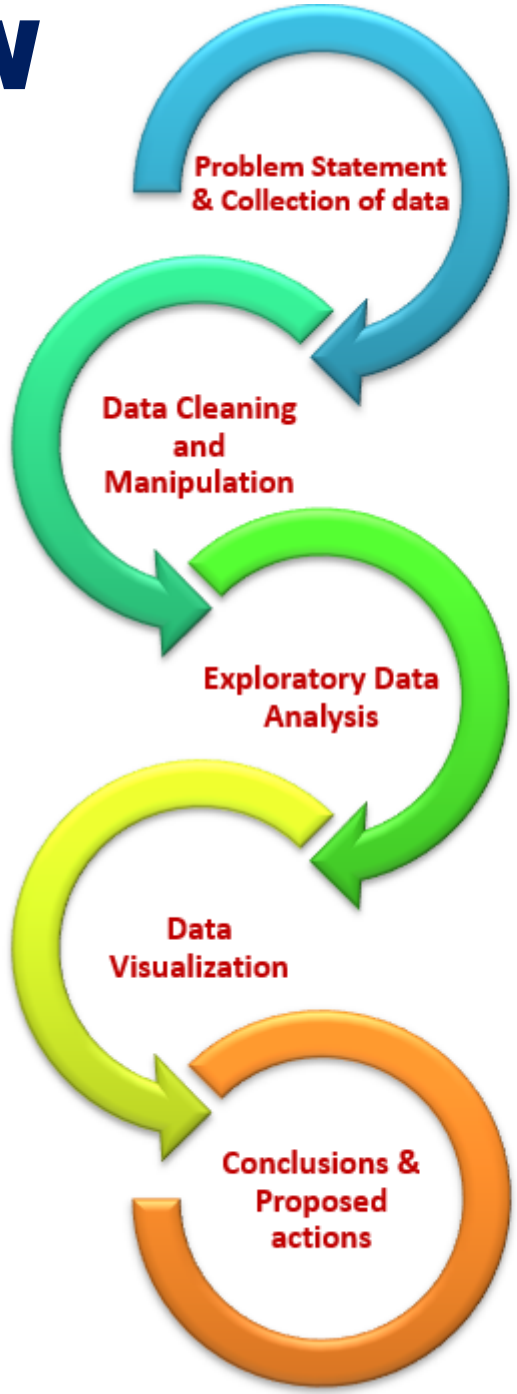
CAPSTONE PROJECT

GLOBAL TERRORISM DATASET (EDA)

INTRODUCTION

We are here to explore a Global terrorism dataset to discover the terrorist attacks across the world. This data set contains Global terrorist attacks information from 1970 to 2017, which includes information such as occurrence of terrorist attack, region, country, state, city, latitude , longitude, Target type, Attack type, Weapon type, Terrorist organization committed, death toll and causalities among other things. We will analyze some important aspects of this Global Terrorism dataset which will help us identify major factors and give us insights to mitigate the terror attacks.

WORK FLOW



- Identify the problem of analysis to be done and then collect the data regarding it
- Raw data will have lot of null values and missing data. So it needs to be cleaned and manipulated as per our analysis
- By Using Python libraries we write code to study and extract results
- The results of our analysis will be displayed in a visualization manner for better readability and understandability
- Based on the EDA results conclusions will be drawn and necessary actions will be proposed to mitigate the risks

For this project we have used this workflow to analyse the Global Terrorism Dataset

DATA INSPECTION & CLEANING

- The Dataset contains 181691 rows and 135 columns.
- There are 106 columns out of 135 columns are having missing values.
- We remove the columns with 85% missing values in the Dataset. After dropping we are left with 47 columns.
- In those 47, We drop some columns which are not useful for our analysis. For instance, country column represent the numerical representation of country, country_txt column represents the name of the country. So, the column ‘country’ is not useful for our analysis. Likewise other columns are as below that we dropped.
  - ‘eventid’, ‘extended’, ’iday’, ‘country’ , ‘targsubtype1’ ,’region’ ,’weaptype1’ , ’weapsubtype1’, specificity, 'vicinity', 'crit1', 'crit2', 'crit3', 'doubtterr’, 'multiple', 'suicide', 'attacktype1', 'targtype1', 'targsubtype1_txt', 'target1’, 'natlty1_txt', 'guncertain1', 'individual', 'weapsubtype1_txt', 'property', 'dbsource', 'INT_LOG', 'INT_IDEO', 'INT_MISC', 'INT_ANY’
- Our Final Dataset will have 181691 rows and 17 columns.

ATTRIBUTE INFORMATION

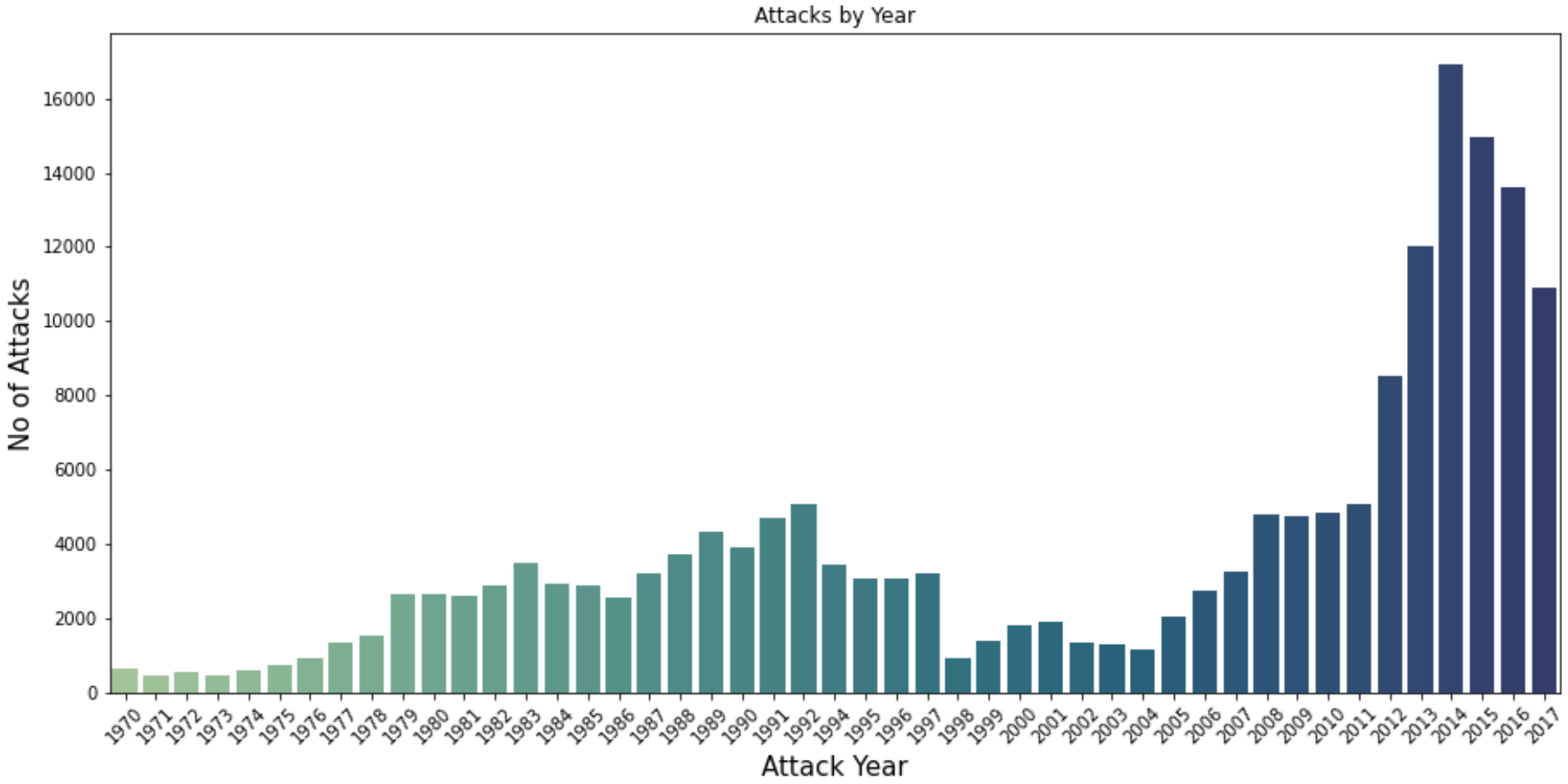
- **iyear :** The Year in which attacks were happened
- **imonth :** The Month in which attacks were happened
- **country_txt :** Name of the country
- **region_txt :** Name of the Region
- **provstate :** Name of the State
- **city:** Name of the City
- **latitude :** Latitude of the terrorist attacked location
- **longitude :** Longitude of the terrorist attacked location
- **attacktype1_txt :** Attack type of the terrorist attack
- **natlty1 :** Indicates victim belongs to which nation
- **success :** (0 or 1) indicates whether attack was failure or success
- **targtype1_txt:** Target type of the terrorist attack
- **gname :** Name of the gang committed the terrorist attack
- **weaptype1_txt:** Type of Weapons used for terrorist attack
- **nkill :** no of people killed during the terrorist attack
- **nwound :** no of people wounded during the terrorist attack
- **Ishostkid:** (0 or 1) Indicates whether hostage is kid or not in the attack

DATA MANIPULATION & HANDLING MISSING VALUES

- If number of kills and Wounds are null, then they got replaced by ‘0’.
- For the missing values in State, City columns, replaced it with ‘Unknown’.
- For the missing values in Target type replaced with ‘Others’.
- For the missing values in ishostkid column, replaced with ‘0’.
- Replaced negative values in ishostkid column with ‘0’.
- Created Victim columns, which is sum of no people killed and wounded.
- Renamed some columns as below for better usability and tracing. They are as follows.
  - 'iyear’: 'Year’, 'imonth’: 'Month’, 'country_txt’: 'Country’, 'region_txt’: 'Region’, 'provstate’: 'State’, 'city’: 'City’, 'attacktype1_txt’: 'Attack_type’, 'targtype1_txt’: 'Target_type’, 'targsubtype1_txt’: 'Target_subtype’, 'gname’: 'Gang_name’, 'weaptype1_txt’: 'Weapon_Type’, 'nkill’: 'Kills’, 'nwound’: 'Wounds'

UNIVARIATE ANALYSIS

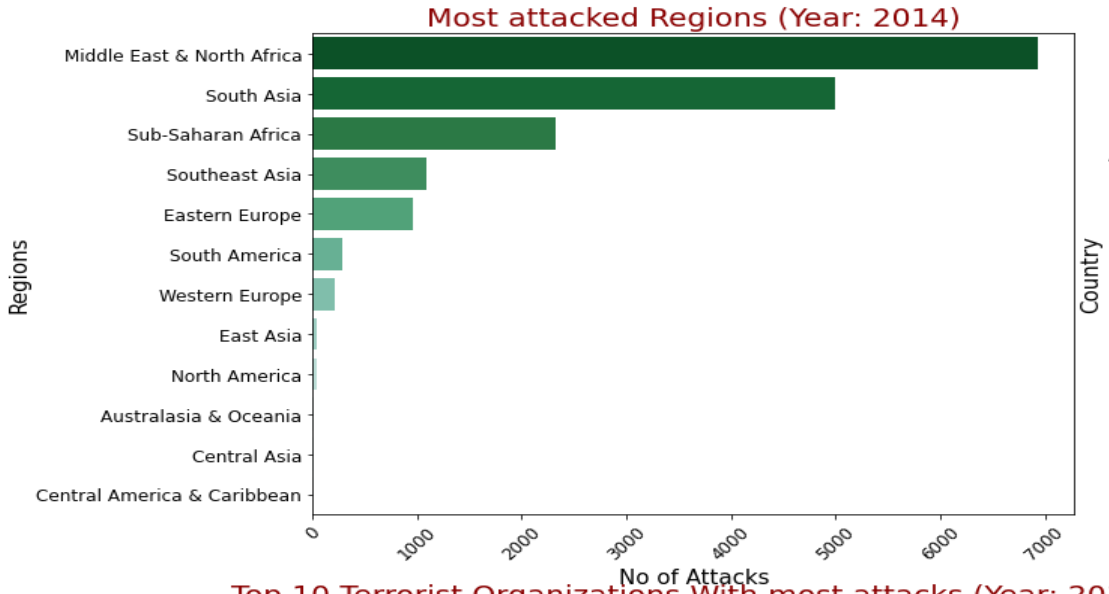
Analysis of Terror attacks year wise:

