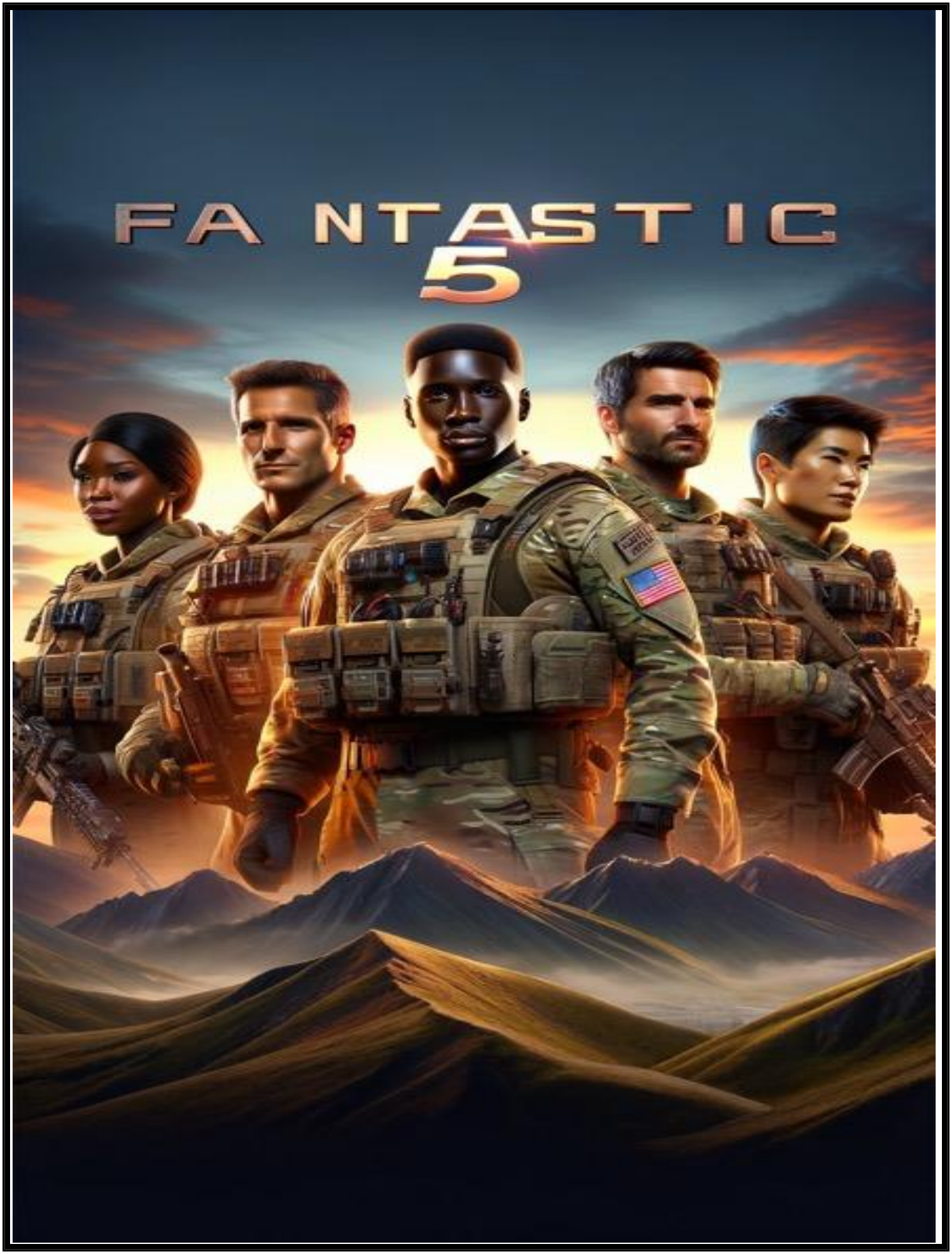

FA NTAS T IC
5

1. Introduction

Domain and Problem Description

- Terrorist activities have a profound impact on global security and require a comprehensive understanding to effectively address them. As security and defense analysts, our mission is to analyze and derive meaningful insights from the Global Terrorism Database (GTD) to address critical inquiries.
- This project will explore the evolution of the frequency of terrorist activities over time and identify regions with divergent trends from global norms. Additionally, we aim to uncover any correlations between the number of incidents and the number of casualties, as well as assess the tactics employed in these attacks. This analysis will provide valuable insights to inform security strategies and global policies.

Tools Used: R and Tableau

Data Collection

- The data for this project has been sourced from the Global Terrorism Database (GTD), a comprehensive open-source database managed by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of Maryland. The GTD contains detailed information on terrorist incidents worldwide from 1970 through 2017. This data is collected from unclassified media articles, providing a reliable foundation for our analysis.

  Global Terrorism Database (kaggle.com)
  Global Terrorism Database (umd.edu)

