

Best places to open Yoga Studio in San Francisco

By Igor Turchenko

Business Problem

We want to open new Yoga Studio in San Francisco, USA. Relevant place should match next criteria:

- To be at least 500 meters away from any existing Yoga Studio
- To be inside given area on a map
- To be near venues which is popular after visiting Yoga

This project would be interested to any stakeholder who would be interested in opening new Yoga Studio in San Francisco. Moreover, given solution could be applied to other input parameters, like another city or another venue category.

Data

In this project I would use open Foursquare API data. Particularly, 2 methods:

- Search, for finding venues of given category near specified point
- NextVenues, for getting insights of which venues are popular to visit after given venue

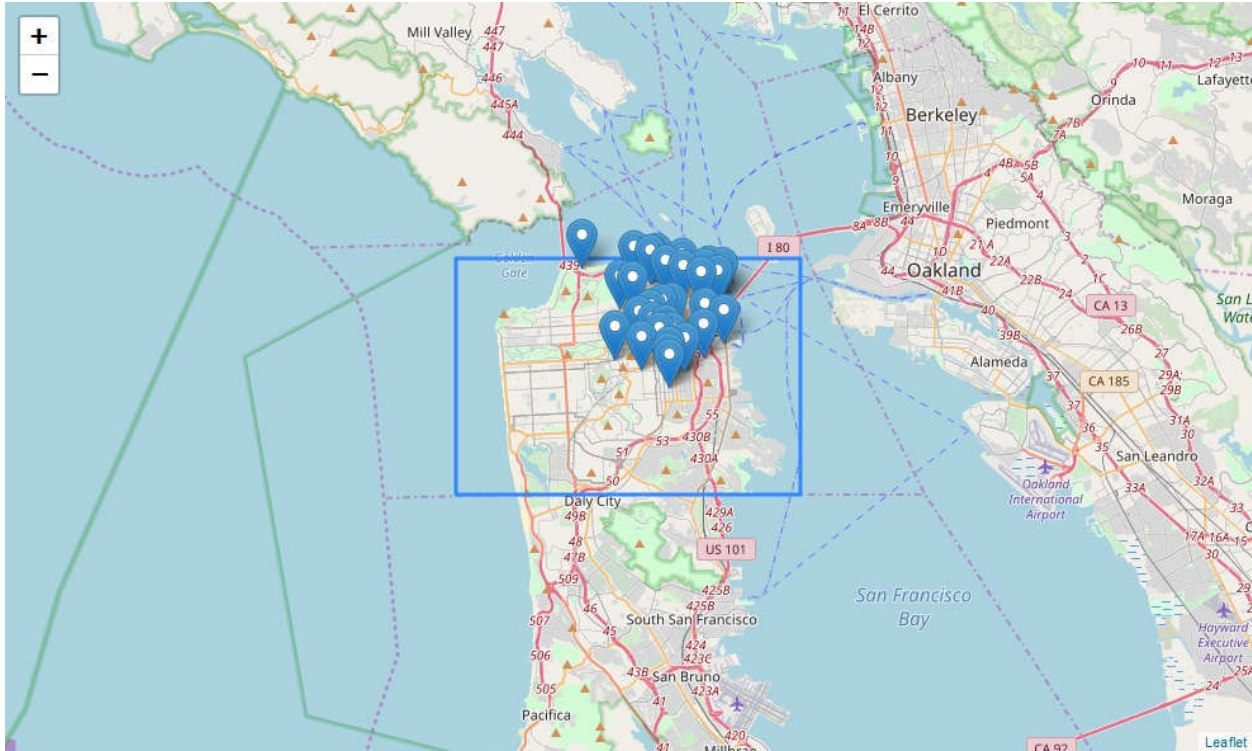
Also I would use Nominatim – service for getting addresses from given latitude and longitude.

To answer main question, “Where are the best places to open Yoga Studio in San Francisco” I should do following steps:

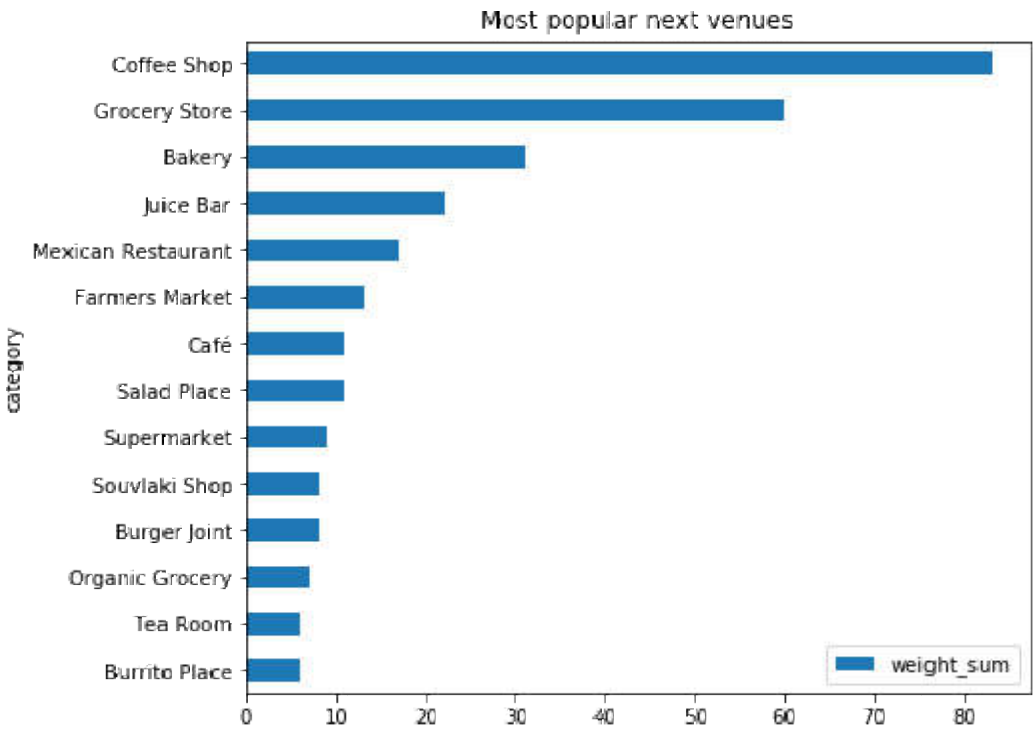
1. Get list of all existing Yoga Studios in San Francisco. Because Foursquare API returns venues of given categories in circular area around given point, I should clean results, so only venues from San Francisco stay.
2. For each given Yoga Studio request popular next venues
3. Calculate most popular categories of them
4. Request list of venues of that categories in San Francisco
5. Generate grid of points inside given area on map (because we are interested in specified area of the city)
6. For each point determine if it acceptable or not, based on distance to nearest existing Yoga Studio. Also I should calculate score for each point based on proximity to popular next venues.
7. Next step would be clusterizing of the appropriate points to determine areas of interest
8. Describing gotten result

Methodology

During research I found that there is 33 Yoga Studios in San Francisco.



Most popular categories among next venues are Coffee Shops, Grocery Stores and Bakeries



For scoring each point in area I used following formula:

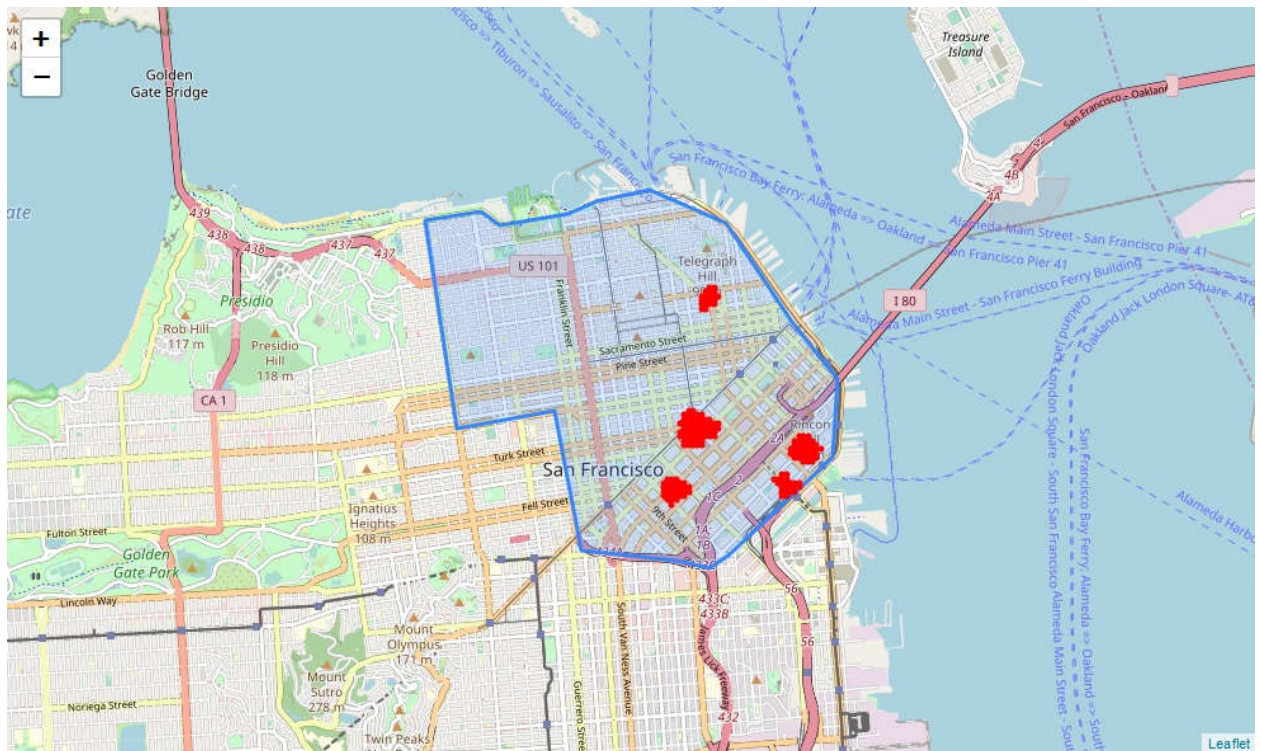
$$mean = \left(\frac{distance_{coffee}}{mean_distance_{coffee}}, \frac{distance_{grocery}}{mean_distance_{grocery}}, \frac{distance_{bakery}}{mean_distance_{bakery}} \right)$$

