

Where to Relocate in California after the COVID-19?

Introduction/Business Problem

COVID-19 is an infectious disease caused by a coronavirus newly discovered in 2019. It is highly contagious and spreads fast from person to person in close contact. After the World Health Organization (WHO) declared the COVID-19 outbreak a global pandemic on March 11, 2020, many regions and areas issued the stay-at-home order, which means a number of non-essential workers need to work from home.

An interesting result of work-from-home during the COVID-19 is that people start to realize the business goes just fine when employees work from home in pajamas. There are 27 United States companies, most in IT, that have announced permanent work-from-home decision by September 2020. As people can work from anywhere, they try to escape from the areas of sky-high home values and choose a place where they can enjoy time with family (and friends after the pandemic).

California is one of those regions. California, especially the Silicon Valley area, has the clusters of world-leading IT companies and its home value is constantly among TOP 3 in the United States. After those companies announced long-terms work-from-home decisions, many employees then decide to relocate to where they can live in the styles that they always want.

The business problem of this project is to find clusters of neighborhoods in northern California that fit different lifestyles for the people considering relocation.

Targeted Audiences/Stakeholders

The primary targeted audience of this project are those who are looking for a place that fit their lifestyles.

Data

Based on the definition of the business problem, the factors that would affect people's decision include:

1. The safeness of the borough based on crime statistics
2. The ranking of public school systems based on school ranking data
3. The medium home values of the borough based on real estate data
4. The most common venues

Data Sources and Methods to Extract the Data

The data sources that will be used for this project:

1. [California Offenses Data by County and City](#): a dataset containing the crimes submitted either through a state Uniform Crime Reporting program or directly to the FBI's program.
2. [US Schools Dataset](#): the dataset taken from the US Department of Homeland Security and contains information of all public and private schools with their geographical distribution.
3. [Housing Data](#): the dataset downloaded from Zillow Research which contains the seasonally adjusted measures of the typical home value in the 35th to 65th percentile range.
4. [Dataset of the Neighborhoods, boroughs, and the most common venues along with co-ordinates](#): This data will be fetched using Four Square API to explore the neighborhood venues.