Project Name - GLOBAL TERRORISM

Project Type - EDA

Contribution - Individual

Project By - Pratiksha Akshay Thorat

Project Summary -

Why I choose this project

I Choose this project to Examine a variety of theoretical and empirical materials needed for analysing pressing questions, relating to war, security and peace.To gain a solid understanding of contemporary security challenges, from environmental degradation to gender based insecurity, and the ability to analyse these issues surrounding security and terrorism.

What is Terrorism

The term "Terrorism" is derived from Latin word "Terror", whic rerfers to "great fear". The word "Terrorism" was first used during French Revolution in the yera 1795. The term was used to refer to intentional or planned use of brutality and violence to create an environment of fright, panic, distress and fear in general public with the sole purpose of establishing a certain political or social narrative. In todays world terrorism is used by different parties in different perspective under different circumstances. Though UN Security Council recognises Terrorism as a threat to pease and security, but fails to define terrorism in any of it resolutions and urges member countries to define Terrorism in their respective national law. Acoordingly different countries define Terrorism & Terrorist differently in their Laws.

How Do Countries Define Terrorism?

US Government/FBI defines and classifies terrorism as below:

International Terrorism:

Violent, criminal acts committed by individuals and/or groups who are inspired by, or associated with , designated foreign terrorist organizations or nations(state-sponsored). We have used Matplotlib and Seaborn libraries to represent our insights meaningfully and draw conclusions.

Through this project we aim to analyse the following:

What the hot zones of terrorism? What causes the strained relationship between countries like Pakistan and Afghanistan? How dreadful were the Boko Harams in Nigeria? How safe id Mumbai after 26/11/2008? Study the patterns of Naxalism in the Indian states. What all security issues and insights you can derive by EDA?

This global Terrorism dataset has the record of attacks in between 1970 to 2017

Lets describe few of the columns which we have considered for our analysis.

In this Global Terrorisim dataset the column iyear represents the year of attack in between 1970-2017, imonth represents the month of the attack & idate represents the day of the terrorist attacks in world wide.

The column country is the country code of the perticular county and the column country_txt shows the corresponding country name.

The column region shows the region code of the targetted region and the column region_txt shows the corresponding region name.

The column provstate shows the impacted state of the counry, city shows the list of cities of the state which have under attack by the terrorist.

Then we have the columns latitude and longitude containing data of the geographical coordinates of the attacks. The column vicinity shows the vicinity of the attacks.

The column location shows the location of the attacks or the place where the terrorists attacked , and the column summary briefs the date of the attack, type of the attack , number of terrorists, type of weapons used, targets and the damages.

The column alternative contains data of the no. of alternative attacks of the the day are made in the terrorist attacks , & the column alternative_txt that shows the type of alternatve attacks in the perticular day in which the terrorist are attacked.

The column success has either 0 or 1 showing success or failure of the attack , and the column sucide contains number of terrorist who died if it was a suicide attack.

Then we have the columns attacktype1 , attacktype2 , attacktype3 which show the code of type of the attack .The column attacktype1_txt shows the text corresponding to the attack type(Assassination,Hostage Taking/Kidnapping,Bombing/Explosion..etc.)

The columns targtype1 shows the code for the type of targets, and the column targtype1_txt shows the text correspoding to targtype1.

The column gname shows the name of the terrorist outfit. The column motive is to show the motive behind the attack.

The column individual shows whether the terrorist carried out the attack in a team or a group.

The columns weaptype1 , weaptype2 ,weaptype3 , weapsubtype1 , weaptype2 , weaptype3 are for the weapon types used in the terrorist attacks, and the columns weaptype1_txt , weaptype2_txt , weaptype3_txt , weapsubtype1_txt , weapsubtype2_txt , weapsubtype3_txt are for the corresponding name of the weapon used (Gasoline or Alcohol,Explosives,Incendiary,Pipe Bombetc.) The column weapdetail provides additional details of the weapon used

The columns ransom,ransompaid contain the randsome amount demand(if any) and how much the authorith have paid(if any).

We as a team believe that meaningful dialogues,inclusive state policies,addressing problems at the root cause can protect us all from the cowardly and ghastly practice of terror mongering called TERRORISM. We as future data scientists and analysts have tried to understand the underlying causes and impacts of these unhumanly acts

GitHub Link -

https://github.com/Pratikshathorat96/EDA-1-Poject-On-Global-Terrorism/blob/main/Pratiksha_Thorat_Capstone_EDA_Submission_on_Global_Terrorism.ipynb

Problem Statement

To the promotion and protection of human rights As a security/defence analyst, try to find out the hot zone of terrorism and finding weapons types used by terrorist.Also finding most active terrorist groups. What all security issues and insights you can derive by EDA?

Define Your Business Objective?

Thousands of researchers, analysts, policy-makers, and students use the Global Terrorism Database(GTD) every day. We aim to better understand the strengths and limitations of the present security agents through an intensive study of this database. The objective is to analyse the causes and consequences of terrorism through a detailed analysis of the GTD.finding hot zones and most active terrorist groups so we can keep eye on them also finding weapon type to band and restrict them. And prevent/stop terrorism to make this world peacfull and harmony.

General Guidelines : -

1 cell hidden

Let's Begin !

1. Know Your Data

Import Libraries

```
#Importing all the libraries required for analysis.
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
import seaborn as sb
import matplotlib.pyplot as mp
```

Dataset Loading

```
# Load Dataset

# Importing and observing data.
data=pd.read_csv("/content/Global Terrorism Data.csv", encoding="latin1")
df=pd.DataFrame(data)
print("Data has been successfully imported")
df.head()
```

```
/usr/local/lib/python3.8/dist-packages/IPython/core/interactiveshell.py:3
  exec(code_obj, self.user_global_ns, self.user_ns)
Data has been successfully imported
```

	eventid	iyear	imonth	iday	approxdate	extended	resolution	co
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	

Dataset First View

Double-click (or enter) to edit

```
# Dataset First Look
df.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
```

Dataset Rows & Columns count

```
# Dataset Rows & Columns count
# overall row and coloumns
overall_number_of_row_and_coloumns = df.shape
print(overall_number_of_row_and_coloumns)
```

```
(181691, 135)
```

```
# number of rows
number_of_rows = len(df.index)
print(number_of_rows)
```

```
181691
```

```
# Total number of columns
number_of_columns = len(df.columns)
print(number_of_columns)
```

```
135
```

```
# column names
df.columns
```

```
Index(['eventid', 'iyear', 'imonth', 'iday', 'approxdate', 'extended',
       'resolution', 'country', 'country_txt', 'region',
       ...
       'addnotes', 'scite1', 'scite2', 'scite3', 'dbsource', 'INT_LOG',
       'INT_IDEO', 'INT_MISC', 'INT_ANY', 'related'],
      dtype='object', length=135)
```

Dataset Information

```
# Dataset Info
df.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
```

Duplicate Values

```
# Dataset Duplicate Value Count
df2 = df.pivot_table(index = ["iyear","imonth","iday","country_txt","region_txt","provstate","c
       "latitude","longitude","location","summary","attacktype1_txt","targtype1_txt",
       "gname","motive","weaptype1_txt","nkill","nwound","addnotes"], aggfunc ='size')
print(df2)
```

```
iyear  imonth  iday  country_txt    region_txt          provstate          city           latitude   longitude    location
1970   1       9     United States  North America       Puerto Rico        Rio Piedras    18.386932  -66.061127   Caparra Shopping
               12    United States  North America       New York           New York City  40.697132  -73.931351   Brooklyn
               19    United States  North America       New Jersey         Jersey City    40.717892  -74.067467   Front of building
                                                        Washington         Seattle        47.610786  -122.331306  Seattle Universit
               25    United States  North America       New York           New York City  40.697132  -73.931351   Manhattan

2017   12      22    Nigeria        Sub-Saharan Africa  Kaduna             Nindem         9.428792    8.333605    The incident occu
               24    Nigeria        Sub-Saharan Africa  Benue              Odonto         7.042064    8.003725    The incident occu
               28    Afghanistan    South Asia          Kabul              Kabul          34.516895   69.147011   The incident occu
                     Philippines    Southeast Asia      Maguindanao        Shariff Aguak  6.862806    124.443649  The incident occu
               31    India          South Asia          Jammu and Kashmir  Lethapora      33.966527   74.964225   The incident occu
Length: 6732, dtype: int64
```

Missing Values/Null Values

```
# Missing Values/Null Values Count
print(df.isnull().sum().sum())
```

```
13853997
```

```
# Visualizing the missing values
import missingno as msno
msno.bar(df)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fbf9deb3160>
```

What did you know about your dataset?

We understood the values in the columns provided. We analysed the dataset and understood that it covers details of all terrorist attacks from 1970 to 2017 globally. The data set provided us values of locations, date, terrorist group responsible, weapons used, targets , casualities etc. We also found some column headings vague and decided to not incluse them in our analysis.

