

# Package: sdamr (via r-universe)

September 3, 2024

**Title** Statistics: Data Analysis and Modelling

**Version** 0.2.0

**Description** Data sets and functions to support the books `` Statistics: Data analysis and modelling" by Speekenbrink, M. (2021) <<https://mspeekenbrink.github.io/sdam-book/>> and `` An R companion to Statistics: data analysis and modelling" by Speekenbrink, M. (2021) <<https://mspeekenbrink.github.io/sdam-r-companion/>>. All datasets analysed in these books are provided in this package. In addition, the package provides functions to compute sample statistics (variance, standard deviation, mode), create raincloud and enhanced Q-Q plots, and expand Anova results into omnibus tests and tests of individual contrasts.

**License** GPL-3

**Encoding** UTF-8

**Language** en-GB

**LazyData** true

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.2.1

**URL** <https://mspeekenbrink.github.io/sdam-r/>

**Depends** R (>= 3.5.0)

**Imports** methods, stats, dplyr, ggplot2, grid, car

**Repository** <https://mspeekenbrink.r-universe.dev>

**RemoteUrl** <https://github.com/mspeekenbrink/sdam-r>

**RemoteRef** HEAD

**RemoteSha** d637c71b139944b2952a5b143335c0920f3a6268

## Contents

anchoring . . . . .	2
---------------------	---

center . . . . .	3
cheerleader . . . . .	4
expand_Anova . . . . .	5
expBelief . . . . .	6
fifa2010 . . . . .	7
fifa2010teams . . . . .	7
GeomFlatViolin . . . . .	8
geom_flat_violin . . . . .	8
gestures . . . . .	9
legacy2015 . . . . .	10
metacognition . . . . .	11
papervotes . . . . .	12
plot_qq_marginals . . . . .	12
plot_raincloud . . . . .	14
position_jitternudge . . . . .	15
redist2015 . . . . .	16
rps . . . . .	17
sample_mode . . . . .	18
sample_sd . . . . .	18
sample_var . . . . .	19
speeddate . . . . .	20
tetris2015 . . . . .	21
trump2016 . . . . .	22
uefa2008 . . . . .	23
<b>Index</b>	<b>24</b>

---

anchoring	<i>Anchoring</i>
-----------	------------------

---

## Description

Numerical judgments of the height of the Mount Everest after a low or high anchor. This dataset comes from the ManyLabs 1 study

## Usage

anchoring

## Format

A data frame with 4632 rows and 5 variables:

**session\_id** Unique identifier for participants  
**sex** Sex of participant (f = female, m = male)  
**age** Age of participant in years  
**citizenship** Country code of citizenship

**referrer** Location of data collection. Site abbreviations used here can be matched up to the full site name in the online supplement <https://osf.io/wx7ck/>

**us\_or\_international** Was the study conducted on a US sample or international sample?

**lab\_or\_online** Was the study conducted online or in-lab?

**anchor** anchor, whether high or low

**everest\_feet** judged height of Mount Everest in feet. Converted from meters if given in meters.

**everest\_meters** judged height of Mount Everest in meters. Only contains values when judgment was actually given in meters.

### Source

<https://osf.io/pqf9r/>. See also Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, S., Bernstein, M. J., . . . , Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. *Social Psychology*, 45(3), 142-152. doi:10.1027/18649335/a000178

---

center

*Mean-centered values*

---

### Description

center computes mean-centered values. It is a convenience wrapper to scale, equal to scale(x, scale=FALSE)

### Usage

```
center(x)
```

### Arguments

x                    Numeric vector

### Value

A numeric vector with mean-centered values

### Examples

```
data(anchoring)
center(anchoring$everest_feet)
```

---

cheerleader	<i>Data from Experiment 1 of Carragher, D.J., Thomas, N.A., Gwinn, O.S. et al. (2019) Limited evidence of hierarchical encoding in the cheerleader effect. Scientific Reports, 9, 9329. <a href="https://doi.org/10.1038/s41598-019-45789-6">https://doi.org/10.1038/s41598-019-45789-6</a></i>
-------------	---

---

## Description

\@format A data frame with 320 observations of 16 variables:

**Participant** (factor) Participant ID

**Age** (numeric) Participant age in years

**Sex** (factor) Participant sex (Male or Female)

**Task** (factor) Identical-Distractors, or Self-Distractors.

**LineClickAccuracy** Measure of average response deviation from the visual analogue scale; scores > +/- 2.00 constitute exclusion.

