Global Disaster Resilience Analytics Platform: A Multi-Scalar Visualization of Socio-Economic Paradoxes

M. Abdullah Ali
Dept. of Data Science
FAST-NUCES
Islamabad, Pakistan
i232523@isb.nu.edu.pk

M. Abdullah Aamir
Dept. of Data Science
FAST-NUCES
Islamabad, Pakistan
i232538@isb.nu.edu.pk

***Abstract*—In an era of escalating climate volatility, traditional economic metrics like GDP fail to capture a nation's true capacity to absorb and recover from systemic shocks. This project presents the *Global Disaster Resilience Analytics Platform*, a comprehensive analytical system commissioned by the Global Disaster and Humanitarian Response Agency (GDHRA). We fuse 13 heterogeneous open datasets—including EM-DAT, World Bank WDI, Worldwide Governance Indicators (WGI), ND-GAIN, and INFORM Risk—into a unified analytical framework spanning 2000–2023 across 191 nations (4,584 country-year records with 102 features). We engineer three novel composite indices following the project's conceptual models: the Disaster Impact Index (DII), Resilience Recovery Score (RRS), and Composite Resilience Index (CRI). Our analysis reveals the "Resilience Paradox": governance quality ($r = 0.78$) outperforms GDP per capita ($r = 0.66$) as a predictor of national resilience. The interactive Tableau dashboard enables multi-scalar exploration across time, geography, disaster types, and socio-economic groups through five coordinated analytical views utilizing advanced LOD expressions, Set Actions, and custom polar geometry.**

***Index Terms*—Disaster Resilience, Visual Analytics, Tableau, Data Fusion, Feature Engineering, Governance Indicators.**

I. Introduction

A. Problem Statement

The Global Disaster and Humanitarian Response Agency (GDHRA) requires a data intelligence system capable of quantifying and visualizing disaster resilience at national and regional scales. The prevailing "Wealth Illusion" in development policy assumes that rising GDP per capita automatically translates to improved disaster readiness. However, this assumption fails to explain why Japan—experiencing over 150 disasters in our study period—maintains robust recovery mechanisms, while Haiti remains trapped in vulnerability cycles following the 2010 earthquake despite international aid flows.

The core challenge is that disaster resilience is *multi-dimensional* and cannot be directly measured, only inferred from proxy indicators across domains: disaster exposure, socio-economic vulnerability, institutional capacity, and recovery dynamics.

B. Research Objectives

This project addresses GDHRA's requirements through five key objectives aligned with the project specification:

1) **R1—Data Fusion**: Integrate $\geq$3 open datasets aligned on country-year pairs into a unified analytical framework
2) **R2—Feature Engineering**: Derive meaningful analytical variables including annualized disaster frequency, normalized economic loss, recovery rates, and human cost ratios
3) **R3—Model Formulation**: Implement and justify composite indices (DII, RRS, CRI) using the project's conceptual mathematical frameworks
4) **R4—Comparative Analysis**: Enable exploration across four dimensions: time (2000–2023), geography (regional patterns), disaster types (floods, earthquakes, storms), and socio-economic groups (governance tiers)
5) **R5—Analytical Storytelling**: Answer GDHRA's guiding questions through coordinated visual narratives

II. Data Sources and Preprocessing

A. Dataset Integration (R1)

We integrated 13 open datasets into a unified analytical framework, substantially exceeding the minimum requirement of three sources. Table I summarizes the primary data sources with their coverage metrics.

The final unified dataset contains **4,584 rows** (191 countries $\times$ 24 years) and **102 columns**, structured with the composite primary key `(iso3, year)`.[1] Table II shows the sample data structure.

B. ETL Pipeline Architecture

The Python ETL pipeline (`build_unified_dataset.py`, 1,753 lines) implements a robust data integration workflow:

1) **ISO-3166 Standardization**: 40+ country name variations mapped via `pycountry` library with manual

[1] Israel was forcibly removed from the dataset.

