

# 36-315 Final Project, Global Terrorism

Team 13, 2020 Spring

Jingyan Xu, Melody Ma, Xinzhe Qi, Yurong Li

May 13, 2020

## Abstract

Terrorism constantly poses threats across the globe. In this project, we analyzed the global terrorism database from Kaggle and applied various data visualization techniques to provide a clear, detailed, and accurate insight into the geographical and temporal distributions of global terrorism. We also examined various variables that are highly related to terroristic attacks, such as weapon type, target type and casualties, and found patterns in them. At the end of this project, we looked closer at India and identified the country's most active terrorism groups and analyzed their behaviors.

## 1. Dataset

The Global Terrorism Dataset(1) consists 181691 rows and 23 columns. Each row corresponds to the information of a terrorism incident that happened around the world. The dataset records terrorism incidents dated from 1970s to year 2017. The main variables we used are as following:

- Location (categorical) is accurate to name of the city, village, or town in which the incident occurred.
- Time of event (qualitative) is accurate to the day in which the incident occurred.
- Type of attack (categorical) is the general method of attack, such as assassination, bombing, hostage etc.
- Type of weapon (categorical) is the main weapon used, such as Firearms, Explosives etc.

- Motive (text) is a column explaining the cause of the event, such as intimidate the police, protest Vietnam War, fight imperialism etc.).
- Ransom (text) is a column explaining any information about non-money demands made by perpetrators, as well as information on conflicting reports of how much money was demanded and/or paid.
- Other interesting variables we looked into include, longitude & latitude (qualitative), attack target (categorical, such as Government, Police etc.).

## 2. Problems

The main research project problems we would like to learn are:

- Which areas are more susceptible to terrorism in the world? Do regions that are vulnerable to terrorism change over years? *See 3.1*
- How does the target type change over time? How do the weapon type and target type distributions look like? Are there any specific weapons that are used for a specific target type? *See 3.2*
- Given the historical data from 1970 to 2017, are terrorism attacks becoming more likely to happen today? *See 3.3*
- What are the common motivations and demands in the terrorism attacks? *See 3.4*
- Case Study: What are the most common terrorism groups and attack locations in India? How do these attacks' severity vary across India? What is the scale of the attack in terms of the number of the terrorists involved? *See 4*

## 3. Methods

### 3.1 Spatial Data Visualization

In order to investigate the global distribution of suicide attacks, we choose to draw an interactive graph, where you can zoom in to see the exact locations of these attacks, or zoom out to see the

aggregated numbers in some larger regions, like Middle East or Southeast Asia. Since the interactive version is only available in html, this report is going to focus on some screenshots from the map.

