5. Interaction Design

Dynamic linking, brushing and filtering in Information Visualization displays

Dr. Thorsten Büring, 22. November 2007, Vorlesung Wintersemester 2007/08

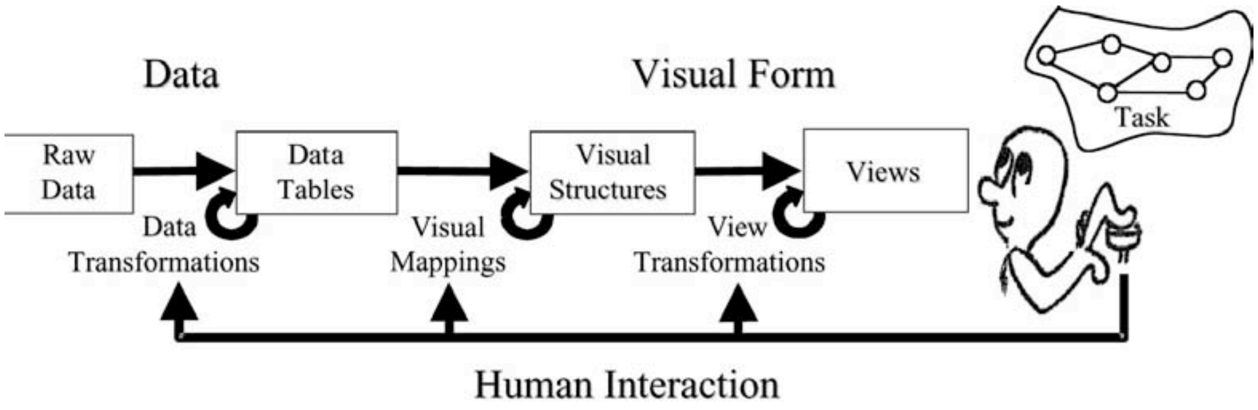
Ludwig-Maximilians-Universität München

Ludwig-Maximilians-Universität München LMU

Outline

Ludwig-Maximilians-Universität München

Reference Model for Visualization



Raw Data: idiosyncratic formats
Data Tables: relations (cases by variables) + meta-data
Visual Structures: spatial substrates + marks + graphical properties
Views: graphical parameters (position, scaling, clipping, …)

Card et al. 1999

InfoVis & Interaction

- Information Visualization research: focus on finding novel visual representations
- Recently one can observe an increasing interest in interaction design, HCI models and evaluation as well as aesthetics
- HCI Interaction models help us to better understand the complex concepts of human-machine communication
- Norman's execution-evaluation cycle (Norman 1988)
  - 1. Establishing the goal
  - 2. Forming the intention
  - 3. Specifying the action sequence
  - 4. Executing the interaction
  - 5. Perceiving the system state
  - 6. Interpreting the system state
  - 7. Evaluating the system state with respect to the goals and intentions

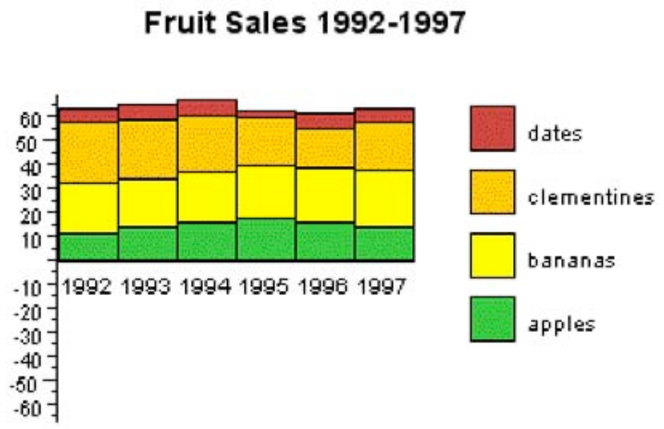
Ludwig-Maximilians-Universität München LMU

Outline

- InfoVis & Interaction
- Direct Manipulation (DM)
- Common Interaction Techniques
  - Brushing
  - Zooming & Panning
  - Dynamic Queries
- Attribute Explorer
- Brushing Histograms vs. DQ
- Dynamic Queries and Movable Filters

Direct Manipulation (DM)

- Shneiderman 1982
- DM features
  - Visibility of the objects of interest
  - Incremental action at the interface with rapid feedback on all actions
  - Reversibility of all actions, so that users are encouraged to explore without penalties
  - Syntactic correctness of all actions, so that every user action is a legal operation
  - Replacement of complex command languages with actions to manipulate directly the visible objects
- DM does not only make interaction easier for novice users but fundamentally extends visualization capabilities
- Simple example: stacked histogram

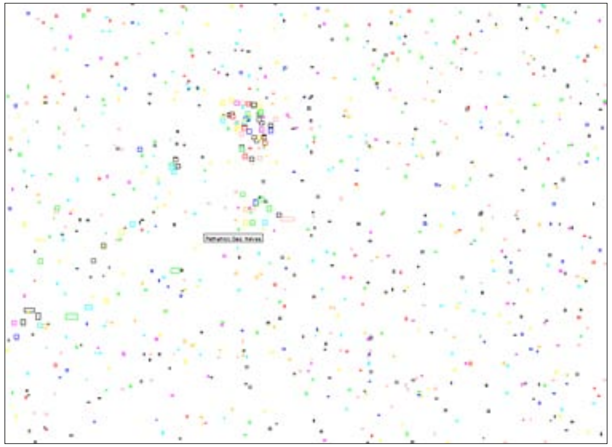


Stacked histogram; how are the banana sales progressing???
http://www.hiraeth.com/alan/topics/vis/hist.html

Ludwig-Maximilians-Universität München

Excentric Labeling

- Another DM example
- Fekete & Plaisant 1999
- Scatterplot with 1,000 marks
- Building plan with offices
- How to label data objects?
- Labeling objectives
  - Readable
  - Non-ambiguous relation to graphical object
  - Does not hide pertinent information
- Excentric labeling approach
  - On-demand labeling of adjacent items in focus
  - Dwell time 1s
  - Cursor-centered circle defines neighborhood region
  - Quick flick of cursor to (temporarily) end labeling mode
- Demo



Ludwig-Maximilians-Universität München LMU

Excentric Labeling

- Usability Evaluation: 8 participants, counter-balanced within-subjects design
- Building map application – "Is <name> in the neighborhood of one of the red dots?" (8 tasks)
- Independent variable: Excentric labeling (without zoom) vs zooming in on dots and labels
- Dependent variables: user performance time, errors
- 60% speed advantage for excentric labels (redraw times for zooming / panning were discarded)
- Small error rate for both tools
- Zoom interface (Observation + think-aloud protocol)
  - Appreciated
  - Felt more confident about their findings
  - Zoom interaction was time-consuming and tiring / navigation problems
- Excentric labeling (Observation + think-aloud protocol)
  - Quickly learned technique and search strategy (hopping in discrete steps)
  - Annoyed by continuous updates of the labels while moving
  - Looked at the same labels several times

Outline

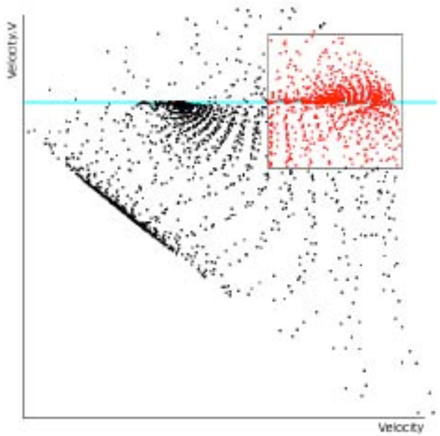
- InfoVis & Interaction
- Direct Manipulation (DM)
- Common Interaction Techniques
  - Brushing
  - Zooming & Panning
  - Dynamic Queries
- Attribute Explorer
- Brushing Histograms vs. DQ
- Dynamic Queries and Movable Filters

Common Interaction Techniques

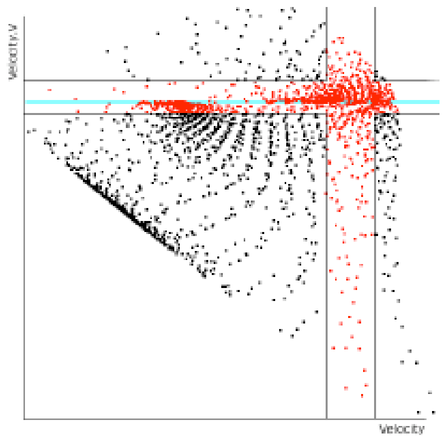
- Details-on-demand
  - Provides improved scalability by displaying information about data case(s) on demand to the user
  - View may move from aggregation of objects to the elements contained
- Direct Walk
  - Linkage between cases
  - Exploring one case may lead to another (e.g. hyperlinks on news page)
- Manipulate View
  - Rearrange view (e.g. move view position, sorting items in a table)
  - Change representation (e.g. from histogram to scatterplot)
- Linking
  - Connection between multiple views of the same data space
  - Updating one view means updating all

Brushing

- Becker & Cleveland 1987
- A collection of dynamic methods for viewing multidimensional data
- Brush is an interactive interface tool to select / mark subsets of data in a single view, e.g. by sweeping a virtual brush across items of interest
- Given linked views (e.g. scatterplot matrix) the brushing can support the identification of correlations across multiple dimensions
- Usually used to visually filter data (via highlighting)
- Additional manipulation / operations may be performed on the subsets (masking, magnification, labeling etc.)
- Different types of brushes (Hauser et al. 2002))
  - Simple brush via sweeping
  - Composite brush: composed multiple single-axis brushes by the use of logical operators
  - Angular brush
  - Smooth brush



