

GLOBAL TERRORISM DASHBOARD

A PROJECT REPORT

Submitted in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION

Submitted by

Kuhu Chawla

(Roll Number: UE195069)

Muskan Gupta

(Roll Number: UE195078)

Pushkar Bhardwaj

(Roll Number: UE195087)

UNDER THE GUIDANCE OF

Ms Daljeet Kaur

(Assistant Professor, ECE, UIET)



UNIVERSITY INSTITUTE OF ENGINEERING AND TECHNOLOGY

PANJAB UNIVERSITY: CHANDIGARH (160025)

March-June 2022

DECLARATION

We hereby certify that the work which is being presented in this project entitled “GLOBAL TERRORISM DASHBOARD”, in partial fulfilment of the requirement for the award of degree of “Bachelor of Engineering in Electronics and Communication” submitted in University Institute of Engineering and Technology, Panjab University, Chandigarh, is an authentic record of our own work carried out under the supervision of Ms Daljeet Kaur, Assistant Professor, ECE UIET.

Place: Chandigarh
Date: 5th July, 2022

Kuhu Chawla (UE195069)

Muskan Gupta (UE195078)

Pushkar Bhardwaj (UE195087)

ABSTRACT

Different regions in the world have been plagued by different types of terror attacks at different times of the year. These terror attacks could keep on happening and list could keep going on if we are not prepared to tackle them. With no uniformity, it can definitely be freaking hard for the World-Leaders to take preventive actions. The Global Terrorism Database (GTD) is an opensource database with information of around 180K terrorist events across the world in the years of 1970-2017. Using the information of the events, for the minor project, we designed and developed an interactive dashboard which explores the dataset in depth. Giving meaning and usable value to the enormous dataset of over half a century by understanding the trends and contouring the patterns therein.

In addition to summarising the terrorism trends over the years, the dashboard provides means to view important statistics of an event and that of a region. Such understanding of the data can be used to form global and local anti-terrorism policies to prevent such events.

With their attacks turning ineffective, terror groups would have no way to keep their finances thriving and with economical failure and tactical failure on ground, the global terrorism can be put down by a large figure.

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of the Project would be incomplete without the mention of the people who made it possible. Lack of their constant guidance and encouragement would have made our efforts go in vain. We consider ourselves privileged to express gratitude and respect towards all those who guided us through the completion of this project.

We convey thanks to Dr Naresh Kumar, coordinator ECE Department and all the faculty members for their constant support. We thank our group incharge Ms. Preeti Gupta for her support. We extend our gratitude to Ms Daljeet Kaur for providing encouragement, constant support and guidance at every step which was of great help in completing this project successfully.

We are grateful to Panjab University for providing us with this great opportunity and facilitating us with all possible resources and guidance.

Kuhu Chawla (UE195069)

Muskan Gupta (UE195078)

Pushkar Bhardwaj (UE195087)

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	Abstract	ii
	Acknowledgement	iii
	List of Figures	iv
1.	INTRODUCTION	7-12
	1.1 Data Analysis	8
	1.2 Data Visualization	9
	1.3 Tableau	10
2.	PROJECT DESCRIPTION	13-15
	2.1 Data Description	13
	2.2 Selected Data Set	13
	2.3 Identify central entities and relationships	14
	2.4 Identify basic Stats that will be useful for processing	14
	2.5 Relation amongst the following parameters	15
3.	METHODOLOGY	16-21
	3.1 Data Retrieval	16
	3.2 Dataset Cleaning up	18
	3.3 Cleaned up Dataset Uploaded to Tableau	19
	3.4 Creation of Worksheets in Tableau	19
	3.5 Final Analysis	19
4.	RESULTS AND DISCUSSION	22-26
	4.1 Trends in Terrorist Events	22
	4.2 Region Wise Statistics	23
	4.3 Information about the Active Terror Groups and Weapons	24

5.	CONCLUSION AND FUTURE SCOPE	27-28
	5.1 Conclusion	27
	5.2 Future Scope	28
	TIMELINE OF PROJECT	29-30
	REFERENCES	31

LIST OF FIGURES

Figure no.	Figure description	Page number
Fig 1.1	Data Visualization	10
Fig 1.2	Process of Analysis	12
Fig 3.1	Block diagram representation of the methodology adopted	16
Fig 3.2.1	Original uncleaned dataset	17
Fig 3.2.2	Original uncleaned dataset	17
Fig 3.3	Screenshot of data cleaning using Python	18
Fig 3.4	Screenshot of cleaned dataset	18
Fig 3.5	Exploratory Analysis	19
Fig 3.6	Exploratory Analysis	19
Fig 3.7	Exploratory Analysis	19
Fig 3.8	Screenshot of Dashboard 1	20
Fig 3.9	Screenshot of Dashboard 2	21
Fig 3.10	Screenshot of Dashboard 3	21
Fig 4.1	Trends in Terrorist Events	23
Fig 4.2	Region wise Statistics	14
Fig 4.3	Active Terror Groups and Weapons	25
Fig 4.4	Count of Attack Type	25
Fig 4.5	Weapons used for attacks in South Asia	26
Fig 4.6	Areas Affected in South Asia	26
Fig 4.7	Attacks over years in South Asia	26

CHAPTER-1

INTRODUCTION TO THE PROBLEM STATEMENT

The word terrorism is associated with acts committed by individuals which are unlawful to demonstrate their power to the world. People who commit these unlawful acts are called terrorists. Terrorism is associated with a negative impact on both society and the people. This leads to loss of life, destruction of families, destruction of infrastructure and destruction of peace in this world.

Over the years terrorism has become more prevalent than ever in this world. The number of terrorist incidents has rapidly increased in the past few years. The ideology behind these attacks is a matter of concern. The act of terrorism which was initially said to be a law and order problem for a few countries has now become a global problem. Countries like the USA, UK, Australia, the European Union etc. have also become a target of these activities.

This project aims to statistically provide a solution to this global problem through the means of constructive, analytical and visual analysis of the past terror activities which has occurred over the years. The primary goal of this project is to analyse the previous trends of incidents to analyse the target areas, weapons used, target personnel and the terror group associated with the attack.

The major output of the project is to give countries a visualised analysis of the terror groups they should be the most concerned about and about what weapon trade they should keep a comprehensive track of. This analysis will help the world leaders get a detailed analysis to be able to create a strategic and defensive plan against these terror activitites and organisations to prevent incidents of destruction to happen in the future.

For the reader to understand this detailed comprehensive report they would need to know the basics of a few concepts discussed below.

1.1 DATA ANALYSIS

Data analysis is the science of examining data to conclude the information to make decisions or expand knowledge on various subjects. It consists of subjecting data to operations. This process happens to obtain precise conclusions to help us achieve our goals, such as operations that cannot be previously defined since data collection may reveal specific difficulties.

The following simple steps help you identify and sort out your data for analysis.

1. Data Requirement Specification - define your scope:

- Define short and straightforward questions, the answers to which you finally need to make a decision.
- Define measurement parameters
- Define which parameter you take into account and which one you are willing to negotiate.
- Define your unit of measurement. Ex – Time, Currency, Salary, and more.

2. Data Collection

- Gather your data based on your measurement parameters.
- Collect data from databases, websites, and many other sources. This data may not be structured or uniform, which takes us to the next step.

3. Data Processing

- Organize your data and make sure to add side notes, if any.
- Cross-check data with reliable sources.
- Convert the data as per the scale of measurement you have defined earlier.
- Exclude irrelevant data.

4. Data Analysis

- Once you have collected your data, perform sorting, plotting, and identifying correlations.

- As you manipulate and organize your data, you may need to traverse your steps again from the beginning. You may need to modify your question, redefine parameters, and reorganize your data.
- Make use of the different tools available for data analysis.

5. Infer and Interpret Results

- Review if the result answers your initial questions
- Review if you have considered all parameters for making the decision
- Review if there is any hindering factor for implementing the decision.
- Choose data visualization techniques to communicate the message better. These visualization techniques may be charts, graphs, color coding, and more.

1.2 DATA VISUALIZATION

Data visualization is the representation of data through use of common graphics, such as charts, plots, infographics, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

Data visualization can be utilized for a variety of purposes, and it's important to note that it is not only reserved for use by data teams. Management also leverages it to convey organizational structure and hierarchy while data analysts and data scientists use it to discover and explain patterns and trends.

Data visualization is commonly used to spur idea generation across teams. They are frequently leveraged during brainstorming or Design Thinking sessions at the start of a project by supporting the collection of different perspectives and highlighting the common concerns of the collective. While these visualizations are usually unpolished and unrefined, they help set the foundation within the project to ensure that the team is aligned on the problem that they're looking to address for key stakeholders.

Data visualization for idea illustration assists in conveying an idea, such as a tactic or process.

Visual discovery and every day data viz are more closely aligned with data teams. While visual discovery helps data analysts, data scientists, and other data professionals identify patterns and trends within a dataset, every day data viz supports the subsequent storytelling after a new insight has been found. Data visualization is a critical step in the data science process, helping teams and individuals convey data more effectively to colleagues and decision makers.

