FIT5147　　Visualization Project

Terrorist Attacks Visualizer and Analyzer

3rd, June 2018

Fong Cheok Kin, 28692381

TABLE OF CONTENT

1. Introduction

In this visualization project, the database from Global Terrorism Database (GTD)[1] is used. The application visualizes the world terrorist events between 1970 and 2016. It contains two parts, with the first part called "Attacks Visualizer" and the other is named "Attacks Analyzer".

The Attacks Visualizer mainly visualizes how terrorist events threatening everyone's life globally. It highlights the victims including the number of deaths and wounded persons in the events. The target audience is the general public, especially who is concerned about safety issue related to terrorist attacks. This visualization is built for general users because it allows them to target their home country or any other place to explore the attacks.
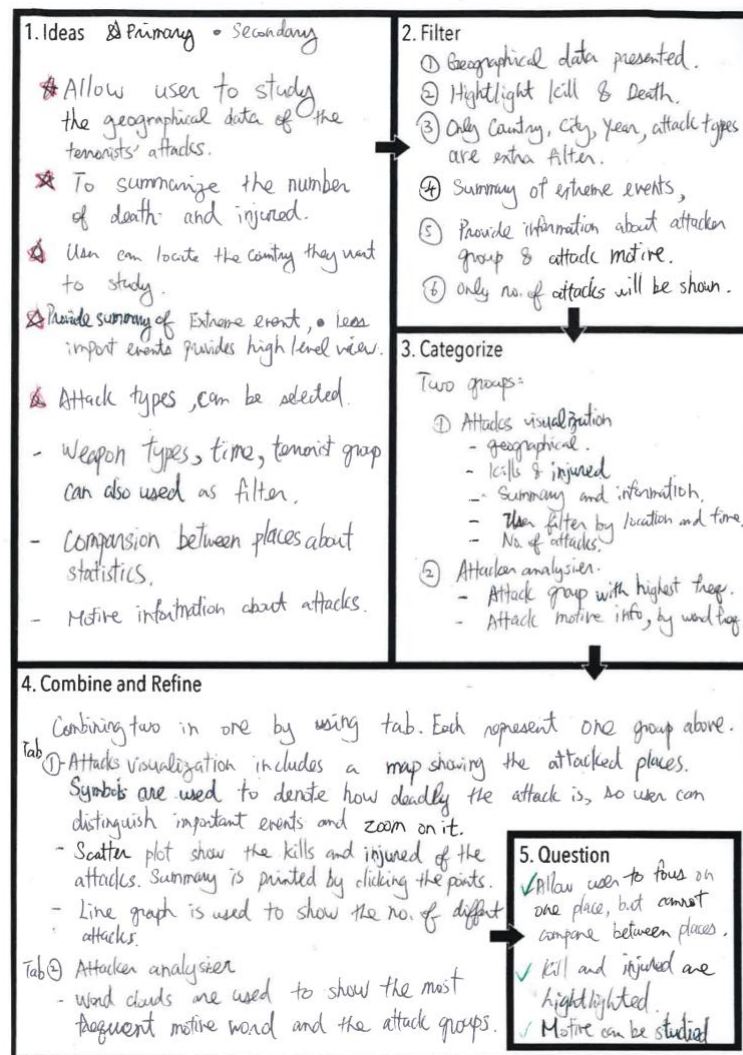
On the other hand, Attacks Analyzer is for those who would like to extract information about the motive and terrorist groups. This visualization allows users to visualize the reported motive in a word cloud, and the names of the related terrorist groups are also highlighted by another cloud. User can identify the threatening organizations and explore their motives based on the clouds. Potential target audience could be government, who can use it to identity the largest local terrorist group, also the groups responsible for certain topic of motive.
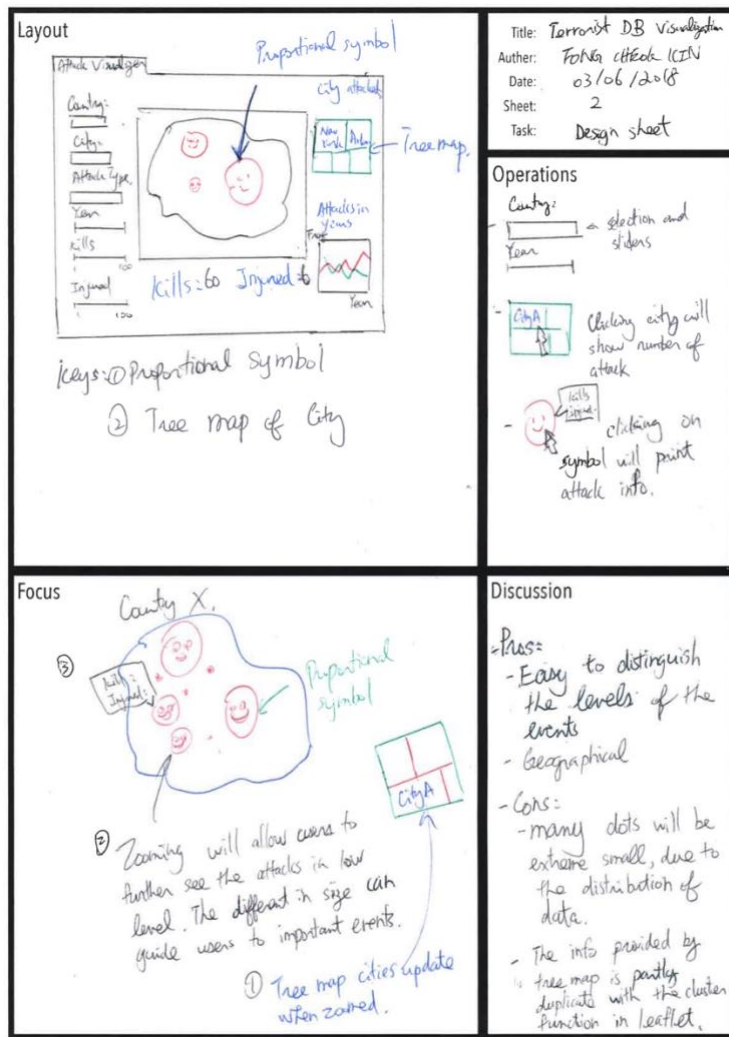
2. Design

