Package 'sae4health'

July 23, 2025

Title Small Area Estimation for Key Health and Demographic Indicators from Household Surveys

Version 1.2.3

Description Enables small area estimation (SAE) of health and demographic indicators in low- and middle-income countries (LMICs). It powers an R 'shiny' application for generating subnational estimates and prevalence maps of 150+ binary indicators from Demographic and Health Surveys (DHS). It builds on the SAE analysis workflow from the 'surveyPrev' package. For documentation, visit <https://sae4health.stat.uw.edu/>. Methodological details can be found at Wu et al. (2025) <doi:10.48550/arXiv.2505.01467>.

License AGPL-3

URL https://sae4health.stat.uw.edu/, https://github.com/wu-thomas/sae4health

BugReports https://github.com/wu-thomas/sae4health/issues

Depends R (>= 4.3)

Imports config, dplyr, DT, ggplot2, golem (>= 0.4.1), grid, gridExtra, htmltools, htmlwidgets, leaflet, R6, sf, sp, shiny (>= 1.7.4), shinyBS, shinydashboard, shinyFeedback, shinyjs, shinyWidgets, SUMMER, surveyPrev, survey, geodata, bookdown, markdown, haven, ggridges, ggthemes, RColorBrewer, viridisLite, scales, patchwork, leaflegend, leafsync, methods, graphics, plotly, readr, sn

Suggests INLA, testthat (>= 3.0.0)

Additional_repositories https://inla.r-inla-download.org/R/testing/

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

NeedsCompilation no

Config/build/clean-inst-doc FALSE

Config/testthat/edition 3

Language en-US

Author Yunhan Wu [cre, aut],
Qianyu Dong [aut],
Zehang R Li [aut],
Jon Wakefield [aut]

Maintainer Yunhan Wu <wu-thomas@outlook.com>

Repository CRAN

Date/Publication 2025-05-07 12:30:02 UTC

Contents

`sae4health-package` *R Shiny App for Small Area Estimation of Health and Demographic Indicators*

Description

The **sae4health** package powers an **R Shiny app** designed for small area estimation (SAE) of health and demographic indicators in low- and middle-income countries (LMICs). It enables subnational estimation and prevalence mapping for more than 150 binary indicators derived from Demographic and Health Surveys (DHS), providing an intuitive interface for public health analysts, policymakers, and researchers.

Details

Built on the **surveyPrev** package, **sae4health** ensures methodological rigor in SAE analysis. It offers guided model selection, automated model fitting, and interactive visualization, making advanced statistical methods accessible to non-experts.

For comprehensive documentation on the **sae4health** project and **web-based app access**, visit: https://sae4health.stat.uw.edu/

The latest development version of the package is maintained at: https://github.com/wu-thomas/sae4health

Citation: Wu, Y., Dong, Q., Xu, J., Li, Z. R., & Wakefield, J. (2025). *sae4health: An R Shiny Application for Small Area Estimation in Low- and Middle-Income Countries*. doi:10.48550/arXiv.2505.01467.

Author(s)

- Yunhan Wu [**Maintainer**] (<wu-thomas@outlook.com>)
- Qianyu Dong (<qdong14@ucsc.edu>)
- Zehang R Li (<lizehang@gmail.com>)
- Jon Wakefield (<jonno@uw.edu>)

References

Wu, Y., Dong, Q., Xu, J., Li, Z. R., & Wakefield, J. (2025). *sae4health: An R Shiny Application for Small Area Estimation in Low- and Middle-Income Countries*. arXiv preprint.doi:10.48550/arXiv.2505.01467

See Also

Getting Started: https://sae4health.stat.uw.edu/overview/project_overview/

Demo and Instruction Video: https://sae4health.stat.uw.edu/overview/youtube_app_demo/

Statistical Methods: https://sae4health.stat.uw.edu/method/approach_overview/

Visualization Gallery: https://sae4health.stat.uw.edu/gallery/visual_overview/

Recent Updates and News: https://sae4health.stat.uw.edu/blog/

---

adm2.link.all *WHO Administrative Level 2 Linkage*

Description

A dataset linking administrative level 2 regions to their corresponding WHO region and country codes.

Usage

```
adm2.link.all
```

Format

A data frame with 1,095 rows and 8 columns:

WHO_REGION WHO-designated region for the country.

ISO.3.DIGIT.COUNTRY.CODE Three-letter ISO country code.

