COSC3000: Data visualisation report

John Owen
Student ID: 43591617

1. Project aim

The project uses data visualisation techniques to analyse two datasets, the UCDP/PRIO Armed Conflict dataset, and the International Crisis Behavior Project dataset. The aim is to gain insights into how wars change over time, the main players involved, and the associations of nuclear weapon capability in conflicts.

2. Motivation

The ever growing need to stop war before it starts and minimise loss of life once it has does not need to be stated. Understanding war is key not only to governments and institutions with the power to control it but also to the citizens who must elect the government officials and political parties that deal with its complexities.

War is usually understood through long texts that describe in detail all recorded events. While this can give a thorough understanding of a particular conflict, exploring war through statistics and graphs has many advantages. Namely,

- Wars can be understood at the macro level, looking at trends across many wars at once.
- Graphics can quickly give a rich summary difficult to describe in words.
- Wars can be distilled to one or two factors that the reader is interested in, allowing a quick comparison of these factors across many wars.

3. Datasets used

A combination of three publicly available data sources were used in the project.

3.1. UCDP/PRIO armed conflict dataset

Contains a list of 259 conflicts ranging from 1946 to 2014 [1, 2, 3]. For each conflict it lists:

- Start and end date.
- Location.
- All states and non-state groups (henceforth collectively referred to as actors) on both sides of the conflict.
- Type of conflict (civil war, international war, other possibilities)
- Rough estimate of death count.

3.2. Gleditch and Ward list of states

Entries in the UCDP/PRIO dataset use numeric IDs to identify countries. This dataset maps those IDs to the full names and three letter country codes of states. For example, "United States of America" has an ID of 2 and a country code of "USA". [4]

3.3. International Crisis Behavior Project datasets

Describes many dimensions of 470 international crises from 1918 to 2014. Interestingly, rather than containing conflicts, it contains military security crises, which the codebook defines as the period following a sudden increase in the presence or threat of violence. This definition will be used in the rest of the report. [5, 6, 7]

The data comes in two datasets which in total have over 160 columns in total, so it is too large to describe here. However, some example information it contains for each crisis:

- What was the cause of the crisis.
- Level of communication between crisis actors.
- Date and location of crisis.
- Crisis management style of crisis actors (e.g, negotiation, economic pressure, violence, etc).
- The role of any states involved in the crisis.
- The role of any global organizations (e.g, branches of the UN).
- Nuclear weapon capability of crisis actors.

4. Method

CSV files and tab separated value (TSV) files were obtained from publicly viewable websites [1, 4, 5]. In the case of the Gleditcsh and Ward list of states, two TSV files were combined into one CSV file for easier data loading. These datasets are located in the `datasets/` directory.

The R language [8] was used primarily in conjunction with its "ggplot2" [9], "network" [10, 11], and "GGally" [12] libraries to generate all graphs. The R code is available in the `code/main.R` script file. The script loads the datasets, converts columns to the correct data type (e.g, "1" is parsed as a number and "1,2,3" is parsed as a vector of numbers), and outputs diagrams. For the network diagrams, the data had to be carefully converted to an adjacency matrix and pruned of boring data (e.g, nodes with no edges).

To build the project, see the `readme.md` file.

5. Results

For all levels of violence, the frequency of crises peaked in the late 70's and drastically dropped afterwards (Fig 1). There are many possible reasons for this, but the most obvious candidate is the decline of the Cold War and the collapse of the Soviet Union.

