

Visualization

COS 323

Many slides based on CHI 2003 tutorial by Marti Hearst
and on Jeff Heer's CS 160 slides (2004)

What is Information Visualization?

“Transformation of the symbolic into the geometric”

(McCormick et al., 1987)

“... finding the artificial memory that best supports our natural means of perception.”

(Bertin, 1983)

“The depiction of information using spatial or graphical representations, to facilitate comparison, pattern recognition, change detection, and other cognitive skills by making use of the visual system.” (Hearst, 2003)

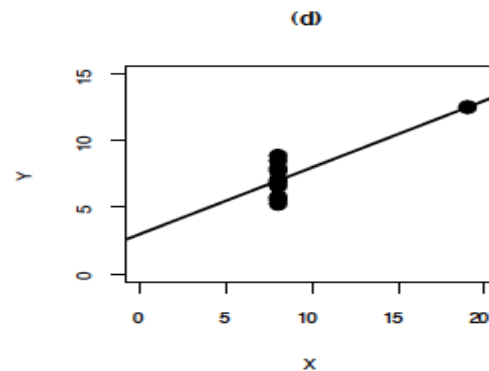
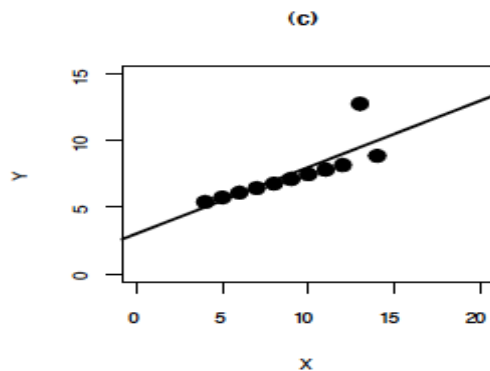
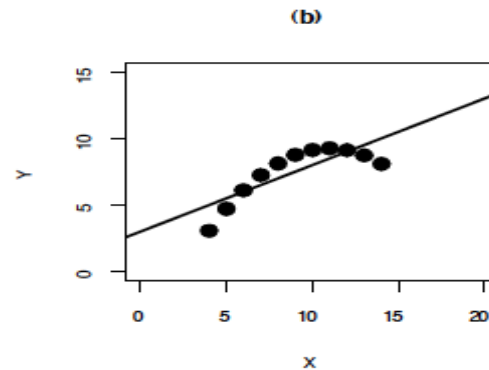
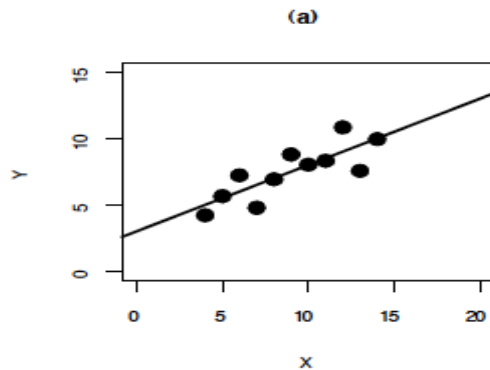
Information Visualization

- Problem
 - How to understand datasets?

Dataset 1		Dataset 2		Dataset 3		Dataset 4	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.50
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Information Visualization

- Problem
 - How to understand datasets?
 - How to understand big, multi-dimensional datasets??



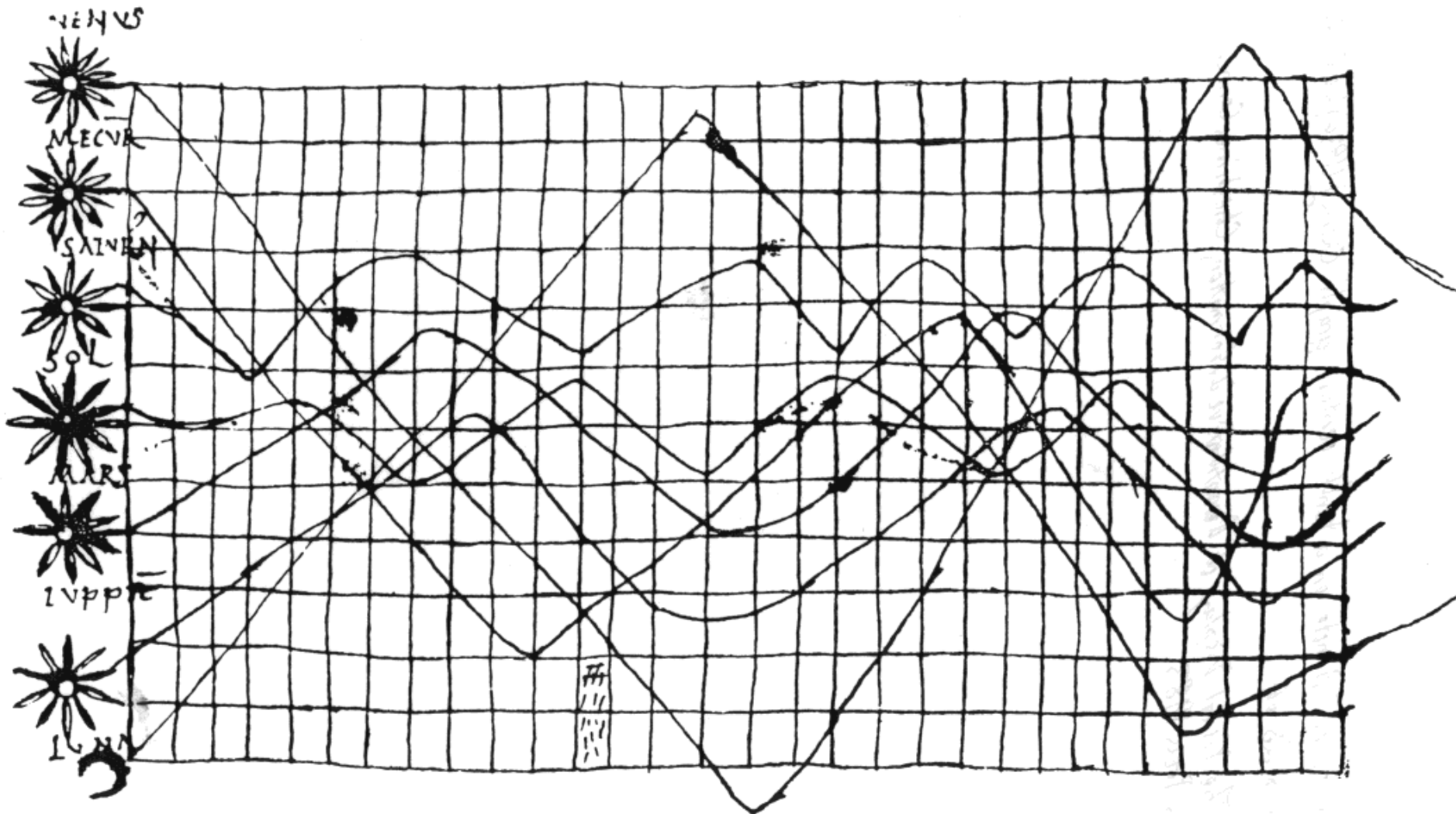
Information Visualization

- Solution
 - Take advantage of the human perceptual system
 - Convert information into a graphical representation.
- Issues
 - How to convert abstract information into graphical form?
 - Do visualizations do a better job than other methods?

Goals of Information Visualization

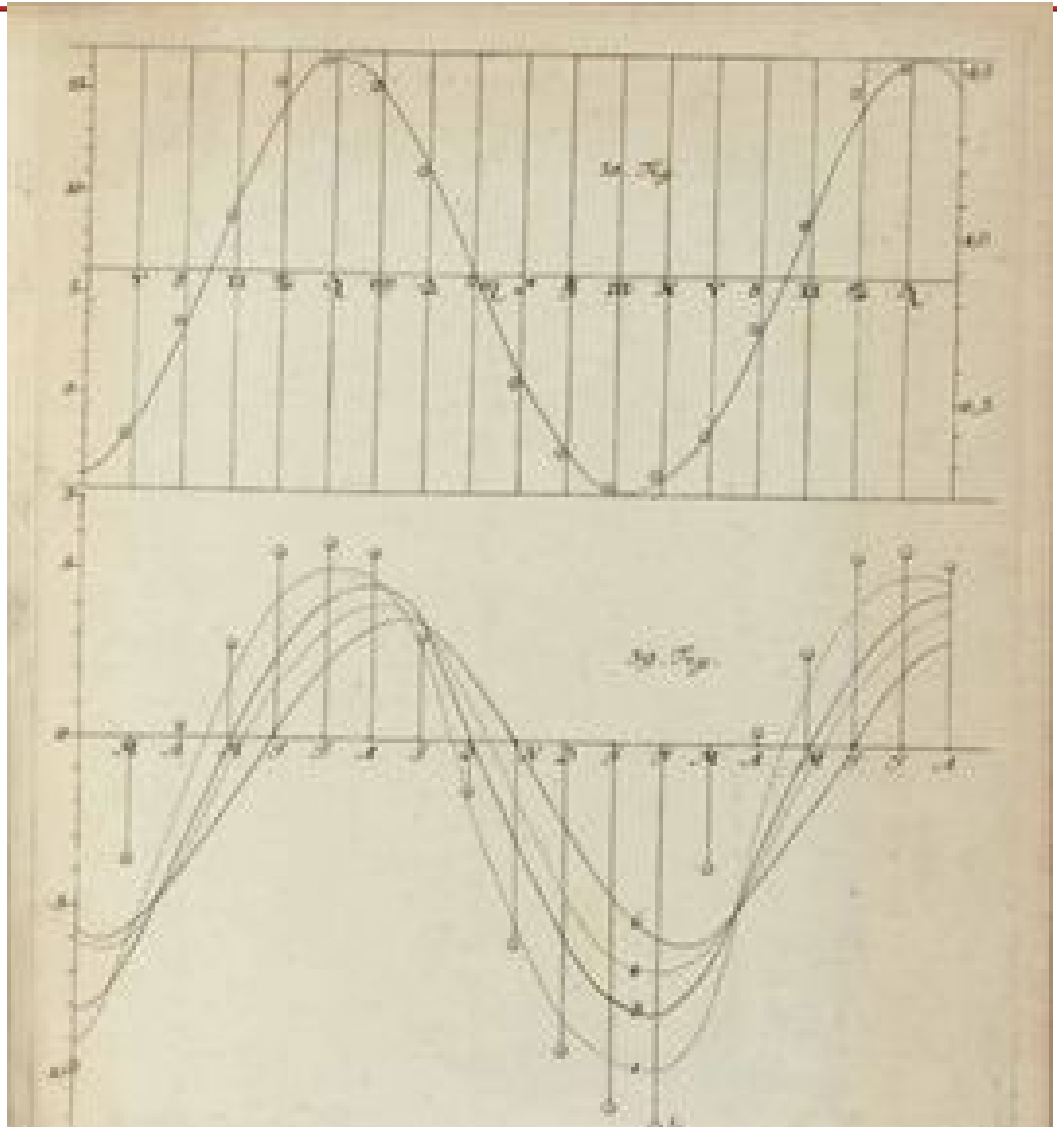
- More specifically, visualization should:
 - Make large datasets coherent
(Present huge amounts of information compactly)
 - Present information from various viewpoints
 - Present information at several levels of detail
(from overviews to fine structure)
 - Support visual comparisons
 - Tell stories about the data

An Early Visualization

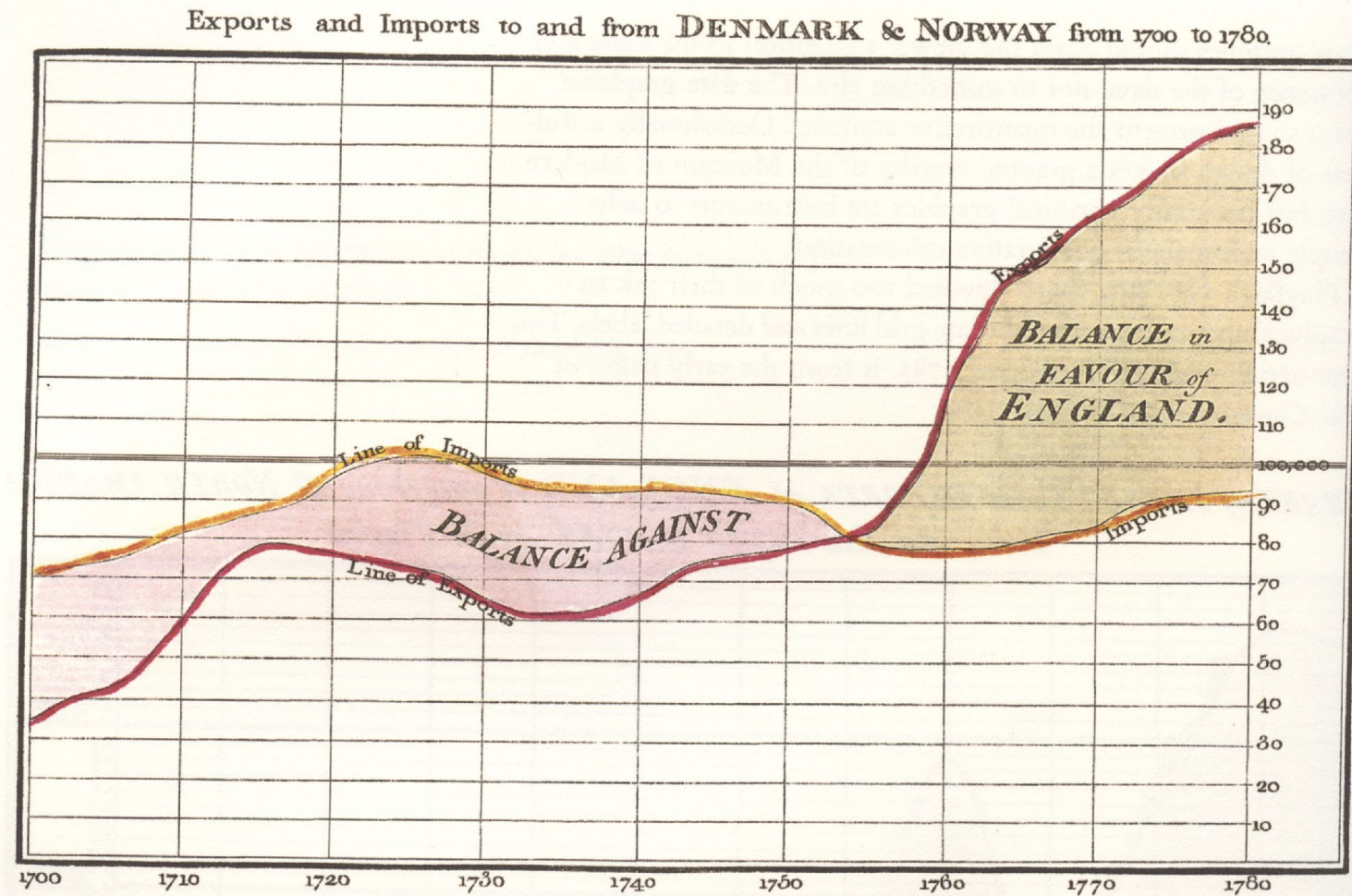


Time series appears in scientific writing in late 1700s

- Lambert's
Pyrometrie,
1779



William Playfair

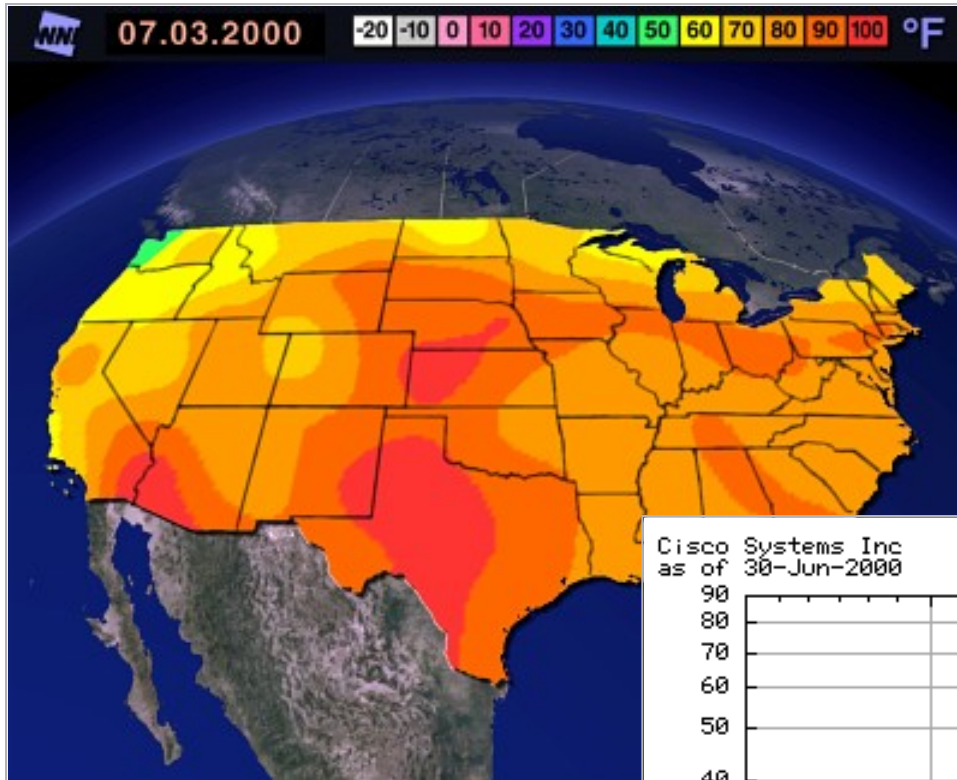


The Bottom line is divided into Years, the Right hand line into £10,000 each.

Published as the Act directs, 14th May 1786, by W^m Playfair

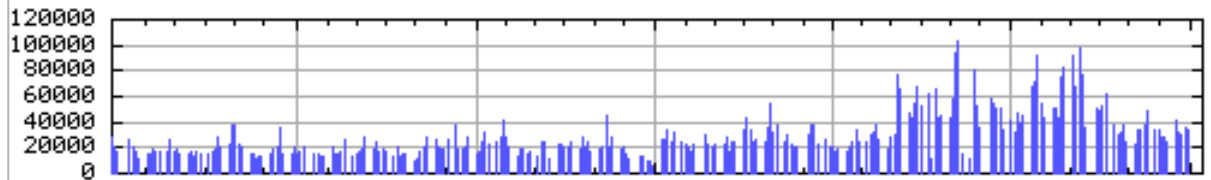
Neele sculpt 352, Strand, London.

Visualization Success Stories



Cisco Systems Inc
as of 30-Jun-2000

Splits: ▼ Mov Avg: 200 day 50 day



Copyright 2000 Yahoo! Inc.

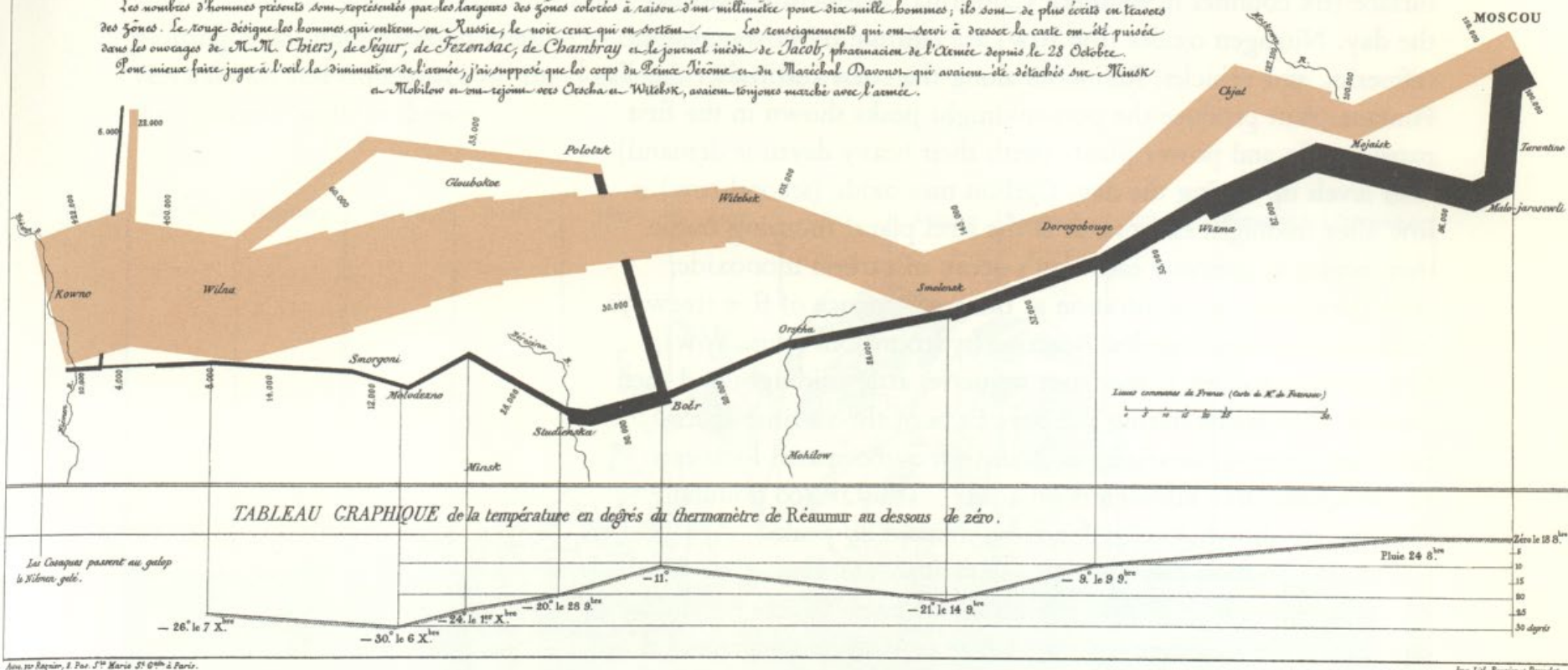
Volume (1000's)

<http://finance.yahoo.com/>

Napoleon's 1812 March by Charles Joseph Minard

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.
Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

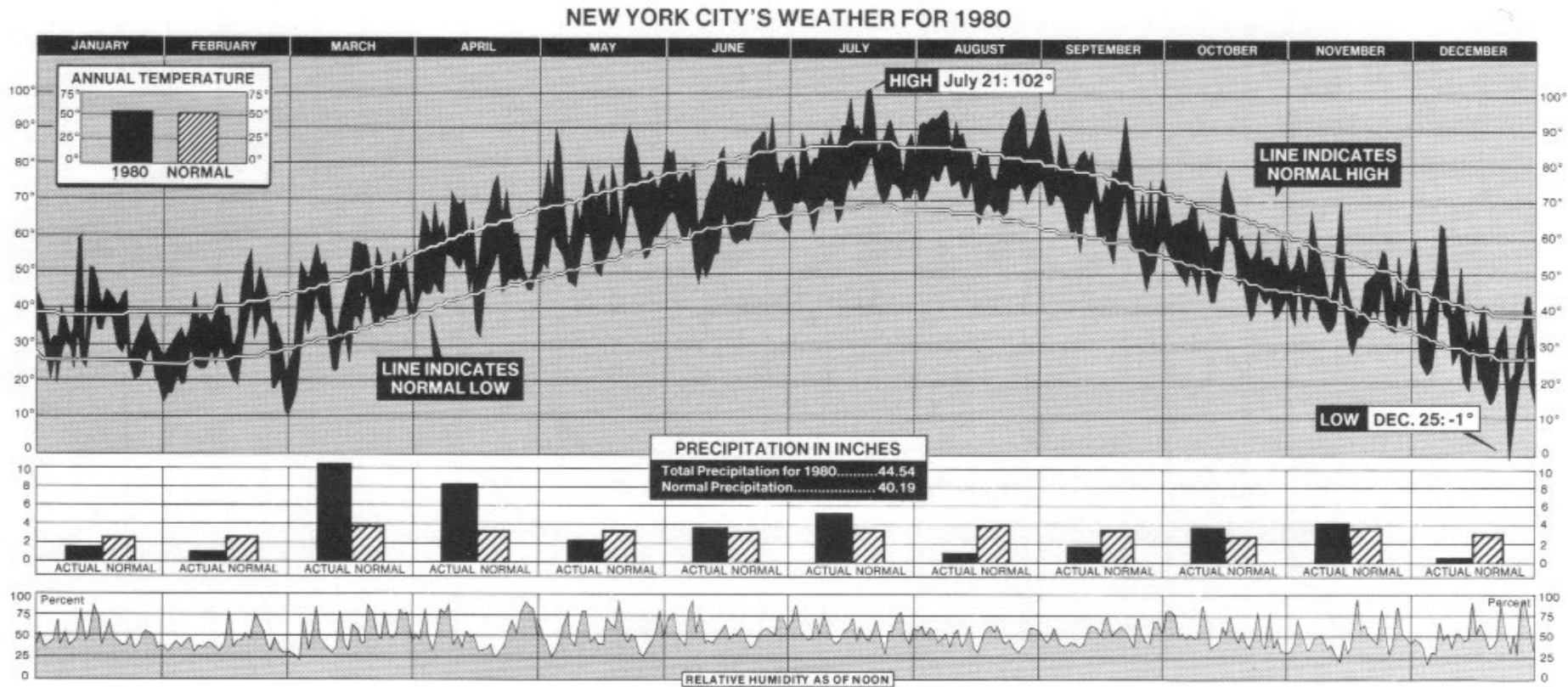
Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilew et qui rejoignent vers Orescha et Witebsk, avaient toujours marché avec l'armée.



Variables shown:

- size of army
- direction
- latitude
- longitude
- temperature
- date

NYC Weather



New York Times, January 11, 1981, p. 32.

1888 numbers

Visualization Success Story

Mystery: what is causing a cholera epidemic in London in 1854?

Visualization Success Story

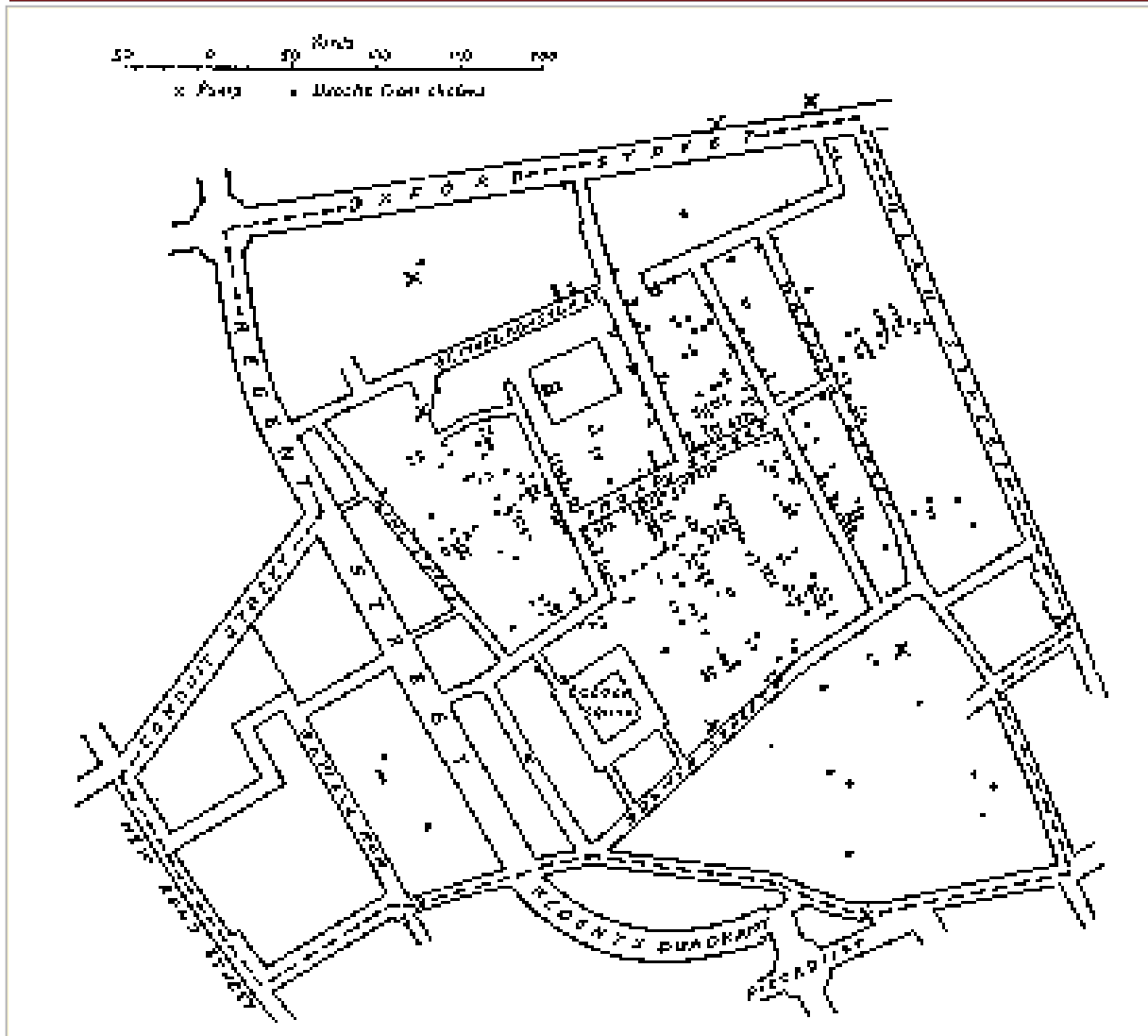


Illustration of John Snow's deduction that a cholera epidemic was caused by a bad water pump, circa 1854.

