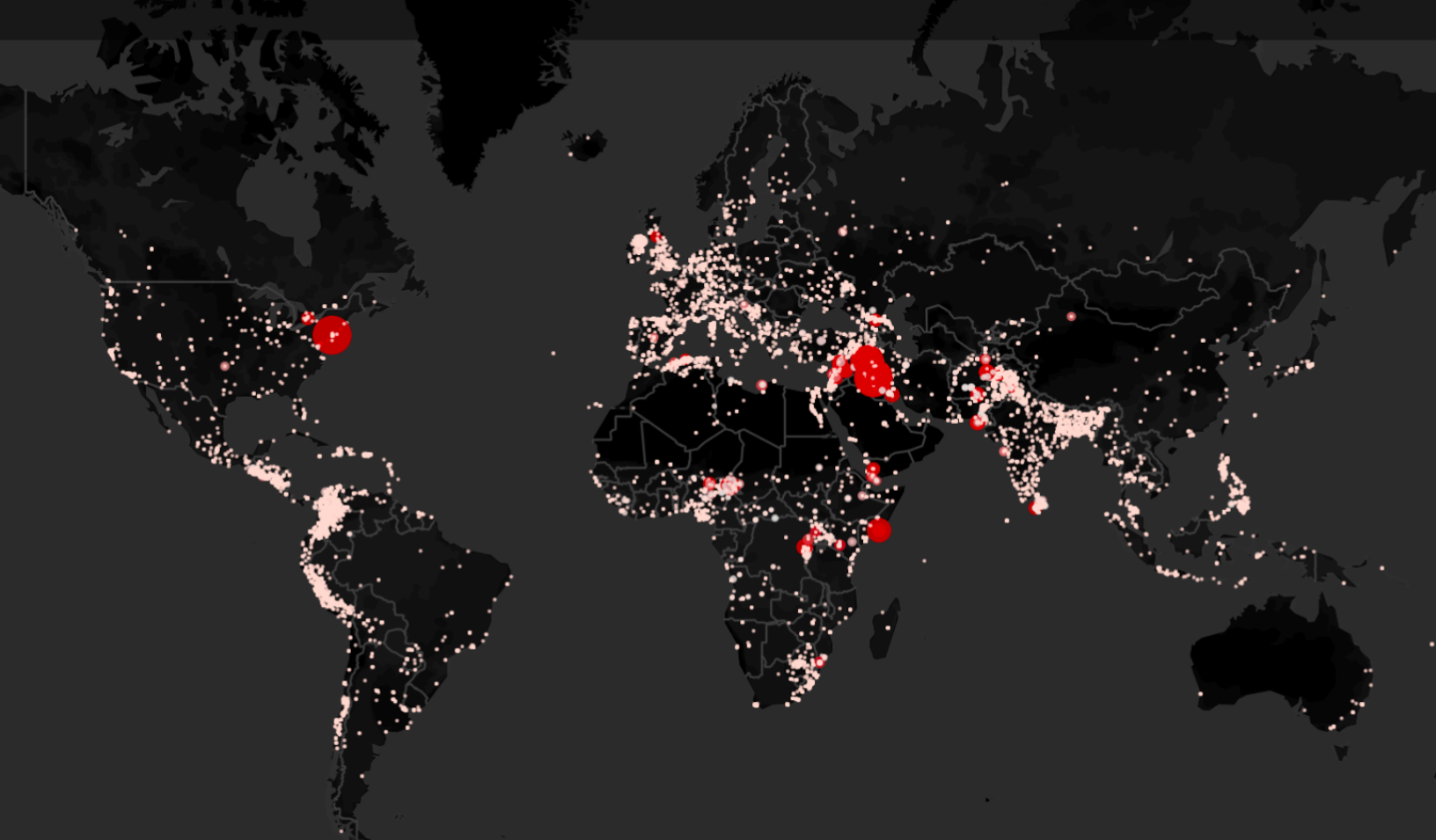


MIS 6380 | SPRING 2021 | GROUP 4

GLOBAL TERRORISM

DATA VISUALIZATION GROUP PROJECT



AKSHATA BHANDIWAD

SHASWATH TAMBAD

JASH KOTHARI

SAHIRA SIAG

SIDDARTH RAO ANNAMANENI

SAI MANISHA MIKKILINENI

Table of Contents

1.	<i>Executive Summary</i>	1
2.	<i>General Introduction:</i>	2
2.1.	Objective	2
2.2.	Hypothesis	2
2.3.	Data Set	2
3.	<i>Data Description :</i>	3
4.	<i>Data Cleaning :</i>	5
5.	<i>Insights and findings:</i>	6
6.	<i>Conclusion</i>	12

1. Executive Summary

Terrorism is a global threat in which thousands of people lose their lives. As there is an increase in such acts around the world, finding the root cause and the places where such acts are happening the most will help in achieving global peace and safety of the people. Through this project, we aim to create an interactive platform to spread the awareness.

The dataset used is Global Terrorism Database, raw dataset had 135 columns with a lot of data points missing. After cleaning the dataset and clearing out empty columns using R, 24 columns were used for our analysis.

The main objective of the study was to analyze the different terrorism events to test our hypothesis. This helped us understand and get insights on major causes of terrorism, accurate patterns, categories of the targets identified, locations of terrorist attacks.