2.1 Design Process

The design of the visualizations makes use of the 5-design sheets methodology. The sheet 1 brainstorms the ideas and 2,3,4 sheets are used to proposed designs. The sheet 5 represents my final selection, which combines part of the features from 2, 3, 4 sheets. Considering the readability, a table will be used for each sheet to clarify and summarize the points in the sheets.
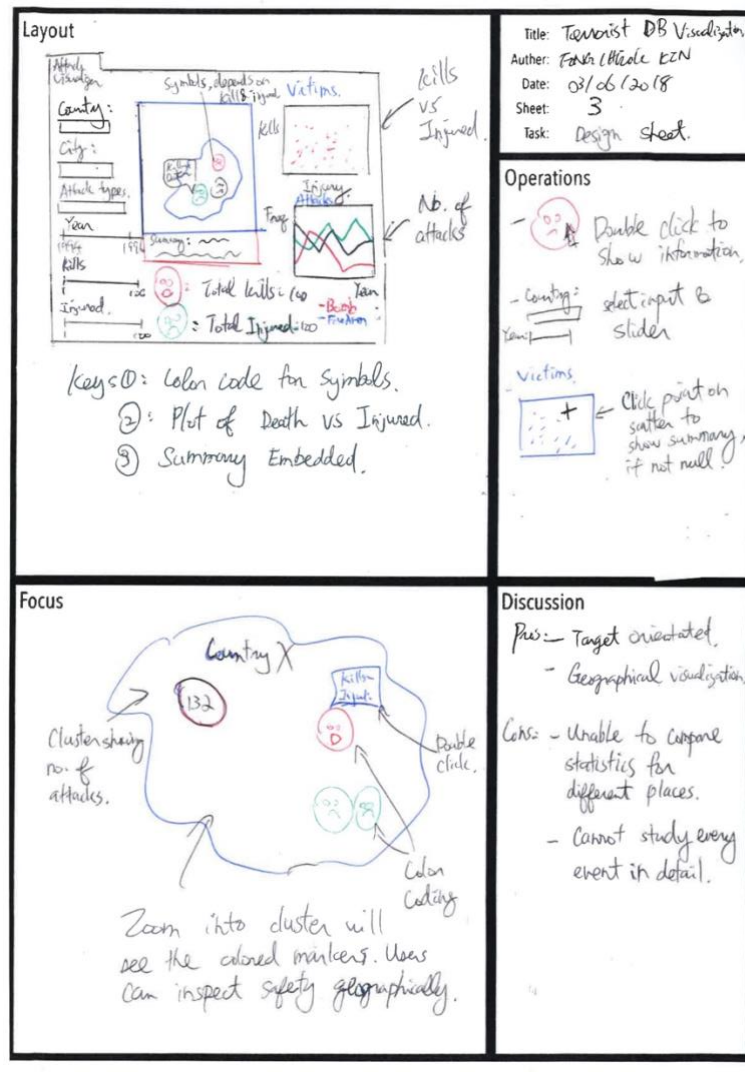
Figure 1Design Sheet 1