2. Understanding Your Variables

```
# Dataset Columns
df.columns
```

```
Index(['eventid', 'iyear', 'imonth', 'iday', 'approxdate', 'extended',
       'resolution', 'country', 'country_txt', 'region',
       ...
       'addnotes', 'scite1', 'scite2', 'scite3', 'dbsource', 'INT_LOG',
       'INT_IDEO', 'INT_MISC', 'INT_ANY', 'related'],
      dtype='object', length=135)
```

```
# Dataset Describe
df.describe(include='all')
```

	eventid	iyear	imonth	iday	approxda
count	1.816910e+05	181691.000000	181691.000000	181691.000000	92
unique	NaN	NaN	NaN	NaN	22
top	NaN	NaN	NaN	NaN	Septemb 18-24, 20
freq	NaN	NaN	NaN	NaN	1
mean	2.002705e+11	2002.638997	6.467277	15.505644	Na
std	1.325957e+09	13.259430	3.388303	8.814045	Na
min	1.970000e+11	1970.000000	0.000000	0.000000	Na
25%	1.991021e+11	1991.000000	4.000000	8.000000	Na
50%	2.009022e+11	2009.000000	6.000000	15.000000	Na
75%	2.014081e+11	2014.000000	9.000000	23.000000	Na
max	2.017123e+11	2017.000000	12.000000	31.000000	Na

11 rows × 135 columns

Variables Description

1. eventid - It contains particular event ID of Terrorist Attack
2. iyear - It contains year of event.
3. imonth - It contains month of event.
4. iday - It contains day of event.
5. approxdate - It contains approximate date in DD/MM/YYYY manner.
6. extended - It contains extended value.
7. resolution - It contains resolution value.
8. country - It contains country name where the acttack happend.
9. country_txt - It contains country name where attack happened.
10. region - it contains region location.
11. succes - It contains succes of attack.
12. addnotes - It contains attack details.
13. scite1 - It contains site details.
14. scite2 - It contains sub site details.
15. scite3 - It contains sub site details.
16. dbsource - It contains mission name.
17. weapontype - It contains weapon type used by terrorists.
18. targettype - It contains target name targeted by terrorists.
19. gname - It contains terrorist organization name.
20. city - It contains city names attacked by terrorist.

Check Unique Values for each variable.

```
# Check Unique Values for each variable.
for i in df.columns.tolist():
  print("No. of unique values in ",i,"is",df[i].nunique(),".")
```

```
No. of unique values in  nperps is 113 .
No. of unique values in  nperpcap is 50 .
No. of unique values in  claimed is 3 .
No. of unique values in  claimmode is 10 .
No. of unique values in  claimmode_txt is 10 .
No. of unique values in  claim2 is 3 .
No. of unique values in  claimmode2 is 9 .
No. of unique values in  claimmode2_txt is 9 .
No. of unique values in  claim3 is 2 .
No. of unique values in  claimmode3 is 8 .
No. of unique values in  claimmode3_txt is 8 .
No. of unique values in  compclaim is 3 .
No. of unique values in  weaptype1 is 12 .
No. of unique values in  weaptype1_txt is 12 .
No. of unique values in  weapsubtype1 is 30 .
No. of unique values in  weapsubtype1_txt is 30 .
No. of unique values in  weaptype2 is 11 .
No. of unique values in  weaptype2_txt is 11 .
No. of unique values in  weapsubtype2 is 28 .
No. of unique values in  weapsubtype2_txt is 28 .
No. of unique values in  weaptype3 is 10 .
No. of unique values in  weaptype3_txt is 10 .
No. of unique values in  weapsubtype3 is 22 .
No. of unique values in  weapsubtype3_txt is 22 .
No. of unique values in  weaptype4 is 5 .
No. of unique values in  weaptype4_txt is 5 .
No. of unique values in  weapsubtype4 is 16 .
No. of unique values in  weapsubtype4_txt is 16 .
No. of unique values in  weapdetail is 19148 .
No. of unique values in  nkill is 205 .
No. of unique values in  nkillus is 31 .
No. of unique values in  nkillter is 96 .
No. of unique values in  nwound is 238 .
No. of unique values in  nwoundus is 44 .
No. of unique values in  nwoundte is 44 .
No. of unique values in  property is 3 .
No. of unique values in  propextent is 4 .
No. of unique values in  propextent_txt is 4 .
No. of unique values in  propvalue is 659 .
No. of unique values in  propcomment is 19157 .
No. of unique values in  ishostkid is 3 .
No. of unique values in  nhostkid is 209 .
No. of unique values in  nhostkidus is 27 .
No. of unique values in  nhours is 35 .
No. of unique values in  ndays is 328 .
No. of unique values in  divert is 143 .
No. of unique values in  kidhijcountry is 217 .
No. of unique values in  ransom is 3 .
No. of unique values in  ransomamt is 429 .
No. of unique values in  ransomamtus is 23 .
No. of unique values in  ransompaid is 156 .
No. of unique values in  ransompaidus is 8 .
No. of unique values in  ransomnote is 387 .
No. of unique values in  hostkidoutcome is 7 .
No. of unique values in  hostkidoutcome_txt is 7 .
No. of unique values in  nreleased is 156 .
No. of unique values in  addnotes is 15429 .
No. of unique values in  scite1 is 83988 .
No. of unique values in  scite2 is 62263 .
```

3. ***Data Wrangling***

Data Wrangling Code

```
# Write your code to make your dataset analysis ready.
```

Note: Since it contains 135 columns. They have a huge proportion in dataset and Learning them doesn't make any sense. So, we will rename the columns name for better understaning and then we will only extract necessary columns.

```
try:
  # Renaming columns
```

```
  terror_master_data = df.rename(columns= {'iyear':'Year','imonth':'Month','iday':'Day','country
  #Removing unwanted columns
  terror_master_data = df[['Year','Month','Day','Country','state','Region','city','latitude','l
except Exception as e:
  print(e)
```

```
df.head()
```

	eventid	Year	Month	Day	approxdate	extended	resolution	count
0	197000000001	1970	7	2	NaN	0	NaN	
1	197000000002	1970	0	0	NaN	0	NaN	1
2	197001000001	1970	1	0	NaN	0	NaN	1
3	197001000002	1970	1	0	NaN	0	NaN	

```
df.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
```

```
df.shape
```

```
(181691, 135)
```

```
df.isnull().sum()
```

```
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
```

Exploratory Data Analysis

```
data.hist(figsize=(40,20))  # This represents  the distribution of data  on each series in the
```

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9ad8c310>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf995358e0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9953fe50>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf994f6e80>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9949ddf0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99459190>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99459280>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf994856d0>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf993dfe20>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9939b250>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf993c9640>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99378730>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99322e50>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf992d95b0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99305cd0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf992bc430>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99266b50>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf9921c2b0>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf992489d0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf991fd130>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf991a8850>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf991d2f70>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf991886d0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99135df0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf990ea550>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99117c70>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf990cd3d0>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99076af0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9902c250>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99057970>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf99000130>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98fba7f0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9b23d340>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9b24e9a0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9da9ce80>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf995551c0>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf994451f0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98e9e460>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98ed2940>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98e86bb0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98e33fa0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98ded3d0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98d9a7c0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98dc1bb0>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf98d70fa0>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98d2a3d0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98d557c0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98d02be0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98cb0fd0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98c69400>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98c187f0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98c46be0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98bf30d0>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf98bae490>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98b5b880>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98b8cc70>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98b37100>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98af3490>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98a9e880>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98acbc70>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98a7a100>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98a334f0>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf989e65e0>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98a10d00>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf989c7460>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98971b80>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf989282e0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98954a00>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9890b160>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf988b58b0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9885ffd0>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf98896730>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9883ee50>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf987f65b0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf987a0cd0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98759430>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98783b50>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9873b1f0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf98757ac0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x7fbf9870f220>,
        <matplotlib.axes._subplots.AxesSubplot object at
0x7fbf986b9940>]],
```

```
dtype=object)
```

```
df["Killed"]=df["Killed"].fillna(0)
df["Wounded"]=df["Wounded"].fillna(0)
df["Casualty"]=df["Killed"]+df["Wounded"]
```

```
df.describe()
```

	eventid	Year	Month	Day	exte
count	1.816910e+05	181691.000000	181691.000000	181691.000000	181691.00
mean	2.002705e+11	2002.638997	6.467277	15.505644	0.04
std	1.325957e+09	13.259430	3.388303	8.814045	0.20
min	1.970000e+11	1970.000000	0.000000	0.000000	0.00
25%	1.991021e+11	1991.000000	4.000000	8.000000	0.00
50%	2.009022e+11	2009.000000	6.000000	15.000000	0.00
75%	2.014081e+11	2014.000000	9.000000	23.000000	0.00
max	2.017123e+11	2017.000000	12.000000	31.000000	1.00

8 rows × 78 columns

Observasion

1. The data consists of terrorist activities ranging from the year:1970 to 2017
2. Maximum number of people killed in an event were: 1570
3. Maximum number of people wounded in an event were: 8191
4. Maximum number of total casualties in an event were: 9574

What all manipulations have you done and insights you found?

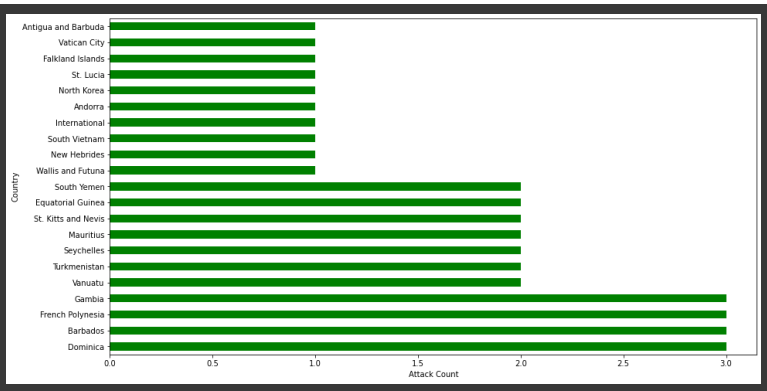
We have changed the column names to more useable names. Also since we found some columns vague and could not understand the values they contained, we filtered our data to include only workable columns

4. Data Vizualization, Storytelling & Experimenting with charts : Understand the relationships between variables

Chart - 1

Frequent terrorist activities may bring instability to a country's economy. Ignoring other factors which countries economy is least effected by terrorism?

