

Terrorism in India (1979-2015) [A data visualization]

- Uday Kamal

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from collections import defaultdict
import geopandas as gpd
```

```
In [2]: df=pd.read_csv('indiadata.csv',encoding="ISO-8859-1")
df.head()
```

iyear	city	country_txt	latitude	longitude	attacktype1_txt	targtype1_txt	targsubtype1_txt	target
1975	Samastipur	India	25.863042	85.781004	Bombing/Explosion	Government (General)	Government Personnel (excluding police, military)	La Naraya Mish and legislative
1997	Unknown	India	33.778175	76.576171	Bombing/Explosion	Transportation	Bus Station/Stop	A bus station Kashmir
1997	Dhalai district	India	23.846698	91.909924	Bombing/Explosion	Military	Military Unit/Patrol/Convoy	Border Patrol Guard

```
In [3]: df = df.loc[:, ~df.columns.str.contains('^Unnamed')]
```

```
In [4]: df.head()
```

```
Out[4]:
```

	iyear	city	country_txt	latitude	longitude	attacktype1_txt	targtype1_txt	target1
0	1975	Samastipur	India	25.863042	85.781004	Bombing/Explosion	Government (General)	Personnel
1	1997	Unknown	India	33.778175	76.576171	Bombing/Explosion	Transportation	Bus
2	1997	Dhalai district	India	23.846698	91.909924	Bombing/Explosion	Military	Unit/Force

```
In [5]: df.columns
```

```
Out[5]: Index(['iyear', 'city', 'country_txt', 'latitude', 'longitude',  
              'attacktype1_txt', 'targtype1_txt', 'targsubtype1_txt', 'target1',  
              'weaptype1_txt', 'weapsubtype1_txt', 'gname', 'motive', 'summary'],  
              dtype='object')
```

```
In [6]: df.dtypes
```

```
Out[6]: iyear          int64  
city              object  
country_txt       object  
latitude          float64  
longitude         float64  
attacktype1_txt   object  
targtype1_txt     object  
targsubtype1_txt  object  
target1          object  
weaptype1_txt     object  
weapsubtype1_txt  object  
gname            object  
motive           object  
summary          object  
dtype: object
```

```
In [7]: df.isna().sum()
```

```
Out[7]: iyear          0
        city          0
        country_txt   0
        latitude      0
        longitude     0
        attacktype1_txt 0
        targtype1_txt 0
        targsubtype1_txt 0
        target1       0
        weaptype1_txt  0
        weapsubtype1_txt 0
        gname         0
        motive        0
        summary       0
        dtype: int64
```

The raw description of the dataset

```
In [8]: df.describe()
```

```
Out[8]:
```

	iyear	latitude	longitude
count	4972.000000	4972.000000	4972.000000
mean	2008.215607	25.197249	84.078899
std	4.349748	5.372288	6.761981
min	1975.000000	8.180270	69.612516
25%	2006.000000	22.282182	77.267544
50%	2009.000000	24.789854	84.476326
75%	2011.000000	26.746018	90.443986
max	2015.000000	34.606944	95.900036

The following dataset contains the information of cities got attacked from respective group corresponding to the year

