EXPLORATORY DATA ANALYSIS(EDA)

PROJECT ON

GLOBAL TERRORISM ANALYSIS

Name of the participants:

CHARAN

Data science trainees,

ALMABETER,BANGLORE

Abstract:

The entire world in today's time is highly undergoing through terrorism. This topic not only includes the incessant obstacles thrusting in the path of global progress but also a deep-down effect on development of every county.

Here we have an impactful dataset drawn from Global Terrorism Database which comprises of data of all the terrorist attacks taken place around the world from 1970 to 2017. It includes more than 180000 attacks and a big range of massacre. It has distressed the political and economic stabilities on worldwide level.

Here our study can help understand what are the basic factors riding behind this growing problem. And some core points are milestones to take a proper action to reduce this problem at some extent.

1.Problem Statement

Data provided here is big source of record of continuous terrorist activities. Here we have to provide a helpful data analysis for having multilateral initiatives in response to terrorism. The followingly depicted points are some major questions which we need to answer are:

1. The most active terrorist group
2. The most used weapon set
3. The most attacked country
4. The most attacked regions
5. The correlation of years and terrorist activities
6. The estimation of people and property loss in each country
7. The calculation of people wounded

And therefore, we need to present a solid data analysis which help people in predicting which country needs more security, which country is safer to live and which kind of weapons and strategies the terrorists use, so that people can be alert.

2.Introduction

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. Terrorist attacks typically involve high lethality and destructive power and directly cause massive casualties and property losses. In addition, they bring tremendous psychological pressure on people. The analysis and prediction of terrorist attacks support targeted attacks on terrorist groups and provide valuable information for antiterrorism and terrorism prevention operations, enabling authorities to find new or hidden terrorists as soon as possible to reduce human and property losses, prevent problems and improve the security and stability of social life. The patterns of attacks planned and carried out by terrorists may seem random on the surface, but in fact, they are typically organized and premeditated actions chosen carefully and deliberately. Moreover, attacks by the same organizations and individuals tend to be substantially related in terms of certain distinguishable characteristics. Therefore, there must be some patterns or informal rules guiding the activities of terrorist organizations. After analysing these characteristic patterns of activity by terrorist organizations, authorities can make more detailed predictions and analyses of terrorist organizations to attack them more accurately and increase the time available for the prevention and prediction of terrorist attacks.

- **Pandas**

We have used pandas for data analysis and manipulation and of better data representation.

- **Matplotlib**

We have used matplotlib for data visualization, graphs & plotting.

- **Seaborn**

We have added seaborn to make data visualisation more colourful and meaningful.

- **Folium**

Folium has been used for drawing world map to locate any desired location within no time.

- **Ipywidgets**

An interactive amazing library 'Ipywidgets' has been used to cater more dynamic representation of our graphs and plots.

4. EXPLORATORY DATA ANALYS

1. DATA LOADING

We have loaded the data into our drive and then finally loaded it to google collab notebook using Pandas library.

Totally we have a huge number of 181691.

2. CLEANING (Null value treatment)

The major part of the dataset is null values which triggers many stumbling blockages in processing the necessary columns, therefore its required to identify the columns containing lumpsum null values and then drop them.

Let's remove the columns which have 'NAN' i.e., no values Remember, it's something else than '0'!!

3. DELETING UNWANTED COLUMNS

In this step we will observe the columns of the dataset minutely and eliminate those columns which don't have much to do with our topic. It will help to smoothen our analysis.

We are almost near to get our final dataset!!

And the columns which may hamper your accuracy are:

extended,'region'specifity', 'vicinty', 'crit1', 'crit2', 'crit3',

'doubtterr', 'multiple','weaptype1','attacktype1','targtype1','targsubtype1','natlty1', 'natlty1_txt','guncertain1', 'individual','weaptype1','property', 'ishostkid', 'dbsource','INT_LOG','INT_IDEO', 'INT_MISC', 'INT_ANY'

So finally, we have the following columns remained with us.

The number of columns is 20!!! How sweet ☺ ☺.

eventid', 'iyear', 'imonth', 'iday', 'country_txt', 'region_txt',

'provstate', 'city', 'latitude', 'longitude', 'success', 'suicide',

'attacktype1_txt', 'targtype1_txt', 'targsubtype1_txt', 'target1',

'gname', 'weaptype1_txt', 'nkill', 'nwound'

4. RENAMING THE COLUMNSs

Now finally we have set all the columns .And now its time to rename the column to make the study easier.

