



```
In [4]: import urllib
import urllib.request
import json
from pandas.io.json import json_normalize
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
!pip install folium
import folium
from IPython.display import display
```

Requirement already satisfied: folium in c:\users\lab306\anaconda3\lib\site-packages (0.11.0)  
Requirement already satisfied: numpy in c:\users\lab306\anaconda3\lib\site-packages (from folium) (1.18.1)  
Requirement already satisfied: requests in c:\users\lab306\anaconda3\lib\site-packages (from folium) (2.22.0)  
Requirement already satisfied: jinja2>=2.9 in c:\users\lab306\anaconda3\lib\site-packages (from folium) (2.11.1)  
Requirement already satisfied: branca>=0.3.0 in c:\users\lab306\anaconda3\lib\site-packages (from folium) (0.4.1)  
Requirement already satisfied: urllib3<1.25.0,!=1.25.1,<1.26,>=1.21.1 in c:\users\lab306\anaconda3\lib\site-packages (from requests->folium) (1.25.8)  
Requirement already satisfied: idna>2.9,>=2.5 in c:\users\lab306\anaconda3\lib\site-packages (from requests->folium) (2.8)  
Requirement already satisfied: certifi>=2017.4.17 in c:\users\lab306\anaconda3\lib\site-packages (from requests->folium) (2019.11.28)  
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in c:\users\lab306\anaconda3\lib\site-packages (from requests->folium) (3.0.4)  
Requirement already satisfied: MarkupSafe=>0.23 in c:\users\lab306\anaconda3\lib\site-packages (from jinja2>=2.9->folium) (1.1.1)

### ทำการเลือก API

- API จุดเดียวโดย: <https://data.police.uk/docs/method/crime-street/>
- โภคภารกิจของกรุงลอนดอนที่ประเทศไทยเดือนกรกฎาคมปีพ.ศ.2020 (JSON) <https://data.police.uk/api/crimes-street/all-crime?lat=52.629729&lng=-1.131592&date=2020-07>
- API สามารถติดต่อได้ที่: [feedback@data.police.uk](mailto:feedback@data.police.uk)

### เรียกดึงข้อมูล API

```
In [5]: #Link: URL file JSON
url = "https://data.police.uk/api/crimes-street/all-crime?lat=52.629729&lng=-1.131592&date=2020-07" #ทำการพิมพ์หรือเลือกคลิกที่ชื่อ URL
response = urllib.request.urlopen(url)
print(response)
raw_json = response.read()
#แปลงไฟล์ JSON ให้มันเป็นรูปแบบซึ่งเป็นประโยชน์ของภาษา Python และเก็บอยู่ใน data
data = json.loads(raw_json)
data
```

```
Out[5]: [{"category": "anti-social-behaviour",
  "location_type": "Force",
  "location": {"latitude": "52.627744",
    "street": {"id": 88202, "name": "On or near Mandora Lane"},
    "longitude": "-1.118922"},
  "context": "",
  "outcome_status": None,
  "persistent_id": '',
  "id": 85258564,
  "location_subtype": '',
  "month": "2020-07"}, {"category": "anti-social-behaviour",
  "location_type": "Force",
  "location": {"latitude": "52.629009",
    "street": {"id": 88345, "name": "On or near Marquis Street"},
    "longitude": "-1.132073"},
  "context": "",
  "outcome_status": None,
  "persistent_id": ''}]
```

```
In [6]: #To json_normalize ใน Pandas เพื่อสร้าง Data Frame ของข้อมูล JSON
df = json_normalize(data)
```

C:\Users\LAB306\anaconda3\lib\site-packages\ipykernel\_launcher.py:2: FutureWarning: pandas.io.json.json\_normalize is deprecated d, use pandas.json\_normalize instead

### ข้อมูลระดับการเกิดอาชญากรรมบนห้องถนนที่ประเทศไทยเดือนกรกฎาคมปีพ.ศ.2020

- Dataframe มีจำนวนrecord ที่

  - category = ประเภทอาชญากรรม
  - location\_type = ประเภทของสถานที่ที่รวมทั้ง Force หรือ BTP,Force ระบุตำแหน่งของสถานที่รวมไปด้วย BTP กองกำลังตำรวจแห่งชาติส่านานกไฟฟ สถานที่ BTP อยู่ในบังคับของกองกำลังตำรวจที่
  - context = สถานที่ที่เดินทางไปในวันอาชญากรรม (ถ้ามี)
  - id = รหัสของอาชญากรรม รหัสนี้ใช้เข้าสู่หน้า API เท่านั้นไม่ใช้ระหว่างสื่อสาร
  - month = เดือนปีของอาชญากรรม
  - location.latitude = Latitude
  - location.longitude = Longitude
  - location.street.id = รหัสของถนนที่สถานที่
  - location.street.name = ชื่อถนนที่เป็นเพียงชื่อของถนนอาชญากรรมเท่านั้น
  - outcome\_status = สถานะของอาชญากรรมที่คลอดของอาชญากรรม
  - persistent\_id = รหัสบุคคล 64 บิตของสำนวนอาชญากรรมที่นิยม
  - outcome\_status.category = สถานะของอาชญากรรมที่คลอดของอาชญากรรม
  - outcome\_status.date = วันที่ผลการบันทึกสถานะล่าสุด
  - location\_subtype = สำหรับสถานที่ BTP ประเภทของสถานที่ที่เป็นที่อยู่อาชญากรรมนั้น

```
In [7]: df #แสดงผลลัพธ์ DataFrame
```