```
# Chart - 1 visualization code
try:
  country_wise_attack_count = df['Country'].value_counts()    # counting number of times each c
  country_wise_attack_count.sort_values(axis=0 , inplace=True,ascending=[True] ) # sorting coun
except Exception as e:
  print(e)
else:
  # Plotting bar graph for 20 countries
  plt.rcParams['figure.figsize']=(15,8)
  country_wise_attack_count[20::-1].plot(kind='barh', color= "green")
  plt.ylabel('Country')
  plt.xlabel('Attack Count')
```



1. Why did you pick the specific chart?

Bar graphs provide a very easy perception. Humans have a better understanding with length differences than areas or angles. Hence, as we wanted to compare the attacks in different countries, we chose bar graph. Horizontal bar was chosen to accomodate more countries in the screen area.

2. What is/are the insight(s) found from the chart?

The chart clearly shows those countries which are least effected by terrorism. We inferred that the common denominator between countries like North Korea, Antigua & Barboda, Vatican city was the fact that these countries do not have very significant religious diversities which may lead to dissent. A strong central governing power also contributes to better law and order.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Yes, the gained insight can help other nations(like RAW,NIA in India) to better fortify their boundries and manage internal affairs better. Unfortunately, this insight can also infer that religious extremism is the leading cause of terrorism.

Chart - 2

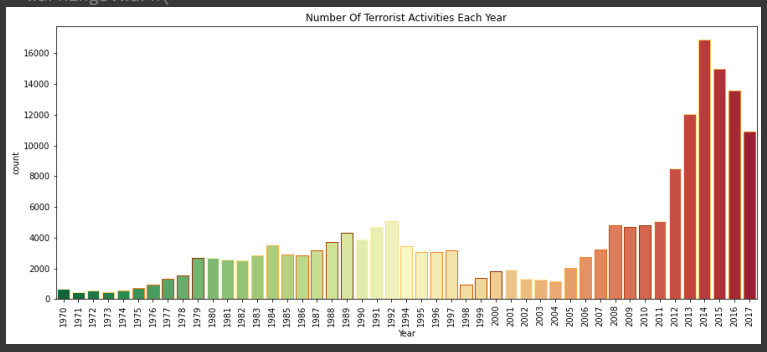
Find out if number of terrorist attacks each year increasing or decreasing

```
# Chart - 2 visualization code
# finding number of attacks in each year
attacks=df["Year"].value_counts(dropna=False).sort_index().to_frame().reset_index().rename(colu
attacks.head()
```

Year	Attacks
1970	651
1971	471
1972	568
1973	473
1974	581

```
# finding if terrorist attacks getting increasing or decreasing year by .
plt.subplots(figsize=(15,6))
sns.countplot('Year',data=data,palette='RdYlGn_r',edgecolor=sns.color_palette("YlOrBr", 10))
plt.xticks(rotation=90)
plt.title('Number Of Terrorist Activities Each Year')
plt.show()
```

```
/usr/local/lib/python3.8/dist-packages/seaborn/_decorators.py:36: FutureWa
  warnings.warn(
```



1. Why did you pick the specific chart?

Bar graphs provide a very easy perception. Humans have a better understanding with length differences than areas or angles. Hence, as we wanted to compare the attacks in different countries, we chose bar graph. Horizontal bar was chosen to accomodate more countries in the screen area.

2. What is/are the insight(s) found from the chart?

As we can see from the graph in 1970 terrorist attacks got started it got similar till 1974 then from 1975 attacks start to increase to;; 1992 then we see sudden drop in attack which shows our stronger defence then from 2005 it starts increasing till 2012.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Yes, the identified pattern can help World peacekeeping agencies to focus on the defence system and investigation system to decrease/ prevent these terrorist attacks. The graph result shows that terrorist attacks increasing day by day world should keep eye on them and increase defence system.

Chart - 3

Each and every life is very important rather than this a mental torture is also really big crime. Govt. should raise a helping hand to the family of dead peoples and wounded peoples for that calculate the total death and wounded number of suffered people.

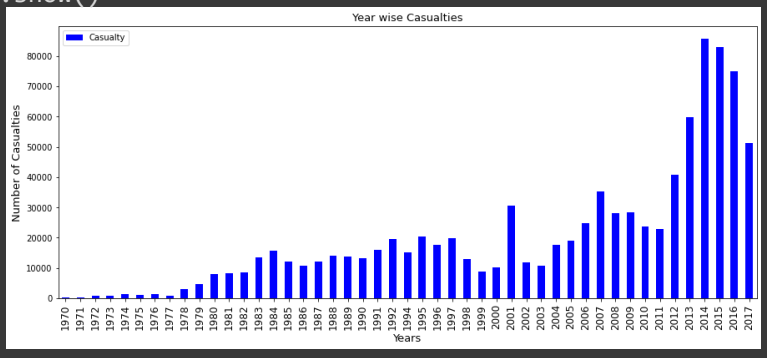
```
# Chart - 3 visualization code
```

```
#Total Casualties (Killed + Wounded) in each Year
yc=df[["Year","Casualty"]].groupby("Year").sum()
yc.head()
```

Year	Casualty
1970	386.0
1971	255.0
1972	975.0
1973	865.0
1974	1404.0

```
# ploting bar for total number of casualities each year
yc.plot(kind="bar",color="blue",figsize=(15,6))
```

```
plt.title("Year wise Casualties",fontsize=13)
plt.xlabel("Years",fontsize=13)
plt.xticks(fontsize=12)
plt.ylabel("Number of Casualties",fontsize=13)
plt.show()
```




```
# calculate killed people in  each year
yk=df[["Year","Killed"]].groupby("Year").sum()
yk.head()
```

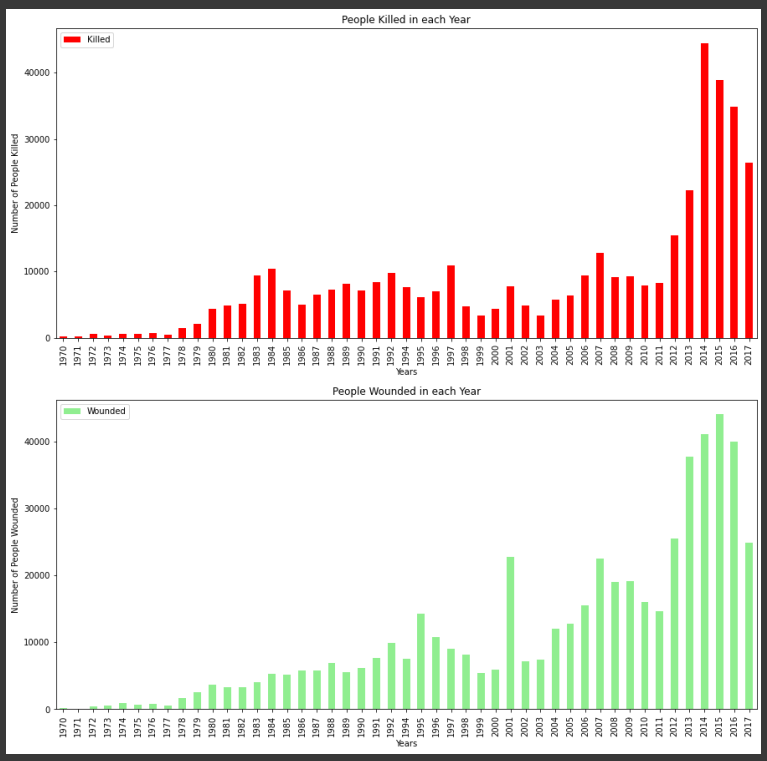
Year	Killed
1970	174.0
1971	173.0
1972	566.0
1973	370.0
1974	539.0

