Package: prcr (via r-universe)

September 15, 2024

Type Package

Title Person-Centered Analysis

Version 0.2.1

Maintainer Joshua M Rosenberg <`jmichaelrosenberg@gmail.com`>

Description Provides an easy-to-use yet adaptable set of tools to conduct person-center analysis using a two-step clustering procedure. As described in Bergman and El-Khouri (1999) <DOI:10.1002/(SICI)1521-4036(199910)41:6%3C753::AID-BIMJ753%3E3.0.CO;2-K>, hierarchical clustering is performed to determine the initial partition for the subsequent k-means clustering procedure.

License MIT + file LICENSE

URL `https://github.com/jrosen48/prcr`

BugReports `https://github.com/jrosen48/prcr/issues`

LazyData TRUE

Imports dplyr, tidyr, ggplot2, tibble, irr, lpSolve, purrr, class, forcats, magrittr

Suggests rmarkdown, knitr, devtools

VignetteBuilder knitr

RoxygenNote 7.0.2

Depends R (>= 2.10)

Repository https://jrosen48.r-universe.dev

RemoteUrl https://github.com/jrosen48/prcr

RemoteRef HEAD

RemoteSha 95db5c839ddfc1f440a78e7776bb5e63c6abf9b5

Contents

`create_profiles_cluster`

Create profiles of observed variables using two-step cluster analysis

---

Description

Create profiles of observed variables using two-step cluster analysis

Usage

```
create_profiles_cluster(
  df,
  ...,
  n_profiles,
  to_center = FALSE,
  to_scale = FALSE,
  distance_metric = "squared_euclidean",
  linkage = "complete"
)
```

Arguments

`df`	with two or more columns with continuous variables
`...`	unquoted variable names separated by commas
`n_profiles`	The specified number of profiles to be found for the clustering solution
`to_center`	Boolean (TRUE or FALSE) for whether to center the raw data with M = 0
`to_scale`	Boolean (TRUE or FALSE) for whether to scale the raw data with SD = 1
`distance_metric`	Distance metric to use for hierarchical clustering; "squared_euclidean" is default but more options are available (see ?hclust)
`linkage`	Linkage method to use for hierarchical clustering; "complete" is default but more options are available (see ?dist)

Details

Function to create a specified number of profiles of observed variables using a two-step (hierarchical and k-means) cluster analysis.

Value

A list containing the prepared data, the output from the hierarchical and k-means cluster analysis, the r-squared value, raw clustered data, processed clustered data of cluster centroids, and a ggplot object.

Examples

```
d <- pisaUSA15
m3 <- create_profiles_cluster(d,
                              broad_interest, enjoyment, instrumental_mot, self_efficacy,
                               n_profiles = 3)
summary(m3)
```

`detect_outliers` *Identifies potential outliers*

Description

Identifies potential outliers

Usage

```
detect_outliers(df, return_index = TRUE)
```

Arguments

`df`	data.frame (or tibble) with variables to be clustered; all variables must be complete cases
`return_index`	Boolean (TRUE or FALSE) for whether to return only the row indices of the possible multivariate outliers; if FALSE, then all of the output from the function (including the indices) is returned

Details

* add an argument to ‘create_profiles_cluster()‘ to remove multivariate outliers based on Hadi’s (1994) procedure

Value

either the row indices of possible multivariate outliers or all of the output from the function, depending on the value of return_index

---

`estimate_r_squared` *Estimates R^2 (r-squared) values for a range of number of profiles*

Description

Estimates R^2 (r-squared) values for a range of number of profiles

Usage

```
estimate_r_squared(
  df,
  ...,
  to_center = FALSE,
  to_scale = FALSE,
  distance_metric = "squared_euclidean",
  linkage = "complete",
  lower_bound = 2,
  upper_bound = 9,
  r_squared_table = TRUE
)
```

Arguments

`df`	with two or more columns with continuous variables
`...`	unquoted variable names separated by commas
`to_center`	(TRUE or FALSE) for whether to center the raw data with M = 0
`to_scale`	Boolean (TRUE or FALSE) for whether to scale the raw data with SD = 1
`distance_metric`	Distance metric to use for hierarchical clustering; "squared_euclidean" is default but more options are available (see ?hclust)
`linkage`	Linkage method to use for hierarchical clustering; "complete" is default but more options are available (see ?dist)
`lower_bound`	the smallest number of profiles in the range of number of profiles to explore; defaults to 2
`upper_bound`	the largest number of profiles in the range of number of profiles to explore; defaults to 9
`r_squared_table`	if TRUE (default), then a table, rather than a plot, is returned; defaults to FALSE

Details

Returns ggplot2 plot of cluster centroids

Value

A list containing a ggplot2 object and a tibble for the R^2 values

---

`pisaUSA15` *student questionnaire data with four variables from the 2015 PISA for students in the United States*

Description

student questionnaire data with four variables from the 2015 PISA for students in the United States

Usage

```
pisaUSA15
```

Format

Data frame with columns #'

CNTSTUID international student ID
SCHID international school ID ...

Source

http://www.oecd.org/pisa/data/

`plot_profiles` *Return plot of profile centroids*

Description

Return plot of profile centroids

Usage

```
plot_profiles(d, to_center = F, to_scale = F)
```

Arguments

`d`	summary data.frame output from create_profiles_cluster()
`to_center`	whether to center the data before plotting
`to_scale`	whether to scale the data before plotting

Details

Returns ggplot2 plot of cluster centroids

Value

A ggplot2 object

`print.prcr` *Prints details of prcr cluster solution*

Description

Prints details of prcr cluster solution

Usage

```
## S3 method for class 'prcr'
print(x, ...)
```

Arguments

`x`	A ‘prcr‘ object
`...`	Additional arguments

Details

Prints details of of prcr cluster solution

`summary.prcr` *Concise summary of prcr cluster solution*

Description

Concise summary of prcr cluster solution

Usage

```
## S3 method for class 'prcr'
summary(object, ...)
```

Arguments

`object`	A ‘prcr‘ object
`...`	Additional arguments

Details

Prints a concise summary of prcr cluster solution

Index