Project Report

CSCI 6406 - VISUALIZATION

Aniket Biswas | B00781091| 10/04/2018

Objective

The main objective of this project was to design, develop and implement a complete visualization system from an existing dataset and also propose an extensive list of prior related work that is similar to this field of research. I had chosen "Global Terrorism" as the topic of my project and planned to implement an interactive and informative visualization of an existing GTD dataset[1]. There has been little empirical research done in this area of visualization and I have focused my efforts in developing a comprehensive and interactive visualization system whilst summarizing an interesting phenomenon and hypothesis.

Abstract

Global Terrorism is a psychological warfare carried out by terrorists, by manipulating the minds of people and changing their behavior in the name of region, religion and creating fear, uncertainty and separation in society. It is of the highest priority to wage wars and prevent terrorism on a global scale. Nothing can justify terrorism. It can only be resolved by international unity and eradication of hatred, oppression and extremism.

"Despite fearful rhetoric to the contrary, terrorism is not a transcendent threat. A terrorist attack cannot possibly destroy our country's way of life; it's only our reaction to that attack that can do that kind of damage" – Bruce Schneier[2].

Terrorism prevention is a major priority among the law enforcement and security of all the countries in the world. Most developed countries are on the verge of completely eradicating terrorism by strengthening their coordination and combat while other countries are still adopting policies and programs to counter extremism and terrorism as the threat of radical attacks still persist and continue to evolve.

In recent years, there has been a lot of work done on supporting policies and practitioners to share their knowledge and address counter-terrorism. Many studies show a lot of structure and patterns which are involved in the number of attacks in numerous cases. My findings partially support a conventional temporal pattern in the number of attacks and the type of executions that has been carried out.

Through the implementation of this visualization system, I have uncovered visual evidence of a hypothesis by exploring the dataspace and employing different visualization techniques to represent historical data. My findings may reflect a different insight on how most people evaluate terrorism in general and indiscriminate it on the basis of borders.

Related Work

The use of visualization to help analysts, developers and statisticians better understand data has been in existence for decades and has grown in popularity recently. There has been significant research in the field of terrorism which further draws on many other disciplines and concerns, namely, History, Sociology, Politics and Anthropology. Adam Roberts explains the study of evaluating the effectiveness of terrorism and its causes in his paper[3]. Prevention is one of the major factors that drive homeland securities to address the issues that help cause it, in the first place. Brent and Jeff used data from the American Terrorism Study (ATS) to address the important issues of how and why terrorist attacks come to the point of failure[4]. Their findings combined with the relationship between the success of an incident and unsecured targets provided a useful crime prevention approach. Several notable findings from this analysis contributed to a better understanding of incident outcomes. Their research provided an impetus to one of my hypothesis.

Guo, Liao and Morgan, in their research[5] presented a unified visualization environment to support the flexibility of spatiomultivariate (SV), spatiotemporal (ST), temporal-multivariate(TV), and spatiotemporal- multivariate (STV) patterns of terrorism incidents. They focused on applying specific data analytics process to develop a holistic visual representation by ordering the countries affected in a reorderable matrix. This again comes with shortcoming of whether that was recorded was actually reported true to its cause and without any bias, the investigation of which has been done in many research[67].

There has also been work done on cyber-terrorism and weblog, to analyze and visualize networks in terrorism and crime related matters. It presents a framework in the form of a fish eye view to explore the different levels of abstraction[8]. Another honorable mention for using a multi-view display coordinated visualization has also been on terrorism and social network analysis and use natural language processing[9]. The team was able to solve the kidnapping case of the VAST challenge 2014[10]. Another visualization system that proposes a portal approach to analyses the dark web for predictive modelling and visualizing terrorist activities and linkages has been done in the University of Arizona[11]. It also uses cross language retrieval components.

Mapping of crime locations in a spatial concentration in maps has been implemented effectively in a number of periodicals and researches. There has been an implementation of hotspot mapping technique for demonstration on a webpage[12] and the impact of boundary elements while identifying spatial clusters on a world map[13] which provides a comprehensive study of clustering data on a world map.

Use of zoomable user interface for representing the analysis of dynamic data is still a work in progress. It portrays an animated bubble chart to represent spot-tracking lens to allow users to examine data by zooming in and out in a still under-rated research area[14].

Dataset and Data

The Global Terrorism Database (GTD) is an open source dataset which is maintained by National Consortium for the Study of Terrorism and Responses to Terrorism (START) containing information of all terrorist attacks around the world from the year 1970 to 2016. This dataset portrays a systematic data of both international and domestic terrorist attacks and casualties (both injuries and fatalities) accumulating to approximately 1,70,000 cases. The data has been collected form unclassified sources and comprises of many attributes such as location, targets, tactics, outcomes, attack type, casualties, group names etc. All data had been collected and retrieved from the GTD website and Kaggle.

