

---

Photo from Wikipedia

# **Battle of the Neighborhoods- Exploring San Francisco**

Coursera project

By Maricar Donor

1. Introduction to
2. Data
3. Methodology .
4. Results section where you discuss the results.
5. Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.
6. Conclusion section

## **1. Identify the Business Problem (Introduction)**

San Francisco is a populous city in Northern California. Together with San Jose, it combines to become the fifth most populous in the United States.

The state is a popular tourist place to visit attractions like the Golden Gate Bridge and Alamo Square Park. It is a city of culture, commerce and a financial center of the state.

The objective of this project is to use location data and Foursquare to determine a city for business opportunity of a bakery. Using the data we will cluster the different venues of a chosen city to find the best place for a bakery.

## **2. Data**

Location data will be acquired using Geocoder with a combination of CVS files. Once geographic data points are acquired, Foursquare API will be used to get the venues for each city. After obtaining the venues data, we will be able to see which cities have bakeries as the top five venues. Geospatial maps will demonstrate clusters of the cities by venues.

	Name	Type	County	Population 2010	Incorporated	Land area	Incorporated	Unnamed
0	Alameda	City	Alameda	73812	10.61	27.5	April 19, 1854	
1	Albany	City	Alameda	18539	1.79	4.6	September 22, 1908	
2	American Canyon	City	Napa	19454	4.84	12.5	January 1, 1992	
3	Antioch	City	Contra Costa	102372	28.35	73.4	February 6, 1872	
4	Atherton	Town	San Mateo	6914	5.02	13.0	September 12, 1923	
...	...	...	...	...	...	...	...	...
96	Vallejo	City	Solano	115942	30.67	79.4	March 30, 1868	
97	Walnut Creek	City	Contra Costa	64173	19.76	51.2	October 21, 1914	
98	Windsor	Town	Sonoma	26801	7.27	18.8	July 1, 1992	
99	Woodside	Town	San Mateo	5287	11.73	30.4	November 16, 1956	
100	Yountville	Town	Napa	2933	1.53	4.0	February 4, 1965	