Through our analysis we found out the following:

1. The number of deaths in Middle East and North Africa region was greater than the South Asian region.
2. Domestic Attacks are higher than the international attack
3. Terrorist attack in the United States have reduced over year due to strict checking at the port of entry

2. General Introduction:

In this visualization, we try to understand how the number of terrorist acts varies around the world and how it has changed over time. To do this, we need a clear and consistent definition of what terrorism is, and how it is different from any other form of violence. This is not straightforward.

Terrorism is defined in the Oxford Dictionary as “the unlawful use of violence and intimidation, especially against civilians, in the pursuit of political aims.” We quickly see that this definition is unspecific and subjective. The issue of subjectivity in this case means that there is no internationally recognized legal definition of terrorism.

The problem is that terrorism is difficult to distinguish from other forms of political violence and violent crime, such as state-based armed conflict, non-state conflict, one-sided violence, hate crime, and homicide. The lines between these different forms of violence are often blurry

In our visualization of global terrorism, we rely strongly on data from the Global Terrorism Database (GTD), which defines terrorism as “acts of violence by non-state actors, perpetrated against civilian populations, intended to cause fear, in order to achieve a political objective.” Its definition excludes violence initiated by governments (state terrorism) and open combat between opposing armed forces, even if they are non-state actors.

2.1. Objective

The main objective of this study is to analyze the different terrorism events that have taken place worldwide, at different times. This will help us understand and get insights on major causes of terrorism, accurate patterns, categories of the targets identified, locations of terrorist attacks. We want to know the answers to questions like Was there an increase or decrease in the number of attacks? Which regions had the most attacks? What were the most targeted region etc.

2.2. Hypothesis

- a. The deaths in Middle-east region are higher than Western countries.
- b. The internationally motivated attacks are highest.
- c. The Global trend matches with trend in the United States.

2.3. Data Set

The selected dataset contains information about each of the terrorist attacks worldwide from 1970 – 2018 with details of exact location, country name, attack type, target location and target of the attack.

3. Data Description :

The raw data set has 135 columns and 191,465 records and can be accessed from Global Terrorism Database: <https://www.kaggle.com/START-UMD/gtd>

The following table describes the variables considered for this project.

	VARIABLE NAME	DESCRIPTION
1.	eventid	Incidents from this file follow a 12-digit Event- ID system. <ul style="list-style-type: none"> First 8 numbers - date recorded “yyyymmdd”. Last 4 numbers – sequential case number for the given day.
2.	Date	This variable gives the exact year, day and month of the incident occurred.
3.	Country	This variable identifies the country where the incident occurred. In the case where the country in which an incident occurred is unidentified it is marked as “unknown”
4.	Region	This variable identifies the region where the incident occurred.
5.	provstate	This variable records the name (at the time of event) of the 1 st order subnational administrative region in which the event occurs.
6.	City	This variable contains the name of the city, village, or town in which the incident occurred.
7.	latitude	This variable records the latitude of the city in which the event occurred.
8.	longitude	This variable records the longitude of the city in which the event occurred.
9.	multiple	This variable gives cases where several attacks are connected, then “Yes” is selected to denote that particular attack was part of “multiple” incidents. 1 = Yes 0 = No
10.	success	1 = Yes - The incident was successful 0 = No - The incident was not successful
11.	suicide	This variable is coded “Yes” in those cases where there is evidence that the perpetrator did not intend to escape from attack alive. 1 = Yes; 0 = No
12.	Attack Type	This variable captures the general method of attack and often reflects the broad class of tactics used.
13.	Target Type	This variable captures the general type of target/victims.

14.	Prep.Group.Name	This variable contains the name of the group that carried out the attack. If no information about the group is available, then it is marked as "unknown"
15.	Claimed	This variable is used to indicate whether a group or person claimed responsibility for the attack. 1 = "Yes" A group or person claimed responsibility for the attack 0 = "No" No claim of responsibility was made. -9 = "Unknown"
16.	Claimed mode	This variable records the modes used by claimants to claim responsibility. When information about mode of the claim is not available it is marked as "Unknown".
17.	Weapon Type	This variable gives information about the weapons used in an attack.
18.	nkill	This variable contains the number of total confirmed fatalities for the incident. This includes all victims and attackers who died as a direct result of the incident.
19.	nkillus	This variable records the number of U.S. victims and U.S. perpetrators who died because of the incident.
20.	nwound	This variable records the number of confirmed non-fatal injuries to both perpetrators and victims.
21.	nwoundus	This variable records the number of confirmed non-fatal injuries to U.S. victims and perpetrators.
22.	Property Damage	This variable is used to list specific details about the property that was damaged in an attack. When the information about the damage not available it is marked as "Unknown".
23.	Prop value	The amount (US Dollars) of property damage is listed in this variable. It is marked as "0" where the information about the damage is "Unknown".
24.	International	This variable is based on the comparison between the nationality of the perpetrator group and location of the attack. 1 = "Yes" The attack was international. 0 = "No" The attack was domestic.

