

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")

In [2]: data = pd.read_csv("EDA(Terrorism).csv",encoding='latin1')

In [3]: data.head()

Out[3]:
   EventId  Year  Month  Day  Country  Region  Provstate  City  Latitude  Longitude  ...  Hostkid  Ransom  Ransomamt  Hostkidoutcome  Hostkidoutcome_txt  rreleasd  INT_LOG  INT_IDEO  INT_MISC  INT_Ap
0  1.970000e+11  1970   7    2  Dominican Republic  CentAS America & Caribbean  NaN  Santo Domingo  18.456792  -69.951164  ...  0.0  0.0  NaN  NaN  NaN  NaN  0  0  0
1  1.970000e+11  1970   0    0  Mexico  North America  FederAS  Mexico city  19.371887  -99.086524  ...  1.0  1.0  80000.0  NaN  NaN  NaN  0  1  1
2  1.970000e+11  1970   1    0  Phil  Southeast Asia  Tariac  Unknown  15.478598  120.599741  ...  0.0  0.0  NaN  2.0  Hostage(s) Released by Perpetrators  -9  -9  1
3  1.970000e+11  1970   1    0  Greece  Western Europe  Attica  Athens  37.997490  23.762728  ...  0.0  0.0  NaN  NaN  NaN  NaN  -9  -9  1
4  1.970000e+11  1970   1    0  Japan  East Asia  Fukouka  Fukouka  33.580412  130.396361  ...  0.0  0.0  NaN  NaN  NaN  NaN  -9  -9  1
5 rows x 60 columns

In [4]: data.columns.values

Out[4]:
array(['EventId', 'Year', 'Month', 'Day', 'Country', 'Region', 'Provstate', 'City', 'Latitude', 'Longitude', 'Hostkid', 'Ransom', 'Ransomamt', 'Hostkidoutcome', 'Hostkidoutcome_txt', 'rreleasd', 'INT_LOG', 'INT_IDEO', 'INT_MISC', 'INT_Ap'], dtype=object)

In [5]: data.rename(columns={'Year':'Year','Month':'Month','Day':'Day','gname':'Group','Country':'Country','Region':'Region','Provstate':'State','City':'City','Latitude':'Latitude','Longitude':'Longitude','Attacktype_txt':'Attacktype','Targetype_txt':'Targetype','weaptype1_txt':'Weapon','nkill':'Kill','mound':'Wound','INT_Ap':'INT_Ap'},dtype=object)

In [6]: data = data[['Year', 'Month', 'Day', 'Country', 'State', 'Region', 'City', 'Latitude', 'Longitude', 'Attacktype', 'Kill', 'Wound', 'Targetype', 'Weapon', 'Ransom']]

In [7]: data.head()

Out[7]:
   Year  Month  Day  Country  State  Region  City  Latitude  Longitude  Attacktype  Kill  Wound  target1  Group  Targetype  Weapon  Ransom
0  1970     7    2  Dominican Republic  NaN  CentAS America & Caribbean  Santo Domingo  18.456792  -69.951164  Assassination  1.0  0.0  Julio Guzman  MANO-D  Private Citizens & Property  Unknown  0.0
1  1970     0    0  Mexico  FederAS  North America  Mexico city  19.371887  -99.086524  Hostage Taking (Kidnapping)  0.0  0.0  Nadine ChavAS, daughter  23rd of September Communist League  Government (Diplomatic)  Unknown  1.0
2  1970     1    0  Phil  Tariac  Southeast Asia  Unknown  15.478598  120.599741  Assassination  1.0  0.0  Employee  Unknown  JournASists & Media  Unknown  0.0
3  1970     1    0  Greece  Attica  Western Europe  Athens  37.997490  23.762728  Bombing/Explosion  NaN  NaN  U.S. Embassy  Unknown  Government (Diplomatic)  Explosives  0.0
4  1970     1    0  Japan  Fukouka  East Asia  Fukouka  33.580412  130.396361  Facility/Infrastructure Attack  NaN  NaN  U.S. Consulate  Unknown  Government (Diplomatic)  Incendiary  0.0
```

```
In [8]: data.shape

Out[8]:
(181691, 17)

In [9]: data.isnull().sum()

Out[9]:
Year          0
Month         0
Day           0
Country       0
State        421
Region       434
City         456
Latitude     457
Longitude    457
Attacktype   0
Kill        16313
Wound       16311
target1      636
Group        0
Targetype    0
Weapon       0
Ransom     194319
dtype: int64