AND-brush

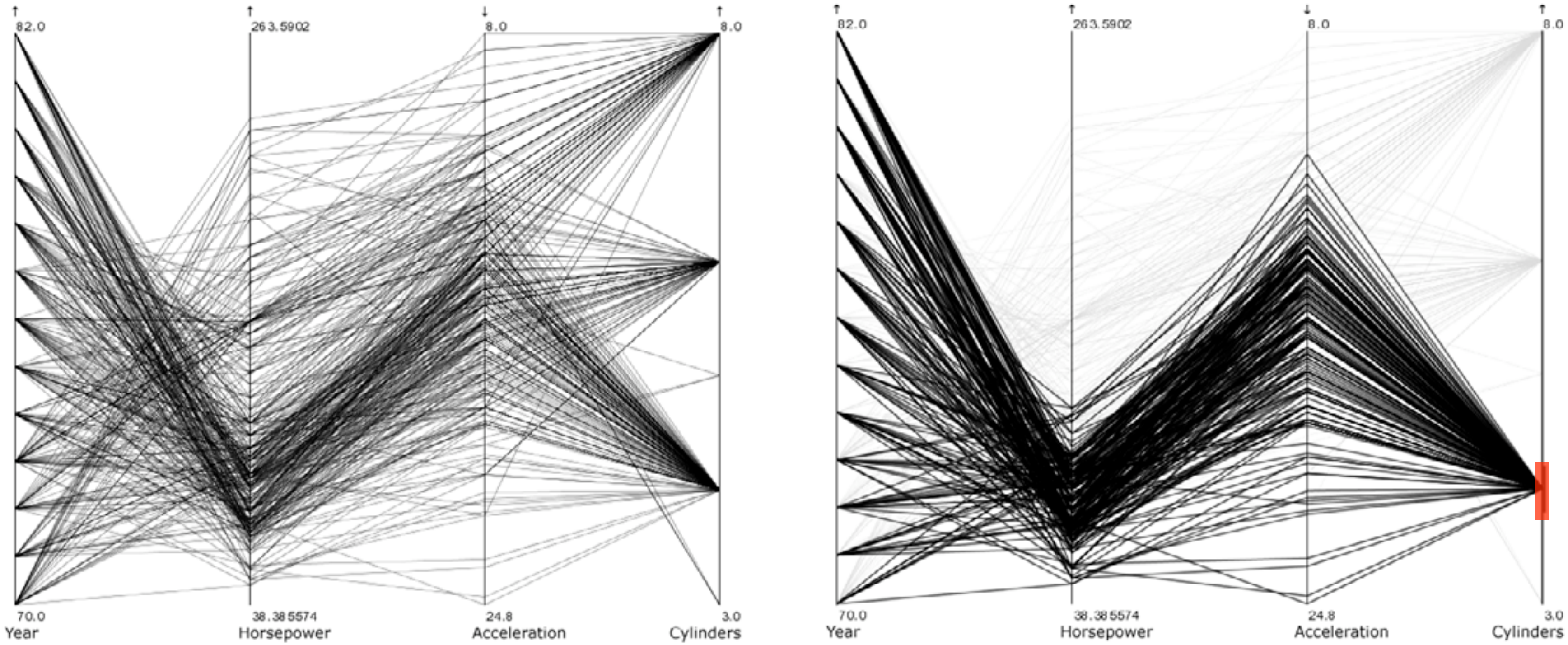


OR-brush

Composite scatterplot brushes - Hauser et al. 2002

Brushing Example

- Brushing one dimension in parallel coordinates to highlight car data objects with 4 cylinders

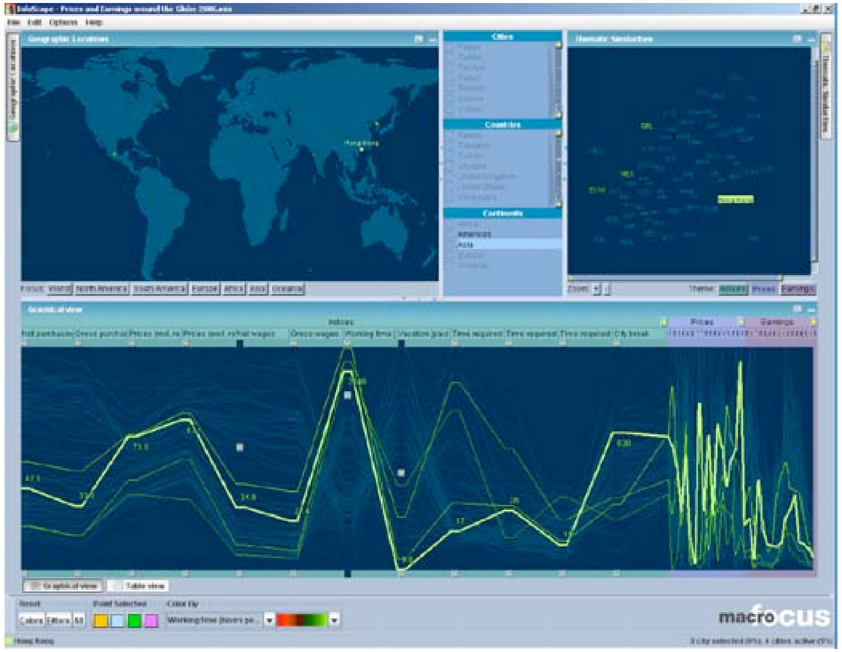


Hauser et al. 2002

Ludwig-Maximilians-Universität München

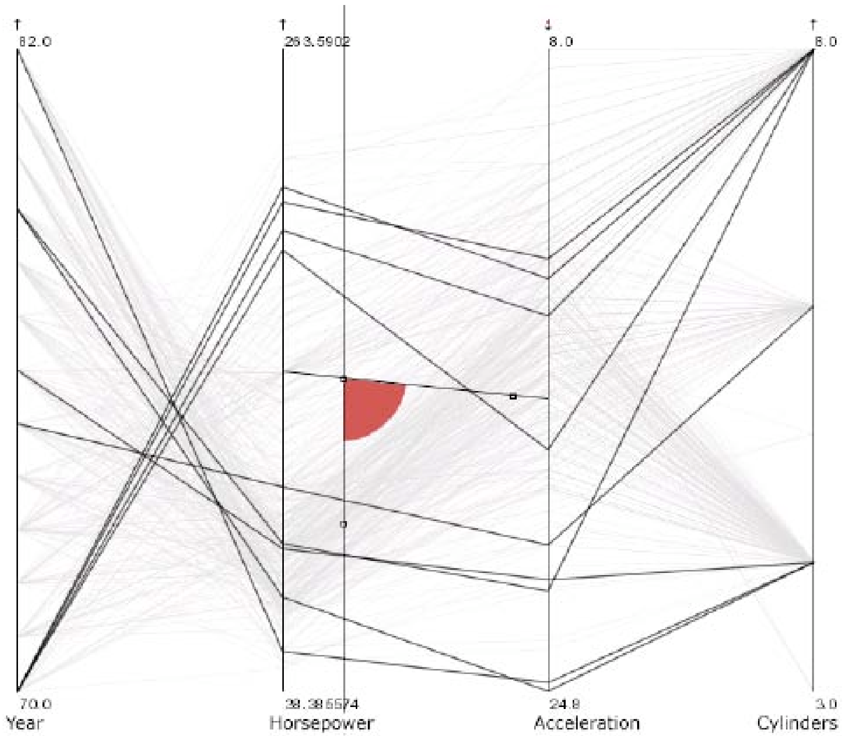
Brushing Example

- Example for composite (AND) brush in Parallel Coordinate Plot – find the cities with high wages, small prices and many paid holiday days
- **Demo** InfoScope: http://www.macrofocus.com/public/products/infoscope.html (free trial and applet)

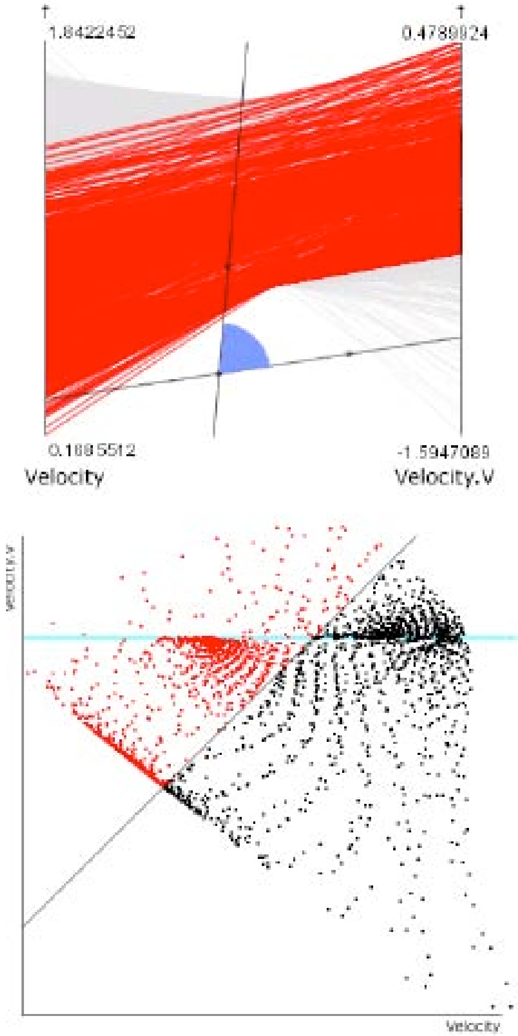


Angular Brush

- Angular brush: brushing by specifying a slope range – highlight correlation and outliers between two dimensions



