	6	Tunjuelito	110611-110621	9.91	199 430	20 124.11
6	7	Bosa	110711-110741	23.93	673 077	28 126.91
7	8	Kennedy	110811-110881	38.59	1 088 443	28 205.31
8	9	Fontibón	110911-110931	33.28	394 648	11 858.41
9	10	Engativá	111011-111071	35.88	887 080	24 723.52
10	11	Suba	111111-111176	100.56	1 218 513	12 117.27
11	12	Barrios Unidos	111211-111221	11.90	243 465	20 459.24
12	13	Teusaquillo	111311-111321	14.19	153 025	10 784
13	14	Los Mártires	111411	6.51	99 119	15 225.65
14	15	Antonio Nariño	111511	4.88	109 176	22 372.12
15	16	Puente Aranda	111611-111631	17.31	258 287	14 921.25
16	17	La Candelaria	111711	2.06	24 088	11 693.2
17	18	Rafael Uribe Uribe	111811-111841	13.83	374 246	27 060.44
18	19	Ciudad Bolívar	111911-111981	130.00	733 859	5 442.83
19	20	Sumapaz	112011-112041	780.96	6 531	8.36

Feature Selection

We need only the boroughs, Postal codes, Post town for further steps. We can drop the locations, dial codes and OS grid.

	Localidad	Códigos_Postales
0	Usaquén	110111-110151
1	Chapinero	110211-110231
2	Santa Fe	110311-110321
3	San Cristóbal	110411-110441
4	Usme	110511-110571

We change the column names for the english language.

	town	post_code
0	Usaquén	110111-110151
1	Chapinero	110211-110231
2	Santa Fe	110311-110321
3	San Cristóbal	110411-110441
4	Usme	110511-110571
5	Tunjuelito	110611-110621
6	Bosa	110711-110741
7	Kennedy	110811-110881
8	Fontibón	110911-110931
9	Engativá	111011-111071
10	Suba	111111-111176
11	Barrios Unidos	111211-111221
12	Teusaquillo	111311-111321
13	Los Mártires	111411
14	Antonio Nariño	111511
15	Puente Aranda	111611-111631
16	La Candelaria	111711
17	Rafael Uribe Uribe	111811-111841
18	Ciudad Bolívar	111911-111981
19	Sumapaz	112011-112041

We need to limit the postal codes of the localities to only one.

