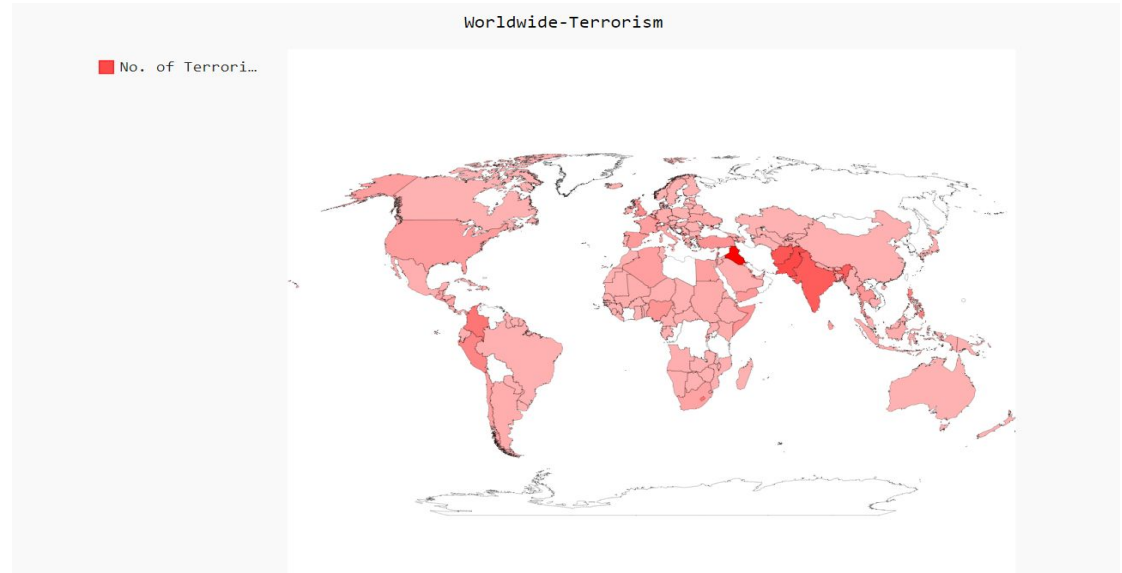


# Capstone Project

## Global Terrorism Analysis

## Let's save the world

- 1) Defining the problem statement.
- 2) Summary of the data on the Global Terrorism.
- 3) Data Preprocessing.
- 4) Data Analysis and Visualization
- 5) Challenges faced.
- 6) Conclusion



# The Problem Statement

Global Terrorism Database is the most comprehensive database of terrorist attacks in the world. This database provides information on the domestic and international terrorist attacks around the world since 1970 and includes more than 180,000 events.

The idea of this project is to use the database and provide the most valuable insights about the terrorism and their groups in the world. To achieve our goal we have divided the analysis into three parts Terrorism on World level, Terrorism in India, Attacks and Damage done by the Terrorist groups.

# Data Summary

The dataset contains data of more than 180000 terrorist events happened since 1970 and has around 135 variables describing each attack. Some of the key attributes consisting those variables which are used for this project are listed below:

Attribute	Description
Date, Month, Year	Calendar details of the event executed by the attackers.
Country_txt	Name of the countries where the event had executed.
Region_txt	Name of the region where the attack happened. Region_txt consists values like East Asia, South Asia, Western Europe, etc
Provstate	Name of the state in which the attack happened.
City	Name of the city where the attack happened.

Attribute	Description
Target_type	Type of target involved in the attack. Target1_txt consists of categorical values like private citizens, military, police, government officials, transportation, education, religious institution, airports, etc
Attack_type	The type of attack happened. Attacktype1_txt consists of categories like explosion, armed assault, assassination, kidnapping, unarmed assaults.
Nkill	Deaths occurred in the attacks by the Terrorist Groups in each country.
Nwound	Citizens injured in the attacks by the Terrorist Groups in each Country.
Weaptype1	Weapons used by the Terrorist Groups in each attack.
GName	List of the Terrorist Groups who took the responsibility of the attack.
Success	'1' if attack was a success. '0' if attack was a failure.
Suicide	'1' if attack was a suicide attempt. '0' if attack was not a suicide attempt.

# Data Preprocessing

Data preprocessing is the first step to be done after collecting data. It is a set of operations performed on the dataset to modify ambiguous data which can be a bottleneck to analytical results. Raw data is simply a collection of related information put together. This data is often unorganized and contains a lot of information which is irrelevant to the project requirements. Data preprocessing methodology helps in converting this raw data into a more meaningful, focused, interpretable and readable format.

The following are the steps used in this project as a part of data preprocessing methodology:

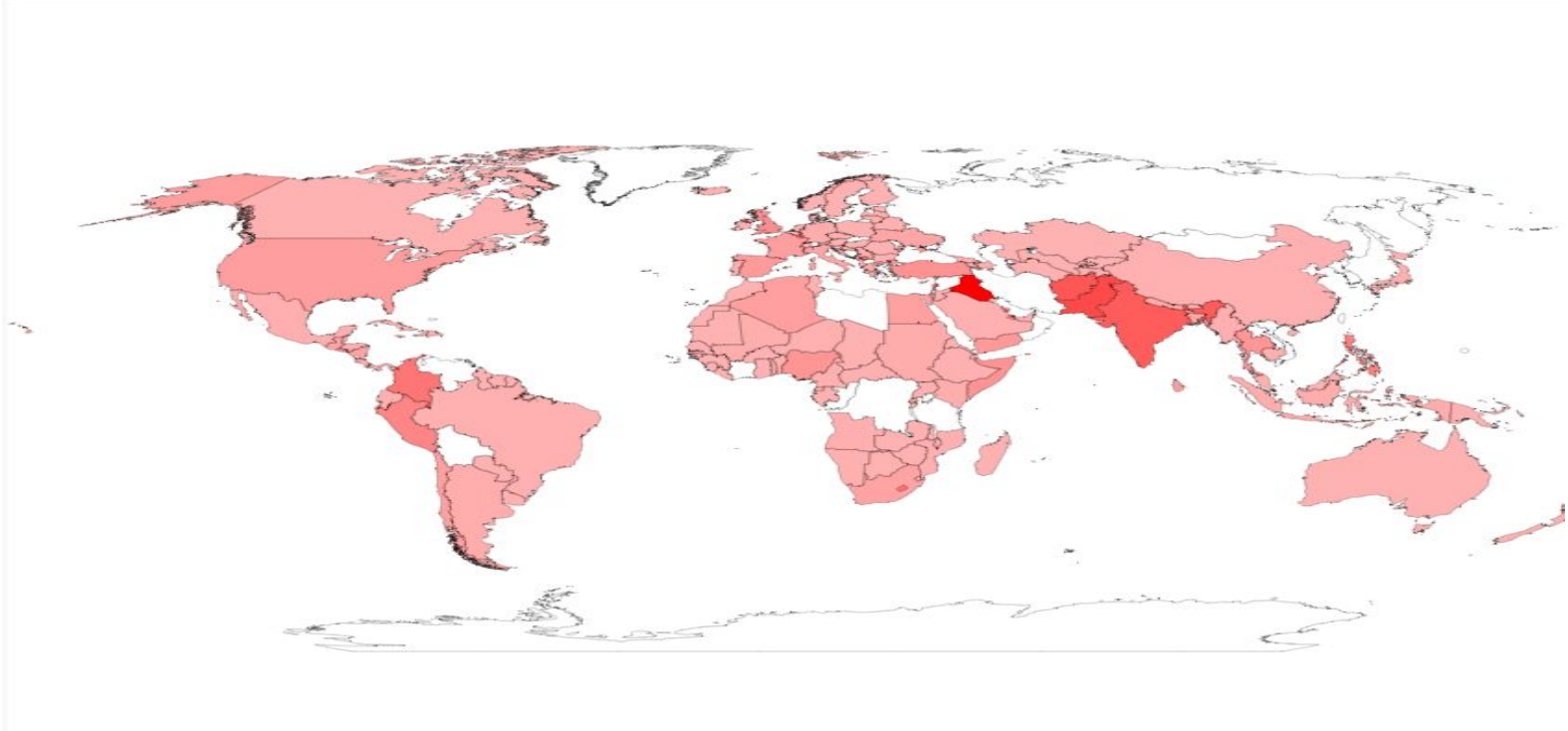
**Data Cleaning:-** Process of handle inconsistencies in data. In terrorism dataset, there are numerous fields like 'motives' or 'responsible organizations' which are missing either due to information not available or that field was not relevant for that specific event. Fields like 'summary', 'claim\_mode', 'claimmode\_txt', 'guncertain', 'nperps' etc. are removed since they are not relevant to the analysis of this project. Fields like 'weapsubtype2', and 'weaptype3\_txt' have more missing values than valid entries. Hence such fields are also removed to reduce complexity.

