

Capstone Project

Global Terrorism Analysis

Team Members

Arun Gupta

Farhat Jahan

Sharaffin B

Shrutika Kawade

Yuvaragavan Palanisamy

Points to Discuss

- Problem Statement
- Abstract
- Data Summary
- Data Pipeline
- Exploratory Data Analysis
- Summary
- Conclusion
- Recommendation

Problem Statement

Explore and analyze the data to discover key findings pertaining to terrorist activities.

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, coercion or intimidation

It is a threat and is the greatest violator of human rights, it challenges the authority of the govt and puts the security and sovereignty of a country in jeopardy.

Our main aim in the project was to explore 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.

Emphasis is given on the **information 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 analysed.**

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

Data Pipeline

Data Processing: In the first part, we have imported necessary libraries and data set. We then used the libraries to understand the data.

Data Cleaning: After understanding the data, we got to know about the null values and discrepancies in the data and the dataset is cleaned.

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

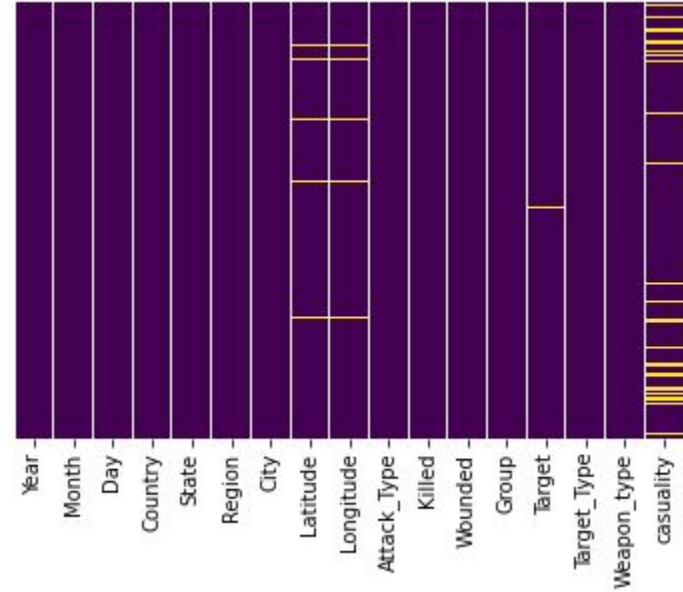


Fig: HeatMap for null values check after cleaning dataset

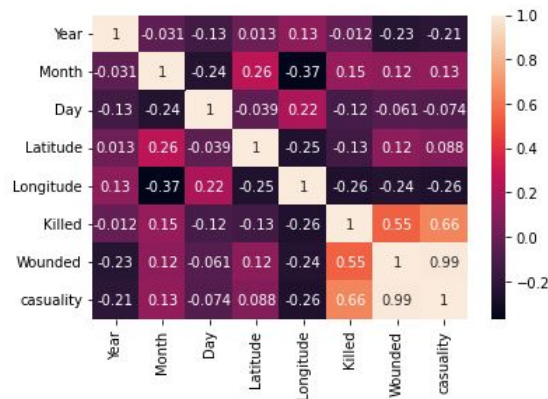
Exploratory Data Analysis

After Cleaning the data set, we did some exploratory data analysis using tables and graphs to derive the observations from the data and get the solution to the problem statement.