```
# calculate wounded people in each reagion
yw=df[["Year","Wounded"]].groupby("Year").sum()
yw.head()
```

Year	Wounded
1970	212.0
1971	82.0
1972	409.0
1973	495.0
1974	865.0

```
# making two graphs of wounded people and another for killed peoples
# first plot the figure
fig=plt.figure()
ax0=fig.add_subplot(2,1,1)
ax1=fig.add_subplot(2,1,2)

# making graph of killed people
yk.plot(kind="bar",color="red",figsize=(15,15),ax=ax0)
ax0.set_title("People Killed in each Year")
ax0.set_xlabel("Years")
ax0.set_ylabel("Number of People Killed")
```

```
# making graph of Wounded people
yw.plot(kind="bar",color="lightgreen",figsize=(15,15),ax=ax1)
ax1.set_title("People Wounded in each Year")
ax1.set_xlabel("Years")
ax1.set_ylabel("Number of People Wounded")

plt.show()
```



1. Why did you pick the specific chart?

I coded two graph in one code to compare total number of death and total number of wounded people.

2. What is/are the insight(s) found from the chart?

In 2007 the terrorist attack make a really hazardous attack that total number of death are at it's peak we get more than 12000+ death rate and in 2001 we get 20000+ wounded people which make a permanant emotional and mental of fear on common people.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Terrorist attacks are increasing day by day govt. and anti terrorist organization should take a look on it.

Chart - 4

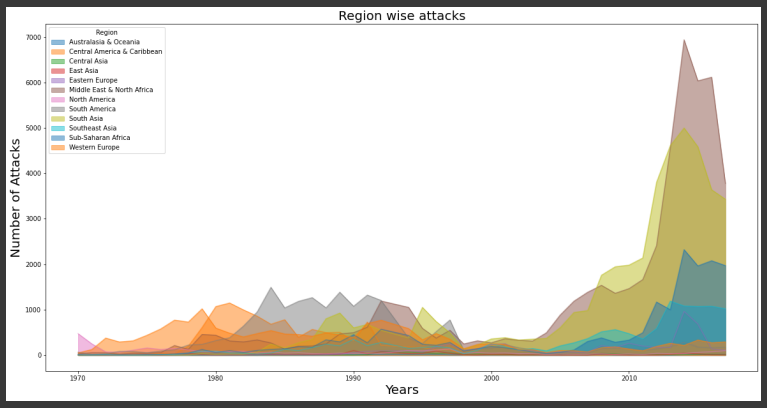
Find out most terrorism affected region so we can declare it red zone so people avoid the visit to those region and we can also inform the regional govt. to improve their defence system.

```
# Chart - 4 visualization code
```

```
#Distribution of Terrorist Attacks over Regions from 1970-2012
reg=pd.crosstab(df.Year,df.Region)
reg.head()
```

Region	Australasia & Oceania	Central America & Caribbean	Central Asia	East Asia	Eastern Europe	Middle East & North Africa	North America	A
Year								
1970	1	7	0	2	12	28	472	
1971	1	5	0	1	5	55	247	
1972	8	3	0	0	1	53	73	
1973	1	6	0	2	1	19	64	

```
# ploting a graph to calculate the affected region esily
reg.plot(kind="area", stacked=False, alpha=0.5,figsize=(20,10))
plt.title("Region wise attacks",fontsize=20)
plt.xlabel("Years",fontsize=20)
plt.ylabel("Number of Attacks",fontsize=20)
plt.show()
```

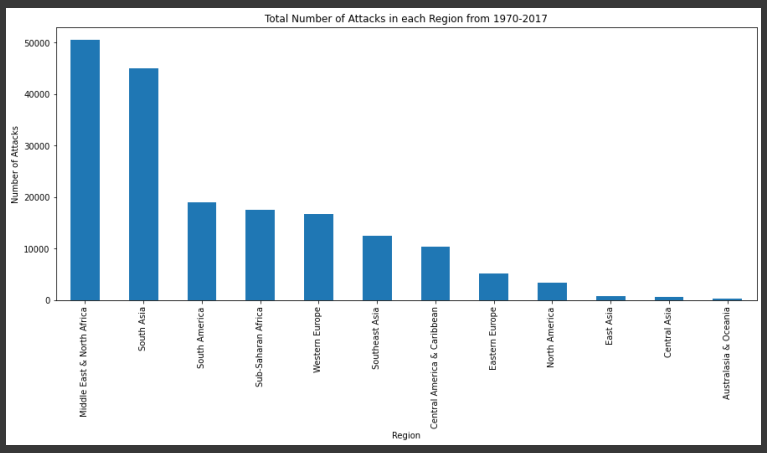



```
#Total Terrorist Attacks in each Region from 1970-2012
regt=reg.transpose()
regt["Total"]=regt.sum(axis=1)
ra=regt["Total"].sort_values(ascending=False)
ra
```

```
Region
Middle East & North Africa     50474
South Asia                     44974
South America                  18978
```

```
Sub-Saharan Africa            17550
Western Europe                16639
Southeast Asia                12485
Central America & Caribbean   10344
Eastern Europe                 5144
North America                  3456
East Asia                       802
Central Asia                    563
Australasia & Oceania           282
Name: Total, dtype: int64
```

```
# ploting a bar graph to understand it by desending order
ra.plot(kind="bar",figsize=(15,6))
plt.title("Total Number of Attacks in each Region from 1970-2017")
plt.xlabel("Region")
plt.ylabel("Number of Attacks")
plt.show()
```

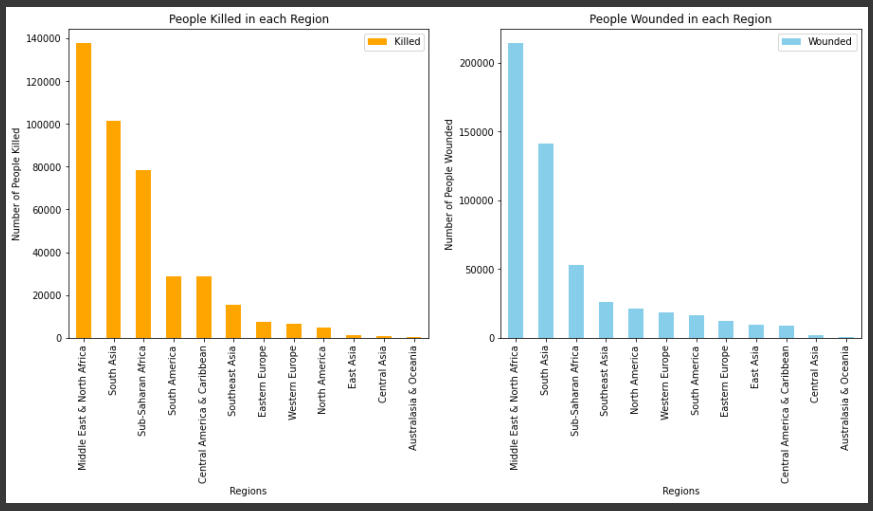



```
# calculating the total number of killed people in each reagion
rk=df[["Region","Killed"]].groupby("Region").sum().sort_values(by="Killed",ascending=False)
print(rk)
# calculating the total number of wounded people in each reagion
rw=df[["Region","Wounded"]].groupby("Region").sum().sort_values(by="Wounded",ascending=False)
print(rw)
```

```
                             Killed
Region
Middle East & North Africa   137642.0
South Asia                   101319.0
Sub-Saharan Africa            78386.0
South America                 28849.0
Central America & Caribbean   28708.0
Southeast Asia                15637.0
Eastern Europe                 7415.0
Western Europe                 6694.0
North America                  4916.0
East Asia                      1152.0
Central Asia                   1000.0
Australasia & Oceania           150.0
                            Wounded
Region
Middle East & North Africa   214308.0
South Asia                   141360.0
Sub-Saharan Africa            52857.0
Southeast Asia                26259.0
North America                 21531.0
Western Europe                18332.0
South America                 16704.0
Eastern Europe                12045.0
East Asia                      9213.0
Central America & Caribbean    8991.0
```

```
Central Asia                  2009.0
Australasia & Oceania          260.0
```

```
# ploting a graph to compare killed people and woundend people by terrorist

fig=plt.figure()
ax0=fig.add_subplot(1,2,1)
ax1=fig.add_subplot(1,2,2)

#Killed
rk.plot(kind="bar",color="orange",figsize=(15,6),ax=ax0)
ax0.set_title("People Killed in each Region")
ax0.set_xlabel("Regions")
ax0.set_ylabel("Number of People Killed")

