Package: mwana (via r-universe)

November 19, 2024

Type Package

Title An Efficient Workflow for Plausibility Checks and Prevalence Analysis of Wasting in R

Version 0.2.0

Description A simple and streamlined workflow for plausibility checks and prevalence analysis of wasting based on the Standardized Monitoring and Assessment of Relief and Transition (SMART) Methodology <`https://smartmethodology.org/`>, with application in R.

License GPL (>= 3)

URL `https://github.com/nutriverse/mwana`, `https://nutriverse.io/mwana`

Imports dplyr (>= 1.1.4), lubridate, nipnTK, rlang, scales, srvyr, stats, zscorer, tibble, methods

Suggests knitr, rmarkdown, quarto, spelling, testthat (>= 3.0.0)

Config/testthat/edition 3

Encoding UTF-8

Language en-GB

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Depends R (>= 4.1)

LazyData true

VignetteBuilder quarto

Config/pak/sysreqs make libicu-dev zlib1g-dev

Repository https://nutriverse.r-universe.dev

RemoteUrl https://github.com/nutriverse/mwana

RemoteRef HEAD

RemoteSha e447f17aa0c83a87162e325479e15c1969a0b432

Contents

`anthro.01` *A sample data of district level SMART surveys with location anonymised*

Description

`anthro.01` is a two-stage cluster-based survey with probability of selection of clusters proportional to the size of the population. The survey employed the SMART methodology.

Usage

```
anthro.01
```

Format

A tibble of 1,191 rows and 11 columns.

Variable	**Description**
area	Location where the survey took place
dos	Survey date
cluster	Primary sampling unit
team	Enumerator IDs
sex	Sex, "m" = boys, "f" = girls
dob	Date of birth
age	Age in months, typically estimated using local event calendars
weight	Weight (kg)
height	Height (cm)
edema	Edema, "n" = no, "y" = yes
muac	Mid-upper arm circumference (mm)

Source

Anonymous

Examples

```
anthro.01
```

`anthro.02` *A sample of an already wrangled survey data*

Description

A household budget survey data conducted in Mozambique in 2019/2020, known as *IOF* (*Inquérito ao Orçamento Familiar* in Portuguese). *IOF* is a two-stage cluster-based survey, representative at province level (admin 2), with probability of the selection of the clusters proportional to the size of the population. Its data collection spans for a period of 12 months.

Usage

```
anthro.02
```

Format

A tibble of 2,267 rows and 14 columns.

Variable	**Description**
province	The administrative unit (admin 1) where data was collected.

strata	Rural and Urban
cluster	Primary sampling unit
sex	Sex, "m" = boys, "f" = girls
age	calculated age in months with two decimal places
weight	Weight (kg)
height	Height (cm)
edema	Edema, "n" = no, "y" = yes
muac	Mid-upper arm circumference (mm)
wtfactor	Survey weights
wfhz	Weight-for-height z-scores with 3 decimal places
flag_wfhz	Flagged observations. 1=flagged, 0=not flagged
mfaz	MUAC-for-age z-scores with 3 decimal places
flag_mfaz	Flagged observations. 1=flagged, 0=not flagged

Source

Mozambique National Institute of Statistics. The data is publicly available at `https://mozdata.ine.gov.mz/index.php/catalog/88#metadata-data_access`. Data was wrangled using this package's wranglers. Details about survey design can be gotten from: `https://mozdata.ine.gov.mz/index.php/catalog/88#metadata-sampling`

Examples

```
anthro.02
```

`anthro.03` *A sample data of district level SMART surveys conducted in Mozambique*

Description

`anthro.03` contains survey data of four districts. Each district data set presents distinct data quality scenarios that requires tailored prevalence analysis approach: two districts show a problematic WFHZ standard deviation whilst the remaining are all within range.

This sample data is useful to demonstrate the use of the prevalence functions on a multiple-area survey data where there can be variations in the rating of acceptability of the standard deviation, hence require different analyses approaches for each area to ensure accurate estimation.

Usage

```
anthro.03
```

Format

A tibble of 943 x 9.

Variable	**Description**
district	The location where data was collected
cluster	Primary sampling unit
team	Survey teams
sex	Sex, "m" = boys, "f" = girls
age	calculated age in months with two decimal places
weight	Weight (kg)
height	Height (cm)
edema	Edema, "n" = no, "y" = yes
muac	Mid-upper arm circumference (mm)

Source

Anonymous

Examples

```
anthro.03
```

`anthro.04` *A sample data of a community-based sentinel site from an anonymized location*

Description

Data was generated through a community-based sentinel site conducted across three provinces. Each province's data set presents distinct data quality scenarios, requiring tailored prevalence analysis:

- "Province 1" has MFAZ's standard deviation and age ratio test rating of acceptability falling within range;
- "Province 2" has age ratio rated as problematic but with an acceptable standard deviation of MFAZ;
- "Province 3" has both tests rated as problematic.

This sample data is useful to demonstrate the use of prevalence functions on a multiple-area survey data where variations in the rating of acceptability of the standard deviation exist, hence require different analyses approaches for each area to ensure accurate estimation.

Usage

```
anthro.04
```

Format

A tibble of 3,002 x 8.

Variable	**Description**
province	location where data was collected
cluster	Primary sampling unit
sex	Sex, "m" = boys, "f" = girls
age	calculated age in months with two decimal places
muac	Mid-upper arm circumference (mm)
edema	Edema, "n" = no, "y" = yes
mfaz	MUAC-for-age z-scores with 3 decimal places
flag_mfaz	Flagged observations. 1=flagged, 0=not flagged

Source

Anonymous

Examples

```
anthro.04
```

`define_wasting` *Define wasting*

Description

Define if a given observation in the data set is wasted or not, and its respective form of wasting (global, severe or moderate) on the basis of z-scores of weight-for-height (WFHZ), muac-for-age (MFAZ), raw MUAC values and combined case-definition.