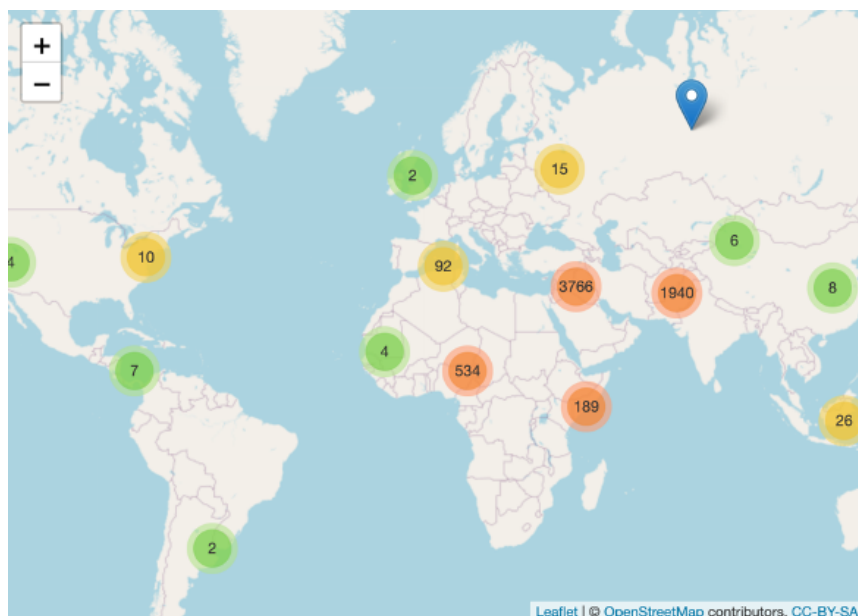


Figure 1

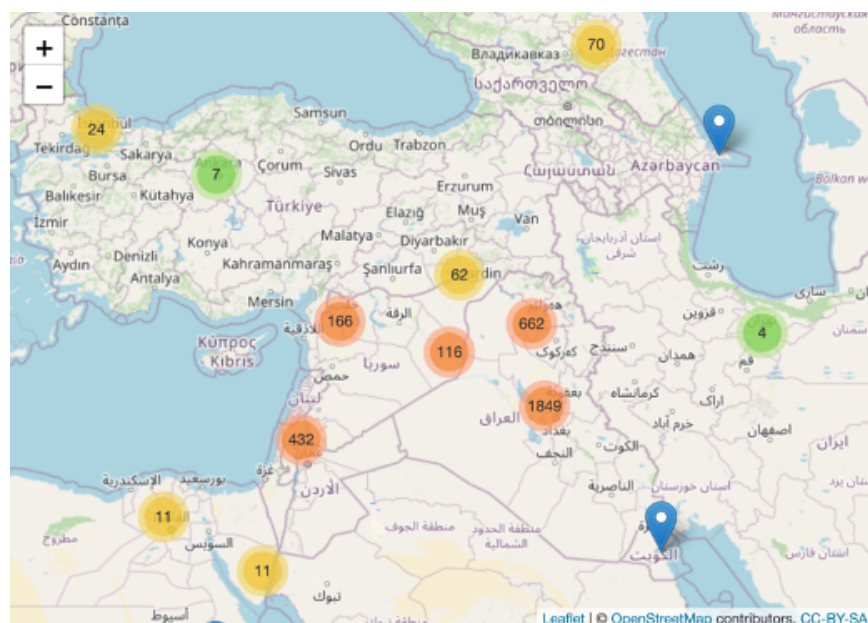


Figure 2

Figure 1 is a general overview of the global distributions with regional aggregations. The colors and numbers here indicate the frequencies of suicide attacks. It can be seen that large numbers of attacks center around the areas of Middle East, and North India/Pakistan. If we zoom in to

see Middle East area (see Figure 2), we can observe that most cases in the Middle East happened inside the borders of Iraq. Moreover, Syria and Israel also contributed a large portion of attacks. In contrast, regions like East Asia and Western Europe appears to have much lower cases than other places.

In order to have an idea of how terrorism changes over time in each region, we plotted a time-series graph counting the number of terrorism attacks in each year.

