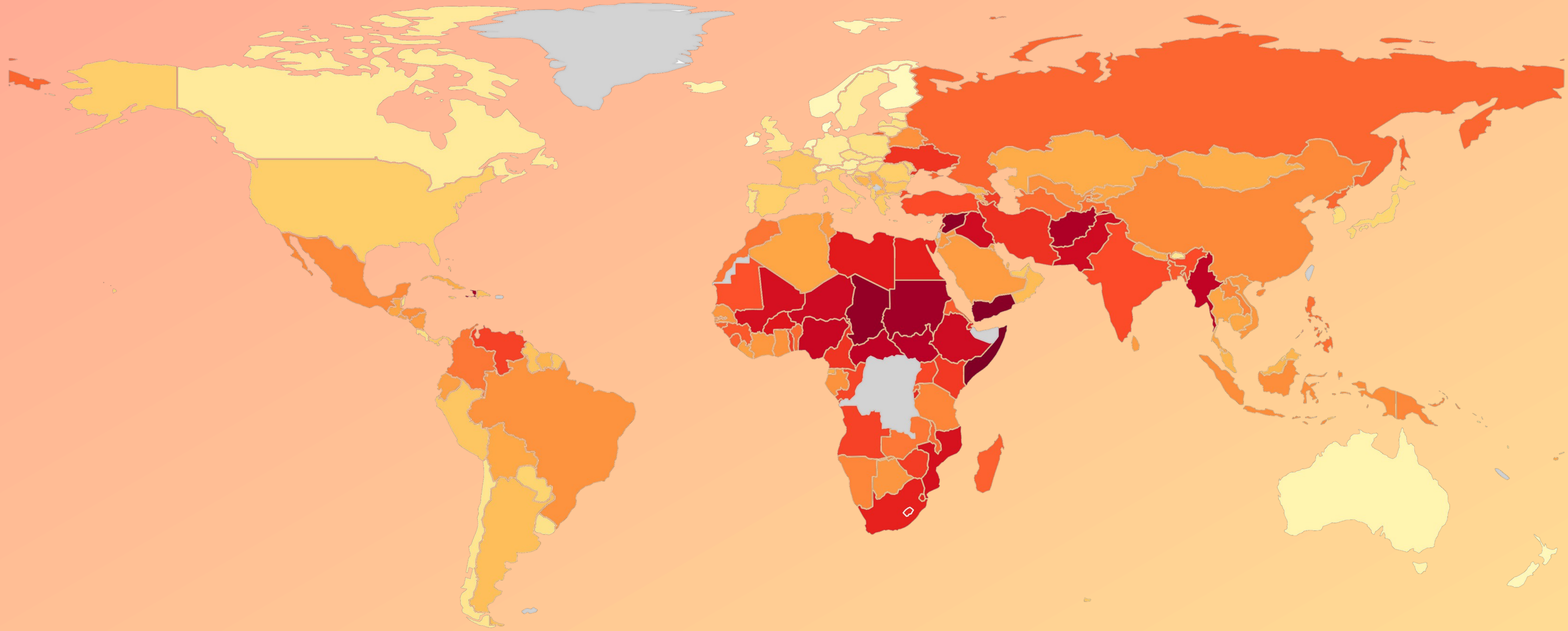


Global Societal Endangerment Index

Arthur Freye

DkIT Data Analysis and Visualisation 2025

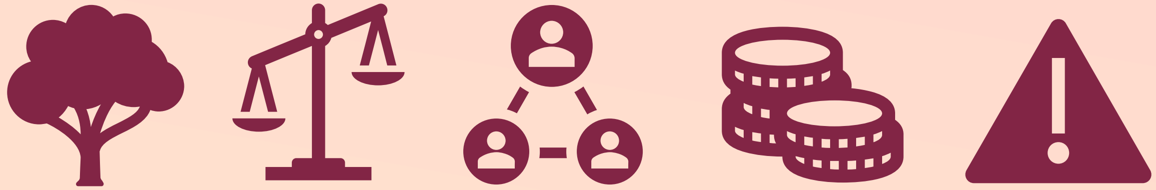
Contents

Executive Summary

GSEI

The Global Societal Endangerment Index (GSEI) is a composite index designed to measure countries' vulnerability to a range of environmental, political, social, economic and other risks and threats. These risks are weighted according to their potential impact on society and the population within each country. The index considers both imminent impacts, such as catastrophes and conflicts, and slower-developing consequences, e.g. increased mortality and instability.

This problem was chosen because the world faces numerous interconnected risks, including climate change, natural disasters, social and political instability, and economic crises, which can have cascading effects, endangering populations globally. A composite index like the GSEI can aid policymakers, businesses, or civil society organisations in better understanding and addressing these challenges by providing a holistic view of a country's vulnerability.

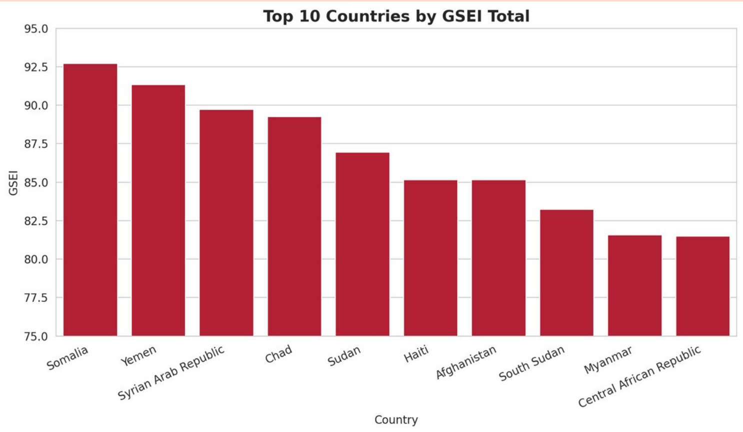


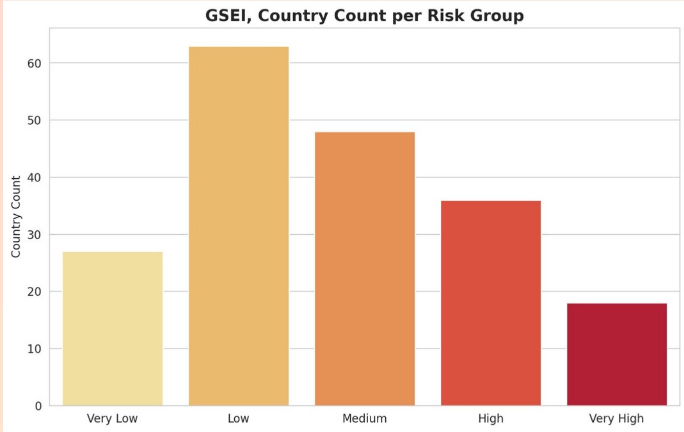
In total, the GSEI comprises 34 indicators, grouped into 5 sub-indicators: Environmental Risks, Political Instability & Governance, Social Vulnerability, Economic Stability & Infrastructure, and Global & Regional Threats. An expert was involved in developing the theoretical framework, notably in determining the weighting of the sub-indicators.

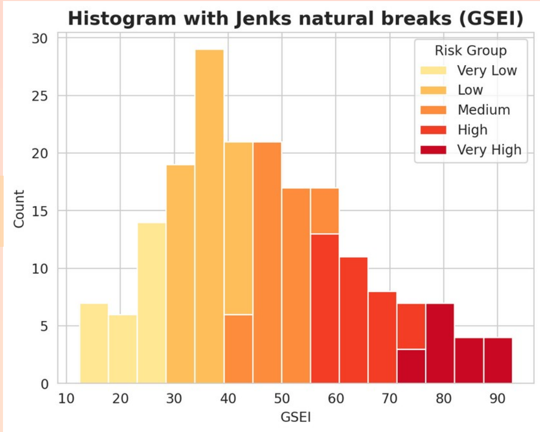
The methodology employed in constructing the GSEI involved using data from reputable sources such as the WHO, World Bank, and UN. To address missing data, clustering was used to impute values in a more nuanced manner. Collinear indicators were dropped or transformed using PCA. To counter the effects of heavily skewed indicators, logarithmic transformation was applied to some variables. Further standardisation, weighting and aggregation finally result in an index that is very comprehensive and allows for informative visualisations. The countries are segmented in 5 risk groups based on their total GSEI score.

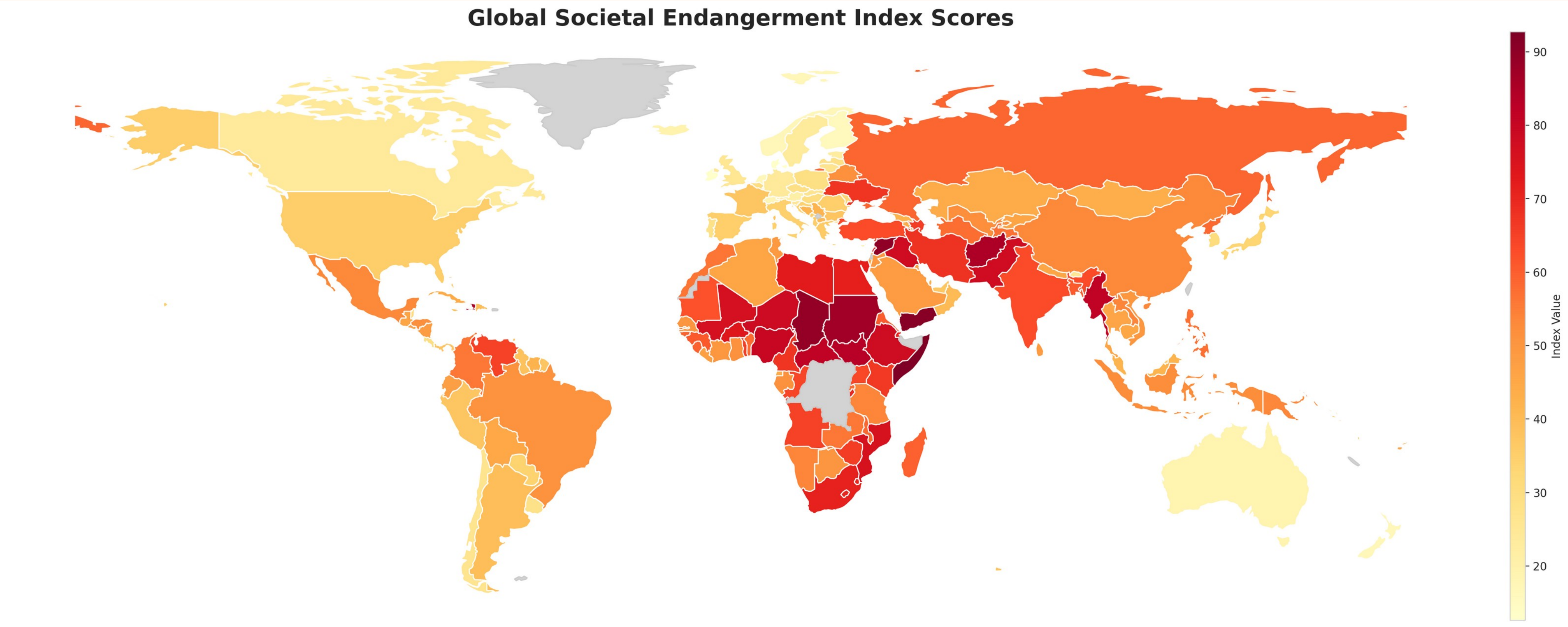
Results

GSEI









Results

The complete list is available on GitHub → GSEI.csv

GSEI

Total

	Country	GSEI
1	Somalia	92.73
2	Yemen	91.36
3	Syria	89.75
4	Chad	89.26
5	Sudan	86.96
6	Haiti	85.18
7	Afghanistan	85.16
8	South Sudan	83.26
9	Myanmar	81.59
10	Central African Rep.	81.50
11	Nigeria	79.70
12	Niger	78.36
13	Ethiopia	78.20
14	Pakistan	77.83
15	Iraq	77.81
16	Mali	76.34
17	Mozambique	75.52
18	Burkina Faso	73.84
19	Libya	72.70
20	Egypt	71.75
21	South Africa	71.38
22	Lebanon	71.34
23	Djibouti	70.67
24	Lesotho	70.36
25	Eswatini	70.16
26	Burundi	69.74
27	Iran	67.78
28	Cameroon	67.51
29	Ukraine	67.51
30	Kenya	66.64
31	Zimbabwe	65.88
32	Azerbaijan	65.62
33	Togo	65.22
34	Venezuela	65.02
35	Angola	64.94

Environmental

	Country	Score
1	Chad	109.97
2	Somalia	108.66
3	South Sudan	91.40
4	Nigeria	89.20
5	Niger	89.13
6	Haiti	84.29
7	Syria	80.86
8	Pakistan	80.48
9	Kuwait	76.54
10	Lesotho	75.84
11	Guinea-Bissau	75.82
12	Yemen	75.44
13	Burkina Faso	75.25
14	Turkmenistan	74.70
15	Iraq	74.29
16	Central African Rep.	72.52
17	Ethiopia	72.52
18	Madagascar	72.25
19	Bahrain	71.39
20	Afghanistan	71.28
21	Bangladesh	70.60
22	Mali	70.39
23	Ukraine	69.83
24	Kenya	69.47
25	Malawi	69.34
26	India	69.02
27	Egypt	67.22
28	Sudan	66.67
29	Iran	65.23
30	Zimbabwe	64.81
31	Libya	64.15
32	Morocco	63.54
33	South Africa	62.92
34	Azerbaijan	61.86
35	Mauritania	61.71

Political

	Country	Score
1	Syria	87.83
2	Afghanistan	81.96
3	Somalia	81.38
4	Myanmar	81.33
5	Sudan	80.58
6	Iran	79.80
7	Yemen	79.77
8	Libya	79.50
9	South Sudan	78.97
10	Mali	77.35
11	China	76.36
12	Ethiopia	76.20
13	Iraq	75.57
14	Central African Rep.	75.22
15	Pakistan	75.10
16	North Korea	74.93
17	Eritrea	74.79
18	Venezuela	73.35
19	Belarus	73.12
20	Egypt	72.98
21	Chad	72.40
22	Russian Federation	72.28
23	Turkmenistan	72.09
24	Tajikistan	70.71
25	Burkina Faso	70.65
26	Azerbaijan	70.37
27	Saudi Arabia	69.68
28	Cameroon	69.65
29	Burundi	69.47
30	Nigeria	67.21
31	Guinea	67.05
32	Bahrain	67.03
33	Equatorial Guinea	67.00
34	Zimbabwe	66.27
35	Djibouti	66.06

Social

	Country	Score
1	Central African Rep.	99.75
2	Niger	99.01
3	Afghanistan	91.33
4	Mozambique	87.68
5	Mali	86.19
6	Sudan	84.52
7	Sierra Leone	83.92
8	Yemen	83.87
9	Burundi	79.40
10	Gambia	79.16
11	Chad	78.92
12	Nigeria	78.60
13	Liberia	78.28
14	Haiti	78.27
15	Myanmar	77.50
16	South Sudan	76.81
17	Madagascar	76.66
18	Guinea	75.16
19	Togo	74.06
20	Congo	73.41
21	Somalia	73.38
22	Benin	73.22
23	Syria	73.05
24	Papua New Guinea	72.47
25	Ethiopia	71.65
26	Iraq	71.07
27	Djibouti	71.04
28	Burkina Faso	71.03
29	Senegal	70.96
30	Guinea-Bissau	69.99
31	Eritrea	69.66
32	Rwanda	68.87
33	Lebanon	68.26
34	Kenya	67.96
35	Libya	67.70

