

GLOBAL TERRORISM ANALYSIS

Report and Data Summary

Team Member's Name, Email and Contribution:

1) Pritam Gangwar :-

- Data sorting.
- Data analysis.
- Approach towards plan.
- Frame work of project.
- Graphical representation.
- Bar plot and Heat map.
- Data visualization.

2) Chaitnya Inamdar :-

- Model presentation.
- Model selection and implementation.
- Cross validation on model.
- Implementation of grid parameters.
- Sorting of values.
- Approach towards multiline graph.

3) Jagdish Tawar :-

- Pi-plot and Heat map.
 - Various model implement.
 - Histogram plot.
 - Suggestion of grid parameters.
 - Project summery template.
 - Analyzing results of model.
-

Introduction

1. The Global Terrorism Database

The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2017. The GTD includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks. The database is maintained by researchers at the National Consortium for the Study of Terrorism and Responses to Terrorism (START), headquartered at the University of Maryland.

2. Characteristics of the Database

Contains information on over 180,000 terrorist attacks

Currently the most comprehensive unclassified database on terrorist attacks in the world

Includes information on more than 88,000 bombings, 19,000 assassinations, and 11,000 kidnappings 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 1998 to 2017 alone

3. Project goals

Read the source and do some quick research to understand more about the dataset and its topic

4. Problem Statement-

Analysing the given Dataset and to extract important data from it to provide required information. Different data frames related to this analysis are –

global_terrorism_df – It contains all the information regarding Total attacks, Countries, Regions of different world, Casualties because of attack, Terrorist Groups from all over world from the year 1970 to 2017.

- country_attacks – It contains the information regarding the Countries in which attack happened.*
- Region_df – It contains different world wide Regions in which Terrorist activities happened. E.g. Eastern Europe, Middle East etc.*
- Attack_type_df – It gives information regarding the type of attack done by Terrorists.*
- Group_df - It gives information regarding different Groups of Terrorists which do attacks globally. E.g.- Taliban, ISIS etc.*
- city_df – It gives information regarding the different cities from countries where the most attacks and Casualties happened.*

Conclusion –

Terrorist attacks are among the causes of national instability. A clear understanding of how this event is occurring will help us to conduct more in-depth investigations. The focus of future research will be on performing a quantitative analysis of the countries in each region to conduct further research. Other future work to be done is the use of Big Data techniques for sentiment analysis, which will extract information from social networks to determine possible threats of cyber terrorism. Thus the investigation would use large volumes of data .

Through this research, it is possible to conclude that the use of Exploratory Data Analysis techniques was able to visualize and predict terrorist attacks. The results section shows that there has been a considerable growth in terrorist attacks since 2010 and that due to the classification models, it was possible to determine the probability of which region and type of attack may occur. Concerning the number of attacks by region, it was obtained that there is a probability that they will happen in the Middle East & North Africa and followed by South Asia.

The GitHub Repo link :-

<https://github.com/pritamgangwar/EDA-capstone-project---1>
