ANL 501 ECA

LIU KAILIANG



2023

Table of Contents

[Executive Summary 2](#_Toc146616510)

[Introduction 2](#_Toc146616511)

[Data 2](#_Toc146616512)

[Exploratory Data Analysis 5](#_Toc146616513)

[Conclusion 17](#_Toc146616514)

### Executive Summary

This report conducts a comprehensive data analysis on emergency food aid and conflict across 42 countries in the Sub-Saharan Africa region from 2002 to 2020. The dataset utilised for the analysis contains information regarding conflict, food aid, disasters, along with political and economic data of the countries. The report is structured into four primary sections. The introduction section gives a brief description of the background, objective analysis processes and key findings of the report. The Data section provides detailed information about the dataset used in this report, the methods used for data processing, and also the data extraction and integration processes of the data series from the WDI database. The Exploratory Data Analysis section uses visualizations to analyse the relationship between conflict and emergency food aid, starting from the overall relationship and then zooming in to specific countries and regions to find possible existing relationships. The final concluding section summarizes the findings and insights of this report.

### Introduction

The conflict and emergency food aid situation in Sub-Saharan Africa(SSA) has been a major challenge for the past 2 decades. The region has experienced a number of conflicts, including civil wars, inter-state wars, and insurgencies. These conflicts have displaced millions of people, disrupted agricultural production, and most importantly, impacting food security and causing poverty to people living in the region. The world has constantly been providing emergency food aid to countries in Sub-Saharan Africa, aiming to ease the negative impacts on food caused by conflict. However, there has been a constant debate about whether emergency food aid could be the fuel to the conflicts going on in the SSA region. Therefore,the objective of the report is discovering the relationship between conflict and emergency food aid through exploratory data analysis, generating insights from data that could potentially answer the question of whether emergency food aid leads to conflicts in the SSA region. The dataset used for the analysis contains information about conflict, food aid, natural disasters and other data related to the political and economical aspects of the SSA region. Analysis of the data starts with discovering the trend, patterns and associations of conflict, natural diaster, and food aid on the macro level.It is then drilled down to specific countries chosen for further analysis of potential factors that may cause emergency food aid to lead to more conflicts.The analysis shows positive relationship between natural disaster and food aid, it also discovers that the impact of emergency food aid on conflict varies among countries and the economic status might be associated with this situation

### Data

The “FoodAid” dataset I utilise for this analysis contains 798 observations from 42 countries in the Sub-Saharan Africa region, spanning from 2002 to 2020. It includes variables related to the impact of civil conflict, natural disasters, food aid information, political status, and some socioeconomic indicators such as GDP Per capita and population. There are some inappropriate datatypes, for example, “overall\_conflict”,‘minor\_conflict’, and “major\_conflict” should not be numeric values as they are binomial data. The “total\_affected\_disasters” variable is derived from summing up the “injured\_disasters”,“affected\_disasters”,and “homeless\_disasters”.

There are no duplicate records in the dataset but several variables have missing values, most notably on\_set\_war, off\_set\_war, battle\_deaths, and civilian\_deaths. Data from 2018 to 2020 of each country are absent in these variables, adding up to 126 NA values in each column. These NA records takes up approximately 16% of the total observations in each column. Removing these rows is not advisable as it could introduce bias. Therefore, I decide to keep these NA values and skip them when used in the analysis. However, having a substantial number of NA records may also lead to inaccurate analysis results. Moreover battle\_deaths and civilian\_deaths contain abnormal amount of “1” values, which could also influence the credibility of the analysis. Additionally, gdp\_per\_capita has 37 missing values, inflation has 34, polity2 has 7, and population has 6 missing values.

#Import the data  
library(readxl)  
FoodAid <- read\_excel("ANL501\_FoodAid.xlsx", sheet = 1)  
  
#check variables and data types  
str(FoodAid)  
#check for NA values  
na\_counts <- colSums(is.na(FoodAid))  
na\_counts   
  
#Check for duplicated record  
sum(duplicated(FoodAid))

##### WDI Data Extraction and Integration

As I am interested to discover if the level of corruption in a country associate with emergency food aid or civil conflict, I extracted the “CC.EST” series from the World Development Incubator Package. According to the World Development Incubator, this series stands for “Control of Corruption: Estimate”, which “captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as “capture” of the state by elites and private interests. “This essentially reflects the level of corruption in the country with -2.5 being most corrupted and 2.5 being the opposite. I want to examine my hypothesis that corruption causes conflict and reduces the positive impact of emergency food aid.

