

# Exploring the Mature Communities in Satellite Cities of Sacramento Metropolitan

## 1. Background Introduction

Sacramento, a city built in the gold rush in 19th century, is located in North California. Sacramento city, together with a number of satellite cities, constitutes the Sacramento metropolitan, which is the fifth largest metropolitan in California. Sacramento is the home to more than 500k people, two commercial airports and the California state government. Sacramento has evolved from a lodging station to an important nexus of the transportation from Western coast to inland American. It is the intersection of two interstate highways, and a great number of trains were dispatched to places all over the USA. In recent years, Sacramento has attracted tremendous attention

Located in the northernmost part of Central Valley, Sacramento is 100 miles away from San Francisco, and 180 miles away from Silicon Valley, the hub of IT giants and thousands of high-technology companies. Due to these geographic advantages, Sacramento has drew extensive attention of companies in Silicon Valley from recent years as an option to relieve the tighter land supplies and increasingly higher operating costs due to the infrastructure of Bay area is becoming saturated and overwhelmed. Nowadays and in the foreseeable future, companies are spreading their new branches in Sacramento areas, moving their manufacture to avoid heavy costs. As a result, Sacramento metropolitan is expected to witness a rapid growth in next decades.

Sacramento city itself has little space due to earlier city plans. Therefore, exploring the neighborhood of Sacramento is of great significance for either companies or their employees who are relocating. Satellite cities, such as Elk grove, Rancho Cordova and Roseville were built for different reasons, and these communities have various degree of development. Using geographic data can reveal if one community is mature or not by gauging the ratio among each category of amenity facilities, such as dining, commerce and leisure, as the ratio tends to become equilibrium as a matter of city development in a given region. These parameters are used to gauge the maturation of a community, which is may be a deal-breaker for employees who could consider relocating and for those who has housing-buying intention.



Figure 1 Skyline of Sacramento

## 2. Data Sources & Treatment

### 2.1 General Workflow

The geographic data were looked up online by the zip code of each community of Sacramento. The raw data contained the coordinate of each zip code and was processed to filter off locations in Sacramento city. Then these coordinates were passed into Foursquare API to retrieve the neighborhood venues. The information was further processed and reorganized filtered off irrelevant venues. The amenities were classified into three categories: dining, commerce and leisure, based on which the venues in these cities were grouped. Top 5 cities with most amenity venues were then selected. The composition of amenity in these cities was analyzed and the maturity of these cities was evaluated as a whole.

### 2.2 Data Processing & Methodology

Neighborhood/Zip code data was found online. These data was input into Excel and converted into csv files. Before being processed, the data was reformatted to make it fit to the following processing.

Meanwhile, based on the zip code, coordinates of the corresponding neighborhood were found in order to retrieve information from Foursquare API. These coordinates were reformatted and split into longitude and latitude, respectively. The coordinates, neighborhood and zip code were combined into a new data frame, followed by removing entries with regarding to Sacramento city, which counts more than 80% of overall data.

The refined data was saved as csv in the name of 'SacNeighProcess.csv'. Coordinates in the data frame were passed into Foursquare API to retrieve venue information of the neighborhood. As over 70% of venues are focused in 5 communities: Elk Grove, Citrus Heights, Rancho Cordova, Carmichael and Folsom, these cities are considered populated and more likely to have mature community, therefore further study was focused on these cities.

The amenity venues in these cities were classified into three categories: dining, commerce and leisure, which are the business that covers routine living necessities. The grouping was achieved based on the presence of a number of key words appended in the list 'value', which are the characteristics of each type of amenity. After that, the number of each type of amenity in these cities was listed and the city with the largest number of amenity and the composition of amenity were visualized by bar charts and pie charts.

## 3. Data Analysis and Discussion

### 3.1 Overall Venues Distribution

1326 venues acquired from Foursquare API are concentrated in the following 10 cities, (Figure 2), which counts for 90% of overall entries (1433 in total).

Neighborhood	Elk Grove	Citrus Heights	Rancho Cordova	Carmichael	Folsom	Fair Oaks	North Highlands	McClellan	Antelope	Orangevale	Total
Count	328	218	205	157	152	72	63	47	43	41	1326

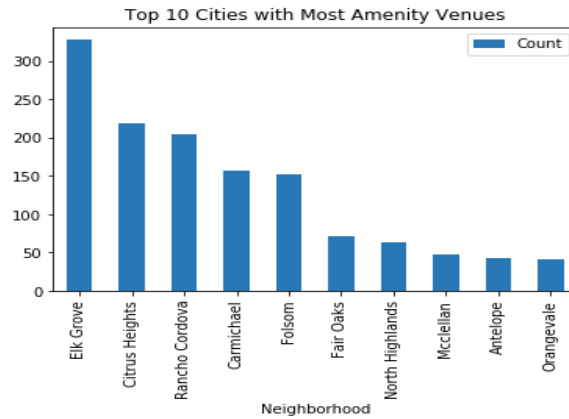


Figure 2 Numbers of Venues in Top 10 Cities

In all, overall 300 venues are found in Elk Grove, which is around 100 more than first runner-up, Citrus Heights. Meanwhile, there is a dramatic drop between Folsom, which ranks the fifth, and Fair Oaks, the sixth place, by around 80 venues. Therefore it is reasonable to consider the first 5 cities as populous and the communities in these cities are considered to be prosperous. The following analysis will be spent on these cities.