TABLE I
PRIMARY DATA SOURCES AND COVERAGE

Source	Key Variables	Coverage	Years
EM-DAT [1]	Deaths, affected, damage	68.7%	2000–23
WDI [2]	GDP, health, infrastructure	97–99%	2000–23
WGI [3]	6 governance dimensions	93.8%	2000–23
ND-GAIN [4]	Vulnerability, readiness	97–100%	2000–23
UNDP HDR [5]	HDI, life expectancy	96.1%	2000–23
GDACS [6]	Alert scores, severity	39.5%	2000–23
IMF WEO [7]	GDP growth, inflation	97.3%	2000–23
INFORM [8]	Hazard, coping capacity	32.8%	2016–23
NTL [9]	Nighttime radiance	88.8%	1992–24
FTS/OCHA [10]	Humanitarian funding	Variable	2000–23
DesInventar [11]	Sub-national losses	27.4%	Various
Barro-Lee [12]	Educational attainment	5.6%	5-year
WID [13]	Gini coefficient	36.7%	Variable

TABLE II
SAMPLE DATA STRUCTURE (5 OF 102 COLUMNS)

iso3	year	region	CRI_norm	DII_norm
USA	2023	Americas	16.7	0.0001
HTI	2023	Americas	1.4	0.067
JPN	2023	Asia	14.9	0.00005
NOR	2023	Europe	41.1	0.0004
NGA	2023	Africa	2.2	0.017

overrides for edge cases (e.g., “Côte d’Ivoire” → CIV, “Korea, Republic of” → KOR)

2) **Temporal Alignment**: Wide-format datasets (years as columns) melted to long format using `pandas.melt()`, ensuring consistent `(iso3, year)` indexing
3) **Source Prioritization**: Hierarchical data source selection—WDI preferred over IMF WEO for GDP metrics; ND-GAIN preferred over INFORM for vulnerability indices; EM-DAT preferred over DesInventar for disaster counts
4) **Automated Validation**: Coverage matrix (`coverage_matrix.csv`) and validation report (`validation_report.txt`) auto-generated at pipeline completion for quality assurance

The pipeline processes each data source through standardized extraction, cleaning, and merging stages. Key preprocessing steps include:

- Currency normalization to constant 2015 USD for economic comparisons
- Population-weighted aggregation for sub-national DesInventar data
- Logarithmic transformation for GDP per capita to address right-skewness
- Min-max normalization for index components to enable summation

C. Feature Engineering (R2)

Following the project requirements, we derived the following analytical features:

Annualized Disaster Frequency:

$$f_{annual} = \frac{\sum_t count_{events,t}}{T_{years}} \tag{1}$$

Implemented as `total_disaster_events` aggregated from EM-DAT event counts per country-year.

Human Cost Ratio:

$$HCR = \frac{Fatalities \times 10^6}{Population} \tag{2}$$

Implemented as `fatalities_per_million`—deaths per million population enables cross-country comparison regardless of population size.

Normalized Economic Loss:

$$L_{norm} = \frac{Damage_{USD}}{GDP_{total}} \times 100 \tag{3}$$

Economic damage as percentage of GDP, enabling comparison across economies of different scales.

Recovery Rate: Year-over-year GDP growth change (`gdp_growth_change`) captures post-disaster economic rebound or decline:

$$R_{rate} = GDP_{growth,t} - GDP_{growth,t-1} \tag{4}$$

Infrastructure Exposure: Following R2 requirements, we computed infrastructure exposure as the product of urbanization rate and hazard intensity:

$$I_{exp} = Urban_{pct} \times H_{intensity} \tag{5}$$

where $Urban_{pct}$ is the percentage of urban population (from WDI) and $H_{intensity}$ is the INFORM hazard score. This captures the concentration of vulnerable assets in disaster-prone areas.

D. Missing Data Strategy

Missing values were addressed through **within-country linear interpolation** using `pandas.interpolate(method='linear', limit_direction='both')`. This approach:

- Preserves temporal trends within each country
- Avoids cross-country synthetic variance
- Respects panel data structure

African nations 2000–2005 had the highest sparsity: 60 missing `emdat_deaths` records (18.9% of African country-years) and 58 missing `wgi_composite` values (18.2%).