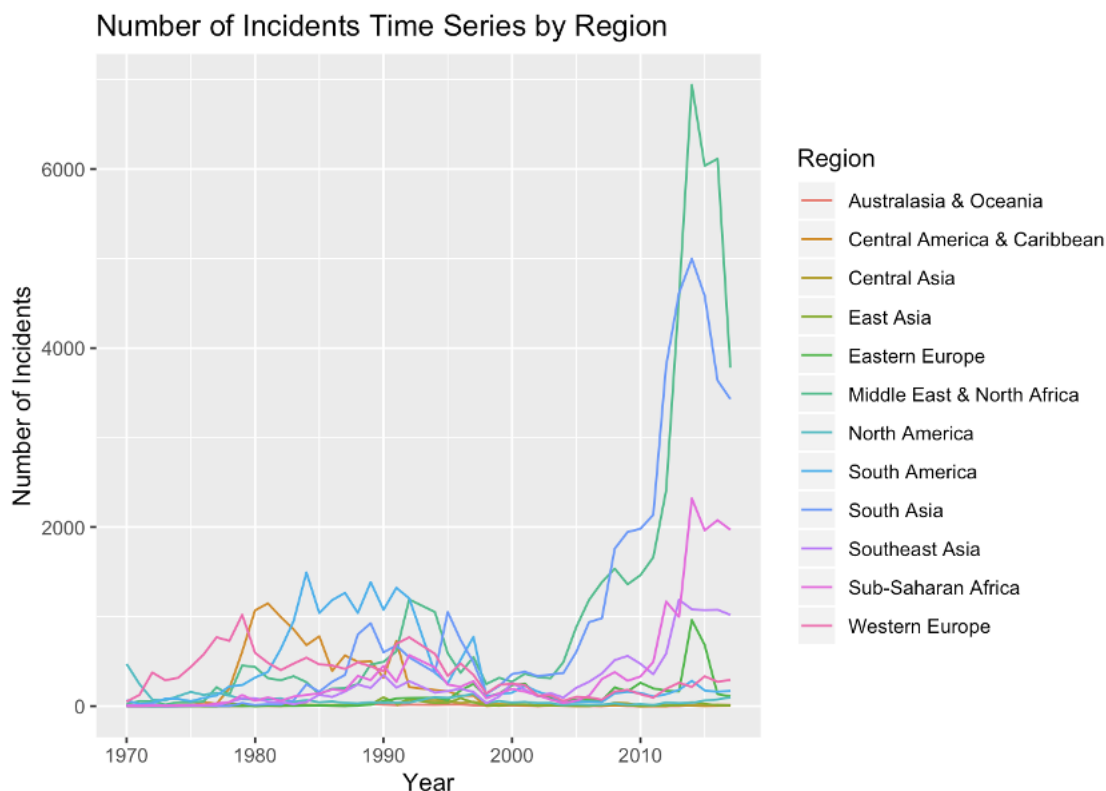


Figure 3

We find that overall, there is an increasing trend in the number of incidents by year in the world. After the year 2000, the number of terrorism events rapidly increases in the Middle East & North Africa, South Asia, and sub-Saharan African. During 1980 and 1997, South America has the most terrorism events. The number of terrorism events in all regions decreased to the lowest points around 1997. After the year 2015, terrorism events decreased rapidly in the Middle East & North Africa and South Asia.

### 3.2 Time Series Modeling

After getting the big picture of terrorism, we want to find whether the target type vary in the terrorist attacks from 1970 to 2017. If so, are there any generalized trends?

First, we created a four-facet plot of seasonal decomposition showing observed, trend, seasonal and random components for the target type. We decided to use monthly data to plot time series, trying to figure out if there are some patterns based on the seasons. Also, to make the graph clearer, we focused on the top eight target types that have the most number of incidents.

