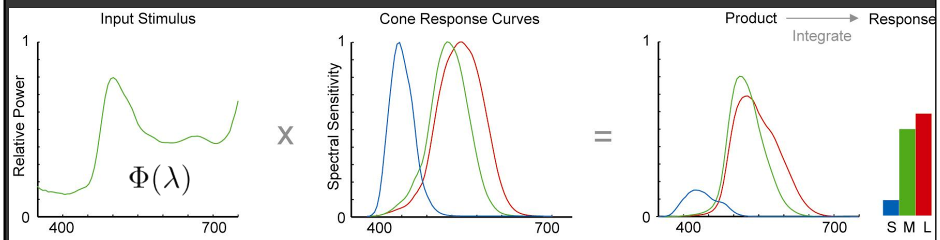
Using Space Effectively: 2D

Maneesh Agrawala

CS 448B: Visualization
Spring 2016

Last Time: Color

Computing Cone Response



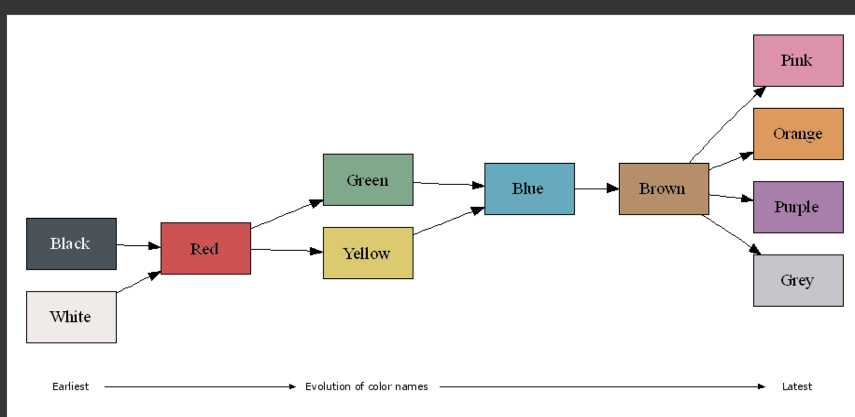
$$L = \int \Phi(\lambda) L(\lambda) \mathrm{d}\lambda$$

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

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

Evolution of Basic Color Terms

Proposed universal evolution across languages



Palette Design + Color Names

Minimize overlap and ambiguity of color names

Color Name Distance

										Salience	Name
0.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	0.20	.47	**blue** 62.9%
1.00	0.00	1.00	0.97	1.00	1.00	1.00	1.00	0.96	1.00	.90	**orange** 93.9%
1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00	0.90	0.99	.67	**green** 79.8%
1.00	0.97	1.00	0.00	1.00	0.95	0.99	1.00	1.00	1.00	.66	**red** 80.4%
0.98	1.00	1.00	1.00	0.00	0.96	0.91	0.97	1.00	0.99	.47	**purple** 51.4%
1.00	1.00	1.00	0.95	0.96	0.00	0.97	0.93	0.98	1.00	.37	**brown** 54.0%
1.00	1.00	1.00	0.99	0.91	0.97	0.00	1.00	1.00	1.00	.58	**pink** 71.7%
1.00	1.00	1.00	1.00	0.97	0.93	1.00	0.00	1.00	1.00	.67	**grey** 79.4%
1.00	0.96	0.90	1.00	1.00	0.98	1.00	1.00	0.00	1.00	.18	**yellow** 31.2%
0.20	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	0.00	.25	**blue** 25.4%
Tableau-10								*Average*	*0.97*	*.52*	

