

# Information Visualizations: Visualizing Wars

**Max Dekoninck**

KU Leuven

max.dekonick@student.kuleuven.be

**Vital D'haveloose**

KU Leuven

vital.dhaveloose@student.kuleuven.be

**Luca Van Deun**

KU Leuven

luca.vandeun@student.kuleuven.be

## ABSTRACT

In this paper, students of the Information Visualization course at the KU Leuven describe the goals of the visualization they created and the most critical design decisions they had to face along the way. The used dataset will be discussed and related works will be analyzed. The visualization will be explained in detail and finally some important takeaways will be examined.

### Author Keywords

Information visualization; course; students; design; dataset.

### ACM Classification Keywords

H.5.0. Information interfaces and presentation (e.g., HCI): General.

## INTRODUCTION

The Information Visualization course at the KU Leuven aims to give students insights into designing their own visualizations. After learning the good and bad practices in the domain of information visualization, students are asked to first design and then create their own visualization in groups of three. This visualization has to be interactive and has to add value to the data. Data-Driven Documents (D3) is the Javascript library that will be used to create the visualization.

The topic chosen by this group is that of 'wars', which is a very broad topic. After a few brainstorm sessions and looking at the available data, the topic is narrowed down to 'interstate wars in the last 3 centuries'. Everyone knows about the World Wars, but there are many conflicts that affected millions of lives that people don't know about. By looking at the past we can better understand the present.

## GOALS AND TARGET AUDIENCE

When creating a visualization it is important to gear it towards the target audience in order to optimally accomplish the visualization's goals.

### Goals

It is impossible to convey the horrors of a war into a simple visualization, but by comparing different conflicts with each other it can help us put them into perspective. The goal of our visualization is to make it easy and intuitive to compare different wars, not only based on the number of casualties, but also their intensity. This is a metric that is further explained in the dataset subsection. Users will also be able to place the different conflicts in space and time, thanks to a timeline and a world map.

### Target Audience

The visualization should appeal to anyone who is interested in history and wants to know more about (the evolution of) wars throughout the last few centuries. Therefore the visualization should not include any history-specific jargon.

### DATASET

In order to create the visualization, the following data is required:

- Name of the conflict
- Start date
- End date
- Involved countries
- Location
- Number of affected people
- Number of casualties

Originally it was unclear if the focus should solely be on wars or if individual battles should also be included. Mixing both would be like comparing apples and oranges. Given the short duration of a battle, the intensity metric would change drastically for a battle and a war, so a clear distinction is required. After some research it became clear that focusing on individual battles would cause problems in terms of data, since they often aren't well documented. It's also difficult to determine when armed combat becomes a battle, since unfortunately that's a daily occurrence in most wars.

The Correlates of War website [1] offered spreadsheets with the conflict names, dates, involved countries and number of casualties. The location can be derived from the involved countries. Only the interstate conflicts were chosen to limit the amount of entry points, which is already close to 100. The wars range from the 19th century up until the 21st. There were a few obvious typographical errors in the dates of the conflicts, like end dates smaller than start dates, that were corrected manually.

The number of affected people is also necessary to calculate our intensity metric. A different dataset is needed in order to obtain this data. The Gapminder website [2] has a very extensive spreadsheet that contains the population for every country going back to the year 1700. There are several missing data points however, which will be linearly interpolated with the previous and next available data points if needed.

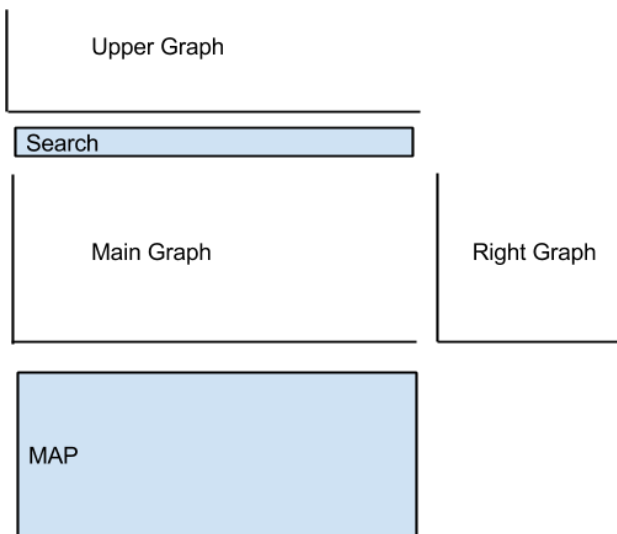
### Formula

The intensity of a war is measured by calculating the average daily odds of dying for someone involved in the war. A shorter war with a lot of casualties will have higher daily odds of death than a longer war where more people were involved with the same number of casualties. In order to calculate this percentage Formula 1 is used. The duration of the war is expressed in number of days and the number of people involved is derived from the populations of the involved countries at the start of the war.

