Capstone Project

"Global Terrorism Analysis"

Abhishek Kirar (abhishekkirar27@gmail.com)

Ajinkya Shingote (shingoteajinkya65@gmail.com)

Ganesh Subramanian (gsubramanian5@gmail.com)

Mayur Chulbhare (chulbharemayur@gmail.com)

Syed Aquib (syedaquib153@gmail.com)

INDEX

OVERVIEW

- **This study examines the difficulties that the world encountered from 1970 to 2017. The Global Terrorism Analysis Project is something We are glad to provide here. The most crucial information regarding the terrorist operations occurring around the world is revealed by this research. Even though this study only includes data from 1970 to 2017, it nonetheless generated a lot of stimulating information regarding the current terrorism situation. Unwrapping the project and discovering the interesting data it has presented in a concise and understandable manner is extremely exciting.**

- **There is a lot of information about terrorism in the dataset, which is fairly extensive. We downloaded the complete Global Terrorism Database dataset, which is accessible via the GTD homepage. It requires 155 MB of disc space and has 1,81,697 terrorist assaults multiplied by 135 characteristics. Noting that it is virtually entirely encoded (strings/long numbers to small numbers) is important. We used the codebook found here to decode the dataset. We eliminated certain columns from the codebook after discovering they were unnecessary or redundant. For more information on our methodology, see the relevant notebook Cleaning Data.**

INITIALIZATION

The goal of this project is to analyze terrorist incidents that occurred worldwide between 1970 and 2017. These are the main goals:

- **To recognize and emphasize the geographic and temporal trends in terrorism**
- **To identify the key factors that contribute to a successful terrorist assault**
- **To give the user the most interactive ability to alter the analysis and study the data**

The project's goal is to understand how terrorism has evolved in the West and whether we should erect high barriers to defend ourselves from potential threats. We decided to focus on a more universal issue since It makes it possible to aggregate data at many different geographic scales, including the world, regions, nations, states, and towns. It has a wide range of interesting characteristics and is quite diverse. It includes geographical and chronological information.

DATA PROCESSING

- In order to accomplish the desired result, this study combines qualitative and quantitative research methodologies.
- Over 180,000 terrorist attacks have been evaluated in total.
- To find the most active and violent terrorist groups, we first do an exploratory data analysis to estimate the impact on a worldwide scale.
- We may be certain that the analysis accurately reflects the current state of affairs.
- Based on the threat level, we identify the major and minor epicenters (most vulnerable regions) and utilize descriptive statistics to determine the characteristics of each group across time and to investigate whether certain groups have a connection to the frequency of fatalities.
- The collection is quite comprehensive and contains a lot of information regarding terrorism.
- We downloaded the entire dataset of information. But our analysis ultimately focused on 21 columns, which contain both quantitative and qualitative data of fundamental importance.
- After the decoding, cleaning, filtering, and encoding operations, we have 181,691 rows x 21 columns.
- Our team spent the most of my time on 21 columns, which contain both numeric and qualitative information of primary importance.

EDA(Exploratory Data Analysis)

- Exploratory Data Analysis, often known as EDA, is the process of interpreting data sets by summarizing their key properties and frequently visualizing them.
- In EDA, plotting options include Histograms, Box plots, Scatter plots, and many more.
- Exploring the data often requires a lot of time.
- Through the EDA process, we can request that the problem statement or definition be defined for our data set, which is highly significant. EDA must always come first when conducting a data analysis.

- Using exploratory data analysis, we can:

- To shed light on a piece of data
- Recognize the fundamental structure
- Extract significant characteristics and the connections that exist between them
- Test the underlying presumptions

EDA FOR GTD

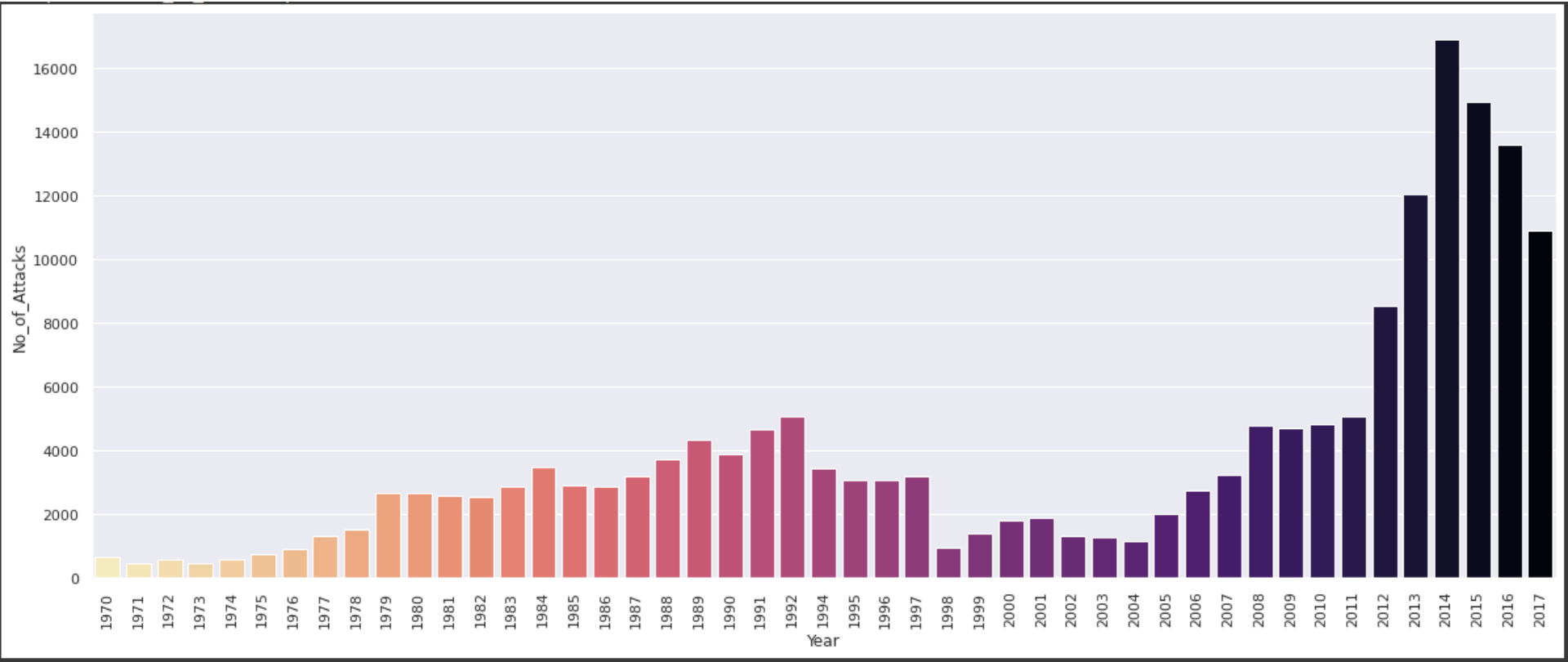
- **In order to identify the hot spot of terrorism, we will conduct exploratory data analysis on the terrorist dataset in this article.**

