# Predicting the best neighborhood to open a restaurant

Agustin Maciel

June 13, 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 powers to generate a few most promising neighborhood based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

### 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 definition of our problem, 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.