Neighborhood Similarity
(Los Angeles vs San Francisco)

Igor Svystun

December 4, 2020

1. Introduction

1.1 Problem

In this capstone project we will compare neighborhoods of the two cities and determine how similar or dissimilar they are. The given model can be applied for comparison with other cities as well.
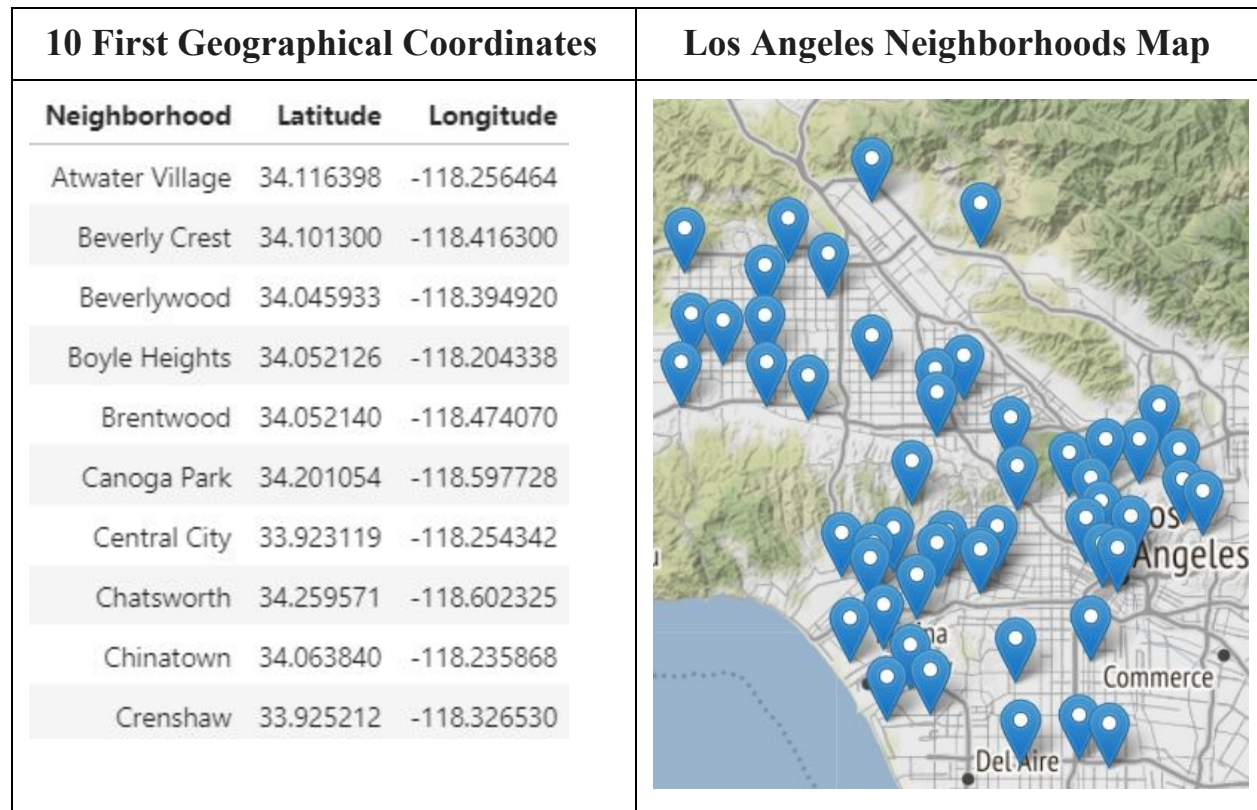
1.2 Target Audience

The target audience of this report would be anyone who wants to move to a similar neighborhood in another city or wants to relocate (open new location of) business (bar, bakery, bookstore etc.) in a similar neighborhood in another city.

2. Data Acquisition and Cleaning

## 2.1.1. Los Angeles Geodata

In this project, there are available geographical coordinates data for 52 neighborhoods in Los Angeles: Atwater Village, Beverly Crest, Beverlywood, Boyle Heights, Brentwood, Canoga Park, Central City, Chatsworth, Chinatown, Crenshaw, Downtown, Eagle Rock, Echo Park, El Sereno, Encino, Financial District, Glassell Park, Granada Hills, Highland Park, Hollywood, Hollywood Hills, Hyde Park, Lakeview Terrace, Los Feliz, Mar Vista, Mid City, Mid-Wilshire, Montecito Heights, North Hills, North Hollywood, Northridge, Palms, Pico-Robertson, Playa Del Rey, Playa Vista, Reseda, Sawtelle, Silver Lake, South Los Angeles, South Park, Studio City, Sylmar, Tarzana, Valley Village, Van Nuys, Venice, West Los Angeles, Westchester, Westlake, Westwood, Winnetka, Woodland Hills.

10 First Geographical Coordinates			Los Angeles Neighborhoods Map
Neighborhood	**Latitude**	**Longitude**	
Atwater Village	34.116398	-118.256464	
Beverly Crest	34.101300	-118.416300	
Beverlywood	34.045933	-118.394920	
Boyle Heights	34.052126	-118.204338	
Brentwood	34.052140	-118.474070	
Canoga Park	34.201054	-118.597728	
Central City	33.923119	-118.254342	
Chatsworth	34.259571	-118.602325	
Chinatown	34.063840	-118.235868	
Crenshaw	33.925212	-118.326530	

Pic. 2.1.1 Los Angeles Neighborhoods

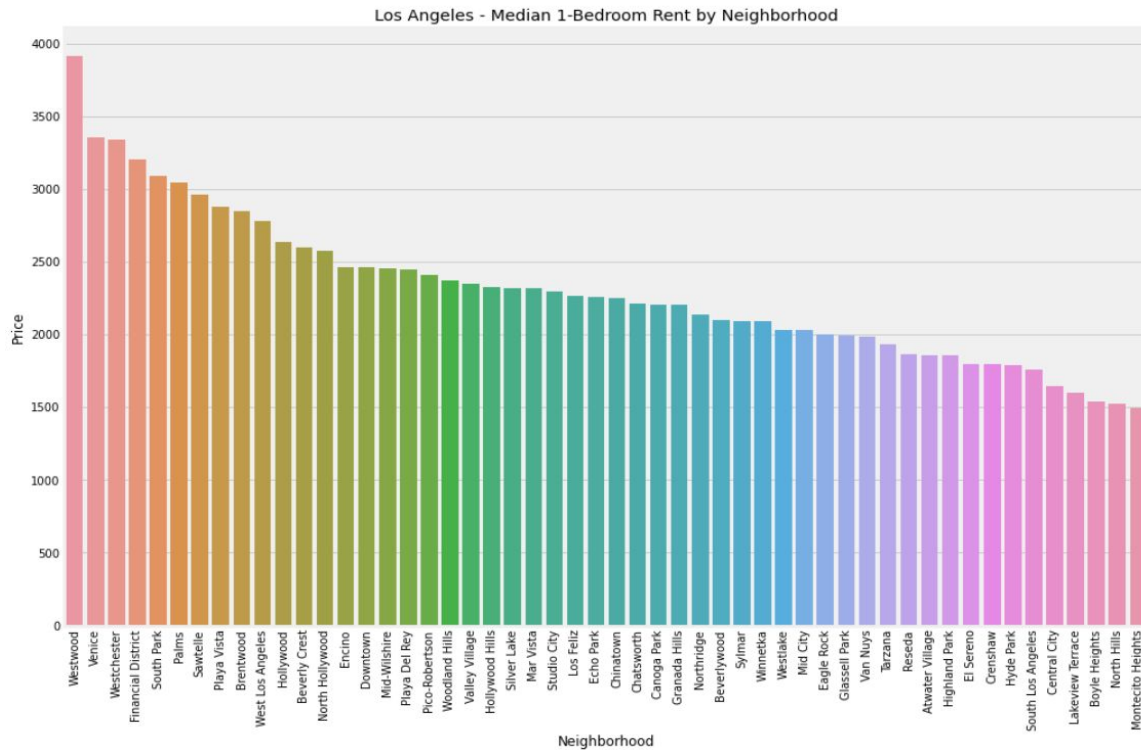## 2.1.2 Los Angeles Crime Data by Neighborhood

