

TASK 4: EXPLORATORY ANALYSIS ON GLOBAL TERRORISM

BY SANA TADVI

```
In [1]: #IMPORTING THE LIBRARIES

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
import time
from IPython.display import clear_output
%matplotlib inline
```

```
In [2]: #LOADING THE DATA

terror=pd.read_csv('globalterrorism.csv',encoding='latin1')
```

```
In [3]: terror.head()
```

Out[3]:

	eventid	iyear	imonth	iday	approxdate	extended	resolution	country	country_txt	regi
0	1970000000001	1970	7	2	NaN	0	NaN	58	Dominican Republic	
1	1970000000002	1970	0	0	NaN	0	NaN	130	Mexico	
2	1970010000001	1970	1	0	NaN	0	NaN	160	Philippines	
3	1970010000002	1970	1	0	NaN	0	NaN	78	Greece	
4	1970010000003	1970	1	0	NaN	0	NaN	101	Japan	

5 rows × 135 columns

```
In [4]: #CHECKING THEGENERAL INFORMATION
terror.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 181691 entries, 0 to 181690
Columns: 135 entries, eventid to related
dtypes: float64(55), int64(22), object(58)
memory usage: 146.9+ MB
```

```
In [5]: terror.columns.values
```

```
Out[5]: array(['eventid', 'iyear', 'imonth', 'iday', 'approxdate', 'extended',
'resolution', 'country', 'country_txt', 'region', 'region_txt',
'provstate', 'city', 'latitude', 'longitude', 'specificity',
'vicinity', 'location', 'summary', 'crit1', 'crit2', 'crit3',
'doubtterr', 'alternative', 'alternative_txt', 'multiple',
'success', 'suicide', 'attacktype1', 'attacktype1_txt',
'attacktype2', 'attacktype2_txt', 'attacktype3', 'attacktype3_txt',
'targtype1', 'targtype1_txt', 'targsubtype1', 'targsubtype1_txt',
'corp1', 'target1', 'natlty1', 'natlty1_txt', 'targtype2',
'targtype2_txt', 'targsubtype2', 'targsubtype2_txt', 'corp2',
```

```
'target2', 'natlty2', 'natlty2_txt', 'targtype3', 'targtype3_txt',
'targsubtype3', 'targsubtype3_txt', 'corp3', 'target3', 'natlty3',
'natlty3_txt', 'gname', 'gsubname', 'gname2', 'gsubname2',
'gname3', 'gsubname3', 'motive', 'guncertain1', 'guncertain2',
'guncertain3', 'individual', 'nperps', 'nperpcap', 'claimed',
'claimmode', 'claimmode_txt', 'claim2', 'claimmode2',
'claimmode2_txt', 'claim3', 'claimmode3', 'claimmode3_txt',
'compclaim', 'weaptype1', 'weaptype1_txt', 'weapsubtype1',
'weapsubtype1_txt', 'weaptype2', 'weaptype2_txt', 'weapsubtype2',
'weapsubtype2_txt', 'weaptype3', 'weaptype3_txt', 'weapsubtype3',
'weapsubtype3_txt', 'weaptype4', 'weaptype4_txt', 'weapsubtype4',
'weapsubtype4_txt', 'weapdetail', 'nkill', 'nkillus', 'nkillter',
'nwound', 'nwoundus', 'nwoundte', 'property', 'propextent',
'propextent_txt', 'propvalue', 'propcomment', 'ishostkid',
'nhostkid', 'nhostkidus', 'nhours', 'ndays', 'divert',
'kidhijcountry', 'ransom', 'ransomamt', 'ransomamtus',
'ransompaid', 'ransompaidus', 'ransomnote', 'hostkidoutcome',
'hostkidoutcome_txt', 'nreleased', 'addnotes', 'scite1', 'scite2',
'scite3', 'dbsource', 'INT_LOG', 'INT_IDEO', 'INT_MISC', 'INT_ANY',
'related'], dtype=object)
```

```
In [6]: #RERENAMING THE COLUMNS IN THE DATA FOR THE PROPER UNDERSTANDING
terror.rename(columns={'eventid':'Eventid', 'iyear':'Year', 'imonth':'Month', 'iday'
                        'extended':'Extended', 'resolution':'Resolution', 'attack
                        'country_txt':'Country', 'region_txt':'Region', 'provstate
                        'city':'City', 'crit1':'Crit1', 'crit2':'Crit2',
                        'crit3':'Crit3', 'multiple':'Multiple', 'success':'Success
                        'targtype1_txt':'Targtype', 'natlty1_txt':'Natlty1', 'natl
                        'natlty3_txt':'Natlty3', 'gname':'Gname', 'gname2':'Gname2
                        'guncertain1':'Guncertain1', 'guncertain2':'Guncertain2',
                        'claimed':'Claimed', 'weaptype1_txt':'Weaptype', 'weapsubt
                        'nkill':'Nkill', 'nkillus':'Nkillus', 'nkillter':'Nkillter
                        'nwoundus':'Nwoundus', 'nwoundte':'Nwoundter', 'property':
                        'propextent_txt':'Propextent', 'propvalue':'Propvalue', 'i
                        'nhostkid':'Nhostkid', 'nhostkidus':'Nhostkidus', 'ransom'
                        'hostkidoutcome':'Hostkidoutcome', 'nreleased':'Nreleased'
```

```
In [9]: #KEEPING 20 COLUMNS IMPORTANT DATA WHICH WILL BE USEFUL FOR THE ANALYSIS

globterror= terror[['Eventid', 'Year', 'Country', 'Region', 'Provstate', 'City', 'Cri
                    'Success', 'Suicide', 'Attacktype', 'Targtype', 'Natlty1', 'Gname',
                    'Claimed', 'Weaptype', 'Nkill', 'Nwound']]

globterror.head(10)
```

```
Out[9]:
```

