Information Visualization
CMSC 838B – Spring 2003

Introduction

Benjamin B. Bederson
Computer Science Dept., Asst. Prof.
Human-Computer Interaction lab, Director
University of Maryland
www.cs.umd.edu/~bederson

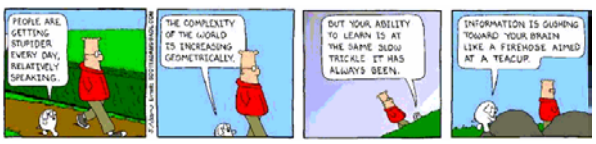
Data Explosion

- Between 1 and 2 exabytes of unique info produced per year
  - 1000000000000000000 (10^{18}) bytes
  - 250 meg for every man, woman and child
  - Printed documents only .003% of total

Peter Lyman and Hal Varian, 2000
Cal-Berkeley, Info Mgmt & Systems
www.sims.berkeley.edu/how-much-info

Data Overload

- Problem: How to make use of the data
  - How do we make sense of the data?
  - How do we harness this data in decision-making processes?
  - How do we avoid being overwhelmed?



The Challenge

- Transform the *data* into *information* (understanding, insight) thus making it useful to people.
- Support specific tasks
- Improve performance as compared to existing mechanisms

Information Visualization

- Provide tools that present data in a way to help people understand and gain insight from it
- Cliches
  - "Seeing is believing"
  - "A picture is worth a thousand words"

"The use of computer-supported, interactive, visual representations of abstract data to amplify cognition."

Main Idea

- Visuals help us think
  - Provide a frame of reference, a temporary storage area
- External cognition
  - Role of external world in thinking and reason
  - Multiplication exercise

Information Visualization

- What is “information”?
  - Items, entities, things which do not have a direct physical correspondence
  - Examples: baseball statistics, stock trends, connections between criminals, car attributes...
- Scientific Visualization
  - Primarily relates to and represents something physical or geometric
  - Examples
    - Air flow over a wing
    - Stresses on a girder
    - Weather over Pennsylvania

Key Attributes

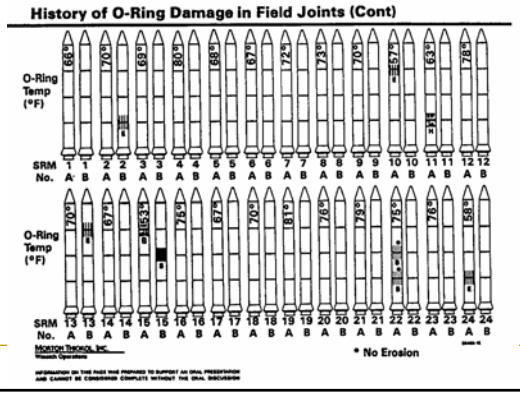
- Scale
  - Challenge often arises when data sets become very large
- Interactivity
  - Want to show multiple different perspectives on the data
- Tasks
  - Want to support specific tasks – not just to create a cool demo
  - Support discovery, decision making, explanation

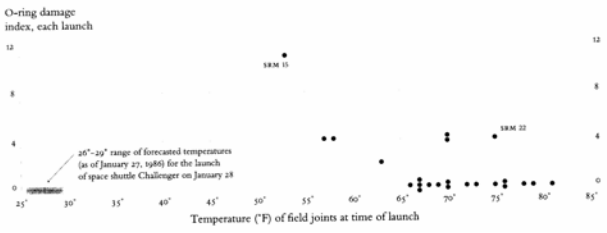
- Which state has highest Income?
- Relationship between Income and Education?
- Outliers?

