UNIVERSITY OF CALIFORNIA

SANTA CRUZ

DATA DRIVEN VISUALIZATION

World of Terrorism

Top 14 Deadliest Terrorist Groups

2000 – 2015

A report submitted in partial satisfaction of the

requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE

by

YIFAN WEI

March 2017

The Project Report

is approved:

_____

Professor Suresh K. Lodha

_____

Professor Jishen Zhao

Table of Contents

Web link

Visualization: https://erfan00.github.io/world_of_terror_6/

GitHub Repository: https://github.com/erfan00/world_of_terror_6

List of Figures

Acknowledgement

I would like to express my special appreciation and thanks to my advisor Professor Dr. Suresh K. Lodha for mentoring me and supporting me to do this project.

Abstract

Terrorism has posed tremendous threats to both individual's life and social civilization. The severity of terrorism motivates people to explore the inner patterns and motivations of terrorist attacks. Global Terrorism Database (GTD), an online open source database that includes more than 150,000 cases of terrorist attacks, has provided a great platform for supporting those who are willing to probe the unknown side of terrorism by extracting and analyzing the data of terrorist activities. Here we implemented a data visualization project of the global terrorist attacks from 2000 to 2015 based on the GTD. The main visualization is based on a terrorist group-driven interface focusing on the detailed information of the top 14 deadliest terrorist groups, which would display the number of casualties caused by one specific terrorist group during this time range. Meanwhile, it also demonstrated the casualty's information concerning top 13 victimized countries from 2000 to 2015.

The original database includes the terrorist attacks from 1970 to 2015. Here we subtracted the datasets from 2000 to 2015 since we focused more on recent terrorist activities. Then we grouped the dataset based on terrorist groups. There are more than 1000 terrorist groups in this dataset. We sorted the terrorist groups' list based on the total number of people killed by each group and selected the top 14 terrorist groups. Also we discovered some geo-patterns of terrorist attacks by analyzing countries with most casualties during 2000 to 2015. The dataset was grouped based on countries and the top 13 countries with most victims were selected.

The purpose of the visualization is to demonstrate the trend of the number of people killed by each terrorist group and the trend of the number of victims in each country. Multi-line charts and stack bar charts were applied to show the changing number of people killed in each year for each terrorist group or country. The original multi-line chart were split into two multi-line charts: the first one contains all the information of the selected terrorist groups or victimized countries, while the second one removed the top 3 groups or countries in order to display the information hidden underneath the main chart. The color scheme of the terrorist groups is as follows: 9 terrorist groups that are related with Islamic were marked red combine with yellow of different shades, and the groups with more killing would have a darker shade while groups with less killings would have a lighter shade. Other groups held different purposes such as civil war or military are marked red combined with purple of different shades. The color scheme of the victimized countries is based on geographical locations: the countries that are belonging to same regions are marked with similar colors with different shades.

The final visualization product revealed that the top 3 deadliest terrorist groups are Taliban, Islamic State of Iraq and the Levant and Boko Haram. And they became more active since 2012. Among the top 14 deadliest terrorist groups, 9 of them are Islamic-related. The top 3 countries that suffered most are Iraq, Afghanistan and Nigeria. The United States suffered from the tragedy of the 911 attacks, which forms a peak on 2001; while for most of other countries, the numbers of people killed are increasing after 2001. For the whole project, it help us to discover the geo-temporal patterns of global terrorist attacks during 2000 to 2015 and to provide some insights on defeating future terrorism.

1. Introduction

The 15th anniversary of 9/11 attacks has once again brought terrorism to the public's mind. 2996 lives perished in this worst terrorist attack in world history and the grief never goes away. Terrorism has been haunting like a ghost in the modern society and turned into one of the major threats to world peace. However, there is still uncertainty concerning the root causes of terrorism and it is really a challenging task to predict the time and locations of terrorist attacks. It always has controversy over the causes of terrorism, which may involve various social, political and economic factors and differ on specific spatial and temporal circumstances [1]. Sociologists and behavioral scientists have provided some insight to understand the causes, tactics, motivations, historical evolution, new adversaries and societal responses of terrorism [2]. However, it is still not enough for foreseeing where and when future incidents might occur. It is essential to gain a thoroughly understanding of the terrorist group emergences and evolutions and also the geographical and historical distributions of terrorist activities [3].

Several organizations have engaged in building terrorism events databases to analyze the tendency of terrorism activities. There are five best-known terrorism events databases: Global Terrorism Database (GTD), International Terrorism: Attributes of Terrorist Events (ITERATE), Rand Database of Worldwide Terrorism Incidents (RDWTI), World Incident Tracking System (WITS), Terrorism in Western Europe: Events Data (TWEED) [4]. Among those databases, Global Terrorism Database (GTD) is considered as the largest, most comprehensive open-source database including information on terrorist attacks that took place around the world between 1970 and 2015, and now it is fully geo-coded that allows users to conduct geo-spatial analysis on terrorism for the past five decades [5]. The GTD contains more than 150,000 terrorist attack cases and includes both domestic and international terrorist incidents that have occurred during 1970 to 2015. More than 75,000 bombings, 17,000 assassinations, and 9,000 kidnappings incidents have been included in the GTD. For each terrorist attack, it provides detailed information about the date, location, target, the weapons used in the incident, the number of death and wounded, and the group or individual responsible for the incident if the info is available. The data source of GTD is generally based on reports from various open media sources and would be verified before entry of the database [5].

It is a challenging task to understand the inner pattern and trends of the GTD since it has multivariate patterns and multi-dimensional datasets. Here are some interesting topics that we want to discover from the GTD: What is the geographical distribution of terrorist attacks during the past 45 years? How does the geographical distribution change over time? What are the major terrorist groups and what are the main areas they intend to attack? What is the number of casualties caused by terrorist attacks in different countries? To gain more insights of those topics, here we implemented a data visualization project using D3 based on the GTD between 2000 and 2015. The purpose of this project is to shed light on our understanding of terrorist activities, to support the research and study of terrorist violence and to provide some insights on defeat of future terrorism.

2. Background and Related Work

2.1. Computational Social Science

With the rapid advances in technology, unprecedented volumes of data have been created revealing both individual and group's behavior in our society. For instance, people send emails to communicate with their colleagues in daily work, make online transactions to get the merchandise that they need, connect with each other and maintain friendship through social networks. All of such activities would generate digital traces that can be translated to comprehensive pictures to assist the understanding of our lives and societies [6]. With thrilling new opportunities brought by increasing volumes of data, a new discipline known as Computational Social Science has been turning into more prominent to move us towards a quantitative comprehension of our intricate social systems [7].

Computational Social Science is defined as "the interdisciplinary investigation of the social universe on many scales, ranging from individual actors to the largest groupings, through the medium of computation" [8]. In general, it refers to an interdisciplinary that makes computational approaches to social sciences. It empowers collecting and analyzing data with an epoch-making breadth and depth and scale [6]. Computational Social Science involves with a board range of fields such as computational economics, computational sociology, anthropology, human geography, psychology and law, and focuses on investigation of social and behavioral relationships through the combinations of algorithmic information extraction, network models, social complexity analysis and various computer simulations [9].

Numerous studies have been done in the field of computational social science. Onnela et al. have figured out the local and the global structure of a society-wide communication network by exploring the communication patterns of millions of phone users [10]. Balcan et al. demonstrated the relationships between human mobility with the spatiotemporal pattern of a global epidemic by analyzing mobility data from 29 countries and integrating a timescale-separation technique in a worldwide-structured meta-population epidemic model [11]. Batty provides new insights into how cities evolve by linking urban economics and transportation behavior to developments in network science, allometric growth and fractal geometry [12]. Chainey et al. summarized the relations between Geographical Information Systems (GIS) and practical criminal justice issues and demonstrated crime mapping can play an important role in crime reduction process [13]. Lillo et al. identified groups of players on the market by subtracting individual strategies from detailed financial data and also demonstrated the players' role in stabilization and destabilization [14]. Dzogang et al. discovered periodic changes in the behavior of a large population by analyzing historical news of United States and United Kingdom from 1836 to 1922 [15]. In general, computational social science has demonstrated powerful capability to assist unveiling the inner patterns of various problems such as society communication network, international epidemics, urban development, crime issues, trends of stock market and the periodic patterns of human behaviors. It is essential to apply the insights of computational social science to get a better understanding of the increasingly complexity of our interconnected global society [7].