'eventid':'event_id',

'iyear':'year',

'imonth':'month',

'iday':'day',

'country_txt':'country',

'region_txt':'region',

'provstate':'state',

'attacktype1_txt':'attack_type',

'targtype1_txt':'targ_type',

'targsubtype1_txt':'targsub_type',

'target1':'main_target',

'gname':'gang_name',

'weaptype1_txt':'weapon_type',

'nkill':'number_of_kill',

'nwound':'number_of_wound'

*5. ANALYSIS OF **MOST** TYPE VARIABLES.*

1. Country
2. Region
3. City
4. Year
5. Month
6. Group
7. Attack types (Strategies / Weapon sets used for attack)

- Country with the most attacks: Iraq
- City with the most attacks: Bhagdad
- Region with the most attacks: Middle East & North Africa
- Year with the most attacks: 2014
- Month with the most attacks: 5
- **Group with the most attacks: Taliban**
- **Most Attack Types: Bombing/Explosion.**

*And here you go !! what can shake your head and make you say just 'wow!' that what you have predicted has definitely foregone in trends. Very recently we came across with overwhelming actions of the **TALIBAN** who had made many **BOMBING EXPLOSIONS** in **KABUL** and overpowered **AFGANISTAN.***

6. DATA VISUALISATION:

6.1 PEOPLE KILLED EVERY YEAR:

Here we have incurved that from year 1970 to 2017 how rapidly graph of peoples' deaths has slayed the surface and in the year 2014 it reached at the topmost.

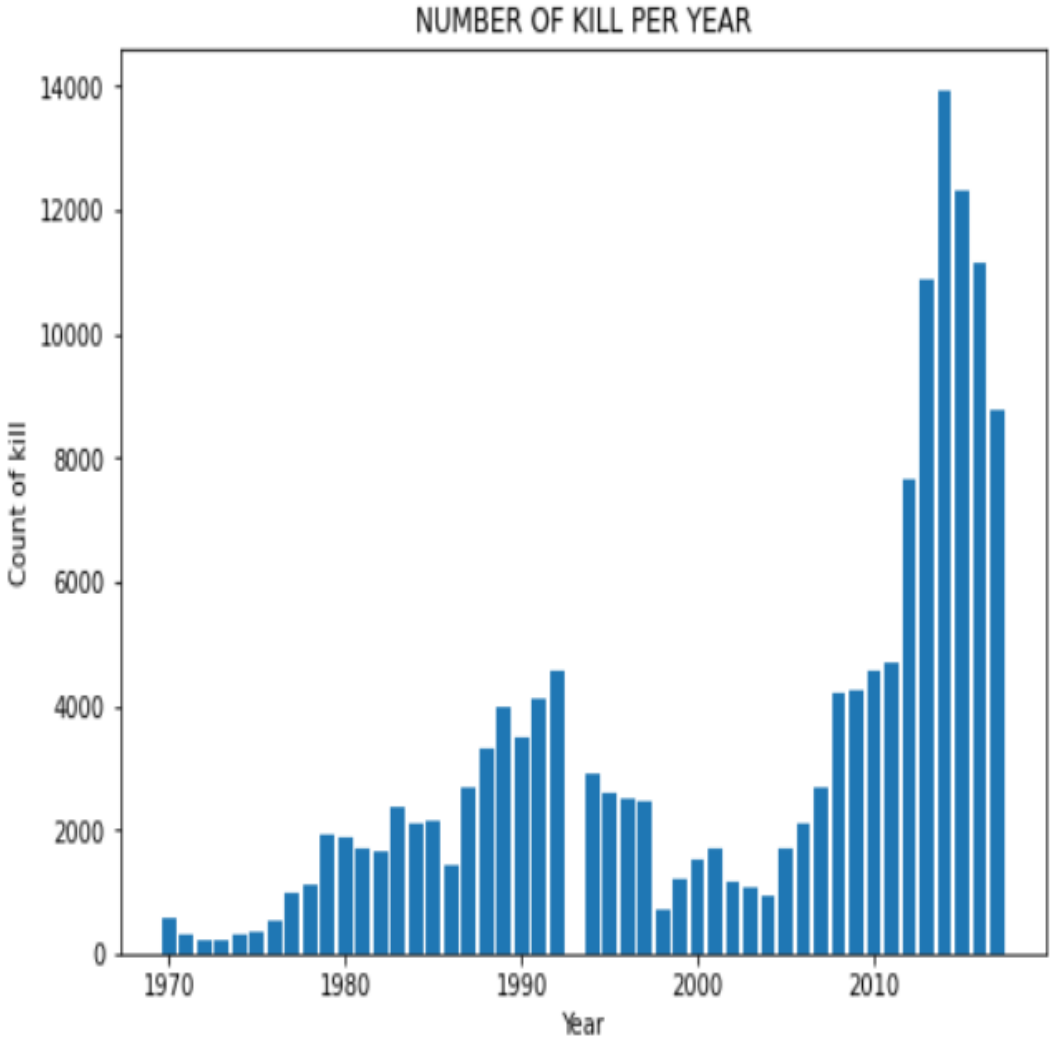
The maxima were more than 14000 a big number!!!

