Recommended areas for French expatriates in San Francisco

DESCRIPTION & DISCUSSION OF THE BACKGROUND

Our company office in San Francisco nave newcomers from Paris office who will stay for 18 months in San Francisco office. As they need to onboard quickly and safely, the company would like to identify and suggest areas for accommodation.

Corporate human resources have given some guidelines regarding safety and also distance to the office as the company cannot finance a corporate vehicle for all of them. From the employee’s perspective, after a survey, it appears that the number one concern was finding a French speaking school for their children. Taking into account all those requirements, we will build an interactive map enhancing most appropriate areas for living

DATA COLLECTION & LIBRARY IMPORT

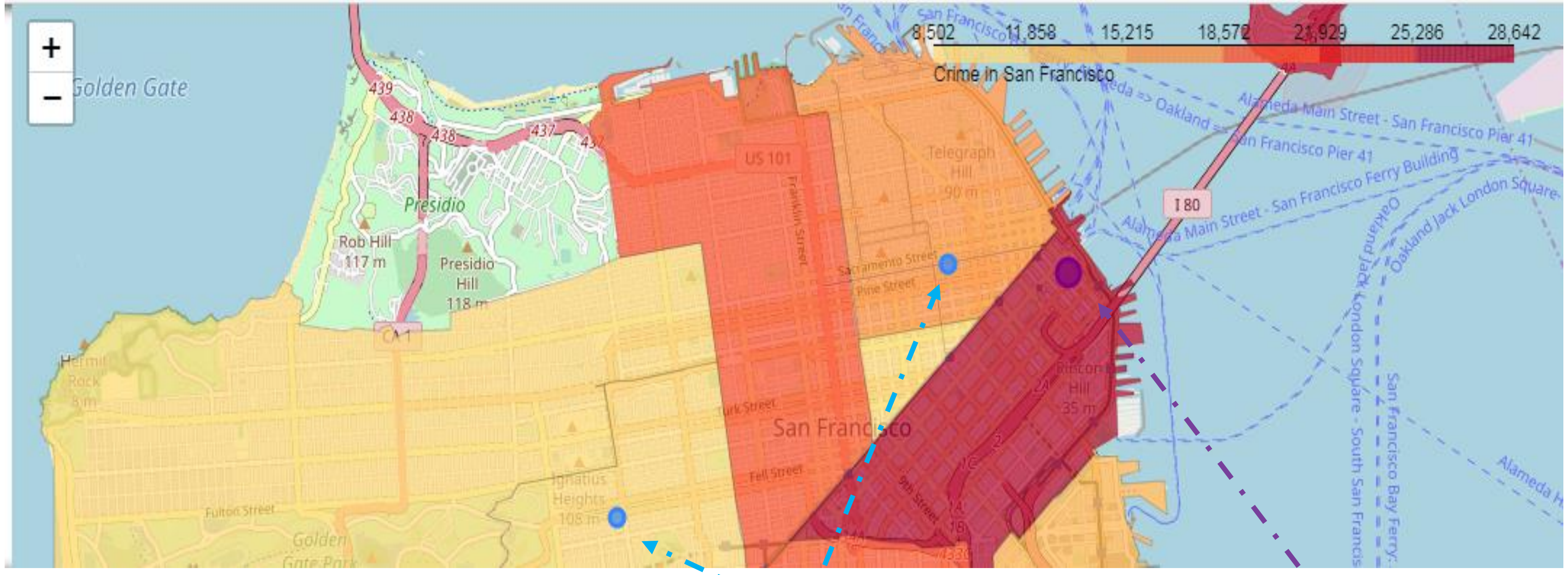
To take into account those 3 criteria, we will use:

- Crime dataset in San Francisco
- Latitude/Longitude for our company office in San Francisco from Foursquare
- Latitude/Longitude main French speaking schools in San Francisco from Foursquare

We will use matplotlib and folium for visualization. With geopy we will be able to measure distances in miles from the office and French schools.

RECOMMENDED AREAS FOR FRENCH EXPATRIATES IN SAN FRANCISCO

DATA EXPLORATION



The two blue points are the French speaking schools and the purple dot is the company office. They are both close to unsafe areas. To enable newcomers to balance those aspects, let's try to cluster the different areas taking into distance and safety considerations

METHODOLOGY

(1) Merge 3 criteria in one dataframe

We build a dataframe merging neighborhood, average crime rate, distance to schools and to the office in miles.

(2) Measure distances

We use geopy to measure the distances to schools and office

(3) Clusterisation with k-means methods

RESULTS

Clusters reconciling proximity and crime. The cluster enables us to see comparable location based on those 3 criteria.

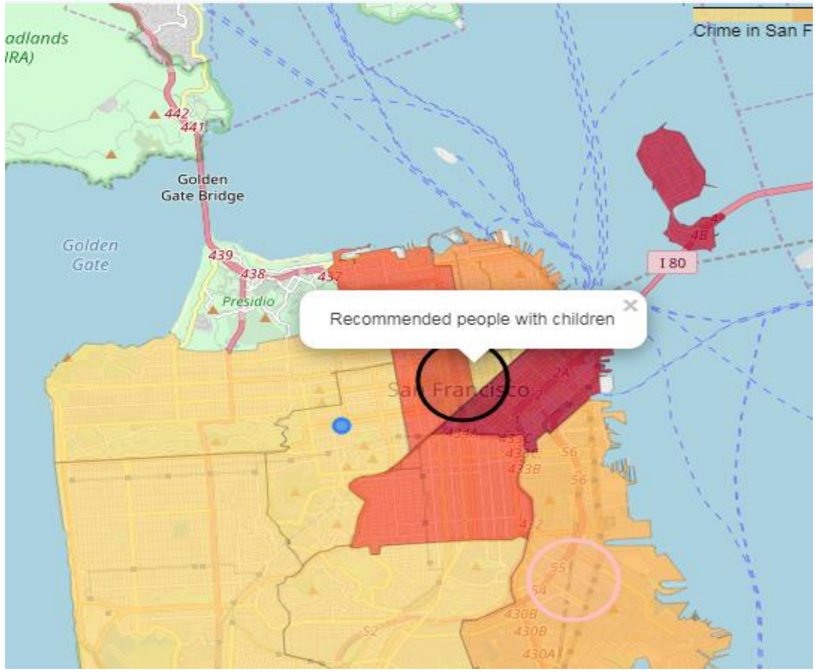
clusters	high school distance	middle school distance	X	Y	office distance	count
0	3.259803	5.262745	-122.447610	37.757474	6.602138	10237.203558
1	3.295506	1.878864	-122.419024	37.780836	3.108539	19145.865390
2	4.633655	1.122557	-122.405254	37.779807	1.636933	28445.000000
3	6.256219	3.590586	-122.393650	37.739968	3.209994	14303.000000

To exclude cluster 0 too far from the office and cluster 2 unsafe

To recommend cluster 1 for families and cluster 3 for employees without children at school

RECOMMENDED AREAS FOR FRENCH EXPATRIATES IN SAN FRANCISCO

RESULTS



DISCUSSION

Additional data could be added to enrich this classification with other important consideration for both corporate and foreign employees such as accommodation price and access to public transportation.

We will be able to use feedbacks based on the location French employees actually rented and use supervised learning methodologies

.