**Excluded** (numeric) Indicator whether participant was excluded from main analysis (0 = no, 1 = yes)

**WhyExcluded** (character) explanation for exclusion

**Item** (factor) Item description

**Response** Attractiveness rating for the target face on a visual analogue scale ranging from “Very Unattractive” (0) to “Very Attractive” (100)

## Usage

cheerleader

## Format

An object of class `data.frame` with 192 rows and 9 columns.

## Source

<https://osf.io/je5u7/>. Carragher, D.J., Thomas, N.A., Gwinn, O.S. et al. (2019) Limited evidence of hierarchical encoding in the cheerleader effect. *Scientific Reports*, 9, 9329 doi:10.1038/s41598019457896.

---

expand_Anova	<i>Expand all contrast terms in car::Anova</i>
--------------	--

---

## Description

expand\_Anova is an experimental function to add more detailed results to those returned by `car::Anova`. In particular, expand\_Anova aims to provide test results for all individual contrasts assigned to the factors in a linear model, in addition to the omnibus tests returned by `car::Anova`.

## Usage

```
expand_Anova(mod, type = c("III", "II", 3, 2), ...)
```

## Arguments

mod	A model of class <code>lm</code> (see <code>?stats::lm</code> )
type	SS Type (see <code>?car::Anova</code> )
...	Further arguments passed to <code>Anova</code>

## Details

This is an experimental function

## Value

Object of class `anova` returned by `car::Anova`

## See Also

[car::Anova\(\)](#) for more information about the Anova tables, and [stats::lm\(\)](#) for information about how to specify the model

## Examples

```
data("tetris2015")
mod <- lm(Days_One_to_Seven_Number_of_Intrusions ~ Condition, data=tetris2015)
car::Anova(mod,type=3) # default type III Anova table
expand_Anova(mod,type=3)
```

---

expBelief                      *Data from Experiment 5 of Gilder, T. S. E., & Heerey, E. A. (2018). The Role of Experimenter Belief in Social Priming. Psychological Science, 29(3), 403–417.*

---

### Description

\@format A data frame with 400 observations of 16 variables:

**pid** Participant ID  
**exptrNum** Experimenter Number  
**age** Participant age in years  
**gender** Participant self-reported gender  
**yearInUni** Year in University  
**ethnicity** Self-reported ethnicity  
**englishFluency** Self-reported English fluency (1=beginner; 7=native language)  
**experimenterBelief** Experimenter Belief (H: High or L: Low)  
**primeCond** Actual Prime Condition (HPP: High-power prime or LPP: low-power prime)  
**powerPRE** Self-reported power BEFORE the manipulation  
**powerPOST** Self-reported power AFTER the manipulation  
**ApproachAdvantage** Approach advantage (Avoid RT - Approach RT; see manuscript)  
**attractive** Rating of experimenter ATTRACTIVENESS  
**competent** Rating of experimenter COMPETENCE  
**friendly** Rating of experimenter FRIENDLINESS  
**trustworthy** Rating of experimenter TRUSTWORTHINESS

### Usage

expBelief

### Format

An object of class `data.frame` with 400 rows and 16 columns.

### Source

<https://osf.io/un4h6/>. See also Gilder, T. S. E., & Heerey, E. A. (2018). The Role of Experimenter Belief in Social Priming. *Psychological Science*, 29(3), 403–417. doi:10.1177/0956797617737128.

---

`fifa2010`*Predictions by Paul the Octopus in the 2010 FIFA World Cup.*

---

**Description**

A dataset containing the predictions and outcomes of matches in the 2010 FIFA European Cup.

**Usage**`fifa2010`**Format**

A data frame with 8 rows and 4 variables:

**Match** countries playing

**Prediction** country predicted to win

**Result** score at the end of the match

**Outcome** whether Paul was correct or incorrect

**Source**

[https://en.wikipedia.org/wiki/Paul\\_the\\_Octopus](https://en.wikipedia.org/wiki/Paul_the_Octopus)

---

`fifa2010teams`*FIFA 2010 team statistics*

---

**Description**

Statistics for all teams playing in the 2010 FIFA world cup.

