Capstone Project

Global Terrorism Analysis

By- Rohit Verma

Points for Discussion

- Data Pipelines
- Data Summary
- Data Explanation
- Data Processing
- Data Preparation & Cleaning
- Data Analysis & Visualization
- Conclusions

DATA PIPELINES

- **Data Processing-1:** In the first process we'have checking the unnecessary features and removed those features which are not relevant for this dataset. Since there are so many columns with null values.
- **Data Processing-2:** In this part, I manually go through all the features selected from first part, then encoded their categorial features, change the columns which containing date time values and also their columns names.
- **EDA:** In this part, we do some exploratory data analysis (EDA) on the necessary data selected in part-1 and part-2 to see their trend.
- **Creating a model:** Finally, In the last but not the least part we create data models. By creating a model we can easily understand the internal point of the data and make the data to be more elaborative and simple to observe. It's start from simple model then one by one it's become complex for the better outputs.

Data Summary

- This dataset gives us exhaustive analysis on the ‘‘**Global Terrorism**’’ since 1970s. The dataset contains 1,80,000 data on different terror activities across the globe and their scale and impact as a whole.

- We will also look into some of the intricate details such as 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.

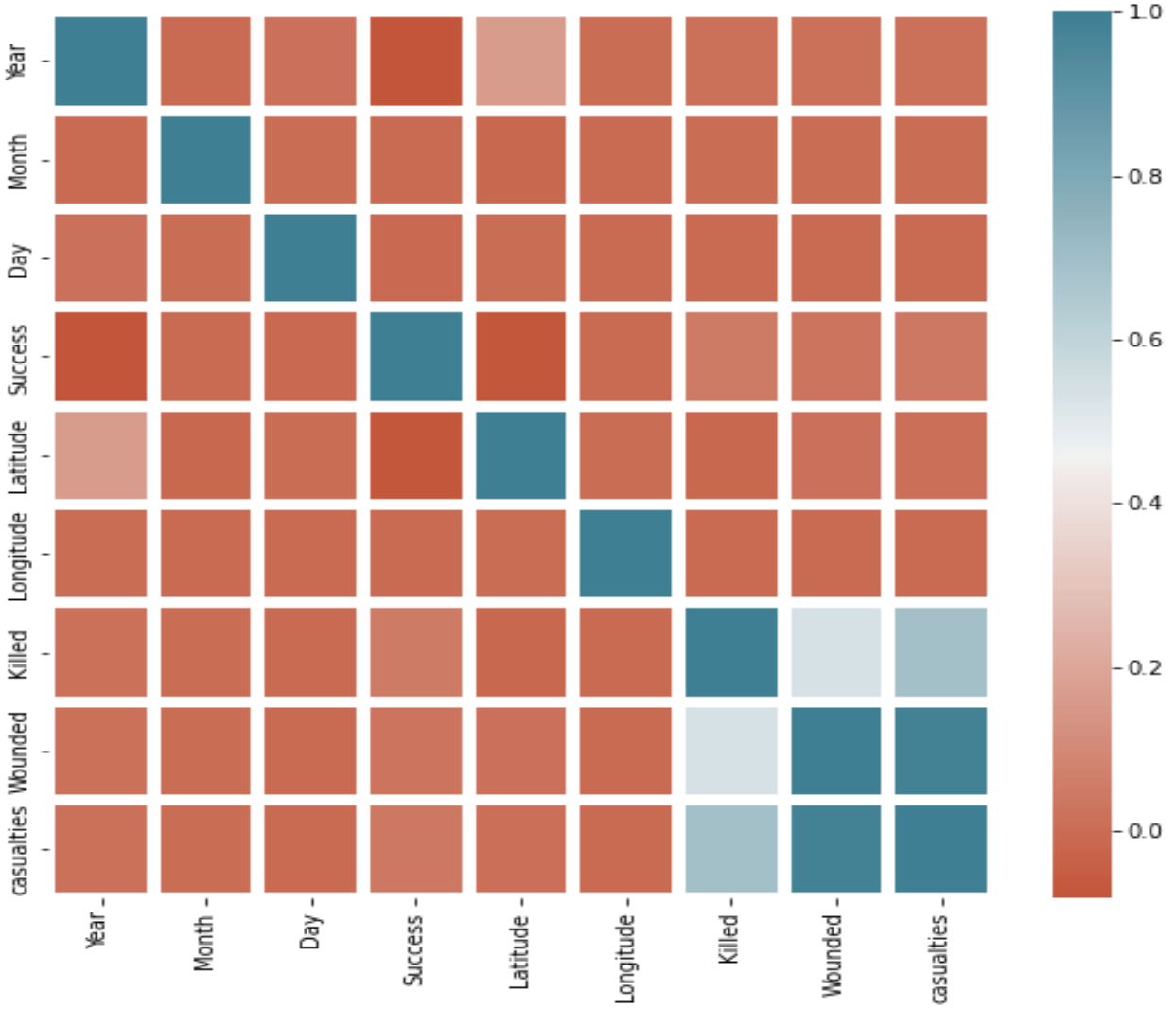
- During 1970-2018 terrorist attacks have greatly shifted from Latin America, Caribbean(LAC) , Europe and Central Asia (ECA) to the Middle East and North Africa (MENA), South Asia (SAS) and sub Sahara Africa (SSA) so that enhanced counterterrorism vigilance and resources are particularly needed in SAS, SSA and MENA.

Data Explanation

The Global Terrorism Dataset consist of

User information:-
year,month,day,success,latitude,longitude,killed,wounded,casualties, etc.

Most of them in my dataset are ‘object’, but few are not.



Data Processing

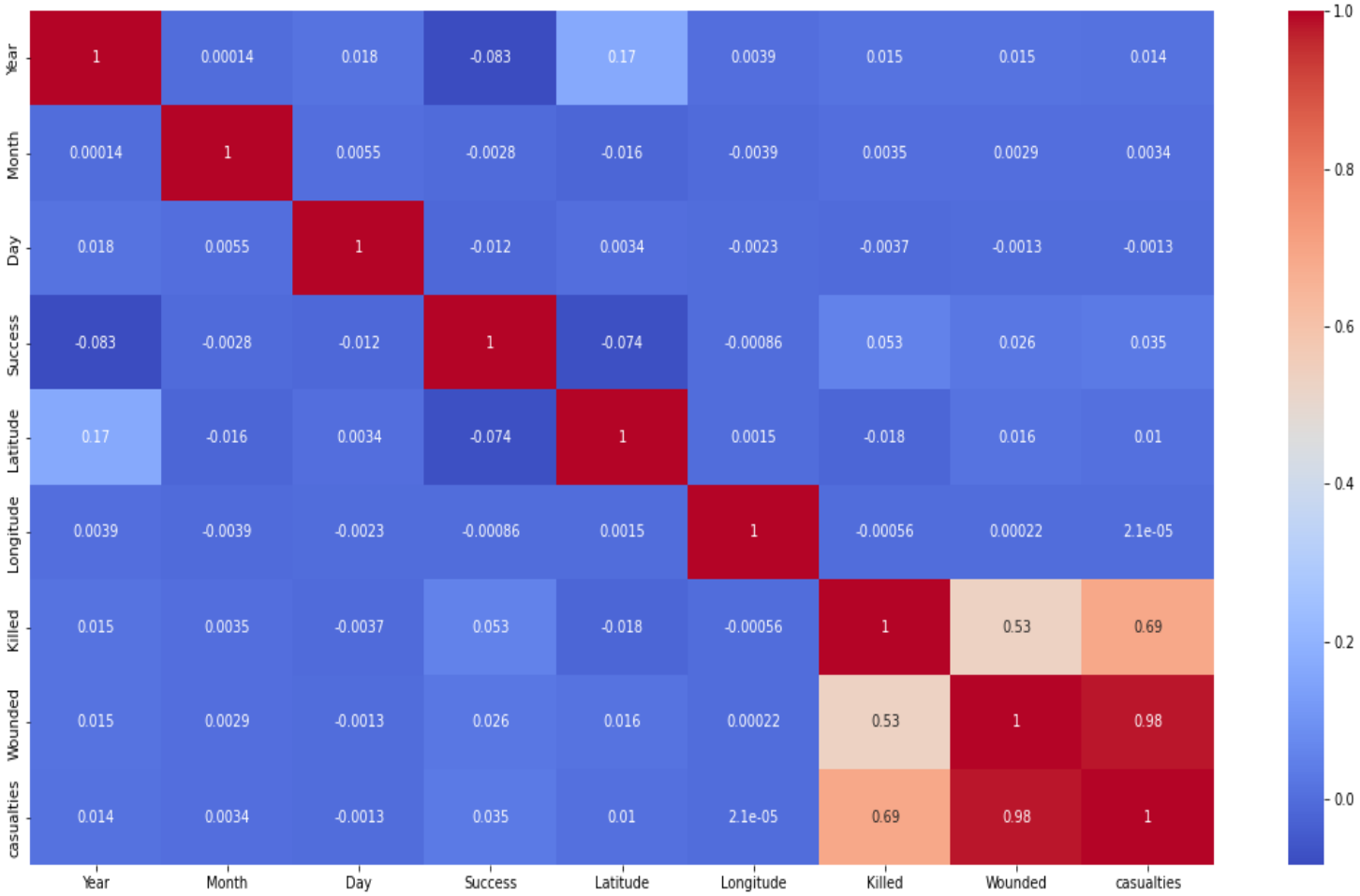
- It is essential to view attribute correlations to select the best features for modeling

