SAFE LOCATIONS

BUSINESS PROBLEM

**San Francisco is a popular destination for tourists and they look for good accommodations
that has a proximity to transport and a variety of amenities like restaurants. Many of them are
unaware of the crimes in the neighborhood. Being unaware and naive roaming in a foreign
country can cause a big problem. They might get mugged or injured**

PROBLEM TO BE RESOLVED

1. **Find hotel**
2. **Amenities near the hotel**
3. **Display crime in the area in a easy to understand manner**
4. **Provide suggestions for sight seeing**

INTRESETED AUDIENCE

**Tourists who are looking for accommodation in San Francisco and require proximity to metro
and different amenities like bars, restaurants, etc. Tourist who is looking for a trendy place to
visit and they can make educated decision dased on crime in the area and proximity to
government authorities and safe places like a cafe where they can ask for help if they get mugged**

FACTS

**A popular website exclaims "With a crime rate of 64 per one thousand residents, San Francisco
has one of the highest crime rates in America compared to all communities of all sizes – from**

the smallest towns to the very largest cities”
SOURCE: https://www.neighborhoodscout.com/ca/san-francisco/crime

DATA SECTION

Data Requirements

Geodata from Foursquare for all the following:
Hotels
Metro
Amenities
Trendy places
Government authorities
Cafes
(Foursquare)
San Francisco Crime Data which depicts all the total crime in the neighborhood
(https://cocl.us/sanfran_crime_dataset)
Geojson file for boundaries
(https://cocl.us/sanfran_geojson/sanfran_geo.json)

Processing

Data processing will be done using Python 3. Foursquare API will be used to get all the required
geographical information. Pandas will be used for manipulation and Folium will be used
mapping.

```
In [1]: import pandas as pd
In [2]: df_sfcrime = pd.read_csv(“https://cocl.us/sanfran_crime_dataset”)
In [3]: df_tmp = df_sfcrime.groupby([‘PdDistrict’]).count().reset_index()
df_tmp.drop([‘Category’,’Descript’,’DayOfWeek’,’Date’,’Time’,
‘Resolution’,’Address’,’X’,’Y’,’Location’,’PdId’], axis=1, inplace=True)
```

df_tmp.rename(columns={'PdDistrict':'Neighbourhood', 'IncidntNum':'Count'}, inplace=True)
In [4]: df_tmp
Out[4]:

	Neighbourhood	Count
0	BAYVIEW	14303
1	CENTRAL	17666
2	INGLESIDE	11594
3	MISSION	19503
4	NORTHERN	20100
5	PARK	8699
6	RICHMOND	8922
7	SOUTHERN	28445
8	TARAVAL	11325
9	TENDERLOIN	9942

Methodology section

The San Francisco data set was a crime statistic data set with many components like X, Coordinates and type of felony, etc. So data was cleaned, removing unnecessary columns and summation of all the crime in the neighborhood. was done.

Then another data set was created with longitude and latitude of all the neighborhoods and merged with the previous dataset externally in excel.

The GeoJSON file has boundary information for making a choropleth map so the crime can be
plotted wit efficiency and have a clean look.

Foursquare API was used to fetch hotels in San Francisco and plotted on a map using folium.

