

Capstone Project - The Battle of Neighborhoods

Yasser Bigdeli

San Francisco, Ca

Location recommendation for a new vegetarian restaurant
opening in San Francisco, California

Contents

1	Business Problem and Background	3
2	Data Source and Description.....	3
3	Methodology	4
4	Exploring the neighborhoods.....	5
5	Neighborhood clustering using K-Mean (K=5)	8
6	Results and Discussion	9
7	Conclusion	9
8	Future works.....	10

1 Business Problem and Background

Restaurants are one of the most important business in the world. However, many restaurants will face a bankruptcy challenge during the first few years of opening. Most of the time, the newly opened restaurants are not able to compete with very successful ones and, so that, can't survive.

The main factor of having a restaurant business successful is the location of the business. It is more likely to fail if your business is very close to the old strong competitors. Thus, finding a right location for a business is very important. The aim of this article is to investigate the San Francisco neighborhoods for the density of the available restaurants at each region, and based on that knowledge, this article will recommend a location for opening a new restaurant for a businessman looking to open a new restaurant (vegetarian) in the area. In the San Francisco city, it is even more important to have a right restaurant in a right place to have a successful business, because the cost in the San Francisco is very high for a new business opening like a vegetarian restaurant.

The business plan is to open a new vegetarian restaurant in the City of San Francisco, California and the following are the very important questions to be answered properly before starting the business:

- The locations of vegetarian restaurants in san Francisco.
- Comparing vegetarian places to the total number of restaurants.
- Where is the densest area for restaurants?
- Where is the densest neighborhood for vegetarian restaurants?
- What is the rating of the Vegetarian restaurants?
- What is the percentages of vegetarian population in the area? (if possible)

2 Data Source and Description

San Francisco neighborhood data is obtained via web scraping (<https://localwiki.org/sf/Neighborhoods>). The geographical longitude and latitudes of 51 San Francisco neighborhoods are obtained using 'geopy.geocoders' python library. The location data is obtained from 'foursquare.com' dataset and is based on the geographical coordinates of the neighborhoods. The foursquare API is used for obtaining the venues and restaurants locations. These neighborhoods are shown on a map in the following figure.

Table 1 San Francisco neighborhoods

	Neighborhood	latitude	longitude
0	Alamo Square	37.776360	-122.434689
1	Bayview	40.772627	-124.183950
2	Bernal Heights	37.741001	-122.414214
3	Buena Vista	37.806532	-122.420648
5	Chinatown	37.794301	-122.406376
6	Civic Center	37.779026	-122.419906
7	Cole Valley	37.765813	-122.449962
8	Cow Hollow	37.797262	-122.436248
9	Crocker Amazon	37.709378	-122.438587
11	Dogpatch	37.760698	-122.389202

