Capstone project: Invest in Bogotá

J. S. Viracachá*.

Decision model of pubs, restaurants and cars park near to Corferias, Bogotá, Colombia.

March, 2020

1. Introduction

Bogotá is the capital and largest city of Colombia, as well as the capital of the department of Cundinamarca. Bogotá had 7,412,566 inhabitants within the city's limits (2018 census) with a population density of approximately 4,310 inhabitants per square kilometer. Only 25,166 people are located in rural areas of Capital District.

Corferias, located in the heart of Bogotá, is known due to is the most important event center in Colombia so, you can find out near to this a large number of hotels and commerce. If somebody is interested in invest in this place, it would be a very good idea if we can advice which would be the better way to put his money: A restaurant, a car parking or a bar and so on. The code would advice also which adress would be better taking into account commerce density.

2. Data

An open-source known as API, which is publicly available in GitHub, is used to retrieve longitude and latitude coordinates of the center of boroughs. The API includes coordinates of all places and boroughs in Bogotá. The information of Corferias is parsed through URL query.

To solve this problem I will use information given by Foursquare relating with restaurants, car parking and bars. In each case I want to use K-means to deal with density problems and to provide stakeholder the best region he can invest his money. For example If you run an initial code of restaurants near to Corferias, you will see that most of this type of commerce is located at western of the event center, so that I forecast that it would be better to setup a restaurant in other location. I will do that with three distinct types of commerce and will compare between these.

3. Methodology

This project has been developed in three steps. First, it was downloaded the data set from foursquare through a python notebook. In this stage it had to be registered the foursquare credentials and to specify the center point where the information had to be pulled (See figures 1, 2 and 3). This center point is Corferias, a convention centre. This place is known due to the commerce surrounding it. The trading models that have been

* jsviracachai@unal.edu.co

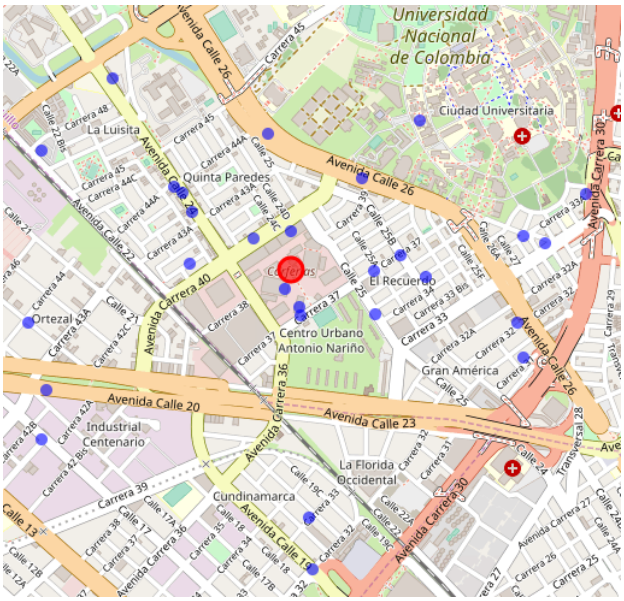

Figura 1: Result of folium function with pubs in Teusaquillo neighbourhood, Bogotá, Colombia.

analyzed were pubs, restaurants and car park.

Second, it was used the k-means method to make clusters to all three business models in order to find the most density regions (See 4, 5 and 6). To do this, was used a radius of 1000 m in the foursquare search. Each point was coloured depending on the cluster it belongs to.

Third, it was computed the density of all centroids to have more tools to analyse and taking a final decision. These densities were estimated with a the equation (1), where N is the number of points in the specific cluster, $\bar{x}$ is the average of the distances to each cluster and an arbitrary adjustment constant α.

$$\rho = \alpha \frac{N}{\bar{x}} \tag{1}$$

4. Results

Taking into account equation 1, have been plotted the clusters density by business model in figures 7, 8 and 9. Every cluster have been charted with and particular color. The cluster centre has been plotted with size its own density with a "X". So that, the bigger the "X", the denser the cluster is. These charts are not in the real scale as the way these were plotted.

Figure 10 shows us all clusters in one chart. Here is possible to see how is the clusters distribution by business model. Finally, in figure 11 we found the contrast among the principal streets, the National University of COlombia and de clusters density. The region colored in violet is the one it would be much preferable to set uo a commerce.