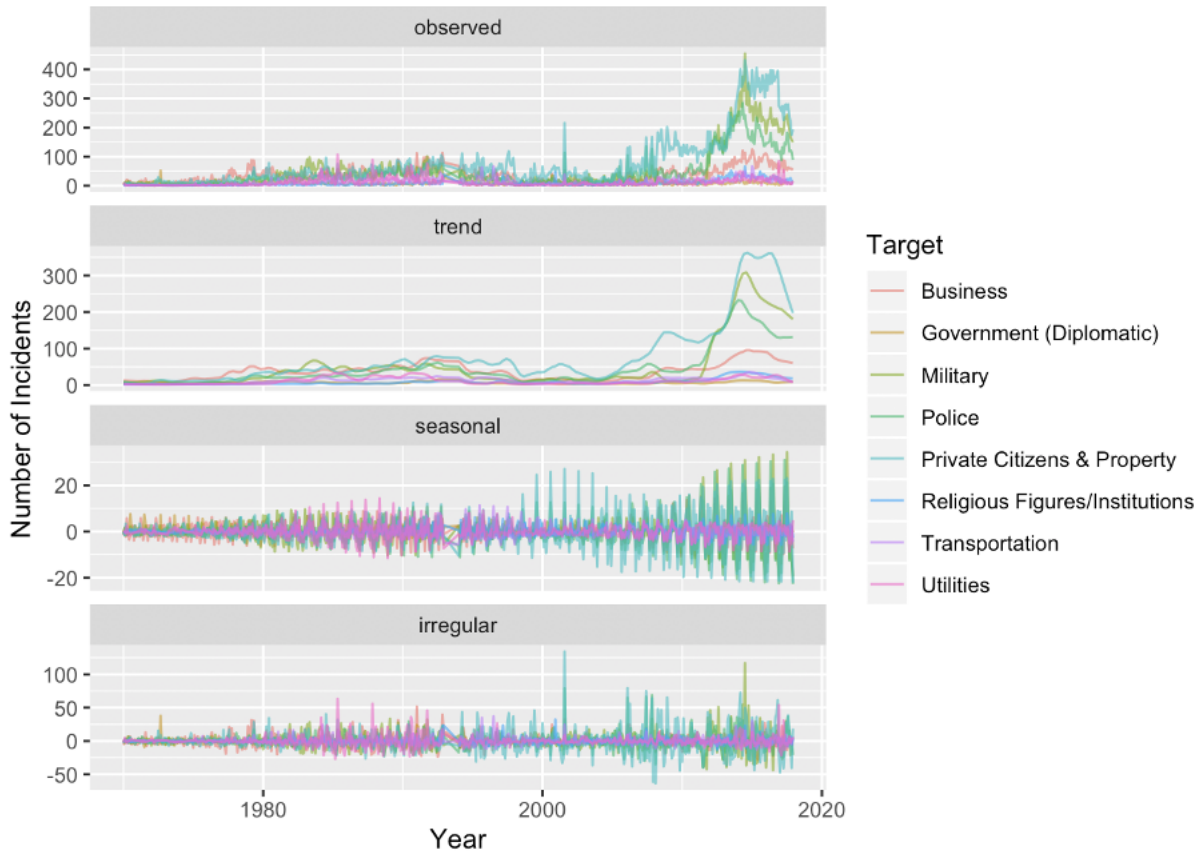


Figure 4

We have the following observations:

- From the trend plots, we find that in general, the number of terrorism for all target types is increasing. Also, there are peaks in the number of terrorism events among ‘Private Citizens & Property’, ‘Military’, and ‘Police’ after 2010. ‘Transportation’ and ‘Utilities’ relatively do not have much growth after 2010. “Business” is also a major target type.

- As for the seasonal plot, there is an obvious pattern after 2008. ‘Private Citizens & Property’, ‘Military’, and ‘Police’ have large variation than other target types. Particularly, the ‘Military’ has the highest variation after 2010. Between 1995 and 2005 year, “Religious figures and Institutions” have more variation. It might be because of the “11 September attack in 2001”, the tension in the Middle East was stressed. It might be due to the same fact, ‘Private Citizens & Property’ have relatively large variations around 2000 and 2005 years. Between 1980 -1990 year, “Utilities” also show some variations. It might due to the Oil Crisis. People at that time believed that “Little Utility respect to the most terrorist actions”(4). Besides, the business also has some variations before the 1980 year. There is a gap in the 1994 year that all targets have few variations. In the last irregular graph, there is a peak in ‘Private Citizens & Property’ around 2001-2002. It might due to the 9/11 Attacks. “Military” also has a peak around 2014-2015.
- In the ”irregular” plot, it does not show a clear trend, which means that they are white noise. It reaches our expectations.

Since we can see the clear trend in “Military”, “Utilities”, “Police”, and “Business” from graphs above, we want to know the auto-correlation among these four targets. We chose lag.max to be 600 lags (around 50 years,  $600 / 12 = 50$ ).

