

Tufte's Design Principles



CS 4460 – Intro. to Information Visualization
October 23, 2017
John Stasko

Please see appropriate
books for missing images

Learning Objectives



- Understand and be able to apply Tufte's principles:
 - Graphical integrity (baselines, size coding)
 - Maximize data-ink ratio
 - Avoid chartjunk
 - Macro/micro-readings
 - Small multiples
 - Minimize/unite grids, labeling, legends
 - Appropriate applications of color

Today's Agenda

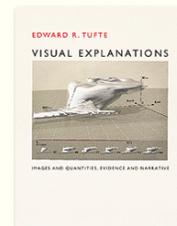
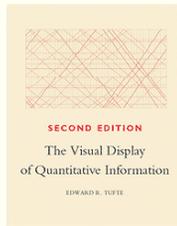


Edward Tufte has written seven books, including *Visual Explanations*, *Envisioning Information*, *The Visual Display of Quantitative Information*, and *Data Analysis for Politics and Policy*. He writes, designs, and self-publishes his books on analytical design, which have received more than 40 awards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

For a moderated forum on analytical design, go to [ASK E.T.](http://www.edwardtufte.com/ask-et)

BOOKS



Fall 2017

CS 4460

3

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 2017

CS 4460

4

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 2017

CS 4460

5

Summary



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

Let's look at each of these

Fall 2017

CS 4460

6

1. Graphical Integrity



- Your graphic should tell the truth about your data

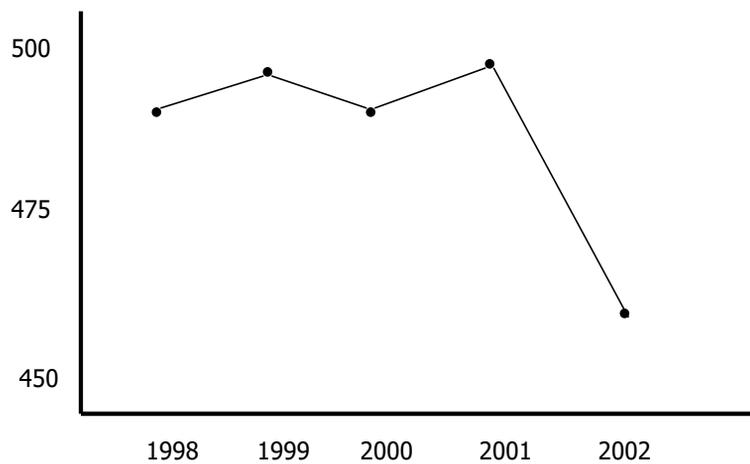
Fall 2017

CS 4460

7

Example

Stock market crash?

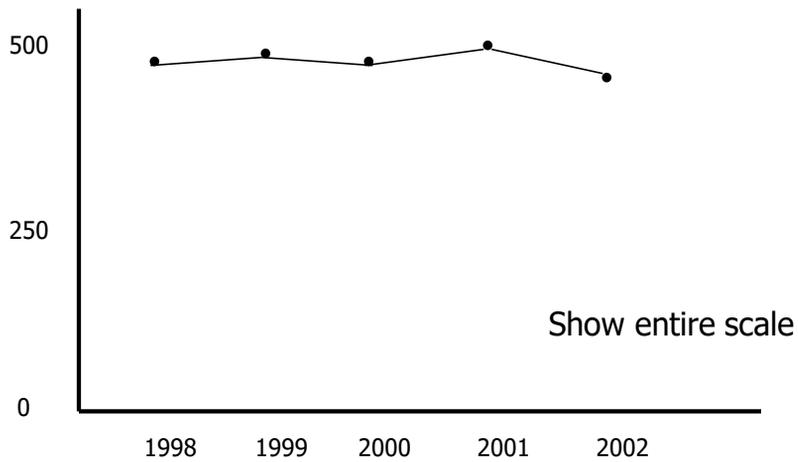


Fall 2017

CS 4460

8

Example

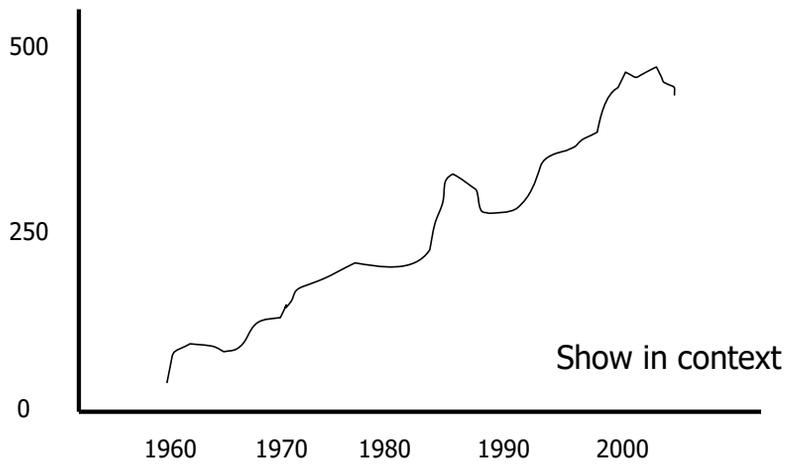


Fall 2017

CS 4460

9

Example



Fall 2017

CS 4460

10

Chart Integrity



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

Fall 2017

CS 4460

11

Vol 1, p. 54 (1)

Where's 0?
Note middle '70

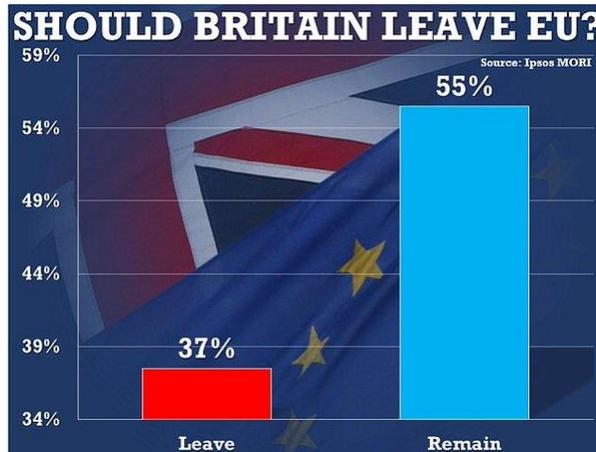


Fall 2017

CS 4460

12

Huge Difference?



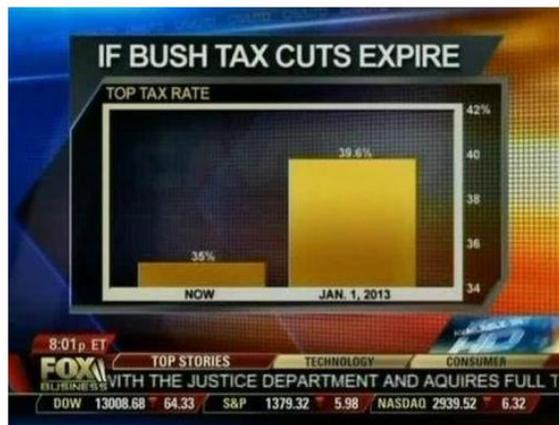
Compare area of right bar to that of the left bar

Fall 2017

CS 4460

13

Huge Difference?



Compare area of right bar to that of the left bar

Fall 2017

CS 4460

14

Vol 1, p 54 (2)



What's being compared?

Fall 2017

CS 4460

15

Vol 1, 57



Scale?

Fall 2017

CS 4460

16

Vol 1, p. 74



Great work!

Fall 2017

CS 4460

17

Vol 1, p. 74



Ahhhh

Show the context

Fall 2017

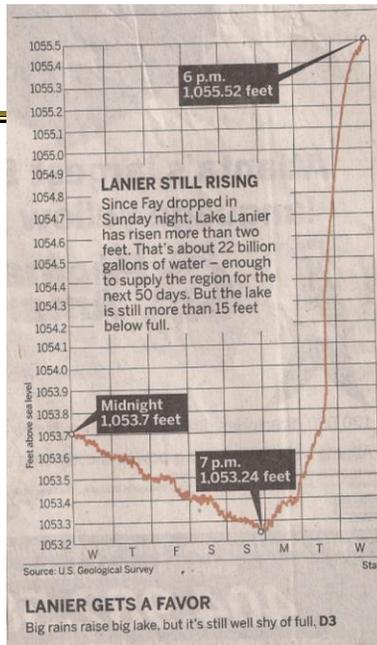
CS 4460

18

Local Example

A huge rise?

Atlanta Journal Constitution
Summer '08



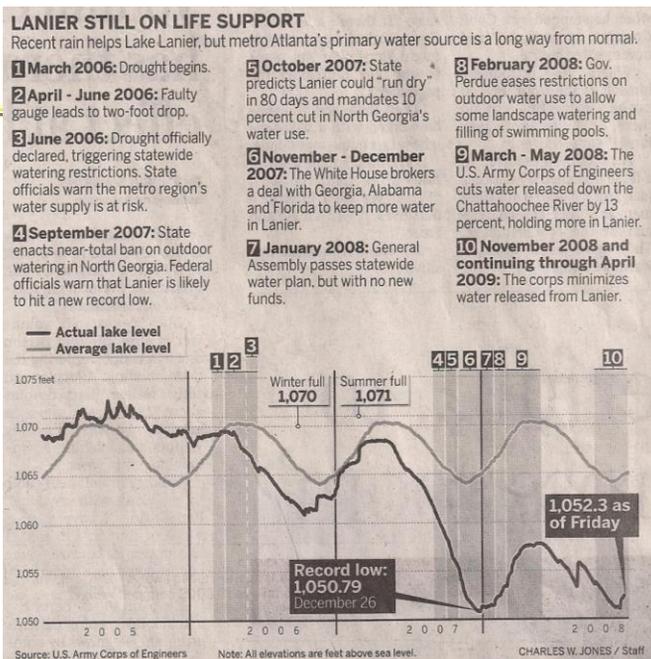
Fall 2017

CS 4460

19

More of the data

Atlanta Journal Constitution
Dec. '08



Fall 2017

CS 4460

20

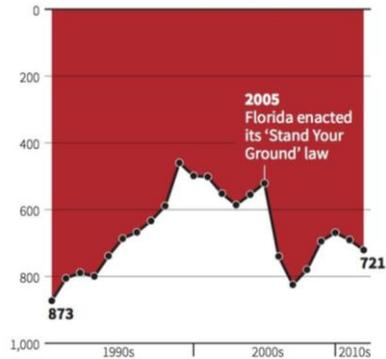
<http://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2>

A Redesign



Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C.Chan 16/02/2014

REUTERS

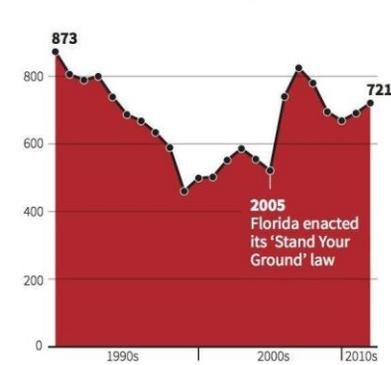
Reuters

Fall 2017

CS 4460

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

P.A. Fildoux and Reuters

21

Watch Size Coding



- Height/width vs. area vs. volume

Fall 2017

CS 4460

22

Vol 1, p. 69



area = value?

Fall 2017

CS 4460

23

Vol 1, p. 62



volume = value?

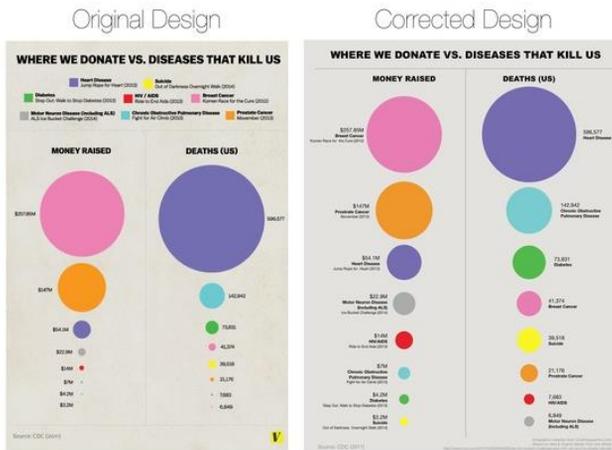
Fall 2017

CS 4460

24

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

Circle Width vs. Area



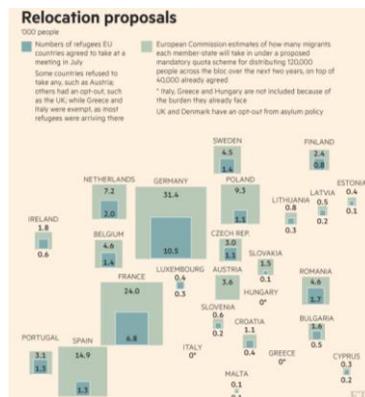
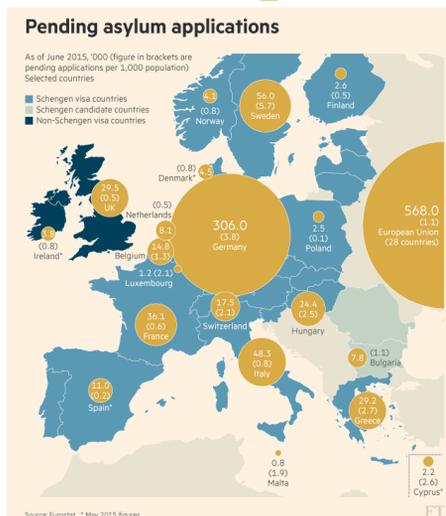
Fall 2017

CS 4460

25

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

Areas? Not Sure



Fall 2017

CS 4460

26

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 = $\frac{4280}{454}$
oil barrels

Fall 2017

CS 4460

27

2. Design Aesthetics



- Set of principles to help guide designers

Fall 2017

CS 4460

28

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 2017

CS 4460

29

Vol 1, p. 94



Good

Bad

Fall 2017

CS 4460

30

Vol 1, p. 30



Outstanding

Fall 2017

CS 4460

31

More...



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

Fall 2017

CS 4460

32

More...



- Maximize data density

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

Quote ...

Fall 2017

CS 4460

33

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 2017

CS 4460

34

Redesign charts



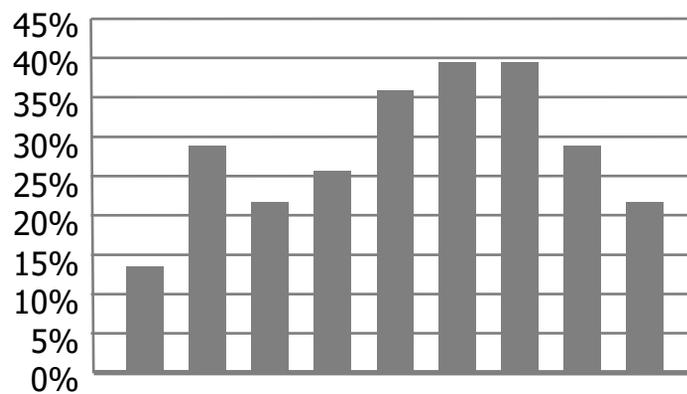
- Bar chart, scatter plot, box plot

Fall 2017

CS 4460

35

Bar chart

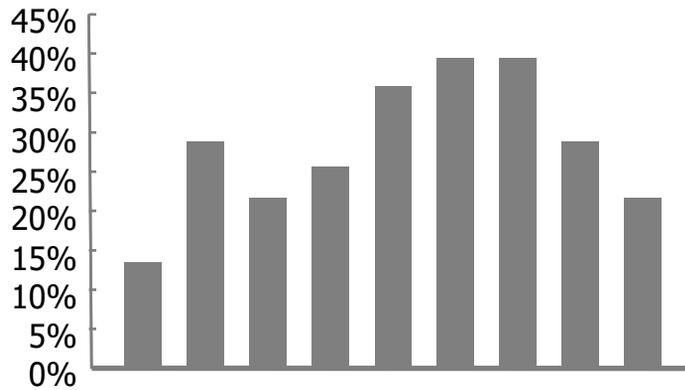


Fall 2017

CS 4460

36

Bar chart

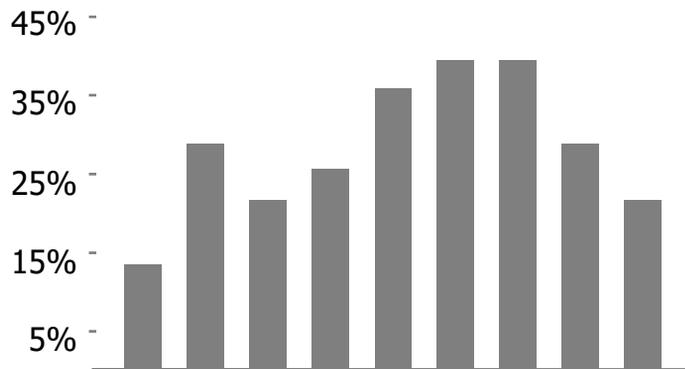


Fall 2017

CS 4460

37

Bar chart

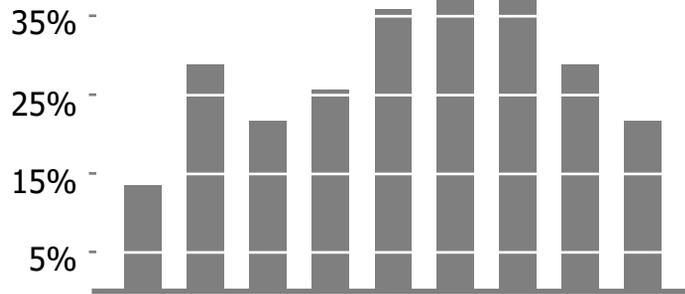


Fall 2017

CS 4460

38

Bar chart

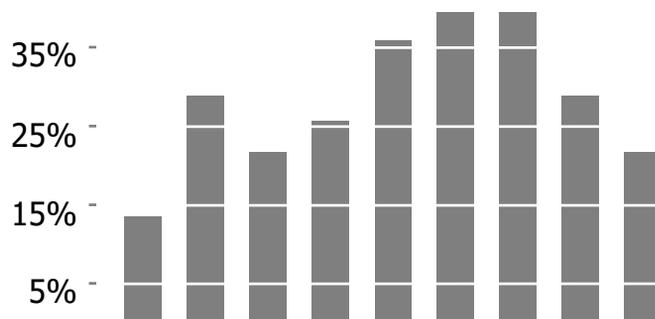


Fall 2017

CS 4460

39

Bar chart

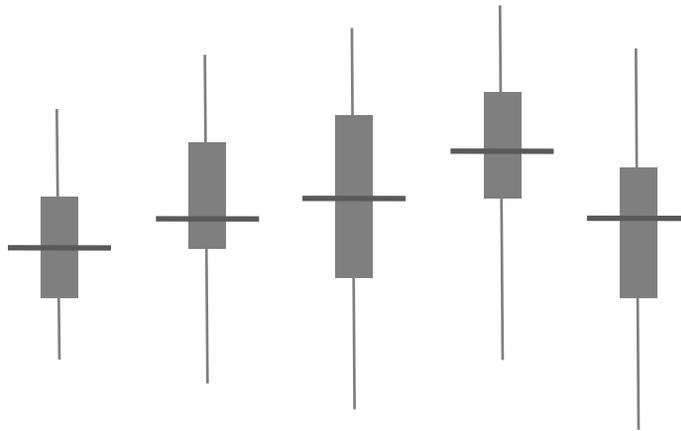


Fall 2017

CS 4460

40

Box plot

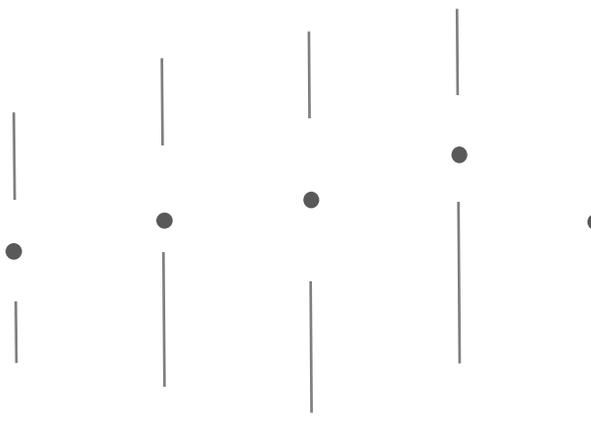


Fall 2017

CS 4460

41

Box plot

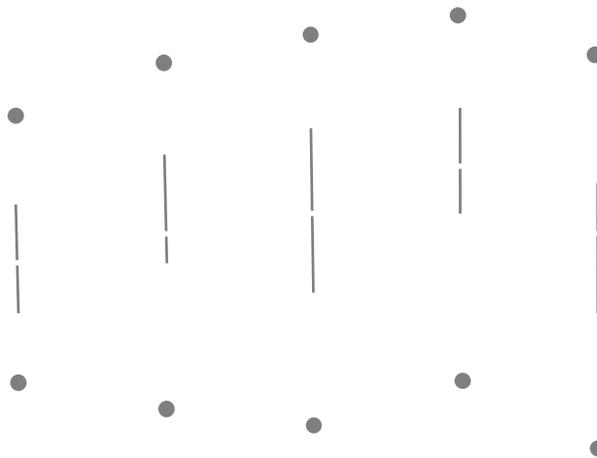


Fall 2017

CS 4460

42

Box plot

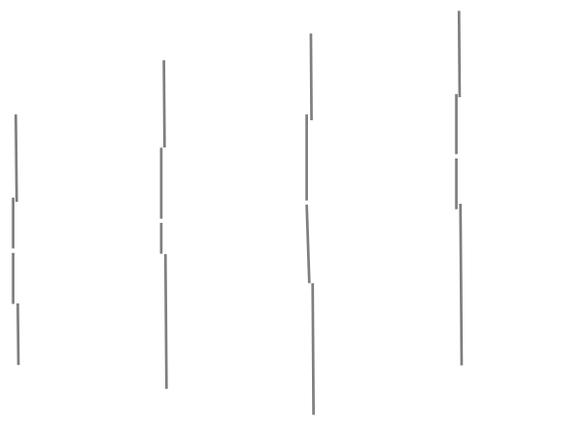


Fall 2017

CS 4460

43

Box plot

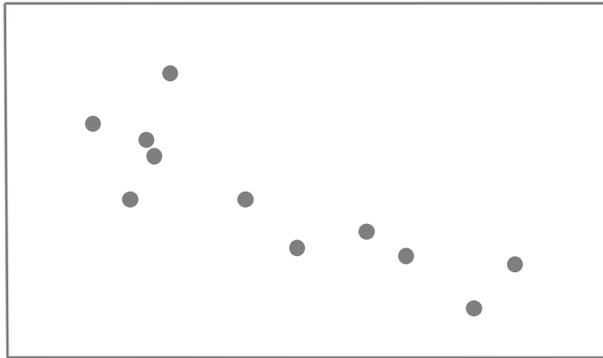


Fall 2017

CS 4460

44

Scatter plot

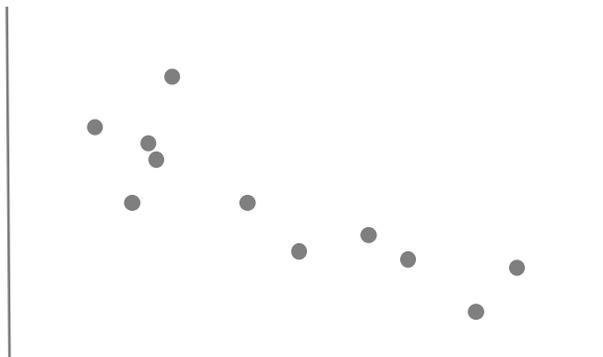


Fall 2017

CS 4460

45

Scatter plot

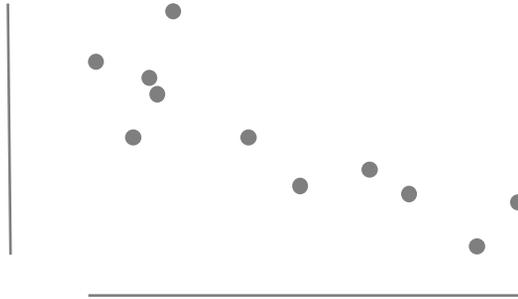


Fall 2017

CS 4460

46

Scatter plot

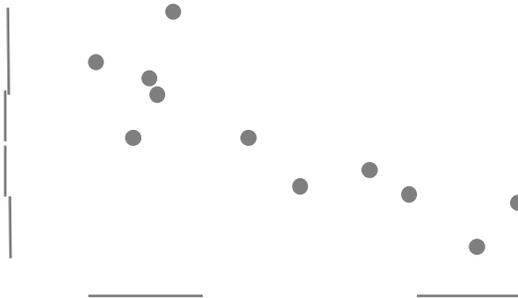


Fall 2017

CS 4460

47

Scatter plot



Fall 2017

CS 4460

48

Design Principles



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

Fall 2017

CS 4460

49

Vol 1, p 108



Fall 2017

CS 4460

50

Vol 2, p.34



A classic

Diamonds Were A
Girl's Best Friend

Fall 2017

CS 4460

51

USA Today

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



The screenshot shows the USA Today Snapshot page. The browser window title is "USA TODAY.com - Snapshots and Interactive Features - Microsoft Internet Explorer". The address bar shows "http://www.usatoday.com/news/snapshot.htm". A "Pop-up blocked" message is visible at the top. The page content includes a "Download My Desktop Post Office" button, a navigation menu with categories like Home, News, Travel, Money, Sports, Life, Tech, and Weather, and a main article titled "Snacks for Santa" with a pie chart showing 48% for Cookies and milk, 29% for Cookies, milk and corns, and 2% for Just Santa. The article is by Anne R. Carey and Adam Carter, dated 12/20/08. A sidebar on the right lists News, Sports, Money, and Life. At the bottom, there is a Symantec advertisement.

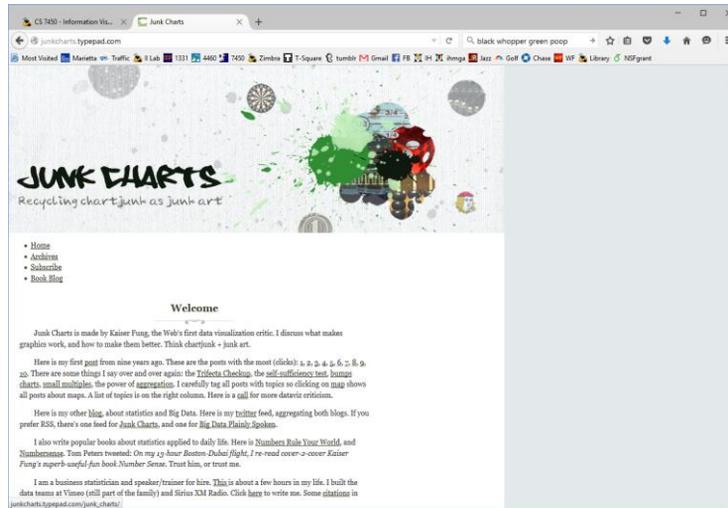
Fall 2017

CS 4460

52

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

Junk Charts Blog



Fall 2017

CS 4460

53

More Thoughts



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

Fall 2017

CS 4460

54

Design Principles



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

Fall 2017

CS 4460

55

Vol 1, p 140



Fall 2017

CS 4460

56

Vol 1, p. 141



US Army Divisions
going to France in
WW I

Leonard P. Ayres
The War with Germany
1919

Fall 2017

CS 4460

57

Vol 2, p. 37

Manhattan 1989
Manhattan Map Company



Fall 2017

CS 4460

58

Vol 2, p. 42



Viet Nam Memorial
in Washington D.C.

Maya Ying Lin

58,000+ dead soldiers

Fall 2017

CS 4460

59

Vol 2, p. 44



Fall 2017

CS 4460

60

Vol 2, p. 43



Names listed
chronologically by death

Fall 2017

CS 4460

61

Design Principles



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

Fall 2017

CS 4460

62

Vol 1, p. 170



23 hours of
LA air pollution

Fall 2017

CS 4460

63

Vol 1, p. 173



Chromosomes of
man, chimpanzee,
gorilla & orangutan

Fall 2017

CS 4460

64

Vol 1, p. 174



Consumer
Reports

Fall 2017

CS 4460

65

Vol 2, p. 68



NY Trains

Fall 2017

CS 4460

66

Vol 2, p. 68



How to draw letters

Fall 2017

CS 4460

67

Vol 2, p. 69



Calligraphy

Fall 2017

CS 4460

68

More Recent Additions



Sparklines

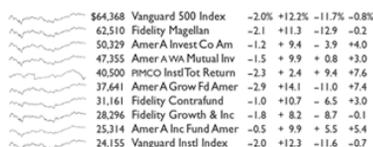
Small, repeated graphics (frequently line graphs)

Fall 2017

CS 4460

69

Sparkline Examples



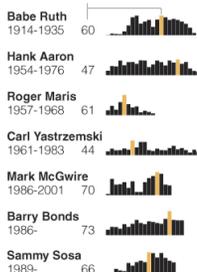
Spikes or trends?

Home-run careers of some heavy hitters. Some players consistently had high home-run years; others had occasionally stellar years. Their best home-run performances are in parentheses.



Source: Baseball Encyclopedia, Baseball Almanac

TOP SEASON



Fall 2017

CS 4460

70

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 2017

CS 4460

71

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 2017

CS 4460

72

Vol 2, p. 54

New Jersey Transit



Before

After

Fall 2017

CS 4460

73

Vol 2, p. 63



Before

After

Fall 2017

CS 4460

74

Vol 3, p. 74



Before

After

Fall 2017

CS 4460

75

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.*”

Fall 2017

CS 4460

76

Proper Color Use



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

Fall 2017

CS 4460

77

Examples



- The bad...

Fall 2017

CS 4460

78

Vol 1, p. 153



Fall 2017

CS 4460

79

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 2017

CS 4460

80

Vol 2, p. 82



Fall 2017

CS 4460

81

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 2017

CS 4460

82

Vol 3, p. 77



Fall 2017

CS 4460

83

Examples



- The good...

Fall 2017

CS 4460

84

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



Fall 2017

CS 4460

85

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 2017

CS 4460

86

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 2017

CS 4460

87

Minard graphic



size of army
direction

latitude
longitude

temperature
date

Fall 2017

CS 4460

88

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 2017

CS 4460

89

Website & Seminar



The screenshot shows a website titled "The Work of Edward Tufte and Graphics Press". The page layout includes a navigation bar with links like HOME, BOOKS, ONE-DAY COURSE, ET NOTEBOOKS, SCULPTURE, PRINTS, POSTERS, GRAPH PAPER, ABOUT ET, and CART. The main content area is divided into several sections:

- THE WORK OF EDWARD TUFT AND GRAPHICS PRESS:** A paragraph describing Edward Tufte as a statistician and artist at Yale University, mentioning his books on data visualization and his modern gallery studio.
- PRESENTING DATA AND INFORMATION:** A section for a one-day course taught by Edward Tufte, listing topics covered and a new, widely-adopted method for presentations.
- ART: ET NOTEBOOKS:** A list of various notebooks and books, including "Hoggen Hill Farms artwork", "Continuous select negative: structure of unknown significance", and "The Tufts Landscape artworks made from charged".
- SCIENCE: ET NOTEBOOKS:** A list of science-related notebooks and books, including "Sentences of the God", "Developing data graphics: graphics theory and practice", and "Artful Forensic Diagrams".

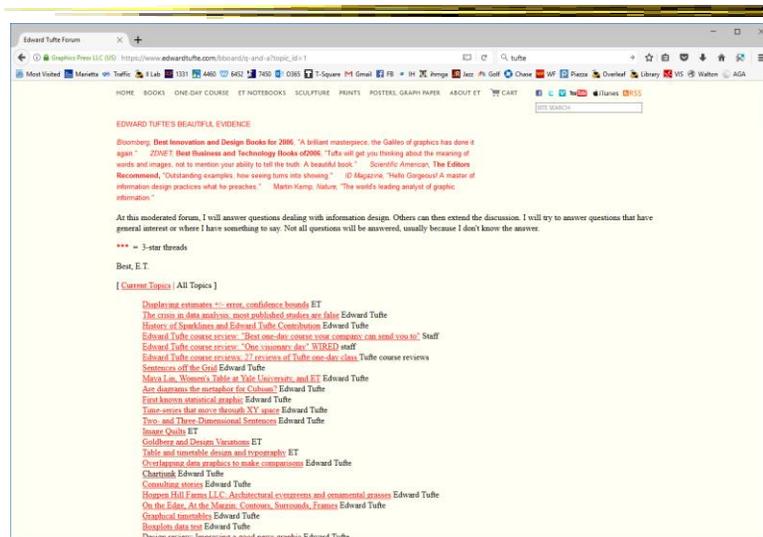
The website also features a video player for "INGE DRUCKREY: TEACHING TO SEE" and a search bar.

Fall 2017

CS 4460

90

Discussion Forum



Fall 2017

CS 4460

91

Interesting Contrast



Nigel Holmes

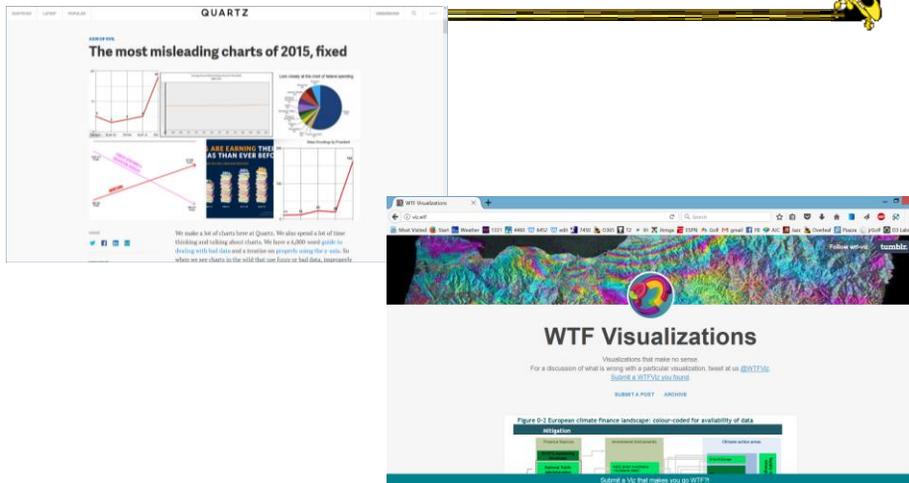
<http://www.nigelholmes.com>

Fall 2017

CS 4460

92

More Bad Examples



<http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/>
<http://viz.wtf/>

Fall 2017

CS 4460

93

Learning Objectives

- Understand and be able to apply Tufte's principles:
 - Graphical integrity (baselines, size coding)
 - Maximize data-ink ratio
 - Avoid chartjunk
 - Macro/micro-readings
 - Small multiples
 - Minimize/unite grids, labeling, legends
 - Appropriate applications of color

Fall 2017

CS 4460

94

Upcoming



- Storytelling/Communication
 - Prep: The Fallen of WWII video
- Lab: D3 Enter, Update, & Exit
 - Prep: Murray, chapter 9

Fall 2017

CS 4460

95

Sources Used



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

Fall 2017

CS 4460

96