Table - StateData ()

State	College Degree %	Per Capita Income
Alabama	20.6%	11486
Alaska	30.3%	17610
Arizona	27.1%	13461
Arkansas	17.0%	10520
California	31.3%	16409
Colorado	33.9%	14821
Connecticut	33.8%	20189
Delaware	27.9%	15854
District of Columbia	36.4%	18881
Florida	24.9%	14698
Georgia	24.3%	13631
Hawaii	31.2%	15770
Idaho	25.2%	11457
Illinois	26.8%	15201
Indiana	20.9%	13149
Iowa	24.5%	12422
Kansas	26.5%	13300
Kentucky	17.7%	11153
Louisiana	19.4%	10635
Maine	25.7%	12957
Maryland	31.7%	17730
Massachusetts	34.5%	17224
Michigan	24.1%	14154
Minnesota	30.4%	14389
Mississippi	19.9%	9648
Missouri	22.3%	12989
Montana	25.4%	11213
Nebraska	26.0%	12452
Nevada	21.5%	15214
New Hampshire	32.4%	15959
New Jersey	30.1%	18714
New Mexico	25.5%	11246
New York	29.6%	16501
North Carolina	24.2%	12885
North Dakota	28.1%	11051
Ohio	22.3%	13461
Oklahoma	22.8%	11893
Oregon	27.5%	13418
Pennsylvania	23.2%	14068
Rhode Island	27.5%	14981
South Carolina	23.0%	11897
South Dakota	24.6%	10661
Tennessee	20.1%	12255
Texas	25.5%	12904
Utah	30.0%	11029
Vermont	31.5%	13527
Virginia	30.0%	15713
Washington	30.9%	14923
West Virginia	16.1%	10520
Wisconsin	24.9%	13276
Wyoming	25.7%	12311

What's the Big Deal?





Presentation is everything!

The Power of Visualization

1. Start out going Southwest on ELLSWORTH AVE Towards BROADWAY by turning right.
2: Turn RIGHT onto BROADWAY.
3. Turn RIGHT onto QUINCY ST.
4. Turn LEFT onto CAMBRIDGE ST.
5. Turn SLIGHT RIGHT onto MASSACHUSETTS AVE.
6. Turn RIGHT onto RUSSELL ST.

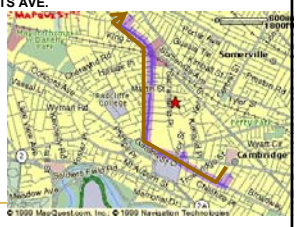
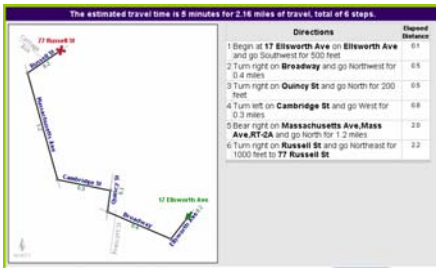


Image from mapquest.com

Slide from Marti Hearst

The Power of Visualization



Slide from Marti Hearst

Tool by Maneesh Agrawala http://graphics.stanford.edu/~maneesh/
Available from www.mapblast.com

Visualization Success Stories

Illustration of John Snow's deduction that a cholera epidemic was caused by a bad water pump, circa 1854.

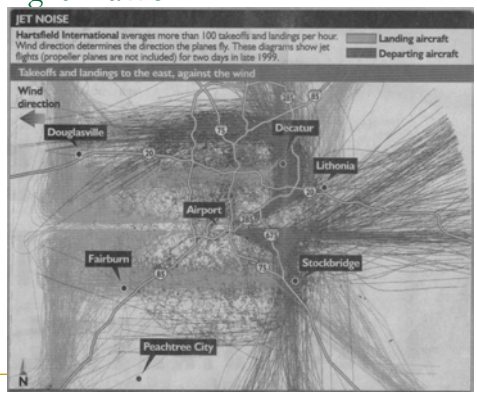
Dots indicate location of deaths.