From above bar chart, we came to know about the total number of terrorist attacks with respect to different years. We can conclude that most attacks were done in the year 2014.

We can see here that attacks slowly started to increase from 1972 until about 1993 where by 2004 we can see a downward trend in the terrorist incidents. Then after 2014 the terrorist activities started increasing till 2014 then slowly decreased.

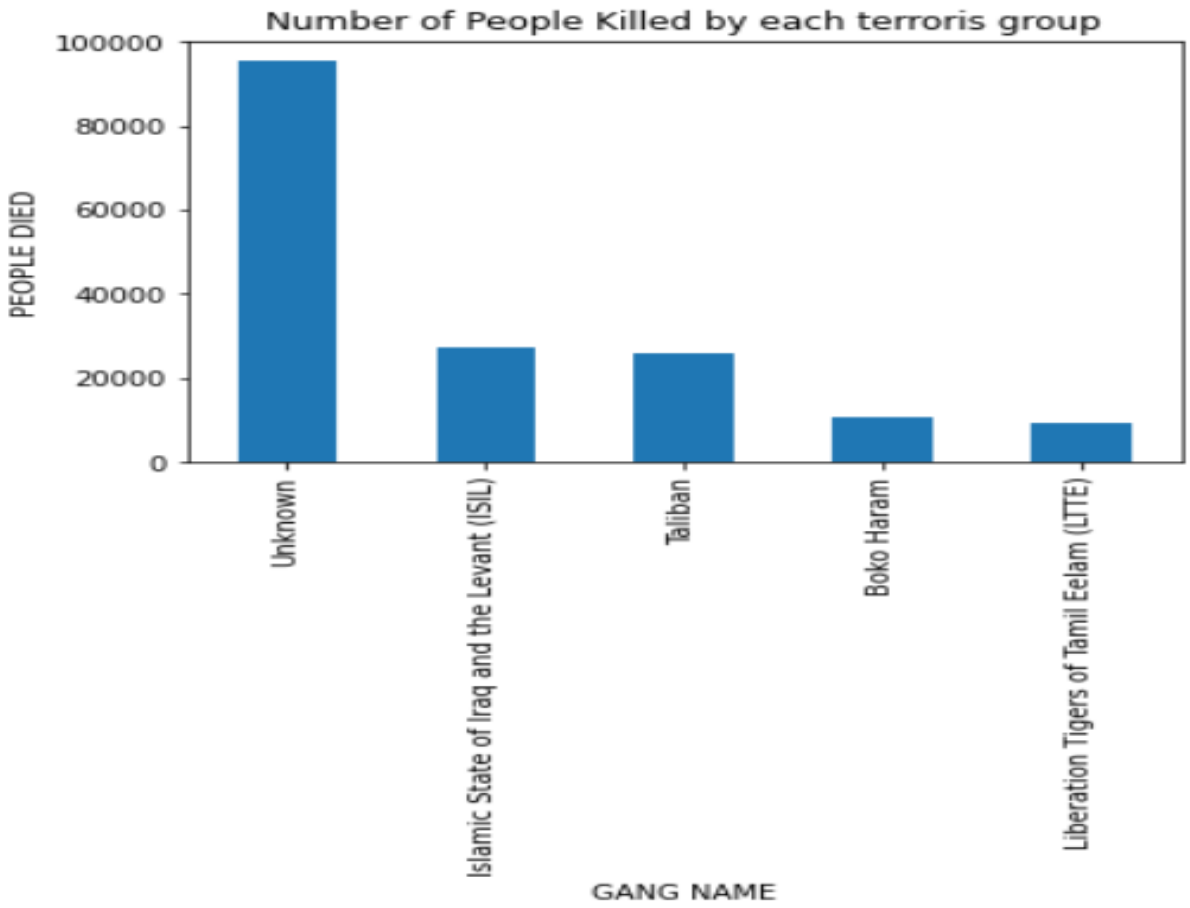
The lines striding upward show that how rapidly the attacks grew between 2010 to 2014.

