

▼ Chapter 8 - Exercise 3: Mexico

```
import folium
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

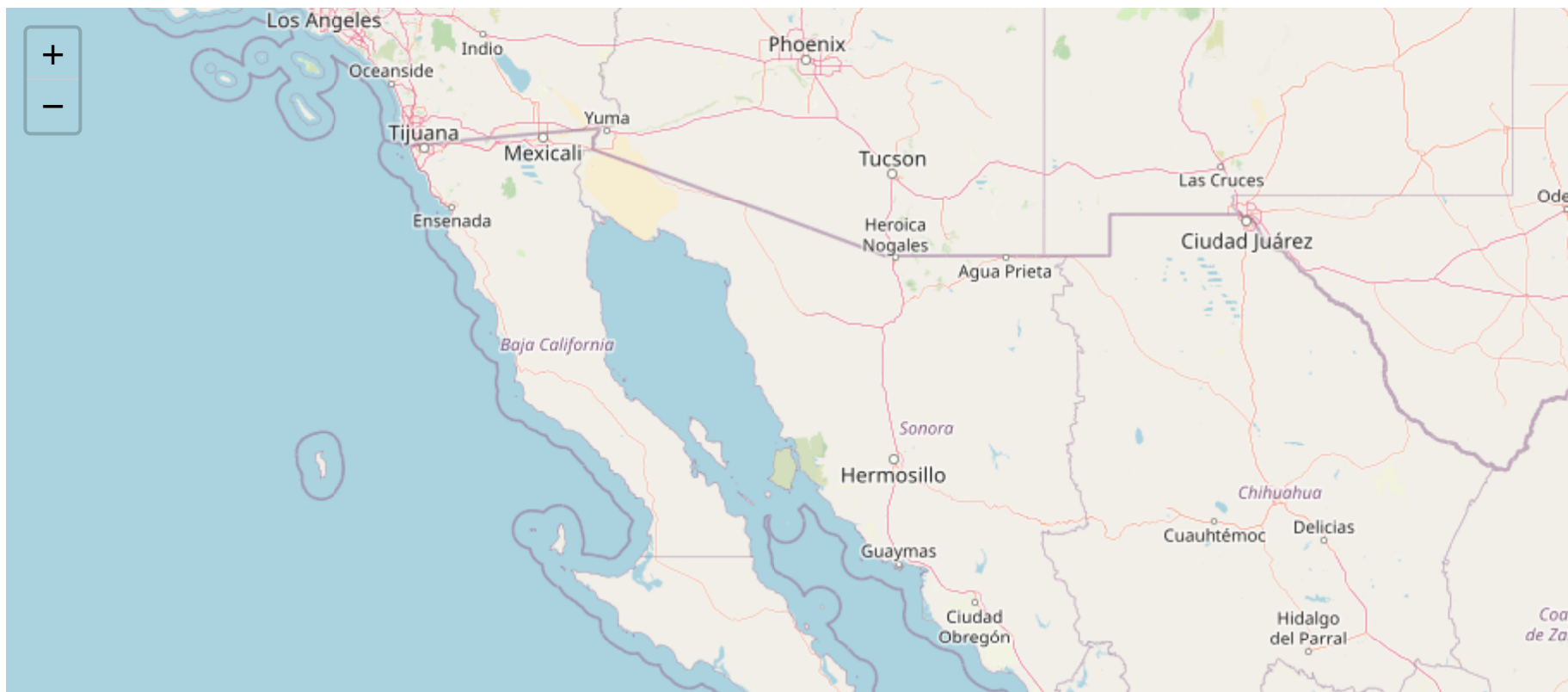
▼ Map: Mexico

```
# Câu 1: Tạo biểu đồ có center là Mexico (location= [23.6345, -102.5528]) với zoom level là 5
mexico_map = folium.Map(location=[23.6345, -102.5528], zoom_start=5)
mexico_map
```



```
# Câu 2: Tạo OpenStreetMap với center là Mexico with zoom level 6.3
```

```
mexico_map = folium.Map(location=[23.6345, -102.5528], zoom_start=6.3, tiles='OpenStreetMap')  
mexico_map
```



▼ Map với Marker

San Francisco Police Department Incidents năm 2016 - được cung cấp từ cổng thông tin dữ liệu công cộng San Francisco. Các sự cố bắt nguồn từ hệ thống báo cáo sự cố tội phạm của Sở cảnh sát San Francisco (SFPD). Được cập nhật hàng ngày, hiển thị dữ liệu cho cả năm 2016. Địa chỉ và vị trí đã được ẩn danh bằng cách di chuyển đến giữa khối (mid-block) hoặc đến một giao lộ (intersection)

