The Purpose of Visualization

Maneesh Agrawala

CS 294-10: Visualization
Fall 2013

How much data (bytes) did we produce in 2011?
2011: 1800 exabytes
10x increase over 5 years

[Image courtesy cabspotting.org]
“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
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.
What is visualization?

“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>8.74</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</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 Linear Regression

\[
\begin{align*}
\mu_X &= 9.0 & \sigma_X &= 3.317 & Y &= 3 + 0.5 \, X \\
\mu_Y &= 7.5 & \sigma_Y &= 2.03 & R^2 &= 0.67
\end{align*}
\]

[Anscombe 73]
Why do we create visualizations?
Why do we create visualizations?

- Help with pattern recognition
- It would be stupid not to use it
- Can show important aspects
- Compresses the data (higher bandwidth)
- Aesthetically pleasing/increases engagement
- Pushing a bias
- Representation similar to original form
- Visual analogies
- More direct perhaps than text

Three functions of visualizations

Record information
- Photographs, blueprints, …

Support reasoning about information (analyze)
- Process and calculate
- Reason about data
- Feedback and interaction

Convey information to others (present)
- Share and persuade
- Collaborate and revise
- Emphasize important aspects of data
Record Information

Answer question

Gallop, Bay Horse “Daisy” [Muybridge 1884-86]
Photographs: Phases of the moon

http://galileo.rice.edu/sci/observations/moon.html

Drawing: Phases of the moon

Galileo’s drawings of the phases of the moon from 1616
http://galileo.rice.edu/sci/observations/moon.html
Other recording instruments

Marey’s sphygmgograph [from Braun 83]

Support Reasoning
Find patterns: New York weather

From the New York Times 1981

Make a decision: Challenger

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]
See data in context: Cholera outbreak

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

See data in context: Cholera outbreak

Used map to support hypothesis Broad St. pump was the cause. [from Tufte 83]
Expand memory: Multiplication

Class Exercise

34
x 72
Expand memory: Multiplication

\[
\begin{array}{c}
34 \\
\times 72 \\
\hline
68 \\
2380 \\
2448
\end{array}
\]

Graphical calculation: Evaporation

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

Graphical calculation: Evaporation

Graphical calculation: Visual proofs

Sum of odd numbers:
\[1 + 3 + 5 + 7 + 9 = 5^2\]

Pythagorean theorem:
Chinese proof by dissection
Convey Information to Others

Present argument

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

Crimean War Deaths [Nightingale 1858]
Inspire

Bones in hand [from 1918 edition]

Double helix model [Watson and Crick 53]

Visualization Research
Challenge

More and more unseen data
- Faster creation and collection

Urban development planning
www.urbansim.org

Fluid flow
ctr.stanford.edu

Simulation
Challenge

More and more unseen data
- Faster creation and collection

Sensing
- Sloan digital sky survey [www.sdss.org]
- Sensor networks [Hill 02] [www.xbow.com]
- Digital photography

Internet
- Photo sharing/annotation [flickr.com]
- Group Authored Encyclopedia [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]
161 exabytes in 2006 [Gantz 07]
1800 exabytes in 2011 [Gantz 11]

Need better tools and algorithms for visually conveying information

The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that’s going to be a hugely important skill in the next decades, … because now we really do have essentially free and ubiquitous data. So the complimentary scarce factor is the ability to understand that data and extract value from it.

Hal Varian, Google’s Chief Economist
*The McKinsey Quarterly*, Jan 2009
Goals of visualization research

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

2. Develop principles and techniques for creating effective visualizations
   - Amplify perception and cognition
   - Strengthen connection between visualization and mental models of data

Topics
1. Data and image models

[Bertin, Graphics and Graphic Information Processing 1981]

2. Visualization Design

Problematic design

Redesign
3. Perception

The psychophysics of sensory function [Stevens 61]

4. Interaction

Oakland Crimespotting (crimespotting.org) [Stamen]
5. Color

[Image: ColorBrewer 5-class diverging RdGy]

[from Cynthia Brewer  http://www.personal.psu.edu/faculty/c/a/cab38/]

6. Building interactive visualizations with D3

[Image: D3: Data Driven Documents [Bostock 2011]]
7. Wrangling, Cleaning and Profiling

Proactive Wrangling [Guo 2011]

8. Spatial Layout

London underground [Beck 33]
9. Collaborative visualization

CommentSpace [Willett 2011]

10. Animation

Animated Transitions [Heer 07]
10. Trees and graphs

Degree-of-Interest Trees [Heer 2004]

11. Conveying structure

Principal Organs & Vascular System [Leonardo da Vinci ca. 1490]
Strange Immersion of Torus in 3-Space [Curtis 92]
12. Identifying design principles

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

Course Mechanics
Course Goals

1. Design, evaluate and critique visualizations
2. Explore data using existing visualization tools
3. Implement interactive data visualizations
4. Gain an overview of research and techniques
5. Develop a substantial visualization project
Textbooks

See also: www.edwardtufte.com

Readings

- Some from textbooks, also many papers
  Username/Password: vis2013/vis2013Readings
- Material in class will be loosely based on readings
- Readings should be read by start of class
- Post discussion comments on class wiki
  Must post by 3pm on day of lecture
  You have 3 passes for the semester

Class home page
http://vis.berkeley.edu/courses/cs294-10-fa13/wiki
Requirements

Class participation (10%)

Assignment 1: Visualization Design (10%)

Assignment 2: Exploratory Data Analysis (15%)

Assignment 3: Creating Interactive Visualization Software (25%)

Final Project (40%)


Final project

- Visualization research project on topic of your choice
- 2nd half of class
- Project write-up in form of a research paper
- Project presentations
  1. Background research on project area
  2. Midway presentation on prototype solutions
  3. Final presentation – exact time to be determined

Projects from previous classes have been published
- IEEE Visualization
- IEEE Information Visualization
- SIGGRAPH
Text Co-Occurrences

Words that co-occur with Israel in the King James Bible [Brandon Liu 2011]

Visualizing Commuting Patterns

Interactive visualization of commuting patterns in San Francisco [Michael Porath 2011]
Assignment 1: Visualization Design

Due by 9am on Sep 11
Information

Journals/person increases 10X every 50 years

[Slide from PARC UIR group]

Information Overload

Journals/person increases 10X every 50 years

[Slide from PARC UIR group]
Assignment 1: Good and Bad Vis.

Find two visualizations one **good** and one **bad**

**Use original sources**
- Journals
- Science magazines
- Newspapers
- Textbooks

**Make wiki page**
- Clearly mark as good or bad
- Provide short explanation
- Be prepared to succinctly describe in class on Wed Jan 27

Due before class Mon Jan 25