A picture containing food

Description automatically generated

**ANL501 Data Visualisation and Storytelling**

*End-of-Course Assessment*

*July 2024 Presentation*

**Tutorial Group:** 01

**Lecturer:** Prof Nicholas Sim

**Submission date:** 30th September 2024

**Done by:**

|  |  |
| --- | --- |
| **Name** | **PI Number** |
| **Chua Yee Huan** | **H1970661** |

The Legacy of Civil War: Landmines, ERWs, and Their Effects on Cambodian Communities

Chua Yee Huan

30 September 2024

## 1. Executive Summary

This study looks at how communities all around Cambodia are still affected by landmines and explosive remnants of war (ERWs). Decades after the nation’s civil war ended, landmines and ERWs still endanger public safety, especially in rural areas where agriculture is a primary livelihood. To shed light on the geographical distribution, trends in casualties, and severity of incidents pertaining to these devices, this investigation makes use of data visualization approaches.

#### Key Findings:

* **Impact of Device**: A direct comparison between landmines and ERWs revealed that the latter typically result in a greater number of victims per occurrence. This implies that even with the success of landmine removal operations, ERWs continue to pose as a substantial threat to the Cambodian communities.
* **Casualty Trends**: The data indicates that victims are still at risk, especially in areas like rice fields, foraging zones, and built-up urban areas, despite the fact that the number of casualties dropped gradually between 2005 and 2013. These areas indicates a high number of amputations, fatalities and injuries, which is quite worrying.
* **Land Use and Casualty Type**: The examination of different casualty categories according to landuse indicates that incidents primarily take place in residential and agricultural regions, putting civilians, particularly farmers and rural residents at greater danger.
* **Geographical Distribution**: Provinces like Battambang, Banteay Meanchey and Pailin had the largest number of landmine and ERW incidences, indicating local hotspots that need targeted demining efforts.

#### Implications:

In order to reduce the risk posed by these war remnants, this study emphasizes the need for ongoing funding for demining operations and community education initiatives. Policymakers, international organizations and non-governmental organizations (NGOs) engaged in landmine clearance and restoration initiatives might utilise the findings as practical insights.

#### Recommendations:

* **Targeted Demining**: To lower the number of incidents caused by mine and ERWs, Cambodia should concentrate demining resources in high-risk provinces which are more vulnerable, such as Battambang, Banteay Meanchey and Pailin.
* **Provide Support for Victims**: Expand access to medical treatment and rehabilitation programs for survivors, with a focus on amputees’ needs.
* **Technological Support**: To expedite the pressing issue to clear impacted regions and bring down casualties number, Cambodia should invest in cutting-edge demining technology and inventions. This is made possible due to the rapid advancement of technology in this day and age.

These proposals seek to resolve the legacy of landmines and ERWs in Cambodia and guarantee safer settings for future generations by concentrating on areas and communities most impacted.

## 2. Introduction

The Cambodian civil war, which spanned from 1960s to 1970s, had a significant and long-lasting effect on the nation. The millions of landmines and explosive remnants of war (ERWs) that are scattered across Cambodia are among the most enduring consequences. Long after the fighting has ended, these concealed threats still significantly jeopardize the security, lifelihood, and progress of Cambodian communities.

A lingering legacy of battle, landmines and ERWs inflict harm, fatalities and psychological anguish on its people. These explosives impede access to land, infrastructure, basic social services, irrigation, and safe drinking water for consumption. Victims are often left crippled, undergo amputation and receive minimal assistance in obtaining job, training, education, and mental and physical health care while they are incapacitated. Victims are typically women, landless, and illiterate. Even while demining efforts have advanced significantly, there are still many obstacles to overcome. Till date, there are still a large number of explosives that remains uncleared.

This report highlights the findings pertaining to landmines and ERWs in Cambodia, as well as the psychological repercussions of these explosives on Cambodian communities.

The 2 datasets that are primarily used to work on this report are:

* **‘mine.shp’**: Contains information on landmine incidents (victim, injure, amputation, killed) that occurred between 2005 and 2013. It also includes information on the type of device that caused the explosion, the incident’s coordinates, date of incident, and other location-related details such as the names of the communes, districts and provinces .
* **‘fear.shp’**: Contains surveys on people’s fears about explosives from 2009 through 2014. The dataset includes locational attributes like the names of the provinces, districts, communes, as well as information about the date of the survey was conducted, the locations of contamination sites, the respondents’ fear levels, their proximity to the contamination sites, and landmine’s class types in which the contamination sites are located.

The source of the respective datasets can be found in [here](https://data.opendevelopmentcambodia.net/dataset/mineerw-casualties-2005--2013) and [here](https://data.opendevelopmentcambodia.net/dataset/erw).

Our analysis reveals that communities located close to minefields or historical battle zones typically have greater levels of fear. Additionally, there is a strong positive correlation between the number of injuries and fatalities in explosive incidents. Areas with more injuries tend to have higher fatalities count, and the degree of amputation adds even another stress aspect. It implies that explosives inflict long-lasting physical and psychological harm on impacted communities in addition to posing an immediate risk.

## 3. Data

The datasets used for this research was collected from Open Development Cambodia (ODC). Mine dataset spans the years 2005 to 2013, while fear dataset spans the years 2009 to 2014. It includes information from across multiple provinces across Cambodia, providing insights into the casualties count, the type of devices which caused the explosions, incident dates, the psychological impact on Cambodian communities, landmine’s class types and their proximity to the contamination sites. The primary goal of the data is to examine the situation regarding landmines and ERWs in Cambodia.