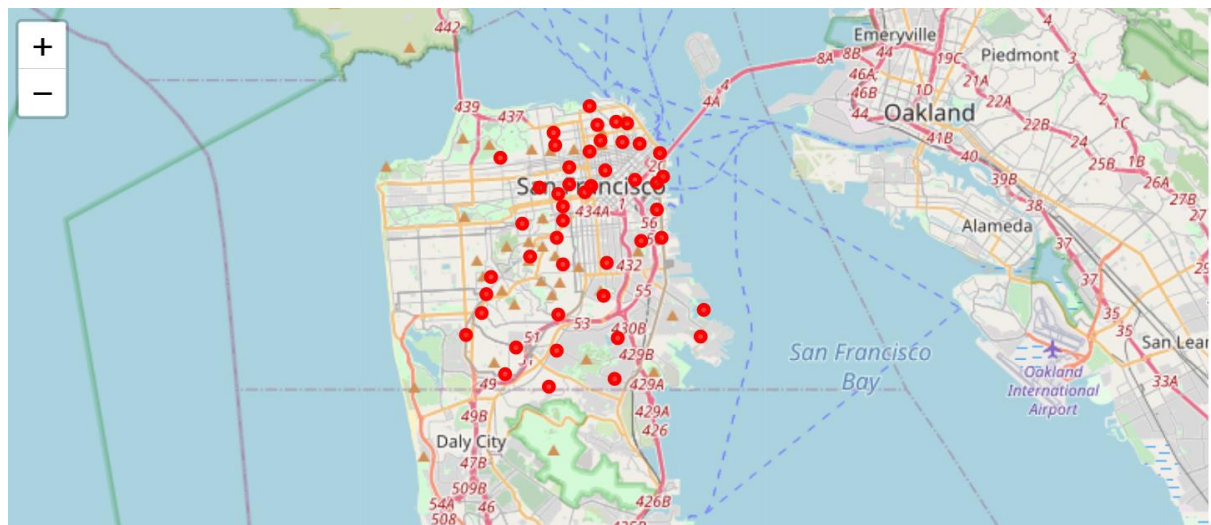


Figure 1 . Map for San Francisco neighborhoods

3 Methodology

The following tasks need to be performed:

1. Find the neighborhoods
2. Obtain the geographical coordinates
3. Get location data from foursquare API
4. Clean the data and create data frame
5. Explore the neighborhoods
6. Identify the number of all types of restaurants locations

7. Identify the number of vegetarian restaurants locations
8. Analyze each neighborhood per restaurants
9. Make a data frame for vegetarian restaurant
10. Group by neighborhood and restaurant types
11. Clustering Neighborhoods using K-means clustering per restaurants
12. Visualization

4 Exploring the neighborhoods

In this section each neighborhood is explored and the venues at the 500m range are identified.

Table 2 Data set for all venues in SF neighbors

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Alamo Square	37.77636	-122.434689	Alamo Square	37.776045	-122.434363	Park
1	Alamo Square	37.77636	-122.434689	Alamo Square Dog Park	37.775878	-122.435740	Dog Run
2	Alamo Square	37.77636	-122.434689	Painted Ladies	37.776120	-122.433389	Historic Site
3	Alamo Square	37.77636	-122.434689	The Independent	37.775573	-122.437835	Rock Club
4	Alamo Square	37.77636	-122.434689	The Mill	37.776425	-122.437970	Bakery
5	Alamo Square	37.77636	-122.434689	Bar Crudo	37.775707	-122.438019	Seafood Restaurant
6	Alamo Square	37.77636	-122.434689	Fool's Errand	37.775512	-122.437961	Bar
7	Alamo Square	37.77636	-122.434689	Nopa	37.774888	-122.437532	New American Restaurant
8	Alamo Square	37.77636	-122.434689	Rare Device	37.775052	-122.437762	Gift Shop

The resultants locations are extracted from the venue data set and the bar graph is developed showing the number of total restaurants at each neighborhood.

Table 3 All type of Restaurants in SF neighborhoods

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
5	Alamo Square	37.77636	-122.434689	Bar Crudo	37.775707	-122.438019	Seafood Restaurant
7	Alamo Square	37.77636	-122.434689	Nopa	37.774888	-122.437532	New American Restaurant
19	Alamo Square	37.77636	-122.434689	jū-ni	37.776743	-122.438770	Sushi Restaurant
21	Alamo Square	37.77636	-122.434689	Tsunami	37.776869	-122.438486	Sushi Restaurant
27	Alamo Square	37.77636	-122.434689	Brenda's Meat & Three	37.778265	-122.438584	Southern / Soul Food Restaurant
30	Alamo Square	37.77636	-122.434689	Zaytoon	37.775185	-122.437896	Mediterranean Restaurant
32	Alamo Square	37.77636	-122.434689	Kung Food	37.777778	-122.438698	Hunan Restaurant
34	Alamo Square	37.77636	-122.434689	Namu Stonepot	37.774763	-122.437780	Korean Restaurant
39	Alamo Square	37.77636	-122.434689	Che Fico	37.777435	-122.438149	Italian Restaurant
41	Alamo Square	37.77636	-122.434689	Saffron Grill	37.776848	-122.437816	Indian Restaurant