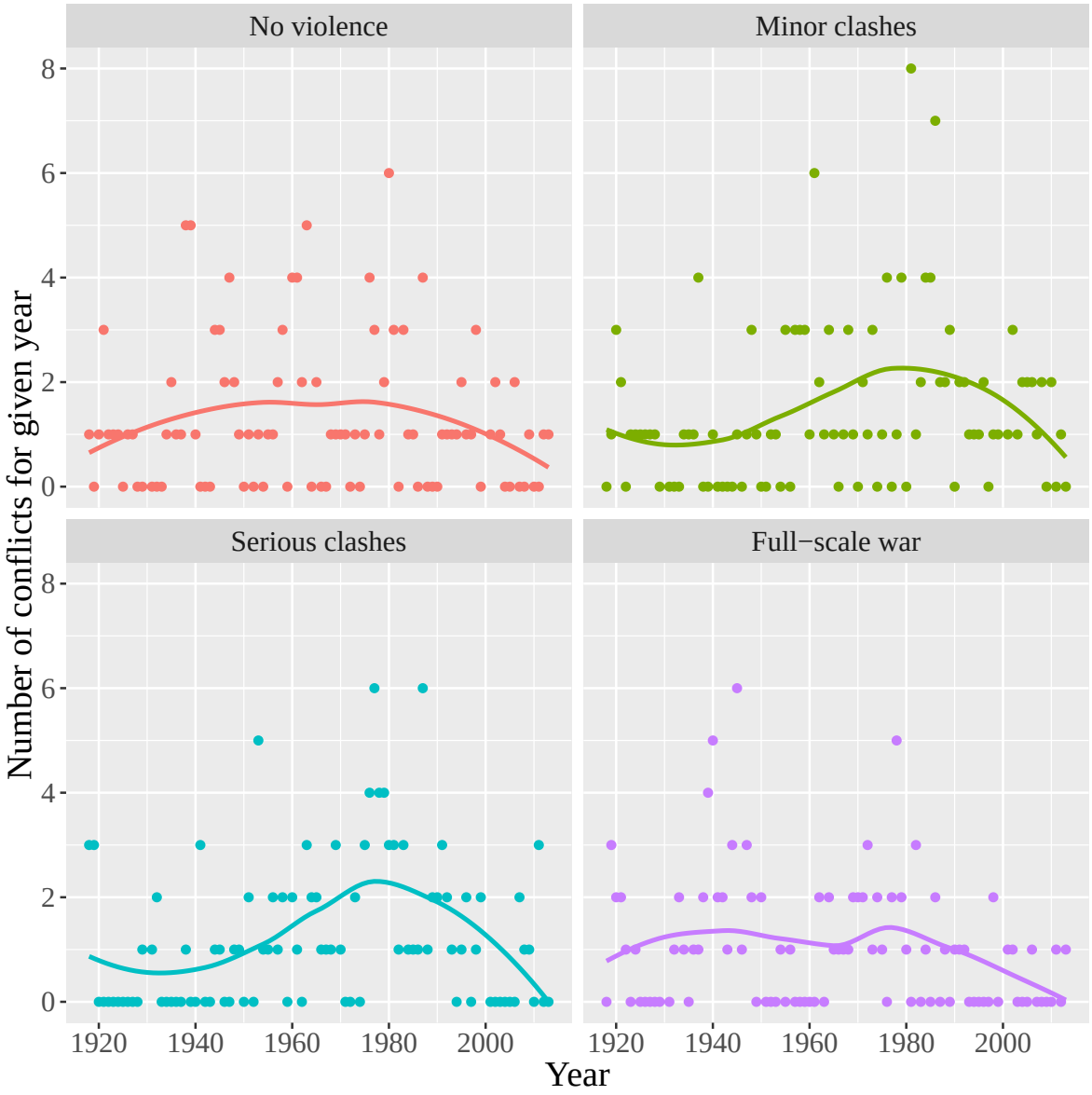

Figure 1: Violence in crises over time. Each dot represents the frequency of crises for a particular year. A Loess curve is fitted to the data to show the overall trend.

An increase in the nuclear capabilities (see Appendix A for precise definitions of the levels of nuclear capabilities) of actors in a crisis is associated with a large increase in the frequency of non-violent crises vs frequency of violent crises (Fig 2). However, there is a negligible difference between the frequency of non-violent crises where actors have the third level of nuclear capability (possession of nuclear weapons) vs the fourth level (second strike capability).