**Data integration:** In this step, conflicts among data are resolved. Different representations of the same data such as multiple subcategories of weapon type (weapsubtype1, weapsubtype2, weapsubtype3) are put together to avoid confusion and duplications. Fields with one to one correspondence like 'country\_code' and 'country\_txt' are mapped to avoid any conflicts.

**Data transformation:** Here data aggregation, generalization, and normalization are performed. Dataset has Nkills and Nwounds which can be shown as a single attribute by the name of the Damage in the analysis. This technique reduces the total number of attributes in the dataset and hence reducing the variability in the data

# Analysis and the observations

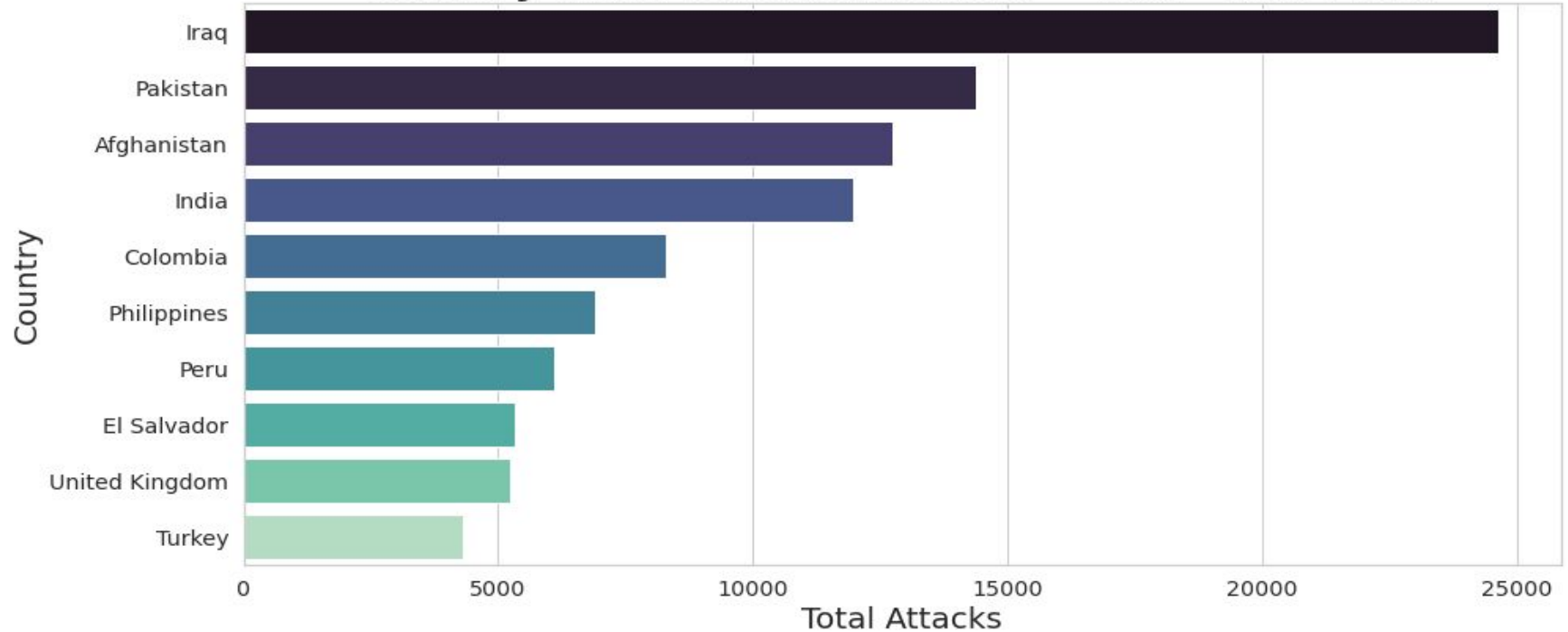
Worldwide-Terrorism



**Asia has the most affected region across the world because of the terrorist activities.**



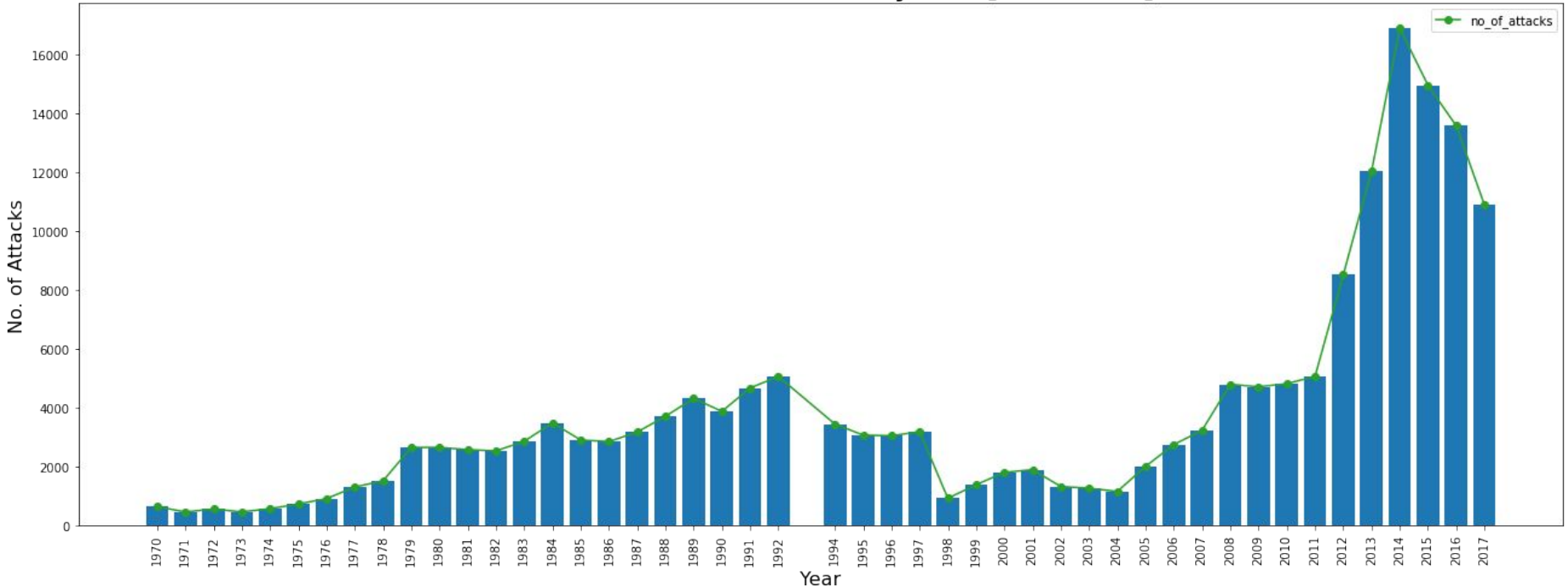
## Country Wise Terrorist Attack - Year 1970-2017



**Iraq is the country that has the most number of attacks in the world between the years 1970 and 2017.**

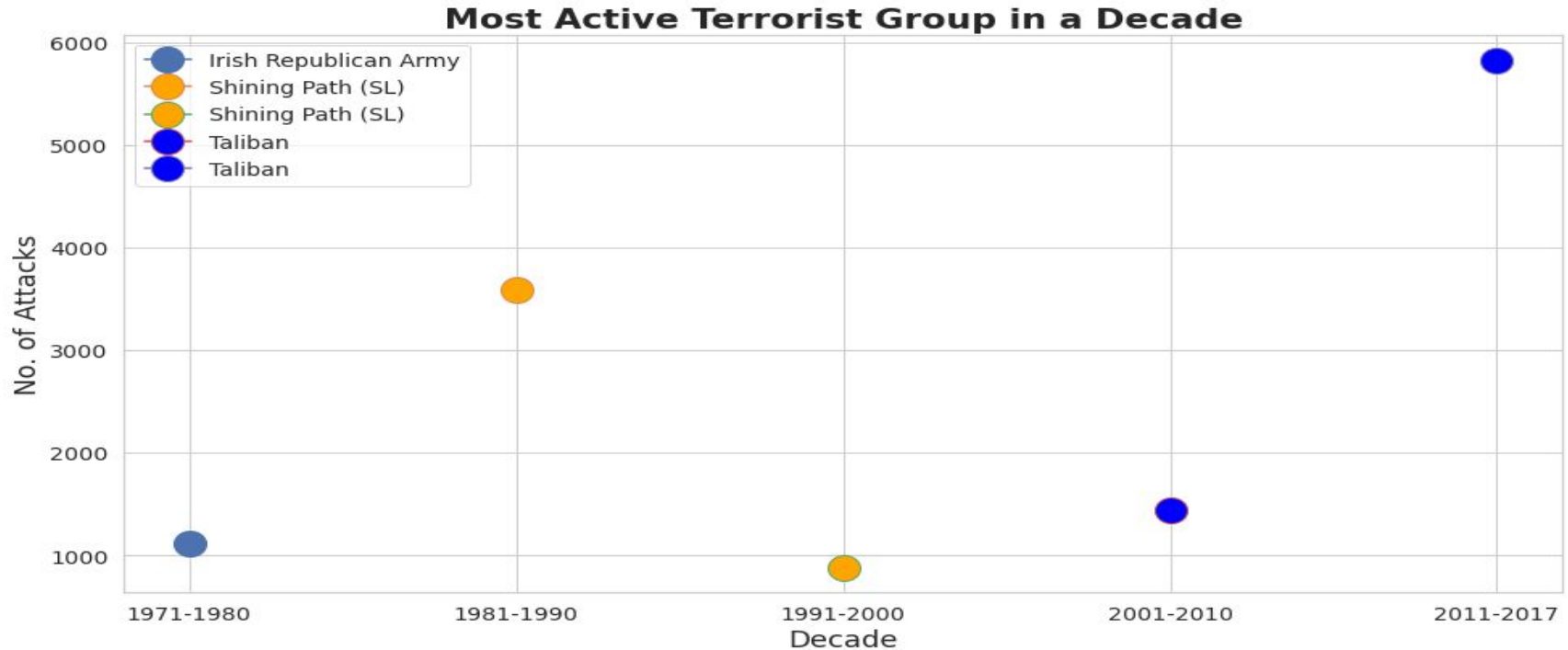
# Yearwise Attack Around the World

Terrorist Attacks Worldwide by Year [1970-2017]



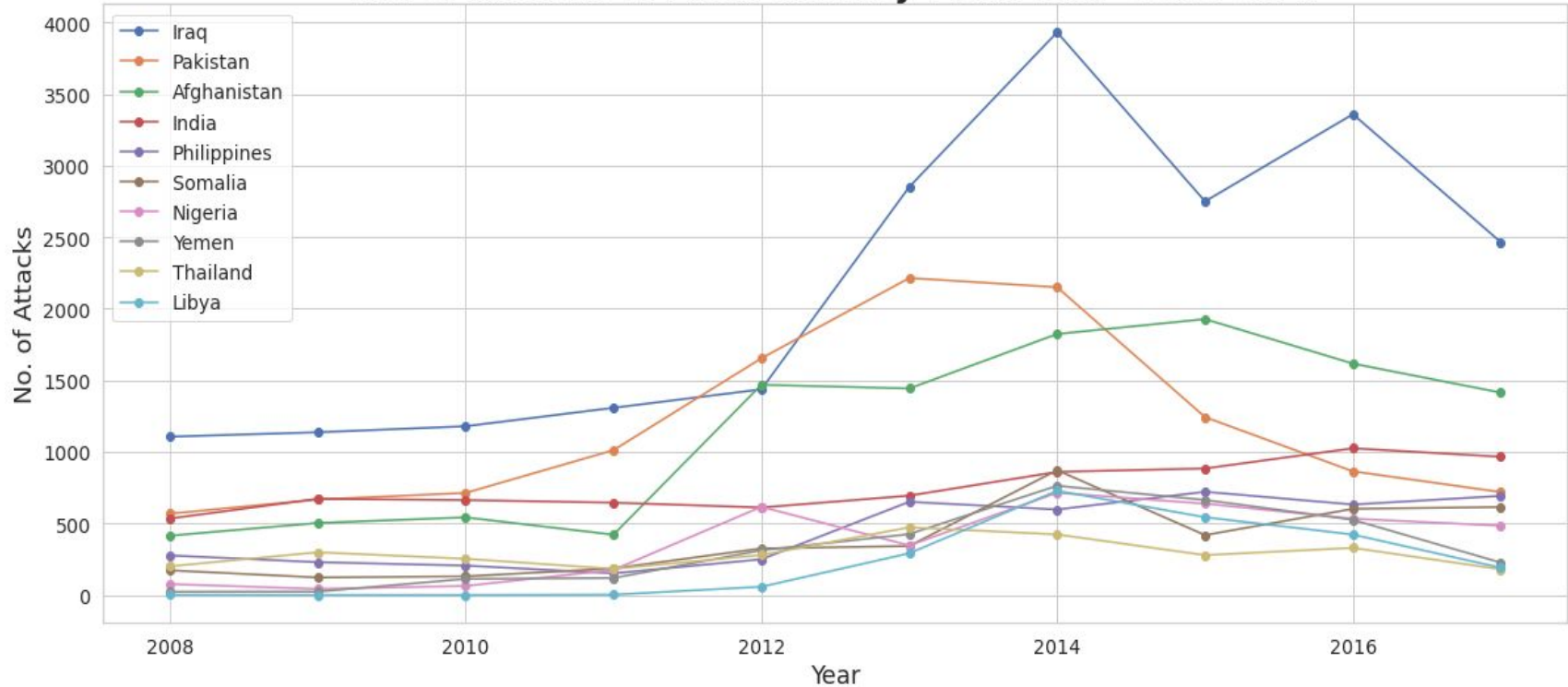
**2014 is the year with most number of attacks i.e 16903 attacks.**

