October 2018: “Top 40” New Packages

2018-11-29

by Joseph Rickert

One hundred eighty-five new packages made it to CRAN in October. Here are my picks for the “Top 40” in eight categories: Computational Methods, Data, Machine Learning, Medicine, Science, Statistics, Utilities, and Visualization.

Computational Methods

compboost v0.1.0: Provides a C++ implementation of component-wise boosting written to obtain high run-time performance and full memory control. The vignette shows how to use the package.

RcppEnsmallen v0.1.10.0.1: Implements an interface to the C++ based Ensmallen mathematical optimization library that provides a simple set of abstractions for writing an objective function to optimize. Optimizers include full-batch gradient descent techniques, small-batch techniques, gradient-free optimizers, and constrained optimization.

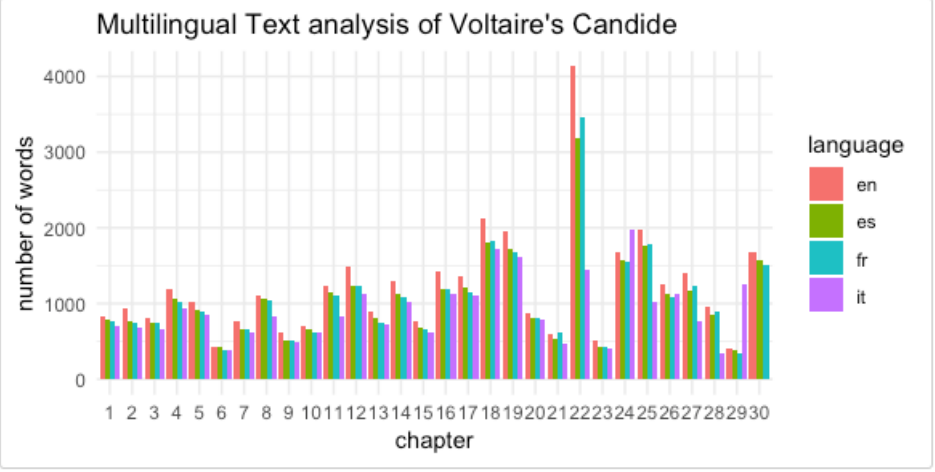
SAMpack v0.1.1: Implements Stochastic Approximation Monte Carlo (SAMC) samplers capable of sampling from multimodal or doubly intractable distributions. See Liang et al (2010) for a complete introduction to the method, and the vignette for an introduction to the package.

Data

crimedata v0.1.0: Provides access to publicly available, police-recorded open crime data from large cities in the United States that are included in the Crime Open Database.

nasapower v1.02: Implements an interface to POWER (Prediction Of Worldwide Energy Resource), NASA’s global meteorology, surface solar energy, and climatology data API. Look here for a quick start.

wikisourcer v0.1.1: Provides access to public domain works from Wikisource, a free library from the Wikimedia Foundation project. See the vignette for a package tutorial.

