Coursera Data Science Capstone Project

Rent and Venues in Palo Alto

Ian Yin

02/17/2020

1 Introduction

The main question that this project aims to determine is the relationship between venues and housing rent in the neighborhood in Palo Alto. In another world, this project aims to answer neighborhoods with what kind of venues have higher rent cost. The answer for this question is especially important for group of people who are looking into purchase of a new house for rent/investment purpose. For this group of people, it is important that the house purchased can generate steady cash flow. With the help of this project, investor can get the idea of rent that they can receive before making a purchase by looking at venues around houses.

2 Data

Foursquare location data and housing rent price in Santa Clara county are used for this project. Rental data for each neighborhood in Santa Clara County can be obtained from "rentcafe.com". Rent price used for this project is the average rent in each neighborhood. Coordinates for each neighborhood are obtained from Google's geocode API. Each neighborhoods' surrounding venues data are obtained from Foursquare location data. With venues data, I can find top 10 venues categories for each neighborhood. With top 10 venues categories data together with average rent data for each neighborhood, I can classify neighborhoods into cluster and find the relationship between these clusters and rent price.

3 Methodology

The objective of is to see neighborhoods with what kind of surrounding venues have highest rent. Therefore, I need to obtain average rental rate data for each neighborhood in Palo Alto. Python request and beautiful soup package are used for web scraping of rental data. By observing rent rate data, I found rent price

in Palo Alto have good amount of variation, which range from \$2693 to \$4086 per month. The mean of the rent data is \$3340, which is close to median of \$3426. Next, using Google's geocoder's API, latitude and longitude data for each neighborhood are obtained. Using coordinate of each neighborhood, Foursquare API is used to get top 50 venues that are within radius of 500 meters for each neighborhood. By taking the mean of the frequency of occurrence under each venue categories for each neighborhood, I am able to determine top 10 venue categories for each neighborhood in Palo Alto.

This study utilizes the K-mean clustering algorithm to answer the main question. It is one of the most popular method to cluster the geographic data. The result of K-Mean clustering allow me to examine what venues are most common in each cluster and their associated rent.
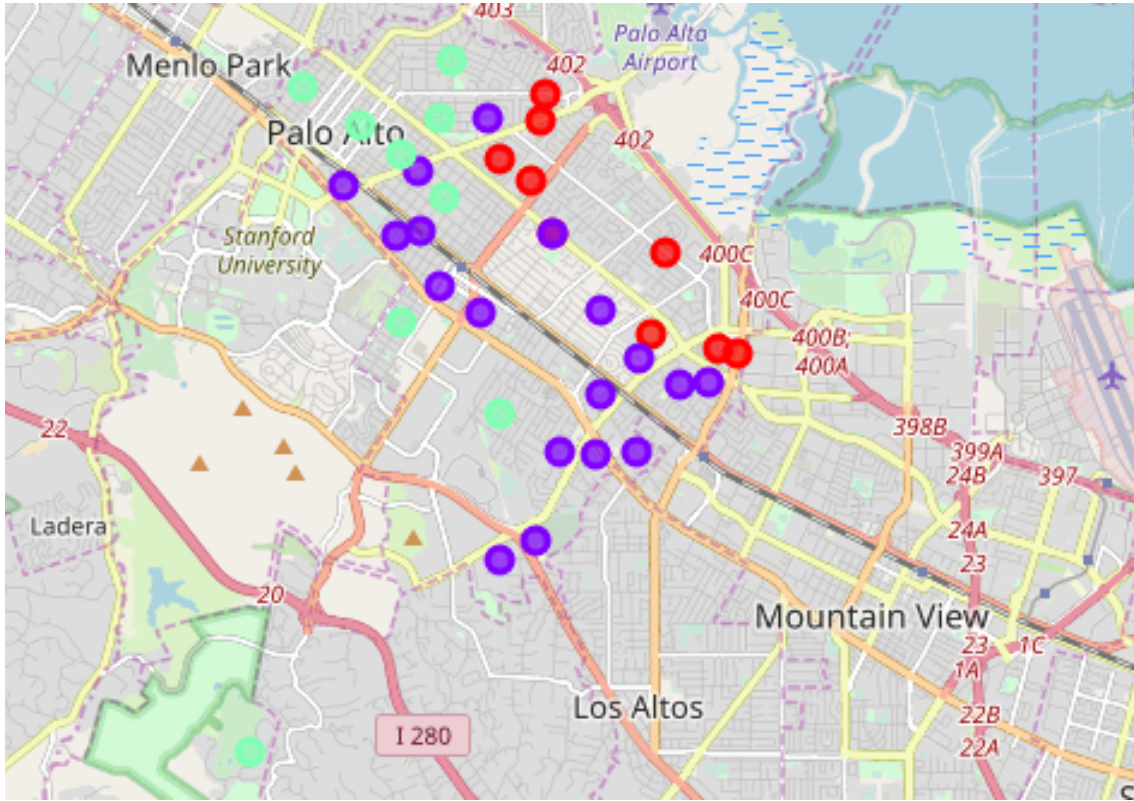
4 Result

