Package 'psychTools'

May 29, 2019

Version 1.9.5.26

Date 2019-05-26

Title Tools to Accompany the 'psych' Package for Psychological
Research

Description Support functions, data sets, and vignettes for the 'psych' package. Contains several of the biggest data sets for the 'psych' package as well as one vignette. A few helper functions for file manipulation are included as well. For more information, see the <https://personality-project.org/r> web page.

License GPL (>= 2)

Imports foreign,psych

Suggests parallel, GPArotation, lavaan

LazyData true

ByteCompile TRUE

URL <https://personality-project.org/r/psych>

   <https://personality-project.org/r/psych-manual.pdf>

NeedsCompilation no

Author William Revelle [aut, cre] (<https://orcid.org/0000-0003-4880-9610>)

Maintainer William Revelle <revelle@northwestern.edu>

Depends R (>= 2.10)

Repository CRAN

Date/Publication 2019-05-29 12:40:03 UTC

R topics documented:

---

ability   *16 ability items scored as correct or incorrect.*

Description

16 multiple choice ability items 1525 subjects taken from the Synthetic Aperture Personality Assessment (SAPA) web based personality assessment project are saved as `iqitems`. Those data are shown as examples of how to score multiple choice tests and analyses of response alternatives. When scored correct or incorrect, the data are useful for demonstrations of tetrachoric based factor analysis `irt.fa` and finding tetrachoric correlations.

Usage

```
data(iqitems)
```

Format

A data frame with 1525 observations on the following 16 variables. The number following the name is the item number from SAPA.

reason.4 Basic reasoning questions

reason.16 Basic reasoning question

reason.17 Basic reasoning question

reason.19 Basic reasoning question

letter.7 In the following alphanumeric series, what letter comes next?

letter.33 In the following alphanumeric series, what letter comes next?

letter.34 In the following alphanumeric series, what letter comes next

letter.58 In the following alphanumeric series, what letter comes next?

matrix.45 A matrix reasoning task

matrix.46 A matrix reasoning task

matrix.47 A matrix reasoning task

matrix.55 A matrix reasoning task

rotate.3 Spatial Rotation of type 1.2

rotate.4 Spatial Rotation of type 1.2

rotate.6 Spatial Rotation of type 1.1

rotate.8 Spatial Rotation of type 2.3

Details

16 items were sampled from 80 items given as part of the SAPA (<https://sapa-project.org>) project (Revelle, Wilt and Rosenthal, 2009; Condon and Revelle, 2014) to develop online measures of ability. These 16 items reflect four lower order factors (verbal reasoning, letter series, matrix reasoning, and spatial rotations. These lower level factors all share a higher level factor ('g').

This data set may be used to demonstrate item response functions, `tetrachoric` correlations, or `irt.fa` as well as `omega` estimates of of reliability and hierarchical structure.

In addition, the data set is a good example of doing item analysis to examine the empirical response probabilities of each item alternative as a function of the underlying latent trait. When doing this, it appears that two of the matrix reasoning problems do not have monotonically increasing trace lines for the probability correct. At moderately high ability (theta = 1) there is a decrease in the probability correct from theta = 0 and theta = 2.

Source

The example data set is taken from the Synthetic Aperture Personality Assessment personality and ability test at <https://sapa-project.org>. The data were collected with David Condon from 8/08/12 to 8/31/12.

Similar data are available from the International Cognitive Ability Resource at <https://icar-project.org>.

References

Revelle, William, Wilt, Joshua, and Rosenthal, Allen (2010) Personality and Cognition: The Personality-Cognition Link. In Gruszka, Alexandra and Matthews, Gerald and Szymura, Blazej (Eds.) Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control, Springer.

Condon, David and Revelle, William, (2014) The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. Intelligence, 43, 52-64.

Examples

```
data(ability)
cs<- psych::cs
keys <- list(ICAR16=colnames(ability),reasoning = cs(reason.4,reason.16,reason.17,reason.19),
  letters=cs(letter.7, letter.33,letter.34,letter.58, letter.7),
  matrix=cs(matrix.45,matrix.46,matrix.47,matrix.55),
  rotate=cs(rotate.3,rotate.4,rotate.6,rotate.8))
psych::scoreOverlap(keys,ability)
  #this next step takes a few seconds
    ability.irt <- psych::irt.fa(ability)
    ability.scores <- psych::score.irt(ability.irt,ability)
```

---

affect | *Two data sets of affect and arousal scores as a function of personality and movie conditions*

Description

A recurring question in the study of affect is the proper dimensionality and the relationship to various personality dimensions. Here is a data set taken from two studies of mood and arousal using movies to induce affective states.

Usage

```
data(affect)
```

Details

These are data from two studies conducted in the Personality, Motivation and Cognition Laboratory at Northwestern University. Both studies used a similar methodology:

Collection of pretest data using 5 scales from the Eysenck Personality Inventory and items taken from the Motivational State Questionnaire (see msq. In addition, state and trait anxiety measures were given. In the "maps" study, the Beck Depression Inventory was given also.

Then subjects were randomly assigned to one of four movie conditions: 1: Frontline. A documentary about the liberation of the Bergen-Belsen concentration camp. 2: Halloween. A horror film. 3: National Geographic, a nature film about the Serengeti plain. 4: Parenthood. A comedy. Each film clip was shown for 9 minutes. Following this the MSQ was given again.

Data from the MSQ were scored for Energetic and Tense Arousal (EA and TA) as well as Positive and Negative Affect (PA and NA).

Study flat had 170 participants, study maps had 160.

These studies are described in more detail in various publications from the PMC lab. In particular, Revelle and Anderson, 1997 and Rafaeli and Revelle (2006). An analysis of these data has also appeared in Smillie et al. (2012).

For a much more complete data set involving film, caffeine, and time of day manipulations, see the msqR data set.

Source

Data collected at the Personality, Motivation, and Cognition Laboratory, Northwestern University.

References

Revelle, William and Anderson, Kristen Joan (1997) Personality, motivation and cognitive performance: Final report to the Army Research Institute on contract MDA 903-93-K-0008

Rafaeli, Eshkol and Revelle, William (2006), A premature consensus: Are happiness and sadness truly opposite affects? Motivation and Emotion, 30, 1, 1-12.

Smillie, Luke D. and Cooper, Andrew and Wilt, Joshua and Revelle, William (2012) Do Extraverts Get More Bang for the Buck? Refining the Affective-Reactivity Hypothesis of Extraversion. Journal of Personality and Social Psychology, 103 (2), 206-326.

Examples

```
data(affect)
psych::describeBy(affect[-1],group="Film")
psych::pairs.panels(affect[14:17],bg=c("red","black","white","blue")[affect$Film],pch=21,
    main="Affect varies by movies ")
psych::errorCircles("EA2","TA2",data=affect,group="Film",labels=c("Sad","Fear","Neutral","Humor")
, main="Enegetic and Tense Arousal by Movie condition")
psych::errorCircles(x="PA2",y="NA2",data=affect,group="Film",labels=c("Sad","Fear","Neutral","
Humor"),  main="Positive and Negative Affect by Movie condition")
```

---

bfi   *25 Personality items representing 5 factors*

Description

25 personality self report items taken from the International Personality Item Pool (ipip.ori.org) were included as part of the Synthetic Aperture Personality Assessment (SAPA) web based personality assessment project. The data from 2800 subjects are included here as a demonstration set for scale construction, factor analysis, and Item Response Theory analysis. Three additional demographic variables (sex, education, and age) are also included.

Usage

```
data(bfi)
data(bfi.dictionary)
```

Format

A data frame with 2800 observations on the following 28 variables. (The q numbers are the SAPA item numbers).

A1 Am indifferent to the feelings of others. (q_146)

A2 Inquire about others' well-being. (q_1162)

A3 Know how to comfort others. (q_1206)

A4 Love children. (q_1364)

A5 Make people feel at ease. (q_1419)

C1 Am exacting in my work. (q_124)

C2 Continue until everything is perfect. (q_530)

C3 Do things according to a plan. (q_619)

C4 Do things in a half-way manner. (q_626)

C5 Waste my time. (q_1949)

E1 Don't talk a lot. (q_712)

E2 Find it difficult to approach others. (q_901)

E3 Know how to captivate people. (q_1205)

E4 Make friends easily. (q_1410)

E5 Take charge. (q_1768)

N1 Get angry easily. (q_952)

N2 Get irritated easily. (q_974)

N3 Have frequent mood swings. (q_1099)

N4 Often feel blue. (q_1479)

N5 Panic easily. (q_1505)

O1 Am full of ideas. (q_128)

O2 Avoid difficult reading material.(q_316)

O3 Carry the conversation to a higher level. (q_492)

O4 Spend time reflecting on things. (q_1738)

O5 Will not probe deeply into a subject. (q_1964)

gender Males = 1, Females =2

education 1 = HS, 2 = finished HS, 3 = some college, 4 = college graduate 5 = graduate degree

age age in years

Details

The first 25 items are organized by five putative factors: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Opennness. The scoring key is created using make.keys, the scores are found using score.items.

These five factors are a useful example of using irt.fa to do Item Response Theory based latent factor analysis of the polychoric correlation matrix. The endorsement plots for each item, as well as the item information functions reveal that the items differ in their quality.

The item data were collected using a 6 point response scale: 1 Very Inaccurate 2 Moderately Inaccurate 3 Slightly Inaccurate 4 Slightly Accurate 5 Moderately Accurate 6 Very Accurate

as part of the Synthetic Aperture Personality Assessment (SAPA https://sapa-project.org) project. To see an example of the data collection technique, visit https://SAPA-project.org or the International Cognitive Ability Resource at https://icar-project.org. The items given were sampled from the International Personality Item Pool of Lewis Goldberg using the sampling technique of SAPA. This is a sample data set taken from the much larger SAPA data bank.

Note

The bfi data set and items should not be confused with the BFI (Big Five Inventory) of Oliver John and colleagues (John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory– Versions 4a and 54. Berkeley, CA: University of California,Berkeley, Institute of Personality and Social Research.)

Source

The items are from the ipip (Goldberg, 1999). The data are from the SAPA project (Revelle, Wilt and Rosenthal, 2010) , collected Spring, 2010 (https://sapa-project.org).

References

Goldberg, L.R. (1999) A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In Mervielde, I. and Deary, I. and De Fruyt, F. and Ostendorf, F. (eds) Personality psychology in Europe. 7. Tilburg University Press. Tilburg, The Netherlands.

Revelle, W., Wilt, J., and Rosenthal, A. (2010) Individual Differences in Cognition: New Methods for examining the Personality-Cognition Link In Gruszka, A. and Matthews, G. and Szymura, B. (Eds.) Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control, Springer.

Revelle, W, Condon, D.M., Wilt, J., French, J.A., Brown, A., and Elleman, L.G. (2016) Web and phone based data collection using planned missing designs. In Fielding, N.G., Lee, R.M. and Blank, G. (Eds). SAGE Handbook of Online Research Methods (2nd Ed), Sage Publcations.

See Also

bi.bars to show the data by age and gender, irt.fa for item factor analysis applying the irt model.

Examples

```
data(bfi)
psych::describe(bfi)
# create the bfi.keys
 bfi.keys <-
  list(agree=c("-A1","A2","A3","A4","A5"),conscientious=c("C1","C2","C3","-C4","-C5"),
extraversion=c("-E1","-E2","E3","E4","E5"),neuroticism=c("N1","N2","N3","N4","N5"),
openness = c("O1","-O2","O3","O4","-O5"))

 scores <- psych::scoreItems(bfi.keys,bfi,min=1,max=6) #specify the minimum and maximum values
```

```
scores
#show the use of the fa.lookup with a dictionary
psych::keys.lookup(bfi.keys,bfi.dictionary[,1:4])
```

---

blant   *A 29 x 29 matrix that produces weird factor analytic results*

Description

Normally, min.res factor analysis and maximum likelihood produce very similar results. This data set (from Alexandra Blant) does not. Warnings are given for the min.res solution, the pa solution, but not the old.min nor the mle solution. Included as a test case for the factor analysis function.

Usage

```
data("blant")
```

Format

The format is: num [1:29, 1:29] 1 0.77 0.813 0.68 0.717 ... - attr(*, "dimnames")=List of 2 ..$: NULL ..$: chr [1:29] "V1" "V2" "V3" "V4" ...

Details

This data matrix was sent by Alexandra Blant as an example of a problem with the minres solution in the `fa` function. The default solution, using fm="minres" issues a warning that the solution has improper factor score weights. This is not the case for the fm="old.min" and fm="mle" options, but is for fm="pa", fm="ols".

The residuals are indeed smaller for fm="minres" than for fm="old.min" or fm="mle".

"old.min" attempts to find the minimum residual but uses the gradient for mle. This was the approach until version 1.7.5 but was changed (see the help page for fa) following extensive communication with Hao Wu.

The problem with this matrix is probably that it is almost singular, with some smcs approaching 1 and the smallest three eigenvalues of .006, .004 and .001.

This problem matrix was provided by Alexandra Blant.