I extracted the data of all countries from the period of 2002 to 2020. I then selected only “country”, “year” and “CC.EST”. To join this data to the main dataset, I first changed the format of the “year” column from the main dataset with the “lubridate” package to match with the corresponding records in the WDI data. I then changed the data type of “year” in both datasets to character as they should not be numeric. Finally, since the WDI data extracted was already in the long format, I simply joined these datasets on the “country” and “year” columns.

#Use the WDI package to fetch the Control of corruption:Estimate variable  
install.packages("WDI",repos = "http://cran.us.r-project.org")

## 程序包'WDI'打开成功，MD5和检查也通过  
##   
## 下载的二进制程序包在  
## C:\Users\Kailiang\AppData\Local\Temp\RtmpCy0hG4\downloaded\_packages里

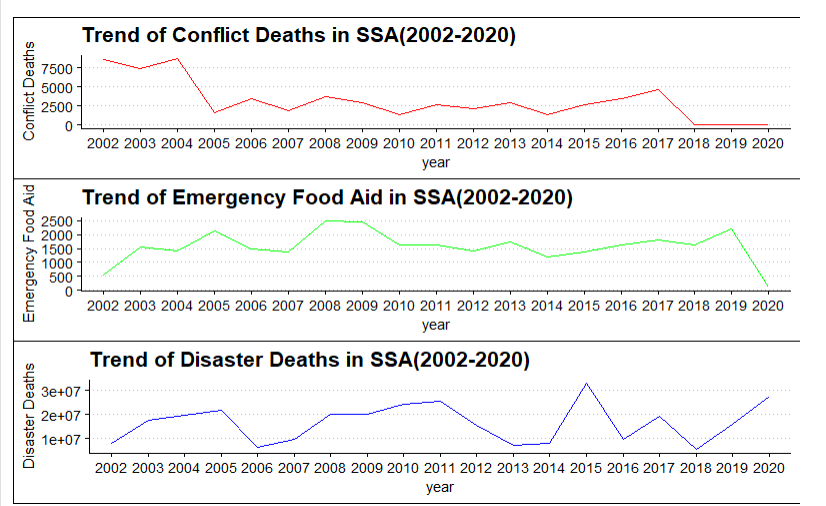
library(WDI)  
wdi\_data <- WDI(country="all", indicator="CC.EST", start=2002, end=2020) #Control of corruption:Estimate  
  
#Select only the country, year, and CC.EST columns  
wdi\_data<-  
 wdi\_data%>%  
 select(country,year,CC.EST)%>%  
 mutate(country = ifelse(country == "Congo, Dem. Rep.", "Democratic Republic of the Congo", country))  
   
#Change the format of the "year" column from the "FoodAid" data frame with the 'lubricate' package to display in yyyy format  
library(lubridate)  
FoodAid$year<-year(FoodAid$year)  
  
#Change the data type of 'year' to character for both tables  
FoodAid$year<-as.character(FoodAid$year)  
wdi\_data$year<-as.character(wdi\_data$year)  
  
#Left join this variable to the food aid data frame using country and year  
FoodAid<- FoodAid %>%  
 left\_join(wdi\_data, by = c("country" = "country", "year" = "year"))  
  
head(FoodAid)