- **Exploratory data analysis is nothing more than examining the provided data to look for patterns, trends, and assumptions.**

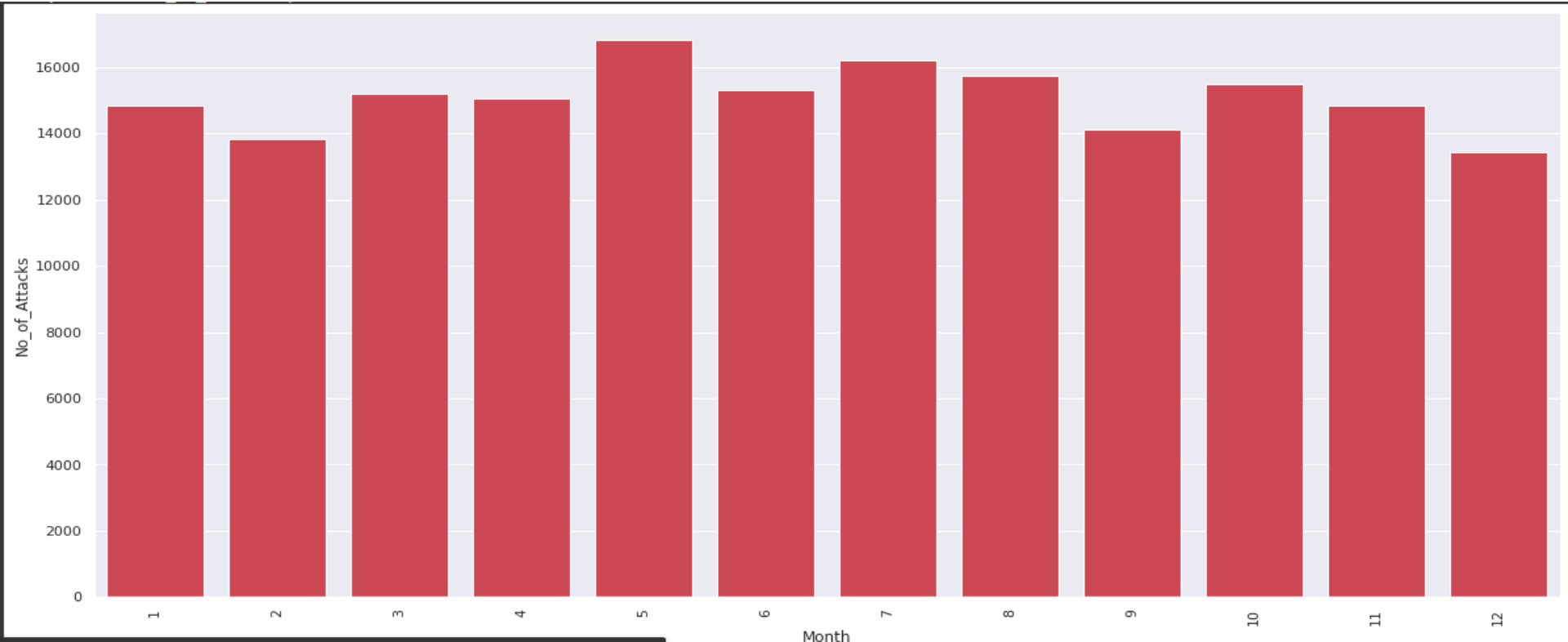
- **There are numerous attributes in this dataset, including years, months, days, gang names, weapon types, target categories, states, regions, nations, and many others.**

- **To better comprehend the Datasets we are given with, we have visualized a variety of topics from the Global Terrorism Database. The Global Terrorism Database's varied characteristics from 1970 to 2017 have been shown graphically. The graphs produced by the GTD after the contents were visualized based on assaults by year, fatalities by year, countries by total attacks, and attacks by kind are shown below.**

