MILESTONE REPORTTHREAT MATRIX FORECASTING TERRORIST INCIDENTS

Vinay Khandelwal, Tanmay Rewari, Kapil Singla, Pavithra Shetty

**Introduction and Our Problem:**

To estimate the chance of a successful terrorist strike based on information gathered from various sources, we will use a dataset on both domestic and foreign terrorism. Information on the date, time, and place of each incident recorded in the Global Terrorism Database is accessible, together with details about the weapons used, the kind of target, the number of victims, and, if validated, the group or individual responsible. We aim to identify variables associated with success, such as attack type (e.g., bombing, assassination), weapon type (e.g., explosive, biological, vehicle), and target (e.g., maritime, diplomatic, private person, police), using data gathered on over 180,000 distinct attacks.

**The Dataset -:**

[**https://www.rand.org/nsrd/projects/terrorism-incidents.html**](https://www.rand.org/nsrd/projects/terrorism-incidents.html)

[**https://reliefweb.int/report/world/global-terrorism-index-2022**](https://reliefweb.int/report/world/global-terrorism-index-2022)

[**https://www.start.umd.edu/gtd/**](https://www.start.umd.edu/gtd/)

**Original Purpose and Publisher:**

The RAND Database of Worldwide Terrorism Incidents (RDWTI) is a valuable resource spanning from 1968 to 2009, offering detailed information on international and domestic acts of terrorism. With contributions from various sponsors, it features over 40,000 coded incidents, demonstrating its remarkable quality and completeness for its time. The database is freely accessible for research and analysis, with attribution to the RAND Database of Worldwide Terrorism Incidents required.

The Global Terrorism Index (GTI) by the Institute for Economics & Peace provides a comprehensive summary of global terrorism trends over the past 14 years, ranking 163 countries based on the impact of terrorism. It relies on data from sources like Terrorism Tracker, which records terrorist events since 2007, encompassing over 60,500 incidents from 2007 to 2021.

The Global Terrorism Database (GTD) is an open-source repository containing information on terrorist events worldwide from 1970 through 2021, with plans for annual updates. Unlike other event databases, the GTD covers domestic, transnational, and international terrorist incidents and now includes over 200,000 cases. Each incident in the GTD is well-documented, providing details on dates, locations, weapons used, casualties, and the responsible group or individual. The data is derived from credible open media sources, and the database is made available by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) to facilitate the study and understanding of terrorist violence.

**Selection of a Database: The Global Terrorism Database (GTD)**

RAND Database of Worldwide Terrorism Incidents (RDWTI) resource spanning from 1968 to 2009,

Global Terrorism Index (GTI) resource spanning from 2007 to 2021 and The Global Terrorism Database (GTD) containing information from 1970 through 2021. Since GTD has highest number of years covered and more than two hundred thousand data points and 135 variables we are choosing GTD as our database to work on.

**Data Descriptions :-**

The GTD was designed to gather a variety of situational variables pertaining to each terrorist incident. Depending on availability of information, the database records up to 135 separate attributes of each incident, including approximately 75 coded variables that can be used for statistical analysis. These variables include not limited to information on location, date, person targeted, corporation or government agency targeted, type of attack, organization responsible, criteria of the attack as an act of war, and the length of attack. The data set has 214,000 plus examples of specific terrorist events that occurred and a wide variety of information gathered on each attack.

**Our Problem - Variables for Study:**-

There are 135 categories of variables in the data collection, such as time, location, and description variables. More recent events will have more information available in the database than events from earlier in the time period, based on the methods of data collecting and publications that are currently available. The following are a few instances of significant variables that were noted about terrorist incidents in the database:

* Region, Nation, Province/State, City, Latitude, Longitude
* Group of Perpetrators Name and attack strategy employed
* Target characteristics (type, subtype, and maximum three targets)
* Weapon type (type and subtype, maximum three weapon kinds) utilized
* Total number of dead (individuals, citizens of the United States, terrorists)
* An indication of the national or international nature of the attack

**Motivation for Selection:-**

The urgent security concerns raised by the recent terrorist acts in Israel, Palestine, and the Ukraine have greatly inspired our team. These attacks have highlighted how urgent it is to improve our knowledge of the elements that lead to these kinds of attacks and create prediction models that can help stop or lessen similar instances in the future.

**Research on Terrorism and Relevant Projects:-**

The issue of terrorism has maintained its prominence on the global stage for several decades, with international organizations and tensions spanning the globe, often resorting to violence as a means to achieve their political or personal objectives. Various entities have collected data on terrorism, though many of them are either private companies or government agencies that tend to keep their data confidential. To facilitate our research, we chose to leverage the Global Terrorism Database (GTD), which provides valuable insights into the codebook, data collection methods, and methodology.

