

# Tufte's Design Principles



CS 7450 - Information Visualization  
 September 17, 2012  
 John Stasko

Please see appropriate  
 books for missing images

## Today's Agenda

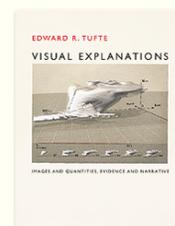
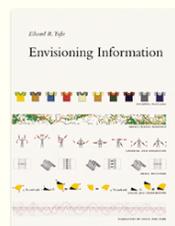
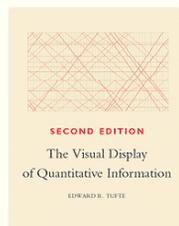


Edward Tufte has written seven books, including *Visual Explanations*, *Envisioning Information*, *The Visual Display of Quantitative Information*, and *Data Analysis for Politics and Policy*. He writes, designs, and self-publishes his books on analytical design, which have received more than 40 awards 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

# 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, synopsisize, 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

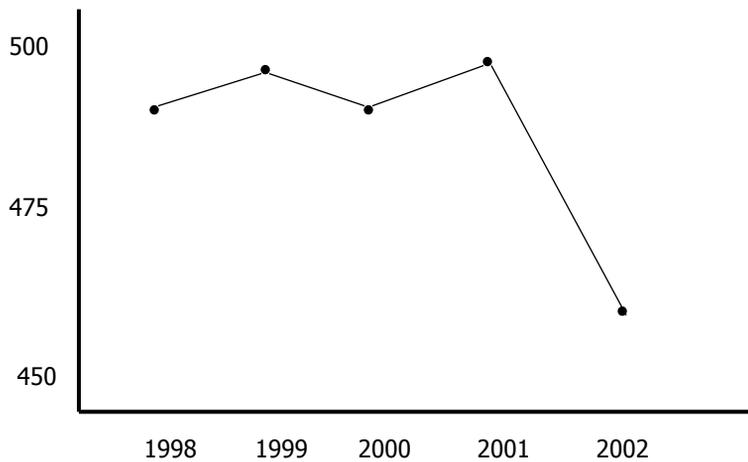
Fall 2012

CS 7450

7

## Example

Stock market crash?

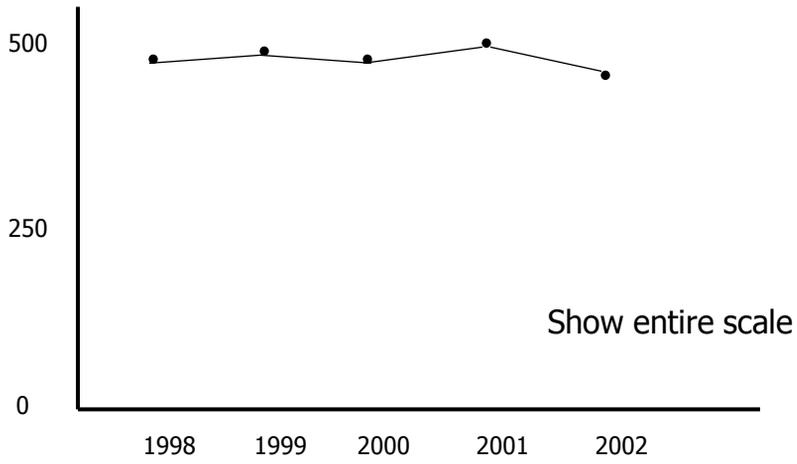


Fall 2012

CS 7450

8

# Example

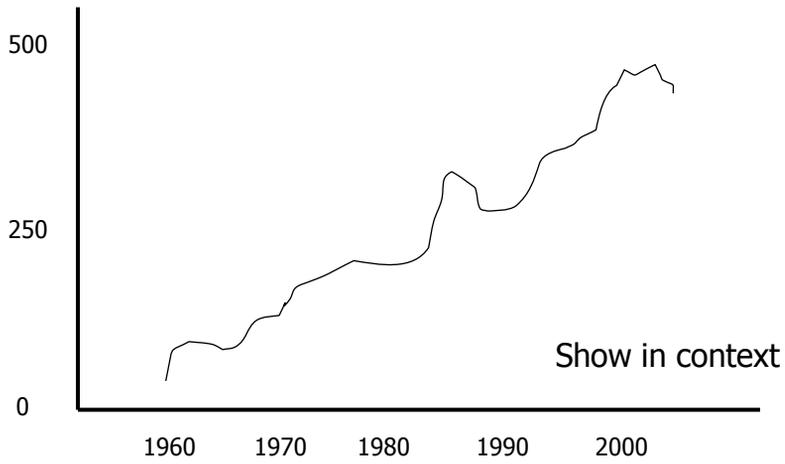


Fall 2012

CS 7450

9

# Example



Fall 2012

CS 7450

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?

Fall 2012

CS 7450

13

---

## Vol 1, 57



Scale?

Fall 2012

CS 7450

14

# Vol 1, p. 61



Scale?

Fall 2012

CS 7450

15

---

# Vol 1, p. 74



Great work!

Fall 2012

CS 7450

16

# Vol 1, p. 74



Ahhhh

Show the context

Fall 2012

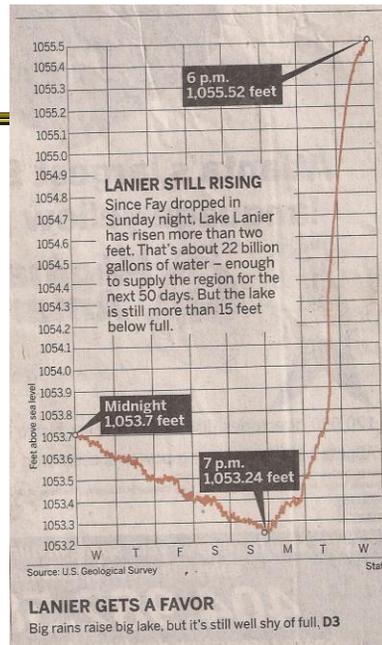
CS 7450

17

## Local Example

A huge rise?

Atlanta Journal Constitution  
Summer '08



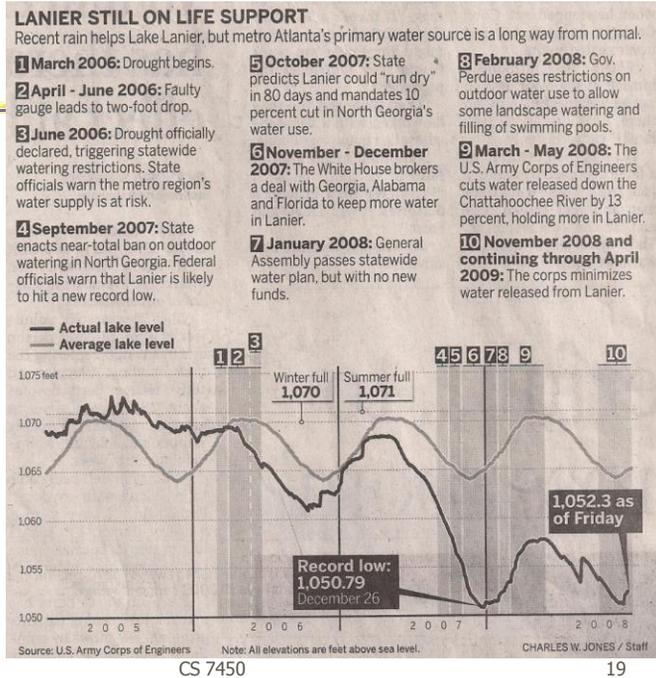
Fall 2012

CS 7450

18

More of the  
data

Atlanta Journal  
Constitution  
Dec. '08



## Watch Size Coding



- Height/width vs. area vs. volume

## Vol 1, p. 69



area = value?

Fall 2012

CS 7450

21

---

## Vol 1, p. 62



volume = value?

