

Global Terrorism Analysis

Pavan Kale

Data science trainees,
AlmaBetter, Bangalore

Abstract:

The Global Terrorism Database (GTD) is the world's most comprehensive, open-source terrorism database. It includes information on more than 180,000 terrorist attacks around the world from 1970 through 2017. It provides a more complete understanding of the dynamics, causes and consequences of terrorism around the world, by allowing its users to analyze patterns such as the frequency of terrorist attacks, geo-spatial patterns of terrorist attacks, the lethality of terrorist attacks, patterns of casualties including injured persons and hostages, the emergence and prevalence of particular tactics and targeting strategies used in terrorist attacks, and the evolution of perpetrators of terrorist attacks.

Keywords: *terrorism, global terrorism database, geo-spatial.*

1. Problem Statement

Terrorist attacks are the challenging issue across the world and need the attention of the practitioners to cope up deliberately predicting the responsible group of an event is a complicated task due to lack of in depth terrorist historical data. Data mining classification techniques are largely used resolve the problem.

Terrorism is a complex political and social phenomenon. Terrorist attacks have a significant threat to the safety and security of the international community and have become one of the greatest obstacles to the sustainable development of global social security. Antiterrorism is an important part of global security governance, which is a sustainability issue that guarantees global security development. At present, terrorist attacks occur frequently, which leads to significant threats and poses a challenge to global social security governance.

2. Introduction

The Global Terrorism Database (GTD) is the most comprehensive unclassified database of terrorist attacks in the world. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) makes the GTD available via this site in an effort to improve understanding of terrorist violence, so that it can be more readily studied and defeated. The GTD is produced by a dedicated team of researchers and technical staff.

The GTD is an open-source database, which provides information on domestic and international terrorist attacks around the world since 1970, and now includes more than 200,000 events. For each event, a wide range of information is available, including the date and location of the incident, the

weapons used, nature of the target, the number of casualties, and when identifiable – the group or individual responsible.

3. Characteristics of GTA

Contain information on over 200,000 terrorist attacks. Currently the most comprehensive unclassified on terrorist attacks in the world. Includes the information on more than 88,000 bombings, 19,000 assassinations and 11,000 wound since 1970, Includes information on at least 45 variables for each case, with more recent incidents including information on more than 120 variables. More than 4,000,000 news articles and 25,000 news sources were reviewed to collect incident data from 1970 to 2017.

4. Effects of Terrorism

The effects of terrorism are:

- **Direct Economic Destruction**

The most immediate and measurable impact of terrorism is physical destruction. Terrorists destroy existing plants, machines, transportation systems, workers, and other economic resources. On smaller scales, acts of terrorism may blow up cafes, churches, or roads.

- **Increased Uncertainty in the Markets**

Even if you do not live anywhere near terrorist attacks, you might still

be negatively impacted indirectly. This is because all kinds of markets hate uncertainty, and terrorism creates a lot of it.

- **Insurance, Trade, Tourism, and FDI**

There are two obvious industries especially vulnerable to the effects of terrorism: insurance and tourism. Not all insurance companies payout in the event of international terrorism or foreign wars, so the impact is likely less than you might first expect. Insurance, Trade, Tourism, and FDI

- **War Is the Health of the State**

There is an old saying in the study of political economy that reads "war is the health of the state." It means that during times of conflict, reactive governments and nervous citizens are far more inclined to give up economic and political freedoms in exchange for security.

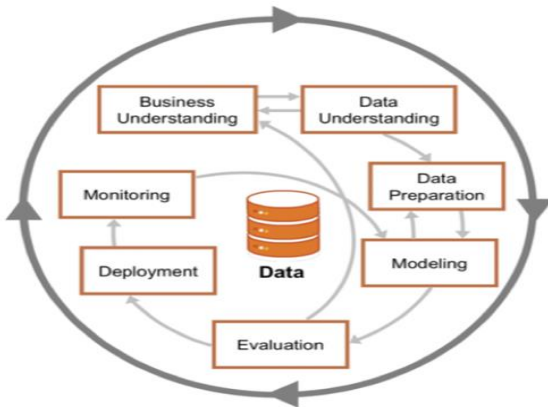
- **Increased Nationalism and Foreign Skepticism**

The final risk to the economy is a political risk. This is already on display in the United States and Europe in 2016, where there has been a rise in skepticism of foreign cultures, businesses, immigrant workers, and refugees. Populist movements already won a victory of sorts in the United Kingdom, where

anti-globalist and anti-trade sentiments helped pass Brexit.

most time-consuming stages in Data Science. However, in order to prevent wrongful predictions, it is important

5. Steps involved:



- **Define Problem Statement**

Before you even begin a Data Science project, you must define the problem you're trying to solve. At this stage, you should be clear with the objectives of your project.