#Wounded
rw.plot(kind="bar",color="skyblue",figsize=(15,6),ax=ax1)
ax1.set_title("People Wounded in each Region")
ax1.set_xlabel("Regions")
ax1.set_ylabel("Number of People Wounded")

plt.show()
```



1. Why did you pick the specific chart?

we choose area type of plot to show region vise attack to simplyfy the results and find a red zone region of terrorist attacks.

2. What is/are the insight(s) found from the chart?

We found out top 10 red zone terrorist attack regions as follows Middle East & North Africa
South Asia
South America
Sub-Saharan Africa
Western Europe
Southeast Asia
Central America & Caribbean
Eastern Europe
North America
East Asia
Central Asia

among them Middle east & North Africa and south Asia are most attacked regions Middle east & North Africa have 50474 attacks with 137642 people died and 214308 number of people got wounded. So here we find top 10 terrorist attacked countries with most hazardous and altra red zone area which are Middle East & North Africa and South Asia which have possibility to get attacked in future also.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

From 2010 the terrorist attacks in Middel east and North Africa got drasticaly increase due to Democracy and civil conflicts are main drivers of MENA terrorism. Govt. of Middel east and North Africa should take action on these. And regional countries should avoid tourist to visit these 10 regions specialy Middile east & North Africa and South Asia.

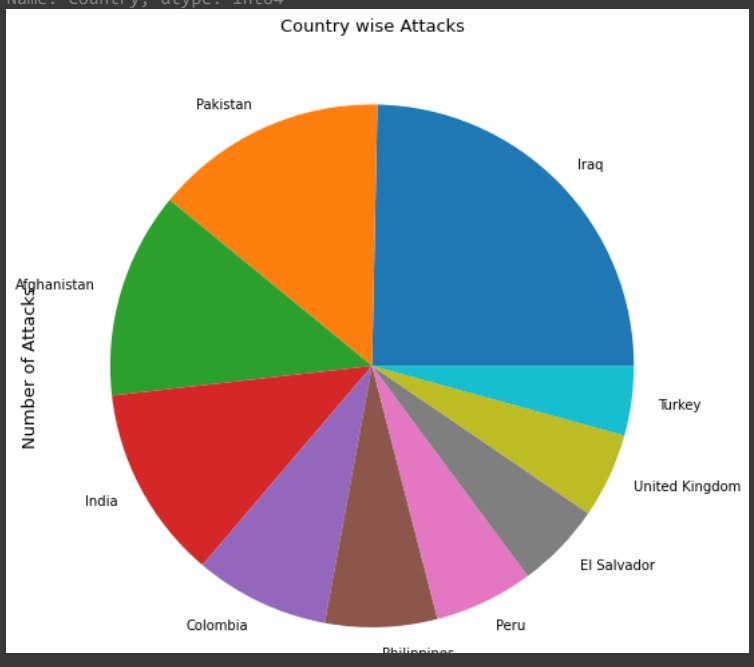
Chart - 5

Just like region countries also should avoid torists to make safty of people and improve their defence . So to increase defence of particular country we should calculate top 10 terrorism affected countries.

```
# Chart - 5 visualization code
```

```
# find Country wise Attacks - Top 10
# Number of Attacks in each Country
ct=df["Country"].value_counts().head(10)
print(ct)
# ploting a grap of it to understand easily
ct.plot(kind="pie",figsize=(20,9))
plt.title("Country wise Attacks",fontsize=13)
plt.xlabel("Countries",fontsize=13)
plt.xticks(fontsize=12)
plt.ylabel("Number of Attacks",fontsize=13)
plt.show()
```

```
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
```



1. Why did you pick the specific chart?

Camparative charts are really easy to understad for humans that's why I choose this specific chart.

2. What is/are the insight(s) found from the chart?

Iraq is the most affected country among all with highest rate of terrorist attacks. Iraq had 24636 attacks. Pakistan Afghanistan And India have similar attacks.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Iraq
Pakistan
Afghanistan
India
Colombia
Philippines
Peru
El Salvador
United Kingdom
Turkey

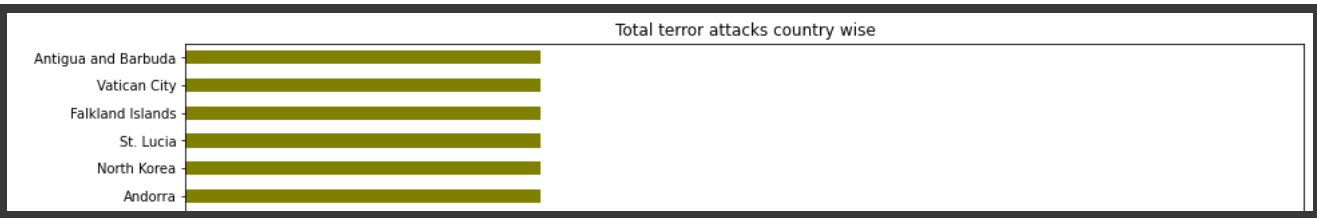
these country should increse their defince system and keep eye on terrorist organizations.

Chart - 6

Frequent terrorist activities may bring instability to a country's economy. Ignoring other factors which countries economy is least effected by terrorism?

```
# Chart - 6 visualization code
```

```
def terror_hot_zones():
  try:
    country_wise_attack_count = df['Country'].value_counts()        #Counting number of times
    country_wise_attack_count.sort_values(axis =0 , inplace=True,ascending=[True] )   #Sorting
  except Exception as e:
    print(e)
  else:
    # Plotting bar graph for 20 countries
    plt.rcParams['figure.figsize']=(15,8)
    country_wise_attack_count[20::-1].plot(kind='barh', color = "olive")
    plt.title('Total terror attacks country wise')
    plt.ylabel('Country')
    plt.xlabel('Attack Count')

terror_hot_zones()
```



1. Why did you pick the specific chart?

Bar graphs provide a very easy perception. Humans have a better understanding with length differences than areas or angles. Hence, as we wanted to compare the attacks in different countries, we chose bar graph. Horizontal bar was chosen to accomodate more countries in the screen area.

2. What is/are the insight(s) found from the chart?

The chart clearly shows those countries which are least effected by terrorism. We inferred that the common denominator between countries like North Korea, Antigua & Barboda, Vatican city was the fact that these countries do not have very significant religious diversities which may lead to dissent. A strong central governing power also contributes to better law and order.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Yes, the gained insight can help other nations(like RAW,NIA in India) to better fortify their boundries and manage internal affairs better. Unfortunately, this insight can also infer that religious extremism is the leading cause of terrorism.

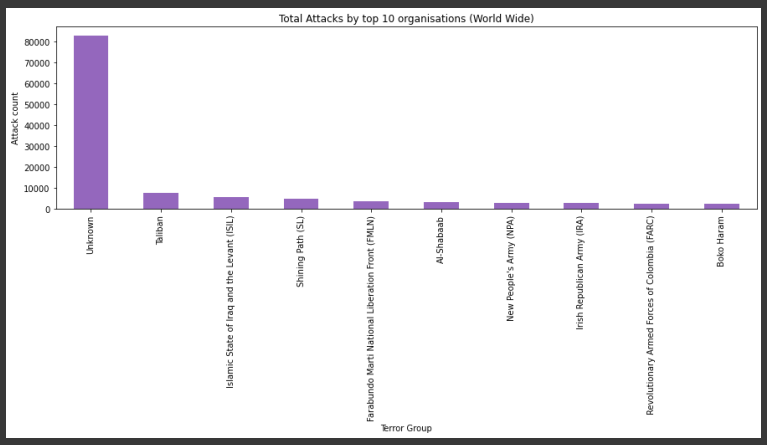
Chart - 7

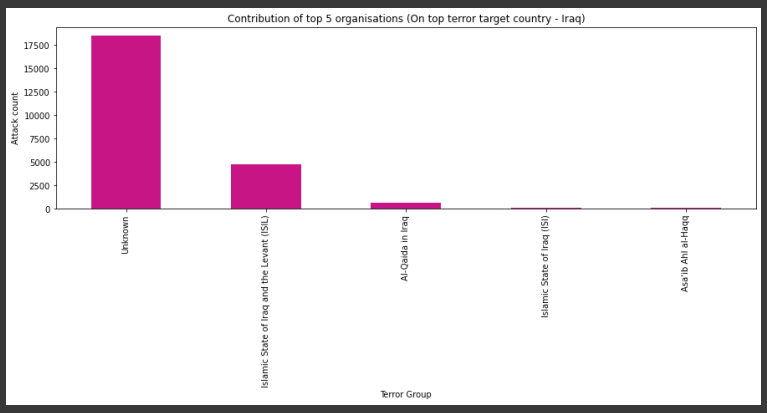
Disturbances in different regions are caused by different(generally local) terrorist organisations. Analyse : i) The main terrorist groups operating worldwide ii) The main terrorist groups operating in the country of most terrorist attacks

```
# Chart - 7 visualization code
```

```
def get_top_terror_groups_in_hot_zones():
  try:
    #Finding top 10 terror outfits of the world
    terror_attack_count = df['Group'].value_counts().head(10)

    #Finding name of the most effected country
    mostEffectedCountry = df['Country'].value_counts().index[0]

    #Segregatting attacks on most effected country
    countries_effect_count = df.loc[(df['Country'] == mostEffectedCountry)]