ADM0_VIZ_NAME Administrative level 0 (country) name for visualization.

ADM1_VIZ_NAME Administrative level 1 (first subnational division) name.

ADM2_VIZ_NAME Administrative level 2 (second subnational division) name.

GUID.LEVEL.0 Global unique identifier for level 0 (country).

GUID.LEVEL.1.. Global unique identifier for level 1 (first subnational division).

GLOBAL.UNIQUE.IDENTIFIER.. Global unique identifier for level 2 (second subnational division).

---

admin_to_num *Convert Administrative Level String to Numeric Code*

Description

This function converts administrative level names (e.g., "National", "Admin-1") into corresponding numerical values.

Usage

```
admin_to_num(admin_level)
```

Arguments

admin_level	A character string representing the administrative level. It can be "National" or "Admin-X" (where X is a positive integer).

Value

An integer representing the numerical level: - "National" is converted to 0. - "Admin-X" is converted to X as an integer. - Returns NULL if the input is invalid.

Examples

```
admin_to_num("National")   # Returns 0
admin_to_num("Admin-1")    # Returns 1
admin_to_num("Admin-2")    # Returns 2
admin_to_num("Invalid")    # Returns NULL
```

DHS.country.meta *DHS Country Metadata*

Description

A dataset containing metadata for countries included in DHS surveys, including standardized country codes across multiple international organizations.

Usage

```
DHS.country.meta
```

Format

A data frame with 60 rows and 12 columns:

UNAIDS_CountryCode Country code used by UNAIDS.

SubregionName Name of the subregion (e.g., West Africa, South Asia).

WHO_CountryCode Country code used by WHO.

FIPS_CountryCode Country code used by FIPS.

UNICEF_CountryCode Country code used by UNICEF.

RegionName Name of the global region (e.g., Africa, Asia).

ISO2_CountryCode Two-letter ISO country code.

ISO3_CountryCode Three-letter ISO country code.

RegionOrder Numeric ordering of the region for visualization.

DHS_CountryCode Unique country code assigned by DHS.

CountryName Full name of the country.

UNSTAT_CountryCode Country code used by UN Statistics Division.

DHS.dataset.meta *DHS Dataset Metadata*

Description

A dataset containing metadata on recodes for DHS surveys supported by the app.

Usage

```
DHS.dataset.meta
```

Format

A data frame with 5,636 rows and 13 columns:

FileFormat Format of the dataset file (e.g., Stata, SPSS, CSV).

FileSize Size of the dataset file in megabytes.

DatasetType Type of dataset/recode (e.g., Household, Individual, Children, Biomarker).

SurveyNum Survey number assigned by DHS.

SurveyId Unique ID of the survey associated with the dataset.

FileType Type of file (e.g., Household Recode, Birth Recode, Men's Recode).

FileDateLastModified Date when the dataset file was last modified.

SurveyType Type of DHS survey (e.g., DHS, AIS, MIS).

SurveyYearLabel Label describing the survey year.

SurveyYear Year in which the DHS survey was conducted.

DHS_CountryCode Unique country code assigned by DHS.

FileName Name of the dataset file.

CountryName Full name of the country associated with the dataset.

DHS.survey.meta *DHS Survey Metadata*

Description

A dataset containing metadata for DHS surveys supported by the app.

Usage

```
DHS.survey.meta
```

Format

A data frame with 153 rows and 30 columns:

ReleaseDate Date the survey data was released.

SurveyId Unique ID assigned to each DHS survey.

MaxAgeWomen Maximum age of surveyed women.

FieldworkStart Start date of survey fieldwork.

MinAgeMen Minimum age of surveyed men.

ImplementingOrg Organization implementing the survey.

SurveyCharacteristicIds IDs of survey characteristics.

SurveyType Type of survey (e.g., DHS, AIS, MIS).

SurveyYearLabel Label describing the survey year.

IndicatorData Boolean. Whether indicator data is available.

MinAgeWomen Minimum age of surveyed women.

SurveyYear Year in which the DHS survey was conducted.

FieldworkEnd End date of survey fieldwork.

DHS_CountryCode DHS country code.

NumberOfSamplePoints Number of sample points in the survey.

CountryName Name of the country where the survey was conducted.

NumberOfWomen Number of women surveyed.

SubregionName Name of the subregion.

MaxAgeMen Maximum age of surveyed men.

SurveyNum Survey number assigned by DHS.