Three clusters are used in K-mean clustering algorithm. and the visualized map of each cluster are presented in the Figure 1 below.

Table 1, 2 and 3 present the top 6 most common venues and associated rent for each neighborhoods in each cluster.

As we can see, the first cluster contains neighborhoods with relatively cheap rent in Palo Alto Area and are represented as red dot in the graph. By observing table, we can see these neighborhoods with low rent are either inconvenient neighborhoods (neighborhoods without much place to drink and eat) or neighborhoods that have unpleasant characteristics (People normally don't want to live next to gas station and playground). Neighborhoods in the second cluster are represented as purple dot in the graph. These are neighborhoods with around the median rent in Palo Alto area. We can see from table that neighborhoods in this cluster are convenient and suitable for living. People living in the second cluster have a lot of place to go to:

Figure 1: Palo Alto Map with Neighborhoods Clusters

Source: calculated by author.

Table 1: Palo Alto Neighborhoods Clusters 1

Rent	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
2693	Playground	Pizza Place	Dog Run	Pet Store	Burrito Place	Salon / Barbershop
2693	Dance Studio	Gym	Automotive Shop	Rental Car Location	Grocery Store	Coffee Shop
2693	Gas Station	Bagel Shop	Chinese Restaurant	Gym	Coffee Shop	Tailor Shop
2693	Mexican Restaurant	ATM	Bagel Shop	Gas Station	Grocery Store	Gym / Fitness Center
2693	Coffee Shop	Gas Station	Shopping Plaza	Gym	Chinese Restaurant	Bagel Shop
2693	Baseball Field	Furniture / Home Store	Intersection	Yoga Studio	Frozen Yogurt Shop	Fried Chicken Joint
2693	Pool	Art Gallery	Art Museum	Garden	Park	Yoga Studio
2862	Pharmacy	Gym	Mexican Restaurant	Bank	Yoga Studio	Shoe Repair
2693	Intersection	Pool	Food Truck	Gym	Park	Yoga Studio
2693	Gym	Automotive Shop	Rental Car Location	Jewish Restaurant	Flower Shop	Fast Food Restaurant

Source: calculated by author.

