

DETERMINING THE BEST LOCATION FOR A NEW-AGE STORE IN THE CITY OF BARRANQUILLA, COLOMBIA.

GREGORY OLIVO

08/07/2020

1. Introduction.

1.1 Background.

is the capital district of Atlántico Department in Colombia. It is located near the Caribbean Sea and is the largest city and second port in the northern Caribbean Coast region; as of 2018 it had a population of 1,206,319 making it Colombia's fourth-most populous city after Bogotá, Medellín and Cali. It's therefore one of the most important cities in Colombia due to its strategic location, at the delta of Magdalena river, making it the only river port in Colombia, which makes it a city appealing for foreign investors in search for starting business there.

1.2 Business Problem.

This project is intended to bring information to **Esoterics inc.**, a small group of entrepreneurs from North America interested in opening a New Age Store in the city of Barranquilla, in Colombia.

The city is relatively new to this sort of things and widely attached to traditions, so in order to determine the correct location, the stakeholders consider that an analysis should be conducted to determine the points of the city in which the trends go in a different direction than traditions.

So, it was necessary to spot those activities held in the city which have a relationship with New Age, and which could be considered as a potential source of purchasers of esoteric goods, given the strategic location, which is to be determined through this study.

2. Data

Thanks to the internet, the location of the city in terms of its geographical coordinates is available, also, local contributors who replicate official information to make it readily available for processing of data are a great source of information, which quality is supported by official records.

So, from a blog containing information about the different Departments of Colombia, a csv file was imported to obtain the following Data frame:

	Code	Municipio	Department	Code	Department	Latitude	Longitude
0	5002	Abejorral		5	Antioquia	5.75	-75.416667
1	5004	Abriaquí		5	Antioquia	6.666667	-76.083333
2	50006	Acacías		50	Meta	3.916667	-73.833333
3	27006	Acandí		27	Chocó	8.333333	-77.166667
4	41006	Acevedo		41	Huila	1.75	-75.916667

Figure 1. Data frame containing information of the Departments and cities in Colombia.

Thereafter, the information related to Barranquilla was found:

	Code	Municipio	Department	Code	Department	Latitude	Longitude
89	8001	Barranquilla		8	Atlántico	10.963889	-74.796389

Figure 2. Data frame containing information of Barranquilla.

The Latitude and Longitude of the city would be later used to draw a map of the city containing the neighborhoods which are candidates for the location of the business.

To obtain the information of the barrios (which will be addressed to as neighborhoods from now on), a .csv file had to be built by hand, since the location services that had the information readily available (such as Google maps API) were charged for, once the file was ready, it was imported to the program:

	Comunidad	Barrio	latitud	longitud
0	Riomar	Adela de Char	11.034698	-74.868855
1	Riomar	Adelita de Char Etp. 2	11.032580	-74.874455
2	Riomar	Altamira	11.006025	-74.825973
3	Riomar	Altos de Riomar	11.015493	-74.824646
4	Riomar	Altos del Limón	11.014751	-74.828707

Figure 3. Data frame containing information of the neighborhoods of Barranquilla.

This information would be later used to draw a map and as reference for the next step, which was the retrieval of venues located within a radius of 500 m for each neighborhood.

Foursquare API was a useful tool to retrieve the location of the venues, this information would be used to build another dataframe:

	Barrio	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adela de Char	11.034698	-74.868855	Aires Y Electricidad Nuñez	11.037871	-74.867362	Construction & Landscaping
1	Adela de Char	11.034698	-74.868855	Enterlock Colombia S.A.S.	11.032169	-74.865724	Furniture / Home Store
2	Adelita de Char Etp. 2	11.032580	-74.874455	Donde Jota	11.033551	-74.870806	Latin American Restaurant
3	Altamira	11.006025	-74.825973	Arabe Internacional	11.006065	-74.826594	Falafel Restaurant
4	Altamira	11.006025	-74.825973	Zahle	11.005083	-74.826653	Middle Eastern Restaurant

Figure 4. Data frame containing information of the venues located near or within each neighborhood.

