

# **"San Francisco safe travel and restaurant visits"**

YK Woo

Capstone project in coursera

# 1. Introduction



A. A description of the problem and a discussion of the background.

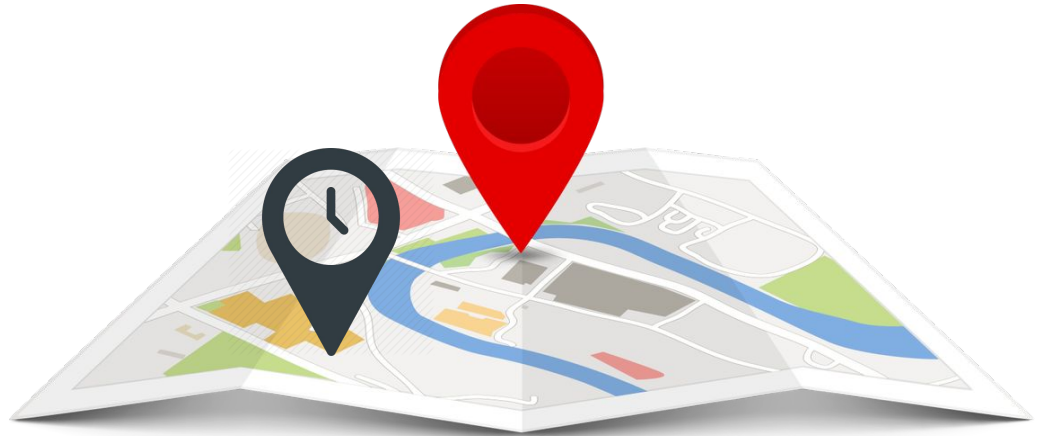
San Francisco is a fairly large city in the United States, the hub of business, finance, and transportation. This city must be an attractive city for many tourists. Unfortunately, a variety of crimes occurred in this city during 2016. I would like to analyze San Francisco's crime patterns by time and region to introduce safer and better neighborhoods for tourists and to recommend good restaurants to visit.

B. A description of the data and how it will be used to solve the problem.

The data can be obtained from the following link. This data includes information such as the time of crime, the type of crime, and the area where the crime occurred. By analyzing and visualizing this data, we will use it as proactive information to help us avoid dangerous risks. In addition, we will use the Foursquare API to add the necessary information based on data such as the location and type of restaurants in the San Francisco area.

The Information that can affect data based on the problems we define :

- Time of Crime
- Day of Crime
- Location of Crime
- Location of Restaurants
- Map of San Francisco



## 2. Import module and data for our project

```
import requests # library to handle requests
import pandas as pd # library for data analysis
import numpy as np # library to handle data in a vectorized manner
import random # library for random number generation
```

```
import matplotlib.pyplot as plt # library for plot visualization
%matplotlib inline # Masic code for plot visualization
```

```
!conda install -c conda-forge geopy --yes # Install for geopy
```

```
from geopy.geocoders import Nominatim # module to convert an address into latitude and longitude values
```

```
# libraries for displaying images
```

```
from IPython.display import Image
```

```
from IPython.core.display import HTML
```

```
# transforming json file into a pandas dataframe library
```

```
from pandas.io.json import json_normalize
```

```
!conda install -c conda-forge folium=0.5.0 --yes # install folium
```

```
import folium # map plotting library
```



## Loading the Crime data of San Francisco

Dataset downloaded and read into a pandas dataframe by URL :

[https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DV0101EN/labs/Data\\_Files/Police\\_Department\\_Incidents\\_-\\_Previous\\_Year\\_\\_2016\\_.csv](https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DV0101EN/labs/Data_Files/Police_Department_Incidents_-_Previous_Year__2016_.csv)