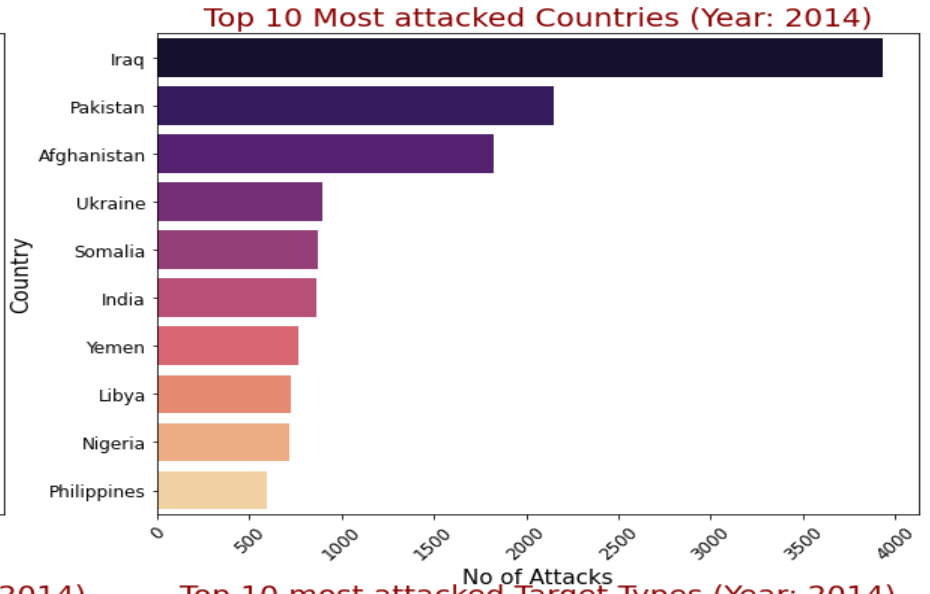


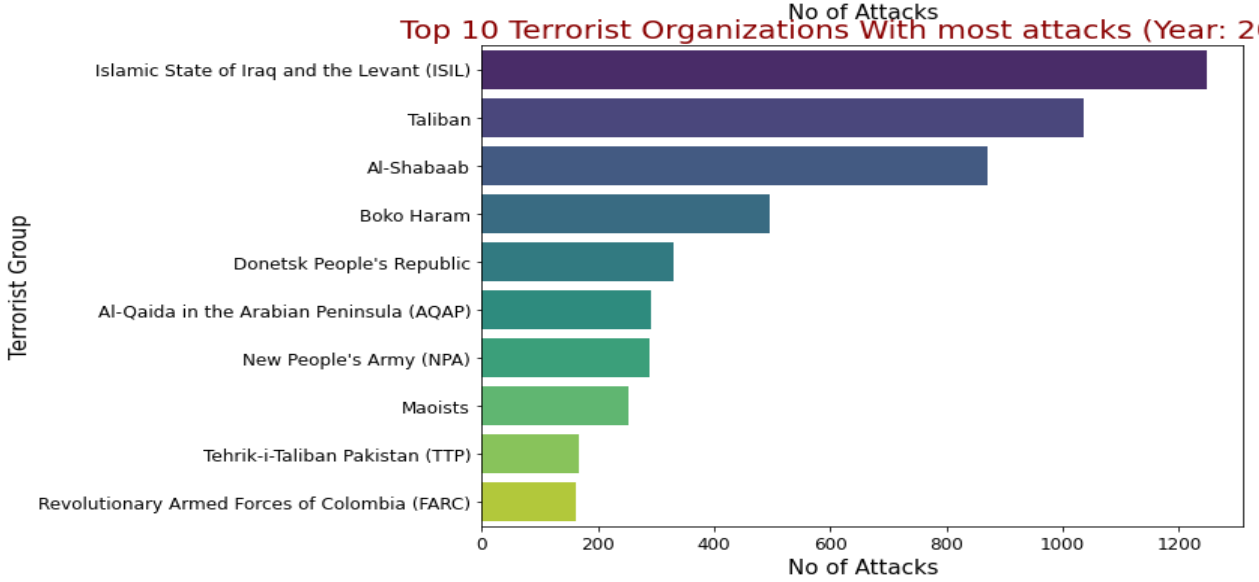
1. Terror attacks were rapidly increasing from the year 2004 to 2014.
2. Last 3 years terror attacks were in decreasing trend

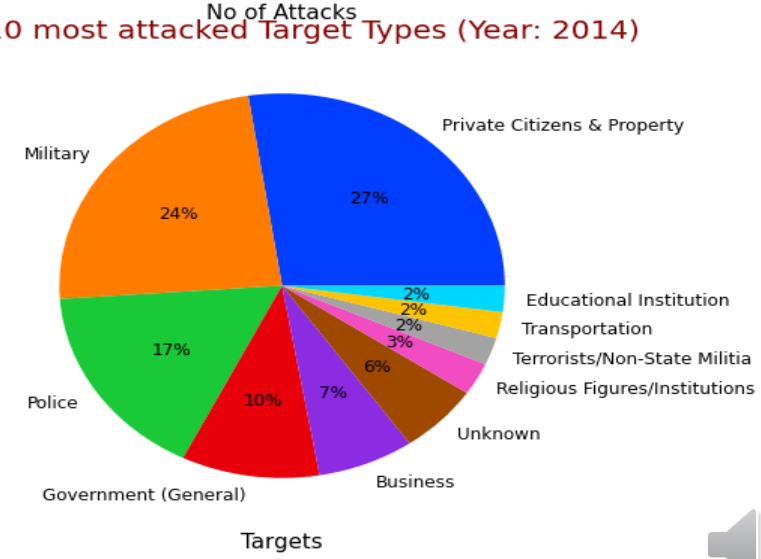
UNIVARIATE ANALYSIS

Analysis of Terror attacks in the Year 2014:



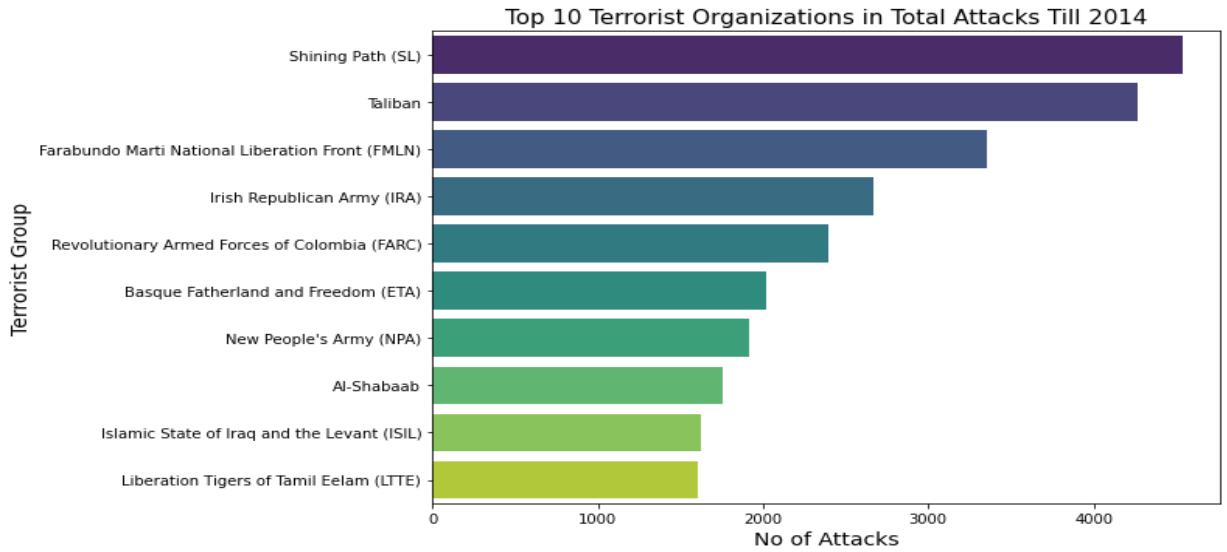


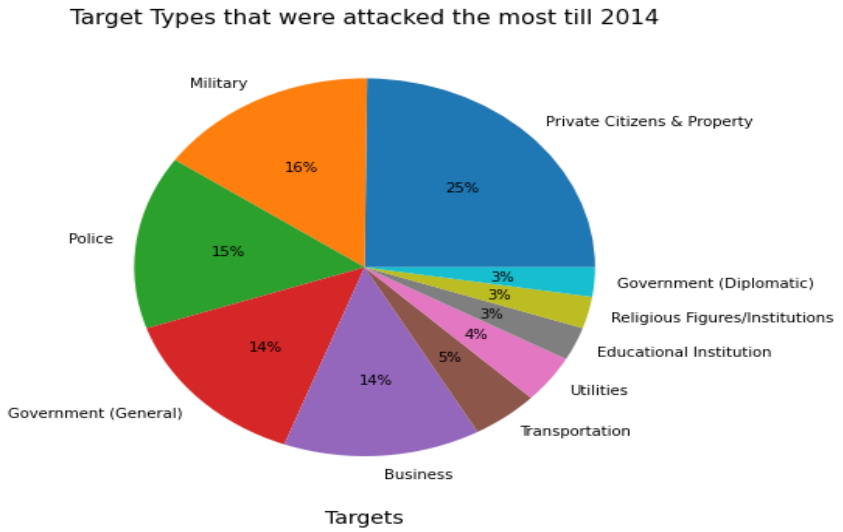


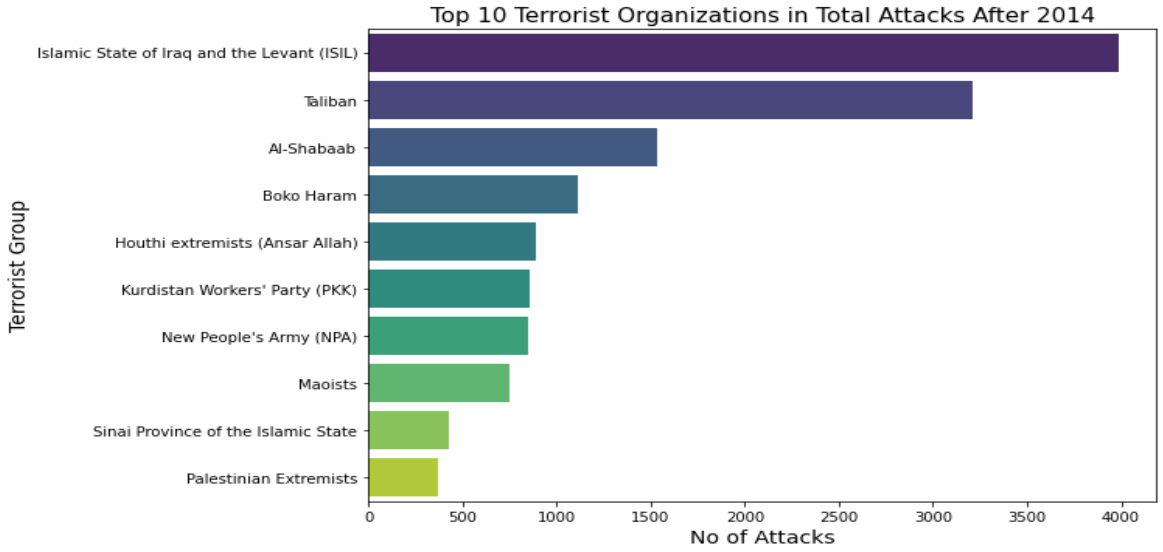


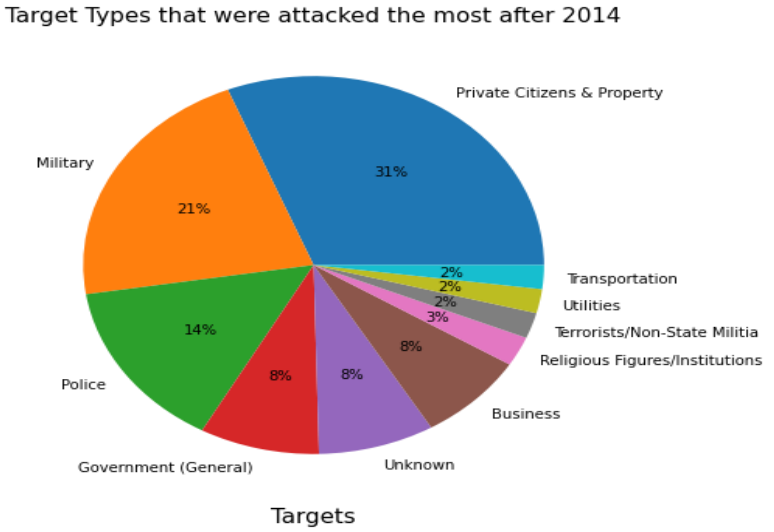
UNIVARIATE ANALYSIS

Analysis of Terrorist Organization & Target type before and after 2014:

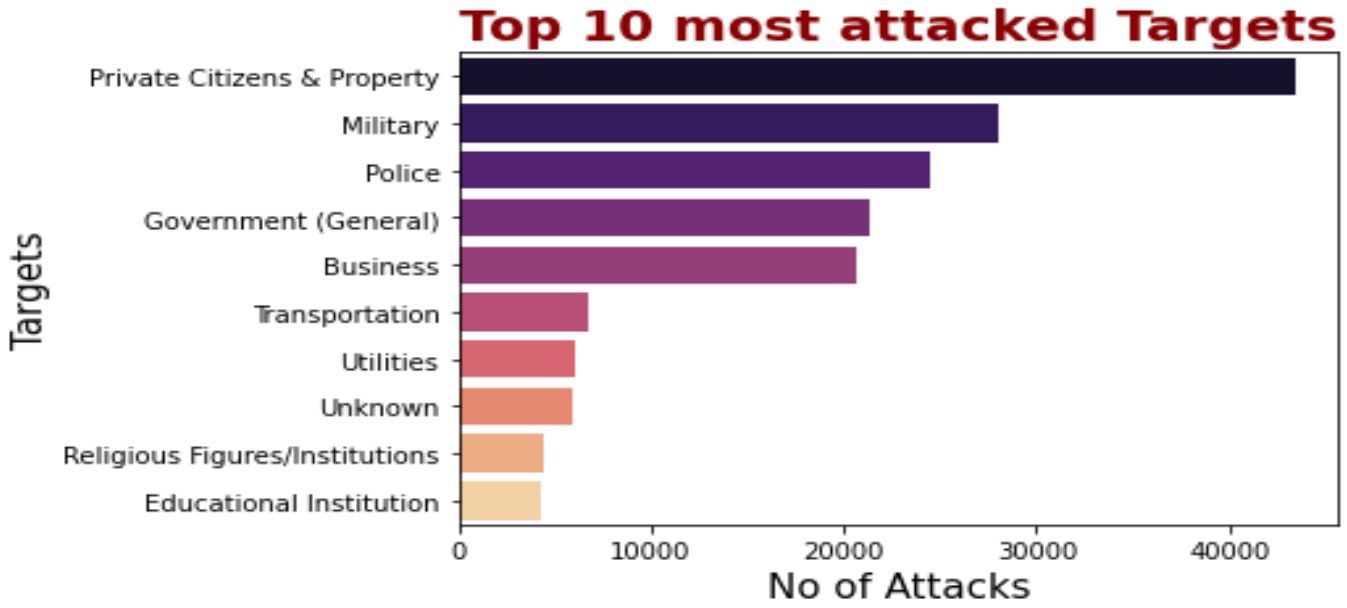


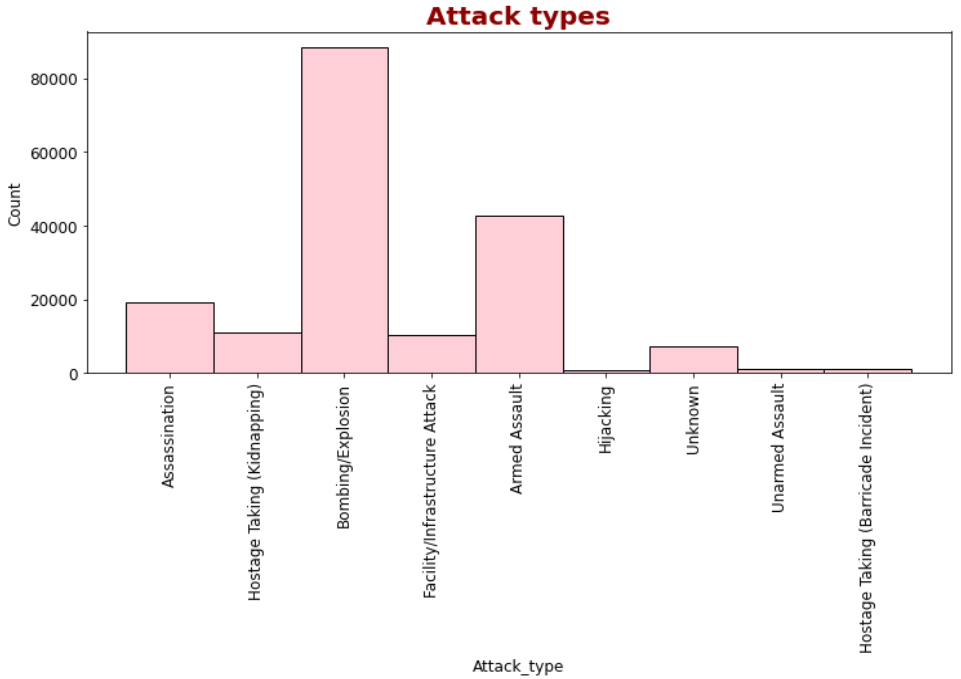


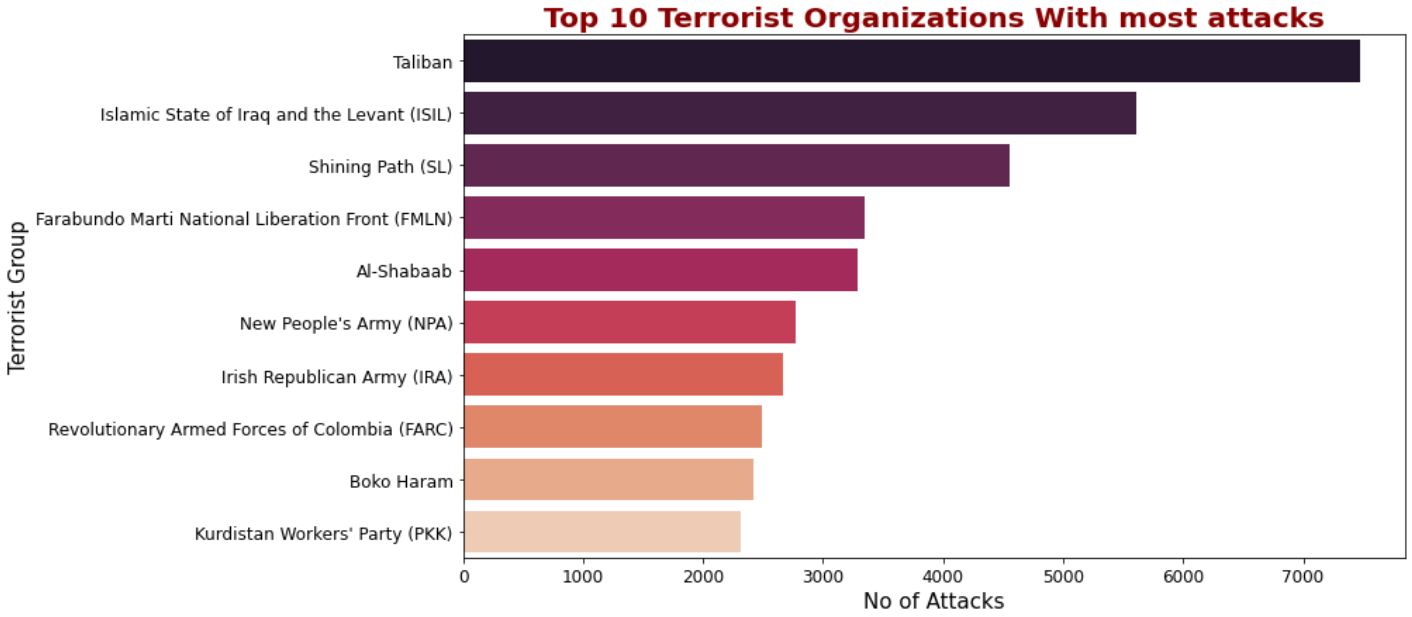


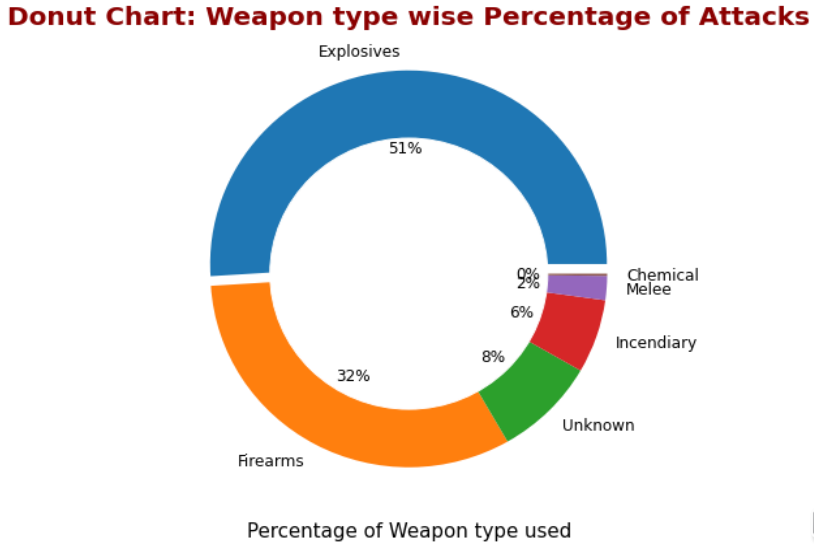


UNIVARIATE ANALYSIS

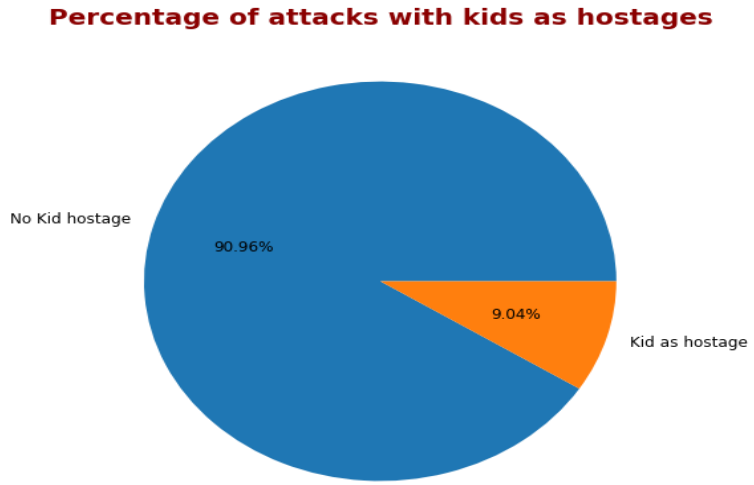


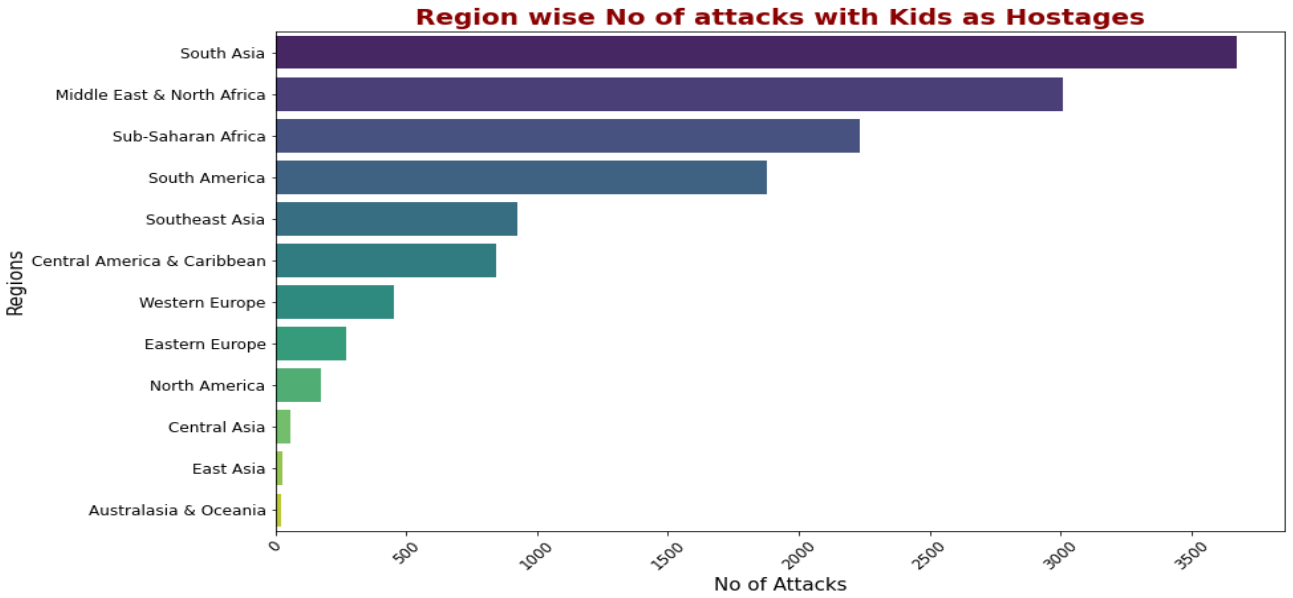