And then submissively they decreased.



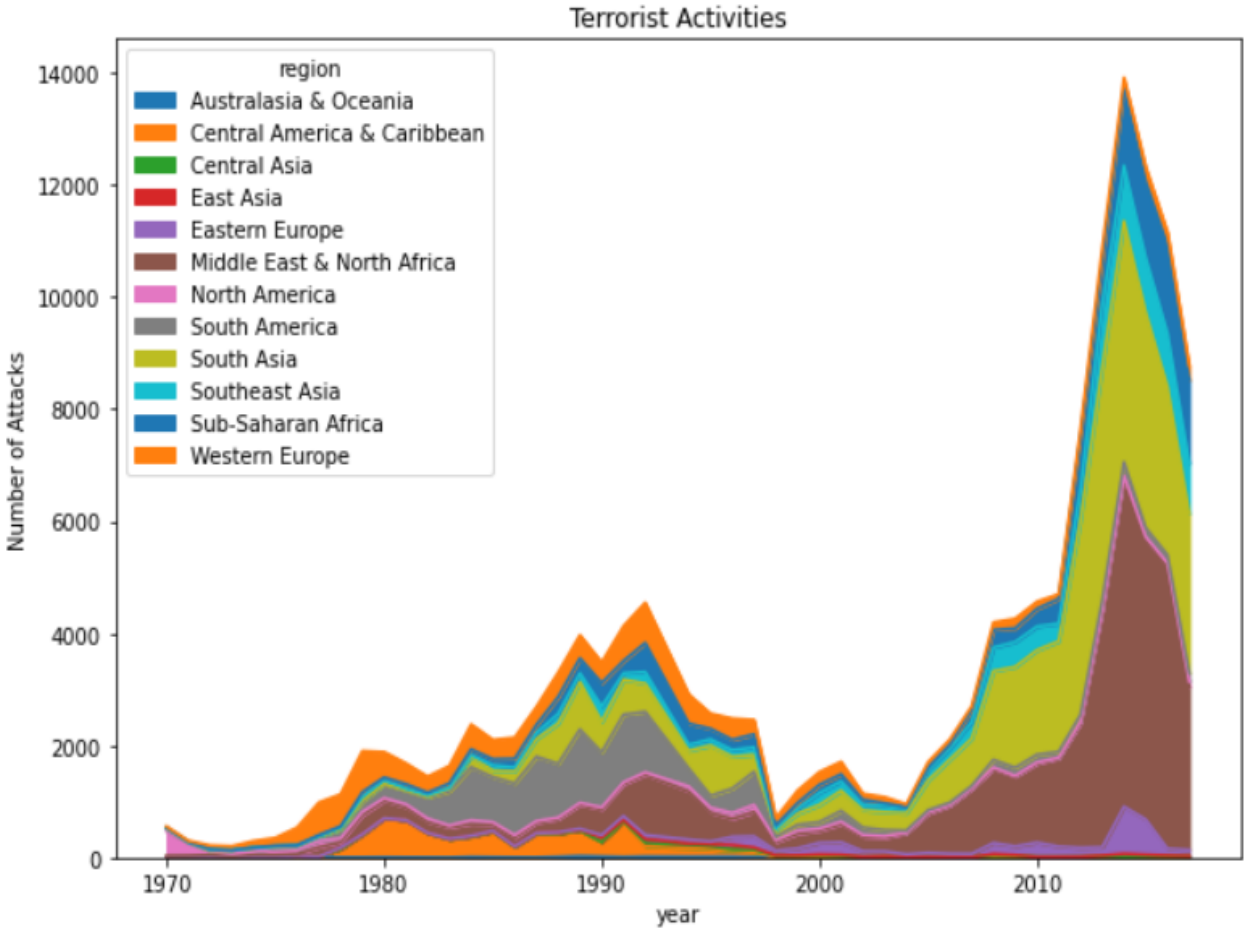
6.2. NUMBER OF PEOPLE KILLED BY EACH TERRORIST GROUP

The increasing graph of peoples killed was reaching high as a resultant of terrorist attacks.
As we analysed the data, we figured out that from above bar graph, which shows the number of people killed by different terrorist groups, we can say that most people were killed by terrorists which is unknown followed by ISIL



