Predicting the Area to Open a French Restaurant in San Francisco

Toan Nguyen

August 26, 2019

1. Introduction

   1.1. Background

   Every year many new restaurants open, and some of them successfully grow, but some of them will die after a while. Choosing a correct area to open a business is the best factor that helps a business successfully. Customers would like to open a French restaurant in San Francisco area. They would like to choose an area in San Francisco whichever has low criminal rate and less competitive.

   1.2. Interest

   Using Marker and Map to group San Francisco criminal markers into different cluster shows criminal rates in different areas of San Francisco. Foursquare API can also be used to pull the list of French restaurants from low crime rate areas of San Francisco. Combination of both techniques will help us to determine which is the best areas that the customers can open his/her French restaurants.

2. Data acquisition and cleaning

   2.1. Data sources

   The first data source for criminal rate from police department incidents (https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-

data/CognitiveClass/DV0101EN/labs/Data_Files/Police_Department_Incidents_-_Previous_Year__2016_.csv). The second data source is from Foursquare API which helps to pull a list of neighbor French restaurants in low crime rate areas.

2.2. Data cleaning

There are around 150,500 crimes, which took place in the year 2016. Limit the data set to 1,000 records of crimes in San Francisco to reduce the cost of computation.

3. Exploratory Data Analysis

3.1. Grouping and map San Francisco criminal incidents:

Using criminal rate from San Francisco police department, we can plot and group the incidents by area in San Francisco. According to the result map (see figure 3.1-1), we can see there are 2 lowest criminal locations that have 2 incidents and 4 incidents. In order to get the latitude and longitude of these locations, we need to set the location data as the label for the map.
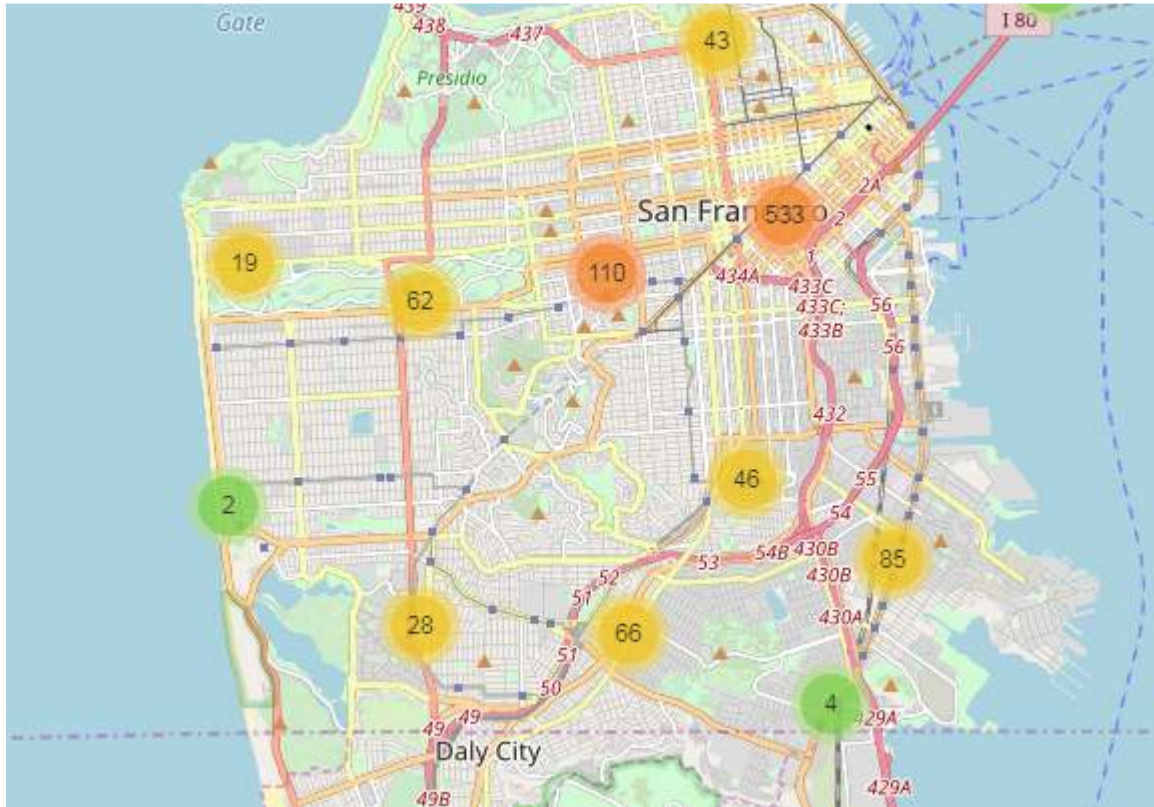
Figure 3.1-1: San Francisco Criminal Rate grouped by areas

