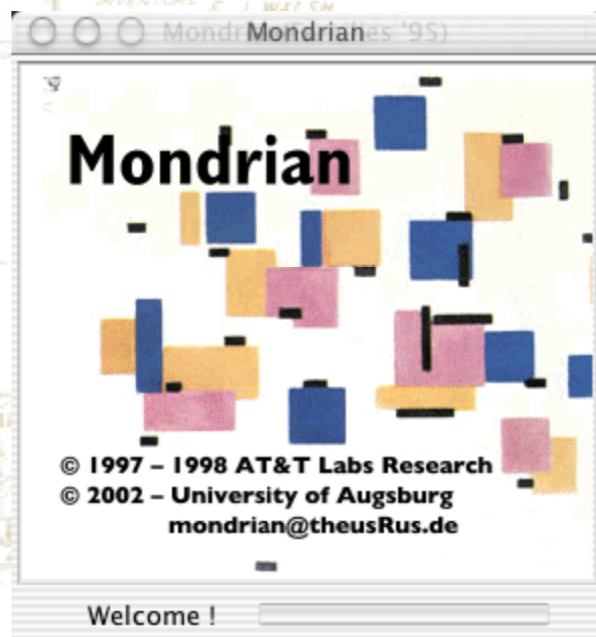


MONDRIAN

Interactive Statistical Graphics in JAVA



martin.theus@math.uni-augsburg.de

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

OUTLINE

- Why JAVA ?

"I like your software, but ... "

- Plots in Mondrian

There is life beyond scatterplots

- Selections

Making selections smart

- Interfacing Databases

When things get really big

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

WHY JAVA ?

JAVA is

- graphical:
 - 2D, Swing, PostScript™ printing
- portable:
 - "write once run everywhere"
 - ISO Standard
- object oriented:
 - if you are afraid of C++: use JAVA
- based on libraries/packages
- developing

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

DESIGN CHOICES:

- Graphs

JFrame

Selectable
Frame

Box Plot

Map

...

- Data

Table

Data Set / JDBC Abstraction

Var 1

Var 2

...

- Flat hierarchies for fast execution!

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

EVENT HANDLING:

- Events in Mondrian

- Selection Event
- DataChanged Event

- Example: Selection Event in Plot 2

Parent:
Event listener / distributor

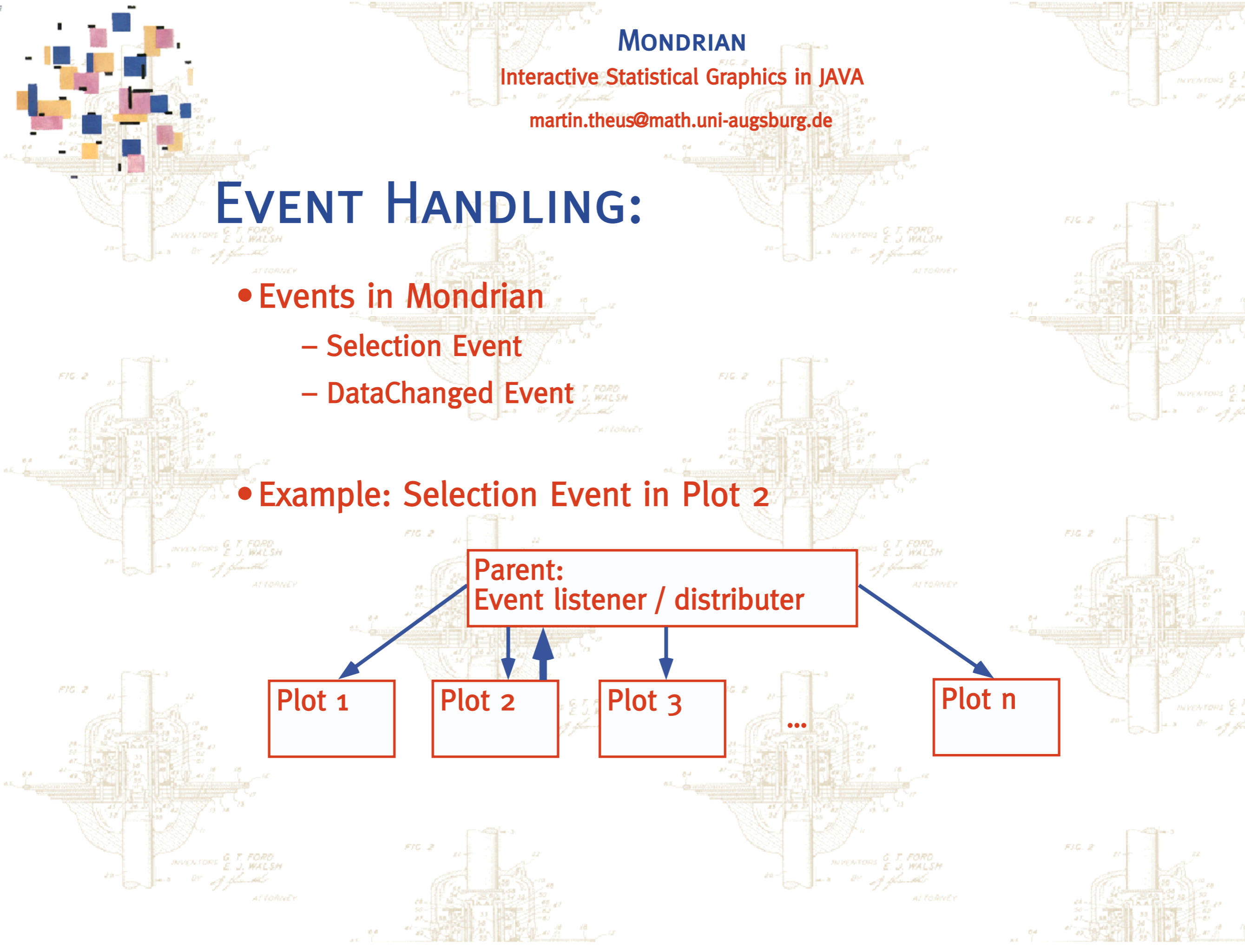
Plot 1

Plot 2

Plot 3

Plot n

...



MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

CONVENTIONS:

Selections:

- click and drag -> create a selection rectangle
- click on selection rectangle handle -> resize this selection
- shift click in selection rectangle -> alter this selection

Interrogation:

- popup-trigger on objects -> interrogation

Alterations:

- meta-click and drag -> zoom in/out
- popup-trigger on background -> alter the plot
- alt-click and drag -> reorder objects

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

All support:

- selection and linked highlighting
- zooming, if there is a suitable coordinate system
- logical zooming, if possible/necessary
- interrogation of objects
- tonal highlighting to handle overplotting

... all plots are “aware” of big ($>10^6$ cases) data sets

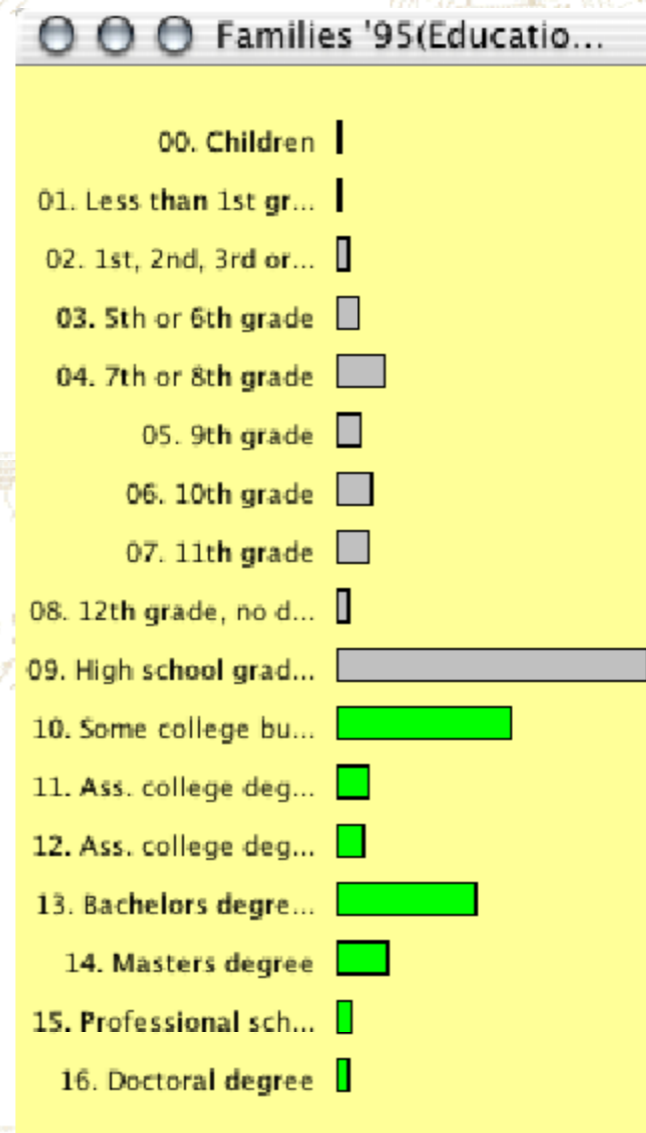
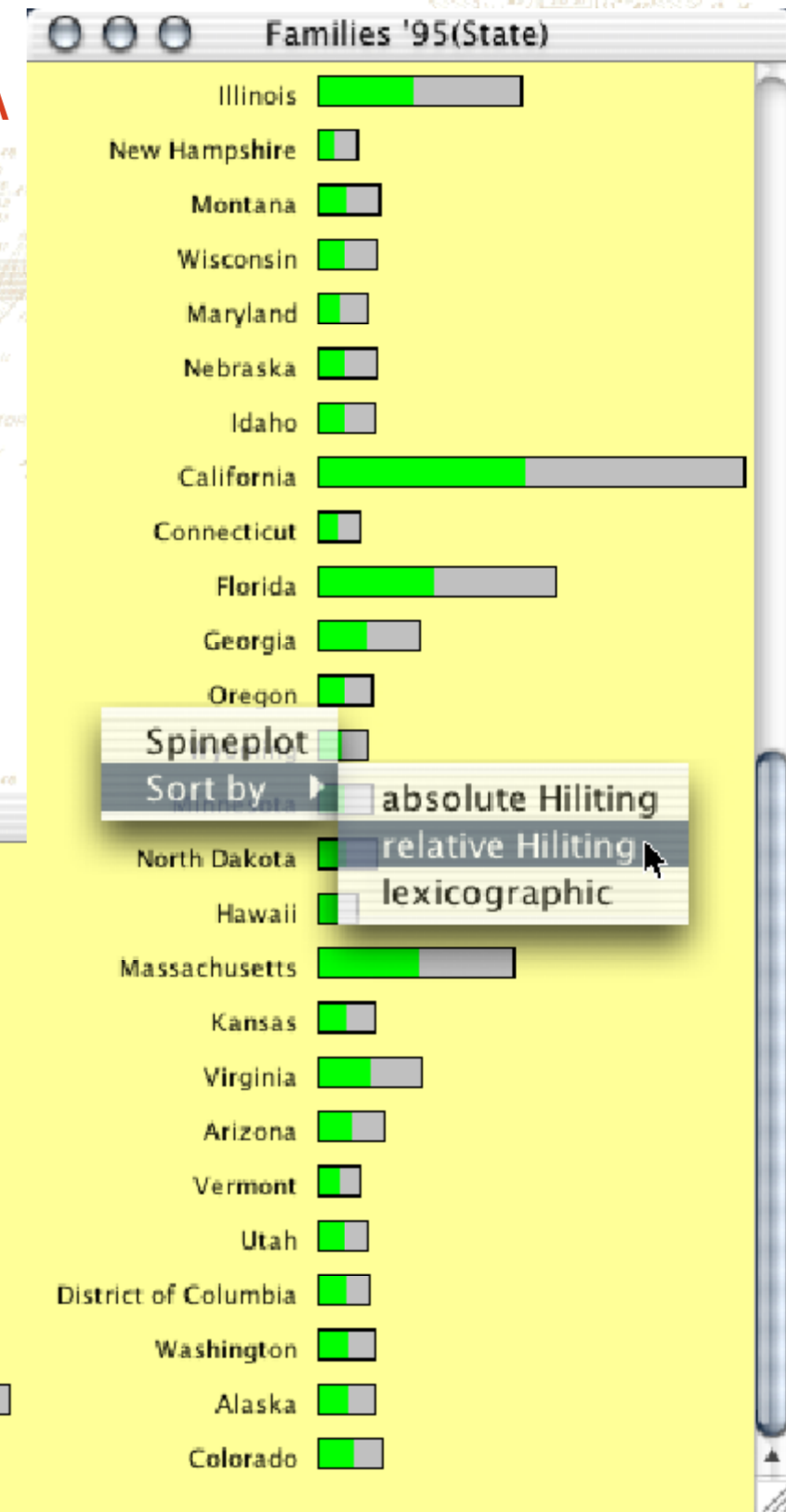
MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

