The Battle of Neighborhoods Capstone Project

JUAN DAVID JARAMILLO

28/04/2020

Table of Contents

Introduction: Business Problem

We will try finding if someone wants to open a new restaurant in Colombia which location is best suited for it keeping in mind the competitors and which income group of people will be attracted most.

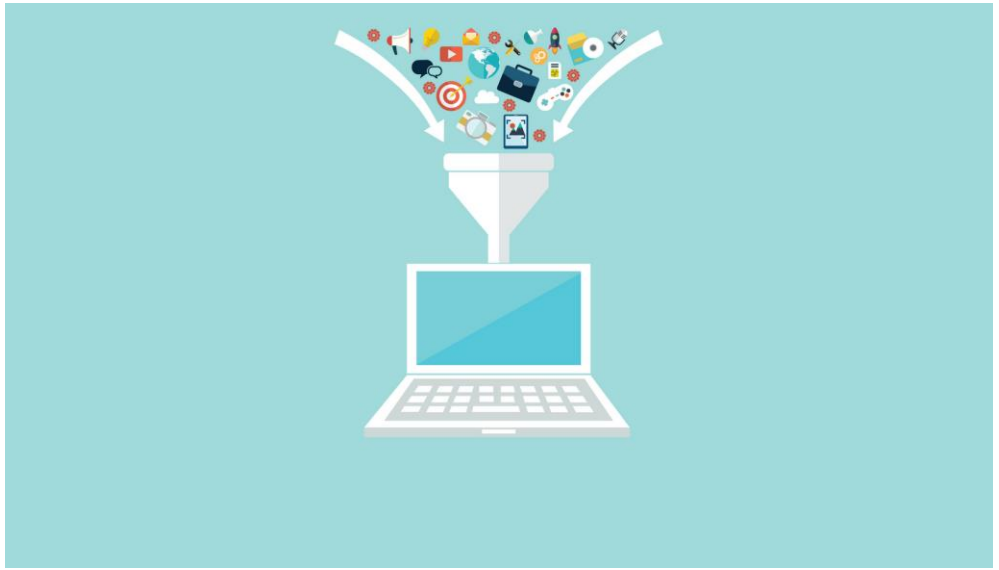
We will consider variables as population of the city, density of American restaurants around and rating given by the people.

We will use our data science powers to generate a few most promising cities based on these criteria.

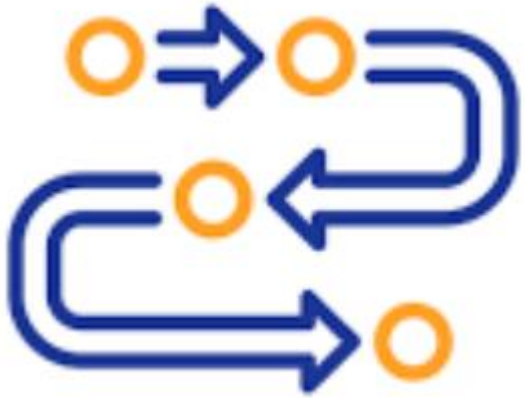
Data

Data on Colombia cities and its population. We will import this database in order to explore it and choose the top 5 cities with the most population.

FourSquare will provide us with information about surrounding restaurants and ratings. By merging data on Colombian restaurants and rating data, we will be able to recommend profitable investments.



Methodology



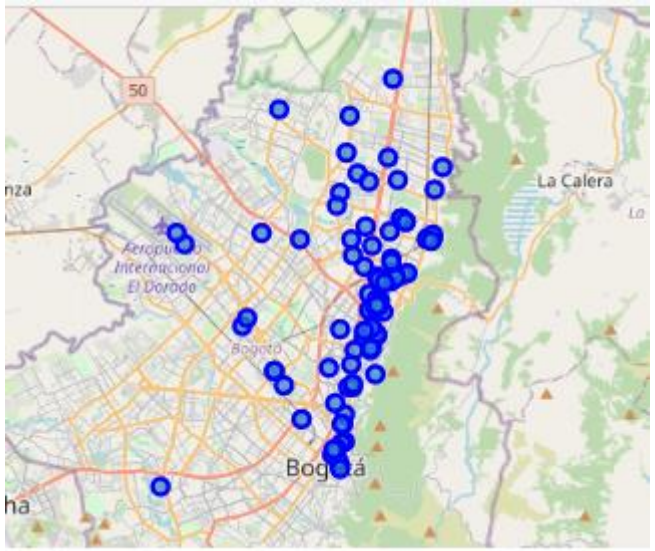
The main target here is to assess which city would have the lowest Aemrican Restaurant density.

I used the FourSquare API through the venues channel. I used the near query to get venues in the cities.