Dimensions of the Data

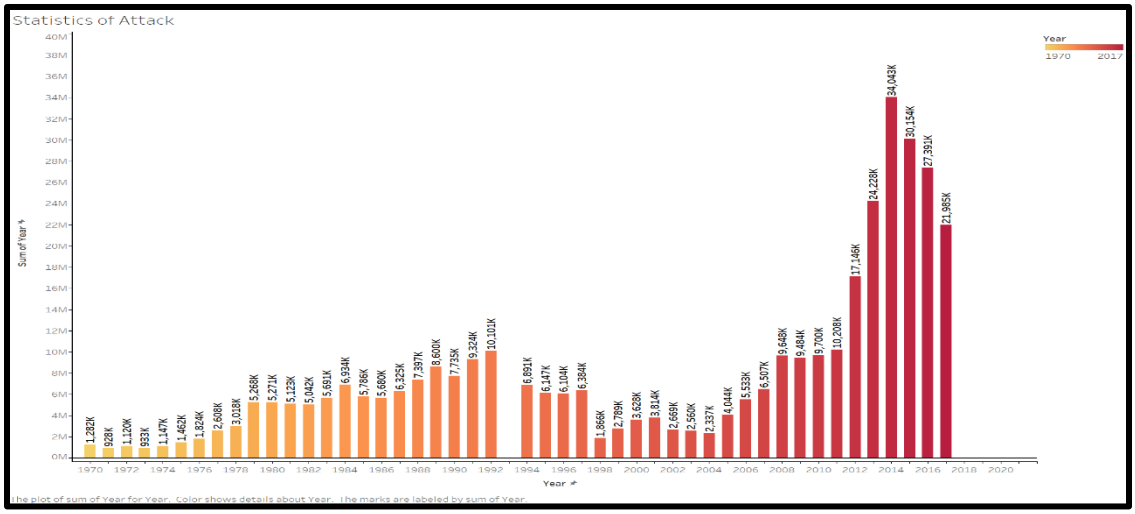
- The dataset consists of an extensive collection of more than 180,000 terrorist incidents, spanning 48 years from 1970 through 2017 (except 1993). It encompasses over 100 variables, including both numerical data, such as casualty counts, and categorical data, such as attack types and target types, locations, etc. The dataset's rich dimensions offer a substantial opportunity for in-depth exploration and analysis.
- Our analysis will center around several key data variables from the GTD. The response variable will be the count of terrorist incidents, while the explanatory variables encompass the year of the incident, location (country, state, region), attack type, target type, weapon type, and the group responsible for the attack. Variables such as the number of casualties (killed and wounded) will allow us to explore the impact of these incidents in greater detail. These variables will be critical in answering our research questions.

Data Cleaning, Pre-processing, and EDA

- Handling Missing Values: We checked for missing values in the entire dataset and found that it contained a significant number of NaN or missing values. To make the data suitable for analysis and visualization, we filled in missing values with relevant placeholders or values such as 'UNKNOWN' or 'Undocumented' for categorical variables. Additionally, we imputed numerical variables with appropriate values like mode or zero. Similar approach was used to resample and missing values in the categorical columns.
- Data Transformation: We performed various data transformations, such as renaming columns for clarity, converting data types, and mapping values to meaningful categories. These transformations improved the readability, reduce clear and usability of the dataset. All the data was transformed to aid in effective visualization with the minimal loss of data whilst data cleaning.
- Cleaning the 'Country' Column: We used the 'country code' package to standardize country names and remove unmatched values, making the 'Country' column consistent and ready for analysis.
- Removing Unmatched Values: We excluded rows with unmatched values that were not valid country names, ensuring the integrity of the 'Country' column.
- The dataset after cleaning, now comprises 181,513 observations across 26 variables (the original containing 181619 rows and 135 columns), spanning the years 1970 to 2017. The data includes information about the year, incident duration, location (city, country, and region), incident success, attack details (tactic, type, and suicide involvement), target details (type and subcategory), terrorist group, motive, weapon information (type and subtype), casualties (killed, wounded, and terrorists killed), property damage, hostage situations (presence and outcome), ransom details (amount and payment), and total casualties.
- Our analysis will center around several key data variables from the GTD. The response variable will be the count of terrorist incidents, while the explanatory variables encompass the year of the incident, location (country, state, region), attack type, target type, weapon type, and the group responsible for the attack. Variables such as the number of casualties (killed and wounded) will allow us to explore the impact of these incidents in greater detail. These variables will be critical in answering our research questions.

