Using Space Effectively: 2D

Maneesh Agrawala

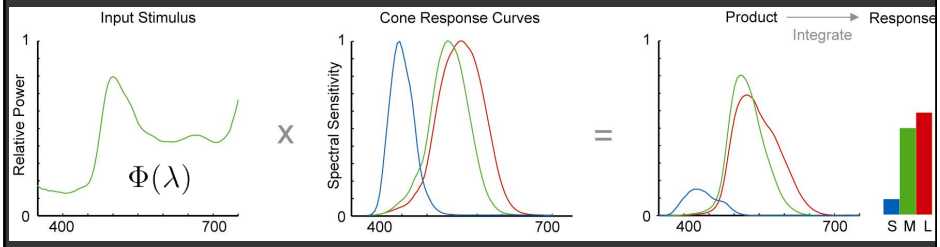
CS 294-10: Visualization
Fall 2013

Last Time: Color

What is Color?

Physical World	Visual System	Mental Models
Lights, surfaces, objects	Eye, optic nerve, visual cortex	Red, green, brown Bright, light, dark, vivid, colorful, dull Warm, cool, bold, blah, attractive, ugly, pleasant, jarring

Computing Cone Response



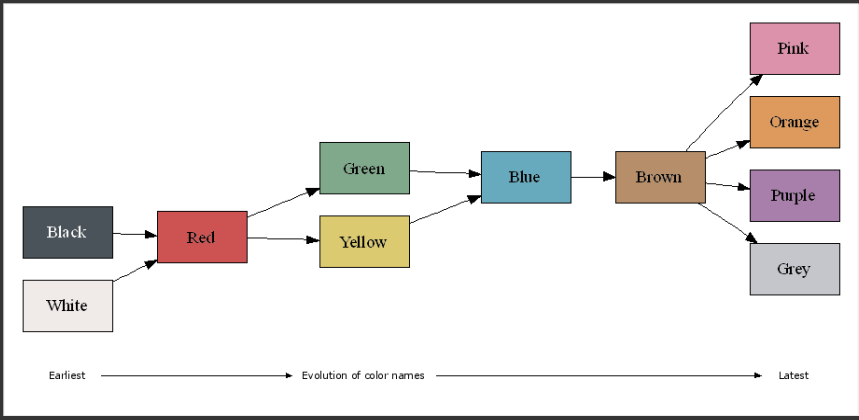
$$L = \int \Phi(\lambda) L(\lambda) \mathrm{d}\lambda$$

$$M = \int \Phi(\lambda) M(\lambda) \mathrm{d}\lambda$$

$$S = \int \Phi(\lambda) S(\lambda) \mathrm{d}\lambda$$

Evolution of Basic Color Terms

Proposed universal evolution across languages



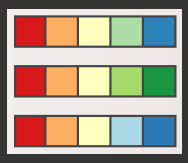
Quantitative color encoding

Sequential color scale

- Constrain hue, vary luminance/saturation
- Map higher values to darker colors

Diverging color scale

- Useful when data has a meaningful “midpoint”
- Use neutral color (e.g., grey) for midpoint
- Use saturated colors for endpoints



Limit number of steps in color to 3-9

Announcements

Assignment 3: Visualization Software

Create a **small** interactive visualization application – you choose data domain and visualization technique.

1. Describe data and storyboard interface
2. Implement interface and produce final writeup
3. Submit the application and a final writeup on the wiki

Can work alone or in pairs

Final write up due before class on **Oct 16, 2013**

Final project

Design new visualization method

- Pose problem, Implement creative solution

Deliverables

- Implementation of solution
- 8-12 page paper in format of conference paper submission
- 1 or 2 design discussion presentations

Schedule

- Project proposal: 10/28
- Project presentation: 11/11-11/13
- Final paper and presentation: 12/2-12/6

Grading

- Groups of up to 3 people, graded individually
- Clearly report responsibilities of each member

Using Space Effectively: 2D

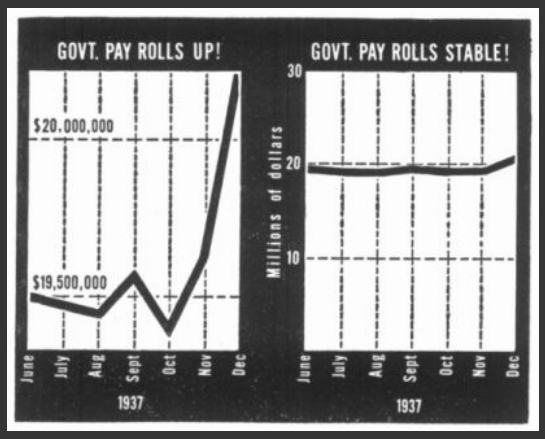
Topics

Displaying data in graphs
Banking to 45 degrees
Fitting data and depicting residuals
Graphical calculations
Zooming and distortion

Graphs and Lines

Effective use of space

Which graph is better?

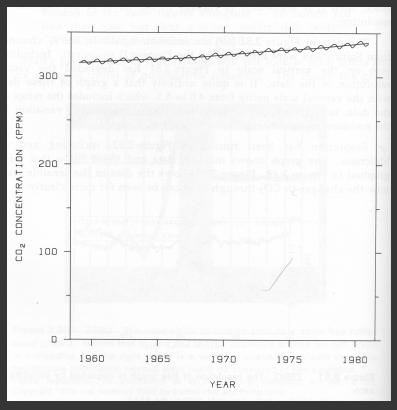


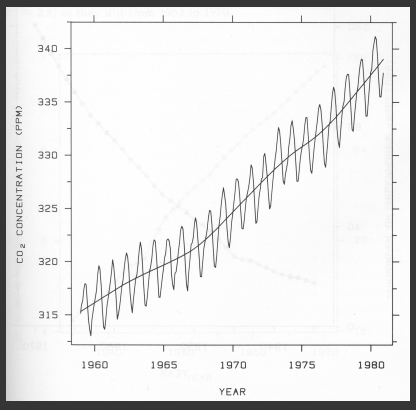
Government payrolls in 1937 [Huff 93]

Aspect ratio

Fill space with data

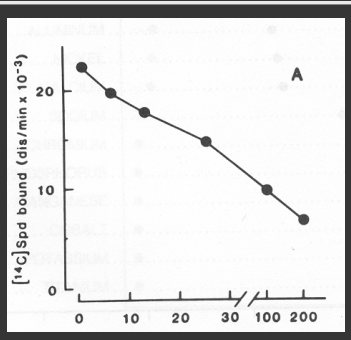
Don't worry about showing zero



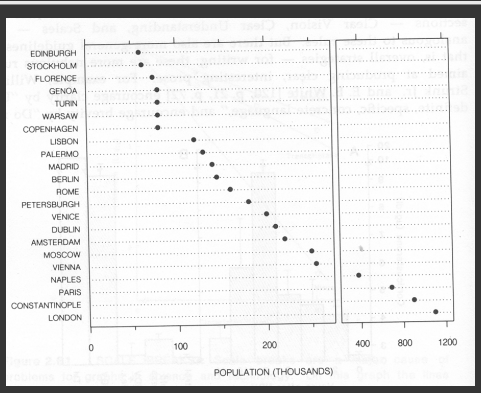


Yearly CO2 concentrations [Cleveland 85]

Clearly mark scale breaks

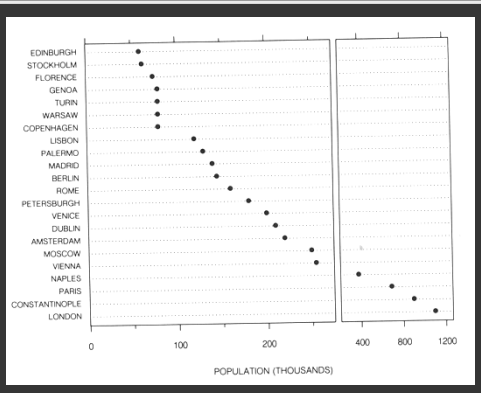


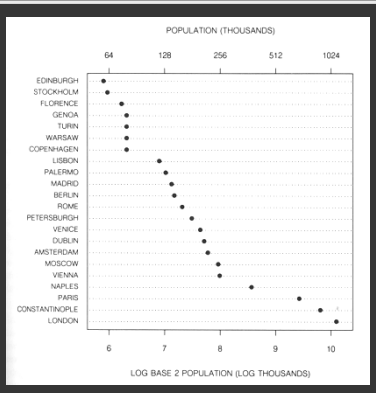
Poor scale break [Cleveland 85]