    #Finding top 5 terror outfits in the most effected country
    countries_effect_count = countries_effect_count["Group"].value_counts()[0:5]

  except Exception as e:
    print(e)


  else:
    #Plotting top 10 terror outfits of the world
    plt.rcParams['figure.figsize']=(15,4)
    terror_attack_count.plot(kind='bar', color = "tab:purple")

    plt.title('Total Attacks by top 10 organisations (World Wide)')
    plt.ylabel('Attack count')
    plt.xlabel('Terror Group')
    plt.show()

    print("\n \n ")
```

```
    #Plotting top terror groups of the most effected country
    countries_effect_count.plot(kind='bar', color = 'mediumvioletred')
    plt.title(f'Contribution of top 5 organisations (On top terror target country - {mostEffect
    plt.ylabel('Attack count')
    plt.xlabel('Terror Group')
    plt.show()

get_top_terror_groups_in_hot_zones()
```





1. Why did you pick the specific chart?

We chose bar graph to show a comparitive study between terror attack counts of the top 10 globally and top 5 terror outfits in a country. The length of the bars gives a relative idea about the count and difference between various terror groups

2. What is/are the insight(s) found from the chart?

As we can see from both the graphs, majority of the terrorist attacks are conducted by 'Unknown' groups world wide and in the most effected nations.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Yes, the identified pattern can help World peacekeeping agencies to focus on identifying and keeping a watch on these 'Unknown' groups. Curbing the advances of these unknown terror outfits can strengthen national security.

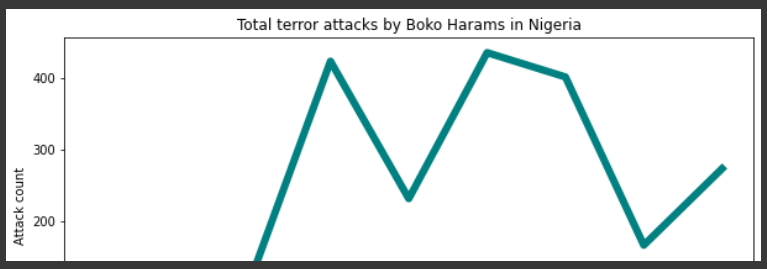
Chart - 8

Analyse the birth and growth of Boko Harams in Nigeria. Who are their primary targets?

```
# Chart - 8 visualization code
```

```
# Chart - 3 visualization code

def get_Nigeria_boko_details():
  try:
    # Filtering those records where Boko Harams have attacked Nigeria
    nigeria_data = df.loc[(df['Country'] == 'Nigeria') & (df['Group'] == 'Boko Haram')]

    #Finding year wise count of the attacks
    year_wise_attack_count = nigeria_data[['Year']].value_counts()

    #Sorting records year wise
    year_wise_attack_count.sort_index(axis =0 , inplace=True)

  except Exception as e:
      print(e)

  else:

    #Plotting line graph to show trend over the years
    plt.rcParams['figure.figsize']=(10,5)
    year_wise_attack_count.plot(kind='line', color = 'teal', linewidth = 6)
    plt.title('Total terror attacks by Boko Harams in Nigeria')

    #Assigning labels for x and y axis
    plt.ylabel('Attack count')
    plt.xlabel('Year')

  try:
    #Finding count of every target type
    primary_targets = nigeria_data[['Target']].value_counts()
    #Sorting values in decreasing order to find most effected targets
    primary_targets.sort_values(axis =0 , inplace=True, ascending=[False])
    #Printing top 3 targets
    print("*****************************************************************")
    print(f"Total attacks by Boko Harams in Nigeria = {len(nigeria_data)}")
    print(primary_targets[0:3])
    print("*****************************************************************")

  except Exception as e1:
    print(e1)

get_Nigeria_boko_details()
```

```
**************************************************************
Total attacks by Boko Harams in Nigeria = 2087
Target
Village      371
Civilians    153
Town          95
dtype: int64
**************************************************************
```



1. Why did you pick the specific chart?

As we wanted to show the trends for different time periods, we chose the line graph. It clearly depicts the rise and fall of the attacks done by the Boko Harams.

2. What is/are the insight(s) found from the chart?

The graph clearly shows how Boko Harams who were a small resilience force from 2002 to 2009, rose as a terror group in Nigeria between the years 2010 and 2012. In 2009, Boko Harams were subjected to excessive use of force by police, which triggered backlashes in the form of bombings and killings. This also questions the planning , proactiveness and policies of the then present government to resolve conflicts by meaningful dialogue.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

This graph highlights 2012 and 2014 as the peak years of terror attacks by Boko Harams.In 2015 President Buhari made meaningful efforts in curbing Boko Harams in Nigeria. The efforts made by the President can be studied and replicated to ensure a positive change in the safety of Nigerians and other countries facing civil unrest worldwide.

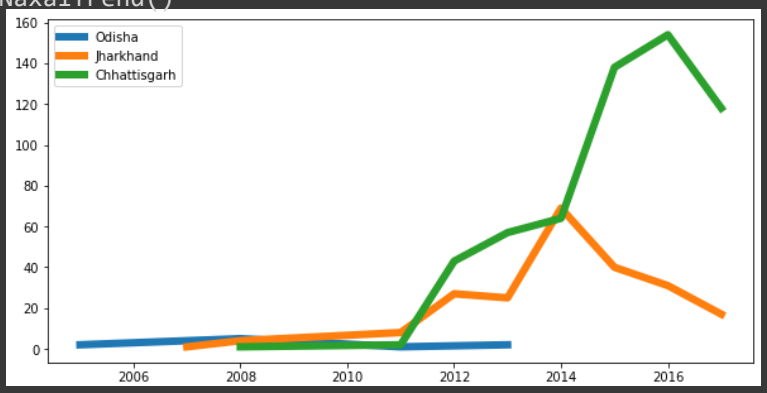
Chart - 9

Compare the success of Odisha, Jharkhand and Chattisgarh governments to reduce naxalism in their regions.

```
# Chart - 9 visualization code
```

```
def getNaxalTrend():
  try:

    #Finding records for the states where Maoists have attacked

    #Odisha
    terror_Od=df.loc[(df['state'] == 'Orissa') & (df['Group'] == 'Maoists')]
    #Jharkhand
    terror_Jh=df.loc[(df['state'] == 'Jharkhand') & (df['Group'] == 'Maoists')]
    #Chhattisgarh
    terror_Ch=df.loc[(df['state'] == 'Chhattisgarh') & (df['Group'] == 'Maoists')]

    #Finding count of attacks by Maoists on the 3 states year wise
    od_count = terror_Od.groupby('Year').size()
    jh_count = terror_Jh.groupby('Year').size()
    ch_count = terror_Ch.groupby('Year').size()

  except Exception as e:
    print(e)

  else:
```

```
    #Plotting line graphs
    plt.plot(od_count, linewidth = 6)
    plt.plot(jh_count, linewidth = 6)
    plt.plot(ch_count, linewidth = 6)
    plt.legend(["Odisha","Jharkhand","Chhattisgarh"])
    plt.show()

getNaxalTrend()
```



1. Why did you pick the specific chart?

We wanted to show a comparision of the trends in maoist activities in the states of Odisha, Jharkhand and Chhattisgarh. Line graphs are a good choice to study trends or changes in patterns.

2. What is/are the insight(s) found from the chart?

While states like Odisha have successfully handled the menace of Naxalism systematically , wiping it out almost completely in the year 2013. We see a significant rise in its neighbouring states in the same year. We infer that political unrest in Jharkhand in 2013 may have been a triggering point.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

The various government efforts like education, monetary grants , benefits on surrender can be further facilitated to decrease Naxalism in the states.

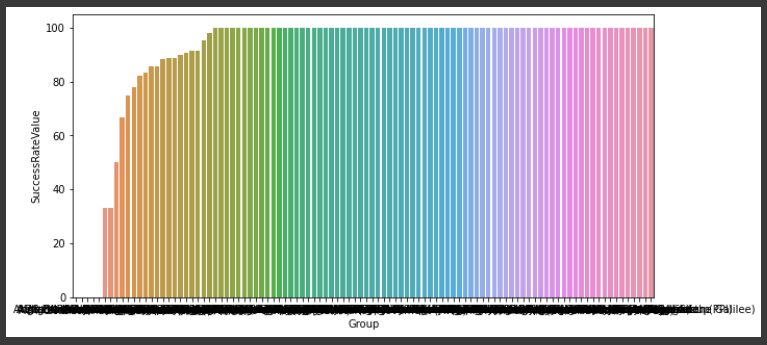
Chart - 10

Terrorist attacks are often well planned and strategically carried out operations. Analyse the success rates of some major terror outfits. Which organisations have the highest success rate?