In [10]: data['Wound'] = data['Wound'].fillna(0)
data['Kill'] = data['Kill'].fillna(0)
data['Casualties'] = data['Wound'] + data['Kill']

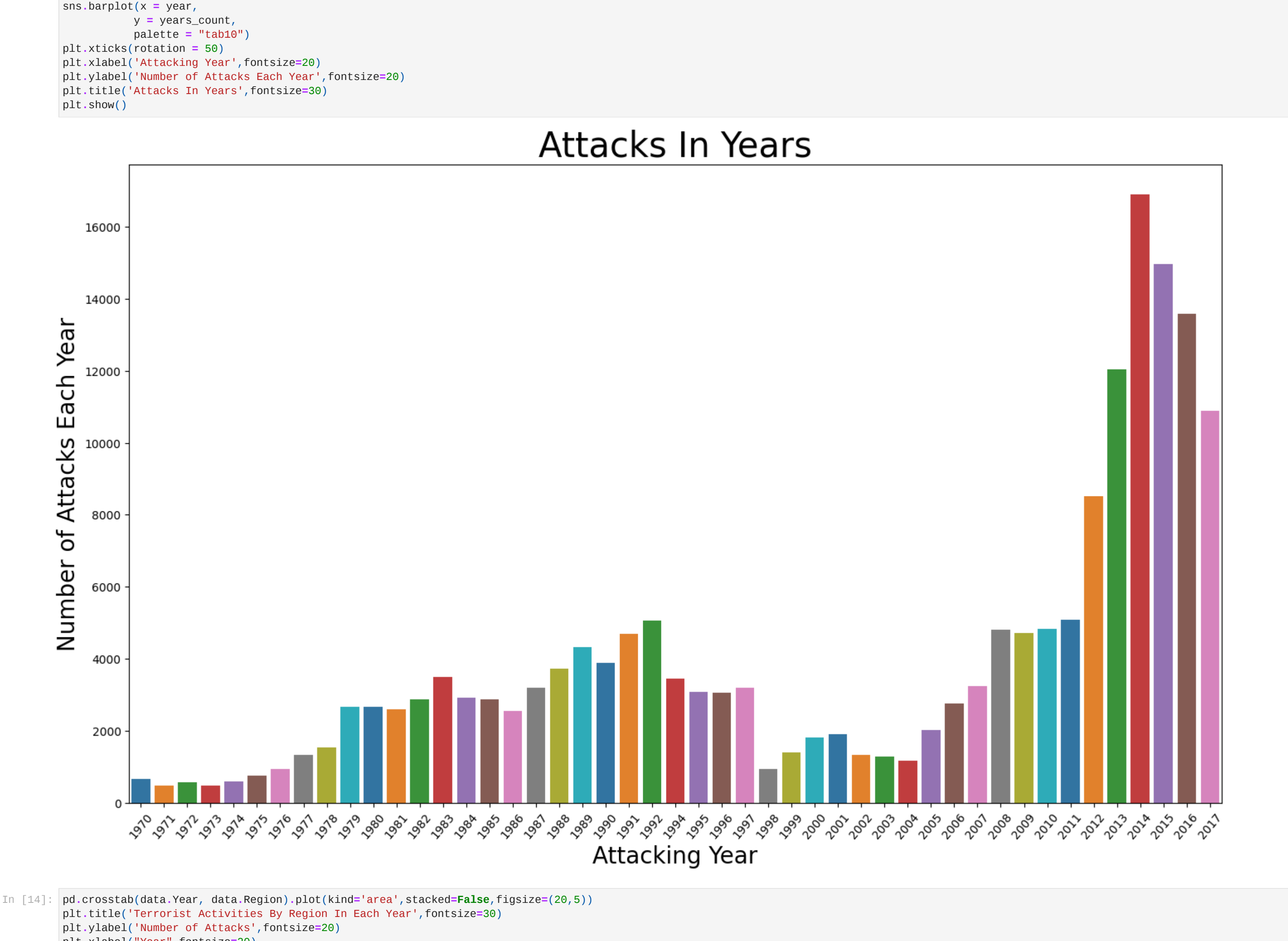
In [11]: data.info()