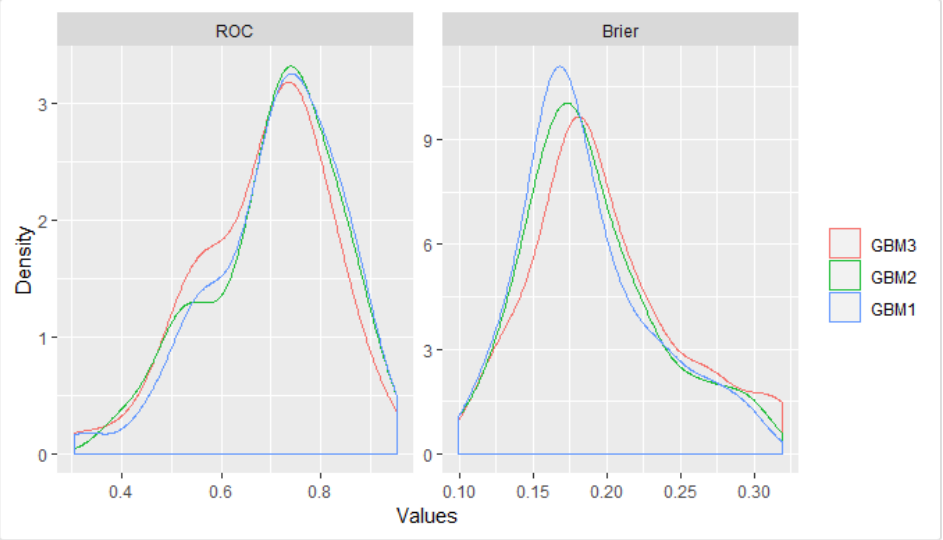


Height: 1080

gcForest v0.2.7: Provides an API interface to the Python implementation of Deep Forest, an alternative to Deep Learning. The algorithm is described in Zhou and Feng (2017), and there is a brief package tutorial.

galgo v1.4: Allows users to build multivariate predictive models from large data sets having a far larger number of features than samples, such as in functional genomics data sets. See Trevino and Falciani (2006) for details.

MachineShop v0.2.0: Provides a common interface for machine learning model fitting, prediction, performance assessment, and presentation of results. There is an Introduction and a note on Implementation Conventions.



mlflow v0.8.0: Provides an interface to `MLflow`, an open-source platform for the complete machine learning life cycle that supports installation, tracking experiments, running projects, and saving models.

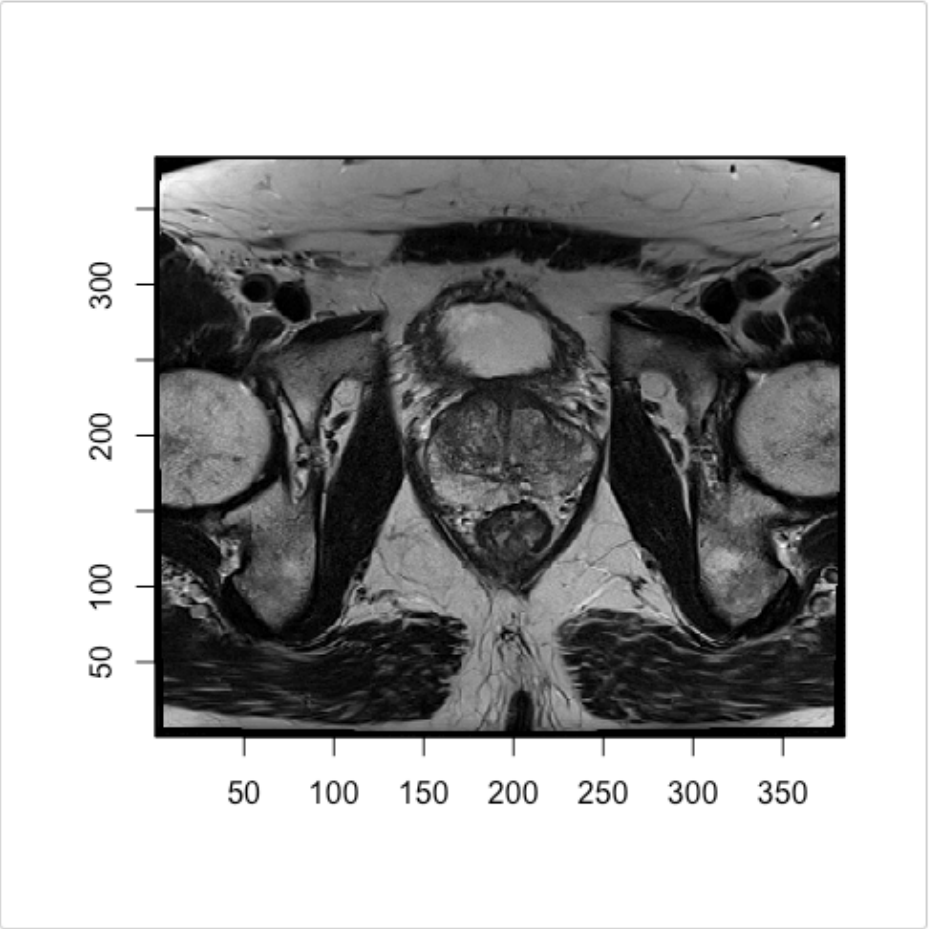
sboost v0.1.0: Provides a fast, C++-based implementation of Freund and Schapire's Adaptive Boosting (AdaBoost) algorithm, and includes methods for classifier assessment, predictions, and cross-validation.