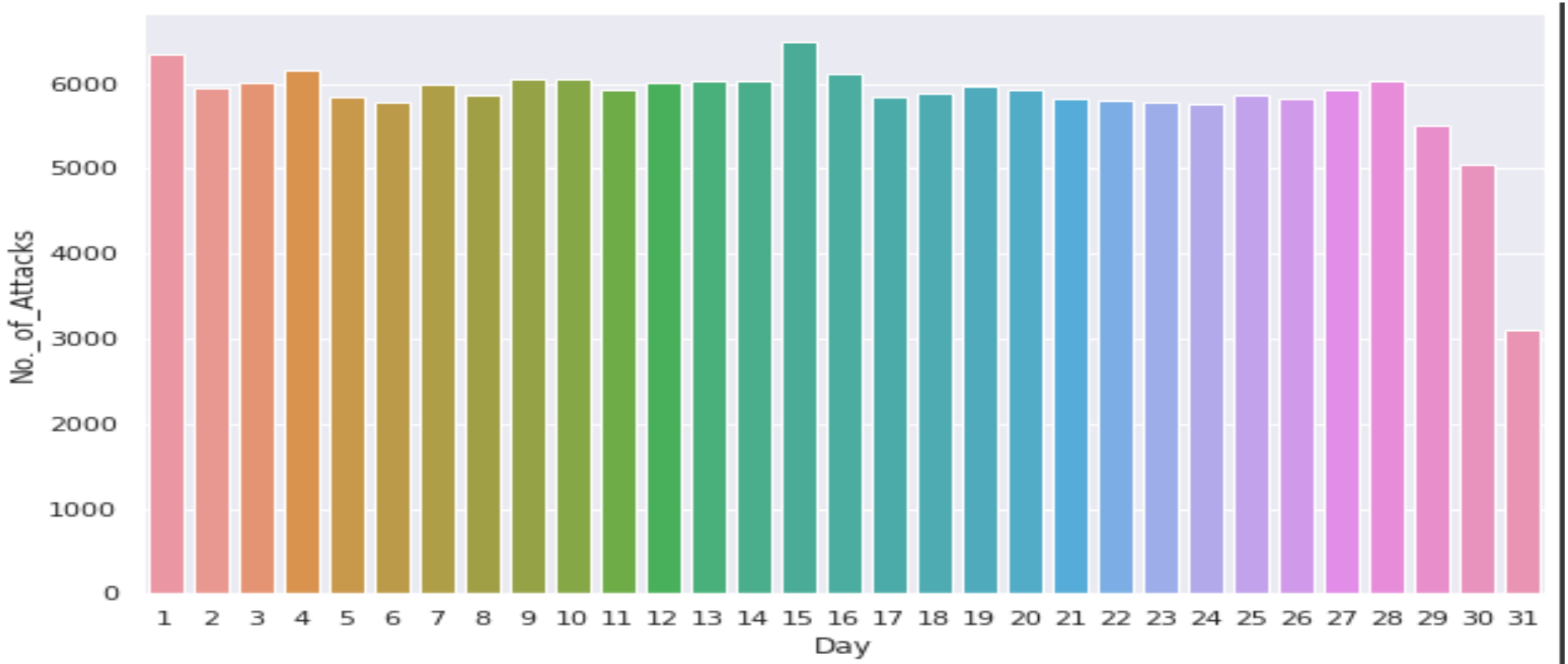
From 1970 to 2017, the number of attacks is depicted in the bar diagram below. From 2011 to 2014, there were more attacks, and then they gradually began to decline, so we may draw the conclusion that they were on the rise. The most attacks were recorded in 2014 as well.



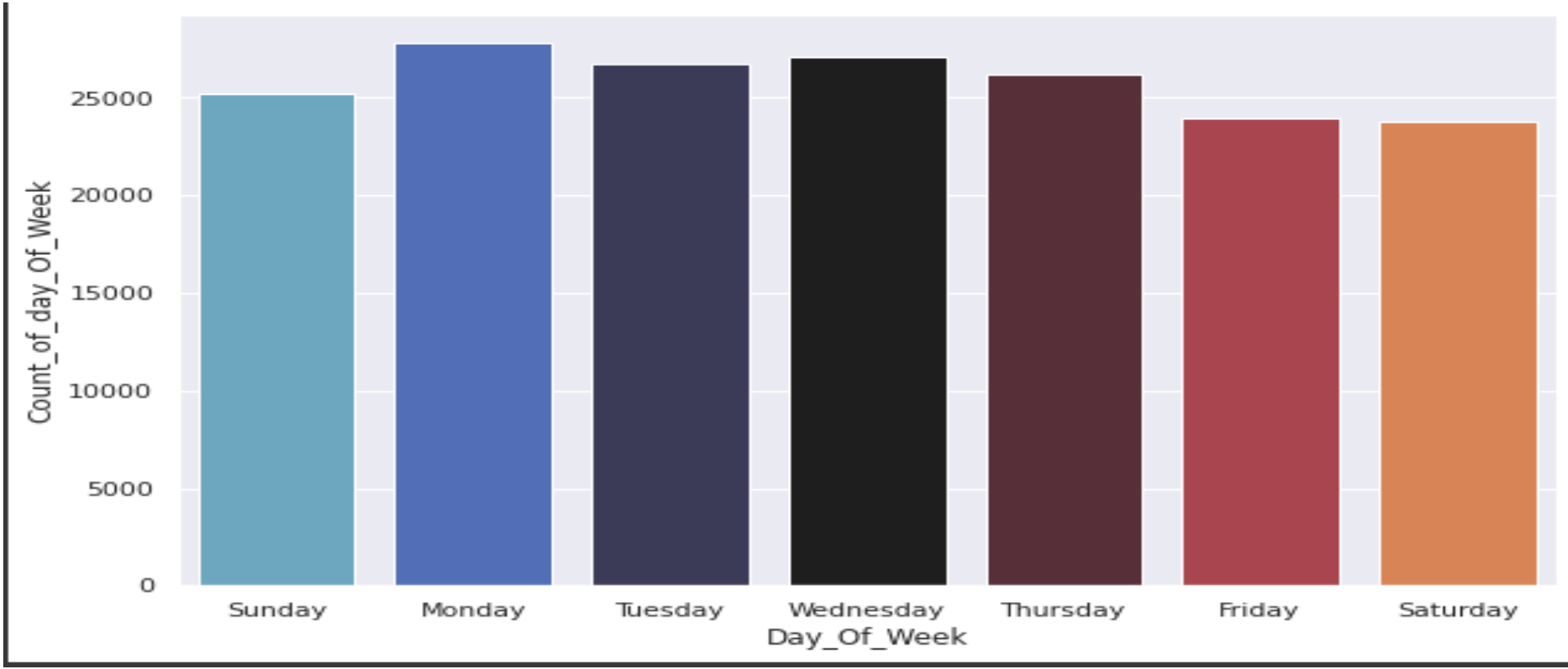
The Bar graph shown below is depicting no. of attacks occurred in the different months through out the year. From the below diagram we can visualize that in month of May and July more attacks occurred as compare to other months of the year.