3.2. Calculation of latitude and longitude of the lowest criminal rate locations:

By clicking on the label with 2 incidents, we can get the location (latitude and longitude). Based on the figure 3.2-1, the location has latitude= 37.7388739744874 and longitude = -122.50648456409. Doing the same for the area with 4 incidents, we can get latitude=37.7119353727691 and longitude = -122.405634443479.
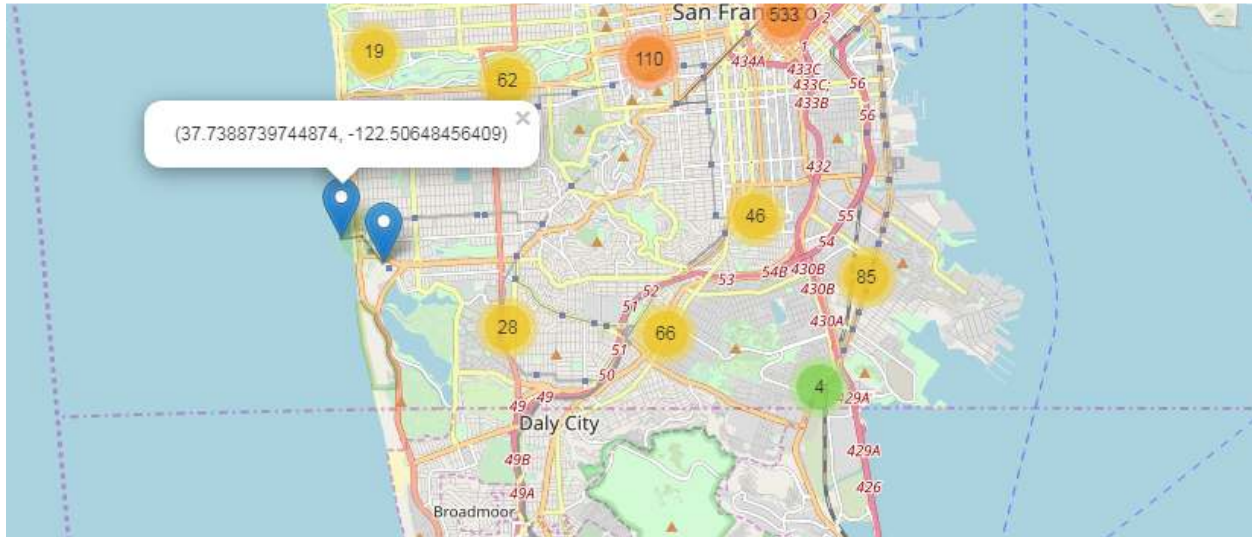
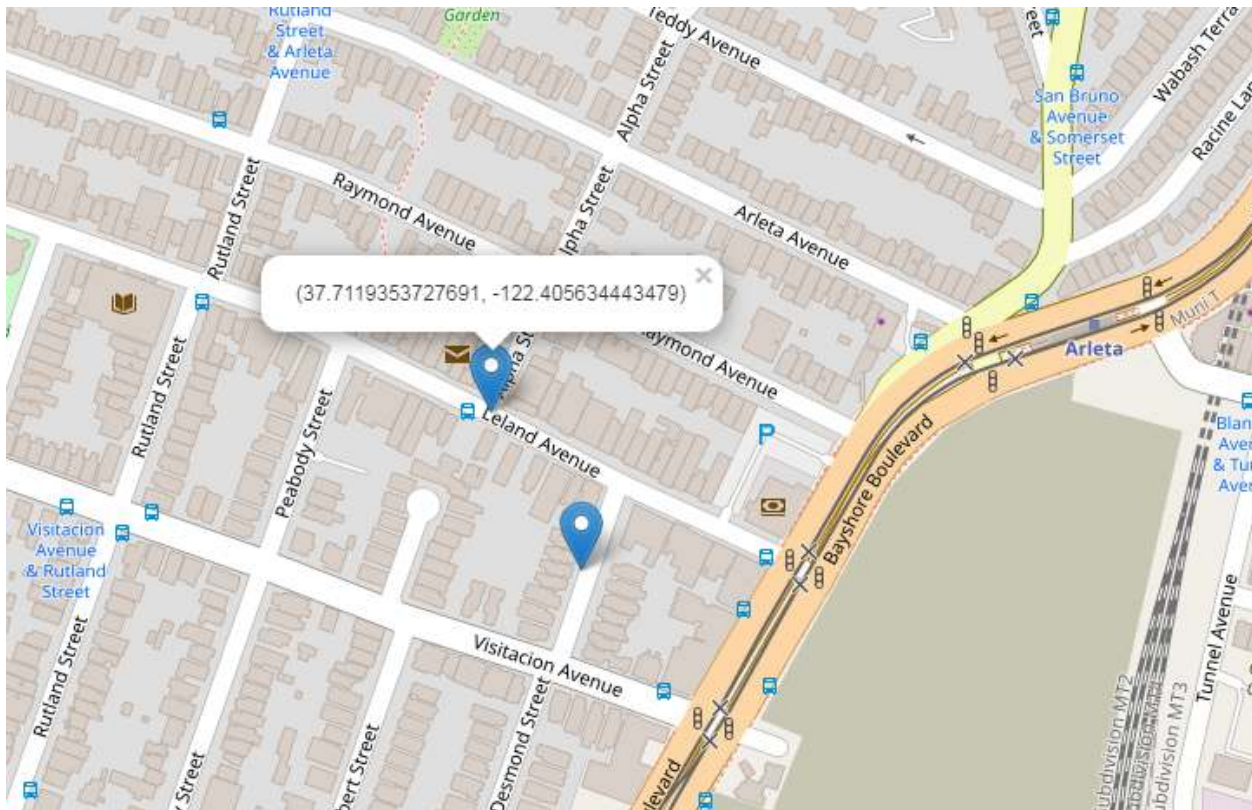Figure 3.2-1: The latitude and longitude of 2-incident location

Figure 3.2-2: Latitude and Longitude of the area with 4 incidents

3.3. Calculation of how many French restaurants around these safe areas:

We will use Foursquare API combining with the latitudes and longitudes of these safe areas in San Francisco to search for a French Restaurants within 3,000 meters. The location with 2 incidents has total 5 French restaurants, whereas the location with 4 incidents has total 3 French restaurants.