Medicine

CoRpower v1.0.0: Provides functions to calculate power for assessment of intermediate biomarker responses as correlates of risk in the active treatment group in clinical efficacy trials, as described in Gilbert et al. (2016). The vignette demonstrates the math.

radtools v1.0.0: Provides a collection of utilities for navigating medical image data in DICOM and NIfTI formats. An emphasis on metadata allows simple conversion of image metadata to familiar R data structures, such as lists and data frames. The vignette shows how to use the package.

Height: 1080



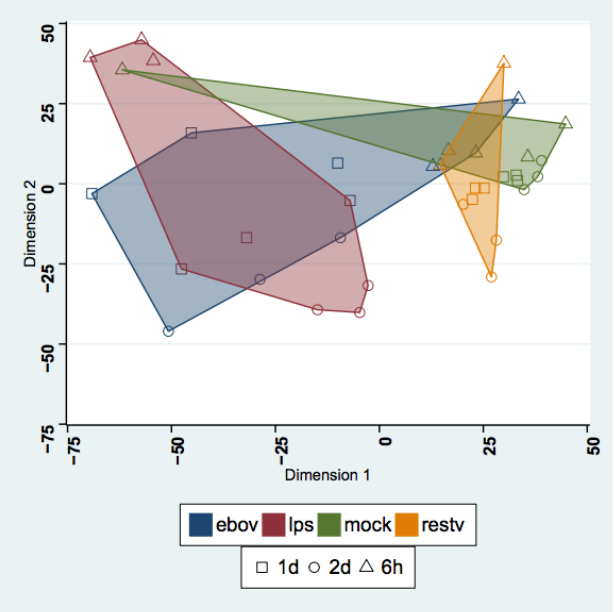
rpact v1.0.0: Provides functions for designing and analyzing confirmatory adaptive clinical trials with continuous, binary, and survival endpoints according to the methods described in the monograph by Wassmer and Brannath (2016). Look here for an overview.

Science

ClimProjDiags v0.0.1: Provides functions for computing metrics and indices for climate analysis, comparing models, and combining them into ensembles. There are vignettes on anomaly agreement, diurnal temperatures, extreme indices, and heat and cold wave duration.

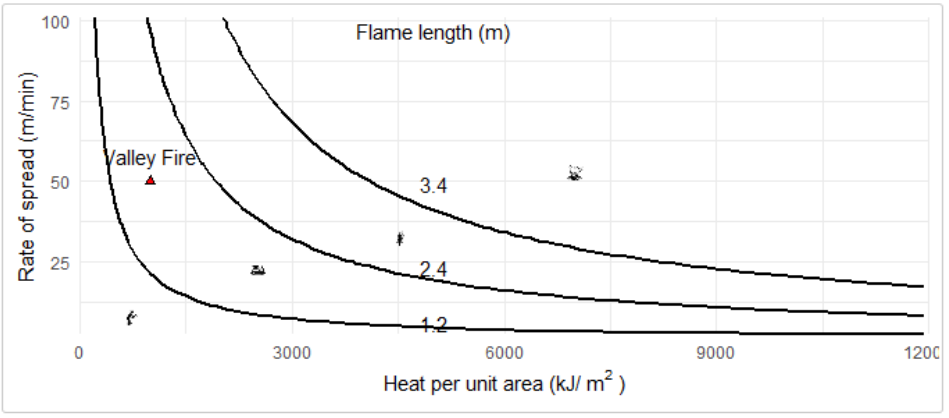
DEVis v1.0.0: Provides a comprehensive tool set for data aggregation, visual analytics, exploratory analysis, and project management that builds upon the Bioconductor DESeq2 differential expression package. The vignette offers a comprehensive introduction.

Height: 1080



epimdr v0.6-1: Provides functions for studying epidemics, including the S(E)IR model, time-series SIR and chain-binomial stochastic models, catalytic disease models, and coupled map lattice models. It is a companion to the book Epidemics: Models and Data in R and the Coursera course Epidemics Massive Online Open Course.