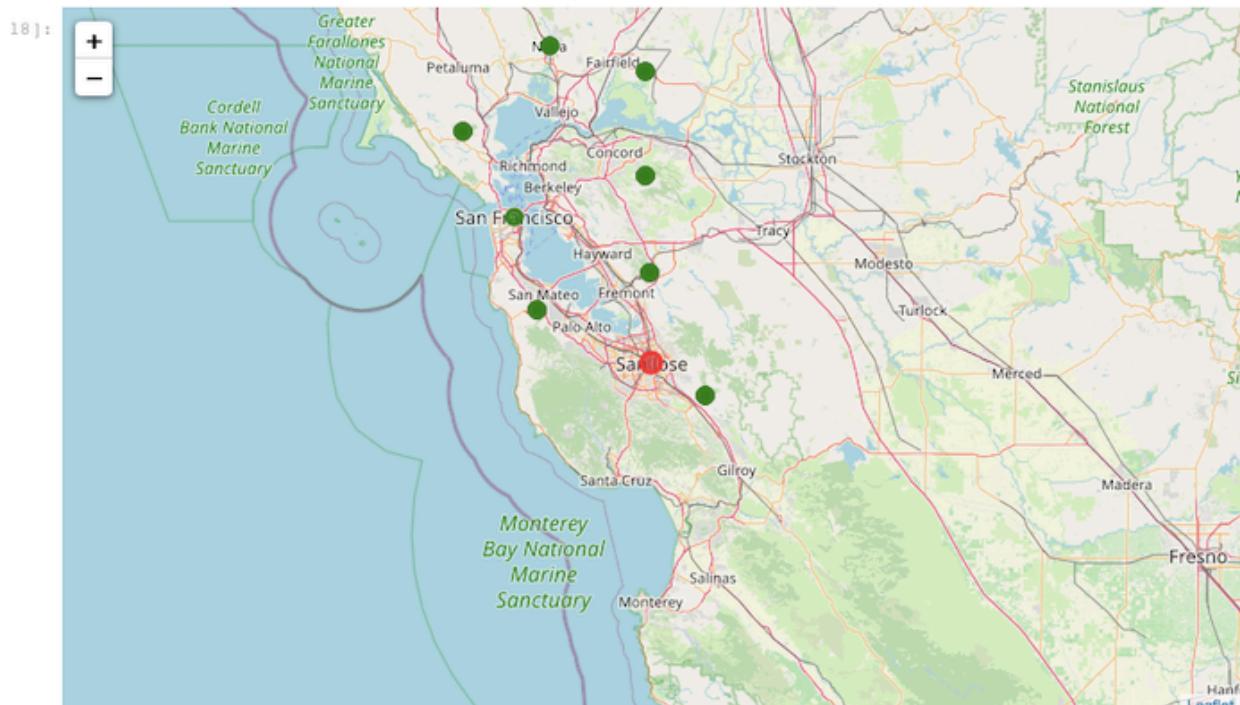
101 rows × 7 columns

( 13 ) :	Name	Type	County	Population 2010
77	San Jose	City	Santa Clara	945942
76	San Francisco	City and county	San Francisco	805235
56	Oakland	City	Alameda	390724
31	Fremont	City	Alameda	214089
84	Santa Rosa	City	Sonoma	167815
...	...	...	...	...
49	Monte Sereno	City	Santa Clara	3341
100	Yountville	Town	Napa	2933
72	Ross	Town	Marin	2415
6	Belvedere	City	Marin	2068
16	Colma	Town	San Mateo	1792

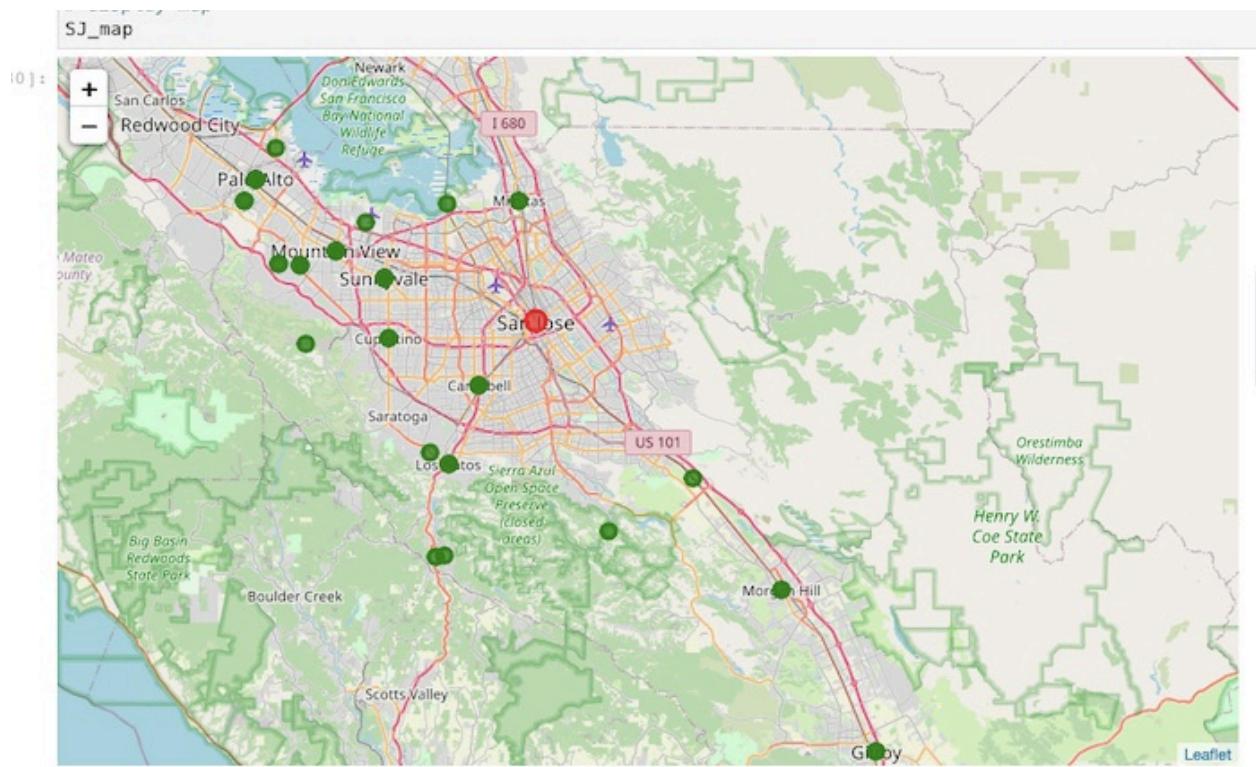
**We will choose the most populous city, San Jose to explore.**

---

## **Map of San Francisco counties. San Jose with red marker and other counties with green marker.**



## Map of San Jose's cities



### 3. Methodology

---

By using Foursquare venues of the cities in San Jose were generated. After the number of venues was obtained, the method of one hot coding was used to get the frequency of the top five for each city.