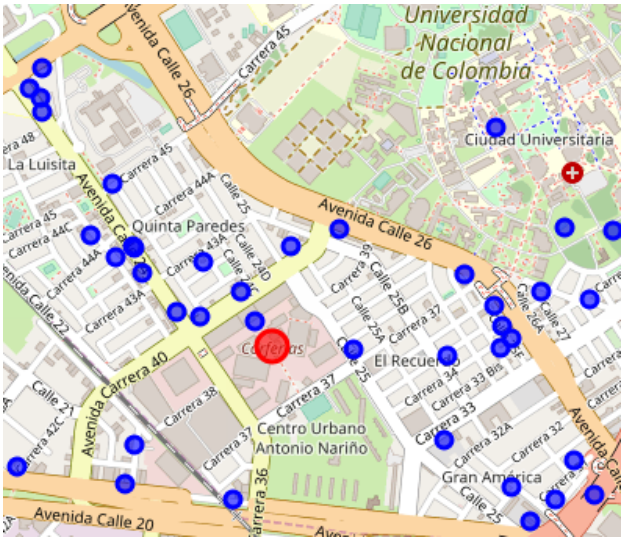

Figura 2: Result of folium function with restaurants in Teusaquillo neighbourhood, Bogotá, Colombia.

Figura 3: Result of folium function with car park in Teusaquillo neighbourhood, Bogotá, Colombia.

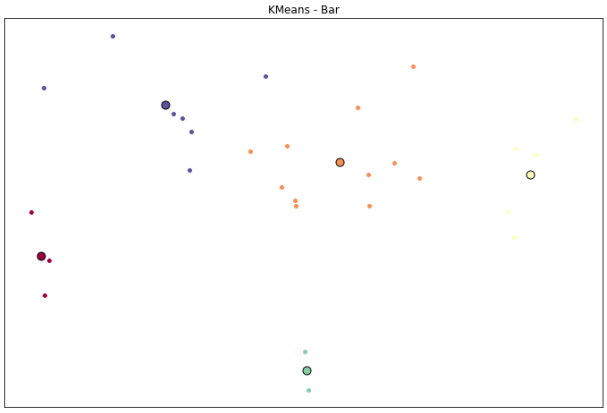

Figura 4: K-means method with pubs in Teusaquillo neighbourhood, Bogotá, Colombia and five centers.

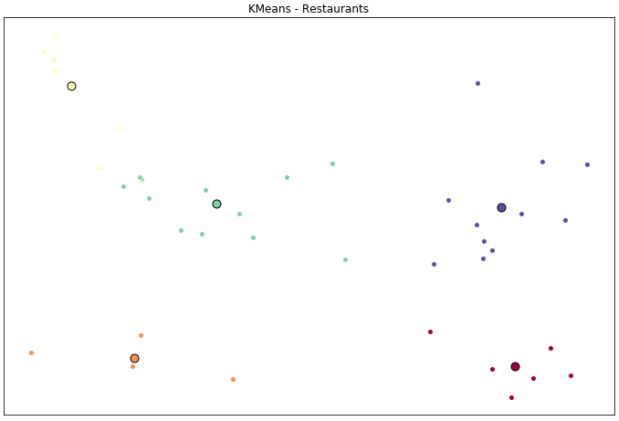

Figura 5: K-means method with restaurants in Teusaquillo neighbourhood, Bogotá, Colombia and five centers.

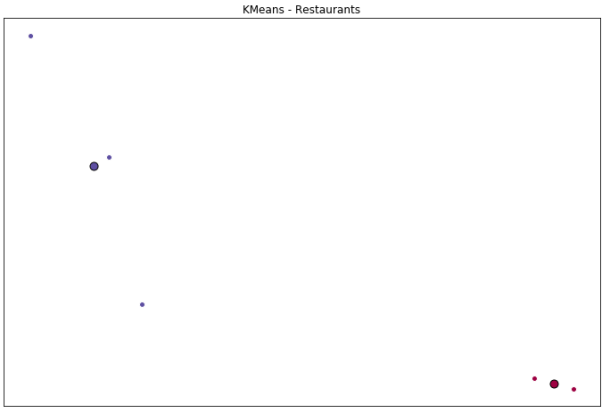

Figura 6: K-means method with park cars in Teusaquillo neighbourhood, Bogotá, Colombia and three centers.

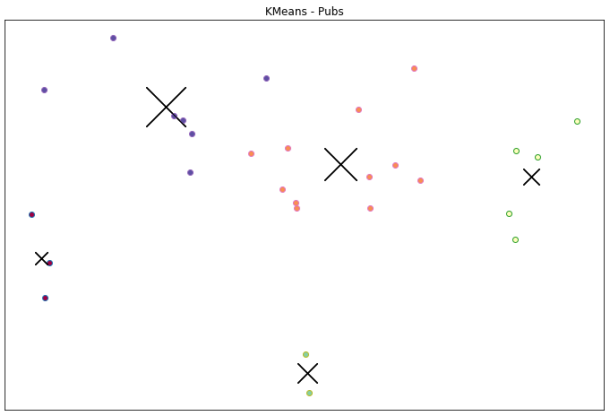

Figura 7: K-means method with pubs in Teusaquillo neighbourhood, Bogotá, Colombia. Centers have been plotted with size depending on the density of each cluster.

