Global Terrorism Analysis

DHARANI NALABOLU

OBJECTIVE

The objective of this project was to gain an understanding of terrorism using Bayesian inference. This analysis specifically looked at the Iraq as the terrorism rate was higher for Iraq over the years.

INTRODUCTION TO GLOBAL TERRORISM DATABASE

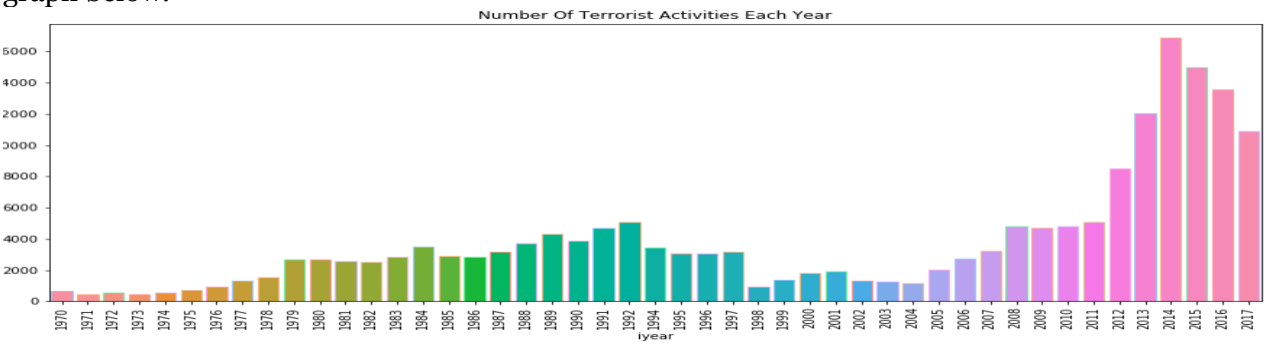
The Global Terrorism Database contains over 170,000 cases of terrorism from around the globe. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) maintains this database. START is headquartered at the University of Maryland. At the time of the submission of this project, 1970 to 2016 was accounted for. The database contains over one hundred and thirty variables including geographic location, type of attack, perpetrators, targets, outcomes, number of fatalities, and motivation of perpetrators. This was a very robust dataset, but it does contain a fair amount of missing information.

METHODS USED :

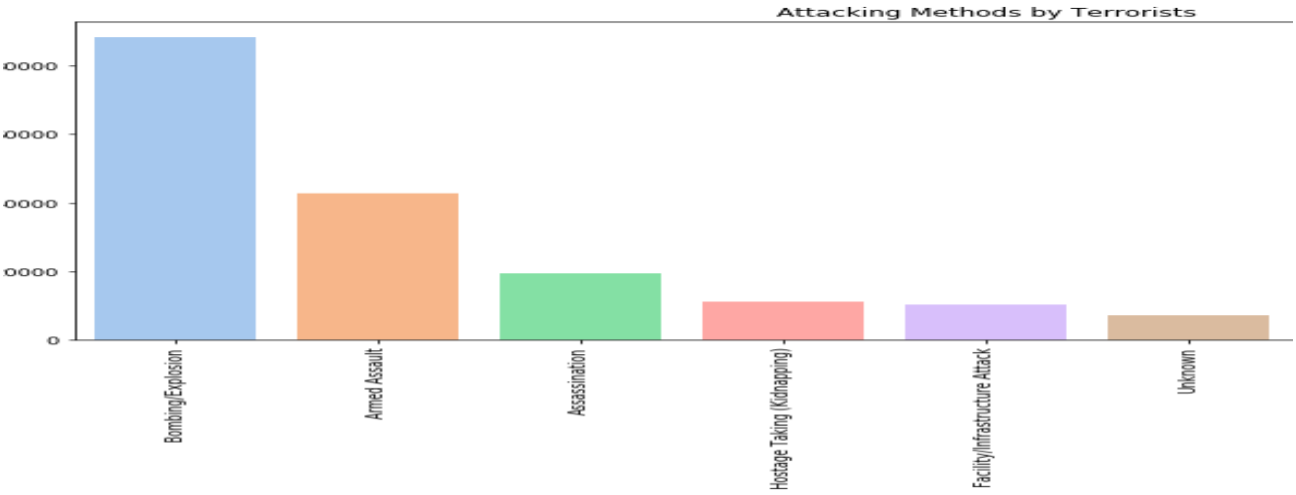
- Natural Language processing
- Bayesian Analysis
- Feature Importance