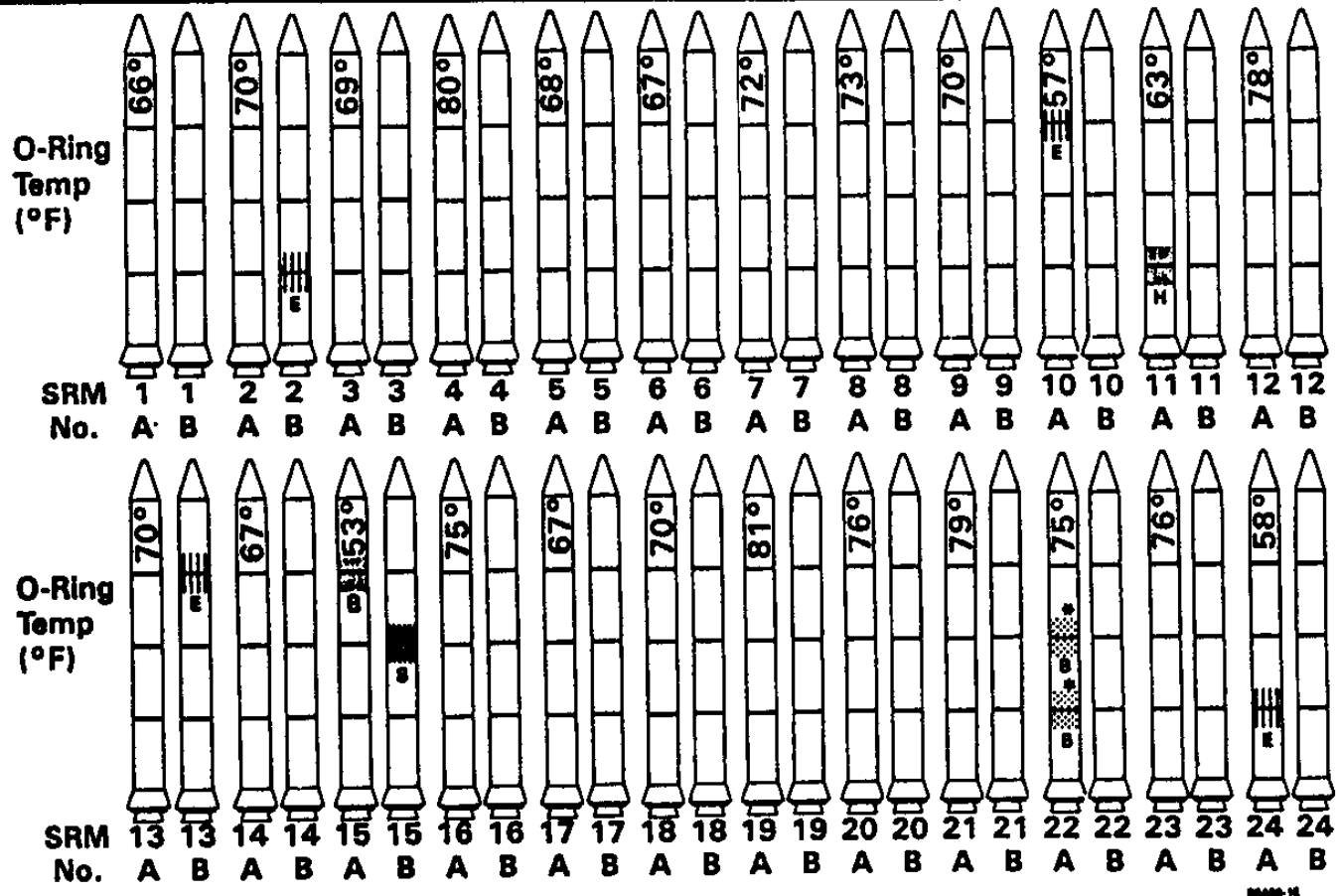
Dots indicate locations of deaths.

Visualization Success Story



Visualization Failure

History of O-Ring Damage in Field Joints (Cont)



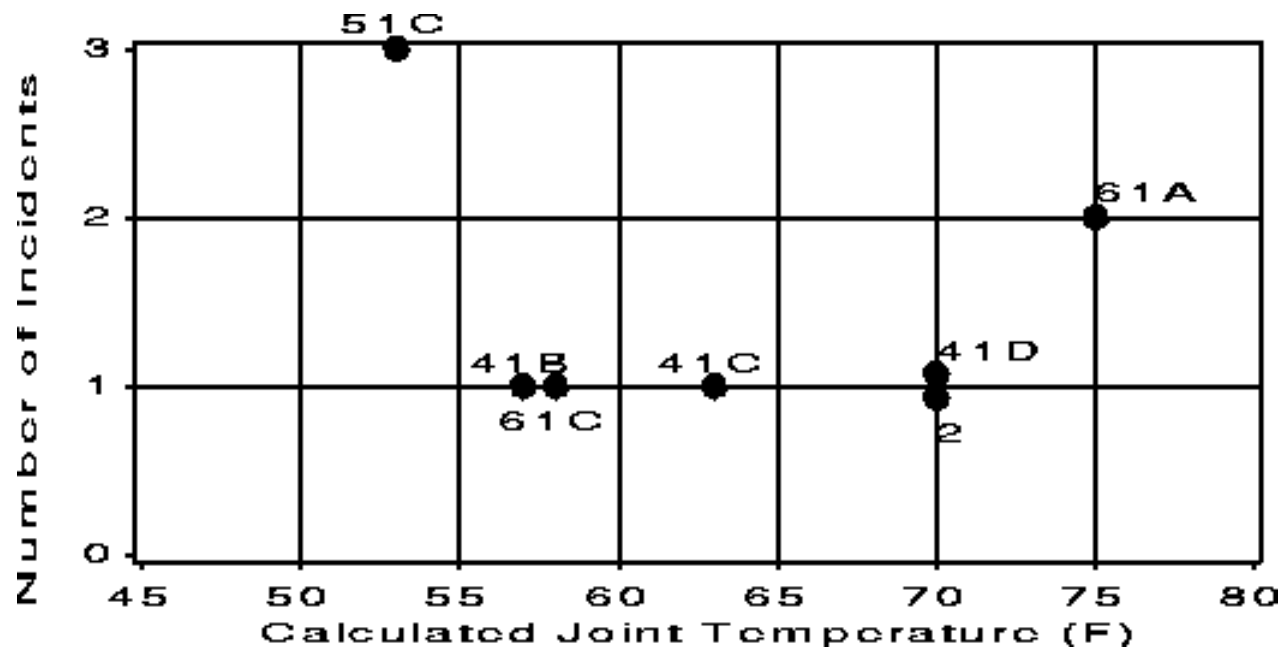
MORTON THOKOL, INC.
Wasatch Operations

* No Erosion

00400-10

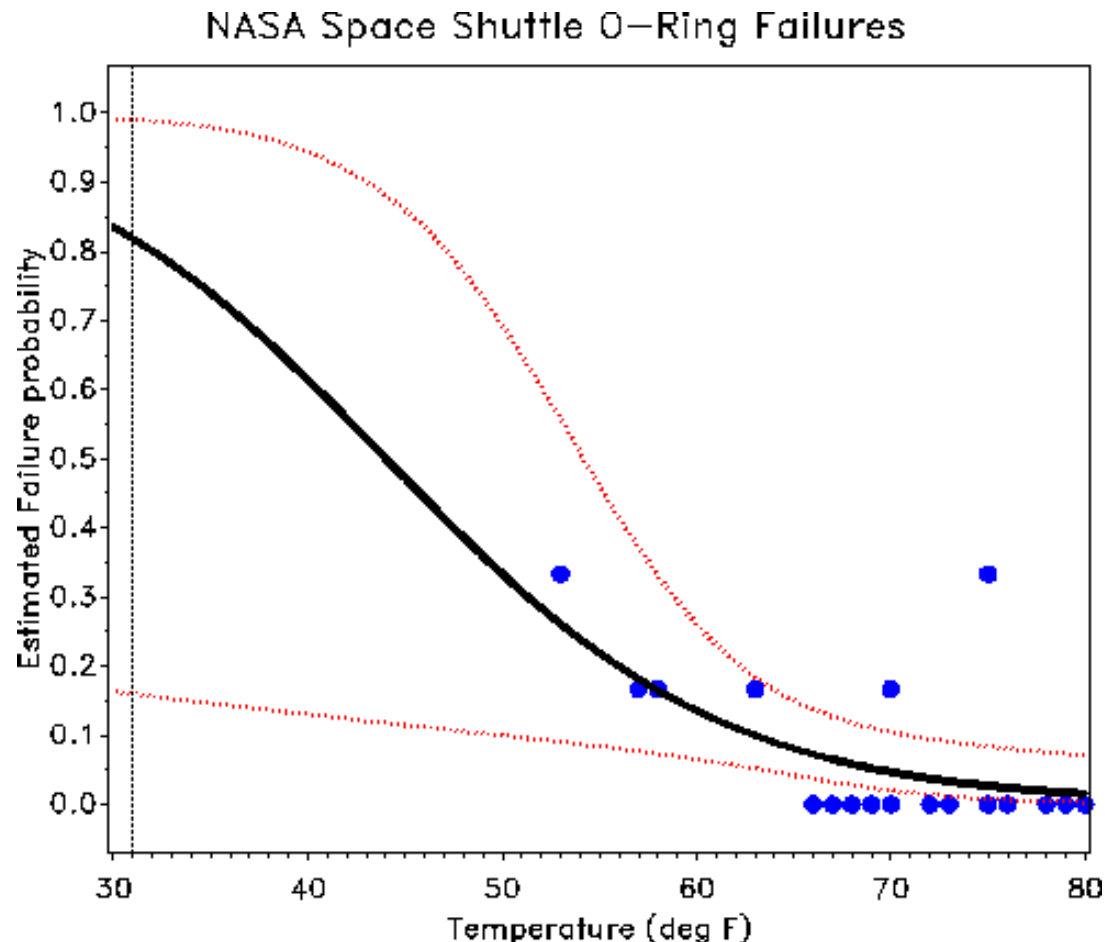
Visualization Failure

The visualization they made...



Visualization Failure

The one they should have made...



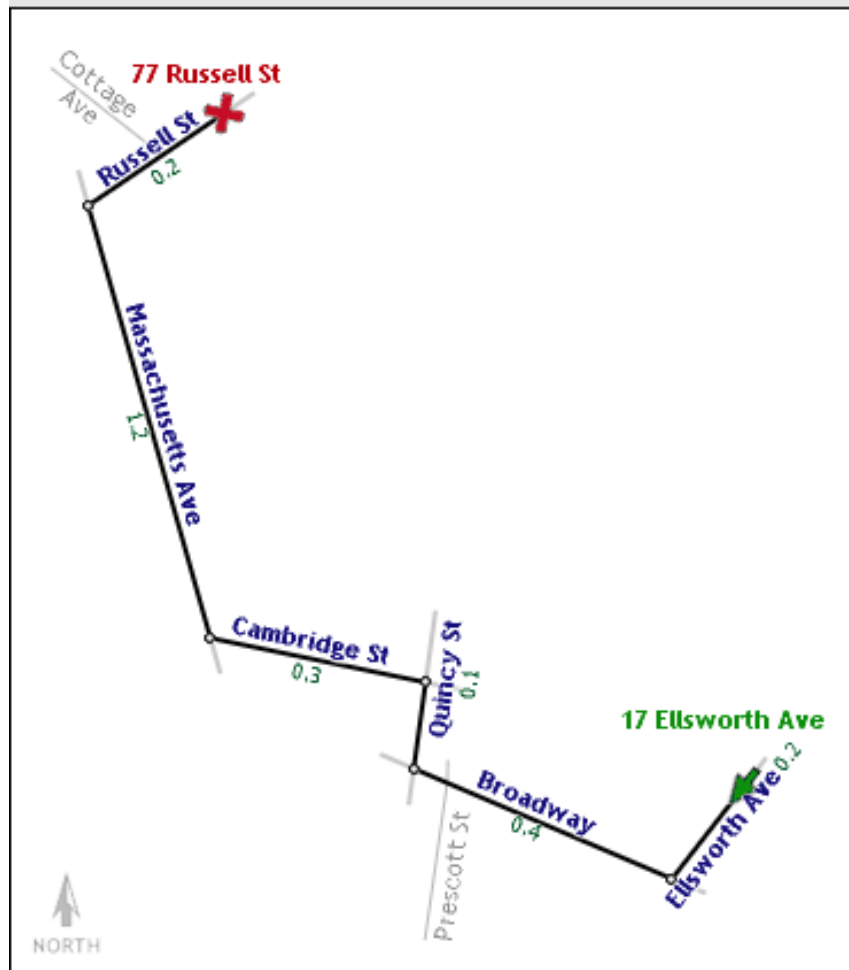
A Map Visualization circa 1999

1. Start out going Southwest on ELLSWORTH AVE
Towards BROADWAY by turning right.
2. Turn RIGHT onto BROADWAY.
3. Turn RIGHT onto QUINCY ST.
4. Turn LEFT onto CAMBRIDGE ST.
5. Turn SLIGHT RIGHT onto
MASSACHUSETTS AVE.
6. Turn RIGHT onto RUSSELL ST.



A more useful visualization

The estimated travel time is 5 minutes for 2.16 miles of travel, total of 6 steps.



Directions		Elapsed Distance
1	Begin at 17 Ellsworth Ave on Ellsworth Ave and go Southwest for 500 feet	0.1
2	Turn right on Broadway and go Northwest for 0.4 miles	0.5
3	Turn right on Quincy St and go North for 200 feet	0.5
4	Turn left on Cambridge St and go West for 0.3 miles	0.8
5	Bear right on Massachusetts Ave, Mass Ave, RT-2A and go North for 1.2 miles	2.0
6	Turn right on Russell St and go Northeast for 1000 feet to 77 Russell St	2.2

Why Visualization?

- Use the eye for pattern recognition; people are good at
 - scanning
 - recognizing
 - remembering images
- Graphical elements facilitate comparisons via
 - length
 - shape
 - orientation
 - texture
- Animation shows changes across time
- Color helps make distinctions
- Aesthetics help maintain interest

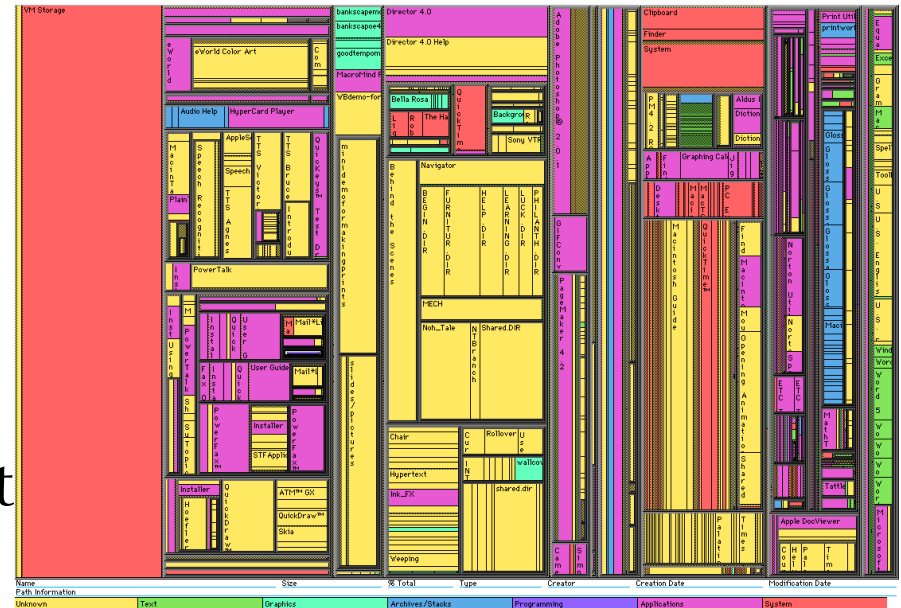
Two Different Primary Goals: Two Different Types of Viz

- Explore/Calculate
 - Analyze
 - Reason about Information
- Communicate
 - Explain
 - Make Decisions
 - Reason about Information

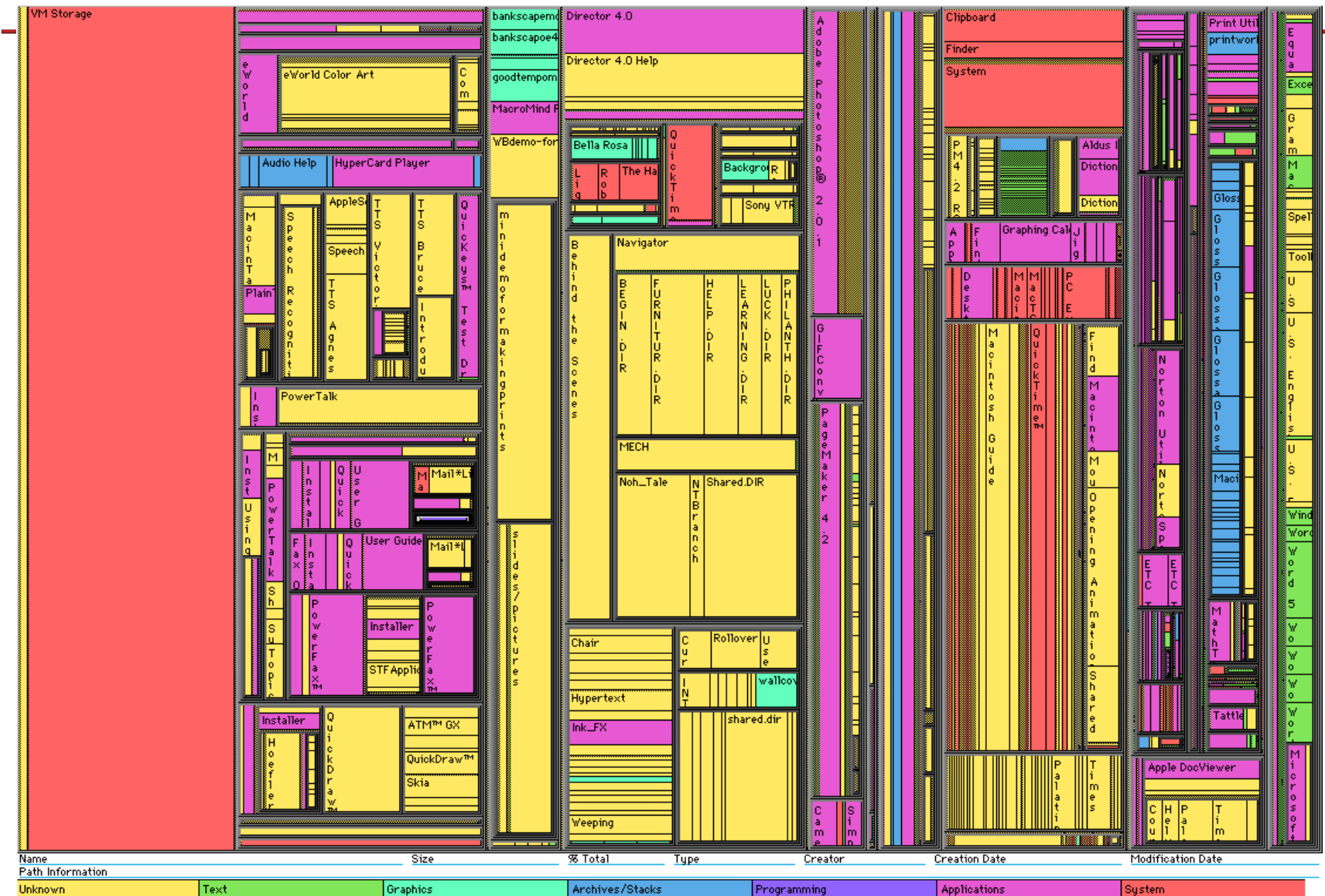
Case Study:

The Journey of the TreeMap

- The TreeMap [Johnson & Shneiderman '91]
- Idea:
 - Show a hierarchy as a 2D layout
 - Fill up the space with rectangles representing objects
 - Size on screen indicates relative size of underlying objects



Early Treemap Applied to File System



Treemap Problems

- Too disorderly
 - Adjacency is meaningless
 - Aspect ratios uncontrolled, leads to lots of skinny boxes that clutter
- Color not (necessarily) used appropriately
- Wrong application
 - Don't need all this to just see the largest files

Successful Application of Treemaps

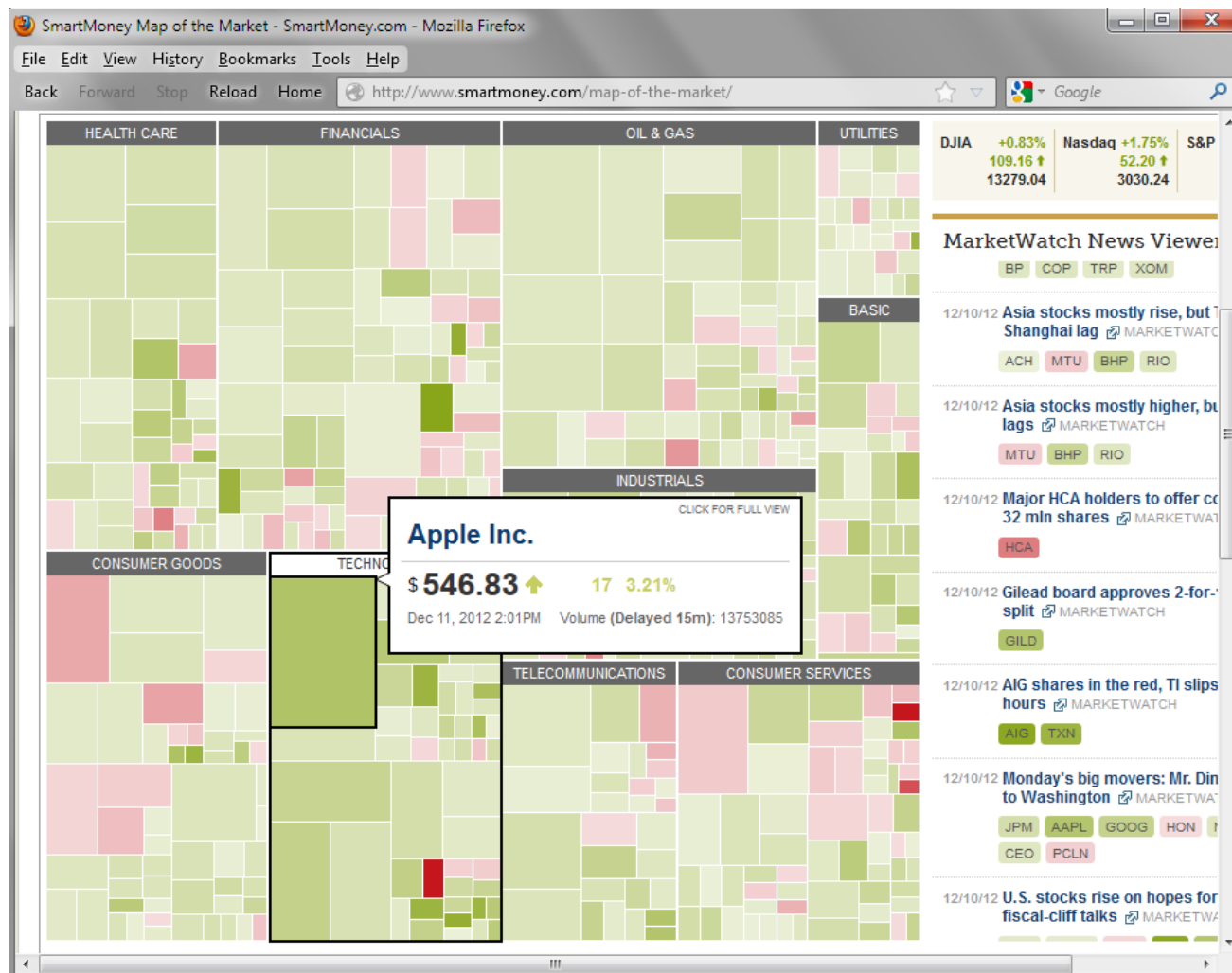
- Think more about the use
 - Break into meaningful groups
 - Fix items into a useful aspect ratio
- Use visual properties (e.g. color) properly
 - Use color to highlight and/or show relationships
- Provide interactivity
 - Access to the real data
 - Makes it into a useful tool

TreeMaps in Action

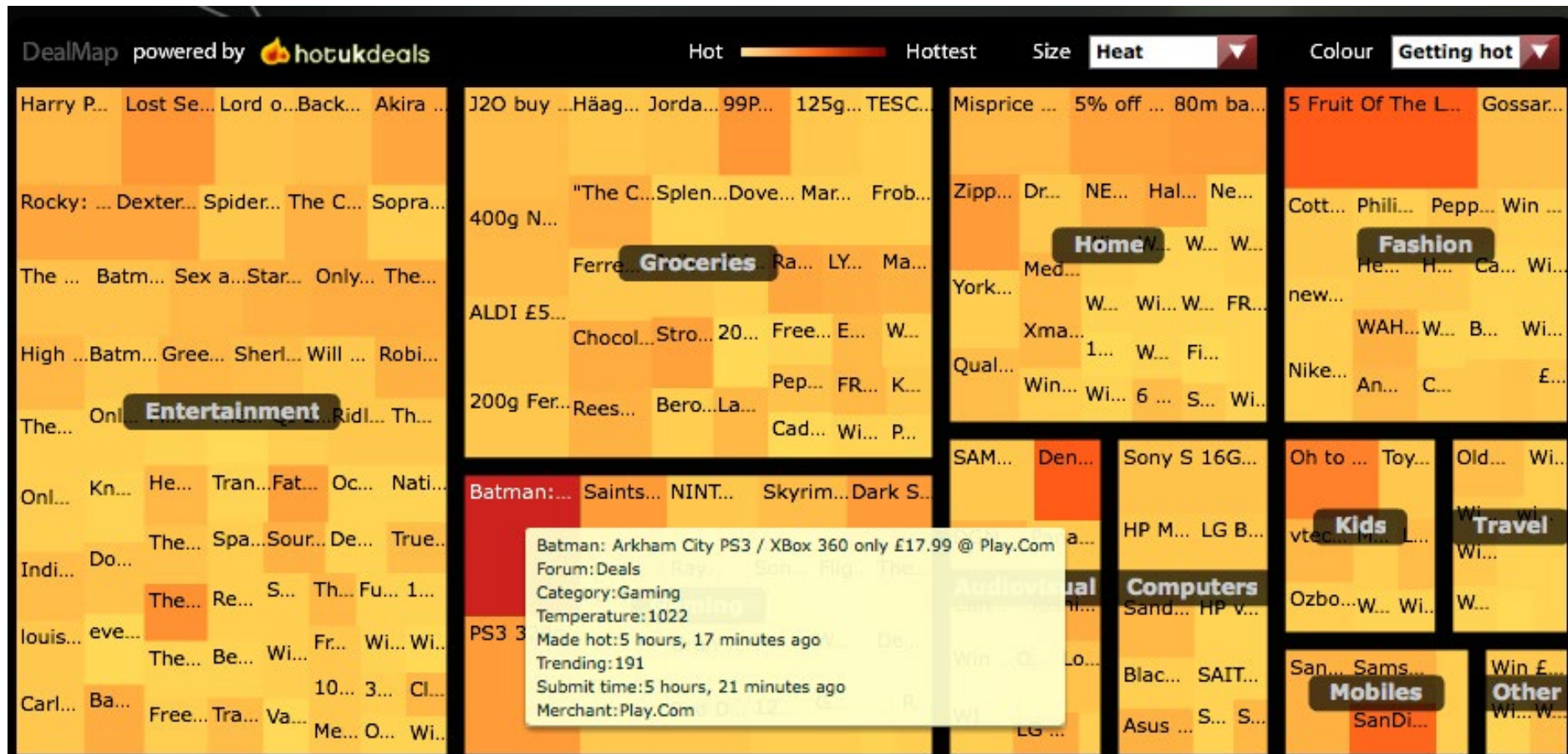
<http://www.smartmoney.com/map-of-the-market>