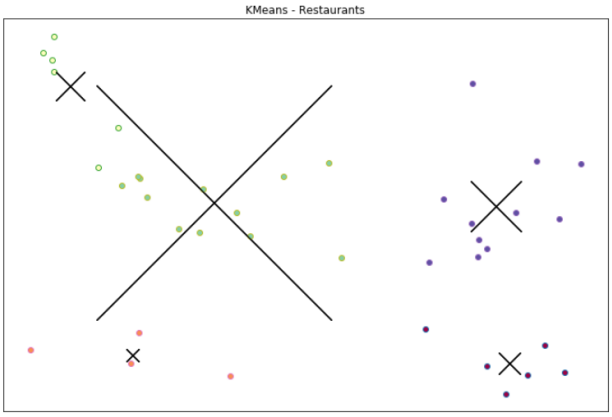

Figura 8: K-means method with restaurants in Teusaquillo neighbourhood, Bogotá, Colombia. Centers have been plotted with size depending on the density of each cluster.

Figura 9: K-means method with cars park in Teusaquillo neighbourhood, Bogotá, Colombia. Centers have been plotted with size depending on the density of each cluster.

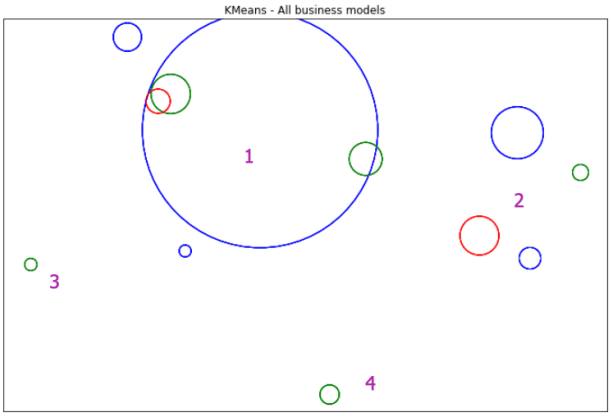

Figura 10: All of cluster centers with their specific density. Blue circles are restaurants clusters. Green circles are pubs clusters. Red circles are car parks. There are 4 super clusters labeled as 1, 2, 3 and 4.

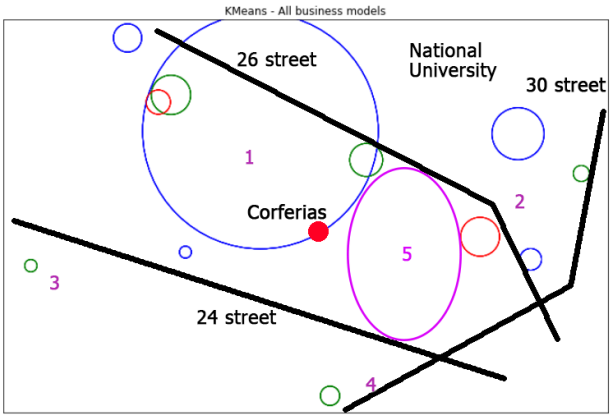

Figura 11: All of cluster centers with their specific density. Blue circles are restaurants clusters. Green circles are pubs clusters. Red circles are car parks. Map of the principal streets. The best zone to invest is the one in violet.

5. Discussion

As you can figure out in plot 10 there are four main regions, two (regions 1 and 2) that are at top and two at bottom (regions 3 and 4). The first two are the bigger ones a the two down the bottom are small density cluster pubs. In figure 11 might be seen the cross of principal streets through the circles. Super cluster 1 is dominated by a high restaurant density (Blue circle) but there are small density points of pubs and car parks. In the second super cluster there is a similar situation but with a density pubs dominating, there are two restaurant clusters and a car park one. In figure 11 shows Corferias as the main point this project turn around (Red dot) and shows at the same time the conclusion of the analysis in region 5 (violet zone). This zone is where business is likely to success due to the density regions. It is not true that there won't already be commerce, but there is in very low density.
It could be done demographic analysis but is not in the scope of this project. This demographic study would be interesting with heat maps where lead to a better precision on boundaries of region 5.
As you can see in figure 11, there are other some regions where density is low moreover. These ones were not chosen because two distinct reasons: first, the empty region among 24 and 26 street (At left of figure 11) was not selected due to is away of Corferias and second, the zone at bottom of the chart between clusters 3 and 4, because inhabitants density and business dynamic there is mean, this is why 3 and 4 clusters are low density ones.
It is visible that high commerce is on 26 street and 30 street between National University and Corferias.

6. Conclusion

As a conclusion was found the zone where is likely your business success (zone 5 in figures 10 and 11). A deeply analysis was done with three different business model: restaurants, pubs and car parks, it means that a deeper study could be done with more types of commerce, with demographic information and with the change of the seasons over the year. It might be done moreover with the possibility to rent commerce places in this zone.
Inside of the 5 region it is advisable set up a business more to the left than right and more to the top than bottom, near to 26 street and National University.