### 3.2 Analyze of Composition of Venues in Selected Cities.

Amenity cities are classified into 3 categories for each city: dining, market and leisure to gauge the cities matureness and livability.

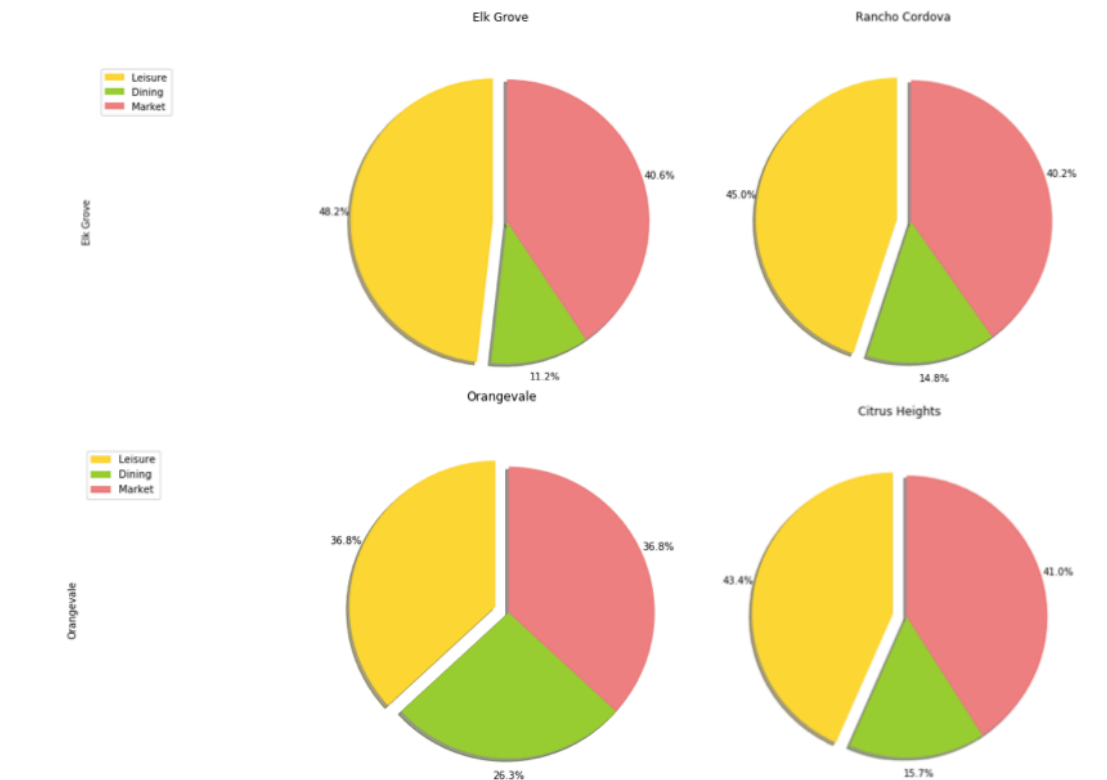


Figure 3 Compositions of Venues in Elk Grove, Rancho Cordova, Orangevale and Citrus Heights

Elk Grove, Rancho Cordova, Citrus Heights and McClellan were selected for analysis. The first three cities are the most prosperous, while Orangevale has

much less amenity venues as a contrast.

In Elk Grove, Rancho Cordova and Citrus Heights, leisure sites, such as coffee shop, teahouse, gym and park consists the majority of amenity venues in all cities, which ranges from 40% to 50%. Commerce facility, such as supermarket, grocery stores and convenient stores, is the second highest, which dining sites, which is the least among three categories, only consists 10% to 15% of overall amenity venues.

In Orangevale however, the ratio is distorted compared to the abovementioned three cities. It has a higher portion of dining facilities, while lower in leisure and market sites. The demographic data can be used to assess the validity of using the category ratio to measure cities' matureness and prosperity. Based on Wikipedia, population in Elk Grove reached 170k in 2017, while Citrus Heights and Rancho Cordova had 100k then. Orangevale on the other hand, cannot even be called community, which had a population less than 30k. It suggests that Elk Grove, Citrus Heights and Rancho Cordova are relatively more prosperous, have more mature community amenity and therefore more livable for persons who wish to relocate their career to Sacramento metropolitan. It also implies that using the ratio of each amenity category could reveal the matureness of a community without knowing the demographic details.

#### **4. Conclusion & Future Direction**

The report examined the distribution of amenity venues in the satellite communities in Sacramento metropolitan. In each community, the venues are divided into three categories and the portion of each category to total number amenity venue in that community is calculated and visualized. It unveils that for a populous and mature community, this ratio seems to fluctuate in a small range. One possible assumption is that the demands for each type of community in less populated region or an emerging city may be significantly insignificant or deviating from its true demands, respectively. As the city's development proceeds, the increase in population will eventually reflect the true demands for each amenity, and the demands will come back to normal track in the long term, therefore brings the ratio to normal range.

It suggests that Elk Grove, Rancho Cordova and Citrus Heights are the top three populous cities around Sacramento. More detailed visualization can be done if suitable geojson can be found to provide more geographic insights of these cities. Also, this report only reflects these cities' development from the past. It has little to do for predicting cities' potential in the future. Using machine learning techniques, and data such as project in progress, distance to city and crime rate, cities development can be predicted, including their afflicted value, such as house pricing and living expense.