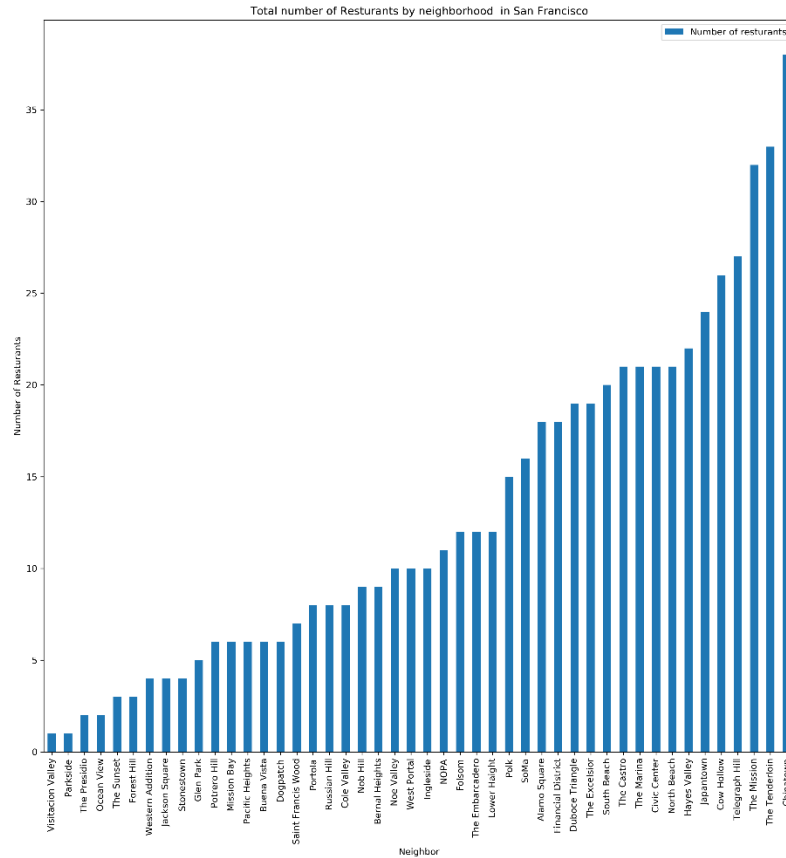


Figure 2 Number of restaurants in each Neighborhood

The vegetarian / Vegan restaurants were extracted out the dataset. The location of them are shown on san Francisco map with green points if Figure 4.

Table 4 Vegetarian restaurants in Sf neighborhoods

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
208	Chinatown	37.794301	-122.406376	Enjoy Vegetarian Restaurant	37.795833	-122.405093	Vegetarian / Vegan Restaurant
258	Chinatown	37.794301	-122.406376	Lucky Creation Restaurant	37.795056	-122.407271	Vegetarian / Vegan Restaurant
299	Civic Center	37.779026	-122.419906	Ananda Fuara	37.777693	-122.416353	Vegetarian / Vegan Restaurant
327	Civic Center	37.779026	-122.419906	Golden Era Vegan	37.781495	-122.416822	Vegetarian / Vegan Restaurant
354	Civic Center	37.779026	-122.419906	Thai Idea Vegetarian Restaurant	37.783395	-122.419141	Vegetarian / Vegan Restaurant
455	Cow Hollow	37.797262	-122.436248	Wildseed	37.797626	-122.432440	Vegetarian / Vegan Restaurant
487	Cow Hollow	37.797262	-122.436248	Vegan Picnic	37.797490	-122.431748	Vegetarian / Vegan Restaurant
1219	Nob Hill	37.794479	-122.415592	Nourish Cafe	37.790529	-122.417296	Vegetarian / Vegan Restaurant
2264	The Marina	37.779026	-122.419906	Ananda Fuara	37.777693	-122.416353	Vegetarian / Vegan Restaurant
2292	The Marina	37.779026	-122.419906	Golden Era Vegan	37.781495	-122.416822	Vegetarian / Vegan Restaurant

