

Package ‘pressfreedom.data’

August 24, 2026

Type Package

Title Download and Process Reporters Without Borders Press Freedom Index Data

Version 0.3.0

Maintainer Peter Baumgartner <petzi53@gmail.com>

Description Download press freedom index data from Reporters Without Borders (RSF) with period-aware encoding handling. Data are downloaded from the RSF website (<<https://rsf.org/en/index>>). Provides infrastructure for data cleaning and ISO 3166 standardization in downstream phases.

License MIT + file LICENSE

URL <https://github.com/petzi53/pressfreedom.data>,
<https://www.peter-baumgartner.net/pressfreedom.data/>

BugReports <https://github.com/petzi53/pressfreedom.data/issues>

Encoding UTF-8

LazyData true

Imports readr, here, rlang, countrycode, dplyr, stringr, cli, fs,
glue, purrr, tibble, stringi

Suggests usethis, knitr, rmarkdown, ggplot2, tidyr, forcats,
patchwork, scales, sf, rnatualearth, testthat (>= 3.0.0), covr

VignetteBuilder knitr

Config/roxygen2/version 8.1.0

Config/testthat/edition 3

Depends R (>= 4.1.0)

NeedsCompilation no

Author Peter Baumgartner [aut, cre] (ORCID:
<<https://orcid.org/0000-0003-4526-8791>>)

Repository CRAN

Date/Publication 2026-08-24 12:40:02 UTC

Contents

rwb_standardized	2
Index	4

rwb_standardized	<i>Standardized Reporters Without Borders Press Freedom Index (2002-2026)</i>
------------------	---

Description

A comprehensive dataset of press freedom indicators from Reporters Without Borders (RSF: Reporters Sans Frontieres, the organization’s French name and legal identity; RWB: Reporters Without Borders, its common English name – used for the ‘rwb_‘ prefix throughout this package), standardized and cleaned by the pressfreedom.data package pipeline.

Usage

rwb_standardized

Format

A data frame with 4,192 rows and 20 columns:

- year_n** Numeric year (2002-2026, excluding 2011)
- iso** ISO 3166-1 alpha-3 country code
- country_en** Standardized country name in English
- score** Press freedom score (0-100; higher = more free). Comparable only within periods.
- rank** Rank within the year (1 = most free)
- political_context** Sub-index: Political context (if available)
- rank_pol** Rank within political context (if available)
- economic_context** Sub-index: Economic context (if available)
- rank_eco** Rank within economic context (if available)
- legal_context** Sub-index: Legal context (if available)
- rank_leg** Rank within legal context (if available)
- social_context** Sub-index: Social context (if available)
- rank_soc** Rank within social context (if available)
- safety** Sub-index: Safety (if available)
- rank_saf** Rank within safety (if available)
- zone** Geographic zone assigned by RSF
- rank_n_1** Previous year’s rank (year_n - 1)
- rank_evolution** Change in rank from previous year
- score_n_1** Previous year’s score (year_n - 1)
- score_evolution** Change in score from previous year

Details

Data Cleaning Pipeline

This dataset is the output of a comprehensive 4-phase data pipeline implemented in the `pressfreedom.data` package:

- **Phase A (Download):** Raw CSV files from RSF website - **Phase B (Normalize):** Column names and data types standardized across periods - **Phase C (Combine):** All periods merged into unified structure - **Phase D (Standardize):** Country names consolidated, ISO codes assigned, duplicates resolved

Important Notes

Audit Trail: The full audit trail (including `'country_name_original'` and `'consolidation_flag'` columns) is preserved in the source RDS file (`'data/processed/rwb_standardized.rds'`) within the `pressfreedom.data` package. This exported dataset contains only the 20 core columns for analysis.

Score Comparability: Scores are only comparable within their respective periods: - **Period 1 (2002-2012):** Non-comparable scores; use ranks for trends - **Period 2 (2013-2021):** Comparable scores (0-100 scale) - **Period 3 (2022-2026):** New methodology; different dimensions tracked

Missing Data: Sub-indices (political, economic, legal, social, safety contexts) are only available in Period 3 (2022-2026). Periods 1-2 have NA values for these columns.

Cyprus: Tracked as two separate entities: - "Cyprus" (ISO: CYP) - Republic of Cyprus - "Northern Cyprus" (ISO: CXX) - Turkish Republic of Northern Cyprus

Source

Reporters Sans Frontieres, <https://rsf.org>

Examples

```
# Load the dataset
data(rwb_standardized)

# Basic summary
head(rwb_standardized)

# Countries included
length(unique(rwb_standardized$country_en))

# Years covered
range(rwb_standardized$year_n)
```

Index

* **datasets**

 rwb_standardized, [2](#)

rwb_standardized, [2](#)