

Package: hmsidwR (via r-universe)

February 19, 2025

Title Health Metrics and the Spread of Infectious Diseases

Version 1.1.2

Description A collection of datasets and supporting functions accompanying Health Metrics and the Spread of Infectious Diseases by Federica Gazzelloni (2024). This package provides data for health metrics calculations, including Disability-Adjusted Life Years (DALYs), Years of Life Lost (YLLs), and Years Lived with Disability (YLDs), as well as additional tools for analyzing and visualizing health data. Federica Gazzelloni (2024) <[doi:10.5281/zenodo.10818338](https://doi.org/10.5281/zenodo.10818338)>.

License MIT + file LICENSE

URL <https://github.com/Fgazzelloni/hmsidwR>,
<https://fgazzelloni.github.io/hmsidwR/>

BugReports <https://github.com/Fgazzelloni/hmsidwR/issues>

Depends R (>= 2.10)

Imports ggplot2, gstat, httr, jsonlite, purrr, showtext, sysfonts, tibble

Suggests devtools, dplyr, geomtextpath, ggthemes, janitor, knitr, lubridate, maps, pkgdown, plotly, readr, readxl, rmarkdown, sessioninfo, sf, stats, testthat (>= 3.0.0), tidyr, tidyverse

Config/testthat/edition 3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

VignetteBuilder knitr

Config/pak/sysreqs libfreetype6-dev libgdal-dev gdal-bin libgeos-dev libpng-dev libssl-dev libproj-dev libsqlite3-dev libudunits2-dev zlib1g-dev

Repository <https://fgazzelloni.r-universe.dev>

RemoteUrl <https://github.com/fgazzelloni/hmsidwr>

RemoteRef HEAD

RemoteSha b218cad3fdd98802635d74b9aa354ca8508a955a

Contents

deaths2019	2
deaths9	3
disweights	4
g7_hmetrics	5
gbd_get_data	6
germany_lungc	7
getunz	7
gho_le_hale	8
gho_lifetables	9
idDALY_map_data	9
id_affected_countries	10
incprev_stroke	11
infectious_diseases	11
kbit	12
rabies	13
sdi90_19	14
spatialdalys2021	15
string_search	16
theme_hmsid	16
Index	18

deaths2019	<i>Dataset: Health Metrics Data - Number of Deaths Due to 9 Causes in 2019</i>
------------	--

Description

A dataset containing the number of Deaths due to 9 causes in 6 regions for 2019.

Usage

```
data(deaths2019)
```

Format

A dataframe with 2754 rows and 7 variables:

The variables are as follows:

location character, France, Germany, Global, Italy, United Kingdom, United States of America

sex character, Female, Male, Both
age character, age groups from <1 to 85+ each 5 years
cause character, Alzheimer's disease and other dementias, Breast cancer, Chronic obstructive pulmonary disease, Colon and rectum cancer, Diabetes and kidney diseases, Lower respiratory infections, Road injuries, Stroke, Tracheal, bronchus, and lung cancer
val numeric, deaths number estimation
upper numeric, upper value estimation
lower numeric, lower value estimation

Source

2019 data from the [IHME](#) website

Examples

```
data(deaths2019)
head(deaths2019)
```

deaths9	<i>Health Metrics Data - Number of Deaths Due to 9 Causes in 6 Locations for the Years 2011 and 2021.</i>
---------	---

Description

Health Metrics Data - Number of Deaths Due to 9 Causes in 6 Locations for the Years 2011 and 2021.

Usage

```
data(deaths9)
```

Format

A dataframe with 5112 rows and 7 variables:

The variables are as follows:

location character, France, Germany, Global, Italy, UK, USA
iso2 character, country code
sex character, female, male, both
age character, 5-year age groups from <5 to 85+
cause character, Alzheimer's disease and other dementias, Breast cancer, Chronic obstructive pulmonary disease, Colon and rectum cancer, Diabetes and kidney diseases, Lower respiratory infections, Road injuries, Stroke, Tracheal, bronchus, and lung cancer
year integer, years 2011 and 2019
dx numeric, deaths number estimation

Source

2021 data from the [IHME](#) website

Examples

```
data(deaths9)
head(deaths9)
```

disweights

Dataset: Health Metrics Data - Disability Weights and Severity in 2019 and 2021

Description

A dataset containing the Disability Weights estimates, upper and lower values, and the Severity level for Stroke, Tuberculosis, and HIV for all countries.

Usage

```
disweights
```

Format

A dataframe with 463 rows and 9 variables:

The variables are as follows:

sequela character, disease sequela

specification character, disease specification

cause1 character, first cause of disease - morbidity

cause2 character, second cause of disease - morbidity

severity character, mild, moderate, severe, mean

dw numeric, disability weights estimation

upper numeric, upper value estimation

lower numeric, lower value estimation

Source

Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 and 2021 Disability Weights. Seattle, United States of America: Institute for Health Metrics and Evaluation (IHME), 2024.

