GLOBAL TERRORISM ANALYSIS

Arun Gupta, Farhat Jahan, Sharaffin B,
Shrutika Kawade, Yuvaragavan Palanisamy

Data science trainees,

AlmaBetter,Bangalore

ABSTRACT

Terrorism is the threatened or actual use of illegal force and violence by a non-state actor to attain a political ,economic, religious or social goal through fear, coercin or intimidation. It is a threat and is the greatest violator of human rights, it challenges the authority of the government and puts the security and sovereignty of a country in jeopardy.

Our main aim in the project was to explore and visualize Global Terrorism Database (GTD), which is maintained by the University of Maryland. The database has information regarding 1,80,000 terrorist events around the world for the years 1970-2017 with detailed attributes like number of deaths, damage, terrorist groups involved, location etc.

.

1. PROBLEM STATEMENT

The GTD contained various columns iyear, imonth, iday, country_txt, region_txt, attacktype1_txt, target1, nkill, nwound, gname, etc. which helped us draw major insights fromthe data set. Our aim here is to understandthe important factors that governs the terrorist attacks in regions

2. INTRODUCTION

Emphasis is given on the information in GTD relating to the time component of the data across the world, region wise or country wise and is visualized into various types of graphs. The location and the details of the terrorist events occurring in a particular year can be seen on graphs. We have done an overall analysis from the data related to attacks. Information related to terrorist groups with highest damage in history, their active years, worst events in history, most affected countries. Other attributes like weapon type, target type, attack type was explored and their distributions are analyzed.

3. DATA SUMMARY

The provided data set has following different columns of over 130 variables. Few of them are given below

a) Attacktype1_txt

The type of attack happened. Attacktype1_txt consists of categories like explosion, armed assault, assassination, kidnapping, unarmed assaults.

b) Target1_txt

Type of target involved in the attack. Target1_txt consists of categorical values like private citizens, military, police, government officials, transportation, education, religious institution, airports, etc.

c) Success

1'if attack was a success. '0'if attack was a failure

d) Region_txt

Name of the region where the attack happened. Region_txt consists values like East Asia, South Asia, Western Europe, etc

e) Day,month,year

Calendar details of the event

f) Nkill

Number of people killed in any event

g) Nwounded

Number of people wounded in any event

h) Property

Total property damage happened in any event

4. STEPS INVOLVED

4.1. Connection with the Data

Here, we are provided with a dataset with more than 180,000 terrorist attacks.

Thus, the dataset is actually a collection of more than 180,000 experiences about worldwide terrorist attacks.

Now, we need to decode the set of experiences to help in building solid future strategies to fight against the terrorist gangs.

At first, we import the libraries or functions for making our journey easy and then get connected to the set of experiences.

4.2. DeeperUnderstanding of the Data

As there is a huge no. of experiences, we take the help of statistics to measure each and every feature in different dimensions and thus step by step, we find the most important features or the exact way to decode the experiences.

"what gets measured gets done

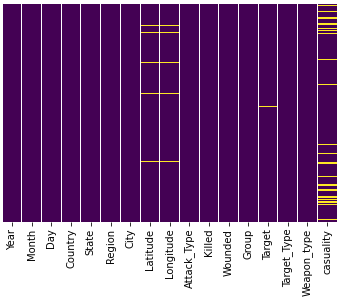
4.3. Cleaning the Data

Now we can create a clean experience set with important features.

In this step, we treat the missing values or "?" values (fig – heatmap for Null values after cleaning)

```
#Heatmap for Null values check after cleaned database
sns.heatmap(dt.isnull() , yticklabels=False , cbar=False , cmap='viridis')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fe102a35a90>
```



4.4. Visualizing the Data

When we know all the important features of our experiences, we can go a step ahead by finding the relationship among features. Here, we take the help of visualization because

"A picture is worth a thousand words"

Technologies Used

Python, Tableau,

Python packages :

Mathplotlib, Pandas, Numpy

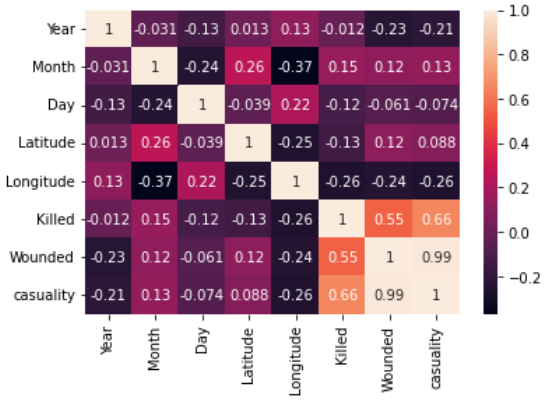
5. EXPLORATORY DATA ANALYSIS

EDA was done to gain understanding on the following Questions and visualized them into various types of graphs

a) Heatmap for the correlation between the values in the dataset

.

```
#Heat map for the correalation between the values
corrMatrix = dt.corr()
sns.heatmap(corrMatrix, annot=True)
plt.show()
```