I started off with pre-processing and cleaning the dataset, since it consisted of various inconsistent and null data. My requirement demanded me carve out different attributes and format/parse them to .csv, .json and tabular formats, to portray it in different types of views. This revised data was then uploaded to public gists on GitHub for it to be accessible publicly and for the system to run efficiently on Chrome and other modern browsers which cannot access local data with JavaScript. The links have been provided along with the code and there is also a separate folder (data) where all the revised data is stored.

The main attributes which were considered for the development of this project are:

1. **iyear** (numeric variable) - This field shows the year in which the incident or attack occurred. In cases of incident(s) occurring over extended periods, the field shows the year when the incident was initiated. There is also a field for the approximate date (day and month), which has not been taken into consideration for this project
2. **country** (location code) – This field identifies the country with its respective code as provided on the GTD. This field had to be preprocessed to match the ISO Numeric Code of each country.
3. **country_text** (categorical variable) – This field contains the name of the country or location where the incident occurred. Some of the values did not match the country names in the GeoJSON format which had to be re-formatted. Many countries were coded "Unknown" or which no longer existed, for instance, "USSR" or "West Germany" which has been assembled as Germany and Russia respectively.
4. **attacktype1_text** (categorical variable) – This field consists of the general mode of attack that was carried out in an incident. For e.g., whether the assault was armed, unarmed, bombing, hijacking, etc. Only one attack type is registered per incident.
5. **gname** (text variable) – This filed portrays the name of the perpetrator group that carried out the attack or claimed responsibility for it.
6. **nkill** (numerical variable) – This field shows the total number of confirmed fatalities that are related to a particular incident. It displayed a value of "-99" for null values which had to be removed so as to not affect the count of this variable.

The overall count, average and sum of data were calculated using basic Excel formulae.

Implementation

I initially intended on creating a choropleth of the world map in either a thematic surface view or a rotating globe containing the information of attacks and their process flow from the source to the victim. I initially worked on the rotating globe view and its transition into a thematic map. Thus, on selecting a certain country of the globe, it would fetch its details and most successful terrorist attacks as a tooltip. This is when I came across certain constraints of visualizing large amount of data inside the tooltip or the fact that certain countries in the world have never been attacked on a global scale. It was one of the first hurdles that I came across while approaching this view.

On further research, I came across a more efficient way to visualize the incidents on a world map. This brings me to my first view, which shows the temporal pattern of data.

1. **Timeline**: It is my first view which portrays color encoded countries on a thematic map according to the number of incidents that has taken place in that region. This mainly signifies the victim country of an attack or the country where the incident has been recorded to have occurred. It acts as the timeline of attacks with a tooltip to identify a country and its count of attacks and a legend to denote the incident count ranges. An interactive bar chart at the bottom also renders the same data with change of time (range slider). A typical screengrab of the year 1990 is shown below (Figure 1).
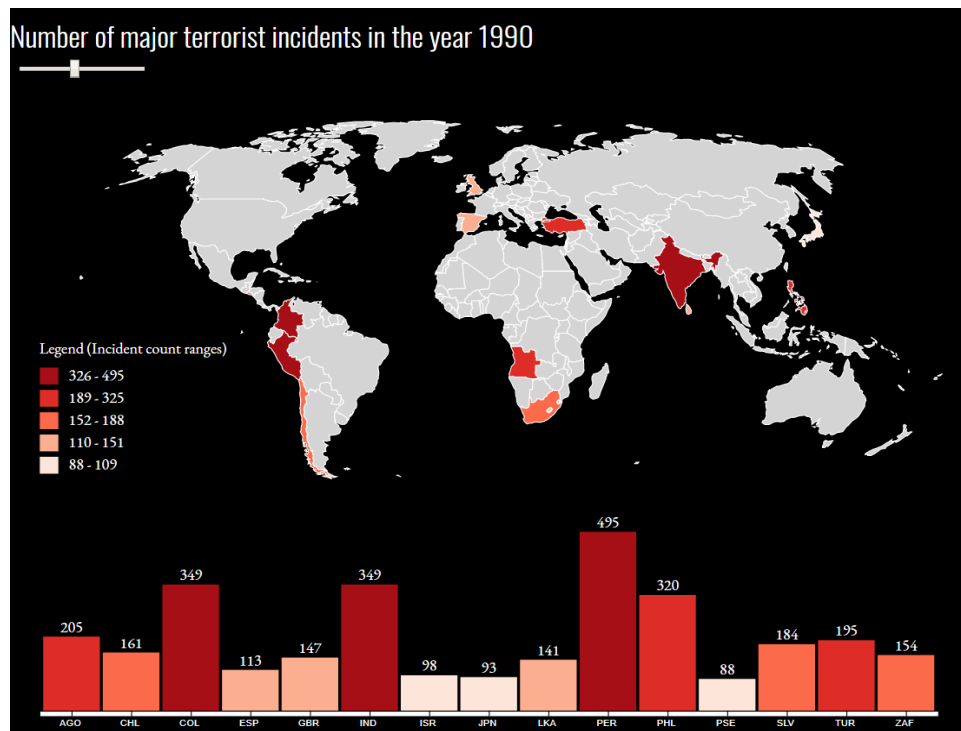