class pandas.core.frame.DataFrame
RangeIndex: 181691 entries, 0 to 181690
Data columns (total 18 columns):
 #   Column  Non-Null Count  Dtype
---  --
 0   Year   181691 non-null      int64
 1   Month  181691 non-null      int64
 2   Day    181691 non-null      int64
 3   Country 181691 non-null      object
 4   State  181270 non-null      object
 5   Region 181691 non-null      object
 6   City   181267 non-null      object
 7   Latitude 177135 non-null     float64
 8   Longitude 177134 non-null     float64
 9   Attacktype 181691 non-null     object
10  Kill   181691 non-null      float64
11  Wound  181691 non-null      float64
12  target1 181691 non-null      object
13  Group   181691 non-null      object
14  Targetype 181691 non-null      object
15  Weapon  181691 non-null      object
16  Ransom  77981 non-null       float64
17  Casualties 181691 non-null   float64
dtypes: float64(8), int64(3), object(9)
memory usage: 25.0+ MB

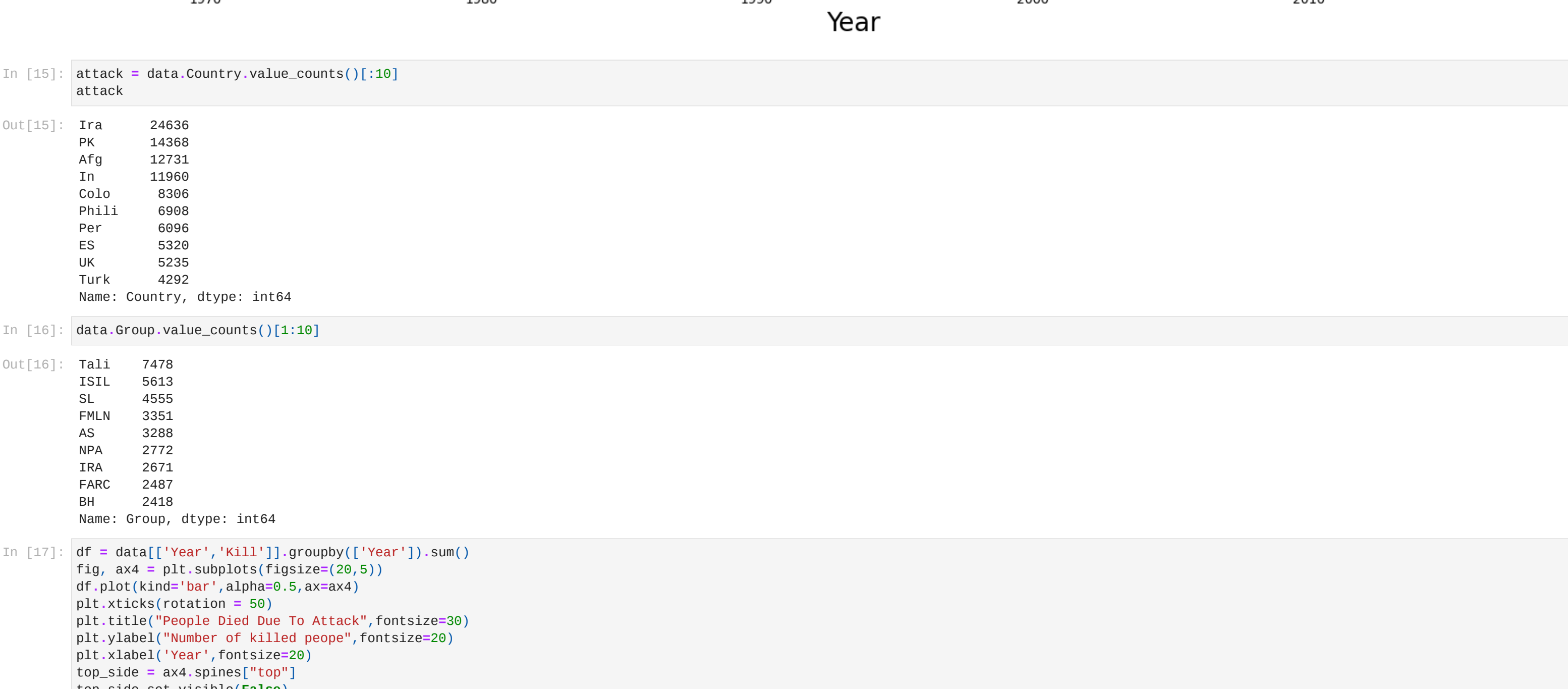
In [12]: data.describe()

