Coursera Capstone Project

Neighborhood Analysis by Venue Ratings

Dwayne Thaele

January 10, 2021

Background and Problem Description

Applications like Yelp and Google Maps are used to identify venues in a city and assess their suitability by ratings submitted by past customers. These applications have proven themselves to be valuable. However, for application users not familiar with a city, they may not obtain full context of the venues and their locations. For example, someone may wish to visit a nice restaurant in a nice area of a city, or avoid bad restaurants in bad parts of a city.

Data Description

Foursquare is a US business that uses detailed location data, along with business, and customer input to "tap into this intelligence to create better customer experiences and smarter business outcomes." See: https://foursquare.com.

Foursquare provides a user-based rating system where Foursquare used feedback from customers, combined with a wealth of other data and artificial intelligence techniques to calculate overall ratings for a venue. Rating are numeric from 0 to 10, 10 being the highest rating; a 0 rating generally used for venues that have not received a rating. For a detailed discussion about Foursquare's rating system, see: https://medium.com/foursquare-direct/finding-the-perfect-10-how-we-developed-the-foursquare-venue-rating-system-c76b08f7b9b3

For this project I'll be using Foursquare's venue ratings and location data that is provided by their API. For a reference about the API, see: https://developer.foursquare.com/docs/apireference/venues/search/

At present, my project will only focus on venues that are restaurants. Depending on complexity I may add additional venues.

Additionally, my project will only focus on the city of San Francisco. However depending on complexity, I may include the cities of Chicago and New York.

Problem Solution

The foundation of my project *Neighborhood Analysis by Venue Ratings*, is an assumption that venues with high ratings are generally in more desirable areas, where venues with lower ratings are generally in lower desirable areas. This serves the primary use case where a visitor, unfamiliar with an area or local venues would like to visit the are and decide on a venue after they arrive and *check-out* the neighborhood. The visitor may even decide to spend more time in the neighborhood if its appealing to them.

I will group and color code venues based on their ratings:

Venue Rating Range	Color Designation
9 - 10	Green
7 - 8	Blue
5 - 6	Yellow
3 - 4	Orange
1 - 2	Red

I will cluster the results and depict them in an overlay of a city map.

The outcome will be a map that reveals the more and least desirable areas based on rating. This provides users with a simple way to determine which area are likely to have the most positive (dining) experience.