EDA PROJECT ON GLOBAL TERRORISM ANALYSIS

Ashwani Kumar, Jyoti Shrivastava
Data Science trainees,
Alma Better, Bangalore

Abstract:

This is the final technical document report of our project title "EDA on Global Terrorism Database GTB " as a part of our data analytic Course at Alma better. The goal of our project is to analyze and predict some important aspects of this Global Terrorism dataset which will help us identify major factors and give us insights to mitigate the terror attacks.

Contents:

We are here to explore a Global terrorism dataset to discover the terrorist attacks across the world. This data set contains Global terrorist attacks information from 1970 to 2017, which includes information such as occurrence of terrorist attack, region, country, state, city, latitude, longitude, Target type, Attack type, Weapon type, Terrorist organization committed, death toll and causalities among other things. We will analyze some important aspects of this Global Terrorism dataset which will help us identify major factors and give us insights to mitigate the terror attacks.

WORK FLOW:

 # Problem statement & collection of data: Identity the problem of analysis to be done and then collection the data regarding it.

 # Data Cleaning & Manipulation: Raw data will have lot of null values and missing data. so it needs to be cleaned and manipulated as per our analysis

 # Exploration Data Analysis: By using python libraries we write code to study and extract result.

 # Data Visualization: The results of our analysis will be displayed in a isualization manner for better readability and understandability.

 # Conclusion & Proposed Actions: Based on the EDA results conclusions will be drawn and necessary actions will be proposed to mitigate the risk.

Approach:

1-Let us first check our dataset's and understand it.
2-Later we will check for any missing data in the data given.
3-We would check the type of data and divide it for our analysis.
4-We checked where there any outlier or unethical data in it if so, we would filer such data for specific analysis.
5-Then do Data analysis by visualization techniques.
6-And then conclude with various outcomes from it.

Dataset's:

The Dataset contains

No. of Rows = 181691

No. of Columns = 135

▪ There are 106 columns out of 135 columns are having missing values.

▪ We remove the columns with 85% missing values in the Dataset. After dropping we are left with 47 columns.

▪ In those 47, We drop some columns which are not useful for our analysis. For instance, country column represent the numerical representation of country, country txt column represents the name of the country. So, the column 'country' is not useful for our analysis. Likewise other columns are as below that we dropped.

➢ 'eventid', 'extended', 'iday', 'country' , 'targsubtype1', 'region', 'weaptype1', 'weapsubtype1', specificity, 'vicinity', 'crit1', 'crit2', 'crit3', 'doubtterr', 'multiple', 'suicide', 'attacktype1', 'targtype1', 'targsubtype1_txt', 'target1', 'natlty1_txt', 'guncertain1', 'individual', 'weapsubtype1_txt', 'property', 'dbsource', 'INT_LOG', 'INT_IDEO', 'INT_MISC', 'INT_ANY'

▪ Our Final Dataset will have 181691 rows and 17 columns.

ATTRIBUTE INFORMATION:-

- **iyear:** The Year in which attacks were happened
- **imonth:** The Month in which attacks were happened
- **country_txt:** Name of the country
- **region_txt:** Name of the Region
- **provstate:** Name of the State
- **city:** Name of the City
- **latitude:** Latitude of the terrorist attacked location
- **longitude:** Longitude of the terrorist attacked location
- **attacktype1_txt:** Attack type of the terrorist attack
- **natlty1:** Indicates victim belongs to which nation
- **success:** (0 or 1) indicates whether attack was failure or success
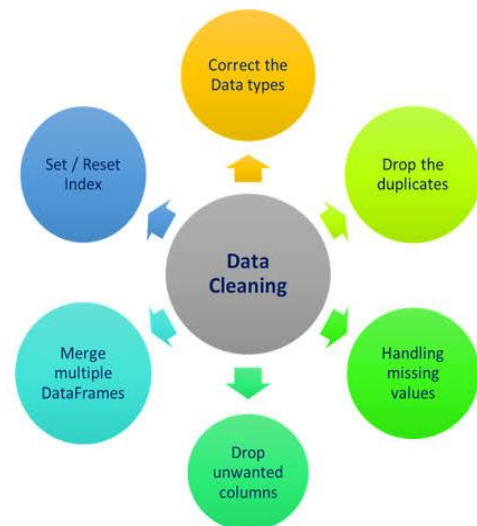- **targtype1_txt:** Target type of the terrorist attack

