THE SPARKS FOUNDATION INTERNSHIP

Task 4:Exploratory Data Analysis Global Terrorism

imports and loading of dataset:

```
In [58]: import pandas as pd
         import numpy as np # linear algebra
         import matplotlib.pyplot as plt
         import matplotlib.patches as mpatches
         import seaborn as sns
         import numpy as np
         import pandas as pd
         import numpy as np # linear algebra
         import matplotlib.pyplot as plt
         import matplotlib.patches as mpatches
         import seaborn as sns
         import numpy as np
         df = pd.read_csv(r"D:\[CN]_Data_Structure\globalterrorismdb_0718dist.csv",encoding="ISO-8859-1")
         print(df)
```

```
D:\anaconda 3\lib\site-packages\IPython\core\interactiveshell.py:3063: DtypeWarning: Columns (4,6,31,33,61,62,63,76,79,
90,92,94,96,114,115,121) have mixed types.Specify dtype option on import or set low_memory=False.
  interactivity=interactivity, compiler=compiler, result=result)
```

```
            eventid  iyear  imonth  iday approxdate  extended resolution  \
0      197000000001   1970       7     2        NaN         0        NaN
1      197000000002   1970       0     0        NaN         0        NaN
2      197001000001   1970       1     0        NaN         0        NaN
3      197001000002   1970       1     0        NaN         0        NaN
4      197001000003   1970       1     0        NaN         0        NaN
...             ...    ...     ...   ...        ...       ...        ...
181686 201712310022   2017      12    31        NaN         0        NaN
181687 201712310029   2017      12    31        NaN         0        NaN
181688 201712310030   2017      12    31        NaN         0        NaN
181689 201712310031   2017      12    31        NaN         0        NaN
181690 201712310032   2017      12    31        NaN         0        NaN

        country          country_txt  region  ... addnotes  \
0            58   Dominican Republic       2  ...      NaN
1           130               Mexico       1  ...      NaN
2           160          Philippines       5  ...      NaN
3            78               Greece       8  ...      NaN
4           101                Japan       4  ...      NaN
...         ...                  ...     ...  ...      ...
181686      182              Somalia      11  ...      NaN
181687      200                Syria      10  ...      NaN
```

```
181688      160        Philippines      5  ...      NaN
181689       92              India      6  ...      NaN
181690      160        Philippines      5  ...      NaN

                                              scite1  \
0                                                NaN
1                                                NaN
2                                                NaN
3                                                NaN
4                                                NaN
...                                              ...
181686  "Somalia: Al-Shabaab Militants Attack Army Che...
181687  "Putin's 'victory' in Syria has turned into a ...
181688  "Maguindanao clashes trap tribe members," Phil...
181689  "Trader escapes grenade attack in Imphal," Bus...
181690  "Security tightened in Cotabato following IED ...

                                              scite2  \
0                                                NaN
1                                                NaN
2                                                NaN
3                                                NaN
4                                                NaN
...                                              ...
181686  "Highlights: Somalia Daily Media Highlights 2 ...
181687  "Two Russian soldiers killed at Hmeymim base i...
181688                                           NaN
181689                                           NaN
181690  "Security tightened in Cotabato City," Manila ...

                                              scite3  \
0                                                NaN
1                                                NaN
2                                                NaN
3                                                NaN
4                                                NaN
...                                              ...
181686  "Highlights: Somalia Daily Media Highlights 1 ...
181687  "Two Russian servicemen killed in Syria mortar...
181688                                           NaN
181689                                           NaN
181690                                           NaN
```

```
                         dbsource  INT_LOG  INT_IDEO  INT_MISC  INT_ANY  related
0                            PGIS        0         0         0        0      NaN
1                            PGIS        0         1         1        1      NaN
2                            PGIS       -9        -9         1        1      NaN
3                            PGIS       -9        -9         1        1      NaN
4                            PGIS       -9        -9         1        1      NaN
...                           ...      ...       ...       ...      ...      ...
181686  START Primary Collection        0         0         0        0      NaN
181687  START Primary Collection       -9        -9         1        1      NaN
181688  START Primary Collection        0         0         0        0      NaN
181689  START Primary Collection       -9        -9         0       -9      NaN
181690  START Primary Collection       -9        -9         0       -9      NaN

[181691 rows x 135 columns]
```