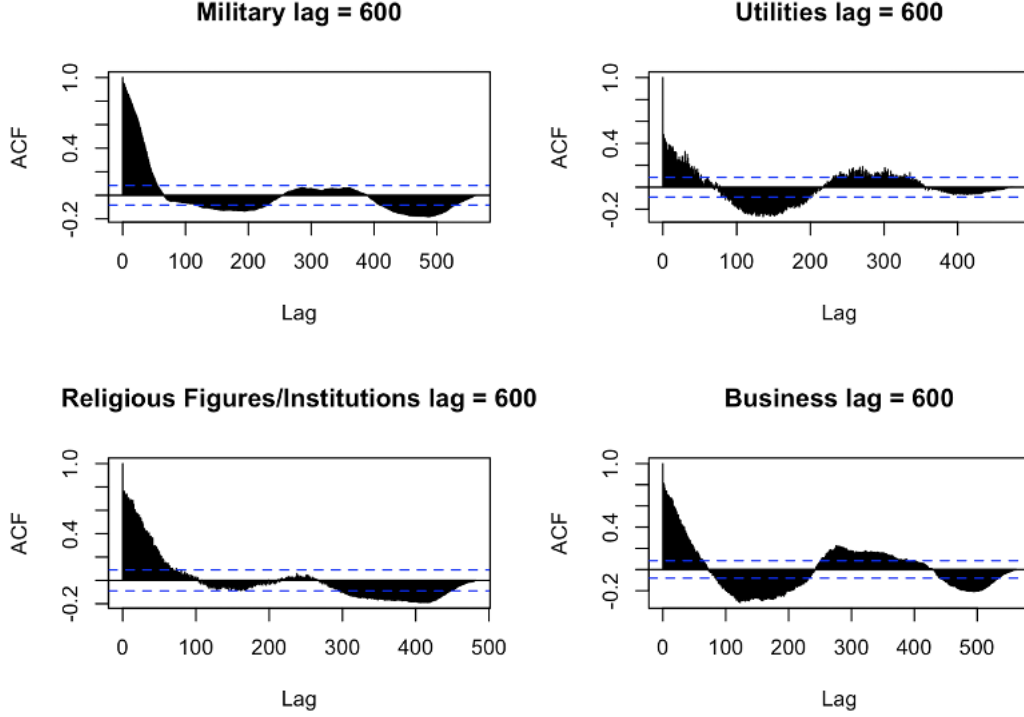


Figure 5

Among these four plots (see Figure 5), there is a much higher positive autocorrelation for lags 0 to 100. All those four autocorrelation plots tend to switch from positive to negative and negative to positive. Also, “Military” and “Business” have most similar patterns among the four plots. The “police” autocorrelation tends to have the longest duration in 0 to 100 and lags 300 to 600. Lags around 400 in the utility graph are not significantly different from zero. Thus, we can see that there is not a clear seasonal trend in one year, but there is a certain trend in one decade in “Business” and “Utilities”.

After carefully examining the target type of attacks alone, we wanted to find some relationship between target type and weapon type by making a conditional bar plot.

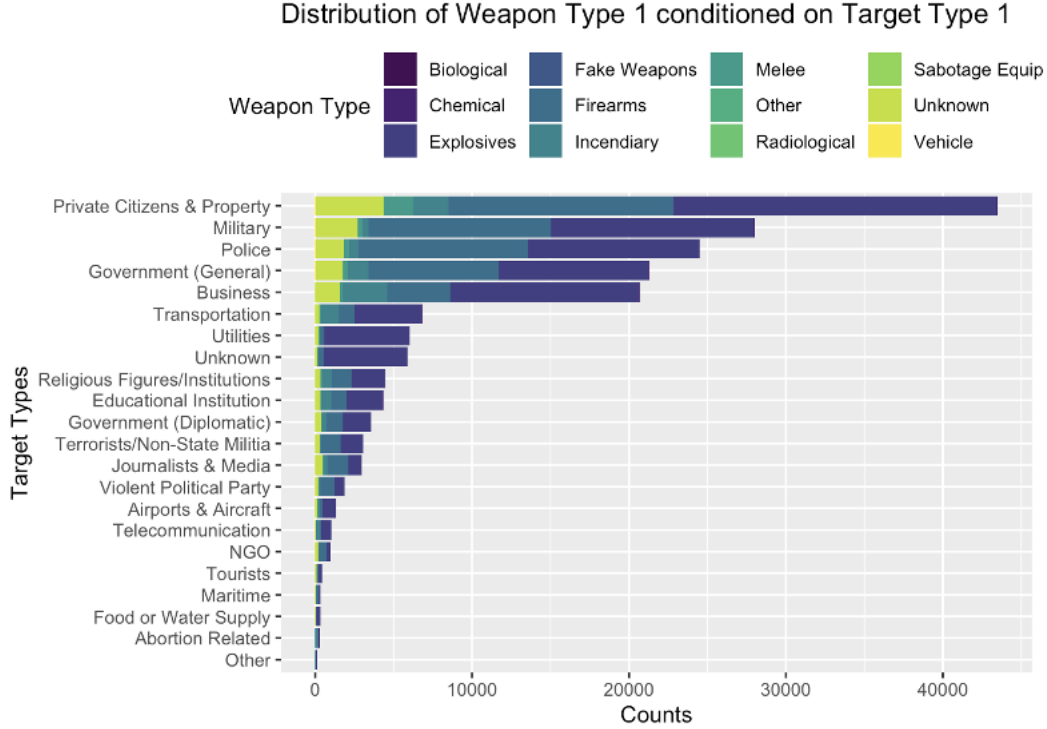


Figure 6

“Weapon Type I” records the major weapon used by the terrorists in the attack; “Target type I” records the major target that terrorists attacked. From Figure 6, we can see that “explosives” and “firearms” take large parts across all target types. “Private Citizens and Property” are the most targeted object among all. We can see that the first weapon that terrorists would choose to attack “utilities” most likely will be explosive. Besides, the “Military” and “Police” have similar ratios for “explosives”, “firearms”, and “unknown” weapons among their total attacks.

