

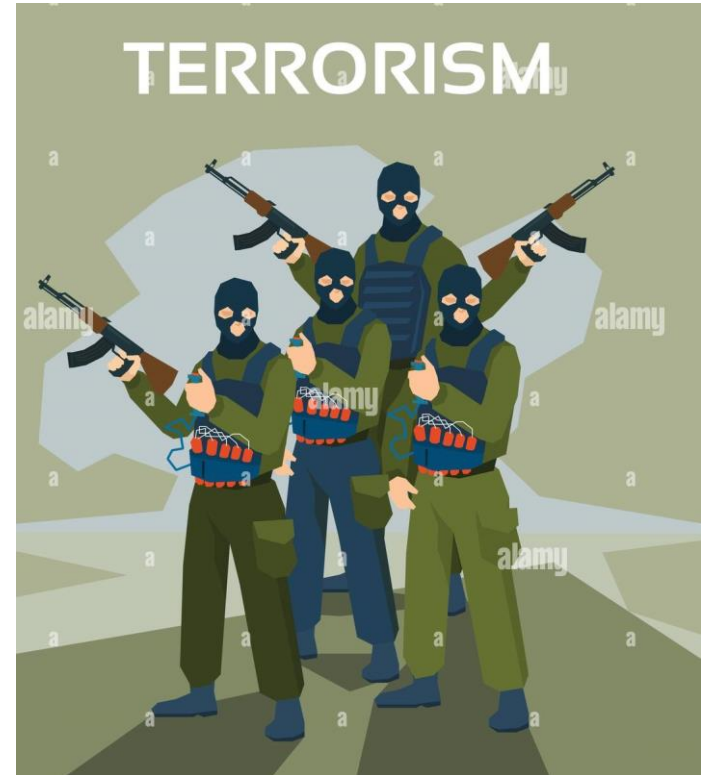
Capstone Project – 1

Global Terrorism Analysis

Presented by:
Pavan Kale

Content

1. Introduction
2. Data Explanation and Preparation
3. Characteristics of GTD
4. Exploratory Data Analysis
5. Data Visualization
6. Conclusion



Introduction

- The Global Terrorism Database (GTD) is the most comprehensive unclassified database of terrorist attacks in the world. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) makes the GTD available via this site in an effort to improve understanding of terrorist violence, so that it can be more readily studied and defeated. The GTD is produced by a dedicated team of researchers and technical staff.
- The GTD is an open-source database, which provides information on domestic and international terrorist attacks around the world since 1970, and now includes more than 200,000 events. For each event, a wide range of information is available, including the date and location of the incident, the weapons used, nature of the target, the number of casualties, and when identifiable the group or individual responsible.

GTD is an open-source database, which provides information on domestic and international terrorist attacks around the world since 1970, and now includes more

Characteristics of GTD

- GTD Contain information on over 200000 terrorist attack. Currently the most comprehensive unclassified on terrorist attacks in the world. Includes the information on more than 88,000 bombings, 19,000 assassinations and 11,000 wound since 1970.
- Our Task is to analyze the terrorist attacks during this 46 year period, which includes 135 different variables and over 181691 incidents.

Summary of Data

This Database consists of information between 181691 Rows and 135 Columns (1970 to 2014).

Some of the variables explained:

year: which contains the year in which the incident occurred.

Country_txt: which identifies the country or location where the incident occurred.

extended: The duration of an incident extended more than 24 hours (1) - Less than 24 hours (0).

multiple: The attack is part of a multiple incident (1) - Not part of a multiple incident (0).

suicide: The incident was a suicide attack (1) - There is no indication of a suicide attack (0).

attacktype_txt: Assassination(1), Hijacking(2), Kidnapping(3), Barricade Incident(4), Bombing/Explosion(5), Armed Assault(6), Unarmed Assault(7), Facility/Infrastructure Attack(8), Unknown(9).

targettype_txt: 22 categories ranging from Business(1), Government(general)(2), Police(3),...Utilities(21), Violent Political Parties(22)

natlty: Nationality of target/victim.

claimed: A group or person claimed responsibility for the attack (1) - No claim of responsibility was made(0).

weaptype: 13 categories ranging from Biological(1), Chemical(2), Radiological(3),...Other(12), Unknown(13)

nkill: Total number of fatalities including all victims and attackers who died as a direct result of the incident.

nkillter: Limited to only perpetrator fatalities.

nwound: Total number of injured victims and attackers.

nwoundte: Total number of injured perpetrators.

Dataset Loading



```
#default it shows first five rows
df.head()
```

	eventid	iyear	imonth	iday	approxdate	extended	resolution	country	country_txt	region	region_txt	provstate	city	latitude	longitude	specificity	vicinity	location
0	197000000001	1970	7	2	NaN	0	NaN	58	Dominican Republic	2	Central America & Caribbean	NaN	Santo Domingo	18.456792	-69.951164	1.0	0	NaN
1	197000000002	1970	0	0	NaN	0	NaN	130	Mexico	1	North America	Federal	Mexico city	19.371887	-99.086624	1.0	0	NaN
2	197001000001	1970	1	0	NaN	0	NaN	160	Philippines	5	Southeast Asia	Tarlac	Unknown	15.478598	120.599741	4.0	0	NaN
3	197001000002	1970	1	0	NaN	0	NaN	78	Greece	8	Western Europe	Attica	Athens	37.997490	23.762728	1.0	0	NaN
4	197001000003	1970	1	0	NaN	0	NaN	101	Japan	4	East Asia	Fukouka	Fukouka	33.580412	130.396361	1.0	0	NaN

location	summary	crit1	crit2	crit3	doubtterr	alternative	alternative_txt	multiple	success	suicide	attacktype1	attacktype1_txt	attacktype2	attacktype2_txt	attacktype3
NaN	NaN	1	1	1	0.0	NaN	NaN	0.0	1	0	1	Assassination	NaN	NaN	NaN
NaN	NaN	1	1	1	0.0	NaN	NaN	0.0	1	0	6	Hostage Taking (Kidnapping)	NaN	NaN	NaN
NaN	NaN	1	1	1	0.0	NaN	NaN	0.0	1	0	1	Assassination	NaN	NaN	NaN
NaN	NaN	1	1	1	0.0	NaN	NaN	0.0	1	0	3	Bombing/Explosion	NaN	NaN	NaN
NaN	NaN	1	1	1	-9.0	NaN	NaN	0.0	1	0	7	Facility/Infrastructure Attack	NaN	NaN	NaN

