**Global Terrorism Analysis**

**Sandeep R**

**Data science trainee**

**AlmaBetter, Bangalore**

**Abstract:**

Global terrorism analysis is a dataset which contains the information of terrorist attacks around the world from 1970 to 2017.The Global terrorist Database includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks. The database is maintained by researchers at the National Consortium for the Study of Terrorism and Responses to Terrorism (START), headquartered at the University of

Maryland

**1.Problem Statement**

Main Aim of the project is to analyze the Dataset and extract the useful information, and to Discover the Key findings with a graphical representation of the Data provided.

Global Terrorism Database is a huge Dataset with More than 1.8 Lakhs Examples and 135 Variables. Where each variables provides different type of information of dataset. Some variables provides the timeframe of the incident such as time, month, date of the attack while other variables provides the severity of the attack such as The total kills, number of people wounded, whether the attack was successful or not, and also we can find the variables which provides the exact location of the incident which is indicated by the latitude and longitude of the place along with the Region of the location. Similiarly there are a lot of variables which provides great such information which we can make use of to Discover the key findings and what measures can be taken to stop such similar attacks.

**2. Introduction**

### Global Terrorism Database contains the information of terrorist attacks around the world from 1970 to 2017. With the given dataset we need to make discoveries while exploring the data and also modifying data such that useful information can be obtained

In This project I will be Reading the complete dataset of Global Terrorism Database with the help of python libraries and exploring dataset by providing the useful information in a graphical representation form. Which includes finding information such as what type of attacks, Where was the most terrorist attacks occurred, Different type of weapons used and many such information can be analyzed.

## **3. Exploratory Data Analysis**

Exploratory Data Analysis [EDA] will be Performed on Global Terrorism Database. With the help of EDA we can get useful insights of database which includes information such as finding the total number of features and records, whether the dataset contains any null values different type of features and also with help of pandas and Matplotlib we can perform feature engineering and get the required results.

While performing EDA we can get to know how the different features are correlated with each other which features is more valuable, Find out the group of terrorist who are more active in which part of the country or region, the motive of their attacks life’s lost due to terror attacks and many such things.

* **Null Values Treatment**

Dataset contains a large number of null values which might tend to disturb our accuracy hence I dropped all the features which contains null values more than 10% of total dataset at the beginning of project In order to get a better result.

After Dropping out features with More Null values there was 45 features out of the 45 features only few columns will be selected with which I will be performing Exploratory Data analysis.

* **Features Selection:**

After Removing all the Features with highest Null values there was 45 features out of which only the required features with the most information were selected which was further used to analyze the data the features selected are:

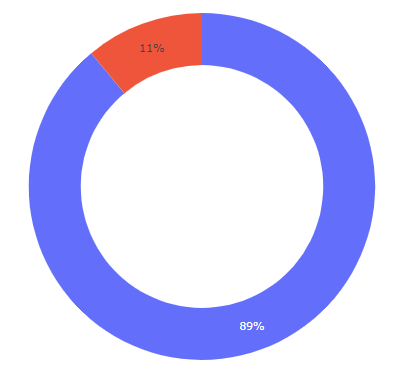
1. eventid: Unique ID of each terrorist attack.
2. iyear: year of terrorist attack.
3. country\_txt: Name of country where the attack took place
4. region\_txt: Name of region where the attack took place.
5. provstate: Name of the state where attack took place.
6. city: Name of city where attack took place
7. latitude: latitude of the terrorist attacked place which provides location information
8. longitude: longitude of the terrorist attacked place which provides location information
9. specificity: specificity gives the information how precise the attack was
10. success: This features indicates whether the terrorist attack was successful or not. It has two values 1 if Successful 0 if Not successful
11. suicide: This features indicates whether the terrorist attack was suicide or not. It has two values 1 if Suicide 0 if Not Suicide.
12. atacktype1\_txt: This features indicates the type of terrorist attack like how the terrorist planned the attack.
13. targettype1: the target of the attack.
14. targetsubtype1: This features indicates what was the main target of the terrorists.
15. gname: Name of the terrorist group
16. weapontype1\_txt: Kind of weapon used for the attack
17. nkill: This features has the info of the total number of deaths due to terrorist attack
18. nowund: This features has the info of the total number of people wounded due to terrorist attack

With the above features only nkill and n wounded are highly correlated because those are the features which are the results of the attack.

And also all the features names are renamed for easy understanding of the Dataset.

# **4. Observations from Exploratory Data Analysis.**

* In the majority of acts of terrorism, the mortality rate and injuries were low, but a small number of actions led to too many deaths and injuries.
* The terrorist activity was most reported after the year 2000. There is a drastic increase in terrorist activity after 2004
* Middle East and North Africa, South Asia has Terrorist incidents which has more than 40000 cases.
* The type of Weapon Type used in in the attack is mostly Explosives and Firearms
* Most of attack type which was used by terrorist was Bombing/Explosion
* Most of the target chosen by terrorist was Private citizens and property which might be because they are the easy target for them. And also to get attention toward them terrorist perform such activities.
* Iraq is the country where most of the attacks took place
* Success and Failure ratio

Success Ration is around 89% whereas the Failure ratio is 11 percent which indicates that most of the attacks planned by the terrorist are well executed.

* Total deaths due to terrorist attack is 411868
* Total people wounded due to this terrorist attack are 523869.
* There's a steady increase in terrorist attack from 2004 to 2014, And Decrease from 2015 to 2017 which indicates that there might be some measures taken by government to prevent the terrorist attack.