Out[12]:
      Year      Month      Day      Latitude      Longitude      Kill      Wound      Ransom      Casualties
mean  181691.000000  181691.000000  181691.000000  177135.000000  1.771340e+05  181691.000000  181691.000000  77381.000000  181691.000000
count  181691.000000  181691.000000  181691.000000  177135.000000  1.771340e+05  181691.000000  181691.000000  77381.000000  181691.000000
min    1970.000000  0.000000  0.000000  15.454613  -8.618590e+07  0.000000  0.000000  -9.000000  0.000000
std    13.259430  3.388303  8.814045  23.498343  -4.586967e+02  2.266860  2.883296  -0.145811  5.150156
qtd    1970.000000  0.000000  0.000000  15.454613  -8.618590e+07  0.000000  0.000000  -9.000000  0.000000
25%    1991.000000  4.000000  8.000000  11.510946  4.545640e+00  0.000000  0.000000  0.000000  0.000000
50%    2009.000000  6.000000  15.000000  31.467463  4.324651e+01  0.000000  0.000000  0.000000  1.000000
75%    2014.000000  9.000000  23.000000  34.685087  6.871033e+01  2.000000  2.000000  0.000000  4.000000
max    2017.000000  12.000000  31.000000  74.633553  1.793667e+02  1570.000000  8191.000000  1.000000  9574.000000
```

```
In [13]: year = data['Year'].unique()
years_count = data['Year'].value_counts(dropna = False).sort_index()
plt.figure(figsize = (17,18))
sns.barplot(x = year,
            y = years_count,
            palette = "tab10")
plt.xticks(rotation = 5)
plt.xlabel('Attacking Year',fontSize=28)
plt.ylabel('Number of Attacks Each Year',fontSize=26)
plt.title('Attacks In Years',fontSize=28)
plt.show()
```



```
In [14]: pd.crosstab(data.Year, data.Region).plot(kind='area',stacked=False,figsize=(20,5))
plt.title('Terrorist Activities By Region In Each Year',fontSize=20)
plt.ylabel('Number of Attacks',fontSize=20)
plt.xlabel('Year',fontSize=28)
plt.show()
```



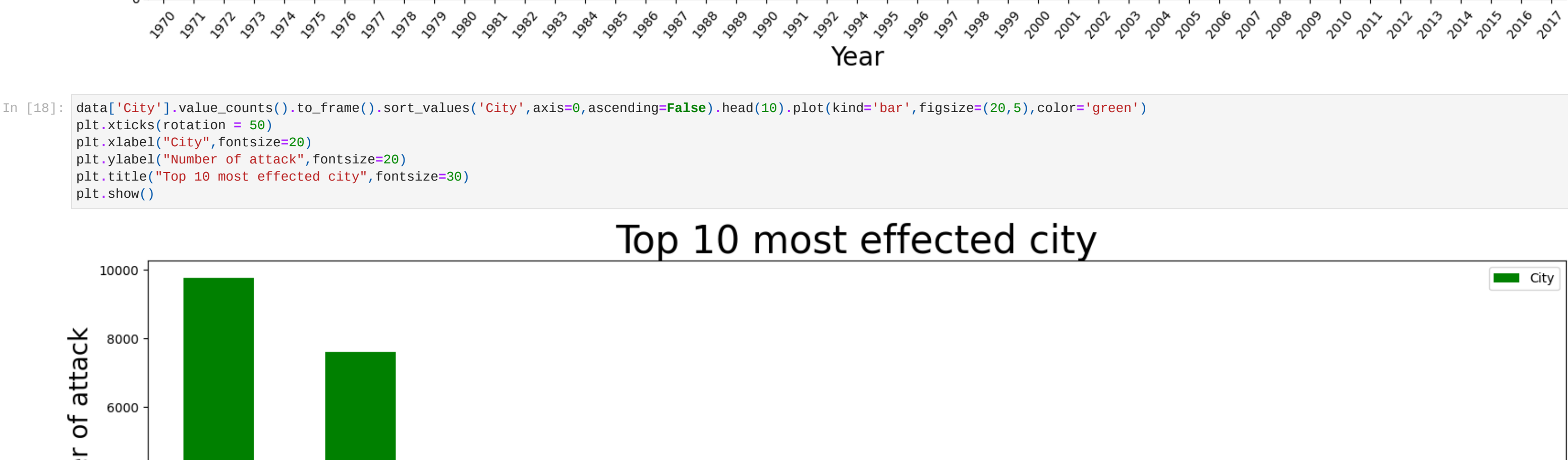
```
In [15]: attack = data.Country.value_counts()[1:10]
attack

