Tufte's Design Principles

CS 4460 – Intro. to Information Visualization
November 18, 2014
John Stasko

Please see appropriate
books for missing images

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 Professor Emeritus at Yale University, where he taught courses in statistical evidence, information design, and interface design. His current work includes landscape sculpture, printmaking, video and a new book.

This website describes Edward Tufte's books, one-day course, and artwork. For further information, call Graphics Press at 203 272-9187, or fax 203 272-8600, or email.

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

BOOKS
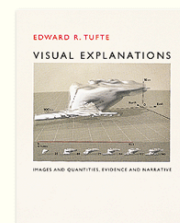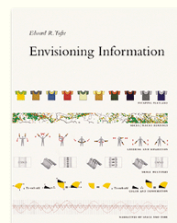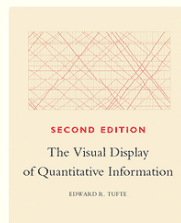
Graphical Excellence

- 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

Fall 2014 CS 4460 3

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

Fall 2014 CS 4460 4

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

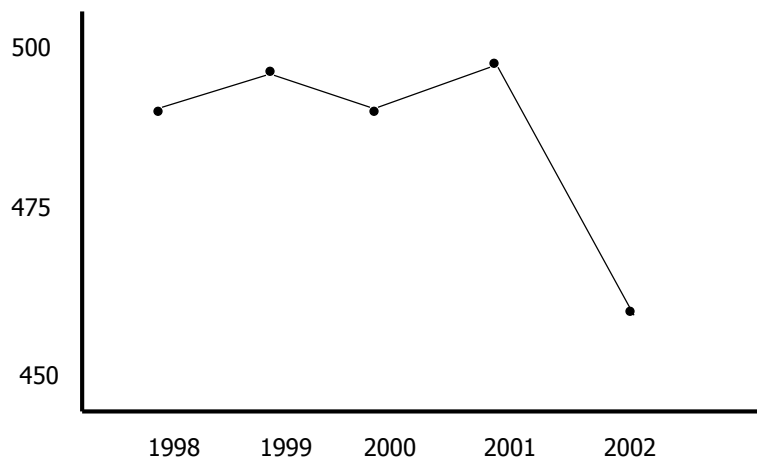- Your graphic should tell the truth about your data

Example   Stock market crash?

Example
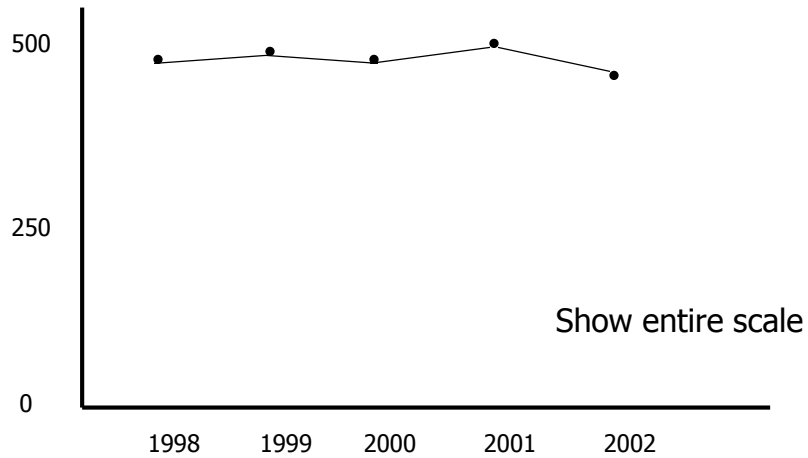

500

250

Show entire scale

0

1998   1999   2000   2001   2002

Fall 2014   CS 4460   9

Example

500

250

Show in context

0

1960   1970   1980   1990   2000
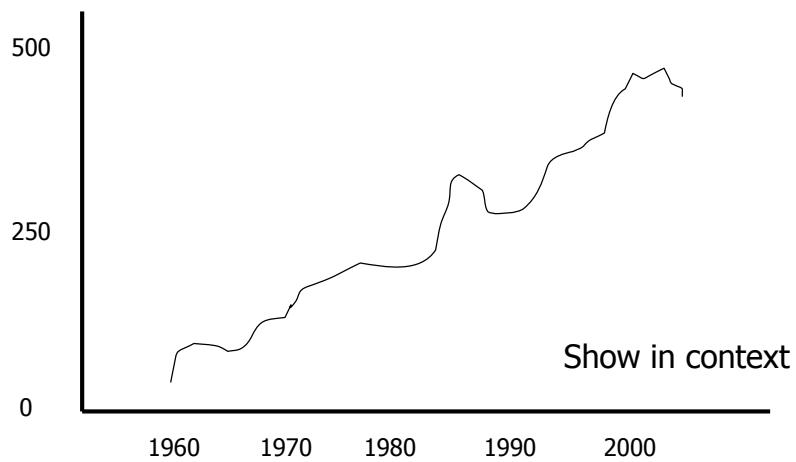
Fall 2014   CS 4460   10

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?

Vol 1, p. 74

Great work!

Vol 1, p. 74

Ahhhh

Show the context

Local Example

A huge rise?

Atlanta Journal Constitution
Summer '08

More of the data

Atlanta Journal Constitution Dec. '08
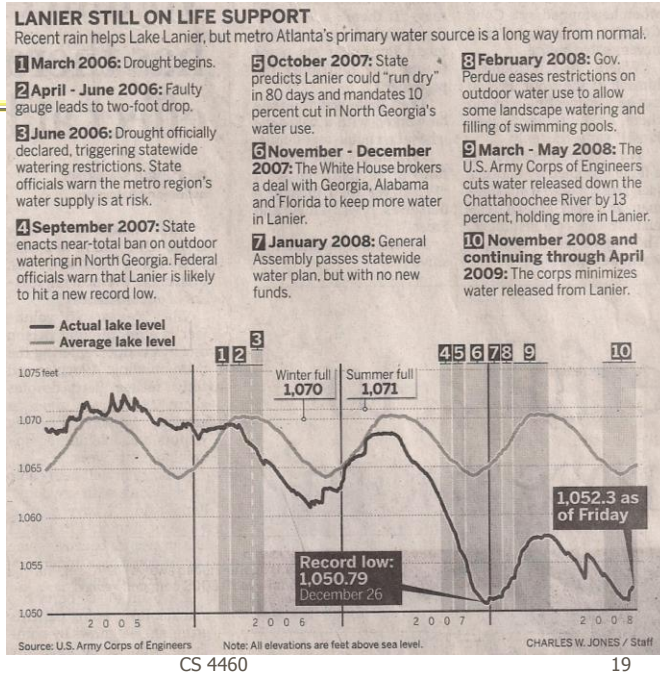
Watch Size Coding

- Height/width vs. area vs. volume

Vol 1, p. 69

area = value?

Vol 1, p. 62

volume = 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 $9.4 = \dfrac{4280}{454}$

2. Design Aesthetics

- Set of principles to help guide designers

Design Principles

- Maximize data-ink ratio

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

= proportion of graphic's ink devoted to the non-redundant display of data-information

Vol 1, p. 94

Good   Bad

Vol 1, p. 30

Outstanding

CS 4460 27

More...

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

CS 4460 28

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
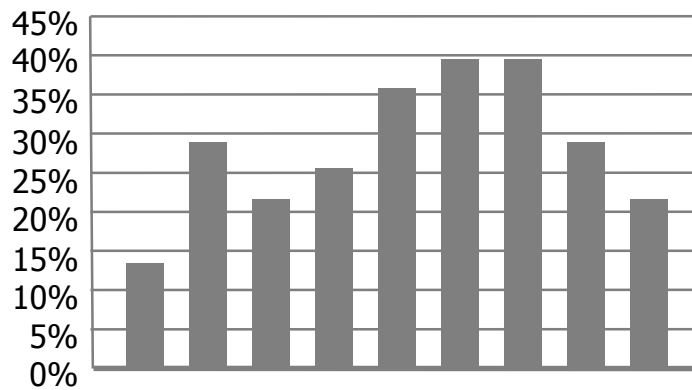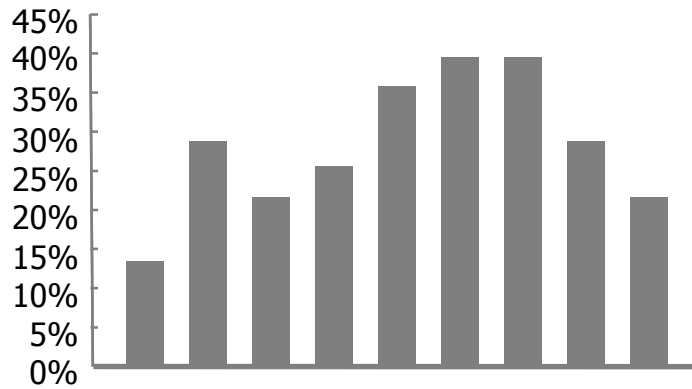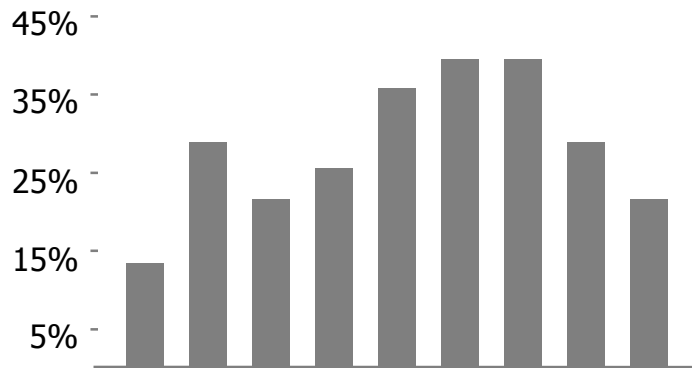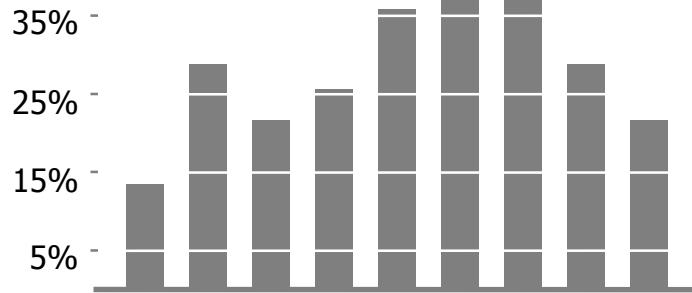
Bar chart

Bar chart

Bar chart

Bar chart

CS 4460

Bar chart
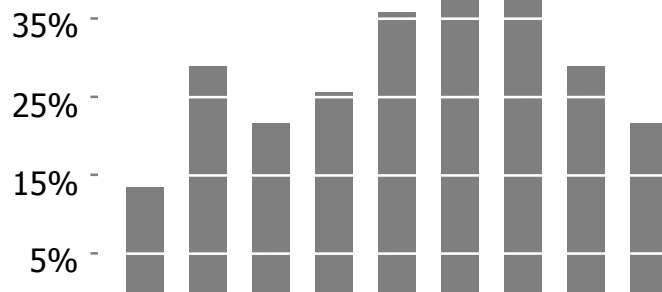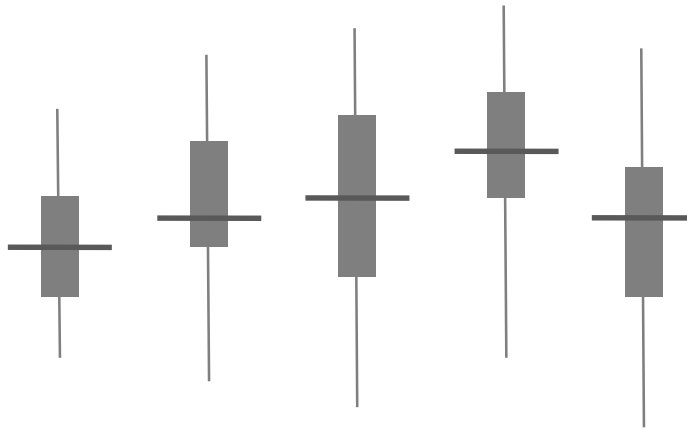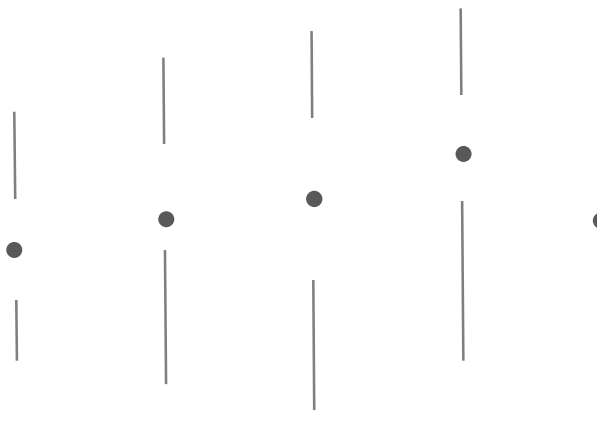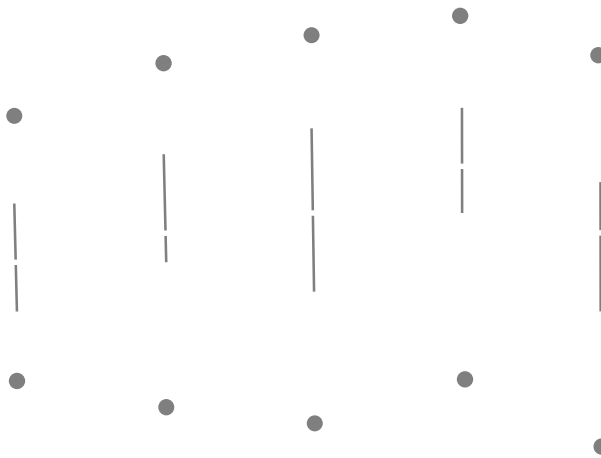
CS 4460

Box plot

Box plot

Box plot

Box plot
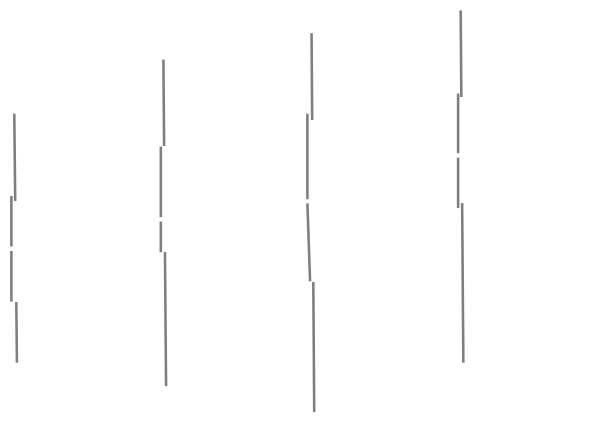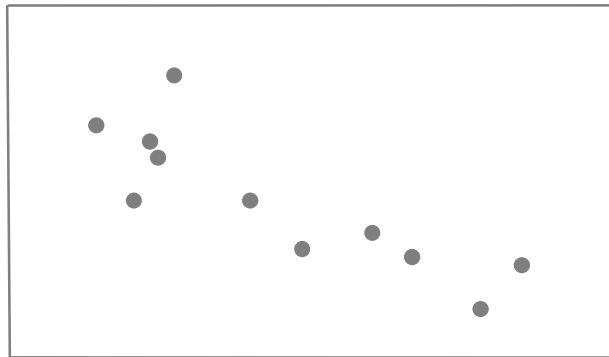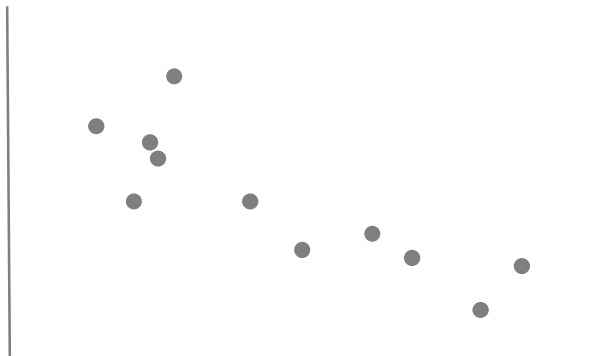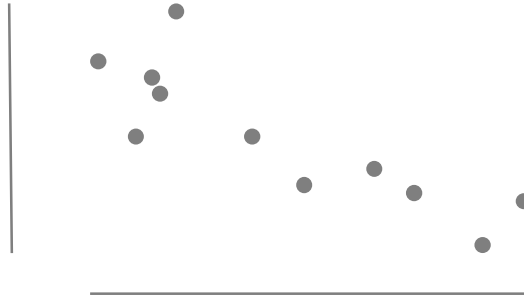
Scatter plot

Scatter plot

Scatter plot

Scatter plot
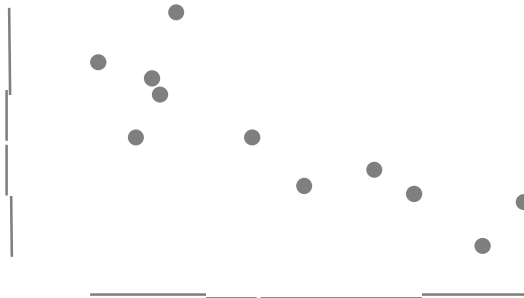
Design Principles

- 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

http://www.usatoday.com/news/snapshot.htm

Junk Charts Blog

`http://junkcharts.typepad.com/`
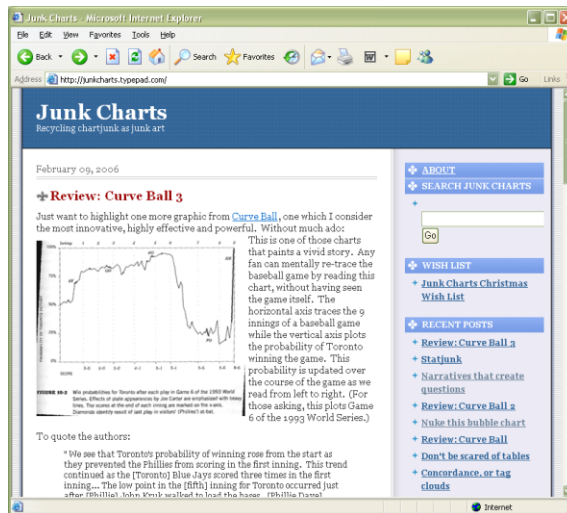
More Thoughts

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

Rethink That?

Compared plain charts to "embellished" charts

Found that the embellished charts were just as good on interpretation accuracy and were recalled better weeks later

Participants also preferred the embellished ones

Some caveats:
   Very simple data
   Very plain plain charts
   Each chart/data is different

My take: It's all about purpose

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

Vol 2, p. 37

Manhattan 1989
Manhattan Map Company

Vol 2, p. 42

Viet Nam Memorial
in Washington D.C.

Maya Ying Lin

58,000+ dead soldiers

Vol 2, p. 44

Vol 2, p. 43

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

Vol 1, p. 173

Chromosomes of
man, chimpanzee,
gorilla & orangutan

Vol 1, p. 174

Consumer
Reports

Vol 2, p. 68

NY Trains

Vol 2, p. 68

How to draw letters

Vol 2, p. 69

Calligraphy

More Recent Additions

Sparklines

Small, repeated graphics (frequently line graphs)

Sparklines: theory and practice

Theory of sparklines (small, intense, simple datawords) along with many practical examples of recent sparkline developments. From Edward Tufte's book *Beautiful Evidence*.

-- Edward Tufte, May 27, 2004

Sparklines: theory and practice

Theory of sparklines (small, intense, simple datawords) along with many practical examples of recent sparkline developments. Excerpts from Edward Tufte, *Beautiful Evidence*. New examples or helpful comments much appreciated.

ET

Sparklines: Intense, Simple, Word-Sized Graphics

THE most common data display is a noun 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

Vol 3, p. 144

Washington Post

Design Principles

- Escape flatland
  - Data is multivariate
  - Doesn't necessarily mean 3D projection

Vol 2, p. 12

Guide for visitors to Ise Shrine, Japan

Vol 2, p. 24

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

Vol 2, p. 54   IBM Series III Copier

Vol 2, p. 61

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

Vol 2, p. 102

Czech air schedule

Vol 2, p. 103

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

Vol 2, p. 54   New Jersey Transit

Before

After

Vol 2, p. 63

Before

After

Vol 3, p. 74

Before   After

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...

Vol 1, p. 153

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 virtuousity, 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…"

Vol 2, p. 82

Vol 2, p. 88

"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."

Vol 3, p. 77

Examples

- The good...

Vol 2, p. 91 & Vol 3, p. 76

Vol 2, p. 80

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 difficultly 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

size of army   latitude   temperature
direction   longitude   date

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

Discussion Forum

Interesting Contrast

Nigel Holmes
http://www.nigelholmes.com

HW 7

- Tableau use & review
- Due Thursday
- Follow the assignment
- Any questions?

Project

- Deliverables
  - Presentation in class (last week)
  - Video
    - 6 minutes max, due on Monday 8th at 5pm
    - Send 1- or 2-page info flier too

Video Advice

- Use Camtasia
  - demo trial version available
- Process
  - 1. Develop script (rehearse timing)
  - 2. Record script
  - 3. Capture video of demo to script
  - 4. Add effects
- You've seen examples all semester
  - eg, http://www.cc.gatech.edu/gvu/ii/videos.html

Fall 2014   CS 4460   107

Upcoming

- Visual Analytics
  - Reading

- Evaluation
  - Reading

Fall 2014   CS 4460   108

Sources Used

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