$$score = \frac{distance_{yoga}}{minYogaRadius} / mean$$

At first I calculate mean of scaled distances from each point to closest venue of matching type. The lower this value – the better, it means that next venues are really close to this point.

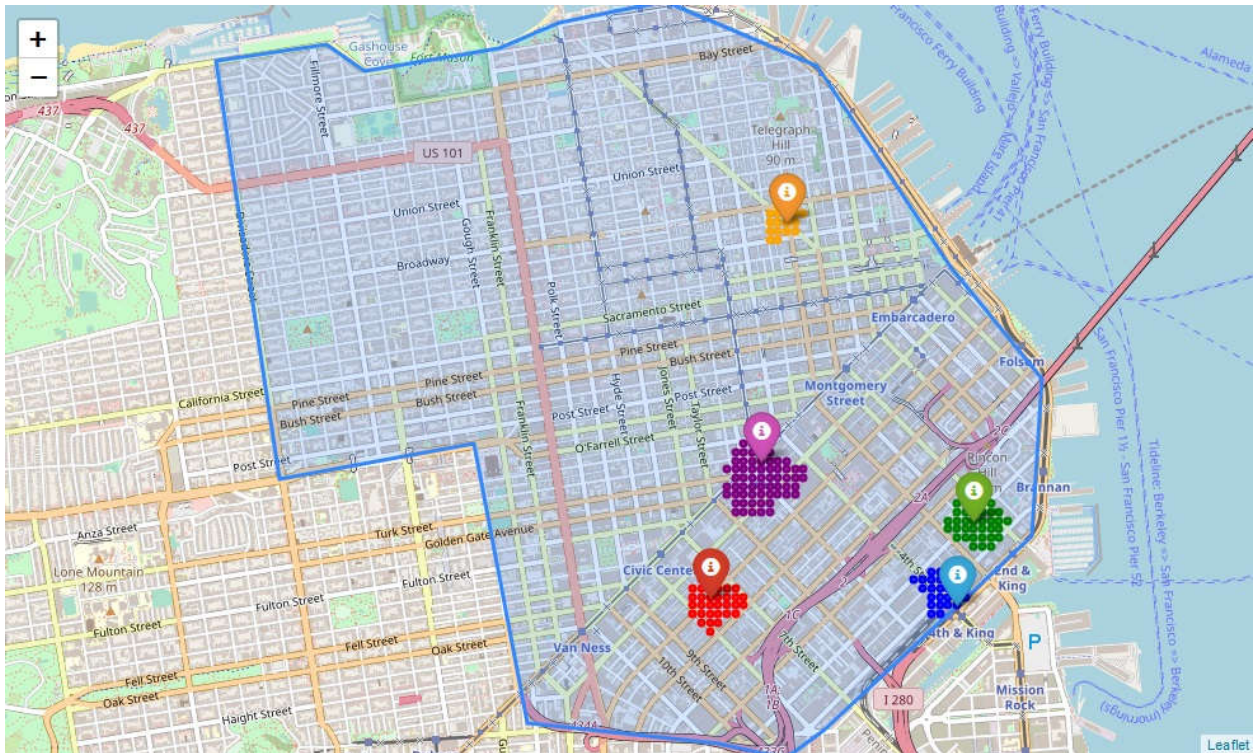
At second I calculate scaled distance to nearest Yoga Studio. This time the higher value – the better.

From all valid points I choose all with score higher than 2 as most prominent. There remain 175 of such points.



For clusterisation task I choose DBSCAN algorithm because it supports non-circular areas.

It works really well:



Results

In given area there are 5 clusters of appropriate points with most prominent (in each cluster) in:

1. Westfield San Francisco Centre, 845, Market Street, Union Square
2. The Beacon, 260, King Street, South Beach
3. Cellarmaker Brewing Co., 1150, Howard Street, West SoMa
4. Lifemark Group, Columbus Avenue, Chinatown
5. Dropbox, 333, Brannan Street, West SoMa

Conclusion

In this project I made research of most prominent places to open new Yoga Studio in San Francisco. However, similar methodology can be applied to venues of other types or placed in different cities.

Thanks for reading!