Capstone Project - The Battle of Neighborhoods (Week 2)

LEVEL OF ACCIDENTALITY IN BOGOTA COLOMBIA

IN THE 2019

Jose Gabriel Gonzalez

Enero 2021

1. INTRODUCTION

Traffic accidents today represent one of the biggest public health problems worldwide. Leaving not alone, losses of loved ones and damages emotional, but also very high economic costs for medical treatments and surgical, material damage. Therefore, the need arises to know what the causes are that originate traffic accidents and the factors that intervene in the development of themselves. In order to be able to generate preventive and corrective measures that decrease the probabilities of its occurrence.

In Bogota in 2019 there were around 34,998 accidents, some slight but others fatal, for which it is necessary to try to establish the multiple causes that can generate an accident. One of the main causes is the consumption of alcoholic beverages, so this work will be based on establishing whether in the neighborhoods of the city of Bogota where most accidents occur is because there are several places where alcoholic beverages can be consumed.

2. DATA

In the city of Bogota, the information on accidents that occurred from 2011 to 2019 is sought. For this study, only data from 2019 will be used. After this, the latitudes and longitudes of each of the neighborhoods will be obtained for thus to be able to classify them better.

In addition, we will use the Foursquare API to obtain the data of the place of each of the neighborhoods of Bogotá.

Foursquare has one of the largest databases of 105+ million places and more than 125,000 developers use this application. Foursquare API provides many

categories of place data, for this study we are interested in knowing the establishments where alcoholic beverages are sold.

The information that was extracted corresponds to the accident rates in the city of Bogota between the years 2011 and 2019. For the analysis, only the data corresponding to the year 2019 were used. In this file we find the neighborhood where the accident occurred, in addition to the class and category of the accident. Additionally, the longitude and latitude data were consulted on the web in order to complement the file.

3. METHODOLOGY SECTION

In this part we are going to perform an analysis of the data and determine if any type of pattern is found. We first carried out a count to determine which of the Bogota neighborhoods is where the highest number of accidents occurred during 2019. These results can be seen in Figure 1.
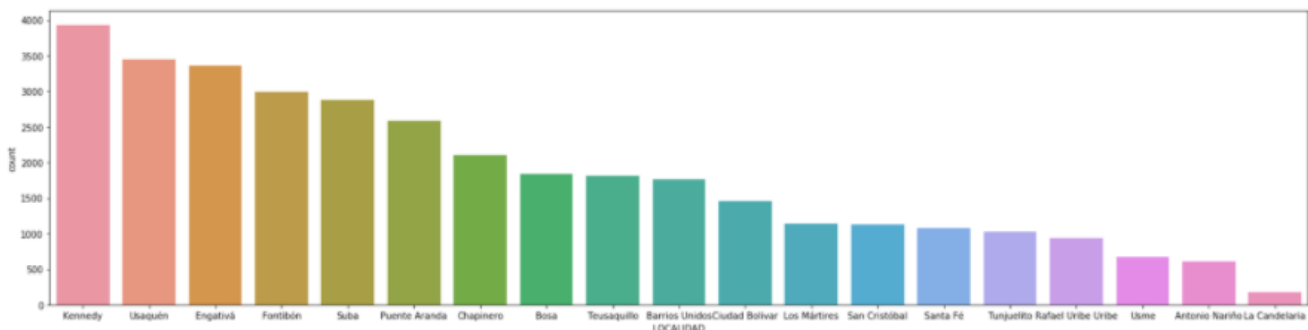

Figure 1: Number of accidents in each of the neighborhoods of Bogota

In the previous figure we can determine that the locality of Kennedy, Usaquen and Engativa are the three neighborhoods where most of the accidents occurred in Bogota, and the peculiarity that exists is that in these neighborhoods there are many places where people can drink drinks

alcoholic. In contrast, the La Candelaria neighborhood is the one with the least number of accidents reported.