SurveyStatus Status of the survey (e.g., Completed, Ongoing).

NumberOfFacilities Number of health facilities surveyed.

UniverseOfMen Population coverage for male respondents.

RegionName Region name for survey coverage.

UniverseOfWomen Population coverage for female respondents.

Footnotes Additional survey notes.

PublicationDate Date when the survey results were published.

NumberofHouseholds Number of households surveyed.

NumberOfMen Number of men surveyed.

GPS_avail Whether GPS coordinates are available for cluster locations.

`DHS_api_est` *DHS API Estimates*

Description

A dataset containing estimated health indicators from the DHS API, including country-level estimates and subgroup breakdowns.

Usage

```
DHS_api_est
```

Format

A data frame with 28,110 rows and 7 columns:

Country Name of the country where the survey was conducted.

Country Code Three-letter ISO country code.

Survey Year Year in which the DHS survey was conducted.

DHS Standard ID Unique identifier for the DHS survey.

Definition Definition of the health indicator being estimated.

Estimate Estimated value of the indicator.

By Variable Label Label describing any subgrouping (e.g., urban/rural, age group).

`DHS_ind_dictionary` *DHS Indicator Dictionary*

Description

A dataset for DHS health indicators, including their DHS official definitions, measurement types, and denominators.

Usage

```
DHS_ind_dictionary
```

Format

A data frame with 4,433 rows and 5 columns:

DHS Standard Indicator ID Unique identifier for each DHS indicator.

Label Short name or label for the indicator.

Full Definition Detailed description of the indicator.

Denominator Explanation of the denominator used in indicator calculation.

Measurement Type Type of measurement (e.g., percentage, count, rate).

`match_all_result` *Matching Results Between DHS and GitHub Indicators, Imported from surveyPrev*

Description

The dataset 'match_all_result' contains the results of matching indicators from DHS (Demographic and Health Surveys) with those extracted from GitHub. The dataset includes similarity scores, recoding names, indicator definitions, and positional information within the matched text.

Usage

```
match_all_result
```

Format

A data frame with 132 rows and 20 variables:

indicator_ID_DHS Character. Unique ID of the indicator in DHS.

DHS_label Character. Label assigned to the indicator in DHS.

DHS_definition Character. Description or definition of the DHS indicator.

Combined Character. Combined text representation for similarity matching.

Similarity Numeric. Similarity score between DHS and GitHub indicators.

indicator_ID_Github Character. Unique ID of the indicator from GitHub.

recode_name Character. Initial recoding name for the indicator.

updated_recode_name Character. Revised recoding name after updates.

indicator_ID_Github_raw Character. Raw version of the GitHub indicator ID.

indicator_def_github Character. Processed indicator definition from GitHub.

indicator_def_github_raw Character. Unprocessed raw definition from GitHub.

indicator_chapter Character. The chapter or category of the indicator.

indicator_R_bundle Character. Associated R bundle or package for processing.

start_position Integer. Start position of the matched pattern in text.

end_position Integer. End position of the matched pattern in text.

matched_pattern Character. The exact text pattern matched between sources.

batch_recode_group Character. Grouping variable for batch recoding.

indicator Character. Final matched indicator name.

X Unknown. This variable needs clarification or may be redundant.

ID_first_two_letters Character. First two letters of the indicator ID, possibly used for grouping or sorting.

natl.WHO.shp *WHO National-Level Shapefile*

Description

A spatial dataset containing national-level WHO boundaries for eight countries, with administrative codes and metadata.

Usage

```
natl.WHO.shp
```

Format

A spatial data frame with 8 rows and 35 columns:

WHO_REGION WHO-designated region for the country.

ISO_3_CODE Three-letter ISO country code.

ADM0_NAME Administrative level 0 (country) name.

ADM0_CODE Administrative level 0 country code. ...

num_to_admin *Convert Numeric Code to Administrative Level String*

Description

This function converts a numerical administrative level into its corresponding string format.

Usage

```
num_to_admin(num)
```

Arguments

num	A single integer representing the administrative level. The value 0 corresponds to "National", while positive integers correspond to "Admin-X".

Value

A character string representing the administrative level: - 0 is converted to "National". - Positive integers are converted to "Admin-X". - Returns NULL if the input is invalid.