- **Barcharts:**
 - horizontal layout for full labels
 - weighted version
 - Spineplot display
 - scrollable
 - sorting of categories by
 - highlighting
 - relative
 - absolute
 - name
 - "hand"



MONDRIAN

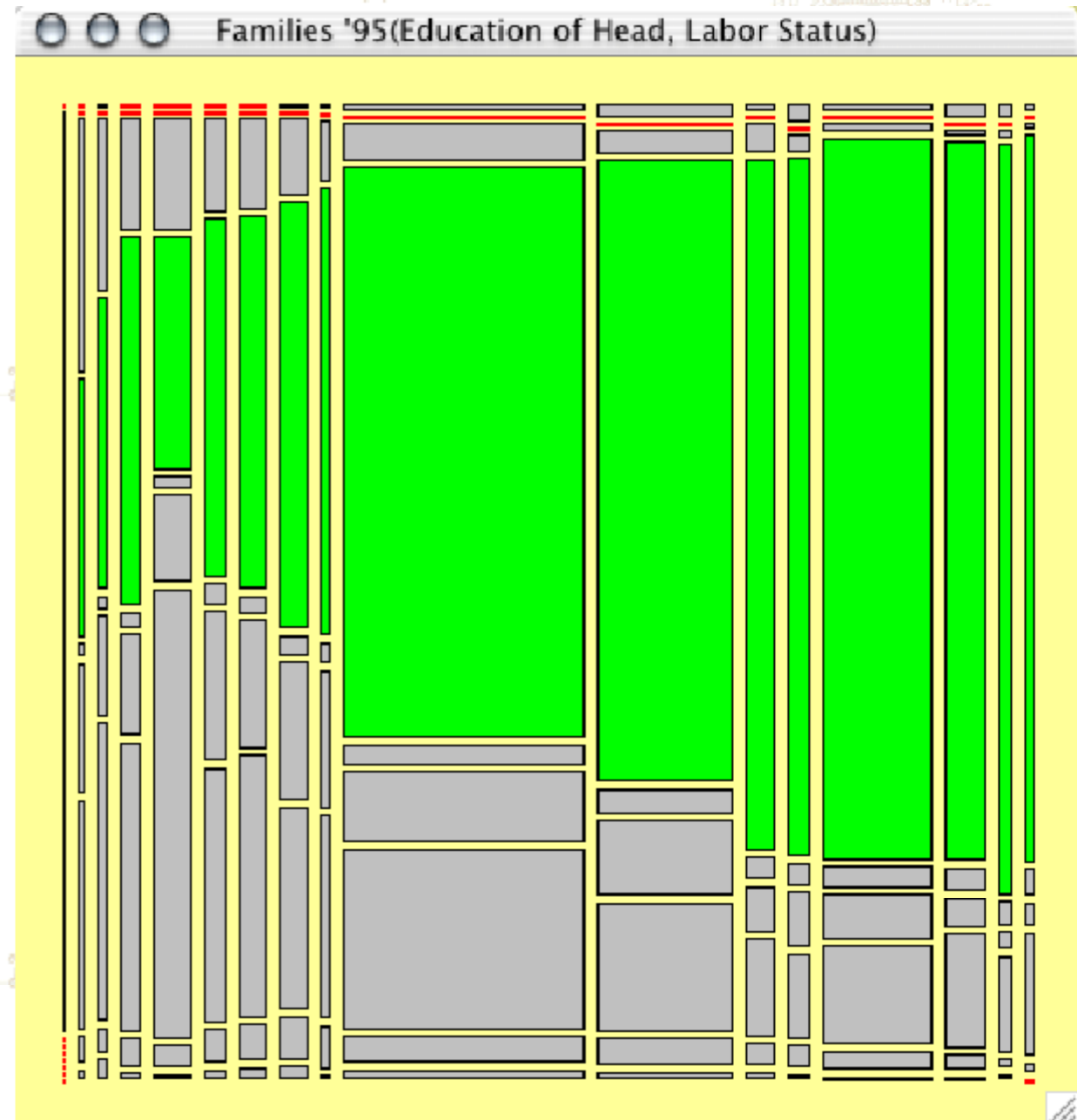
Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

- Mosaic Plots:

- interactive reordering (easy keyboard controls)
- empty cells on higher levels are not subdivided any further
- no labeling in favor to flexible interrogation
- interactive modeling of loglinear models
- weighted version