Los Angeles crime data was scraped from **xtown.la** .
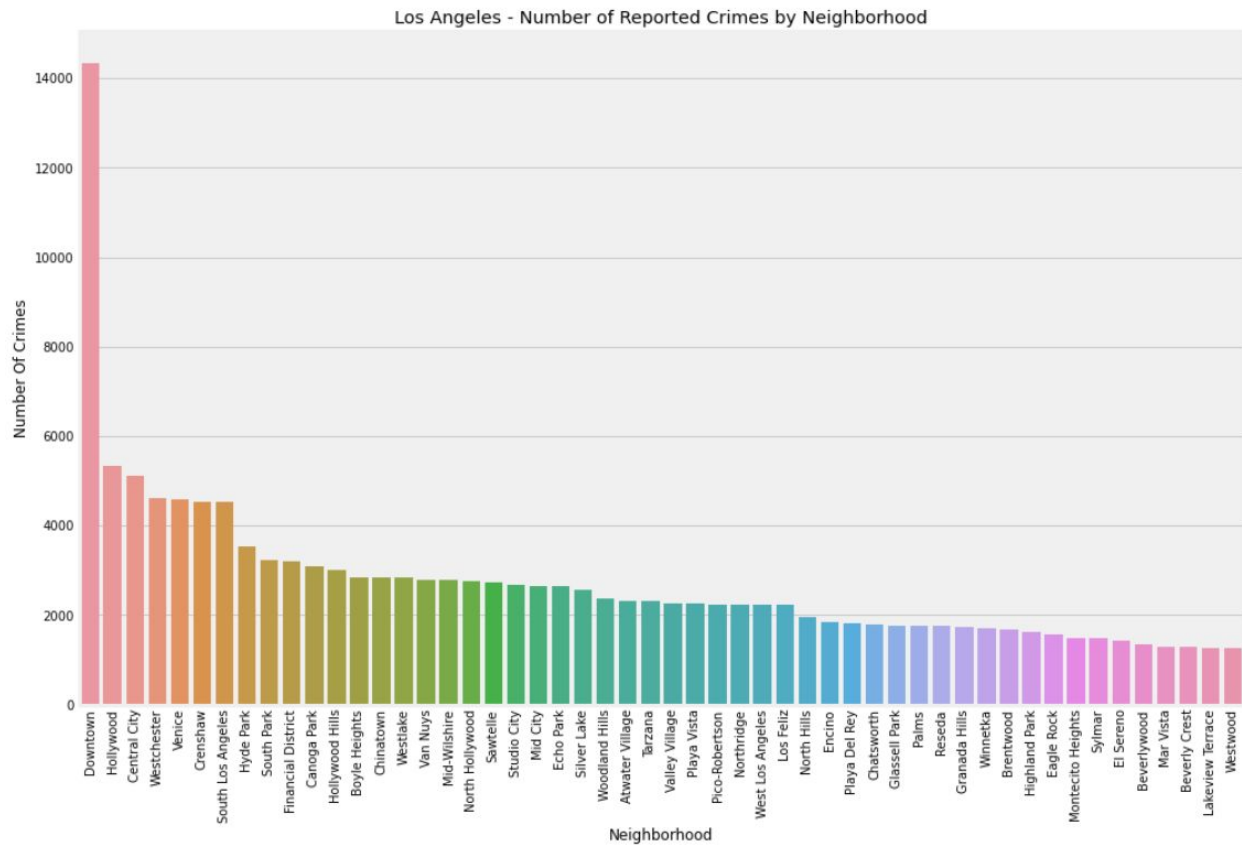
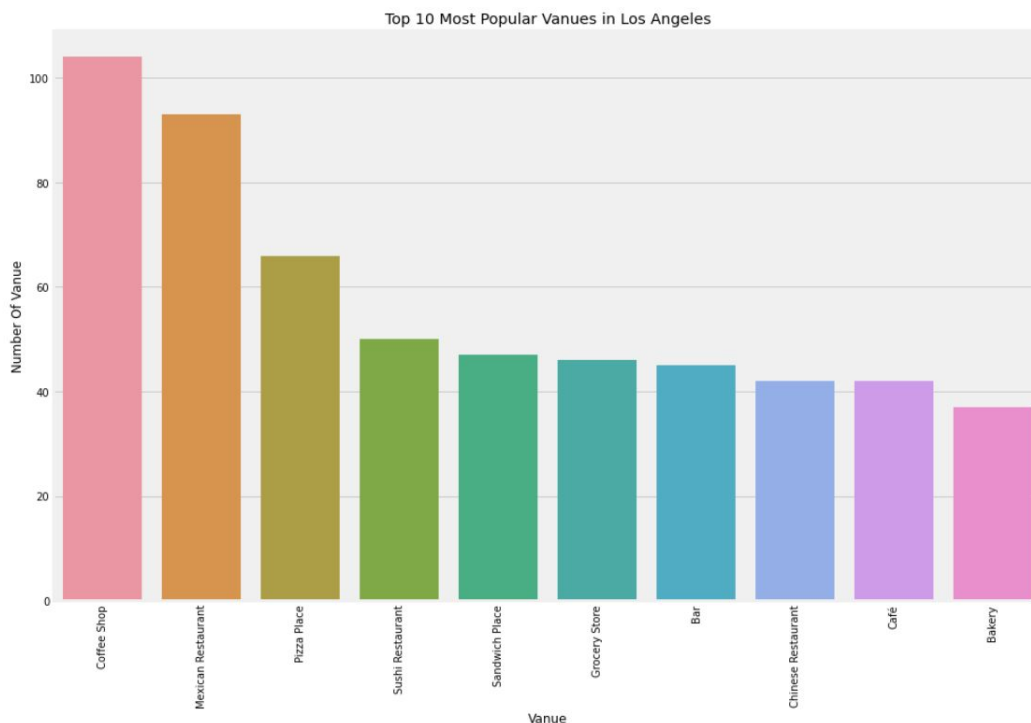**Pic 2.1.2** Los Angeles Crime Data by Neighborhood

2.1.3 Los Angeles Average Rent by Neighborhood

Los Angeles average 1-bedroom rent was scraped from rentjungle.com .

Pic 2.1.3 Los Angeles Average Rent by Neighborhood
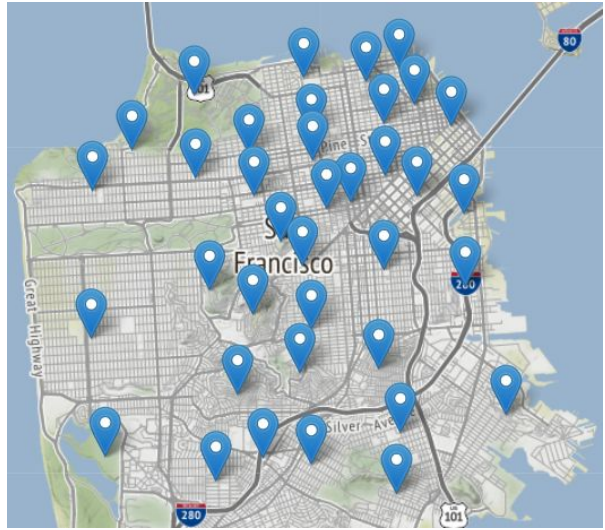
2.1.4 Los Angeles Foursquare Data

Pic 2.1.4 Los Angeles Top 10 Most Popular Venue Categories

2.2.1. San Francisco Geodata

In this project there are available geographical coordinates data for 37 neighborhoods in San Francisco: Bayview Hunters Point, Bernal Heights, Castro/Upper Market, Chinatown, Excelsior, Financial District/South Beach, Glen Park, Haight Ashbury, Hayes Valley, Inner Richmond, Inner Sunset, Japantown, Lakeshore, Lone Mountain/USF, Marina, Mission, Mission Bay, Nob Hill, Noe Valley, North Beach, Oceanview/Merced/Ingleside, Outer Mission, Outer Richmond, Pacific Heights, Portola, Potrero Hill, Presidio, Presidio Heights, Russian Hill, Seacliff, South Of Market, Sunset/Parkside, Tenderloin, Twin Peaks, Visitacion Valley, West of Twin Peaks, Western Addition.

10 First Geographical Coordinates	San Francisco Neighborhoods Map
<table><tr><th>Neighborhood</th><th>Latitude</th><th>Longitude</th></tr><tr><td>Bayview Hunters Point</td><td>37.730769</td><td>-122.382009</td></tr><tr><td>Bernal Heights</td><td>37.740512</td><td>-122.416067</td></tr><tr><td>Castro/Upper Market</td><td>37.762519</td><td>-122.436733</td></tr><tr><td>Chinatown</td><td>37.797113</td><td>-122.406529</td></tr><tr><td>Excelsior</td><td>37.720120</td><td>-122.434370</td></tr><tr><td>Financial District/South Beach</td><td>37.792437</td><td>-122.396564</td></tr><tr><td>Glen Park</td><td>37.739886</td><td>-122.437455</td></tr><tr><td>Haight Ashbury</td><td>37.767617</td><td>-122.442736</td></tr><tr><td>Hayes Valley</td><td>37.774691</td><td>-122.430372</td></tr><tr><td>Inner Richmond</td><td>37.781425</td><td>-122.465499</td></tr></table>	

Pic. 2.2.1 San Francisco Neighborhoods
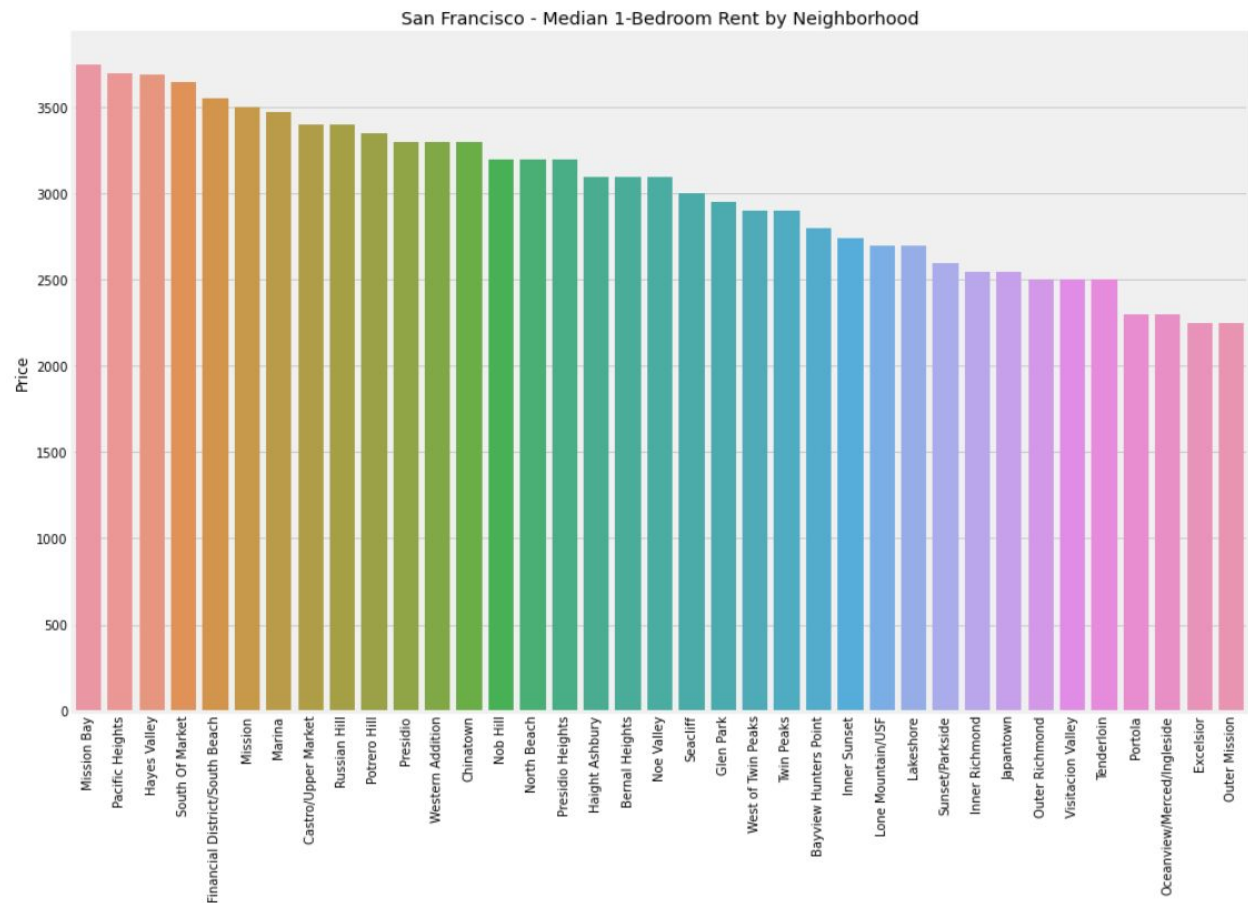
2.2.2 San Francisco Crime Data by Neighborhood