## Pandas dataframe of San Francisco Criminal Issue

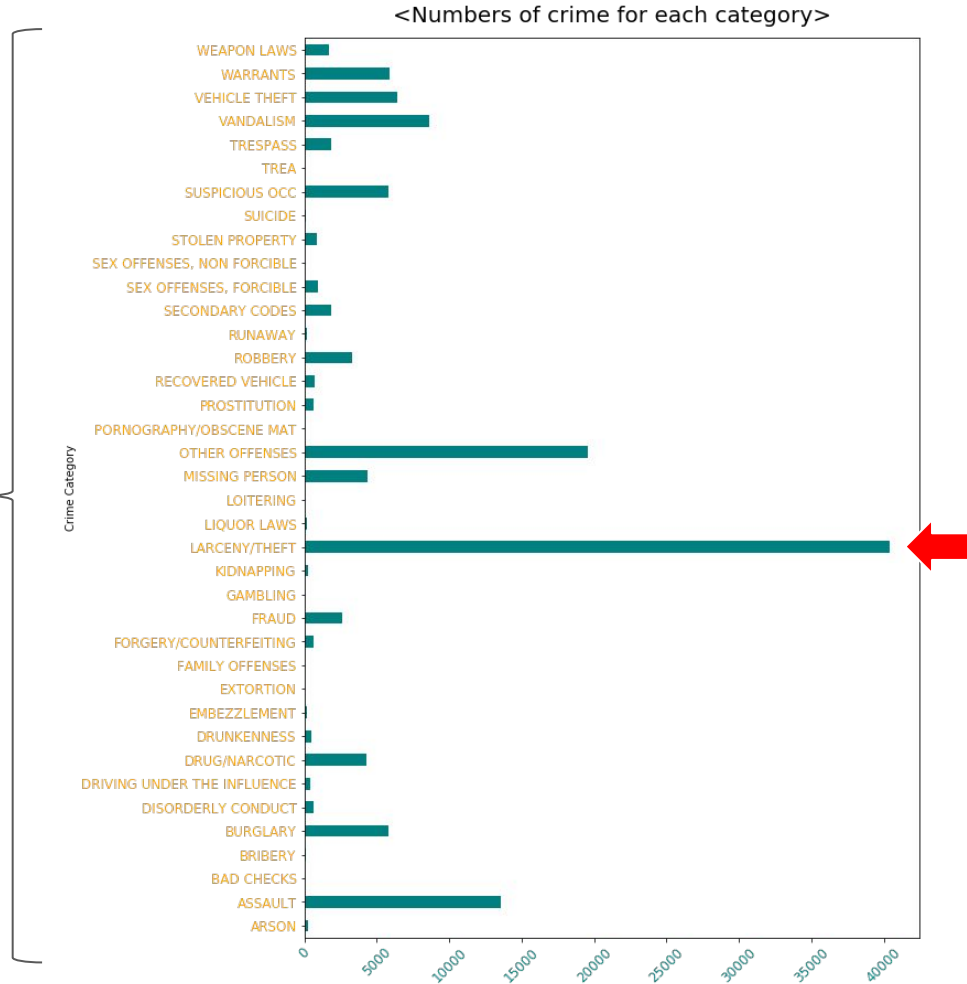
	IncidentNum	Category	Descript	DayOfWeek	Date	Time	PdDistrict	Resolution	Address	X	Y
0	120058272	WEAPON LAWS	POSS OF PROHIBITED WEAPON	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARREST, BOOKED	800 Block of BRYANT ST	-122.403405	37.775421
1	120058272	WEAPON LAWS	FIREARM, LOADED, IN VEHICLE, POSSESSION OR USE	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARREST, BOOKED	800 Block of BRYANT ST	-122.403405	37.775421
2	141059263	WARRANTS	WARRANT ARREST	Monday	04/25/2016 12:00:00 AM	14:59	BAYVIEW	ARREST, BOOKED	KEITH ST / SHAFTER AV	-122.388856	37.729981
5	160002869	ASSAULT	BATTERY	Friday	01/01/2016 12:00:00 AM	21:35	NORTHERN	NONE	1700 Block of BUSH ST	-122.426077	37.788019
6	160003130	OTHER OFFENSES	PAROLE VIOLATION	Saturday	01/02/2016 12:00:00 AM	00:04	SOUTHERN	ARREST, BOOKED	MARY ST / HOWARD ST	-122.405721	37.780879