## # A tibble: 6 × 26  
## country year overall\_conflict minor\_conflict major\_conflict onsetwar  
## <chr> <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 Angola 2002 1 1 0 0  
## 2 Angola 2003 0 0 0 0  
## 3 Angola 2004 1 1 0 1  
## 4 Angola 2005 0 0 0 0  
## 5 Angola 2006 0 0 0 0  
## 6 Angola 2007 1 1 0 1  
## # ℹ 20 more variables: offsetwar <dbl>, battle\_deaths <dbl>,  
## # civilian\_deaths <dbl>, death\_disasters <dbl>, injured\_disasters <dbl>,  
## # affected\_disasters <dbl>, homeless\_disasters <dbl>,  
## # total\_affected\_disasters <dbl>, total\_affected\_othercountries <dbl>,  
## # total\_affected\_neighbours <dbl>, total\_affected\_non\_neighbours <dbl>,  
## # nda\_other\_region <dbl>, emergency\_food\_aid <dbl>,  
## # non\_emergency\_food\_aid <dbl>, total\_aid <dbl>, gdp\_per\_capita <dbl>, …

### Exploratory Data Analysis

##### Analysing Trends

To start the exploratory data analysis, I analysed the overall trend of conflict, emergency food aid and natural disaster in the SSA region by creating line plots to reveal the time series trend. The purpose of visualising this is to find out any similarities and differences in the progress of each variable and whether there is an association between these variables. I chose the sum of battle death and civilian deaths as the conflict indicator. Although 16% of these data are missing, they are still suitable variables as they imply the degree of seriousness of the conflicts in a country, and no data is missing from the rest of the years. An alternate variable, overall\_conflict could be used as a reference to infer the conflict situation from 2017 to 2020. I used total\_affected\_disaster as the natural disaster indicator as it is an aggregation of various impacts caused by disasters.”Emergency\_food\_aid” is used to plot the time series plot for emergency food aid.



Trend of Conflict Deaths in SSA (2002-2020):

The plot shows fluctuations in the number of conflict deaths over the years. There is a noticeable declining trend stating from its peak in the year 2002 till the year 2005, conflict deaths then keep at a relatively lower level from 2005 till 2014. The line starts to go upwards from 2014 to 2017. Although data for the year 2018,2019,2020 is missing, the over\_conflict data displays an upward trend within this period.

Trend of Disaster Impacts in SSA (2002-2020):

The number of people affected by disasters also shows fluctuations over the years by the fluctuations are relatively more violent. There are noticeable spikes in the years 2005, 2011, and 2015 and 2020, which means that major natural disasters could have occurred in these years. Other years show relatively lower numbers of people affected by disasters.

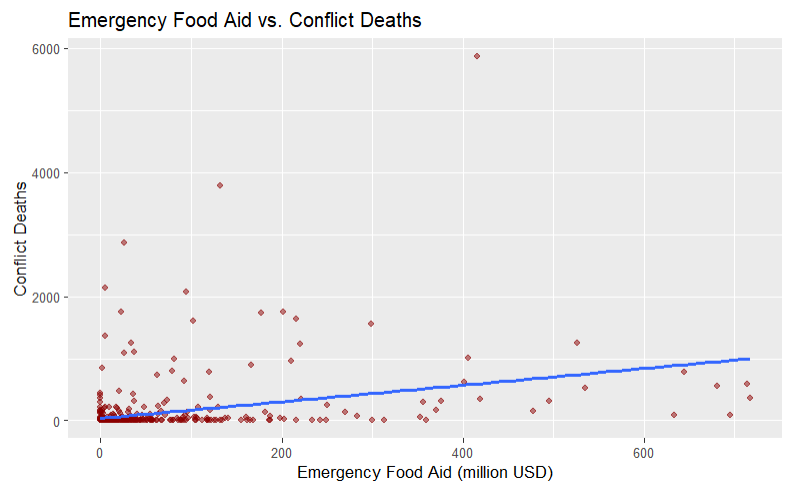
Trend of Emergency Food Aid in SSA (2002-2020):