Usage

```
define_wasting(
  df,
  zscores = NULL,
  muac = NULL,
  edema = NULL,
  .by = c("zscores", "muac", "combined")
)
```

Arguments

`df`	A data set object of class `data.frame` to use. It must have been wrangled using this package's wrangling functions for WFHZ or MUAC, or both (for combined) as appropriate.
`zscores`	A vector of class `double` of WFHZ or MFAZ values. If the class does not match the expected type, the function will stop execution and return an error message indicating the type of mismatch.
`muac`	A vector of class `integer` or `numeric` of raw MUAC values in millimeters. If the class does not match the expected type, the function will stop execution and return an error message indicating the type of mismatch.
`edema`	A vector of class `character` of edema. Default is `NULL`. If the class does not match the expected type, the function will stop execution and return an error message indicating the type of mismatch. Code values should be "y" for presence and "n" for absence of bilateral edema. If different, the function will stop execution and return an error indicating the issue.
`.by`	A choice of the criterion by which the case-definition should done. Choose `zscores` for WFHZ or MFAZ, `muac` for raw MUAC and `combined` for combined.

Value

Three vectors named `gam`, `sam` and `mam`, of class `numeric`, same length as inputs, containing dummy values: 1 for case and 0 for not case. This is added to `df`. When `combined` is selected, vector's names become `cgam`, `csam` and `cmam`.

Examples

```
## Case-definition by z-scores ----
z <- anthro.02 |>
  define_wasting(
    zscores = wfhz,
    muac = NULL,
    edema = edema,
    .by = "zscores"
  )
head(z)

## Case-definition by MUAC ----
m <- anthro.02 |>
  define_wasting(
    zscores = NULL,
    muac = muac,
    edema = edema,
    .by = "muac"
  )
head(m)

## Case-definition by combined ----
c <- anthro.02 |>
```

```
  define_wasting(
    zscores = wfhz,
    muac = muac,
    edema = edema,
    .by = "combined"
  )
head(c)
```

`flag_outliers` *Identify, flag outliers and remove them*

Description

Identify outlier z-scores for weight-for-height (WFHZ) and MUAC-for-age (MFAZ) following the SMART methodology. The function can also be used to detect outliers for height-for-age (HFAZ) and weight-for-age (WFAZ) z-scores following the same approach.

For raw MUAC values, outliers constitute values that are less than 100 millimeters or greater than 200 millimeters.

Removing outliers consist in setting the outlier record to `NA` and not necessarily to delete it from the data set. This is useful in the analysis procedures where outliers must be removed, such as the analysis of the standard deviation.

Usage

```
flag_outliers(x, .from = c("zscores", "raw_muac"))

remove_flags(x, .from = c("zscores", "raw_muac"))
```

Arguments

`x`	A vector of class `numeric` of WFHZ, MFAZ, HFAZ, WFAZ or raw MUAC values. The latter should be in millimeters. If the class is different than expected, the function will stop execution and return an error message indicating the type of mismatch.
`.from`	A choice between `zscores` and `raw_muac` for where outliers should be detected and flagged from.

Details

For z-score-based detection, flagged records represent outliers that deviate substantially from the sample's z-score mean, making them unlikely to reflect accurate measurements. For raw MUAC values, flagged records are those that fall outside the acceptable fixed range. Including such outliers in the analysis could compromise the accuracy and precision of the resulting estimates.

The flagging criterion used for raw MUAC values is based on a recommendation by Bilukha, O., & Kianian, B. (2023).

Value

A vector of the same length as `x` for flagged records coded as `1` for is a flag and `0` not a flag.

References

Bilukha, O., & Kianian, B. (2023). Considerations for assessment of measurement quality of mid-upper arm circumference data in anthropometric surveys and mass nutritional screenings conducted in humanitarian and refugee settings. *Maternal & Child Nutrition*, 19, e13478. Available at `https://doi.org/10.1111/mcn.13478`

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: `https://smartmethodology.org`.

Examples

```
## Sample data of raw MUAC values ----
x <- anthro.01$muac

## Apply the function with `.from` set to "raw_muac" ----
m <- flag_outliers(x, .from = "raw_muac")
head(m)

## Sample data of z-scores (be it WFHZ, MFAZ, HFAZ or WFAZ) ----
x <- anthro.02$mfaz

# Apply the function with `.from` set to "zscores" ----
z <- flag_outliers(x, .from = "zscores")
tail(z)

## With `.from` set to "zscores" ----
z <- remove_flags(
  x = wfhz.01$wfhz,
  .from = "zscores"
)

head(z)

## With `.from` set to "raw_muac" ----
m <- remove_flags(
  x = mfaz.01$muac,
  .from = "raw_muac"
)

tail(m)
```

`get_age_months` *Calculate child's age in months*

Description

Calculate child's age in months based on the date of birth and the date of data collection.

Usage

```
get_age_months(dos, dob)
```

Arguments

`dos`	A vector of class `Date` for the date of data collection. If the class is different than expected, the function will stop execution and return an error message indicating the type of mismatch.
`dob`	A vector of class `Date` for the child's date of birth. If the class is different than expected, the function will stop execution and return an error message indicating the type of mismatch.

Value

A vector of class `numeric` for child's age in months. Any value less than 6.0 and greater than or equal to 60.0 months will be set to `NA`.

Examples

```
## Take two vectors of class "Date" ----
surv_date <- as.Date(
  c(
    "2024-01-05", "2024-01-05", "2024-01-05", "2024-01-08", "2024-01-08",
    "2024-01-08", "2024-01-10", "2024-01-10", "2024-01-10", "2024-01-11"
  )
)
bir_date <- as.Date(
  c(
    "2022-04-04", "2021-05-01", "2023-05-24", "2017-12-12", NA,
    "2020-12-12", "2022-04-04", "2021-05-01", "2023-05-24", "2020-12-12"
  )
)

## Apply the function ----
get_age_months(
  dos = surv_date,
  dob = bir_date
)
```

`mfaz.01` *A sample MUAC screening data from an anonymized setting*

Description

A sample MUAC screening data from an anonymized setting

Usage

```
mfaz.01
```

Format

A tibble with 661 rows and 4 columns.