	category	location_type	context	outcome_status	persistent_id	id	location_subtype	month	location.latitude
0	anti-social-behaviour	Force		NaN		85258564		2020-07	52.6277
1	anti-social-behaviour	Force		NaN		85258574		2020-07	52.6299
2	anti-social-behaviour	Force		NaN		85258590		2020-07	52.6428
3	anti-social-behaviour	Force		NaN		85258593		2020-07	52.6274
4	anti-social-behaviour	Force		NaN		85259287		2020-07	52.6373
...	...	...	...	...	...	...	...	...	...
1097	other-crime	Force		NaN f5b7c056c2b17571520d40309765c85766d82404dd6e9.		85265170		2020-07	52.6257
1098	other-crime	Force		NaN 6cf0c6cb73688a872db33ecb0de17c02a703e17cad...		85262079		2020-07	52.6271
1099	other-crime	Force		NaN 520fa25027db0439da587b8674bb38b9a04df5db124213...		85261664		2020-07	52.6159
1100	other-crime	Force		NaN 7076eaafb9cfbd90bd140a89245905c0b4522e2c1982d...		85265197		2020-07	52.6257
1101	other-crime	Force		NaN be71c29274a5759f752cac6091d5f8c7e72db5ba5323...		85263144		2020-07	52.6276

```

ก่อนปรับปรุงแต่งข้อมูล
• เข้ามายังใน Dataframe เป็นไฟล์CSV ชื่อไฟล์ Alter_clean_Data

]: df.to_csv("before_clean_Data.csv") #Export ข้อมูลDataFrame ที่ปรับแต่งไว้แล้ว
print(df)

          category location_type context outcome_status \
0    anti-social-behaviour     Force      NaN        NaN \
1    anti-social-behaviour     Force      NaN        NaN \
2    anti-social-behaviour     Force      NaN        NaN \
3    anti-social-behaviour     Force      NaN        NaN \
4    anti-social-behaviour     Force      NaN        NaN \
...           ...       ...      ...       ... \
1097   other-crime        Force      NaN        NaN \
1098   other-crime        Force      NaN        NaN \
1099   other-crime        Force      NaN        NaN \
1100   other-crime        Force      NaN        NaN \
1101   other-crime        Force      NaN        NaN \
                                             persistent_id   id \
0                                         85258564 \
1                                         85258574 \
2                                         85258590 \
3                                         85258593 \
4                                         85259287 \
...           ...       ...      ... \
1097   f5b7c056c2b175152d40309765c85766d82404dd6e9  85265170 \
1098   6cf0d8c6cb73688a872d33ecbd85de17c02a783e17cad...  85262079 \
1099   520f2a5027db439da507f8674b8b9a04df5db124123...  85261664 \
1100   7076eaaf9cfbd90b140a5924595c0b4522e2c1982fd...  85265197 \
1101   be71c292f74a575f9752c6091f58c7e727db5a53223...  85263144 \
                                             location_subtype month location.latitude location.street.id \
0                               2020-07      52.627744             883202 \
1                               2020-07      52.627112             883345 \
2                               2020-07      52.642857             884332 \
3                               2020-07      52.627470             883260 \
4                               2020-07      52.637327             883382 \
...           ...       ...      ...       ... \
1097   2020-07      52.625777             882372 \
1098   2020-07      52.627112             882209 \
1099   2020-07      52.615943             882342 \
1100   2020-07      52.625777             882372 \
1101   2020-07      52.627692             883257 \
                                             location.street.name location.longitude \
0                                On or near Mandora Lane            -1.118922 \
1                                On or near Marquis Street            -1.132873 \
2                                On or near Calgary Road            -1.125903 \
3                                On or near Princess Road East           -1.1244809 \
4                                On or near Camden Street            -1.127374 \
...           ...       ...      ... \
1097  On or near Further/Higer Educational Building           -1.125181 \
1098  On or near Narborough Road            -1.150477 \
1099  On or near Euston Street            -1.132600 \
1100  On or near Further/higher Educational Building           -1.125181 \
1101  On or near Prison                  -1.132545 \
                                             outcome_status.category outcome_status.date \
0                           NaN           NaN        NaN \
1                           NaN           NaN        NaN \
2                           NaN           NaN        NaN \
3                           NaN           NaN        NaN \
4                           NaN           NaN        NaN \
...           ...       ...      ... \
1097   Under investigation        2020-07 \
1098  Unable to prosecute suspect        2020-07 \
1099  Under investigation        2020-07 \
1100  Under investigation        2020-07 \
1101  Under investigation        2020-07 \

```

แสดงจำนวนเก้าของข้อมูล Missing value, Data type ของแต่ละคอลัมน์

```
In [9]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1102 entries, 0 to 1101
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype  
 ---  -- 
 0   category        1102 non-null    object  
 1   location_type   1102 non-null    object  
 2   context          1102 non-null    object  
 3   outcome_status   0 null         float64 
 4   persistent_id   1102 non-null    object  
 5   id               1102 non-null    int64  
 6   location_subtype 1102 non-null    object  
 7   month            1102 non-null    object  
 8   location.latitude 1102 non-null    object  
 9   location.street_id 1102 non-null    int64  
 10  location.street.name 1102 non-null    object  
 11  location.longitude 1102 non-null    object  
 12  outcome_status.category 934 non-null    object  
 13  outcome_status.date  934 non-null    object  
dtypes: float64(1), int64(2), object(11)
memory usage: 128.7 KB
```

