Package: adventr (via r-universe)

September 6, 2024

Type Package

Title Interactive R Tutorials to Accompany Field (2016), ``An Adventure in Statistics''

Version 0.1.10

Maintainer Andy Field <andyf@sussex.ac.uk>

Description Interactive 'R' tutorials written using 'learnr' for Field (2016), `` An Adventure in Statistics'', <ISBN:9781446210451>. Topics include general workflow in 'R' and 'Rstudio', the 'R' environment and 'tidyverse', summarizing data, model fitting, central tendency, visualising data using 'ggplot2', inferential statistics and robust estimation, hypothesis testing, the general linear model, comparing means, repeated measures designs, factorial designs, multilevel models, growth models, and generalized linear models (logistic regression).

License GPL-3

Depends R (>= 3.6.0), learnr

LazyData true

Imports BayesFactor, boot, dplyr, car, effsize, effects, forcats, GGally, ggplot2, Hmisc, lm.beta, nlme, readr, robust, sandwich, sjstats, tidyr, WRS2

URL http://milton-the-cat.rocks/home/adventr.html, https://github.com/profandyfield/adventr

BugReports https://github.com/profandyfield/adventr/issues

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

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

RemoteUrl https://github.com/profandyfield/adventr

RemoteRef HEAD

RemoteSha 2d894ca2d41fafdd9df1e903e6f9e48e739024e5

Contents

---

adventr   *adventr: tutorials in R for An Adventure in Statistics (Field, 2016).*

Description

The adventr package contains a series of tutorials that accompany my books An Adventure in Statistics and Discovering statistics using R. These tutorials contain abridged sections from my books so there are some copyright considerations but I offer them under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. (Basically you can use them for teaching and non-profit activities but not meddle with them or claim them as your own work.)

Who is the package aimed at?

Anyone teaching from or reading An Adventure in Statistics or Discovering statistics using R might find them useful.

What is covered?

The tutorials *do not* teach the background theory (the books do): it is assumed you have either attended my lecture or read the relevant chapter in the aforementioned books (or someone else's) The aim of this tutorial is to augment the theory that you already know by guiding you through fitting linear models using R and RStudio and asking you questions to test your knowledge along the way.

Running a tutorial

To run a tutorial execute:

```
learnr::run_tutorial("name_of_tutorial", package = "adventr")
```

replacing *name_of_tutorial* with the name of the tutorial (see below). For example, to run the tutorial on data basics execute:

```
learnr::run_tutorial("adventr_02", package = "adventr")
```

A list of tutorials

- **adventr_02**: Data basics in R and RStudio. This tutorial looks at the basic concepts in R, such as installing packages, setting the working directory, getting help, getting data into R, creating variables, creating data frames and tibbles, reading in data files, manipulating variables within a tibble/data frame.

- **adventr_03**: Summarizing data (introducing ggplot2)

- **adventr_04**: Fitting models (central tendency)

- **adventr_05**: Presenting data (summarizing groups and more ggplot2)

- **adventr_08**: Inferential statistics and robust estimation (covers Chapter 9 too)

- **adventr_11** Hypothesis testing

- **adventr_14**: The general linear model

- **adventr_15**: Comparing two means

- **adventr_15_rm**: Comparing two means (repeated measures)

- **adventr_16**: Comparing several means

- **adventr_16_rm**: Comparing several means (repeated measures)

- **adventr_17**: Factorial designs

- **adventr_mlm**: Multilevel models (not covered in the book)

- **adventr_growth**: Growth models (not covered in the book)

- **adventr_log**: Logistic regression (not covered in the book)

References

- Field, A. P. (2016). *An adventure in statistics: the reality enigma*. London: Sage.

- Field, A. P., Miles, J. N. V., & Field, Z. C. (2012). *Discovering statistics using R: And sex and drugs and rock 'n' roll*. London: Sage.

---

alice_dat	*Alice's gene data*

Description

A dataset relating to the following fictitious experiment: Participants who had facial burns were recruited. One-week pre-test, a gene C-10XFMG (the so-called chameleon gene) was introduced into all participants.Two days prior to test, a genetic toggle switch was introduced into half of the participants. Half of the participants were asked to look at a photograph of themselves from before their injury and to imagine the cells in their faces changing to become like the photo. The remainder acted as a control and were asked to look at a picture of a same-sex stranger and to try to change their face to become the person in the picture. All participants looked at the picture for 6 sessions of 20 minutes each. At the end of the sessions their faces were scanned into a computer and compared to the face in the photograph. Facial recognition software produced a precise resemblance measure as a percentage (100 in the photograph, 0 The data are taken from Chapter 17 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

    alice_dat

Format

A tibble with 32 rows and 4 variables:

id Participant ID

picture A factor describing whether a participant tried to resemble a picture of themselves (Self), or a matched stranger (Other)

gene A factor describing whether a participant received the toggle switch (C-gene + Toggle) or not (C-gene)

resemblance How closely their face resembled the picture (100% = the participants face is exactly like the face in the photograph, 0% the person in the photo bears no resemblance at all to the participant)