Variable	**Description**
sex	Sex, "m" = boys, "f" = girls
months	calculated age in months with two decimal places
edema	Edema, "n" = no, "y" = yes
muac	Mid-upper arm circumference (mm)

Source

Anonymous

Examples

```
mfaz.01
```

`mfaz.02` *A sample SMART survey data with MUAC*

Description

A sample SMART survey data with MUAC

Usage

```
mfaz.02
```

Format

A tibble with 303 rows and 7 columns.

Variable	**Description**
cluster	Primary sampling unit
sex	Sex, "m" = boys, "f" = girls
age	calculated age in months with two decimal places
edema	Edema, "n" = no, "y" = yes
mfaz	MUAC-for-age z-scores with 3 decimal places
flag_mfaz	Flagged observations. 1=flagged, 0=not flagged

Source

Anonymous

Examples

```
mfaz.02
```

`mw_check_ipcamn_ssreq` *Check whether IPC Acute Malnutrition (IPC AMN) sample size requirements were met*

Description

Evidence on the prevalence of acute malnutrition used in the IPC AMN can come from different sources: surveys, screenings or community-based surveillance system. The IPC set minimum sample size requirements for each source. This function helps in verifying whether those requirements were met or not depending on the source.

Usage

```
mw_check_ipcamn_ssreq(df, cluster, .source = c("survey", "screening", "ssite"))
```

Arguments

`df`	A data set object of class `data.frame` to check.
`cluster`	A vector of class `integer` or `character` of unique cluster or screening or sentinel site IDs. If a `character` vector, ensure that names are correct and each name represents one location for accurate counts. If the class does not match the above expected type, the function will stop execution and return an error message indicating the type of mismatch.
`.source`	The source of evidence. A choice between "survey" for representative survey data at the area of analysis; "screening" for screening data; "ssite" for community-based sentinel site data.

Value

A summary table of class `data.frame`, of length 3 and width 1, for the check results. `n_clusters` is for the total number of unique clusters or screening or site IDs; `n_obs` for the correspondent total number of children in the data set; and `meet_ipc` for whether the IPC AMN requirements were met.

References

IPC Global Partners. 2021. *Integrated Food Security Phase Classification Technical Manual Version 3.1.Evidence and Standards for Better Food Security and Nutrition Decisions*. Rome. Available at: `https://www.ipcinfo.org/ipcinfo-website/resources/ipc-manual/en/`.

Examples

```
mw_check_ipcamn_ssreq(
  df = anthro.01,
  cluster = cluster,
  .source = "survey"
)
```

`mw_estimate_prevalence_combined`

Estimate the prevalence of combined wasting

Description

Estimate the prevalence of wasting based on the combined case-definition of weight-for-height z-scores (WFHZ), MUAC and/or edema. The function allows users to get the prevalence estimates in accordance with the complex sample design properties; this includes applying survey weights when needed or applicable. Before estimating, the function evaluates the quality of data by calculating and rating the standard deviation of WFHZ and MFAZ, as well as the p-value of the age ratio test. Prevalence will be calculated only when the rating of all test is as not problematic concurrently. If either of them is problematic, it cancels out the analysis and `NA`s get thrown.

Outliers are detected in both WFHZ and in MUAC data set (through z-scores) based on SMART flags get excluded prior being piped into the actual prevalence analysis workflow.

Usage

```
mw_estimate_prevalence_combined(df, wt = NULL, edema = NULL, .by = NULL)
```

Arguments

`df`	A data set object of class `data.frame` to use. This must have been wrangled using this package's wrangling functions for both WFHZ and MUAC data sequentially. The order does not matter. Note that MUAC values should be converted to millimeters after using the MUAC wrangler. If this is not done, the function will stop execution and return an error message. Moreover, the function uses a

	variable called `cluster` where the primary sampling unit IDs are stored. Make sure to rename your cluster ID variable to `cluster`, otherwise the function will error and terminate the execution.
`wt`	A vector of class `double` of the final survey weights. Default is `NULL` assuming a self-weighted survey, as in the ENA for SMART software; otherwise a weighted analysis is computed.
`edema`	A vector of class `character` of edema. Code will be "y" for presence and "n" for absence of bilateral edema. Default is `NULL`.
`.by`	A vector of class `character` or `numeric` of the geographical areas or respective IDs for where the data was collected and for which the analysis should be summarised at.

Details

A concept of "combined flags" is introduced in this function. It consists of defining as flag any observation that is flagged in either `flag_wfhz` or `flag_mfaz` vectors. A new column `cflags` for combined flags is created and added to `df`. This ensures that all flagged observations from both WFHZ and MFAZ data are excluded from the prevalence analysis.

A glimpse on how `cflags` *are defined:*

flag_wfhz	**flag_mfaz**	**cflags**
1	0	1
0	1	1
0	0	0

Value

A summarised table of class `data.frame` for the descriptive statistics about combined wasting.

Examples

```
## When .by and wt are set to NULL ----
mw_estimate_prevalence_combined(
  df = anthro.02,
  wt = NULL,
  edema = edema,
  .by = NULL
)

## When wt is not set to NULL ----
mw_estimate_prevalence_combined(
  df = anthro.02,
  wt = wtfactor,
  edema = edema,
  .by = NULL
)
```

`mw_estimate_prevalence_mfaz` *Estimate the prevalence of wasting based on z-scores of muac-for-age (MFAZ)*

Description

Calculate the prevalence estimates of wasting based on z-scores of muac-for-age and/or bilateral edema. The function allows users to get the prevalence estimates calculated in accordance with the complex sample design properties; this includes applying survey weights when needed or applicable.