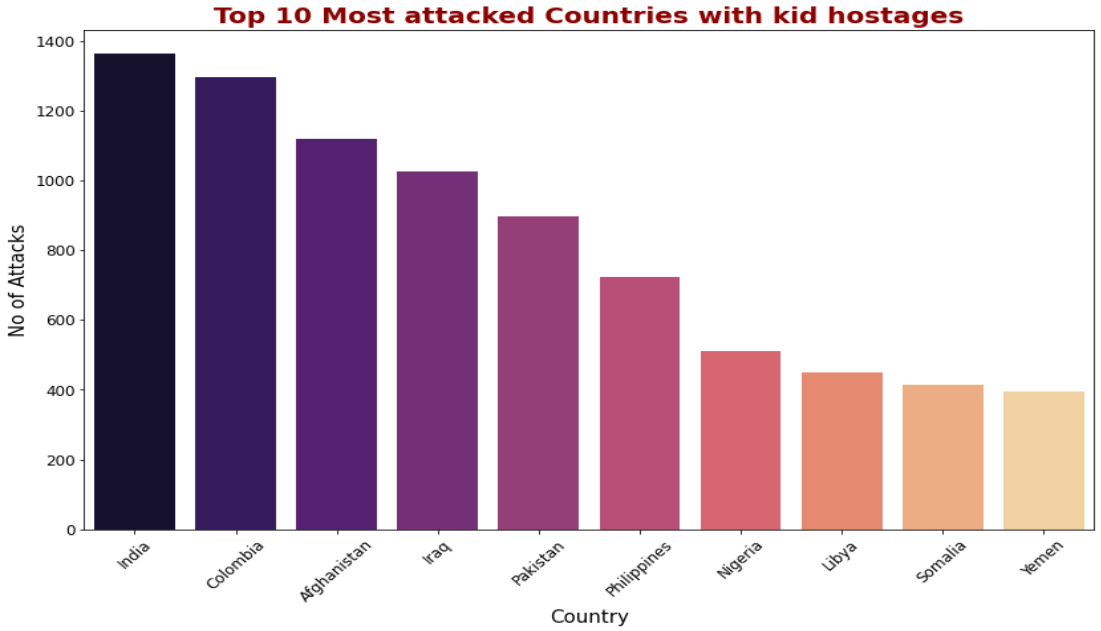


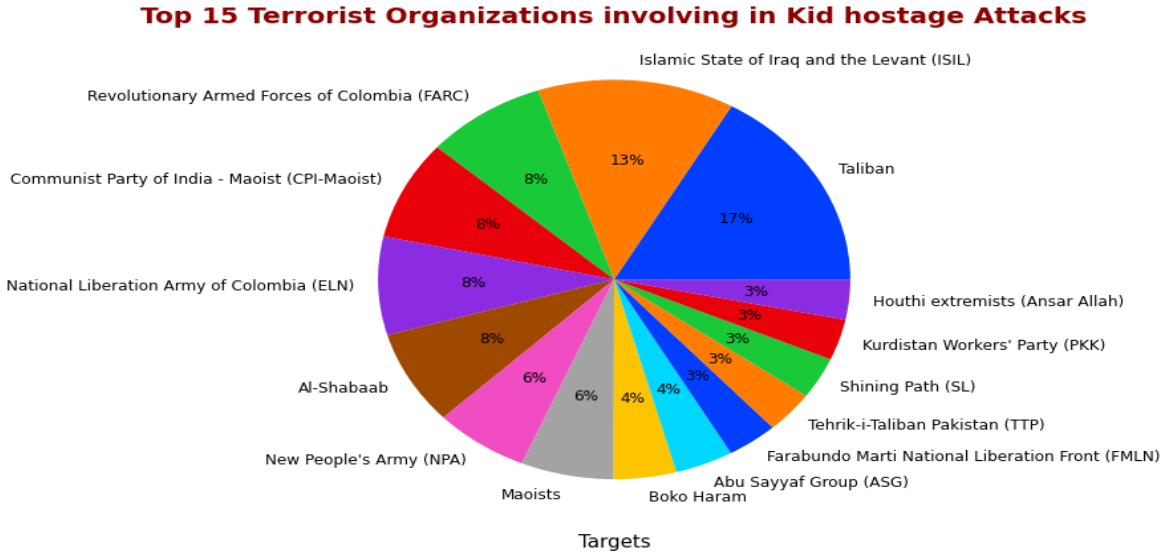


UNIVARIATE ANALYSIS

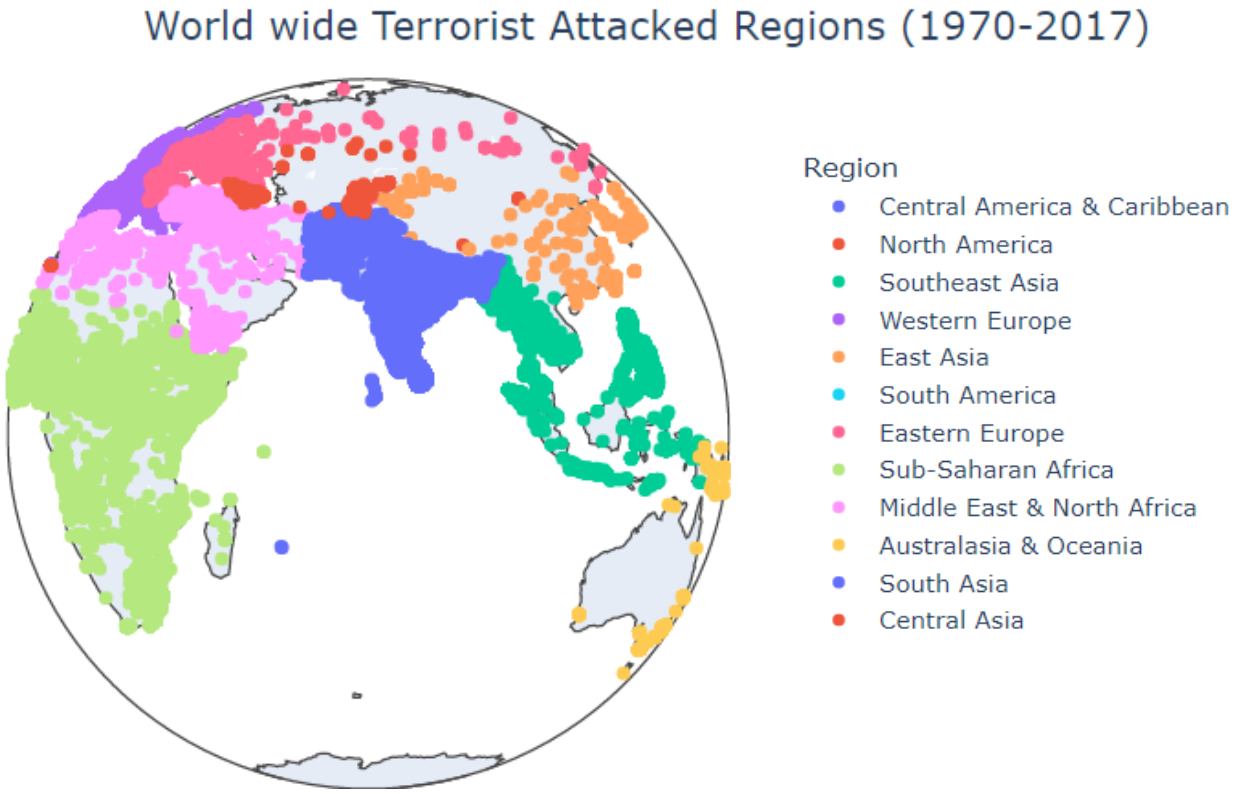


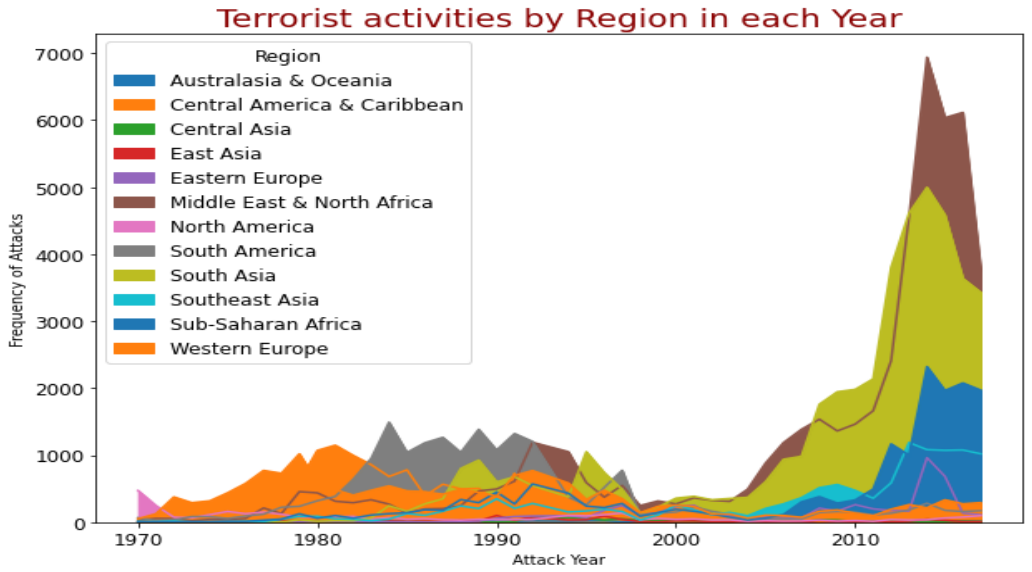


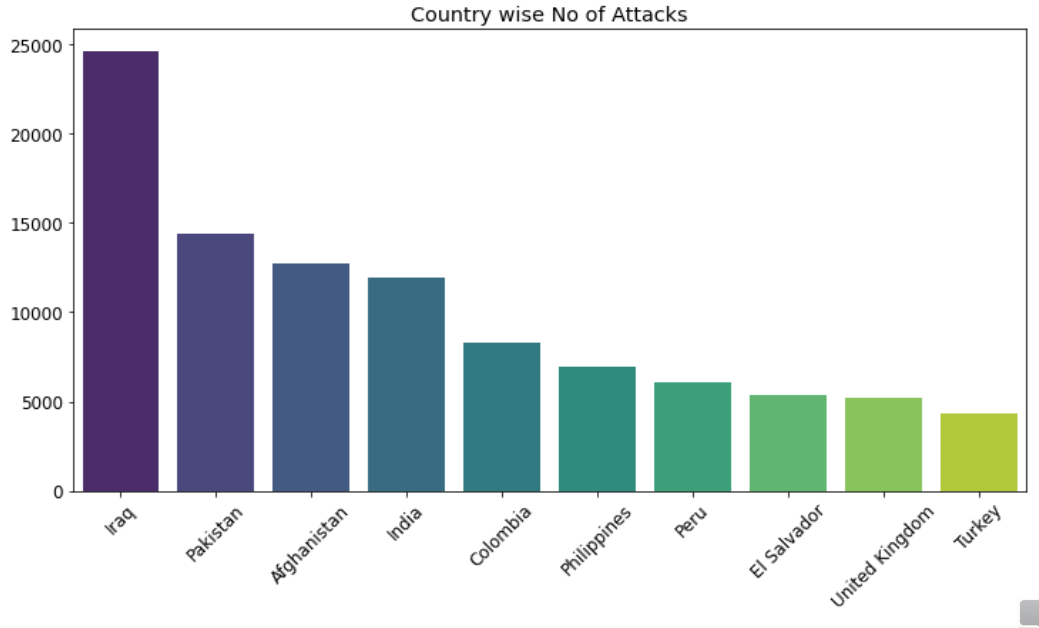




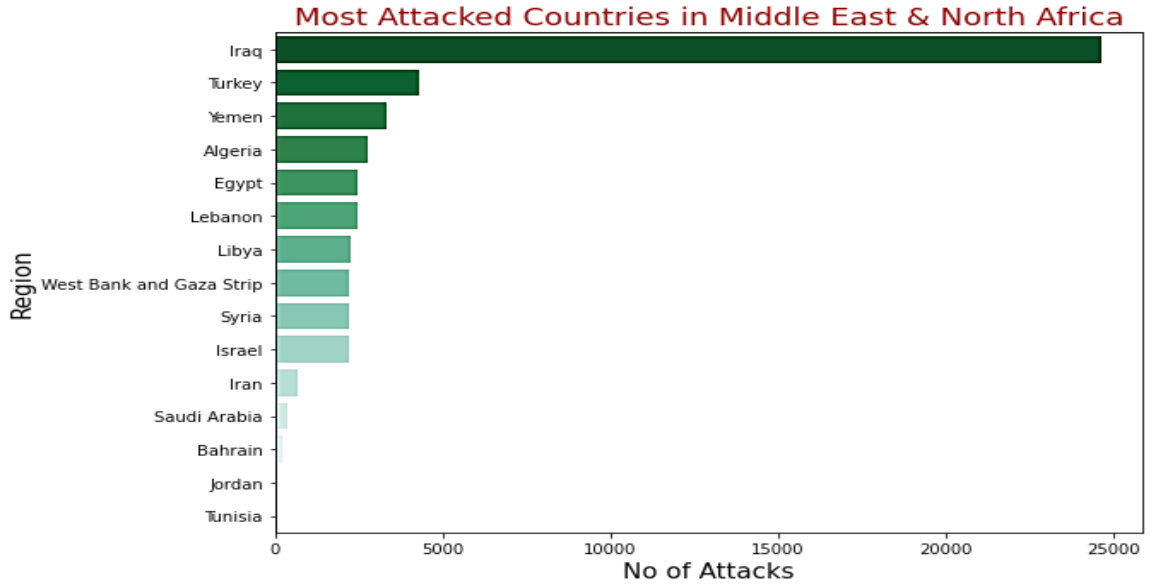
UNIVARIATE ANALYSIS

