



## hD Visualization

CMSC 436/636

**Penny Rheingans**

University of Maryland, Baltimore County

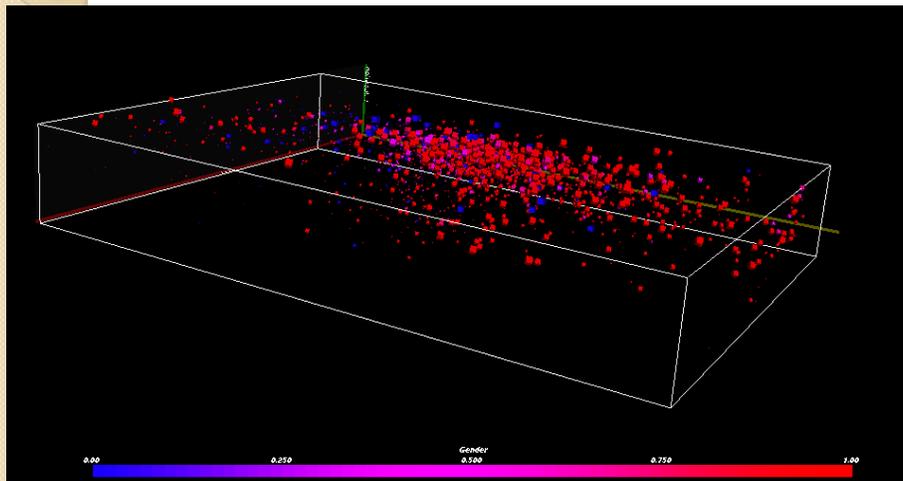


## Information Visualization

- **Characteristics**
  - no spatial context for data
  - generally multivariate data
  - generally high-dimensional
- **Implications**
  - need to assign spatial location to data element
  - need to combine visual vocabulary elements
  - need to reduce dimensionality before visualizing

## Visual Vocabulary ( Revisited)

- Position
- Color
- Density
- Glyphs
- Motion
- Interaction

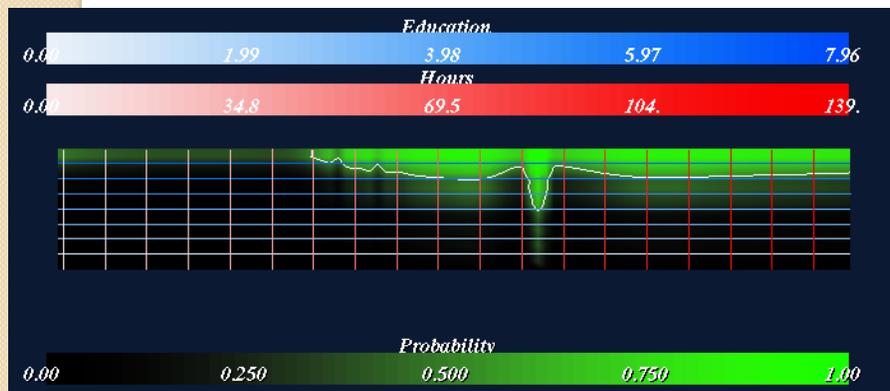


## Dimension Reduction

- Reduce high variable data space to two or three dimensional display space
- Some approaches
  - Feature selection
  - Combined dimensions
  - Similarity clustering