Fall 2012

CS 7450

22

## Measuring Misrepresentation



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

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

$$\text{p.62} \quad 9.4 = \frac{4280}{454}$$

## 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

Fall 2012

CS 7450

27

---

## More...



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

Fall 2012

CS 7450

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



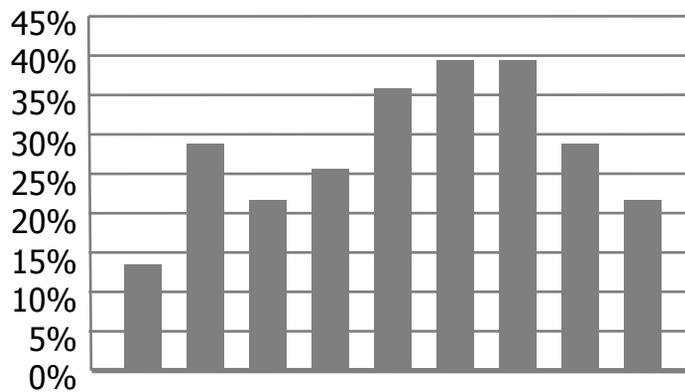
- Bar chart, scatter plot, box plot

Fall 2012

CS 7450

31

## Bar chart

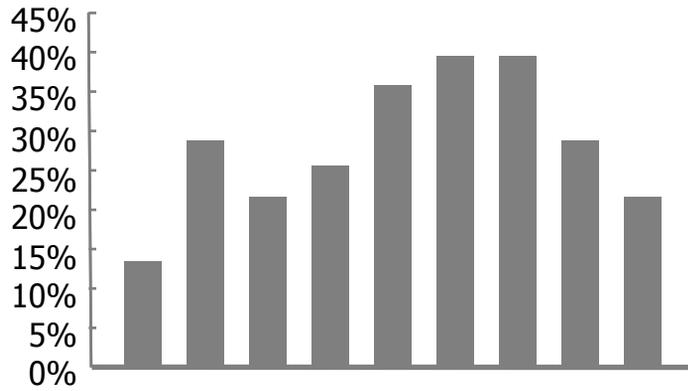


Fall 2012

CS 7450

32

# Bar chart

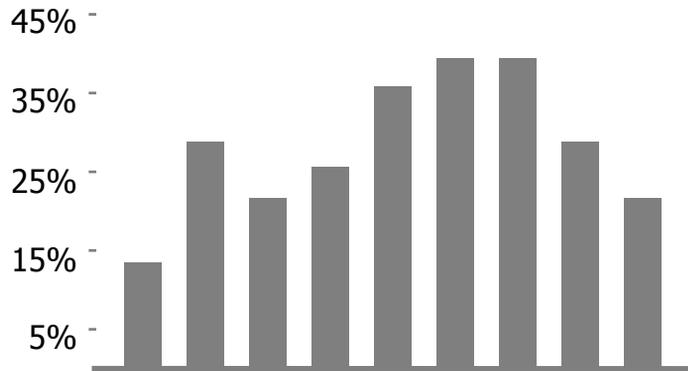


Fall 2012

CS 7450

33

# Bar chart

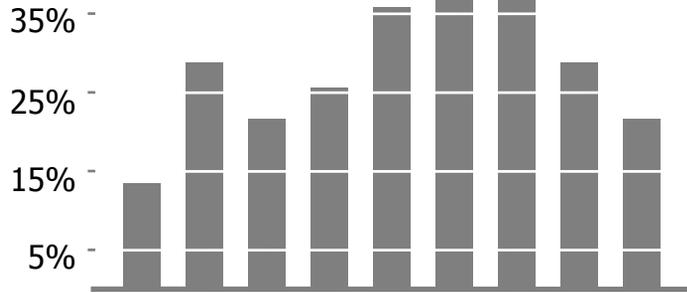


Fall 2012

CS 7450

34

# Bar chart

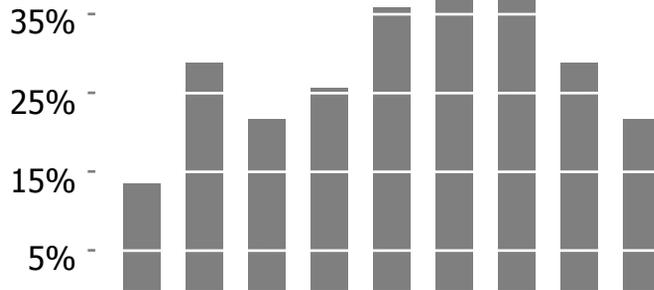


Fall 2012

CS 7450

35

# Bar chart

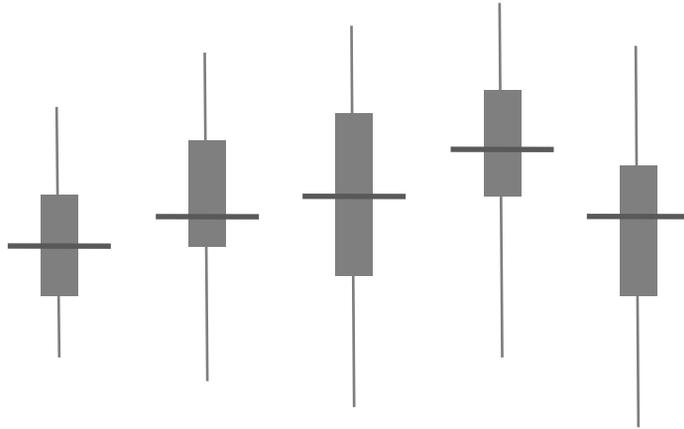


Fall 2012

CS 7450

36

# Box plot

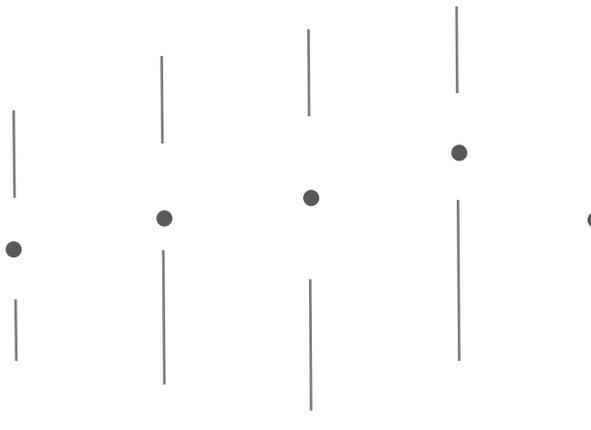


Fall 2012

CS 7450

37

# Box plot

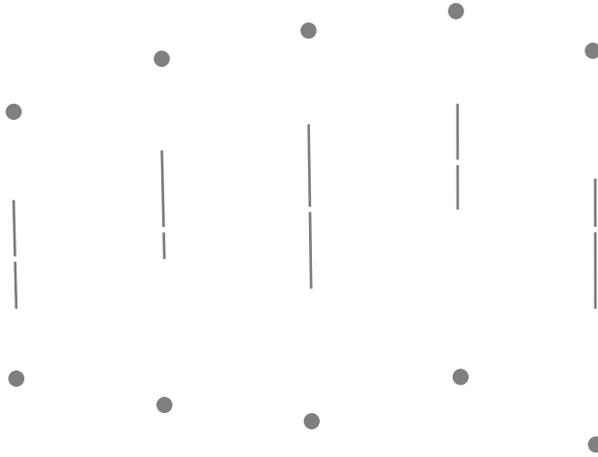


Fall 2012

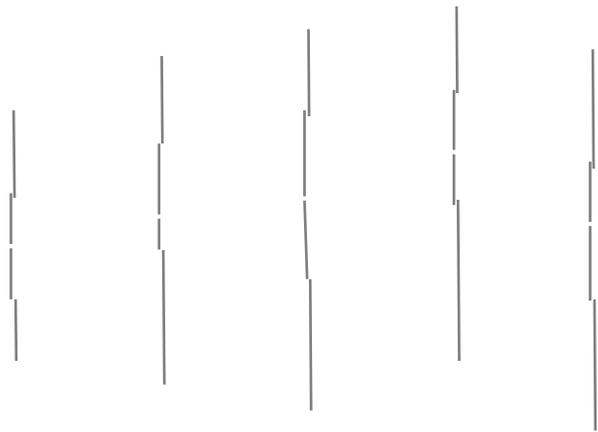
CS 7450

38

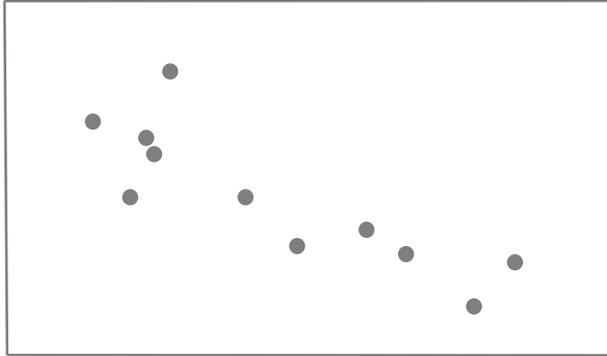
# Box plot



# Box plot



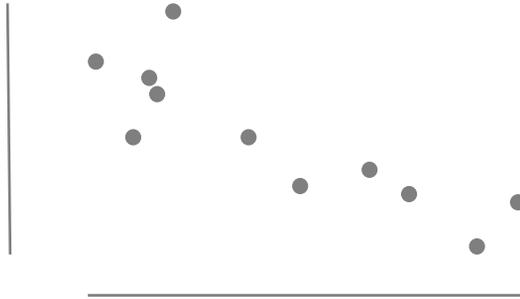
# Scatter plot



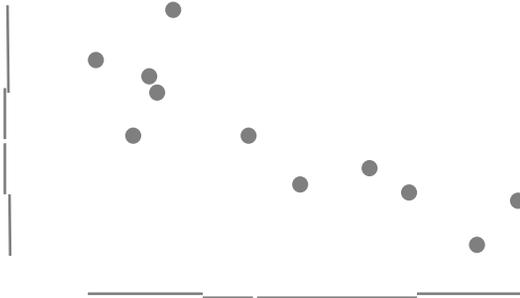
# Scatter plot



# 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

Fall 2012

CS 7450

47

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

## USA Today



USA TODAY Snapshot

On Christmas Eve, 48% of adults who say they leave snacks for Santa set out cookies and milk.

Snack	Percentage
Cookies and milk	48%
Cookies, milk and cornets	29%
Cookies	2%

By Anne R. Capp and Ann-Cate. USA TODAY Source: Opinion Research Corp. No. Leland

Advertisement

symantec

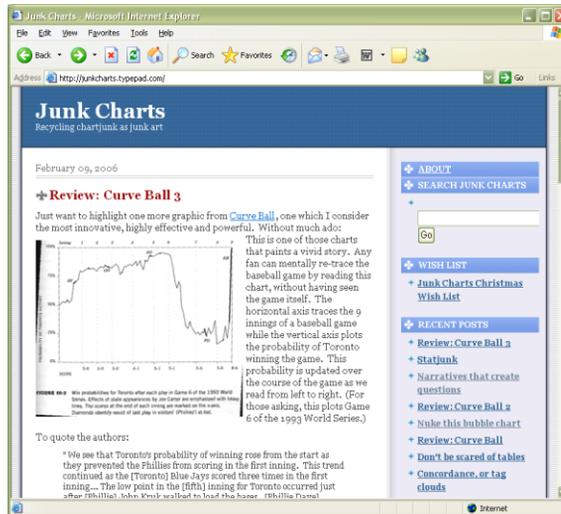
Fall 2012

CS 7450

48

http://junkcharts.typepad.com/

# Junk Charts Blog



Fall 2012

CS 7450

49

## More Thoughts



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

Fall 2012

CS 7450

50

# Rethink That?



CHI 2010 - Graphs April 10-15, 2010, Atlanta, GA, USA

## Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandyk, Carl Gutwin, Aaron Gervet, David McDine, Christopher Brooks  
Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada  
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca, aaron.gervet@usask.ca, david@cs.usask.ca, cbrooks@usask.ca

**ABSTRACT**  
Cautions for designing information charts often state that the presentation should reduce "chart junk" or "visual embellishments" that are not essential to understanding the data. In contrast, some graphic chart designers argue the presented data is distilled and enhanced imagery, raising the question of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both comprehension accuracy and long-term recall. We found that people's accuracy in describing the embellished charts was no worse than the plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be published in this style, our results question some of the premises of the minimalist approach to chart design.

**Author Keywords:**  
Charts, information visualization, imagery, memorability.

**ACM Classification Keywords:**  
H.1.1. Information interfaces and presentation (e.g., HCI); M.0.0.0. Miscellaneous

**General Terms:**  
Design, Human Factors

**INTRODUCTION**  
Many experts in the area of chart design, such as Edward Tufte, criticize the inclusion of visual embellishments in charts and graphics. Their guidelines for good chart design often require that the addition of chart junk, decorations and other kinds of non-essential imagery, to a chart can make comprehension more difficult and can obscure trends from the data [2]. This minimalist perspective advocates plain and simple charts that minimize the proportion of data out- or in- the way to the data used to represent data.