Source

Alexandra Blant, personal communication

Examples

```
data(blant)
#compare
f5 <- psych::fa(blant,5,rotate="none")  #the default minres
f5.old <- psych::fa(blant,5, fm="old.min",rotate="none") #old version of minres
f5.mle <- psych::fa(blant,5,fm="mle",rotate= "none") #maximum likelihood
#compare solutions
psych::factor.congruence(list(f5,f5.old,f5.mle))
#compare sums of squared residuals
sum(residuals(f5,diag=FALSE)^2,na.rm=TRUE) # 1.355489
sum(residuals(f5.old,diag=FALSE)^2,na.rm=TRUE) # 1.539757
sum(residuals(f5.mle,diag=FALSE)^2,na.rm=TRUE) # 2.402092

#but, when we divide the squared residuals by the original (squared) correlations, we find
#a different ordering of fit
f5$fit     # 0.9748177
f5.old$fit  # 0.9752774
f5.mle$fit   # 0.9603324
```

---

blot   *Bond's Logical Operations Test – BLOT*

Description

35 items for 150 subjects from Bond's Logical Operations Test. A good example of Item Response Theory analysis using the Rasch model. One parameter (Rasch) analysis and two parameter IRT analyses produce somewhat different results.

Usage

```
data(blot)
```

Format

A data frame with 150 observations on 35 variables. The BLOT was developed as a paper and pencil test for children to measure Logical Thinking as discussed by Piaget and Inhelder.

Details

Bond and Fox apply Rasch modeling to a variety of data sets. This one, Bond's Logical Operations Test, is used as an example of Rasch modeling for dichotomous items. In their text (p 56), Bond and Fox report the results using WINSTEPS. Those results are consistent (up to a scaling parameter) with those found by the rasch function in the ltm package. The WINSTEPS seem to produce difficulty estimates with a mean item difficulty of 0, whereas rasch from ltm has a mean difficulty of -1.52. In addition, rasch seems to reverse the signs of the difficulty estimates when reporting the coefficients and is effectively reporting "easiness".

However, when using a two parameter model, one of the items (V12) behaves very differently.

This data set is useful when comparing 1PL, 2PL and 2PN IRT models.

Source

The data are taken (with kind permission from Trevor Bond) from the webpage https://www.winsteps.com/BF3/bondfox3.htm and read using read.fwf.

References

T.G. Bond. BLOT:Bond's Logical Operations Test. Townsville, Australia: James Cook University. (Original work published 1976), 1995.

T. Bond and C. Fox. (2007) Applying the Rasch model: Fundamental measurement in the human sciences. Lawrence Erlbaum, Mahwah, NJ, US, 2 edition.

See Also

See also the `irt.fa` and associated plot functions.

Examples

```
data(blot)

 #ltm is not required by psychTools, but if available, may be run to show a Rasch model

#do the same thing with functions in psych
blot.fa <- psych::irt.fa(blot)  # a 2PN model
plot(blot.fa)
```

---

burt   *11 emotional variables from Burt (1915)*

Description

Cyril Burt reported an early factor analysis with a circumplex structure of 11 emotional variables in 1915. 8 of these were subsequently used by Harman in his text on factor analysis. Unfortunately, it seems as if Burt made a mistake for the matrix is not positive definite. With one change from .87 to .81 the matrix is positive definite.

Usage

```
data(burt)
```

Format

A correlation matrix based upon 172 "normal school age children aged 9-12".

Sociality Sociality

Sorrow Sorrow

Tenderness Tenderness

Joy Joy

Wonder Wonder

Elation Elation

Disgust Disgust

Anger Anger

Sex Sex

Fear Fear

Subjection Subjection

Details

The Burt data set is interesting for several reasons. It seems to be an early example of the organizaton of emotions into an affective circumplex, a subset of it has been used for factor analysis examples (see Harman.Burt, and it is an example of how typos affect data. The original data matrix has one negative eigenvalue. With the replacement of the correlation between Sorrow and Tenderness from .87 to .81, the matrix is positive definite.

Alternatively, using cor.smooth, the matrix can be made positive definite as well, although cor.smooth makes more (but smaller) changes.

Source

(retrieved from the web at https://www.biodiversitylibrary.org/item/95822#790) Following a suggestion by Jan DeLeeuw.

References

Burt, C.General and Specific Factors underlying the Primary Emotions. Reports of the British Association for the Advancement of Science, 85th meeting, held in Manchester, September 7-11, 1915. London, John Murray, 1916, p. 694-696 (retrieved from the web at https://www.biodiversitylibrary.org/item/95822#790)

See Also

Harman.Burt in the Harman dataset and cor.smooth

Examples

```
data(burt)
eigen(burt)$values  #one is negative!
burt.new <- burt
burt.new[2,3] <- burt.new[3,2] <- .81
eigen(burt.new)$values  #all are positive
bs <- psych::cor.smooth(burt)
round(burt.new - bs,3)
```

---

cities   *Distances between 11 US cities*

Description

Airline distances between 11 US cities may be used as an example for multidimensional scaling or cluster analysis.

Usage

```
data(cities)
```

Format

A data frame with 11 observations on the following 11 variables.

ATL Atlana, Georgia

BOS Boston, Massachusetts

ORD Chicago, Illinois

DCA Washington, District of Columbia

DEN Denver, Colorado

LAX Los Angeles, California

MIA Miami, Florida

JFK New York, New York

SEA Seattle, Washington

SFO San Francisco, California

MSY New Orleans, Lousianna

Details

An 11 x11 matrix of distances between major US airports. This is a useful demonstration of multiple dimensional scaling.

city.location is a dataframe of longitude and latitude for those cities.

Note that the 2 dimensional MDS solution does not perfectly capture the data from these city distances. Boston, New York and Washington, D.C. are located slightly too far west, and Seattle and LA are slightly too far south.

Source

https://www.timeanddate.com/worldclock/distance.html

Examples

```
data(cities)
city.location[,1] <- -city.location[,1]

#an overlay map can be added if the package maps is available

libary(maps)
map("usa")
title("MultiDimensional Scaling of US cities")
points(city.location)

plot(city.location, xlab="Dimension 1", ylab="Dimension 2",
   main ="Multidimensional scaling of US cities")
city.loc <- cmdscale(cities, k=2) #ask for a 2 dimensional solution  round(city.loc,0)
city.loc <- -city.loc
 city.loc <- psych::rescale(city.loc,apply(city.location,2,mean),apply(city.location,2,sd))
points(city.loc,type="n")
text(city.loc,labels=names(cities))
```

---

cubits	*Galton's example of the relationship between height and 'cubit' or forearm length*

Description

Francis Galton introduced the 'co-relation' in 1888 with a paper discussing how to measure the relationship between two variables. His primary example was the relationship between height and forearm length. The data table (cubits) is taken from Galton (1888). Unfortunately, there seem to be some errors in the original data table in that the marginal totals do not match the table.

The data frame, heights, is converted from this table.

Usage

```
data(cubits)
```

Format

A data frame with 9 observations on the following 8 variables.

16.5 Cubit length < 16.5

16.75 16.5 <= Cubit length < 17.0

17.25 17.0 <= Cubit length < 17.5

17.75 17.5 <= Cubit length < 18.0

18.25 18.0 <= Cubit length < 18.5

18.75 18.5 <= Cubit length < 19.0

`19.25` 19.0 <= Cubit length < 19.5

`19.75` 19.5 <= Cubit length

Details

Sir Francis Galton (1888) published the first demonstration of the correlation coefficient. The regression (or reversion to mediocrity) of the height to the length of the left forearm (a cubit) was found to .8. There seem to be some errors in the table as published in that the row sums do not agree with the actual row sums. These data are used to create a matrix using `table2matrix` for demonstrations of analysis and displays of the data.

Source

Galton (1888)

References

Galton, Francis (1888) Co-relations and their measurement. Proceedings of the Royal Society. London Series,45,135-145,

See Also

`table2matrix`, `table2df`, `ellipses`, `heights`, `peas`,`galton`

Examples

```
data(cubits)
cubits
heights <- psych::table2df(cubits,labs = c("height","cubit"))
psych::ellipses(heights,n=1,main="Galton's co-relation data set")
psych::ellipses(jitter(heights$height,3),jitter(heights$cubit,3),pch=".",
     main="Galton's co-relation data set",xlab="height",
     ylab="Forearm (cubit)") #add in some noise to see the points
psych::pairs.panels(heights,jiggle=TRUE,main="Galton's cubits data set")
```

---

cushny | *A data set from Cushny and Peebles (1905) on the effect of three drugs on hours of sleep, used by Student (1908)*

Description

The classic data set used by Gossett (publishing as Student) for the introduction of the t-test. The design was a within subjects study with hours of sleep in a control condition compared to those in 3 drug conditions. Drug1 was 06mg of L Hscyamine, Drug 2L and Drug2R were said to be .6 mg of Left and Right isomers of Hyoscine. As discussed by Zabell (2008) these were not optical isomers. The detal1, delta2L and delta2R are changes from the baseline control.

Usage

```
data(cushny)
```

Format

A data frame with 10 observations on the following 7 variables.

Control Hours of sleep in a control condition

drug1 Hours of sleep in Drug condition 1

drug2L Hours of sleep in Drug condition 2

drug2R Hours of sleep in Drug condition 3 (an isomer of the drug in condition 2

delta1 Change from control, drug 1

delta2L Change from control, drug 2L

delta2R Change from control, drug 2R

Details

The original analysis by Student is used as an example for the t-test function, both as a paired t-test and a two group t-test. The data are also useful for a repeated measures analysis of variance.

Source

Cushny, A.R. and Peebles, A.R. (1905) The action of optical isomers: II hyoscines. The Journal of Physiology 32, 501-510.

Student (1908) The probable error of the mean. Biometrika, 6 (1) , 1-25.

References

See also the data set sleep and the examples for the t.test

S. L. Zabell. On Student's 1908 Article "The Probable Error of a Mean" Journal of the American Statistical Association, Vol. 103, No. 481 (Mar., 2008), pp. 1- 20

Examples

```
data(cushny)
with(cushny, t.test(drug1,drug2L,paired=TRUE)) #within subjects

psych::error.bars(cushny[1:4],within=TRUE,ylab="Hours of sleep",xlab="Drug condition",
        main="95% confidence of within subject effects")
```

---

df2latex   *Convert a data frame, correlation matrix, or factor analysis output to*
                            a LaTeX table

Description

A set of handy helper functions to convert data frames or matrices to LaTeX tables. Although
Sweave is the preferred means of converting R output to LaTeX, it is sometimes useful to go directly
from a data.frame or matrix to a LaTeX table. cor2latex will find the correlations and then create
a lower (or upper) triangular matrix for latex output. fa2latex will create the latex commands for
showing the loadings and factor intercorrelations. As the default option, tables are prepared in an
approximation of APA format.

Usage

```
df2latex(x,digits=2,rowlabels=TRUE,apa=TRUE,short.names=TRUE,font.size ="scriptsize",
      big.mark=NULL,drop.na=TRUE, heading="A table from the psych package in R",
   caption="df2latex",label="default", char=FALSE,
    stars=FALSE,silent=FALSE,file=NULL,append=FALSE,cut=0,big=0)
cor2latex(x,use = "pairwise", method="pearson", adjust="holm",stars=FALSE,
      digits=2,rowlabels=TRUE,lower=TRUE,apa=TRUE,short.names=TRUE,
    font.size ="scriptsize", heading="A correlation table from the psych package in R.",
     caption="cor2latex",label="default",silent=FALSE,file=NULL,append=FALSE,cut=0,big=0)
fa2latex(f,digits=2,rowlabels=TRUE,apa=TRUE,short.names=FALSE,cumvar=FALSE,
      cut=0,big=.3,alpha=.05,font.size ="scriptsize",
      heading="A factor analysis table from the psych package in R",
      caption="fa2latex",label="default",silent=FALSE,file=NULL,append=FALSE)
omega2latex(f,digits=2,rowlabels=TRUE,apa=TRUE,short.names=FALSE,cumvar=FALSE,cut=.2,
       font.size ="scriptsize",
       heading="An omega analysis table from the psych package in R",
      caption="omega2latex",label="default",silent=FALSE,file=NULL,append=FALSE)

irt2latex(f,digits=2,rowlabels=TRUE,apa=TRUE,short.names=FALSE,
      font.size ="scriptsize", heading="An IRT factor analysis table from R",
      caption="fa2latex",label="default",silent=FALSE,file=NULL,append=FALSE)
ICC2latex(icc,digits=2,rowlabels=TRUE,apa=TRUE,ci=TRUE,
   font.size ="scriptsize",big.mark=NULL, drop.na=TRUE,
    heading="A table from the psych package in R",
  caption="ICC2latex",label="default",char=FALSE,silent=FALSE,file=NULL,append=FALSE)
```

Arguments

x	A data frame or matrix to convert to LaTeX. If non-square, then correlations will be found prior to printing in cor2latex
digits	Round the output to digits of accuracy. NULL for formatting character data
rowlabels	If TRUE, use the row names from the matrix or data.frame

short.names	Name the columns with abbreviated rownames to save space
apa	If TRUE formats table in APA style
cumvar	For factor analyses, should we show the cumulative variance accounted for?
font.size	e.g., "scriptsize", "tiny" or anyother acceptable LaTeX font size.
heading	The label appearing at the top of the table
caption	The table caption
lower	in cor2latex, just show the lower triangular matrix
f	The object returned from a factor analysis using fa or irt.fa.
label	The label for the table
big.mark	Comma separate numbers large numbers (big.mark=",")
drop.na	Do not print NA values
method	When finding correlations, which method should be used (pearson)
use	use="pairwise" is the default when finding correlations in cor2latex
adjust	If showing probabilities, which adjustment should be used (holm)
stars	Should probability 'magic astericks' be displayed in cor2latex (FALSE)
char	char=TRUE allows printing tables with character information, but does not allow for putting in commas into numbers
cut	In omega2latex, df2latex and fa2latex, do not print abs(values) < cut
big	In fa2latex and df2latex boldface those abs(values) > big
alpha	If fa has returned confidence intervals, then what values of loadings should be boldfaced?
icc	Either the output of an ICC, or the data to be analyzed.
ci	Should confidence intervals of the ICC be displayed
silent	If TRUE, do not print any output, just return silently – useful if using Sweave
file	If specified, write the output to this file
append	If file is specified, then should we append (append=TRUE) or just write to the file

Value

A LaTeX table. Note that if showing "stars" for correlations, then one needs to use the siunitx package in LaTex. The entire LaTeX output is also returned invisibly. If using Sweave to create tables, then the silent option should be set to TRUE and the returned object saved as a file. See the last example.