San Francisco crime data was scraped from **data.sfgov.org** .
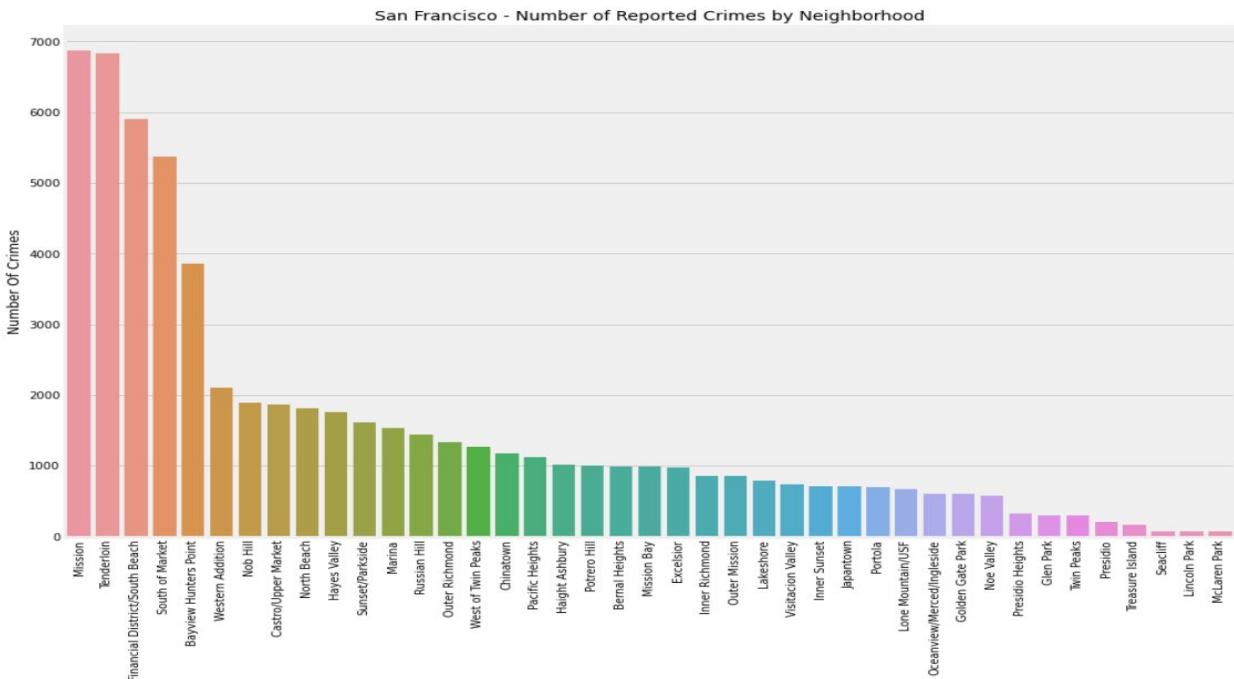
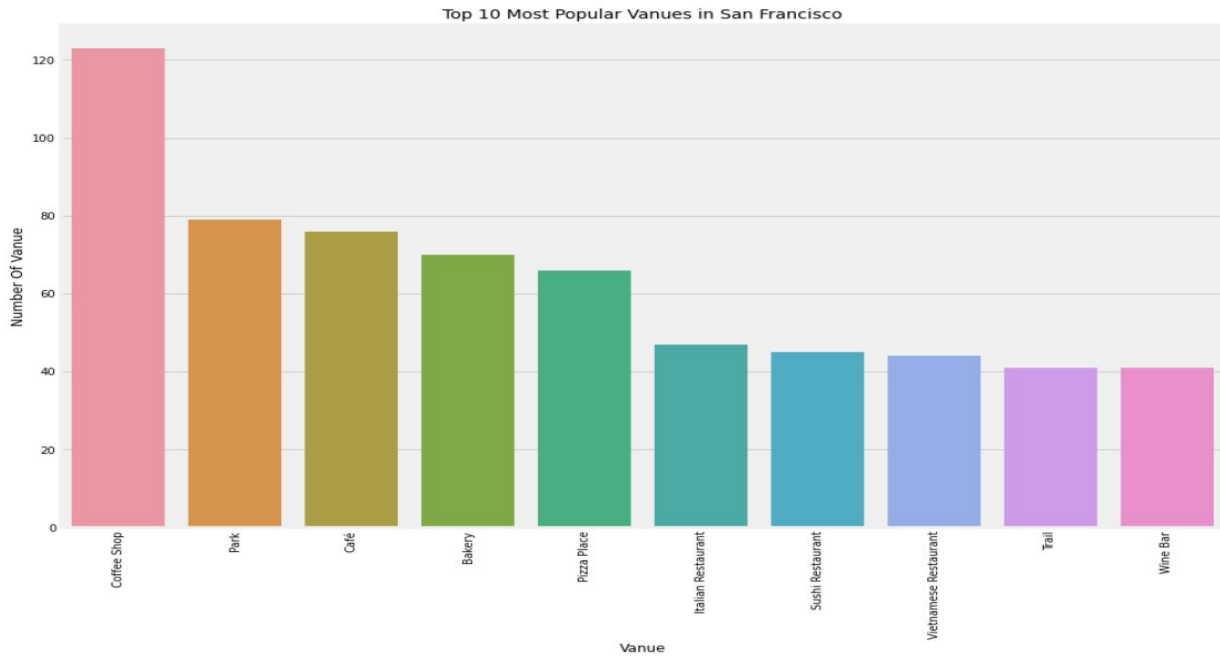**Pic 2.2.2** San Francisco Crime Data by Neighborhood

2.2.3 San Francisco Average Rent by Neighborhood

San Francisco average 1-bedroom rent was scraped from rentjungle.com .

Pic 2.1.3 San Francisco Average Rent by Neighborhood

2.2.4 San Francisco Foursquare Data

Pic 2.2.4 San Francisco Top 10 Most Popular Venue Categories

2.3 Feature Selection

After data fetching from FourSquare, there were some symmetric differences in venue categories fetched for both cities. In our model, we will use venue categories that intersect in both cities. There are 220 such categories: Sandwich Place, Accessories Store, Electronics Store, Women's Store, Greek Restaurant, ATM, Athletics & Sports, Furniture / Home Store, Restaurant, Juice Bar, Breakfast Spot, Toy / Game Store, Mexican Restaurant, Wine Shop, Kitchen Supply Store, Pedestrian Plaza, Salad Place, Chinese Restaurant, Tiki Bar, Pool, Miscellaneous Shop, Construction & Landscaping, Poke Place, Plaza, Baseball Field, Taco Place, French Restaurant, Brazilian Restaurant, Men's Store, Sake Bar, Grocery Store, Middle Eastern Restaurant, Sushi Restaurant, Farmers Market, Metro Station, Salon / Barbershop, Ice Cream Shop, Hotel, General Entertainment, Recreation Center, Speakeasy, Flower Shop, Hawaiian Restaurant, Smoke Shop, Gym, Flea Market, Music Store, Diner, Intersection, Food, Art Gallery, Southern / Soul Food Restaurant, Light Rail Station, Sculpture Garden, Nightclub, Big Box Store, Cupcake Shop, Opera House, Asian Restaurant, Seafood Restaurant, Board Shop, Shoe Store, Performing Arts Venue, Golf Course, Cheese Shop, Music Venue, Bakery, Dance Studio, Gastropub, Filipino Restaurant, Food & Drink Shop, Bar, Burrito Place, Supermarket, Tanning Salon, Sporting Goods Shop, German Restaurant, Wine Bar, Health & Beauty Service, Taiwanese Restaurant, Gourmet Shop, Dim Sum Restaurant, Motel, Art Museum, Trail, Paper / Office

