Aid Worker Attacks Report

Eric Liu

1. INTRODUCTION

The prevalence of danger to aid workers around the world is something that is often overlooked when discussing foreign attacks. In recent years, these people who provide assistance in humanitarian relief are often subject to acts of violence during their service. The Aid Worker Security Database (AWSD) allows users to take a detailed analysis of the many attacks that have affected workers since 1997. With over 3200 recorded incidents and a multitude of variables, the database could be an important piece in providing safety to aid workers in the future.

The goal of the following report is to provide some semblance of a pattern of these incidents where we're able to locate common areas or time periods of attacks. Extracting this type of valuable information from this database could potentially prevent further danger for these workers.

2. DATA

The dataset we are using is from the Aid Worker Security Database and is collected from various public sources which are then compiled into a complete spreadsheet. The information is then verified and then crosschecked across humanitarian agencies to ensure valid data. In total, there are 3511 incidents that occurred between 1997 and 2022 where an aid worker has been attacked in some form. Additionally, there are 41 columns that describe the incident including gender, location, and year. These columns contain both numerical and categorical values so data preprocessing is needed before an analysis can be done on the incidents.

3. DATA CLEANUP

Removing unnecessary columns and values is the first step in our data analysis process. To do

this, we should first run an info function to determine which variables are not needed.

```
$ Year                      : chr  "#date+year" "1997" "1997" "1997" ...
$ Month                     : chr  "#date+month" "" "1" "1" ...
$ Day                       : chr  "#date+day" "" "" "" ...
$ Country.Code              : chr  "#country+code" "" "KH" "RW" ...
$ Country                   : chr  "#country+name+i_en" "" "Cambodia" "Rwanda" ...
$ Region                    : chr  "#adm1+name+i_en" "" "Banteay Meanchey" "Northern" ...
$ District                  : chr  "#adm2+name+i_en" "" "" "Musanze" ...
$ City                      : chr  "#adm3+name+i_en" "" "" "Ruhengeri" ...
$ UN                        : chr  "#affected+UN" "0" "0" "0" ...
$ INGO                      : chr  "#affected+INGO" "0" "0" "4" ...
$ ICRC                      : chr  "#affected+ICRC" "2" "1" "0" ...
$ NRCS.and.IFRC             : chr  "#affected+IFRC" "0" "0" "0" ...
$ NNGO                      : chr  "#affected+NGO" "0" "0" "0" ...
$ Other                     : chr  "#affected+other" "0" "0" "0" ...
$ Nationals.killed          : chr  "#affected+killed+national" "1" "1" "0" ...
$ Nationals.wounded         : chr  "#affected+injured+national" "0" "0" "0" ...
$ Nationals.kidnapped       : chr  "#affected+kidnapped+national" "0" "0" "0" ...
$ Total.nationals           : chr  "#affected+total+national" "1" "1" "0" ...
$ Internationals.killed     : chr  "#affected+killed+international" "0" "0" "3" ...
$ Internationals.wounded    : chr  "#affected+injured+international" "1" "0" "1" ...
$ Internationals.kidnapped: chr  "#affected+kidnapped+international" "0" "0" "0" ...
$ Total.internationals      : chr  "#affected+total+international" "1" "0" "4" ...
$ Total.killed              : chr  "#affected+total+killed" "1" "1" "3" ...
$ Total.wounded             : chr  "#affected+total+injured" "1" "0" "1" ...
$ Total.kidnapped           : chr  "#affected+total+kidnapped" "0" "0" "0" ...
$ Total.affected            : chr  "#affected+total" "2" "1" "4" ...
$ Gender.Male               : chr  "#affected+total+m" "0" "0" "3" ...
$ Gender.Female             : chr  "#affected+total+f" "0" "0" "1" ...
$ Gender.Unknown            : chr  "#affected+total+i" "2" "1" "0" ...
$ Means.of.attack           : chr  "#event+means+i_en" "Unknown" "Unknown" "Shooting" ...
$ Attack.context            : chr  "#event+type+i_en" "Unknown" "Unknown" "Raid" ...
$ Location                  : chr  "#loc+type+i_en" "Unknown" "Unknown" "office/compound" ...
$ Latitude                  : chr  "#geo+lat" "" "14.07092855" "-1.49984" ...
$ Longitude                 : chr  "#geo+lon" "" "103.099916" "29.63497" ...
$ Motive                    : chr  "#x_event_motive" "Unknown" "Unknown" "Unknown" ...
$ Actor.type                : chr  "#group+type+i_en" "Unknown" "Unknown" "Unknown" ...
$ Actor.name                : chr  "#group+name+i_en" "Unknown" "Unknown" "Unknown" ...
$ Details                   : chr  "#description+i_en" "1 ICRC international staff wounded and 1
```

