Package ‘random.cdisc.data’

March 27, 2024

Type Package

Title Create Random ADaM Datasets

Version 0.3.15

Date 2024-03-26

Description
A set of functions to create random Analysis Data Model (ADaM) datasets and cached dataset. ADaM dataset specifications are described by the Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model Team.

License Apache License 2.0

Depends R (>= 3.6)

Imports checkmate (>= 2.1.0), dplyr (>= 1.1.2), lifecycle (>= 1.0.3), lubridate (>= 1.7.10), magrittr (>= 1.5), rlang (>= 1.1.0), stringr (>= 1.4.1), tibble (>= 3.2.1), tidyr (>= 1.1.4), yaml (>= 2.1.19)

Suggests diffdf, knitr (>= 1.42), rmarkdown (>= 2.19), testthat (>= 3.0.4), withr (>= 2.0.0)

VignetteBuilder knitr

RdMacros lifecycle

Config/Needs/verdepcheck mllg/checkmate, tidyverse/dplyr, r-lib/lifecycle, tidyverse/lubridate, tidyverse/magrittr, r-lib/rlang, tidyverse/stringr, tidyverse/tibble, tidyverse/tidyr, yaml=vubiostat/r-yaml, gowerc/diffdf, yihui/knitr, rstudio/rmarkdown, r-lib/testthat, r-lib/withr

Config/Needs/website insightsengineering/nesttemplate

Config/testthat/edition 3

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 7.3.1
R topics documented:

random.cdisc.data-package ........................................ 3
apply_metadata ......................................................... 4
cadab ................................................................. 5
cadae ................................................................. 5
cadaette ............................................................... 5
cadcm ................................................................. 6
caddv ................................................................. 6
cadeg ................................................................. 6
cadex ................................................................. 7
cadhy ................................................................. 7
cadlb ................................................................. 7
cadmh ................................................................. 8
cadpc ................................................................. 8
cadpp ................................................................. 8
cadqlqc .............................................................. 9
cadqs ................................................................. 9
cadrs ................................................................. 9
cadsl ................................................................. 10
cadsub ............................................................... 10
cadr ................................................................. 10
cadtte ............................................................... 11
cadvs ................................................................. 11
mutate_na .......................................................... 11
radab ................................................................. 12
raed ................................................................. 14
radaette ............................................................. 16
radcm ............................................................... 17
raddv ............................................................... 19
rade ................................................................. 20
radex ............................................................... 22
Description

Package to create random SDTM and ADAM datasets.

Author(s)

Maintainer: Joe Zhu <joe.zhu@roche.com>

Authors:

- Pawel Rucki <pawel.rucki@roche.com>
- Nick Paszty <npaszty@gene.com>
- Jana StoiLOVA <jana.stoiLOva@roche.com>
- Davide Garolini <davide.garolini@roche.com>
- Emily de la Rua <emily.de_la_rua@contractors.roche.com>
- Christopher DiPietrantonio
- Adrian Waddell <adrian.waddell@gene.com>

Other contributors:

- F. Hoffmann-La Roche AG [copyright holder, funder]
apply_metadata

Description

Apply label and variable ordering attributes to domains.

Usage

apply_metadata(
  df,
  filename,
  add_adsl = TRUE,
  adsl_filename = "metadata/ADSL.yml"
)

Arguments

df (data.frame)
Data frame to which metadata is applied.

filename (yaml)
File containing domain metadata.

add_adsl (logical)
Should ADSL data be merged to domain.

adsl_filename (yaml)
File containing ADSL metadata.

Value

Data frame with metadata applied.

Examples

seed <- 1
adsl <- radsl(seed = seed)
adsub <- radsub(adsl, seed = seed)
yaml_path <- file.path(path.package("random.cdisc.data"), "inst", "metadata")
adsl <- apply_metadata(adsl, file.path(yaml_path, "ADSL.yml"), FALSE)
adsub <- apply_metadata(
  adsub, file.path(yaml_path, "ADSUB.yml"), TRUE,
  file.path(yaml_path, "ADSL.yml")
)
cadab

Cached ADAB

Description
Cached ADAB data generated with seed = 1

Usage
data(cadab)

Format
An object of class tbl_df (inherits from tbl, data.frame) with 6916 rows and 21 columns.

cadae

Cached ADAE

Description
Cached ADAE data generated with seed = 1

Usage
data(cadae)

Format
An object of class tbl_df (inherits from tbl, data.frame) with 1934 rows and 92 columns.

cadaette

Cached ADAETTE

Description
Cached ADAETTE data generated with seed = 1

Usage
data(cadaette)

Format
An object of class tbl_df (inherits from tbl, data.frame) with 3600 rows and 66 columns.
cadcm  
*Cached ADCM*

**Description**
Cached ADCM data generated with seed = 1

**Usage**
data(cadcm)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 3685 rows and 83 columns.

caddv  
*Cached ADDV*

**Description**
Cached ADDV data generated with seed = 1

**Usage**
data(caddv)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 119 rows and 66 columns.

cadeg  
*Cached ADEG*

**Description**
Cached ADEG data generated with seed = 1

**Usage**
data(cadeg)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 13600 rows and 88 columns.
cadex

---

**Cadex**

**Description**
Cached ADEX data generated with seed = 1

**Usage**
data(cadex)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 6400 rows and 79 columns.

---

cadhy

**Description**
Cached ADHY data generated with seed = 1

**Usage**
data(cadhy)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 20000 rows and 71 columns.

---

cadlb

**Description**
Cached ADLB data generated with seed = 1

**Usage**
data(cadlb)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 8400 rows and 102 columns.
**cadmhc**  
*Cached ADMH*

**Description**  
Cached ADMH data generated with seed = 1

**Usage**  
data(cadmh)

**Format**  
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 1934 rows and 67 columns.

---

**cadpc**  
*Cached ADPC*

**Description**  
Cached ADPC data generated with seed = 1

**Usage**  
data(cadpc)

**Format**  
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 6640 rows and 72 columns.

---

**cadpp**  
*Cached ADPP*

**Description**  
Cached ADPP data generated with seed = 1

**Usage**  
data(cadpp)

**Format**  
An object of class `data.frame` with 26268 rows and 68 columns.
**cadqlqc**

* Cached ADQLQC data generated with seed = 1

**Usage**

```
data(cadqlqc)
```

**Format**

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 116803 rows and 50 columns.

