Global Terrorism: An Exploratory Data Analysis

Nidhi Pandey (Data science trainee)
AlmaBetter

Abstract

Terrorism has been a major threat to people's lives for centuries. This has recently emerged to be one of the most powerful and violent acts to suppress the free will of any thinking entity by means of intimidation. Judging global peace by just the number of attacks can often be one-dimensional and would potentially lead one to massively underestimate an imminent threat. As the well-known adage goes, it's the punch you don't see coming that knocks you out. Hence, there is an inherent need to study the past in order to avoid repeating it.

In this project, we will be performing Exploratory Data Analysis on Global Terrorism Database collected by the National Consortium for the Study of Terrorism and Responses to Terrorism (START).

1. Problem Statement

The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2017. The GTD 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.

The objectives of the projects are:

a) To import the necessary modules and datasets.
b) To perform data wrangling on the datasets to perform Data Analysis successfully
c) To obtain useful insights from the dataset through Data Exploration and Visualization techniques.

2. Introduction

The Global Terrorism Database has provided a vast record of terror attacks with diverse list of specifications for each record. In order to prevent the data's own details obscuring it, this project views the subject of the matter through four lenses:

- Organization
- Regional & Global Trends
- Type of Attack
- Target Type

The main factors chosen to determine the magnitude of the attack are the number of fatalities and number of wounded victims.
Insights from the data were obtained using data visualisations like bar plot, pie plot, strip plot, scatter plot and line plot.

3. Data

The Global Terrorism Database contains a record of over 180,000 terror attacks from 1970 through 2017 with 137 attributes. These attributes were cut down to 20. The attributes are listed below.

- eventid: Unique ID number assigned each terror attack.
- iyear: The year the terror attack took place.
- country_txt: Name of the country the terror attack took place in.
- region_txt: Name of the subcontinent the terror attack took place in.
- city: Name of the city the terror attack took place in
- attacktype1, attacktype1_txt: Numerical Encoding and the corresponding type of attack that took place.
- weaptype1weaptype1_txt: Numerical Encoding and the type of weapon used to propagate the attack.
- targtype1, targtype1_txt: Numerical Encoding and the type of target the attack was perpetrated on.
- nwound: Number of victims wounded from the terror attack.
- gname: Name of the Terrorist Group that perpetrated the attack.
- claimed: Information on whether or not the attack was claimed.
- nkill: Number of Fatalities of the terror attack.
- crit1: Information if the attack had a political, economic, religious, or social goal.
- crit2: Information if the attack had an intention to coerce, intimidate or publicize to a larger audience.
- crit3: Information if the attack was outside international humanitarian law.
- success: Specifies if the attack took place
- suicide: Specifies if the attack was perpetrated by suicide

4. Methodology

The steps taken to arrive at the required data frame are listed below.

4.1. Importing Required Libraries and Dataset

In this step, we import the numpy and pandas modules to do operations on a large amount of data in a quick amount of time and to effectively perform database functions. The encoding format while importing the dataset had to be set to 'ISO-8859-1' instead of the default 'UTF-8'. We take help of the 'read_csv' function in the pandas module for performing this task.

4.2. Selecting Attributes

The Global Terrorism comes with GTD Codebook which is a descriptive guide elaborating all the features of the database. This Codebook provides the information that has been used to pick 20 out 137 attributes of the database. The selection of the attributes was done based on the relevance to the problem, ease to use and the degree of availability.

4.3. Handling Null Values

Presence of null Values was detected using the 'isna( )' function from the pandas module. The Null values were observed only in the numerical columns of the dataset: 'nkill' and 'nwound'. These null values were effectively replaced by their respective median values using the 'fillna( )' function from the pandas module.

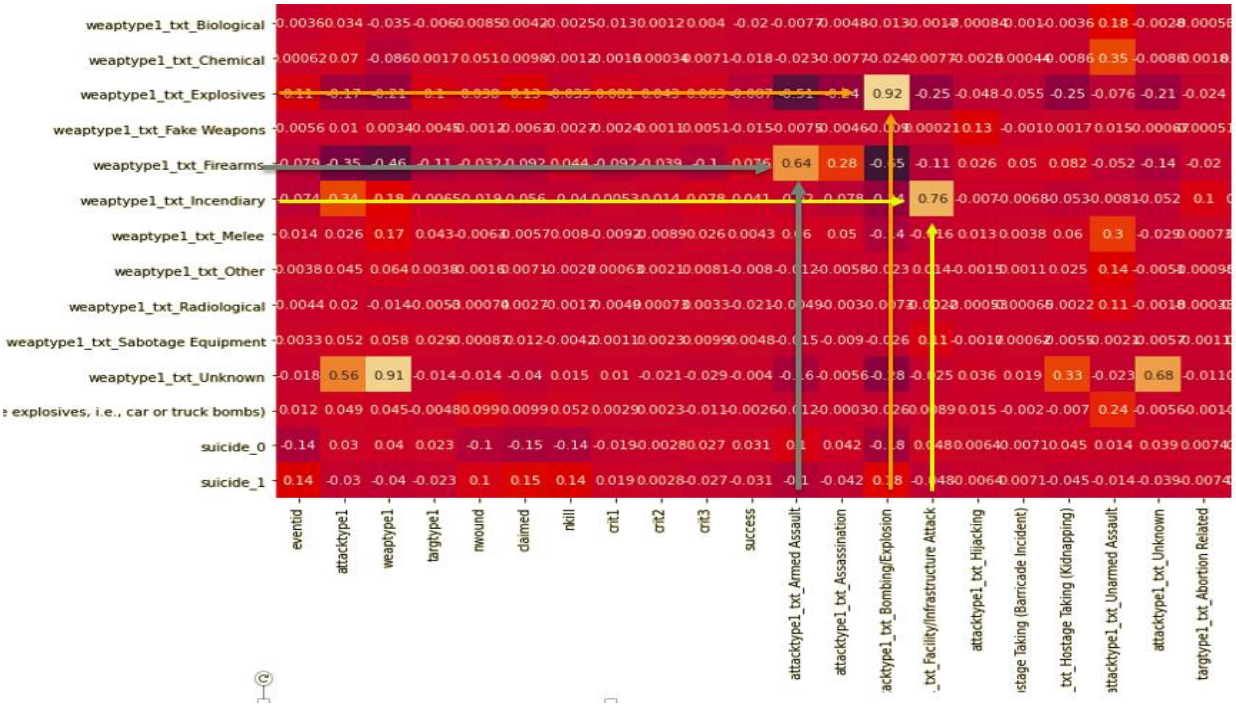

Fig 4.4.1: Correlation between weapons and types of attacks

4.4. Data Exploration

Once the required dataset was cleaned and obtained, the data exploration stage begins to find outliers and notable correlations. These insights are used to suggest what to look for when performing analysis.