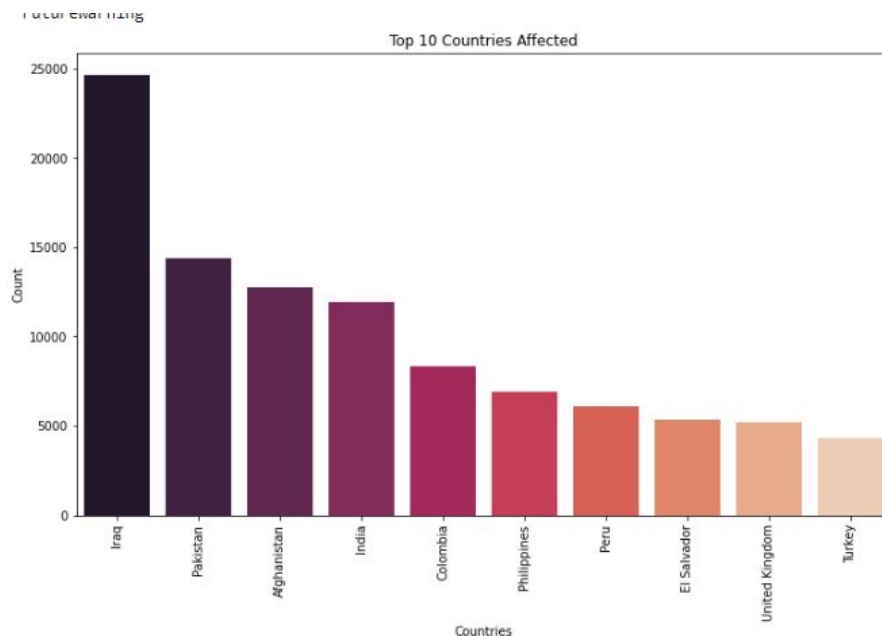
Exploratory Data Analysis

1. Heatmap for the correlation between the values in the dataset

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



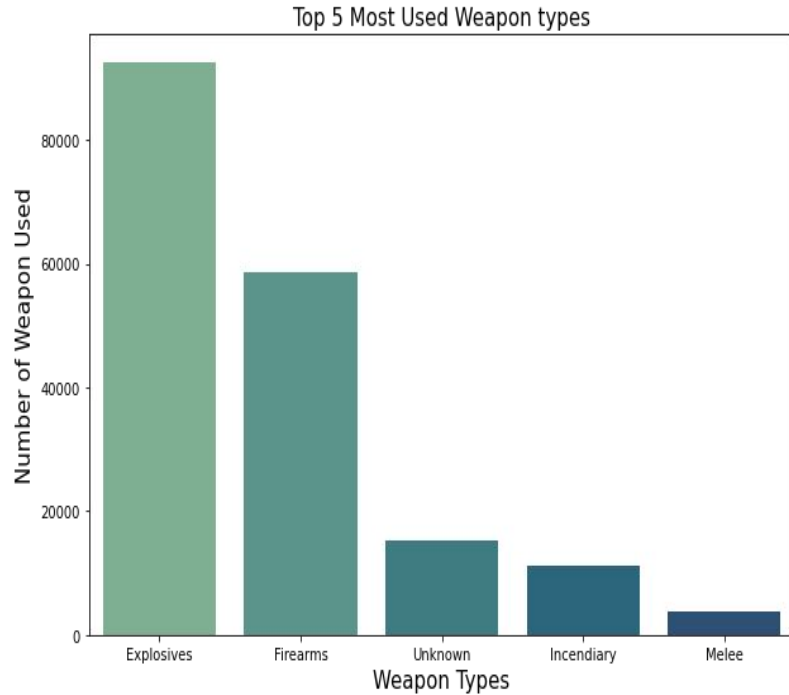
2. The top countries affected by terrorism