http://vis.stanford.edu/color-names

Palette Design + Color Names

Minimize overlap and ambiguity of color names

Color Name Distance

										Salience	Name
0.00	1.00	1.00	0.89	0.07	1.00	0.35	0.99	1.00	0.89	.30	**blue** 50.5%
1.00	0.00	0.99	1.00	1.00	0.92	1.00	0.84	0.98	0.99	.21	**red** 27.8%
1.00	0.99	0.00	1.00	0.98	1.00	1.00	1.00	0.17	1.00	.34	**green** 36.8%
0.89	1.00	1.00	0.00	0.98	1.00	0.71	0.93	1.00	0.32	.55	**purple** 67.3%
0.07	1.00	0.98	0.98	0.00	1.00	0.36	1.00	0.97	0.95	.20	**blue** 36.6%
1.00	0.92	1.00	1.00	1.00	0.00	1.00	0.97	0.99	1.00	.39	**orange** 51.9%
0.35	1.00	1.00	0.71	0.36	1.00	0.00	0.95	0.92	0.42	.13	**blue** 15.7%
0.99	0.84	1.00	0.93	1.00	0.97	0.95	0.00	0.98	0.85	.16	**pink** 29.4%
1.00	0.98	0.17	1.00	0.97	0.99	0.92	0.98	0.00	0.97	.12	**green** 21.7%
0.89	0.99	1.00	0.32	0.95	1.00	0.42	0.85	0.97	0.00	.30	**purple** 23.9%
Excel-10								*Average*	*0.87*	*.27*	

http://vis.stanford.edu/color-names

Hints for the colorist

Use only a few colors (~6 ideal)
Colors should be distinctive and named
Strive for color harmony (natural colors?)
Use cultural conventions; appreciate symbolism
Beware of bad interactions (red/blue etc.)
Get it right in black and white
Respect the color blind

Announcements

Assignment 3: Dynamic Queries

Create a **small** interactive dynamic query application similar to Homefinder, but for SF Crime Data.

1. 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 **May 4, 2016**

Using Space Effectively: 2D

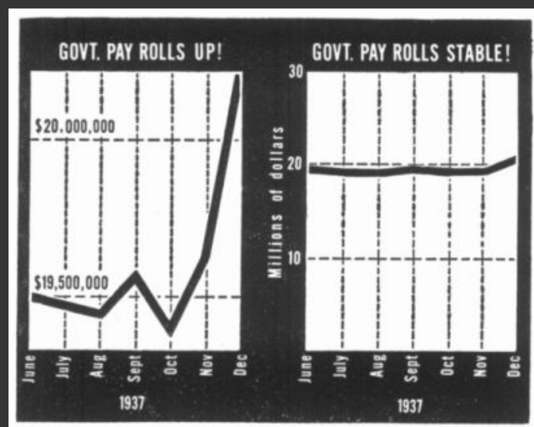
Topics

Displaying data in graphs
Selecting aspect ratio
Fitting data and depicting residuals
Graphical calculations
Zooming and Focus + Context
Cartographic distortion

Graphs and Lines

Effective use of space

Which graph is better?

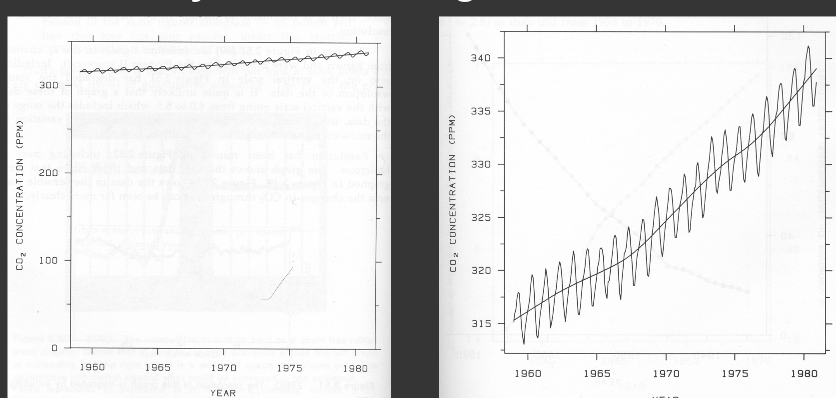


Government payrolls in 1937 [Huff 93]