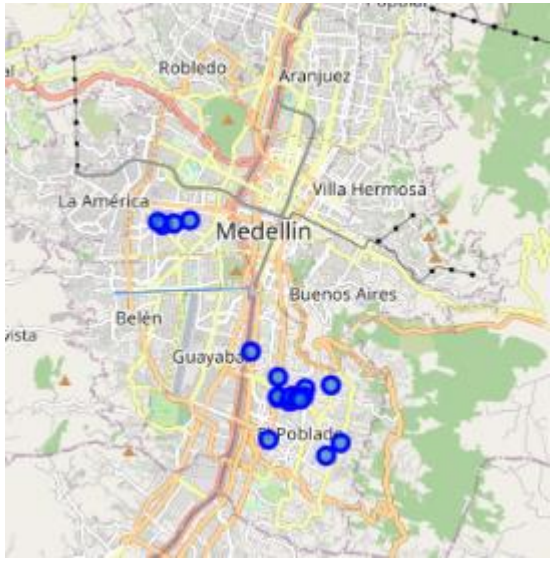
Also, I use the CategoryID to set it to show only American Restuarant type Places.

Each city will be considered as an independent cluster, and the cost function will be to reduce the eucledian distance.

Results and Discussion



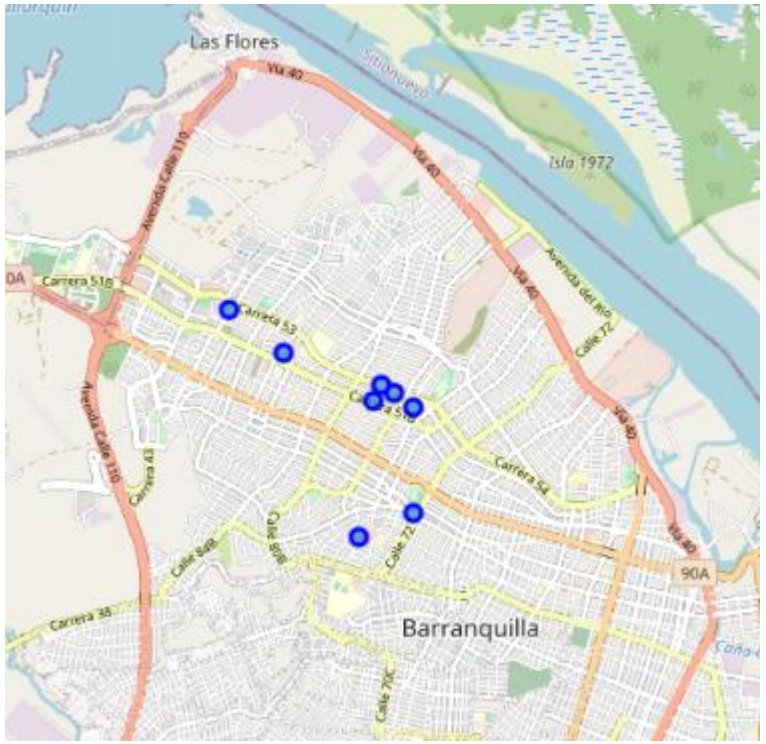
Bogotá



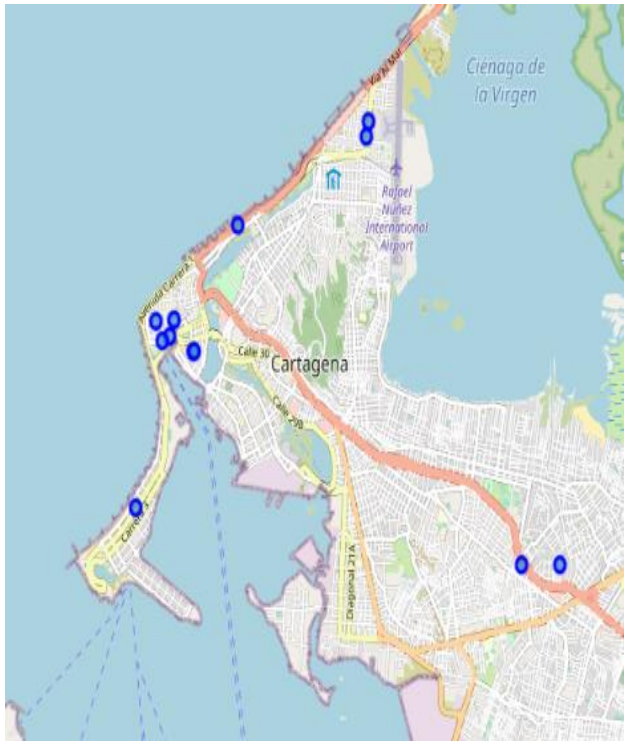
Medellín

As we can see, Bogota es the biggest city in terms of population however by far it has 100 or more restaurants all over de city, thus the market shar will be really small and would have to invest a lot of money to outstand the other good restaurants.
Medellin has the best function in terms of Population/Density. Thus it was considered as one of the possible cities.

