Analysing a Crime dataset using Folium

OVERVIEW

A project that is intended in providing a better understanding of creating maps and visualize. Usage of libraries Pandas.
In the course of completing the project, you use folium and create a Choropleth map.

Problem Statement

Creation of maps with markers to explore crime rate in San Francisco, California.
Eventually creation of a Chloropleth map to visualize the crime rate in San Francisco.

Software Requirements

1. Programming Language : Python

2. Environment: Jupyter Notebooks / Google Colab

3. Database: CSV(export type)

4. Operation System: Windows XP or above

5. Libraries Used: Pandas,Folium

6.Datasets used:

SanFrancisco:

https://cocl.us/sanfran_crime_dataset(https://cocl.us/sanfran_crime_dataset)

Geojson file: https://cocl.us/sanfran_geojson (https://cocl.us/sanfran_geojson)

1. Open a New Notebook and import the required libraries and read the csv file

```
import pandas as pd
df_sfcrime = pd.read_csv("*insert the path of CSV file")
```

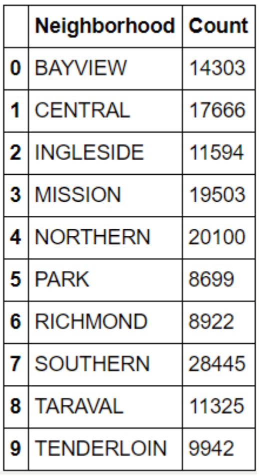
- Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures.Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

- df_sfcrime reads the data from the dataset sf_crime.

2. Set the dataset accordingly

```
    df_tmp = df_sfcrime.groupby(['PdDistrict']).count().reset_i
    ndex()
    df_tmp.drop(['Category','Descript','DayOfWeek','Date','Time
    ', 'Resolution','Address','X','Y','Location','PdId'], axis=
    1, inplace=True)
   df_tmp.rename(columns={'PdDistrict':'Neighborhood', 'IncidtN
um':'Count'}, inplace=True)
     df_tmp
```

```
 Output:
```



	Neighborhood	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

- The groupby operation is used for splitting the object, applying the function and combining the results.When computing large amounts of data the groupby function is very useful.

- df_tmp.drop drops the column.

- df_tmp.rename renames the column.

3. Importing Folium

```
import folium
sf_geo = r'*insert path of the GeoJSON correctly*'
print('Folium installed and imported!')
```

- In this step we import the folium library.

- Folium makes it easy to visualize data that’s been manipulated in Python on an interactive leaflet map. It enables both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map.

4. Creating a plain map(Choropleth) of Sanfrancisco:

```
sf_map = folium.Map(location=[37.773972, -122.431297], zoom_sta
rt=12) #, tiles='Mapbox Bright')
sf_map.choropleth(
    geo_data=sf_geo,
    data=df_tmp,
    columns=['Neighborhood','Count'],
    key_on='feature.properties.DISTRICT',
    fill_color='YlOrRd',
    fill_opacity=0.7,
    line_opacity=0.2,
    legend_name='San Francisco Crimes'
)
```

- A choropleth map is a type of thematic map in which a set of pre-defined areas is colored or patterned in proportion to a statistical variable that represents an aggregate summary of a geographic characteristic within each area, such as population density or per-capita income.

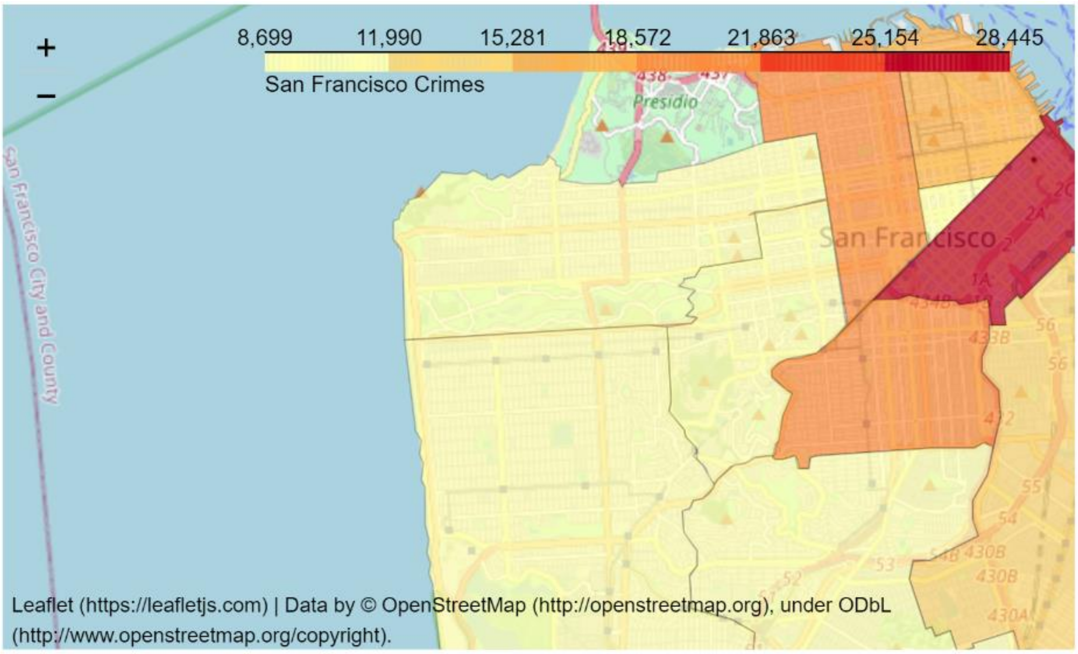
- Lets plot a cholropleth map for the state of San Francisco.

- While plotting a map, we have the mention the starting zoom level so that the user vision is pointed clearly at the respected place. Foilum provides three parameters : zoom_start,min_zoom and max_zoom.

- fill_color is used to fill the background color of the map. fill_opacity is used to adjust the opacity of the visual .

- legend_name displays the required map name.

Output:



+
−
8,699
11,990
15,281
18,572
21,863
25,154
28,445
San Francisco Crimes
Presidio
San Francisco
San Francisco City and County
Leaflet (https://leafletjs.com) | Data by © OpenStreetMap (http://openstreetmap.org), under ODbL (http://www.openstreetmap.org/copyright).

5. Conclusion

We created a Choropleth with markers while exploring and visualising the crime rate in San Franciso.