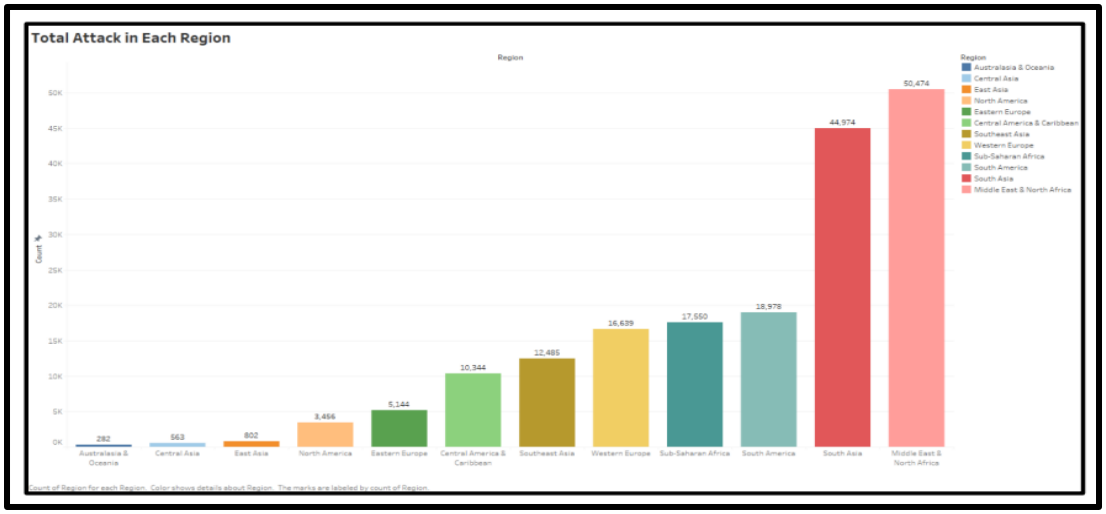
2. Visualizations

1. Total Attacks in each year



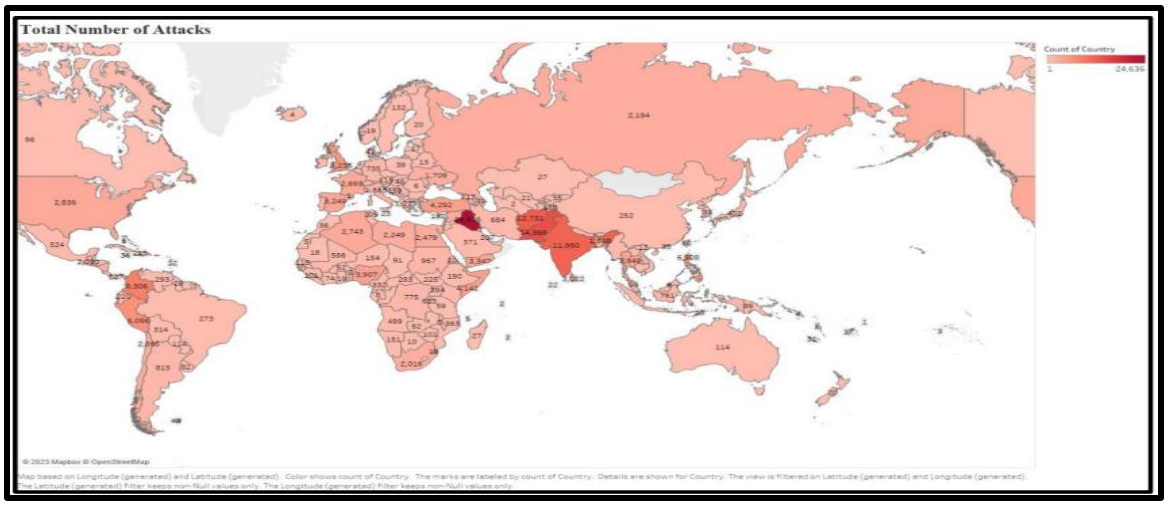
- During the whole timeline between 1970 to 2010, we can observe that there is a significant increase in the terrorist activities in the past few years. There is a sudden spike in the number of attacks since 2006, reaching as many as 34M attacks in 2014. The graph shows a void space in 1930 since no data is available for that year.

2. Total attacks in each Region



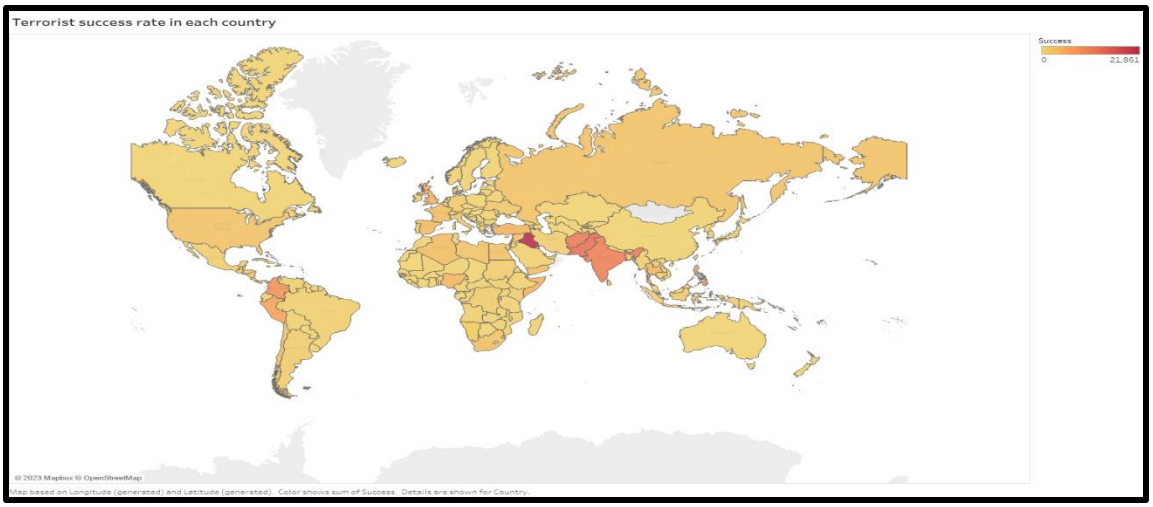
- In the bar graph depicting total attacks in each of the twelve world regions, it is evident that the Middle East and North Africa region has experienced the highest number of attacks, totaling 50,474 incidents.
- On the contrary, Australasia and Oceania stand out as the region with the lowest number of attacks, recording a comparatively minimal count of 282 incidents.
- This data not only highlights the stark contrast in the security landscape among global regions but also emphasizes the varying degrees of susceptibility to terrorism worldwide.

3. Nations that experienced the highest number of attacks since 1970



- In the choropleth representation of the Global Terrorism Database spanning the years 1970 to 2017, it is evident that Iraq has experienced the highest number of terrorist attacks, totaling 24,636 incidents.
- Following closely is Pakistan, ranking second with 14,368 recorded attacks. This data highlights the significant impact of terrorism on these nations over the specified period.
- The disparity in attack counts underscores the varying degrees of vulnerability and resilience among countries in the face of terrorism, offering valuable insights into the global landscape of security challenges.

