Coursera Capstone Project

A new restaurant in Quito: What is the best location of the city?

Jorge Luis Vega

January 14, 2021

1. Introduction

1.1 Background

In this project we will try to find an optimal location for opening a new restaurant in the different locations of Quito (North, South or Centre of the city).

Quito, the capital of Ecuador, is a city that is characterized by its gastronomy. Most of the citizens know which are the best restaurants in the city and it's due to their ratings and reviews than others can go to these restaurants hoping good food and a good service.

Apps like Google Maps allow identity good restaurants through ratings submitted by previous customers.

We will use our data science criteria to find out what is the best part of the city for opening a new restaurant.

1.2 Problem

When someone open a new restaurant in a new place, it’s important to know if the place it’s very visited and has good ratings, because if the place has not good restaurants around, some people will prefer go to other parts of the city, and it would not be a good part or zone for opening a new restaurant.

2. Data acquisition and cleaning

2.1 Data sources

Factors that will influence for opening a new restaurant are:

- The number of restaurants in the three parts of the city
- The amount of ratings
- The number of reviews

We decided to use a data base of the neighborhoods in Quito to obtain the different restaurants around of them. Following data sources will be needed to extract / generate the required information:

- Number of restaurants in the city and their location using Foursquare API.
- Ratings and reviews of restaurant throughout the city using Google Maps.

2.2 Data analysis

We use geopy library to get the latitude and longitude values of Quito City. Then, we upload the data base of Quito's neighborhoods for getting their coordinates and finding restaurants nearby each one.

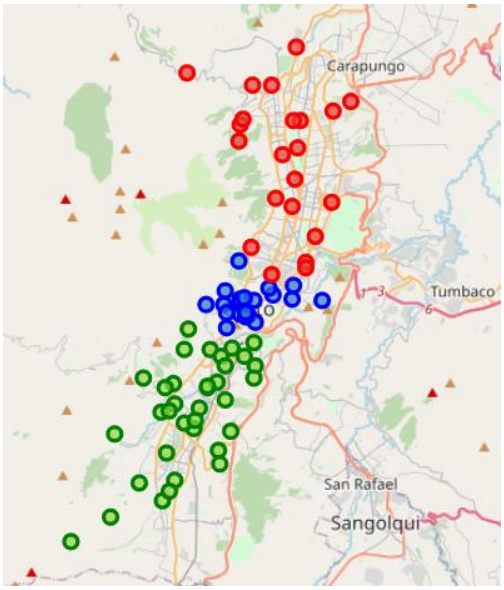

Figure 1. Quito's map with its neighborhoods.

In the Figure 1 we can see Quito's map using "folium" dependence, and ubicating each neighborhood in the map.

- Neighborhoods in the North are represented by color red
- Neighborhoods in the South are represented by color green
- Neighborhoods in the Centre are represented by color blue

We found 191 restaurants or related around Quito. However, using this Foursquare API "free-account" we can't see the reviews and ratings, so we are going to use data got from Google Maps. Data got from Foursquare will help us to know the number of restaurants in each part of the city, and data got from Google Maps will help us to know the ratings and reviews. So, we upload the data base of Quito's restaurants from Google Maps.

In Table 1 we can first five restaurants in Quito. We can see the ratings, reviews, and coordinates.

Now we have the database for analyzing and getting the best part of the city for opening a new restaurant.

Table 1. First five restaurants in Quito (Google Maps data base)

Name	Rating	Review	Latitude	Longitude
Café Plaza Grande	4.5	8	-0.219723	-78.512227
La Colmena	4.2	167	-0.219759	-78.513558
San Ignacio Restaurant	4.4	387	-0.221389	-78.513862
La Casona De La Ronda	4.6	174	-0.224727	-78.513618
Casa Los Geranios	4.5	390	-0.225019	-78.513576

3. Methodology

In this project we collected information of restaurants in different neighborhoods located in the three parts of the city (North, South and Centre). We will limit our analysis to find the best part of the city, and then we will find out what neighborhoods are more suitable for opening a new restaurant.

In the first step we have to relate the data base of restaurants (restaurants location) with each part of the city, using the data base of the neighborhoods. Second step in our analysis will be to find out the relation between the ratings and the reviews. In third and final step we will focus on most promising areas with better ratings and reviews. We will present on the map these locations and the neighborhoods more suitable for opening a new restaurant.

3.1 Analysis

We found what restaurants correspond to each neighborhood.

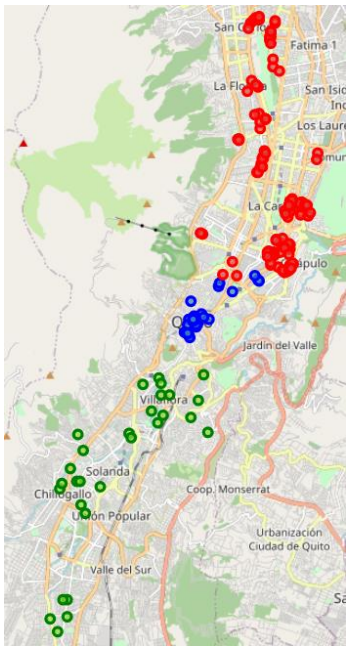

Figure 2. Restaurants located in the three parts of Quito.

