task_1(2)

July 23, 2022

1 Terrorism Analysis

1.0.1 importing header files

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```
data = pd.read_csv('terrorism.csv',  encoding="ISO-8859-1", low_memory=False)
```

[3]:
```
data.head()
```

[3]:
```
        eventid  iyear  imonth  iday approxdate  extended resolution  country  \
0  197000000001   1970       7     2        NaN         0        NaN       58
1  197000000002   1970       0     0        NaN         0        NaN      130
2  197001000001   1970       1     0        NaN         0        NaN      160
3  197001000002   1970       1     0        NaN         0        NaN       78
4  197001000003   1970       1     0        NaN         0        NaN      101

          country_txt  region  … addnotes scite1 scite2  scite3  dbsource  \
0  Dominican Republic       2  …      NaN    NaN    NaN     NaN      PGIS
1              Mexico       1  …      NaN    NaN    NaN     NaN      PGIS
2         Philippines       5  …      NaN    NaN    NaN     NaN      PGIS
3              Greece       8  …      NaN    NaN    NaN     NaN      PGIS
4               Japan       4  …      NaN    NaN    NaN     NaN      PGIS

   INT_LOG  INT_IDEO INT_MISC INT_ANY  related
0        0         0        0       0      NaN
1        0         1        1       1      NaN
2       -9        -9        1       1      NaN
3       -9        -9        1       1      NaN
4       -9        -9        1       1      NaN

[5 rows x 135 columns]
```

[4]:
```
data.tail()
```

[4]:
```
             eventid  iyear  imonth  iday approxdate  extended resolution  \
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  \
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  \
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  \
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  \
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
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

[5 rows x 135 columns]
```

Renaming columns correctly

[5]:

```
data.rename(columns={'iyear':'year','imonth':'month','extended':
 ↪'extended','iday':'day','country_txt':'country_name','provstate':
 ↪'state','region_txt':'region_name','attacktype1_txt':'attacktype','target1':
 ↪'target','nkill':'killed','nwound':'wounded','summary':'summary','gname':
 ↪'group','targetype1_txt':'target_type','weaptype1_txt':
 ↪'weapon_type','motive':'motive','city':'city','latitude':
 ↪'latitude','longitude':'longitude'},inplace=1)
```

[6]:
```
data.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: 187.1+ MB
```

[7]:
```
data.head()
```

[7]:

	eventid	year	month	day	approxdate	extended	resolution	country	\
0	197000000001	1970	7	2	NaN	0	NaN	58	
1	197000000002	1970	0	0	NaN	0	NaN	130	
2	197001000001	1970	1	0	NaN	0	NaN	160	
3	197001000002	1970	1	0	NaN	0	NaN	78	
4	197001000003	1970	1	0	NaN	0	NaN	101	

	country_name	region	…	addnotes	scite1	scite2	scite3	dbsource	\
0	Dominican Republic	2	…	NaN	NaN	NaN	NaN	PGIS	
1	Mexico	1	…	NaN	NaN	NaN	NaN	PGIS	
2	Philippines	5	…	NaN	NaN	NaN	NaN	PGIS	
3	Greece	8	…	NaN	NaN	NaN	NaN	PGIS	
4	Japan	4	…	NaN	NaN	NaN	NaN	PGIS	

	INT_LOG	INT_IDEO	INT_MISC	INT_ANY	related
0	0	0	0	0	NaN
1	0	1	1	1	NaN
2	-9	-9	1	1	NaN
3	-9	-9	1	1	NaN
4	-9	-9	1	1	NaN

[5 rows x 135 columns]

[8]:
```
data.shape
```

[8]:
```
(181691, 135)
```

[9]:
```
print(data.keys())
```

```
Index(['eventid', 'year', 'month', 'day', 'approxdate', 'extended',
       'resolution', 'country', 'country_name', 'region',
```

```
       …
       'addnotes', 'scite1', 'scite2', 'scite3', 'dbsource', 'INT_LOG',
       'INT_IDEO', 'INT_MISC', 'INT_ANY', 'related'],
      dtype='object', length=135)
```

[10]:
```
data.describe
```

[10]:
```
<bound method NDFrame.describe of              eventid  year  month  day
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_name  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]>
```

1.1 Null checking

[11]:
```
data.isnull().sum()
```

[11]:
```
eventid            0
year               0
```

```
month              0
day                0
approxdate    172452
                …
INT_LOG            0
INT_IDEO           0
INT_MISC           0
INT_ANY            0
related       156653
Length: 135, dtype: int64
```

[12]:
```
print(data['country'])
```