`g7_hmetrics`*Dataset: Health Metrics Data - G7 Countries*

Description

A subset of data from the IHME GBD on Deaths, Disability-Adjusted Life Years (DALYs), Years of Life Lost (YLLs), and Years Lived with Disability (YLDs), Incidence and Prevalence, age standardized for all causes and respiratory infections and tuberculosis. For years 2010, 2019 and 2021.

Usage`g7_hmetrics`**Format**

A dataframe with 3402 rows and 9 variables:

The variables are as follows:

measure character, metric name

location character, country

sex character, Female, Male, Both

cause character, all causes, and respiratory infections and tuberculosis

year integer, year

val numeric, estimated values

upper numeric, estimated upper values

lower numeric, estimated lower values

Details

Locations available are Global, Canada, France, Germany, Italy, Japan, UK, and US.

Source

<https://vizhub.healthdata.org/gbd-results/>

`gbd_get_data`*Title: gbd_get_data - Fetch Data from GBD API*

Description

This function fetches data from the GBD API. To use this function, you need to have an API key. You can get the key by registering on the [IHME-API](#) website.

Usage

```
gbd_get_data(url, key, endpoint, ...)
```

Arguments

<code>url</code>	The base URL of the API.
<code>key</code>	The API key for authorization.
<code>endpoint</code>	The specific endpoint to retrieve data from.
<code>...</code>	Additional query parameters such as <code>location_id</code> , <code>year</code> , etc.

Value

A data frame or list of results from the API.

Examples

```
## Not run:
# This is a dontrun example because it requires an API KEY.
url <- "https://api.healthdata.org/sdg/v1"
key <- "YOUR-KEY"
endpoint <- "GetResultsByIndicator"

data <- gbd_get_data(url,
                    key,
                    endpoint,
                    indicator_id="1001",
                    location_id= c("29", "86", "102"),
                    year="2019",
                    limit = 10)

## End(Not run)
```

`germany_lungc`*Dataset: Health Metrics Data - Germany lungcancer Deaths 2019*

Description

A dataset containing deaths number due to lungcancer in Germany 2019.

Usage`germany_lungc`**Format**

A dataframe with 48 rows and 8 variables:

The variables are as follows:

age character, age groups from 10-14 to 85+ each 5 years

sex character, both, male, female

prevalence numeric, prevalence rate estimation due to lungcancer

prev_upper numeric, upper value estimation

prev_lower numeric, lower value estimation

dx numeric, deaths rate estimation due to lungcancer

dx_upper numeric, upper value estimation

dx_lower numeric, lower value estimation

Source

2019 data from the [IHME](#) website

`getunz`*Download, Unzip and Read Data: getunz*

Description

Download, Unzip and Read Data: getunz

Usage`getunz(url)`**Arguments**

`url` A url string for a .zip file.

Value

A dataframe object from a zipped file. Particular useful For downloading data from IHME GBD Results: "https://vizhub.healthdata.org/gbd-results/". The function takes the url, creates a temp directory, unzip the file, if more than one csv files is available, it lists the files, and reads them.

Select a dataset from the IHME GBD results and download it. You will receive an email with a url. Use the url to download the data.

Examples

```
## Not run:
# This is a dontrun example because it requires a valid url.
url <- "https://www.healthdata.org/.../some-file.zip"
getunz(url)

## End(Not run)
```

gho_le_hale

Dataset: Global Health Observatory (GHO) - Countries Life Expectancy and Healthy Life Expectancy(HALE) 2000-2019

Description

A dataset containing World countries Life Expectancy and HALE from 2000 to 2019.

Usage

```
gho_le_hale
```

Format

A dataframe with 8784 rows and 6 variables:

The variables are as follows:

indicator character, Healthy life expectancy (HALE) at age 60 (years), Healthy life expectancy (HALE) at birth (years), Life expectancy at age 60 (years), Life expectancy at birth (years)

year numeric, from 2000 to 2019

region character, 6 World regions: Africa, Americas, Eastern Mediterranean, Europe, South-East Asia, and Western Pacific

country character, 183 World countries

sex character, both, male, female

value numeric, value of the indicator

Source

WHO

gho_lifetables	<i>Dataset: Global Health Observatory (GHO) Life tables: WHO Global Life table values</i>
----------------	---

Description

A dataset containing the Global region Life tables from 2000 to 2019.