2.2. Data Journalism

Now we are living in a digital world that almost everything can be described with numbers, which is transforming traditional journalism into a new media approach – Data Journalism. Data Journalism can be simply explained as journalism done with data, which can help a journalist tell a compelling story through engaging info-graphics [16]. Raw data is always confusing, boring and puzzling to the public's mind. Therefore, It has a growing value to gathering, filtering, visualizing the raw data and re-organizing it into a more comprehensive way. Currently, pioneers like the New York Times, the Guardian and the Texas Tribune have demonstrated how data-driven stories can provide deeper insights into the inner patterns of our rapidly changing global society [16]. Many visualization tools have been created and accelerated the development of Data Journalism.

2.2.1. D3

Data-Driven Documents (D3) is a JavaScript library for making representation-transparent approach to dynamic and interactive data visualizations for the web [17]. According to the official definition, "D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation [18]". Specifically, D3 enables users to directly inspect and manipulate DOM by selectively binding input data to arbitrary document elements and dynamically transforming the generation and modifications of content [17]. The feature of representational transparency makes D3 more expressive and better integrate with developer tools. D3 also retains features of high efficiency, powerful declarative components and simplified debugging [17]. D3 is the successor of Protovis. In 2009, Bostock et al. presented Protovis - a graphic toolkit for visualization by composing simple graphical primitives [19]. In 2011, Bostock et al. released the first version of D3 with comparable notational efficiency [17]. Currently, many data visualization sites use D3 as the core toolkit to build compelling visualizations. The New York Times is one of the best examples of D3 usage in digital media. Fig. 1 demonstrated a D3 example of an interactive map of the presidential election result of 2016 [20].

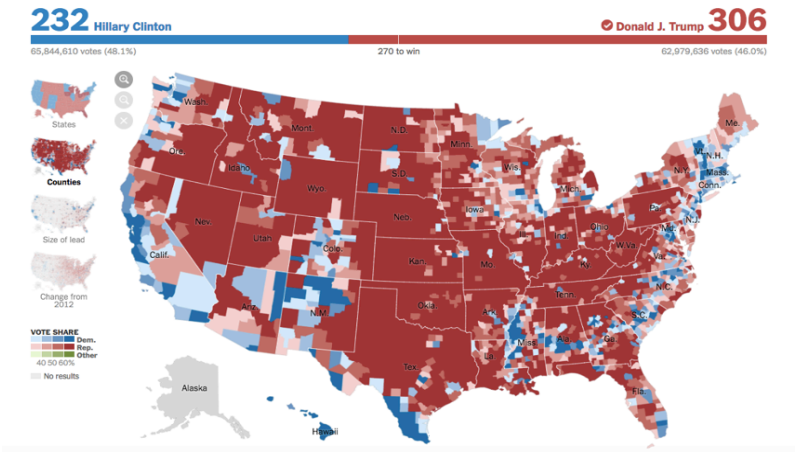


Fig 1. A D3 example of 2016 presidential election results built by The New York Times [20].

2.2.2. Visualization Tools

According to Wikipedia, there are more than 40 popular JavaScript charting frameworks (Fig 2.) [21].

Framework	Line	Timeline	Scatter	Area	Pie	Donut	Bullet	Radar	Funnel	Gantt	Network	Grouped	Stacked	Negative	Discrete	Horizontal	Legends	Mouse Over	onClick	HTML5 Canvas	SVG	VML	AxisXY
amCharts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
AnyChart	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
C3.js	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes
CanvasJS	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	
canvasXpress	Yes	No	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	Yes	
Chartist	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	
Chart.js	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Chart Builder By Livegap	Yes	No	No	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	
Charts 4 PHP	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	No	
Cytoscape.js	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	Yes	Yes	Yes	No	No	Yes
D3.js	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
dc	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes	No	Yes	No	
DevExtreme	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	
DHTMLX Charts	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	
dimple	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Dojo Charting	Yes	No	Yes	Yes	Yes	No	No	No	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Dygraphs	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	
Factmint Charts	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	
Flot Charts	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No	No	No	Yes	Yes	Yes	Yes	No	No	
FusionCharts	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
Flotr2	Yes	No	Yes	No	Yes	No	No	Yes	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	
Google Chart Tools	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	
gRaphael	Yes	No	Yes	No	Yes	No	No	No	No	No	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No	Yes	No	
Highcharts, Highstock	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
JenScript	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes
jqPlot	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	
KoolChart	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
MetricsGraphics	Yes	No	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
NextCharts	Yes	No	No	Yes	Yes	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
NVD3	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	
OLAPCharts	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	
plotly.js	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	
PlusCharts	Yes	No	No	Yes	Yes	No	No	No	No	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	
RGraph	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	
rickshaw	Yes	No	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes	No	No	No	Yes	Yes	No	No	Yes	No	
Shield UI	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	
Syncfusion	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TeeChart for Javascript	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	
Vaadin Charts	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
VanCharts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	
Webcharts	Yes	Yes	Yes	No	No	No	Yes	No	No	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	No	
Webix UI	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
xcharts	Yes	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes	Yes	No	Yes	No	
YUI Charts	Yes	No	No	Yes	Yes	No	No	No	No	No	No	Yes	Yes	No	No	No	Yes	Yes	No	Yes	Yes	Yes	
ZingChart	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZoomCharts	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	

Fig 2. Comparison of JavaScript charting frameworks [21]

It is really challenging to find and choose an appropriate visualization tool since there are numerous JavaScript libraries available here. If we want to choose the right one, first we need to know what kind of approaches each framework applied to. Basically, there are two approaches: declarative approach and imperative approach. Most of JavaScript charting libraries apply the declarative paradigm, which means the users first define the result they want to end up with and then the library provides an interactive chart based on the user's requirements. For example, amCharts, Chart.js, Highcharts and FusionCharts are following the declarative approach. On the other hand, D3 or Paper.js apply the imperative paradigm, which means the users need to write code step by step to achieve the final visualization effects [22].

Here are some brief introductions of several recent popular visualization tools: Google Chart is a flexible library with large amount of commented code that enables users to embed HTML5 / SVG into web page, it is a fantastic tool for beginners with great tutorials so that users can create visualization step by step; FusionCharts is a scalable tool which supports jQuery and Angular and contains over 90 charts and 1000 maps, it is a more complete solution that would meet enterprise requirements; D3 has the most complete features and can create compelling user customized visualizations, it is more suitable for hardcore charting experts; Dygraphs can plot large data sets without slowing down, however for users who pursue a nice UI design, it may not be a good choice; dc.js are proficient in making interactive dashboards, which is good for creating a dashboard of relational charts [23].

2.3. Principles of Visualization

There are some important principles to follow when we are planning to do a data visualization project. First, the data must be trustworthy and should be presented without lies, distortion or fancy but meaningless effects. Second, Any unnecessary or redundant graphical design should be removed to create a clean and distraction-free visualization. Data-ink, which is the non-erasable core of the graphic, should be kept and all non-data-ink and redundant data-ink should be removed [24]. Third, a good visualization should be self-explanatory, all related information need to be provided in the graph to help the audiences understand the visualization.

2.4. Background on Research Topic

Terrorism has been pervasive during the past decades. In response, there is a growing interest in collecting and analyzing terrorism event databases. An open source global terrorism database (GTD), maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START), has been released since 2007 [25]. It includes more than 150,000 terrorism cases both domestically and internationally from 1970 through 2015. It originated from the incident data between 1970 and 1997 collected by the Pinkerton Global Intelligence Service (PGIS), and then continued with the efforts of several collection Intuitions such as CETIS, ISVG and START [5]. Current data collection is done by START combining automated and manual data collection strategies. The collection process begins with scanning over and filtering one million media articles and then being reviewed by the GTD team manually to identify attacks to be added to the GTD.

Sometimes it is hard to define terrorism. According to the GTD's definition, a terrorist attack is "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion or intimidation". In order to be included in GTD, an incident must be intentional and entail some level of violence or immediate threat of violence; also the perpetrators of the incidents must be sub-national actors [5]. Besides, it must meet at least two of the following three criteria: "1. The act must be aimed at attaining a political, economic, religious, or social goal; 2. There must be evidence of an intention to coerce, intimidate, or convey some other message to a larger audience (or audiences) than the immediate victims; 3. The action must be outside the context of legitimate warfare activities"[5]. The variables of the GTD include GTD ID and date, incident information, incident location, attack information, weapon information, target/victim information, perpetrator information, casualties and consequences.