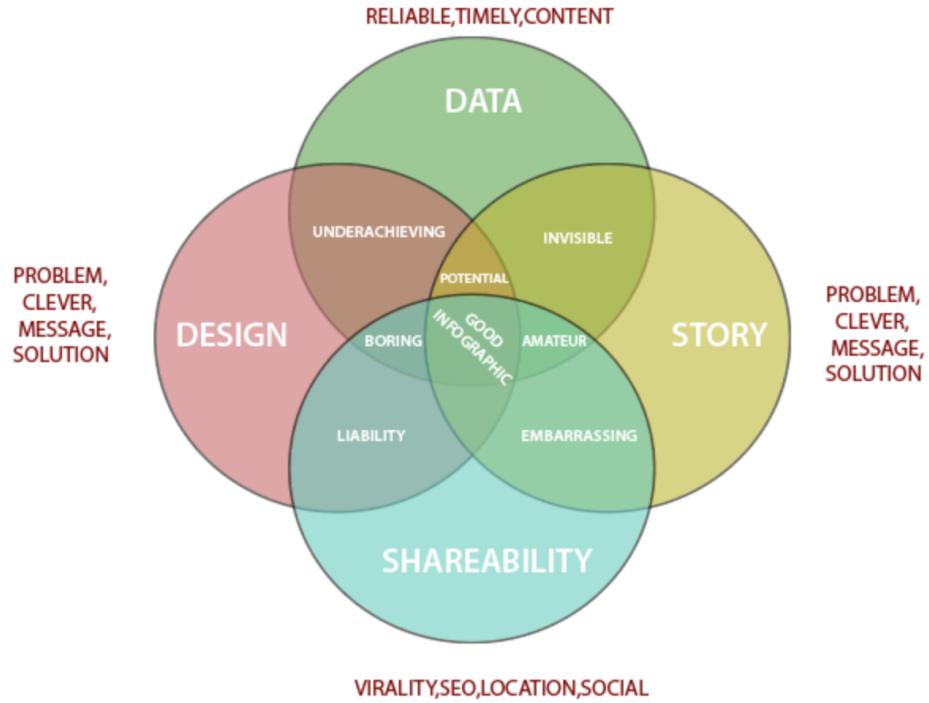


Fig 1.1 Data Visualization

1.3 TABLEAU

Tableau is the rapidly growing visualization tool used for different business applications. It is high in demand software due to its easy access to use. It is widely used in Business Intelligence (BI) processes. Tableau is crafted in such a way that all kinds of charts, plots, and graphs with different designs can be positioned simultaneously for visualization.

If we talk about the data type and structure and their access in tableau, it can work on any kind of data type either data is structured, unstructured datasets with the accessing of any sort of programming language as R, Python, SAS, etc.

A person with a non-technical background can easily work on Tableau as it doesn't require any technical or programming knowledge. Researchers described Tableau as a "highly interactive and intuitive visual-based exploration experience for business users to easily access, prepare and analyze their data without the need for coding"

Tableau has plenty of easily accessible functions that can create highly simplified graphs or charts for any set of complex data. A business analyst can investigate any pattern, insight, flow, or trends from visually available data and hence predict or conclude for any business problem.

Tableau provides solutions for all kinds of industries, departments, and data environments. Following are some unique features which enable Tableau to handle diverse scenarios.

- Self-Reliant – Tableau does not need a complex software setup. The desktop version which is used by most users is easily installed and contains all the features needed to start and complete data analysis.
- Visual Discovery – The user explores and analyzes the data by using visual tools like colors, trend lines, charts, and graphs. There is very little script to be written as nearly everything is done by drag and drop.
- Blend Diverse Data Sets – Tableau allows you to blend different relational, semistructured and raw data sources in real time, without expensive up-front integration costs. The users don't need to know the details of how data is stored.
- Architecture Agnostic – Tableau works in all kinds of devices where data flows. Hence, the user need not worry about specific hardware or software requirements to use Tableau.
- Real-Time Collaboration – Tableau can filter, sort, and discuss data on the fly and embed a live dashboard in portals like SharePoint site or Salesforce. You can save your view of data and allow colleagues to subscribe to your interactive dashboards so they see the very latest data just by refreshing their web browser.
- Centralized Data – Tableau server provides a centralized location to manage all of the organization's published data sources. You can delete, change permissions, add tags, and manage schedules in one convenient location. It's easy to schedule extract refreshes and

- Manage them in the data server. Administrators can centrally define a schedule for extracts on the server for both incremental and full refreshes.

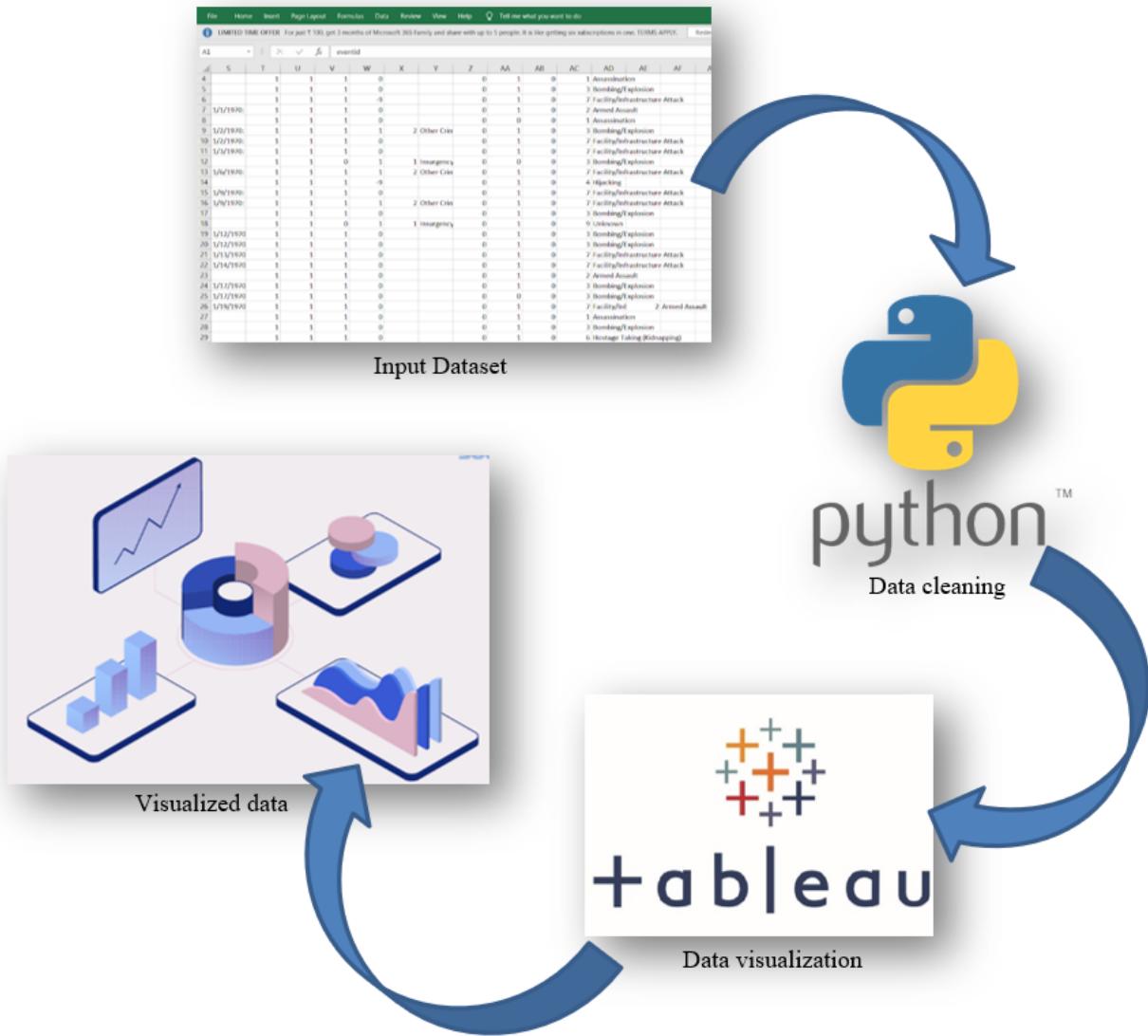


Fig 1.2: Process of Data Analysis

CHAPTER 2

PROJECT DESCRIPTION

2.1 Data Description:

The Global Terrorism Dataset (GTD) [1] is an open-source database maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of Maryland (UMD). This dataset recorded information on 181,693 terrorist attacks all over the world from 1970 to 2017. The START organization defines terrorism as follows: "The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation." Thus, following the definition, the GTD database covers events such as assassination, kidnapping, armed assault in rallies, attacks on religious group etc, apart from all the global major terror events.

2.2 Selected Data set:

1. Size: 181,693 rows, 160 MB
Original Format: CSV
2. Is the data at rest? Is it Streaming?
Data is at rest.
3. Is the data Time Variant data?
Yes.
4. Is Time an essential component of the desired data processing?
Yes, time is an essential component of data processing.

Identifying trends in terrorism is central to facilitating data exploration of GTD. In the final model, a page is dedicated for this purpose which can be explored in the 'Trends' section. In this section, the trends in the number of attacks and deaths worldwide and for each individual country over the years can be observed. This helps in understanding of the political, social or economic standing of a country at a given period of time and how the policies of countries is helping in either curbing or fostering terrorist attacks.

As an example, from the visualisation in Figure 1, it can be seen that there was a sudden peak in terrorism in the year 2014, due to the large number of terror events occurring in Iraq (Figure 2). The visualization over time has a potential to be useful for experts to chain events together through time and form policies based on historical events.