<Each column's explain>

- 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

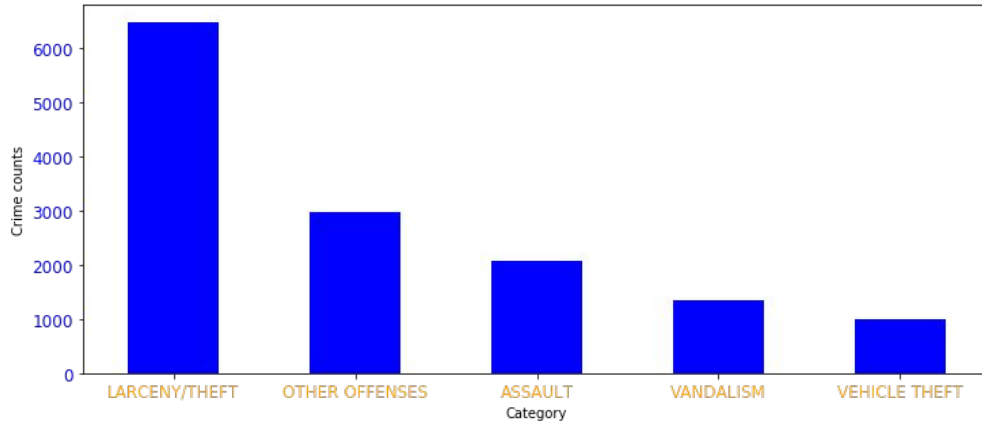
### 3. Analysis of Data

39 kinds of crimes

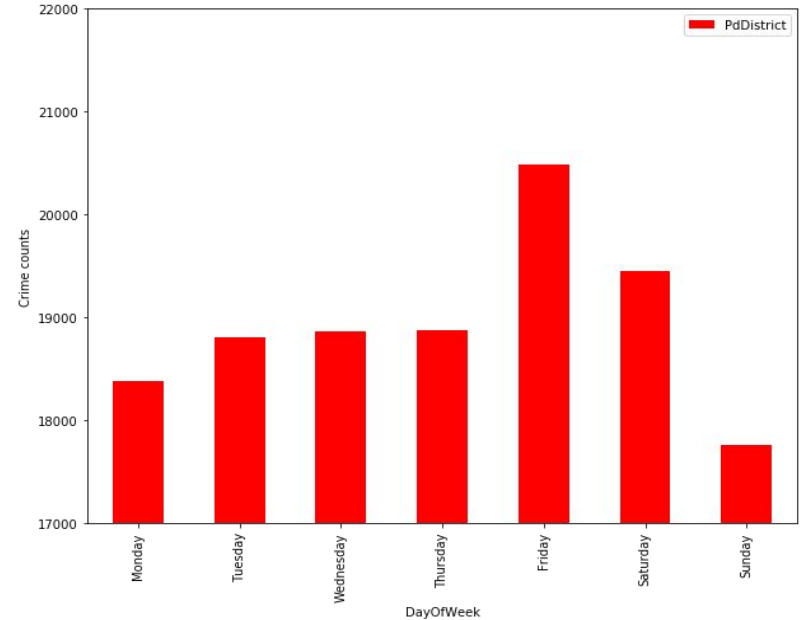