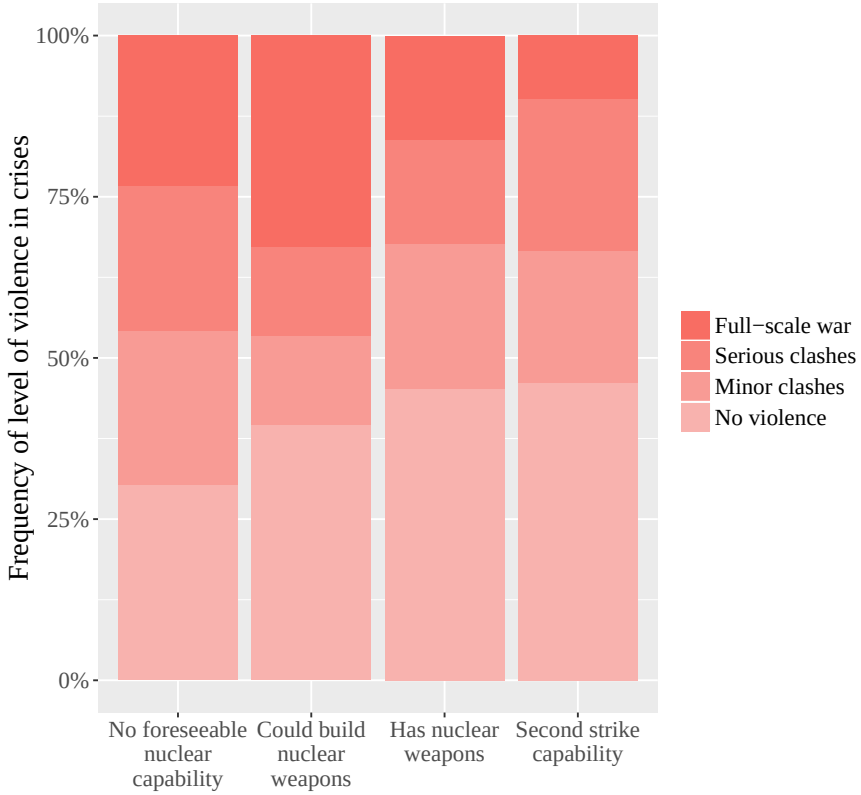

Figure 2: Country nuclear capability compared with crisis violence. See Appendix A for precise definitions of the nuclear capability levels.

The possession of nuclear weapons can decrease violent conflicts because states will avoid direct conflict to avoid triggering a mutually destructive nuclear war, as seen in the Cold War [13]. However, a hypothesis shared by some is that beyond a certain level of nuclear capability, further bolstering a nuclear arsenal doesn't lead to a decrease in conflict [14]. While this graph shouldn't be taken as evidence of that claim, this is the result that would be expected if it were true.

Most countries that have fought wars, do so with only a small number of allies, generally one or two. A small number of countries (Malia, USA, Afghanistan, Iraq) have fought in wars with many allies. (Fig 3)