4. Data Cleaning :

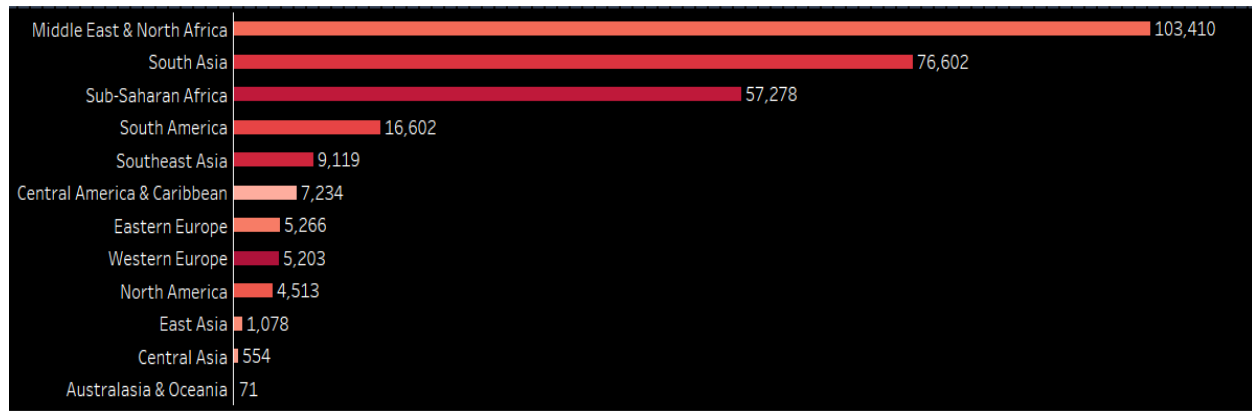
Data cleaning and manipulation is done in order to address the anomalies within the data set. The data cleaning process was carried out 'R' software. Following steps were followed:

- The raw data had nearly 135 attributes to define a particular terrorist event (record). The attributes required for the project were shortlisted.
- Each column was carefully analyzed of the missing records.
- Information of about the date of the terrorist attack was captured in four different columns in the raw data. Only one column 'Date' was retained and the other redundant columns were dropped.
- The columns with country code, region code, state code and city code were also dropped. This exact information has been captured by Country, Region, State and City columns.
- The records with missing location attributes – latitude and longitude were dropped.
- All the information about multiple attacks has been captured by the 'Multiple Attack' column. The other columns with details of the incidents in case of multiple attacks were also dropped.
- Only the Main Attack type, Weapon used, Target type and the terrorist group involved for the first attack has been considered for this study.
- 78% of the records in the property value column of the raw data set was missing. But since this would be important to understand the impact of terrorist attack, data imputation was carried out. Median value of every 'Property Damage' group has been imputed for the respective missing property value.
- The columns with more than 90% of missing records were dropped for this project.

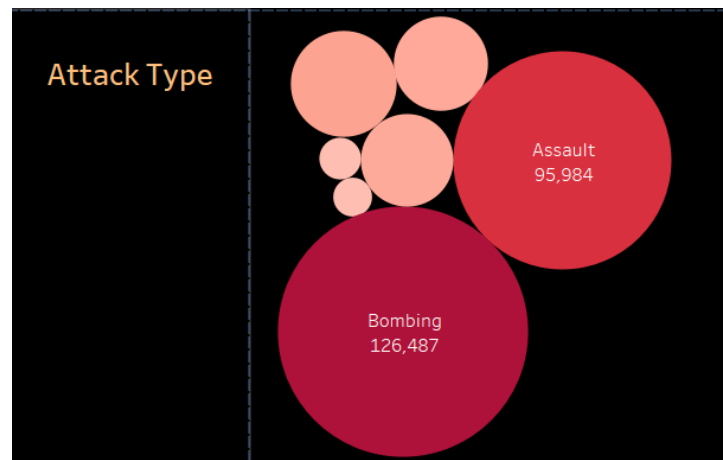
After the data preprocessing, the cleaned dataset was taken up for this project visualization using tableau software.

5. Insights and findings:

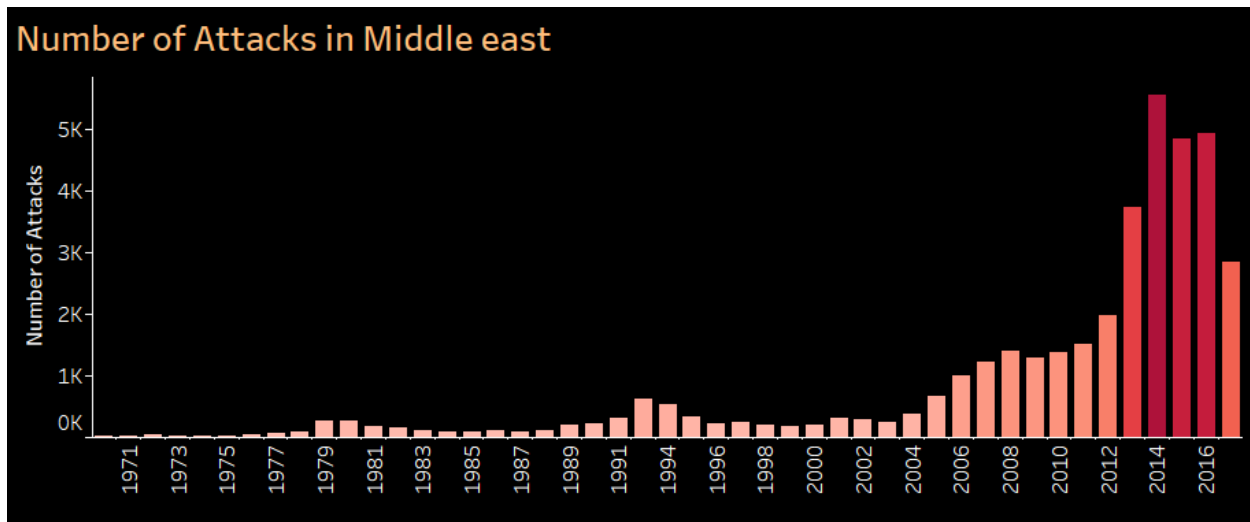
1. Middle East and North Africa region has the greatest number of people killed in terrorist attacks from 1970 to 2017. Although the number of attacks in this time range were similar in Middle East and South Asian regions, more number of people were killed in Middle East.