Usage

gho_lifetables

Format

A dataframe with 1995 rows and 5 variables:

The variables are as follows:

indicator character, Tx - person-years lived above age x, ex - expectation of life at age x, lx - number of people left alive at age x, nLx - person-years lived between ages x and x+n, nMx - age-specific death rate between ages x and x+n, ndx - number of people dying between ages x and x+n, nqx - probability of dying between ages x and x+n

year numeric, from 2000 to 2019

age character, from <1 to 85+ each 5 years

sex character, both, male, female

value numeric, value of the tables

Source

WHO

idDALY_map_data	<i>Dataset: Health Metrics Data - Simple Feature Collection Average Disability-Adjusted Life Years (DALYs) per 100,000 population from 1990 to 2021</i>
-----------------	---

Description

Dataset: Health Metrics Data - Simple Feature Collection Average Disability-Adjusted Life Years (DALYs) per 100,000 population from 1990 to 2021

Usage

idDALY_map_data

Format

A Simple feature collection with 1402 rows and 4 variables:

group double, country's polygon

location_name character, 200 Countries affected by Infectious Diseases

DALYs double, Average DALYs per 100,000 population from 1990 to 2021

geometry POLYGON

Source

2021 data from the [IHME](#) website

id_affected_countries *Dataset: Health Metrics Data - Infectious Diseases 1980-2021*

Description

A dataset containing average values for deaths rates, Disability-Adjusted Life Years (DALYs), Years of Life Lost (YLLs), and Years Lived with Disability (YLDs) due to 37 infectious diseases from 1980 to 2012 for all countries.

Usage

id_affected_countries

Format

A dataframe with 3066 rows and 6 variables:

The variables are as follows:

location_name character, list of countries

year numeric, from 1980 to 2021

DALYs numeric, DALYs for 100 000

YLLs numeric, YLLs for 100 000

YLDs numeric, YLDs for 100 000

Deaths numeric, deaths rate

Source

[IHME](#) website

incprev_stroke	<i>Global Region Health Metrics Data - Incidence and Prevalence for Stroke 2019 and 2021 Numbers - 5-year age groups from <1 to 85+ and both Location available Global</i>
----------------	---

Description

Global Region Health Metrics Data - Incidence and Prevalence for Stroke 2019 and 2021 Numbers - 5-year age groups from <1 to 85+ and both Location available Global

Usage

incprev_stroke

Format

A dataframe with 228 rows and 7 variables:

The variables are as follows:

measure character, metric name

sex character, female, male, both

age character, age groups from <1 to 85+ each 5 years

year integer, years 2019 and 2021

val numeric, estimated values

upper numeric, estimated upper values

lower numeric, estimated lower values

Source

<https://vizhub.healthdata.org/gbd-results/>

infectious_diseases	<i>Dataset: Health Metrics Data - Infectious Diseases 1980-2021</i>
---------------------	---

Description

A dataset containing Deaths rates, Disability-Adjusted Life Years (DALYs), Years of Life Lost (YLLs), and Years Lived with Disability (YLDs), Prevalence and Incidence due to Infectious Diseases form 1980 to 2021 for Lesotho, Eswatini, Malawi, Central African Republic, and Zambia.

Usage

infectious_diseases

Format

A dataframe with 7470 rows and 10 variables:

The variables are as follows:

year numeric, from 1980 to 2021

location_name character, list of countries

location_id numeric, list of countries by id

cause_name character, type of infectious disease

Deaths numeric, deaths rate

DALYs numeric, DALYs for 100 000

YLDs numeric, YLDs for 100 000

YLLs numeric, YLLs for 100 000

Prevalence numeric, prevalence rate

Incidence numeric, incidence rate

val numeric, estimated values

Source

[IHME](#) website

kbfir

Kriging Best Fit: kbfir - Fit variogram models and kriging models to spatial data and select the best model based on the metrics values

Description

Kriging Best Fit: kbfir - Fit variogram models and kriging models to spatial data and select the best model based on the metrics values

Usage

```
kbfir(response, formula, data, models, initial_values)
```

Arguments

response	A character string specifying the response variable
formula	A formula object specifying the model to fit: response ~ predictors
data	A simple feature object containing the variables in the formula
models	A list of characters vector specifying the variogram models to fit
initial_values	A list of named numeric vectors specifying the initial values for the variogram models: psill, range, nugget

Value

A list with two elements: `all_models` and `best_model`

Examples

```
## Not run:
# This is a dontrun example because it requires a spatial data object(data_sf).
# Try different initial values for fitting the variogram models
initial_values <- list(
  list(psill = 1, range = 100000, nugget = 10),
  list(psill = 0.5, range = 50000, nugget = 5),
  list(psill = 2, range = 150000, nugget = 15)
)

# Set some models to fit
models <- c("Sph", "Exp", "Gau")

# Select Best: Fit variogram models and kriging models
result <- hmsidwR::kbfit(response = "response",
  formula = response ~ predictor1 + predictor2,
  data = data_sf,
  models = c("Sph", "Exp", "Gau", "Mat"),
  initial_values = initial_values)

result$all_models
result$best_model

## End(Not run)
```

rabies

Dataset: Health Metrics Data - Rabies Deaths and DALYs from 1980 to 2021

Description

A subset of data from the IHME GBD on Disability-Adjusted Life Years (DALYs) and Deaths due to All Causes and Rabies. Locations available are Global Region and Asia.