---

**cadqs**

* Cached ADQS data generated with seed = 1

**Usage**

```
data(cadqs)
```

**Format**

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 14000 rows and 73 columns.

---

**cadrs**

* Cached ADRS data generated with seed = 1

**Usage**

```
data(cadrs)
```

**Format**

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 3200 rows and 65 columns.
**cadsl**  
*Cached ADSL*

**Description**
Cached ADSL data generated with seed = 1

**Usage**
data(cadsl)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 400 rows and 55 columns.

---

**cadsub**  
*Cached ADSUB*

**Description**
Cached ADSUB data generated with seed = 1

**Usage**
data(cadsub)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 2000 rows and 65 columns.

---

**cadtr**  
*Cached ADTR*

**Description**
Cached ADTR data generated with seed = 1

**Usage**
data(cadtr)

**Format**
An object of class `data.frame` with 2800 rows and 76 columns.
cadtte

*Cached ADTTE*

**Description**
Cached ADTTE data generated with seed = 1

**Usage**
data(cadtte)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 2000 rows and 67 columns.

---

cadvs

*Cached ADVS*

**Description**
Cached ADVS data generated with seed = 1

**Usage**
data(cadvs)

**Format**
An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 16800 rows and 87 columns.

---

mutate_na

*Replace Values with NA*

**Description**

[Stable]
Replace column values with NAs.

**Usage**
mutate_na(ds, na_vars = NULL, na_percentage = 0.05)
Arguments

**ds** (data.frame)
Any data set.

**na_vars** (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
- **seed** (numeric)
  The seed to be used for this element - can be NA.
- **percentage** (proportion)
  Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

**na_percentage** (proportion)
Default percentage of values to be replaced by NA.

Value
dataframe without NA values.

---

**radab**

*Anti-Drug Antibody Analysis Dataset (ADAB)*

Description

[Stable]
Function for generating a random Anti-Drug Antibody Analysis Dataset for a given Subject-Level Analysis Dataset and Pharmacokinetics Analysis Dataset.

Usage

```r
radab(
adsl,
adpc,
constants = c(D = 100, ka = 0.8, ke = 1),
paramcd = c("R1800000", "RESULT1", "R1800001", "RESULT2", "ADASTAT1", "INDUCD1", "ENHANC1", "TRUNAFF1", "EMERNEG1", "EMERPOS1", "PERSADA1", "TRANADA1", "BFLAG1", "TIMADA1", "ADADUR1", "ADASTAT2", "INDUCD2", "ENHANC2", "EMERNEG2", "EMERPOS2", "BFLAG2", "TRUNAFF2"),
"Treatment Emergent - Negative, Neutralizing Antibody",
"Treatment Emergent - Positive, Neutralizing Antibody",
"Baseline, Neutralizing Antibody", "Treatment unaffected, Neutralizing Antibody"),
avalu = c("titer", "", "titer", "", "", "", "", "", "", "", "", "", "weeks",
"weeks", "", "", "", "", "", "", "")
seed = NULL,
na_percentage = 0,
na_vars = list(AVAL = c(NA, 0.1)),
cached = FALSE
)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).
adpc (data.frame)
Pharmacokinetics Analysis Dataset.
constants (character vector)
Constant parameters to be used in formulas for creating analysis values.
paramcd (character vector)
Parameter code values.
param (character vector)
Parameter values.
avalu (character)
Analysis value units.
seed (numeric)
Seed to use for reproducible random number generation.
na_percentage (proportion)
Default percentage of values to be replaced by NA.
na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
  • seed (numeric)
    The seed to be used for this element - can be NA.
  • percentage (proportion)
    Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.
cached boolean whether the cached ADAB data cadab should be returned or new data should be generated. If set to TRUE then the other arguments to radab will be ignored.

Details

One record per study per subject per parameter per time point: "R1800000", "RESULT1", "R1800001", "RESULT2".
Value

data.frame

Examples

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adpc <- radpc(adsl, seed = 2, duration = 9 * 7)
adab <- radab(adsl, adpc, seed = 2)
adab
```

radae **Adverse Event Analysis Dataset (ADAE)**

Description

[Stable]

Function for generating random Adverse Event Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

```r
radae(
adsl,
max_n_aes = 10L,
lookup = NULL,
lookup_aag = NULL,
seed = NULL,
na_percentage = 0,
na_vars = list(AEBODSYS = c(NA, 0.1), AEDECOD = c(1234, 0.1), AETOXGR = c(1234, 0.1)),
cached = FALSE
)
```

Arguments

- **adsl** *(data.frame)*
  Subject-Level Analysis Dataset (ADSL).
- **max_n_aes** *(integer)*
  Maximum number of AEs per patient. Defaults to 10.
- **lookup** *(data.frame)*
  Additional parameters.
- **lookup_aag** *(data.frame)*
  Additional metadata parameters.
- **seed** *(numeric)*
  Seed to use for reproducible random number generation.
na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
• seed (numeric)
  The seed to be used for this element - can be NA.
• percentage (proportion)
  Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached boolean whether the cached ADAE data cadae should be returned or new data should be generated. If set to TRUE then the other arguments to radae will be ignored.

Details
One record per each record in the corresponding SDTM domain.
Keys: STUDYID, USUBJID, ASTDTM, AETERM, AESEQ

Value
data.frame

Examples

adsl <- radsl(N = 10, study_duration = 2, seed = 1)
adae <- radae(adsl, seed = 2)
adae

# Add metadata.
aag <- utils::read.table(
  sep = "", header = TRUE,
  text = paste(
    "NAMVAR,SRVVAR,GRPTYPE,REFNAME,REFERTERM,SCOPE",
    "CQ01NAM,AEDECOD,CUSTOM,D.2.1.5.3/A.1.1.1.1 AESI,dcd D.2.1.5.3," ,
    "CQ01NAM,AEDECOD,CUSTOM,D.2.1.5.3/A.1.1.1.1 AESI,dcd A.1.1.1.1," ,
    "SMQ01NAM,AEDECOD,SMQ,C.1.1.1.3/B.2.2.3.1 AESI,dcd C.1.1.1.3,BROAD",
    "SMQ01NAM,AEDECOD,SMQ,C.1.1.1.3/B.2.2.3.1 AESI,dcd B.2.2.3.1,BROAD",
    "SMQ02NAM,AEDECOD,SMQ,Y.9.9.9.9/Z.9.9.9.9 AESI,dcd Y.9.9.9.9,NARROW",
    "SMQ02NAM,AEDECOD,SMQ,Y.9.9.9.9/Z.9.9.9.9 AESI,dcd Z.9.9.9.9,NARROW", sep = "\n"
  ), stringsAsFactors = FALSE
)
adae <- radae(adsl, lookup_aag = aag)

with(
  adae,
  cbind(}
radaette

Time to Adverse Event Analysis Dataset (ADAETTE)

Description

[Stable]
Function to generate random Time-to-AE Dataset for a given Subject-Level Analysis Dataset.