Several related works have been done based on the GTD. Jones et al. developed a novel uncertainty visualization techniques to identify geo-spatial trends within the GTD [26]. Wang et al. proposed a visual analytical system focusing on depicting one of the five W's (who, what, where, when, and why) of the terrorist activities included in the GTD [27]. Lee presented a web-based visualization of the GTD's temporal data [28]. Spaaij presented the global patterns, motivations and prevention of lone wolf terrorism by analyzing the GTD [29]. Drakos demonstrated terrorism would exert a significant negative impact on daily stock market based on the data sample of 22 countries of the GTD from 1994 to 2004 [30].

3. Research

3.1. Purpose of visualization

Terrorism is the use of intentionally indiscriminate violence to achieve a political, religious or ideological goal [31]. According to the definition of The Global Terrorism Database (GTD), a terrorist attack is "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation" [5]. More than 150,000 terrorist attacks have occurred from 1970 through 2015 and claimed over hundred thousands lives [5]. To gain a better understanding of terrorist violence, especially on the recent terrorist groups that have the most threats to human's life, we proposed a data visualization project "World of Terrorism" of terrorist attacks happened during 2000 to 2015.

3.2. User Profile

Our target audience is the general public who want to know more information, specifically about the distribution and density of global terrorist attacks happened during the past 16 years.

3.3. Data Design

3.3.1. Data Curation

Here we use the dataset of Global Terrorism Database that contains the detailed information of global terrorist attacks collected from 2000 to 2015, mainly focusing on these data variables: Name, EventId, Year, Month, Day, Country, Nkill, Gname, with the explanation of different variables is demonstrated in Fig 3.

Name	What it represents
EventId	Identical id of each terrorist attack
Year	Which Year did it occur?
Month	Which month did it occur?
Day	Which day did it occur?
Country	Which country it occurred in?
Nkill	Number of people killed in the incident
Gname	Name of the group that carried out the attack

Fig 3. Data Variables selected from the GTD with explanation

3.3.2. Data Scraping and Data Wrangling

For the whole dataset, it ranges from 1970 to 2015. Here we selected the subset of the whole dataset from 2000 to 2015 since we care more about the recent terrorist activities. After that, we group the dataset based on terrorist groups, and sorted the list based on the number of people killed by the specific terrorist group. Here we selected the top 14 terrorist groups that caused the most of total casualties during 2000 to 2015. These groups are: Taliban, Islamic State of Iraq and the Levant, Boko Haram, Al-Shabaab, Tehrik-i-Taliban Pakistan, Al-Qaida in Iraq, Al-Qaida, Al-Qaida in the Arabian Peninsula, Al-Nusrah Front, Lord's Resistance Army, Liberation Tigers of Tamil Eelam, Revolutionary Armed Forces of Colombia, Communist Party of India – Maoist, Fulani Militants. Besides, we grouped the dataset based on countries that have been attacked. The top 13 countries with most causality are Iraq, Syria, Yemen, Algeria, Afghanistan, Pakistan, India, Sri Lanka, Nigeria, Somalia, Sudan, Congo, Philippines, Russia, and United States.

3.4. Visualization Design

The default page would contain four multi-line charts and two stack bar charts. On the left column, the chart on the top would be a multi-line chart represented the information of top 14 terrorists groups. The x-axis is the year that the attacks occurred and the y-axis is the number of people killed by the specific group. The second chart is a similar multi-line chart, except that it removes the top 3 terrorist groups. The third chart is a stack bar chart that showed the number of killed people on each year for all the 14 terrorist groups. For each stack in the bar, it represents the total number of incidents claimed by a terrorist group on a specific year. Here we select the top 14 deadliest terrorist groups, for all the other terrorist groups, we put them into one group named other, and for the incidents that not claimed with any terrorist groups, we put them into another group named unknown. On the right column, the chart on the top would be a multi-line chart represented the information of top 13 victimized countries. The x-axis is the year that the attacks occurred and claimed lives and the y-axis is the number of people killed on that year in specific country. The second chart is similar, except that it removes the top 3 victimized countries. The third chart is a stack bar chart that showed the number of killed people on each year for all the 13 countries.

For the choice of color, the main background color would be white. For the multiline chart, each terrorist group would have a specific color, and the group belonging to same region would have similar color with different shades. For the stack bar chart representing the information of the terrorist groups, it would have a color scheme as follows: the 9 terrorist groups that are related with Islamic would be marked red with different shades, and the groups with more killing would have a darker red while groups with less killings would have a lighter red. The 9 terrorist groups related with Islamic are Taliban, Islamic State of Iraq and the Levant, Boko Haram, Al-Shabaab, Tehrik-i-Taliban Pakistan, Al-Qaida in Iraq, Al-Qaida, Al-Qaida in the Arabian Peninsula, Al-Nusrah Front. For other 5 groups, three of them are related to civil war or military that are marked black with different shades: Lord's Resistance

Army, Liberation Tigers of Tamil Eelam, Revolutionary Armed Forces of Colombia. For the Communist Party of India – Maoist, it is marked with dark blue, and Fulani Militants is marked with dark brown. For the stack bar chart representing the information of victimized countries, the color schema is as follows: the countries that are belonging to same regions are marked with similar colors with different shades. For example, Iraq, Syria, Yemen and Algeria are marked purple of different shades, and the darker ones means this country has more victims. Similarly, Afghanistan, Pakistan, India and Sri Lanka are marked orange with different shades, while Nigeria, Somalia, Sudan and Congo are marked blue-grey with different shades. For other countries, they are isolated so that each of them is marked with different colors.

The final product is shown as follows. The default page of the top row is shown as Fig 4.

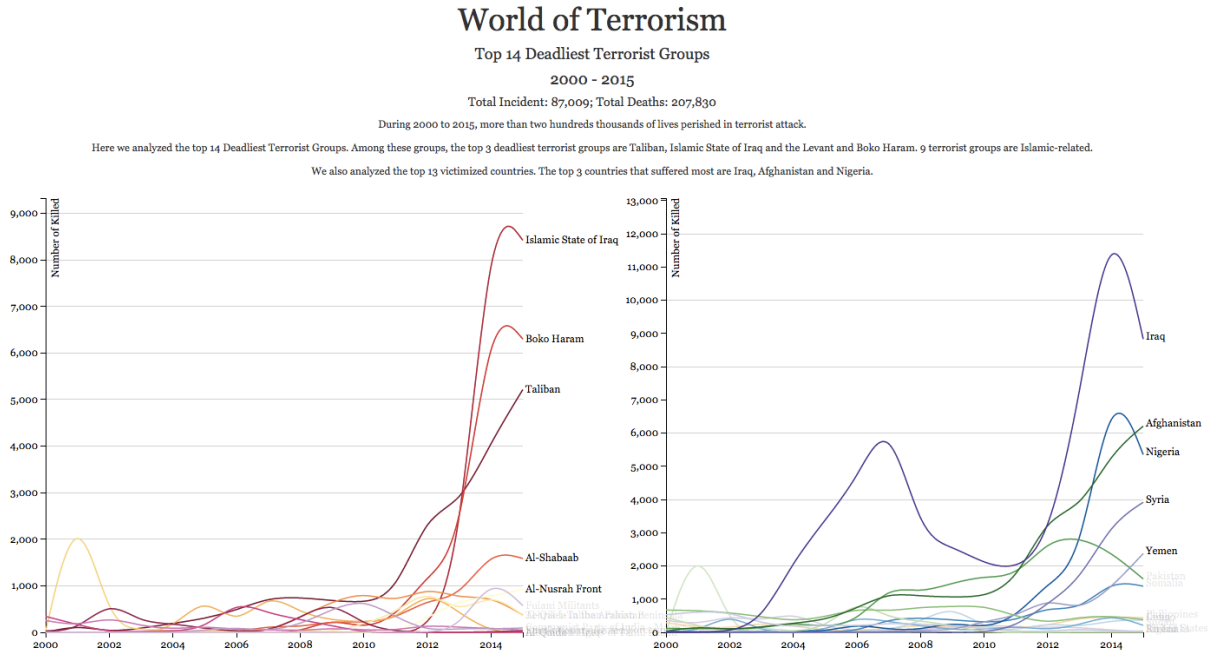


Fig 4. Top row of the visualization

Once scrolled down, it will show the second row of the graph shown as Fig 5.