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

---

calcite_dat	*Calcite mask data*

Description

A dataset relating to the following fictitious experiment, which is an extension of the invisibility paste data described elsewhere: "... we asked whether we could manipulate the face perceived by the participant when viewing someone wearing calcite invisibility paste. Participants were all so-called Chippers, who had ID chips installed in their brains. Each participant met 40 people who had calcite paste on their face. For each of these people we took a photo of their face (the 'actual' face), and found a photo of a different face matched for sex and age (the 'different' face). We could send these images to the participant's visual cortex using their ID chip. For 20 of the encounters an image of the person's 'actual' face was transmitted to the participant's ID chip during the encounter, for the remaining 20 encounters the 'different' face was sent to the chip. During each encounter the participant was shown the photos of the actual and different face and pointed to the one depicting whom they thought they had just met. As in Experiment 1, participants scored a point for each person they correctly identified: a score of 0 would mean that they never chose the correct photo, a score of 20 would mean that they always chose the correct photo." The data are taken from Chapter 15 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
calcite_dat
```

Format

A tibble with 20 rows and 3 variables:

id Participant ID

actual How many of the 20 people the participant correctly identified when a picture of the 'actual' person was transmitted to the participant's ID chip

different How many of the 20 people the participant correctly identified when a picture of a 'different' person was transmitted to the participant's ID chip

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

garlic_dat   *Zombie garlic data*

Description

A dataset relating to whether zombies can be repelled by garlic. It is taken from Chapter 14 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
garlic_dat
```

Format

A tibble with 10 rows and 3 variables:

id Victim ID

latency The time taken for a zombie to approach the victim in seconds

garlic The number of bulbs of garlic worn by the victim

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

ha_dat   *Subset of Ha et al. (2010) data*

Description

A dataset containing a random sample of 20 female participants from Ha, T., Overbeek, G., & Engels, R. C. M. E. (2010). Effects of Attractiveness and Social Status on Dating Desire in Heterosexual Adolescents: An Experimental Study. Archives of Sexual Behavior, 39(5), 1063-1071. doi: 10.1007/s10508-009-9561-z. It is usedin Chapter 3 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
ha_dat
```

Format

A tibble with 20 rows and 13 variables:

id Participant ID

age Age of the participant, years

Sex Biological sex of the individual

hi_salary Rating (out of 10) of the importance of a high salary in a potential romantic partner

fin_ed Rating (out of 10) of the importance of a potential romantic partner having finished their education

kind Rating (out of 10) of the importance of a potential romantic partner being kind

humour Rating (out of 10) of the importance of a potential romantic partner having a sense of humour

ambitious Rating (out of 10) of the importance of a potential romantic partner being ambitious

wants_child Rating (out of 10) of the importance of a potential romantic partner wanting children

romantic Rating (out of 10) of the importance of a potential romantic partner being romantic

attractive Rating (out of 10) of the importance of a potential romantic partner being physically attractive

creativity Rating (out of 10) of the importance of a potential romantic partner being creative

honest Rating (out of 10) of the importance of a potential romantic partner being honest

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

implant_dat   *Memory implanting data*

Description

A dataset relating to the following fictitious experiment: "We wanted to see whether our failure to replace memories was due to the type of conflicting information used. Five participants engaged in three different scripted encounters with strangers during which 10 unusual things happened. After each encounter participants were asked to privately recall what had happened. During the period of recall for one encounter they were left alone (control), during the recall of another their ID chip was used to implant a verbal description of the encounter in which the 10 unusual events were replaced with different events, and during the recall of the final encounter their ID chip was used to implant visual images of 10 different things to what they actually experienced. Each participant, therefore, experienced three encounters and a different type of recall manipulation followed each one; the type of manipulation associated with a particular encounter was counterbalanced. Finally, participants recalled the event for a second time and we measured how many of the 10 unusual events that they originally experienced that they could recall." The data are taken from Chapter 16 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
implant_dat
```

Format

A tibble with 5 rows and 4 variables:

id Participant ID

no_interference How many of the 10 unusual events were remembered when there was no interference

verbal How many of the 10 unusual events were remembered when verbal information was used as interference

visual How many of the 10 unusual events were remembered when visual images were used as interference

