Global Terrorism

Agam Singh, Jyoti Patel,
Data science trainees,
AlmaBetter,
Bangalore

Abstract:
The Global Terrorism Database (GTD) is an open-source database including information on terrorist events worldwide from 1970 through 2017 (with annual updates planned for the future). Unlike many other event databases, the GTD includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 200,000 cases.

The Global Terrorism Database (GTD) is a database on terrorist attacks around the world from 1970 through 2017 containing over 150,000 observations. Researchers at the National Consortium maintain the database for the Study of Terrorism and Responses to Terrorism (START), headquartered at the University of Maryland.

This project entails an exploratory data analysis (EDA) conducted on the Global Terrorism Database, aiming to extract valuable insights and information concerning patterns and trends in global terrorism worldwide. The dataset utilized in this analysis encompasses details about terrorist groups, the most dangerous terrorist organizations, regions significantly affected by terrorist activities, as well as states, cities, and specific locations where terrorist incidents occur frequently, such as military installations, police facilities, government premises, and private properties. This comprehensive dataset sheds light on the global impact of terrorism and provides a yearly account of the lives lost due to such acts of terror spanning from 1970 to 2017.

While analyzing this dataset we found that the most terrorist-affected regions were the Middle East, Africa, and South Asia. The most dangerous and leading terrorist groups, including the Taliban, and the Islamic State of Iraq, are responsible for most of all terrorist deaths.

1. Introduction

This dataset has around 181691 observations in it with 135 columns and it is a mix of categorical and numeric values. The dataset contains a large number of columns but here are some important columns.

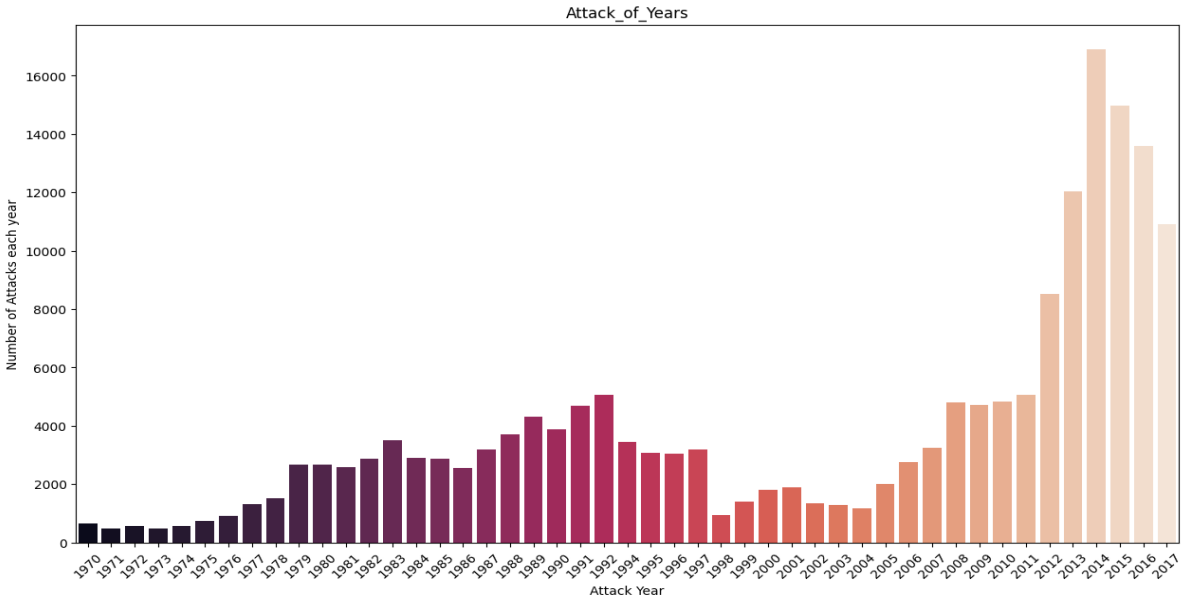
- **event_Id:** This variable consists of numerical values and all values are unique.
- **Year:** This variable consists of numerical values and each value represents the particular Year.
- **Month:** This variable consists of numerical values and each value represents the particular Month.
- **Day:** This variable consists of numerical values and each value represents a particular day.
- **Country:** This variable consists of categorical values and each value represents the Year.
- **region:** This variable consists of categorical values and each value represents a particular region.
- **state:** This variable consists of categorical values and each value represents a particular state.
- **City:** This variable consists of categorical values and each value represents a particular City.
- **latitude:** This variable consists of numerical values and each value represents the latitude of the specific location of terrorist attacks.
- **longitude:** This variable consists of numerical values and each value represents the latitude of the specific location of terrorist attacks.
- **location:** This variable consists of categorical values and each value represents a particular location of terrorist attacks.
- **AttackType:** This variable consists of categorical values and each value represents a particular type of terrorist attack.
- **Target_type:** This variable consists of categorical values and each value represents a particular type of terrorist attack target.
- **Gang_name:** This variable consists of categorical values and each value represents a particular type of terrorist Gang name.
- **Killed:** This variable consists of numerical values and each value represents the particular number of people killed.

● **Weapon_type:** This variable consists of categorical values and each value represents a particular type of Weapon used in terrorist attacks.
● **property:** This variable consists of categorical values and each value represents a particular type of property attacked by terrorists.
● **motive:** This variable consists of categorical values and each value represents a motive for attacks.