Out[15]:
Ira      24636
PK       14368
AFG      12731
IN       13968
Colo     8386
Phil     6988
Per      6096
ES       5320
UK       5226
Turk     4292
Name: Country, dtype: int64

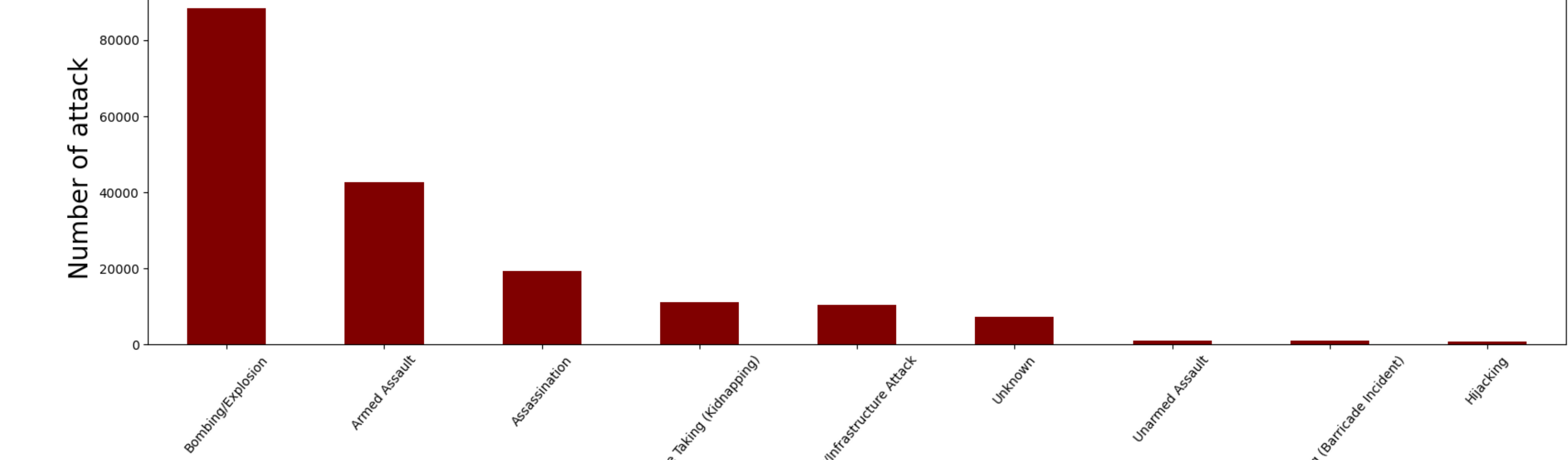
In [16]: data.Group.value_counts()[1:10]

Out[16]:
Tal     7478
ISIL    5633
SI       4550
FNLN    3351
AS       2908
NPA      2772
IRA      2671
FARC     2487
BH       2418
Name: Group, dtype: int64