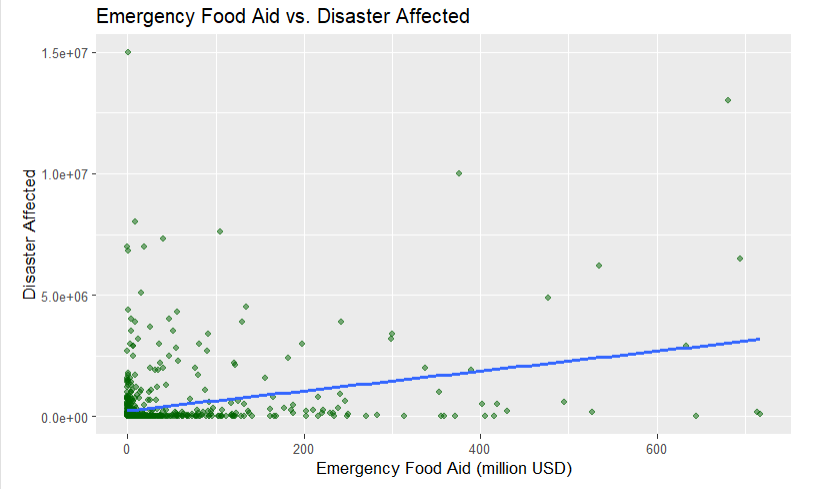
The trend of emergency food aid appears to be relatively stable with slightly less fluctuations. There is an increasing trend in emergency food aid stating from the year 2002 till the year 2005. There is a significant decrease in emergency food aid in the year 2020 possibly due to COVID-19.

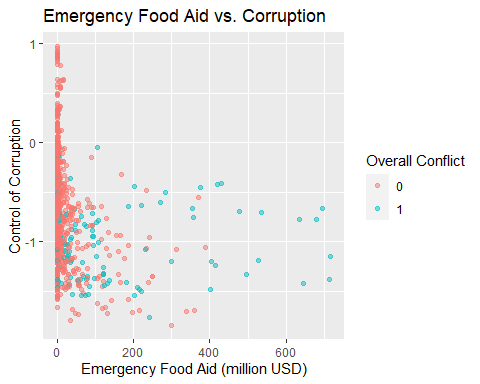
It is interesting to observe that the trend of conflict deaths and the trend of emergency food aid appear to be inversely proportionate to each other during the period from 2002 to 2005, but it is too early to draw the conclusion that conflict and emergency food aid are negatively related as this characteristic is hard to observe in the following years, where the trend lines of conflict death and emergency food aid display similar pattern. It is also observed that the shapes of the trend line of disaster impacts and emergency food aid are similar before they get very different from 2014 onwards.

##### Correlation Between Conflict, Food Aid and Natural Disaster:

To further discover any associations among the variables, I plotted several scatter plots aiming to discover any observable relationships. Scatter plots are suitable for visualising relationships between variables. Apart from the conflict, emergency food aid and natural disasters variables, I am also interested in finding out whether there is a relationship between corruption control and emergency food aid received. To plot conflict vs emergency food aid, I created the variable “conflict\_death\_sum” in the “Food Aid” data set by adding up battle deaths and civilian deaths with the mutate function.







Looking at the 2 scatter plots demonstrating food aid vs conflict deaths and food aid vs disaster affected as a whole, they do not seem to display obvious relationship between conflict and emergency food aid and also the relationship between natural disaster and emergency food aid. Majority data points cluster around the origin and along either of the x and y axis. The reason for this could be due to the quality of the data source or uneven distribution of conflict, food aid and disaster. However, data points that are not clustered around the origin shows a general downward trend.

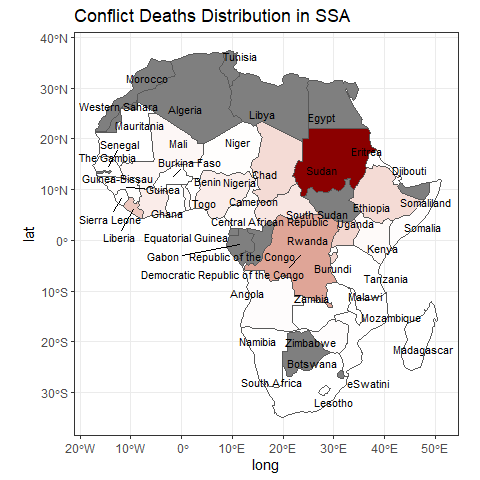
The Food Aid vs Corruption scatter plot displays a clearer pattern. the data points are coloured by “overall\_conflict” to find out the association between conflict and corruption. Although many data points are clustered along the y axis, It can be observed from the plot that countries which has negative corruption control scores are likely to receive more food aid. Moreover, these countries are more prone towards conflict as all of the blue points are below 0 corruption control estimate.