	Eventid	Year	Country	Region	Provstate	City	Crit1	Crit2	Crit3	Success
0	197000000001	1970	Dominican Republic	Central America & Caribbean	NaN	Santo Domingo	1	1	1	1
1	197000000002	1970	Mexico	North America	Federal	Mexico city	1	1	1	1
2	197001000001	1970	Philippines	Southeast Asia	Tarlac	Unknown	1	1	1	1
3	197001000002	1970	Greece	Western Europe	Attica	Athens	1	1	1	1
4	197001000003	1970	Japan	East Asia	Fukouka	Fukouka	1	1	1	1

	Eventid	Year	Country	Region	Provstate	City	Crit1	Crit2	Crit3	Success
5	197001010002	1970	United States	North America	Illinois	Cairo	1	1	1	1
6	197001020001	1970	Uruguay	South America	Montevideo	Montevideo	1	1	1	0
7	197001020002	1970	United States	North America	California	Oakland	1	1	1	1
8	197001020003	1970	United States	North America	Wisconsin	Madison	1	1	1	1
9	197001030001	1970	United States	North America	Wisconsin	Madison	1	1	1	1

In [10]: `globterror.shape`

Out[10]: (181691, 20)

In [11]: `globterror.isnull().sum()`

```
Out[11]: Eventid      0
Year      0
Country    0
Region     0
Provstate  421
City      434
Crit1      0
Crit2      0
Crit3      0
Success    0
Suicide    0
Attacktype  0
Targtype   0
Natlty1    1559
Gname      0
Guncertain1 380
Claimed    66120
Weaptype   0
Nkill      10313
Nwound     16311
dtype: int64
```

In [12]: `globterror.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 181691 entries, 0 to 181690
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Eventid         181691 non-null  int64
1   Year            181691 non-null  int64
2   Country         181691 non-null  object
3   Region          181691 non-null  object
4   Provstate       181270 non-null  object
5   City            181257 non-null  object
6   Crit1           181691 non-null  int64
7   Crit2           181691 non-null  int64
8   Crit3           181691 non-null  int64
9   Success         181691 non-null  int64
10  Suicide         181691 non-null  int64
11  Attacktype      181691 non-null  object
12  Targtype        181691 non-null  object
```

```
13 Natlty1      180132 non-null object
14 Gname        181691 non-null object
15 Guncertain1  181311 non-null float64
16 Claimed      115571 non-null float64
17 Weaptype     181691 non-null object
18 Nkill        171378 non-null float64
19 Nwound       165380 non-null float64
dtypes: float64(4), int64(7), object(9)
memory usage: 21.5+ MB
```

```
In [13]: globterror.describe()
```

Out[13]:

	Eventid	Year	Crit1	Crit2	Crit3	Success	
count	1.816910e+05	181691.000000	181691.000000	181691.000000	181691.000000	181691.000000	1
mean	2.002705e+11	2002.638997	0.988530	0.993093	0.875668	0.889598	
std	1.325957e+09	13.259430	0.106483	0.082823	0.329961	0.313391	
min	1.970000e+11	1970.000000	0.000000	0.000000	0.000000	0.000000	
25%	1.991021e+11	1991.000000	1.000000	1.000000	1.000000	1.000000	
50%	2.009022e+11	2009.000000	1.000000	1.000000	1.000000	1.000000	
75%	2.014081e+11	2014.000000	1.000000	1.000000	1.000000	1.000000	
max	2.017123e+11	2017.000000	1.000000	1.000000	1.000000	1.000000	

