

# **Information Visualization: Principles, Promise, and Pragmatics**

Marti Hearst

CHI 2003 Tutorial

1

## **Agenda**

- Introduction
- Visual Principles
- Design Guidelines
- Using Visualization in Problem Solving
- Visualizing Documents & Search
- Design Exercise
- Wrap-Up

2

## What is Information Visualization?

visualize: to form a mental image or vision of ...  
visualize: to imagine or remember as if actually seeing.

American Heritage dictionary, Concise Oxford dictionary

3

## Purposes of Visualization

Explore/Calculate

Analyze

Reason about Information

Communicate

Explain

Make Decisions

Reason about Information

4

## Why Visualization?

Use the eye for pattern recognition; people are good at  
scanning  
recognizing  
remembering images

Graphical elements facilitate comparisons via  
length  
shape  
orientation  
texture

Animation shows changes across time  
Color helps make distinctions  
Aesthetics make the process appealing

5

## Goals of Information Visualization

More specifically, visualization should:

- Make large datasets coherent  
(Present huge amounts of information compactly)
- Present information from various viewpoints
- Present information at several levels of detail  
(from overviews to fine structure)
- Support visual comparisons
- Tell stories about the data

6

## The Need for Critical Analysis

- We see many creative ideas, but they often fail in practice
- The hard part: how to apply it judiciously
  - Inventors usually do not accurately predict how their invention will be used
- This tutorial will emphasize
  - Getting past the coolness factor
  - Examining usability studies

7

## Case Study: The Journey of the Treemap

- The Treemap (**Johnson & Shneiderman '91**)
- Idea:
  - Show a hierarchy as a 2D layout
  - Fill up the space with rectangles representing objects
  - Size on screen indicates relative size of underlying objects.

8

## Early Treemap Applied to File System



9

## Treemap Problems

- Too disorderly
  - What does adjacency mean?
  - Aspect ratios uncontrolled leads to lots of skinny boxes that clutter
- Color not used appropriately
  - In fact, is meaningless here
- Wrong application
  - Don't need all this to just see the largest files in the OS

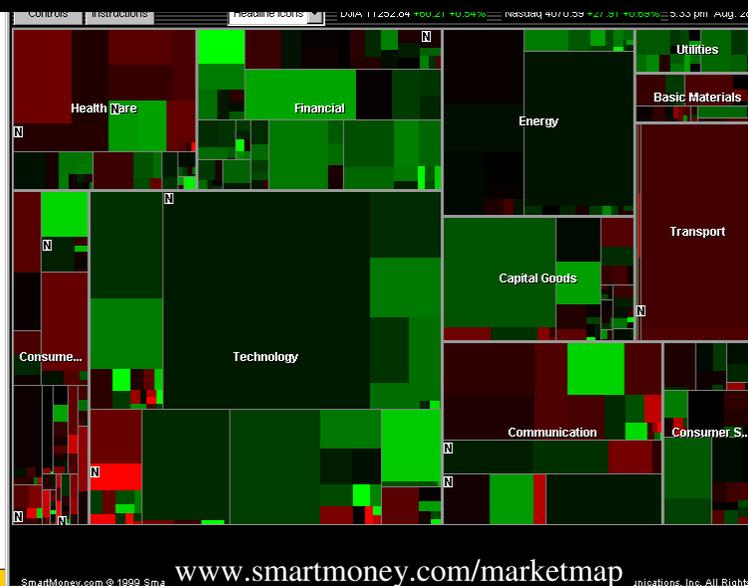
10

## Successful Application of Treemaps

- Think more about the use
  - Break into meaningful groups
  - Fix these into a useful aspect ratio
- Use visual properties properly
  - Use color to distinguish meaningfully
    - Only two colors: can distinguish one thing from another
    - Amount isn't very important
- Provide excellent interactivity
  - Access to the real data
  - Makes it into a useful tool

11

## A Good Use of TreeMaps and Interactivity



12

## Analysis vs. Communication

- MarketMap use of treemaps allows for sophisticated analysis
- Peets use of treemaps is more for presentation and communication
- This is a key contrast

13

## Open Issues

- Does visualization help?
  - The jury is still out
  - Still supplemental at best for text collections
    - A correlation with spatial ability
    - Learning effects: with practice ability on visual display begins to equal that of text
- Does visualization sell?
  - Jury is still out on this one too!
- This is a *hot* area! More ideas will appear!

14

## Key Questions to Ask about a Viz

1. What does it teach/show/elucidate?
2. What is the key contribution?
3. What are some compelling, *useful* examples?
4. Could it have been done more simply?
5. Have there been usability studies done?  
What do they show?

15

## What we are *not* covering

- Scientific visualization
- Statistics
- Cartography (maps)
- Education
- Games
- Computer graphics in general
- Computational geometry

16

## Promising Techniques

- Standard Techniques
- Brushing, Linking, and Dynamic Highlighting
- Animation
- Overview + Detail
- Color for Accent, Selection, Distinguishing
- Providing Multiple Views and Models
- Aesthetics!