```
In [9]: df_year=df.groupby(['iyear'])['city','gname'].sum()
```

In [10]: df_year

Out[10]:

iyear	city		gname
1975	Samastipur		Ananda Marga
1997	UnknownDhalai district	Hyderabad	Muslim RebelsNational Liberation Front of Trip...
1998	New DelhiWandhama	Nalbari districtCoimbatoreCoi...	UnknownUnknownUnited Liberation Front of Assam...
1999	PattanManoharpur	AllahabadNarayanpurTuilaphaiBh...	UnknownVishwa Hindu Parishad (VHP)Vishwa Hindu...
2000	UdhampurKakotibari	SrinagarNew DelhiUnknownKagh...	Lashkar-e-Taiba (LeT)United Liberation Front o...
2001	KeshpurPoonch	SrinagarDhalai districtSoporeDhal...	Communist Party of India- MarxistHizbul Mujahi...
2002	BandiporaYaripora	PathankotPoonch DistrictSrina...	UnknownUnknownUnknownLashkar-e-Taiba (LeT)Unkn...

The following data shows the number of attacks in each year

In [11]: year=df['iyear'].value_counts()
year

Out[11]:

2009	671
2010	655
2011	641
2008	508
2013	264
2015	243
2001	234
2014	234
2012	224
2003	196
2002	181
2000	179
2006	166
2007	149
2005	143
1999	111
2004	108
1998	61
1997	3
1975	1

Name: iyear, dtype: int64

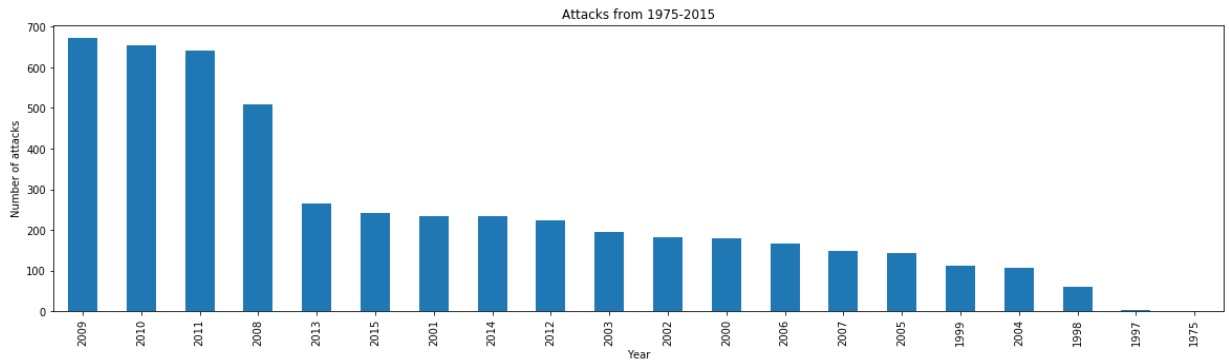
In [12]: print('Year with maximum number of attacks:',year.idxmax())
print('Number of maximum attacks in that year:',year.max())

Year with maximum number of attacks: 2009
Number of maximum attacks in that year: 671

This is a bar graph showing the trend of attacks in each year

```
In [13]: year.plot(kind='bar',figsize=(20,5))
plt.xlabel('Year')
plt.ylabel('Number of attacks')
plt.title('Attacks from 1975-2015')
#plt.ylim(0,20)
```

Out[13]: Text(0.5, 1.0, 'Attacks from 1975-2015')



Following data shows the list of top-10 cities attacked from 1979 to 2015.

```
In [14]: city=df['city'].value_counts()[:10]
city
```

```
Out[14]: Imphal          230
Srinagar             222
Unknown              137
Guwahati              61
Sopore                40
New Delhi             38
Latehar district      37
West Midnapore district 36
Malkangiri district   34
Anantnag              32
Name: city, dtype: int64
```

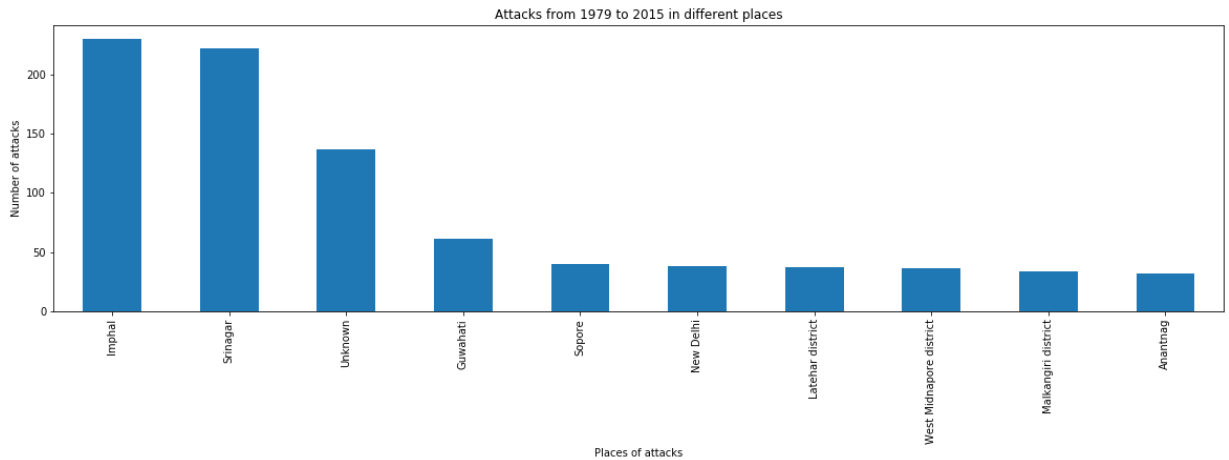
```
In [15]: print('Place with maximum number of attacks:',city.idxmax())
print('Number of attacks in that city:',city.max())
```

```
Place with maximum number of attacks: Imphal
Number of attacks in that city: 230
```

This is a bar graph showing the trend of top-10 attacks in these cities.