Well marked scale break [Cleveland 85]

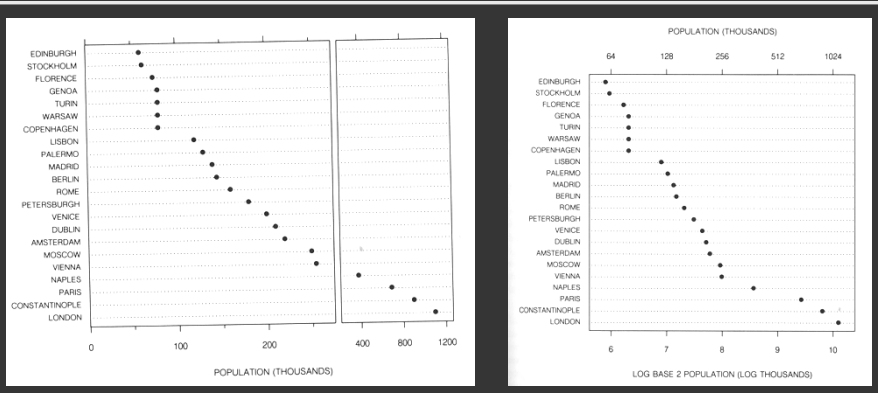
Scale break vs. Log scale





[Cleveland 85]

Scale break vs. Log scale

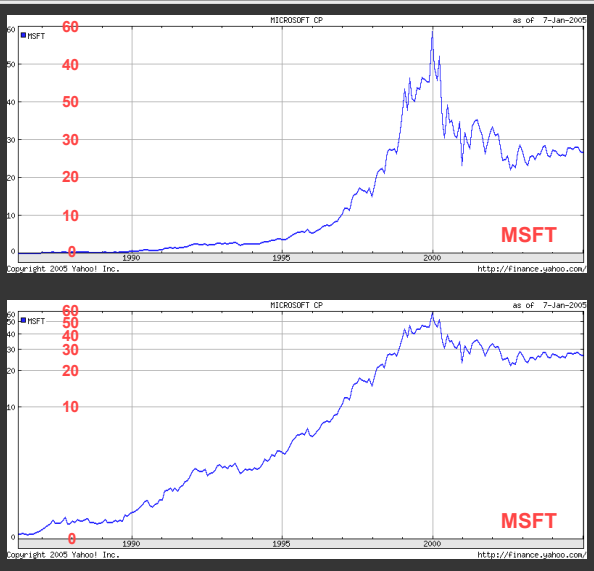


[Cleveland 85]

Both increase visual resolution

- **Log scale - easy comparisons of all data**
- **Scale break – more difficult to compare across break**

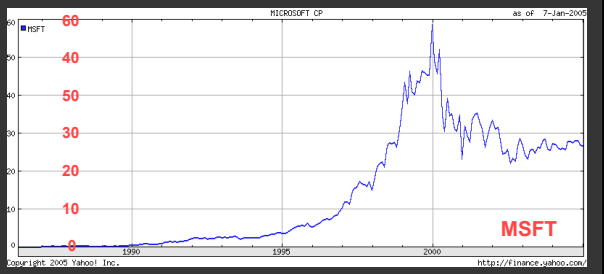
Linear scale vs. Log scale



Linear scale vs. Log scale

Linear scale

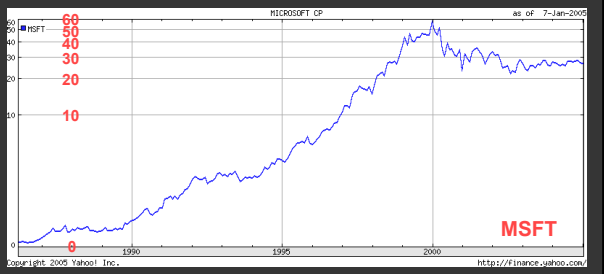
- Absolute change



Log scale

- Small fluctuations
- Percent change

  d(10,20) = d(30,60)



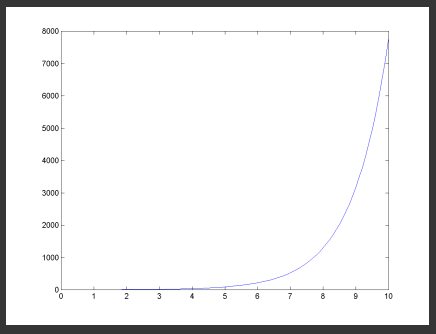
Semilog graph: Exponential growth

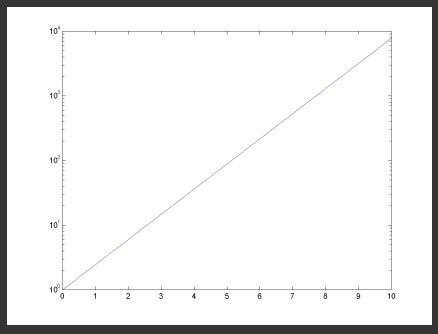
Exponential functions ($y = ka^{mx}$) transform into lines

$\log(y) = \log(k) + \log(a)mx$

Intercept: $\log(k)$

Slope: $\log(a)m$





$y = 6^{0.5x}$, slope in semilog space: $\log(6)*0.5 = 0.3891$

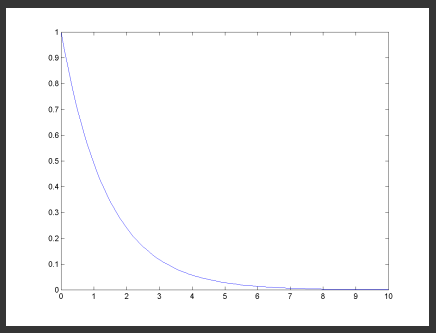
Semilog graph: Exponential decay

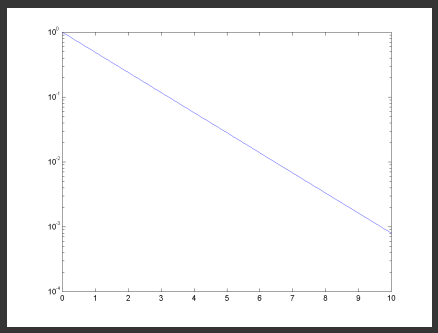
Exponential functions ($y = ka^{mx}$) transform into lines

$\log(y) = \log(k) + \log(a)mx$

Intercept: $\log(k)$

Slope: $\log(a)m$





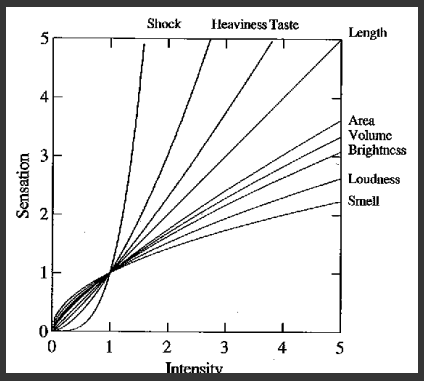
$y = 0.5^{2x}$, slope in semilog space: $\log(0.5)*2 = -0.602$

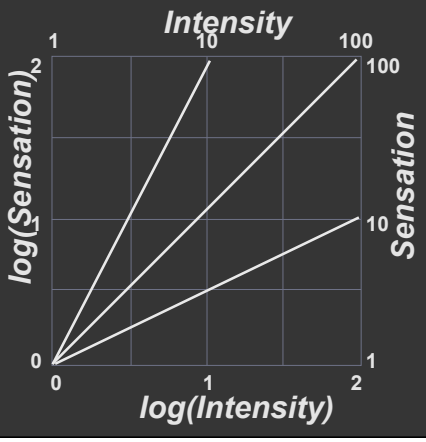
Log-Log graph

Power functions ($y = kx^{a}$) transform into lines

Example - Steven's power laws:

$S = kI^{p} \rightarrow \log S = \log k + p \log I$



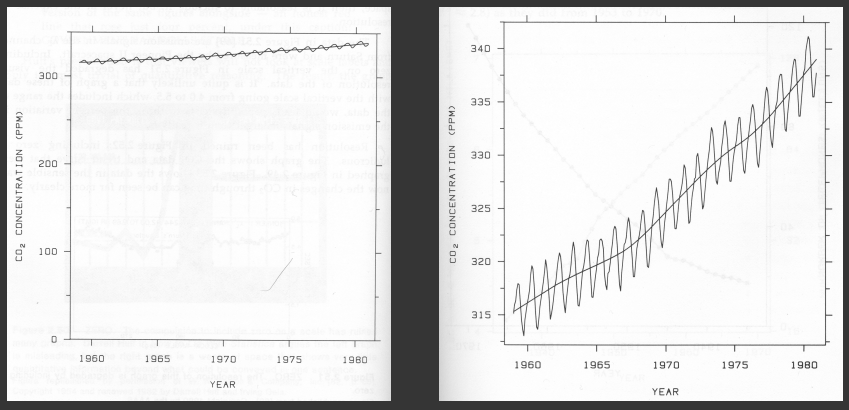


Banking to 45 Degrees

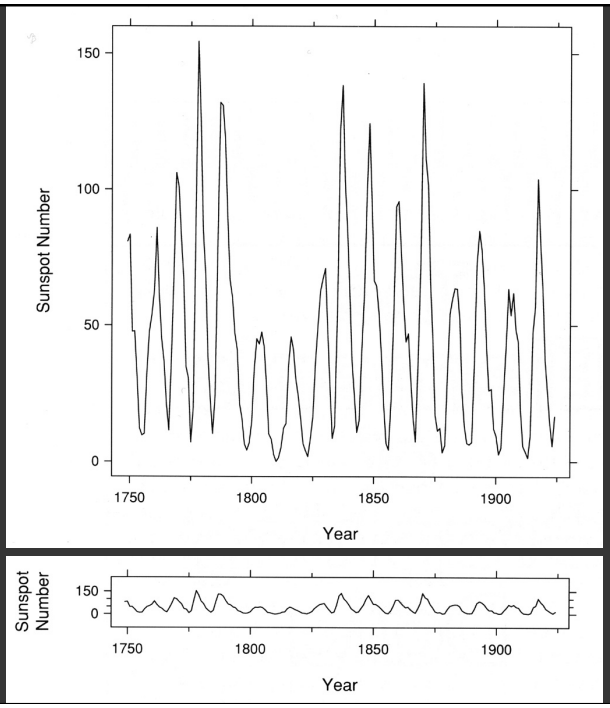
Aspect ratio

Fill space with data

Don't worry about showing zero



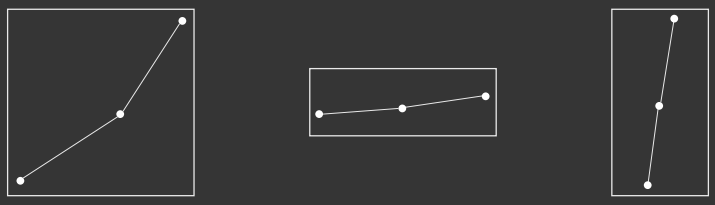
Yearly CO2 concentrations [Cleveland 85]



William S. Cleveland
The Elements of Graphing Data

Banking to 45° [Cleveland]

To facilitate perception of trends, maximize the discriminability of line segment orientations



Two line segments are maximally discriminable when their average absolute angle is 45°

Optimize the *aspect ratio* to bank to 45°

Aspect-ratio banking techniques

Median-Absolute-Slope

$$\alpha = \text{median} \mid s_i \mid R_x / R_y$$

Average-Absolute-Slope

$$\alpha = \text{mean} \mid s_i \mid R_x / R_y$$

Has Closed Form Solution

Average-Absolute-Orientation

Unweighted

$$\sum_i \frac{\mid \theta_i(\alpha) \mid}{n} = 45°$$

Weighted

$$\frac{\sum_i \mid \theta_i(\alpha) \mid l_i(\alpha)}{\sum_i l_i(\alpha)} = 45°$$

Max-Orientation-Resolution

Global (over all i, j s.t. i≠j)

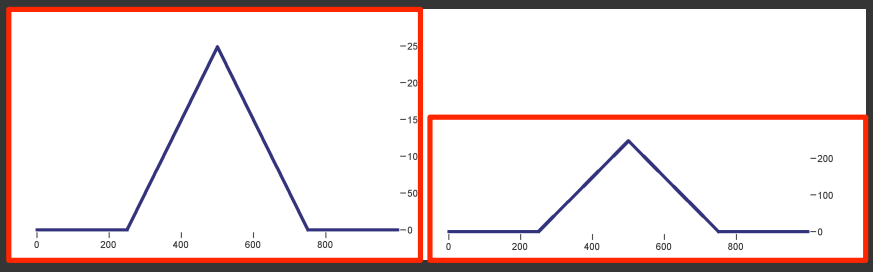
$$\sum_i \sum_j \mid \theta_i(\alpha) - \theta_j(\alpha) \mid^2$$

Local (over adjacent segments)

$$\sum_i \mid \theta_i(\alpha) - \theta_{i+1}(\alpha) \mid^2$$

Requires Iterative Optimization

Slopeless line culling



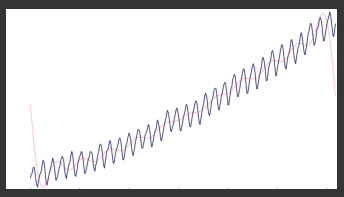
Standard, Aspect Ratio = 1.97

Culled, Aspect Ratio = 4.00

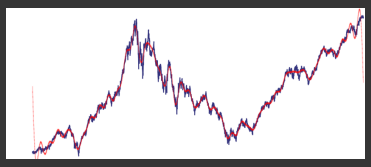
Exclude line segments with zero or infinite slope

Comparison data sets

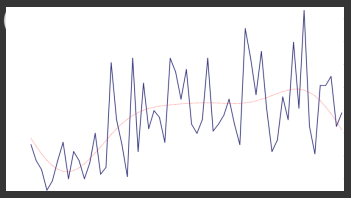
CO_2 Measurements (co2)



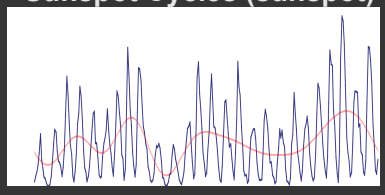
PRMTX Mutual Fund (prmtx)



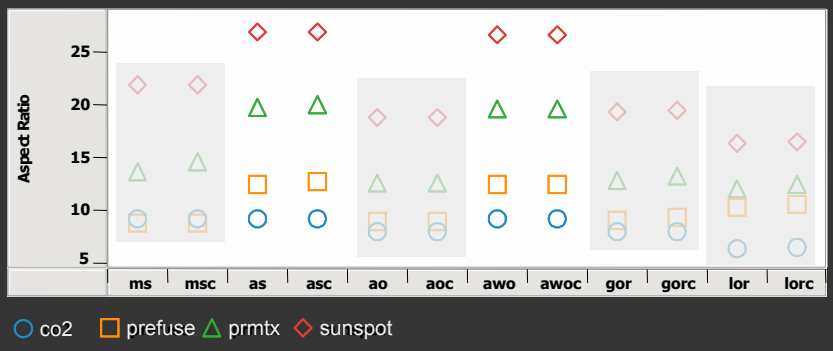
Prefuse Downloads



Sunspot Cycles (sunspot)