```
In [14]: #CREATE AND ADD COLUMN DAMAGE (Number of Dead OR Injured people) by adding Nkill and
globterror['Damage'] = globterror['Nkill'] + globterror['Nwound']
globterror.describe()
```

Out[14]:

	Eventid	Year	Crit1	Crit2	Crit3	Success	
count	1.816910e+05	181691.000000	181691.000000	181691.000000	181691.000000	181691.000000	1
mean	2.002705e+11	2002.638997	0.988530	0.993093	0.875668	0.889598	
std	1.325957e+09	13.259430	0.106483	0.082823	0.329961	0.313391	
min	1.970000e+11	1970.000000	0.000000	0.000000	0.000000	0.000000	
25%	1.991021e+11	1991.000000	1.000000	1.000000	1.000000	1.000000	
50%	2.009022e+11	2009.000000	1.000000	1.000000	1.000000	1.000000	
75%	2.014081e+11	2014.000000	1.000000	1.000000	1.000000	1.000000	
max	2.017123e+11	2017.000000	1.000000	1.000000	1.000000	1.000000	

```
In [17]: #CHECKING THE HIGHEST ATTACKS IN COUNTRIES AND REGION

print('Country with the highest number of Terror Attacks:',globterror['Country'].
print('Regions with the highest number of Terror Attacks:',globterror['Region'].v
print('Maximum n0. of people were killed by a single terrorist attack are',globte
      'people that took place in',globterror.loc[globterror['Nkill'].idxmax()].Count

Country with the highest number of Terror Attacks: Iraq
Regions with the highest number of Terror Attacks: Middle East & North Africa
Maximum n0. of people were killed by a single terrorist attack are 1570.0 people tha
t took place in Iraq
```

In []:

In []:

Data analysis of Global Terrorism from 1970 to 2017 (EDA)

In [19]:

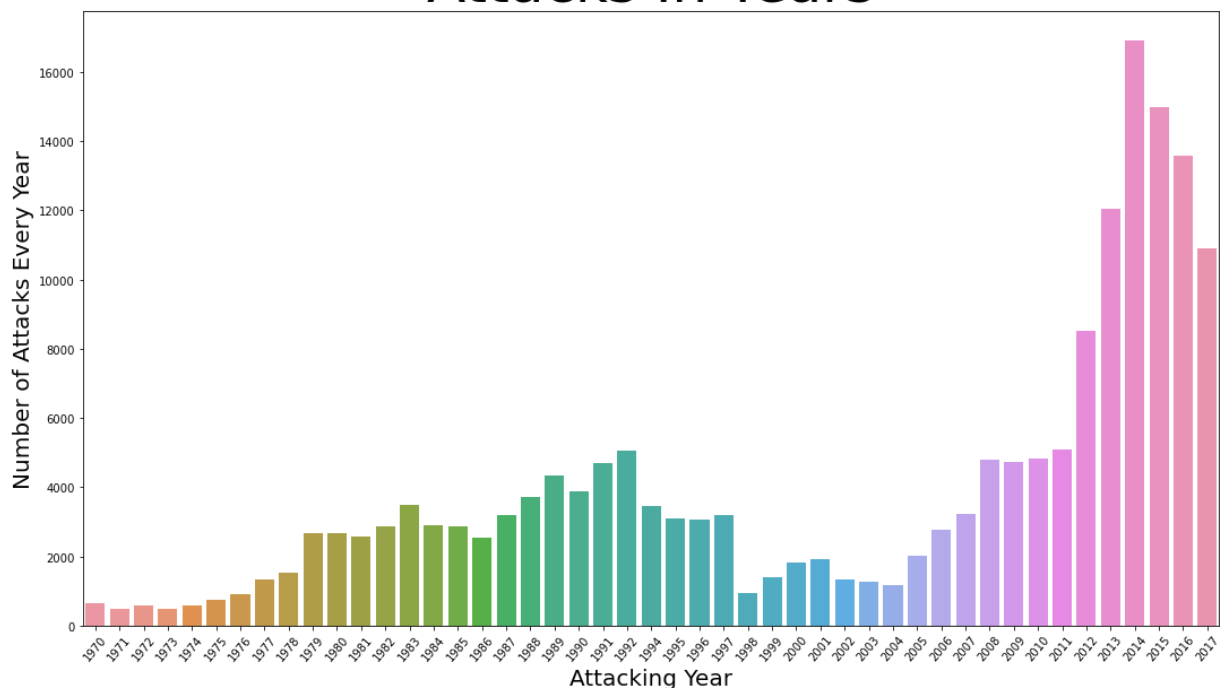
#GRAPH SHOWING ATTACKS IN YEAR

```

year = globterror['Year'].unique()
years_count = globterror['Year'].value_counts(dropna = False).sort_index()
plt.figure(figsize = (18,10))
sns.barplot(x = year,
            y = years_count)
plt.xticks(rotation = 50)
plt.xlabel('Attacking Year',fontsize=20)
plt.ylabel('Number of Attacks Every Year',fontsize=20)
plt.title('Attacks In Years',fontsize=50)
plt.show()

```

Attacks In Years



Terror activities in Countries, Cities and Region

In [20]:

#COUNTRIES EDA ANALYSIS