2.3 Identify central entities and relationships

The central entities in GTD are:

1. Incident
2. Location
3. Time of Occurrence
4. Target/Victim
5. Perpetrator
6. Casualties and Consequences Relationships:
 - Incident occurred at Location.
 - Incidents occurred at time (year, month, day) and lasted for how long.
 - Perpetrator carried out the incident.
 - Target/Victim casualties or injuries by perpetrator.
 - Incident caused casualties.
 - Nationality (Location) of the Target/Victim

2.4 Identify basic Stats that will be useful for processing

1. No. of Attributes = 135
2. No. of Records = 181,693
3. No. of Terrorist groups = 3,537
4. No. of related events: 25,032

The dataset logs list of related events. For each event, an attribute contains a list of event ids of events which are related to the particular event. One example of related events is the terror events which took done the World Trade Center in New York City, United States on 9/11. An event logs the details of the first plane crashing into the towers and a separate event entry logs

the second plane crashing into the second tower, minutes after the first one. Both events are marked as related in the database.

2.5 Relation amongst the following parameters:

- Location, Attack type, Target type, Weapon type (Political/Religious): Identify the political, social scenario of a country.
- Country, success of attack: How successful a country is in curbing a threat.
- Victim type (Tourists/Govt/ Police), casualties
- Terrorist group, attack type, country attacked: Identify characteristics of a terrorist group.
- Terrorist group, motive: Identify the ideology or motive of a terrorist group.
- Event, casualty: Related events and their damage and extent

CHAPTER 3

METHODOLOGY

The project has been divided into 5 key steps:

1. Dataset Retrieval
2. Cleaning up of dataset and general Exploratory Data Analysis Using Python
3. Cleaned up dataset uploaded to Tableau
4. WorkSheets and Dashboards built
5. Final Analysis

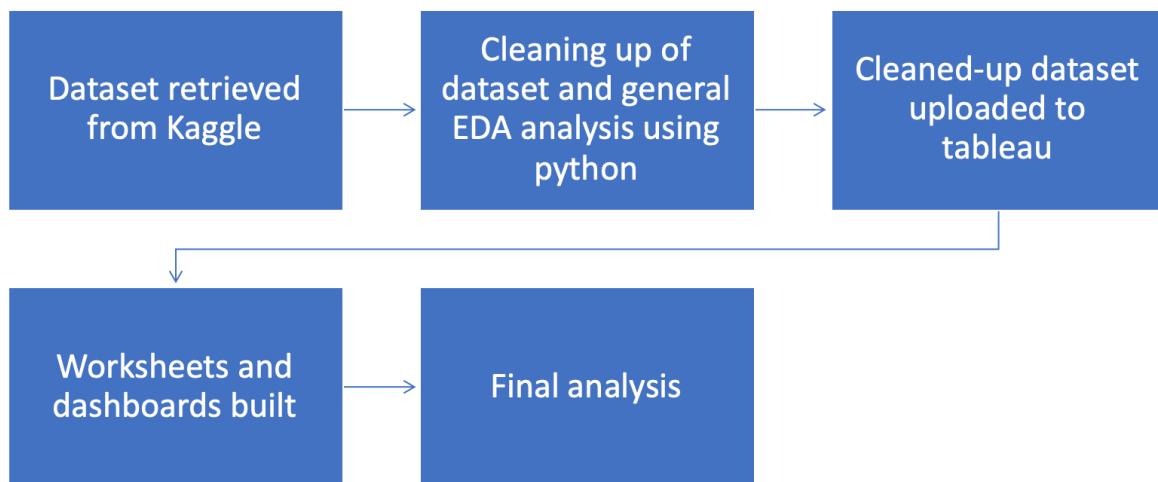


Fig 3.1: Block diagram representation of the methodology adopted

3.1 Dataset Retrieval

The Global Terrorism Dataset (GTD) [1] is an open-source database maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of

Maryland (UMD). This dataset recorded information on 181,693 terrorist attacks all over the world from 1970 to 2017.

At the time building of the project the GTD dataset was restricted for download from the respective website due to the maintenance of the dataset as well as addition of new figures from the year 2018 and 2019. Due to this reason we retrieved the older version of the dataset from Kaggle for our analysis. In this project, the Dataset used is from the year 1970 to 2017.

The retrieved dataset is in the form of csv a screenshot of which is attached below in Fig 3.2 below.

globalterrorism data base Original - Excel																											
Muskan Gupta MG																											
File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do																											
<i>LIMITED TIME OFFER For just ₹ 100, get 3 months of Microsoft 365 Family and share with up to 5 people. It is like getting six subscriptions in one. TERMS APPLY.</i> Redeem now																											
A1	eventid	year	imonth	iday	approxdat	extended	resolution	country	country_b	region	region_txt	provstate	city	latitude	longitude	specificity	vicinity	location	summary	crit1	crit2	crit3					
1	1.97E+11	1970	7	2		0		58 Dominican	2 Central America & Ca	Santo Don	18.45679	-69.9512	1	0						1	1						
2	1.97E+11	1970	0	0		0		130 Mexico	1 North Am Federal	Mexico cit	19.37189	-99.0866	1	0						1	1						
3	1.97E+11	1970	1	0		0		160 Philippines	5 Southeast	Tarlac	15.4786	120.5997	4	0						1	1						
4	1.97E+11	1970	1	0		0		78 Greece	8 Western E	Attica	37.99749	23.76273	1	0						1	1						
5	1.97E+11	1970	1	0		0		101 Japan	4 East Asia	Fukouka	33.58041	130.3964	1	0						1	1						
6	1.97E+11	1970	1	0		0		217 United Sta	1 North Am Illinois	Cairn	37.00511	-89.1763	1	0						1/1/1970:	1	1					
7	1.97E+11	1970	1	1		0		218 Uruguay	3 South Am Montevide	Montevide	-34.89192	-56.1872	1	0						1	1						
8	1.97E+11	1970	1	2		0		217 United Sta	1 North Am California	Oakland	37.79193	-122.226	1	0						0 Edes Subst	1/2/1970:	1	1				
9	1.97E+11	1970	1	2		0		217 United Sta	1 North Am Wisconsin	Madison	43.07659	-89.4125	1	0						1/2/1970:	1	1					
10	1.97E+11	1970	1	2		0		217 United Sta	1 North Am Wisconsin	Madison	43.07295	-89.3867	1	0						1/3/1970:	1	1					
11	1.97E+11	1970	1	3		0		217 United Sta	1 North Am Wisconsin	Madison	43.07295	-89.3867	1	0							1	1					
12	1.97E+11	1970	1	1		0		217 United Sta	1 North Am Wisconsin	Baraboo	43.4685	-89.7443	1	0							1	1					
13	1.97E+11	1970	1	6		0		217 United Sta	1 North Am Colorado	Denver	39.75897	-104.876	1	0						1/6/1970:	1	1					
14	1.97E+11	1970	1	8		0		98 Italy	8 Western E	Lazio	41.89096	12.49007	1	0							1	1					
15	1.97E+11	1970	1	9		0		217 United Sta	1 North Am Michigan	Detroit	42.33169	-83.0479	1	0						1/9/1970:	1	1					
16	1.97E+11	1970	1	9		0		217 United Sta	1 North Am Puerto Rico	Rio Piedra	18.38693	-66.0611	1	0						0 Caparra St	1/9/1970:	1	1				
17	1.97E+11	1970	1	10		0		499 East Germ	9 Eastern Eu	Berlin	52.50153	13.40185	1	0							1	1					
18	1.97E+11	1970	1	11		0		65 Ethiopia	11 Sub-Saharan	Unknown	5	0								1	1						
19	1.97E+11	1970	1	12		0		217 United Sta	1 North Am New York	New York	40.69713	-73.9314	1	0						Brooklyn	1/12/1970	1	1				
20	1.97E+11	1970	1	12		0		217 United Sta	1 North Am Puerto Ric	Rio Grande	18.38	-65.8309	1	0						1/12/1970	1	1					
21	1.97E+11	1970	1	13		0		217 United Sta	1 North Am Washington	Seattle	47.61079	-122.331	1	0							1/13/1970	1	1				
22	1.97E+11	1970	1	14		0		217 United Sta	1 North Am Illinois	Champaign	40.11675	-88.2393	1	0							Champaign	1/14/1970	1	1			

Fig 3.2.1: Screenshot of the original uncleaned dataset having many columns without much information