MONDRIAN

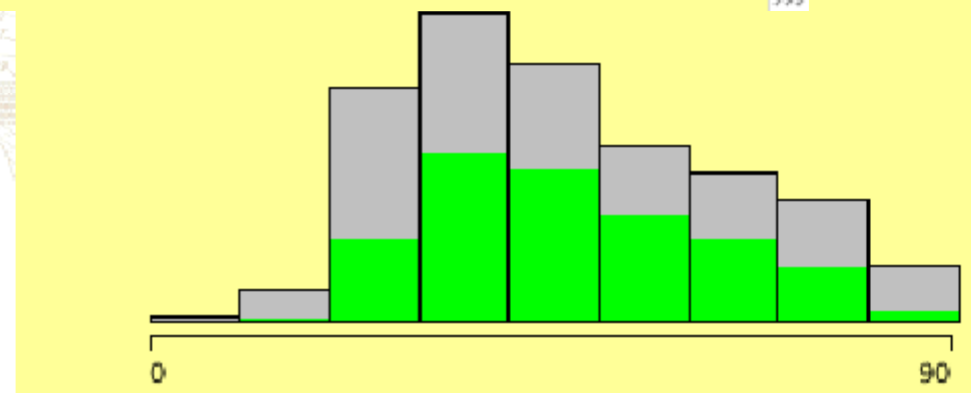
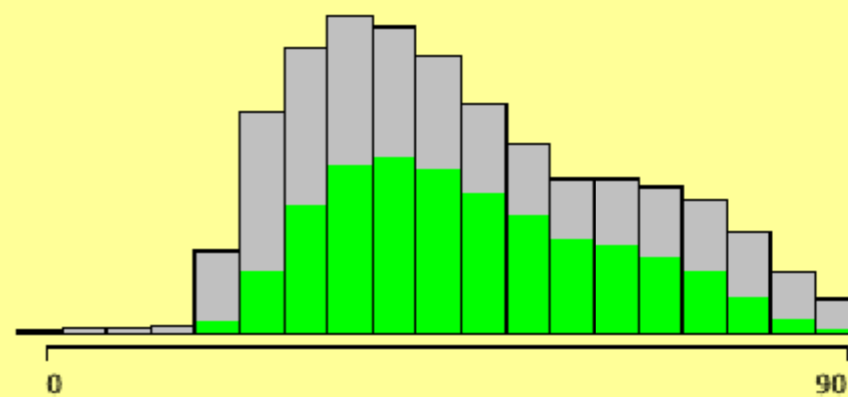
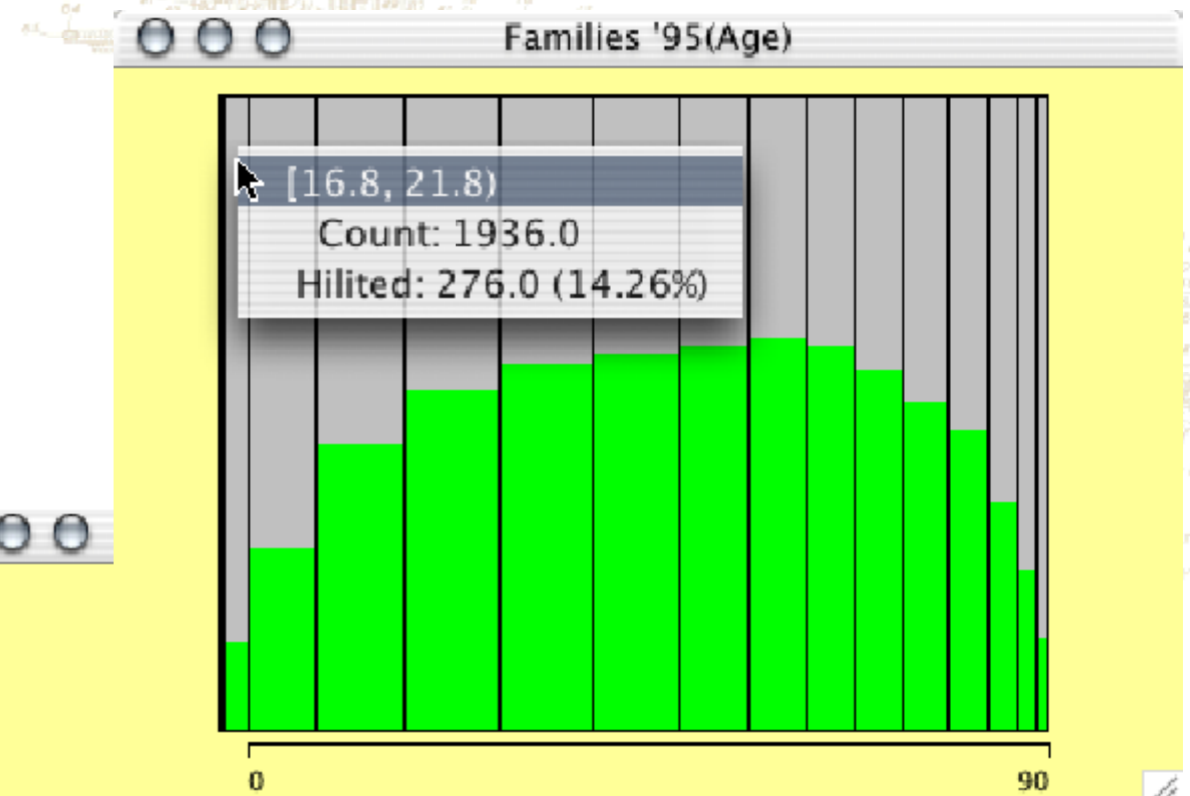
Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

- Histogram / Spinogram

- Start point and bin width can quickly be changed by simple keyboard controls
- Exact values can be set via sliders
- Spinogram generalizes from Spineplot
- Attention: x-axis no longer “simple” to interpret



MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

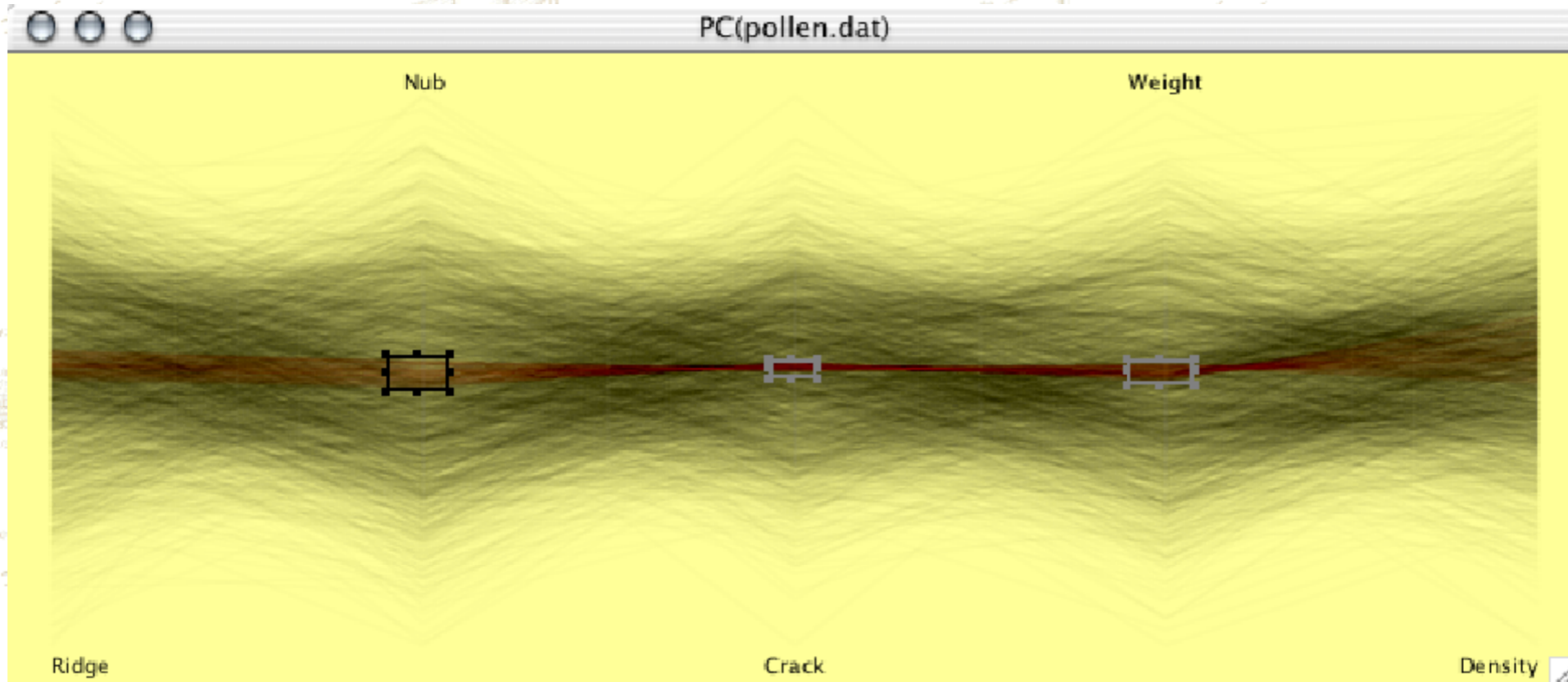
- Parallel Coordinates / Boxplots:

- Common/Individual scale of axes

- Lines/Boxes or both (i.e. only selected lines) can be displayed

- Axis can be reordered

- “tonal highlighting” via alpha-channel to estimate density



MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

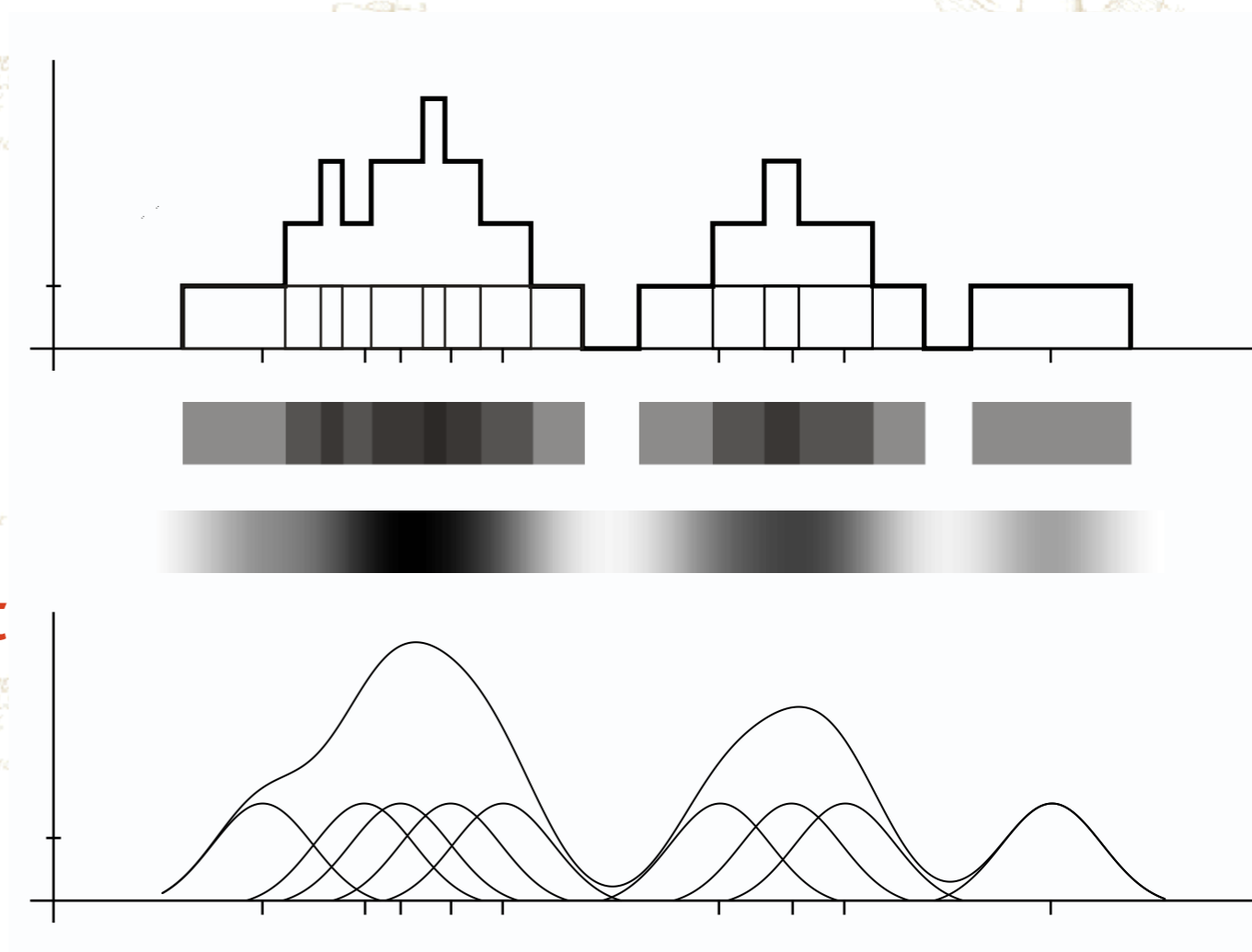
PLOTS IN MONDRIAN

- Tonal Highlighting

- Implemented in Mondrian, Chrystal Vision and MANET
- “Cheap” version of density estimation
- Uses alpha-channel of graphics system
- For larger samples sub-optimality is less important

EXAMPLE

tonal highlighting



kernel estimate (normal)