Aspect ratio

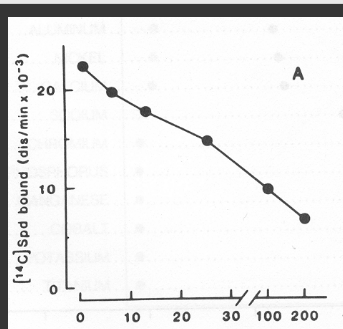
Fill space with data

Don't worry about showing zero

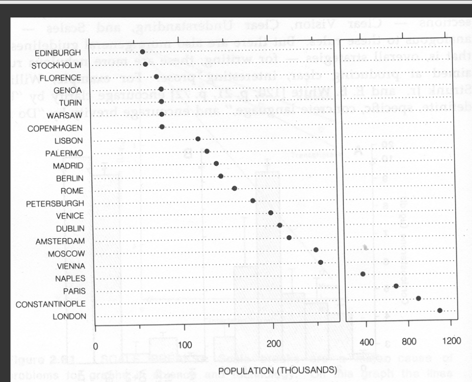


Yearly CO2 concentrations [Cleveland 85]

Clearly mark scale breaks

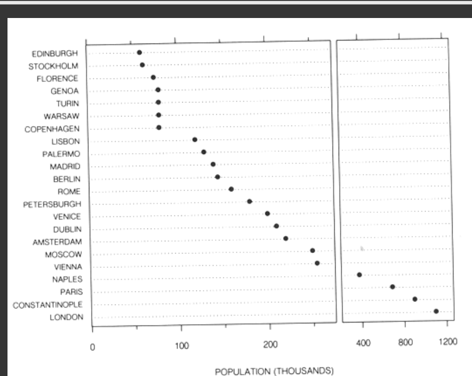


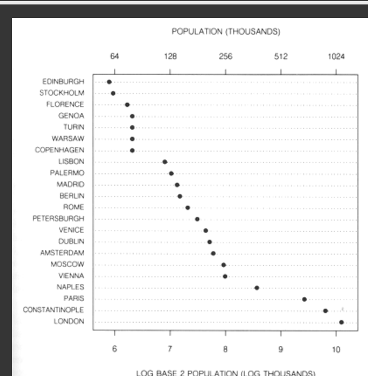
Poor scale break [Cleveland 85]



Well marked scale break [Cleveland 85]

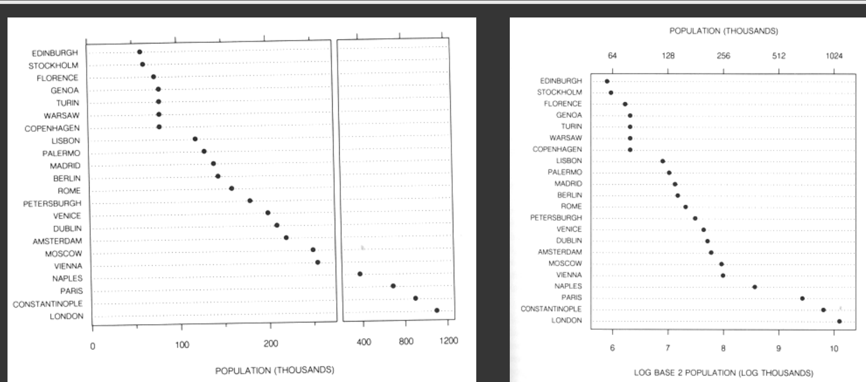
Scale break vs. Log scale





[Cleveland 85]

Scale break vs. Log scale

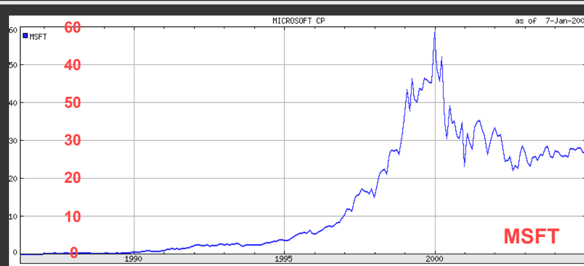


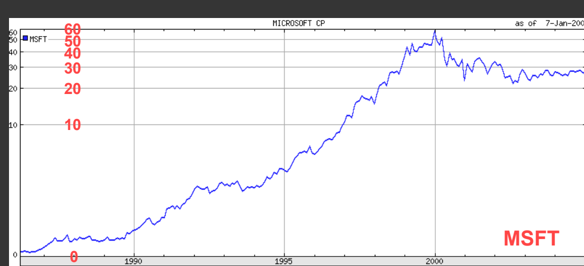
[Cleveland 85]

Both increase visual resolution

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

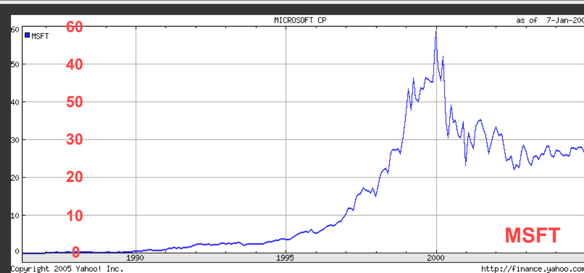
Linear scale vs. Log scale





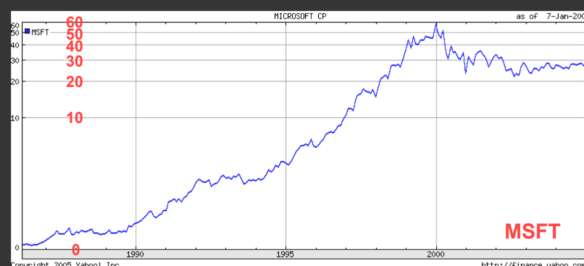
Linear scale vs. Log scale

Linear scale
- Absolute change



Log scale
- Small fluctuations
- Percent change

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



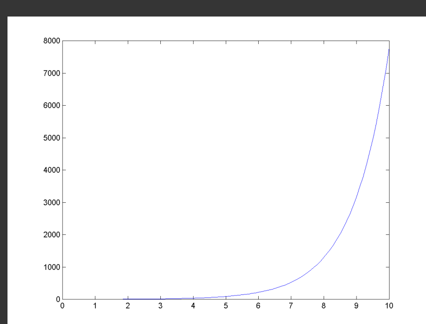
Semilog graph: Exponential growth

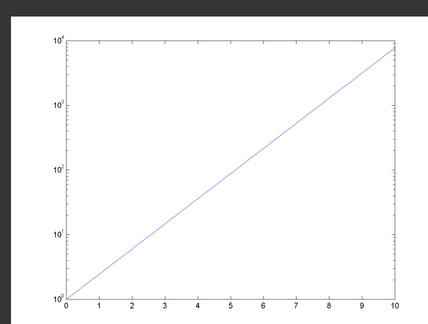
Exponential functions ($y = ka^{mx}$) transform into lines

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

Intercept: $\log(k)$

Slope: $\log(a)m$





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

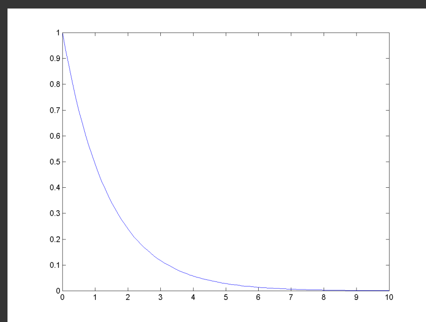
Semilog graph: Exponential decay

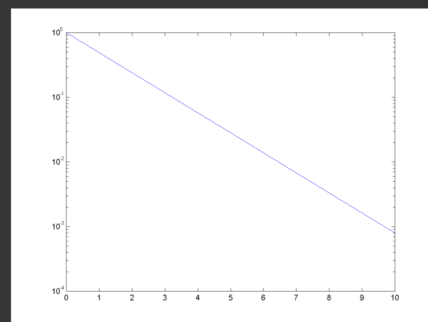
Exponential functions ($y = ka^{mx}$) transform into lines