- The visual to the right conveys the correlations of the remaining attributes

  Color = +/- Correlation

  Size = Intensity of Correlation

Correlation between columns in dataset



Data Processing (Cont'd)

terror_data.count()

- There are 181691 rows and 135 columns
- It shows clearly that the mean of the countries with crime rate is approximately 132

```
terror_data.count
```

```
<bound method DataFrame.count of             eventid  iyear  imonth  iday  ... INT_IDEO  INT_MISC INT_ANY  related
0       197000000001   1970       7     2  ...        0         0       0      NaN
1       197000000002   1970       0     0  ...        1         1       1      NaN
2       197001000001   1970       1     0  ...       -9         1       1      NaN
3       197001000002   1970       1     0  ...       -9         1       1      NaN
4       197001000003   1970       1     0  ...       -9         1       1      NaN
...              ...    ...     ...   ...  ...      ...       ...     ...      ...
181686  201712310022   2017      12    31  ...        0         0       0      NaN
181687  201712310029   2017      12    31  ...       -9         1       1      NaN
181688  201712310030   2017      12    31  ...        0         0       0      NaN
181689  201712310031   2017      12    31  ...       -9         0      -9      NaN
181690  201712310032   2017      12    31  ...       -9         0      -9      NaN

[181691 rows x 135 columns]>
```

Here i can cleaerly saw that the mean of the countries with crime rate is approximately 132 and total is 181691.

Data Preparation & Cleaning

The dataset has 135 columns with different types of data and rows. This is huge for our exploration so we will get rid of some of columns based on some criteria like:

Rate of missing data

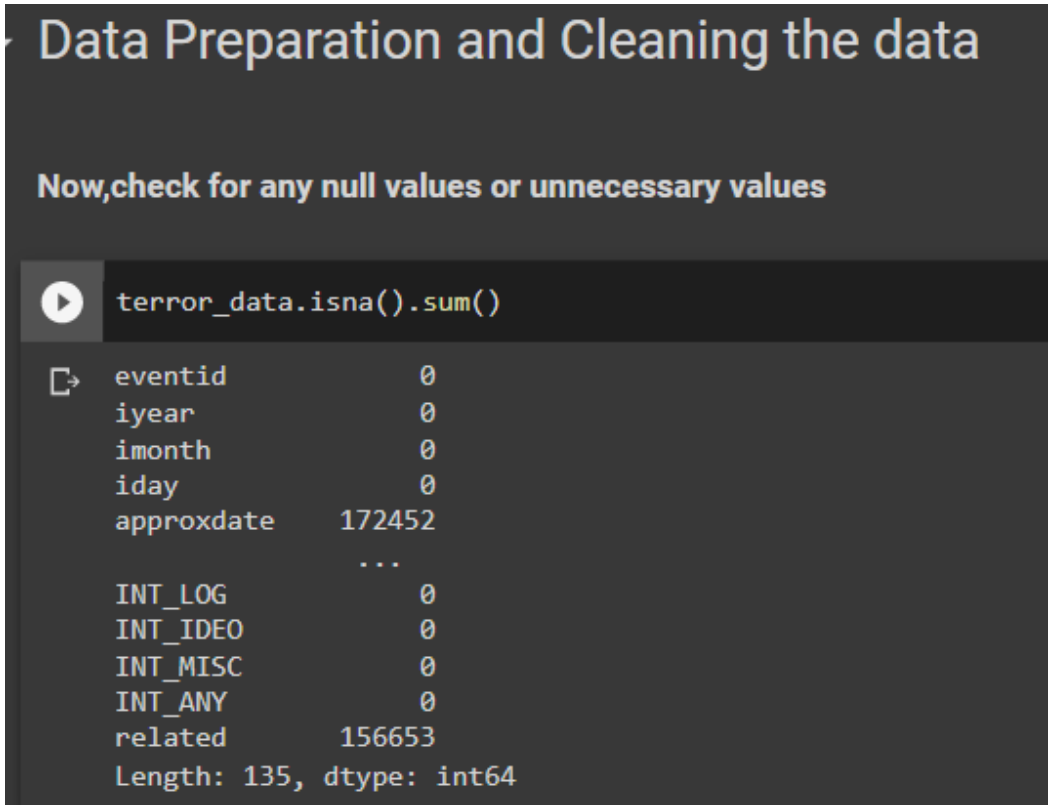
Reccurence of info() by columns

Utility of info() by columns

Data Preparation & Cleaning (Cont'd)

- We start our exploration by reducing the number of columns based on the previous criteria
- Checking null or unnecessary values

Null values

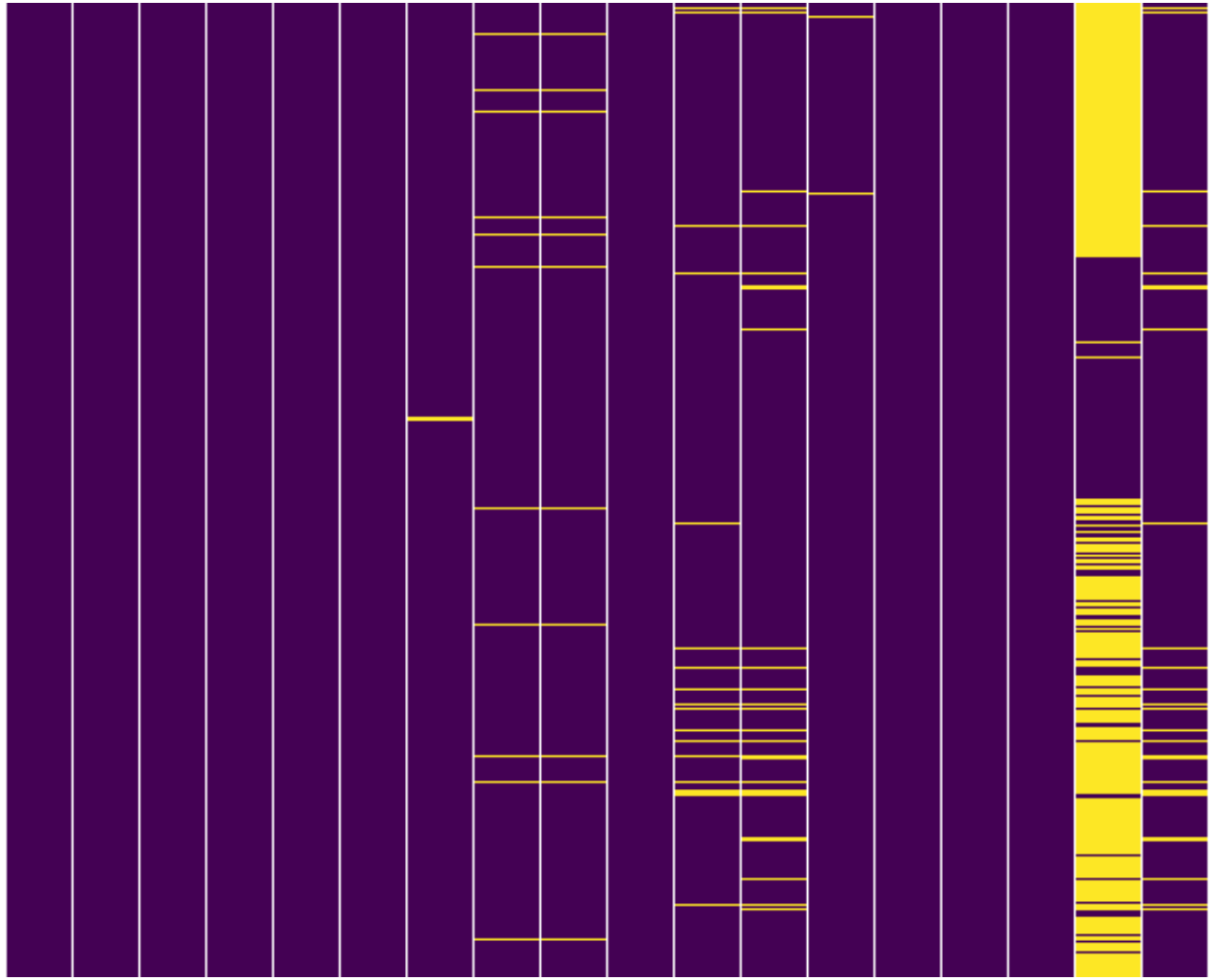


Data Prepration & Cleaning (Cont'd)