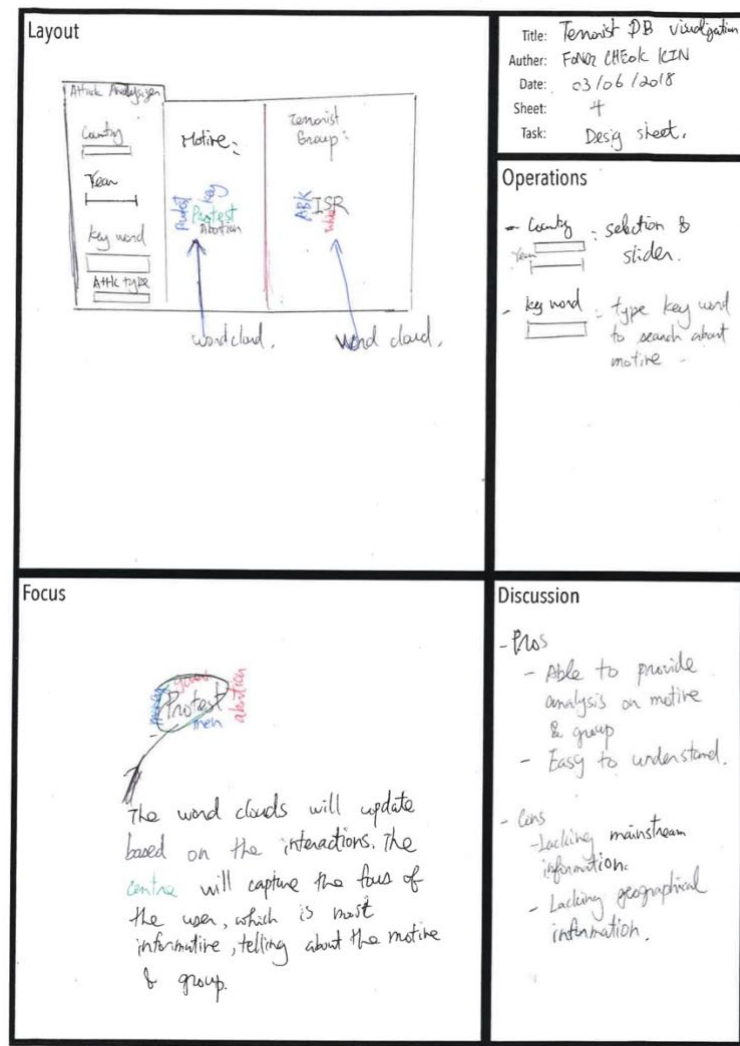|	
1. Ideas	- Allow user to study geographical data of terrorists' attacks - Summarize the number of deaths and injured - Locate the countries - Provide summary of extreme events, if recorded - Attack types can be selected - Weapon types, time, terrorist group as filters - Comparison between places about statistics - Attack frequency - Motive information can be explored
2. Filter	- Geographical data - Highlight kill and injured - Only use country, city, year, attack types as selection filters - Summary of extreme events - Provide information about terrorist group and motive - Number of attacks will be plotted
3. Categorize	1st Group (about attacks): - Geographical data - Kills and injured - summary of attack - Number of attacks 2st Group (about attacker): - Motive information - Terrorist group name
4. Combine and Refine	One visualization will be used to present one group. Each visualisation includes the ideas in its group. The two visualizations are then combined into one using tabs. - Tab1 - Attacks visualization includes a map (geographical), scatter (kill and injured), texts (summary), and line graph (number of attacks) - Tab2 - Attacks analyser includes two word clouds to show frequent motive word and terrorist group.
5. Question	The visualization answers: - which locations are dangerous. - how many people are killed and wounded. - motives about the attacks and the corresponding terrorist groups.

