# README for the Reproducibility Package for

# Public Institutions Global Observatory: Annual Review 2025

## Overview

This reproducibility package is designed to recreate all analyses featured in the Public Institutions Global Observatory (PIGO) Annual Review 2025, a publication produced by the EGVPI. It contains all the necessary data transformation scripts, custom functions, and plotting scripts required to replicate the report’s figures precisely.

The package is structured to align with the report's chapters:

* Chapter 1: Workflows in this section enable the reproduction of analyses examining global and regional patterns and trends in institutional capacity.
* Chapter 2: This chapter's contents allow for a more granular, cluster-level analysis of operational engagements and the evidence based on institutional capacity. Additionally, it includes scripts to reproduce the display of selected highlights from World Bank lending operations and knowledge products focused on institutions.

It is important to note that the PIGO Annual Review 2025 is intended for internal use by World Bank staff. Therefore, the report’s structure and findings are specifically tailored to meet the needs of staff, ensuring it is fit-for-purpose and provides actionable insights.

## Data Availability

The full documentation is available in the [PIGO\_Data\_Availability\_Statement.Rmd](https://github.com/WB-PIDA-Data-Science-Shop/pigographs/blob/main/documentation/PIGO_Data_Availability_Statement.Rmd) file, classified by chapter. Nonetheless, here is an overview.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Dataset | | Link | Year | Access |
| CLIAR Dashboard | | cliar\db\_variables.xlsx | <https://w0lxdrconn01.worldbank.org/CLIAR/>  Codebook elaborated by the EGVPI team | May 2025 | Limited |
| CLIAR Dashboard | | cliar\dynamic\_ctf\_032525.dta | <https://w0lxdrconn01.worldbank.org/CLIAR/> | March 2025 | Limited |
| CLIAR Dashboard | | cliar\static\_ctf\_032525.dta | <https://w0lxdrconn01.worldbank.org/CLIAR/> | March 2025 | Limited |
| World Bank | | wb\client\_countries.csv | Internally shared | February 2025 | Limited |
| World Bank | | wb\CLASS.xlsx | <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> | February 2025 | Limited |
| World Bank | P\_Data\_Extract\_From\_World\_Development\_Indicators updated.xlsx | | <https://databank.worldbank.org/source/world-development-indicators> | March 2025 | Public |
| World Bank | boost\BOOST\_country\_data.xlsx | | Internally shared | October 2024 | Limited |
| World Bank | boost\budget\_priorities.xlsx | | Internally processed | November 2024 | Limited |
| Global Survey of Public Servants | gsps/GSPS\_Indicators\_Dataset\_11\_10\_23.xlsx | | <https://www.globalsurveyofpublicservants.org/sites/default/files/2023-11/GSPS_Indicators_Dataset_11_10_23.xlsx> | May 2025 | Limited |
| GovTech Maturity Index | wb/dgss.csv | | <https://datacatalogfiles.worldbank.org/ddh-published/0037889/DR0065450/WBG_GovTech20Dataset_Dec2020.xlsx> | May 2025 | Limited |

## Statement about Rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the LICENSE.txt file.

## Instructions for Replicators

As outlined in the project's README.Rmd, replicating the results involves two key stages:

**1. Restore Package Dependencies:**

1. Ensure you are using the precise package versions required for this project.
2. Open the pigographs.Rproj file in RStudio.
3. Execute the command renv::restore() in the RStudio console. This will install the exact package versions specified in the renv.lock file.

**2. Execute the Full Workflow:**

1. Data Transformation:
   1. Confirm that you have received and placed the input/ folder provided by the EGVPI team in the project's root directory.
   2. Source each of the raw data preprocessing scripts located in the data-raw/source/ subdirectory. Execute these scripts sequentially.
2. Data Analysis & Visualization:
   1. Source all analysis and plotting scripts found within the analysis/ subdirectory.
   2. Carefully verify that each plot generated in the figures/ folder matches the corresponding figure in the PIGO Annual Review exactly.

**To perform both stages conveniently, open the project's README.Rmd file and simply knit it.** For detailed information on each exhibit, please refer to the List of Exhibits.

## Software Requirements

R: Version 4.3 or higher

Dependency Management: renv for managing package dependencies

Development Tools: devtools for package loading

## Memory, Runtime, and Storage Requirements

Memory: No minimum required, but 16 GB is recommended for optimal performance.

Run Time: ~5 minutes on recommended hardware.

## Folder Structure

The repository is organized as follows:

├── data-raw/ # Raw data and initial processing scripts.

│ ├── source # Scripts for transforming raw data

│ ├── input # Raw data files (provided by the EGVPI Team)

│ └── output # Transformed data for ad-hoc analysis

├── analysis/ # Scripts for analysis and plotting.