Therefore, we can see that after 2010, the number of terrorism for each type of event is increasing, especially for ‘Private Citizens & Property’, ‘Military’, and ‘Police’. ‘Private Citizens & Property’, ‘Military’, and ‘Police’ have large variation than other target types. Explosives and firearms were the major choices across all target types.

### 3.3 A Study on the Number of Attacks

After doing the time series analysis on weapon type and target type, we are wondering whether terrorism attacks become more likely to happen as the world population increases. Therefore, we decided to explore the relationship between the number of attacks and world population of each



year. we fit a linear regression by setting the number of attacks to be the response variable, and world population to be the predictor variable.

```
Call:
lm(formula = attack_counts ~ PopTotal, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-4576  -1925    446   1134   9342

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -8354.1318  2093.8994  -3.990 0.000241 ***
PopTotal      2.1816    0.3661   5.959 3.59e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Notice that the p-value for coefficient is small than 0.05, which means we have enough evidence to reject the null hypothesis that the two variables are independent. As we can see from the R output above, there is a positive relationship between these two variables with slope 2.18. This indicates that if the world population increases by 1 million, there would be 2.18 more number of attacks.

However, by drawing a dot plot for the model (not included in the report due to the limit of graphs), we think linear regression might not be the most suitable model to use, since there are some points after world population reaches 6 billion that deviated from the linear line. As the world population reaches 7 billion, the number of attacks increases dramatically, and after the world population reaches 7.25 billion, the number of attacks tends to drop back. So in the future, we might use polynomial regression models to fit the points better.

### 3.4 Text Data Analysis

Different motives generate different terrorism attacks. Next, we conducted text data analysis for the motives of terror groups.



We are also interested in the specific ransom requests made by perpetrators, as a supplemental data to understand motive above. “Ransomnote” text is about the requests made by the perpetrators. As we can see that not all perpetrators have ransom requests. Or it could be the case that the requests are not known to the outside world, so “unknown” really stands out in the word cloud. Other than that, large part of ransom is about “hostage”, “kidnap”, money “exchange”, and releasing “prisoner”. With some deeper exploration with the data (maybe too deep), we actually come across some very interesting (if not hilarious) ransom notes. Here are some quotes: “...The assailants also demanded that the victim’s family provide them with four Nokia handsets in exchange for her release.” “The assailants requested either a ransom or the apples stocked in the truck.” Obviously, these do not coordinate with the general view we usually perceive ransom requests, but we suppose the lesson here, if any, is that we should not bring our preconceptions into data analysis; otherwise, we may only end up with the conclusions we would like to see.