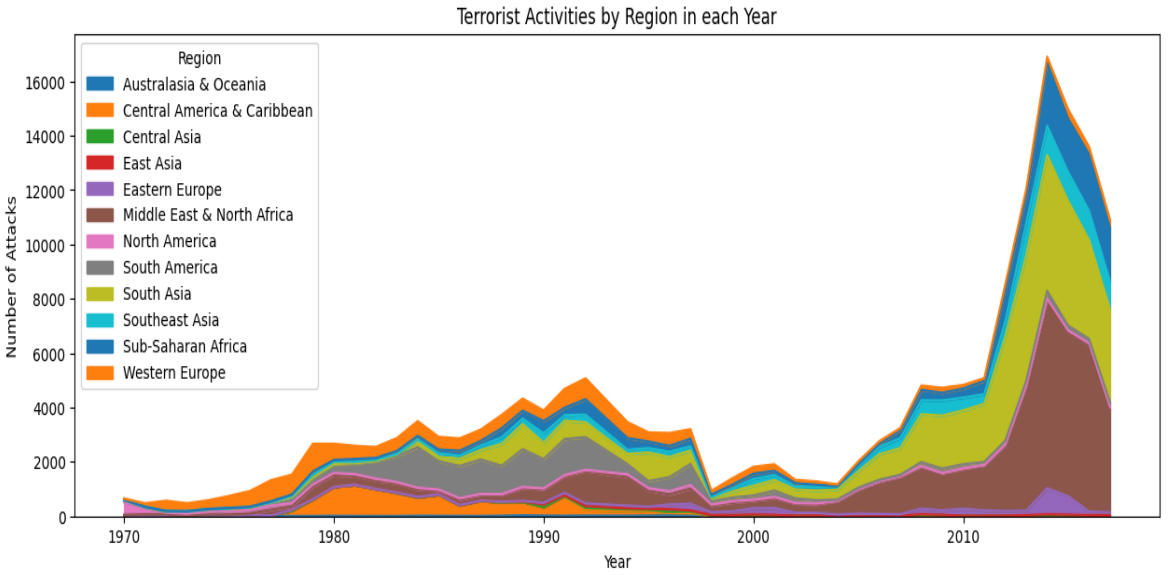
This exploratory data analysis (EDA) project on the Global Terrorism Database reveals crucial insights into the nature and impact of terrorism worldwide. By examining patterns and trends spanning from 1970 to 2017, the analysis sheds light on various aspects of global terrorism. The dataset provides information on terrorist groups, identifies the most dangerous ones, and highlights the regions, states, and cities most affected by terrorist activities. Furthermore, it delves into the specific locations targeted by terrorists, including military installations, police facilities, government premises, and private properties. The dataset also quantifies the human cost of terrorism, capturing the number of lives lost each year. Overall, this EDA project provides a comprehensive understanding of the magnitude and consequences of terrorism on a global scale.

2. Problem Statement

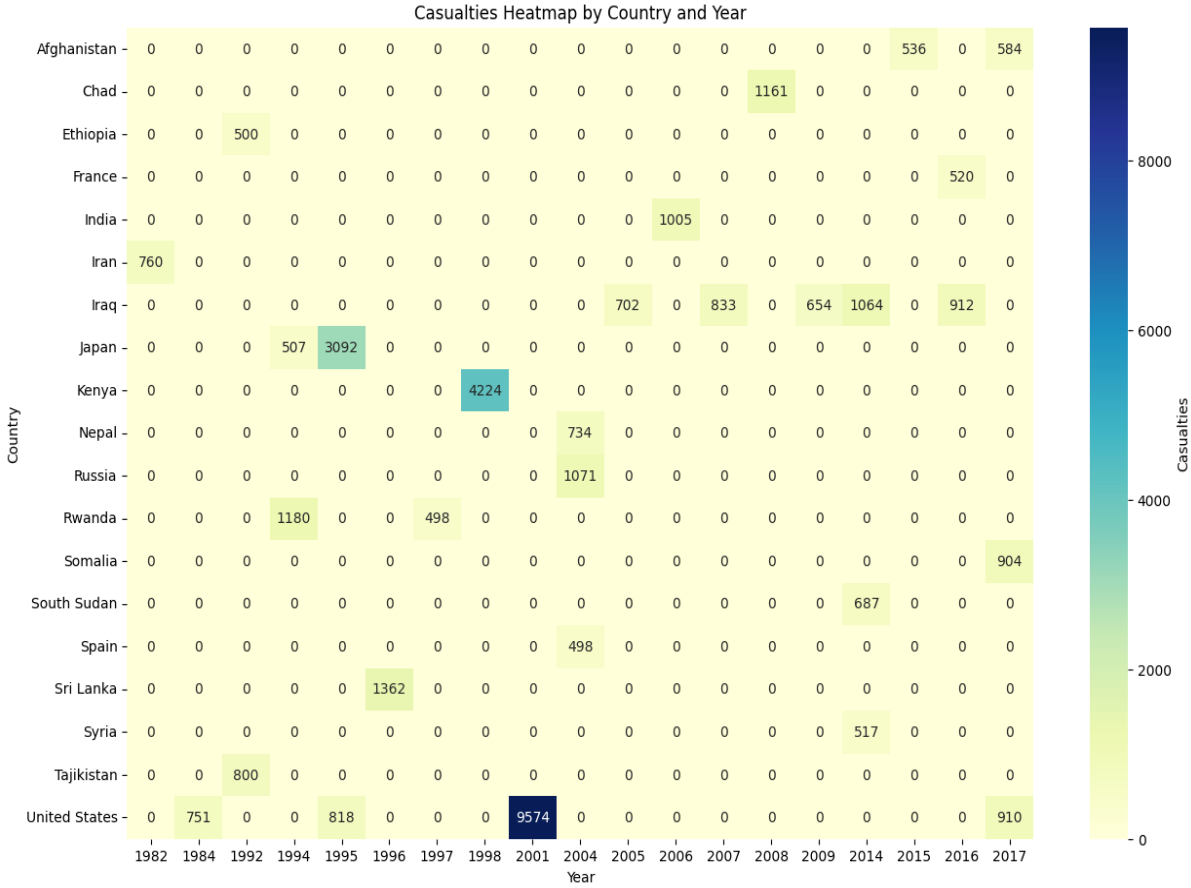
➢ Number of Terrorist Activities Each Year.



