A black rectangle with a black background

Description automatically generated with low confidenceLogo, company name

Description automatically generated

DAEN 690

Project Report

Namratha Alapati

Sanisha Kolanu

Priyadarsini Deverapalli

Akash Reddy Tummala

Ninad Bhavesh Daftary

Spring 2024

Terrorism Dashboard

**About the Cover**

Professor Berlin is an instructor at the George Mason University College of Engineering and Computing, Volgenau School of Engineering, MS Data Analytics Engineering (DAEN) program. He began working with the DAEN program as an adjunct faculty member in 2012 and became a fulltime faculty member in 2016. He is a passionate contributor to the program and a devoted mentor to his students.

His passion for new value creation is built on over 50 years of professional experience – innovating and advocating for innovators applying leading-edge digital solutions to mission challenges. He has served with outstanding teams in various roles, including senior strategy executive, consultant, and mentor; applied information and systems technologist; collaborative leader; computer scientist, and public policy entrepreneur.

He serves as a strategy advisor and mentor to public and private sector innovators and entrepreneurs and as a public speaker (emerging challenges, innovation opportunities, and ethics). His core interests include public policy, high-performance computing, cyber, emerging big data, health informatics, and digital economy and governance challenges.

In addition to teaching and mentoring, Professor Berlin seeks new engagements with high-quality, core-value-centered innovation teams – collaborating to address societal and market challenges with cyber-physical and policy innovation. Specifically, sustainable solutions can be delivered at the intersection of innovative value creation, human aspiration, and strategic vision.

Professor Berlin is a graduate of the US Air Force Academy with a Bachelor of Science in Computer Science and Mathematics as well as a graduate of the University of Texas at Austin with a Master of Arts in Computer Science. He is also attended the USAF Air Command and Staff College for leadership and strategy training.

Contents

Table of Contents

[Section 1: Problem Definition 3](#_Toc165843838)

[1.1 Background 3](#_Toc165843839)

[1.2 Problem Space 6](#_Toc165843840)

[1.3 Research 7](#_Toc165843841)

[1.4 Solution Space 8](#_Toc165843842)

[1.5 Project Objectives 9](#_Toc165843843)

[1.6 Primary User Stories 9](#_Toc165843844)

[1.6.1 Policy Maker 10](#_Toc165843845)

[1.6.2 Law Enforcement Officer 10](#_Toc165843846)

[1.6.3 Humanitarian Aid Planner 10](#_Toc165843847)

[1.7 Product Vision 10](#_Toc165843848)

[1.7.1 Scenario #1 10](#_Toc165843849)

[1.7.2 Scenario #2 10](#_Toc165843850)

[Section 2: Datasets 11](#_Toc165843851)

[2.1 Overview 11](#_Toc165843852)

[2.2 Field Descriptions 11](#_Toc165843853)

[2.3 Data Context 22](#_Toc165843854)

[2.4 Data Conditioning 22](#_Toc165843855)

[2.5 Data Quality Assessment 22](#_Toc165843856)

[2.6 Other Data Sources 23](#_Toc165843857)

[2.7 Storage Medium 24](#_Toc165843858)

[2.8 Storage Security 24](#_Toc165843859)

[2.9 Storage Costs 24](#_Toc165843860)

[Section 3: Algorithms & Analysis / ML Model Exploration & Selection 25](#_Toc165843861)

[3.1 Solution Approach 25](#_Toc165843862)

[3.1.1 Systems Architecture 25](#_Toc165843863)

[3.1.2 Systems Security 26](#_Toc165843864)

[3.1.3 Systems Data Flows 26](#_Toc165843865)

[3.1.4 Algorithms & Analysis 27](#_Toc165843866)

[Section 4: Visualizations / ML Model Training, Evaluation, & Validation 27](#_Toc165843867)

[4.1 Overview 27](#_Toc165843868)

[4.2 Visualizations 28](#_Toc165843869)

[Section 5: Findings 53](#_Toc165843870)

[Section 6: Summary 53](#_Toc165843871)

[Section 7: Future Work 53](#_Toc165843872)

[Appendix A: Glossary 55](#_Toc165843873)

[Appendix B: GitHub Repository 56](#_Toc165843874)

[Overview 56](#_Toc165843875)

[GitHub Repository Link 56](#_Toc165843876)

[GitHub Repository Contents 56](#_Toc165843877)

[Appendix C: Risks 57](#_Toc165843878)

[Sprint 1 Risks 57](#_Toc165843879)

[Sprint 2 Risks 58](#_Toc165843880)

[Sprint 3 Risks 59](#_Toc165843881)

[Sprint 4 Risks 60](#_Toc165843882)

[Sprint 5 Risks 61](#_Toc165843883)

[Appendix D: Agile Development 63](#_Toc165843884)

[Scrum Methodology 63](#_Toc165843885)

[Sprint 1 Analysis 63](#_Toc165843886)

[Sprint 2 Analyis 63](#_Toc165843887)

[Sprint 3 Analysis 64](#_Toc165843888)

[Sprint 4 Analysis 64](#_Toc165843891)

[Sprint 5 Analysis 64](#_Toc165843892)

[Appendix E: Automatic dataset Download 66](#_Toc165843893)

Table of Figures

[Figure 1 Act of Terrorism [3] 3](#_Toc165234954)

[Figure 2 Remains of a twin tower after the 9/11 attack. [7] 4](#_Toc165234955)

[Figure 3 Dataflow diagram 25](#_Toc165234956)

[Figure 4 World map presenting number of terrorism incidents by color. 28](#_Toc165234957)

[Figure 5 Casualities stacked bargraph of each organization. 29](#_Toc165234958)

[Figure 6 Line graph of number of incidents for each category. 30](#_Toc165234959)

[Figure 7 Line graph of facilities targeted in each incident. 30](#_Toc165234960)

[Figure 8 Line graph of tactics used in each incident. 31](#_Toc165234961)

[Figure 9 Line graph of victime targeted in each incident. 31](#_Toc165234962)

[Figure 10 Line graph of logistic used in each incident. 32](#_Toc165234963)

[Figure 11 Line graph of weapon used in each incident. 32](#_Toc165234964)

[Figure 12 Line graph of victims affected by the each incident. 33](#_Toc165234965)

[Figure 13 Bar graph comparing the categories of Iran backed and Jihadist organizations. 34](#_Toc165234966)

[Figure 14 Iran backed and Jihadist target facilities preferences. 34](#_Toc165234967)

[Figure 15 Iran backed and Jihadist tactics preferences. 35](#_Toc165234968)

[Figure 16 Iran backed and Jihadist logistics preferences. 35](#_Toc165234969)

[Figure 17 Iran backed and Jihadist weapon preferences. 36](#_Toc165234970)

[Figure 18 Iran backed and Jihadist's victims. 36](#_Toc165234971)

[Figure 19 Tactics used by lone actors. 37](#_Toc165234972)

[Figure 20 Victims of Lone Actors 38](#_Toc165234973)

[Figure 21 Wepons used by Lone Actors 38](#_Toc165234974)

[Figure 22 Tactics used by Jihadist 39](#_Toc165234975)

[Figure 23 Victims of Jihadist 39](#_Toc165234976)

[Figure 24 Weapons used by Jihadist 40](#_Toc165234977)

[Figure 25 Tactics used by Right-Wing Extremists 40](#_Toc165234978)

[Figure 26 Victims og Right-Wing Extremists 41](#_Toc165234979)

[Figure 27 Wepons used by Right-Wing Extremists 41](#_Toc165234980)

[Figure 28 Tactics used by Vigilante. 42](#_Toc165234981)

[Figure 29 Victims of Vigilante. 42](#_Toc165234982)

[Figure 30 Weapons used by Vigilante. 43](#_Toc165234983)

[Figure 31 Correlation between terrorist organization and the tactics they used. 44](#_Toc165234984)

[Figure 32Correlation between other terrorist organization and the tactics they used. 44](#_Toc165234985)

[Figure 33 Top factors contributing to severity of terrorist attacks. 45](#_Toc165234986)

[Figure 34 Top 10 factors for severity of terrorist attacks. 46](#_Toc165234987)

[Figure 35 Correlation between people killed factors contributing it. 46](#_Toc165234988)

[Figure 36 Correlation between victims and tactics. 47](#_Toc165234989)

[Figure 37 Correlation between weapons and tactics. 48](#_Toc165234990)

[Figure 38 Home page of dashboard. 49](#_Toc165234991)

[Figure 39 State map 49](#_Toc165234992)

[Figure 40 Sate information. 50](#_Toc165234993)

[Figure 41 List of terrorist organizations. 50](#_Toc165234994)

[Figure 42 List for type of casualities. 51](#_Toc165234995)

[Figure 43 Information with respect to the type of casuality. 51](#_Toc165234996)

[Table 1: Glossary Table 55](#_Toc165234997)

[Table 2: Sprint 1 Risks 57](#_Toc165234998)

[Table 3: Sprint 2 Risks 58](#_Toc165234999)

[Table 4: Sprint 3 Risks 59](#_Toc165235000)

[Table 5: Sprint 4 Risks 61](#_Toc165235001)

[Table 6: Sprint 5 Risks 62](#_Toc165235002)

[Figure 44: Sprint project dates. 63](#_Toc165235003)

**This page intentionally left blank**

Abstract

This project utilizes the extensive dataset from the Global Terrorism Trends and Analysis Center (GTTAC) from the year 2018. It aims to provide insights into global terrorism trends, focusing on evolving tactics, targets, and operational capabilities of terrorist groups. The analysis identifies hotspots of terrorist activities, particularly in Afghanistan and Syria, and compares various perpetrator groups, including Iran-backed, Jihadist organizations, lone actors, right-wing extremists, and vigilante groups. By applying advanced data analytics and Tableau visualizations, the study examines incident patterns and operational strategies. Notably, religious, and ethnonationalist/separatist groups exhibit significant operational capacity. The research also explores the correlation between perpetrator types and tactics used in terrorist attacks, highlighting the increasing reliance on exploitative methods by jihadist groups. Additionally, the project aims to create a publicly accessible General Terrorism Dashboard for real-time data visualization, enhancing insights into terrorist threats globally.

The analysis reveals a significant operational capacity among religious and ethnonationalist/separatist groups. These groups outperform others in terms of tactics, weapons usage, and impact on victims, indicating their extensive operational reach and activity level. The research also investigates the relationship between perpetrator types and the tactics employed in terrorist attacks. Notably, there is a positive correlation between jihadist groups and exploitative tactics, suggesting an increasing reliance on such methods as their frequency grows.

A crucial outcome of this study is the initial development of a future publicly accessible General Terrorism Dashboard. This dashboard provides real-time visualizations of GTTAC data, translating intricate terrorism trends into clear, actionable insights for scholarly analysis. By enhancing the accessibility and usability of terrorism data, this initiative offers comprehensive insights to counter terrorist threats. Our research underscores the intricate global terrorism landscape and emphasizes the indispensable role of data-driven analysis in navigating this challenging domain. These valuable perspectives contribute to the ongoing discourse on terrorism worldwide.

**This page intentionally left blank**

Report

# Problem Definition

## Background

The act of causing physical, property damage or loss of life using violence with the aim to intimidate or stir fear in public for their religious, political and their ideological beliefs which is performed by an individual or an organization who may have an affinity with the government of other country/state. The acts of terrorism were present throughout the history of mankind, but it was coined as the word terrorism in the 1790s during French Revolution. [1] [2]

A large explosion in a city

Description automatically generated

Figure 1 Act of Terrorism [3]

Different people and countries have different definitions of terrorism, and the term "terrorist" that goes along with it is arbitrary and context dependent. When individuals rebel against repressive regimes, this relativity is especially clear since what one person views as a justifiable fight for freedom is another person's definition of terrorism. The proverb "One man's terrorism is another man's freedom," which Gerald Seymour eloquently expressed in his 1976 novel "Harry's Game," captures this duality. The quote highlights the subjective aspect of characterizing acts of violence or resistance. It also highlights the significance of considering varying viewpoints when assessing the characteristics and intentions of political movements or individuals involved in such actions. Title 22, Section 2656f, U.S. Code, criteria is used to classify an incident as a terrorist act, if any violent act committed by nonstate actors with the principal intent of advancing social, political, economic, or religious objectives is considered a terrorist act. Also, there must be unmistakable proof that the conduct was performed with the goal of intimidating, coercing, or sending a message to a group of people other than the direct victims. In addition, the conduct must specifically target non-combatants and take place outside the bounds of international humanitarian law. [4] [5]

The most recent Annual Annex of Statistical Information highlights the widespread threat posed by international terrorism, listing 42,836 occurrences over the previous five years that have tragically claimed 134,118 lives. Aside from these terrifying incidents that were documented during this time, the September 11, 2001, attack on the Twin Towers in the United States stands out as particularly noteworthy. This terrorist attack, which was masterminded by the Islamic extremist organization al-Qaeda, is regarded as one of the saddest and most impactful incidents in contemporary history, having taken the lives of roughly 2,977 individuals. The attack's severe loss and lingering effects bring to light the continued difficulties in combating and stopping terrorism across the globe. [6]