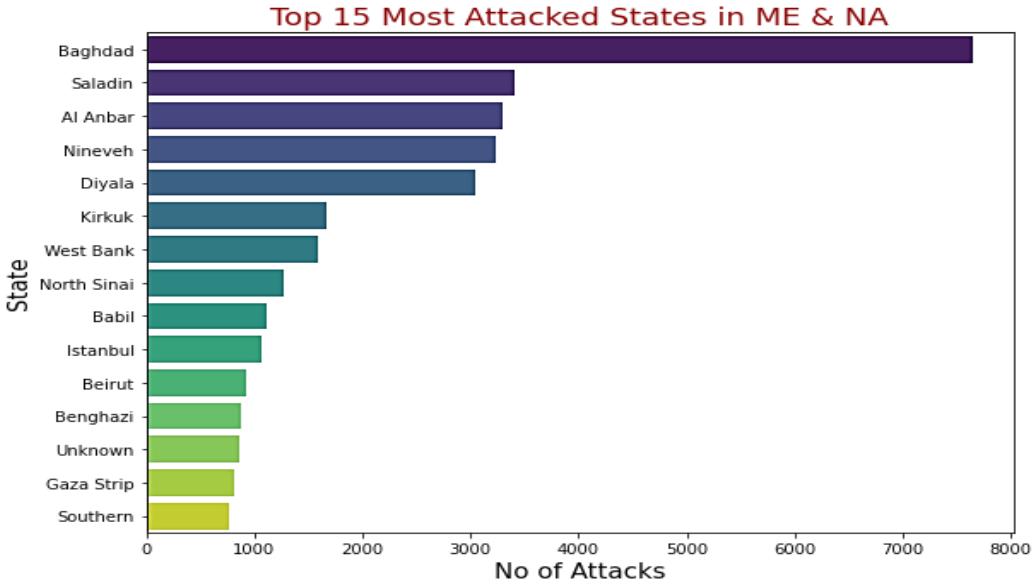


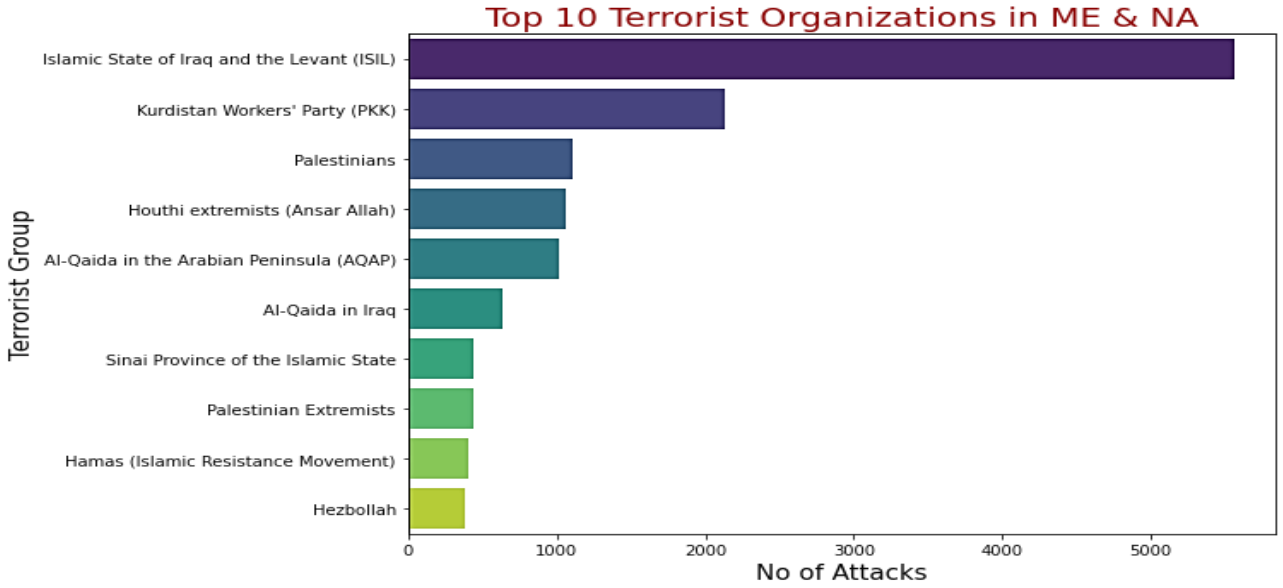


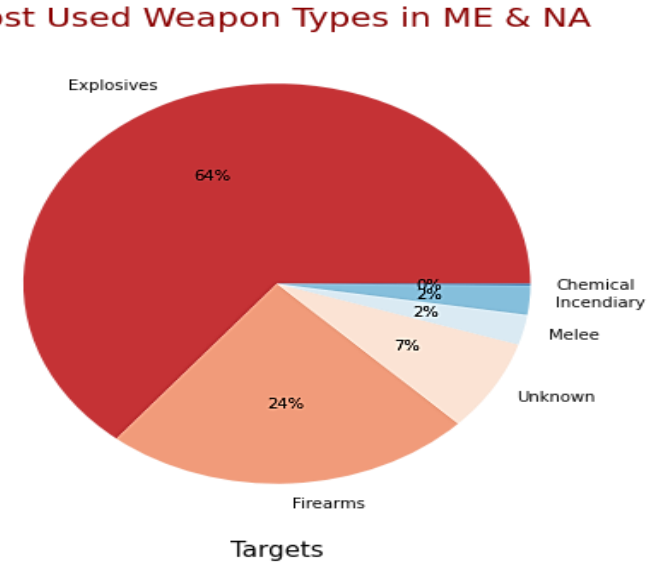


UNIVARIATE ANALYSIS

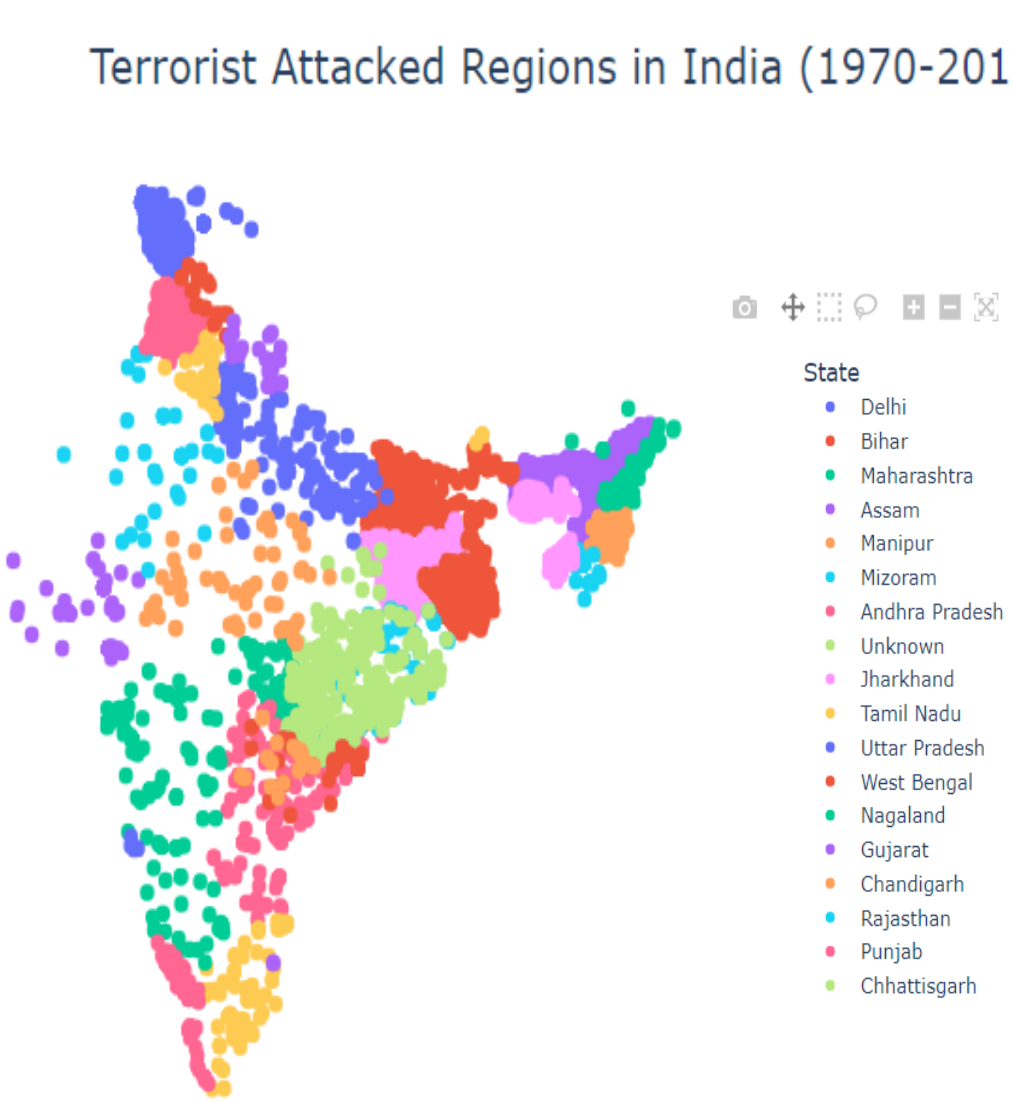


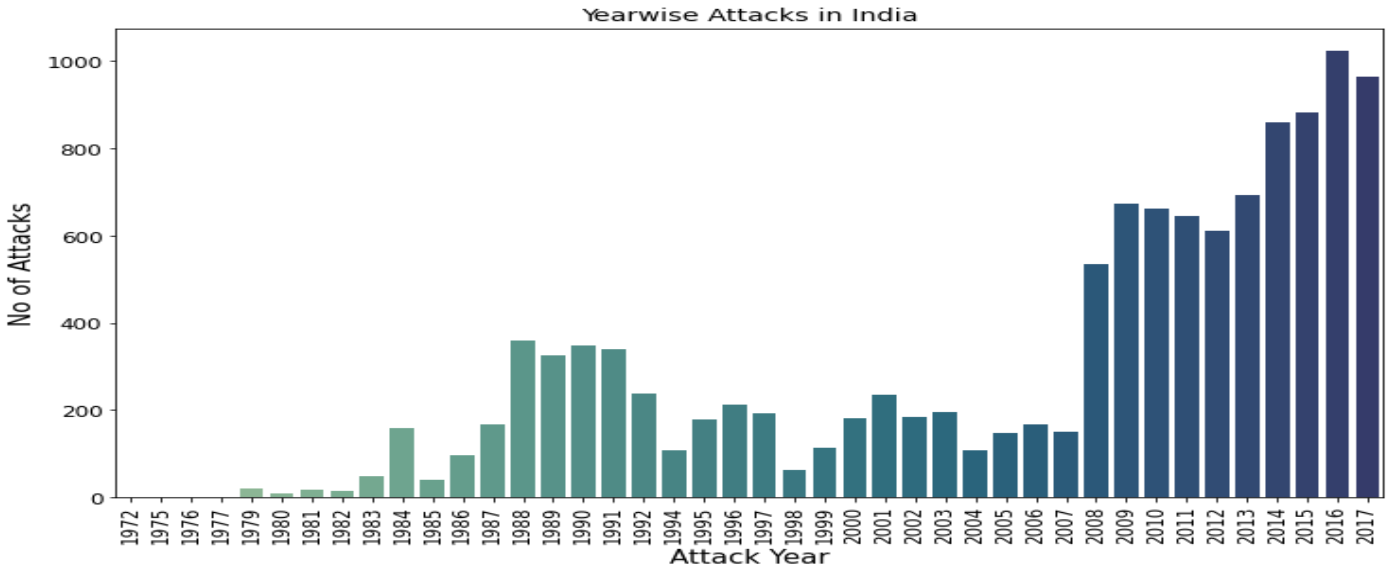


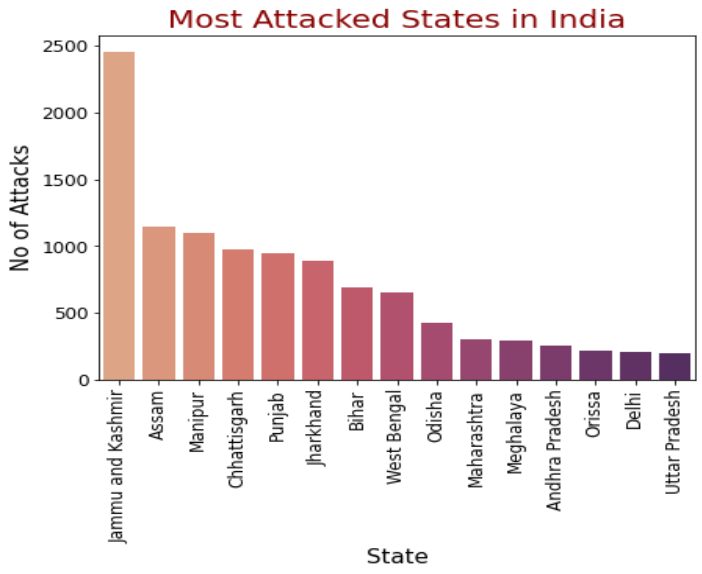