City	
Alviso	6
Campbell	90
Coyote	3
Cupertino	60
East Palo Alto	11
Gilroy	10
Holy City	4
Los Altos	90
Los Altos Hills	2
Los Gatos	33
Milpitas	60
Moffett Field	7
Monte Sereno	1
Monte Vista	4
Morgan Hill	60
Mountain View	180
New Almaden	4
Palo Alto	240
Redwood Estates	5

Number count of venues per city.

## Frequency of top 5 venues per city

----Alviso----

	venue	freq
0	Music	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Post Office	0.0
4	Pool	0.0

----Campbell----

	venue	freq
0	Ice Cream Shop	0.33
1	Bakery	0.33
2	Bookstore	0.33
3	Airport	0.00
4	Optical Shop	0.00

----Coyote----

	venue	freq
0	Pizza Place	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Post Office	0.0
4	Pool	0.0

----Cupertino----

	venue	freq
0	Japanese Restaurant	0.5
1	Bakery	0.5
2	Airport	0.0
3	Middle Eastern Restaurant	0.0
4	Monument / Landmark	0.0

----East Palo Alto----

	venue	freq
0	Sushi Restaurant	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Pool	0.0
4	Plaza	0.0

----Gilroy----

	venue	freq
0	Deli / Bodega	0.5
1	Farmers Market	0.5
2	Airport	0.0
3	Mexican Restaurant	0.0
4	Pool	0.0

----Holy City----

	venue	freq
0	Coffee Shop	1.0
1	Mediterranean Restaurant	0.0
2	Post Office	0.0
3	Pool	0.0

4 Plaza 0.0

----Los Altos----

	venue	freq
0	Ice Cream Shop	0.33
1	Pizza Place	0.33
2	Bakery	0.33
3	Airport	0.00
4	Optical Shop	0.00

----Los Altos Hills----

	venue	freq
0	Spanish Restaurant	0.5
1	Bookstore	0.5
2	Airport	0.0
3	Mediterranean Restaurant	0.0
4	Pool	0.0

----Los Gatos----

	venue	freq
0	Hotel	0.33
1	Breakfast Spot	0.33
2	Grocery Store	0.33
3	Optical Shop	0.00
4	Middle Eastern Restaurant	0.00

----Milpitas----

	venue	freq
0	Restaurant	0.5
1	Cocktail Bar	0.5
2	Rental Car Location	0.0
3	Post Office	0.0
4	Pool	0.0

----Moffett Field----

	venue	freq
0	American Restaurant	1.0

1	Massage Studio	0.0
2	Pool	0.0
3	Plaza	0.0
4	Pizza Place	0.0

----Monte Sereno----

	venue	freq
0	Gourmet Shop	1.0
1	Airport	0.0
2	Massage Studio	0.0
3	Pool	0.0
4	Plaza	0.0

----Monte Vista----

	venue	freq
0	Spanish Restaurant	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Pool	0.0
4	Plaza	0.0

----Morgan Hill----

	venue	freq
0	Toy / Game Store	0.5
1	Mexican Restaurant	0.5
2	Airport	0.0
3	Massage Studio	0.0
4	Pool	0.0

----Mountain View----

	venue	freq
0	Farmers Market	0.17
1	Coffee Shop	0.17
2	Deli / Bodega	0.17
3	Japanese Restaurant	0.17
4	Bakery	0.17

----New Almaden----

	venue	freq
0	Park	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Post Office	0.0
4	Pool	0.0

----Palo Alto----

	venue	freq
0	History Museum	0.12
1	Park	0.12
2	Mexican Restaurant	0.12
3	Farmers Market	0.12
4	French Restaurant	0.12

----Permanente----

	venue	freq
0	American Restaurant	1.0
1	Massage Studio	0.0
2	Pool	0.0
3	Plaza	0.0
4	Pizza Place	0.0

----Redwood Estates----

	venue	freq
0	Farmers Market	1.0
1	Airport	0.0
2	Mediterranean Restaurant	0.0
3	Pool	0.0
4	Plaza	0.0

----Stanford----

	venue	freq
0	American Restaurant	0.5
1	Greek Restaurant	0.5
2	Massage Studio	0.0
3	Pool	0.0