Attribute Information

```
#total null values in the column  
df.isna().sum()
```

eventid	0
Year	0
Month	0
Day	0
approxdate	172452
extended	0
resolution	179471
country	0
Country	0
region	0
Region	0
State	421
city	434
Latitude	4556
Longitude	4557
specificity	6
vicinity	0
location	126196
Summary	66129
crit1	0
crit2	0
crit3	0
doubtterr	1
alternative	152680
alternative_txt	152680
multiple	1
Success	0
suicide	0
attacktype1	0
Attack_type	0
attacktype2	175377

attacktype2_txt	175377
attacktype3	181263
attacktype3_txt	181263
targetype1	0
Target_type	0
targsubtype1	10373
targsubtype1_txt	10373
corp1	42550
Target	636
natlty1	1559
natlty1_txt	1559
targetype2	170547
targetype2_txt	170547
targsubtype2	171006
targsubtype2_txt	171006
corp2	171574
target2	170671
natlty2	170863
natlty2_txt	170863
targetype3	180515
targetype3_txt	180515
targsubtype3	180594
targsubtype3_txt	180594
corp3	180665
target3	180516
natlty3	180544
natlty3_txt	180544
Group_name	0
gsubname	175801
gname2	179678
gsubname2	181531

```
#null values percentage  
df.isna().sum()*100/len(df)
```

eventid	0.000000
Year	0.000000
Month	0.000000
Day	0.000000
approxdate	94.914993
extended	0.000000
resolution	98.778145
country	0.000000
Country	0.000000
region	0.000000
Region	0.000000
State	0.231712
city	0.238867
Latitude	2.507554
Longitude	2.508104
specificity	0.003302
vicinity	0.000000
location	69.456385
Summary	36.396409
crit1	0.000000
crit2	0.000000
crit3	0.000000
doubtterr	0.000550
alternative	84.032781
alternative_txt	84.032781
multiple	0.000550
Success	0.000000
suicide	0.000000
attacktype1	0.000000
Attack_type	0.000000
attacktype2	96.524869

attacktype2_txt	96.524869
attacktype3	99.764435
attacktype3_txt	99.764435
targetype1	0.000000
Target_type	0.000000
targsubtype1	5.709144
targsubtype1_txt	5.709144
corp1	23.418882
Target	0.350045
natlty1	0.858050
natlty1_txt	0.858050
targetype2	93.866510
targetype2_txt	93.866510
targsubtype2	94.119136
targsubtype2_txt	94.119136
corp2	94.431755
target2	93.934757
natlty2	94.040431
natlty2_txt	94.040431
targetype3	99.352747
targetype3_txt	99.352747
targsubtype3	99.396228
targsubtype3_txt	99.396228
corp3	99.435305
target3	99.353298
natlty3	99.368708
natlty3_txt	99.368708
Group_name	0.000000
gsubname	96.758232
gname2	98.892075
gsubname2	99.911938
gname3	99.821675
gsubname3	99.988992
Motive	72.171984

Data Processing

So many columns have null values so we removed that columns.

```
'''In this dataset most of the columns have null values and that not contribute into the analysis so we removed the null values columns
and keep only necessary columns from the dataset.'''
column_to_keep = ['Year', 'Month', 'Day', 'State', 'Region', 'Latitude', 'Longitude', 'Attack_type', 'Country', 'Success', 'Target', 'Killed', 'Wounded', 'Summary',
                  'Group_name', 'Target_type', 'Weapon_type', 'Motive']
```

```
df = df[column_to_keep]
```

```
df.head(5)
```

	Year	Month	Day	State	Region	Latitude	Longitude	Attack_type	Country	Success	Target	Killed	Wounded	Summary	Group_name	Target_type	Weapon_type	Motive
0	1970	7	2	NaN	Central America & Caribbean	18.456792	-69.951164	Assassination	Dominican Republic	1	Julio Guzman	1.0	0.0	NaN	MANO-D	Private Citizens & Property	Unknown	NaN
1	1970	0	0	Federal	North America	19.371887	-99.086624	Hostage Taking (Kidnapping)	Mexico	1	Nadine Chaval, daughter	0.0	0.0	NaN	23rd of September Communist League	Government (Diplomatic)	Unknown	NaN
2	1970	1	0	Tarlac	Southeast Asia	15.478598	120.599741	Assassination	Philippines	1	Employee	1.0	0.0	NaN	Unknown	Journalists & Media	Unknown	NaN
3	1970	1	0	Attica	Western Europe	37.997490	23.762728	Bombing/Explosion	Greece	1	U.S. Embassy	NaN	NaN	NaN	Unknown	Government (Diplomatic)	Explosives	NaN
4	1970	1	0	Fukouka	East Asia	33.580412	130.396361	Facility/Infrastructure Attack	Japan	1	U.S. Consulate	NaN	NaN	NaN	Unknown	Government (Diplomatic)	Incendiary	NaN