```

attack = globterror.Country.value_counts()[:10]
attack

```

```

Out[20]: Iraq          24636
Pakistan    14368
Afghanistan 12731
India       11960
Colombia     8306
Philippines  6908

```

```

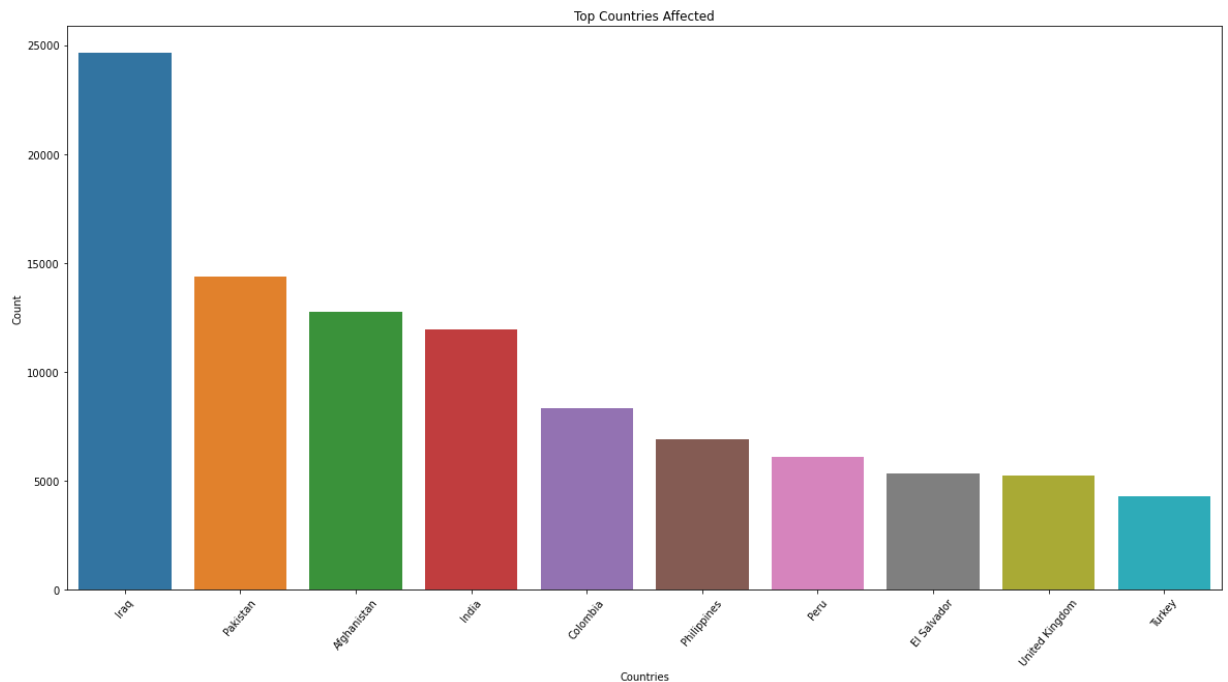
Peru                6096
El Salvador         5320
United Kingdom      5235
Turkey              4292
Name: Country, dtype: int64

```

```

In [33]: plt.subplots(figsize=(20,10))
sns.barplot(globterror['Country'].value_counts()[:10].index,globterror['Country'].va
plt.title('Top Countries Affected')
plt.xlabel('Countries')
plt.ylabel('Count')
plt.xticks(rotation = 50)
plt.show()

```

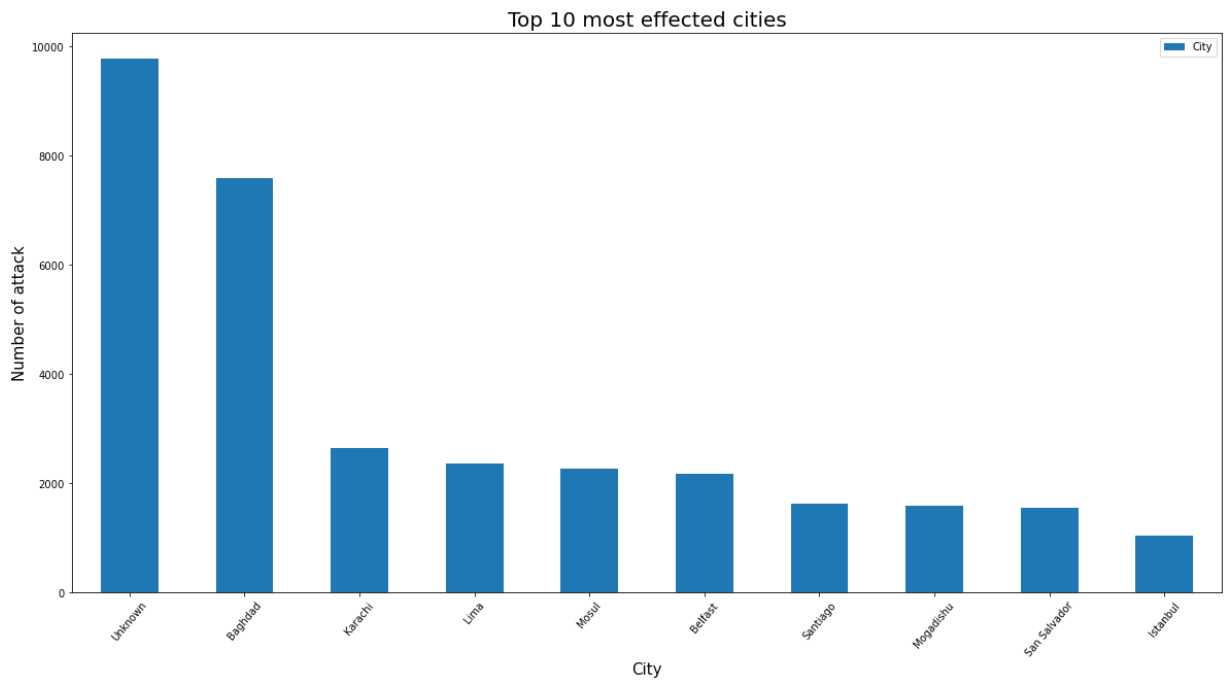