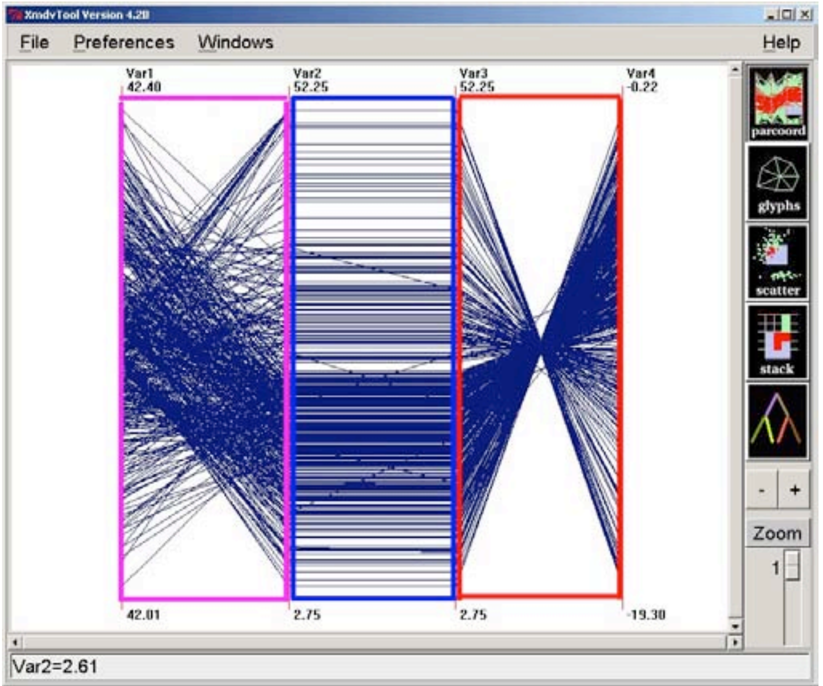
Hauser et al. 2002

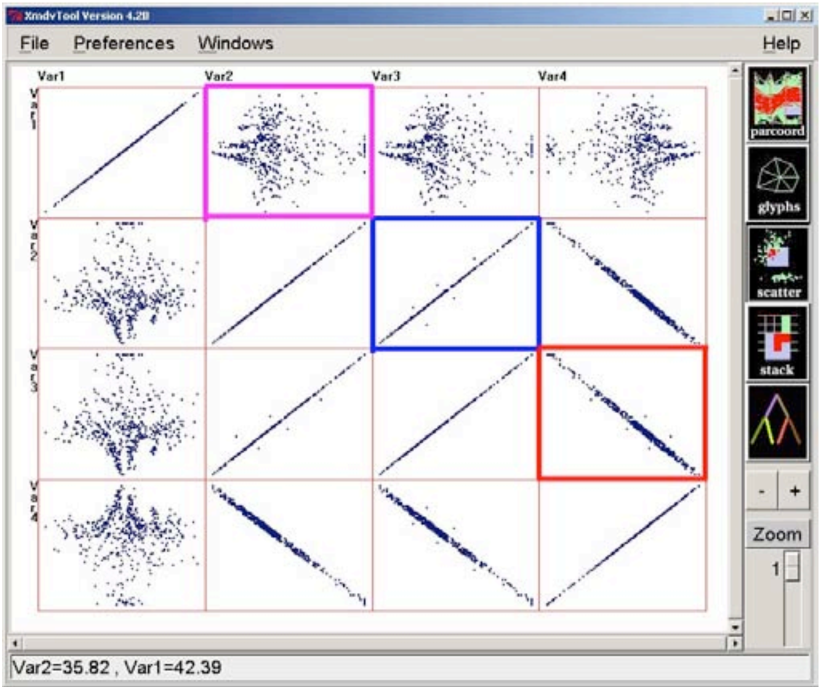


angular brush

PCP - Correlation Patterns

- Var1-Var2 has no correlation; Var2-Var3 has very strong positive correlation; Var3-Var4 has very strong negative (inverse) correlation



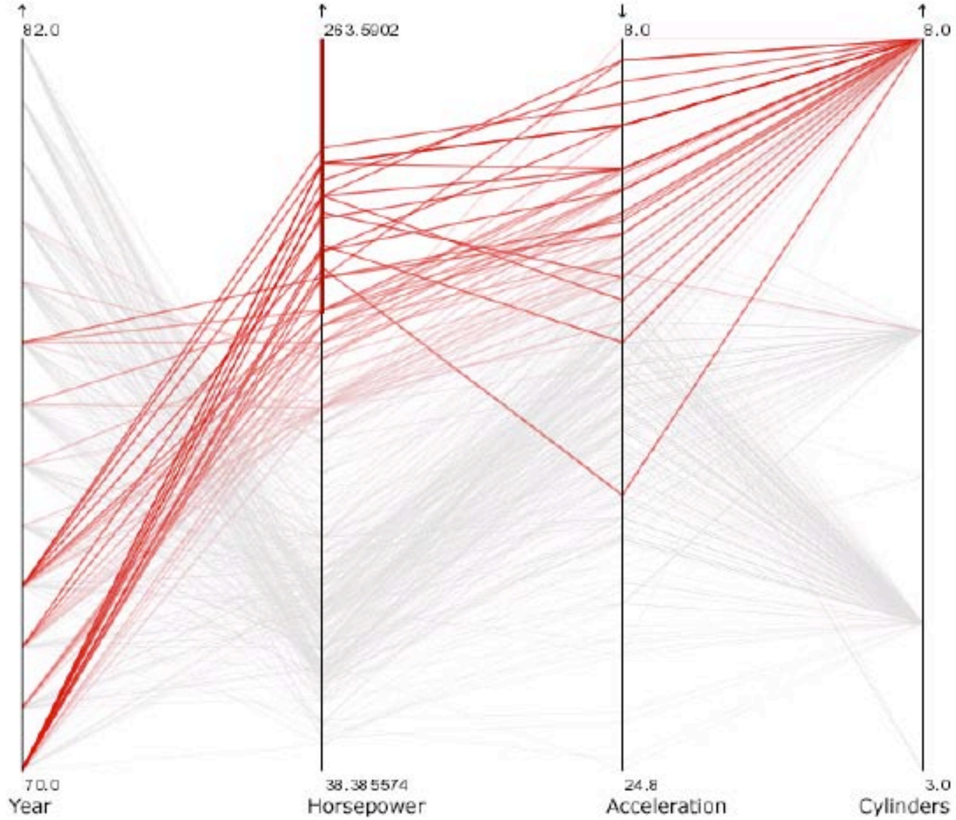


http://www.evl.uic.edu/aej/526/kyoung/Training-parallelcoordinate.html

Ludwig-Maximilians-Universität München

Smooth Brush

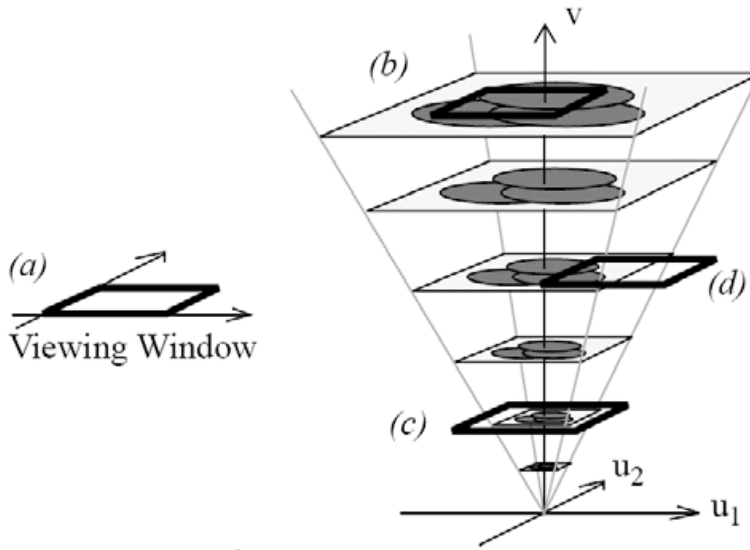
- Non-binary brushing
- Degree-of-interest defined by distance to brushed range
- Decreasing degree is mapped to decreasing drawing intensity



Hauser et al. 2002

Zooming & Panning

- Moving from overview to detail: another way to filter data / focus on a subset of data
- Scale and translation of the viewport
- Geometrical versus semantic zooming
- Topic of a lecture to come…



Furnas & Bederson 1995

Dynamic Queries

- Shneiderman 1994
- Explore and search databases
- SQL example: SELECT customer_id, customer_name, COUNT(order_id) as total FROM customers INNER JOIN orders ON customers.customer_id = orders.customer_id GROUP BY customer_id, customer_name HAVING COUNT(order_id) > 5 ORDER BY COUNT(order_id) DESC
- Problems
  - Takes time to learn
  - Takes time to formulate and reformulate
  - User must know what she is looking for – only exact matches
  - Lots of ways to fail
  - SQL error messages helpful?
  - Zero hits – what component is to be changed?