```
In [16]: city.plot(kind='bar',figsize=(20,5))
plt.xlabel('Places of attacks')
plt.ylabel('Number of attacks')
plt.title('Attacks from 1979 to 2015 in different places')
```

Out[16]: Text(0.5, 1.0, 'Attacks from 1979 to 2015 in different places')



Following contains the dataset with information of year of attack and the city corresponding to the group responsible for it.

```
In [17]: df_group=df.groupby(['gname'])['iyear','city'].max()
df_group[:10]
```

Out[17]:

	iyear	city
gname		
A'chik Matgrik Elite Force (AMEF)	2015	Wageasi
Achik National Cooperative Army (ANCA)	2014	Siju
Achik National Liberation Army (ANLA)	2015	Wageasi
Achik National Volunteer Council (ANVC)	2001	Songsak
Achik National Volunteer Council-B (ANVC-B)	2015	Sasatgre
Achik Songna An'pachakgipa Kotok (ASAK)	2015	Rongge
Achik Tiger Force	2014	Tura
Adivasi Cobra Militants of Assam (ACMA)	2011	Kokrajhar
Adivasi National Liberation Army (ANLA)	2014	Udalguri
Adivasi People's Army (APA)	2013	Kokrajhar district

```
In [18]: len(df['city'].unique())
```

```
Out[18]: 2228
```

```
In [19]: len(df['gname'].unique())
```

```
Out[19]: 158
```

This is a list containing the information of number of top-10 attacks done by each group.

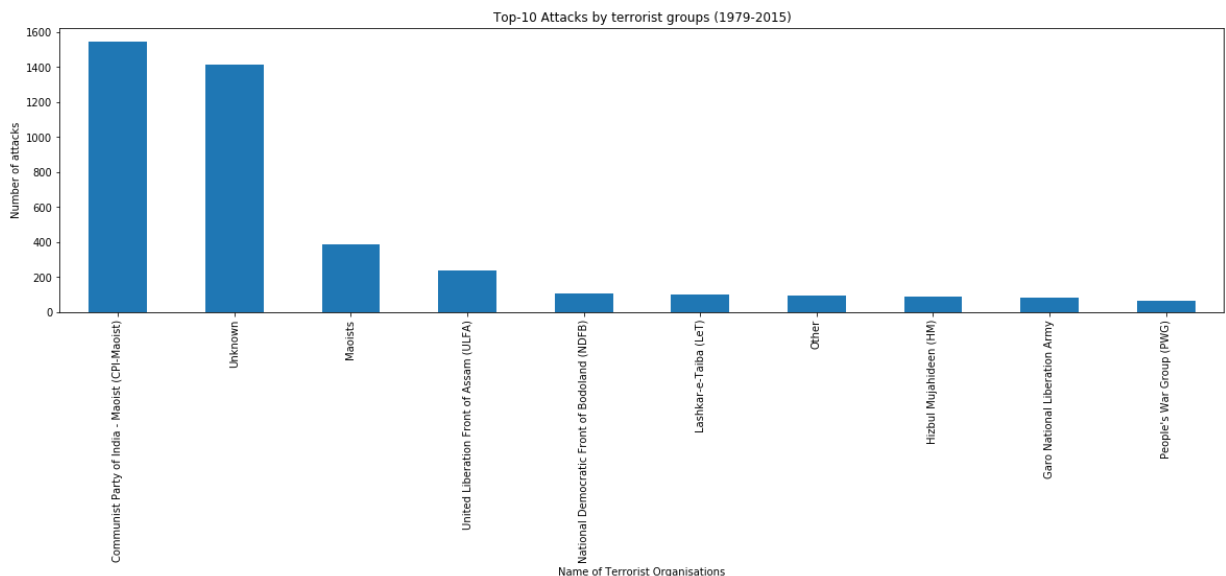
```
In [67]: group=df['gname'].value_counts()[:10]
group
```

```
Out[67]: Communist Party of India - Maoist (CPI-Maoist)    1547
Unknown                                                  1415
Maoists                                                  385
United Liberation Front of Assam (ULFA)                 238
National Democratic Front of Bodoland (NDFB)           106
Lashkar-e-Taiba (LeT)                                   100
Other                                                    92
Hizbul Mujahideen (HM)                                  88
Garo National Liberation Army                           82
People's War Group (PWG)                               62
Name: gname, dtype: int64
```

Bar graph containing the trend of attacks by each group.

