

Correlates of War Lab

Jacob Freiheit

September 2023

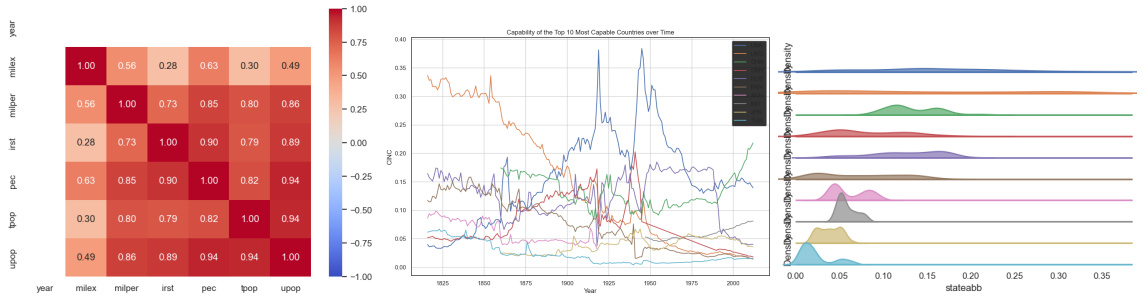
1 Introduction

The main idea behind the project is to keep accurate record of international relations between the countries of the world. This project includes databases on international war contribution, losses of personnel in international and civil conflict and more, all with the intent to have the public work with accurate information in regards to the subject of conflict. This project is spearheaded by a number of directors, each appointed to a 5-year term by an Advisory Board. This makes the project more commercial than academic, though each director is involved in academia in some way.

Each dataset highlights different capabilities and tendencies of the countries of the world. One such dataset is the Militarized Interstate Disputes

dataset, which shows which countries are fighting or threatening to fight each other. Another is the National Material Capabilities dataset, which details the military resources countries have at their disposal. There are many others, including World Religion Data, Formal Alliances, and Territorial Change datasets as well. Updates 4.0 and 5.0 give are merged information with previous versions, and ease of access by allowing more platforms. In 4.0, additional categories were included, which required an update be made for information going all the way back to the 1820s. 5.0 takes this updated dataset and adds a "2017 Update" at the conclusion of each component section. This is to provide information about areas of the dataset that require additional context given the 2017 update.

2 Data



The National Material Capabilities (NMC) dataset summarizes 6 variables. These are: Military Expenditure (milex), Military Personnel (milper), Energy Consumption (pec), Iron and Steel Production (irst), Urban Population (upop), and Total Population (tpop).

Based on what I have seen in the Ridgeline plot on the top 10 countries' capabilities, countries seem to have a distribution set to the left, meaning the majority of the top 10 countries have significantly more datapoints that indicate lower capability relative to the rest of the top 10. This tells us that the most capable countries (the US, the UK, and China) are much more capable than the other countries even within the top 10. (The colors of the rows on the Ridgeline plot correspond to the legend on the line graph) There are some variables that are correlated simply by nature. These variables become clear when we view the heatmap. upop and tpop are very heavily correlated. This might simply be because any country that has a lot of people living in cities has a lot of people. In addition to this, Iron and Steel production seem to be correlated by nature with the energy figure, as the more iron and steel produced, the more energy was used.

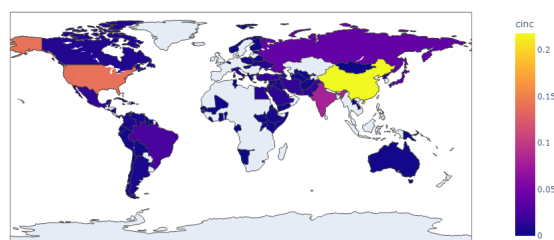
3 Analysis

When calculating the smallest Euclidean and Manhattan distances, I found that the most similar countries were Honduras, Paraguay, Tuvalu, Saint Lucia, Nauru, and Vanuatu. These countries also have the lowest CINC values of the countries of the world.

A possible explanation for why the most similar countries also have the lowest CINC values may be that countries are only similar because their COW numbers are almost zero. Countries that have higher CINC values can often achieve those values in different ways. Maybe one country has more personnel, while another manages to achieve high CINC through steel production. The countries would be very different in terms of euclidean and manhattan distances despite both having high CINC. When a CINC value is very low, you can only achieve this by having very low values for variables upon which CINC relies. It follows that these countries would be very similar to each other.

4 Discussion

CINC Values by Country (Year 2012)



The areas with by far the biggest differences in terms of National Capability are going to be around the few countries with relatively high CINC in 2012. Countries like the US and China are the biggest examples of this, where all the surrounding countries have significantly less CINC than these two.