Ludwig-Maximilians-Universität München LMU

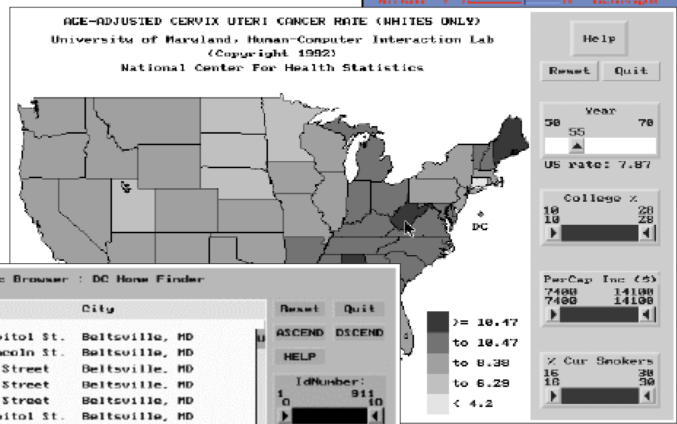
Dynamic Queries

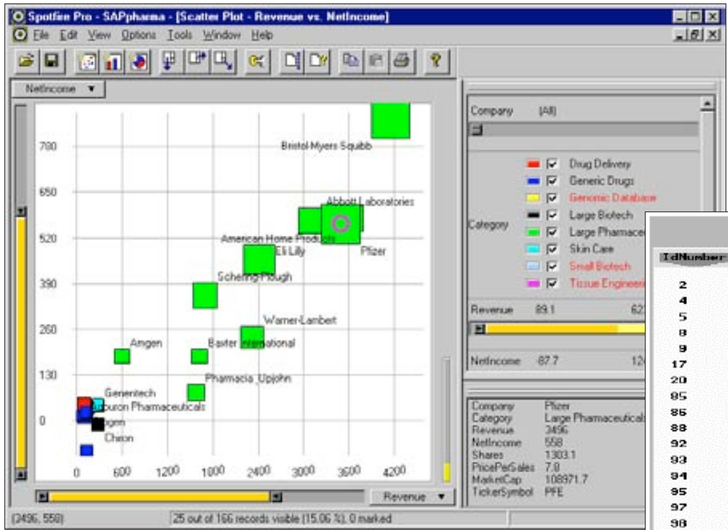
- Based on Direct Manipulation (DM)
- DM principles with regard to Dynamic Queries
  - Visual presentation of the query's components
  - Visual presentation of results
  - Rapid, incremental, and reversible control of the query
  - Selection by pointing, not typing
  - Immediate, continuous feedback
- Implementation approach
  - Graphical query formulation: Users formulate queries by adjusting sliders, pressing buttons, bounding box selection...
  - Search results displayed are continuously updated (< 100 ms)

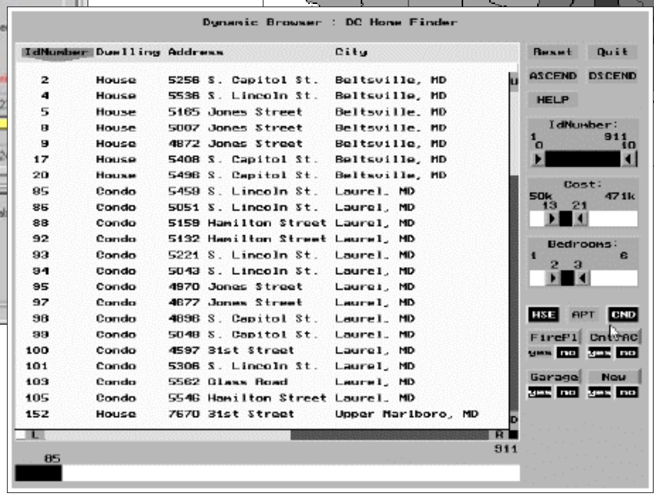
Ludwig-Maximilians-Universität München

Examples

- Visual representations of data to query?
- Some examples: geographic data, starfields, tables etc.





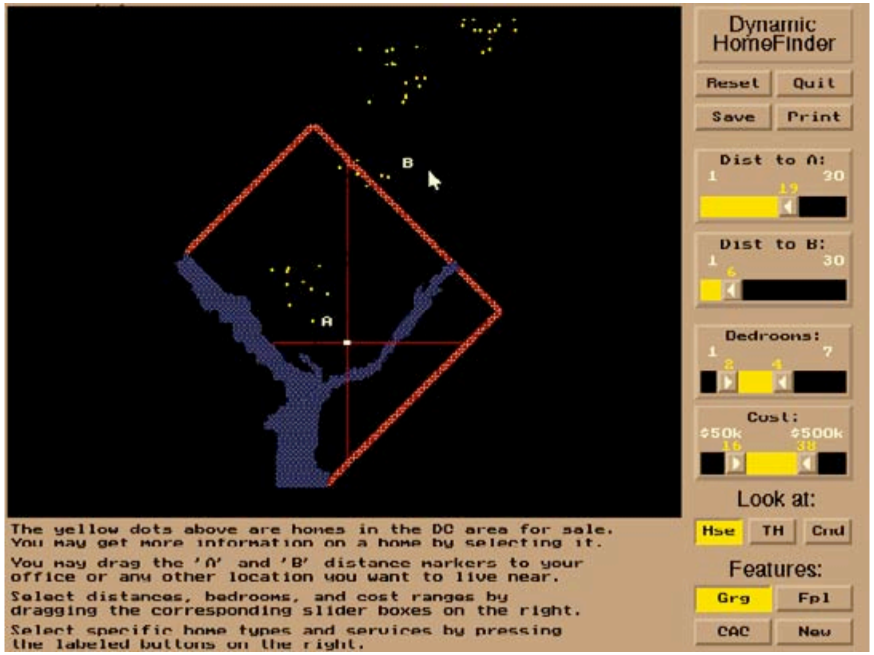


Shneiderman 1994

Ludwig-Maximilians-Universität München LMU

HomeFinder

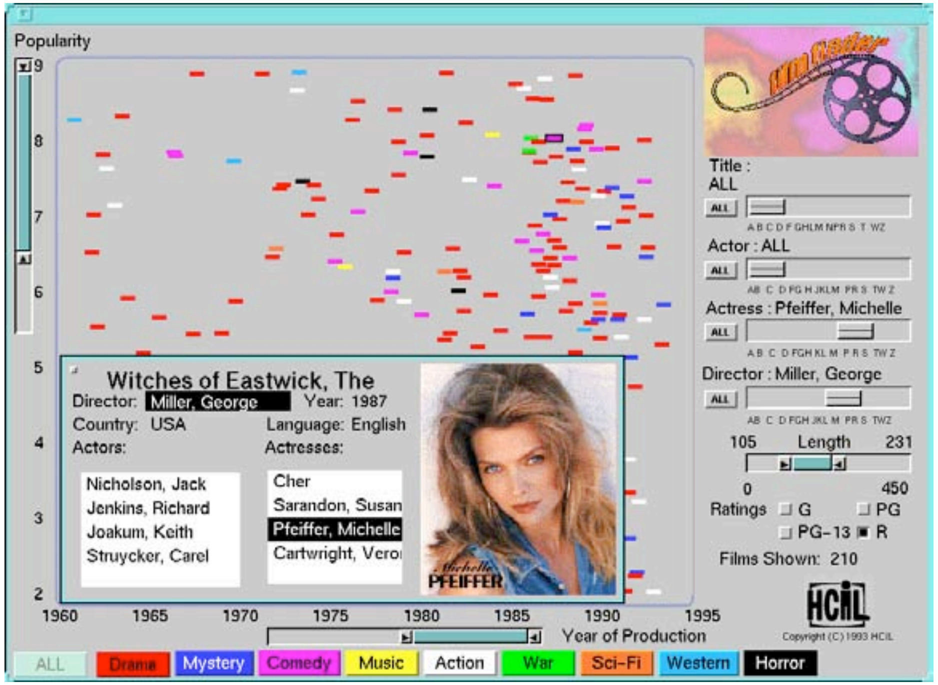
- One of the first DQ interfaces
- Williamson & Shneiderman **1983(!)**
- Demo



Ludwig-Maximilians-Universität München LMU

FilmFinder

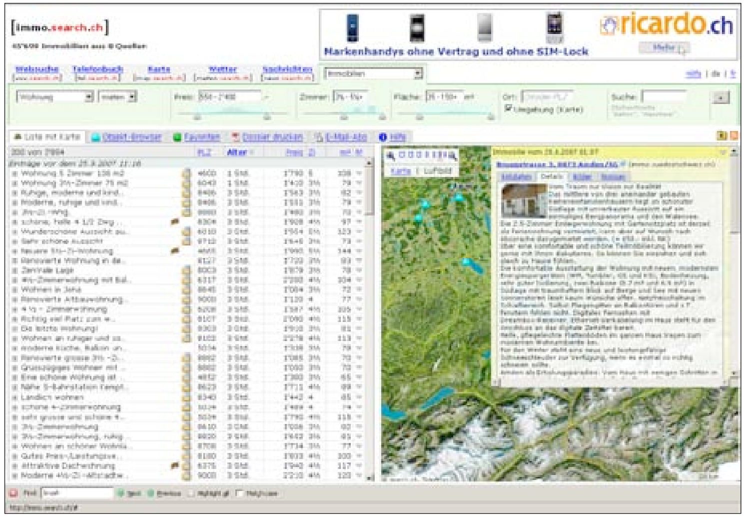
- Ahlberg & Shneiderman 1994
- Movie

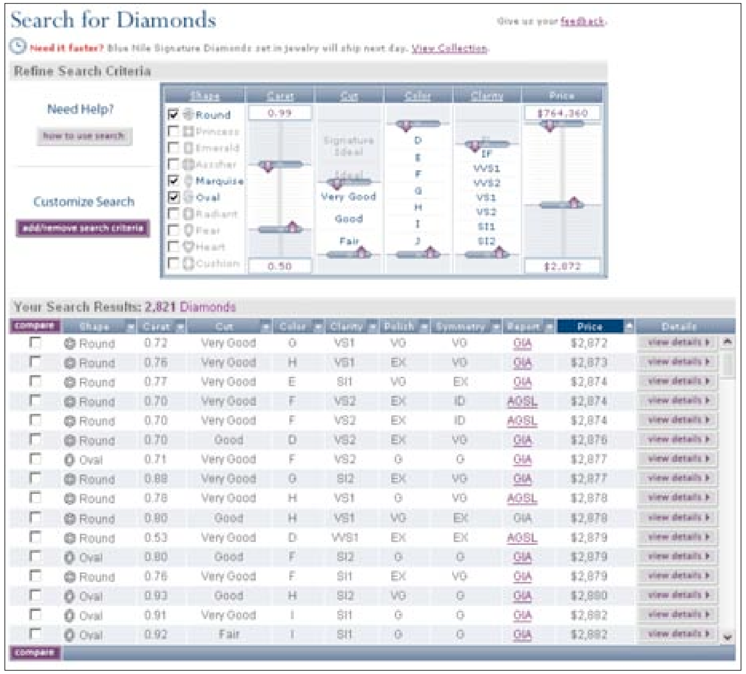


Ludwig-Maximilians-Universität München LMU

Dynamic Queries Online

- Online examples: immo.search.ch and diamond search (http://www.bluenile.com)