- **Data Collection**

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

- **Null values Treatment**

Our dataset contains a large number of null values which might tend to disturb our accuracy hence we dropped them.

Data cleaning is the process of removing redundant, missing, duplicate and unnecessary data. This stage is considered to be one of the

- **Data Analysis and Exploration**

Once you're done cleaning the data, it is time to get the inner Sherlock Holmes out. At this stage in a Data Science life-cycle, you must detect patterns and trends in the data. This is where you retrieve useful insights and study the behavior of the data. At the end of this stage, you must start to form hypotheses about your data and the problem you are tackling.

- **Data Modelling**

This stage is all about building a model that best solves your problem. A model can be a Machine Learning Algorithm that is trained and tested using the data. This stage always begins with a process called Data Splicing, where you split your entire data set into two proportions. One for training the model (training data set) and the other for testing the efficiency of the model (testing data set).

- **Optimization and Deployment**

This is the last stage of the Data Science life-cycle. At this stage, you must try to improve the efficiency of the data model, so that it can make more accurate predictions. The end goal is to deploy the model into production or production-like environment for final user acceptance. The users must validate the performance of the models and if

there are any issues with the model then they must be fixed in this stage.

6. Exploratory Data

Analysis:

I. Incident by Year

Large number of Terrorist activity happen in 2014 and 2015. And least activity was recorded in 1971. We begin our review of the GTD database with event counts by year.

Table 1. Distribution of Incidents for Years 1970-2017.

Year	Frequency
1970	651
1971	471
1972	568
1973	473
1974	581
1975	740
1976	923
1977	1319
1978	1526
1979	2662
1980	2662
1981	2586
1982	2544
1983	2870
1984	3495
1985	2915
1986	2860
1987	3183
1988	3721
1989	4324
1990	3887
1991	4683
1992	5071

1994	3456
1995	3081
1996	3058
1997	3197
1998	934
1999	1395
2000	1814
2001	1906
2002	1333
2003	1278
2004	1166
2005	2017
2006	2758
2007	3242
2008	4805
2009	4721
2010	4826
2011	5076
2012	8522
2013	12036
2014	16903
2015	14965
2016	13587
2017	10900
Total	1,181,691

II. Type of Attacks

Various types happen from 1970 to 2017 and below we discuss the highest frequency attack type. Bombings and Armed Assaults attacks were by far the most common, jointly accounting for more than 75 percent of all incidents. The next most common event type was the assassination, account for over 10% of total incidents. Kidnappings were far less common, account for a little more than 6% of total events. Aerial hijackings, Barricade for less than 1% of total cases. Some attacks type are not identified which account for 4%.

Table 2. Distribution of Incidents by Type of Attack.

Type	Frequency	Percent
Armed Assault	42669	23.48%
Assassination	19312	10.63%
Bombing/Explosion	88255	48.57%
Facility/Infrastructure Attack	10356	5.70%
Hijacking	659	0.36%
Barricade Incident	991	0.55%
Kidnapping	11158	6.14%
Unarmed Assault	1015	0.56%
Unknown	7276	4.00%

III. Region

Table 7 gives the distribution of events across regions. The table shows that the Middle East & North Africa, was by far the most common region for terrorist events, accounting for about 28% of all events. Following South Asia having 25%, South America-10% and the Sub-Saharan Africa -9%. Accounting for another i.e. Western Europe, Southeast Asia, Central America & Caribbean, having one-twentieth of all events. Many fewer events were attributed to Eastern Europe, North America, East Asia, Central Asia, Australasia & Oceania whose regional total was just under three percent.

Table 3. Distribution of Incidents by Region of the World.