Figure 2 Design Sheet 2

1. Layout	Key elements:
	• Proportional symbol map, size dependent to the number of victims.
	• Tree map showing cities attacked.
	• Line graph showing attacks.
2. Operation	• Country, city, attack type, year, death, and injured provided for user inputs.
	• Clicking the tree map will show summary statistics about attacks.
	• Clicking the symbol on the geographical map will pop up information about attacks.
3. Focus	• The tree map and the geographical map are the focus.
	• Zooming into the map further shows how attacks are distributed on the map. The size of the symbols also guides users to locate deadly events.
	• Tree map and the map will update automatically based on the country selected.
4. Discussions	Pros:
	• Easy to distinguish the levels of the events.
	• Geographical representation, so it is target-orientated.
	Cons:
	• Many extreme small points due to skewed distribution.
	• Information provided by the tree map is partly duplicate by the clusters on map.
	• Motive and terrorist group information is not included

Figure 3 Design Sheet 3


Figure 3 Design Sheet 3

1. Layout	Key elements: • Colour codes are used to define the level of attacks on map. Level is dependent on the number of victims. • Scatter plot is used to plot the death and injured. • Line graph shows number of attacks.
2. Operation	• Country, city, attack type, year, death, and injured provided for user inputs. • Clicking the points on scatter plot will print summary and motive. • Clicking the symbols on the map also provide high level information.
3. Focus	• The scatter and the geographical map are focus. • Zoom into the map will split the clusters and coloured symbols will be plotted. • Scatter and map will update automatically based on the country selected.
4. Discussions	Pros: • Geographical representation • Colours help users to distinguish levels of the events. • Summary of extreme events can be printed Cons: • Cannot compare statistics for different places easily. • Motive and terrorist group information is not included.

Figure 4 Design sheet 4

Figure 4 Design sheet 4

1. Layout	Key elements:
	• Word Cloud 1: Showing words about the motive of the attacks.
	• Word Cloud 2: Showing related terrorist groups.
2. Operation	• Country, year, attack type, keyword provided for user inputs
	• User can enter any word in text box to search certain motive word.
3. Focus	• After user inputs, the central areas of the word clouds are the most important information about the motives and the related terrorist groups. User can observe the pattern for different inputs.
4. Discussions	Pros:
	• Able to provide analysis on motive and terrorist groups.
	• Interface is straight forward, easy to understand.
	Cons:
	• Does not include mainstream information, such as date, victims.
	• Lacking geographical information.