Before estimating, the function evaluates the quality of data by calculating and rating the standard deviation of z-scores of MFAZ. If rated as problematic, the prevalence is estimated based on the PROBIT method.

Outliers are detected based on SMART flags and get excluded prior prevalence analysis.

Usage

```
mw_estimate_prevalence_mfaz(df, wt = NULL, edema = NULL, .by = NULL)
```

Arguments

`df`	A data set object of class `data.frame` to use. This must have been wrangled using this package's wrangling function for MUAC data. The function uses a variable name called `cluster` where the primary sampling unit IDs are stored. Make sure to rename your cluster ID variable to `cluster`, otherwise the function will error and terminate the execution.
`wt`	A vector of class `double` of the final survey weights. Default is `NULL` assuming a self weighted survey, as in the ENA for SMART software; otherwise, when a vector of weights if supplied, weighted analysis is done.
`edema`	A vector of class `character` of edema. Code should be "y" for presence and "n" for absence of bilateral edema. Default is `NULL`.
`.by`	A vector of class `character` or `numeric` of the geographical areas or respective IDs for where the data was collected and for which the analysis should be summarized at.

Value

A summarized table of class `data.frame` of the descriptive statistics about wasting.

Examples

```
## When .by = NULL ----
mw_estimate_prevalence_mfaz(
  df = anthro.04,
  wt = NULL,
```

```
  edema = edema,
  .by = NULL
)

## When .by is not set to NULL ----
mw_estimate_prevalence_mfaz(
  df = anthro.04,
  wt = NULL,
  edema = edema,
  .by = province
)
```

---

mw_estimate_prevalence_muac

Estimate the prevalence of wasting based on MUAC for survey data

---

Description

Calculate the prevalence estimates of wasting based on MUAC and/or bilateral edema. Before estimating, the function evaluates the quality of data by calculating and rating the standard deviation of z-scores of muac-for-age (MFAZ) and the p-value of the age ratio test; then it sets the analysis path that best fits the data:

- If all tests are rated as not problematic, a normal analysis is done.
- If standard deviation is not problematic and age ratio test is problematic, prevalence is age-weighted. This is to fix the likely overestimation of wasting when there are excess of younger children in the data set.
- If standard deviation is problematic and age ratio test is not, or both are problematic, analysis gets cancelled out and `NA`s get thrown.

Outliers are detected based on SMART flags on the MFAZ values and then get excluded prior being piped into the actual prevalence analysis workflow.

Usage

```
mw_estimate_prevalence_muac(df, wt = NULL, edema = NULL, .by = NULL)

mw_estimate_smart_age_wt(df, edema = NULL, .by = NULL)
```

Arguments

`df`	A data set object of class `data.frame` to use. This must have been wrangled using this package's wrangling function for MUAC data. Make sure MUAC values are converted to millimeters after using the wrangler. If this is not done, the function will stop execution and return an error message. The function uses a variable name called `cluster` where the primary sampling unit IDs are stored. Make sure the data set has this variable and its name has been renamed to `cluster`, otherwise the function will error and terminate the execution.

`wt`	A vector of class `double` of the final survey weights. Default is `NULL` assuming a self weighted survey, as in the ENA for SMART software; otherwise, when a vector of weights if supplied, weighted analysis is done.
`edema`	A vector of class `character` of edema. Code should be "y" for presence and "n" for absence of bilateral edema. Default is `NULL`.
`.by`	A vector of class `character` or `numeric` of the geographical areas or respective IDs for where the data was collected and for which the analysis should be summarized at.

Value

A summarized table of class `data.frame` of the descriptive statistics about wasting.

References

SMART Initiative (no date). *Updated MUAC data collection tool*. Available at: `https://smartmethodology.org/survey-planning-tools/updated-muac-tool/`

See Also

`mw_estimate_smart_age_wt()` `mw_estimate_prevalence_mfaz()` `mw_estimate_prevalence_screening()`

Examples

```
## When .by = NULL ----
mw_estimate_prevalence_muac(
  df = anthro.04,
  wt = NULL,
  edema = edema,
  .by = NULL
)

## When .by is not set to NULL ----
mw_estimate_prevalence_muac(
  df = anthro.04,
  wt = NULL,
  edema = edema,
  .by = province
)

## An application of `mw_estimate_smart_age_wt()` ----
.data <- anthro.04 |>
  subset(province == "Province 2")

mw_estimate_smart_age_wt(
  df = .data,
  edema = edema,
  .by = NULL
)
```

`mw_estimate_prevalence_screening` *Estimate the prevalence of wasting based on MUAC for non survey data*

Description

It is common to estimate prevalence of wasting from non survey data, such as screenings or any other community-based surveillance systems. In such situations, the analysis usually consists only in estimating the point prevalence and the counts of positive cases, without necessarily estimating the uncertainty. This is the job of this function.

Before estimating, it evaluates the quality of data by calculating and rating the standard deviation of z-scores of muac-for-age (MFAZ) and the p-value of the age ratio test; then it sets the analysis path that best fits the data.

- If all tests are rated as not problematic, a normal analysis is done.
- If standard deviation is not problematic and age ratio test is problematic, prevalence is age-weighted. This is to fix the likely overestimation of wasting when there are excess of younger children in the data set.
- If standard deviation is problematic and age ratio test is not, or both are problematic, analysis gets cancelled out and `NA`s get thrown.

Outliers are detected based on SMART flags on the MFAZ values and then get excluded prior being piped into the actual prevalence analysis workflow.

