

README for Data for Better Governance: Building Government Analytics Ecosystems in Latin America and the Caribbean

Overview

The “Data for Better Governance: Building Government Analytics Ecosystems in Latin America and the Caribbean” report includes an assessment of the current state of data analytics using MIS data, as well as the enabling conditions across the region from a survey conducted in 20 countries in the LAC region, with the findings detailed in Chapter 4 of the report. The findings of the survey are presented in Chapter 4.

The code in the folder generates the figures presented in Chapter 4 of the report. Additionally, the package includes a set of two excel files to generate the two figures presented in Chapter 1.

The folder contains three directories:

1. Data: This directory contains an input and output folder. The input folder contains the datasets used in this report (survey of data analytics in 6 MIS) and the output folder contains the generated final datasets and figures generated throughout the code.
2. Do-files: This directory contains the master do-file “Master” that imports, cleans and harmonized the datasets as well as creates the figures for the report.
3. Excel: This directory contains the excel documents with data and analysis used to generate figures 1.1 and 1.2 of the report.

Data Availability

Data cannot be publicly shared in the RP website due to confidentiality agreements with participating countries.

The data were shared exclusively to replicators. The graphs in the report accurately replicate those generated from the STATA do files, albeit with design adjustments for presentation.

The complete datasets in dta format are included in the Reproducibility Package for verification purposes under the Data>Input folder as follow:

- EDMIS.dta
- HealthMIS.dta
- HRMIS.dta
- PFMIS.dta
- ProcurementMIS.dta
- TaxesMIS.dta

- indcap_data.dta

The classification of the datasets is:

- All data are publicly available.
- Some data cannot be made publicly available.
- No data can be made publicly available.

Data Sources

We conducted a survey in 20 countries in the LAC region. The survey comprised seven questionnaires: six MIS-level questionnaires tailored to six different MIS (EdMIS, HealthMIS, HRMIS, e-Procurement, PFMIS, and TaxMIS), and an additional country-level “capabilities questionnaire” that focused on government initiatives to enhance the analytical skills of public servants. Over 100 government officials responded, providing valuable insights from their respective fields.

The survey data was collected and stored on SurveyCTO, the platform used to administer and manage the survey process. The input data was downloaded from there.

Instructions for Replicators

1. Run the do-file “Master.do”. When doing so, it is only necessary to replace the global main with the desired working directory in code line 10 (the do-files called by the Master do-file rely only on that path, so no further changes are necessary).

*Figures 1.1 and 1.2 of the report use data from CEPALSTAT that is not contained in the raw datasets. This data and the numbers necessary to replicate these two figures are included inside the “Excel” folder. This data has been manually added and included in the package for transparency.

Note on Country Representation Across Figures

When reviewing Figures 4.2, 4.11, and 4.13, it is important to note that the representation of countries is not consistent across these figures. For instance, “Country 1” in Figure 4.2 may correspond to a different country than “Country 1” in Figure 4.11 or Figure 4.13.

This variation arises because the figures are generated independently from distinct subsets of the data. Users should exercise caution when interpreting results across these figures and avoid assuming that the same country is represented consistently by the same label or identifier.

List of Exhibits

The provided code reproduces:

- All numbers provided in text in the paper
- All tables and figures in the paper

- □ Selected tables and figures in the paper, as explained and justified below

Exhibit name	Output filename	Script	Note
Figure 1.1	LAC GOV. Analytics Figure 1.1.xlsx	Excel file	Found in the excel file (sheet “analysis”)
Figures 1.2	LAC GOV. Analytics Figure 1.2a.xlsx	Excel file	Found in the excel file (sheet “Trends”)
Figures 4.1 to 4.14	Figure_4.1_png-Figure_4.14_png	Graphs_for_report	

Requirements

Computational Requirements

The program runs in Stata (graphs 4.1-4.14) and uses the mrtab dependency. For graphs 1.1 and 1.2, the figures are generated using excel.

Software Requirements

- **Stata version 15**
 - mrtab

Memory and Runtime and Storage Requirements

The code takes around 15 minutes to run and reproduce all the figures.

Folder Structure

The shared folder’s structure for a reproducibility package looks like this:

```
Data
├── Input
└── Output
Code
├── Master.do
└── Create input/`MIS'_analysis.do
```