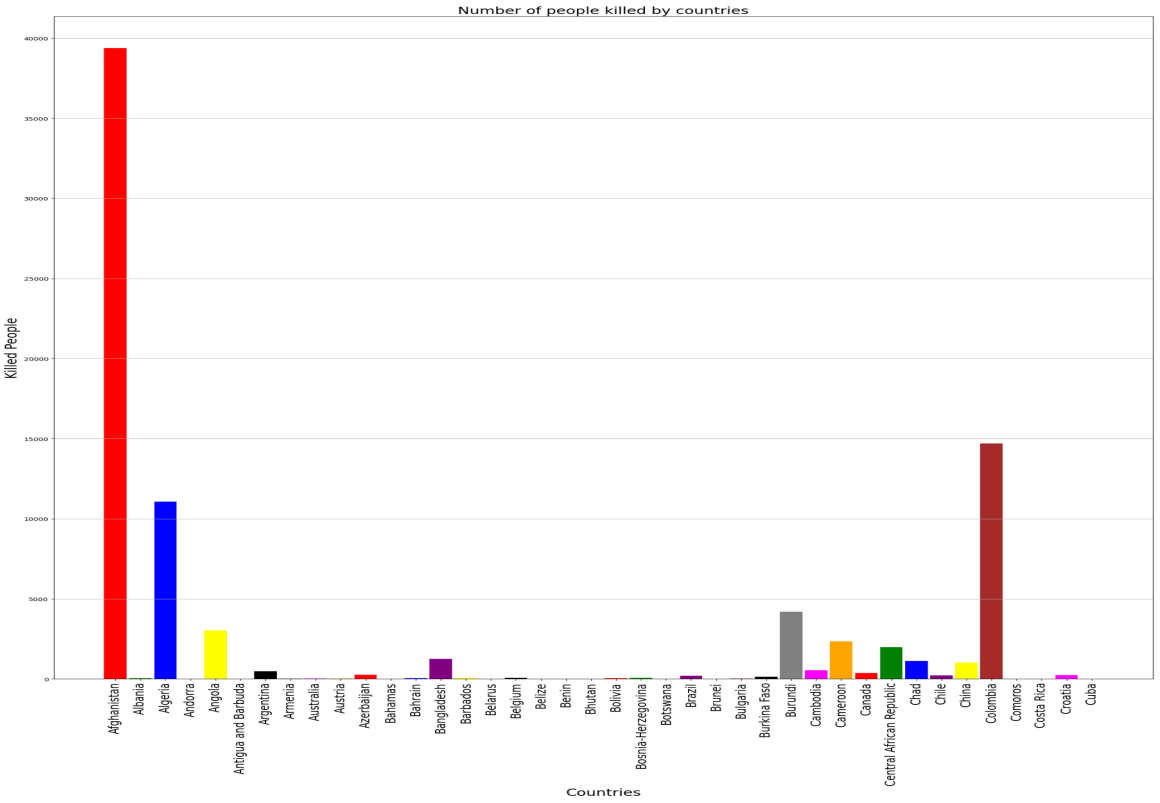
➢ Terrorist Activities by Region in Each Year.