Author(s)

William Revelle with suggestions from Jason French and David Condon and Davide Morselli

See Also

The many LaTeX conversion routines in Hmisc.

Examples

```
df2latex(psych::Thurstone,rowlabels=FALSE,apa=FALSE,short.names=FALSE,
        caption="Thurstone Correlation matrix")
df2latex(psych::Thurstone,heading="Thurstone Correlation matrix in APA style")

df2latex(psych::describe(psych::sat.act)[2:10],short.names=FALSE)
cor2latex(psych::Thurstone)
cor2latex(psych::sat.act,short.names=FALSE)
fa2latex(psych::fa(psych::Thurstone,3),heading="Factor analysis from R in quasi APA style")


 #If using Sweave to create a LateX table as a separate file then set silent=TRUE
#e.g.,
#LaTex preamble
#....
#<<print=FALSE,echo=FALSE>>=
#f3 <- fa(Thurstone,3)
#fa2latex(f3,silent=TRUE,file='testoutput.tex')
#@
#
#\input{testoutput.tex}
```

---

dfOrder   *Sort (order) a dataframe or matrix by multiple columns*

Description

Although order will order a vector, and it is possible to order several columns of a data.frame by specifying each column individually in the call to order, dfOrder will order a dataframe or matrix by as many columns as desired.

Usage

```
dfOrder(object, columns,absolute=FALSE,ascending=TRUE)
```

Arguments

object	The data.frame to be sorted
columns	Column numbers to use for sorting. If positive, then they will be sorted in increasing order. If negative, then in decreasing order
absolute	If TRUE, then sort the absolute values
ascending	By default, order from smallest to largest.

Details

This is just a simple helper function to reorder data.frames. Originally developed to organize IRT output from the ltm package. It is a basic add on to the order function.

(Completely rewritten for version 1.8.1.)

Value

The original data frame is now in sorted order.

Author(s)

William Revelle

See Also

Other useful file manipulation functions include `read.file` to read in data from a file or `read.clipboard` from the clipboard, `fileScan`, `filesList`, `filesInfo`, and `fileCreate`

`dfOrder` code is used in the `test.irt` function to combine ltm and `sim.irt` output.

Examples

```
set.seed(42)
x <- matrix(sample(1:4,64,replace=TRUE),ncol=4)
dfOrder(x)  # sort by all columns
dfOrder(x,c(1,4))  #sort by the first and 4th column
x.df <- data.frame(x)
dfOrder(x.df,c(1,-2))  #sort by the first in increasing order,
   #the second in decreasing order
```

---

epi   *Eysenck Personality Inventory (EPI) data for 3570 participants*

Description

The EPI is and has been a very frequently administered personality test with 57 measuring two broad dimensions, Extraversion-Introversion and Stability-Neuroticism, with an additional Lie scale. Developed by Eysenck and Eysenck, 1964. Eventually replaced with the EPQ which measures three broad dimensions. This data set represents 3570 observations collected in the early 1990s at the Personality, Motivation and Cognition lab at Northwestern. An additional data set (epiR) has test and retest information for 474 participants. The data are included here as demonstration of scale construction and test-retest reliability.

Usage

```
data(epi)
data(epi.dictionary)
data(epiR)
```

Format

A data frame with 3570 observations on the following 57 variables.

id The identification number within the study

time First (group testing) or 2nd time (before a lab experiment) for the epiR data set.

study Four lab based studies and their pretest data

V1 a numeric vector

V2 a numeric vector

V3 a numeric vector

V4 a numeric vector

V5 a numeric vector

V6 a numeric vector

V7 a numeric vector

V8 a numeric vector

V9 a numeric vector

V10 a numeric vector

V11 a numeric vector

V12 a numeric vector

V13 a numeric vector

V14 a numeric vector

V15 a numeric vector

V16 a numeric vector

V17 a numeric vector

V18 a numeric vector

V19 a numeric vector

V20 a numeric vector

V21 a numeric vector

V22 a numeric vector

V23 a numeric vector

V24 a numeric vector

V25 a numeric vector

V26 a numeric vector

V27 a numeric vector

V28 a numeric vector

V29 a numeric vector

V30 a numeric vector

V31 a numeric vector

V32 a numeric vector

V33 a numeric vector

V34 a numeric vector

V35 a numeric vector

V36 a numeric vector

V37 a numeric vector

V38 a numeric vector

V39 a numeric vector

V40 a numeric vector

V41 a numeric vector

V42 a numeric vector

V43 a numeric vector

V44 a numeric vector

V45 a numeric vector

V46 a numeric vector

V47 a numeric vector

V48 a numeric vector

V49 a numeric vector

V50 a numeric vector

V51 a numeric vector

V52 a numeric vector

V53 a numeric vector

V54 a numeric vector

V55 a numeric vector

V56 a numeric vector

V57 a numeric vector

Details

The original data were collected in a group testing framework for screening participants for subsequent studies. The participants were enrolled in an introductory psychology class between Fall, 1991 and Spring, 1995.

The actual items may be found in the `epi.dictionary`.

The structure of the E scale has been shown by Rocklin and Revelle (1981) to have two subcomponents, Impulsivity and Sociability. These were subsequently used by Revelle, Humphreys, Simon and Gilliland to examine the relationship between personality, caffeine induced arousal, and cognitive performance.

The epiR data include the original group testing data and matched data for 474 participants collected several weeks later. This is useful for showing that internal consistency estimates (e.g. `alpha` or `omega`) can be low even though the test is stable across time. For more demonstrations of the distinction between immediate internal consistency and delayed test-retest reliability see the `msqR` and `sai` data sets and `testRetest`.

Source

Data from the PMC laboratory at Northwestern.

References

Eysenck, H.J. and Eysenck, S. B.G. (1968). Manual for the Eysenck Personality Inventory.Educational and Industrial Testing Service, San Diego, CA.

Rocklin, T. and Revelle, W. (1981). The measurement of extraversion: A comparison of the Eysenck Personality Inventory and the Eysenck Personality Questionnaire. British Journal of Social Psychology, 20(4):279-284.

Examples

```
data(epi)
epi.keys <- list(E = c("V1",  "V3",  "V8",  "V10", "V13", "V17", "V22", "V25", "V27", "V39",
  "V44", "V46", "V49", "V53", "V56", "-V5", "-V15", "-V20", "-V29", "-V32", "-V34","-V37",
   "-V41", "-V51"),
N = c( "V2",  "V4",  "V7",  "V9", "V11", "V14", "V16", "V19", "V21", "V23", "V26", "V28",
"V31", "V33", "V35", "V38", "V40","V43", "V45", "V47", "V50", "V52","V55", "V57"),
L = c("V6",   "V24", "V36", "-V12", "-V18", "-V30", "-V42", "-V48", "-V54"),
Imp = c( "V1",  "V3",  "V8",  "V10", "V13", "V22", "V39", "-V5", "-V41"),
Soc = c( "V17", "V25", "V27", "V44", "V46", "V53", "-V11", "-V15", "-V20",
"-V29", "-V32", "-V37", "-V51")
)
scores <- psych::scoreItems(epi.keys,epi)

keys <- psych::make.keys(epi,epi.keys)   #the old way of making keys is to make a matrix
psych::fa.lookup(keys[,1:3],epi.dictionary) #show the items and keying information

#a variety of demonstrations (not run) of test retest reliability versus alpha versus omega

E <- psych::selectFromKeys(epi.keys$E)

psych::omega(epi[E])  #to show the low omega but high alpha of Extraversion
psych::testRetest(epiR,select=E)  #test retest of the extraversion scale (.82) is higher than
                       #alpha  for either the first (.77) or second administration (.74)
Imp <- psych::selectFromKeys(epi.keys$Imp)
psych::testRetest(epiR,select=Imp)  #Similarly test retest = .68 but alpha = .48 and .50.
Soc <- psych::selectFromKeys(epi.keys$Soc)
psych::testRetest(epiR,select=Soc)  # test retest =.83, alpha = .76, .75
N <- psych::selectFromKeys(epi.keys$N)
psych::testRetest(epiR, select=N) #Test retest = .8, alpha = .81, .80
```

---

epi.bfi | *13 personality scales from the Eysenck Personality Inventory and Big 5 inventory*

Description

A small data set of 5 scales from the Eysenck Personality Inventory, 5 from a Big 5 inventory, a Beck Depression Inventory, and State and Trait Anxiety measures. Used for demonstrations of correlations, regressions, graphic displays.

Usage

```
data(epi.bfi)
```

Format

A data frame with 231 observations on the following 13 variables.

epiE EPI Extraversion

epiS EPI Sociability (a subset of Extraversion items

epiImp EPI Impulsivity (a subset of Extraversion items

epilie EPI Lie scale

epiNeur EPI neuroticism

bfagree Big 5 inventory (from the IPIP) measure of Agreeableness

bfcon Big 5 Conscientiousness

bfext Big 5 Extraversion

bfneur Big 5 Neuroticism

bfopen Big 5 Openness

bdi Beck Depression scale

traitanx Trait Anxiety

stateanx State Anxiety

Details

Self report personality scales tend to measure the "Giant 2" of Extraversion and Neuroticism or the "Big 5" of Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness. Here is a small data set from Northwestern University undergraduates with scores on the Eysenck Personality Inventory (EPI) and a Big 5 inventory taken from the International Personality Item Pool.

Source

Data were collected at the Personality, Motivation, and Cognition Lab (PMCLab) at Northwestern by William Revelle)

References

https://personality-project.org/pmc.html

Examples

```
data(epi.bfi)
psych::pairs.panels(epi.bfi[,1:5])
psych::describe(epi.bfi)
```

---

galton   *Galton's Mid parent child height data*

Description

Two of the earliest examples of the correlation coefficient were Francis Galton's data sets on the relationship between mid parent and child height and the similarity of parent generation peas with child peas. This is the data set for the Galton height.

Usage

```
data(galton)
```

Format

A data frame with 928 observations on the following 2 variables.

parent Mid Parent heights (in inches)

child Child Height

Details

Female heights were adjusted by 1.08 to compensate for sex differences. (This was done in the original data set)

Source

This is just the galton data set from UsingR, slightly rearranged.

References

Stigler, S. M. (1999). Statistics on the Table: The History of Statistical Concepts and Methods. Harvard University Press. Galton, F. (1886). Regression towards mediocrity in hereditary stature. Journal of the Anthropological Institute of Great Britain and Ireland, 15:246-263. Galton, F. (1869). Hereditary Genius: An Inquiry into its Laws and Consequences. London: Macmillan.

Wachsmuth, A.W., Wilkinson L., Dallal G.E. (2003). Galton's bend: A previously undiscovered nonlinearity in Galton's family stature regression data. The American Statistician, 57, 190-192.

See Also

The other Galton data sets: heights, peas,cubits

Examples

```
data(galton)
psych::describe(galton)
 #show the scatter plot and the lowess fit
psych::pairs.panels(galton,main="Galton's Parent child heights")
#but this makes the regression lines look the same
psych::pairs.panels(galton,lm=TRUE,main="Galton's Parent child heights")
 #better is to scale them
psych::pairs.panels(galton,lm=TRUE,xlim=c(62,74),ylim=c(62,74),
               main="Galton's Parent child heights")
```

---

heights   *A data.frame of the Galton (1888) height and cubit data set.*

Description

Francis Galton introduced the 'co-relation' in 1888 with a paper discussing how to measure the relationship between two variables. His primary example was the relationship between height and forearm length. The data table (cubits) is taken from Galton (1888). Unfortunately, there seem to be some errors in the original data table in that the marginal totals do not match the table.

The data frame, heights, is converted from this table using table2df.

Usage

```
data(heights)
```

Format

A data frame with 348 observations on the following 2 variables.

height Height in inches

cubit Forearm length in inches

Details

Sir Francis Galton (1888) published the first demonstration of the correlation coefficient. The regression (or reversion to mediocrity) of the height to the length of the left forearm (a cubit) was found to .8. The original table cubits is taken from Galton (1888). There seem to be some errors in the table as published in that the row sums do not agree with the actual row sums. These data are used to create a matrix using table2matrix for demonstrations of analysis and displays of the data.

Source

Galton (1888)

References

Galton, Francis (1888) Co-relations and their measurement. Proceedings of the Royal Society. London Series,45,135-145,

See Also

table2matrix, table2df, cubits, ellipses, galton

Examples

```
data(heights)
psych::ellipses(heights,n=1,main="Galton's co-relation data set")
```

---

income   *US family income from US census 2008*

Description

US census data on family income from 2008

Usage

```
data(income)
```

Format

A data frame with 44 observations on the following 4 variables.

value lower boundary of the income group

count Number of families within that income group

mean Mean of the category

prop proportion of families

Details

The distribution of income is a nice example of a log normal distribution. It is also an interesting example of the power of graphics. It is quite clear when graphing the data that income statistics are bunched to the nearest 5K. That is, there is a clear sawtooth pattern in the data.

