Restaurants in Glendale, CA

By Amir Tayyebian

10/2019

IBM Data Science Specialization

Capstone Project Part I

**Introduction**

1. **Background**

Glendale California is a city in Los Angeles county with a population of over 200,000, due to the cities’ large Armenian population, Middle Eastern are also widely available. Believing that too much competition exists to open a Middle Eastern restaurant, our client is looking to open either a Thai, Mexican or Italian restaurant in Glendale. Our client is primarily concerned about the number of competitors for each type of restaurant. Additionally, our client wants to know which Glendale zip code (other than heavily residential zip codes 91208 and 91207) is best suited for this new restaurant

This report will first discuss the data requirements and methodology used to make this recommendation.

1. **Problem**

The purpose of this project is to answer the following questions:

1. Can Foursquare location data be used to provide a recommendation for our client regarding which of the three cuisine types should they choose.
2. Can Foursquare location data be used to provide a recommendation for which zip code is best suited for this new restaurant
3. **Interest**

This project will be of interest particularly to anyone who is trying to open a restaurant in Glendale CA. Additionally, anyone who is looking to open a restaurant anywhere in the United States can use the methodology in this project for similar insights.

**Data**

1. **Requirements**

There are several data entities that is required for this project. First, a list of all the zip codes in Glendale, CA is needed, along with the longitude and latitude values for the city of Glendale and each of these zip codes. This data will be used alongside the Foursquare API to gather data on the venues in the city of Glendale in order to decide which restaurant type should be chosen. After the restaurant type is decided, additional data is required to find the best zip code to open the restaurant in. This data will be included population data to find the number of potential customers. Population data will be combined with additional Foursquare API data on the different venues in each zip code, and the ratings for these venues.

1. **Sources**

United States Zip code data is publicly available in csv format, and can be downloaded by clicking the following link:

<https://simplemaps.com/data/us-zips>.

1. **Cleaning**

Relevant data entities in this file includes Zip Codes, City, Population per zip code, and the longitude and Latitude values for each zip code. These columns will be stripped from the csv file in Python to create the primary database for this project. The Foursquare API will also be cleaned to include only venues that fall under “Thai” and “Italian” and “Mexican” restaurants located inside Glendale. Additionally, these venues will be grouped by zip code for more detailed analysis.

Foursquare venue data was gathered by separately searching the queries Mexican, Italian, and Thai, with the three resulting data frames cleaned and ultimately joined to create the 37 resulting venues. Additionally, the data was cleaned in order to find non-restaurant venues that may have been included in the data.