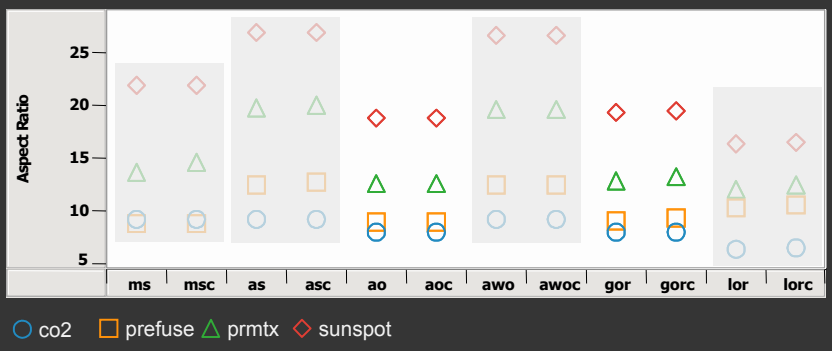


Comparison (Results)



Average-slope (as) and Average-weighted-orientation (awo) provide similar ratios

Comparison (Results)



Average-orientation (ao) and Global-orientation-resolution (gor) provide similar ratios

Discussion

Due to computational complexity…

Prefer avg-slope to avg-weighted-orient

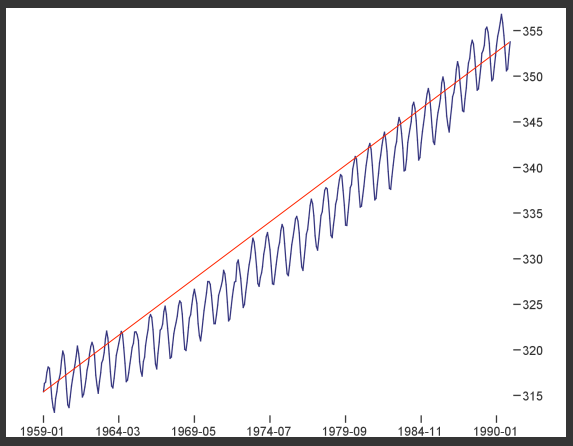
Prefer avg-orient to global-orient-resolution

But due to perceptual effectiveness… ?

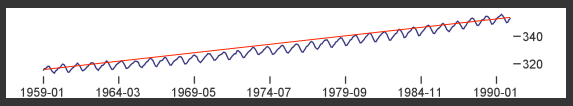
Cleveland recommends weighted-avg-orient

But, goal is to maximize discriminability

Perceptual experiments needed to clarify



Aspect Ratio = 1.17



Aspect Ratio = 7.87

CO_2 Measurements

William S. Cleveland
Visualizing Data

Multi-Scale Banking to 45°

Goal

Optimized aspect ratios for varying scales

Approach

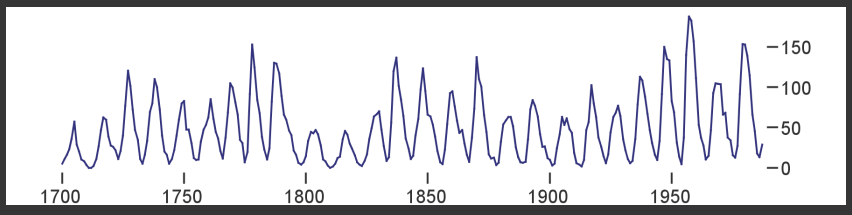
Identify Scales of Interest
Generate Scale-Specific Trend Lines
Bank Trend Lines to 45°
Filter Resulting Aspect Ratios

Multi-Scale Banking to 45°

Use Spectral Analysis to identify trends

- Find strong frequency components
- Lowpass filter to create trend lines

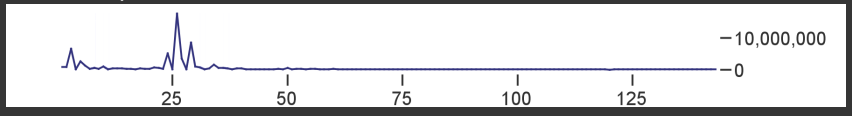
Compute Power Spectrum



Take Discrete Fourier Transform

Compute squared magnitudes

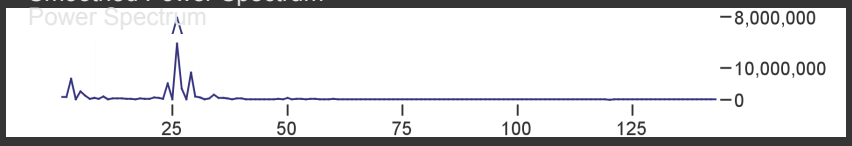
Power Spectrum



Smooth the Spectrum

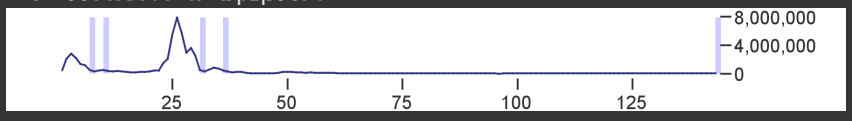
Convolve with Gaussian filter
window size = 3, $\sigma = 1$

Smoothed Power Spectrum

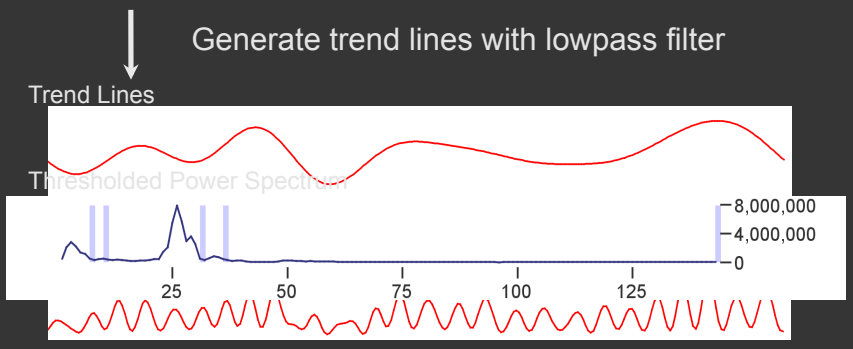


Threshold the Spectrum

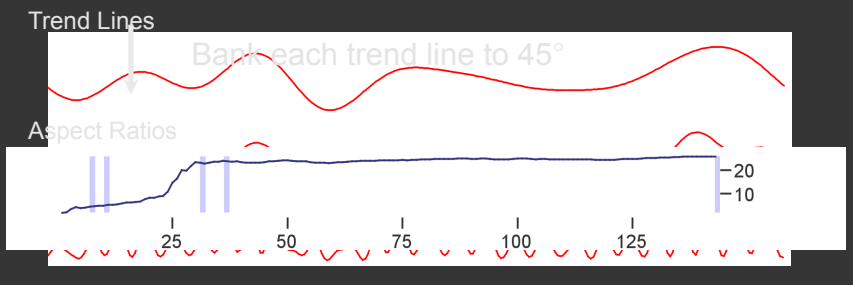
Threshold at mean of power spectrum
Retain last values of contiguous runs