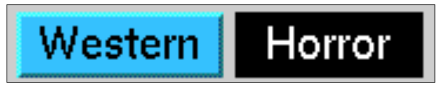


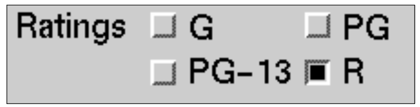


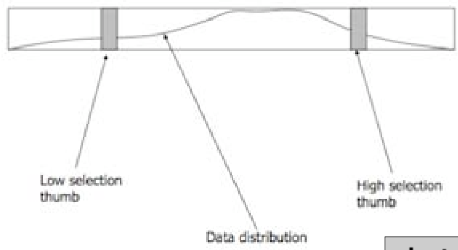
Ludwig-Maximilians-Universität München LMU

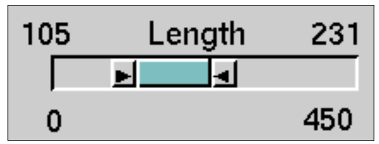
Dynamic Query Controls

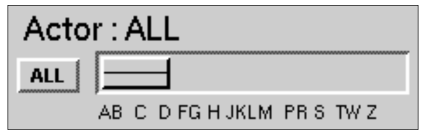
- Check boxes and buttons (Nominal with low cardinality)
- Sliders and range slider (ordinal and quantitative data)
- Alphaslider (ordinal data) (Ahlberg & Shneiderman 1994)
    - Small-sized widget to search sorted lists
    - Online-text output
    - Two-tiled slider thumb for dragging operations with different granularities
    - Letter index visualizing the distribution of initial letters – jump to a position in the slider
    - Locating an items out of a list of 10,000 items ~ 28s for novice users
    - Pros and cons to text entry?
- Redesigned Alphalsider for PDAs / MP3 player - **movie**
- Extend data sliders with data visualization (Eick 1994)

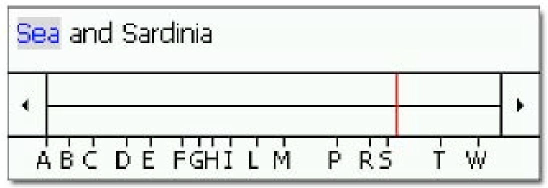












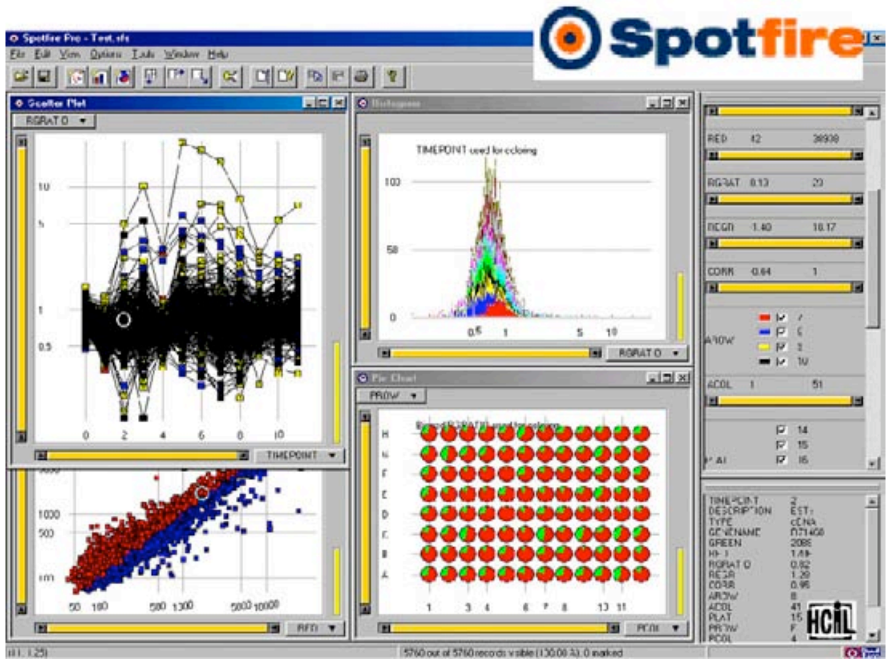
Ludwig-Maximilians-Universität München LMU

Summary Dynamic Queries

- Users can rapidly, safely playfully explore a data space – no false input possible
  - Users can rapidly generate new queries based on incidental learning
  - Visual representation of data supports data exploration
  - Analysis by continuously developing and testing hypotheses (detect clusters, outliers, trends in multivariate data)
  - Provides straightforward undo and reversing of actions
- Potential problems with DQ as implemented in the FilmFinder?
  - Limit of query complexity – filters are always conjunctive
  - Performance is limited for very large data sets and client / server applications
  - Controls require valuable display space
  - Information is pruned
  - Only single range queries and single selection in the alphaslider

Dynamic Queries

- Starfield displays and Dynamic Queries provided the basis for SpotFire
- Christopher Ahlberg
  - 1991: Visiting student from Sweden at the HCIL University of Maryland
  - 1996: Founder of SpotFire
  - 2007: SpotFire was sold for 195 Mio. $
- Well done!



Outline

- InfoVis & Interaction
- Direct Manipulation (DM)
- Common Interaction Techniques
  - Brushing
  - Zooming & Panning
  - Dynamic Queries
- Attribute Explorer
- Brushing Histograms vs. DQ
- Dynamic Queries and Movable Filters

Ludwig-Maximilians-Universität München

Attribute Explorer

- Tweedie et al. 1994
- Example for DQ, brushing & linking and fuzzy search
- Linked histograms to search and explore multivariate data
- Filtering data via range sliders
- Color-coding to highlight and discriminate data cases across views
- Sensitivity information: visualizes how well data cases meet the filter requirements
- Particularly useful for zero-hits situations
- Movie

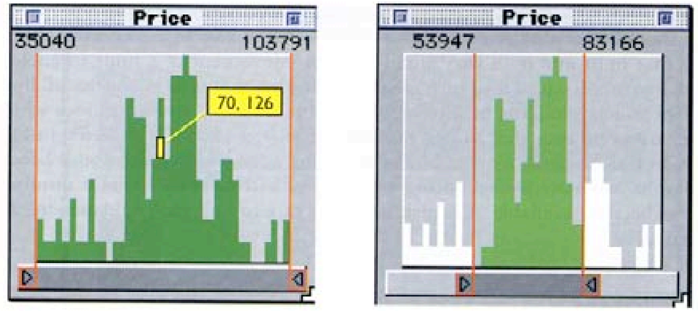
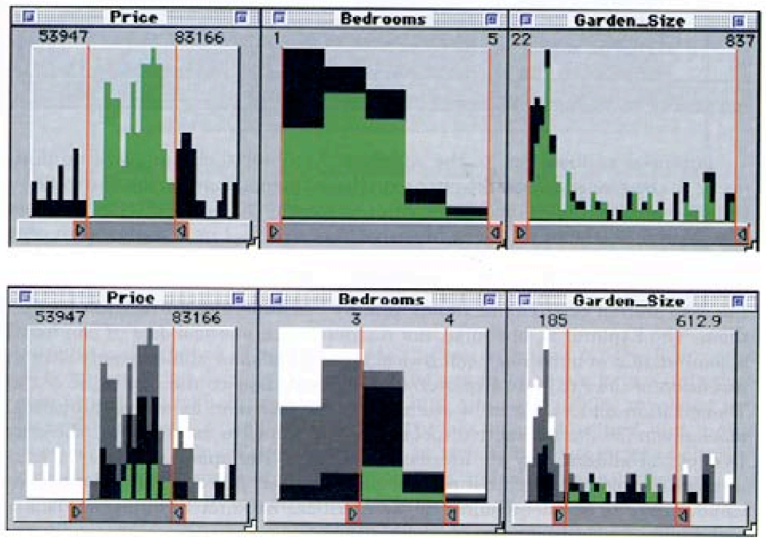


FIGURE 3.55 A histogram representing the prices of a collection of houses. The contribution of one house is shown in yellow