**Usage**`fifa2010teams`**Format**

A data frame with 11 variables and 32 rows

**nr** Unique numeric identifier for each team

**team** Name of the team (i.e. country)

**matches\_played** Number of matches played

**goals\_for** Total goals counted against their opponents

**goals\_scored** Total goals scored against their opponents

**goals\_against** Goals counted against the team  
**penalty\_goal** Number of penalty goals scored  
**own\_goals\_for** Number of own goals  
**yellow\_cards** Number of yellow cards  
**indirect\_red\_cards** Number of indirect red cards  
**direct\_red\_cards** Number of direct red cards

### Source

FIFA website. <https://www.fifa.com/worldcup/archive/southafrica2010/statistics/teams/goal-scored> and <https://www.fifa.com/worldcup/archive/southafrica2010/statistics/teams/disciplinary>

---

GeomFlatViolin	<i>Flat violin geometry</i>
----------------	-----------------------------

---

### Description

Flat violin geometry

---

geom_flat_violin	<i>Half violin plot</i>
------------------	-------------------------

---

### Description

Half violin plot

### Usage

```
geom_flat_violin(
  mapping = NULL,
  data = NULL,
  stat = "ydensity",
  position = "dodge",
  trim = TRUE,
  scale = "area",
  show.legend = NA,
  inherit.aes = TRUE,
  ...
)
```



**Arguments**

mapping	The mapping
data	data.frame
stat	statistic (don't change)
position	position dodge
trim	Logical
scale	Scale (don't change)
show.legend	Logical
inherit.aes	Logical
...	other arguments

**Value**

A layer for a `ggplot2::ggplot` object, similar to e.g. `ggplot2::geom_violin`.

**Source**

[urlhttps://gist.github.com/dgrtwo/eb7750e74997891d7c20](https://gist.github.com/dgrtwo/eb7750e74997891d7c20)

**See Also**

[ggplot2::geom\\_violin\(\)](#), which provided the basis of this function.

**Examples**

```
library(ggplot2)
data(diamonds)
ggplot(diamonds, aes(cut, carat)) + geom_flat_violin() + coord_flip()
```

---

gestures

*Data from Winter, B., & Burkner, P. (2021) Poisson regression for linguists: A tutorial introduction to modelling count data with brms. Language and Linguistics Compass, 15, e12439*  
 Rhref<https://doi.org/10.1111/lnc3.12439>  
[doi:10.1111/lnc3.12439](https://doi.org/10.1111/lnc3.12439)

---

**Description**

\@format A data frame with 54 observations of 6 variables:

**ID** (factor) Participant ID

**context** (factor) Whether talking to a friend or professor

**duration** (numeric) Duration of the interaction

**language** (factor) Language spoken: Catalan or Korean

**gender** (factor) Participant gender (M = male, F = female)

**gestures** (numeric) number of gestures in the interaction.

**Usage**

gestures

**Format**

An object of class `data.frame` with 54 rows and 6 columns.

**Source**

<https://osf.io/6j8kc>.

---

legacy2015

*Legacy motives and pro-environmental behaviour*

---

**Description**

Legacy motives and pro-environmental behaviour

**Usage**

legacy2015

**Format**

A data frame with 245 rows and 9 variables:

**id** (numeric) ID variable relating to the original dataset

**sex** (character) biological sex of participant (male or female)

**age** (numeric) age in years

**legacy** (numeric) Sverage of 8 items reflecting legacy motivation, on a scale from 1 (Not at all) to 6 (A great amount)

**belief** (numeric) average of 5 items reflecting belief in climate change, on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree)

**intention** (numeric) average of 8 items reflecting intention to act in a pro-environmental way, on a scale from 1 (Never) to 6 (All the time)

**education** (numeric) Level of education, 1 = 8th grade or less, 2 = Some high school, 3 = Graduated high school, 4 = Some college or technical school, 5 = Graduated college or technical school, 6 = Post-graduate

**income** (numeric) Approximate household income, 1 = less than \$20K, 2 = \$20K-\$35K, 3 = \$35K-\$50K 4 = \$50K-\$75K, 5 = \$75K-100K, 6 = more than 100K

**donation** (numeric) Donation of possible bonus payment, between \$0 and \$10

**Source**

Harvard DataVerse <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27740&version=1.0>

**Examples**

```
## Not run:
# this dataset was processed from the raw data as follows:
tdat <- read.csv("legacy study - pilot_data.csv")

## End(Not run)
```

---

metacognition	<i>Data from Rausch, M. &amp; Zehetleitner, M. (2016) Visibility is not equivalent to confidence in a low contrast orientation discrimination task. Frontiers in Psychology, 7, p. 591. Rhref<a href="https://doi.org/10.3389/fpsyg.2016.00591">https://doi.org/10.3389/fpsyg.2016.00591</a></i>
---------------	--

---

**Description**

\@format A data frame with 7560 observations of 10 variables:

**id** (factor) Participant ID

**age** (numeric) Participant age in years

**sex** (factor) Participant sex (male or female)

**block** (numeric) number of the test block (from 1 to 9). Practice block is excluded.

**trial** (numeric) number of trial (between 1 and 42) within a block.

**tilt** (numeric) whether grating is horizontal (0) or vertical (90)

**contrast** (numeric) contrast of grating shown

**correct** (numeric) Whether identified title was correct (1) or not (0)

**visibility** (numeric) Rated visibility of the stimulus, on a scale between 0 () and 100 ()

**confidence** (numeric) Rated confidence in tilt identification, on a scale between 0 () and 100 ()

**Usage**

```
metacognition
```

**Format**

An object of class `data.frame` with 7560 rows and 10 columns.

**Source**

<https://osf.io/vk6fe/>. Rausch, M. & Zehetleitner, M. (2016) Visibility is not equivalent to confidence in a low contrast orientation discrimination task. *Frontiers in Psychology*, 7, p. 591 [doi:10.3389/fpsyg.2016.00591](https://doi.org/10.3389/fpsyg.2016.00591).

---

`papervotes`      *Data based on a post-election survey by YouGov (see <https://yougov.co.uk/topics/politics/articles-reports/2017/06/13/how-britain-voted-2017-general-election>). Note that the data was recreated by combining frequency and percentage results reported in [https://d25d2506sfb94s.cloudfront.net/cumulus\\_uploads/document/smo1w49ph1/InternalResults\\_170613\\_2017Election\\_Demographics\\_W.pdf](https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/smo1w49ph1/InternalResults_170613_2017Election_Demographics_W.pdf). Due to rounding and other potential inconsistencies, this data set will likely differ from the actual results.*

---

### Description

\@format A data frame with 90 observations of 3 variables:

**newspaper** (factor) Reported newspaper read most often

**vote** (factor) Which party voted on (including "did not vote")

**n** (numeric) Number of people in the survey who responded with that combination of newspaper and vote

### Usage

```
papervotes
```

### Format

An object of class `data.frame` with 90 rows and 3 columns.

### Source

[https://d25d2506sfb94s.cloudfront.net/cumulus\\_uploads/document/smo1w49ph1/InternalResults\\_170613\\_2017Election\\_Demographics\\_W.pdf](https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/smo1w49ph1/InternalResults_170613_2017Election_Demographics_W.pdf).

---

`plot_qq_marginals`      *Q-Q plots with distributions in the margins*

---

### Description

`plot_qq_marginals` creates an enhanced Q-Q plot with the observed and theoretical distributions shown in the margins of the plot.

**Usage**

```
plot_qq_marginals(  
  x,  
  breaks = "Sturges",  
  newpage = TRUE,  
  xlab = "Observed Quantiles",  
  ylab = "Theoretical quantiles",  
  xlim = grDevices::extendrange(c(min(x), max(x))),  
  ylim = NULL,  
  main = NULL,  
  sub = NULL,  
  axes = TRUE,  
  border = TRUE,  
  ...  
)
```

**Arguments**

x	A numeric vector
breaks	How to compute breakpoints for the histogram. See <code>?hist</code>
newpage	(logical) Should the plot be plotted on a new page?
xlab	Label for x-axis
ylab	Label for y-axis
xlim	Range of x values shown
ylim	Range of y values shown
main	Main title
sub	Subtitle
axes	(logical) Draw axes?
border	(logical) Draw a border?
...	Further arguments

**Value**

No return value. The function adds a plot to the active graphics window.

**Examples**

```
data(anchoring)  
plot_qq_marginals(anchoring$everest_feet)
```

---

plot\_raincloud      *Create a raincloud plot*

---

## Description

plot\_raincloud creates a raincloud plot to display the distribution of data by a combination of a boxplot, a kernel density plot, and a scatterplot. The boxplot includes the median (displayed as a horizontal line) and the mean (displayed as a point). It does not indicate potential outliers, as these can be seen in the scatter plot. The kernel density plot provides a nonparametric estimate of the distribution. The scatterplot depicts all values in y with random jittering on the x-axis. The data can be grouped by supplying a grouping factor in the groups argument, in which case multiple raincloud plots are shown side by side. As plot\_raincloud provides a ggplot2::ggplot object, it can be combined with further layers and functionality from the ggplot2 package.

## Usage

```
plot_raincloud(data, y, horizontal = FALSE, groups, point_size = 0.5, ...)
```

## Arguments

data	Data.frame (or tibble)
y	The unquoted name of the variable in data for which to create the raincloud plot
horizontal	(logical) change the orientation of the plot
groups	An unquoted name of grouping variable in data (ideally a factor)
point_size	Size of the jittered points
...	Other arguments, passed to ggplot(aes(...))