The all.income set is interpolates intervening values for 100-150K, 150-200K and 200-250K

Source

US Census: Table HINC-06. Income Distribution to $250,000 or More for Households: 2008

https://www.census.gov/hhes/www/cpstables/032009/hhinc/new06_000.htm

Examples

```
data(income)
with(income[1:40,], plot(mean,prop, main="US family income for 2008",xlab="income",
        ylab="Proportion of families",xlim=c(0,100000)))
with (income[1:40,], points(lowess(mean,prop,f=.3),typ="l"))
psych::describe(income)


with(all.income, plot(mean,prop, main="US family income for 2008",xlab="income",
                ylab="Proportion of families",xlim=c(0,250000)))
with (all.income[1:50,], points(lowess(mean,prop,f=.25),typ="l"))
```

---

iqitems   *16 multiple choice IQ items*

Description

16 multiple choice ability items taken from the Synthetic Aperture Personality Assessment (SAPA) web based personality assessment project. The data from 1525 subjects are included here as a demonstration set for scoring multiple choice inventories and doing basic item statistics. For more information on the development of an open source measure of cognitive ability, consult the readings available at the https://personality-project.org.

Usage

```
data(iqitems)
```

Format

A data frame with 1525 observations on the following 16 variables. The number following the name is the item number from SAPA.

reason.4 Basic reasoning questions

reason.16 Basic reasoning question

reason.17 Basic reasoning question

reason.19 Basic reasoning question

letter.7 In the following alphanumeric series, what letter comes next?

letter.33 In the following alphanumeric series, what letter comes next?

letter.34 In the following alphanumeric series, what letter comes next

letter.58 In the following alphanumeric series, what letter comes next?

matrix.45 A matrix reasoning task

matrix.46 A matrix reasoning task

matrix.47 A matrix reasoning task

`matrix.55` A matrix reasoning task

`rotate.3` Spatial Rotation of type 1.2

`rotate.4` Spatial Rotation of type 1.2

`rotate.6` Spatial Rotation of type 1.1

`rotate.8` Spatial Rotation of type 2.3

Details

16 items were sampled from 80 items given as part of the SAPA (<https://sapa-project.org>) project (Revelle, Wilt and Rosenthal, 2009; Condon and Revelle, 2014) to develop online measures of ability. These 16 items reflect four lower order factors (verbal reasoning, letter series, matrix reasoning, and spatial rotations. These lower level factors all share a higher level factor ('g'). Similar data are available from the International Cognitive Abiity Resource at <https://icar-project.org> .

This data set and the associated data set (`ability` based upon scoring these multiple choice items and converting them to correct/incorrect may be used to demonstrate item response functions, `tetrachoric` correlations, or `irt.fa` as well as `omega` estimates of of reliability and hierarchical structure.

In addition, the data set is a good example of doing item analysis to examine the empirical response probabilities of each item alternative as a function of the underlying latent trait. When doing this, it appears that two of the matrix reasoning problems do not have monotonically increasing trace lines for the probability correct. At moderately high ability (theta = 1) there is a decrease in the probability correct from theta = 0 and theta = 2.

Source

The example data set is taken from the Synthetic Aperture Personality Assessment personality and ability test at <https://sapa-project.org>. The data were collected with David Condon from 8/08/12 to 8/31/12.

References

Condon, David and Revelle, William, (2014) The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. Intelligence, 43, 52-64.

Revelle, W., Wilt, J., and Rosenthal, A. (2010) Individual Differences in Cognition: New Methods for examining the Personality-Cognition Link In Gruszka, A. and Matthews, G. and Szymura, B. (Eds.) Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control, Springer.

Revelle, W, Condon, D.M., Wilt, J., French, J.A., Brown, A., and Elleman, L.G. (2016) Web and phone based data collection using planned missing designs. In Fielding, N.G., Lee, R.M. and Blank, G. (Eds). SAGE Handbook of Online Research Methods (2nd Ed), Sage Publcations.

Examples

```
data(iqitems)
iq.keys <- c(4,4,4, 6,  6,3,4,4,   5,2,2,4,   3,2,6,7)
```

```
psych::score.multiple.choice(iq.keys,iqitems)   #this just gives summary statisics
#convert them to true false
iq.scrub <- psych::scrub(iqitems,isvalue=0)  #first get rid of the zero responses
iq.tf <-  psych::score.multiple.choice(iq.keys,iq.scrub,score=FALSE)
                #convert to wrong (0) and correct (1) for analysis
psych::describe(iq.tf)
#see the ability data set for these analyses
#now, for some item analysis
iq.irt <- psych::irt.fa(iq.tf)  #do a basic irt
iq.sc <- psych::scoreIrt(iq.irt,iq.tf)  #find the scores
op <- par(mfrow=c(4,4))
irt.responses(iq.sc[,1], iq.tf)
op <- par(mfrow=c(1,1))
```

---

msq	*75 mood items from the Motivational State Questionnaire for 3896 participants*

Description

Emotions may be described either as discrete emotions or in dimensional terms. The Motivational State Questionnaire (MSQ) was developed to study emotions in laboratory and field settings. The data can be well described in terms of a two dimensional solution of energy vs tiredness and tension versus calmness. Additional items include what time of day the data were collected and a few personality questionnaire scores.

Usage

```
data(msq)
```

Format

A data frame with 3896 observations on the following 92 variables.

active a numeric vector

afraid a numeric vector

alert a numeric vector

angry a numeric vector

anxious a numeric vector

aroused a numeric vector

ashamed a numeric vector

astonished a numeric vector

at.ease a numeric vector

at.rest a numeric vector

attentive a numeric vector

blue a numeric vector

bored a numeric vector

calm a numeric vector

cheerful a numeric vector

clutched.up a numeric vector

confident a numeric vector

content a numeric vector

delighted a numeric vector

depressed a numeric vector

determined a numeric vector

distressed a numeric vector

drowsy a numeric vector

dull a numeric vector

elated a numeric vector

energetic a numeric vector

enthusiastic a numeric vector

excited a numeric vector

fearful a numeric vector

frustrated a numeric vector

full.of.pep a numeric vector

gloomy a numeric vector

grouchy a numeric vector

guilty a numeric vector

happy a numeric vector

hostile a numeric vector

idle a numeric vector

inactive a numeric vector

inspired a numeric vector

intense a numeric vector

interested a numeric vector

irritable a numeric vector

jittery a numeric vector

lively a numeric vector

lonely a numeric vector

nervous a numeric vector

placid a numeric vector

`pleased` a numeric vector

`proud` a numeric vector

`quiescent` a numeric vector

`quiet` a numeric vector

`relaxed` a numeric vector

`sad` a numeric vector

`satisfied` a numeric vector

`scared` a numeric vector

`serene` a numeric vector

`sleepy` a numeric vector

`sluggish` a numeric vector

`sociable` a numeric vector

`sorry` a numeric vector

`still` a numeric vector

`strong` a numeric vector

`surprised` a numeric vector

`tense` a numeric vector

`tired` a numeric vector

`tranquil` a numeric vector

`unhappy` a numeric vector

`upset` a numeric vector

`vigorous` a numeric vector

`wakeful` a numeric vector

`warmhearted` a numeric vector

`wide.awake` a numeric vector

`alone` a numeric vector

`kindly` a numeric vector

`scornful` a numeric vector

`EA` Thayer's Energetic Arousal Scale

`TA` Thayer's Tense Arousal Scale

`PA` Positive Affect scale

`NegAff` Negative Affect scale

`Extraversion` Extraversion from the Eysenck Personality Inventory

`Neuroticism` Neuroticism from the Eysenck Personality Inventory

`Lie` Lie from the EPI

`Sociability` The sociability subset of the Extraversion Scale

`Impulsivity` The impulsivity subset of the Extraversions Scale

MSQ_Time Time of day the data were collected

MSQ_Round Rounded time of day

TOD a numeric vector

TOD24 a numeric vector

ID subject ID

condition What was the experimental condition after the msq was given

scale a factor with levels msq r original or revised msq

exper Which study were the data collected: a factor with levels AGES BING BORN CART CITY COPE
    EMIT FAST Fern FILM FLAT Gray imps item knob MAPS mite pat-1 pat-2 PATS post RAFT
    Rim.1 Rim.2 rob-1 rob-2 ROG1 ROG2 SALT sam-1 sam-2 SAVE/PATS sett swam swam-2 TIME
    VALE-1 VALE-2 VIEW

Details

The Motivational States Questionnaire (MSQ) is composed of 72 items, which represent the full af-
fective space (Revelle & Anderson, 1998). The MSQ consists of 20 items taken from the Activation-
Deactivation Adjective Check List (Thayer, 1986), 18 from the Positive and Negative Affect Sched-
ule (PANAS, Watson, Clark, & Tellegen, 1988) along with the items used by Larsen and Diener
(1992). The response format was a four-point scale that corresponds to Russell and Carroll's (1999)
"ambiguous–likely-unipolar format" and that asks the respondents to indicate their current standing
("at this moment") with the following rating scale:
0——————-1——————-2——————-3
Not at all A little Moderately Very much

The original version of the MSQ included 70 items. Intermediate analyses (done with 1840 subjects)
demonstrated a concentration of items in some sections of the two dimensional space, and a paucity
of items in others. To begin correcting this, 3 items from redundantly measured sections (alone,
kindly, scornful) were removed, and 5 new ones (anxious, cheerful, idle, inactive, and tranquil) were
added. Thus, the correlation matrix is missing the correlations between items anxious, cheerful,
idle, inactive, and tranquil with alone, kindly, and scornful.

Procedure. The data were collected over nine years, as part of a series of studies examining the
effects of personality and situational factors on motivational state and subsequent cognitive perfor-
mance. In each of 38 studies, prior to any manipulation of motivational state, participants signed a
consent form and filled out the MSQ. (The procedures of the individual studies are irrelevant to this
data set and could not affect the responses to the MSQ, since this instrument was completed before
any further instructions or tasks). Some MSQ post test (after manipulations) is available in affect.

The EA and TA scales are from Thayer, the PA and NA scales are from Watson et al. (1988). Scales
and items:

Energetic Arousal: active, energetic, vigorous, wakeful, wide.awake, full.of.pep, lively, -sleepy,
-tired, - drowsy (ADACL)

Tense Arousal: Intense, Jittery, fearful, tense, clutched up, -quiet, -still, - placid, - calm, -at rest
(ADACL)

Positive Affect: active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud,
strong (PANAS)

Negative Affect: afraid, ashamed, distressed, guilty, hostile, irritable , jittery, nervous, scared, upset (PANAS)

The PA and NA scales can in turn can be thought of as having subscales: (See the PANAS-X) Fear: afraid, scared, nervous, jittery (not included frightened, shaky) Hostility: angry, hostile, irritable, (not included: scornful, disgusted, loathing guilt: ashamed, guilty, (not included: blameworthy, angry at self, disgusted with self, dissatisfied with self) sadness: alone, blue, lonely, sad, (not included: downhearted) joviality: cheerful, delighted, energetic, enthusiastic, excited, happy, lively, (not included: joyful) self-assurance: proud, strong, confident, (not included: bold, daring, fearless) attentiveness: alert, attentive, determined (not included: concentrating)

The next set of circumplex scales were taken (I think) from Larsen and Diener (1992). High activation: active, aroused, surprised, intense, astonished Activated PA: elated, excited, enthusiastic, lively Unactivated NA : calm, serene, relaxed, at rest, content, at ease PA: happy, warmhearted, pleased, cheerful, delighted Low Activation: quiet, inactive, idle, still, tranquil Unactivated PA: dull, bored, sluggish, tired, drowsy NA: sad, blue, unhappy, gloomy, grouchy Activated NA: jittery, anxious, nervous, fearful, distressed.

Keys for these separate scales are shown in the examples.

In addition to the MSQ, there are 5 scales from the Eysenck Personality Inventory (Extraversion, Impulsivity, Sociability, Neuroticism, Lie). The Imp and Soc are subsets of the the total extraversion scale.

Source

Data collected at the Personality, Motivation, and Cognition Laboratory, Northwestern University.

References

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark (Ed.), Review of personality and social psychology, No. 13. Emotion (pp. 25-59). Thousand Oaks, CA, US: Sage Publications, Inc.

Rafaeli, Eshkol and Revelle, William (2006), A premature consensus: Are happiness and sadness truly opposite affects? Motivation and Emotion, 30, 1, 1-12.

Revelle, W. and Anderson, K.J. (1998) Personality, motivation and cognitive performance: Final report to the Army Research Institute on contract MDA 903-93-K-0008. (https://www.personality-project.org/revelle/publications/ra.ari.98.pdf).

Thayer, R.E. (1989) The biopsychology of mood and arousal. Oxford University Press. New York, NY.

Watson,D., Clark, L.A. and Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6):1063-1070.

See Also

msqR for a larger data set with repeated measures for 3032 participants measured at least once, 2753 measured twice, 446 three times and 181 four times. affect for an example of the use of some of these adjectives in a mood manipulation study.

make.keys, scoreItems and scoreOverlap for instructions on how to score multiple scales with and without item overlap. Also see fa and fa.extension for instructions on how to do factor analyses or factor extension.