```
0          58
1         130
2         160
3          78
4         101
         …
181686    182
181687    200
181688    160
181689     92
181690    160
Name: country, Length: 181691, dtype: int64
```

1.1.1 Finding "Unknown"s in group name

Thease must be eliminated

[13]:
```
len(data[data['group']=="Unknown"])/len(data)*100
```

[13]:
```
45.561970598433604
```

[14]:
```
len(data['group'].unique())
```

[14]:
```
3537
```

2 World

2.1 Country, State, City

[15]:
```
print(data['country_name'].value_counts().head(10))
```

```
Iraq              24636
Pakistan          14368
Afghanistan       12731
India             11960
Colombia           8306
Philippines        6908
```

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

[16]:
```
print(data['state'].value_counts().head(10))
```

```
Baghdad               7645
Northern Ireland      4498
Unknown               4290
Balochistan           3710
Saladin               3411
Al Anbar              3299
Nineveh               3241
Sindh                 3206
Khyber Pakhtunkhwa    3084
Diyala                3041
Name: state, dtype: int64
```

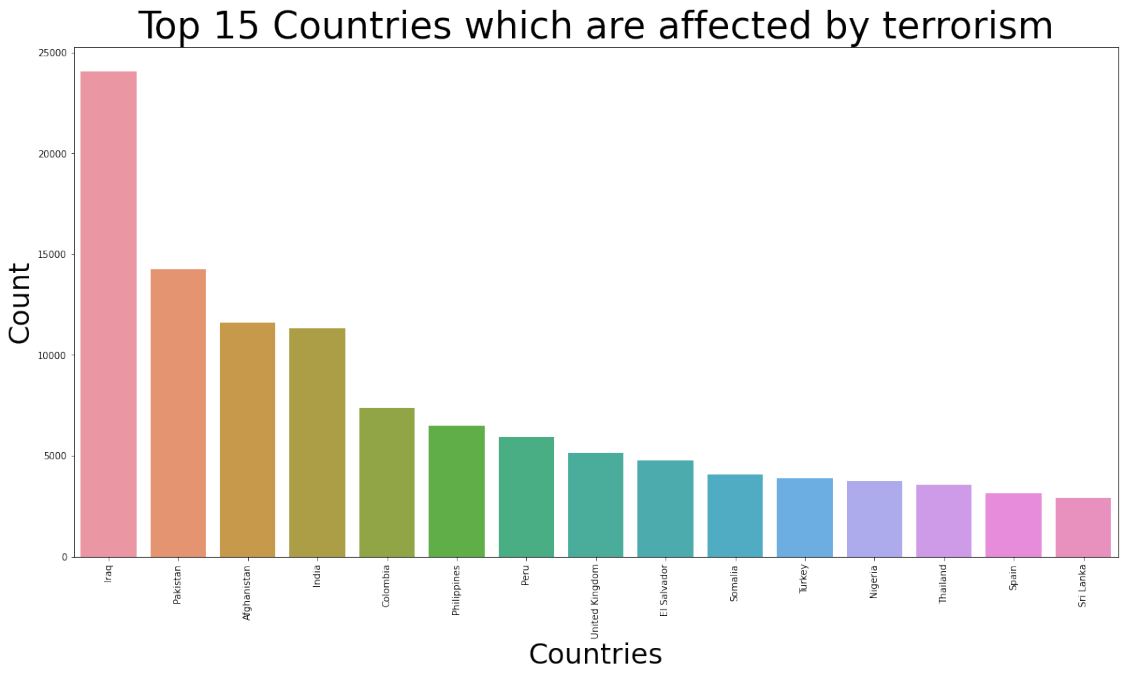
[17]:
```
data = data[data['city']!='Unknown']
##removed unknown from city
print(data['city'].value_counts().head(10))
```

