Information Visualization

John Stasko
Spring 2007

Agenda

• Why visualization?
• Definitions
• Examples
Exercise

- House directions

Data Explosion

- Society is more complex
  - There simply is more “stuff”

- Computers, internet and web give people access to an incredible amount of data
  - news, sports, financial, purchases, etc...
Data Overload

- Confound: 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 Problem

Web, Books, Papers, Game scores, Scientific data, Biotech, Shopping People Stock/finance News

**Data**  →  **Data Transfer**

**How?**

*Vision:* 100 MB/s  
*Ears:* <100 b/s
Telepathy  
Haptic/tactile  
Smell  
Taste

Two slides courtesy of Chris North
Human Vision

- Highest bandwidth sense
- Fast, parallel
- Pattern recognition
- Pre-attentive
- Extends memory and cognitive capacity
- People think visually

Impressive. Let's use it!

Want More Evidence?

Questions:

- Which state has the highest income?
- Is there a relationship between income and education?
- Are there any outliers?

<table>
<thead>
<tr>
<th>State</th>
<th>Income 1</th>
<th>Income 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>24,5%</td>
<td>6756</td>
</tr>
<tr>
<td>Alaska</td>
<td>28,5%</td>
<td>4168</td>
</tr>
<tr>
<td>Arizona</td>
<td>27,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Arkansas</td>
<td>25,5%</td>
<td>5678</td>
</tr>
<tr>
<td>California</td>
<td>32,5%</td>
<td>9876</td>
</tr>
<tr>
<td>Colorado</td>
<td>25,5%</td>
<td>6758</td>
</tr>
<tr>
<td>Connecticut</td>
<td>31,5%</td>
<td>1234</td>
</tr>
<tr>
<td>Delaware</td>
<td>31,5%</td>
<td>1234</td>
</tr>
<tr>
<td>District of Columbia</td>
<td>31,5%</td>
<td>1234</td>
</tr>
<tr>
<td>Florida</td>
<td>28,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Georgia</td>
<td>24,5%</td>
<td>1234</td>
</tr>
<tr>
<td>Hawaii</td>
<td>25,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Idaho</td>
<td>25,5%</td>
<td>7896</td>
</tr>
<tr>
<td>Illinois</td>
<td>24,5%</td>
<td>1234</td>
</tr>
<tr>
<td>Indiana</td>
<td>26,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Iowa</td>
<td>24,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Kansas</td>
<td>25,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Kentucky</td>
<td>17,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Louisiana</td>
<td>14,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Maine</td>
<td>15,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Maryland</td>
<td>17,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>14,5%</td>
<td>7456</td>
</tr>
<tr>
<td>Minnesota</td>
<td>18,5%</td>
<td>1234</td>
</tr>
</tbody>
</table>

Example courtesy of Chris North
Visualize the Data

Even Tougher?

- What if you could only see 1 state’s data at a time? (e.g. Census Bureau’s website)
- What if I read the data to you?
Exercise Redux

- An interesting query...

- People work differently

Our Challenge

- Transform data into information (understanding, insight) thus making it useful to people
Visualization

- Often thought of as process of making a graphic or an image
- Really is a cognitive process
  - Form a mental image of something
  - Internalize an understanding
- “The purpose of visualization is insight, not pictures”
  - Insight: discovery, decision making, explanation

Main Idea

- Visually help us think
  - Provide a frame of reference, a temporary storage area
  - “Seeing is believing”
  - “A picture is worth a thousand words”
- External cognition aid
  - Role of external world in thinking and reason
  - An illustrative example
Examples

- Images
  - Are these static pictures information visualizations?

Information Visualization

- What is “information”?
  - Items, entities, things which do not have a direct physical correspondence
  - Notion of abstractness of the entities is important too
  - Examples: baseball statistics, stock trends, connections between criminals, car attributes...
Information Visualization

• What is “visualization”?
  – The use of computer-supported, interactive visual representations of data to amplify cognition.
  • From [Card, Mackinlay Shneiderman ’98]

Two Key Attributes

• Scale
  – Challenge often arises when data sets become very large

• Interactivity
  – Want to show multiple different perspectives on the data
Domains for Info Vis

- Text
- Statistics
- Financial/business data
- Internet information
- Software
- ...

Components of Study

- Data analysis
  - Data items with attributes or variables
  - Generate data tables
- Visual structures
  - Spatial substrate, marks, graphical properties of marks
- UI and interaction
- Analytic tasks to be performed
  - Browse, correlate, identify, associate...
More Examples

- Seeing is believing...

Excel

Get rid of those darn 3D bars!
USA Today Graphics

To the breaking point of stress
People who frequently experience stress in daily life:
- Men
  - 18-49 years: 40%
  - 50+ years: 23%
- Women
  - 18-49 years: 59%
  - 50+ years: 35%

Source: Gallup poll

Atlanta Flight Traffic

Atlanta Journal
April 30, 2000
In Living Color

Maxim Magazine, July '01

Country Music

Figure 34. States Mentioned in Country-Music Lyrics
Note: The size of each state is proportional to the number of times it is mentioned.
London Subway

www.thetube.com

True Geography

www.kottke.org/plus/misc/images/tubagao.gif
Easy Walking Lines Added
Napoleonic March

From E. Tufte
The Visual Display of Quantitative Information

Minard graphic

size of army
direction
latitude
longitude
temperature
date
NYC Weather

2220 numbers

Data Table Format

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Case_1</th>
<th>Case_2</th>
<th>Case_3</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable_1</td>
<td>Value_{11}</td>
<td>Value_{21}</td>
<td>Value_{31}</td>
<td></td>
</tr>
<tr>
<td>Variable_2</td>
<td>Value_{12}</td>
<td>Value_{22}</td>
<td>Value_{32}</td>
<td></td>
</tr>
<tr>
<td>Variable_3</td>
<td>Value_{13}</td>
<td>Value_{23}</td>
<td>Value_{33}</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Think of as a function

\[ f(\text{case}_1) = \langle \text{Val}_{11}, \text{Val}_{12}, \ldots \rangle \]

Time series data a special case

Data Structure

- Sometimes the data has additional structure
  - Network/graph data
  - Hierarchical data
  - Important meta-data

True InfoVis Examples

- Systems – Key part of information visualization is the interactive capability (view different perspectives on data)
Baby Names

Viewing historical trends in baby names

http://babynamewizard.com/namevoyager/

Spotfire

www.spotfire.com
Table Lens

Tasks in Info Vis

• Search (not so much)
  – Finding a specific piece of information
    • How many games did the Braves win in 1995?
    • What novels did Ian Fleming author?

• Browsing (much more)
  – Look over or inspect something in a more casual manner, seek interesting information
    • How did the Falcons season go last year?
    • What’s a good car to buy?
Tasks in Info Vis

- Analysis & exploration
  - Comparison-Difference
  - Outliers, Extremes
  - Patterns
- Assimilation
- Monitoring
- Awareness
- Presentation

Case Study

- Understanding hierarchies

- Learn about some InfoVis techniques
Hierarchies

• Definition
  – Data repository in which cases are related to subcases
  – Can be thought of as imposing an ordering in which cases are parents or ancestors of other cases

Hierarchies in the World

• Pervasive
  – Family histories, ancestries
  – File/directory systems on computers
  – Organization charts
  – Animal kingdom: Phylum,..., genus,...
  – Object-oriented software classes
  – ...

• Hierarchies often represented as trees
Representations

Space-Filling Representation

Each item occupies an area

Children are “contained” under parent

One example
Treemap

- Space-filling representation developed by Shneiderman and Johnson, Vis '91
- Children are drawn inside their parent
- Alternate horizontal and vertical slicing at each successive level
- Use area to encode other variable of data items

Example
Example

Example
Treemap

- Example

Directories

SequoiaView

www.win.tue.nl/sequoiaview/

File visualizer built using cushion treemap notion

Demo
Map of the Market

www.smartmoney.com/marketmap

Sunburst

- Visualizing file and directory structures
- Root dir at center
- Color - file type
- Angle - file/dir size

Demo

6750-Spr '07
InfoVis Techniques

• Aggregation
  – Accumulate individual elements into a larger unit to be presented as some whole

• Overview & Detail
  – Provide both global overview and detail zooming capabilities

• Focus + Context
  – Show details of one or more regions in a more global context (e.g., fisheye)

InfoVis Techniques

• Drill-down
  – Select individual item or smaller set of items from a display for a more detailed view/analysis

• Brushing
  – Select or designate/specify value, then see pertinent items elsewhere on the display
CS 7450
Spring term

Course foci
- Examine research ideas
- Work with commercial systems
- Assignments and term project

HW 4

- Find an InfoVis-style graphic
- Critique the graphic (+/-) 1-page
- Due next Thursday
Upcoming

- WWW design and evaluation
- Embodied agents