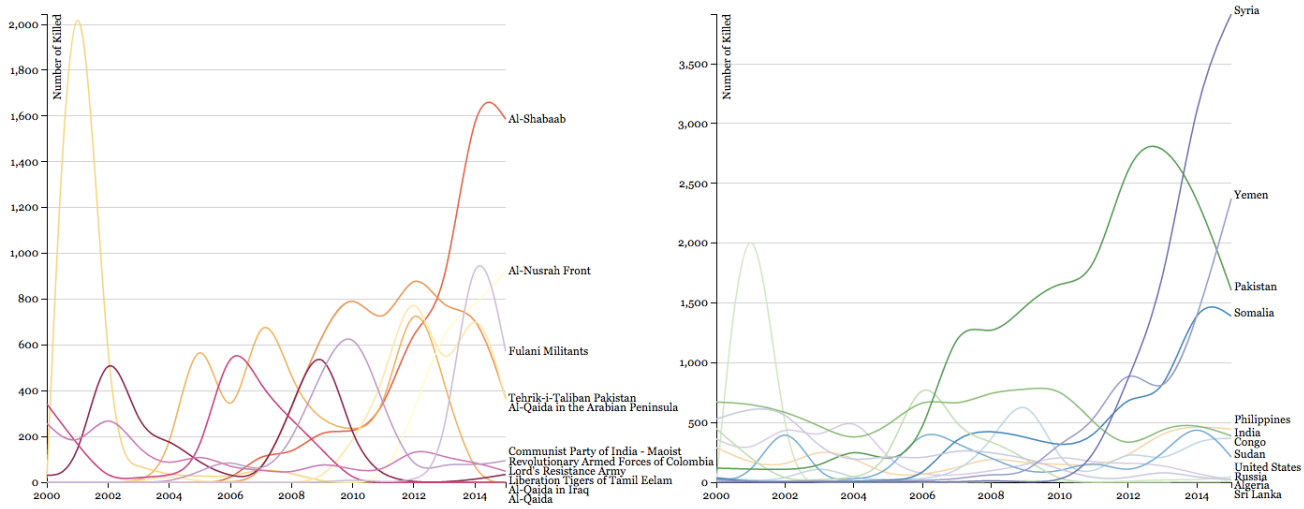


Fig 5. Second row of the visualization

On the bottom is the third part of the visualization shown as Fig 6.

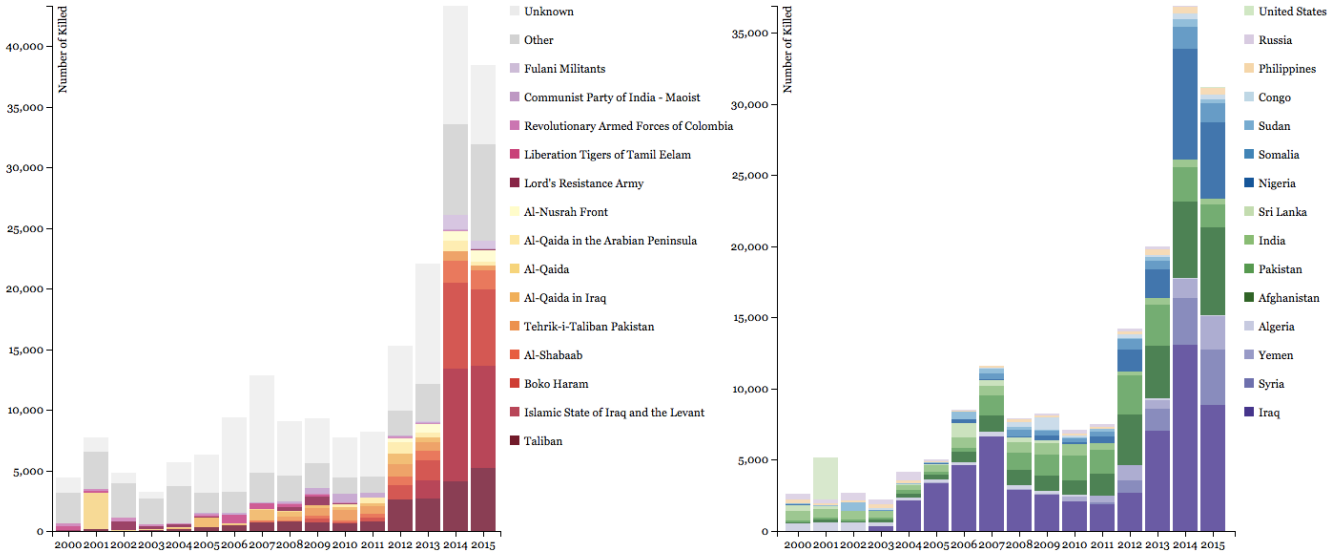


Fig 6. Third row of the visualization

3.5. Interaction Design

For the multiline chart, once hovering on the line, the stroke size of the line would be larger and the related group name or country name would appear at the end of the line (Fig 7).

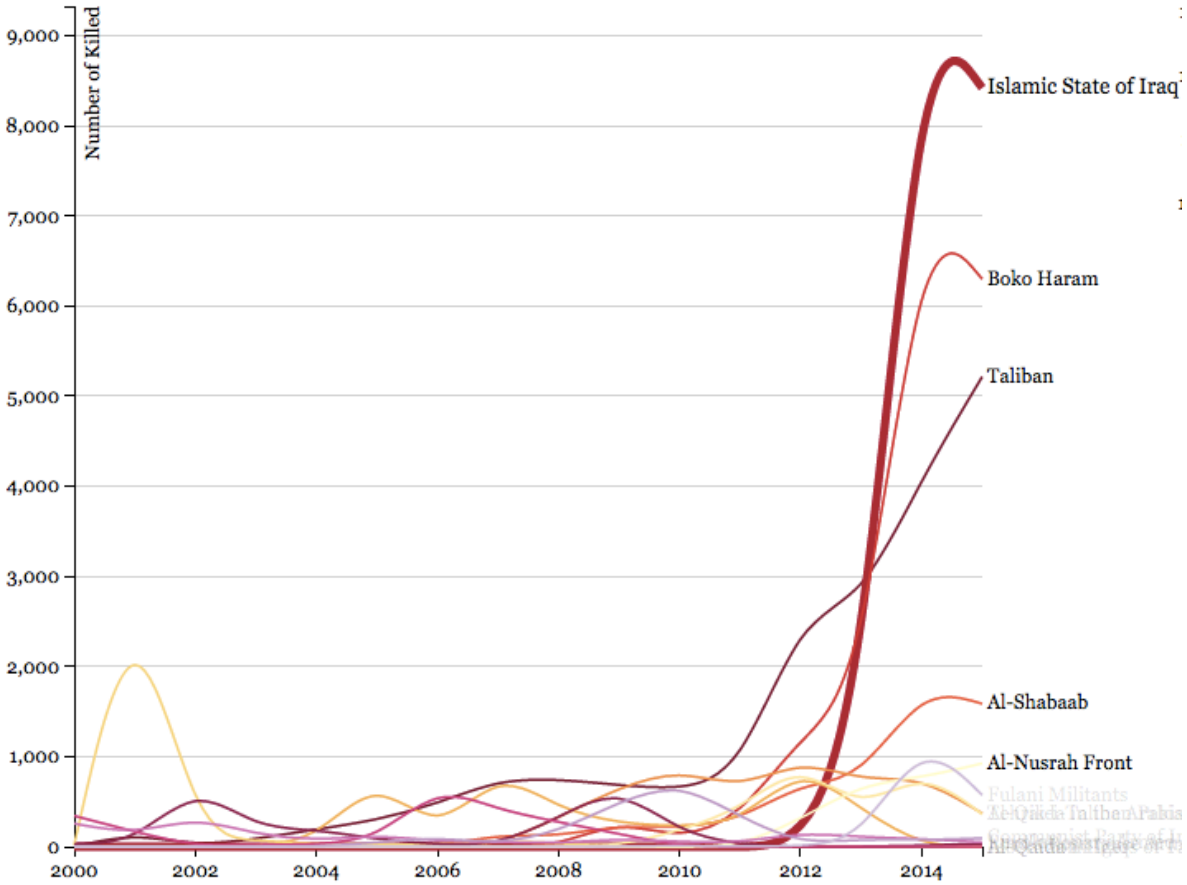


Fig 7. Interactive design of the multi-line chart

For the stack bar chart, different legends would be added on the margin of the chart. Once hovering on a specific legend, the related rectangles that have the same terrorist group name or country name would be highlighted. A tooltip is added when the mouse hovering on a stack on the stack bar chart, which displays the name of the terrorist group and the number of killed on this selected year (Fig 8).