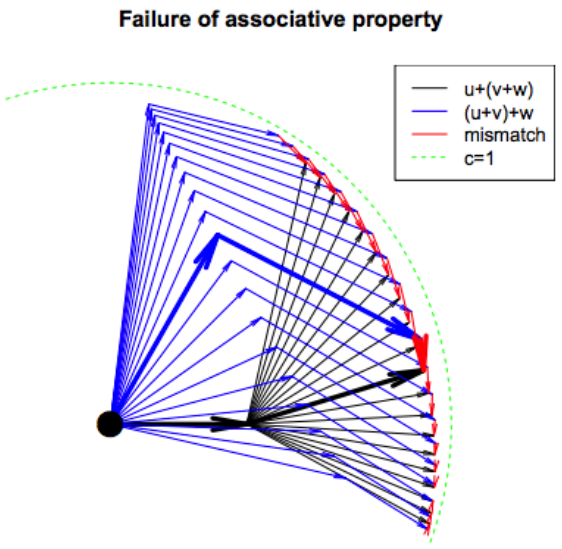
firebehavior v0.1.1: Implements fire behavior prediction models, including those documented in Scott & Reinhardt (2001) and Alexander et al. (2006). The vignette is informative.



Fire characteristics chart

lorentz v1.0.0: Provides the functionality to work with Lorentz transforms and the gyrogroup structure in Special Relativity.

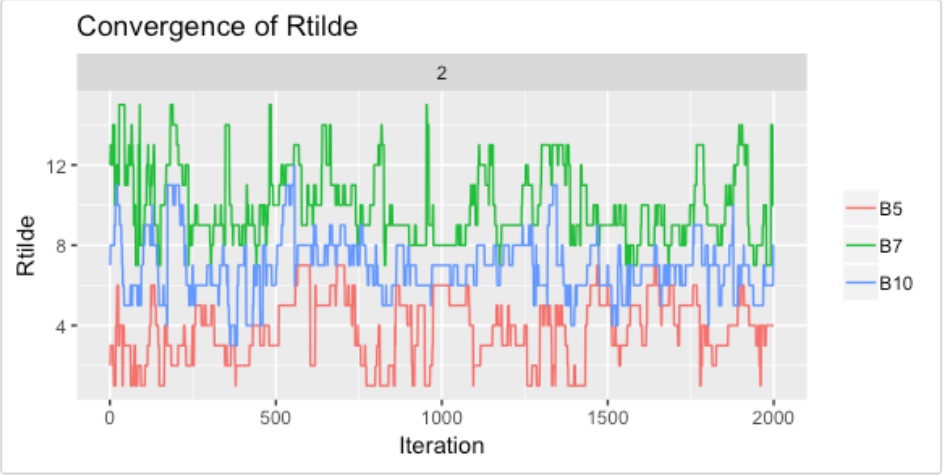
Height: 1080



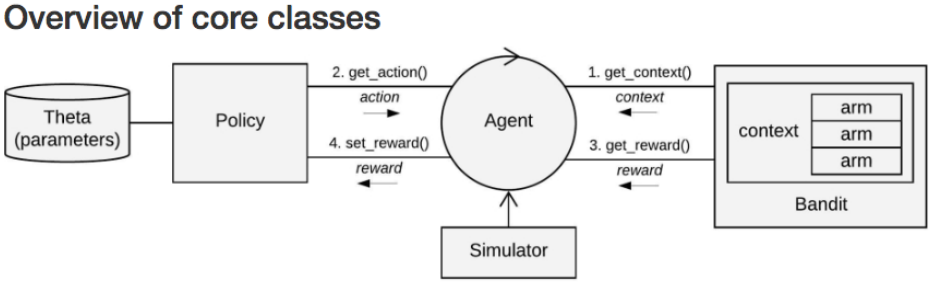
pubchunks v0.1.0: Provides functions for extracting chunks of XML from scholarly articles without having to know how to work with XML. See README to get going.