17

## Standard Techniques

- It's often hard to beat:
  - Line graphs, bar charts
  - Scatterplots (or Scatterplot Matrix)
  - Tables
- A Darwinian view of visualizations:
  - Only the fittest survive
  - We are in a period of great experimentation; eventually it will be clear what works and what dies out.
- A bright spot:
  - Enhancing the old techniques with interactivity
  - Example: Spotfire
    - Adds interactivity, color highlighting, zooming to scatterplots
  - Example: TableLens / Eureka
    - Adds interactivity and length cues to tables

18

## We Use Animation to...

- Tell stories / scenarios: cartoons
- Illustrate dynamic process / simulation
- Create a character / an agent
- Navigate through virtual spaces
- Draw attention
- Delight

Slide by Saifon Obromsook & Linda Harjono

19

## Problem Solving

- A Detective Tool for Multidimensional Data
  - Inselberg on using Parallel Coordinates
- Analyzing Web Clickstream Data
  - Brainerd & Becker, Waterson et al.
- Information Visualization for Pattern Detection
  - Carlson & Konstan on Periodic Data
- Visualization vs. Analysis
  - Comments by Wesley Johnson of Chevron

20

## Visualization vs. Analysis?

- Applications to data mining and data discovery.
- Wesley Johnson '02:
  - Visualization tools are helpful for exploring hunches and presenting results
    - Examples: scatterplots
  - The WRONG primary tool when the goal is to find a good classifier model in a complex situation
  - Need:
    - Solid insight into the domain and problem
    - Tools that visualize several alternative models.
    - Emphasize "model visualization" rather than "data visualization"

21

## IR Infovis Meta-Analysis

(Empirical studies of information visualization:  
a meta-analysis, Chen & Yu IJHCS 53(5),2000)

- Goal
  - Find invariant underlying relations suggested collectively by empirical findings from many different studies
- Procedure
  - Examine the literature of empirical infoviz studies
    - 35 studies between 1991 and 2000
    - 27 focused on information retrieval tasks
    - But due to wide differences in the conduct of the studies and the reporting of statistics, could use only 6 studies

22

## IR Infovis Meta-Analysis

(Empirical studies of information visualization:  
a meta-analysis, Chen & Yu IJHCS 53(5),2000)

- **Conclusions:**
  - IR Infoviz studies not reported in a standard format
  - Individual cognitive differences had the largest effect
    - Especially on accuracy
    - Somewhat on efficiency
  - Holding cognitive abilities constant, users did better with simpler visual-spatial interfaces
  - The combined effect of visualization is not statistically significant

23

## What Works for Search?

- Hearst, M, et al., Finding the Flow in Web Site Search, CACM 45(9), 2002; Chapter 10 of Modern Information Retrieval, Baeza-Yates & Ribiero-Neto (Eds).
- Color highlighting of query terms in results listings
- Sorting of search results according to important criteria (date, author)
- Grouping of results according to well-organized category labels.
- Only if highly accurate:
  - Spelling correction/suggestions
  - Simple relevance feedback (more-like-this)
  - Certain types of term expansion
- Note: most don't benefit from visualization!

24

## Comparing 3 Commercial Systems

- Alfred Kobsa, An Empirical Comparison of Three Commercial Information Visualization Systems, INFOVIS'01.

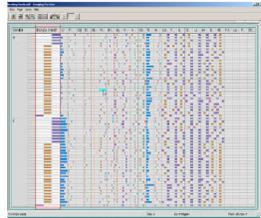


Figure 1a. A screenshot from Eureka that shows how a user might solve the question "Did males cheat more on their girlfriends than females on their boyfriends?" (an actual question used in the experiment). After grouping the attribute "Gender" and sorting the column "Did you cheat?", a user can compare the number of "Yes" entries and then find that more females than males indicated having cheated. (One male and one female gave no answers.)

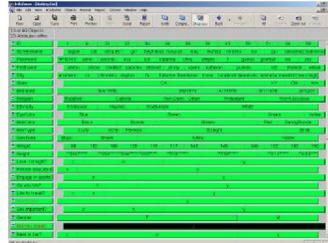


Figure 1b. This screenshot shows one possible way for solving the same problem in InfoZoom, specifically in its overview mode. After clicking at, and thereby zooming into, the "Yes" entries in the attribute "Did you cheat?", users can see from the length of the bars in the Gender category that females indicated more frequently having cheated than males.

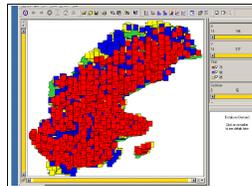


Figure 2. Spotfire's graphical representation of female-male connections through a scatterplot diagram.

25

## Holistic Design Goals for Information Visualization

- Tailor to the application and the domain
- Create highly interactive and integrated systems
- Embed the visualization within a larger application
- Provide alternative views

26