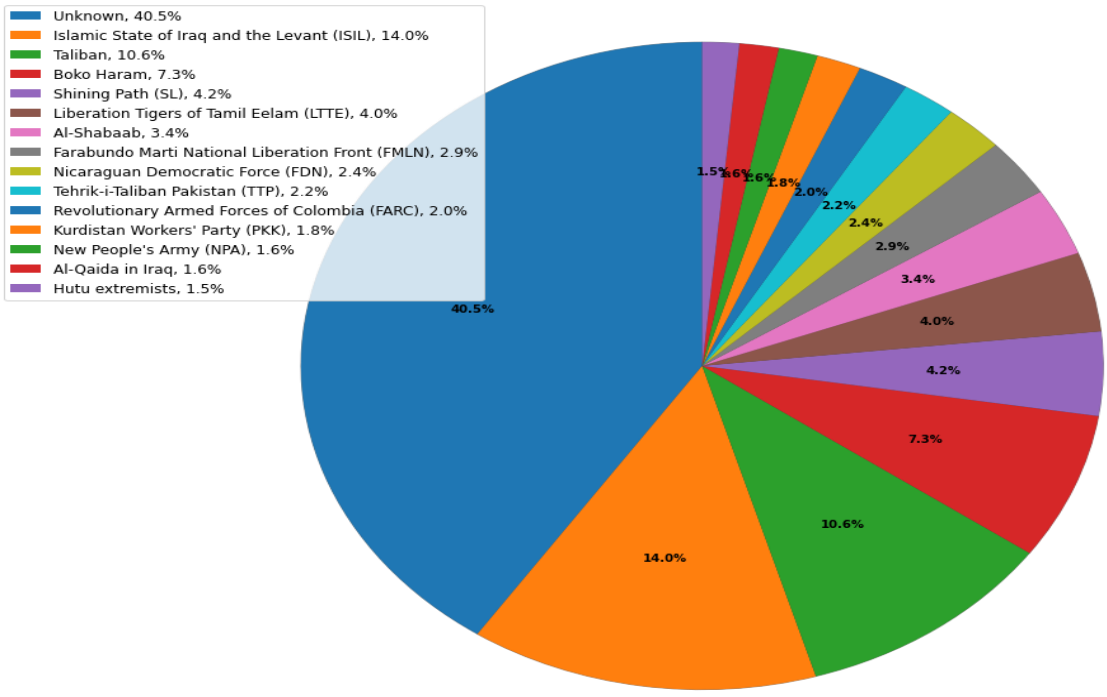
6.3. Terrorist Activities (Number of Attacks vs Year)

The below graph is plotted using an area type of graph. It shows the total number of attacks in corresponding years and regions. Most attacks were made in 2014 on Australasia & Oceania.
In the 1970s we can see that most terrorist attacks occurred in Western Europe, followed by Latin and South America in the 1980s.But then there is an overall decline in the number of attacks between the mid-1990s and 2003.