```
# Chart - 10 visualization code
```

```
def getTerrorGroupSuccessRate():
  try:

    #Counting total attacks by specific groups
    totalAttacks = df.groupby('Group')['Group'].count()

    #Counting success of specific groups
    success = df.groupby('Group')["Success"].sum()

    #merging dataframes on Terror group names(Index)
    successRate =  pd.merge(totalAttacks[0:100], success, how='inner', left_index=True, right_i

    # Calculating success rate : Success/Total * 100
    successRate ["SuccessRateValue"] = (successRate["Success"]/successRate["Group"])*100
    successRate.sort_values(by ='SuccessRateValue', inplace=True)
```

```
  except Exception as e:
    print(e)

  else:
    #Plotting graph depicting success rate of 100 terror groups
    sns.barplot(x = successRate.index, y =successRate['SuccessRateValue'] )

getTerrorGroupSuccessRate()
```



1. Why did you pick the specific chart?

We wanted to show a comparitive study between the success rates of 100 terrorist organisations worldwide. To picturize their values in a more unstandable form, we chose a bar graph

2. What is/are the insight(s) found from the chart?

The chart showcases 100% accuracy for majority of the terror groups chosen. It shows how accurately and strategically these attacks are planned.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

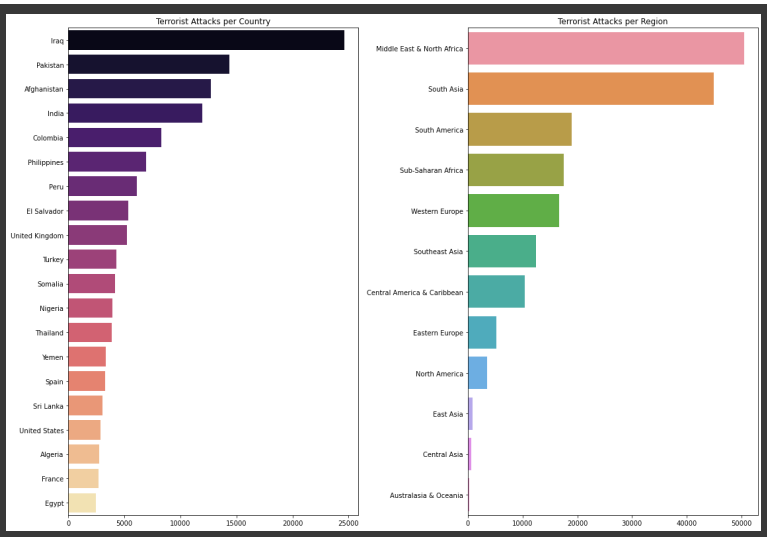
National security agencies can use this data to examine the reason for high success rates and also for low success rates of a few groups. The difference in their operating strategy can be understood to study patterns and foil more attacks in the future.

Chart - 11

The way a terrorist attack is carried out differs from organisation to organisation. Compare the various attack types and their frequency?

```
# Chart - 11 visualization code
```

```
# Chart - 8 visualization code

fig,axes = plt.subplots(figsize=(16,11),nrows=1,ncols=2)
sns.barplot(x = data['Country'].value_counts()[:20].values, y = data['Country'].value_counts()[
            ax=axes[0],palette = 'magma');
axes[0].set_title('Terrorist Attacks per Country')
sns.barplot(x=data['Region'].value_counts().values,y=data['Region'].value_counts().index,
            ax=axes[1])
axes[1].set_title('Terrorist Attacks per Region')
fig.tight_layout()
plt.show()
```



▼ 1. Why did you pick the specific chart?

We wanted to show the contribution of each weapon in the overall weapons used. Hence to show the distribution out of total, we used a pie chart.

▼ 2. What is/are the insight(s) found from the chart?

The chart depicts how out of all the weapons used , almost 50% of attacks are done using explosives. Firearms are also a major weapon used by terrorists.

▼ 3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

Yes this analysis can be used by security agencies to impose strict ban and increase vigilence on movement on particular substances especially the raw materials used to build explosives.

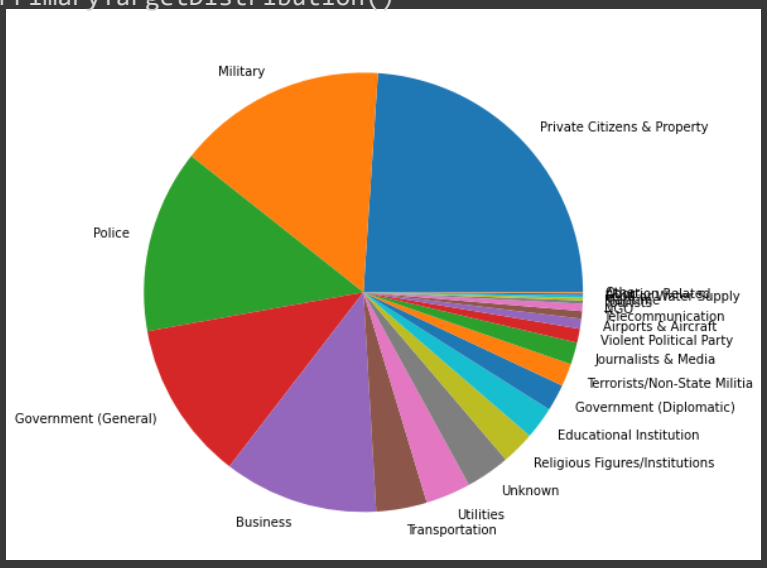
▼ Chart - 12

Although terrorism has no religion, but terrorist organisations have specific targets. Who are the most vulnerable targets? Civilians, military or politicians?

```
# Chart - 12 visualization code
```

```
def getPrimaryTargetDistribution():
  try:
    #Counting terrorism targets
    primary_target = terror_master_data.Target_type.value_counts()
  except Exception as e:
    print(e)
  else:
    #Plotting pie chart
```

```
    plt.rcParams['figure.figsize']=(15,8)
    plt.pie(primary_target, labels = primary_target.index)
    plt.show()

getPrimaryTargetDistribution()
```



1. Why did you pick the specific chart?

To show distribution of social groups out of the total casualities/targets, we chose a pie chart

2. What is/are the insight(s) found from the chart?

The primary target of terrorists are civilians and private property amounting to aoproximately 25%. Military, Police and Government officials are almost at equal risk.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

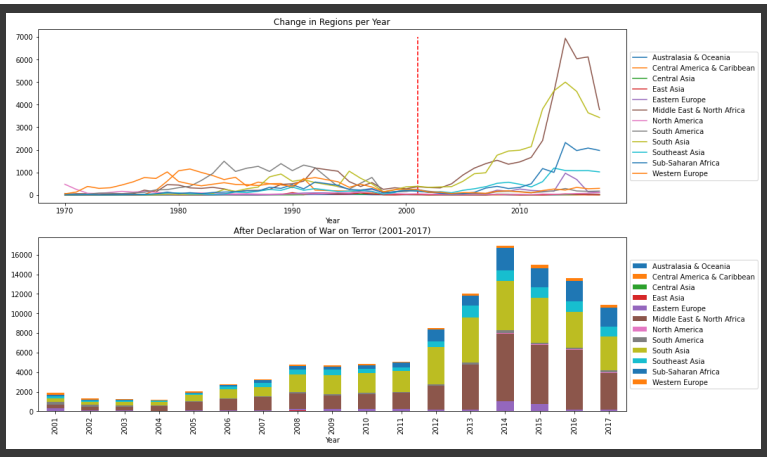
This insight can be used to improve security in public places. The distribution clearly shows that since the terrorists have no personal disputes with the civilians the purpose is primarily to create fear.

Chart - 13

Make a closer look at trend Before and after the War on Terror.

```
# Chart - 13 visualization code
```

```
# Now we will look closer at trend Before and after the War on Terror
data_after = data[data['Year']>=2001]
fig,ax = plt.subplots(figsize=(15,10),nrows=2,ncols=1)
ax[0] = pd.crosstab(data.Year,data.Region).plot(ax=ax[0])
ax[0].set_title('Change in Regions per Year')
ax[0].legend(loc='center left',bbox_to_anchor = (1,0.5))
ax[0].vlines(x=2001,ymin=0,ymax=7000,colors='red',linestyles='--')
pd.crosstab(data_after.Year,data_after.Region).plot.bar(stacked=True,ax=ax[1])
ax[1].set_title('After Declaration of War on Terror (2001-2017)')
ax[1].legend(loc='center left',bbox_to_anchor = (1,0.5))
plt.show()
```



1. Why did you pick the specific chart?

To understand trend Before and after the War on Terror.

2. What is/are the insight(s) found from the chart?

From the first plot, it is very noticeable that the terrorism landscape before and after the War on Terror is vastly different. Before 2001, the regions were much closer together in terms of activity, eventually all dropping to a minimum in 2000. After 2001, the Middle East and South Asia have dictated the rise in terrorism numbers, with a significant increase in Sub-Saharan Africa as well

3. Will the gained insights help creating a positive business impact?

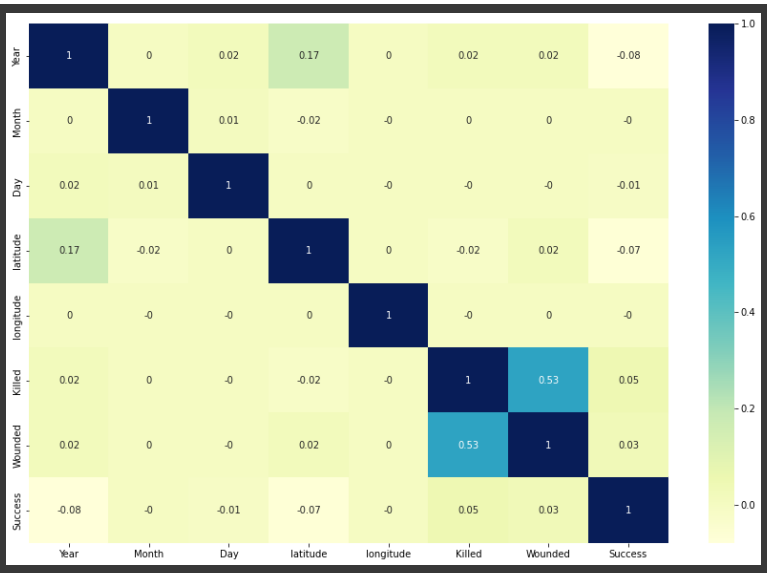
Are there any insights that lead to negative growth? Justify with specific reason.

From the first plot, it is very noticeable that the terrorism landscape before and after the War on Terror is vastly different. Before 2001, the regions were much closer together in terms of activity, eventually all dropping to a minimum in 2000. After 2001, the Middle East and South Asia have dictated the rise in terrorism numbers, with a significant increase in Sub-Saharan Africa as well

Chart - 14 - Correlation Heatmap

Analyse the data correlation between various attributes from the Global Terrorism Database.