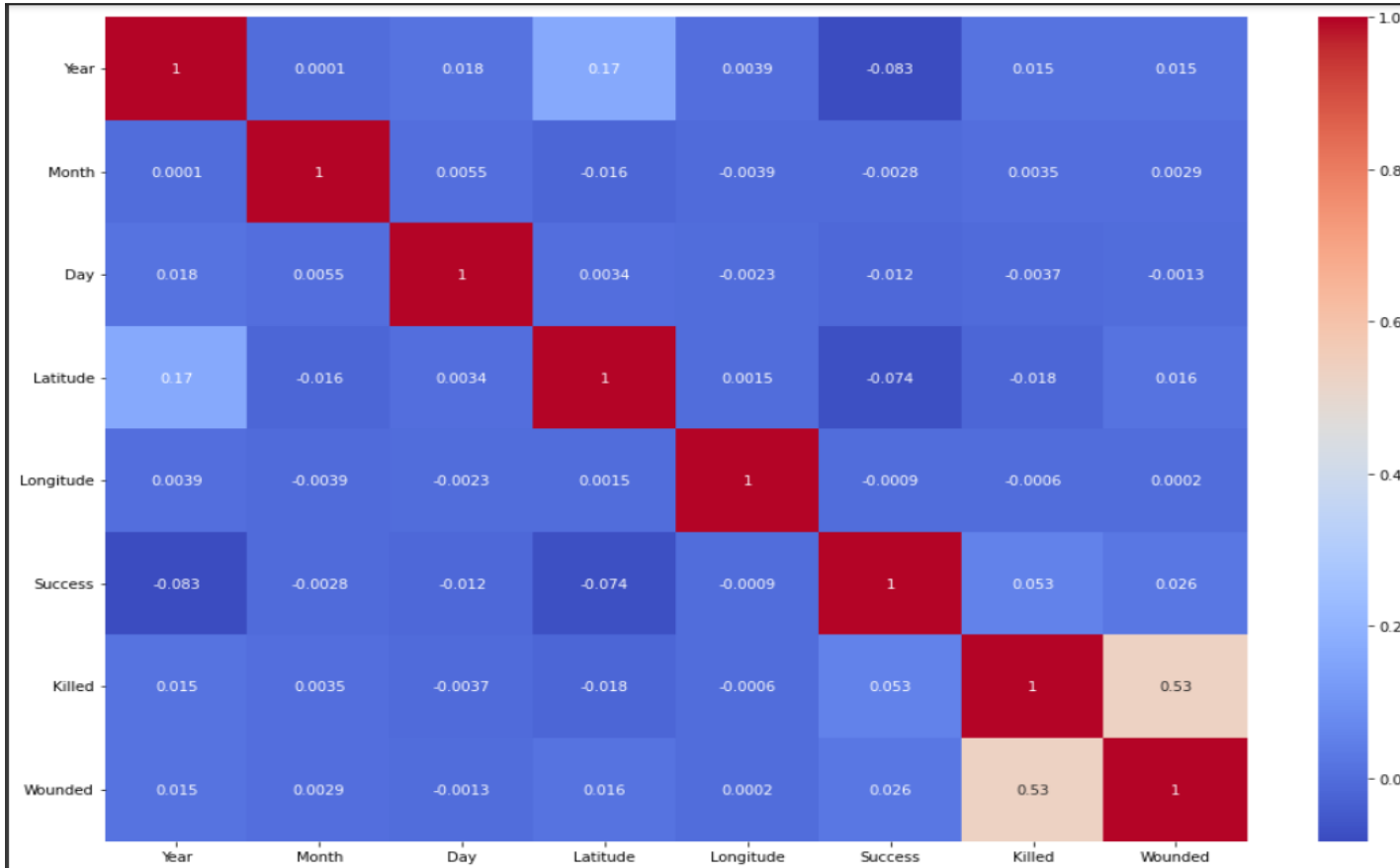
Descriptive Statistics

This is the Summary of new database.

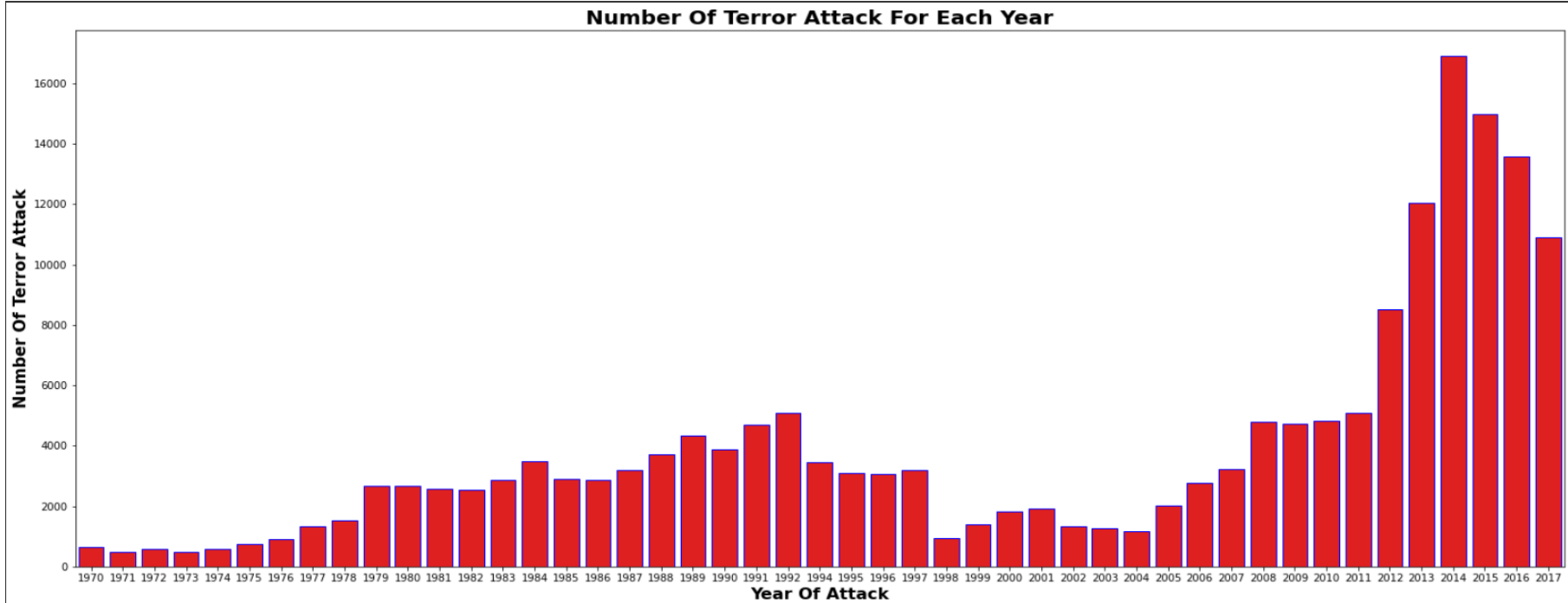
```
#summary of a dataset  
df.describe()
```