Figure 1: Timeline displaying the major terrorist attacks over the years on a map

This is when I noticed a structural and temporal pattern in the number of attacks that has taken place starting from the year 1970 right up to 2016 and its outspread on the map.

a) **Hypothesis 1**: There is an obvious trend in the number of incidents occurring in countries that gained independence in the latter half of the 20th century as opposed to the already developed countries. There is also rising tendency and gradual shift of the number of attacks from the Western Hemisphere to the African, Middle Eastern and Asian regions near the equator.
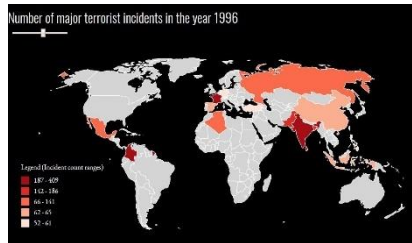

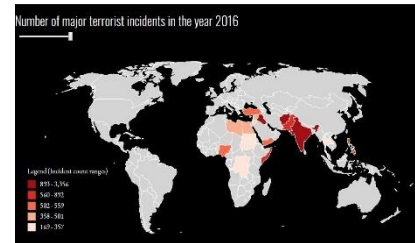*Figure 2: Year 1973*   *Figure 3: Year 1996*   *Figure 4: Year 2016*

It is pretty evident from Figures 1-3, that there has been a radical shift in the number of terrorist incidents from the West to the East. In Figure 1 (year 1973), there were more attacks in the United States and Great than the rest of the world combined. 1970s was the decade when the Apartheid and Black Power Revolution came into the picture. It was also the year when The Nuclear Non-Proliferation Treaty was put into effect (March, 1970). Surprisingly, this was a peace treaty where nearly 200 countries joined hands to work towards complete disarmament of nuclear weapons and warfare. Some of the incidents included the terrorist attacks on the Munich Olympic Games, US-Vietnam War and US-Cambodia combats that came to an end after 12 years.

Figure 3 shows a change in the world terrorism scene with the more developed countries, such as the USA and Great Britain, showing lesser signs of major incidents on their homeland. 1996 was the year when major terrorist groups "Taliban" and "al-Qaeda" were founded with fundamentalists and extremists capturing Afghanistan. This was also the year when US declared war on Iraq and several other major incidents in the Middle East. This also attributes to the fact that the developed countries were fighting an uphill battle in mitigating similar types of attacks with increase of combat and security.

Figure 4 portrays a present-day scenario which shows the rise of major terrorist groups in regions of Syria, Iraq, Pakistan, Afghanistan and India. This was the decade when ISIS came into the picture and major terrorist incidents in Iraq and Syria took place. The top 20 countries with highest rated incidents in this decade were mostly developing or third world countries whilst the developed countries enjoyed a negligible number of cases. This pattern attributes to how the outgrowth of population explosion, poverty and a lack of justice played an infamous role in aggravating the number of incidents. The attacks are also becoming more centralized which could provide an impetus to curb them frantically.

2. **<u>Casualties</u>**: This is my second view which displays the attack_type attribute with the help of a pie chart that mainly examines the impact of different types of attack and the rate of its success. This visualization has portrayed an interesting find which shows how likely is an incident going to be successful in the mortality of human lives.