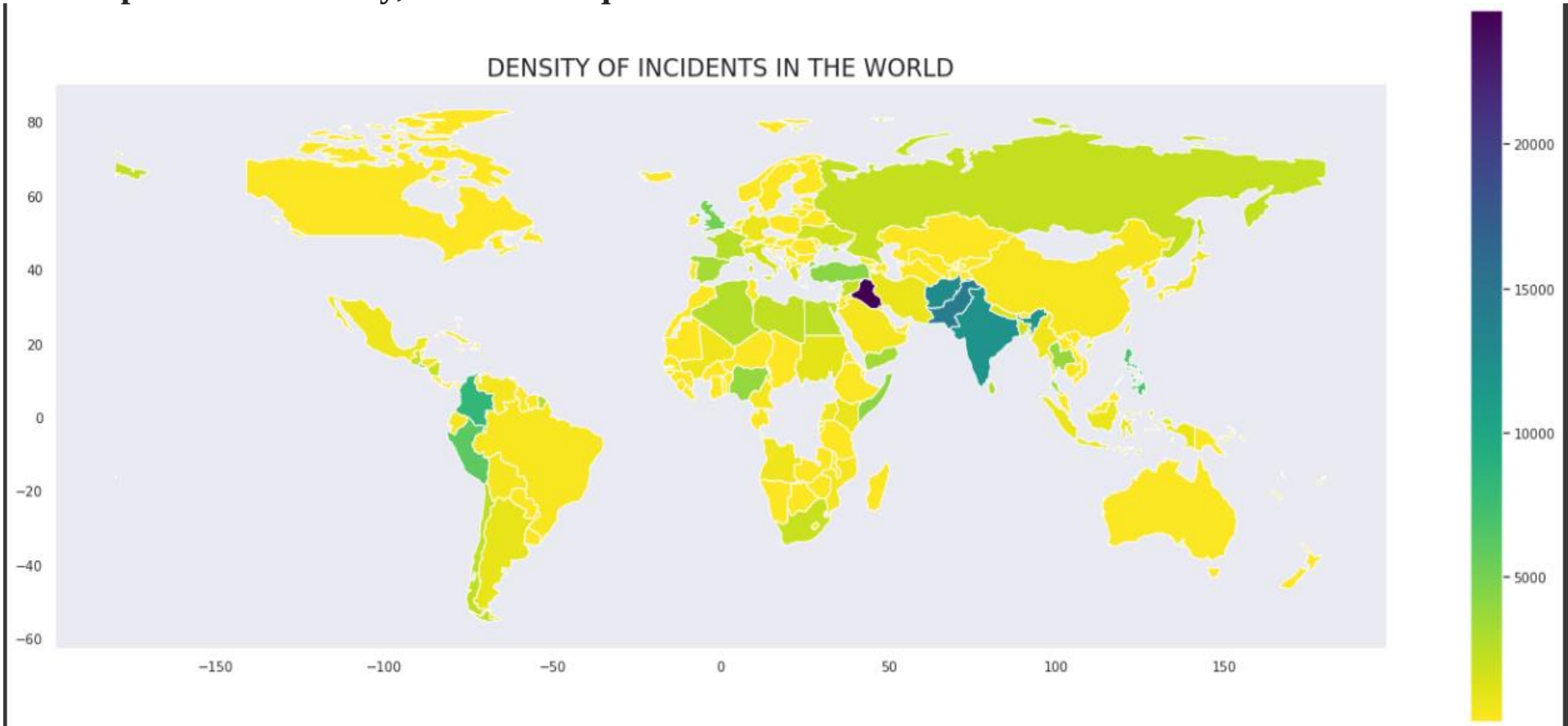
The Bar graph shown below is depicting no. of attacks occurred on the different dates through out the month. From the below diagram we can visualize that more attacks occurred on start and mid of the month as compare to any other date of that particular month. At the end of the month no. of attacks were the lowest.



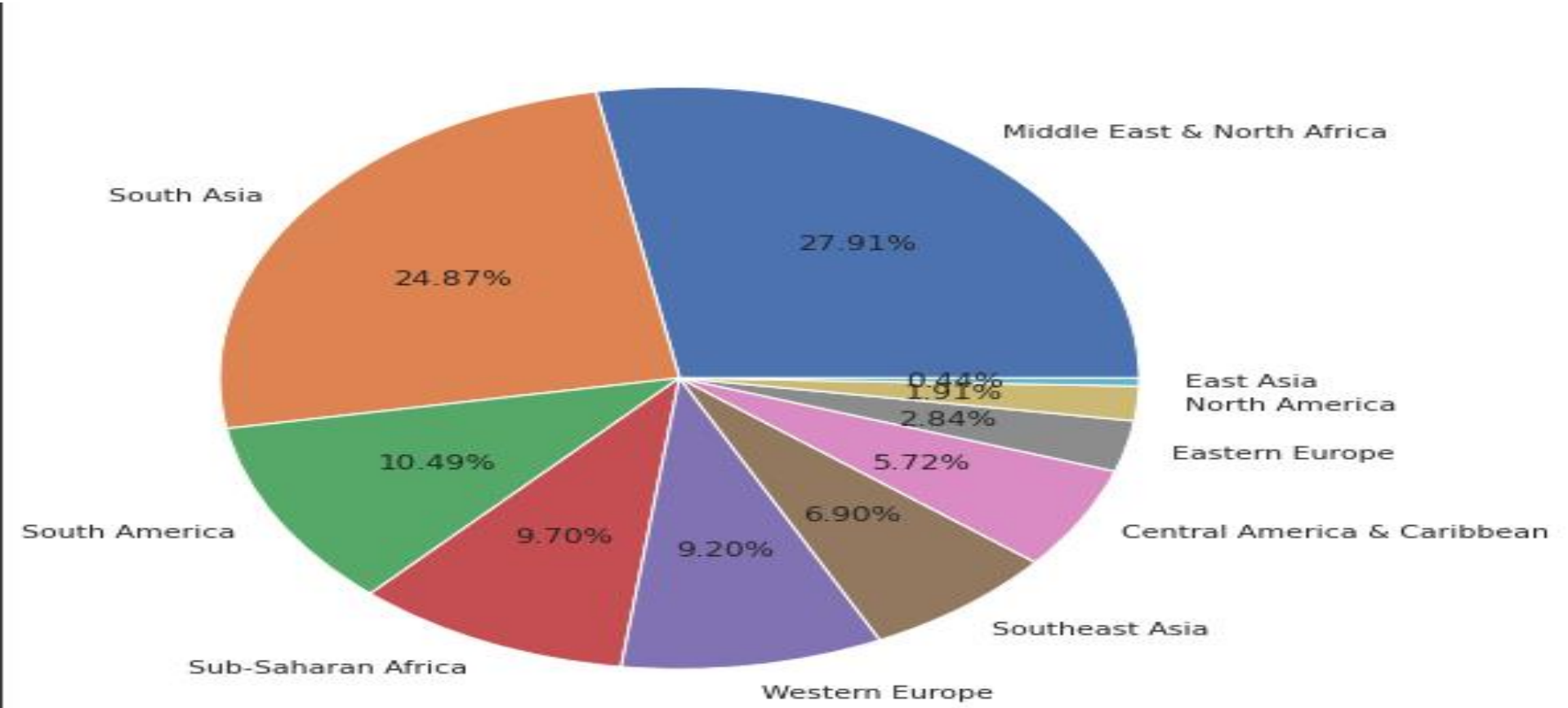
The Bar graph shown below is depicting that more no. of attack occurred on weekdays as compare to the weekend and on Monday.



