The design of data visualizations

CS 48N, Winter 2011

Marc Levoy
Computer Science Department
Stanford University

Why display data graphically?

"The purpose of computing is insight, not numbers."
- Richard Hamming

- computers produce too much data
- bandwidth of human visual channel is high
- large portion of brain devoted to visual processing
- well developed languages for visual communication
- AI versus IA

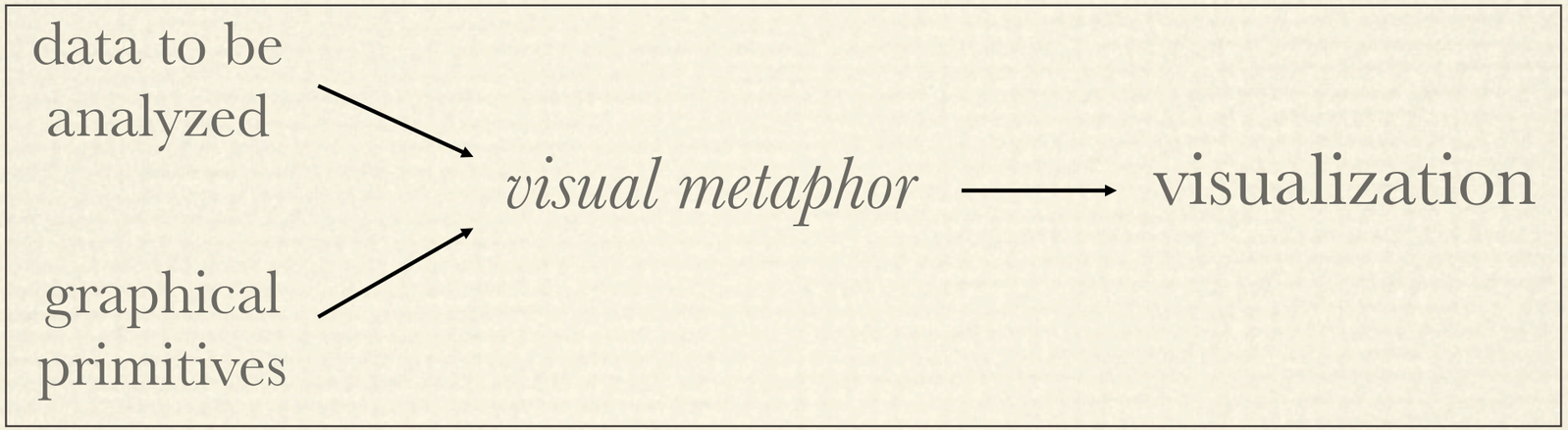
Uses of data visualization

- data exploration
- hypothesis testing
- steering of computations
- program / performance debugging
- presentation
- education

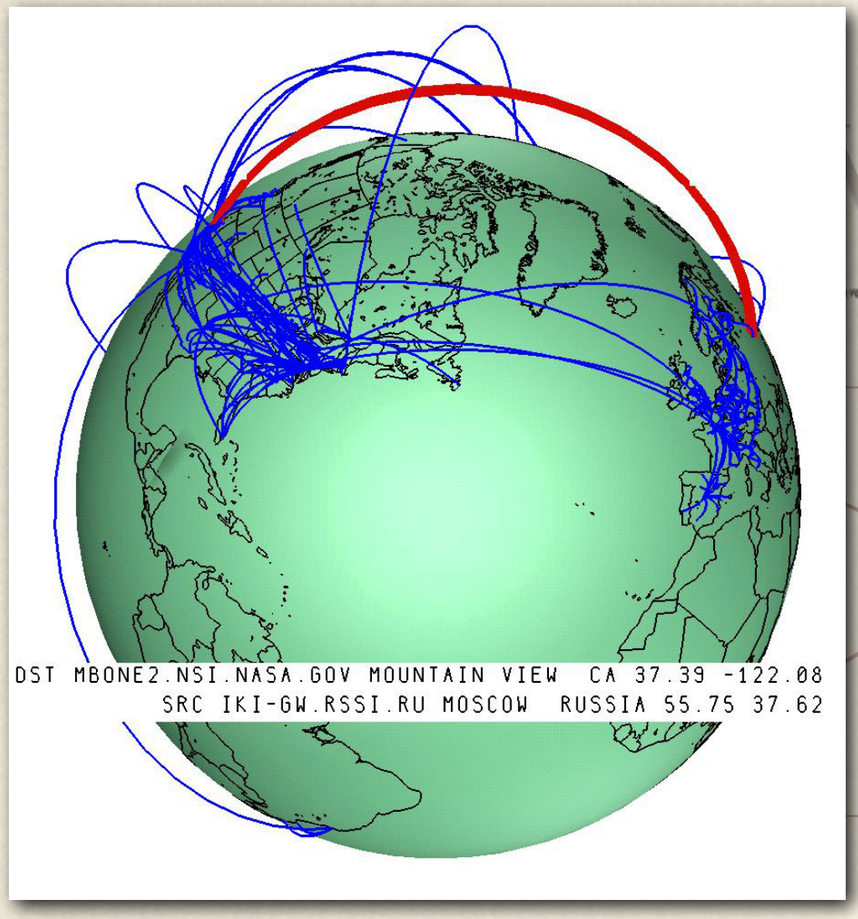
Steps in data exploration

- thinking about the data
- building a visualization
- searching a parameter space
- looking for patterns
- looking for associations
- finding errors
- discovering the unexpected

Visual metaphors



Visual metaphors

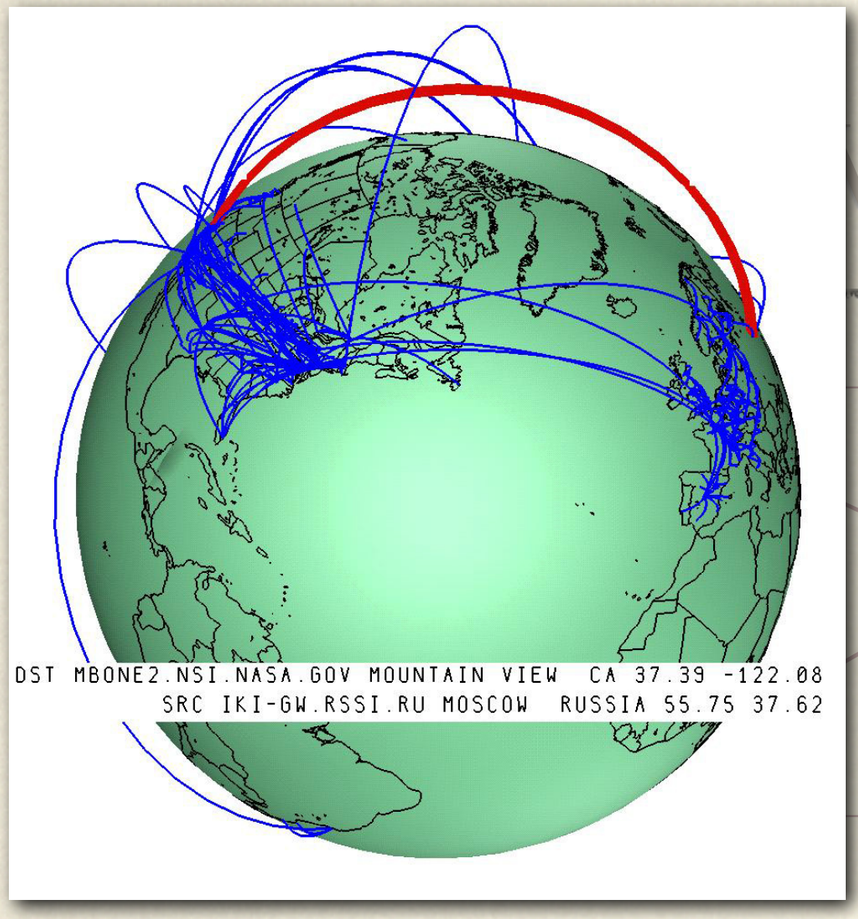


visual metaphor ⟶ visualization

(T. Munzner)

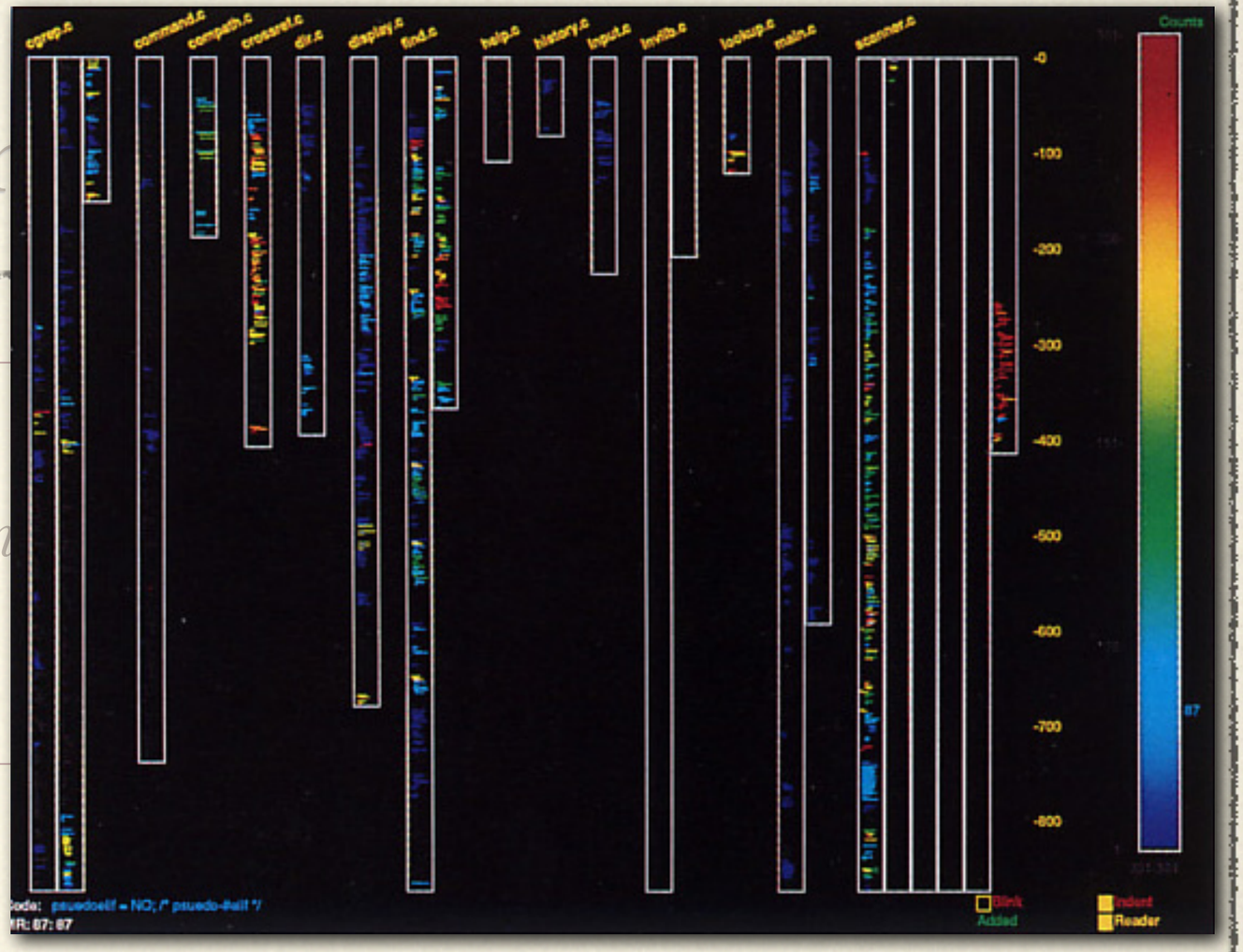
- Some obvious metaphors
  - World Wide Web = display of globe

Visual m



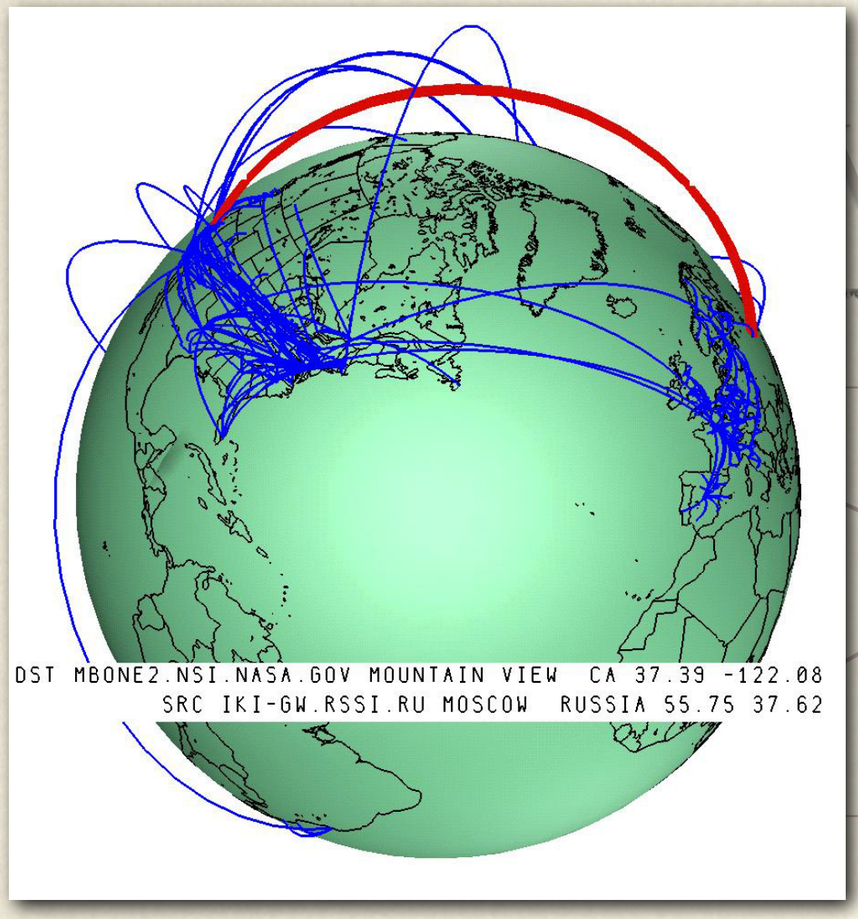
(T. Munzner)

visual m

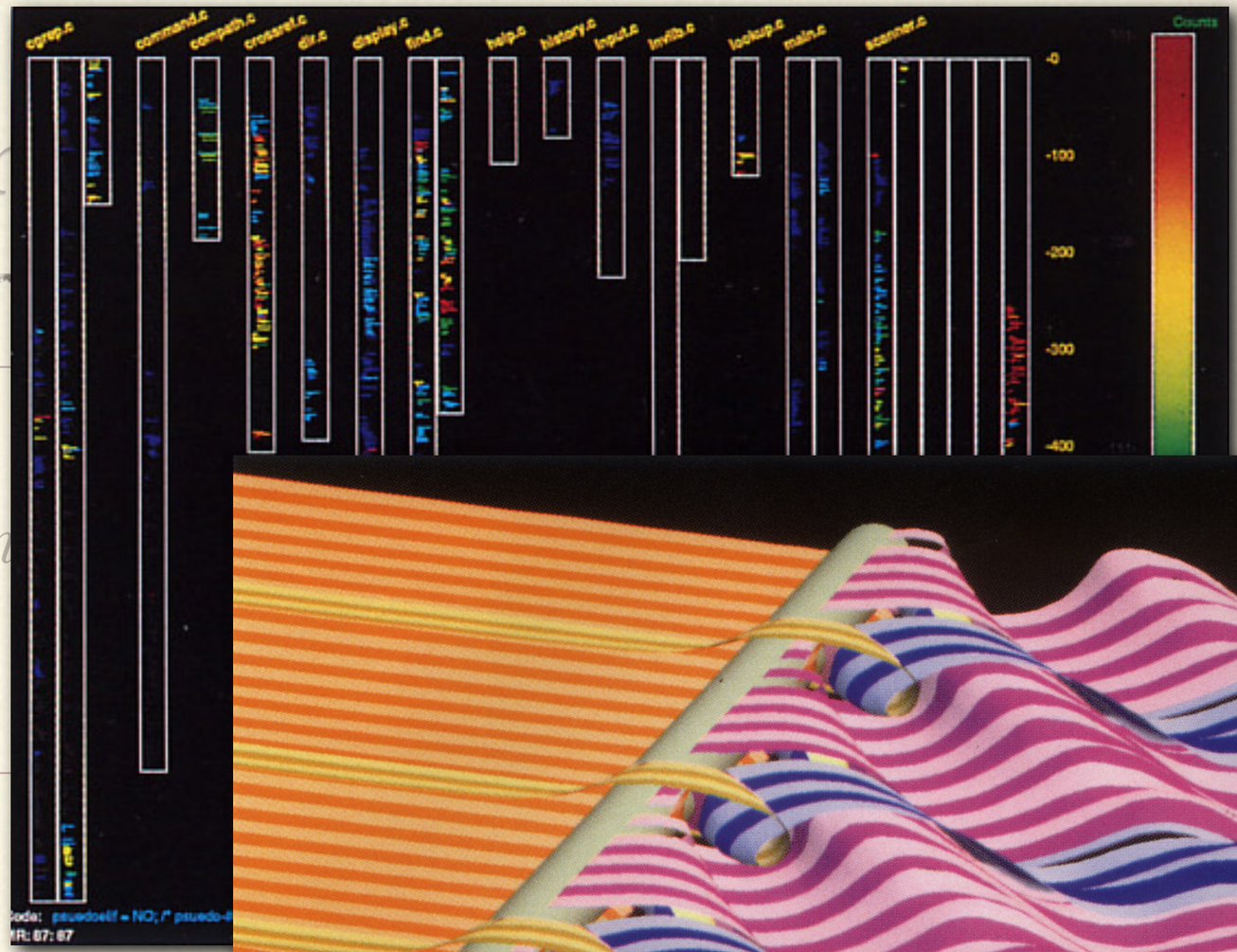


(S.G. Eick)