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

Intercept: $\log(k)$

Slope: $\log(a)m$





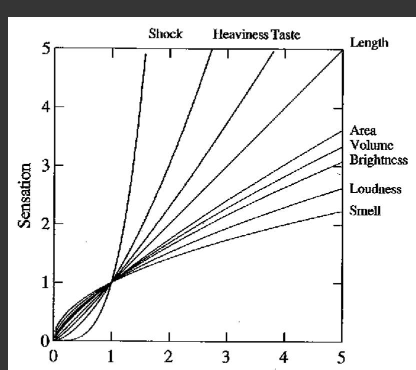
$y = 0.5^{2x}$, slope in semilog space: $\log(0.5)*2 = -0.602$

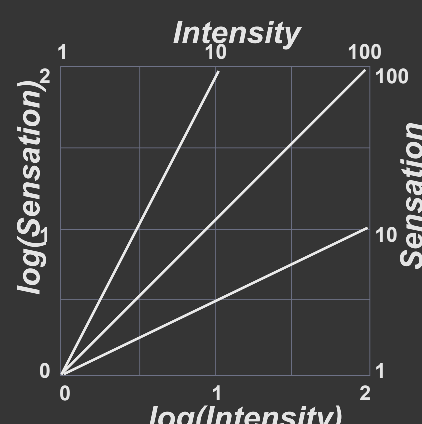
Log-Log graph

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

Example - Steven's power laws:

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



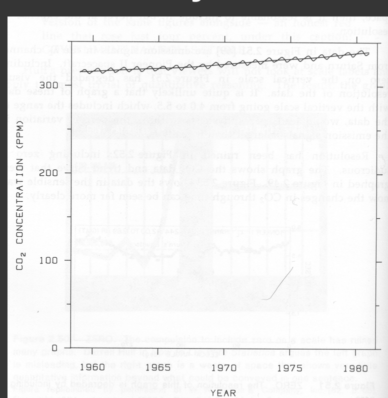


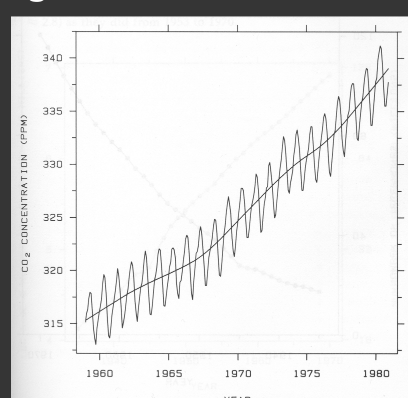
Selecting Aspect Ratio

Aspect ratio

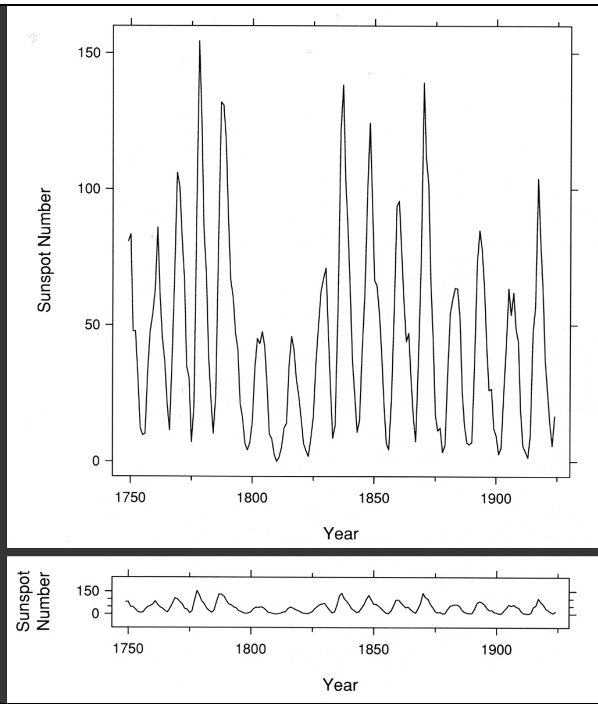
Fill space with data

Don't worry about showing zero





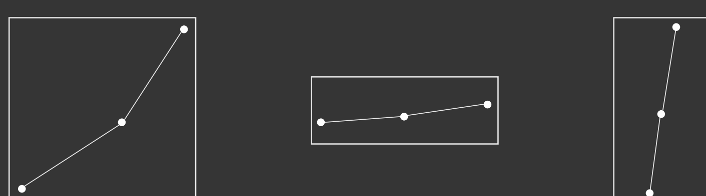
Yearly CO2 concentrations [Cleveland 85]



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 avg. absolute angle between them 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

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

Weighted

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

Max-Orientation-Resolution

Global (over all i, j s.t. i≠j)

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

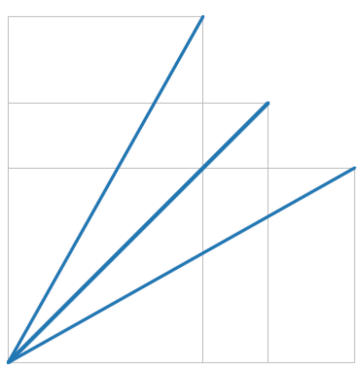
Local (over adjacent segments)

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

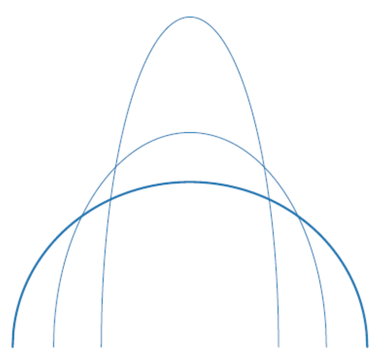
Requires Iterative Optimization

An alternate approach:

Minimize arc length (hold area constant)



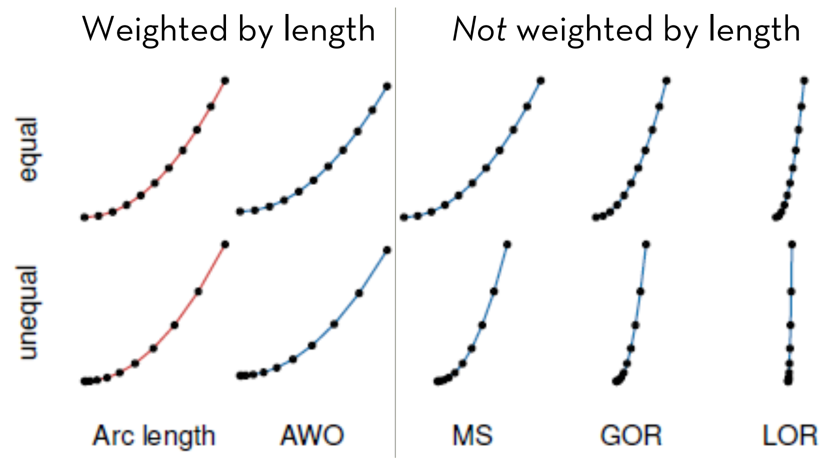
Straight line -> 45 deg



Ellipse -> Circle

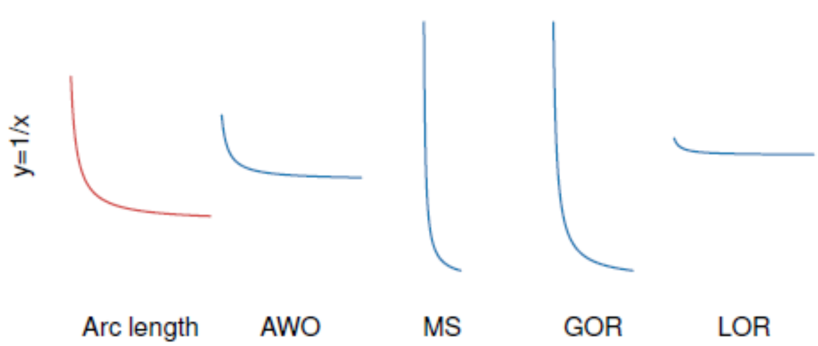
[Talbot et al, 2011]

Parameterization invariance



[Talbot et al, 2011]

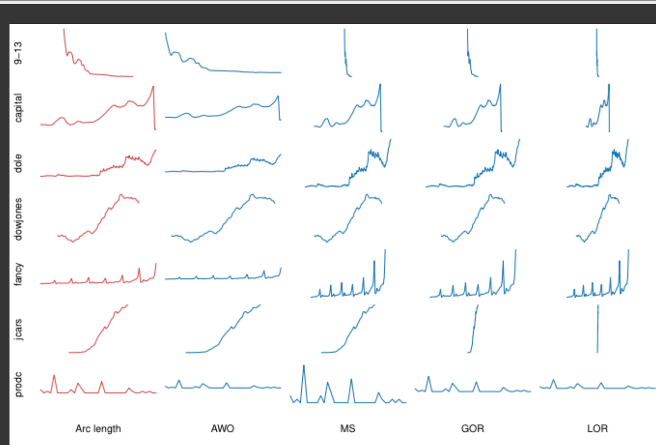
Robustness: banking y = 1 / x



[Talbot et al, 2011]

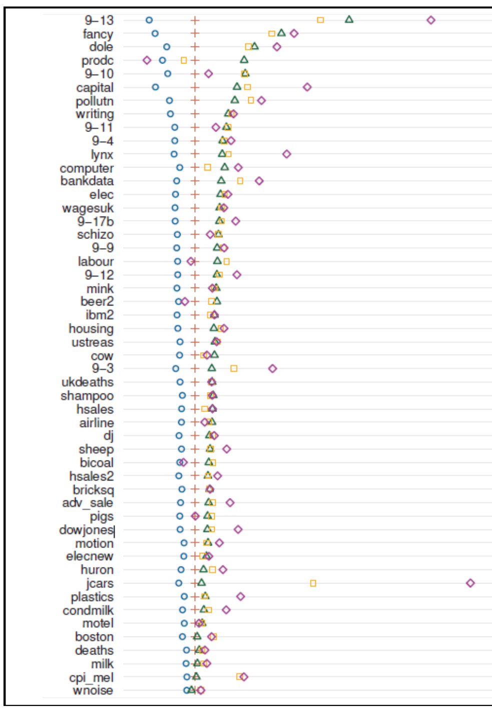
Arc-length based aspect ratio

$$\min_{a\in(0,\infty)} \sum_{i=1}^{N} ||\frac{\Delta x_i}{\sqrt{a}}, \sqrt{a}\Delta y_i||$$



Arc-length parameterization independent of spacing of points
Result are better than GOR and LOR and close to Cleveland's AWO

[Talbot 11]

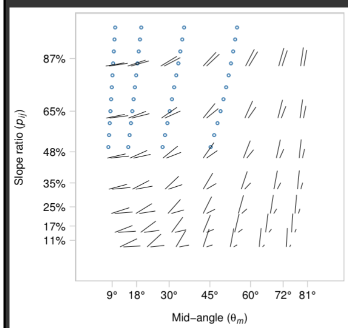


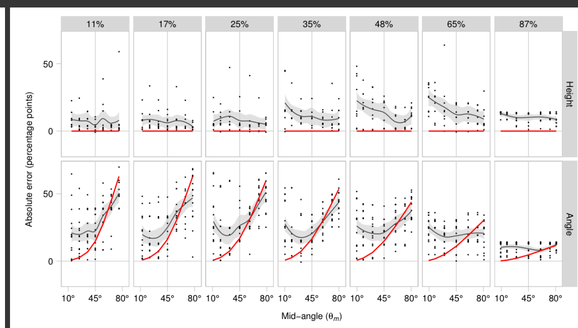
Compromise

Arc-length banking produces aspect ratios in-between those produced by other methods.