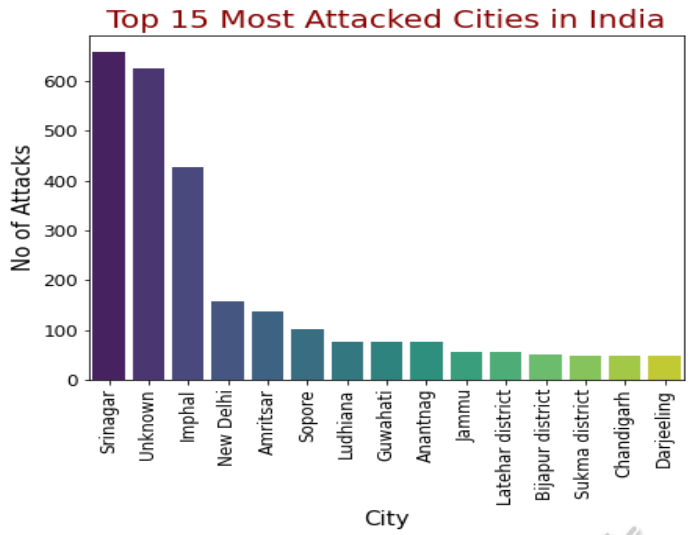


UNIVARIATE ANALYSIS



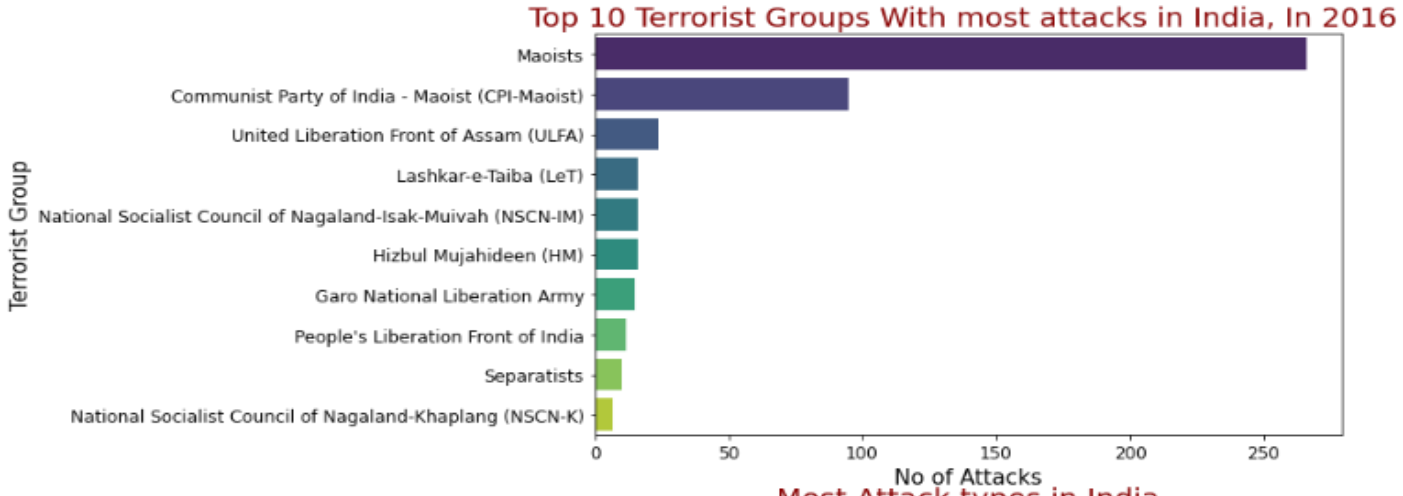


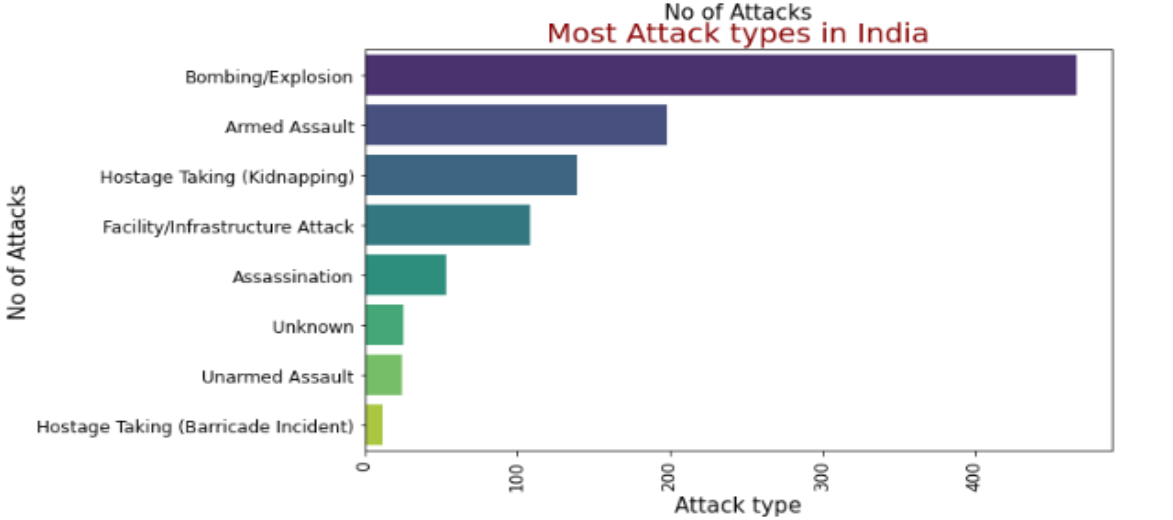


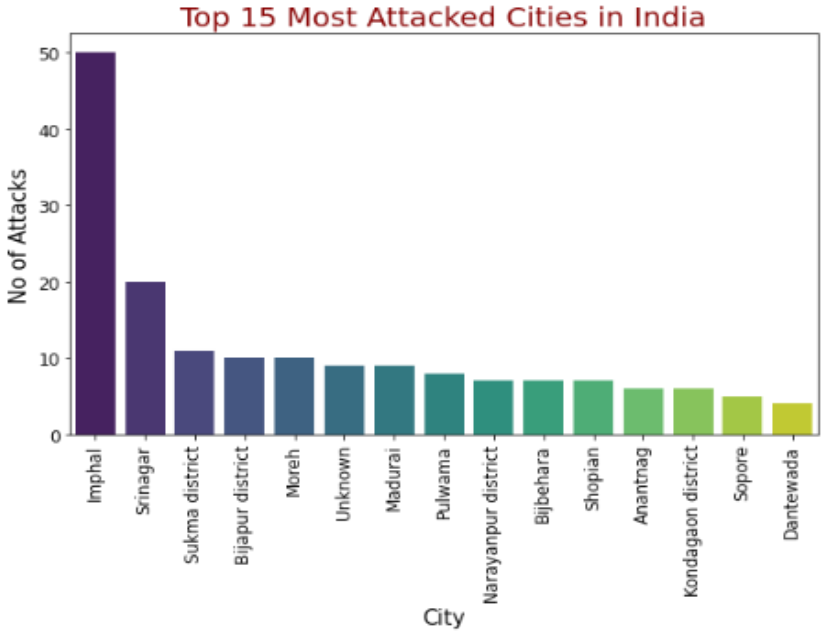


UNIVARIATE ANALYSIS

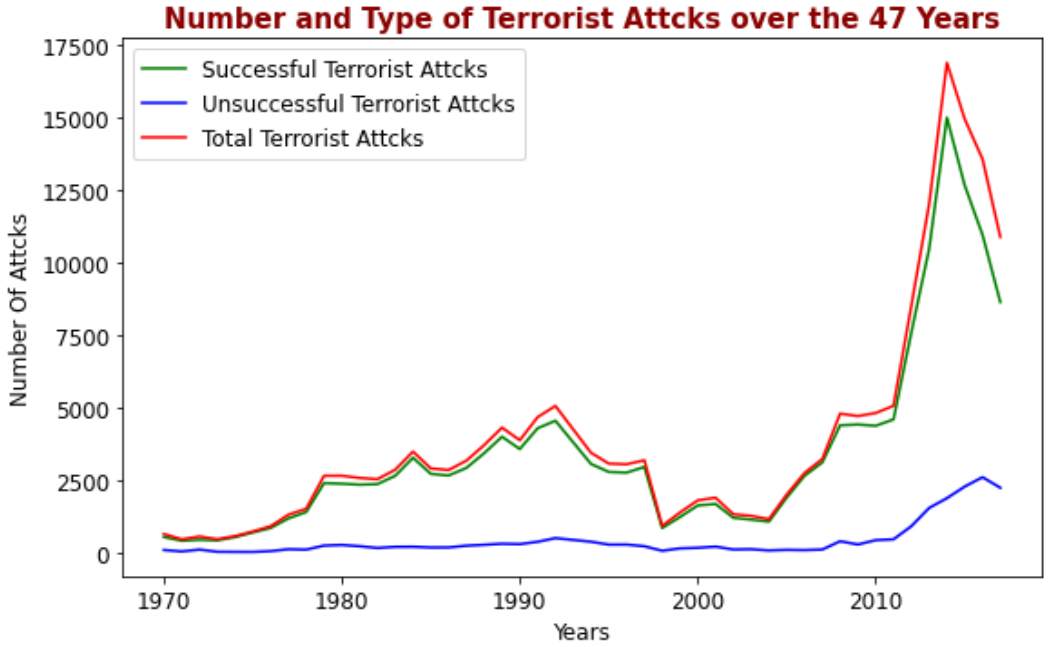
Analysis of Terrorist attacks in India in the Year 2016