## Value

An object of class gg, i.e. a ggplot object from the ggplot2 package

## Source

Allen M, Poggiali D, Whitaker K et al. Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res 2019, 4:63 ([doi:10.12688/wellcomeopenres.15191.1](https://doi.org/10.12688/wellcomeopenres.15191.1))

## See Also

[ggplot2::ggplot\(\)](#) for information about ggplot objects, [ggplot2::theme\(\)](#) for information about changing various aspects of the plot, and [ggplot2::facet\\_wrap\(\)](#) and [ggplot2::facet\\_grid\(\)](#) for creating multiple raincloud plots for different levels of grouping factors beyond those specified in groups.

**Examples**

```

data(anchoring)
plot_raincloud(anchoring,y=everest_feet)
plot_raincloud(anchoring,y=everest_feet,groups=anchor)
plot_raincloud(anchoring,y=everest_feet,groups=anchor) +
  ggplot2::facet_wrap(~us_or_international) +
  ggplot2::ylab("How high is Mount Everest (in feet)?")

```

---

position\_jitternudge *Simultaneously nudge and jitter*

---

**Description**

Simultaneously nudge and jitter

**Usage**

```

position_jitternudge(
  jitter.width = NULL,
  jitter.height = 0,
  nudge.x = 0,
  nudge.y = 0,
  seed = NA
)

```

**Arguments**

jitter.width	degree of jitter in x direction. Defaults to 40% of the resolution of the data.
jitter.height	degree of jitter in y direction. Defaults to 0.
nudge.x	the amount to nudge in the x direction.
nudge.y	the amount to nudge in the y direction.
seed	Optional seed for the random jitter

**Value**

Positions for data in a `ggplot2::ggplot` object, similar to e.g. `ggplot2::position_jitter`

**See Also**

[ggplot2::position\\_jitter\(\)](#), which is the basis of this function.

**Examples**

```

library(ggplot2)
dsub <- diamonds[ sample(nrow(diamonds), 1000), ]
ggplot(dsub, aes(x = cut, y = carat, fill = clarity)) +
  geom_boxplot(outlier.size = 0) +
  geom_point(pch = 21, position = position_jitterdodge())

```

redist2015

*Redistribution of wealth***Description**

It is generally found that wealthy people tend to be more opposed to policies to reduce wealth inequalities. This may be unsurprising from a classical economic standpoint, because the material burden of the redistribution of wealth will fall on wealthier people. Wealthier people are also more likely than poorer people to adopt political ideologies that oppose redistribution policies. Dawtry, Sutton, and Sibley (2015) investigated whether, in addition to such factors, “social-sampling processes” lead wealthier people to oppose redistribution policies. Social sampling is the idea that people (partly) base inferences on their social surroundings. Wealthier people tend to live in more affluent areas and move in wealthier social circles. This may bias their view of the world, where wealthier people estimate the general population to be wealthier (with less of a gap between the wealthy and the poor) than it really is.

**Usage**

redist2015

**Format**

A data frame with 305 rows and 12 variables:

**id** unique ID number for each participant

**gender** only "male" or "female" could be answered by the looks of it

**age** participant age in years

**income** yearly household income (in units of \$1,000)

**pol\_att** political leaning from 1="Extremely Liberal" to 9="Extremely Conservative"

**PD\_mean** estimate average household income in the general US population

**PD\_gini** GINI index computed for a subjective distribution of wealth in the general US population. The GINI index is a measure of wealth inequality; higher numbers mean more inequality

**PD\_fair** Answer to the question "To what extent do you feel that household incomes are fairly–unfairly distributed across the US population?" on a scale from 1="Extremely Fair" to 9 = "Extremely Unfair".

**PD\_sat** Answer to the question "How satisfied–dissatisfied are you with the way in which household incomes are distributed across the US population?" on a scale from 1="Extremely satisfied" to 9="Extremely dissatisfied".

**SC\_mean** estimate average household income in the participant’s social circles

**SC\_gini** (subjective) inequality in the participant’s social circles

**redist** support for wealth redistribution policies (average of four items, higher scores indicate stronger support).



## Details

In Experiment 1a of Dawtry, Sutton, and Sibley (2015), they assessed income and opinions for n=305 online U.S. participants recruited via Amazon's Mechanical Turk.