Economic

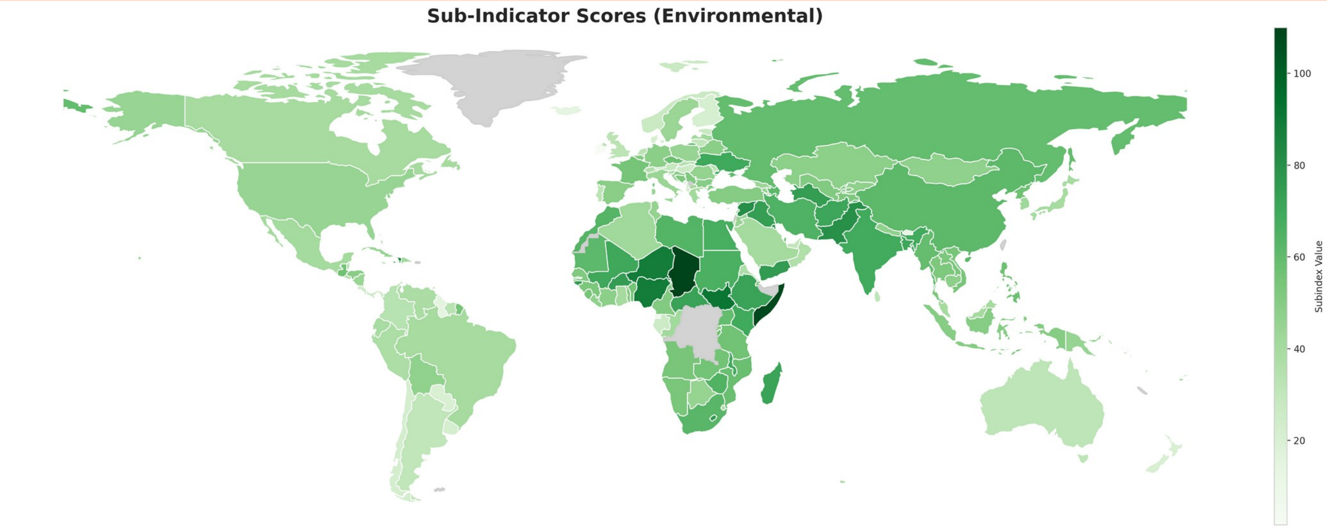
	Country	Score
1	Djibouti	96.80
2	Yemen	93.73
3	Lebanon	93.14
4	Eswatini	85.47
5	Egypt	85.29
6	Syria	83.07
7	Congo	81.24
8	Suriname	81.10
9	South Africa	80.88
10	Sri Lanka	80.70
11	Sudan	79.32
12	Colombia	79.13
13	Angola	78.18
14	Saint Vincent	77.17
15	Libya	75.80
16	Haiti	75.13
17	Afghanistan	72.70
18	Türkiye	71.60
19	Jordan	69.30
20	Tunisia	69.25
21	Venezuela	68.32
22	Somalia	67.56
23	Palau	67.41
24	Lesotho	67.07
25	Barbados	67.03
26	Mongolia	66.88
27	Argentina	66.06
28	Namibia	65.61
29	Chad	63.98
30	North Macedonia	63.77
31	Ecuador	62.91
32	Pakistan	62.51
33	Jamaica	62.33
34	Ghana	62.22
35	Guyana	62.18

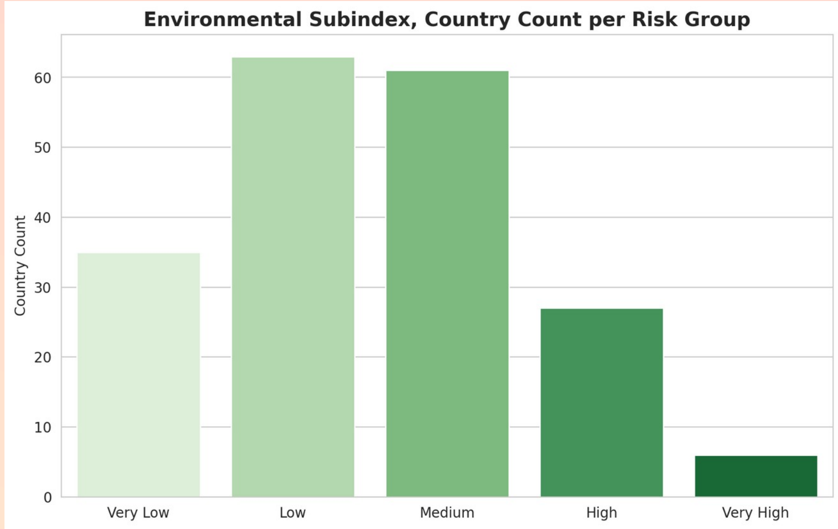
Threats

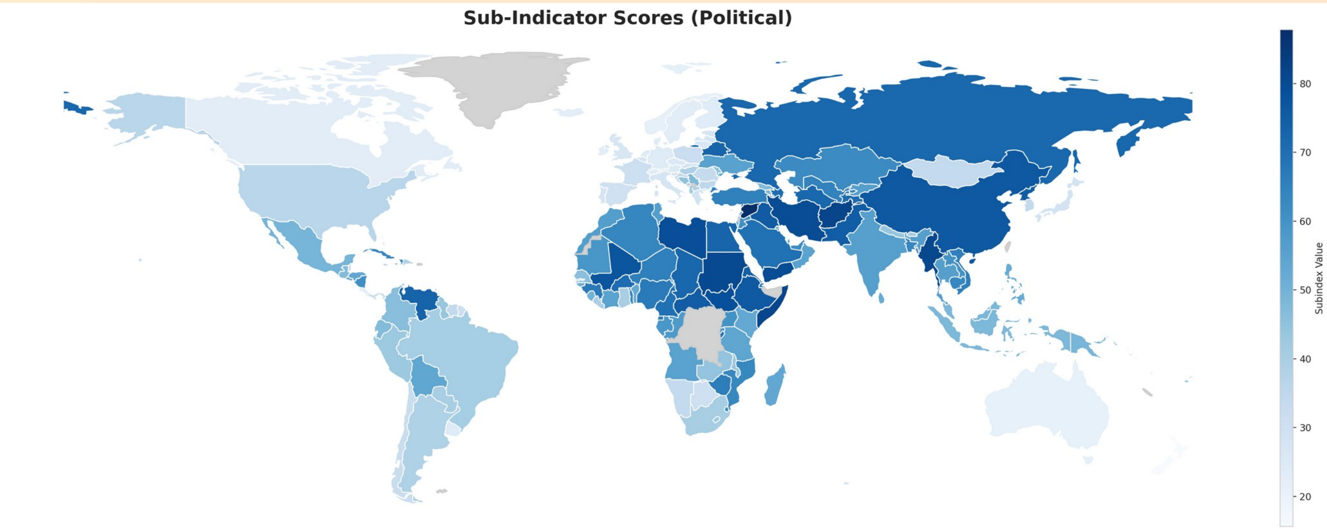
	Country	Score
1	South Africa	94.70
2	Myanmar	89.49
3	Haiti	86.58
4	Mozambique	86.31
5	Burkina Faso	86.30
6	Ethiopia	84.63
7	Ukraine	84.28
8	Nigeria	83.62
9	Sudan	83.09
10	Central African Rep.	82.80
11	Eswatini	82.65
12	Somalia	82.09
13	Lesotho	82.01
14	Mali	81.08
15	Yemen	80.58
16	Uganda	77.98
17	Chad	77.71
18	Mexico	77.37
19	Niger	77.29
20	Kenya	76.90
21	Syria	76.80
22	Cameroon	76.56
23	India	75.94
24	Togo	75.40
25	Pakistan	75.39
26	Philippines	73.54
27	South Sudan	73.11
28	Iraq	72.79
29	Brazil	71.72
30	Afghanistan	71.63
31	Azerbaijan	70.70
32	Indonesia	70.16
33	Burundi	68.10
34	Botswana	68.03
35	Russian Federation	67.33

