Where to find our code: https://github.com/raahilsha/cs171-pr-alexander

Process Book

Alexander: The *What If?* of War

Motivation

~~Wars haunted Humanity at least since the emergence of the civilization. They constitute a somewhat puzzling phenomenon: obviously, trade and peace is much better than the risk of death and destruction, and self-interested agents shouldn't choose war. As we now know, even chimpanzees have wars [1], and there seems to be something fundamentally wrong at least with some primates which causes them to be aggressive towards different groups of their own species.~~

Wars are not something people think about on a daily basis. We are confident that transforming the data about armed conflicts into knowledge and communicating this knowledge is a great application of the skills we learned in CS171. Our purpose is 2-fold: we want to communicate the price the Humanity pays for wars, and encourage people to look for patterns in the complicated relationships between wars, economic power, political alliances and history.

Skills - data analysis, statistical analysis, gamification, visualization

Inspiration:

This visualization of US drone attacks in Pakistan appeals to us very much. We think it is a great example of the important message, which is communicated in a clear and beautiful way.

The visualization of US population density shown in the "Maps" lecture inspired some of the ideas for advanced features.

This visualization of arms transfer between countries over time, while not perfect, also seemed interesting.

Project Evolution

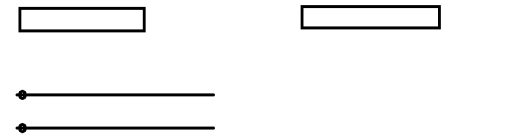
Our initial ideas was to construct a data-based game, where players could select 2 or more countries at a particular year and see the progression and the results of a hypothetical armed conflict between these countries and their allies. We put a lot of emphasis on building a good

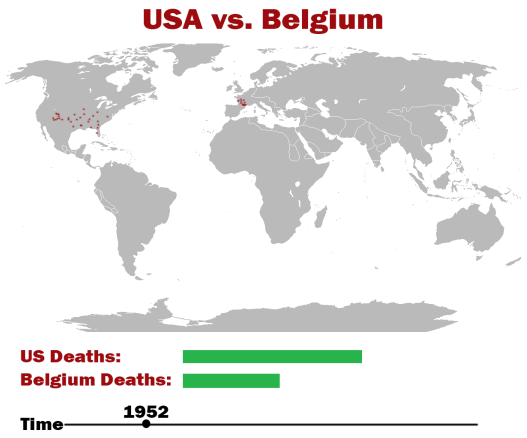
statistical model to predict the outcomes of wars based on the historical data about military power, economics and political alliances.

However, while we took a closer look at the data and tried testing simplest models to predict war outcomes, this task turned out to be very complicated: there seem to be a lot of variables at play, and we don't have data about most of them. Our optimistic assumptions about the relationships between military power and war victories turned out to be false, and the composition and changes in alliances (treaties, non-aggression pacts, defense alliances, etc.) turned out to be complicated and require a lot of research and wrangling to be ready for the use in a prediction model.

Our Teaching Fellow suggested that we limit the statistical part of the project and confine ourselves to a few features of the data that can be cleaned and used relatively easily, and instead focus our attention on the design of the most effective ways to communicate our message.