```
In [21]: group.plot(kind='bar',figsize=(20,5))
plt.xlabel('Name of Terrorist Organisations')
plt.ylabel('Number of attacks')
plt.title('Top-10 Attacks by terrorist groups (1979-2015)')
```

```
Out[21]: Text(0.5, 1.0, 'Top-10 Attacks by terrorist groups (1979-2015)')
```

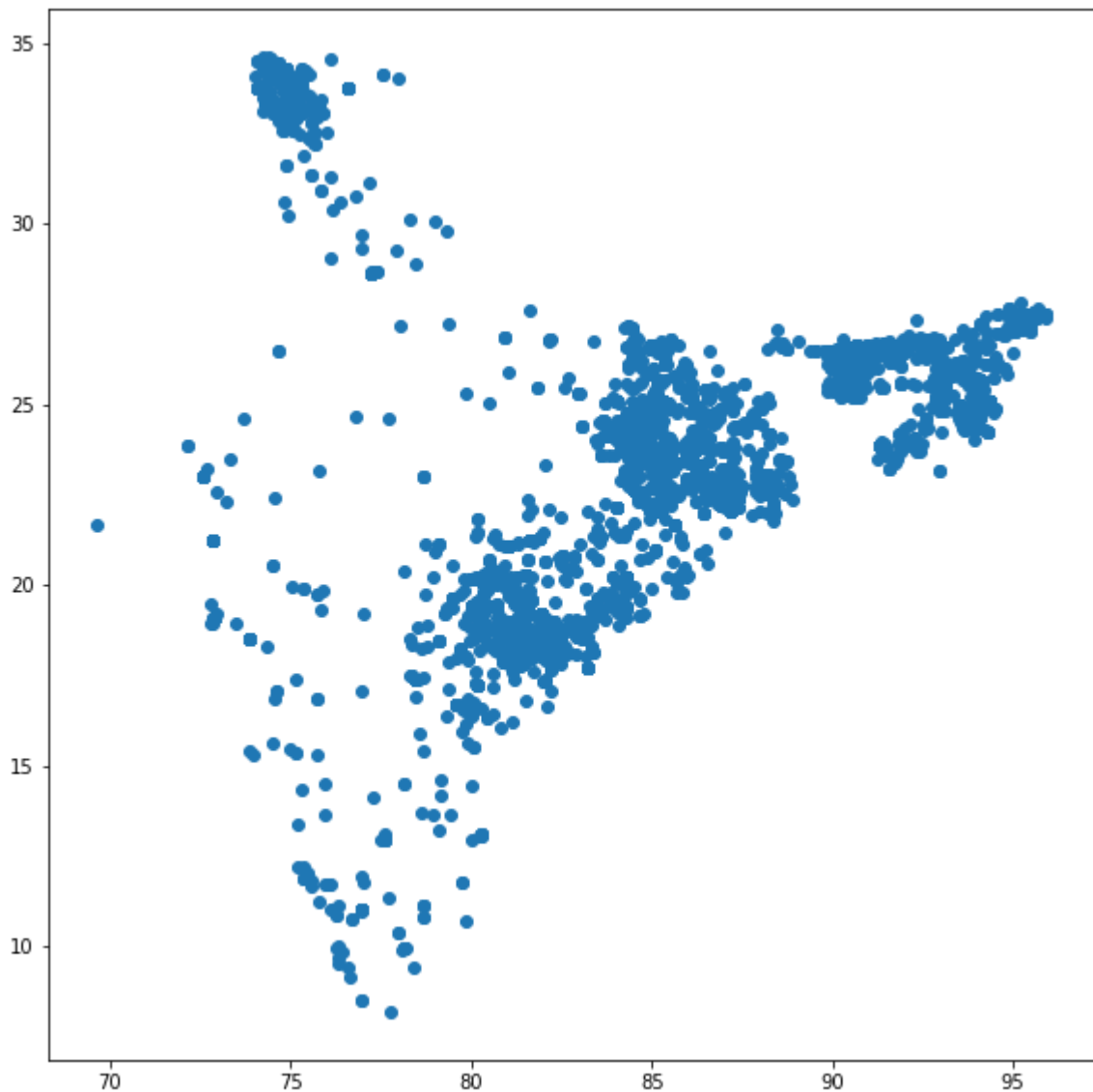


```
In [22]: gdf01=gpd.GeoDataFrame(df,geometry=gpd.points_from_xy(df['longitude'],df['latitude'],gdf01)
gdf01
#this is a geo dataframe
```

...