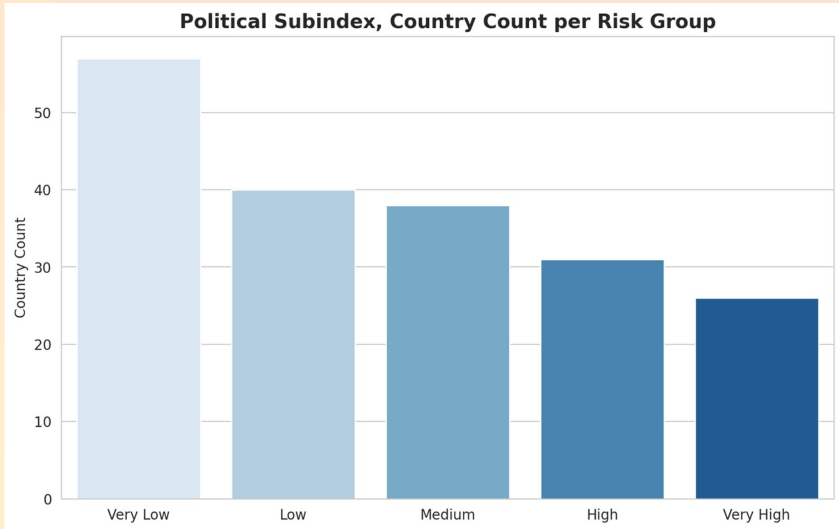
Results

GSEI



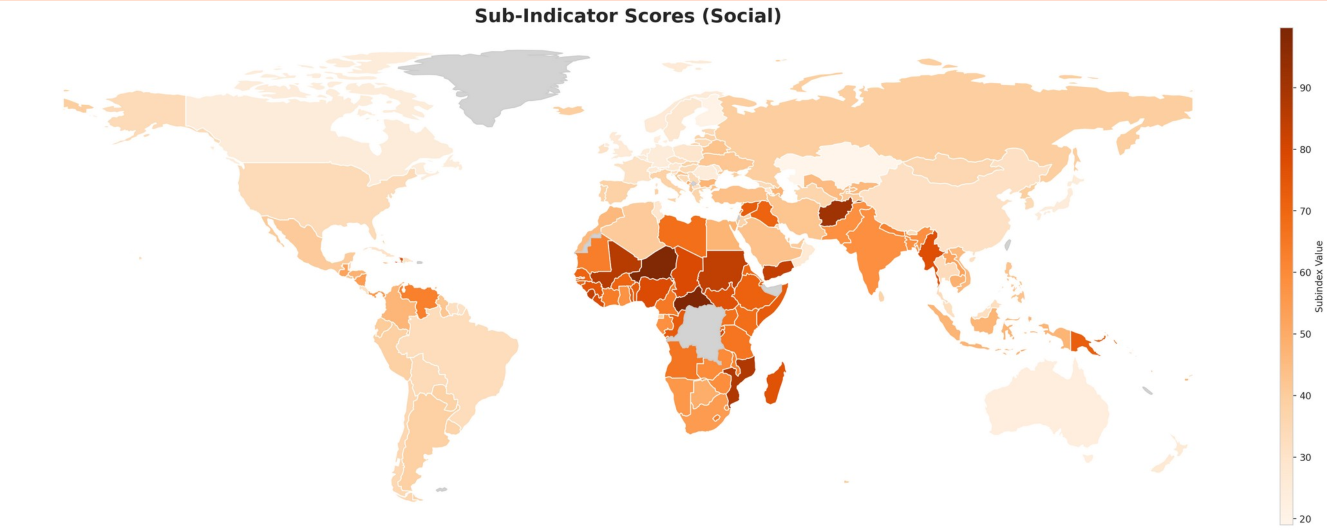


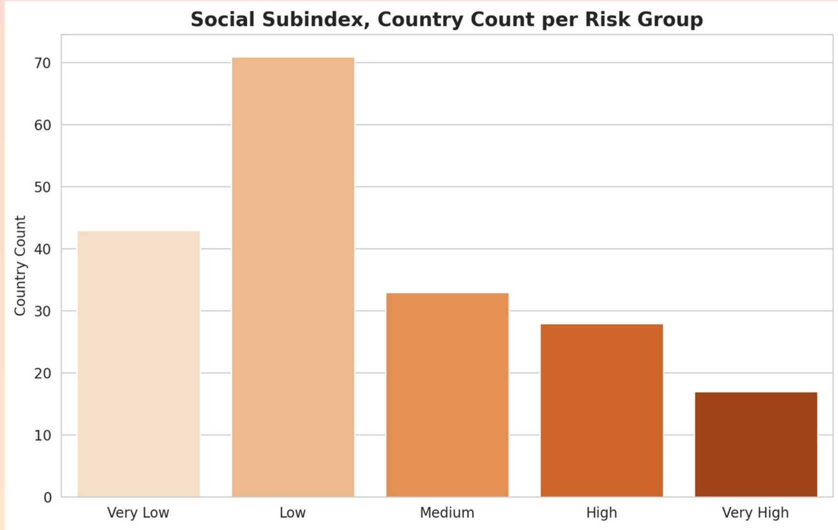


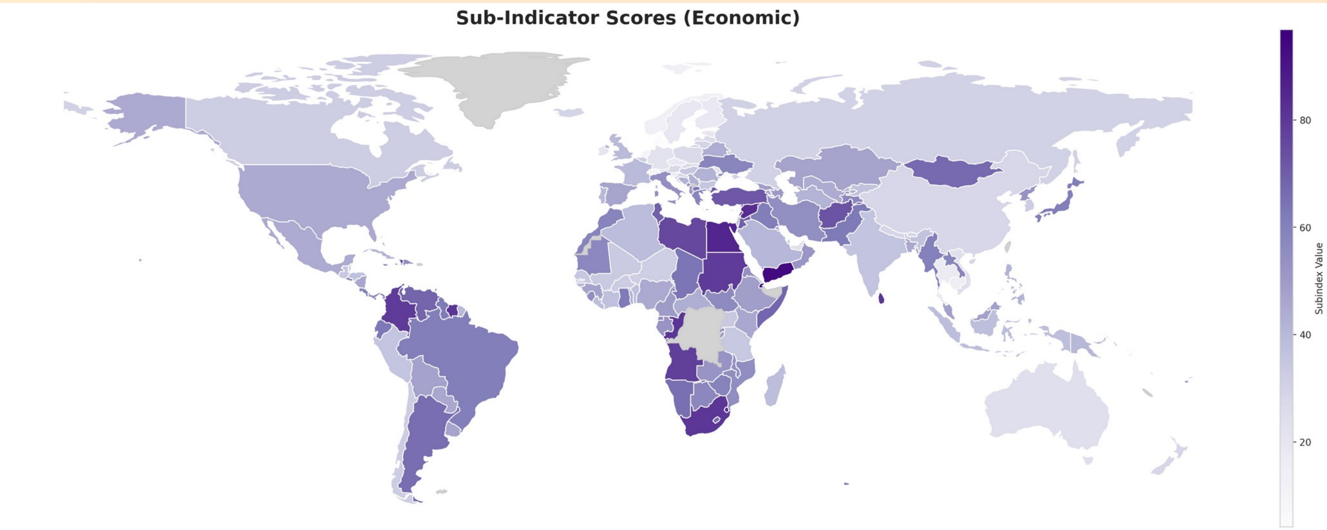


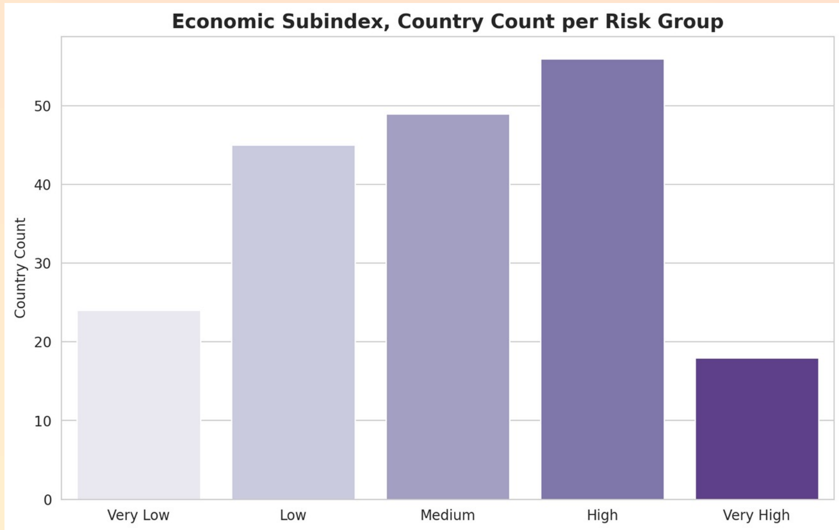
Results

GSEI



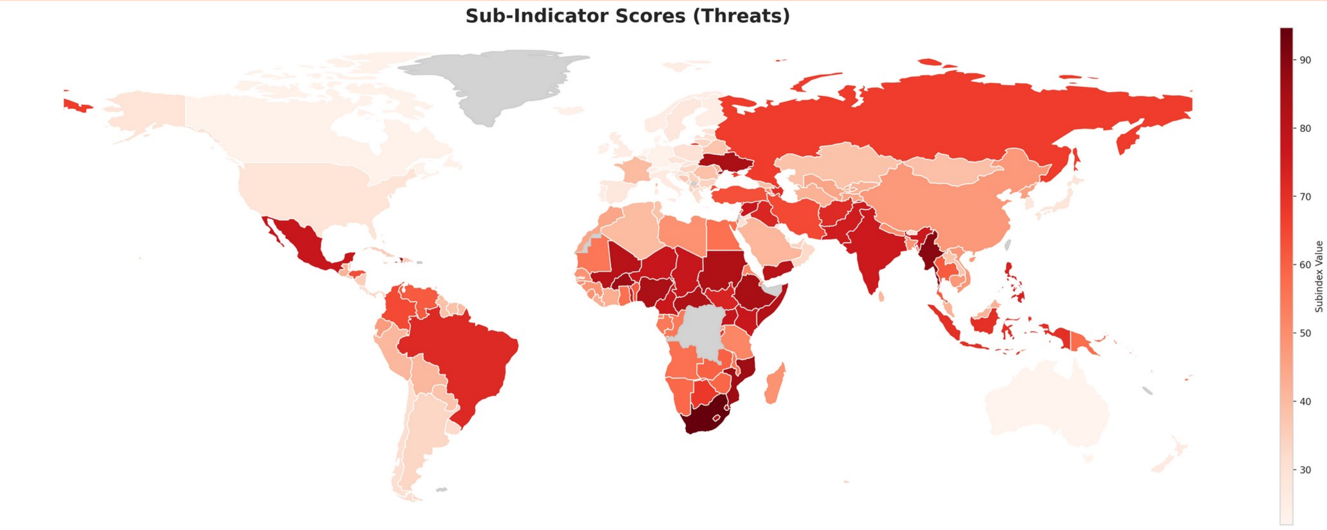


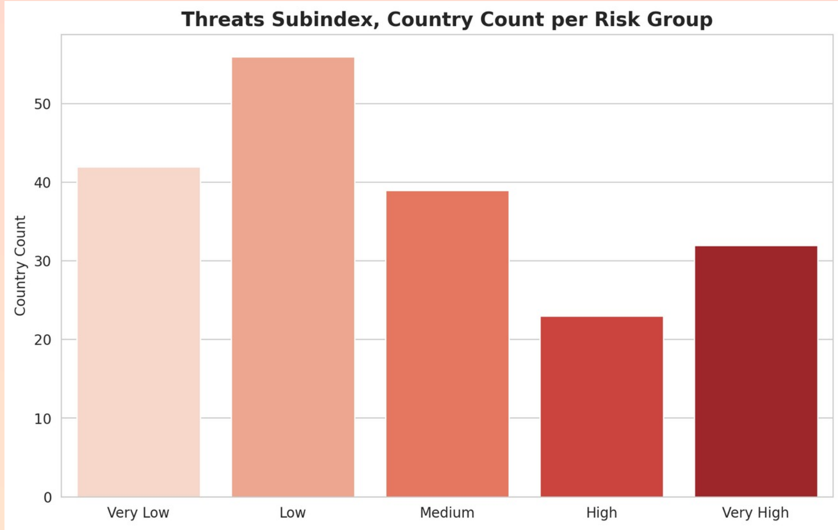




Results

GSEI

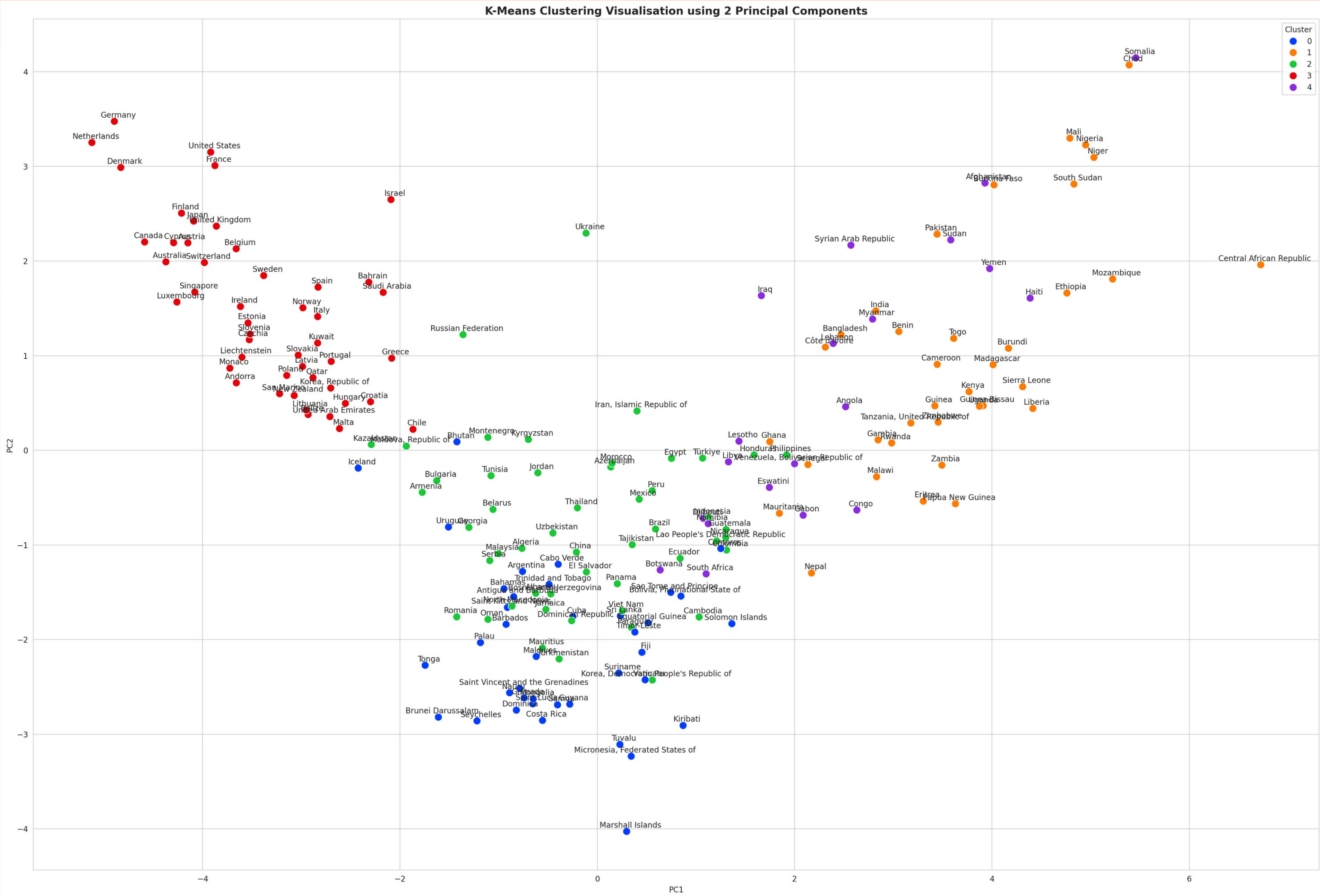




Total		Environmental		Political		Social		Economic		Threats	
Country	GSEI	Country	Score	Country	Score	Country	Score	Country	Score	Country	Score
176 Canada	24.14	176 Uruguay	23.00	176 Uruguay	24.43	176 United Kingdom	26.27	176 Germany	22.32	176 New Zealand	26.30
177 Malta	24.00	177 Denmark	21.74	177 Netherlands	24.13	177 Netherlands	25.74	177 Luxembourg	22.18	177 Malta	26.30
178 Tonga	23.92	178 Palau	21.63	178 Finland	24.00	178 Marshall Islands	25.58	178 Moldova	22.15	178 Liechtenstein	25.98
179 Sweden	23.70	179 Finland	21.62	179 Denmark	23.94	179 Brunei	25.48	179 Liechtenstein	21.29	179 United Kingdom	25.71
180 Estonia	22.93	180 Malta	21.18	180 Canada	23.51	180 Japan	25.28	180 Bahrain	20.92	180 Norway	25.51
181 Austria	20.98	181 Paraguay	21.04	181 Sweden	23.28	181 Germany	25.26	181 Ireland	20.75	181 Portugal	25.36
182 Iceland	19.63	182 New Zealand	20.93	182 Switzerland	23.27	182 Romania	25.14	182 Sweden	19.32	182 Finland	25.08
183 Australia	18.98	183 Albania	20.64	183 Costa Rica	23.09	183 Canada	25.08	183 Estonia	19.26	183 Switzerland	24.54
184 Luxembourg	18.29	184 Barbados	19.51	184 Iceland	22.96	184 Seychelles	24.76	184 Slovenia	18.68	184 Ireland	24.27
185 Switzerland	18.01	185 Tuvalu	15.50	185 Ireland	22.91	185 Denmark	23.76	185 Finland	18.48	185 Iceland	23.64
186 Norway	17.03	186 Guyana	15.49	186 Norway	21.93	186 Australia	23.70	186 Czechia	18.34	186 Canada	23.51
187 Netherlands	16.96	187 Iceland	15.05	187 Tuvalu	21.58	187 Dominica	20.67	187 Thailand	17.76	187 Germany	23.37
188 New Zealand	16.90	188 Liechtenstein	14.61	188 Barbados	21.46	188 Samoa	20.67	188 Cambodia	15.24	188 Australia	22.68
189 Finland	15.88	189 Montenegro	13.73	189 Australia	21.45	189 Finland	20.40	189 Norway	13.79	189 Denmark	22.62
190 Ireland	12.72	190 Bhutan	13.15	190 Luxembourg	21.29	190 Kazakhstan	18.98	190 Netherlands	10.11	190 Luxembourg	22.39
191 Denmark	12.61	191 Equatorial Guinea	7.63	191 Liechtenstein	19.55	191 Tonga	16.45	191 Switzerland	8.32	191 Netherlands	21.99
192 Liechtenstein	12.39	192 Ireland	1.59	192 New Zealand	15.63	192 Liechtenstein	13.96	192 Denmark	4.22	192 Austria	21.90

Results

GSEI



Global Societal Endangerment Index (GSEI)

Technical Report

Arthur Freye
Dep. of Mathematics, Natural Sciences and Informatics
Technische Hochschule Mittelhessen
University of Applied Sciences
35390 Gießen, Germany
arthur.freye@mni.thm.de

March 30, 2025

Abstract

The Global Societal Endangerment Index (GSEI) is a composite index to measure the vulnerability of countries to multidimensional risks and threats. These are weighted regarding their impact on society and population in each country. More imminent impact can be in form of catastrophes, conflicts or other severe societal shocks, while slower consequences are also taken into account, e.g. increased mortality or instability. The GSEI comprises 34 indicators, grouped into 5 sub-indicators: Environmental Risks, Political Instability & Governance, Social Vulnerability, Economic Stability & Infrastructure, and Global & Regional Threats. An expert was involved in developing the theoretical framework, notably in determining the weighting of the sub-indicators. The results compare well to similar indicators and allow for very informative visualisation.

Keywords Composite Index · Risk Index · Multivariate Analysis · PCA · Clustering · Visualisation

Accompanying Code Repository

There is a repository on GitHub that contains the code, data and graphics used to build the index and write this report. Mentions of "notebook [...]" throughout the report reference the python notebooks there. The final index scores are also available as `GSEI.csv`. The repository can be accessed using this link: https://github.com/arth3mis/global-societal-endangerment-index.

Preliminary note about generative AI

While it is very helpful to discuss ideas, methodology and get coding support, the tools would often overlook important aspects, even if they were mentioned right before. Also, as discussed in class, the scope of a well-developed composite index is still beyond AI skills. Even on a technical level, there is too much to keep in mind and the AI tools will just suggest the general steps it knows, which necessitates attention and questioning of the generated answers. As seen in the two chats I used for this project (ChatGPT[1] and Gemini[2]), I often intervened, corrected or expressed my doubts about the presented solutions. I also suspect that the phrasing of prompts (e.g. positive/negative framing of a concept by me) can influence the AI's opinion. They did tell me when I was wrong, but only in clear cases. I would not want to learn data analysis solely with these tools.

[1] https://chatgpt.com/share/67d22aaa-9880-8005-a37e-40a6d3811138
[2] https://g.co/gemini/share/9c12a83dba8f

Contents

1 Theoretical Framework

1.1 Background and Problem Definition

The world is facing a multitude of risks and threats, from climate change and natural disasters to social and political instability and economic crises. These risks are interconnected and can exacerbate each other, leading to cascading effects that endanger the population of countries and world regions. A composite index that captures the multidimensional nature of this endangerment can help policymakers, businesses, and civil society organisations better understand and address the challenges they face. By combining multiple indicators into a single score, the index provides a holistic view of a country's vulnerability and enables comparisons across different dimensions of risk.

This is especially important in a globalised world where the impacts of risks in one country can quickly spread to others through trade, migration, and other channels. NGOs and international humanitarian organisations can use the index to prioritise their interventions and allocate resources more effectively to the countries and regions that are most at risk. This is also where I got the idea from: The IRC (International Rescue Committee) publishes an annual Emergency Watchlist [1] where it details the countries that face the biggest risks of humanitarian crises. This is a list, not an index, and it focuses on a few major aspects that are directly linked to the IRC's work. I want to create an index that builds on this idea, but with a broader scope, capturing a wider range of risks and threats. This report details the development of the Global Societal Endangerment Index (GSEI), which aims to be a comprehensive index as described above.

1.2 Data Domains and Justification

There are five main domains of data that are chosen to represent the dimensions of endangerment of countries: environmental (also appears as 'env' throughout this report), political ('pol'), social ('soc'), economic ('eco'), and global risks and threats ('threat'). These are going to be sub-indicators for the total GSEI.

Environmental Risks are an important factor because they affect the health and well-being of populations, the stability of ecosystems, and the sustainability of economic activities. Countries that are highly vulnerable to environmental risks may face challenges such as water scarcity, food insecurity, and natural disasters.

Political instability and governance are important because they can lead to unrest, violence, and conflict within and between countries (see also Appendix A). Countries with weak governance structures may struggle to respond effectively to crises and may be more susceptible to external influences.

Social vulnerability is a key factor in determining how well populations can cope with and recover from crises. Factors such as poverty, inequality, and access to education and healthcare can significantly impact a country's resilience to risks. Generally, this is a more long-term measure, as many indicators in this category are not directly linked to imminent crises, but contribute to a slow decay of society, which must not be neglected.

Economic stability and infrastructure are crucial for a country's ability to withstand crises and support their population, especially the most vulnerable lower class. Countries with strong economies and well-developed infrastructure are better equipped to respond to shocks and provide support, mainly in monetary form.

Global and regional threats are important because they can have far-reaching impacts on countries and regions, affecting trade, migration, and security. These threats can include geopolitical tensions and pandemics.

In the GSEI, I want to include general and some specific data, e.g. related to the COVID-19 pandemic. All data should be up-to-date (preferably from 2020 or newer) and reliable, coming from reputable sources such as the World Bank, WHO, United Nations, and other international organisations. The sources and indicators are selected with current trends (e.g. cybersecurity) and future risks (e.g. conflict risk, extreme temperatures) in mind. The data is also comprehensive, covering a wide range of risks and threats that countries face. This makes the data I'm using to build the index appropriate for the task, as they are widely used in other research and analysis.

Indicators that measure quantities with more imminent and severe impact on society should receive a higher weight than those indicating slower developments like general health, increasing mortality, inequality or instability. Regarding the weighting of the sub-indicators of the final index, I have consulted a friend of mine who is an expert in international politics and strategies, Salah Alnachawati. His full written assessment can be found in Appendix A. Salah suggested that the political dimension should be weighted higher than the others (28%), as a functioning government is often the most important factor in determining a country's vulnerability to risks and threats. As further rankings, he suggests (in decreasing order): Economic (23%), Threats (22%), Social (15%), Environmental (12%).

While Salah makes very good points, as a political scientist he has a bias towards political data. I will weight the environmental category higher, since I also set a focus of the index on imminent crises, which are often environmental in nature. This is also supported by the current trends in global risks, where climate change and natural disasters are becoming more frequent and severe. The INFORM risk index is focused on natural crises and disasters, and it encompasses similar categories to mine [2]. Joining the expert opinion with the research and related indices, these weightings are chosen:

- Environmental Risks: 20%
- Political Instability & Governance: 25%
- Social Vulnerability: 15%
- Economic Stability & Infrastructure: 18%
- Global & Regional Threats: 22%

2 Data Selection

The following list contains all indicators that are selected for the GSEI, sorted by sub-indicator and further grouped by source, with links to the bibliography. Note that this symbol ↓ shows that the indicator is inversely related to endangerment and to the index, which is dealt with in section 5.

- **Environmental Risks**
  - Climate Change [3]
    - Maximum relative temperature change (°C) in the last 10 years, compared to a 1951-1980 baseline
  - Exposure to Natural Disasters (earthquakes, floods, hurricanes) [4]
    - Total affected population
    - Total damage in US$ (adjusted)
  - Air and Water Pollution Levels [5]
    - Years of lost life due to unsafe water, sanitation, and handwashing
    - Years of lost life due to air pollution
  - Water Scarcity & Food Security [6]
    - ↓ Renewable internal freshwater resources per capita (cubic meters)
    - ↓ People using safely managed drinking water services (% of population)
    - Prevalence of moderate or severe food insecurity in the population (%)
    - Prevalence of undernourishment (% of population)
- **Political Instability & Governance**
  - Governance Quality [7]
    - ↓ Corruption Control
    - ↓ Rule of Law
    - ↓ Political Stability
    - ↓ Government Effectiveness
    - ↓ Regulatory Quality
    - ↓ Voice and Accountability
  - Regime Type [8]
    - ↓ Regime type in most recent year available (Democracy/Autocracy)
- **Social Vulnerability**
  - Poverty [6]
    - Population percentage below societal poverty line
  - Health System Strength [9]

    - ↓ Current Health Expenditure (CHE) per Capita in US$
    - Rest of the World (RoW) as % of Current Health Expenditure (CHE)
  - Access to Education [10]
    - ↓ Population age 15+ literacy rate
    - Population age 15+ with no education
    - ↓ Government expenditure on education
    - ↓ Lower secondary school completion rate
  - Crime & Violence Rates [11]
    - Global Organised Crime Index
- **Economic Stability & Infrastructure**
  - Debt & Economic Resilience [**debt**]
    - General government debt as % of GDP
    - ↓ Total reserves (% of total external debt)
  - Income Inequality [6]
    - Gini Coefficient
    - Income share held by highest 10%
  - Inflation [6]
    - Inflation, consumer prices (annual %)
  - Unemployment [6]
    - Unemployment, total (% of total labour force)
    - Unemployment, youth (% ages 15-24)
  - Energy Security & Infrastructure [6]
    - ↓ Access to electricity (% of population)
    - ↓ Renewable energy consumption (% of total final energy consumption)
  - Internet Access & Security [6]
    - Individuals using the Internet (% of population)
    - Secure Internet servers (per 1 million people)
- **Global & Regional Threats**
  - Geopolitical Tensions [2]
    - Projected Conflict Probability
    - Current Conflict Intensity
  - Pandemic Preparedness & Disease Burden [9] [5]
    - ↓ Expenditure on COVID-19 per Capita in US$
    - Prevalence of infectious diseases (4 single indicators)

During data collection I noticed that some indicators are missing data for some countries in recent years, which will require increased effort in the imputation step, since regional data is not easy to impute without creating misrepresentations (e.g. by just taking global average).