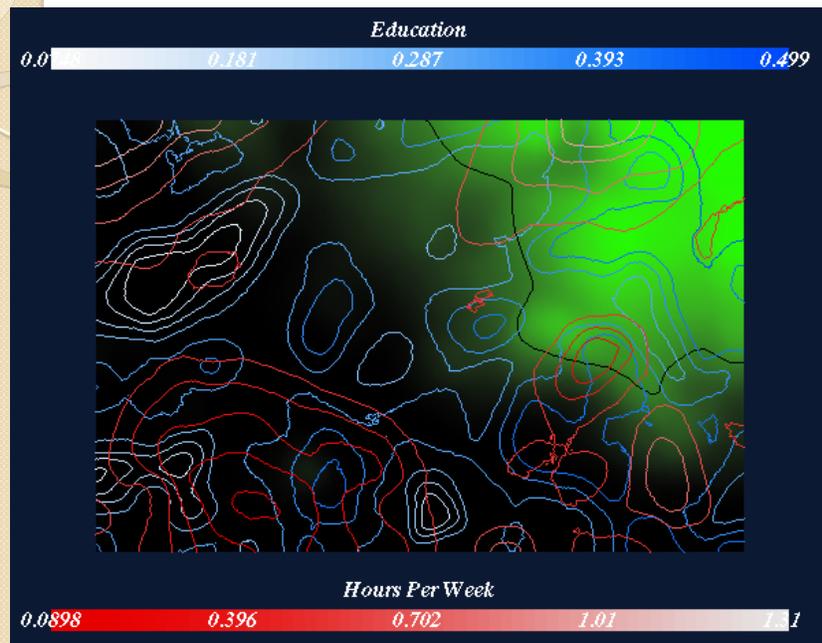
## Feature Selection

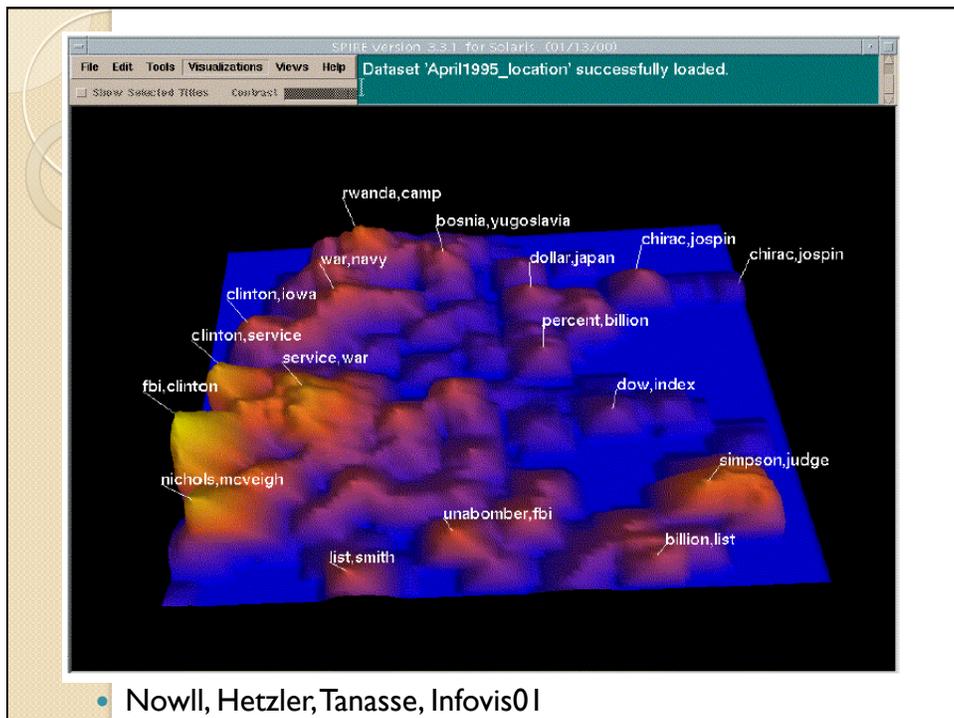
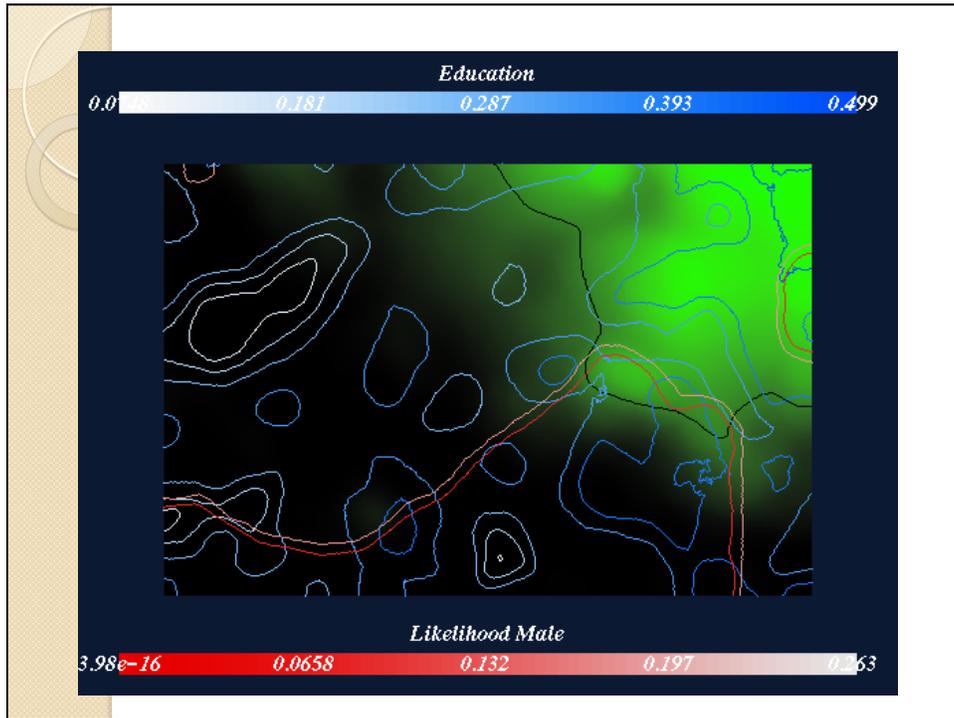
- Advantages: simple to perform  
intuitive to understand  
continuous process