	Year	Month	Day	Latitude	Longitude	Success	Killed	Wounded
count	181691.000000	181691.000000	181691.000000	177135.000000	1.771340e+05	181691.000000	171378.000000	165380.000000
mean	2002.638997	6.467277	15.505644	23.498343	-4.586957e+02	0.889598	2.403272	3.167668
std	13.259430	3.388303	8.814045	18.569242	2.047790e+05	0.313391	11.545741	35.949392
min	1970.000000	0.000000	0.000000	-53.154613	-8.618590e+07	0.000000	0.000000	0.000000
25%	1991.000000	4.000000	8.000000	11.510046	4.545640e+00	1.000000	0.000000	0.000000
50%	2009.000000	6.000000	15.000000	31.467463	4.324651e+01	1.000000	0.000000	0.000000
75%	2014.000000	9.000000	23.000000	34.685087	6.871033e+01	1.000000	2.000000	2.000000
max	2017.000000	12.000000	31.000000	74.633553	1.793667e+02	1.000000	1570.000000	8191.000000

Correlation Analysis

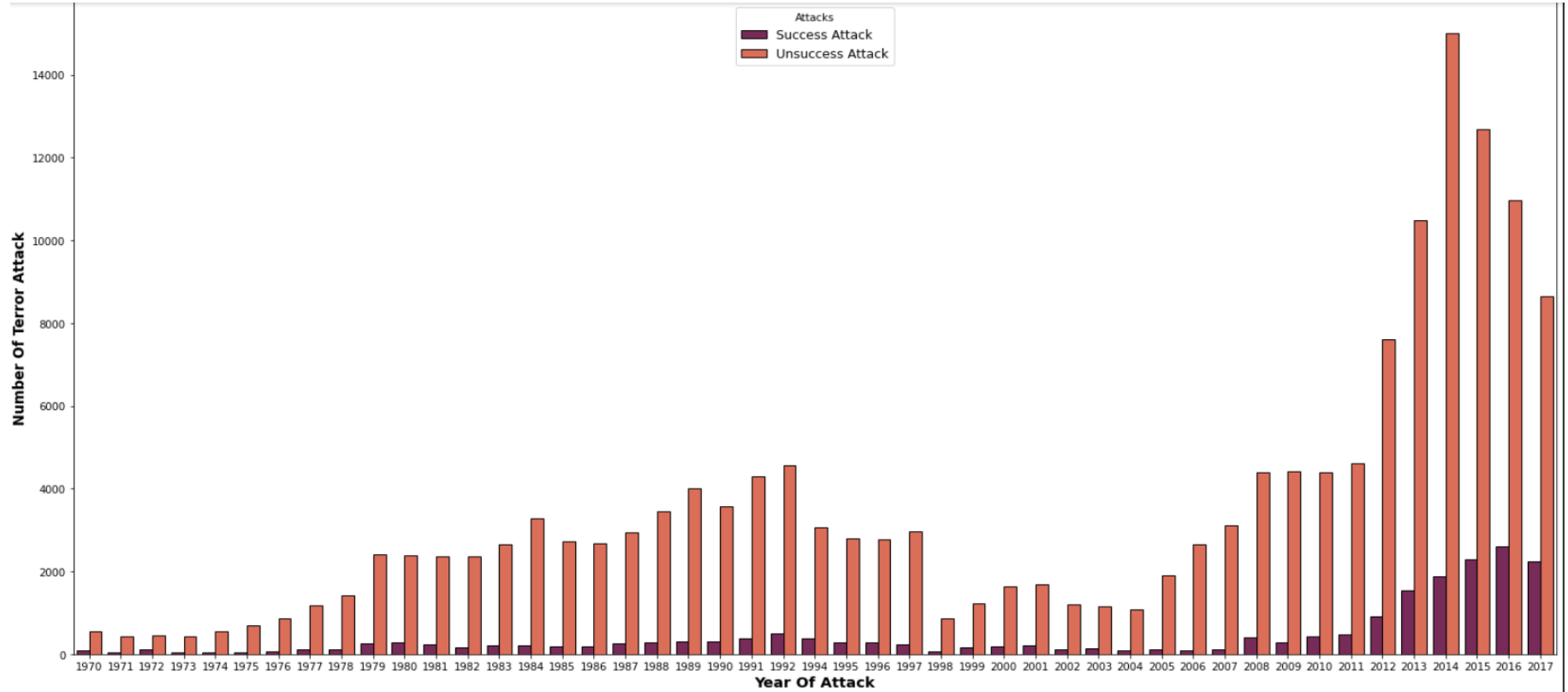


Exploratory Data Analysis



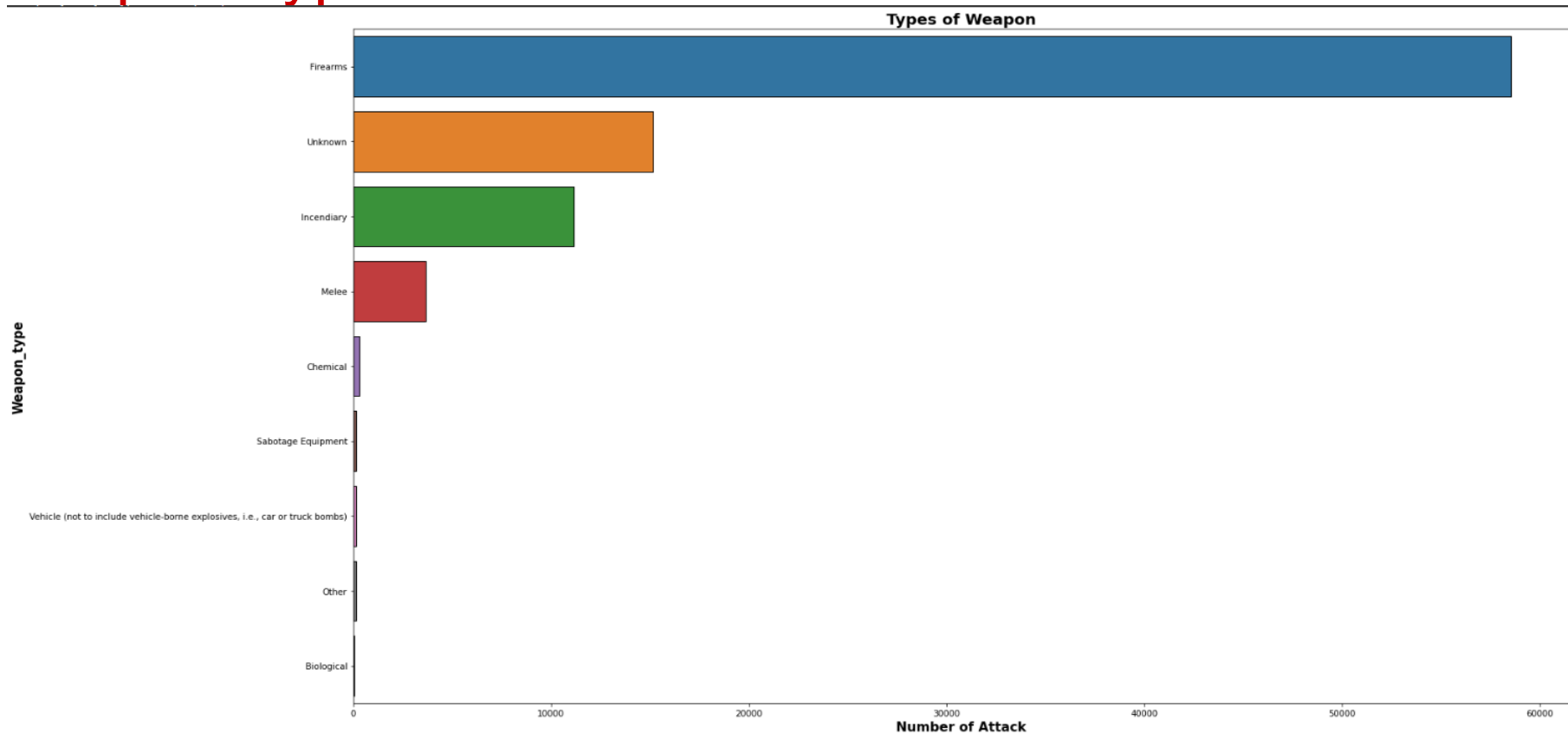
In Year 2014 and 2015 has large number of Terrorist Activities happen.

Number of Success Attacks Each Year



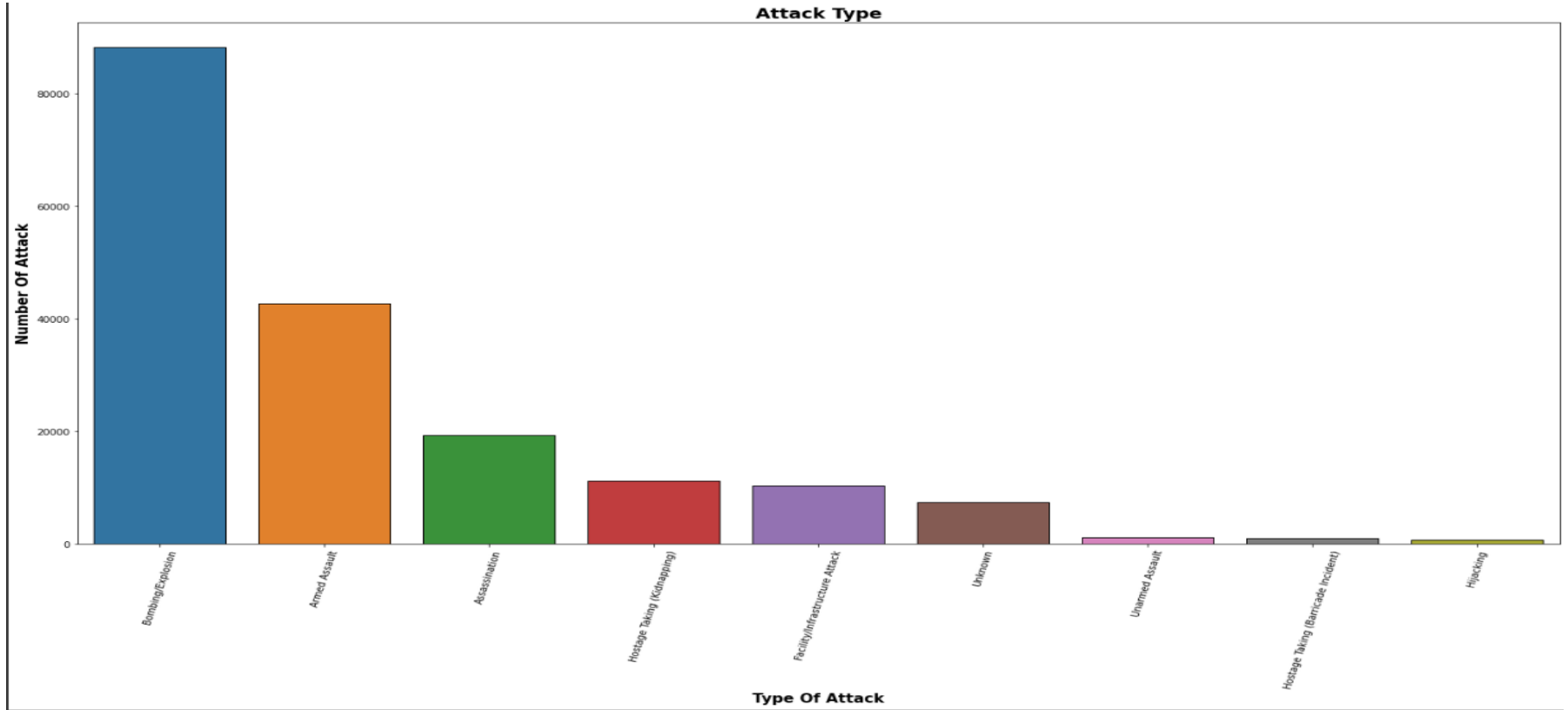
In 2016 Most Successful Terrorist Attack done by terrorist.

