Market Research Express

By Diego Lombardi (diego.lombardi.boscan@gmail.com)

Introduction/Business Problem

Market Research is an area that requires an important effort to gather information for each client. On many occasions these clients want a "first vision" of their market. These requests are usually accompanied by a sense of urgency. In this sense, I have designed this application to be able to know the environment of any market in a few minutes, particularly elements of the geographical distribution of the premises, the average distance of the chosen point (indicating the degree of density), as well as a brief analysis of the predominant names (as a basis for future Branding).

To run this application, you only need to have Foursquare credentials, and the location you want to evaluate.

Data

It works with the data provided in Foursquare, specifically: name of the place, location (latitude and longitude), category, and distance of the selected point. After testing the application's functionality, we take as an example the city of Calgary, specifically we take as reference the following address: '101 9 Ave SW, Calgary, AB T2P 1J9, Canada'. This is result is show below.

	Venue	Venue Latitude	Venue Longitude	Evaluation	Distance
0	Cafe Koi	51.042854	-114.065565	This spot is popular	221
1	Blink	51.045422	-114.063733	This spot is popular	227
2	LOCAL on 8th Ave	51.045784	-114.068404	This spot is popular	166
3	Cactus Club Cafe	51.045966	-114.070166	This spot is popular	276
4	Vintage Chophouse and Tavern	51.042637	-114.068286	This spot is popular	252

We established a radius of 700 meters, initially we work in radius of 2,500 meters and 1,500 meters. However, friends living in Calgary recommended us to reduce the radius as it is a small city, and with a certain commercial density. After connecting the application with Foursquare, and processing the data, this result was obtained, a dataframe of 100 venues and 53 categories.

Methodology

Once the data was obtained, the next step was to make a map with the identified locations. In this way it was expected to have a visual idea of how the premises are distributed in the selected area. This is shown below.
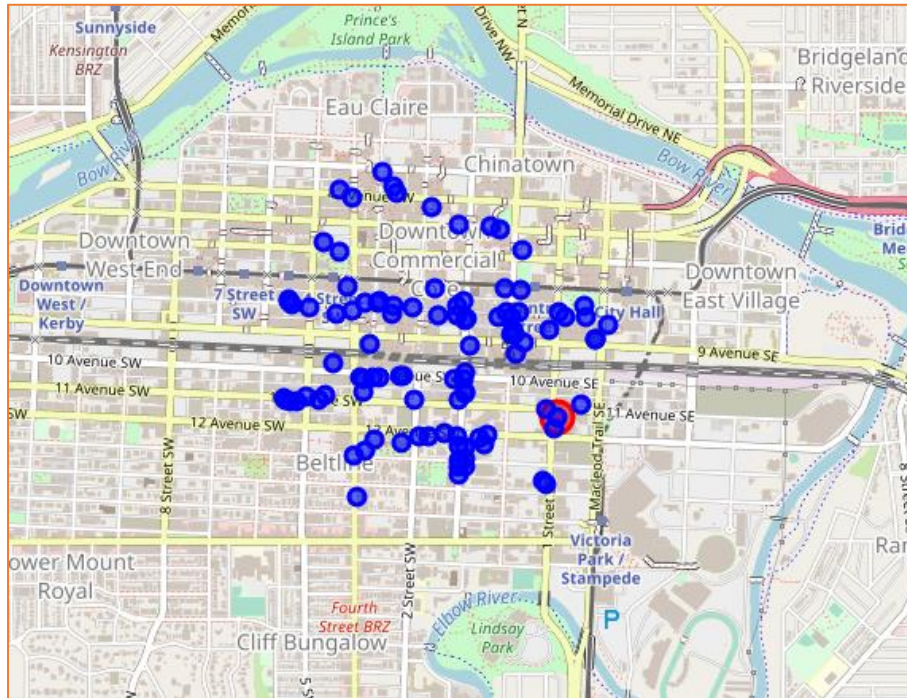
Folium was used for this phase, so you can have an initial map of the premises. Despite being an initial phase, observing the map has a first interesting result, it is the clear trend of the location of the premises to the left of the selected point.

From the methodological point of view, we worked with techniques to make a market analysis. First, we worked with distance, a variable that is used to measure the accessibility of the respective premises analyzed. In addition, the average distance of all the premises considered was taken, which was used as a measure of density; and, on the other hand, the distribution of these distances was analyzed, to know if they tended towards the selected point or if, on the contrary, they tended to move away.