Heatmap of null values

Columns with null values

- State
- City
- Latitude
- Longitude
- Kill
- Wound
- Summary
- Motive



Data Preparation & Cleaning (Cont'd)

- Now we will removing the null values which make the data more complex & unrelevant
- By dropping the unnecessary columns from the dataset and taking the only relevant necessary features which is required to analyzing the data model

```
terror_data=terror_data.copy()
terror_data.dropna(axis=1,how='all',thresh=None,subset=None,inplace=False)
```

	eventid	iyear	imonth	iday	approxdate	extended	resolution	country	country_txt	region	region_txt	provstate	city	latitude	longitude	s
0	197000000001	1970	7	2	NaN	0	NaN	58	Dominican Republic	2	Central America & Caribbean	NaN	Santo Domingo	18.456792	-69.951164	
1	197000000002	1970	0	0	NaN	0	NaN	130	Mexico	1	North America	Federal	Mexico city	19.371887	-99.086624	
2	197001000001	1970	1	0	NaN	0	NaN	160	Philippines	5	Southeast Asia	Tarlac	Unknown	15.478598	120.599741	
3	197001000002	1970	1	0	NaN	0	NaN	78	Greece	8	Western Europe	Attica	Athens	37.997490	23.762728	
4	197001000003	1970	1	0	NaN	0	NaN	101	Japan	4	East Asia	Fukouka	Fukouka	33.580412	130.396361	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
181686	201712310022	2017	12	31	NaN	0	NaN	182	Somalia	11	Sub-Saharan Africa	Middle Shebelle	Ceelka Geelow	2.359673	45.385034	

Data Analysis & Visualization

Analyzing the data:

Terrorism attacks and their activities includes:

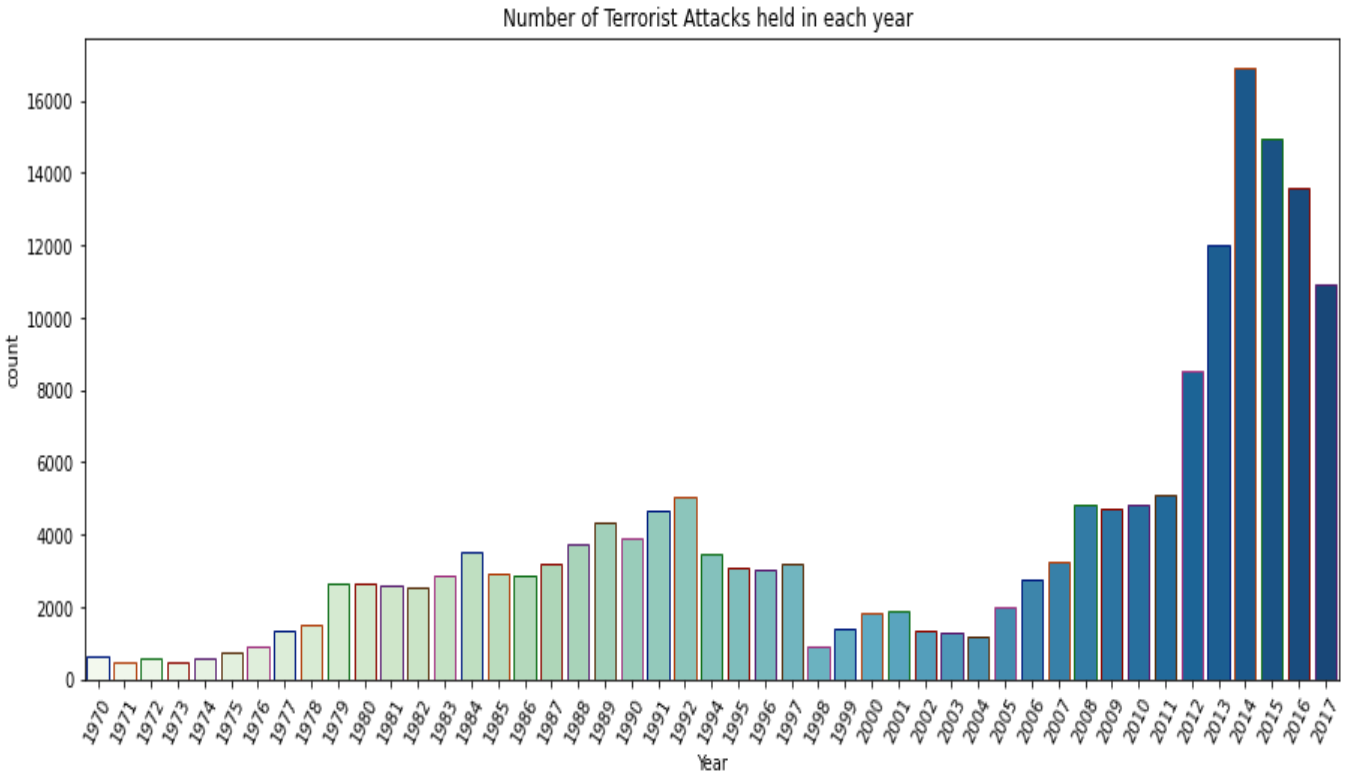
- Motive, Methods, Target
- Hot Zone Areas
- Terrorism By Country
- Groups
- Analysis on India

(Motive)

- While analyzing the data we find that the motive behind the terrorist attacks and terrorism are unknown.
- By seeing this we get on the point that we are missing something behind finding the motive of terrorist in the real life

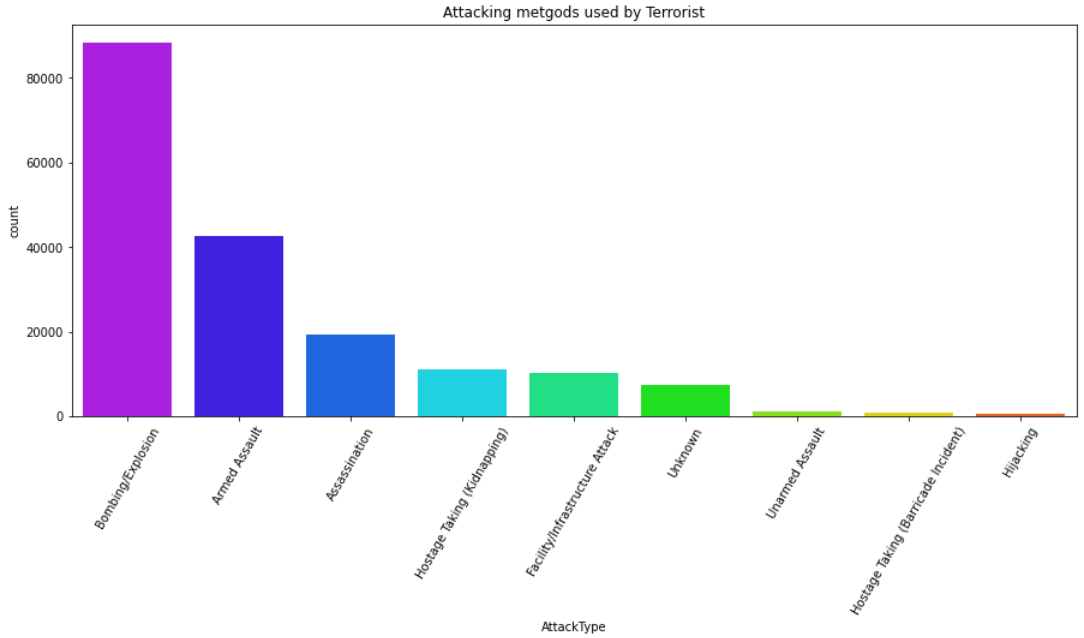
(Activities Vs Year)

- This graph shows number of terrorist attacks held in each year and their intensity of enlarging
- Here we can see that the terrorist attacks since 1970 are increasing rapidly by each year one by one upto 2014 then it’s gradually drop from 2014 to 2017 and may be it drop further which is a good factor for a world



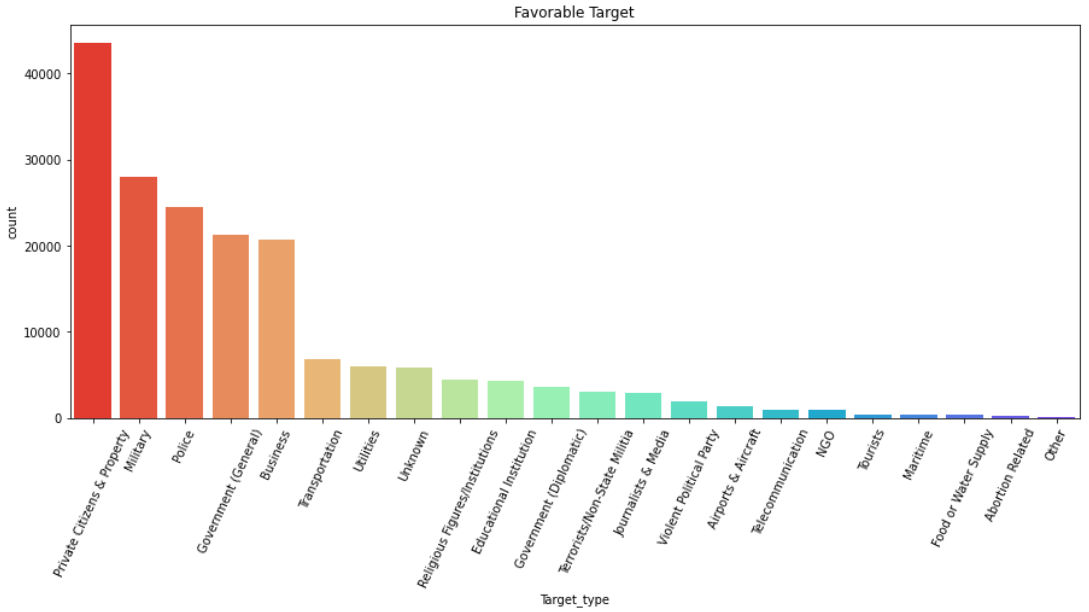
(Methods Vs Target)

Attacking methods used by terrorists



Bombing/Explosives

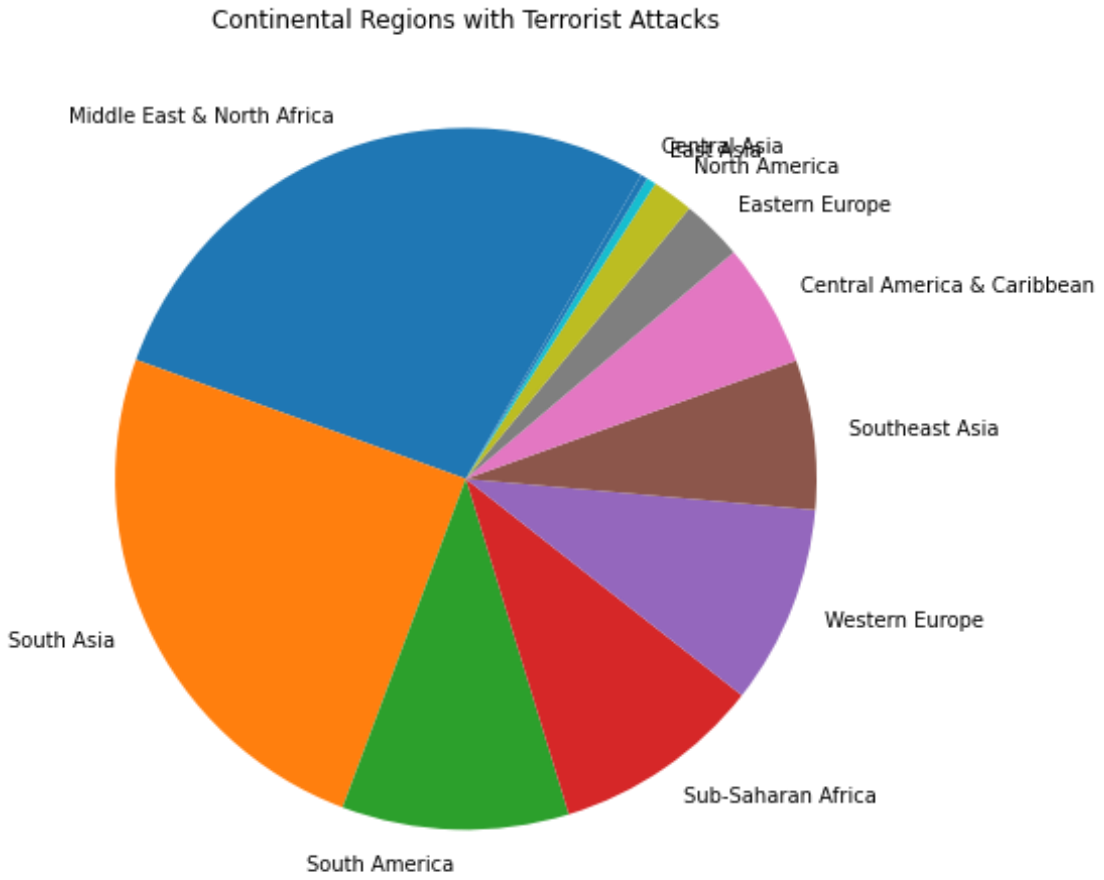
Who being the Targets?



Citizen & Property

Main Hot Zone Areas

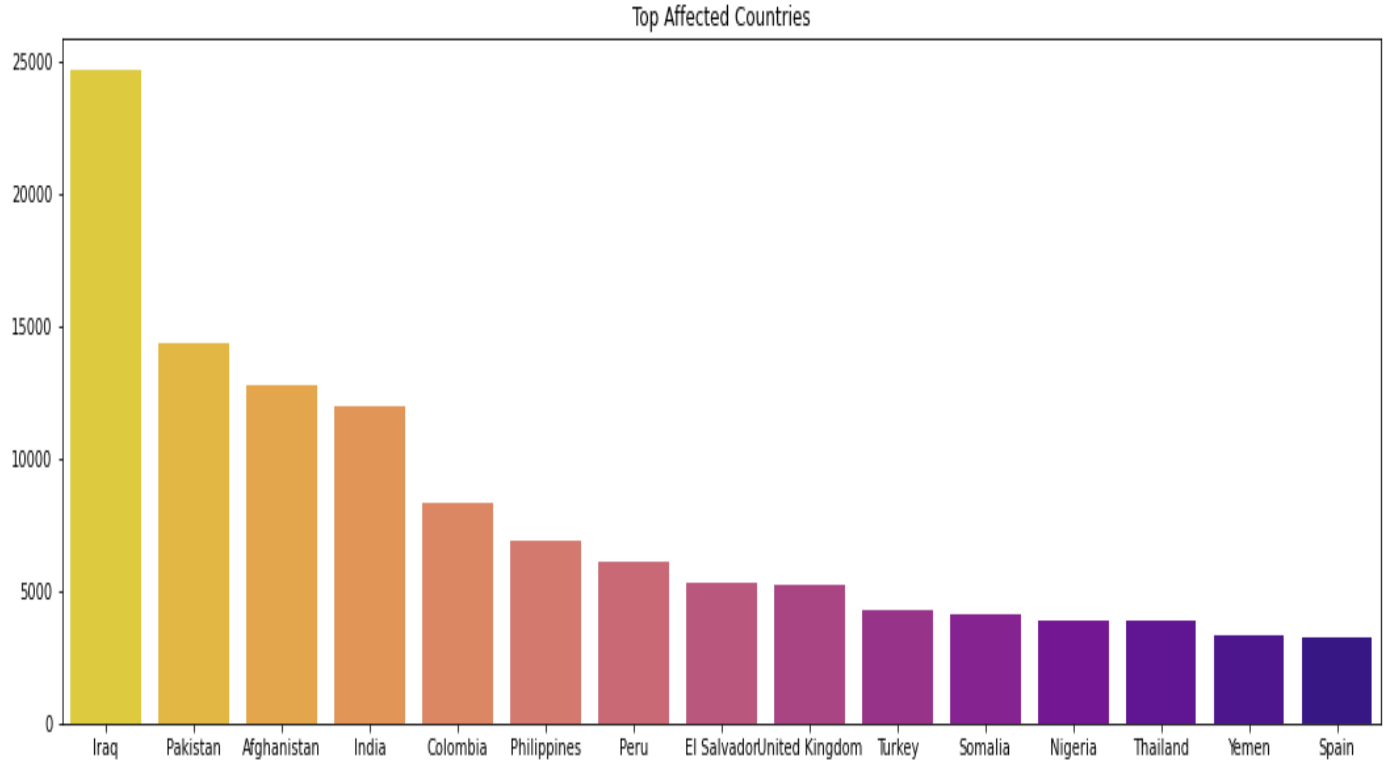
- The most active area **Middle East & North America**
- South America, Middle East and North Africa, Central America were the first regions for terrorist attacks



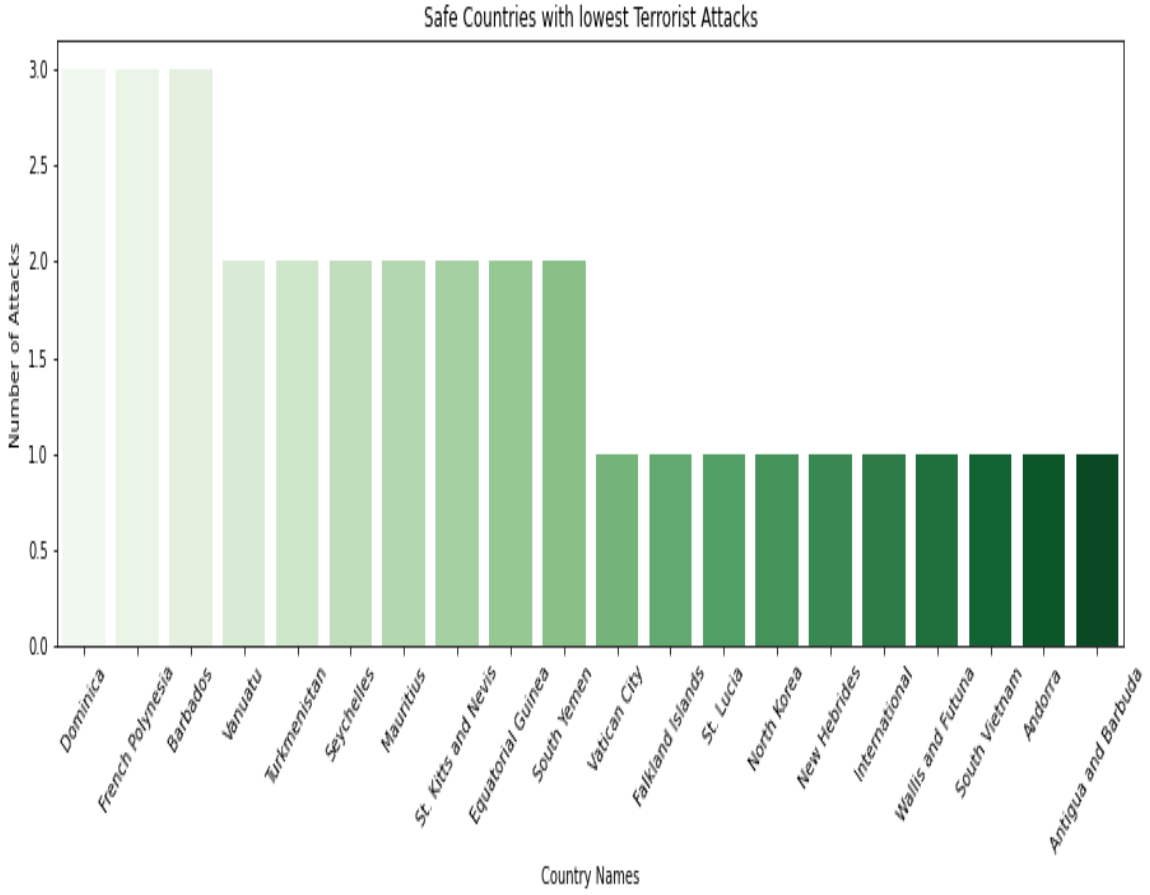
Terrorism By Country

- **Iraq** stand first in terms of terrorism and terrorist activities based on the data from 1970 to 2017

- India's rank falling from 8th in 2017 and 2018 to 7th in 2019 which reflecting the worsening of the overall security situation in the country on the ground of terrorism

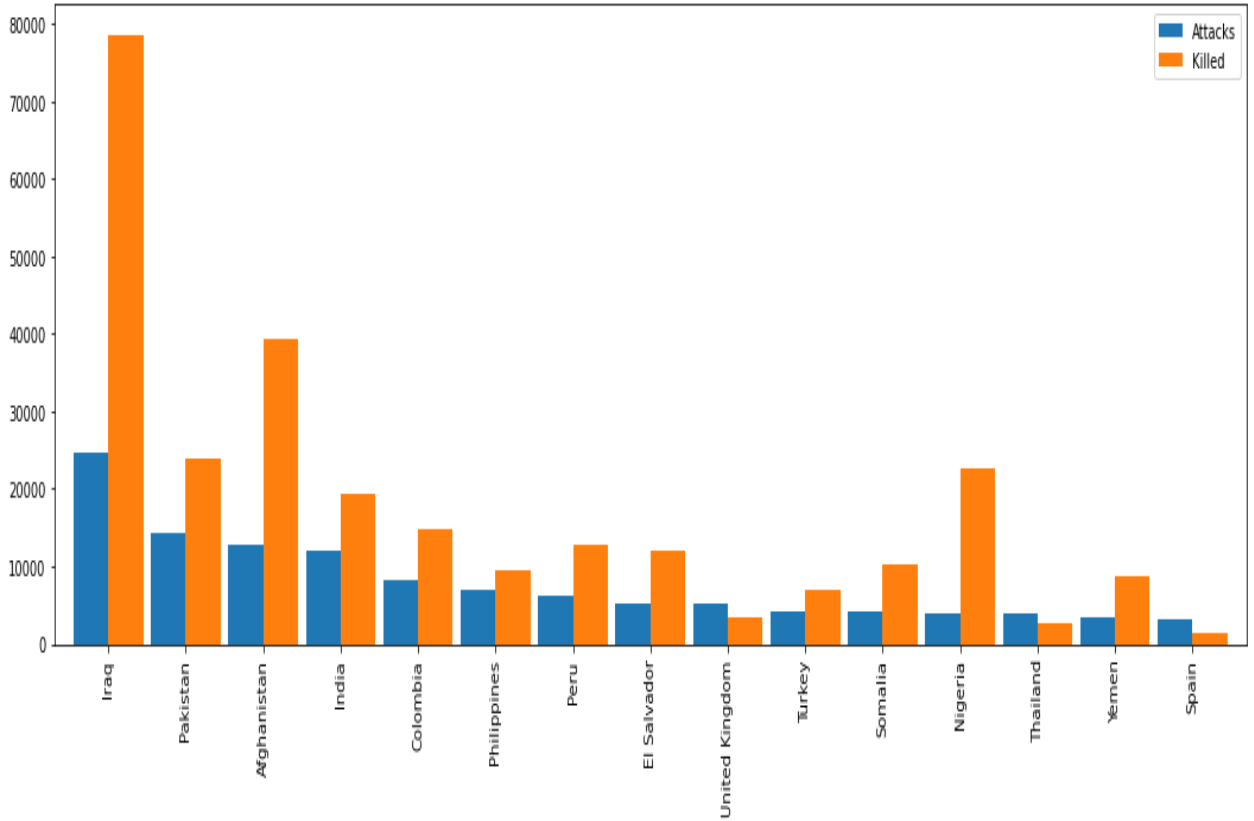


(Countries which are safe from terror attacks)



Dominica, French Polynesia, Barbados

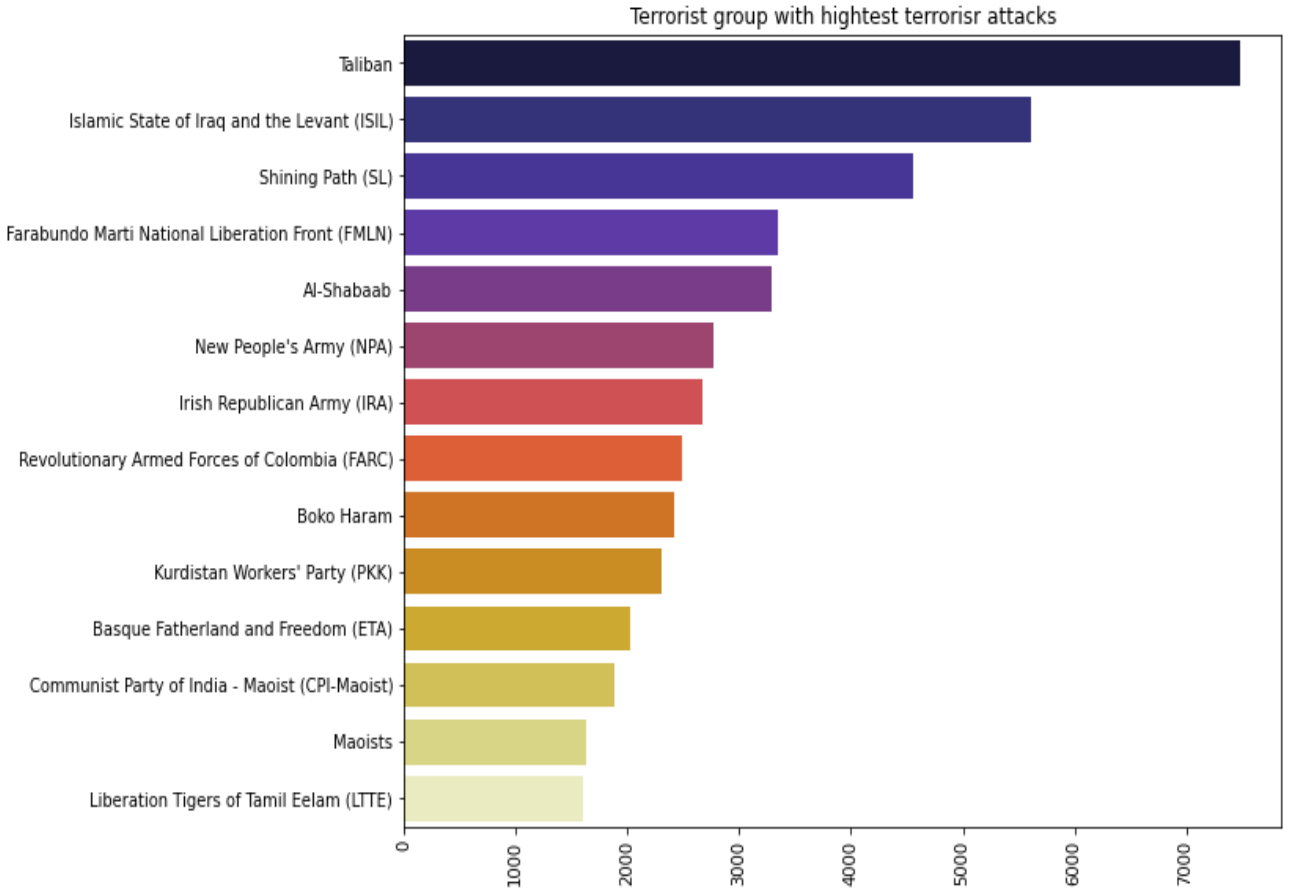
(No of Attacks Vs Killed)



Iraq- approx 80000 killed & 25000 attacks

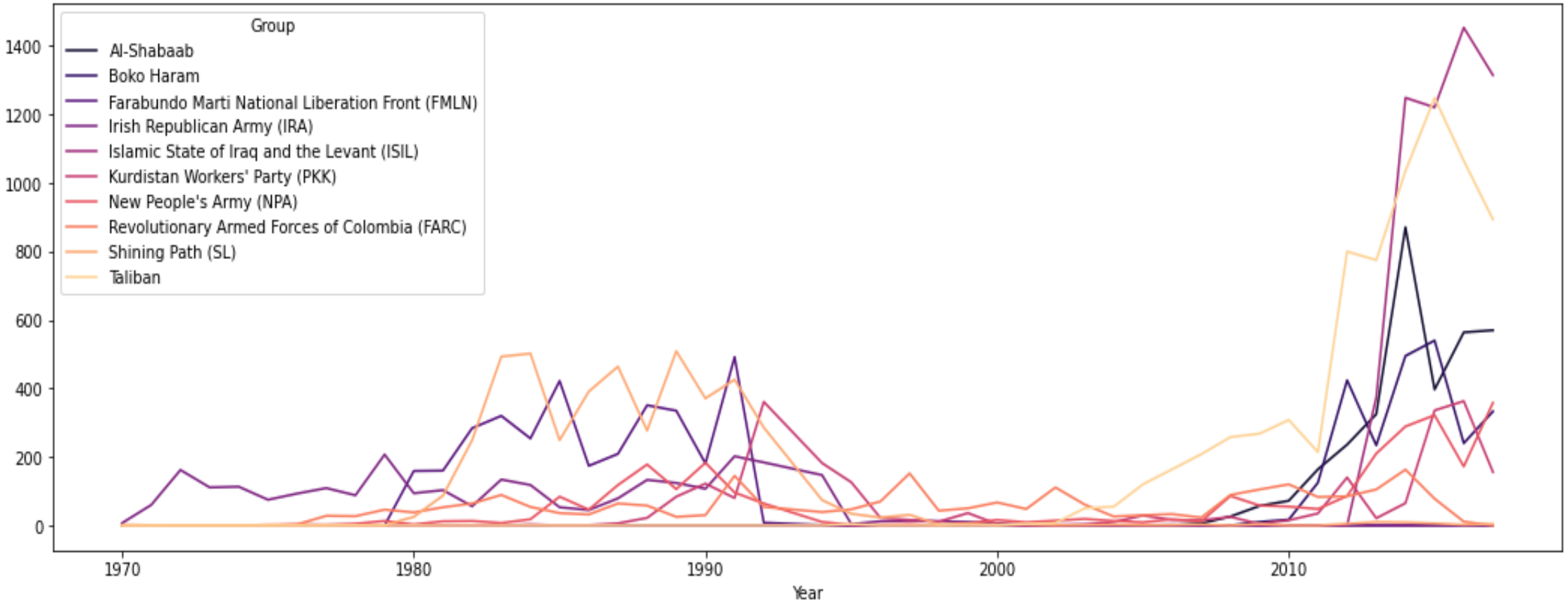
Most active Notorious Groups

Taliban is the most active group which involves in most of the terrorist attacks



(Top 10 Terrorist Group Activities)

After 2010 the terrorist activities were increasing and in 2014 maximum terrorist attacks took place by the top terrorist groups



Analysis on India

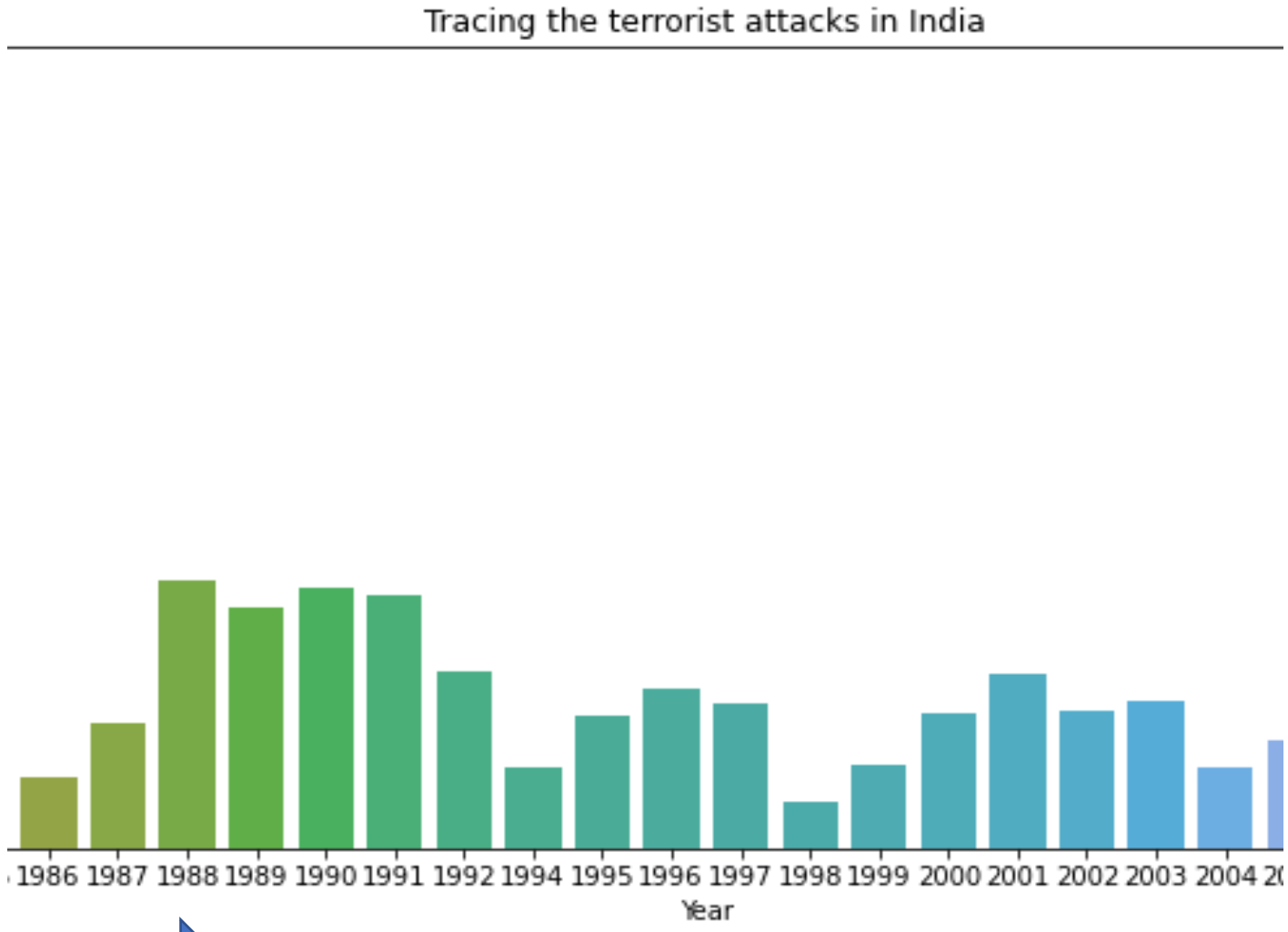
- India’s tryst with terrorism and violent extremism can be traced back partly to the religion based partition in 1947
- **The primary causes of terrorism and insurgency in India** are based on political, religious, ethnic, ideological, identity driven, linguistic or socio-economic grievances.

```
India=terror_data[terror_data.Country=='India']
India.head(5)
```