The Global Terrorism Database, hosted at the University of Maryland, offers a comprehensive resource for understanding terrorism trends and patterns. By examining their online supporting documents, we gained essential knowledge about their data collection techniques and the methodology used to classify incidents as acts of terrorism.

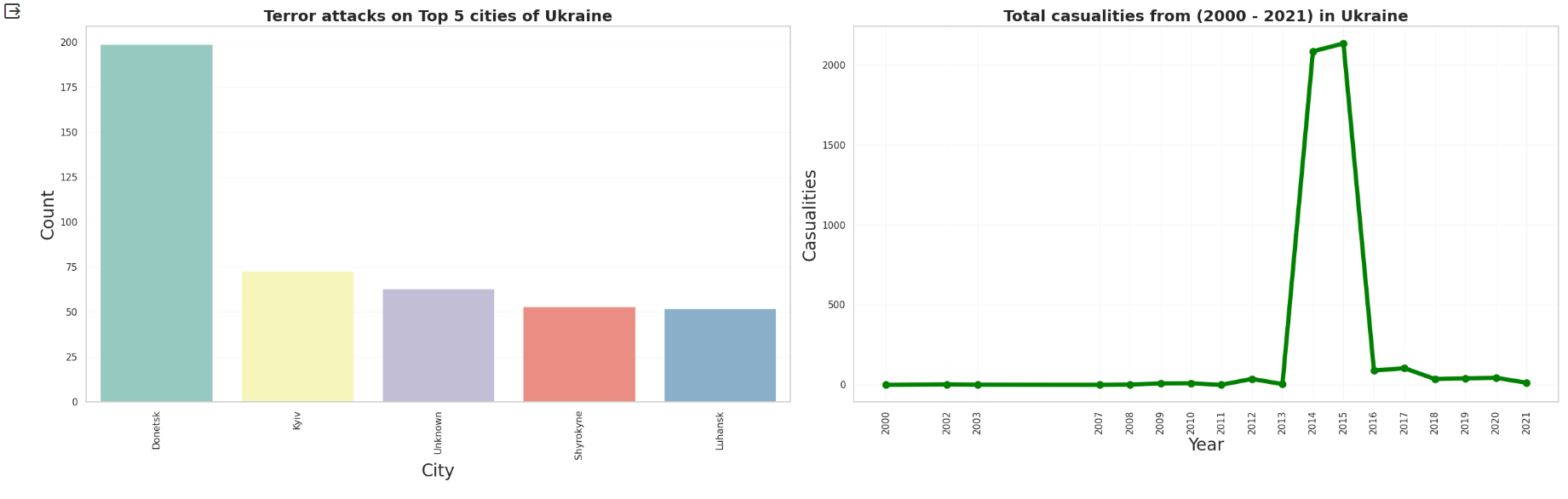
**Other Projects and Works:-**

In addition to the GTD, we also turned to other authoritative sources to gain a deeper understanding of current terrorism trends and insights into specific issues related to different countries. The U.S. Department of State and the National Counter Terrorism Center proved to be invaluable resources. These sources provide country-specific information on terrorist issues, shedding light on evolving trends. We believe that this background research will enhance our understanding of various aspects of terrorism, including the types of attacks, the regional organizations responsible for these acts, and the potential for their success.