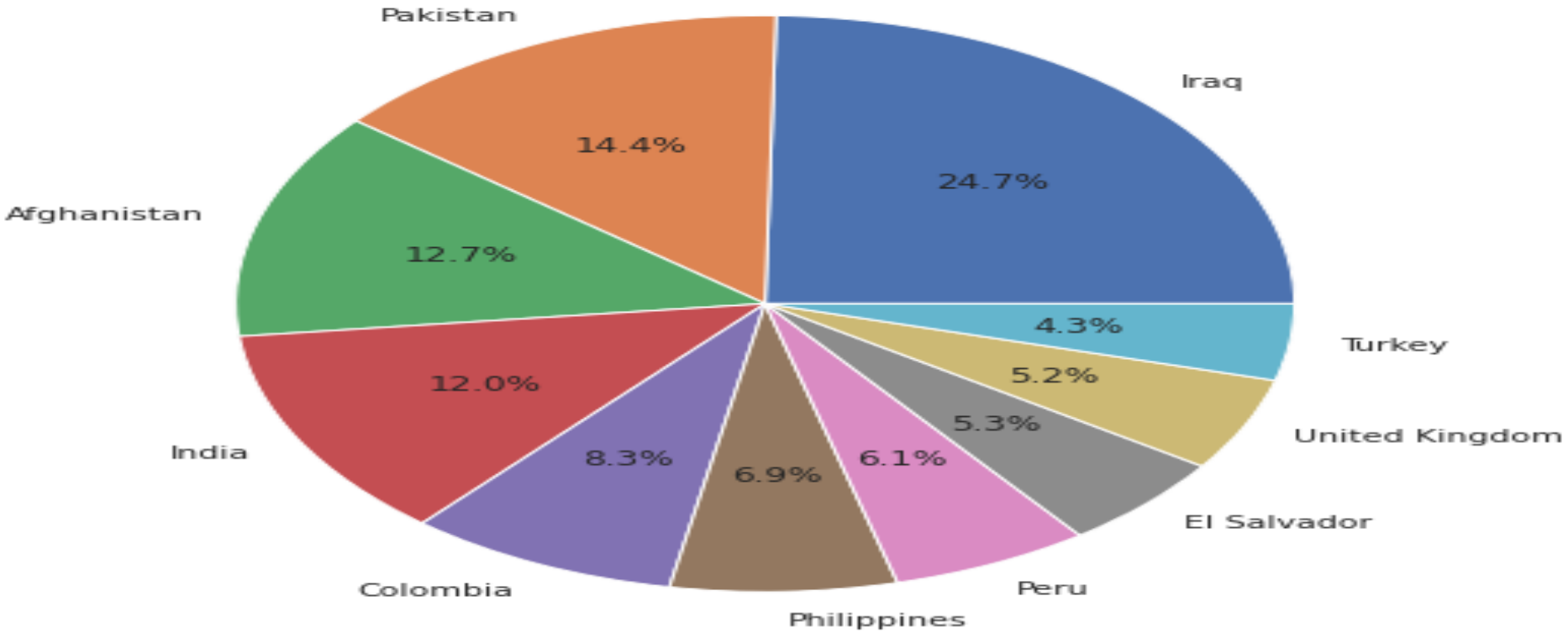
The darkest spot shown in the below map is depicting that the most no. of terrorist activities occurred in that particular country, which is Iraq.



The Pie chart shown below is saying about the top 10 affected regions by the terrorist activities.