To examine whether the messiness in these scatter plots is due to uneven distributions of conflict, emergency food aid and disaster, I would like to visualise them in the next part of the analysis.

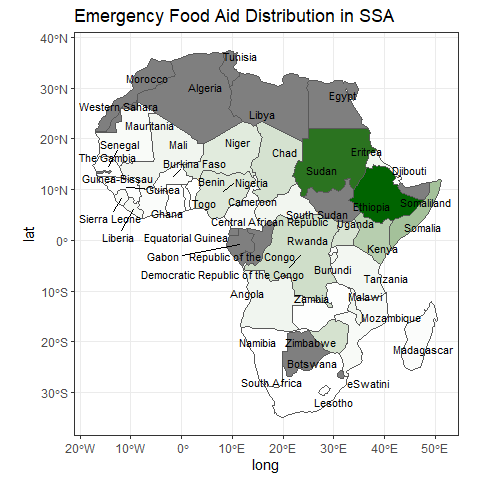
##### Geographical Distribution of Conflict, Disaster and Food Aid

I would like to discover the distributions of these variables with the combination of bar charts and spacial plots. Bar charts are very suitable for identifying the top few countries with high level of conflicts, disasters and emergency food aid whereas spacial plots clearly display the geographical locations of those countries, allowing me to identify potential clusters of conflicts, natural disasters or emergency food aid.

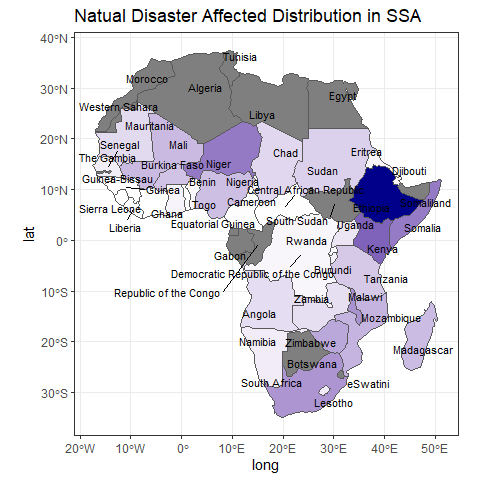
To plot theses visualisations, I grouped the FoodAid data frame by country and aggregated the sum of selected variables



This plot visualises the geographical distribution of accumulated conflict deaths in SSA.The colour scale represents the total number of conflict deaths in each country, the darker the colour, the higher the number of conflict deaths. It is observed that the distribution is uneven, most conflict deaths come from Sudan and its neighbouring countries in the central Africa region. Several countries in the Western Africa region also had noticeable amount of conflict deaths. Little data can be observed in the Southern Africa region.

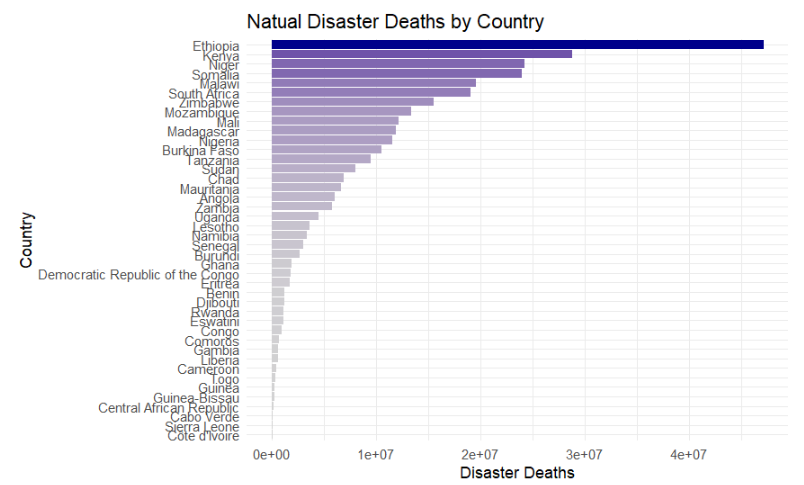
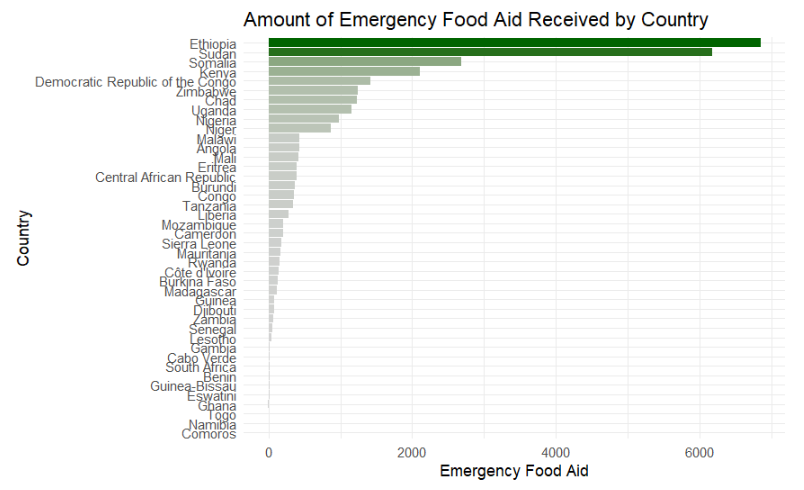
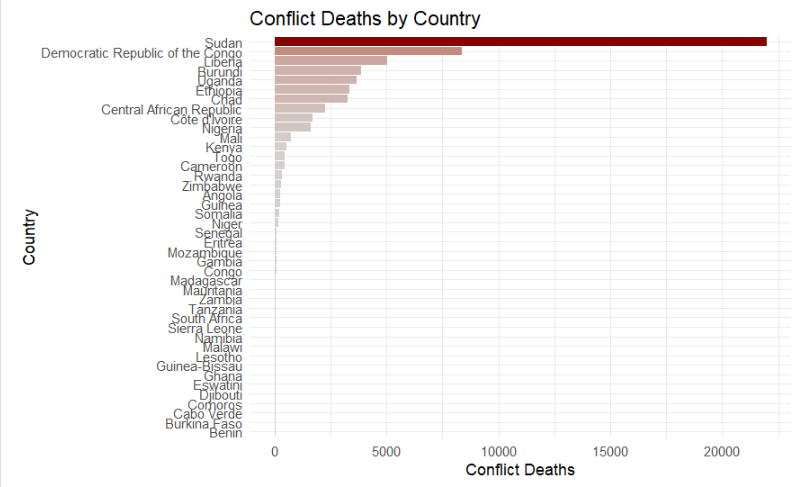


This plot visualises the geographical distribution of accumulated emergency food aid in SSA.The colour scale represents the total amount of food aid received in each country from 2002 to 2020, the darker the colour, the higher the amount of food aid received. The plot shows that the highest levels of emergency food aid are distributed in Ethiopia and Sudan. Emergency food aid is relatively evenly distributed compared to conflict deaths, but several countries in the Southern Africa and Western Africa regions generally receive less to no food aid.



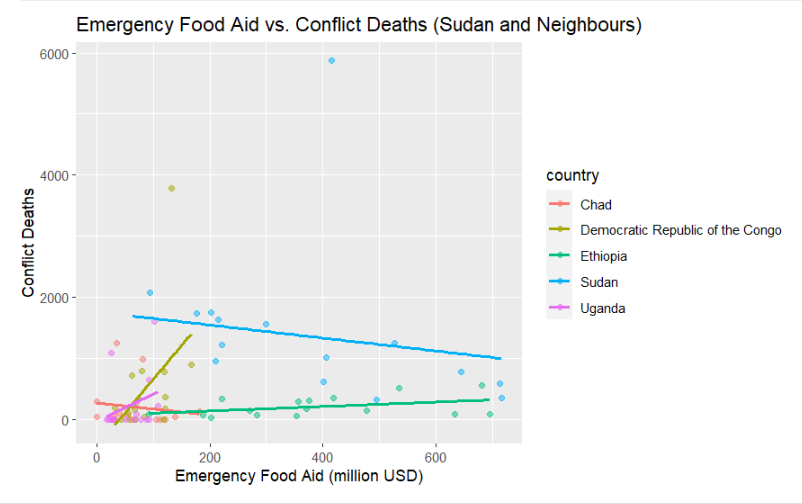
This plot visualises the geographical distribution of the accumulated number of people affected by natural disasters in the SSA region. The colour scale represents the total number of people affected by natural disasters in each country from 2002 to 2020.The darker the colour, the higher the number. The plot shows that natural disasters have caused major damages to Ethiopia where the colour is the darkest.

By looking across these 3 spacial plots and comparing them, it is found that Sudan and Ethiopia, being countries with the highest level of conflict deaths and natural disaster impacts respectively, received highest level of food aid. This information helps me to locate the countries to focus on in the subsequent analysis.

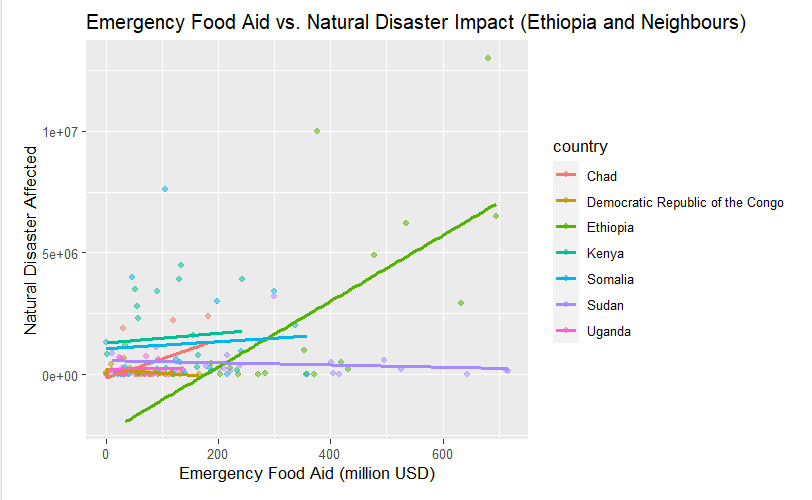
 These bar chats, ordered in descending order, supports the spacial plots by clearly showing what are the top countries in each plot. Having this information revealed, I can conduct analysis on the country level to discover the relationship between conflict, natural disasters and emergency food aid in key countries.

##### Country-level Analysis

In this part of the data analysis, I will first focus on conflict and food aid in several countries as the sample. Countries selected are Sudan, Ethiopia, Democratic Republic of the Congo, Chad, and Uganda. These countries are closely located and they take up a substantial amount of conflict deaths and emergency food aid in the region. I will plot a scatter plot to start my analysis.

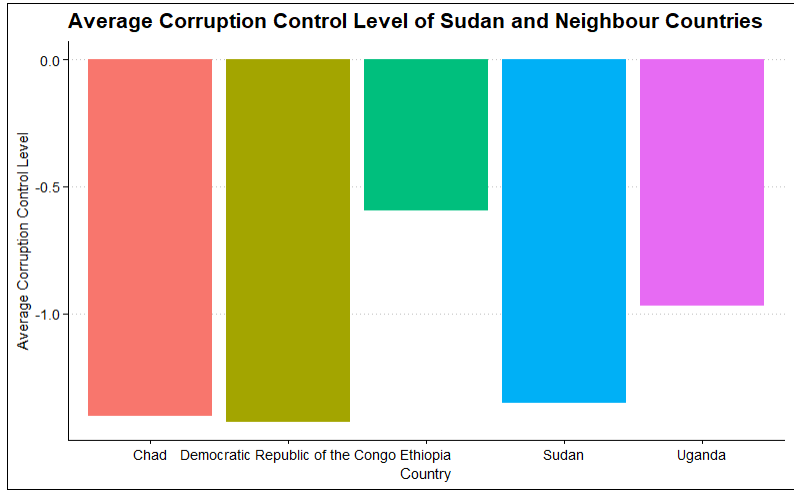


It is shown in this plot that the relationship between conflict deaths and emergency food aid differs by country. In Sudan and Chad, amount of conflict deaths seems to negatively relates to emergency food aid. On the other hand, amount of conflict deaths seems to positively relates to emergency food aid received in Democratic Republic of the Congo, Ethiopia and Uganda. This implies that impact of emergency food aid on conflicts can be different in different countries, there are other factors to consider such as natural disaster.



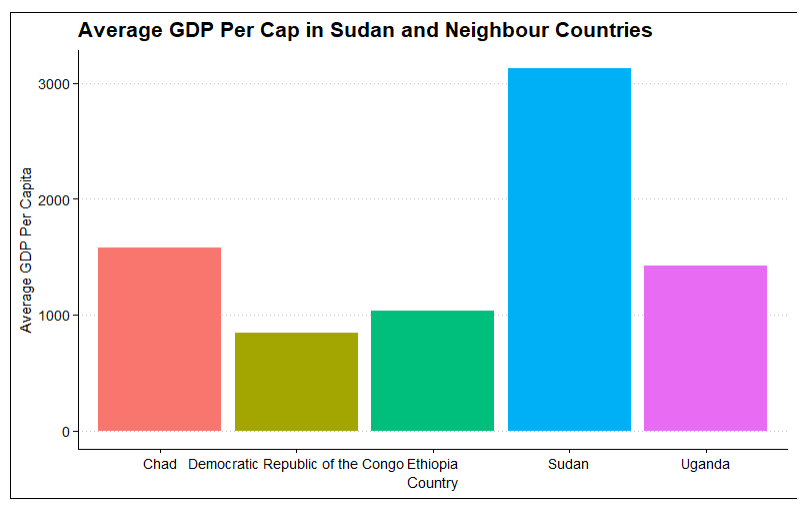
To find out the relationship between natural disaster and emergency food aid, the plot uses data from Ethiopia and its neighbouring countries which take up a substantial proportion of natural disaster affected and emergency food aid in the region. Countries used to plot the previous plot are also included to find out whether natural disaster is a potential cause of the differences found earlier. This plot reveals a weak positive relationship between emergency food aid and Natural Disaster Affected in these countries. Overall, more food aid is provided to areas suffering bigger losses from natural disasters and no clear sign shows that natural disaster in the country influences whether emergency food aid reduces conflict or leads to more conflict.

There might be many factors that determine whether food aid leads to more conflicts in a country I am interested in corruption control. It is found from the corruption\_foodaid\_scatter plot that corrupted countries are prone toward conflicts and emergency food aid often flows to them. Does more corrupted countries have reduced effect of food aid on conflict?



To answer this question, I have plotted a bar chart showing the average control of corruption estimate score of the sample countries that I selected for analysing the relationship between conflict and food aid. I have identified that Democratic Republic of the Congo Ethiopia, and Uganda displays proportional relationships to different extends. However, in this plot, their CC.EST scores do not imply that they are more corrupted than the other 2 countries.

There are other variables that could be used for further analysis. Apart from corruption control, I am also interested in whether food aid leads to more conflict in poorer countries.



The plot demonstrates the average GDP per capita of the selected countries. It is observed from the plot that Democratic Republic of the Congo, Ethiopia and Uganda have relatively lesser GDP per capita compared to Sudan and Chad. This observation could potentially reflect that in poorer countries, it is possible that food aid could lead to more conflicts. However, this observation cannot represent the entire SSSA region as it is only derived from data representing parts of the region

### Conclusion

The analysis in the report finds out that the association between the overall trend of conflict and emergency food aid, as well as natural disasters, expresses different and similar characteristics at different points of time. It is also found that the distribution of conflicts is uneven, with a large proportion of conflicts locating in Sudan and its neighbouring countries. Although there is not a noticeable correlation between conflicts and emergency food aid of the entire SSA region overall, by zooming into these areas, both proportional and inverse proportional relationships of food aid and conflict is observed. Thus, it can be concluded that whether emergency food aid leads to more conflict in the SSA region differs in different countries. While corruption may not a driven factor, one potential factor could be the economic status of that country. Countries with lower GDP per capita are like to have more conflicts when receiving emergency food aid. However, there are still insufficient evidences to support the statement that emergency food aid leads to more conflict, A more extensive exploration into additional factors is necessary for a comprehensive and conclusive understanding.