Weapon Type



Firearms weapon have used most times.

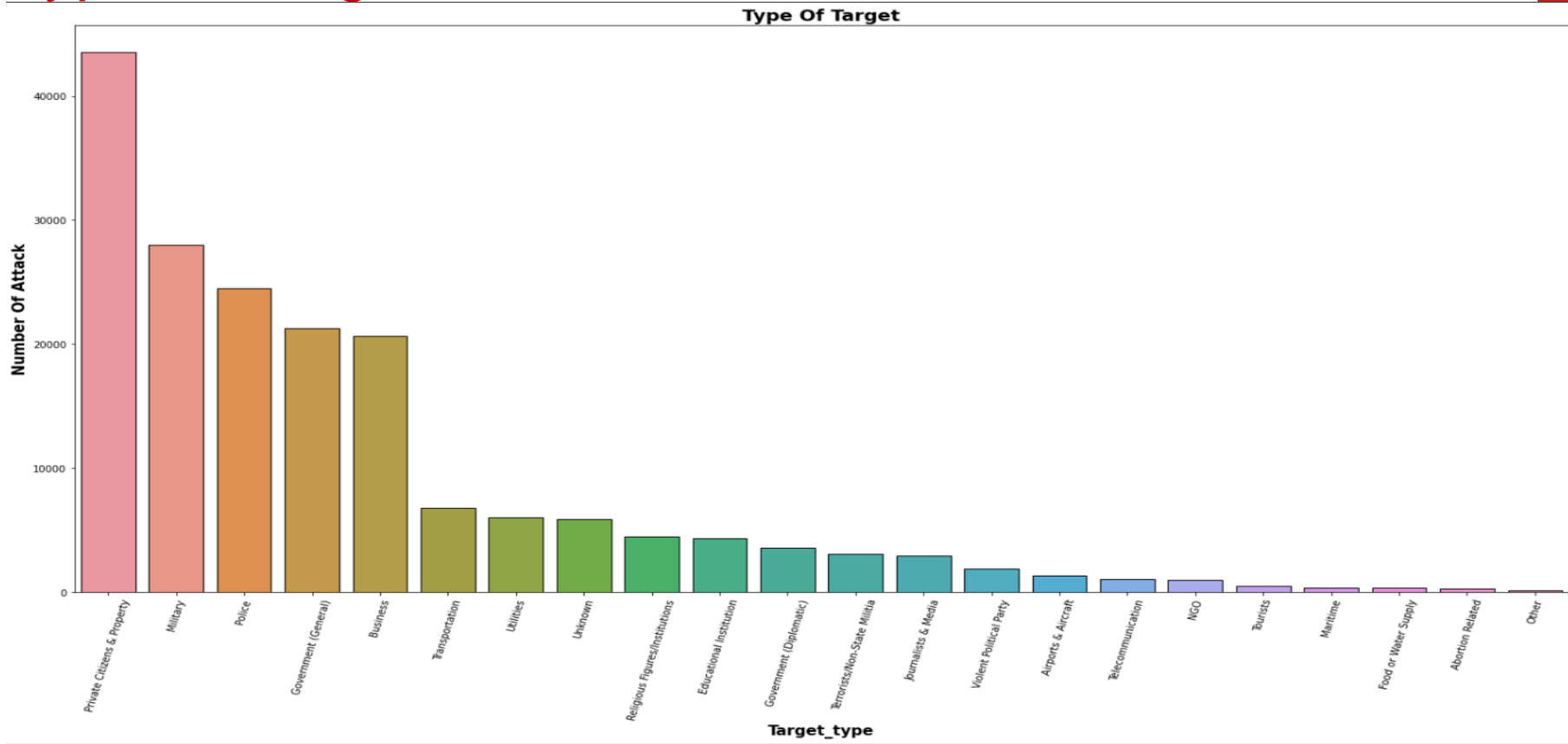
Type of Attack



Most of Attack are from Bombing/Explosion and Armed Assault.