Supplies Store, Cocktail Bar, Butcher, BBQ Joint, Tea Room, Thai Restaurant, Boutique, Clothing Store, Indie Theater, Japanese Restaurant, Dive Bar, Escape Room, American Restaurant, Shipping Store, Movie Theater, Café, Bank, Bookstore, Yoga Studio, Optical Shop, Spiritual Center, Cycle Studio, Historic Site, Latin American Restaurant, Deli / Bodega, Burger Joint, Tennis Court, Convenience Store, Garden Center, Playground, South American Restaurant, Hotel Bar, Pizza Place, Frozen Yogurt Shop, Shopping Mall, Fast Food Restaurant, Coworking Space, Market, Mountain, Indian Restaurant, Cuban Restaurant, Vietnamese Restaurant, Concert Hall, Gift Shop, Basketball Court, Library, Track, Eastern European Restaurant, Comic Shop, Smoothie Shop, Residential Building (Apartment / Condo), Dessert Shop, Mobile Phone Shop, Mediterranean Restaurant, Rock Club, Chocolate Shop, Pharmacy, Alternative Healer, Garden, Candy Store, Pet Store, Noodle House, Roof Deck, Bagel Shop, Cosmetics Shop, Financial or Legal Service, Massage Studio, Business Service, Argentinian Restaurant, Park, Hobby Shop, Supplement Shop, Video Store, Spa, Dumpling Restaurant, Beer Garden, Gym / Fitness Center, Tunnel, Hardware Store, Food Truck, Martial Arts School, Reservoir, Steakhouse, Gay Bar, Jewelry Store, Pub, Szechuan Restaurant, Caribbean Restaurant, Baby Store, Monument / Landmark, Bowling Alley, Lounge, Spanish Restaurant, Sports Bar, Hot Dog Joint, New American Restaurant, Marijuana Dispensary, Creperie, Arts & Crafts Store, Wings Joint, Vegetarian / Vegan Restaurant, Bubble Tea Shop, Bus Station, Nail Salon, Department Store, Ramen Restaurant, Theater, Udon Restaurant, Lake, Gas Station, Coffee Shop, Korean Restaurant, Donut Shop, Discount Store, Organic Grocery, Dog Run, Health Food Store, Liquor Store, Lingerie Store, Beer Store, Bus Stop, Office, Brewery, Arcade, Record Shop, Fried Chicken Joint, Pilates Studio, Beer Bar, Thrift / Vintage Store, Museum, Jazz Club, Tailor Shop, Italian Restaurant, Indie Movie Theater, Comedy Club.

3. Methodology

3.1 Classification Model

K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. Concatenated venues data from both cities were fitted to the K-means model.

3.2 Optimal Number of Clusters

The Elbow method shows an optimal number of clusters should be between 3 and 5.

The Elbow Method showing the optimal k
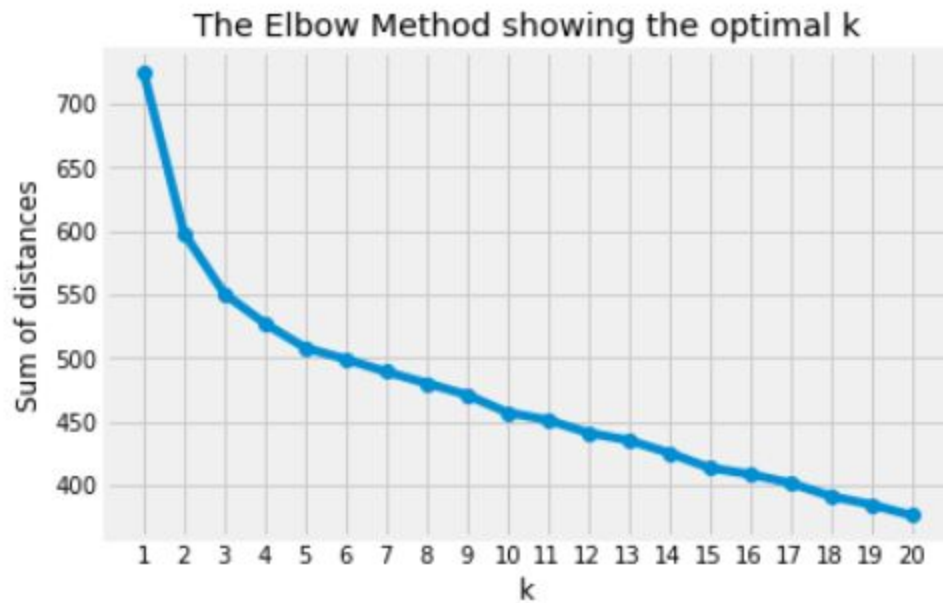
After building a dendrogram , we will choose 5 as the optimal number of clusters.
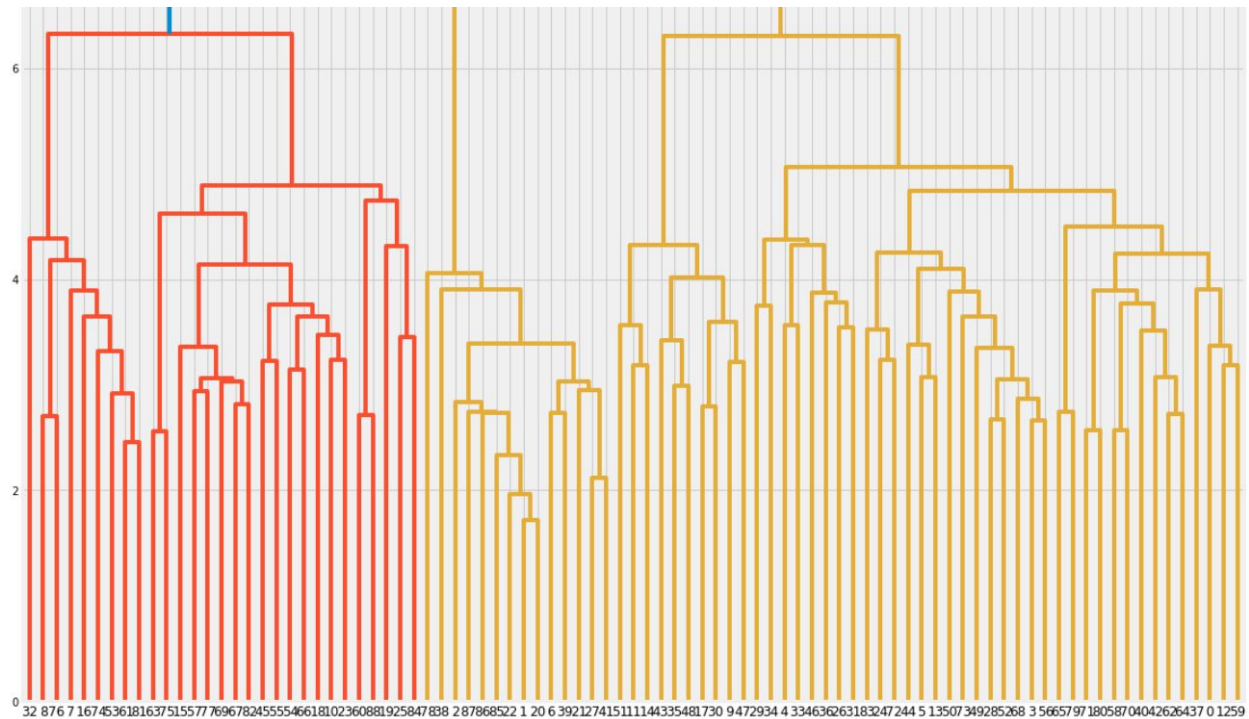
4. Results

Los Angeles has representatives of all clusters and San Francisco has only 4 clusters.

Cluster	Los Angeles Neighborhoods	San Francisco Neighborhoods
1	Beverly Crest, Beverlywood, Central City, Hollywood Hills, Hyde Park, Lakeview Terrace, Montecito Heights, South Los Angeles, South Park, Sylmar	Presidio, Twin Peaks, Visitacion Valley, West of Twin Peaks
2	Chatsworth, Chinatown, Glassell Park, Pico-Robertson	Bernal Heights, Inner Richmond, Japantown, Outer Richmond, Portola, Seacliff, Sunset/Parkside
3	Crenshaw, Eagle Rock, Encino, Granada Hills, North Hollywood, Northridge, Reseda, Valley Village, Van Nuys, Westchester, Westlake, Woodland Hills	
4	Downtown, Financial District, Highland Park, Hollywood, Los Feliz, Mid City, Venice	Castro/Upper Market, Chinatown, Financial District/South Beach, Hayes Valley, Marina, Mission, Nob Hill, Pacific Heights, Potrero Hill, South Of Market, Tenderloin, Western Addition
5	Atwater Village, Boyle Heights, Brentwood, Canoga Park, Echo Park, El Sereno, Mar Vista, Mid-Wilshire, North Hills, Palms, Playa Del Rey, Playa Vista, Sawtelle, Silver Lake, Studio City, Tarzana, West Los Angeles, Westwood, Winnetka	Bayview Hunters Point, Excelsior, Glen Park, Haight Ashbury, Inner Sunset, Lakeshore, Lone Mountain/USF, Mission Bay, Noe Valley, North Beach, Oceanview/Merced/Ingleside, Outer Mission, Presidio Heights, Russian Hill