A building with a few debris

Description automatically generated with medium confidence

Figure 2 Remains of a twin tower after the 9/11 attack. [7]

In 2022, the landscape of global terrorism changed dramatically from 2012, with four terrorist groups accounting for roughly half of all deaths, up from 7% a decade ago. Despite an overall decrease in the number of terrorism-related fatalities, attacks were more deadly, with an average of 1.7 deaths per occurrence. Notably, terrorist groups increasingly avoided taking responsibility, hampering efforts to monitor their activities. Key players in this shifting dynamic included the Taliban, which transitioned from insurgency to becoming state actors in Afghanistan after capturing Kabul in 2021. ISIS remained a considerable danger, expanding into Sub-Saharan Africa's Sahel region, while Al-Shabaab increased its activities in East Africa, particularly Somalia. The Balochistan Liberation Army's quest of regional autonomy resulted in a substantial increase in terrorism-related deaths, while Jamaat Nusrat Al-Islam wal Muslimeen had a steady but decreasing impact in the Sahel. These trends highlight the complexities and shifting nature of global terrorism, necessitating sophisticated approaches to handle growing threats and mitigate their impact on impacted regions. [8]

The first center in the USA to investigate and comprehend the intricate connections between corruption, international crime, and terrorism is the Terrorism, Transnational Crime and Corruption Center (TraCCC). TraCCC is committed to increasing the body of knowledge in these fields by instruction, training, research, and active involvement in the creation of policy. The Anti-Illicit Trade Institute (AITI), a prestigious program that tackles and opposes illicit trade using cutting-edge strategies, is also hosted by TraCCC. Under the leadership of Development Service Group (DSG), the TraCCC’s staff and students are working on five-year project called The Global Terrorism Trends and Analysis Center (GTTAC). The primary objective of this project is to provide data and analysis to the U.S. State Department’s congressionally mandated Statistical Annex for the Annual Country Reports on Global Terrorism. [9] [10]

Through the Worldwide Incidents Tracking System, the National Counterterrorism Center, a division of the Office of the Director of National Intelligence, gathered data on terrorist incidents across the globe from 2004 to 2011. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) was brought in by the Department of State in June 2012 to carry out this duty. Through the use of its Global Terrorism Database, STARTT produced the Annex of Statistical Information on Terrorism till 2017. DSG and its subcontractor, the TraCCC at George Mason University, were given the contract by the Department of State in August 2018. The GTTAC is presently run by the international research firm DSG, which also edits the Country Reports on Terrorism and produces the Annex of Statistical Information. [11]

In order to identify possible terrorist situations, GTTAC serves as a platform for gathering information from five extensive open-source multimedia data aggregators. Using ontologies for different aspects of terrorist attacks, GTTAC does data processing, text analysis, predictive modeling, and feature extraction using open-source technology tools based on Python on a Linux platform. A proficient, multilingual team of data analysts completes incident generation and validation after automated procedures, yielding an extensive database of international terrorism incidents. This database serves as the foundation for a site created for the Department of State, which includes the Annex of Statistical Information that goes with the yearly Country Reports on Terrorism that are required by Congress. [12]

To provide a more impartial and truthful account of terrorist attacks, especially in cases when governments have a motive to minimize the severity of these acts, the TraCCC employs a methodical approach to gathering data. Understanding that governments may downplay the number of incidents or damages done by these incidents to not appear weak on security or to protect their tourism sector, TraCCC only gets its news from unbiased news sources that are not connected to any government agencies. In order to ensure a more transparent and all-encompassing assessment of the extent and impact of terrorism worldwide, free from potential official interference, this data is verified by students affiliated with TraCCC. On top of that according to Title 22, Section 2656f, U.S. Code, if the violent actions are performed by the people of government or by the order of government itself then those incidents are not considered into dataset. For example, in Afghanistan, when the Taliban took over, all the incidents from that point onwards are not recorded into dataset as those acts are committed by their ruling government.

All this data on terrorism incidents collected can be used to perform data analysis, by analyzing data on terrorist attacks provides several benefits that are critical for effective counterterrorism efforts. Policymakers and security agencies get vital insights into emerging terrorist methods, favored targets, and geographical concentrations of activity by detecting patterns and trends across time. This insight helps to build focused strategies and provide resources where they are most needed. The ability to identify hazards and implement early warning systems improves readiness, allowing for proactive measures to avert or lessen future threats. Furthermore, data analysis promotes evidence-based policymaking, enabling the development and evaluation of counter-terrorism initiatives. It also promotes international collaboration by offering a common platform for information exchange, resulting in a better knowledge of global terrorist networks.

Furthermore, data analysis helps with resource allocation, ensuring the efficient deployment of security troops and the implementation of counter-terrorism actions. Understanding terrorist groups' motivations and ideology is critical for designing focused counter-radicalization strategies. Furthermore, honest communication based on reliable facts raises public awareness, dispels myths, and builds support for counterterrorism activities. Authorities can improve counter-terrorism efforts over time by measuring the impact and effectiveness of their programs. Finally, new techniques like as predictive modeling improve the ability to identify future dangers, allowing for a proactive and preventive strategy to managing developing risks. In summary, data analysis is a foundation for evidence-based decision-making, resource optimization, and international collaboration in the complex terrain of counterterrorism.

## Problem Space

Following the impactful events of the 9/11 attacks, the ongoing challenge of terrorism requires continuous vigilance and adaptability. Despite a respite from major attacks on US soil in recent years, the specter of terrorism remains a global concern, particularly in regions like Africa, the Middle East, and Asia. To unravel the complexities of this multifaceted threat from 2018 to 2023, our research relies on the wealth of information meticulously recorded in the GTTAC Record of Incident Database (GRID).

The global landscape of terrorism is marked by a myriad of organizations, each with its unique dynamics and operational reach. Al-Qaeda, a prominent jihadist network, has evolved and expanded its influence across diverse regions, including Yemen, Syria, Afghanistan, Nigeria, and Mozambique. Despite debates surrounding its power, Al-Qaeda has showcased adaptability by forming affiliations with entities like Al Shabaab, JNIM, and HTS. Concurrently, the Islamic State of Iraq and Syria (ISIS), known for its brutal tactics and aspiration to establish a caliphate, has affiliates operating globally. From Boko Haram in Africa, particularly Nigeria, to the Taliban in South Asia with a stronghold in Afghanistan and parts of Pakistan, each organization exhibits distinct characteristics. Hezbollah, rooted in Lebanon and backed by Iran, focuses on regional conflicts, particularly against Israel. This includes an assessment of their operational capabilities, and an understanding of the tactics they employ such as bombings, kidnappings, executions, hijacking, cyber-attacks, only property damage, and guerrilla warfare. [13] [14]

Various terrorist organizations recognize the influence of digital platforms. In addition to Al-Qaeda and ISIS, groups like Hamas and Hezbollah have also utilized internet propaganda to disseminate their ideologies and advance their agendas. These organizations leverage social media platforms to reach a global audience, enabling them to recruit new members and propagate their messages. Understanding the tactical nuances of each organization is crucial for developing comprehensive counterterrorism strategies that address the multifaceted nature of the global terrorist threat. [13] [15]

Previous group work analysis centered around dissecting specific groups like Al-Qaeda and ISIS, comparing their incident trends, casualty counts, and target preferences. It primarily used descriptive statistics to track changes in attack numbers and perpetrator involvement over time. This year, however, we're diving deeper. We're expanding the scope to encompass the entire global terrorist landscape, analyzing various groups' operational capacities across tactics, targets, and victim types. We're also looking at dynamic changes from 2018 to 2023, examining how tactics, weapons, and victim profiles have evolved. Additionally, we're comparing broader ideological groups like Iran-backed and jihadist organizations, investigating their similarities and differences. By exploring the motivations and ideologies of lone actors, analyzing the relationship between perpetrator types and attack methods, and even examining how weapon and target choices influence tactics, this year's research aims to illuminate not just the "what" but also the "why" and "how" behind terrorist activities, ultimately contributing to more informed and effective counterterrorism strategies.

Our quest for knowledge doesn't end there. We delve into the factors that influence the severity of attacks, seeking to understand what makes some incidents deadlier than others. Weapon and target selection also come under scrutiny, as we explore how these choices might influence the tactics employed by terrorist organizations.

Ultimately, this research aims to illuminate the intricate web of global terrorism, providing valuable insights for policymakers, security analysts, and researchers alike. By unraveling the threads of this complex phenomenon, we can better equip ourselves to counter its threats and build a safer future for all.

## Research

Integrating a diverse range of sources, including academic research, fieldwork, and government records, is pivotal in constructing a nuanced understanding of terrorism and transnational crime. Scholarly articles, grounded in extensive research and fieldwork, delve into the motivations, ideologies, and operational strategies of terrorist organizations, providing a comprehensive view of their underpinnings and methodologies. These studies are instrumental in exploring the complex dynamics of terrorism, offering insights into its causes, the psychological profile of terrorists, and the socio-political contexts that foster such activities. References to academic work, such as "Understanding Terrorism: Challenges, Perspectives, and Issues" by Gus Martin [16], provide theoretical frameworks and empirical analysis that contribute to the broader discourse on counterterrorism strategies and policy implications.

Moreover, government publications, notably those from the Global Terrorism Trends and Analytics Center (GTTAC), like the Annex of Statistical Information, serve as a crucial resource for current data on terrorist incidents, tactics, and counterterrorism operations. These documents offer a macroscopic view of terrorism's global landscape, detailing the activities of major terrorist groups, their geographic spread, and the international efforts aimed at curbing their influence. The GTTAC reports are invaluable for researchers and policymakers alike, offering a data-driven basis for understanding trends in terrorism and the effectiveness of various counterterrorism measures [17].

The synthesis of academic insights and government-compiled data enables a holistic approach to studying terrorism. It facilitates a multidimensional analysis that encompasses both the theoretical aspects of terrorism and the practical challenges of combating it on the ground. This integrated approach is crucial for developing effective security policies and strategies, as it combines the depth of academic research with the breadth of real-world data. For instance, the work of Bruce Hoffman in "Inside Terrorism" [18] offers an in-depth analysis of the evolution of terrorism and complements the statistical data provided by government reports, thereby enriching our understanding of both the historical context and the current state of global terrorism.

The latest figures from the GTI 2023, we find that attacks have become more deadly, with the lethality of attacks increasing by 26%. Islamic State remains the deadliest terrorist group, with activities in 21 countries, reflecting the changing dynamics of global terrorism [19]. The Sahel has become particularly impacted, representing 43% of global terrorism deaths [20]. These figures underscore the evolving nature of terrorism, necessitating adaptive and informed counterterrorism strategies.

Our research methodology unfolds in a stepwise manner, guided by a meticulously outlined plan that aligns with the project's objectives, available resources, and predefined timelines. This structured approach not only facilitates a systematic exploration of the terrorism landscape but also ensures the coherence and effectiveness of our research endeavors. Central to our project is the analysis of terrorism's global footprint, drawing upon GTTAC's rich repository of data to uncover patterns, trends, and insights that span across different regions and terrorist entities.

The project's analytical framework is bolstered by scholarly contributions, notably those by Dr. Mahmut Cengiz, whose research delves into the comparative analysis of threats posed by distinct terrorist organizations. Through a critical examination of literature and empirical data, we aim to unravel the complex dynamics of power and control that underpin the operations of these groups. Dr. Cengiz's work, among others, provides a critical lens through which we assess the strategic motivations and operational tactics of terrorist organizations, enriching our understanding of their foundational drivers [21].

Furthermore, the project recognizes the importance of geopolitical factors in shaping the landscape of terrorism. By analyzing regional trends and the influence of international politics on terrorism, the project offers a nuanced perspective on how global and local dynamics converge to facilitate or hinder terrorist activities. This analysis is particularly relevant in the context of understanding how state-sponsored terrorism and the international arms trade contribute to the proliferation of terrorism. The project's findings underscore the complexity of combating terrorism, which requires not only military and security-based approaches but also diplomatic and socio-economic interventions [22].

The application of data analytics and dashboard stands out as a cornerstone of the project. This technological approach enables the identification of significant predictors of terrorism, such as the type of weapon used, the nationality of the perpetrators, and the location of attacks.

In conclusion, the project exemplifies the power of data analytics in understanding and combating terrorism. By synthesizing data from GTTAC with advanced analytical methods, the project illuminates the multifaceted nature of terrorism and provides actionable insights professionals. The project's comprehensive approach, which integrates ideological, operational, and geopolitical analyses, contributes significantly to the field of terrorism studies. It highlights the need for a multifaceted strategy in addressing terrorism, one that combines intelligence-led initiatives with efforts to address the root causes of radicalization and violence [23].

## Solution Space

In this project, our solution space is carefully crafted to align with the overarching goals of identifying, assessing, and visualizing global terrorism trends. A central focus is placed on robust data management and processing, involving the acquisition of extensive terrorism data from the GTTAC Data Portal [9], supplemented by additional data as needed. To ensure accessibility and security, a third-party Development Service Group (DSG) [24] maintains, integrates, and stores this data in a physical storage solution, utilizing on-premises servers.