#### Key Variables:

* **Landuse**: A categorical variable representing the type of land usage.
* **ERW\_TYPE**: A categorical variable representing the type of ERW which caused the explosion.
* **MINE\_TYPE**: A categorical variable representing the type of MINE which caused the explosion.
* **Fear\_Level**: A categorical variable representing the respondent’s fear level on these explosion.
* **Proximity**: A categorical variable representing the respondent’s proximity to the contamination sites.

#### General Issues

During data pre-processing stage, it was observed that there were a total of 61 empty values under the column, ERW\_TYPE, which we will replace these with “Unknown”. In addition, province names of Cambodia were either spelled incorrectly or requires fixing of spaces, so we standardized these names across both datasets. Lastly, it was noted that there was quite a large number of respondents who had left their fear level as blank in their survey responses, which we will replace these blank fields with “Unknown”.

#### General Observations

* Preliminary analysis showed that ERWs cause higher number of victims as compared to mines. Based on the trendlines plotted from 2005 to 2013, the number of victims who suffered from ERW-related casualties or mine-related casualties indicate a steady decline in incidents over the years, probably due to demining activities and increased public awareness. Nonetheless, the fact that victims continue to appear annually, particularly between 2011 and 2013, indicates that mines and ERWs continue to pose as a dangerous threat, particularly in regions with challenging terrain or little access to demining resources.

## 4. Main Discussions

### 4.1 Exploratory Data Analysis

The two datasets are both in shape files, so in order to import and read both files, I will use the sf library, i.e. install.packages(“sf”) and use read\_sf to parse the shape files into a sf data frame. I will store my datasets as mine.df and fear.df.

# Importing raw data files  
  
mine.df <- sf::read\_sf(dsn = "./fear/mine.shp")  
  
fear.df <- sf::read\_sf(dsn = "./mine/khm\_blscontaminationp\_gov\_cmaa.shp")

After my datasets have been imported successfully, I will move on to perform data analysis, data management, check for any missing values, and consider how I want to handle these missing values separately for both of my datasets - mine and fear.

It is noted that in the last column of the dataset, the geometry variable given to us is not in the latitude-longitude coordinate system, so we will have to convert it using st\_transform before we are able to read the sf data frame in geom\_sf.

Likewise, I will repeat similar steps for my fear dataset. Noticed that for village code in the fear dataset, it is originally “chr”, so I will convert it into “num” instead, using as.numeric().

Next, I will do some data cleaning to both my mine and fear datasets, before I can start to conduct my data visualization to draw insights for my research. Noticed that from the original datasets given, province names slightly differ from across both datasets. One of the dataset has all the province names capitalized while the other is not. In addition, some contains spelling errors or have to fix spacing issues. I will rectify these issues in my data cleaning step. I will convert all the province names in both datasets into lowercase characters and remove any leading/trailing spaces between characters before correcting the names to standardize across both datasets.

## [1] "Mondol Kiri" "Kampong Cham" "Banteay Mean Chey"   
## [4] "Otdar Mean Chey" "Siem Reap" "Preah Vihear"   
## [7] "Pursat" "Battambang" "Krong Pailin"   
## [10] "Kaoh Kong" "Kampot" "Kampong Speu"   
## [13] "Kandal" "Kracheh" "Kampong Thom"   
## [16] "Kampong Chhnang" "Stueng Traeng" "Svay Rieng"   
## [19] "Takaev" "Prey Veaeng" "Phnom Penh"   
## [22] "Rotanak Kiri" "Krong Preah Sihanouk"

## [1] "ODDAR MEANCHEY" "BATTAMBANG" "PREAH VIHEAR" "PAILIN"   
## [5] "PURSAT" "SIEMREAP" "BANTEAY MEANCHEY" "KAMPONG THOM"   
## [9] "KRATIE" "KAMPONG CHHNANG" "PHNOM PENH" "KAMPONG CHAM"   
## [13] "KAMPONG SPEU" "SVAY RIENG" "PREY VENG" "KANDAL"   
## [17] "KAMPOT" "TAKEO" "KEP" "MONDUL KIRI"   
## [21] "STUNG TRENG" "KOH KONG" "SIHANOUKVILLE" "RATANAK KIRI"

## [1] "mondulkiri" "kampong cham" "banteay meanchey" "oddar meanchey"   
## [5] "siem reap" "preah vihear" "pursat" "battambang"   
## [9] "pailin" "koh kong" "kampot" "kampong speu"   
## [13] "kandal" "kratie" "kampong thom" "kampong chhnang"   
## [17] "stung treng" "svay rieng" "takeo" "prey veng"   
## [21] "phnom penh" "ratanakiri" "preah sihanouk"

## [1] "oddar meanchey" "battambang" "preah vihear" "pailin"   
## [5] "pursat" "siem reap" "banteay meanchey" "kampong thom"   
## [9] "kratie" "kampong chhnang" "phnom penh" "kampong cham"   
## [13] "kampong speu" "svay rieng" "prey veng" "kandal"   
## [17] "kampot" "takeo" "kep" "mondulkiri"   
## [21] "stung treng" "koh kong" "preah sihanouk" "ratanakiri"

## [1] "mondulkiri" "kampong cham" "banteay meanchey" "oddar meanchey"   
## [5] "siem reap" "preah vihear" "pursat" "battambang"   
## [9] "pailin" "koh kong" "kampot" "kampong speu"   
## [13] "kandal" "kratie" "kampong thom" "kampong chhnang"   
## [17] "stung treng" "svay rieng" "takeo" "prey veng"   
## [21] "phnom penh" "ratanakiri" "preah sihanouk"

### 4.2 Data Visualization Analysis Insights

This section contains visual insights drawn from both mine and fear datasets. It allows stakeholders to come to a conclusion more intuitively, since visuals help to condense large amount of data into bite-sized information.