Examples

```
data(msq)
   #in in the interests of time
#basic descriptive statistics
psych::describe(msq)

#score them for 20 short scales -- note that these have item overlap
#The first 2 are from Thayer
#The next 2 are classic positive and negative affect
#The next 9 are circumplex scales
#the last 7 are msq estimates of PANASX scales (missing some items)
keys.list <- list(
EA = c("active", "energetic", "vigorous", "wakeful", "wide.awake", "full.of.pep",
       "lively", "-sleepy", "-tired", "-drowsy"),
TA =c("intense", "jittery", "fearful", "tense", "clutched.up", "-quiet", "-still",
       "-placid", "-calm", "-at.rest") ,
PA =c("active", "excited", "strong", "inspired", "determined", "attentive",
          "interested", "enthusiastic", "proud", "alert"),
NAf =c("jittery", "nervous", "scared", "afraid", "guilty", "ashamed", "distressed",
          "upset", "hostile", "irritable" ),
HAct = c("active", "aroused", "surprised", "intense", "astonished"),
aPA = c("elated", "excited", "enthusiastic", "lively"),
uNA = c("calm", "serene", "relaxed", "at.rest", "content", "at.ease"),
pa = c("happy", "warmhearted", "pleased", "cheerful", "delighted" ),
LAct = c("quiet", "inactive", "idle", "still", "tranquil"),
uPA =c( "dull", "bored", "sluggish", "tired", "drowsy"),
naf = c( "sad", "blue", "unhappy", "gloomy", "grouchy"),
aNA = c("jittery", "anxious", "nervous", "fearful", "distressed"),
Fear = c("afraid" , "scared" , "nervous" , "jittery" ) ,
Hostility = c("angry" ,  "hostile", "irritable", "scornful" ),
Guilt = c("guilty" , "ashamed" ),
Sadness = c( "sad"  , "blue" , "lonely",  "alone" ),
Joviality =c("happy","delighted", "cheerful", "excited", "enthusiastic", "lively", "energetic"),
Self.Assurance=c( "proud","strong" , "confident" , "-fearful" ),
Attentiveness = c("alert" , "determined" , "attentive" )
#, acquiscence = c("sleepy" ,  "wakeful" ,  "relaxed","tense")
#dropped because it has a negative alpha and throws warnings
   )

msq.scores <- psych::scoreItems(keys.list,msq)

#show a circumplex structure for the non-overlapping items
fcirc <- psych::fa(msq.scores$scores[,5:12],2)
psych::fa.plot(fcirc,labels=colnames(msq.scores$scores)[5:12])

#now, find the correlations corrected for item overlap
msq.overlap <- psych::scoreOverlap(keys.list,msq)
#a warning is thrown by smc  because of some NAs in the matrix
```

```
f2 <- psych::fa(msq.overlap$cor,2)
psych::fa.plot(f2,labels=colnames(msq.overlap$cor),
      title="2 dimensions of affect, corrected for overlap")

#extend this solution to EA/TA  NA/PA space
fe  <- psych::fa.extension(cor(msq.scores$scores[,5:12],msq.scores$scores[,1:4]),fcirc)
psych::fa.diagram(fcirc,fe=fe,
          main="Extending the circumplex structure to  EA/TA and PA/NA ")

#show the 2 dimensional structure
f2 <- psych::fa(msq[1:72],2)
psych::fa.plot(f2,labels=colnames(msq)[1:72],
     title="2 dimensions of affect at the item level",cex=.5)

#sort them by polar coordinates
round(polar(f2),2)
```

msqR	*75 mood items from the Motivational State Questionnaire for 3032 unique participants*

Description

Emotions may be described either as discrete emotions or in dimensional terms. The Motivational State Questionnaire (MSQ) was developed to study emotions in laboratory and field settings. The data can be well described in terms of a two dimensional solution of energy vs tiredness and tension versus calmness. Alternatively, this space can be organized by the two dimensions of Positive Affect and Negative Affect. Additional items include what time of day the data were collected and a few personality questionnaire scores. 3032 unique participants took the MSQ at least once, 2753 at least twice, 446 three times, and 181 four times. The 3032 participants also took the [sai](sai) state anxiety inventory at the same time. Some studies manipulated arousal by caffeine, others manipulations included affect inducing movies.

Usage

```
data("msqR")
```

Format

A data frame with 6411 observations on the following 88 variables.

active a numeric vector

afraid a numeric vector

alert a numeric vector

alone a numeric vector

angry a numeric vector

aroused a numeric vector

ashamed a numeric vector

astonished a numeric vector

at.ease a numeric vector

at.rest a numeric vector

attentive a numeric vector

blue a numeric vector

bored a numeric vector

calm a numeric vector

clutched.up a numeric vector

confident a numeric vector

content a numeric vector

delighted a numeric vector

depressed a numeric vector

determined a numeric vector

distressed a numeric vector

drowsy a numeric vector

dull a numeric vector

elated a numeric vector

energetic a numeric vector

enthusiastic a numeric vector

excited a numeric vector

fearful a numeric vector

frustrated a numeric vector

full.of.pep a numeric vector

gloomy a numeric vector

grouchy a numeric vector

guilty a numeric vector

happy a numeric vector

hostile a numeric vector

inspired a numeric vector

intense a numeric vector

interested a numeric vector

irritable a numeric vector

jittery a numeric vector

lively a numeric vector

lonely a numeric vector

nervous a numeric vector

placid a numeric vector

pleased a numeric vector

proud a numeric vector

quiescent a numeric vector

quiet a numeric vector

relaxed a numeric vector

sad a numeric vector

satisfied a numeric vector

scared a numeric vector

serene a numeric vector

sleepy a numeric vector

sluggish a numeric vector

sociable a numeric vector

sorry a numeric vector

still a numeric vector

strong a numeric vector

surprised a numeric vector

tense a numeric vector

tired a numeric vector

unhappy a numeric vector

upset a numeric vector

vigorous a numeric vector

wakeful a numeric vector

warmhearted a numeric vector

wide.awake a numeric vector

anxious a numeric vector

cheerful a numeric vector

idle a numeric vector

inactive a numeric vector

tranquil a numeric vector

kindly a numeric vector

scornful a numeric vector

Extraversion Extraversion from the EPI

Neuroticism Neuroticism from the EPI

Lie Lie from the EPI

Sociability Sociability from the EPI

Impulsivity Impulsivity from the EPI

gender 1= male, 2 = female (coded on presumed x chromosome). Slowly being added to the data set.

TOD Time of day that the study was run

drug 1 if given placebo, 2 if given caffeine

film 1-4 if given a film: 1=Frontline, 2= Halloween, 3=Serengeti, 4 = Parenthood

time Measurement occasion (1 and 2 are same session, 3 and 4 are the same, but a later session)

id a numeric vector

form msq versus msqR

study a character vector of the experiment name

Details

The Motivational States Questionnaire (MSQ) is composed of 75 items, which represent the full affective space (Revelle & Anderson, 1998). The MSQ consists of 20 items taken from the Activation-Deactivation Adjective Check List (Thayer, 1986), 18 from the Positive and Negative Affect Schedule (PANAS, Watson, Clark, & Tellegen, 1988) along with the affective circumplex items used by Larsen and Diener (1992). The response format was a four-point scale that corresponds to Russell and Carroll's (1999) "ambiguous–likely-unipolar format" and that asks the respondents to indicate their current standing ("at this moment") with the following rating scale:
0————————-1————————-2————————-3
Not at all A little Moderately Very much

The original version of the MSQ included 70 items. Intermediate analyses (done with 1840 subjects) demonstrated a concentration of items in some sections of the two dimensional space, and a paucity of items in others. To begin correcting this, 3 items from redundantly measured sections (alone, kindly, scornful) were removed, and 5 new ones (anxious, cheerful, idle, inactive, and tranquil) were added. Thus, the correlation matrix is missing the correlations between items anxious, cheerful, idle, inactive, and tranquil with alone, kindly, and scornful.

2605 individuals took Form 1 version, 3806 the Form 2 version. 3032 people (1218 form 1, 1814 form 2) took the MSQ at least once. 2086 at least twice, 1112 three times, and 181 four times.

To see the relative frequencies by time and form, see the first example.

Procedure. The data were collected over nine years in the Personality, Motivation and Cognition laboratory at Northwestern, as part of a series of studies examining the effects of personality and situational factors on motivational state and subsequent cognitive performance. In each of 38 studies, prior to any manipulation of motivational state, participants signed a consent form and in some studies, consumed 0 or 4mg/kg of caffeine. In caffeine studies, they waited 30 minutes and then filled out the MSQ. (Normally, the procedures of the individual studies are irrelevant to this data set and could not affect the responses to the MSQ at time 1, since this instrument was completed before any further instructions or tasks. However, caffeine does have an effect.) The MSQ post test following a movie manipulation) is available in affect as well as here.

The XRAY study crossed four movie conditions with caffeine. The first MSQ measures are showing the effects of the movies and caffeine, but after an additional 30 minutes, the second MSQ seems

to mainly show the caffeine effects. The movies were 9 minute clips from 1) a BBC documentary on British troops arriving at the Bergen-Belsen concentration camp (sad); 2) an early scene from Halloween in which the heroine runs around shutting doors and windows (terror); 3) a documentary about lions on the Serengeti plain, and 4) the "birthday party" scene from Parenthood.

The FLAT study measured affect before, immediately after, and then after 30 minutes following a movie manipulation. See the affect data set.

To see which studies used which conditions, see the second and third examples.

The EA and TA scales are from Thayer, the PA and NA scales are from Watson et al. (1988). Scales and items:

Energetic Arousal: active, energetic, vigorous, wakeful, wide.awake, full.of.pep, lively, -sleepy, -tired, - drowsy (ADACL)

Tense Arousal: Intense, Jittery, fearful, tense, clutched up, -quiet, -still, - placid, - calm, -at rest (ADACL)

Positive Affect: active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong (PANAS)

Negative Affect: afraid, ashamed, distressed, guilty, hostile, irritable , jittery, nervous, scared, upset (PANAS)

The PA and NA scales can in turn can be thought of as having subscales: (See the PANAS-X) Fear: afraid, scared, nervous, jittery (not included frightened, shaky) Hostility: angry, hostile, irritable, (not included: scornful, disgusted, loathing guilt: ashamed, guilty, (not included: blameworthy, angry at self, disgusted with self, dissatisfied with self) sadness: alone, blue, lonely, sad, (not included: downhearted) joviality: cheerful, delighted, energetic, enthusiastic, excited, happy, lively, (not included: joyful) self-assurance: proud, strong, confident, (not included: bold, daring, fearless) attentiveness: alert, attentive, determined (not included: concentrating)

The next set of circumplex scales were taken from Larsen and Diener (1992). High activation: active, aroused, surprised, intense, astonished Activated PA: elated, excited, enthusiastic, lively Unactivated NA : calm, serene, relaxed, at rest, content, at ease PA: happy, warmhearted, pleased, cheerful, delighted Low Activation: quiet, inactive, idle, still, tranquil Unactivated PA: dull, bored, sluggish, tired, drowsy NA: sad, blue, unhappy, gloomy, grouchy Activated NA: jittery, anxious, nervous, fearful, distressed.

Keys for these separate scales are shown in the examples.

In addition to the MSQ, there are 5 scales from the Eysenck Personality Inventory (Extraversion, Impulsivity, Sociability, Neuroticism, Lie). The Imp and Soc are subsets of the the total extraversion scale based upon a reanalysis of the EPI by Rocklin and Revelle (1983). This information is in the msq data set as well.

Note

In December, 2018 the caffeine, film and personality conditions were added. In the process of doing so, it was discovered that the EMIT data had been incorrectly entered. This has been fixed.

Source

Data collected at the Personality, Motivation, and Cognition Laboratory, Northwestern University.

References

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark (Ed.), Review of personality and social psychology, No. 13. Emotion (pp. 25-59). Thousand Oaks, CA, US: Sage Publications, Inc.

Rafaeli, Eshkol and Revelle, William (2006), A premature consensus: Are happiness and sadness truly opposite affects? Motivation and Emotion, 30, 1, 1-12.

Revelle, W. and Anderson, K.J. (1998) Personality, motivation and cognitive performance: Final report to the Army Research Institute on contract MDA 903-93-K-0008. (https://www.personality-project.org/revelle/publications/ra.ari.98.pdf).

Smillie, Luke D. and Cooper, Andrew and Wilt, Joshua and Revelle, William (2012) Do Extraverts Get More Bang for the Buck? Refining the Affective-Reactivity Hypothesis of Extraversion. Journal of Personality and Social Psychology, 103 (2), 206-326.

Thayer, R.E. (1989) The biopsychology of mood and arousal. Oxford University Press. New York, NY.

Watson,D., Clark, L.A. and Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6):1063-1070.

See Also

`msq` for 3896 participants with scores on five scales of the EPI. `affect` for an example of the use of some of these adjectives in a mood manipulation study.

`make.keys`, `scoreItems` and `scoreOverlap` for instructions on how to score multiple scales with and without item overlap. Also see `fa` and `fa.extension` for instructions on how to do factor analyses or factor extension.

Given the temporal ordering of the `sai` data and the `msqR` data, these data are useful for demonstrations of `testRetest` reliability. See the examples in `testRetest` for how to combine the `sai tai` and `msqR` datasets.