Our analytical framework employs a combination of statistical analysis to identify trends and patterns, predictive modeling for forecasting, and qualitative analysis to understand terrorist narratives and motivations.

A pivotal component of our solution is the creation of an interactive and user-friendly dashboard. We utilize tools like Tableau [25], chosen for their compatibility with project goals. As we rely on physical storage, the data is stored in dedicated hardware. This dashboard is more than a visualization tool; it serves as a portal to deep insights, offering real-time data visualization capabilities for dynamic and up-to-date information presentation.

Security and compliance are paramount, with stringent measures in place to protect sensitive data and adhere to legal and ethical standards. The user interface is designed for private local accessibility, catering to a diverse audience, including clients, research teams, and policymakers.

The solution's architecture is scalable and capable of handling increasing data volumes and user traffic. The DSG [24] has a plan for ongoing maintenance and feature enhancements. Regular data updates are facilitated through an automated system downloading data from the GTTAC portal [9] every Monday morning, ensuring the dashboard remains current without relying on cloud-based solutions.

Finally, the project underscores the significance of insightful reporting and knowledge dissemination. Comprehensive reports generated from the dashboard findings will be crucial for academic, policymaking, and awareness purposes, making a substantial contribution to the field of terrorism studies. This solution aims not only to address the core research questions but also to provide impactful insights informing counterterrorism strategies, and enhancing public understanding of global terrorism trends.

## Project Objectives

To deepen understanding of the intricate dynamics within transnational organized crime, including its ties to corruption and terrorism, extensive research and analysis will be conducted. This endeavor will explore the interconnectedness of various criminal activities such as money laundering and illicit trade with terrorism, while also delving into factors like state fragility and their repercussions on global security. By uncovering patterns and root causes of organized crime, this project aims to shed light on its role within social conflicts, providing valuable insights to policymakers, analysts, and researchers involved in counterterrorism and crime prevention efforts.

Examining the dataset will reveal recurring themes and patterns in terrorist incidents, considering elements like the scene, employed strategies, and casualty numbers. Understanding these fundamental patterns will enable decision-makers and security organizations to craft focused counterterrorism plans, allocate resources efficiently, and implement preventive measures to mitigate potential risks.

Assessing the claimed responsibility and attribution accuracy for terrorist incidents recorded in the dataset is paramount. This evaluation allows decision-makers to arrive at informed conclusions about the perpetrators of terrorist acts, aiding in resource allocation for combating emerging threats and crafting evidence-based counterterrorism policies.

Analyzing the dataset to ascertain details such as date, location, strategies employed, and casualty numbers related to terrorist incidents is crucial for comprehending the dynamics, patterns, and trends of terrorist activity. This understanding is essential for developing effective counterterrorism strategies and allocating resources more efficiently among policymakers, analysts, and security agencies.

Utilizing data analytics techniques to enhance counterterrorism strategies and operations is a key objective. Leveraging data analytics tools and methodologies enables the analysis of global terrorism trends, identification of emerging threats, and evaluation of counterterrorism measures' effectiveness. Predictive modelling to anticipate future terrorist activities supports proactive intervention strategies.

Evaluating the impact of terrorism and transnational crime on socioeconomic development aims to understand how these phenomena exacerbate social inequality, impede development initiatives, and fuel economic instability. By examining the relationships between crime, terrorism, and socioeconomic conditions, proactive measures can be taken to lessen their negative effects on societies worldwide.

Strengthening international coordination and cooperation in the fight against terrorism and transnational crime is essential. This involves assessing current frameworks and procedures, identifying opportunities for improvement, and addressing gaps in data exchange, legal frameworks, and operational coordination. Practical recommendations will be developed to enhance information sharing, expedite joint operations, and bolster capacity-building initiatives, leading to a more cohesive strategy against transnational threats.

## Primary User Stories

As per the project’s objectives, the following user stories have been created:

### Policy Maker

As a Policy Maker, I will analyze the data of past incidents to make policies which will effectively allocate resources to prevent any further terrorism incidents.

### Law Enforcement Officer

As a Law Enforcement Officer, I will analyze the data which will help me in investigating and help in solving future terrorism incidents.

### Humanitarian Aid Planner

As a Humanitarian Aid Planner, whenever there is an incident, I can look into this data to determine the extent of damage done so that I can effectively allocate resources for the aid.

## Product Vision

### Scenario #1

A governmental body is actively working to improve national security by honing its policies and tactics to more effectively counter transnational threats. Drawing on knowledge gained from a thorough investigation, the organization incorporates conclusions clarifying the complex relationships among international organized crime, corruption, and terrorism into its framework for formulating policy. By incorporating these insights, the agency will be able to create more sophisticated and effective strategies. The agency can also predict and anticipate future terrorist activities by using advanced data analytics techniques, which helps to prioritize proactive interventions and optimize resource allocation. The result of this coordinated effort is a strengthened national security apparatus that can counter new threats in a fast-changing environment.

### Scenario #2

In the context of promoting international cooperation, a prestigious multilateral organization is committed to fostering member state collaboration to tackle global security challenges. In keeping with this promise, the group hosts cooperative seminars and conferences that promote the sharing of cutting-edge strategies and best practices for countering international threats. Moreover, consistent with knowledge gained from an extensive research project, the organization designs its capacity-building activities to target vulnerabilities in partner countries. The effectiveness and applicability of the help provided are guaranteed by this strategic alignment. The result of these efforts is an improved response on a global scale to transnational threats, which in turn leads to increased security and resilience on a global scale.

# Datasets

## Overview

The Global Terrorism Tracking and Analysis Center (GTTAC) has taken an open-source approach, making data on terrorist incidents publicly available via its site. The initial stage for GTTAC team students is to review publications to find situations that match particular criteria. These requirements include the involvement of non-state or non-government actors, the targeting of unarmed civilians, and the act being carried out for political, social, or religious purposes. Once recognized, students upload incident information onto an online form called the GTTAC Record of Incident Database (GRID).

Students complete the form and submit it for verification, which is reviewed by a GTTAC team professor. The data is included to the dataset once the professor has acknowledged it. The database is managed by the DSG team and can be downloaded for free via the GTTAC webpage. There are three datasets available: incidents, perpetrators, and sources. The events dataset contains information on when and where terrorist incidents happened, with each row indicating a separate incident. The perpetrators dataset focuses on individuals who have committed terrorist crimes, as well as their linked groups. Both datasets additionally include information on the weapons used, the tactics used, the target and affected victims, and the targeted facilities. The source dataset includes information about the news stories from which incident details were taken, as well as the date of publication. To aid comprehension, a codebook is given, which includes a brief description of each column in the dataset. This complete technique offers transparency and accessibility while analyzing global terrorist trends.

## Field Descriptions

**Weapon Types Classification System**

**Level 1 Classification (Parent Category):**

This level categorizes weapons into broad groups based on their general type or mode of operation. Examples might include:

• Firearms

• Explosives

• Improvised Explosive Device

• Unmanned Aerial Vehicle (UAV)

• Incendiary

• Melee

• Other

• Unknown

**Level 2 Classification (Specific Attribute):**

Under each broad category in Level 1, Level 2 provides more specific details, identifying particular types or models of weapons within the general category. For instance:

**Firearms**

• Small Arms (Handguns, Long rifles, Intermediate Caliber rifles)

• Full-automatic (machine Guns, RPK, RPD)

• Improvised/Homemade

• Heavy Weaponry (Anti-tank rifle, .50 Caliber rifle, large Caliber rifle)

• Unclear/Unknown Firearm

**Explosives**

• Unclear/Unknown Explosives

• Rocket Artillery (Not RPG)

• Mortars / non-rocket artillery

• Conventional explosives/ Munitions (C4, Dynamite)

• Grenades (RPG, Non-RPG, Hand grenades)

• Mines (Anti-personnel, Anti-vehicle)

• Missiles

**Improvised Explosive Devices**

• Suicide Vest IED (Suicide bomber)

• Non-Suicide Vest IED

**Unmanned Aerial Vehicle (UAV)**

• Unclear/ Unknown UAV

• Military UAV

• Commercial UAV

**Incendiary**

• Unclear/Unknown incendiary

• Molotov Cocktail/Petrol Bomb

• Gasoline/Oil/Diesel

• Manufactured Incendiary Device/Weapon

**Melee**

• Unclear/Unknown Melee

• Knives/Swords/Machetes

• Blunt Objects

**Other**

• Fake Weapons

• Vehicle

• Chemical, Biological, Radiological, Nuclear (CBRN/WMD)

**Data Types and Contents:**

Each weapon type is represented by a binary variable (1 = used in an attack, 0 = not used) in the dataset.

**Column Names:**

Column names should be descriptive of the weapon type, such as "weapon\_lvl1\_fi", " weapon\_lvl1\_ot", " weapon\_lvl1\_ua", with further specificity in Level 2 classifications like "weapon\_lvl2\_fi\_02", "weapon\_lvl2\_ua\_03", etc.

**Tactics Types Classification System**

**Level 1 Classification (Parent Category)**

This level includes broad categories of tactics used in terrorist attacks. Examples include:

• Trauma

• Assault

• Covert

• Coordinated

• Exploitive

• Unknown

**Level 2 Classification (Specific Attribute)**

Under each Level 1 category, Level 2 provides finer granularity, specifying types of tactics within the broader category. For example:

**Trauma**

• Gender-based violence and rape

• Executions

• Bodily Mutilations

• Suicide Bombing

• Kidnapping

• Car-Ramming

• Hostage taking

• Assassinations

• Stabbing

• Hijacking

**Assault**

• Shooting

• Storming/Rapid Assault

• Drive-By

• Bombing (car bomb, pipe bomb, TNT)

• Unarmed Assault (sticks or rocks)

**Covert**

• Booby-Trapping

• Ambush

• Sabotage

• Planting Mines/IEDs

• Disguise/Infiltration

**Coordinated**

• One Country (Coordinated in one country)

• Multiple Countries (Coordinated in multiple countries)

• One location (Coordinated in one location)

• Multiple locations (Coordinated across multiple locations)

• One team (Coordinated with one team of perpetrators)

• Multiple teams (Coordinated across multiple teams of perpetrators)

**Exploitive**

• Use of foreign fighters as perpetrators

• Use of female fighters as perpetrators

• Use of children fighters as perpetrators

• Involvement of illicit trade

• Use of extortion practices (ransom, extortion, robbery)

• Use of corruption (bribery, ill-gotten, tenders, illegal permits)

• Use of money laundering processes

• Property damage only

• Cyber attack

**Data Types and Contents:**

Each tactic type is represented by a binary variable (1 = used in an attack, 0 = not used) in the dataset.

**Column Names:**

Column names should be descriptive of the tactic type, such as "tactic\_lvl1\_tr", " tactic\_lvl1\_co", " tactic\_lvl1\_ex", with further specificity in Level 2 classifications like "tactic\_lvl2\_tr\_02", "tactic\_lvl2\_ex\_10", etc.

**Logistics Types Classification System**

**Level 1 Classification (Parent Category)**

These are broad categories within the logistics sector, focusing on the types of operations or services involved.

• Vehicle

• Mail/Postage

• Other

• Unknown

• None (No logistics types was found or used)

**Level 2 Classification (Specific Attribute)**

Under each Level 1 category, Level 2 provides more detail about specific operations or logistics types.

**Vehicle**

• Plane

• Helicopter

• Civilian Vehicle

• Military Vehicle (Jeep, Car, Truck)

• Military Armed Vehicle (Tank, Armour car)

• Improvised Armored/Technical Fighting Vehicle (Unarmored military pickup truck )

• UAV

• Public transportation/Government Vehicle

• Boat

• Bicycles and Motorcycles

• Unclear/Unknown

**Mail/Postage**

• Targeted

• Random Targets

**Other**

• Hand propelled (thrown or dropped)

• Kite

• Ballon

• Slingshot

• Aerosols

• Animal Borne

**Data Types and Contents:**

Each logistics type is represented by a binary variable (1 = used in an attack, 0 = not used) in the dataset.

**Column Names:**

Column names should be descriptive of the logistics type, such as "logistic\_lvl1\_ve", " logistic\_lvl1\_mp", " logistics\_lvl1\_ot", with further specificity in Level 2 classifications like "logistic\_lvl2\_ve\_05", "logistic\_lvl2\_mp\_01", etc.

**Facility Types Classification System**

**Level 1 Classification (Parent Category):**

These are broad categories within the facility sector, focusing on the types of buildings involved.

• Commercial

• Culture

• Government

• Infrastructure

• Military Facility

• Miscellaneous

• Unclear/Unknown

• None (If there was no specific facility type)

**Level 2 Classification (Specific Attribute):**

Under each Level 1 category, Level 2 provides more detail about specific building or facility types.