In [59]: `df.head()`

Out[59]:

	eventid	iyear	imonth	iday	approxdate	extended	resolution	country	country_txt	region	...	addnotes	scite1	scite2	scite3	dbsource
0	197000000001	1970	7	2	NaN	0	NaN	58	Dominican Republic	2	...	NaN	NaN	NaN	NaN	PGIS
1	197000000002	1970	0	0	NaN	0	NaN	130	Mexico	1	...	NaN	NaN	NaN	NaN	PGIS
2	197001000001	1970	1	0	NaN	0	NaN	160	Philippines	5	...	NaN	NaN	NaN	NaN	PGIS
3	197001000002	1970	1	0	NaN	0	NaN	78	Greece	8	...	NaN	NaN	NaN	NaN	PGIS
4	197001000003	1970	1	0	NaN	0	NaN	101	Japan	4	...	NaN	NaN	NaN	NaN	PGIS

5 rows × 135 columns

```
In [60]: df.rename(columns={'iyear':'Year','imonth':'Month','city':'City','iday':'Day','country_txt':'Country','region_txt':'Regi
         df['Casualities'] = df.Killed + df.Wounded
         df=df[['Year','Month','Day','Country','Region','City','latitude','longitude','AttackType','Killed','Wounded','Casualitie
         
         df.head()
```

Out[60]:

	Year	Month	Day	Country	Region	City	latitude	longitude	AttackType	Killed	Wounded	Casualities	Target	G
0	1970	7	2	Dominican Republic	Central America & Caribbean	Santo Domingo	18.456792	-69.951164	Assassination	1.0	0.0	1.0	Julio Guzman	MAI
1	1970	0	0	Mexico	North America	Mexico city	19.371887	-99.086624	Hostage Taking (Kidnapping)	0.0	0.0	0.0	Nadine Chaval, daughter	2: Septe Comn Le
2	1970	1	0	Philippines	Southeast Asia	Unknown	15.478598	120.599741	Assassination	1.0	0.0	1.0	Employee	Unk
3	1970	1	0	Greece	Western Europe	Athens	37.997490	23.762728	Bombing/Explosion	NaN	NaN	NaN	U.S. Embassy	Unk
4	1970	1	0	Japan	East Asia	Fukouka	33.580412	130.396361	Facility/Infrastructure Attack	NaN	NaN	NaN	U.S. Consulate	Unk

In [61]:
```python
df.isnull().sum()
```

Out[61]:
```
Year                 0
Month                0
Day                  0
Country              0
Region               0
City               434
latitude          4556
longitude         4557
AttackType           0
Killed           10313
Wounded          16311
Casualities      16874
Target             636
Group                0
Target_type          0
Weapon_type          0
dtype: int64
```

Most Attack-prone regions of the world:

In [62]:
```python
print(f"The highest terrorist attacks were commited in {df.Country.value_counts().index[0]} with {df.Country.value_count

print('Top 4 countries with highest terrorist attacks following iraq are:')
for i in range(1,5):
    print(f"{i+1}. {df.Country.value_counts().index[i]} with {df.Country.value_counts()[i]} attacks")
```

```
The highest terrorist attacks were commited in Iraq with 24636 attacks
Top 4 countries with highest terrorist attacks following iraq are:
2. Pakistan with 14368 attacks
3. Afghanistan with 12731 attacks
4. India with 11960 attacks
5. Colombia with 8306 attacks
```

Timeline of Terror attacks,over the past 45 years:

In [63]:
```python
plt.subplots(figsize=(15,3))
sns.countplot('Year',data=df,palette='YlOrBr')
plt.xticks(rotation=90)
plt.title('Number Of Terrorist Attacks According to Each Year(1970-2017)')
plt.show()
```
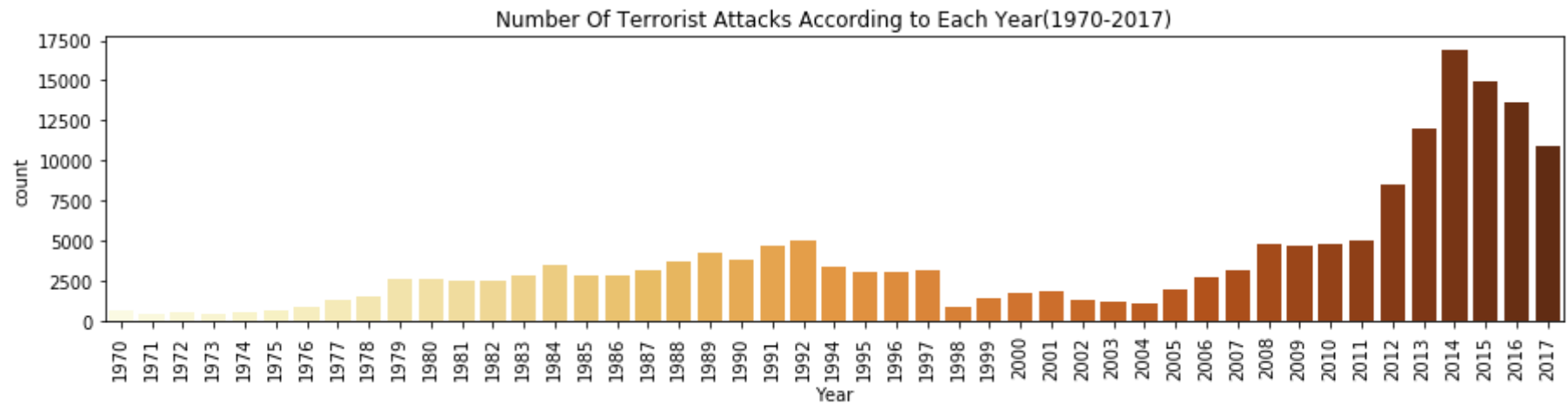

Timeline of Death:

In [64]:
```python
plt.subplots(figsize=(15,5))
year_cas = df.groupby('Year').Casualities.sum().to_frame().reset_index()
year_cas.columns = ['Year','Casualities']
sns.barplot(x=year_cas.Year, y=year_cas.Casualities, palette='RdYlBu')
plt.xticks(rotation=80)
plt.title('Number Of Deaths Each Year')
plt.show()
```