4

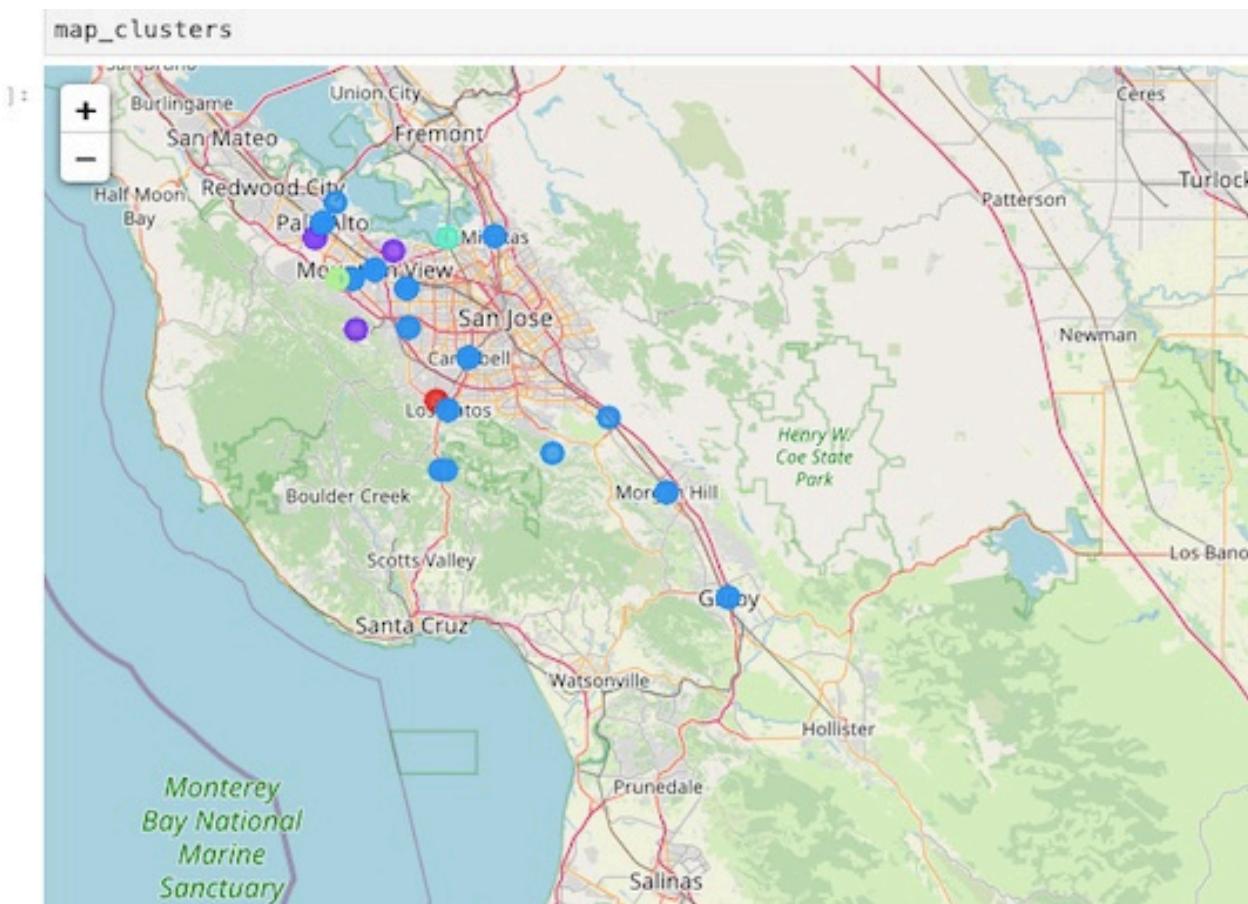
Plaza 0.0

----Sunnyvale----

	venue	freq
0	Hotel	0.17
1	Grocery Store	0.17
2	Gourmet Shop	0.17
3	Toy / Game Store	0.17
4	Cocktail Bar	0.17

## 4. Results

**By using K means the data was used in clustering San Jose cities by venues and a map of San Jose clustering was generated. Clusters displayed the popular venues of cities that are similar.**



## Cluster One -Red(Gourmet Shop/Yoga)

City	location	point	latitude	longitude	altitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Monte Sereno	(Monte Sereno, Santa Clara County, California,...	(37.236333, -121.992458, 0.0)	37.236333	-121.992458	0.0	0	Gourmet Shop	Yoga Studio	Falafel Restaurant	Comedy Club	Concert Hall