## Source

<https://osf.io/3mftr/>. See also Dawtry, Rael J., Robbie M. Sutton, and Chris G. Sibley. 2015. "Why Wealthier People Think People Are Wealthier, and Why It Matters: From Social Sampling to Attitudes to Redistribution." *Psychological Science* 26 (9): 1389–1400. doi:10.1177/0956797615586560.

---

rps	<p><i>Data from Experiment 1 in Guennouni, I., Speekenbrink, M. (2022). Transfer of learned opponent models in repeated games. Computational Brain and Behaviour, 5, 326–342</i>  <i>Rhref</i><a href="https://doi.org/10.1007/s42113-022-00133-6">https://doi.org/10.1007/s42113-022-00133-6</a>  <i>doi:10.1007/s42113022001336. Participants (n=52) each play 50 rounds of Rock-Paper-Scissors against an AI player who either adopts a "level-1" or "level-2" strategy. A level-1 strategy assumes the opponent will repeat their last action, and chooses the action that beats this. A level-2 strategy assumes the opponent adopts a level-1 strategy, and chooses the action that beats this. On 10% of rounds, the AI players pick a random action. On the remainder, they act according to their strategy.</i></p>
-----	--

---

## Description

\@format A data frame with 2600 observations of 6 variables:

**id** (factor) Participant ID

**ai\_strategy** (factor) Strategy adopted by AI player

**round** (numeric) Round number (between 1 and 50)

**human\_action** (factor) Action taken by human (rock, paper, or scissors)

**ai\_action** (factor) Action taken by AI (rock, paper, or scissors)

**score** (numeric) Outcome for human player, where 1 indicates a win, -1 a loss, and 0 a tie

## Usage

rps

## Format

An object of class `data.frame` with 2600 rows and 6 columns.

## Source

Guennouni, I., Speekenbrink, M. (2022). Transfer of learned opponent models in repeated games. *Computational Brain and Behaviour*, 5, 326–342. doi:10.1007/s42113022001336

---

sample_mode	<i>Compute a sample mode</i>
-------------	------------------------------

---

**Description**

sample\_mode computes the sample mode, i.e. the value in `x` with the highest frequency of occurrence. If there are multiple modes, the mode that occurs first in `x` is returned, with a warning that lists the other modes found.

**Usage**

```
sample_mode(x)
```

**Arguments**

<code>x</code>	Numeric vector
----------------	----------------

**Value**

A single numeric value equal to the sample mode

**Examples**

```
data(anchoring)
sample_mode(anchoring$everest_feet)
# Multiple modes give a warning:
sample_mode(c(3,3,3,1,1,1,2,2,2))
```

---

sample_sd	<i>Compute the sample standard deviation</i>
-----------	--

---

**Description**

sample\_sd computes the sample standard deviation, i.e. the square root of the sum of squared deviations of `x` from the mean divided by the total number of observations. This is in contrast to `sd`, which computes an unbiased estimate of the standard deviation (i.e. it divides the sum of squared deviations by `n - 1`).

**Usage**

```
sample_sd(x, na.rm = FALSE)
```

**Arguments**

<code>x</code>	Numeric vector
<code>na.rm</code>	(logical) Should missing values be removed?

**Value**

A single numeric value equal to the sample variance

**Examples**

```
data(anchoring)
sample_sd(anchoring$everest_feet)
```

---

sample_var	<i>Compute the sample variance</i>
------------	------------------------------------

---

**Description**

sample\_var computes the sample variance, i.e. the sum of squared deviations of  $x$  from the mean divided by the total number of observations. This is in contrast to `var`, which computes an unbiased estimate of the variance (i.e. it divides the sum of squared deviations by  $n - 1$ ).

**Usage**

```
sample_var(x, na.rm = FALSE)
```

**Arguments**

<code>x</code>	Numeric vector
<code>na.rm</code>	(logical) Should missing values be removed?

**Value**

A single numeric value equal to the sample variance

**Examples**

```
data(anchoring)
sample_var(anchoring$everest_feet)
```

speeddate

*Speed dating***Description**

A subset of cases (wave 6-9) and variables (see below) from an experiment on speed dating. by Columbia Business School professors Ray Fisman and Sheena Iyengar for their paper Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment.