From Visual Explanations by Edward Tufte, Graphics Press, 1997

Slide from Marti Hearst

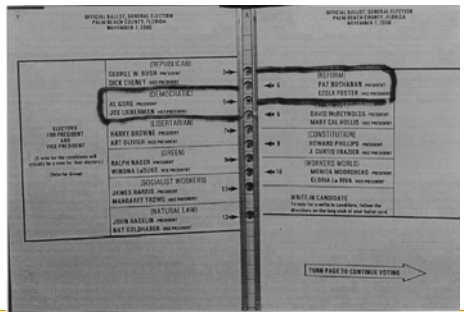
Examples - static

Atlanta Flight Traffic

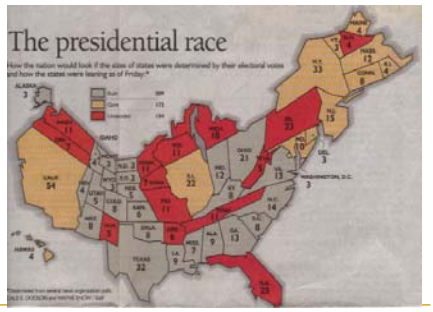


Atlanta Journal
April 30, 2000

2000 Election Ballot



Electoral College



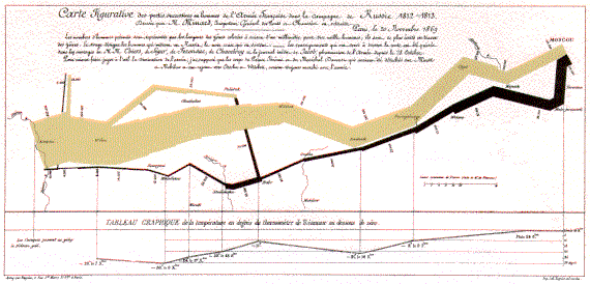
Atlanta Journal
November 5, 2000

London Subway

www.londontransport.co.uk/tube

Napolean's March

From E. Tufte
The Visual Display of Quantitative Information



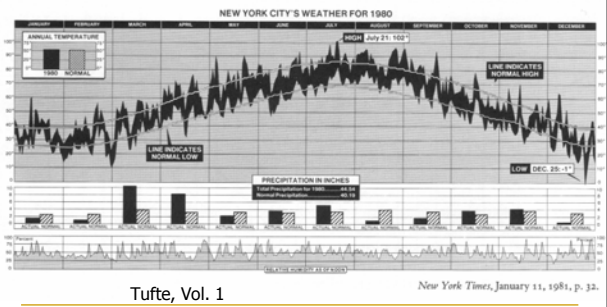
Minard graphic

size of army
direction
latitude
longitude
temperature
date

Example

NYC weather

2220 numbers



Tufte, Vol. 1

New York Times, January 11, 1981, p. 32.

Examples - interactive

StarTree

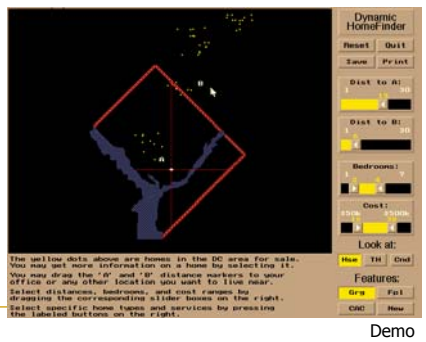
Hyperbolic tree

www.inxight.com

Demo

HomeFinder

HCIL
Univ. Maryland
1992



Demo

So Why Vision?

- Why not “perceptualization”?
  - Sonification
  - Touchification
  - Smellification
  - Tastification
- Bandwidth, bandwidth, bandwidth

Tasks in Info Vis

- Search
  - Finding a specific piece of information
    - How many games did the Braves win in 1995?
    - What novels did Ian Fleming author?
- Browsing
  - Look over or inspect something in a more casual manner, seek interesting information
    - Learn about crystallography
    - What has Jane been up to lately?