Number Of Deaths Each Year
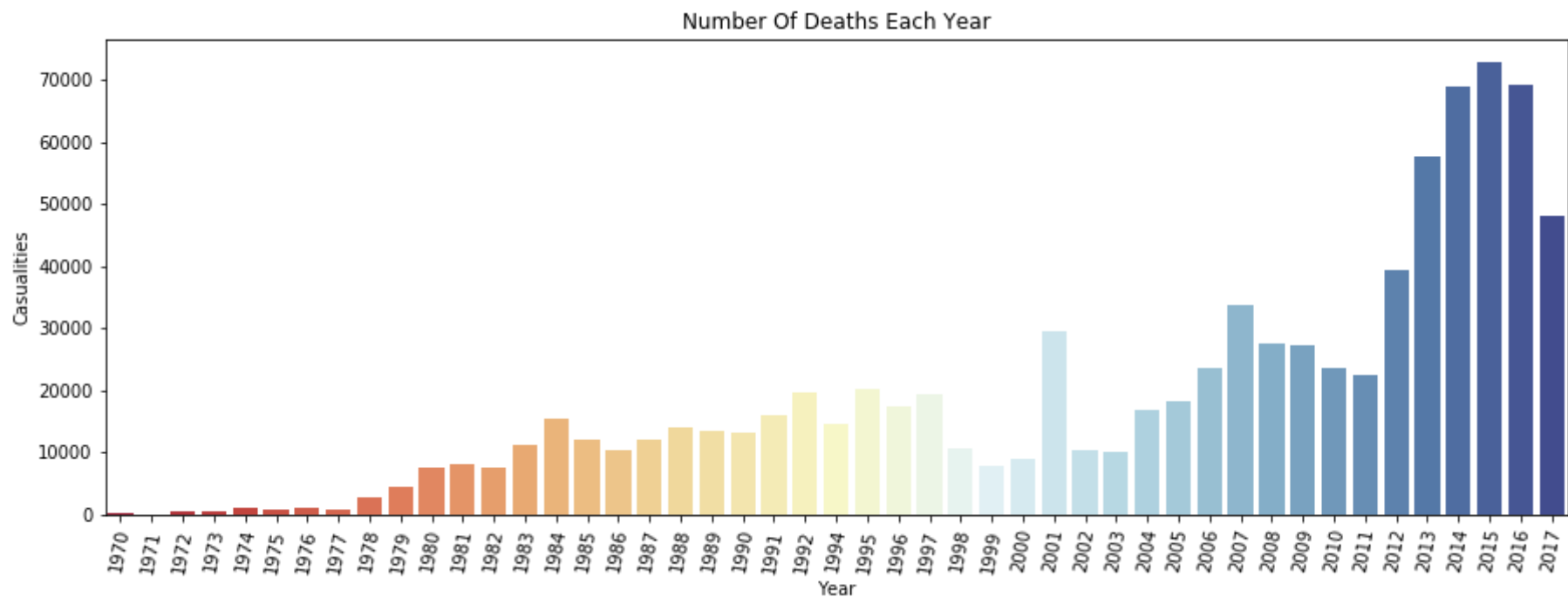
Nation-wise total attacks:

```
In [65]: plt.subplots(figsize=(15,3))
country_attacks = df.Country.value_counts()[:15].reset_index()
country_attacks.columns = ['Country', 'Total Attacks']
sns.barplot(x=country_attacks.Country, y=country_attacks['Total Attacks'], palette= 'YlGnBu',edgecolor=sns.color_palette
plt.xticks(rotation=50)
plt.title('Number Of Total Attacks in Each Nation')
plt.show()
```
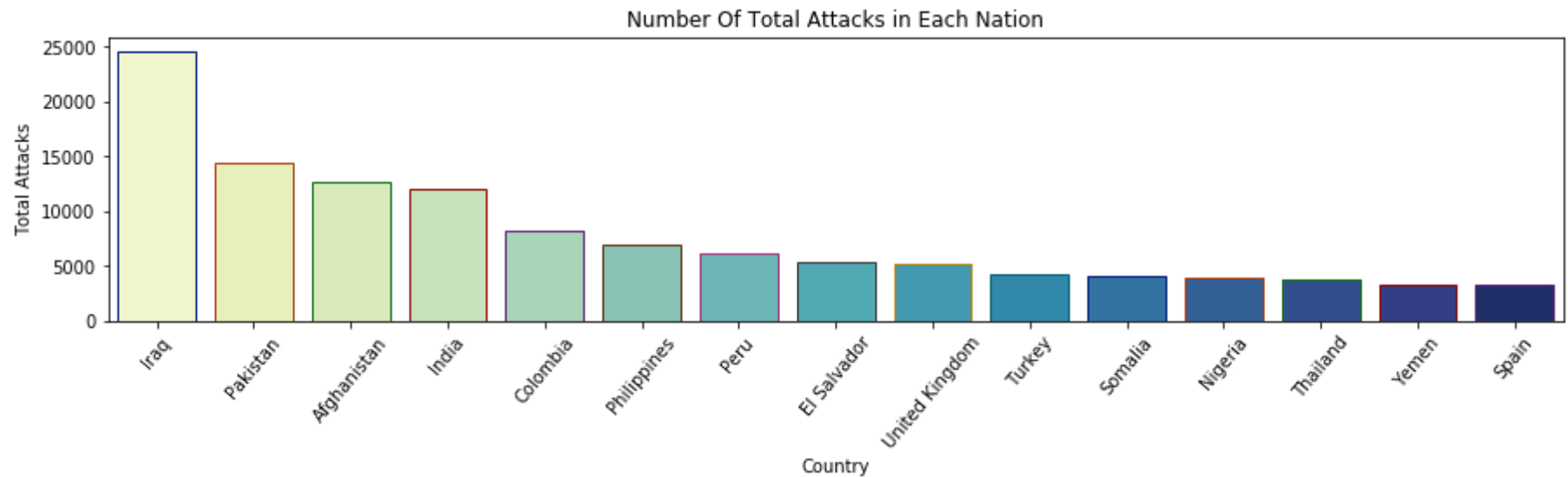

Terrain wise attack:

In [66]:
```python
region_attacks = df.Region.value_counts().to_frame().reset_index()
region_attacks.columns = ['Region', 'Total Attacks']
plt.subplots(figsize=(15,8))
sns.barplot(x=region_attacks.Region, y=region_attacks['Total Attacks'], palette='Accent', edgecolor=sns.color_palette('d
plt.xticks(rotation=46)
plt.title('Number Of Total Attacks in Each Terrrain')
plt.show()
```

In [67]:
```python
group_attacks = df.Group.value_counts().to_frame().drop('Unknown').reset_index()[:16]
group_attacks.columns = ['Terrorist Organizations', 'Total Attacks']
group_attacks
```

Out[67]:

	Terrorist Organizations	Total Attacks
0	Taliban	7478
1	Islamic State of Iraq and the Levant (ISIL)	5613
2	Shining Path (SL)	4555
3	Farabundo Marti National Liberation Front (FMLN)	3351
4	Al-Shabaab	3288
5	New People's Army (NPA)	2772
6	Irish Republican Army (IRA)	2671
7	Revolutionary Armed Forces of Colombia (FARC)	2487
8	Boko Haram	2418
9	Kurdistan Workers' Party (PKK)	2310
10	Basque Fatherland and Freedom (ETA)	2024
11	Communist Party of India - Maoist (CPI-Maoist)	1878
12	Maoists	1630
13	Liberation Tigers of Tamil Eelam (LTTE)	1606
14	National Liberation Army of Colombia (ELN)	1561
15	Tehrik-i-Taliban Pakistan (TTP)	1351

```
In [68]: group_attacks = df.Group.value_counts().to_frame().drop('Unknown').reset_index()[:16]
         group_attacks.columns = ['Terrorist Group', 'Total Attacks']
         plt.subplots(figsize=(10,6))
         sns.barplot(y=group_attacks['Terrorist Group'], x=group_attacks['Total Attacks'], palette='Pastel2',
                     )

         plt.title('Number Of Total Attacks by Terrorist Group')
         plt.show()
```
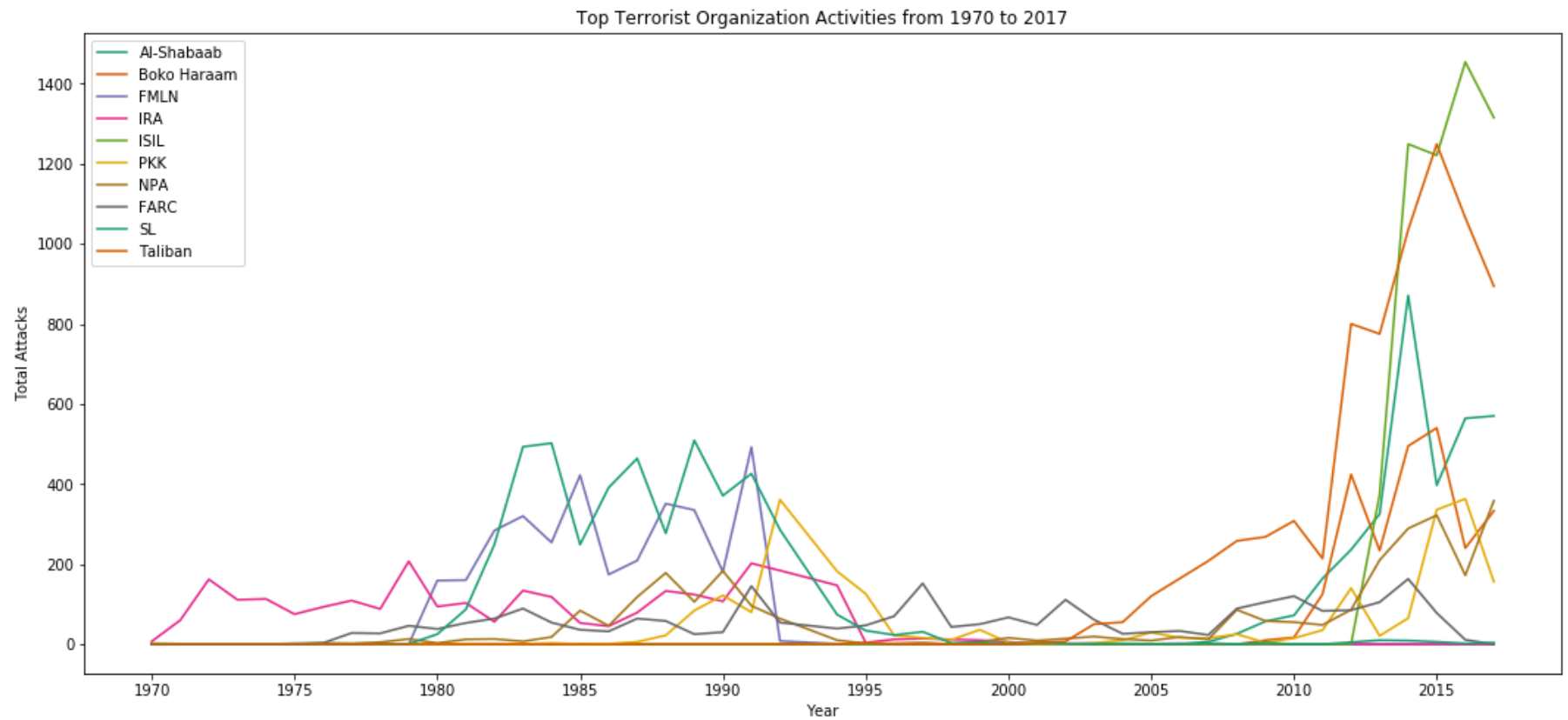