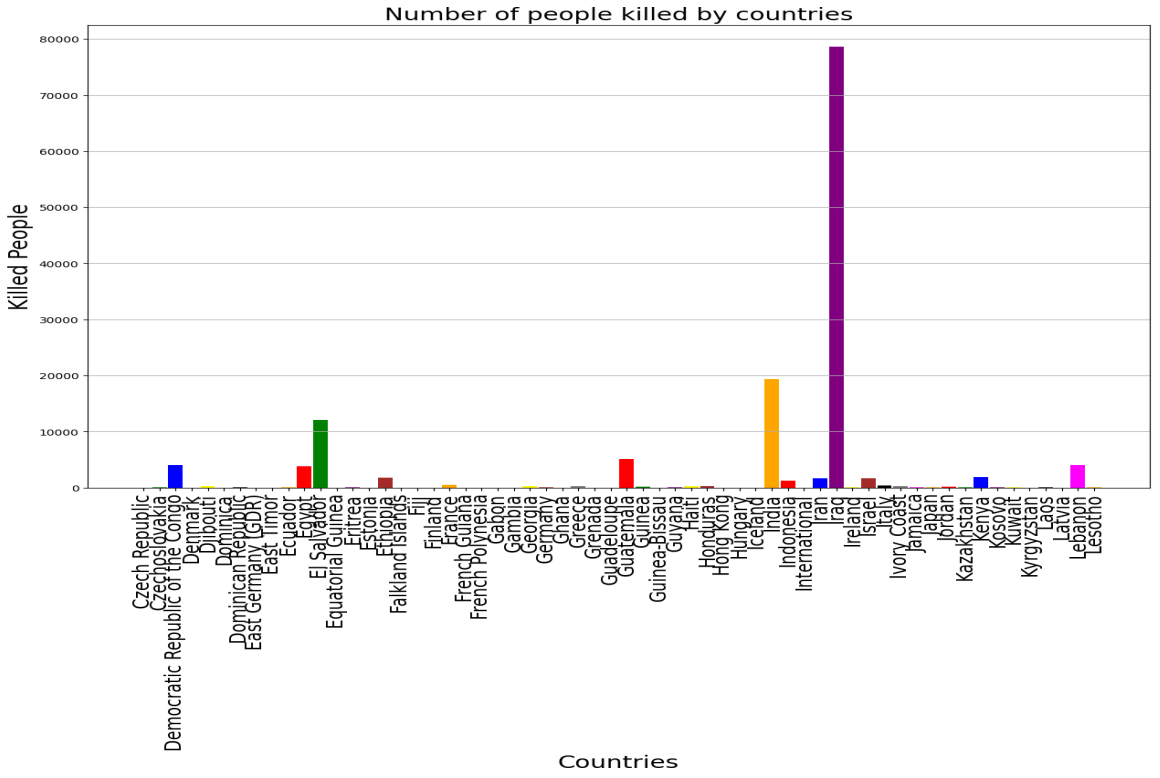


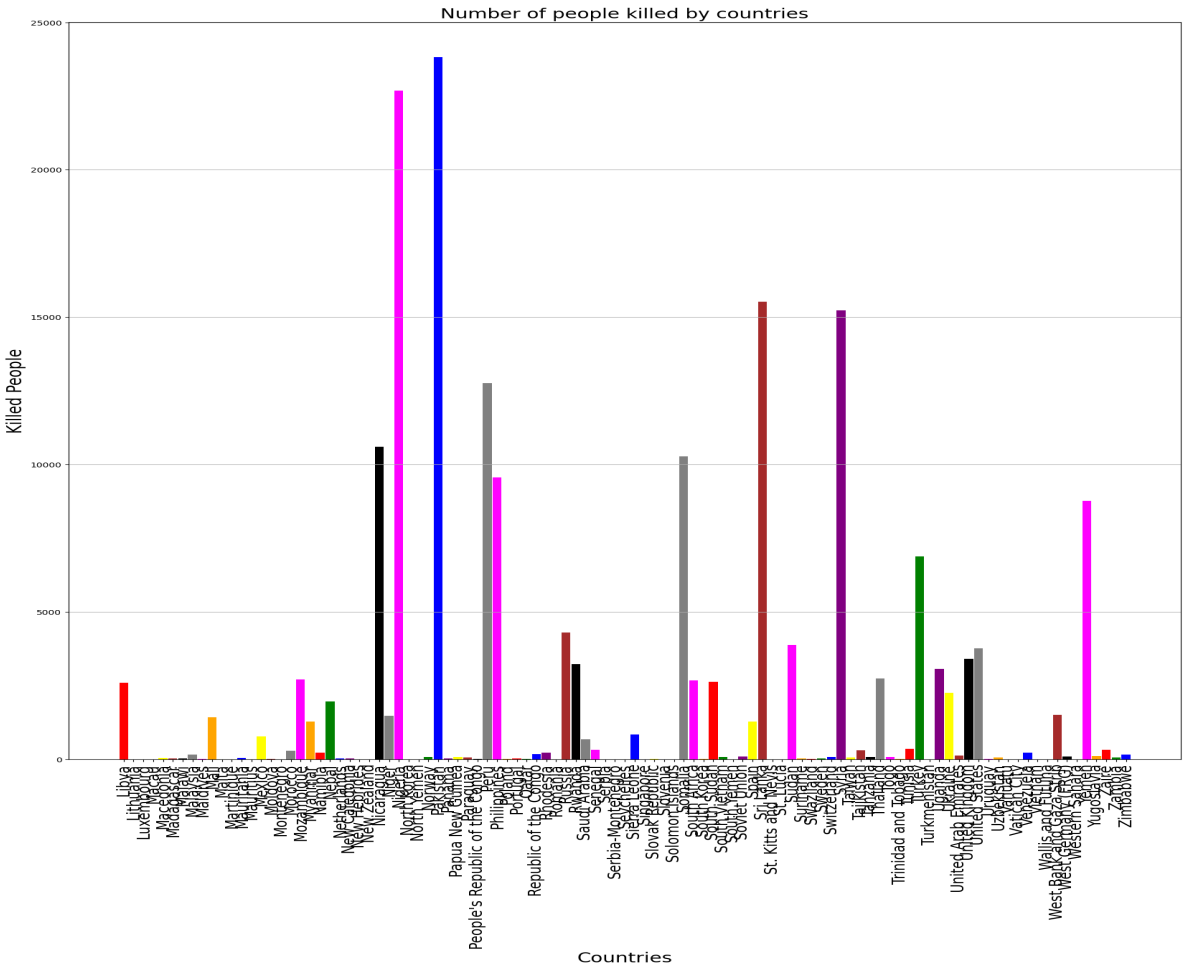
- Correlation between different countries' Casualties with respect to Years, using HeatMap.



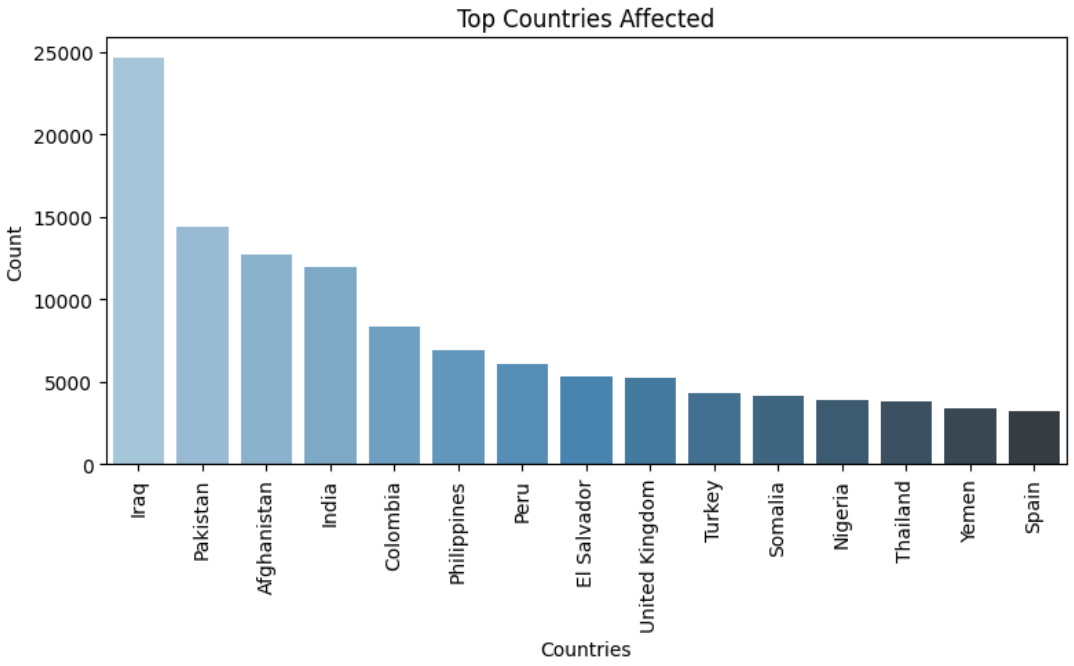
➢ Number of People Killed in Terrorist Attacks across Different Countries.