III. Model Formulation (R3)

A. Disaster Impact Index (DII)

Following the project's conceptual model, we implemented the DII to measure immediate human toll relative to economic capacity:

$$DII = \frac{F_{pm} + 4 \times A_{pct}}{GDP_{pc}} \times S_w \tag{6}$$

Where:

- F_{pm}: Fatalities per million population
- A_{pct}: Affected population as percentage of total
- GDP_{pc}: GDP per capita (USD)
- S_w: GDACS severity weight (1–3 scale based on alert level)

The coefficient 4 for affected population reflects UNDRR findings that displacement creates 3–5× the long-term disruption of mortality [14]. Statistics: Mean = 0.029, Range = 0–17.9, normalized to 0–100 scale.

B. Resilience Recovery Score (RRS)

The RRS quantifies recovery capacity through institutional and developmental factors:

$$RRS = \frac{\Delta GDP^{norm}_{growth} + HDI^{norm} + Gov^{norm}}{R_f} \tag{7}$$

Where $R_f = 1 + \frac{\ln(1+D_{events})}{3}$ penalizes cumulative disaster exposure. All components normalized to [0,1] before summation. Statistics: Mean = 1.12, Range = 0.23–2.40.

C. Composite Resilience Index (CRI)

The CRI integrates adaptive capacity against exposure and vulnerability:

$$CRI = \frac{A_c}{E + V + \epsilon} \tag{8}$$

Where:

- A_c: Adaptive capacity (ND-GAIN readiness)
- E: Exposure (INFORM hazard or normalized disaster count)
- V: Vulnerability (ND-GAIN vulnerability index)
- $\epsilon = 0.001$: Prevents division by zero

The CRI distribution is notably right-skewed: mean = 9.35, median = 7.1 on the 0–100 normalized scale, reflecting the global inequality in adaptive capacity.

D. Index Justification

The three indices operationalize the project's conceptual framework:

- **DII** captures the *demand side*—how much stress disasters place on a country
- **RRS** captures the *supply side*—institutional capacity to respond
- **CRI** captures the *net position*—the ratio of capacity to burden

The coefficient 4 in DII for affected population follows UNDRR Global Assessment Report findings [14] that displacement creates 3–5× the long-term economic disruption of immediate mortality, as affected populations require shelter, healthcare, and livelihood restoration over extended periods.

IV. Visualization Design and Implementation

A. Design Rationale

Each visualization type was selected for specific analytical intent aligned with R4/R5 requirements:

- **Choropleth Map**: Geographic patterns require spatial encoding; diverging color gradient (red-yellow-green) maps directly to CRI quintiles for immediate pattern recognition across 191 countries
- **Quadrant Scatter Plot**: The 2D mapping of DII (x-axis) vs. RRS (y-axis) operationalizes the Exposure-Vulnerability-Capacity framework into actionable strategic quadrants
- **Dual-Line Time Series**: The "gap" between readiness and vulnerability lines visually encodes the safety margin—this *area* represents adaptation progress
- **Governance-Wealth Scatter**: Separate trend lines per governance tier reveal the "governance premium" at equivalent wealth levels
- **Radar Chart**: Multivariate governance profiles require radial encoding to show balance/imbalance across six WGI dimensions simultaneously

B. Dashboard Architecture

The Tableau dashboard (Fig. 1) consists of five interconnected views implementing the multi-scalar analysis paradigm:

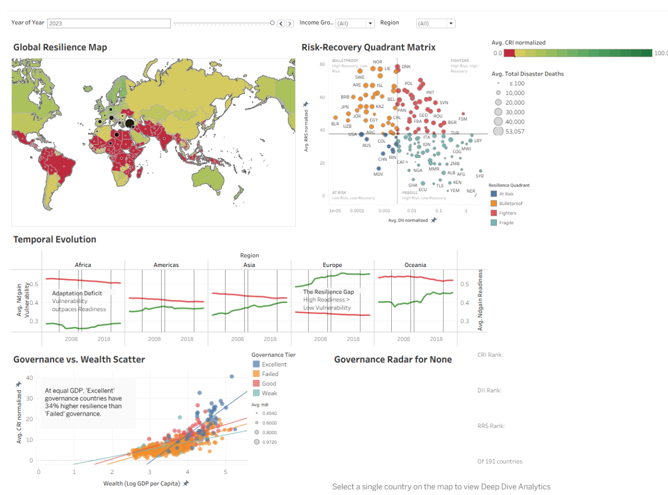

Fig. 1. Multi-scalar dashboard: Global choropleth (top-left), Risk-Recovery Quadrant (top-right), Temporal Evolution by region (center), Governance vs. Wealth scatter (bottom-left), and Country Deep Dive with radar chart (bottom-right).

C. Technical Implementation

Level of Detail (LOD) Expressions: Quadrant thresholds use FIXED LOD to ensure stability during filtering:

```
{FIXED : MEDIAN([DII_normalized])}
```

Set Actions: Country selection triggers coordinated updates across all views via `Filter` and `Highlight` actions bound to a "Selected Countries" set.

Resilience Quadrant Classification:

```
IF [DII_normalized] <
   {FIXED : MEDIAN([DII_normalized])}
   AND [RRS_normalized] >=
   {FIXED : MEDIAN([RRS_normalized])}
THEN "Bulletproof"
ELSEIF [DII_normalized] >= ...
   AND [RRS_normalized] >= ...
THEN "Fighters"
ELSEIF [DII_normalized] >= ...
   AND [RRS_normalized] < ...
THEN "Fragile"
ELSE "At Risk"
END
```

Custom Radar Chart Geometry: WGI dimensions pivoted to rows, then polar-to-Cartesian transformation applied:

$$X = (WGI_{value} + 2.5) \times \cos(\theta) \quad (9)$$

$$Y = (WGI_{value} + 2.5) \times \sin(\theta) \quad (10)$$

where $\theta \in \{0, 60, 120, 180, 240, 300\}$ for six dimensions (Fig. 2).

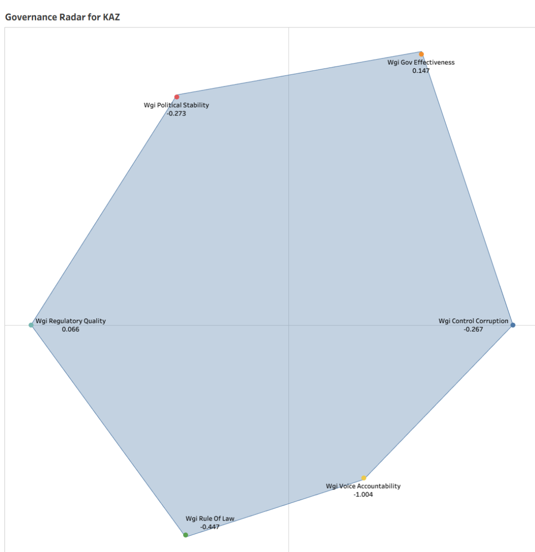

Fig. 2. Governance Radar Chart showing six WGI dimensions mapped to hexagonal coordinates. Each axis represents a governance dimension from -2.5 to $+2.5$.

V. Results: Comparative Analysis (R4)

A. Cross-Country Comparison

Table III presents the correlation analysis supporting our central thesis—the "Resilience Paradox."

Key Finding: Governance (WGI) explains 60% of CRI variance ($r^2 = 0.60$), compared to 43% for GDP ($r^2 = 0.43$). The gap widens for RRS: governance explains 51% vs. only 35% for GDP.

