IBM Applied Data Science Capstone
Lets eat in Los Angeles CA
Analyzing the Neighborhoods to start a new Restaurant

By Frank Andrade

Introduction

As one of the largest cities in the United States, Los Angeles is well known for its Film industry, celebrities, and entertainment. From enjoying time on the sanding beaches of the Pacific Ocean to nightlife excursions in the clubs to major sporting events, not least of which are the scheduled Summer Olympics of 2028, there are many different aspects of Los Angeles that will draw tourists from near and far for many years to come. With so many people traveling to and around Los Angeles, there is a relatively high demand for various restaurants, especially with how many cultures are meeting in one city. The goal of this project is to determine any possible locations to start a restaurant or eatery to meet this demand. The main focus will be to scrap and analysis data to provide any potential entrepreneurs and/or business owners with information on any possible future investments

Data Collection

The information that is needed to complete this project will be collected from multiple sources.

The data that is required for this project are as follows:
1) Neighborhood data on Los Angeles
2) Geographical coordinate of Los Angeles and all nearby neighborhoods
3) Neighborhood Venue Data in Los Angeles

Neighborhood Data

The data for the LA neighborhoods will be gathered and scraped from the following site using the read_html() method to read the data into a panda data frame: https://en.wikipedia.org/wiki/List_of_districts_and_neighborhoods_in_Los_Angeles. The reasoning for using this site is that Wikipedia has a plethora of comprehensive and detailed tables on various cities, including LA. Using the Wikipedia page for LA will allow for easy scraping of the site via the read_html() method of panda.

Geographical Coordinates

The GeopPy library in python will be used to gather Los Angeles's geoograpical coordinates. This will be important when using the python's Folium library to plot a map of Los Angeles. To acquire the latitude and longitude data points for different neighborhoods within Los Angeles, the geocoder library will be used. These data points will then be compared to data scraped from Wikipedia and updated if the difference exceeds the code parameters before using the Folium library for plotting the coordinates.

Venue Data

Foursquare API will be utilized to extract and provide the venue data. The data will contain venue recommendations for different neighborhoods of Los Angeles, which can be used to study popular venues in the various neighborhoods.