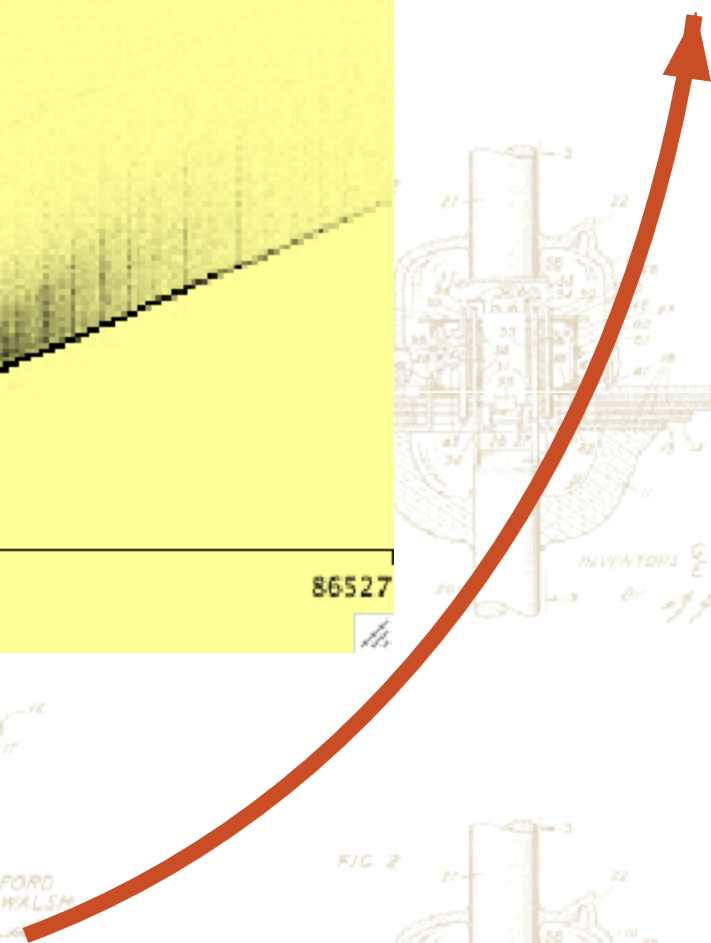
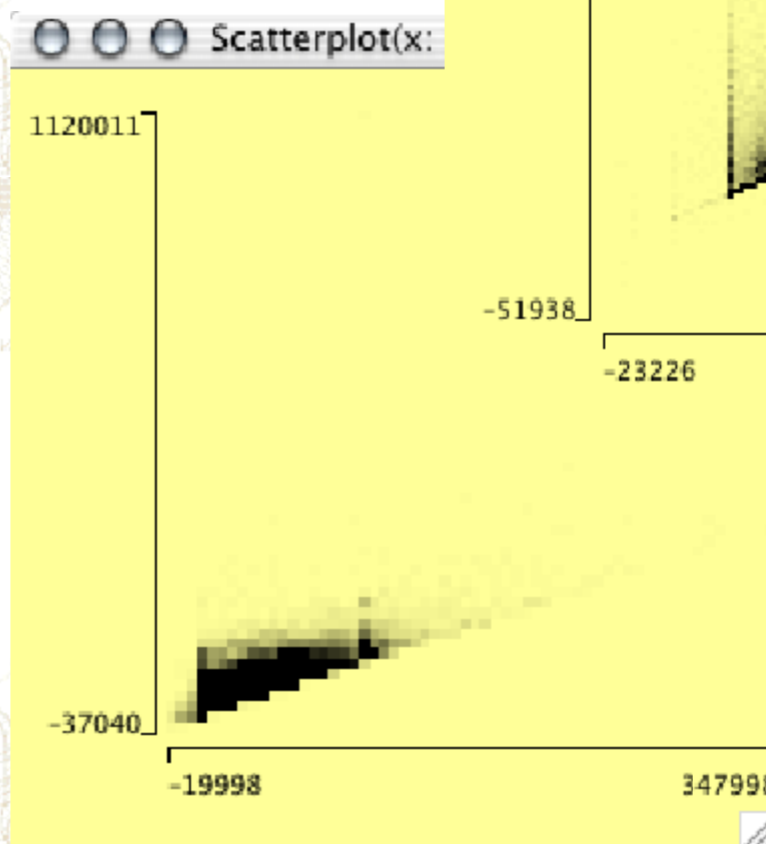
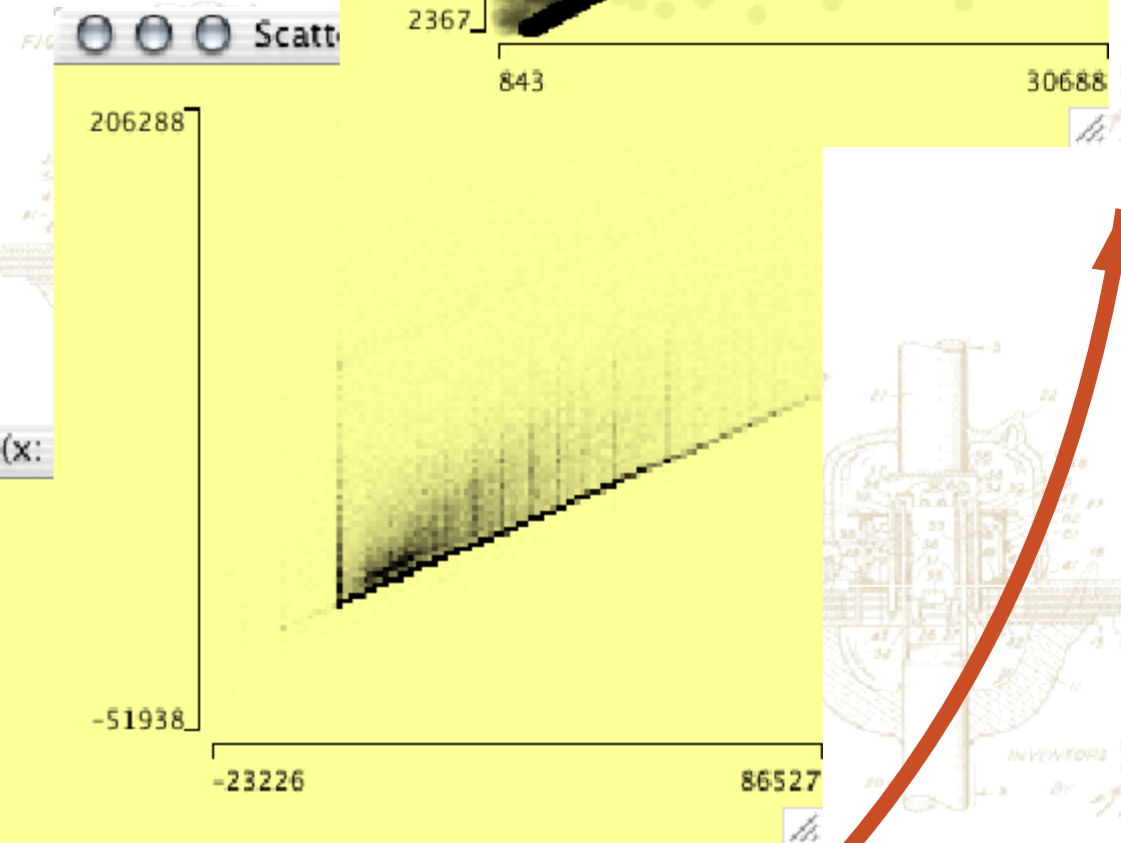
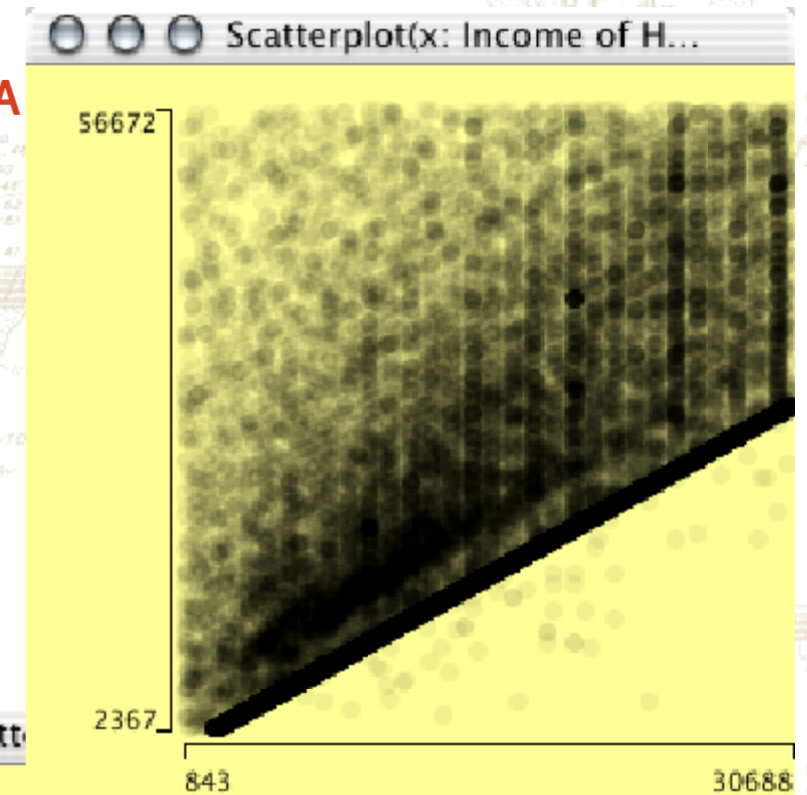
MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

