What is visualization?

Definition [www.oed.com]

1. The action or fact of visualizing; the power or process of forming a mental picture or vision of something not actually present to the sight; a picture thus formed.

2. The action or process of rendering visible.

“Transformation of the symbolic into the geometric” [McCormick et al. 1987]

“... finding the artificial memory that best supports our natural means of perception.” [Bertin 1967]

“The use of computer-generated, interactive, visual representations of data to amplify cognition.” [Card, Mackinlay, & Shneiderman 1999]
<table>
<thead>
<tr>
<th>Set A</th>
<th>Set B</th>
<th>Set C</th>
<th>Set D</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>10</td>
<td>8.04</td>
<td>10</td>
<td>9.14</td>
</tr>
<tr>
<td>8</td>
<td>6.95</td>
<td>8</td>
<td>8.14</td>
</tr>
<tr>
<td>13</td>
<td>7.58</td>
<td>13</td>
<td>7.74</td>
</tr>
<tr>
<td>9</td>
<td>8.81</td>
<td>9</td>
<td>8.77</td>
</tr>
<tr>
<td>11</td>
<td>8.33</td>
<td>11</td>
<td>9.26</td>
</tr>
<tr>
<td>14</td>
<td>9.96</td>
<td>14</td>
<td>8.1</td>
</tr>
<tr>
<td>6</td>
<td>7.24</td>
<td>6</td>
<td>6.13</td>
</tr>
<tr>
<td>4</td>
<td>4.26</td>
<td>4</td>
<td>3.1</td>
</tr>
<tr>
<td>12</td>
<td>10.84</td>
<td>12</td>
<td>9.11</td>
</tr>
<tr>
<td>7</td>
<td>4.82</td>
<td>7</td>
<td>7.26</td>
</tr>
<tr>
<td>5</td>
<td>5.68</td>
<td>5</td>
<td>4.74</td>
</tr>
</tbody>
</table>

Summary Statistics
- $u_X = 9.0$, $\sigma_X = 3.317$
- $u_Y = 7.5$, $\sigma_Y = 2.03$

Linear Regression
- $Y' = 3 + 0.5X$
- $R^2 = 0.67$

[Anscombe 73]
Why do we create visualizations?

- Answer questions (or discover them)
- Make decisions
- See data in context
- Expand memory
- Support graphical calculation
- Find patterns
- Present argument or tell a story
- Inspire
Three functions of visualizations

**Record**: store information
- Photographs, blueprints, ...

**Analyze**: support reasoning about information
- Process and calculate
- Reason about data
- Feedback and interaction

**Communicate**: convey information to others
- Share and persuade
- Collaborate and revise
- Emphasize important aspects of data

---

Record Information

Galileo’s drawings of the phases of the moon from 1616
[galileo.rice.edu/sci/observations/moon.html](http://galileo.rice.edu/sci/observations/moon.html)

Drawing: Phases of the moon

Gallop, Bay Horse “Daisy” [Muybridge 1884-86]
Other recording instruments

Marey’s sphygmnograph [from Braun 83]
Support Reasoning

Make a decision: Challenger

Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

2 of 13 pages of material faxed to NASA by Morton Thiokol [from Tufte 1997]
Make a decision: Challenger

Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

Data in context: Cholera outbreak

In 1854 John Snow plotted the position of each cholera case on a map. [from Tufte 83]

Data in context: Cholera outbreak

Used map to hypothesize that pump on Broad St. was the cause. [from Tufte 83]

Expand memory: Multiplication

- Class Exercise
Expand memory: Multiplication

\[
\begin{array}{c}
34 \\
\times 72
\end{array}
\]

\[
\begin{array}{c}
\phantom{34} \\
\phantom{\times 72}
\end{array}
\]

\[
\begin{array}{c}
68 \\
2380 \\
2448
\end{array}
\]

Expand memory: Multiplication

Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]

Calculation: Evaporation

Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]
Find patterns: NYC weather

From the New York Times 1981

The most powerful brain?

The Dragons of Eden [Carl Sagan]

The Elements of Graphing Data [Cleveland]
Convey Information to Others

Inspire

Bones in hand [from 1918 edition]  Double helix model [Watson and Crick 53]

"to affect thro' the Eyes what we fail to convey to the public through their word-proof ears"

1856 "Coxcomb" of Crimean War Deaths, Florence Nightingale
Three functions of visualizations

**Record**: store information
- Photographs, blueprints, ...

**Analyze**: support reasoning about information
- Process and calculate
- Reason about data
- Feedback and interaction

**Communicate**: convey information to others
- Share and persuade
- Collaborate and revise
- Emphasize important aspects of data

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>More and more unseen data</td>
<td>More and more unseen data</td>
</tr>
<tr>
<td>Faster creation and collection</td>
<td>Faster creation and collection</td>
</tr>
</tbody>
</table>
Challenge

- More and more unseen data
- Faster creation and collection

Sloan digital sky survey
www.sdss.org

Sensor networks [Hill 02]
www.xbow.com

Digital photography

Challenge

- More and more unseen data
- Faster creation and collection
- Faster dissemination

Sensing

Photo sharing/annotation
flickr.com

Group Authoring
wikipedia.org

Map of the Internet [Cheswick 99]
research.lumeta.com

Challenge

More and more unseen data
- Faster creation and collection
- Faster dissemination

5 exabytes of new information in 2002 [Lyman 03]
- 37,000 Libraries of Congress
161 exabytes in 2006 [Gantz 07]

Necessitates better tools and algorithms for visually conveying information

2010: 1,200 exabytes
10x increase over 5 years

Gantz et al, 2008, 2010
Attention

“What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.”

Herb Simon as quoted by Hal Varian
Scientific American September 1995

Goals of visualization research

1 Understand how visualizations convey information
   - What do people perceive/comprehend?
   - How do visualizations correspond with mental models?

2 Develop principles and techniques for creating effective visualizations and supporting analysis
   - Amplify perception and cognition
   - Strengthen tie between visualization and mental models

Course Topics

Data and Image Models

Sémiologie Graphique [Bertin 67]
Visualization (Re-)Design

Graphical Perception

Color

Interaction
**Animation**

Animated transitions in statistical data graphics [Heer & Robertson 07]

**Visualization Toolkits**

**Using Space Effectively**

Dymaxion Maps [Fuller 46]

**Graphs and Trees**

Degree-OF-Interest Trees [Heer & Card 04]
Graphs and Trees

Text Visualization

Identifying Design Principles

Collaboration and History

Testing effectiveness of 3 types of assembly instructions [Heiser 04]

Where have all the dentists gone?
Course Mechanics

Course Goals

1 Evaluate and critique visualization designs
2 Implement interactive data visualizations
3 Gain an overview of research & techniques
4 Develop a substantial visualization project

Instructors

Jeffrey Heer
Assistant Professor, Computer Science
Office Hours: Tu 11-12, 375 Gates
http://vis.stanford.edu/jheer

Vadim Ogievetsky
MS Student, Computer Science
Office Hours: W 12:30-2, Th 3:30-5, Bytes Cafe
http://vadim.ogievetsky.com/
Textbooks

See also: www.edwardtufte.com

Readings

Some from textbooks, also many papers
Material in class will be loosely based on readings
Readings should be read by start of class
Post discussion comments on class wiki
Comments must be posted within 1 day of lecture

Course wiki: http://cs448b.stanford.edu/
Use your SUNet login and password.

Requirements

Class participation (10%)
A1: Visualization Design (10%)
A2: Exploratory Data Analysis (15%)
A3: Interactive Visualization Software (25%)
FP: Final Project (40%)

Final Project

Visualization research project on topic of your choice
Project write-up in form of a 8 page research paper
Two project presentations
  1. Background research and project proposal (Nov 16 + 18)
  2. Final presentation - exact time to be determined

Projects from previous classes have been published
  • IEEE Visualization
  • IEEE Information Visualization
  • SIGGRAPH
Protovis: A Graphical Toolkit for Visualization
Mike Bostock
Assignment 1: Visualization Design

Design a static visualization for a small, given data set.
You are assisting a manager in a tech company. You’ve been asked to create a simple means to help her compare the sales of Android phones (both actual and predicted) to other smartphones. She wants to know how the competitive landscape may be shifting.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sales (in units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>75.5</td>
</tr>
<tr>
<td>2007</td>
<td>31.0</td>
</tr>
<tr>
<td>2007</td>
<td>13.3</td>
</tr>
<tr>
<td>2007</td>
<td>14.7</td>
</tr>
<tr>
<td>2007</td>
<td>0.0</td>
</tr>
<tr>
<td>2007</td>
<td>10.9</td>
</tr>
<tr>
<td>2007</td>
<td>70.1</td>
</tr>
<tr>
<td>2008</td>
<td>85.5</td>
</tr>
<tr>
<td>2008</td>
<td>51.4</td>
</tr>
<tr>
<td>2008</td>
<td>15.8</td>
</tr>
<tr>
<td>2008</td>
<td>4.6</td>
</tr>
<tr>
<td>2008</td>
<td>15.5</td>
</tr>
<tr>
<td>2008</td>
<td>80.5</td>
</tr>
<tr>
<td>2008</td>
<td>16.5</td>
</tr>
<tr>
<td>2008</td>
<td>24.9</td>
</tr>
<tr>
<td>2008</td>
<td>4.9</td>
</tr>
<tr>
<td>2008</td>
<td>8.8</td>
</tr>
<tr>
<td>2009</td>
<td>15.4</td>
</tr>
<tr>
<td>2009</td>
<td>187.7</td>
</tr>
<tr>
<td>2009</td>
<td>45.9</td>
</tr>
<tr>
<td>2009</td>
<td>41.8</td>
</tr>
<tr>
<td>2009</td>
<td>32.7</td>
</tr>
<tr>
<td>2009</td>
<td>12.3</td>
</tr>
<tr>
<td>2009</td>
<td>16.8</td>
</tr>
<tr>
<td>2009</td>
<td>441.3</td>
</tr>
<tr>
<td>2009</td>
<td>63.2</td>
</tr>
</tbody>
</table>

Questions?