Capstone project: Invest in Bogotá

Juan Sebastián Viracachá Izquierdo
jsviracachai@unal.edu.co

Contenido

Introduction

¿Where?

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.

Where

Figura:

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.

Data

Figura:

Methodology

- First: it was downloaded the data set from foursquare through a python notebook.
- 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 figure 4).
- 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.

Density equation

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

Results

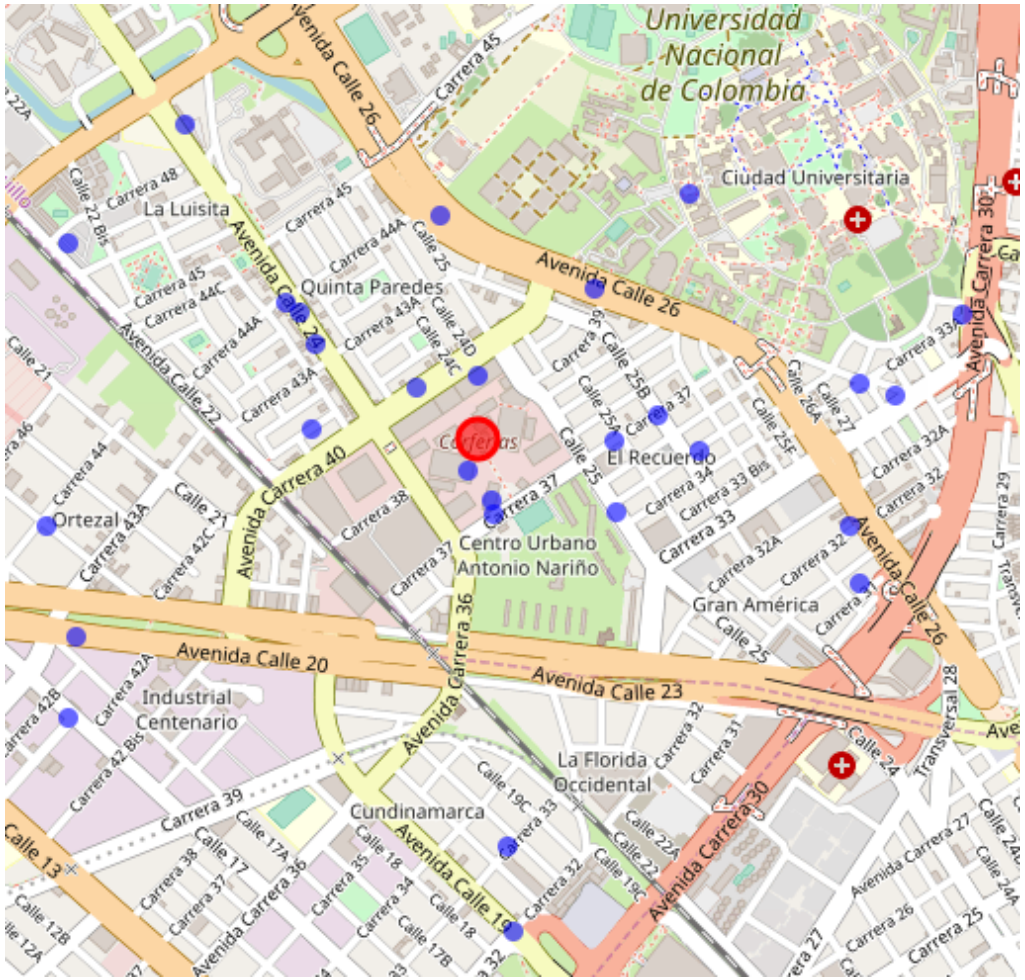

Figura: Result of folium function with pubs in Teusaquillo neighbourhood, Bogotá, Colombia.

Results

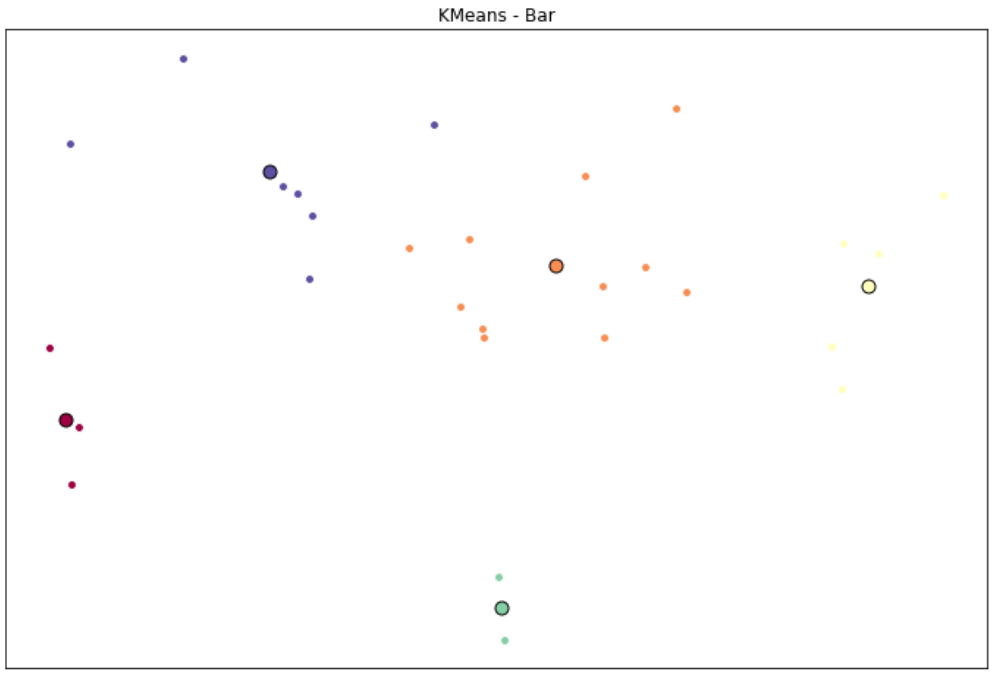

Figura: K-means method with pubs in Teusaquillo neighbourhood, Bogotá, Colombia and five centers.