In the following figure an analysis was made of which are the most common class of accidents that occurred during the year 2019 and we can conclude that the collision with around 85.4% and run over with 10.3% were the most common classes of the traffic accidents in Bogota.
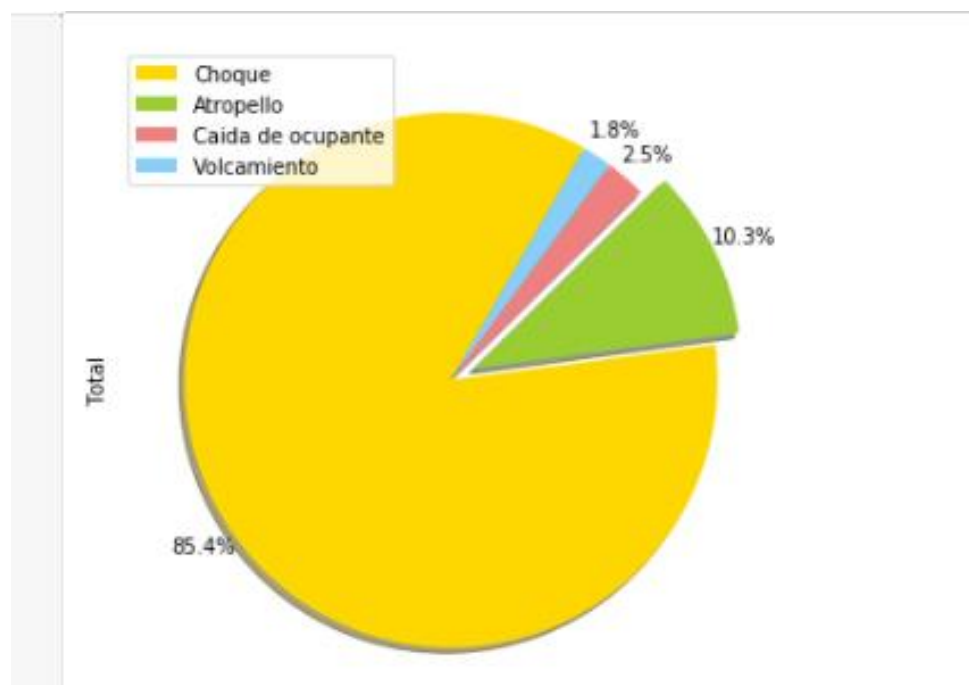

Figure 2: Percentages of accident classes in Bogota in 2019

Finally, an analysis was carried out by the number of accidents that occurred per month in 2019. What can be concluded is that although there are months in which there were many accidents, the values vary between 3205 and 2694 accidents, so These are not figures in which too much can be concluded.
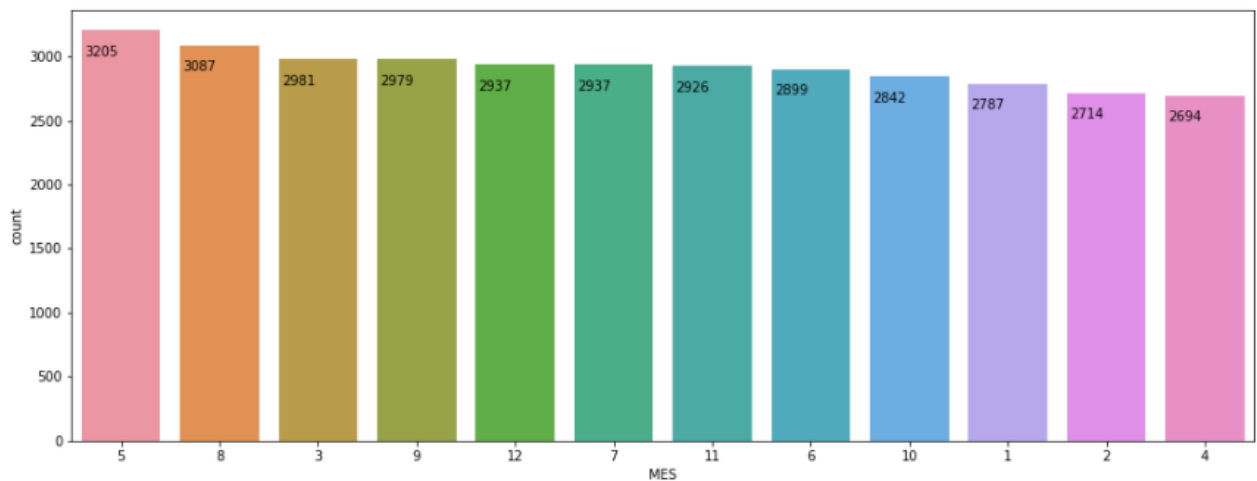

Figure 3: Number of accidents in each of the months of the year 2019

4. RESULTS

In this section we are going to review the results that were generated after having made a classification of each of the neighborhoods of Bogota and thus be able to determine if the neighborhoods where there are large amounts of businesses that sell alcoholic beverages are where the greatest amount occurred. of accidents in the city of Bogota. Now we are going to analyze each of the generated clusters and analyze them.

CLUSTER 1:

	NEIGHBOURHOOD	LATITUD	LONGITUD	TOTAL_ACCIDENTS	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
11	Rafael Uribe Uribe	4.5703	-74.1093	935	0	Arepa Restaurant	Donut Shop	Comfort Food Restaurant	Concert Hall	Convenience Store	Creperie	Cultural Center

Figure 4: First cluster after analyzing the data.

For this first cluster, it can be concluded that in the Antonio Nariño neighborhood, where there were few accidents in 2019, there are not many places where alcoholic beverages are sold, so it can be said that for this reason there are not many accidents.