3. Data Analysis

3.1 Methodology

The scope of this project is to retrieve the best location for a New Age store, under certain criteria.

Once all the information was obtained, as well as the location of venues surrounding the different neighborhoods it was necessary to get the neighborhoods which met the desired criteria: surrounded within a radius of 500m by places which activity is somehow aligned with New Age trends, for that purpose, the stakeholders considered that these activities could be:

Reading: bookstores offer a variety of texts for public, even in the esoteric area, therefore, people into esoteric readings may be potential customers for our new store.

Veganism: there are trends within certain kinds of spiritualities that suggest that the sacrifice of animals is a crime against life and therefore, eating their meat is not good. So, according to the stakeholders' criteria, such places as vegan restaurants may be attended by people who are aligned with such philosophies, and therefore be potential customers for the store.

Fitness: fitness centers offer yoga within their service packages, though it is mostly offered within the frame of physical wellbeing, it encloses a spiritual background, so, there is a potential customer in a yoga practitioner according to the stakeholders.

So, the filtering of the information was done bearing in mind the above stated, the result was another data frame:

	Barrio	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
37	Altos de Riomar	11.015493	-74.824646	Librería Nacional S.A.	11.013726	-74.827062	Bookstore
42	Altos de Riomar	11.015493	-74.824646	Panamericana	11.013303	-74.828388	Bookstore
96	Altos del Limón	11.014751	-74.828707	Panamericana	11.013303	-74.828388	Bookstore
118	Altos del Limón	11.014751	-74.828707	Librería Nacional S.A.	11.013726	-74.827062	Bookstore
182	Andalucía	11.015520	-74.816503	Centro Médico Deportivo Body & Soul	11.013104	-74.813725	Gym / Fitness Center
326	La Floresta	11.023139	-74.812633	Springfield Gym	11.024397	-74.812476	Gym / Fitness Center
338	Las Tres Avemarias	11.020626	-74.810417	Springfield Gym	11.024397	-74.812476	Gym / Fitness Center
350	Paraiso	11.014722	-74.811007	Centro Médico Deportivo Body & Soul	11.013104	-74.813725	Gym / Fitness Center

Figure 5. Data frame containing the filtered information according to the stakeholders' criteria.

Further filtering was conducted to obtain more insight on the information, so the next step was to determine how often a category occurred in each neighborhood, this information would be useful for further wrangling and subsequently running the model:

	Barrio	Bookstore	Gym / Fitness Center	Vegetarian / Vegan Restaurant
0	Alfonso López	0.000000	1.0	0.000000
1	Altos de Riomar	1.000000	0.0	0.000000
2	Altos del Limón	1.000000	0.0	0.000000
3	Andalucía	0.000000	1.0	0.000000
4	Betania	0.000000	1.0	0.000000
5	California	0.000000	0.0	1.000000
6	El Carmen	0.000000	1.0	0.000000
7	El Prado	0.666667	0.0	0.333333
8	El Recreo	0.000000	1.0	0.000000

Figure 6. Data frame containing the mean occurrence of a category for each neighborhood.

Then, it was necessary to rank the categories from the 1st to the 3rd for each neighborhood:

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Alfonso López	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
1	Altos de Riomar	Bookstore	Vegetarian / Vegan Restaurant	Gym / Fitness Center
2	Altos del Limón	Bookstore	Vegetarian / Vegan Restaurant	Gym / Fitness Center
3	Andalucía	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
4	Betania	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore

Figure 7. Data frame containing the categories ranked per occurrence for each neighborhood.

Once the filtering was conducted, the next step was to run a clustering analysis on the neighborhoods obtained to determine how they are like each other and choose a good location among the options.