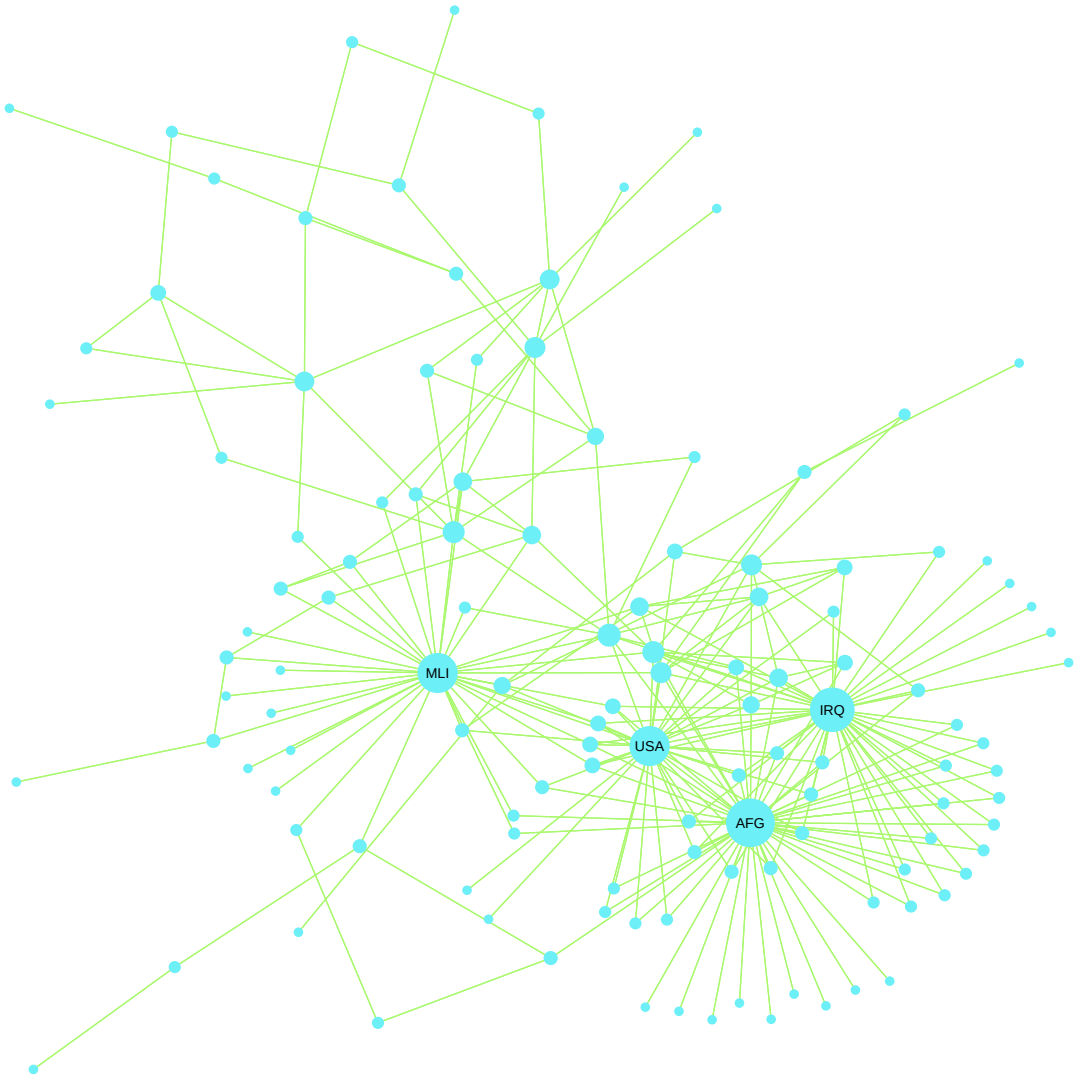

Figure 3: A network plot of countries that have fought as allies. Each node is a country and every edge represents an instance where two countries have fought on the same side. Countries that have not had any allies in conflicts they have fought (i.e, nodes with no edges) are not displayed. Four countries which have fought with a far greater number allies have been labeled. These countries are Mali (MLI), The United States of America (USA), Afghanistan (AFG), and Iraq (IRQ). The size of each node is determined by the number of edges it has.

Most countries that have fought wars, have done so with a few opponents. A few countries (Iraq, North Korea, Russia, China) have directly fought wars with many opponents (Fig 4).

