# Visualizing Armed Conflicts

*Alan Peral and Kadija Yilla*

*12/15/2017*

## Introduction

Since the end of the cold war in 1991, 2016 has been the fifth most violent year in the world (Dupuy et al. 2015). Currently in 2017, there are more than 40 active conflicts in the world. Armed conflicts severely affect the lives of individuals in those countries. Most of the data online outlines the armed conflict and casualties of the many occurring wars. Others also include an interactive feature of the global conflict to observe the scale of the conflict with respect to other nations. One special interactive map illustrates how various armed conflict affect US interests. Despite all the information surrounding global conflicts, none accurately depict how the conflict has affected the lives of the individuals alive and living amidst the war.

War impacts people profoundly, and yet countries like the US insist on engaging in numerous armed conflicts regardless of the outcome. Westernized people, on account of the privilege awarded by not being engaged in the conflict on a daily basis, rarely understand the impact of war on the population. How has the war impacted education? Are more males leaving school as a result of the conflict? What about women? How has the war impacted the need to have equal education for women? How has the conflict caused the governmental expenditures to changes? Is there more funding for the various industries or is a majortiy of the funding being directed towards the conflict? In addition, do the citizens have access to water, electricity, housing, and other basic human necesities?
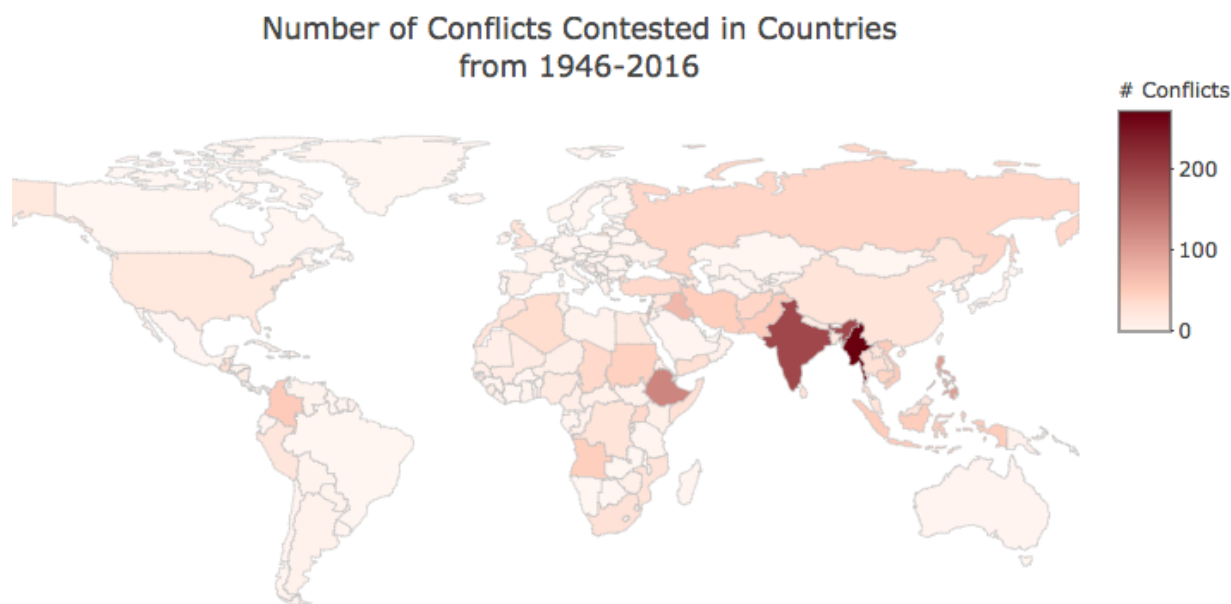


Figure 1: Concentration of conflicts worldwide. This is an example of one tab of our application.

Our application, shown above, offers a tool to explore the relationship between quality of life and armed conflict in various countries. Although we don't think we can answer in full the questions above, our application offers a glimpse to what those answers can look like for various countries relative to each other. In addition, we are interested in evaluating whether there is a difference in locations of armed conflicts in western countries versus the global south and the network of warring sides. We intergrated the wealth of

information available for armed conflicts and various findings on coutry-level data to provide a synthesized view of how armed conflicts affect various nations.