## Most active terrorist group in a decade.



**Taliban is the most active terrorist group from the last two decades i.e from 2001-2017.**

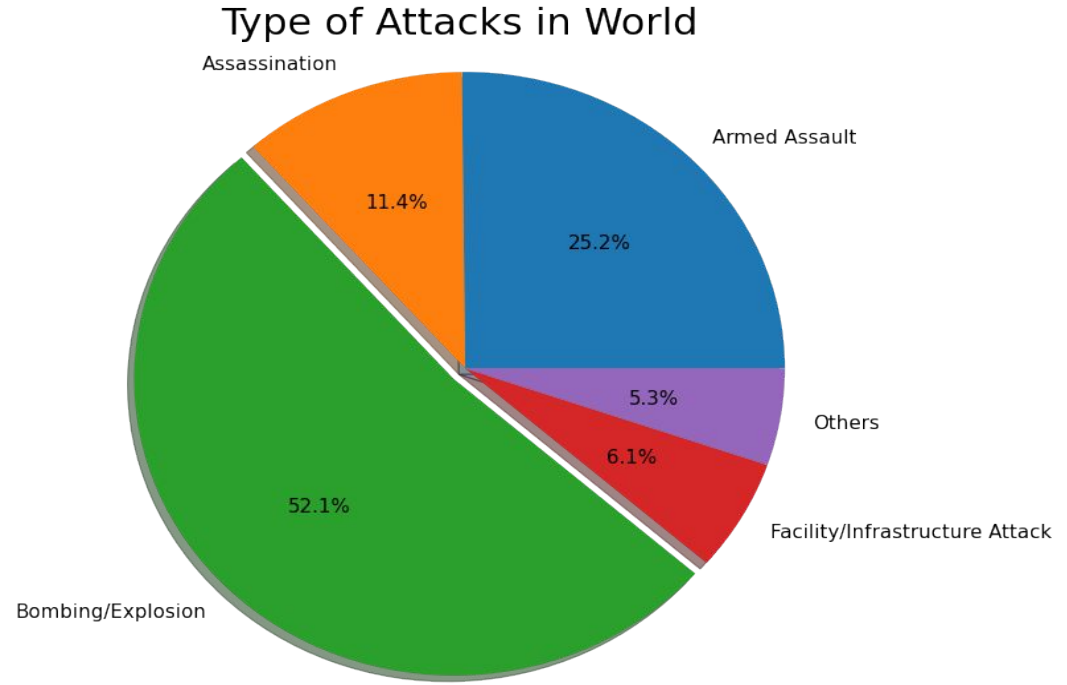
## No. of Attacks in Each Country From Year 2008-2017



From 2008 till 2012 attacks were executed on the constant rate each year in the countries but after 2012 we have seen the rapid rise in the no. of events in Iraq.

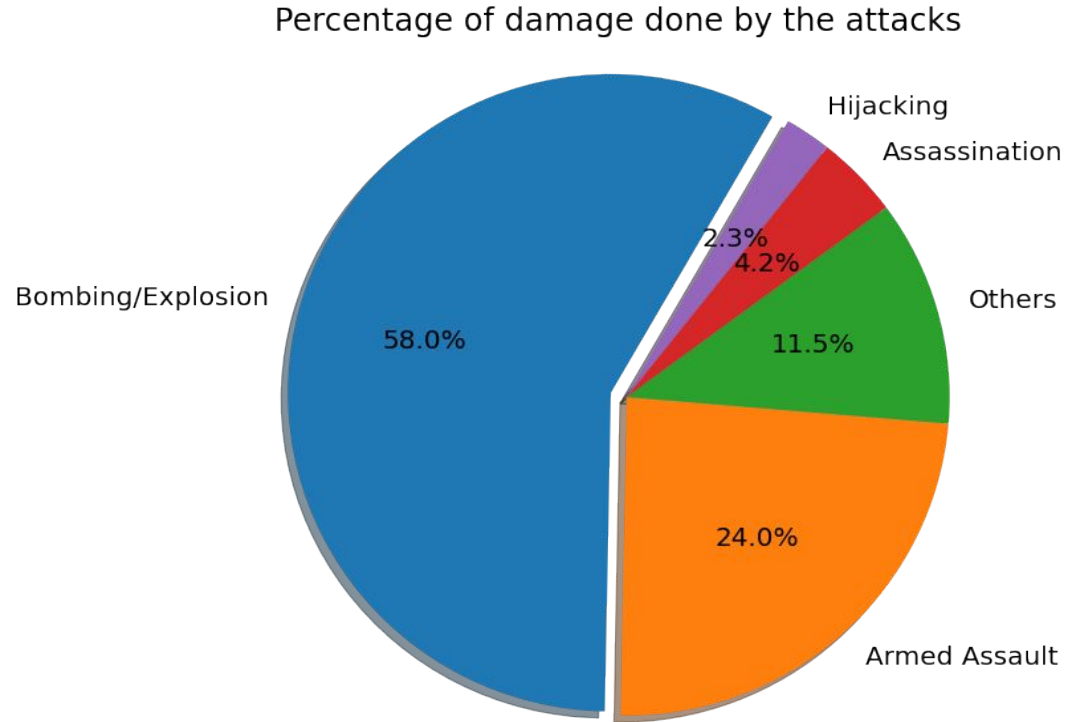
## Type of attacks executed by the Terrorist Groups and their percentage in the world from 1970 to 2017

Bombing & Explosions are the most dangerous type of attack used by the Terrorist Groups to do damage in the countries around the World.