### 3. Analysis of Data : Top 5 , by day

<Top 5 crimes in Friday's crime cases>

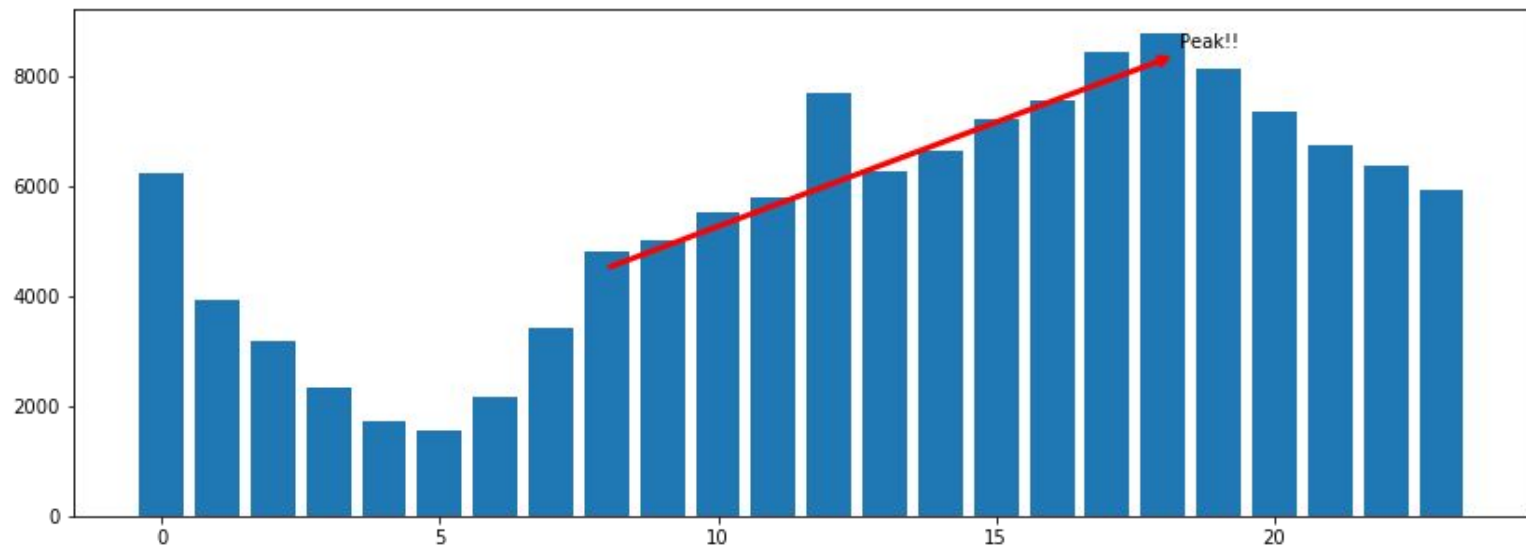


<Numbers of crime for each days>