**Usage**

speeddate

**Format**

A data frame with 1562 rows and 32 variables:

**iid** (numeric) unique ID variable of participant

**pid** (numeric) unique ID variable of date partner

**gender** (character) gender of participant (male or female)

**age** (numeric) age in years

**date\_like** (numeric) how much they like their date partner in general (between 1 and 10)

**other\_like** (numeric) how much their date partner likes them (between 1 and 10)

**date\_want** do they want to go on another date with their date partner? (1 = yes, 0 = no)

**other\_want** does their date partner want to go on another date with them? (1 = yes, 0 = no)

**match** do they both want to go on another date with each other? (1 = yes, 0 = no)

**self\_attr** how attractive do they think they are? (between 1 and 10)

**self\_sinc** how sincere do they think they are? (between 1 and 10)

**self\_intel** how intelligent do they think they are? (between 1 and 10)

**self\_fun** how much fun do they think they are? (between 1 and 10)

**self\_amb** how ambitious do they think they are? (between 1 and 10)

**other\_attr,other\_sinc,other\_intel,other\_fun,other\_amb** how attractive etc does their date partner think they are? (between 1 and 10)

**other\_shar** how much does their date partner think they share hobbies and interests? (between 1 and 10)

**date\_attr,date\_sinc,date\_intel,date\_fun,date\_amb,date\_shar** how do they rate their date partner's attractiveness etc? (between 1 and 10)

**self\_imp\_attr,self\_imp\_sinc,self\_imp\_intel,self\_imp\_fun,self\_imp\_amb,self\_imp\_shar** how important do they find attractiveness etc in a partner? (between 1 and 10)

**other\_imp\_attr,other\_imp\_sinc,other\_imp\_intel,other\_imp\_fun,other\_imp\_amb,other\_imp\_shar** how important does their date partner find attractiveness etc? (between 1 and 10)

**Source**

Kaggle <https://www.kaggle.com/annavictoria/speed-dating-experiment>

tetris2015

*Tetris and intrusive memories***Description**

Tetris and intrusive memories

**Usage**

tetris2015

**Format**

A data frame with 72 rows and 28 variables:

**Condition** (factor) Condition: Control, Tetris\_Reactivation, Tetris, or Reactivation**Time\_of\_Day** Time of day participant commenced experiment, either "morning" or "afternoon"**BDI\_II** Beck Depression Inventory-II (BDI-II): Total score**STAI\_T** Spielberger State-Trait Anxiety Trait scale (STAD): Total score**pre\_film\_VAS\_Sad** Self-rated level of Sadness: Pre-film VAS mood. VAS = visual analogue scale. All VAS mood scales anchored from "not at all" to "extremely" in response to the question "Right at this very moment I am feeling". Composite for pre-film mood calculated by summing the six pre-film VAS mood ratings**pre\_film\_VAS\_Hopeless** Self-rated level of Hopelessness: Pre-film VAS mood**pre\_film\_VAS\_Depressed** Self-rated level of Depressed: Pre-film VAS mood**pre\_film\_VAS\_Fear** Self-rated level of Fear: Pre-film VAS mood**pre\_film\_VAS\_Horror** Self-rated level of Horror: Pre-film VAS mood**pre\_film\_VAS\_Anxious** Self-rated level of Anxiousness: Pre-film VAS mood**post\_film\_VAS\_Sad** Self-rated level of Sadness: Post-film VAS mood. Composite for post-film mood calculated by summing the six post-film VAS mood ratings**post\_film\_VAS\_Hopeless** Self-rated level of Hopelessness: Post-film VAS mood**post\_film\_VAS\_Depressed** Self-rated level of Depressed: Post-film VAS mood**post\_film\_VAS\_Fear** Self-rated level of Fear: Post-film VAS mood**post\_film\_VAS\_Horror** Self-rated level of Horror: Post-film VAS mood**post\_film\_VAS\_Anxious** Self-rated level of Anxious: Post-film VAS mood**Attention\_Paid\_to\_Film** Attention paid to the film rating: How much attention did you pay to the film from 0-not at all to 10-extremely**Post\_film\_Distress** Post film distress rating: How distressing did you find the film from 0-not at all to 10-extremely**Day\_Zero\_Number\_of\_Intrusions** Day 0: Number of image-based intrusive memories in the Intrusion Diary (pre-intervention)

**Days\_One\_to\_Seven\_Number\_of\_Intrusions** Days 1-7: Number of image-based intrusive memories in the Intrusion Diary (post-intervention)

**Visual\_Recognition\_Memory\_Test** Visual recognition memory test score: Number of correct responses (out of 22)

**Verbal\_Recognition\_Memory\_Test** Verbal recognition memory test score: Number of correct responses (out of 32)

**Number\_of\_Provocation\_Task\_Intrusions** Intrusion Provocation Task (IPT): Number of image-based intrusive memories during 2min laboratory task on Day 7

**Diary\_Compliance** Diary compliance rating - indicate how accurate you think your diary is from 1 - not at all accurate to 10 extremely

**IES\_R\_Intrusion\_subscale** Impact of Event Scale-Revised (IES-R): Intrusion Subscale

**Tetris\_Total\_Score** Tetris game play computer score total - cumulative (sum of all games). Only participants who played Tetris have data relating to Tetris\_Total\_Score

**Self\_Rated\_Tetris\_Performance** Self-rated Tetris performance: How difficult or easy did you find the game you just played. Only participants who played Tetris have data relating to Self\_Rated\_Tetris\_Performance.