```
Baghdad         7589
Karachi         2652
Lima            2359
Mosul           2265
Belfast         2171
Santiago        1621
Mogadishu       1581
San Salvador    1558
Istanbul        1048
Athens          1019
Name: city, dtype: int64
```

[18]:
```
plt.figure(figsize = (20, 10))
sns.barplot(data['country_name'].value_counts()[:15].index,␣
 ↪data['country_name'].value_counts()[:15].values)
plt.title('Top 15 Countries which are affected by terrorism', fontsize = 40)
plt.xlabel('Countries', fontsize = 30)
plt.ylabel('Count', fontsize = 30)
plt.xticks(rotation=90)
plt.show()
```

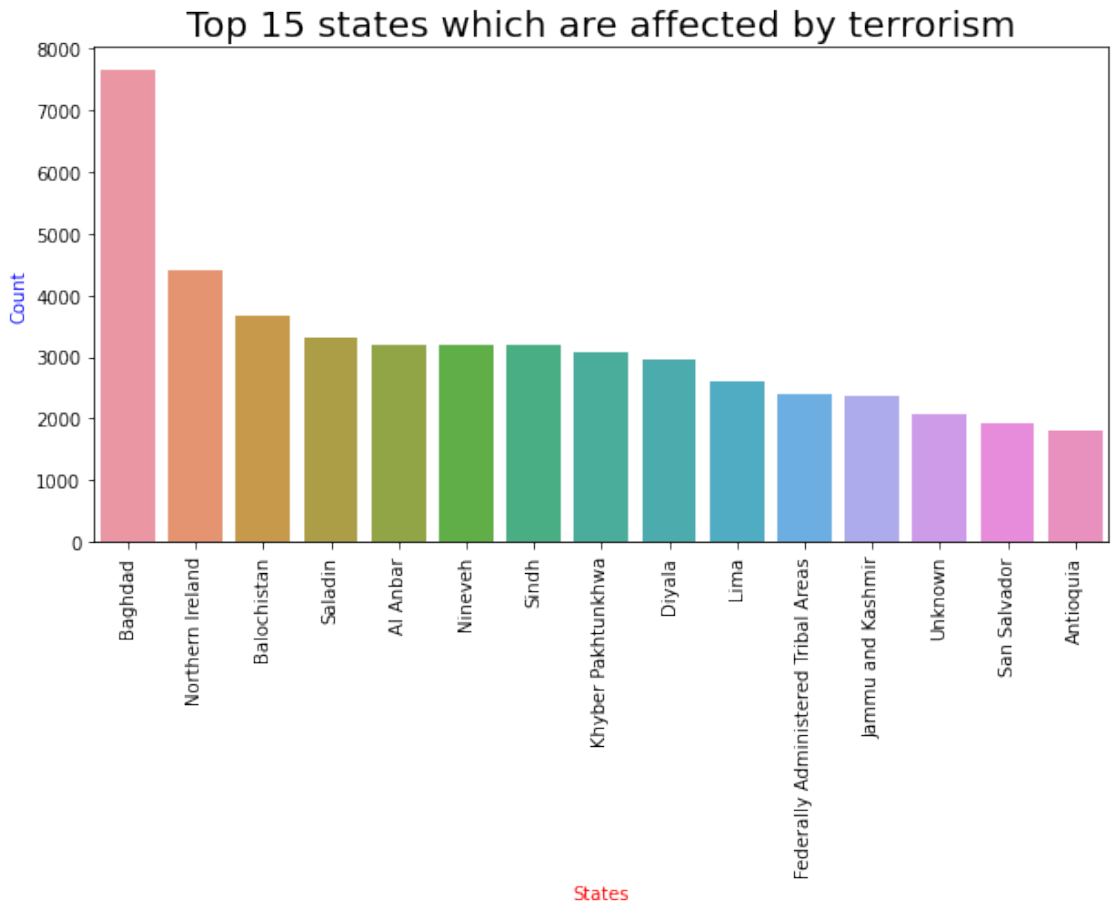
```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
```

```
result in an error or misinterpretation.
  warnings.warn(
```



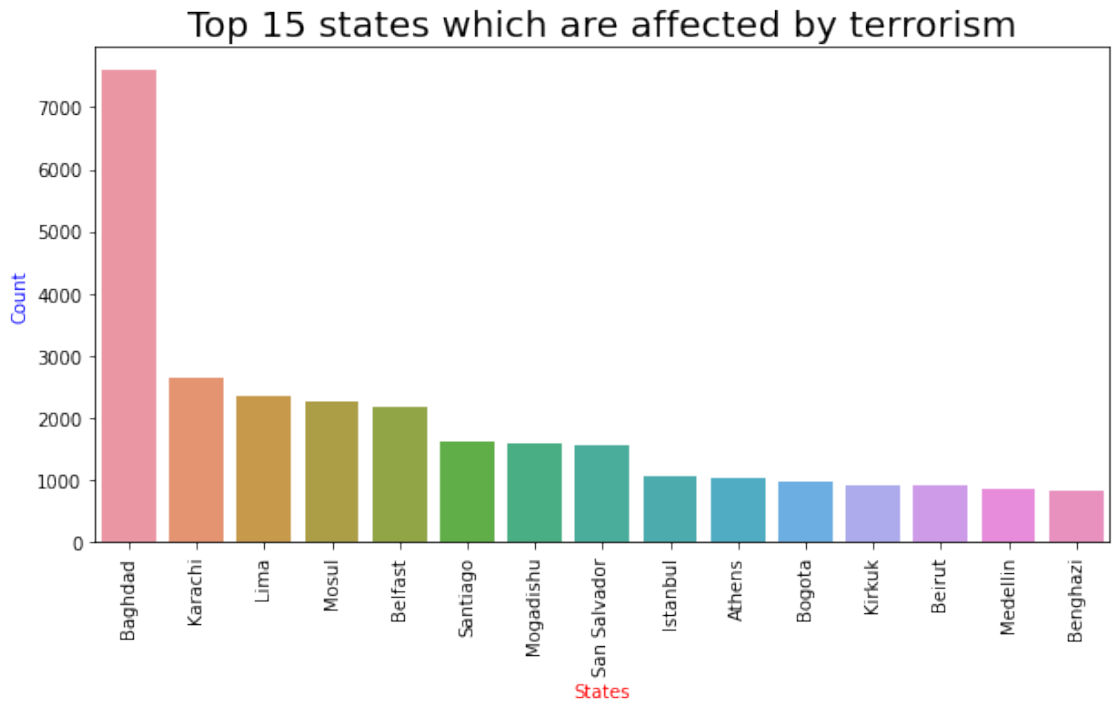
[19]:
```
plt.figure(figsize = (10, 5))
sns.barplot(data['state'].value_counts()[:15].index, data['state'].
  ↪value_counts()[:15].values)
plt.title('Top 15 states which are affected by terrorism', fontsize = 20)
plt.xlabel('States', fontsize = 10, color ='red')
plt.ylabel('Count', fontsize = 10, color ='blue')
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



