Battle of the Neighborhoods – IBM Capstone Project

Where to develop a Karaoke Bar

Introduction

A small investment company inquires where they should best allocate their resources to develop a Karaoke-themed bar. They are looking for an area where starting a new business will attract the right customers and generate the highest revenue.

Background

They've been racking their brains trying to pinpoint a city, but the group just can't come to a consensus. Luckily, they heard about you through a mutual colleague that you specialize in data analytics and could come up with best possible solutions based on data. What they're looking for. Generally speaking, they're looking to define customers spending patterns in various localities to determine the best U.S. based location to set up their karaoke business.

Business Problem

In order to help the investment company from maximizing profits and minimizing loss, I will definitely be incorporating data from websites to define which areas are the most densely populated, have higher than average incomes, use the Foursquare API to obtain venues in various U.S. cities, then categorize the venues based on proportional weights to designate value. Once the city(s) is determined, we can narrow in on distinct neighborhoods and obtain visual representation of the coordinates. In particular, we want to focus on the venues listed for each locality (using Foursquare API) and determine why a certain location could be considered more appealing.

Data

Data will be scraped from (https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population) and (https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income) to obtain a pandas dataframe using the 'BeautifulSoup' library discussed in the modules. Clean the data and pre-process for exploratory data analysis. The two Wikipedia websites should provide enough information to determine which location would provide the best potential return on investment for the company.