

# **Predicting the best place to open a Beer Bar in Milan**

Claudio Brutti

November/December, 2019

## **1. Introduction**

### **1.1 Background**

Food and drink venues are very popular in Italy, and in Milan, finding a good place where to open a venue, without being overwhelmed by a crowd of similar venues in the near, it's very hard.

### **1.2 Problem**

The client has already opened successful beer bars in other towns and knows that the correct location is a key factor to have a good choice to ensure a enough customers to run the business. He wants to analyze the Milan city neighborhoods and use this information in order to find the best place.

### **1.3 Interest**

TBD

## **2. Data acquisition and cleaning**

### **2.1 Data sources**

To group the city locations we decided to use the postal codes (CAP) that divides Milan. A CSV file with streets and their correspondent CAPs was previously produced, to have a way to correctly place a CAP center in the city map.

A Foursquare data connection was used to retrieve venues in the proximity of each CAP center.

### **2.2 Data cleaning**

A geocoder from geopy library helped to localize a center point for a group of streets belonging to the same CAP.

For each of these CAPs we asked Foursquare to retrieve the venues in the nearby. Only the venues where drinking was involved were considered.

### **2.3 Feature selection**

We were only interested in the quantity of venues grouped by category in the nearby of a CAP center, so no features other than coordinates and category name were extracted.

## **3. Exploratory Data Analysis**

### **3.1 Calculation of target variable**

TBD

### **3.2 Relationship between ... ....**

TBD

## **4. Predictive Modeling**

TBD

## **5. Conclusions**

TBD

## **6. Future directions**

TBD