$$\text{averageDailyOddsOfDeath} = \frac{\text{nbVictims}}{\text{nbInvolved} \cdot \text{duration}} \cdot 1$$

**Formula 1.** The ‘average daily odds of death’ denote the chance that a person involved in the war is likely to die on any given day.

There are some considerations with this newly created metric however. The change in population from the start of the war until the end is not taken into account. If for example the population of a country involved in a war would grow, the average daily odds of death would slightly decrease, since the number of people involved increases. On the flipside the people that die during the war are no longer part of the population at the end of it. Therefore only looking at the population at the end of the war, or taking an average between the start and the end would also be suboptimal, since the victims should obviously be part of the number of people involved.



**Figure 1.** An overview of the visualization, which is divided into multiple graphs (placeholder).

### VISUALIZATION OVERVIEW

The visualization is a combination of multiple graphs, as shown in Figure 1. The main graph shows a timeline on the x-axis and the average daily odds of death on the y-axis using a logarithmic scale. Wars are indicated on this graph as horizontal lines, starting at their start date and stopping at their end date. The upper graph also shows a timeline on the x-axis and shows the number of casualties on the y-axis. The graph on the right shows the average daily odds of death on the y-axis and the number of casualties on the x-axis. A world map on the bottom shows the geographical location of the different wars. Finally there’s also a search bar that allows users to search for a specific war. More details can be found in the Visualization and Interaction subsection.

### RELATED WORKS

A similar visualization is ‘Poppy Field’ [3] as seen in Figure 2. It shows the wars from the 20th century onwards represented as poppies (the symbol of the First World War). The x-axis is also a timeline, while the y-axis shows the duration of the different wars with a logarithmic scale. The surface area of the flowers represent the number of casualties. The user can filter by continent, choose ranges for the number of casualties and choose a start and end year. Clicking on a flower shows some information about the war, including the duration, the number of fatalities, the continents where it took place, the participants and a link to the data source and some additional notes.

The purpose of this visualization is literally explained: “The First World War was initially dubbed the war to end all wars. Sadly, the world has witnessed more than 200 conflicts in the 100 years since. Poppy Field reflects on the human cost of war from the beginning of the 20th Century.”

One critique that can be given on this visualization is that it uses a surface area to express the number of casualties in each war. Humans are not very accurate when it comes to comparing surface areas [4]. It’s difficult to know precisely how much larger the number of casualties are for certain wars compared to others.

### VISUALIZATION AND INTERACTION

The visualization tries to make it easy and intuitive for users to find and compare different wars. In order to make searching up the wars more convenient, there is a search bar that auto-completes the user’s input. Once a war is entered the visualization will highlight the selected war on the timeline and automatically zoom in on it. The highlight remains if a user enters a second war in the search bar, in order to be able to compare the two entered conflicts.



**Figure 2: The Poppy Field visualization.**

The main graph determines the behavior of the other two graphs and the world map. The user can control the timeline of the main graph by clicking and dragging it left or right and by zooming in or out by scrolling. This functionality is made clear by displaying a short animation when starting up the visualization. The cursor also changes when hovering over the main graph, to further clarify that the graph is interactive. Figure 3 shows the cursor that appears.



**Figure 3. The cursor changes when hovering over the main graph, indicating that it can be manipulated.**

The supporting graphs adjust according to what wars are visible in the main graph. If the user zooms in on the main graph, the amount of information gets narrowed down. The information shown in the upper graph follows the timeline of the main graph, but shows the number of casualties instead of the average daily odds of death. This means the horizontal positioning for a specific war will be the same in both graphs. The graph on the right shows the average daily odds of death and the number of casualties, so the vertical positioning of a war is the same in the main graph and the one on the right. The world map also only shows the conflicts that are visible on the timeline of the main graph.

