Tufte’s Design Principles

CS 4460/7450 - Information Visualization
February 17, 2009
John Stasko

Today’s Agenda

Edward Tufte has written seven books, including Visual Explanations: Envisioning Information, The Visual Display of Quantitative Information, and Data Analysis for Politics and Policy. He writes, designs, and self-publishes his books on analytical design, which have received more than 40 awards for content and design. He is a Professor Emeritus at Yale University, where he taught courses in statistical evidence, information design, and interface design. His current work includes landscape sculpture, graphic design, and a new book.

This website describes Edward Tufte’s books, one-day courses, and artwork. For further information, call Graphics Press at 203-272-9187, or fax 203-272-8808, or email.

For a moderated forum on analytical design, go to ASK E.T.
Envisioning Information

- Let’s hear your views on the book...

Graphical Excellence

- Principles
  - Graphical excellence is the well-designed presentation of interesting data---a matter of *substance*, of *statistics*, and of *design*
  - Graphical excellence consists of complex ideas communicated with clarity, precision and efficiency

According to Tufte
Graphical Excellence

- Principles
  - Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
  - Graphical excellence is nearly always multivariate
  - And graphical excellence requires telling the truth about the data

Leveraging Human Capabilities

- Data graphics should complement what humans do well

“We thrive in information-thick worlds because of our marvelous and everyday capacities to select, edit, single out, focus, organize, condense, reduce, boil down, choose, categorize, catalog, classify, list, abstract, scan, look over, sort, integrate, blend, inspect, filter, lump, skip, smooth, chunk, average, approximate, cluster, aggregate, outline, summarize, itemize, review, dip into, flop through, browse, glance into, leaf through, skim, refine, enumerate, glean, synopsize, winnow the wheat from the chaff, and separate the sheep from the goats.” Vol.2, page 50
Summary

• 1. Tell the truth
  – Graphical integrity
• 2. Do it effectively with clarity, precision...
  – Design aesthetics

Let’s look at each of these

1. Graphical Integrity

• Your graphic should tell the truth about your data
Example

Stock market crash?

- 500
- 475
- 450

1998 1999 2000 2001 2002

Example

Show entire scale

- 500
- 250
- 0

1998 1999 2000 2001 2002

Spring 2009
CS 4460/7450
**Example**

- Show in context

**Chart Integrity**

- Where’s baseline?
- What’s scale?
- What’s context?
Vol 1, p. 54 (1)

Where's 0?
Note middle '70

Vol 1, p 54 (2)

What's being compared?
Vol 1, 57

Scale?

Vol 1, p. 61

Scale?
Great work!

Ahhhh

Show the context
Local Example

A huge rise?

Atlanta Journal Constitution
Summer ‘08

More of the data

Atlanta Journal Constitution
Dec. ‘08
Watch Size Coding

- Height/width vs. area vs. volume

Vol 1, p. 69

area = value?
Measuring Misrepresentation

- Visual attribute value should be directly proportional to data attribute value

\[
\text{Lie factor} = \frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}
\]

\[
p.62 \quad 9.4 = \frac{4280}{454}
\]
2. Design Aesthetics

- Set of principles to help guide designers

Design Principles

- Maximize data-ink ratio

Data ink ratio = $\frac{\text{Data ink}}{\text{Total ink used in graphic}}$

= proportion of graphic’s ink devoted to the non-redundant display of data-information
More...

- Above all else, show the data
- Maximize the data-ink ratio
- Erase non-data-ink
- Erase redundant data-ink
- Revise and edit

More...

- Maximize data density

\[
\text{data density of graphic} = \frac{\text{number of entries in data matrix}}{\text{area of data graphic}}
\]

Quote ...
Maximize Data Density

“Data-rich designs give a context and credibility to statistical evidence. Low-information designs are suspect: what is left out, what is hidden, why are we shown so little? High-density graphics help us to compare parts of the data by displaying much information within the view of the eye: we look at one page at a time and the more on the page, the more effective and comparative our eye can be. The principle, then, is:

Maximize data density and the size of the data matrix, within reason.”

Vol 1, p 168

Redesign charts

• Bar chart, scatter plot, box plot
  (See drawings)
Design Principles

• Avoid chartjunk
  – Extraneous visual elements that detract from message

Vol 1, p 108
Vol 2, p.34

A classic

Diamonds Were A Girl’s Best Friend

USA Today

Junk Charts Blog

http://junkcharts.typepad.com/

Great narrative: Vol.2, bottom page 33-34
Design Principles

- Utilize multifunctioning graphical elements (macro/micro readings)
  - Graphical elements that convey data information and a design function

Vol 1, p 140
Vol 1, p. 141

US Army Divisions
going to France in
WW I

Leonard P. Ayres
_The War with Germany_
1919

Vol 2, p. 36

Michel E. Turgot
Louis Bretz

_Plan de Paris_
1739
Viet Nam Memorial in Washington D.C.

Maya Ying Lin

58,000+ dead soldiers
Names listed chronologically by death
Design Principles

- Use small multiples
  - Repeat visually similar graphical elements nearby rather than spreading far apart

Vol 1, p. 170

23 hours of LA air pollution
Chromosomes of man, chimpanzee, gorilla & orangutan

Consumer Reports
NY Trains

Vol 2, p. 68

How to draw letters
Spring 2009  
CS 4460/7450  
53

Vol 2, p. 69

Calligraphy

More Recent Additions

Sparklines

Small, repeated graphics (frequently line graphs)

Sparklines: theory and practice

Theory of sparklines (small, simple, data-viz) along with many practical examples of recent sparkline developments. From Edward Tufte’s book Beautiful Evidence.