Results

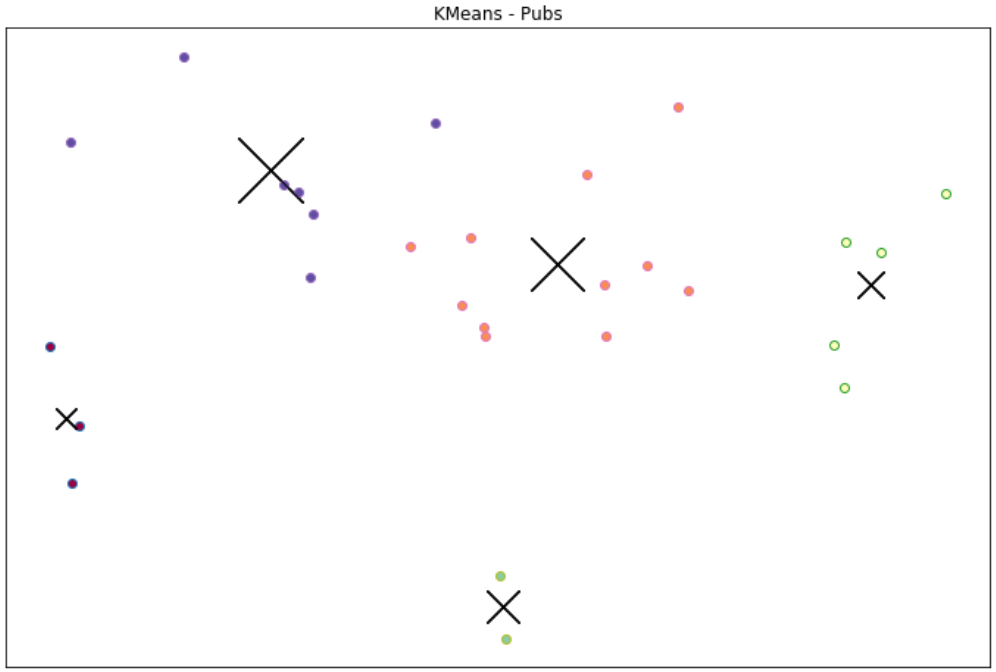

Figura: K-means method with pubs in Teusaquillo neighbourhood, Bogotá, Colombia. Centers have been plotted with size depending on the density of each cluster.

Results

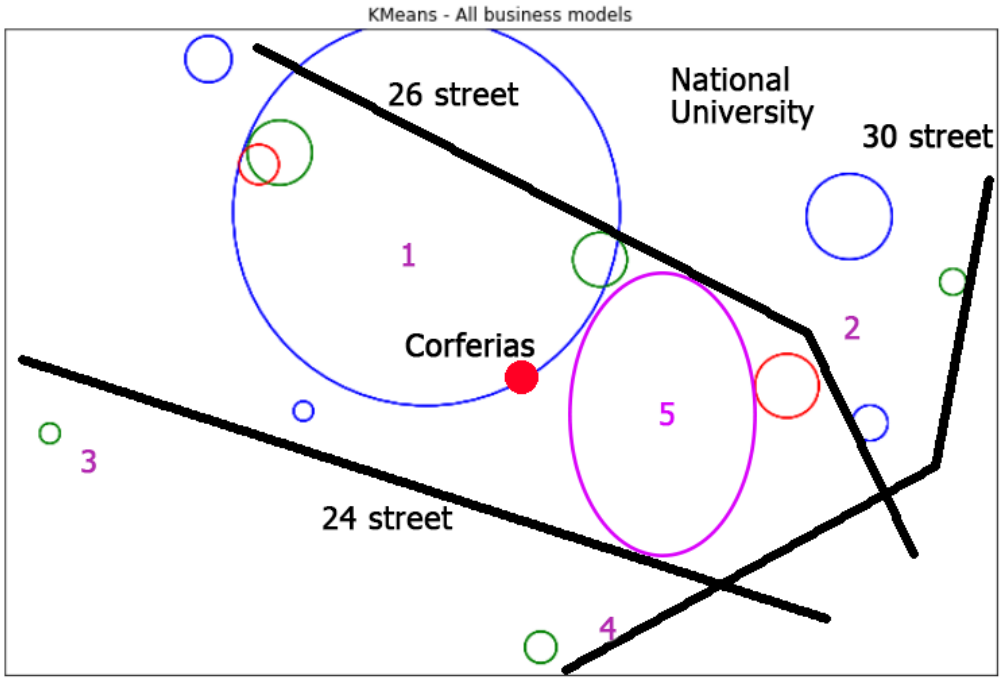

Figura: All of cluster centers with their specific density. Blue circles are restaurants clusters. Green circles are pubs clusters. Red circles are car parks. Map of the principal streets. The best zone to invest is the region 5, the one in violet.

Discussion

As you can see in figure 6, 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 6) 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.

Discussion

As you can see in figure 6, 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 6) 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.

Conclusions

As a conclusion was found the zone where is likely your business success (zone 5 in figure 6). 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.