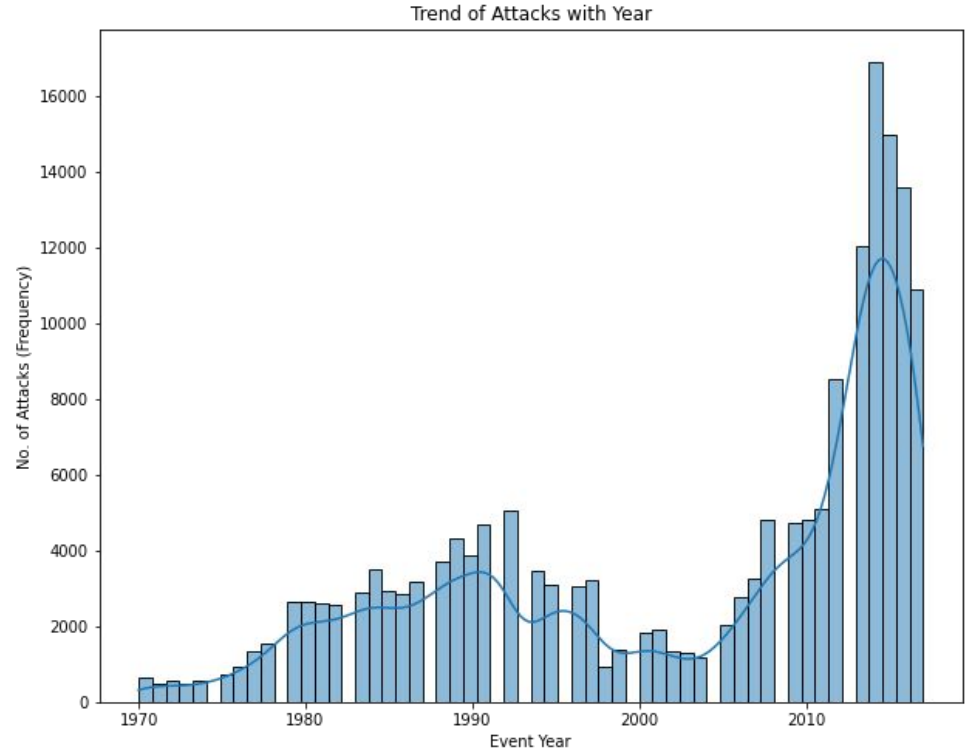
3. Scatter plot to visualize the top targets which are attacked by terrorists



