

R code explanation

Author: Monica Cueto Tapia

Position: Impact Evaluation Consultant - IDB Invest

The R code I've provided serves as a comprehensive script designed for systematically retrieving and loading financial data from various countries. Organized into distinct sections corresponding to each nation, the script utilizes techniques like web scraping, file downloads, and data update handling to accomplish its tasks. Initially, I set up directories for efficient data loading and storage and load essential R packages like dplyr, glue, here, and httr to streamline data processing.

In executing country-specific procedures, such as those for Bolivia, I employ web scraping to extract PDF URLs from the ASOBAN website, followed by downloading and table extraction from these PDFs. Similarly, procedures for Brazil involve downloading ZIP files from the Banco Central do Brasil website, while Chile's data is obtained in Excel format from the Superintendencia de Bancos e Instituciones Financieras site. The script's versatility is evident as I repeat similar steps for countries like Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Nicaragua, Panama, Paraguay, Peru, and Venezuela. To ensure successful data updates, I monitor the update status for each country and compile the information into a data frame. Additionally, I generate a comprehensive table summarizing the loading status for each nation to enhance transparency and provide an accessible overview. It's important to acknowledge that the effectiveness of the script hinges on data availability and website structures, and adhering to the terms of use for scraped data sources is crucial. For production deployment, I recommend incorporating error handling and logging mechanisms to bolster robustness.