Researchers who adopt a hard stance of all or part of this will be surprised to discover that a general consensus has formed that the use of such visual embellishments is not as detrimental to comprehension and long-term recall as is often claimed. In fact, as we review in the introduction to this paper, some research has shown that the use of such visual embellishments can actually improve comprehension and long-term recall.

Copyright 2010 ACM 978-1-60558-324-1/10/04...\$5.00.

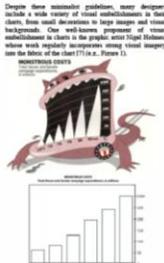


Figure 1. A chart for Fishes (7) (above), and a 'plain' version. These kinds of charts appear regularly in many mass-media publications, but the widespread use of embellished design raises questions about whether the minimalist position on chart design is really the better approach. The issue is particularly an unmet first, whether visual embellishments do in fact cause comprehension problems, and second, whether the embellishments may provide additional information that is valuable for the reader. For example, the added visual imagery in a fish-themed chart could draw

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

Fall 2012

CS 7450

51

# Design Principles



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

Fall 2012

CS 7450

52

## Vol 1, p 140



Fall 2012

CS 7450

53

---

## Vol 1, p. 141



US Army Divisions  
going to France in  
WW I

Leonard P. Ayres  
*The War with Germany*  
1919

Fall 2012

CS 7450

54

## **Vol 2, p. 36**



Michel E. Turgot  
Louis Bretz

*Plan de Paris*  
1739

Fall 2012

CS 7450

55

---

## **Vol 2, p. 37**

Manhattan 1989  
Manhattan Map Company



Fall 2012

CS 7450

56

## **Vol 2, p. 42**



Viet Nam Memorial  
in Washington D.C.

Maya Ying Lin

58,000+ dead soldiers

Fall 2012

CS 7450

57

---

## **Vol 2, p. 44**



Fall 2012

CS 7450

58

## 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

Fall 2012

CS 7450

61

---

## **Vol 1, p. 173**



Chromosomes of  
man, chimpanzee,  
gorilla & orangutan

Fall 2012

CS 7450

62

# Vol 1, p. 174



Consumer  
Reports

Fall 2012

CS 7450

63

---

# Vol 2, p. 68



NY Trains

Fall 2012

CS 7450

64

## **Vol 2, p. 68**



How to draw letters

Fall 2012

CS 7450

65

---

## **Vol 2, p. 69**



Calligraphy

Fall 2012

CS 7450

66

## More Recent Additions



### 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

### 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. 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:

Fall 2012

CS 7450

67

## Sparkline Examples



Fall 2012

CS 7450

68

# 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

---

Fall 2012

CS 7450

71

# Vol 2, p. 12



Guide for visitors to Ise Shrine, Japan

---

Fall 2012

CS 7450

72

## Vol 2, p. 24



### Timetable for Java railroad line

Fall 2012

CS 7450

73

---

## Vol 3, p. 90

Music history



### Steve Chapple and Reebee Garofalo

Fall 2012

CS 7450

74

# 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



---

Fall 2012

CS 7450

77

## Design Principles



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

---

Fall 2012

CS 7450

78

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



Life of a beetle

L. Hugh Newman

Fall 2012

CS 7450

79

---

## **Vol 2, p. 102**



Czech air schedule

Fall 2012

CS 7450

80

## Vol 2, p. 103



China railway  
timetable

Fall 2012

CS 7450

81

## 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

Fall 2012

CS 7450

82

## 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

---

Fall 2012

CS 7450

83

## Vol 2, p. 54

New Jersey Transit



Before

After

---

Fall 2012

CS 7450

84

## Vol 2, p. 63



Before

After

Fall 2012

CS 7450

85

---

## Vol 3, p. 74



Before

After

Fall 2012

CS 7450

86

## 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 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..."

Fall 2012

CS 7450

91

## Vol 2, p. 82



Fall 2012

CS 7450

92

## 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."

Fall 2012

CS 7450

93

---

## Vol 3, p. 77



Fall 2012

CS 7450

94

# Examples



- The good...

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



## Vol 2, p. 80



Swiss Mountain Map

Fall 2012

CS 7450

97

## 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

Fall 2012

CS 7450

98

# 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**

Fall 2012

CS 7450

99

# Minard graphic



size of army  
direction

latitude  
longitude

temperature  
date

Fall 2012

CS 7450

100

# 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

Fall 2012

CS 7450

101

# Website & Seminar



The Work of Edward Tufte and Graphics Press - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.edwardtufte.com/tufte/

HOME BOOKS COURSES POSTERS AND GRAPH PAPER FINE ART SCULPTURE ASK E.T. SHOPPING CART! CHECKOUT

ET WRITINGS, ARTWORKS, NEWS DOGS & OTHERS OF GRAPHICS PRESS POWERPOINT ESSAY

THE WORK OF EDWARD TUFTTE AND GRAPHICS PRESS  
GRAPHICS PRESS LLC, P.O. BOX 430, CHESHIRE, CT 06410 800 822-2454 FAX 203 272-8600

PRESENTING DATA AND INFORMATION:  
A ONE-DAY COURSE TAUGHT BY EDWARD TUFTTE

Topics covered in this one-day course include:

- fundamental strategies of analytical design
- evaluating evidence used in presentations
- statistical data tables, graphics, and semi-graphics
- business, scientific, research, and financial presentations
- complexity and clarity
- effective presentations: on paper and in person
- interface design
- use of PowerPoint, video, overheads, and handouts
- multi-media, internet, and websites
- credibility of presentations
- animation and scientific visualizations
- many practical examples

Durham, North Carolina [March 9, 2009](#)  
[March 10, 2009](#)

Atlanta, Georgia [March 12, 2009](#)  
[March 13, 2009](#)

Arlington, Virginia [April 6, 2009](#)  
[April 7, 2009](#)  
[April 8, 2009](#)

Seattle, Washington [April 28, 2009](#)  
[April 29, 2009](#)

Portland, Oregon [May 1, 2009](#)

"One visionary day...the insights of this class lead to new levels of understanding both for creators and viewers of visual displays." [Wired](#)

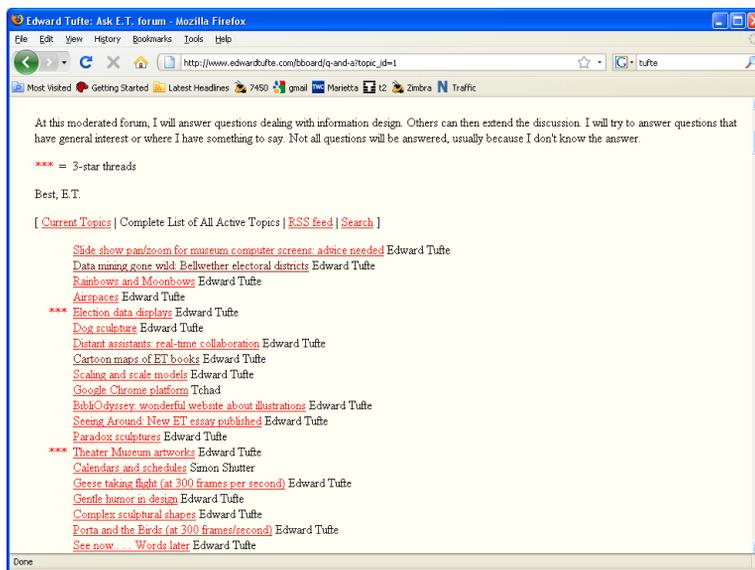
"The Leonardo da Vinci of data" [THE NEW YORK TIMES](#)

Fall 2012

CS 7450

102

# Discussion Forum



Fall 2012

CS 7450

103

# Interesting Contrast



Nigel Holmes

<http://www.nigelholmes.com>Good gallery: <http://simplecomplexity.net/nigel-holmes-gallery/>

Fall 2012

CS 7450

104

## HW 4



- Data analysis with Many Eyes
- Due next Monday

---

Fall 2012

CS 7450

105

## Upcoming



- Few's Design Principles
  - Reading
  - Few chapters 5-12
- InfoVis systems & toolkits
  - Reading
  - Viegas et al, '07

---

Fall 2012

CS 7450

106

## Sources Used

---



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