Region	Frequency	Percent
Middle East & North Africa	50474	27.78%
South Asia	44974	24.75%
South America	18978	10.45%
Sub-Saharan Africa	17550	9.66%
Western Europe	16639	9.16%
Southeast Asia	12485	6.87%

Central America & Caribbean	10344	5.69%
Eastern Europe	5144	2.83%
North America	3456	1.90%
East Asia	802	0.44%
Central Asia	563	0.31%
Australasia & Oceania	282	0.16%

IV. Weapons Used

The type of weapon used was recorded in different cases. The data entry interface was designed to accept up to four different categories of weapon used in each incident in order to account for multiple weapon types used in a single event. We have coded the specific information in these fields into 21 general weapon categories. For 45 example, specific weapon details in the database such as automatic pistols, submachine guns, AK-47's, M-16's and others were categorized as "Automatic Weapons." Table 4 shows the total distribution of weapon categories by combining all four of the weapon fields from the database.

Table 4. Distribution of Incidents by Weapon Type

Weapon	Frequency	Percent
Explosives	92426	50.87%
Firearms	58524	32.21%
Unknown	15157	8.34%
Incendiary	11135	6.13%
Melee	3655	2.01%
Chemical	321	0.18%
Sabotage Equipment	141	0.08%
Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)	136	0.07%
Other	114	0.06%
Biological	35	0.02%
Fake Weapons	33	0.02%
Radiological	14	0.01%

V. Terrorist Group

There are currently 3,537 distinct terrorist groups in the GTD data. In addition, some group names listed in the database are given as generic descriptions of actors, such as “rebels” or “student protesters.” Researchers are defining decision rules using dummy variable coding to incorporate these types of groups as well. Table 5. Shows Distribution of Incidents by (Top-10) Terrorist Group. Taliban and ISIL are the most active Terrorist Groups. Taliban and ISIL account more than 5500 attacks each, then follows SL, FMLN and Al-Shabaab.

Table 5. Distribution of Incidents by (Top-10) Terrorist Group

Terrorist Group	Frequency
Taliban	7478
Islamic State of Iraq and the Levant (ISIL)	5613
Shining Path (SL)	4555
Farabundo Marti National Liberation Front (FMLN)	3351
Al-Shabaab	3288
New People's Army (NPA)	2772
Irish Republican Army (IRA)	2671
Revolutionary Armed Forces of Colombia (FARC)	2487
Boko Haram	2418
Kurdistan Workers' Party (PKK)	2310
Basque Fatherland and Freedom (ETA)	2024

VI. Country

The database includes 205 distinct countries. The country listing also includes separately

Northern Ireland from the rest of the United Kingdom and Corsica from France. In addition, the political circumstances of other countries have changed over time. In every case of political change, we have tried to match the incident to the country name in effect at the time of the incident.

Table 6. Distribution of Incidents by (Top-10) Country

Country	Frequency
Iraq	24636
Pakistan	14368
Afghanistan	12731
India	11960
Colombia	8306
Philippines	6908
Peru	6096
El Salvador	5320
United Kingdom	5235
Turkey	4292

VII. Number of Killed

Among the incidents in which someone was killed, the largest number of killed in one year 2014 was 15800. The Average people kill in each attack during 1970-2017 is 10.

Table 7. Distribution of number of people killed per year.

Year	Killed
1970	174
1971	173
1972	566
1973	370
1974	539
1975	617
1976	674

1977	456
1978	1459
1979	2100
1980	4400
1981	4851
1982	5136
1983	9444
1984	10450
1985	7094
1986	4976
1987	6482
1988	7208
1989	8152
1990	7148
1991	8429
1992	9742
1994	7690
1995	6103
1996	6966
1997	10924
1998	4688
1999	3393
2000	4403
2001	7729
2002	4805
2003	3317
2004	5743
2005	6331
2006	9380
2007	12824
2008	9157
2009	9273
2010	7827
2011	8246
2012	15497
2013	22273
2014	44490
2015	22273
2016	34871
2017	26445

8 Conclusion:

- Iraq ranked first on global terrorist activity followed by Pakistan, Afghanistan then India and so on.
- Most Targeted Areas are Private Citizens & Property, Military, Police and so on.
- Global Terror attack rise sharply from 2011 and Maximum Attacks are in 2014 and Maximum people where died in 2014.
- Taliban and ISIL are the most active terrorist Groups.

References-

1. www.analyticsvidhya.com
2. www.start.umd.edu/gtd.com/
3. <https://www.kaggle.com/>
4. <https://learn.almabetter.com/>