iPlots 2.0

Tobias Wichtrey
Alexander Gouberman
Simon Urbanek
Martin Theus

iPlots: Motivation

- R is good at managing
  - data
  - models
  - (static) graphics

  but is less strong in exploratory data analysis
- Interactive Statistical Graphics (ISG) is good at
  - supporting exploratory analyses
  - checking data quality
  - revealing structure in data

  but can not be automated or scripted
- Solution: Bring both tools/paradigms together

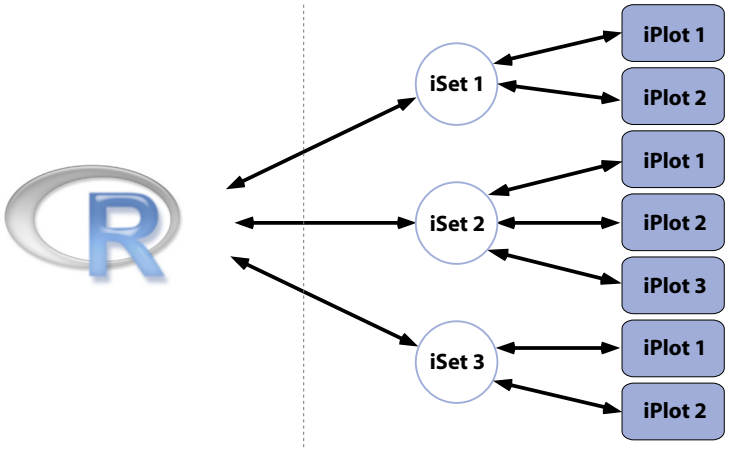
Bringing Interactive Graphics and R together

- Different ways of bringing ISG and R together

1. **Run two applications in parallel**
   pros: full feature-set of both applications available
   cons: two different user interfaces, coupling relatively loose
   example: ggobi
2. **Use R as stat-computing engine**
   pros: no need to learn R, only one interface
   cons: only packaged functionality, no extensibility
   example: KLIMT, Mondrian (all via Rserve)
3. **Add interactive plots within R**
   pros: one interface, still "just" R, flat learning curve
   cons: can not be implemented using R graphics
   example: iPlots

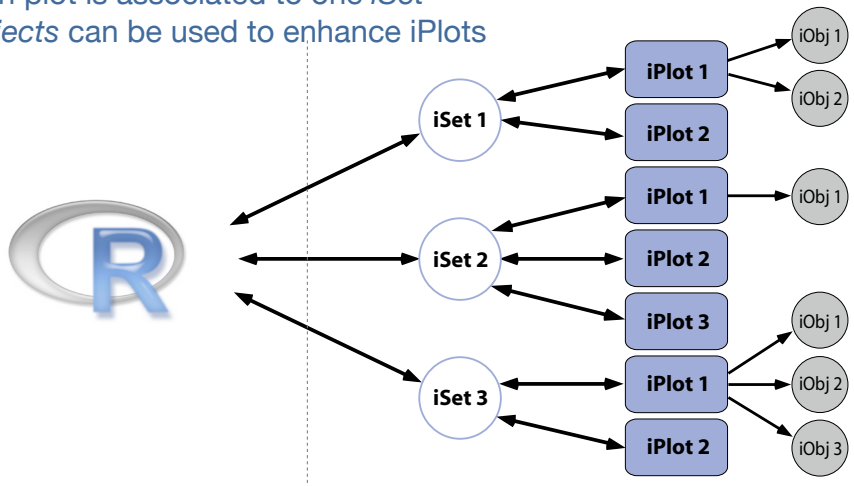
iPlots Internals

- iPlots use JAVA to achieve interactivity
- Data is stored in so called *iSets*
- Each plot is associated to one *iSet*



iPlots Internals

- iPlots use JAVA to achieve interactivity
- Data is stored in so called *iSets*
- Each plot is associated to one *iSet*
- *iObjects* can be used to enhance iPlots



iPlots: Past

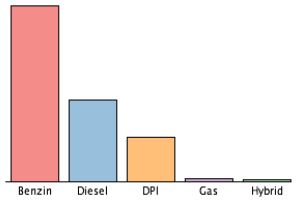
- The first version of iPlots was presented at the DSC meeting in 2003.
- Features of Version “1.0”
  - implemented basic plots
    - histogram
    - barplot
    - scatterplot
  - defined API
    - as similar to existing R functions as sensible to flatten the learning curve
    - handling of *iSets* and *iObjects*
  - available from RoSuDa repository
  - “proof of concept”

What is new in iPlots 2.0?