Tasks in Info Vis

- Analysis
  - Comparison-Difference
  - Outliers, Extremes
  - Patterns
- Assimilation
- Monitoring
- Awareness

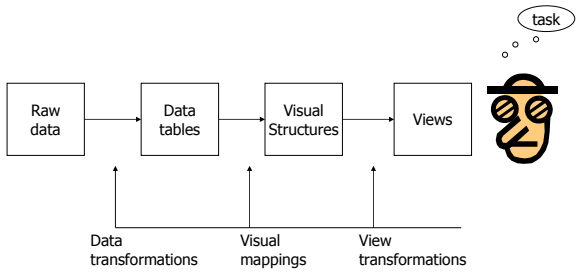
Knowledge Crystallization – Work Process

- Information foraging
- Search for schema (representation)
- Instantiate schema
- Problem solve to trade off features
- Search for a new schema that reduces problem to a simple trade-off
- Summarize and communicate

How Vis Amplifies Cognition

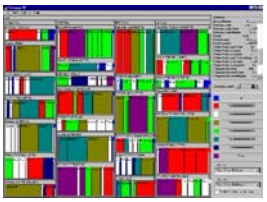
- Increasing memory and processing resources available
- Reducing search for information
- Enhancing the recognition of patterns
- Enabling perceptual inference operations
- Using perceptual attention mecahnisms for monitoring
- Encoding info in a manipulable medium

Data Process



The Need for Critical Analysis

- We see many creative ideas, but they often don't really work
- This course will emphasize
  - Getting past the coolness factor
  - Examining usability studies
- Example: Treemaps (www.cs.umd.edu/hcil/treemaps)
  - Show a hierarchy as a 2D layout
  - Size on screen indicates relative size of underlying objects



Early Treemap file browser

Slide from Marti Hearst

Treemap Problems

- Too disorderly
  - What does adjacency mean?
  - Large aspect ratios lead to skinny boxes that clutter
- Color difficult to understand
- What are the tasks?
  - Don't need all this to just see the largest files in the OS
  - But are there tasks where this would be appropriate?

Slide from Marti Hearst

Successful Application of Treemaps

- Think more about the use
  - Break into meaningful groups
  - Improve aspect ratio
- Use visual properties properly
  - Use color to distinguish meaningfully
- Provide excellent interactivity
  - Access to the real data
  - Makes it into a useful tool

Slide from Marti Hearst

A Good Use of TreeMaps and Interactivity

www.smartmoney.com/marketmap

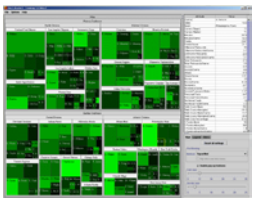
Slide from Marti Hearst

Treemaps in Peets site

www.peets.com/tast/11/coffee_selector.asp
www.peets.com/tast/12/tea_selector.asp

Treemap 3

- HCIL's latest
  - Control over the data and mappings
  - Control over the color
  - Better layout algorithms
  - Better interaction



www.cs.umd.edu/hcil/treemap3 - the software
www.cs.umd.edu/hcil/treemaps - the HCIL Treemap story

Course Administration

- Look at Syllabus
- Readings
  - Everyone reads every paper every class – no kidding
  - Everyone is prepared to talk about every paper every class – no kidding
- First homework due next week
- WAM accounts next week

How to Prepare for Readings

- What is the problem (specifically what tasks does it solve)?
- What assumptions are made?
- Who are the intended users of the research?
- Have those users been involved in the design or evaluation of the work (i.e., is the solution usable?)
- Is the solution scalable (how much data does it work with)?
- Is the solution generalizable (does the solution work in other domains)?
- What is the key contribution?

Research Class

- Creativity
- No “right” answer
- Reasoning/argument is more important
- Self motivation
- Open ended
- Contribute to the state-of-the-art

Class Project

- Build a new visualization
- Evaluation
- Groups 2-4
- Choose topic
- Literature review
- Design it
- Build it
- Evaluate it
- Write a paper about it
- Give a presentation.

Question to think about

- Is a spreadsheet a visualization?