- Some obvious metaphors
  - World Wide Web = display of globe
  - code profiling = temperature map



(T. Munzner)



(J. Helman)

- Some obvious metaphors
  - World Wide Web = display of globe
  - code profiling = temperature map
- Some non-obvious metaphors
  - fluid flow over time = 3D “strip surface”

The design of visualizations

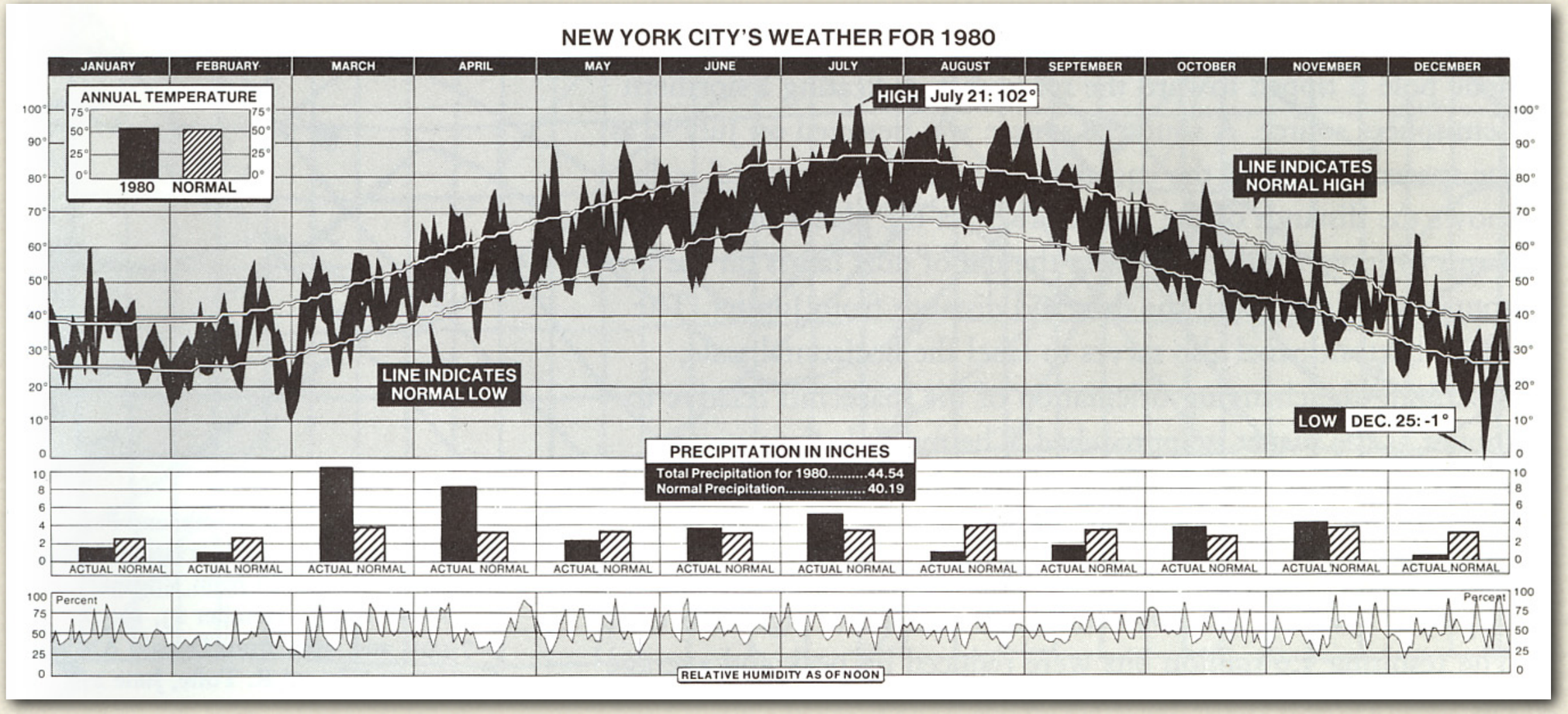
- More variables at once
- Micro/macro readings
- Small multiples
- Graphic design
- Things to avoid

More variables at once

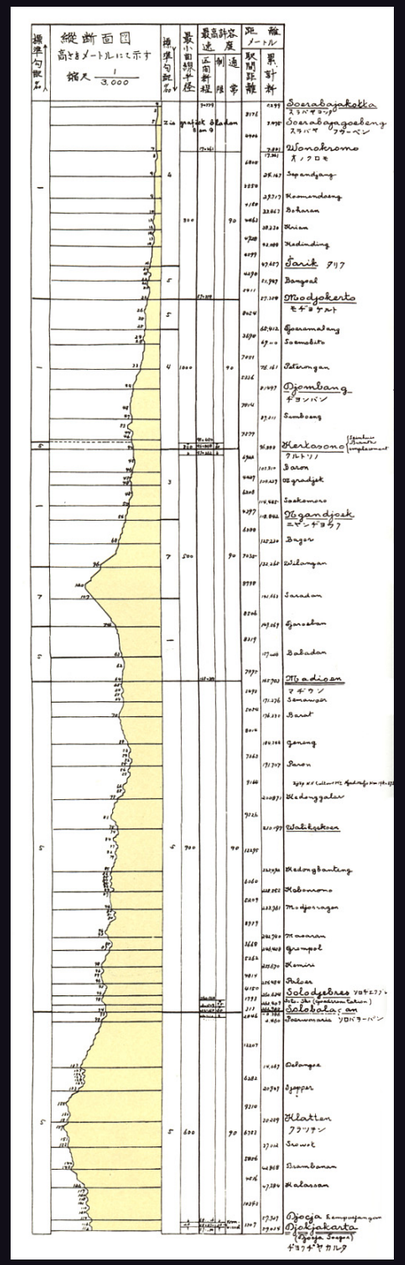
- parallel plots
- drop variables
- combine variables
- use more visual cues
- compress scales

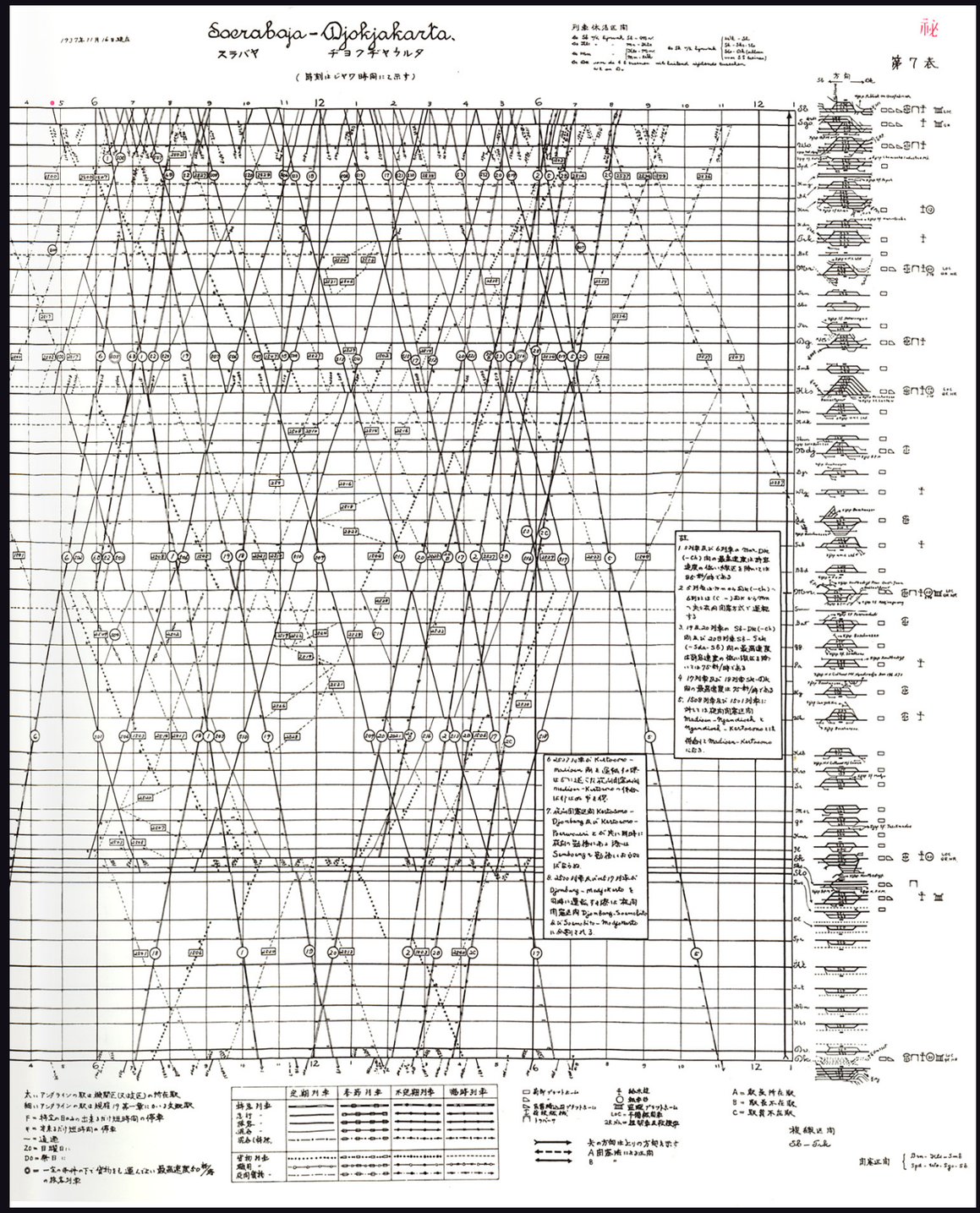
More variables at once

- parallel plots



(from Tufte)