EXPLORATORY DATA ANALYSIS

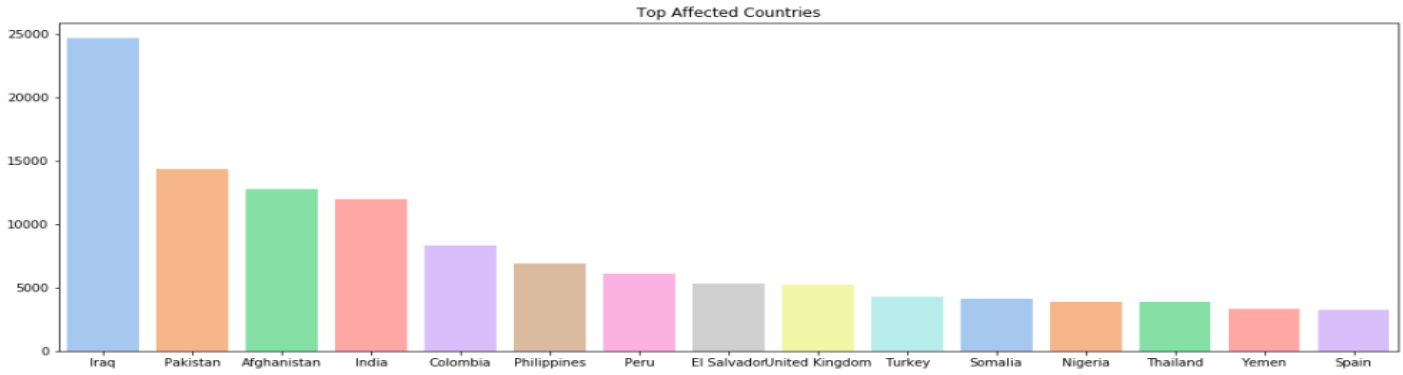
The majority of the data is analyzed through visualizations in python. We have interpreted the global terrorism worldwide from 1970 to 2016 and noticed that there is a significant increase since 2010. It is depicted in the graph below.



Then the we further went into the analysis of the weapons they used and the countries that were affected the most .



Bombing and explosion was the most used technique as the number of people killed through bombings are high compared to others.



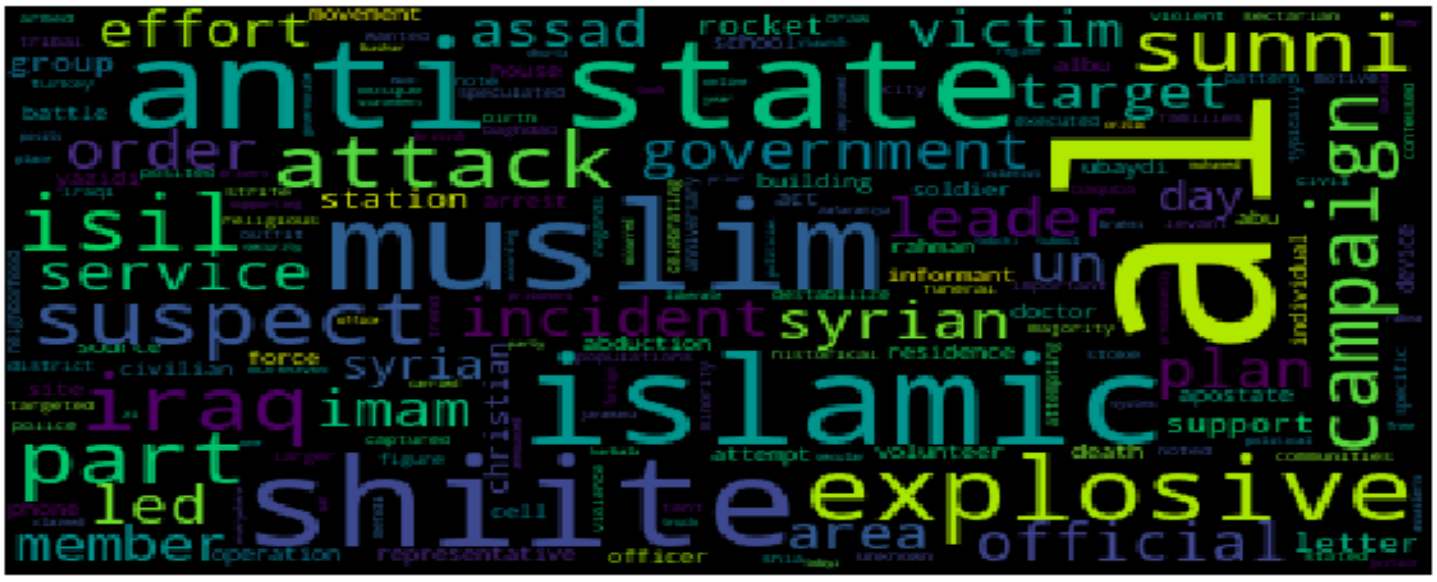
Iraq is the top affected country from 1970 to 2016 irrespective of increase / decrease of terrorism so to understand the motives and terrorism in Iraq we have moved forward with **natural language processing** and **feature importance** to find the importance features .