[20]:
```
plt.figure(figsize = (10, 5))
sns.barplot(data['city'].value_counts()[:15].index, data['city'].
 ↪value_counts()[:15].values)
plt.title('Top 15 states which are affected by terrorism', fontsize = 20)
plt.xlabel('States', fontsize = 10, color ='red')
plt.ylabel('Count', fontsize = 10, color ='blue')
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



3 India

3.1 State wise analysis

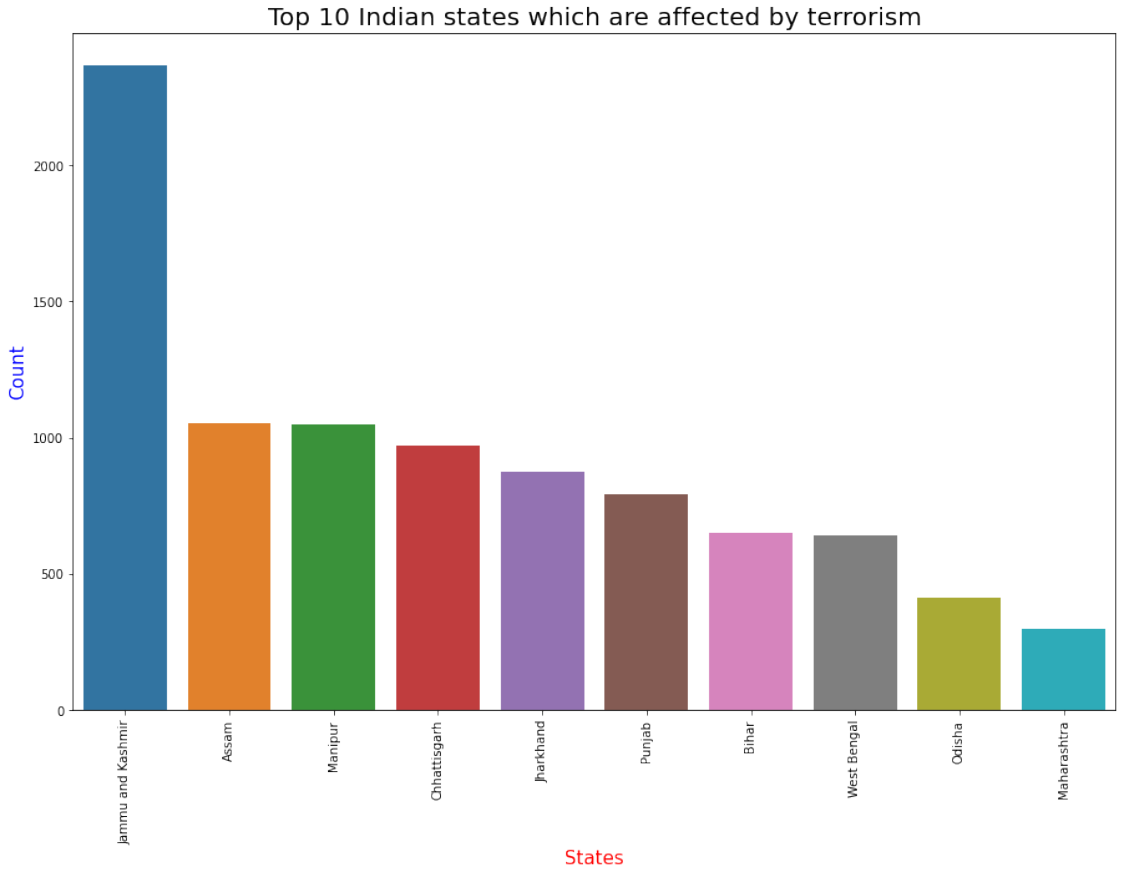
[21]:
```
data_state = data[data['country_name']=='India']['state']
data_state.value_counts()[:10]
```

[21]:
```
Jammu and Kashmir    2368
Assam                1054
Manipur              1049
Chhattisgarh          970
Jharkhand             873
Punjab                794
Bihar                 652
West Bengal           643
Odisha                414
Maharashtra           299
Name: state, dtype: int64
```

[22]:
```
plt.figure(figsize = (15, 10))
sns.barplot(data_state.value_counts()[:10].index, data_state.value_counts()[:
 ↪10].values)
plt.title('Top 10 Indian states which are affected by terrorism', fontsize = 20)
plt.xlabel('States', fontsize = 15, color ='red')
```