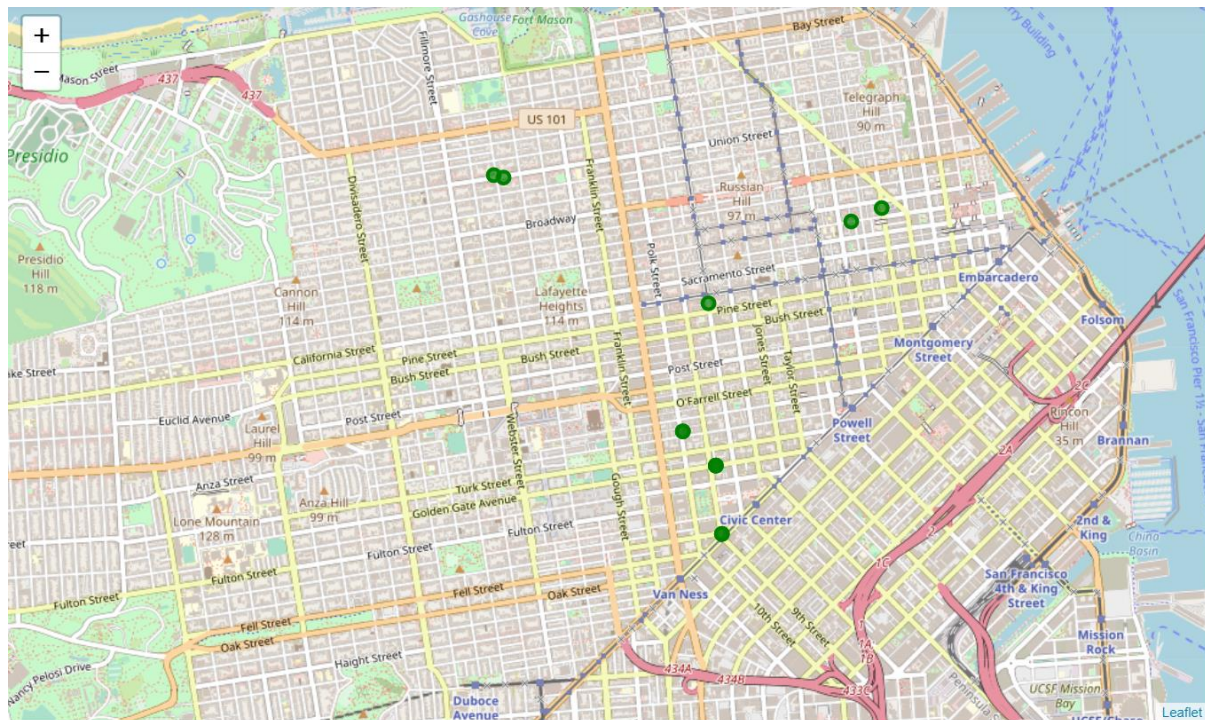


Figure 3 vegetarian restaurant locations

The neighborhood where at least one vegetarian/vegan restaurant is located in are compared to the total number of restaurants located at that specific neighborhood.