ตรวจสอบพบว่ามีค่า NaN

- គម្រោង  
▪ outcome\_status  
▪ outcome\_status.category  
▪ outcome\_status.date

```
In [10]: df.isnull().sum() #ตรวจสอบจำนวนค่าNaNของแต่ละ colum
```

```
Out[10]: category          0
         location_type       0
         context            0
         outcome_status     1102
         persistent_id      0
         id                 0
         location_subtype    0
         month              0
         location.latitude   0
         location.state_id   0
         location.street_name 0
         location.longitude   0
         outcome_status.category 168
         outcome_status.date   168
         dtype: int64
```

### แสดงจำนวนประเภทอาชญากรรมแต่ละประเภท

```
In [11]: df[["category"]].value_counts() # เผยค่าจำนวนบันทึกของกลุ่มภารกิจในใจดีคนดี
```

```
Out[11]: violent-crime           365  
anti-social-behaviour        168  
public-order                  134  
criminal-damage-arson       84  
other                         61  
shoplifting                   55  
burglary                      53  
durgeling                     53
```

```

vehicle-crime      36
bicycle-theft     37
possession-of-weapons 22
robbery           19
other-crime       14
theft-from-the-person 13
Name: category, dtype: int64

```

#### แสดงจำนวนชื่อสถานที่เกิดอาชญากรรม

```
In [12]: df["location.street.name"].value_counts() #แสดงจำนวนสถานที่เกิดอาชญากรรมที่เกิดขึ้นมากที่สุด
```

location.street.name	Count
On or near Supermarket	42
On or near Petrol Station	37
On or near Parking Area	27
On or near Leicester Royal Infirmary	21
On or near Euston Street	20
..	..
On or near Regent Street	1
On or near Noble Street	1
On or near Archdeacon Lane	1
On or near Leicester Station	1
On or near Nugent Street	1

Name: location.street.name, Length: 295, dtype: int64

#### แสดงสถานะแต่ละประเภทอาชญากรรม

- มีค่า NaN อยู่จำนวน 168 ค่า

```
In [13]: df["outcome_status.category"].value_counts() #จำนวนผลการปฏิบัติสืบสวนของแต่ละประเภทอาชญากรรม
```

outcome_status.category	Count
Investigation complete; no suspect identified	386
Unable to prosecute suspect	293
Under investigation	153
Awaiting court outcome	74
Local resolution	46
Further investigation is not in the public interest	24
Formal action is not in the public interest	17
Action to be taken by another organisation	9
Offender given a caution	7
Further action is not in the public interest	3
Suspect charged as part of another case	3

Name: outcome\_status.category, dtype: int64

#### แสดงประเภทของสถานที่ตำรวจทั้ง Force หรือ BTP

- Force ระบุสถานที่ของตำรวจทั่วไปทั้งหมด
- BTP ลงทุนที่ตำรวจที่สำเร็จในไฟฟ้าสถานที่ BTP อยู่ในขอบเขตของตำรวจทั่วไปทั้งหมด

```
In [14]: df["location_type"].value_counts() #แสดงจำนวนประเภทของสถานที่ตำรวจ
```

location_type	Count
Force	1099
BTP	3

Name: location\_type, dtype: int64

#### แสดงจำนวนชื่อคอลัมน์ที่ซ้ำกันใน Dataframe

- ไม่มีจำนวนแผลงแผลงที่ซ้ำกัน

```
In [15]: df.duplicated().sum() #จำนวนชื่อคอลัมน์ที่ซ้ำกันทั้งหมด
```

Out[15]: 0

#### ลบคอลัมน์ที่ไม่ได้ใช้

- context
- outcome\_status
- persistent\_id
- location\_subtype
- outcome\_status.date

```
In [16]: df.drop(["context","outcome_status","persistent_id","location_subtype","outcome_status.date"],axis= 1, inplace=True) #ลบคอลัมน์ที่ไม่ได้ใช้
```

```
In [17]: df #แสดงรีวิวคอลัมน์ DataFrame จาก 14 คอลัมน์  จะเห็นว่าคอลัมน์ที่ไม่ได้ใช้ถูกลบออกไปแล้ว
```

category	location_type	id	month	location.latitude	location.streetId	location.street.name	location.longitude	outcome_status.category	
0	anti-social-behaviour	Force	85258564	2020-07	52.627744	883202	On or near Mandora Lane	-1.118922	NaN
1	anti-social-behaviour	Force	85258574	2020-07	52.629909	883345	On or near Marquis Street	-1.132073	NaN
2	anti-social-behaviour	Force	85258590	2020-07	52.642857	884332	On or near Calgary Road	-1.125003	NaN
3	anti-social-behaviour	Force	85258593	2020-07	52.627470	883260	On or near Princess Road East	-1.124409	NaN
4	anti-social-behaviour	Force	85259287	2020-07	52.637327	883382	On or near Camden Street	-1.127374	NaN
..	..	..	..	..	..	..	..	..	..
1097	other-crime	Force	85265170	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1098	other-crime	Force	85262079	2020-07	52.627112	882209	On or near Narborough Road	-1.150477	Unable to prosecute suspect
1099	other-crime	Force	85261664	2020-07	52.615943	882342	On or near Euston Street	-1.132600	Under investigation
1100	other-crime	Force	85265197	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1101	other-crime	Force	85263144	2020-07	52.627692	883257	On or near Prison	-1.132545	Under investigation