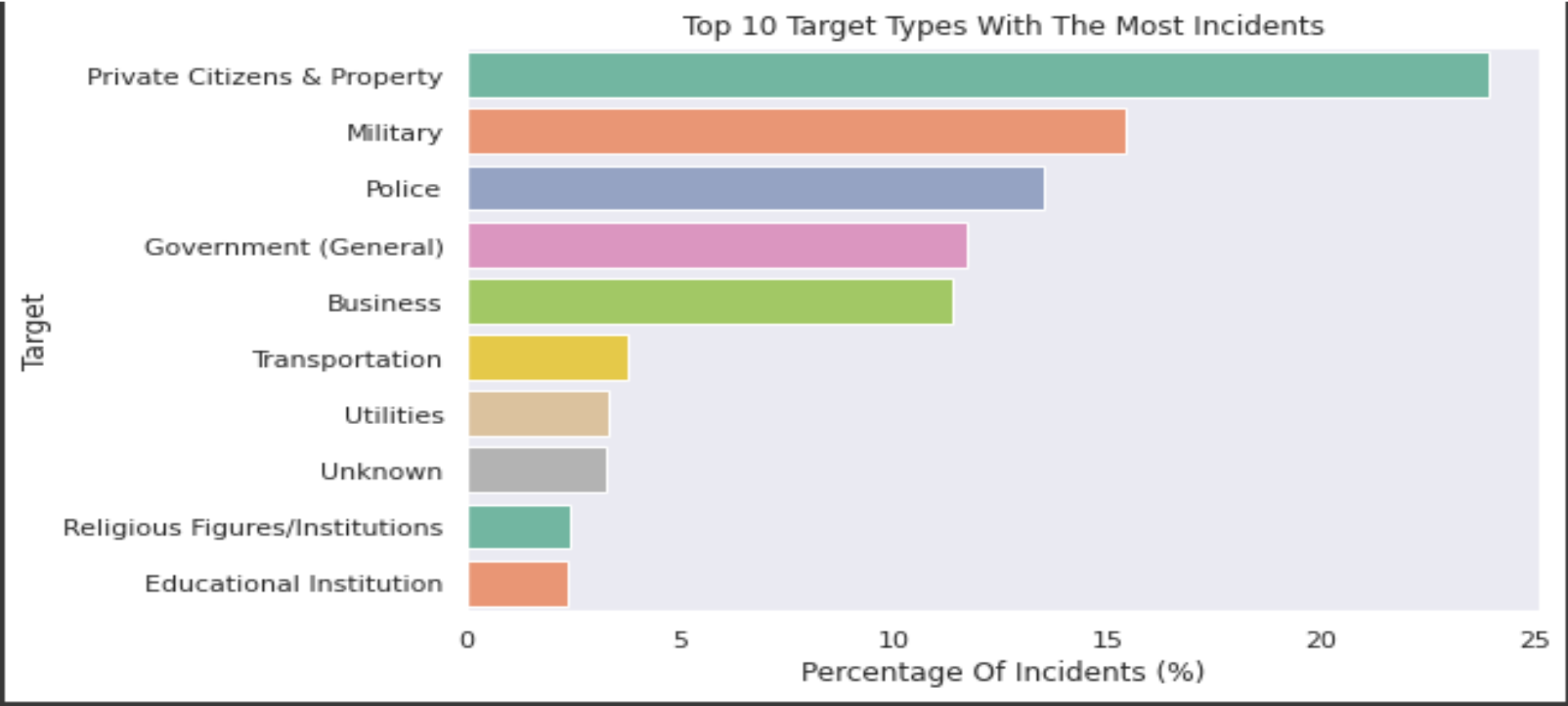


The Pie chart shown below is saying about the top 10 affected countries around the world by the terrorist activities.

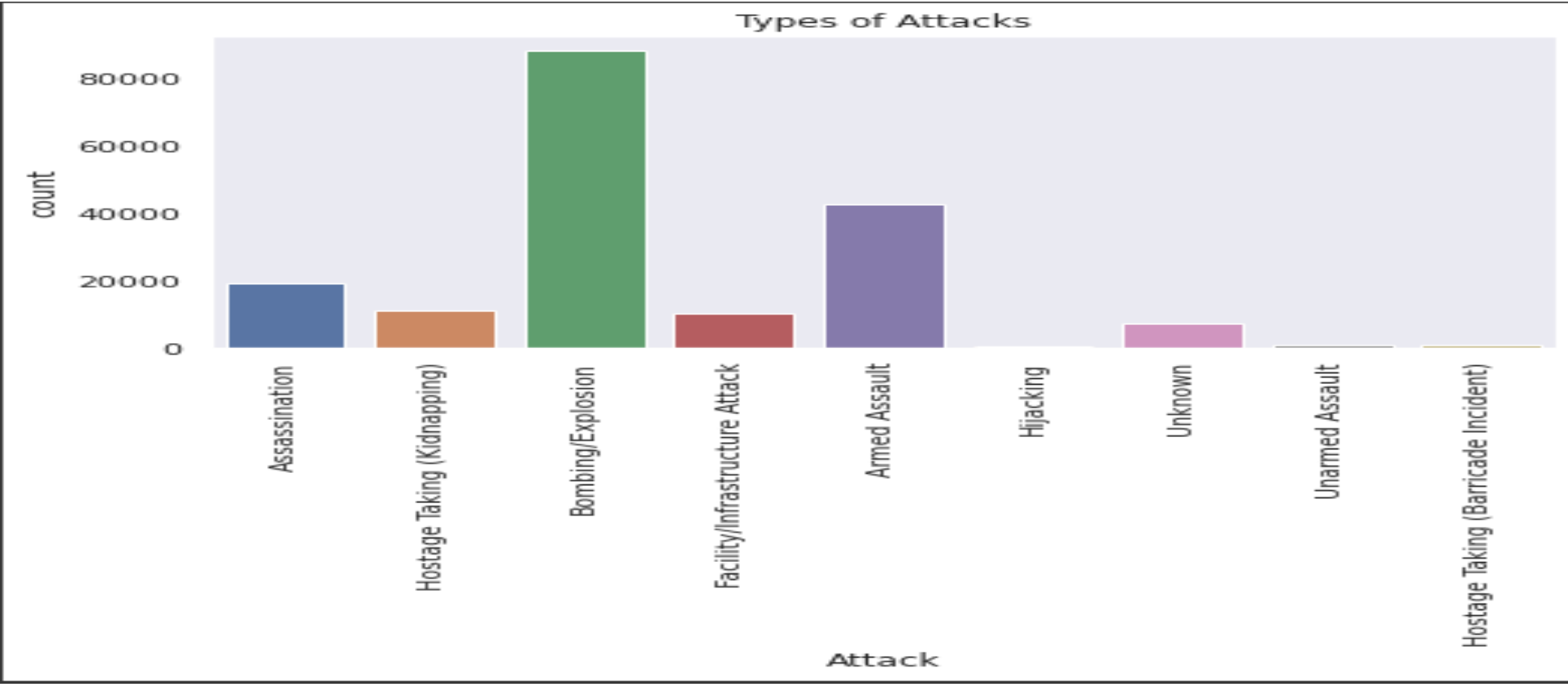


The picture below is depicting the different purposes behind any attack performed by terrorist group or by any individual. But in most of the attacks there were no specific motive behind this sinful and violent act.