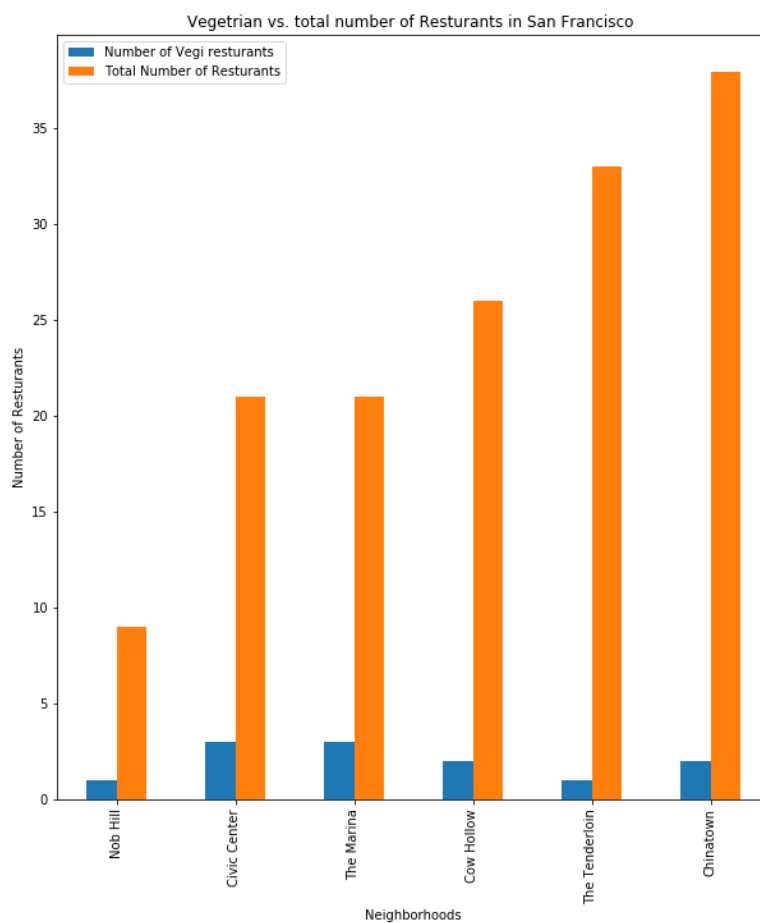


Figure 4 Number of Vegetarian Restaurants in SF Neighborhoods

5 Neighborhood clustering using K-Mean (K=5)

The 10 most common restaurants were ranged for each neighborhood as listed in the following table. The data set is used for clustering neighborhood using K-mean clustering into 5 categories. The class of clustering is also added to the data frame. The clustering results are shown on San Francisco map as shown in figure 5.

Table 10 Most common restaurants and identified clustering per neighborhood data frame

Neighborhood	latitude	longitude	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	7th Most Common Venue	8th Most Common Venue
Alamo Square	37.776360	-122.434689	1.0	Indian Restaurant	Seafood Restaurant	Ethiopian Restaurant	Sushi Restaurant	Hunan Restaurant	Mediterranean Restaurant	Mexican Restaurant	New American Restaurant
Bernal Heights	37.741001	-122.414214	2.0	Italian Restaurant	American Restaurant	Asian Restaurant	Mexican Restaurant	Indian Restaurant	Caribbean Restaurant	New American Restaurant	Peruvian Restaurant
Buena Vista	37.806532	-122.420648	2.0	Seafood Restaurant	Brazilian Restaurant	Fast Food Restaurant	Cantonese Restaurant	Korean Restaurant	Austrian Restaurant	Filipino Restaurant	Japanese Restaurant
Chinatown	37.794301	-122.406376	1.0	Chinese Restaurant	New American Restaurant	Italian Restaurant	Vietnamese Restaurant	Restaurant	Dim Sum Restaurant	Sushi Restaurant	Szechuan Restaurant
Civic Center	37.779026	-122.419906	1.0	Vegetarian / Vegan Restaurant	French Restaurant	Sushi Restaurant	Restaurant	Vietnamese Restaurant	Southern / Soul Food Restaurant	Mediterranean Restaurant	Mexican Restaurant
Cole Valley	37.765813	-122.449962	2.0	Vietnamese Restaurant	Thai Restaurant	Caribbean Restaurant	Mexican Restaurant	Tapas Restaurant	Middle Eastern Restaurant	Mediterranean Restaurant	Indian Restaurant
Cow Hollow	37.797262	-122.436248	1.0	Italian Restaurant	Mexican Restaurant	French Restaurant	American Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	Sushi Restaurant	Caribbean Restaurant
Dogpatch	37.760698	-122.389202	1.0	Restaurant	Latin American Restaurant	Italian Restaurant	Southern / Soul Food Restaurant	Sushi Restaurant	Vietnamese Restaurant	Greek Restaurant	Filipino Restaurant
Duboce Triangle	37.767138	-122.432230	2.0	New American Restaurant	Seafood Restaurant	Mexican Restaurant	Sushi Restaurant	Vietnamese Restaurant	Indian Restaurant	Mediterranean Restaurant	Middle Eastern Restaurant
Financial District	37.793647	-122.398938	2.0	Restaurant	Japanese Restaurant	Mediterranean Restaurant	New American Restaurant	Dim Sum Restaurant	Latin American Restaurant	Mexican Restaurant	Middle Eastern Restaurant