Figure 5 Design Sheet 5
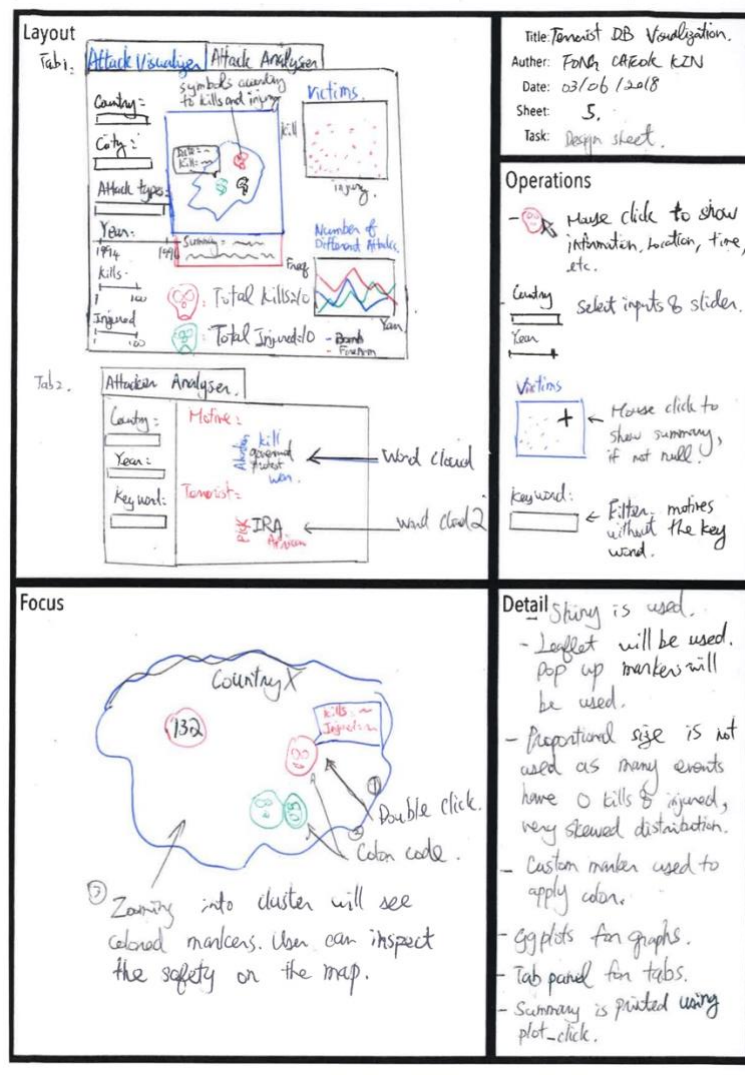
Selection and combination for Sheet 5:

For the geographical map, a mix of coloured and fixed size symbols are chosen instead of proportional symbols, as the median of death of the attacks is 1. Most of the points will become extremely small and equal size, which loses the significance of proportional size. Secondly, the scatter plot between death and wounded is chosen instead of the tree map. Considering the theme of this visualization, which aims to highlight the victims in the events, the scatter has an advantage on it and also allows user to observe extreme events. Lastly, the design of sheet 3 is also valuable in term of motive analysis, so it is embedded in the visualisation using a tab.

The following table summarizes my realization sheet 5.

1. Layout	Key elements: Tab one: • One geographical map using a mixed of coloured and fixed size symbols • One scatter plot showing death and injured. • One line graph showing number of attacks in years. • One text box printing details of selected attack. Tab two: • One word cloud for motive word • One word cloud for reported terrorist group
2. Operation	Tab one: • Country, city, attack type, year, death, and injured provided for user inputs. • Clicking points on scatter plot will print summary and motive. • Clicking the symbols on map also provides high level information. Tab two: • Country, year, attack type, keyword provided for user inputs • User can enter any word in text box to search certain motive word.
3. Focus	Tab one: • The scatter and the geographical map are the focus. • Zoom into the map will split the clusters and coloured symbols will be plotted. Tab two: • Central area of word cloud highlight important words and groups.

4. Detail	• Core tools would be shiny, leaflet, and ggplot for graphics.
	• Theme colour is black to match the topic.
	• Customized marker is used to apply colour and symbolize victims.
	• Tab panel is used to combine tabs.
	• Summary printed using "plot_click" function
	• 4 levels of colour codes
	    o level 4 is the highest, attacks with equal to or more than 10 kills.
	    o level 3, attacks with equal to or more than 1 kill.
	    o level 2, attacks with no kill but injured.
	    o level 1, attacks with no kill or injured.

2.2   Theme and symbols

The topic of the visualization is terrorist attacks and it is designed to highlights the victims in the events. Black is chosen as the theme colour throughout the design, as it represents a feeling of danger and fear. The cyborg theme is chosen because it uses black and grey colours, which matches the theme. On the other side, the tile of the leaflet map uses the Stamen design for a similar reason.