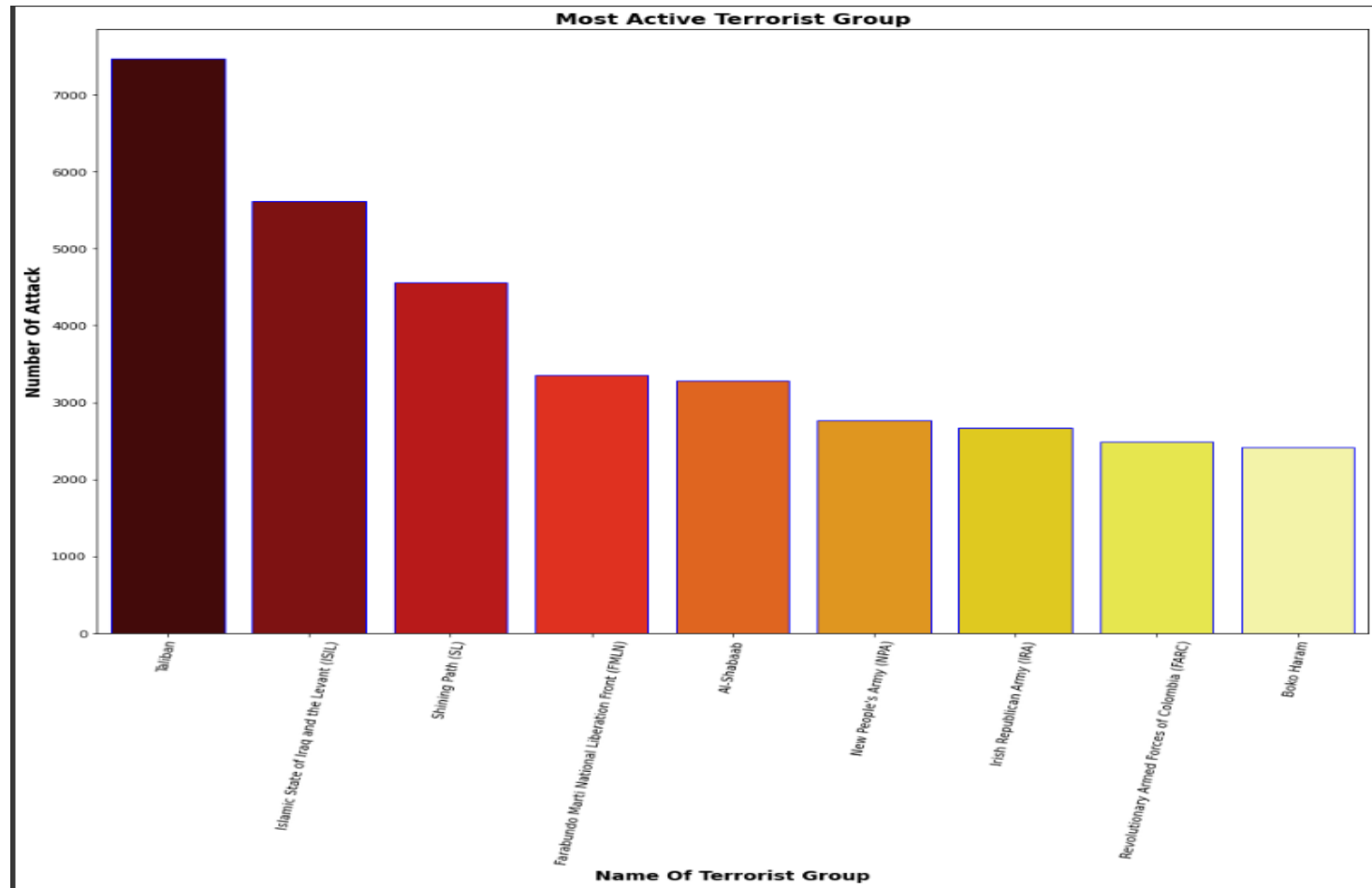
Type of Target

AI



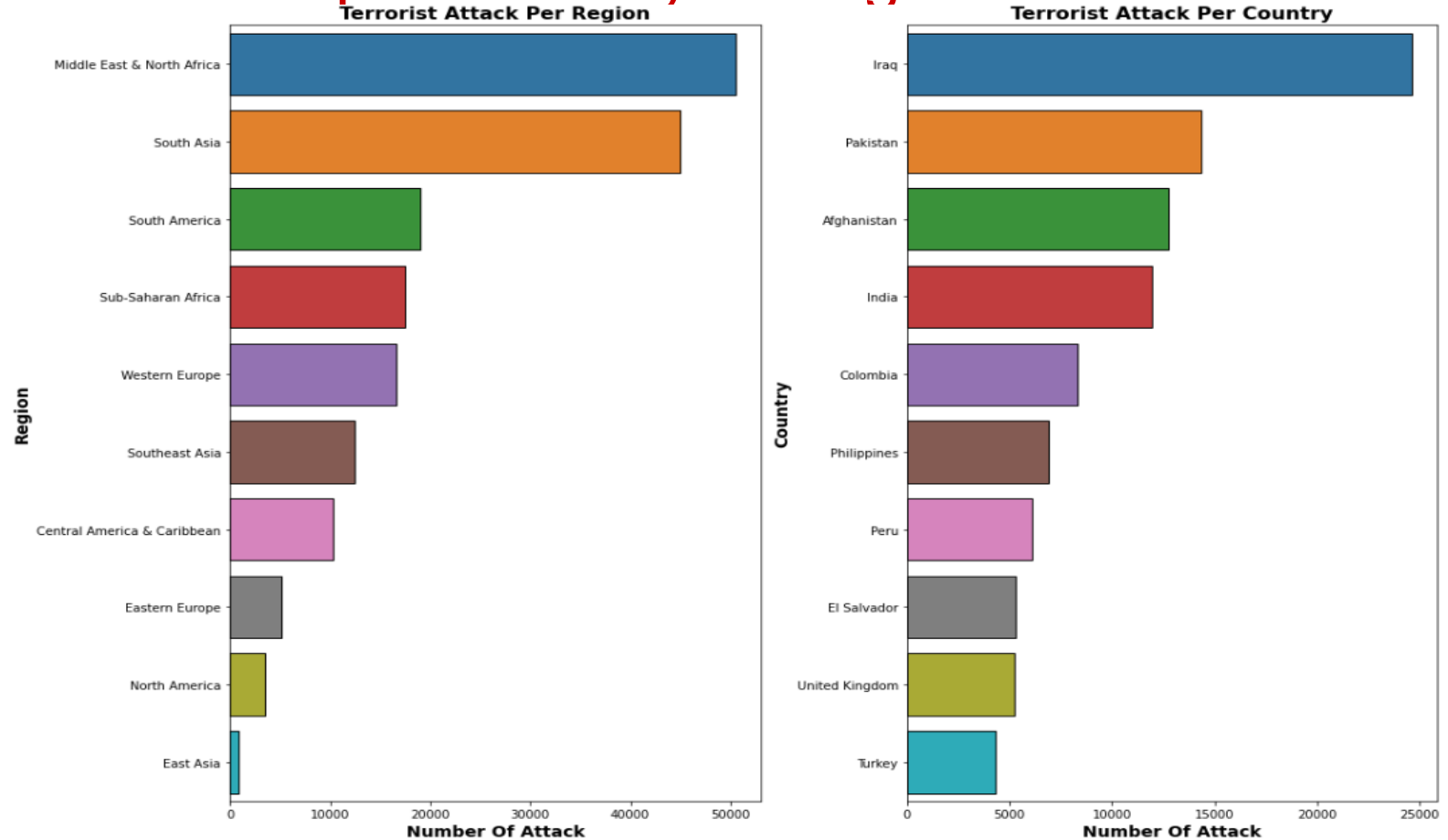
Main Target of Terrorist is Private Citizens & Property and Military.