```
plt.ylabel('Count', fontsize = 15, color ='blue')
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



3.2 Regions wise analysis

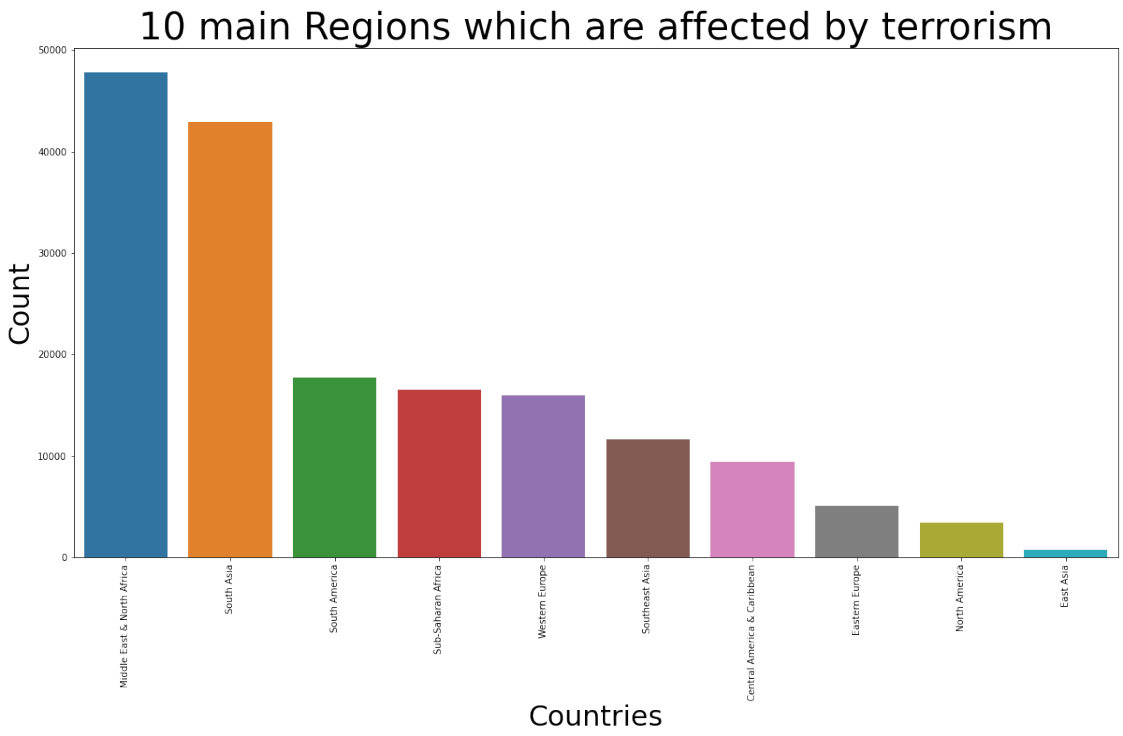
[23]:
```
print(data['region_name'].value_counts().head(10))
```

```
Middle East & North Africa     47809
South Asia                     42936
South America                  17744
Sub-Saharan Africa             16554
Western Europe                 15936
```

```
Southeast Asia                  11617
Central America & Caribbean      9361
Eastern Europe                   5026
North America                    3398
East Asia                         723
Name: region_name, dtype: int64
```

