Deciding potential locations to start a food-related business in Rosario, Santa Fe, Argentina

Manavella, Matias Nicolas

April 16, 2020

1. Introduction

Rosario is one of the largest and economically most significant cities in Argentina. Rosario is located strategically at the side of the Parana River, controlling the access and navigation from Buenos Aires to the heart of the country.

Being such an important city, and having two top public universities of great category, Rosario has an ever growing amount of citizens, with a large percentage of young people, mostly students. This gives place to great opportunities in the food and entertainment areas.

1.1 Problem

This project will make an analysis of the city, searching for potential business opportunities in some of the districts of Rosario, located in the central Area of the city, where most people spend their free and work time. The main objective will be to identify the lack of some kind of venues in each district.

1.2 Interest

Stakeholders for this project would be the investors group, ready to start a new business in the central area of Rosario.

2. Data acquisition and cleaning

2.1 Data sources

We will use data from two main sources. First, the official web site of the city hall will provide the necessary information about the district's shapes and location. We will download a .csv file directly from the link of the website, storing it into a pandas dataframe.

Secondly we will use the geojson information to calculate the geographical center of each district, and use this to get the venues in the area, using the Foursquare app.

2.2 Data cleansing

Once we have all the for each district, the first problem we encountered was that the geoJson file was store separately in different rows of the dataframe. We had to make our way into making a string variable that would be readable by the folium library.

We also had to clean the data from Foursqare. For starter, the information comes into a geoJson file all together. We used the tool provided by the pandas library, json_normalize, which organize the data into a data frame. Second problem here was, that the categories row is, in fact, a complete dictionary with a lot of data we don't need, so we used a function to separate the only value of the category dictionary that was useful for us, the name. Once we had the info of the venue, the category and the location, we add the info of the district to which the venue belongs.

Being a city from a country in development, the information is not always complete. In this case, many categories were unknown, and this could cause several problems in the classification algorithms. So the third step was to clean all the rows, where the category data was missing.

With all of this done, we had our data ready to be analyzed.

3. Methodology

In this we will direct our efforts to search, identify and filter the venues in each district. We divided our analysis in the following steps:

1. First, we obtained the information about each district, directly from the city hall web page. We used this data to get information about the shape and distribution of all the districts in Rosario.

2. Having all the information about the districts, we used the geographical information provided by the city hall to choose the geographical center of each district. After doing this, we used this center to get the venues in the area.

3. Third we will filter the venues-related information, getting only relevant data about food related stores. This kind of venues will provide the information we need to start our food related business.

4. Finally we will analyze the frequency of each kind of store in each relevant area, and use this information for our startup.

4. Analysis

At the moment, we count we the following data:

- Distribution of districts in Rosario, and geographical info about the ones in the central area of Rosario.
- Venues in the district and its category.

The analysis will start by filtering the venues to be analyzed. First we selected from the Foursqare documentation, key words to filter every venue, which were not related with food. The set of key words was the following:

```
['bar', 'bakery', 'beer', 'cafe', 'café', 'coffee', 'food', 'ice cream
', 'izza', 'retaurant']
```

We ended up with a list of food related venues in the central districts of Rosario. What we must do now is get the dummies for each category, and group all the venues according to its district. We take the mean value of the total and finally, we get the frequency of each venue in each district.

5. Results

The result of the top five most frequent venues per area was the following:

```
---- District: C1----
                             venue  freq
0                    Ice Cream Shop  0.23
1                            Bakery  0.18
2                Salon / Barbershop  0.14
3                       Coffee Shop  0.09
4   Vegetarian / Vegan Restaurant  0.05
5                               Bar  0.05


---- District: C2----
                         venue  freq
0           Salon / Barbershop  0.12
1                  Beer Garden  0.12
2                       Bakery  0.12
3        Argentinian Restaurant  0.06
4                  Gaming Cafe  0.06
5            Mexican Restaurant  0.06
```

```
---- District: C3----
                     venue  freq
0               Restaurant  0.23
1                     Café  0.23
2    Argentinian Restaurant  0.08
3          Food & Drink Shop  0.08
4         Spanish Restaurant  0.08
5               Pizza Place  0.08


---- District: C4----
                     venue  freq
0                     Café  0.21
1    Argentinian Restaurant  0.14
2      Fast Food Restaurant  0.14
3          Food & Drink Shop  0.14
4         Salon / Barbershop  0.14
5          Italian Restaurant  0.07


---- District: C5----
                     venue  freq
0    Argentinian Restaurant  0.18
1                Food Truck  0.18
2            Ice Cream Shop  0.09
3                    Bakery  0.09
4               Pizza Place  0.09
5               Coffee Shop  0.05


---- District: C6----
                     venue  freq
0                    Bakery  0.31
1                       Bar  0.15
2        Salon / Barbershop  0.15
3          Food & Drink Shop  0.15
4                      Café  0.08
5                Restaurant  0.08


---- District: C7----
                     venue  freq
0               Pizza Place  0.15
1                Restaurant  0.15
2               Coffee Shop  0.15
3                    Bakery  0.10
4          Food & Drink Shop  0.10
5    Argentinian Restaurant  0.05
```