4

Table 2: Palo Alto Neighborhoods Clusters 2

Rent	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
3426	Intersection	Pizza Place	Sandwich Place	Sushi Restaurant	Coffee Shop	Park
3426	Café	Intersection	Plaza	Food Truck	Financial or Legal Service	Frozen Yogurt Shop
3426	French Restaurant	Mexican Restaurant	Italian Restaurant	Bank	Organic Grocery	Japanese Restaurant
3426	Intersection	Indian Restaurant	Dog Run	Playground	Café	Park
3426	Garden	Intersection	Pool	Theater	Dog Run	Financial or Legal Service
3426	Rental Car Location	Hotel	American Restaurant	Park	Pharmacy	Gym / Fitness Center
3426	Dance Studio	Grocery Store	Automotive Shop	Nail Salon	Shopping Plaza	Chinese Restaurant
3426	Dance Studio	Bookstore	Pizza Place	Deli / Bodega	Pool	Coffee Shop
3426	Coffee Shop	American Restaurant	Grocery Store	Spa	Bakery	Burger Joint
3426	Hotel	Gym / Fitness Center	Spa	Park	Chinese Restaurant	Bistro
3426	Yoga Studio	Café	Rest Area	Garden	Indonesian Restaurant	Farmers Market
3426	Hotel	Rental Car Location	American Restaurant	Gym / Fitness Center	Park	Street Food Gathering
3426	Japanese Restaurant	Rest Area	Optical Shop	Automotive Shop	Park	Financial or Legal Service
3426	Pharmacy	Gym	Mexican Restaurant	Bank	Yoga Studio	Shoe Repair
3426	Optical Shop	Recreation Center	Park	Steakhouse	College Soccer Field	Rugby Pitch
3426	Pool	Yoga Studio	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant
3426	Mobile Phone Shop	IT Services	Mexican Restaurant	Deli / Bodega	Rental Car Location	Outdoor Sculpture
3426	Garden	Pool	Art Museum	Tennis Court	Park	Financial or Legal Service

Source: calculated by author.

Table 3: Palo Alto Neighborhoods Clusters 3

Rent	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
3859	Professional & Other Places	Park	Yoga Studio	Financial or Legal Service	Frozen Yogurt Shop	Fried Chicken Joint
3859	Park	Gas Station	Mattress Store	Arts & Crafts Store	Chinese Restaurant	Sushi Restaurant
3820	Park	Theater	Performing Arts Venue	Pool	Tennis Court	Campground
3820	Pizza Place	Park	Yoga Studio	Financial or Legal Service	Frozen Yogurt Shop	Fried Chicken Joint
3820	Park	Yoga Studio	Tapas Restaurant	Pastry Shop	Creperie	Coffee Shop
3820	Garden	Rest Area	Israeli Restaurant	Intersection	Fried Chicken Joint	French Restaurant
4086	Golf Course	Financial or Legal Service	Furniture / Home Store	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant
3820	Yoga Studio	Elementary School	Park	Massage Studio	Garden	Historic Site
3820	Coffee Shop	Electronics Store	French Restaurant	Frozen Yogurt Shop	Greek Restaurant	Shoe Store

Source: calculated by author.

eat, drink and have fun. What's more important, there are not much unpleasant venues in this cluster. The third cluster which are represented by green dot in the graph has the highest rent. Venues like park, garden and gold course are most common venues in this cluster. Therefore, neighborhoods in this cluster have good view, fresh air and less crowed. Neighborhoods in the third cluster are the most suitable for living and working, which is not surprising to have the highest rent cost.

5 Discussion

The result in the above section suggest three things. First, neighborhoods that are inconvenient and have unpleasant surrounding for living have lowest rent. Inconvenience here means lack of place like restaurant coffee shop for people to hang out. Unpleasant surroundings include place with unpleasant smell and noise like gas station and playground. Second, neighborhoods located in the convenient area have around the average rent. These areas are most like downtown and commercial district of the city. There are plenty of restaurant, coffee shop and recreation place for people live here to go to, but these locations can be noisy and less secure, which have negative impact on rent price. Moreover, places for rental in these neighborhoods are most likely apartments, which is another factor for lower rent. Last, neighborhoods near park and professional places have highest rent price. These neighborhoods have best view with some distant away from the busiest part of the city. These neighborhoods are most suitable for living because they can be the most quiet and secure part of the town. Places for rental in this area are most like houses with good amount of living space. Therefore, rentals in these neighborhoods have highest rate.

6 Conclusion

Based on the result of discussion, we can conclude that if an investor wants to make a real estate purchase to collect rent in Palo Alto, the best choice is to make a purchase in the neighborhood with parks and professional places. If the above neighborhoods are out of option, the next optimal choice is to invest in neighborhoods that near commercial center with plenty of venues for people living here to eat, drink and have fun. Investors need to avoid neighborhoods that are distant away from commercial area with unpleasant living features like gas stations and playground