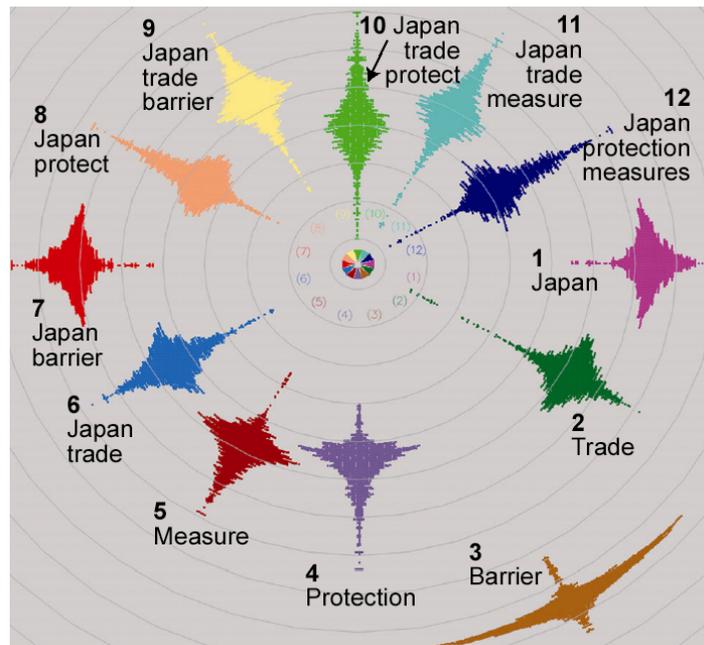
## Similarity Clustering

- Self-organizing maps (SOM, Kohonen97) used to preserve locality in projection
  - instance set used to train map, starting from random codebook vectors
  - each instance mapped to most similar map position and used to reinforce local tendency
  - result is highly non-linear projection where neighboring regions in display space tend to correspond to neighboring regions in data space

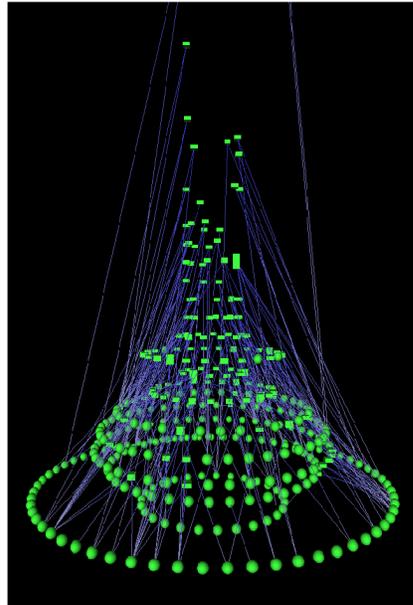
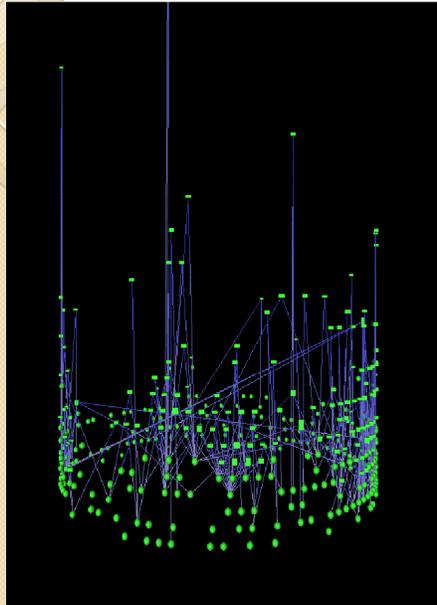