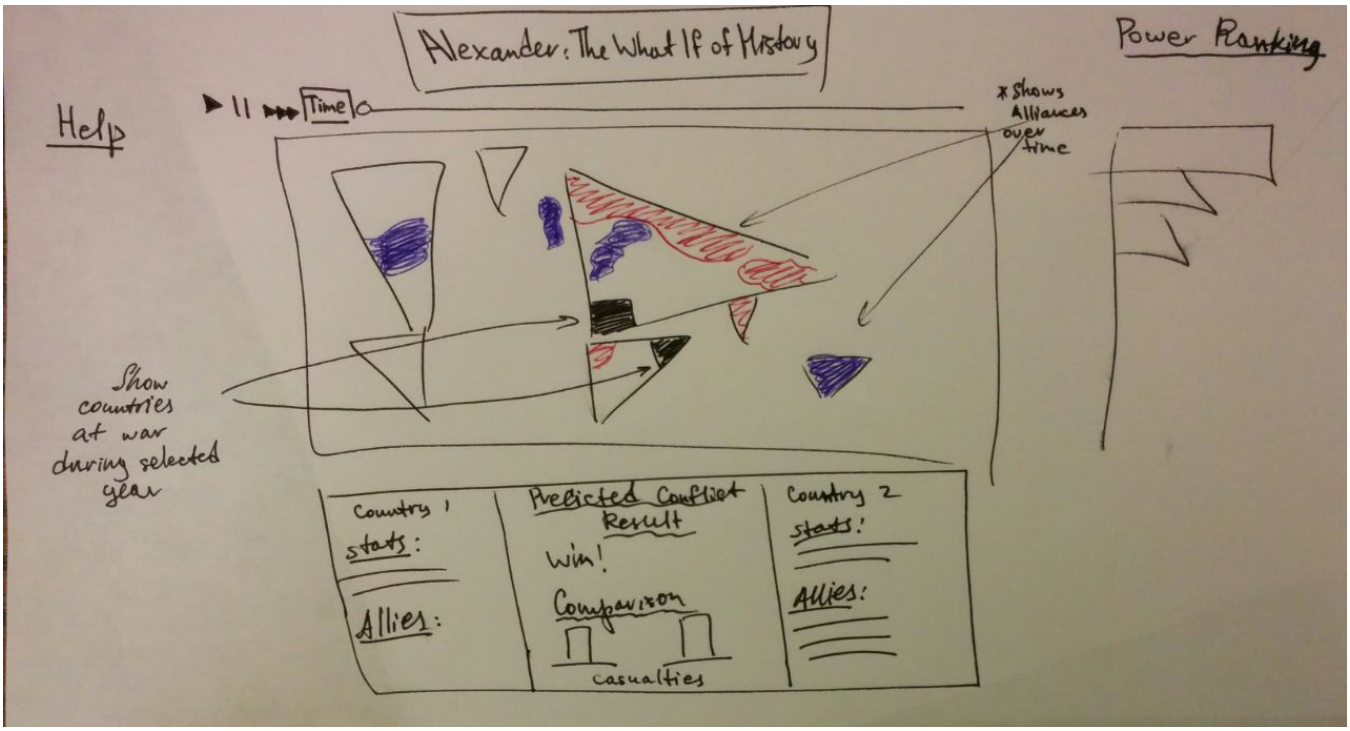
Because of this, we conducted the review of the interface of our game and realized that it was not very effective: we initially planned that the game would consist of 2 screens: one for the war setup, and the second for the demonstration of the course of the war. Such a setup would not be fully utilizing the amazing opportunities for interaction provided by web-based visualization.





Because of this, we decided to make simulation an optional additional simulation feature, and instead focus on visualizing complicated relationships between alliances, power metrics and war outcomes throughout history.

Through many sketches and discussions we came to the following basic multi-view layout, which we are currently implementing:



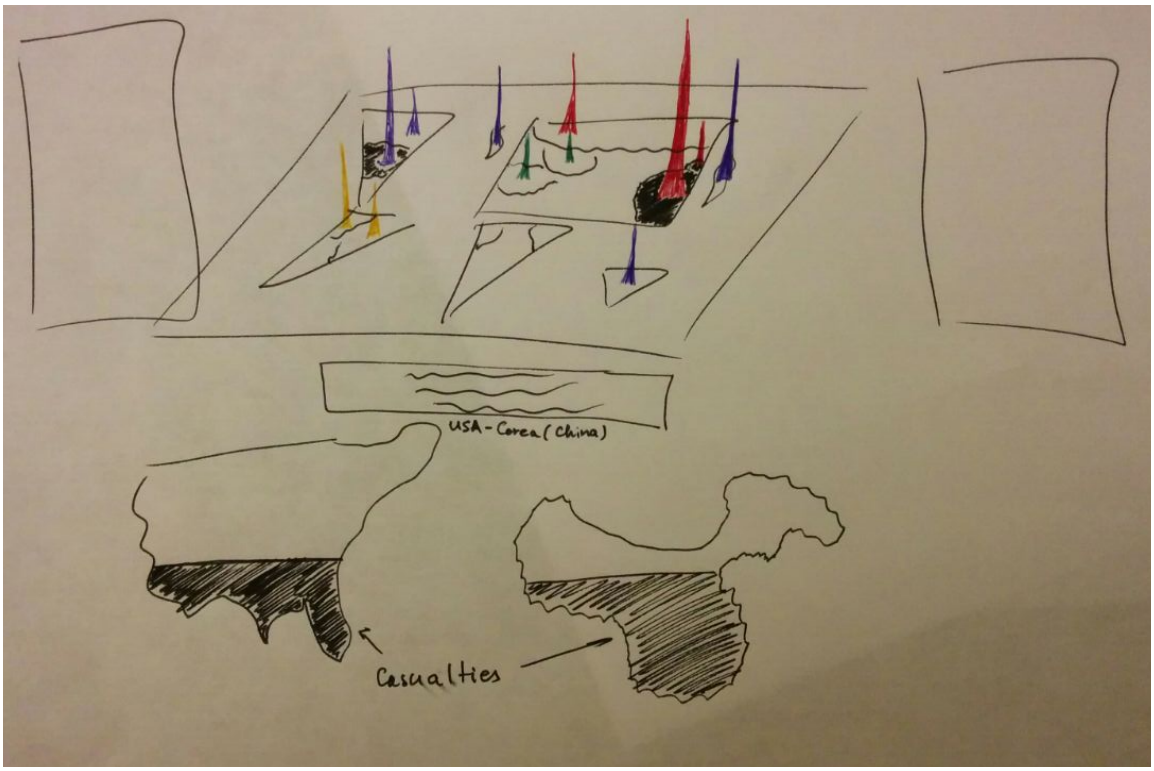
The must-have features now include:

- An interactive map, which allows the user to select countries
- A bottom panel which represents historical statistics about countries
- A bar chart of top-10 countries by National Power Index
- Highlighting the countries in war at the selected year

Optional features include:

- Incorporating the data about military alliances and international trade and finding a way to visualize it over time
- A special view showing the simulated characteristics of the world or of the selected countries if some or all wars did not take place
- A pseudo-3D version of the power ranking bar-graph that allows the display of the power data on the map naturally. This may also include some version of more imaginative area-based graphs to show casualties or other war related

statistics(inspired by this):



Questions:

a. Our initial questions included:
    i. What's the history of wars? - tell the story
    ii. What's the price of wars? - show clearly
    iii. how are power and alliances related to war outcomes? - give opportunity to see patterns
b. Novel questions that we discovered:
    i. What if there were no wars?

Data

All the data needed to implement basic features is collected and is now being processed:

1. Power index - includes military expenditure, military personnel, energy consumption, iron and steel production, urban population, and total population (since 1812)[1]
2. Formal Alliances[2]

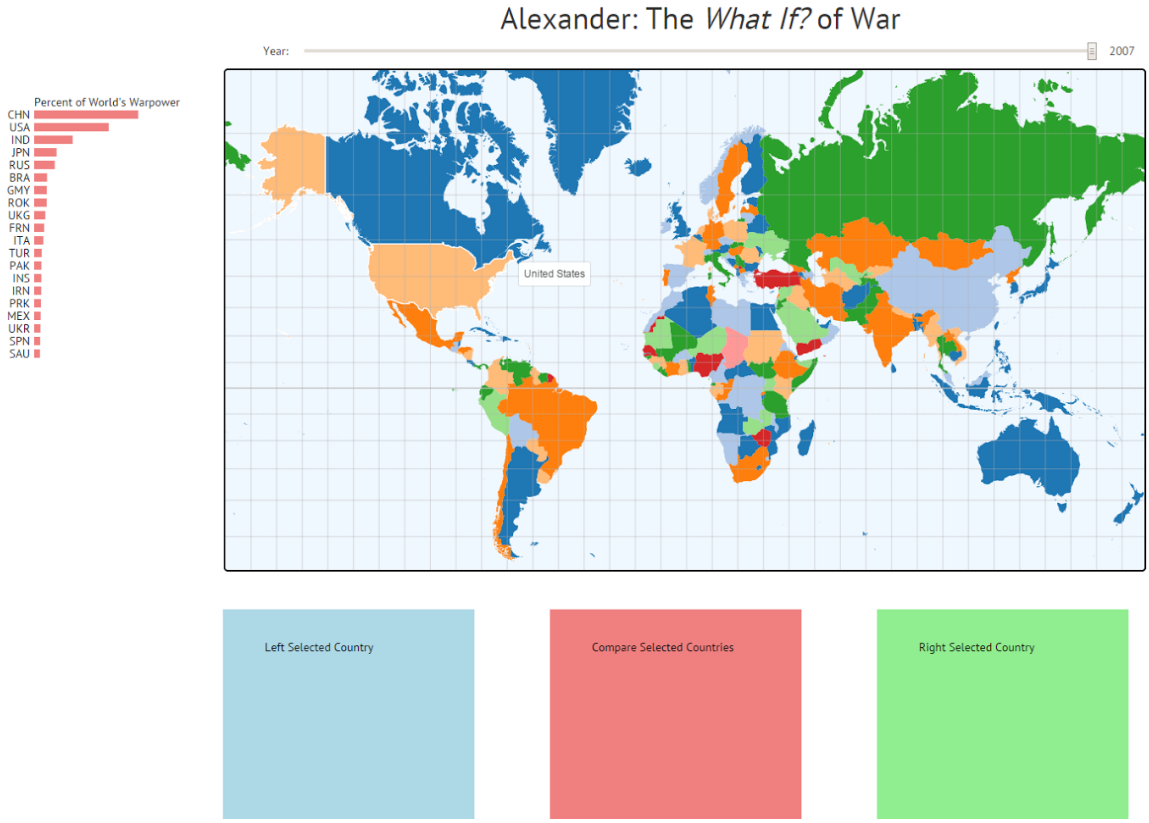
[1] *Singer, J. David, Stuart Bremer, and John Stuckey. (1972). "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." in Bruce Russett (ed) Peace, War, and Numbers, Beverly Hills: Sage, 19-48.*

