Predicting the Area to Open a French Restaurant in San Francisco

Toan Nguyen

August 18, 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 10,000 records of crimes in San Francisco to reduce the cost of computation.