Usage

```
mw_estimate_prevalence_screening(df, muac, edema = NULL, .by = NULL)
```

Arguments

`df`	A data set object of class `data.frame` to use. This must have been wrangled using this package's wrangling function for MUAC data. Make sure MUAC values are converted to millimeters after using the wrangler. If this is not done, the function will stop execution and return an error message with the issue.
`muac`	A vector of raw MUAC values of class `numeric` or `integer`. The measurement unit of the values should be millimeters. If any or all values are in a different unit than the expected, the function will stop execution and return an error message indicating the issue.
`edema`	A vector of class `character` of edema. Code should be "y" for presence and "n" for absence of bilateral edema. Default is `NULL`. If class, as well as, code values are different than expected, the function will stop the execution and return an error message indicating the issue.
`.by`	A vector of class `character` or `numeric` of the geographical areas or respective IDs for where the data was collected and for which the analysis should be summarized at.

Value

A summarized table of class `data.frame` of the descriptive statistics about wasting.

References

SMART Initiative (no date). *Updated MUAC data collection tool*. Available at: https://smartmethodology.org/survey-planning-tools/updated-muac-tool/

See Also

`mw_estimate_prevalence_muac()` `mw_estimate_smart_age_wt()`

Examples

```
mw_estimate_prevalence_screening(
  df = anthro.02,
  muac = muac,
  edema = edema,
  .by = province
)

## With `edema` set to `NULL` ----
mw_estimate_prevalence_screening(
  df = anthro.02,
  muac = muac,
  edema = NULL,
  .by = province
)

## With `.by` set to `NULL` ----
mw_estimate_prevalence_screening(
  df = anthro.02,
  muac = muac,
  edema = NULL,
  .by = NULL
)
```

`mw_estimate_prevalence_wfhz`

Estimate the prevalence of wasting based on z-scores of weight-for-height (WFHZ)

Description

Calculate the prevalence estimates of wasting based on z-scores of weight-for-height and/or bilateral edema. The function allows users to get the prevalence estimates calculated in accordance with the complex sample design properties; this includes applying survey weights when needed or applicable.

Before estimating, the function evaluates the quality of data by calculating and rating the standard deviation of z-scores of WFHZ. If rated as problematic, the prevalence is estimated based on the PROBIT method.

Outliers are detected based on SMART flags and get excluded prior being piped into the actual prevalence analysis workflow.

Usage

```
mw_estimate_prevalence_wfhz(df, wt = NULL, edema = NULL, .by = NULL)
```

Arguments

`df`	A data set object of class `data.frame` to use. This must have been wrangled using this package's wrangling function for WFHZ data. The function uses a variable name called `cluster` where the primary sampling unit IDs are stored. Make sure to rename your cluster ID variable to `cluster`, otherwise the function will error and terminate the execution.
`wt`	A vector of class `double` of the final survey weights. Default is `NULL` assuming a self weighted survey, as in the ENA for SMART software; otherwise, when a vector of weights if supplied, weighted analysis is done.
`edema`	A vector of class `character` of edema. Code should be "y" for presence and "n" for absence of bilateral edema. Default is `NULL`.
`.by`	A vector of class `character` or `numeric` of the geographical areas or respective IDs for where the data was collected and for which the analysis should be summarised at.

Value

A summarised table of class `data.frame` of the descriptive statistics about wasting.

Examples

```
## When .by = NULL ----
### Start off by wrangling the data ----
data <- mw_wrangle_wfhz(
  df = anthro.03,
  sex = sex,
  weight = weight,
  height = height,
  .recode_sex = TRUE
)

### Now run the prevalence function ----
mw_estimate_prevalence_wfhz(
  df = data,
  wt = NULL,
  edema = edema,
  .by = NULL
)
```

```
## Now when .by is not set to NULL ----
mw_estimate_prevalence_wfhz(
  df = data,
  wt = NULL,
  edema = edema,
  .by = district
)

## When a weighted analysis is needed ----
mw_estimate_prevalence_wfhz(
  df = anthro.02,
  wt = wtfactor,
  edema = edema,
  .by = province
)
```

`mw_neat_output_mfaz` *Clean and format the output table returned from the MFAZ plausibility check for improved clarity and readability*

Description

Clean and format the output table returned from the MFAZ plausibility check for improved clarity and readability. It converts scientific notations to standard notations, round values and rename columns to meaningful names.

Usage

```
mw_neat_output_mfaz(df)
```

Arguments

`df`	An object of class `data.frame` returned by this package's plausibility checker for MFAZ data, containing the summarized results to be formatted.

Value

A `data.frame` object of the same length and width as `df`, with column names and values formatted for clarity and readability.

Examples

```
## First wrangle age data ----
data <- mw_wrangle_age(
  df = anthro.01,
  dos = dos,
  dob = dob,
  age = age,
```

```
  .decimals = 2
)

## Then wrangle MUAC data ----
data_mfaz <- mw_wrangle_muac(
  df = data,
  sex = sex,
  age = age,
  muac = muac,
  .recode_sex = TRUE,
  .recode_muac = TRUE,
  .to = "cm"
)

## Then run plausibility check ----
pl <- mw_plausibility_check_mfaz(
  df = data_mfaz,
  flags = flag_mfaz,
  sex = sex,
  muac = muac,
  age = age
)

## Now neat the output table ----
mw_neat_output_mfaz(df = pl)
```

`mw_neat_output_muac` *Clean and format the output table returned from the MUAC plausibility check for improved clarity and readability.*

Description

Clean and format the output table returned from the plausibility check of raw MUAC data for improved clarity and readability. It converts scientific notations to standard notations, round values and rename columns to meaningful names.

Usage

```
mw_neat_output_muac(df)
```

Arguments

`df`	An object of class `data.frame` returned by this package's plausibility checker for raw MUAC data, containing the summarized results to be formatted.

Value

A `data.frame` object of the same length and width as `df`, with column names and values formatted for clarity and readability.

