## **Declaration on Plagiarism**

|  |  |
| --- | --- |
| **Name:** | Aditya Gupta |
| **Student Number:** | 19210195 |
| **Programme:** | Master’s in Computing (Cloud) |
| **Module Code:** | CA682 |
| **Assignment Title:** | Data Visualisation |
| **Submission Date:** | 16 Dec 2019 |
| **Module Coordinator:** | Dr Suzanne Little |

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. I have read and understood the Assignment Regulations. I have identified and included the source of all facts, ideas, opinions, and viewpoints of others in the assignment references. Direct quotations from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged, and the source cited are identified in the assignment references. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

I have read and understood the referencing guidelines found at <http://www.dcu.ie/info/regulations/plagiarism.shtml>, <https://www4.dcu.ie/students/az/plagiarism> and/or recommended in the assignment guidelines.

Name: Aditya Gupta Date: 16/12/2019

**Global Terrorism Trends**

**Abstract**

**What is the question or story you are trying to tell?**

Terrorism is an act, which aims to create fear among ordinary people by illegal means. It is a threat to humanity. Yet compared to most types of crime, terrorism poses unique data collection challenges. As a result, even basic descriptive questions about terrorism have been difficult or impossible to answer like Terror attacks in each country, relationship between number of wounds and number of victims killed and the most commonly weapon type used. So, I am aiming to visualise these results so that it allows people to track and analyse the terrorism trends across the world.

**What is the conclusion that you reached?**

So, after exploring the dataset in an unstructured way to uncover initial patterns and characteristics I am aiming to visualise the terrorism trends across the world with the help of line chart and map in which I will be going to show the relationship between number of people killed and wounded that had occurred due to terror attack and also I am going to visualise the terror attack in each country from 1970 to 2018.

**1. Dataset**

**Where/how did you retrieve it or them?**

The data was taken from National Consortium for the Study of Terrorism and Responses to Terrorism website <https://www.start.umd.edu/research-projects/global-terrorism-database-gtd>. I just signed up on this webpage and downloaded the dataset or it can be downloaded from Kaggle as well.

**Describe the data - size (GB or attributes), number of rows, attributes, data types present? What aspects (if any) of big data (volume, variety, velocity) are present in your data.**

This Global Terrorism Dataset (GTD) include more than 190,000 international and domestic terrorist attacks that occurred worldwide from 1970 to 2018. The GTD defines terrorist attacks as religious, social goal through fear, coercion, intimidation or acts by non-state actors involving the threatened or actual use of illegal force or violence to attain a political. The size of dataset is 90Mb and the attributes such as weapon type, attack type, country, year, fatalities, number of persons wounded etc.