Generate Trend Lines



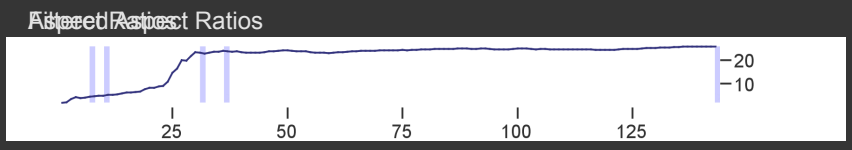
Bank Trend Lines to 45°



Filter Aspect Ratios

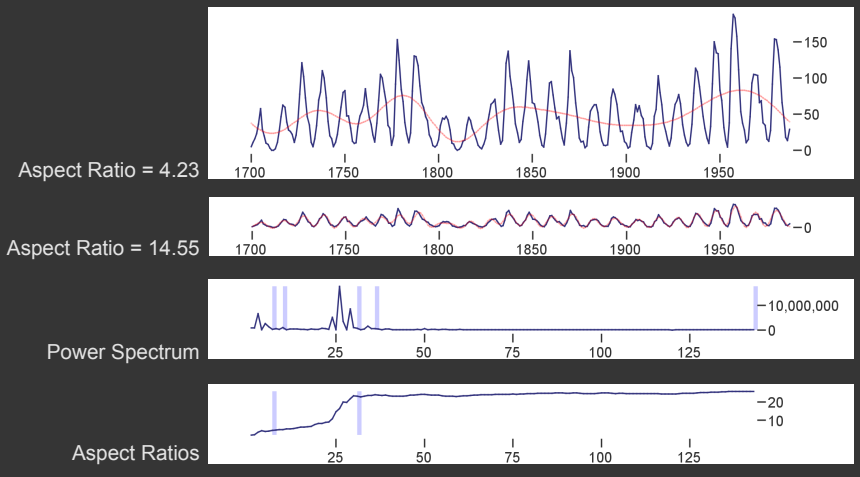
Filter similar aspect ratios

Keep if $\alpha_{i+1} > c\alpha_i$ (c=1.25 by default)



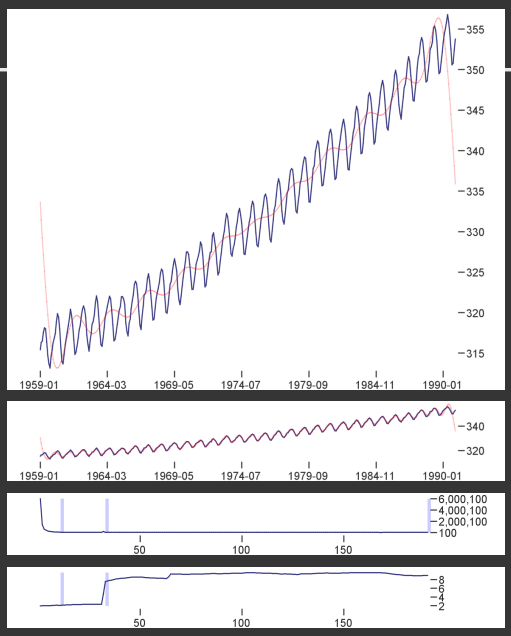
Sunspot Cycles

Yearly values 1700-1987



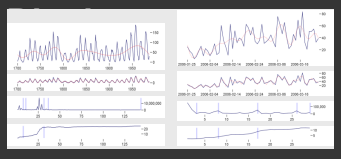
CO_2

Monthly concentrations from the Mauna Loa Observatory, 1950-1990



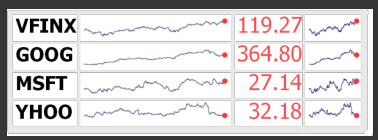
Applications

Small Multiples

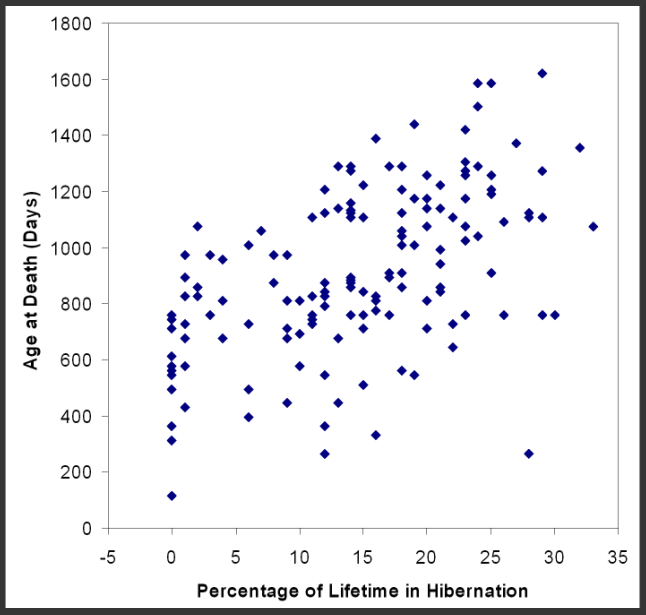


Trend Explorer

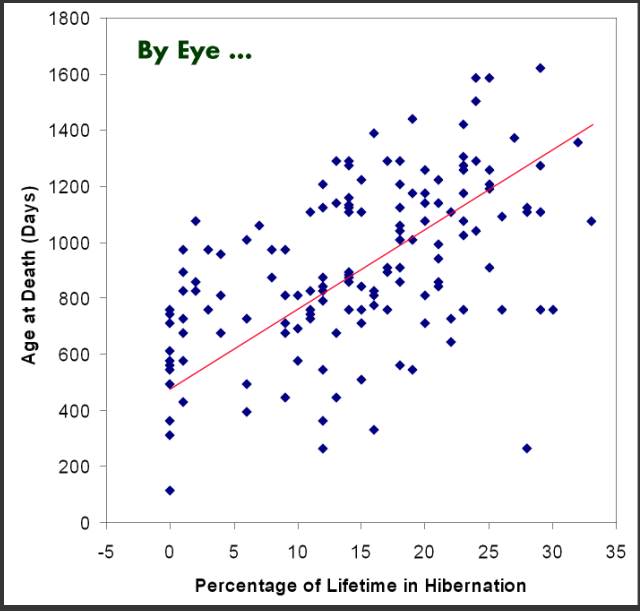
Sparklines



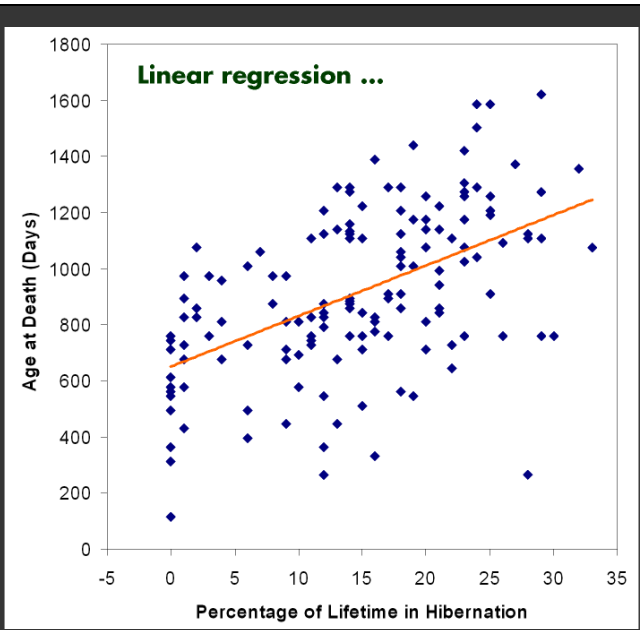
Fitting the Data



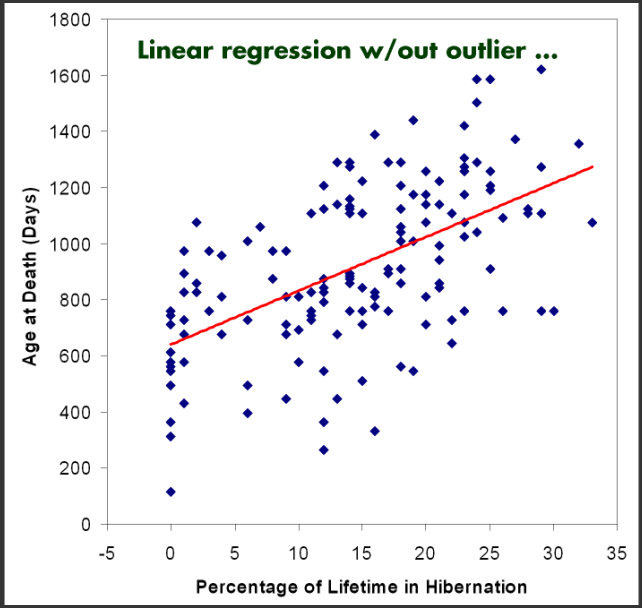
[The Elements of Graphing Data. Cleveland 94]



[The Elements of Graphing Data. Cleveland 94]



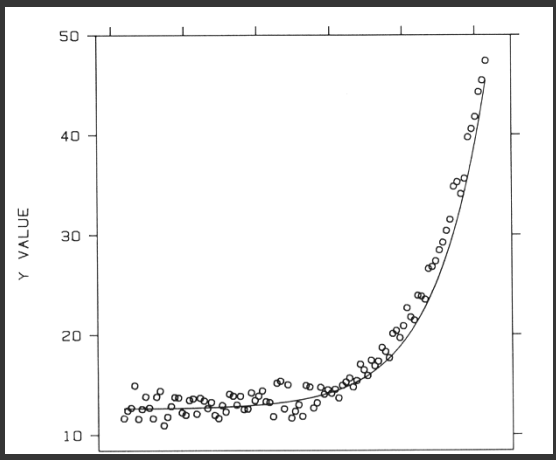
[The Elements of Graphing Data. Cleveland 94]



[The Elements of Graphing Data. Cleveland 94]

Transforming data

How well does curve fit data?