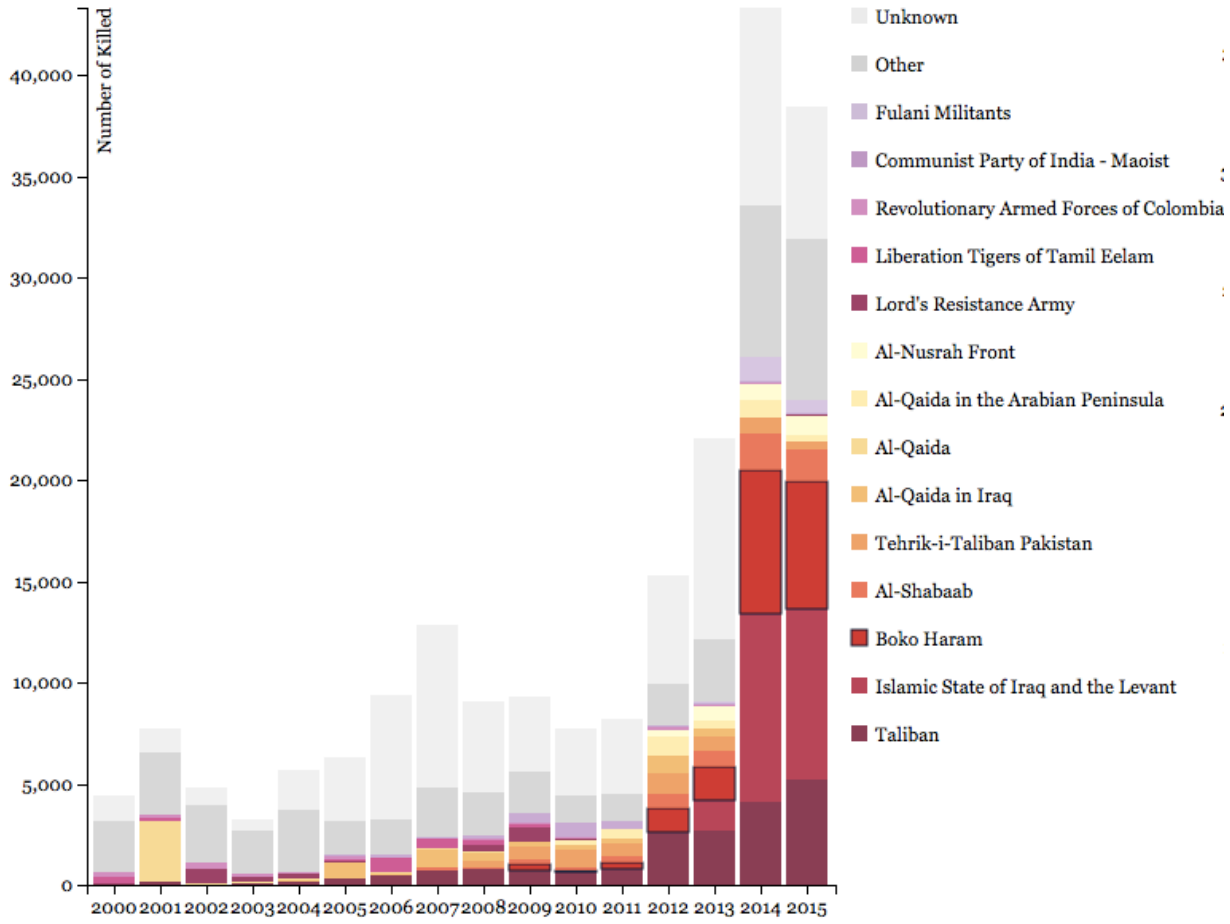


Fig 8. Interactive design of the stack bar chart.

3.6. Problems and Troubleshooting

The whole dataset contains hundreds of variables. The biggest problem is how to grasp the information that we want. Our initial idea is to use the whole dataset that contains all the information from 1970 to 2015. However, one problem is that we can't guarantee the data of old date are accurate and there are a lot of changes of names of countries and terrorist groups that may cause confusion. Another reason is that we care about the terrorist activities that happened recently to help us understand and prevent the future terror attacks. Therefore, we only selected the data from 2000 to 2015 as our main dataset. For the variables that we selected, we proposed to use number of killed and number of wounded as the indicator of the severity of the incident. However, we figured out that the data of wounded might be not accurate due to the lack of information. For example, the wounded data of 911 attacks is missing. Therefore we only use number of killed as the indicator of the severity of the terror attacks.

The UI of the final product is quite different from the original UI we proposed. The original one would contain a world map that demonstrated the specific locations of the accidents occurred. After built the first demo, it seems that the map would hide some important information that we want to demonstrate. The second version would be two stack bar charts representing the information of terrorist groups and victimized countries. However, some terrorist groups have larger number of casualties that causes the hidden information of some other terrorist groups that have smaller number of casualties. Therefore, we added multiline chart to show the trend of number of killed of each group. After that, we figured out some information of the groups with less number of killed are still hidden due to the comparable larger number of the top 3 groups, therefore we create additional charts to show the information hidden underneath by removing the top 3 groups.

Another problem is how to choose appropriate color. Here we have 14 terrorist groups and 13 victimized countries. That means there are too many colors on the page, which may lead to confusion. Therefore, we tried to group these terrorist groups and countries based on some other rules. For example, the terrorist groups that are related to Islamic are marked with similar color. The countries that are located within the same region are marked with similar color of different shades. This would reduce the over redundant effects caused by too many colors.

4. Conclusions and Future Work

The Global Terrorism Database is a comparable complete open source library that contains information of terrorist attacks both domestically and internationally from 1970 to 2015. It is a really interesting topic to explore this database and discover inner patterns in it. Here we implemented a data visualization project to figure out the geo-temporal pattern of global terrorist attacks from 2000 to 2015. We find that more than two hundreds thousands of lives perished in terrorist attack during 2000 to 2015. Among the top 14 deadliest terrorist groups, the top 3 deadliest terrorist groups are Taliban, Islamic State of Iraq and the Levant and Boko Haram. 9 terrorist groups are Islamic-related. Among the top 13 victimized countries, the three countries that suffered most are Iraq, Afghanistan and Nigeria. More future work would be done, such as the implementation of visualization effects between the terrorist groups and the victimized countries, improving the rendering efficiency and do data analysis to gain some inner sights from our visualization results.

Bibliography

[1] Guo, Diansheng, Ke Liao, and Michael Morgan. "Visualizing patterns in a global terrorism incident database." *Environment and Planning B: Planning and Design* 34.5 (2007): 767-784.

[2] Hoffman, Bruce. Inside terrorism. Columbia University Press, 2006.

[3] Webb, Jennifer J., and Susan L. Cutter. "The geography of US terrorist incidents, 1970–2004." *Terrorism and Political Violence* 21.3 (2009): 428-449.

[4] Lum C, Kennedy L W. Evidence-based counterterrorism policy[M]//Evidence-based counterterrorism policy. Springer New York, 2012: 3-9.

[5] National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2016). Global Terrorism Database [Data file]. Retrieved from https://www.start.umd.edu/gtd

[6] Lazer, David, et al. "Life in the network: the coming age of computational social science." *Science (New York, NY)* 323.5915 (2009): 721.

[7] Conte, Rosaria, et al. "Manifesto of computational social science." *The European Physical Journal Special Topics* 214.1 (2012): 325-346.

[8] Cioffi-Revilla, Claudio. *Introduction to computational social science: principles and applications*. Springer Science & Business Media, 2013.

[9] Cioffi-Revilla, Claudio. "Bigger Computational Social Science: Data, Theories, Models, and Simulations--Not Just Big Data." *Theories, Models, and Simulations--Not Just Big Data (May 24, 2016)* (2016).

[10] Onnela, J-P., et al. "Structure and tie strengths in mobile communication networks." *Proceedings of the National Academy of Sciences* 104.18 (2007): 7332-7336.

[11] Balcan, Duygu, et al. "Multiscale mobility networks and the spatial spreading of infectious diseases." *Proceedings of the National Academy of Sciences* 106.51 (2009): 21484-21489.

[12] Batty, Michael. "The size, scale, and shape of cities." *science* 319.5864 (2008): 769-771.