[24]:
```
plt.figure(figsize = (20, 10))
sns.barplot(data['region_name'].value_counts()[:10].index, data['region_name'].
 ↪value_counts()[:10].values)
plt.title('10 main Regions which are affected by terrorism', fontsize = 40)
plt.xlabel('Countries', fontsize = 30)
plt.ylabel('Count', fontsize = 30)
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



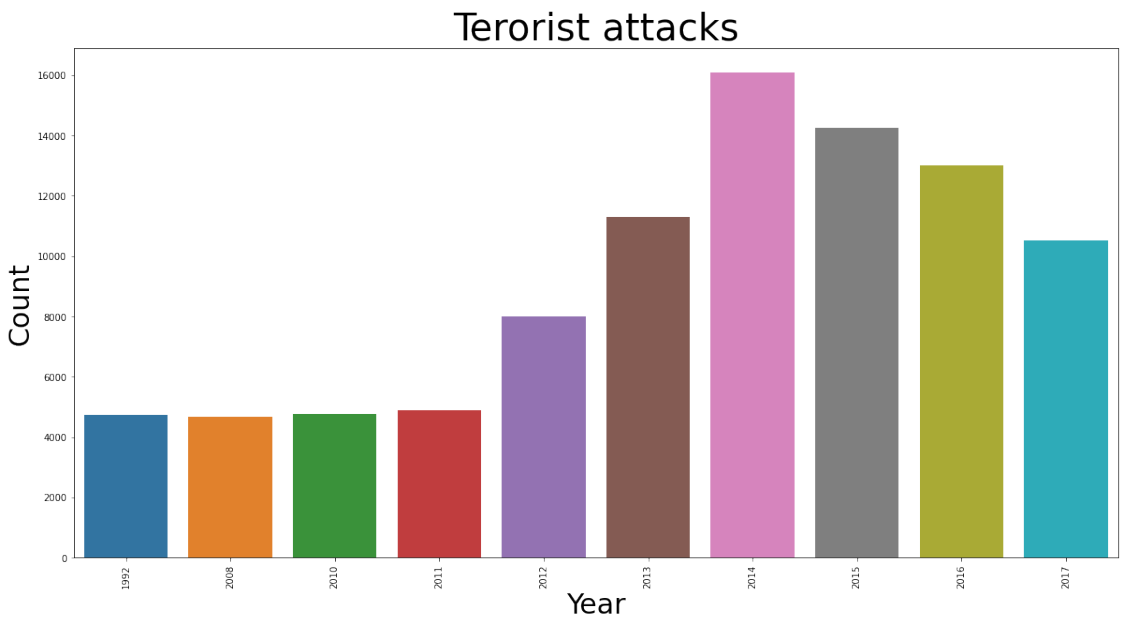
3.2.1 Years with most terrorist activities

[25]:
```
print(data['year'].value_counts().head(15))
```

```
2014    16103
2015    14243
2016    12999
2013    11311
2017    10536
2012     8008
2011     4906
2010     4766
1992     4730
2008     4688
2009     4667
1991     4313
1989     4125
1990     3666
1988     3553
Name: year, dtype: int64
```

[26]:
```
plt.figure(figsize = (20, 10))
sns.barplot(data['year'].value_counts()[:10].index, data['year'].
 ↪value_counts()[:10].values)
plt.title('Terorist attacks', fontsize = 40)
plt.xlabel('Year', fontsize = 30)
plt.ylabel('Count', fontsize = 30)
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



4 Most involved terrorist groups

[27]:
```
print(data['group'].value_counts().head(15))
```

```
Unknown                                                79026
Taliban                                                 7016
Islamic State of Iraq and the Levant (ISIL)             5438
Shining Path (SL)                                       4425
Al-Shabaab                                              3201
Farabundo Marti National Liberation Front (FMLN)        2888
Irish Republican Army (IRA)                             2592
New People's Army (NPA)                                 2547
Boko Haram                                              2356
Revolutionary Armed Forces of Colombia (FARC)           2155
Kurdistan Workers' Party (PKK)                          1993
Basque Fatherland and Freedom (ETA)                     1949
Communist Party of India - Maoist (CPI-Maoist)          1849
Maoists                                                 1607
Liberation Tigers of Tamil Eelam (LTTE)                 1568
Name: group, dtype: int64
```