- **gname:** Name of the gang committed the terrorist attack
- **weaptype1_txt:** Type of Weapons used for terrorist attack
- **nkill:** no of people killed during the terrorist attack
- **nwound:** no of people wounded during the terrorist attack
- **Ishostkid:** (0 or 1) Indicates whether hostage is kid or not in the attack
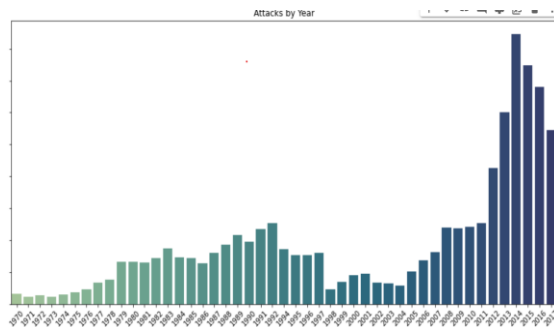
Exploratory Data Analysis:

Now we know that we are ready for an exploration of our data, we can make a rule that we are going to be working from left to right. The reason some may prefer to do this is due to its set approach - some datasets have a substantial number of attributes; plus, this way we will remember to explore each column individually to make sure we learn as much as we can about our dataset's.
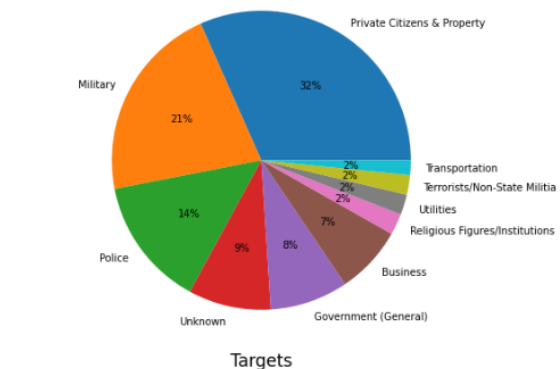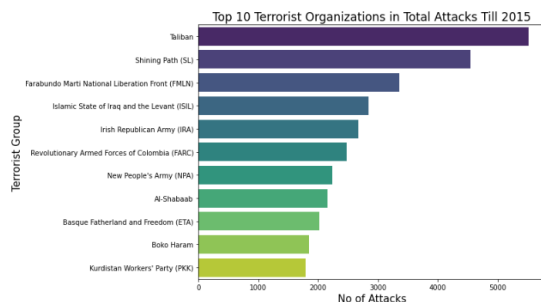
Observation 1

1. Terror attacks were rapidly increasing from the year 2004 to 2015.

2. Last 3 years terror attacks were in decreasing trends

4. In 2015, 'Private Citizens & Property' ,'Military' and 'Police' are the 3 main targets in terror attacks
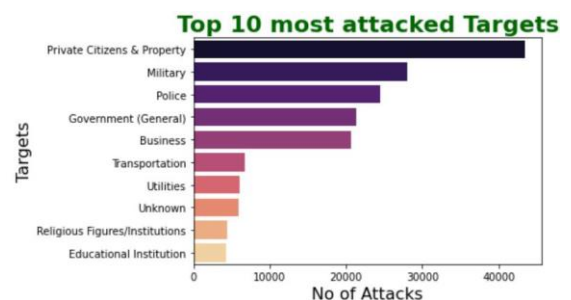
Observation 2

1. In 2015, The most attacked region is 'Middle East & North Africa' then followed by South Asia
2. In 2015, The most attacked Country is 'Iraq' followed by 'Pakistan' and 'Afghanistan'
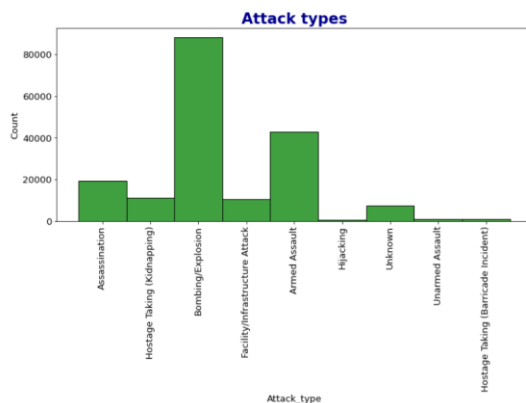3. In 2015, 'ISIL' is the Terrorist Organization with most attacks in the world.