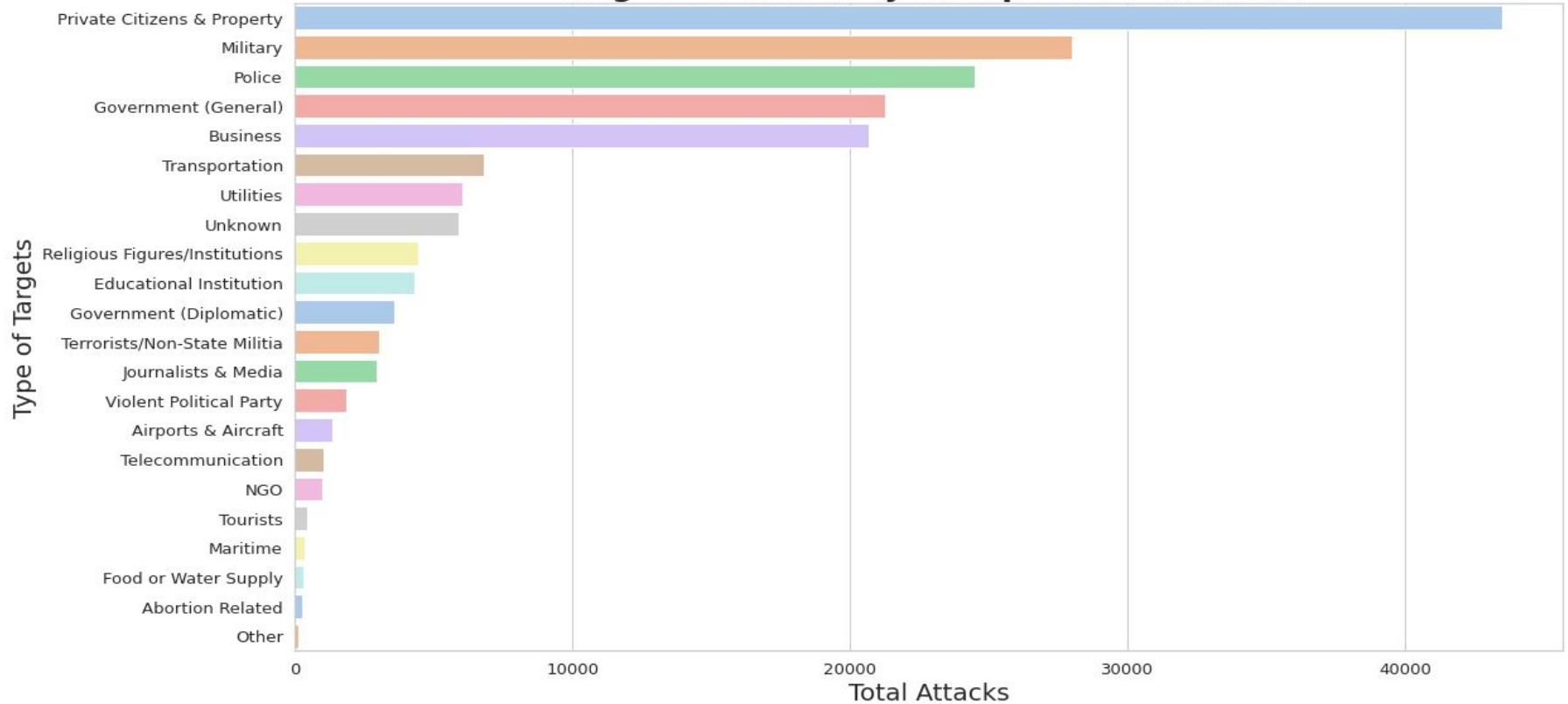


## Percentage of damage done in the world by different attacks

Bombing and armed assault has done the maximum damage in terms of deaths in the World from 1970 to 2017