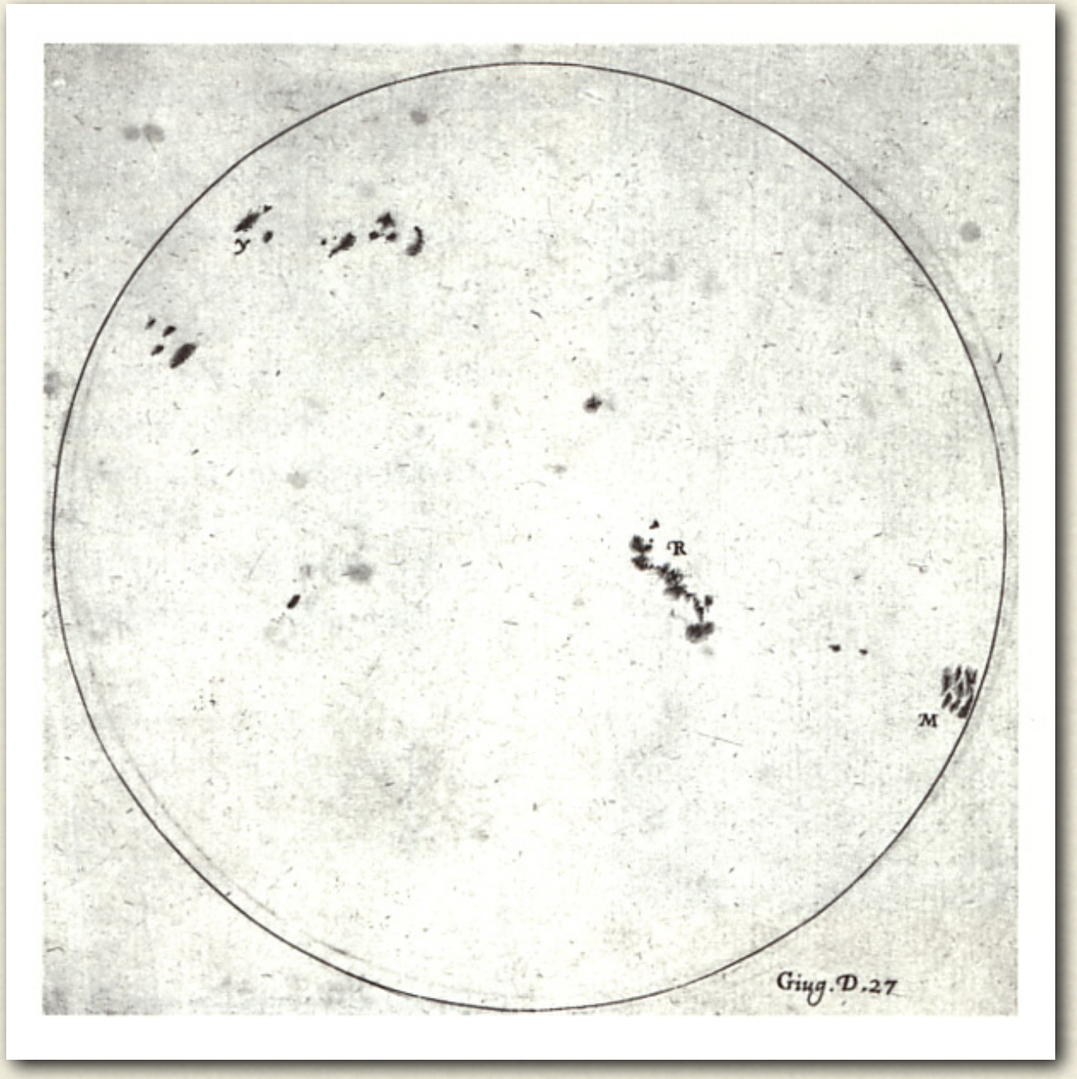




Soerabaja-Djokjakarta train line

Example: sunspots

1613 Galileo

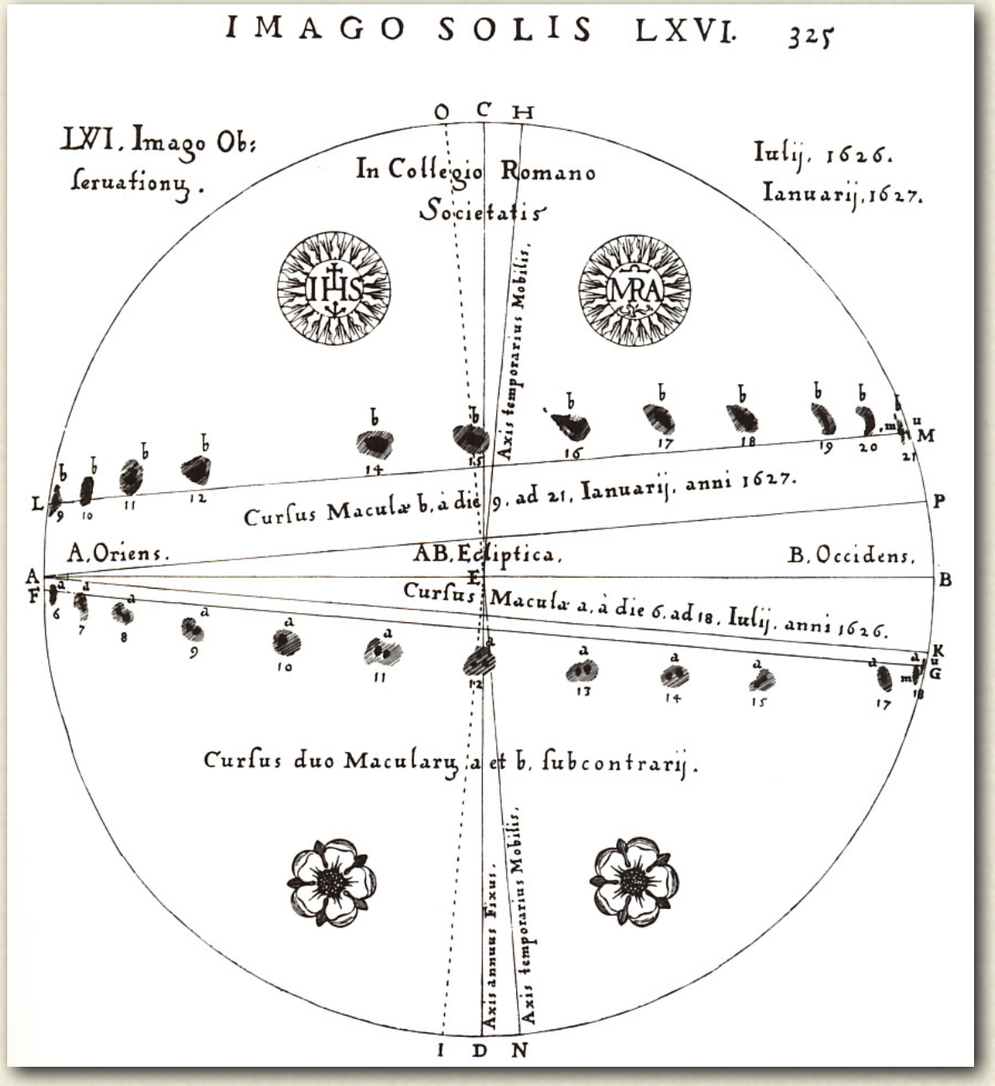


Example: sunspots

1613 Galileo

1630 Scheiner

* superimposed multiples, showing path



Examp

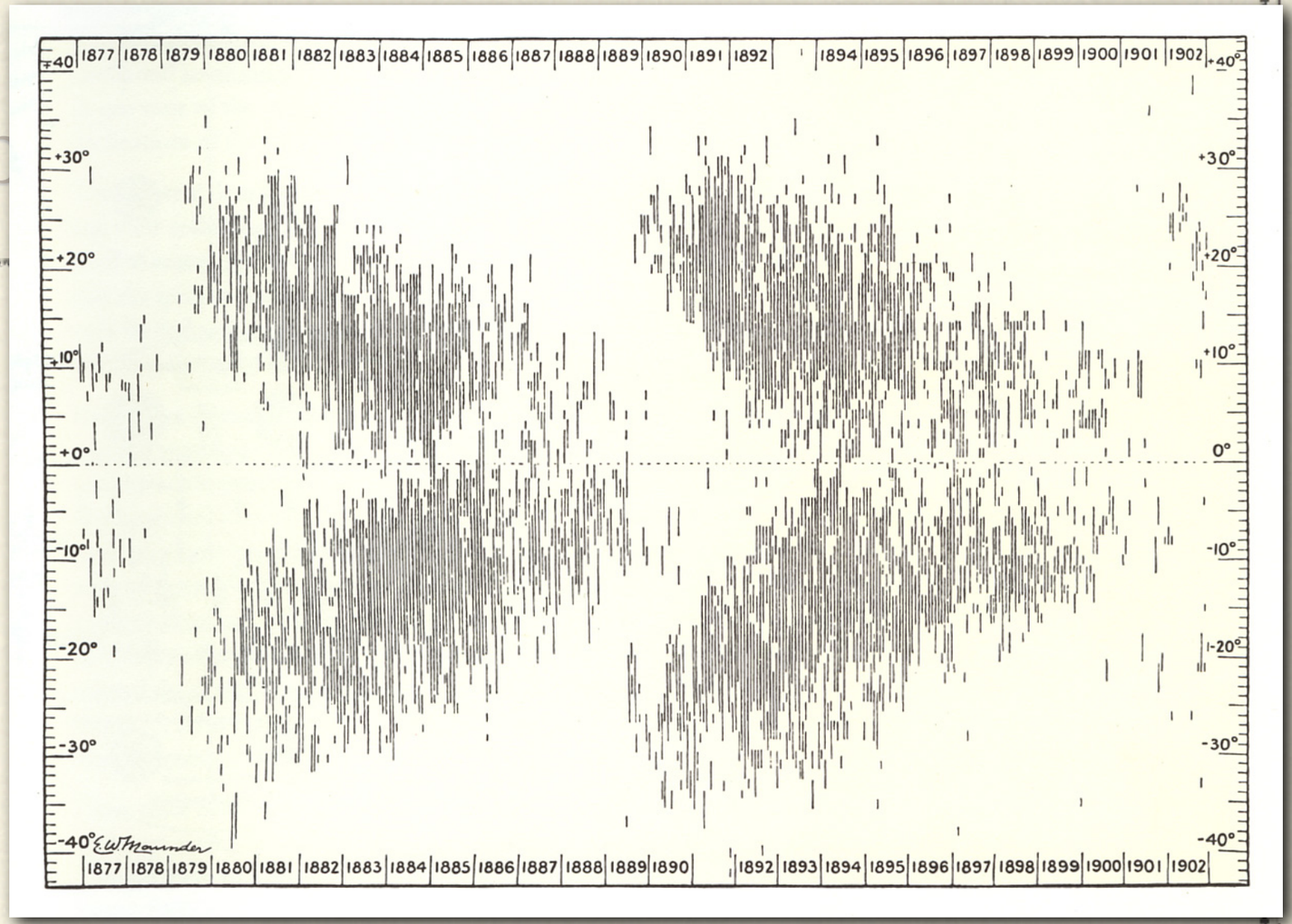
1613 Galileo

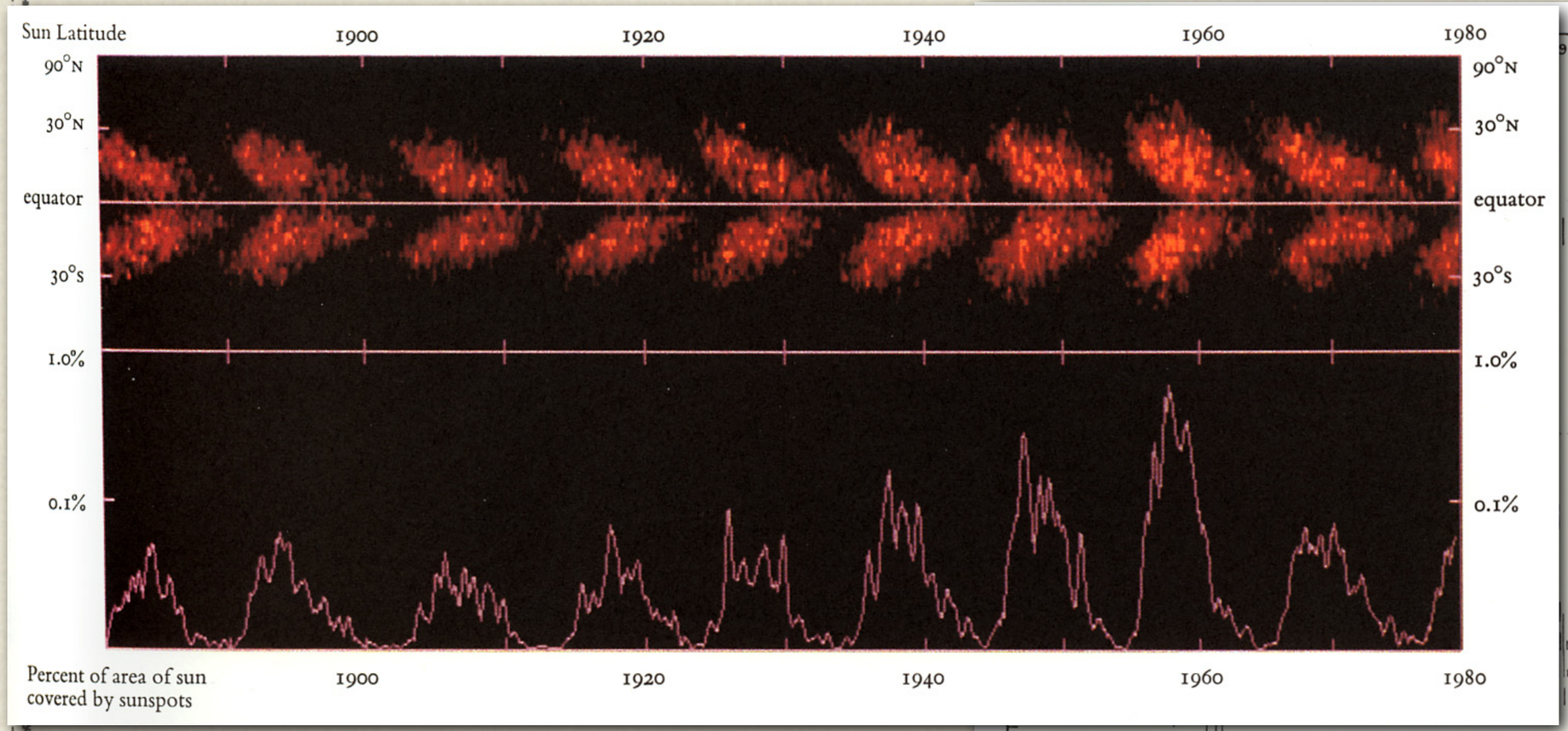
1630 Scheiner

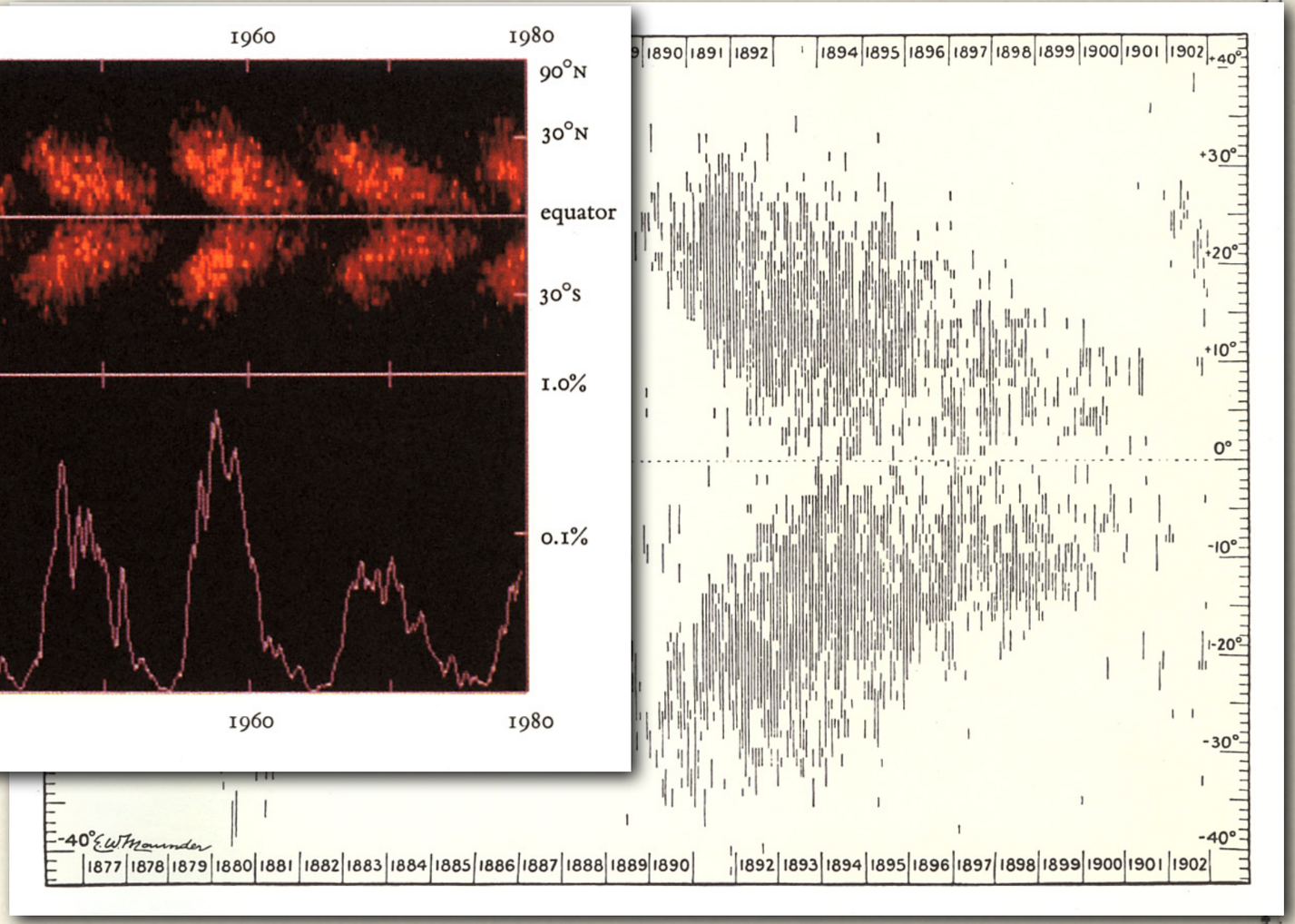
* superimposed multiples, showing path

1904 Maunder

* horizontal extent dropped
* vertical extent displayed as scatterplot
* time axis compressed
* latitude transformed to heliocentric coordinates







1980 Hathaway

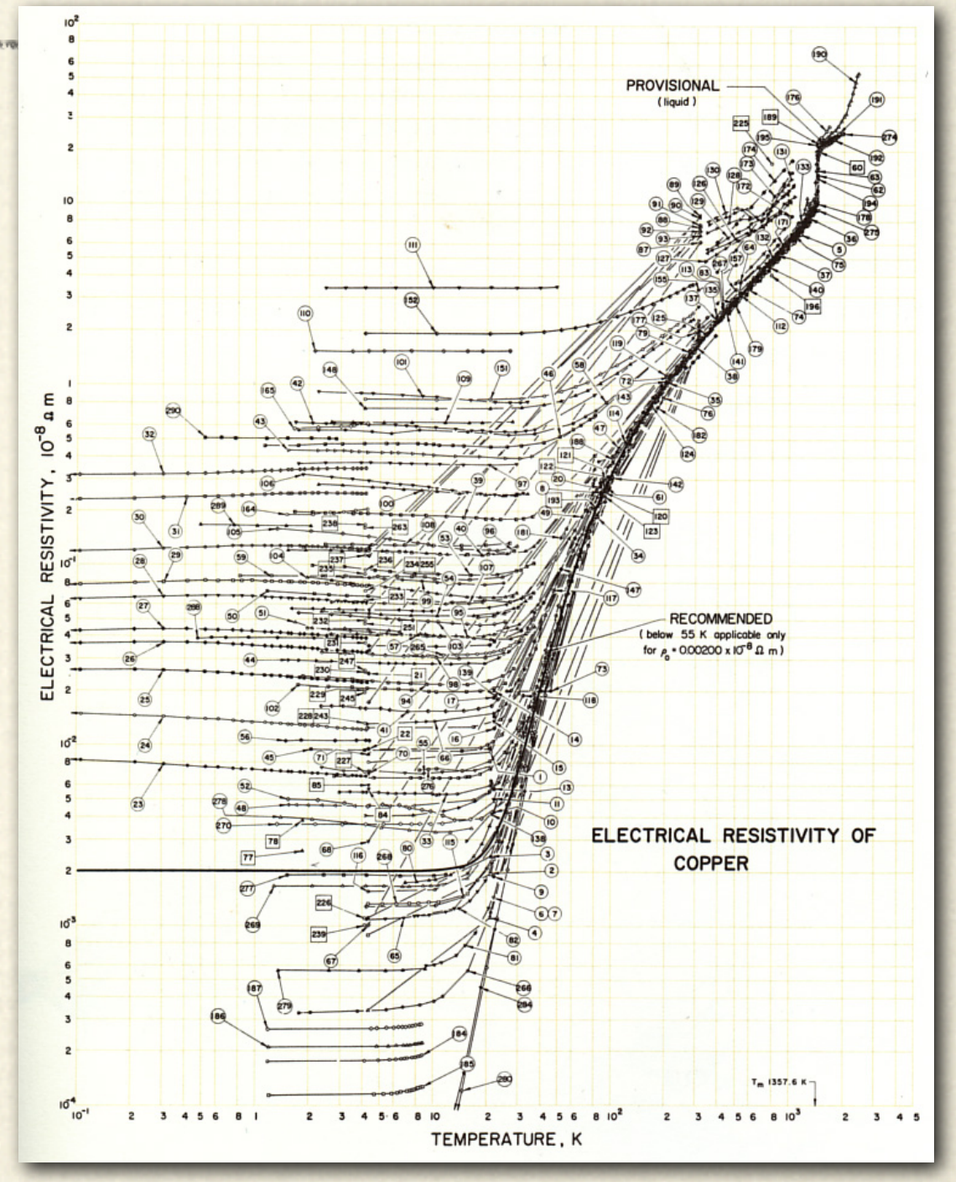
* non-linear latitude axis
* sunspot area coded using color
* time axis compressed more
* total area displayed as parallel plot

Micro/macro readings

- motivations
  - interpretation arises directly from data
  - allows personal readings
- techniques
  - choose the right level of abstraction
  - careful handling of line weights
  - subtle labeling

Micro/macro readings

- motivations
  - interpretation arises directly from data
  - allows personal readings
- techniques
  - choose the right level of abstraction
  - careful handling of line weights
  - subtle labeling



(from Tufte)

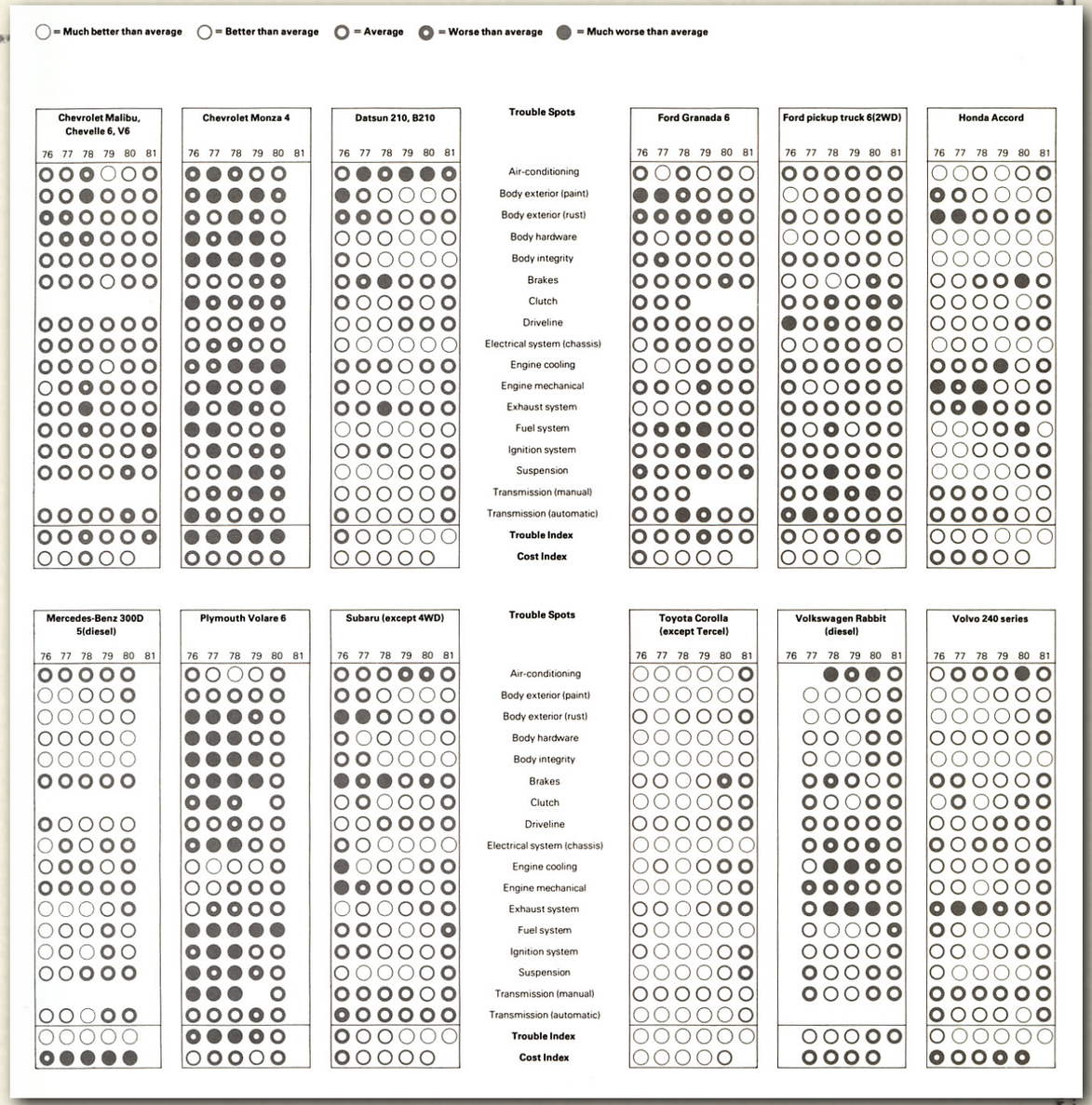
(U.S. Census Bureau)

Small multiples

- motivations
  - used to show the “outermost” variable
  - engages human ability to find differences
- techniques
  - position slices within an eyespan
  - distort non-critical aspects to allow this
  - draw non-varying aspects identically

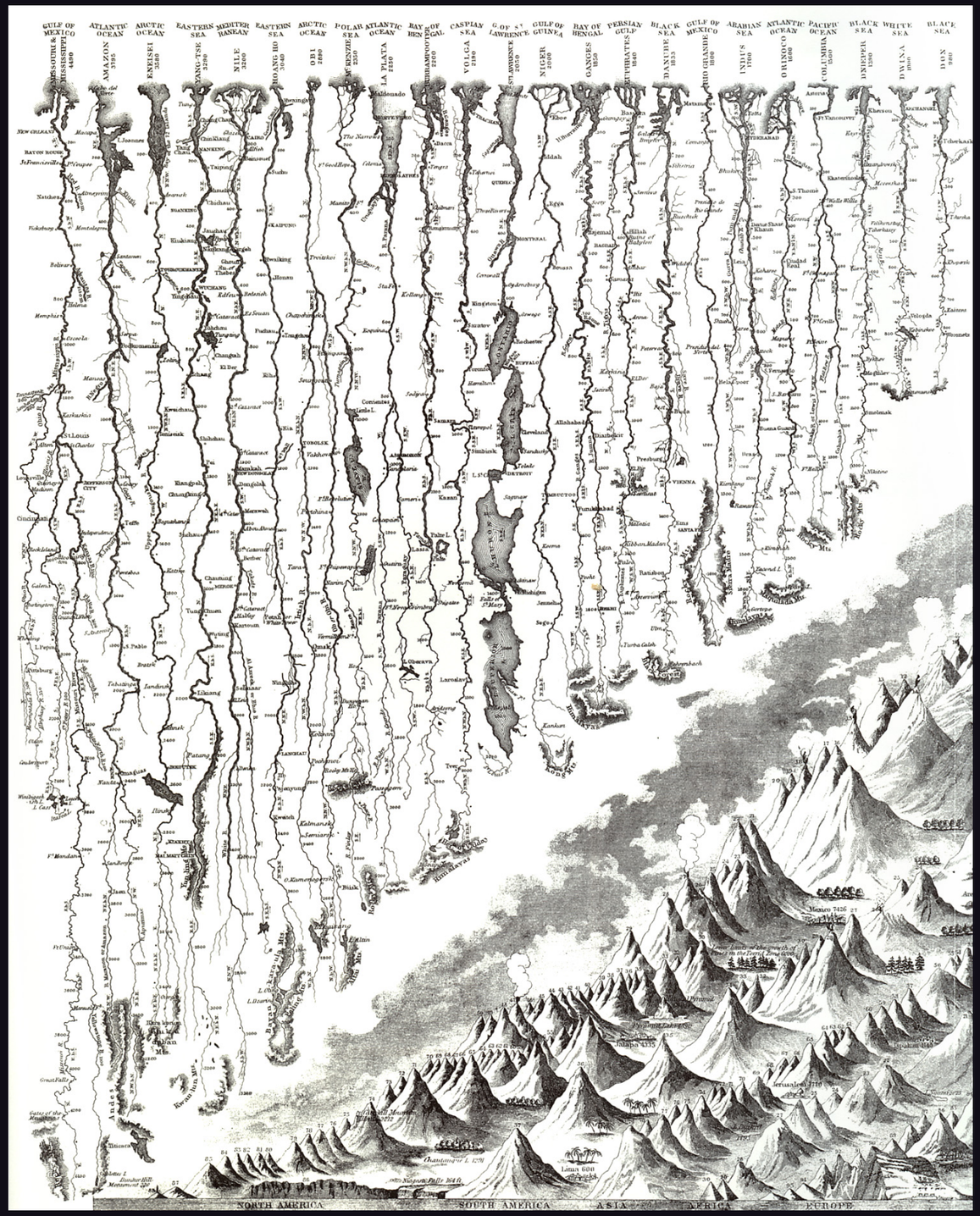
Small multiples

- motivations
  - used to show the “outermost” variable
  - engages human ability to find differences
- techniques
  - position slices within an eyespan
  - distort non-critical aspects to allow this
  - draw non-varying aspects identically



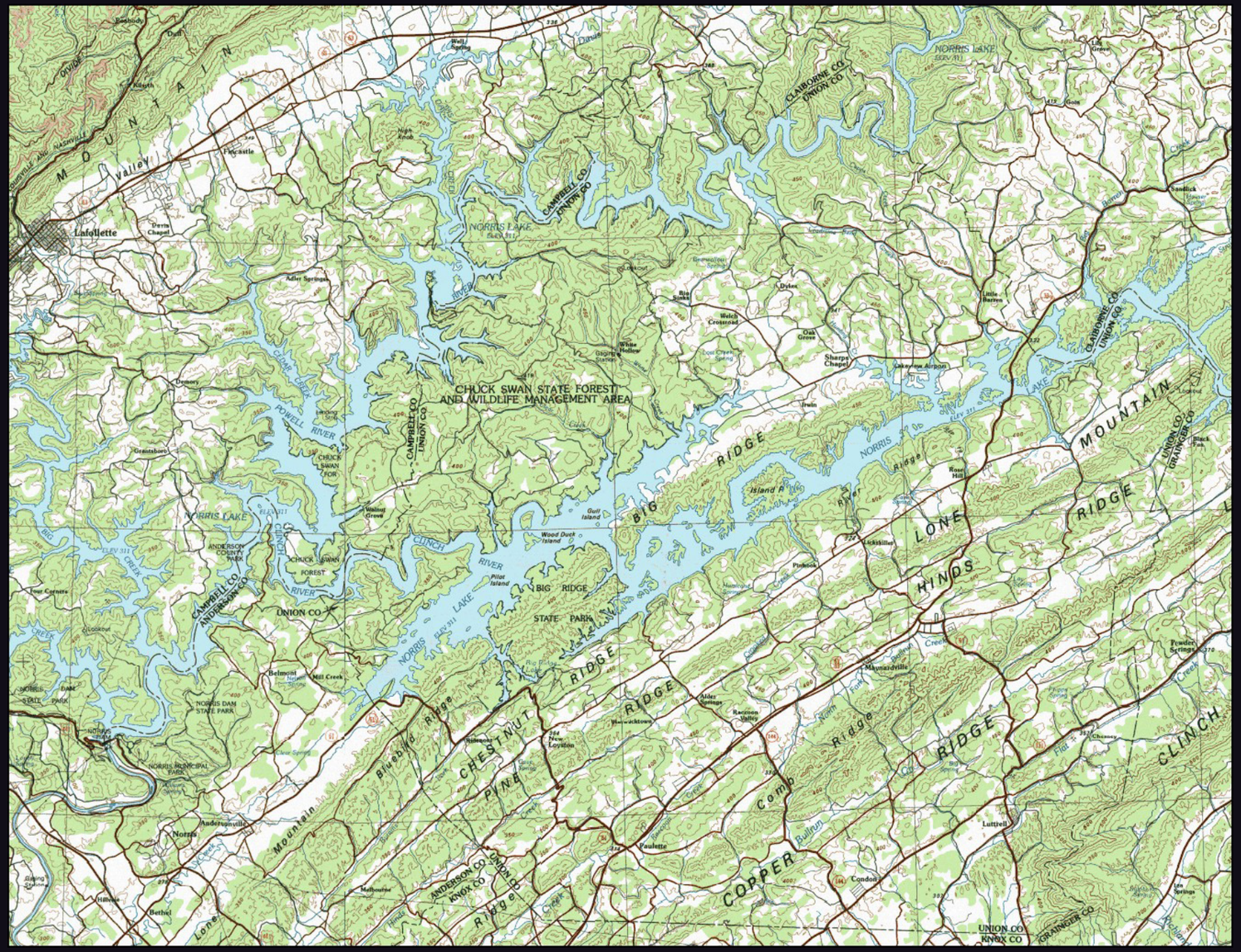
(Consumer Reports)

Joseph Colton
Rivers of the world
(1864)

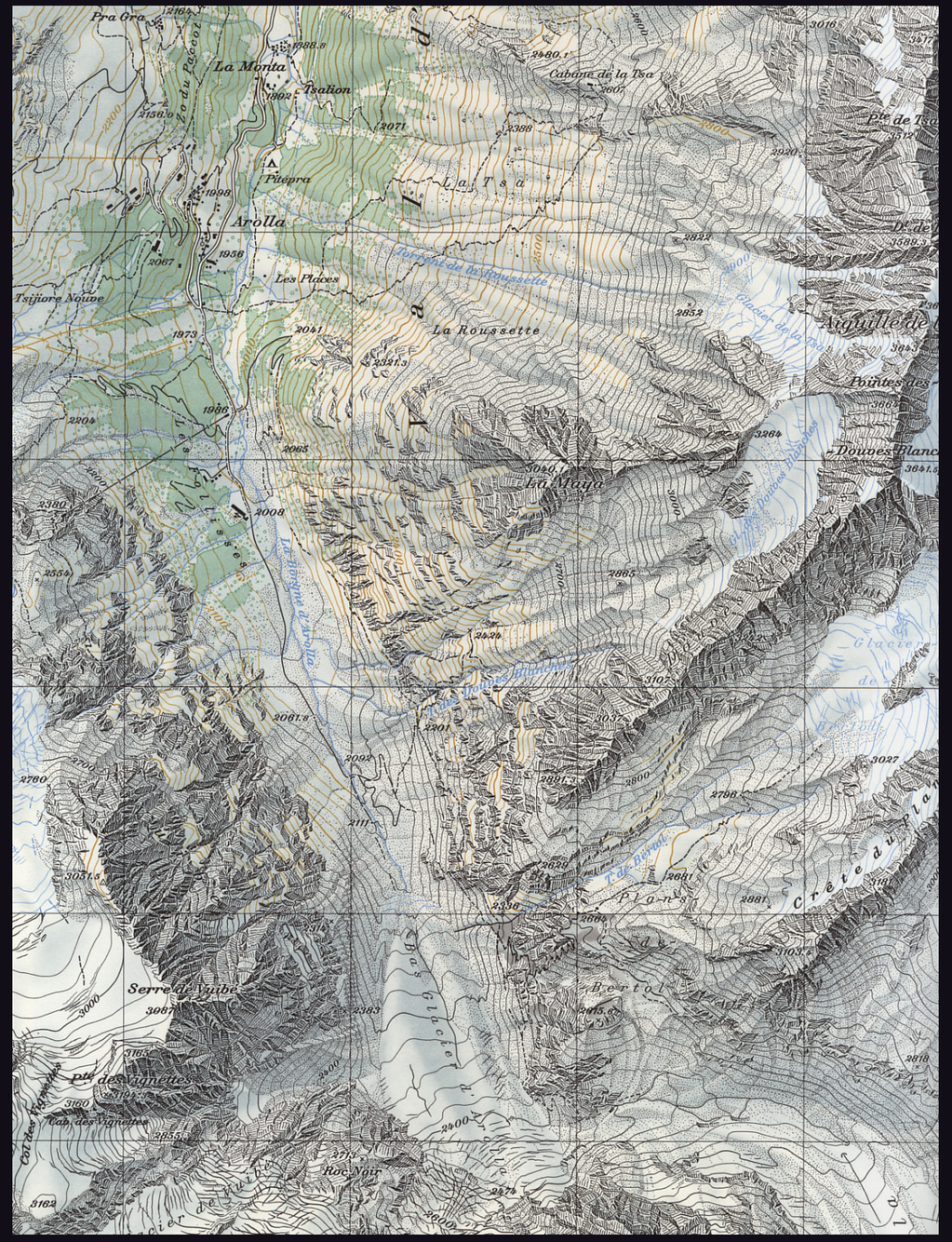


Graphic design

- layering and separation
  - spatial separation
  - variations in density
  - line weight
  - color