Table 4 Neighborhoods grouped by clusters

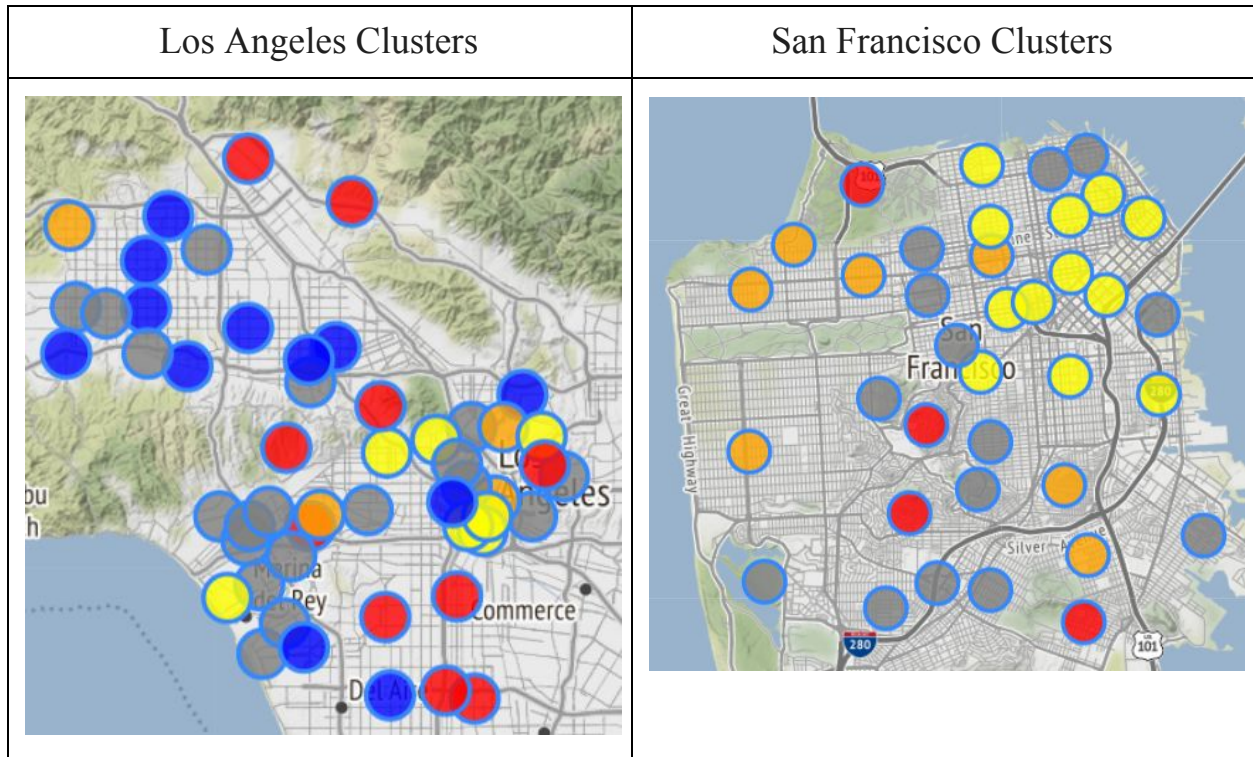|	Los Angeles Clusters	San Francisco Clusters

Fig. 4 Neighborhoods grouped by clusters

4.1 First Cluster

As we can see from Table 4.1 the first cluster is more like general entertainment (parks, playgrounds, trails etc.) with low average crime rate (CR) and moderate average rent rate (RR) . Average CR in Los Angeles is higher within the cluster than in San Francisco. Average RR in Los Angeles is lower within the cluster than in San Francisco. This caused by two neighborhoods (South Los Angeles and Central City) that are even more dissimilar to others neighborhoods based on top 10 venues categories (Pic. 4.1.1).

	1	2	3	4	5	6	7	8	9	10	CR	RR
LA	Park	Grocery Store	Food	Gas Station	Pizza Place	Fast Food Restaurant	Mexican Restaurant	Taco Place	Trail	Light Rail Station	0.49	0.61
SF	Monument / Landmark	Convenience Store	Vietnamese Restaurant	Café	Dumpling Restaurant	Pool	General Entertainment	Reservoir	Tunnel	Spa	0.3	0.77

Table 4.1 Average Top 10 venues, crime rate and rent rate in the first cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Lakeview Terrace	Garden	Food & Drink Shop	Park	Trail	Comedy Club	Filipino Restaurant	Health & Beauty Service	Wine Bar	German Restaurant	Sporting Goods Shop	0.235989	0.476460
Beverly Crest	Playground	Park	Trail	Comedy Club	Gastropub	Health & Beauty Service	Wine Bar	German Restaurant	Sporting Goods Shop	Tanning Salon	0.242924	0.774732
Beverlywood	Food Truck	Burrito Place	Boutique	Historic Site	Park	Women's Store	Filipino Restaurant	Taiwanese Restaurant	Health & Beauty Service	Wine Bar	0.253608	0.625745
Sylmar	Pizza Place	Business Service	Food	Mexican Restaurant	Park	Comedy Club	Food & Drink Shop	Bar	Burrito Place	Supermarket	0.278538	0.624255
Montecito Heights	Food	Lake	Donut Shop	Gas Station	Pizza Place	Park	Food & Drink Shop	Bar	Burrito Place	Supermarket	0.279288	0.444875
Hollywood Hills	Mountain	Trail	Park	Comedy Club	Health & Beauty Service	Wine Bar	German Restaurant	Sporting Goods Shop	Tanning Salon	Supermarket	0.565886	0.693683
South Park	Grocery Store	Taco Place	Shopping Mall	Video Store	Park	Mexican Restaurant	Health & Beauty Service	Wine Bar	German Restaurant	Filipino Restaurant	0.604873	0.920739
Hyde Park	Motel	Bookstore	Performing Arts Venue	Fast Food Restaurant	Taco Place	Grocery Store	Convenience Store	Chinese Restaurant	Food	Pizza Place	0.662231	0.533671
South Los Angeles	Light Rail Station	Intersection	Bus Stop	Fast Food Restaurant	American Restaurant	Shipping Store	Wings Joint	Seafood Restaurant	Tanning Salon	Supermarket	0.847235	0.523838
Central City	Liquor Store	Gas Station	Burger Joint	Breakfast Spot	Mexican Restaurant	Grocery Store	Park	German Restaurant	Wine Bar	Filipino Restaurant	0.955951	0.490465

Pic 4.1.1 The First Cluster Neighborhoods in Los Angeles

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Presidio	Tunnel	Museum	General Entertainment	Garden Center	Brewery	Plaza	Sporting Goods Shop	Gift Shop	Bowling Alley	Movie Theater	0.094787	0.880000
Twin Peaks	Tailor Shop	Monument / Landmark	Reservoir	Trail	Playground	Supermarket	Filipino Restaurant	Food & Drink Shop	Bar	Burrito Place	0.141232	0.773333
Visitacion Valley	Convenience Store	Garden	Pool	Vietnamese Restaurant	Coffee Shop	Park	Trail	Comedy Club	Food & Drink Shop	Wine Bar	0.346445	0.666667
West of Twin Peaks	Spa	Dumpling Restaurant	Monument / Landmark	Breakfast Spot	Sushi Restaurant	Bar	Vietnamese Restaurant	Convenience Store	Grocery Store	Café	0.604265	0.773333

Pic 4.1.2 The First Cluster Neighborhoods in San Francisco

4.2 Second Cluster

As we can see from Table 4.2 the second cluster is more like an International/Asian kitchen with moderate average CR and RR.

	1	2	3	4	5	6	7	8	9	10	CR	RR
LA	Bakery	Sushi Restaurant	Beer Bar	Recreation Center	Filipino Restaurant	BBQ Joint	Café	Chinese Restaurant	Bar	Sandwich Place	0.4	0.66
SF	Bakery	Chinese Restaurant	Tea Room	Korean Restaurant	Bubble Tea Shop	Southern / Soul Food Restaurant	BBQ Joint	Sushi Restaurant	Flower Shop	Vietnamese Restaurant	0.43	0.71

Table 4.2 Average Top 10 venues, crime rate and rent rate in the second cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Glassell Park	Pharmacy	Filipino Restaurant	Café	Bakery	Restaurant	Performing Arts Venue	Sandwich Place	Shipping Store	Bar	Pool	0.332234	0.594756
Chatsworth	Supermarket	Beer Bar	Flower Shop	Rock Club	BBQ Joint	Creperie	Hotel	Sushi Restaurant	Spa	Mexican Restaurant	0.333083	0.660012
Pico-Robertson	Deli / Bodega	Salon / Barbershop	Alternative Healer	Spiritual Center	Arcade	Jewelry Store	French Restaurant	Sushi Restaurant	Recreation Center	Thai Restaurant	0.419669	0.717521
Chinatown	Chinese Restaurant	Jazz Club	Dim Sum Restaurant	Plaza	Bakery	Record Shop	Recreation Center	Brewery	Tea Room	Vietnamese Restaurant	0.534021	0.670739

Pic 4.2.1 The Second Cluster Neighborhoods in Los Angeles

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Seacliff	Southern / Soul Food Restaurant	Flea Market	Sushi Restaurant	Garden Center	Tea Room	Café	Korean Restaurant	Restaurant	Athletics & Sports	BBQ Joint	0.037915	0.800000
Portola	Library	Recreation Center	Bubble Tea Shop	Chinese Restaurant	Brewery	Dim Sum Restaurant	Tea Room	Vietnamese Restaurant	Bakery	BBQ Joint	0.332701	0.613333
Japantown	Udon Restaurant	Tea Room	Bakery	Ice Cream Shop	Jazz Club	Ramen Restaurant	Escape Room	Cosmetics Shop	Creperie	Pub	0.336493	0.680000
Inner Richmond	Bakery	Sushi Restaurant	Wine Shop	Chinese Restaurant	Southern / Soul Food Restaurant	Korean Restaurant	Japanese Restaurant	Bubble Tea Shop	Farmers Market	Dim Sum Restaurant	0.408531	0.680000
Bernal Heights	Butcher	Cocktail Bar	Brewery	Gourmet Shop	Bakery	Southern / Soul Food Restaurant	Flea Market	Record Shop	Indian Restaurant	Latin American Restaurant	0.472038	0.826667
Outer Richmond	Tiki Bar	Sporting Goods Shop	Music Store	Chinese Restaurant	Flower Shop	Café	Record Shop	Bakery	Noodle House	Dim Sum Restaurant	0.634597	0.666667
Sunset/Parkside	Baseball Field	Taiwanese Restaurant	Bar	Market	Korean Restaurant	Chinese Restaurant	Dumpling Restaurant	Bakery	Pharmacy	Italian Restaurant	0.764929	0.693333

Pic 4.2.2 The Second Cluster Neighborhoods in San Francisco

4.3 Fourth Cluster

As we can see from Table 4.3 the fourth cluster is more like a Financial/Business Offices with high average CR and RR.