**Tetris\_Demand\_Rating** Demand rating: How much did you think Tetris after a distressing film would increase or decrease intrusive memories of the film: -10: extremely decrease, to +10: extremely increase

#### Source

<https://osf.io/ideta/>. See also James et al., 'Computer Game Play Reduces Intrusive Memories of Experimental Trauma via Reconsolidation-Update Mechanisms'.

---

trump2016

*Trump votes in 2016 for 50 US states and the District of Columbia*

---

#### Description

Trump votes in 2016 for 50 US states and the District of Columbia

#### Usage

trump2016

#### Format

A data frame with 4632 rows and 5 variables:

**state** Name of the state

**hate\_groups** Number of hate groups in the state in 2016 as reported by the Southern Poverty Law Center (<https://www.splcenter.org/hate-map>)

**population** Number of citizens in the state in 2016

**hate\_groups\_per\_million** Number of hate groups per million citizens  
**percent\_bachelors\_degree\_or\_higher** Percentage of citizens with a bachelor's degree of higher  
**percent\_in\_poverty** Percentage of citizens below the poverty threshold  
**percent\_Trump\_votes** Percentage of votes for Trump in the 2016 elections

### Source

CSI Without Dead Bodies "Hate Groups and Trump's Vote%: Predictive effect present when education and poverty are considered" <https://web.archive.org/web/20210414051437/https://www.csiwithoutdeadbodies.com/2017/02/hate-groups-and-trumps-vote-predictive.html>

---

uefa2008

*Predictions by Paul the Octopus in the 2008 UEFA Cup.*

---

### Description

A dataset containing the predictions and outcomes of matches in the 2008 UEFA European Cup.

### Usage

uefa2008

### Format

A data frame with 6 rows and 4 variables:

**Match** countries playing

**Prediction** country predicted to win

**Result** score at the end of the match

**Outcome** whether Paul was correct or incorrect

### Source

[https://en.wikipedia.org/wiki/Paul\\_the\\_Octopus](https://en.wikipedia.org/wiki/Paul_the_Octopus)

# Index

## \* datasets

- anchoring, [2](#)
- cheerleader, [4](#)
- expBelief, [6](#)
- fifa2010, [7](#)
- fifa2010teams, [7](#)
- GeomFlatViolin, [8](#)
- gestures, [9](#)
- legacy2015, [10](#)
- metacognition, [11](#)
- papervotes, [12](#)
- redist2015, [16](#)
- rps, [17](#)
- speeddate, [20](#)
- tetris2015, [21](#)
- trump2016, [22](#)
- uefa2008, [23](#)

## \* position adjustments

- position\_jitternudge, [15](#)

anchoring, [2](#)

car::Anova(), [5](#)

center, [3](#)

cheerleader, [4](#)

expand\_Anova, [5](#)

expBelief, [6](#)

fifa2010, [7](#)

fifa2010teams, [7](#)

geom\_flat\_violin, [8](#)

GeomFlatViolin, [8](#)

gestures, [9](#)

ggplot2::facet\_grid(), [14](#)

ggplot2::facet\_wrap(), [14](#)

ggplot2::geom\_violin(), [9](#)

ggplot2::ggplot(), [14](#)

ggplot2::position\_jitter(), [15](#)

ggplot2::theme(), [14](#)

legacy2015, [10](#)

metacognition, [11](#)

papervotes, [12](#)

plot\_qq\_marginals, [12](#)

plot\_raincloud, [14](#)

position\_jitternudge, [15](#)

PositionJitternudge (GeomFlatViolin), [8](#)

redist2015, [16](#)

rps, [17](#)

sample\_mode, [18](#)

sample\_sd, [18](#)

sample\_var, [19](#)

speeddate, [20](#)

stats::lm(), [5](#)

tetris2015, [21](#)

trump2016, [22](#)

uefa2008, [23](#)