CLUSTER 2:

	NEIGHBOURHOOD	LATITUD	LONGITUD	TOTAL_ACCIDENTS	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
7	Kennedy	4.6435	-74.1533	3935	1	BBQ Joint	Donut Shop	Breakfast Spot	Movie Theater	Fast Food Restaurant	Coffee Shop	Thai Restaurant
17	Usaquén	4.7352	-74.0183	3450	1	Pub	Burger Joint	Vegetarian / Vegan Restaurant	Dessert Shop	Comfort Food Restaurant	Concert Hall	Convenience Store
5	Engativá	4.6975	-74.1123	3366	1	Shopping Mall	Fried Chicken Joint	Fast Food Restaurant	Bowling Alley	Vegetarian / Vegan Restaurant	Comfort Food Restaurant	Concert Hall
6	Fontibón	4.6785	-74.1410	2998	1	Restaurant	Food & Drink Shop	Fried Chicken Joint	Notary	Mexican Restaurant	Shopping Mall	Park
14	Suba	4.7569	-74.0888	2874	1	Cocktail Bar	Steakhouse	BBQ Joint	Bar	Lake	Donut Shop	Concert Hall
10	Puente Aranda	4.6125	-74.1066	2585	1	Park	Neighborhood	Bakery	Ice Cream Shop	South American Restaurant	Shopping Mall	Department Store
3	Chapinero	4.6459	-74.0634	2102	1	Restaurant	Gay Bar	Coffee Shop	Nightclub	Arepa Restaurant	Bar	Breakfast Spot
15	Teusaquillo	4.6419	-74.0829	1810	1	Gym / Fitness Center	Pool	Event Space	Soccer Stadium	Soccer Field	Restaurant	Bowling Alley
1	Barrios Unidos	4.6787	-74.0787	1769	1	Fried Chicken Joint	Sandwich Place	Fast Food Restaurant	Mexican Restaurant	Creperie	College Bookstore	Coffee Shop
9	Los Mártires	4.6045	-74.0900	1143	1	Shopping Mall	Clothing Store	Skate Park	Boutique	Department Store	Comfort Food Restaurant	Concert Hall
16	Tunjuelito	4.5823	-74.1321	1034	1	Fast Food Restaurant	Bakery	Multiplex	Vegetarian / Vegan Restaurant	Donut Shop	Concert Hall	Convenience Store
0	Antonio Nariño	4.5901	-74.1034	614	1	BBQ Joint	Restaurant	Seafood Restaurant	Clothing Store	Comfort Food Restaurant	Food Court	Convenience Store
8	La Candelaria	4.5944	-74.0741	187	1	Café	Restaurant	History Museum	Latin American Restaurant	Italian Restaurant	Mexican Restaurant	Argentinian Restaurant

Figure 5: Second cluster after analyzing the data.

For this second cluster, it is determined that in most neighborhoods where most traffic accidents occurred during 2019 there are mainly places where alcoholic beverages are sold, although in others where many accidents occur, there are not many places of these.

CLUSTER 3:

	NEIGHBOURHOOD	LATITUD	LONGITUD	TOTAL_ACCIDENTS	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
18	Usme	4.4703	-74.1174	673	2	Southern / Soul Food Restaurant	Vegetarian / Vegan Restaurant	Donut Shop	Comfort Food Restaurant	Concert Hall	Convenience Store	Creperie

Figure 6: Third cluster after analyzing the data.

For this third cluster, it is determined that in the Usme neighborhood there are not many accidents and there are not many places where alcoholic beverages are sold.

CLUSTER 4:

	NEIGHBOURHOOD	LATITUD	LONGITUD	TOTAL_ACCIDENTS	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8t Co
12	San Cristóbal	4.5646	-74.0833	1135	3	Playground	Coffee Shop	College Bookstore	Comfort Food Restaurant	Concert Hall	Convenience Store	Creperie	

Figure 7: Fourth cluster after analyzing the data.

For this fourth cluster, it is determined that in the San Cristobal neighborhood there are many accidents, but the curious thing is that there are not many places where alcoholic beverages are sold.

CLUSTER 5:

	NEIGHBOURHOOD	LATITUD	LONGITUD	TOTAL_ACCIDENTS	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
2	Bosa	4.6171	-74.1899	1841	4	Park	Supermarket	Dessert Shop	College Bookstore	Comfort Food Restaurant	Concert Hall	Convenience Store

Figure 8: Fifth cluster after analyzing the data.

For this fifth cluster, it is determined that in the neighborhood of Bosa there are many accidents but according to Foursquare data there are not many places where alcoholic beverages are sold.

5. CONCLUSION

According to the separation of the clusters, it can be determined that in some places where many accidents occur, it is because there are many places where alcoholic beverages can be consumed, but in other cases where the accident rate is higher, there is not a great variety of places. where you can drink alcoholic beverages so it would not be relevant to think that there is a correlation. Additionally, it may also happen that there is not enough data in the Foursquare app to be able to carry out a more in-depth analysis.