The second regret with which we worked was with the names of the premises, when analyzing the names certain tendencies in the use of the words can be observed, which can give indications to
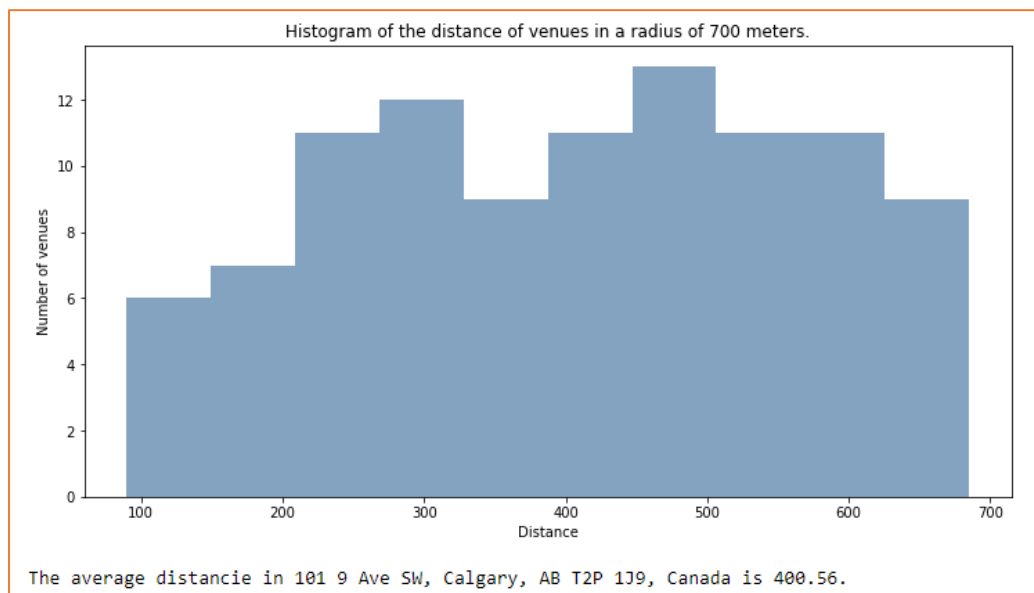
evaluate if a specific name that is wanted to develop is in agreement with the environment or not. To do this analysis he worked with WordCloud.

Finally, the premises were segmented according to categories. Although Foursquare already generates certain categories, they were very broad for the analysis to be performed. In this sense they were reduced from 53 to 7, for this purpose a function was created to generate the corresponding classes.
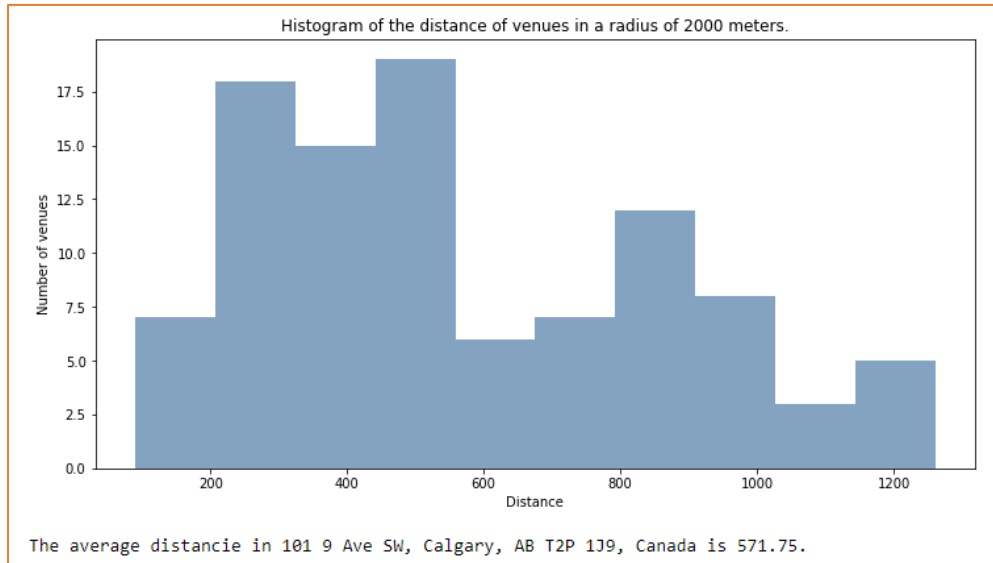
Results

As noted in the previous section, the first aspect evaluated was distance. In this sense, the average distance of all the premises considered with respect to the selected point was estimated, being the same as 400 mts.

This result was complemented with a histogram showing the distribution of distances. If the histogram tends to the left it indicates that the distances must be close, if on the contrary it tends to the right it would be in the presence of a tendency to move away. The results are shown below.

Histogram of the distance of venues in a radius of 700 meters.