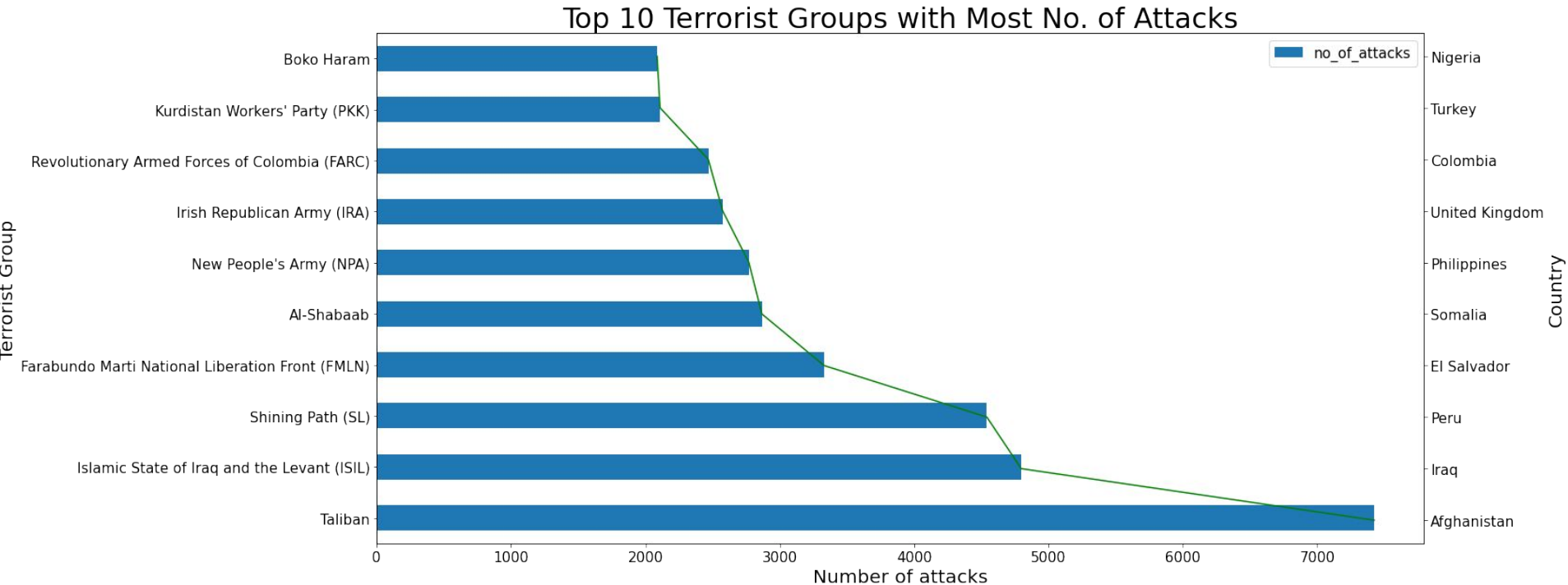


## Targets selected by Groups in the attacks



**Terrorist Groups are more likely to attack on the private properties like Malls or on Citizens.**

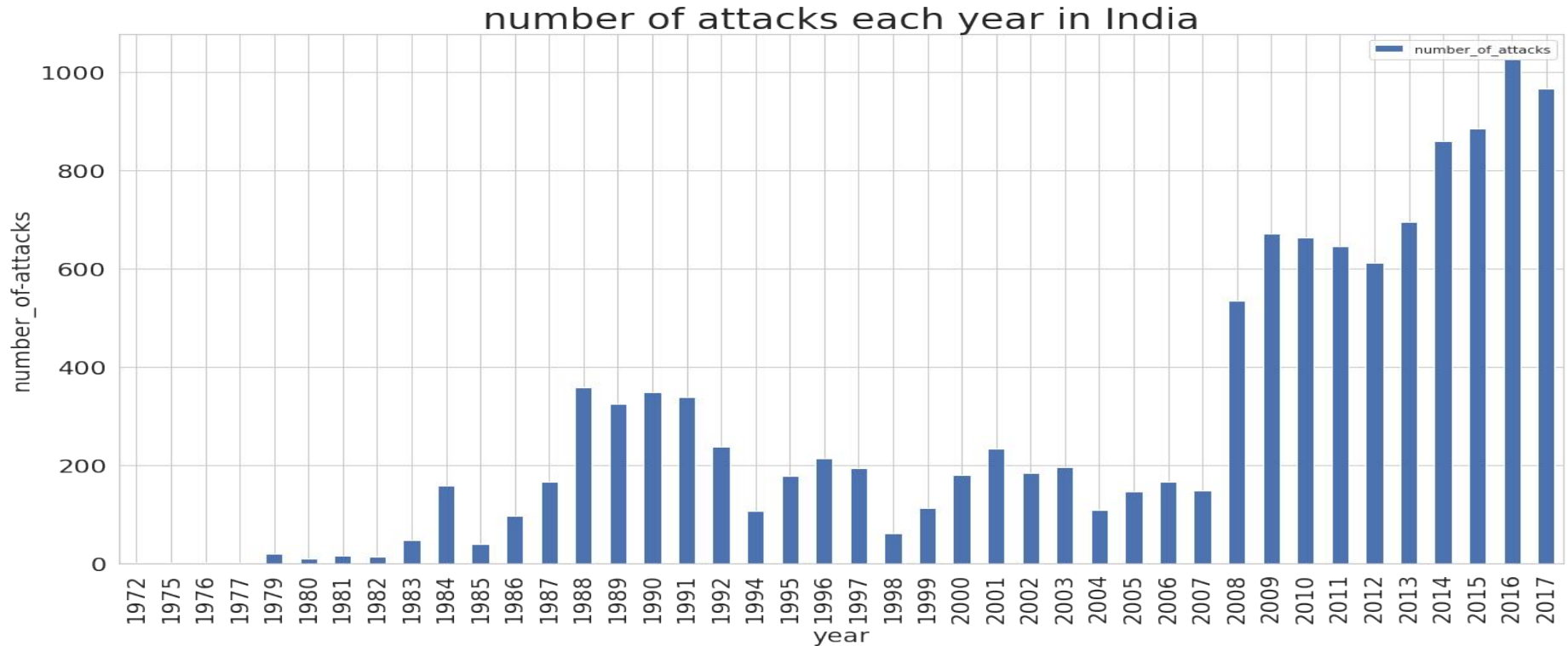
# Top Ten Rivalries



**Taliban has the highest number of attacks till date and they mostly operate in Afghanistan.**

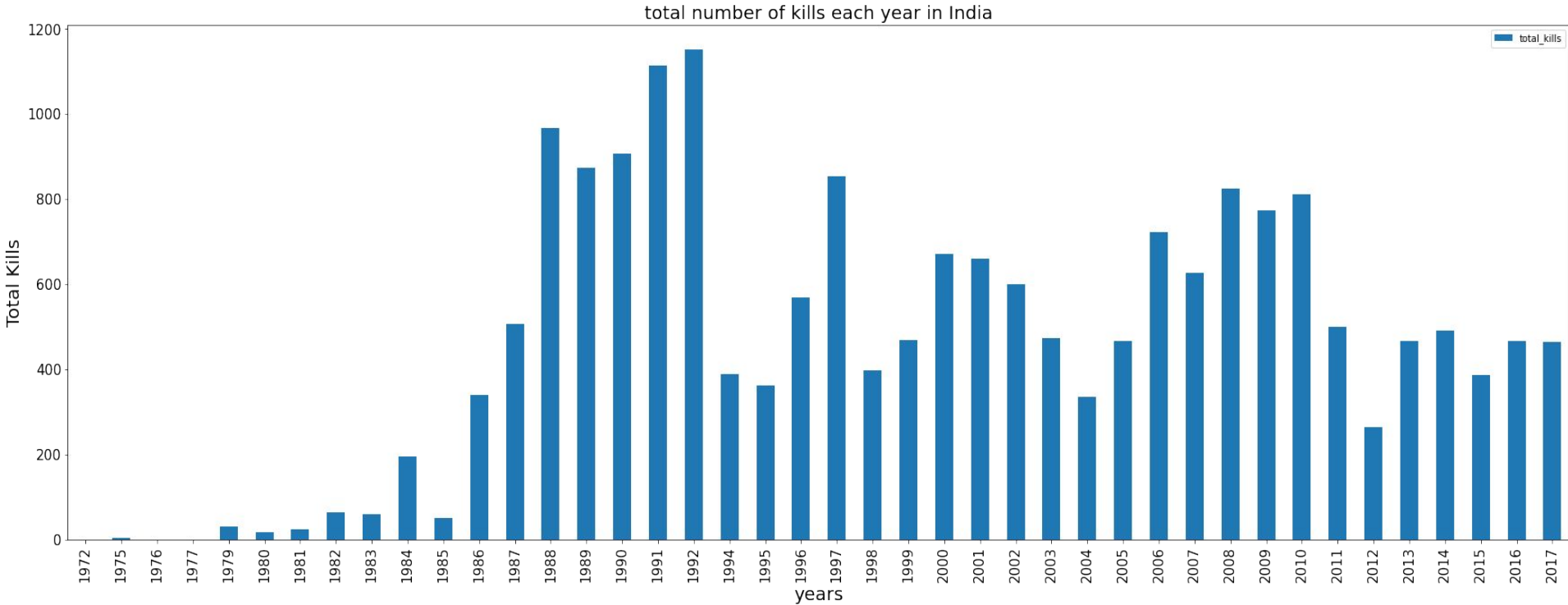