```
# Câu 1: Đọc dữ liệu Police_Department_Incidents_-_Previous_Year__2016.csv và lưu vào df_incidents
df_incidents = pd.read_csv('data\Police_Department_Incidents_-_Previous_Year__2016_.csv')
```

So each row consists of 13 features:

1. IncidentNum: Incident Number

2. **Category:** Category of crime or incident
3. **Descript:** Description of the crime or incident
4. **DayOfWeek:** The day of week on which the incident occurred
5. **Date:** The Date on which the incident occurred
6. **Time:** The time of day on which the incident occurred
7. **PdDistrict:** The police department district
8. **Resolution:** The resolution of the crime in terms whether the perpetrator was arrested or not
9. **Address:** The closest address to where the incident took place
10. **X:** The longitude value of the crime location
11. **Y:** The latitude value of the crime location
12. **Location:** A tuple of the latitude and the longitude values
13. **PdId:** The police department ID

```
# Cho biết thông tin của df_incidents
df_incidents.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150500 entries, 0 to 150499
Data columns (total 13 columns):
IncidentNum    150500 non-null int64
Category       150500 non-null object
Descript       150500 non-null object
DayOfWeek      150500 non-null object
Date           150500 non-null object
Time           150500 non-null object
PdDistrict     150499 non-null object
Resolution     150500 non-null object
Address        150500 non-null object
X              150500 non-null float64
Y              150500 non-null float64
Location       150500 non-null object
PdId           150500 non-null int64
dtypes: float64(2), int64(2), object(9)
memory usage: 14.9+ MB
```

```
# Hiển thị 5 dòng dữ liệu đầu của df_incidents
df_incidents.head()
```

	IncidentNum	Category	Descript	DayOfWeek	Date	Time	PdDistrict	Resolution	Address	X
0	120058272	WEAPON LAWS	POSS OF PROHIBITED WEAPON	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARREST, BOOKED	800 Block of BRYANT ST	-122.403405
1	120058272	WEAPON LAWS	FIREARM, LOADED, IN VEHICLE, POSSESSION OR USE	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARREST, BOOKED	800 Block of BRYANT ST	-122.403405
2	141059263	WARRANTS	WARRANT ARREST	Monday	04/25/2016 12:00:00 AM	14:59	BAYVIEW	ARREST, BOOKED	KEITH ST / SHAFTER AV	-122.388856
					04/25/2016				JONES ST	

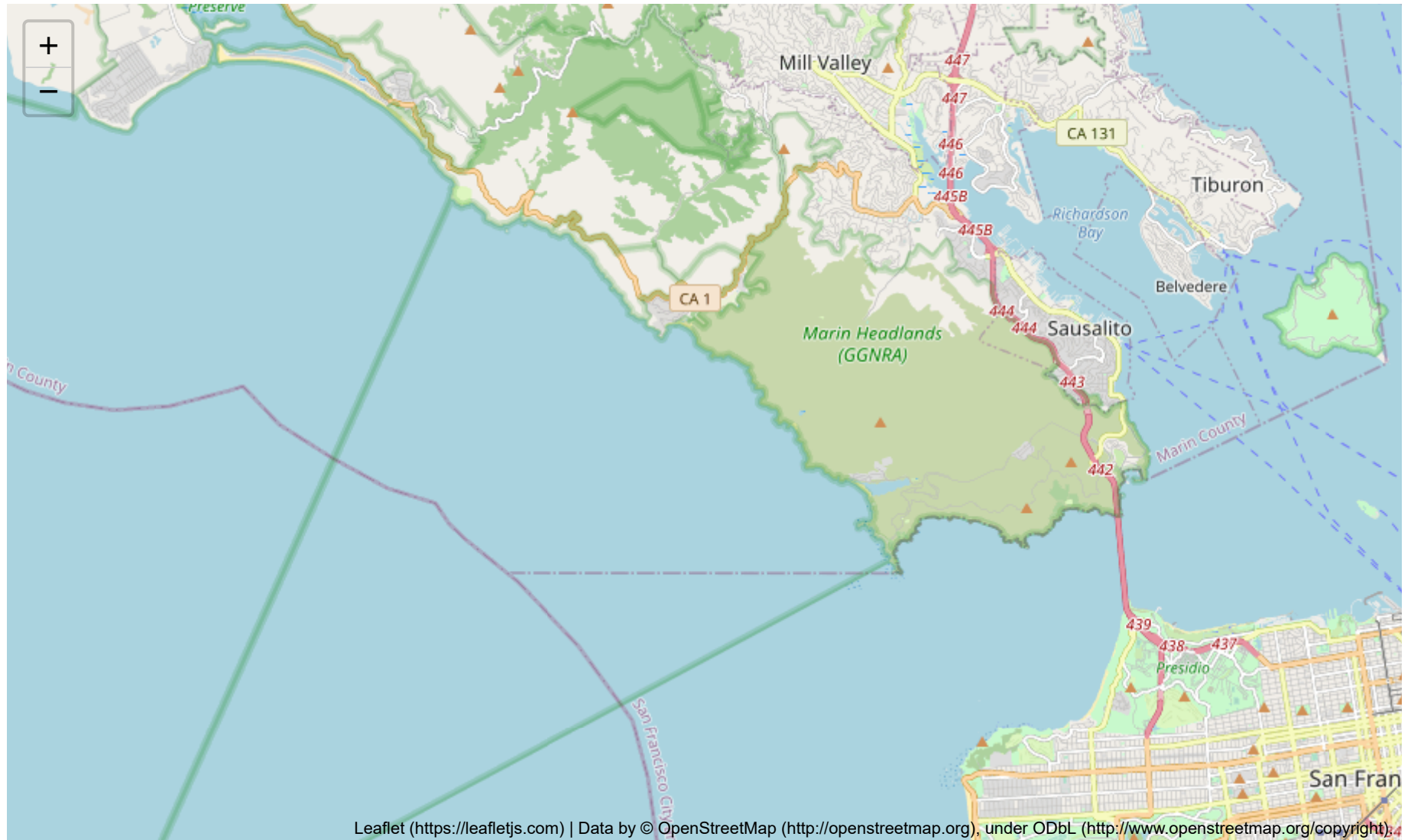
```
# Cho biết kích thước của df_incidents
df_incidents.shape
```