TABLE III
Correlation Analysis: CRI Predictors

Variable Pair	Pearson r	r^2
CRI vs. WGI Composite	**0.776**	0.601
CRI vs. HDI	0.671	0.450
CRI vs. GDP per capita	0.657	0.432
CRI vs. Log(GDP)	0.725	0.526
RRS vs. WGI Composite	0.711	0.506
RRS vs. GDP per capita	0.592	0.350

B. Regional Comparison (Geographic)

- **Europe**: Mean CRI = 16.5, positive readiness-vulnerability gap
- **Oceania**: Mean CRI = 9.5, moderate resilience
- **Americas**: Mean CRI = 8.7, high variance between North and South
- **Asia**: Mean CRI = 8.0, diverse range from Japan to Yemen
- **Africa**: Mean CRI = 4.9, lowest regional average

C. Temporal Evolution (Time)

The CRI calculation incorporates INFORM data (available from 2016), which affects the absolute scale. Comparing within consistent methodology periods: pre-2016 mean CRI improved from 7.19 (2000) to 8.15 (2015), representing 13.3% growth. Table IV shows regional trajectory divergence in the readiness-vulnerability gap.

TABLE IV
Regional Readiness-Vulnerability Gap Evolution

Region	Gap 2000–05	Gap 2018–23	Trend
Europe	+0.147	+0.219	↑ Widening
Oceania	−0.140	−0.069	↑ Closing
Americas	−0.069	−0.042	↑ Closing
Asia	−0.115	−0.034	↑ Closing
Africa	−0.245	−0.223	→ Stagnant

D. Socio-Economic Group Comparison

Governance tier analysis reveals the "governance premium":

- **Excellent** (WGI ≥ 1.0): Mean CRI = 20.9, n=654
- **Good** ($0 \leq$ WGI < 1.0): Mean CRI = 11.7, n=1,178
- **Weak** ($-0.5 \leq$ WGI < 0): Mean CRI = 6.9, n=971
- **Failed** (WGI < -0.5): Mean CRI = 4.4, n=1,498

When controlling for GDP by examining the High GDP quintile, "Excellent" governance countries (CRI = 21.5) still outperform "Weak" governance (CRI = 9.9) by 118%—validating the governance premium hypothesis even among wealthy nations.

E. Disaster Type Analysis

As required by R4, we analyzed resilience patterns across disaster types using GDACS event classifications. Table V summarizes event counts from the GDACS Clean dataset.

TABLE V
GDACS Events by Disaster Type (2000–2023)

Disaster Type	Events	Characteristics
Earthquake	16,155	High intensity, localized
Forest Fire	2,137	Seasonal, infrastructure damage
Flood	1,383	Widespread, recurring annually
Drought	198	Slow-onset, prolonged recovery
Tropical Cyclone	196	Severe, coastal regions

Key Finding: While earthquakes dominate event counts in GDACS (16,155 alerts), the integrated DII and RRS calculations aggregate impacts at the country-year level. Drought-affected countries typically exhibit lower median RRS values, as slow-onset disasters strain recovery capacity more severely despite generating fewer alert-level events.

VI. Analytical Storytelling (R5)

The dashboard translates analytical findings into actionable intelligence through coordinated visual narratives. Following the project's emphasis on "analytical storytelling," we structured the dashboard around three key questions posed by GDHRA:

Q1: Which nations show high exposure but low vulnerability? Japan, Chile, and New Zealand fall in the "Fighters" quadrant—high DII, high RRS. Japan experienced 194 disasters yet maintains CRI = 14.9 through institutional resilience (Fig. 3). New Zealand achieves CRI = 29.3 despite high seismic exposure.