Figure 6 Shiny Cyborg theme   *Figure 7 Leaflet Stamen Tiles*
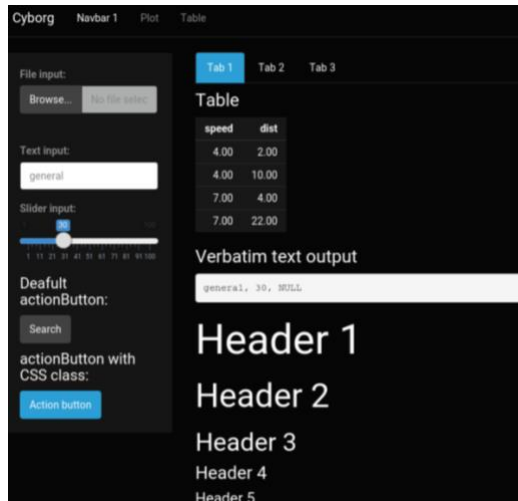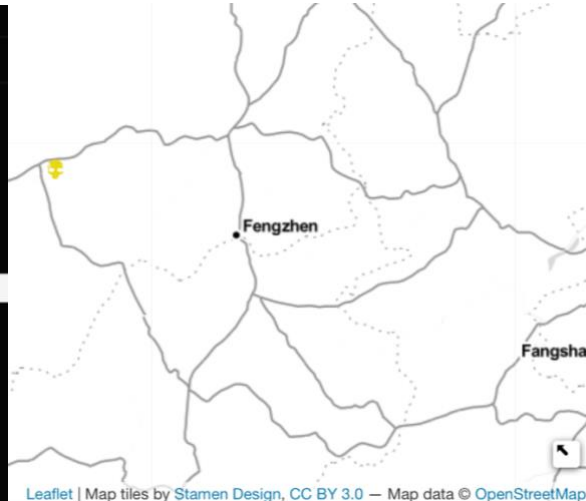

The symbols representing the attacks are customized by using skeleton head, to deliver a message that the events are deadly. The colour of different levels changed from red to green, as red usually represents the most dangerous level. In implementation, the size of red skeletons is also adjusted to a bit larger to emphasize the killing events.

Figure 8 Event symbols
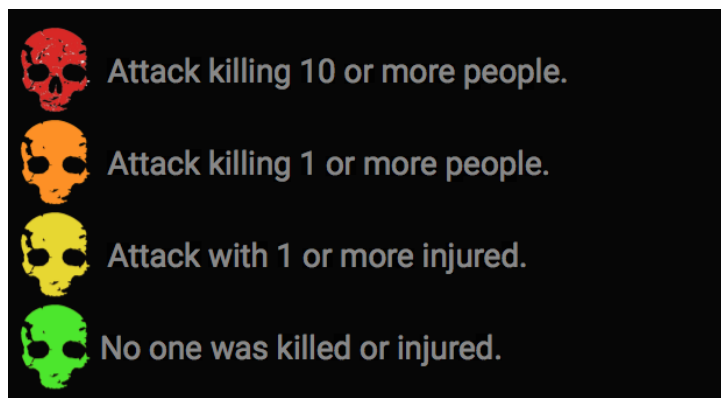
3. Implementation

3.1 Software and system

The visualization is built by R script, the detail of the libraries used will be mentioned in next section. One important note is that the source code is developed in mac OS system, part of the coding involves fixing errors caused by encoding issue. Therefore, the source code is not runnable in windows machines.

3.2 Libraries and usage

The following libraries are used:

shiny, scales, shinythemes, leaflet, ggplot2, dplyr, tm, wordcloud, RColorBrewer are used.

shiny: to build the interactive interface for user inputs such as country, and year. This system is highly interactive with users. It is chosen because input widgets can be easily applied by shiny. Besides, the map is plotted by leaflet and can be integrated with shiny.

leaflet: to plot the geographical location of the attacks on a map. customized markers and pop up are also plotted by leaflet. it works well together with shiny, and provide powerful features such as clusters and mini-map.

ggplot2: to plot scatter plot and line graph. It is easy to plot data stored in a data frame of R, the ggplot graphs also work together with the interactive functions provided by shiny, such as plot_click.

tm, **wordcloud**: wordcloud is used to generate word cloud, tm is used to pre-process the text such as removing stop words.

dplyr: to mutate data frame easily.

scales, **shinythemes**, **RColorBrewer**: these three libraries are used for adjusting styles. scales, shinythemes, RColorBrewer are used for catering plots, system interface, word cloud style respectively.