## Cluster Two- Purple (American and Greek Restaurants)

City	location	point	latitude	longitude	altitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Moffett Field	(Building 126, Cummins Avenue, Ames Research C...	(37.41129325, -122.05416980310586, 0.0)	37.411293	-122.054170	0.0	1	American Restaurant	Yoga Studio	Farmers Market	Concert Hall	Convenience Store
Stanford	(Stanford, Santa Clara County, California, Uni...	(37.427467, -122.1702445, 0.0)	37.427467	-122.170244	0.0	1	American Restaurant	Greek Restaurant	Yoga Studio	Falafel Restaurant	Concert Hall
Stanford	(Stanford, Santa Clara County, California, Uni...	(37.427467, -122.1702445, 0.0)	37.427467	-122.170244	0.0	1	American Restaurant	Greek Restaurant	Yoga Studio	Falafel Restaurant	Concert Hall
Permanente	(Permanente, Stephen E. Abbotts Trail, Loyola, ...	(37.3189, -122.1111, 0.0)	37.318900	-122.111100	0.0	1	American Restaurant	Yoga Studio	Farmers Market	Concert Hall	Convenience Store

## Cluster Three- Blue (Bakery, Pizza, and Ice Cream)

	City	location	point	latitude	longitude	altitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Los Altos	(Los Altos, Santa Clara County, California, Un...)	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0	2	Bakery	Pizza Place	Ice Cream Shop	Dessert Shop	Dumpling Restaurant
2	Los Altos	(Los Altos, Santa Clara County, California, Un...)	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0	2	Bakery	Pizza Place	Ice Cream Shop	Dessert Shop	Dumpling Restaurant
3	Los Altos	(Los Altos, Santa Clara County, California, Un...)	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0	2	Bakery	Pizza Place	Ice Cream Shop	Dessert Shop	Dumpling Restaurant
3	Mountain View	(Mountain View, Santa Clara County, California...)	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0	2	Coffee Shop	Japanese Restaurant	Breakfast Spot	Deli / Bodega	Bakery
7	Mountain View	(Mountain View, Santa Clara County, California...)	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0	2	Coffee Shop	Japanese Restaurant	Breakfast Spot	Deli / Bodega	Bakery
3	Mountain View	(Mountain View, Santa Clara	(37.3893889, -122.0832101,	37.389389	-122.083210	0.0	2	Coffee	Japanese	Breakfast	Deli /	Bakery

## Cluster Four-Dark Green (Music)

	City	location	point	latitude	longitude	altitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
	Alviso	(Alviso, San Jose, Santa Clara County, Califor...)	(37.426051, -121.9752373, 0.0)	37.426051	-121.975237	0.0	3	Music Venue	Yoga Studio	Falafel Restaurant	Concert Hall	Convenience Store

## Cluster Five-Light Green (Spanish Restaurant, Bookstore)

City	location	point	latitude	longitude	altitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
Los Altos Hills	(Los Altos Hills, Santa Clara County, California...)	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0	4	Spanish Restaurant	Bookstore	Yoga Studio	Falafel Restaurant	Concert Hall	Conv
Los Altos Hills	(Los Altos Hills, Santa Clara County, California...)	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0	4	Spanish Restaurant	Bookstore	Yoga Studio	Falafel Restaurant	Concert Hall	Conv
Monte Vista	(Monte Vista, Alta, Placer County, California,...)	(39.1865648, -120.8327162, 0.0)	39.186565	-120.832716	0.0	4	Spanish Restaurant	Yoga Studio	Cocktail Bar	Comedy Club	Concert Hall	Conv

## 5. Discussion

The cities that venues included bakery as the top five were Cupertino, Campbell, Los Altos, and MountainView.

Frequencies of bakery venues/number of total venues in each city:

1. Cupertino= 0.5/90
2. Campbell=0.33/60
3. Los Altos=0.33/90
4. Mountain View=0.17/180

## **6. Conclusion**

---

CVS files were used along with Python libraries to get location point along with population of cities in San Francisco. Being the most populous city from the data, we explored San Jose. Using the Foursquare API we were able to determine the top 5 venues in each city. By clustering the cities we saw the similarities by venues. With Cupertino being the highest frequency along with the number of venues, it is the best city to open a bakery!