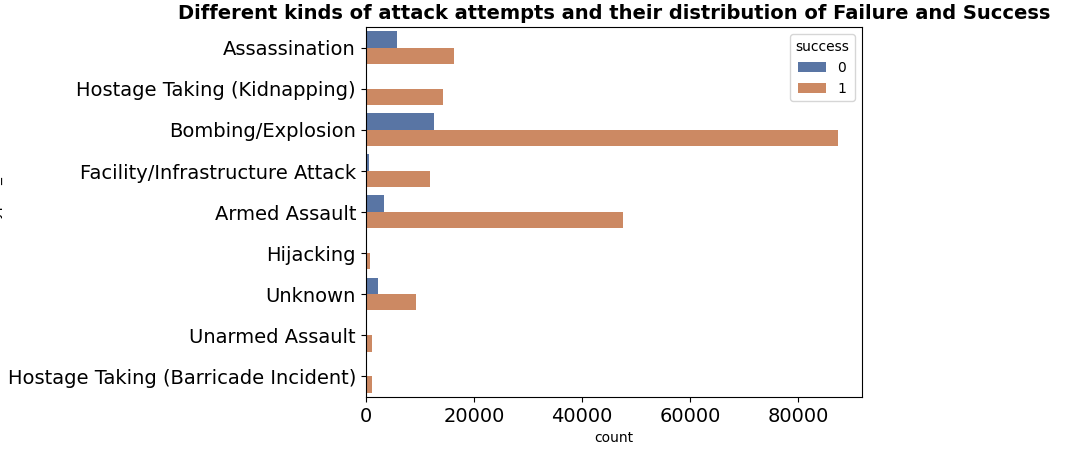
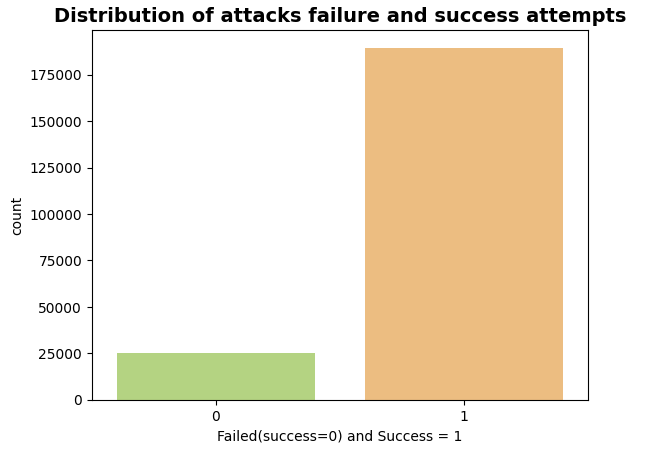
By tapping into these resources, we aim to gather information that will be reflected in our dataset. Key topics discussed in our research include countries experiencing turmoil, such as Syria and Somalia, as well as insights into the preferences and activities of organizations like Al-Qaeda, al-Shabaab, and Hezbollah, which are on the rise. These elements contribute to a comprehensive understanding of the landscape of terrorism, enabling us to make informed and data-driven analyses.

**Exploratory Data Analysis:-**

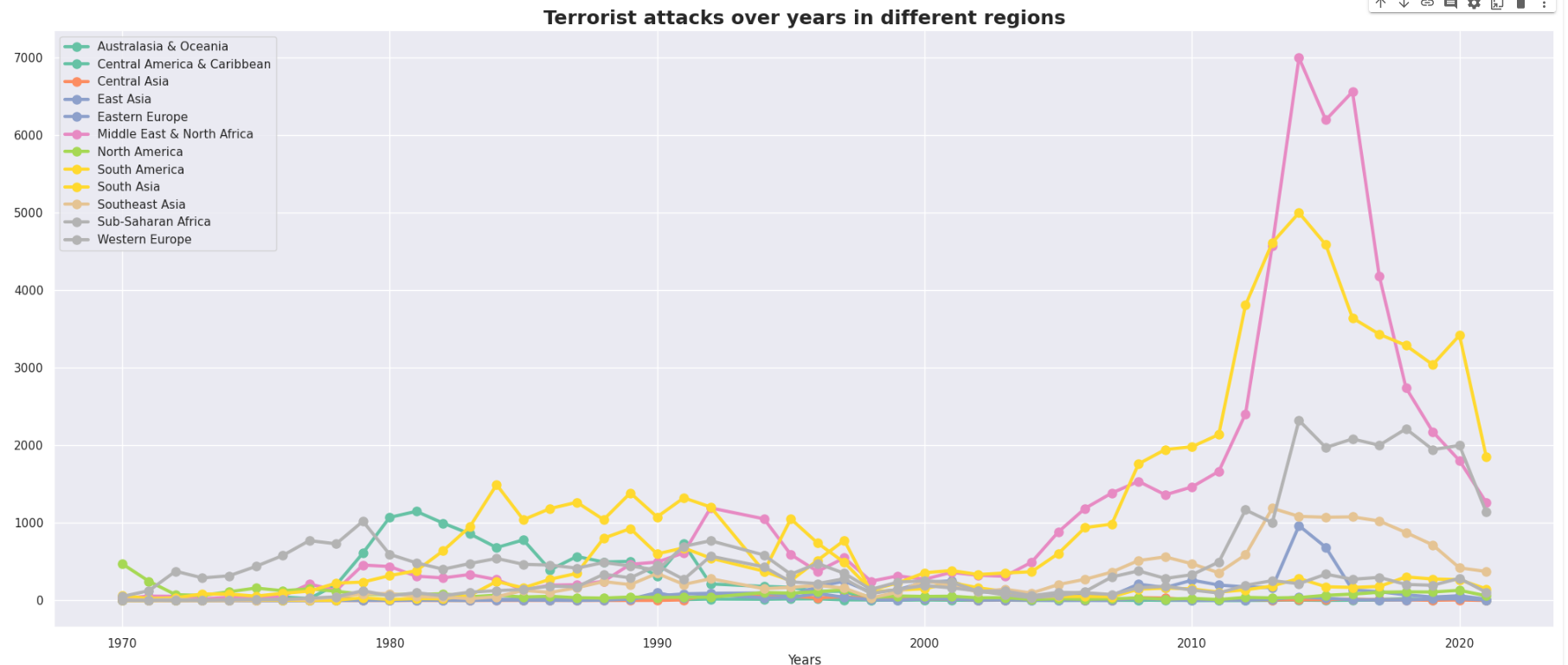
In the course of exploring and analyzing the Terrorism Threat Matrix dataset, a comprehensive data cleaning process was undertaken to ensure the integrity and reliability of the information. The dataset spans multiple years, providing a historical perspective on global terrorist activities.