6.4 Percentage of attacks by each gang.

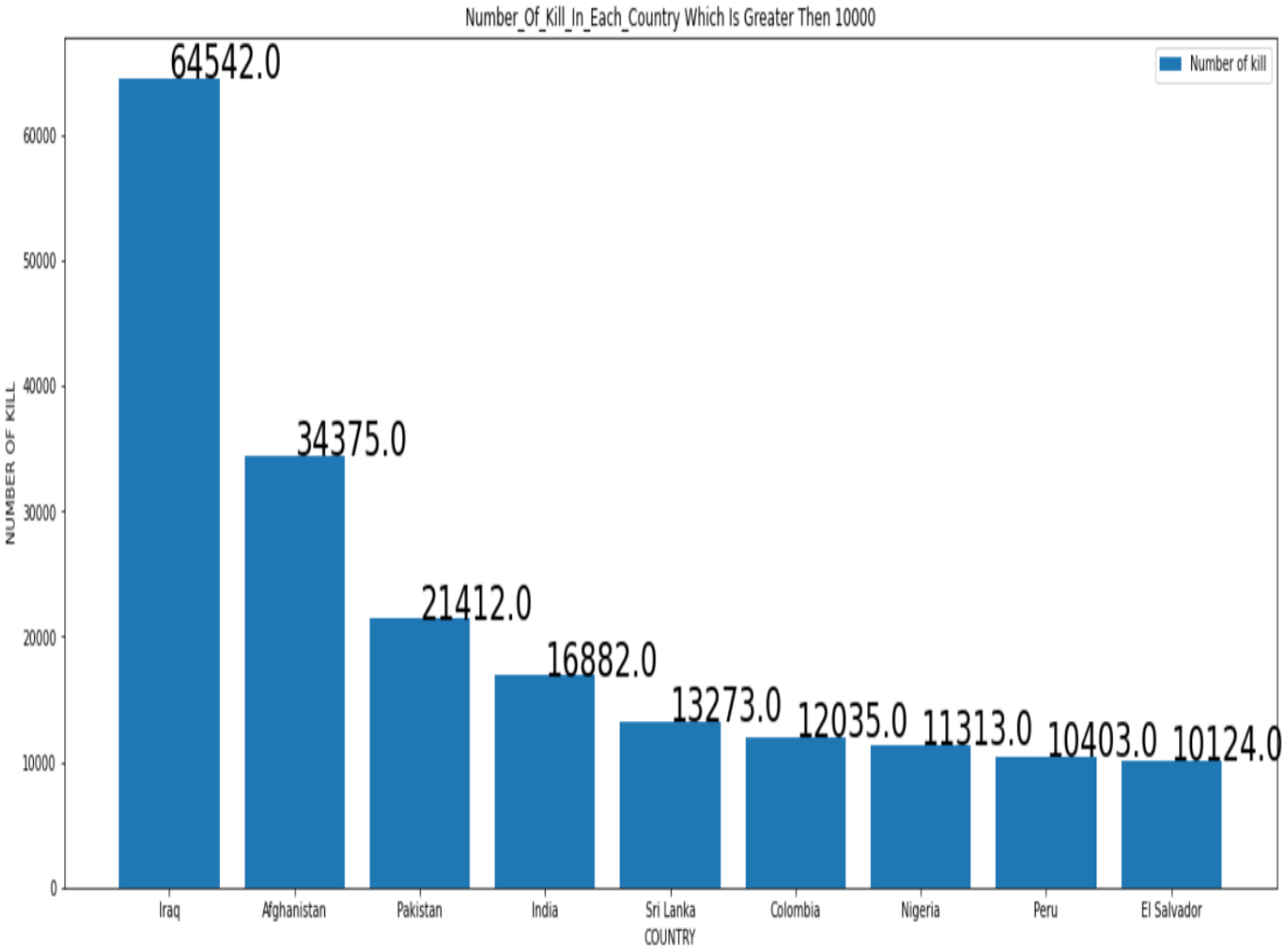
The following figure is plotted using a pie chart which shows the percentage of attacks made by different terrorist groups out of total attacks.



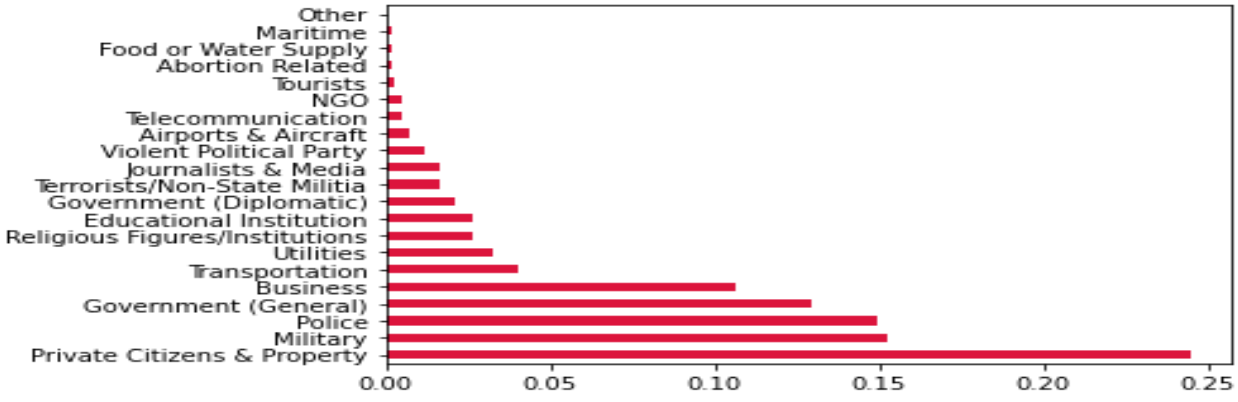
6.5 Maximum Number of kill in each country which is greater than 10000

This bar graph shows us the total number of people killed in the attacks in different countries…

And here we have taken top 10 countries which has death rate above 10000.