Examples

```
data(msqR)
table(msqR$form,msqR$time) #which forms?
table(msqR$study,msqR$drug) #Drug studies
table(msqR$study,msqR$film) #Film studies
table(msqR$study,msqR$TOD) #To examine time of day


#score them for 20 short scales -- note that these have item overlap
#The first 2 are from Thayer
#The next 2 are classic positive and negative affect
#The next 9 are circumplex scales
#the last 7 are msq estimates of PANASX scales (missing some items)
keys.list <- list(
EA = c("active", "energetic", "vigorous", "wakeful", "wide.awake", "full.of.pep",
       "lively", "-sleepy", "-tired", "-drowsy"),
TA =c("intense", "jittery", "fearful", "tense", "clutched.up", "-quiet", "-still",
       "-placid", "-calm", "-at.rest") ,
```

```
    PA =c("active", "excited", "strong", "inspired", "determined", "attentive",
            "interested", "enthusiastic", "proud", "alert"),
    NAf =c("jittery", "nervous", "scared", "afraid", "guilty", "ashamed", "distressed",
            "upset", "hostile", "irritable" ),
    HAct = c("active", "aroused", "surprised", "intense", "astonished"),
    aPA = c("elated", "excited", "enthusiastic", "lively"),
    uNA = c("calm", "serene", "relaxed", "at.rest", "content", "at.ease"),
    pa = c("happy", "warmhearted", "pleased", "cheerful", "delighted" ),
    LAct = c("quiet", "inactive", "idle", "still", "tranquil"),
    uPA =c( "dull", "bored", "sluggish", "tired", "drowsy"),
    naf = c( "sad", "blue", "unhappy", "gloomy", "grouchy"),
    aNA = c("jittery", "anxious", "nervous", "fearful", "distressed"),
    Fear = c("afraid" , "scared" , "nervous" , "jittery" ) ,
    Hostility = c("angry" ,  "hostile", "irritable", "scornful" ),
    Guilt = c("guilty" , "ashamed" ),
    Sadness = c( "sad"  , "blue" , "lonely",  "alone" ),
    Joviality =c("happy","delighted", "cheerful", "excited", "enthusiastic", "lively", "energetic"),
    Self.Assurance=c( "proud","strong" , "confident" , "-fearful" ),
    Attentiveness = c("alert" , "determined" , "attentive" ))

    #acquiscence = c("sleepy" ,  "wakeful" ,  "relaxed","tense")
    #Yik Russell and Steiger list the following items
    Yik.keys <- list(
    pleasure =psych::cs(happy,content,satisfied, pleased),
    act.pleasure =psych::cs(proud,enthusiastic,euphoric),
    pleasant.activation = psych::cs(energetic,full.of.pep,excited,wakeful,attentive,
        wide.awake,active,alert,vigorous),
    activation = psych::cs(aroused,hyperactivated,intense),
    unpleasant.act = psych::cs(anxious,frenzied,jittery,nervous),
    activated.displeasure =psych::cs(scared,upset,shaky,fearful,clutched.up,tense,
        ashamed,guilty,agitated,hostile),
    displeaure =psych::cs(troubled,miserable,unhappy,dissatisfied),
    Ueactivated.Displeasure = psych::cs(sad,down,gloomy,blue,melancholy),
    Unpleasant.Deactivation = psych::cs(droopy,drowsy,dull,bored,sluggish,tired),
    Deactivation =psych::cs( quiet,still),
    pleasant.deactivation = psych::cs(placid,relaxed,tranquil, at.rest,calm),
    deactived.pleasure =psych::cs( serene,soothed,peaceful,at.ease,secure)
    )

    #of these 60 items, 46 appear in the msqR
    Yik.msq.keys <- list(
    Pleasure =psych::cs(happy,content,satisfied, pleased),
    Activated.Pleasure =psych::cs(proud,enthusiastic),
    Pleasant.Activation = psych::cs(energetic,full.of.pep,excited,wakeful,attentive,
        wide.awake,active,alert,vigorous),
    Activation = psych::cs(aroused,intense),
    Unpleasant.Activation = psych::cs(anxious,jittery,nervous),
    Activated.Displeasure =psych::cs(scared,upset,fearful,
            clutched.up,tense,ashamed,guilty,hostile),
    Displeaure = psych::cs(unhappy),
    Deactivated.Displeasure = psych::cs(sad,gloomy,blue),
    Unpleasant.Deactivation = psych::cs(drowsy,dull,bored,sluggish,tired),
    Deactivation =psych::cs( quiet,still),
```

```
Pleasant.Deactivation = psych::cs(placid,relaxed,tranquil, at.rest,calm),
Deactivated.Pleasure =psych::cs( serene,at.ease)
)
yik.scores <- psych::scoreItems(Yik.msq.keys,msqR)
yik <- yik.scores$scores
f2.yik <- psych::fa(yik,2) #factor the yik scores
psych::fa.plot(f2.yik,labels=colnames(yik),title="Yik-Russell-Steiger circumplex",cex=.8,
      pos=(c(1,1,2,1,1,1,3,1,4,1,2,4)))


msq.scores <- psych::scoreItems(keys.list,msqR)

#show a circumplex structure for the non-overlapping items
fcirc <- psych::fa(msq.scores$scores[,5:12],2)
psych::fa.plot(fcirc,labels=colnames(msq.scores$scores)[5:12])


#now, find the correlations corrected for item overlap
msq.overlap <- psych::scoreOverlap(keys.list,msqR)
f2 <- psych::fa(msq.overlap$cor,2)
psych::fa.plot(f2,labels=colnames(msq.overlap$cor),
            title="2 dimensions of affect, corrected for overlap")

#extend this solution to EA/TA  NA/PA space
fe  <- psych::fa.extension(cor(msq.scores$scores[,5:12],msq.scores$scores[,1:4]),fcirc)
psych::fa.diagram(fcirc,fe=fe,main="Extending the circumplex structure to  EA/TA and PA/NA ")

#show the 2 dimensional structure
f2 <- psych::fa(msqR[1:72],2)
psych::fa.plot(f2,labels=colnames(msqR)[1:72],title="2 dimensions of affect at the item level")

#sort them by polar coordinates
round(psych::polar(f2),2)

#the msqR and sai data sets have 10 overlapping items which can be used for
#testRetest analysis.  We need to specify the keys, and then choose the appropriate
#data sets
sai.msq.keys <- list(pos =c( "at.ease" ,  "calm" , "confident", "content","relaxed"),
  neg = c("anxious", "jittery", "nervous" ,"tense"  ,   "upset"),
  anx = c("anxious", "jittery", "nervous" ,"tense", "upset","-at.ease" ,  "-calm" ,
  "-confident", "-content","-relaxed"))

select <- psych::selectFromKeys(sai.msq.keys$anx)
#The following is useful for examining test retest reliabilities
msq.control <- subset(msqR,is.element( msqR$study , c("Cart", "Fast", "SHED", "SHOP")))
msq.film <- subset(msqR,(is.element( msqR$study ,  c("FIAT", "FILM","FLAT","MIXX","XRAY"))
    & (msqR$time < 3) ))

msq.film[((msq.film$study == "FLAT") & (msq.film$time ==3)) ,] <- NA
msq.drug <- subset(msqR,(is.element( msqR$study ,  c("AGES","SALT", "VALE", "XRAY")))
   &(msqR$time < 3))

msq.day <- subset(msqR,is.element( msqR$study ,  c("SAM", "RIM")))
```

---

neo   *NEO correlation matrix from the NEO_PI_R manual*

Description

The NEO.PI.R is a widely used personality test to assess 5 broad factors (Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness) with six facet scales for each factor. The correlation matrix of the facets is reported in the NEO.PI.R manual for 1000 subjects.

Usage

```
data(neo)
```

Format

A data frame of a 30 x 30 correlation matrix with the following 30 variables.

N1 Anxiety

N2 AngryHostility

N3 Depression

N4 Self-Consciousness

N5 Impulsiveness

N6 Vulnerability

E1 Warmth

E2 Gregariousness

E3 Assertiveness

E4 Activity

E5 Excitement-Seeking

E6 PositiveEmotions

O1 Fantasy

O2 Aesthetics

O3 Feelings

O4 Ideas

O5 Actions

O6 Values

A1 Trust

A2 Straightforwardness

A3 Altruism

A4 Compliance

A5 Modesty

A6 Tender-Mindedness

C1 Competence

C2 Order

C3 Dutifulness

C4 AchievementStriving

C5 Self-Discipline

C6 Deliberation

Details

The past thirty years of personality research has led to a general consensus on the identification of major dimensions of personality. Variously known as the "Big 5" or the "Five Factor Model", the general solution represents 5 broad domains of personal and interpersonal experience. Neuroticism and Extraversion are thought to reflect sensitivity to negative and positive cues from the environment and the tendency to withdraw or approach. Openness is sometimes labeled as Intellect and reflects an interest in new ideas and experiences. Agreeableness and Conscientiousness reflect tendencies to get along with others and to want to get ahead.

The factor structure of the NEO suggests five correlated factors as well as two higher level factors. The NEO was constructed with 6 "facets" for each of the five broad factors.

For a contrasting structure, examine the items of the link{spi} data set (Condon, 2017).

Source

Costa, Paul T. and McCrae, Robert R. (1992) (NEO PI-R) professional manual. Psychological Assessment Resources, Inc. Odessa, FL. (with permission of the author and the publisher)

References

Condon, D. (2017) The SAPA Personality Inventory:An empirically-derived, hierarchically-organized self-report personality assessment model

Digman, John M. (1990) Personality structure: Emergence of the five-factor model. Annual Review of Psychology. 41, 417-440.

John M. Digman (1997) Higher-order factors of the Big Five. Journal of Personality and Social Psychology, 73, 1246-1256.

McCrae, Robert R. and Costa, Paul T., Jr. (1999) A Five-Factor theory of personality. In Pervin, Lawrence A. and John, Oliver P. (eds) Handbook of personality: Theory and research (2nd ed.) 139-153. Guilford Press, New York. N.Y.

Revelle, William (1995), Personality processes, Annual Review of Psychology, 46, 295-328.

Joshua Wilt and William Revelle (2009) Extraversion and Emotional Reactivity. In Mark Leary and Rick H. Hoyle (eds). Handbook of Individual Differences in Social Behavior. Guilford Press, New York, N.Y.

Joshua Wil and William Revelle (2016) Extraversion. In Thomas Widiger (ed) The Oxford Handbook of the Five Factor Model. Oxford University Press.

Examples

```
data(neo)
n5 <- psych::fa(neo,5)
neo.keys <- psych::make.keys(30,list(N=c(1:6),E=c(7:12),O=c(13:18),A=c(19:24),C=c(25:30)))
n5p <- psych::target.rot(n5,neo.keys) #show a targeted rotation for simple structure
n5p
```

---

peas   *Galton's Peas*

Description

Francis Galton introduced the correlation coefficient with an analysis of the similarities of the parent and child generation of 700 sweet peas.

Usage

```
data(peas)
```

Format

A data frame with 700 observations on the following 2 variables.

parent The mean diameter of the mother pea for 700 peas

child The mean diameter of the daughter pea for 700 sweet peas

Details

Galton's introduction of the correlation coefficient was perhaps the most important contribution to the study of individual differences. This data set allows a graphical analysis of the data set. There are two different graphic examples. One shows the regression lines for both relationships, the other finds the correlation as well.

Source

Stanton, Jeffrey M. (2001) Galton, Pearson, and the Peas: A brief history of linear regression for statistics intstructors, Journal of Statistics Education, 9. (retrieved from the web from https://www.amstat.org/publications/jse. reproduces the table from Galton, 1894, Table 2.

The data were generated from this table.

References

Galton, Francis (1877) Typical laws of heredity. paper presented to the weekly evening meeting of the Royal Institution, London. Volume VIII (66) is the first reference to this data set. The data appear in

Galton, Francis (1894) Natural Inheritance (5th Edition), New York: MacMillan).

See Also

The other Galton data sets: `heights`, `galton`,`cubits`

Examples

```
data(peas)
psych::pairs.panels(peas,lm=TRUE,xlim=c(14,22),ylim=c(14,22),main="Galton's Peas")
psych::describe(peas)
psych::pairs.panels(peas,main="Galton's Peas")
```

---

psychTools	*psychTools: datasets and utility functions to accompany the psych package*

Description

PsychTools includes the larger data sets used by the `psych` package and also includes a few general utility functions such as the `read.file` and `read.clipboard` functions. The data sets ara made available for demonstrations of a variety of psychometric functions.

Details

See the various helpfiles listed in the index or as links from here. Also see the main functions in the psych package `00.psych-package`.