Figure 9 library usage

12

4. User Guide

Since the GTD has 170,350 records in total. It is found that leaflet has its upper limit for loading the markers on the map. Therefore, users must not select a large data size for the visualization and should follow the user guide, although prompt is given when data size is more than 62000 items. (The threshold is approximated by testing). Besides, selecting a large data set will degrade the processing speed.

4.1   Attacks Visualizer

When user launches the visualization, the default interface is shown below. The next step is to follow the selection procedure in order to explore the visualization.

Figure 10 Attack Visualizer Default interface
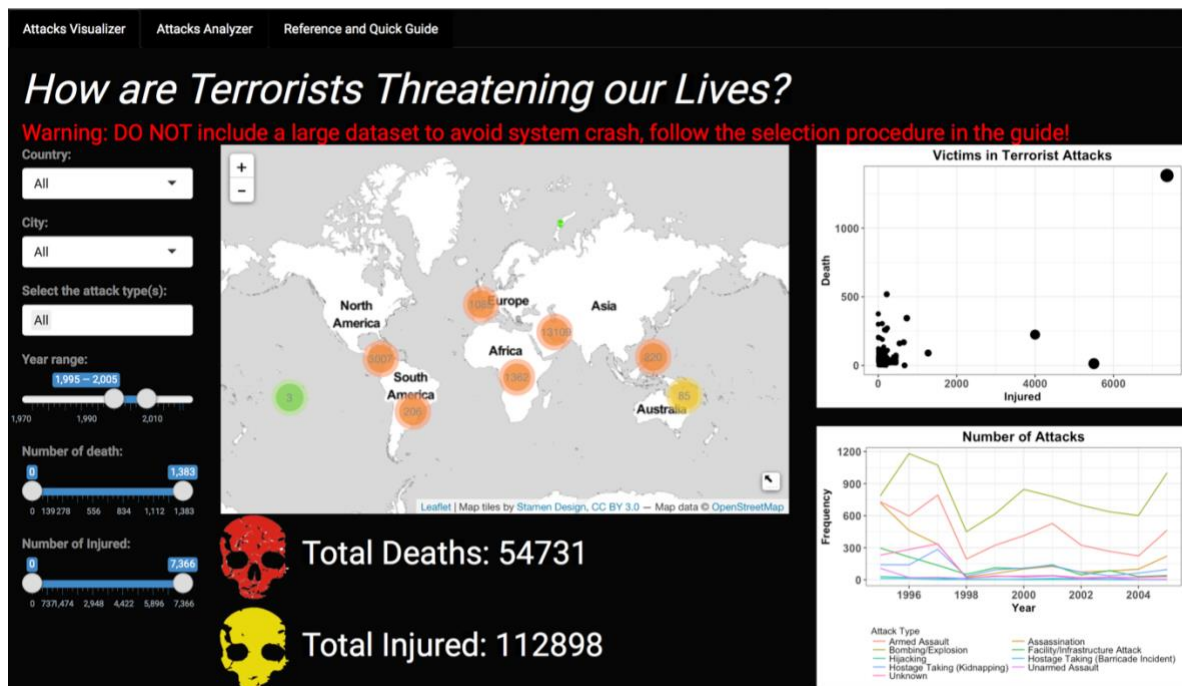
Selection procedure:

1. Select the Country you want to explore (Recommended)

2. Select the City if applicable

3. Select Attack type(s)

4. Adjust the time period (within 10 years is recommended), death and injured based on your preference.

After selection, the number about the victims at the bottom will be updated. The attacks on the map will be shown. Users can zoom into the locations they are more interested in by double clicking the map. Also, when they click a skeleton, pop up will show high-level information about the attack.

Figure 11 Attacks Visualizer Interface   *Figure 12 Attacks Visualizer, Pop up*

The scatter shows the number of death and injured, whereas the line graph shows the number attacks (based on attack types). The size of the scatter points is proportional to the sum of kill and wounded. If a user wants to study an extreme event in the plot, **clicking the data point** will print the details.