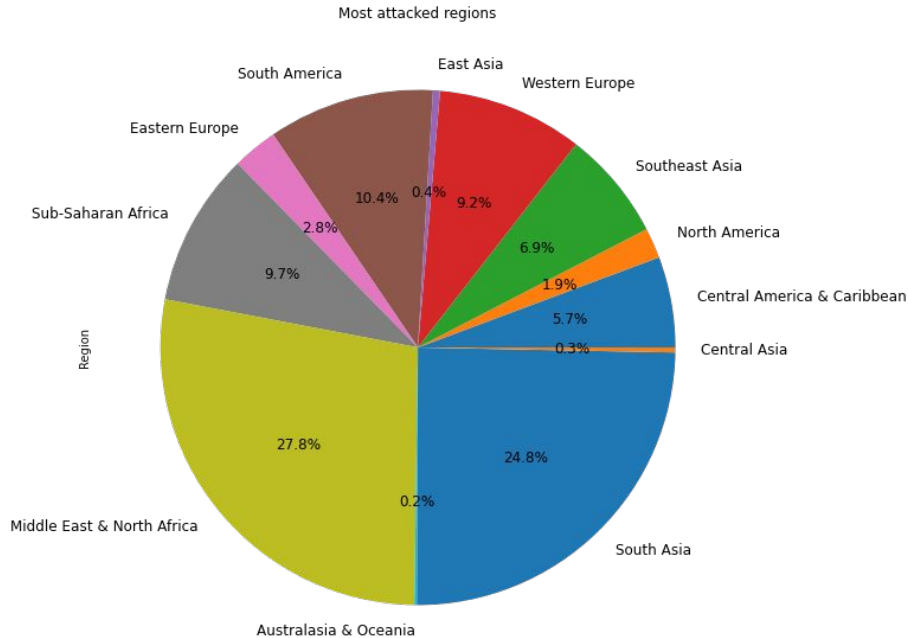
4. The most commonly used weapons



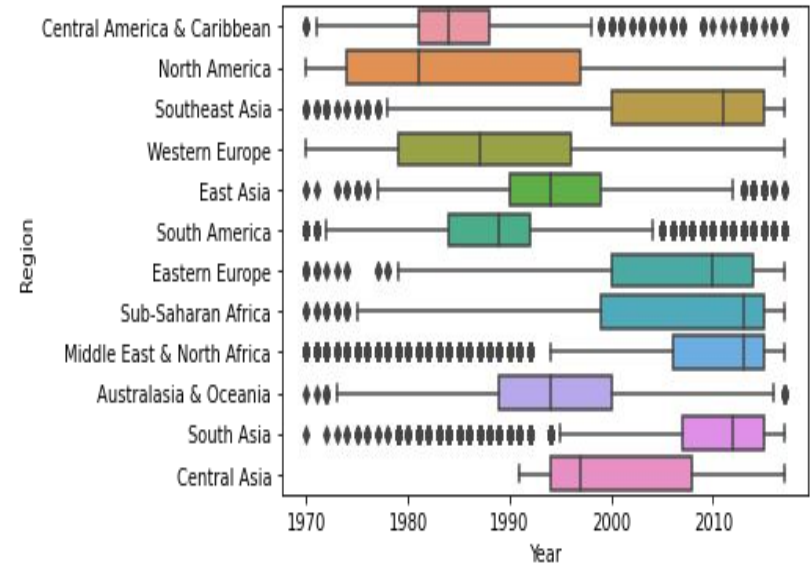
5. The No of attacks happened with respect to the Years



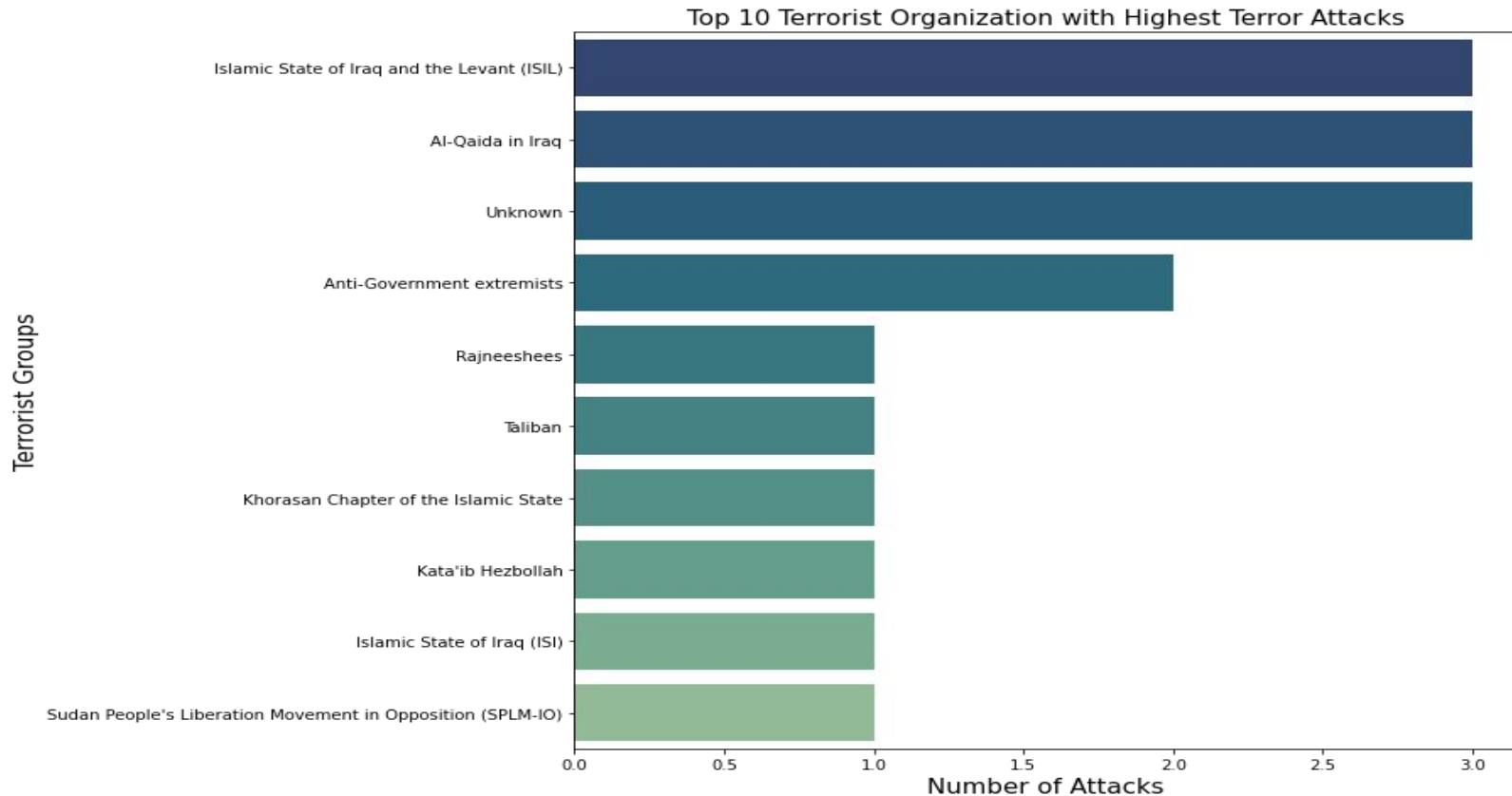
6. The most attacked Regions in the form of a pie chart with its percentage