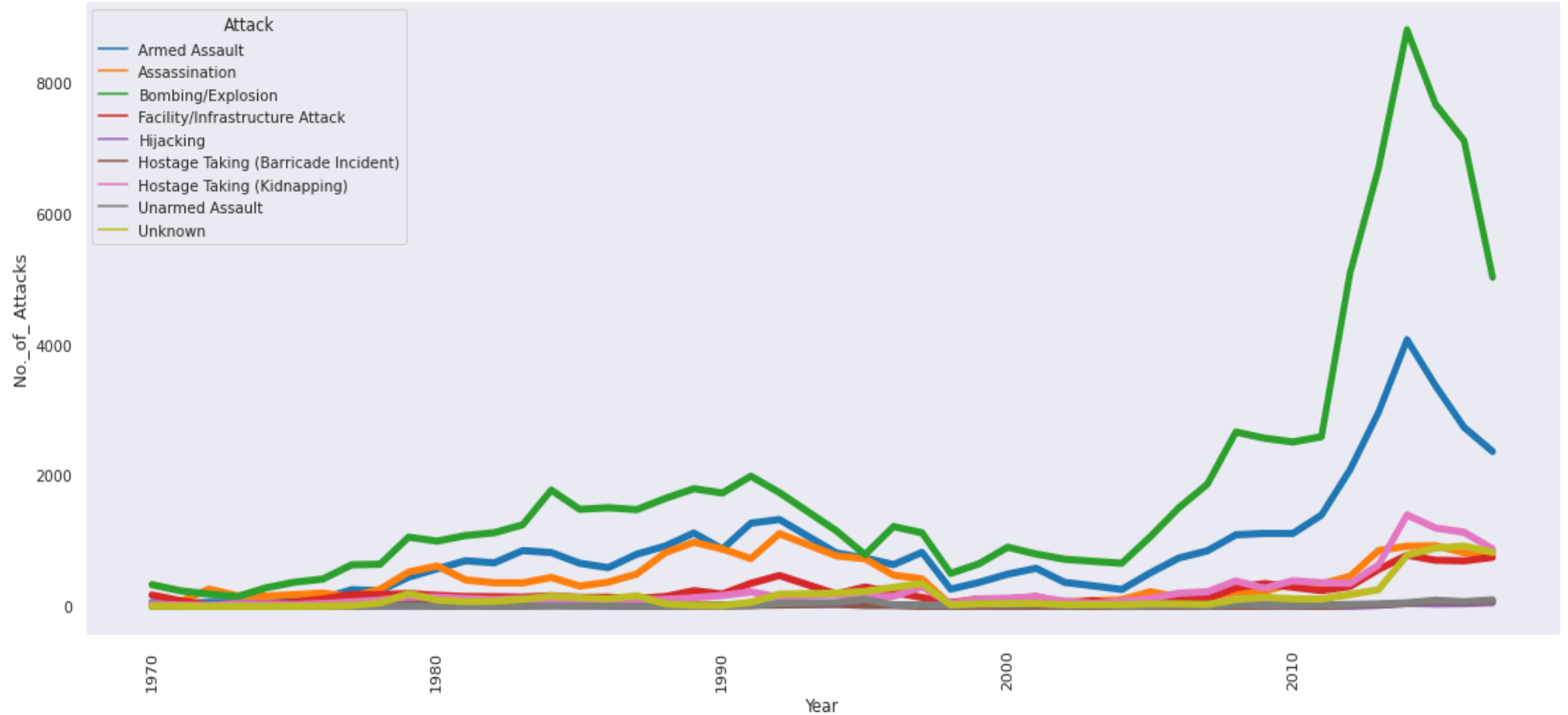
Below graph showing the top 10 targets of terrorists in which Private citizens and properties is on the top of the list.



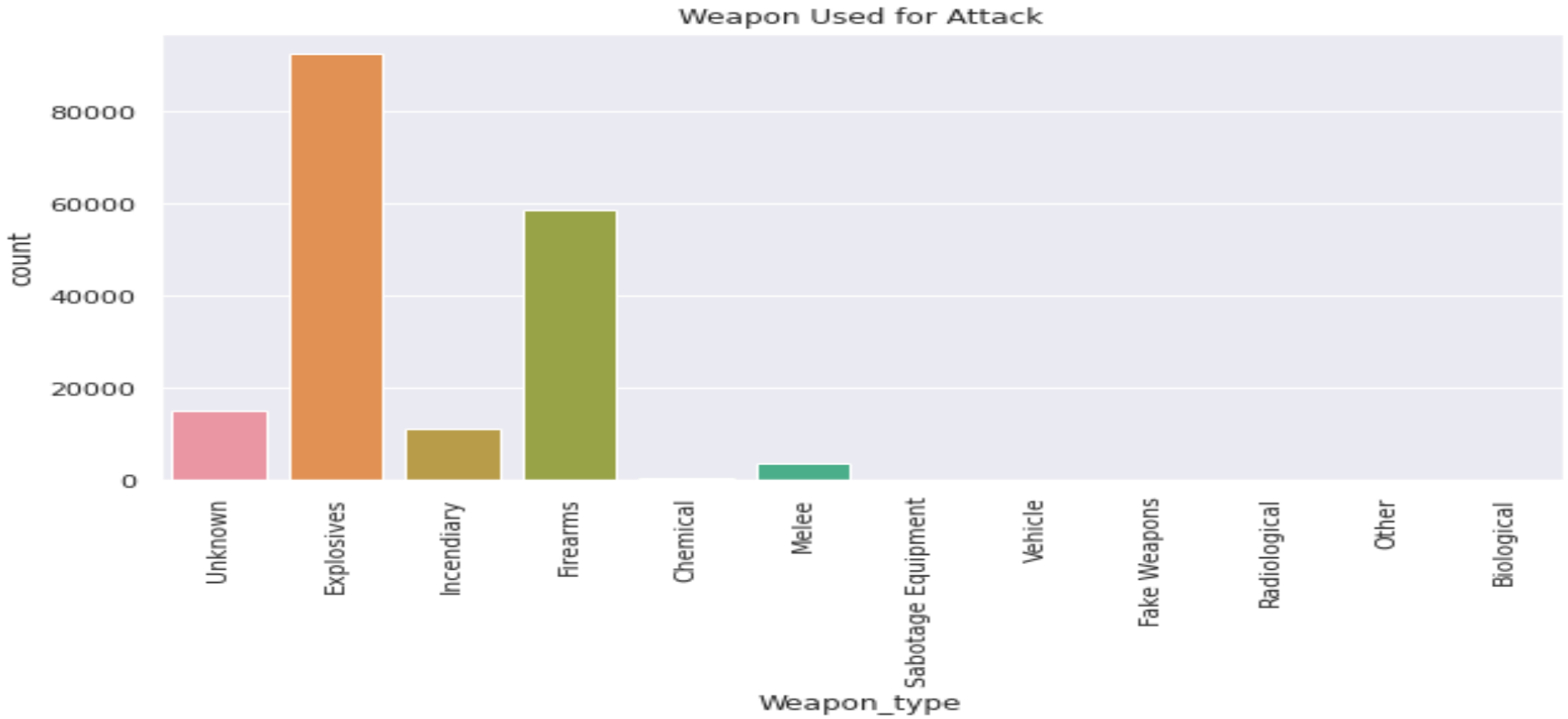
The total attacks are shown by kind in the below bar diagram. Armed Assault is clearly the second most popular attack type, behind bombing and explosions, based on the overall number of attacks. Attacks over the years 1970 to 2017 totaled more than 100,000.



