



# FAMILY FRIENDLY NEIGHBORHOODS IN SAN FRANCISCO, CA

Created by Oxana Matveyuk



# Objective

San Francisco is a popular tourist destination, a big financial center, famous historical and culture place and has unique unforgettable landscape.

But family people with children prefer to live outside the city.

Our goal is to identify family friendly neighborhoods of San Francisco.

# Data

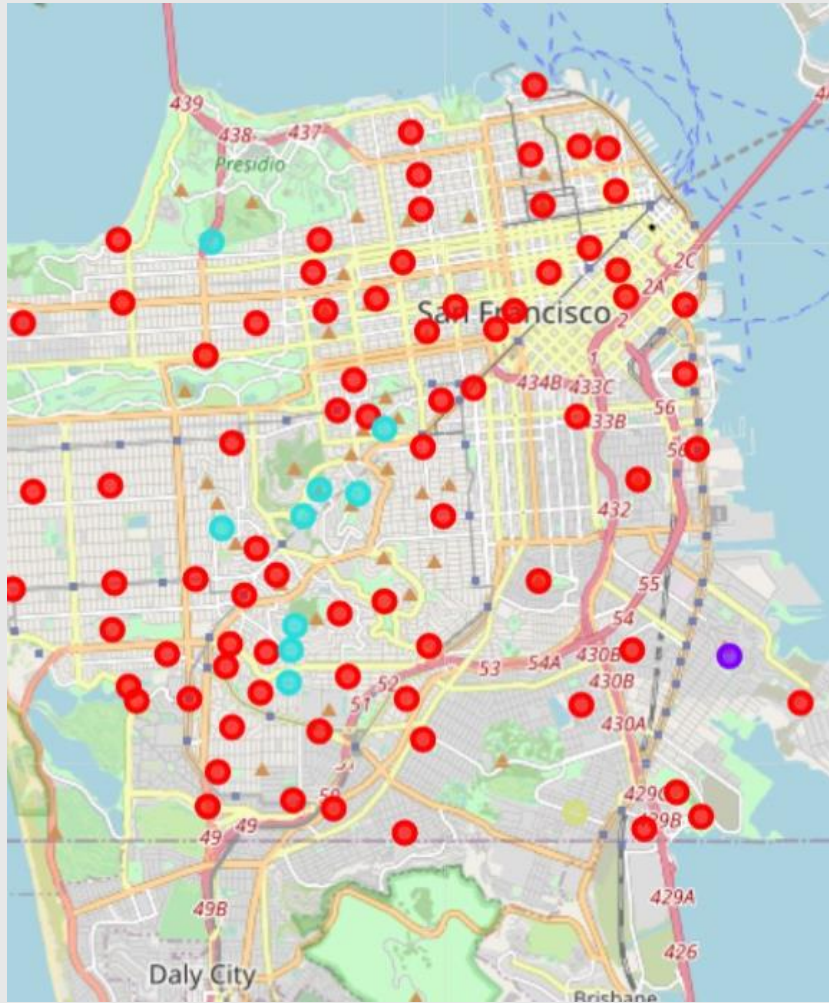
- Neighborhoods of San Francisco (real estate site): <http://www.houseofkinoko.com/district-guide/>. San Francisco has 10 districts. Each neighborhood belongs one district.
- Venues of each neighborhoods: <https://api.foursquare.com>. Venues includes restaurants, parks, schools, culture spots, business locations which help us to define a type/style of neighborhoods.
- geopy library to find coordinates of each neighborhoods. Unfortunately, sometimes geopy finds coordinates not correctly. In this case we will find significant points for such neighborhoods using Google and use them as main point to find coordinates.

# Model

- San Francisco has 10 districts with 88 neighborhoods.
- There are 3772 venues total which belong to 335 unique categories.
- Create vector for each vector which contains the mean of the frequency for each category.
- Use unsupervised machine learning algorithm – K-Means Clustering with  $k=4$  to divide neighborhoods on clusters
- Make decision based on context of each cluster.



# Results



Cluster 0 (red): 77 neighborhoods

- A lot of bars, restaurants
- Shopping
- Friendly for tourists
- Match to young people which spend almost their time on work or with friends

Cluster 1 (violet): 1 neighborhood

- the most isolated neighborhood
- Very industrial

Cluster 2 (cyan): 9 neighborhoods

- A lot of parks, trails
- Sport activities
- Good transportation

Cluster 3 (yellow): 1 neighborhood

- Isolated
- Residential

# Conclusion

## Advantages:

Model defines that Cluster 2 is the most beneficial for families. It contains many parks and sports spots. Most of the neighborhoods in this cluster is located in the city center (the most neighborhoods belong DISTRICT 4 and DISTRICT 5), in close proximity to cultural sites and restaurants/shopping.

## Disadvantages:

Unfortunately, the data does not have enough information about schools and childcare which we can use for future improvements of the model.



# Family friendly neighborhoods

- Clarendon Heights
- Corona Heights
- Golden Gate Heights
- Lake Mountain
- Midtown Terrace
- Sherwood Forest
- Twin Peaks
- Westwood Highlands
- Westwood Park