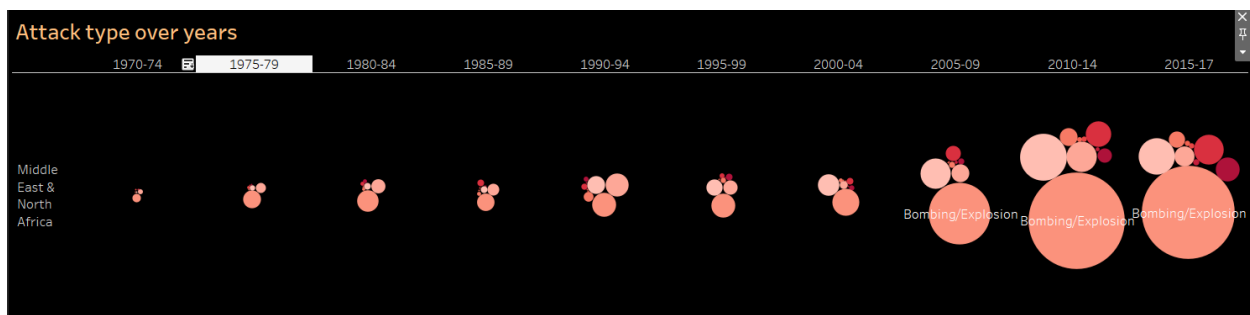
2. Bombings have killed the most number of people over the world, followed by Assault.



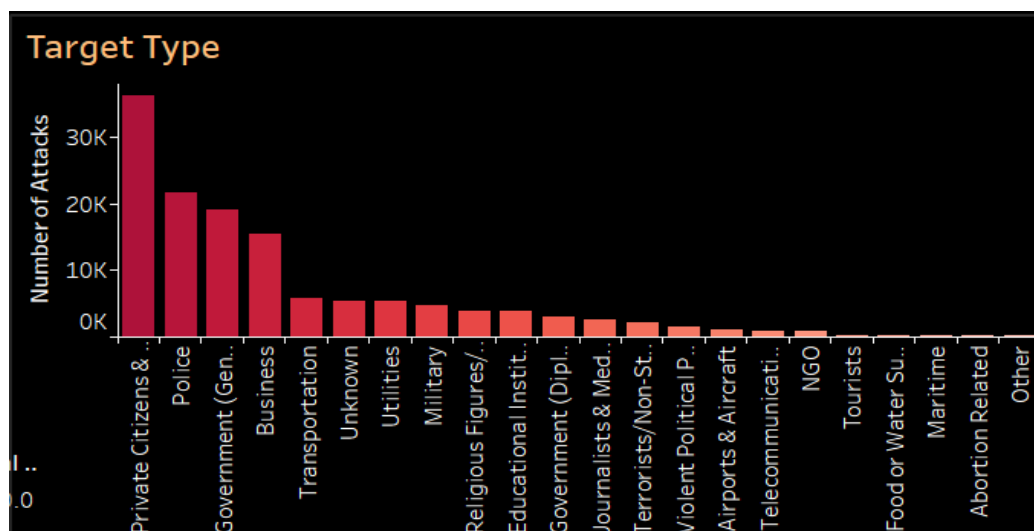
- Number of deaths drastically increased in the Middle East region after 2003 due to increase in number of terrorist attacks and bombings during wars.



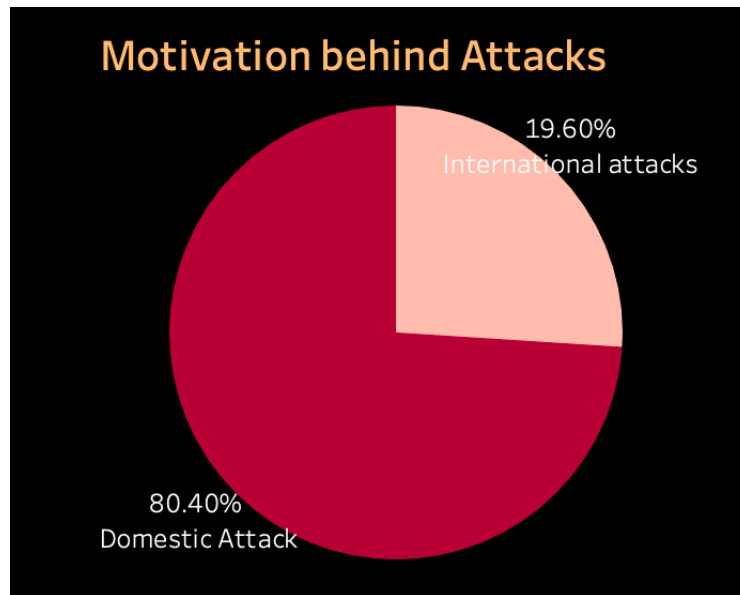
- Bombings is the most common attack type in the Middle East, as well as other regions, since 2005.