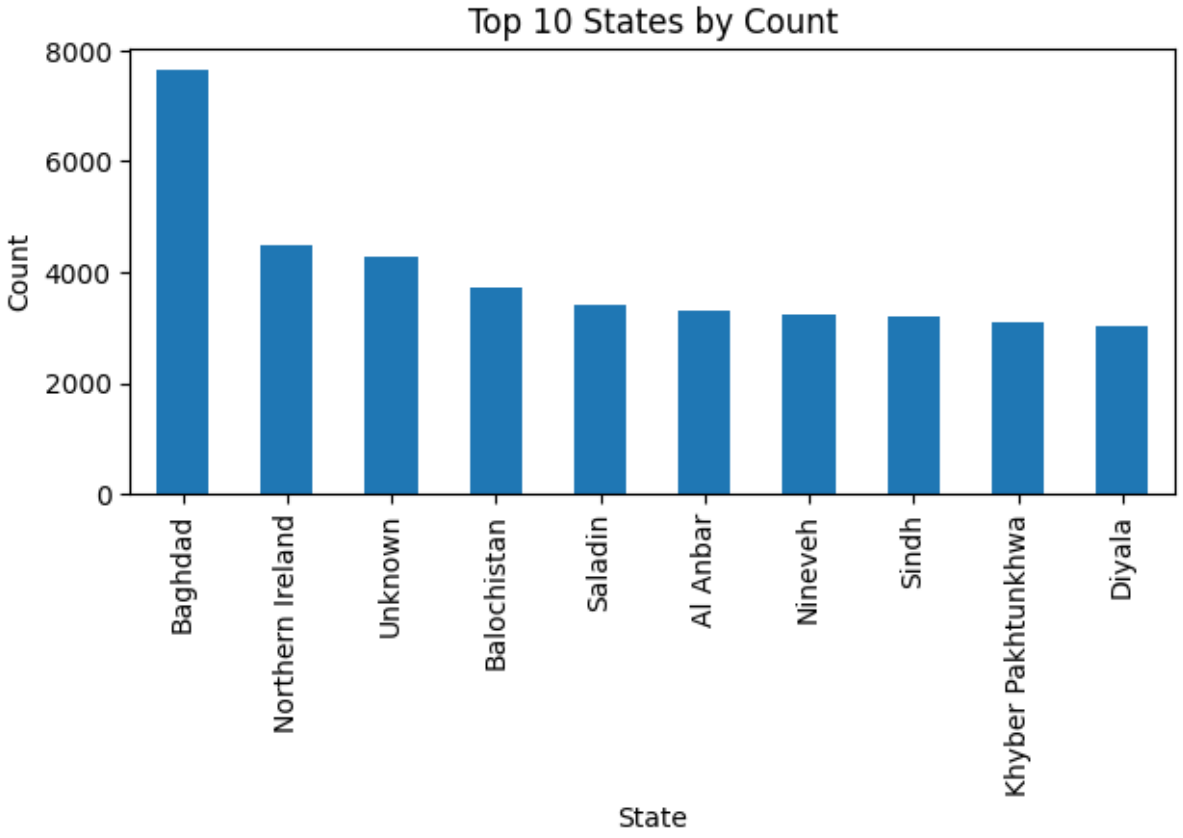




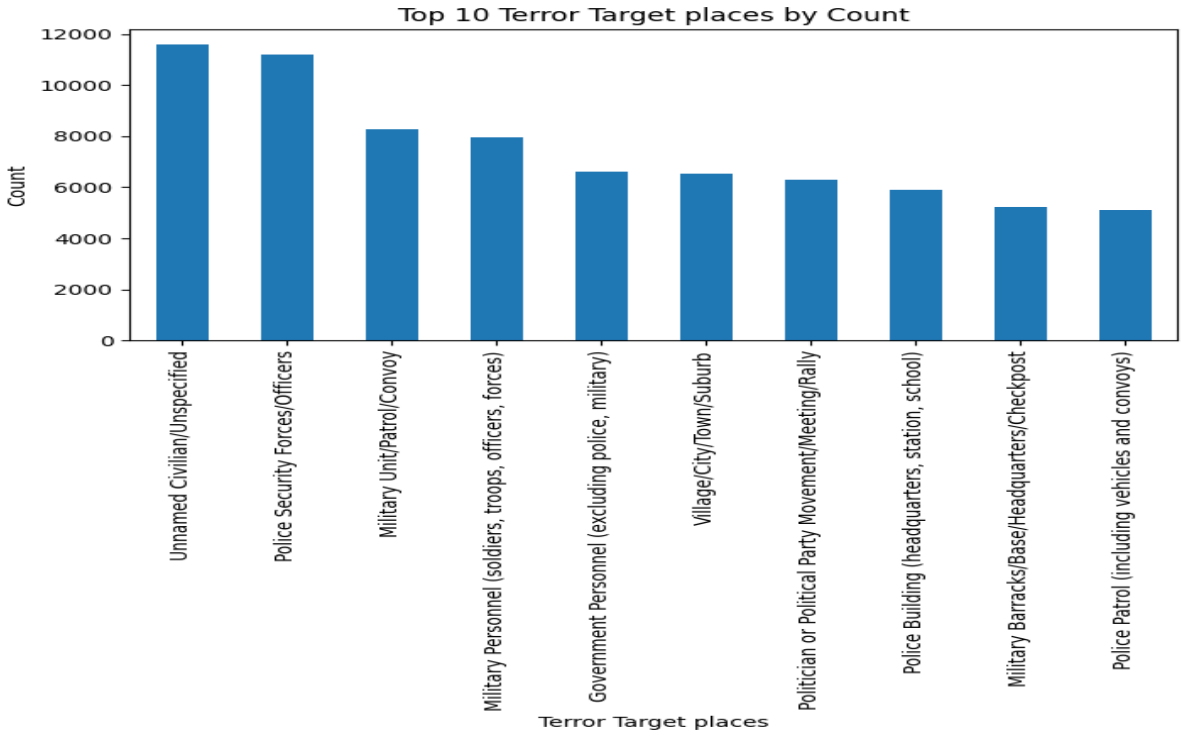
➢ Top Countries Affected by Terror Attacks.