1102 rows x 9 columns

#### เบื้องต้นของคอลัมน์ outcome\_status.category เป็นค่า NaN

- กรณีไม่สามารถหาสถานที่ของประเภทอาชญากรรม
- หมายความว่า NaN เป็นค่าความ The value could't determined

```
In [18]: df["outcome_status.category"].fillna("The value could't determined",inplace = True) #แทนที่ค่า NaN ด้วยชื่อความThe value could't determined
```

category	location_type	id	month	location.latitude	location.streetId	location.street.name	location.longitude	outcome_status.category	
0	anti-social-behaviour	Force	85258564	2020-07	52.627744	883202	On or near Mandora Lane	-1.118922	The value could't determined
1	anti-social-behaviour	Force	85258574	2020-07	52.629909	883345	On or near Marquis Street	-1.132073	The value could't determined
2	anti-social-behaviour	Force	85258590	2020-07	52.642857	884332	On or near Calgary Road	-1.125003	The value could't determined
3	anti-social-behaviour	Force	85258593	2020-07	52.627470	883260	On or near Princess Road East	-1.124409	The value could't determined
4	anti-social-behaviour	Force	85259287	2020-07	52.637327	883382	On or near Camden Street	-1.127374	The value could't determined
..	..	..	..	..	..	..	..	..	..
1097	other-crime	Force	85265170	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation

1098	other-crime	Force	85262079	2020-07	52.627112	882209	On or near Narborough Road	-1.150477	Unable to prosecute suspect
1099	other-crime	Force	85261664	2020-07	52.615943	882342	On or near Euston Street	-1.132600	Under investigation
1100	other-crime	Force	85265197	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1101	other-crime	Force	85263144	2020-07	52.627692	883257	On or near Prison	-1.132545	Under investigation

1102 rows x 9 columns

#### เปลี่ยนชื่อคอลัมน์ให้อ่านง่าย些

```
In [19]: df_2 = df.rename(columns = {"id": "ID", "category": "Category",  
                           "location_type": "Location_type",  
                           "month": "Month",  
                           "location_street_id": "Street_id",  
                           "location_street_name": "Location_Street_name",  
                           "location.latitude": "Latitude",  
                           "location.longitude": "Longitude",  
                           "outcome_status.category": "Status_category"}) #ใช้คำสั่ง rename เพื่อเปลี่ยนชื่อคอลัมน์
```

In [20]: df\_2

Out[20]:

	Category	Location_type	ID	Month	Latitude	Street_id	Location_Street_name	Longitude	Status_category
0	anti-social-behaviour	Force	85258564	2020-07	52.627744	883202	On or near Mandora Lane	-1.118922	The value couldn't determined
1	anti-social-behaviour	Force	85258574	2020-07	52.629909	883345	On or near Marquis Street	-1.132073	The value couldn't determined
2	anti-social-behaviour	Force	85258590	2020-07	52.642857	884332	On or near Calgary Road	-1.125003	The value couldn't determined
3	anti-social-behaviour	Force	85258593	2020-07	52.627470	883260	On or near Princess Road East	-1.124409	The value couldn't determined
4	anti-social-behaviour	Force	85259287	2020-07	52.637327	883382	On or near Camden Street	-1.127374	The value couldn't determined
...	...	...	...	...	...	...	...	...	...
1097	other-crime	Force	85265170	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1098	other-crime	Force	85262079	2020-07	52.627112	882209	On or near Narborough Road	-1.150477	Unable to prosecute suspect
1099	other-crime	Force	85261664	2020-07	52.615943	882342	On or near Euston Street	-1.132600	Under investigation
1100	other-crime	Force	85265197	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1101	other-crime	Force	85263144	2020-07	52.627692	883257	On or near Prison	-1.132545	Under investigation

1102 rows x 9 columns

#### แสดงจำนวนข้อมูลแກ้ว่าซ้ำกันแล้วก็ลบไป

- อย่างเช่นจะลบ掉ที่ไม่ซ้ำกันทุกอย่างที่มีความแตกต่างกัน

```
In [21]: len(df_2[df_2.duplicated(keep = False)]) #เมื่อเราลบซ้ำกันแล้วก็จะเหลือแค่ 0 รายการ
```

Out[21]: 0

#### แสดงจำนวนค่าNaNแต่ละคอลัมน์แล้วจำกัดผลลัพธ์ของแต่ละคอลัมน์

- แทนที่ NaN ด้วยคำสั่ง Status\_category ถ้าซ้ำกัน The value couldn't determined
- หากไม่มีค่าใน Dataframe ให้เป็น NaN

```
In [22]: df_2.isnull().sum() #เมื่อเราลบซ้ำกันแล้วก็จะเหลือแค่ 0 รายการ
```

Out[22]: Category 0  
Location\_type 0  
ID 0  
Month 0  
Latitude 0  
Street\_id 0  
Location\_Street\_name 0  
Longitude 0  
Status\_category 0  
dtype: int64

#### หลังจากปรับแต่งข้อมูลเสร็จ

- export ที่บันทึก Dataframe เป็นไฟล์ CSV ชื่อไฟล์ After\_clean\_Data

```
In [23]: df_2.to_csv("After_clean_Data.csv") #Export Dataframe to file.csv  
print(df_2)
```