```

In [32]: #CITIES EDA ANALYSIS

globterror['City'].value_counts().to_frame().sort_values('City',axis=0,ascending=False)
plt.xticks(rotation = 50)
plt.xlabel("City",fontsize=15)
plt.ylabel("Number of attack",fontsize=15)
plt.title("Top 10 most effected cities",fontsize=20)
plt.show()

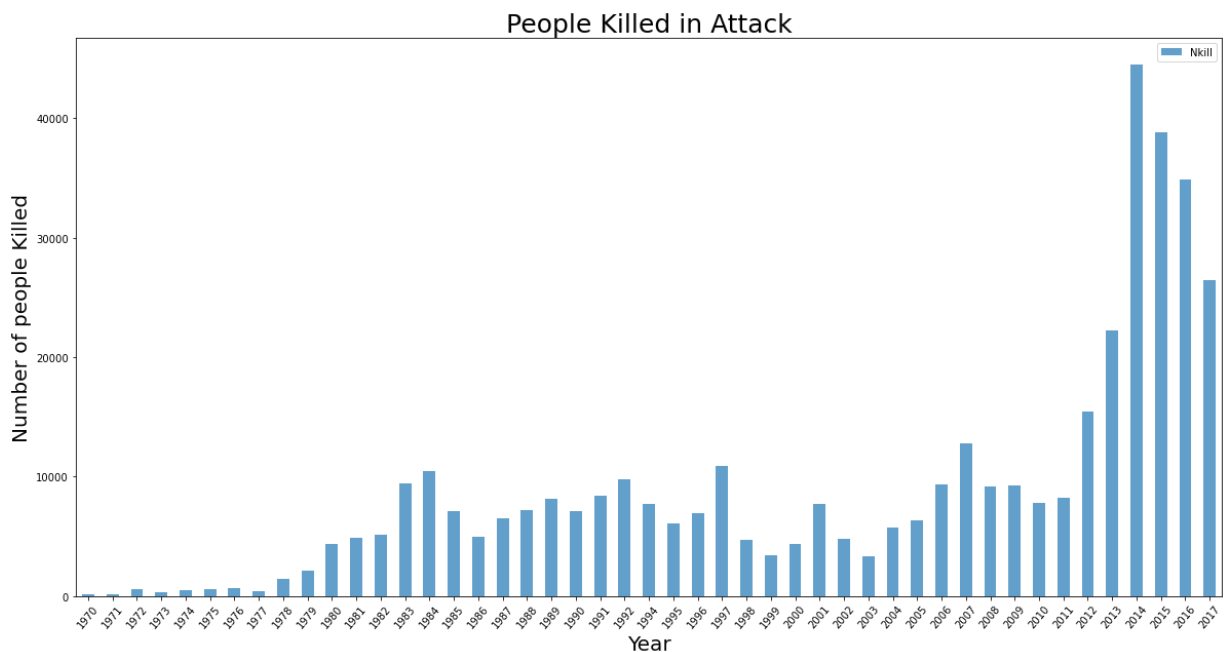
```



Number of People Killed

```
In [38]: df= globterror[['Year','Nkill']].groupby(['Year']).sum()
fig, ax4 = plt.subplots(figsize=(20,10))
df.plot(kind='bar',alpha=0.7,ax=ax4)
plt.xticks(rotation = 50)
plt.title("People Killed in Attack",fontsize=25)
plt.ylabel("Number of people Killed",fontsize=20)
plt.xlabel('Year',fontsize=20)
```

Out[38]: Text(0.5, 0, 'Year')



Top attacker Groups EDA