**2. Data Exploration, Processing, Cleaning and/or Integration [½ page]**

**What did you need to do to prepare the dataset(s) to create your graph/chart?**

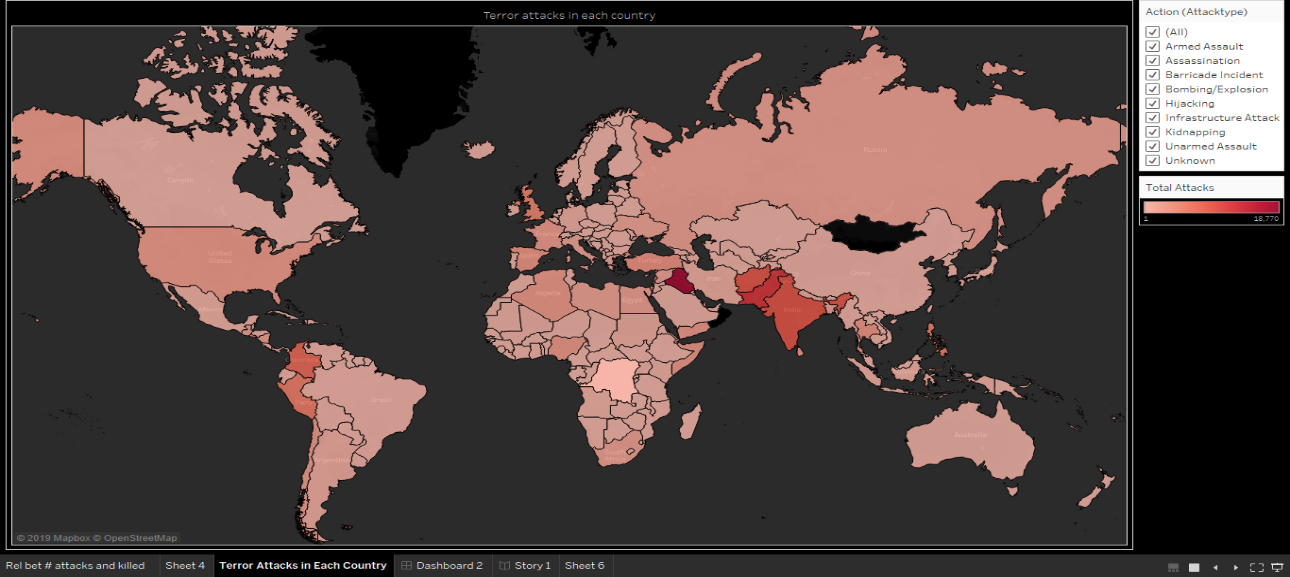
After downloading the dataset, I just explored the large data in an unstructured way to uncover initial patterns, characteristics and point of interest. This helped me to create broad picture of important trends and major points to study and visualise it. I then further used OpenRefine (formerly called as Google Refine) tool to clean the messy dataset and get rid of unwanted columns. During cleaning of the dataset, I reduced the number of columns to one that are needed for visualising the important trends and trimmed the values to remove additional spaces. Furthermore, I removed the data which is erroneous and replaced NA values with 0 in numerical columns. Then finally using Tableau, I just uploaded the dataset and then further visualised it.

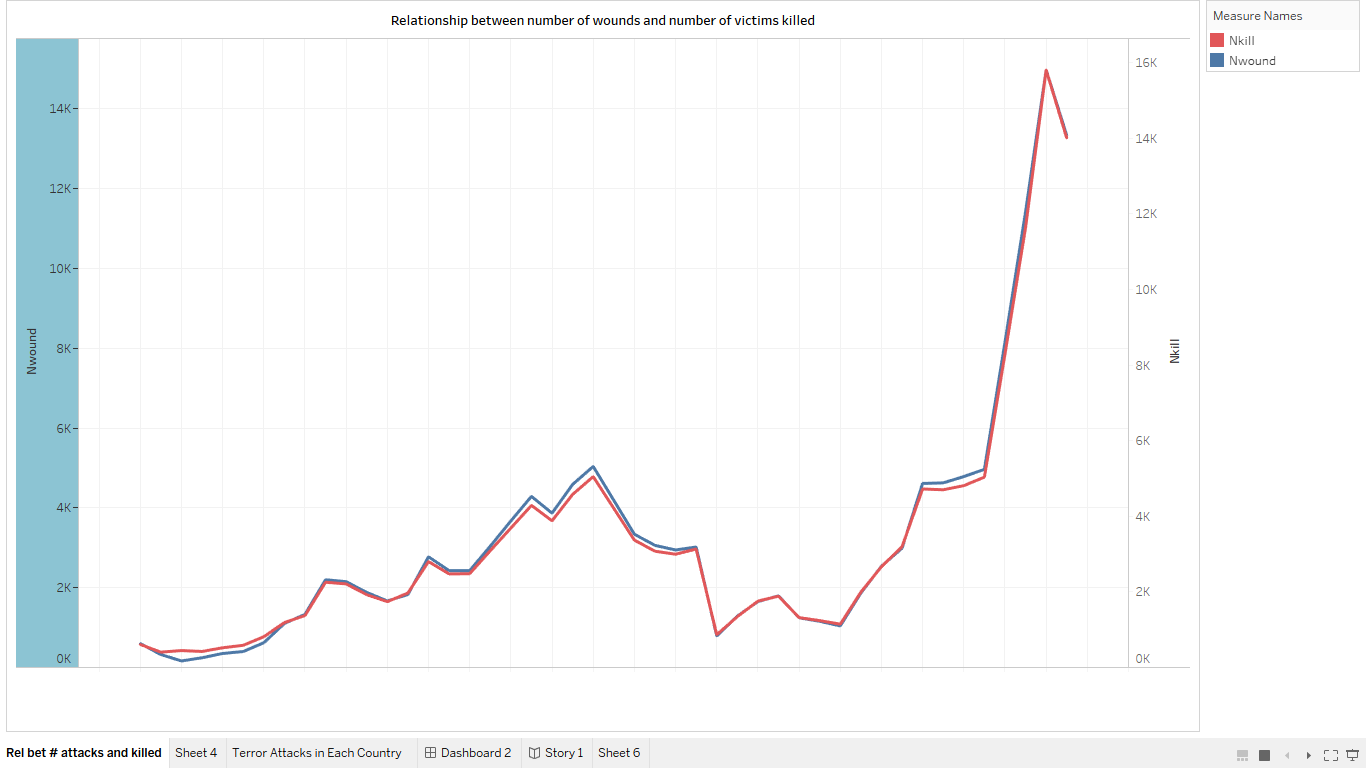
**How did you choose the attributes to visualise?**