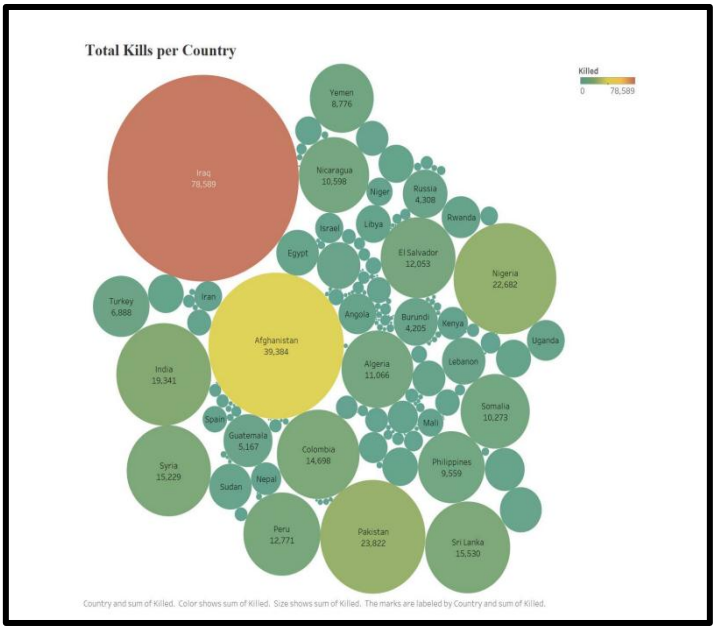
4. Terrorism success rate in each country



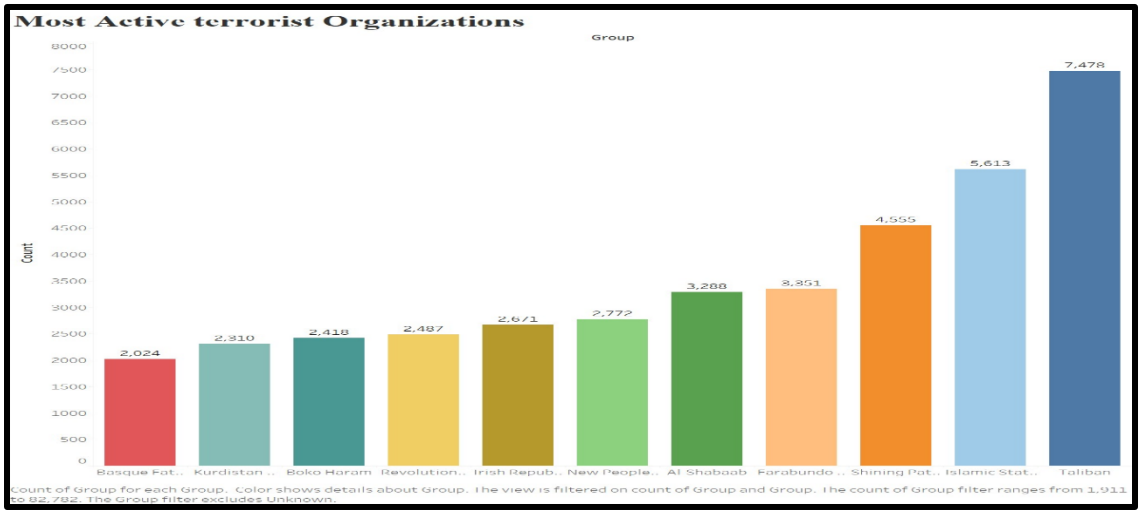
- It can be inferred from the above visualization that; the terrorists have been executing strikes at a significant rate in the middle east and south Asian regions. Specifically in the countries India, Pakistan, Afghanistan, and Iraq.
- Iraq that has experienced the highest number of terrorist attacks, totaling 24,636 incidents as seen in the previous visualization, has suffered a very significant activity: with the terrorists actively executing as many as 21,861 strikes.

5. Total deaths per Country

- The success rates of the terrorists led to huge tolls of deaths. Over the time, several countries suffering from those attacks left many killed, wounded and some stranded due to loss of property.
- Here, the above graph enables us to visualize the total number of people killed in each country by Terrorist Attacks. The size of the bubble represents the number of deaths. As we can see, Iraq has the biggest bubble with the highest death toll of 78,589 followed by Afghanistan with a death toll of 39,384.
- We can notice how the bubble is smaller for lower death toll countries like Burundi, Russia, Yemen etc.