[13] Chainey, Spencer, and Jerry Ratcliffe. *GIS and crime mapping*. John Wiley & Sons, 2013.

[14] Lillo, Fabrizio, et al. "Specialization and herding behavior of trading firms in a financial market." *New Journal of Physics* 10.4 (2008): 043019.

[15] Dzogang, Fabon, et al. "Discovering Periodic Patterns in Historical News." *PloS one* 11.11 (2016): e0165736.

[16] Gray, Jonathan, Lucy Chambers, and Liliana Bounegru. *The data journalism handbook*. " O'Reilly Media, Inc.", 2012.

[17] Bostock, Michael, Vadim Ogievetsky, and Jeffrey Heer. "D³ data-driven documents." *IEEE transactions on visualization and computer graphics* 17.12 (2011): 2301-2309.

[18] Bostock, Michael. "D3. js." *Data Driven Documents* (2012).

[19] Bostock, Michael, and Jeffrey Heer. "Protovis: A graphical toolkit for visualization." *IEEE transactions on visualization and computer graphics* 15.6 (2009): 1121-1128.

[20] The New York Times. “Presidential Election Results: Donald J. Trump Wins” http://www.nytimes.com/elections/results/president.

[21] Comparison of JavaScript charting frameworks. https://en.wikipedia.org/wiki/Comparison_of_JavaScript_charting_frameworks

[22] Chaidas, Angelos. Choosing a JavaScript charting library in 2016. https://www.developereconomics.com/javascript-charting-library-2016

[23] Watkins, Christopher. 12 Best Charting Libraries for Web Developers. http://blog.udacity.com/2016/03/12-best-charting-libraries-for-web-developers.html

[24] Tufte, Edward R., and P. R. Graves-Morris. *The visual display of quantitative information*. Vol. 2. No. 9. Cheshire, CT: Graphics press, 1983.

[25] LaFree, Gary, and Laura Dugan. "Introducing the global terrorism database." *Terrorism and Political Violence* 19.2 (2007): 181-204.

[26] Jones, Josh, et al. "Visualizing uncertainty for geographical information in the global terrorism database." *SPIE Defense and Security Symposium*. International Society for Optics and Photonics, 2008.

[27] Godwin, Alex, et al. "Visual analysis of entity relationships in the Global Terrorism Database." *SPIE Defense and Security Symposium*. International Society for Optics and Photonics, 2008.

[28] Lee, Joonghoon. "Exploring global terrorism data: a web-based visualization of temporal data." *Crossroads* 15.2 (2008): 7-14.

[29] Spaaij, Ramon. *Understanding lone wolf terrorism: Global patterns, motivations and prevention*. Springer Science & Business Media, 2011.

[30] Drakos, Konstantinos. "Terrorism activity, investor sentiment, and stock returns." *Review of Financial Economics* 19.3 (2010): 128-135.

[31] Drakos, Konstantinos. "Terrorism activity, investor sentiment, and stock returns." *Review of Financial Economics* 19.3 (2010): 128-135.

Appendix I

Data Sets

The whole original dataset can be downloaded from the Global Terrorism Database as an excel file, the link is https://www.start.umd.edu/gtd/.

The table of the top 13 victimized countries is as follows:

year	Sri Lanka	Congo	Sudan	United States	Russia	Philippines	Algeria	Somalia	Yemen	India	Syria	Nigeria	Pakistan	Afghanistan	Iraq
2000	449	40	27	0	362	290	528	36	19	671	0	20	118	38	10
2001	194	15	13	3003	228	154	624	3	2	658	0	3	109	174	9
2002	3	16	588	4	512	125	615	6	4	593	0	28	105	74	10
2003	23	154	0	0	328	288	269	7	16	472	0	28	119	163	347
2004	30	22	28	0	607	206	166	6	0	334	4	41	304	275	2146
2005	125	22	54	0	156	49	228	20	12	462	0	19	150	367	3380
2006	1012	23	484	1	57	60	185	22	14	722	5	254	314	731	4612
2007	410	72	317	0	55	115	279	408	24	626	0	82	1406	1197	6648
2008	352	319	205	2	95	220	249	436	74	763	18	72	1184	1089	2863
2009	201	824	77	18	141	152	204	381	47	774	0	316	1487	1064	2583
2010	2	118	81	4	231	155	109	295	345	812	0	117	1695	1058	2071
2011	0	50	193	0	160	127	25	344	461	484	163	447	1655	1521	1864
2012	0	288	60	7	161	210	30	775	1056	264	876	1504	2783	3521	2688
2013	0	150	220	22	148	432	101	646	624	467	1558	2013	2874	3696	7038
2014	3	366	546	19	67	472	37	1581	1348	490	3300	7773	2412	5413	13075
2015	3	367	210	44	21	444	21	1389	2373	387	3916	5351	1606	6208	8831

The table of top 14 deadliest terrorist groups is as follows:

year	Revolutionary Armed Forces of Colombia	Fulani Militants	Liberation Tigers of Tamil Eelam	Communist Party of India - Maoist	Lord's Resistance Army	Al-Nusrah Front	Al-Qaida in the Arabian Peninsula	Al-Qaida	Al-Qaida in Iraq	Tehrik-i-Taliban Pakistan	Al-Shabaab	Boko Haram	Islamic State of Iraq and the Levant	Taliban	Other	Unknown
2000	256	0	342	0	30	0	0	19	0	0	0	0	0	0	2517	1253
2001	133	0	164	0	36	0	0	2995	0	0	0	0	0	153	3078	1174
2002	335	0	3	0	699	0	0	100	0	0	0	0	0	5	2823	828
2003	132	0	23	0	201	0	0	73	0	0	0	0	0	86	2134	620
2004	65	0	24	0	193	0	33	28	69	0	0	0	0	192	3061	2046
2005	131	0	75	45	80	0	0	29	806	0	0	0	0	293	1657	3179
2006	59	0	684	108	23	0	4	20	99	0	0	0	0	476	1736	6152
2007	55	0	405	27	15	0	0	63	869	14	135	0	0	745	2483	8021
2008	34	0	275	163	322	0	0	41	416	324	112	0	0	750	2134	4522
2009	91	0	143	500	680	0	6	0	258	658	240	304	0	683	2013	3695
2010	53	13	0	722	153	0	178	0	223	863	206	72	0	627	1331	3278
2011	40	0	0	329	20	0	408	1	257	634	294	322	0	811	1359	3717
2012	159	6	0	25	0	303	963	0	917	974	706	1227	0	2574	2068	5397
2013	106	66	0	86	0	695	370	0	414	728	737	1713	1439	2714	3104	9899
2014	88	1230	0	72	12	765	866	0	0	780	1777	7093	9326	4095	7465	9763
2015	48	572	0	94	34	924	360	0	0	368	1586	6299	8420	5215	7998	6504

Appendix II: As Needed

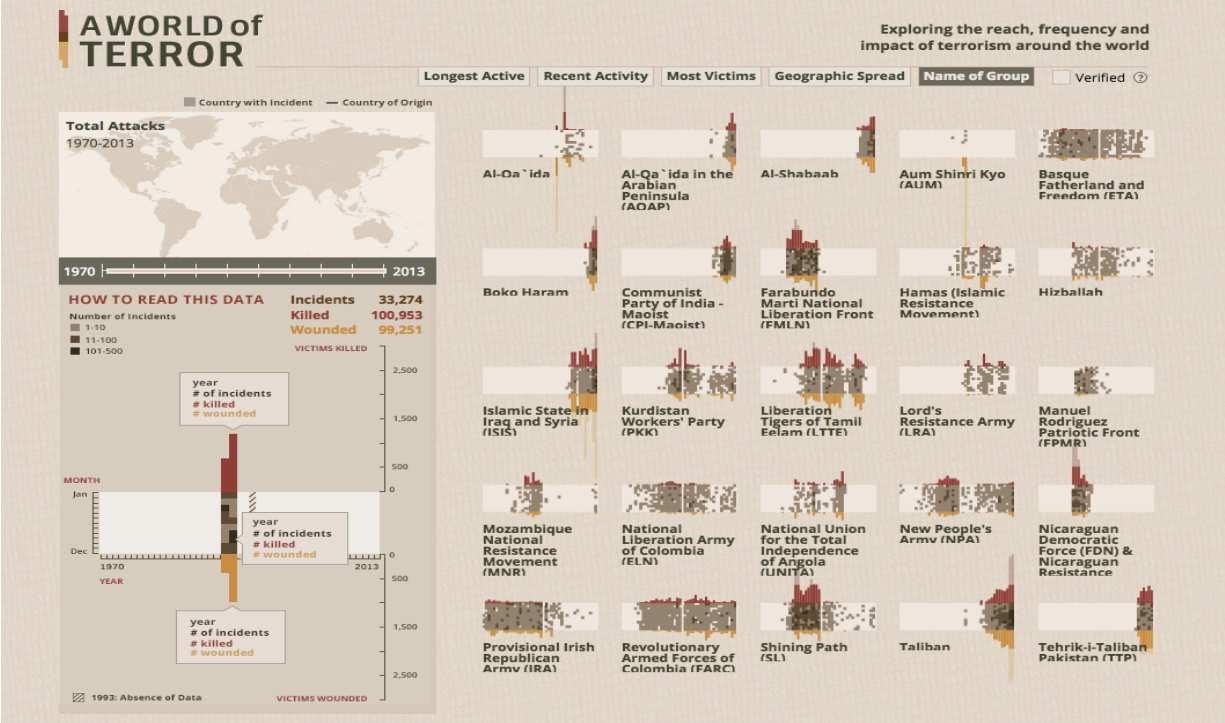
Code sample.

