

Finding the best place to open a restaurant

Agustin Maciel

June 11, 2020

1. Introduction

1.1 Background

Understanding the best place to open a restaurant is critical to the future success of the business. Choosing the wrong place can make it a failure as not many customers will come in, and we need to consider the competition as well.

Montevideo is the capital city of Uruguay, a small country in South America. As a resident of this city, I decided to use it in my project. The city is divided into 62 neighborhoods in total.

Being a small city, we need to choose carefully where we want to open our food venue. When we think of it by the stakeholders, we expect from them to prefer the neighborhood where there is a lot of population, and where the income is higher as well.

We also need to consider the neighborhood where there is not as much competition as other neighborhoods.

1.2 Problem

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening a restaurant in Montevideo, Uruguay.

We will first analyze the neighborhoods to identify the median income and total population in order to create a ranking based on these metrics. Then we will try to detect which neighborhood has the least number of food

venues and from there we will choose a location that is not already too crowded with restaurants.

We will use our data science skills to generate a few most promising neighborhood based on these criteria.

1.3 Interest

Obviously, the stakeholders would be very interested in accurate prediction of the best place to open this business, for competitive advantage and business value.

2. Data acquisition and cleaning

2.1 Data sources

Based on the problem, the factors that will influence our decision are:

- median income in the neighborhood
- population in the neighborhood
- number of food venues in the neighborhood

Following data sources will be needed to extract/generate the required information:

- median income will be obtained using data from the National Statistics Institute [Neighborhoods income](#)
- population will be obtained using data from the National Statistics Institute [Neighborhoods population](#)
- number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
- Segmentation of the neighborhoods within the city will be obtained using a [JSON file](#) found on GitHub

2.2 Data cleaning

Data has been modified and cleaned as the original source provided the raw data in a format that was not compatible and also unnecessarily big considering our scope.

3. Methodology

In this project directed our efforts on finding a neighborhood in Montevideo that has a high median income, a high population, and a low number of food venues.

We limited our analysis to areas 800 meters around the neighborhood center as defined by the geospatial coordinates in the json file.

In the first step we collected the required **data: population and median income by neighborhood**.

In the second step in our analysis we calculated and explored the number of '**food venues**' in the top neighborhoods defined in the first step.

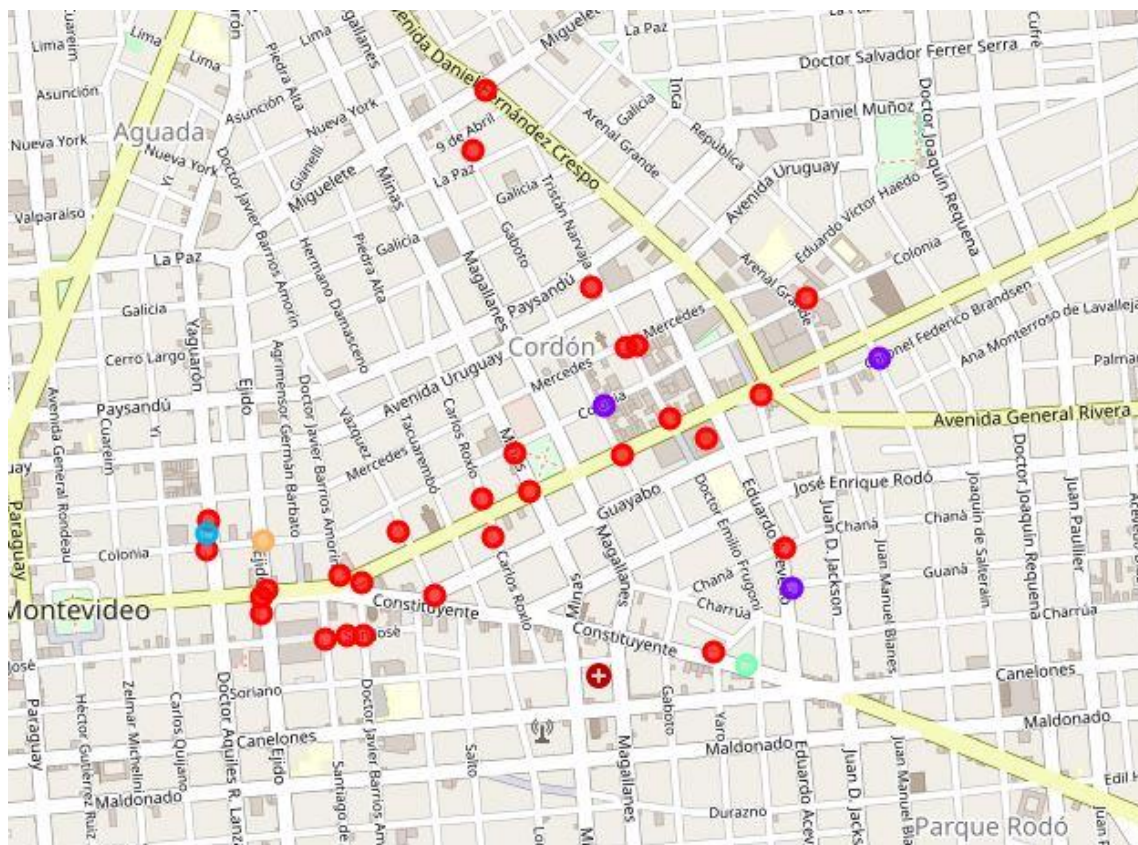
In the third and final step we focused on the best neighborhood and created **clusters of locations of food venues**.

We created clusters (using **k-means clustering**) of those locations to identify general zones which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

4. Results and Discussion

Our analysis shows that although the neighborhood with the most median income in Montevideo is "Pocitos", this is also the neighborhood with the most food venues from our selection, so we focused our attention in the second neighborhood, "Cordon", which offers a combination of closeness to the city center, strong ranking regarding population and median income and a relative low number of food venues. After directing our attention to this more narrow area of interest we clustered the food venues on this neighborhood to create zones of interest which contain greatest number of location candidates.

Result of all this is the following map where we can see all the food venues locations and this will be used as a starting point by the stakeholders in order to select the optimal location based on their criteria.



5. Conclusion

Purpose of this project was to identify areas in Montevideo low number of food venues in a high populated and strong median income neighborhood in order to aid stakeholders in narrowing down the search for an optimal location for a new restaurant.

By calculating the top neighborhoods regarding median income and population we have first identified a subset of strong candidates, and then leveraged data from Foursquare to identify how many food venues exist in these top neighborhoods. Once we had the best neighborhood which satisfied our requirements we used machine learning to cluster the food venues locations in order to create major zones of interest (containing greatest number of potential locations) and those were plotted on a map to be used as starting points for final "street level" exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on their criteria and taking into consideration additional factors like attractiveness of each location, proximity to major roads, real estate availability, prices, etc.