In the Figure 2 we can see the restaurants in the neighborhoods in the Quito's map.

- Restaurants in the North are represented by color red
- Restaurants in the South are represented by color green
- Restaurants in the Centre are represented by color blue

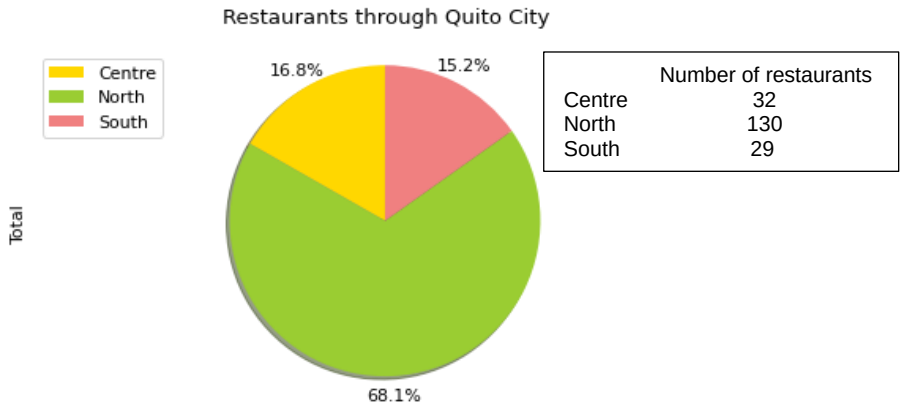

Figure 3. Number of restaurants in the three parts of Quito.

In the Figure 3 we can see the number of restaurants located in each part of the city, and the percentage of each part.

We can realize that in the North of the city exist more restaurants than the other parts.

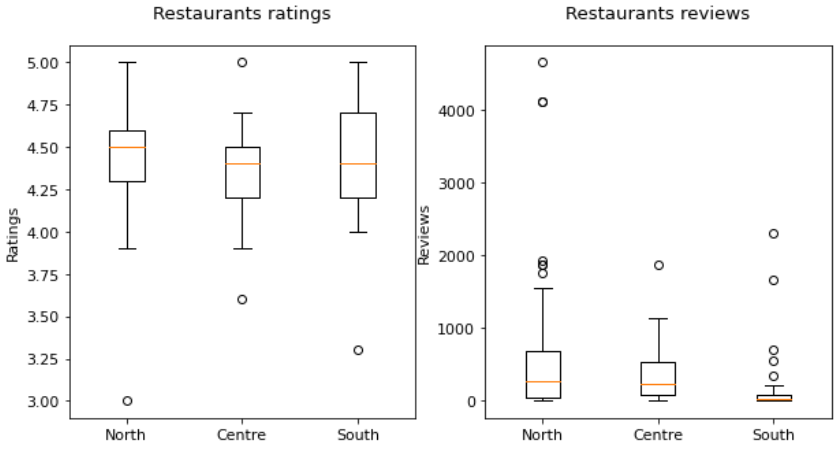

Figure 4. Box plots of ratings and reviews based on the parts of the city.

Table 2. Description of main features of each group.

	Restaurants	Rating (mean)	Review (mean)
Quito North	130	4.45	738
Quito Center	32	4.36	358
Quito South	29	4.42	224

Now, based on Figure 4 and Table 2, we can see that the north of the city has the best mean on ratings and reviews. We can conclude that this is the best part of the city for opening a new restaurant. However, we need to find the best part of the north of the city

We use K-means clustering to find the group of restaurants with the best qualities.

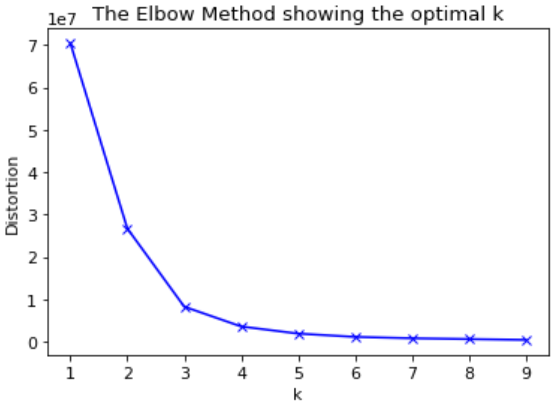

Figure 5. The elbow method to find an optimal k.

In the Figure 5, we can see the elbow method to find the number of clusters. We can observe that the “elbow” is the number 3 which is optimal for this case. Now we can run a K-means using as number of clusters the number 3. We can appreciate in Figure 6 each cluster ubicated in Quito.