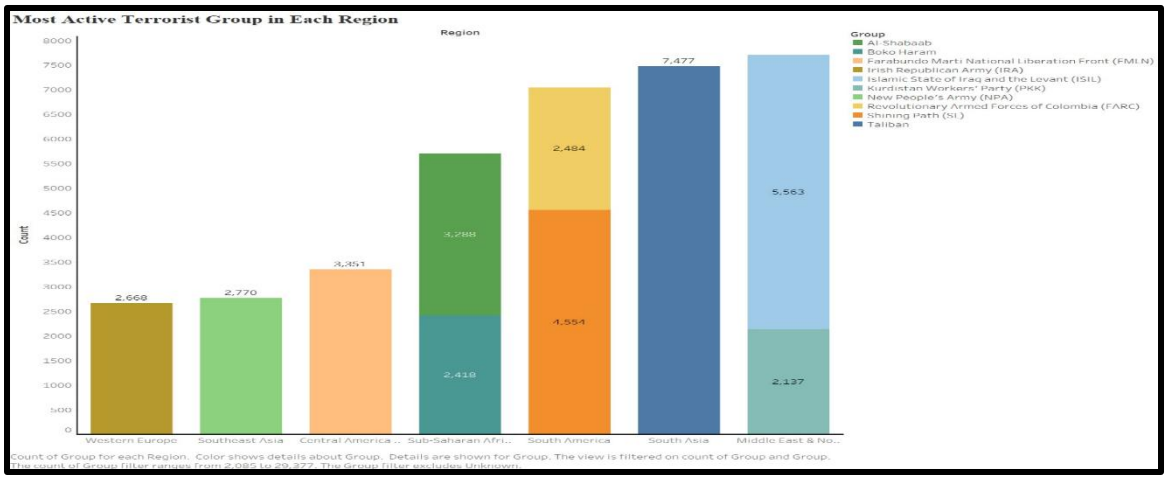


6. Most Active Terrorist Organizations



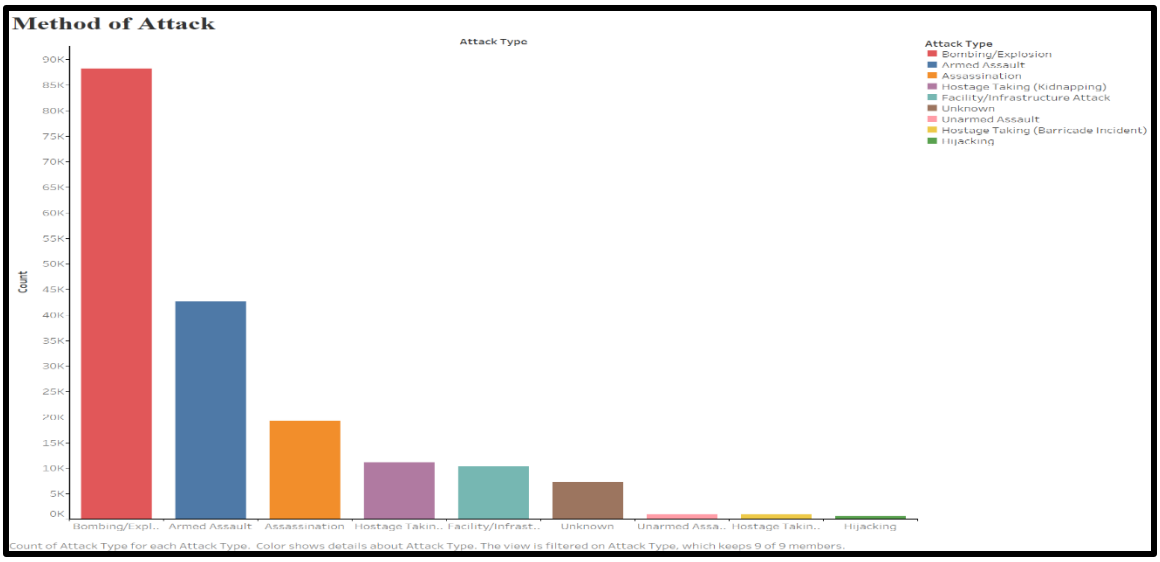
- Terrorist Groups Who Caused the Most Casualties: The above visualization sheds light on the most lethal terrorist groups in terms of causing casualties. Taliban have inflicted the highest number of casualties, followed by "Islamic State of Iraq and the Levant (ISIL)" amongst the known ones.
- Understanding which groups have caused the most harm is essential for prioritizing efforts to counteract and combat these organizations effectively. This assists in focusing counterterrorism resources where they are needed most and in minimizing the human cost of terrorist activities.

7. Most active Terrorist Group in Each Region



- This interactive visualization presents the most active terrorist groups in each region, such as South Asia, Western Europe, Southeast Asia, Central America, Sub-Saharan Africa, and South America; based on their reported counts.
- For example, the Taliban stands out as the most active terrorist group in South Asia, whereas in Central America, the FMLN holds the position of the most active terrorist group, leading 3351 attacks. Similarly, in the Middle East and North African region, both ISIL and PKK are actively involved in acts of terrorism.

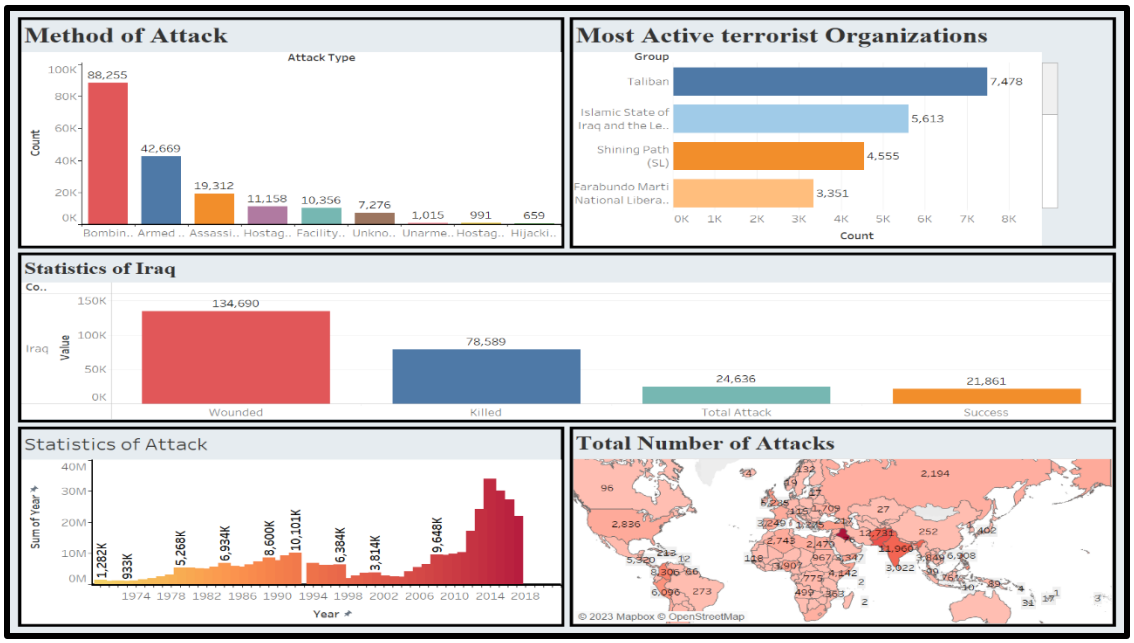
8. Method of Attack



- The above visualization presents a comprehensive view of the number of attacks by different attack types in various regions. It is evident that the type of attacks varies significantly across regions.
- For instance, "Bombing/Explosion" is the most common attack type, while "Armed Assault" is prevalent. Demonstrating the variation in attack types, this visualization reveals distinct preferences. It underscores regional differences in attack methodologies.