**Commercial**

• Individual retail

• Offices

• Tourist Facilities

**Culture**

• Entertainment/Sports

• Media

• Religious sites

• Tourist sites

**Government**

• Buildings/Vehicles

• Diplomatic

• Law enforcements & Intelligence

**Infrastructure**

• Agriculture

• Civilian Aviation

• Communications, Cyber/Digital

• Construction & Factory

• Education

• Healthcare Facilities

• Land Transport

• Maritime & Fishing

• Private property/Residences

• Public places

• Relief

• Utilities & Mining

**Military Facility**

• Military Facility -Foreign

• Military Facility -Multi-National

• Military Facility -National

• Military Facility -US

**Miscellaneous**

• Political organizations

• Property of PEPs (Politically exposed persons)

• Perpetrators

• Pro-Government forces

• Violent non-state actors

**Data Types and Contents:**

Each facility type is represented by a binary variable (1 = used in an attack, 0 = not used) in the dataset.

**Column Names:**

Column names should be descriptive of the facility type, such as "facility\_lvl1\_co", " facility\_lvl1\_mi", " facility\_lvl1\_in", with further specificity in Level 2 classifications like "facility\_lvl2\_ms\_03", "facility\_lvl2\_go\_01", etc.

**Victim Types Classification System**

Each record has two victim variables: **Actual Victim** and **Intended Victim**. The actual Victim is the person whom the perpetrators actually attacked, injured, and/or killed. The intended Victim is the person whom the perpetrators targeted.

**Level 1 Classification (Parent Category):**

This level includes broad categories of victims that might be affected in incidents, such as:

• Groups

• Government

• General Population

• Military

• Political

• Professions

• Other

• None (No victims)

• Foreign Nationals

**Level 2 Classification (Specific Attribute):**

Under each Level 1 category, Level 2 provides finer granularity, specifying types of victims within the broader category. For example:

**Groups**

• Identified by Gender

• Identify by Race

• LGBTQIA

• Refugees/IDPS

• Children/Students

• Identify by religion (Muslim, Hindu, Jewish, Christian, Buddhist, and other)

**Government**

• Civil servants & Local government

• Foreign Diplomatic personnel

• High-level officials/head of state

• Intelligence (non-military)

• Judiciary

• Law enforcement

• Legislators

**Military**

• Military – Foreign (doesn’t include NATO or US)

• Military – Multi-national (NATO, ATMIS, but not US)

• Military - national

• Military - US

**Political**

• Candidates

• Election Personnel

• Parties

• Political Demonstrations

**Professions**

• Agriculture

• Commerce & Trade

• Cultural Figures

• Education

• Journalism & Media

• Maritime & Fisheries

• NGO & Relief

• Private Security

• Tourism

• Transportation

• Religious Figures

• Utilities and Minning

• Medical (Abortion related, emergency-related responders, non-emergency)

**Other**

• Processions (Funerals, Weddings, Religious)

• Labour Unions

• Perpetrators

• Pro-Government forces

• Other Violent non-state actors

**Data Types and Contents:**

For tracking purposes, each victim type can be represented similarly with a binary variable (1 = affected, 0 = not affected) in the dataset.

**Column Names:**

Column names should be descriptive of the victim type, such as "victim\_lvl1\_pr", " victim\_lvl1\_gp ", " victim\_lvl1\_ci ", with specific attributes in Level 2 classifications like " victim\_lvl2\_pr\_06 ", " victim\_lvl2\_mi\_03 ", " victim\_lvl2\_go\_02\* ", etc.

**Incident Identification:**

In this we get the details of the incident by using their unique 9 digits number. The incident also has a 7-digit number if the incident has multiple events happened within the incident and the multiple events are linked to the sequential incident numbers.

**Data Type:**

Each type is represented by a Numeric Variable.

**Columns Name:**

Columns name are the descriptive of Incident Identification, such as unique\_incident\_id, incident\_id, incident\_seq.

**Source and Publication Information:**

In this we get the information that from where the data is collected like the article, date it is published.

**Data Type:**

Here the data is represented by string type and Date variable.

**Column Names:**

Column names are the descriptive forms of source and publication Information, such as headline, publication\_date, publication\_source.

**Incident Summary:**

We get the information of the incident description, and it contains information if the incident report is drawn from a multiple source or a single source.

**Data Type:**

Type of data that is used in this section is string and binary variable.

**Column Names:**

Column names are the description of the Incident summary, such as summary, is\_multi\_source\_corr.

**Perpetrator Information:**

This tells us about the information if the perpetrator has claimed the event of incident or the Incident, along with the unique code assigned to it and the name.

**Icat- Perpetrator Category**

* Anarchist
* Enthnonationalist/ Separatist
* Iranian-Backed
* Left-Wing/Revolutionary Extremists
* Right-Wing Extremists
* Religious (Other- Hindu, Buddhist, Jewish, Christian)
* Religious (Jihadist)
* Single Issue
* Unknown
* Vigilante (Pro-Government)

**Data Type:**

Here the data formats are Binary, Categorical, and String variable

**Column Names:**

Column names are the descriptive form of the perpetrators, such as claimed, perp\_cd, perp\_name, Icat.

**Casualty Information:**

This gives us information about the number of people killed including the victims and perpetrators. Also, a separate information on number of US citizens killed, wounded, hostage, Kidnapped, perpetrators wounded and killed. Here in these columns the default value is 0 if there is no record of killed.

**Data Type:**

Each type is represented by a Numeric Scale Variable.

**Column Names:**

Column names are the descriptive of Casualty Information, such as num\_killed, num\_killed\_perp, num\_killed\_us, num\_wounded, num\_wounded\_perp, num\_wounded\_us, num\_hostkid, num\_hostkid\_us.

## Data Context

The Global Terrorism Trends and Analysis Center (GTTAC) [1] datasets used in our project have a comprehensive framework for categorizing and documenting terrorist incidents provided by the GRID Codebook [1]. It includes detailed data dictionaries for incident and perpetrator files, specifying variables such as incident identification, dates, summaries, locations, perpetrator information, casualty figures, weapon types, and logistical details. This framework is critical for the accurate and consistent recording of global terrorism incidents, facilitating analysis, research, and policy development aimed at understanding and combating terrorism. It also outlines the definitions, inclusion criteria for terrorist incidents, and the categorization of incidents with terrorist attributes (VITA) [1]. It covers specific types of incidents such as mass shootings, state violence, and violence by non-state actors, along with guidelines for classifying incidents in conflict zones. For our analysis of the project, we consider information on regions, countries, and perpetrator groups with significant operational capabilities, focusing on their preferred weapon types, tactics, targets, logistics, and victim types.

## Data Conditioning

The data conditioning for our terrorism dashboard project involved a thorough cleaning of the GTTAC Database to remove duplicates and rectify missing or inconsistent entries. We are planning to transform data formats for compatibility and merge incident, perpetrator, and source files for a comprehensive analysis. Ensuring high data quality and maintaining detailed documentation of our process were key to providing a reliable foundation for our Tableau dashboard.

## Data Quality Assessment

GRID\_PERPS

|  |  |
| --- | --- |
|  |  |
| Completeness | Check crucial fields like "perp\_name" and "num\_killed\_perp" for any missing or null values and they don’t have any missing or null values in them. |
| Consistency | Ensure that data is formatted consistently and meets domain standards. Check for uniform date formats in the incident\_date field and codes for is\_ied field. |
| Uniqueness | Ensure that each record in the dataset is distinct using unique identifier perp\_incident\_id. |
| Integrity | Verify dataset links and connections, particularly between characteristics like "perp\_incident\_id" and "unique\_incident\_id." |
| Conformity | Determine if data meets set criteria. Compare numeric ranges (e.g., num\_killed\_perp, num\_hostkid), date ranges, and categorical values (e.g., tactical\_lvl2\_as\_03, victim\_lvl1\_ci) using criteria. |
| Accuracy | Validate data values in fields such as "num\_killed\_perp" and "num\_hostkid" to ensure accuracy and eliminate mistakes. |

GRID\_INCIDENTS

|  |  |
| --- | --- |
| Completeness | Check crucial fields like "incident\_id" and "incident\_month" for any missing or null values and they don’t have any missing or null values in them. |
| Consistency | Ensure that data is formatted consistently and meets domain standards. Check for uniform date formats in the incident\_date field and codes for is\_multi\_dayfield. |
| Uniqueness | Ensure that each record in the dataset is distinct using unique identifier unique\_incident\_id. |
| Integrity | Verify dataset links and connections, particularly between characteristics like "unique\_incident\_id" and "incident\_id." |
| Conformity | Determine if data meets set criteria. Compare numeric ranges (e.g., num\_killed\_us, num\_hostkid\_us), date ranges, and categorical values (e.g., facility\_lvl2\_cu\_02, intended\_victim\_lvl1\_ci) using criteria. |
| Accuracy | Validate data values in fields such as "num\_killed\_us" and "num\_hostkid\_us" to ensure accuracy and eliminate mistakes. |

GRID\_SOURCES

|  |  |
| --- | --- |
| Completeness | Check crucial fields like "unique\_incident\_id" and “incident\_id” for any missing or null values and they don’t have any missing or null values in them. |
| Consistency | Ensure that data is formatted consistently and meets domain standards. Check for uniform date formats in the publication\_date field. |
| Uniqueness | Ensure that each record in the dataset is distinct using unique identifier unique\_incident\_id. |
| Integrity | Verify dataset links and connections, particularly between characteristics like "unique\_incident\_id" and "incident\_id." |
| Conformity | Determine if data meets set criteria. Compare numeric range like incident\_seqand date like publication\_date. |
| Accuracy | Validate data values in fields such as "publication\_date”, “publication” and “headline” to ensure accuracy and eliminate mistakes. |

## Other Data Sources

Our project solely relies on data acquired from the Global Terrorism Trends and Analysis Center (GTTAC), a collaborative effort between the Development Services Group (DSG) and the Terrorism, Transnational Crime and Corruption Center (TraCCC) of George Mason University, as the subprime. Also, the database contains information from famous publications like “BBC Monitoring”, “The Times of Israel”, “New York Times”, “Khaaama Press”, and many more. By focusing solely on the GTTAC dataset, we ensure data accuracy, also reducing the risk of inconsistencies or errors that can be generated by merging the datasets. It also reduces the chances of having biases inherent in different data collection methodologies found in diverse datasets.

## Storage Medium

Our project relies on dedicated hardware storage mediums to store GRID datasets, prioritizing high levels of data security and fast, reliable access within a localized setting. By storing data on dedicated hardware within personal computers, we ensure physical control over storage media, thereby minimizing the risk of unauthorized external access and protecting sensitive information, particularly related to terrorism incidents within GRID datasets. This setup also guarantees swift and dependable access to datasets, reducing dependencies on external services and internet connectivity and mitigating potential downtime. While there is an initial investment in dedicated hardware, the predictable data storage needs of our project make this approach cost-effective over time. In conclusion, leveraging dedicated hardware storage solutions offers a comprehensive solution that meets our project's key requirements, ensuring data security, accessibility, and efficiency.

## Storage Security

A few key procedures must be followed to guarantee the security of a manually downloaded dataset on a local system. First, keeping the dataset in a secure folder or directory, away from places that are open to the public or from directories with lax permissions, helps limit access to authorized users only, lowering the possibility of exposure to unauthorized parties. The system on which the dataset is stored must have strong password protection as the first line of defense against unwanted access. Frequent backups kept in different, safe places offer protection against data loss or corruption, guaranteeing the availability and integrity of the data. Additionally, to fix known vulnerabilities and fortify defenses against possible threats, it's crucial to keep software including the operating system, encryption tools, and security software up to date with the most recent patches and updates. In addition to digital security measures, physical security measures reduce the possibility of unauthorized access and tampering, such as keeping the system in a locked cabinet or room. By permanently deleting private data from the storage device, secure dataset deletion helps to prevent unwanted access and data breaches. Finally, educating users about fundamental security best practices gives them the ability to identify and successfully reduce possible security threats.

## Storage Costs

When a dataset is downloaded manually and saved for free on a laptop's hardware, it means that keeping the data on the device has no direct financial consequences. There are no extra costs associated with using the laptop's internal storage, such as the hard drive or solid-state drive, beyond the cost of the laptop itself.

# Algorithms & Analysis / ML Model Exploration & Selection

## Solution Approach

### Systems Architecture

A diagram of a computer system

Description automatically generated

Figure 3 Dataflow diagram

The architecture of the system is designed to create a seamless workflow that effectively manages and processes data to generate visual representations. The Development Service Group (DSG) initiates this process by gathering data from various global sources. Concurrently, the Global Terrorism Trends and Analysis Center (GTTAC) Team engages with the database to commence the data ingestion operations into grid , which include the critical tasks of validation and publication of the data into the GTTAC Portal.

In the primary phase, the DSG is responsible for compiling and depositing the data into a central repository. This repository serves as the initial storage point, ensuring raw data is accessible for future processing steps. The central database is interfaced with the GTTAC Portal, which acts as a conduit for the GTTAC Team to validate the data and incorporate it into Grid datasets as new information becomes available. Parallelly, data is downloaded into a channeled Local Desktop by an automated Python script to download data.

Once the data is cleaned and processed, it undergoes Data Integration, where it is likely integrated with other relevant data sets by joining with their foreign keys to provide a unified view. Following integration, the data is available for Data Visualization & Analysis, a stage that transforms raw data into actionable insights through various visualization techniques and analytical processes.