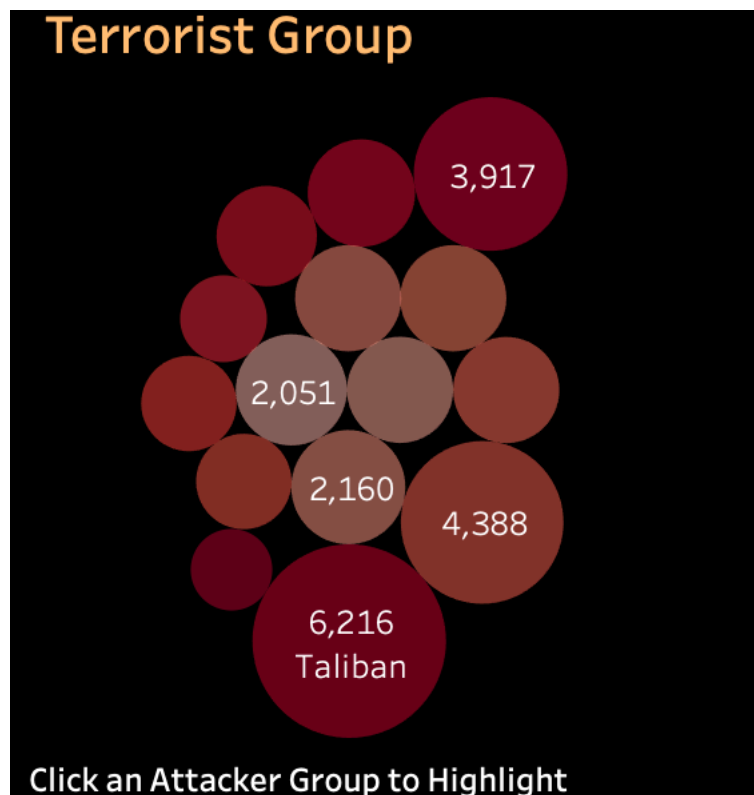
- The most common target type of the terrorist attacks over the years has been Private Citizens & Property, followed by Police.



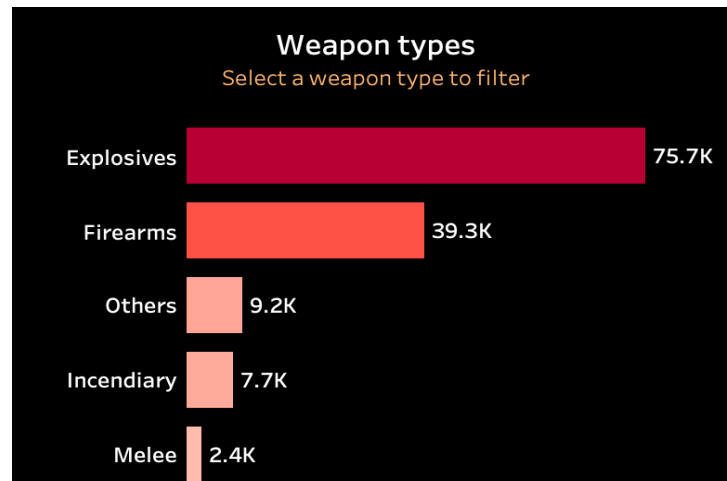
- Domestic attacks, at approximately 80.4%, are more than internationally motivated attacks worldwide.



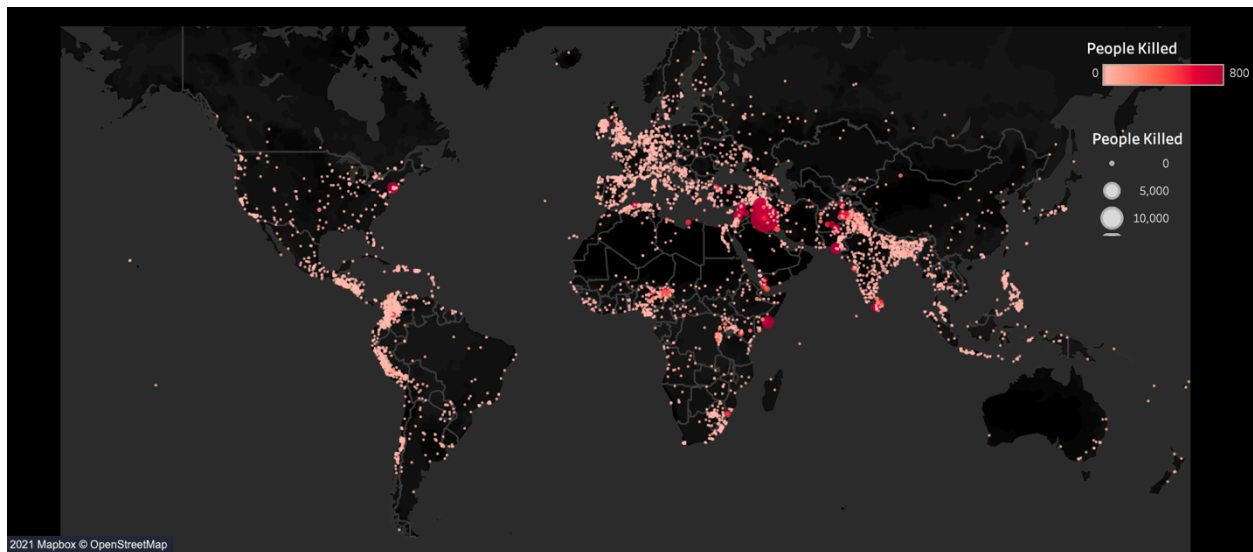
- Taliban has been the most active terrorist group with 6216 total number of attacks. This is followed by ISIL with 4388 attacks over the years.