Statistics

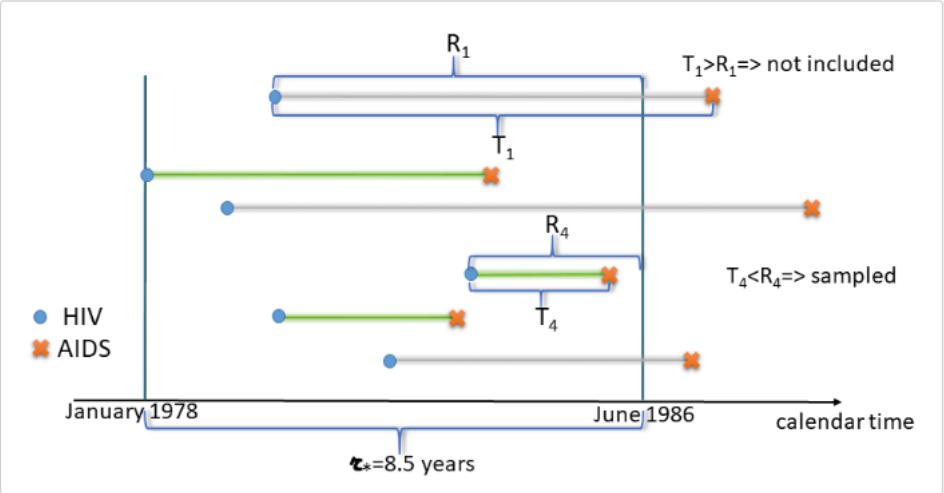
BayesMallows v0.1.1: Implements the Bayesian version of the Mallows rank model (Vitelli et al. (2018) (http://jmlr.org/papers/v18/15-481.html). The vignette provides the details.



contextual v0.9.1: Facilitates the simulation and evaluation of context-free and contextual multi-Armed Bandit policies or algorithms to ease the implementation, evaluation, and dissemination of both existing and new bandit algorithms and policies. See the Getting Started Guide and this list of posts for more information.



Height: 1080

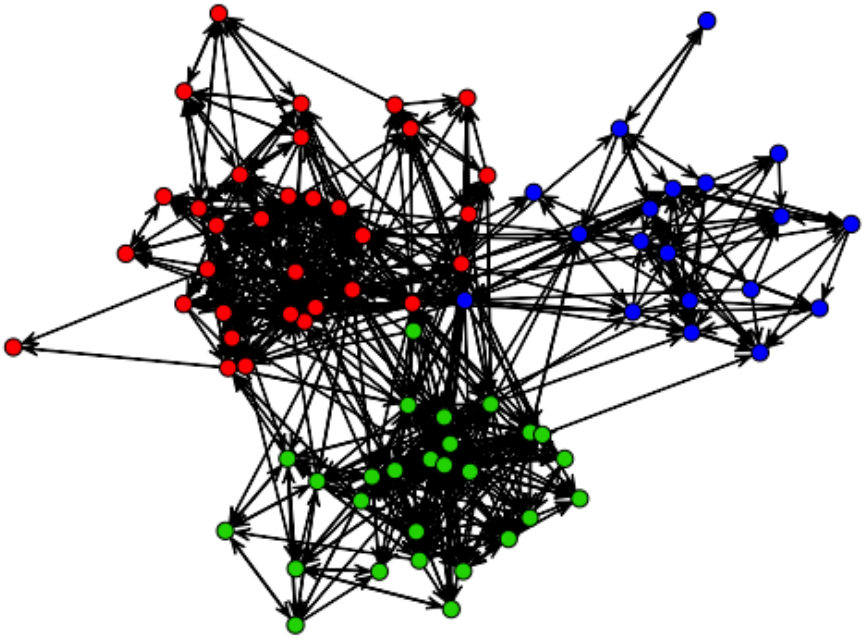


crossrun v0.1.0: Estimates the joint distribution of number of crossings and the longest run in a series of independent Bernoulli trials. There is a vignette.

logisticRR v0.2.0: Asserting that relative risk is often of interest in public health, this package provides functions to return adjusted relative risks from logistic regression model under potential confounders. The vignette does the math.

lognorm v0.1.3: Estimates the distribution parameters and computes moments and other basic statistics of the lognormal distribution Limpert al. (2001), and also provides an approximation to the distribution of the sum of several correlated lognormally distributed variables Lo (2013). There is a vignette on Aggregating Correlated Random Variables and another on Approximating Sums.