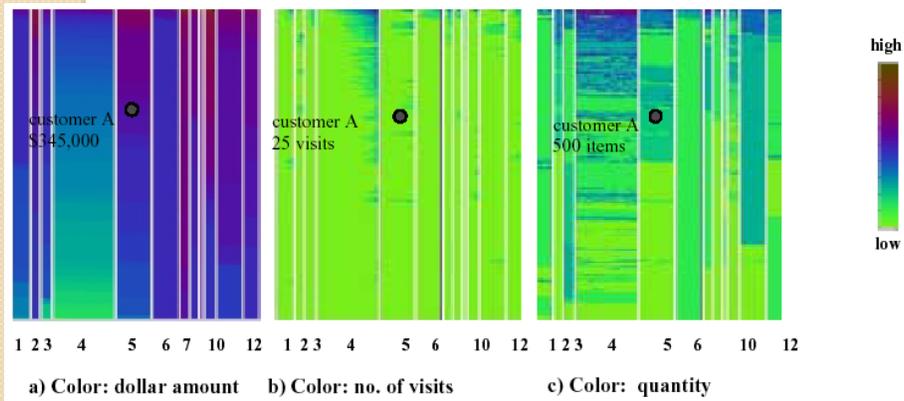
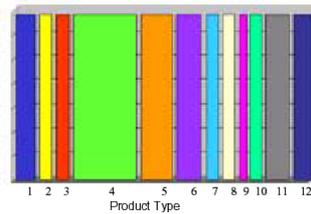
# Non-Euclidean Layouts



• Havre, Hetzler, Perrine, Jurrus, Miller, Infovis01



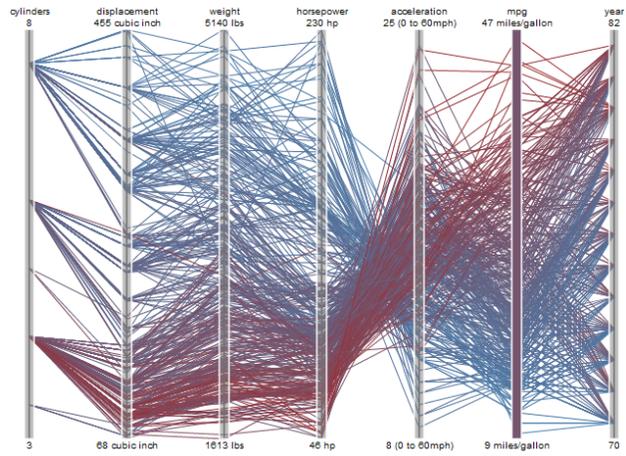
Yang-Pelaez and Flowers '00



• Keim, Hao, Ladisch, Hsu, Dayal, Infovis01

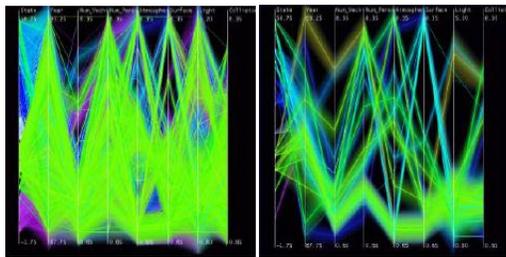
## Parallel Coordinates

- Layout
  - Each variable becomes a vertical axis
  - Each data point becomes a line from axis to axis