With this information we find that several columns contain large descriptions, unknown values,

or repetitive information. Seeing as this is the case, we can then remove these columns that are

not necessary for our analysis.

Description: df [3,511 x 19]

|	Year <dbl>	UN <dbl>	INGO <dbl>	Other <dbl>	Nationals.killed <dbl>	Nationals.wounded <dbl>	Nationals.kidnapped <dbl>	Total.nationals <dbl>	Internationals.killed <dbl>
2	1997	0	0	0	1	0	0	1	0
3	1997	0	0	0	1	0	0	1	0
4	1997	0	4	0	0	0	0	0	3
5	1997	4	0	0	0	0	4	4	0
6	1997	0	1	0	0	0	0	0	1
7	1997	1	0	0	1	0	0	1	0
8	1997	0	0	0	10	0	0	10	0
9	1997	3	0	0	1	2	0	3	0
10	1997	1	0	0	1	0	0	1	0
11	1997	0	1	0	1	0	0	1	0

1-10 of 3,511 rows | 1-10 of 19 columns   Previous 1 2 3 4 5 6 … 100 Next

Now we have a smaller dataset with 19 columns but is now much easier to do analysis on given

that they are now all numerical values. Visualizations and interpretations can now be created

using this dataset as we move onto the next part of the analysis process.

4. ANALYSIS

With the new dataset that we've finished cleanup on, we can create various visualizations that will help in determining information on where and when attacks happen. First a histogram is made depicting what years had the highest occurrence of attacks on workers.
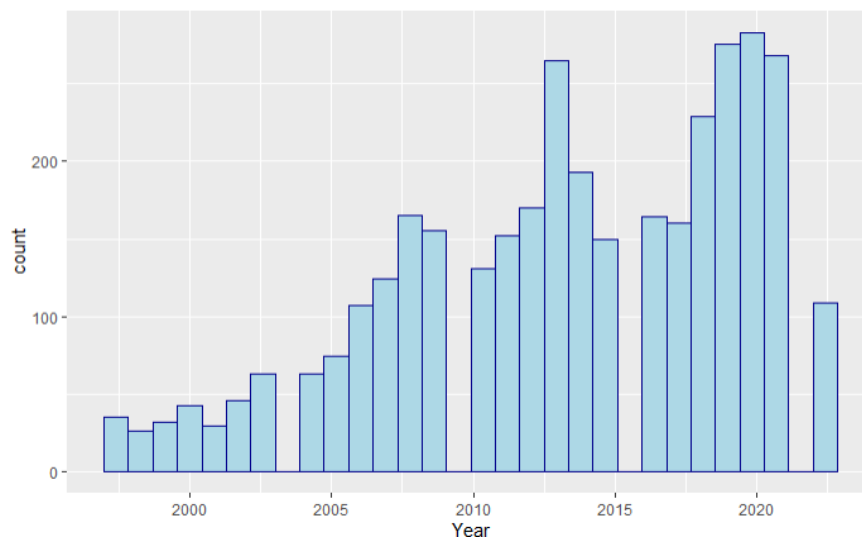

Based on this graph, we can tell that has been a steady increase in the violent attacks on workers since the 1990s with it peaking just a few years ago in 2020. Perhaps as the pandemic entered the world, more humanitarian workers needed to be sent out to help foreign nations which, in turn, led to a higher number of attacks being recorded.