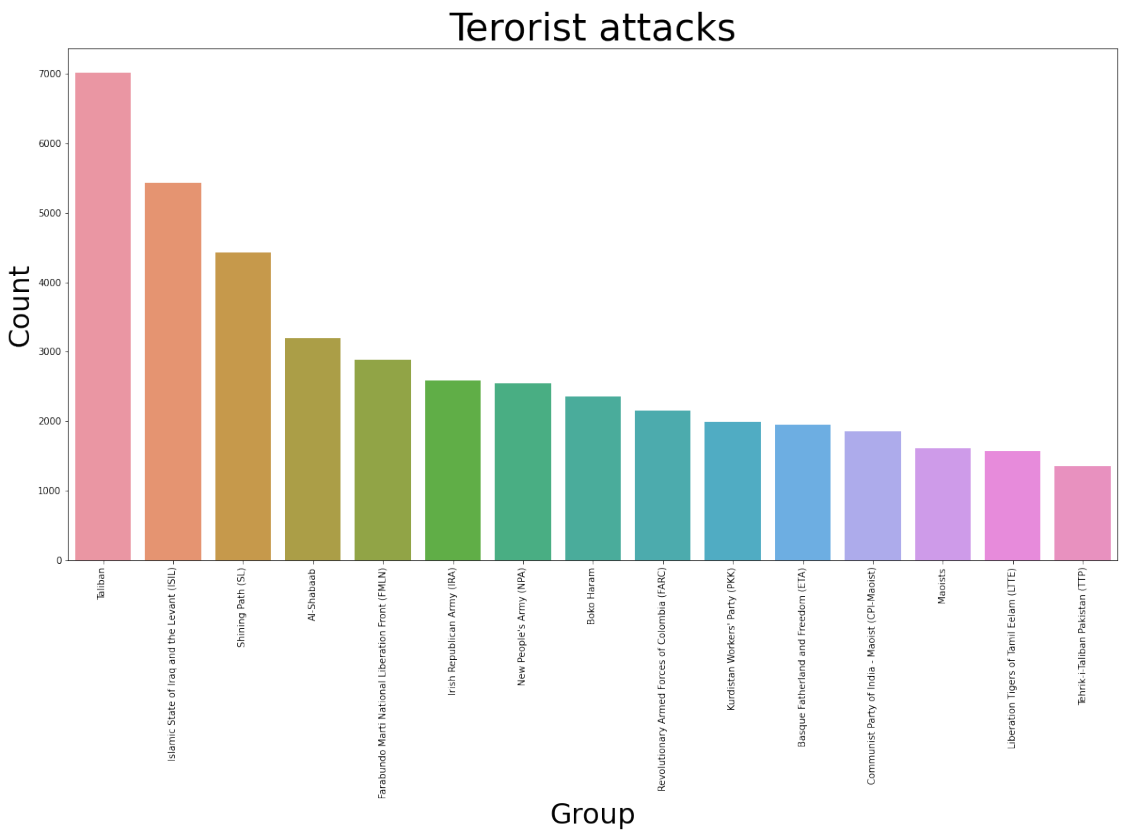
4.0.1 Data shows Unknown groups are most involved

[28]:
```
data = data[data['group']!='Unknown']
print(data['group'].value_counts().head(15))
```

```
Taliban                                                    7016
Islamic State of Iraq and the Levant (ISIL)                5438
Shining Path (SL)                                          4425
Al-Shabaab                                                 3201
Farabundo Marti National Liberation Front (FMLN)           2888
Irish Republican Army (IRA)                                2592
New People's Army (NPA)                                    2547
Boko Haram                                                 2356
Revolutionary Armed Forces of Colombia (FARC)              2155
Kurdistan Workers' Party (PKK)                             1993
Basque Fatherland and Freedom (ETA)                        1949
Communist Party of India - Maoist (CPI-Maoist)             1849
Maoists                                                    1607
Liberation Tigers of Tamil Eelam (LTTE)                    1568
Tehrik-i-Taliban Pakistan (TTP)                            1347
Name: group, dtype: int64
```

[29]:
```
plt.figure(figsize = (20, 10))
sns.barplot(data['group'].value_counts()[:15].index, data['group'].
 ↪value_counts()[:15].values)
plt.title('Terorist attacks', fontsize = 40)
plt.xlabel('Group', fontsize = 30)
plt.ylabel('Count', fontsize = 30)
plt.xticks(rotation=90)
plt.show()
```

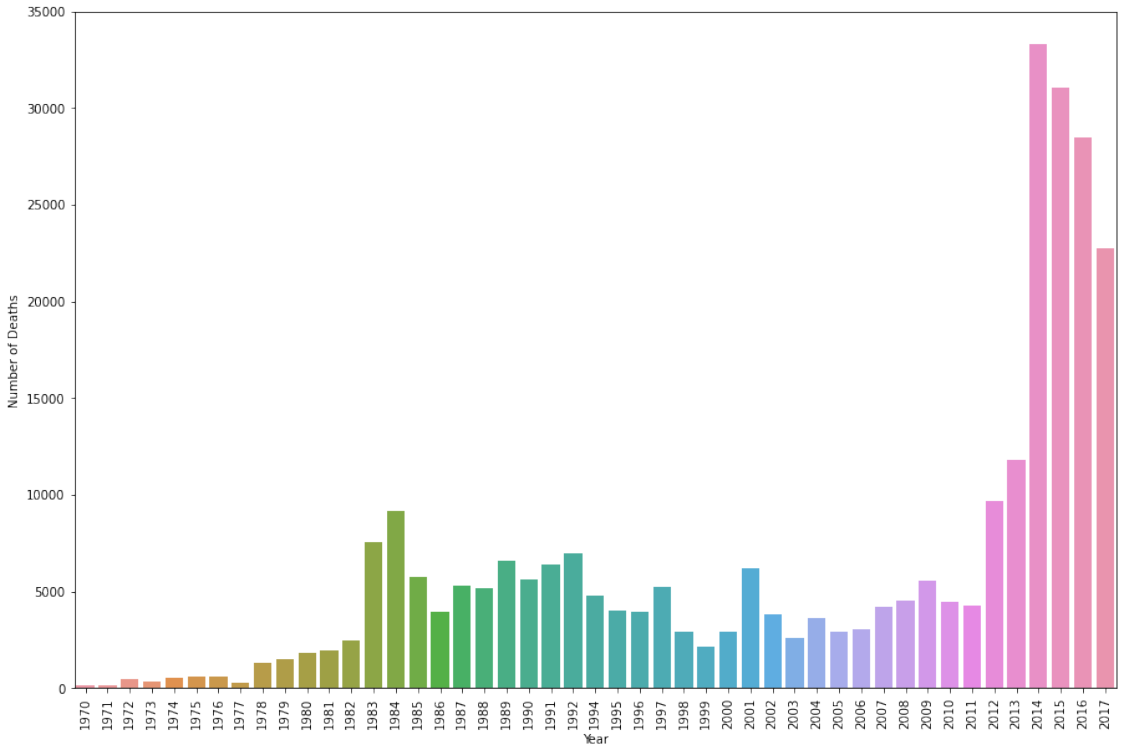
```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



4.0.2 Number of deaths

[30]:
```
data['killed'].fillna(0)
no_of_kills = data.groupby('year')['killed'].sum()
plt.subplots(figsize=(15,10))
plot1 = sns.barplot(no_of_kills.index,no_of_kills)
plot1.set_xlabel("Year")
plot1.set_ylabel("Number of Deaths")
plot1.set_xticklabels(no_of_kills.index,rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