The summaries of the data's central tendency, dispersion and shape of data distribution are obtained to get an idea about the numerical columns of the dataset.

Due to the presence of categorical data, traditional correlation methods did not yield useful results. This obstacle was tackled using One Hot Encoding with the 'get dummies( )' function from the pandas module. One Hot Encoding finds every unique value in a categorical column and these unique values into separate columns with binary values to indicate the category (1, if the row belongs to the category and 0, if not).
Seaborn module's heatmap function is used to visualize the correlation between columns of the dataset. Relations between Type of Attack and Type of Weapon can be noted from the Correlation Heatmap given in Fig. 4.4.1.

5. Analysis

5.1 Most Notorious Organization

The dataset is grouped by the Terrorist Group that is listed to have perpetrated the crime and their part in the terror attack and casualties are analysed.

The most notorious terrorist organization is being assessed by the number of attacks initiated and the number of lives taken.

The most notorious terrorist organizations are

1. Islamic State of Iraq and the Levant (ISIL)
2. Taliban
3. Boko Haram
4. Shining Path (SL)

These organizations alone have contributed to more than one third of the total fatalities and almost 50 percent of the total terror attacks in the world.

Furthermore, when we plot the top 10 terrorist groups' activity from years 1917 through 2017 against the number of fatalities lost that year; we can get a better understanding about the rise and fall of different terrorist groups

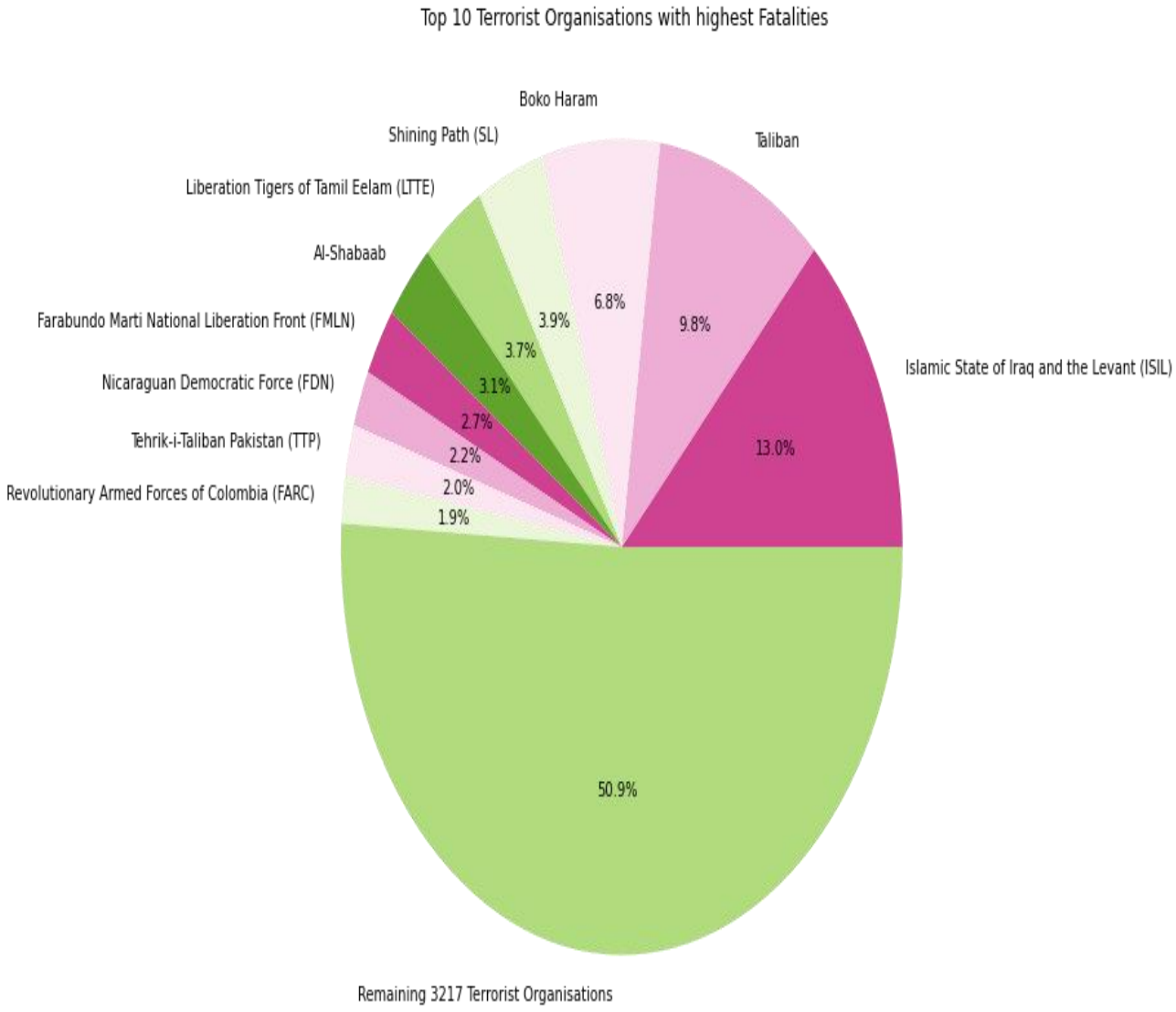

Fig 5.1.1: Top 10 Terrorist Organization with Highest fatalities

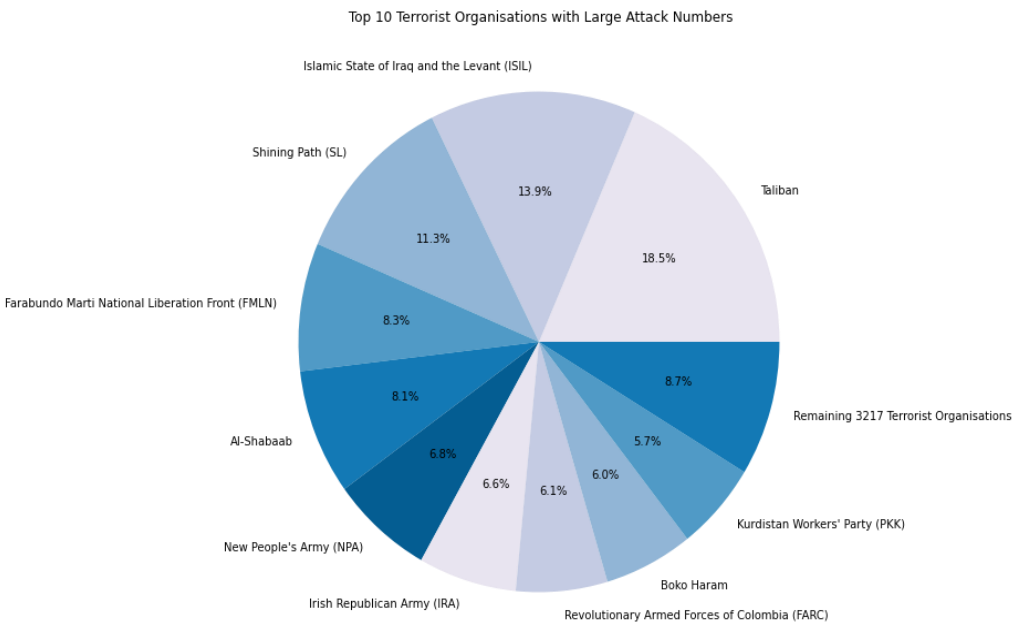

Fig 5.1.2: Top 10 terrorist organizations with large attack numbers

5 .2 Visualizing relation between number Fatalities and Wounded victims

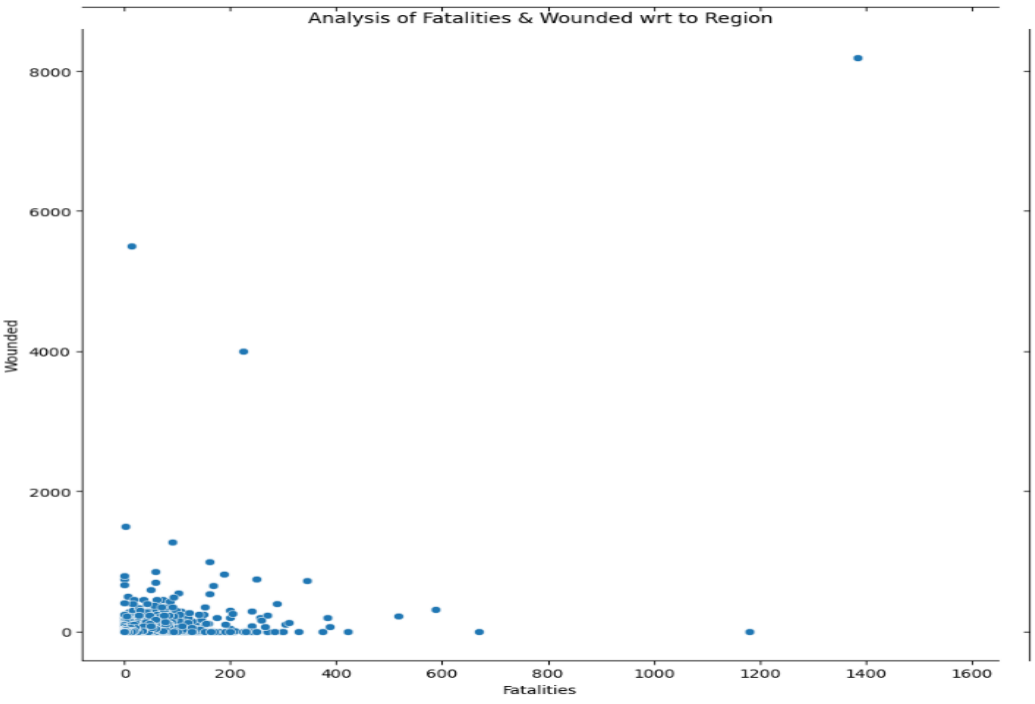

Fig 5.2.1 Fatalities and no. of Wounded people

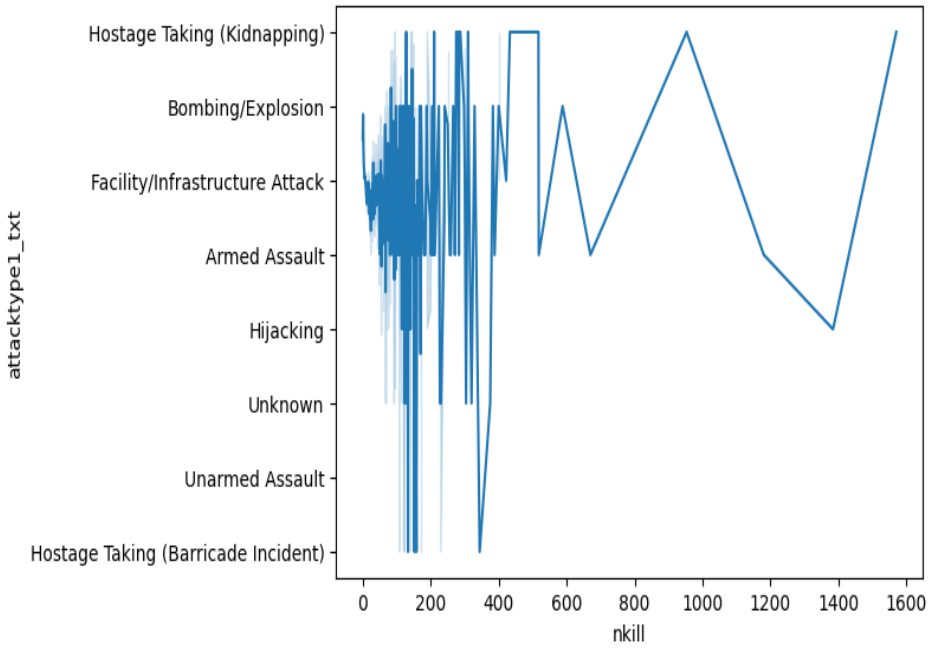

Fig 5.2.2 This Fig. shows the type of attack and the no. of people killed during this attack

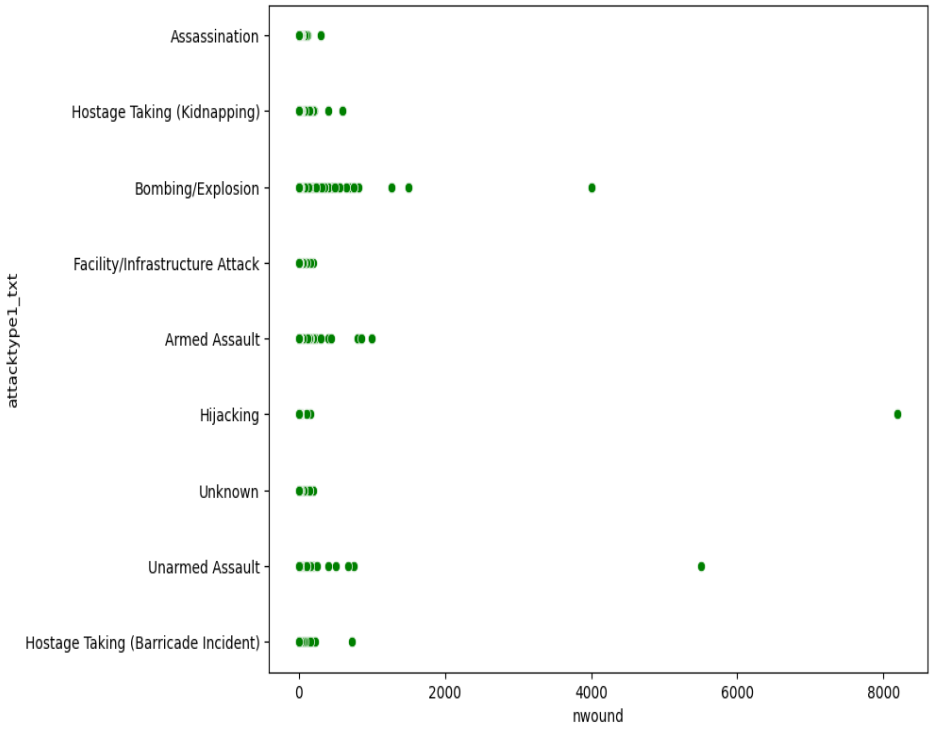

Fig 5.2.3 Attack type and no. of Wounded people

From fig 5.2.1, 5.2.2 and 5.2.3 it is clear that: -

- The graphic shows the breakdown of the total number of assaults' fatalities and wounded victims. From this, we might infer that most attacks result in less than 200 fatalities and 1,000 injured people.
- we have observe from the above graph that the maximum number of kills has been due to Bombing/Explosion, Facility and infrastructure attack , and Armed assault.
- From the chart more than 8000 fatalities are caused during these attacks. Most of it happens because of Hijacking, unarmed assault and bombing/explosion.

5.3 Analysis of Global terrorist attack trend

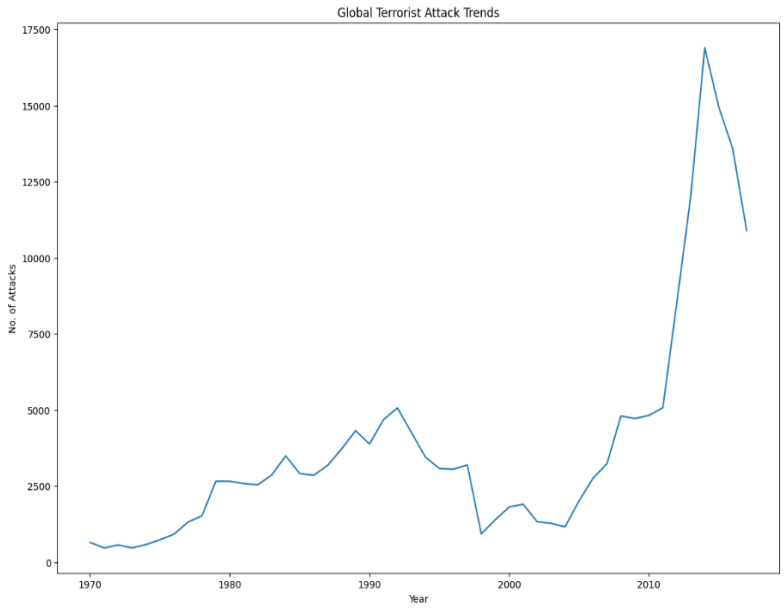

Fig 5.3.1Global Terrorist Attack Trend

Global & Regional Trends

Fatalities of terror attacks in different regions of the world are plotted against time to gain insights about Regional Trends. This is compared to the terror attack trends of the whole world.

Upon doing so we found that Global Terrorism started increasing to an all-time high from 2011. This peaked in the year 2014 and started dipping since (Fig. 5.2.1).

From fig 5.3.1

- From the above chart we can see the growth in global terrorism trend since 1970 it grown in between those years as we observe through the chart. And it is at its

peaks after 2010 and Global Terrorism started increasing to an all-time high from 2011. This peaked in the year 2014 and started dipping ever since.

5.4 Type of Attack

In this section, terror attacks will be analyzed by the type of attack perpetrated. We'll be looking at attacks that are perpetrated frequently and the ones that have claimed the most lives (Fig. 5.3.1). The 3 most perpetrated types of attack are:

- Bombing/Explosion (more than twice as frequent than Armed Assassination
- Armed Assault (twice as frequent than Assassinations)
- Assassination

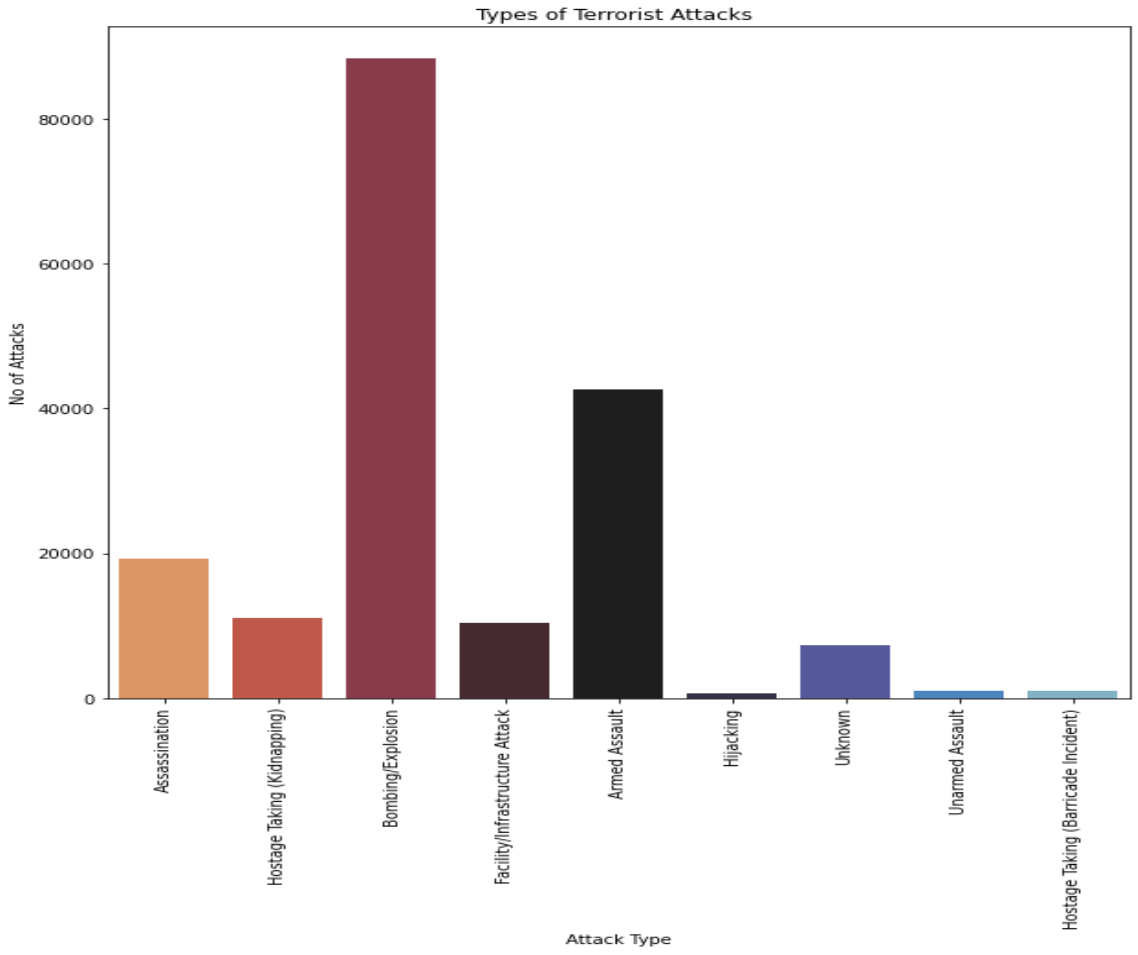

Fig 5.4.1: Types of terrorist attacks

The frequency of the top 3 attacks increases by two folds as we move up the ladder

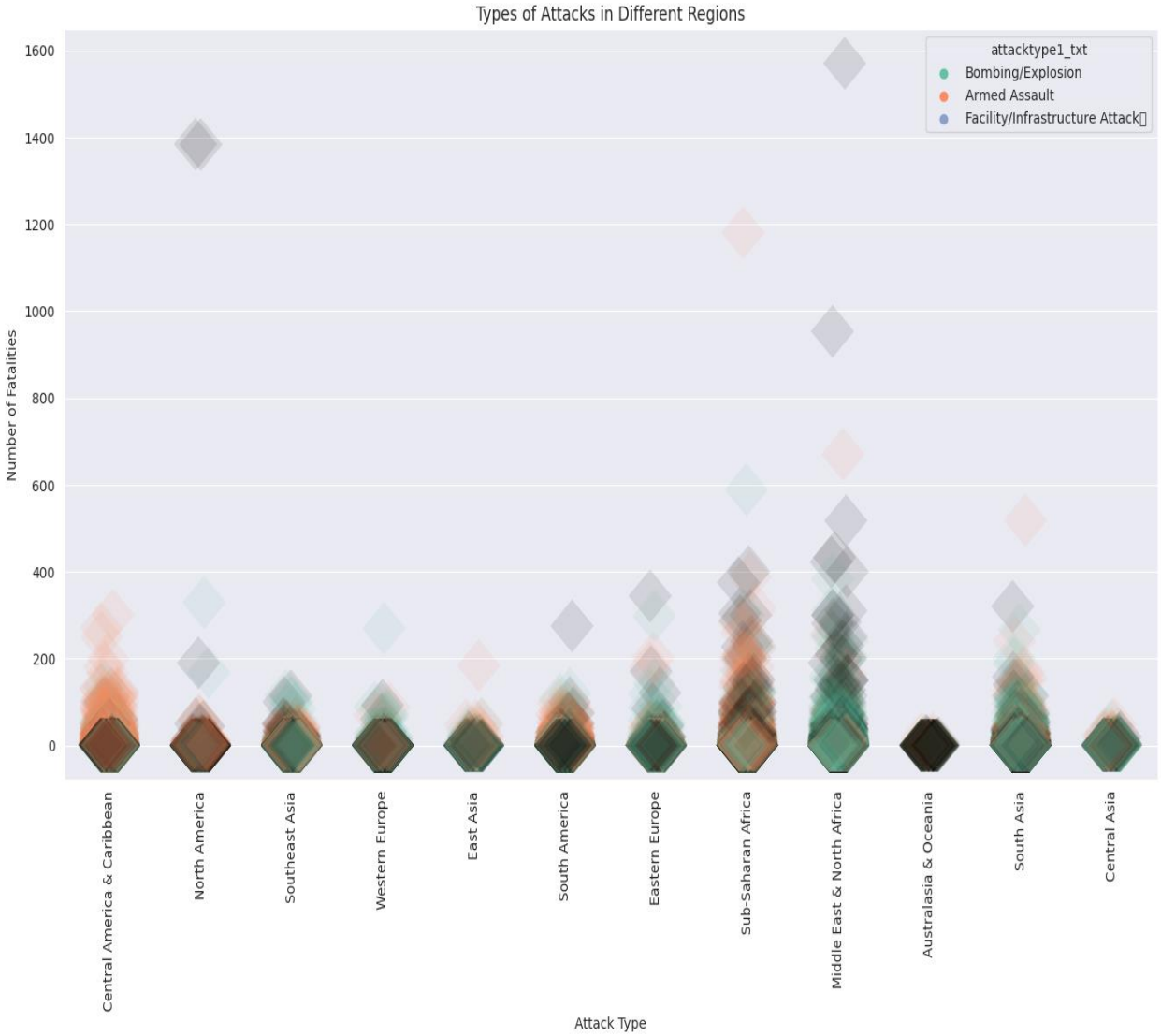

Fig 5.4.2: Types of attacks in different regions

Armed Assault is a dominant type of attack in regions like
1) Central America & Caribbean
2) Sub-Saharan Africa and
3) North America
(Presence of dense orange)

Bombings/Explosion is a dominant type of attack in regions like
1) Middle East & North Africa
2) South Asia and
3) Eastern Europe
(Presence of dense green)

5.5 Type of Target Analysis

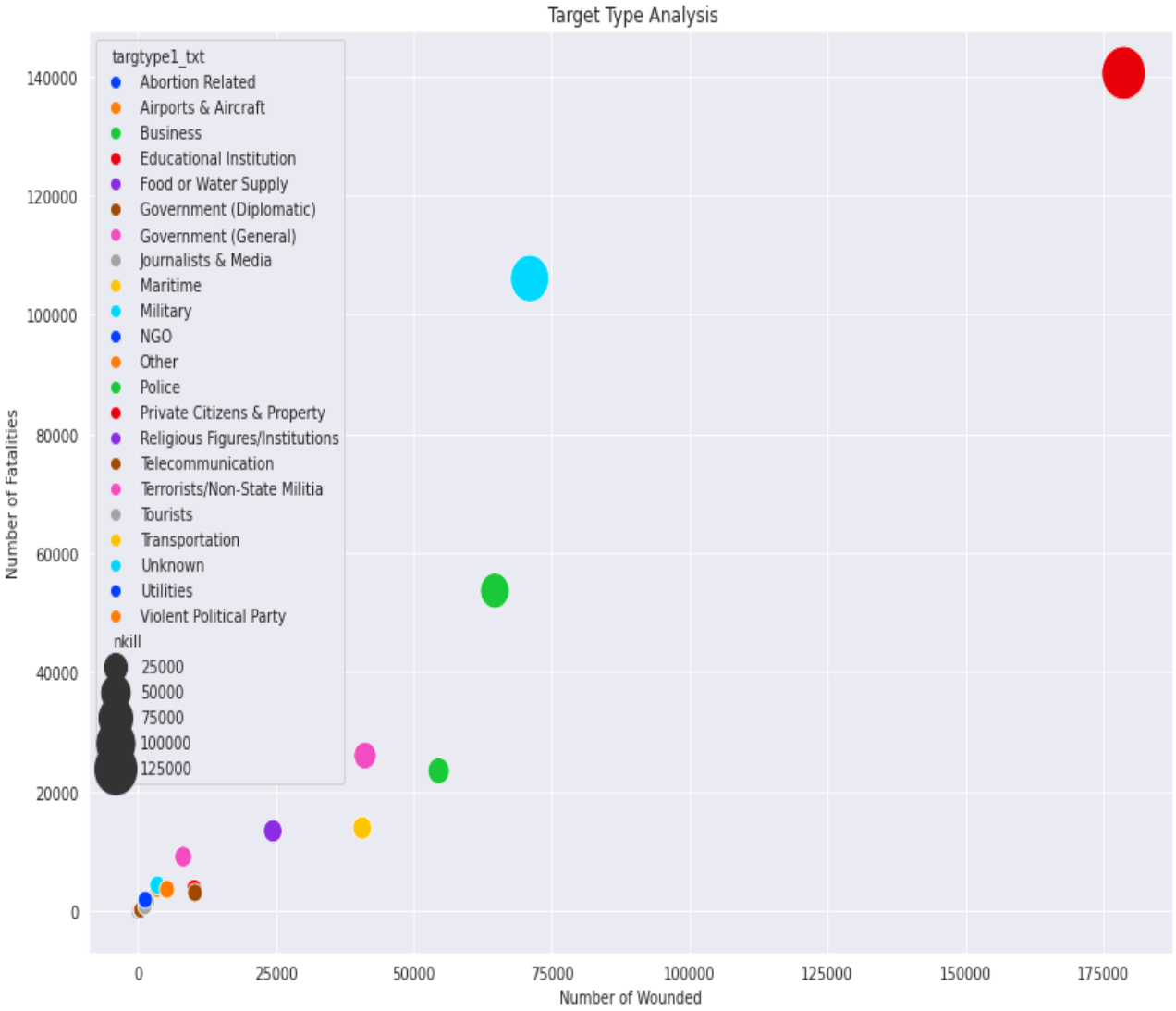

Fig 5.5.1: Analysing Target Types by Fatalities and Wounded

From above chart we can say that the most targets are

1) Educational institution

2)Private citizen and properties

3)Police

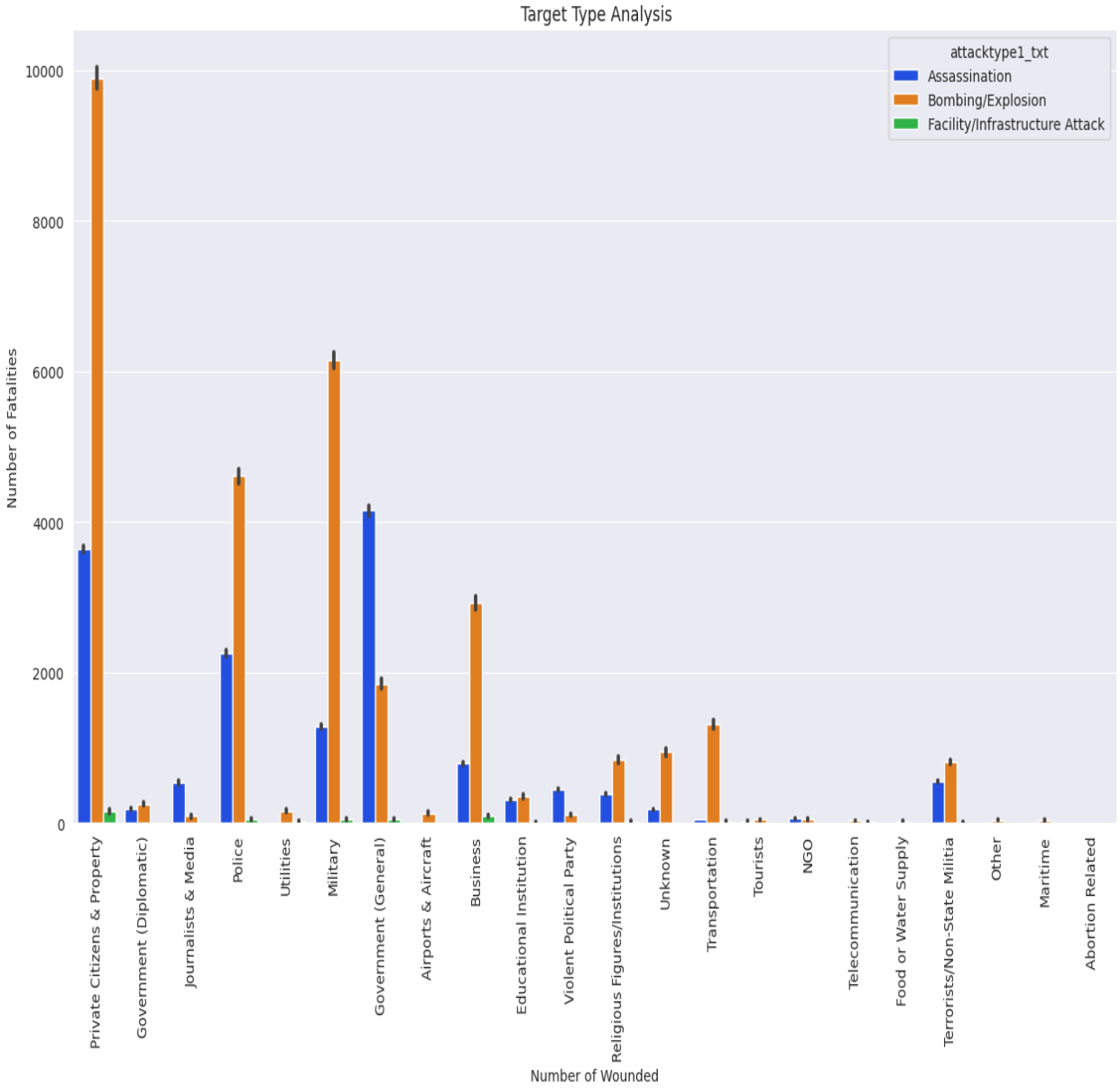

Fig 5.5.2 Target Type Analysis

The most common attacks are

1)Assassination
2)Bombing/Explosion
3)Facilities and infrastructure attack
Assassination is the most common type attack perpetrated on Government officials and Media Journalists.

Bombings are frequently used to thwart Businesses and Religious Figures/Institutions.

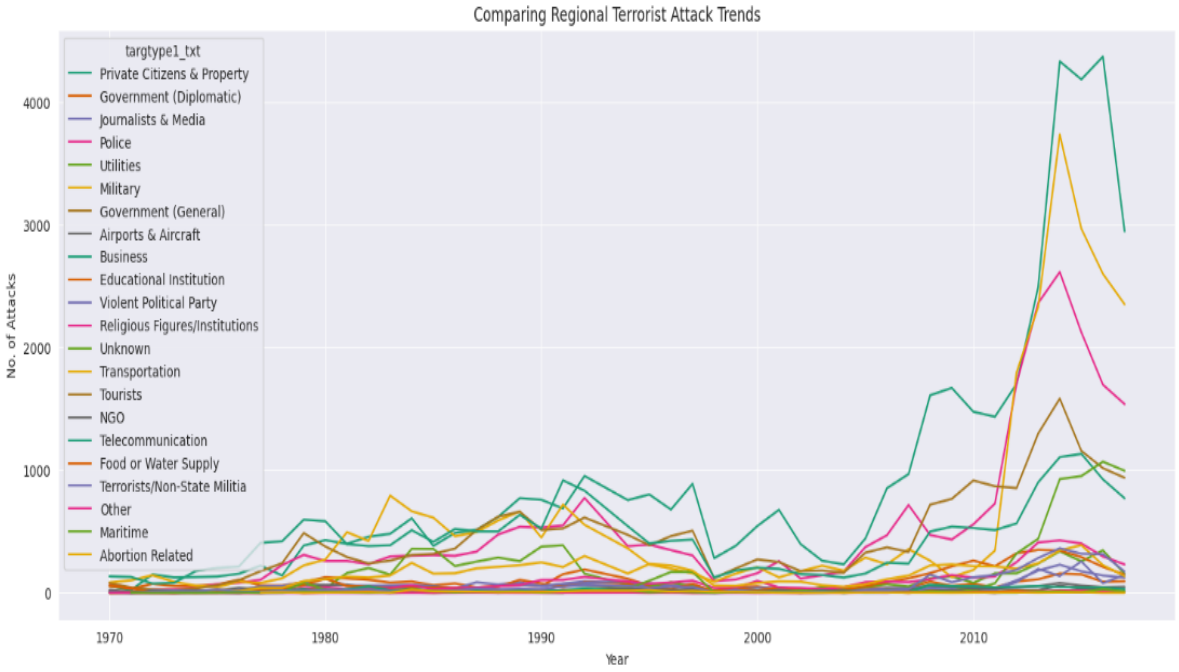

Fig 5.5.3 comparing most common target types in regional attack trend

By comparing regional terrorist attack, we got to know that the most target types are

1)Private citizen and Property

2)Military

3)Police

7. Correlation Heatmap

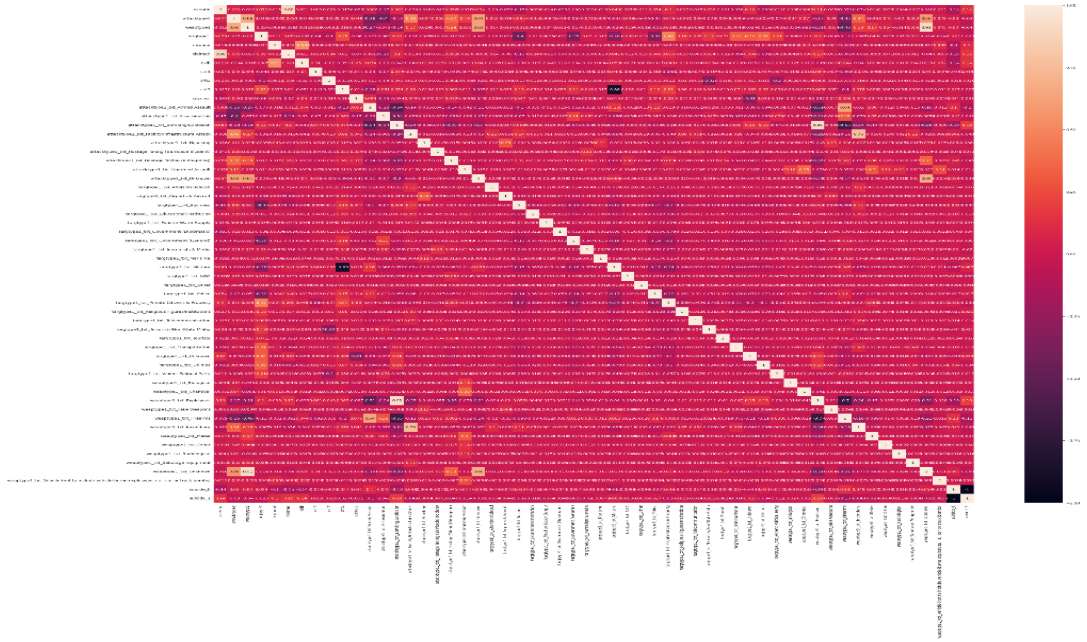

Fig 6.1 Correlation Heatmap

- We observe positive correlation between weapons and types of attacks
- Use of Explosives and Bombings/Explosions (0.91)
- Use of Firearms and Armed Assault (0.68)
- Use of Incendiary and Facility/Infrastructure Attack (0.71)

7. Conclusion

- ***We got to know about most affected target types are***
  - Private Citizens and property
  - Police
  - Military

- **Analysis of the types of attacks and number of causalities**
  - Bombing/Explosion (more than twice as frequent than Armed Assault)
  - Armed Assault (twice as frequent than Assassinations)
  - Assassination - *More than 80000 causalities have been happened due to these attacks*

- ***Analysis of Top organization with large attacks***
  - Taliban
  - 2)Islamic state of Iraq and the Levant
  - 3)shining Path (SL)

  The Civil War between Iraq and Islamic State of Iraq and the Levant (ISIL) has claimed 1570 fatalities, the largest taken in any terror attack

- **Most Notorious Terror Organisations**
  - Taliban
  - Boko Haram
  - Islamic State of Iraq and the Levant (ISIL)
  - Shining Path (SL)

  These organisations alone have contributed to more than one third of the total fatalities and almost 50% of the total attacks

- **Common Attacks Perpetrated on different Targets**
  - Assassination
  - Bombing/Explosion
  - Facilities and infrastructure attack

  Assassination is the most common type attack perpetrated on Government officials and Media Journalists. Bombings are frequently used to thwart Businesses and Religious Figures/Institutions.

- **Visualizing the extent of the commonly perpetrated attacks in different regions of the country**
  - *Armed Assault is a dominant type of attack in regions like*
  - Central America & Caribbean

- Sub-Saharan Africa and
- North America

❖ ***Bombings/Explosion is a dominant type of attack in regions like***
- Middle East & North Africa
- South Asia and
- Eastern Europe

❖ **Analysis of Global terrorist attack trend**
From the above chart we can see the growth in global terrorism trend since 1970 it grown in between those years as we observe through the chart. And it is at its peaks after 2010 and Global Terrorism started increasing to an all-time high from 2011. This peaked in the year 2014 and started dipping ever since.

❖ **Analysis of the number of fatalities**
From the chart more than 8000 people were wounded during these attacks. Most of it happens because of "Bombing/Explosion".

❖ **Correlation/Heatmap**
- We observe positive correlation between weapons and types of attacks
- Use of Explosives and Bombings/Explosions (0.91)
- Use of Firearms and Armed Assault (0.68)

Use of Incendiary and Facility/Infrastructure Attack (0.71)

8. References

1. GTD (Global Terrorism Database) - Codebook: Inclusion Criteria and Variables, October 2019