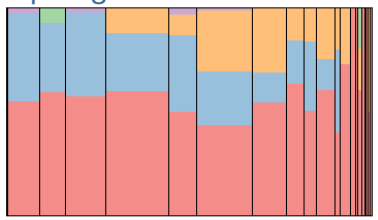
- Extensions to existing plots:
  - Histogram / Spinogram
  - Barplot / Spineplot
- New (multivariate) Plots
  - (parallel) Boxplots (y by x)
  - Parallel Coordinate Plots
  - Mosaic Plots (and its variants)
- New Features
  - Color Brushing
  - Better control through R calls
- OpenGL support to speed up glyph-based plots
- Custom plots allow creation of new interactive plots

Extensions to existing Plots

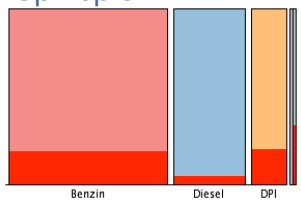
- Conditional plots for continuous and categorical data

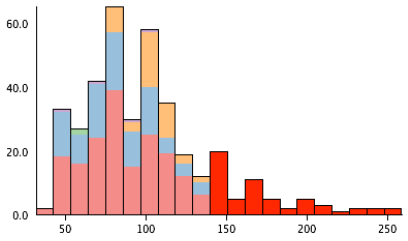


Spinogram



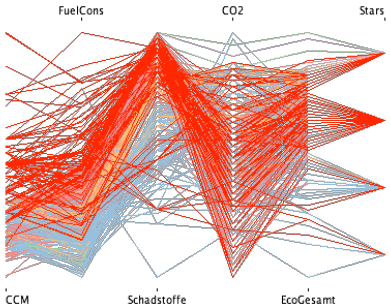
Spineplot



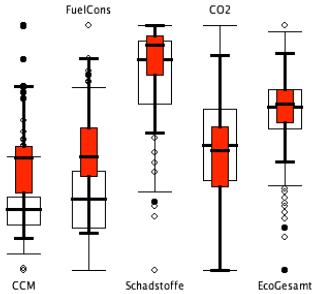


New Multivariate Plots

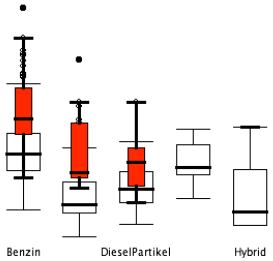
Parallel Coordinates



Parallel Boxplots

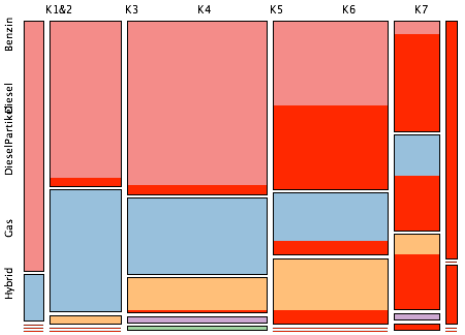


Boxplot y by x

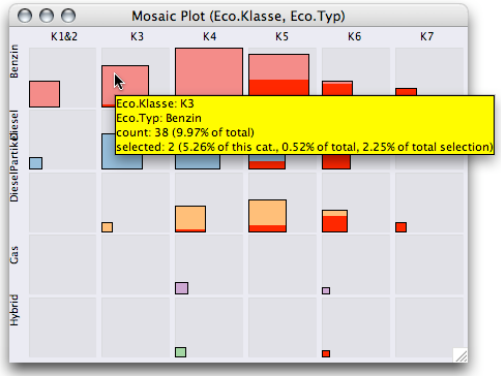


New Multivariate Plots

Mosaic Plot



Fluctuation Diagram



- Further variations include
  - Same Binsize
  - Multiple Barchart
  - Double Decker Plot

New Features

- Color Brushing, both
  - Quantitative and
  - Qualitative
- Extended Queries
  All objects – points, lines, axes, plot-canvases – can be queries. Results of extended queries can even be user defined.
- Full Parameter control from R
- α blending is implemented for all-glyph based plots to get crude density estimations and handle larger data decently.

AWT vs. 2D vs. OpenGL

- Java is platform independent, but graphics rendering is still done by the CPU (as of Version 5.0, 6.0, …)
- iPlots support three different “graphics” engines
  - AWT
  - Swing
  - OpenGL
- OpenGL speeds up glyph-based plot by factor
  - 2-3 point based plots
  - ~10 for line based plots
- Specific timings may vary, essential improvement is to push the rendering from the CPU to the GPU.

Custom Plots

- iPlots 2.0 support several standard plots which are defined on the JAVA side
- In an extensible environment like R, we want to be able to build new plot, defined by R code.
- iPlots 2.0 expose the plot primitives (elementary objects like points, lines/polygons, bars, …) defined on the JAVA side within R.
- These plot primitives know about:
  - selection
  - highlighting
  - queries
- See also the Focus Session on Friday 15:00 - 18:30.

Conclusions

- iPlots 2.0 now feature the full set of statistical standard graphics.
- Advanced features like color brushing and extended queries
- Custom plots offer new perspective in prototyping and developing new interactive applications.
- Soon available on CRAN
- Still need a Logo? Any ideas?