2.1 Resolving Inconsistencies

The collection step also involves re-structuring many of the datasets to make them compatible with each other. For example, all World Bank indicators are exported with Country-Indicator pairs as rows. I need to pivot them to have one row per country with all indicators as columns. Since I use one variable per indicator (not multiple years), this is the preferred format.

The country names also don't match across all datasets, so I need to harmonise them to ensure that the data can be merged correctly. Important note: The country mapping is done solely based on inspection of the datasets, without any attached political motivation.

Other source data need decisions that have semantic implications are as follows:

1. The EM-DAT natural disaster data is already in the right format, but I need to aggregate it to the country level, as it is currently at the event level.

2. Temperature Change: The raw data is available for 1961-2022, but I will use the maximum value of the latest 10 years for each country. This is chosen to emphasise recent development and focus on extreme events rather than averages, since extreme events pose significant danger to the population.

3. Many datasets are incomplete for recent years. In notebook 2.1, I created overviews for the availability of data for each indicator and year, so that I can make informed decisions about which years to include in the index, balancing recency and completeness of the data. The exact years chosen for each indicator are documented in the notebook.

4. For the education indicators, some selected years are quite old (2010-2019) compared to the other indicators used in the index. This is due to the low availability of data for the education indicators. However, education (or lack thereof) has effects over decades, so the selected years are still relevant for the index.

2.2 Quality Checking

Given the reputable sources, the indicators are considered to be analytically sound. All indicators are measurable and consist of continuous data, so there will be no problem with PCA later (which is not really applicable to categorical data).

To check the country coverage of the loaded indicators, availability is calculated per indicator and per country. Since there are 249 countries in the initial list used, there are likely ones where so many data are missing that the index cannot represent them well. These countries should be excluded from the index to increase its statistical validity and reduce the amount of missing data. The same should be done for indicators that have too many missing values, as they would not contribute much to the index. However, some indicators have important information even if they are missing for many countries. I set a baseline of 50% availability/country coverage for comparison. The calculation results of the whole preprocessed dataset show that 55 countries have less than 50% availability, and 10 indicators have less than 50% country coverage.

My first step here is to restrict the country list to UN member states, which reduces the number of countries to 193. This is a common practice in global indices. After the reduction, only 1 country has less than 50% availability, and 9 indicators have less than 50% country coverage (see Table 1). This shows that the country list reduction fits with my data, as e.g. small island states and disputed territories are often not included in global datasets. This shows that the country list reduction fits with my data, as e.g. small island states and disputed territories are often not included in global datasets.

Table 1: Country Coverage per Indicator, showing only entries with below 50% availability.

Indicator	**Available Countries**	**Availability (%)**
RoW Health Expenditure (%)	40	21
COVID-19 Expenditure per Capita (US$)	57	30
Population below Poverty Line (%)	71	37
Gini Coefficient	71	37
Income Share Top 10%	71	37
Lower Secondary Completion Rate (%)	72	38
Literacy Rate (%)	77	40
Government Debt (% of GDP)	82	43
Total Reserves (% of External Debt)	95	49

This is acceptable for now, I will come back to this when issues arise during imputation or multivariate analysis. Inspecting the distribution boxplot (see Figure 1) shows that most indicators have a lot of outliers (outside the 1.5 inter-quantile range (IQR)), which is expected given the wide range of values across countries. The outliers do give relevant and real information, so I will not remove them from the dataset. For example, Ukraine's current conflict intensity is far beyond all other countries, which should not be ignored in an index that measures societal endangerment. However, the outliers are important to consider when normalising the data (see section 5).

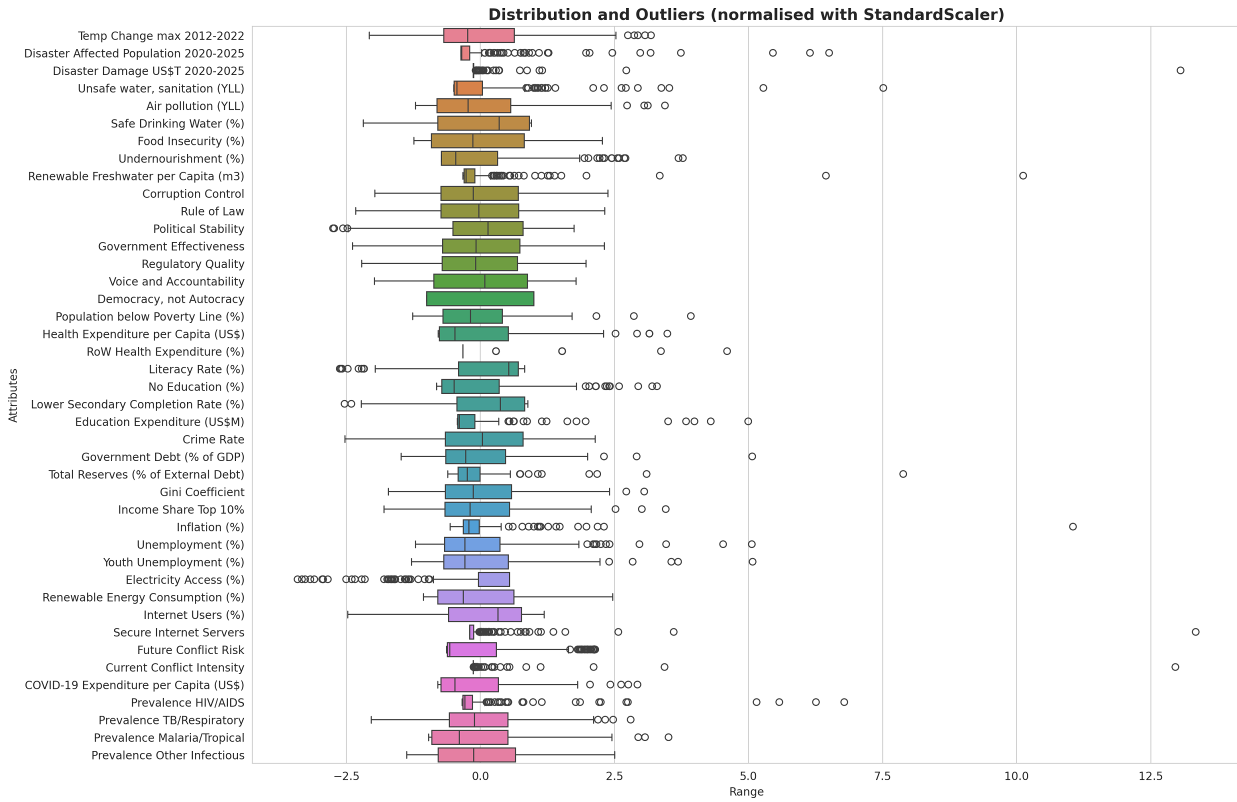


Figure 1: Boxplot of each indicator's distribution (scaled with z-score), showing that many outliers exist in the data.

3 Missing Data Imputation

19% of data are missing in the dataset. This is a high percentage, but it is not unexpected given the complexity of the index and the wide range of indicators used. The missing data are not missing completely at random (MCAR), as some countries have more missing values than others. However, the missing data do not appear to be strictly systematic (MNAR), as the missing values are spread across different indicators and countries. As seen in notebook 2.2, The 25th percentile of per-indicator availability is 64% while the 25th percentile of per-country availability is 74%; this suggests that missingness is not solely a country-level phenomenon. While MNAR cannot be ruled out, this finding supports the argument that the data are missing at random (MAR), which is a less severe form of missingness. This is important because it means that the imputation will not introduce (as much) bias into the index if a valid method is chosen.

Without MCAR data, and also judging from a logical point of view, filling missing values with the global average is not a good idea. This would introduce much bias into the index, as the global average may not be representative of the missing values. Instead, I will use a clustering method to group countries based on their similarity in the available data, and then impute the missing values with the average of the countries in the same cluster. "Cluster Analysis serves as [...] a method for selecting groups of countries for the imputation of missing data with a view to decreasing the variance of the imputed values" [12, p. 26]. This method is more likely to preserve the structure of the data and reduce the bias introduced by imputation.

The chosen clustering method for imputation is hierarchical clustering. This method is well-suited for handling missing data and allows for a visual representation of the clusters, which helps in identifying a suitable cluster number by checking the vertical distances between branchings in the dendrogram, which represent similarity (see Figure 2). I chose Ward's method for linkage. This method minimises the variance within clusters when forming them. The methods discussed in class were also tested (single, complete, average linkage), but yielded worse results in terms of cluster sizes and interpretability.

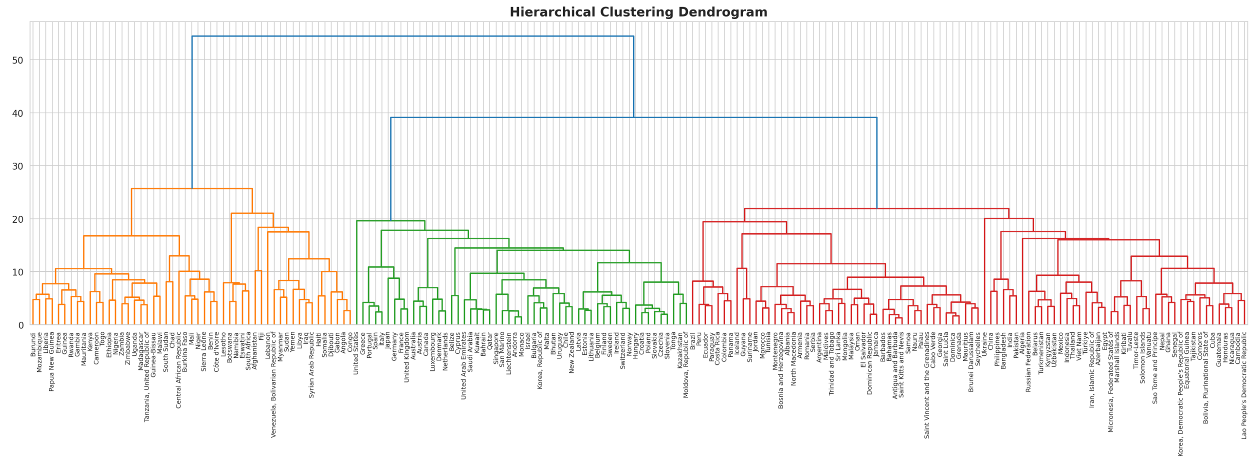


Figure 2: Result of the hierarchical clustering, showing that the countries can be best grouped in 3-6 clusters.

Upon applying hierarchical clustering, the maximum number of clusters with each one still having a reasonable size is 5. This achieves a good balance between data availability in each cluster, the number of clusters, and the interpretability of the results. The clusters are based on the available data, so they may not correspond to any real-world groupings of countries. However, the goal here is to impute missing values, not to create meaningful, human-interpretable clusters (which is similar to the loss of interpretability in PCA).

After the imputation, 3 variables are still missing data for a whole cluster, so I take the global average for these variables. This is a simple imputation method which should only be used when the values are missing (completely) at random (MCAR). As a whole cluster is missing, it is likely that the values are not MCAR here. However, I argue that the method is okay to use because a very low number values is still missing: $3 \cdot 21 = 63$, which is only 0.7% of the $42 \cdot 192 = 8064$ values in the whole selected dataset. During multivariate analysis, one of these indicators will be removed because of multicollinearity, which further reduces the values imputed with this rough technique.

4 Multivariate Analysis

I consider a correlation > 0.8 as severe multicollinearity which should be dealt with first [13]. But all correlations > 0.7 indicate problematic multicollinearity, so I will try to eliminate them [14]. Based on [15], I found a good way to display collinearities between independent variables as a heatmap matrix. This is shown in Figure 3.

I have 81 variable pairs with a correlation above 0.7, and 25 variable pairs with a correlation above 0.8 (see Table 2). This is quite severe, so further analysis is needed. I don't want to remove too many variables, as this would reduce the richness of the index.

Table 2: Top 10 indicator pairs with the highest correlations.

Indicator 1	**Indicator 2**	**Correlation (abs)**
Gini Coefficient	Income Share Top 10%	0.98
Rule of Law	Regulatory Quality	0.94
Corruption Control	Regulatory Quality	0.94
Rule of Law	Government Effectiveness	0.94
Government Effectiveness	Regulatory Quality	0.93
Unemployment (%)	Youth Unemployment (%)	0.93
Corruption Control	Rule of Law	0.91
Corruption Control	Government Effectiveness	0.89
Safe Drinking Water (%)	Lower Secondary Completion Rate (%)	0.84
Lower Secondary Completion Rate (%)	Internet Users (%)	0.83

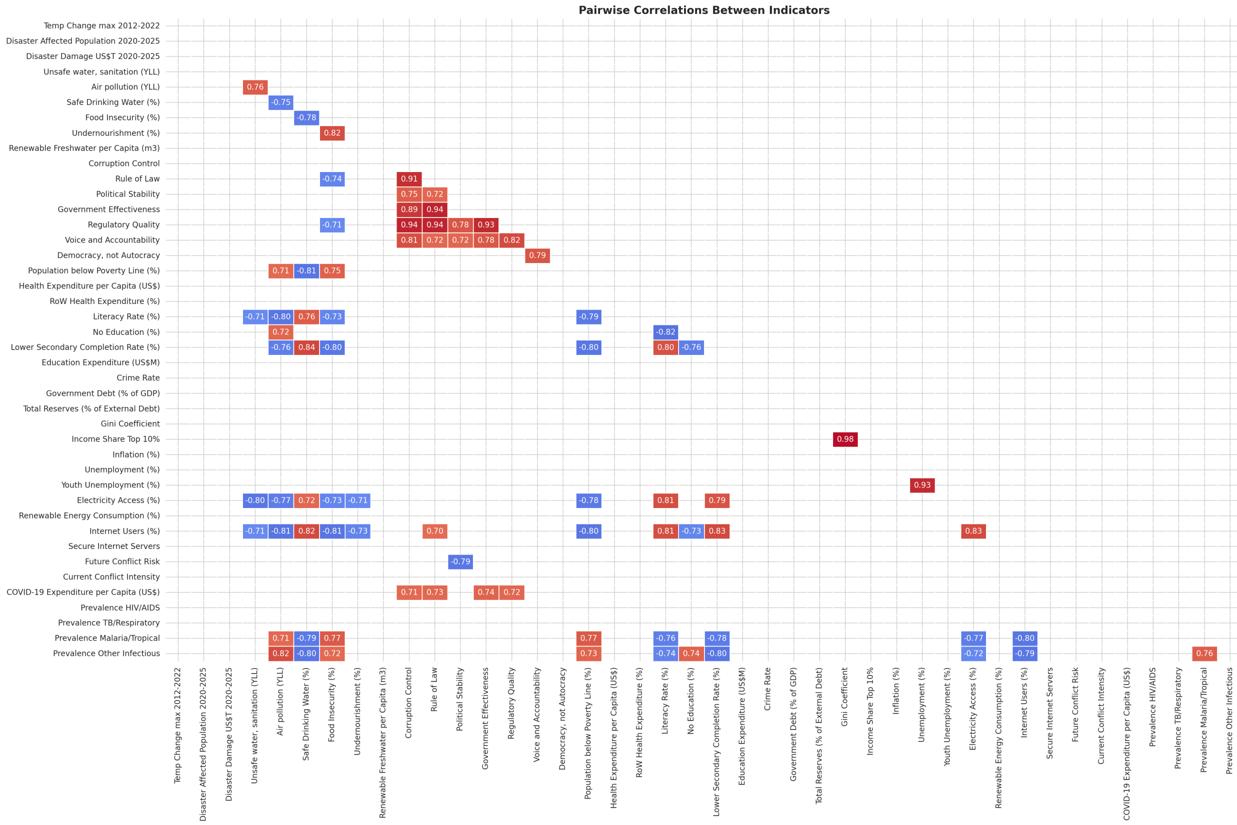


Figure 3: Correlation matrix between indicators, showing all values with an absolute > 0.7, and coloured as a heatmap.

My investigation of the coloured correlation matrix and the pairwise list lead to the following decisions:

- Gini Coefficient and Income Share Held by Highest 10% have extremely high correlated (0.98). This makes sense, as they both measure income inequality. I will only keep the Gini Coefficient.
- Unemployment (%) and Youth Unemployment (%) are highly correlated (0.93). This is also expected, as youth unemployment is a subset of total unemployment. Youth unemployment is more relevant to societal endangerment, so I will remove Unemployment Total from the dataset.
- Internet Users (%) is highly correlated with many other indicators, so I will use Electricity Access (%) as a proxy for this, since they are also causally related (logical deduction).
- In a second iteration, Electricity Access is also removed, now “Population below Poverty Line” acts as a highly correlated proxy for both this and Internet Users.
- I will remove Lower secondary school completion rate, and Literacy Rate, “No education” is suitable as a proxy here.
- Food Insecurity and Undernourishment are highly correlated (0.82). I will remove Food Insecurity because it also highly correlates with 10 other indicators.
- For 3 groups of indicators, I will apply PCA to reduce the dimensionality and remove multicollinearity. These are described below.