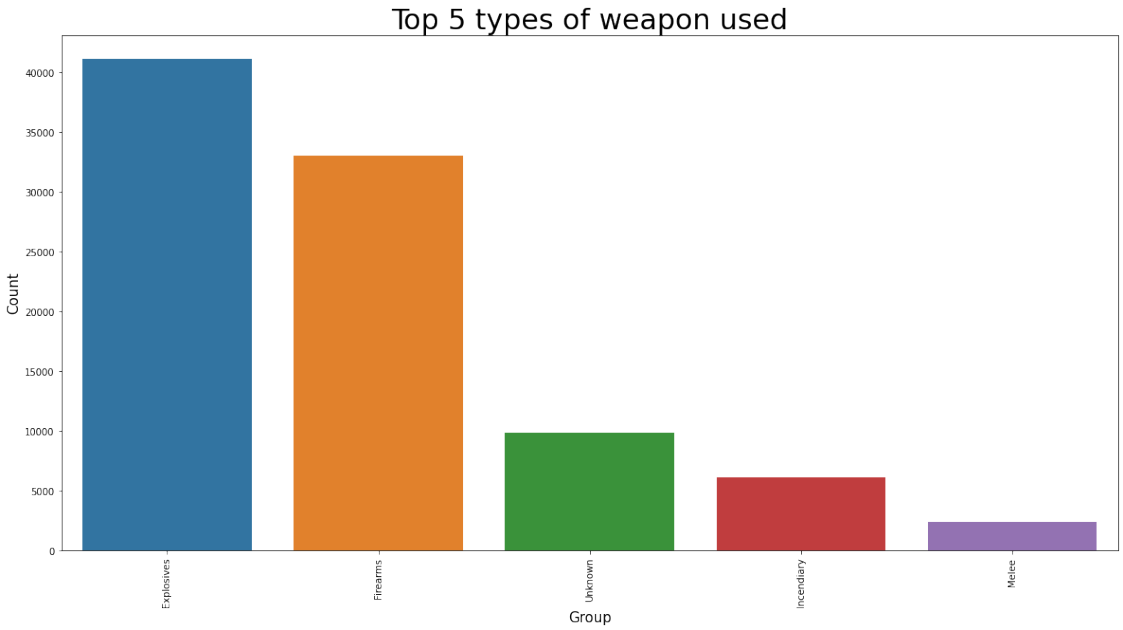
[31]:
```
print(data['weapon_type'].value_counts().head(15))
```

```
Explosives
41068
Firearms
33031
Unknown
9852
Incendiary
6132
Melee
2374
Chemical
136
Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)
104
Sabotage Equipment
83
Other
80
Fake Weapons
16
Biological
```

```
11
Radiological
3
Name: weapon_type, dtype: int64
```

[32]:
```
plt.figure(figsize = (20, 10))
sns.barplot(data['weapon_type'].value_counts()[:5].index, data['weapon_type'].
 ↪value_counts()[:5].values)
plt.title('Top 5 types of weapon used', fontsize = 30)
plt.xlabel('Group', fontsize = 15)
plt.ylabel('Count', fontsize = 15)
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Sushan Shivagiri\AppData\Local\Programs\Python\Python310\lib\site-
packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables
as keyword args: x, y. From version 0.12, the only valid positional argument
will be `data`, and passing other arguments without an explicit keyword will
result in an error or misinterpretation.
  warnings.warn(
```



[ ]: