

Tufte's Design Principles



CS 7450 - Information Visualization
 October 14, 2015
 John Stasko

Please see appropriate books for missing images

Today's Agenda

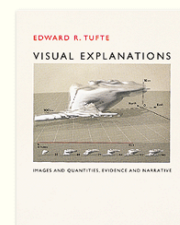
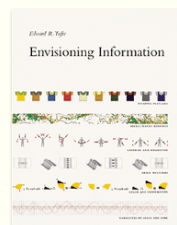
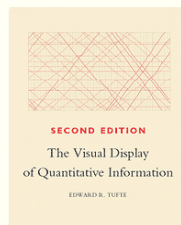


Edward Tufte has written seven books, including *Visual Explanations*, *Envisioning Information*, *The Visual Display of Quantitative Information*, and *Data Analysis for Politics and Policy*. He writes, designs, and self-publishes his books on analytical design, which have received more than 40 awards 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 2015

CS 7450

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 2015

CS 7450

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, synopsise, 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

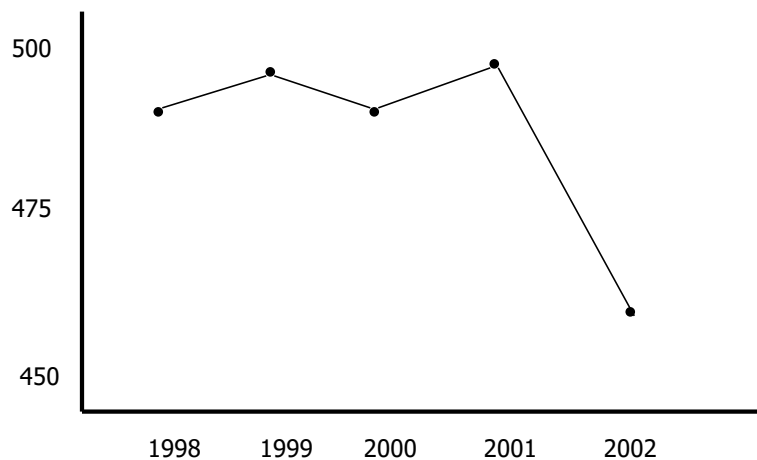
Fall 2015

CS 7450

7

Example

Stock market crash?

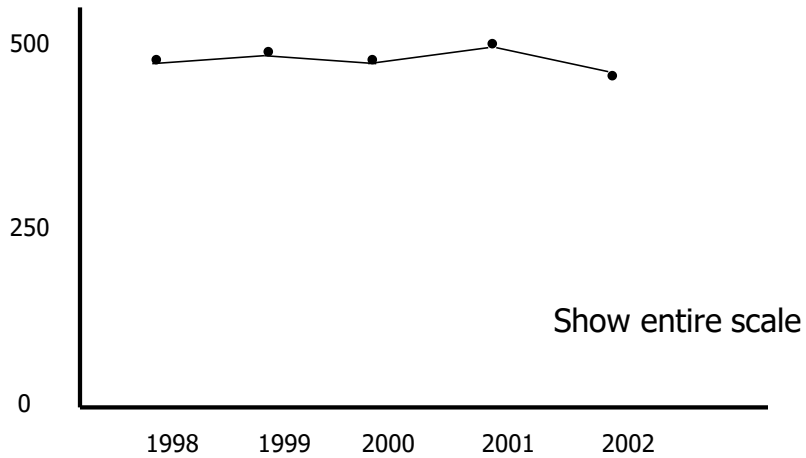


Fall 2015

CS 7450

8

Example

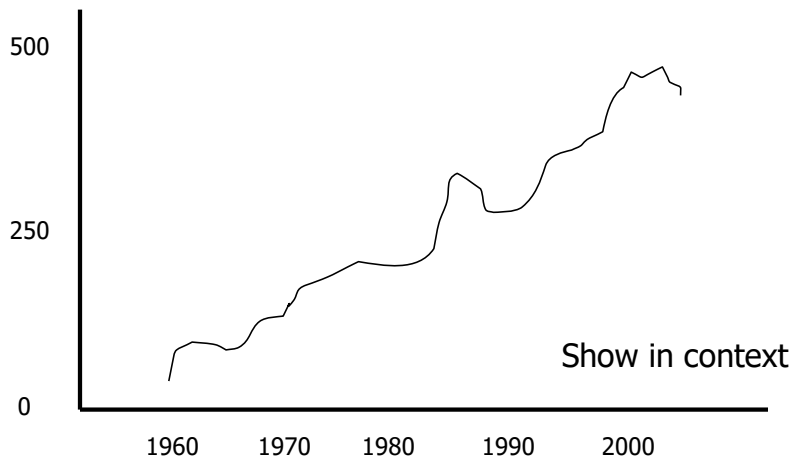


Fall 2015

CS 7450

9

Example



Fall 2015

CS 7450

10

Chart Integrity



- Where's baseline?
- What's scale?
- What's context?

Fall 2015

CS 7450

11

Vol 1, p. 54 (1)

Where's 0?
Note middle '70



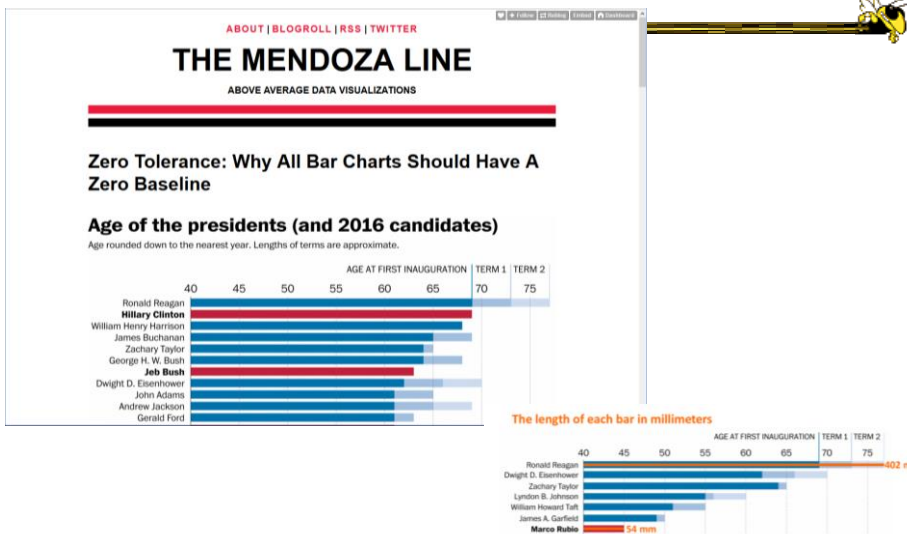
Fall 2015

CS 7450

12

http://themendozaonline.org/post/118146423986/zero-tolerance-why-all-bar-charts-should-have-a

Baseline?

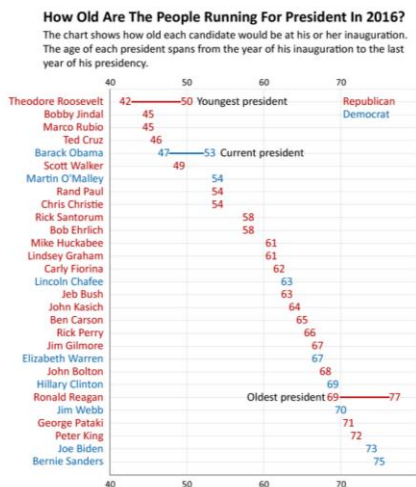


Fall 2015

CS 7450

13

Suggested Redo



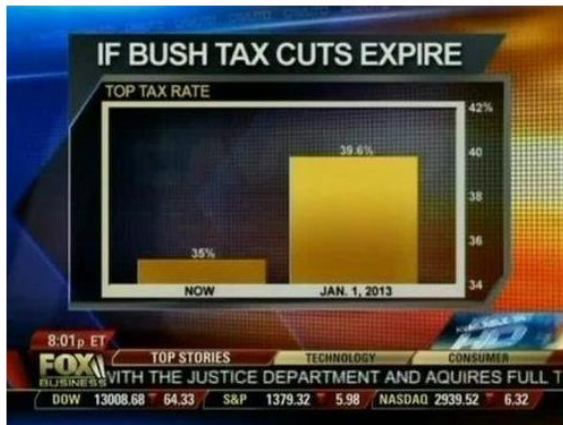
Suggests using dot plots when there is a big range down to zero

Fall 2015

CS 7450

14

Huge Difference?



Compare area of right bar to that of the left bar

Fall 2015

CS 7450

15

Vol 1, p 54 (2)



What's being compared?

Fall 2015

CS 7450

16

Vol 1, 57



Scale?

Fall 2015

CS 7450

17

Vol 1, p. 61



Scale?

Fall 2015

CS 7450

18

Vol 1, p. 74



Great work!

Fall 2015

CS 7450

19

Vol 1, p. 74



Ahhhh

Show the context

Fall 2015

CS 7450

20

Local Example

A huge rise?

Atlanta Journal Constitution
Summer '08



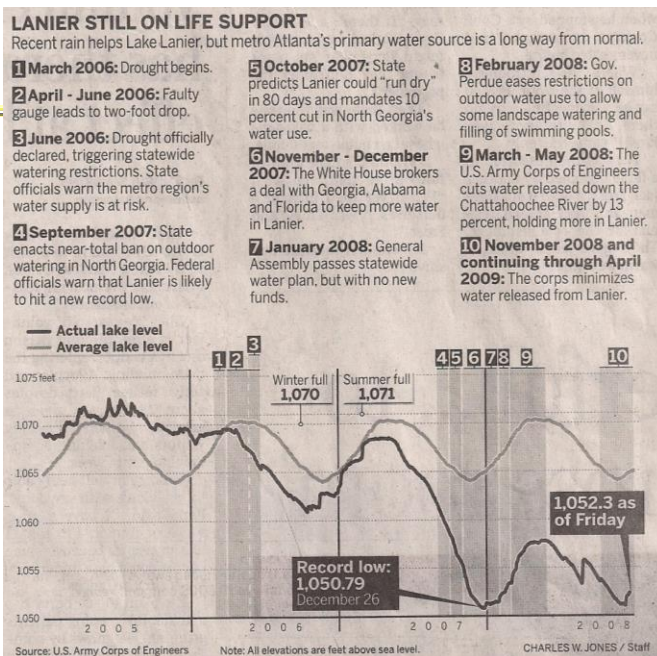
Fall 2015

CS 7450

21

More of the data

Atlanta Journal Constitution
Dec. '08

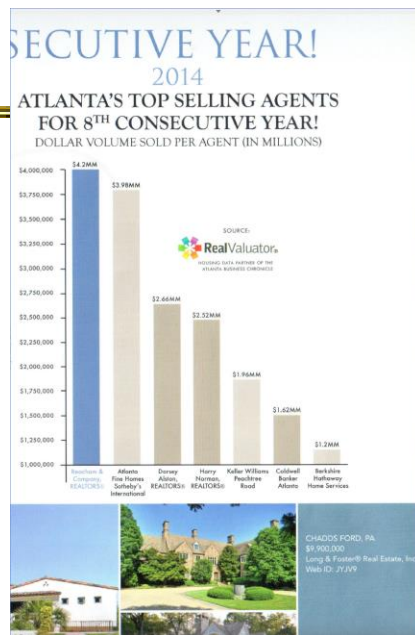


Fall 2015

CS 7450

22

Baselines & heights?



Fall 2015

CS 7450

23

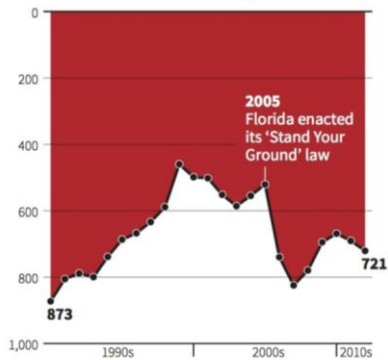
<http://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2>

A Redesign



Gun deaths in Florida

Number of murders committed using firearms

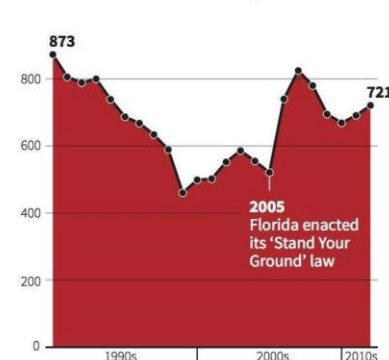


Source: Florida Department of Law Enforcement
C. Chan 16/02/2014



Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

P.A. Follows and Reuters

Fall 2015

CS 7450

24

Watch Size Coding



- Height/width vs. area vs. volume

Fall 2015

CS 7450

25

Vol 1, p. 69



area = value?

Fall 2015

CS 7450

26

Vol 1, p. 62



volume = value?

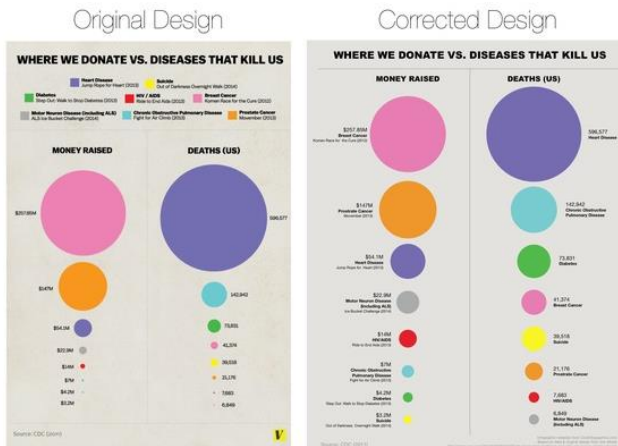
Fall 2015

CS 7450

27

http://www.huffingtonpost.com/randy-krum/false-visualizations-when_b_5736106.html

Circle Width vs. Area



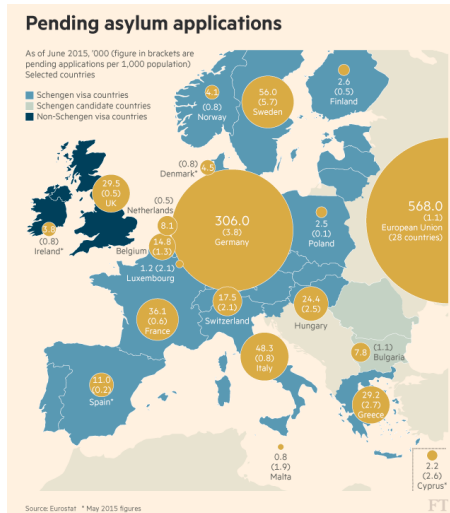
Fall 2015

CS 7450

28

http://www.ft.com/intl/cms/s/2/7f7e0d28-5225-11e5-8642-453585f2cfcf.html

Areas? Not Sure



Europe

Home UK* World Companies* Markets* Global Economy* List* Comment*

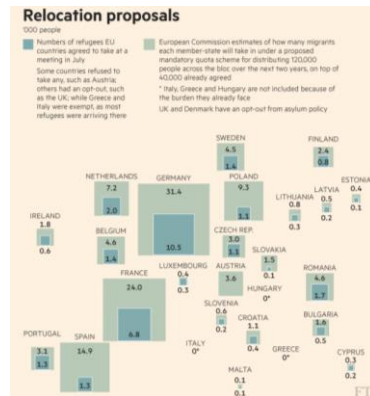
Asia Asia Pacific* Europe* Latin America* Middle East & North Africa* UK* US & Canada*

Last updated September 4, 2015 12:10 pm

Migrant crisis explained in numbers

Share Author alerts Print Clip

The EU is struggling to respond to a surge of desperate migrants, thousands of whom have perished in their attempts to seek a better life in Europe. Where are they going and where are they coming from?



Fall 2015

CS 7450

29

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 $9.4 = \frac{4280}{454}$

Fall 2015

CS 7450

30

2. Design Aesthetics



- Set of principles to help guide designers

Fall 2015

CS 7450

31

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

Fall 2015

CS 7450

32

Vol 1, p. 94



Good

Bad

Fall 2015

CS 7450

33

Vol 1, p. 30



Outstanding

Fall 2015

CS 7450

34

More...



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

Fall 2015

CS 7450

35

More...



- Maximize data density

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

Quote ...

Fall 2015

CS 7450

36

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

Fall 2015

CS 7450

37

Redesign charts



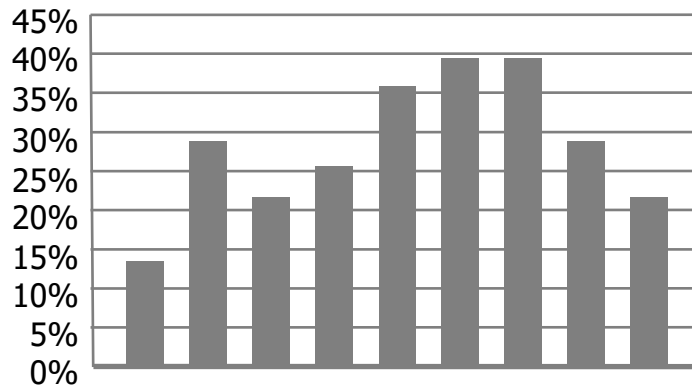
- Bar chart, scatter plot, box plot

Fall 2015

CS 7450

38

Bar chart

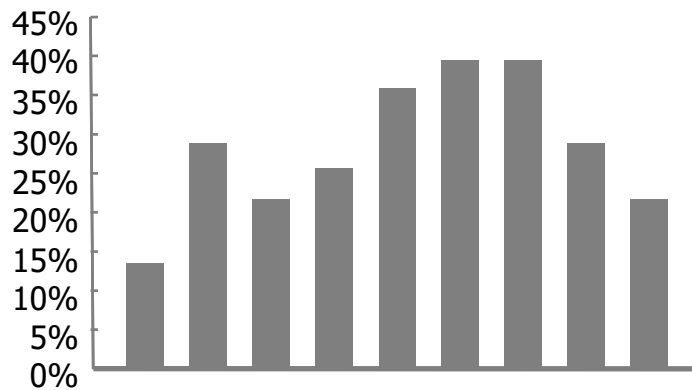


Fall 2015

CS 7450

39

Bar chart

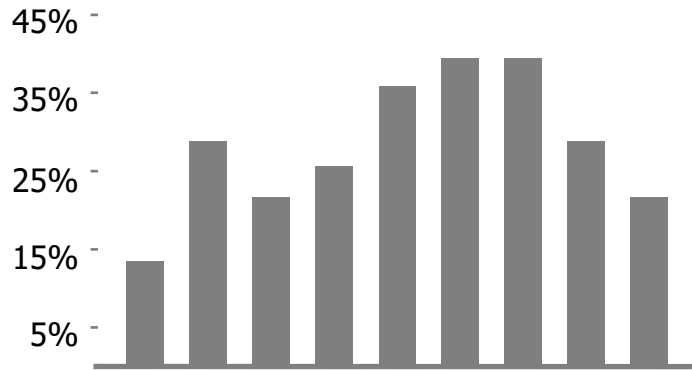


Fall 2015

CS 7450

40

Bar chart

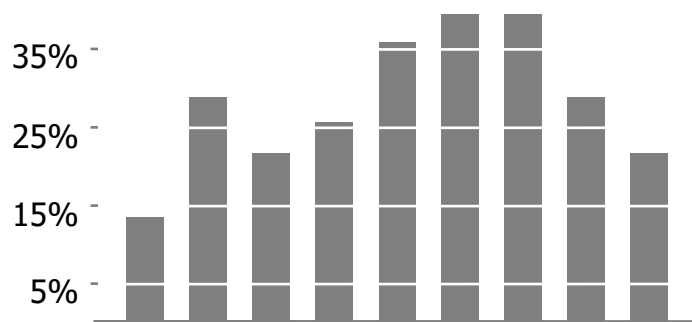


Fall 2015

CS 7450

41

Bar chart

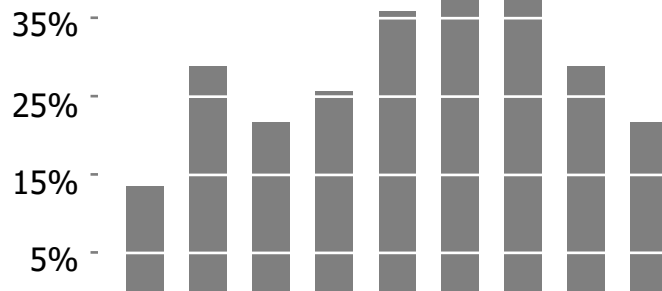


Fall 2015

CS 7450

42

Bar chart

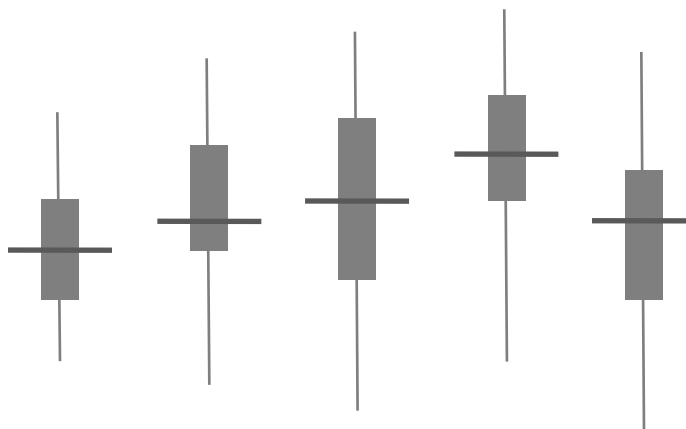


Fall 2015

CS 7450

43

Box plot

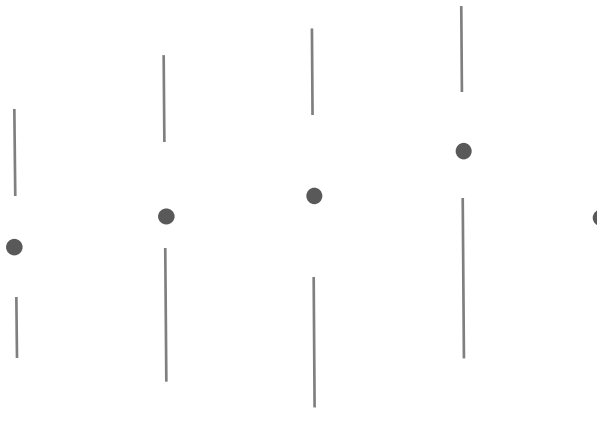


Fall 2015

CS 7450

44

Box plot

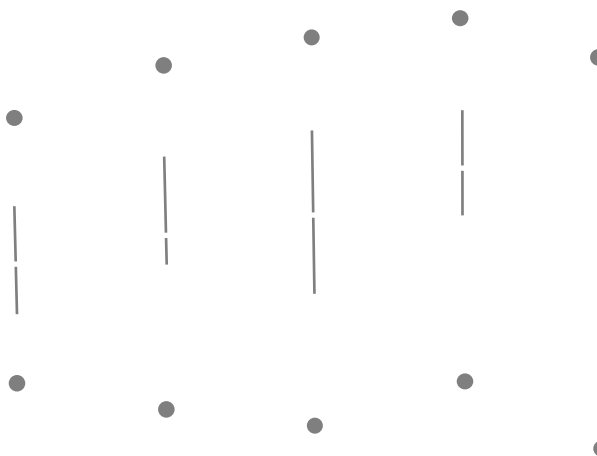


Fall 2015

CS 7450

45

Box plot

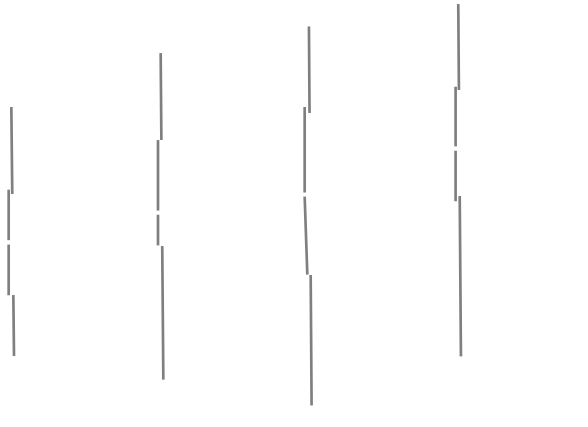


Fall 2015

CS 7450

46

Box plot

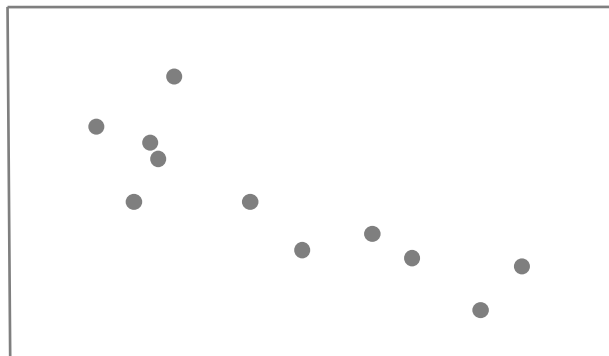


Fall 2015

CS 7450

47

Scatter plot

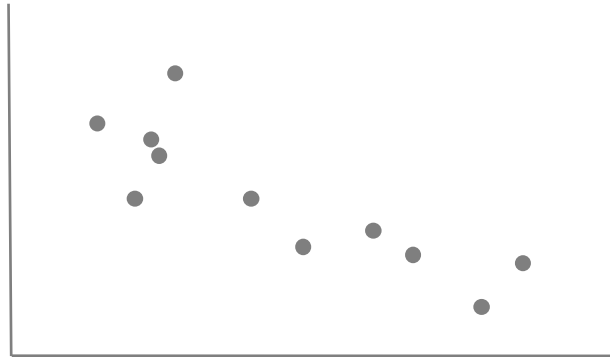


Fall 2015

CS 7450

48

Scatter plot

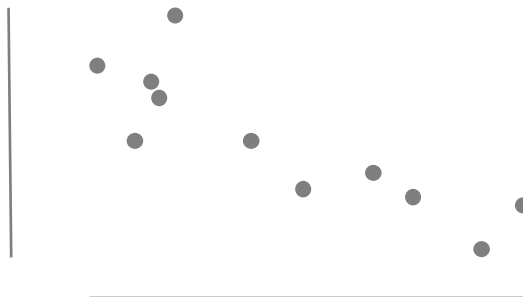


Fall 2015

CS 7450

49

Scatter plot



Fall 2015

CS 7450

50

Scatter plot



Fall 2015

CS 7450

51

Design Principles



- **Avoid chartjunk**
 - Extraneous visual elements that detract from message

Fall 2015

CS 7450

52

Vol 1, p 108



Fall 2015

CS 7450

53

Vol 2, p.34



A classic

Diamonds Were A
Girl's Best Friend

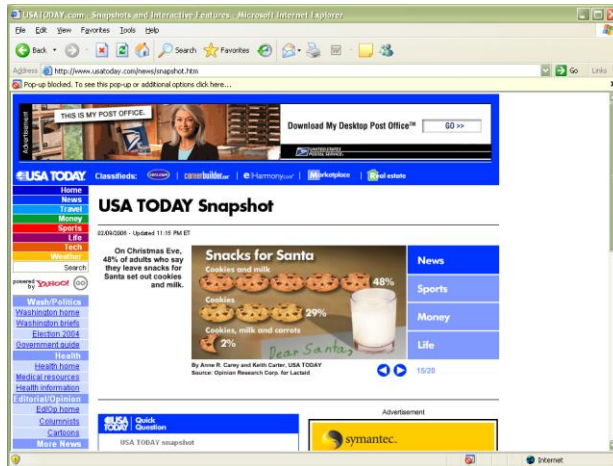
Fall 2015

CS 7450

54

USA Today

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



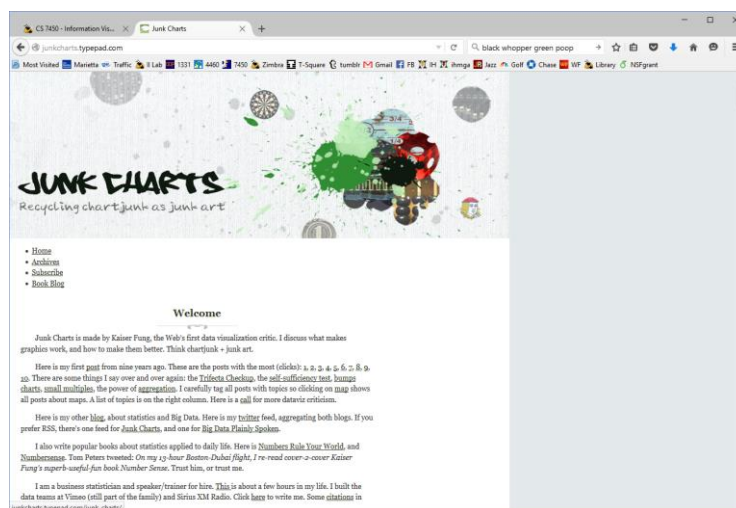
Fall 2015

CS 7450

55

Junk Charts Blog

<http://junkcharts.typepad.com/>



Fall 2015

CS 7450

56

More Thoughts



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

Fall 2015

CS 7450

57

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 Green, David McDias, 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.green@usask.ca, dmndr@usask.ca, cb95@usask.ca

ABSTRACT
 Guidelines for designing information charts often state that the presenter should include their goal – visual embellishments that are not essential to understanding the data. In contrast, some propose chart designers strip the presented data to denoted and enhance memory, citing the existence of shallow but deeper memory as determined by 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-four-week gap was significantly better. Although we are cautious about recommending that all charts be presented in this style, our results suggest some of the promises of the minimalist approach to chart design.

Author keywords:
 Chart, information visualization, imagery, memorability.

ACM Classification Keywords:
 H.3.0. Information systems and presentation (e.g., HCI); I.2.10. Visualization.

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 graphs; their guidelines for good chart design often suggest that the addition of chart junk obscures and other kinds of non-essential imagery, in a chart can make comprehension more difficult and slow down readers from the data [22]. This minimalist perspective advocates plain and simple charts that maximize the proportion of data ink – or the ink to the chart used to represent data.

Researchers in media design in broad range of all or part of the work the period of computer use in general within the period has been to see how the inclusion of graphics in general affects the perceived time taken to read the data and the perceived time taken to read the data. In some cases, the inclusion of graphics has been shown to be helpful in some cases, but in other cases, it has been shown to be harmful. In some cases, the inclusion of graphics has been shown to be helpful in some cases, but in other cases, it has been shown to be harmful. In some cases, the inclusion of graphics has been shown to be helpful in some cases, but in other cases, it has been shown to be harmful.

Designs that minimize problems, many designers include a wide variety of visual embellishments in their charts, from visual decorations to large images and visual backgrounds. Our work focuses on the presence of visual embellishments in charts in the popular, online image gallery, whose work supports comparisons among visual imagery and the focus of the chart (Fig. 1, Figure 1).

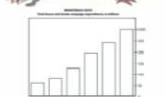


Figure 1. A chart by Bateman [2] (above), and a plain version. These kinds of charts appear regularly in many mass-media publications, and the widespread use of embellished designs raises questions about whether the minimalist position on chart design is really the better approach. This figure in particular are varied, 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 Tufte-style 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 2015

CS 7450

58

Memorability



Fall 2015

CS 7450

59

Mechanical Turk study

- Color and human recognizable objects enhance memorability
- Common graphs are less memorable than unique visualization types

But memorability is only one dimension of utility ... and perhaps not one of the most important ones

Design Principles



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

Fall 2015

CS 7450

60

Vol 1, p 140



Fall 2015

CS 7450

61

Vol 1, p. 141



US Army Divisions
going to France in
WW I

Leonard P. Ayres
The War with Germany
1919

Fall 2015

CS 7450

62

Vol 2, p. 36



Michel E. Turgot
Louis Bretz

Plan de Paris
1739

Fall 2015

CS 7450

63

Vol 2, p. 37

Manhattan 1989
Manhattan Map Company



Fall 2015

CS 7450

64

Vol 2, p. 42



Viet Nam Memorial
in Washington D.C.

Maya Ying Lin

58,000+ dead soldiers

Fall 2015

CS 7450

65

Vol 2, p. 44



Fall 2015

CS 7450

66

Vol 2, p. 43



Names listed
chronologically by death

Fall 2015

CS 7450

67

Design Principles



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

Fall 2015

CS 7450

68

Vol 1, p. 170



23 hours of
LA air pollution

Fall 2015

CS 7450

69

Vol 1, p. 173



Chromosomes of
man, chimpanzee,
gorilla & orangutan

Fall 2015

CS 7450

70

Vol 1, p. 174



Consumer
Reports

Fall 2015

CS 7450

71

Vol 2, p. 68



NY Trains

Fall 2015

CS 7450

72

Vol 2, p. 68



How to draw letters

Fall 2015

CS 7450

73

Vol 2, p. 69



Calligraphy

Fall 2015

CS 7450

74

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 2015

CS 7450

75

Sparkline Examples



Fall 2015

CS 7450

76

Design Principles



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

Space shuttle case we discussed first day

Fall 2015

CS 7450

77

Vol 3, p. 144



Washington Post

Fall 2015

CS 7450

78

Design Principles



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

Fall 2015

CS 7450

79

Vol 2, p. 12



Guide for visitors to Ise Shrine, Japan

Fall 2015

CS 7450

80

Vol 2, p. 24



Timetable for Java railroad line

Fall 2015

CS 7450

81

Vol 3, p. 90

Music history



Steve Chapple and Reebee Garofalo

Fall 2015

CS 7450

82

Design Principles



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

Fall 2015

CS 7450

83

Vol 2, p. 54

IBM Series III Copier



Fall 2015

CS 7450

84

Vol 2, p. 61



Fall 2015

CS 7450

85

Design Principles



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

Fall 2015

CS 7450

86

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



Life of a beetle

L. Hugh Newman

Fall 2015

CS 7450

87

Vol 2, p. 102



Czech air schedule

Fall 2015

CS 7450

88

Vol 2, p. 103



China railway
timetable

Fall 2015

CS 7450

89

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 2015

CS 7450

90

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 2015

CS 7450

91

Vol 2, p. 54

New Jersey Transit



Before

After

Fall 2015

CS 7450

92

Vol 2, p. 63



Before

After

Fall 2015

CS 7450

93

Vol 3, p. 74



Before

After

Fall 2015

CS 7450

94

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...

Fall 2015

CS 7450

97

Vol 1, p. 153



Fall 2015

CS 7450

98

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 2015

CS 7450

99

Vol 2, p. 82



Fall 2015

CS 7450

100

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 2015

CS 7450

101

Vol 3, p. 77



Fall 2015

CS 7450

102

Examples



- The good...

Fall 2015

CS 7450

103

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



Fall 2015

CS 7450

104

Vol 2, p. 80



Swiss Mountain Map

Fall 2015

CS 7450

105

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 2015

CS 7450

106

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 2015

CS 7450

107

Minard graphic



size of army
direction

latitude
longitude

temperature
date

Fall 2015

CS 7450

108

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 2015

CS 7450

109

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. SHIPPING 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." wirepo

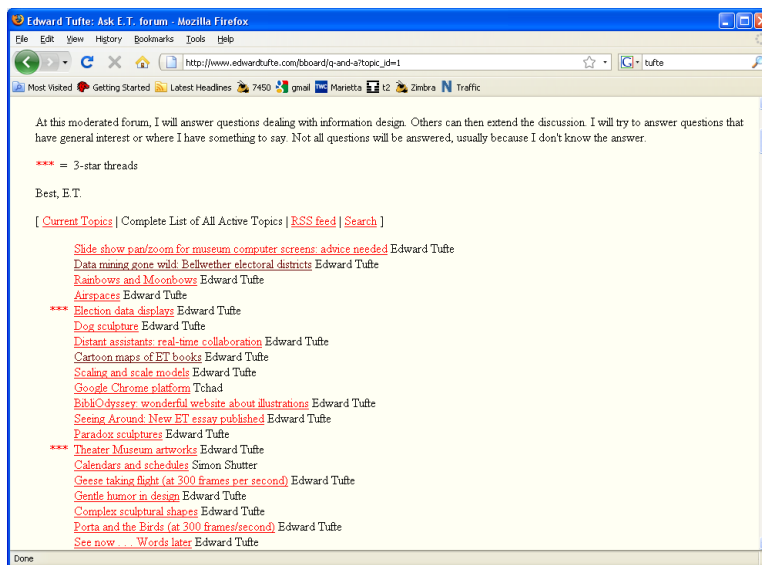
"The Leonardo da Vinci of data" THE NEW YORK TIMES

Fall 2015

CS 7450

110

Discussion Forum



Fall 2015

CS 7450

111

Interesting Contrast



Nigel Holmes

<http://www.nigelholmes.com>

Fall 2015

CS 7450

112

HW 5



- Design problem
- Text and document collection
 - Amazon TV reviews
 - Narrative text paired with some other attributes
- Due on Monday 19th
- Bring two copies

Fall 2015

CS 7450

113

Project Design Document



- Expand on problem, data, and tasks
- Show different design ideas
- Which one(s) are you choosing?
 - Explain

Fall 2015

CS 7450

114

Upcoming



- Text & documents 1
 - Reading
 - Viegas & Wattenberg, '08

- Text & documents 2
 - Reading
 - Keim & Oelke, '08

Fall 2015

CS 7450

115

Sources Used



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

Fall 2015

CS 7450

116