

Package ‘datarium’

July 22, 2025

Type Package

Title Data Bank for Statistical Analysis and Visualization

Version 0.1.0

Date 2019-05-19

Description Contains data organized by topics: categorical data, regression model, means comparisons, independent and repeated measures ANOVA, mixed ANOVA and ANCOVA.

License GPL-2

Encoding UTF-8

LazyData true

Depends R (>= 3.1.0)

BugReports <https://github.com/kassambara/datarium/issues>

RoxygenNote 6.1.1

Suggests knitr

NeedsCompilation no

Author Alboukadel Kassambara [aut, cre]

Maintainer Alboukadel Kassambara <alboukadel.kassambara@gmail.com>

Repository CRAN

Date/Publication 2019-05-21 07:00:07 UTC

Contents

anxiety	2
depression	2
genderweight	3
headache	4
heartattack	4
housetasks.raw	5
jobsatisfaction	6
marketing	6
mice	7

mice2	7
performance	8
properties	8
selfesteem	9
selfesteem2	9
stress	10
titanic.raw	10
weightloss	11

Index	12
--------------	-----------

anxiety	<i>Anxiety Data for Two-Way Mixed ANOVA</i>
---------	---

Description

The data provide the anxiety score, measured at three time points, of three groups of individuals practicing physical exercises at different levels (grp1: basal, grp2: moderate and grp3: high)

Two-way mixed ANOVA can be used to evaluate if there is interaction between group and time in explaining the anxiety score.

Usage

```
data("anxiety")
```

Format

A data frame with 45 rows and 5 columns.

Examples

```
data(anxiety)
head(as.data.frame(anxiety))
```

depression	<i>Depression Data for Two Way Mixed ANOVA</i>
------------	--

Description

The data correspond to an experiment in which a treatment for depression is studied. Two groups of patients (1: control / 2: treatment) have been followed at five different times (0: pre-test, 1: one month post-test, 3: 3 months follow-up and 6: 6 months follow-up). The dependent variable is a depression score.

Repeated measures ANOVA can be performed in order to determine the effect of the treatment and the effect of time on the depression score.

Usage

```
data("depression")
```

Format

A data frame with 24 rows and 5 columns.

Examples

```
data(depression)
head(as.data.frame(depression))
```

genderweight

Weight Data By Gender for Two-Samples Mean Test

Description

Contains the weights by sex (M for male; F for female). The question is whether the average women's weight differs from the average men's weight?

A two-samples independent t-test can be performed to answer to this question.

Usage

```
data("genderweight")
```

Format

A data frame with 40 rows and 3 columns

Examples

```
data(genderweight)
head(as.data.frame(genderweight))
```

headache

Headache Data for Three Way ANOVA

Description

Measures of cholesterol concentration in 72 participants treated with three different drugs. The aim is to examine the potential of new class of drugs in lowering the cholesterol concentration and consequently reducing heart attack.

A pharmaceutical company tested three treatments for migraine headache sufferers. 72 participants were enrolled in the experiments. The aim is to examine the potential of new class of treatments in lowering the pain score associated with the migraine headache episode.

The participants include 36 males and 36 females. Males and females were further (equally) subdivided into whether they were at low or high risk of migraine headache.

This data set is suited for three way Anova test.

It contain the following variables:

- gender, which has two categories: "male" and "female";
- risk which has two levels: "low" and "high"
- treatment, which has three categories: "X", "Y" and "Z".

Usage

```
data("headache")
```

Format

A data frame with 72 rows and 4 columns.

Examples

```
data(headache)
head(as.data.frame(headache))
```

heartattack

Heart Attack Data for Three Way ANOVA

Description

Measures of cholesterol concentration in 72 participants treated with three different drugs. The aim is to examine the potential of new class of drugs in lowering the cholesterol concentration and consequently reducing heart attack.

The participants include 36 males and 36 females. Males and females were further (equally) subdivided into whether they were at low or high risk of heart attack.

This data set is suited for three way Anova test.

It contain the following variables:

- gender, which has two categories: "male" and "female";
- risk which has two levels: "low" and "high"
- drug, which has three categories: "A", "B" and "C".

Usage

```
data("heartattack")
```

Format

A data frame with 72 rows and 5 columns.

Examples

```
data(heartattack)
head(as.data.frame(heartattack))
```

housetasks.raw

Housetasks

Description

A data frame containing the frequency of execution of 13 house tasks in the couple.

Usage

```
data("housetasks.raw")
```

Format

A data frame with 1744 rows and 2 columns (tasks and status).

Examples

```
data(housetasks.raw)
table(housetasks.raw)
```

jobsatisfaction	<i>Job Satisfaction Data for Two-Way ANOVA</i>
-----------------	--

Description

Contains the job satisfaction score organized by gender and education level.

Usage

```
data("jobsatisfaction")
```

Format

A data frame with 58 rows and 3 columns.

Examples

```
data(jobsatisfaction)
head(as.data.frame(jobsatisfaction))
```

marketing	<i>Marketing Data Set</i>
-----------	---------------------------

Description

A data frame containing the impact of three advertising medias (youtube, facebook and newspaper) on sales. Data are the advertising budget in thousands of dollars along with the sales. The advertising experiment has been repeated 200 times.

Usage

```
data("marketing")
```

Format

A data frame with 200 rows and 4 columns.

Examples

```
data(marketing)
res.lm <- lm(sales ~ youtube*facebook, data = marketing)
summary(res.lm)
```

mice

Mice Weight Data for One Sample Mean Test

Description

Contains the weight of 10 mice. The question is whether the average weight of the mice differs from 25g.

A one sample t-test can be performed to answer to this question.

Usage

```
data("mice")
```

Format

A data frame with 10 rows and 2 columns

Examples

```
data(mice)
head(as.data.frame(mice))
```

mice2

Mice Weight Data for Paired-Samples Mean Test

Description

contains the weight of 10 mice before and after the treatment.

A paired-samples t-test can be performed to answer to this question.

Usage

```
data("mice2")
```

Format

A data frame with 10 rows and 3 columns

Examples

```
data(mice2)
head(as.data.frame(mice2))
```

performance

Performance Data for Three-Way Mixed ANOVA

Description

Contains the performance score measures of participants at two time points. The aim of this study is to evaluate the effect of gender and stress on performance score. The three-way mixed ANOVA test can be used to investigate this question.

The data include two between-subjects factors (gender and stress) and one within-subject factor (time, repeated measures).

Usage

```
data("performance")
```

Format

A data frame with 24 rows and 5 columns.

Examples

```
data(performance)
head(as.data.frame(performance))
```

properties

Properties Data for Chi-square Test of Independence

Description

Contains the type of properties and the buyer types. Buyer categories are: "single male", "single female", "married couple" and "family".

The type of property these buyers purchased were sorted into four categories: "flat", "bungalow" (i.e., a one-storey home), "detached house" and "terrace" (i.e., a block of adjoining houses).

Chi-square test of independence can be used to assess the association between the type of buyer who purchases a property and the type of property that is purchased.

Usage

```
data("properties")
```

Format

A data frame with 333 rows and 2 columns.

Examples

```
data("properties")
head(as.data.frame(properties))
```

selfesteem

Self-Esteem Score Data for One-way Repeated Measures ANOVA

Description

The dataset contains 10 individuals' self-esteem score on three time points during a specific diet to determine whether their self-esteem improved.

One-way repeated measures ANOVA can be performed in order to determine the effect of time on the self-esteem score.

Usage

```
data("selfesteem")
```

Format

A data frame with 10 rows and 4 columns.

Examples

```
data(selfesteem)
head(as.data.frame(selfesteem))
```

selfesteem2

Self Esteem Score Data for Two-way Repeated Measures ANOVA

Description

Data are the self esteem score of 12 individuals enrolled in 2 successive short-term trials (4 weeks) - control (placebo) and special diet trials.

The self esteem score was recorded at three time points: at the beginning (t1), midway (t2) and at the end (t3) of the trials.

The same 12 participants are enrolled in the two different trials with enough time between trials.

Two-way repeated measures ANOVA can be performed in order to determine whether there is interaction between time and treatment on the self esteem score.

Usage

```
data("selfesteem2")
```

Format

A data frame with 24 rows and 5 columns.

Examples

```
data(selfesteem2)
head(as.data.frame(selfesteem2))
```

stress	<i>Stress Data for Two-Way ANCOVA</i>
--------	---------------------------------------

Description

Researchers want to evaluate the effect of a new "treatment" and "exercise" on the stress score reduction after adjusting for "age".

Two-way ANCOVA can be performed in order to determine whether there is interaction between exercise and treatment on the stress score.

Usage

```
data("stress")
```

Format

A data frame with 60 rows and 5 columns.

Examples

```
data(stress)
head(as.data.frame(stress))
```

titanic.raw	<i>Survival of Passengers on the Titanic</i>
-------------	--

Description

Survival of passengers on the Titanic. This data set provides information on the fate of passengers on the fatal maiden voyage of the ocean liner 'Titanic'. Columns are economic status (Class), Sex, Age and Survived.

Usage

```
data("titanic.raw")
```

Format

A data frame with 2201 rows and 4 columns.

Examples

```
data(titanic.raw)
data("titanic.raw")
with(titanic.raw, table(Class, Survived))
```

weightloss

Weight Loss Score Data for Three-way Repeated Measures ANOVA

Description

A researcher wanted to assess the effects of Diet and Exercises on weight loss in 10 sedentary males. The participants were enrolled in four trials: (1) no diet and no exercises; (2) diet only; (3) exercises only; and (4) diet and exercises combined.

Each participant performed all four trials. The order of the trials was counterbalanced and sufficient time was allowed between trials to allow any effects of previous trials to have dissipated (i.e., a "wash out" period).

Each trial lasted nine weeks and the weight loss score was measured at the beginning of each trial (t1), at the midpoint of each trial (t2) and at the end of each trial (t3).

Three-way repeated measures ANOVA can be performed in order to determine whether there is interaction between diet, exercises and time on the weight loss score.

Usage

```
data("weightloss")
```

Format

A data frame with 48 rows and 6 columns.

Examples

```
data(weightloss)
head(as.data.frame(weightloss))
```

Index

anxiety, 2

depression, 2

genderweight, 3

headache, 4

heartattack, 4

housetasks.raw, 5

jobsatisfaction, 6

marketing, 6

mice, 7

mice2, 7

performance, 8

properties, 8

selfesteem, 9

selfesteem2, 9

stress, 10

titanic.raw, 10

weightloss, 11