Data sets from the SAPA/ICAR project:

`ability`	16 ICAR ability items scored as correct or incorrect for 1525 participants.
`iqitems`	multiple choice IQ items (raw responses)
`affect`	Two data sets of affect and arousal scores as a function of personality and movie conditions
`bfi`	25 Personality items representing 5 factors from the SAPA project for 2800 participants
bfi.dictionary	Dictionary of the bfi
`epi`	Eysenck Personality Inventory (EPI) data for 3570 participants
epi.dictionary	The items for the epi
`epi.bfi`	13 personality scales from the Eysenck Personality Inventory and Big 5 inventory
`epiR`	474 participants took the epi twice
`msq`	75 mood items from the Motivational State Questionnaire for 3896 participants
`msqR`	75 mood items from the Motivational State Questionnaire for 3032 unique participants
`tai`	Trait Anxiety data from the PMC lab matching the sai sample. 3032 unique subjects
`sai`	State Anxiety data from the PMC lab over multiple occasions. 3032 unique subjects.
sai.dictionary	items used in the sai
`spi`	4000 cases from the SAPA Personality Inventory including an item dictionary and scoring keys.
spi.dictionary	The items for the spi
spi.keys	Scoring keys for the spi

Historically interesting data sets

burt	11 emotional variables from Burt (1915)
galton	Galtons Mid parent child height data
heights	A data.frame of the Galton (1888) height and cubit data set
cubits	Galtons example of the relationship between height and cubit or forearm length
peas	Galtons Peas
cushny	The data set from Cushny and Peebles (1905) on the effect of three drugs on hours of sleep, used by Student (1908

Miscellaneous example data sets

blant	A 29 x 29 matrix that produces weird factor analytic results
blot	Bonds Logical Operations Test - BLOT
cities	Distances between 11 US cities
city.location	and their geograpical location
income	US family income from US census 2008
all.income	US family income from US census 2008
neo	NEO correlation matrix from the NEO_PI_R manual
Schutz	The Schutz correlation matrix example from Shapiro and ten Berge
veg	Paired comparison of preferences for 9 vegetables (scaling example)

Functions to convert various objects to latex

fa2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table
df2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table
ICC2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table
irt2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table
cor2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table
omega2latex	Convert a data frame, correlation matrix, or factor analysis output to a LaTeX table

File manipulation functions

fileCreate	Create a file
fileScan	Show the first few lines of multitple files
filesInfo	Show the information for all files in a directory
filesList	Show the names of all files in a directory

dfOrder Sorts a data frame

File input/output functions

read.clipboard	Shortcuts for reading from the clipboard or a file
read.clipboard.csv	
read.clipboard.fwf	
read.clipboard.lower	
read.clipboard.tab	

                    read.clipboard.upper
                    read.file   Read a file according to its suffix
                    read.file.csv
                    read.https
                    write.file   Write data to a file
                    write.file.csv

Examples

```
psych::describe(ability)
```

---

read.file   *Shortcuts for reading from the clipboard or a file*

Description

Input from a variety of sources may be read. data.frames may be read from files with suffixes of .txt, .text, .TXT, .dat, .DATA,.data, .csv, .rds, rda, .xpt, or .sav (i.e., data from SPSS sav files may be easily read.) Data exported by JMP or EXCEL in the csv format are also able to be read. Fixed Width Files saved in .txt mode may be read if the widths parameter is specified. Files saved with writeRDS have suffixes of .rds or Rds, and are read using readRDS. Files associated with objects with suffixes .rda and .Rda are loaded. The default values for read.spss are adjusted for more standard input from SPSS files. Input from the clipboard is easy but a bit obscure, particularly for Mac users. read.clipboard and its variations are just an easier way to do so. Data may be copied to the clipboard from Excel spreadsheets, csv files, or fixed width formatted files and then into a data.frame. Data may also be read from lower (or upper) triangular matrices and filled out to square matrices. write.file will prompt for a file name (if not given) and then write or save to that file depending upon the suffix (text, txt, or csv will call write.table, R, or r will dput, rda, Rda will save, Rds,rds will saveRDS).

Usage

```
read.file(file=NULL,header=TRUE,use.value.labels=FALSE,to.data.frame=TRUE,sep=",",
widths=NULL,f=NULL, filetype=NULL,...)
 #for .txt, .text, TXT, .csv, .sav, .xpt, XPT,  R, r, Rds, .rds, or .rda,
# .Rda, .RData, .Rdata, .dat and .DAT  files

read.clipboard(header = TRUE, ...)   #assumes headers and tab or space delimited
read.clipboard.csv(header=TRUE,sep=',',...)   #assumes headers and comma delimited
read.clipboard.tab(header=TRUE,sep='\t',...)   #assumes headers and tab delimited
 #read in a matrix given the lower off diagonal
 read.clipboard.lower(diag=TRUE,names=FALSE,...)
 read.clipboard.upper(diag=TRUE,names=FALSE,...)

#read in data using a fixed format width (see read.fwf for instructions)
read.clipboard.fwf(header=FALSE,widths=rep(1,10),...)
```

```
read.https(filename,header=TRUE)

read.file.csv(file=NULL,header=TRUE,f=NULL,...)
write.file(x,file=NULL,row.names=FALSE,f=NULL,...)
write.file.csv(x,file=NULL,row.names=FALSE,f=NULL,...)
```

Arguments

header	Does the first row have variable labels (generally assumed to be TRUE).
sep	What is the designated separater between data fields? For typical csv files, this will be a comma, but if commas designate decimals, then a ; can be used to designate different records.
diag	for upper or lower triangular matrices, is the diagonal specified or not
names	for read.clipboard.lower or upper, are colnames in the the first column
widths	how wide are the columns in fixed width input. The default is to read 10 columns of size 1.
filename	Name or address of remote https file to read.
...	Other parameters to pass to read
f	A file name to read from or write to. If omitted, `file.choose` is called to dynamically get the file name.
file	A file name to read from or write to. (same as f, but perhaps more intuitive) If omitted and if f is omitted,then `file.choose` is called to dynamically get the file name.
x	The data frame or matrix to write to f
row.names	Should the output file include the rownames? By default, no.
to.data.frame	Should the spss input be converted to a data frame?
use.value.labels	
	Should the SPSS input values be converted to numeric?
filetype	If specified the reading will use this term rather than the suffix.

Details

A typical session of R might involve data stored in text files, generated online, etc. Although it is easy to just read from a file (particularly if using `file.choose`, copying from the file to the clipboard and then reading from the clipboard is also very convenient (and somewhat more intuitive to the naive user). This is particularly convenient when copying from a text book or article and just moving a section of text into R.)

The `read.file` function combines the `file.choose` and either `read.table`, `read.fwf`, `read.spss` or `read.xport`(from foreign) or `load` or `readRDS` commands. By examining the file suffix, it chooses the appropriate way to read. For more complicated file structures, see the foreign package. For even more complicated file structures, see the rio or haven packages.

Note that `read.file` assumes by default that the first row has column labels (header =TRUE). If this is not true, then make sure to specify header = FALSE. If the file is fixed width, the assumption

is that it does not have a header field. In the unlikely case that a fwf file does have a header, then you probably should try fn <- file.choose() and then my.data <- read.fwf(fn,header=TRUE,widths= widths)

Further note: If the file is a .Rda, .rda, etc. file, the read.file command will load this file and return the name of the file. In this case, it is necessary to either assign the output (the file name) to an object that has a different name than any of the objects in the file, or to call read.file() without any specification.

If the file has no suffix the default action is to quit with a warning. However, if the filetype is specified, it will use that type in the reading (e.g. filetype="txt" will read as text file, even if there is no suffix.)

If the file is specified and has a prefix of http:// https:// it will be downloaded and then read.

Currently supported input formats are

.sav	SPSS.sav files
.csv	A comma separated file (e.g. from Excel or Qualtrics)
.txt	A typical text file
.TXT	A typical text file
.text	A typical text file
.data	A data file
.dat	A data file
.rds	A R data file
.Rds	A R data file (created by a write)
.Rda	A R data structure (created using save)
.rda	A R data structure (created using save)
.RData	A R data structure (created using save)
.rdata	A R data structure (created using save)
.R	A R data structure created using dput
.r	A R data structure created using dput
.xpt	A SAS data file in xport format
.XPT	A SAS data file in XPORT format

The foreign function read.spss is used to read SPSS .sav files using the most common options. Just as read.spss issues various warnings, so does read.file. In general, these can be ignored. For more detailed information about using read.spss, see the help pages in the foreign package.

If you have a file written by JMP, you must first export to a csv or text file.

The write.file function combines the file.choose and either write.table or saveRDS. By examining the file suffix, it chooses the appropriate way to write. For more complicated file structures, see the foreign package, or the save function in R Base. If no suffix is added, it will write as a .txt file. write.file.csv will write in csv format to an arbitrary file name.

Currently supported output formats are

.csv	A comma separated file (e.g. for reading into Excel)
.txt	A typical text file
.text	A typical text file
.rds	A R data file

.Rds	A R data file (created by a write)
.Rda	A R data structure (created using save)
.rda	A R data structure (created using save)
.R	A R data structure created using dput
.r	A R data structure created using dput

Note that new=TRUE option in write.file works only in R.app and not in RStudio. To create a new file using RStudio (or on a PC) you can use the link{fileCreate} function first.

read.clipboard was based upon a suggestion by Ken Knoblauch to the R-help listserve.

If the input file that was copied into the clipboard was an Excel file with blanks for missing data, then read.clipboard.tab() will correctly replace the blanks with NAs. Similarly for a csv file with blank entries, read.clipboard.csv will replace empty fields with NA.

read.clipboard.lower and read.clipboard.upper are adapted from John Fox's read.moments function in the sem package. They will read a lower (or upper) triangular matrix from the clipboard and return a full, symmetric matrix for use by factanal, fa , ICLUST, pca. omega , etc. If the diagonal is false, it will be replaced by 1.0s. These two function were added to allow easy reading of examples from various texts and manuscripts with just triangular output.

Many articles will report lower triangular matrices with variable labels in the first column. read.clipboard.lower will handle this case. Names must be in the first column if names=TRUE is specified.

Other articles will report upper triangular matrices with variable labels in the first row. read.clipboard.upper will handle this. Note that labels in the first column will not work for read.clipboard.upper. The names, if present, must be in the first row.

read.clipboard.fwf will read fixed format files from the clipboard. It includes a patch to read.fwf which will not read from the clipboard or from remote file. See read.fwf for documentation of how to specify the widths.

Value

the contents of the file to be read or of the clipboard.

Author(s)

William Revelle

Examples

```
#All of these functions are meant for interactive Input
#Because these are dynamic functions, they need to be run interactively and
# can not be shown as examples.
#Thus they are not to be tested by CRAN

if(interactive()) {
 my.data <- read.file()  #search the directory for a file and then read it.
                         #return the result into an object
#or, if the file is a rda, etc. file
my.data <- read.file()  #return the path and instructions of how to load
  # without assigning a value.
```

```
filesList() #search the system for a particular file and then list all the files in that directory
fileCreate() #search for a particular directory and create a file there.
write.file(Thurstone) #open the search window, choose a location and name the output file,
# write the data file (e.g., Thurstone ) to the file chosen

#the example data set from read.delim in the readr package to read a remote csv file
my.data <-read.file("https://github.com/tidyverse/readr/raw/master/inst/extdata/mtcars.csv")


#These functions read from the local clipboard and thus are interactive
my.data <- read.clipboard()   #space delimited columns
my.data <- read.clipboard.csv()  # , delimited columns
my.data <- read.clipboard.tab()  #typical input if copied from a spreadsheet
my.data <- read.clipboad(header=FALSE)  #data start on line 1
my.matrix <- read.clipboard.lower()
}
```

---

sai   *State Anxiety data from the PMC lab over multiple occasions.*

Description

State Anxiety was measured two-three times in 11 studies at the Personality-Motivation-Cognition laboratory. Here are item responses for 11 studies (9 repeated twice, 2 repeated three times). In all studies, the first occasion was before a manipulation. In some studies, caffeine, or movies or incentives were then given to some of the participants before the second and third STAI was given. In addition, Trait measures are available and included in the tai data set (3032 subjects).

Usage

```
data(sai)
data(tai)
data(sai.dictionary)
```

Format

A data frame with 3032 unique observations on the following 23 variables.

id a numeric vector

study a factor with levels ages cart fast fiat film flat home pat rob salt shedshop xray

time 1=First, 2 = Second, 3=third administration

TOD TOD (time of day 1= 8:50-9:30 am,2 = 1=3 pm, 3= 7:-8pm

drug drug (placebo (0) vs. caffeine (1))

film film (1=Frontline (concentration camp), 2 = Halloween 3= National Geographic (control), 4-Parenthood (humor)

anxious anxious

at.ease at ease

calm calm

comfortable comfortable

confident confident

content content

high.strung high.strung

jittery jittery

joyful joyful

nervous nervous

pleasant pleasant

rattled over-excited and rattled

regretful regretful

relaxed relaxed

rested rested

secure secure

tense tense

upset upset

worried worried

worrying worrying

Details

The standard experimental study at the Personality, Motivation and Cognition (PMC) laboratory (Revelle and Anderson, 1997) was to administer a number of personality trait and state measures (e.g. the epi, msq, msqR and sai) to participants before some experimental manipulation of arousal/effort/anxiety. Following the manipulation (with a 30 minute delay if giving caffeine/placebo), some performance task was given, followed once again by measures of state arousal/effort/anxiety.

Here are the item level data on the sai (state anxiety) and the tai (trait anxiety). Scores on these scales may be found using the scoring keys. The affect data set includes pre and post scores for two studies (flat and maps) which manipulated state by using four types of movies.

In addition to being useful for studies of motivational state, these studies provide examples of test-retest and alternate form reliabilities. Given that 10 items overlap with the msqR data, they also allow for a comparison of immediate duplication of items with 30 minute delays.

Studies CART, FAST, SHED, RAFT, and SHOP were either control groups, or did not experimentally vary arousal/effort/anxiety.

AGES, CITY, EMIT, RIM, SALT, and XRAY were caffeine manipulations between time 1 and 2 (RIM and VALE were repeated day 1 and day 2)

FIAT, FLAT, MAPS, MIXX, and THRU were 1 day studies with film manipulation between time 1 and time 2.

SAM1 and SAM2 were the first and second day of a two day study. The STAI was given once per day. MSQ not MSQR was given.

VALE and PAT were two day studies with the STAI given pre and post on both days

RIM was a two day study with the STAI and MSQ given once per day.

Usually, time of day 1 = 8:50-9am am, and 2 = 7:30 pm, however, in rob, with paid subjects, the times were 0530 and 22:30.

Source

Data collected at the Personality, Motivation, and Cognition Laboratory, Northwestern University, between 1991 and 1999.

References

Charles D. Spielberger and Richard L. Gorsuch and R. E. Lushene, (1970) Manual for the State-Trait Anxiety Inventory.

