

# EGIM08 - Coursework 1

Prajwal Bharadwaj - 2337862

## 1 Goal

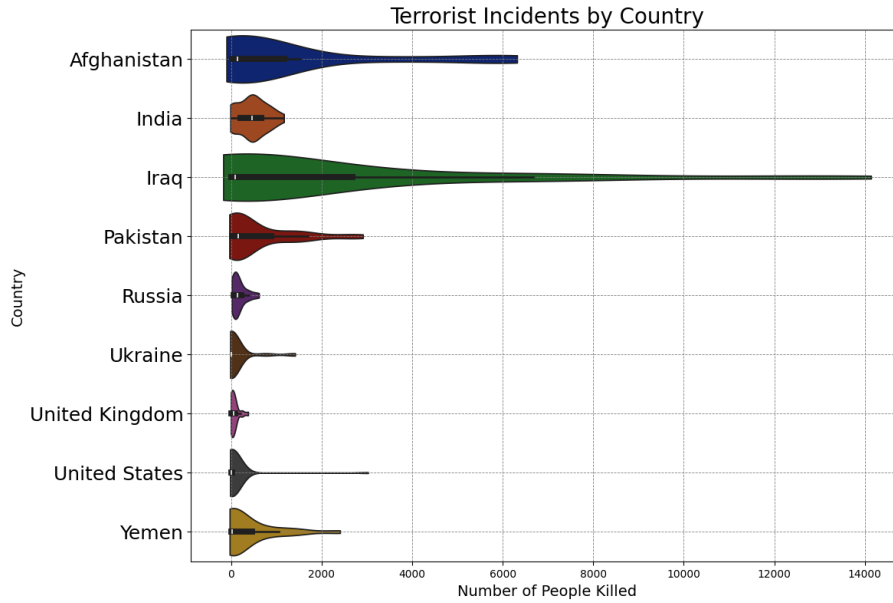


Figure 1: Illustration of Highest death toll faced by Middle Eastern countries

This visualization aims to quantify the distribution of Total Death toll in different regions and countries. Trends are now plotted using Violin plots to understand more insights of Terrorist incidents over the Top 9 active countries in terrorism.

## 2 Insights

The violin plot in the Figure 1 describes the number of people killed in each of those mentioned countries and these countries are indeed selected from the Regions of Middle East and North Africa, Western Europe, Eastern Europe, South Asia, where selected countries from these regions had significant Death toll numbers. Violin plots is a combination of KDE distribution and Box plot indicating the Inner-quartile ranges and density distributions giving a deeper insight on the statistical parameters of Death toll faced by each of these countries. From the plot, Iraq has significantly been affected with the death toll peaking at nearly 14000 and its inner

quartile ranges are nearly 3000 deaths and USA and UK has significantly lower death tolls as compared to the Middle East and North Africa regions.

### 3 Data Abstraction

- **Dataset Type :** Tabular Data - Comma separated Value File
- **Items:** Each Item is a terror event consisting of data related to information of Terrorism event
- **Attributes:** For the below visualization of Terrorism death toll for each of the mentioned countries, a data-frame consisting of **Countries** which is a Categorical attribute type and **Total Killed** which is a quantitative attribute are computed and plotted as a Violin Plot by utilizing the seaborn library from Python.  
This information is stored as Attributes for each Item.

### 4 Task abstraction

- **Marks:** KDE Plot  
In a violin plot, the main marks are the kernel density estimation (KDE) curves, which depict the distribution of data.
- **Channels:** Colors  
Different colors to categorize each countries Death Toll
- **Users:** General public, researcher, journalist, and policy maker.
- **Actions :**  
High level → Present the Data  
Mid Level → Browse the Data  
Low Level → Compare the Data
- The analysis of total death toll for each country is represented by the mean death toll displayed within the violin plot as a box plot.
- Each violin plot includes a box plot with its own inner quartile ranges:
- The first quartile represents the 25th percentile of the death toll.
- The upper quartile indicates the 75th percentile of the death toll.
- The extreme points represent the whiskers, indicating extreme values of death toll for each country. Additionally, the violin plot provides a density estimate of the death toll distribution, indicating the frequency of death toll.

- **Target:** Comparison: → Comparison: Violin Plots enable effective comparison of death tolls for each country using: Inner quartile ranges of the box plot Density distribution indicating frequency of death toll occurrences. For Iraq and Pakistan: Whisker point for Iraq is around 6500, while for Afghanistan, it is approximately 1800. Data points beyond the whisker points may be considered outliers, indicating extreme values. Analysis of quartile ranges: USA, UK, Russia, and Ukraine exhibit significantly lower involvement in terrorism incidents, reflected by reduced death toll. Trend of increasing death tolls: Suggests these countries may harbor terrorists or have citizens actively participating in terrorism.

## 5 Additional Data source

No additional data source has been utilized except the Input Data file from the Kaggle Website

Dataset : Global Terrorism Dataset

Link : <https://www.kaggle.com/datasets/START-UMD/gtd>

## PEER FEEDBACK

**Goals and insights:** 60% - *Unclear description of the goals or insights. Lack of correspondence between both.*

- The goal could be confusing to the users as it lacks some specificity regarding the time frame of the data. You could rephrase the goal to be more focused.

**Data abstraction:** 100% - *The description completely corresponds to the data and the vis. Description of dataset and data types included and clearly explained.*

- The information provided is clear and accurate.

**Task abstraction:** 60% - *Task abstractions are described in detail with some flaws or misunderstandings of the task abstractions. Description of marks and channels.*

- A description is accurate. However, you could have made the 'Actions' section even stronger by adding additional relevant actions like 'analyse' in high-level actions, 'Search' in mid-level actions, 'identify', and 'compare' in low-level actions.

**Image of the vis:** 60% - *The image is of appropriate quality, but it is unclear how the stated insights could be drawn from the vis.*

- The image is of appropriate quality. While the insights section mentions the features of the violin plot, it does not fully explain how all these elements contribute to the specific comparisons made. Also, the inner quartile ranges shown in the image do not provide us with accurate information.