Package ‘barcode’

February 19, 2015

Version 1.1
Date 2012-02-19
Title Barcode distribution plots
Author John W. Emerson and Walton A. Green and John A. Hartigan
Maintainer John W. Emerson <john.emerson@yale.edu>
Depends grid, lattice
Enhances gpairs
Description This package includes the function \{code\{barcode\}\}, which produces a histogram-like plot of a distribution that shows granularity in the data.
License GPL (>= 2)
Copyright (C) 2012 John W. Emerson and Walton A. Green and John A. Hartigan
Repository CRAN
Date/Publication 2012-02-20 15:08:35
NeedsCompilation no

R topics documented:

barcode ..................................................... 1
NewHavenResidential ...................................... 4

Index 6

barcode  Barcode plots

Description

Produce barcode plot(s) of the given (grouped) values.
Usage

barcode(x, outer.margins = list(bottom = unit(2, "lines"),
                            left = unit(2, "lines"),
                            top = unit(2, "lines"),
                            right = unit(2, "lines")),
        horizontal = TRUE, xlim = NULL, nint = 0, main = "", xlab = "",
        labelloc = TRUE, axisloc = TRUE, labelouter = FALSE,
        newpage = TRUE, fontsize = 9, psrtsize = unit(0.25, "char"),
        ptpch = 1, bcspace = NULL, use.points = FALSE, buffer = 0.02,
        log = FALSE, outerbox = TRUE)

barcode.panel(x, horizontal = TRUE, xlim = NULL, labelloc = TRUE, axisloc = TRUE,
               labelouter = FALSE, nint = 0, fontsize = 9,
               psrtsize = unit(0.25, "char"), ptpch = 1, bcspace = NULL,
               xlab = "", xlaboffset = unit(2.5, "lines"),
               use.points = FALSE, buffer = 0.02, log = FALSE)

Arguments

x a vector of values for which the barcode is desired, or a list of such vectors for
"side-by-side" barcodes. Matrices are coerced to data frames and treated as lists
NA's are allowed in the data.

outer.margins a list of length 4 with units as components named bottom, left, top, and right,
giving the outer margins. Defaults to two lines of text.

horizontal logical indicating the barcode orientation; the default, TRUE, produces horizontal
barcodes.

xlim the x limits (xmin, xmax) of the plot; the default, NULL, uses the range of the
full data, range(unlist(x)), plus the multiplicative buffer.

nint default, 0, uses no "binning"— i.e., the barcode presents the exact measure-
ments, to the precision of the data set; nint=100 uses roughly 100 "bins" in
constructing the barcode; fewer bins give a more histogram-like plot.

main the plot title.

xlab the axis label for the quantitative measurements.

labelloc for the location of the factor labels of the barcodes; default TRUE may also be
specified as 'left' or 'top' (having similar results but relating to the horizontal
alignment); values 'right' or 'bottom' are available as alternatives to FALSE.

axisloc for the location of the quantitative axis labels; default, TRUE, may also be spec-
ified as 'left' or 'top' (having similar results but relating to the horizontal
alignment); values 'right' or 'bottom' are available as alternatives to FALSE.

labelouter default, FALSE, positions all labels within the viewport; TRUE forces the barcodes
to the edge of the viewport, with the labels outside the viewport. May be of use
to advanced users.

newpage default, TRUE, creates the barcodes in a new graphics device instead of adding
the plot to the current viewport.

fontsize for the size of the axis and factor labels.
barcode

ptsize for the size of the plotted points.
ptpch for the type of plotted points.
bcpspace indicates the proportion of total available space occupied by the barcode part of
the displays. Can range from 0 to 1; reasonable values seem to be between 0.1 and 0.5.
use.points default FALSE uses segments instead of points in the histogram-style display.
xlab.offset used for tuning the position of the label of the quantitative variable; needs to be
a unit.
buffer an additional proportion of empty space added to the right and left of the bar-
code, to avoid having the maximum and minimum on the frame of the plot.
log if TRUE, use the log scale for the y-axis of the histogram-like part of the barcodes.
outerbox if TRUE, plot a box around the display.

Details

The barcode plot aids in comparing distributions. It shares some of the characteristics of side-by-
side histograms or boxplots, and of rugs or stripplots. We have found it particularly useful with
clumped data, when other methods obscure detail.

Note

John Hartigan designed and implemented an early version of the barcode plot. The implementa-
tion provided here uses grid graphics, adds some useful options, and is better suited for general
distribution.

Author(s)

John W. Emerson and Walton A. Green and John A. Hartigan

References


See Also

YaleToolkit, gpairs, rug, stripplot

Examples

# Simulate some data:
x <- list(Rounded.2=round(rnorm(500, 2, 1),2),
  SmallerLevel=round(rnorm(100), rnorm(100,4,1)),
  LargerBivariateRounded.4=round(c(rnorm(500), rnorm(500,3,1)),4))

barcode(x)
barcode(x, main="Different orientation", horizontal=FALSE)
data(NewHavenResidential)
**NewHavenResidential**

`barcode(split(NewHavenResidential$dep, NewHavenResidential$zone),
  xlab="Percent Depreciation",
  main=paste("New Haven Residential Depreciation by Residential Zone",
              "RS = Single Family, RM = Mixed Residential", sep="\n"))`

---

**New Haven, CT Residential Property Data**

**Description**

Selected characteristics of a set of small residential properties in New Haven, CT (excluding larger multi-family properties and apartment buildings).

**Usage**

data(NewHavenResidential)

**Format**

A data frame with 18221 observations on the following 8 variables.

- `totalCurrVal` the 2006 assessed value of the property
- `livingArea` the living area in square feet
- `dep` the amount of depreciation, as a percent
- `size` the size of the land, in acres
- `zone` the residential zone, a factor with levels `other rm rs`
- `acType` whether the property has central air conditioning: a factor with levels `AC No AC`
- `bedrms` the number of bedrooms
- `bathrms` the number of bathrooms

**Details**

The data have been cleaned somewhat, with emphasis on somewhat. For example, there is a property (a very nice one), which has an extremely low assessed value, given its characteristics. It happens to straddle the border between New Haven and Hamden, and so it pays only a proportion of its property taxes to the City of New Haven.

**Source**

John W. Emerson, from the City of New Haven’s property database, which contains more than 27,000 property records (including, for example, the New Haven Airport) and many more variables than included here.
Examples

# This example is excluded from running automatically in the checks
# because it takes a little while to produce.

## Not run:
  data(NewHavenResidential)
gpairs(NewHavenResidential)

## End(Not run)
Index

*Topic datasets
  NewHavenResidential, 4

*Topic multivariate
  barcode, 1
  gpairs, 3
  NewHavenResidential, 4
  rug, 3
  stripplot, 3
  YaleToolkit, 3