In [17]: df = data[['Year','Kill']].groupby(['Year']).sum()
fig, ax4 = plt.subplots(figsize=(20,5))
df.plot(kind='bar',alpha=0.5,ax=ax4)
plt.xticks(rotation = 5)
plt.title('People Died Due To Attack',fontSize=30)
plt.ylabel('Number of killed people',fontSize=28)
plt.xlabel('Year',fontSize=20)
top_side = ax4.spines['top']
top_side.set_visible(False)
right_side = ax4.spines['right']
right_side.set_visible(False)
```



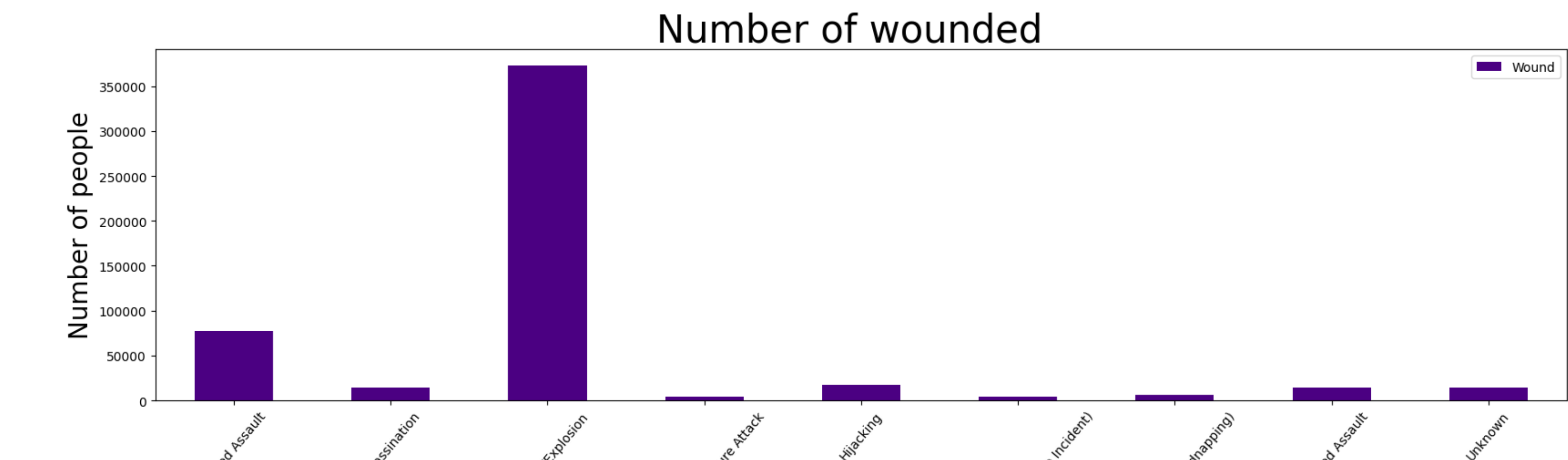
```
In [18]: data['City'].value_counts().to_frame().sort_values('City',axis=0,ascending=False).head(10).plot(kind='bar',figsize=(20,5),color='green')
plt.xticks(rotation = 5)
plt.title('Top 10 most effected city',fontSize=28)
plt.ylabel('Number of attack',fontSize=28)
plt.xlabel('City',fontSize=20)
plt.show()
```



```
In [19]: data['Attacktype'].value_counts().plot(kind='bar',figsize=(20,5),color='Maroon')
plt.xticks(rotation = 5)
plt.xlabel('Attacktype',fontSize=30)
plt.ylabel('Number of attack',fontSize=28)
plt.title('Name of attacktype',fontSize=28)
plt.show()
```



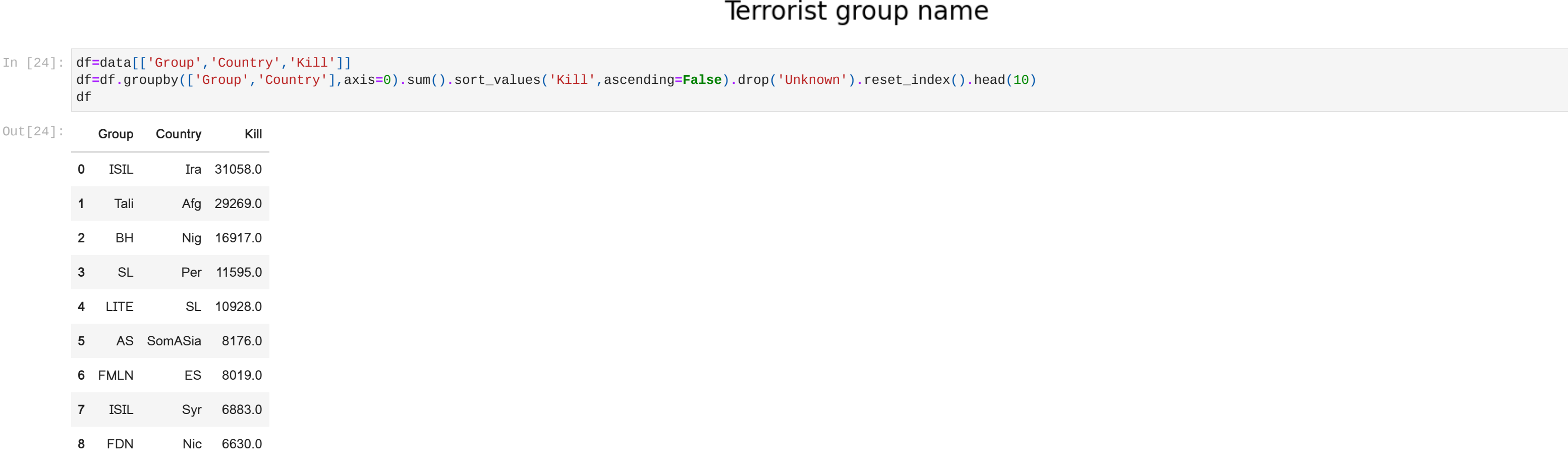
```
In [20]: data[['Attacktype', 'Kill']].groupby(['Attacktype'],axis=0).sum().plot(kind='bar',figsize=(20,5),color='Orange')
plt.xticks(rotation=5)
plt.title('Number of killed ',fontSize=30)
plt.ylabel('Number of people',fontSize=28)
plt.xlabel('Attack type',fontSize=30)
plt.show()
```