Furthermore, now that we know 2020 was the year where attacks occurred the most, we can use the sort function on the table to determine which location was the most prevalent in the dataset.

```{r}
tail(names(sort(table(df2$Country))), 1)
```

```
[1] "Afghanistan"
```

Seeing as Afghanistan is the country with most attacks on aid workers, a plausible explanation could be due to the many casualties that were seen from the War in Afghanistan which occurred throughout the 21st century. Now that we understand the location and period where many of the attacks occurred, we can then focus on which groups of people were most likely to be targeted.
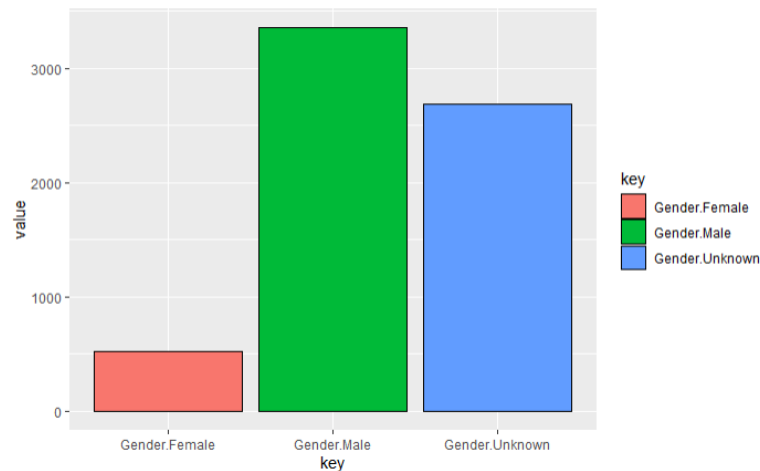
 From this graph, we can see that there is an overwhelmingly larger number of men being attacked in comparison to women. This may be due to several possible reasons, such as an increasing number of men being sent out into foreign nations over the decades or perhaps cultural differences within these countries. Given this distribution of male to female, the unrecorded genders in blue would likely follow the same trend.

  Now that we found the groups of people that are at increased risk of attack, we may be curious as to which factors affect this outcome the most. For this, we can create a correlation heatmap that will show the relevance of different variables.
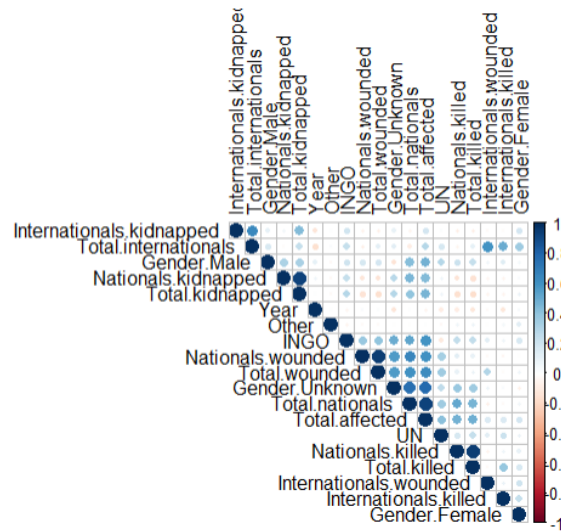
Using this chart, we can see that highly correlated variables have a darker shade of blue. There are several correlations that may be of interest to this analysis. This includes nationals kidnapped with INGO, and UN with nationals wounded. While there may be other higher correlations in the chart, these two seem to make the most sense as to how they are connected. INGO and the UN are both organizations who send out the aid workers to other countries. With the large number of attacks throughout the years, it appears that these two are most clearly linked with kidnappings and bodily harm. This may be due to large amount of people sent by the organizations during wartime and national conflicts.

5. CONCLUSION

Based on the analysis of common factors that surround an attack on aid workers, we have come to several conclusions which may be useful in predicting future incidents. When focusing on the location and time of an attack, we find that the country Afghanistan and the year 2020 had attacks happen most frequently. This could be due to several reasons, such as the war in

Afghanistan where many troops had to be deployed. COVID could have been a big factor in the trend toward attacks in recent years, as many countries needed foreign assistance.

On the other hand, when we look at the groups of people that tended to be involved in the acts, we find that an overwhelming amount had men as the victims. In addition, INGO and the UN appeared to be the two organizations where incidents occurred with most frequency. This may be since they contain a large number of workers which they send during worldwide conflict. The majority of these workers may be male, as what our analysis has suggested.

Now that we have very clear evidence that these certain groups of people are commonly affected by foreign attacks on aid workers, we must work on preventing the trend of attacks in the future. Knowing these factors is the first step in prevention and based on our analysis, males from the UN and INGO may be the most vulnerable.