	Category	Location_type	ID	Month	Latitude	Street_id	Location_Street_name	Longitude	Status_category
0	anti-social-behaviour	Force	85258564	2020-07	52.627744	883202	On or near Mandora Lane	-1.118922	The value couldn't determined
1	anti-social-behaviour	Force	85258574	2020-07	52.629909	883345	On or near Marquis Street	-1.132073	The value couldn't determined
2	anti-social-behaviour	Force	85258590	2020-07	52.642857	884332	On or near Calgary Road	-1.125003	The value couldn't determined
3	anti-social-behaviour	Force	85258593	2020-07	52.627470	883260	On or near Princess Road East	-1.124409	The value couldn't determined
4	anti-social-behaviour	Force	85259287	2020-07	52.637327	883382	On or near Camden Street	-1.127374	The value couldn't determined
...	...	...	...	...	...	...	...	...	...
1097	other-crime	Force	85265170	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1098	other-crime	Force	85262079	2020-07	52.627112	882209	On or near Narborough Road	-1.150477	Unable to prosecute suspect
1099	other-crime	Force	85261664	2020-07	52.615943	882342	On or near Euston Street	-1.132600	Under investigation
1100	other-crime	Force	85265197	2020-07	52.625777	882372	On or near Further/higher Educational Building	-1.125181	Under investigation
1101	other-crime	Force	85263144	2020-07	52.627692	883257	On or near Prison	-1.132545	Under investigation

[1102 rows x 9 columns]

#### แสดงจำนวนแต่ละประเภทของอาชญากรรม

```
In [24]: count_category = df_2["Category"].value_counts() #นับจำนวนของประเภทของอาชญากรรมในแต่ละคอลัมน์ Category
```

Out[24]: violent-crime 365  
anti-social-behaviour 168

```

public-order          134
criminal-damage-arson   84
other-theft           61
shoplifting            53
burglary               53
drugs                  52
vehicle-crime         37
bicycle-theft          27
possession-of-weapons    22
robbery                 19
other-crime             14
theft-from-the-person     13
Name: Category, dtype: int64

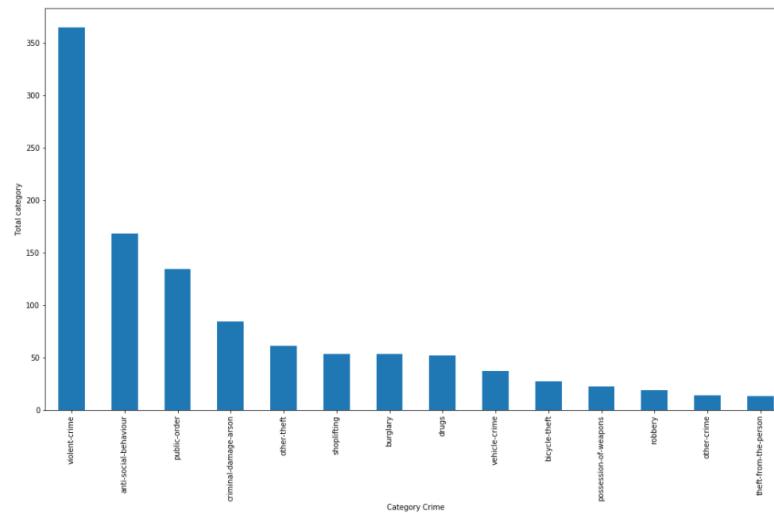
```

#### แสดงกราฟ bar chart ประเภทดับการเกิดอาชญากรรมทั้งหมดบนห้องถนนที่ประเทศไทยเดือนกรกฎาคม ปี พ.ศ.2020

- จำนวนที่เกิดอาชญากรรมมากที่สุดคืออาชญากรรมประเภท violent-crime 365 ครั้ง
- จำนวนที่เกิดอาชญากรรมน้อยที่สุดคืออาชญากรรมประเภท theft-from-the-person 13 ครั้ง

```
In [25]: my_plot = count_category.plot(kind="bar", figsize=(18, 10)) #plt.count_category ภาพลักษณ์กราฟบาร์ chart
my_plot.set_xlabel("Category Crime") #กราฟจะมีตัวบ่งบอก X
my_plot.set_ylabel("Total category") #กราฟจะมีตัวบ่งบอก Y
```

```
Out[25]: Text(0, 0.5, 'Total category')
```



(เมื่อจากอาชญากรรมประเภท violent-crime อาชญากรรมทุกประเภทเกิดขึ้นเยอะที่สุด และดูว่าอาชญากรรมประเภท violent-crime เกิดขึ้นที่สถานที่ในเมืองที่สูงที่สุด 10 อันดับคน)

```
In [26]: lomax_cat = df_2[df_2['Category'] == 'violent-crime']['Location_Street_name'].value_counts().head(10) #เพื่อ category มากที่สุด violent
lomax_cat
```