7. Boxplot visualizing the region of attack and the timeline



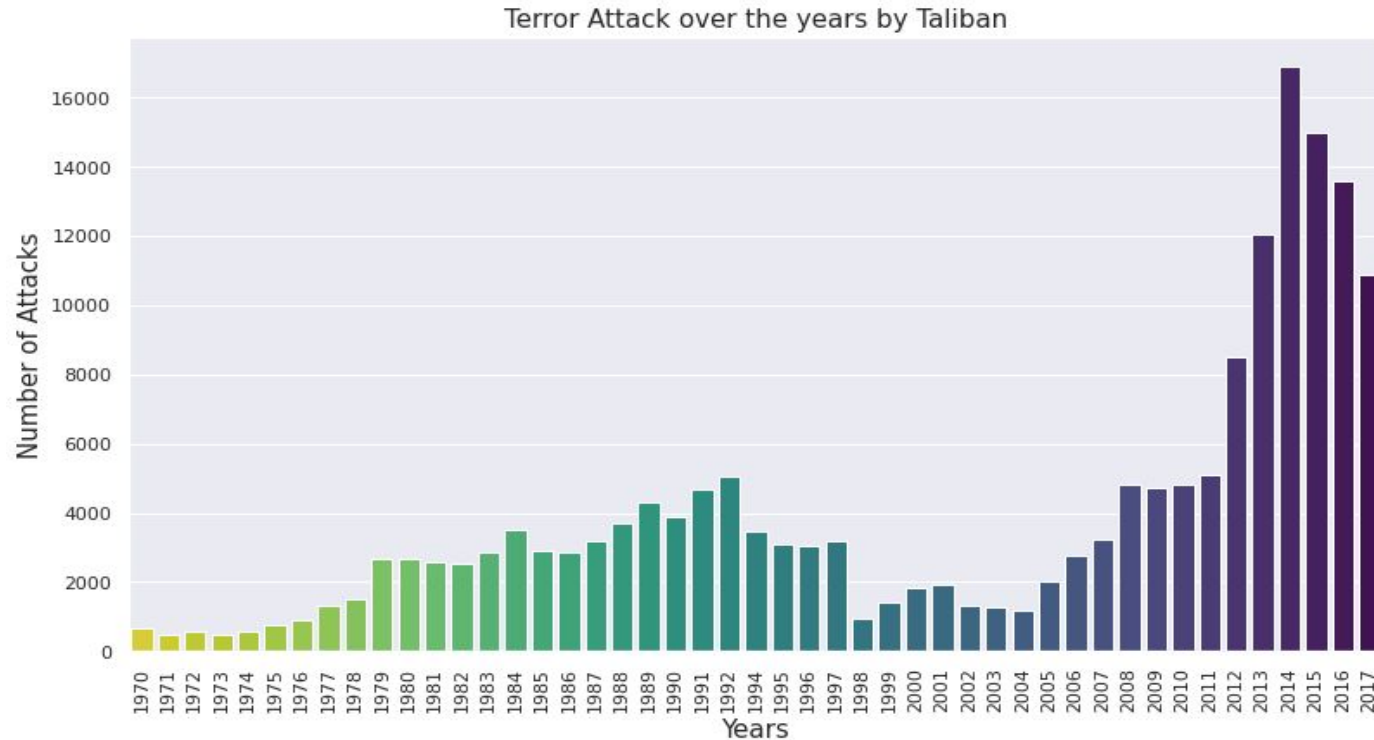
8. The top Terrorist organisations with Highest terror attacks