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

jig_dat   *JIG:SAW employee data.*

Description

A dataset containing information about the characteristics of different groups of employees within the corporation called JIG:SAW and non-employees. It is taken from Chapter 5 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
jig_dat
```

Format

A tibble with 240 rows and 7 variables:

id employee ID

employee whether or not the employee works for JIG:SAW

job_type Categories of employee

footspeed Footspeed of the individual, miles per hour

strength Maximal push force of the individual, Newtons

vision Visual accuity

sex Biological sex of the individual

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

mask_dat	*Invisibility paste data*

Description

A dataset relating to the following fictitious experiment: "Twenty-eight participants were tested individually. Each participant sat on a revolving chair in the middle of a circle of 20 people. The 20 people all wore a light deflecting aid on their face. For half of the participants (N = 14) the 20 people wore a standard calcite-based mask, whereas the other half (N = 14) saw people wearing 'invisibility paste', a transparent cream containing calcite micro-crystals. Participants looked at each person and were given two photographs, one depicting the person they were looking at and one depicting a different person matched for age and sex. Participants had to decide which person they were looking at, before turning to the next person and being given two different photographs. Participants scored a point for each person they identified correctly: a score of 0 would mean they always chose the incorrect photo, a score of 20 would mean that they always chose the correct photo." The data are taken from Chapter 15 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
mask_dat
```

Format

A tibble with 28 rows and 3 variables:

id Participant ID

mask Whether the person was assigned to the calcite mask condition or the invisibility paste condition

recognition How many of the 20 people they correctly identified

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

memory_dat	*Memory erasing data*

Description

A dataset relating to the following fictitious experiment: "Thirty-six participants with ID chips were tested individually. All participants met a stranger who had a 5-minute scripted conversation with them containing 10 critical pieces of information. A week later, the participants were asked to recall the encounter for 5 minutes, and then after a 10-minute break wrote everything that they could remember about the original encounter. During the 5-minute initial recall, one of three things happened: the control group (N = 12) received no intervention; the erase group (N = 12) had a pulse of electricity sent to their brain via their ID chip; and the replace group (N = 12) had conflicting verbal descriptions sent to their ID chip. The outcome was how many of the 10 critical pieces of information the participants wrote down after the recall phase." The data are taken from Chapter 16 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
memory_dat
```

Format

A tibble with 36 rows and 7 variables:

id Participant ID

group To which of the no intervention, memory erase or memory replace conditions the participant was assigned

recall How many of the 10 critical pieces of information the participant identified

erase_dummy Dummy variable that compares the erase condition to no intervention

replace_dummy Dummy variable that compares the replace condition to no intervention

contrast_1 Variable that uses contrast codes to compare the memory conditions (combined) to no intervention

contrast_2 Variable that uses contrast codes to compare the erase condition to the replace condition

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

mem_cov_dat   *Memory erasing data with covariate*

Description

A dataset relating to the following fictitious experiment from Chapter 16 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage: "Thirty-six participants with ID chips were tested individually. All participants met a stranger who had a 5-minute scripted conversation with them containing 10 critical pieces of information. A week later, the participants were asked to recall the encounter for 5 minutes, and then after a 10-minute break wrote everything that they could remember about the original encounter. During the 5-minute initial recall, one of three things

happened: the control group (N = 12) received no intervention; the erase group (N = 12) had a pulse of electricity sent to their brain via their ID chip; and the replace group (N = 12) had conflicting verbal descriptions sent to their ID chip. The outcome was how many of the 10 critical pieces of information the participants wrote down after the recall phase."

Usage

```
mem_cov_dat
```

Format

A tibble with 36 rows and 7 variables:

id Participant ID

recall How many of the 10 critical pieces of information the participant identified

group To which of the no intervention, memory erase or memory replace conditions the participant was assigned

working_mem Variable that measures working memory capacity as the individual's digit span

Details

This version of the data includes an additional variable that measured working memory in each participant (as their digit span).

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

ras_dat *Alice's RAS scores*

Description

A dataset containing Alice's ratings on the Relationship Assessment Scale (RAS) over 10 weeks. It is taken from Chapter 4 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
ras_dat
```

Format

A tibble with 10 rows and 2 variables:

week The week in which the rating was given

rating Alice's score on the relaationship assessment scale

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

---

recovery_dat *Zombie recovery data*

Description

A dataset not from Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage, but continuing the theme of the story. At the end of the book it is revealed that Alice used her C-gene therapy to restore the code 1318 workers to a human state. This dataset relates to her third attempt. It contains data from 400 code 1318 workers. Workers were randomly assigned to two arms of a trial (C-gene therapy and C-gene therapy with a toggle switch). They were also given one of five different doses of C-gene: 0 (no dose) or 1, 2, 3, or 4 doses. The outcome was whether (or not) the worker had fully recovered after the treatment.