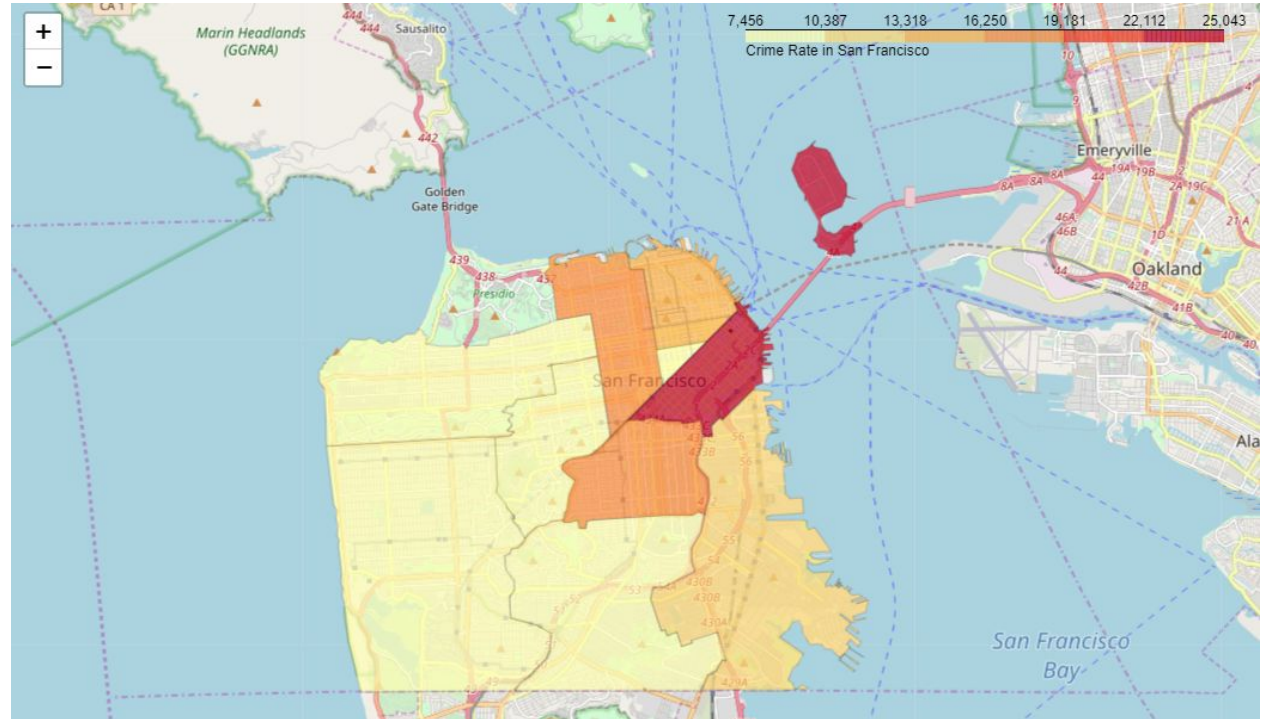


### 3. Analysis of Data : by time

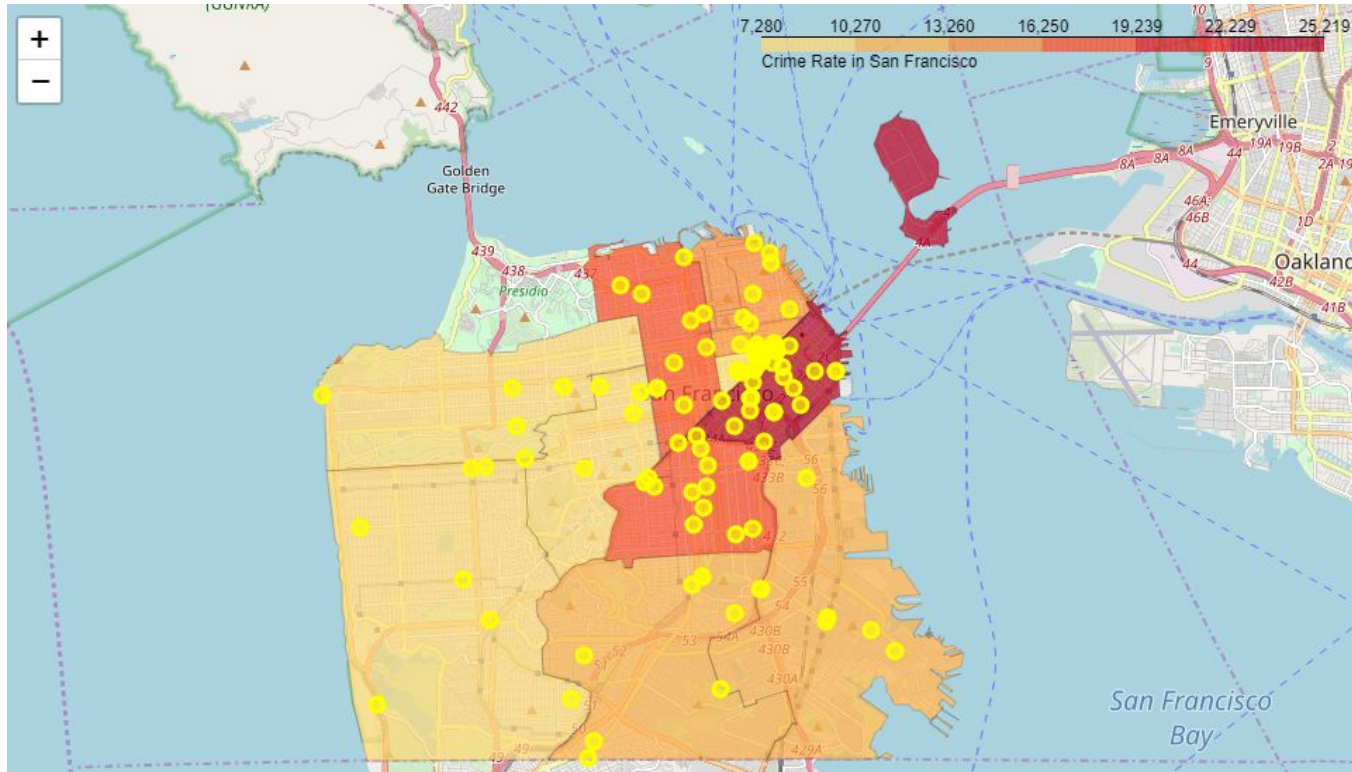


### 3. Analysis of Data : by location on map

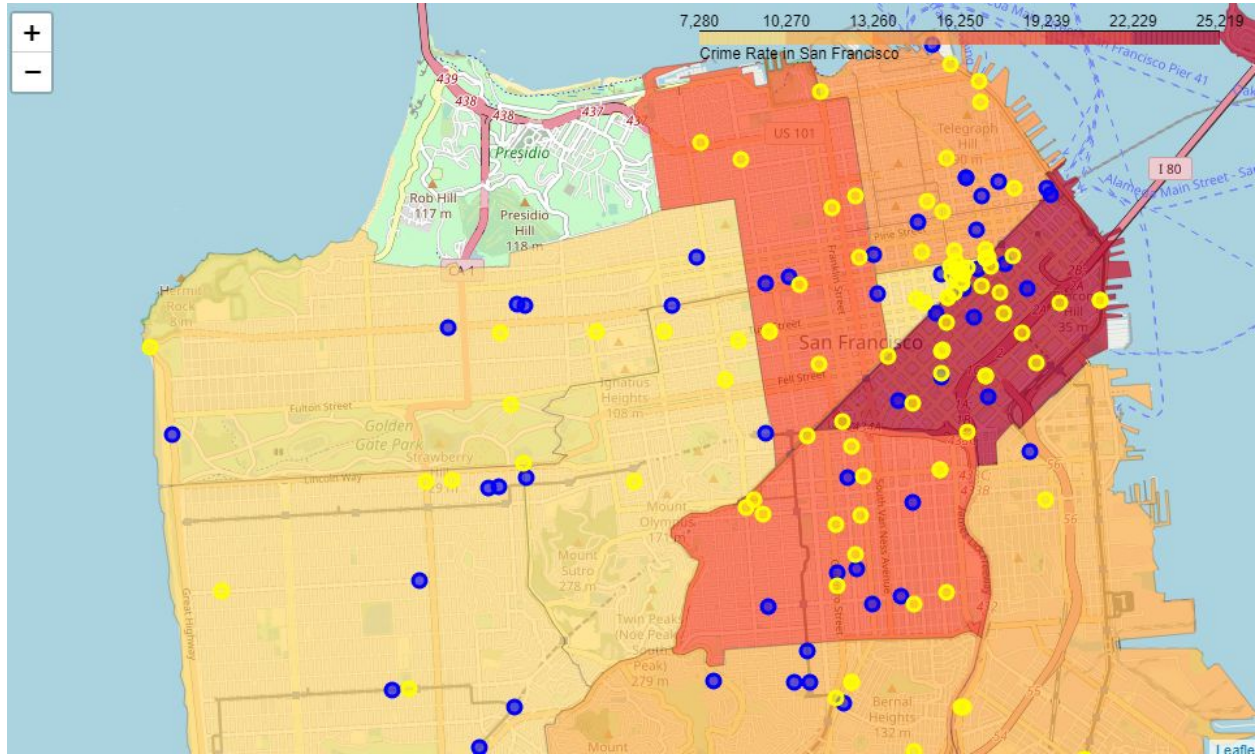
	Neighborhood	Count
0	SOUTHERN	25043
1	NORTHERN	18047
2	MISSION	17242
3	CENTRAL	15356
4	BAYVIEW	12991
5	INGLESIDE	10358
6	TARAVAL	9901
7	TENDERLOIN	8563
8	RICHMOND	7676
9	PARK	7456