## 4. Case Study: Terrorism in India

We noticed that India is one of the countries most impacted by terrorism (See Figure 1). Unlike those countries in the middle east or north Africa, which are heavily threatened by terrorism, India has a stable political environment and a steady-paced economic development. So we decided to take a closer look at India and attempted to understand more about the terrorism threats it faces.

We drew a tree-map first, which provides a hierarchical view of the most active terror groups in India and their operation locations. In the graph, we also incorporate the number of killings in each region (shown by the colors), and the number of perpetrators attended (shown by the size of the squares).

*Notice that the number after state name is the number of perpetrators from different terror groups in each region.*

Terrorism in India - size is proportional to the number of perpetrators

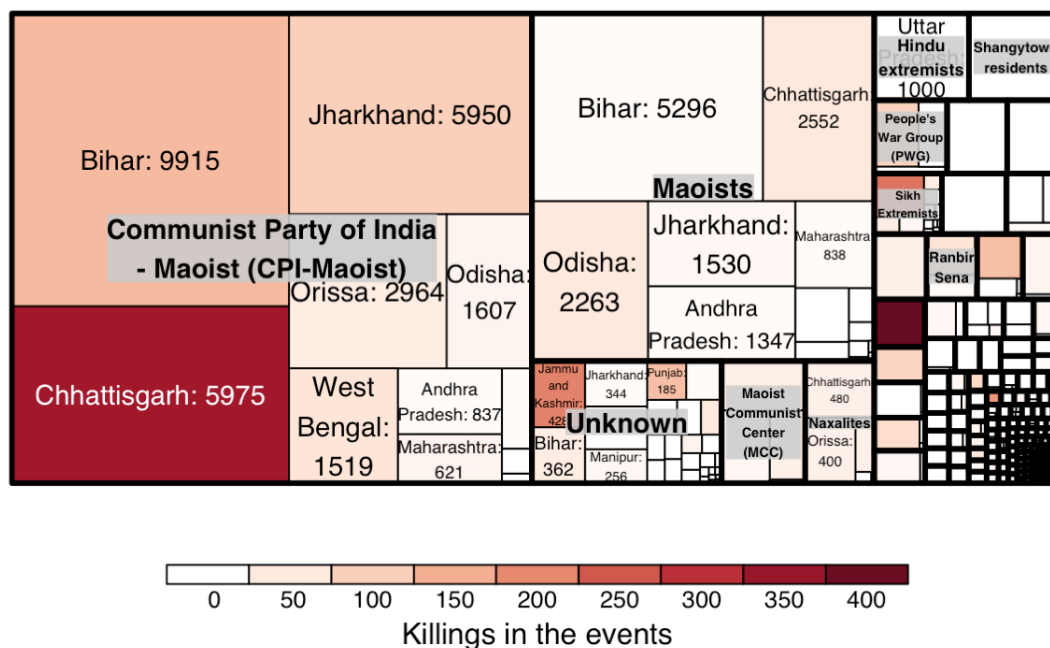


Figure 9

We have 4 major findings:

- India faces many different types of terror groups, including those for political goals (e.g., CPI-Maoist, Maoists), religious goals (e.g., Hindu extremists) and ideological goals (e.g., GNLF).
- Political terrorism groups dominate almost 2/3 of all the terrorists' attacks. Among them, the CPI-Maoist and the Maoists are responsible for most of attacks. This suggests that most terrorism attacks in India are results of political disputes, rather than a product of poverty.
- The location of the events concentrate in 3 states, Bihar, Jharkhand and Chhattisgarh. Although these 3 regions have the most number of perpetrators, Jammu and Kashmir have the most number of killings.
- The Maoists group has lots of perpetrators but relatively few killings. On contrast, the attacks happened in Jammu and Kashmir have the most killings but a small number of perpetrators. This may due to the difference in the forms of terrorism in the two regions. The Maoists is a dissent terrorism group, aiming to overthrow the government through people's war (3). However, Jammu and Kashmir, due to historical reasons, have been the center of political

and military conflicts between India and Pakistan for over half of the century, and the attacks were accused as state-sponsored terrorism (2).

To examine the forth point further, we drew a dot map by attack types.

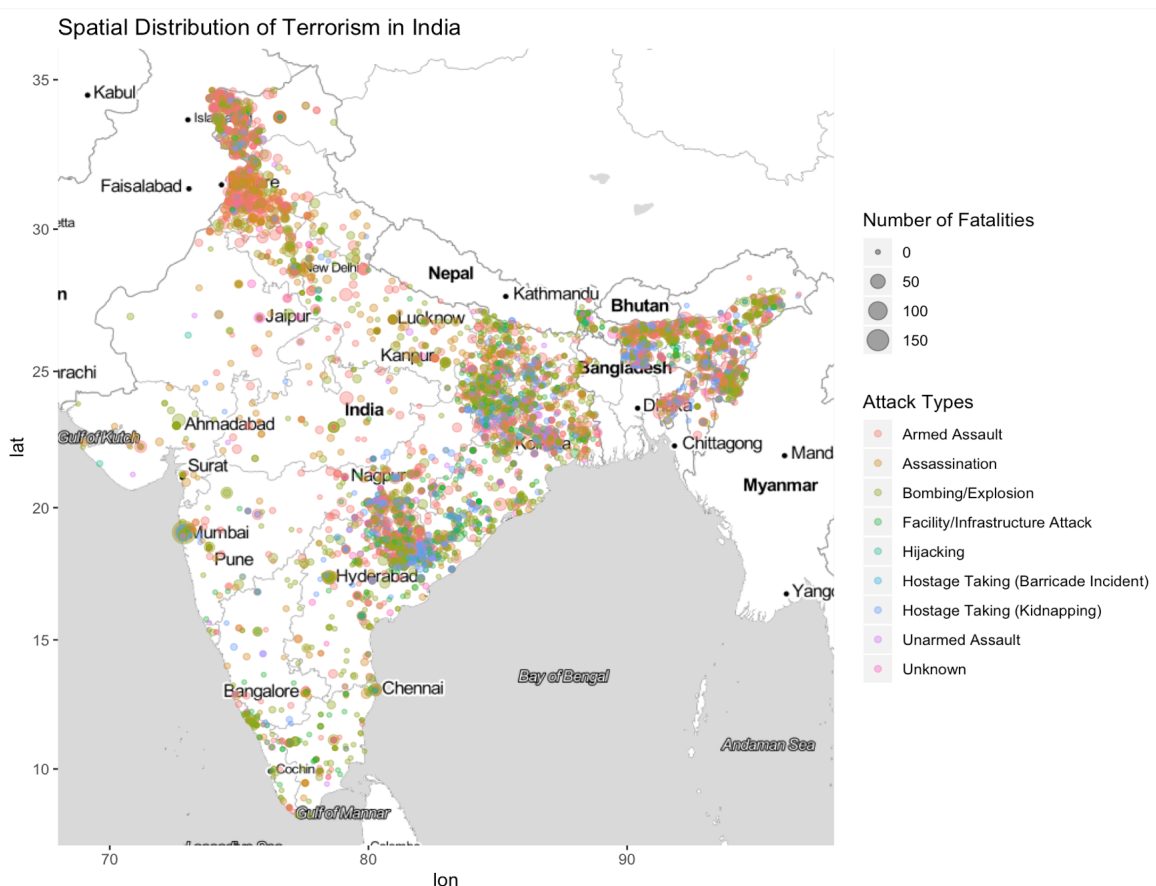


Figure 10

We can observe that attack types are mostly armed assaults in Jammu and Kashmir regions (the north corner), but the attack types of the Maoists' active locations (the east regions) are mostly bombing/explosion, facility/infrastructure attacks and hostage taking. The former is more like military conflicts, while the latter mainly aims at damaging infrastructure. This verifies the differences in casualties and perpetrators that we observed previously between the Maoists' locations and Jammu and Kashmir.

