

# Package ‘ibawds’

March 9, 2021

**Type** Package

**Title** Functions and Datasets for the Data Science Course at IBAW

**Version** 0.1.2

**Author** Stefan Lanz

**Maintainer** Stefan Lanz <slanz1137@gmail.com>

**Description** A collection of useful functions and datasets for the Data Science Course at IBAW in Lucerne.

**License** MIT + file LICENSE

**URL** <https://github.com/stibu81/ibawds>

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**Depends** R (>= 3.6.0), dslabs

**Imports** stats, utils, ggplot2, stringr, magrittr

**Suggests** tidyverse, rmarkdown, caret, reshape2, lubridate, ggrepel, writexl, cowplot, DT, gutenbergr, tidytext, rvest, Lahman, pdftools, HistData, titanic, BiocManager, usethis, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-03-09 08:50:02 UTC

## R topics documented:

bills	2
define_latex_stats	2
distribution_plot	3
install_ibawds	4
mtcars2	5
rescale	5
seatbelts	6

**Index**

7

---

bills	<i>Summarised Data on Restaurant Bills</i>
-------	--------------------------------------------

---

**Description**

Summary of data on restaurant bills from the dataset `reshape2::tips`. Labels are in German.

**Usage**

```
bills
```

**Format**

A data frame with 8 rows and 4 variables:

**sex** sex of the bill payer

**time** time of day

**smoker** whether there were smokers in the party

**mean\_bill** mean of all the bills in dollars

---

define_latex_stats	<i>Define LaTeX commands for statistical symbols</i>
--------------------	------------------------------------------------------

---

**Description**

Add the definitions for various useful LaTeX equation symbols for statistics to a RMarkdown document..

**Usage**

```
define_latex_stats()
```

**Details**

Run this function from within a code chunk in a RMarkdown document with options `results = "asis"` and `echo = FALSE` (see "Examples"). It only works for pdf output.

It defines the following macros:  $\backslash E$ ,  $\backslash P$ ,  $\backslash Var$ ,  $\backslash Cov$ ,  $\backslash Cor$ ,  $\backslash SD$ ,  $\backslash SE$ ,  $\backslash Xb$ ,  $\backslash Yb$ .

**Value**

The function returns NULL invisibly. The command definitions are output as a side effect.

## Examples

```
## Not run:  
# add this code chunk to a RMarkdown document  
```${r results = "asis", echo = FALSE}  
  define_latex_stats()  
```\n\n## End(Not run)
```

---

distribution\_plot

*Plot Density and Distribution Function With Markings*

---

## Description

Create plots of the density and distribution functions of a probability distribution. It is possible to mark points and shade the area under the curve.

## Usage

```
distribution_plot(  
  fun,  
  range,  
  ...,  
  points = NULL,  
  var = "x",  
  title = "Verteilungsfunktion"  
)
```

```
density_plot(  
  fun,  
  range,  
  ...,  
  from = NULL,  
  to = NULL,  
  points = NULL,  
  var = "x",  
  title = "Dichte"  
)
```

## Arguments

fun	a density or distribution function that takes quantiles as its first argument.
range	numeric vector of length two giving the range of quantiles to be plotted.
...	further arguments that are passed to fun().
points	numeric vector giving quantiles where the function should be marked with a red dot.

var	character giving the name of the quantile variable. This is only used to label the axes.
title	character giving the title of the plot
from, to	numeric values giving start and end of a range where the area under the density will be shaded. If only one of the two values is given, the shading will start a negative infinity or go until positive infinity, respectively.

**Value**

a ggplot object

**Examples**

```
# plot density of the normal distribution
density_plot(dnorm, c(-5, 7),
             mean = 1, sd = 2,
             to = 3)

# plot distribution function of the Poisson distribution
distribution_plot(ppois, c(0, 12),
                 lambda = 4,
                 points = c(2, 6, 10),
                 var = "y")
```

---

install\_ibawds

*Install the R-Packages Required for the Course*


---

**Description**

A number of R-packages are used in the courses and the video lectures. They are also dependencies of this package. Use `install_ibawds()` to install the packages that are not yet installed.

**Usage**

```
install_ibawds(just_print = FALSE)
```

**Arguments**

just_print	logical. If TRUE, the function will just print a message with the packages that need to be installed (if any) and stops without installing them.
------------	--------------------------------------------------------------------------------------------------------------------------------------------------

**Details**

This function checks whether all the packages that `ibawds` depends on, imports or suggests are installed. A message informs the user about missing packages and asks, whether they should be installed. If the process is aborted, no packages are installed.

**Value**

Invisible logical indicating whether package installation was successful. TRUE is returned also when all required packages were already installed.

---

mtcars2	<i>Dataset mtcars without row names</i>
---------	-----------------------------------------

---

**Description**

In the [mtcars](#) dataset, the names of the car models are stored as row names. However, when working with [ggplot2](#) and other packages from the [tidyverse](#), it is convenient to have all data in columns. `mtcars2` is a variant of `mtcars` that contains car models in a column instead of storing them as row names.

**Usage**

```
mtcars2
```

**Format**

A data frame with 32 rows and 12 variables. The format is identical to [mtcars](#) and details can be found in its documentation. The only difference is that the car model names are stored in the column `model` instead of row names.

---

rescale	<i>Rescale Mean And/Or Standard Deviation of a Vector</i>
---------	-----------------------------------------------------------

---

**Description**

Rescale Mean And/Or Standard Deviation of a Vector

**Usage**

```
rescale(x, mu = mean(x), sigma = sd(x))
```

**Arguments**

<code>x</code>	numeric vector
<code>mu</code>	numeric value giving the desired mean
<code>sigma</code>	numeric value giving the desired standard deviation

**Details**

By default, mean and standard deviation are not changed, i.e., `rescale(x)` is identical to `x`. Only if a value is specified for `mu` and/or `sigma` the mean and/or the standard deviation are rescaled.

**Value**

a numeric vector with the same length as `x` with mean `mu` and standard deviation `sigma`.

**Examples**

```
x <- runif(1000, 5, 8)

# calling rescale without specifying mu and sigma doesn't change anything
all.equal(x, rescale(x))

# change the mean without changing the standard deviation
x1 <- rescale(x, mu = 3)
all.equal(mean(x1), 3)
all.equal(sd(x1), sd(x))

# rescale mean and standard deviation
x2 <- rescale(x, mu = 3, sigma = 2)
all.equal(mean(x2), 3)
all.equal(sd(x2), 2)
```

---

seatbelts

*Road Casualties in Great Britain 1969-84*

---

**Description**

Extract of the data in the [Seatbelts](#) dataset as a data frame. The original dataset is a multiple time series (class `mts`). Labels are in German.

**Usage**

```
seatbelts
```

**Format**

A data frame with 576 rows and 3 variables:

**date** data of the first data of the month for which the data was collected.

**seat** seat where the persons that were killed or seriously injured were seated. One of "Fahrer" (driver's seat), "Beifahrer" (front seat), "Rücksitz" (rear seat).

**victims** number of persons that were killed or seriously injured.

# Index

## \* datasets

bills, [2](#)

mtcars2, [5](#)

seatbelts, [6](#)

bills, [2](#)

define\_latex\_stats, [2](#)

density\_plot (distribution\_plot), [3](#)

distribution\_plot, [3](#)

install\_ibawds, [4](#)

mtcars, [5](#)

mtcars2, [5](#)

rescale, [5](#)

reshape2::tips, [2](#)

Seatbelts, [6](#)

seatbelts, [6](#)