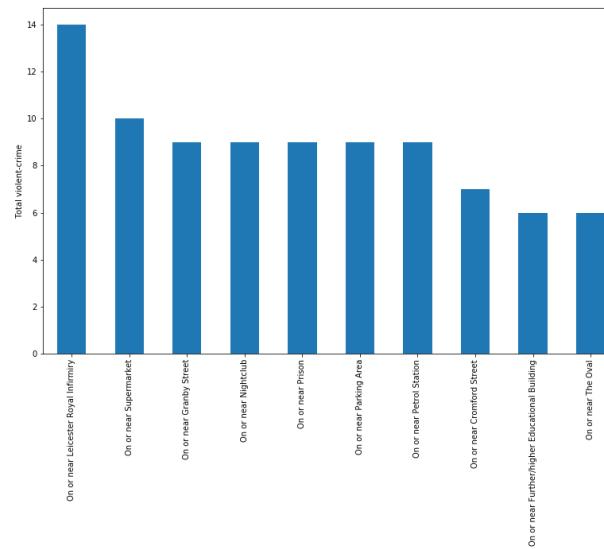
```
Out[26]: On or near Leicester Royal Infirmary      14
On or near Supermarket                         10
On or near Granby Street                      9
On or near Nightclub                           9
On or near Prison                            9
On or near Parking Area                       9
On or near Petrol Station                     9
On or near Cromford Street                   7
On or near Further/higher Educational Building 6
On or near The Oval                           6
Name: Location_Street_name, dtype: int64
```

#### แสดงกราฟ bar chart ของสถานที่ที่เกิดอาชญากรรมประเภท violent-crime มากที่สุด 10 อันดับ บนห้องถนนที่ประเทศไทยเดือนกรกฎาคม ปี พ.ศ. 2020

- สถานที่ที่เกิดอาชญากรรมมากที่สุดใน 10 อันดับคือ On or near Leicester Royal Infirmary 14 ครั้ง
- สถานที่ที่เกิดอาชญากรรมน้อยที่สุดใน 10 อันดับคือ On or near The Oval, On or near Further/higher Educational Building 6 ครั้ง

```
In [27]: my_plot = lomax_cat.head(10).plot(kind="bar", figsize=(13, 8)) #นำสถานที่เกิดอาชญากรรมประเภท violent-crime มากที่สุด 10 อันดับมาภาพลักษณ์กราฟบาร์
my_plot.set_xlabel("Location Name") #กราฟจะมีตัวบ่งบอก X
my_plot.set_ylabel("Total violent-crime") #กราฟจะมีตัวบ่งบอก Y
```

```
Out[27]: Text(0, 0.5, 'Total violent-crime')
```



```
In [28]: count_status = df_2["Status_category"].value_counts() #นับจำนวนสถานะของคดี
count_status
```

Status Category	Count
Investigation complete; no suspect identified	306
Unable to prosecute suspect	292
The value couldn't be determined	168
Under investigation	153
Awaiting court outcome	74
Local resolution	46
Further investigation is not in the public interest	24
Formal action is not in the public interest	17
Action to be taken by another organisation	9
Offender given a caution	7
Further action is not in the public interest	3
Suspect charged as part of another case	3

Name: Status\_category, dtype: int64

แสดงจำนวนประเภทอาชญากรรมที่มีสถานะ Investigation complete; no suspect identified

- อาชญากรรมที่มีสถานะการสอบสวนเสร็จสมบูรณ์และไม่พบผู้ต้องสงสัย

```
In [29]: status_cat = df_2[df_2["Status_category"] == 'Investigation complete; no suspect identified'][['Category']].value_counts() #นับจำนวน
status_cat
```

Category	Count
violent-crime	56
public-order	49
criminal-damage-arson	38
shoplifting	36
other-theft	36
burglary	31
vehicle-crime	29
bicycle-theft	24
theft-from-the-person	7
robbery	6
possession-of-weapons	2
drugs	1

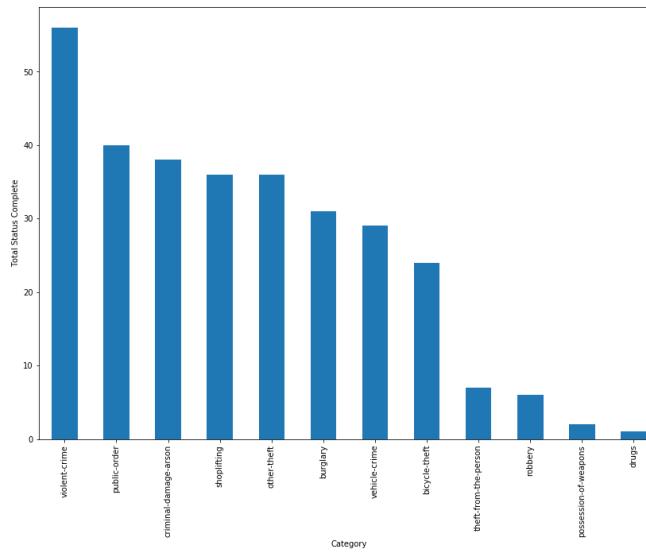
Name: Category, dtype: int64

แสดงกราฟ bar chart ประเภทอาชญากรรมที่มีสถานะการสอบสวนที่เสร็จสมบูรณ์และไม่พบผู้ต้องสงสัย

- ประเภทอาชญากรรมที่มีสถานะการสอบสวนเสร็จสมบูรณ์และที่สุดคือ violent-crime 56 ครั้ง
- ประเภทอาชญากรรมที่มีสถานะการสอบสวนเสร็จสมบูรณ์น้อยที่สุดคือ drugs 1 ครั้ง

```
In [30]: my_plot = status_cat.plot(kind="bar", figsize=(14, 10)) #พื้นที่กราฟ bar chart และจำนวนประเภทอาชญากรรมที่มีสถานะการสอบสวนที่เสร็จสมบูรณ์และไม่พบผู้ต้องสงสัย
my_plot.set_xlabel("Category") #กำหนดชื่อแกน x
my_plot.set_ylabel("Total Status Complete") #กำหนดชื่อแกน y
```

Out[30]: Text(0, 0.5, 'Total Status Complete')



แสดงจำนวนสถานที่ที่เกิดอาชญากรรมมากที่สุด 10 อันดับ

```
In [31]: count_lname = df_2["Location_Street_name"].value_counts() #นับจำนวนสถานที่หรือถนนที่เกิดอาชญากรรม
count_lname.head(10) #แสดงเฉพาะ 10 อันดับสถานที่ที่มีจำนวนที่เกิดอาชญากรรมมากที่สุด
```

Location Street Name	Count
On or near Supermarket	42
On or near Petrol Station	37
On or near Parking Area	27
On or near Leicester Royal Infirmary	21
On or near Euston Street	20
On or near Nightclub	19
On or near Granby Street	15
On or near Tudor Road	14
On or near Prison	14
On or near Further/higher Educational Building	14

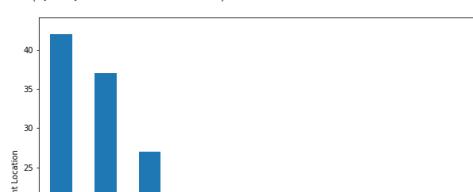
Name: Location\_Street\_name, dtype: int64

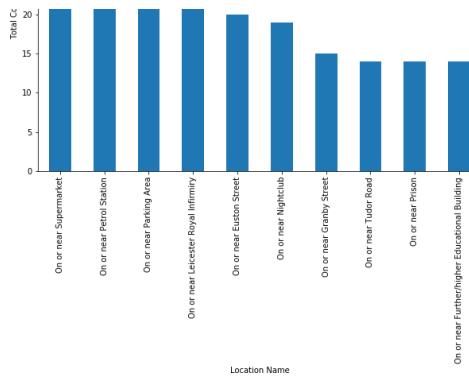
แสดงกราฟ bar chart สถานที่หรือถนนที่เกิดอาชญากรรมมากที่สุด 10 อันดับบนห้องถนนที่ประเทศไทย

- สถานที่ที่เกิดอาชญากรรมมากที่สุดใน 10 อันดับคือ On or near Supermarket 42 ครั้ง
- สถานที่ที่เกิดอาชญากรรมน้อยที่สุดใน 10 อันดับคือ On or near Further/higher Educational Building, On or near Tudor Road, On or near Prison 14 ครั้ง

```
In [32]: my_plot = count_lname.head(10).plot(kind="bar", figsize=(10, 8)) #นำจำนวนสถานที่ที่เกิดอาชญากรรมมากที่สุด 10 อันดับมาสร้างกราฟบาร์
my_plot.set_xlabel("Location Name") #กำหนดชื่อแกน x
my_plot.set_ylabel("Total Count Location") #กำหนดชื่อแกน y
```

Out[32]: Text(0, 0.5, 'Total Count Location')





เมืองจาก On or near Supermarket เป็นสถานที่ที่เกิดอาชญากรรมมากที่สุด 42 ครั้ง จากกราฟด้านบน

- ซึ่งบ้านเดียวที่คุณจะพบมากที่สุดที่ On or near Supermarket

```
In [33]: location_born_ct = df_2[df_2['Location_Street_name'] == 'On or near Supermarket']['Category'].value_counts() #ผลลัพธ์จำนวนบ้านที่อยู่ใกล้ Supermarket
location_born_ct
```

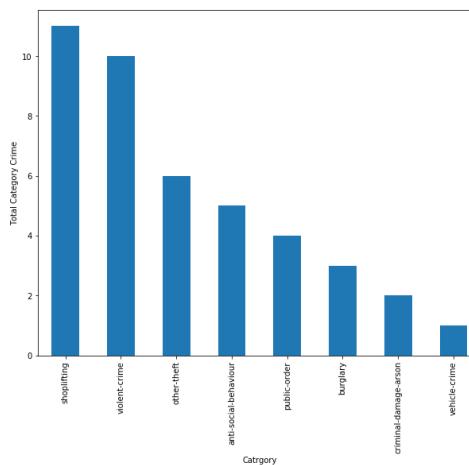
```
Out[33]:
shoplifting      11
violent-crime   10
other-theft      6
anti-social-behaviour  5
public-order     4
burglary         3
criminal-damage-arson  2
vehicle-crime    1
Name: Category, dtype: int64
```

แสดงกราฟ bar chart จำนวนประเภทของอาชญากรรมที่เกิดขึ้นที่ On or near Supermarket บนห้องถนนที่ประทศอังกฤษเดือนกรกฎาคม ปี พ.ศ. 2020

- ประเภทอาชญากรรมที่เกิดขึ้นมากที่สุดคือ shoplifting 11 ครั้ง
- ประเภทอาชญากรรมที่เกิดขึ้นน้อยที่สุดคือ vehicle-crime 1 ครั้ง

```
In [34]: my_plot = location_born_ct.plot(kind="bar",figsize=(10, 8)) #ภาพ柱状圖 Location_born_ct หวานด้วยกราฟบาร์
my_plot.set_xlabel("Category") #กำหนดชื่อแก้ไข x
my_plot.set_ylabel("Total Category Crime") #กำหนดชื่อแก้ไข y
```