Most Active Terrorist Organization



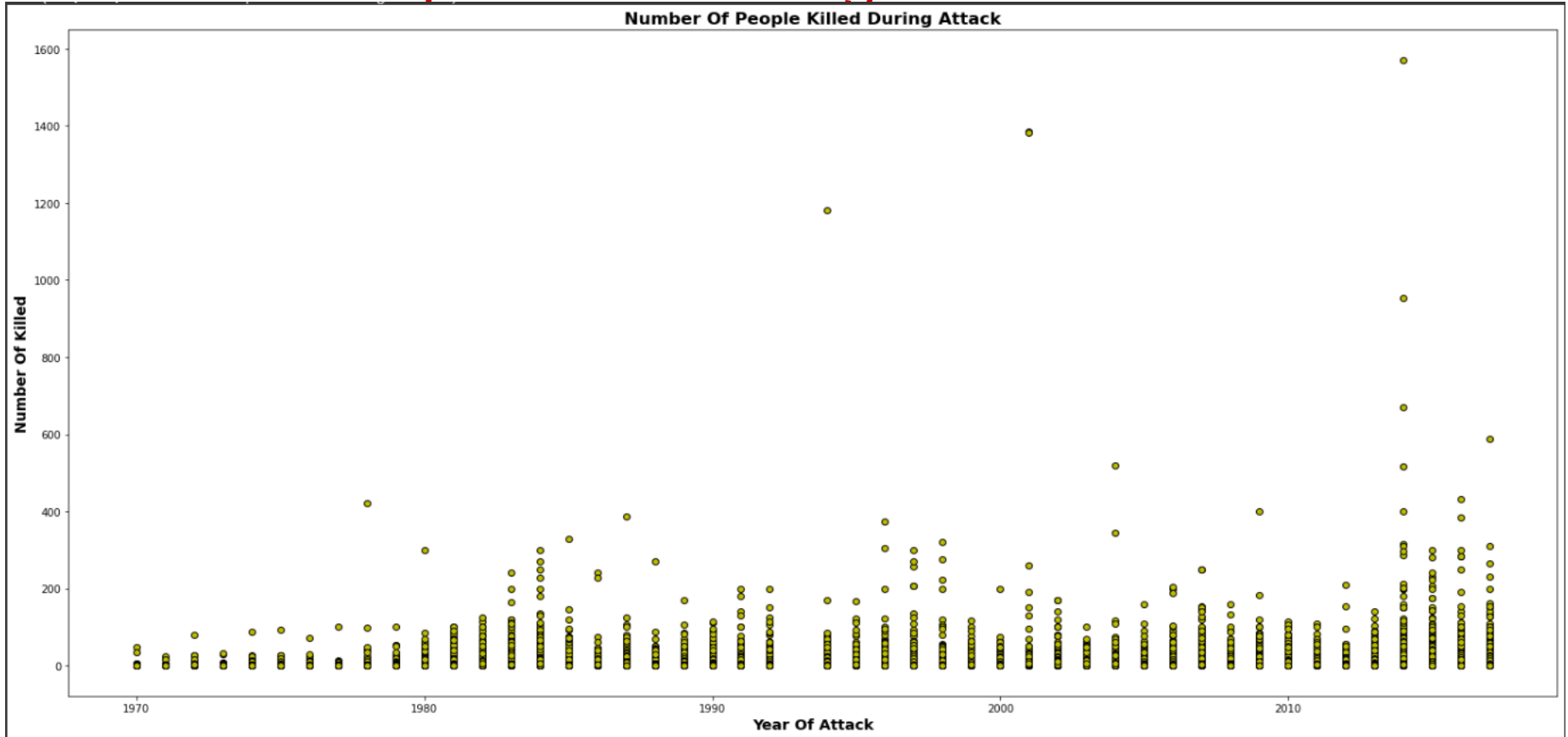
Taliban and ISIL are most Active Terrorist Group

Terrorist Attack per Country & Region



Maximum Terrorist Activities happen in Iraq and in Middle East & North Africa region.

Number of People Killed during the attack



In 2014 maximum people are died in terrorist attacks.

Conclusion

1. Most Active Terrorist Organization are Taliban and Islamic State of Iraq and the Levant(ISIL).
2. In 2014 Maximum people died in Terrorist Attack.
3. Maximum Number of 'Firearms Weapon' are used for Terrorist Attack.
4. Iraq ranked first on global terrorist activity and Middle East & North Africa and South
5. Private Citizens & Property and Military is the main Target of Terrorism.
Asia has most Terrorist region.

Thank You