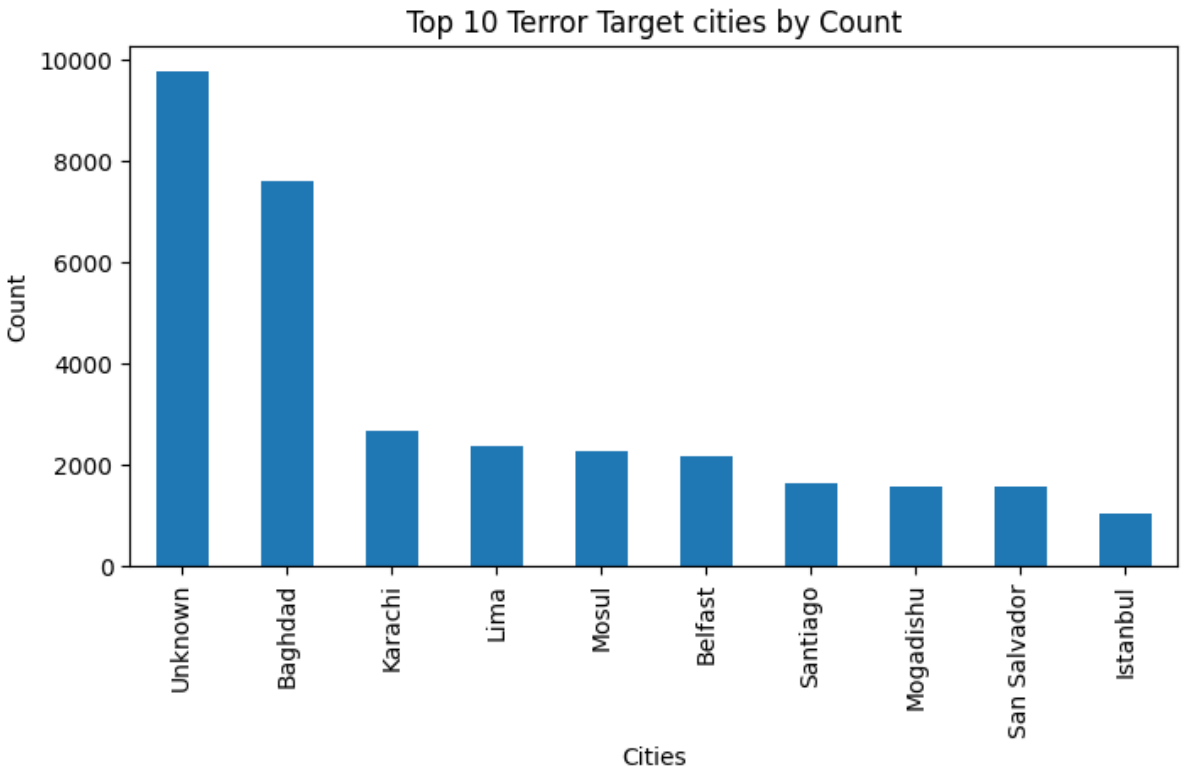
➢ Top 10 states most impacted by terrorism.