3.2 Running the Model

3.2.1 Clustering

Clustering is one of the most common exploratory data analysis technique used to get an intuition about the structure of the data. It can be defined as the task of identifying subgroups in the data such that data points in the same subgroup (cluster) are very similar while data points in different clusters are very different. In other words, we try to find homogeneous subgroups within the data such that data points in each cluster are as similar as possible according to a similarity measure such as euclidean-based distance or correlation-based distance.

So, to achieve the clustering of neighborhoods, a K-Means model was run. This is where the information of the location of the city, its neighborhoods and the venues were used, to draw the map.

The results:

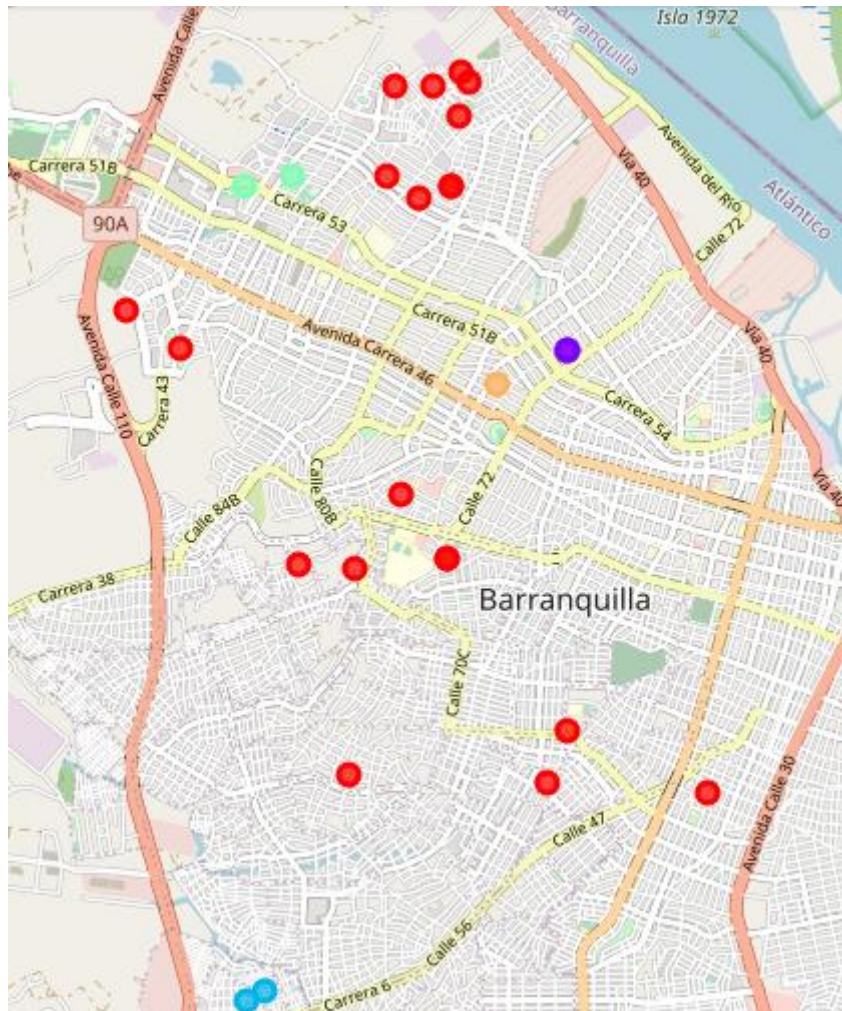


Figure 8. Map of Barranquilla containing the neighborhoods as clustered.

The map of Barranquilla with colored dots, each color corresponds to a cluster, neighborhoods corresponding to each cluster are shown in the following figures:

Cluster 1 (Red):

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
6	Andalucía	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
14	La Floresta	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
16	Las Tres Avemarias	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
17	Paraiso	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
19	San Marino	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
20	San Salvador	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
24	Solaire	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
28	Villa del Este	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
31	Villas del Puerto	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
36	Betania	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
47	El Recreo	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
49	El Tabor	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
64	Miramar	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
101	Alfonso López	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore
115	El Carmen	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Bookstore

Figure 9. Data frame containing the neighborhoods belonging to the first cluster.

