Introduction / Business Problem

Suppose you are looking to open a restaurant in San Francisco. San Francisco is huge, so you might be looking to narrow down which general neighborhoods might be suitable. To find a suitable neighborhood, we will mainly look at general restaurant prices in all neighborhoods of San Francisco, with the idea that if you want to establish a fine dining restaurant, you may not want to do so in a neighborhood where restaurants tend to be very cheap, or vice versa.

As an extension, suppose you own a restaurant chain with existing locations in certain neighborhoods. What neighborhoods might be best to expand to?

We will help to provide an answer to this business case by clustering neighborhoods in San Francisco by restaurant pricing. By doing so, we can see what cluster your new restaurant would belong to and thus find suitable neighborhoods for it.

Data

First, we will need to identify all coordinates of neighborhoods in San Francisco. Using those coordinates, we will use FourSquare's Places API, namely the Venue group and Explore endpoint to find all restaurants/food locations within a certain radius of each neighborhood that are classified in each of FourSquare's four price buckets: $ to $$$$. We will then process this data, analyze it, looking at the characteristics of restaurants within each neighborhood, and finally provide a k-means clustering of the neighborhoods by restaurant price.