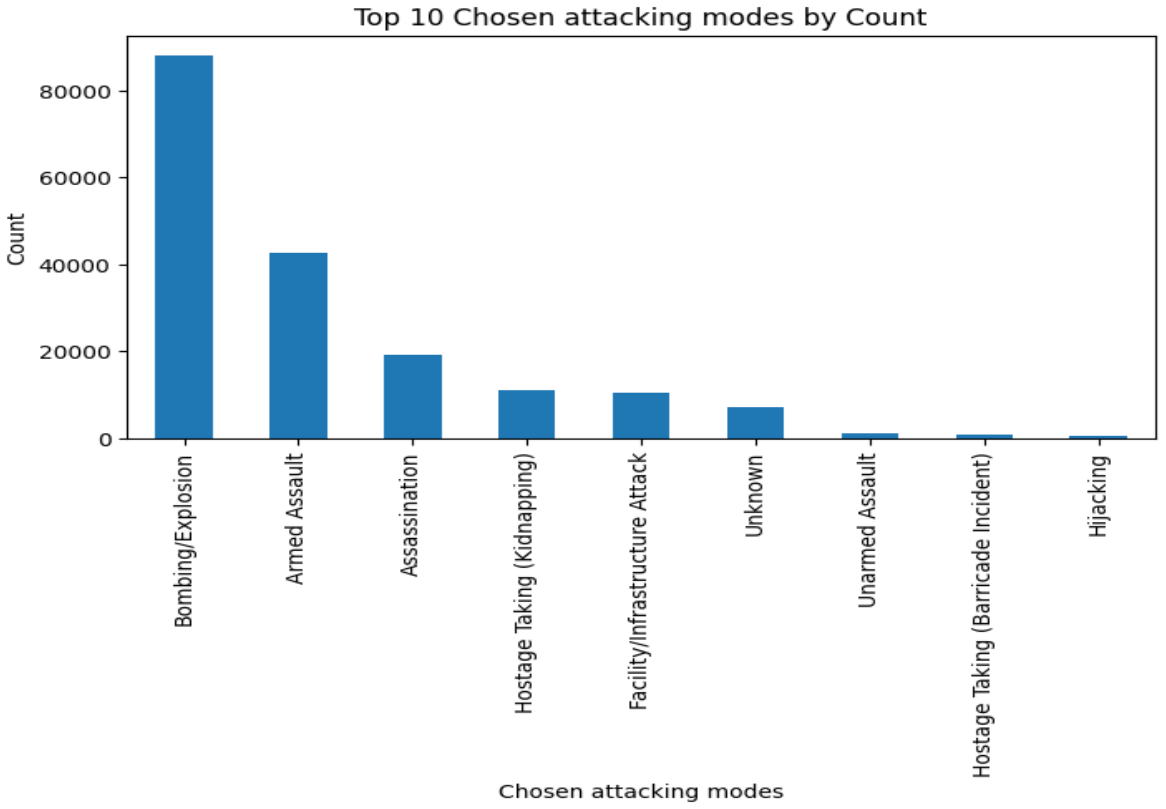
➢ Top 10 targeted locations by terror attacks.



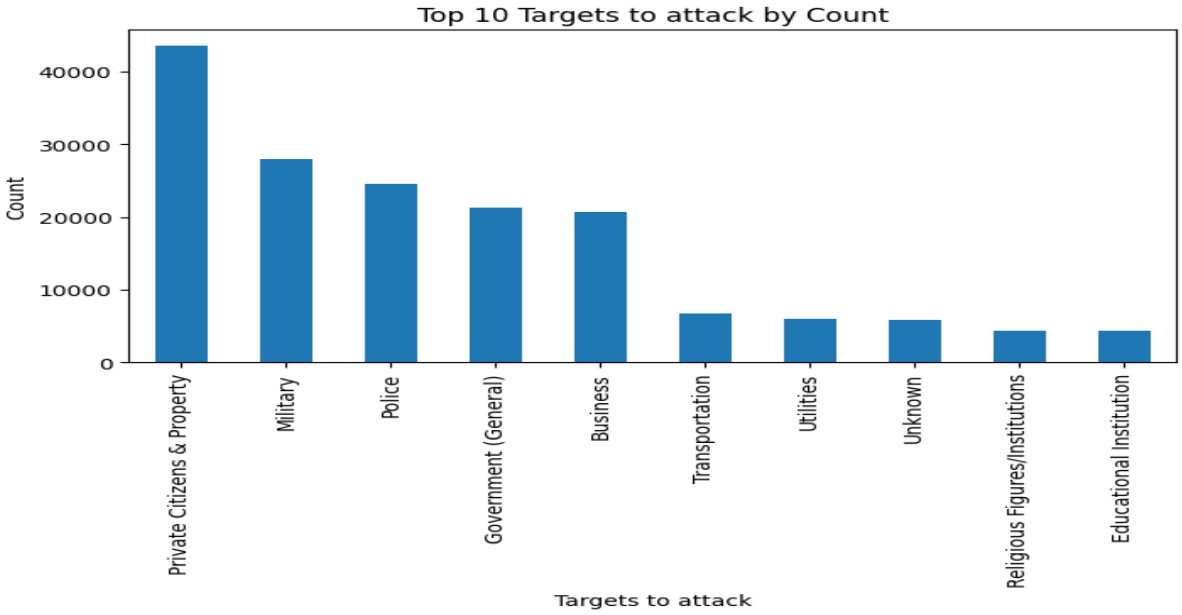
➢ Top 10 cities most impacted by terrorism.