#### 4.2.1 Cambodia Map Overview

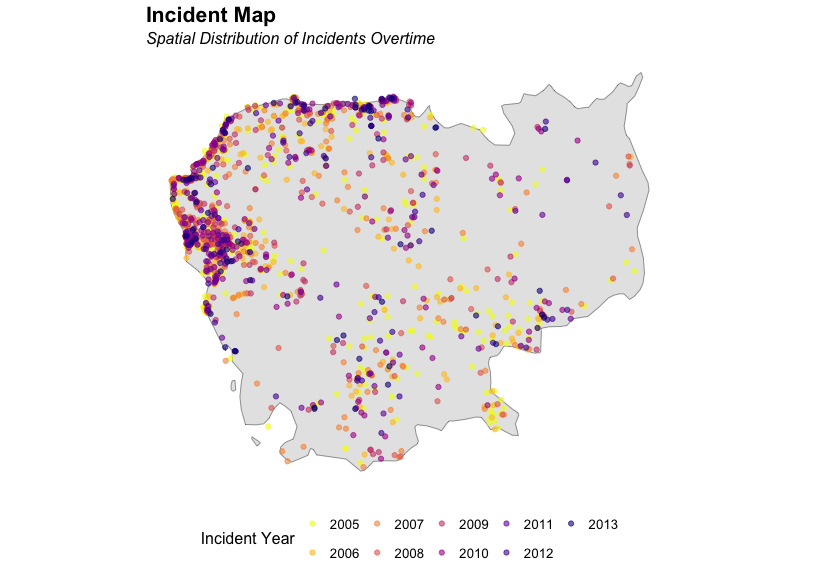
The image below shows an overview of Cambodia map plotted using its geographic coordinates, which is fetched from get\_googlemap. To be able to retrieve this map, I have obtained my Google API key prior to this. This map functions as a broad geographic reference, to tell us which are the neighboring countries of Cambodia, based on its longitude and latitude.

## # A tibble: 1 × 2  
## lon lat  
## <dbl> <dbl>  
## 1 105. 12.6



#### 4.2.2 Cambodia Incident Map Overtime

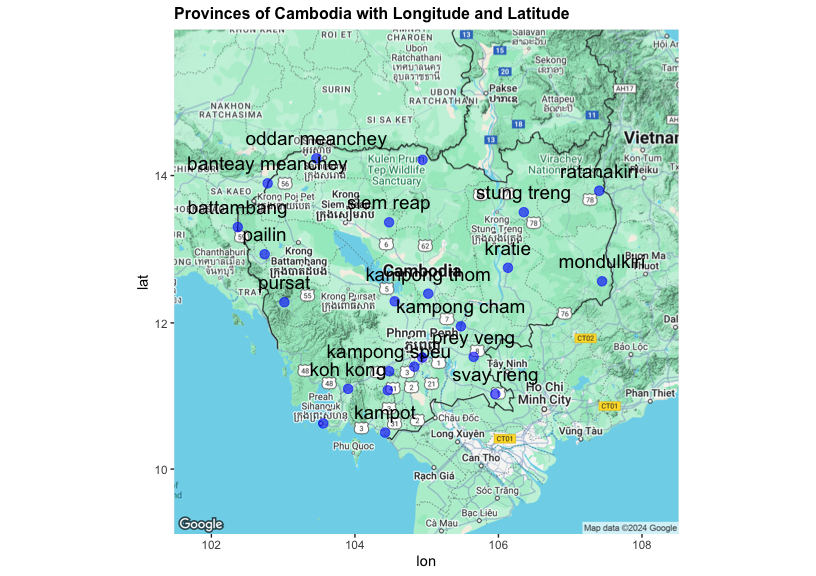
With the help of a color-coded legend at the bottom, this map shows the spatial distribution of incidences associated with explosive ordnance - mines and ERWs spanning across several years throughout Cambodia. To be able to plot this, I will first extract out the Year information and pass it into a new column, ‘Year’.



As indicated in the legend, the dots represent separate incidents when the incident occurred, with color corresponding to the different years in which it happened. There seems to be a high concentration of incidents in several locations of Cambodia, especially so in the Northwest.

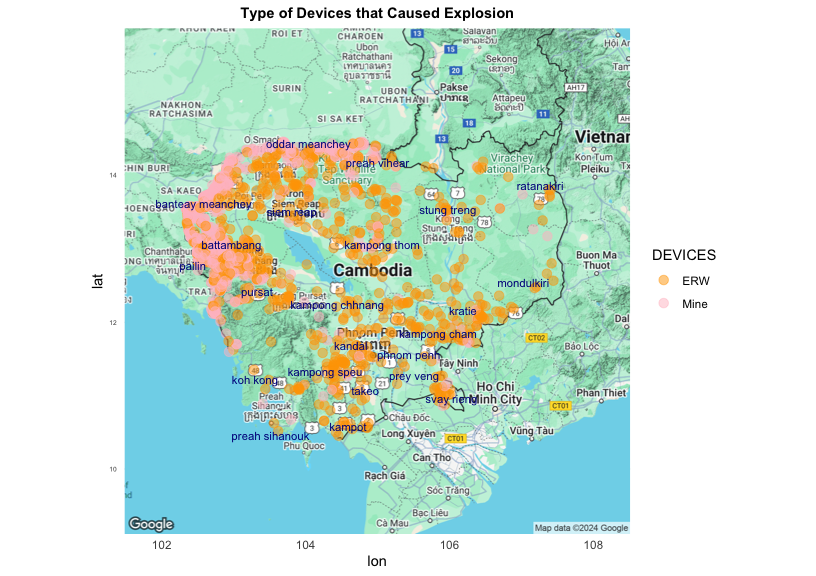
#### 4.2.3 Provinces of Cambodia Map Overview

The provinces of Cambodia are displayed on this map, with each one designated with a blue point that indicates its unique location based on latitude and longitude. These labeling of Cambodia’s provinces enables readers to recognize particular areas on the map.