<http://www.dealmapper.co.uk/>

A Good Use of TreeMaps and Interactivity



Deal Mapper



Analysis vs. Communication

- MarketMap's use of TreeMaps allows for sophisticated **analysis**
- DealMapper's use of TreeMaps is more for **presentation** and communication

Visual Principles

Visual Principles

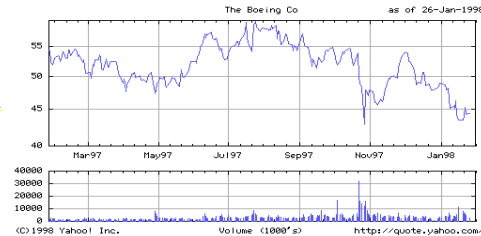
- Types of visual representations
- Pre-attentive Properties
- Relative expressiveness of visual cues
- Visual illusions
- Tufte's notions
 - Graphical Excellence
 - How to Lie with Visualization
 - Data-Ink Ratio Maximization

References for Visual Principles

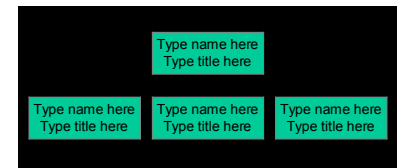
- Kosslyn: Types of Visual Representations
- Lohse et al: How do people perceive common graphic displays
- Bertin, MacKinlay: Perceptual properties and visual features
- Tufte/Wainer: How to mislead with graphs

Types of Symbolic Displays

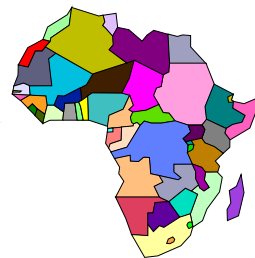
- Graphs



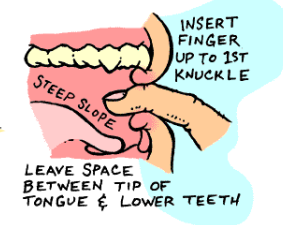
- Charts



- Maps



- Diagrams

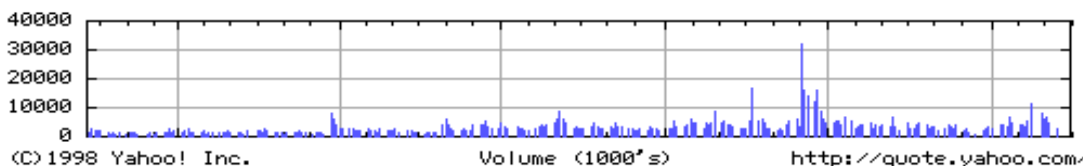


Types of Symbolic Displays

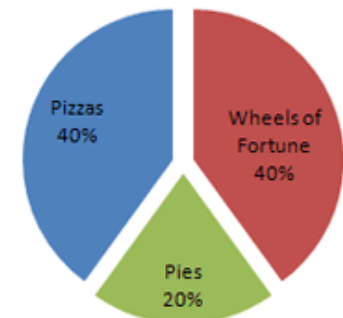
Graphs

- requires scale
- values associated by symmetric “paired with” relation
- “Greater” is represented by bigger, higher, longer, etc.
- Examples: scatter-plot, bar-chart, pie chart

The Boeing Co as of 26-Jan-1998



Things That Look Like Pie Charts

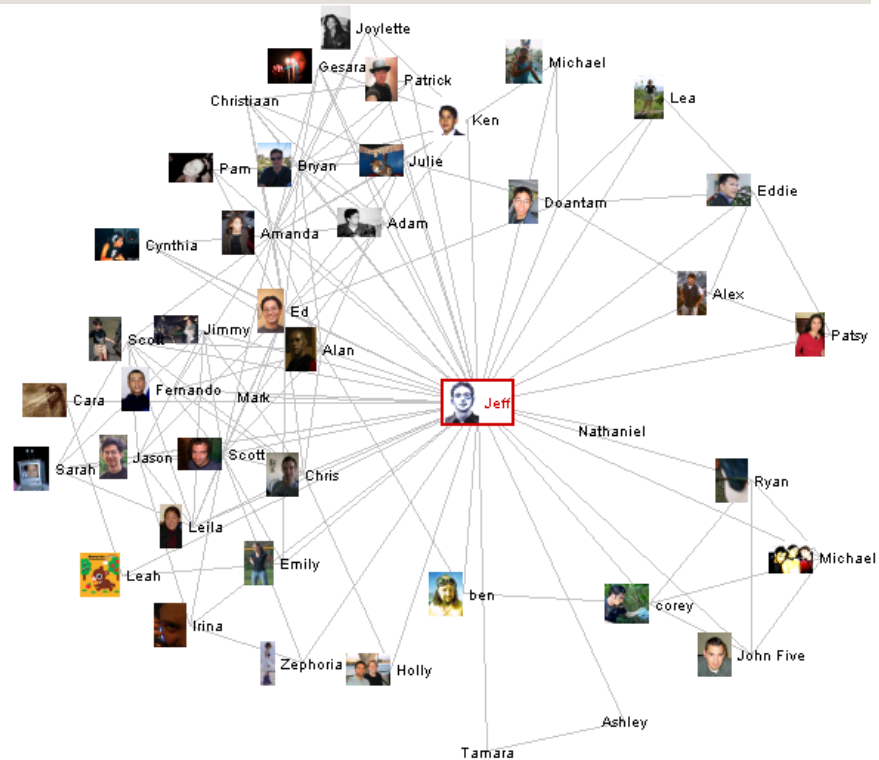


*Not Charted: Peace Symbols

Types of Symbolic Displays

Charts

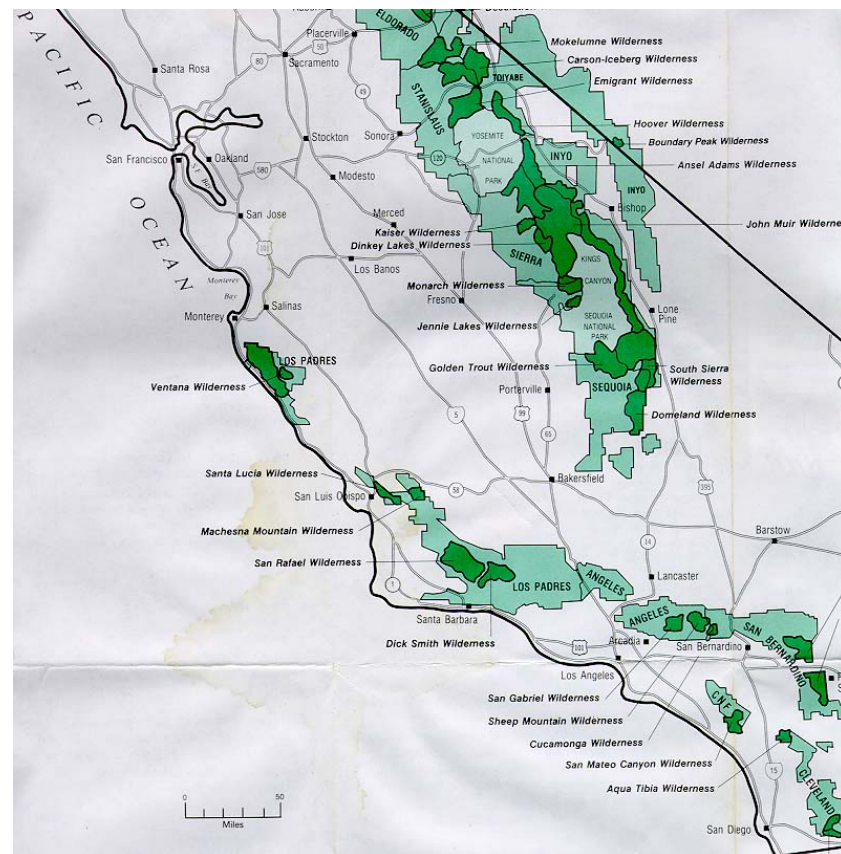
- discrete relations among discrete entities
- structure relates entities to one another
- lines and relative position serve as links
- Examples:
 - family tree,
 - flow chart,
 - network diagram



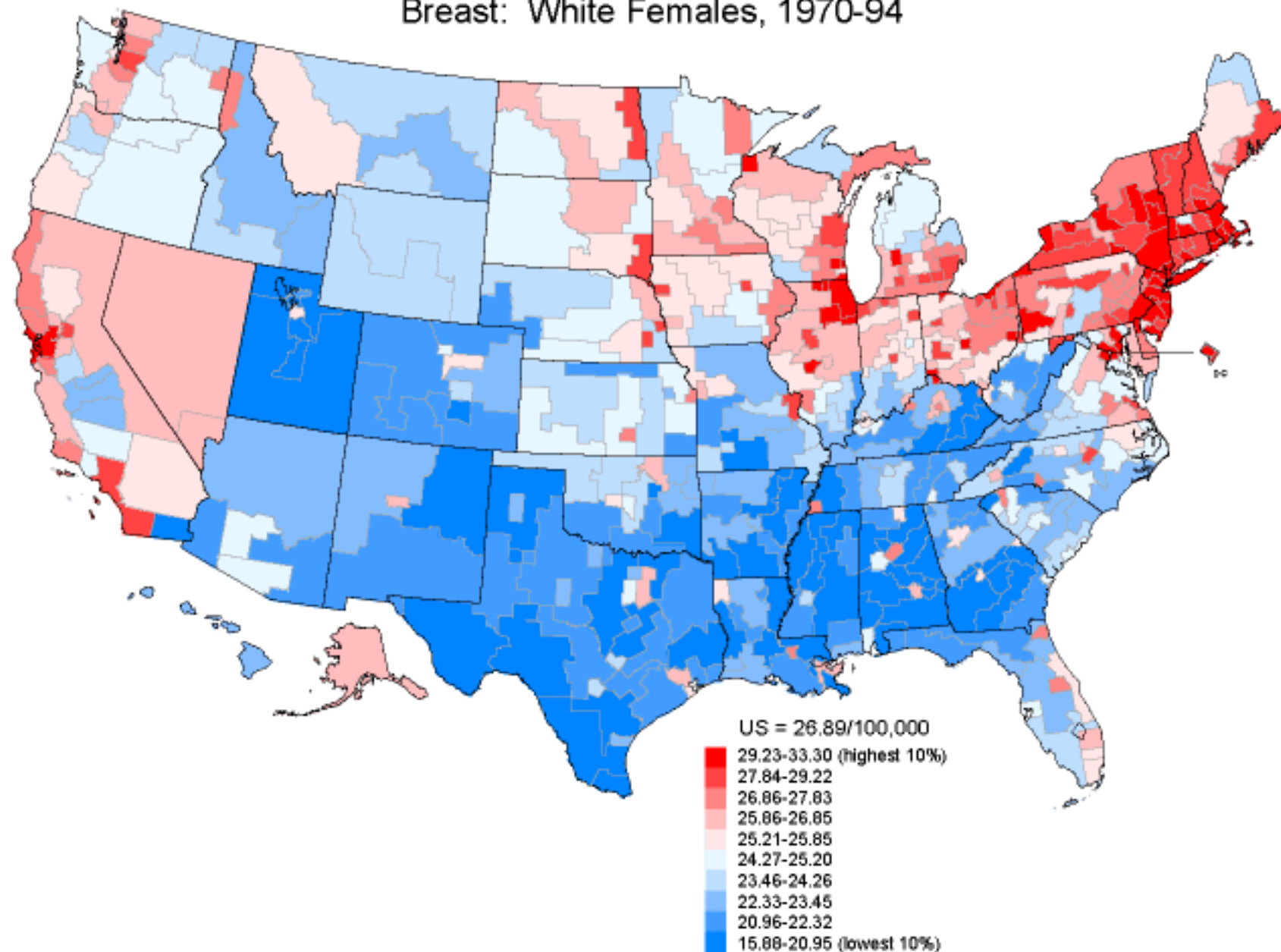
Types of Symbolic Displays

Maps

- internal relations determined (in part) by the spatial relations of what is pictured
- labels paired with locations
- Examples:
 - physical maps,
 - topographic maps,
 - political maps,
 - maps of census data

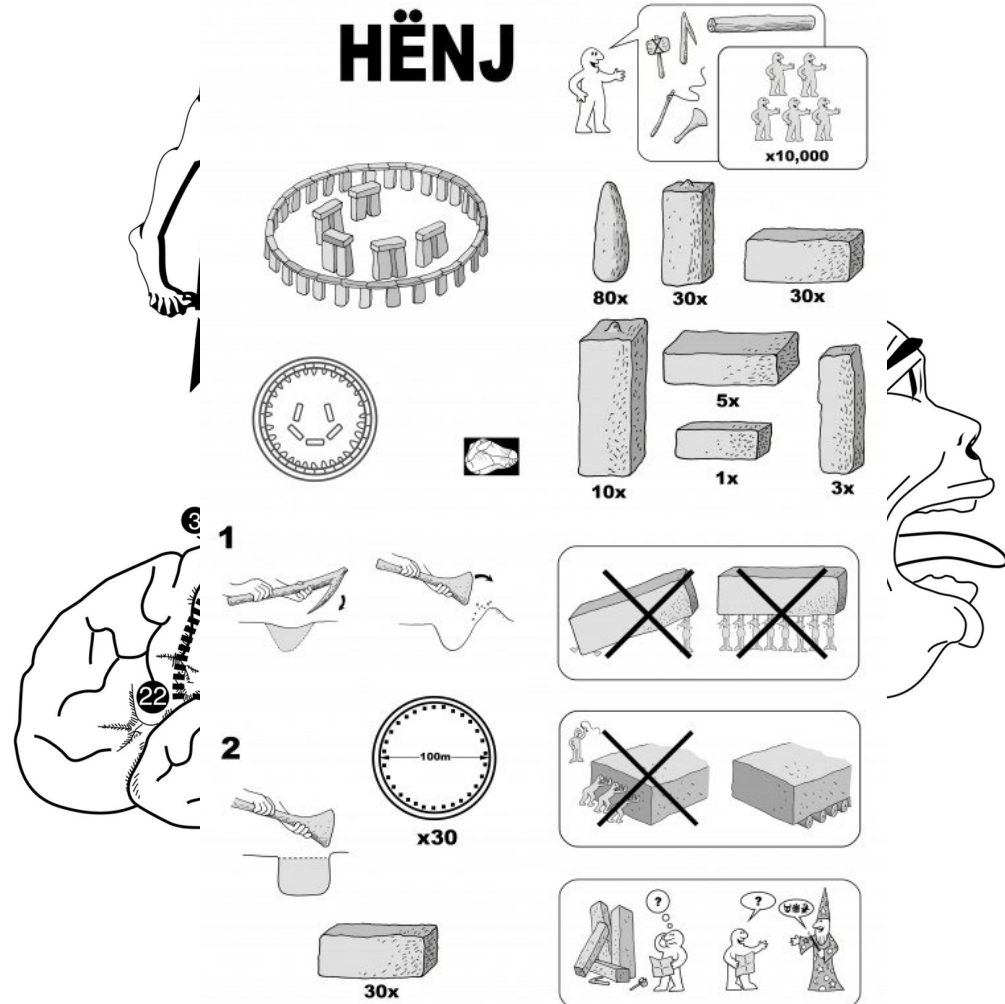


Cancer Mortality Rates by State Economic Area
Breast: White Females, 1970-94



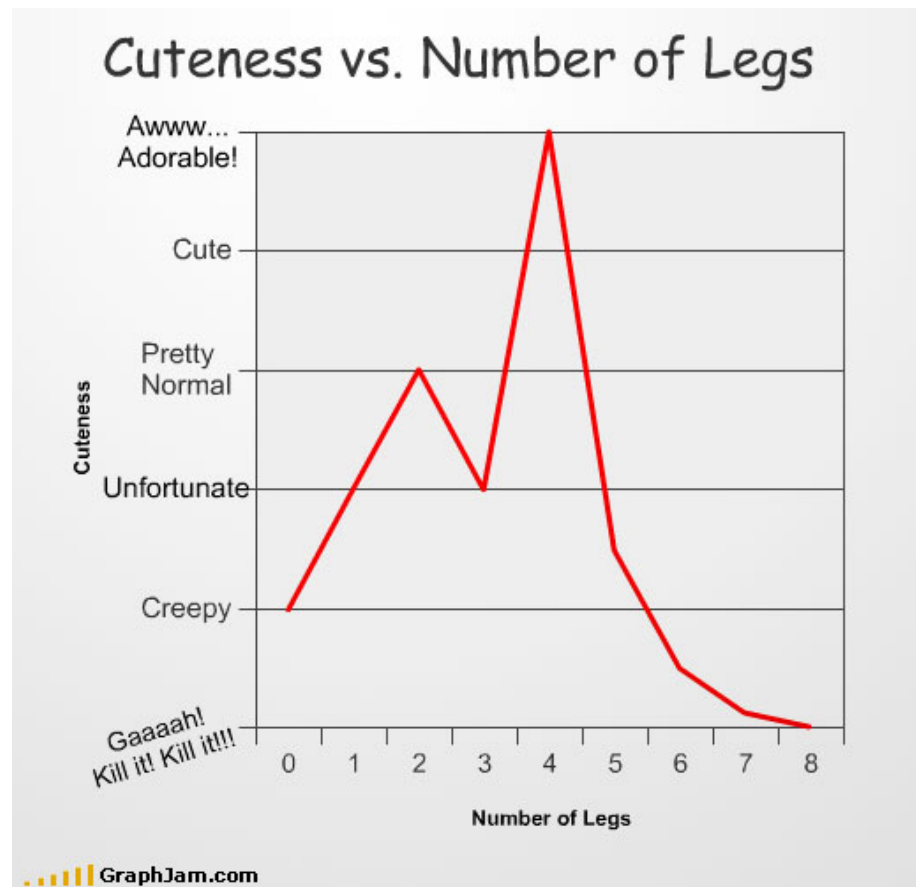
Types of Symbolic Displays

- Diagrams
 - schematic pictures of objects or entities
 - parts are symbolic (unlike photographs)
 - Examples:
 - how-to illustrations, figures in a manual



Anatomy of a Graph [Kosslyn 89]

- Background
- Framework
 - kinds of measurements, scale, ... (outer framework)
 - Includes grid lines (inner)
- Specifier
 - conveys information about entities represented by framework
 - maps parts of the framework to other parts of the framework
 - e.g., the red line ----->
- Labels
 - title, axes, tic marks, ...



Basic Types of Data

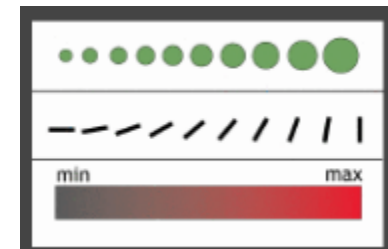
- Nominal (qualitative)
 - (no inherent order)
 - city names, types of diseases, ...



- Ordinal (qualitative)
 - (ordered, but not at measurable intervals)
 - first, second, third, ...
 - cold, warm, hot

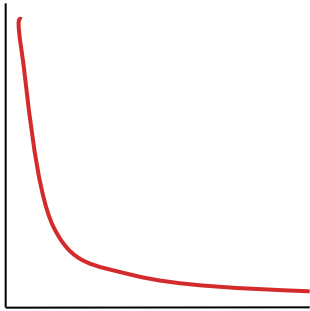


- Interval (quantitative)
 - list of integers or reals



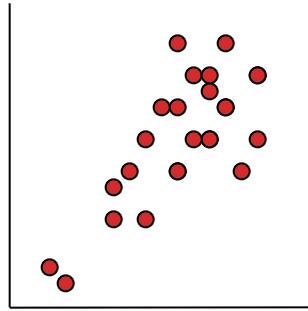
Common Graph Types

of accesses



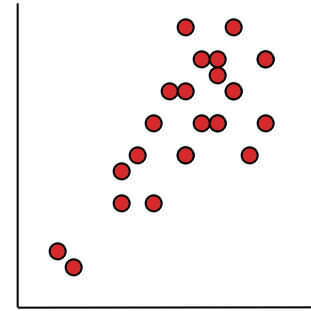
URL

of accesses



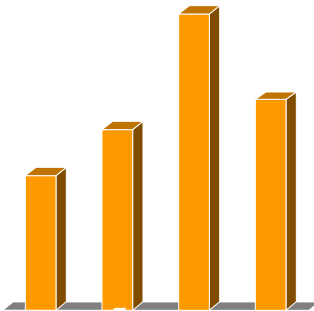
length of access

length of page



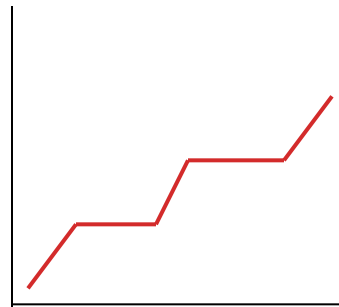
length of access

length of access



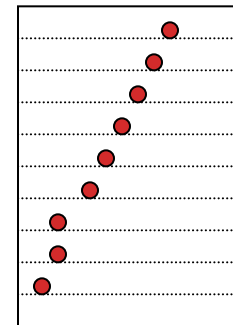
length of page

of accesses



days

url 1
url 2
url 3
url 4
url 5
url 6
url 7



of accesses

When to use which type?

- Line graph
 - x-axis requires quantitative variable
 - Variables have contiguous values
 - familiar/conventional ordering among ordinals
- Bar graph
 - comparison of relative point values
- Scatter plot
 - convey overall impression of relationship between two variables
- Pie Chart?
 - Emphasizing differences in proportion among a few numbers

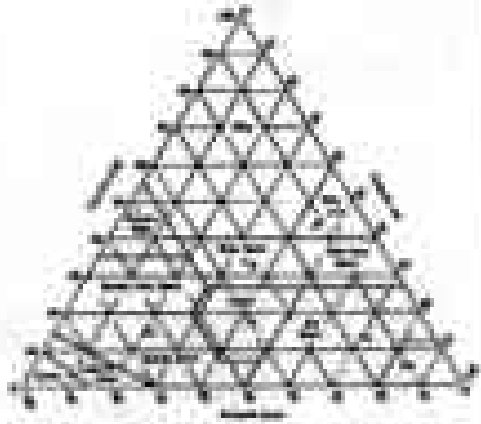
Classifying Visual Representations

- Lohse et al, 1994: **How do people** categorize visual representations?
- Participants sorted 60 visual representations into categories and assigned labels from Likert scales
- Experimenters clustered the results various ways.

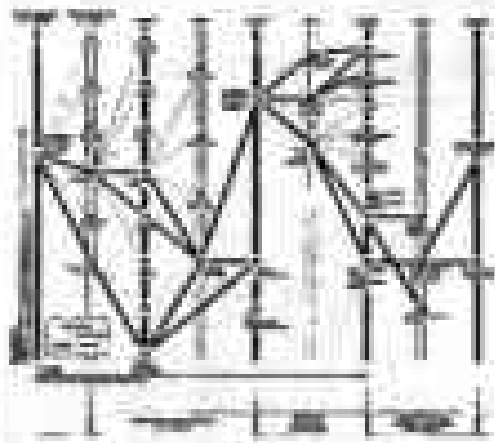
Lohse, G L; Biolsi, K; Walker, N and H H Rueter. "A Classification of Visual Representations." CACM, Vol. 37, No. 12, pp 36-49, 1994.

Subset of Example Visual Representations

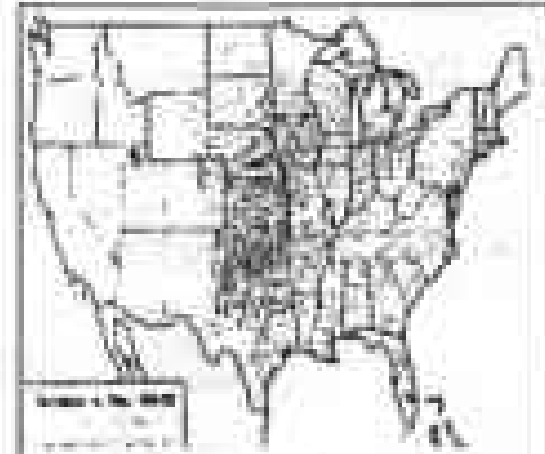
From Lohse et al. 94



1. soil triangle



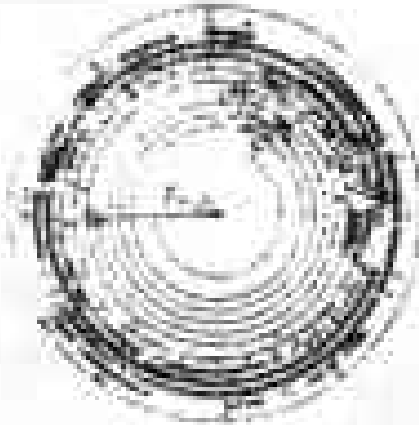
2. missile crisis



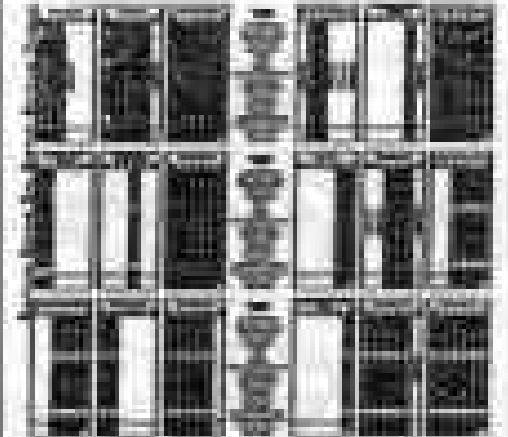
3. USA tornadoes



5. microscope



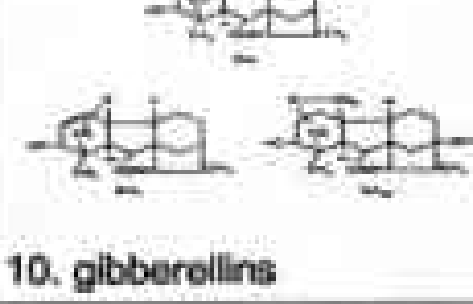
6. circular tree diagram



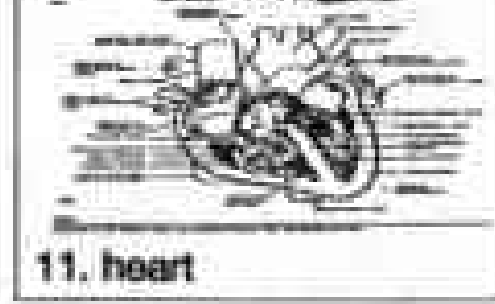
7. auto repair records



Lake



10. gibberellins



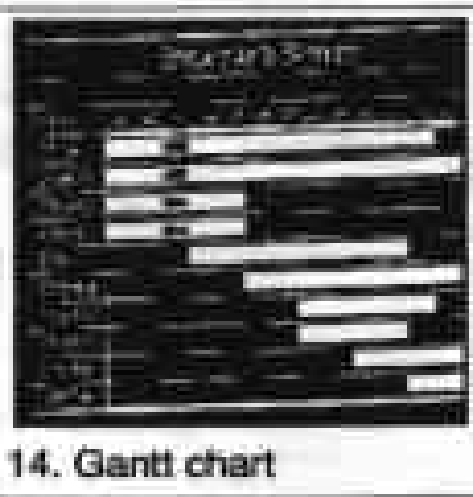
11. heart



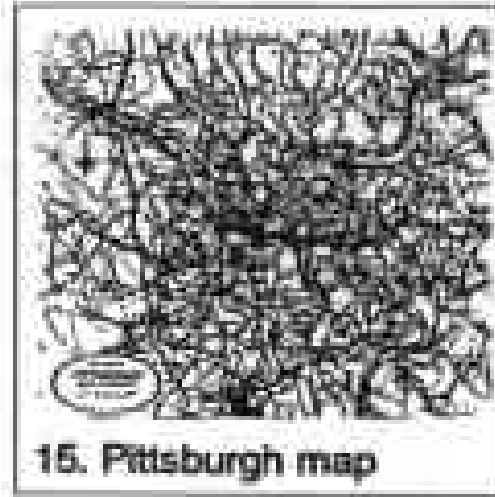
12. microbiology



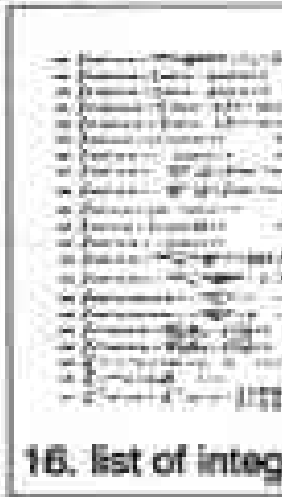
spreadsheet budget



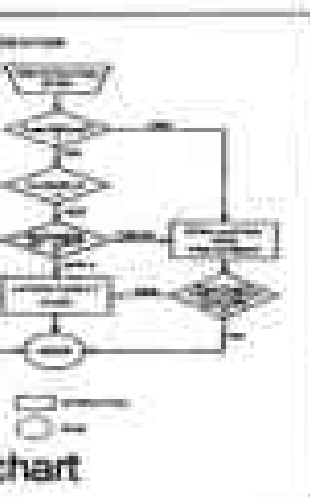
14. Gantt chart



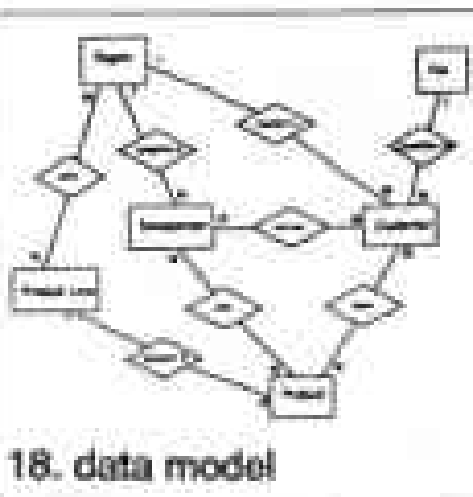
15. Pittsburgh map



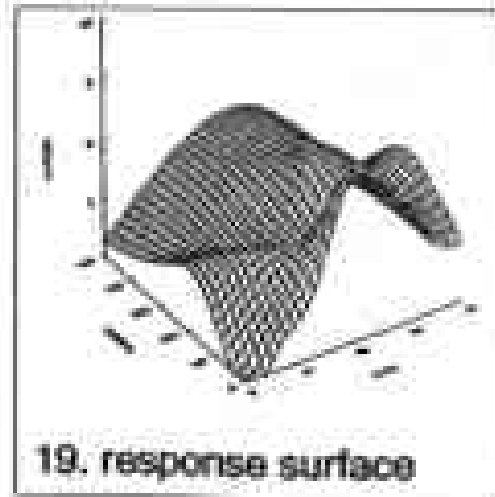
16. list of integers



chart



18. data model



19. response surface



20. wheelbarrow

Likert-style Scales Used

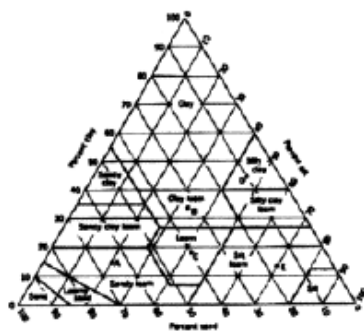
- Spatial -- nonspatial
- Hard to understand – easy to understand
- attractive -- unattractive
- conveys a lot of information – conveys a little information
- emphasizes whole – emphasizes parts
- ... (10 used total)

