Capstone Project - 1

Exploratory Data Analysis on Global Terrorism Database

Team Members:
Aehteshaam Shaikh
Aniket Borde

AI

Problem Statement :

The Global Terrorism Database (GTD) is a database of terrorist incidents from 1970 to 2017 maintained by START.
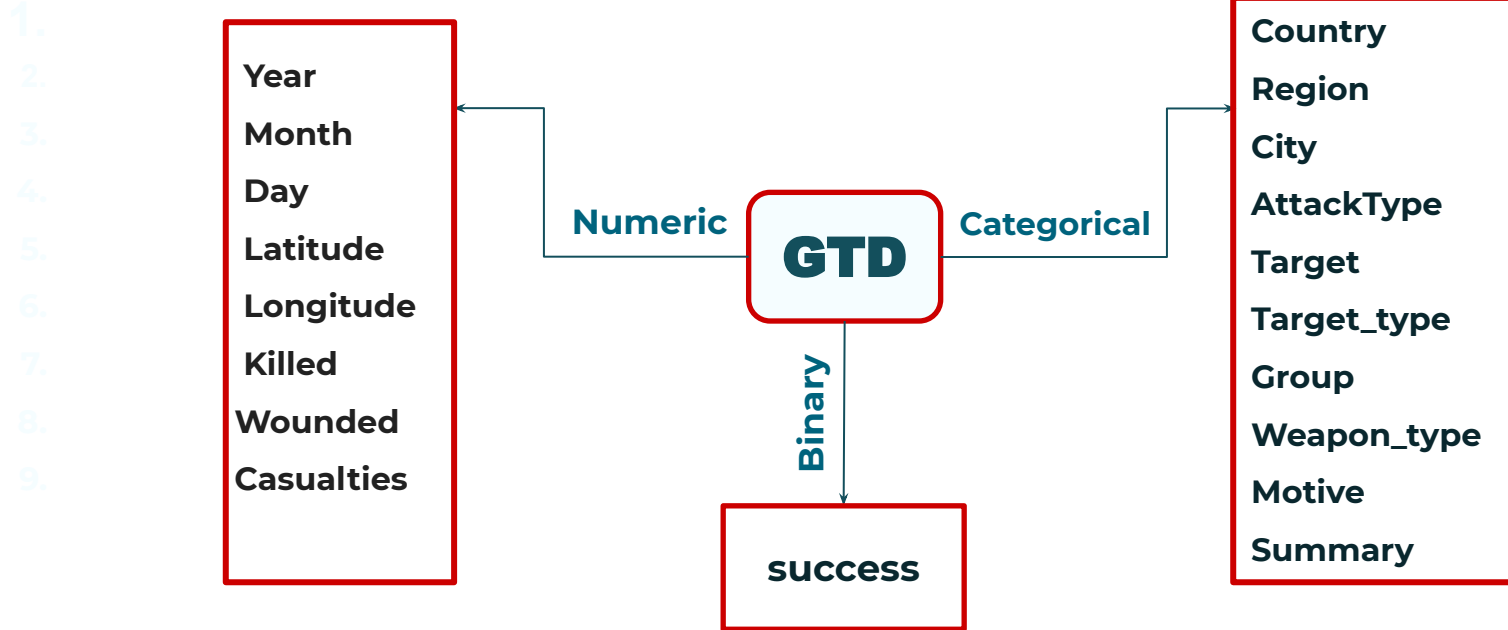
The dataset has about 1,80,000+ rows each containing information for each Attack. Our objective is to use this data, explore it, and generate insights from it.

Data Cleaning:

The data set consist of around 1,81,691 rows and 135 columns most of which have null values more than 75%. Hence, those columns have been dropped and the Data frame has been reduced to 19 columns.

Again, Null Values from these 19 columns have been replaced with suitable values.

Data Summary:

Numeric

Year
Month
Day
Latitude
Longitude
Killed
Wounded
Casualties

GTD

Categorical

Country
Region
City
AttackType
Target
Target_type
Group
Weapon_type
Motive
Summary

Binary

success

Data Summary:

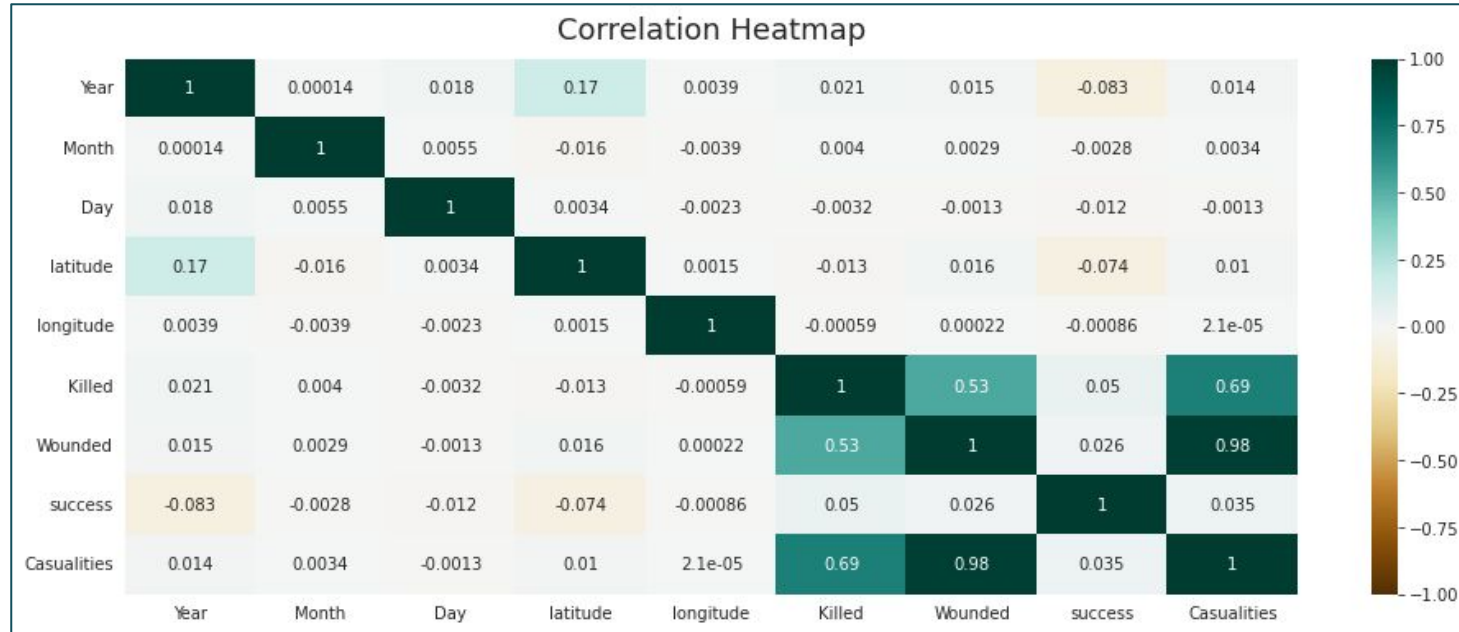| Sr	Column Name	Description
1	**eventid**	A 12-digit Event ID system. First 8 numbers – date recorded "yyyymmdd". Last 4 numbers – sequential case number for the given day (0001, 0002 etc).
2	**Year**	This field contains the year in which the incident occurred.
3	**Month**	This field contains the number of the month in which the incident occurred
4	**Day**	This field contains the numeric day of the month on which the incident occurred
5	**Country**	This field identifies the country or location where the incident occurred
6	**Region**	This field identifies the region in which the incident occurred
7	**city**	Name of the city, village, or town in which the incident occurred

Data Summary:

| Sr	Column Name	Description
8	Latitude	The latitude of the city in which the event occurred
9	Longitude	The longitude of the city in which the event occurred
10	success	Whether the attack was successful or not
11	AttackType	The general method of attack and broad class of tactics used.
12	Target_type	The general type of target/victim
13	Group	The name of the group that carried out the attack
14	Weapon_type	General type of weapon used in the incident

Data Summary:

| Sr	Column Name	Description
15	**Killed**	The number of total confirmed fatalities for the incident
16	**Wounded**	The number of people injured in the incident
17	**Motive**	Motive of the attack
18	**Summary**	Summary of the attack

AI

Visualizing the Data

Correlation plot : This plot is used to investigate the dependence between multiple variables at the same time and to highlight the most correlated variables in a data table. In this visual, correlation coefficients are colored according to the value.

Correlation Heatmap

|	Year	Month	Day	latitude	longitude	Killed	Wounded	success	Casualities
Year	1	0.00014	0.018	0.17	0.0039	0.021	0.015	-0.083	0.014
Month	0.00014	1	0.0055	-0.016	-0.0039	0.004	0.0029	-0.0028	0.0034
Day	0.018	0.0055	1	0.0034	-0.0023	-0.0032	-0.0013	-0.012	-0.0013
latitude	0.17	-0.016	0.0034	1	0.0015	-0.013	0.016	-0.074	0.01
longitude	0.0039	-0.0039	-0.0023	0.0015	1	-0.00059	0.00022	-0.00086	2.1e-05
Killed	0.021	0.004	-0.0032	-0.013	-0.00059	1	0.53	0.05	0.69
Wounded	0.015	0.0029	-0.0013	0.016	0.00022	0.53	1	0.026	0.98
success	-0.083	-0.0028	-0.012	-0.074	-0.00086	0.05	0.026	1	0.035
Casualities	0.014	0.0034	-0.0013	0.01	2.1e-05	0.69	0.98	0.035	1

The Success Rate of the terrorist attacks is 89%, which is a big issue of concern for the safety of citizens and organisations who are responsible for the safety of everyone.

Success Rate of Terrorist Attacks
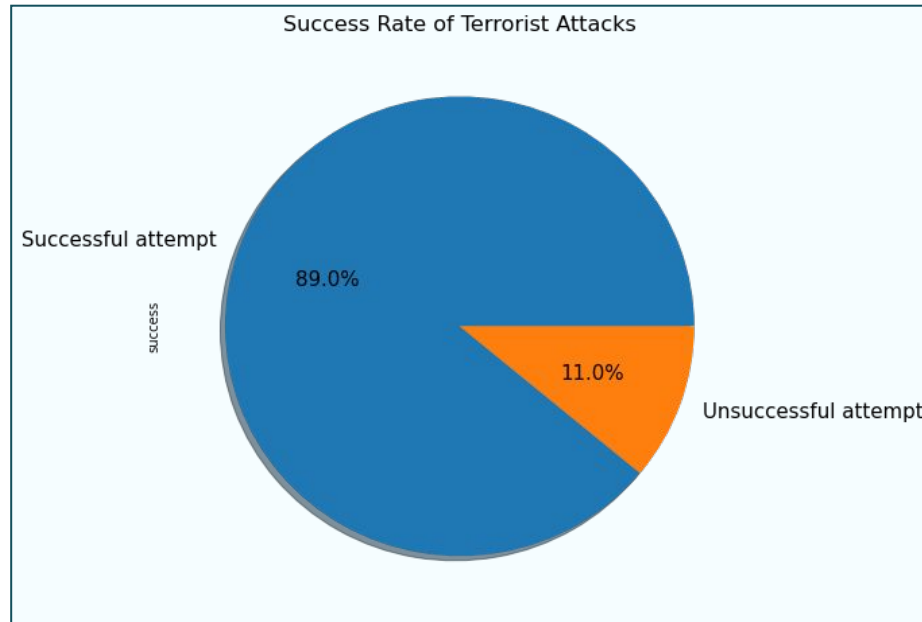
Middle East and North Africa are the most terrorism prone regions followed by South Asia. The Australia and Oceania region has experienced fewer terrorist attacks.
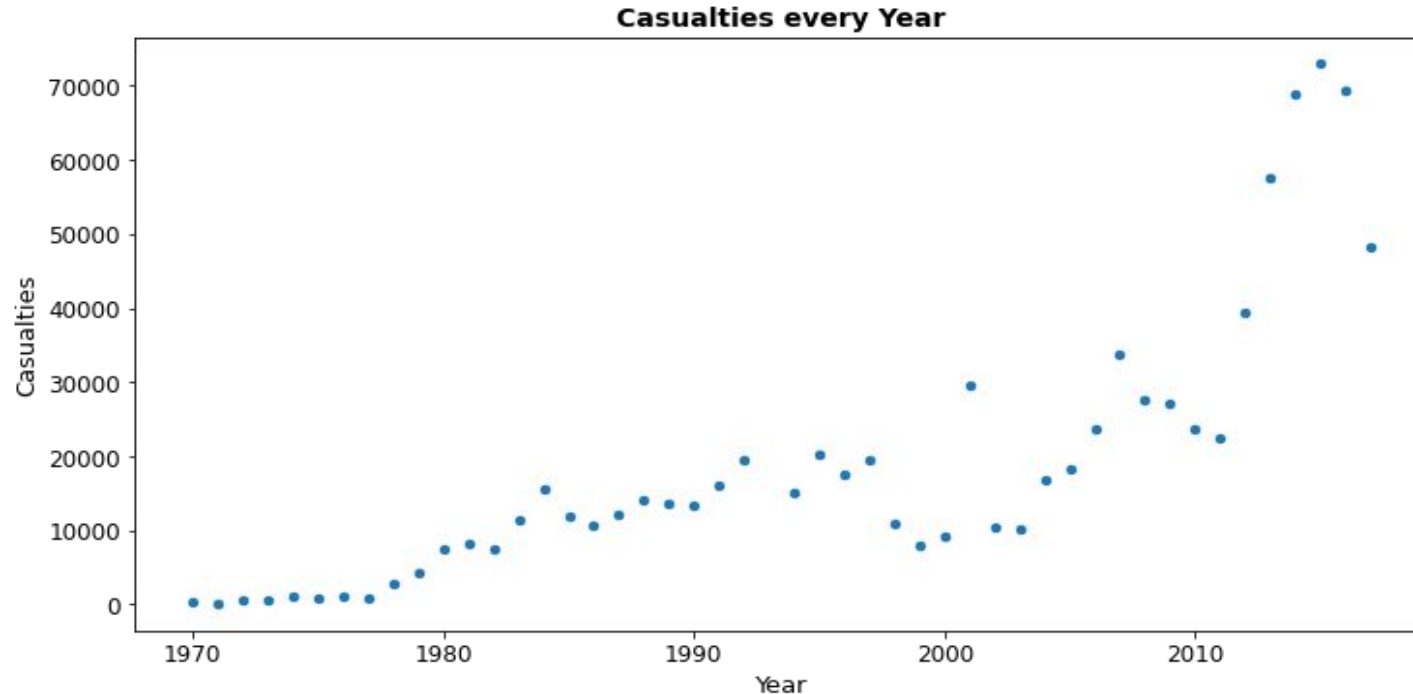

Number Of Terrorist Activities By Region
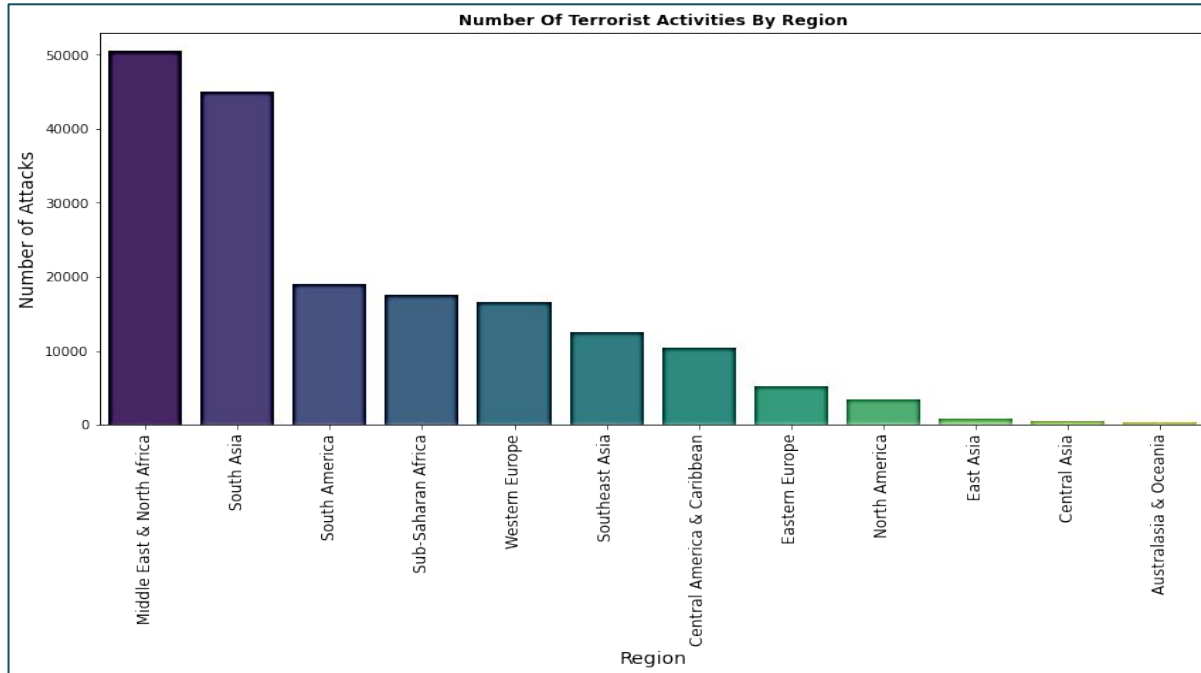
From the list of Top 10 Terror prone countries , the most number of attacks took place till now are in Iraq (~ 25,000) which is huge in (~ 50 years) i.e. on average ~ 500 attacks per year.
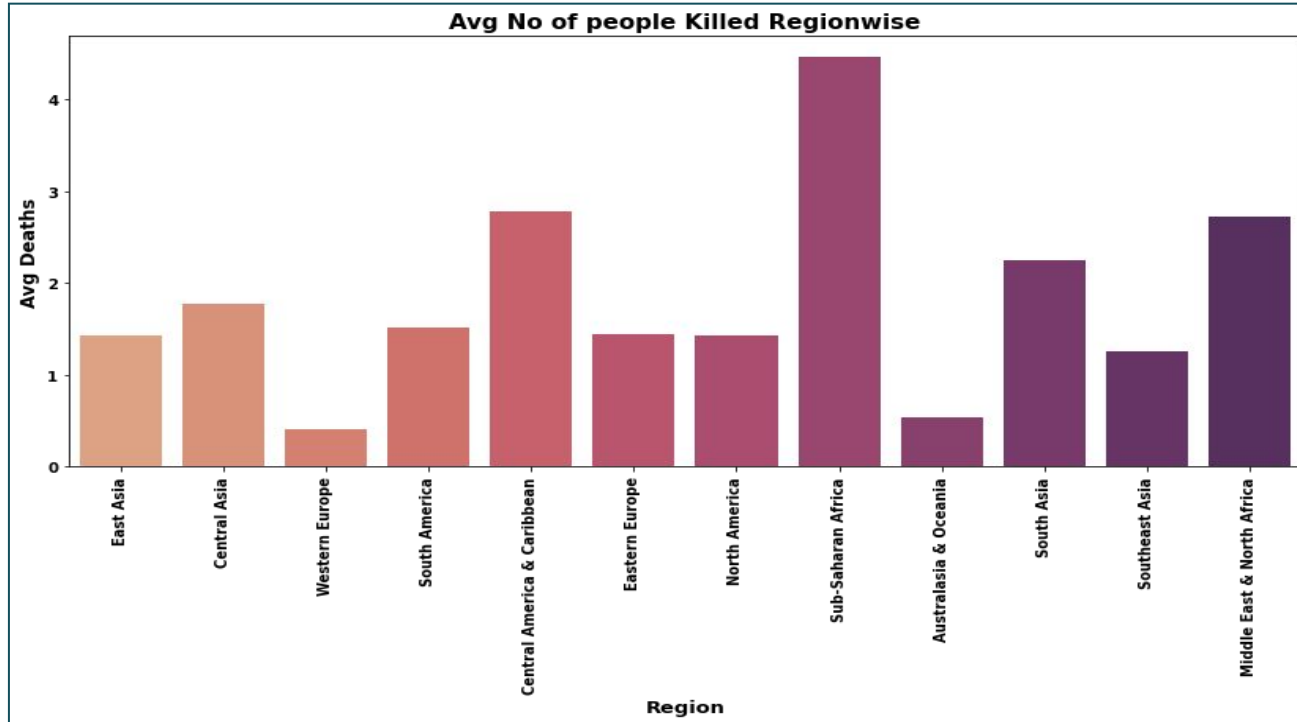

Number Of Terrorist Activities By Country
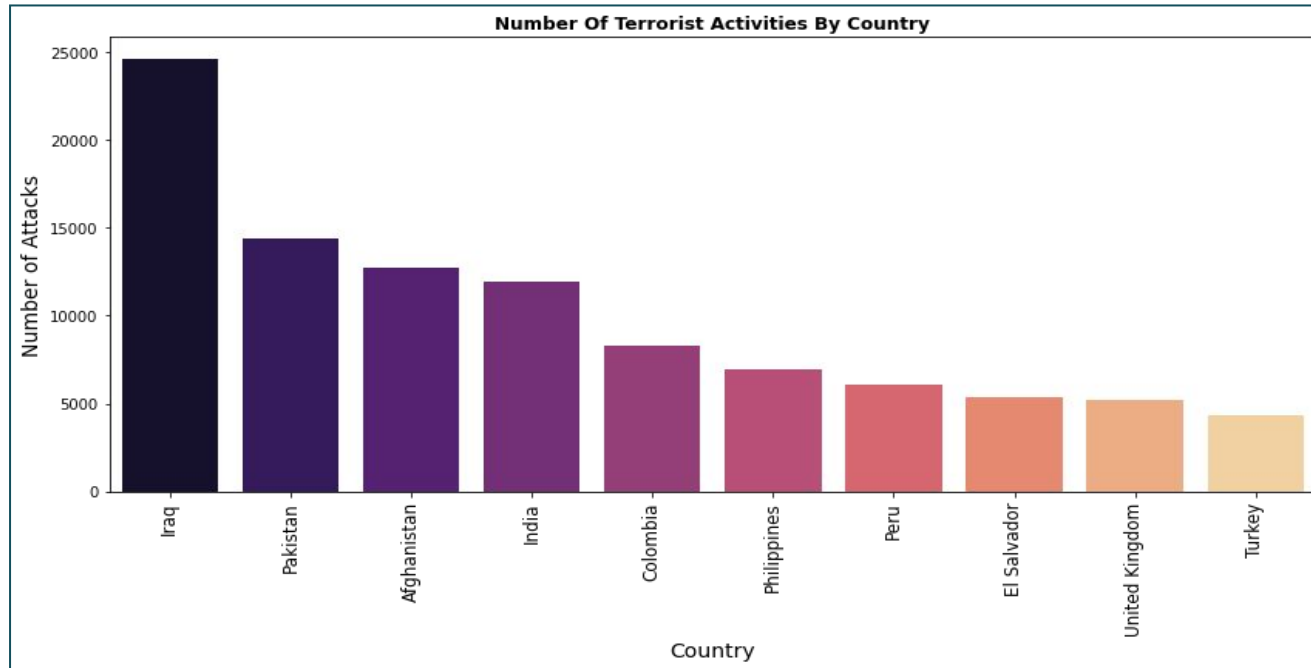
The People affected by these Terrorist attacks have been increased exponentially from the year 2000 to 2015.

Maximum Casualties were happened in the Year 2015.
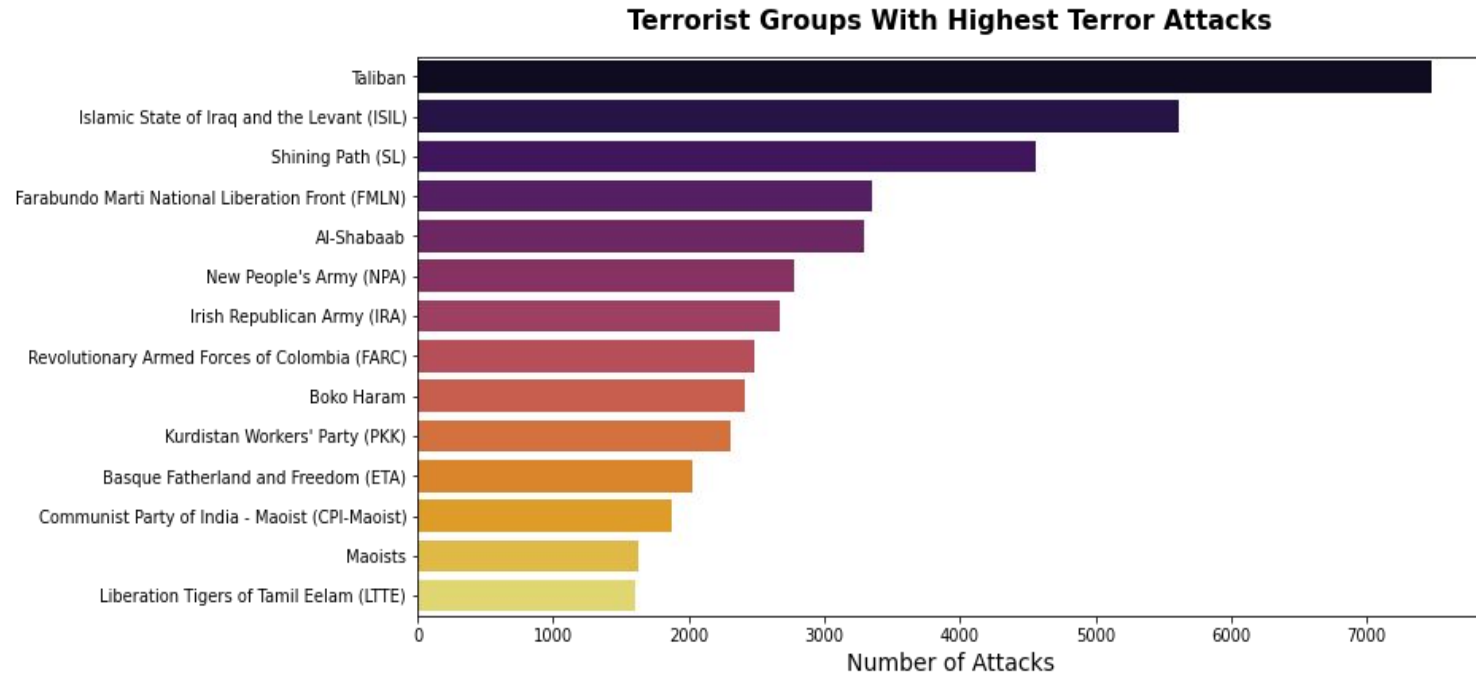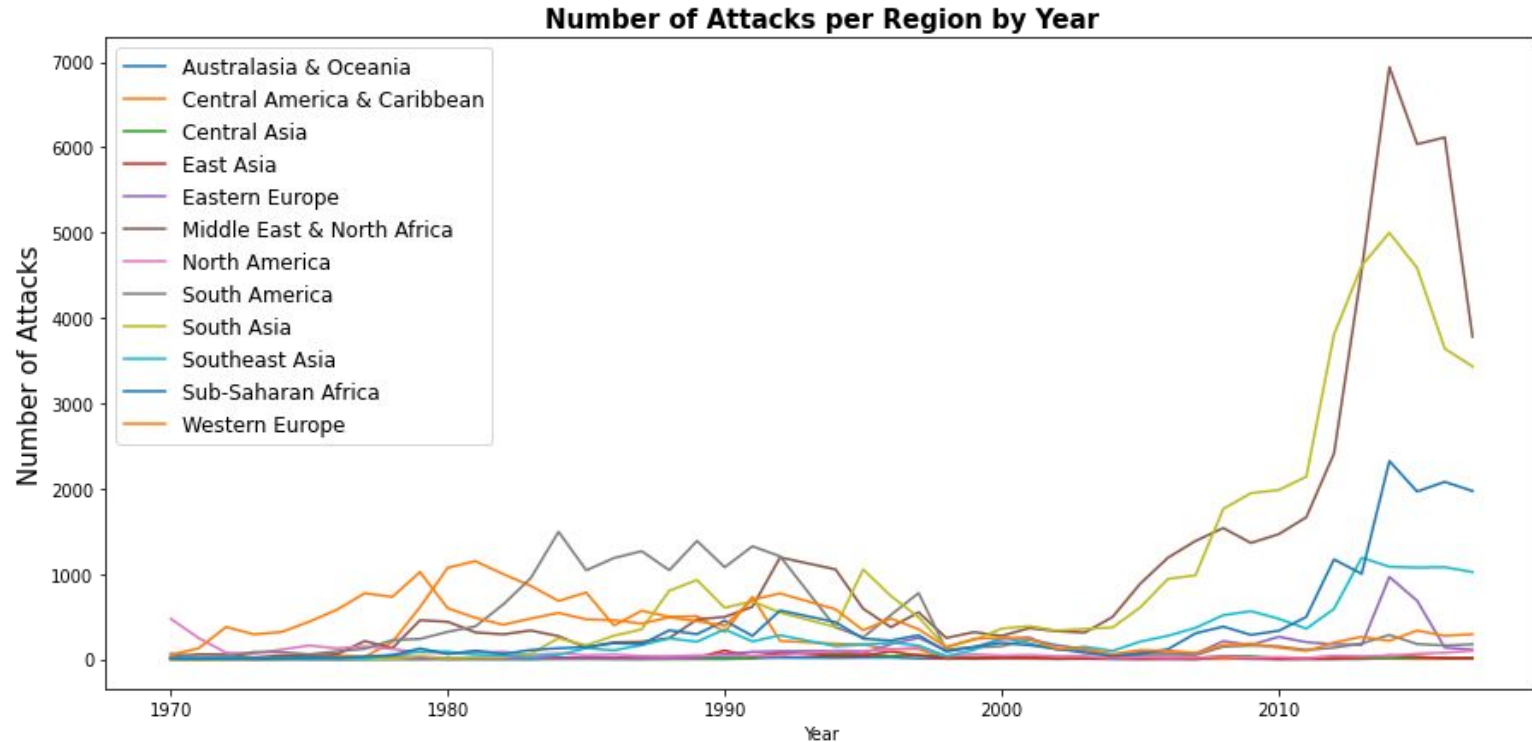
Casualties every Year

Every terror attack in Sub-Saharan Africa claims over 5 lives on average.

Avg No of people Killed Regionwise

The terrorist Groups that have led to most number of terrorist activities are: Taliban, Islamic State of Iraq and The Levant and Shining Path.
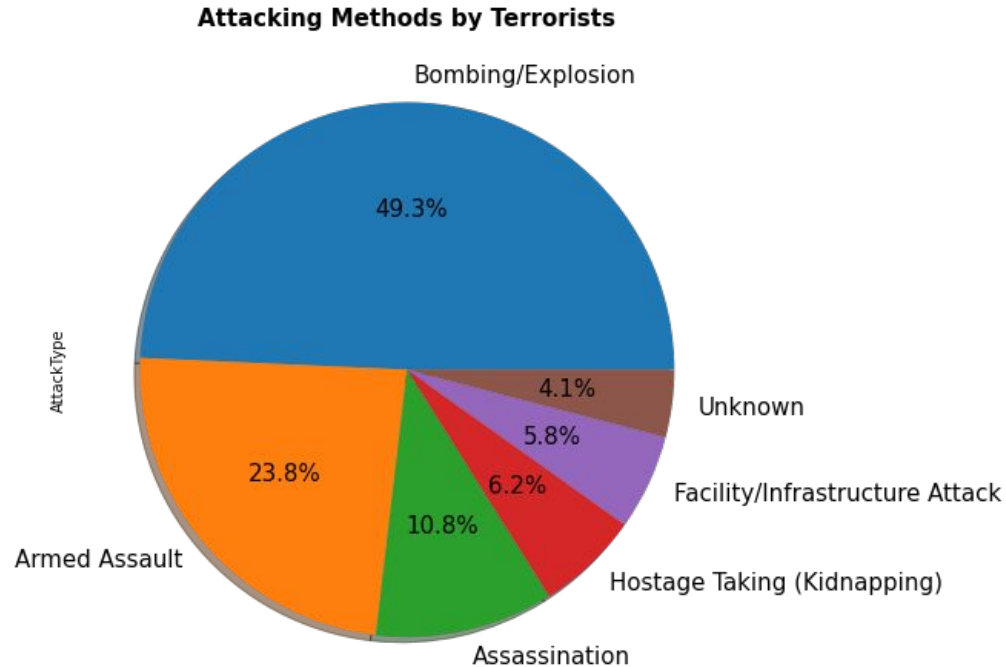

Terrorist Groups With Highest Terror Attacks

Group	Number of Attacks (approx.)
Taliban	7500
Islamic State of Iraq and the Levant (ISIL)	5600
Shining Path (SL)	4500
Farabundo Marti National Liberation Front (FMLN)	3300
Al-Shabaab	3300
New People's Army (NPA)	2700
Irish Republican Army (IRA)	2600
Revolutionary Armed Forces of Colombia (FARC)	2400
Boko Haram	2400
Kurdistan Workers' Party (PKK)	2300
Basque Fatherland and Freedom (ETA)	2000
Communist Party of India - Maoist (CPI-Maoist)	1900
Maoists	1600
Liberation Tigers of Tamil Eelam (LTTE)	1600

This graph confirms that Attacks increased exponentially after 2002 mainly in Middle East & North Africa and South Asia

**Number of Attacks per Region by Year**

Legend:
- Australasia & Oceania
- Central America & Caribbean
- Central Asia
- East Asia
- Eastern Europe
- Middle East & North Africa
- North America
- South America
- South Asia
- Southeast Asia
- Sub-Saharan Africa
- Western Europe

X-axis: Year
Y-axis: Number of Attacks

Scatter Map box is used to determine the Attack type in the specified city with number of kills

AttackType=Assassination
AttackType=Hostage Taking (Kidnapping)
AttackType=Bombing/Explosion
AttackType=Facility/Infrastructure Attack
AttackType=Armed Assault
AttackType=Hijacking
AttackType=Unknown
AttackType=Unarmed Assault
AttackType=Hostage Taking (Barricade Incident)

© CARTO

Most prominent attack mode of terrorist is Bomb/Explosion and Armed Assault

Attacking Methods by Terrorists

- Bombing/Explosion — 49.3%
- Armed Assault — 23.8%
- Assassination — 10.8%
- Hostage Taking (Kidnapping) — 6.2%
- Facility/Infrastructure Attack — 5.8%
- Unknown — 4.1%

AttackType

It is evident that Private Citizens & Property, Military and Police are most affected by the attacks

**Favorite Targets**
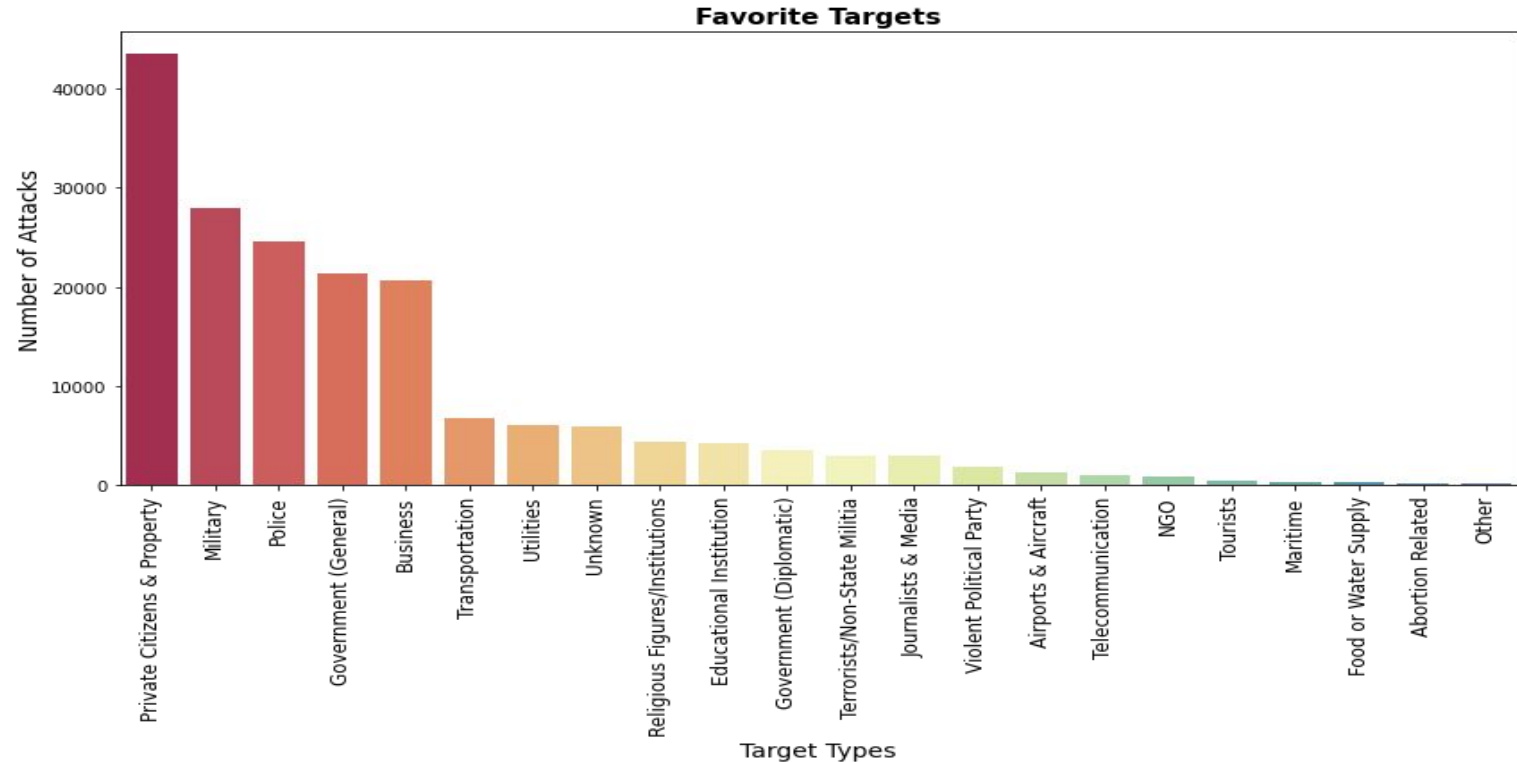
Analysis on India:

- **Maximum Casualties happened in an attack were : 1005
  which took place in Mumbai in the year 2006 by the Group Lashkar-e-Taiba (LeT)**

- **Maximum people killed in an attack were : 188.0 which took place in Mumbai in
  the year 2006**

- **Top 4 Indian Cities with maximum
  number of people killed in Terrorist
  attacks are shown in the table**

| Sr	City	Killed
1	Srinagar	540
2	Mumbai	460
3	New Delhi	309
4	Amritsar	284

- **After Year 2007 the no of attacks are increased much in India**
- **Year 2016 has the highest number of attack in India with total number of 1025 attacks in a year.**

Number of Attacks Each Year in India
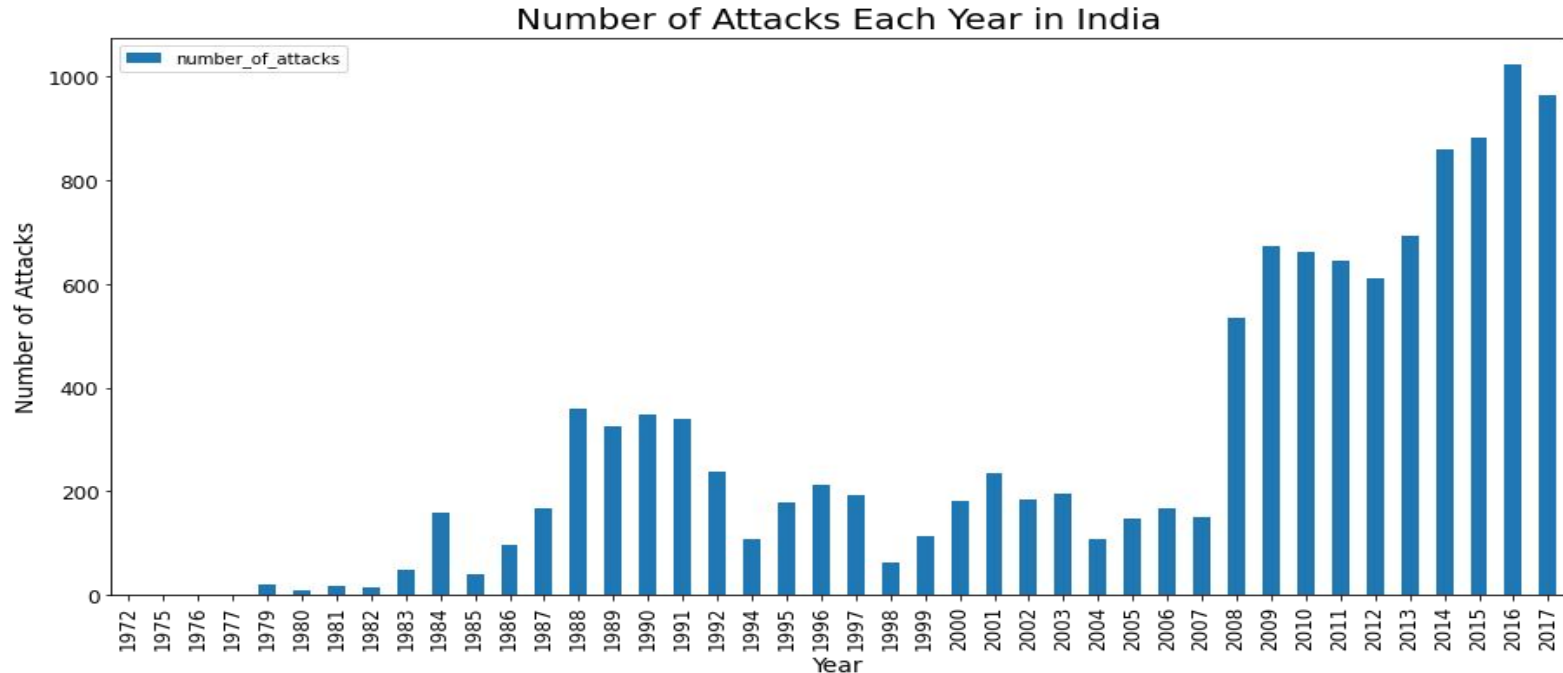
Most Number of People were Killed In Year 1992

Killed every Year
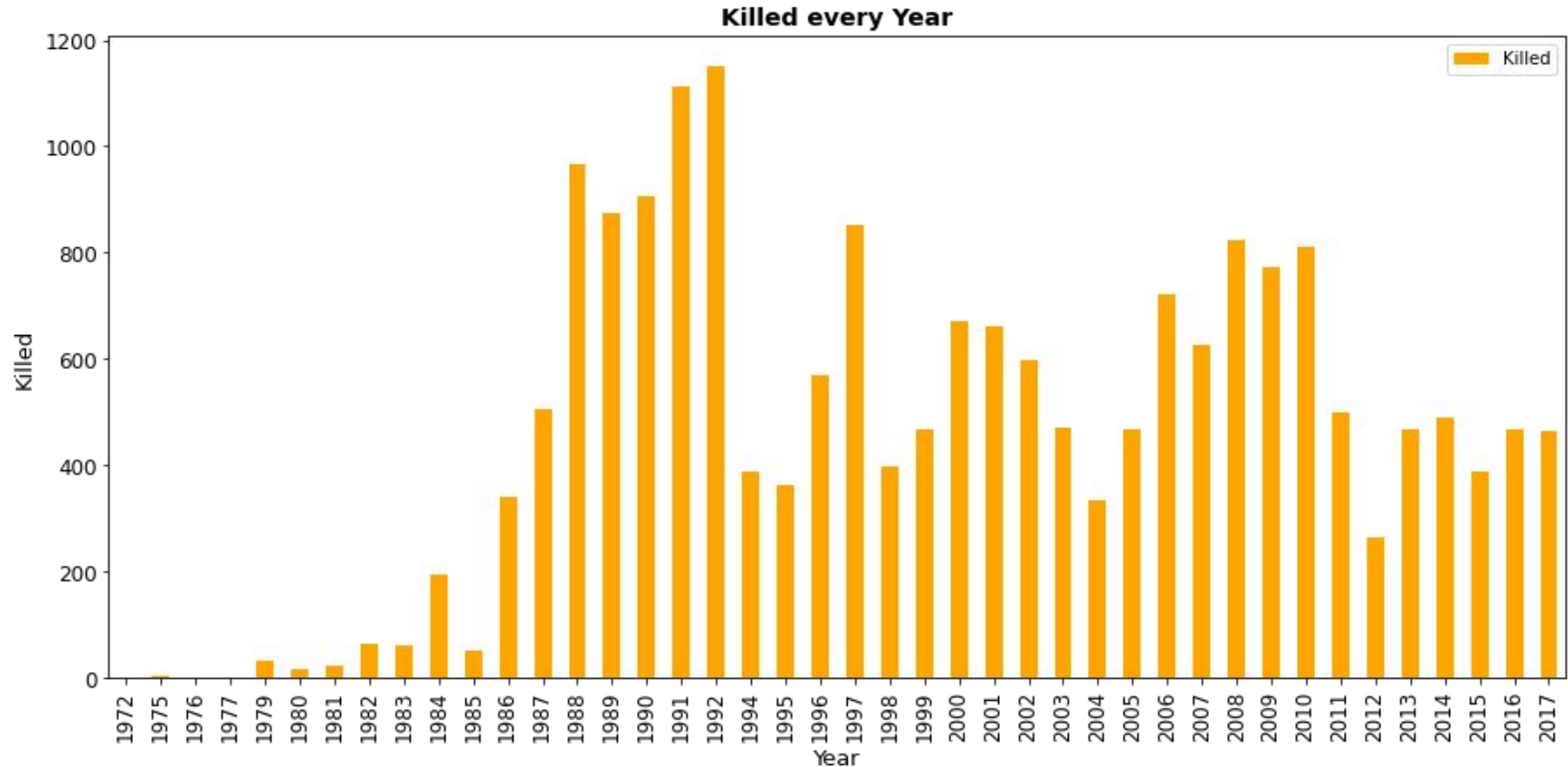
Most No of Kills by various Terrorist Groups in India

Terrorist Groups with Highest No of Kills
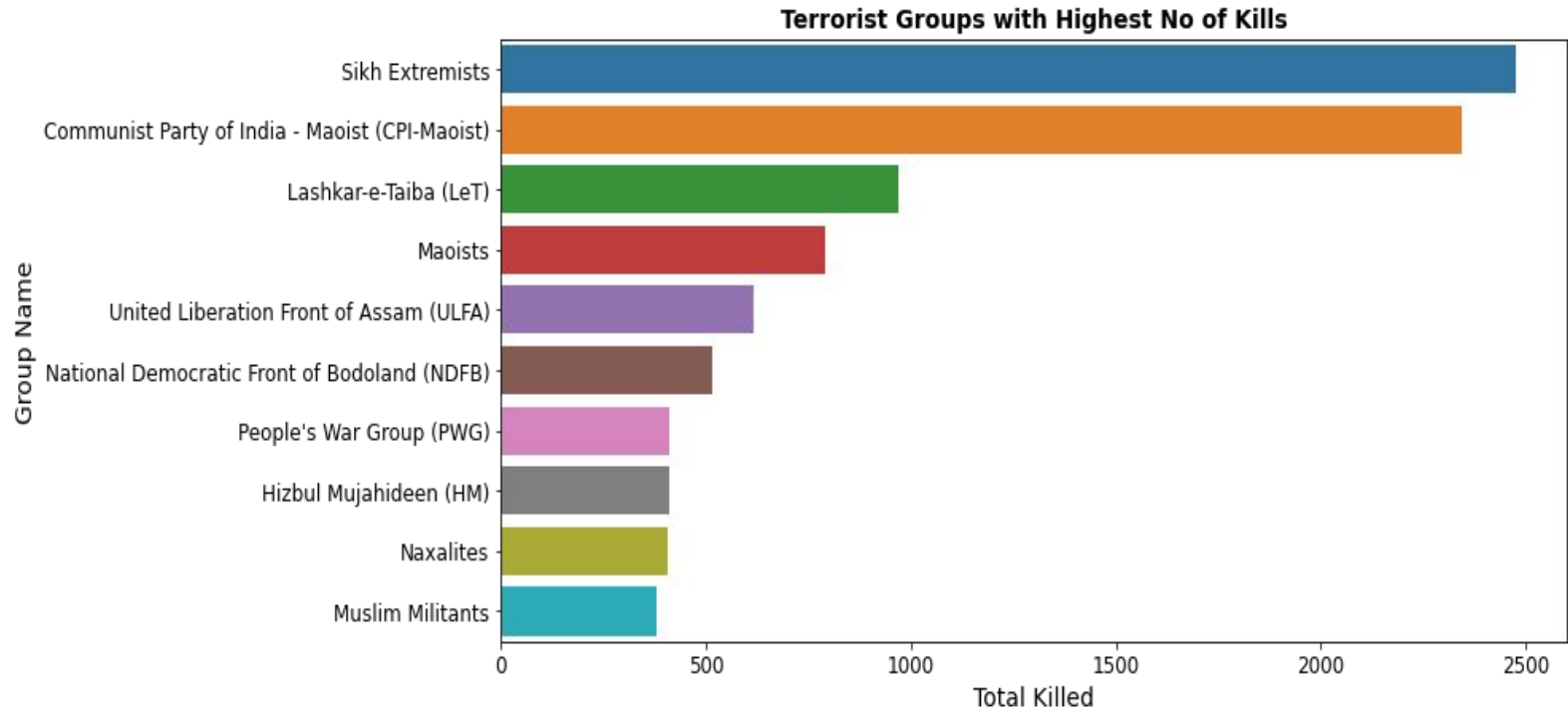
Conclusion:

- For some unknown reason the dataset doesn't provide any data for the Year 1993, also from this we can't conclude that no attack has happened in this Year.

- Since the data shows the Terrorist activities and Casualties are decreasing in late years, we can optimistically say that the world can witness more and more reduction in these activities in near future.