8. Globally and in the United States, most common weapon type used are Explosives.

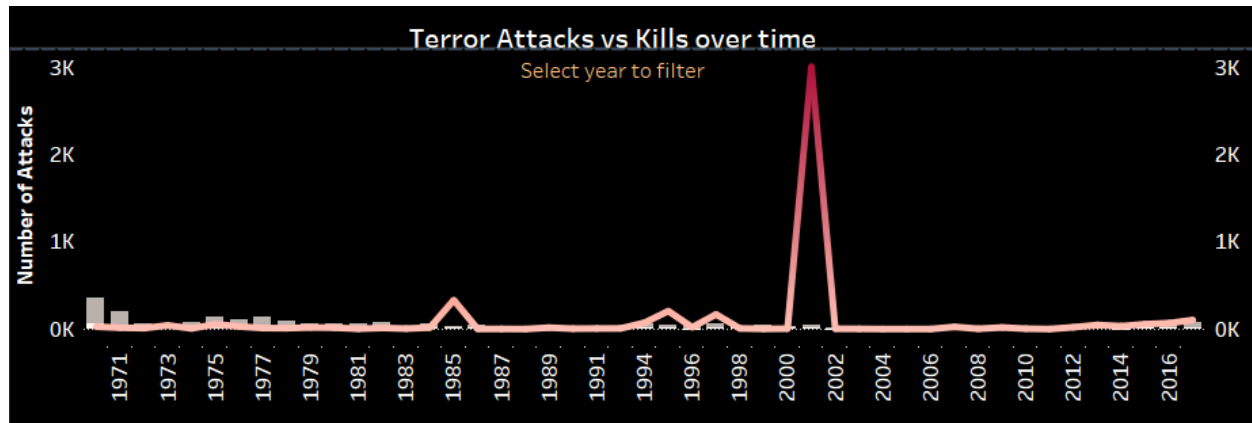


9. One of the catastrophic attacks where more than 800 people were killed is in the United States.

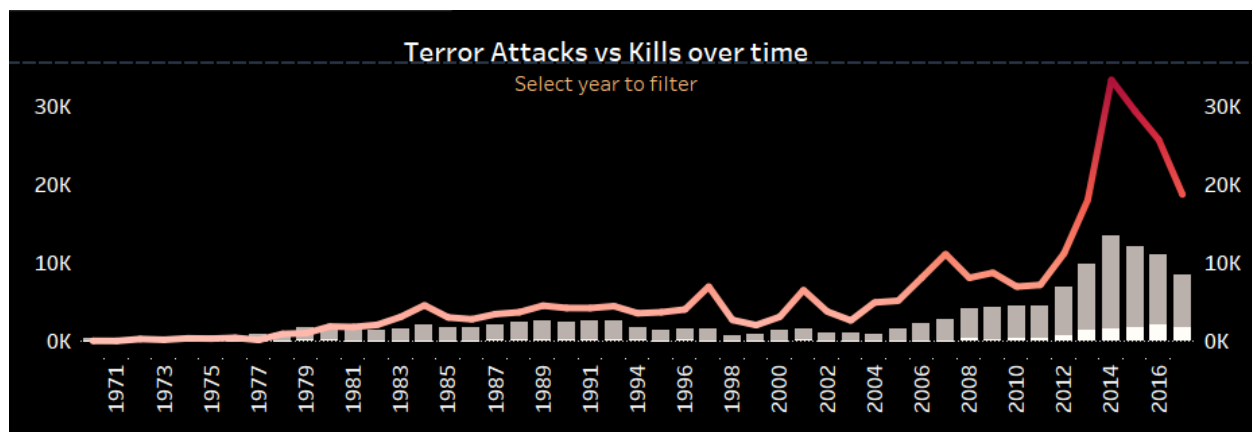


10. Number of terrorist attacks have decreased over the years in the United States unlike the global trend. There is a spike in the number of people killed in the USA in 2001. This is because of the 9/11 attack.