The average distancie in 101 9 Ave SW, Calgary, AB T2P 1J9, Canada is 400.56.
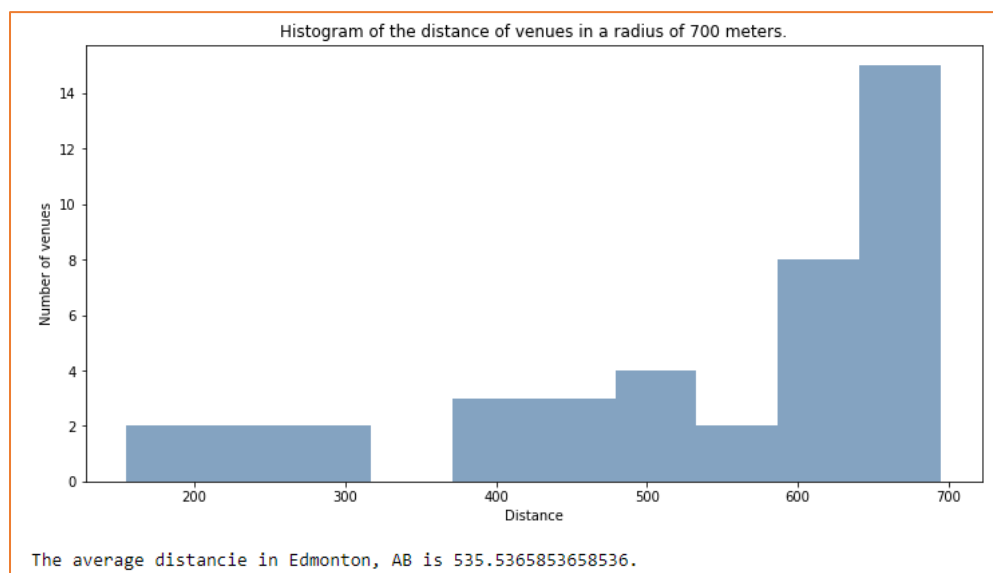
As can be seen in the graph, the tendency is towards the center, which suggests a homogeneous distribution and towards a central value. To deepen more on this aspect two additional exercises were done, the first one was to extend the radius. In this sense, the radius was extended to 2,000 meters and the following result was obtained.

Histogram of the distance of venues in a radius of 2000 meters.

The average distancie in 101 9 Ave SW, Calgary, AB T2P 1J9, Canada is 571.75.

As can be seen, the average distance did not increase significantly, reaching 571 meters. This aspect is corroborated in the following histogram, which shows a tendency towards the left, suggesting that the premises do tend towards the selected point.

In order to make a comparison, the city center of Edmonton was taken, it should be noted that this point was not due to any specific aspect and that it may not represent the point equivalent to that selected in Calgary. In the same way the analysis was done, and the following result was obtained.

Histogram of the distance of venues in a radius of 700 meters.

The average distancie in Edmonton, AB is 535.5365853658536.

As it is observed, in this case the tendency of the premises is to move away from the selected point. In fact, the average distance is quite close to the imposed limit of a radius of 700 meters.

The second aspect that was analyzed were the names of the premises, for that we worked with WordCloud. The objective of this analysis was to determine if there was any pattern in the names, so that it serves as a tool for the naming phase with respect to the Branding of any company that wants to install in the area, or that wants to improve its brand. The result is shown below.

As you can see, there is a certain predominance of bars, cafes, and taverns. However, as will be seen later in the analysis by categories, restaurants have more diverse names and at least the word restaurant appears less. What seems clear is that the name "Calgary" predominates in the names of the selected premises.

The last aspect that was worked with was the segmentation of the premises into a smaller number of categories, going from 53 to 7. Being these: Restaurants, Café, Bar, Art, Fitness, Store. Below is the result of the number of premises according to the categories created.

In the selected area restaurants predominate, followed by bars. In this sense, this suggests that the selected area is suitable for locals to eat or share drinks. To deepen this aspect, a map of the area was made showing three categories: restaurants (blue), bars (green), and coffee (orange). Then the result.
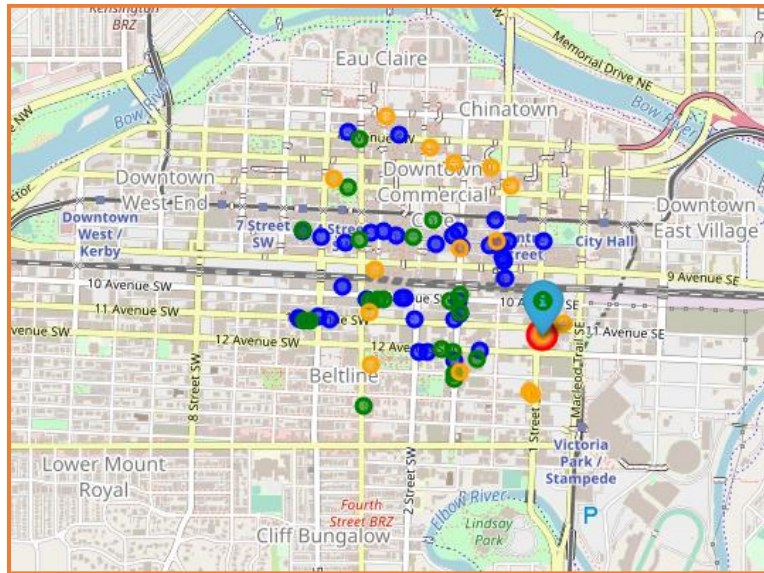
The graph shows that most restaurants and bars are near Avenue 9. Restaurants tend to be on Avenue 8, while bars on Avenue 10. Coffes are more scattered.

Discussion and Conclusion

As it has been shown in this project, by having a specific location it is possible to perform a general analysis of the commercial characteristics of the area. Specifically, you can show aspects such as the location of premises according to category, the distance from the selected point and if they tend to be close or, on the contrary, to move away. An analysis of the names of the premises was also made as a support tool for the branding of the business that you want to develop.

The main value of this tool is that it can be applied in any location, in a simple way, and with valuable results for any market study.