Observation 3

1. The most attacked target is 'Private Citizens & Property' which is approximately 25%

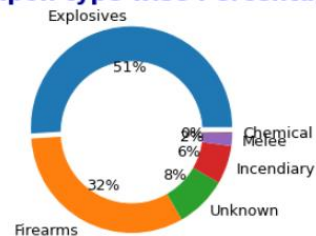2. Military, Police, Government and Business are of 10-20% attacked Targets

Observation 4

Attack types

1. 50% of the Attacks happened with Bombing/Explosions
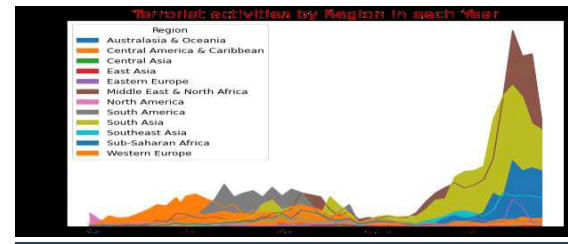2. 10-20% of the Attacks happened with Armed assault and Assassination. Rest

in between 0-6%

Observation 5

Weapon type wise Percentage of Attacks

1. Explosive were used for more then 50% of attacks.
2. 6-30% of the attacks happened with fire arms and incendiary. Rest in between 0-2%
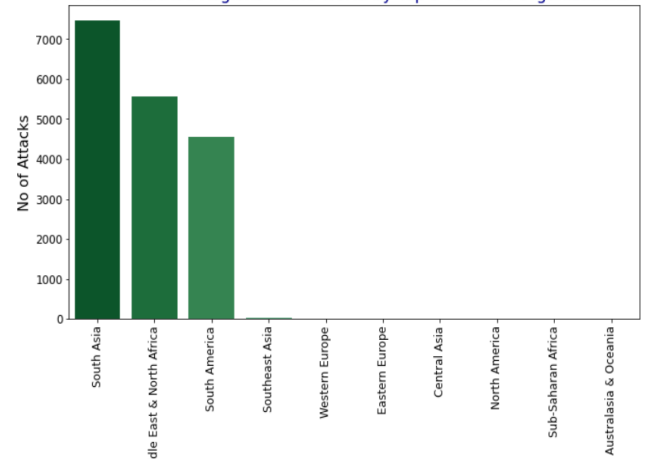
Observation 6

1. Middle East & North Africa is the region with most number of Terror attacks in the world.
2. South Asia is the second most attacked region
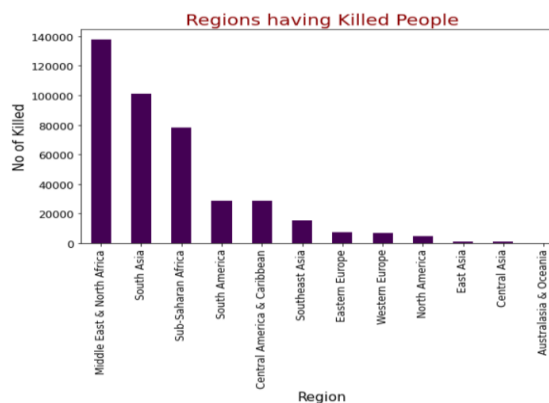
Observation 7

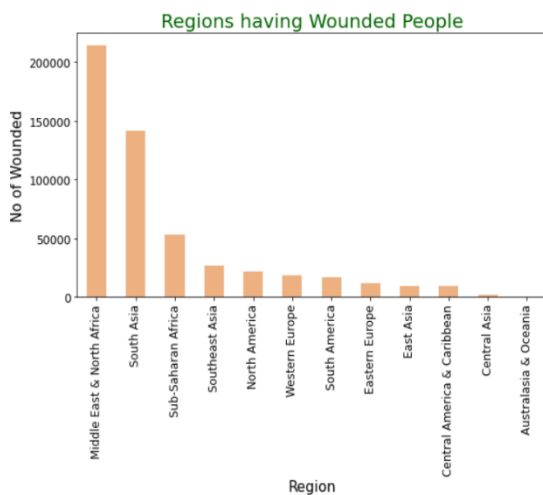Most Attacked Regions in the World by Top 3 Terrorist Organizations

1. These top 3 organizations attacks were increased from the year 2013 to 2015 which can be corelated to the withdrawal of US soldiers from Afghanistan.