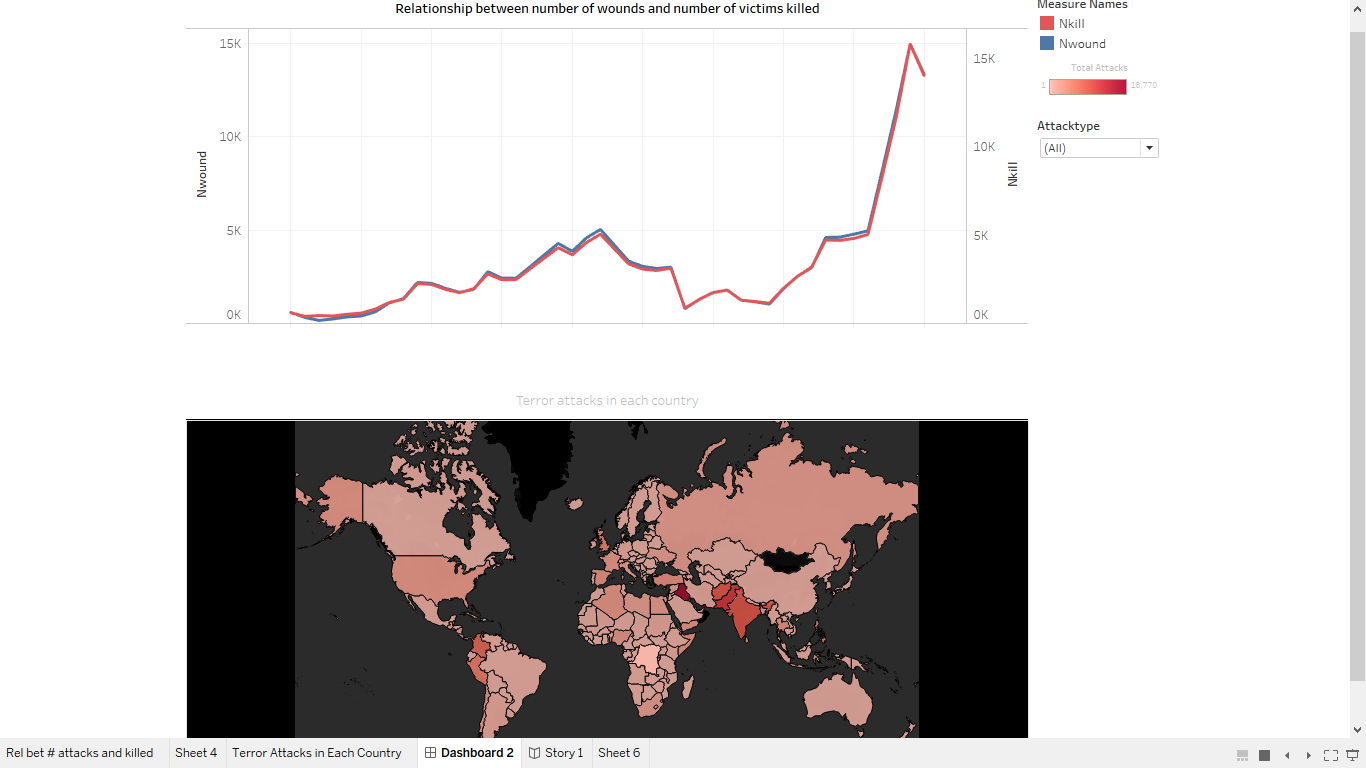
I choose the attributes based on attack types and number of kills and wounded in every country and number of terror attacks that took place at different places in the world.

