AIT614: Big Data Essentials

**Global Terrorism Analysis**

**Project Progress Report**

**By**

**Snigdha Chekoti**

**Sai Hemanth Nirujogi**

****

**George Mason University, Fairfax VA**

**Introduction**

Terrorist attacks around the world for the past have been increased for the past decade. Terrorism has no accurate definition and defining it is getting harder, because of the terrorism has become a urban term for different acts of violence which are not considered as terrorism before.

“2017 is the second year in a row, the total number of deaths caused by terrorist attacks has been decreased. The reduction in deaths is encouraging, but 2016 was also the third deadliest year since 2000” [[1]](https://reliefweb.int/report/world/global-terrorism-index-2017)

**Problem Statement**

In this project, we are going to analyze the global terrorism database to generate summary statistics and visualizations. The variables with most effect on the incident will be determined from the analysis. We will determine the weights of attributes to understand the importance on the attack, which helps to gain success rates of the attack. From this study of the database, future work can be implemented with counter measures in reduction of terrorism in coming years.

**Dataset**

The global terrorism dataset is an open-source dataset curated with information from different open media sources with systematic data on all the terrorist incidents that have happened from 1970 through 2016. This database consists information on the terrorist attacks occured all over the world including domestic as well as international terrorist attacks that have happened during the time period. This database is created and maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START), located at University of Maryland.

* The dataset includes more than 170,000 cases of terrorist attacks.
* For each incident, details about the incident like date, time, location (with latitude and longitude), type of weapons used, target, number of deaths and is the incident a group or individual agenda.
* Information about more than 83,000 bombings, 11,000 kidnappings and 18,000 assassinations is included in the dataset.
* Each terrorist incident at least has information on 45 variables. [[2]](https://www.kaggle.com/START-UMD/gtd)

**Appendix**

|  |  |
| --- | --- |
| 1. iyear - Year | 1. imonth - Month |
| 1. Iday - Day | 1. approxdate - Approximate Date |
| 1. country - Country | 1. country\_txt-Country |
| 1. region - Region | 1. Region\_txt - Region |
| 1. provstate - Province | 1. vicinity - Vicinity |
| 1. specificity - Geocoding Specificity | 1. attacktype - Attack Type |
| 1. success -Successful Attack | 1. weapsubtype1- Weapon Sub-type |
| 1. weaptype1 - Weapon Type | 1. weapdetail- Weapon Details |
| 1. targtype1 - Target/Victim Type | 1. targsubtype1 - Target Sub-Type |
| 1. killnumb - No. of. Deaths | 1. summary - Incident Summary |