Number Of Total Attacks by Terrorist Group
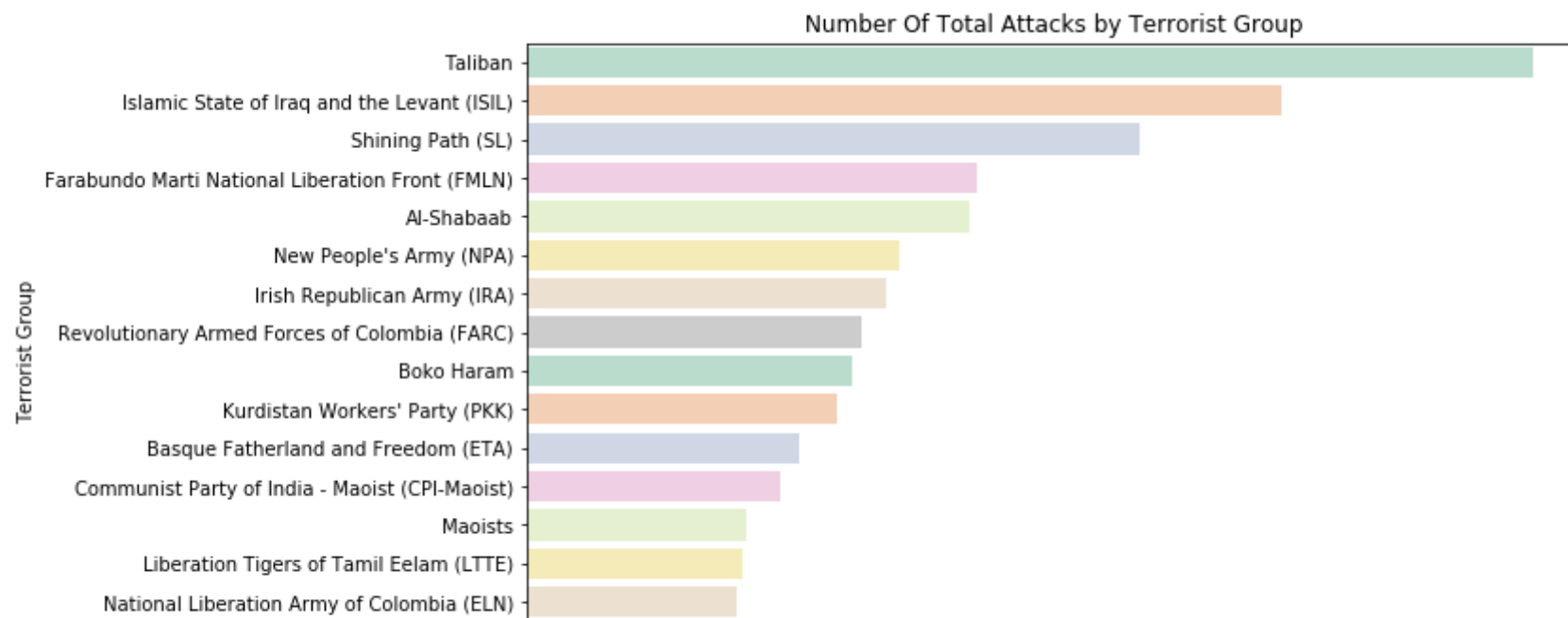
Terrorist Organizations Attack Trend

In [69]:
```python
groups_10 = df[df.Group.isin(df.Group.value_counts()[1:11].index)]
pd.crosstab(groups_10.Year, groups_10.Group).plot(color=sns.color_palette('Dark2', 10))
fig=plt.gcf()
fig.set_size_inches(18,8)
plt.xticks(range(1970, 2017, 5))
plt.ylabel('Total Attacks')
plt.title('Top Terrorist Organization Activities from 1970 to 2017')
plt.legend(labels=['Al-Shabaab',
                   'Boko Haraam',
                   'FMLN',
                   'IRA',
                   'ISIL',
                   'PKK',
                   'NPA',
                   'FARC',
                   'SL',
                   'Taliban'], loc='upper left')
plt.show()
```

In [ ]: