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>
<thead>
<tr>
<th>State</th>
<th>College Exp (%)</th>
<th>Per Capita Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>20.2%</td>
<td>17,682</td>
</tr>
<tr>
<td>Alaska</td>
<td>64.6%</td>
<td>18,762</td>
</tr>
<tr>
<td>Arizona</td>
<td>17.1%</td>
<td>18,181</td>
</tr>
<tr>
<td>Arkansas</td>
<td>17.9%</td>
<td>17,782</td>
</tr>
<tr>
<td>California</td>
<td>14.6%</td>
<td>18,061</td>
</tr>
<tr>
<td>Colorado</td>
<td>13.3%</td>
<td>18,106</td>
</tr>
<tr>
<td>Connecticut</td>
<td>63.3%</td>
<td>18,111</td>
</tr>
<tr>
<td>Delaware</td>
<td>15.3%</td>
<td>15,994</td>
</tr>
<tr>
<td>District of Columbia</td>
<td>59.2%</td>
<td>15,657</td>
</tr>
<tr>
<td>Florida</td>
<td>31.2%</td>
<td>13,911</td>
</tr>
<tr>
<td>Georgia</td>
<td>23.1%</td>
<td>13,020</td>
</tr>
<tr>
<td>Hawaii</td>
<td>24.2%</td>
<td>13,457</td>
</tr>
<tr>
<td>Idaho</td>
<td>25.5%</td>
<td>17,122</td>
</tr>
<tr>
<td>Illinois</td>
<td>19.5%</td>
<td>12,124</td>
</tr>
<tr>
<td>Indiana</td>
<td>29.0%</td>
<td>13,146</td>
</tr>
<tr>
<td>Iowa</td>
<td>25.2%</td>
<td>13,003</td>
</tr>
<tr>
<td>Kansas</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Kentucky</td>
<td>19.4%</td>
<td>12,125</td>
</tr>
<tr>
<td>Louisiana</td>
<td>19.4%</td>
<td>13,005</td>
</tr>
<tr>
<td>Maine</td>
<td>26.2%</td>
<td>15,714</td>
</tr>
<tr>
<td>Maryland</td>
<td>25.3%</td>
<td>12,118</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>25.2%</td>
<td>11,996</td>
</tr>
<tr>
<td>Michigan</td>
<td>25.3%</td>
<td>11,996</td>
</tr>
<tr>
<td>Minnesota</td>
<td>26.2%</td>
<td>15,714</td>
</tr>
<tr>
<td>Missouri</td>
<td>19.4%</td>
<td>12,125</td>
</tr>
<tr>
<td>Montana</td>
<td>25.2%</td>
<td>13,003</td>
</tr>
<tr>
<td>Nebraska</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Nevada</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>24.2%</td>
<td>17,148</td>
</tr>
<tr>
<td>New Jersey</td>
<td>26.2%</td>
<td>15,714</td>
</tr>
<tr>
<td>New York</td>
<td>25.3%</td>
<td>12,118</td>
</tr>
<tr>
<td>North Carolina</td>
<td>25.2%</td>
<td>11,996</td>
</tr>
<tr>
<td>Ohio</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Oregon</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>South Carolina</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Tennessee</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Texas</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Utah</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Virginia</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Washington</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
<tr>
<td>West Virginia</td>
<td>25.2%</td>
<td>11,997</td>
</tr>
</tbody>
</table>
What's the Big Deal?

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.

Presentation is everything!
The Power of Visualization

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.


Examples - static
Napoleon’s March

From E. Tufte
The Visual Display of Quantitative Information

Minard graphic

<table>
<thead>
<tr>
<th>size of army</th>
<th>direction</th>
<th>latitude</th>
<th>longitude</th>
<th>temperature</th>
<th>date</th>
</tr>
</thead>
</table>

Example

NYC weather
2220 numbers

Tufte, Vol. 1
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 mechanisms for monitoring
- Encoding info in a manipulable medium

Data Process

- Raw data → Data tables → Visual Structures → Views
- Data transformations → Visual mappings → View transformations
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

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?

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
A Good Use of TreeMaps and Interactivity

www.smartmoney.com/marketmap

Slide from Marti Hearst

Treemaps in Peets site

www.peets.com/soap/11/coffee_selector.asp
www.peets.com/soap/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?