Examples

```
## First wranlge MUAC data ----
df_muac <- mw_wrangle_muac(
  df = anthro.01,
  sex = sex,
  muac = muac,
  age = NULL,
  .recode_sex = TRUE,
  .recode_muac = FALSE,
  .to = "none"
)

## Then run the plausibility check ----
pl_muac <- mw_plausibility_check_muac(
  df = df_muac,
  flags = flag_muac,
  sex = sex,
  muac = muac
)

## Neat the output table ----

mw_neat_output_muac(df = pl_muac)
```

`mw_neat_output_wfhz` *Clean and format the output table returned from the WFHZ plausibility check for improved clarity and readability*

Description

Clean and format the output table returned from the WFHZ plausibility check for improved clarity and readability. It converts scientific notations to standard notations, round values and rename columns to meaningful names.

Usage

```
mw_neat_output_wfhz(df)
```

Arguments

`df`	An object of class `data.frame` returned by this package's plausibility checker for WFHZ data, containing the summarized results to be formatted.

Value

A `data.frame` object of the same length and width as `df`, with column names and values formatted for clarity and readability.

Examples

```
## First wrangle age data ----
data <- mw_wrangle_age(
  df = anthro.01,
  dos = dos,
  dob = dob,
  age = age,
  .decimals = 2
)

## Then wrangle WFHZ data ----
data_wfhz <- mw_wrangle_wfhz(
  df = data,
  sex = sex,
  weight = weight,
  height = height,
  .recode_sex = TRUE
)

## Now run the plausibility check ----
pl <- mw_plausibility_check_wfhz(
  df = data_wfhz,
  sex = sex,
  age = age,
  weight = weight,
  height = height,
  flags = flag_wfhz
)

## Now neat the output table ----
mw_neat_output_wfhz(df = pl)
```

`mw_plausibility_check_mfaz`

Check the plausibility and acceptability of muac-for-age z-score (MFAZ) data

---

Description

Check the overall plausibility and acceptability of MFAZ data through a structured test suite encompassing sampling and measurement-related biases checks in the data set. The test suite in this function follows the recommendation made by Bilukha, O., & Kianian, B. (2023) on the plausibility of constructing a comprehensive plausibility check for MUAC data similar to WFHZ to evaluate its acceptability when the variable age exists in the data set.

The function works on a data frame returned from this package's wrangling function for age and for MFAZ data.

Usage

```
mw_plausibility_check_mfaz(df, sex, muac, age, flags)
```

Arguments

`df`	A data set object of class `data.frame` to check.
`sex`	A vector of class `numeric` of child's sex.
`muac`	A vector of class `numeric` of child's MUAC in centimeters.
`age`	A vector of class `double` of child's age in months.
`flags`	A vector of class `numeric` of flagged records.

Details

Whilst the function uses the same test checks and criteria as that of WFHZ in the SMART plausibility check, the percent of flagged data is evaluated using a different cut-off points, with a maximum acceptability of 2.0%, as shown below:

Excellent	**Good**	**Acceptable**	**Problematic**
0.0 - 1.0	>1.0 - 1.5	>1.5 - 2.0	>2.0

Value

A summarized table of class `data.frame`, of length 17 and width 1, for the plausibility test results and their respective acceptability ratings.

References

Bilukha, O., & Kianian, B. (2023). Considerations for assessment of measurement quality of mid-upper arm circumference data in anthropometric surveys and mass nutritional screenings conducted in humanitarian and refugee settings. *Maternal & Child Nutrition*, 19, e13478. `https://doi.org/10.1111/mcn.13478`

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: `https://smartmethodology.org`.

See Also

`mw_wrangle_age()` `mw_wrangle_muac()` `mw_stattest_ageratio()` `flag_outliers()`

Examples

```
## First wrangle age data ----
data <- mw_wrangle_age(
  df = anthro.01,
  dos = dos,
  dob = dob,
  age = age,
  .decimals = 2
```

```
)

## Then wrangle MUAC data ----
data_muac <- mw_wrangle_muac(
  df = data,
  sex = sex,
  age = age,
  muac = muac,
  .recode_sex = TRUE,
  .recode_muac = TRUE,
  .to = "cm"
)

## And finally run plausibility check ----
mw_plausibility_check_mfaz(
  df = data_muac,
  flags = flag_mfaz,
  sex = sex,
  muac = muac,
  age = age
)
```

`mw_plausibility_check_muac` *Check the plausibility and acceptability of raw MUAC data*

Description

Check the overall plausibility and acceptability of raw MUAC data through a structured test suite encompassing sampling and measurement-related biases checks in the data set. The test suite in this function follows the recommendation made by Bilukha, O., & Kianian, B. (2023).

Usage

```
mw_plausibility_check_muac(df, sex, muac, flags)
```

Arguments

`df`	An object of class `data.frame` to check. It must have been wrangled using this package's wrangling function for MUAC.
`sex`	A vector of class `numeric` of child's sex.
`muac`	A vector of class `double` of child's MUAC in centimeters.
`flags`	A vector of class `numeric` of flagged records.

Details

Cut-off points used for the percent of flagged records:

Excellent	**Good**	**Acceptable**	**Problematic**
0.0 - 1.0	>1.0 - 1.5	>1.5 - 2.0	>2.0

Value

A summarized table of class `data.frame`, of length 9 and width 1, for the plausibility test results and their respective acceptability ratings.

References

Bilukha, O., & Kianian, B. (2023). Considerations for assessment of measurement quality of mid-upper arm circumference data in anthropometric surveys and mass nutritional screenings conducted in humanitarian and refugee settings. *Maternal & Child Nutrition*, 19, e13478. `https://doi.org/10.1111/mcn.13478`

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: `https://smartmethodology.org`.