# Data

The first data set used is from Kaggle titled "Country Socioeconomic Status Scores." This data set includes the overall score of socioeconomic status by country for every decade since 1880. Socioeconomic Status (SES) Scores measure the accesibility of an individual, household, or community to collective resources. This measure incoprorates, income, wealth, health, and occupation (Psaki et al. 2014). For each country, this indivudialzied measure is aggregated as a whole to provide the measure for the country.

```
## # A tibble: 5 x 6
##    wbid      country  year   SES gdppc popshare
##    <chr>       <chr> <int> <dbl> <dbl>    <dbl>
## 1   AFG Afghanistan  2010  5.68  1663  0.00415
## 2   AFG Afghanistan  2000  2.06   565  0.00331
## 3   AFG Afghanistan  1990  1.27   604  0.00235
## 4   AFG Afghanistan  1980  3.47   690  0.00306
## 5   AFG Afghanistan  1970  3.47   709  0.00310
```

The second set is world data from the World Bank. The data was accessed from an r package `WDI`. Using the r package we choose 17 indicators which we believed were good measures of the well-being of the individuals in a country. The indicators were as follows:

- health expenditure, total (% of GDP)
- fertility rate, tota (births per total)
- life expectancy at birth, female (years)
- life expectancy at birth, total (years)
- mortality rate, under-5 (per 1000 live births)
- children in employment, total (% of children 7-14)
- labor force, female (% of total labor force)
- labor force participation rate, female (% of female population ages 15+)(modeled ILO estimate) indicator
- GINI index (World Bank estimate) indicator
- Refugee population by country or territory of origin
- Refugee population by country or territory of asylum
- Improved sanitation facilities (% of population with access) indicator
- Improved water source (% of population with access) indicator
- Access to electricity (% of population) indicator
- Population living in slums (% of urban population) indicator

```
##       country year fetility_rate life_expectancy_female
## 1 Arab World 1960          6.95                   47.9
## 2 Arab World 1961          6.97                   48.5
## 3 Arab World 1962          6.99                   49.1
## 4 Arab World 1963          7.01                   49.7
## 5 Arab World 1964          7.02                   50.3
```

The third and final set is data about armed conflicts from the Uppsala Conflict Data Program (UCDP) at the department of Peace and Conflict Research, Uppsala University and the Centre for the Study of Civil War at the Peace Research Institute Oslo (PRIO). The UCDP/PRIO Armed Conflict Dataset (Version 17.1) represents both internal and external conflict from 1946 to the present. It documents the two sides in the conflict, as well as the location of the conflict, and indicators of intensity and number of fatalities. Armed conflict is defined as (2002):

> "A contested incompatibility that concerns government and/or territory where the use of armed

> force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in a calendar year" – Gleditsch et al. (2002)

```
## # A tibble: 5 x 4
##               location                 sidea
##                  <chr>                 <chr>
## 1              Bolivia Government of Bolivia
## 2              Bolivia Government of Bolivia
## 3              Bolivia Government of Bolivia
## 4 Cambodia (Kampuchea)  Government of France
## 5 Cambodia (Kampuchea)  Government of France
## # ... with 2 more variables: `side b` <chr>, year <int>
```

**Wrangling**

Using the three seperate sets, we were able to construct a comprehensive data set of all necessary information using data wrangling techniques. First we merged the SES scores set, **SES.data**, with the WBD data, **WBD.data2** by merging based on the country first and then the year. Before the merge was successful, the country names had to be reconciled because of subtle differences in naming techniques. For example, "Korea, Dem. People's Rep." in the WBD set had to be renamed to "North Korea" as that was the name in the SES set. We choose to rename the countries based on names we thought the general public would be more familiar with.

```
##        country unid wbid  SES
## 1 Afghanistan    4  AFG 38.0
## 2 Afghanistan    4  AFG 29.6
## 3 Afghanistan    4  AFG 28.1
## 4 Afghanistan    4  AFG 27.0
## 5 Afghanistan    4  AFG 27.0
```