After the removal of statistically problematic indicators and excluding the ones to be transformed, the correlation matrix looks a lot better (see Figure 4).

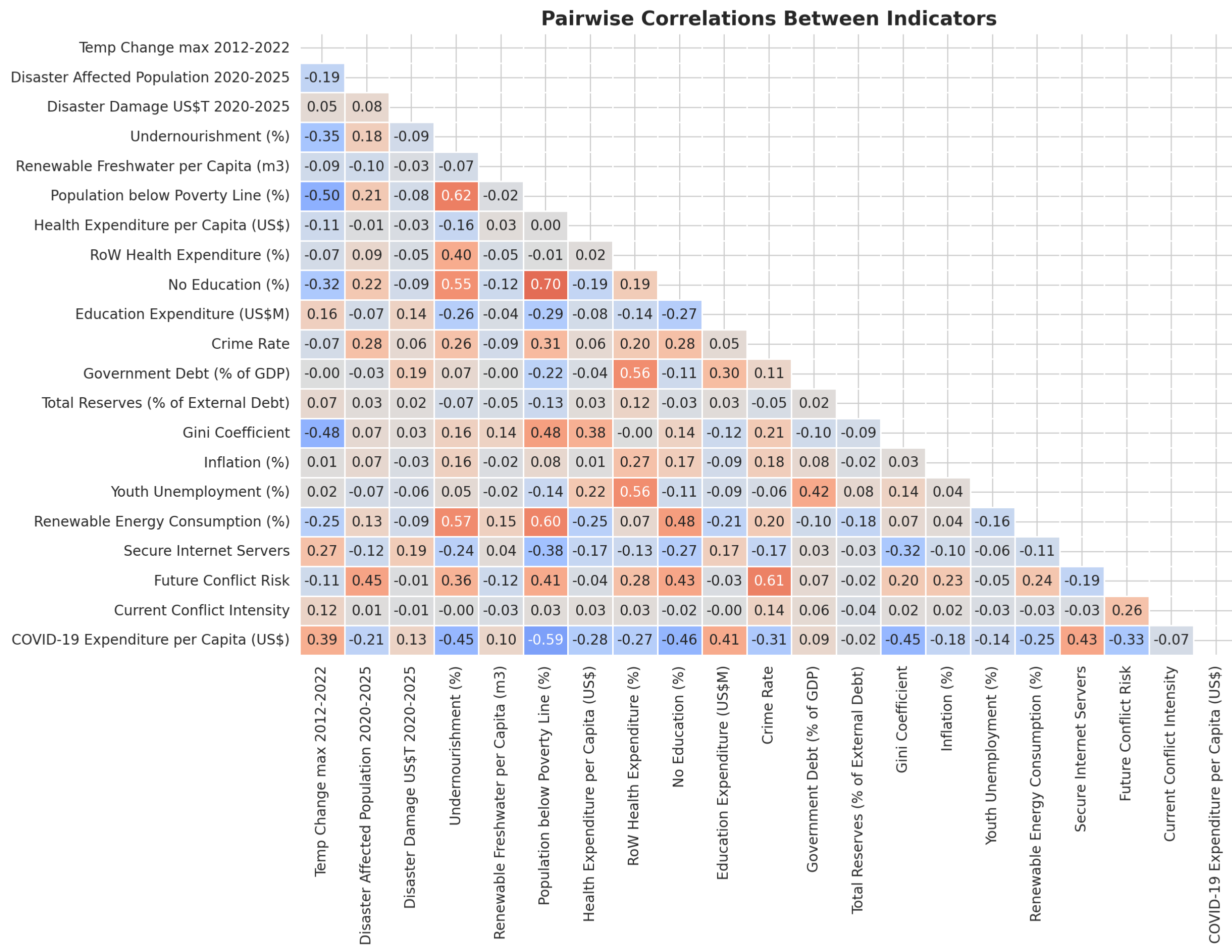


Figure 4: Correlation matrix between indicators, showing all values and coloured as a heatmap.

4.1 PCA

PCA is applied to the whole dataset first to find noteworthy properties about it, then it is used to transform groups with high multicollinearity. PCA already needs normalisation in order not to face disproportional variance along different axes.

PCA on the whole dataset reveals information about the variance structure of the data. For 99% variance, 34 out of 42 components are needed. The first PC explains 38%, then there is a steep drop-off as seen in Figure 5.

This suggests that the data is complex and that there are multiple independent patterns of variability in the data, which is expected given the wide range of indicators used in the index. The PCA also shows that the data is not easily reducible to a few components, warranting a composite index to capture the multidimensional nature of societal endangerment. The component loadings support this (see Figure 6), as they are spread across many indicators [16], with no single indicator dominating any component (the maximum is 62%, and this is only on the 7th PC).

These indicator groups are transformed using PCA:

- All political data
- Water & Air
- Infectious diseases

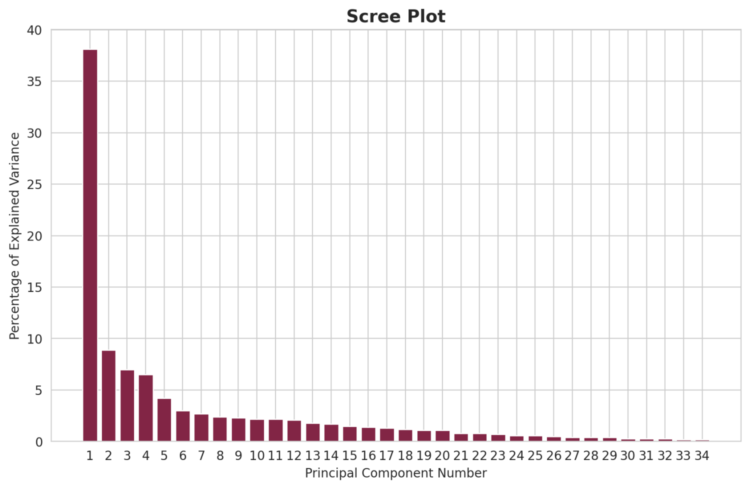


Figure 5: Scree plot showing the explained variance of each principal component (whole dataset).

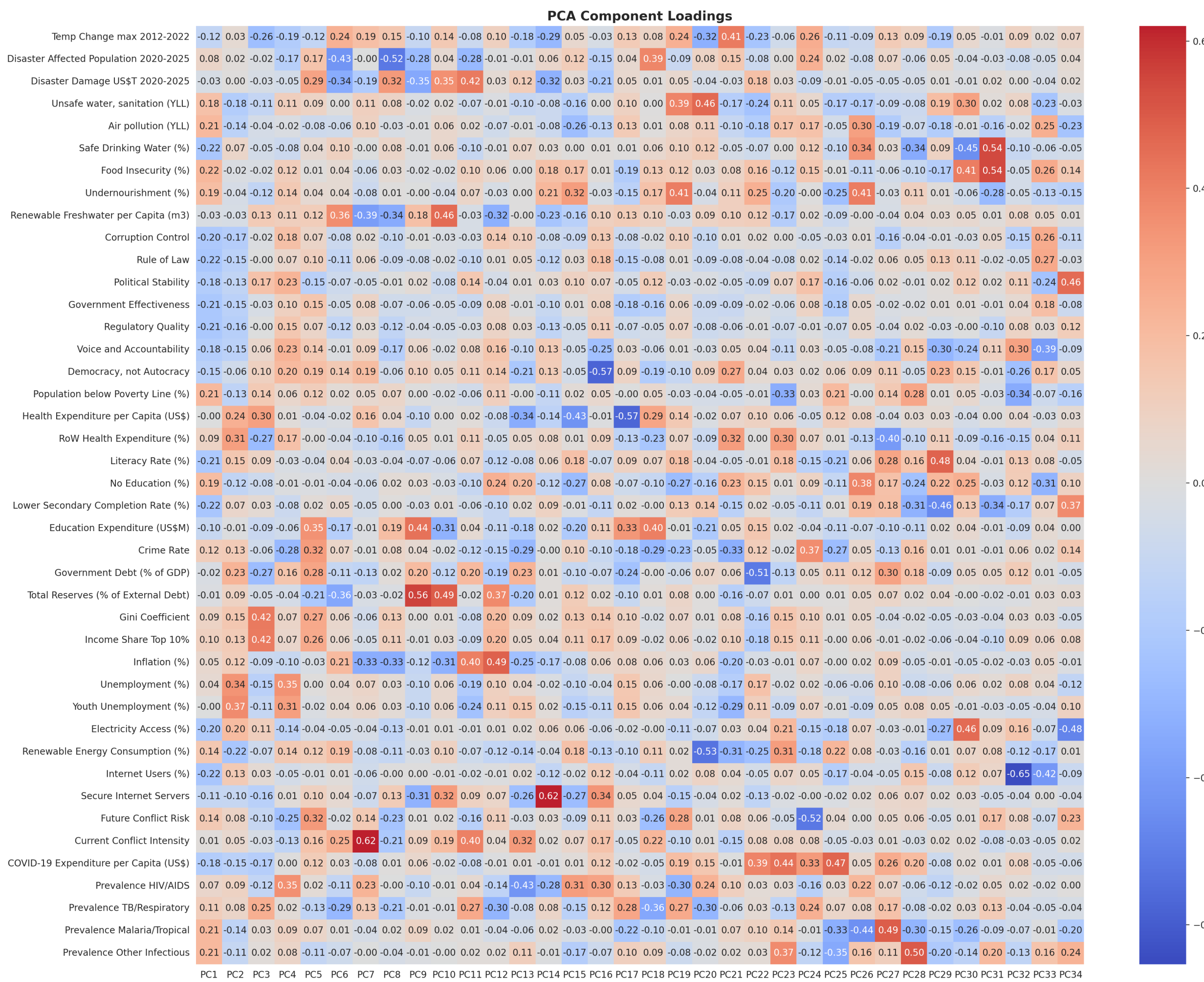


Figure 6: Component loadings for PCA on the whole dataset, coloured as a heatmap.

For Water & Air, there are mixed relations to the resulting index within the group. "Safe drinking water" is inversely related to endangerment. So I have to check the loadings in PC1 to see if the inversely related variable has a negative loading and the others positive ones. That is the case (see Figure 7), so the PCA results can be used for the transformation of the group. Since PC2 and PC3 cover different dimensions and more "abstract" statistical constructs, they don't need to be checked in this way.

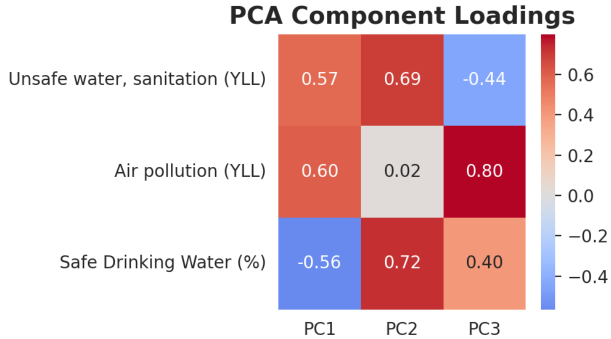
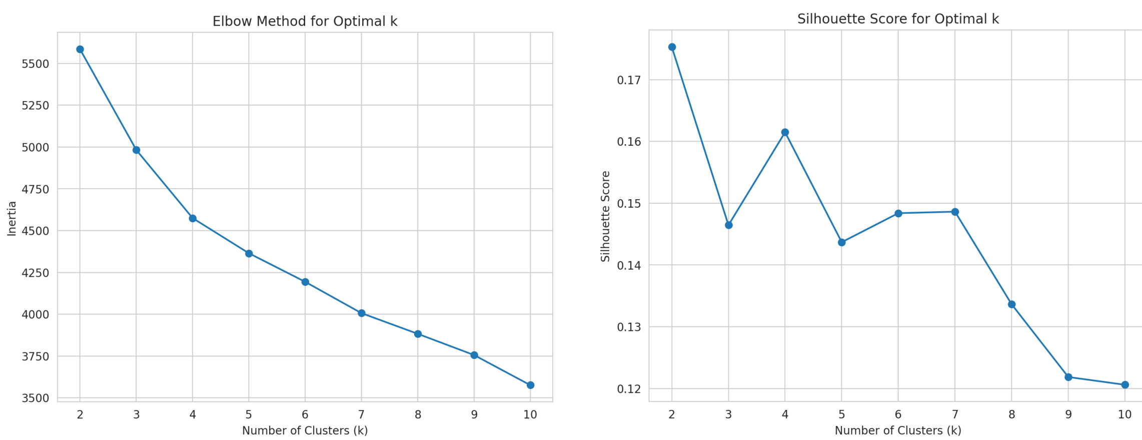


Figure 7: Component loadings for PCA on the group of air and water pollution indicators.

I decide to keep almost all the principal components, as the groups are rather small (3-7 variables) and I want to retain a high proportion of explained variance (99%).

4.2 Clustering

K-means clustering [17] is used to group countries based on their similarity regarding endangerment. This can help identify patterns in the data and inform the weighting of the sub-indices. Possible to investigate are the cluster centres, i.e the mean values of each indicator for each cluster; cluster member counts, mainly to check the balance or find unusually small/large clusters; and the countries in each cluster. Additionally, the clusters can be visualised when adding PCA to reduce dimensionality.



(a) Elbow plot

(b) Silhouette plot

Figure 8: Plots to choose the cluster number k.

The best cluster count can be found by trying different ones and plotting metrics (see Figure 8). The silhouette score suggests $k = 4$, but the elbow plot and resulting cluster sizes lead me to choose 5 clusters. This yields decently distinguishable results in the cluster centres. Then, a visualisation is done with 2 PCs in 2D (see Figure 9). There is good separation visible between the clusters, which suggests that the clustering algorithm has successfully identified distinct groups of countries based on their vulnerability to different types of endangerment. It can be seen that countries with generally higher wealth and development status (which are not direct indicators) are grouped closely together and quite far off from the other clusters. This is expected, as wealthier countries are generally better equipped to deal with risks and threats. The other clusters are still distinguishable, but appear more mixed, which may reflect the complex interplay of different

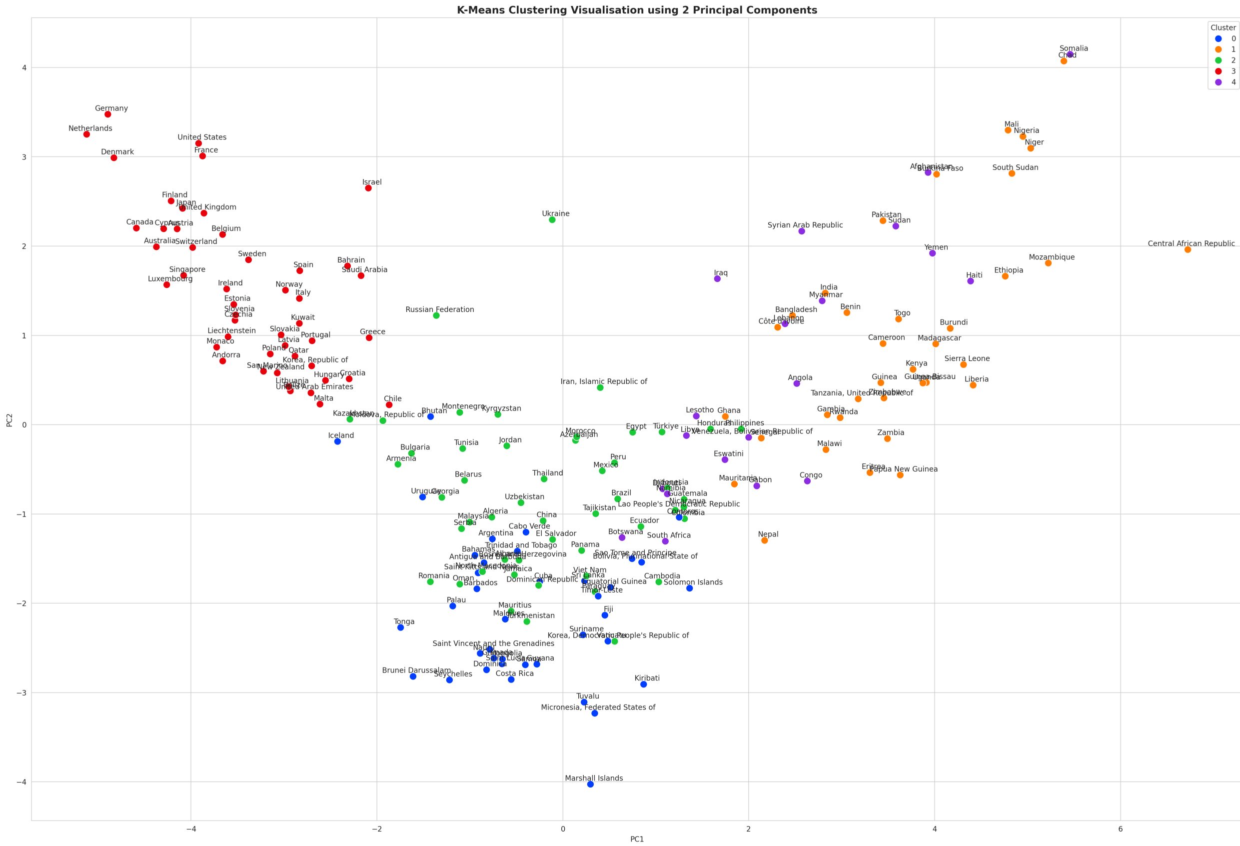


Figure 9: Result of K-Means clustering, visualised using the first two principal components of the data. Some clusters are already well separated despite the 30% explained variance of the PCs.

factors that contribute to societal endangerment. PC1 and PC2 can only explain around 30% of the variance in this case.