├── R/ # Custom R functions.

├── documentation/ # Reproducibility documentation.

├── data/ # Processed data for analysis.

├── figures/ # Final generated plots.

├── man/ # R package manual pages.

├── renv/ # R package dependency management.

├── README.Rmd # Main file to run the reproducibility workflow.

├── pigographs.Rproj # RStudio project file.

├── .gitignore # Files ignored by Git.

├── .Rbuildignore # Files ignored when building.

├── .Rprofile # Project-specific R startup settings.

├── DESCRIPTION # Project metadata.

├── LICENSE # Licensing information.

├── LICENSE.md # License in Markdown format.

├── NAMESPACE # Package namespace definition.

└── spielplatz/ # Sandbox/experimentation area.

## List of Exhibits

|  |  |  |  |
| --- | --- | --- | --- |
| Exhibit name | Output filename | Script | Note |
| Figure 1. Country-Level Institutional Assessment and Review (CLIAR) Framework | NA | NA | Image.  Source: [CLIAR Methodology note](https://w0lxdrconn01.worldbank.org/CLIAR/)) |
| Figure 2. Quality of institutions and levels of economic development are positively correlated | GDPpercapitaPPPconstant2021international\_main | correlation-income-risk.R |  |
| Figure 3. Higher levels of institutional capacity are associated with lower levels of poverty | Povertyheadcountratioat215aday2017PPPofpopulation\_main | correlation-income-risk.R |  |
| Figure 4. Higher levels of institutional capacity are correlated with lower country risk | CountryRisk\_vs\_CTF\_main | correlation-income-risk.R |  |
| Figure 5. Regional variation in the policy arena (political and social) | 05-poli-arena-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 5. Regional variation in the policy arena (political and social) | 05-social-arena-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 6. Regional variation in institutions at the center of government | 06-center-of-gov-digital-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 6. Regional variation in institutions at the center of government | 06-center-of-gov-hrm-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 7. Regional Variation in Degree of Integrity and Transparency and Accountability | 07-integrity-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 7. Regional Variation in Degree of Integrity and Transparency and Accountability | 07-transp-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 8. Regional variation in specific thematic areas: Justice and Energy and Environment | 08-enviroment-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 8. Regional variation in specific thematic areas: Justice and Energy and Environment | 08-justice-regional-dumbbells.png | regional-cluster-dumbbells.R |  |
| Figure 9. Change in political institutions over time (2018 vs 2022) | 09-political-cluster-change.png | clusters-over-time.R |  |
| Figure 10. Change in social institutions over time (2018 vs 2022) | 10-social-cluster-change.png | clusters-over-time.R |  |
| Figure 11. Change in public human resource management institutions over time (2018-2022) | 11-hrm-cluster-change.png | clusters-over-time.R |  |
| Figure 12. Change in degree of integrity over time (2018-2022) | 12-integrity-cluster-change.png | clusters-over-time.R |  |
| Figure 13. Regional variation in the share of workers with tertiary education in the public and private sectors | NA | NA | Image. Source:  Worldwide Bureaucracy Indicators |
| Figure 14. Global datasets on public financial management | NA | NA | Hyperlinks references, not applicable |
| Figure 15. Regional variation in capital budget execution rate in service delivery (2019) | 15-service-delivery-education-budget-execution.png | budget-execution-by-region.R |  |
| Figure 15. Regional variation in capital budget execution rate in service delivery (2019) | 15-service-delivery-health-budget-execution.png | budget-execution-by-region.R |  |
| Figure 16. Global datasets on government technologies | NA | NA | Hyperlinks references, not applicable |
| Figure 17. Share of countries with operational management information systems, by region | 17-coverage\_information\_systems.png | analyze\_information\_systems.R |  |
| Figure 18. Global datasets on management practices | NA | NA | Hyperlinks references, not applicable |
| Figure 19. Institutional-level variation in institutional target-setting (Guatemala and Ghana) | 19\_mgmt\_target.png | gsps-management.R |  |
| Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (a) | GDPpercapitaPPPconstant2021international\_annex | correlation-income-risk.R |  |
| Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (b) | Povertyheadcountratioat215aday2017PPPofpopulation\_annex | correlation-income-risk.R |  |
| Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (c) | CountryRisk\_vs\_CTF\_annex | correlation-income-risk.R |  |
| Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (d) | Povertyheadcountratioat215aday2017PPPofpopulation\_annex\_nossa | correlation-income-risk.R |  |
| Annex 2. Change in Sector-specific Institutions (a) - justice | annex-a-justice-cluster-change.png | clusters-over-time.R |  |
| Annex 2. Change in Sector-specific Institutions (b) - environmental | annex-b-enviromental-cluster-change.png | clusters-over-time.R |  |