Usage

radaette(
  adsl,
  event_descr = NULL,
  censor_descr = NULL,
  lookup = NULL,
  seed = NULL,
  na_percentage = 0,
  na_vars = list(CNSR = c(NA, 0.1), AVAL = c(1234, 0.1)),
  cached = FALSE
)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

event_descr (character vector)
Descriptions of events. Defaults to NULL.

censor_descr (character vector)
Descriptions of censors. Defaults to NULL.

lookup (data.frame)
Additional parameters.

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:

- seed (numeric)
The seed to be used for this element - can be NA.
radcm

- percentage (proportion)
  Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached  boolean whether the cached ADAETTE data cadaette should be returned or new data should be generated. If set to TRUE then the other arguments to radaette will be ignored.

Details

Keys: STUDYID, USUBJID, PARAMCD

Value
data.frame

Author(s)
Xiuting Mi

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)

adaette <- radaette(adsl, seed = 2)

adaette

radcm

Previous and Concomitant Medications Analysis Dataset (ADCM)

Description

[Stable]
Function for generating random Concomitant Medication Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

radcm(
adsl,
max_n_cms = 10L,
lookup = NULL,
seed = NULL,
na_percentage = 0,
na_vars = list(CMCLAS = c(NA, 0.1), CMDECOD = c(1234, 0.1), ATIREL = c(1234, 0.1)),
who_coding = FALSE,
cached = FALSE
)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

max_n_cms (integer)
Maximum number of concomitant medications per patient. Defaults to 10.

lookup (data.frame)
Additional parameters.

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:

- seed (numeric)
The seed to be used for this element - can be NA.
- percentage (proportion)
Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

who_coding (flag)
Whether WHO coding (with multiple paths per medication) should be used.

cached boolean whether the cached ADCM data cadcm should be returned or new data should be generated. If set to TRUE then the other arguments to radcm will be ignored.

Details

One record per each record in the corresponding SDTM domain.

 Keys: STUDYID, USUBJID, ASTDTM, CMSEQ

Value
data.frame

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adcm <- radcm(adsl, seed = 2)
adcm

adcm_who <- radcm(adsl, seed = 2, who_coding = TRUE)
adcm_who
**raddv**

Protocol Deviations Analysis Dataset (ADDV)

**Description**

[Stable]

Function for generating random Protocol Deviations Analysis Dataset for a given Subject-Level Analysis Dataset.

**Usage**

```r
raddv(
  adsl,  
  max_n_dv = 3L,  
  p_dv = 0.15,  
  lookup = NULL,  
  seed = NULL,  
  na_percentage = 0,  
  na_vars = list(ASTDT = c(seed = 1234, percentage = 0.1), DVCAT = c(seed = 1234, percentage = 0.1)),  
  cached = FALSE
)
```

**Arguments**

- **adsl** (data.frame): Subject-Level Analysis Dataset (ADSL).
- **max_n_dv** (integer): Maximum number of deviations per patient. Defaults to 3.
- **lookup** (data.frame): Additional parameters.
- **seed** (numeric): Seed to use for reproducible random number generation.
- **na_percentage** (proportion): Default percentage of values to be replaced by NA.
- **na_vars** (list): A named list where the name of each element is a column name of `ds`. Each element of this list should be a numeric vector with two elements:
  - **seed** (numeric): The seed to be used for this element - can be NA.
  - **percentage** (proportion): Percentage of elements to be replaced with NA. If NA, `na_percentage` is used as a default.
cached  boolean whether the cached ADDV data caddv should be returned or new data should be generated. If set to TRUE then the other arguments to raddv will be ignored.

Details

One record per each record in the corresponding SDTM domain.

Keys: STUDYID, USUBJID, ASTDT, DVTERM, DVSEQ

Value

data.frame

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
addv <- raddv(adsl, seed = 2)
addv

radeg  ECG Analysis Dataset (ADEG)

Description

[Stable]

Function for generating random dataset from ECG Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

radeg(
adsl,
egcat = c("INTERVAL", "INTERVAL", "MEASUREMENT", "FINDING"),
param = c("QT Duration", "RR Duration", "Heart Rate", "ECG Interpretation"),
paramcd = c("QT", "RR", "HR", "ECGINTP"),
paramu = c("msec", "msec", "beats/min", ""),
visit_format = "WEEK",
n_assessments = 5L,
n_days = 5L,
max_n_eg = 10L,
lookup = NULL,
seed = NULL,
na_percentage = 0,
na_vars = list(ABLFL = c(1235, 0.1), BASE = c(NA, 0.1), BASEC = c(NA, 0.1), CHG = c(1234, 0.1), PCHG = c(1234, 0.1)),
cached = FALSE)
Arguments

- `adsl` (data.frame): Subject-Level Analysis Dataset (ADSL).
- `egcat` (character vector): EG category values.
- `param` (character vector): Parameter values.
- `paramcd` (character vector): Parameter code values.
- `paramu` (character vector): Parameter unit values.
- `visit_format` (character): Type of visit. Options are "WEEK" and "CYCLE".
- `n_assessments` (integer): Number of weeks or cycles.
- `n_days` (integer): Number of days in each cycle (only used if `visit_format` is "CYCLE").
- `max_n_eg` (integer): Maximum number of EG results per patient. Defaults to 10.
- `lookup` (data.frame): Additional parameters.
- `seed` (numeric): Seed to use for reproducible random number generation.
- `na_percentage` (proportion): Default percentage of values to be replaced by NA.
- `na_vars` (list): A named list where the name of each element is a column name of `ds`. Each element of this list should be a numeric vector with two elements:
  - `seed` (numeric): The seed to be used for this element - can be NA.
  - `percentage` (proportion): Percentage of elements to be replaced with NA. If NA, `na_percentage` is used as a default.
- `cached` (boolean): Whether the cached ADEG data `cadeg` should be returned or new data should be generated. If set to TRUE then the other arguments to `radeg` will be ignored.