5 Normalisation

From the outlier analysis in step 2 (section 2), I know that all indicators have up to 17% values outside the 1.5 IQR (Interquartile Range). Those that don't are mostly removed or transformed in step 4 (section 4). A usual way is to standardise all variables to a mean of 0 and a standard deviation of 1 (z-score normalisation). This is done to ensure that all variables are on the same scale in the index and e.g. for PCA (as used in the previous step). For my data with quite extreme outliers, I considered and tried to use the RobustScaler from scikit-learn, which scales the data according to the median and the IQR instead of the mean and standard deviation. This is more robust to outliers, but my indicators have such a wide range of values that the IQR is not a good measure for scaling. The data is still not close to normally distributed after RobustScaler (see Figure 10), and outlier countries dominate the index. Thus, I will use z-score scaling for the final normalisation of the data, but apply a log-transformation to variables that are heavily skewed (apparent from the histograms in Figure 11 where a large spike occurs around 0).

Indicators can be log-transformed before scaling to reduce the skewness of the distribution and make the data more normally distributed [18]. This skewness happens e.g. for expenditure variables, as wealthier countries spend exponentially more on health or education, or inversely for inflation. For global country data, strongly right- or left-tailed distributions are very common due to the gap between e.g. developed and developing countries [12]. I will use the `log1p` function from Numpy, which adds 1 to the data before taking the logarithm to avoid issues with zero values.

Another important aspect in this step is to invert variables where higher values are better (e.g. corruption control, health expenditure, literacy rate), so that the index is consistent in its interpretation (higher values

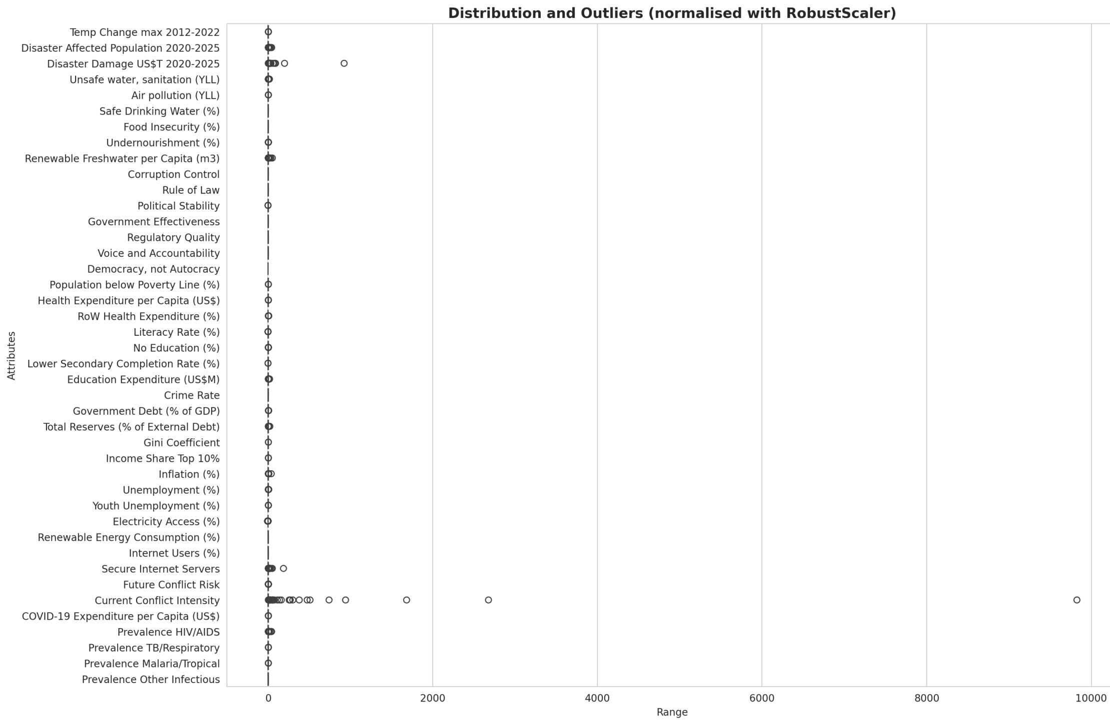


Figure 10: Boxplot of each indicator's distribution (scaled with 1.5 IQR scaling).

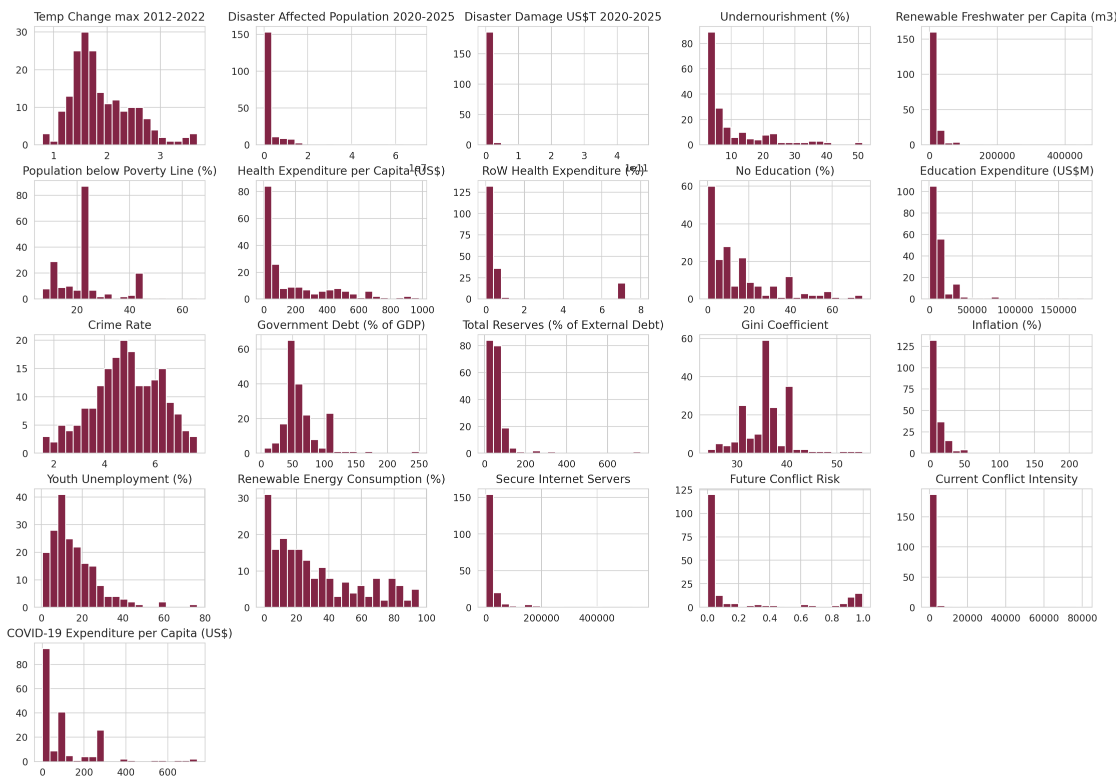


Figure 11: Histograms of all indicators (excluding PCs) before normalisation.

always indicate higher endangerment). Invert these variables before applying normalisation (but after log-transformation):

- Renewable Freshwater per Capita (m3)
- Renewable Energy Consumption (%)
- Total Reserves (% of External Debt)
- Secure Internet Servers
- Education Expenditure (US$M)
- Health Expenditure per Capita (US$)
- COVID-19 Expenditure per Capita (US$)

I tried the idea of scaling some variables with a Min-Max scaler, since they are already bounded (e.g. percentages). But this creates discrepancies with others that are scaled with z-score, and the histograms are the same when using z-score on these. So I scale all indicators using z-score normalisation (StandardScaler). Checking the histograms as they are shown in Figure 12, the distributions look less skewed. The variances are within range, all data is on the same scale and can now be weighted and aggregated to form the composite index.

Figure 12: Histograms of all indicators (excluding PCs) after normalisation.

Principal components are not re-scaled in this step because they were already normalised for the PCA. They show up in the variance check as having potentially too high/low variance (for earlier/later PCs respectively), but this is actually good since they represent multiple indicators in decreasing order of explanatory power, so different variance already provides their final weighting (see section 6).

6 Weighting and Aggregation

Within each sub-indicator, the weights of variables that pose more imminent and severe dangers to society are increased according to the theoretical framework (see section 1). These are:

- Temp Change max 2012-2022: a direct measure of climate change, which has strong and direct impact on health and well-being today.
- Disaster Affected Population 2020-2025: a direct measure of the impact of natural disasters on populations
- Population below Poverty Line (%): the most imminent social vulnerability indicator out of the ones included in this category.
- Secure Internet Servers: The selected economic indicators are all important, but this one is the most relevant to the current global situation. It is a direct measure of a country's ability to protect its digital infrastructure and citizens from cyber threats, which are becoming increasingly common and severe.
- Current Conflict Intensity: Most increased in the threat category as it is always connected with violence and death, which is the most severe form of endangerment.
- Future Conflict Risk and COVID-19 Expenditure per Capita (US$): Increased because they are more topical than disease prevalence (the remaining PCA group in this sub-indicator).

The weights for each sub-indicator are discussed in the theoretical framework section. They are based on expert opinion and research. Principal components will receive equal weighting among each other, since they are already weighted by their inherent variance. The full list of weights can be found in notebook 6.1.

For Aggregation, there are two main methods: arithmetic mean and geometric mean. The arithmetic mean is more intuitive and easier to understand, but the geometric mean is more robust to outliers and skewed distributions. Additionally, it outputs positive values, which is preferable for the index as endangerment should not have negative values. For geometric mean, another preparation step is needed, as it cannot handle negative values. To solve this, the formula $\mathrm{df}_{shifted} = \mathrm{df} - \mathrm{df.min}() + 1$ is applied. After aggregating the indicators into sub-indicators, these are normalised again (z-score) to ensure that they are on the same scale before being combined into the composite index. I found that the geometric mean condenses the index values too much, so I will use the arithmetic mean instead. It works well with standardised data, which are present here.

All steps in order:

1. Shift the data to only have positive values.
2. Aggregate the indicators into sub-indicators using the weighted geometric mean.
3. Normalise the sub-indicators using z-score normalisation.
4. Aggregate the sub-indicators into the composite index using the weighted geometric mean.
5. Shift the values by the overall minimum to be positive only
6. Multiply by a chosen constant to get a more readable index value (roughly 0-100).

As a final step, I group index ranges to 5 categories (like [2]). This is done with the Jenks natural breaks method [19]. I call them risk groups, with values "Very Low", "Low", "Medium", "High", "Very High". The distribution is shown in Figure 13.

7 Link to other Indicators

There are many other indices that measure similar concepts, often with a specific focus or primary dimension. I will compare the GSEI with some of these indices to see how close they are. The results are shown in Table 3

Table 3: Comparison to other indices using correlations. "Sub." refers to a sub-indicator correlation with the entire other index.

Other Index	Score Corr.	Ranking Corr.	Max Sub. Corr.	Min Sub. Corr.
World Risk Index 2024	0.3	0.35	threat (0.35)	eco (0.14)
IRC Watchlist 2025	–	0.17	pol (0.4)	threat (0.01)
INFORM Risk Index 2025	0.87	0.83	threat (0.86)	eco (0.37)

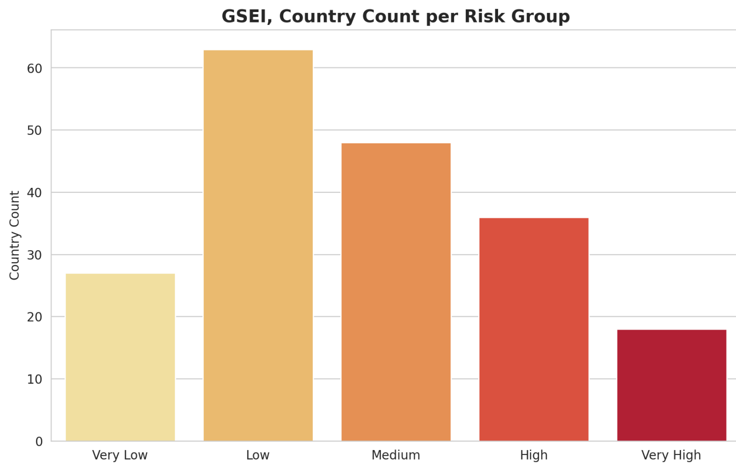


Figure 13

7.1 World Risk Index

The World Risk Index (WRI) measures the vulnerability of countries to natural disasters [20]. It correlates with GSEI by 0.3, which is moderately low (see Table 3). The ordinal ranking correlation is 0.35, which is a bit higher, but still low. This is expected, as the WRI has a narrower focus, and different indicators and methodology.

7.2 IRC Emergency Watchlist

The International Rescue Committee (IRC) is an NGO that deploys humanitarian help worldwide [21]. It publishes an annual watchlist with the most critical countries to support [1]. This list inspired the development of my index. The watchlist lays a strong focus on conflicts and diplomacy, but incorporates all 5 categories of the GSEI to an extent. Unfortunately, it is not a complete global index. It only consists of the 20 countries identified as the most vulnerable, and only the first 10 are in order. This limits the comparability severely.

The results (see Table 3) show only a 0.17 correlation with the GSEI, which is very low. This is to be expected, as the number of samples is just 18 in this case (GSEI is missing 2 countries that are on the watchlist). However, another finding was that the political dimension of GSEI correlates with the Watchlist by 0.4, which is moderate but much higher than 0.17. This suggests that the political dimension can better predict of the countries that are on the Watchlist, which matches the general alignment of the IRC. Political data possibly have the highest weight in the Watchlist, since they match the GSEI better than the rest of the dimensions.

7.3 INFORM Risk Index

The INFORM Risk index is developed by the European Commission [2]. Driven by the measurement of susceptibility to natural disasters, this should be the closest comparison to GSEI, as it is a very comprehensive index that factors in many of the variables also used in the GSEI. The comparison (see Table 3) shows that the correlation is very high (0.87 on scores and 0.83 on ordinal ranking). A good visualisation of this correlation can be seen in Figure 14.

Comparing the world map visualisations of INFORM Risk (Figure 15) and GSEI (Figure 16), they are quite similar, but with a few distinct exceptions. To pick a well-known example of a more recent, European situation: In the INFORM map, Ukraine and Russia are on a similar risk level. This is probably due to the primarily environmental scoring of INFORM, but with conflict intensity and humanitarian crises in mind, Ukraine should be shown to be more endangered. In my index and map, Ukraine is way higher up, so the GSEI successfully captures the other dimensions I added and weighted to achieve this.

This shows that the GSEI is a good representation of the risks and threats that countries face, as I consider the INFORM Risk Index to be an excellent and highly respectable piece of work. I will conclude from this: The many hours of work I put into thoroughly building the index and exploring the statistical implications of my applied methods were worth it, even beyond the big learning effects for myself.