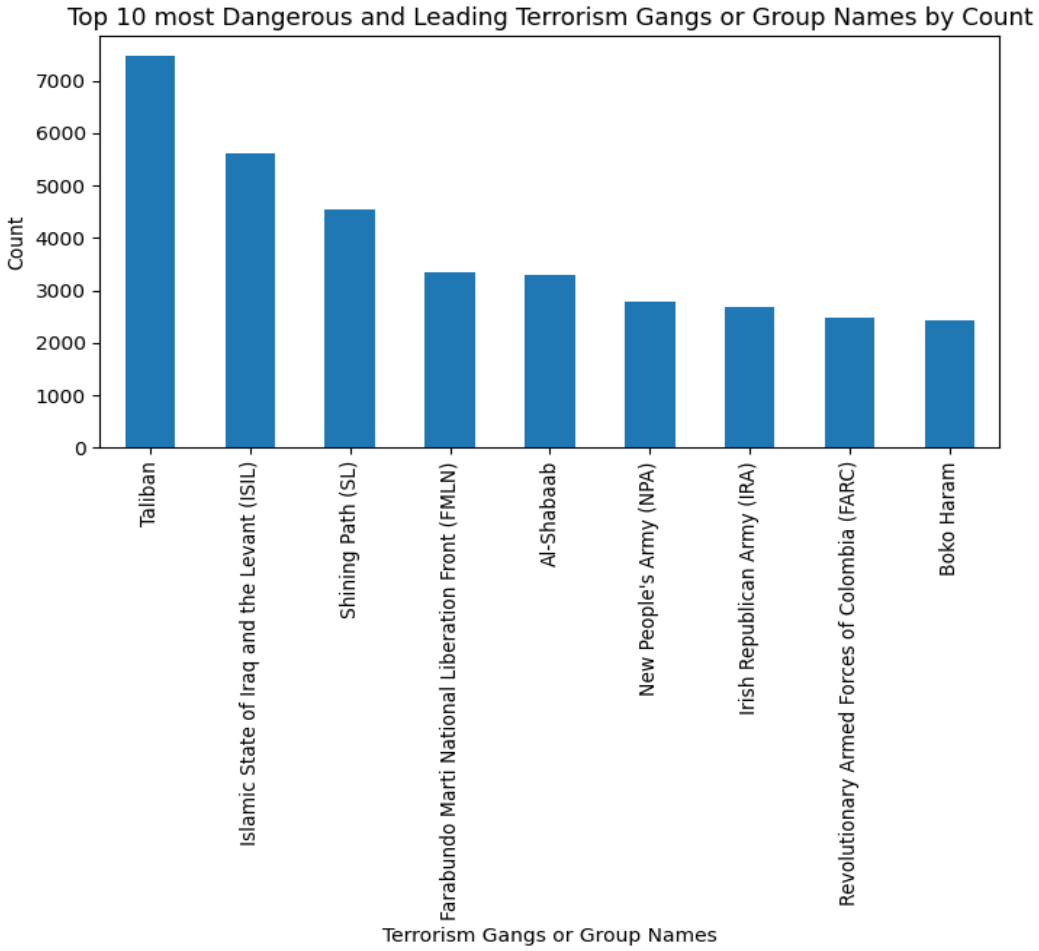
➢ Top 10 chosen methods of attack.



➢ Top targets of terrorist attacks.



➢ Top Most Dangerous and Leading Terrorism Gangs or Group Names.



3. Steps involved:

- **Exploratory Data Analysis**

  After loading the dataset we performed this method, aiming to extract valuable insights and information concerning patterns and trends in global terrorism worldwide. The dataset utilized in this analysis encompasses details about terrorist groups, the most dangerous terrorist organizations, regions significantly affected by terrorist activities, as well as states, cities, and specific locations where terrorist incidents occur frequently, such as military installations, police facilities, government premises, and private properties. This comprehensive dataset sheds light on the global impact of terrorism and provides a yearly account of the lives lost due to such acts of terror spanning from 1970 to 2017.

- **Null values Treatment**

  Our dataset contains a large number of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project in order to get a better result.

- **Data Exploration**

  Exploring the dataset to understand the structure of the data, identify any patterns or trends, and detect outliers.

- **Feature Engineering**

  Creating new features from the existing data to gain more insights and improve the predictive power of the model.

- **Data Visualization**

  Creating various charts and graphs to visualize the data and communicate the insights gained from the analysis.

- **Statistical Analysis**

  Conducting statistical tests to validate the insights and hypotheses generated from the data.

- **Standardization of features**

  Our main motive through this step was to scale our data into a uniform format that would allow us to utilize the data in a better way while performing fitting and applying different algorithms to it. The basic goal was to enforce a level of consistency or uniformity to certain practices or operations within the selected environment.

4. Conclusion:

The exploratory data analysis (EDA) conducted on the Global Terrorism Database has provided significant insights into the patterns and trends of global terrorism. The analysis revealed that 2014 recorded the highest number of terrorist activities, while 1971 had the lowest incidence. Western Europe emerged as the region with the highest number of terrorist activities, whereas Australasia & Oceania experienced the lowest.

One striking finding was that in 2014, a majority of terrorist attacks (84%) were concentrated in the Afghanistan, Iraq, and Israel regions. Subsequently, regions such as the Philippines, Malaysia, parts of Europe (such as Ukraine), and the African continent witnessed significant terrorist activity. The Middle East and North Africa, currently plagued by wars and terrorist attacks, emerged as the epicenter of these incidents. The impact of terrorism in the Middle East and North Africa regions has been devastating, with Muslims being disproportionately affected. Analysis of the dataset highlighted Iraq, Afghanistan, and Pakistan as the most severely impacted countries, reflecting the concentrated toll on Muslim-majority nations. Cities also faced the brunt of terrorist activities. Baghdad stood out as the most affected city, followed by Northern Ireland, Balochistan, Saladin, and others. These cities bore witness to significant incidents of terrorism, leaving a lasting impact on their communities.

Examining the target locations of terrorist attacks revealed the range and nature of these acts. The top 10 targets included unnamed civilians, police security forces/officers, military units/patrols/convoys, military personnel, government personnel, villages/cities/towns/suburbs, politicians or political party movements/meetings/rallies, police buildings, military barracks/bases/headquarters/checkpoints, and police patrols.

The chosen modes of attack were also identified, with bombing or explosions being the most commonly employed method, followed by armed assaults, assassinations, and various other tactics. These findings shed light on the strategies and techniques employed by terrorist groups to inflict harm and create fear.

Furthermore, the dataset provided valuable insights into the most dangerous and leading terrorist groups, including the Taliban, Islamic State of Iraq and the Levant (ISIL), Shining Path (SL), Farabundo Marti National Liberation Front (FMLN), Al-Shabaab, and others. Understanding the prominence and activities of these groups is crucial for formulating effective counterterrorism measures.

This EDA project on the Global Terrorism Database has revealed crucial information about the geographical and temporal patterns of terrorism. It has shed light on the regions and cities most affected by terrorist activities, the preferred targets of these attacks, the methods employed, and the leading terrorist groups. This comprehensive understanding of global terrorism is essential for devising strategies to mitigate its impact and promote global security.