### 3. Analysis of Data : Based on the **LARCENY/THEFT Crime**



### 3. Analysis of Data : Based on the **LARCENY/THEFT Crime** and **Restaurants**



### 3. Analysis of Data

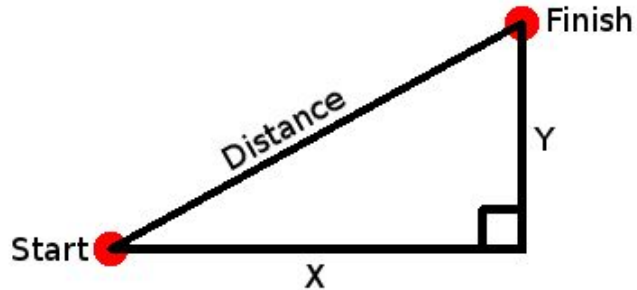
The calculated distance was set as a boundary value of 0.003 (90-100 meters) based on a GPS error of about 30 meters at 0.001 unit.

#### Two dimensions [\[edit\]](#)

In the [Euclidean plane](#), if  $\mathbf{p} = (p_1, p_2)$  and  $\mathbf{q} = (q_1, q_2)$  then the distance is given by

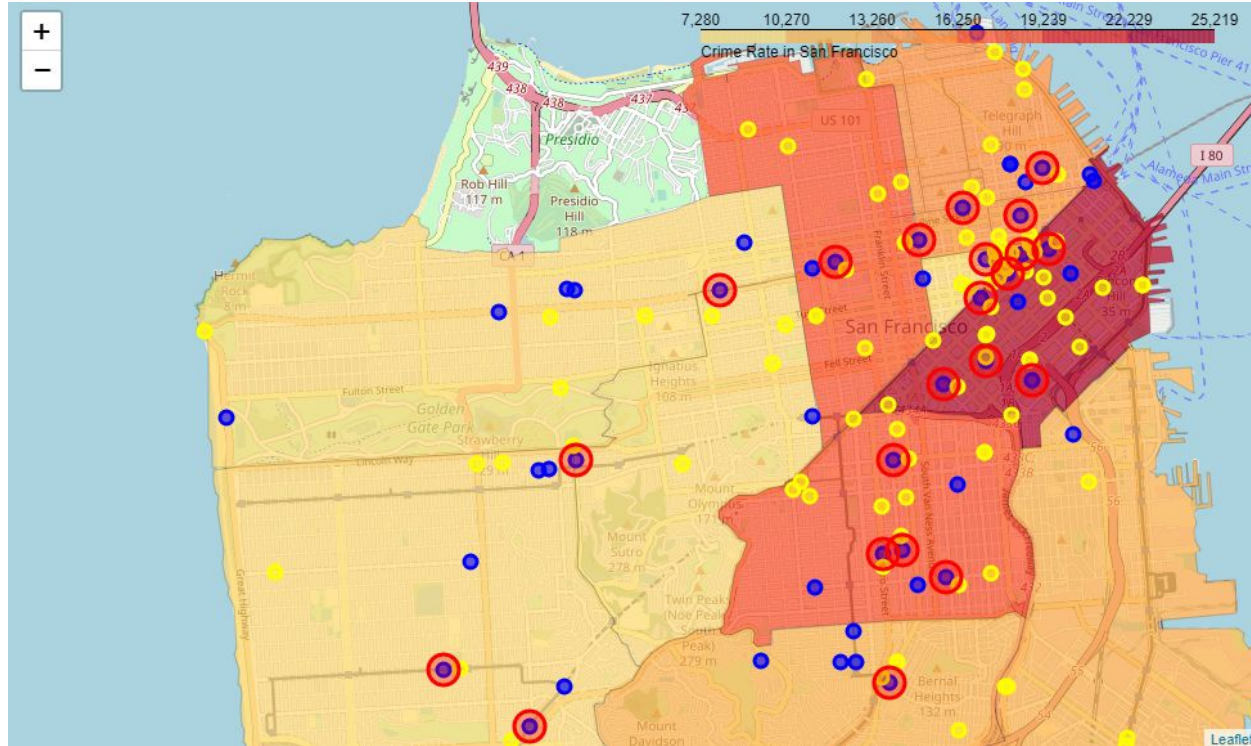
$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