```
In [23]: gdf01.plot(figsize=(20,10))
#Now we have to overlap this data in the map of india.
```

```
Out[23]: <matplotlib.axes._subplots.AxesSubplot at 0x1d22f4c6848>
```



```
In [24]: gpd.datasets
```

```
Out[24]: <module 'geopandas.datasets' from 'C:\\Users\\RAVI\\Anaconda3\\lib\\site-packages\\geopandas\\datasets\\__init__.py'>
```



```
In [25]: world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))
ax=world.plot(figsize=(20,10))
ax.axis('off')
```

...

```
In [28]: fig,ax=plt.subplots(figsize=(20,10))
gdf01.plot(cmap='Greens',ax=ax)
world.boundary.geometry.plot(color=None,edgecolor='k',ax=ax)
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.title('Terrorist attacks in India (1979-2015)')
```

...

```
In [29]: asia=world[world['continent']=='Asia']
asia
```

...

```
In [30]: India=asia[asia['name']=='India']
```

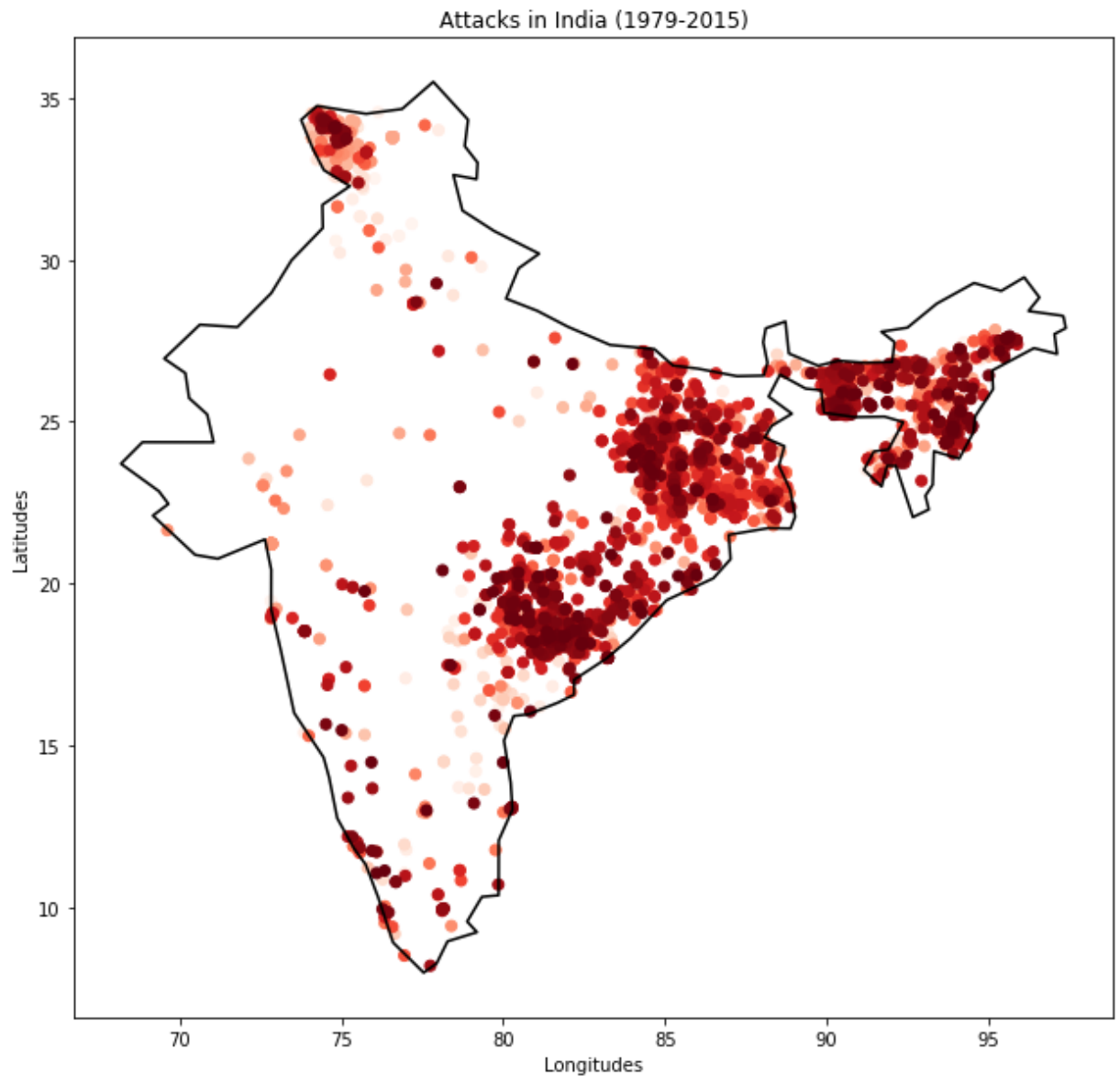
```
In [31]: a=gdf01[gdf01['country_txt']=='India']
```