USGS map, Union County, Tennessee



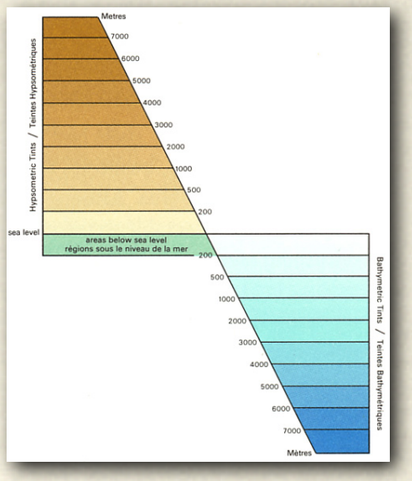
Imhof, Swiss topographic map

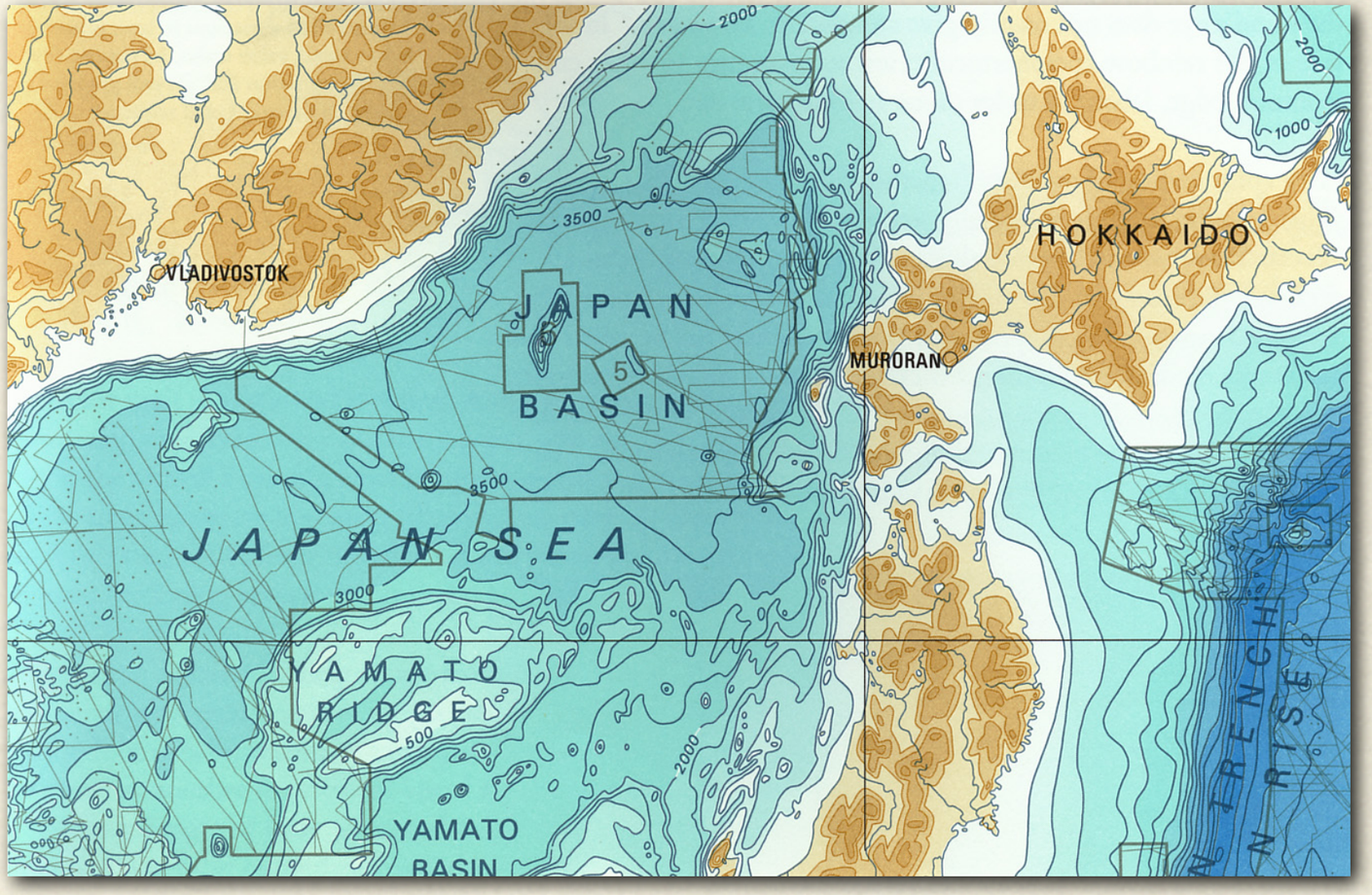
Graphic design

- layering and separation
  - spatial separation
  - variations in density
  - line weight
  - color
- color as a variable
  - pseudo-color scales

Graphic design

- layering and separation
  - spatial separation
  - variations in density
  - line weight
  - color
- color as a variable
  - pseudo-color scales

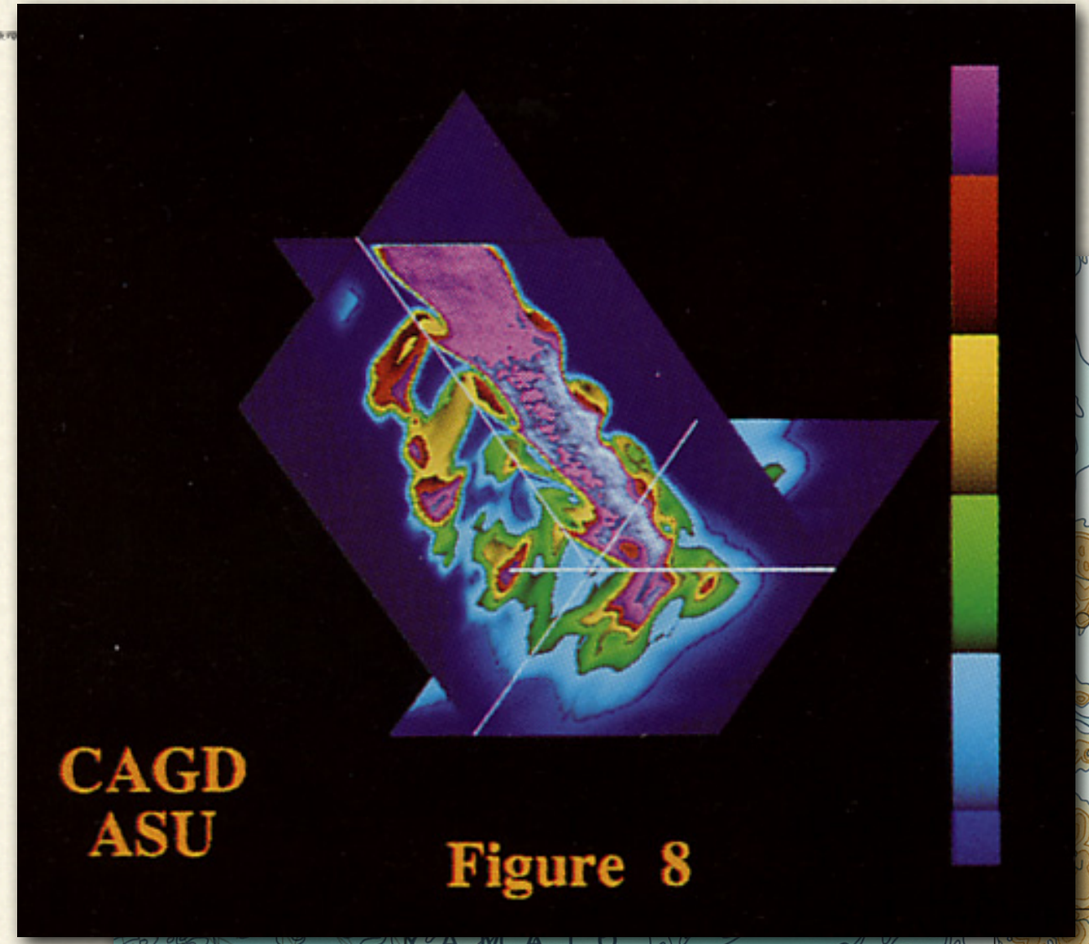


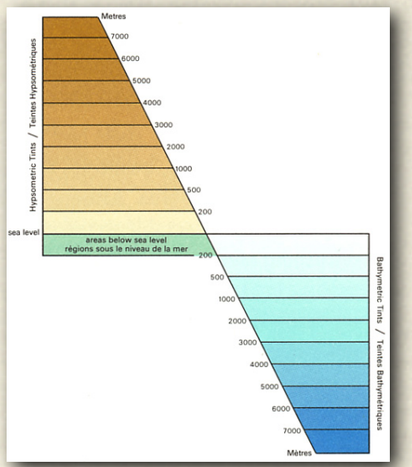


(from Tufte)

Graphic design

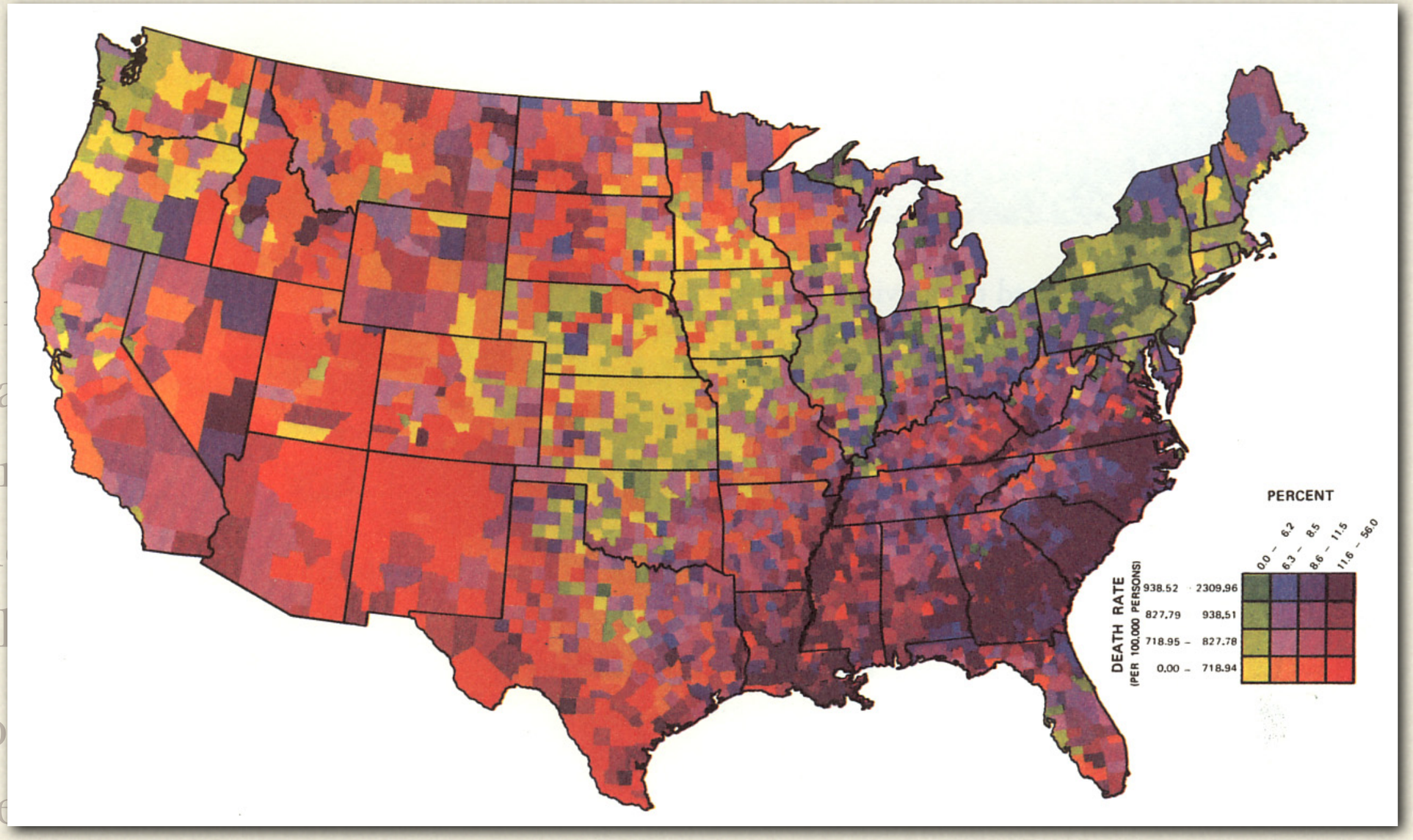
- layering and separation
  - spatial separation
  - variations in density
  - line weight
  - color
- color as a variable
  - pseudo-color scales





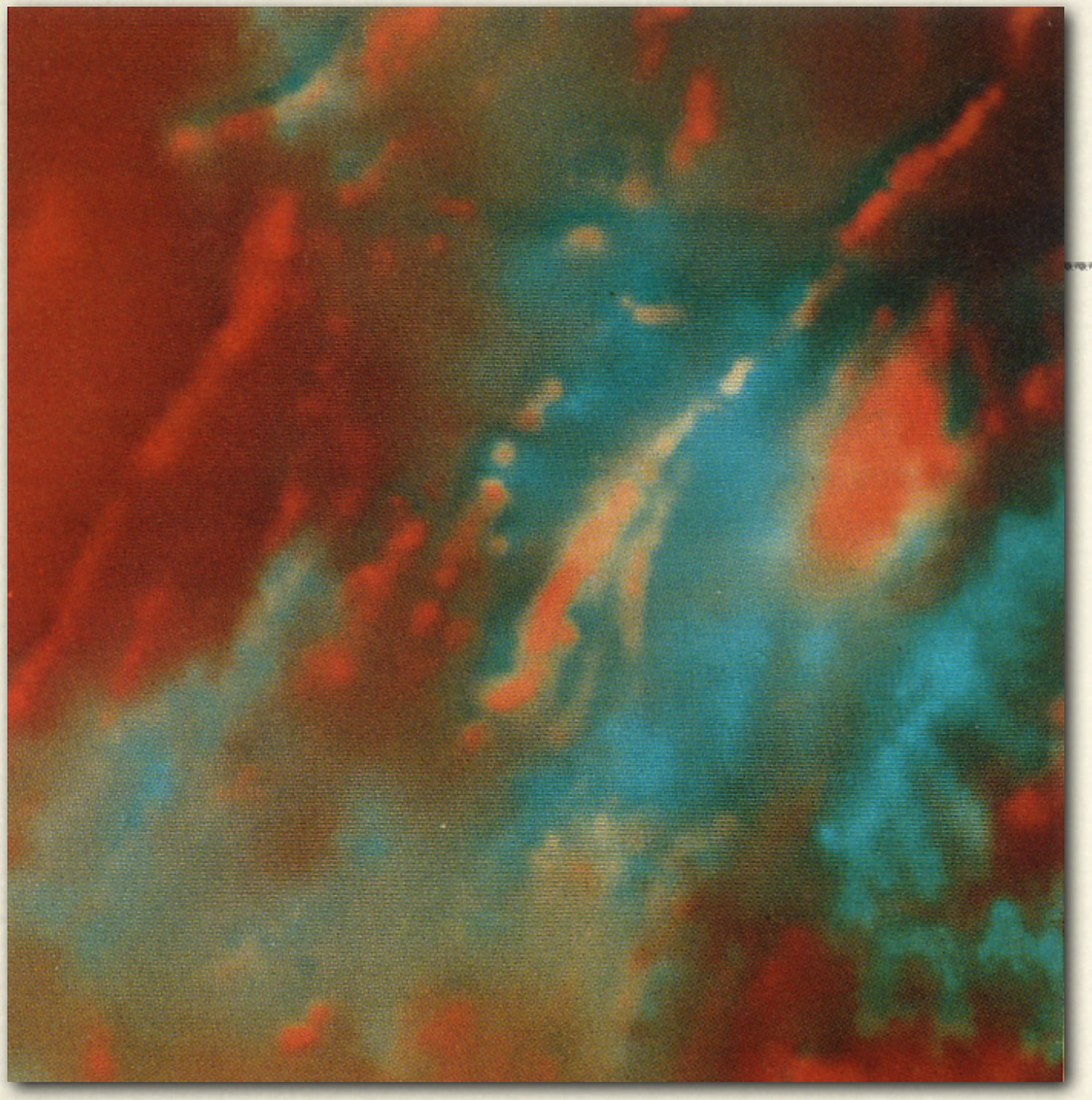
(from Tufte)

- layer
  - spa
  - var
  - lin
  - col
- colo
  - pse
  - 2D and 3D colorspaces

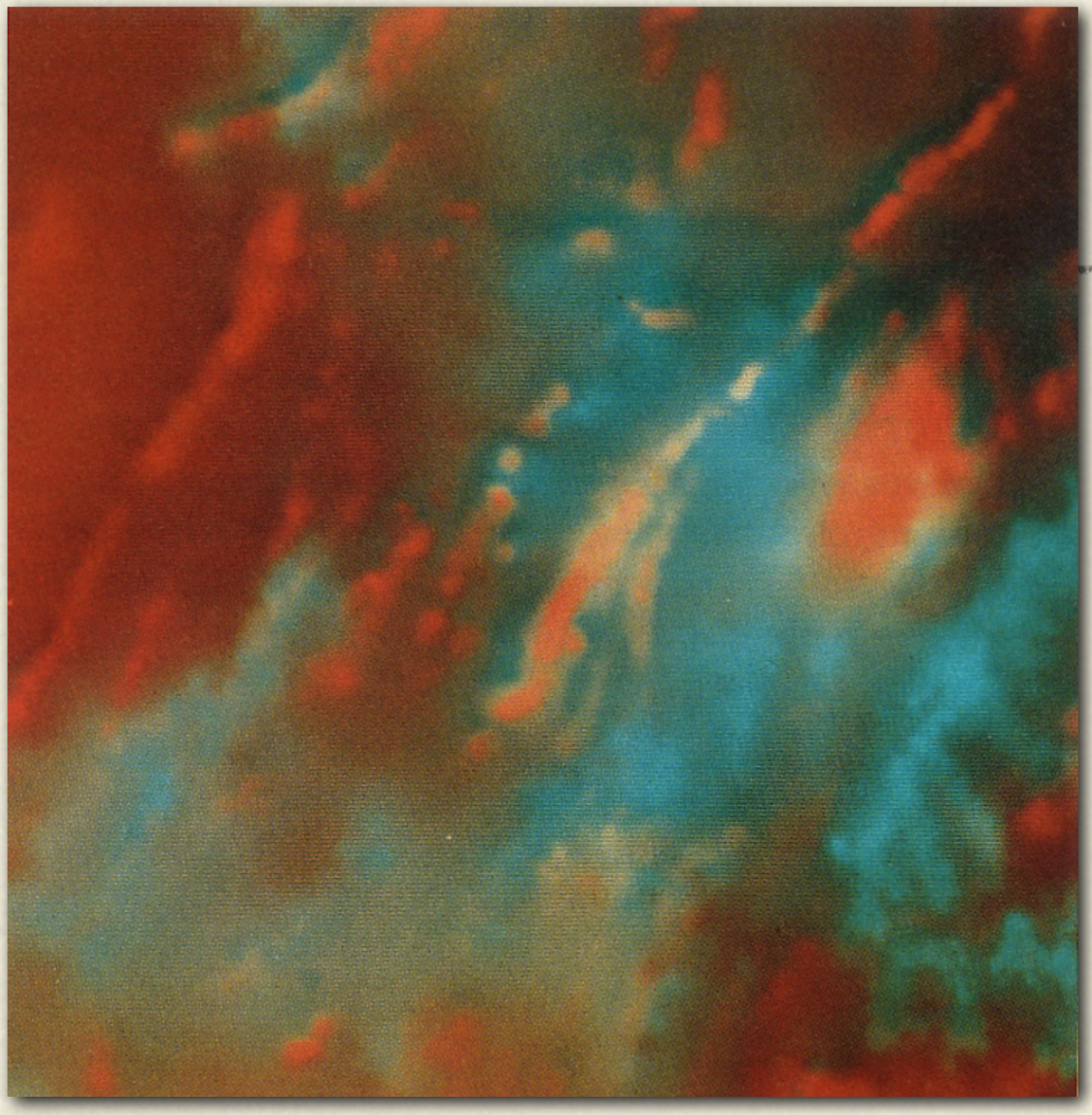


(from Tufte)

design

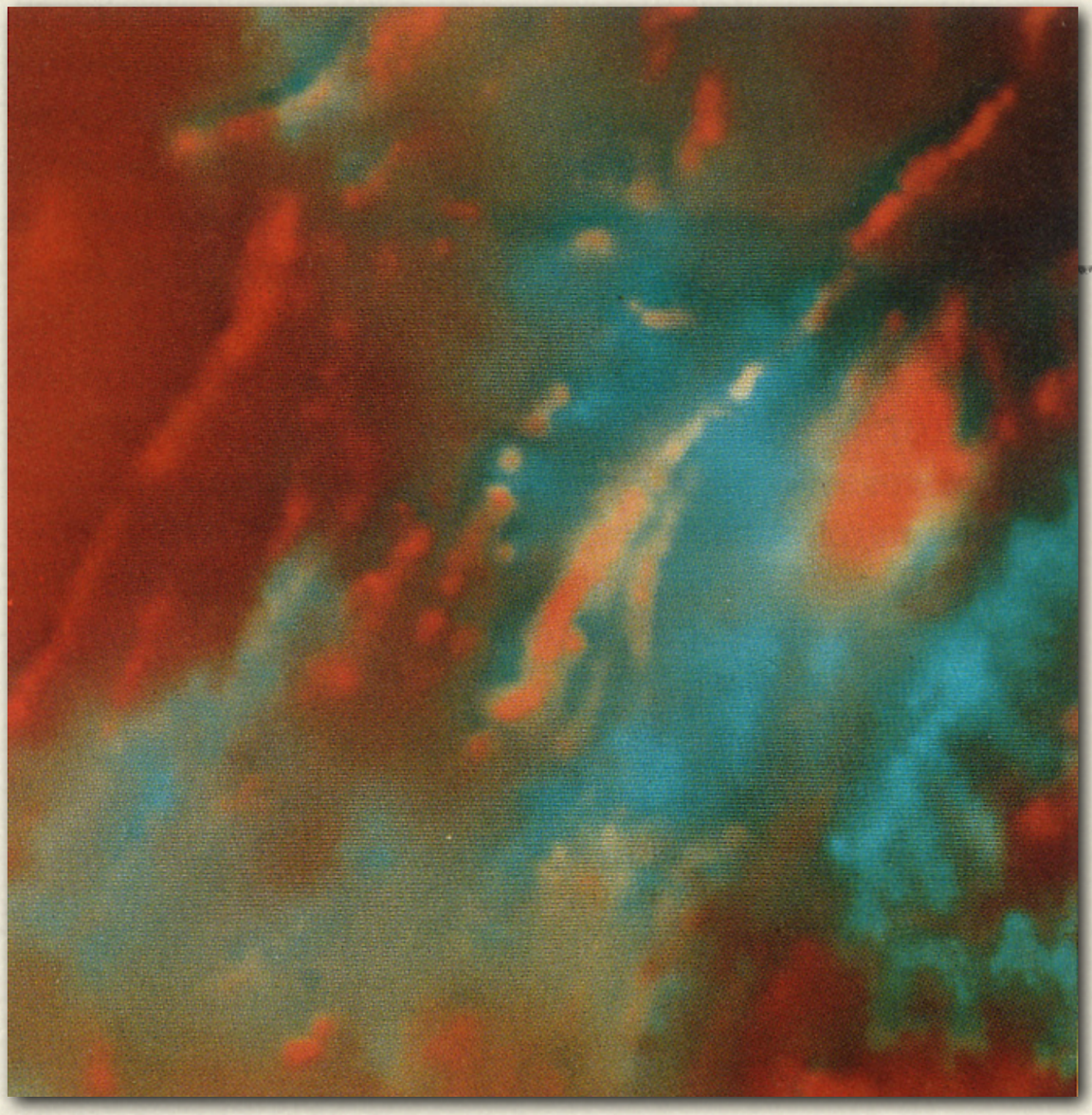


* 2D and 3D colorspaces



(P.K. Robertson)

* 2D and 3D colorspaces



(P.K. Robertson)

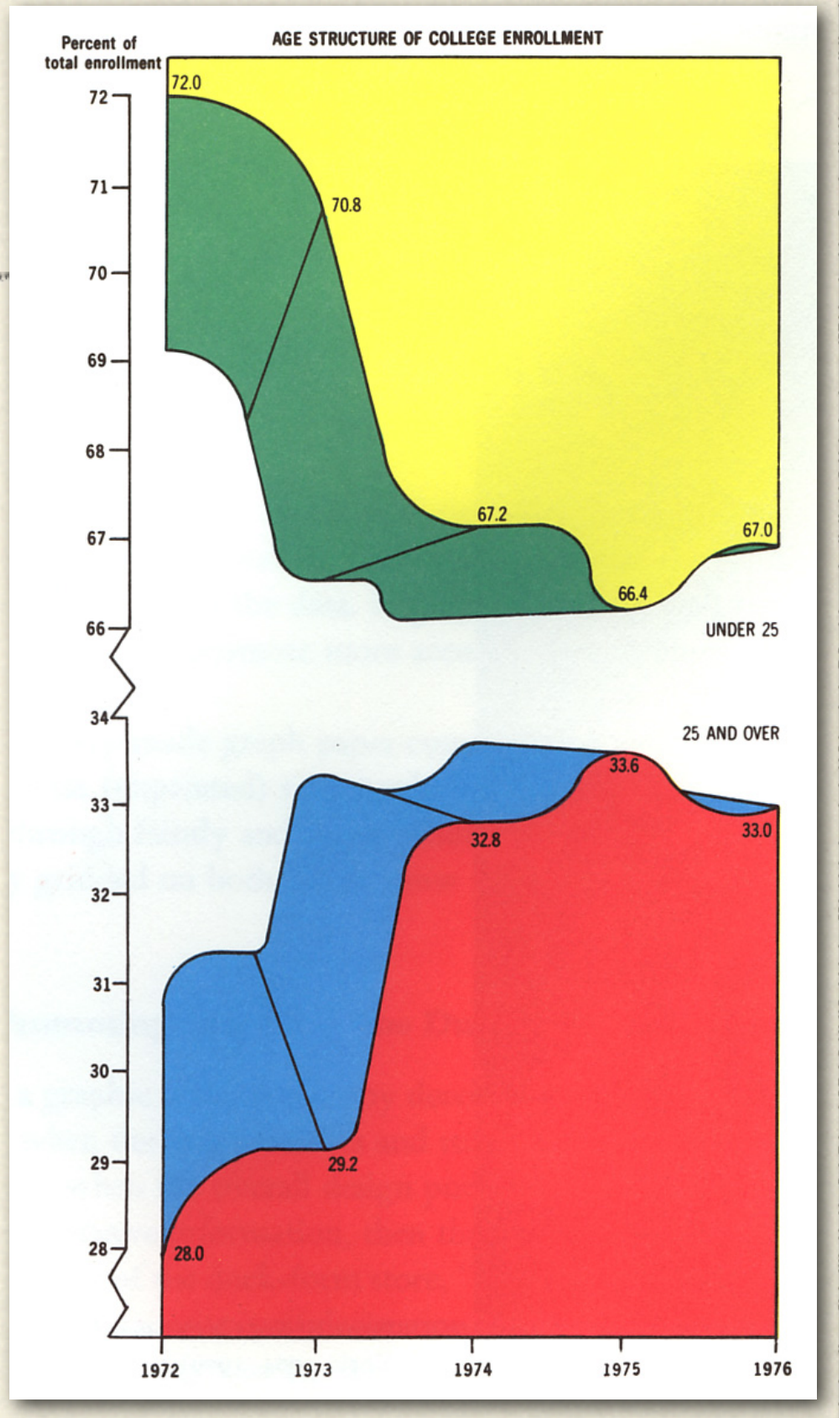
* 2D and 3D colorspaces
* other shading parameters

Things to avoid

- clutter
- chartjunk
- vibrating colors
- poor choice of scales
- inappropriate use of 3D
- etc.

Things to avoid

- clutter
- chartjunk
- vibrating colors
- poor choice of scales
- inappropriate use of 3D
- etc.



(from Tufte)

Can we automate the creation of visualizations?

Can we automate the creation of visualizations?

- yes
  - there are a limited number of datatypes and visual metaphors for displaying them
  - spectrum of solutions from menu-driven to completely automatic
  - Excel does it

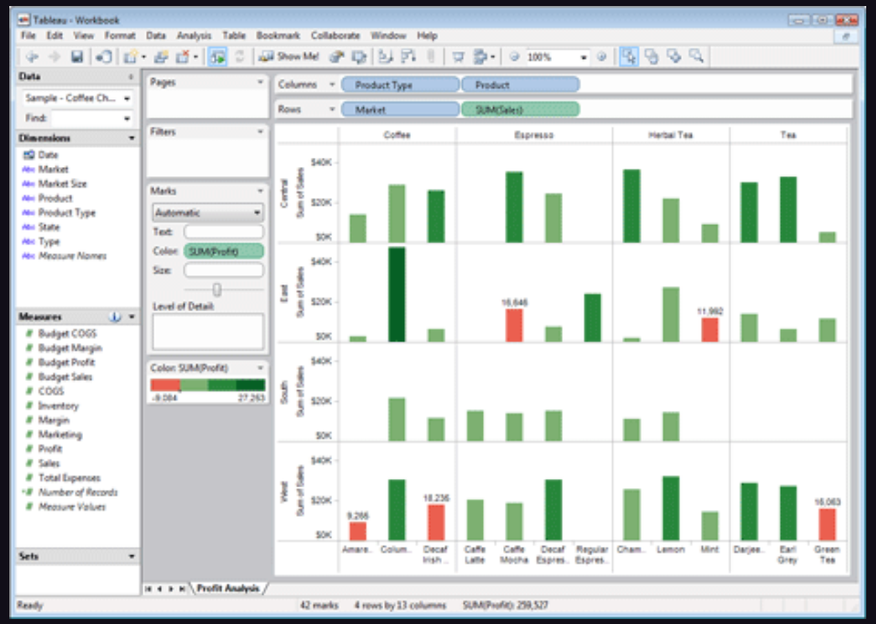
- no
  - effective visualizations must be designed
  - the design process itself leads to insight
  - Excel does it poorly

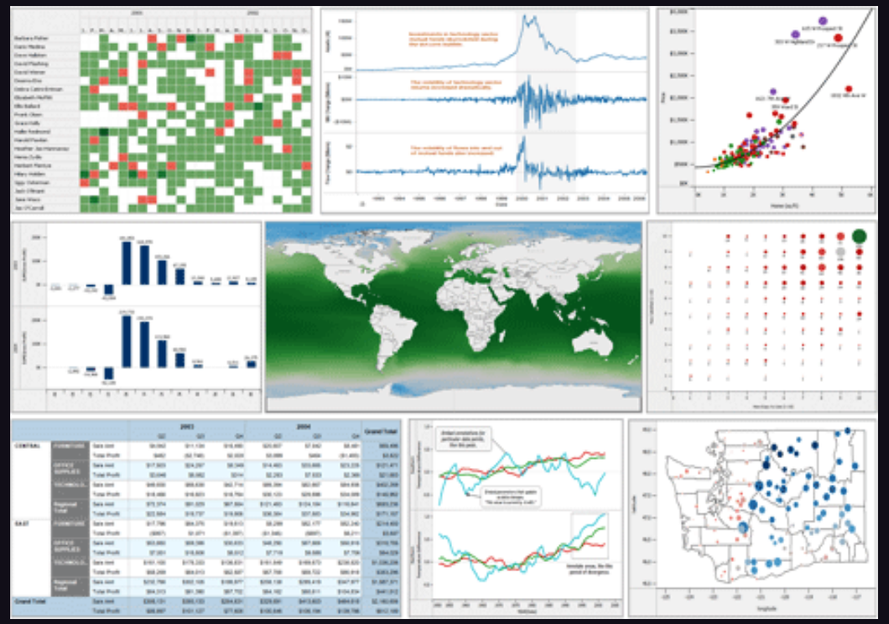
Taxonomies of visualization techniques

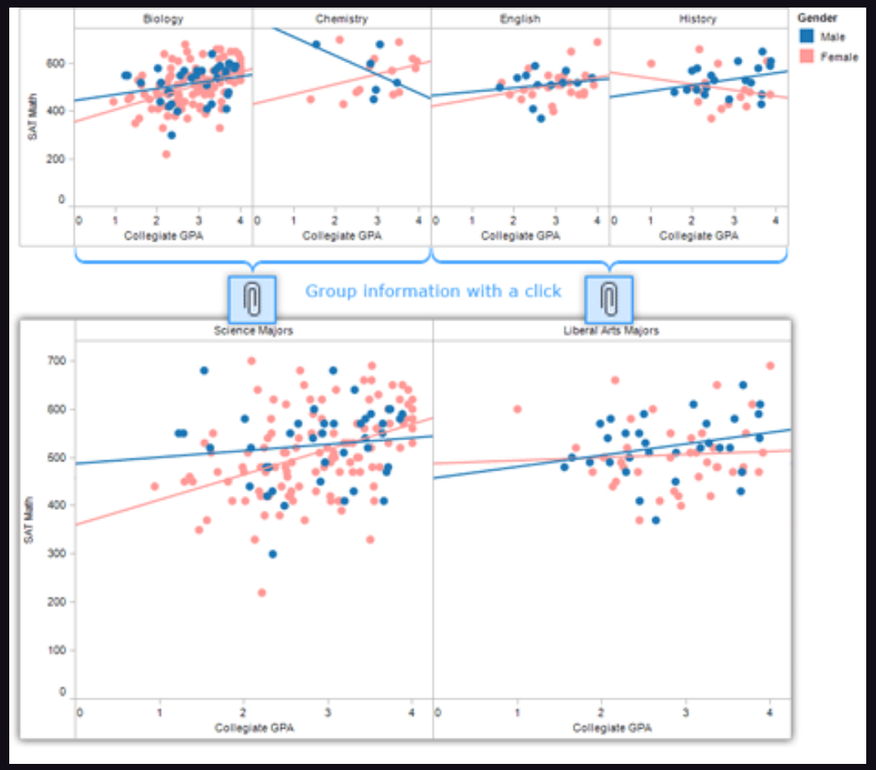
- by data type
  - number of independent variables (domain)
  - number of dependent variables (range)
  - discrete vrs continuous domain
  - binary vrs multivalued vrs continuous range