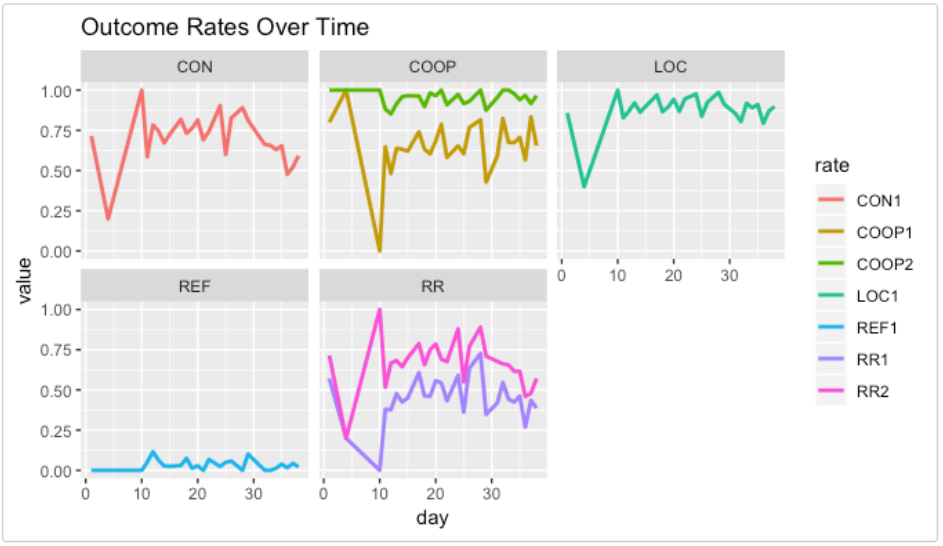
lolog v1.1: Provides functions to estimate Latent Order Logistic (LOLOG) Models for Networks, and also visual diagnostics and goodness of fit metrics are provided. See Fellows (2018) for a detailed description of the methods. One vignette works through an example, and another introduces lolog models.



matrixNormal v0.0.0: Provides the functions to compute densities, probabilities, and random deviates of the Matrix Normal distribution. See Iranmanesh et.al. (2010)

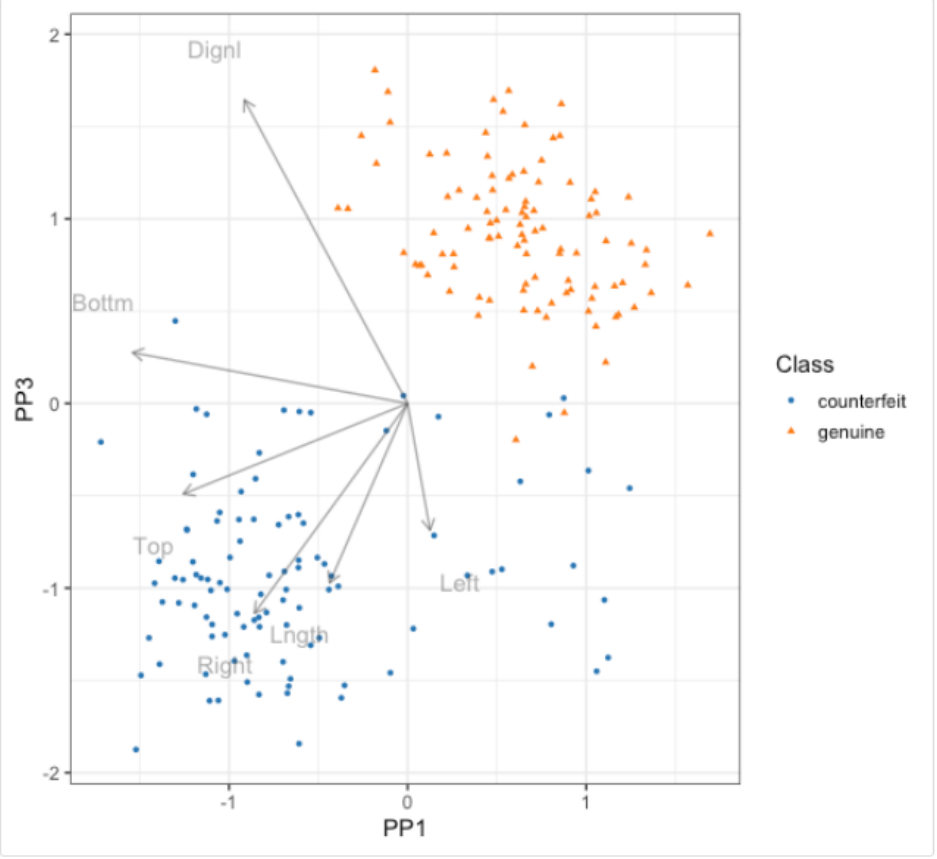
outcomerate v1.0.1: Implements standardized survey outcome rate functions, including the response rate, contact rate, cooperation rate, and refusal rate that allow researchers to measure the quality of survey data using standards published by the American Association of Public Opinion Research. For details, see AAPOR (2016). The vignette