	1	2	3	4	5	6	7	8	9	10	CR	RR
LA	Bar	Clothing Store	Mediterranean Restaurant	Juice Bar	Gastropub	Ice Cream Shop	Italian Restaurant	Concert Hall	Coffee Shop	Vegetarian / Vegan Restaurant	0.9	0.75
SF	Wine Bar	Cocktail Bar	Bakery	Beer Garden	Gym / Fitness Center	Office	Coffee Shop	Ice Cream Shop	Italian Restaurant	French Restaurant	1.48	0.9

Table 4.3 Average Top 10 venues, crime rate and rent rate in the fourth cluster

Pic 4.3.1 The Fourth Cluster Neighborhoods in Los Angeles

Pic 4.3.1 The Fourth Cluster Neighborhoods in San Francisco

4.4 Fifth Cluster

As we can see from Table 4.4 the fifth cluster is more like a suburban with moderate average CR and high average RR.

	1	2	3	4	5	6	7	8	9	10	CR	RR
LA	Coffee Shop	Gym / Fitness Center	Farmers Market	Sandwich Place	Pharmacy	Italian Restaurant	Bank	Grocery Store	Furniture / Home Store	Burger Joint	0.4	0.71
SF	Coffee Shop	Cosmetics Shop	Sandwich Place	Bakery	Sushi Restaurant	American Restaurant	Furniture / Home Store	Café	Grocery Store	Wine Bar	0.51	0.77

Table 4.4 Average Top 10 venues, crime rate and rent rate in the fifth cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Westwood	Residential Building (Apartment / Condo)	Bank	Bus Station	New American Restaurant	Hotel	Italian Restaurant	Coffee Shop	Food Truck	Fast Food Restaurant	German Restaurant	0.235239	1.166567
Mar Vista	Gourmet Shop	Martial Arts School	Noodle House	Music Store	Japanese Restaurant	Market	Taiwanese Restaurant	Accessories Store	Furniture / Home Store	Indian Restaurant	0.243299	0.689809
El Sereno	Brazilian Restaurant	ATM	South American Restaurant	Thrift / Vintage Store	Restaurant	Bank	Liquor Store	Mexican Restaurant	Grocery Store	0.267666	0.536353
Brentwood	Juice Bar	Track	Shoe Store	Gym / Fitness Center	Italian Restaurant	Farmers Market	Wine Shop	Pharmacy	Athletics & Sports	Massage Studio	0.313027	0.847735
Winnetka	Frozen Yogurt Shop	Department Store	Diner	Bowling Alley	Toy / Game Store	Hobby Shop	Hot Dog Joint	Restaurant	Hotel	Burger Joint	0.319213	0.623063
Palms	Organic Grocery	Hawaiian Restaurant	Sports Bar	Farmers Market	Ramen Restaurant	BBQ Joint	Italian Restaurant	Vegetarian / Vegan Restaurant	Sandwich Place	Grocery Store	0.329709	0.907926
Playa Del Rey	Athletics & Sports	Speakeasy	Optical Shop	Massage Studio	Gym / Fitness Center	Dog Run	Indian Restaurant	Mediterranean Restaurant	Deli / Bodega	Baseball Field	0.342643	0.728546
North Hills	Sculpture Garden	Golf Course	Financial or Legal Service	Dance Studio	Coffee Shop	Asian Restaurant	Seafood Restaurant	Health & Beauty Service	Wine Bar	German Restaurant	0.366448	0.494112
West Los Angeles	Sake Bar	Middle Eastern Restaurant	Beer Store	Indie Movie Theater	Szechuan Restaurant	Sushi Restaurant	Japanese Restaurant	Garden Center	Farmers Market	Gastropub	0.410932	0.828963
Playa Vista	Dog Run	Cycle Studio	Electronics Store	Pub	Nail Salon	Recreation Center	Hardware Store	Hawaiian Restaurant	Farmers Market	Gastropub	0.422868	0.856973
Tarzana	Greek Restaurant	Brazilian Restaurant	Health & Beauty Service	Music Venue	Deli / Bodega	Poke Place	Rock Club	Bank	Italian Restaurant	0.432615	0.574791
Atwater Village	Pilates Studio	Indie Theater	Record Shop	Gym / Fitness Center	Mediterranean Restaurant	Gourmet Shop	Thrift / Vintage Store	Farmers Market	Cuban Restaurant	Restaurant	0.434864	0.553337
Silver Lake	Accessories Store	Board Shop	Record Shop	Furniture / Home Store	Salon / Barbershop	Recreation Center	Ice Cream Shop	Health & Beauty Service	Plaza	Discount Store	0.480787	0.690703
Echo Park	Art Museum	Health Food Store	Beer Store	Comic Shop	Arts & Crafts Store	Record Shop	Vegetarian / Vegan Restaurant	Smoke Shop	Pet Store	Latin American Restaurant	0.497095	0.673421
Studio City	Argentinian Restaurant	Tanning Salon	Ice Cream Shop	Sushi Restaurant	Cycle Studio	Cheese Shop	Arts & Crafts Store	Salad Place	Cupcake Shop	Smoke Shop	0.503280	0.683254
Sawtelle	Movie Theater	Baby Store	Greek Restaurant	Indian Restaurant	Martial Arts School	Ramen Restaurant	Yoga Studio	Market	Athletics & Sports	Supermarket	0.509653	0.882598
Mid-Wilshire	Escape Room	Caribbean Restaurant	Art Gallery	Ramen Restaurant	Sake Bar	Nail Salon	Concert Hall	Spa	Hawaiian Restaurant	Sandwich Place	0.519963	0.732122
Boyle Heights	Fried Chicken Joint	Mexican Restaurant	Coffee Shop	Mediterranean Restaurant	Pharmacy	Sandwich Place	Intersection	Burger Joint	Thai Restaurant	Seafood Restaurant	0.534208	0.457688
Canoga Park	Thrift / Vintage Store	Big Box Store	Pharmacy	Bakery	Flower Shop	Bank	Furniture / Home Store	Pet Store	South American Restaurant	Spa	0.577882	0.657926

Pic 4.4.1 The Fifth Cluster Neighborhoods in Los Angeles

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Glen Park	Cheese Shop	Alternative Healer	Library	Dim Sum Restaurant	Dive Bar	Dog Run	Gift Shop	Grocery Store	Furniture / Home Store	Coffee Shop	0.142180	0.786667
Presidio Heights	Furniture / Home Store	Health Food Store	Health & Beauty Service	Cosmetics Shop	Miscellaneous Shop	Supplement Shop	Pub	Wine Bar	Hotel Bar	Baby Store	0.155450	0.853333
Noe Valley	Gift Shop	Wine Shop	Cheese Shop	Pub	Bakery	Greek Restaurant	Library	Bookstore	Gourmet Shop	Health & Beauty Service	0.276303	0.826667
Oceanview/Merced/Ingleside	Thai Restaurant	Intersection	Liquor Store	Light Rail Station	Japanese Restaurant	Grocery Store	Café	Coffee Shop	Playground	Park	0.286256	0.613333
Lone Mountain/USF	Health & Beauty Service	Women's Store	Supplement Shop	Café	Pub	Wine Bar	Furniture / Home Store	Salon / Barbershop	Cosmetics Shop	Coffee Shop	0.317536	0.720000
Inner Sunset	Salad Place	Vegetarian / Vegan Restaurant	Sandwich Place	Cosmetics Shop	Ice Cream Shop	Middle Eastern Restaurant	Tennis Court	Bus Stop	Sushi Restaurant	Trail	0.337441	0.730667
Lakeshore	Sandwich Shop	Sushi Restaurant	Tennis Court	American Restaurant	Gym	Lake	Pizza Place	Golf Course	Park	Filipino Restaurant	0.373934	0.720000
Outer Mission	Dive Bar	Metro Station	Pedestrian Plaza	Poke Place	BBQ Joint	Bubble Tea Shop	Grocery Store	Dessert Shop	Baseball Field	Bank	0.407583	0.600000
Excelsior	Latin American Restaurant	Restaurant	Pharmacy	Bakery	Basketball Court	Bubble Tea Shop	Gastropub	Chinese Restaurant	Mobile Phone Shop	Deli / Bodega	0.465877	0.600000
Mission Bay	Beer Garden	Tanning Salon	New American Restaurant	Coffee Shop	Southern / Soul Food Restaurant	Pub	Wine Bar	Food Truck	Basketball Court	Deli / Bodega	0.468720	1.000000
Haight Ashbury	Lingerie Store	Thrift / Vintage Store	Gift Shop	Jazz Club	Gastropub	Shoe Store	Ice Cream Shop	Dive Bar	Board Shop	Cheese Shop	0.481991	0.826667
Russian Hill	Comedy Club	Chocolate Shop	Beer Garden	Candy Store	Art Museum	Wine Shop	Ice Cream Shop	Italian Restaurant	Wine Bar	Sushi Restaurant	0.684360	0.906667
North Beach	Comedy Club	Candy Store	Bakery	Argentinian Restaurant	Sake Bar	Italian Restaurant	Diner	Deli / Bodega	Ice Cream Shop	Seafood Restaurant	0.856872	0.853333
Bayview Hunters Point	Asian Restaurant	Coffee Shop	Construction & Landscaping	Gym	Park	Comedy Club	Filipino Restaurant	Health & Beauty Service	Wine Bar	German Restaurant	1.828436	0.746667

Pic 4.4.2 The Fifth Cluster Neighborhoods in San Francisco

4.5 Third Cluster

As we can see from Table 4.5 the third cluster is similar to cluster 5 but with higher average CR and lower average RR. There are no such representatives in San Francisco. Research showed that even by reducing the number of clusters from 5 to 4, we will still have 4 clusters in Los Angeles and 3 in San Francisco. With 3 clusters, the cluster's description will not be so detailed.