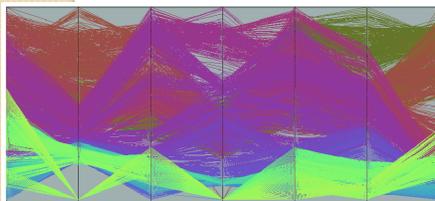


## Parallel Coordinates Issues

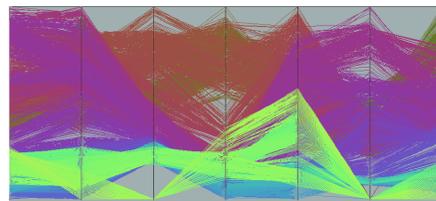
- Density of items



- Order of axes



a. Sequential Arrangement

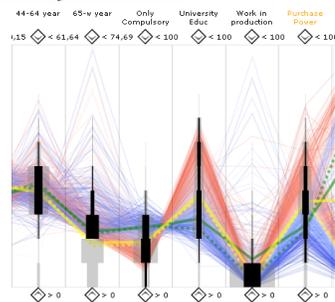
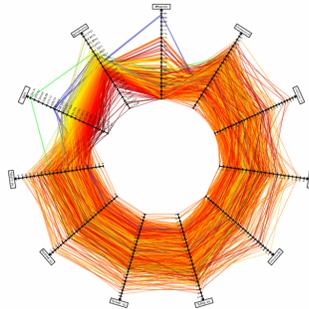


b. Similarity Arrangement

Ankerst, Berchtold, and Keim '98

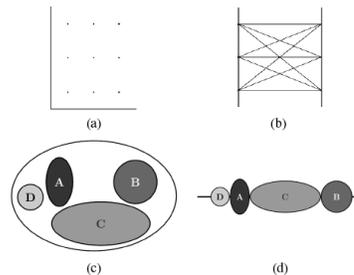
## Parallel Coordinates Variants

- Other layouts
- Summary statistics



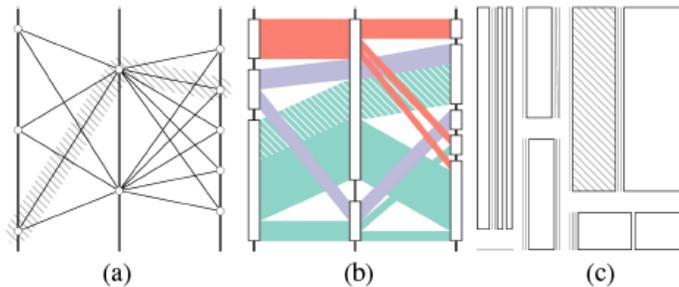
## Parallel Sets

- Fabian Bendix, Robert Kosara, Helwig Hauser, Infovis 05
- Builds on parallel coordinates to better handle categorical data
  - Discrete
  - Small number values
  - No implicit relationship
- Interaction



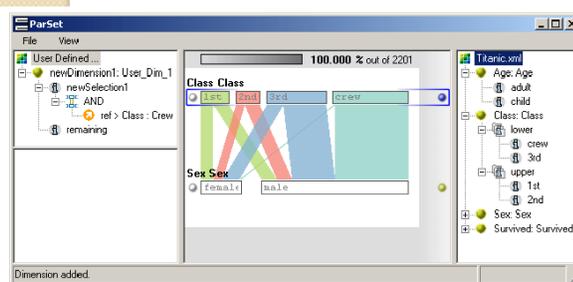
## Based on Parallel Coordinates

- Each attribute value becomes an axis position
  - Can modify to show frequency
  - Compare with Mosaic Display



## Basic Approach

- Potentially bin continuous attributes
- One active attribute determines color
- Use crosstabulation to order bins



(a)

Class	Sex		
	female	male	
first	145 44.6%	180 55.4%	325
	30.8% 6.6%	10.4% 8.2%	14.8%
second	106 37.2%	179 62.8%	285
	22.6% 4.8%	10.4% 8.1%	12.9%
third	196 27.8%	510 72.2%	706
	41.7% 8.9%	29.5% 23.2%	32.1%
crew	23 2.6%	862 97.4%	885
	4.9%	49.8% 39.1%	40.2%
	470	1731	2201
	21.4%	78.6%	100%

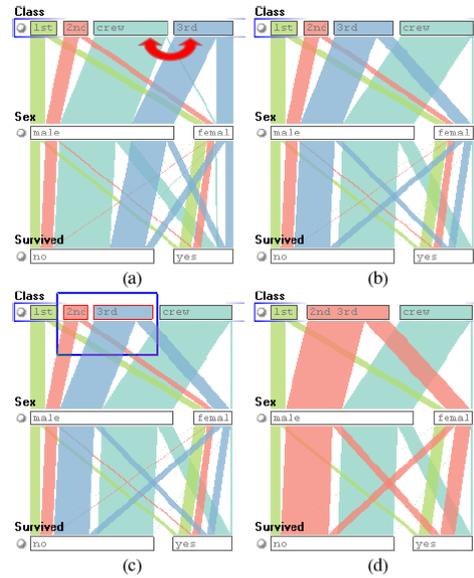
	40%	32%	13%	15%
	9%	-3%	-3%	-5%

(b)