#### 4.2.4 Type of device that caused the explosion in Cambodia

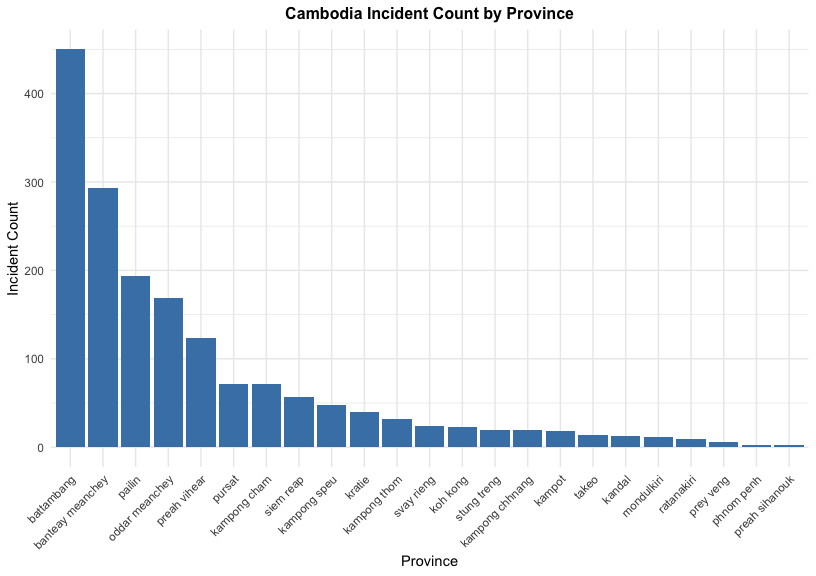
It shows the spatial distribution of the type of devices, whether is it caused by landmines or ERWs which resulted in an explosion in Cambodia.



The Northwest of Cambodia, including the provinces of Battambang, Pailin, Banteay Meanchey and Oddar Meanchey exhibit a discernible cluster of incidents. There are a lot of incidents on mines and ERWs that happened in these areas. Other than those provinces mentioned, there are also quite a number of incidents that occurred in other regions such as Kampong Thom and Siem Reap. It suggests that there seems to be more cases of explosions caused by ERWs as compared to mines, since the number of orange dots stands out more.

#### 4.2.5 Bar Plot to show Incident Count by Province

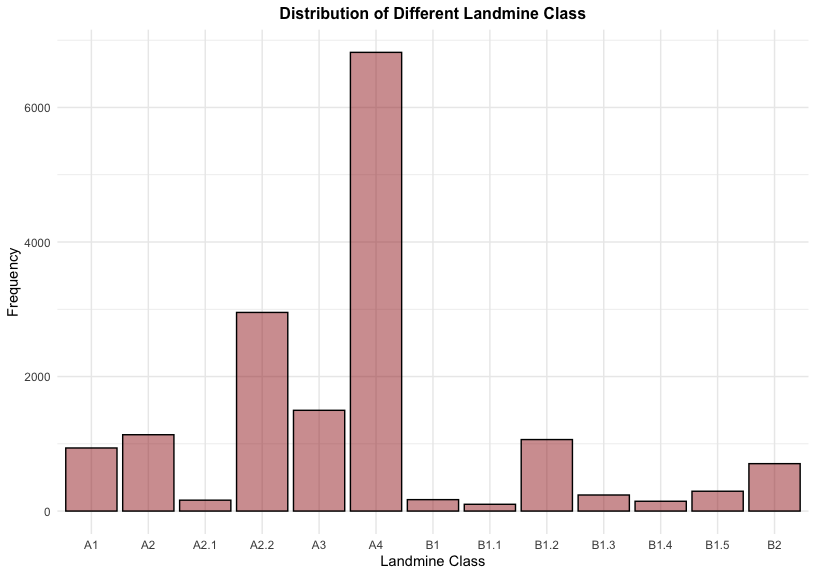
This bar plot illustrates the total number of incidents each province experience ERW and landmine incidents in the different provinces of Cambodia. I intentionally rearranged the incident count from the largest number to the lowest, for better visibility on the province that encounters the most number of incident counts.



This bar plot is ranked from the highest number to the lowest number of incident count. It is observed that Battambang has the highest incident count, followed by Banteay Meanchey and Pailin. It means that these 3 provinces experience the greatest hit from these explosives.

#### 4.2.6 Bar Plot to show Distribution of Different Landmine Class

This bar plot shows the frequency of different landmine classes found in Cambodia. The x-axis shows a unique landmine class, whereas the y-axis shows the number of occurrences for each particular landmine class.

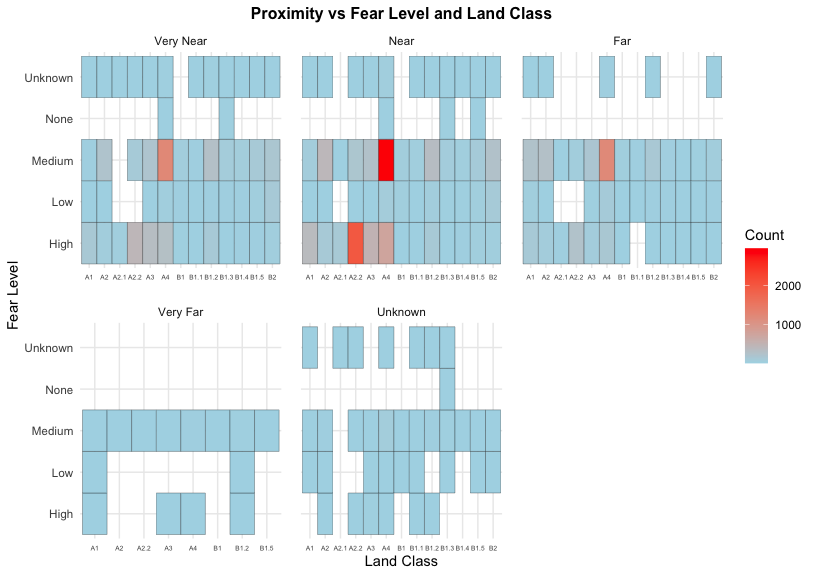


It is noted that this particular landmine class, A4, is the most prevalent since it has the highest frequency (tallest bar). It might suggest that, in comparison to other classes of landmines, this one was planted or found more frequently.

This could help focus demining efforts in places where specific classes are more common, particularly if some landmine kinds are more harmful or challenging to remove.

#### 4.2.7 Heatmap with Facets

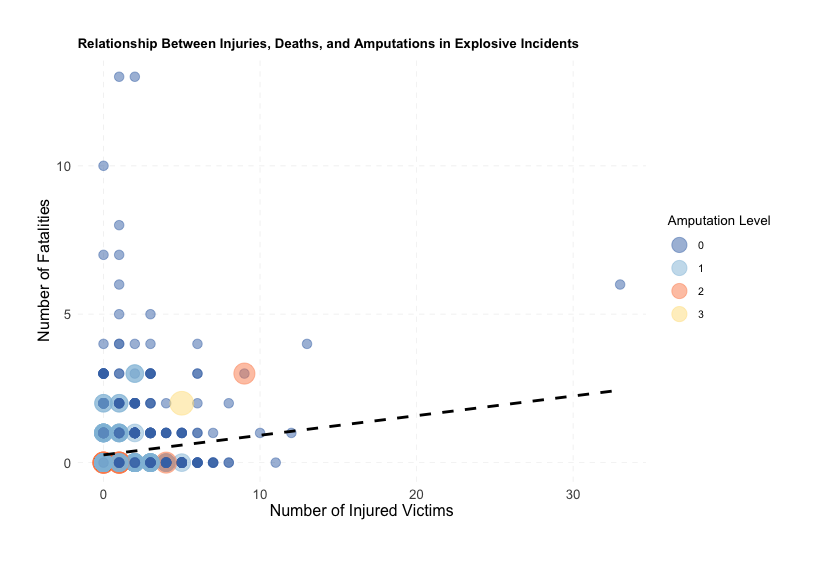
This heatmap matrix illustrates the correlation between 3 variables - proximity to danger, fear level and various landmine classes. It is divided into sections based on proximity to a potentially dangerous landmine area (very near, near, far, very far and unknown), land class (A1, A2, B4, etc) and fear level (none, low, medium, high and unknown).



Based on the heatmap above, we can conclude that for areas deemed as very near and near to landmine areas correspond to higher fear levels, as seen from medium to high categories. A2 and A4 landmine classes for example, emerge more frequently in the near proximity and elicit higher degree of fear level as compared to other proximities. Heatmap offers us a detailed look at how individual’s anxiety levels vary according to their proximity to potential landmine danger and the type of landmine they might face. Higher incident counts highlighted in red suggests that more attention should be channeled, where demining and intervention activities may need to be prioritized in order to lower risk and anxiety.

#### 4.2.8 Bubble Plot on Incident Severity - Scatter Plot with point size to represent Amputation

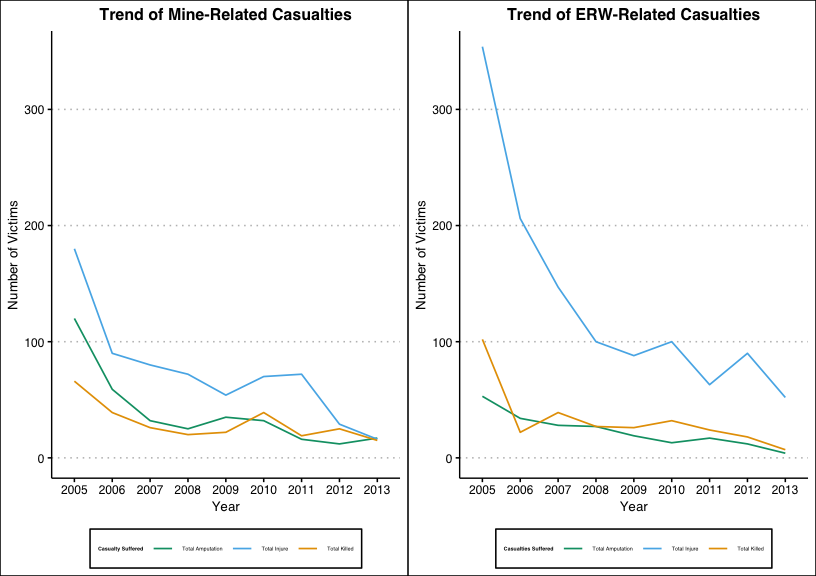
This scatter plot shows how the number of injured victims in explosive incidents correlates with fatalities and the level of amputations.



The dotted trend line shows a weak positive correlation between the number of fatalities and injuries, since the trendline is not steep. Majority of the points are clustered around the lower end of both axes, which implies that most of the incidents result in low injuries and death counts. There are also not much incidents which requires victim to undergo amputation, as there are not a lot of visible orange and yellow bubbles on the scatter plot.

#### 4.2.9 Trend lines to view how Mine and ERW Incidents have evolved over time

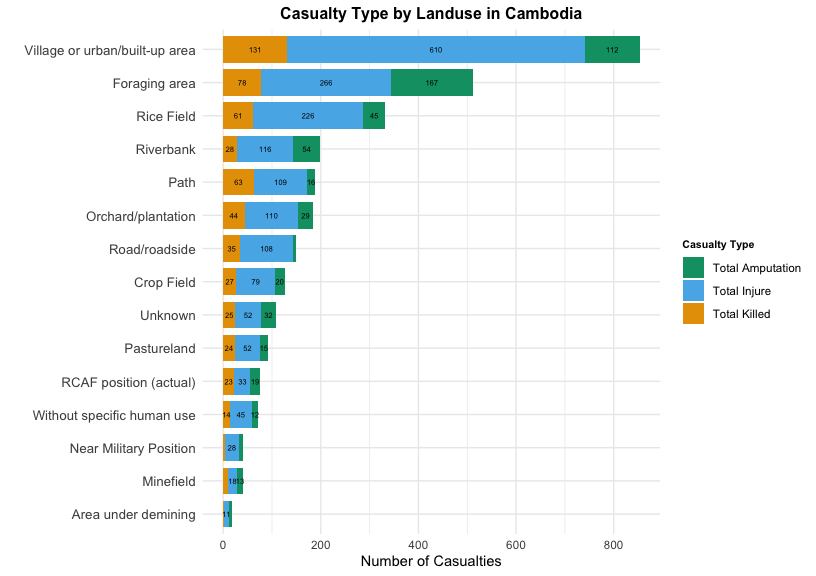
I purposely placed these two graphs next to each other and modified the y-axis to be the same for better comparison. The graphs displays the trends in mine-related and ERW-related casualties in Cambodia from 2005 to 2013. It shows the number of victims who were killed, got injured, or requires amputation, with all 3 degree of casualties each represented by a different color line.



Generally, the number of victims who suffered from casualties (amputations, injuries and deaths) have been showing a sharp decreasing trend over the years for both ERW and landmine incidents, which might suggest that the improvements are due to the enhanced effort to increase demining activities, better public awareness and higher emphasize on safety precautions. Nonetheless, there is still a low number of injuries, amputations and fatalities that continue to occur, indicating that these problems still exist, but on a much lesser scale.

#### 4.2.10 Bar Plots to Show Casualty Type by Landuse

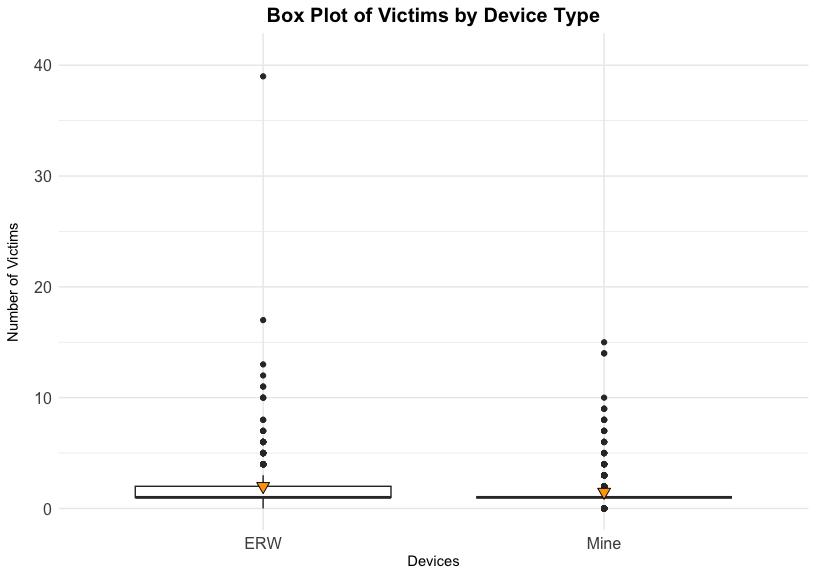
The distribution of casualties in Cambodia by kind of landuse is shown using the bar plot below, which is segregated by three different categories, namely: total amputation, total injure and total killed. We can conclude that the top three most dangerous landuse areas are located at village or urban/built-up area, followed by foraging area and subsequently rice field. These 3 landuse have resulted in the greatest casualties numbers. For this plot, I have omitted several values (mainly for casualty type: Total Amputation and Total Killed) because the number is very small and too insignificant to be included.



The regions of Cambodia with the highest death toll from landmine or ERW incidents are reflected in this plot. According to the bar plot, high-risk places include village or urban/built-up areas, foraging areas and rice fields. The most prevalent type of casualty is an injury, followed by amputation and death. The plot reiterates the existing dangers that mines and ERWs present to various landuse types in Cambodia, particularly in regions that are inhabited or used for agriculture purposes.

#### 4.2.11 Box Plot of Victims by Device Type

The box plot below shows the distribution of victims according to the types of explosive devices - mines and ERWs. The image shows that the median number of victims in both incidents is close to zero. The interquartile range is also very small, which implies that majority of the incidents has resulted in very few victims.



Even though the median and IQR in both devices suggests that it has resulted in very few victims, it is observed that there is an extreme outlier in ERW, with close to 40 victims in one particular incident, whereas landmine’s highest number of victims got hurt in any incident is at most, 15. This suggest that ERW incidents tend to result in a wider spread in the number of victims being harmed.

## 5. Conclusion

This research reveals that the aftermath from Cambodian civil war, despite having happened many years ago, still lingers and remain even till today. Analysis conducted on landmines and ERWs incidents in Cambodia reveals the disastrous effect that these incidents continue to have on the Cambodian communities. Even though the trend overtime indicates a decline in the total number of victims, landmines and ERWs continue to pose serious risks in several areas, such as Battambang, Banteay Meanchey and Pailin, where these provinces experience disproportionately high incident counts, which indicates the necessity for ongoing demining operations in these areas.

The breakdown of casualties by type of landuse emphasizes how vulnerable humans still are, especially in urban or rural settings. Communities that depend on agriculture for livelihood are especially vulnerable, as they are more susceptible to injuries, amputations and fatalities in places like villages, rice fields and foraging regions. Moreover, the box plot analysis of victims by device type further emphasizes that ERW accidents typically results in more severe implications in terms of the number of victims compared to landmine accidents, although both types of devices continue to have a negative impact on Cambodian communities.

This investigation serves as a potent reminder of the lasting effects of the Cambodian civil war. Even though the number of incidents have been on a decreasing trend overtime, demining operations and awareness campaigns must still continue because of the prevalent existence of these devices, especially in rural and high-risk areas. Targeted intervention might be necessary, especially in high-incidence provinces.

## 6. Appendix

#### Issue 1 - Missing ERW\_TYPE Values in mine.df

There are a total of 61 rows of NA fields where DEVICES is “ERW” and ERW\_TYPE is NA.

I will replace NA fields with “Unknown”.

updated.mine.df <- mine.df %>%  
 mutate(ERW\_TYPE = if\_else(DEVICES == "ERW" & is.na(ERW\_TYPE), "Unknown", ERW\_TYPE))

#### Issue 2 - Geometry variable not in latitude-longitude coordinate in mine.df

The geometry variable given in mine.dataset is not in the latitude-longitude coordinate.

I will convert it into latitude-longitude coordinate using st\_transform().

updated.mine.df$geometry <- st\_transform(updated.mine.df$geometry, crs = 4326)

#### Issue 3 - Convert data type for ‘VilCode’ in fear.df from “chr” to “num”

The data type for ‘VilCode’ in fear.df is originally “chr”.

I will convert ‘VilCode’ data type in fear.df from “chr” to “num”. I will have to put options(scipen = 999) to set options to display numbers without scientific notation.

options(scipen = 999) # Set options to display numbers without scientific notation  
fear.df$VilCode <- as.numeric(fear.df$VilCode) # Converting fear.df dataset VilCode from "chr" to "num"  
str(fear.df)

## sf [16,224 × 12] (S3: sf/tbl\_df/tbl/data.frame)  
## $ BLS\_Code : chr [1:16224] "BS/CMAA/00001" "BS/CMAA/00002" "BS/CMAA/00003" "BS/CMAA/00004" ...  
## $ District : chr [1:16224] "Anlong Veaeng" "Anlong Veaeng" "Anlong Veaeng" "Anlong Veaeng" ...  
## $ Province : chr [1:16224] "ODDAR MEANCHEY" "ODDAR MEANCHEY" "ODDAR MEANCHEY" "ODDAR MEANCHEY" ...  
## $ Commune : chr [1:16224] "Trapeang Tav" "Trapeang Tav" "Trapeang Tav" "Trapeang Tav" ...  
## $ VilCode : num [1:16224] 22010302 22010302 22010302 22010302 22010302 ...  
## $ Village : chr [1:16224] "Ta Dev" "Ta Dev" "Ta Dev" "Ta Dev" ...  
## $ Fear\_Level: chr [1:16224] "Medium" "Medium" "Medium" "Medium" ...  
## $ Land\_Class: chr [1:16224] "A3" "A2" "A2" "A4" ...  
## $ Proximity : chr [1:16224] "Very Near" "Very Near" "Near" "Near" ...  
## $ Operators : chr [1:16224] "CMAC" "CMAC" "CMAC" "CMAC" ...  
## $ SurveyDate: Date[1:16224], format: "2009-07-13" "2009-07-13" ...  
## $ geometry :sfc\_POINT of length 16224; first list element: 'XY' num [1:2] 384965 1569740  
## - attr(\*, "sf\_column")= chr "geometry"  
## - attr(\*, "agr")= Factor w/ 3 levels "constant","aggregate",..: NA NA NA NA NA NA NA NA NA NA ...  
## ..- attr(\*, "names")= chr [1:11] "BLS\_Code" "District" "Province" "Commune" ...

#### Issue 4 - Missing Fear\_Level Values in fear.df

The column, ‘Fear\_Level’ contains certain rows with “(blank)” values, which will give rise to incorrect interpretations if I were to drop these rows.

fear.df$Fear\_Level %>% unique()

## [1] "Medium" "(blank)" "Low" "High" "None"

I will replace the “(blank)” fields with the word, “Unknown”.

updated.fear.df <- fear.df %>%  
 mutate(Fear\_Level = if\_else(Fear\_Level == "(blank)", "Unknown", Fear\_Level))

#### Issue 5 - Geometry variable not in latitude-longitude coordinate in fear.df

Similar to Issue 2 identified earlier, the geometry variable given in fear.dataset is not in the latitude-longitude coordinate.

I will convert it into latitude-longitude coordinate using st\_transform().

updated.fear.df$geometry <- st\_transform(updated.fear.df$geometry, crs = 4326)

#### Issue 6 - Province names differ across both datasets

The names of the provinces in Cambodia differs across both mine and fear datasets. Fear dataset are all capitalized while mine dataset consists of first letter capitalized and the rest small letters. It is also observed that the province names either have incorrect spellings or requires fixing of spaces.

I will convert the names into lowercase letters, remove any leading/trailing spaces between characters. I will also standardize and fix the spelling or spacing issues for both datasets.

new.mine.df$Province <- tolower(trimws(new.mine.df$Province)) # Convert everything into lowercase and remove any leading/trailing spaces between characters  
new.fear.df$Province <- tolower(trimws(new.fear.df$Province))  
  
# Creating named vector with correct spellings for Cambodia's province names  
province\_names\_corrections <- c(  
 "mondol kiri" = "mondulkiri", # Correct to "mondulkiri"  
 "mondul kiri" = "mondulkiri", # Correct to "mondulkiri"  
 "kampong cham" = "kampong cham", # Already correct  
 "banteay mean chey" = "banteay meanchey", # Fix spacing  
 "otdar mean chey" = "oddar meanchey", # Correct to "oddar meanchey"  
 "siem reap" = "siem reap", # Already correct  
 "siemreap" = "siem reap", # Fix to "siem reap"  
 "preah vihear" = "preah vihear", # Already correct  
 "pursat" = "pursat", # Already correct  
 "battambang" = "battambang", # Already correct  
 "krong pailin" = "pailin", # Correct to "pailin"  
 "kaoh kong" = "koh kong", # Correct to "koh kong"  
 "kampot" = "kampot", # Already correct  
 "kampong speu" = "kampong speu", # Already correct  
 "kandal" = "kandal", # Already correct  
 "kracheh" = "kratie", # Correct to "kratie"  
 "kampong thom" = "kampong thom", # Already correct  
 "kampong chhnang" = "kampong chhnang", # Already correct  
 "stueng traeng" = "stung treng", # Correct to "stung treng"  
 "svay rieng" = "svay rieng", # Already correct  
 "takaev" = "takeo", # Correct to "takeo"  
 "prey veaeng" = "prey veng", # Correct to "prey veng"  
 "phnom penh" = "phnom penh", # Already correct  
 "rotanak kiri" = "ratanakiri", # Correct to "ratanakiri"  
 "ratanak kiri" = "ratanakiri", # Correct to "ratanakiri"  
 "krong preah sihanouk" = "preah sihanouk", # Correct to "preah sihanouk"  
 "sihanoukville" = "preah sihanouk" # Correct to "preah sihanouk"  
)

#### Feature Engineering 1

The original mine dataset provides me with the incident dates of the incident, in this format, YYYY-MM-DD. I will extract out only the ‘Year’ information, put it under a new column, so that I can plot my spatial plot distribution of incidents overtime.

# Spatial Plot Distribution of Incidents Overtime from cam.map  
new.mine.df$Incident\_D <- as.Date(new.mine.df$Incident\_D) # Change data type for Incident\_D into date type  
new.mine.df$Year <- format(new.mine.df$Incident\_D, "%Y") # Extract Year as character

#### Feature Engineering 2

Earlier on I have already converted my geometry points into the latitude-longitude coordinate system. I will then extract these points and create new separate columns, one for “longitude” and the other “latitude”. This is because when I am performing my Geospatial Analysis, Google Maps require latitude and longitude as independent inputs to plot points on the map. Compatibility with these tools, which assume separate handling of each dimension, is ensured by splitting coordinates.

# Converting geometry variable into the lat-long coordinate system using st\_transform  
updated.fear.df$geometry <- st\_transform(updated.fear.df$geometry, crs = 4326)  
  
print(class(updated.fear.df$geometry)) # To check if my geometry column is set up properly to represent spatial point data  
  
coordinates.fear.df <- st\_coordinates(updated.fear.df$geometry) # Extract coordinates  
coordinates.fear.df <- as.data.frame(coordinates.fear.df) # Convert coordinates.fear.df into dataframe  
colnames(coordinates.fear.df) <- c("Longitude", "Latitude") # Rename the columns of the extracted data  
  
new.fear.df <- cbind(updated.fear.df, coordinates.fear.df) # Merge coordinates back into original fear.df dataset

## 

## References

World Bank. (2024). Cambodia Overview. Retrieved from <https://www.worldbank.org/en/country/cambodia/overview>.

Open Development Cambodia. (n.d.). Landmines, UXO and Demining. Retrieved from <https://opendevelopmentcambodia.net/topics/landmines-uxo-and-demining/#ref-73803-6>

Open Development Cambodia. (n.d.). Mine/ERW Casualties 2005–2013. Retrieved from <https://data.opendevelopmentcambodia.net/dataset/mineerw-casualties-2005--2013>

Open Development Cambodia. (n.d.). Explosive Remnants of War (ERW). Retrieved from <https://data.opendevelopmentcambodia.net/dataset/erw>

Open Development Cambodia. (n.d.). Administrative Boundaries of Cambodia (2014). Retrieved from <https://data.opendevelopmentcambodia.net/dataset/administrative-boundaries-of-cambodia-2014>

## Appendix