FIGURE 3.56 Limits on Price identify a subset of houses, coded green



Spence 2004

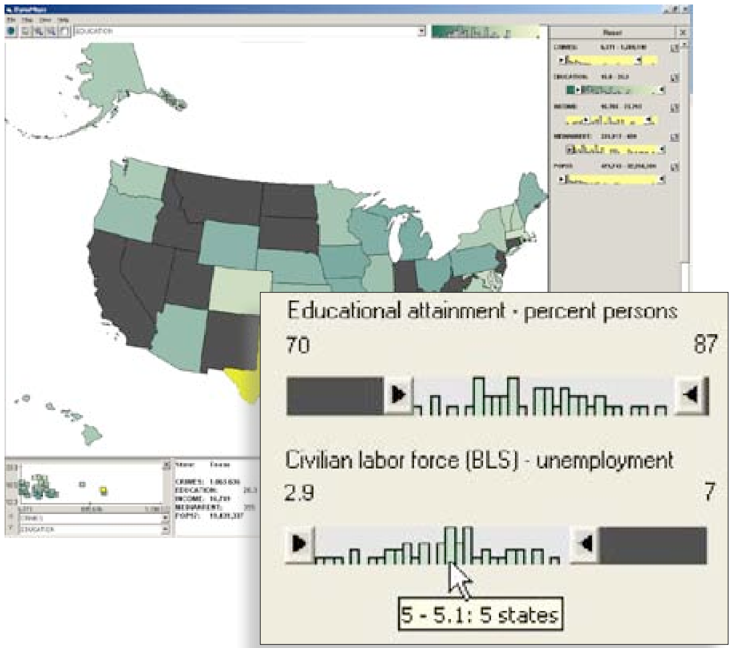
Outline

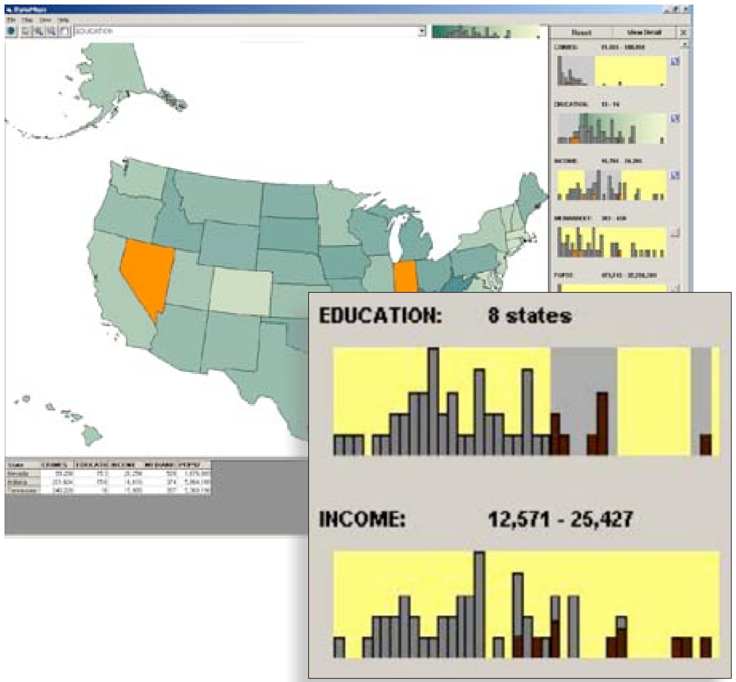
- InfoVis & Interaction
- Direct Manipulation (DM)
- Common Interaction Techniques
  - Brushing
  - Zooming & Panning
  - Dynamic Queries
- Attribute Explorer
- Brushing Histograms vs. DQ
- Dynamic Queries and Movable Filters

Ludwig-Maximilians-Universität München LMU

Brushing Histograms vs. DQ

- Usability experiment - Li & North 2003
- Demo





Ludwig-Maximilians-Universität München LMU

Brushing Histograms vs. DQ

- Differences of the approaches
  - Filtering vs. highlighting
  - Single range vs. multiple ranges query
  - One directional vs. bi-directional interaction
- Usability evaluation with 36 students
  - Independent variables: type of query tool, type of task
  - Dependent variables: user performance time, errors, user satisfaction ratings
  - Within subjects, counterbalanced design

Li & North 2003

Task Name	Description
Single range	Finding states within a single range for a given attribute. Example: How many states have the population between 20 and 25 millions in 1996?
Multiple ranges	Finding states within multiple ranges for a given attribute. Example: List the number of states with population in the following ranges: 6.3 – 10 millions, 6.3 – 14 millions and 6.5 – 18 millions.
Multiple criteria	Finding states according to different ranges on multiple attributes. Example: How many states have the number of farms within 28,000 – 85,000 and the population more than 10 millions?
Attribute correlation	Discovering the correlation between two attributes. Example: What's the relationship between educational attainment and personal income? Potential answers include: no relationship, direct proportion or inverse proportion.
Compare	Comparing states according to multiple criteria. Example: Given three states, which one has the lowest median rent?
Evaluate trend	Evaluating the trend of a particular state in the global context. Example: What kind of state is Florida in the United States? The potential answer could be that Florida had relatively higher population and median level of income compared with other states.

Brushing Histograms vs. DQ

- Dynamic Query Sliders
  - More efficient for simple range and criteria tasks
  - Users found them easier to control and less confusing (no additional feedback on the other histograms)
- Brushing histograms
  - Faster for complex tasks: trend evaluation, attributes relation and compare tasks
  - Took more time to learn, problems with accuracy of interaction

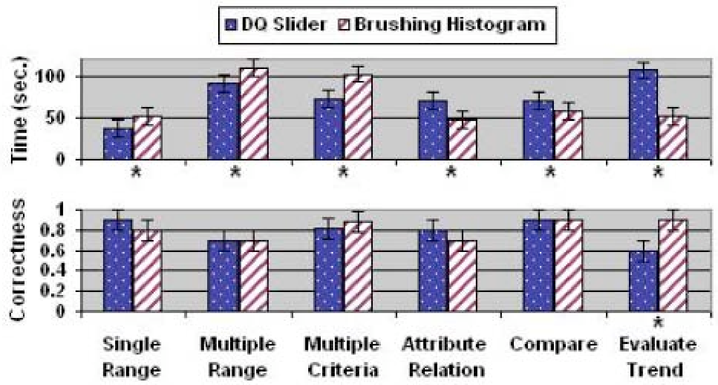


Figure 11: Mean user performance time and correctness for each task and query tool. Asterisks indicate significant difference at $p<0.05$. Correctness: 1 = right, 0 = wrong

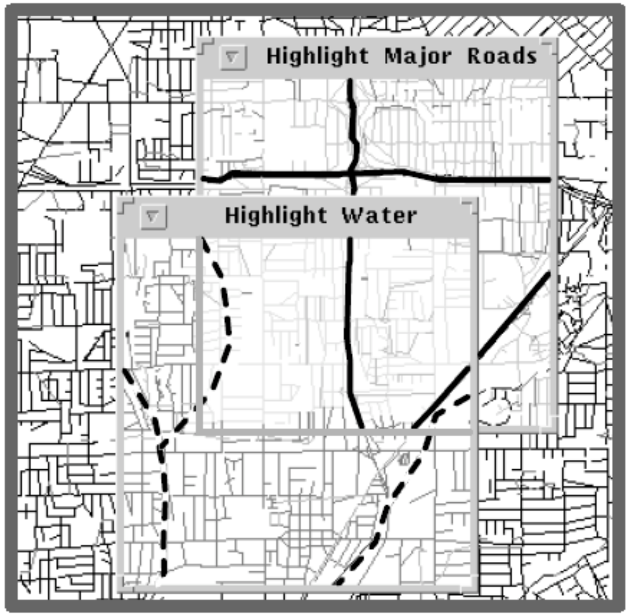
Li & North 2003

Outline

- InfoVis & Interaction
- Direct Manipulation (DM)
- Common Interaction Techniques
  - Brushing
  - Zooming & Panning
  - Dynamic Queries
- Attribute Explorer
- Brushing Histograms vs. DQ
- Dynamic Queries and Movable Filters

Dynamic Queries and Movable Filters

- Fishkin and Stone 1995
- Dynamic Queries (DQ)
  - Disjunctive queries can only be performed by sequential querying
  - Effect of DQ is global – no way to limit filtering to only a portion of the data
  - Number of possible queries is fixed in advance
- Combine approach with magic lens filters
  - Arbitrarily-shaped region with an operator that manipulates the view of underlying objects
  - Filters are spatially bounded – global context is maintained
  - Filters that overlap compose their effects in the overlap region



Stone et al. 1994

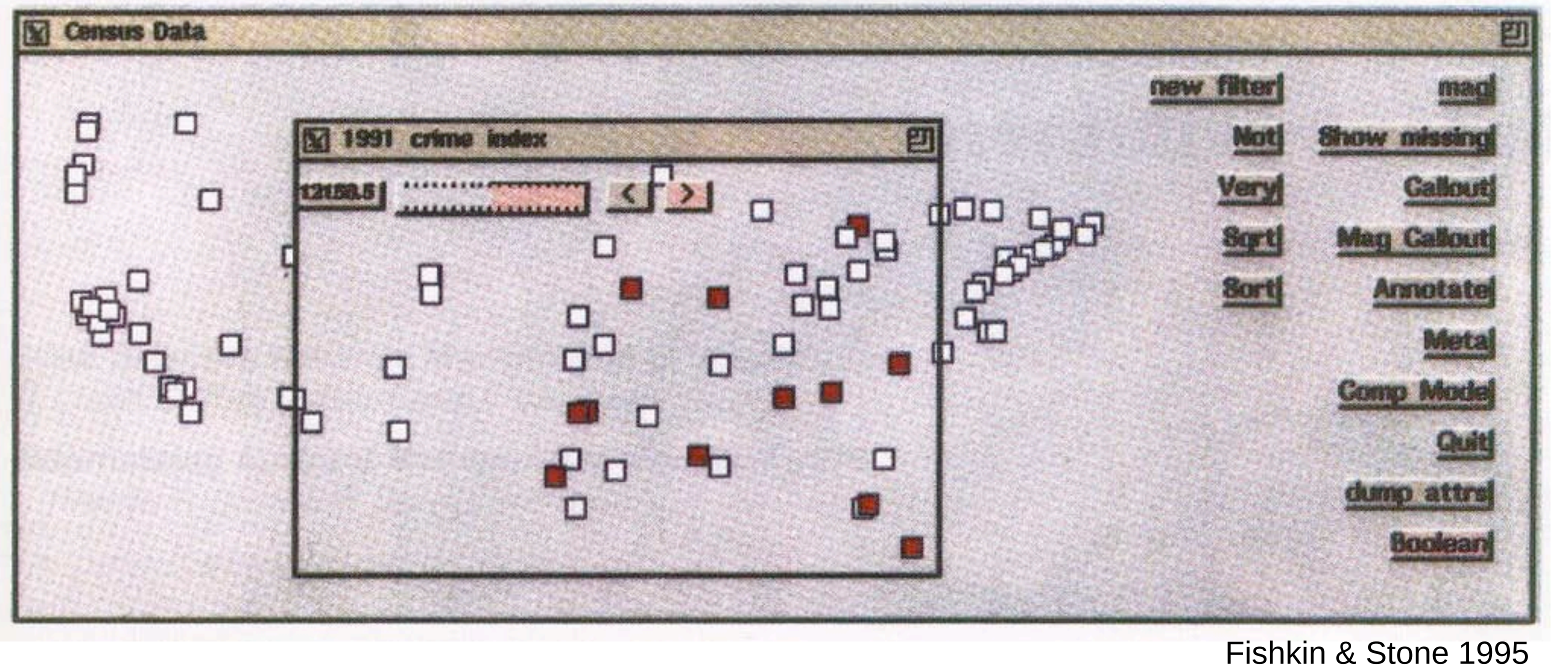
Idea & Implementation

- Each lens acts as a filter that screens on some attribute of the data
- Lens components
  - Filtering function (what to filter)
  - Composition mode (how to combine the filter result with lenses underneath, i.e. AND, OR, NOT)
- Composition modes are implemented as buttons on the lens
- Grouping: Replace a stack of lenses by a single compound lens, which also has a composition mode
- Compound lenses may contain other compound lenses
- Boolean queries and grouping allow queries of arbitrary complexity
- Multiple concurrent queries on different portions of the data space