Cluster 2 (Dark blue)

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
69	San Francisco	Vegetarian / Vegan Restaurant	Bookstore	Gym / Fitness Center
70	Santa Ana	Vegetarian / Vegan Restaurant	Bookstore	Gym / Fitness Center

Figure 10. Data frame containing the neighborhoods belonging to the second cluster.

Cluster 3 (Sea blue)

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
104	California	Vegetarian / Vegan Restaurant	Gym / Fitness Center	Bookstore
120	El Romance	Vegetarian / Vegan Restaurant	Gym / Fitness Center	Bookstore

Figure 11. Data frame containing the neighborhoods belonging to the third cluster.

Cluster 4 (light green)

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
3	Altos de Riomar	Bookstore	Vegetarian / Vegan Restaurant	Gym / Fitness Center
4	Altos del Limón	Bookstore	Vegetarian / Vegan Restaurant	Gym / Fitness Center

Figure 12. Data frame containing the neighborhoods belonging to the fourth cluster.

Cluster 5 (orange)

	Barrio	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
46	El Prado	Bookstore	Vegetarian / Vegan Restaurant	Gym / Fitness Center

Figure 13. Data frame containing the neighborhoods belonging to the fifth cluster.

4. Discussion

As can be seen, the neighborhoods were segmented in 5 clusters, and each with its particularity. There are different locations for converging places, but further analysis now should be made relying on in-city experience, since there are other factors that may be influential for the location of a new store, such as socio-demographical factors, security, etc. It also can be seen that Barranquilla is quite new to the issue, while in other cities around the world there are places directly categorized as yoga centers or such, our city is just taking its first steps towards that through the inclusion of that discipline in a sportily manner. There are fewer places where the occurrences of bookstores are the highest, which tells a bit about the population appeal for reading on paper books, which can also be an indicator of their intellectual inclination. In order to make a choice, it is necessary to determine which of the three categories influences the most the final decision: while the occurrence of gym/fitness places is greater in opposition to that of bookstores, it must be said that gyms occurrence shouldn't be considered as the primary factor for determining the location of the business, due to the fact that such places not always involve yoga as part of their agenda and not all of them offer it with a slight spiritual bias. Bookstores are not quite frequent around the city, same for vegan/vegetarian restaurants. Therefore, given the variety of customer a bookstore has, it shouldn't be considered in the first place, rather in the second place

after vegan/vegetarian restaurants, which are mostly attended either by people with any health problems or people with certain beliefs which makes them more aligned with the type of store which is to be started.

5. Conclusion

From this analysis, it can be concluded that there are many possibilities for starting the business, but retrieving additional information about the city, asking locals and such, it was advised that the most suitable areas of the city for commercial activity are the center-northern area and the north, towards the city has developed greatly in the past years, also the majority of the wealthy population is clustered. This is aligned with the analysis ran in this project, since if we split the city in two halves (north and south), it can be appreciated that the most of the clusters resulting from the study appear more in the northern half than in the southern half. Also, given the fact that was stated in the discussion in regards to which category should be primary to make the final decision of which neighborhood is the best for starting business, **Altos de Riomar** and **Altos del Limón** appear as the best candidates, but another factor to be considered is a socio-demographical one, which is how far it is from the south of the city, so El Prado appears as another good option in this regard, which like **Altos de Riomar** and **Altos del Limón** has bookstores and vegetarian/vegan restaurants occurring in the 1st and 2nd place respectively, but has a special advantage: it is closer to the center of the city, a traditional commercial area, and frequented by people from both the north and south. Therefore, **El Prado** will be considered as the most suitable location for the store.