Figure 13 Attacks Visualizer, Print Summary

4.2   Attacks Analyzer

User can switch to Attacks Analyzer by clicking the tab bar on the top, and the default interface prints two word clouds. Users can select the motive based on the country, attack type, and time criteria. For example, if a government wants to know the local threatening organizations, they can select their home country.

Figure 14 Attacks Analyzer Default Interface

14

Lastly, users can further enter any key word (can be based on observation from the word clouds). For example, if "war" is observed, user can further find out the motive words and terrorist groups that are related to "war".

Figure 15 Attacks Analyzer, key word
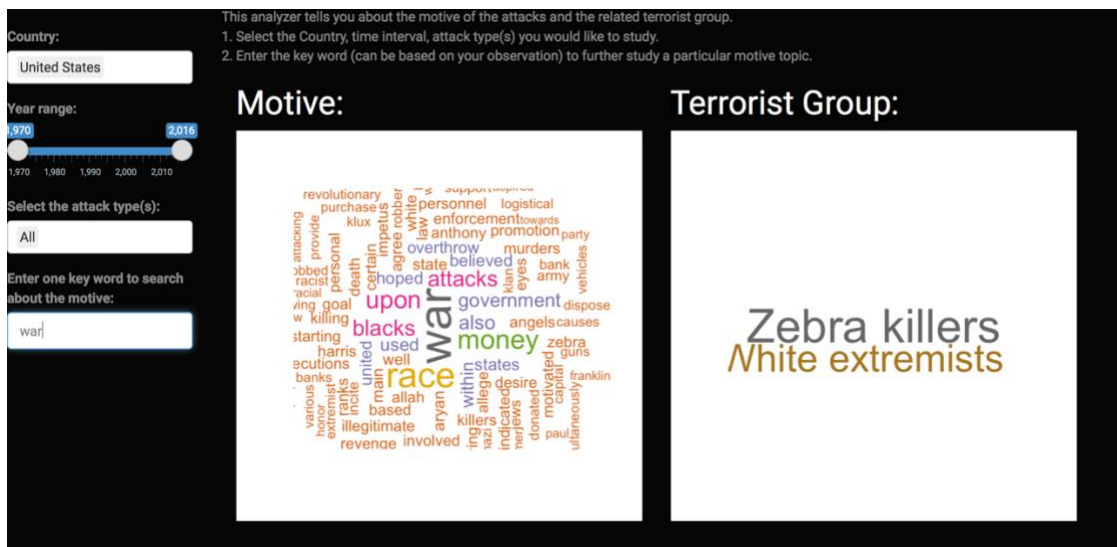
5. Summary & Discussion

Two visualizations are built in this project. The Attacks Visualizer is mainly used to study the events geographically and show how those events took away people's lives. It can also provide the detail about deadly events. This visualization delivers mainstream information that general public would like to know. On the other side, the Attacks Analyzer includes information that general public may not be interested in. It is intended for people who want to study the motives or the terrorist groups about the attacks.

The Attacks Visualizer runs relatively slow due to the large data size, and the leaflet map crashes when users try to include too much data. It is challenging in the design of restrictions on user input widgets. On one side the information and flexibility provided to user have to be maximized, meanwhile the size of datasets has to be controlled to avoid system crash. The crash problem might be possibly solved by using more power packages that accept larger limits, or by adopting a better design that could restrict data size without reducing flexibility significantly. The process speed might also be improved by a better algorithm in accessing data.

Another problem is the coding part of the "renderUI", of which the input choice depends on another input, as a result filters cannot apply directly in the reactive data. Being unfamiliar to shiny, parts of the code are repeated in every renderUI function to solve the problem. This reduces the readability of the codes. The structure of the files can also be improved, instead of putting all in one. It is believed to have a method that can store variables globally so that city names and country names can still be accessible by the ui, so as to improve the overall file structure.

6. References

1. The National Consortium for the Study of Terrorism and Responses to Terrorism, *Global Terrorism Database Code book: Inclusion Criteria and Variables*, Retrieved from http://www.start.umd.edu/gtd/downloads/Codebook.pdf