globalterrorism data base Original - Excel																									
Muskan Gupta MG																									
File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do																									
<i>LIMITED TIME OFFER For just ₹ 100, get 3 months of Microsoft 365 Family and share with up to 5 people. It is like getting six subscriptions in one. TERMS APPLY.</i> Redeem now																									
S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO			
4	1	1	1	0		0	1	0	1	0	1	Assassination						10 Journalists	54 Radio Jour	Voice of A Employee	217 Ur				
5	1	1	1	0		0	1	0	1	0	3	Bombing/Explosion						7 Governme	46 Embassy/Consulate	U.S. Emba	217 Ur				
6	1	1	1	-9		0	1	0	1	0	7	Facility/Infrastructure Attack						7 Governme	46 Embassy/Consulate	U.S. Cons	217 Ur				
7	1/1/1970:	1	1	1	0		0	1	0	2	Armed Assault						3 Police	22 Police	Bull Cairo Polic	Cairo Polic	217 Ur				
8	1	1	1	0		0	0	0	0	0	1	Assassination						3 Police	25 Police Secu	Uruguay Juan Mari	Uruguay	218 Ur			
9	1/2/1970:	1	1	1	1	2	Other Crim	0	1	0	3	Bombing/Explosion						21 Utilities	107 Electricity	Pacific Gas Edes Subst	Electricity	217 Ur			
10	1/2/1970:	1	1	1	0		0	1	0	7	Facility/Infrastructure Attack						4 Military	28 Military	R&O.T.C. R.O.T.C. o	Military	217 Ur				
11	1/3/1970:	1	1	1	0		0	1	0	7	Facility/Infrastructure Attack						2 Governme	21 Governme	Selective S Selective S	Selective S	217 Ur				
12	1	1	0	1	1	1	Insurgency	0	0	0	3	Bombing/Explosion						4 Military	27 Military	Barracks/Bas Badger Arr	Military	217 Ur			
13	1/6/1970:	1	1	1	1	2	Other Crim	0	1	0	7	Facility/Infrastructure Attack						4 Military	28 Military	Re Army Recr Army Recr	Re Army Recr	217 Ur			
14	1	1	1	-9		0	1	0	4	4	Hijacking						6 Airports &	42 Aircraft	(n)Trans Wor Flight 802	Aircraft	217 Ur				
15	1/9/1970:	1	1	1	0		0	1	0	7	Facility/Infrastructure Attack						2 Governme	21 Governme	U.S. Gover Packard Pr	U.S. Gover	217 Ur				
16	1/9/1970:	1	1	1	1	2	Other Crim	0	1	0	7	Facility/Infrastructure Attack						1 Business	7 Retail	Gro American Baker's Stc	Retail	217 Ur			
17	1	1	1	0		0	1	0	3	Bombing/Explosion						2 Government	2 Government (General)	Jurists Ball	Jurists Ball	217 Ur					
18	1	1	0	1	1	1	Insurgency	0	1	0	9	Unknown						4 Military	34 Military	Pe U.S. Army Soldier	Pe U.S. Army	217 Ur			
19	1/12/1970	1	1	1	0		0	1	0	3	Bombing/Explosion						8 Education	49 School	Un High Schoo James Mac	Education	217 Ur				
20	1/12/1970	1	1	1	0		0	1	0	3	Bombing/Explosion						1 Business	4 Multiatio	General El General El	General El	217 Ur				
21	1/13/1970	1	1	1	0		0	1	0	7	Facility/Infrastructure Attack						1 Business	7 Retail	Gro Fusion's De Fusion's De	Retail	217 Ur				
22	1/14/1970	1	1	1	0		0	1	0	7	Facility/Infrastructure Attack						3 Police	22 Police	Bull Police Dep Champaigr	Bull Police	217 Ur				
23	1	1	1	1	0		0	1	0	2	Armed Assault						8 Education	49 School	University/Ed Secondary	University/Ed Secondary	218 Ur				
24	1/17/1970	1	1	1	0		0	1	0	3	Bombing/Explosion						8 Education	49 School	Un Seattle Un Liberal Art	Liberal Art	217 Ur				
25	1/17/1970	1	1	1	0		0	0	0	3	Bombing/Explosion						4 Military	28 Military	Re R.O.T.C. Air Force F	Air Force F	217 Ur				
26	1/19/1970	1	1	1	0		0	1	0	2	Facility/Inf	2 Armed Assault						22 Violent Po	110 Party Offic	Black Pant Headquart	Black Pant Headquart	217 Ur			
27	1	1	1	0		0	1	0	1	0	1	Assassination						7 Governme	46 Embassy/C British con Bodyguard	JUSMAG H	Bodyguard	216 Gr			
28	1	1	1	0		0	1	0	3	Bombing/Explosion						7 Governme	46 Embassy/Consulate	J USMAG H	J USMAG H	217 Ur					
29	1	1	1	0		0	0	1	0	6	Hostage Taking (Kidnapping)						1 Business	7 Retail	Gro Father ow Leon Jacol	Rather ow Leon Jacol	222 Ve				
30	1/22/1970	1	1	1	0		0	1	0	3	Bombing/Explosion						14 Private Cit	76 House/API	Private residence	Private residence	217 Ur				

Fig 3.2.2: Screenshot of the original uncleaned dataset having many columns without much information

3.2 Dataset Cleaning up

The retrieved dataset had the following characteristics when uploaded to the Jupyter Notebook:

Size: 181,691 rows

Columns: 135

DataTypes: Float64(55), Int64(22), Object(58)

Memory Usage: 187.1 Mb

The screenshot shows a Jupyter Notebook interface with the following details:

- Header:** jupyter globalTerrorismData Last Checkpoint: 20/05/2022 (autosaved) | Logout | Trusted | Python 3
- In [1]:** Import statements:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go
import seaborn as sns
```
- In [3]:** Data loading:

```
data = pd.read_csv("/Users/kuhuchawla/Downloads/globalterrorismdb_0718dist.csv", encoding='ISO-8859-1')
```

A warning message is displayed below the code:

```
/Users/kuhuchawla/opt/anaconda3/lib/python3.7/site-packages/IPython/core/interactiveshell.py:3063: DtypeWarning: Columns (4,6,31,33,61,62,63,76,79,90,92,94,96,114,115,121) have mixed types. Specify dtype option on import or set low_memory=False.
    interactivity=interactivity, compiler=compiler, result=result)
```
- In [4]:** Data preview:

```
data.head()
```

Out [4]: A Pandas DataFrame with 5 rows and 135 columns. The columns include eventid, iyear, imonth, iday, approdate, extended, resolution, country, country_txt, region, ..., addnotes, scite1, scite2, scite3, dbsource, and INT_LOG.
- Bottom:** 5 rows x 135 columns

Fig 3.3 Screenshot of data cleaning using Python

The null values were identified and columns with relevant information were extracted.

The columns were renamed for easier analysis.

The Cleaned Up data is Obtained.

The screenshot shows a Microsoft Excel spreadsheet with the following details:

- Header:** File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do
- Sheet:** A181692 | 181690
- Columns:** A, B, C, D, E, F, G, H, I, J
- Rows:** 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, ...
- Content:** The data is organized into several columns:
 - Year, Month, Day (B, C, D)
 - Country (E)
 - state (F)
 - Region (G)
 - city (H)
 - latitude (I)Specific rows include:
 - Row 2: 1970, 7, 2, Dominican Republic, Central America & Caribbean, Santo Domingo, 18.45679
 - Row 3: 1970, 0, 0, Mexico, Federal, North America, Mexico city, 19.37189
 - Row 4: 1970, 1, 0, Philippines, Tarlac, Southeast Asia, Unknown, 15.4786
 - Row 5: 1970, 1, 0, Greece, Attica, Western Europe, Athens, 37.99749
 - Row 6: 1970, 1, 0, Japan, Fukuoka, East Asia, Fukuoka, 33.58041
 - Row 7: 1970, 1, 1, United States, Illinois, North America, Cairo, 37.00511
 - Row 8: 1970, 1, 2, Uruguay, Montevideo, South America, Montevideo, -34.8912
 - Row 9: 1970, 1, 2, United States, California, North America, Oakland, 37.79193
 - Row 10: 1970, 1, 2, United States, Wisconsin, North America, Madison, 43.07659
 - Row 11: 1970, 1, 3, United States, Wisconsin, North America, Madison, 43.07295
 - Row 12: 1970, 1, 1, United States, Wisconsin, North America, Baraboo, 43.4685
 - Row 13: 1970, 1, 6, United States, Colorado, North America, Denver, 39.75897
 - Row 14: 1970, 1, 8, Italy, Lazio, Western Europe, Rome, 41.89096
 - Row 15: 1970, 1, 9, United States, Michigan, North America, Detroit, 42.33169
 - Row 16: 1970, 1, 9, United States, Puerto Rico, North America, Rio Piedras, 18.38693
 - Row 17: 1970, 1, 10, East Germany (GDR), Berlin, Eastern Europe, Berlin, 52.50153
 - Row 18: 1970, 1, 11, Ethiopia, Unknown, Sub-Saharan Africa, Unknown, ...

Fig 3.4 Screenshot of cleaned dataset

The dataset now cleaned up is used for exploratory analysis using python where we analyze the data using graphs, charts, pie charts etc. to get an idea about what attributes are present.

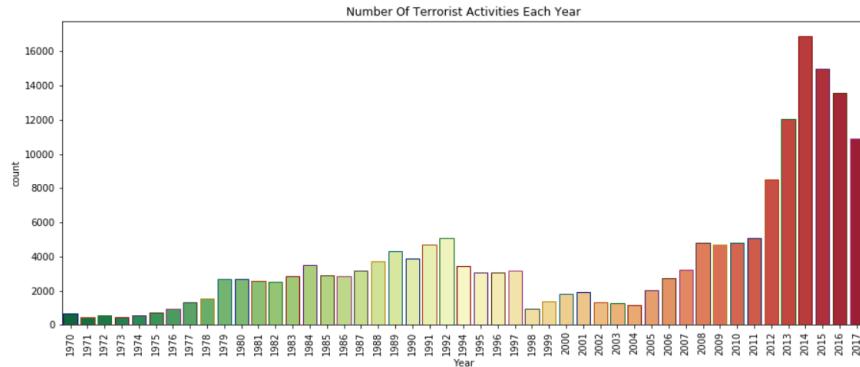


Fig 3.5