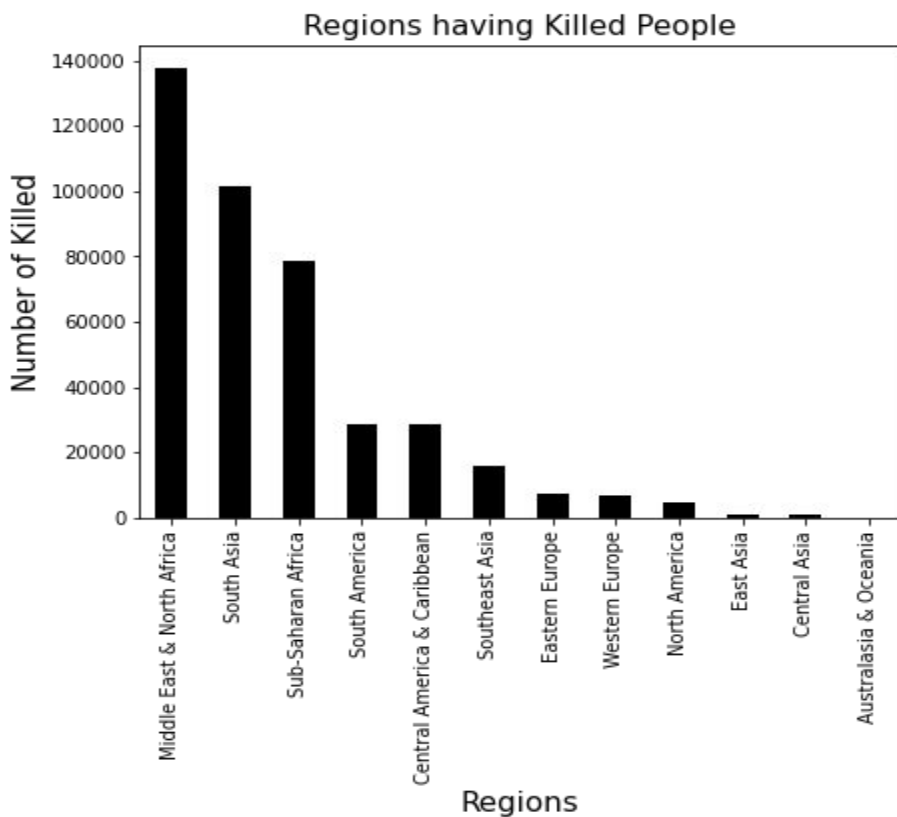
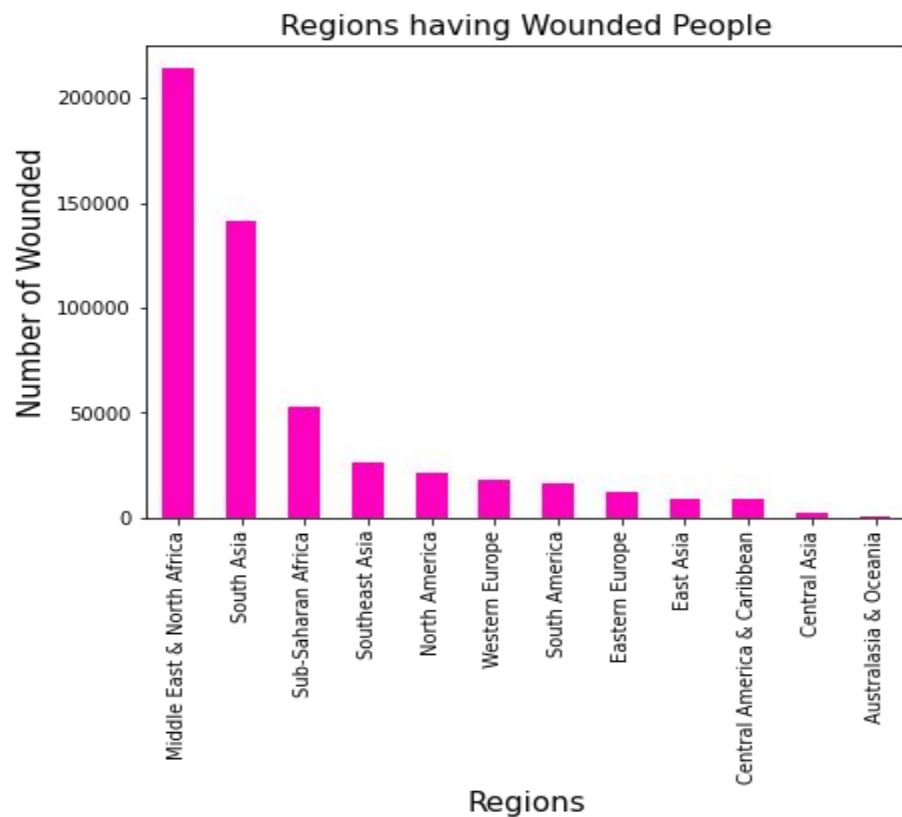
9. Al Qaida the organisation had more number of kills