Examples

```
num_to_admin(0)    # Returns "National"
num_to_admin(1)    # Returns "Admin-1"
num_to_admin(2)    # Returns "Admin-2"
```

ref_tab_22 *Reference Table for Original 22 Indicators*

Description

A dataset containing the original 22 health and demographic indicators supported in the app.

Usage

```
ref_tab_22
```

Format

A data frame with 22 rows and 13 columns:

ID Unique identifier for the indicator.

Description Short label for the indicator.

Full_definition Detailed definition of the indicator.

Topic General category or theme of the indicator.

Chap_abbrev Abbreviation of the chapter where the indicator appears.

IR Logical. Whether the indicator is available in the Individual Recode dataset.

PR Logical. Whether the indicator is available in the Household Members dataset.

KR Logical. Whether the indicator is available in the Children's Recode dataset.

BR Logical. Whether the indicator is available in the Birth Recode dataset.

HR Logical. Whether the indicator is available in the Household Recode dataset.

MR Logical. Whether the indicator is available in the Men's Recode dataset.

AR Logical. Whether the indicator is available in the AIDS Indicator Survey dataset.

CR Logical. Whether the indicator is available in the Couple's Recode dataset.

ref_tab_all *Combined Indicators Reference Table*

Description

A dataset containing all 153 indicators supported in the app, combining the original 22 indicators with the newly added ones.

Usage

```
ref_tab_all
```

Format

A data frame with 153 rows and 15 columns:

ID Unique identifier for the indicator.

Description Short label for the indicator.

Full_definition Detailed definition of the indicator.

Topic General category or theme of the indicator.

Chap_abbrev Abbreviation of the chapter where the indicator appears.

IR Logical. Whether the indicator is available in the Individual Recode dataset.

PR Logical. Whether the indicator is available in the Household Members dataset.

KR Logical. Whether the indicator is available in the Children's Recode dataset.

BR Logical. Whether the indicator is available in the Birth Recode dataset.

HR Logical. Whether the indicator is available in the Household Recode dataset.

MR Logical. Whether the indicator is available in the Men's Recode dataset.

AR Logical. Whether the indicator is available in the AIDS Indicator Survey dataset.

CR Logical. Whether the indicator is available in the Couple's Recode dataset.

Chapter Chapter reference from the DHS reports.

Title Title of the section where the indicator appears.

---

ref_tab_new *Newly Added Indicators Reference Table*

Description

A dataset containing newly added health and demographic indicators in the app after its initial release.

Usage

```
ref_tab_new
```

Format

A data frame with 134 rows and 13 columns:

ID Unique identifier for the indicator.

Description Short label for the indicator.

Full_definition Detailed definition of the indicator.

Topic General category or theme of the indicator.

Chap_abbrev Abbreviation of the chapter where the indicator appears.

IR Logical. Whether the indicator is available in the Individual Recode dataset.

PR Logical. Whether the indicator is available in the Household Members dataset.

KR Logical. Whether the indicator is available in the Children's Recode dataset.

BR Logical. Whether the indicator is available in the Birth Recode dataset.

HR Logical. Whether the indicator is available in the Household Recode dataset.

MR Logical. Whether the indicator is available in the Men's Recode dataset.

AR Logical. Whether the indicator is available in the AIDS Indicator Survey dataset.

CR Logical. Whether the indicator is available in the Couple's Recode dataset.

`run_app` *Run the Shiny Application*

Description

Run the Shiny Application

Usage

```
run_app(
  onStart = NULL,
  options = list(),
  enableBookmarking = NULL,
  uiPattern = "/",
  ...
)
```

Arguments

`onStart`	A function that will be called before the app is actually run. This is only needed for `shinyAppObj`, since in the `shinyAppDir` case, a `global.R` file can be used for this purpose.
`options`	Named options that should be passed to the `runApp` call (these can be any of the following: "port", "launch.browser", "host", "quiet", "display.mode" and "test.mode"). You can also specify `width` and `height` parameters which provide a hint to the embedding environment about the ideal height/width for the app.
`enableBookmarking`	Can be one of `"url"`, `"server"`, or `"disable"`. The default value, `NULL`, will respect the setting from any previous calls to `enableBookmarking()`. See `enableBookmarking()` for more information on bookmarking your app.
`uiPattern`	A regular expression that will be applied to each `GET` request to determine whether the `ui` should be used to handle the request. Note that the entire request path must match the regular expression in order for the match to be considered successful.
`...`	arguments to pass to golem_opts. See '?golem::get_golem_options' for more details.

Index