```
In [32]: asia=world[world['continent']=='Asia']
```

Following scatter plot shows the distribution of attacks throughout the country from 1979 to 2015.

```
In [34]: fig,ax=plt.subplots(figsize=(20,10))
a.plot(cmap='Reds',ax=ax)
India.boundary.geometry.plot(color=None,edgecolor='k',ax=ax)
plt.xlabel('Longitudes')
plt.ylabel('Latitudes')
plt.title('Attacks in India (1979-2015)')
```

```
Out[34]: Text(0.5, 1, 'Attacks in India (1979-2015)')
```



```
In [38]: print('Types of wepaons used :',df['weaptype1_txt'].unique())
print(len(df['weaptype1_txt'].unique()))
```

```
Types of wepaons used : ['Explosives/Bombs/Dynamite' 'Melee' 'Firearms' 'Incend
iary' 'Unknown'
'Chemical' 'Sabotage Equipment'
'Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)']
8
```

Following is the list containing the number of times each of the following weapon was used.

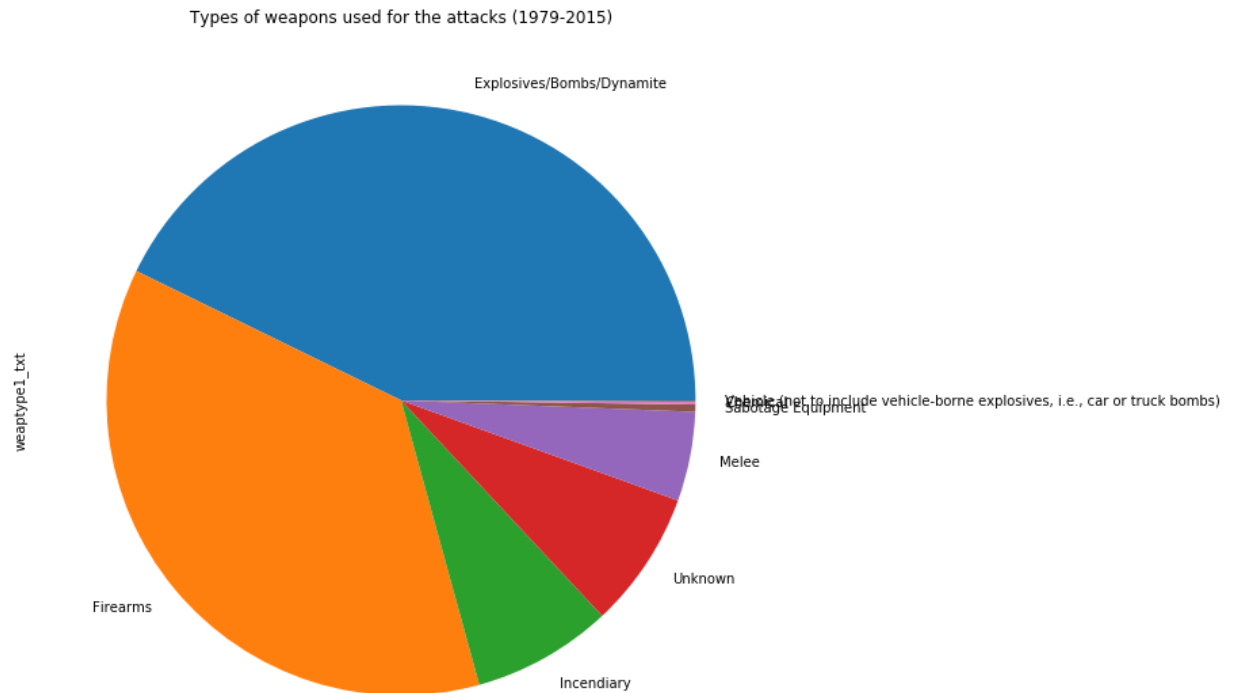
```
In [39]: weap=df['weaptype1_txt'].value_counts()
weap
```