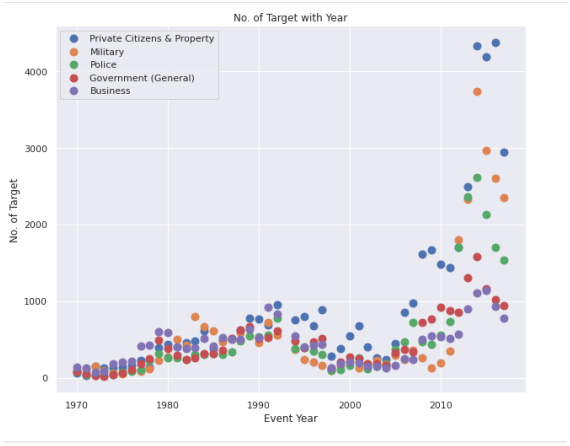
b) The top countries affected by terrorism

Iraq, Pakistan and Afaganistan were the regions badly hit by terrorism

c) The top 30 Worst terrorist attacks with country, year and casualities

d) The top targets attacked

Private citizens and Property were the most preferred targets



e) The most commonly used weapons

Explosives and firearms were used a lot by the terrorist outfits

f) The No of attacks happened with respect to the Years

g) The most attacked Regions in the form of a pie chart with its percentage

h) boxplot visualizing the region of attack and the timeline

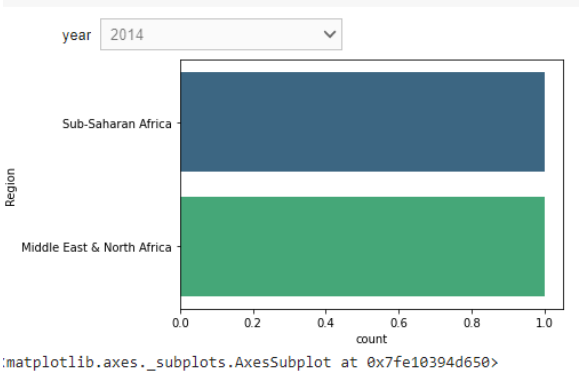
i) The top Terrorist organisations with Highest terror attacks

j) Al Qaida the organisation had more number of kills

k) Terror Attacks by Taliban being the largest terror group

l) The number of people killed by Taliban over the years

m) To check which region was affected by terrorism each year



n) **A bivariate analysis of people killed and wounded in the regions of Terror**
o) **Plotting number count for kill and wound with line**
p) **Word cloud map for the names of the states affected by terrorism**

6. OBSERVATIONS

After performing the Exploratory Data Analysis, we get the following insights from the data:

- Private Citizens and Property were attacked most followed by Military,Police, Government and so on.
- Middle East & North Africa was most affected among the top affected region.
- Iraq was the country which was most affected by terror attacks and had maximum number of killed and wounded people.
- Most of the people in this region were either wounded or killed.
- The State and City that was most affected was Baghdad
- In the last decade 2014 had the greatest number of terror attacks. That was around 16500 attacks during this year which means on average 45 attacks per day
- The most common attack type was Bombing/Explosion.
- Explosives have been consistently the most popular weapon of choice for terrorists.
- Taliban became more active since 2012 and they are responsible for the most of the terror attack

7) CONCLUSION

The goal of this project was to understand andinterpret the nature of terrorism. We perceived the START dataset through visualdesigns. A visualization which can be used to calculate the total number of attacks, totalkillcountsandlocationbasedontheselected regionandyearprovidesinteractiveinterfaceto explore this dataset. We understood various patterns, trends and correlation interrorismthroughvisualinterpretationanditsprovidedexplanation.This work can be used by curious civilians, security relatedpolicy-makers,internationalorganizationshostingworldwideevents,foreigninvestorsandacademicresearchersforthepurposeofunderstanding terrorismand itsnature

8. RECOMMENDATIONS:

Since Private Citizens and Property are being targeted consistently so stronger security and surveillance should be provided, especially in the dense populated regions. More surveillance is required especially in the Middle East & North African Regions. Strict border policy should be implemented to prevent the movement of explosives between the regions.

References-

1. MachineLearningMastery
2. GeeksforGeeks
3. Stackoverflow