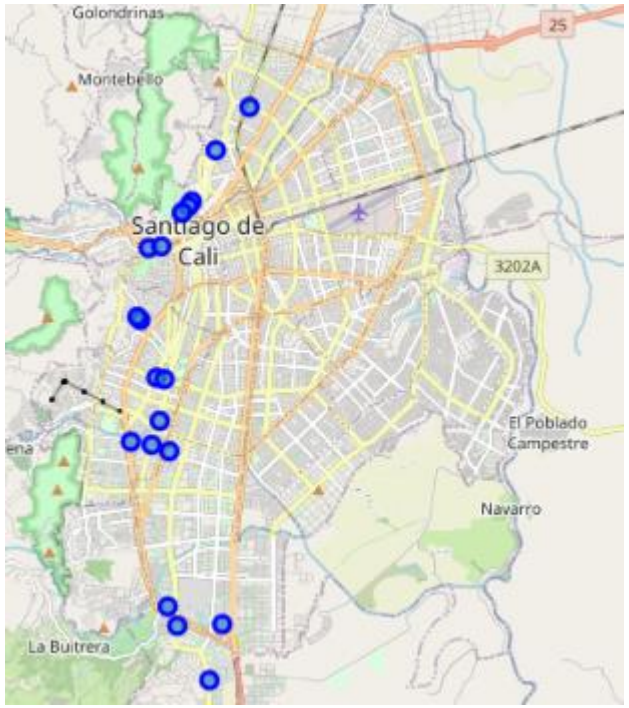
Results and Discussion



Barranquilla

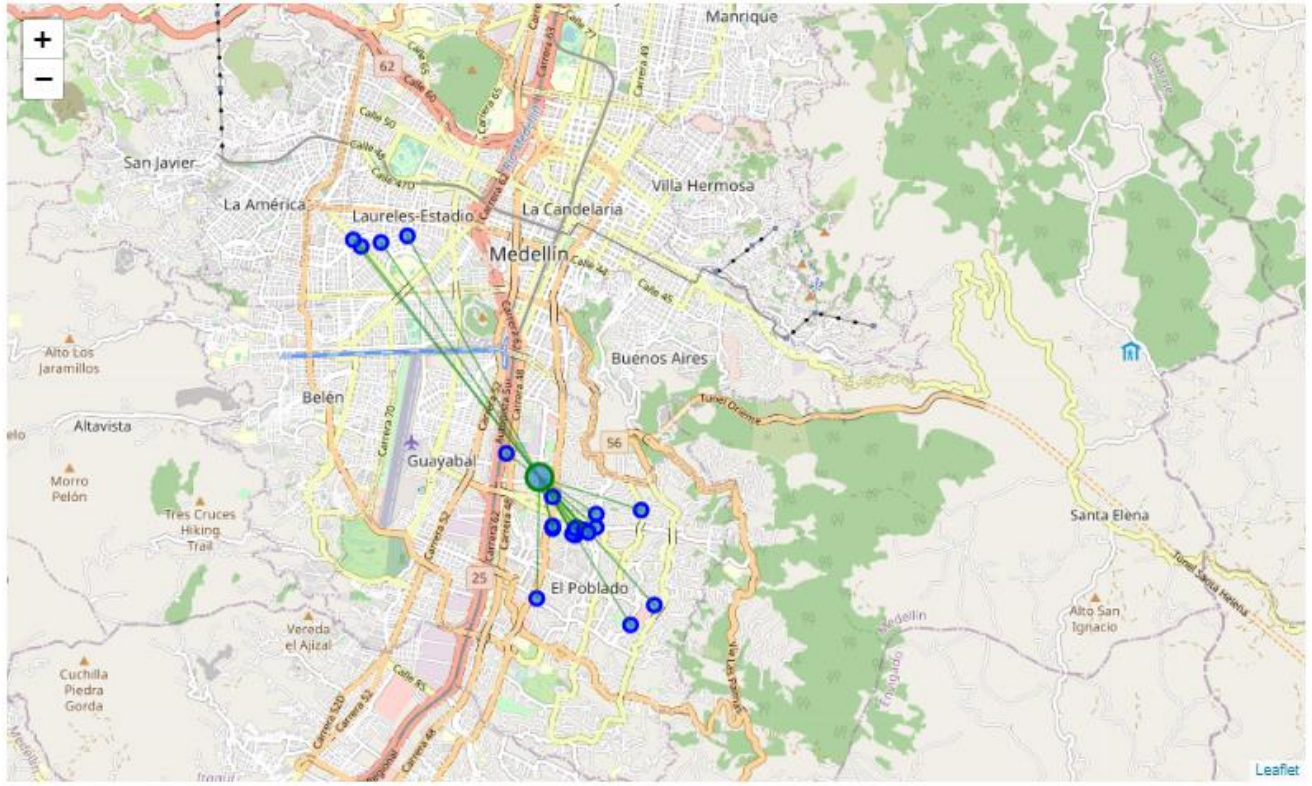


Cartagena



Cali

Results and Discussion



From the results of the clustering algorithm, it was determined that neighborhoods corresponding to cluster 3 (Medellin) were the best choice for opening an American restaurant in Colombia based on the normalized spending power and population.

Conclusion

Opening a restaurant is a complex task that can lead to a large monetary loss if not done properly. Thus, extensive research about the area would greatly increase the likelihood of the restaurant succeeding. From the project above, I demonstrated the workflow necessary for a client to determine what area the restaurant should open. For specifically, I determined that the optimal location to open an American Restaurant in Colombia, would be in Medellin.