Observation 8

1. Middle East & North Africa has the most killed people(>1.4 Lacs) and wounded people (>2 Lacs)
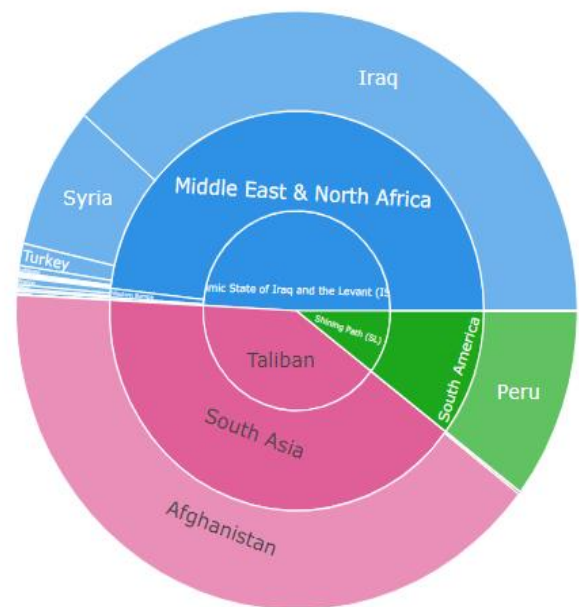

Regions having Killed People

Observation 9

Regions having Wounded People

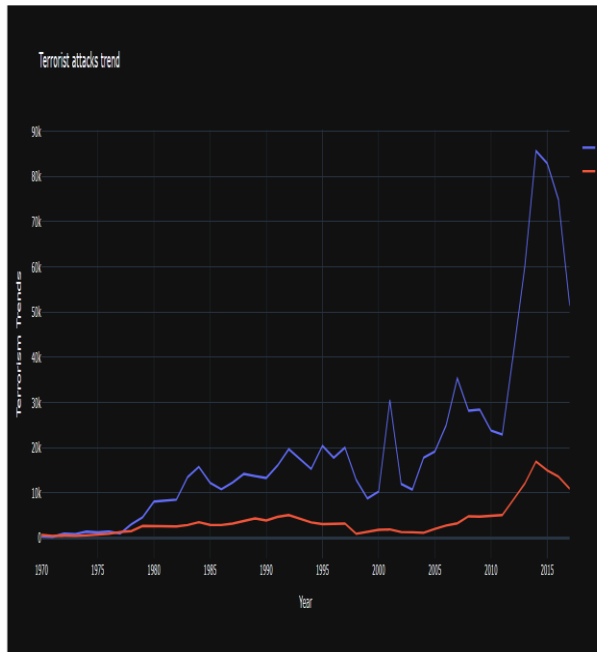2 . Region wise rate of Kills & wounded people are similar to each other.

Observation 10

1. Taliban's were most active in South Asia region and Afghanistan is the country which suffered most attacks by this group.

2. ISIL were most active in Middle East & North Africa region. Iraq, Syria and Turkey are the top 3 attacked countries by them.

Observation 11

1. Year wise no of attacks and victims rate are proportional to each other.

Conclusion

1. Terrorist attacks reached a peak during 2014 and then in 2015, started to drop. The drop in the number of terror attacks is not a sign of improvement in security checks in the society, as the ability to carry out a successful terror attack has not reduced significantly.
2. The middle East & North Africa is the most affected region in terms of the number of terror attacks among all the regions followed by South Asia, South America and Sub-Saharan Africa.
3. Iraq is the most affected nation of all and it has suffered the highest number of attacks followed by Pakistan, Afghanistan and India.
4. Most of the attacks are either bombing or explosions followed by firearms and the most used weapons in the terror attacks are explosives and assault weapons.
5. Taliban and ISIL are the most active terrorist organisations in the world and Afghanistan is the country that has suffered the most because of Taliban's.
6. In 2016, India suffered the highest number of attacks, which was peaked during Aug-2016. Most attacks were happened in Srinagar (Jammu & Kashmir). Maoists were responsible for most of the attacks.
7. Terrorist organisations and attacks are slowly moving to underdeveloped and developing countries with large populations below the poverty level in south Asia and Africa.
8. This creates a cycle of terrorism as poverty and unemployment help radical ideologists lure youth towards terrorism, discouraging investors which in turn effects the earnings of local governments and leads to further poverty and unemployment.

ACKNOWLEDGEMENT

References

1 https://www.geeksforgeeks.org/

2.https://www.almabetter.com/

    (notes)

3. some others reference

4. https://www.python.org.com