Analysis and Discussion:

- Focusing on a single country, for instance IRAQ, which has been severely affected by terrorism with a total of 24,636 attacks out of which 21,861 were not countered effectively and destructive.

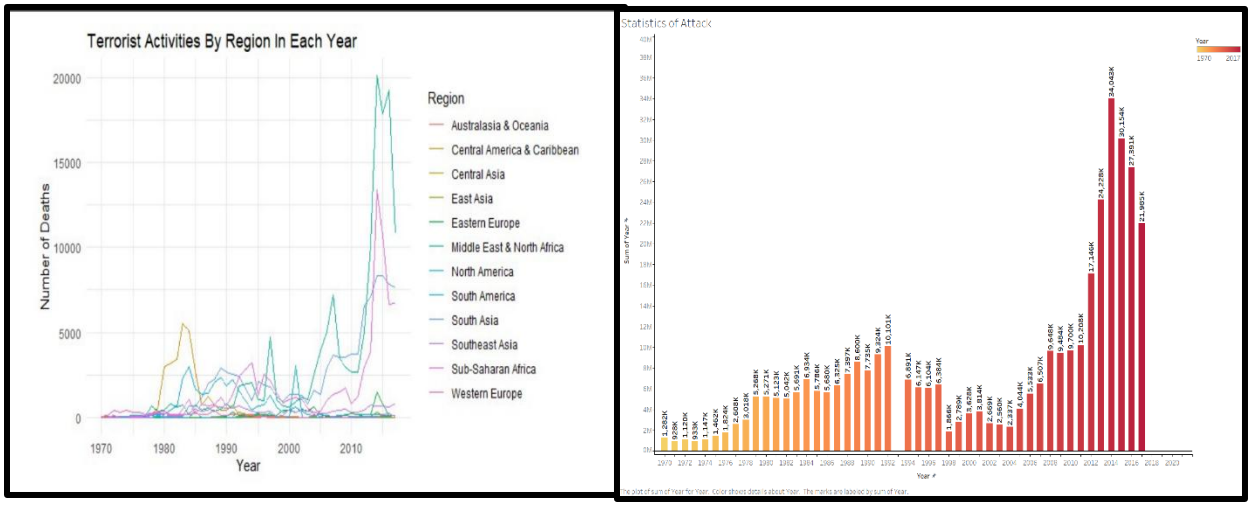


- These attacks left about 78,589 dead and 134,690 wounded.
- The attacks have increased drastically over the time with a spike in attacks from about 2010.
- The most active terrorist attacks across the globe were carried out by Taliban with 7478 attacks and 5613 attacks from ISIL.
- Bombing and Armed assault were the most adopted methodologies for attacks by terrorists tallying up to 130,924 such attacks.

This project aims to comprehensively analyze and derive insights from the Global Terrorism Database (GTD) to address crucial inquiries regarding terrorist activities and their impact on global security. The dataset, spanning 48 years from 1970 through 2017, contains over 180,000 incidents with rich dimensions, allowing for in-depth exploration. The tools used for analysis are R and Tableau. The data, sourced from the GTD, underwent thorough cleaning, pre-processing, and exploratory data analysis (EDA) to ensure reliability and usability. The analysis focuses on key variables such as the count of terrorist incidents, year, location, attack type, target type, weapon type, and the responsible group. Visualizations highlight trends, regional variations, nations with the highest attacks, success rates, total deaths, active terrorist organizations, and attack methods. Notable findings include a surge in attacks since 2006, the Middle East and North Africa experiencing the highest incidents, Iraq suffering most attacks, and the Taliban causing the most casualties. The project's insights contribute to informed security strategies and global policies.