Details

One record per subject per parameter per analysis visit per analysis date.

Keys: STUDYID, USUBJID, PARAMCD, BASETYPE, AVISITN, ATPTN, DTYPE, ADTM, EGSEQ, ASPID

Value

data.frame
Author(s)

tomlinsj, npaszty, Xuefeng Hou, dipietrc

Examples

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)

adeg <- radeg(adsl, visit_format = "WEEK", n_assessments = 7L, seed = 2)
adeg

adeg <- radeg(adsl, visit_format = "CYCLE", n_assessments = 2L, seed = 2)
adeg
```

radex  Exposure Analysis Dataset (ADEX)

Description

[Stable]

Function for generating random Exposure Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

```r
radex(
adsl,  
param = c("Dose administered during constant dosing interval",  
"Number of doses administered during constant dosing interval",  
"Total dose administered", "Total number of doses administered"),  
paramcd = c("DOSE", "NDOSE", "TDOSE", "TNDOSE"),  
paramu = c("mg", "", "mg", "");  
parcat1 = c("INDIVIDUAL", "OVERALL"),  
parcat2 = c("Drug A", "Drug B"),  
visit_format = "WEEK",  
n_assessments = 5L,  
n_days = 5L,  
max_n_exs = 6L,  
lookup = NULL,  
seed = NULL,  
na_percentage = 0,  
na_vars = list(AVAL = c(NA, 0.1), AVALU = c(NA), 0.1),  
cached = FALSE  
)
```
Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

param (character vector)
Parameter values.

paramcd (character vector)
Parameter code values.

paramu (character vector)
Parameter unit values.

parcat1 (character vector)
Dose amount categories. Defaults to "Individual" and "Overall".

parcat2 (character vector)
Types of drug received. Defaults to "Drug A" and "Drug B".

visit_format (character)
Type of visit. Options are "WEEK" and "CYCLE".

n_assessments (integer)
Number of weeks or cycles.

n_days (integer)
Number of days in each cycle (only used if visit_format is "CYCLE").

max_n_exs (integer)
Maximum number of exposures per patient. Defaults to 6.

lookup (data.frame)
Additional parameters.

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
- seed (numeric)
The seed to be used for this element - can be NA.
- percentage (proportion)
Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached boolean whether the cached ADEX data cadex should be returned or new data should be generated. If set to TRUE then the other arguments to radex will be ignored.

Details

One record per each record in the corresponding SDTM domain.

Keys: STUDYID, USUBJID, EXSEQ, PARAMCD, PARCAT1, ASTDTM, AENDTM, ASTDY, AENDY, AVISITN, EXDOSFRQ, EXROUTE, VISIT, VISITDY, EXSTDTC, EXENDTC, EXSTDY, EXENDY
Value

data.frame

Examples

```r
adsl <- radsl(N = 10, study_duration = 2, seed = 1)
adex <- radex(adsl, seed = 2)
adex
```
Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

param (character vector)
Parameter values.

paramcd (character vector)
Parameter code values.

seed (numeric)
Seed to use for reproducible random number generation.

cached boolean whether the cached ADHY data cadhy should be returned or new data should be generated. If set to TRUE then the other arguments to radhy will be ignored.

Details

One record per subject per parameter per analysis visit per analysis date.

Keys: STUDYID, USUBJID, PARAMCD, AVISITN, ADTM, SRCSEQ

Value
data.frame

Author(s)
wojciakw

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adhy <- radhy(adsl, seed = 2)
adhy
Description

[Stable]

Function for generating a random Laboratory Data Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

```r
radlb(
adsl,
  lbcat = c("CHEMISTRY", "CHEMISTRY", "IMMUNOLOGY"),
  param = c("Alanine Aminotransferase Measurement", "C-Reactive Protein Measurement",
            "Immunoglobulin A Measurement"),
  paramcd = c("ALT", "CRP", "IGA"),
  paramu = c("U/L", "mg/L", "g/L"),
  aval_mean = c(18, 9, 2.9),
  visit_format = "WEEK",
  n_assessments = 5L,
  n_days = 5L,
  max_n_lbs = 10L,
  lookup = NULL,
  seed = NULL,
  na_percentage = 0,
  na_vars = list(LOQFL = c(NA, 0.1), ABLFL2 = c(1234, 0.1), ABLFL = c(1235, 0.1), BASE2 =
                 c(NA, 0.1), BASE = c(NA, 0.1), CHG2 = c(1235, 0.1), PCHG2 = c(1235, 0.1), CHG =
                 c(1234, 0.1), PCHG = c(1234, 0.1)),
  cached = FALSE)
)
```

Arguments

- **adsl** (data.frame)
  Subject-Level Analysis Dataset (ADSL).
- **lbcat** (character vector)
  LB category values.
- **param** (character vector)
  Parameter values.
- **paramcd** (character vector)
  Parameter code values.
- **paramu** (character vector)
  Parameter unit values.
radlb

aval_mean  (numeric vector)
Mean values corresponding to each parameter.

visit_format  (character)
Type of visit. Options are "WEEK" and "CYCLE".

n_assessments  (integer)
Number of weeks or cycles.

n_days  (integer)
Number of days in each cycle (only used if visit_format is "CYCLE").

max_n_lbs  (integer)
Maximum number of labs per patient. Defaults to 10.

lookup  (data.frame)
Additional parameters.

seed  (numeric)
Seed to use for reproducible random number generation.

na_percentage  (proportion)
Default percentage of values to be replaced by NA.

na_vars  (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
  • seed (numeric)
    The seed to be used for this element - can be NA.
  • percentage (proportion)
    Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached  boolean whether the cached ADLB data cadlb should be returned or new data should be generated. If set to TRUE then the other arguments to radlb will be ignored.

Details
One record per subject per parameter per analysis visit per analysis date.

Keys: STUDYID, USUBJID, PARAMCD, BASETYPE, AVISITN, ATPTN, DTYPE, ADTM, LBSEQ, ASPID

Value
data.frame

Author(s)
tomlinsj, npaszty, Xuefeng Hou

Examples
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adlb <- radlb(adsl, visit_format = "WEEK", n_assessments = 7L, seed = 2)
adlb <- radlb(adsl, visit_format = "CYCLE", n_assessments = 2L, seed = 2)
adlb

radmh  
*Medical History Analysis Dataset (ADMH)*

**Description**

[Stable]

Function for generating a random Medical History Analysis Dataset for a given Subject-Level Analysis Dataset.

**Usage**