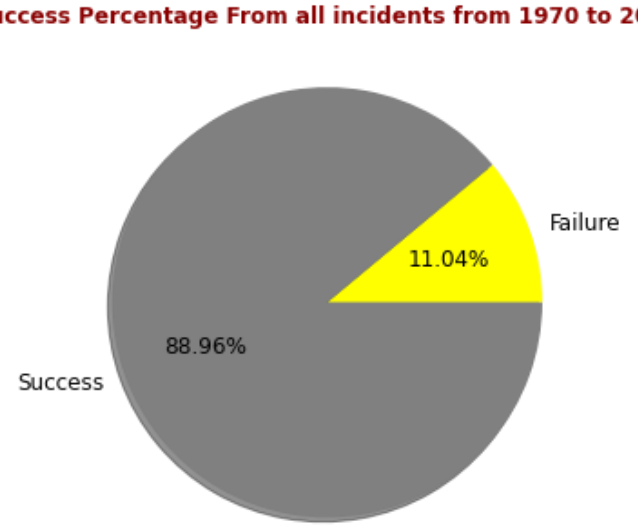


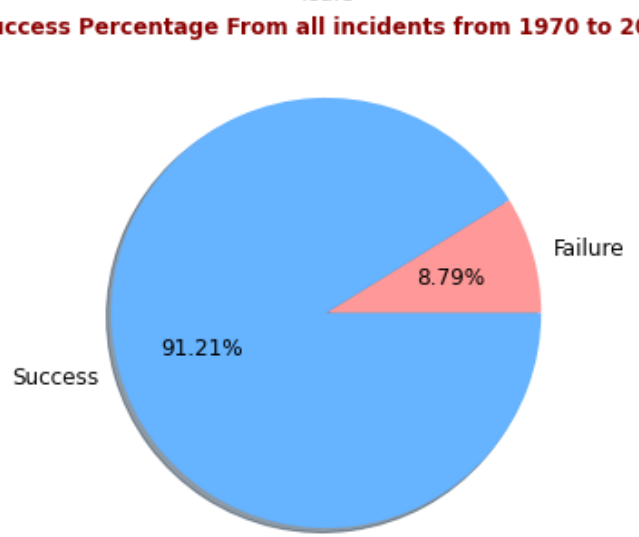


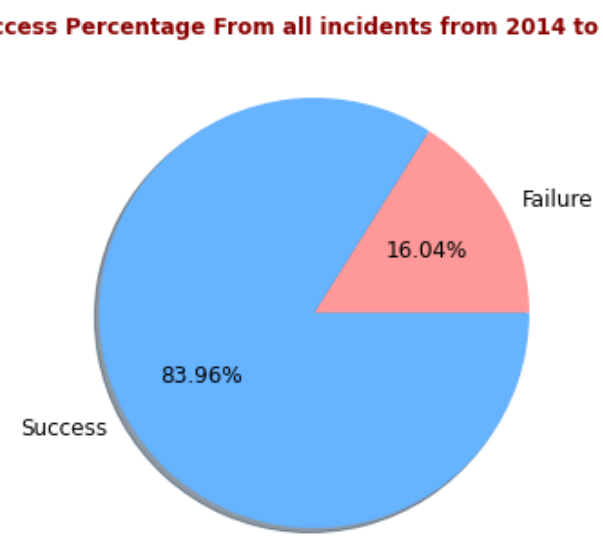


UNIVARIATE ANALYSIS

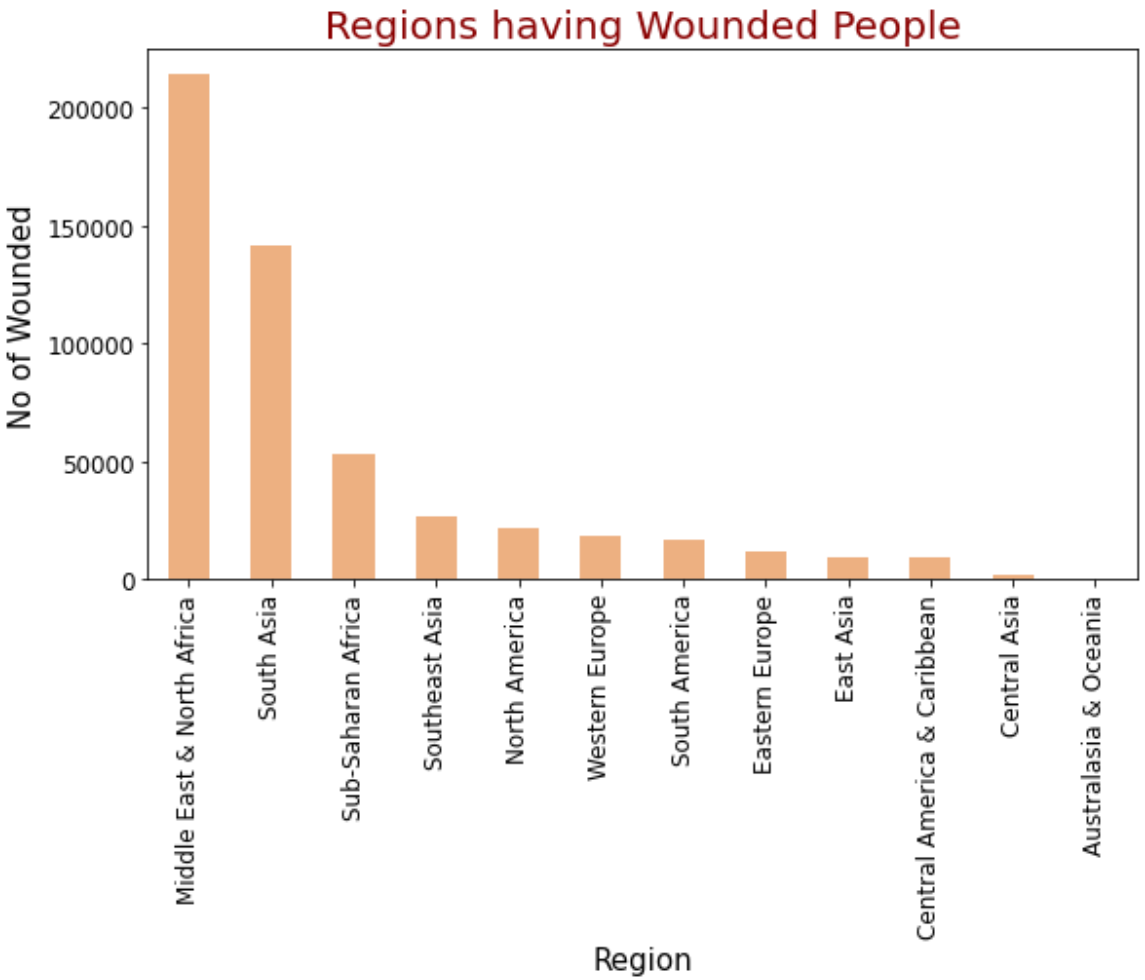


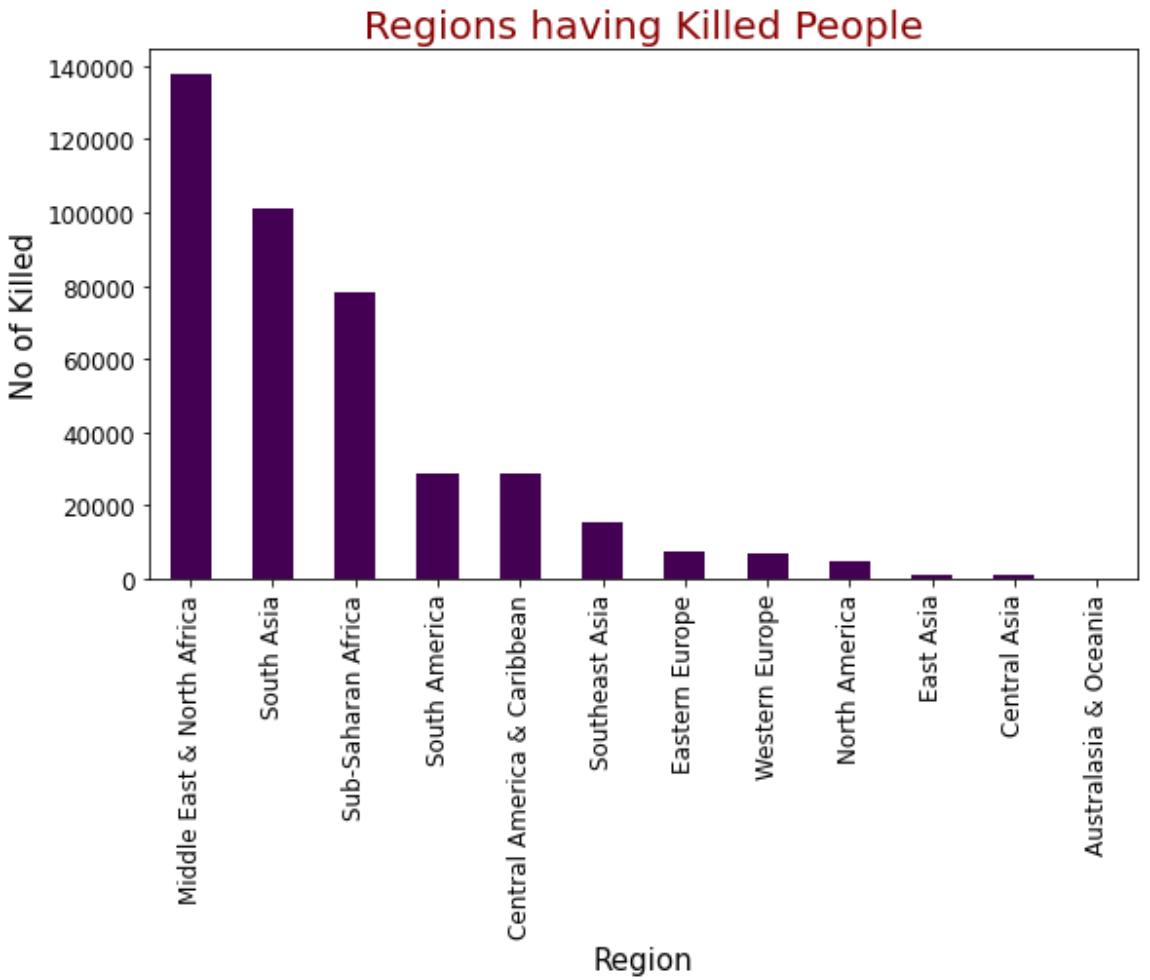






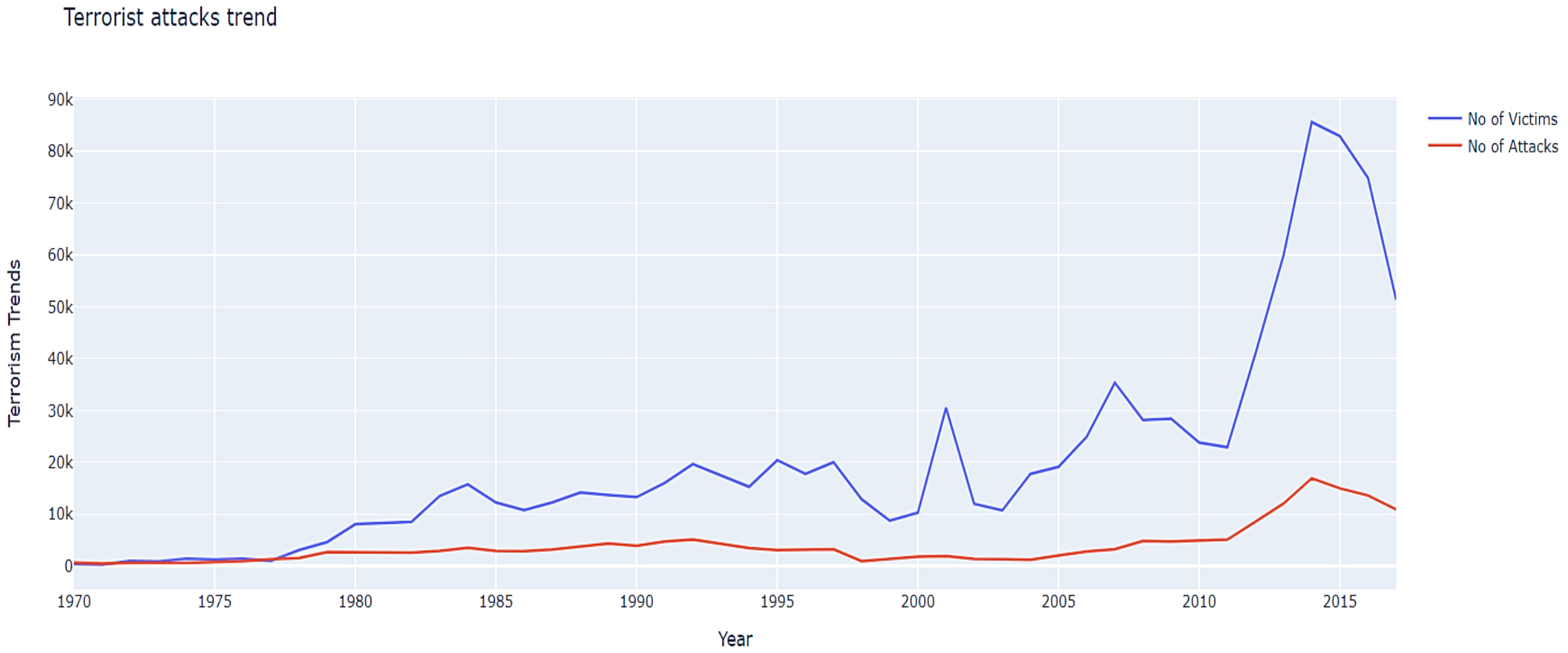
BIVARIATE ANALYSIS





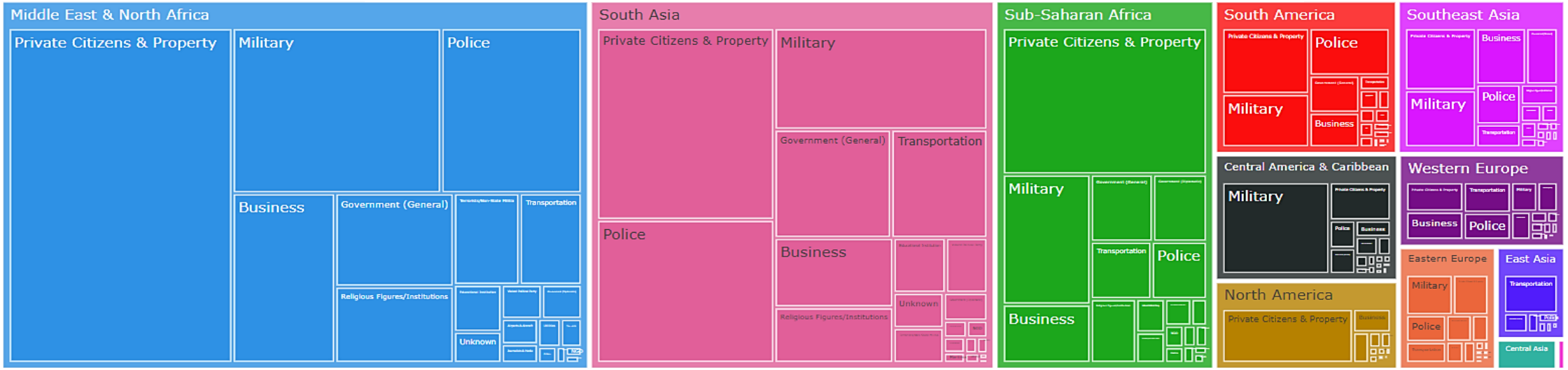
BIVARIATE ANALYSIS

Analysis of No of Terrorist attacks Vs No of Victims

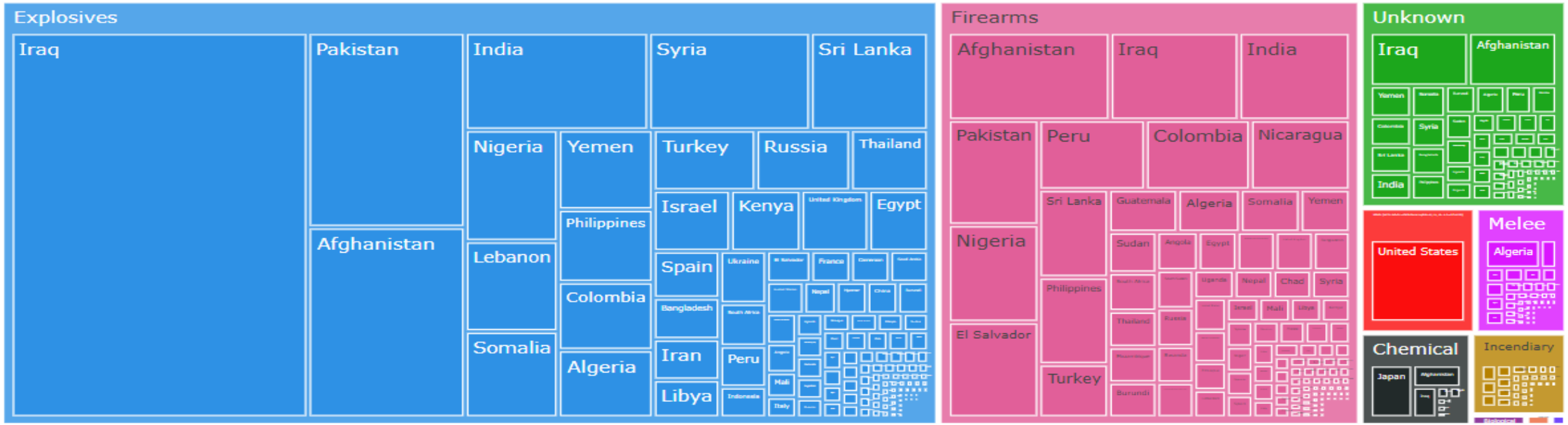


BIVARIATE ANALYSIS

Treemap Chart: Grouping the Target type groups in different regions

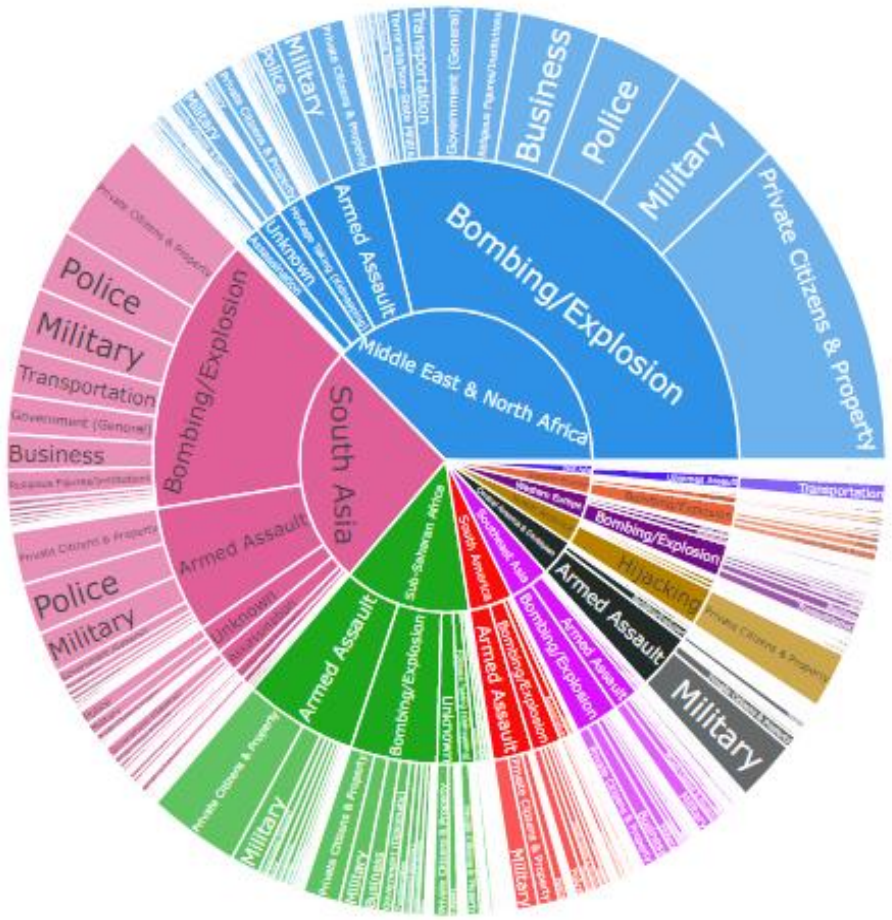


Treemap Chart: Grouping the Weapon type wise attacks with respect to country

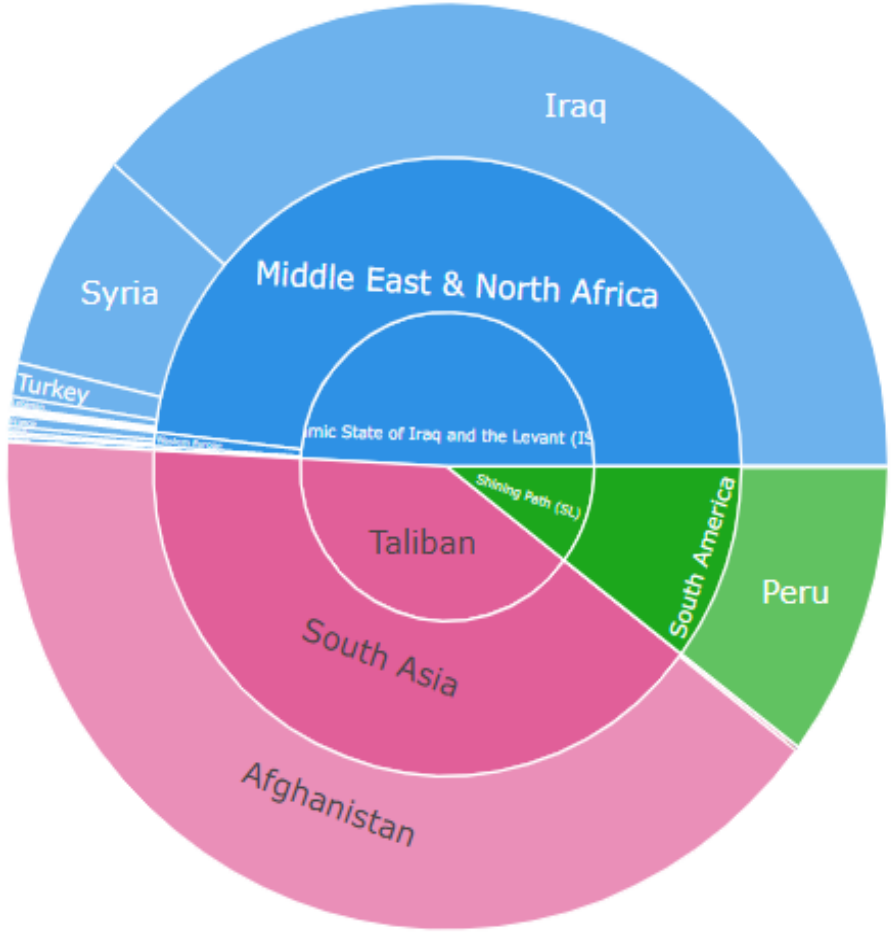


MULTIVARIATE ANALYSIS

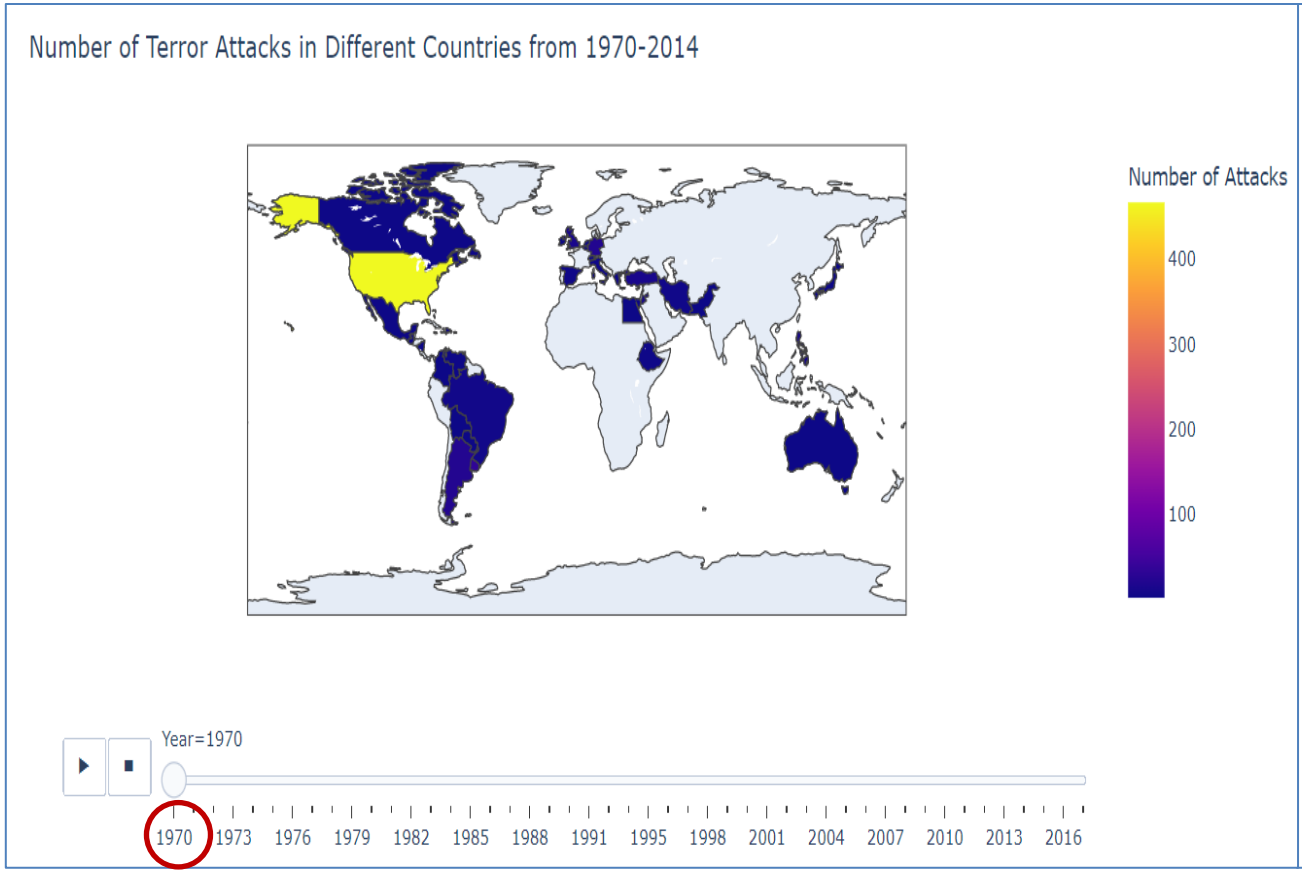
Sunburst Chart: Grouping the Region wise Attack Type & Target types

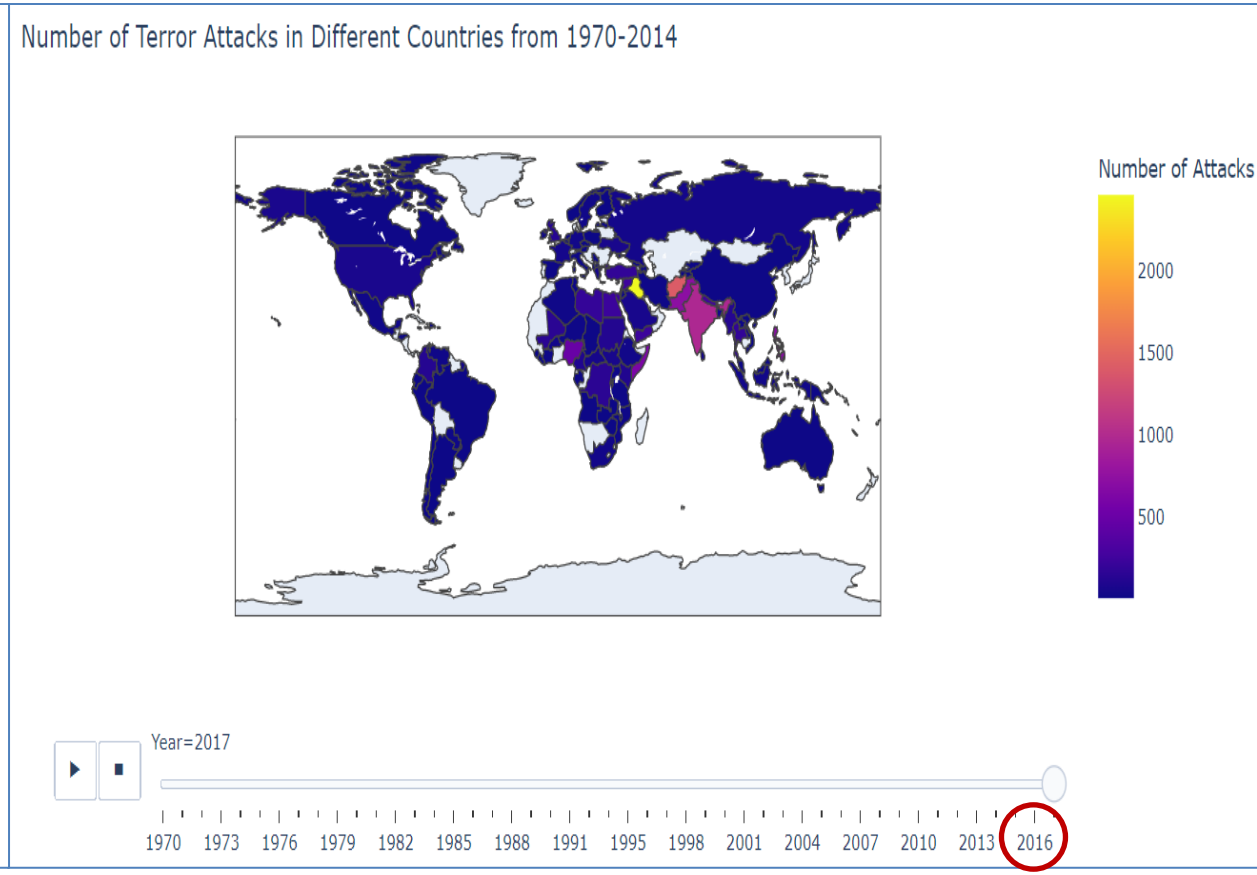


Sunburst Chart: Grouping the Top 3 gangs attacks with respect to region & Country



TIMELINE ANALYSIS OF NO. OF ATTACKS COUNTRY WISE FROM 1970 TO 2017





CONCLUSION

- Terrorist attacks reached a peak during 2014 and then in 2015 started to drop. The Drop in the Number of Terror Attacks is not a sign of improvement in security checks in society as the ability to carry out a successful terror attack has not reduced significantly.
- The Middle East & North Africa is the most affected region in terms of the number of terror attacks among all the regions followed by South Asia, South America, and Sub-Saharan Africa.
- Iraq is the most affected nation of all, and it has suffered the highest number of attacks followed by Pakistan, Afghanistan, and India.