Next we take the established **WBD.SES** set and merge with the **conflict.data** set based on the "country"/"location." In order to create a tidy data set upon merge, we had to clean-up the conflict set for the "location" variable. For some conflicts in which the conflict occured in different locations globally, we had to seperate the countries in the "location" field and recreate the data for the seperately valued contries. To achieve this we had to create a user-function that replicated the speicified joined countries and duplicated the values in one step. For example, the India vs. Pakistan War of 1965 occured in both India and Pakistan but was represented as "India, Pakistan".

```
## # A tibble: 5 x 4
##            location               sidea               `side b`  year
##               <chr>               <chr>                  <chr> <int>
## 1 India, Pakistan Government of India Government of Pakistan  1948
## 2 India, Pakistan Government of India Government of Pakistan  1964
## 3 India, Pakistan Government of India Government of Pakistan  1965
## 4 India, Pakistan Government of India Government of Pakistan  1971
## 5 India, Pakistan Government of India Government of Pakistan  1984
```

This was instead changed to:

```
## # A tibble: 5 x 4
##   location               sidea               `side b`  year
##      <chr>               <chr>                  <chr> <int>
## 1    India Government of India Government of Pakistan  1948
## 2 Pakistan Government of India Government of Pakistan  1948
## 3    India Government of India Government of Pakistan  1964
## 4 Pakistan Government of India Government of Pakistan  1964
```

```
## 5    India Government of India Government of Pakistan  1965
```

In addition, we had to conduct a few name changes to reconcile the differences in the set in order to have a bigger join. The final set which we will be working with is displayed below. The set excludes the regions created in the WBD set for summary purposed such as "Small States" and "Low Middle Income" countries. This set is comprehensive as it includes the SES score for the decade, other country-specific indicators, and information about the conflict for a specific year.

```
##          location year                    sidea
## 1    Afghanistan 1990 Government of Afghanistan
## 2         Albania 1990                     <NA>
## 3         Algeria 1990                     <NA>
## 4 American Samoa 1990                      <NA>
## 5         Andorra 1990                     <NA>
##   labor_force_Particpation_rate   SES
## 1                         16.44  1.27
## 2                         53.15 72.88
## 3                          9.93 56.71
## 4                            NA    NA
## 5                            NA    NA
```

# Results

The entire shiny application can be accessed by visiting: https://yillak.shinyapps.io/armed_conflicts2/

There are a wide range of observations seen through the various apps on the page. When viewing the map graph we see that the countries with a majority of the conflicts are in the global south. The two countries with the most conflicts are India and Mynamar. Doing further research with our data, we observe that the United Kingdom is the first western country in the list and that occurs in the 26th place, having 29 conflicts occuring in that location. But in total, the UK has been involved with 203 armed conflicts globally. Similarly the United States has been in 173 conflicts but only 21 of those have occured within the country. On the other hand, India has been involved with 218 conflicts but 194 have occured in India. This strengthens our hypothesis that although most conflicts might involve western countries, they are not fought in those countries and so most of the effects of the war will be seen not in the western country, but in the country on the other side of the war.

| Country | Number of Conflicts (on location and off) | Number of Conflicts on Location |
|---|---|---|
| United States | 173 | 21 |
| United Kingdom | 203 | 29 |
| Myanmar | 272 | 271 |
| India | 218 | 194 |
| France | 177 | 5 |
| Ethiopia | 203 | 127 |

Using these preliminary results we move on to the "plot" tab to see how the various indicators for quality of life affect the countries at war. We observe that westernized countries, exemplified by the United States and the United Kingdom, score rather highly in most of the variables. They have an above average SES score even in years following huge conflicts. On the other hand, India and Myanmar have SES scores below 25 and are showing no trends of growth over the last 25 years (figure 2). In other indicators such as health expenditure, the trends are as previously described.

One special indicator of interest is the refugee origin/asylum. We see that refugee origin is low, close to zero, for various westernized countries (figure 3). But we see this number fluctuate for other war-torn coutnries. We see a different effect when looking at refugee asylum. Most refugees turn to western countries when seeking asylum (figure 4). This is particularly interesting because the western countries are in as many conflicts, if not more, than the countries in which the individuals are coming from. But because the conflicts aren't occuring in those nations directly, the location is seen as safer.



Figure 2: SES indicator (score from 1 to 100)

The final aspect of the app is a network graph connecting warring sides. When we unlimit the graph to include all entities regardless of number of conflicts, we observe that there is a connection between all of them through other entities. But on the other extreme, by increasing the number of conflicts, we observe two disconnected graphs. The two graphs have node centers of India and Myanmar. This makes sense because they have had the most conflicts and, interestingly enough, have not been in direct conflict with each other.