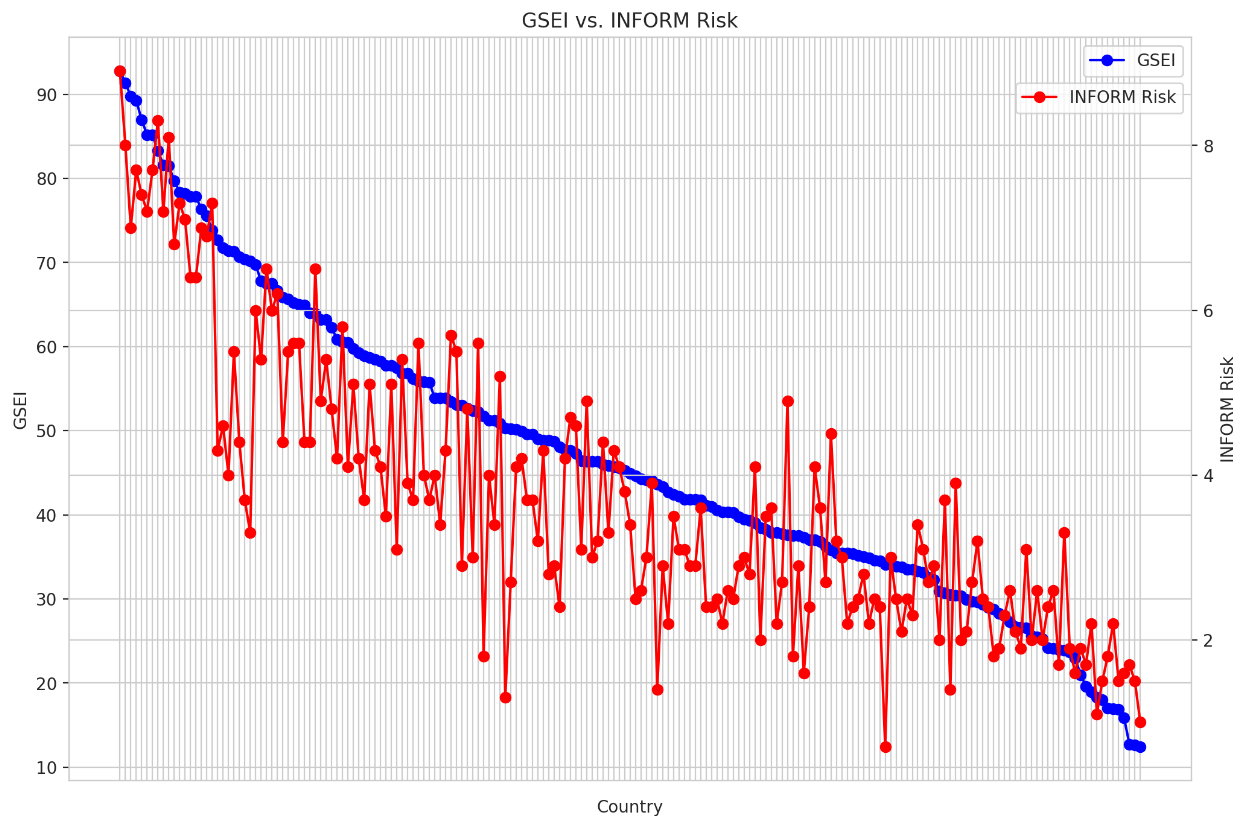


Figure 14: Line plot of sorted GSEI values (blue), with the corresponding INFORM Risk Index value for each point (red).

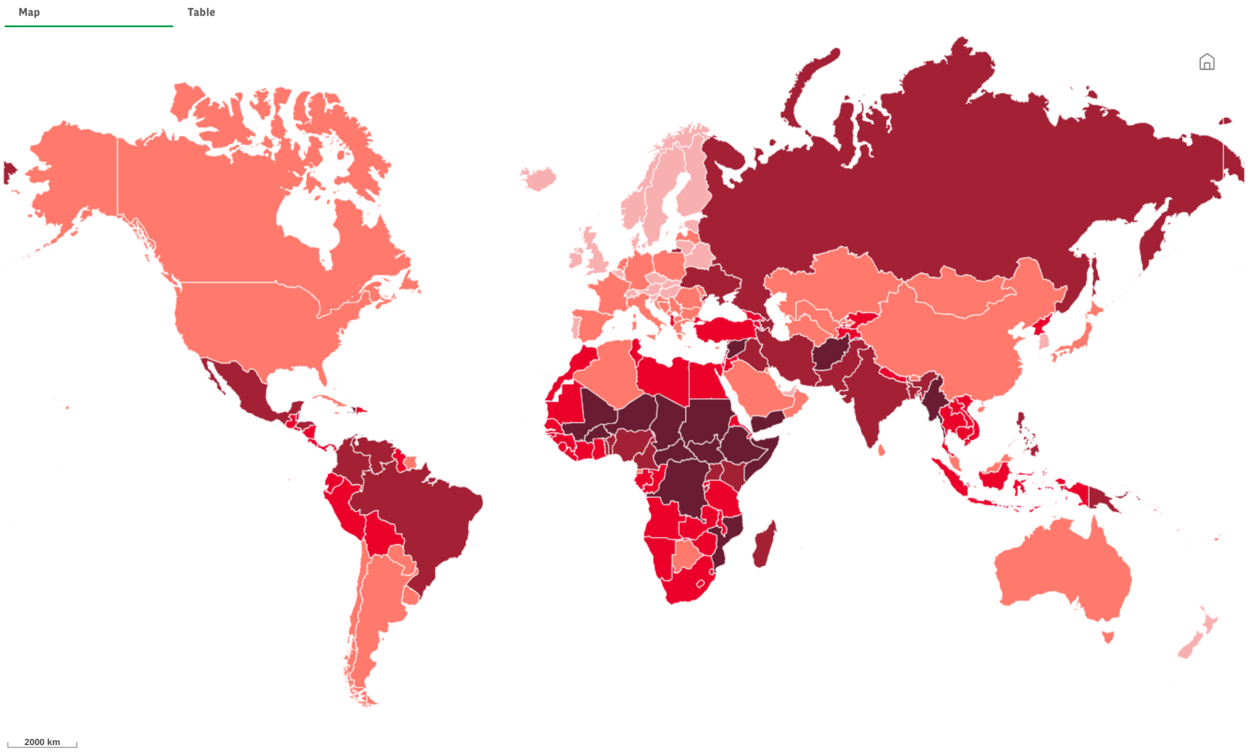


Figure 15: INFORM Risk Index world map, colour mapped from highest scores (dark red) to lowest (light red).

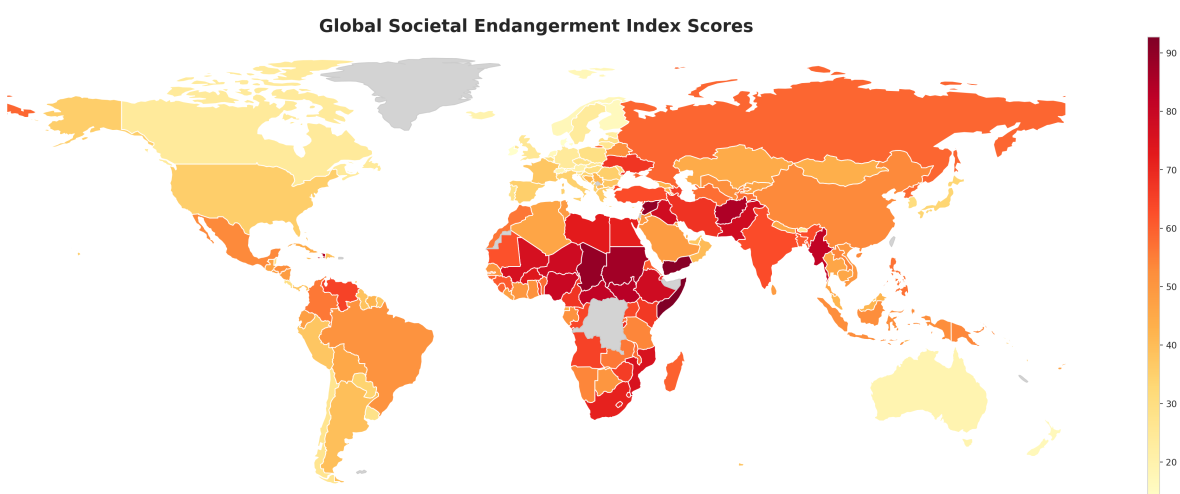


Figure 16: World map showing the country scores (as colours) of the Global Societal Endangerment Index.

A Expert Opinion

Salah Alnachawati, political scientist and strategy expert.

The government is the primary entity responsible for implementing decisions, adhering to recommendations, and securing public confidence. A government that does not function properly or lacks the trust of its citizens is incapable of addressing risks effectively. Instead, it becomes one of the most significant threats to the very nation it governs. Consequently, governance holds a weight of 28% in the risk index.

The economy represents a fundamental pillar of governance and serves as a primary metric for assessing the success of government policies. It directly influences the lives of the governed, as a successful government is capable of fostering economic growth, which in turn stabilizes the dynamics between the rulers and the ruled. Economic stability prevents external hostile interventions and mitigates the risk of social unrest. For these reasons, the economy ranks second, accounting for 23% of the risk index.

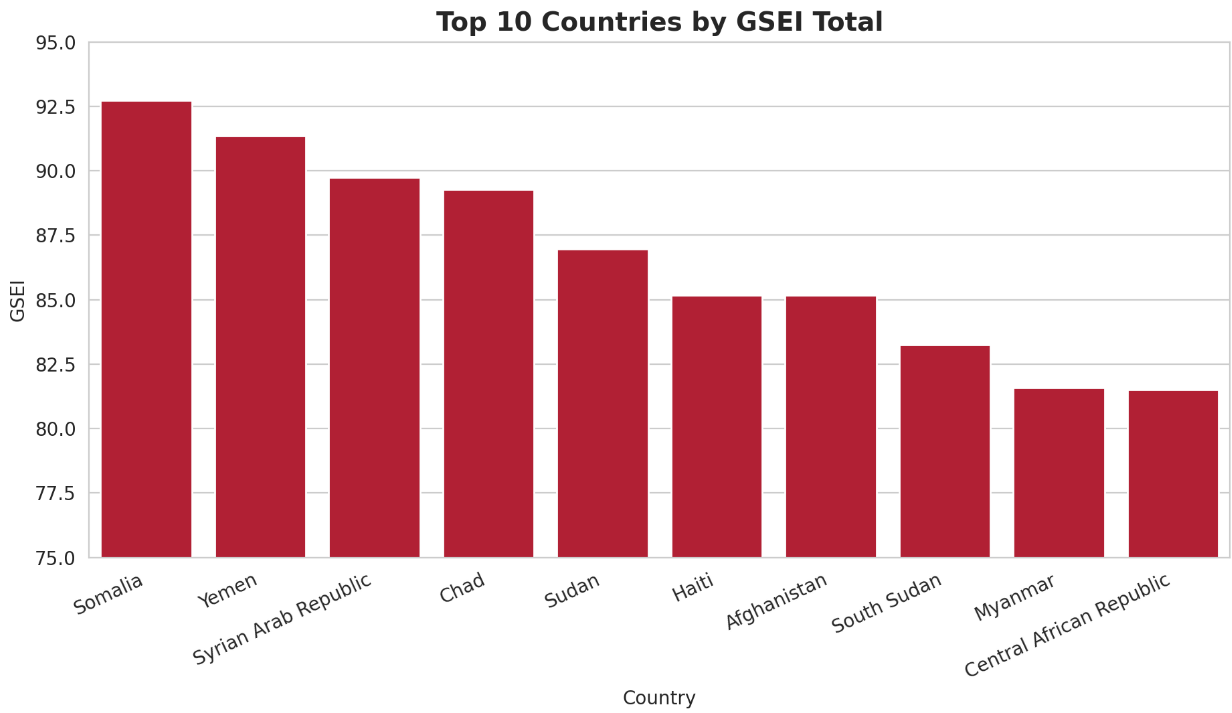
Closely following the economy are geopolitical tensions in the country's surrounding region, which account for 22% of the index. Neighbouring conflicts often translate into geopolitical crises, leading to the deprivation of crucial resources, partnerships, or geographic advantages. These factors negatively impact economic growth, governance stability, and public confidence. Such circumstances may drive the government in one of two directions: either externalizing the crisis, potentially escalating geopolitical tensions into military confrontations, or intensifying domestic polarization between supporters and opponents of the government's regional policies.

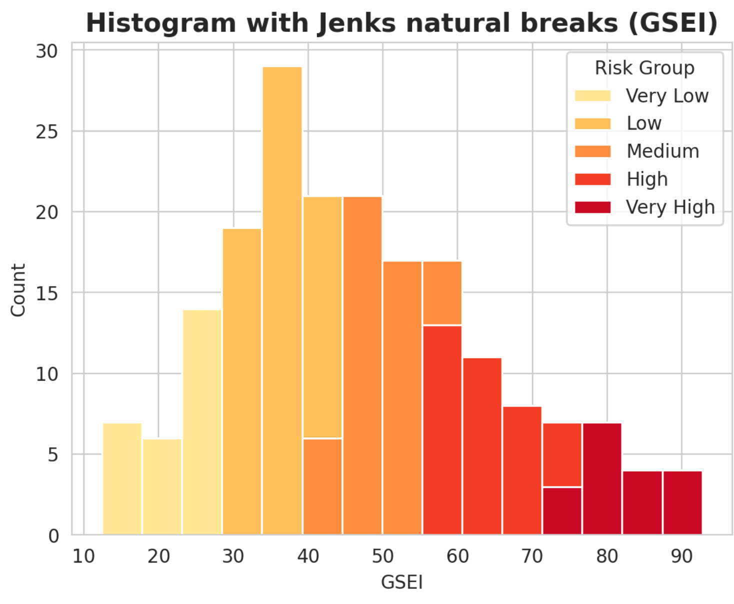
Social factors contribute approximately 15% to the risk index. Social crises are often the result of the previously mentioned high-impact factors rather than independent causes. For instance, societal polarization tends to increase when a community faces the repercussions of specific government policies, such as adopting a hostile stance toward a neighbouring country or allowing an uncontrolled influx of refugees without a broadly accepted and justified rationale.

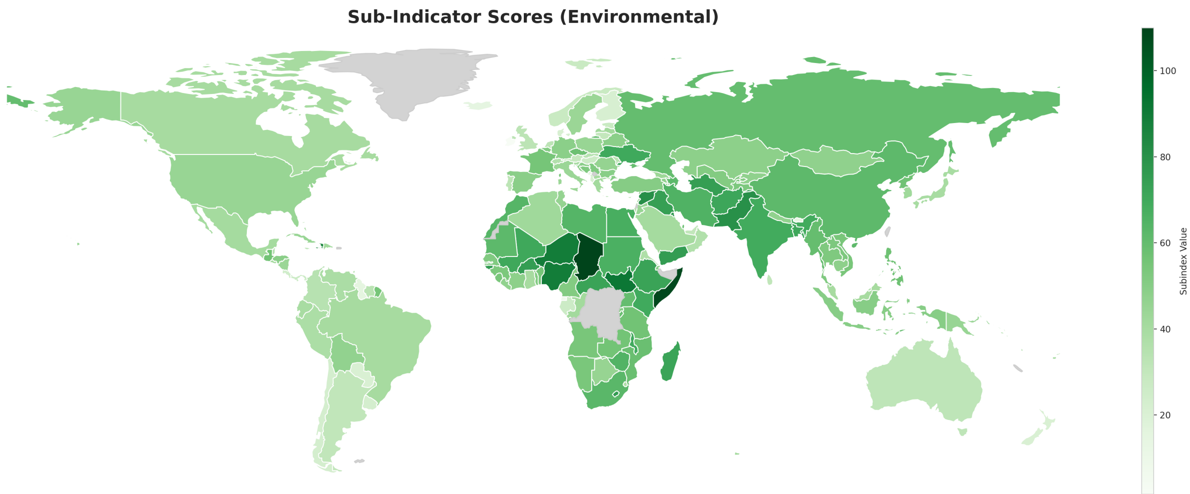
Despite its significance, environmental factors account for only 12% of the risk index. This is due to the inherent difficulty for a single government to address global environmental challenges such as climate change. Additionally, mitigating natural disasters often requires resources that only highly affluent and technologically advanced states can afford. The unpredictability of natural disasters further complicates preparedness efforts. Moreover, rallying domestic public opinion in favour of international cooperation for environmental protection can be challenging, particularly when such initiatives entail increased taxation or sovereign debt. While issues such as access to clean drinking water or the outbreak of infectious diseases undoubtedly constitute national crises, they may be more appropriately categorized under separate indicators, given the differing temporal scales of their occurrence and response mechanisms.

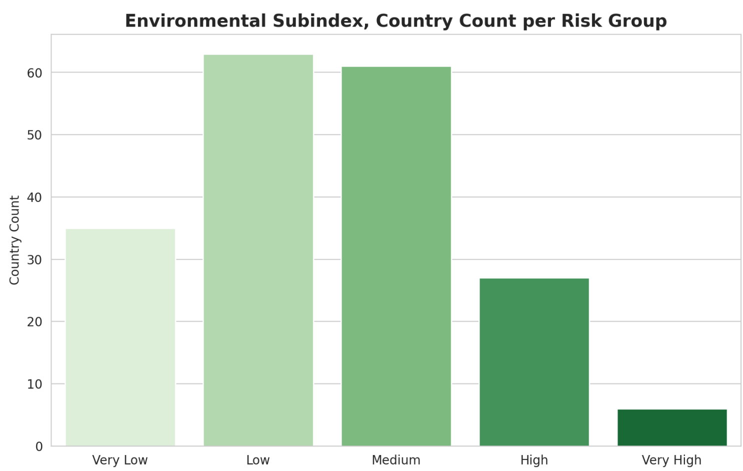
B Visualisations

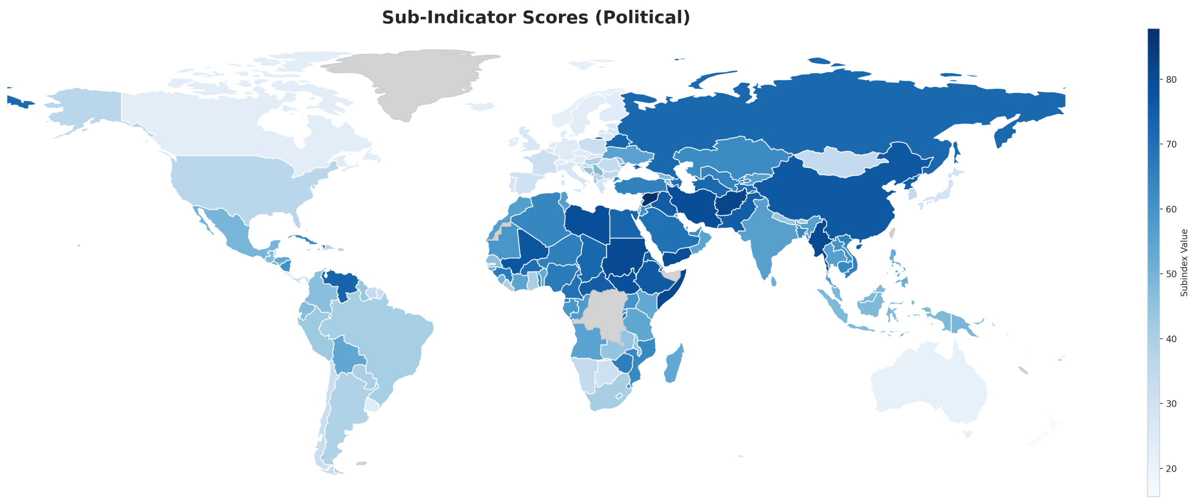
For completeness of this technical report, the generated visualisations are appended here. The structure of this report and the visualisations is in part inspired by [22].