```
plt.subplots(figsize=(15,6))
year_cas = data.groupby('Year').Casualties.sum().to_frame().reset_index()
year_cas.columns = ['Year','Casualties']
sns.barplot(x=year_cas.Year, y=year_cas.Casualties, palette='RdYlGn_r', edgecolor=sns.color_palette('dark',10))
plt.title('Number Of Casualties Each Year')
plt.show()
```

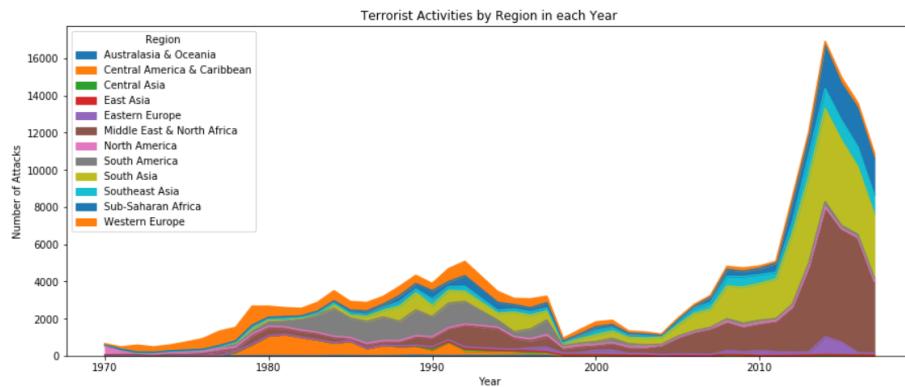


Fig 3.6

```
plt.subplots(figsize=(15,6))
sns.barplot(data['Country'].value_counts()[:15].index,data['Country'].value_counts()[:15].values,palette='mako')
plt.title('Top Countries Affected')
plt.xlabel('Countries')
plt.ylabel('Count')
plt.xticks(rotation= 90)
plt.show()
```

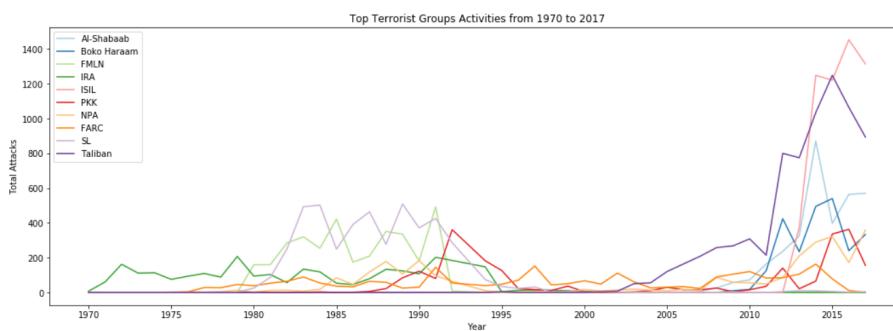


Fig 3.7

```
# Total Number of people killed in terror attack
killData = data.loc[:, 'Killed']
print('Number of people killed by terror attack:', int(sum(killData.dropna())))# drop the NaN values
```

Number of people killed by terror attack: 411868