The process includes a feedback loop from Data Visualization & Analysis to Data Integration and Data Cleaning & Processing, indicating iterative refinement based on the review and adaptation of the insights generated.

Should the updates from the feedback loop prove satisfactory, the process moves to the Deployment phase. This stage implies the operationalization of the data-driven insights, potentially through implementing changes or informing decision-making in the relevant operational environment.

This architecture outlines a comprehensive system designed to facilitate the continuous transformation of terrorism data into valuable insights, emphasizing iterative improvement and operational deployment of terrorism trends visualizations.

### Systems Security

Placing a high priority on data accuracy and security, we managed all research tools and databases on our personal systems, thereby mitigating the risk of data breaches from third-party providers or exposure to unauthorized external access. This self-contained approach to system management is foundational to our workflow as a data analyst, incorporating robust security measures at every level. Encrypting our device's hard drive, for instance, ensures that the data remains secure, even in scenarios of physical loss or theft. Regular updates to the operating system and applications were made, as they often include patches for vulnerabilities that could otherwise be exploited.

When it comes to handling data within Python and Tableau, we adhere to secure coding principles and leverage built-in access controls, respectively. These practices, along with a disciplined backup strategy involving encrypted cloud storage, form a comprehensive security framework. This not only ensures the protection of valuable data but also enhances the efficiency and integrity of our analytical processes.

### Systems Data Flows

The Development Services Group (DSG) utilizes various intelligence techniques to monitor and evaluate terrorism. This includes gathering information from publicly accessible sources like news articles, social media posts, and official websites, known as open-source intelligence. DSG employs robust data analysis techniques to identify patterns and trends in terrorist activity. This analysis involves advanced statistical methods and machine learning algorithms to discern significant insights from the vast amount of data collected.

The analysis conducted by DSG informs critical counterterrorism initiatives such as threat assessments, target identification, and support for law enforcement agency investigations. By identifying emerging threats and assessing the severity of existing ones, DSG plays a crucial role in enhancing national security measures. Once the analysis is completed, the data is transmitted to the GTTAC team. GTTAC (Global Terrorism Trends and Analysis Canter) is responsible for managing the database, editing the annual Country Reports on Terrorism, and preparing the Annex of Statistical Information on Terrorism for the Country Report. This involves organizing and categorizing the data according to relevant parameters, ensuring its accuracy and consistency.

After the data undergoes verification and validation processes to ensure its reliability and integrity, it is integrated into the GTTAC datasets stored in the GTTAC portal. This portal serves as a centralized repository for terrorism-related data, accessible to authorized personnel for research, analysis, and reporting purposes. Access to the datasets is facilitated through an automated process using a Selenium-based Python script. This script streamlines the data retrieval process by automatically filling out access forms and downloading the datasets from the GTTAC portal.

The downloaded datasets are then imported into the tabular server for visualization purposes. Dashboards are subsequently created based on these visualizations.

### Algorithms & Analysis

The script is designed to automate the process of downloading data from the GTTAC portal using Selenium. Selenium is a technology that automates web browser actions. The script was developed using Python, it begins by configuring a Chrome WebDriver and browsing to the GTTAC portal's URL. It then fills out all the mandatory text fields in the form, checks the necessary checkboxes, and starts a download by clicking a submit button while simultaneously managing errors with a try-except block. While waiting for the download to complete, the script checks frequently until a predetermined period is reached. This shows how Selenium interacts with web items, manages wait durations, fills out the text fields, and performs clicks to automate downloading procedures. Furthermore, the script is designed to work on both macOS and Windows operating systems, demonstrating its portability across several platforms to automate the importing of data into local desktop.

# Visualizations / ML Model Training, Evaluation, & Validation

## Overview

In Sprint 4, the team made significant strides in developing the terrorism trends dashboard by incorporating iterative feedback to fine-tune data visualizations, automating data integration for real-time updates, and optimizing performance to manage extensive datasets efficiently. Emphasizing a user-centric approach, stakeholder feedback was integral to refining the dashboard's functionality and interface. Concurrently, proactive risk management strategies were implemented to address potential challenges such as user adoption and project scope expansion. This phase of development culminated in an enhanced dashboard with improved usability, setting the stage for subsequent advancements in functionality and integration with broader data sources in the forthcoming sprint.

## Visualizations

* What regions/countries record the most terrorist incidents?

A map of the world

Description automatically generated

Figure 4 World map presenting number of terrorism incidents by color.

The summary analysis of above Visualization indicates that Afghanistan and Syria are among the countries with the highest number of terrorist incidents in certain years, with Afghanistan recording 1,767 incidents in 2019 and 1,715 incidents in 2020, and Syria recording 1,482 incidents in 2021, 1,426 incidents in 2023, and 1,413 incidents in 2020. This information correlates with the map visualization, where the countries with the darkest shades like red would represent the highest frequency of incidents. These figures highlight regions that have experienced significant terrorist activities in those years.

* What terrorist perpetrator types have the most operational capacity regarding tactic, weapon, and victim types?

A graph of a bar chart

Description automatically generated with medium confidence

Figure 5 Casualities stacked bargraph of each organization.

The series of bar charts depict the operational capacity of various groups, categorized by their motivations (e.g., religious, ethnonationalist, left-wing, right-wing extremism) and delineated by the types of activities they engage in, such as tactics, weapons, and victim impact.

The first chart, "Combined Impact by Group," appears to aggregate the impacts in terms of people killed, wounded, and kidnapped. The second chart, "Operational Capacity by Group: Tactics, Weapons, and Victims," breaks down this impact into separate components to distinguish between the number of victims and the types of tactics and weapons used. The third and fourth charts show detailed counts of wounded and killed, respectively, for each group category.

From these visualizations, we can observe that certain group categories, like religious and ethnonationalist/separatists, have significantly higher operational capacities than others, as evidenced by the higher bars across all aspects: tactics, weapons, and victims. This suggests these groups are not only more active but also have a wider reach in terms of operational methods and their effects on victims. Conversely, groups categorized as vigilantes/pro-government, unknown, or right-wing extremists show a considerably lower operational capacity.

The consistent feature across all charts is the prominence of the first two group categories, which could indicate that religious and ethnonationalist/separatist motivations are significant drivers of high-impact activities. It's also notable that there's a substantial number of incidents labeled as 'Unknown,' which points to several activities where the perpetrators are not identified, leaving a gap in the data.

These charts would be extremely useful in security, policy-making, and strategic planning, as they provide clear indications of where resources may be needed most to prevent or respond to such activities and could guide further research into the methods and motivations of the most active groups.

* How have terrorist tactics, targets, weapons, facilities, logistics, and victim types changed from 2018 to 2023?

A graph of a line graph

Description automatically generated with medium confidence

Figure 6 Line graph of number of incidents for each category.

A graph of a graph showing the number of people attacked by terrorist

Description automatically generated with medium confidence

Figure 7 Line graph of facilities targeted in each incident.

A graph of a number of people

Description automatically generated

Figure 8 Line graph of tactics used in each incident.

A graph of a number of people

Description automatically generated

Figure 9 Line graph of victime targeted in each incident.

A graph of a line graph

Description automatically generated with medium confidence

Figure 10 Line graph of logistic used in each incident.

A graph of a number of people

Description automatically generated with medium confidence

Figure 11 Line graph of weapon used in each incident.

A graph of different colored lines

Description automatically generated

Figure 12 Line graph of victims affected by the each incident.

To acquire a better understanding of the growth of terrorist methods, including tactics, objectives, weapons, facilities, logistics, and types of victims, we used line graphs to show the number of events in each area from 2018 to 2023. These graphs serve as a visual aid for identifying patterns and variations in the frequency of occurrences related with each type of terrorist activity. By mapping these episodes year after year, we may see whether there has been an increase or reduction in activity, which provides insights into terrorists' evolving tastes and approaches.

These line graphs are especially useful for examining the dynamics within each category, demonstrating how the type and scope of terrorist actions have changed over time. For example, an increased trend in the line linked with a given category, such as logistics or weapons, indicates that terrorist groups are relying or emphasizing that element more heavily. In contrast, a negative trend indicates a reduction in prominence or utilization. By painstakingly scrutinizing these trends, we can detect patterns that may not be immediately obvious, providing a more nuanced picture of terrorist operations.

Furthermore, by linking these observable trends with key global events, we might deduce plausible causes of the shifts in terrorist activity. Understanding the impact of key political, social, or environmental events on terrorism allows us to better forecast how future attacks may unfold. This comprehensive approach not only improves our current understanding of terrorism, but it also helps to foresee future patterns, offering significant insights for the creation of counter-terrorism tactics and regulations.

* What are the similarities and distinctions between Iran-backed and jihadist terrorist groups?

A graph with colorful bars and numbers

Description automatically generated

Figure 13 Bar graph comparing the categories of Iran backed and Jihadist organizations.

A graph of a number of people

Description automatically generated with medium confidence

Figure 14 Iran backed and Jihadist target facilities preferences.

A graph with text and numbers

Description automatically generated with medium confidence

Figure 15 Iran backed and Jihadist tactics preferences.

A graph with colorful bars

Description automatically generated with medium confidence

Figure 16 Iran backed and Jihadist logistics preferences.

A graph with text and numbers

Description automatically generated with medium confidence

Figure 17 Iran backed and Jihadist weapon preferences.

A graph with colorful lines and numbers

Description automatically generated

Figure 18 Iran backed and Jihadist's victims.

To compare the similarities and distinctions between the Iran backed and jihadist terrorist groups, a detailed analytical technique was used, which included creating side-by-side bar graphs for each group. These graphs depict a variety of characteristics, including the number of victims affected, facilities targeted, intended victims, logistics used, strategies employed, and weapons used. This systematic approach enables a direct visual comparison of both organizations' operational characteristics, offering a clear understanding of their methods and strategic preferences.

The bar graphs are designed so that the length of each bar corresponds to the number of incidences linked with each category, allowing for a clear contrast between Iran-backed groups and jihadists. By comparing the relative lengths of the bars for each group within a category, one may see not only the overall operational differences but also the peculiarities of each category. This granular research uncovers the distinct qualities and preferences underlying their operational tactics, target selection, and logistical approaches, providing insights into tactical and strategic frameworks.

Beyond revealing differences, this comparative research allows for an assessment of each group's strengths and operational capacities. Understanding the frequency and preference for specific tactics or targets provides insight into the groups' operational priorities and capabilities. As a result, our research not only illuminates the operational distinctions and similarities between Iran-backed and jihadist groups, but also serves as a framework for the development of specific counter-terrorism policies. Policymakers and security professionals can devise more effective methods to counter the unique hazards posed by these entities if they thoroughly grasp each group's operating model.

* What tactics, weapon types, and victim types have lone actors, jihadists, Right-Wing Extremists and Vigilante (Pro-Government) used?

A graph of different colored squares

Description automatically generated

Figure 19 Tactics used by lone actors.

A graph of different colored squares

Description automatically generated

Figure 20 Victims of Lone Actors

A graph with different colored squares

Description automatically generated

Figure 21 Wepons used by Lone Actors

A graph of different colored bars

Description automatically generated

Figure 22 Tactics used by Jihadist

A graph of different colored bars

Description automatically generated

Figure 23 Victims of Jihadist

A graph of different colored bars

Description automatically generated

Figure 24 Weapons used by Jihadist

A graph of a bar chart

Description automatically generated with medium confidence

Figure 25 Tactics used by Right-Wing Extremists

A graph with different colored bars

Description automatically generated

Figure 26 Victims og Right-Wing Extremists

A screenshot of a graph

Description automatically generated

Figure 27 Wepons used by Right-Wing Extremists

A graph of different colored bars

Description automatically generated

Figure 28 Tactics used by Vigilante.

A graph of different colored bars

Description automatically generated

Figure 29 Victims of Vigilante.

A graph with different colored bars

Description automatically generated

Figure 30 Weapons used by Vigilante.

The data presented across the multiple bar charts reveals trends and characteristics of incidents attributed to different actors, such as lone actors, jihadists, right-wing extremists, and vigilante pro-government groups from 2018 to 2023. For lone actors, there appears to be a diverse use of tactics with assault, explosives, and unknown methods being common. The victim categories primarily include the general population and military personnel, suggesting a non-discriminatory targeting pattern. Weapon choices are varied, with a notable presence of firearms and explosive devices.

Jihadist incidents, by contrast, show a high number of assaults and explosive-related tactics with a preference for targeting the general population and government entities. The use of firearms and explosives as weapon types is predominant, reflecting a consistent approach over the years. Right-wing extremists demonstrate a marked increase in incidents in 2021, primarily using explosive and assault tactics. The general population is the most affected victim category, followed by political targets, suggesting a more focused intent. The use of firearms is significantly high, alongside notable incidents involving explosives and melee weapons. Vigilante pro-government groups exhibit a different pattern, with a peak in activity in 2022. Their victimization primarily affects the general population and political figures, indicating specific agendas. Explosives and firearms are their weapons of choice, pointing to organized and deliberate methods. Overall, while each group shows a distinct modus operandi, commonalities exist in the frequent use of firearms and explosives, as well as the predominant targeting of the general population. The data indicates that while tactics and weapons may vary, the impact on civilians remains a consistent and concerning thread throughout the observed period.