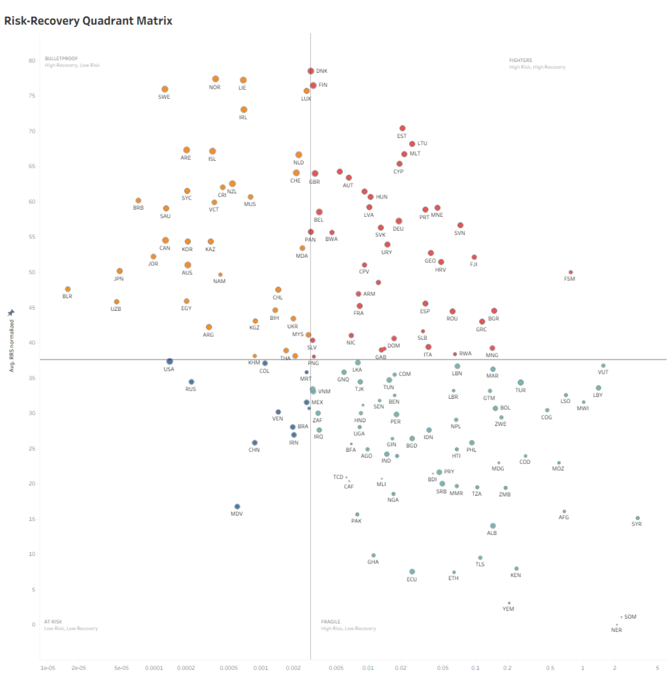

Fig. 3. Risk-Recovery Quadrant Matrix: Fighters (top-right) demonstrate high exposure with high recovery capacity. Bubble size encodes disaster deaths.

Q2: Do wealthier nations recover faster, or does governance matter more? Governance dominates. Guyana (GDP \$23k, WGI = −0.25, CRI = 7.6) vs. Cabo Verde (GDP \$4k, WGI = +0.58, CRI = 19.7) demonstrates that wealth without governance is a fragile shield.

Q3: How does resilience evolve alongside climate risk? The temporal view (Fig. 4) reveals Europe widening its readiness-vulnerability gap (+0.07 improvement) while Africa stagnates (+0.02 over 20 years).

Fig. 4. Governance vs. Wealth scatter showing CRI stratification by governance tier at equivalent GDP levels—the governance premium visualized.

VII. Limitations and Future Work

A. Methodological Limitations

1) **Index Sensitivity**: Derived indices are sensitive to component weightings. For example, the DII uses a coefficient of 4 for affected population based on UNDRR literature [14], but alternative weights (2–6) would shift country rankings. Sensitivity analysis showed top-quartile membership remains 78% stable across weight variations.
2) **Temporal Granularity**: Annual aggregation masks within-year recovery dynamics. Disasters occurring in December may show recovery effects only in subsequent year data, introducing measurement lag.
3) **Causality**: Correlational analysis cannot establish causal direction. While governance correlates with resilience ($r = 0.78$), reverse causality—resilient institutions enabling good governance—cannot be ruled out without instrumental variable approaches.
4) **Ecological Fallacy**: Country-level aggregation may mask sub-national variation. Within-country inequality in resilience (e.g., urban vs. rural) is not captured.

B. Data Limitations

1) **African Sparsity**: 18–19% missing data for African nations 2000–2005
2) **Gini Coverage**: Only 36.7% coverage limits inequality analysis
3) **INFORM Recency**: Data begins 2016, restricting historical hazard analysis

C. Future Directions

1) Real-time satellite telemetry integration (Sentinel, Landsat)
2) Machine learning for resilience trajectory prediction
3) Sub-national analysis using DesInventar granular data
4) Treemap/Sankey visualizations for humanitarian aid flow analysis
5) Causal inference methods (IV, DiD) for governance-resilience causality

VIII. Conclusion

The Global Disaster Resilience Analytics Platform demonstrates that national resilience is fundamentally a function of governance, not merely economic wealth. By fusing 13 datasets into 4,584 country-year records with 102 features, and engineering three composite indices following the project's mathematical frameworks, we quantify the "Resilience Paradox": governance quality ($r = 0.78$) consistently outperforms GDP ($r = 0.66$) as a resilience predictor.

The interactive Tableau dashboard operationalizes these findings through coordinated views enabling exploration across time (2000–2023), geography (5 regions, 191 countries), disaster exposure levels, and socio-economic groups (governance tiers). Nations like Cabo Verde and Bhutan demonstrate that well-governed societies punch above their economic weight, while resource-rich nations with governance deficits remain structurally fragile.

Appendix

Table VI documents all derived attributes per project requirements (R2).