Usage

recovery_dat

Format

A tibble with 400 rows and 4 variables:

id Participant ID

treatment character vector that codes which arm of the trial the participant was randomized to (no toggle switch or toggle switch)

dose The dose of C-gene used during treatment (0 doses, 1, 2, 3, or 4 doses)

recovered Character vector that expresses whether or not the participant had recovered following treatment (Recovered or Not Recovered)

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

---

rehab_dat   *Zombie rehabilitation data*

Description

A dataset not from Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage, but continuing the theme of the story. At the end of the book it is revealed that Alice used her C-gene therapy to restore the code 1318 workers to a human state. This dataset relates to her first attempt. It contains data from 190 code 1318 workers treated at 10 different clinics. Workers were randomly assigned to two arms of the trial (wait list vs.C-gene therapy) and the outcome was how much they resembled their pre-zombie state (as a percentage). The zombification occurred as a result of these workers entering JIG:SAWs genetic enhancement programme (GEP), and additional variables quantify how long it has been since they entered this programme (i.e., how long have they been a zombie) and whether they were originally exposed to the low- or high-intensity GEP.

Usage

```
rehab_dat
```

Format

A tibble with 190 rows and 6 variables:

id Participant ID

id_clin Factor that codes which of 10 clinics the participant attended

intervention Factor that codes which arm of the trial the participant was randomized to (0 = wait list, 1 = gene therapy)

resemblance How closely their face resembled their pre-zombified state (100% = the participants face is exactly like their original face, 0% the person bears no resemblance to their pre-zombified face)

gep_level Factor that codes whether the participant was originally exposed to the low- (0) or high-intensity (1) genetic enhancement program at JIG\:SAW

tse_months How many months since the participant completed the genetic enhancement program that resulted in their zombification

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

rehab_growth_dat   *Zombie rehabilitation longitudinal data*

Description

A dataset not from Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage, but continuing the theme of the story. At the end of the book it is revealed that Alice used her C-gene therapy to restore the code 1318 workers to a human state. This dataset relates to her second attempt. It contains data from 141 code 1318 workers measured at four timepoints (baseline and 1, 6, and 12 month follow-up). Workers were randomly assigned to two arms of the trial (wait list vs. C-gene therapy) and the outcome was how much they resembled their pre-zombie state (as a percentage).

Usage

```
rehab_growth_dat
```

Format

A tibble with 564 rows and 5 variables:

id Participant ID

intervention character vector that codes which arm of the trial the participant was randomized to (wait list or gene therapy)

resemblance How closely their face resembled their pre-zombified state (100% = the participants face is exactly like their original face, 0% the person bears no resemblance to their pre-zombified face)

time Character vector that expresses when resemblance was measured as "t0" (baseline), "t1" (1 month follow-up), "t6" (6-month follow up) and "t12" (12-month follow-up)

time_num integer vector exptressing time in months from baseline (0, 1, 6, 12) as a number

Source

https://www.discoveringstatistics.com/books/an-adventure-in-statistics/

---

taser_dat   *Zombie immobility data*

Description

A dataset relating to the degree to which zombies can be rendered immobile by repetitive transcranial magnetic stimulation (rTMS) and zapping with a taser. It is taken from Chapter 14 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
taser_dat
```

Format

A tibble with 156 rows and 4 variables:

id Victim ID

immobility The duration of the zombie's immobility, seconds

r_tms The frequency of rTMS, Hz

taser The output of the taser, kV

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

---

teddy_dat *Teddy therapy data.*

Description

A dataset based on two ficticious studies looking at whether cuddling a teddy bear (compared to the cardboard box that the teddy was packaged in) affects self-reported self-esteem. It is taken from Chapter 10 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
teddy_dat
```

Format

A tibble with 220 rows and 4 variables:

id Participant ID

group whether the participant cuddled a teddy bear or the cardboard box that contained the teddy bear

study_n Factor that distinguishes the two studies in the data set. The first was based on a total N of 20 and the second was based on a total N of 200

self_esteem Self-reported self-esteem scores

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

| zhang_female_dat | *Subset of Zhang data (female sample N = 20)* |

Description

A dataset extracted from Ha, T., Overbeek, G., & Engels, R. C. M. E. (2010). Effects of Attractiveness and Social Status on Dating Desire in Heterosexual Adolescents: An Experimental Study. Archives of Sexual Behavior, 39(5), 1063-1071. doi:10.1007/s10508-009-9561-z. The study looked at accuracy on a maths test when performed under the participant's own name or a fake name. It is taken from Chapter 15 of Field, A. P. (2016). An adventure in statistics: the reality enigma. London: Sage.

Usage

```
zhang_female_dat
```

Format

A tibble with 20 rows and 2 variables:

name Specifies whether participants completed the test under their own name or a fake name

accuracy Accuracy on a maths test, percent

Source

<https://www.discoveringstatistics.com/books/an-adventure-in-statistics/>

Index