**Introduction / Business Problem**

The San Francisco Bay Area is a popular destination for people who want to live in a diverse cosmopolitan location with significant job opportunities and a comfortable climate. With years of development and growth, many cities have sprung up around San Francisco, contributing to a suburban sprawl that offers vastly distinctive options for potential residents. However, with the sheer number of cities, it is also increasingly difficult to research and find the best city or cities around which to focus a new home search. Data Science can be used collect and analyze data from disparate sources to arrive at a short-list of cities based on potential homeowner preferences.

**Data**

This analysis uses data from the following two sources.

**Wikipedia** – This is a good source of reference data compiled by an extensive user base. The data is often displayed in table format from a page written in html. A number of web-scraping techniques can be used to extract the data required for analysis. This analysis uses the html parser BeautifulSoup for this purpose.

1. Cities in the San Francisco Bay Area (sample screenshot below) (<https://en.wikipedia.org/wiki/List_of_cities_and_towns_in_the_San_Francisco_Bay_Area>)

A screenshot of a cell phone

Description automatically generated

1. Crime rates for cities in the San Francisco Bay Area (sample screenshot below) (<https://en.wikipedia.org/wiki/California_locations_by_crime_rate>)

A screenshot of a cell phone

Description automatically generated

**FourSquare** – This is a good source of information about venues of multiple types located around a specified latitude/longitude (sample screenshot below).

A screenshot of a cell phone

Description automatically generated

Potential residents are likely to be interested in knowing about neighborhood venues when they decide where to relocate, so this information can be used to characterize each Bay Area city and to help focus a search on cities with desirable venues.