Contents

[Objective: 1](#_Toc82282605)

[About the Data: 1](#_Toc82282606)

[Source: 1](#_Toc82282607)

[Parameters: 1](#_Toc82282608)

[Importance of this data: 2](#_Toc82282609)

[Strengths & Weaknesses of this data: 2](#_Toc82282610)

[Technology: 3](#_Toc82282611)

# Objective:

Create a public visualization using the [Aid Worker Security Database (AWSD](https://aidworkersecurity.org/about)). The primary audience for this visualization is…

*Target Audience:*

a candidate or existing aid worker researching what countries they’d like to serve in

*Deliverables:*

*to help the target audience make their decision this project will deliver…*

1. see a visualization of the types of violence that have occurred in that country

2. a comparison of a country to other countries they indicated they might like to serve in

3. a visualization of the top 10 highest violence countries and type of violence

4. a visualization of the top 10 lowest violence countries and type of violence

# About the Data:

Source:The [AWSD](https://aidworkersecurity.org/about) is a global compilation of reports on major security incidents involving deliberate acts of violence affecting aid workers.

Parameters:Below is a summary based on the “About” page at the above link.

*"Major incidents"* are defined as killings, kidnappings, and attacks that result in serious injury.

*"Aid workers"* are defined as the employees and associated personnel of not-for-profit aid agencies that provide material and technical assistance in humanitarian relief contexts.

*For each incident, the information recorded is:*

* Date
* Country and specific location, including geocodes
* Number of aid workers affected (victims)
* Sex of victims
* Institutional affiliation of victims (UN/Red Cross/NGO/other)
* Type of staff (national or international)\*
* Outcome of the incident (victims killed/wounded/kidnapped)
* Means of violence (e.g. shooting, IED, aerial bombardment)
* Context of attack (ambush, armed incursion, etc.)
* Summary of incident (public details)

## Importance of this data:

This data can be used for a number of purposes, and in fact has been, as demonstrated by the variety of reports ([link to reports](https://aidworkersecurity.org/reports)) and [graphics](https://aidworkersecurity.org/incidents/report/location) created. Note that though there are a variety of graphics, they appear to be one graph per page, versus the approach this project will take to create a visualization with multiple charts to answer the primary audience’s questions.

## Strengths & Weaknesses of this data:

*Strengths:* This data is available as a csv file spanning the years 1997 through 2021. There are over 3,000 incidents and though there is missing data it appears to be minimal as shown in the table below…

There are 3,247 incidents (rows) in the csv file…

| *feature* | *values\_count* | *missing\_values\_count* |
| --- | --- | --- |
| City | 2385 | 862 |
| District | 2574 | 673 |
| Region | 2892 | 355 |
| Day | 2933 | 314 |
| Month | 3201 | 46 |
| Country Code | 3213 | 34 |
| NRCS and IFRC | 3237 | 10 |
| ICRC | 3237 | 10 |
| NNGO | 3237 | 10 |
| Actor type | 3240 | 7 |
| Country | 3241 | 6 |
| Latitude | 3241 | 6 |
| Longitude | 3241 | 6 |
| Actor name | 3246 | 1 |
| Source | 3246 | 1 |
| Verified | 3246 | 1 |

These remaining columns have no missing values:

| *feature* | *feature* |
| --- | --- |
| Total kidnapped | Internationals Kidnapped |
| Details | Year |
| Location | Internationals Killed |
| Attack context | Total nationals |
| Means of attack | Nationals Kidnapped |
| Gender Unknown | Nationals Wounded |
| Gender Female | Nationals Killed |
| Gender Male | Other |
| Total affected | INGO |
| Incident ID | UN |
| Total wounded | Internationals Wounded |
| Total killed |  |

*Weaknesses:*

The only noted weakness so far is that a number of the features appear redundant. This will be addressed via feature reduction and including only features needed for the stated objectives.

*Additional exploration and cleaning of data will occur if this proposal is approved to move forward with.*

# Technology:

These are the technologies under consideration to build this interactive visualization…

|  |  |  |
| --- | --- | --- |
| *Preferred Order* | *Technology* | *Comments* |
| 1 | Plotly in Python with Dash | Given the challenges I’m encountering with DATA608 assignment #4, I’m not sure if I can identify how to apply this approach to meet this project’s scope. |
| 2 | Data Studio | My employer is promoting the use of Google’s Data Studio. Using this technology could help me be more productive at my work. |
| 3 | Tableau | Given past experience with this technology, I’m confident Tableau has the capabilities and I have the skills to build this visualization |