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

[Heatmap diagram with Education, Hours, and Probability axes]
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
Nowll, Hetzler, Tanasse, Infovis01
Non-Euclidean Layouts

- Havre, Hetzler, Perrine, Jurrus, Miller, Infovis01
Yang-Pelaez and Flowers '00

Keim, Hao, Ladisch, Hsu, Dayal, Infovis'01

a) Color: dollar amount  b) Color: no. of visits  c) Color: quantity

- Keim, Hao, Ladisch, Hsu, Dayal, Infovis'01
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
Interaction

- Operations
  - Reorder
  - Merge categories

Interaction

- Operations
  - Filter
  - Highlight
Parallel Sets Example: Voting