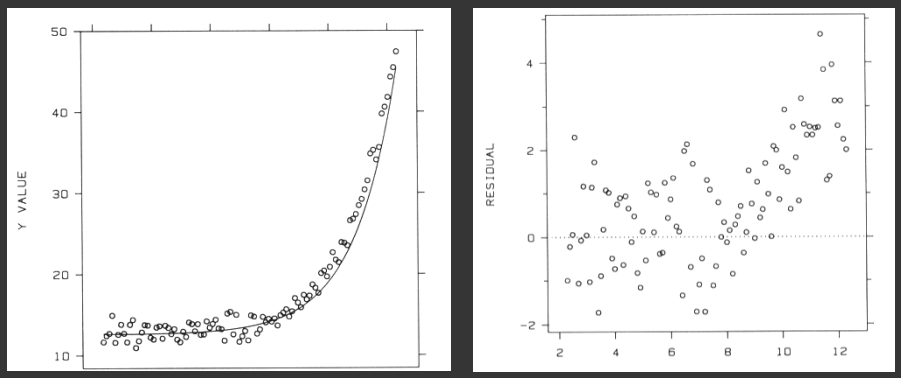


[Cleveland 85]

Transforming data

Residual graph

- Plot vertical distance from best fit curve
- Residual graph shows accuracy of fit



[Cleveland 85]

Most powerful brain?

Microsoft Excel - animal.xls

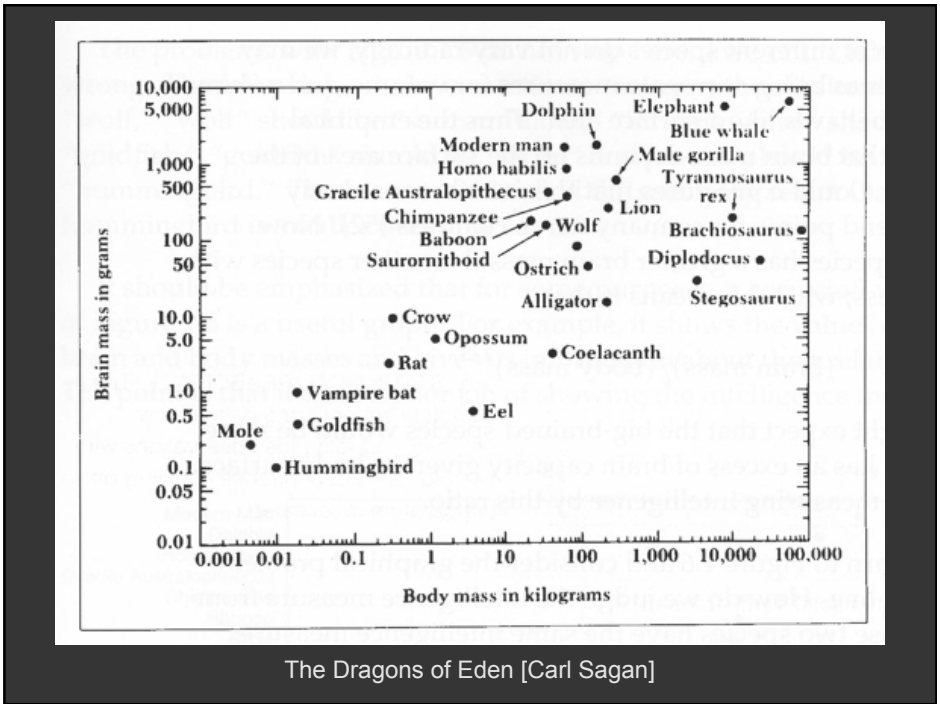
File Edit View Insert Format Tools Data Window Help

A1 ID

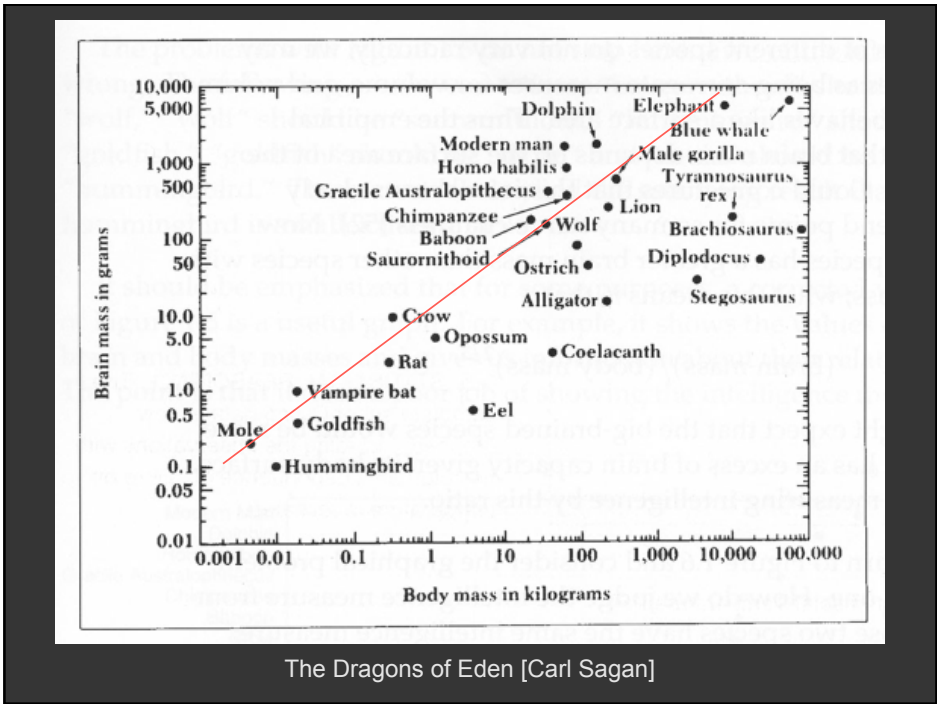
	A	B	C	D	E
1	ID	Name	Body Weight	Brain Weight	
2	1	Lesser Short-tailed Shrew	5	0.14	
3	2	Little Brown Bat	10	0.25	
4	3	Mouse	23	0.3	
5	4	Big Brown Bat	23	0.4	
6	5	Musk Shrew	48	0.33	
7	6	Star Nosed Mole	60	1	
8	7	Eastern American Mole	75	1.2	
9	8	Ground Squirrel	101	4	
10	9	Tree Shrew	104	2.5	
11	10	Golden Hamster	120	1	
12	11	Mole Rate	122	3	
13	12	Galago	200	5	
14	13	Rat	280	1.9	
15	14	Chinchilla	425	6.4	
16	15	Desert Hedgehog	550	2.4	
17	16	Rock Hyrax (a)	750	12.3	
18	17	European Hedgehog	785	3.5	
19	18	Tenrec	900	2.6	
20	19	Arctic Ground Squirrel	920	5.7	
21	20	African Giant Pouched Rat	1000	6.6	
22	21	Guinea Pig	1040	5.5	
23	22	Mountain Beaver	1350	8.1	
24	23	Slow Loris	1400	12.5	
25	24	Genet	1410	17.5	
26	25	Phalanger	1620	11.4	

animal

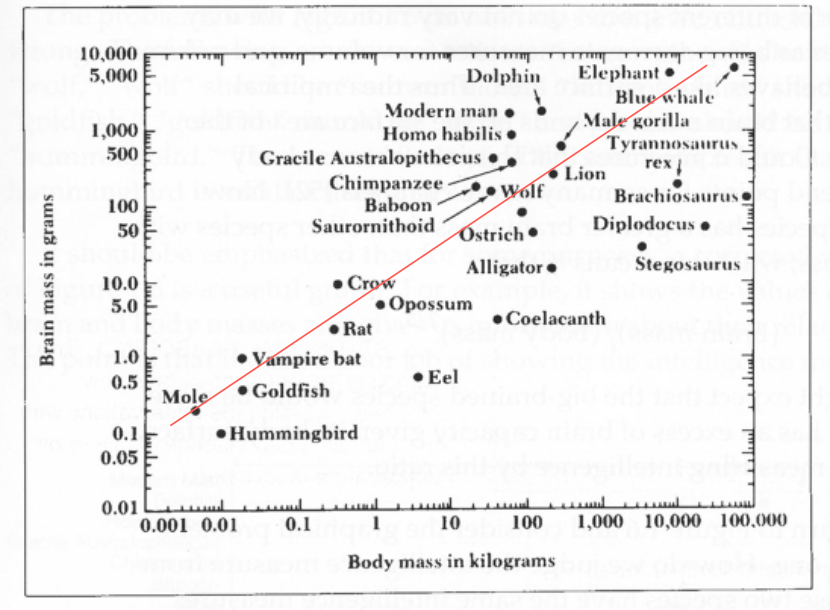
Ready



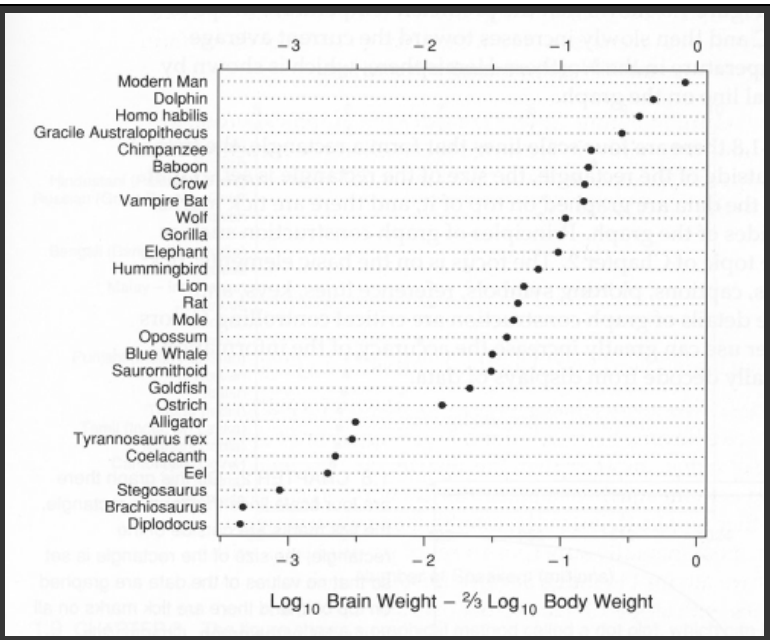
The Dragons of Eden [Carl Sagan]



The Dragons of Eden [Carl Sagan]



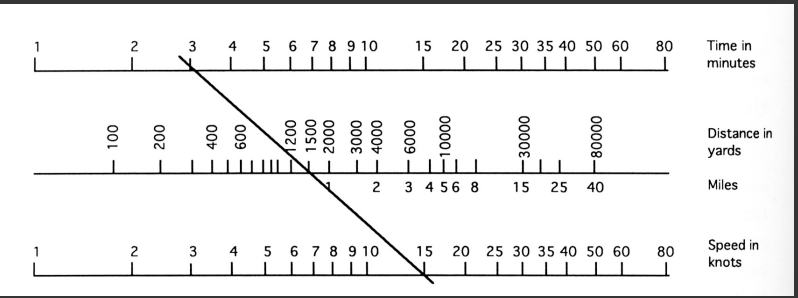
The Dragons of Eden [Carl Sagan]



The Elements of Graphing Data [Cleveland]

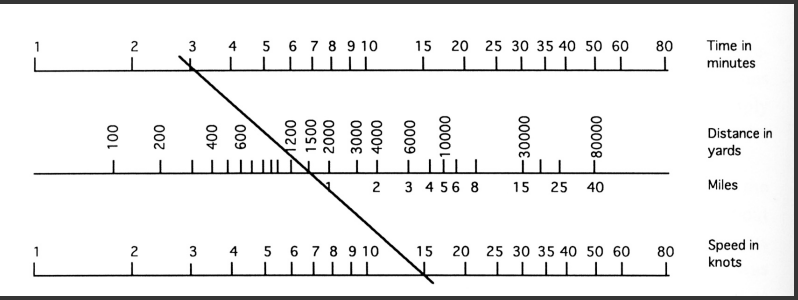
Graphical Calculations

Nomograms



Sailing: The Rule of Three

Nomograms

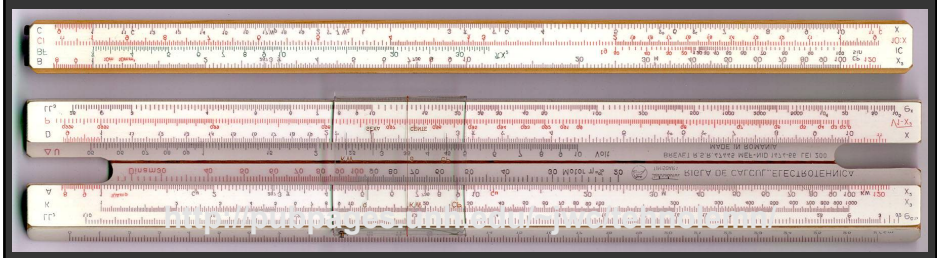


1. Compute in any direction; fix n-1 params and read nth param
2. Illustrate sensitivity to perturbation of inputs
3. Clearly show domain of validity of computation

Theory

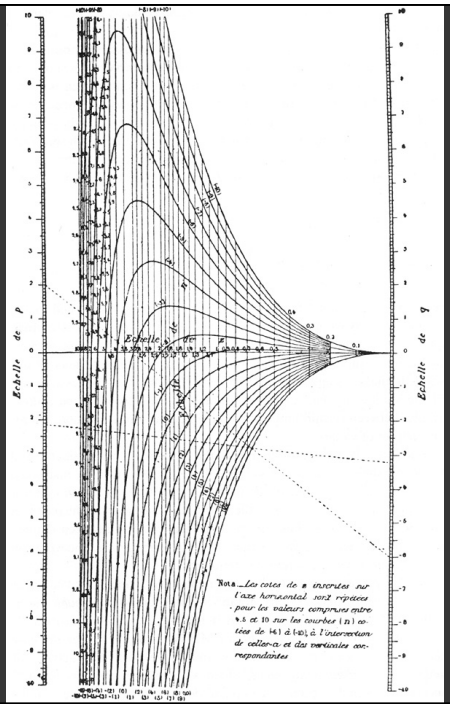
$$\begin{vmatrix} x_1(u) & y_1(u) & w_1(u) \\ x_2(v) & y_2(v) & w_2(v) \\ x_3(s,t) & y_3(s,t) & w_3(s,t) \end{vmatrix} = 0$$