	1	2	3	4	5	6	7	8	9	10	CR	RR
LA	Mobile Phone Shop	Discount Store	Pharmacy	Sandwich Place	Supermarket	Bank	Smoke Shop	Music Venue	Pet Store	Chinese Restaurant	0.49	0.67
SF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 4.5 Average Top 10 venues, crime rate and rent rate in the third cluster

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	CrimeRate	RentRate
Eagle Rock	Wine Shop	Board Shop	Bubble Tea Shop	Farmers Market	American Restaurant	Ramen Restaurant	Music Venue	Mobile Phone Shop	Dessert Shop	Pharmacy	0.291284	0.595650
Granada Hills	Smoke Shop	Cupcake Shop	Middle Eastern Restaurant	Discount Store	Pharmacy	Poke Place	Supermarket	Bank	Furniture / Home Store	Salad Place	0.327085	0.655840
Reseda	Arts & Crafts Store	Greek Restaurant	Pharmacy	Noodle House	Thrift / Vintage Store	Vietnamese Restaurant	Discount Store	Music Venue	Supermarket	Gift Shop	0.328772	0.556317
Encino	Yoga Studio	Miscellaneous Shop	Steakhouse	Supplement Shop	Board Shop	Pet Store	ATM	Health & Beauty Service	Japanese Restaurant	Sushi Restaurant	0.348070	0.733015
Northridge	Comic Shop	Smoke Shop	Middle Eastern Restaurant	Hot Dog Joint	Burger Joint	Chinese Restaurant	Dessert Shop	Discount Store	Mobile Phone Shop	BBQ Joint	0.419306	0.637068
Valley Village	Cuban Restaurant	Discount Store	Music Venue	Mobile Phone Shop	Dessert Shop	Eastern European Restaurant	Supermarket	Arts & Crafts Store	Cosmetics Shop	Sushi Restaurant	0.425117	0.699642
Woodland Hills	Supplement Shop	Bagel Shop	Mediterranean Restaurant	Ice Cream Shop	Sushi Restaurant	Smoke Shop	Indian Restaurant	Spa	Health & Beauty Service	Sports Bar	0.444799	0.706794
North Hollywood	Theater	Greek Restaurant	Smoke Shop	Hardware Store	Big Box Store	Sports Bar	Farmers Market	Hawaiian Restaurant	Gastropub	Music Venue	0.519025	0.766389
Van Nuys	Hot Dog Joint	Mobile Phone Shop	Cuban Restaurant	BBQ Joint	Furniture / Home Store	Convenience Store	Chinese Restaurant	Sandwich Place	Latin American Restaurant	Shoe Store	0.520150	0.592074
Westlake	Clothing Store	Hardware Store	Music Venue	Discount Store	Bakery	Pharmacy	Bagel Shop	Metro Station	Supermarket	Cocktail Bar	0.531584	0.606079
Crenshaw	Pet Store	Basketball Court	Smoothie Shop	Cosmetics Shop	Hardware Store	Big Box Store	Mobile Phone Shop	Discount Store	Pharmacy	Department Store	0.848360	0.534267
Westchester	Paper / Office Supplies Store	Record Shop	Furniture / Home Store	Smoothie Shop	Health & Beauty Service	Sandwich Place	Mobile Phone Shop	Discount Store	Pharmacy	Bakery	0.866354	0.994636

Pic 4.5.1 The Third Cluster Neighborhoods in Los Angeles

5. Neighborhood Similarity

In this section, we will measure how neighborhoods are similar within the cluster. We will add 10 points if two neighborhoods have the same venue category and in the same place and 10 - (difference in places) if two neighborhoods have the same venue category but not in the same place based on the top 10 most common venues. Given results can be helpful for those who want to move to similar neighborhoods or define highly competitive neighborhoods in some venues categories.

Example.

	1	2	3	4	5	6	7	8	9	10
N1	Garden	Food & Drink Shop	Park	Trail	Comedy Club	Filipino Restaurant	Health & Beauty Service	Wine Bar	German Restaurant	Sporting Goods Shop
N2	Convenience Store	Garden	Pool	Vietnamese Restaurant	Coffee Shop	Park	Trail	Comedy Club	Food & Drink Shop	Wine Bar

Garden - 9 points, Food & Drink Shop - 3, Park - 7, Trail - 7, Comedy Club - 7, Filipino Restaurant - 0, Health & Beauty Service - 0, Wine Bar - 8, German Restaurant - 0, Sporting Goods Shop - 0. Score = 41.
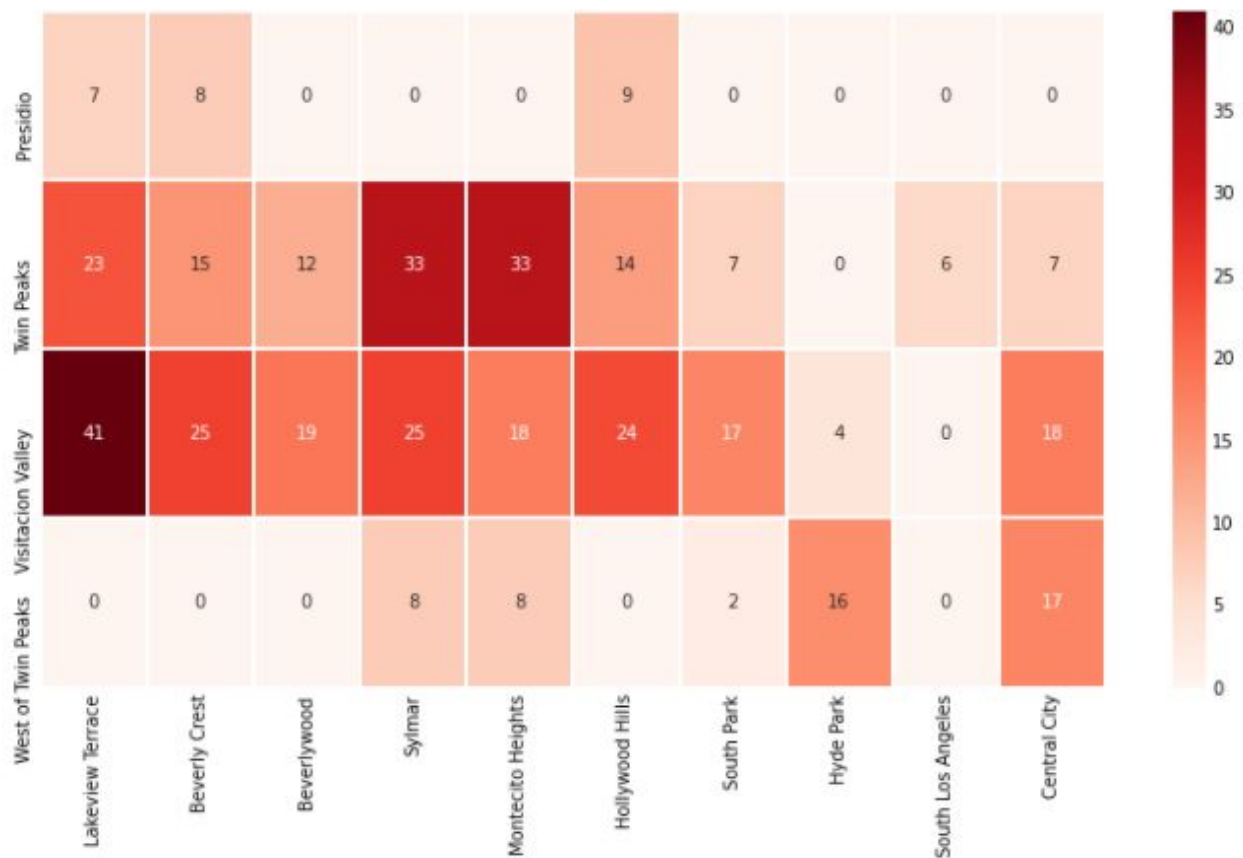
Pic. 5.1 Similarity in the First Cluster
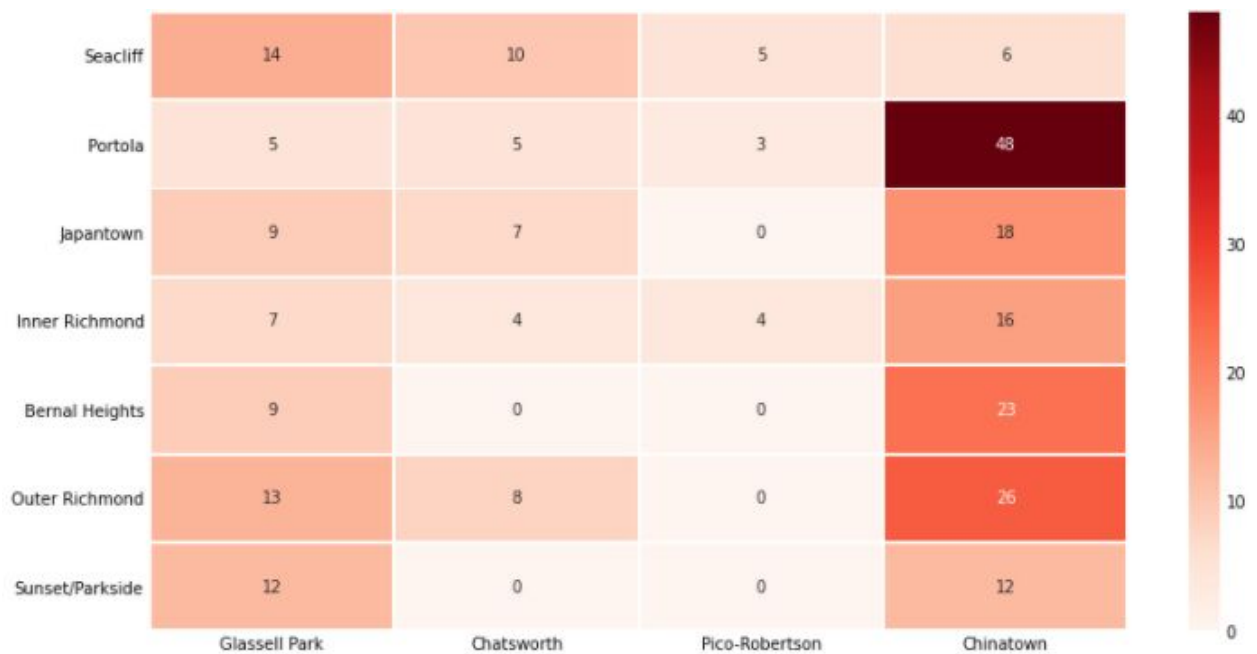
Pic. 5.2 Similarity in the Second Cluster