Some results



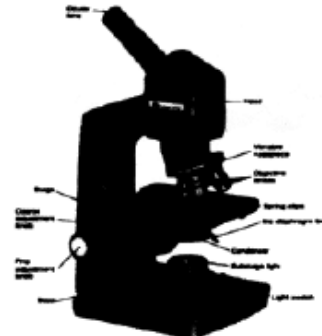
6. circular tree diagram

hard to understand 1.7



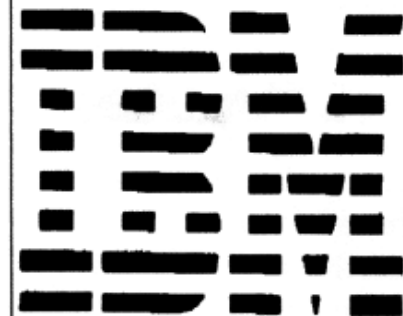
1. soil triangle

hard to understand 3.3



5. microscope

easy to understand 8.2

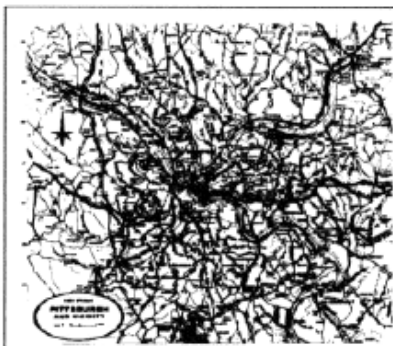


4. IBM logo

easy to understand 8.4

35. periodic table

conveys a lot 1.7



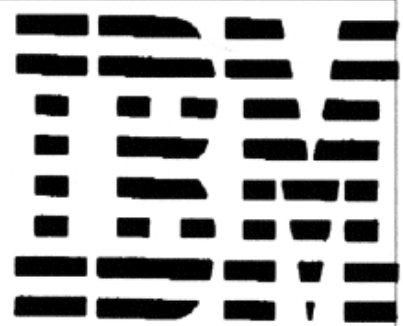
15. Pittsburgh map

conveys a lot 2.1



49. robot

conveys a little 7.3



4. IBM logo

conveys a little 7.7

Interesting Findings

Lohse et al. 94

- Photorealistic images were least informative
 - Echoes results in icon studies – better to use less complex, more schematic images
- Graphs and tables are the most self-similar categories
 - Results in the literature comparing these are inconclusive
- Temporal data more difficult to show than cyclic data

Properties of Visual Perception

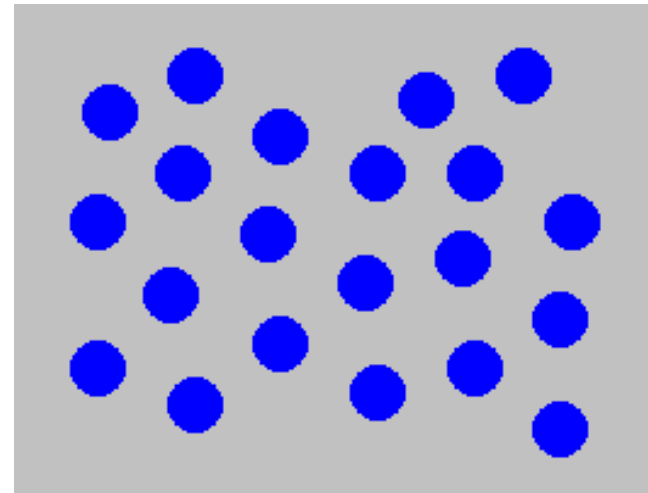
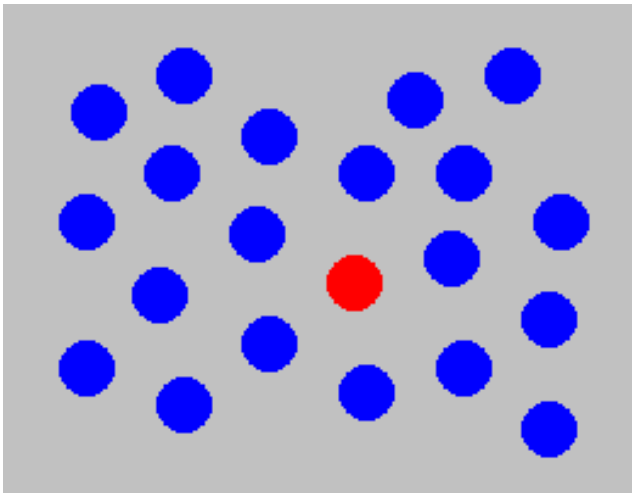
Visual Properties

- Preattentive Processing
- Accuracy of Interpretation of Visual Properties
- Illusions and the Relation to Graphical Integrity

Preattentive Processing

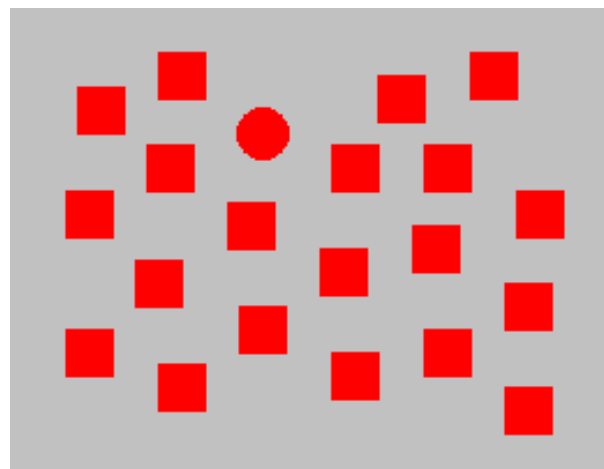
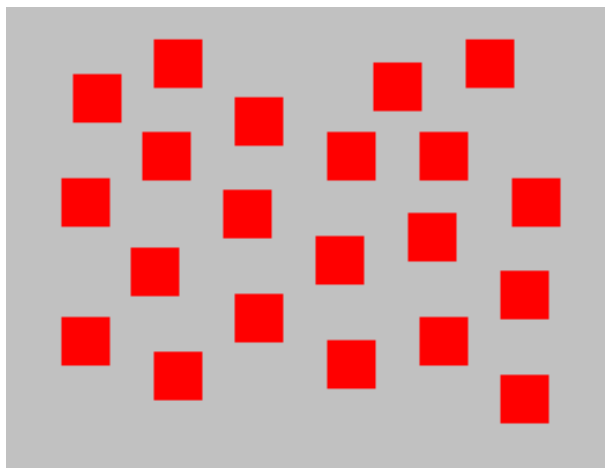
- Some properties are processed preattentively (without need for focusing attention).
- Important for design of visualizations
 - what can be perceived immediately
 - what properties are good discriminators
 - what can mislead viewers

Example: Color Selection



Viewer can rapidly and accurately determine whether the target (red circle) is present or absent. Difference detected in **color**.

Example: Shape Selection

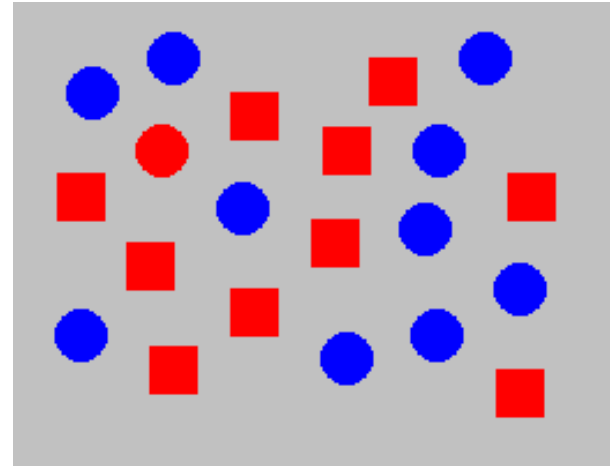
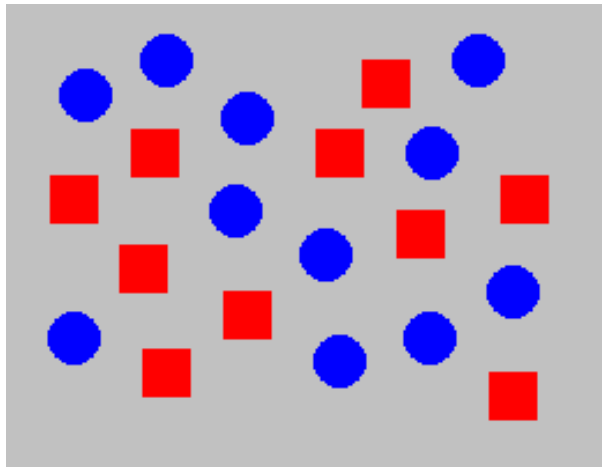


Viewer can rapidly and accurately determine whether the target (red circle) is present or absent. Difference detected in **form** (curvature)

Pre-attentive Processing

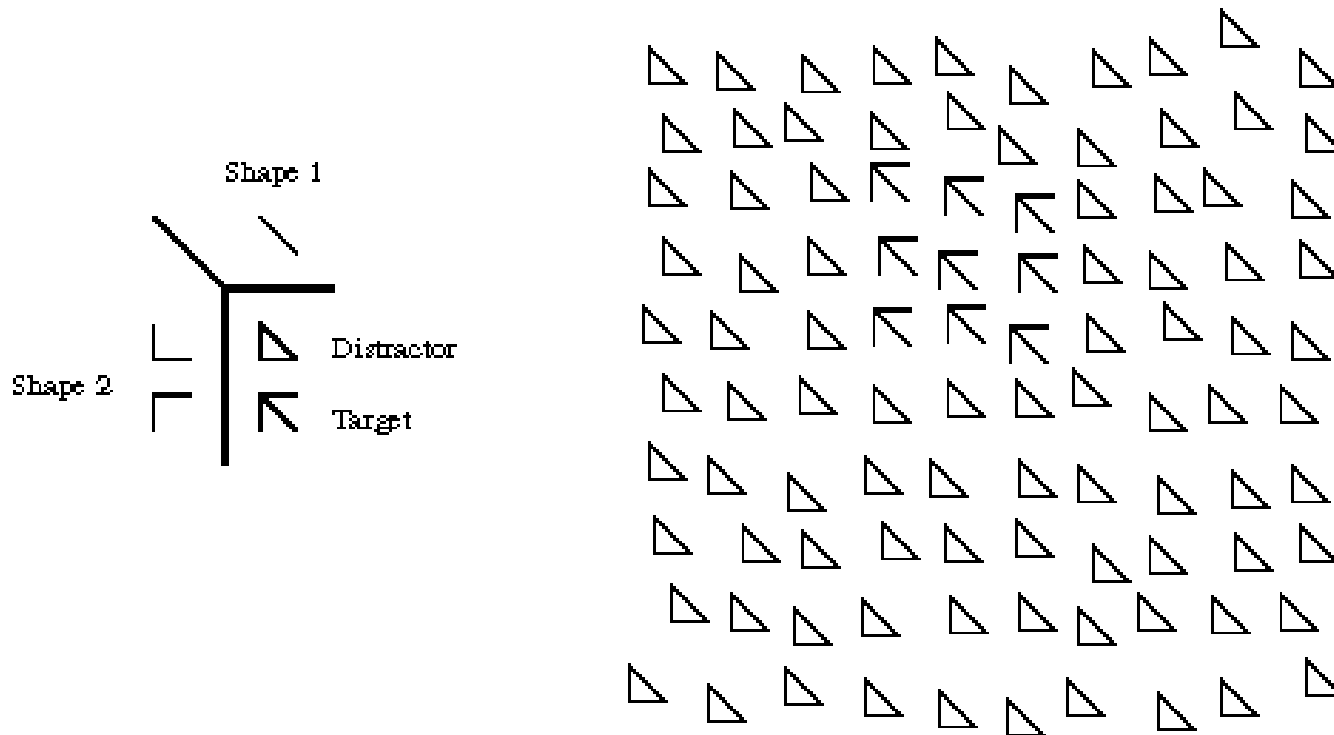
- $< 200\text{--}250$ ms qualifies as pre-attentive
 - eye movements take at least 200ms
 - yet certain processing can be done very quickly, implying low-level processing in parallel
- If a decision takes a fixed amount of time regardless of the number of distractors, it is considered to be preattentive

Example: Conjunction of Features



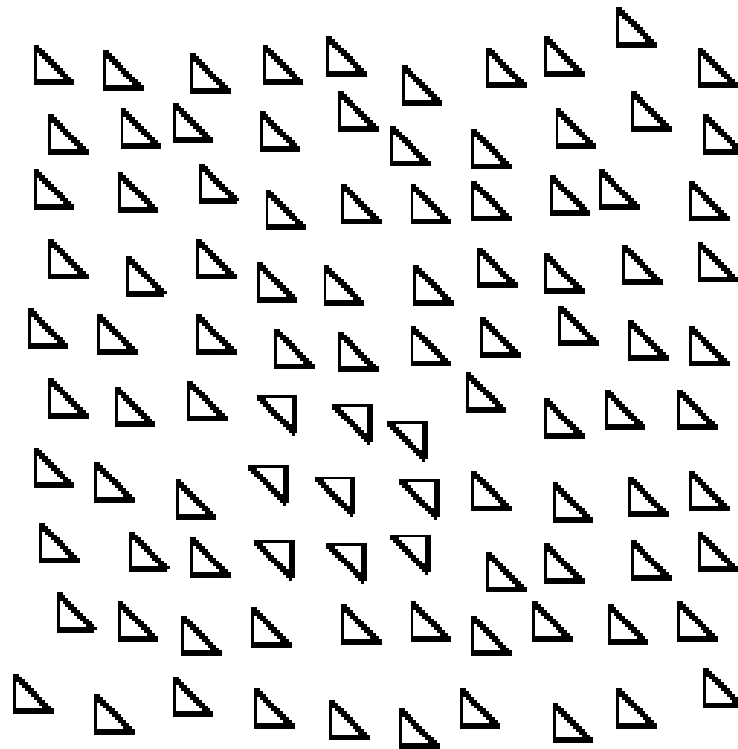
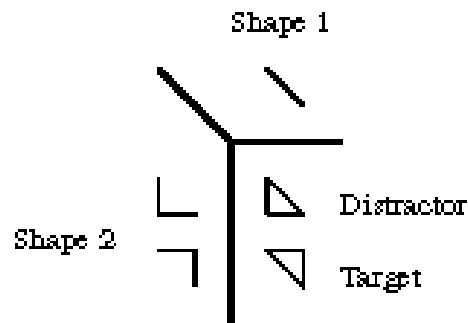
Viewer *cannot* rapidly and accurately determine whether the target (red circle) is present or absent when **target has two or more features, each of which are present in the distractors**. Viewer must search sequentially.

Example: Emergent Features



Target has a unique feature with respect to distractors (open sides) and so the group can be detected preattentively.

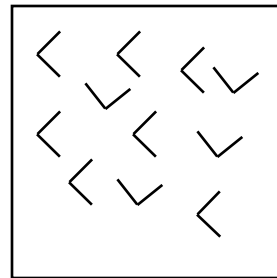
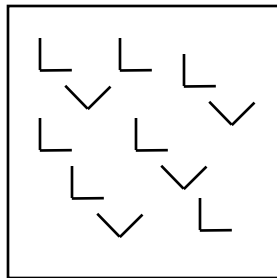
Example: Emergent Features



Target does **not** have a unique feature with respect to distractors and so the group **cannot** be detected preattentively.

Asymmetric and Graded Preattentive Properties

- Some properties are **asymmetric**
 - a sloped line among vertical lines is preattentive
 - a vertical line among sloped ones is not
- Some properties have a **gradation**
 - some more easily discriminated among than others



SUBJECT PUNCHED QUICKLY OXIDIZED TCEJBUS DEHCNUP YLKCIUQ DEZIDIXO
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC
GOVERNS PRECISE EXAMPLE MERCURY SNREVOG ESICERP ELPMAXE YRUCREM
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
GOVERNS PRECISE EXAMPLE MERCURY SNREVOG ESICERP ELPMAXE YRUCREM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC
SUBJECT PUNCHED QUICKLY OXIDIZED TCEJBUS DEHCNUP YLKCIUQ DEZIDIXO
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC

Text NOT Preattentive

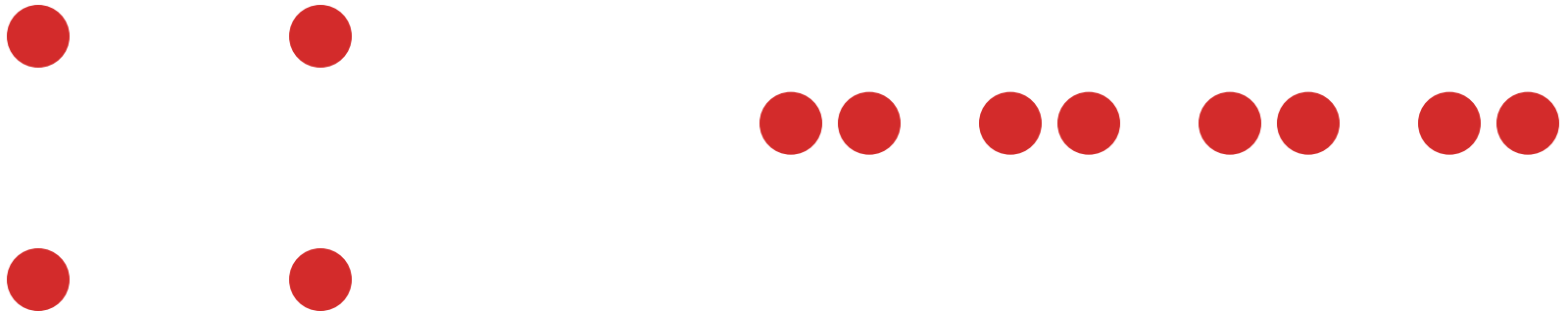
SUBJECT PUNCHED QUICKLY OXIDIZED TCEJBUS DEHCNUP YLKCIUQ DEZIDIXO
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC
GOVERNS PRECISE EXAMPLE MERCURY SNREVOG ESICERP ELPMAXE YRUCREM
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
GOVERNS PRECISE EXAMPLE MERCURY SNREVOG ESICERP ELPMAXE YRUCREM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC
SUBJECT PUNCHED QUICKLY OXIDIZED TCEJBUS DEHCNUP YLKCIUQ DEZIDIXO
CERTAIN QUICKLY PUNCHED METHODS NIATREC YLKCIUQ DEHCNUP SDOHTEM
SCIENCE ENGLISH RECORDS COLUMNS ECNEICS HSILGNE SDROCER SNMULOC

Preattentive Visual Properties [Healey 97]

length	Triesman & Gormican [1988]
width	Julesz [1985]
size	Triesman & Gelade [1980]
curvature	Triesman & Gormican [1988]
number	Julesz [1985]; Trick & Pylyshyn [1994]
terminators	Julesz & Bergen [1983]
intersection	Julesz & Bergen [1983]
closure	Enns [1986]; Triesman & Souther [1985]
colour (hue)	Nagy & Sanchez [1990, 1992]; D'Zmura [1991] Kawai et al. [1995]; Bauer et al. [1996]
intensity	Beck et al. [1983]; Triesman & Gormican [1988]
flicker	Julesz [1971]
direction of motion	Nakayama & Silverman [1986]; Driver & McLeod [1992]
binocular lustre	Wolfe & Franzel [1988]
stereoscopic depth	Nakayama & Silverman [1986]
3-D depth cues	Enns [1990]
lighting direction	Enns [1990]

Gestalt Properties

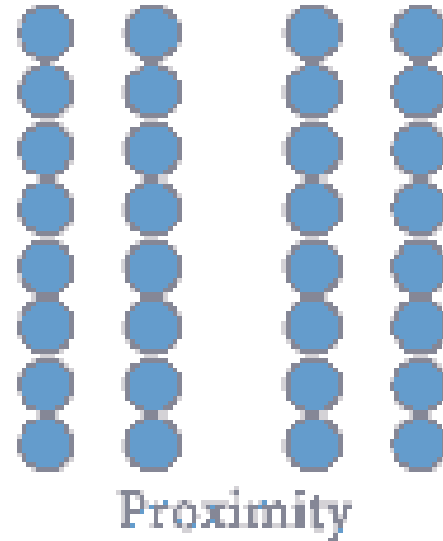
- *Gestalt*: essence or shape of an entity's complete form
- Idea: forms or patterns transcend the stimuli used to create them
 - Why do patterns emerge? Under what circumstances?
- Visual system uses **Proximity, similarity, closure, symmetry, common fate, continuity**



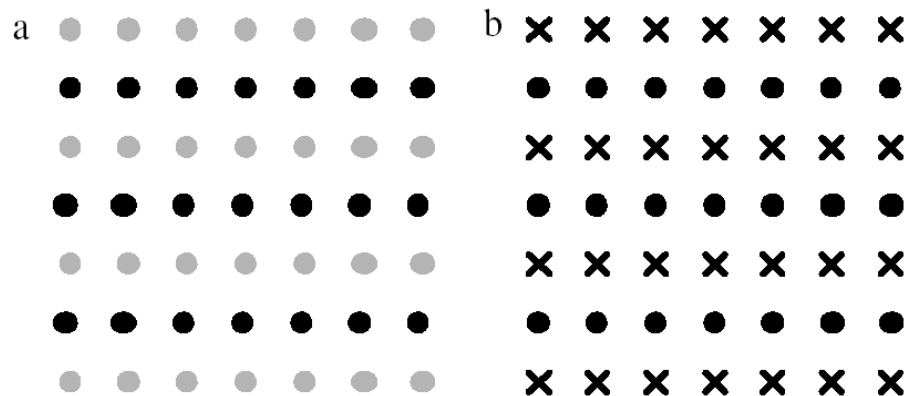
Why perceive pairs vs. triplets?

More Gestalt Laws

- Law of Proximity
 - Stimulus elements that are close together will be perceived as a group

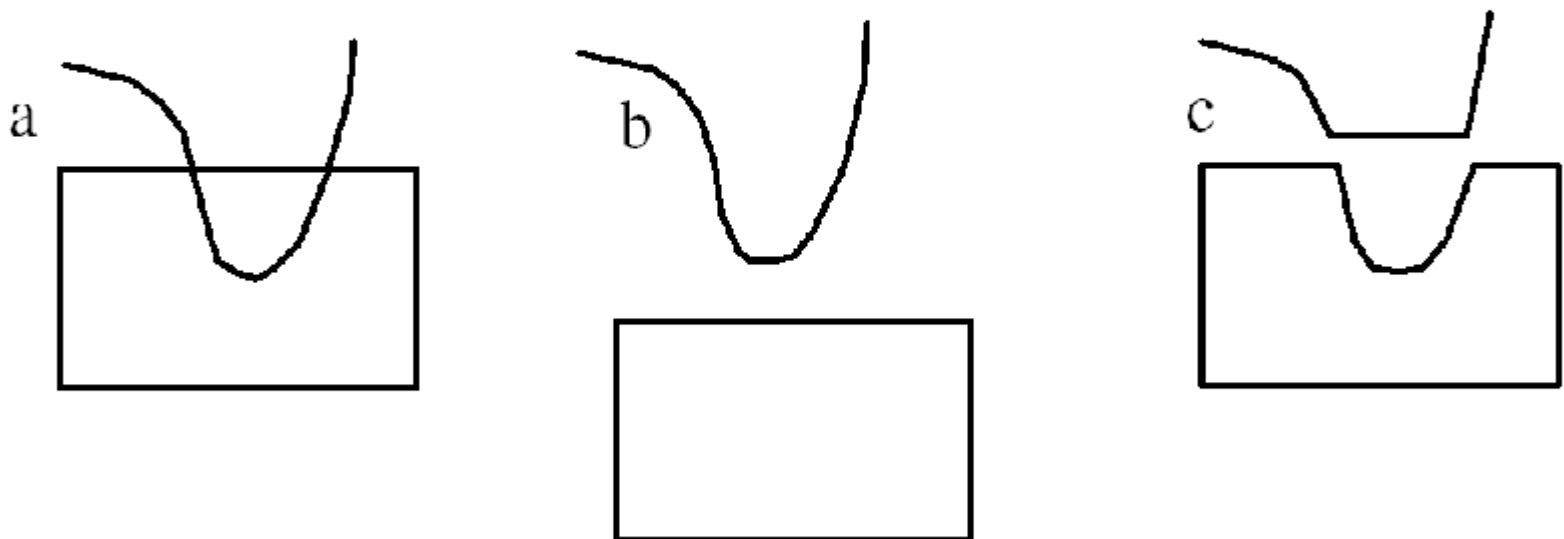


- Law of Similarity



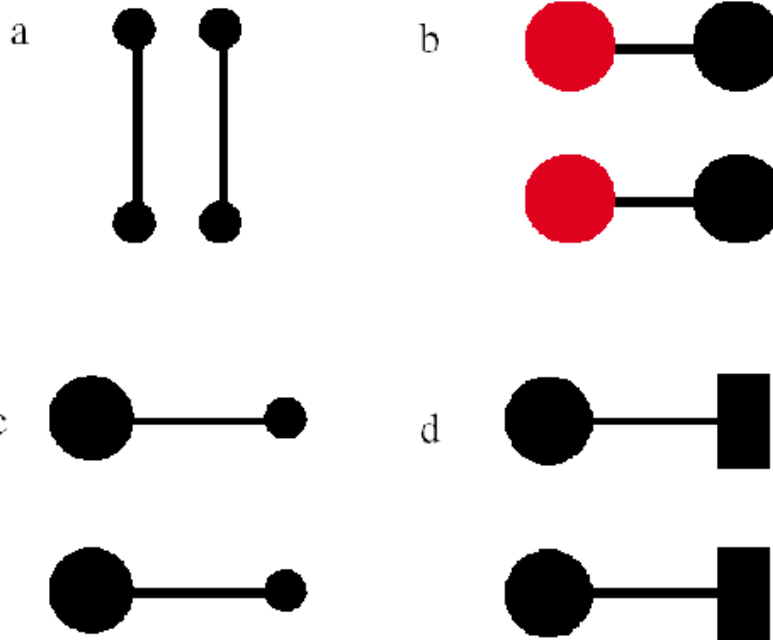
Continuity

smooth not abrupt change
overrules proximity



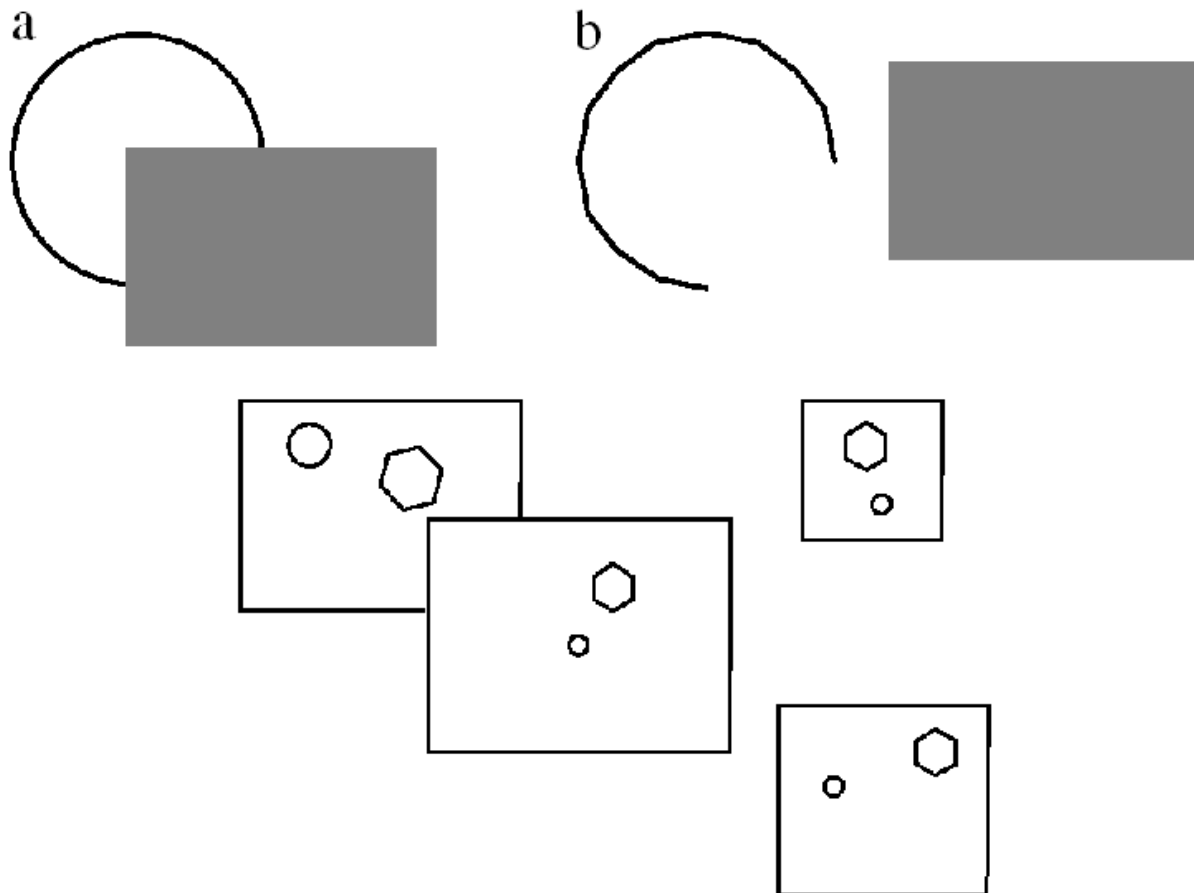
Connectedness

can overrule size, shape



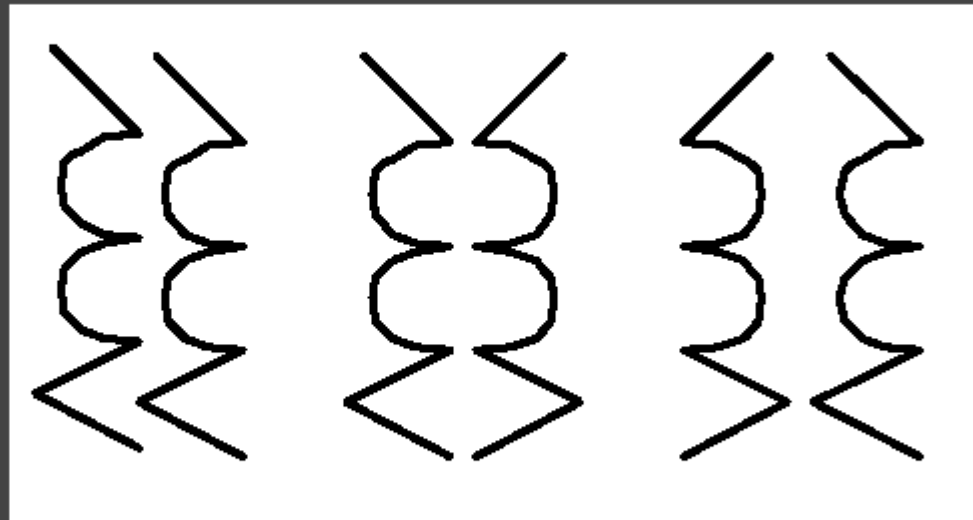
Closure

overrules proximity



Symmetry

emphasizes relationships

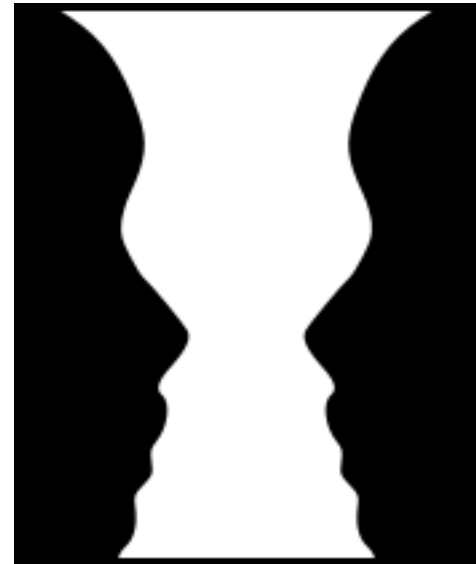


More Gestalt Laws

- Law of Common Fate
 - move a subset of objects among similar ones and they will be perceived as a group
 - example: <http://www.andyrutledge.com/common-fate.php>
- More info:
http://en.wikipedia.org/wiki/Principles_of_grouping

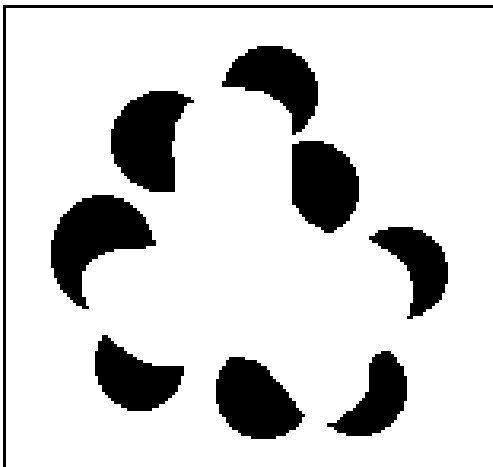
Gestalt Laws of Perceptual Organization [Kaufman 74]

- Figure and Ground
 - We take elements that are similar in appearance/shape and group them together as a whole
 - Contrast is a natural means to perceive difference/similarity
 - Illusions exploit ambiguity in figure/ground relationships:
multistability

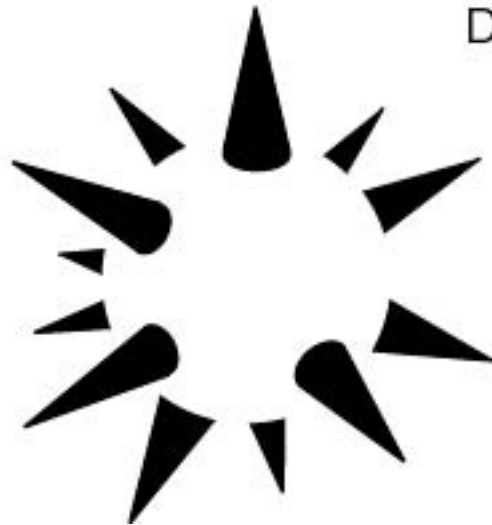


Gestalt Laws of Perceptual Organization [Kaufman 74]

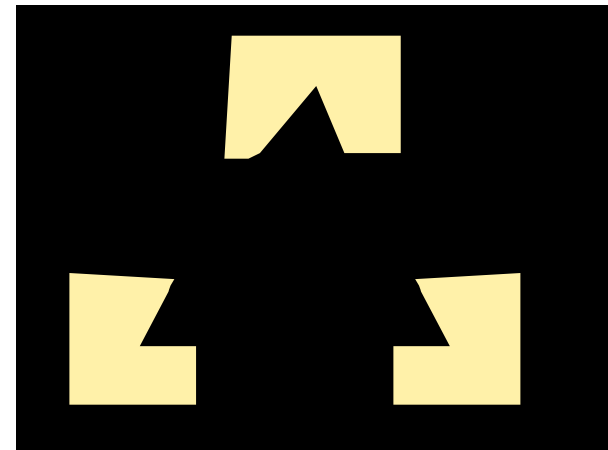
- Reification: Experienced percept contains more explicit spatial information than the sensory stimulus itself:



C



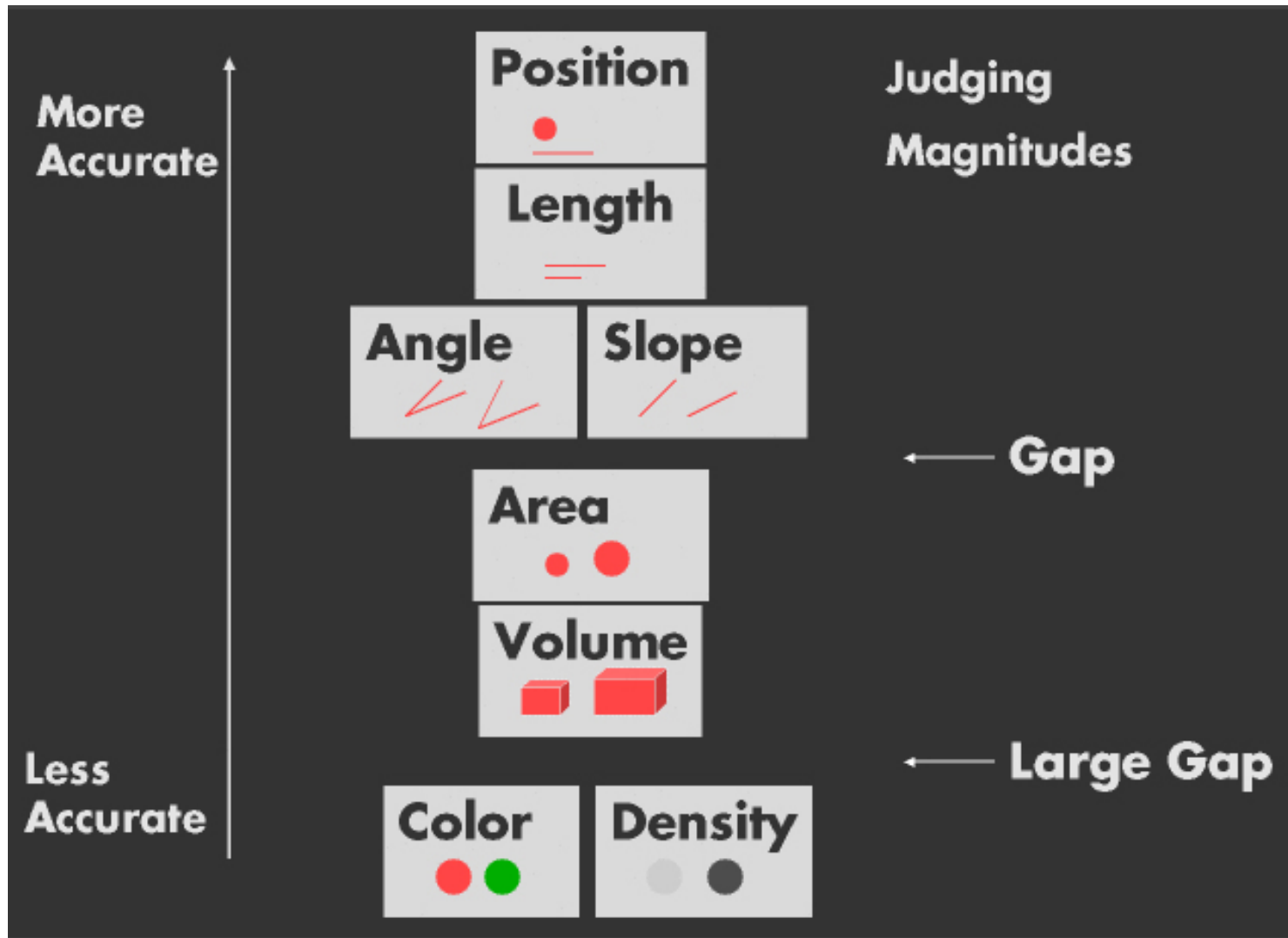
D



Which Visual Properties are Appropriate for Which Information Types?

Accuracy Ranking of Quantitative Perceptual Tasks

Estimated; only pairwise comparisons have been validated
[Mackinlay 88 from Cleveland & McGill]



Interpretations of Visual Properties

Some properties discriminated more accurately but have no intrinsic meaning [Senay & Ingatious 97, Kosslyn, others]

- Density (Greyscale)
Darker → More
- Size / Length / Area
Larger → More
- Position
Leftmost → first, Topmost → first
- Hue
??? no intrinsic meaning
- Slope
??? no intrinsic meaning

Ranking of Applicability of Properties for Different Data Types

[Mackinlay 88, Not Empirically Verified]

Quantitative	Ordinal	Nominal
Position	Position	Position
Length	Density	Color Hue
Angle	Color Saturation	Texture
Slope	Color Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Color Saturation
Color Saturation	Length	Shape
Color Hue	Angle	Length

Visual Illusions

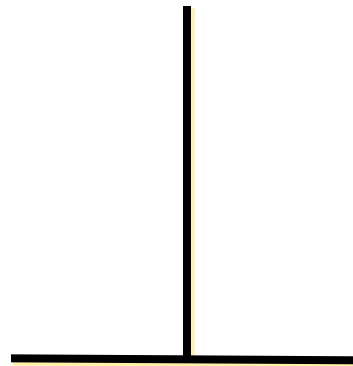
- People don't perceive length, area, angle, brightness the way they "should"
- Some illusions have been reclassified as systematic **perceptual errors**
 - e.g., brightness contrasts (grey square on white background vs. on black background)
 - partly due to increase in our understanding of the relevant parts of the visual system
- Nevertheless, the visual system does some really unexpected things

Illusions of Linear Extent

- Mueller-Lyon (off by 25-30%)



- Horizontal-Vertical

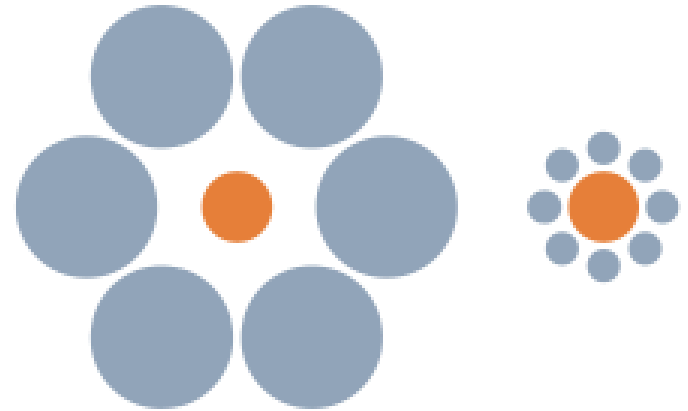


Illusions of Area

- Delboeuf Illusion



- Ebbinghaus illusion



- Height of 4-story building overestimated by approximately 25%
 - Worse for real objects!

Tufte's Principles

Tufte's Principles of Graphical Excellence

“Graphical excellence is the well-designed presentation of interesting data – a matter of *substance*, of *statistics*, and of *design*.”

- consists of complex ideas communicated with clarity, precision and efficiency
- is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
- requires telling the truth about the data

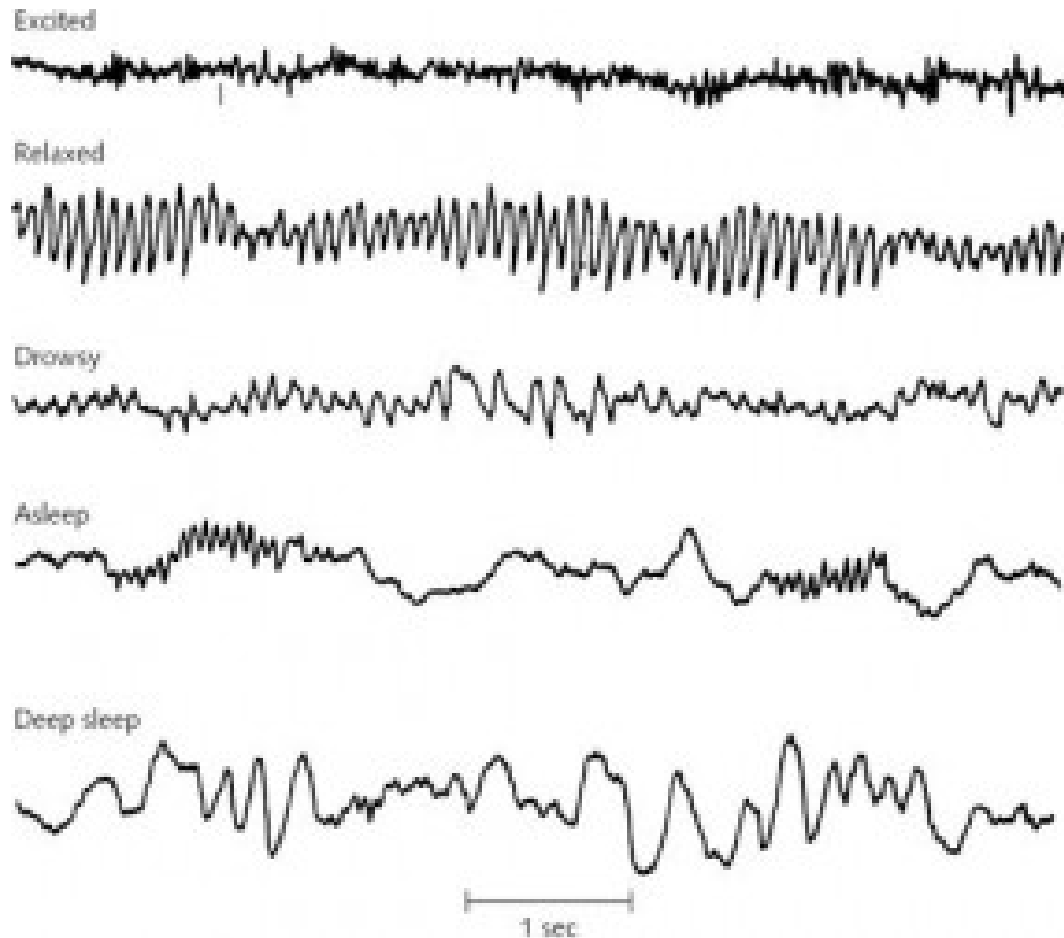
Tufte's Principles of Graphical Excellence

- Data Ink, multifunctioning graphical elements, high data density and small multiples, aesthetics

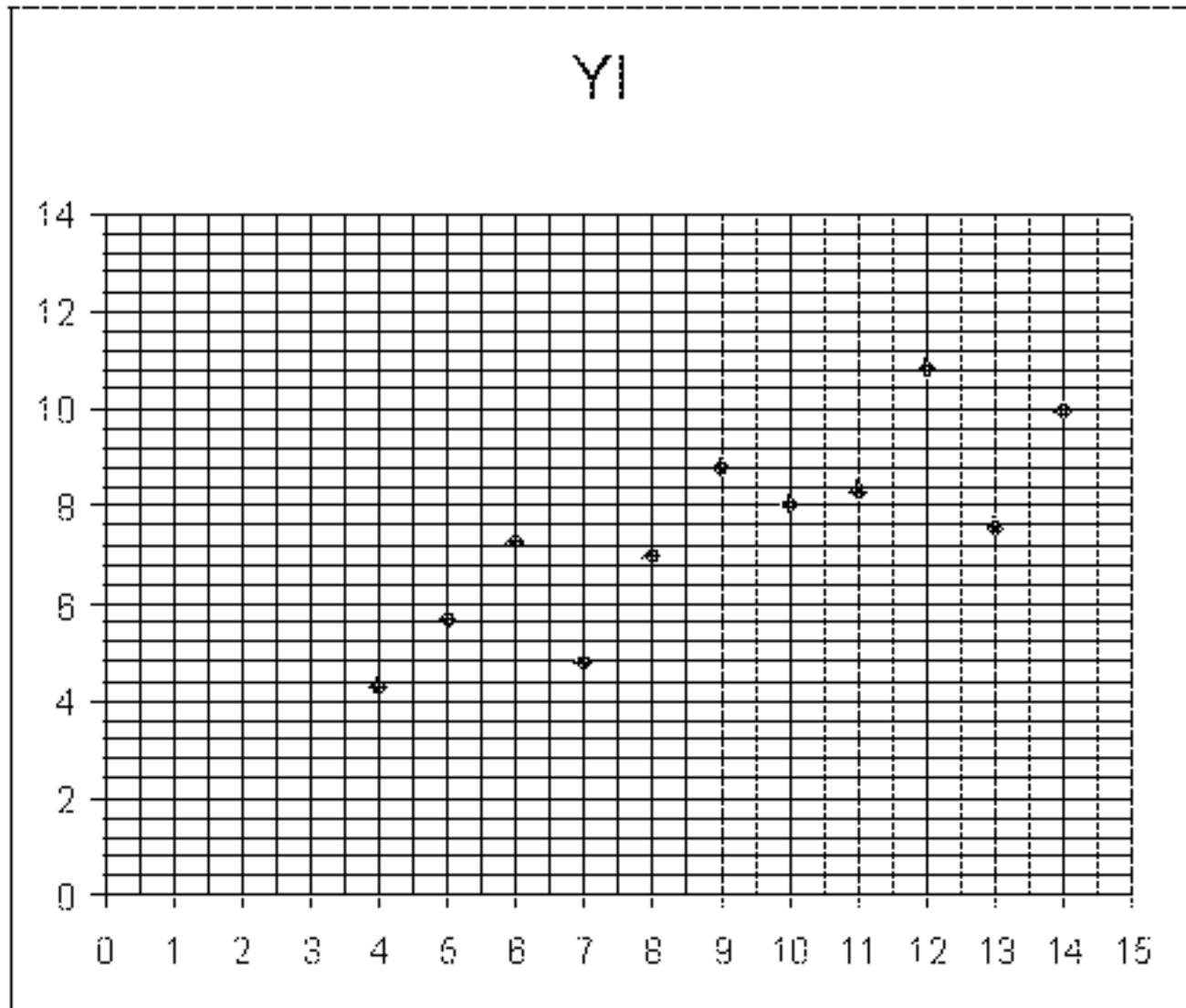
Data Ink

- **Above all else, show the data.**
- “A large share of ink on a graphic should represent data-information, the ink changing as the data change. Data-ink is the non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers represented.”

A high data-ink ratio



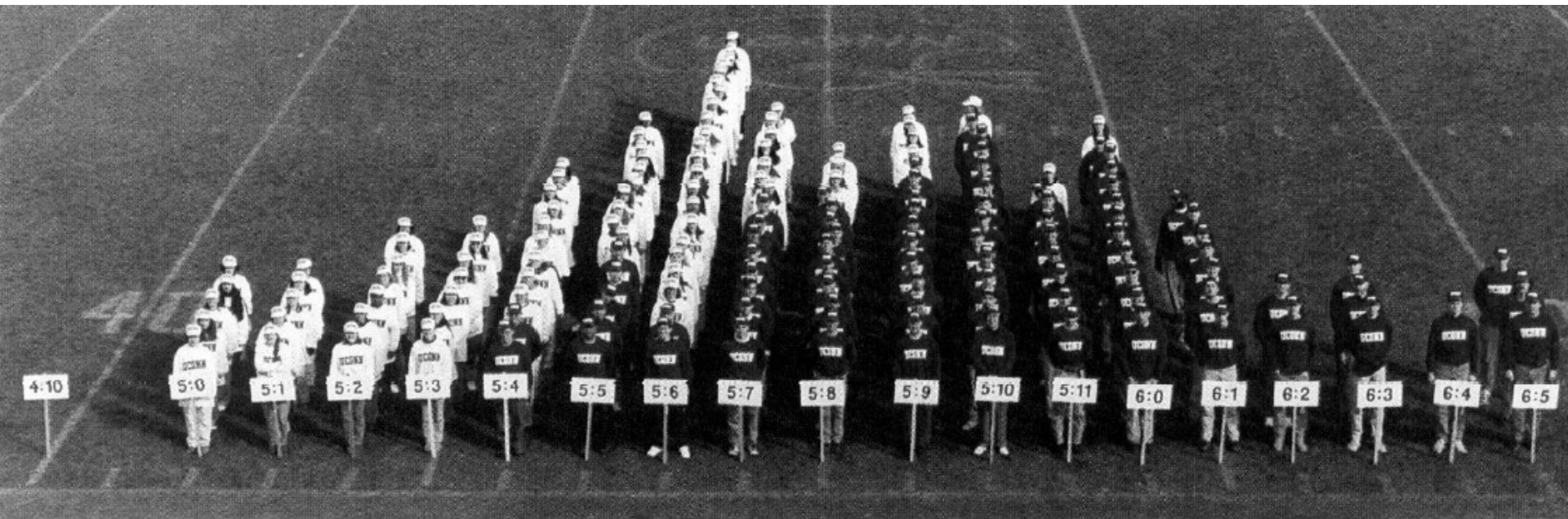
A low data-ink ratio



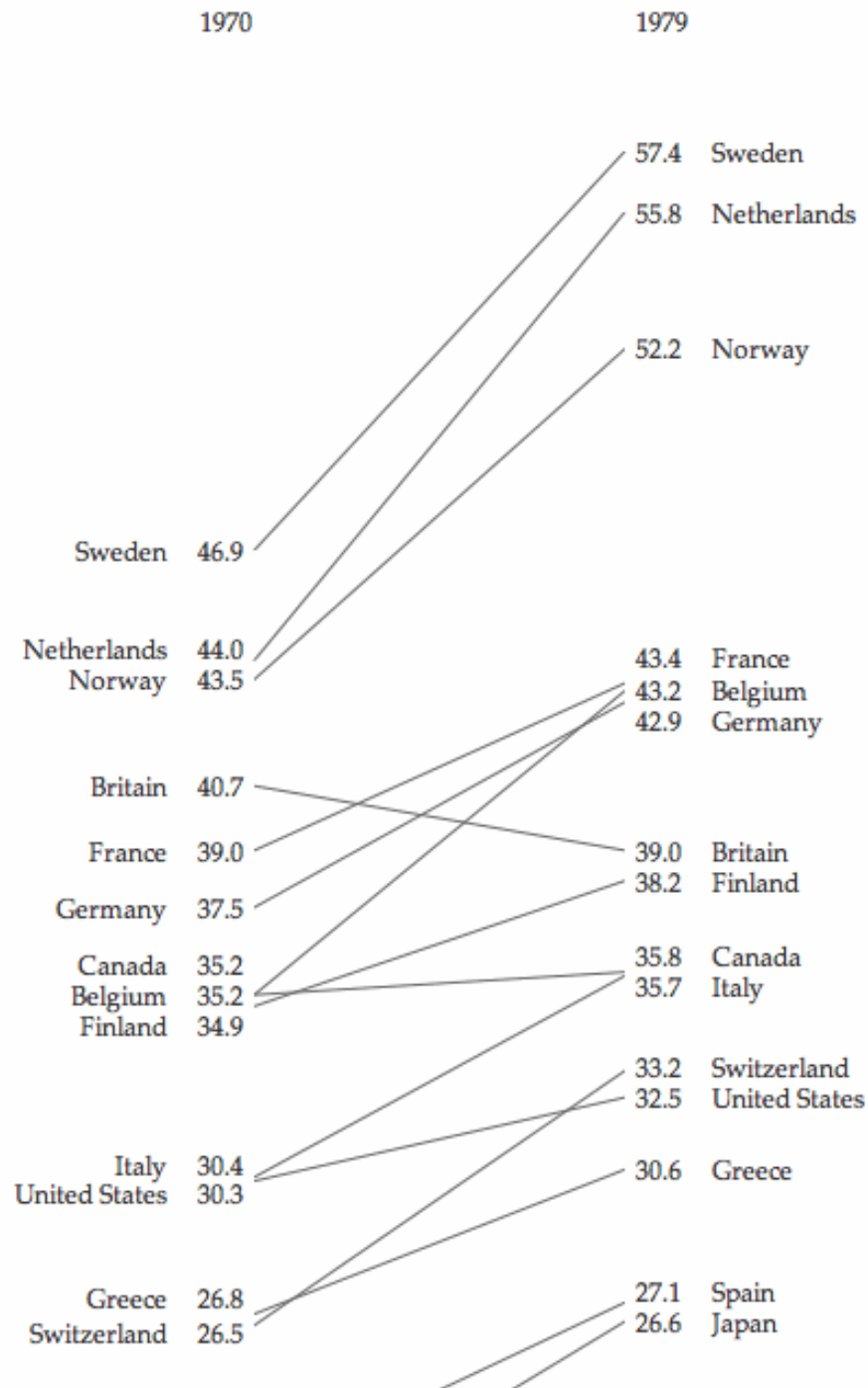
Multifunctioning Graphical Elements

- **Mobilize every graphical element, perhaps several times over, to show the data.**
- “The same ink should often serve more than one graphical purpose. A graphical element may carry data information and also perform a design function usually left to non-data ink. Or it might show several different pieces of data.”