```
Out[39]: Explosives/Bombs/Dynamite
2127
Firearms
1815
Incendiary
382
Unknown
374
Melee
243
Sabotage Equipment
20
Chemical
7
Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)
4
Name: weaptype1_txt, dtype: int64
```

Pie chart containing the visualization of the above list.

```
In [40]: weap.plot(kind='pie',figsize=(20,10))  
plt.title('Types of weapons used for the attacks (1979-2015)')
```

Out[40]: Text(0.5, 1.0, 'Types of weapons used for the attacks (1979-2015)')



```
In [41]: df1=df[df['iyear']==2015]
```

This dataframe contains information of attacks limited to only 2015.

In [42]: df1

iyear	city	country_txt	latitude	longitude	attacktype1_txt	targettype1_txt	targetsubtype1
2015	Kottapalli	India	18.832981	80.776640	Armed Assault	Private Citizens & Property	Lab (General)/Occupied
2015	Kannur district	India	12.166372	75.334212	Facility/Infrastructure Attack	Business	Construction
2015	Keriaguda	India	17.985198	81.582525	Hostage Taking (Kidnapping)	Private Citizens & Property	Lab (General)/Occupied

In [43]: df1.describe()

Out[43]:

	iyear	latitude	longitude
count	243.0	243.000000	243.000000
mean	2015.0	22.506835	85.500270
std	0.0	4.536694	6.156386
min	2015.0	9.829387	74.461538
25%	2015.0	19.132886	80.993779
50%	2015.0	23.755579	84.354205
75%	2015.0	25.043653	91.887771
max	2015.0	34.285891	95.663547

In [44]: d=df1['summary'].to_list()

In [45]: d

...

```
In [46]: date=[]
for i in d:
    i=i.split(':')
    date.append(i[0])
```

In [47]: date

...

```
In [58]: dates=defaultdict(int)
        for di in date:
            dates[di]+=1
```

```
In [64]: dates.keys()
```

...

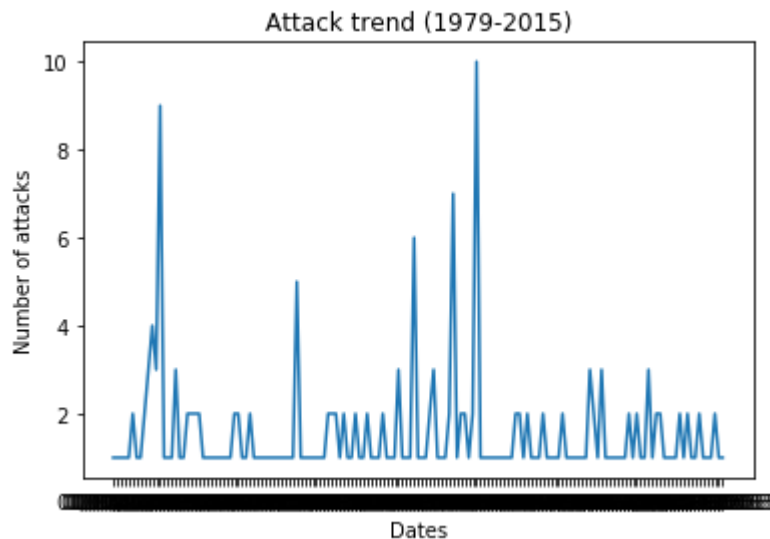
```
In [50]: dates.values()
```

...

Study of the pattern of the number of attacks in 2015.

