Finding the Best Area to Start a Restaurant Business in San Francisco

IBM Data Science Professional Certificate
Capstone Project

Tianfang Chen
Oct. 2019

Introduction

Restaurant is one of the most common small business entrepreneurs start. It is also one of the business with least technical barrier of entry. At the same time, the restaurant business has very low profit margins, with the average number between three to five percent. Also, opening a new restaurant requires quite an amount of initial capital investment if renting or buying a venue. Therefore, it is very important to start off on the right foot by choosing the best area to open the restaurant.

Objective

This project aims to offer an infographic view of the demographics and restaurant competitions in each neighborhood of San Francisco. Anyone who wants to start a new restaurant in San Francisco can use the report or the interactive tool on the Jupyter Notebook as a guide to find the optimal place to start a restaurant based on the two elements.

Data Description

- Demographics data from San Francisco Planning Department (https://default.sfplanning.org/publications_reports/SF_NGBD_SocioEconomic_Profiles/2012-2016_ACS_Profile_Neighborhoods_Final.pdf)
  - total population, race percentages, and median household income for each neighborhood
- Restaurant competition data from Foursquare API
  - amount of different types of restaurants within each neighborhood

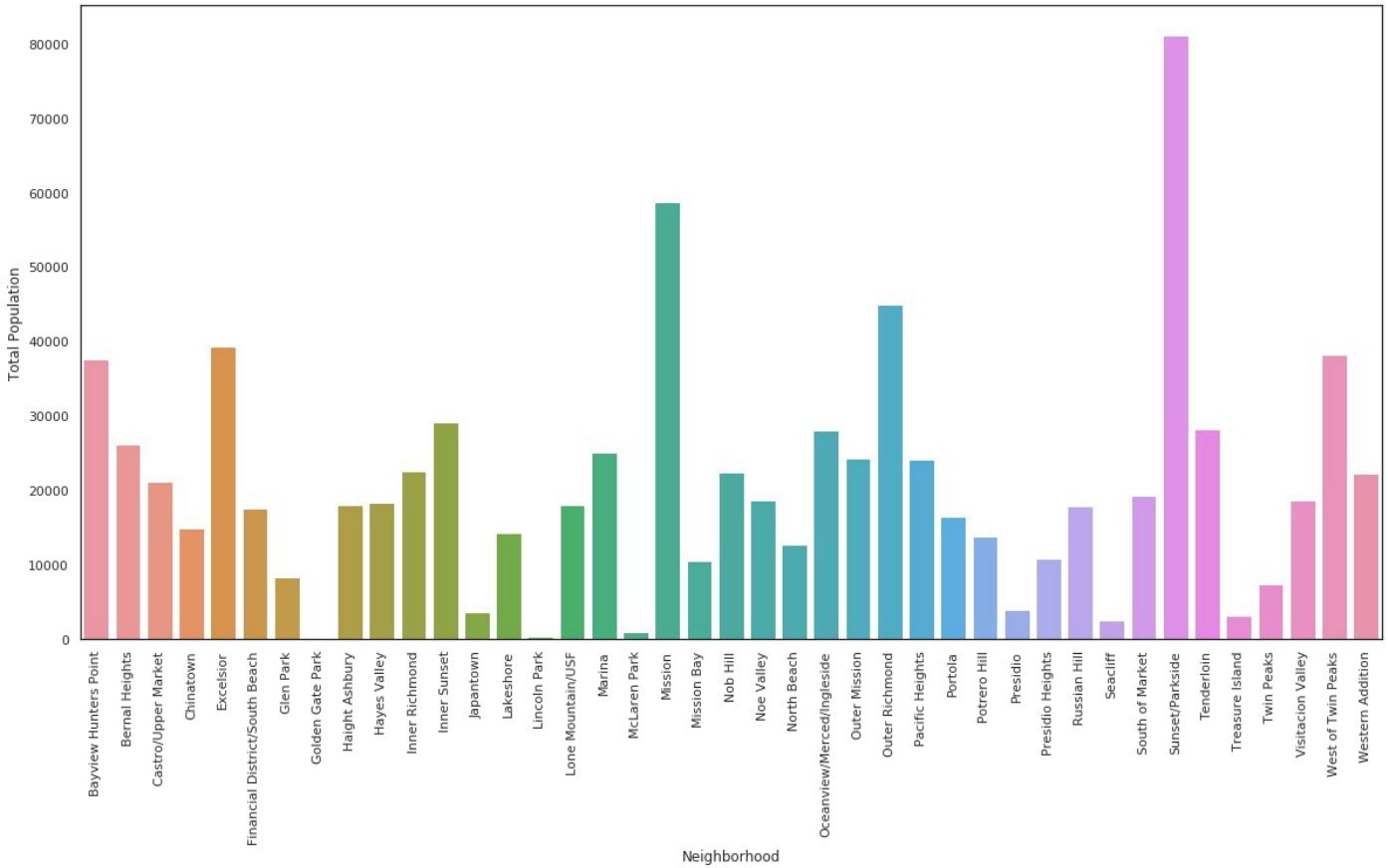
Methodology

Data Acquisition and Cleaning

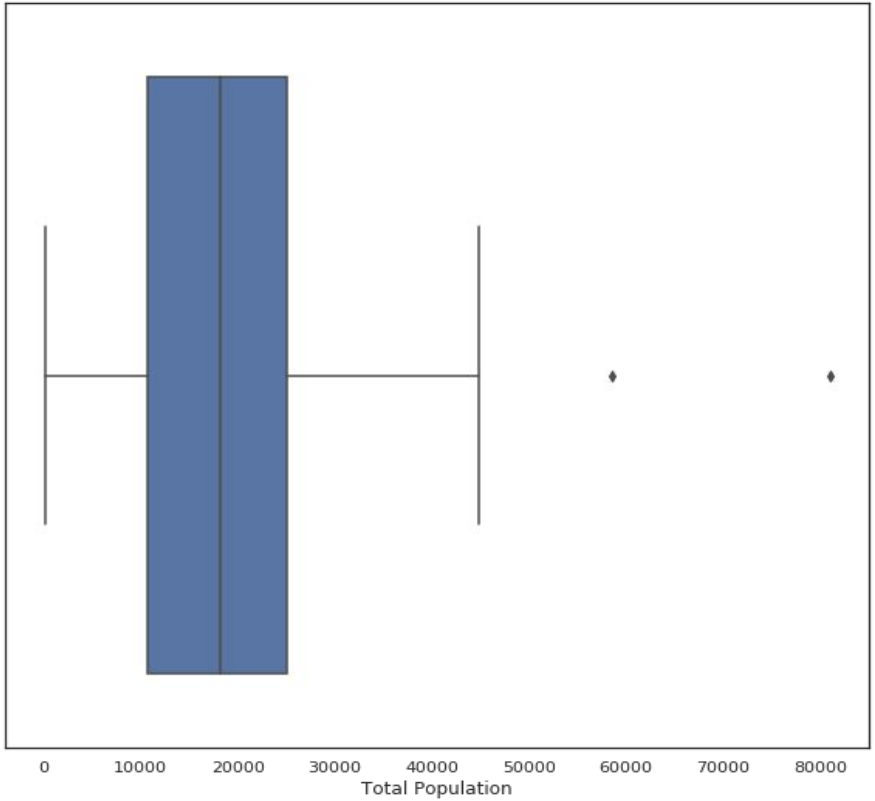
Exploratory Data Analysis

Creating Interactive Map

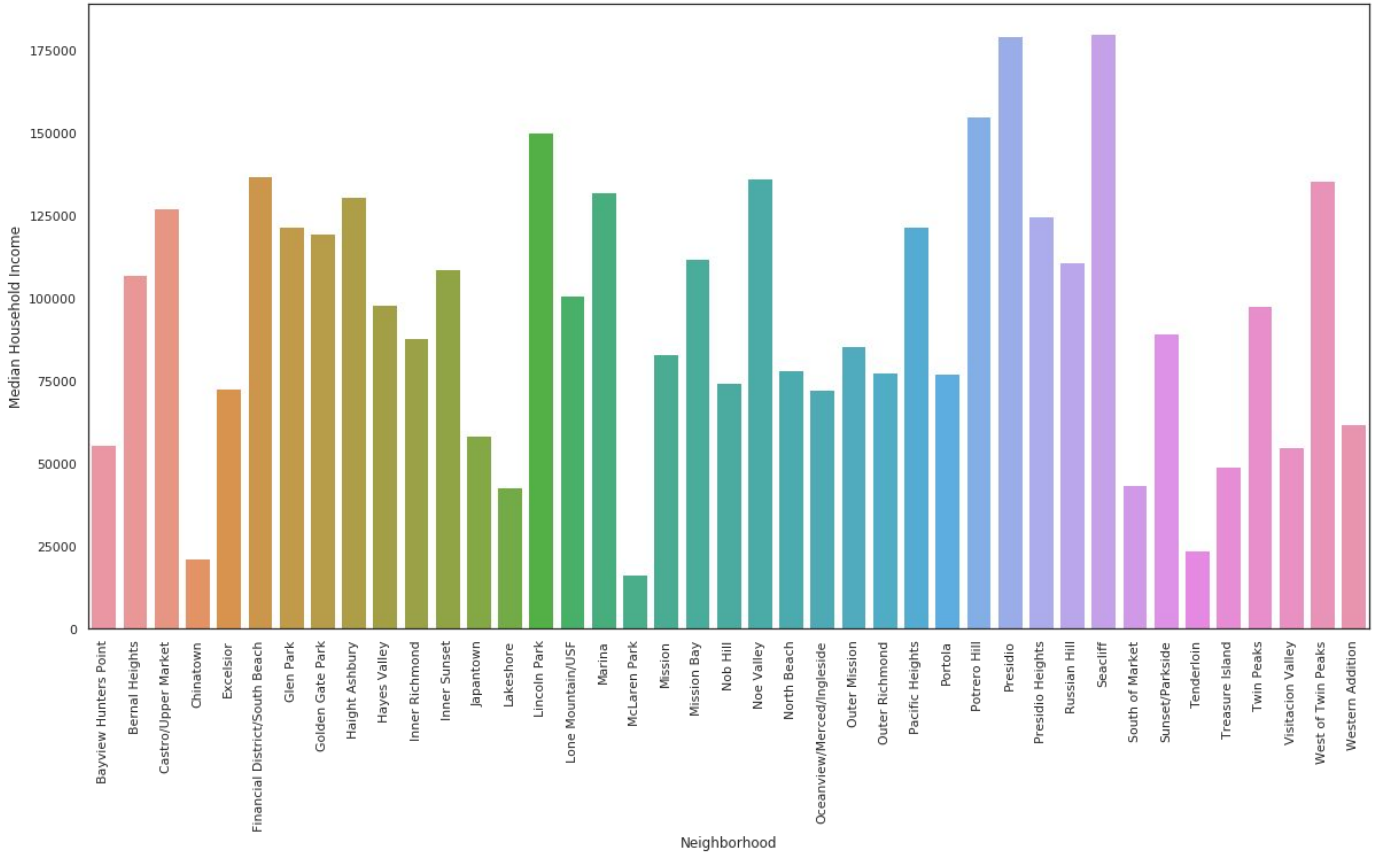
Results: Population



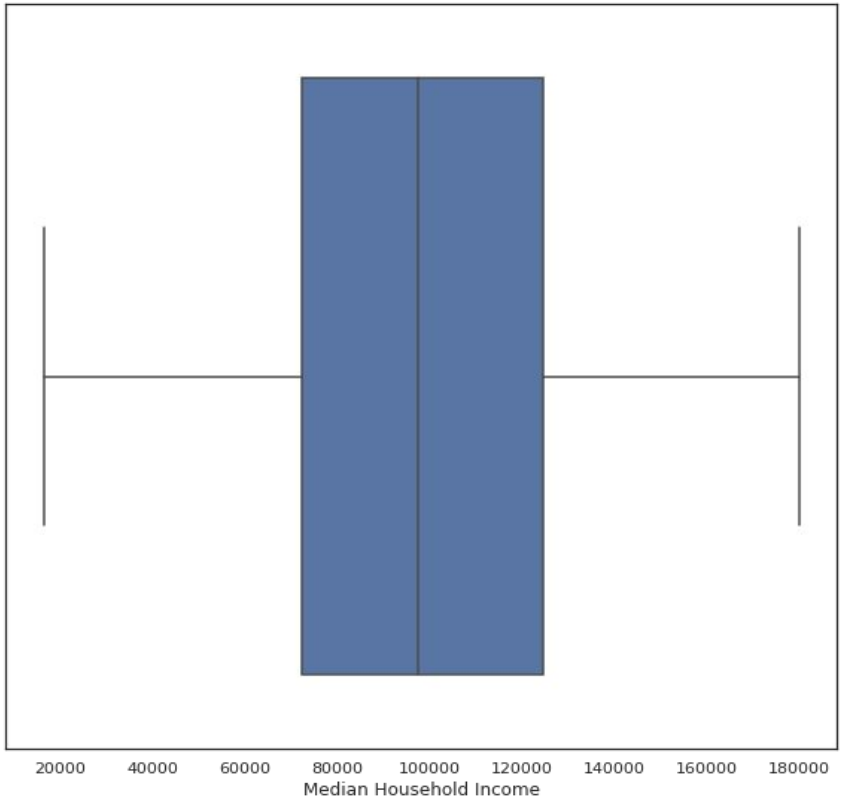
Results: Population



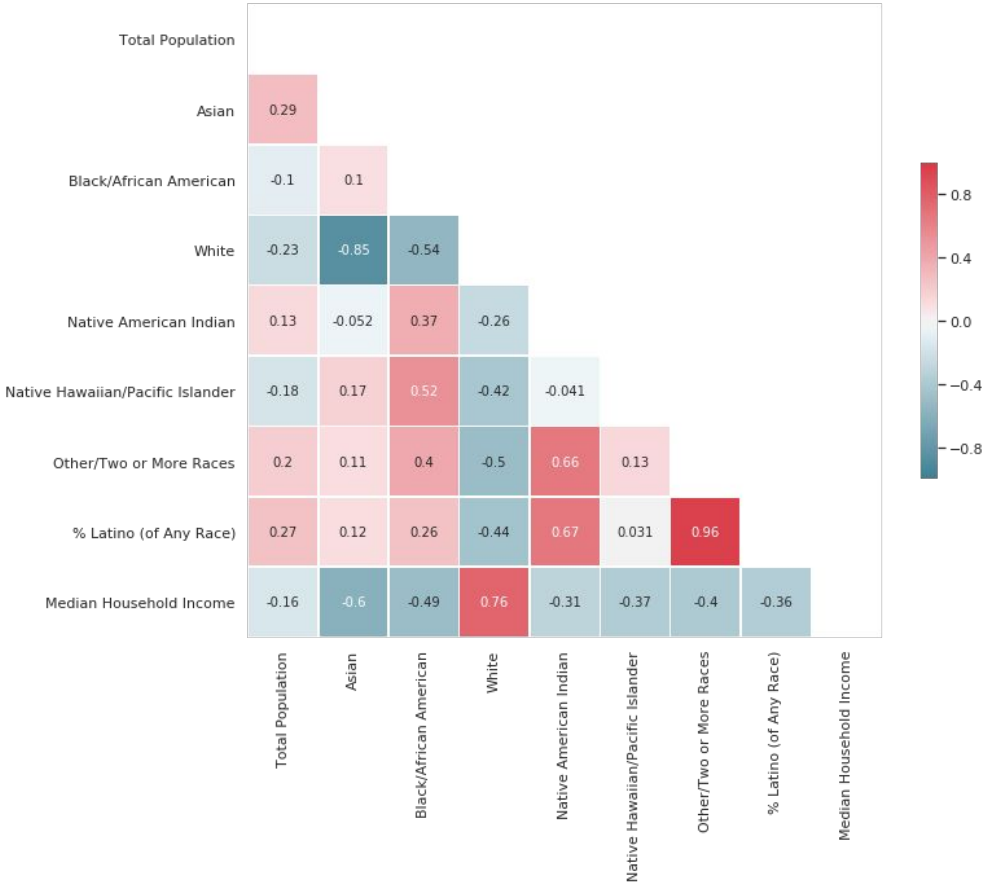
Results: Income



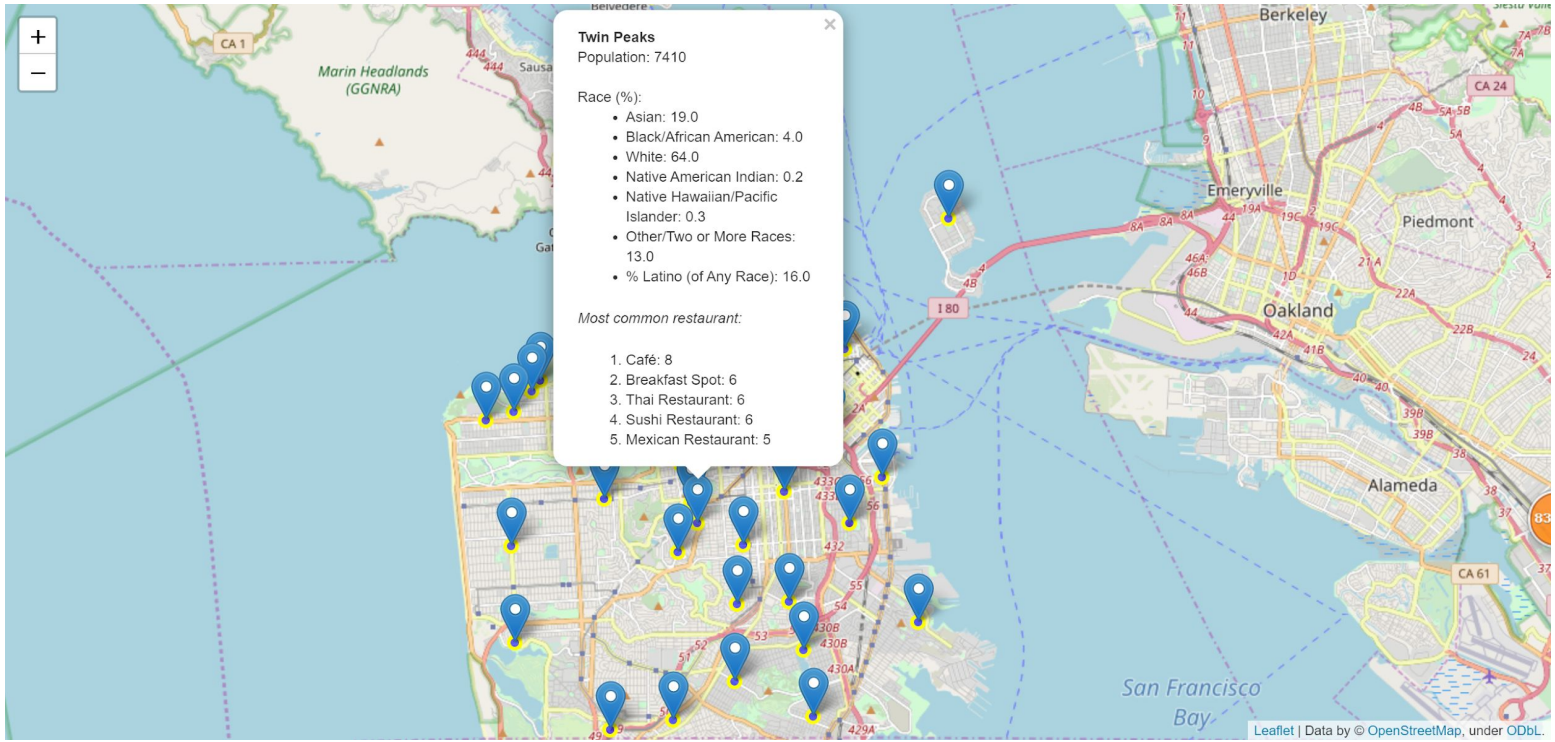
Results: Income



Results: Correlation



Results: Interactive Map



Discussion: Limitation

- The use of one mile radius for finding restaurants in the neighborhood. The latitude and longitude are not at the center of each neighborhood, and the neighborhoods' boundaries are not one mile circles centered at the latitudes and longitudes.
- Foursquare API has a limit of 100 for each request. The maximum amount of restaurants retrieved for each neighborhood is 100 while there might be more than that in the area.

Discussion: Future Improvement

- Circumventing the two limitations can be done by picking enough amount of distanced latitude and longitude pairs on the map, and finding restaurants near each location pairs. The method would generate a list of all the restaurants in San Francisco. And then the locations of each restaurant would be compared with the boundary of each neighborhood to create accurate lists of restaurants in each neighborhood.
- Develop a search algorithm--when a prospective restaurant owner searches for a type of restaurant, the tool will give the recommendation automatically.

Conclusion

In this project, we have acquired and cleaned data from online sources, performed exploratory data analysis, and created the interactive map in Jupyter notebook. Prospective restaurant owners can visualize relevant information of each neighborhood with the interactive map, and therefore choose the best place to start the restaurant easier and better.