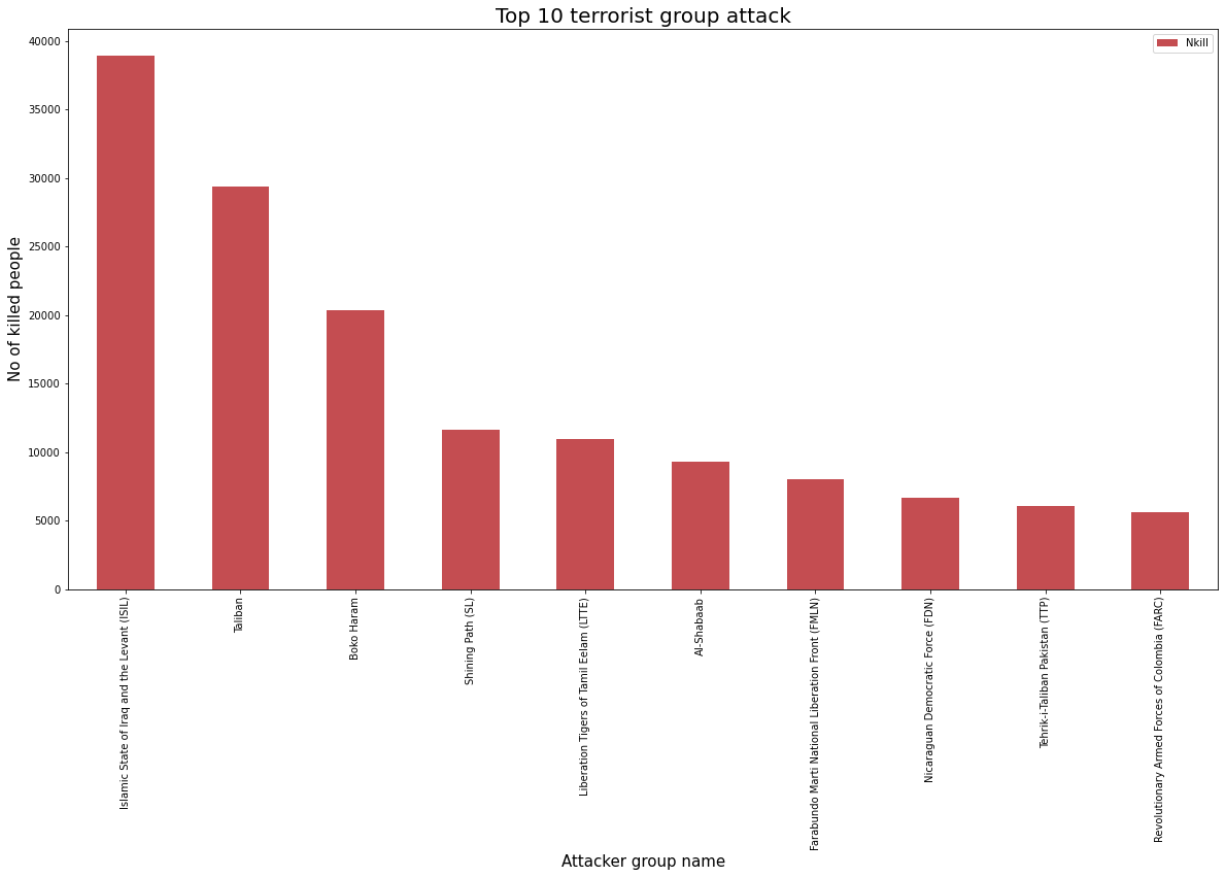
```
In [39]: Tclan=globterror[['Gname','Country','Nkill']]
Clan=Tclan.groupby(['Gname','Country'],axis=0).sum().sort_values('Nkill',ascending=F)
Clan
```

Out[39]:

	Gname	Country	Nkill
0	Islamic State of Iraq and the Levant (ISIL)	Iraq	31058.0
1	Taliban	Afghanistan	29269.0
2	Boko Haram	Nigeria	16917.0
3	Shining Path (SL)	Peru	11595.0
4	Liberation Tigers of Tamil Eelam (LTTE)	Sri Lanka	10928.0
5	Al-Shabaab	Somalia	8176.0
6	Farabundo Marti National Liberation Front (FMLN)	El Salvador	8019.0
7	Islamic State of Iraq and the Levant (ISIL)	Syria	6883.0
8	Nicaraguan Democratic Force (FDN)	Nicaragua	6630.0
9	Tehrik-i-Taliban Pakistan (TTP)	Pakistan	6014.0

In [42]:

```
globterror[['Gname', 'Nkill']].groupby(['Gname'], axis=0).sum().drop('Unknown').sort_v
plt.title("Top 10 terrorist group attack", fontsize=20)
plt.xlabel("Attacker group name", fontsize=15)
plt.ylabel("No of killed people", fontsize=15)
plt.show()
```



In [44]:

```
# NUMBER OF KILLS IN COUNTRIES
countryKill = globterror.pivot_table(columns='Country', values='Nkill', aggfunc='sum')
countryKill
```

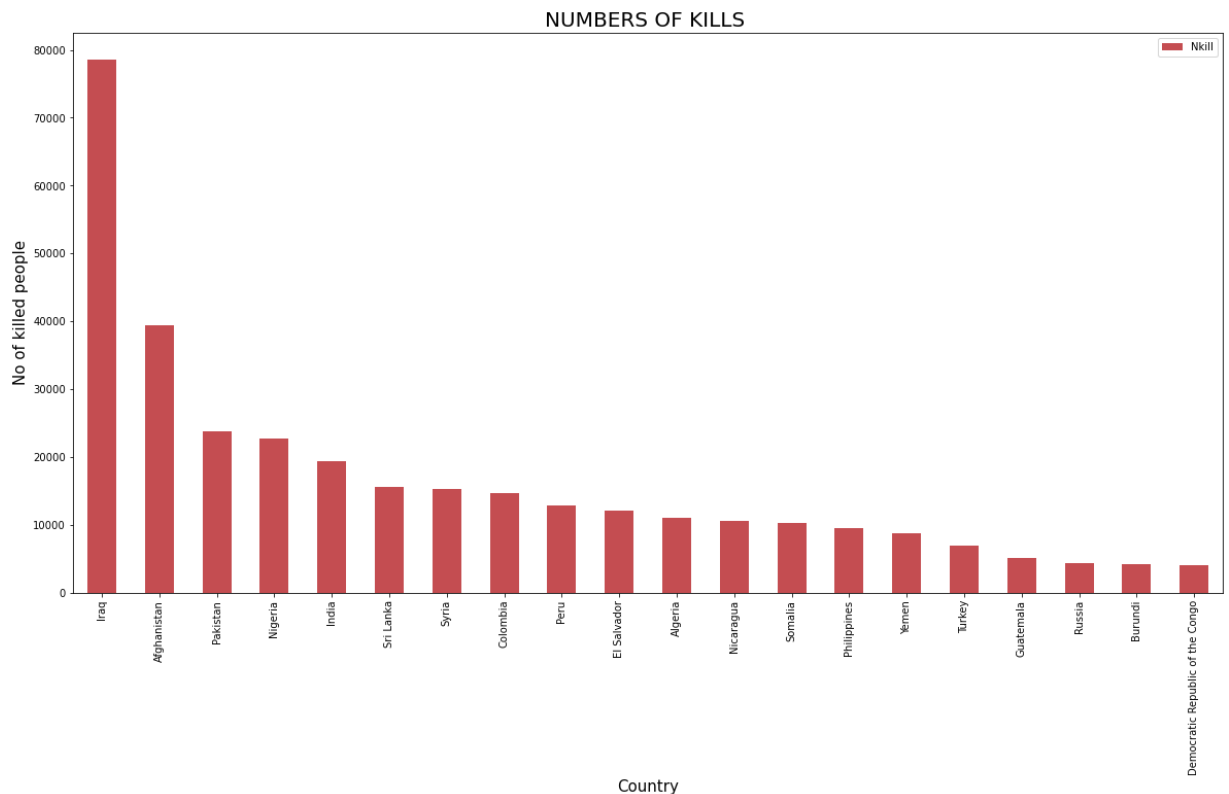
Out[44]:

Country	Afghanistan	Albania	Algeria	Andorra	Angola	Antigua and Barbuda	Argentina	Armenia	Australia
---------	-------------	---------	---------	---------	--------	---------------------	-----------	---------	-----------

Country	Afghanistan	Albania	Algeria	Andorra	Angola	Antigua and Barbuda	Argentina	Armenia	Australia
Nkill	39384.0	42.0	11066.0	0.0	3043.0	0.0	490.0	37.0	23.0

1 rows × 205 columns