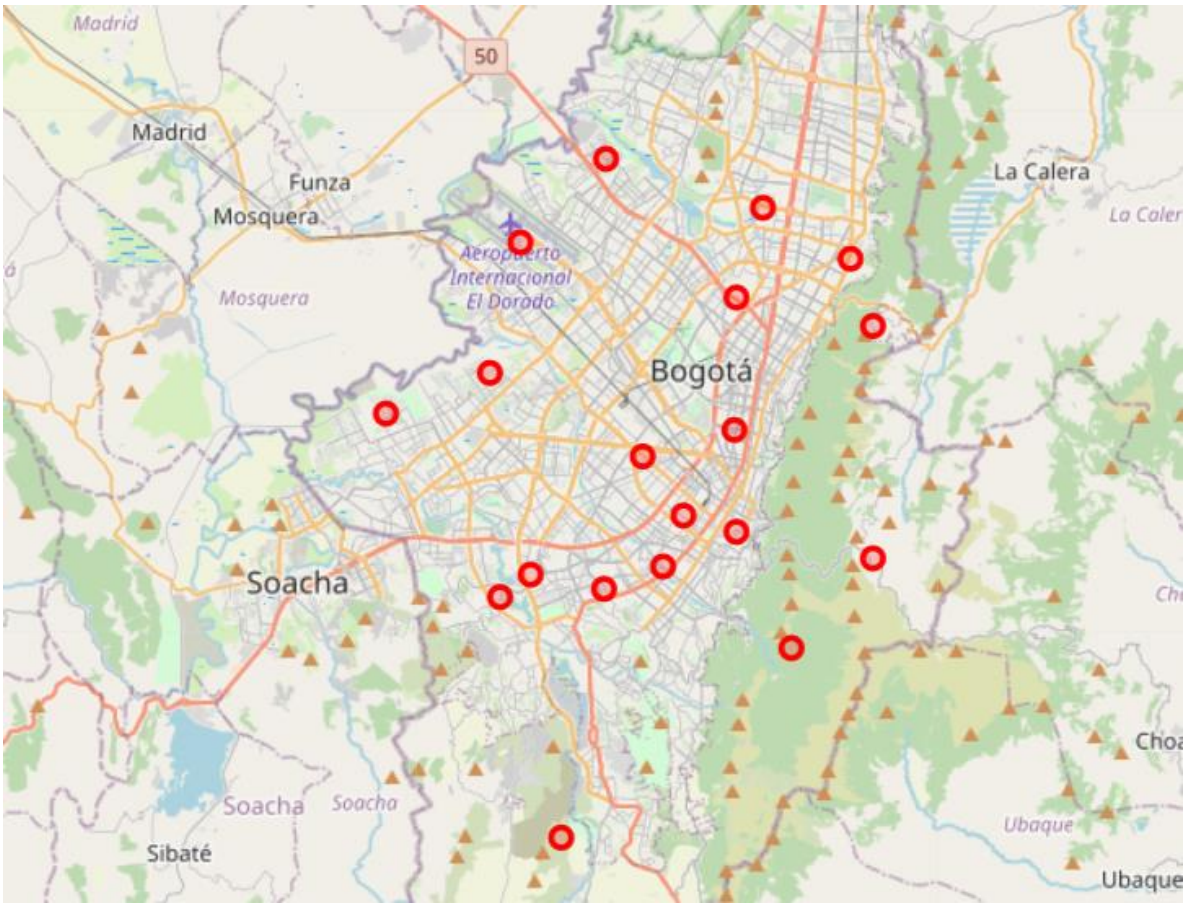
	town	post_code
0	Usaquén	110111
1	Chapinero	110211
2	Santa Fe	110311
3	San Cristóbal	110411
4	Usme	110511
5	Tunjuelito	110611
6	Bosa	110711
7	Kennedy	110811
8	Fontibón	110911
9	Engativá	111011
10	Suba	111111
11	Barrios Unidos	111211
12	Teusaquillo	111311
13	Los Mártires	111411
14	Antonio Nariño	111511
15	Puente Aranda	111611
16	La Candelaria	111711
17	Rafael Uribe Uribe	111811
18	Ciudad Bolívar	111911
19	Sumapaz	112011

Geolocations of the localities of Bogota

ArcGis API

We need to get the geographical co-ordinates for the neighbourhoods to plot out map. We will use the arcgis package to do so.

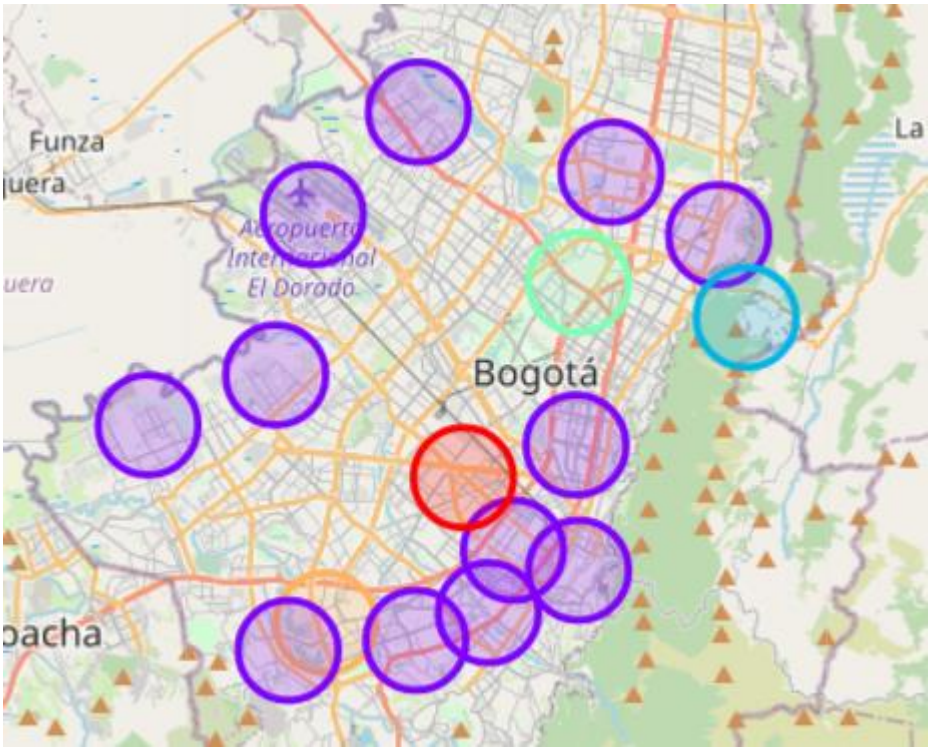
	town	post_code	latitude	longitude
0	Usaquén	110111	4.691000	-74.033448
1	Chapinero	110211	4.667800	-74.025910
2	Santa Fe	110311	4.590255	-74.025630
3	San Cristóbal	110411	4.560310	-74.053523
4	Usme	110511	4.496341	-74.130810
5	Tunjuelito	110611	4.584725	-74.141170
6	Bosa	110711	4.638727	-74.189920
7	Kennedy	110811	4.652555	-74.155299
8	Fontibón	110911	4.695895	-74.144576
9	Engativá	111011	4.724539	-74.115505
10	Suba	111111	4.708096	-74.063170
11	Barrios Unidos	111211	4.677830	-74.071605
12	Teusaquillo	111311	4.633320	-74.072908
13	Los Mártires	111411	4.604256	-74.089915
14	Antonio Nariño	111511	4.587130	-74.096930
15	Puente Aranda	111611	4.624250	-74.103724
16	La Candelaria	111711	4.599095	-74.071979
17	Rafael Uribe Uribe	111811	4.579730	-74.116825
18	Ciudad Bolívar	111911	4.577028	-74.151465
19	Sumapaz	112011	4.223150	-74.185031