Usage

```
rabies
```

Format

A dataframe with 296 rows and 7 variables:

The variables are as follows:

measure character, metric name

location character, country

cause character, cause
year integer, year
val numeric, estimated values
upper numeric, estimated upper values
lower numeric, estimated lower values

Source

<https://www.healthdata.org/>

sdi90_19	<i>Dataset: Health Metrics Data - Socio-Demographic Index (SDI) for 1990 and 2019</i>
----------	---

Description

A subset of data from the IHME GBD containing location, year and estimated values of the SDI for the years 1990 and 2019.

Usage

sdi90_19

Format

A dataframe with 20010 rows and 3 variables:

The variables are as follows:

location character, country
year integer, year
val numeric, estimated values

Source

<healthdata.org>

spatialdalys2021	<i>Health Metrics Data - Disability-Adjusted Life Years (DALYs) Estimations for 204 countries in 2021 with spatial information.</i>
------------------	---

Description

Health Metrics Data - Disability-Adjusted Life Years (DALYs) Estimations for 204 countries in 2021 with spatial information.

Usage

```
data(spatialdalys2021)
```

Format

A dataframe with 92862 rows and 7 variables:

The variables are as follows:

location character, France, Germany, Global, Italy, UK, USA, ...

value double, DALYs number estimation

lower_bound double, DALYs number estimation lower bound

upper_bound double, DALYs number estimation upper bound

long double, longitude

lat double, latitude

group double, polygons' group

Source

2021 data from the [IHME](#) website

Examples

```
data(spatialdalys2021)
head(spatialdalys2021)
```

string_search	<i>Scan all folders and files to find a string: string_search</i>
---------------	---

Description

Scan all folders and files to find a string: string_search

Usage

```
string_search(path = ".", pattern, string)
```

Arguments

path	If NULL, the current directory is used
pattern	A regular expression pattern such as '\\.R\$'
string	A string such as 'metric'

Value

A character vector with the names of the files that contain the string

Examples

```
string_search(path=".", "\\\.R$", "metric")  
  
# function string_search
```

theme_hmsid	<i>Custom ggplot2 theme function</i>
-------------	--------------------------------------

Description

Custom ggplot2 theme function

Usage

```
theme_hmsid(  
  base_size,  
  text_size,  
  subtitle_size,  
  subtitle_margin,  
  plot_title_size,  
  plot_title_margin,  
  ...  
)
```


Arguments

<code>base_size</code>	base font size
<code>text_size</code>	plot text size
<code>subtitle_size, subtitle_margin</code>	plot subtitle size and margin
<code>plot_title_size, plot_title_margin</code>	plot title size and margin
<code>...</code>	Other arguments passed to <code>theme_hmsid</code>

Value

A customized theme for a ggplot object.

Examples

```
library(ggplot2)
dat <- data.frame(
  x = seq_along(1:5),
  y = rnorm(n = 5, mean = 0.5, sd = 1)
)
dat |>
  ggplot(aes(x = x, y = y)) +
  geom_line() +
  hmsidwR::theme_hmsid()
```

Index

* datasets

- deaths2019, [2](#)
- deaths9, [3](#)
- disweights, [4](#)
- g7_hmetrics, [5](#)
- germany_lungc, [7](#)
- gho_le_hale, [8](#)
- gho_lifetables, [9](#)
- id_affected_countries, [10](#)
- idDALY_map_data, [9](#)
- incprev_stroke, [11](#)
- infectious_diseases, [11](#)
- rabies, [13](#)
- sdi90_19, [14](#)
- spatialdalys2021, [15](#)

- deaths2019, [2](#)
- deaths9, [3](#)
- disweights, [4](#)

- g7_hmetrics, [5](#)
- gbd_get_data, [6](#)
- germany_lungc, [7](#)
- getunz, [7](#)
- gho_le_hale, [8](#)
- gho_lifetables, [9](#)

- id_affected_countries, [10](#)
- idDALY_map_data, [9](#)
- incprev_stroke, [11](#)
- infectious_diseases, [11](#)

- kbfit, [12](#)

- rabies, [13](#)

- sdi90_19, [14](#)
- spatialdalys2021, [15](#)
- string_search, [16](#)

- theme_hmsid, [16](#)