See Also

`mw_wrangle_muac()` `flag_outliers()`

Examples

```
## First wranlge MUAC data ----
df_muac <- mw_wrangle_muac(
  df = anthro.01,
  sex = sex,
  muac = muac,
  age = NULL,
  .recode_sex = TRUE,
  .recode_muac = FALSE,
  .to = "none"
)

## Then run the plausibility check ----
mw_plausibility_check_muac(
  df = df_muac,
  flags = flag_muac,
  sex = sex,
  muac = muac
)
```

`mw_plausibility_check_wfhz` *Check the plausibility and acceptability of weight-for-height z-score (WFHZ) data*

Description

Check the overall plausibility and acceptability of WFHZ data through a structured test suite encompassing sampling and measurement-related biases checks in the data set. The test suite, including the criteria and corresponding rating of acceptability, follows the standards in the SMART plausibility check. The only exception is the exclusion of MUAC checks. MUAC is checked separately using more comprehensive test suite as well.

The function works on a data frame returned from this package's wrangling function for age and for WFHZ data.

Usage

```
mw_plausibility_check_wfhz(df, sex, age, weight, height, flags)
```

Arguments

`df`	A data set object of class `data.frame` to check.
`sex`	A vector of class `numeric` of child's sex.
`age`	A vector of class `double` of child's age in months.
`weight`	A vector of class `double` of child's weight in kilograms.
`height`	A vector of class `double` of child's height in centimeters.
`flags`	A vector of class `numeric` of flagged records.

Value

A summarized table of class `data.frame`, of length 19 and width 1, for the plausibility test results and their respective acceptability rates.

References

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: https://smartmethodology.org.

See Also

`mw_plausibility_check_mfaz()` `mw_plausibility_check_muac()` `mw_wrangle_age()`

Examples

```
## First wrangle age data ----
data <- mw_wrangle_age(
  df = anthro.01,
  dos = dos,
  dob = dob,
  age = age,
  .decimals = 2
)

## Then wrangle WFHZ data ----
data_wfhz <- mw_wrangle_wfhz(
  df = data,
  sex = sex,
  weight = weight,
  height = height,
  .recode_sex = TRUE
)

## Now run the plausibility check ----
mw_plausibility_check_wfhz(
  df = data_wfhz,
  sex = sex,
  age = age,
  weight = weight,
  height = height,
  flags = flag_wfhz
)
```

`mw_stattest_ageratio` *Test for statistical difference between the proportion of children aged 24 to 59 months old over those aged 6 to 23 months old*

Description

Calculate the observed age ratio of children aged 24 to 59 months old over those aged 6 to 23 months old and test if there is a statistical difference between the observed and the expected.

Usage

```
mw_stattest_ageratio(age, .expectedP = 0.66)
```

Arguments

`age`	A vector of class `numeric` of child's age in months. If different than expected, the function will stop execution and return an error message indicating the type of mismatch.

`.expectedP`	The expected proportion of children aged 24 to 59 months old over those aged 6 to 23 months old. This is estimated to be 0.66.

Details

This function should be used specifically when assessing the quality of MUAC data. For age ratio test of children aged 6 to 29 months old over 30 to 59 months old, as performed in the SMART plausibility check, use `nipnTK::ageRatioTest()` instead.

Value

A vector of class `list` of three statistics: `p` for p-value of the statistical difference between the observed and the expected proportion of children aged 24 to 59 months old over those aged 6 to 23 months old; `observedR` and `observedP` for the observed ratio and proportion respectively.

References

SMART Initiative. *Updated MUAC data collection tool*. Available at: `https://smartmethodology.org/survey-planning-tools/updated-muac-tool/`

Examples

```
mw_stattest_ageratio(
  age = anthro.02$age,
  .expectedP = 0.66
)
```

`mw_wrangle_age` *Wrangle child's age*

Description

Wrangle child's age for downstream analysis. This includes calculating age in months based on the date of data collection and the child's date of birth, and setting to `NA` the age values that are less than 6.0 and greater than or equal to 60.0 months old.

Usage

```
mw_wrangle_age(df, dos = NULL, dob = NULL, age, .decimals = 2)
```

Arguments

`df`	A data set of class `data.frame` to wrangle age from.
`dos`	A vector of class `Date` for date of data collection from the `df`. Default is `NULL`.
`dob`	A vector of class `Date` for child's date of birth from the `df`. Default is `NULL`.

`age`	A vector of class `numeric` of child's age in months. In most cases this will be estimated using local event calendars; in some other cases it can be a mix of the former and the one based on the child's date of birth and the date of data collection.
`.decimals`	The number of decimals places to which the age should be rounded. Default is 2.

Value

A `data.frame` based on `df`. The variable `age` will be automatically filled in each row where age value was missing and both the child's date of birth and the date of data collection are available. Rows where `age` is less than 6.0 and greater than or equal to 60.0 months old will be set to `NA`. Additionally, a new variable for `df` named `age_days`, of class `double`, will be created.

Examples

```
## A sample data ----
df <- data.frame(
  surv_date = as.Date(c(
    "2023-01-01", "2023-01-01", "2023-01-01", "2023-01-01", "2023-01-01"
  )),
  birth_date = as.Date(c(
    "2019-01-01", NA, "2018-03-20", "2019-11-05", "2021-04-25"
  )),
  age = c(NA, 36, NA, NA, NA)
)

## Apply the function ----
mw_wrangle_age(
  df = df,
  dos = surv_date,
  dob = birth_date,
  age = age,
  .decimals = 3
)
```

`mw_wrangle_muac` *Wrangle MUAC data*

Description

Calculate z-scores for MUAC-for-age (MFAZ) and identify outliers based on the SMART methodology. When age is not supplied, wrangling will consist only in detecting outliers from the raw MUAC values. The function only works after the age has been wrangled.