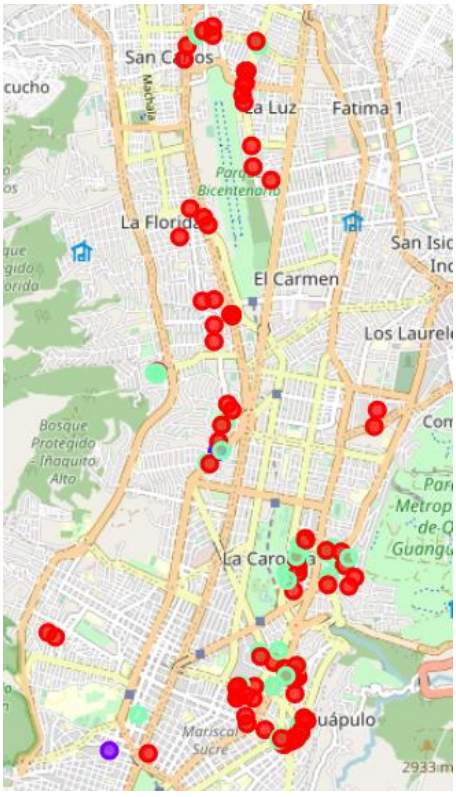

Figure 6. Cluster in Quito’s map.

Table 3. Features of each cluster

	Number of restaurants	Mean Rating	Mean Review
Cluster 1	90	4.43	184
Cluster 2	3	4.30	4289
Cluster 3	37	4.49	1023

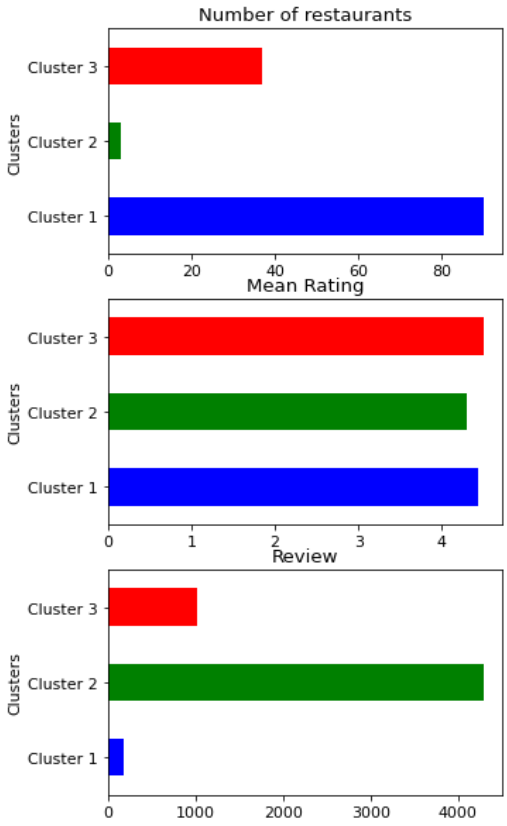

Figure 7. Bar chart of clusters in Quito.

We can notice, based on Figure 7 and Table 3, these results:

- Cluster 1 does not have enough reviews despite of having a good number of ratings and restaurants.
- Cluster 2 does not have enough restaurants despite of having a good number of ratings and reviews.
- Cluster 3 is a middle point between Cluster 1 and Cluster 2 because it has a good number of restaurants and reviews, and the best Mean Rating Thus, Cluster 3 is the best option. Now we have to rank the neighborhoods where we can open a new restaurant based on this cluster

Table 4. the best neighborhoods in the north of Quito for opening a new restaurant.

	Number of restaurants	Mean Rating	Mean Review
Neighborhood			
Quito Norte	8.0	4.48	818
Rumiñahui	5.0	4.62	727
Bellavista	14.0	4.42	726
Guápulo	20.0	4.43	718
Iñaquito	5.0	4.40	581
Quito Tennis	1.0	4.50	568
El Inca	8.0	4.42	545
Kennedy	3.0	4.56	464
El Batán	2.0	4.40	349
La González Suárez	17.0	4.28	266

Finally, in the Table 4, we can see the best neighborhoods in the north of Quito that would be good option for opening a new restaurant.

4. Results and Discussion

Our analysis shows that there are a large number of restaurants throughout Quito, however the classification to find a suitable place in the three parts of the city (north, south and center) were made taking into account the ratings and reviews made by different people.

It was observed that the most desirable area to open a restaurant is the north of the city, since it has a balance between the number of restaurants, ratings and reviews.

On the other hand, a total of 13 neighborhoods were found in the north of the city, which are a good option to open a new restaurant, but the best 10 were chosen.

5. Conclusion

As proposed at the beginning of this project, the different restaurants throughout the city of Quito were examined. Classifying each of these restaurants based on a database of the neighborhoods of Quito. In this way, the restaurants were grouped by sectors (north, south and center) facilitating the analysis. Through the graphs of pie and box plots, it was concluded that the north of the city is the most desirable area to open a new restaurant. However, by using K-means, I can get more in-depth. Resulting in the different neighborhoods where we can open restaurants based on the ratings and reviews.