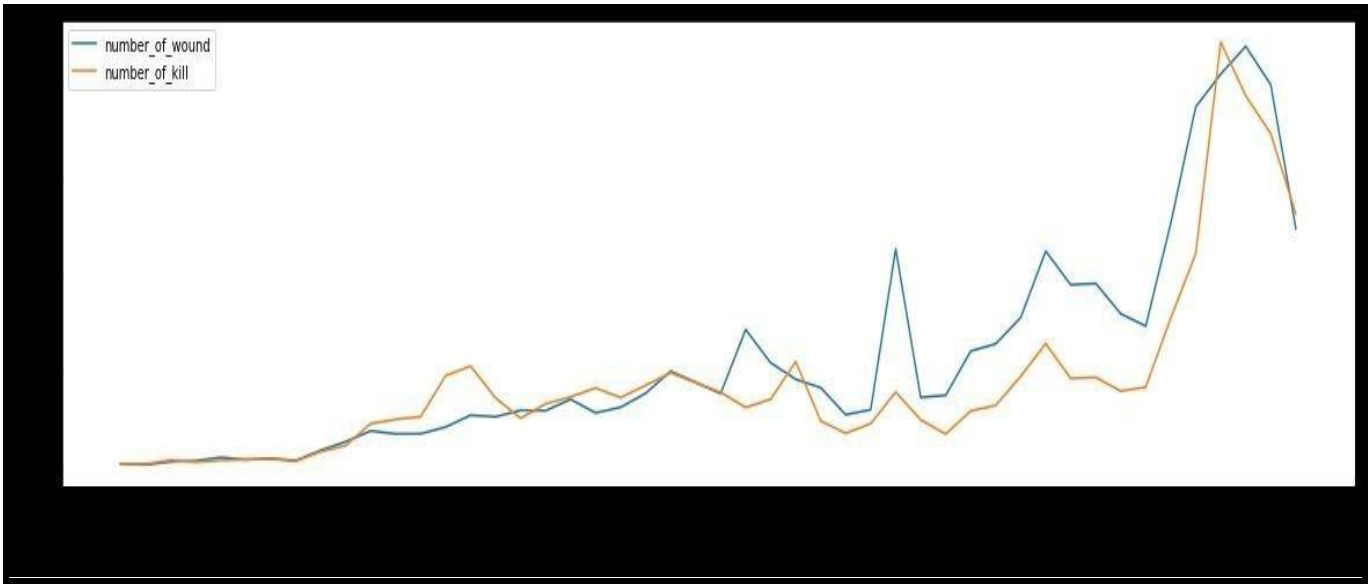


6.6 Target values



From above graph, which depicts that most of the terrorist gangs targeted the private citizens and property followed by military, police.