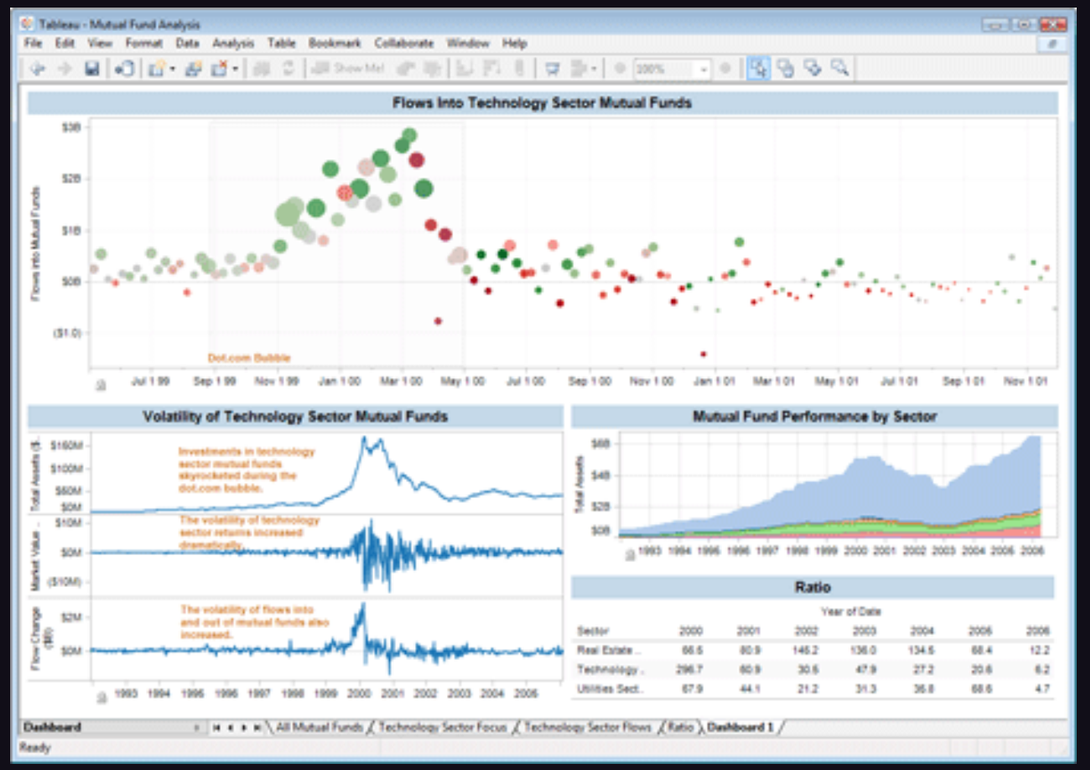
- by visual metaphor
  - Bertin's *Semiology*
  - Keller and Keller's *Visual Cues*

Tableau Software



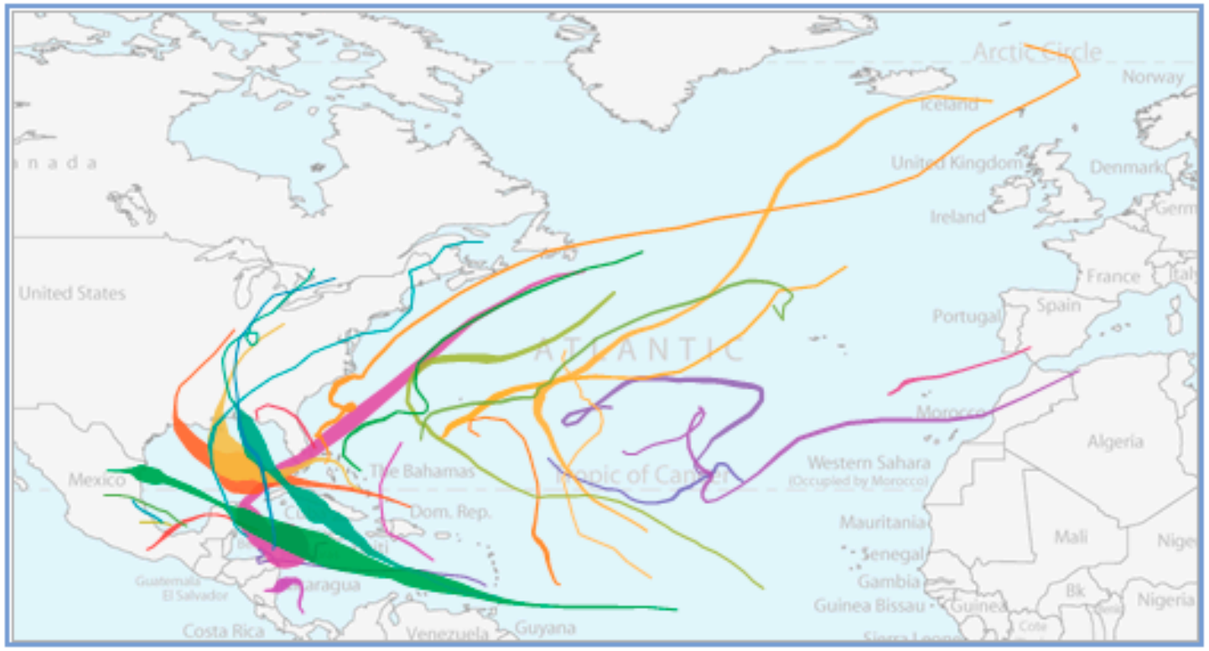






The 2005 Hurricane Season

Tableau Software



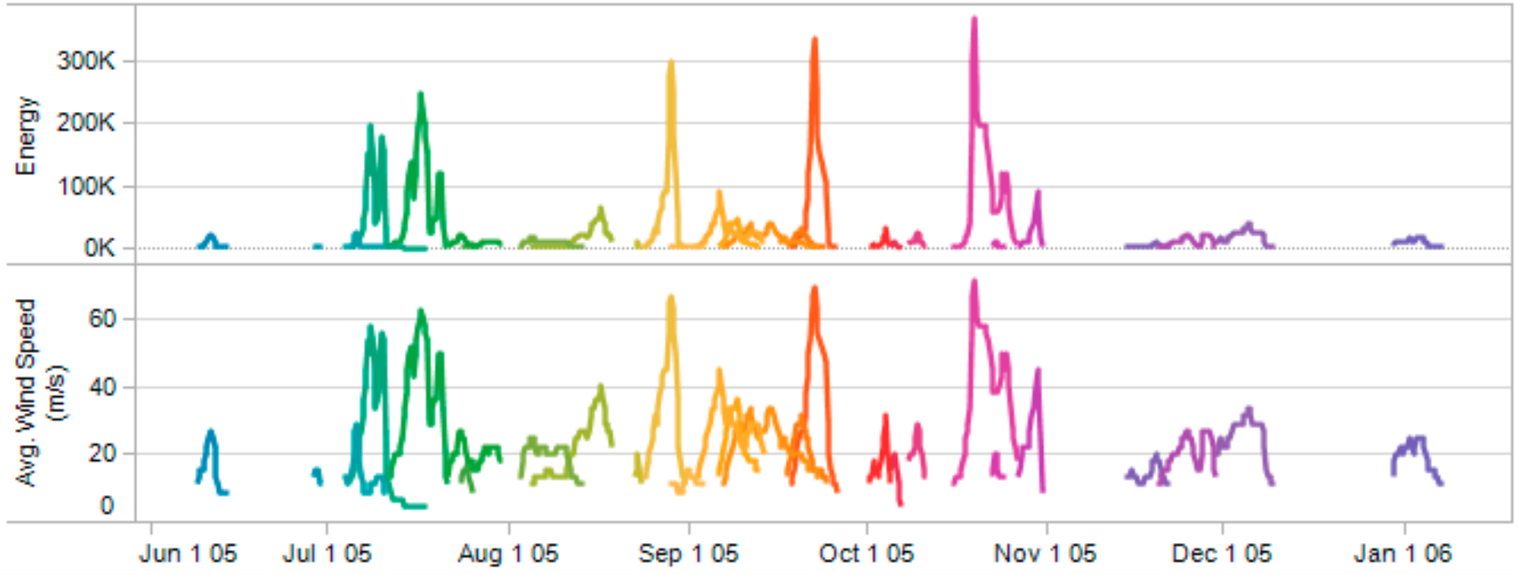
Select Basin:

- (•) Atlantic
- ( ) Eastern Pacific
- ( ) Western Pacific

Click to Highlight:

- ARLENE
- BRET
- CINDY
- DENNIS
- EMILY
- FRANKLIN
- GERT
- HARVEY
- IRENE
- JOSE
- KATRINA
- LEE
- MARIA
- NATE
- OPHELIA
- PHILIPPE
- RITA
- STAN
- TAMMY
- VINCE
- WILMA
- ALPHA
- BETA
- DELTA
- EPSILON
- GAMMA
- ZETA

Energy and Wind Speed by Time



Bertin's taxonomy

- imposition
  - diagrams
  - networks
  - maps
  - symbols

  ×

  - arrangement
  - rectilinear
  - circular
  - orthogonal
  - polar
- retinal variables
  - size
  - value
  - texture
  - color
  - orientation
  - shape

  ×

  - association
  - selection
  - order
  - quantity

Keller and Keller's taxonomy

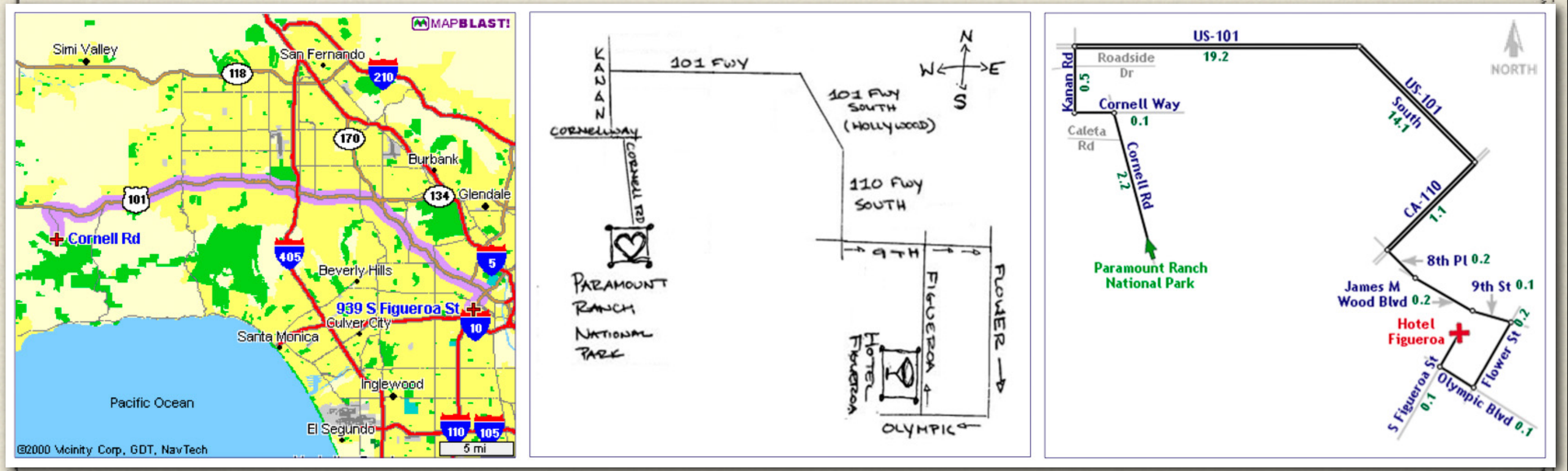
Actions	×	Data
identify		scalar
locate		nominal
distinguish		direction
categorize		shape
cluster		position
rank		region
compare		structure
associate		
correlate		

Functions of one variable

- 1D -> 1D
  - graph

Functions of one variable

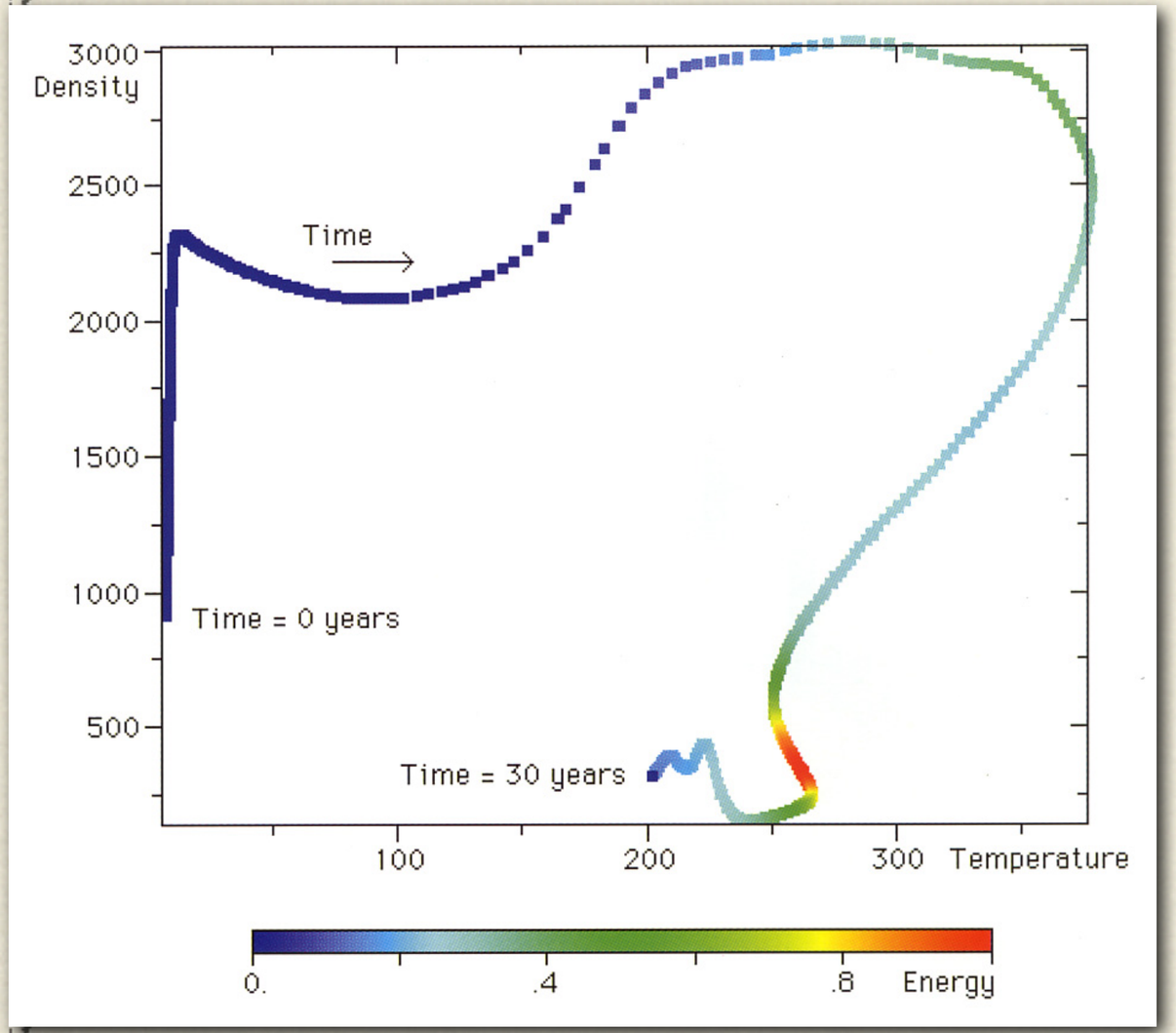
- 1D -> 1D
  - graph
- 1D -> 2D
  - parallel graphs
  - parametric line on the plane
  - animating arrow in the plane



(M. Agrawala)

- 1D -> 2D
  - parallel graphs
  - parametric line on the plane
  - animating arrow in the plane

Functions of one variable



- 1D -> 3D
  - more parallel graphs
  - colored parametric line on the plane
  - parametric line in space
  - animating 3-vector
  - etc.

(from Keller)

(FLASH DEMO)

http://graphics.stanford.edu/courses/cs178/applets/locus.html

Functions of two variables

- 2D -> 1D

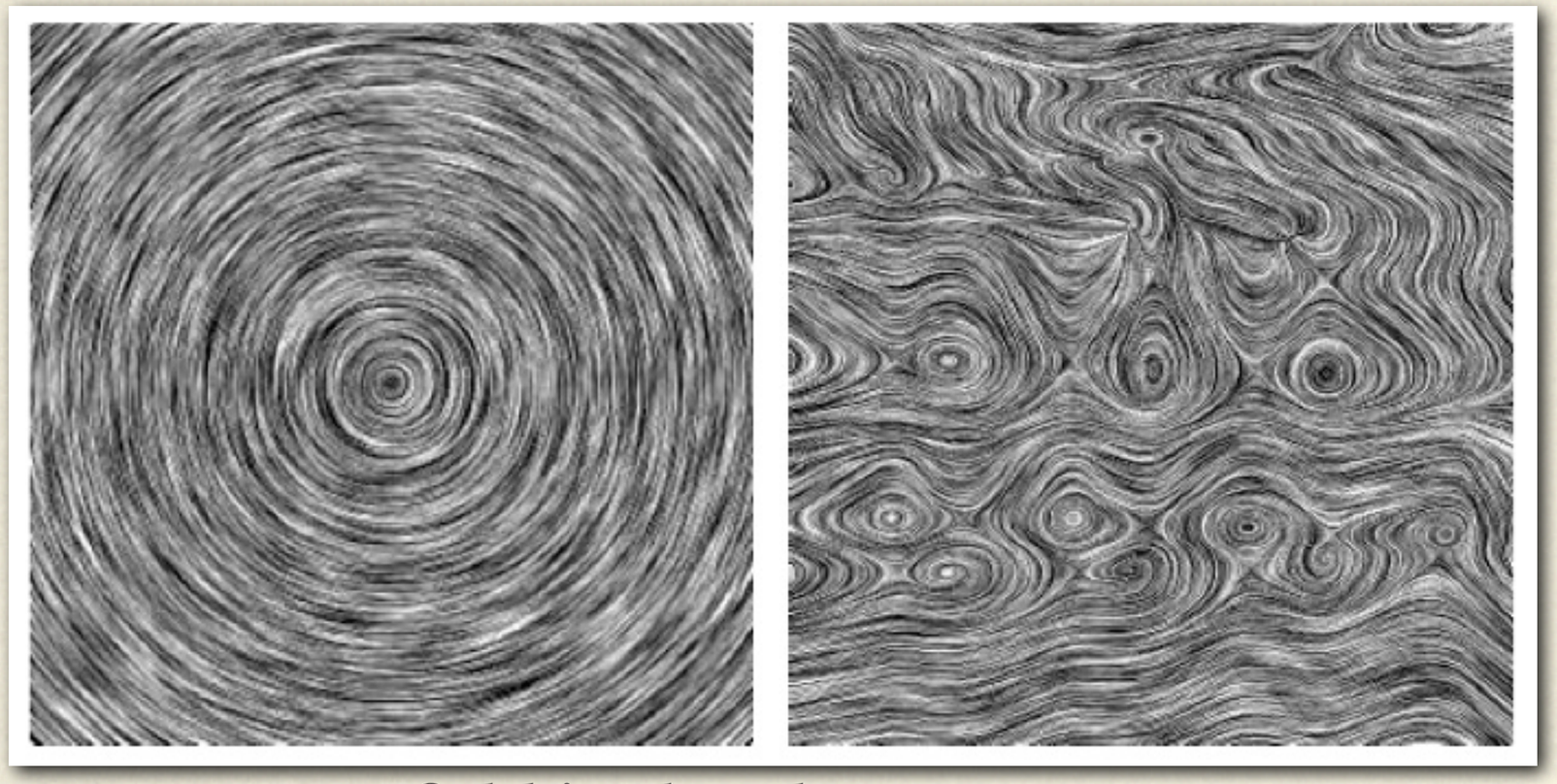
Functions of two variables

- 2D -> 1D
  - grayscale image
  - elevation map
  - animating graph

Functions of two variables

- 2D -> 1D
  - grayscale image
  - elevation map
  - animating graph
- 2D -> 2D
  - color image
  - grayscale + elevation map
  - vector field in the plane
  - parametric warp of the plane
  - animating parametric line in the plane

Functions of two variables



* vector field in the plane
* parametric warp of the plane
* animating parametric line in the plane

Functions of two variables

- 2D -> 1D
  - grayscale image
  - elevation map
  - animating graph
- 2D -> 2D
  - color image
  - grayscale + elevation map
  - vector field in the plane
  - parametric warp of the plane
  - animating parametric line in the plane
- 2D -> 3D
  - 3-vector field in the plane
  - parametric surface in space
  - etc.

Functions of two variables

- 2D -
  - * gra
  - * ele
  - * an
- 2D -
  - * col
  - * gra
  - * vec
  - * pa
  - * an plane

- 2D -> 3D
  - * 3-vector field in the plane
  - * parametric surface in space
  - * etc.

Functions of three variables

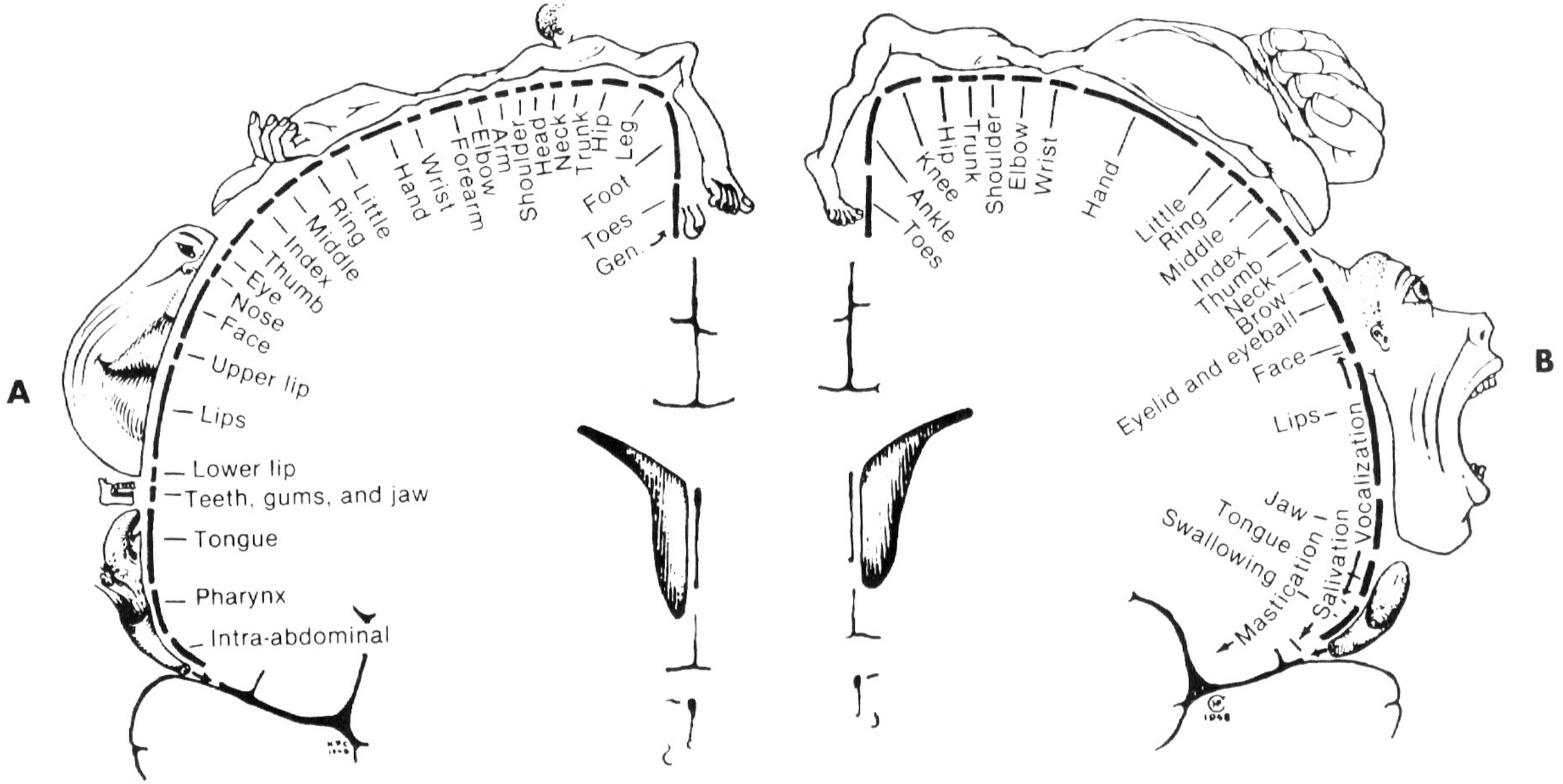
- 3D -> 1D
  - volume rendering
  - animating grayscale image
  - animating elevation map
- 3D -> 2D
  - color volume rendering
  - animating elevation + color map
  - animating vector field in the plane
  - animating parametric warp of the plane
- 3D -> 3D, 4D
  - 3-vector field in space
  - parametric warp of space
  - animating 3-vector field in the plane
  - animating parametric surface in space
  - etc.

Higher dimensions

- domain
  - multiples
  - animation
  - slices
  - level sets
  - projections
- range
  - selection
  - texture attributes
  - glyphs

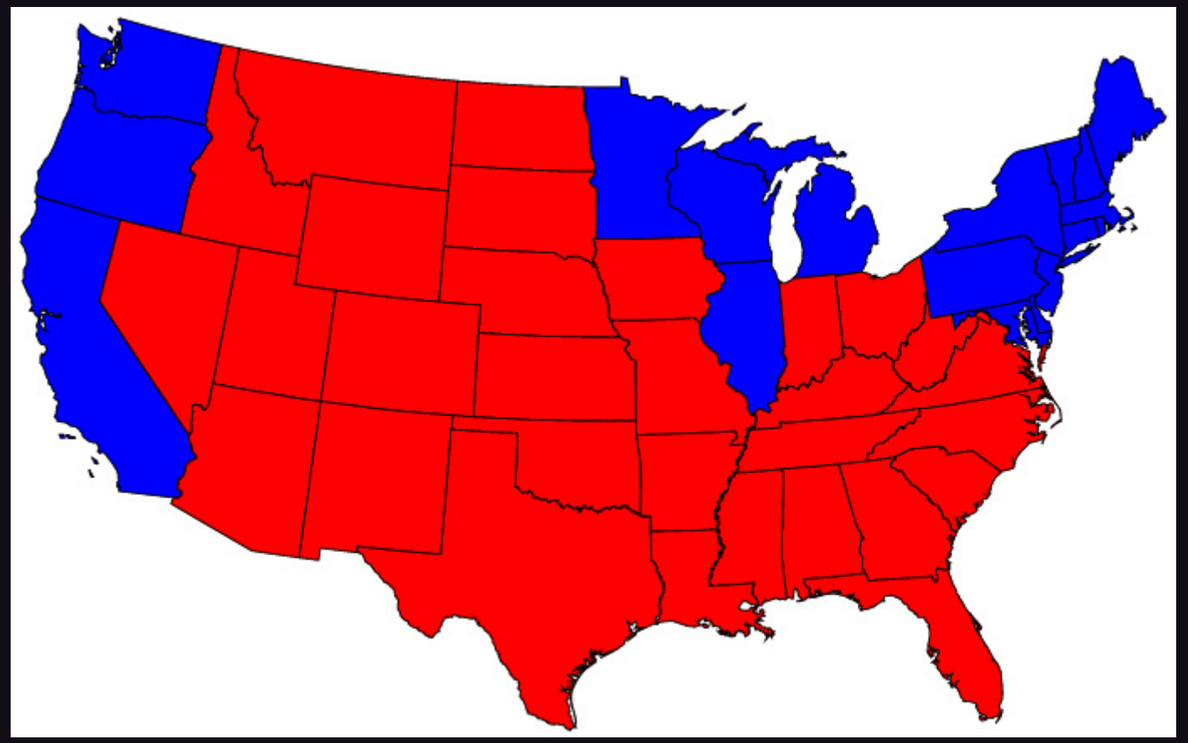
Discrete functions

- 1D -> 1D
  - bar graph
- 1D -> 2D
  - parallel bar graphs
  - row of arrows
- 2D -> 1D
  - table of glyphs
  - chartmaps
  - cartograms
  - animating bar graph
- 2D -> 2D
  - fancier glyphs
  - etc.

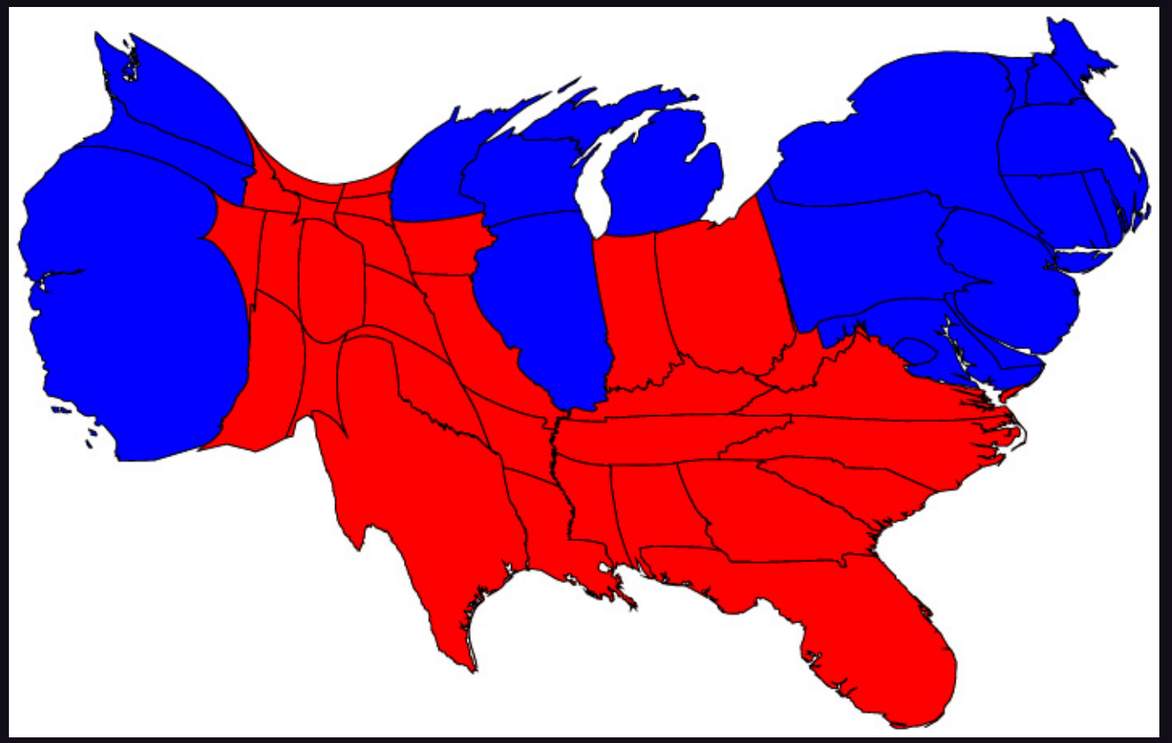


Wilder Penfield, sensory and motor humonculi (c. 1951)

Wilder Penfield, sensory and motor humonculi (c. 1951)



map



cartogram

Results of the U.S. 2004 presidential election

Other issues

- computational versus physical space
- displaying context and labeling
- displaying error bounds
- controlling error
- interaction

References

- H. Robin, *The Scientific Image*, 1994.
- Jacques Bertin, *The Semiology of Graphics*, 1967.
- John W. Tukey, *Exploratory Data Analysis*, 1977.
- Peter Keller and Mary Keller, *Visual Cues*, 1993.
- Richard Gallagher, *Computer Visualization*, 1995.
- William Schroeder et al., *The Visualization Toolkit*, 1996.
- Al Globus and E. Raible, *13 Ways to Say Nothing with Scientific Visualization*, 1992.