	Year	Month	Day	Country	Success	Region	City	Latitude	Longitude	AttackType	Killed	Wounded	Target	Group	Target_type	Weapon_type	Motive	casualties
1186	1972	2	22	India	1	South Asia	New Delhi	28.585836	77.153336	Hijacking	0.0	0.0	B-747	Palestinians	Airports & Aircraft	Explosives	NaN	0.0
2764	1975	1	2	India	1	South Asia	Samastipur	25.863042	85.781004	Bombing/Explosion	4.0	0.0	Lalit Narayan Mishra and a legislator	Ananda Marga	Government (General)	Explosives	Unknown	4.0
3857	1976	5	26	India	1	South Asia	New Delhi	28.585836	77.153336	Bombing/Explosion	0.0	0.0	New Delhi airport	Unknown	Airports & Aircraft	Explosives	NaN	0.0
5327	1977	9	28	India	1	South Asia	Bombay	19.075984	72.877656	Hijacking	0.0	0.0	DC-8	Japanese Red Army (JRA)	Airports & Aircraft	Firearms	NaN	0.0
7337	1979	1	13	India	1	South Asia	Unknown	26.200605	92.937574	Armed Assault	NaN	NaN	patrol	Naga People	Police	Firearms	NaN	NaN

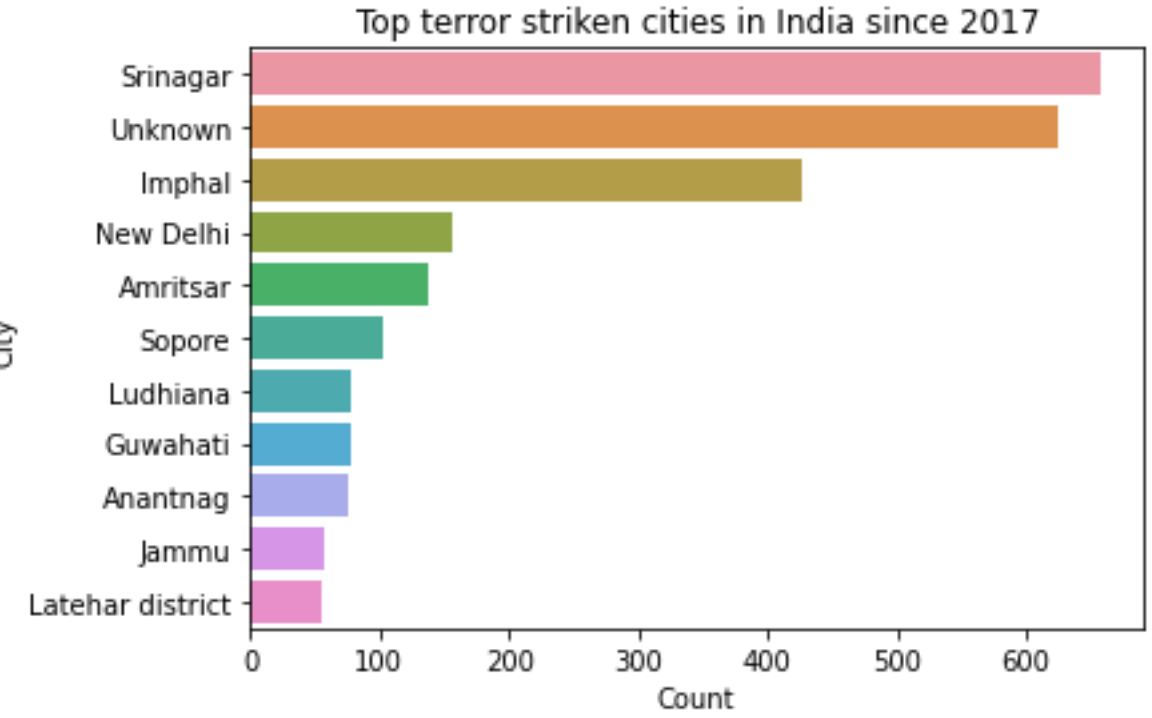
(Terrorist Attacks in India)

- There are 10280 successful terrorist strike has taken place in India while 1680 attempts failed from 1970-2017
- Since 2008 there has been an significant rise in the number of terror attempts in India



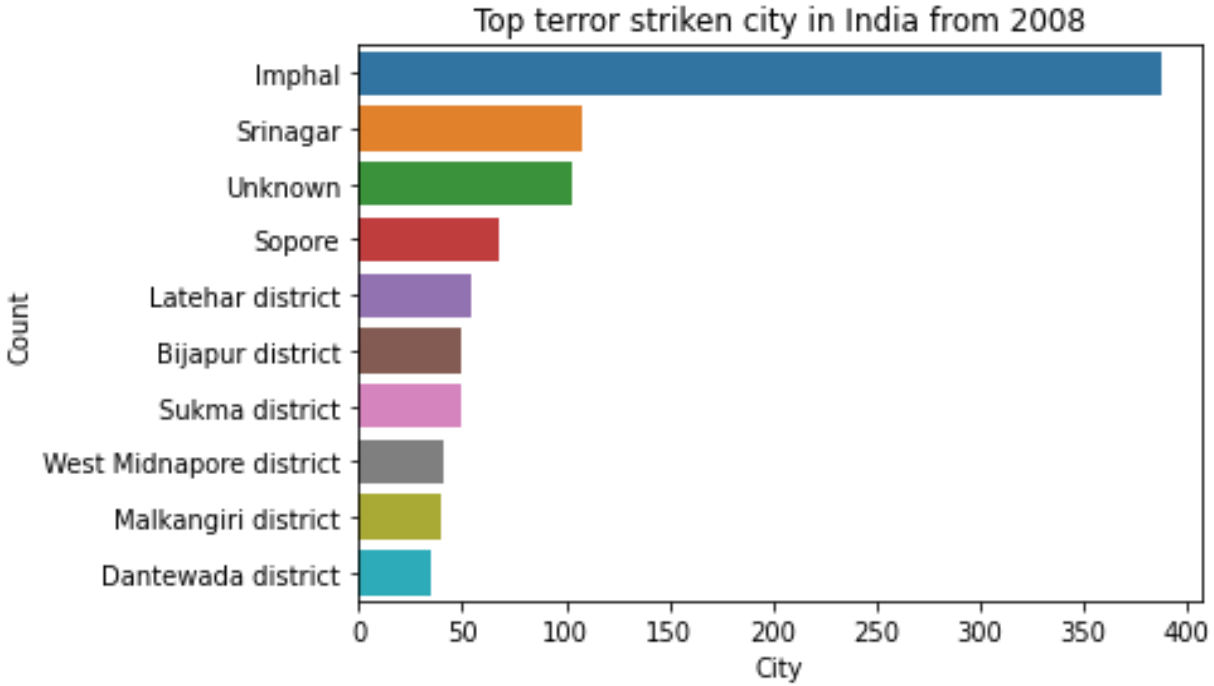
(Terror Strike cities in India since 2017)

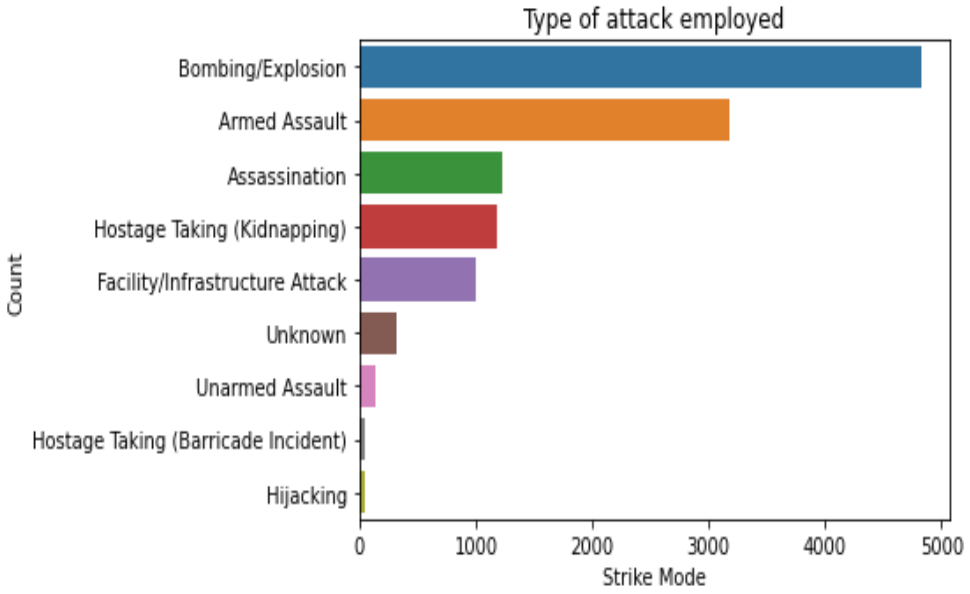
- Srinagar in North and Imphal in northeast are most affected

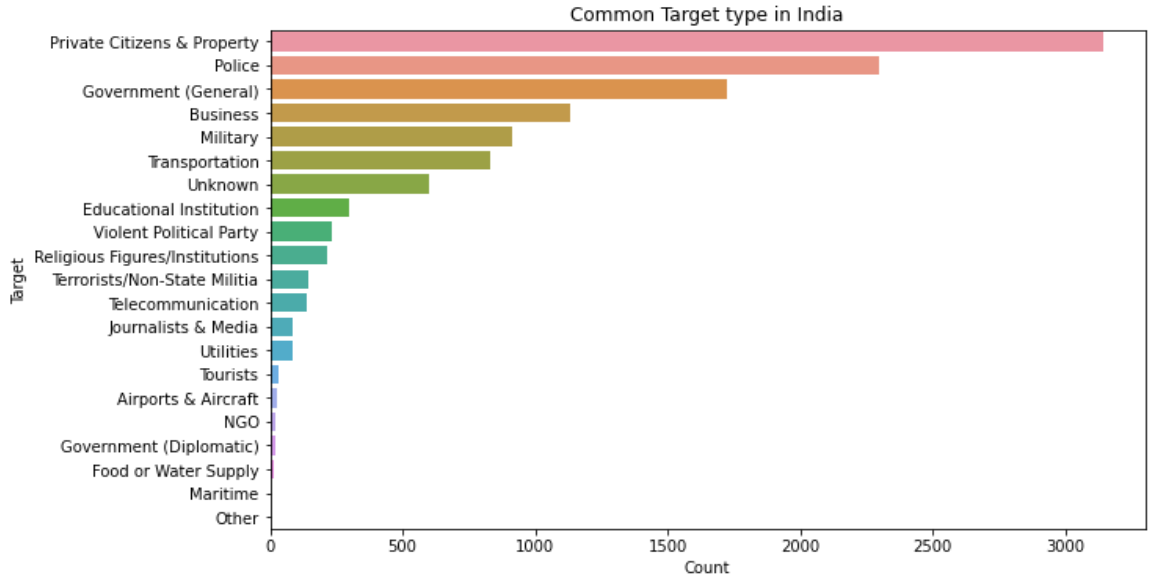


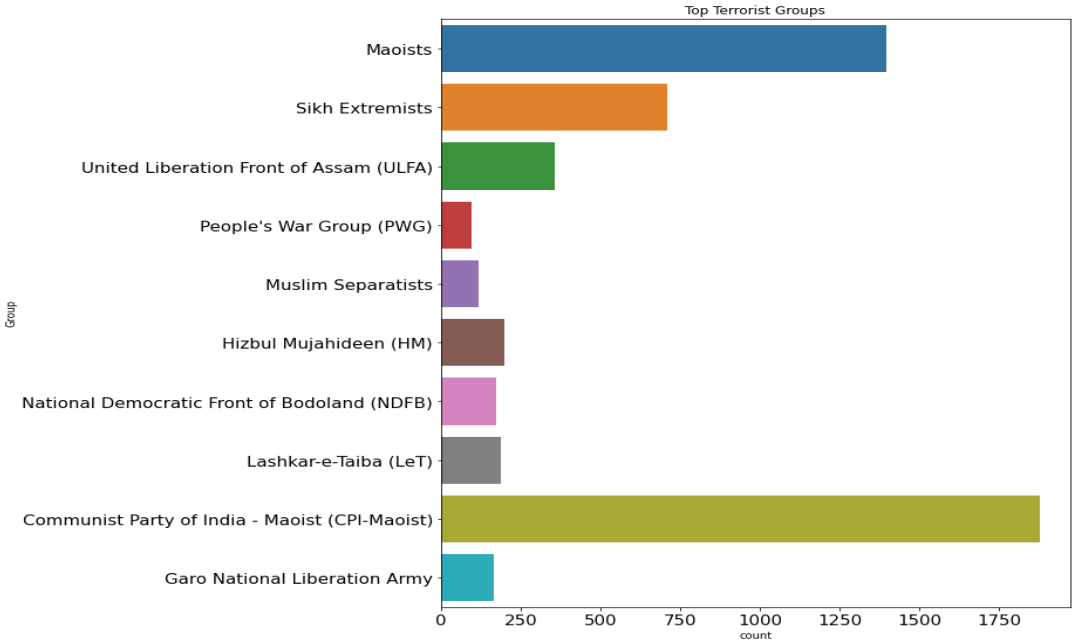
(Terror Strike cities in India from 2008)

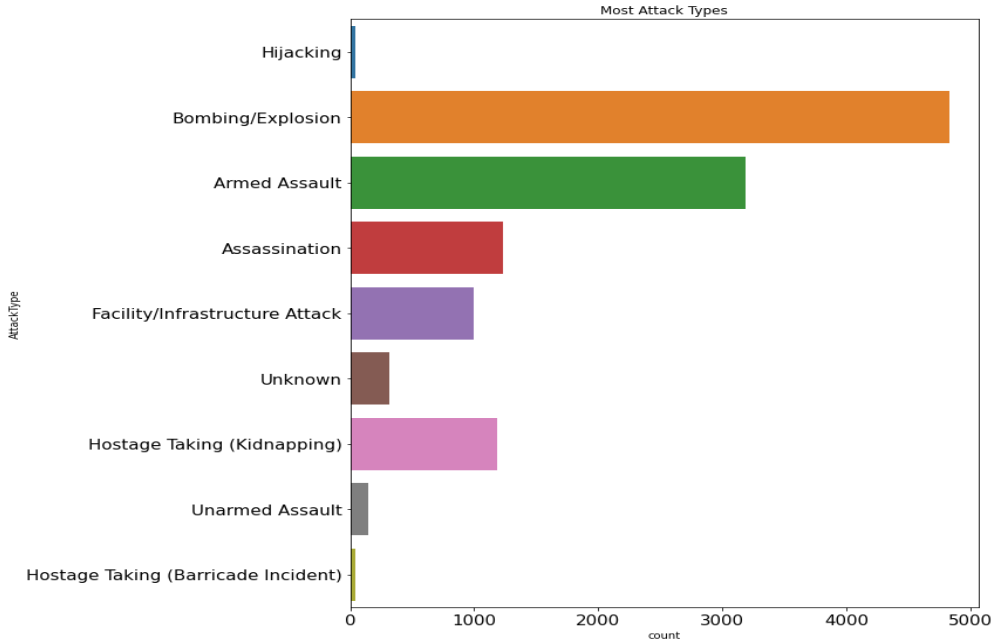
- Imphal in Northeast India has encountered lots of terror strikes in the last decades




Type of attack employed
Bombing/Explosion
Armed Assault
Assassination
Hostage Taking (Kidnapping)
Facility/Infrastructure Attack
Unknown
Unarmed Assault
Hostage Taking (Barricade Incident)
Hijacking
Count
0
1000
2000
3000
4000
5000
Strike Mode


Common Target type in India
Private Citizens & Property
Police
Government (General)
Business
Military
Transportation
Unknown
Educational Institution
Violent Political Party
Religious Figures/Institutions
Terrorists/Non-State Militia
Telecommunication
Journalists & Media
Utilities
Tourists
Airports & Aircraft
NGO
Government (Diplomatic)
Food or Water Supply
Maritime
Other
Target
0
500
1000
1500
2000
2500
3000
Count


Top Terrorist Groups
Maoists
Sikh Extremists
United Liberation Front of Assam (ULFA)
People's War Group (PWG)
Muslim Separatists
Hizbul Mujahideen (HM)
National Democratic Front of Bodoland (NDFB)
Lashkar-e-Taiba (LeT)
Communist Party of India - Maoist (CPI-Maoist)
Garo National Liberation Army
Group
0
250
500
750
1000
1250
1500
1750
count


Most Attack Types
Hijacking
Bombing/Explosion
Armed Assault
Assassination
Facility/Infrastructure Attack
Unknown
Hostage Taking (Kidnapping)
Unarmed Assault
Hostage Taking (Barricade Incident)
AttackType
0
1000
2000
3000
4000
5000
count

Conclusion

- The world security level is degrading into much lower levels in the recent times.
- The main method in terrorist attacks is Explosives and arms.
- Mostly we cannot find a terror attack free country.
- Prime target for terrorist attacks are government and military.
- Terrorist groups are choosing sleeper cells as main weapon for attacking.
- Mostly unknown groups are killing most of the people.
- Mostly any terror attack, are having a most successfull success rate.

Q & A