## No. of attack happened each year in India



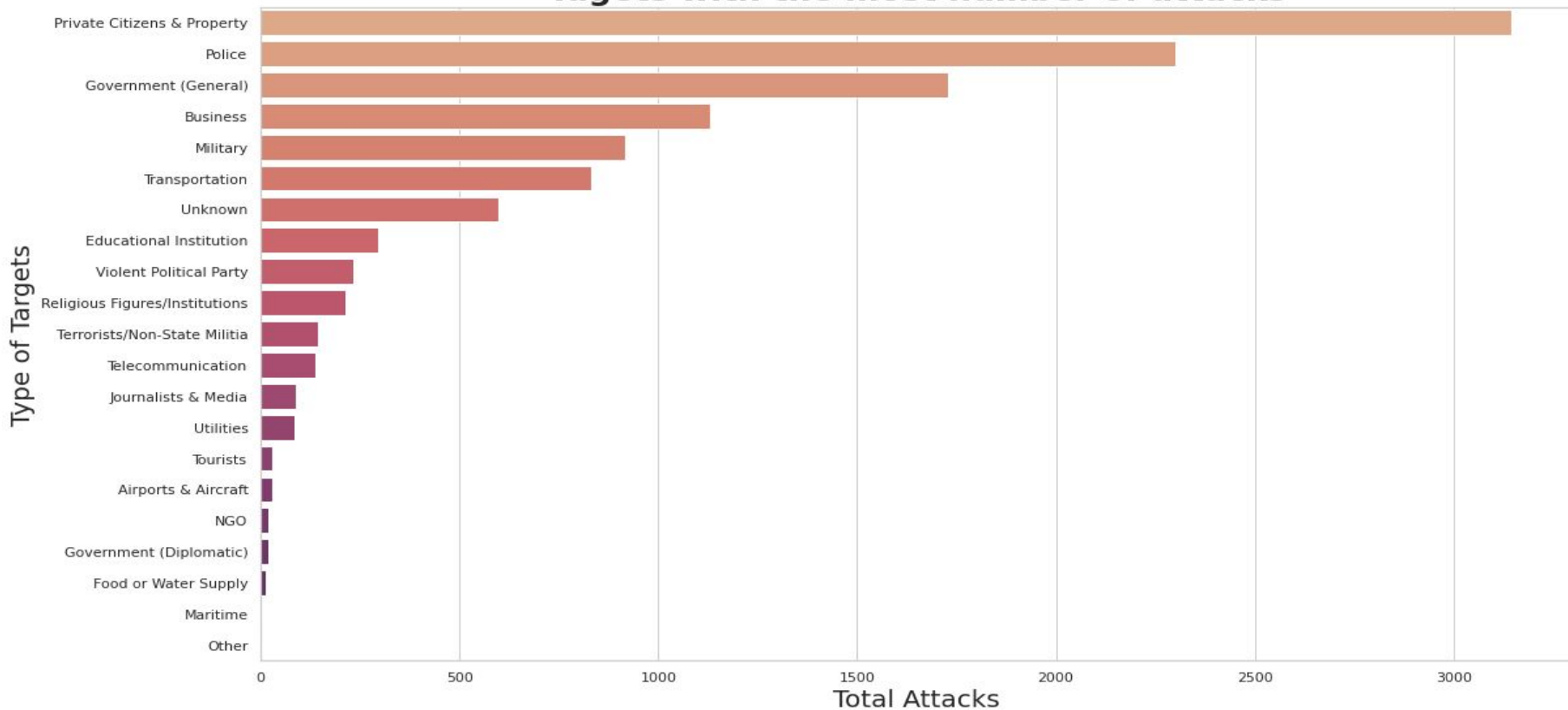
**Year 2016 has the highest number of attack in India with total number of 1025 attacks in a year.**

# No. of citizens killed each year in India



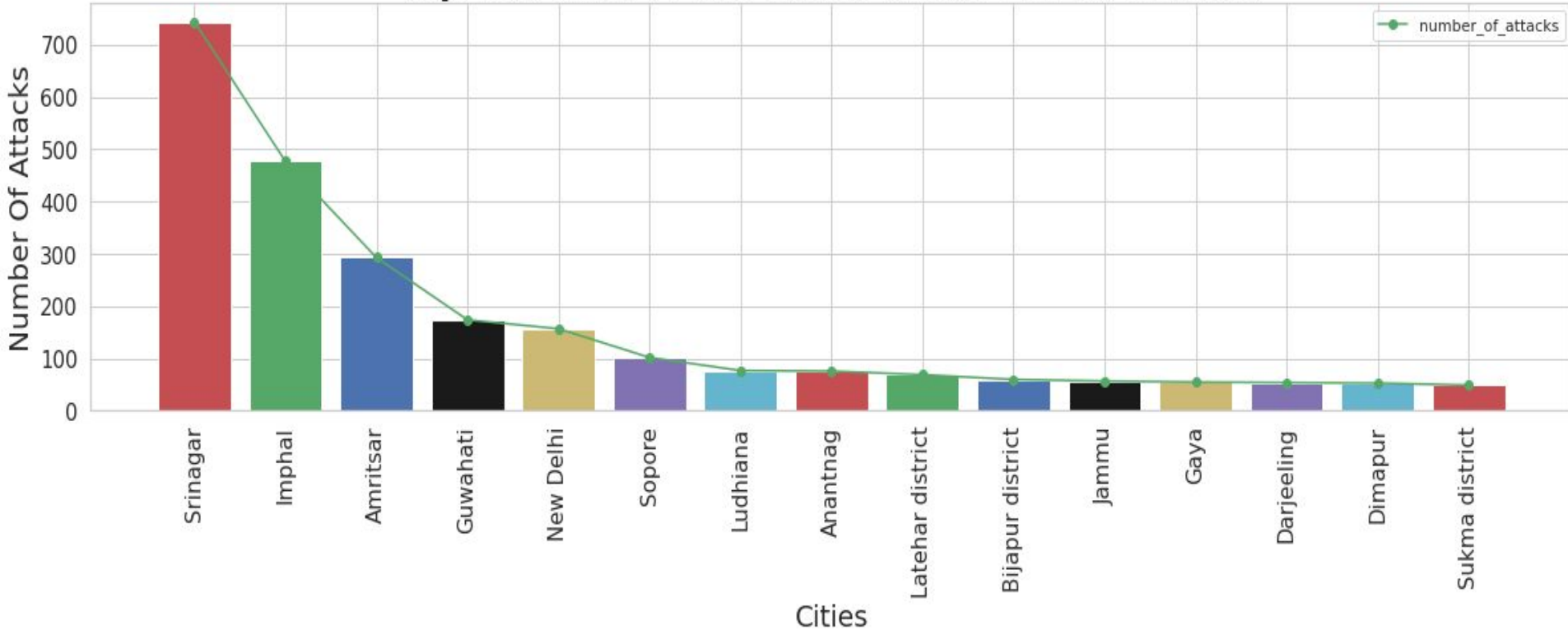
1992 was the year where the most number of citizen were killed with total number is 1152.