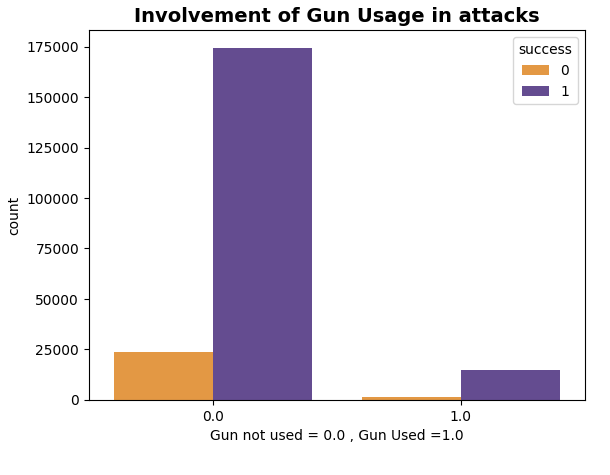
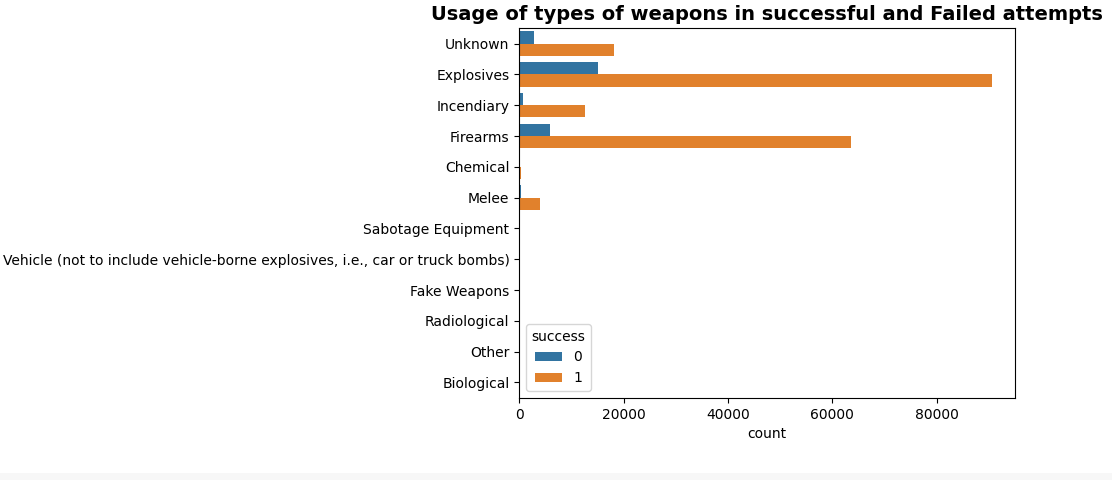
The initial data exploration unveiled several intriguing patterns and trends. Notably, the geographical distribution of terrorist incidents across regions revealed varying levels of threat, with certain regions experiencing higher frequencies than others. The temporal analysis further highlighted shifts in threat dynamics over the years, underscoring the evolving nature of global terrorism.



One key focus of the analysis was the correlation between different features, such as the relationship between attack types and casualty counts. This revealed valuable insights into the tactics employed by terrorist groups and their corresponding impact on human casualties. Additionally, the examination of specific countries and their respective threat landscapes offered a nuanced understanding of the geopolitical factors influencing terrorism.



A significant aspect of the analysis was the identification of factors contributing to the success or failure of terrorist incidents. By examining variables such as the use of firearms and the success of attacks, a clearer picture emerged regarding the modus operandi of different groups.

In conclusion, the EDA process has not only provided a comprehensive overview of the Terrorism Threat Matrix dataset but has also laid the groundwork for more advanced analyses. The insights gained from this exploratory phase will inform subsequent modeling and predictive efforts, contributing to a more robust understanding of terrorism dynamics and aiding in the development of effective counterterrorism strategies.