For the whole code, the link is https://github.com/erfan00/world_of_terror_6

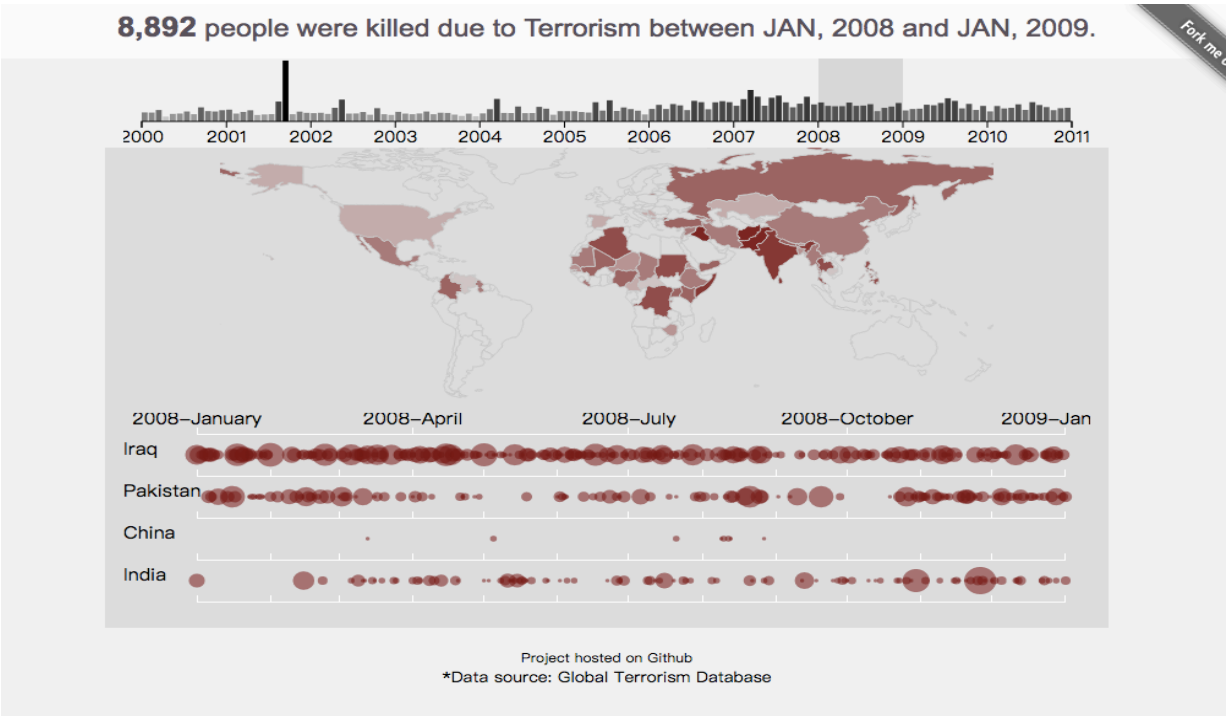
```
var margin = {top: 20, right: 280, bottom: 30, left: 80},
  width  = 830 - margin.left - margin.right,
  height = 600  - margin.top  - margin.bottom;

var x = d3.scale.ordinal()
  .rangeRoundBands([0, width], .1);

var y = d3.scale.linear()
  .rangeRound([height, 0]);

var xAxis = d3.svg.axis()
  .scale(x)
  .orient("bottom");

var yAxis = d3.svg.axis()
  .scale(y)
  .orient("left");


var color = d3.scale.ordinal()
        .range(["#5B121D","#6D1623","#7F1929","#911D2E", "#A32034", "#B5243A",  "#C7273F", "#CC3A50", "#D14E61",
        "#252525", "#525252", "#737373","#46465A", "#716168","#d3d3d3","#ededed"]);


var svg = d3.select("#stack_bar_chart").append("svg")
  .attr("width",  width  + margin.left + margin.right)
  .attr("height", height + margin.top  + margin.bottom)
.append("g")
  .attr("transform", "translate(" + margin.left + "," + margin.top + ")");


d3.csv("data/fixed_year_data_output_3_reverse.csv", function (error, data) {

    var labelVar = 'year';
    var varNames = d3.keys(data[0]).filter(function (key) { return key !== labelVar;});
    color.domain(varNames);

    data.forEach(function (d) {
      var y0 = 0;
      d.mapping = varNames.map(function (name) {
        return {
          name: name,
          label: d[labelVar],
          y0: y0,
          y1: y0 += +d[name]
        };
      });
      d.total = d.mapping[d.mapping.length - 1].y1;
    });

    x.domain(data.map(function (d) { return d.year; }));
    y.domain([0, d3.max(data, function (d) { return d.total; })]);
```

Relevant visualization links:

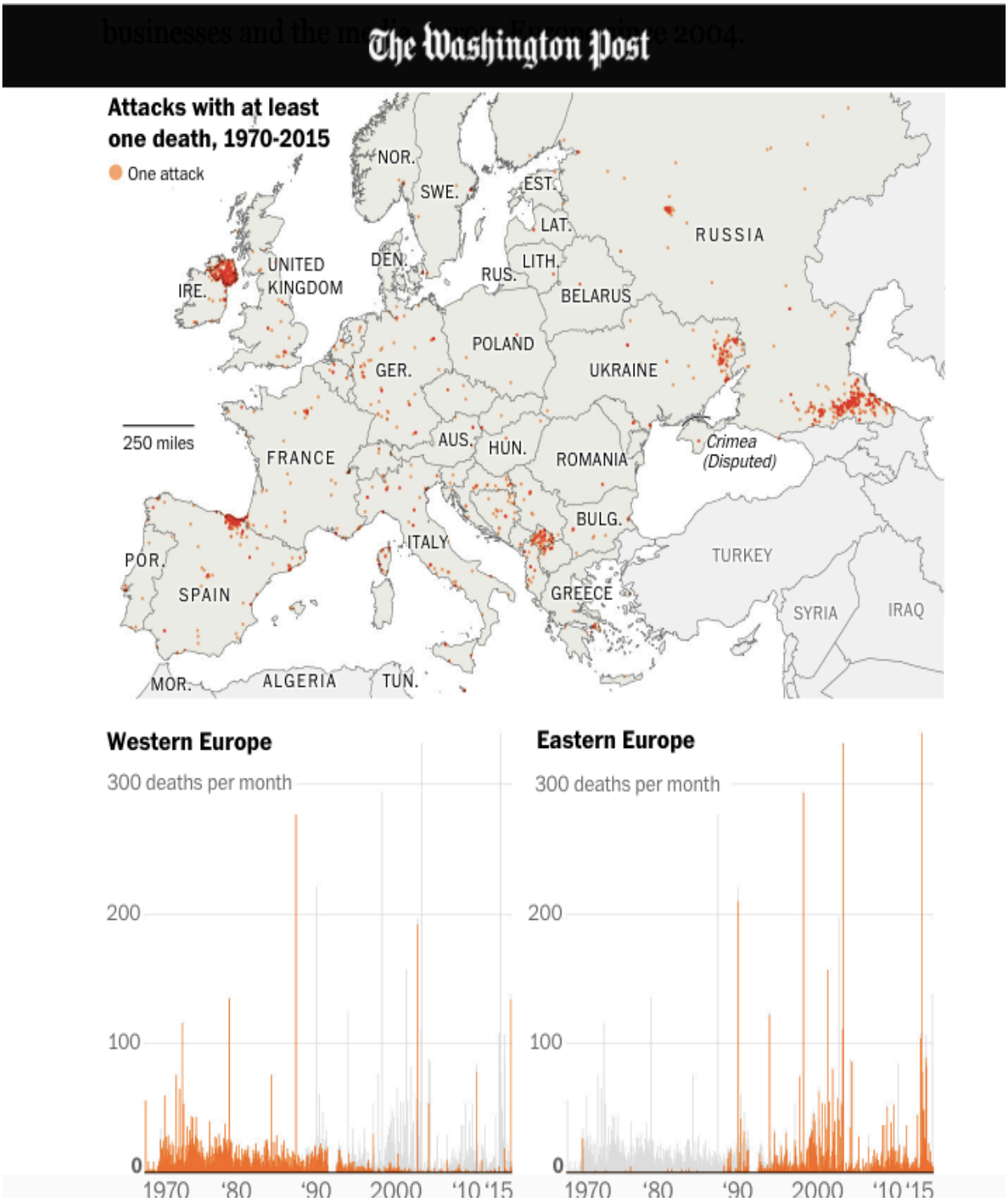
A World of Terror: http://terror.periscopic.com/