This is equivalent to the [Pythagorean theorem](#).





### 3. Analysis of Data : Distance in **100 m** between the **LARCENY/THEFT Crime** and **Restaurants**



## 4. Result review

**San Francisco crimes are classified into 39 types.**

**The most common type of crime is LARCENY / THEFT.**

**More crimes occurred on Fridays and Saturdays than on weekdays**

**More crimes occurred in the afternoon than in the morning**

**The highest number of crimes occurred at 6 pm, dinner time**

**Extraction the restaurant and the location of the LARCENY / THEFT are less than 100 meters**



## 5.Conclusion

**Based on the information from these studies, tourists hope to enjoy safe restaurant visits and enjoy San Francisco more.**



- List of restaurants in the area with high risk of theft
  - Panchita's Pupuseria Restaurant #2
  - Tú Lan Restaurant
  - May Lee Restaurant
  - Tuba - Authentic Turkish Restaurant
  - Restaurant Anzu
  - Basil Thai Restaurant & Bar
  - MKT Restaurant and Bar
  - Tony's Cable Car Restaurant
  - The Restaurant Collection Under The Dome
  - Akiko's Restaurant & Sushi Bar
  - Sun Rise Restaurant
  - Heung Yuen Restaurant
  - Big 4 Restaurant
  - Don Ramon's Mexican Restaurant
  - Mars Bar & Restaurant
  - Ryoko's Japanese Restaurant & Bar
  - San Wang Restaurant
  - Clay Oven Indian Restaurant
  - Nena's Restaurant
  - Sai's Vietnamese Restaurant
  - New Ming's Restaurant
  - 8 Immortals Restaurant
- List of restaurants in the area with low risk of theft
  - San Tung Chinese Restaurant 山東小館 (San Tung Chi...)
  - Cafe Bakery & Restaurant
  - Kabuto Restaurant
  - Eric's Restaurant
  - New Eritrea Restaurant & Bar
  - Henry's Hunan Restaurant
  - The Chieftain Irish Pub & Restaurant
  - Turtle Tower Restaurant
  - Evergreen Garden Restaurant
  - Bullshead Restaurant
  - City View Restaurant
  - Pa'ina Lounge & Restaurant
  - Alice's Restaurant
  - Taiwan Restaurant 台灣飯店
  - Kamei Restaurant Supply
  - 750 Restaurant & Bar
  - Trimark Economy Restaurant Fixtures
  - Chava's Mexican Restaurant
  - Z & Y Restaurant
  - Kezar Bar & Restaurant
  - Osha Thai Restaurant & Lounge
  - One Market Restaurant
  - Pier Market Seafood Restaurant
  - Chavita's Restaurant
  - NOVY Restaurant
  - Canton Dim Sum & Seafood Restaurant
  - Kufu-ya Japanese Restaurant