- Scatterplots
 - Fixed/free aspect ratio
 - Flipping of axes
 - Logical zooming, i.e. for large dataset, data is binned.
 - Tonal highlighting if points are plotted.



MONDRIAN

Interactive Statistical Graphics in JAVA

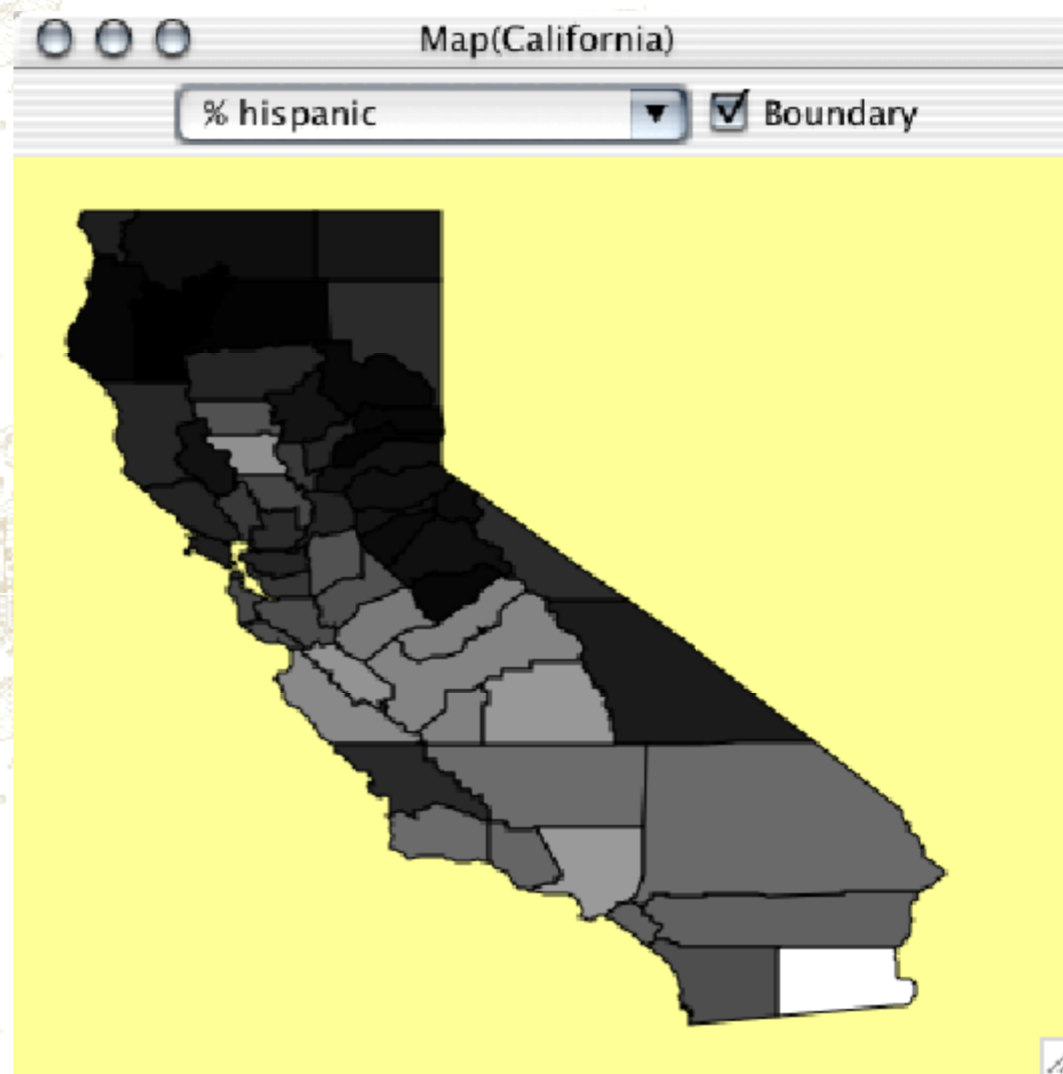
martin.theus@math.uni-augsburg.de

PLOTS IN MONDRIAN

- Maps

- Choropleth maps

- w/out boundaries



MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

SELECTIONS

- Standard

New selections replace the old selection set. (e.g. GGobi)