[2] *Gibler, Douglas M. 2009. International military alliances, 1648-2008. CQ Press.*

3. Inter-State War Data: including the battle-related combatant fatalities suffered by the state, year in which the war took place and a lot of other features (see ipython data processing notebook for details)[3]

Implementation: Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.

Progress at First Milestone



- Map in the center
  - Allows user to hover over country to display tooltip
  - Will be used in future to control the three visualizations at the bottom
  - Optional Feature: Highlight countries currently involved in war on the world map
- Time slider at the top
  - Allows the user to select a time from 1816 to 2007
  - Updates all of the other visualization views based on the time
- Bar chart at right
  - Displays top 20 countries in terms of % of world's warpower
    - Composite Index of National Capability

[3] *Sarkees, Meredith Reid and Frank Wayman (2010). Resort to War: 1816 - 2007. Washington DC: CQ Press.*

    - Will be updated in future to show the country's color, which will make it more clear when one country grows stronger than another
- Blank SVGs for the three views at bottom
    - The left view will allow the user to click on a country to permanently select it
    - The rightmost view will update itself when the user hovers over a country
    - The center view will display statistics that reflect the "clicked" country and the "hovered" country
- Data-wise, we finalized collecting all the data we will need for our project
    - Most of the data is processed in some form
    - Country index data is completely processed
- You can find our data analysis code in /data_processing/processing_code.html on our project Github

Evaluation:

We went through a lot of ideas for visualization, and learned how hard it is to come up with anything worth mentioning:

Our original idea to perform a war simulation didn't pan out in the end, as we were unable to find a correlation between a country's warpower and who actually won the war using simple statistical analysis. As we did not want to create a more complex model to determine the victor of a war (as this would take too much time), we instead decided to move on to a different visualization based on the same idea.

Notes:

TF feedback summary:

- Limit number of variables in model and check correlations (in process)
- Think a lot more about visualization and how to represent the war in time - more sketches, what visual variables, thorough justifications for design decisions (in process)

Design Studio Feedback summary:

- Make sure to consider the possibility of 2 allies fighting
- Color different sides of the war
- Additional options:
    - Complicate simulation more
    - Zoomable map
    - Allow more user interaction with simulation progress

Tasks for now:

1. Clean and wrangle the data, try to find correlations between features and outcomes (Grigory)
2. More sketches and thinking about vis design (several views?)
3. Tomorrow: ~working vis and good-looking process book with lots of text and pictures that make sense

Possible work division:

- Visualization design and thinking about statistical model - both of us
- Data cleaning, reshaping, stat model building - mainly Grigory in python
- Implementing visualization - more Raahil?