```
In [21]: data[['Attacktype', 'Wound']].groupby(['Attacktype'],axis=0).sum().plot(kind='bar',figsize=(20,5),color='Indigo')
plt.xticks(rotation=5)
plt.title('Number of wounded ',fontSize=30)
plt.ylabel('Number of people',fontSize=28)
plt.xlabel('Attack type',fontSize=30)
plt.show()
```



```
In [22]: data['Group'].value_counts().to_frame().drop('Unknown').head(10).plot(kind='bar',color='Grey',figsize=(20,5))
plt.xticks(rotation=5)
plt.title('Top 10 terrorist group attack',fontSize=30)
plt.xlabel('Terrorist group name',fontSize=30)
plt.ylabel('Attack number',fontSize=28)
plt.show()
```



```
In [23]: data[['Group','Kill']].groupby(['Group'],axis=0).sum().drop('Unknown').sort_values('Kill',ascending=False).head(10).plot(kind='bar',color='crimson',figsize=(20,5))
plt.xticks(rotation=5)
plt.title('Top 10 terrorist group attack',fontSize=30)
plt.xlabel('Terrorist group name',fontSize=30)
plt.ylabel('No of killed people',fontSize=28)
plt.show()
```



```
In [24]: df=data[['Group','Country','Kill']]
df=df.groupby(['Group','Country'],axis=0).sum().sort_values('Kill',ascending=False).drop('Unknown').reset_index().head(10)
df

Out[24]:
   Group  Country  Kill
0  ISIL  Ira  31058.0
1  Tal  Afg  29269.0
2  BH  Nig  16917.0
3  SL  Per  11595.0
4  LITE  SL  10928.0
5  AS  SomASia  8176.0
6  FNLN  ES  8019.0
7  ISIL  Sy  6883.0
8  FDN  Nic  6630.0
9  TTP  PK  6014.0

In [25]: Kill = data.loc[:, 'Kill']
print("Number of people killed by terror attack:",int(sum(Kill.dropna())))
```

```
In [26]: typeKill = data.pivot_table(columns='Attacktype', values='Kill', aggfunc='sum')
typeKill

Out[26]:
Attacktype  Armed Assault  Assassination  Bombing/Explosion  Facility/Infrastructure Attack  Hijacking  Hostage Taking (Barricade Incident)  Hostage Taking (Kidnapping)  Unarmed Assault  Unknown
Kill          160297.0      24920.0      157321.0      3642.0      3718.0      4478.0      24231.0      880.0      32381.0

In [27]: countryKill = data.pivot_table(columns='Country', values='Kill', aggfunc='sum')
countryKill

Out[27]:
Country  Albania  Algeria  Alg  Andorra  Angola  Antigua and Barbuda  Argentina  Armenia  AustrASia  Austria  ...  Vietnam  WAsia and Futuna  West Bank and Gaza Strip  West Germany (FRG)  Western Sahara  Yemen  Yugoslavia  Zaire  Zambia  Zimbabwe
Kill          4.0      11066.0  39384.0      0.0      3043.0      0.0      490.0      37.0      23.0      30.0  ...  1.0      0.0      1500.0      97.0      1.0      8776.0      119.0      324.0      70.0      194.0
1 rows x 42 columns

In [28]: print("This was an Exploratory Data Analysis - Terrorism by Ayushi Lanjewar")
This was an Exploratory Data Analysis - Terrorism by Ayushi Lanjewar

In [ ]:
```