This graph portrays three aspects of the attack type –

a) The number of attacks of a certain type which signifies the number of attempts.
b) The number of casualties which denotes the total number of successful killings.
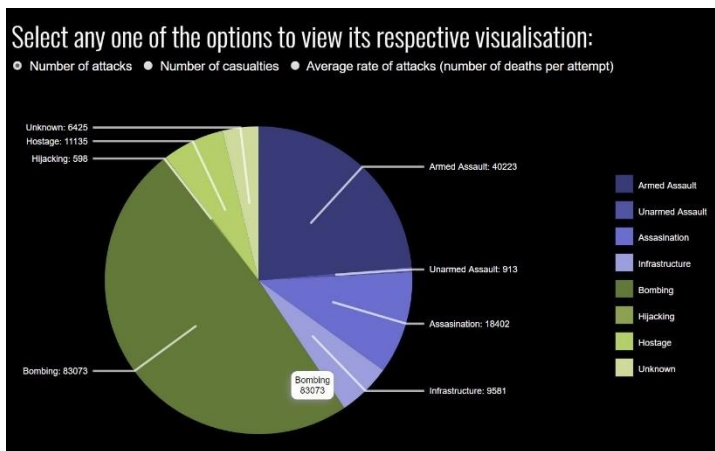c) The average success rate of an attempt (what type fails/foils and what succeeds).

Figure 5: Pie chart highlighting incidents of Bombing

Figure 5 shows a classic case of how Bombing attributes to 83073 number of incidents while High jacking, only a mere 598 times. This is not highly unlikely for terrorists to resort more to explosives to destroy a greater amount of habitation as it contains greater potential energy to destroy greater number of lives. This when compared to Hijacking, does not produce the same or greater amount of destruction or damage per se.

Now, looking into Figure 6, on the average rate of attacks and its success rate, we can find out that Bombing, even though, has a greater number of attempts and feasibly higher destructive and eliminating power, it only accounts for 1.83 of deaths. This means that for every 83073 attempts made by means of bombing, 1.83 people died per attempt. This, when compared to only 588 cases of hijacking accounted for 6.65 kills per attempt. So, isn't this a comprehensive estimate of where security needs to focus?!
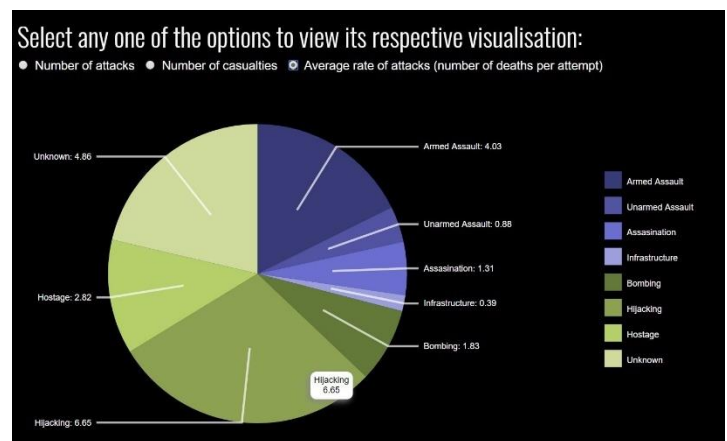

Figure 6: Pie chart demonstrating the success of attacks

b) **Hypothesis 2:** Hijacking a flight is approximately 5 times deadlier than bombing it! Hijacking accounts for more kills per incident as compared to any other attacks.

3. **Perpetrators**: This view comprises of the perpetrators that were involved in all of these attacks and the groups who claimed responsibility for most of the attacks. Most of the times, it is difficult to ascertain as to who is actually responsible for an attack as there may be various groups or individuals who might collectively claim one or even deny it. There are approximately 3000 perpetrator groups to have been in existence in the last 50 years. I have displayed the Top 50 deadliest perpetrator group names who are responsible for the death of millions of people worldwide.
I have used a bubble chart to display data in the area of circles with the diameter signifying the number of incidents that they have claimed responsibility for. I have also portrayed it tabular form below the chart to ascertain it by rank.

Figure 7: Bubble chart displaying the Top 50 deadliest terrorist organizations

Group Name	Incident Count
Shining Path (SL)	4523
Farabundo Marti National Liberation Front (FMLN)	3351
Irish Republican Army (IRA)	2673
Taliban	2421
Revolutionary Armed Forces of Colombia (FARC)	2129
Basque Fatherland and Freedom (ETA)	2027
Liberation Tigers of Tamil Eelam (LTTE)	1606
Communist Party of India - Maoist (CPI-Maoist)	1464
New People's Army (NPA)	1413
National Liberation Army of Colombia (ELN)	1300
Palestinians	1139
Nicaraguan Democratic Force (FDN)	895
Manuel Rodriguez Patriotic Front (FPMR)	830
Tehrik-i-Taliban Pakistan (TTP)	753

Figure 8: Tabular list of Top 50 perpetrator groups and terrorist organizations

Dependencies

There are various dependencies that have been utilized during the course of my project.

