

RMB: Visualizing categorical data with Relative Multiple Barcharts

Alexander Pilhöfer

1. Alexander Pilhöfer

Department of Computer Oriented Statistics and Data Analysis

Institute of Mathematics

University of Augsburg

86135 Augsburg, Germany

E-mail: alexander.pilhoefer@math.uni-augsburg.de

URL: <http://www.rosuda.org>

Keywords: categorical data, visualization, multiple barcharts, logistic regression

This talk will present a new graphic for displaying categorical data called “Relative Multiple Barcharts” (`rmb-plot`) and its implementation in the *R* package **extracat**. It is a new attempt to enrich the family of mosaicplots by combining the most important advantages of multiple barcharts (see [Hofmann, 2000](#)) and classical mosaicplots (see [Friendly, 1994](#)) in one display. The intention of `rmb-plots` is to precisely display relative frequencies of a target variable for each combination of explanatory variables divided over a grid-like graphical display and, simultaneously, their corresponding weights. The breakup of absolute frequencies into conditional distributions and weights is a common procedure in many methodologies for categorical data analysis, such as generalized linear models or correspondence analysis, but even so there seems to be a lack of graphical solutions for exploratory as well as illustrative purposes. After a brief introduction to the concepts of the plot the talk will focus on the implementation in *R*. Using the well-known Copenhagen housing dataset for the examples the talk will first present the main variants of the plot such as a generalized version of spineplots as well as the most important options like color schemes, ceiling censored zooming and residual-based shadings according to Poisson models or logistic regression models. Moreover an interactive version of the graphic based on the **iWidgets** package will be presented which provides interactive controls for all important options as well as a connection to classical mosaicplots in the **vcd** ([Hornik et al., 2006](#)) package.

References

- Friendly, M. (1994). Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association* 89, 190–200.
- Hofmann, H. (2000). Exploring categorical data: Interactive mosaic plots. *Metrika* 51(1), 11–26.
- Hornik, K., A. Zeileis, and D. Meyer (2006). The strucplot framework: Visualizing multi-way contingency tables with `vcd`. *Journal of Statistical Software* 17.
- Pilhoefer, A. and A. Unwin (2010). Multiple barcharts for relative frequencies and parallel coordinates plots for categorical data - package `extracat`. *Journal of Statistical Software*. submitted.