### Important design decisions

One of the first decisions that had to be made was choosing on what aspect of wars we would focus. Given the broad nature of the topic, narrowing it down helps make more informed decisions. As alluded to in the introduction we chose to forego individual battles and focus on full scale wars. Battles aren't always very distinct and finding data like the number of casualties for a specific battle is rather an exception. Interstate conflicts were chosen over intrastate conflicts, since interstate conflicts include the recognizable wars that people can already relate to. It also helps limit the number of entry points that were already close to 100 with interstate wars alone. Adding too many wars to the timeline causes too much overlap

Originally the number of casualties was shown on the main graph, as depicted in Figure 4. The surface area of a circle indicates the number of casualties. The main problem with this approach is the same that is brought up in the critique of the Poppy Field visualization in the Related Works subsection. It is difficult to compare surface areas and determine the exact difference. Overlapping circles also make it more difficult to distinguish what circle belongs to which war. This first attempt tried to show too much information in too little space. Splitting up the main graph in multiple graphs helps the clarity of the visualization.

## REFERENCES

1. The Correlates of War Project.  
<http://www.correlatesofwar.org/data-sets/COW-war/>.
2. Gapminder Data.  
<http://www.gapminder.org/data/>.

3. Poppy Field Visualization.  
<http://www.poppyfield.org/>.
4. Mackinlay, Jock. "Automating the design of graphical presentations of relational information." *Acm Transactions On Graphics (Tog)* 5.2 (1986): 110-141.

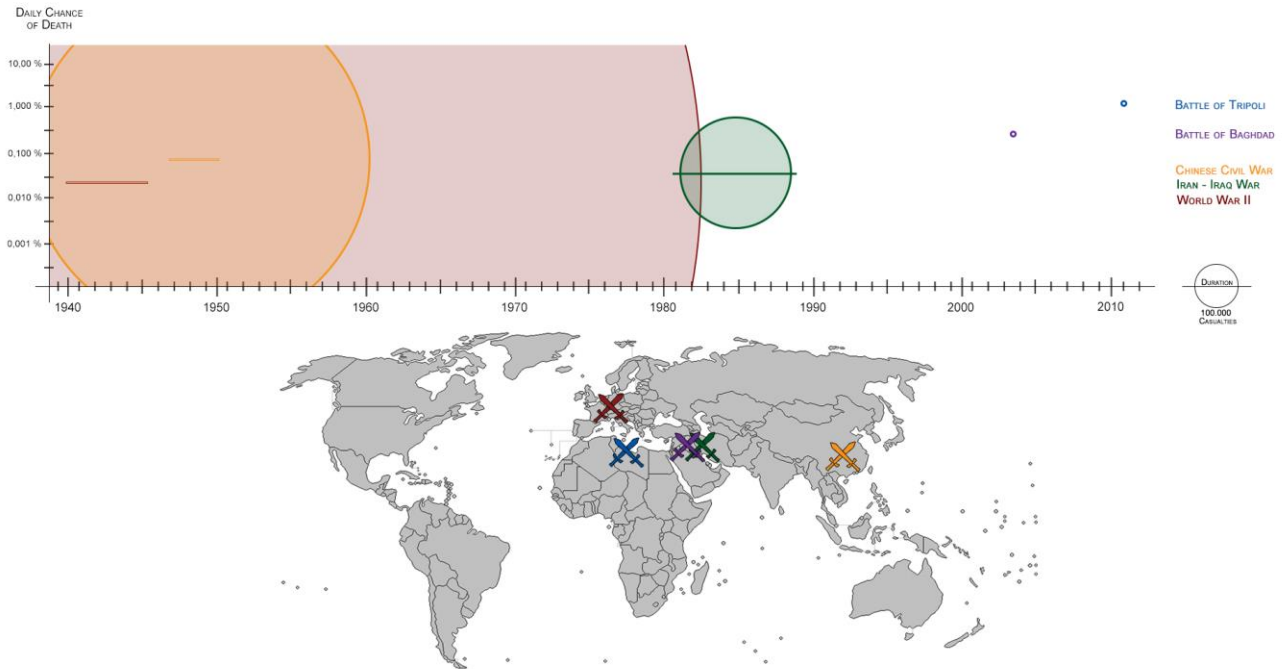


Figure 4: The initial design of the visualization used the surface area of circles to indicate the number of casualties.