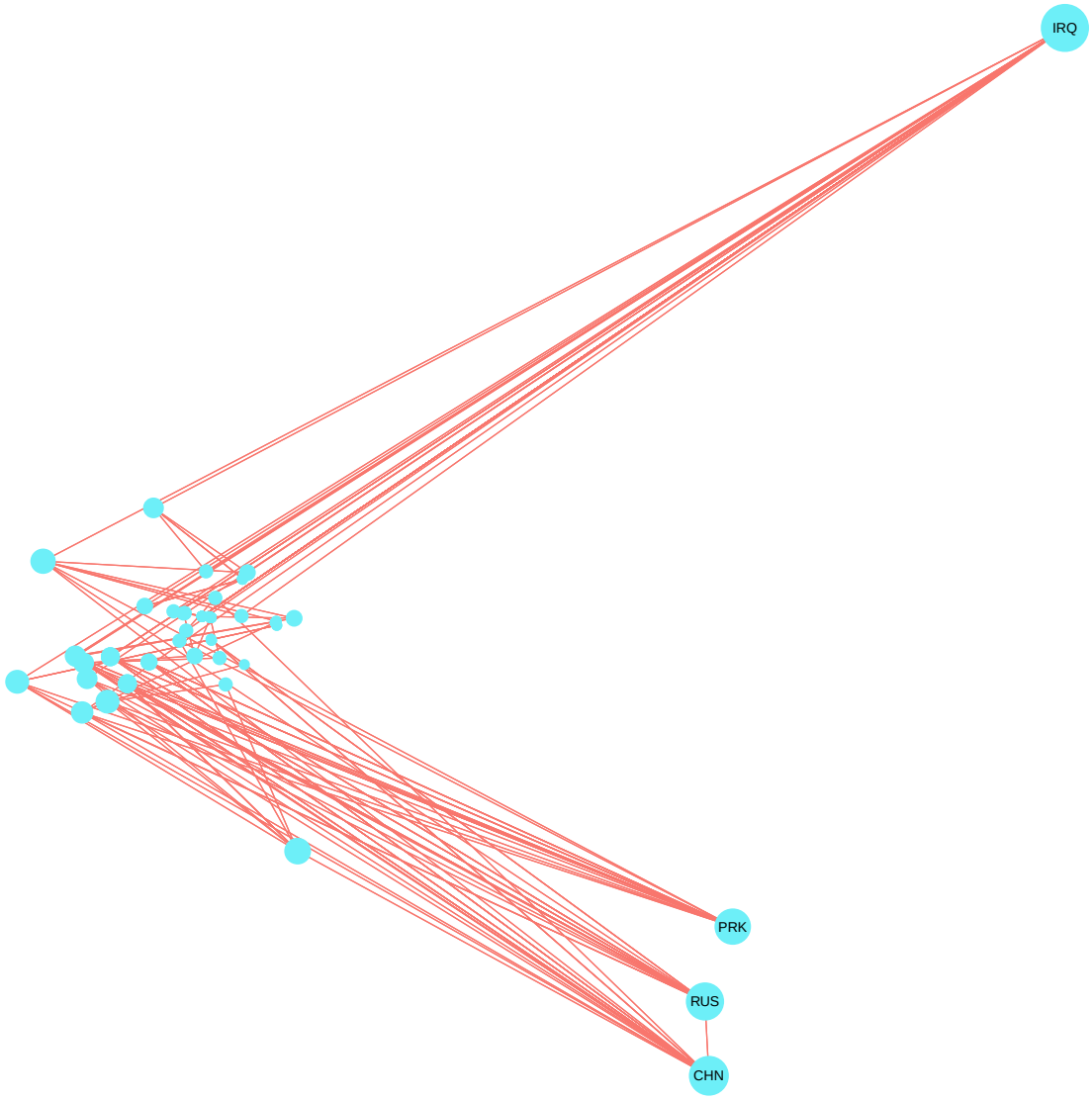

Figure 4: A network plot of countries that have fought as enemies. Each node is a country and every edge represents an instance where two countries have fought on opposing sides. Countries that have not been involved in any conflicts (i.e, nodes with no edges) are not displayed. Four countries which have fought with a far greater number of enemies have been labeled. These countries are Iraq (IRQ), North Korea (PRK), Russia (RUS), and China (CHN). The size of each node is determined by the number of edges it has.

Interestingly, despite the USA's military power and its involvement in many theatres of conflict, it doesn't appear on the list of top countries. This is because after World War II, due to fears that a direct conflict with the Soviet Union could lead to a nuclear apocalypse, the USA engaged less often in combat and more often participated indirectly in proxy wars; often choosing to equip and train governments and rebel groups instead. [13].

6. Appendix

A. Nuclear capability levels

Taken directly from the ICB dataset codebook [6].

No (foreseeable) nuclear capability The actor did not possess a nuclear capability with any operational military significance when the crisis began; moreover, the international consensus at the time was that it could not develop or acquire such capability within five years.

Foreseeable nuclear capability The actor could develop or acquire operational nuclear military capability within five years of the beginning of the crisis.

Possession of nuclear capability the actor had nuclear military capability (weapons) and delivery means but no second-strike capability.

Developed nuclear capability, with second strike capability Superpower or great power with ability to absorb a first strike and retaliate.

References

[1] Nils Petter Gleditsch et al. *UCDP/PRIO Armed Conflict Dataset*. http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset. June 2015.

[2] Nils Petter Gleditsch et al. *UCDP/PRIO Armed Conflict Dataset Codebook*. Version 4-2015. Uppsala Conflict Data Program and Centre for the Study of Civil Wars, International Peace Research Institute, Oslo. June 2015.

[3] Nils Petter Gleditsch et al. "Armed Conflict 1946-2001: A New Dataset". In: 39.5 (2002).

[4] Gleditsch, Kristian S., and Michael D. Ward. *List of independent states dataset*. http://privatewww.essex.ac.uk/~ksg/statelist.html. Mar. 2013.

[5] Michael Brecher et al. *ICB Project datasets*. http://sites.duke.edu/icbdata/data-collections/. 2016.

[6] Michael Brecher et al. *International Crisis Behavior Data Codebook*. Version 11. 2016.

[7] Michael Brecher and Jonathan Wilkenfeld. "A Study of Crisis". In: *University of Michigan Press* (2000).

[8] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2014. URL: http://www.R-project.org/.

[9] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN: 978-0-387-98140-6. URL: http://ggplot2.org.

[10] Carter T. Butts. *network: Classes for Relational Data*. R package version 1.13.0. The Statnet Project (http://statnet.org). 2015. URL: http://CRAN.R-project.org/package=network.

[11] Carter T. Butts. "network: a Package for Managing Relational Data in R." In: *Journal of Statistical Software* 24.2 (2008). URL: http://www.jstatsoft.org/v24/i02/paper.

[12] Barret Schloerke et al. *GGally: Extension to ggplot2*. R package version 1.0.1. 2016. URL: http://CRAN.R-project.org/package=GGally.

[13] Robert Wilde. "Mutually Assured Destruction". In: *About Education. About.com* (Apr. 2015). URL: http://europeanhistory.about.com/od/glossary/g/glmad.htm.

[14] Henry Reich. *Why You Should Care About Nukes*. Youtube. Mar. 2016. URL: https://www.youtube.com/watch?v=qRnU0bqsyq0.