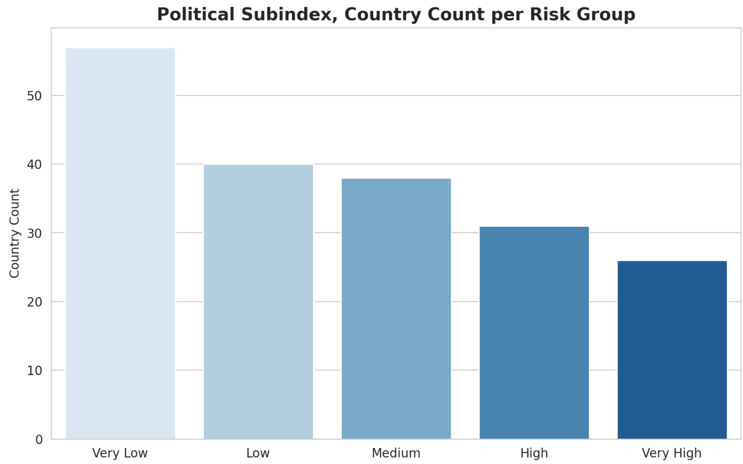


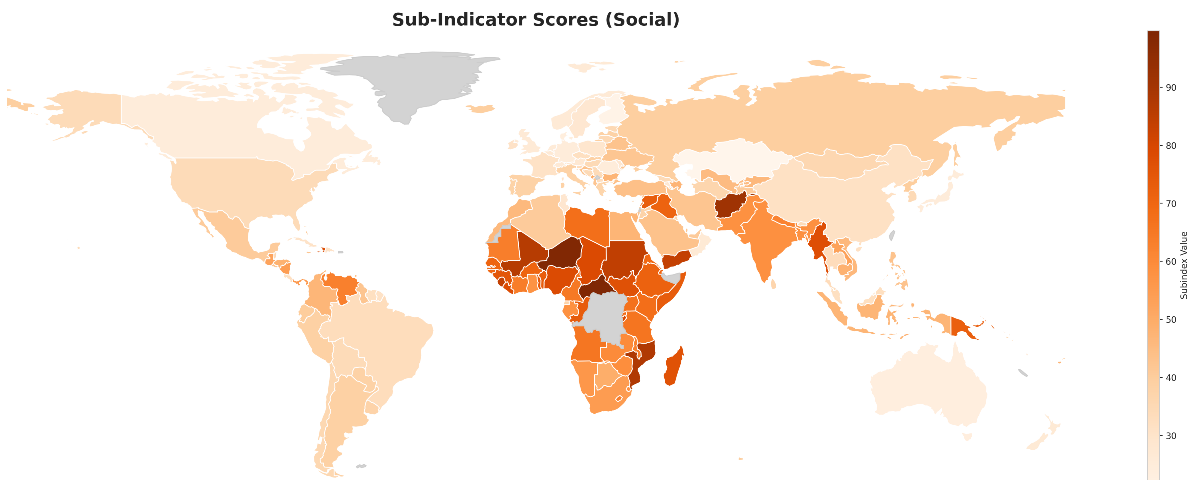


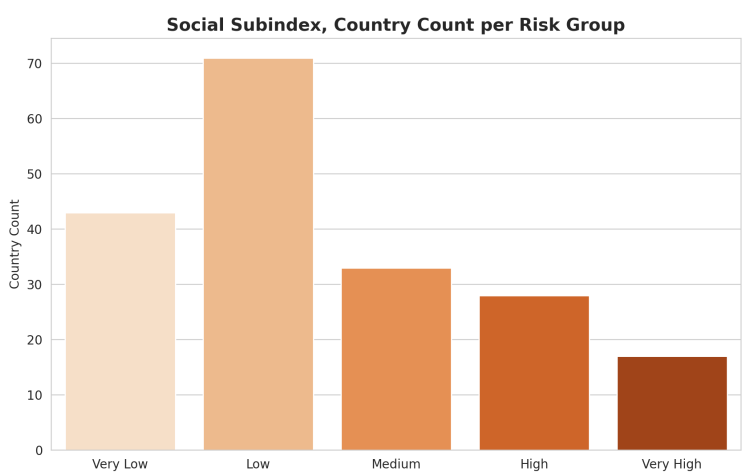

Sub-Indicator Scores (Environmental)
Subindex Value
100
80
60
40
20

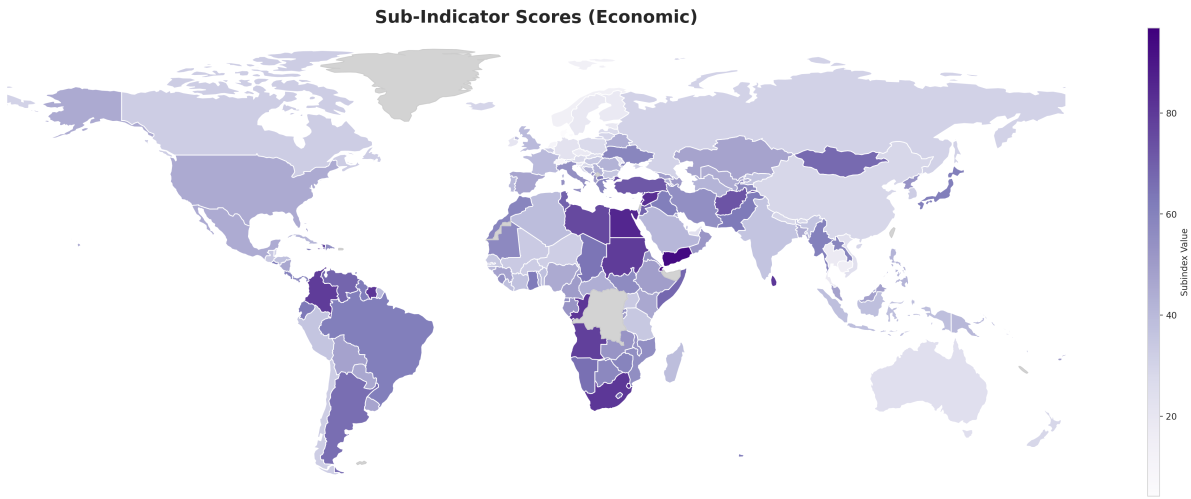

Environmental Subindex, Country Count per Risk Group
Country Count
60
50
40
30
20
10
0
Very Low
Low
Medium
High
Very High

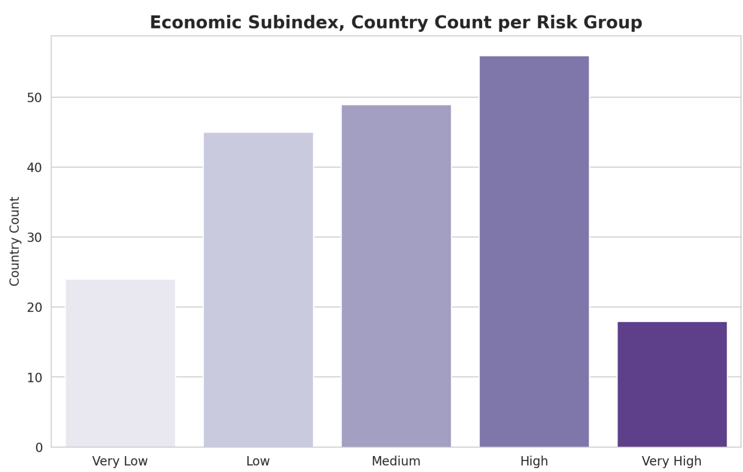

Sub-Indicator Scores (Political)
Subindex Value
80
70
60
50
40
30
20

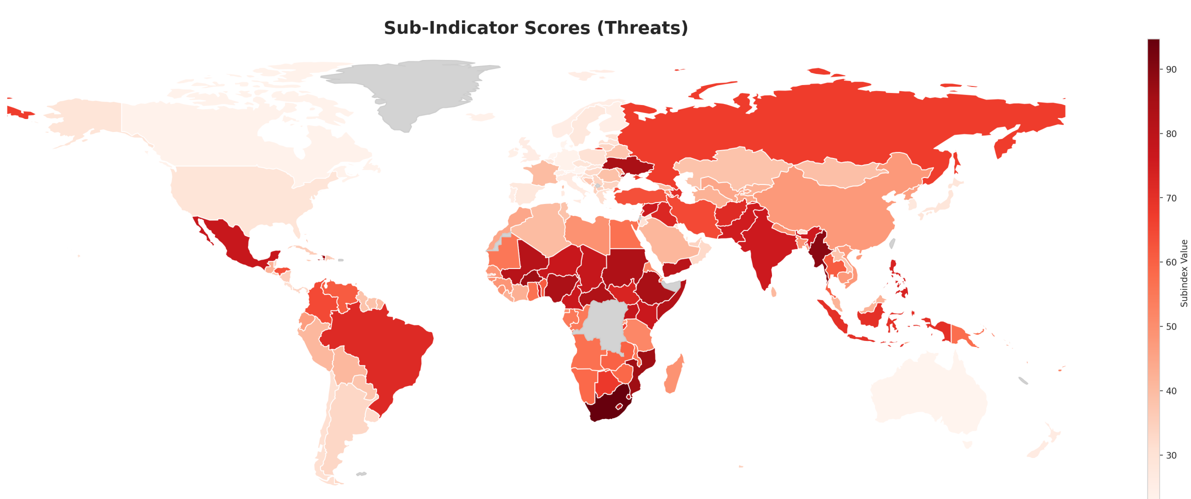

Political Subindex, Country Count per Risk Group
Country Count
0
10
20
30
40
50
Very Low
Low
Medium
High
Very High

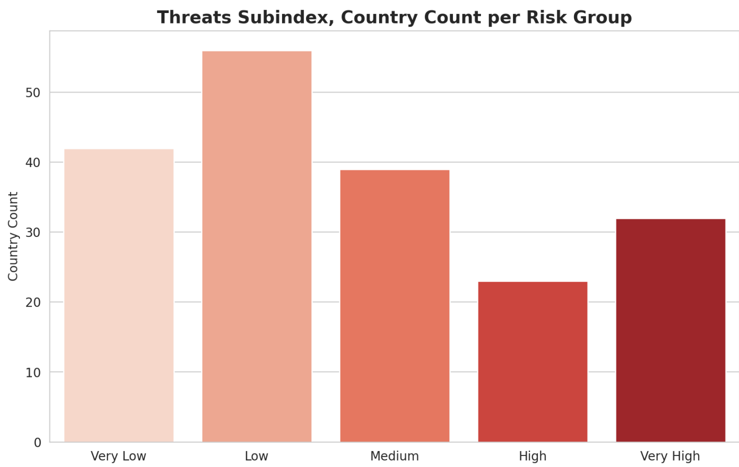

Sub-Indicator Scores (Social)
Subindex Value
20
30
40
50
60
70
80
90


Social Subindex, Country Count per Risk Group
Country Count
0
10
20
30
40
50
60
70
Very Low
Low
Medium
High
Very High


Sub-Indicator Scores (Economic)
80
60
40
20
Subindex Value


Economic Subindex, Country Count per Risk Group
Country Count
0
10
20
30
40
50
Very Low
Low
Medium
High
Very High


Sub-Indicator Scores (Threats)
90
80
70
60
50
40
30
Subindex Value


Threats Subindex, Country Count per Risk Group
Country Count
0
10
20
30
40
50
Very Low
Low
Medium
High
Very High

References

[1] International Rescue Committee (IRC). *2025 Emergency Watchlist*. URL: https://www.rescue.org/eu/watchlist2025 (visited on 03/30/2025).

[2] European Commission DRMKC Disaster Risk Management Knowledge Centre. *INFORM Risk Index*. 2025. URL: https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Risk (visited on 03/30/2025).

[3] Tarun Mugesh. *Climate change Indicators*. 2022. URL: https://www.kaggle.com/datasets/tarunrm09/climate-change-indicators (visited on 03/30/2025).

[4] Centre for Research on the Epidemiology of Disasters (CRED) and University of Louvain (UCLouvain). *The International Disaster Database*. 2025. URL: https://public.emdat.be/ (visited on 03/30/2025).

[5] Institute for Health Metrics and Evaluation (IHME). *Global Burden of Disease (GBD) study*. 2021. URL: https://vizhub.healthdata.org/gbd-results/ (visited on 03/30/2025).

[6] World Bank Group. *World Development Indicators*. 2023. URL: https://databank.worldbank.org/source/world-development-indicators (visited on 03/30/2025).

[7] World Bank Group. *Worldwide Governance Indicators*. 2023. URL: https://databank.worldbank.org/source/worldwide-governance-indicators (visited on 03/30/2025).

[8] Varieties of Democracy (V-Dem). *The Episodes of Regime Transformation (ERT) dataset*. 2023. URL: https://www.v-dem.net/data/ert-dataset/ (visited on 03/30/2025).

[9] World Health Organization (WHO). *Global Health Expenditure Database*. 2022. URL: https://apps.who.int/nha/database (visited on 03/30/2025).

[10] World Bank Group. *Education Statistics - All Indicators*. 2023. URL: https://databank.worldbank.org/source/education-statistics-%5e-all-indicators (visited on 03/30/2025).

[11] Global Initiative Against Transnational Organized Crime (GI-TOC). *The Global Organized Crime Index report*. 2023. URL: https://ocindex.net/downloads (visited on 03/30/2025).

[12] Nardo M et al. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. 2nd ed. Paris (France): OECD publishing, 2008. ISBN: 978-92-64-04345-9. URL: www.oecd.org/publishing,http://213.253.134.43/oecd/pdfs/browseit/3008251E.PDF.

[13] Simon Fraser University. *Chapter 8: Multicollinearity*. 2016. URL: https://www.sfu.ca/~dsignori/buec333/lecture%2016.pdf (visited on 03/30/2025).

[14] Carsten F. Dormann et al. "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance". In: *Ecography* 36 (Jan. 2013), pp. 27–46. DOI: 10.1111/j.1600-0587.2012.07348.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1600-0587.2012.07348.x (visited on 03/30/2025).

[15] European Commission and Joint Research Centre. *The Global Conflict Risk Index 2022: revised data and methods*. eng. LU: Publications Office, 2022. DOI: 10.2760/041759. URL: https://drmkc.jrc.ec.europa.eu/initiatives-services/global-conflict-risk-index\#documents/1435/list (visited on 03/30/2025).

[16] IBM. *What is principle component analysis (PCA)?* 2023. URL: https://www.ibm.com/think/topics/principal-component-analysis (visited on 03/30/2025).

[17] Sai Praneeth. *Customer Segmentation: Kmeans Clustering*. Aug. 2021. URL: https://medium.com/analytics-vidhya/customer-segmentation-kmeans-clustering-7670e317b54c (visited on 03/26/2025).

[18] Robert M West. "Best practice in statistics: The use of log transformation". In: *Annals of Clinical Biochemistry* 59.3 (May 2022), pp. 162–165. DOI: 10.1177/00045632211050531. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9036143/ (visited on 03/26/2025).

[19] Rehan Ahmad. *Jenks Natural Breaks — Best range finder algorithm*. Aug. 2019. URL: https://medium.com/analytics-vidhya/jenks-natural-breaks-best-range-finder-algorithm-8d1907192051 (visited on 03/30/2025).

[20] Institute for International Law of Peace and Armed Conflict (IFHV) and United Nations. *WorldRiskIndex*. 2024. URL: https://data.humdata.org/dataset/worldriskindex (visited on 03/30/2025).

[21] International Rescue Committee (IRC). *Who we are*. URL: https://www.rescue.org/eu/who-we-are (visited on 03/30/2025).

[22] Economist Intelligence Unit and Siemens. *European Green City Index*. 2009. URL: https://assets.new.siemens.com/siemens/assets/api/uuid:fddc99e7-5907-49aa-92c4-610c0801659e/european-green-city-index.pdf (visited on 03/30/2025).