```
In [52]: globterror[['Country','Nkill']].groupby(['Country'],axis=0).sum().sort_values('Nkill')
plt.title("NUMBERS OF KILLS ",fontsize=20)
plt.xlabel("Country",fontsize=15)
plt.ylabel("No of killed people",fontsize=15)
plt.show()
```

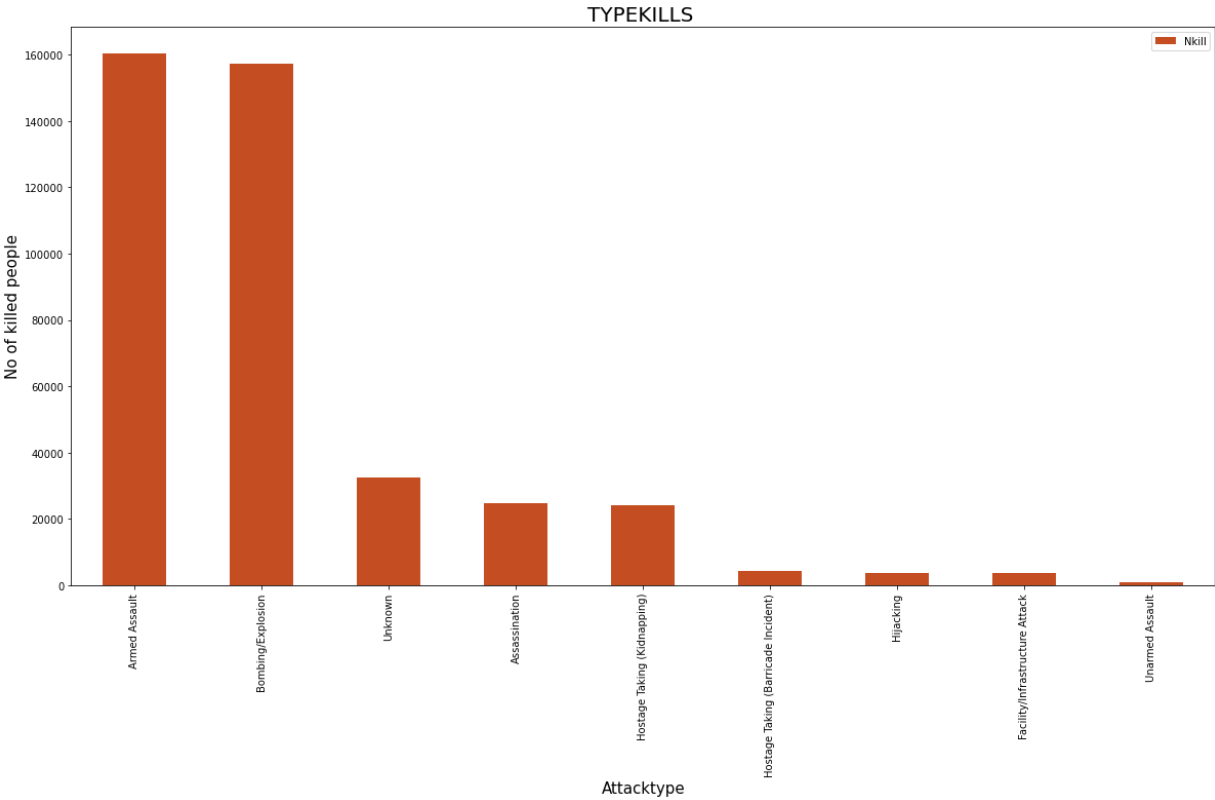


```
In [54]: # WEAPON KILLS
typeKill = globterror.pivot_table(columns='Attacktype', values='Nkill', aggfunc='sum')
typeKill
```

```
Out[54]:
```

Attacktype	Armed Assault	Assassination	Bombing/Explosion	Facility/Infrastructure Attack	Hijacking	Hostage Taking (Barricade Incident)
Nkill	160297.0	24920.0	157321.0	3642.0	3718.0	4478.0

```
In [73]: globterror[['Attacktype','Nkill']].groupby(['Attacktype'],axis=0).sum().sort_values('Nkill')
plt.title("TYPEKILLS ",fontsize=20)
plt.xlabel("Attacktype",fontsize=15)
plt.ylabel("No of killed people",fontsize=15)
plt.show()
```



```
In [60]: pplkill = globterror.loc[:, 'Nkill']
print('Number of people killed by terror attack:', int(sum(pplkill.dropna())))
```

Number of people killed by terror attack: 411868

Safest countries from 2000 to 2017 which has 0 dead/injured people caused by Terrorist Attacks

```
In [62]: SAFE = globterror[globterror['Year'] > 1999][["Country", "Damage"]].groupby('Country')
In [64]: SAFE[SAFE["Damage"]==0]
```

Out[64]:

	Country	Damage
9	Bahamas	0.0
14	Belize	0.0
34	Cyprus	0.0
64	Iceland	0.0
105	New Zealand	0.0
117	Portugal	0.0
129	Slovenia	0.0
139	Swaziland	0.0
147	Togo	0.0
160	Vietnam	0.0

Above countries are safest amongst all with zero terror attacks

Conclusion

1. Number of people killed by terror attack: 411868
2. Country with the highest number of Terror Attacks: Iraq (Maximum nO. of people were killed by a single terrorist attack are 1570.0 people that took place in Iraq)
3. Regions with the highest number of Terror Attacks: Middle East & North Africa
4. City with the most attacks: Baghdad
5. Year with the most attacks: 2014
6. Group with the most attacks: Taliban
7. Most Attack Types: Bombing/Explosion

In []: