Global Terrorism Analysis by EDA

Prepared by
SAAQUIB MUSTAFA
ARABINDA PATTNAIK
SAHIL KOLAMBKAR
RAHUL MUKHERJEE
SANDEEP

Abstract:

Terrorism is a tactic that has been used by radical groups for decades to achieve their goals. The purpose of terrorism is to create fear and panic in order to establish power and control over a society. Terrorist attacks are not limited to any one type of group, but they are more often carried out by different organizations.

Our EDA can make us understand about the role of terrorism in different region of the world, how much it destroyed our economy and most importantly, the loss of innocent people life for different countries. Also, we will study the different patterns of a terrorist organization, how they attack, which weapons they use, etc.

1.Problem Statement

The Global terrorism database (GTD) is an open source database including information on terrorist attacks around the world from 1970 to 2017.

The GTD includes systematic data on domestic as well as international terrorist attacks that have been occurred during this time period.

Through this database we can find out some of relevant queries regarding the behaviour of terrorist organization, like how they will attack, what weapons they use maximum time, which region of this world faced the most attacks as well as we can find which countries suffered the most.

By analyzing this database we can figure out the recent activities of terrorist organization in our country India. Which terrorist organization is more active in India in the recent years?

Our security agencies could use this data to avoid further terrorist attacks by neutralizing them.

The main objective is to do Exploratory Data Analysis (EDA) on the dataset which was provided, to conclude the facts about the certain features which are mentioned.

The Global Terrorism dataset has around 181,691 observations in it with 135 columns and it is a mix between categorical and numeric values**.**

From this large amount of date we filtered out some important columns, with which we are going to analyse our problems. These are the required columns data mention below.

1. eventid: It is the unique ID of each entries in database.

2. iyear: It is the year in which the terror attack occurred

3. country_txt:It is the name of the country in which the attack happened

4. attacktype1_txt: Inside this column we will find the type of attack.

5. targettype1_txt: In this column we will find what terrorists aim to target mainly.

6.gname: Thename of terrorist organization.

7. nkill: Number of casualties happened in their respective attack

8. nwound: It gives us information about the minimum number of wounded people during that attack.

9. weapontype1_txt: It gives information about the type of weapons used by terrorists.

10. weapsubtype1_txt: In this column we can find the particular weapon name which terrorist used

11. ransomamt: Through this column, we can find the ransom amount demanded by terrorists.

12. propvalue: In this column we will find the value of properties damaged after the terrorist attack.

13. motive: In this column, terrorists' motive behind their attack described.

14. region_txt: In this column we will find the region of this world affected by terrorism.

These 14 columns mention above are to be considered for our further analyse.

2. <u>Steps involved:</u>

- **Data Wrangling:**
  After loading the dataset we performed this method by cleaning, organizing, and selecting those particular columns which makes us to understand the data clearly. This process helped us to tackle the unwanted data, to produce accurate results, to make better decision.

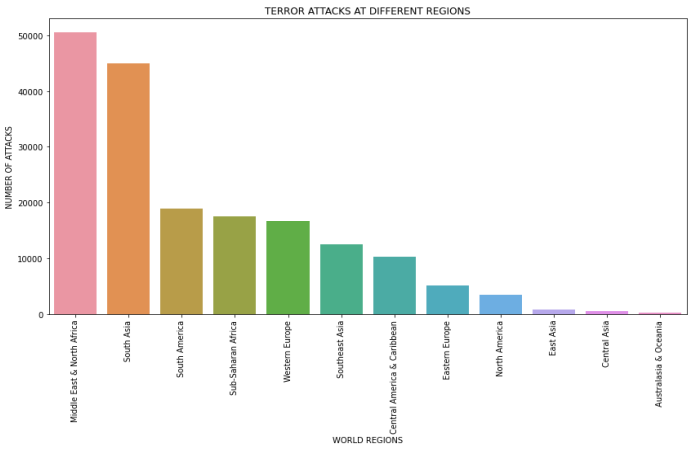
- **Exploratory Data Analysis (EDA):**
  After the data wrangling step we performed EDA by comparing different parameters which are involved in the dataset. EDA help us to find the different relations among the parameters. It involves the visualization of the data by comparing the different parameters to find out the best among all.

3. EDA:

In the process of understanding the data, first of all we found out the region in the world most affected by terrorism. For which we took the **region_txt**column and by applying the value**_count()** method we got our required result.

```
region_attacks = terrorism_df.sort_values("region_txt", ascending=False)
region_attacks_counts = region_attacks['region_txt'].value_counts()
```

`terrorism_df`is our main dataframe variable containing the original CSV file.



(Fig-1)

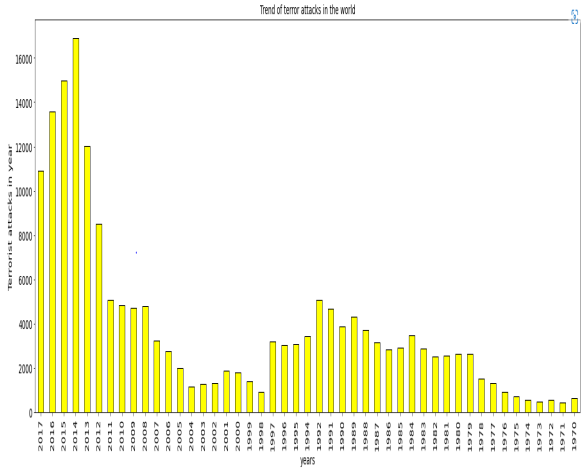
After plotting the required graph shown above **(Fig-1)**.It is clearly seen that

Middle east and North Africa are the worst affected by terrorism activities, about more than 50,000 recorded terrorist activities, followed by South Asia and South America.

When we came to check the frequency of terrorist attacks per year in the world, we have to use the `iyear` column to count their values to find out how many terrorists attack reported per year. To get the result we used this code mention below.

```
freq_attacks_w=terrorism_df['iyear'].value_counts().sort_index(ascending=False)
```

When we plot this output in a bar graph, this is the result we got.



(Fig-2)

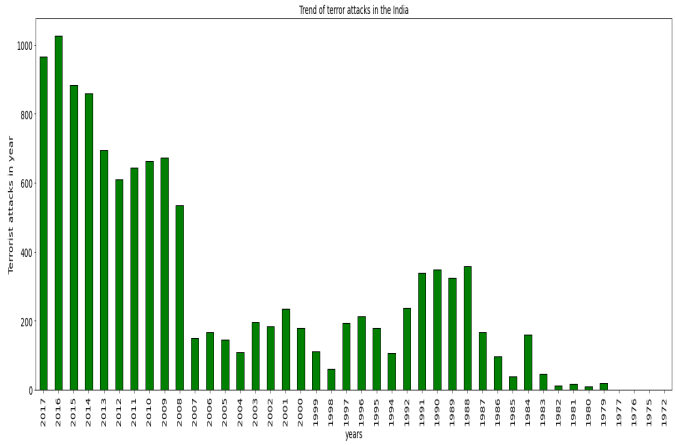
The year 2014 turned out to be the year in which maximum number of terrorist attack reported in the whole world followed by 2015 in 2nd position and 2016 became the 3rd year in which maximum attacks occurred respectively.

When we came to see the number for our country India, we have to put a condition in our code to find out the number of terrorist attacks happened in India only.

So we used this code to filtered out India data from the main DataFrame.

```
terrorism_india=terrorism_df[terrorism_df['country_txt']=='India']
freq_attacks_i=terrorism_india['iyear'].value_counts().sort_index(ascending=False)
```

By plotting the output data in graph, we found this result.



(Fig-3)

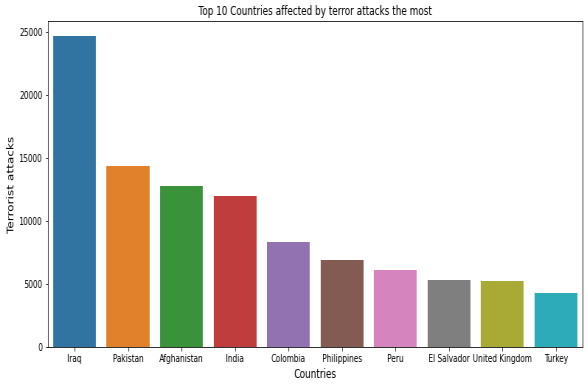
In the year 2016 India reported the maximum number of terrorist attacks which is more than a 1000 number, followed by 2017 and 2015 became the 3rd year in which maximum number of attacks reported respectively.

In the recent years Terrorists are became more active in all over the world, which might be sponsored terrorism, but this terrorism doesn't see any boundaries. Over the years each and every country faced terrorism in their own way and faced many difficulties. So let's find out top 10 countries mostly affected by terrorism in the recent years.

To get the required result we used this code mentioned below.

```
affected_w=terrorism_df['country_txt'].value_count
top_ten_affected=affected_w.iloc[:10]
```

By using the output we found out our required result shown below.



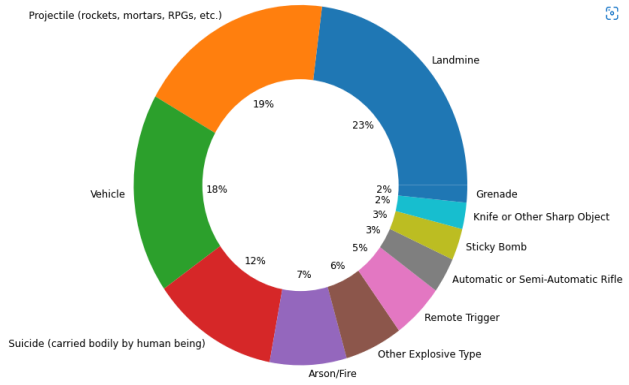
(Fig-4)

By analyzing the above graph we can clearly see Iraq faced the maximum number of terrorist attack followed by Pakistan and Afghanistan.
India is the 4th nation who faced the maximum number of terrorist attack.

When we look into the different terrorist organizations, we found out there are some particular terrorist organizations are more active in the recent years, like Taliban, Al-Qaida, ISIS, etc.
So, let's discuss briefly about Taliban, like what weapon they use maximum time and their way of attack.
This following code will help us to find the required details about the particular weapons used by Taliban for their operation.

```
taliban_df=terrorism_df[terrorism_df['gname']=='Taliban']
taliban_df_weapons_used=taliban_df['weapsubtype1_txt'].value_counts().reset_index()
count = list(taliban_df_weapons_used['weapsubtype1_txt'])
count1=count[2:13]
count1=np.array(count1)
mylabels = list(taliban_df_weapons_used['index'])
mylabels1=mylabels[2:13]
```

By plotting this, we found this result.



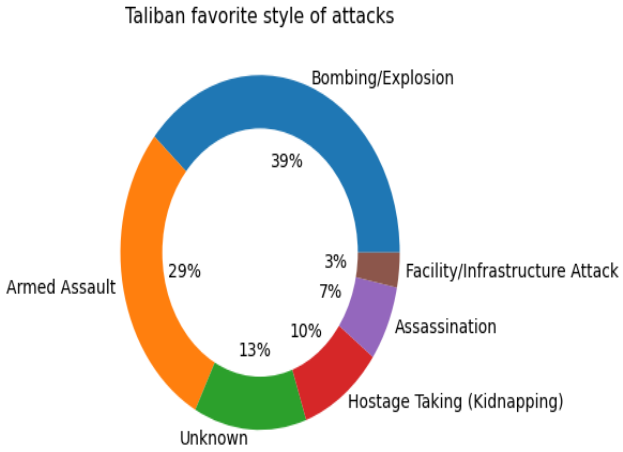
(Fig-5)

In this above pie chart **(Fig-5)**. It was found out that Taliban is using a maximum number of landmines and projectile armours like rockets, mortars, and RPGs.

When we analise their weapon choice, we need to identify how they are using these weapons to execute their operations. So by using **attacktype1_txt** column we found out their way of terror operation.
This is the following code we used to find out our result.

```
favorite_style=taliban_df['attacktype1_txt'].value_counts().reset_index()
favorite_style_name=list(favorite_style['index'])
favorite_style_name1=favorite_style_name[:6]
favorite_style_count=list(favorite_style['attacktype1_txt'])
favorite_style_count1=favorite_style_count[:6]
```

By plotting the above output we found out this result.



(Fig-6)

By looking at the chart (Fig-6). We can clearly see that Taliban's way of conducting any attack includes the Bombing/Explosion, followed by armed assault.

Every Terror attack leaves behind a huge amount of financial loss to the country, which may affect their economy deeply.
So we found out the monetary loss occurred to the world due to terrorism per year.

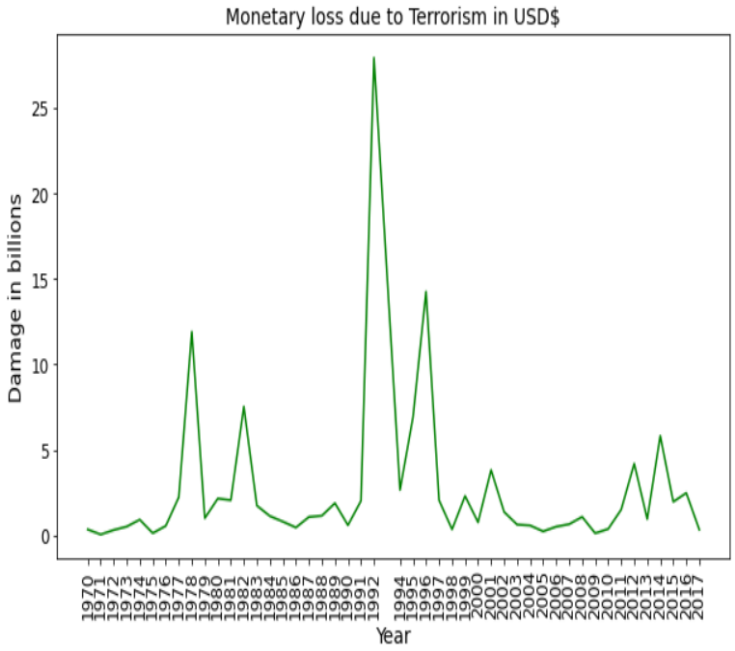
```
propvalue=terrorism_df.groupby('iyear')['propvalue'].sum()
ransomvalue=terrorism_df.groupby('iyear')['ransomamt'].sum()
total_sum=(propvalue+ransomvalue)/100000000
```

By using the above code we can calculate the loss of amount per year.

By creating a line graph using the below code, we will easily identify each year data.

```
total_sum.plot(kind='line',fontsize=12,color='green')
plt.rcParams['figure.figsize']=(14,7)
plt.title('Monetary loss due to Terrorism in USD$',fontsize=14)
plt.ylabel('Damage in billions',fontsize=14)
plt.xlabel('Year',fontsize=14)
plt.xticks(list(total_sum.index),rotation=90)
```

The output result is shown below.



(Fig-7)

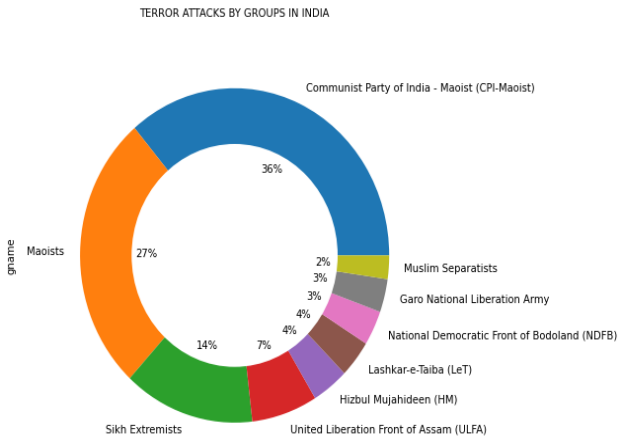
By analyzing the above graph we can identify that between the year 1991 to 1998 the monetary loss was at peak due to terrorism, but particularly in the year 1992 the world witnessed the highest amount of monetary loss due to terror attack which is approximately $25 Billion USD.

In the past years there are many terror organization emerged in our country India. So we will evaluate which terror organization is more actively operating in our country.

To evaluate the data we used the following code mentioned below.

```
terrorism_in_india = terrorism_df[terrorism_df['country_txt'] == 'India']
group_name_local = terrorism_in_india.sort_values("gname", ascending=False)
group_name_count_1 = group_name_local['gname'].value_counts()[1:10]
```

After getting the output result we plot that data in a chart mentioned below.



(Fig-8)

CPI-Maoist turn out to be the major terrorist organization in India, who conducted the maximum number of terrorist attacks followed by Maoists and Sikh extremists respectively.

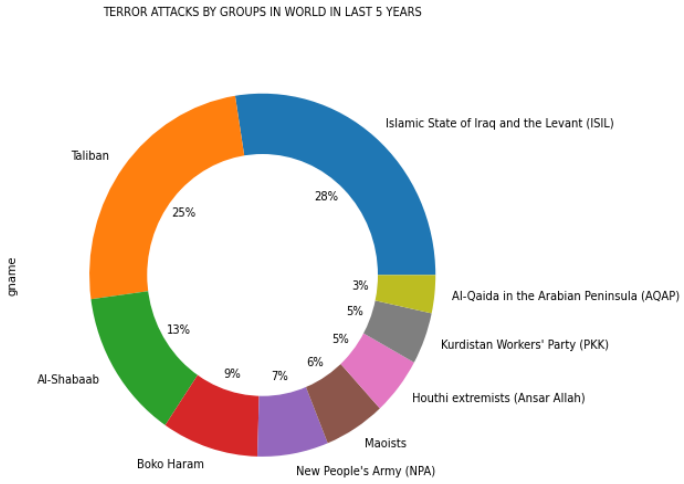
By the years some terror organization weakening their strength but on the other hand some became more powerful. To tackle this situation, we must be aware of the current threat to the world as well as India by analyzing the last 5 years of data.

By using the respective codes shown below we will find out the current threat to the world as well as to India.

```
group_name_local = last_5_yr.sort_values("gname", ascending=False)
group_name_count = group_name_local['gname'].value_counts()[1:10]
```

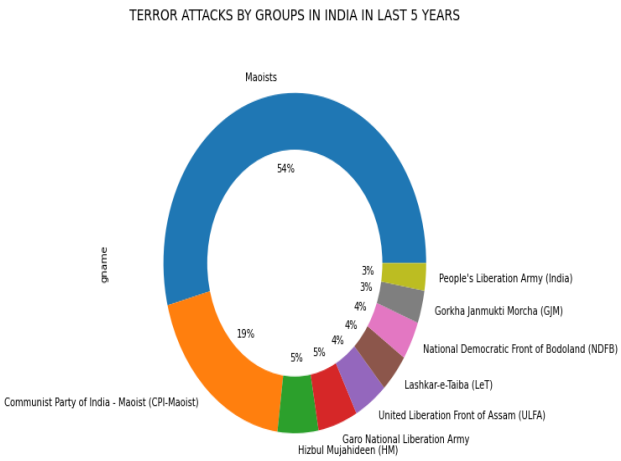
```
groups_attacked_in_india_in_last_5_yr = last_5_yr[last_5_yr['country_txt'] == 'India']
group_name_local_in = groups_attacked_in_india_in_last_5_yr.sort_values("gname", ascending=False)
group_name_count_in = group_name_local_in['gname'].value_counts()[1:10]
```

By plotting the above two outputs individually, we got our required result, which is shown below.



(Fig-9)

By analyzing this chart we can see that ISIL and Taliban are the two major organizations who were more active in last 5 year.



(Fig-10)

By analyzing the data shown in above chart, which is showing the organizations which are more active in India in the last 5 year, Maoists turned out to be the major terror organization who are more active in last 5 year compare to other organization followed by CPI-Maoists and Hizbul Mujahedeen respectively.

Conclusion:

It is evident from the trend analysis that since 1971 there has been significant increase in terror attacks globally. Terrorist Groups like ISIL, Taliban, Al-Shabaab, Boko Haram, NPA, and Maoist are causing havoc throughout the world by bombing, armed assaults & kidnapping. However, in recent times there has been slight decrease in Terrorist Attacks.

Middle East and North African countries like Iraq, Somalia, Nigeria, Yemen, Libya, Egypt, etc are highly affected by Terrorism followed by South Asian countries like Afghanistan, Pakistan & India.

Terrorism is acknowledged as a global threat and India is no exception.

India States like J&K, Chhattisgarh, Manipur, Jharkhand, Assam, Bihar, Orissa, etc are worst effected due to terrorism. J&K suffers for cross border terrorism where as states like Chhattisgarh, Manipur, Jharkhand, Assam, Bihar, Orissa, etc are impacted by Naxal movement.

In India terrorist groups like Maoist, CPI Maoist, Hizbul Mujahideen, Garo National Liberation Army, ULFA, LeT and active as per recent data (2013-2017) and they resort to bombings, armed assault, kidnapping, assassination, etc. Their primary targets are Private Citizens & Property, Police Forces, Government Assets, Business and Military by using Explosives, Grenade, Setting properties on fire, Landmine, etc.

References-

- https://www.fbi.gov/investigate/terrorism
- https://www.mea.gov.in/articles-in-indian-media.htm?dtl/15985/Terrorism+and+the+law++I
- https://www.dhs.gov/sites/default/files/publications/start_global-terrorism-database_coe-factsheet_1611-508.pdf