Ludwig-Maximilians-Universität München LMU

Simple Range Filter

- Example: US census data, each box represents a city (position mapped to physical location)
- Lens filter (Crime index 1999) covers the center of the country
- Slider to manipulate the value of the query
- Arrow buttons show the direction of the query, i.e. screen data for less than or greater than the slider value
- Red-coded cities pass the filter



Fishkin & Stone 1995

Composition Modes

- AND (conjunctive), OR (disjunctive)
- SELF: lens only displays the effect of its own filter; other lenses are ignored
- NOP: filter effect of lens is disabled

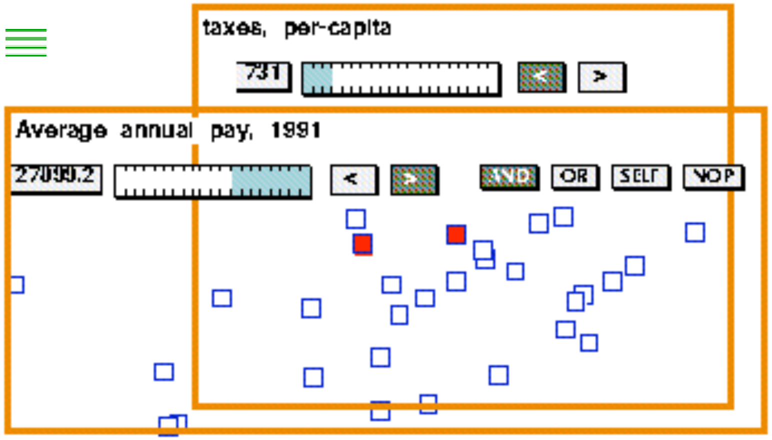


Figure 1(a) High salaries AND low taxes.

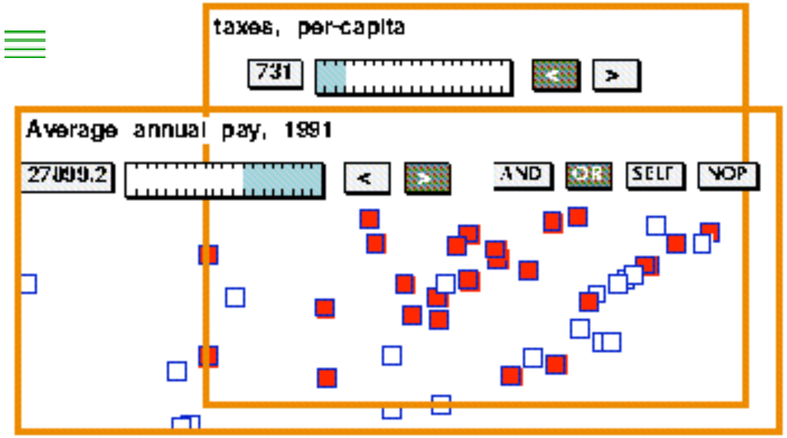


Figure 1(b) High salaries OR low taxes. Both conjunctive (AND) and disjunctive (OR) queries are incorporated in our system.

Fishkin & Stone 1995

Alternate Views

- Lenses to generate alternate views of the data
- Magnification, verbal description, sorted views etc.
- Cities listed without boxes are missing the value for the filter attribute (missing data)

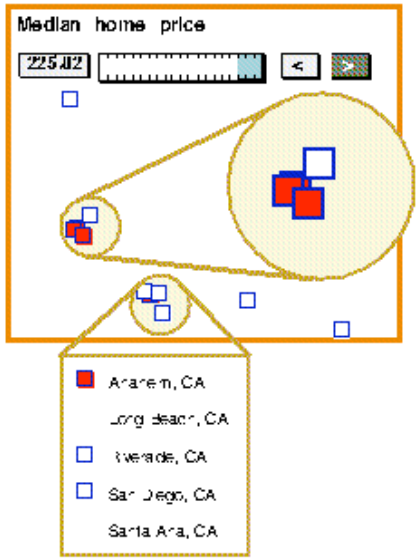


Figure 2. Semantic filters can be augmented with visual filters. Here, a magnifying lens and a call-out lens show clumped cities while maintaining context elsewhere.

Fishkin & Stone 1995

Ludwig-Maximilians-Universität München LMU

Local Effects

- “Which cities in California and Texas have relatively low housing prices?”
- With Dynamic Queries we can filter data by global prices (range slider affecting the entire starfield)
- Problem: Houses on the west coast are typically more expensive than houses in the midwest!
- Movable filters allow for concurrent queries on the two areas

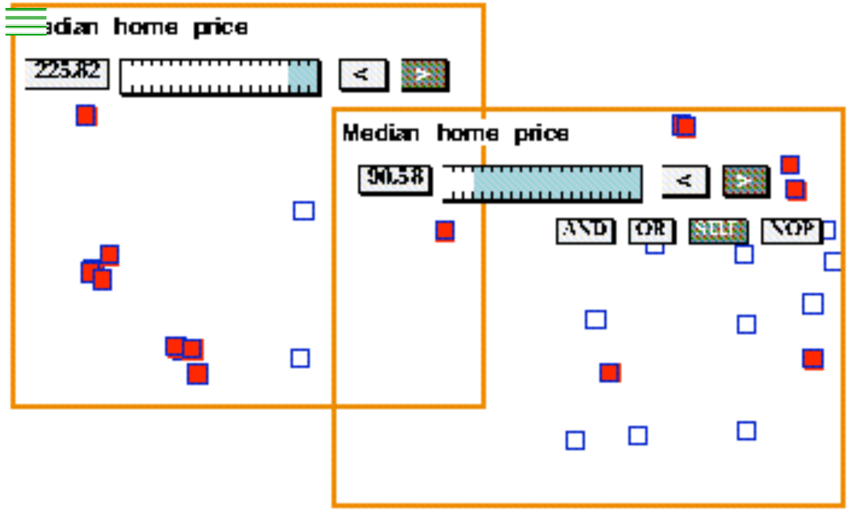


Figure 3. To find relatively high housing prices in California and Texas, two different filters are positioned simultaneously.

Fishkin & Stone 1995

Ludwig-Maximilians-Universität München

Real-Valued Queries

- Assign a real valued score (range [0...1]) to each datum
  - Cases with a score of 0 fail the filter entirely
  - Cases with intermediate scores partly satisfy the filter
  - Cases with a score of 1 entirely satisfy the filter
- The higher the score the higher is the box filled with red color

Fishkin & Stone 1995

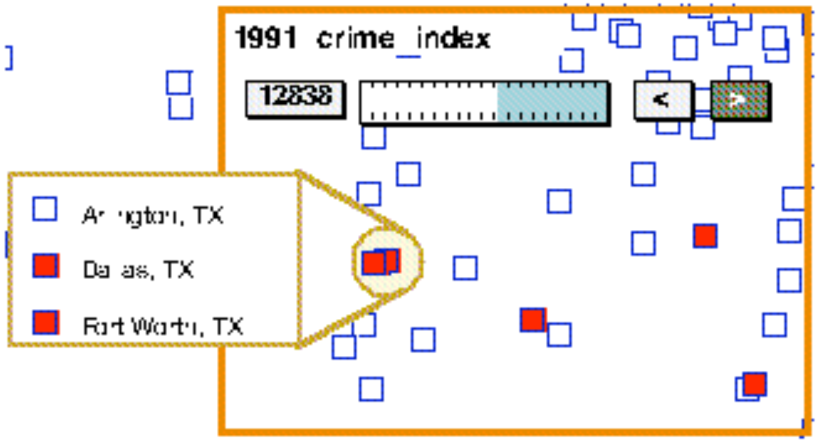


Figure 4(a) boolean query on crime rate for three cities in Texas.