```
# Correlation Heatmap visualization code

def findDataCorr():
  # Correlation Heatmap visualization code
  plt.figure(figsize=(15,10))
  #This shows how much related is one parameter to the other in the dataset.
  try:
    sns.heatmap(np.round(terror_master_data.corr(),2),annot=True, cmap='YlGnBu')
  except Exception as e:
    print(e)


findDataCorr()
```



1. Why did you pick the specific chart?

Correlation heatmaps can be used to find potential relationships between variables and to understand the strength of these relationships. In addition, correlation plots can be used to identify outliers and to detect linear and nonlinear relationships.

2. What is/are the insight(s) found from the chart?

We inferred that deaths and wounded have a correlation of 0.53. Also the success rate of an attack is not correlated with either month and longitude of the attack at all.

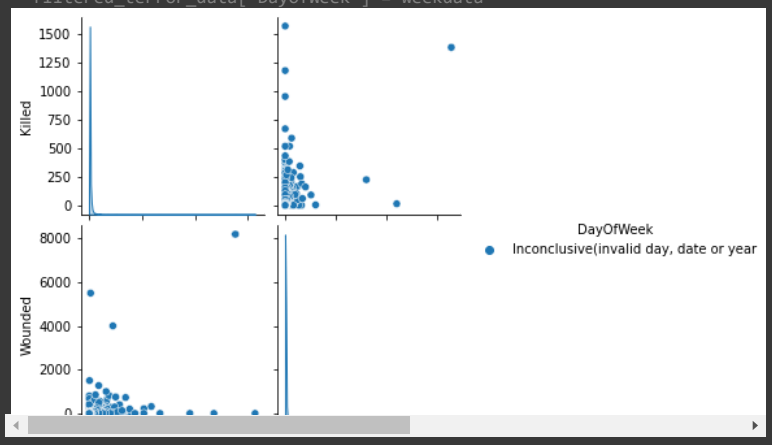
Chart - 15 - Pair Plot

Terrorism have no impathy with people they attack and kill people. find out is their any attack day pattern in which terrorist attack happens most.

```
# Pair Plot visualization code

def findMostAttackedDayOfWeek():
  #Using only required columns
  filtered_terror_data = df[['Killed','Wounded','Year','Month','Day']]
  #Creating a blank column
  filtered_terror_data['DayOfWeek']=[""]*len(filtered_terror_data)
  weekdays = ["Monday","Tuesday","Wednesday","Thursday","Friday","Saturday","Sunday","Inconclus
  weekdata =[]

  #Iterating through the rows of the data frame
  for x in filtered_terror_data.iterrows():
    x= list(x)

    try:
      d = date(int(x[1][2]) , int(x[1][3]) ,int(x[1][4])).weekday()
      weekdata.append(weekdays[d])
    except:
      weekdata.append("Inconclusive(invalid day, date or year")

  filtered_terror_data['DayOfWeek'] = weekdata

  filtered_terror_data = filtered_terror_data[['DayOfWeek', 'Killed','Wounded']]

  #Plotting pair plot
  sns.color_palette("tab10")
```

```
  sns.pairplot(filtered_terror_data,hue='DayOfWeek')

findMostAttackedDayOfWeek()
```

```
<ipython-input-59-0aa4777ad32f>:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-doc
  filtered_terror_data['DayOfWeek']=[""]*len(filtered_terror_data)
<ipython-input-59-0aa4777ad32f>:21: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-doc
  filtered_terror_data['DayOfWeek'] = weekdata
```



1. Why did you pick the specific chart?

The Seaborn Pairplot allows us to plot pairwise relationships between variables within a dataset. This creates a nice visualisation and helps us understand the data by summarising a large amount of data in a single figure.

2. What is/are the insight(s) found from the chart?

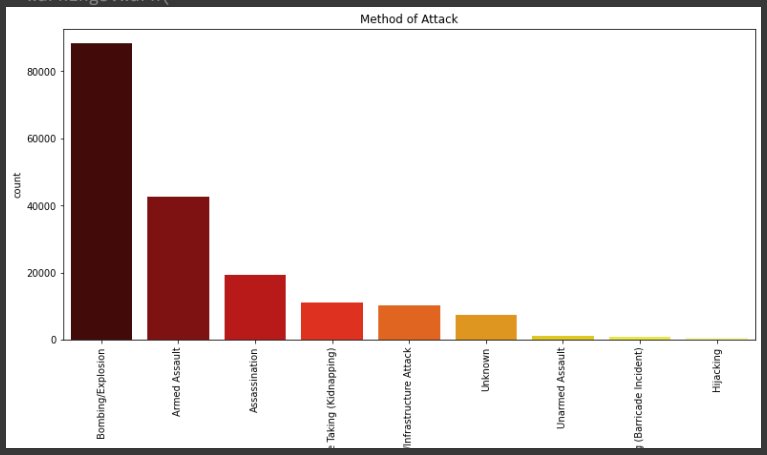
We found that most of the attacks were conducted around weekends, i.e Mondays and Fridays. So the security agencies could look into the movement of people and goods move vigilantly on weekends.

Chart - 16

To decrease or to stop terrorist attacks which weapons and their raw materials should we ban\restrict ?

```
# getting types of attacks by terrorists
plt.figure(figsize=(13,6))
sns.countplot(data['AttackType'],order=data['AttackType'].value_counts().index,
              palette='hot')
plt.xticks(rotation=90)
plt.xlabel('Method')
plt.title('Method of Attack')
plt.show()
```

```
/usr/local/lib/python3.8/dist-packages/seaborn/_decorators.py:36: FutureWa
  warnings.warn(
```



1. Why did you pick the specific chart?

To get highest and lowest used weapons for attack by terrorist.

2. What is/are the insight(s) found from the chart?

Since from the above chart it is clear that Bombing/Explosion method was mostly used.

3. Will the gained insights help creating a positive business impact?

Are there any insights that lead to negative growth? Justify with specific reason.

from the above chart it is clear that Bombing/Explosive are the highly used attack weapons by terrorist so the govt. and world anti terror organization and human right should ban and restrict the Bombing/explosive and their raw material.

5. Solution to Business Objective

What do you suggest the client to achieve Business Objective ?

Explain Briefly.

1. The Middle East and North Africa regions are the most targeted so the Govt. of that region should increase their defence and investigation departments and also should ban bombing/explosion raw material.
2. Iraq country is the most targeted so the Govt. of that country should increase their defence and investigation departments and also should ban bombing/explosion raw material.
3. Make people aware about terrorism.
4. Anti terrorism organizations and defence department should keep an eye on Taliban and ISIL which are most active organisations.
5. World should make a strong law act and actions against the Terrorism,

Conclusion

1. Attacks has increased but number of people killed manier times as attack happened.
2. Iraq has the most attacks.
3. The Middle East and North Africa Regions has most taregeted.
4. Maximum number of attacks are from Bombing/Explosions.
5. There are maximum number of attacks in Private citizens and Property.
6. Taliban and ISIL has a most active organisation.
7. It is evident form the trend analysis that since 1971 there has been significant increase in terror attacks globally. Terrorist Groups like ISIL, taliban, Al-Shabaab , BOko Haram, NPA, assassination, etc. However, in recent times there has been slight decrease in terrorist attacks.
8. We need to undeerstand that every human live is precious and we should take all efforts to curb terrorism and sponsors of terrorism. Development of both socio econimic and educational are the only permanent solution to this problem.
9. We should make common people aware about the terrorism.

Hurrah! You have successfully completed your EDA Capstone Project !!!

✓ 0s completed at 1:20 PM