```
In [66]: x=list(dates.keys())
        y=list(dates.values())
        plt.xlabel('Dates')
        plt.ylabel('Number of attacks')
        plt.title('Attack trend (1979-2015)')
        plt.plot(x,y)
```

```
Out[66]: [<matplotlib.lines.Line2D at 0x1d23340b808>]
```



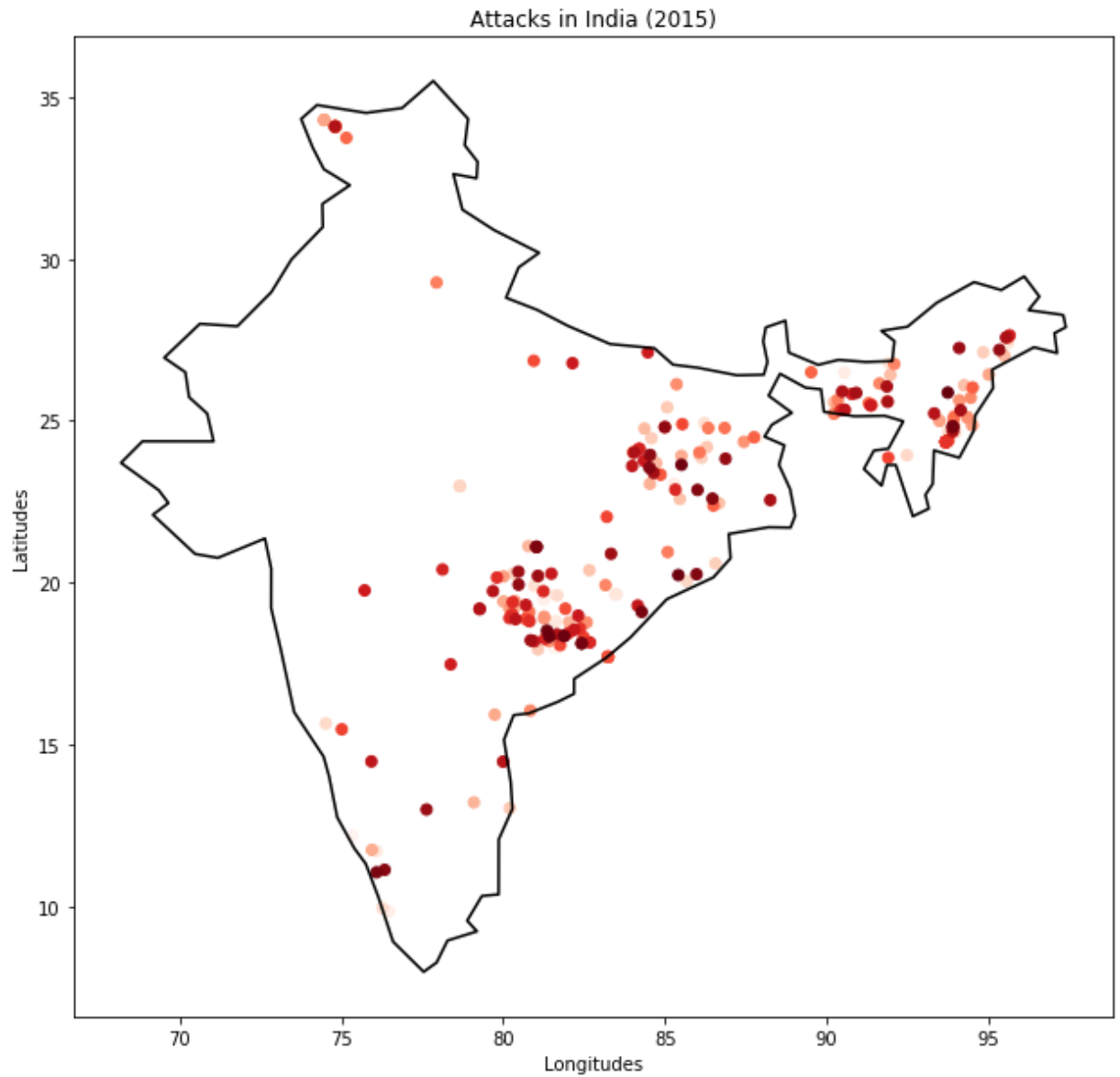
```
In [53]: gdf02=gpd.GeoDataFrame(df1,geometry=gpd.points_from_xy(df1['longitude'],df1['lat:
        gdf02.plot(figsize=(20,10))
        a=gdf02[gdf02['country_txt']=='India']
        asia=world[world['continent']=='Asia']
```

...

Distribution of attacks in each region in India (2015).

```
In [68]: fig,ax=plt.subplots(figsize=(20,10))
a.plot(cmap='Reds',ax=ax)
India.boundary.geometry.plot(color=None,edgecolor='k',ax=ax)
plt.xlabel('Longitudes')
plt.ylabel('Latitudes')
plt.title('Attacks in India (2015)')
```

```
Out[68]: Text(0.5, 1, 'Attacks in India (2015)')
```



```
In [ ]:
```