Revelle, William and Anderson, Kristen Joan (1997) Personality, motivation and cognitive performance: Final report to the Army Research Institute on contract MDA 903-93-K-0008

Rafaeli, Eshkol and Revelle, William (2006), A premature consensus: Are happiness and sadness truly opposite affects? Motivation and Emotion, 30, 1, 1-12.

Smillie, Luke D. and Cooper, Andrew and Wilt, Joshua and Revelle, William (2012) Do Extraverts Get More Bang for the Buck? Refining the Affective-Reactivity Hypothesis of Extraversion. Journal of Personality and Social Psychology, 103 (2), 206-326.

Examples

```
data(sai)

table(sai$study,sai$time)  #show the counts for repeated measures

#Here are the keys to score the sai total score, positive and negative items
sai.keys <- list(sai = c("tense","regretful" , "upset", "worrying", "anxious", "nervous" ,
"jittery" , "high.strung", "worried" , "rattled","-calm",
"-secure","-at.ease","-rested","-comfortable", "-confident" ,"-relaxed" , "-content" ,
"-joyful", "-pleasant"  ) ,
sai.p = c("calm","at.ease","rested","comfortable", "confident", "secure" ,"relaxed" ,
        "content" , "joyful", "pleasant" ),
sai.n = c( "tense" , "anxious", "nervous" , "jittery" , "rattled",     "high.strung",
         "upset", "worrying","worried","regretful" )
)

#using the is.element function instead of the %in% function
#just get the control subjects
control <- subset(sai,is.element(sai$study,c("Cart", "Fast", "SHED", "RAFT", "SHOP")) )

#pre and post drug studies
drug <- subset(sai,is.element(sai$study, c("AGES","CITY","EMIT","SALT","VALE","XRAY")))

#pre and post film studies
```

```
film <- subset(sai,is.element(sai$study, c("FIAT","FLAT", "MAPS", "MIXX") ))

#this next set allows us to score those sai items that overlap with the msq item sets
msq.items <- c("anxious", "at.ease" ,"calm", "confident","content", "jittery",
 "nervous" ,  "relaxed" ,  "tense"  ,  "upset" ) #these overlap with the msq

sai.msq.keys <- list(pos =c( "at.ease" ,  "calm" , "confident", "content","relaxed"),
  neg = c("anxious", "jittery", "nervous" ,"tense"  ,   "upset"),
  anx = c("anxious", "jittery", "nervous" ,"tense", "upset","-at.ease" ,  "-calm" ,
  "-confident", "-content","-relaxed"))
sai.not.msq.keys <- list(pos=c(  "secure","rested","comfortable" ,"joyful" , "pleasant" ),
    neg=c("regretful","worrying", "high.strung","worried", "rattled" ),
    anx = c("regretful","worrying", "high.strung","worried", "rattled",    "-secure",
    "-rested", "-comfortable", "-joyful",  "-pleasant" ))
sai.alternate.forms <- list( pos1 =c( "at.ease","calm","confident","content","relaxed"),
  neg1 = c("anxious", "jittery", "nervous" ,"tense"  ,   "upset"),
  anx1 = c("anxious", "jittery", "nervous" ,"tense", "upset","-at.ease" ,  "-calm" ,
       "-confident", "-content","-relaxed"),
  pos2=c(  "secure","rested","comfortable" ,"joyful" , "pleasant" ),
  neg2=c("regretful","worrying", "high.strung","worried", "rattled" ),
  anx2 = c("regretful","worrying", "high.strung","worried", "rattled", "-secure",
    "-rested", "-comfortable", "-joyful",  "-pleasant" ))

sai.repeated <- c("AGES","Cart","Fast","FIAT","FILM","FLAT","HOME","PAT","RIM","SALT",
    "SAM","SHED","SHOP","VALE","XRAY")
sai12 <- subset(sai,is.element(sai$study,  sai.repeated))  #the subset with repeated measures

#Choose those studies with repeated measures by :
sai.control <- subset(sai,is.element(sai$study, c("Cart", "Fast", "SHED", "SHOP")))
sai.film <- subset(sai,is.element(sai$study, c("FIAT","FLAT") )  )
sai.drug <- subset(sai,is.element(sai$study, c("AGES",  "SALT", "VALE", "XRAY")))
sai.day <- subset(sai,is.element(sai$study, c("SAM", "RIM")))
```

Schutz	*The Schutz correlation matrix example from Shapiro and ten Berge*

Description

Shapiro and ten Berge use the Schutz correlation matrix as an example for Minimum Rank Factor Analysis. The Schutz data set is also a nice example of how normal minres or maximum likelihood will lead to a Heywood case, but minrank factoring will not.

Usage

```
data("Schutz")
```

Format

The format is: num [1:9, 1:9] 1 0.8 0.28 0.29 0.41 0.38 0.44 0.4 0.41 0.8 ... - attr(*, "dim-names")=List of 2 ..$:1] "Word meaning" "Odd Words" "Boots" "Hatchets" ... ..$: chr [1:9] "V1" "V2" "V3" "V4" ...

Details

These are 9 cognitive variables of importance mainly because they are used as an example by Shapiro and ten Berge for their paper on Minimum Rank Factor Analysis.

The solution from the [fa](#) function with the fm='minrank' option is very close (but not exactly equal) to their solution.

This example is used to show problems with different methods of factoring. Of the various factoring methods, fm = "minres", "uls", or "mle" produce a Heywood case. Minrank, alpha, and pa do not.

See the blant data set for another example of differences across methods.

Source

Richard E. Schutz,(1958) Factorial Validity of the Holzinger-Crowdeer Uni-factor tests. Educational and Psychological Measurement, 48, 873-875.

References

Alexander Shapiro and Jos M.F. ten Berge (2002) Statistical inference of minimum rank factor analysis. Psychometrika, 67. 70-94

Examples

```
data(Schutz)
psych::corPlot(Schutz,numbers=TRUE,upper=FALSE)

f4min <- psych::fa(Schutz,4,fm="minrank")  #for an example of minimum rank factor Analysis
#compare to
f4 <- psych::fa(Schutz,4,fm="mle")  #for the maximum likelihood solution which has a Heywood case
```

---

spi	*A sample from the SAPA Personality Inventory including an item dictionary and scoring keys.*

Description

The SPI (SAPA Personality Inventory) is a set of 135 items primarily selected from International Personality Item Pool (ipip.ori.org). This is an example data set collected using SAPA procedures the sapa-project.org web site. This data set includes 10 demographic variables as well. The data set with 4000 observations on 145 variables may be used for examples in scale construction and validation, as well as empirical scale construction to predict multiple criteria.

Usage

```
data("spi")
data(spi.dictionary)
data(spi.keys)
```

Format

A data frame with 4000 observations on the following 145 variables. (The q numbers are the SAPA item numbers).

age Age in years from 11 -90

sex Reported biological sex (coded by X chromosones => 1=Male, 2 = Female)

health Self rated health 1-5: poor, fair, good, very good, excellent

p1edu Parent 1 education

p2edu Parent 2 education

education Respondents education: less than 12, HS grad, current univ, some univ, associate degree, college degree, in grad/prof, grad/prof degree

wellness Self rated "wellnes" 1-2

exer Frequency of exercise: very rarely, < 1/month, < 1/wk, 1 or 2 times/week, 3-5/wk, > 5 times/week

smoke never, not last year, < 1/month, <1/week, 1-3 days/week, most days, up to 5 x /day, up to 20 x /day, > 20x/day

ER Emergency room visits none, 1x, 2x, 3 or more times

q_253 see the spi.dictionary for these items (q_253

q_1328 see the dictionary for all items q_1328)

Details

Using the data contributed by about 125,000 visitors to the https://SAPA-project.org website, David Condon has developed a hierarchical framework for assessing personality at two levels. The higher level has the familiar five factors that have been studied extensively in personality research since the 1980s – Conscientiousness, Agreeableness, Neuroticism, Openness, and Extraversion. The lower level has 27 factors that are considerably more narrow. These were derived based on administrations of about 700 public-domain IPIP items to 3 large samples. Condon describes these scales as being "empirically-derived" because relatively little theory was used to select the number of factors in the hierarchy and the items in the scale for each factor (to be clear, he means relatively little personality theory though he relied on quite a lot of sampling and statistical theory). You can read all about the procedures used to develop this framework in his book/manual. If you would like to reproduce these analyses, you can download the data files from Dataverse (links are also provided in the manual) and compile this script in R (he used knitR). Instructions are provided in the Preface to the manual.

This small subset of the data is provided for demonstration purposes.

Source

https://sapa-project.org/research/SPI/SPIdevelopment.pdf.

References

Condon, D. (2017) The SAPA Personality Inventory:An empirically-derived, hierarchically-organized self-report personality assessment model

Examples

```
data(spi)
data(sapa.dictionary)
psych::bestScales(spi, criteria="health",dictionary=spi.dictionary)

sc <- psych::scoreVeryFast(spi.keys,spi) #much faster scoring for just scores
sc <- psych::scoreItems(spi.keys,spi)  #gives the alpha reliabilities and various stats

psych::corPlot(sc$corrected,numbers=TRUE,cex=.4,xlas=2,min.length=6,
     main="Structure of SPI (disattenuated r above the diagonal)")
```

Utility   *Useful utility functions for file/directory exploration and manipulation.*

Description

Wrappers for dirname, file.choose, readLines. file.create, file.path to be called directly for listing directories, creating files, showing the files in a directory, and listing the content of files in a directory. `fileCreate` gives the functionality of `file.choose`(new=TRUE). `filesList` combines file.choose, dirname, and list.files to show the files in a directory, `fileScan` extends this and then returns the first few lines of each readable file

Usage

```
fileScan(f = NULL, nlines = 3, max = NULL, from = 1, filter = NULL)
filesList(f=NULL)
filesInfo(f=NULL,max=NULL)
fileCreate(newName="new.file")
```

Arguments

f	File path to use as base path (will use file.choose() if missing. If f is a directory, will list the files in that directory, if f is a file, will find the directory for that file and then list all of those files.)
nlines	How many lines to display
max	maximum number of files to display
from	First file (number) to display
filter	Just display files with "filter" in the name
newName	The name of the file to be created.

Details

Just a collection of simple wrappers to powerful core R functions. Allows the user more direct control of what directory to list, to create a file, or to display the content of files. The functions called include file.choose, file.path, file.info,file.create, dirname, and dir.exists. All of these are very powerful functions, but not easy to call interactively.

fileCreate will ask to locate a file using file.choose, set the directory to that location, and then prompt to create a file with the new.name. This is a workaround for file.choose(new=TRUE) which only works for Macs not using R.studio.

filesInfo will interactively search for a file and then list the information (size, date, ownership) of all the files in that directory.

filesList will interactively search for a file and then list all the files in same directory.

Note

Work arounds for core-R functions for interactive file manipulation

Author(s)

William Revelle

See Also

read.file to read in data from a file or read.clipboard from the clipboard. dfOrder to sort data.frames.

Examples

```
if(interactive()) {
#all of these require interactive input and thus are not given as examples

fileCreate("my.new.file.txt")
filesList()   #show the items in the directory where a file is displayed
fileScan() #show the content of the files in a directory
#or, if you have a file in mind
 f <- file.choose()  #go find it
filesList(f)
fileScan(f)
}
```

---

vegetables   *Paired comparison of preferences for 9 vegetables*

Description

A classic data set for demonstrating Thurstonian scaling is the preference matrix of 9 vegetables from Guilford (1954). Used by Guiford, Nunnally, and Nunally and Bernstein, this data set allows for examples of basic scaling techniques.

Usage

```
data(vegetables)
```

Format

A data frame with 9 choices on the following 9 vegetables. The values reflect the perecentage of times where the column entry was preferred over the row entry.

Turn Turnips

Cab Cabbage

Beet Beets

Asp Asparagus

Car Carrots

Spin Spinach

S.Beans String Beans

Peas Peas

Corn Corn

Details

Louis L. Thurstone was a pioneer in psychometric theory and measurement of attitudes, interests, and abilities. Among his many contributions was a systematic analysis of the process of comparative judgment (thurstone, 1927). He considered the case of asking subjects to successively compare pairs of objects. If the same subject does this repeatedly, or if subjects act as random replicates of each other, their judgments can be thought of as sampled from a normal distribution of underlying (latent) scale scores for each object, Thurstone proposed that the comparison between the value of two objects could be represented as representing the differences of the average value for each object compared to the standard deviation of the differences between objects. The basic model is that each item has a normal distribution of response strength and that choice represents the stronger of the two response strengths. A justification for the normality assumption is that each decision represents the sum of many independent inputs and thus, through the central limit theorem, is normally distributed.

Thurstone considered five different sets of assumptions about the equality and independence of the variances for each item (Thurston, 1927). Torgerson expanded this analysis slightly by considering three classes of data collection (with individuals, between individuals and mixes of within

and between) crossed with three sets of assumptions (equal covariance of decision process, equal correlations and small differences in variance, equal variances).

This vegetable data set is used by Guilford and by Nunnally to demonstrate Thurstonian scaling.

Source

Guilford, J.P. (1954) Psychometric Methods. McGraw-Hill, New York.

References

Nunnally, J. C. (1967). Psychometric theory., McGraw-Hill, New York.

Revelle, W. An introduction to psychometric theory with applications in R. (in preparation), Springer. https://personality-project.org/r/book

See Also

thurstone

Examples

```
data(vegetables)
psych::thurstone(veg)
```

Index