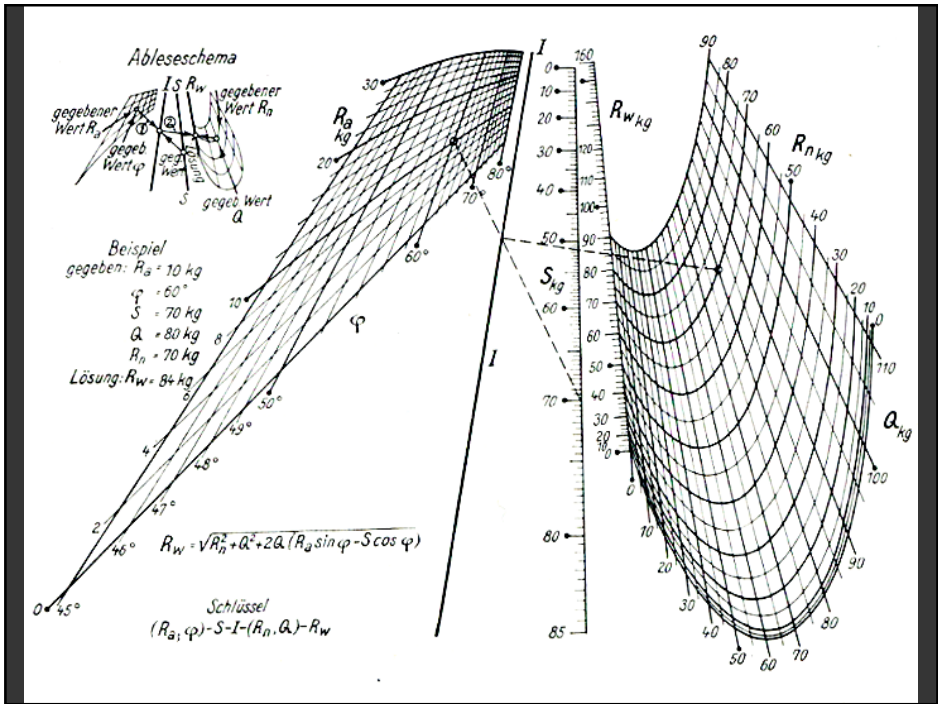
Slide rule



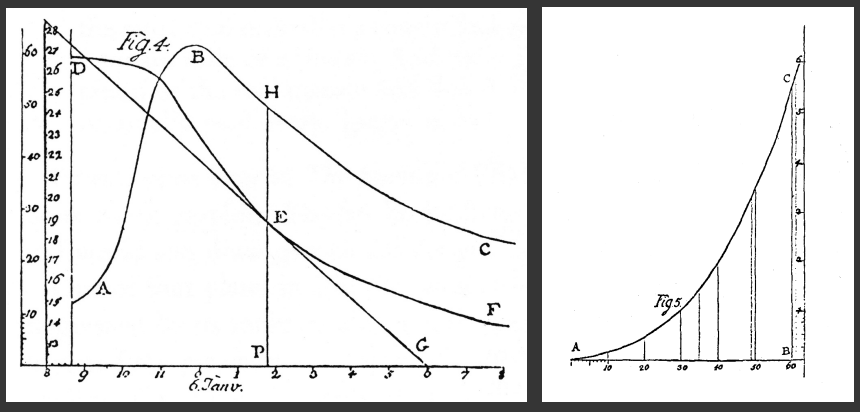
Model 1474-66 Electrotechnica 18 Scales

Tehnolemn Timisoara Slide Rule Archive
http://pubpages.unh.edu/~jwc/tehnolemn/

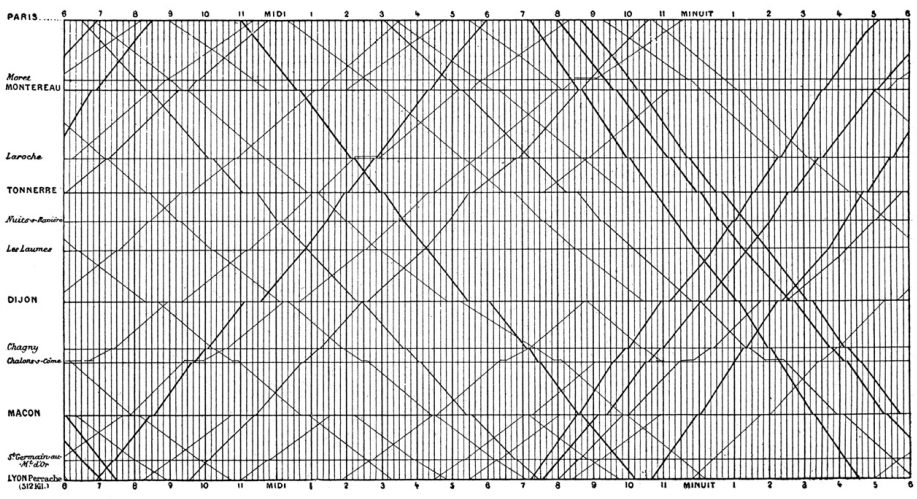




Lambert's graphical construction

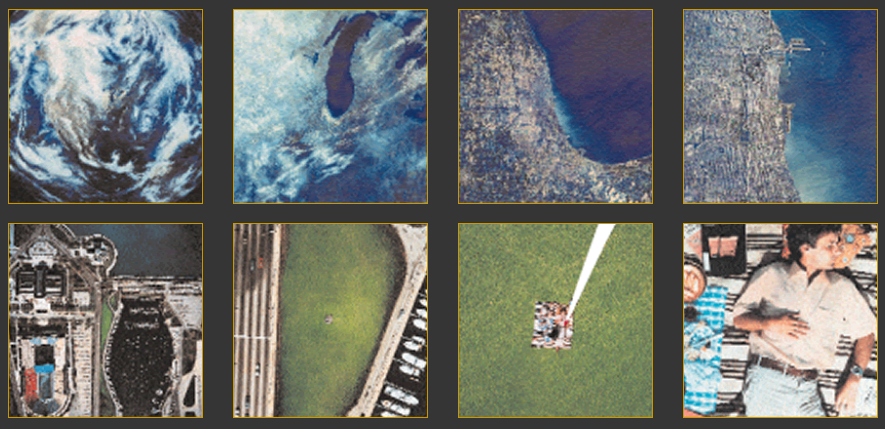


Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]



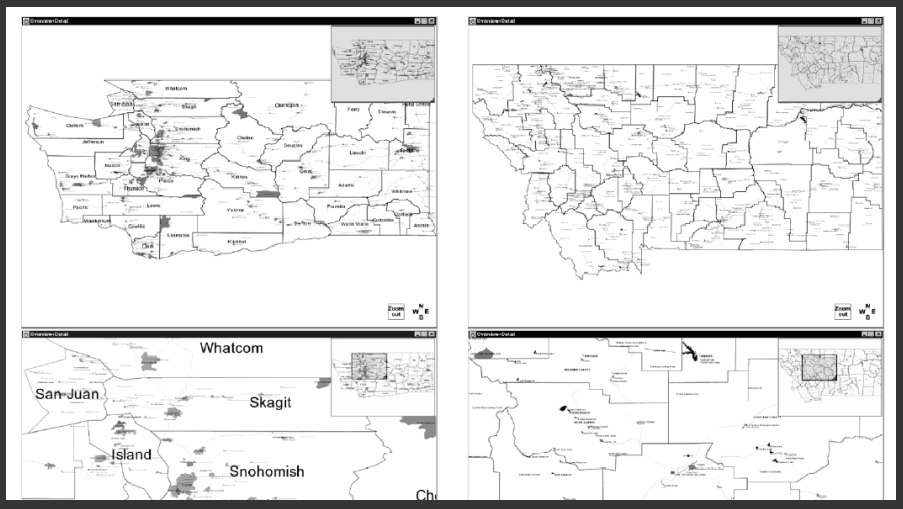
E. J. Marey, *La Méthode Graphique* (Paris, 1885), p. 20. The method is attributed to the French engineer, Ibry.

Zooming



Eames' Powers of Ten [http://www.powersof10.com/film]

Overview + details



[Hornbaek et al. 2002]