6.7 Number of kill vs Number of wounded

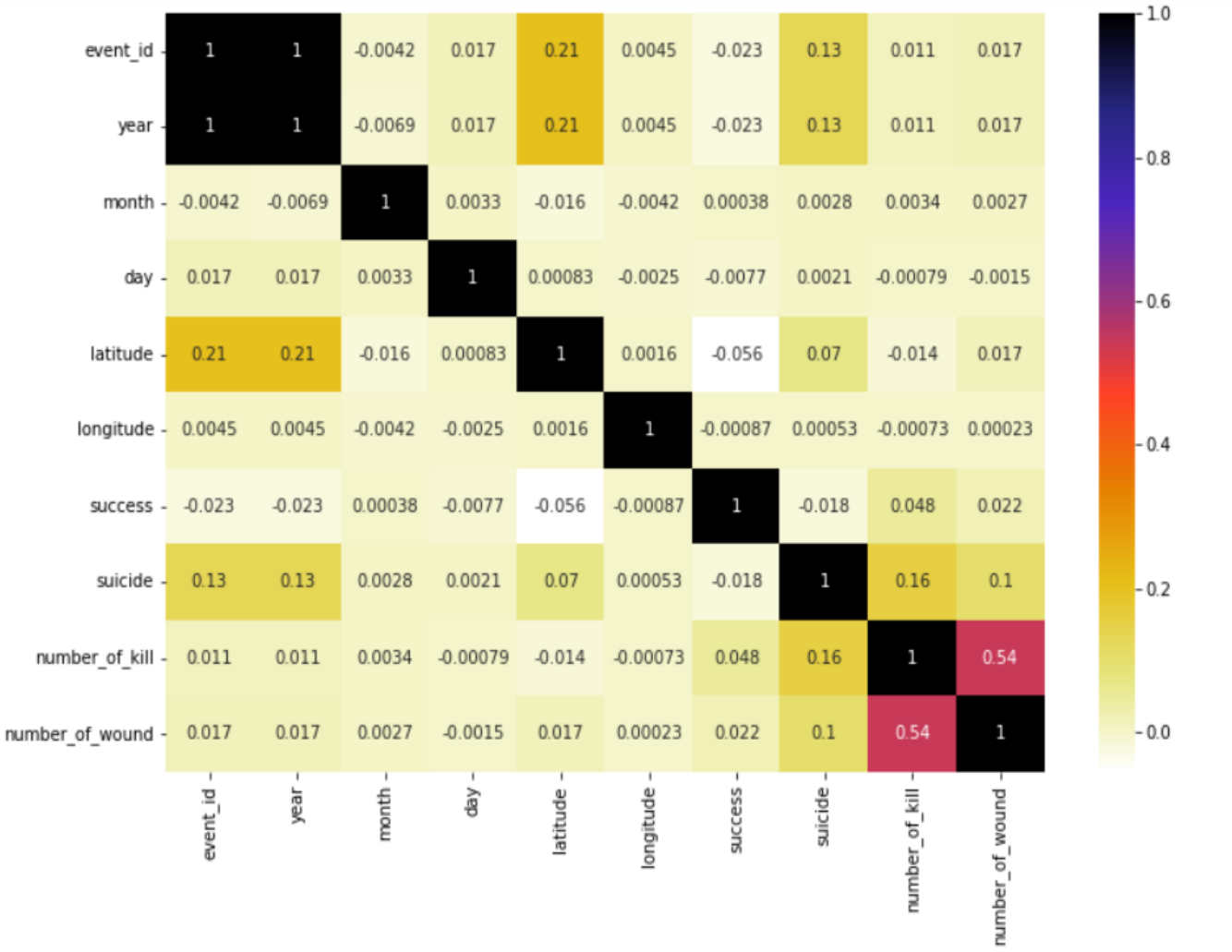


The above graph shows the graph of deaths and people wounded. With every year the rate of attacks increased and it resulted in increasing death rate and wounded peoples. In the beginning the kill rate was lower than wound rate but in the ending years the death rate rapidly made a big arch marking a big range of death rate in proportion to wound rate.

6.8 Correlation

The below heatmap of correlation between the independent variables. The matrix shows the coefficients in the squared form coloured as per the intensity scale. The map has positive covariance as with increase in one variable another is also increasing.

As per the increase in years suicides are also increasing. But with each increase of years number of deaths are intensively growing.



Conclusion

From the analysis, we came to know that attacks slowly started increasing from 1972 until 1993.Then the rise took a great leap from 2011 to 2014 and then slowly decreased. The most attacks were done in May 2014. The country with the most attacks is Iraq. In the 1970s we can see that most terrorist attacks occurred in Western Europe followed by Latin and South America in the 1980s. Most of the terrorist gangs targeted the private citizens and property followed by the military, police and government. In most attacks the weapons used by terrorists were bombing/Explosion followed and armed assault. Maximum terrorist attacks are unknown. These unknown attacks made a huge loss and this is the only reason the attacks consistently kept their track and caused a great human and property loss as they were not recognized. The terrorist group causing most attacks is ISIL as per the records then Taliban.

Our country India was also badly affected by this massacre. Total 10491 people were killed. Whereas Pakistan stands on the second number in the global human loss. 12756 people were killed in these attacks in Pakistan. Whereas in Afghanistan 10871 people were killed.

Overall, the middle east area was broadly affected and was looted. Whereas the UK was the safest country.

To get rid of such a huge problem it's very important to understand key factors threshing towards this problem. As it has been a kind of deadlock which is pushing the developing countries down and making them handicapped from standing in the rows of the other developed countries.