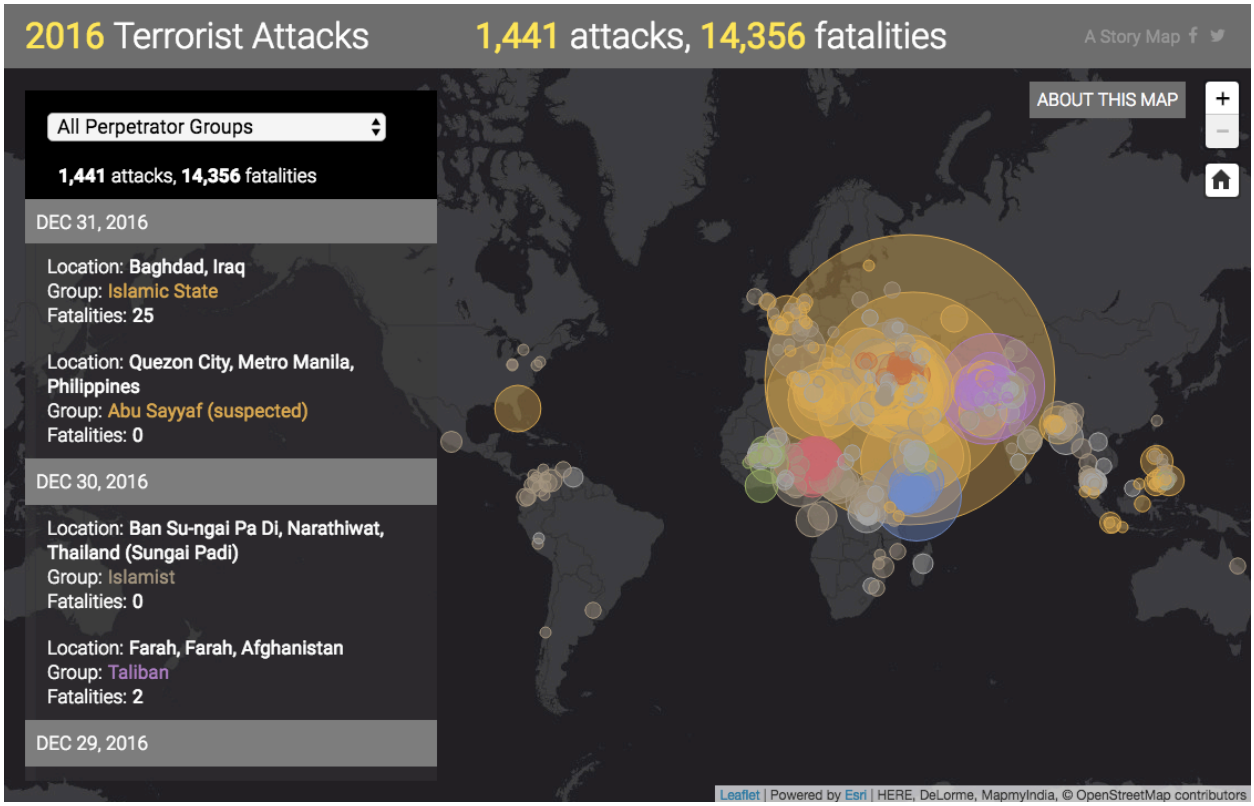
Global Terrorism Visualization: http://cse512-14w.github.io/a3-chaoyu-aniket/



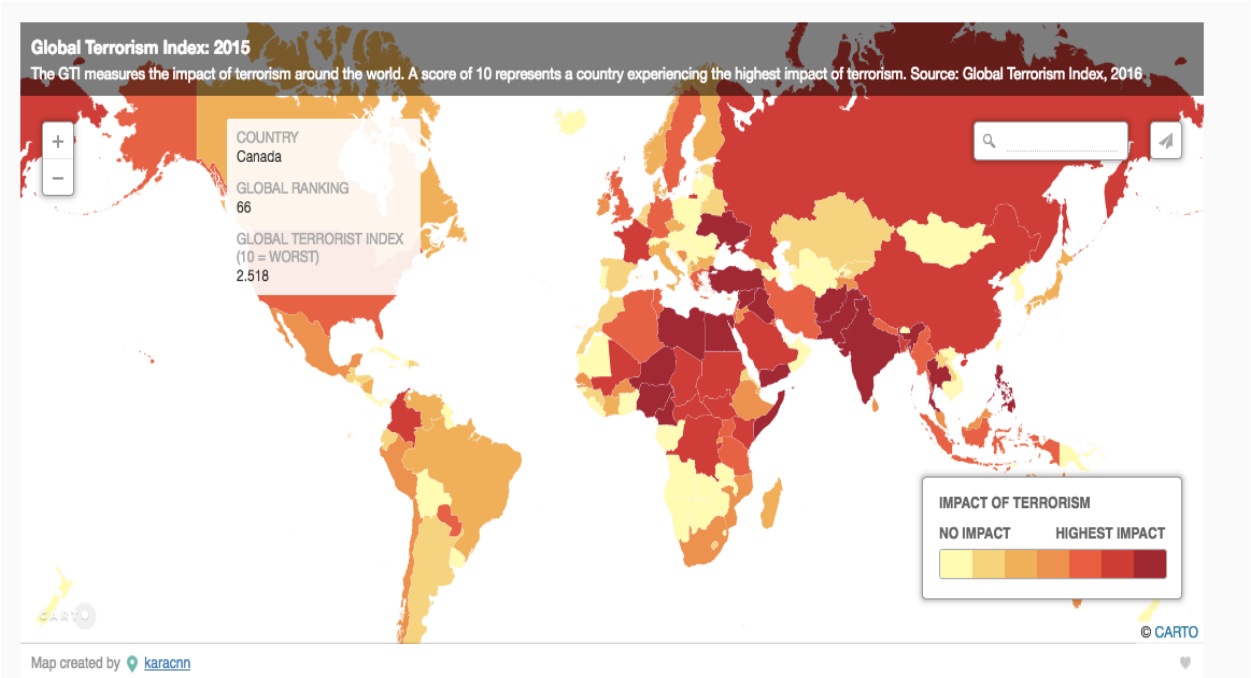
45 years of terrorist attacks in Europe, visualized: https://www.washingtonpost.com/graphics/world/a-history-of-terrorism-in-europe/



2016 Terrorist Attacks: https://storymaps.esri.com/stories/terrorist-attacks/?year=2016



Global Terrorism Report: http://www.cnn.com/2016/11/16/world/global-terrorism-report/



Related Terrorist Databases:

Terrorist Database Name	Link
Global Terrorism Database (GTD)	https://www.start.umd.edu/gtd/
International Terrorism: Attributes of Terrorist Events (ITERATE)	http://library.duke.edu/data/collections/iterate
Rand Database of Worldwide Terrorism Incidents (RDWTI)	http://www.rand.org/nsrd/projects/terrorism-incidents.html
Terrorism in Western Europe: Events Data (TWEED)	http://folk.uib.no/sspje/tweed.htm