Pic. 5.3 Similarity in the Fourth Cluster

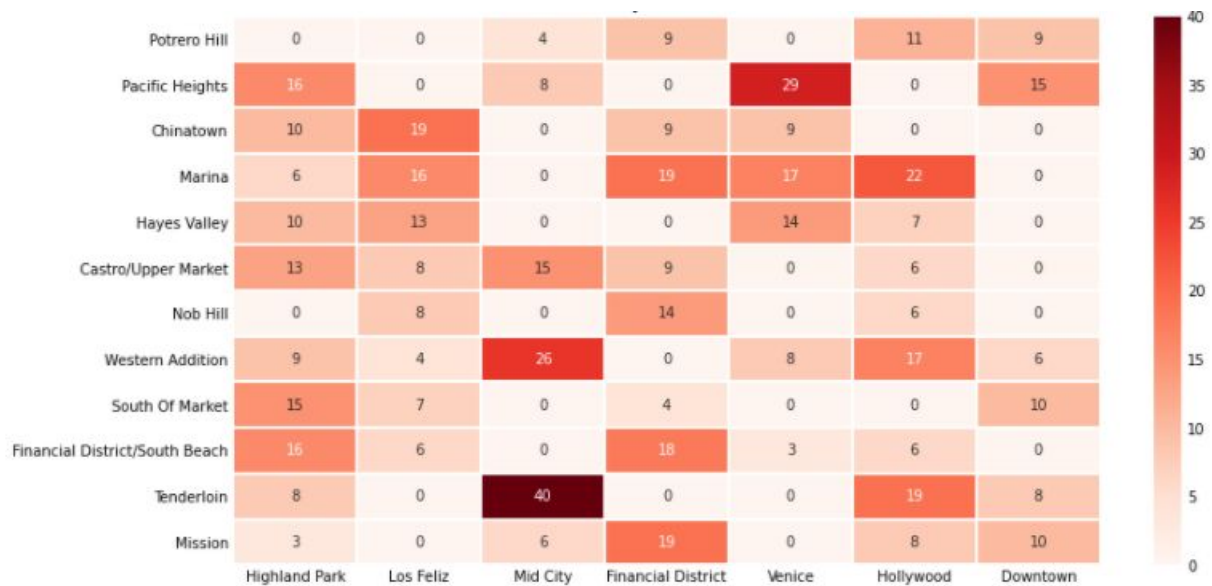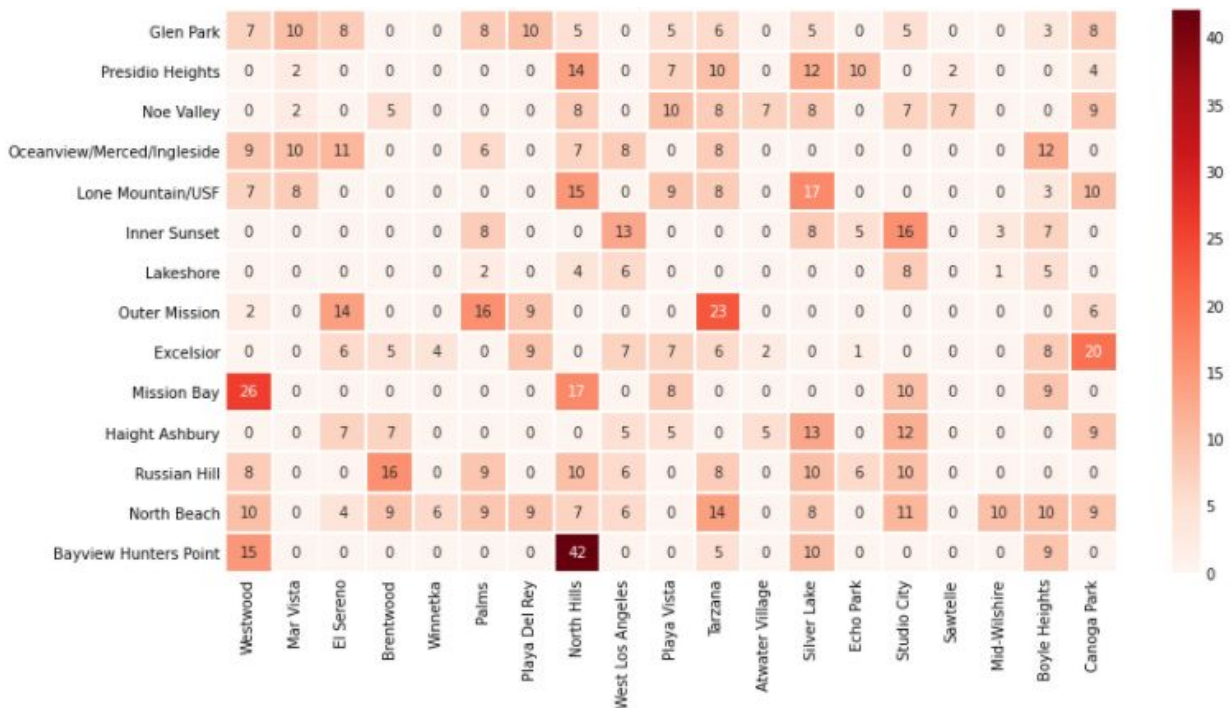|	Highland Park	Los Feliz	Mid City	Financial District	Venice	Hollywood	Downtown
Potrero Hill	0	0	4	9	0	11	9
Pacific Heights	16	0	8	0	29	0	15
Chinatown	10	19	0	9	9	0	0
Marina	6	16	0	19	17	22	0
Hayes Valley	10	13	0	0	14	7	0
Castro/Upper Market	13	8	15	9	0	6	0
Nob Hill	0	8	0	14	0	6	0
Western Addition	9	4	26	0	8	17	6
South Of Market	15	7	0	4	0	0	10
Financial District/South Beach	16	6	0	18	3	6	0
Tenderloin	8	0	40	0	0	19	8
Mission	3	0	6	19	0	8	10

Pic. 5.4 Similarity in the Fifth Cluster

	Westwood	Mar Vista	El Sereno	Brentwood	Winnetka	Palms	Playa Del Rey	North Hills	West Los Angeles	Playa Vista	Tarzana	Atwater Village	Silver Lake	Echo Park	Studio City	Sawtelle	Mid-Wilshire	Boyle Heights	Canoga Park
Glen Park	7	10	8	0	0	8	10	5	0	5	6	0	5	0	5	0	0	3	8
Presidio Heights	0	2	0	0	0	0	0	14	0	7	10	0	12	10	0	2	0	0	4
Noe Valley	0	2	0	5	0	0	0	8	0	10	8	7	8	0	7	7	0	0	9
Oceanview/Merced/Ingleside	9	10	11	0	0	6	0	7	8	0	8	0	0	0	0	0	0	12	0
Lone Mountain/USF	7	8	0	0	0	0	0	15	0	9	8	0	17	0	0	0	0	3	10
Inner Sunset	0	0	0	0	0	8	0	0	13	0	0	0	8	5	16	0	3	7	0
Lakeshore	0	0	0	0	0	2	0	4	6	0	0	0	0	0	8	0	1	5	0
Outer Mission	2	0	14	0	0	16	9	0	0	0	23	0	0	0	0	0	0	0	6
Excelsior	0	0	6	5	4	0	9	0	7	7	6	2	0	1	0	0	0	8	20
Mission Bay	26	0	0	0	0	0	0	17	0	8	0	0	0	0	10	0	0	9	0
Haight Ashbury	0	0	7	7	0	0	0	0	5	5	0	5	13	0	12	0	0	0	9
Russian Hill	8	0	0	16	0	9	0	10	6	0	8	0	10	6	10	0	0	0	0
North Beach	10	0	4	9	6	9	9	7	6	0	14	0	8	0	11	0	10	10	9
Bayview Hunters Point	15	0	0	0	0	0	0	42	0	0	5	0	10	0	0	0	0	9	0

6. Conclusion

In this capstone project, we defined 5 major clusters in Los Angeles area, 4 of which are presented in San Francisco area. Clusters 1 and 4 have a significant amount of very similar neighborhoods in both cities based on the top 10 most common venues. Also as we can see that averages crime and rent rates are fairly

similar within a cluster, which suggests that we can estimate rent price and crime rate by knowing venues located in the given neighborhood.