Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email, and Contribution:

Contributor Roles :

1. Jayanth V (Email – arambamvj@gmail.com) :

1) Data cleaning:
   a) Dropping columns with max null values.
   b) Dropping duplicate rows in the Data Frame.
   c) Replacing NAN values with 0 for the essential columns.

2) Correlation graph between columns.

3) Maximum attacks (year-wise).

4) Statistical graph for max attacks in consecutive years.

5) Region-wise attacks.

6) Terrorist Groups involved.

7) Countries with the most attacks.

8) Most attacked city.

9) Weapons used to cause maximum casualties.

2. Gowthaam Kumarasamy (Email - gowthaam02@gmail.com) :

1) Casualties in those attacks.

2) Number of killings in the region and the country.

3) Number of Wounded individuals in the region and the country.

4) Top 10 years with most casualties.

5) Most used Weapons.

6) Frequent target types.

7) Common methods of attacking.

8) Number of suicides during those attacks.

9) Success Rate of attacks.

10) Entire analysis of the year 2014.

Please paste the GitHub Repo link and Google drive link.

GitHub Link:

https://github.com/jayV1999/Global-Terrorism-Analysis.git

Google Drive Link:

https://drive.google.com/drive/folders/1TV1IVolcYOrUcNsOua_JoyH09izTb4OH?usp=sharing

Please write a summary of your Capstone project and its components. Describe the problem statement, your approaches, and your conclusions. (200-400 words)

The Global Terrorism Data (GTD) is a collection of data containing specific information on the terrorist attacks that took place between 1970 and 2017. This data was put out by the organization called START (Study of Terrorism and Responses to Terrorism). We have used this dataset for performing an EDA.

At first, we performed Data cleaning by creating a data frame and removing unnecessary columns with null values, dropping the duplicate rows in this dataset, and started getting basic insight from the cleaned dataset.

In this cleaned dataset we have focused only on the quantitative analysis such as regions, countries, and cities with the maximum number of attacks and visually represented them by plotting these data in a graph and also found out the number of casualties in these attacks.

We found the weapons used for these attacks and also the number of casualties caused by these weapons. This overall general analysis helped us identify the killing rate and wounded rate in a particular region, the weapons used, common attack types, their most common targets, and the success rate of these attacks. we have used that to plot the bar graph and pie chart to give much insight into this visually.

And finally, we found out that 2014 was the year in which most attacks took place and we made a detailed analysis considering this year alone.