```
(150500, 13)
```

```
# Câu 2: Rút trích dữ liệu: có đến 150.500 tội phạm, diễn ra vào năm 2016.
# Tạo bộ dữ liệu mới chỉ lấy 100 tội phạm đầu tiên trong bộ dữ liệu df_incidents
limit = 100
df_incidents = df_incidents.iloc[0:limit, :]
# Cho biết kích thước của df_incidents
df_incidents.shape
```

```
(100, 13)
```

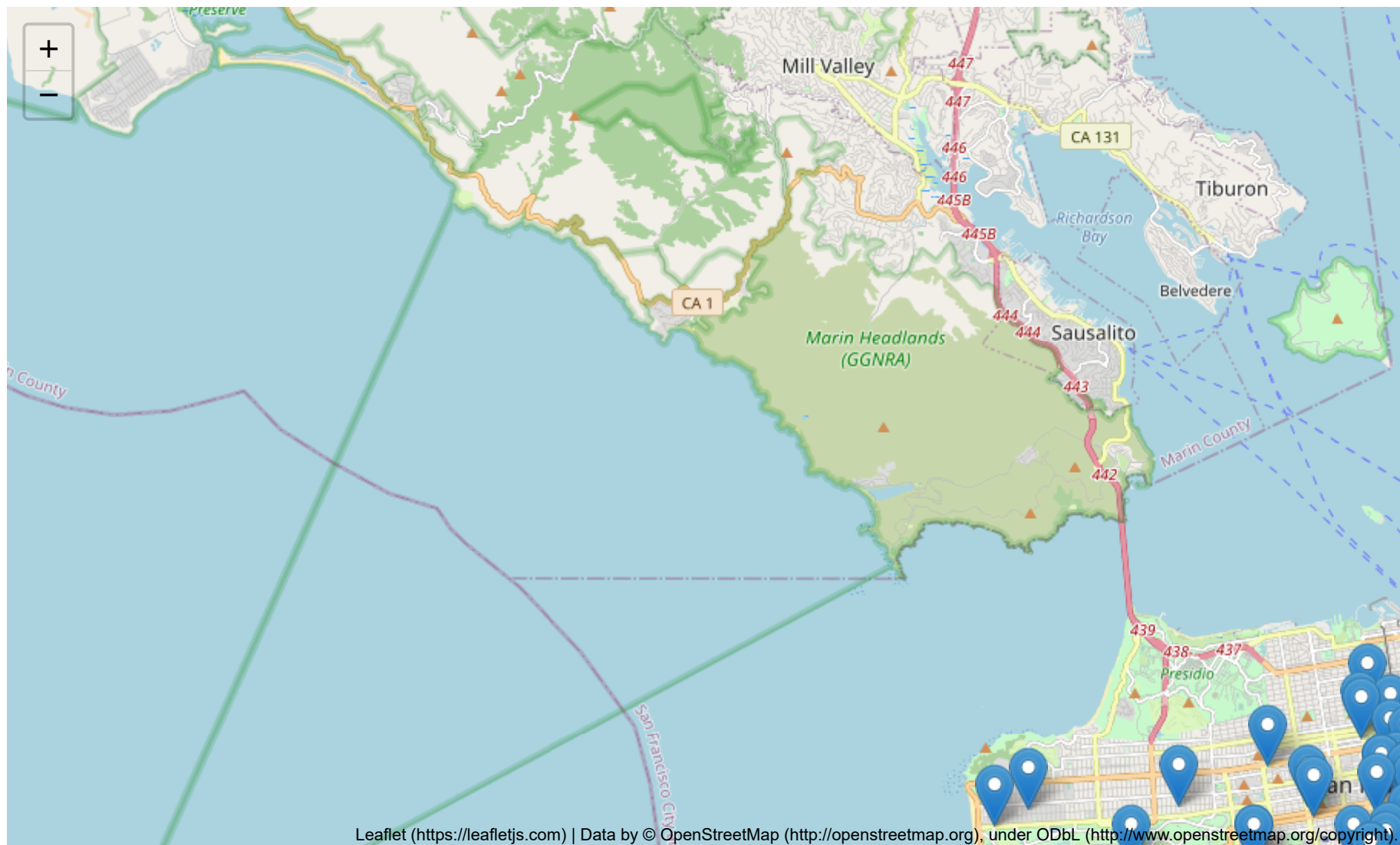
```
# Câu 3: Tạo biểu đồ có center là San Francisco (location= [37.77, -122.42]) với zoom level là 12
sanfran_map = folium.Map(location=[37.77, -122.42], zoom_start=12)
sanfran_map
```



```
# Câu 4: Đưa 100 điểm tội phạm lên bản đồ (với marker)
incidents = folium.map.FeatureGroup()

for lat, lng, in zip(df_incidents.Y, df_incidents.X):
    incidents.add_child(
        folium.features.Marker(
            [lat, lng],
            radius=5,
            color='yellow',
            fill=True,
            fill_color='blue',
            fill_opacity=0.6
        )
    )

sanfran_map.add_child(incidents)
```

```
# Câu 5: Thêm pop-up text sẽ được hiển thị thông tin Category khi người dùng di chuyển chuột qua chuột qua marker
incidents = folium.map.FeatureGroup()

for lat, lng, in zip(df_incidents.Y, df_incidents.X):
    incidents.add_child(
        folium.features.Marker(
            [lat, lng],
            radius=5,
            color='yellow',
            fill=True,
            fill_color='blue',
            fill_opacity=0.6
        )
    )

latitudes = list(df_incidents.Y)
longitudes = list(df_incidents.X)
labels = list(df_incidents.Category)

for lat, lng, label in zip(latitudes, longitudes, labels):
    folium.Marker([lat, lng], popup=label).add_to(sanfran_map)