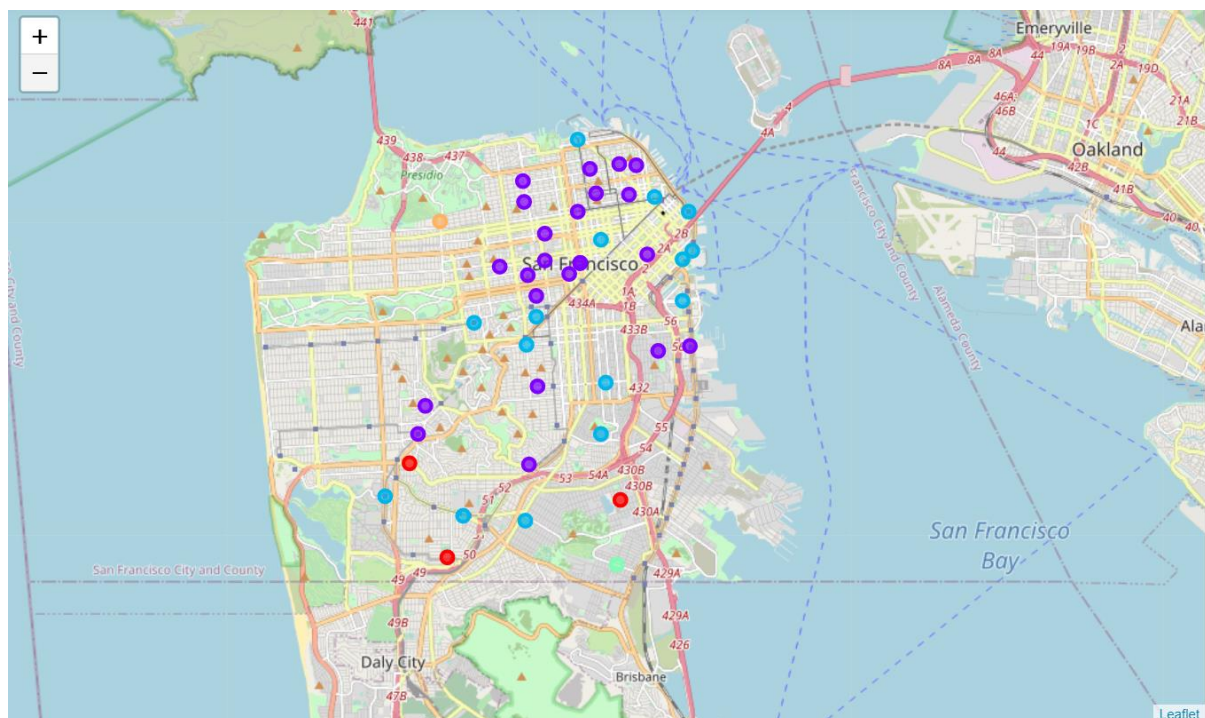


Figure 5 Class of clusters location on map (similar colors represent similar clusters)

To see more details on clustering please see the Appendix A.

6 Results and Discussion

From analyzing data, it is noticed that only 7 neighborhoods have vegetarian restaurants available. Therefore, opening a new vegetarian restaurant, in general, would be beneficial. It is argued that, the best neighborhood to open vegetarian place would be the neighborhood that has many restaurants types except vegetarian. In this case we recommend “The Mission” neighborhood. There are more factors to consider to get better recommendations, however, it is out of scope of this article.

