Package: scagnostics (via r-universe)

September 5, 2024

Version 0.2-6

Title Compute scagnostics - scatterplot diagnostics

Author Lee Wilkinson <leland.wilkinson@gmail.com> and Anushka Anand

Maintainer Simon Urbanek <simon.urbanek@r-project.org>

Depends rJava

Description Calculates graph theoretic scagnostics. Scagnostics describe various measures of interest for pairs of variables, based on their appearance on a scatterplot. They are useful tool for discovering interesting or unusual scatterplots from a scatterplot matrix, without having to look at every individual plot.

License BSD_2_clause + file LICENSE

URL https://www.rforge.net/scagnostics/, https://www.cs.uic.edu/~wilkinson/

Repository https://s-u.r-universe.dev

RemoteUrl https://github.com/s-u/scagnostics

RemoteRef HEAD

RemoteSha e36116b82c9faa7c4602bf58b1ad985fa800fd57

Contents

`scagnostics` *Calculcate scagnostics for pairs of variables*

Description

Scagnostics (scatterplot diagnostics) summarize potentially interesting patterns in 2d scatterplots.

Usage

```
scagnostics(x, ...)
scagnosticsOutliers(scagnostics)
scagnosticsExemplars(scagnostics)
scagnosticsGrid(scagnostics)
```

Arguments

`x`	object to calculate scagnostics on: a vector, a matrix or a data.frame
`...`	...
`scagnostics`	objects returned from the `scagnostics` function

Value

`scagnostics` returns a vector (for a pair) or a matrix consisting of scagnostics (rows) by variable pairs (columns).

`scagnostics.outliers` and `scagnostics.exemplars` return a logical vector.

`scagnostics.grid` returns a data frame with columns `x` and `y` which as a pair define the index of variables corresponding to the entries in the scagnostics matrix. Hence it has as many rows as there are columns in the scagnostics matrix and each row defines one pair of variables.

Author(s)

Lee Wilkinson <leland.wilkinson@gmail.com>, Anushka Anand (University of Illinois at Chicago), R package: Simon Urbanek <simon.urbanek@r-project.org>

References

Wilkinson L., Anand, A., and Grossman, R. (2006). *High-Dimensional visual analytics: Interactive exploration guided by pairwise views of point distributions*. IEEE Transactions on Visualization and Computer Graphics, November/December 2006 (Vol. 12, No. 6) pp. 1363-1372.

Wilkinson L., Anand, A., and Grossman, R. (2005). *Graph-Theoretic Scagnostics*. Proceedings of the 2005 IEEE Symposium on Information Visualization, p. 21.

https://www.cs.uic.edu/~wilkinson/

Examples

```
# tesing various forms
scagnostics(1:10, 1:10)
scagnostics(rnorm(100), rnorm(100))
scagnostics(as.matrix(mtcars))

# more real use on a dataset
s <- scagnostics(mtcars)

# look at outliers
o <- scagnosticsOutliers(s)
o[o]
# one outlier, let's plot it
g <- scagnosticsGrid(s)
go <- g[o,]
plot(mtcars[[go$x]], mtcars[[go$y]], pch=19,
     xlab=names(mtcars)[go$x], ylab=names(mtcars)[go$y])

# find and plot exemplars
e <- scagnosticsExemplars(s)
e[e]
ge <- g[e,]
par(mfrow = c(2,2))
for (i in 1:dim(ge)[1])
  plot(mtcars[[ge$x[i]]], mtcars[[ge$y[i]]], pch=19,
       xlab=names(mtcars)[ge$x[i]], ylab=names(mtcars)[ge$y[i]])
```

Index