Height: 1080



parmsurvfit v0.0.1: Fits right-censored data to a given parametric distribution, and produces summary statistics, hazard, cumulative hazard and probability plots, and the Anderson-Darling test statistic. There is a vignette.

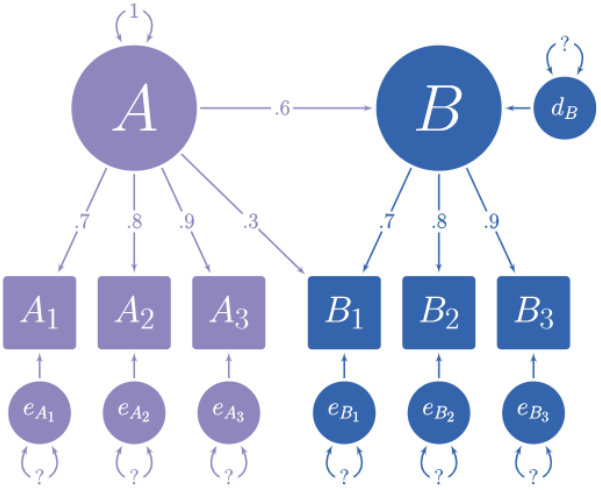
ppgmmga v1.0.1: Implements a Projection Pursuit algorithm for dimension reduction based on Gaussian Mixture Models. The vignette provides a quick tour of the package.



RcppDist v0.1.1: Provides additional statistical distributions that can be called from C++ when writing code using Rcpp or RcppArmadillo. See the vignette for a list of the distributions supported.

simstandard v0.2.0: Enables the creation of simulated data from structural equation models with standardized loading. The vignette shows how to use the package.

Height: 1080



Utilities

carrier v0.1.0: Enables users to create functions that are isolated from their environment. These isolated functions, also called crates, print at the console with their total size and can be easily tested locally before being sent to a remote.

carbonate v0.1.0: Implements an interface to carbon.js, which allows developers to create images of source code. There is a vignette on Tests and Coverage.

generics v0.0.1: In order to reduce potential package dependencies and conflicts, `generics` provides a number of commonly used S3 generics.