```
# attackData
attackData = data.loc[:, 'AttackType']
typeKillData = pd.concat([attackData, killData], axis=1)
typeKillFormatData = typeKillData.pivot_table(columns='AttackType', values='Killed', aggfunc='sum')
typeKillFormatData
```

3.3 Cleaned up Dataset Uploaded to Tableau

The data cleaned using python was then uploaded to tableau which was meant to be our intended software for a better visual analysis as well as a comprehensive and detailed analysis with interactive features such as Bar Charts, Line Charts, Bubble Charts, Maps, etc.

3.4 Creation of Worksheets in Tableau

Tableau worksheets have been created. We have used various plots to depict the data:

1. Bar Charts
2. Line Charts
3. Bubble Charts
4. Maps

3.5 Final Analysis

The worksheets obtained have been combined to form dashboards. We have made interactive dashboards which have the ability to change according to the region selected

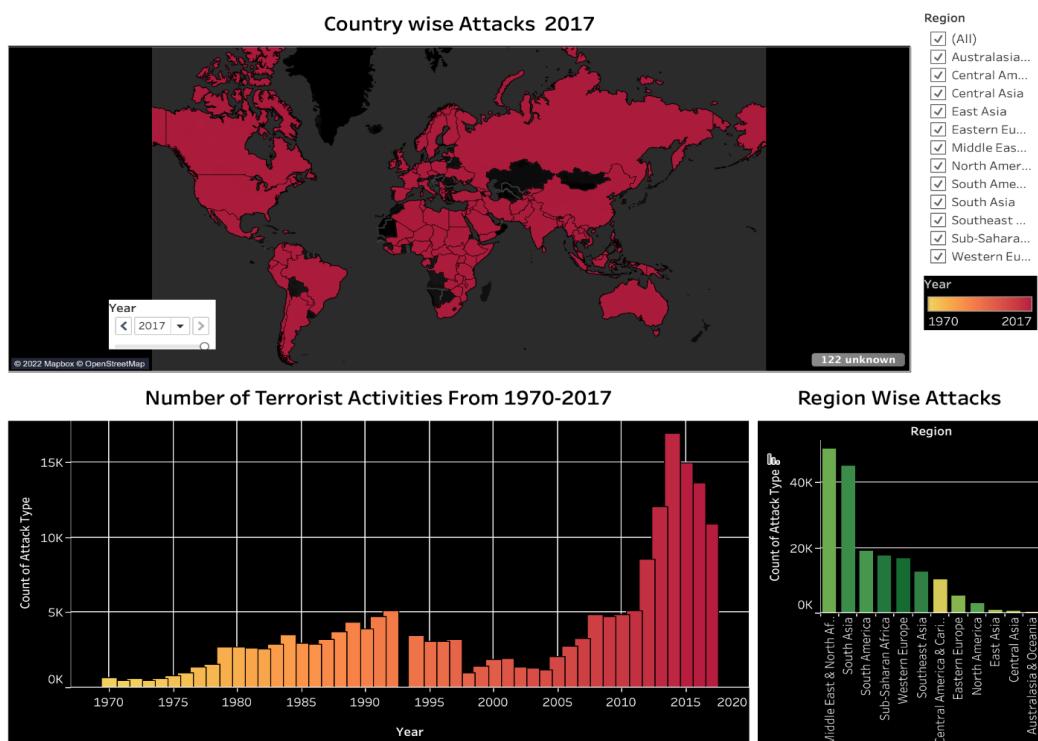


Fig 3.8 Screenshot of Dashboard 1

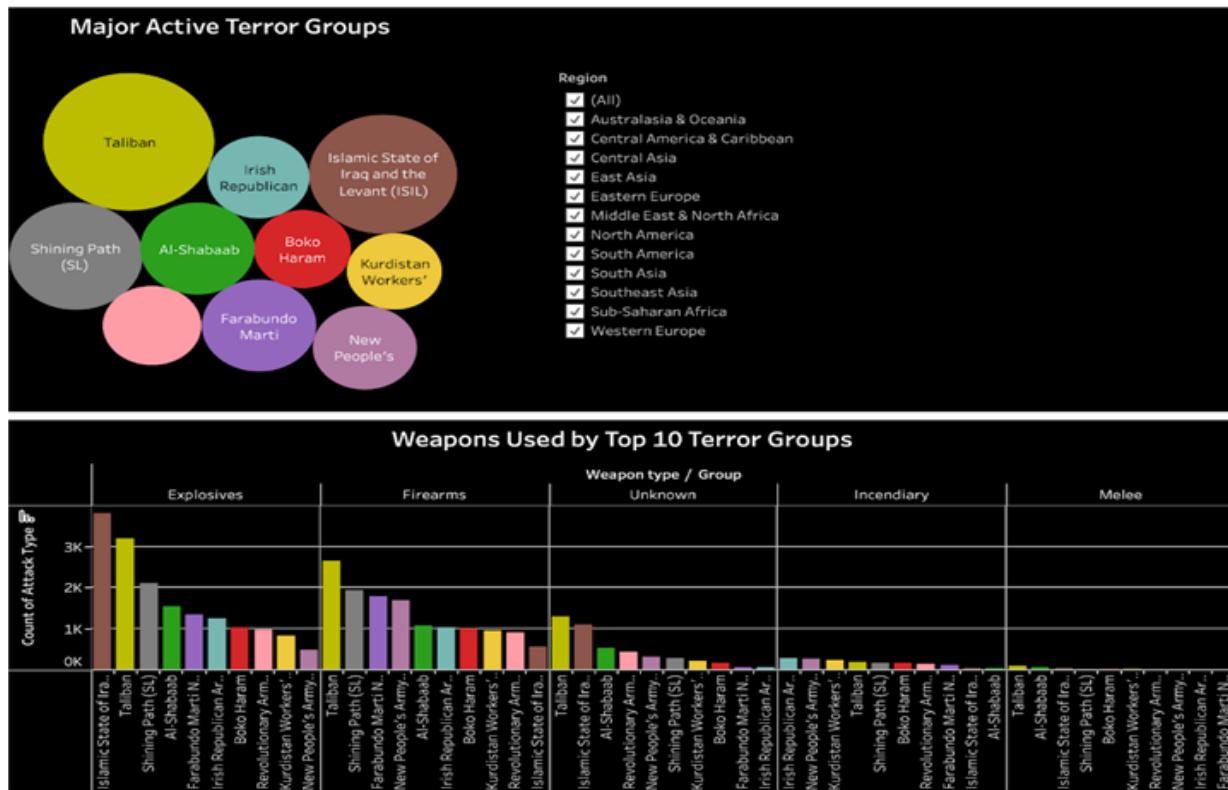


Fig 3.9 Screenshot of Dashboard 2

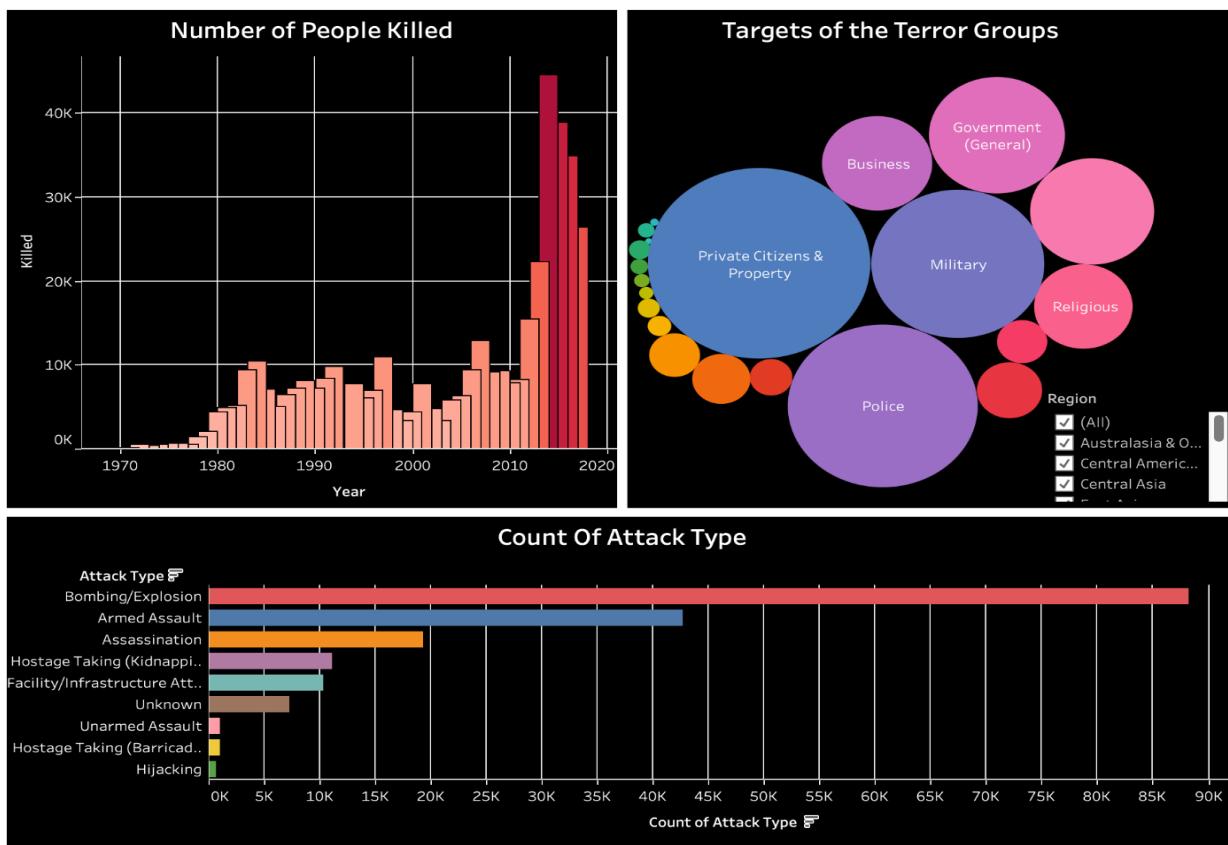


Fig 3.10 Screenshot of Dashboard 3

CHAPTER 4

RESULTS AND DISCUSSIONS

We have obtained three interactive dashboards showing the global terrorism data from the year 1970 to 2017. The interactive dashboards are based on regions i.e we are able to showcase data from different regions and analyse each region individually as well as a whole.

The dataset has been divided into 12 regions:

1. Australasia and Oceania
2. Central America and Caribbean
3. Central Asia
4. East Asia
5. Eastern Europe
6. Middle East and North Africa
7. North America
8. South America
9. South Asia
10. Southeast Asia
11. Sub-Saharan Africa
12. Western Europe

Since we can select the region which we want to analyze we have taken the South Asian region as our primary region of study.

4.1 Trends in Terrorist Events

We can see that in 1970s there were very few or negligible terror attacks in the South Asian Region. The number of terrorist activities started to significantly increase from 2005 with the

lowest number of incidents being 604. The South Asian Region recorded the highest number of incidents in the year 2014 with 4,998 attacks. The year 2014 is also categorised as the year with the occurrence of most terror attacks around the globe equating to 16,903 attacks.

The highest number of terror activities combined from the year 1970 to 2017 is in the Middle East and North African region. Thus, this region is regarded as the most unsafe region in the world.

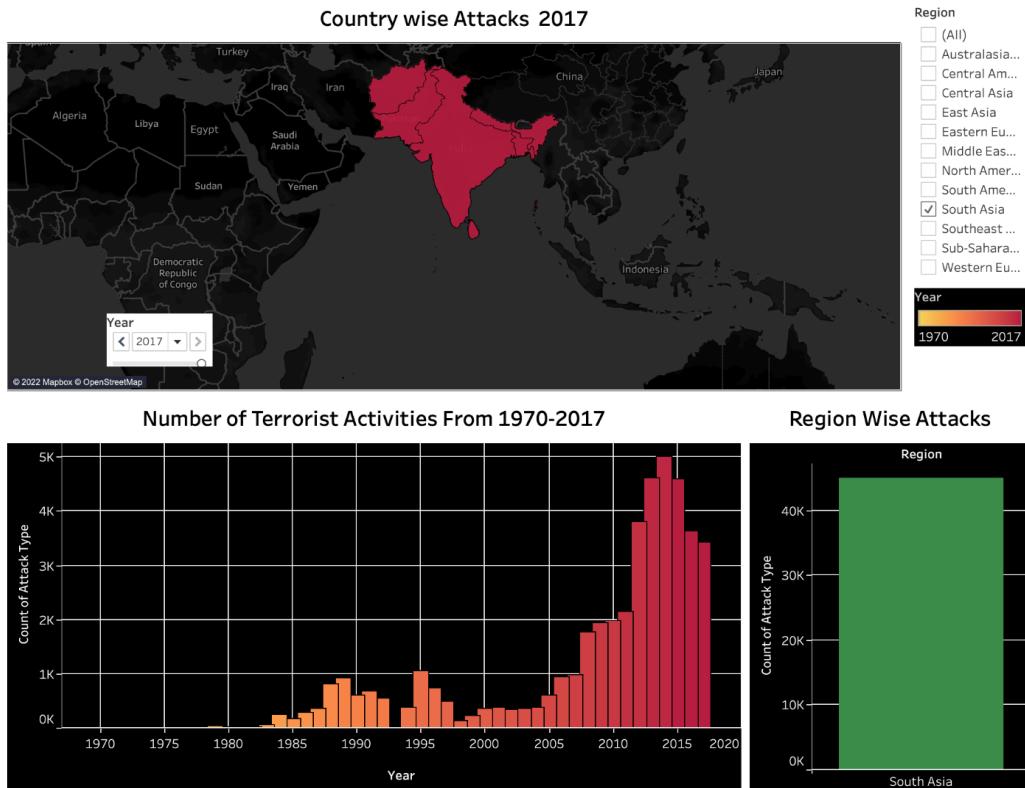


Fig 4.1: Trends in Terrorist Events

4.2 Region Wise Statistics

1. Top 3 types of attacks across all regions: Bombing/Explosion, Armed/Unarmed Assault, and Facility/Infrastructure Attack
2. Top 3 most frequently used weapons: Explosives, Firearms, and Incendiary
3. Most affected types of targets
 - North America and Western Europe: Business, Private Citizen & Property
 - South Asia: Private Citizen & Property, Police

- East Asia: Government, Business
- Middle East and North Africa: Private Citizen & Property, Military

For the South Asian Region we can see that the major attack type is Explosion followed by Armed Assault and Assassination. The major targets of this region are private citizens followed by the police and then the military.