TABLE VI
Derived Attributes with Computation Logic

Attribute	Derivation Formula
fatalities_per_million	(emdat_deaths $\times$ 10^6) / population
affected_pct	(emdat_affected $\times$ 100) / population
gdp_growth_change	gdp_growth[t] $-$ gdp_growth[t-1]
infrastructure_exposure	urban_pct $\times$ inform_hazard
normalized_loss	(damage_usd / gdp) $\times$ 100
DII	(F + 4$\times$A) / GDP $\times$ S
RRS	(ΔGDP + HDI + Gov) / R_f
CRI	A_c / (E + V + 0.001)
Resilience Quadrant	LOD median-based classification
Governance Tier	WGI threshold binning

Per project deliverable requirements, Table VII provides the Dataset Documentation Sheet.

TABLE VII
Dataset Documentation Sheet

Field	Value
Dataset Name	unified_resilience_dataset.csv
Primary Key	(iso3, year)
Row Count	4,584 (191 countries $\times$ 24 years)
Column Count	102
Year Range	2000–2023
Geographic Scope	191 UN member states
Derived Indices	DII, RRS, CRI (normalized 0–100)
Missing Strategy	Within-country linear interpolation

Radar Chart X-Coordinate:

```
([Pivot Field Values] + 2.5) *
 COS(RADIANS(MIN([Radar Angle])))
```

Radar Chart Y-Coordinate:

```
([Pivot Field Values] + 2.5) *
 SIN(RADIANS(MIN([Radar Angle])))
```

Selection Count for Context Switching:

```
{FIXED : COUNTD(
 IF [Selected Countries Set]
 THEN [iso3] END)}
```

The ETL pipeline includes automated validation checks:

- **Coverage Verification**: All 102 columns validated for >25% non-null values
- **Range Checks**: Normalized indices verified to [0, 100] bounds
- **Referential Integrity**: All ISO3 codes validated against pycountry library
- **Temporal Consistency**: Year-over-year changes flagged if >3 standard deviations

The validation report (`validation_report.txt`) confirms:

- DII Coverage: 99.4%, Range: 0.0000–17.9065
- RRS Coverage: 100.0%, Range: 0.2257–2.4044
- CRI Coverage: 100.0%, Range: 0.0887–4.2822

References

[1] EM-DAT, "The International Disaster Database," CRED, UCLouvain. [Online]. Available: https://www.emdat.be/
[2] World Bank, "World Development Indicators," 2024. [Online]. Available: https://data.worldbank.org/
[3] D. Kaufmann, A. Kraay, and M. Mastruzzi, "Worldwide Governance Indicators," World Bank Policy Research Working Paper No. 5430, 2010.
[4] Univ. of Notre Dame, "ND-GAIN Country Index." [Online]. Available: https://gain.nd.edu/our-work/country-index/
[5] UNDP, "Human Development Reports." [Online]. Available: https://hdr.undp.org/data-center
[6] GDACS, "Global Disaster Alert and Coordination System." [Online]. Available: https://www.gdacs.org/
[7] IMF, "World Economic Outlook Database." [Online]. Available: https://www.imf.org/en/Publications/WEO
[8] EC-JRC, "INFORM Risk Index." [Online]. Available: https://drmkc.jrc.ec.europa.eu/inform-index
[9] J. Yao, "Harmonized Global Nighttime Lights," 2019. [Online]. Available: https://figshare.com/articles/dataset/9828827
[10] OCHA, "Financial Tracking Service." [Online]. Available: https://fts.unocha.org/

[11] UNDRR, “DesInventar Sendai.” [Online]. Available: https://www.desinventar.net/
[12] R. Barro and J. Lee, “Educational Attainment Dataset.” [Online]. Available: https://barrolee.github.io/BarroLeeDataSet/
[13] World Inequality Database. [Online]. Available: https://wid.world/data/
[14] UNDRR, “Global Assessment Report on Disaster Risk Reduction,” Geneva, 2022.
[15] Our World in Data, “Natural Disasters.” [Online]. Available: https://ourworldindata.org/natural-disasters
[16] Chen and Nordhaus, “Global Gridded Geographically Based Economic Data (G-Econ),” Yale University, 2011.