Star Glyph

General Overview

- Each line represents a dimension
- Angles between lines are constant
- Values of variables represented by length of lines. (Polygon is visualization aid only.)
- Useful up to approximately 20 dimensions.
- Some tools allow screen placement to represent two additional variables.

Purpose in Data Mining

- Useful for recognizing progressions
  - Changes over time
  - Relationship between several variables (two lines seem to grow together, or inversely.)
- Not useful for precise values.
- Used when a general idea of relation or change among a great many variable is desired.
- Generally an initial tool, that can lead to more precise inquiries among specific variables of interest.

http://davis.wpi.edu/~xmdv/vis_starglyph.html

- From Xmdv Unix tool
- Protrusion at 257° shows cleared homicides - evolves into a concavity over time.
- Of course we can’t see which variable are which here.
- Xmdv provides a separate panel that defines this.
- It was not clear how color is used in the Xmdv tool.
Comparison of Star Glyph and Scatterplot (same data)

Example: Animal features

- Statistica v.5 (1997)
- One glyph per animal
- Each branch is a physical attribute
- Attributes are boolean
  - Branch: attribute present

An Empirical Evaluation of Chernoff Faces, Star Glyphs, Spatial Visualizations for Binary Data. *

- Found both star glyphs and Chernoff faces representing boolean valued data led to slow, inaccurate answers with low response confidence.
- However, boolean (or incremental) value comparison is a poor task for star glyphs.


Tree Map

http://www.hivegroup.com/
General Overview

- Professor Shneiderman (U of Maryland)
  - Developed to show use of disc space by 14 researchers
- Uses visual elements to display and organize data:
  - Size - the larger the area, the larger the variable value
  - Color - keys rectangles to variables
  - Arrangement - all small boxes inside a larger box represent the tree originating at that branch.
  - Shape - generally rectangular
- Appears to be primarily commercial.

Purpose in Data Mining

- To facilitate “comprehensive understanding of complex structures.”
- The idea is that color coding, spatial organization, and size will show relationships like:
  - Clients use 90% of storage space for old email
  - Inexpensive coffee has more “full-bodied” options than expensive coffee

Purchasing Coffee online

Example : Purchasing Coffee

- Coffee can be organized by:
  - Price, Acidity, Body, Category
- Color can be:
  - Body or Acidity
- Hovering over any square presents a text description of that coffee.

- I liked being able to organize the coffee as desired. I could quickly find a full-bodied coffee under $12/lb.
- But did I need a tree map to do it?
- Or was it simply the ability to organize the data according to my own desires that was useful?

Flaws in the Paradigm

- These rectangles all represent $11.95.
- Color and shape affect perception of area.
- “Squarifying” is announced as a major advance.
  * Sami Lais, 6/02/001
  ComputerWorld

Another Example:

- What the heck is that thing?! --->
- How many of Tufte’s principles does it violate?
- How can we figure out what each stripe is?
- And how about that “reflection”?!?

“Star Glyphs seem like they would actually be useful in situations where one is exploring the relationship between variables, especially as it changes over time.

- Tree maps were a really creative idea when developed. I like the original concept, but was not impressed with any representation I found with many variables, even though that is the supposed strength of the tool.

Summary