- Most of the attacks are either bombing or explosions followed by firearms and the most used weapons in terror attacks are Explosives and assault weapons.
- Taliban and ISIL are the most active terrorist organizations in the world and Afghanistan is the country that has suffered the most because of Taliban's.
- In 2016, India suffered the highest number of attacks, which was peaked during Aug-2016. Most attacks were happened in Srinagar (Jammu & Kashmir). Maoists were responsible for most of the attacks.
- Terror Organizations and Attacks are slowly moving to Underdeveloped and Developing Countries with large populations below the poverty level in South Asia and Africa.
- This creates a cycle of terrorism as poverty and unemployment help radical ideologists lure youth towards terrorism, discouraging investors which in turn affects the earnings of local governments and leads to further poverty and unemployment.

MITIGATION MEASURES

- Defining Terrorism on a global stage and adopting an International Framework to tackle terrorism.
- Creating a joint task force to improve Intelligence Sharing and sharing best practices to combat terrorism.
- Countries should take measures against illegal immigration by implementing better tracking systems.
- By imposing strict regulations on the access and sale of Dangerous weapons and firearms and tracing sources of such weapons.
- Adopting a global framework and implementing a tough action plan against terror financers.
- Education in population control and improving conditions to encourage employment opportunities might reduce terror activities in the long run.
- Strict policies against disinformation and radicalization need to be adopted, including media & internet.

THANK YOU