Natural Language Processing :

The analysis initially concentrated on finding the motives of most prevalent terrorism group with most number of attacks which is ISIL in Iraq. The following are the motives :

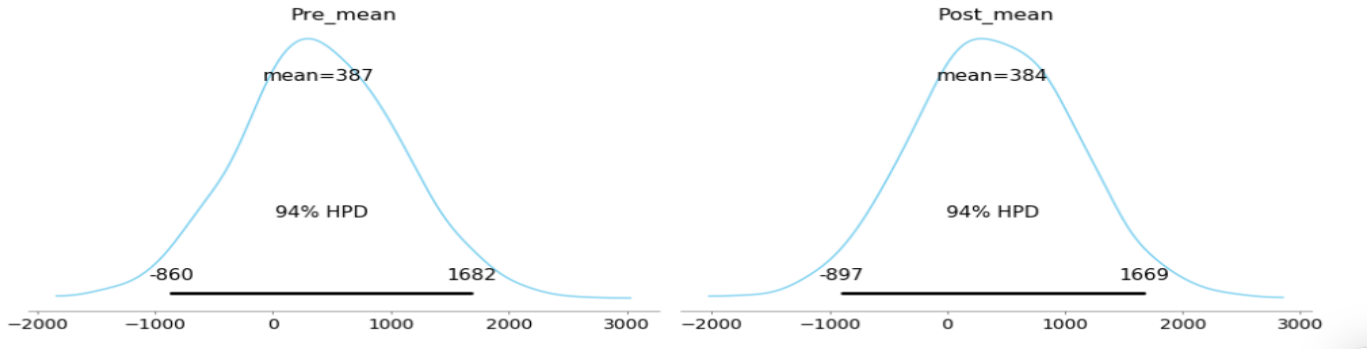
```
The specific motive is unknown; however, sources noted that the victims were attempting to leave the area at the time
of the attack.
25
According to an online statement, Islamic State of Iraq and the Levant (ISIL) claimed responsibility for the inciden
t, stating that the attack was in retaliation for the Iraqi government's alleged "crimes" against the minority Sunni
community. Moreover, the armed Islamist outfit claims that the incident was part of its newly launched offensive camp
aign called "Harvesting the Soldiers".    23
In a statement posted online, Islamic State of Iraq and the Levant (ISIL) claimed responsibility and stated that the
attacks were in response to Iraqi security forces arresting hundreds of Muslims.
23
The specific motive is unknown; however, sources stated that the victims were accused of working with security person
nel.
20
```

A **wordcloud** has been created out of the motive to find the most used by the terrorist groups while sending the message. The wordcloud as follows:



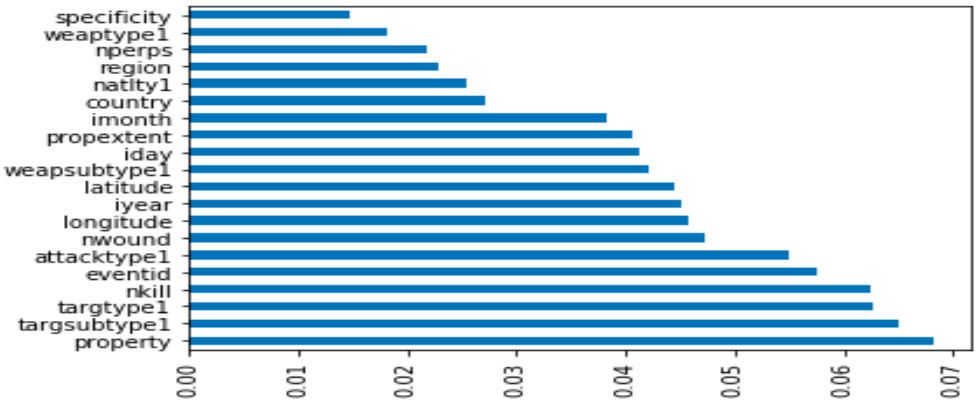
Bayesian Inference

The goal of the Bayesian analysis was to determine if there was a statistical between the number of terrorist bombings in the Iraq from pre-2010 and post-2010. From a global view, bombings increased drastically since 2010. The prior probability of terror attacks in the Iraq was a simple average the number of bombings between 1970 and 2009. This turned out to be around 387 bombings per year. The average of bombings per year between 1970 and 2009 was a logical prior assumption for Bayesian statistics because this would have been the only information available at the onset of 2009. If we were making the analysis in 2010, we would want to see whether or not one year of increased terror attacks was statistically different than the previous years. It can be seen that the mean pre and post -2010 did not change slightly which makes Iraq being the most attacked country .



FEATURE IMPORTANCE :

Feature Importance provides a score that indicates how useful or valuable each **feature** was in the construction of the boosted decision trees within the model. The more an attribute is used to make key decisions with decision trees, the higher its relative importance. The importance are listed below . These features will help us in understanding the target variable in a good way . The relative importance is selected according to the relative importance it is to the target variable.



CONCLUSION & RECOMMENDATIONS:

- Terrorism in Iraq has been the highest throughout the years.
- Addressing the motives of ISIL in Iraq will decrease terrorism in Iraq.
- Bayesian inference can be applied in understanding and comparing terrorism for various countries, continents etc.
- Adfuller test to find the missing data in 1993.