* Is there a positive correlation between the terrorist perpetrator types and the attack types?

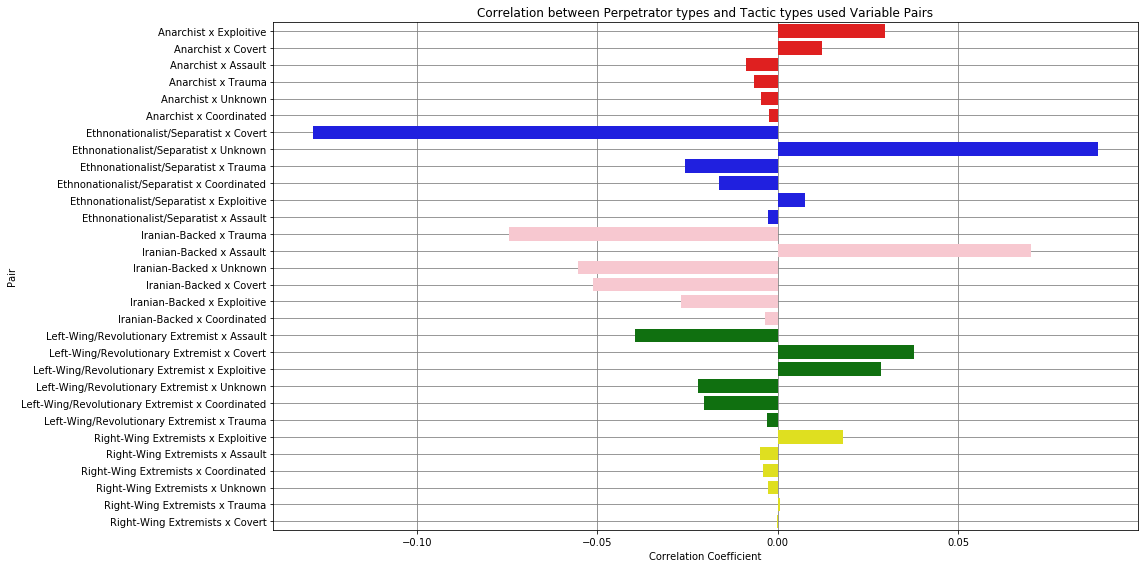


Figure 31 Correlation between terrorist organization and the tactics they used.

A screen shot of a graph

Description automatically generated

Figure 32Correlation between other terrorist organization and the tactics they used.

The above graphs represent correlation coefficients between various perpetrator types and tactic types used in terrorist attacks. The vertical axis lists pairs of categorical variables — perpetrator categories are paired with different tactic types. The horizontal axis represents the correlation coefficient, with positive values indicating a direct correlation and negative values an inverse correlation.

In the first graph, we observe that religious (jihadist) groups have a positive correlation with exploitive tactics, suggesting that as the frequency of jihadist groups increases, there's a tendency for exploitive tactics to be used more often. Conversely, vigilante (pro-government) groups show a negative correlation with coordinated tactics, indicating that these groups are less likely to use coordinated tactics.

The second graph demonstrates similar relationships for other perpetrator types. For instance, ethnonationalist/separatist groups show a strong positive correlation with unknown tactics, suggesting that these groups are more likely to use tactics that are not well-defined or classified in the dataset. On the other hand, right-wing extremists have a negative correlation with covert tactics, implying that these types of attacks are less common among right-wing extremist activities.

* What factors affect the severity of terrorist attacks? (severity is measured by the number of people killed in an incident)

A graph of a number of people

Description automatically generated

Figure 33 Top factors contributing to severity of terrorist attacks.

A graph of orange bars

Description automatically generated

Figure 34 Top 10 factors for severity of terrorist attacks.

A graph showing a bar chart

Description automatically generated with medium confidence

Figure 35 Correlation between people killed factors contributing it.

In an effort to unpack the factors that elevate the severity of terrorist attacks, where severity is quantified by the number of fatalities, our analysis pursued a data-driven approach. The aim was to discern which aspects of an attack—ranging from the type of weapons used to the nature of the target—most significantly impact the death toll.

Leveraging a detailed terrorism dataset, we employed an analytical technique to compute the average number of kills associated with various attack vectors: weapons, tactics, targets, whether the attack was claimed, and whether the incident spanned multiple days or locations. The data was then meticulously segmented, and point-biserial correlations were calculated to determine the strength and direction of the relationships between these vectors and the number of fatalities.

The outcome of this intricate analysis was then graphically presented in two comprehensive bar charts, which eloquently depict the top contributing factors to attack severity. The charts reveal that 'Suicide Bombing' stands out as the most lethal tactic, with 'Multi-Day' incidents coming in second, indicating a disturbing trend where prolonged attacks result in higher fatalities. Other significant factors included the use of 'Ambush' tactics and 'Military Facility' as a target, both of which were associated with greater average deaths than other factors.

This visual and analytical scrutiny not only identifies the methods and targets that are most likely to amplify casualties but also provides a critical lens through which researchers can examine the dynamics of terrorist attacks. Such insights are invaluable for informing policy and shaping counter-terrorism strategies, aiming to mitigate the impacts of these tragic events. The report serves as a resource for researchers by furnishing them with a clearer picture of how particular strategies and target selections by terrorists correlate with the devastating impact of their attacks.

* Do weapon and target types influence the tactic type of a terrorist organization?

A graph of a project

Description automatically generated with medium confidence

Figure 36 Correlation between victims and tactics.

A graph with multiple colored bars

Description automatically generated with medium confidence

Figure 37 Correlation between weapons and tactics.

This analysis was centered around the pressing question: "Do weapon and target types influence the tactic type of a terrorist organization?" the study employed point-biserial correlation to delve into the intricate relationship between the choice of weapons, the selection of targets, and the employment of tactics within terrorist operations. This approach enabled the identification of patterns that might suggest preferred or avoided combinations of these elements in terrorist strategies.

The findings from the analysis were visually represented through bar charts, which clearly illustrated the correlations between the variables. Positive correlations indicated a tendency to pair specific weapons with certain tactics, suggesting these combinations might be strategically preferred due to their effectiveness or other operational advantages. Conversely, negative correlations highlighted less favored pairings, possibly due to operational constraints or strategic considerations that lead organizations to avoid them.

This insightful analysis offers a clearer understanding of the strategic considerations underlying terrorist operations. It unveils not just the preferences but also the aversions that could define the operational scope of these groups. For terrorism researchers, this study provides a valuable lens through which to view the operational dynamics of terrorist organizations, contributing significantly to the body of knowledge that underpins effective counter-terrorism strategies.

* Dashboard to show information regarding terrorism incident.

A map of the world

Description automatically generated

Figure 38 Home page of dashboard.

A map of the world

Description automatically generated

Figure 39 State map

A screenshot of a computer

Description automatically generated

Figure 40 Sate information.

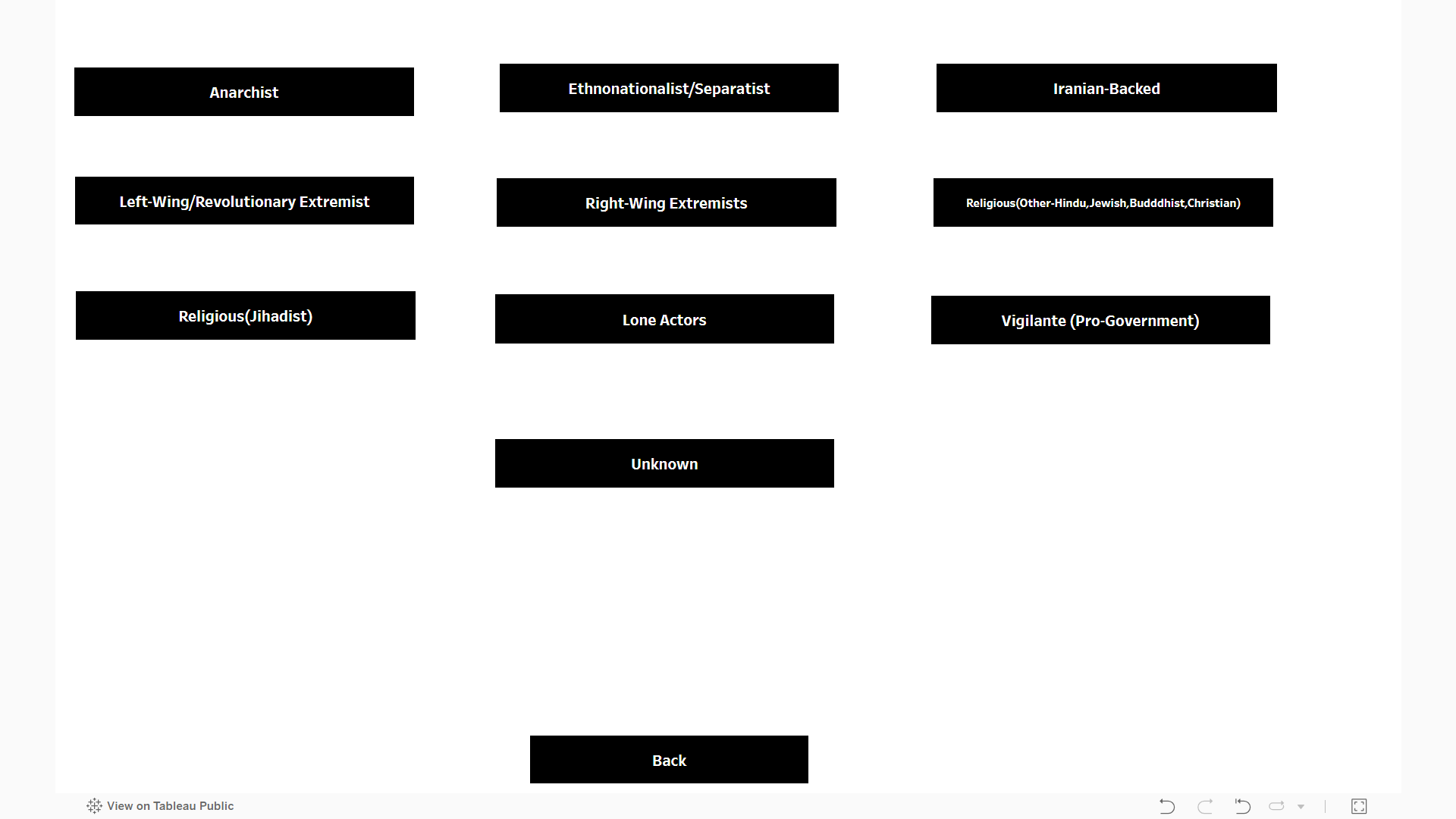


Figure 41 List of terrorist organizations.

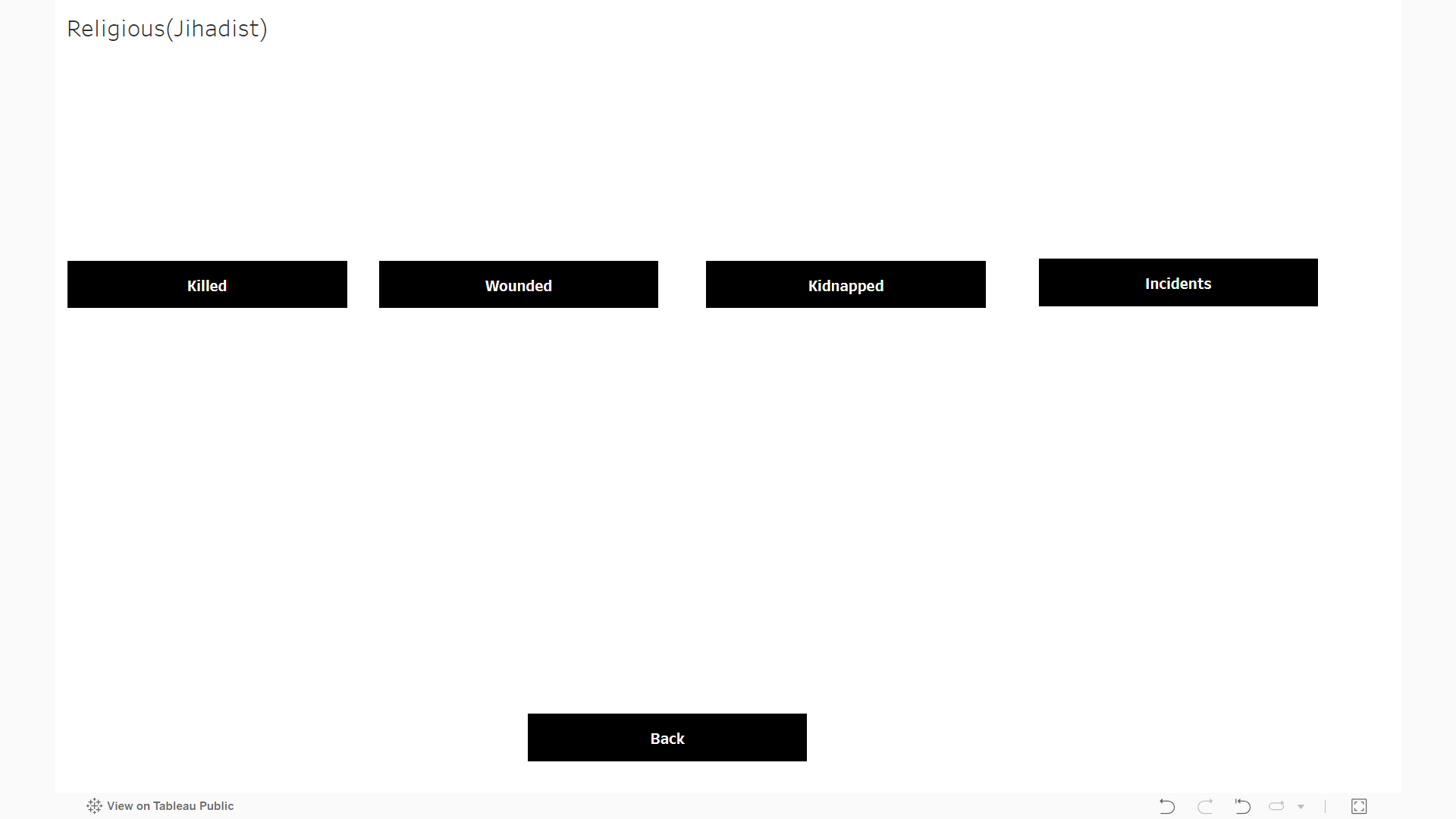


Figure 42 List for type of casualities.