Data-built Data Measures



Government tax collections

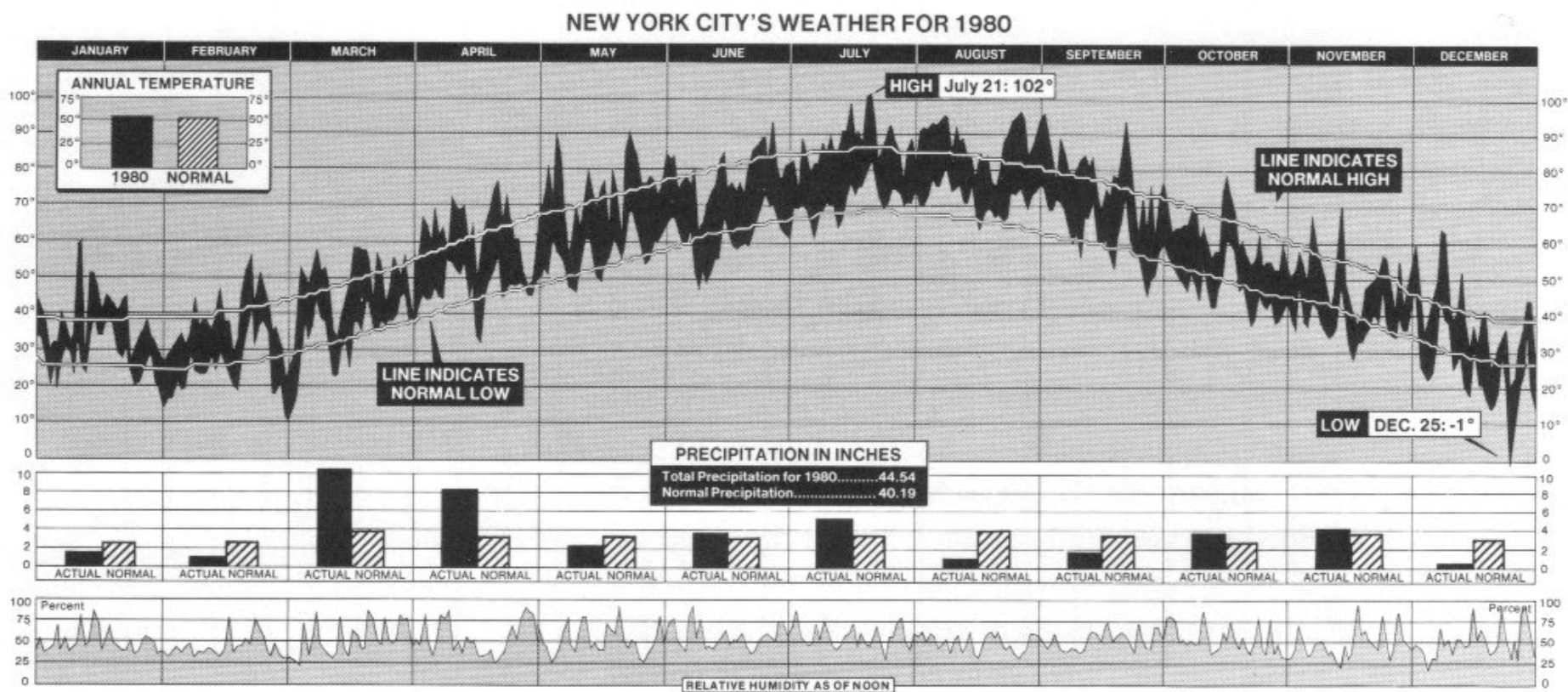


High Data Density and Small Multiples

- Take advantage of eye's ability to detect large amounts of information in small spaces
- Data density = (# entries in data matrix) / area of data graphic

Data-dense

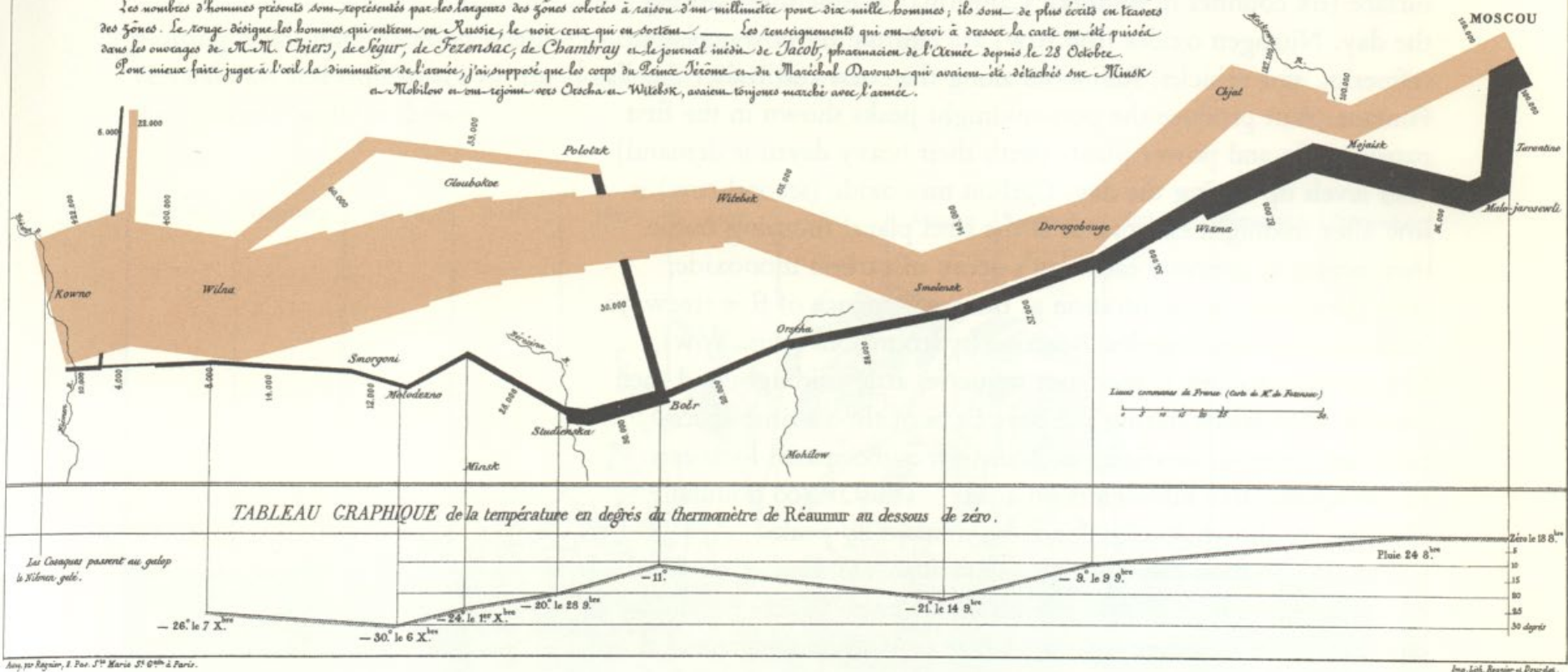
- 181 numbers per square inch:



Napoleon's 1812 March by Charles Joseph Minard

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.
Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilew et qui rejoignent vers Orescha et Witebsk, avaient toujours marché avec l'armée.

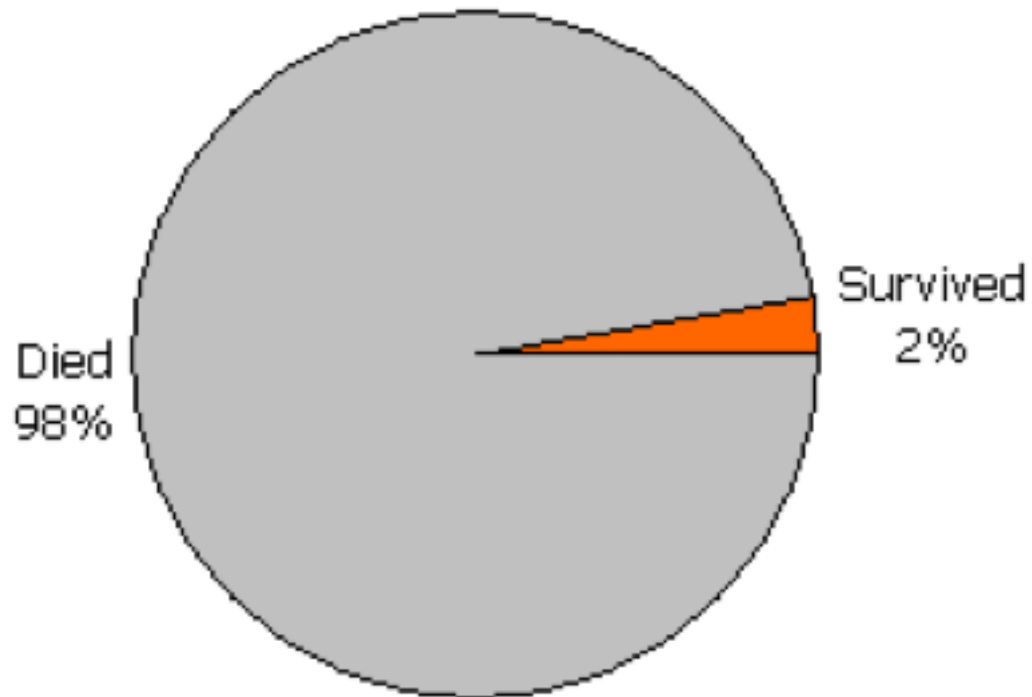


Variables shown:

- size of army
- direction
- latitude
- longitude
- temperature
- date

Executive Summary

Napoleon's Russian Campaign



Small Multiples

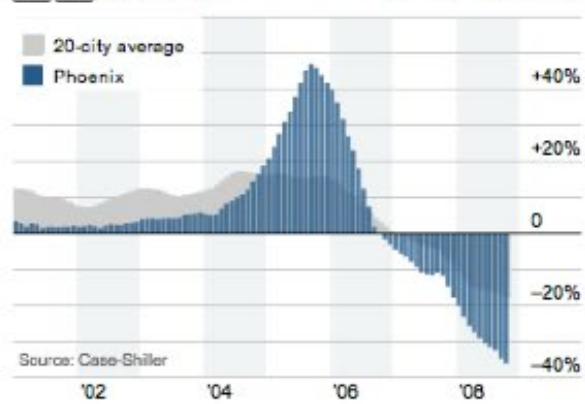
- “At the heart of quantitative reasoning is a single question: **Compared to what?** Small multiple designs, multivariate and data bountiful, answer directly by visually enforcing comparisons of changes, of the differences among objects, of the scope of alternatives. For a wide range of problems in data presentation, small multiples are the best design solution.”

Small Multiples

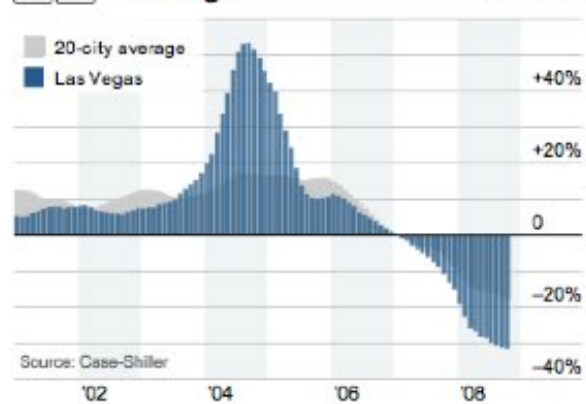
Change in Home Prices (year over year)

From New York Times

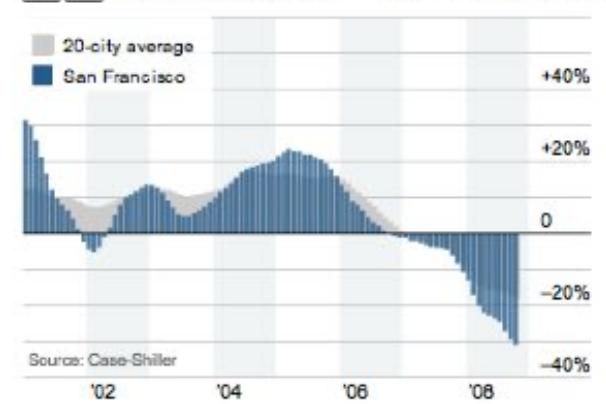
Phoenix OCT. '08: -36.2%



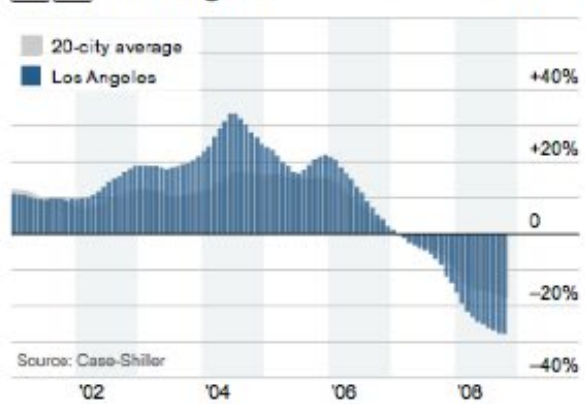
Las Vegas OCT. '08: -31.7%



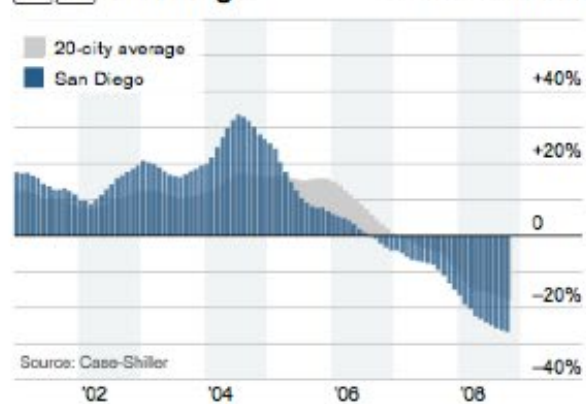
San Francisco OCT. '08: -31.0%



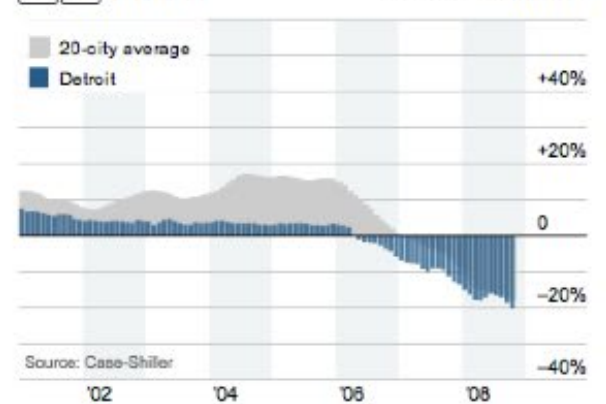
Los Angeles OCT. '08: -27.9%



San Diego OCT. '08: -26.7%

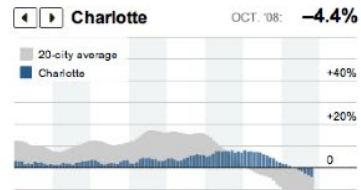
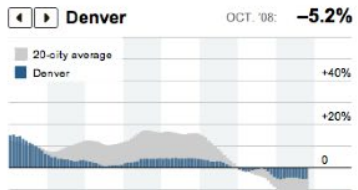
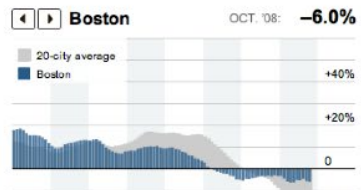
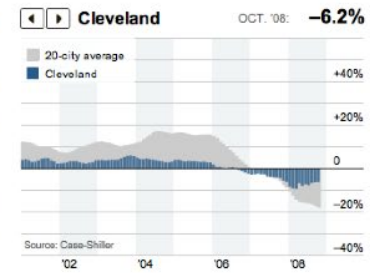
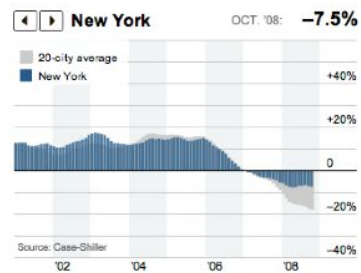
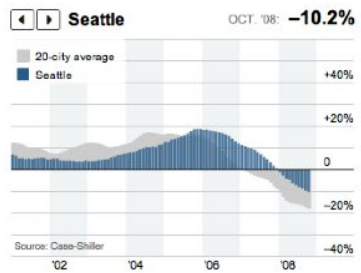
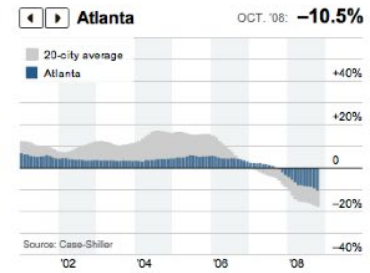
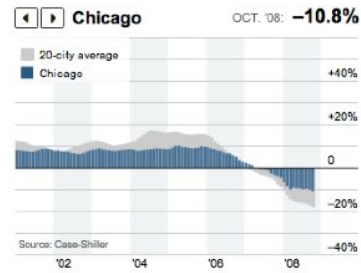
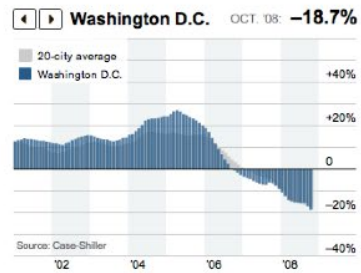
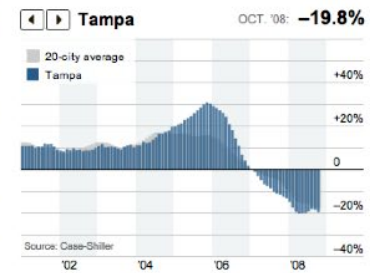
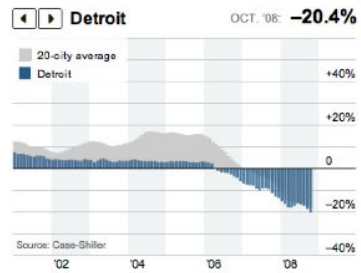
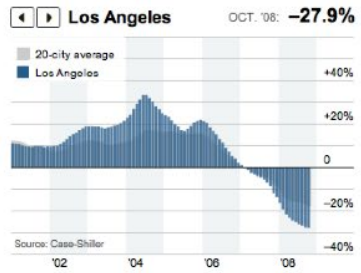
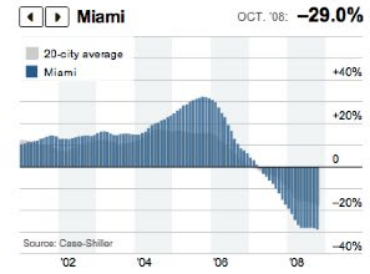
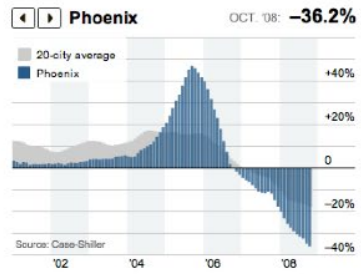


Detroit OCT. '08: -20.4%



Change in Home Prices (year over year)

From New York Times Economix blog





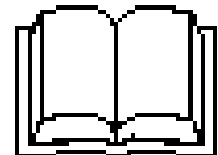
Aesthetics

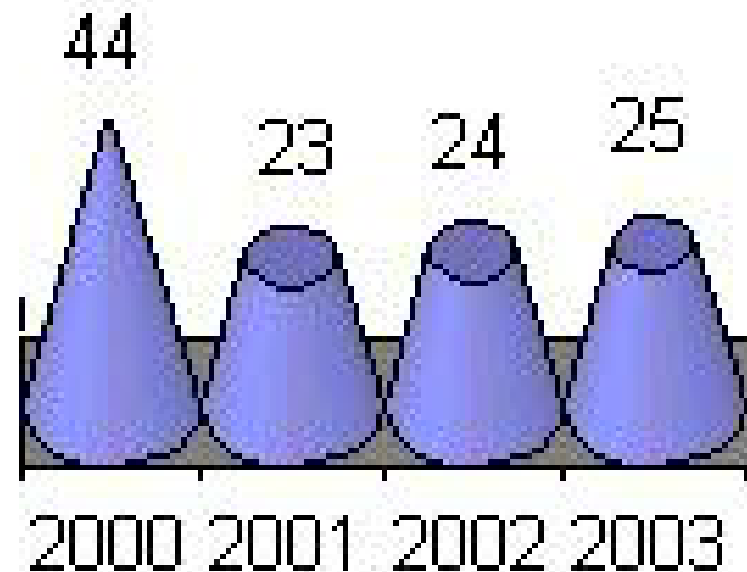
- “Graphical elegance is often found in simplicity of design and complexity of data.”
 - Sometimes tables are simplest and most appropriate representation
- Graphics should be wider than they are tall
 - Eye is good at detecting deviations from the horizon

Tufte's Take on Graphical Integrity

- Some lapses intentional, some not
- Lie Factor =
$$\frac{\text{size of effect in graph}}{\text{size of effect in data}}$$
- Visualizations mislead using area
- Visualizations mislead using perspective
- Visualizations mislead by leaving out important context
- Lack of taste and aesthetics

How to Lie With Visualizations

Year	Books circulated
2001	100 
2002	141 
2003	200 

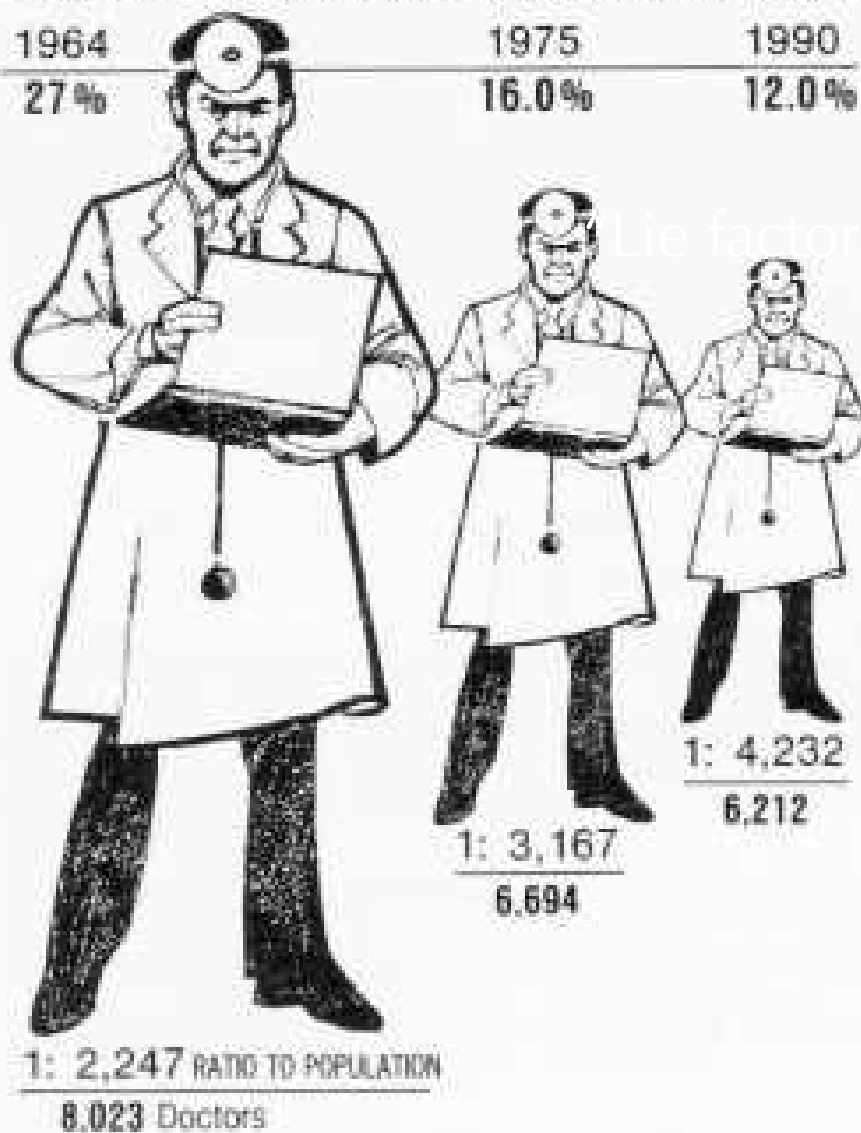


THE SHRINKING FAMILY DOCTOR

In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27 %	16.0 %	12.0 %





1958 — EISENHOWER: \$1.00



1963 — KENNEDY: 94c



1968 — JOHNSON: 83c



1973 — NIXON: 64c



1978 — CARTER: 44c

Purchasing Power of the Diminishing Dollar

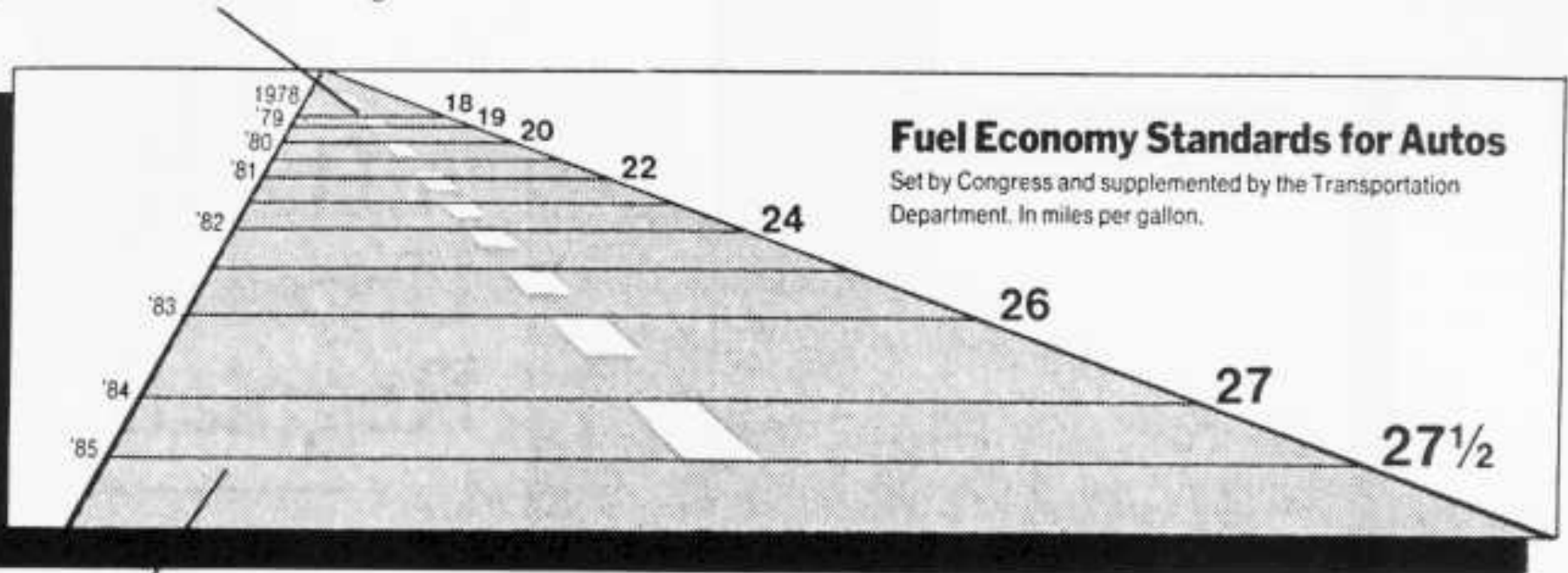
Source: Labor Department

Washington Post, October 25, 1978, p. 1.

dimensions

How to Lie With Visualizations

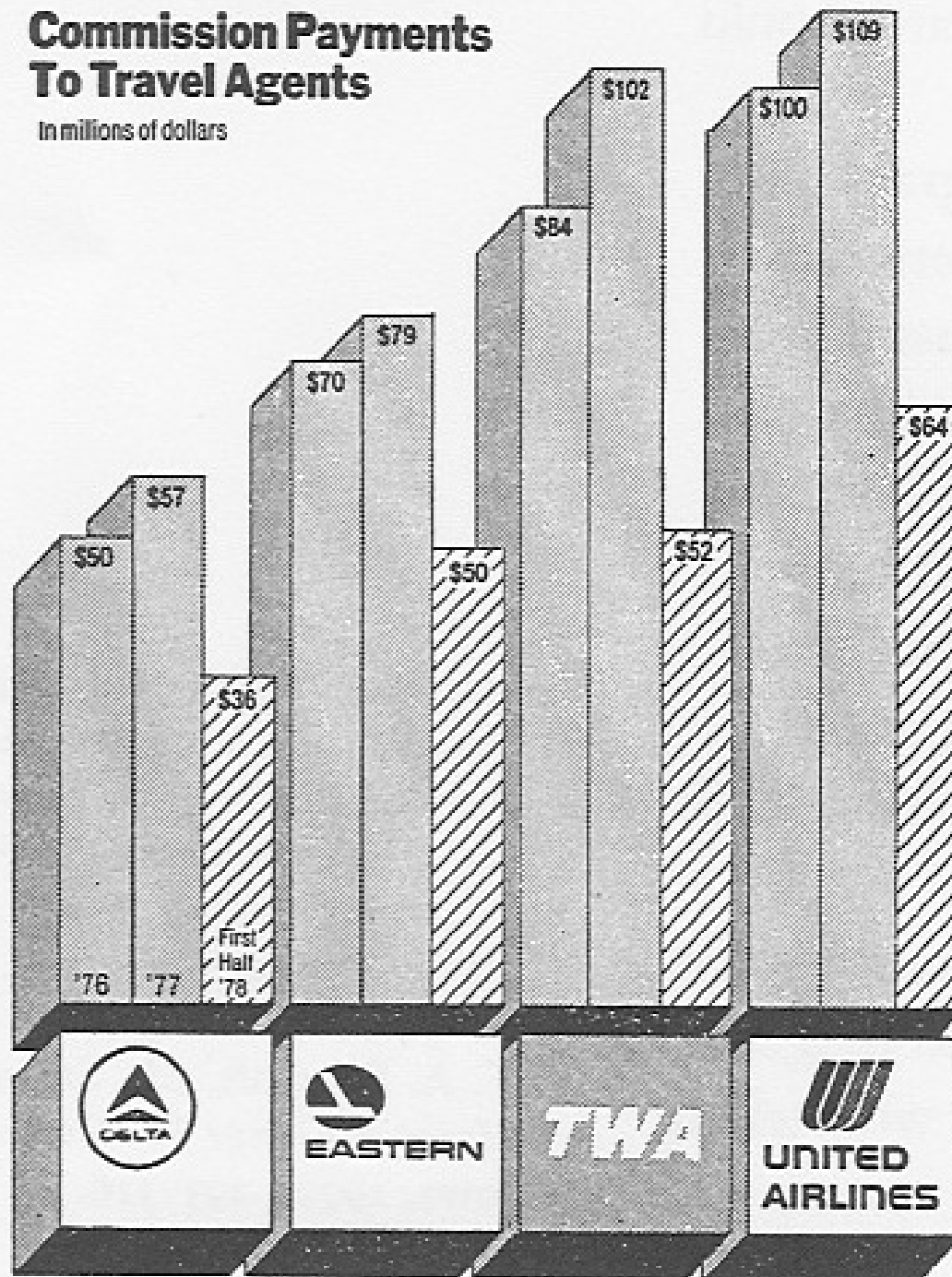
This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



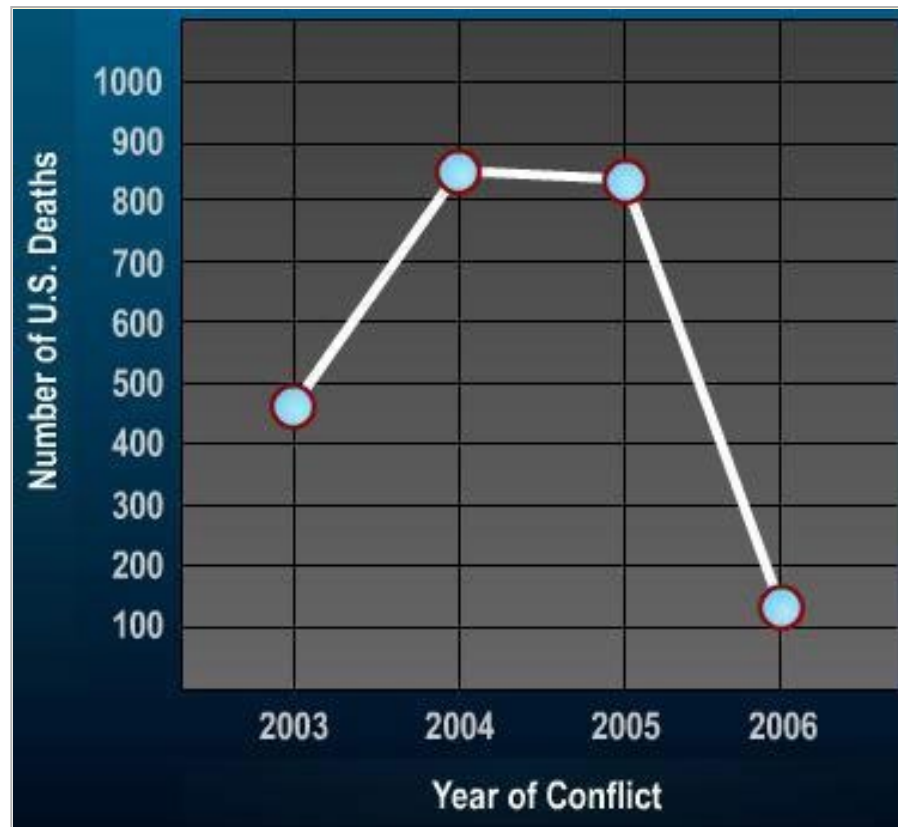
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

Commission Payments To Travel Agents

In millions of dollars

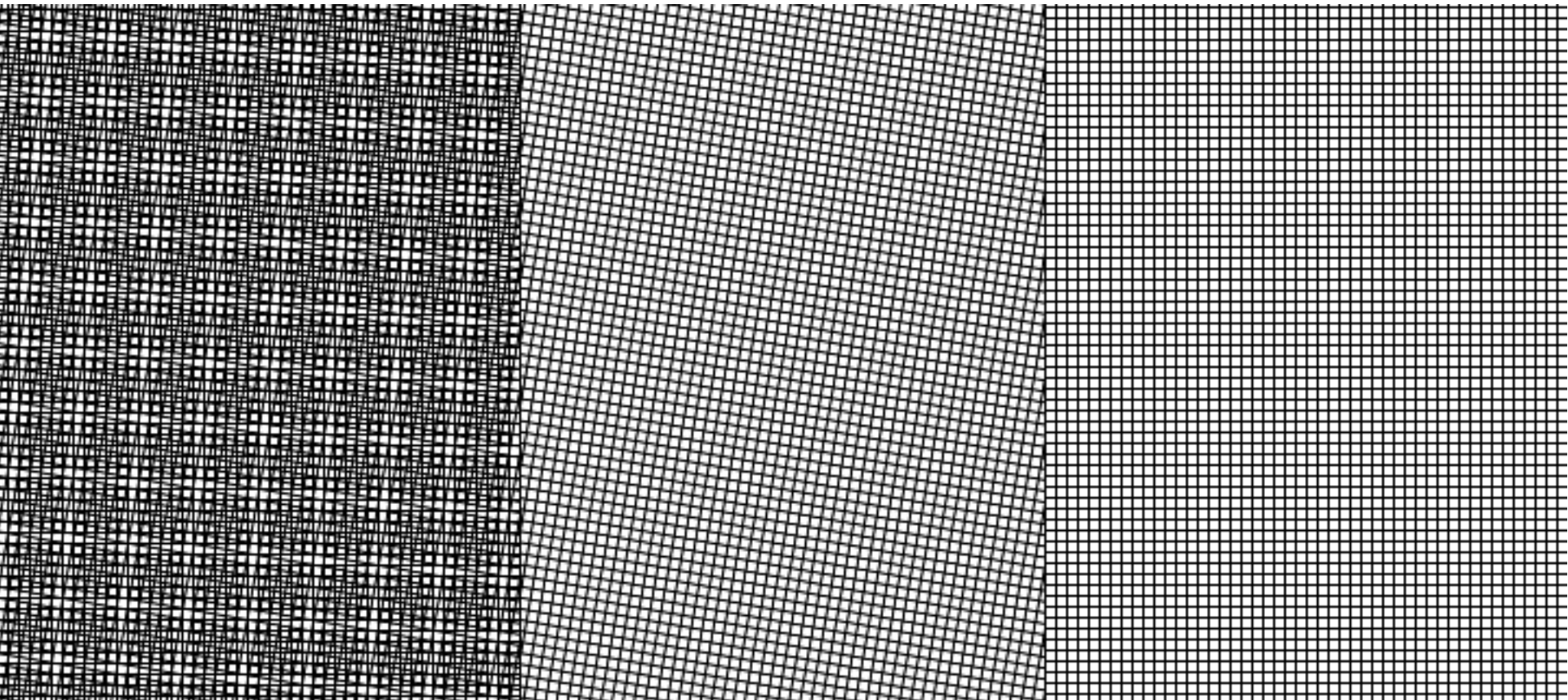


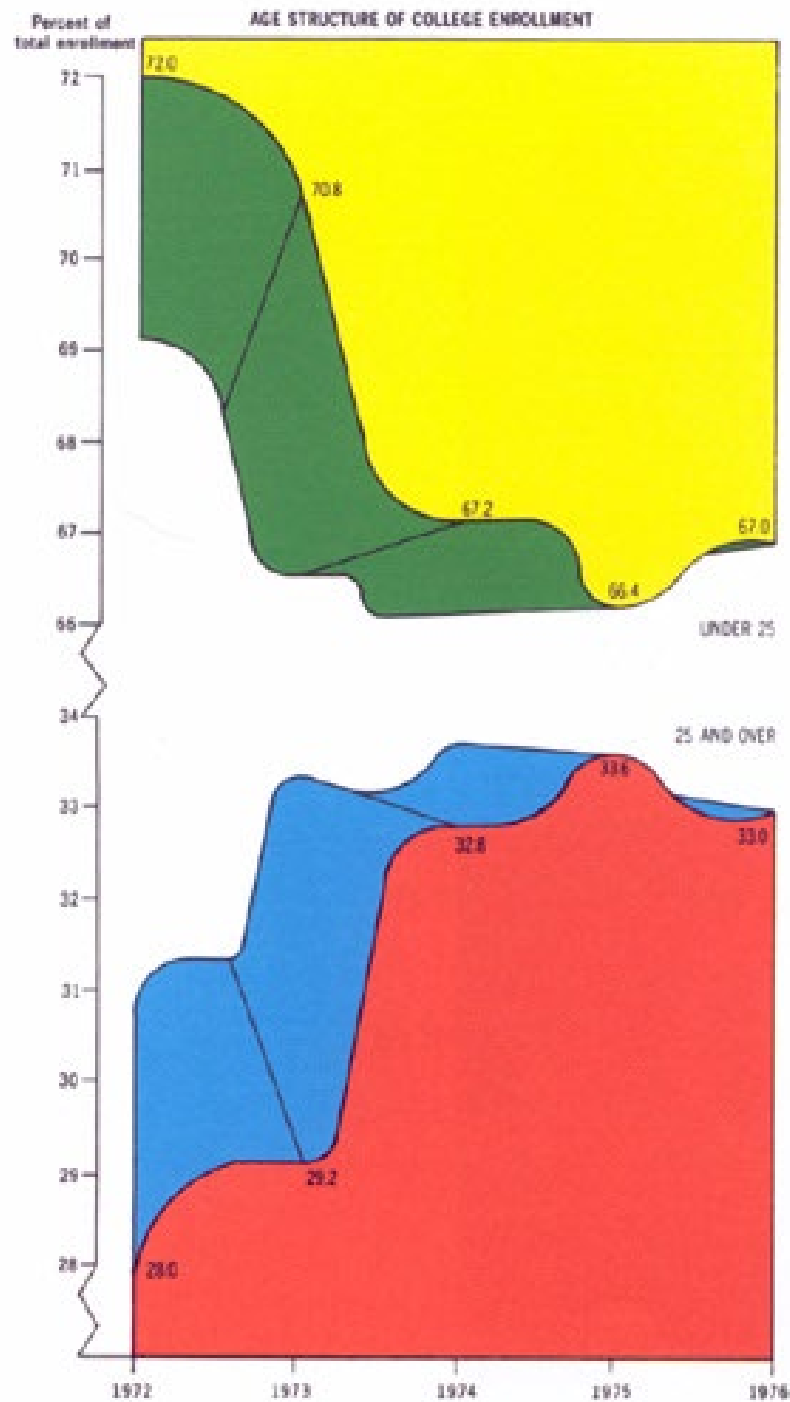
Recent Example



Data Ink and “Chart Junk”

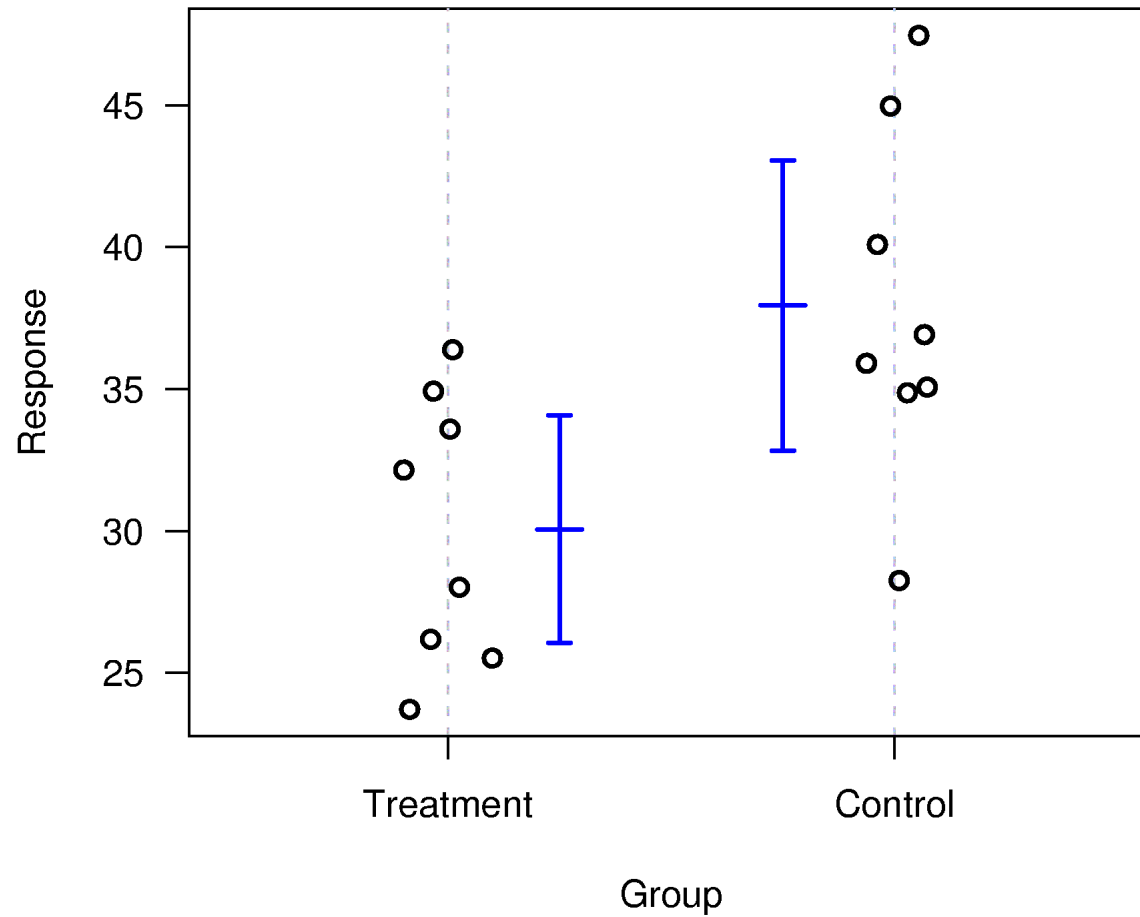
- Junk can mislead, confuse, obscure



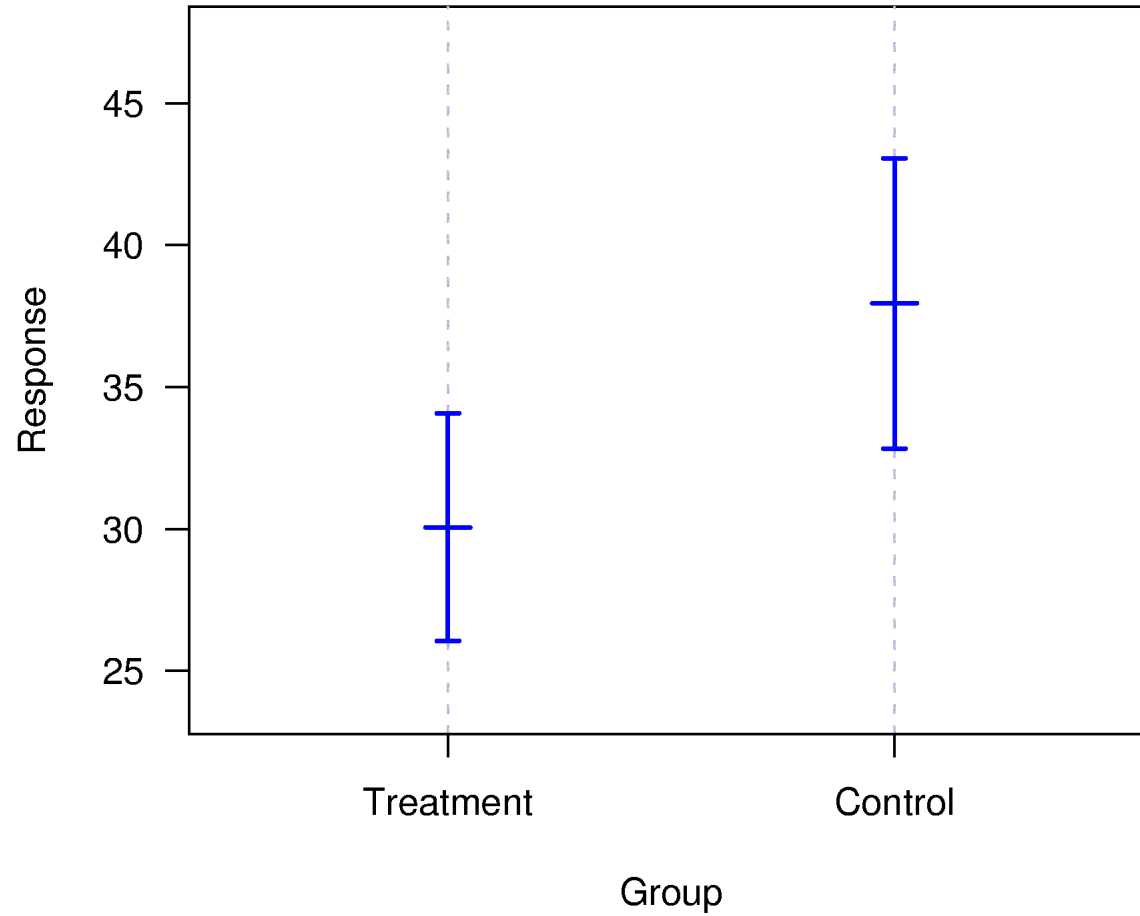


Wrap-up: Excel Charts and You

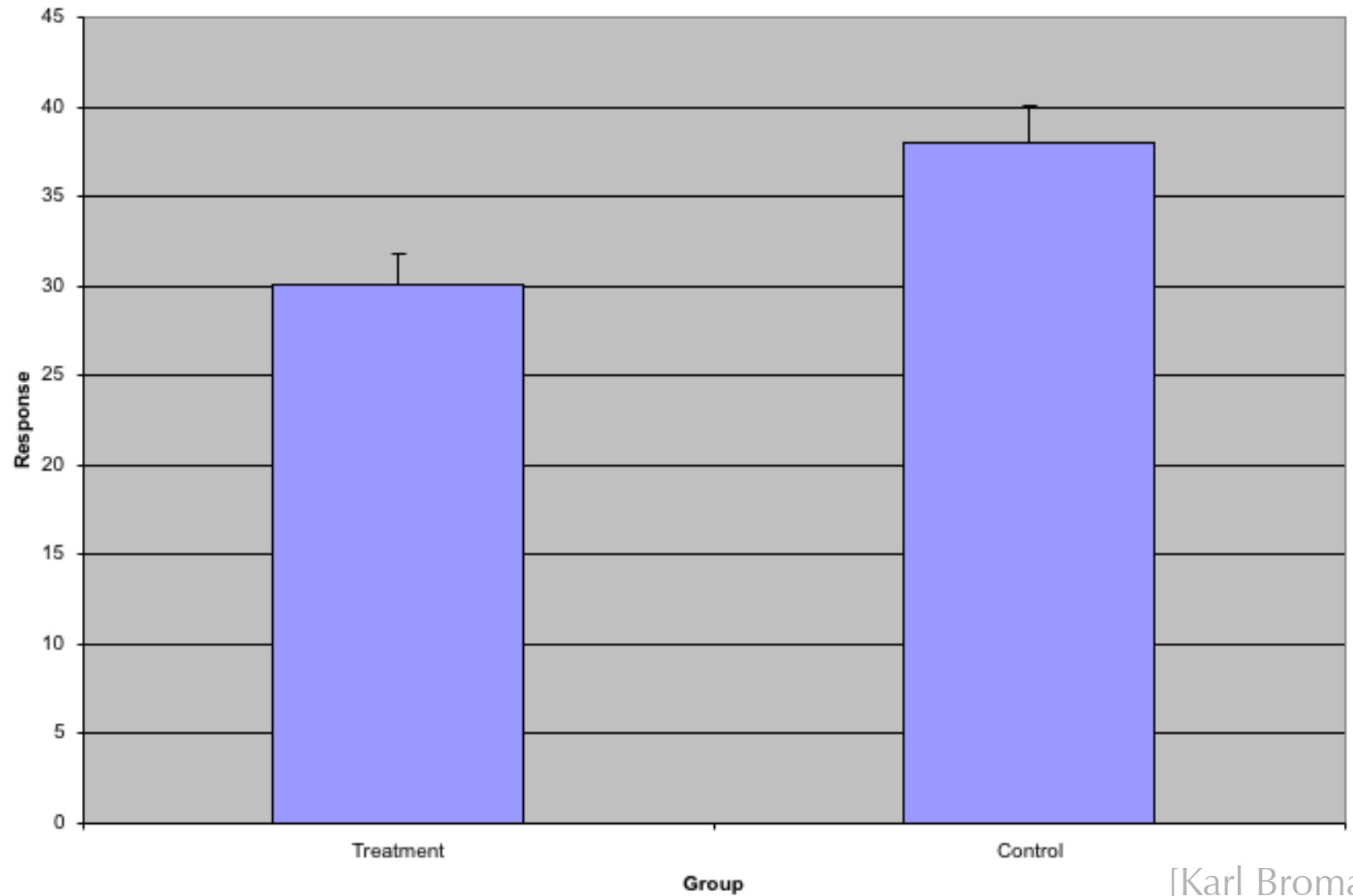
Example 1



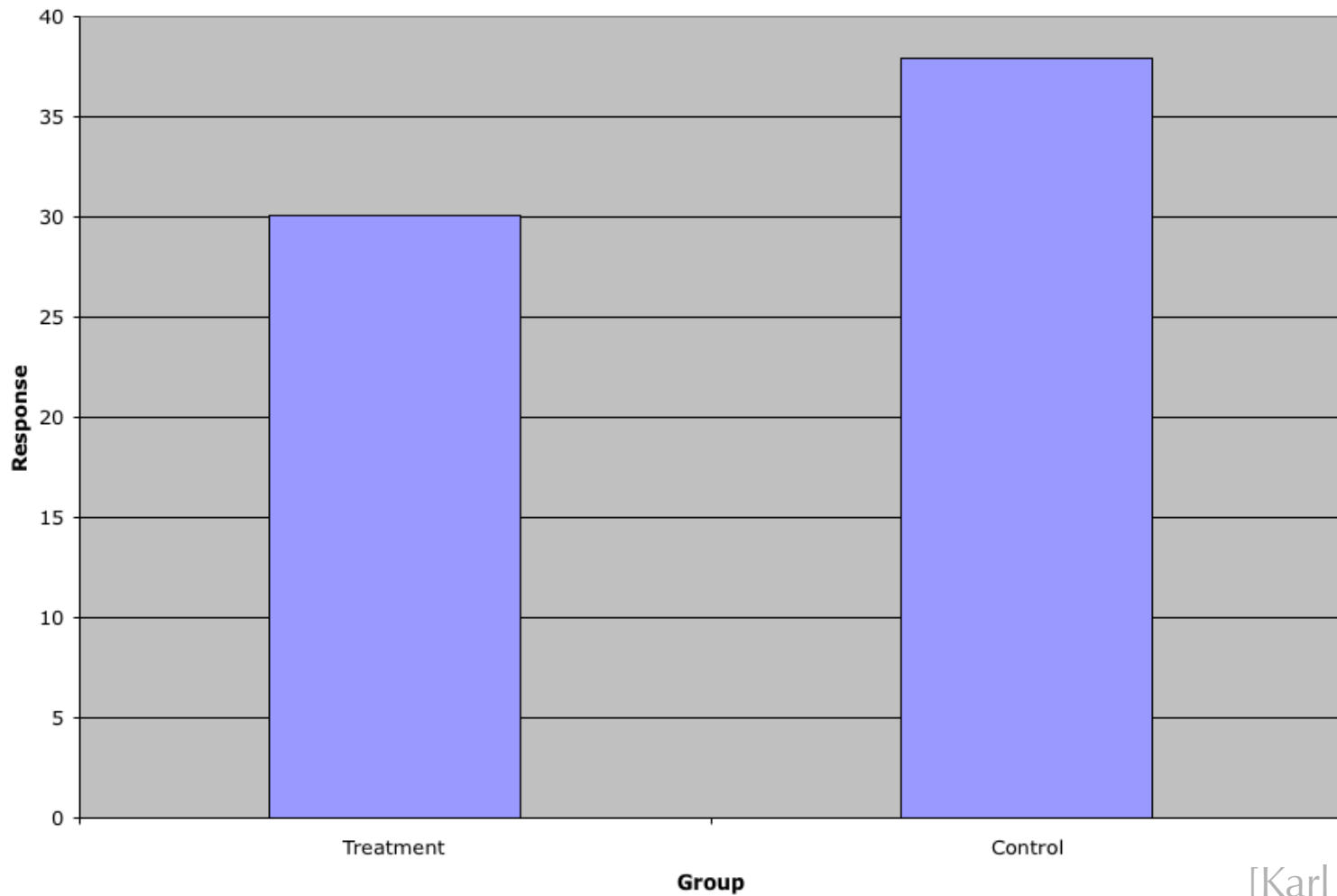
Example 1



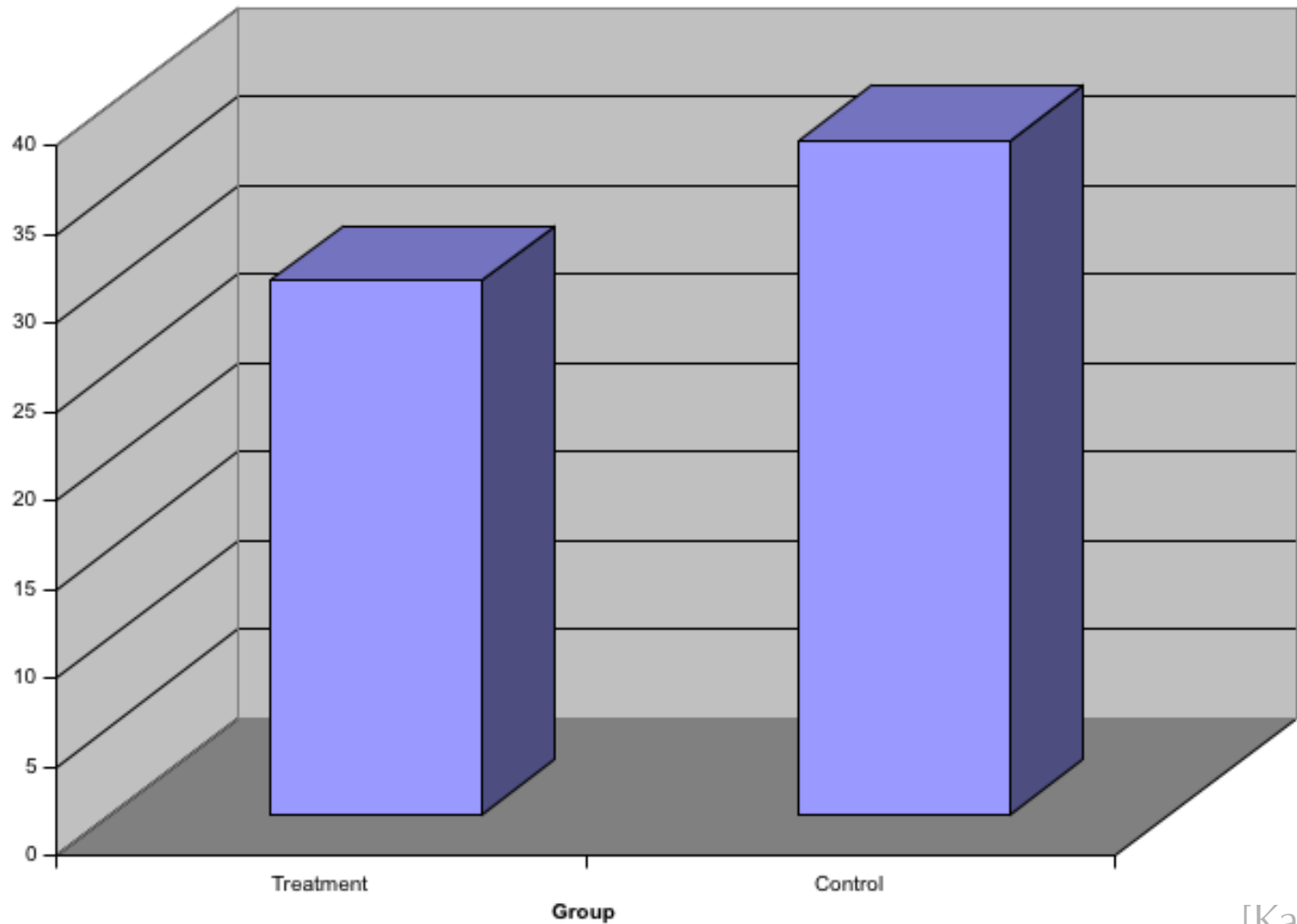
Example 1



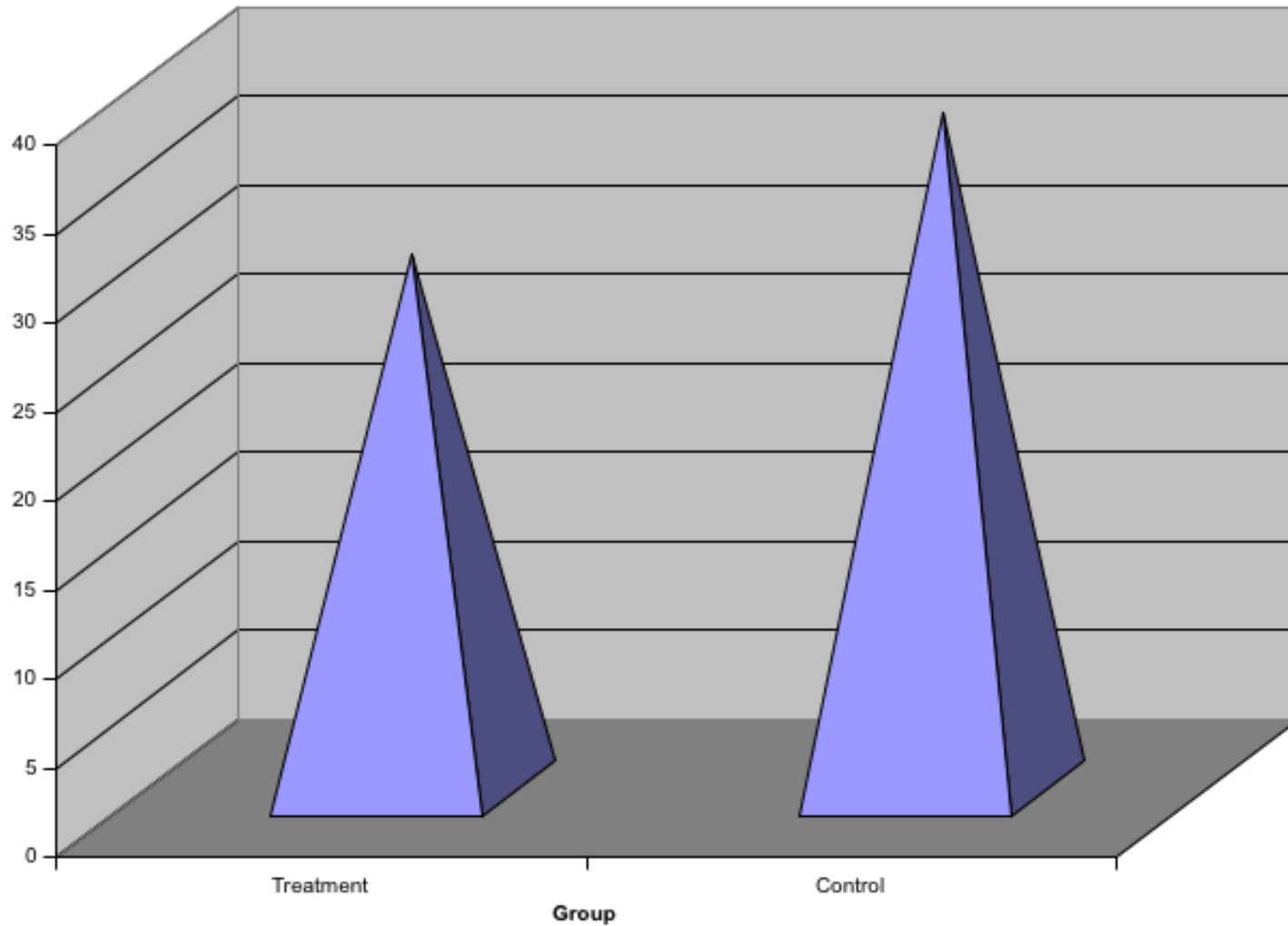
Example 1



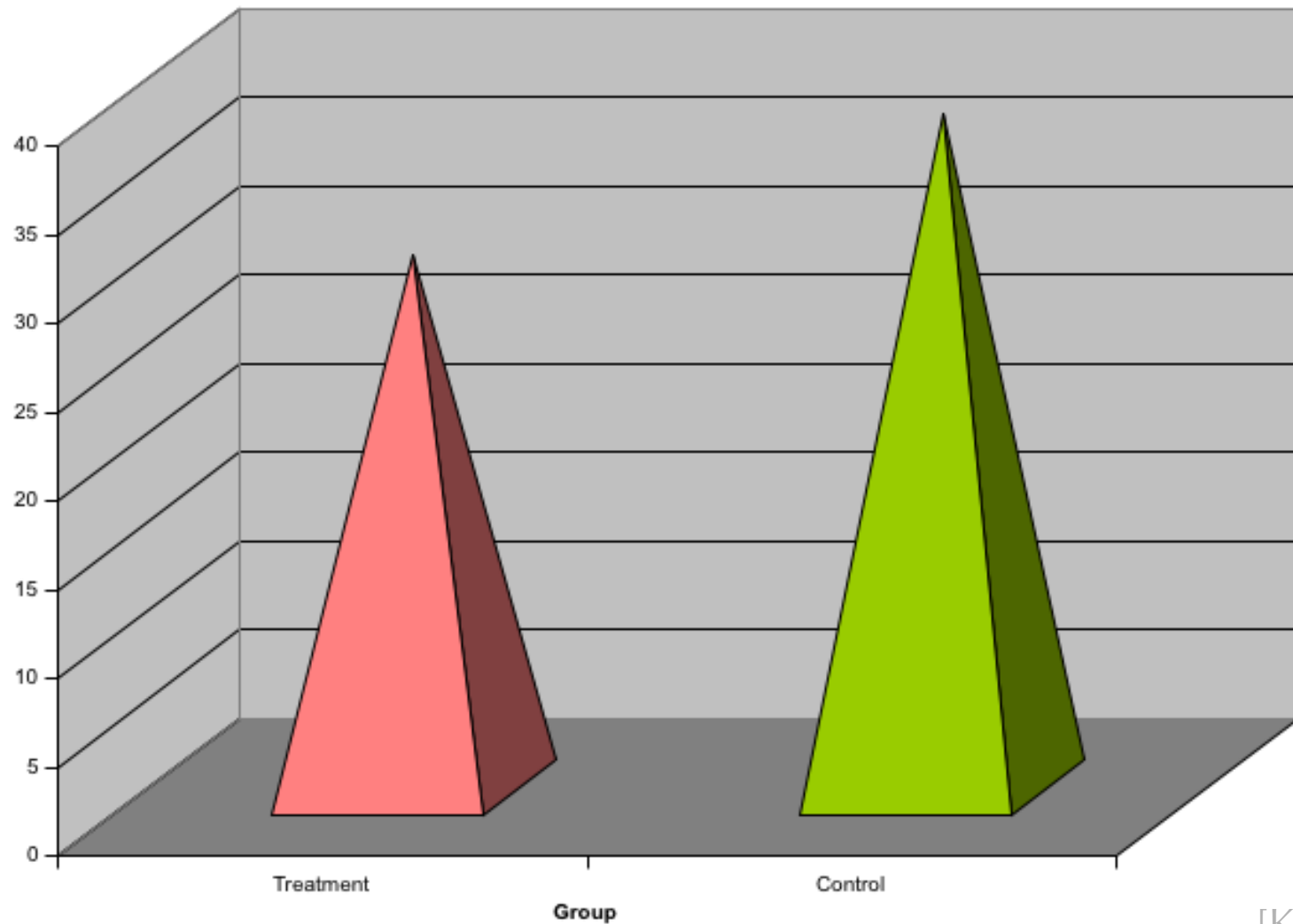
Example 1



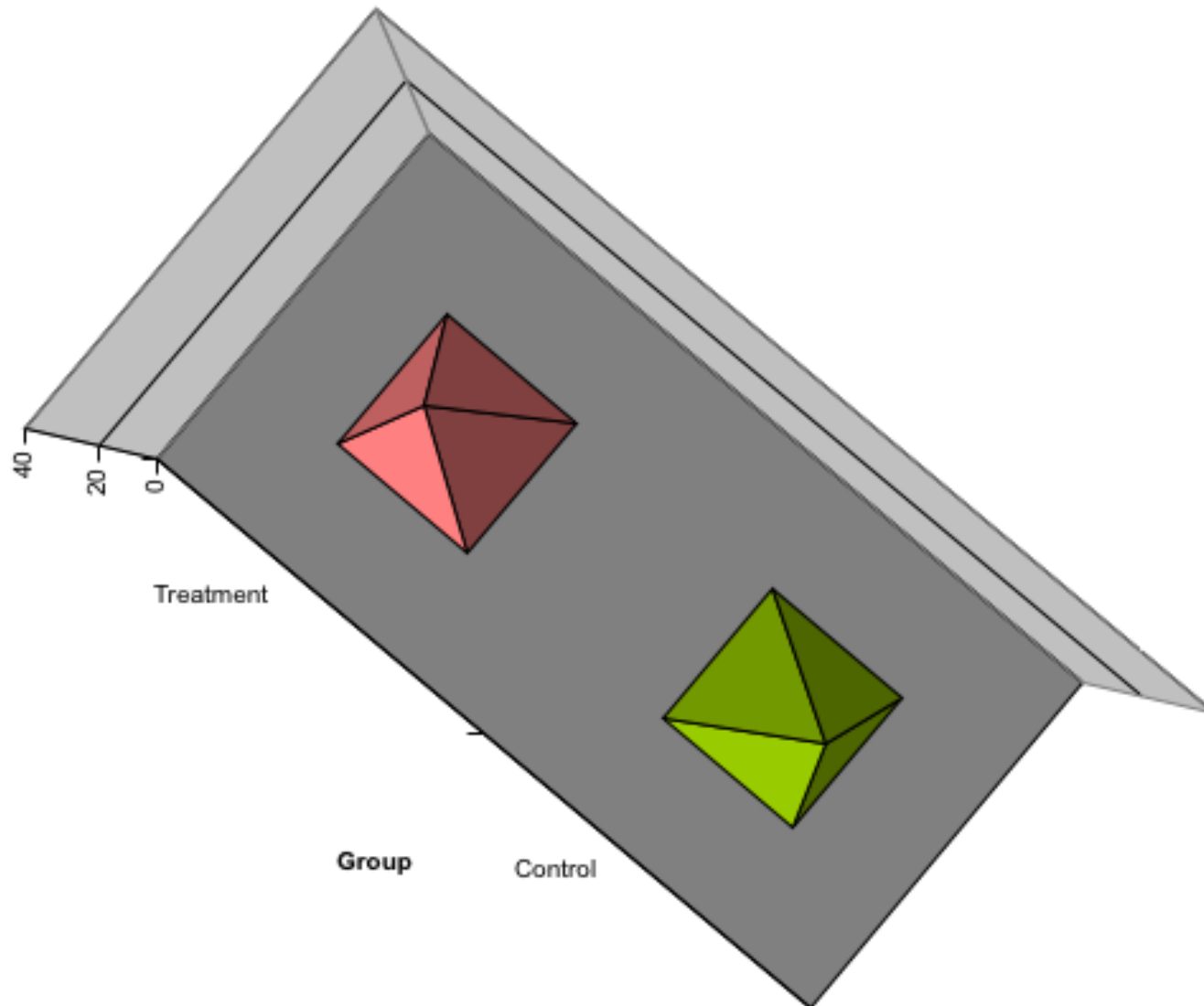
Example 1



Example 1



Example 1



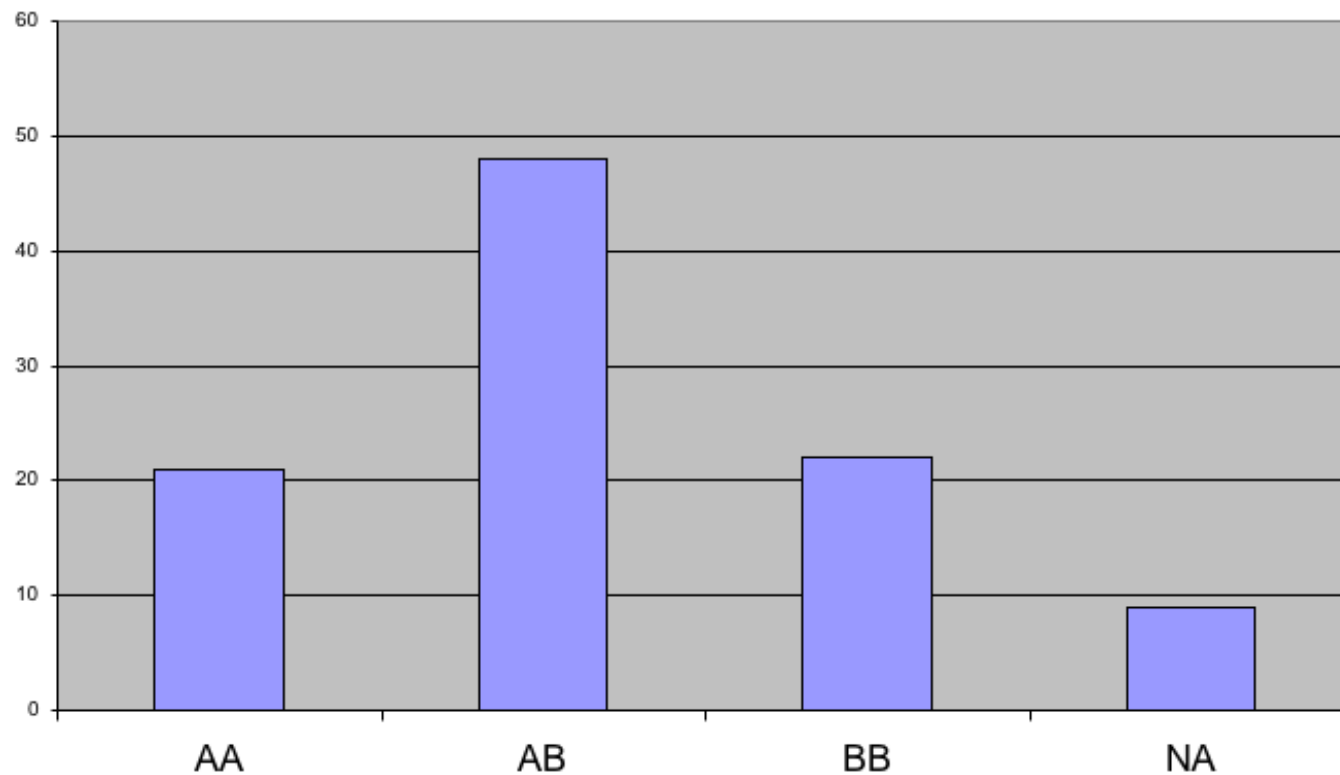
Example 2

Distribution of genotypes

AA	21%
AB	48%
BB	22%
missing	9%

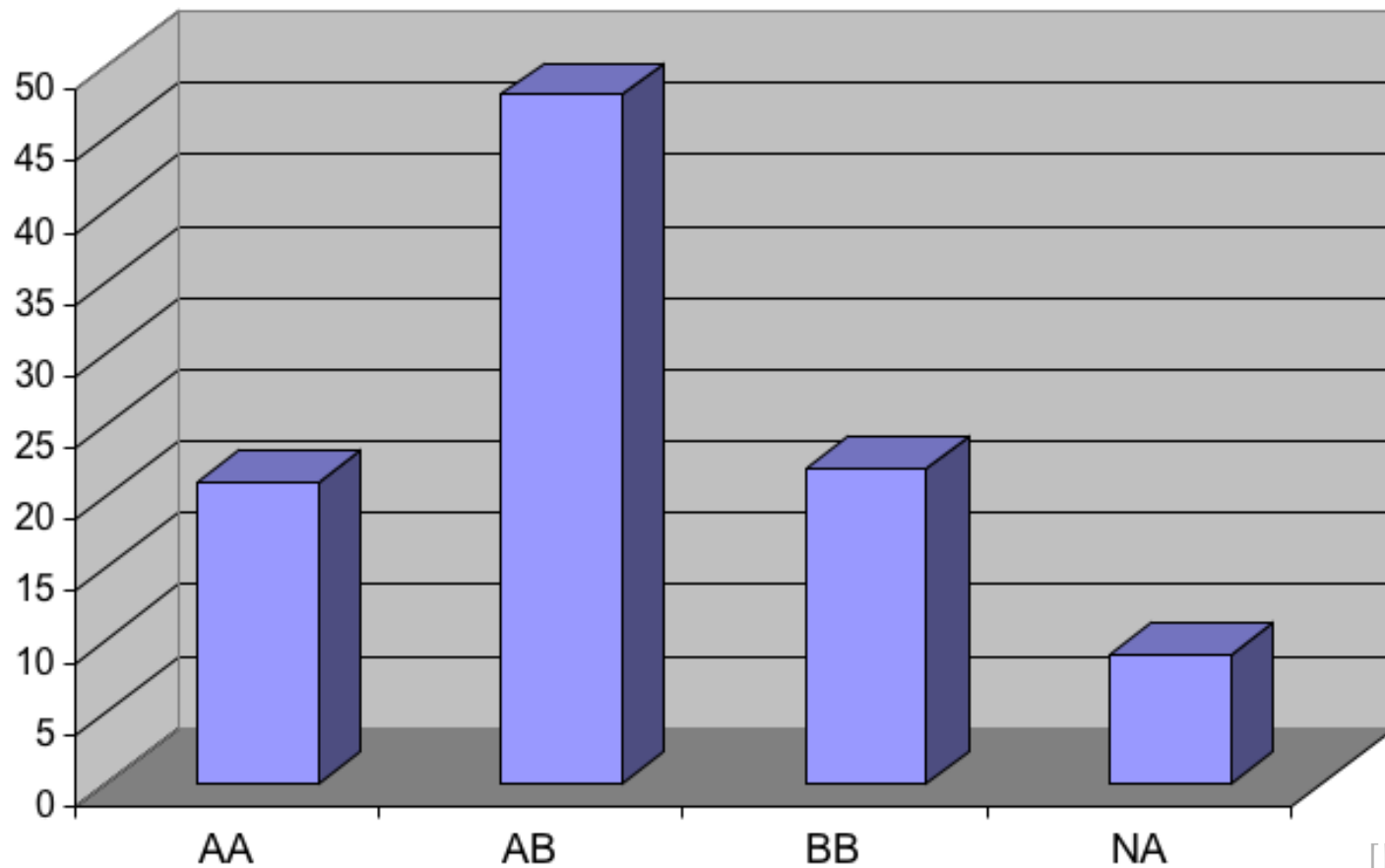
Example 2

Distribution of genotypes

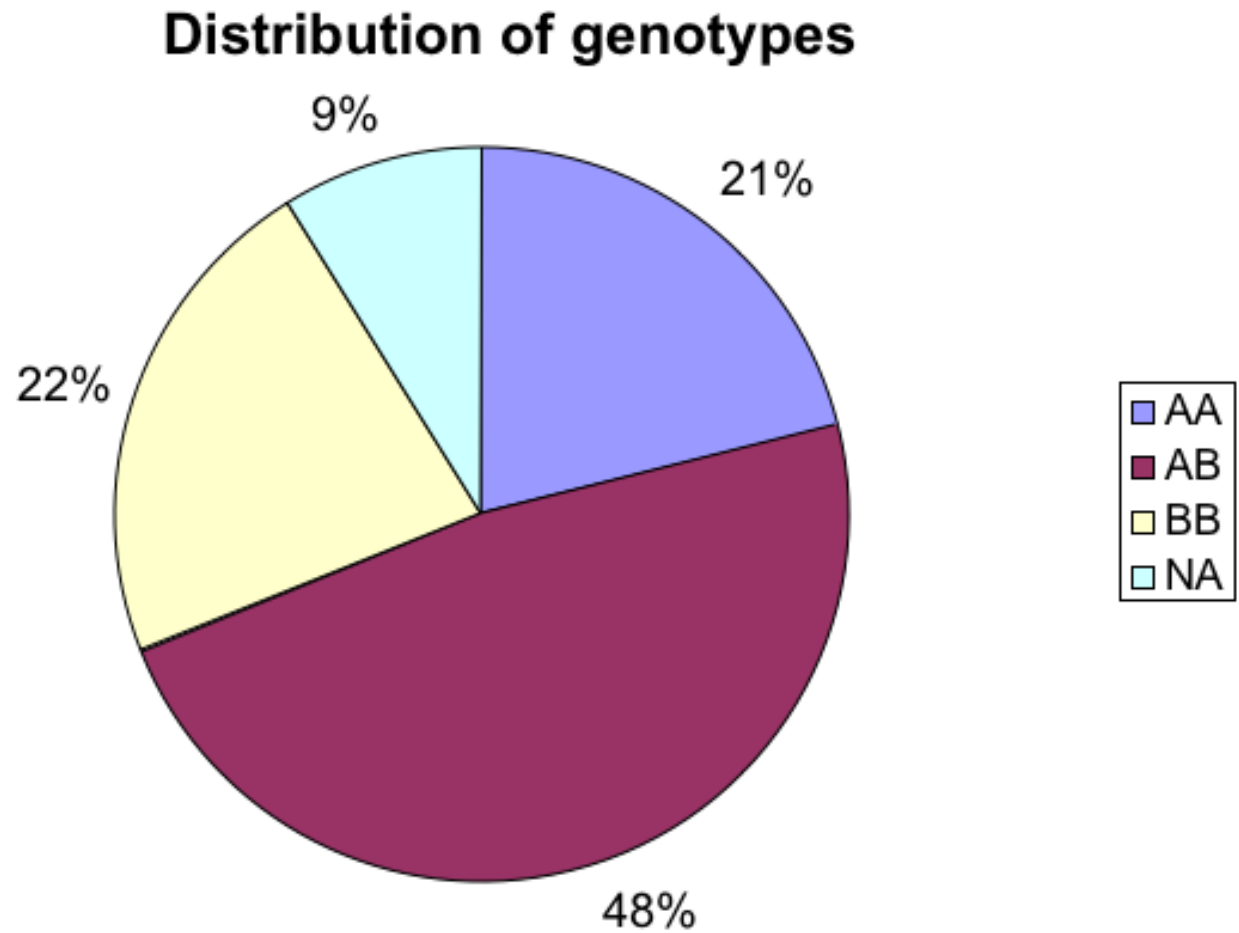


Example 2

Distribution of genotypes

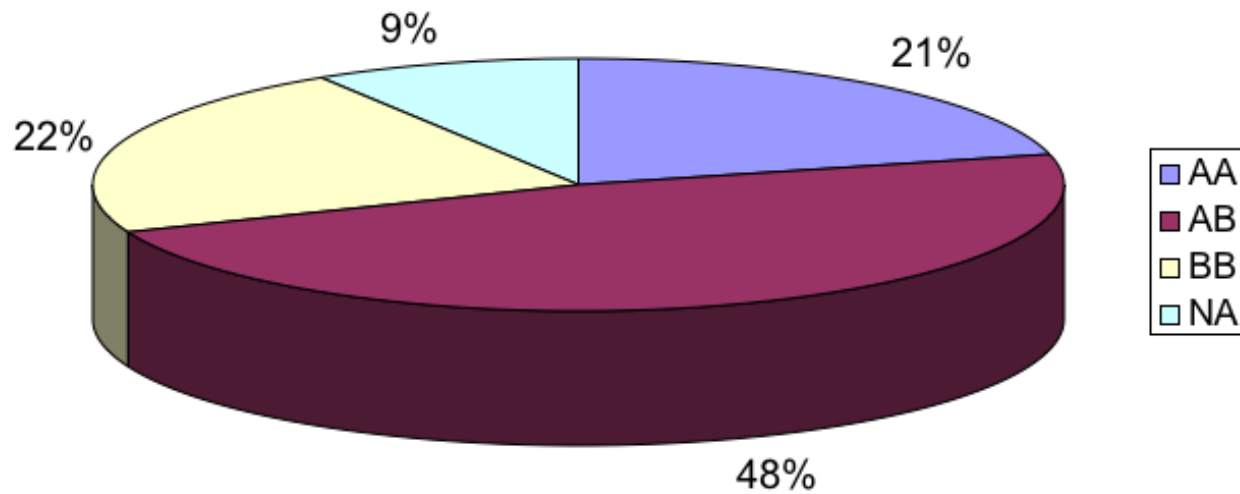


Example 2



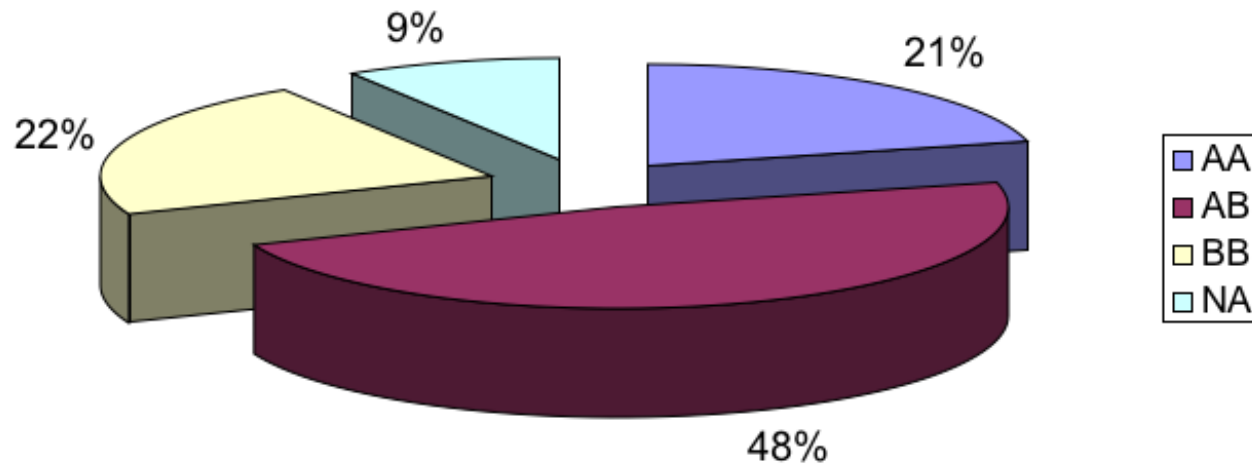
Example 2

Distribution of genotypes



Example 2

Distribution of genotypes



Resources

- <http://flowingdata.com/2010/12/14/10-best-data-visualization-projects-of-the-year-%E2%80%93-2010/>
- <http://www.tableausoftware.com/products/digital/#tour>
- <http://www-958.ibm.com/software/data/cognos/manyeyes/>