Data Cleaning, Preprocessing and Time Scale Specific Visualization

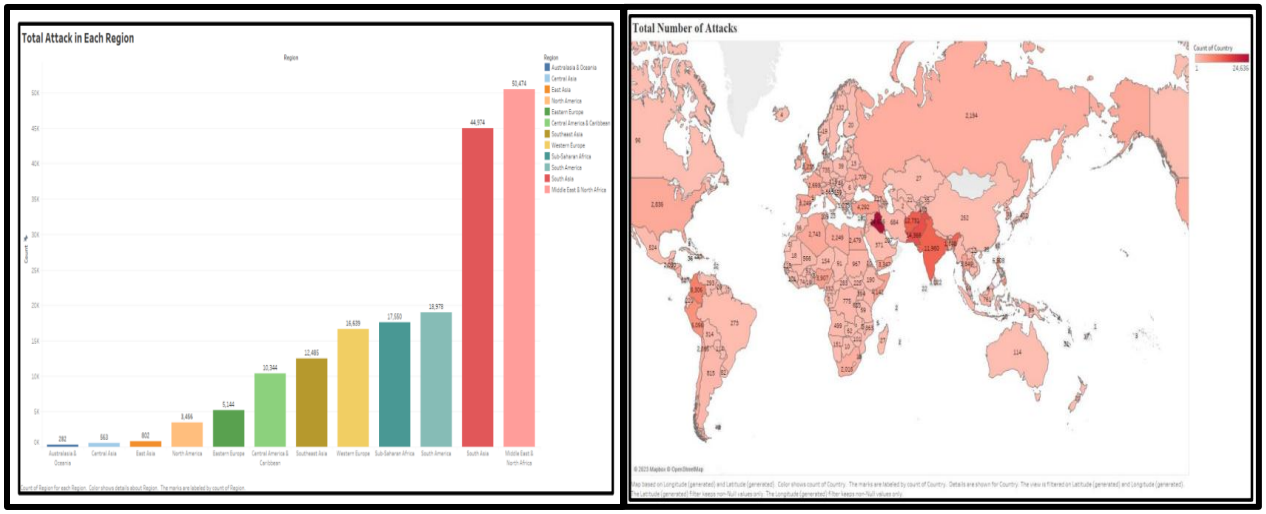
In my role as a team member for the Global Terrorism Project, I took charge of visualizing the total attacks over time. My primary focus was on analyzing the graph covering the years 1970 to 2010. To effectively communicate the significant changes in attack frequencies, I implemented a different scale on the axes. This not only highlighted trends but also allowed for a clearer representation of variations. Additionally, I tackled the challenging issue of missing data by collaborating with team members in the cleaning and preprocessing stages.



This experience significantly enhanced my skills in visualization techniques in R alongside Tableau. I learned the importance of selecting appropriate scales to ensure a more accurate representation of data trends. The logarithmic scale proved crucial in emphasizing the variations in attack frequencies. Moreover, addressing missing data showcased the challenges involved in working with large datasets, providing valuable insights into data cleaning techniques and the importance of data integrity.

Regional Disparities in Terrorism Incidents

In my contribution to the Global Terrorism Project, my focus shifted towards creating visualizations that depicted the total attacks in different regions. The key visualization was a bar graph showcasing the twelve world regions, emphasizing the stark contrasts in the security landscape. Notably, Australasia and Oceania stood out with a minimal count of incidents, providing a unique perspective on regional susceptibility to terrorism.

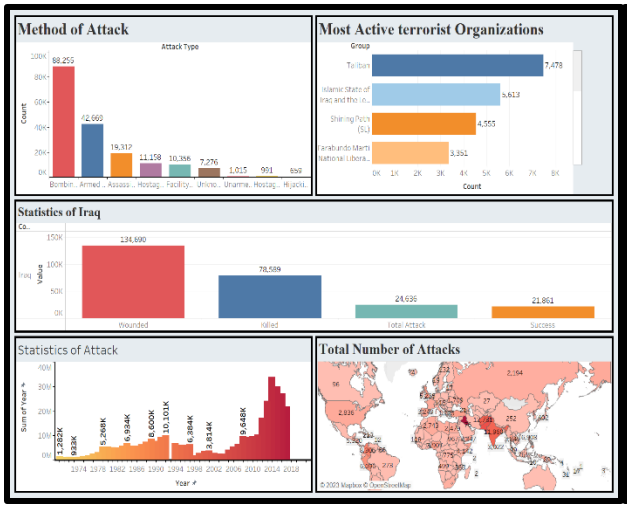


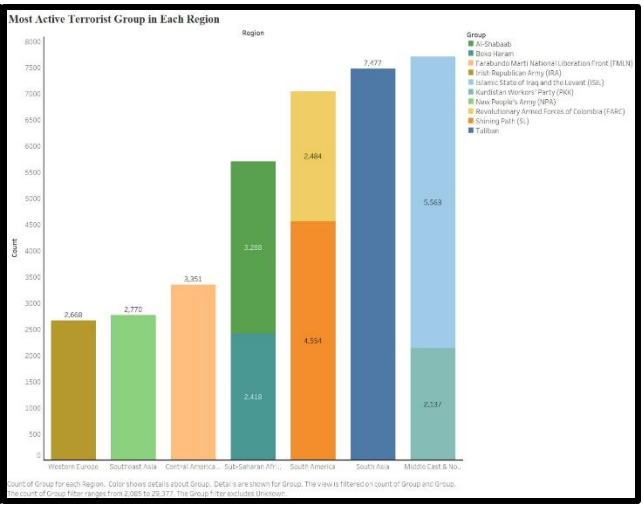
This task deepened my proficiency in geographical visualizations and allowed me to grasp the varying degrees of vulnerability to terrorism globally.

Reflecting on this experience, I realized the significance of region-specific analyses in understanding the global terrorism landscape. The visualization brought attention to the unequal distribution of terrorist incidents, underlining the need for targeted security measures in specific regions. This project expanded my skills in creating visualizations that provide nuanced insights into complex global issues.

Country-Specific Visualizations

For the Global Terrorism Project, my primary responsibility involved developing visualizations related to nations that experienced the highest number of attacks. The focal point was the overall representation of the Global Terrorism Database, spanning the years 1970 to 2017. The visualization underscored Iraq as the most affected country, requiring meticulous handling of data cleaning challenges. This included standardizing country names and addressing unmatched values to ensure data integrity.