Usage

```
mw_wrangle_muac(
  df,
  sex,
  muac,
  age = NULL,
  .recode_sex = TRUE,
  .recode_muac = TRUE,
  .to = c("cm", "mm", "none"),
  .decimals = 3
)
```

Arguments

`df`	A data set object of class `data.frame` to wrangle data from.
`sex`	A `numeric` or `character` vector of child's sex. Code values should only be 1 or "m" for males and 2 or "f" for females. Make sure sex values are coded in either of the aforementioned before calling the function. If input codes are different than expected, the function will stop execution and return an error message with the type of mismatch.
`muac`	A vector of class `numeric` of child's age in months. If the class is different than expected, the function will stop execution and return an error message indicating the type of mismatch.
`age`	A vector of class `numeric` of child's age in months.
`.recode_sex`	Logical. Set to `TRUE` if the values for `sex` are not coded as 1 (for males) or 2 (for females). Otherwise, set to `FALSE` (default).
`.recode_muac`	Logical. Set to `TRUE` if the values for raw MUAC should be converted to either centimeters or millimeters. Otherwise, set to `FALSE` (default)
`.to`	A choice of the measuring unit to which the MUAC values should be converted; "cm" for centimeters, "mm" for millimeters and "none" to leave as it is.
`.decimals`	The number of decimals places the z-scores should have. Default is 3.

Value

A data frame based on `df`. New variables named `mfaz` and `flag_mfaz`, of child's MFAZ and detected outliers, will be created. When age is not supplied, only `flag_muac` variable is created. This refers to outliers detected based on the raw MUAC values.

References

Bilukha, O., & Kianian, B. (2023). Considerations for assessment of measurement quality of mid-upper arm circumference data in anthropometric surveys and mass nutritional screenings conducted in humanitarian and refugee settings. *Maternal & Child Nutrition*, 19, e13478. https://doi.org/10.1111/mcn.13478

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: https://smartmethodology.org.

See Also

flag_outliers() remove_flags() mw_wrangle_age()

Examples

```
## When age is available, wrangle it first before calling the function ----
w <- mw_wrangle_age(
  df = anthro.02,
  dos = NULL,
  dob = NULL,
  age = age,
  .decimals = 2
)

### Then apply the function to wrangle MUAC data ----
mw_wrangle_muac(
  df = w,
  sex = sex,
  age = age,
  muac = muac,
  .recode_sex = TRUE,
  .recode_muac = TRUE,
  .to = "cm",
  .decimals = 3
)

## When age is not available ----
mw_wrangle_muac(
  df = anthro.02,
  sex = sex,
  age = NULL,
  muac = muac,
  .recode_sex = TRUE,
  .recode_muac = TRUE,
  .to = "cm",
  .decimals = 3
)
```

`mw_wrangle_wfhz` *Wrangle weight-for-height data*

Description

Calculate z-scores for weight-for-height (WFHZ) and identify outliers based on the SMART methodology.

Usage

```
mw_wrangle_wfhz(df, sex, weight, height, .recode_sex = TRUE, .decimals = 3)
```

Arguments

`df`	A data set object of class `data.frame` to wrangle data from.
`sex`	A `numeric` or `character` vector of child's sex. Code values should only be 1 or "m" for males and 2 or "f" for females. Make sure sex values are coded in either of the aforementioned before to call the function. If input codes are neither of the above, the function will stop execution and return an error message with the type of mismatch.
`weight`	A vector of class `double` of child's weight in kilograms. If the input is of a different class, the function will stop execution and return an error message indicating the type of mismatch.
`height`	A vector of class `double` of child's height in centimeters. If the input is of a different class, the function will stop execution and return an error message indicating the type of mismatch.
`.recode_sex`	Logical. Set to `TRUE` if the values for `sex` are not coded as 1 (for males) or 2 (for females). Otherwise, set to `FALSE` (default).
`.decimals`	The number of decimals places the z-scores should have. Default is 3.

Value

A data frame based on `df`. New variables named `wfhz` and `flag_wfhz`, of child's WFHZ and detected outliers, will be created.

References

SMART Initiative (2017). *Standardized Monitoring and Assessment for Relief and Transition*. Manual 2.0. Available at: `https://smartmethodology.org`.

See Also

`flag_outliers()` `remove_flags()`

Examples

```
mw_wrangle_wfhz(
  df = anthro.01,
  sex = sex,
  weight = weight,
  height = height,
  .recode_sex = TRUE,
  .decimals = 2
)
```

`recode_muac` *Convert MUAC values to either centimeters or millimeters*

Description

Convert MUAC values to either centimeters or millimeters as required. Before to covert, the function checks if the supplied MUAC values are in the opposite unit of the intended conversion. If not, execution stops and an error message is returned.

Usage

```
recode_muac(x, .to = c("cm", "mm"))
```

Arguments

`x`	A vector of raw MUAC values. The class can either be `double` or `numeric` or `integer`. If different than expected, the function will stop execution and return an error message indicating the type of mismatch.
`.to`	A choice between `cm` (centimeters) and `mm` (millimeters) for the measuring unit to convert MUAC values to. Before to execute the conversion, the function checks if values are in the opposite unit; in case not, the execution stops and an error message is returned. Strive to address the error and try again.

Value

A `numeric` vector of the same length as `x`, with values converted to the chosen measuring unit.

Examples

```
## Recode from millimeters to centimeters ----
muac_cm <- recode_muac(
  x = anthro.01$muac,
  .to = "cm"
)
head(muac_cm)

## Using the `muac_cm` object to recode it back to "mm" ----
muac_mm <- recode_muac(
  x = muac_cm,
  .to = "mm"
)
tail(muac_mm)
```

`wfhz.01` *A sample SMART survey data with WFHZ standard deviation rated as problematic*

Description

A sample SMART survey data with WFHZ standard deviation rated as problematic

Usage

```
wfhz.01
```

Format

A tibble with 303 rows and 6 columns.

Variable	**Description**
cluster	Primary sampling unit
sex	Sex, "m" = boys, "f" = girls
age	calculated age in months with two decimal places
edema	Edema, "n" = no, "y" = yes
wfhz	MUAC-for-age z-scores with 3 decimal places
flag_wfhz	Flagged observations. 1=flagged, 0=not flagged

Source

Anonymous

Examples

```
wfhz.01
```

Index