Lecture 34 of 41

Visualization, Part 1 of 3: Data (Quantities & Evidence)

William H. Hsu
Department of Computing and Information Sciences, KSU

KSOL course pages: http://bit.ly/hGvXlH / http://bit.ly/eVizrE
Public mirror web site: http://www.kddresearch.org/Courses/CIS636
Instructor home page: http://www.cis.ksu.edu/~bhsu

Readings:
Last class: Chapter 15, Eberly 2e – see **http://bit.ly/ieUq45**; **Ray Tracing Handout**
Today: **Tufte Handout 1**
Next class: **Ray Tracing Handout**
Wikipedia, *Visualization*: **http://bit.ly/gVxRFp**
Wikipedia, *Data Visualization*: **http://bit.ly/9icAZk**

Lecture Outline

- **Reading for Last Class: Chapter 15, Eberly *2e*; Ray Tracing Handout**
- **Reading for Today: Tufte Handout**
- **Reading for Next Class: Ray Tracing Handout**
- **Last Time: Ray Tracing 2 of 2**
  - **Stochastic & distributed RT**
    - **Stochastic (local) *vs.* distributed (nonlocal) randomization**
    - **"Softening" shadows, reflection, transparency**
  - **Hybrid global illumination: RT with progressive refinement radiosity**
- **Today: Visualization Part 1 of 3 – Scientific, Data, Information Vis**
  - **What is visualization?**
  - **Tufte 1: *The Visual Display of Quantitative Information, 2 e***
    - **Basic statistical & scientific visualization techniques**
    - **Graphical integrity *vs.* lie factor ("How to lie with statisticsvis")**
    - **Graphical excellence vs. chartjunk**
    - **Data-ink, data-ink ratio (& "data-pixels")**

Where We Are

21	**Lab 4a: Animation Basics**	**Flash animation handout**
22	Animation 2: Rotations; Dynamics, Kinematics	Chapter 17, esp. §17.1 – 17.2
23	Demos 4: Modeling & Simulation; Rotations	Chapter 10[1], 13[2], §17.3 – 17.5
24	**Collisions 1: axes, OBBs, Lab 4b**	**§2.4.3, 8.1, GL handout**
25	Spatial Sorting: Binary Space Partitioning	Chapter 6, esp. §6.1
26	**Demos 5: More CGA; Picking; HW/Exam**	**Chapter 7[2]; § 8.4**
27	**Lab 5a: Interaction Handling**	**§ 8.3 – 8.4; 4.2, 5.0, 5.6, 9.1**
28	Collisions 2: Dynamic, Particle Systems	§ 9.1, particle system handout
	Exam 2 review; Hour Exam 2 (evening)	**Chapters 5 – 6, 7[2] – 8, 12, 17**
29	**Lab 5b: Particle Systems**	**Particle system handout**
30	**Animation 3: Control & IK**	**§ 5.3, CGA handout**
31	Ray Tracing 1: intersections, ray trees	Chapter 14
32	Lab 6a: Ray Tracing Basics with POV-Ray	RT handout
33	Ray Tracing 2: advanced topic survey	Chapter 15, **RT handout**
34	Visualization 1: Data (Quantities & Evidence)	Tufte handout (1)
35	**Lab 6b: More Ray Tracing**	**RT handout**
36	Visualization 2: Objects	Tufte handout (2 & 4)
37	Color Basics; Term Project Prep	**Color handout**
38	Lab 7: Fractals & Terrain Generation	**Fractals/Terrain handout**
39	**Visualization 3: Processes; Final Review 1**	Tufte handout (3)
40	**Project presentations 1; Final Review 2**	**–**
41	Project presentations 2	–
	Final Exam	**Ch. 1 – 8, 10 – 15, 17, 20**

Lightly-shaded entries denote the due date of a written problem set; heavily-shaded entries, that of a machine problem (programming assignment); blue-shaded entries, that of a paper review; and the green-shaded entry, that of the term project.

Green, **blue** and **red** letters denote **exam review**, **exam**, and **exam solution review** dates.

Acknowledgements: Statistical & Data Visualization

Edward R. Tufte

Professor Emeritus of Political Science, Statistics, & Computer Science
Yale University
http://bit.ly/gKhM0G
http://www.edwardtufte.com

Yale University

Leland Wilkinson

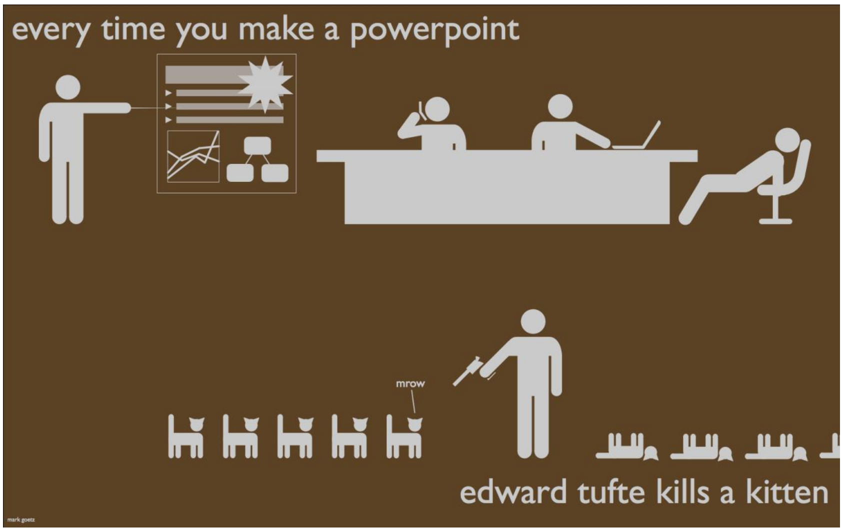
Executive Vice President, SYSTAT Software
Adjunct Professor of Computer Science, University of Illinois at Chicago
Adjunct Professor of Statistics, Northwestern University
http://www.cs.uic.edu/~wilkinson/

SYSTAT

UIC

5

Oops!

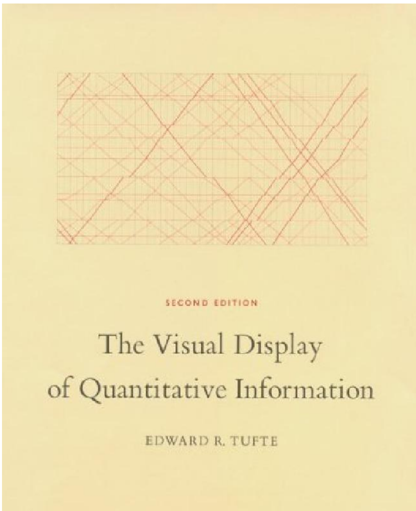


♥ 2009 M. A. Goetz

Mark Goetz: A blog of technology, usability, and sensemaking – http://bit.ly/cBMKR3

6

Tufte 1: The Visual Display of Quantitative Information, 2e



Tufte, E. R. (2001). *The Visual Display of Quantitative Information.* Cheshire, CT, USA: Graphics Press.

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

Apologia Pro PowerPoint Sua

- Tufte's Criticisms of Microsoft *PowerPoint* (Summarized in *Wikipedia*)
  - Used to guide, reassure presenters, rather than enlighten audience
  - Unhelpfully simplistic tables, charts (due in part to low-res displays)
  - Outliners may arrange ideas in necessarily deep hierarchy
    - Not visually retained: must be repeated on each slide
    - Artifact of "outline", "overview" format
  - Enforcement of linear progression
  - Poor design: typography, chart layout, use of templates, defaults
  - Simplistic thinking due to ideas being squashed into bulleted lists
    - Discontinuity of stories: beginning, middle, end → points
    - Cognitive load on reader: illusion of objectivity, neutrality
- Some (Though Not All) Problems with *PowerPoint* Avoidable by Design
- Garbage In, Garbage Out (Wikipedia: http://bit.ly/ff3EZ)
- Tufte, *The Cognitive Style of PowerPoint*: http://bit.ly/TYch
- Wikipedia Synopsis of Tufte's Critique: http://bit.ly/8XQFZm

Definition: Visualization

- *Wikipedia*: "Using Images, Diagrams, Animations to Communicate"
  - Images: illustrations; photographs, especially modified photos
  - Diagrams: structural diagrams, blueprints, plots & charts
  - Animations: based on simulation or other specifications
- Includes, But Not Limited to, Statistical Graphics
- Kinds of Visualization (Often Abbreviated "Vis" *cf.* IEEE InfoVis)
  - Scientific: transformation, representation of data for exploration
  - Data: schematic form
    - *e.g.*, relational database form (tuples of attribute values)
    - "Data vis" often synonymous with "statistical vis"
  - Information: spectrum from "raw data" to "info", "knowledge"
    - Premise: info more structured, organized, abstract than data
    - Emphasis on computational tools
    - Working with (especially analyzing) large data sets

Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Definition: Graphical Excellence

- Complex Ideas
- Communicated with
  - Clarity
  - Precision
  - Efficiency

Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Graphical Excellence: Graphical Displays Should… ("DO's")

- 1. Show the data
- 2. Induce the reader to think about the *substance* rather than about
  - Methodology
  - Graphic design
  - Technology of graphic production
  - Something else
- 3. Avoid distorting what the data have to say
- 4. Present many numbers in a small space
- 5. Make large data sets coherent
- 6. Encourage the eye to compare different pieces of data
- 7. Reveal the data at different levels of detail, broad to fine
- 8. Serve a clear purpose: description, evaluation, tabulation, decoration
- 9. Be closely integrated with statistical and verbal descriptions of data
- Rules were made to be broken ("What's different on this slide?")

Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

11

Graphics Reveal Data [1]: Limitations of Descriptive Statistics

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

→

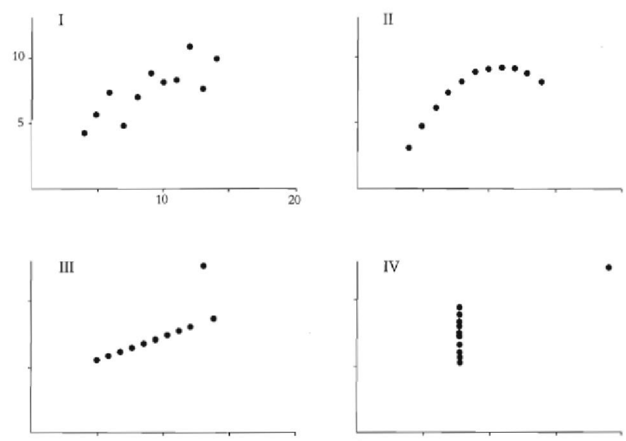
$N = 11$
mean of X's $= 9.0$
mean of Y's $= 7.5$
equation of regression line: $Y = 3 + 0.5X$
standard error of estimate of slope $= 0.118$
$t = 4.24$
sum of squares $X - \overline{X} = 110.0$
regression sum of squares $= 27.50$
residual sum of squares of $Y = 13.75$
correlation coefficient $= .82$
$r^2 = .67$

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

12

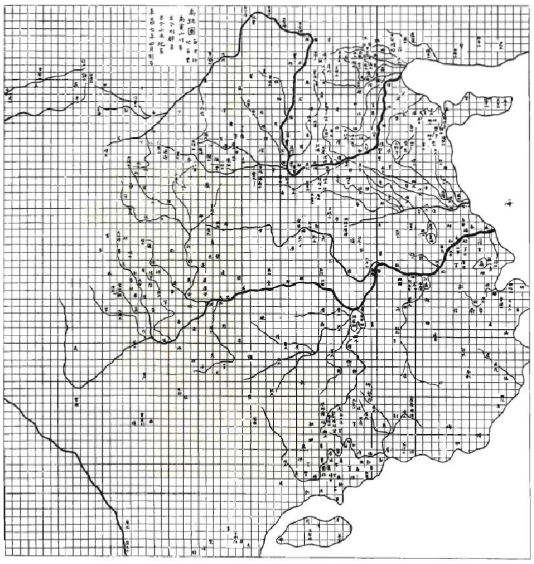
Graphics Reveal Data [2]: Differences in Data Shown by Vis

F. J. Anscombe, "Graphs in Statistical Analysis," *American Statistician*, 27 (February 1973), 17–21.



The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Earliest Grid Map: Song Dynasty, 960 – 1279 CE)



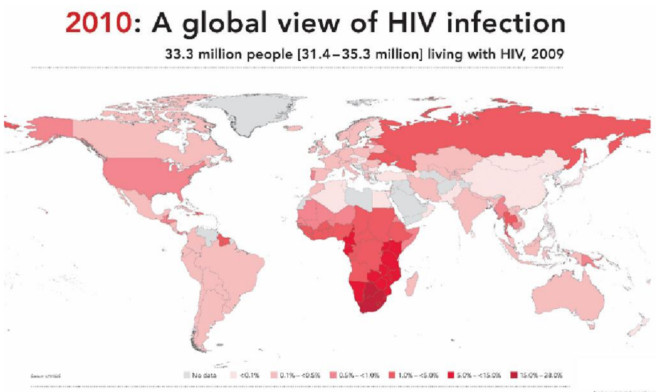
E. Chavannes, "Les Deux Plus Anciens Spécimens de la Cartographie Chinoise," *Bulletin de l'École Française de l'Extrême Orient*, 3 (1903), 1–35, Carte B.

This grid map, compiled c. 1100 CE (carved in stone c. 1137 CE) uses a grid of ~100 里 (*li*) to the square, ~42km in Han dynasty standard units (415.8m per *li*)

Shown:
Major rivers and tributaries

Thematic Maps & Other Data Maps

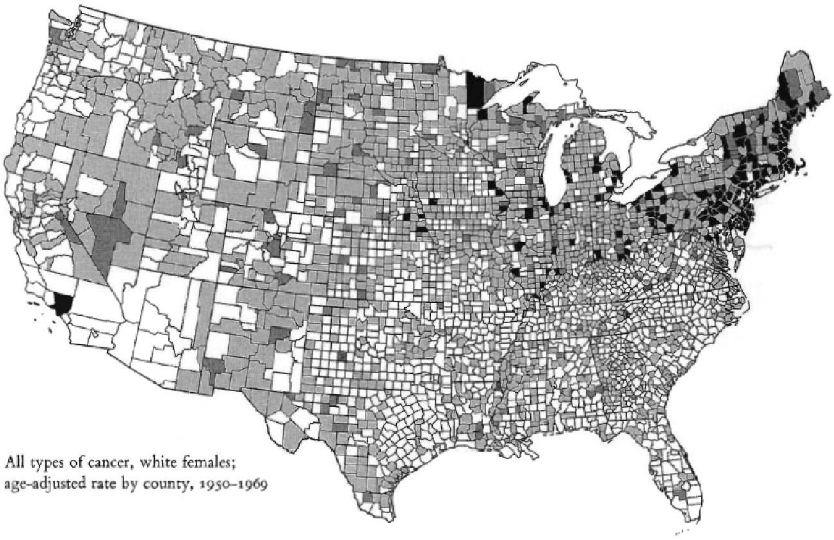
- **Data Maps: Visual Presentation of Variables over Region (*e.g.*, Spatial)**
- **Thematic Map: Shows Topic (Theme) Referenced by Geographic Area**
- **Wikipedia: http://en.wikipedia.org/wiki/Thematic_map**
- **Example: 2010 UNAIDS Report on Global HIV Infection Rates**



UNAIDS

World Health Organization

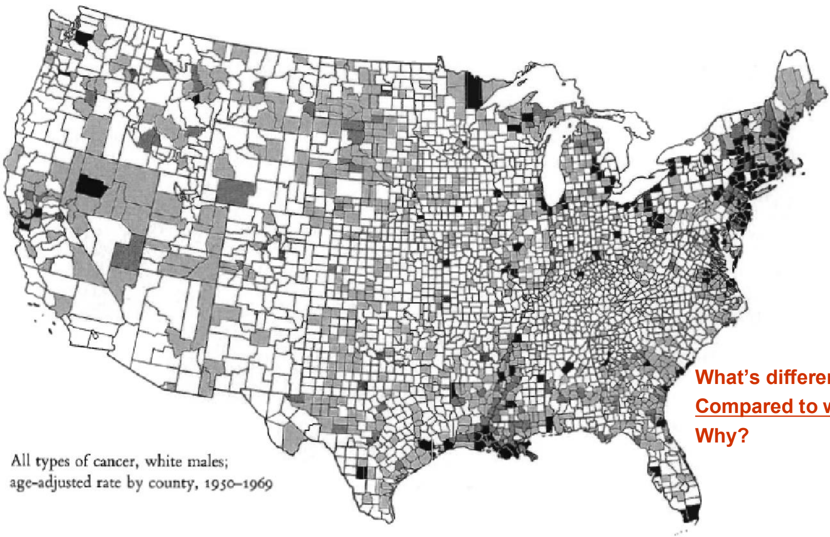
Example – Cancer Rates by County [1]: White Females, All Types of Cancer



All types of cancer, white females;
age-adjusted rate by county, 1950–1969

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Example – Cancer Rates by County [2]: White Males, All Types of Cancer

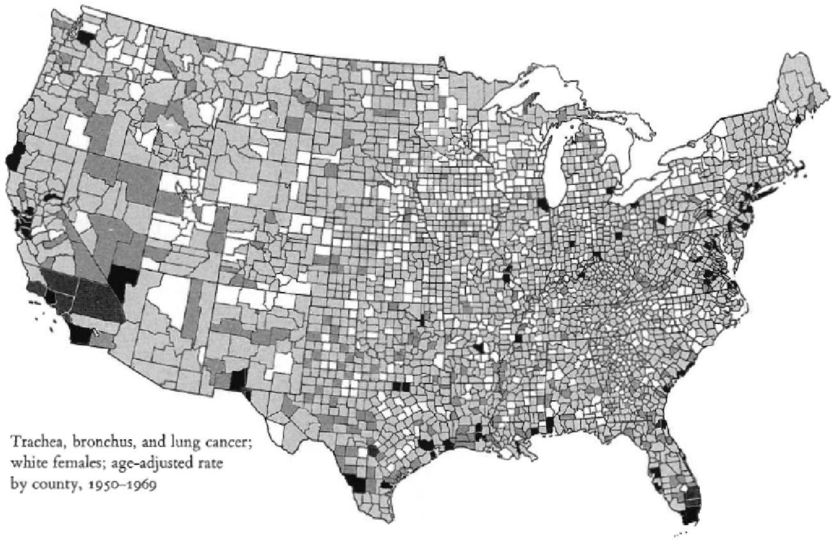


All types of cancer, white males;
age-adjusted rate by county, 1950–1969

What's different?
Compared to what?
Why?

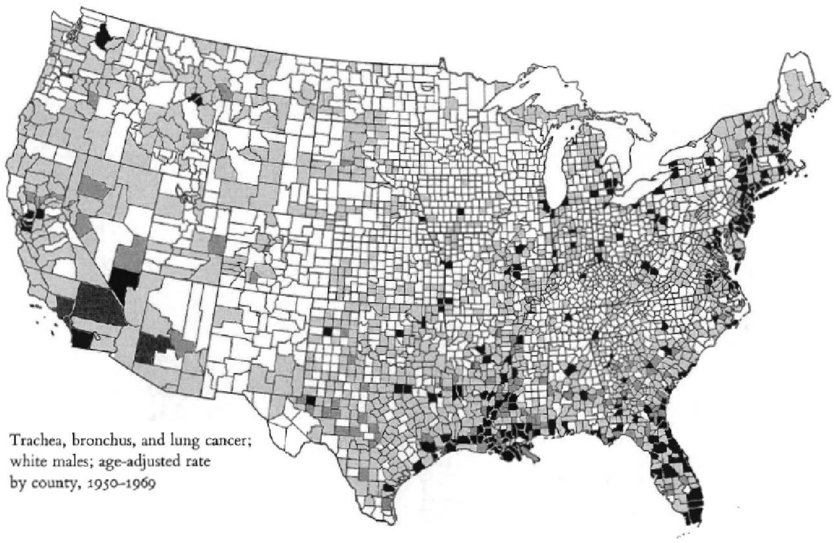
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Example – Cancer Rates by County [3]: White Females, Respiratory Cancers



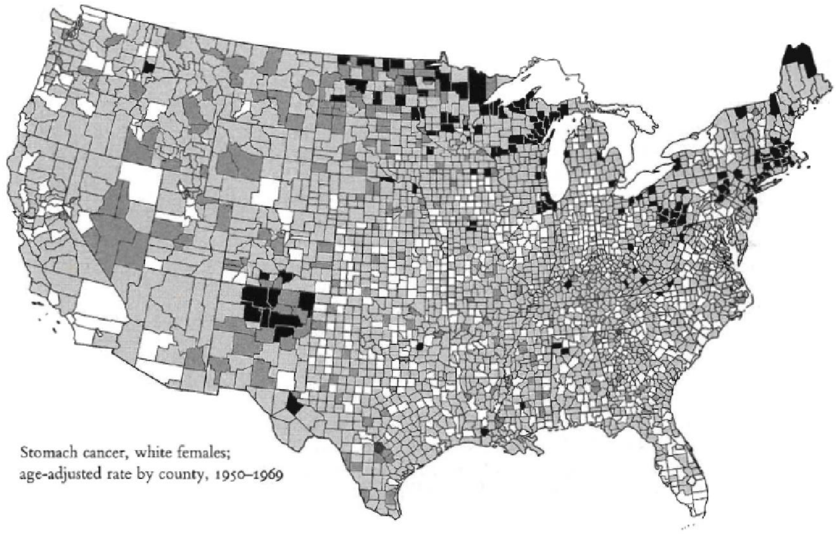
Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Example – Cancer Rates by County [4]: White Males, Respiratory Cancers



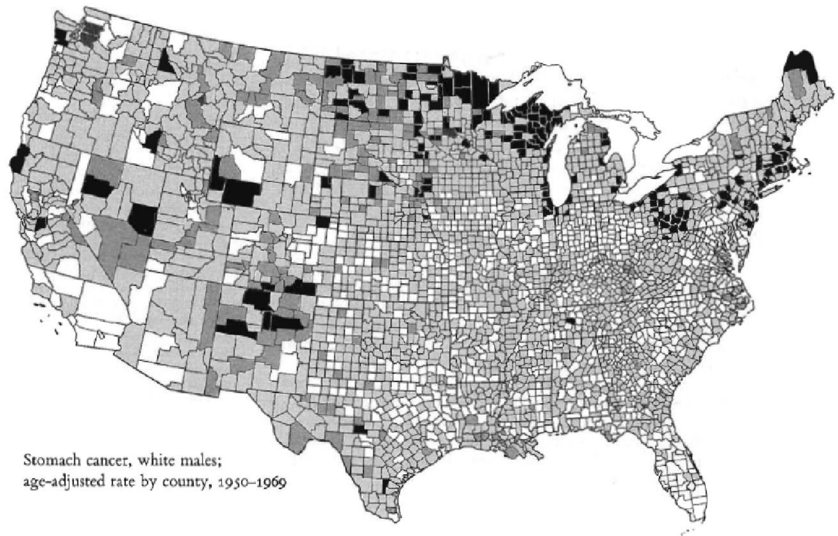
Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Example – Cancer Rates by County [5]: White Females, GI Cancers



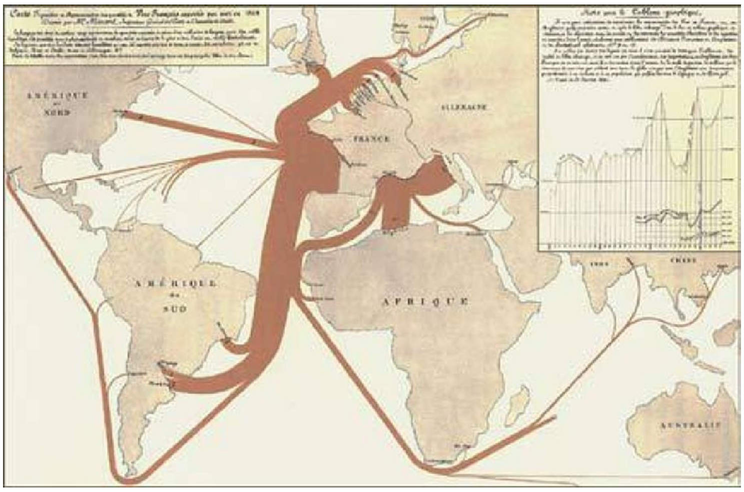
Stomach cancer, white females;
age-adjusted rate by county, 1950–1969

Example – Cancer Rates by County [6]: White Males, GI Cancers



Stomach cancer, white males;
age-adjusted rate by county, 1950–1969

Flow Maps [1]: Minard, 1865 – French Wine Exports



Minard, C. J. (1865). *Carte figurative et approximative des quantités de vin français exportés par mer en 1864.* Lithograph (835 x 547). Retrieved from: http://bit.ly/hTz4ST

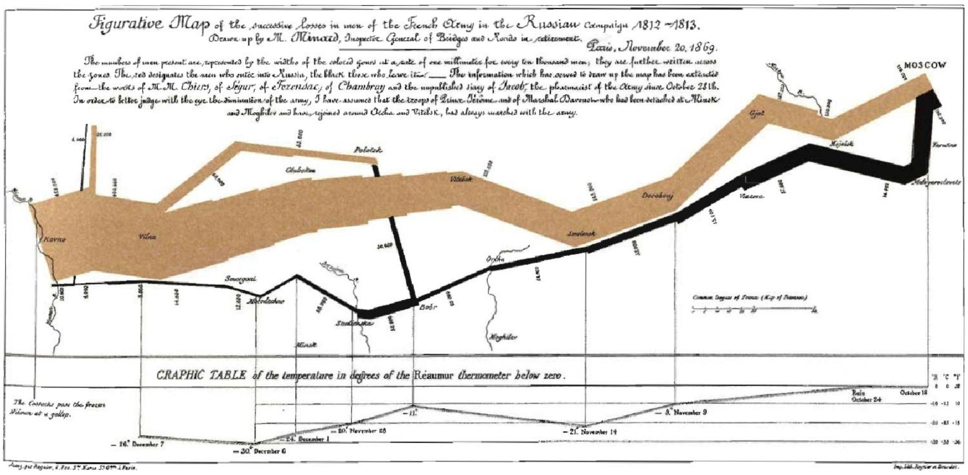
Wikipedia:
Flow Map – http://bit.ly/dJgJCY
Thematic Map – http://bit.ly/enon7w

Reprinted by National Visualization and Analysis Center (NVAC), Pacific Northwest National Lab (PNL): http://bit.ly/egbQUJ

Charles Joseph Minard, *Tableaux Graphiques et Cartes Figuratives de M. Minard, 1845–1869,* a portfolio of his work held by the Bibliothèque de l'École Nationale des Ponts et Chaussées, Paris.

Inspired by book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Flow Maps [2]: Minard, 1869 – Napoleon in Russia



Minard, C. J. (1869). *Carte figurative des pertes successives en hommes de l'Arm ée Française dans la campagne de Russie 1812-1813.* Lithograph, 62 x 30 cm.

Translation by E. R. Tufte. French version mirrored at: http://bit.ly/6MHa87
Wikipedia: *Charles Joseph Minard* – http://bit.ly/4zZ8nQ

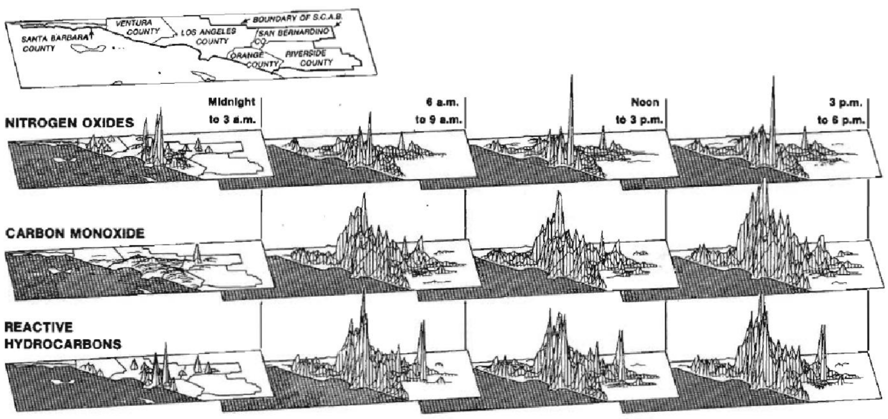
Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Flow Maps [3]: Minard's Map of Napoleon's March

- 2-D Map
- 6 Scalar Dimensions
  - Size of army
  - Location on 2-D surface by date (compare: latitude & longitude)
  - Direction of movement
  - Date
  - Temperature (referenced by position & date)
- How to Represent 6 Dimensions in 2
  - Size – width of line and written besides army (main camp) position
  - Location – (*x*, *y*) coordinate on map; align with timeline on bottom
  - Direction – color, arrow
  - Date – timeline on bottom
  - Temperature – next to date on timeline (today: brush-over tooltip)

Adapted from book material © 2001 E. R. Tufte, Yale University
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

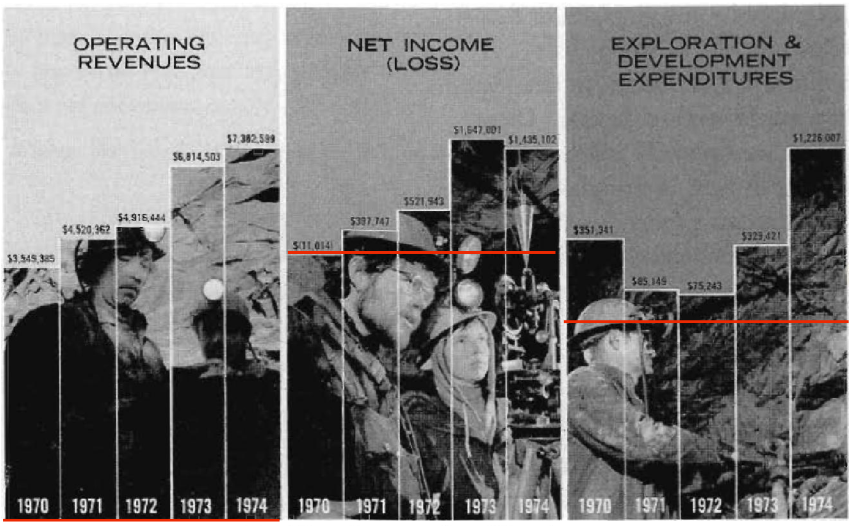
Preview: Small Multiple – Air Pollution Map



p. 42 (Tufte 1 *aka* VDQI, 1e & 2e)

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

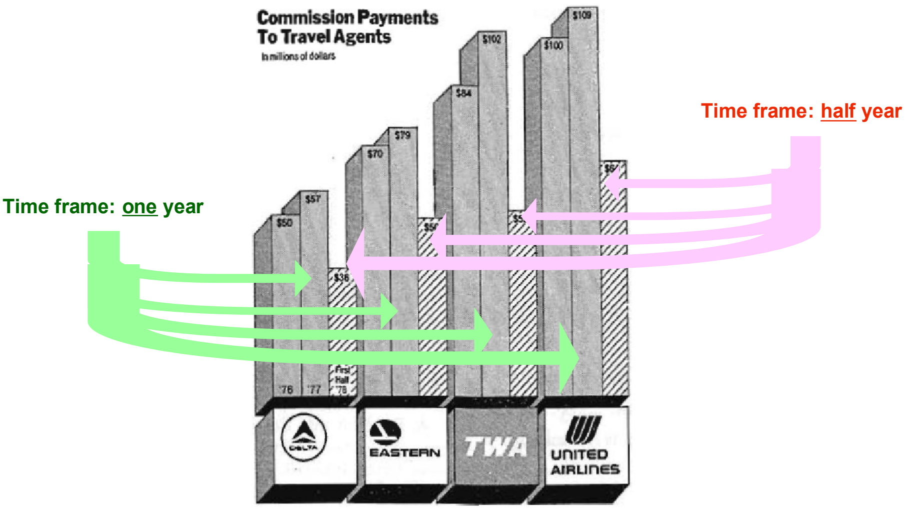
Graphical Integrity [1]: Consistency in Labeling, Baselines



Day Mines, Inc., *1974 Annual Report*, p. 1.

The Visual Display of Quantitative Information, **Graphics Press: http://bit.ly/16Se1**

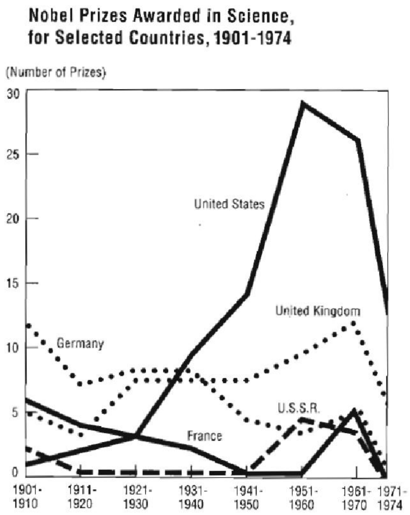
Graphical Integrity [2]: Consistency in Time (Independent Axis)

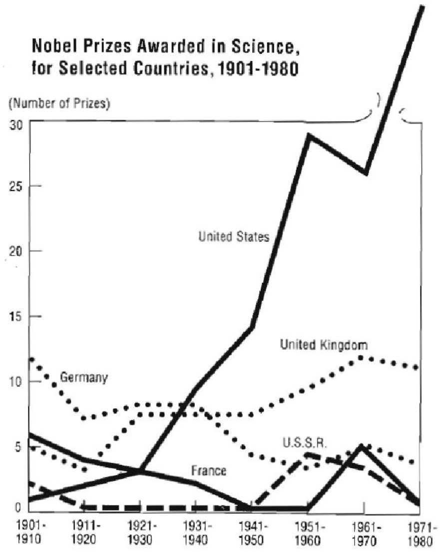


New York Times, August 8, 1978, p. D-1.

The Visual Display of Quantitative Information, **Graphics Press: http://bit.ly/16Se1**

Graphical Integrity [3]: Dangers of Partial Annual Data





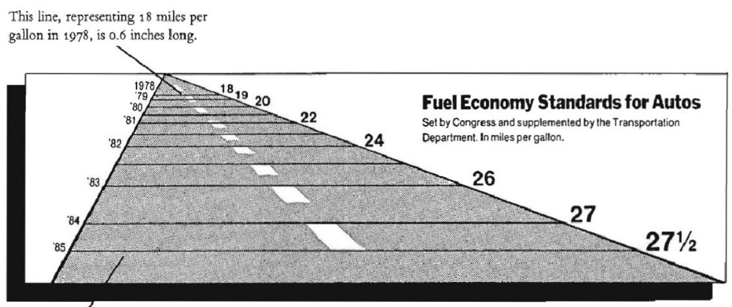
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Lie Factor [1]: Definition & Example

- **Definition**

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

- **Goal: Keep Close to 1 (*e.g.*, in interval [0.95, 1.05])**



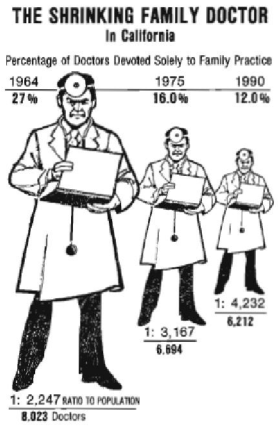
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

$$\text{Lie Factor} = \frac{5.3 - 0.6}{0.6} = \frac{783}{53} = 14.8$$

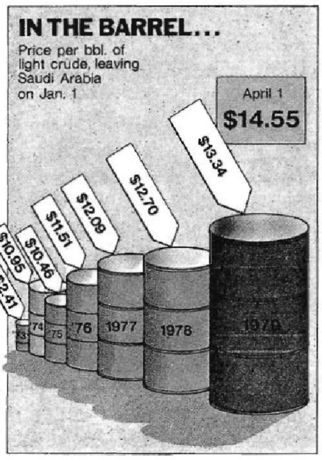
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

29

Lie Factor [2]: Examples – Excess Dimensions (2- & 3-D)



Los Angeles Times, August 5, 1979, p. 3.



Time, April 9, 1979, p. 57.

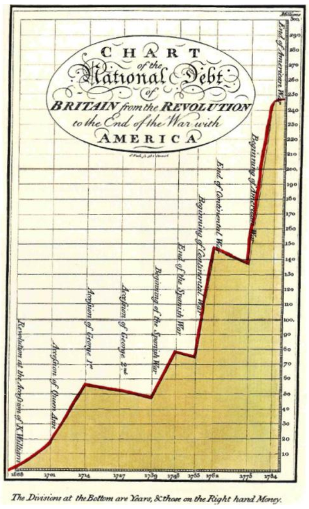
Lie Factor: ~1 based on height, 2.8 based on area

Lie Factor: 9.4 based on area, 59.4 based on volume!

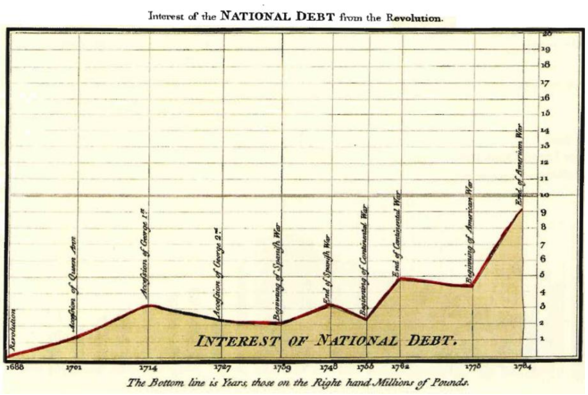
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

30

Graphical Integrity [4]: Need for Data Normalization



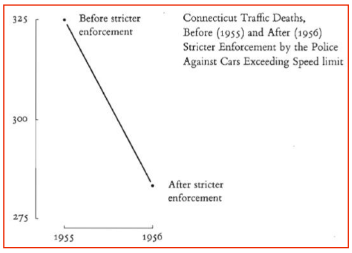
Playfair (1786): Nominal Money



Playfair (1786): Real (Standardized, *i.e.*, Inflation-Adjusted *aka* Deflated) Money

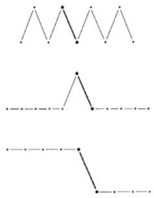
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Graphical Integrity [5]: Context – "Compared to What?"

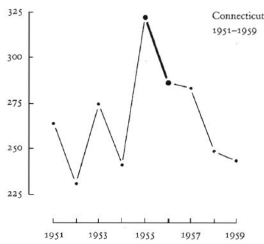


Imagine the very different interpretations other possible time-paths surrounding the 1955–1956 change would have:

What If…?



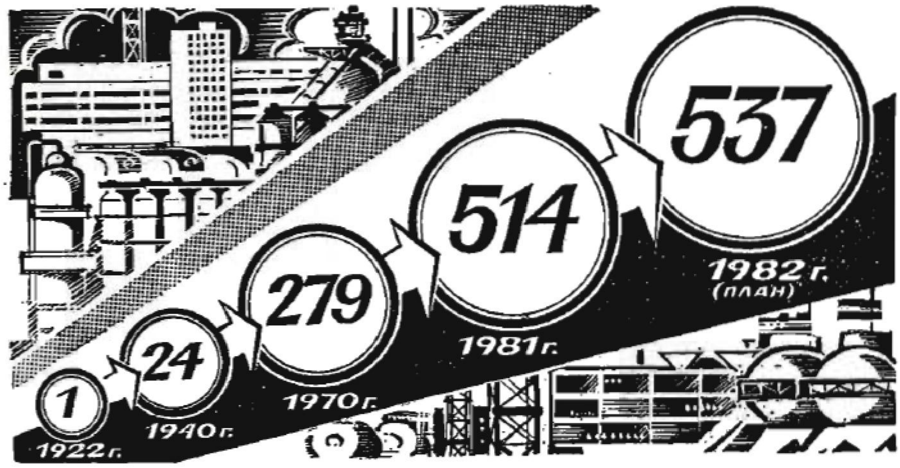
A few more data points add immensely to the account:



Donald T. Campbell and H. Laurence Ross, "The Connecticut Crackdown on Speeding: Time Series Data in Quasi-Experimental Analysis," in Edward R. Tufte, ed., *The Quantitative Analysis of Social Problems* (Reading, Mass., 1970), 110–125.

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Graphical Integrity [6]: Pravda School of Ordinal Graphics



Рост продукции промышленности [1922 г. = 1].

Pravda, May 24, 1982, p. 2.

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Graphical Integrity: Tufte's Six Principles ("DO's")

- See also *How to Lie With Statistics* (Huff, 1984): http://bit.ly/3wAgS0
- 1. Make Representation of Numbers Proportional to Quantities
  - Ratio of size to numerical value should be close to 1
  - As physically measured on surface of graphic
- 2. Use Clear, Detailed, Thorough Labeling
  - Don't introduce or propagate graphical distortion, ambiguity
  - Write out explanations of the data *on the graphic itself*
  - Label important events in the data
- 3. Show Data Variation, Not Design Variation
- 4. Use Standardized (*e.g.*, Inflation-Adjusted) Units, Not Nominal
- 5. Depict *N* Data Dimensions with ≤ *N* Variable Dimensions
  - Don't use more than *N* information-carrying dimensions for *N*-D data
  - When graphing data in *N*-D, use *N*-D ratio (see #1 above)
- 6. Quote Data in Full Context (Don't Quote Out of Context)

The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Data-Ink

Above all else show the data.
– Edward Tufte

- Data-Ink: Non-Erasable Core of A Graphic
  - Non-redundant ink
  - Arranged in response to variation in numbers represented
- Data Density
  - Amount of usable information per unit (space, ink, time, *etc.*)
  - Want: higher data density as function of resource
  - Example: data-ink ratio
- Non-Data-Ink Can Be Erased to Improve Data Density

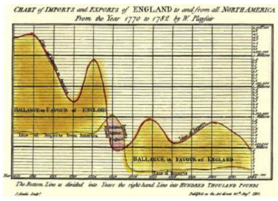
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Data-Ink Ratio

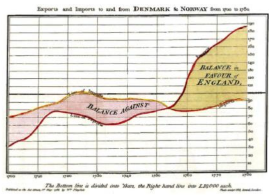
$$\text{Data-ink ratio} = \frac{\text{data-ink}}{\text{total ink used to print the graphic}}$$

$$= \text{proportion of a graphic's ink devoted to the non-redundant display of data-information}$$

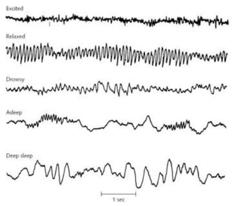
$$= 1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$$



Low Data-Ink Ratio (Playfair, 1785):
Charts with Many Grid Lines & Detailed Labels



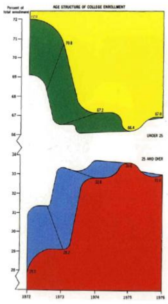
Intermediate Data-Ink Ratio (Playfair, 1786):
Conventional Charts – Grid Thinned



High Data-Ink Ratio:
Electroencephalogram (EEG) Signals

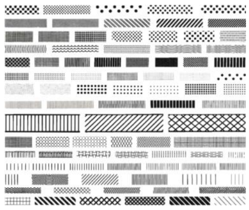
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Chartjunk

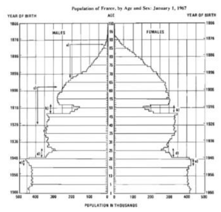


This chart uses five colors, three dimensions, and two parts to show only five data points!

Redundancy in
- Symmetry
- Superfluous use of perspective and color



Wikipedia, *Moiré pattern*: http://bit.ly/fssTem



© 1973 H. S. Shyrock & J. S. Siegel

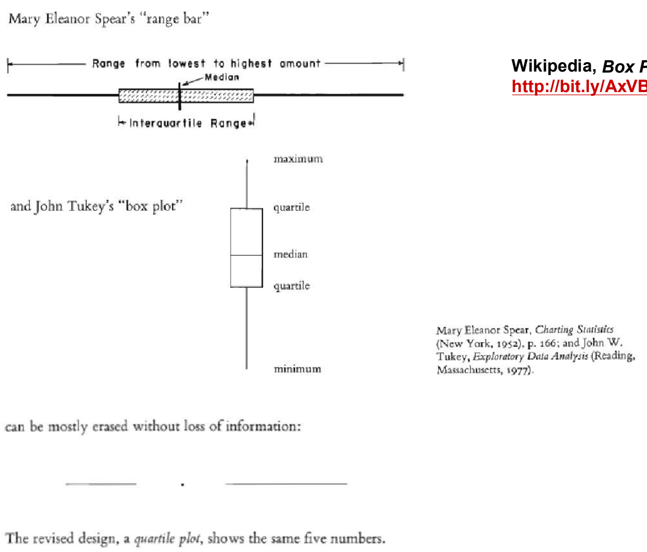
Big Duck © 2000 E. R. Tufte

"Duck" here refers to self-promoting decorative graphics.

Forgo chartjunk, including moiré vibration, the grid, and the duck. – Edward Tufte

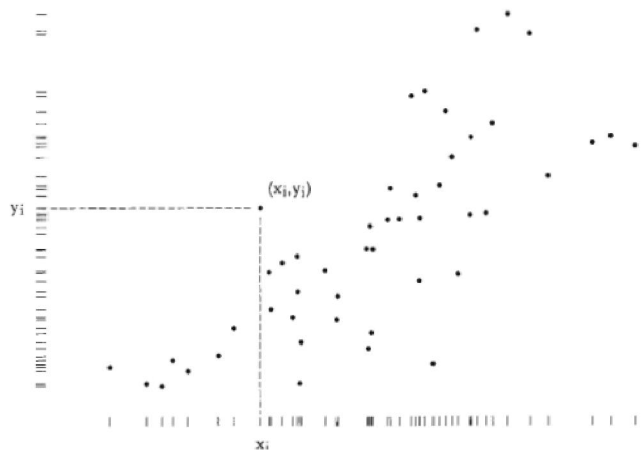
The Visual Display of Quantitative Information, Graphics Press: http://bit.ly/16Se1

Statistical Graphics [1]: Box plots (Spear, 1958; Tukey, 1977)



Wikipedia, *Box Plot*:
http://bit.ly/AxVBN

Statistical Graphics [2]: Dot-Dash Scatter Plots (Tufte, 1983)



Wikipedia, *Scatter Plot*:
http://bit.ly/hkGYri

In a dot-dash plot, sometimes called a Tufte scatterplot, the axes are replaced with marginal distributions (the projection of the bivariate scatter down to one variable)

See: http://bit.ly/f6yDru
© 2008 University of Michigan

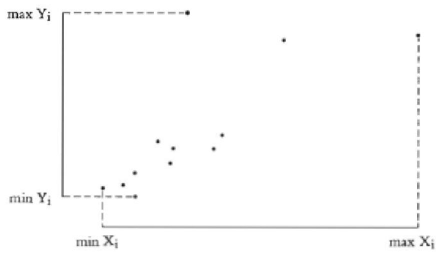
Statistical Graphics [3]: Stemplots (Bowley, c. 1900)

- Also Known As Stem-and-Leaf Plots (Wikipedia: http://bit.ly/adrgEM)
- Construct Distribution of Variable Using Numbers Themselves
- Construction Algorithm
  - Sort data in ascending order – example:

    `44 46 47 49 63 64 66 68 68 72 72 75 76 81 84 88 106`
  - Designate meaning of stems, leaves
    - Leaves: suffixes, to right of vertical line – usually last digit
    - Stems: prefixes, to left of vertical line – usually all other digits
  - Group data by common stem (prefix) – example:

```
4  | 4 6 7 9
5  |
6  | 3 4 6 8 8
7  | 2 2 5 6
8  | 1 4 8
9  |
10 |
6 key: 6|3=63, leaf unit: 1.0, stem unit: 10.0
```

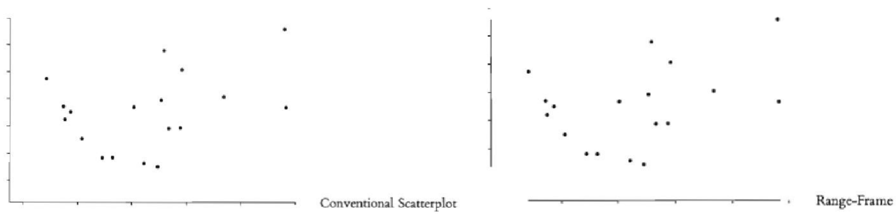

Statistical Graphics [4]: Range-Frame Plot



The min and max values for each variable are made implicit by erasing the axes below the min and above the max value.

This is an example of how erasing non-data ink can lead to greater data density and data-ink ratio.

Nothing but the tails of the frame need change:

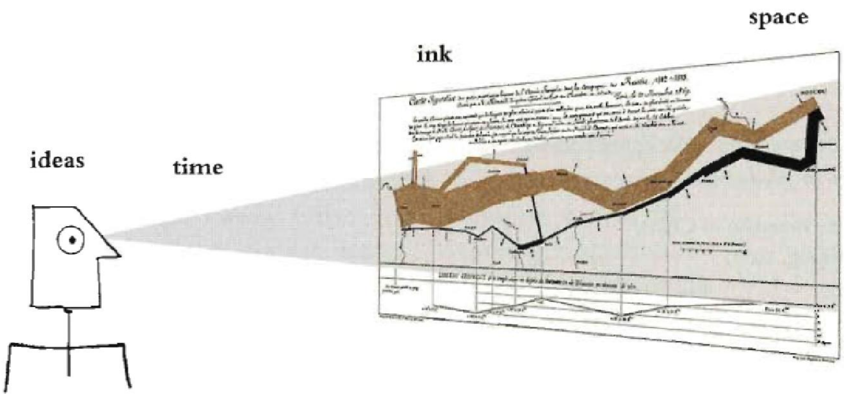


Graphical Excellence – Synopsis [1]

- Well-Designed Presentation of Interesting Data
  - Substance – accurate, precise data, labels, other original content
  - Statistics – analytical content, summative & descriptive
  - Design – non-redundant, concise presentation of multiple variables
- Complex Ideas Communicated with
  - Clarity – ease of interpretation and understanding
  - Precision – ability to reconstruct original data, process
  - Efficiency – rendering time *vs.* reading & comprehension time
- Gives to Viewer
  - Greatest number of ideas – data
  - In shortest time – "ink ratio" really rate per time (cognitive effort)
  - With least ink – filled space, pixels, primitives, rendered objects
  - In smallest space – total size of graphic, page, viewport, window

Graphical Excellence – Synopsis [2]

- Nearly Always Multivariate
- Requires Telling Truth about The Data (Graphical Integrity)



p. 51 (Tufte 1 *aka* VDQI, 1e & 2e)

43

Summary

- Reading for Last Class: Chapter 15, Eberly *2*^{e}; Ray Tracing Handout
- Reading for Today: Tufte Handout
- Reading for Next Class: Ray Tracing Handout
- Last Time: Ray Tracing 2 of 2
  - Stochastic & distributed RT
  - Hybrid RT (for specular reflectance) & radiosity (for diffuse)
- Today: Visualization Part 1 of 3 – Statistical, Scientific, Data/Info Vis
  - Tufte 1: *The Visual Display of Quantitative Information, 2*^{e}
  - Graphical integrity
    - Lack: lie factor ("How to lie with statisticsvisualization")
    - Desiderata: transparency; labeled axes, clear comparisons
    - "Show variation in data, not presentation"
  - Graphical excellence
    - Lack: chartjunk
    - Desiderata: data-ink, data-ink ratio (& "data-pixels")

44

Terminology

- Visualization: Using Images, Diagrams, Animations to Communicate
  - Scientific: transformation, representation of data for exploration
  - Statistical / data: info in schematic form (attributes, variables)
  - Information: computational tools; analyzing large, abstract data sets
- Statistical Visualization Techniques
  - Boxplot *aka* range bar, box-and-whisker diagram: mean, quartiles
  - Dot-dash plot *aka* Tufte scatterplot *aka* scatter plot with Tufte axes
  - Stemplots *aka* stem-and-leaf display: prefix (stem), suffix (leaf)
  - Range-frame plot: erase axes outside range (min/max x, y)
- Tufte 1: *The Visual Display of Quantitative Information, 2*^{e}
  - Graphical integrity: accurate, truthful visual communication
  - Example of lapse in graphical integrity: lie factor (distortion ratio)
  - Data-ink ratio: quantity of usable/accessible info per unit of "ink"
  - Graphical excellence: high data-ink ratio, no wasted axes
  - Antithesis of graphical excellence: chartjunk (visual clutter)