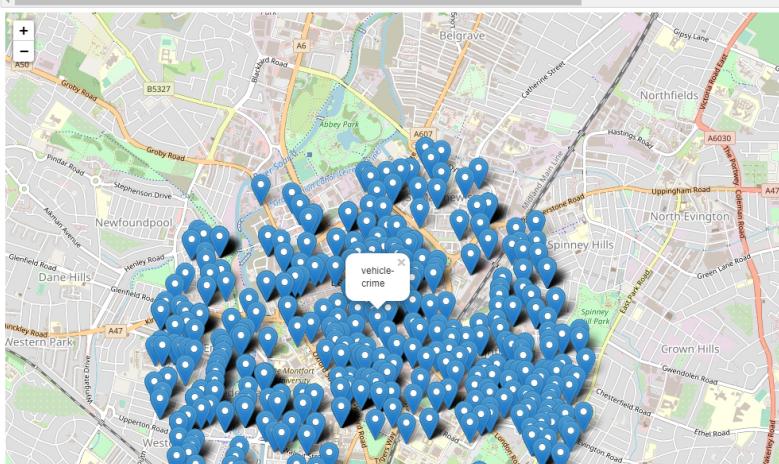
```
Out[34]: Text(0, 0.5, 'Total Category Crime')
```



แสดงกราฟที่การเกิดอาชญากรรมที่ประเภทอังกฤษเดือนกรกฎาคม ปี พ.ศ. 2020

- ระบุจุดที่เกิดเหตุอาชญากรรม
- แสดงประเภทของอาชญากรรมที่ลงบนเมืองคือคลิกที่ปุ่มMark

```
In [35]: mapp = folium.Map(location=[52.638452, -1.128697],zoom_start=14) #แสดงรูปที่ตั้วภาพรวมของที่ที่เกิดเหตุที่เกิดอาชญากรรมและใช้เวลาที่zoom_start
tooltip = 'Click me!'
for i in range(len(df_2)): #ทำการตั้วปุ่มลงบนที่ที่เกิดเหตุใน Dataframe
    #ทำการตั้วปุ่มลงบนที่ที่เกิดเหตุใน latitude,Longitude ที่มีอยู่ใน df_2 และใช้ปุ่มนี้ category เมื่อกดปุ่มนี้แล้วล็อกที่เมืองรวมและตั้วปุ่มมาที่ที่เกิดเหตุ
    folium.Marker([df_2['Latitude'].iloc[i], df_2['Longitude'].iloc[i]], popup=df_2['Category'].iloc[i], tooltip=tooltip).add_to(mapp) #ทำให้มันตั้ว
display(mapp) #ทำให้มันแสดง
```



แสดง ratio ของแต่ละประเภทอาชญากรรม

```
In [36]: category_ratio = df_2[\"Category\"].value_counts(normalize = True) # คำนวณอัตราของแต่ละประเภทของ Category มากที่สุดคราวน์
category_ratio
```

```
Out[36]: violent-crime      0.331216
anti-social-behaviour    0.152450
public-order              0.121597
criminal-damage-arson   0.076225
other-theft               0.055354
shoplifting                0.048894
burglary                  0.048894
drugs                      0.047187
vehicle-crime             0.033575
bicycle-theft              0.024501
possession-of-weapons     0.019964
robbery                    0.017241
other-crime                0.012704
theft-from-the-person      0.011797
Name: Category, dtype: float64
```

```
In [37]: category_ratio = df_2[\"Category\"].value_counts(normalize = True).reset_index(name=\"ratio\") # คำนวณและบันทึกอัตราของแต่ละประเภทของ Category
total_ratio = category_ratio[\"ratio\"]*0.05, \"index\"]
# คำนวณอัตราของแต่ละประเภทของ Category และบันทึกในตัวแปร total_ratio
df_2.loc[df_2[\"Category\"].isin(total_ratio), \"Category\"] = \"Other Offenses\" # กำหนดชื่อ Category ใหม่ให้เป็นชื่อ \"Other Offenses\"
df_2[\"Category\"].value_counts(normalize = True) # คำนวณอัตราของแต่ละประเภทของ Category
```

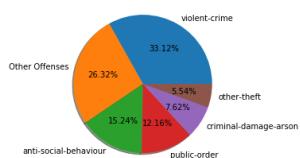
```
Out[37]: violent-crime      0.331216
Other Offenses            0.263158
anti-social-behaviour    0.152450
public-order              0.121597
criminal-damage-arson   0.076225
other-theft               0.055354
Name: Category, dtype: float64
```

แสดงกราฟ Pie chart อัตราการเกิดอาชญากรรมของแต่ละประเภท

โดยในประเภทที่มีอัตราการเกิดน้อยกว่า 5% นำไปรวมอยู่ใน Other Offenses

- อาชญากรรมที่มีอัตราการเกิดมาที่สุดคือ violent-crime 33.12%
  - อาชญากรรมที่มีอัตราการเกิดน้อยที่สุดคือ other-theft 5.54%

```
In [38]: pie_cat = ['violent-crime','Other Offenses',"anti-social-behaviour",'public-order ','criminal-damage-arson','other-theft'] #ปัจจุบัน ratio = [0.332126, 0.263158, 0.152450, 0.121597, 0.076225, 0.055354] #ratio ของแต่ละประเภทอาชญากรรม fig, ax1 = plt.subplots(figsize=(10, 6)) ax1.pie(ratio, labels= pie_cat, autopct='%.2F%%', shadow=True) #เพื่อกราฟจะเป็นรูปทรงพิเศษ chart ให้ดูง่ายมากขึ้นการใช้ratioเป็นค่าที่ต้องมีผลลัพธ์
```



ស្រុបល