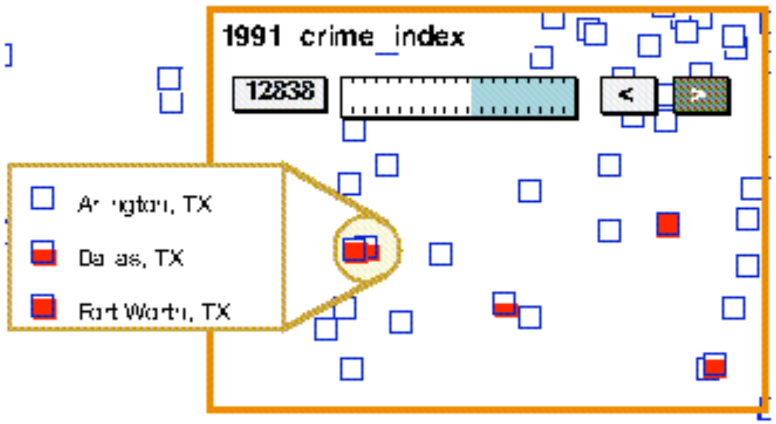
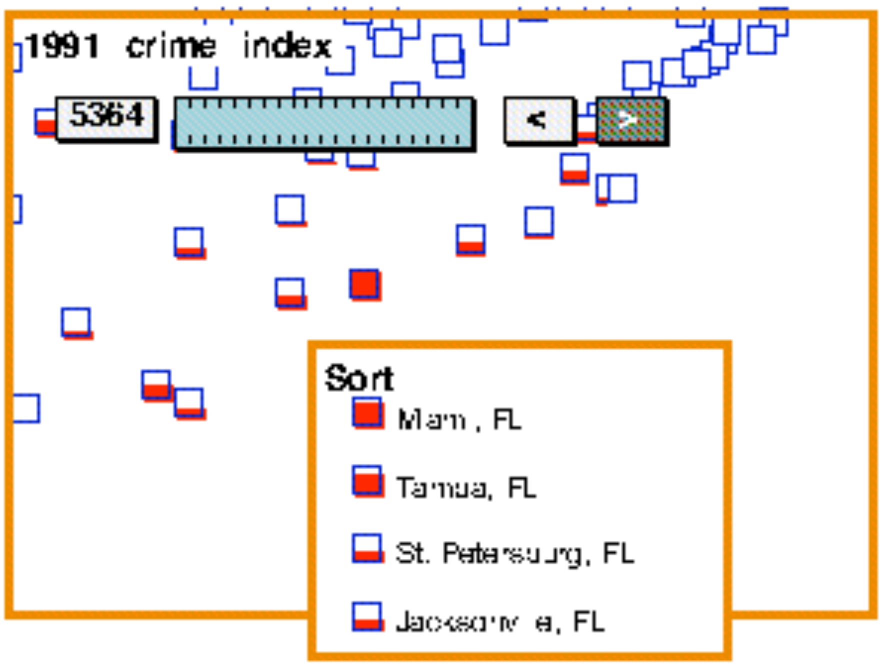


Figure 4(b) Real-valued query on crime rate for the same cities. Extending our filters from boolean-valued to real-valued allows distinctions to be maintained.

Ludwig-Maximilians-Universität München LMU

Real-Valued Queries

- Real-valued query lens overlaid by a sorting lens



Fishkin & Stone 1995

Figure 5. A sorting lens sorts cities by crime rate in Florida.

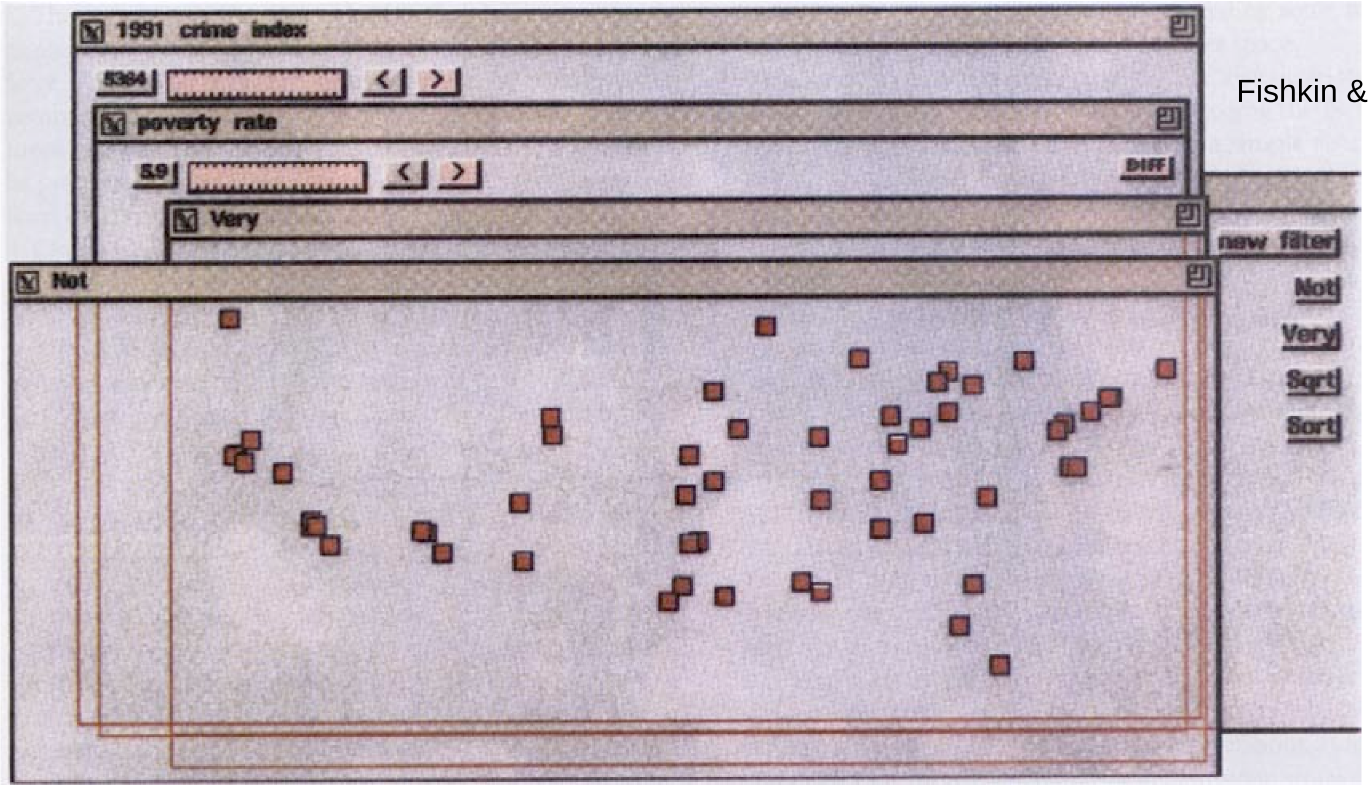
Real-valued Composition Modes

- Real-valued composition modes and operators
  - MIN and MAX: work the same on 0.0 and 1.0 as AND and OR on 0 and 1, but can also incorporate values in-between
  - NOT: returns for each case 1.0 minus the case's input value
  - Fuzzy logic operators: e.g. "very", "somewhat", "more or less"
  - Mathematical operators: e.g. difference ,log
- Example
  - Is crime rate and poverty positively correlated?
  - One real-valued crime filter
  - One real-valued poverty filter - composition operator DIFF (absolute value of the difference between the two filter outputs)
  - VERY filter (where are the differences very different? Very(x) defined as x^2)
  - NOT filter (where are the differences NOT very different?)

Ludwig-Maximilians-Universität München LMU

Dynamic Queries and Movable Filters

- Filter result: the redder the city, the greater the extent to which poverty and crime rates are NOT VERY DIFFerent – strong positive correlation between poverty and crime rate



Fishkin & Stone 1995

Ludwig-Maximilians-Universität München

Dynamic Queries and Movable Filters

- Missing data is visualized by a special lens

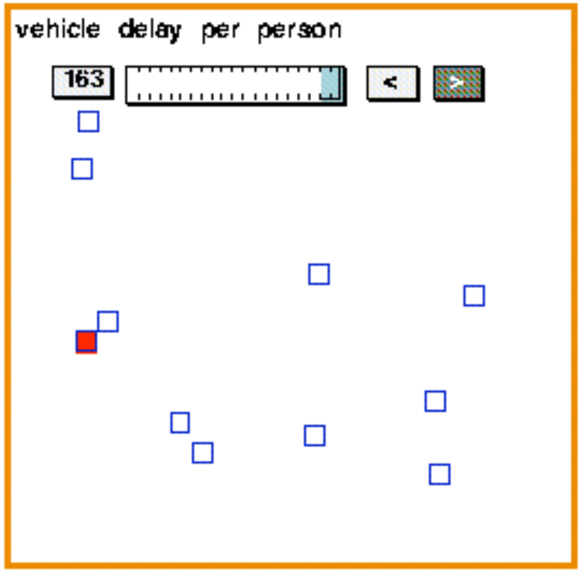


Figure 6(a). A filter finds only one city (San Francisco) with a high score.

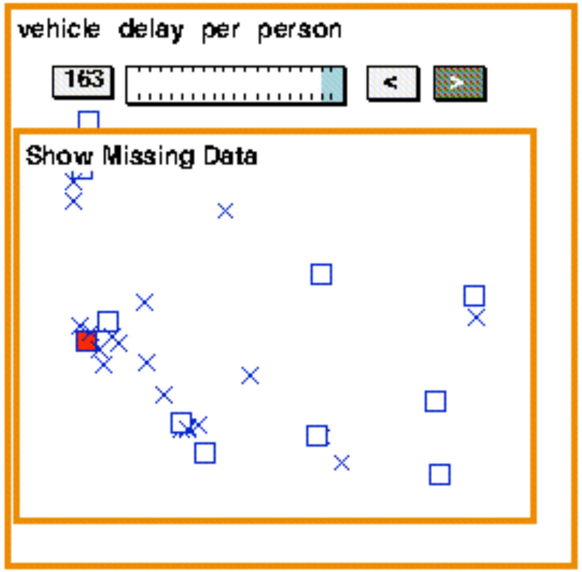


Figure 6(b). A missing data lens shows that attribute values are missing for many cities. Cities with missing data are marked with an 'X'.

Fishkin & Stone 1995

Ludwig-Maximilians-Universität München LMU

Additional Sources

- Alan Dix et al.: Human Computer Interaction., 3. Auflage, 2003.
- Lecture material CS 7450 John Stasko, 2006