1. D3.js: JavaScript library used for creation of SVG and all the graphs.
2. The map has been created using DataMaps and Natural Earth.
3. Data has been retrieved from START website.
4. Others – geoProjection, scaleChromatic, topojson, jquery, modernizr.
5. The default webpage has been retrieved from a sample template from Blacktie.
6. A running internet connection is required. Default home page is index.html.
7. Various other dependencies have been referenced in a separate file, References.txt.

Future Work

There were a lot of other attributes which I could not lay my hands on. Those variables might contain other riveting patterns and theories which might suggest a better approach to understanding terrorism and supporting counter terrorism. Although this project has been an important first step towards that goal, it has not been without certain limitations.

There are other attempts that has been made in adopting a counter terrorism policy and related work in this field has shown promising research opportunities by analyzing temporal data. My understanding of the logistics of how terrorism works has increased a great deal after this, and I am planning to hold on to this approach of examining incidents and perpetrators to understand a pattern of what other characteristics of a country (such as GDP, educational levels, military expenditure) has, on successfully triggering an attack.

Conclusion

My initial plan was to just provide an informative overview of data in the form of graphs by flattening facts and numbers. After successful execution of the graphs and charts, I discovered interesting hypothesis and trends by studying the different views. All the three views portray different meanings and signify different attributes of terrorism as a subject.

Visualization is basically a pictorial representation of already known facts. While exploring the dataspace, there was no evidence as to what might be uncovered after visualizing this data. But, after the completion of this system, I stumbled upon more interesting facts that were previously unknown to me. This goes to show how visualization is a vital aspect of data analytics which mostly goes unexplored between sheets of rows and columns.

"The greatest value of picture is when it forces us to notice what we never expected to see" –
John W. Turkey

References

[1] National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2013). GTD [Data file]. Retrieved from http://www.start.umd.edu/gtd"

[2] Bruce Schneier Quotes. (n.d.). BrainyQuote.com. Retrieved April 10, 2018, from Web site: https://www.brainyquote.com/quotes/bruce_schneier_443192

[3] Adam Roberts (2014) Terrorism Research: Past, Present, and Future, Studies in Conflict & Terrorism,38:1, 62-74, DOI: 10.1080/1057610X.2014.976011.

[4] Brent R. Klein, Jeff Gruenewald, Brent L. Smith, Opportunity, Group Structure, Temporal Patterns, and Successful Outcomes of Far-Right Terrorism Incidents in the United States, Crime & Delinquency, Vol 63, Issue 10, pp. 1224 – 1249

[5] Diansheng Guo, Ke Liao, Michael Morgan, Visualizing Patterns in a Global Terrorism Incident Database, Environment and Planning B: Urban Analytics and City Science, Vol 34, Issue 5, pp. 767 – 784.

[6] Suat Cubukcu, Brian Forst, Measuring Terrorism, Homicide Studies, Vol 22, Issue 1, pp. 94 – 116, Published December 20, 2017.

[7] Jeff Gruenewald, A Comparative Examination of Homicides Perpetrated by Far-Right Extremists, Homicide Studies, Vol 15, Issue 2, pp. 177 – 203.

[8] C. C. Yang and T. D. Ng, "Terrorism and Crime Related Weblog Social Network: Link, Content Analysis and Information Visualization," 2007 IEEE Intelligence and Security Informatics, New Brunswick, NJ, 2007, pp. 55-58.

[9] R. Gao, H. Tao, H. Chen, W. Wang and J. Zhang, "Multi-view display coordinated visualization design for crime solving analysis: Vast challenge 2014: Honorable mention for effective use of coordinated visualizations," 2014 IEEE Conference on Visual Analytics Science and Technology (VAST), Paris, 2014, pp. 321-322.

[10] Visual Analytics Community(online), http://vacommunity.org/VAST+Challenge+2014

[11] G. Lai, E. Reid, J. Qin, H. Chen and Y. Zhou, "Studying the presence of terrorism on the web: an knowledge portal approach," Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '05), Denver, CO, 2005, pp. 402-402.

[12] Guiyun Zhou, Jiayuan Lin and Wenfeng Zheng, "A web-based geographical information system for crime mapping and decision support," 2012 International Conference on Computational Problem-Solving (ICCP), Leshan, 2012, pp. 147-150.

[13] Y. Zhang and R. Maciejewski, "Quantifying the Visual Impact of Classification Boundaries in Choropleth Maps," in IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 1, pp. 371-380, Jan. 2017.

[14] Y. Hu, T. Polk, J. Yang, Y. Zhao and S. Liu, "Spot-tracking lens: A zoomable user interface for animated bubble charts," 2016 IEEE Pacific Visualization Symposium (PacificVis), Taipei, 2016, pp. 16-23.