

1. Introduction/Business Problem

Buying a house is an extremely daunting task. You want to first consider where your house is located. This type of research considers different both economic and environmental factors, such as the proximity to hospitals, restaurants, and parks, to name a few. Once you have found a desirable location, you must find a house that matches your desired expectations. This process is lengthy and difficult. This capstone project aims to help customers narrow their search by providing a tool that allows them to compare and contrast four different locations, based on their surrounding environmental factors. This will aim to minimize the difficulties presented in the first step of buying a house in four locations in the Bay Area – Palo Alto, San Francisco, Redwood City, and Atherton.

2. Description of Data

2.1. Foursquare API data

By using the latitude and longitude coordinates of a particular location, the Foursquare Places API is able to provide a list of venues associated to the location. In order to obtain a list of venues within a specified area, we use the *explore* endpoint from the API via an HTTP request, which provides us with a JSON object. This JSON object gives details on the *location* (coordinate of each venue) and the different types of *categories*. This provides valuable information regarding various types of venues (e.g. entertainment, dining, etc.) surrounding a particular location.

2.2. Neighborhood Geographic

OpenCage Geocoder (opencagedata.com) provides an API to convert coordinates to and from places. This allows easy access to get the longitude and latitude details of the four locations.

2.3. Real Estate History

Real estate details, such as the average price of a house in a location and the housing market increase or decrease from year to year, can be collected from Zillow (Zillow.com) to make more informed decisions in creating this tool.

2.4. Zip Code Information

To retrieve further insights into each zip code, a special python database named uszipcode (pypi.org/project/uszipcode/) was used.