**Limitations**

Although our data and app allows us to answer some of our questions, it's important to understand the limitations in the set before we, or anyone, can make generalizations. The most important limitation is in the conflict set. As stated in the code book, one biasedness is as follows:

> "Apart from that, the dataset only includes information when we are quite confident that it is correct. The bias produced by this approach is against the inclusion of conflicts in the earlier decades and in the less-developed world. An armed conflict in a developed country in the 1990s is more likely to be recorded than a conflict in a less developed country in the 1950s." – Gleditsch et al. (2002)

This means that some conflict which occured in the middle of the 20th century might be omitted for lack of verifiability. This would be especially the case for less developed countries. The definition of conflict, as mentioned previously, also causes bias towards which conflicts are used. This will cause our analysis to be incorrect in some sense. For example, because of its definition, it might include singular attacks in which over 25 people are killed. One instance is the Omagh bombing in Northern Ireland in 1998. This event, and others like it, would not have significant impact on any of the country-wide indicators used in the plot.

Another limitation will be in the amount of missing information from the World Bank Data set. None of the
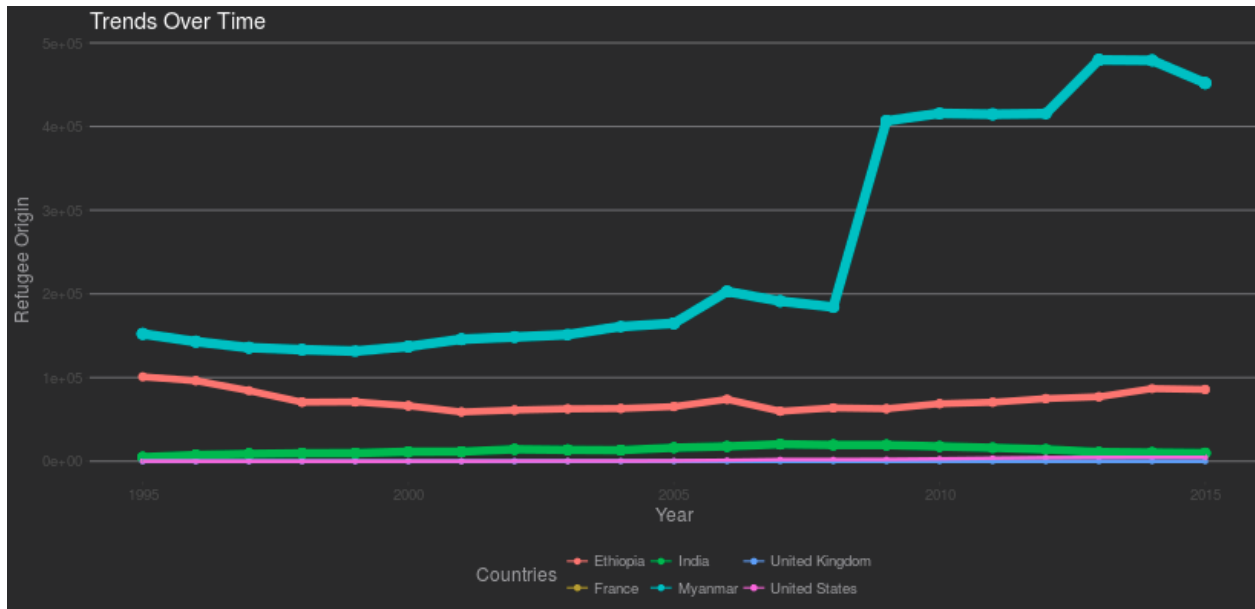
Figure 3: Refugee Origin (% of the country population leaving)
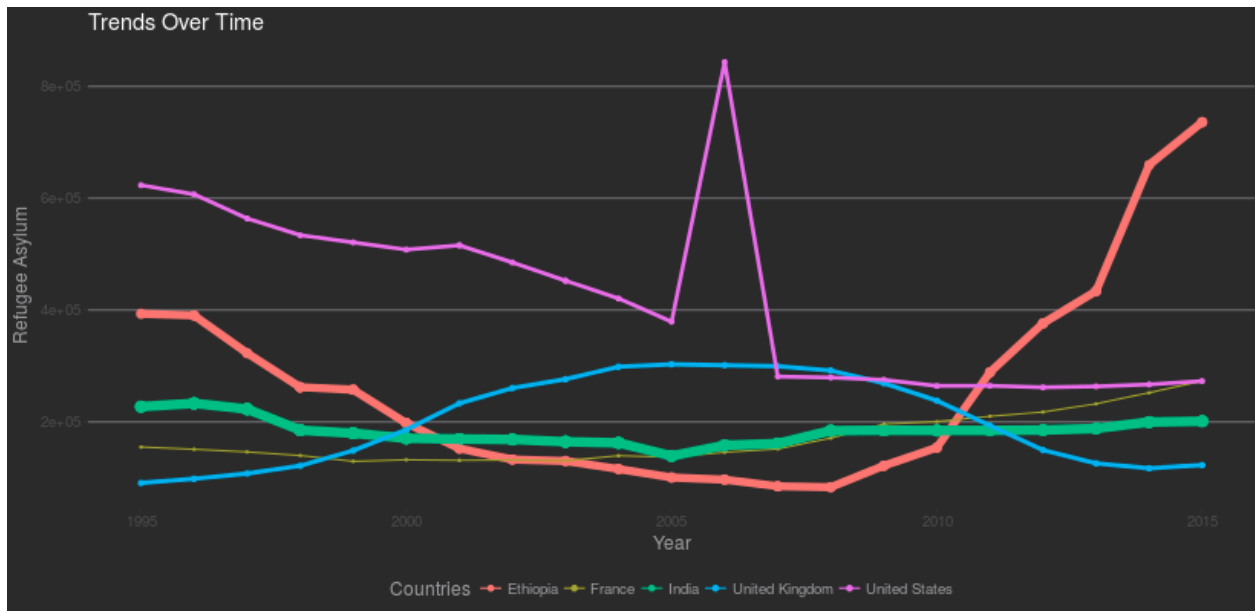


Figure 4: Refugee Asylum (% of the country population arriving)

variables selected included information for every country in every year. Because of this, there is gaps in our data and thus also in our analysis.

The last noteworthy limitation is in the set-up of the conflict set. To get a better analysis of our questions, it would have been better to analyze the sides of conflict as well as the location. With this change, we can see which sides are involved and reach more straightforward conclusions about how westernized countries differ from the less developed countries in terms of conflict.

# Conclusion

By merging three different data sets, we were able to create a shiny app to address a couple of key questions: How do various governments and entities connect in their war history? How do war-torn countries, countries in which the conflict occur, differ in statistics from western countries? We were able to visualize all the historical entities that have been at war and create a network that allows us to see all these connections. By varing the amount of conflicts per entity we saw different variations of the network that we can draw conclusions from. In addition, a majority of conflicts occur in the global south, as opposed to the western countries who are also involved in hundreds of conflicts.

Moving forward, the next step would be to recreate the app using an amended set. In this set, each observation would be one side involved in the conflict, as opposed to the location of the conflict. This will expand our knowledge of the questions stated above.

Overall, this was a very fun project to undertake. We learned a lot along the way while encountering and solving many challenges. We were able to apply some techniques we learned and for that we thank Professor Hardin for giving us the tools to analyze data. In addition, we learned how to make an interactive application from scratch to create a successful, publishable app. We look forward to continuing to expand our knowledge of data science and to utilizing it to answer thought-provoking questions.

# References

Dupuy, Kendra, Scott Gates, Håvard Mokleiv Nygård, Ida Rudolfsen, Siri Aas Rustad, Håvard Strand, and Henrik Urdal. 2015. "Trends in Armed Conflict, 1946–2016." Peach Research Institute OSLO. https://www.prio.org/utility/DownloadFile.ashx?id=1373&type=publicationfile.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39 (5).

Psaki, Stephanie R, and Jessica C Seidman, Mark Miller, Michael Gottlieb, Zulfiqar A Bhutta, Tahmeed Ahmed, AM Shamsir Ahmed, et al. 2014. "Measuring Socioeconomic Status in Multicountry Studies: Results from the Eight-Country MAL-ED Study." *Population Health Metrics* 12 (1). Springer Nature. doi:10.1186/1478-7954-12-8.