## 5. Conclusions

In this project, we identified both the spatial and temporal distributions of global terrorism. We observed that most attacks happened in the Middle East. We also identified the year 2000 as a turning point for the number of terrorism events in different regions. Before 2000, the events mainly happened in South America, but after 2000, the number of events surged in Middle East & North Africa, South Asia and sub-Saharan African. After observing this phenomenon, we focused on how target type changes over the years and concluded that "Business" and "Utilities" have seasonal trend in terms of terrorism event number, which can be potentially explained by trading conflicts such as the Oil Crisis. Later, we looked at weapon type and target type and discovered how weapon type changes according to the change in target type. Meanwhile, we analyzed the top motivations and ransom notes for these attacks. We noticed that political words and religious words are frequently used, suggesting terrorism attacks are highly related to political and religious conflicts. In the case study, we recognized the major terror groups in India and observed how the number of perpetrators and the number of killings distribute across the country. Based on our analysis and some background information, we reached the conclusion that political dispute is the major cause of terrorism in India.

## 6. Future Work

We can conduct more case studies on different countries to generate a better idea on the global terrorism pattern. We have already seen how the case study on India gives us more insights into the country's specific situation. Given the different political environment and history that each country has, we want to have a closer look at other cases, such as developed countries, developing countries, politically unstable countries, etc. Because of a lack of time to examine each case in details, this task is left for future work. Meanwhile, we would like to update the graphs based on data from 2017 to 2020, in order to have a more accurate view of the most recent international terrorism trends. In addition, if we have more time, we will explore more variables and incorporate them into our research, such as incident summary, ransom amount, unaffiliated individuals, etc.

## References

- [1] Global Terrorism Dataset, Kaggle,  
<https://www.kaggle.com/START-UMD/gtd>
- [2] Pakistan and state-sponsored terrorism, Wikipedia,  
[https://en.wikipedia.org/wiki/Pakistan\\_and\\_state-sponsored\\_terrorism](https://en.wikipedia.org/wiki/Pakistan_and_state-sponsored_terrorism)
- [3] Naxalite-Maoist insurgency, Wikipedia,  
[https://en.wikipedia.org/wiki/Naxalite\0T1\textendashMaoist\\_insurgency](https://en.wikipedia.org/wiki/Naxalite\0T1\textendashMaoist_insurgency)
- [4] The Sociology and Psychology of Terrorism: Who Becomes a Terrorist and Why?, Rex A. Hudson, September 1999, Library of Congress  
<https://fas.org/irp/threat/frd.html>