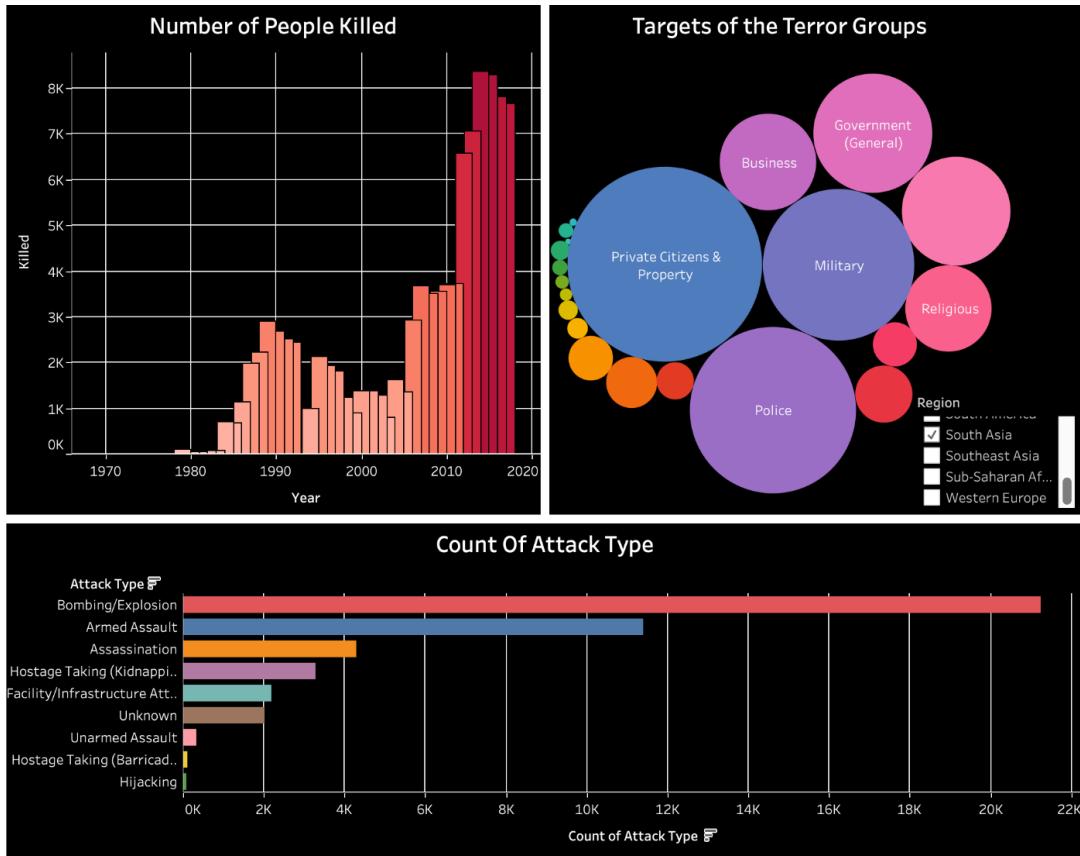


Fig 4.2: Region Wise Statistics

4.3 Information About the Active Terror Groups and Weapons

Most Active Terror Groups in the regions:

1. Middle East & North America: Islamic State of Iraq and the Levant (ISIL), Kurdistan Workers Party (PKK)

2. North America and Western Europe: Irish Republican Army(IRA),Kurdistan Workers Party(PKK)
3. South Asia: Taliban
4. East Asia: Irish Republic Army

Weapons Used by the Active Terror Groups Region Wise:

1. Middle East & North America:Explosives, Firearms, Incendiary
2. North America and Western Europe: Explosives, Firearms, Incendiary,Melee
3. South Asia: Explosives, Firearms, Incendiary
4. Australasia and Oceania: Explosives

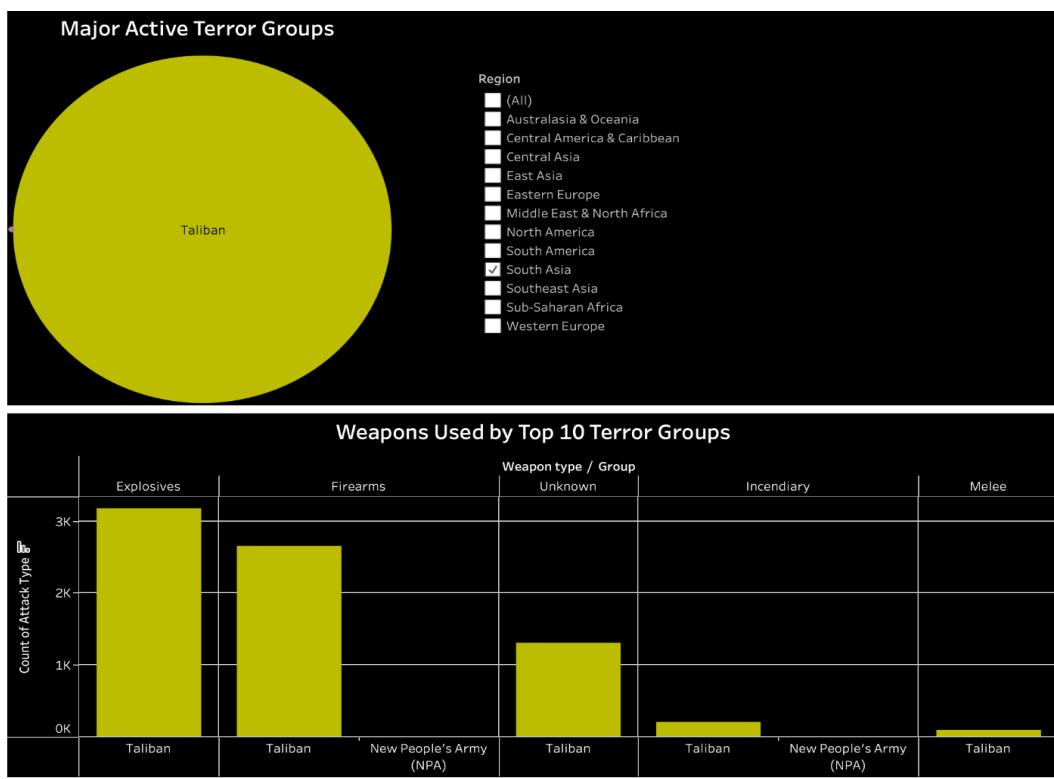


Fig 4.3: Active Terror Groups and Weapons

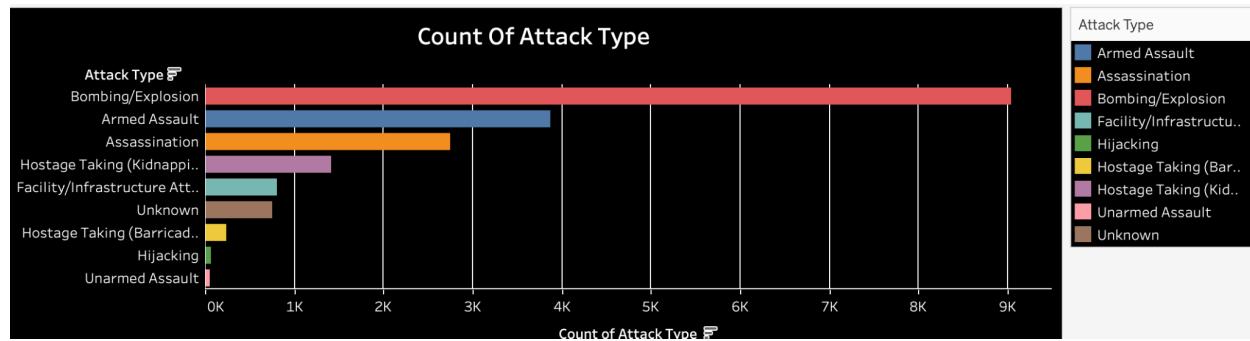


Fig 4.4: Count of Attack type

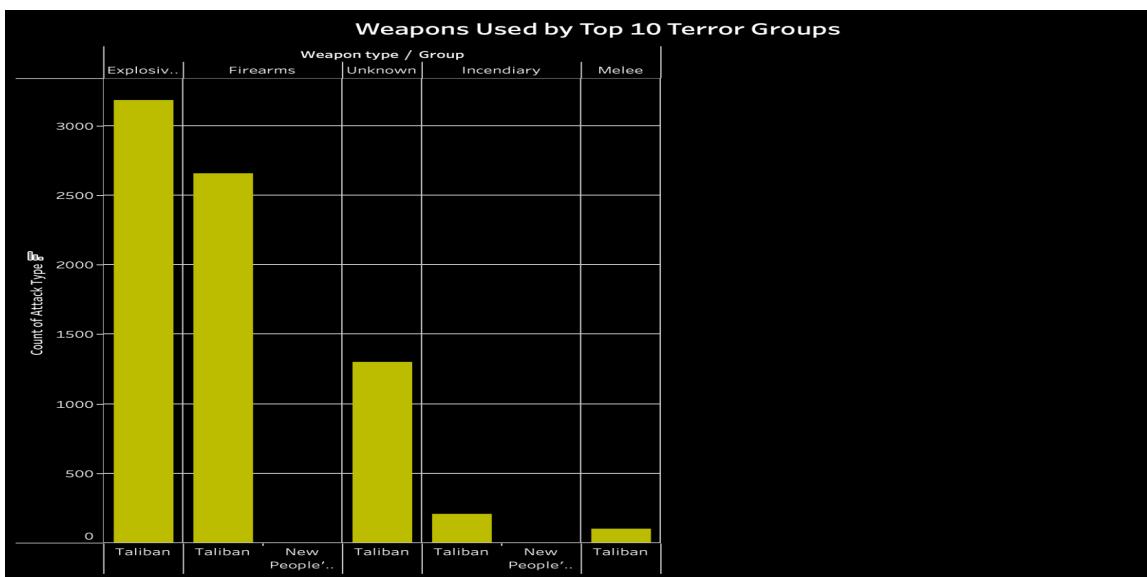


Fig 4.5: Weapons used for attacks in South Asia

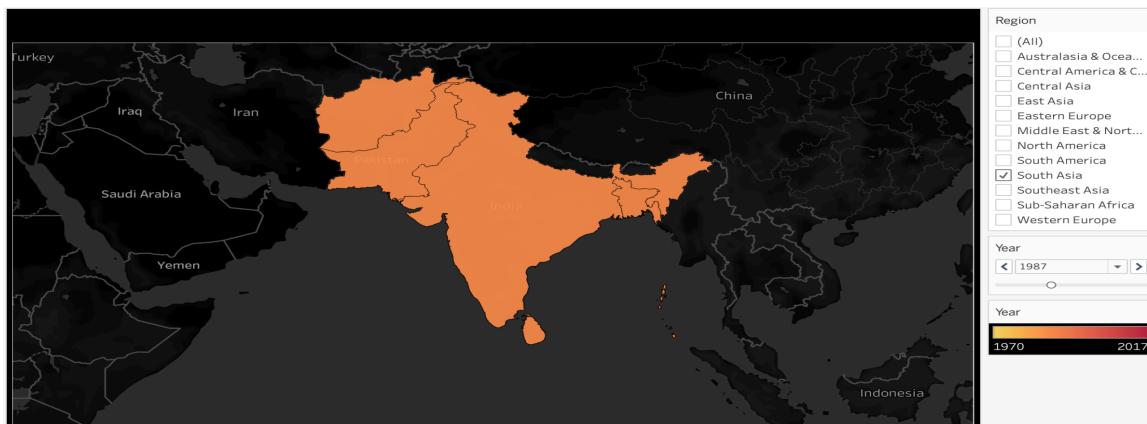


Fig 4.6: Areas Affected in South Asia

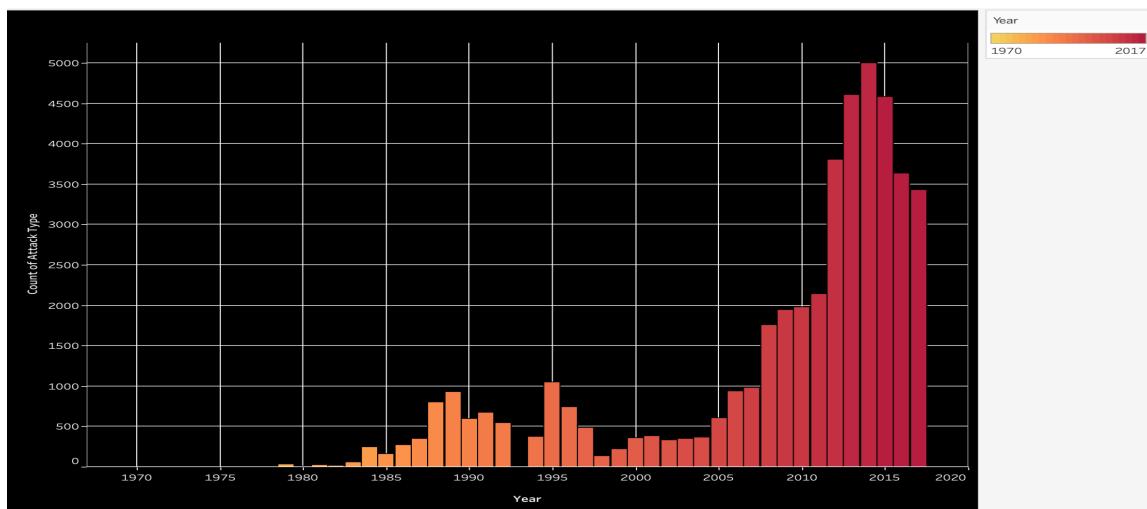


Fig 4.7: Attacks over years in South Asia

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

The Global Terrorism Dashboard (GTD) is a report with an aim to examine trends and to help inform a positive and practical debate about the future of terrorism and the required policy responses. It provides a comprehensive summary of the key global trends and patterns in terrorism since 1970. It is an attempt to systematically label the nations of the world according to terrorist activities within them. The index combines a number of factors associated with terrorist attacks to build an explicit picture of the impact of terrorism, illustrating trends, and providing a data series for analysis by researchers and policymakers.

The Dashboard is based on data from the Global Terrorism Database (GTD) which is collected and collated by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of Maryland. The GTD has codified over 190,000 cases of terrorism, it covers 180+ countries, consisting 99.7% of the world's population.

The aim is to examine trends and to help inform a positive and practical debate about the future of terrorism and the required policy responses.

The following points could be concluded from our analysis:

- From the Dashboard, we are able to analyse which regions should focus on which terror group.
- We are also able to give an output of which weapon trade the specific region should focus on.
- It can be seen that in year 2011 terror attacks had spiked-up and reached its crescendo around 2014-15, killing and devastating many individuals.

It was largely due to the Arab-spring which engulfed around that time in a spark to free countries from dictatorship that actually gave way to the creation of many small, but heavily equipped militant groups, which grew in size over a short period of time and

swept away the liveliness from the mind and maps of regions of North Africa and the middle east.

- The sub-Saharan and the middle east was enwrapped by ISIL and ISIS.
- The region of South Asia has resulted in the most number of attacks, by number, with Taliban being the top conspirator.