Figure 43 Information with respect to the type of casuality.

We used tableau application to created a dashboard as it has effective and easy to use tool for the dashboard creation. In the home page we have created a map in which each country is colored by the number of incidents and when the mouse hovers over a country, few stats are displayed regarding the incidents occurred in that country, then by clicking on a country we get the country map with its state and providence displaying information similar to the world map and further clicking on a state or providence we move to next page with additional information for that state.

On the right side, we have added a button called perpetrator group which upon clicking will take the user to the list of terrorist organizations page when clicking on particular terrorist organization will lead to a page that contains 4 different buttons (Killed, Wounded, Kidnapped, and Incidents) upon click a button will lead to sheets that contains a country map affected by that particular organizations and 3 bar graphs (Victim, Weapon, and Tactic) that particular organization has used or affected.

# Findings

In a comprehensive analysis of global terrorism trends from 2018 to 2023, sophisticated visualizations revealed key insights: Afghanistan and Syria experienced the highest number of terrorist incidents, highlighting their geopolitical instability. Religious and ethnonationalist/separatist groups were found to have significantly higher operational capabilities, often resulting in severe casualties and extensive damage. The study observed a dynamic shift in terrorist tactics, targets, weapons, facilities, logistics, and victim types, indicating an evolution toward more sophisticated and varied attack strategies. Correlation analyses demonstrated specific relationships between perpetrator types and attack tactics, with religious (jihadist) groups favoring exploitive tactics. Factors such as the use of suicide bombings and military facilities as targets were associated with higher fatality rates, underscoring the need for focused security measures. Additionally, the analysis provided insights into the strategic preferences and aversions of various terrorist groups, aiding in the formulation of nuanced counterterrorism policies. Predictive analysis based on historical data offered forecasts of potential future threats, enhancing proactive counterterrorism planning and resource allocation. These findings not only enhance understanding of terrorism’s evolving patterns but also support the development of more effective counterterrorism strategies, emphasizing the crucial role of advanced data analytics in global security.

# Summary

The report provides an extensive analysis of global terrorism trends from 2018 to 2023, utilizing advanced visualizations to explore and interpret complex data patterns. This analysis covers a broad spectrum of terrorism-related activities, including changes in tactics, targets, and operational strategies of terrorist organizations worldwide. The methodologies employed range from detailed statistical analyses to predictive modeling, aiming to enhance the predictive accuracy and understanding of terrorism dynamics. The findings from these methodologies inform strategic counterterrorism decisions and policies, underscoring the importance of sophisticated data analytics in addressing global security challenges. Overall, the report exemplifies the integration of technology and analytics in modern counterterrorism efforts, offering insights that are crucial for developing effective security strategies and interventions.

# Future Work

* Custom User Interface Design: Enable dashboard customization, allowing users to tailor the data and visualizations to their specific research needs or operational focus.​.
* Real-Time Data Processing and Visualization: Implement systems capable of handling dashboard and visualizing data in real-time, allowing for instantaneous reflection of changes and updates in the data sources.
* Predictive Analytics and Forecasting: Develop and integrate more sophisticated algorithms for data analysis, including machine learning models that can predict future trends and identify patterns not immediately apparent to human analysts.​
* Intelligent Automation: Incorporate AI to automate the discovery of insights, generation of visualizations.
* Continuous User-Driven Development: Establish a continuous development cycle that is driven by user feedback, ensuring the dashboard evolves based on the required demands to meet changing needs.
* Multilingual Support: Ensure the dashboard is fully functional in multiple languages, with the ability to automatically translate data and insights to reach a global audience.
* Advanced Custom Reporting: Develop a highly flexible reporting system that allows users to create and share custom reports easily, with sophisticated data drill-down capabilities.

Appendix

Appendix A: Glossary

|  |  |
| --- | --- |
| Term | Definition |
| GTTAC | Global Terrorism and Analysis Center |
| ISIS | Islamic State of Iraq and The Levant |
| TraCCC | Terrorism, Transnational Crime and Corruption Center |
|  |  |
|  |  |
|  |  |
|  |  |

Table 1: Glossary Table

Appendix B: GitHub Repository

Overview

This GitHub repository hosts our DAEN 690 Master's final project, dedicated to developing a comprehensive terrorism dashboard using GTTAC datasets. Our objective is to address specific research questions, leveraging data visualization and analytical techniques to uncover patterns, trends, and insights into global terrorism activities. This interactive platform is designed to aid researchers, policymakers, and security experts in understanding the dynamics of terrorism, facilitating informed decision-making and strategic planning. Through this project, we aim to contribute significantly to the field of terrorism studies, enhancing the capabilities for monitoring, analysis, and response to terrorism worldwide.

GitHub Repository Link

The URL for the repository

<https://github.com/anamratha16/Optimus.git>

GitHub Repository Contents

A: Team Optimus Final Showcase Presentation

B: Team Optimus Report

C: Codes to all Research Questions

D: Tableau Code files S

Appendix C: Risks

Sprint 1 Risks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability | Impact | Mitigation |
| Inadequate Domain Knowledge | Understanding the correct definition of terrorism without misconception. The team had to look more into domain area to get a better understanding | Medium | Medium | Carefully went through articles and learned from them. |
| Tunnel vision | We focused too much on option one that we failed to see its short comings. | Medium | High | Explore other option rather than the first option |
| Cost Estimation | We failed to realize that some products used in our solution would cost more than expected to use them. | Low | Medium | Explore other products which are free to use or costs very less. |
|  |  |  |  |  |
|  |  |  |  |  |

Table 2: Sprint 1 Risks

Upon evaluating the project's characteristics and associated data, we identified several risks that needed careful consideration. One significant concern was the potential for “Inadequate Domain Knowledge”, leading to a lack of clarity in defining terrorism without misconceptions. To address this, the team proactively engaged in an extensive exploration of the domain, dedicating time to carefully review articles authored by experts. This approach proved essential in gaining a complete understanding of the project's context, allowing us to navigate the complexities of terrorism analysis.

The building of a project solution is a broad vision process. Since, it comprises multiple options to choose from, before considering a fixated option. The risk of “Tunnel vision” then emerged, as the team initially was fixated on one option, neglecting to adequate its shortcomings. This concentration on a single solution posed a medium probability of occurrence posing a high impact on the project. To address this risk, the team consciously broadened the perspective in building a solution model considering options beyond the initial choice. By promoting a more comprehensive evaluation of potential solutions, the team aimed to mitigate the adverse impact of tunnel vision and ensure a more informed decision-making process.

The total cost estimation of the products used in building the solution model was initially not considered as a primary impact, later then we saw the risk of “Cost estimation” emerging, as the chosen technology tools to build the solution are very expensive more than the affordable budget to the project. This ignorance of cost posed a medium probability of occurrence posing a high impact on the project. The team addressed this risk by considering more affordable solution tools and exploring alternative technology options. This proactive approach aimed to mitigate the high impact of cost estimation, ensuring a more budget-friendly selection of service providers for the project.

Sprint 2 Risks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability | Impact | Mitigation |
| Semantic Ambiguity | The potential for confusion or misunderstanding regarding the meaning of data attributes, labels, or terms within a dataset. | Medium | High | GTTAC Manual has access to a detailed references that can serve as a reliable source of information for enhancing the understanding of the data. |
| Incomplete data | Incomplete data refers to missing or unavailable information in the dataset, often represented as null values which can occur for various reasons, such as errors in data collection, data entry mistakes, or the absence of information for certain records. | Medium | High | Implementing data imputation techniques to fill in missing values with estimated or calculated values. |
| Expired Website Certificates | While automating an activity on a website, if any of its security certificates expire then it will interrupt the entire automation process. | Low | High | Design the scripts to navigate through the certification expire page or ignore it. |
|  |  |  |  |  |
|  |  |  |  |  |

Table 3: Sprint 2 Risks

Semantic ambiguity is a significant risk in projects and technical communications because it allows for misunderstandings caused by ambiguous or poorly defined terms. This risk, with a medium chance and a high impact, can cause project delays, implementation problems, and fails to reach objectives. Mitigation measures include developing a single language and a detailed reference guide to ensure that all stakeholders have a consistent understanding, lowering the risk of misinterpretations and misalignments between project needs and delivered results.

In information-driven contexts, incomplete data, defined as lacking information critical for analysis and decision-making, represents a medium likelihood but high effect risk. Incomplete data sets can result in erroneous analysis, misleading conclusions, and flawed judgments, all of which have an impact on strategic directions and operational efficiency. Implementing data imputation techniques and developing strong data gathering standards are critical mitigating measures for ensuring data integrity and facilitating informed decision-making.

Expired website certificates pose a low likelihood but high effect danger, especially for automated web-based activities. These certificates are required for safe web connections, and their expiration can interrupt automated services by causing security warnings. Mitigation entails creating automation scripts that can handle such alerts, as well as automating certificate monitoring and renewal to prevent expiration, ensuring that automated procedures run continuously and without substantial delays.

Sprint 3 Risks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability | Impact | Mitigation |
| Ineffective Data  Representation | This risk involves choosing visualizations that do not accurately or effectively communicate the underlying data, potentially misleading viewers. | High | High | Educate the team on basic data visualization principles, use simple and universally understood visualizations, and conduct user testing for feedback. |
| Complexity | The visualization becomes too complicated for the intended audience to understand, reducing its effectiveness. | Medium | High | Aim for simplicity and clarity, use legends and explanatory text where necessary, and seek to remove non-essential information that does not contribute to the research question. |
| Data Preparation and Transformation Limitations | when it comes to data manipulation tasks or the need to process grid datasets efficiently, Tableau's built-in capabilities are falling short. This includes scenarios where advanced detailed row-level data manipulation is required. | Medium | High | Use of External Programming Skills (Python) to solve research questions. |
|  |  |  |  |  |
|  |  |  |  |  |

Table 4: Sprint 3 Risks

In the process of evaluating the project parameters and the depth of our data, we have identified crucial risks that require strategic management. The risk of "Ineffective Data Representation" stood out prominently, characterized by the use of visualizations that might not accurately communicate the underlying data, potentially leading viewers astray. This risk, with its high probability and impact, was mitigated by reinforcing the team's understanding of foundational data visualization principles. By adopting straightforward and universally understandable visualizations and incorporating iterative user testing for feedback, we significantly minimized the potential for misrepresentation.

Concurrently, we faced the risk of "Complexity" within our visualizations, which bore a medium chance of occurrence yet posed a high impact on the project's clarity. To counteract this, we embraced a philosophy of simplicity, integrating clear legends and explanatory texts, and were diligent in excising extraneous information. By refining our visual approach, we ensured that every piece of information presented was directly relevant to the research question, thereby maintaining the integrity and effectiveness of our communication.

Our project's success hinges on the ability to meticulously prepare and transform data, tasks for which Tableau's inherent features occasionally prove insufficient, particularly with grid datasets demanding granular row-level manipulation. The probability of encountering such limitations is assessed at a medium level, given the complexity and variety of our data. Nonetheless, the impact is considered high, as the effectiveness of our visualizations and the accuracy of insights drawn are contingent upon the quality of these preparatory stages. To proactively mitigate this risk, we've incorporated Python into our workflow, utilizing its robust programming capabilities to fill the gaps Tableau leaves in data manipulation. Python's extensive libraries and functions enable us to perform complex data transformations and preparations that are essential for answering our research questions comprehensively. This strategic integration of Python's versatility with Tableau's visualization strengths forms a cohesive analytical toolkit, ensuring no data challenge goes unmet.