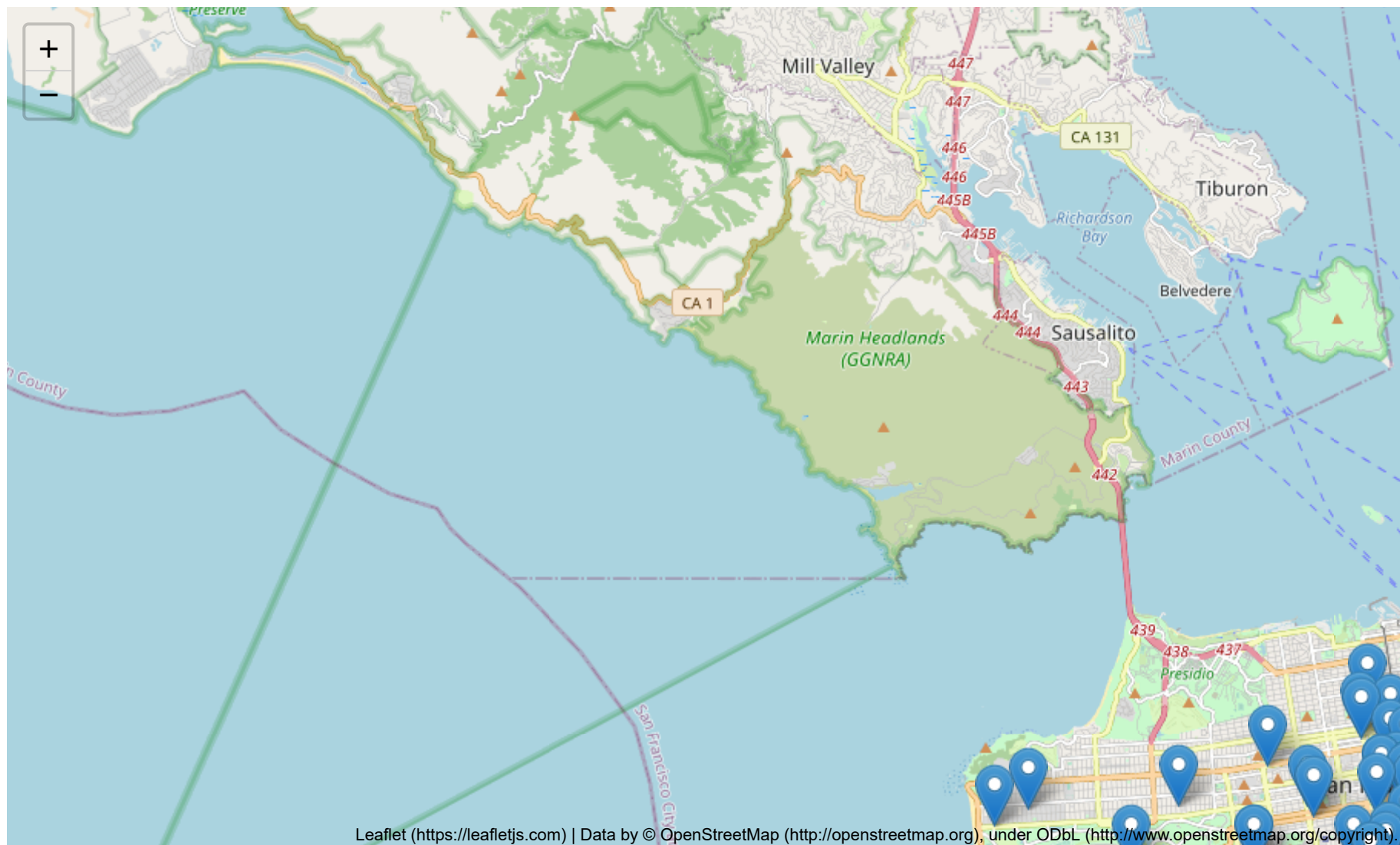
sanfran_map.add_child(incidents)
```




Câu 6: Để cho bản đồ khỏi rối, bỏ đi các location marker, và chỉ thêm text vào từng circle marker
 sanfran_map = folium.Map(location=[37.77, -122.42], zoom_start=12)

```
for lat, lng, label in zip(df_incidents.Y, df_incidents.X, df_incidents.Category):
    folium.features.Marker(
        [lat, lng],
        radius=5,
        color='yellow',
        fill=True,
        popup=label,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(sanfran_map)
```

sanfran_map



```
# Câu 7: Nhóm các markers vào các cluster. Mỗi cluster sẽ hiển thị số lượng các tội phạm trong mỗi neighborhood.
# Gợi ý: Sử dụng MarkerCluster object và thêm tất cả các data point trong dataframe vào object này

from folium import plugins

sanfran_map = folium.Map(location = [37.77, -122.42], zoom_start = 12)

incidents = plugins.MarkerCluster().add_to(sanfran_map)

for lat, lng, label, in zip(df_incidents.Y, df_incidents.X, df_incidents.Category):
    folium.Marker(
        location=[lat, lng],
        icon=None,
        popup=label,
    ).add_to(incidents)

sanfran_map
```