Attacks and kills for USA:

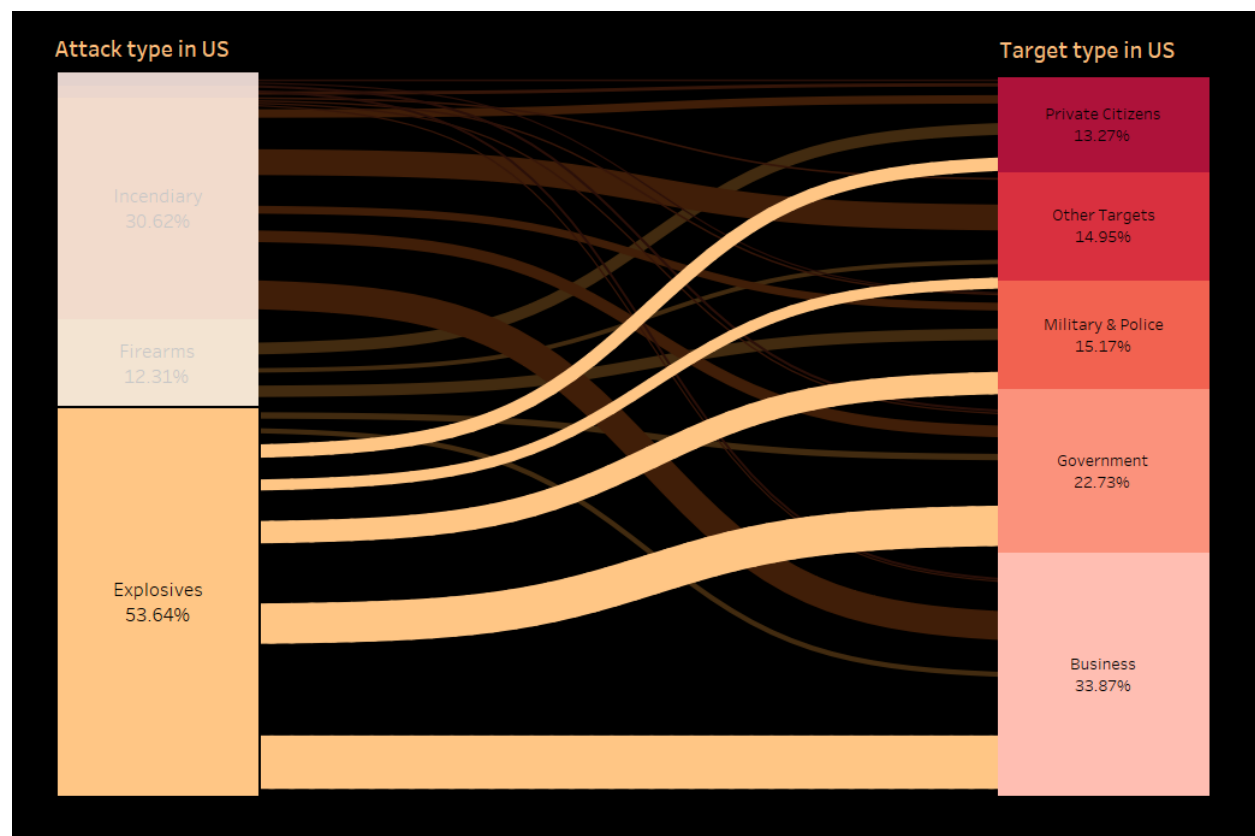


Attacks and kills worldwide:



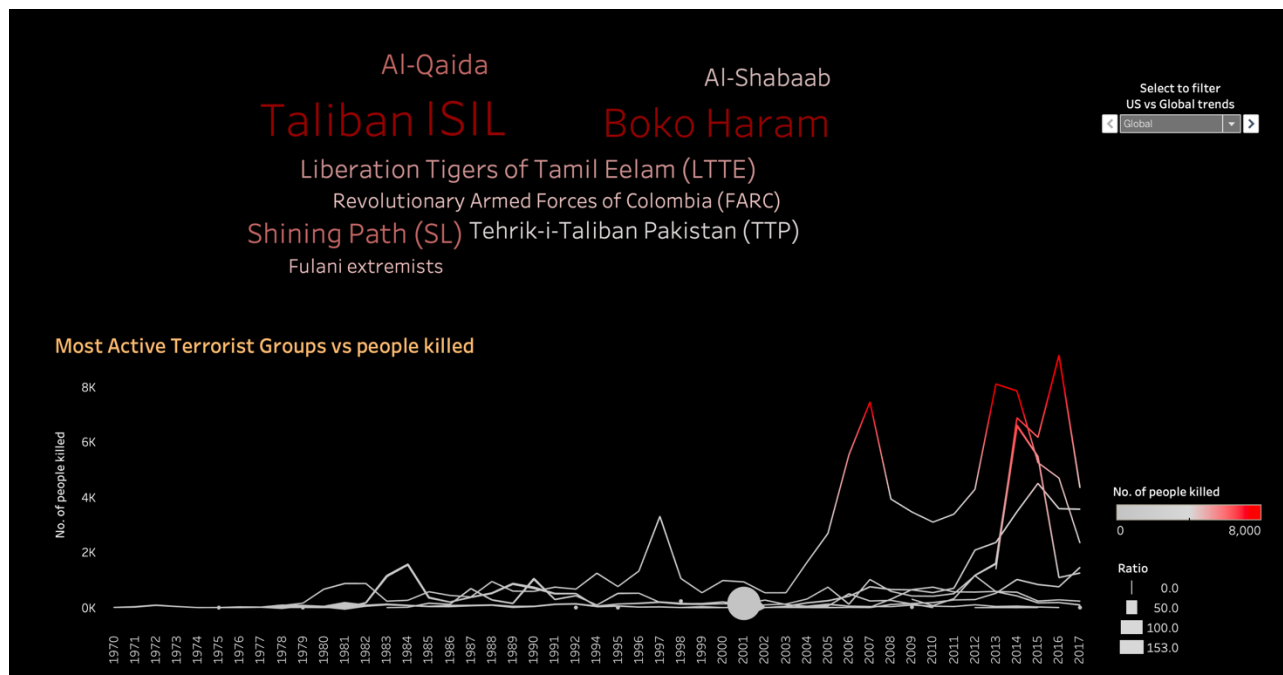
11. Globally, Private citizens and Police are the common target types. While in the United States, Business and Government are the most common target type attacked using Explosives.