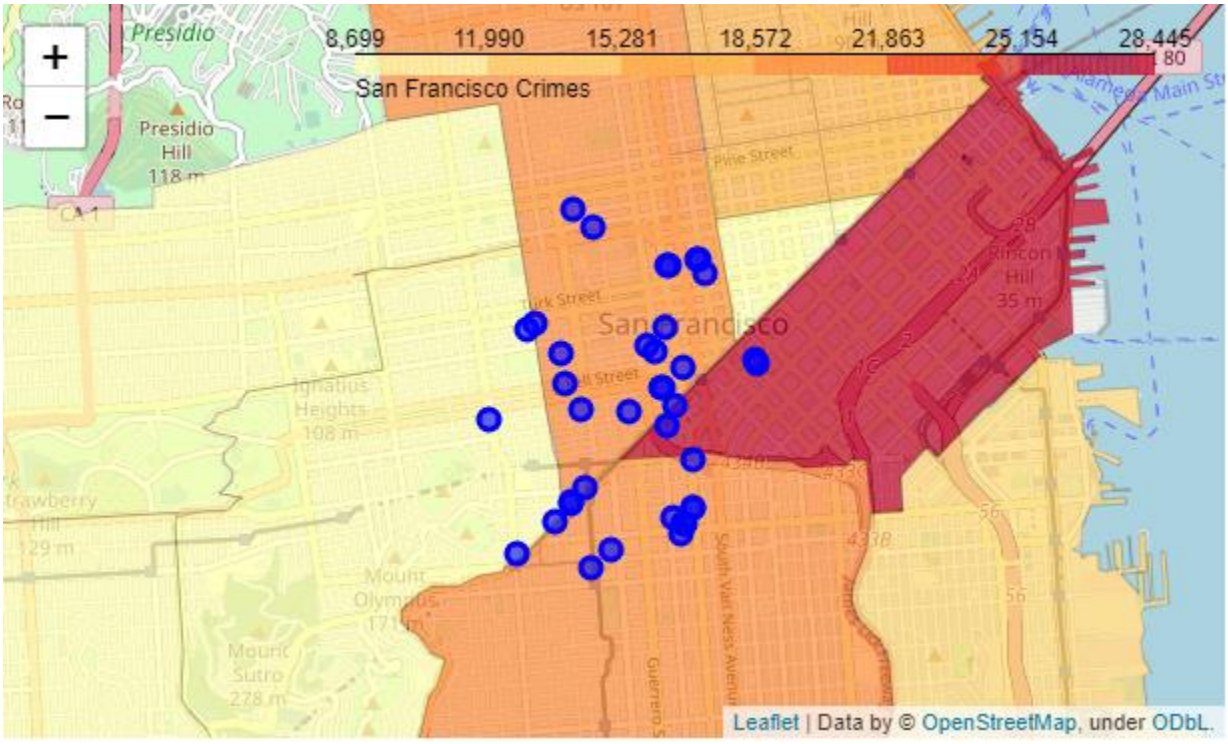
Foursquare API was used to gather information about the amenities in San Francisco and then it was clustered together using Kmeans. Clustering was done so that for each neighborhood a cluster of hotspots with lots of the amenities could be crated and was mapped in folium. This was done so that tourists can make get a rough sense of where all the amenities are in a neighborhood.

As transport in important and SF having metro and bus facilities throughout the city, it becomes an important factor for some to choose a hotel with proximity to the station.

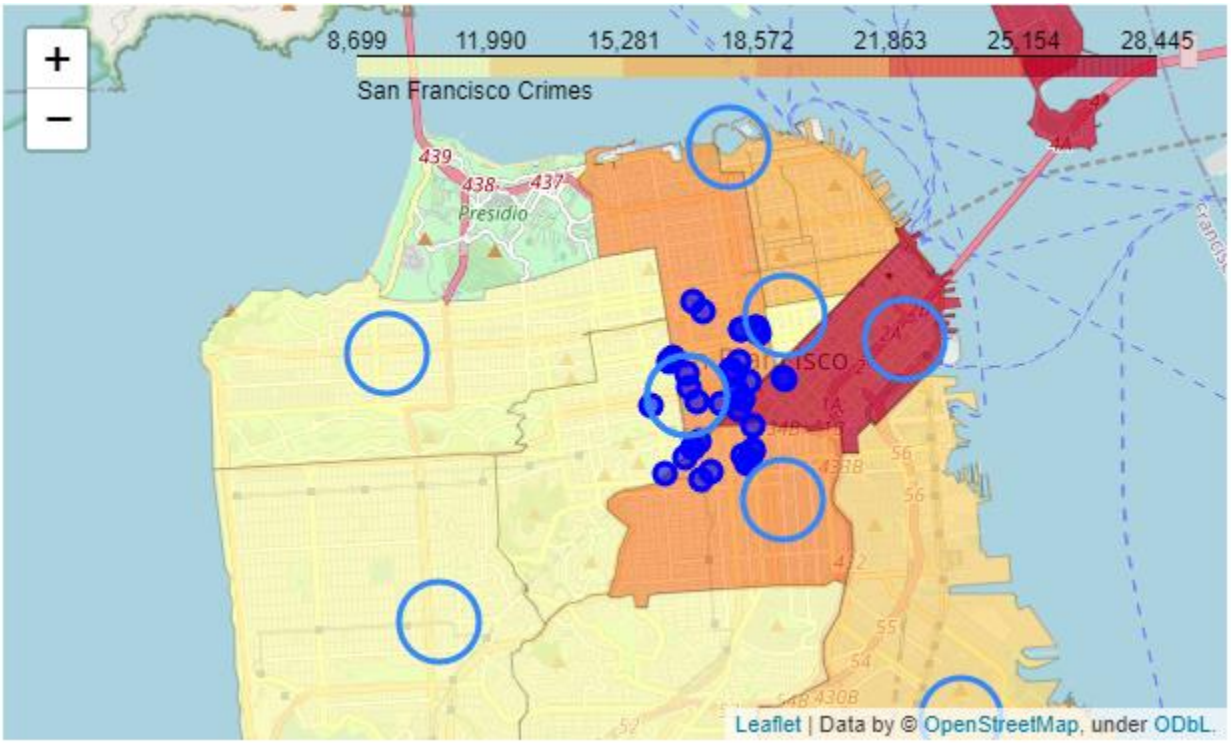
After making a decision Foursquare API was to get sightseeing the location in SF and after a location was decided police station and safe location near it are also displayed.