The **less frequent** venues in each area were:

```
---- District:C1----
                          venue  freq
0        Argentinian Restaurant   0.0
1            Spanish Restaurant    0.0
2     South American Restaurant   0.0
3            Mexican Restaurant    0.0
4            Italian Restaurant    0.0
5                 Internet Cafe    0.0
6                   Gaming Cafe    0.0
7                    Food Court    0.0
8             Food & Drink Shop    0.0
9                     Hotel Bar    0.0
10        Fast Food Restaurant    0.0
11            Fondue Restaurant    0.0
```

```
---- District: C2----
                             venue  freq
0    Vegetarian / Vegan Restaurant    0.0
1                Sushi Restaurant     0.0
2              Spanish Restaurant     0.0
3       South American Restaurant    0.0
4                       Cafeteria    0.0
5                            Café    0.0
6                     Coffee Shop    0.0
7                      Restaurant    0.0
8                     Pizza Place    0.0
9                      Food Court    0.0
10                     Food Truck    0.0
11                Tapas Restaurant    0.0
```

```
---- District: C3----
                          venue  freq
0                     Hotel Bar    0.0
1            Italian Restaurant    0.0
2                 Internet Cafe    0.0
3            Salon / Barbershop    0.0
4              Tapas Restaurant    0.0
5                   Gaming Cafe    0.0
6                    Food Truck    0.0
7                    Food Court    0.0
8     Latin American Restaurant    0.0
```

```
9     South American Restaurant    0.0
10          Fast Food Restaurant    0.0
11              Sushi Restaurant    0.0



---- District: C4----
                            venue  freq
0                       Hotel Bar   0.0
1                Sushi Restaurant   0.0
2              Spanish Restaurant   0.0
3       South American Restaurant   0.0
4                      Restaurant   0.0
5                     Pizza Place   0.0
6              Mexican Restaurant   0.0
7       Latin American Restaurant   0.0
8                   Internet Cafe   0.0
9                  Ice Cream Shop   0.0
10               Tapas Restaurant   0.0
11                   Gaming Cafe   0.0



---- District: C5----
                            venue  freq
0                       Hotel Bar   0.0
1                Sushi Restaurant   0.0
2              Spanish Restaurant   0.0
3              Mexican Restaurant   0.0
4       Latin American Restaurant   0.0
5              Italian Restaurant   0.0
6                   Internet Cafe   0.0
7                Tapas Restaurant   0.0
8                     Gaming Cafe   0.0
9   Vegetarian / Vegan Restaurant   0.0
10               Fondue Restaurant   0.0
11           Fast Food Restaurant   0.0
```

```
---- District: C6----
                             venue  freq
0          Argentinian Restaurant   0.0
1                Sushi Restaurant   0.0
2              Spanish Restaurant   0.0
3       South American Restaurant   0.0
4                     Pizza Place   0.0
5              Mexican Restaurant   0.0
6        Latin American Restaurant  0.0
7              Italian Restaurant   0.0
8                   Internet Cafe   0.0
9                Tapas Restaurant   0.0
10                   Gaming Cafe   0.0
11                     Hotel Bar   0.0


---- District: C7----
                               venue  freq
0                         Hotel Bar   0.0
1                  Sushi Restaurant   0.0
2                Spanish Restaurant   0.0
3         South American Restaurant   0.0
4                 Salon / Barbershop  0.0
5         Latin American Restaurant   0.0
6                     Internet Cafe   0.0
7                    Ice Cream Shop   0.0
8                  Tapas Restaurant   0.0
9                       Gaming Cafe   0.0
10                       Food Truck   0.0
11   Vegetarian / Vegan Restaurant   0.0
```

6. Conclusion

In this study we were able to clearly identify the presence of different food related business in the central area of Rosario, a city of great economic importance in Argentina. With the data obtained, we can observe, for example, that Ice Cream shops are very popular in districts C1,2,3,5 and 6, but there is a lack of ice cream shops in districts C4 and C7. Further analysis of this kind might be done with the info obtained.

Analyzing further each district, and taking into consideration additional factors which may be related to how attractive each area is, how many people transit each day in the district, etc, may lead the stakeholders to make a final decision.