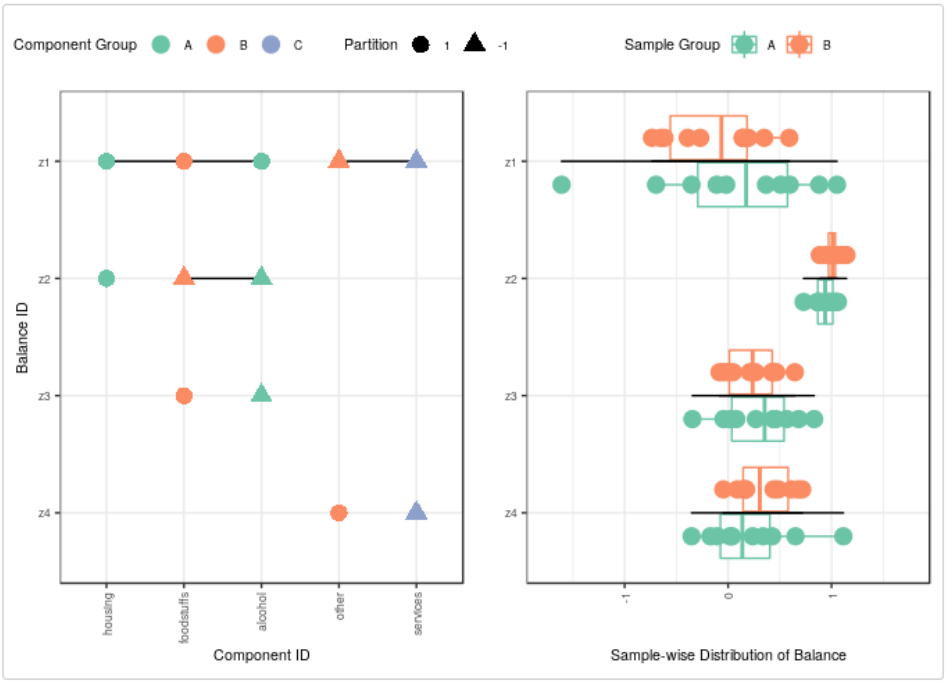
REPLesentR v0.3.0: Allows users to create presentations and display them inside the R `REPL` (console). Supports `RMarkdown` and other text format.

stationery v0.98.5.5: Provides templates, guides, and scripts for writing documents in `LaTeX` and `R markdown` to produce guides, slides, and reports; and includes several vignettes to assist new users of literate programming. There is an Overview, a vignette on R Markdown Basics, and another on R Markdown HTML, and a comparison between Sweave and Knitr code chunks.

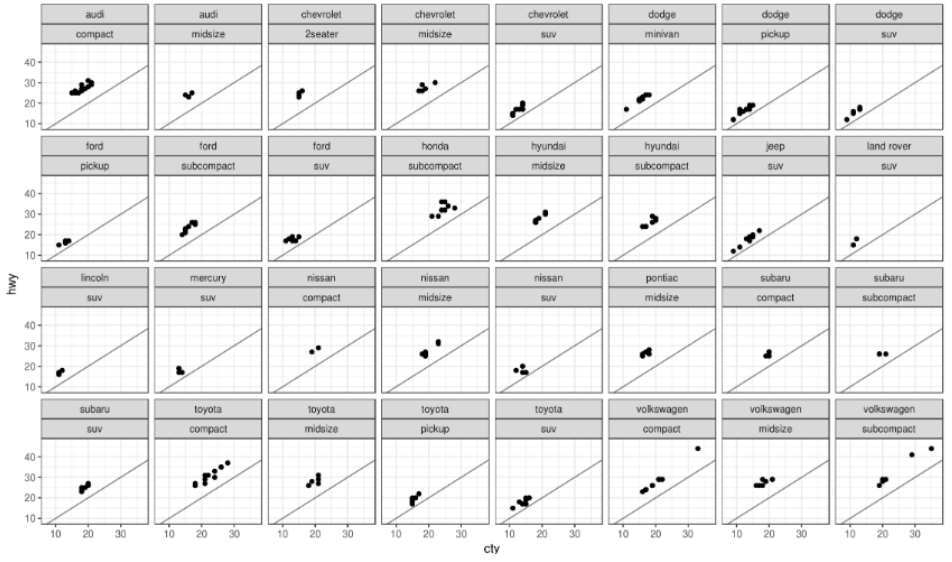
Visualization

balance v0.1.6: Provides an alternative scheme for visualizing balances (used in compositional data analysis) as described in Quinn (2018), as well as a method for principal balance analysis. See the vignette for details.

Height: 1080



trelliscopejs v0.1.14: Provides methods that make it easy to create a Trelliscope display specification for TrelliscopeJS, including high-level functions for creating displays from within `dplyr` or `ggplot2` workflows. There is a vignette on trelliscope Documentation and a trelliscope Package Reference.



R Language · R Packages · Opinion R Language · R Packages · CRAN Comments Share

OLDER
Slack and Plumber, Part Two

NEWER
Statistics in Glaucoma: Part I

Height: 1080