**3. Visualisation**

**Screenshot or image of visualisation**







**Explain your choice of chart or graph type - what relationship or data type are you showing?**

Line charts shows a better way of analysing a visualisation and can be very attractive to users, also line charts made easy for audience of any kind to quickly understand the graph and its attributes. It can also expose average trends, to assist the reader make predictions or projections for future outcomes. I used map as a chart because it looks way easier to identify the countries and world map is something that can reach out to any kind of audience irrespective of their age or education and map provides an interactive aspect to the visualisation.

**Design choices - justify your use of colour, shapes, marks, layout, structure, font, labels?**

Colours plays and important role in visualisation not only it distinguishes strips but also make the data visualisation appealing and more enjoyable. I used colours like red and blue in my line chart and shades of red in map to represent kills and wounds and the country which is most prone to terrorist attacks and using different colours the graph is easier to understand and analyse. When the user hovers the mouse pointer over a stripe the targeted stripe’s border is coloured white and its width is thickened. Within few seconds looking at the graph anyone can analyse the graph.

**Any interactivity or animation and how it helps answer your question?**

As you can see on the right side of visualisation is the drop-down menu which as default selected as “all”. Whenever you select any country from here, the main visualisation filters out the stripes to only show those that have a prefix in them that matches the content in drop down menu. Moreover, when a user’s mouse hovers over a stripe, the name of country along with number of person’s killed or wounded will be shown and if that stripe is clicked on the name of the selected stripe is shown on drop-down menu and the visualisation shows only the selected stripe in map.

**List of tools or libraries used**

Tools Used: -

1. OpenRefine
2. Tableau

**4. Conclusion**

**Critically analyse the outcome of your visualisation**.

The visualisation is based on the terrorist attacks in every country based on different attributes like attack types, number of victims wounded or killed, etc. The outcome of the visualisation was to visualise the terrorist attacks from 1970 to 2018 and to analyse the relationship between number of kills and number of wounds in that time period. It supports mouse interaction to explore the data and is easily extendable to support different types of data. It has a drop-down filter through which one can easily filter out the results of that country and can explore more about number of victims killed or wounded.

**Were there aspects that you think could be improved upon?**

I also wanted to add the most active terrorism groups and their presence and activity in every city and wanted to show that on the world map also I would like to predict which of the cities will be the next target using my visualisation so as to make it more helpful and interesting for someone who wanted to study about Global Terrorism. And several improvements that may be made in the visualization, stripes that have wide region may be labelled. Comparing several attributes are not present and the solution for this is to have checkboxes through which user can select or deselect items to make stripes visible or not.

**Were there effects or functionality that you were technically unable to achieve?**

I tried the visualisation to look more user interactive, so I tried to add animations in the visualisations for example I wanted to show the top 15 most attack prone country with respect to kills and wounds that changes every year but was unable to achieve.

**References**

<https://www.start.umd.edu/research-projects/global-terrorism-database-gtd>

<http://www.cs.umd.edu/hcil/gtd/gtd/gtd.html>

<https://chartio.com/learn/charts/essential-chart-types-for-data-visualization/>