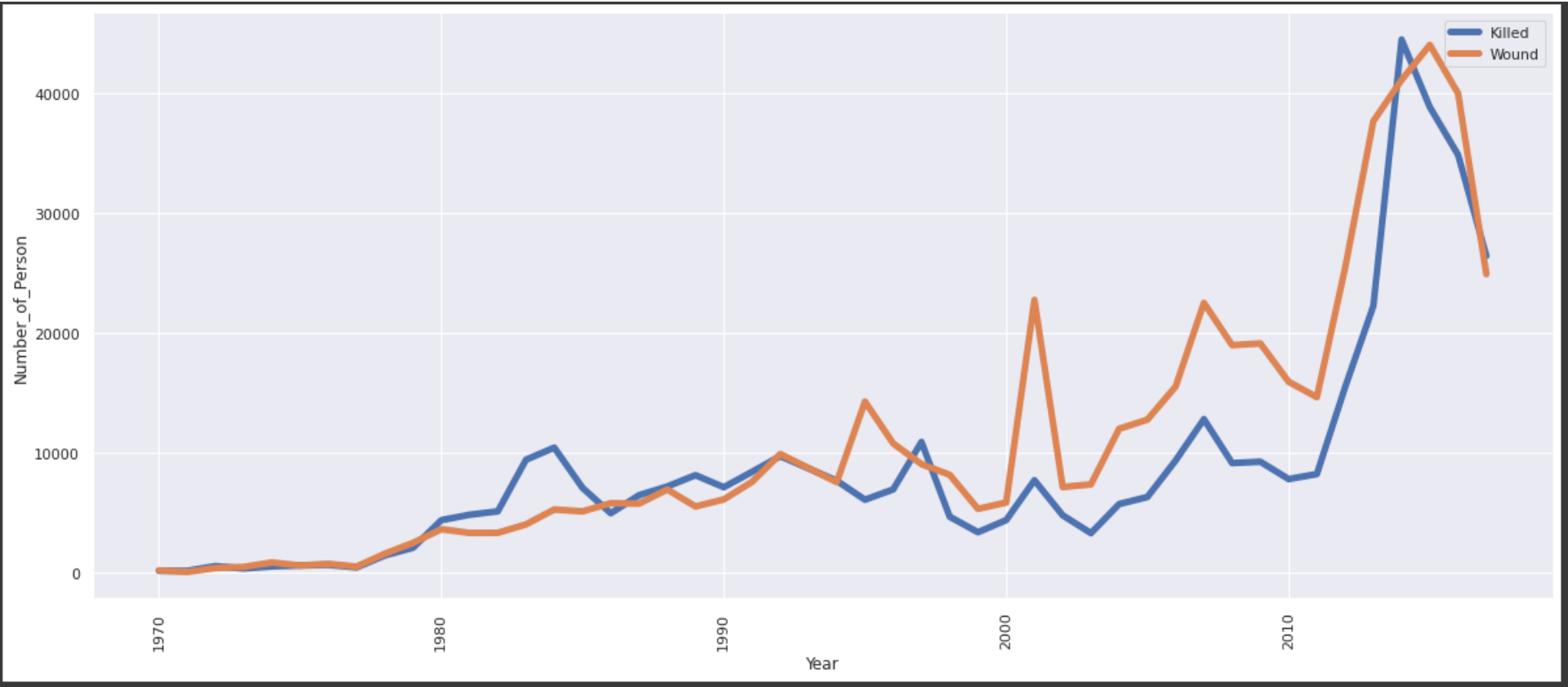
The types of attack over the years are shown in the below bar diagram. There were sudden rise in bombing and armed assault during the year 2010 -15 and then there were also a sudden fall as compare to the other types of attack.



Attacks between 1970 and 2017 are categorized in the graph above according to the type of weapon used. Undoubtedly, the most popular weapon kinds are explosives and firearms.



Injuries and fatalities for each year between 1970 and 2017 are depicted in the above bar graph. The years 2011 to 2017 clearly saw the highest number of injuries and fatalities.



CONCLUSION

- Upon completion of EDA using various types of graphs, the majority of attacks used explosives to initiate and weapons to finish
- The number of attacks peaked in 2014, then fell in 2015
- When compared to attacks beginning in 1970, the most recent six years had the greatest rating
- But starting in 2014, the number started to decrease. Attacks peaked on start and mid of the month but very low during the end of the month
- Almost every day makes the same amount of contribution
- Iraq tops the list of countries with the most attacks, followed by Pakistan, Afghanistan, and India
- South Asia comes ranked second among all the regions, followed by the Middle East and North Africa.
- The military comes in second place, with the majority of attacks going towards private persons' property

CHALLENGES

- Recognize the importance of each column
- Choose the proper graph to show the varying patterns throughout
- Considering a crucial analytical query
- Discover a strategy to consider and address the crucial question
- A few code and mapping issues
- It was necessary to manage a sizable amount of data while being careful not to overlook anything, no matter how irrelevant
- High calculation time

END-NOTES

- By using a dataset, we have learned how to conduct exploratory data analysis
- In this project, we examined the terrorism data set to determine the country's hotspots
- Finally, we discuss visualizations
- I sincerely hope that this essay on exploratory data analysis was helpful to you