## Targets with the most number of attacks



**Private Citizen and Property are the top target with total of 3144 attacks.**

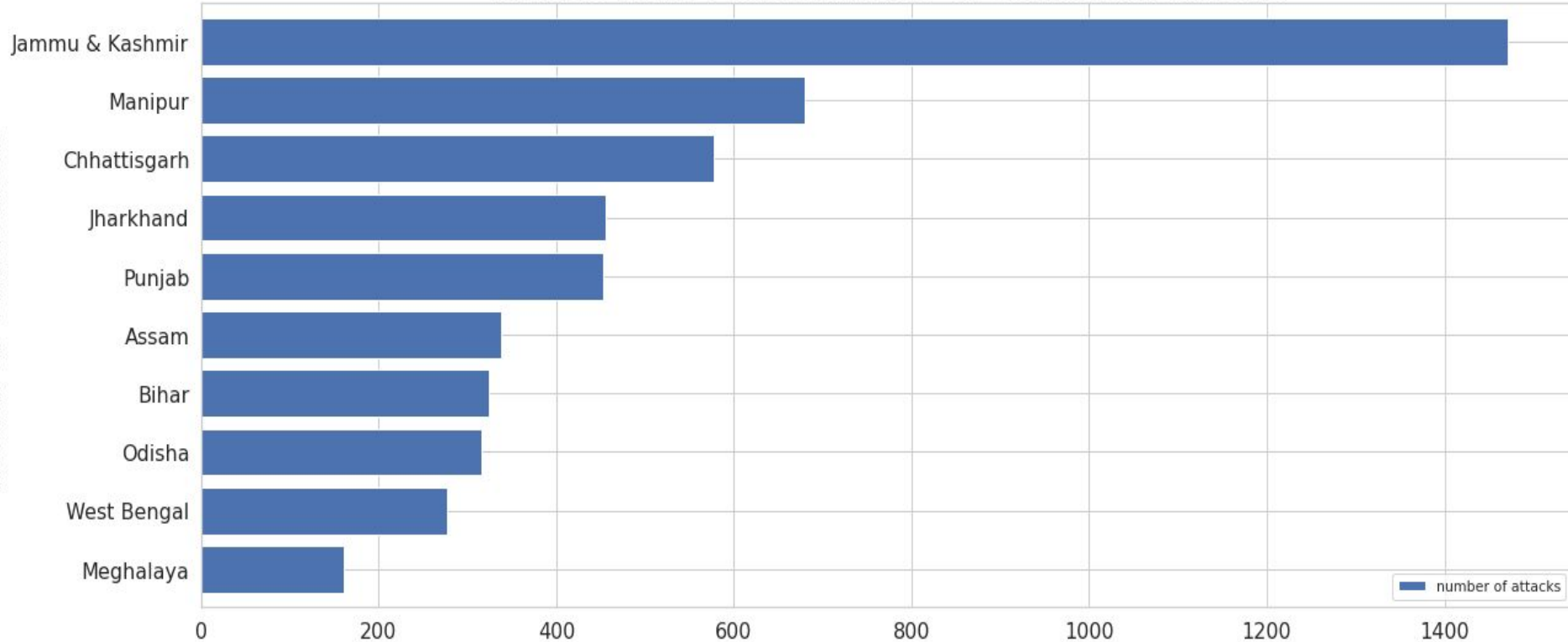
## Top Terrorist Prone Cities (1970-2017) in India



**Top Terrorist Prone City is Srinagar with total of 744 attacks.**

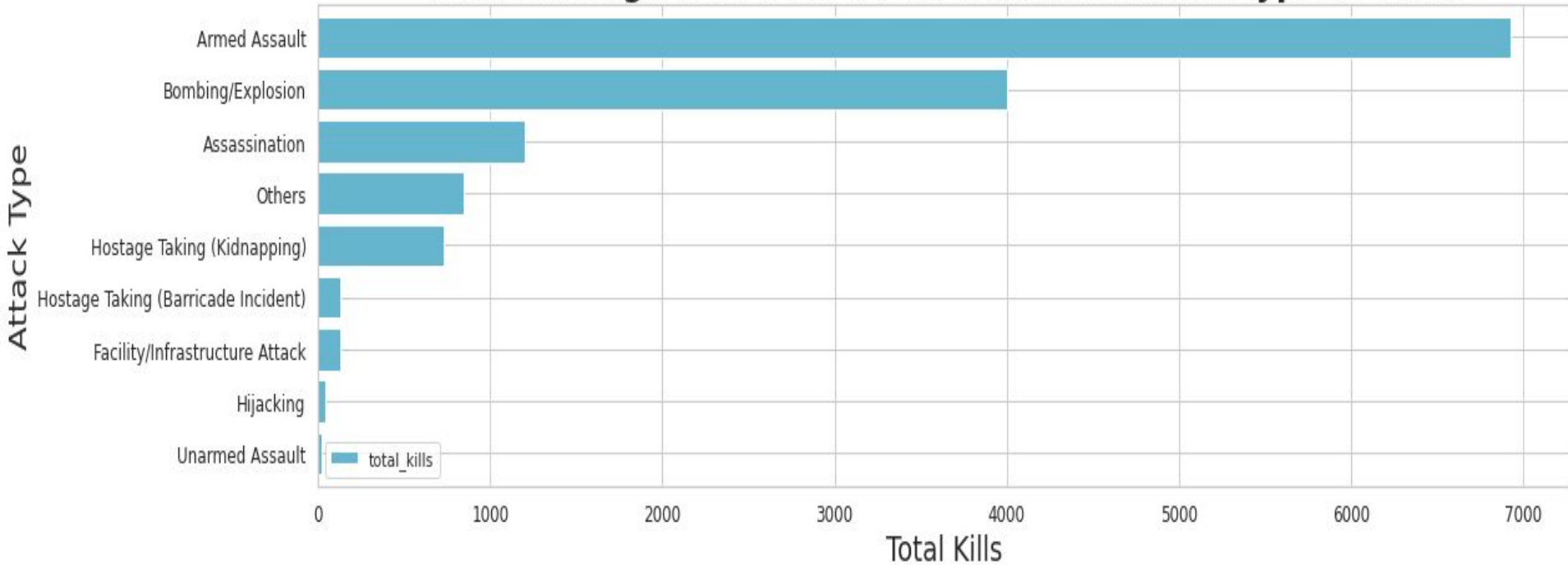
## States With Most Number of Terrorist Attacks

Number of Attacks



**Jammu and Kashmir has the highest number of attacks i.e 1471 attacks from the year 1970-2017.**

## Total Damage (1970-2017) vs Terrorist Attack Type in India



**Armed Assault has done higher amount(nearabout-7000 killings) of Damage**

# Challenges Encountered

- **Collecting meaningful data:** With so much data available, it's difficult to dig down and access the insights that are needed most.
- **Unknown Values:** Dealing with the unknown values in the data is a task in itself to analyze in the best way.
- **Visual representation of data:** To understand data it often needs to be visually presented in graphs or charts. It's difficult to build them manually. Taking the time to pull information from multiple areas is a time consuming.
- **Collate the work:** Integrating everyone's work into a single space took us a little more time than expected in a very well understandable way.

# Conclusion

1. Maximum number of attacks are 16903 across the world in the year 2014.
2. Taliban group has executed the maximum number of attacks i.e 7360 in the last two decades.
3. Asia is the most affected region by terrorist attacks with Iraq having maximum number of attacks.
4. Bombing/ Explosion are the most preferred attack type and the same did the maximum damage too.
5. 2016 has seen the maximum number of attacks in India but the maximum number of people were killed in 1992.
6. In India, J&K has been the most affected state and Srinagar having the highest number of attacks i.e 749.