```r
radmh(
adsl,
max_n_mhs = 10L,
lookup = NULL,
seed = NULL,
na_percentage = 0,
na_vars = list(MHBODSYS = c(NA, 0.1), MHDECOD = c(1234, 0.1)),
cached = FALSE
)
```

**Arguments**

- `adsl` *(data.frame)*
  Subject-Level Analysis Dataset (ADSL).
- `max_n_mhs` *(integer)*
  Maximum number of MHs per patient. Defaults to 10.
- `lookup` *(data.frame)*
  Additional parameters.
- `seed` *(numeric)*
  Seed to use for reproducible random number generation.
- `na_percentage` *(proportion)*
  Default percentage of values to be replaced by NA.
- `na_vars` *(list)*
  A named list where the name of each element is a column name of `ds`. Each element of this list should be a numeric vector with two elements:
  - `seed` *(numeric)*
    The seed to be used for this element - can be NA.
  - `percentage` *(proportion)*
    Percentage of elements to be replaced with NA. If NA, `na_percentage` is used as a default.
radpc

boolean whether the cached ADMH data cadmh should be returned or new data should be generated. If set to TRUE then the other arguments to radmh will be ignored.

Details

One record per each record in the corresponding SDTM domain.

Keys: STUDYID, USUBJID, ASTDTM, MHSEQ

Value

data.frame

Examples

adsl <- radsl(N = 10, study_duration = 2, seed = 1)
admh <- radmh(adsl, seed = 2)
admh

radpc Pharmacokinetics Analysis Dataset (ADPC)

Description

[Stable]

Function for generating a random Pharmacokinetics Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

radpc(
adsl,
avalu = "ug/mL",
constants = c(D = 100, ka = 0.8, ke = 1),
duration = 2,
seed = NULL,
na_percentage = 0,
na_vars = list(AVAL = c(NA, 0.1)),
cached = FALSE
)
Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

avalu (character)
Analysis value units.

constants (character vector)
Constant parameters to be used in formulas for creating analysis values.

duration (numeric)
Duration in number of days.

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
  • seed (numeric)
    The seed to be used for this element - can be NA.
  • percentage (proportion)
    Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached boolean whether the cached ADPC data cadpc should be returned or new data should be generated. If set to TRUE then the other arguments to radpc will be ignored.

Details

One record per study, subject, parameter, and time point.

Value
data.frame

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adpc <- radpc(adsl, seed = 2)
adpc

adpc <- radpc(adsl, seed = 2, duration = 3)
adpc
Pharmacokinetics Parameters Dataset (ADPP)

Description

[Stable]
Function for generating a random Pharmacokinetics Parameters Dataset for a given Subject-Level Analysis Dataset.

Usage

```r
radpp(
adsl,  
paramcd = c("AUCIFO", "CMAX", "CLO", "RMAX", "TON", "RENALCL", "RENALCLD", "RCAMINT", "RCPCINT"),  
param = c("AUC Infinity Obs", "Max Conc", "Total CL Obs", "Time of Maximum Response", "Time to Onset", "Renal CL", "Renal CL Norm by Dose", "Amt Rec from T1 to T2", "Pct Rec from T1 to T2"),  
paramu = c("day*ug/mL", "ug/mL", "ml/day/kg", "hr", "hr", "L/hr", "L/hr/mg", "mg", "%"),  
aval_mean = c(200, 30, 5, 10, 0.05, 0.005, 1.5613, 15.65),  
visit_format = "CYCLE", 
n_days = 2L,  
seed = NULL,  
na_percentage = 0,  
na_vars = list(AVAL = c(NA, 0.1)),  
cached = FALSE)
```

Arguments

- **adsl**: (data.frame) Subject-Level Analysis Dataset (ADSL).
- **ppcat**: (character vector) Categories of parameters.
- **ppspec**: (character vector) Specimen material types.
- **paramcd**: (character vector) Parameter code values.
- **param**: (character vector) Parameter values.
- **paramu**: (character vector) Parameter unit values.
### radqlqc

**val_mean** (numeric vector)
Mean values corresponding to each parameter.

**visit_format** (character)
Type of visit. Options are "WEEK" and "CYCLE".

**n_days** (integer)
Number of days in each cycle (only used if visit_format is "CYCLE").

**seed** (numeric)
Seed to use for reproducible random number generation.

**na_percentage** (proportion)
Default percentage of values to be replaced by NA.

**na_vars** (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
- **seed** (numeric)
  The seed to be used for this element - can be NA.
- **percentage** (proportion)
  Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

**cached** boolean whether the cached ADPP data cadpp should be returned or new data should be generated. If set to TRUE then the other arguments to radpp will be ignored.

### Details

One record per study, subject, parameter category, parameter and visit.

### Value
data.frame

### Examples

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adpp <- radpp(adsl, seed = 2)
adpp
```

---

**radqlqc**

**EORTC QLQ-C30 V3 Analysis Dataset (ADQLQC)**

### Description

[Stable]
Function for generating a random EORTC QLQ-C30 V3 Analysis Dataset for a given Subject-Level Analysis Dataset.
Usage

radqlqc(adsl, percent, number, seed = NULL, cached = FALSE)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

percent (numeric)
Completion - Completed at least y percent of questions, 1 record per visit

number (numeric)
Completion - Completed at least x question(s), 1 record per visit

seed (numeric)
Seed to use for reproducible random number generation.

cached boolean whether the cached ADQLQC data cadqlqc should be returned or new data should be generated. If set to TRUE then the other arguments to radqlqc will be ignored.

Details

Keys: STUDYID, USUBJID, PARCAT1N, PARAMCD, BASETYPE, AVISITN, ATPTN, ADTM, QSSEQ

Value
data.frame

Examples

adsl <- radsl(N = 10, study_duration = 2, seed = 1)
adqlqc <- radqlqc(adsl, seed = 1, percent = 80, number = 2)
adqlqc

radqs Questionnaires Analysis Dataset (ADQS)

Description

[Stable]
Function for generating a random Questionnaires Analysis Dataset for a given Subject-Level Analysis Dataset.
Usage

```r
radqs(
adsl,
    param = c("BFI All Questions", "Fatigue Interference",
              "Function/Well-Being (GF1,GF3,GF7)", "Treatment Side Effects (GP2,C5,GP5)",
              "FKSI-19 All Questions"),
    paramcd = c("BFIALL", "FATIGI", "FKSI-FWB", "FKSI-TSE", "FKSIALL"),
    visit_format = "WEEK",
    n_assessments = 5L,
    n_days = 5L,
    seed = NULL,
    na_percentage = 0,
    na_vars = list(LOQFL = c(NA, 0.1), ABLFL2 = c(1234, 0.1), ABLFL = c(1235, 0.1), CHG2 =
                    c(1235, 0.1), PCHG2 = c(1235, 0.1), CHG = c(1234, 0.1), PCHG = c(1234, 0.1)),
    cached = FALSE
)
```

Arguments

- **adsl** (data.frame): Subject-Level Analysis Dataset (ADSL).
- **param** (character vector): Parameter values.
- **paramcd** (character vector): Parameter code values.
- **visit_format** (character): Type of visit. Options are "WEEK" and "CYCLE".
- **n_assessments** (integer): Number of weeks or cycles.
- **n_days** (integer): Number of days in each cycle (only used if visit_format is "CYCLE").
- **seed** (numeric): Seed to use for reproducible random number generation.
- **na_percentage** (proportion): Default percentage of values to be replaced by NA.
- **na_vars** (list): A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:
  - **seed** (numeric): The seed to be used for this element - can be NA.
  - **percentage** (proportion): Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.
- **cached** (boolean): Whether the cached ADQS data cadqs should be returned or new data should be generated. If set to TRUE then the other arguments to radqs will be ignored.
**Details**

One record per subject per parameter per analysis visit per analysis date.

Keys: STUDYID, USUBJID, PARAMCD, AVISITN

**Value**

data.frame

**Author(s)**
npaszty

**Examples**

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adqs <- radqs(adsl, visit_format = "WEEK", n_assessments = 7L, seed = 2)
adqs
adqs <- radqs(adsl, visit_format = "CYCLE", n_assessments = 3L, seed = 2)
adqs
```

---

**radrs**  
*Tumor Response Analysis Dataset (ADRS)*

**Description**

[Stable]

Function for generating a random Tumor Response Analysis Dataset for a given Subject-Level Analysis Dataset.

**Usage**

```r
radrs(  
adsl,  
avalc = NULL,  
lookup = NULL,  
seed = NULL,  
na_percentage = 0,  
na_vars = list(AVISIT = c(NA, 0.1), AVAL = c(1234, 0.1), AVALC = c(1234, 0.1)),  
cached = FALSE)
```
Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

avalc (character vector)
Analysis value categories.

lookup (data.frame)
Additional parameters.

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:

- seed (numeric)
  The seed to be used for this element - can be NA.
- percentage (proportion)
  Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached boolean whether the cached ADRS data cadrs should be returned or new data should be generated. If set to TRUE then the other arguments to radrs will be ignored.

Details

One record per subject per parameter per analysis visit per analysis date. SDTM variables are populated on new records coming from other single records. Otherwise, SDTM variables are left blank.

Keys: STUDYID, USUBJID, PARAMCD, AVISITN, ADT, RSSEQ

Value
data.frame

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adrs <- radrs(adsl, seed = 2)
adrs
radsaftte

**Time to Safety Event Analysis Dataset (ADSAFTTE)**

### Description

Function to generate random Time-to-Safety Event Dataset for a given Subject-Level Analysis Dataset.

### Usage

```r
radsaftte(adsl, ...)  
```

### Arguments

- `adsl` *(data.frame)*  
  Subject-Level Analysis Dataset (ADSL).
- `...`  
  Additional arguments to be passed to `radaette`

### Details

**Keys:** STUDYID, USUBJID, PARAMCD

### Value

*data.frame*

### Examples

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adlsaftte <- radsaftte(adsl, seed = 2)
adlsaftte
```

### radsl

**Subject-Level Analysis Dataset (ADSL)**

### Description

[Stable]  

The Subject-Level Analysis Dataset (ADSL) is used to provide the variables that describe attributes of a subject. ADSL is a source for subject-level variables used in other analysis data sets, such as population flags and treatment variables. There is only one ADSL per study. ADSL and its related metadata are required in a CDISC-based submission of data from a clinical trial even if no other analysis data sets are submitted.
Usage

radsl(  
  N = 400,  
  study_duration = 2,  
  seed = NULL,  
  with_trt02 = TRUE,  
  na_percentage = 0,  
  na_vars = list(AGE = NA, SEX = NA, RACE = NA, STRATA1 = NA, STRATA2 = NA, BMRKR1 = 
      c(seed = 1234, percentage = 0.1), BMRKR2 = c(1234, 0.1), BEP01FL = NA),  
  ae_withdrawal_prob = 0.05,  
  cached = FALSE  
)

Arguments

N               (numeric)  
Number of patients.

study_duration (numeric)  
Duration of study in years.

seed            (numeric)  
Seed to use for reproducible random number generation.

with_trt02     (logical)  
Should period 2 be added.

na_percentage   (proportion)  
Default percentage of values to be replaced by NA.

na_vars         (list)  
A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:

  • seed (numeric)  
    The seed to be used for this element - can be NA.
  • percentage (proportion)  
    Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

ae_withdrawal_prob (proportion)  
Probability that there is at least one Adverse Event leading to the withdrawal of a study drug.

cached          boolean whether the cached ADSL data cadsl should be returned or new data should be generated. If set to TRUE then the other arguments to radsl will be ignored.

Details

One record per subject.

Keys: STUDYID, USUBJID
Value
data.frame

Examples

adsl <- radsl(N = 10, study_duration = 2, seed = 1)
adsl

adsl <- radsl(
  N = 10, seed = 1,
  na_percentage = 0.1,
  na_vars = list(
    DTHDT = c(seed = 1234, percentage = 0.1),
    LSTALVDT = c(seed = 1234, percentage = 0.1)
  )
)
adsl

adsl <- radsl(N = 10, seed = 1, na_percentage = .1)
adsl

radsu
Subcategory Analysis Dataset (ADSUB)

Description

[Stable]
Function for generating a random Subcategory Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

radsu(
adsl,
  param = c("Baseline Weight", "Baseline Height", "Baseline BMI", "Baseline ECOG",
            "Baseline Biomarker Mutation"),
  paramcd = c("BWGHTSI", "BHGHTSI", "BBMISI", "BECOG", "BBMRKR1"),
  seed = NULL,
  na_percentage = 0,
  na_vars = list(),
  cached = FALSE
)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).
param (character vector) Parameter values.

paramcd (character vector) Parameter code values.

seed (numeric) Seed to use for reproducible random number generation.

na_percentage (proportion) Default percentage of values to be replaced by NA.

na_vars (list) A named list where the name of each element is a column name of ds. Each element of this list should be a numeric vector with two elements:

  • seed (numeric)
    The seed to be used for this element - can be NA.
  • percentage (proportion)
    Percentage of elements to be replaced with NA. If NA, na_percentage is used as a default.

cached boolean whether the cached ADSUB data cadsub should be returned or new data should be generated. If set to TRUE then the other arguments to radsub will be ignored.

Details

One record per subject.

Keys: STUDYID, USUBJID, PARAMCD, AVISITN, ADTM, SRCSEQ

Value
data.frame

Author(s)
tomlinsj, npaszty, Xuefeng Hou, dipietrc

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adsub <- radsub(adsl, seed = 2)
adsub
### radtr

* **Tumor Response Analysis Dataset (ADTR)**

#### Description

**[Stable]**

Function for generating a random Tumor Response Analysis Dataset for a given Subject-Level Analysis Dataset.

#### Usage

```r
radtr(
adsl,
param = c("Sum of Longest Diameter by Investigator"),
paramcd = c("SLDINV"),
seed = NULL,
cached = FALSE,
...
)
```

#### Arguments

- `adsl` *(data.frame)*
  Subject-Level Analysis Dataset (ADSL).
- `param` *(character vector)*
  Parameter values.
- `paramcd` *(character vector)*
  Parameter code values.
- `seed` *(numeric)*
  Seed to use for reproducible random number generation.
- `cached` *(boolean)*
  Whether the cached ADTR data should be returned or new data should be generated. If set to `TRUE` then the other arguments to `radtr` will be ignored.
- `...` Additional arguments to be passed to `radrs`.

#### Details

One record per subject per parameter per analysis visit per analysis date.

**Keys:** STUDYID, USUBJID, PARAMCD, BASETYPE, AVISITN, DTYPE

#### Value

-data.frame
Author(s)

tomlinsj, npaszty, Xuefeng Hou, dipietrc

Examples

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adtr <- radtr(adsl, seed = 2)
adtr
```

radtte | Time-to-Event Analysis Dataset (ADTTE)

Description

[Stable]

Function for generating a random Time-to-Event Analysis Dataset for a given Subject-Level Analysis Dataset.

Usage

```r
radtte(
adsl,
  event_descr = NULL,
  censor_descr = NULL,
  lookup = NULL,
  seed = NULL,
  na_percentage = 0,
  na_vars = list(CNSR = c(NA, 0.1), AVAL = c(1234, 0.1), AVALU = c(1234, 0.1)),
  cached = FALSE
)
```

Arguments

- `adsl` (data.frame): Subject-Level Analysis Dataset (ADSL).
- `event_descr` (character vector): Descriptions of events. Defaults to `NULL`.
- `censor_descr` (character vector): Descriptions of censors. Defaults to `NULL`.
- `lookup` (data.frame): Additional parameters.
- `seed` (numeric): Seed to use for reproducible random number generation.
- `na_percentage` (proportion): Default percentage of values to be replaced by NA.
radvs

na_vars (list)
A named list where the name of each element is a column name of ds. Each
element of this list should be a numeric vector with two elements:
  • seed (numeric)
    The seed to be used for this element - can be NA.
  • percentage (proportion)
    Percentage of elements to be replaced with NA. If NA, na_percentage is
    used as a default.

cached boolean whether the cached ADTTE data cadtte should be returned or new
data should be generated. If set to TRUE then the other arguments to radtte will
be ignored.

Details

 Keys: STUDYID, USUBJID, PARAMCD

Value
data.frame

Examples

adsl <- radsl(N = 10, seed = 1, study_duration = 2)
adtte <- radtte(adsl, seed = 2)
adtte

radvs Vital Signs Analysis Dataset (ADVS)

Description

 [Stable]
Function for generating a random Vital Signs Analysis Dataset for a given Subject-Level Analysis
Dataset.

Usage

radvs(
adsl,
  param = c("Diastolic Blood Pressure", "Pulse Rate", "Respiratory Rate",
            "Systolic Blood Pressure", "Temperature", "Weight"),
  paramcd = c("DIABP", "PULSE", "RESP", "SYSBP", "TEMP", "WEIGHT"),
  paramu = c("Pa", "beats/min", "breaths/min", "Pa", "C", "Kg"),
  visit_format = "WEEK",
  n_assessments = 5L,
  n_days = 5L,
)
seed = NULL,
na_percentage = 0,
na_vars = list(CHG2 = c(1235, 0.1), PCHG2 = c(1235, 0.1), CHG = c(1234, 0.1), PCHG =
c(1234, 0.1), AVAL = c(123, 0.1), AVALU = c(123, 0.1)),
cached = FALSE)

Arguments

adsl (data.frame)
Subject-Level Analysis Dataset (ADSL).

param (character vector)
Parameter values.

paramcd (character vector)
Parameter code values.

paramu (character vector)
Parameter unit values.

visit_format (character)
Type of visit. Options are "WEEK" and "CYCLE".

n_assessments (integer)
Number of weeks or cycles.

n_days (integer)
Number of days in each cycle (only used if visit_format is "CYCLE").

seed (numeric)
Seed to use for reproducible random number generation.

na_percentage (proportion)
Default percentage of values to be replaced by NA.

na_vars (list)
A named list where the name of each element is a column name of ds. Each
element of this list should be a numeric vector with two elements:

- seed (numeric)
The seed to be used for this element - can be NA.

- percentage (proportion)
Percentage of elements to be replaced with NA. If NA, na_percentage is
used as a default.

cached boolean whether the cached ADVS data cadvs should be returned or new data
should be generated. If set to TRUE then the other arguments to radvs will be
ignored.

Details

One record per subject per parameter per analysis visit per analysis date.

Keys: STUDYID, USUBJID, PARAMCD, BASETYPE, AVISITN, ATPTN, DTYPE, ADTM, VSSEQ, ASPID
**Value**

data.frame

**Author(s)**
npaszty

**Examples**

```r
adsl <- radsl(N = 10, seed = 1, study_duration = 2)
advs <- radvs(adsl, visit_format = "WEEK", n_assessments = 7L, seed = 2)
advs
advs <- radvs(adsl, visit_format = "CYCLE", n_assessments = 3L, seed = 2)
advs
```

**Description**

Assign values to a related variable within a domain.

**Usage**

```r
rel_var(df, var_name, related_var, var_values = NULL)
```

**Arguments**

- `df` *(data.frame)*: Data frame containing the related variables.
- `var_name` *(character)*: Name of variable related to `rel_var` to add to `df`.
- `related_var` *(character)*: Name of variable within `df` with values to which values of `var_name` must relate.
- `var_values` *(any)*: Vector of values related to values of `related_var`.

**Value**

`df` with added factor variable `var_name` containing `var_values` corresponding to `related_var`. 
Examples

```r
# Example with data.frame.
params <- c("Level A", "Level B", "Level C")
adlb_df <- data.frame(
  ID = 1:9,
  PARAM = factor(
    rep(c("Level A", "Level B", "Level C"), 3),
    levels = params
  )
)
rel_var(
  df = adlb_df,
  var_name = "PARAMCD",
  var_values = c("A", "B", "C"),
  related_var = "PARAM"
)

# Example with tibble.
adlb_tbl <- tibble::tibble(
  ID = 1:9,
  PARAM = factor(
    rep(c("Level A", "Level B", "Level C"), 3),
    levels = params
  )
)
rel_var(
  df = adlb_tbl,
  var_name = "PARAMCD",
  var_values = c("A", "B", "C"),
  related_var = "PARAM"
)
```

Description

Replace Values in a Vector by NA

Usage

```r
replace_na(v, percentage = 0.05, seed = NULL)
```

Arguments

- `v` (any)
  Vector of any type.
percentage (proportion)
Value between 0 and 1 defining how much of the vector shall be replaced by `NA`. This number is randomized by +/- 5% to have full randomization.

seed (numeric)
Seed to use for reproducible random number generation.

Value
The input vector `v` where a certain number of values are replaced by `NA`.

---

**rtexp**  
*Truncated Exponential Distribution*

**Description**

[Stable]
This generates random numbers from a truncated Exponential distribution, i.e. from $X \mid X > l$ or $X \mid X < r$ when $X \sim \text{Exp}(\text{rate})$. The advantage here is that we guarantee to return exactly `n` numbers and without using a loop internally. This can be derived from the quantile functions of the left- and right-truncated Exponential distributions.

**Usage**

```r
rtexp(n, rate, l = NULL, r = NULL)
```

**Arguments**

- `n` (numeric)
  Number of random numbers.
- `rate` (numeric)
  Non-negative rate.
- `l` (numeric)
  Positive left-hand truncation parameter.
- `r` (numeric)
  Positive right-hand truncation parameter.

**Value**

The random numbers. If neither `l` nor `r` are provided then the usual Exponential distribution is used.
Examples
x <- stats::rexp(1e6, rate = 5)
x <- x[x > 0.5]
hist(x)

y <- rexp(1e6, rate = 5, l = 0.5)
hist(y)

z <- rexp(1e6, rate = 5, r = 0.5)
hist(z)

rtpois Zero-Truncated Poisson Distribution

Description
[Stable]
This generates random numbers from a zero-truncated Poisson distribution, i.e. from $X \mid X > 0$ when $X \sim \text{Poisson}(\lambda)$. The advantage here is that we guarantee to return exactly $n$ numbers and without using a loop internally. This solution was provided in a post by Peter Dalgaard.

Usage
rtpois(n, lambda)

Arguments

n (numeric)
Number of random numbers.

lambda (numeric)
Non-negative mean(s).

Value
The random numbers.

Examples
x <- rpois(1e6, lambda = 5)
x <- x[x > 0]
hist(x)

y <- rtpois(1e6, lambda = 5)
hist(y)
**sample_fct**  
*Create a Factor with Random Elements of x*

**Description**  
Sample elements from x with replacement to build a factor.

**Usage**  
sample_fct(x, N, ...)

**Arguments**  
- **x**  
  (character vector or factor)  
  If character vector then it is also used as levels of the returned factor. If factor then the levels are used as the new levels.
- **N**  
  (numeric)  
  Number of items to choose.
- **...**  
  Additional arguments to be passed to sample.

**Value**  
A factor of length N.

**Examples**  
```r  
sample_fct(letters[1:3], 10)  
sample_fct(iris$Species, 10)  
```

---

**var_relabel**  
*Primary Keys: Labels*

**Description**  
Relabel a subset of variables in a data set.

**Usage**  
var_relabel(x, ...)

**Arguments**  
- **x**  
  (data.frame)  
  Data frame containing variables to which labels are applied.
- **...**  
  (named character)  
  Name-Value pairs, where name corresponds to a variable name in x and the value to the new variable label.
visit_schedule

Description
Create a visit schedule as a factor.

Usage
visit_schedule(visit_format = "WEEK", n_assessments = 10L, n_days = 5L)

Arguments
visit_format (character)
Type of visit. Options are "WEEK" and "CYCLE".
n_assessments (integer)
Number of weeks or cycles.
n_days (integer)
Number of days in each cycle (only used if visit_format is "CYCLE").

Details
X number of visits, or X number of cycles and Y number of days.

Value
A factor of length n_assessments.

Examples
visit_schedule(visit_format = "WEEK", n_assessments = 10L)
visit_schedule(visit_format = "CYCLE", n_assessments = 5L, n_days = 2L)
Index

* datasets
  cadab, 5
cadae, 5
cadaette, 5
cadcm, 6
caddv, 6
cadeg, 6
cadex, 7
cadhy, 7
cadlb, 7
cadmh, 8
cadpc, 8
cadpp, 8
cadqlqc, 9
cadqs, 9
cadrs, 9
cadsl, 10
cadsub, 10
cadtr, 10
cadtte, 11
cadvs, 11
  mutate_na, 11
  radab, 12
  radae, 14
  radaette, 16
  radcm, 17
  raddv, 19
  radeg, 20
  radex, 22
  radhy, 24
  radlb, 26
  radmh, 28
  radpc, 29
  radpp, 31
  radqlqc, 32
  radqs, 33
  radrs, 35
  radsaftte, 37
  radsl, 37
  radsub, 39
  radtr, 41
  radtte, 42
  radvs, 43
  random.cdisc.data
    (random.cdisc.data-package), 3
    random.cdisc.data-package, 3
    rel_var, 45
    replace_na, 46
    rtexp, 47
    rtpois, 48
  sample_fct, 49
  var_relabel, 49
  visit_schedule, 50

apply_metadata, 4