RESULTS

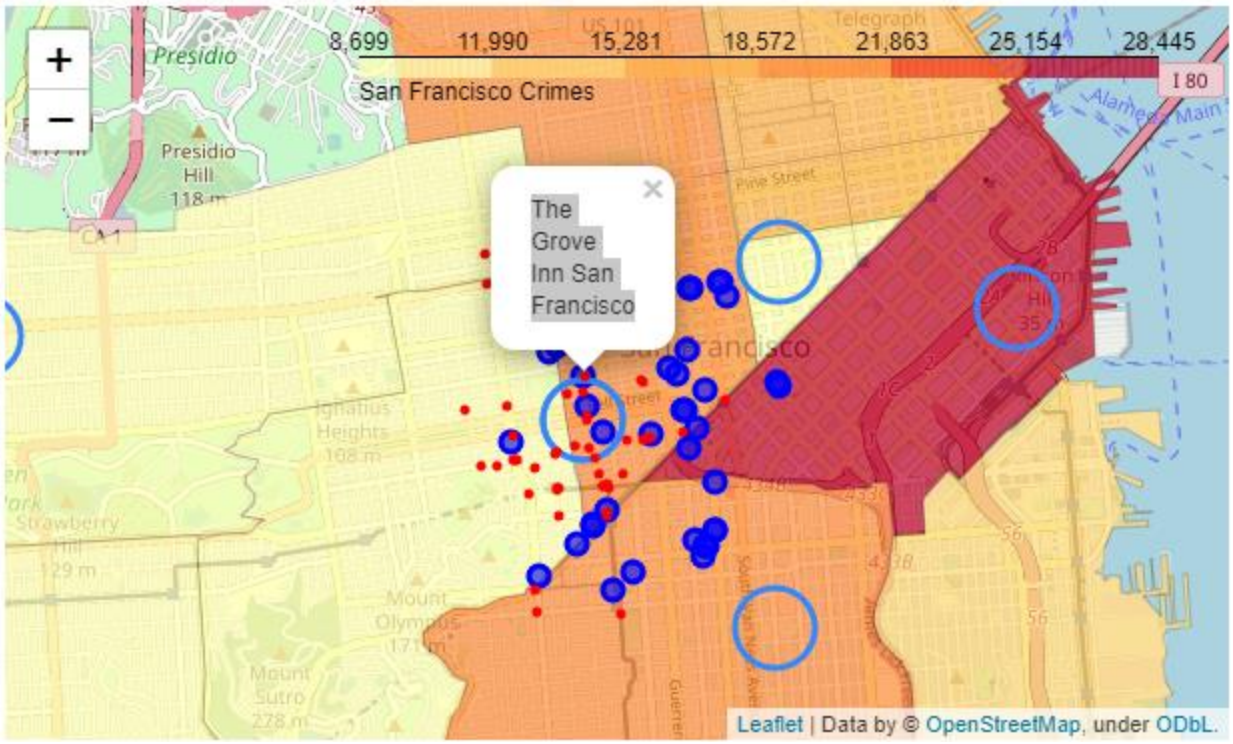
ALL HOTELS WERE PLOTTED CORRECTLY



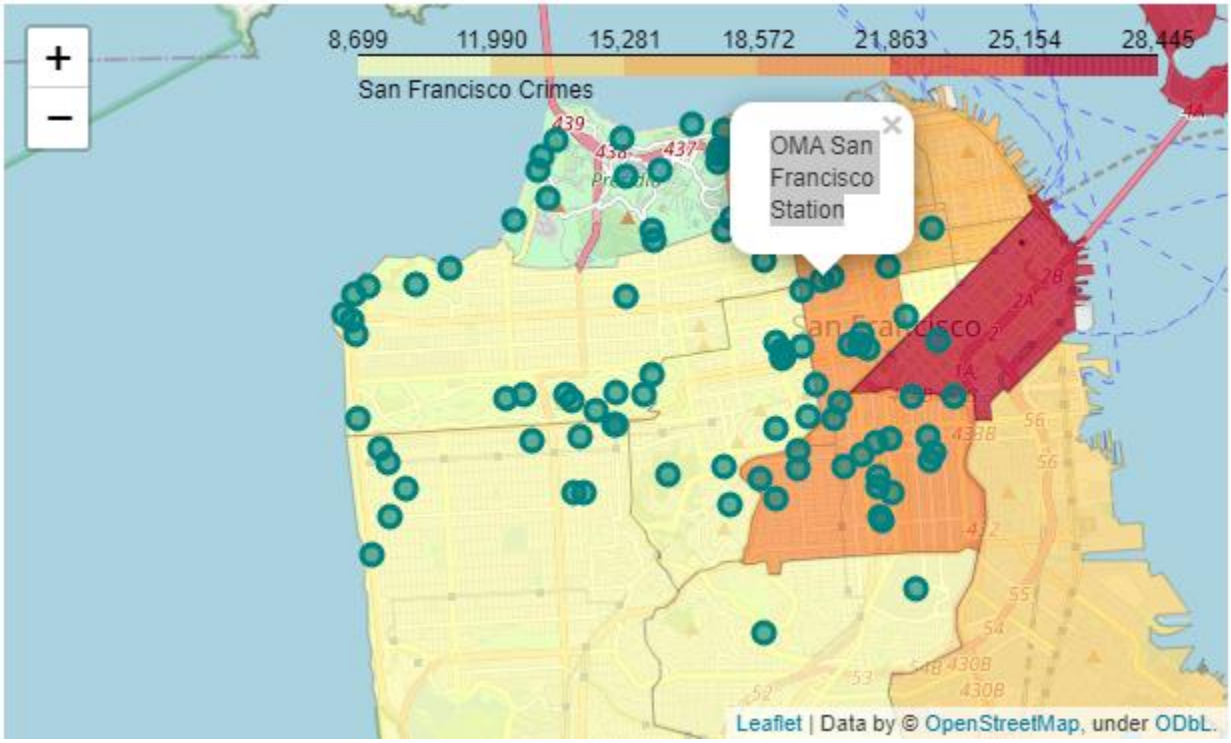
CLUSTERS WERE CREATED AND THEN WERE PLOTTED



METRO STATIONS WERE PLOTTED



TRENDY LOCATIOS WERE PLOTTED



THE POPUP MARKER IS THE HOTEL AND THE SELECTED LOCATION AND THE GREEN DOTARE SAFE LOCATIONS AND BLUE ARE POLICE STATION

So the tourist can make a informed decision on what hotel to choose and what place to visit so they can reduce their chances of being mugged or a victim of a crime

DISCUSSION

As SF has a high crime rate, tourist can feel unsafe wandering in the streets and as they travel across SF they should have peace that they are going in a low crime area or something along these lines.

It is also important to be prepared if anything happens to the tourists. They should no they location of police stations or places where they can feel safe and get help.

CONCLUSION

As anyone would like to feel safe in a city this program helps them by recommending a good hotel based on machine learning.
Nice places to visit and if anything goes sideways also provides knowledge about police and
safe locations

THANK YOU!