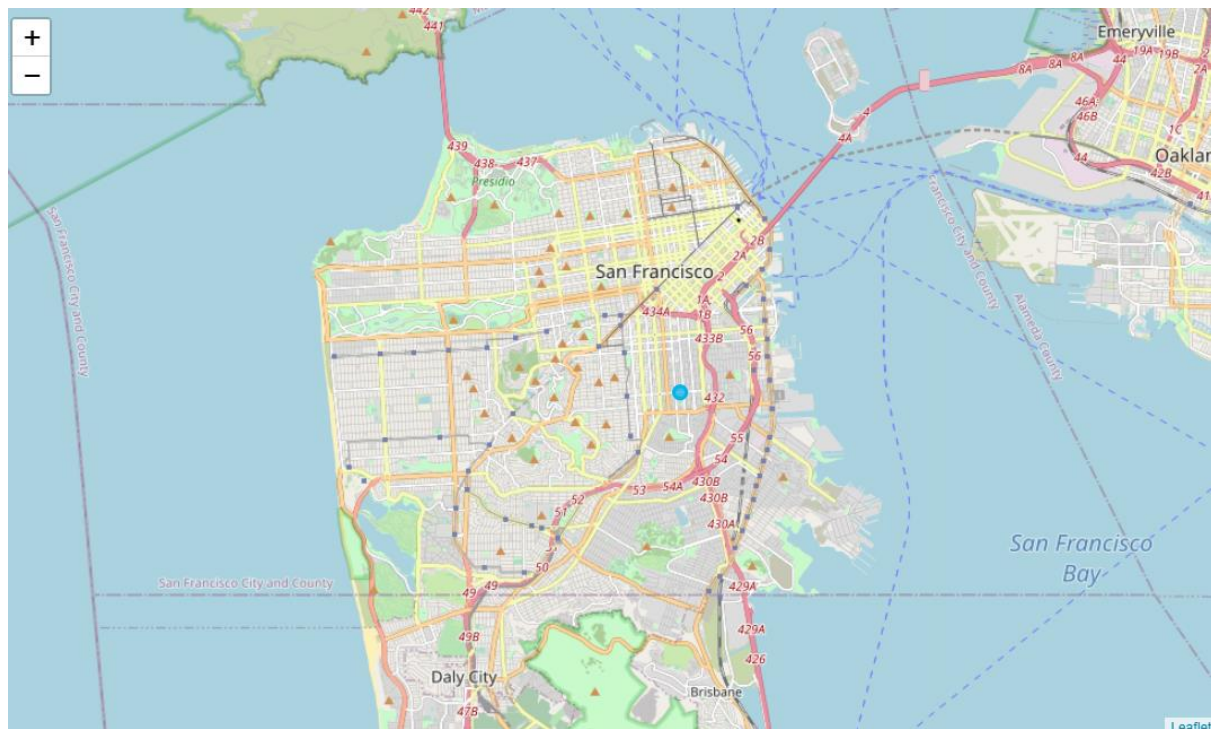


Figure 6 The recommended location (The Mission) for opening a Vegetarian restaurant

7 Conclusion

For a businessman, the most important factor is choosing a right place to start the business. This is more crucial for restaurant business. This article tries to recommend a good location at the city of San Francisco, California for a vegetarian eating place.

After analyzing all the neighborhoods and exploring the available restaurant and, in specific, vegetarian / vegan eating places, the ‘The Mission’ neighborhood is recommended for opening new restaurant for vegetarians. Based on location data analyses, this neighborhood is the dense area of eating places, where, however, there is no single vegetarian restaurant in this region. Therefore, makes it a perfect place for opening a new restaurant for vegetarians. The neighborhoods are also classified based on the frequency of restaurants available in the area using K-Means clustering methodology. This could be used for further analyses which is out of this articles scope.

8 Future works

This project could be well expanded for many other types of business and sectors. There are a few tides that can be taken into exploration:

1. Where to open a fast-food restaurant
2. Where to open coffee shop
3. Where a police station might be needed to open
4. To spot the locations with higher crime rates

Appendix A:

To have access to the Jupyter notebook version (.ipynb format) of this article please see the following link:

https://github.com/yasser64b/Coursera_Capstone/blob/master/Coursera%20Capstone%20Project_Final.ipynb