Sprint 4 Risks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability | Impact | Mitigation |
| Lack of User Adoption | Users may not find the dashboard intuitive or useful, leading to low adoption rates, wasted resources, and mismatched features with user needs. | High | High | Involve end-users in the design process to understand their requirements and preferences. Provide training and support to ensure users understand how to use the dashboard effectively. Continuously gather feedback and make improvements based on user input. |
| Expansion of Scope | The project scope may expand beyond initial requirements, leading to delays, budget overruns, often without proper evaluation of their impact on timeline, budget, and resources. | Medium | High | Clearly define project objectives, requirements, and deliverables upfront. Establish change control processes to manage any changes to the scope, timeline, or budget effectively. Prioritize features based on their impact and feasibility. |
| Performance Issues | Performance issues can arise from factors such as inefficient data connections, excessive data volume, or overly complex visualizations. | Medium | High | Optimize data connections, use data extracts where necessary, and simplify complex calculations or visualizations. |
|  |  |  |  |  |
|  |  |  |  |  |

Table 5: Sprint 4 Risks

In developing the visualizations, we identified several potential risks that could impact the project's success. A significant concern is the Lack of User Adoption. There is a risk that users may not engage with the dashboard if it’s not intuitive or aligns with their needs, which could lead to low adoption rates and wasted resources. To address this, we involved our client in the design process, provided comprehensive training, and set up a system to continuously integrate feedback for improvements. Another challenge is the Expansion of Scope, where the project could potentially extend beyond the initial requirements, causing delays. We tackled this by clearly setting project objectives and deliverables from the start and maintaining strict control over any changes in scope. Furthermore, Performance Issues such as inefficient data connections or complex visualizations could undermine the dashboard’s effectiveness. To mitigate these issues, we optimized data connections, utilized data extracts, and simplified visualizations to ensure the dashboard operates smoothly and efficiently.

Sprint 5 Risks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability | Impact | Mitigation |
| Misinterpreting Graphs | Drawing conclusion from the graphs without proper background check or knowledge cause misinformation. | Medium | High | Communicating with a domain expert and verify the finding of the graph will prevent in drawing wrong conclusions. |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Table 6: Sprint 5 Risks

During the final analysis, we encountered the risk of misinterpreting graphs, potentially leading to misinformation. To overcome this challenge, we established a mitigation strategy. We ensure the accuracy of our conclusions by communicating with domain experts to verify the findings of the graph. This approach significantly reduces the risk of drawing incorrect conclusions and spreading misinformation.

Appendix D: Agile Development

Scrum Methodology

Figure 44: Sprint project dates.

Our team initially, faced some hurdles in adapting to the YouTrack. We quickly unified by setting a dedicated time for daily scrums. These 15–30-minute meetings became our communication hub, fostering transparency and accountability. We assigned tasks, shared updates, and actively sought feedback, enabling us to achieve sprint goals efficiently.

Sprint 1 Analysis

It is focused on problem definition and plays a crucial role in establishing the foundational framework for the entire project. During this phase, we collaborate closely with the client to conduct an in-depth requirements analysis, aiming to further refine the problem statement and project goals initially outlined in the project catalog. Additionally, the team engages in comprehensive research within the domain, seeking to deepen their understanding of the specific knowledge relevant to the project. The goal of Sprint 1 is to not only solidify the project's objectives but also to identify a viable solution approach. This involves the creation of a detailed solution design and systems architecture, laying the groundwork for subsequent sprints that will delve into dataset analysis, algorithm development, and the eventual generation of visualizations or implementation, depending on the project's nature. Sprint 1 is pivotal in shaping the strategic direction and conceptual framework that will guide the project's progression.

Sprint 2 Analyis

In Sprint 2, our team's primary goals were to thoroughly review the provided datasets to identify the attributes needed. Additionally, we carefully checked the datasets for null values and discussed potential strategies to reduce their impact on the data's accuracy. We focused on streamlining the process of downloading data by writing Python scripts for both Windows and Mac OS. Furthermore, we explored different visualization strategies intended to elucidate each research question to enhance our findings' significance and comprehensibility.

We strengthened our data processing workflow's resilience and effectiveness by proactively identifying and managing null values and automating data download processes. Furthermore, the clarity and significance of our findings were enhanced by our focus on visualization techniques, which made sure that the insights from our analyses were effectively communicated. Because of the flexibility of the Scrum framework, we were able to adjust our goals as we gained new knowledge, which made leading each Sprint both rewarding and challenging. By demonstrating our dedication to careful data management and presentation, these methodological developments provide a strong platform for our ongoing pursuit of excellence.

Sprint 3 Analysis

In Sprint 3, our team's primary aim is to go with a systematic approach for this project analysis with the goal of increasing the effectiveness of our data visualization efforts. At first, we focused on utilizing Tableau to improve dashboards and visualizations. During this phase, we thoroughly investigated Tableau's features, delving into both its advantages and disadvantages. We gained invaluable knowledge about the platform's features and limitations during this process, which enabled us to modify our research questions. Through enhancing our Tableau skills, we were able to improve not only the visuals that we produced, but also the analytical queries that we conducted. This allowed us to develop a deeper comprehension of the data dynamics that underlie our visualizations.

After exploring Tableau, we expanded our project into Python to continue developing visualizations. With this switch, we were able to take advantage of Python's many libraries and scripting powers to produce dynamic and personalized visuals. We also did a lot of research at the same time to investigate different ways to improve the research questions and dashboards. By means of thorough investigation and testing, we discovered a multitude of approaches and strategies targeted at improving our dashboard's user interface and expanding the scope of our research questions. This multidisciplinary approach put our project at the forefront of innovation in data-driven research while also expanding our technical toolkit and enhancing the scholarly rigor of our academic pursuits.

Sprint 4 Analysis

In Sprint 4 of this project, the primary focus was on implementing and developing a comprehensive dashboard to provide actionable insights into terrorism trends and patterns. The team made significant progress by incorporating feedback from previous sprints, enhancing data visualizations of each research question analysis, and automating data integration processes for more dynamic and robust dashboard functionality.

Additionally, the team had regular meetings and discussions held to align the dashboard's development with user needs and to integrate essential feedback for further improvements. Technical enhancements were also made, including optimizing data connections and simplifying complex visualizations to enhance performance and efficiency, especially important given the extensive datasets involved.

By the end of Sprint 4, Team Optimus had significantly refined the terrorism dashboard, enhancing its functionality and user interface. The improvements in automation and data processing supported a more effective and efficient tool for analyzing terrorism trends. Looking ahead, the team plans to continue refining the dashboard based on ongoing client’s feedback and to explore further integration of additional data sources and analytical features, ensuring the dashboard remains a cutting-edge tool for stakeholders requiring timely and accurate information to make informed decisions.

Sprint 5 Analysis

Sprint 5 was all about polishing the final report format and crafting impactful slides for our showcase presentation. We closely followed the professor's guidance from Blackboard and feedbacks to ensure everything aligned perfectly with our research questions and project objectives. To make the presentation extra clear for the audience, we used a strategic approach of highlighting key elements in the slides.

We started by building a compelling title slide, then dove into sections that introduced our partner organizations and clearly defined the problem we were tackling. Following that, we showcased the Analytics we performed, and the Visualizations we created to present the findings. We also carefully modified a couple graphs in response to comments we received. These adjustments were made with the intention of enhancing the data and analysis's readability, precision, and aesthetic appeal in order to present the results in a way that is more intelligent and professional.

Appendix E: Automatic dataset Download

**Automation Script setup:**

**Step 1:** Create a folder to store GTTAC’s downloaded zip file (GRID\_as\_csv\_tables.zip).

**Step 2:** Create a folder to store the extracted zip file.

**Step 3:** Create a folder to store the automation script file (GTTAC\_Auto\_Download.py).

**Step 4:** Right click on GTTAC\_Auto\_Download.py and select “Edit with notepad.”

**Step 5:** Make changes to the script according to Note1, Note2 and Note3 of the script and save the file.

**Setting up task scheduler:**

1. **Open Task Scheduler:** Click on start, type “Task Scheduler” in search bar and select Task Scheduler App.
2. **Creating Basic Task:** In Task Scheduler application right click on Task Scheduler Library and select “Create Basic Task”.
3. **About Task:** Write the name and description of task and click on Next.
4. **Setting up repeat schedule:** To run the script every week, select “Weekly” option and click on Next.
5. **Setting up time:** To run the script at a particular time and at a particular day of the week, enter the start data and time as show in the below image. To run the script only once a week, enter 1 at Recur every. Then select Monday option and click on Next.
6. **Setting up task to perform:** Select “Start a program” option (default) and click on Next.
7. **Selecting task:** Enter the path of the script or use browse button to select the script and click on Next.
8. **Complete task creation**: Check the details of the task and click on Finish.

Reference

|  |  |
| --- | --- |
| [1] | B. Herre, V. Samborska, H. Ritchie, J. Hasell, E. Roser and M. Mathieu, "Terrorism," Our World in Data, [Online]. Available: https://ourworldindata.org/terrorism. |
| [2] | J. P. Jenkins, "Terrorism," Britannica, 25 January 2024. [Online]. Available: https://www.britannica.com/topic/terrorism. |
| [3] | I. Boutchalern. [Online]. Available: https://www.vecteezy.com/photo/22330405-terrorist-incident-background-anti-terrorism-day-concept-generative-ai. |
| [4] | S. R., "One man's terrorist is another man's freedom fighter.," Linkedin, 22 October 2023. [Online]. Available: https://www.linkedin.com/pulse/one-mans-terrorist-another-freedom-fighter-sanjay-rai. |
| [5] | Development Services Group, Inc., "Annex of Statistical Information 2022," Bethesda, 2023. |
| [6] | "9/11 FAQs," 911 Memorial & Museum, [Online]. Available: https://www.911memorial.org/911-faqs. |
| [7] | WikiImages, pixabay, [Online]. Available: https://pixabay.com/photos/world-trade-center-twin-towers-67695/. |
| [8] | J. Ellis, "Deadliest Terrorist Groups in 2023," Vision of Humanity, [Online]. Available: https://www.visionofhumanity.org/deadliest-terrorist-groups-in-2023/. |
| [9] | P. Mahmut Cengiz, "Global terrorism tends and analytics center," 28 September 2023. [Online]. Available: https://gttac.com/about\_gttac/. [Accessed 11 february 2024]. |
| [10] | "Welcome to TraCCC," George Mason University, [Online]. Available: https://traccc.gmu.edu. |
| [11] | "METHODOLOGY," GTTAC, [Online]. Available: https://gttac.com/methodology/. |
| [12] | "ABOUT GTTAC," GTTAC, [Online]. Available: https://gttac.com/about\_gttac/. |
| [13] | "Information on more than 200,000 terrorist attacks," Global Terrorism Database, [Online]. Available: https://www.start.umd.edu/gtd/ . [Accessed 28 January 2024]. |
| [14] | "You searched for terrorism," Vision of Humanity, [Online]. Available: https://www.visionofhumanity.org/?s=Terrorism. [Accessed 28 January 2024]. |
| [15] | "Home Page," Counter Extremism Project, [Online]. Available: https://www.counterextremism.com/. [Accessed 28 January 2024]. |
| [16] | G. Martin, "Understanding Terrorism: Challenges, Perspectives, and Issues," *6th ed. Sage Publications,* 2021. |
| [17] | "Global Terrorism Trends and Analysis Center," George Mason University, [Online]. Available: https://traccc.gmu.edu/projects/current/global-terrorism-trends-and-analysis-center/. |
| [18] | B. Hoffman, "Inside Terrorism," *3rd ed. Columbia University Press,* 2017. |
| [19] | Institute for Economics & Peace, "Global Terrorism Index 2023: Measuring the Impact of Terrorism," Institute for Economics & Peace, Sydney, Australia, 2023. |
| [20] | Vision of Humanity, "Countries most impacted by terrorism," Institute for Economics & Peace, Sydney, Australia, 2023. |
| [21] | M. Cengiz, "Comparative Analysis of ISIS and Al-Qaeda: Strategic Objectives, Operational Tactics, and Threat Assessment," *Journal of Terrorism Studies,* vol. V, no. 2, pp. 34-48, 2022. |
| [22] | S. Patel, "Geopolitics of Terrorism: A Regional Analysis," *Geopolitical Review,* vol. XIX, no. 3, pp. 275-290, 2023. |
| [23] | T. Johnson, "The Multifaceted Approach to Counterterrorism," *Security and Strategy Journal,* vol. XXII, no. 2, pp. 111-127, 2023. |
| [24] | Development Services Group, Inc., "Development Services Group, Inc.," State Gov, Maryland, 2023. |
| [25] | Salesforce, tableau from, "Tableau Desktop," TABLEAU SOFTWARE, LLC, A SALESFORCE COMPANY., [Online]. Available: https://www.tableau.com/products/desktop. [Accessed 11 February 2024]. |

**This page intentionally left blank**