Thick population being a favourable condition for the terror attacks, the south Asian regions of Afghanistan, Pakistan and India are susceptible but with the help of an analysis such as the one we have done vigilant security forces working day in and day out, and the intelligence bureau, the attacks have reduced in the past years

- Terrorism in the West declined substantially, since 2012.
- Deaths from terrorism decreased as compared to their peak in 2014.
- Sub-Saharan Africa accounted for most of global terrorism deaths.
- Taliban is the world's deadliest terror group in 2017.
- South Asia recorded the largest deterioration, followed by Central America and the Caribbean.
- However, Central America and the Caribbean remains the region with the lowest impact of terrorism, a position it has held for the past 17 years.
- Abnormality in dataset: Data for the year 1993 is unavailable.

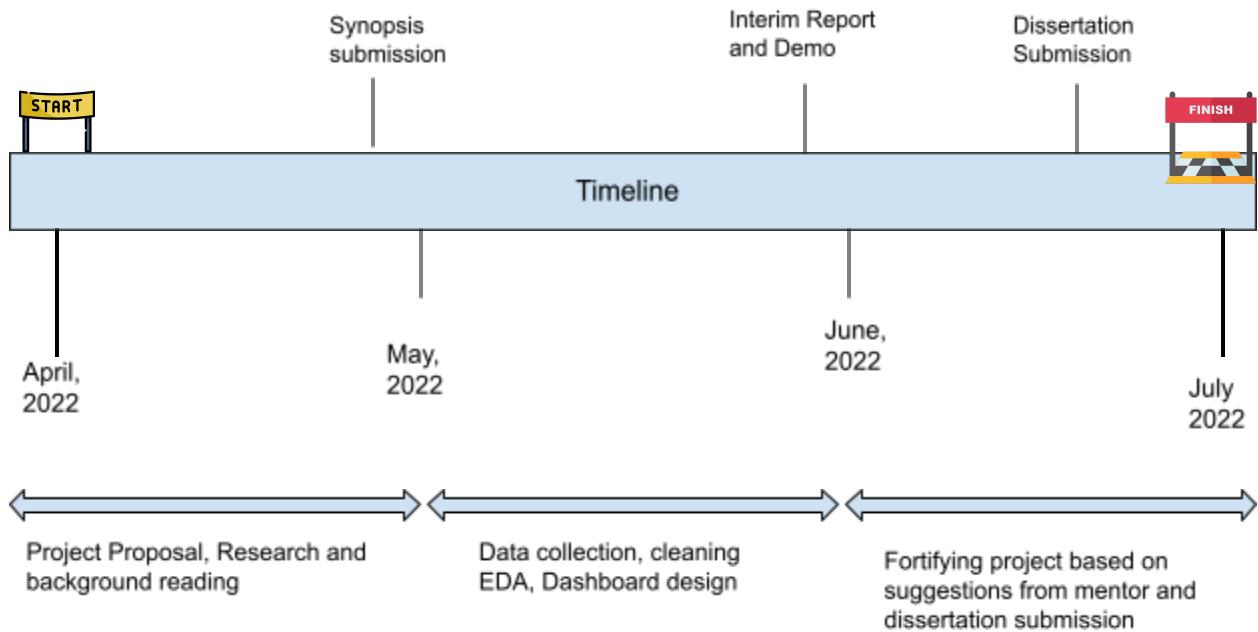
5.2 Future scope

A good analysis can be done with a large data set, so this GTD shall be instrumental in curbing terror attacks for a long time and regular updates will make it rather more reliable, reducing the impact of terror attacks in the years to come.

Apart from just terrorism analysis, the knowledge and implementation of data analytics can help companies to spruce up their marketing strategies by keeping a watch on what people like.

TIMELINE

The project can be said to have been made in three separate phases of its working where our team performed based on the targets ahead of the self decided deadlines. Starting off from April, when the project idea of Global Terrorism Dashboard using Tableau was formally adopted, the team started exploration for the acquisition of all prerequisites and all plausible implementations, market value and future scope of the project. After we had enough knowledge and resources available at our disposal, we submitted a synopsis of the project that we were intending to work on over the next couple of months. With all checkboxes ticked and ideas approved, the team drove into an extensive exploration for the best dataset available which could be coupled with this project. Global Terrorism Database [1] by the University of Maryland was then chosen. The dataset was a conglomerate of hundreds of columns and many null values which required cleaning before any other advancements.



The dataset was cleaned using python and as a “code-of-conduct” for any other data analytics project, we ran our data through an exploratory data analysis (EDA) followed by which, the real, visible, part of the project began - the dashboard design. Come June, we prepared the project upto 90% of its completion before the Mid-term evaluation (Interim report and demo), presented

before our mentor. Seeking her guidance and possible add-ons, we fortified our project based on the suggestions from our mentor.

At last, the report submission and final project presentation was done.

REFERENCES

- [1] The data of the terrorism dataset collected from Kaggle is maintained by University of Maryland, United States. Link: <https://www.start.umd.edu/gtd/>
- [2] GeeksForGeeks: <https://www.geeksforgeeks.org/introduction-to-tableau/>
- [3] Research Papers:
 - (i) <https://www.tandfonline.com/doi/full/10.1080/09546550701246817> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5462416/>
 - (ii) <https://www.tandfonline.com/doi/full/10.1080/09546550701246817>
 - (iii) <https://www.janes.com/capabilities/military-threat-intelligence/terrorism-and-insurgency>