- Advanced

Existing selections can be combined with further selections via AND, OR, XOR, NOT, ... (e.g. DataDesk)

- Selection Sequences

Any element in a hierarchical sequence of selections is stored and can be altered. (e.g. MANET)

- Smart Selection Sequences

Store selections in "data coordinates" rather than in "screen coordinates" (e.g. Mondrian)

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

SMART SELECTION SEQUENCES

- Selection Rectangle
 - Indicates selection region
 - Can be used as brush
 - Resizable at 8 handles
 - Maintained via pop-ups
 - Fast deletion with Backspace (Undo)



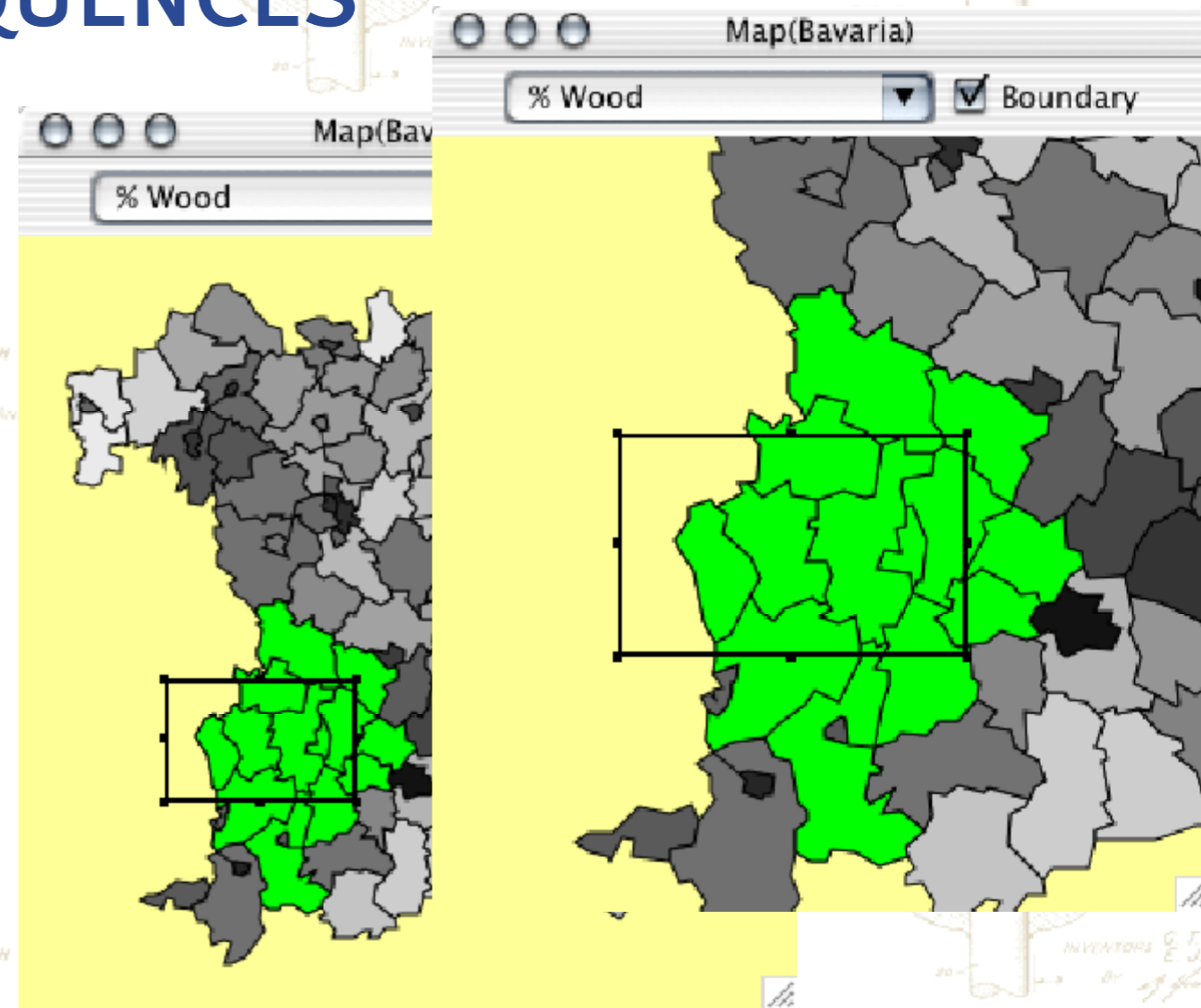
MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

SMART SELECTION SEQUENCES

- Transformation invariant
 - Location is determined from the data
- Multiple Selections per Window
 - Any window can hold as many selection rectangles as necessary



MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

INTERFACING DATABASES

- Why?

- Data can be too big to handle in memory
- Data may be transactional
- Many people may work on the same data simultaneously
- Corporate data always sits in databases

HOW TO CONNECT?

The screenshot shows a 'DB Connection' dialog box with the following fields and buttons:

Driver:	org.gjt.mm.mysql.Driver	Load
URL:	jdbc:mysql://jetta.math.uni-augsburg.de:3306/mysql	
User:	theusm	Pwd:
DB:	Not Connected	Connect
Table:	Choose DB	
Select		Cancel

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

INTERFACING DATABASES

- Consequence

- Data is no longer available in memory, but ...
- can only be retrieved from the database
- Selections must be expressed in database queries – usually SQL
- Only plots based on summaries are feasible, except for ...
- hot selection plots, only containing selected data
- Things will be less efficient, but at least effective
- Make sure your DB and tables are “in a good shape”

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

INTERFACING DATABASES

- Problems

- Interactions might be too slow to be “interactive”
- What happens to plots if the underlying data changes in the DB?
- We usually can't write to the DBs

But:

- Results are still better than a “simple” text query

MONDRIAN

Interactive Statistical Graphics in JAVA

martin.theus@math.uni-augsburg.de

SUMMARY:

- JAVA is the best choice for platform independence, especially when dealing with graphics
- Proper handling of categorical data is important
- Plots must scale up to more than just a thousand cases.
- Selection techniques get more and more important when dealing with large datasets
- Direct database connections are important to complement standard file based data sources