(c)

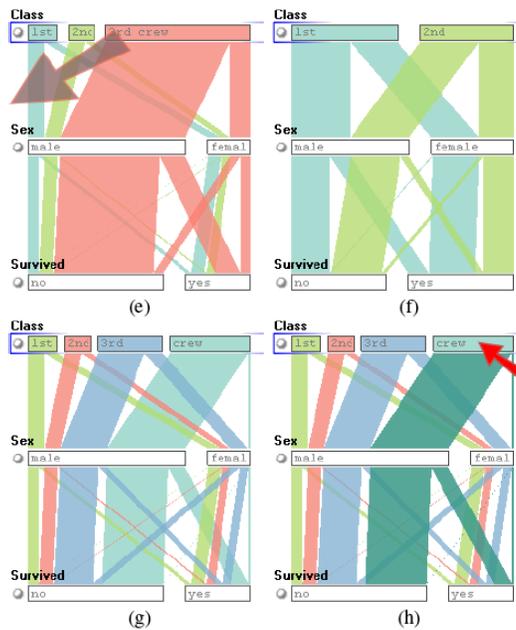
## Interaction

- Operations
  - Reorder
  - Merge categories

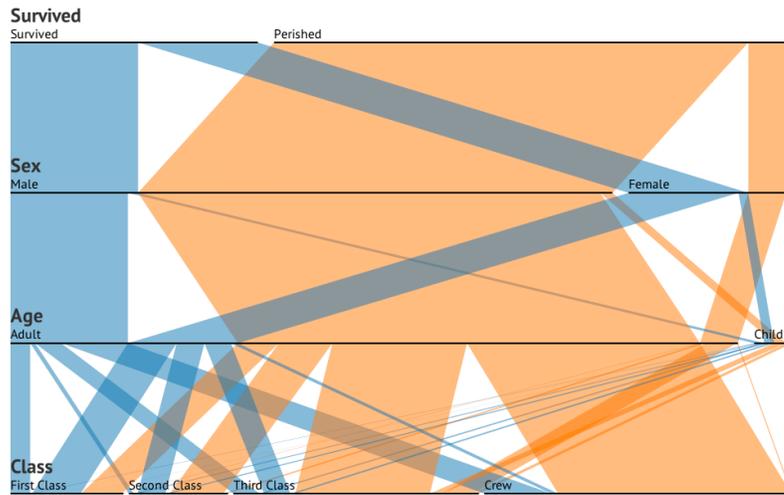


## Interaction

- Operations
  - Filter
  - Highlight

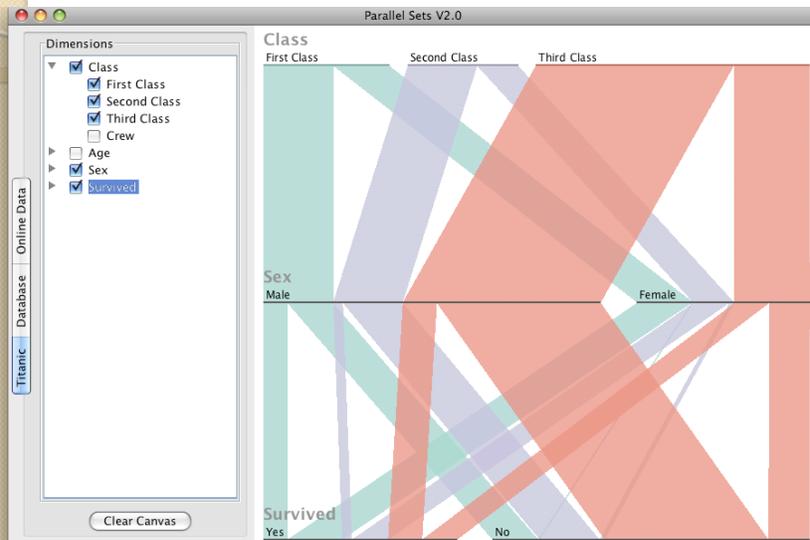


# Parallel Sets Example: Titanic



<http://www.jasondavies.com/parallel-sets/>

# Parallel Sets Example: Titanic



# Parallel Sets Example: Voting