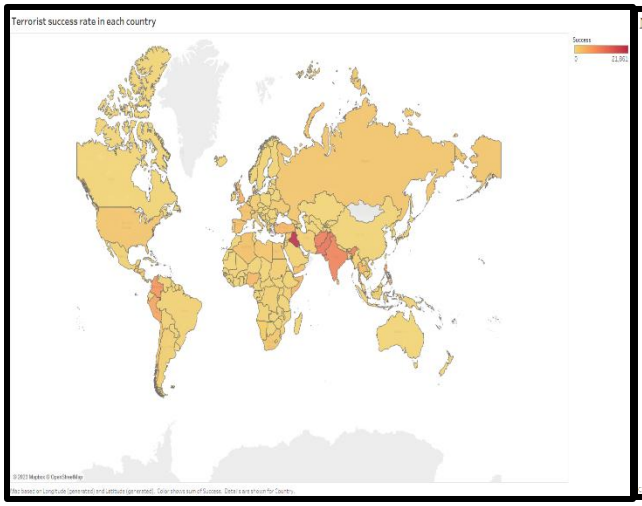


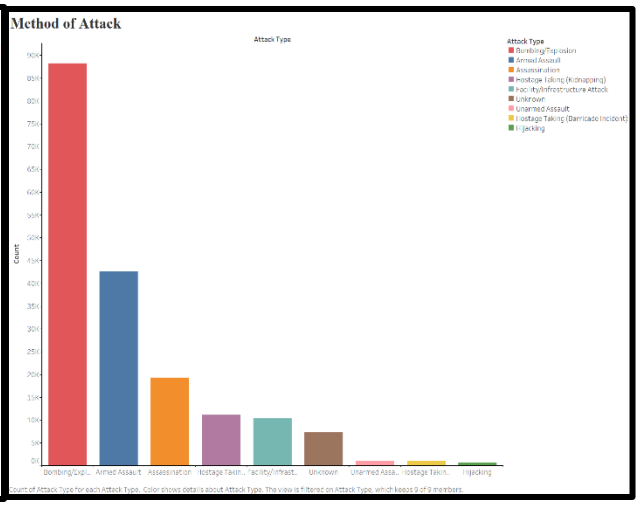
The experience provided valuable insights into the disparities in attack counts among countries, emphasizing the varying degrees of vulnerability and resilience. The process of handling data cleaning challenges further enriched my understanding of ensuring data reliability for meaningful analyses. This project highlighted the critical role of meticulous data preparation in extracting accurate and insightful information from large-scale datasets.

Terrorism Success Rates and Variable Analytics

In my role for the Global Terrorism Project, my focus centered on visualizations related to the success rate of terrorism in each country. Interpreting the data, I highlighted significant terrorist activity in the Middle East and South Asian regions, particularly in countries like India, Pakistan, Afghanistan, and Iraq. The visualizations effectively underscored the scale of terrorist strikes, with Iraq experiencing a substantial number of attacks.

This task deepened my understanding of regional variations in terrorism and emphasized the importance of visualizing success rates for effective analysis. The visualizations provided valuable insights into the

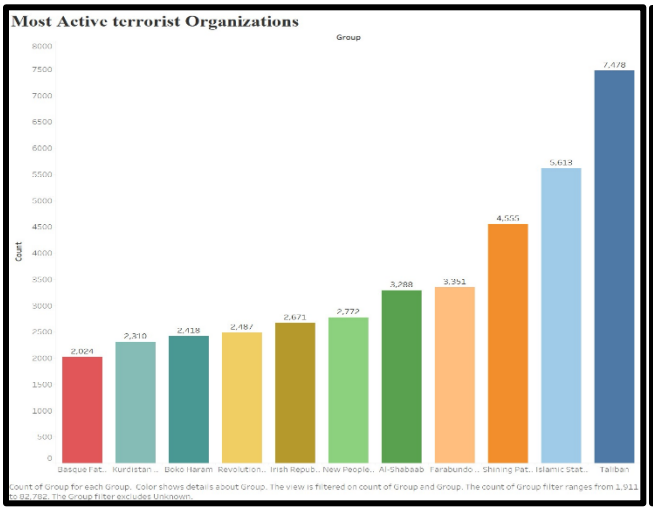


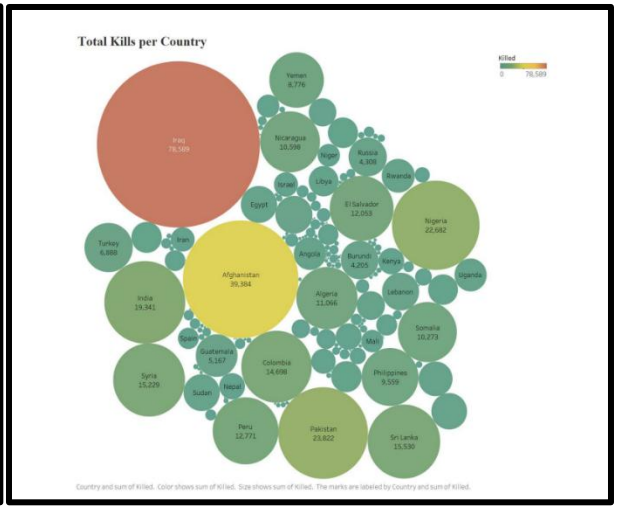


geopolitical landscape of terrorism, aiding in the identification of high-risk areas. This project significantly advanced my skills in translating complex data into meaningful visual representations that contribute to a deeper understanding of global issues.

Conveying the Human Cost Through Total Deaths Visualizations

As a team member on the Global Terrorism Project, my primary responsibility was to contribute to visualizations portraying the total deaths per country. The key visualization, a bubble chart, effectively communicated the toll of terrorist attacks, with Iraq bearing the highest death toll followed by Afghanistan. This visualization provided a poignant perspective on the human cost of terrorism.





This broadened my skills in creating impactful visualizations that go beyond raw numbers, conveying the emotional and societal impact of terrorism. The bubble chart not only showcased the disparities in death tolls but also highlighted the countries most affected. This project reinforced the importance of conveying complex information in a visually accessible manner and contributing to a more comprehensive understanding of global security challenges.