	Year	Month	Day	Country	State	Region	City	Latitude	Longitude	Attack_Type	Killed	Wounded	Group	Target	Target_Type	Weapon_type	casualty
73127	2001	9	11	United States	New York	North America	New York City	40.697132	-73.931351	Hijacking	1383	8191	Al-Qaida	Passengers and crew members on United Airlines...	Private Citizens & Property	Vehicle (not to include vehicle-borne explosiv...	9574.0
73126	2001	9	11	United States	New York	North America	New York City	40.697132	-73.931351	Hijacking	1384	8190	Al-Qaida	Passengers and crew members on American Airlin...	Private Citizens & Property	Vehicle (not to include vehicle-borne explosiv...	9574.0
68071	1998	8	7	Kenya	Nairobi	Sub-Saharan Africa	Nairobi	-1.285180	36.821107	Bombing/Explosion	224	4000	Al-Qaida	United States Embassy in Nairobi, Kenya	Government (Diplomatic)	Explosives	4224.0
76326	2004	3	11	Spain	Madrid	Western Europe	Madrid	40.465595	-3.696263	Bombing/Explosion	73	450	Al-Qaida	Indiscriminate civilians on commuter trains in...	Transportation	Explosives	523.0

10. Terror Attacks by Taliban being the largest terror group



11. A bivariate analysis of people killed and wounded in the regions of Terror





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
- Al_Qaida Most mischievous gang in early 2000's and have done most number kill till date and the weapon used for killing is only two Explosives and vehicle-borne explosive.

Conclusion

The goal of this project was to understand and interpret the nature of terrorism. We perceived the START dataset through visual designs. A visualization which can be used to calculate the total number of attacks, Total kill counts and location based on the selected region and year provides interactive interface to explore this dataset. We understood various patterns, trends and correlation in terrorism through visual interpretation and its provided explanation. This work can be used by curious civilians, security related policy-makers ,international organizations hosting world wide events, foreign investors and academic researchers for the purpose of understanding terrorism and its nature.

Recommendation

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.