-- Edward Tufte, May 27, 2004

Sparklines: Intense, Simple, Word-Sized Graphics

Two more common data displays: a area followed by some numbers. For example, a medical patient’s current level of glucose is typically reported in a clinical record as a word and number.
Sparkline Examples

Design Principles

- Show mechanism, process, dynamics, and causality
  - Cause and effect are key
  - Make graphic exhibit causality

Space shuttle case we discussed first day
Design Principles

- Escape flatland
  - Data is multivariate
  - Doesn’t necessarily mean 3D projection
Guide for visitors to Ise Shrine, Japan

Timetable for Java railroad line
Vol 3, p. 90

Music history

Steve Chapple and Reebee Garofalo

Design Principles

- Utilize layering and separation
  - 1+1 = 3 or more
  - Good or bad
Design Principles

• Utilize narratives of space and time
  – Tell a story of position and chronology through visual elements

Vol 1, p.43 & Vol 2, p 110

Life of a beetle
L. Hugh Newman
Czech air schedule

China railway timetable
Design Principles

- Content is king
  - Quality, relevance and integrity of the content is fundamental
  - What’s the analysis task? Make the visual design reflect that
  - Integrate text, chart, graphic, map into a coherent narrative

Graph and Chart Tips

- Avoid separate legends and keys -- Just have that information in the graphic
- Make grids, labeling, etc., very faint so that they recede into background
Before

After

New Jersey Transit

Vol 2, p. 54

Spring 2009 CS 4460/7450 71

Vol 2, p. 63

Spring 2009 CS 4460/7450 72
Using Color Effectively

- "The often scant benefits derived from coloring data indicate that even putting a good color in a good place is a complex matter. Indeed, so difficult and subtle that avoiding catastrophe becomes the first principle in bringing color to information: *Above all, do no harm.*"
Proper Color Use

- To label
- To measure
- To represent or imitate reality
- To enliven or decorate

Examples

- The bad...
Description

"...despite its clever and multifunctioning data measure, formed by crossing two four-colored grids, this is a puzzle graphic. Deployed here, in a feat of technological virtuosity, are 16 shades of color spread on 3,056 counties, a monument to a sophisticated computer graphics system. But it is surely a graphic experienced verbally not visually. Over and over, the viewers must run little phrases through their minds, trying to maintain the right pattern of words to make sense of the visual montage: "Now let's see, purple represents counties where there are both high levels of male cardiovascular disease mortality and 11.6 to 56.0 percent of the households have more than 1.01 persons per room..."
“Color’s multidimensionality can also enliven and inform what users must face at computer terminals, although some color applied to display screens has made what should be a straight-forward tool into something that looks like a grim parody of a video game.”
Examples

- The good...
Vol 2, p. 91 & Vol 3, p. 76

Swiss Mountain Map
**Guides for Enhancing Visual Quality**

- Attractive displays of statistical info
  - have a properly chosen format and design
  - use words, numbers and drawing together
  - reflect a balance, a proportion, a sense of relevant scale
  - display an accessible complexity of detail
  - often have a narrative quality, a story to tell about the data
  - are drawn in a professional manner, with the technical details of production done with care
  - avoid content-free decoration, including chartjunk

---

**Information Overload**

What about confusing clutter? Information overload? Doesn't data have to “boiled down” and “simplified”? These common questions miss the point, for the quantity of detail is an issue completely separate from the difficulty of reading. *Clutter and confusion are failures of design, not attributes of information.* Often the less complex and less subtle the line, the more ambiguous and less interesting is the reading. Stripping the detail out of data is a style based on personal preference and fashion, considerations utterly indifferent to substantive content.  

*Vol. 2, p. 51*
Minard graphic

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

Graphical Displays Should

- Show the data
- Induce the viewer to think about substance rather than about methodology, graphic design the technology of graphic production, or something else
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with statistical and verbal descriptions of a data set
Website & Seminar

Presenting data and information:
A one-day course taught by Edward Tufte
Topics covered in the one-day workshop include:
- Fundamental strategies of analytical design
- Evaluating evidence used in presentations
- Statistical data tables, graphics, and tables
- Business, scientific, research, and financial presentation
- Replication and clarity
- Effective presentation on paper and screens
- Visual design
- Use of fonts, rules, color, and illustration
- Multimodal, instant, and textual
- Concepts of presentation
- Sensation and cognitive aesthetics
- Recent graphical examples

"Our workshop will bring the insights of the best human information design to the forefront of your experience."- Edward Tufte

Spring 2009 CS 4460/7450

Discussion Forum

As the semester continues, I will answer questions dealing with information design. Others can then retell the discussion. I will try to answer questions that have general interest or when I have something to say. Not all questions will be answered, usually because I don't know the answer.

*** = 3 star formula

Spring 2009 CS 4460/7450
Upcoming

- Time series data
  - Papers to discuss
    Aigner

Sources Used

E. Tufte, *The Visual Display of Quantitative Information*
E. Tufte, *Envisioning Information*
E. Tufte, *Visual Explanations*