Target types	
Select a target type to filter	
Private Citizens & Property	36,235
Military & Police	26,409
Government	22,073
Business	15,531
Others	10,844
Transportation	6,816
Utilities	6,624
Religious	3,961
Educational Institution	3,875
Journalists & Media	2,472

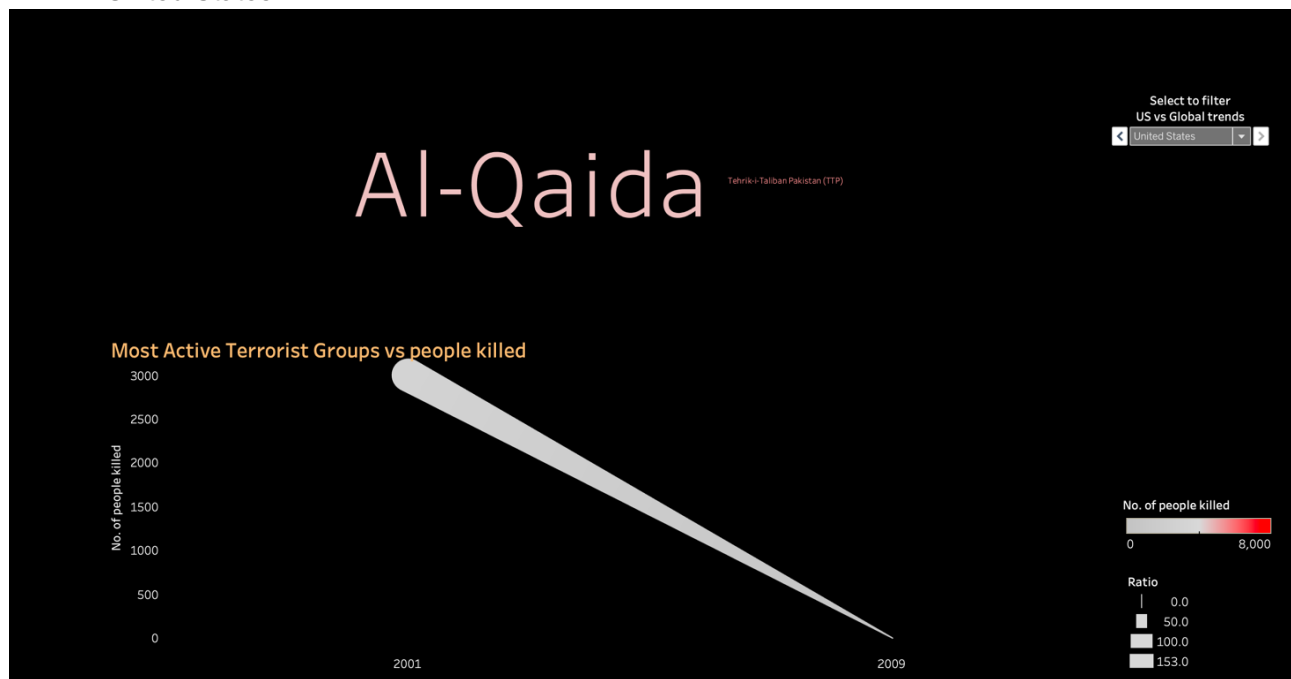


12. Although Al-Qaida has done less than 10% of the terrorist attacks globally, they are the most active terrorist group in the United States with 99.24% attacks and 3000 people killed.

Globally:



United States:



6. Conclusion

With this project we analyzed global trends in terrorist attacks showing most active groups over the years, which countries had highest kills and attacks, comparison between international and domestic attacks, and trends in US. After this analysis, we concluded that since 2004, the number of terrorist attacks have increased globally and over the years nearly 90% of terrorist attacks have been successful. With most attacks in these locations, one should avoid staying around landmarks, police stations and government bureau. You should also watch out for masked people and people with huge bags or suspicious handheld objects. If you encounter a terrorist attack, leave the scene as soon as possible or take cover. With such studies, a data-driven understanding of the nature of terrorist attacks could help inform and prioritize prevention efforts at point of entry into the country and reduce the number of attacks which have been continuously increasing since 2000.