Grouping by Venue Categories

Venue Category	Localities	Latitude	Longitude	Venue
Airport	Fontibón	4.695895	-74.144576	Aeropuerto Internacional El Dorado (BOG) (Aero...
Airport Lounge	Fontibón	4.695895	-74.144576	Sala VIP LATAM
Airport Service	Fontibón	4.695895	-74.144576	International Connections Security
Airport Terminal	Fontibón	4.695895	-74.144576	Gate 92
Arcade	La Candelaria	4.599095	-74.071979	Teatro Odeon
...	...	...	...	...
Sushi Restaurant	Usaquén	4.691000	-74.033448	Sushisan Usaquén
Theater	La Candelaria	4.599095	-74.071979	Teatro Colón
Theme Restaurant	Fontibón	4.695895	-74.144576	Andrés Paradero
Toy / Game Store	Engativá	4.724539	-74.115505	Pepe Ganga Outlet
Vegetarian / Vegan Restaurant	La Candelaria	4.599095	-74.071979	Quinua y Amaranto

After One Hot Encoding / Visualization clusters



Results and Discussion¶

The towns of Bogota have a diverse number of places to which you could go, the number of towns is few, despite the grouping that it can be seen that Bogota is a multicultural area, with few towns but with a great variety of activities that They can be done depending on what the person is looking for, its restaurants are divided mainly into typical Colombian food, French, Chinese, Italian and sometimes thematic restaurants, it has many museums and green areas that people seem to like very much, many bars and for people looking to exercise all towns have at least one gym.

Conclusion

The purpose of this project was to explore the localities of the city of Bogota and see how attractive it is for tourists, people who live and also for people who would like to live in Bogota. We explore the city based on the zip codes of the localities and then we extrapolate the common places present in each of the neighborhoods and finally we conclude with the grouping of similar neighborhoods.

We were able to see that each of the towns in the city of Bogota has a wide variety of experiences to offer that are unique in their own way. The cultural diversity is quite evident, which also gives the feeling of inclusion thanks to its section on multinational culture.

Not all towns seem to offer a vacation getaway or romantic getaway with many places to explore, beautiful landscapes, and a wide variety of cultures. But if there is a large amount that could end up in a great experience to spend a short vacation with a pleasant memory not only for its museums, restaurants and parks, but also for its culture, tourist sites and the wonderful Colombian coffee.

References

1. The Battle of Neighbourhood — A Tale of Two Cities by Thomas George
2. Foursquare API
3. ArcGIS API
4. https://es.wikipedia.org/wiki/Anexo:Localidades_de_Bogot%C3%A1