[Talbot et al, 2011]

Perceptual model based aspect ratio





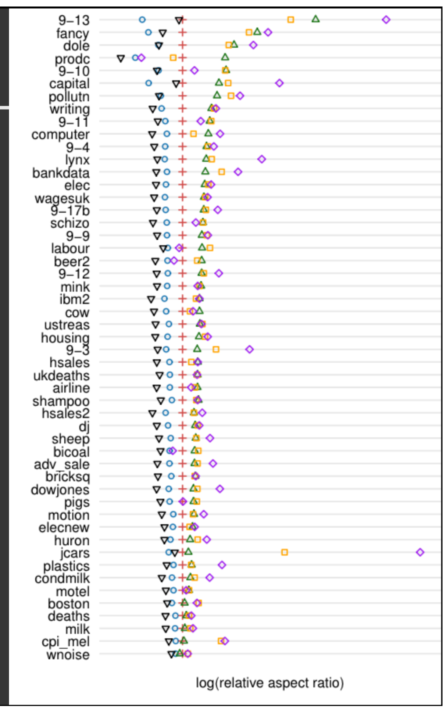
Ask people to estimate slope ratios for different conditions
Use data to fit a model derived from perceptual theory

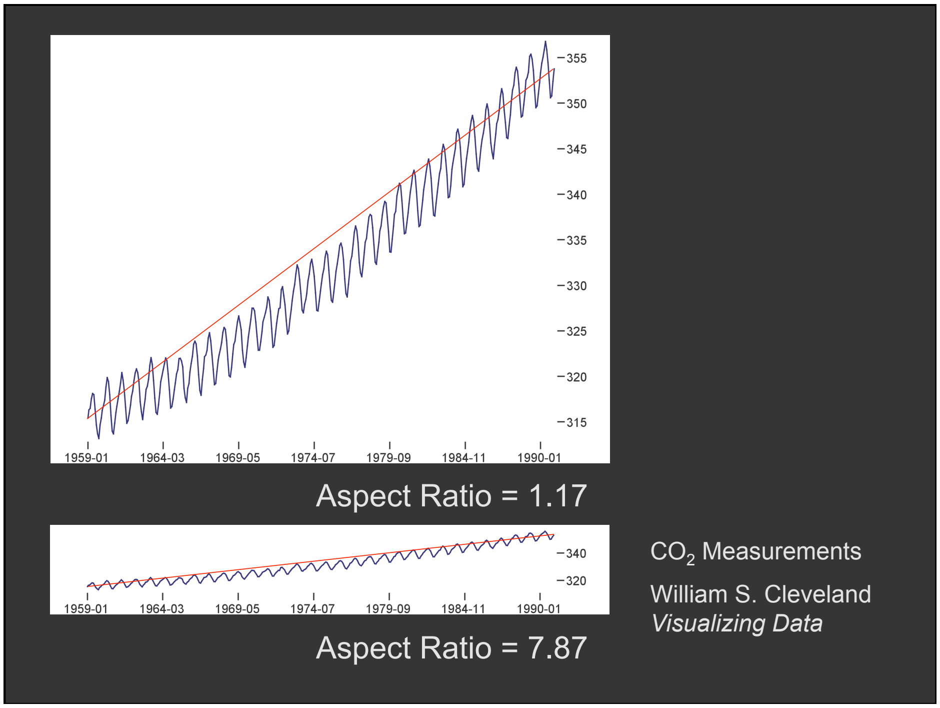
$$\hat{p}_{ij} = \begin{cases} \frac{\sin(\theta_i)\, l_i}{\sin(\theta_j)\, l_j} \times 100 & +\gamma & +\varepsilon^h_{ij} & \text{if HEIGHT} \\ \frac{\theta_i}{\theta_j} \times 100 & +(\mu + \beta\theta_m) & +\varepsilon^a_{ij} & \text{if ANGLE} \end{cases}$$

[Talbot 12]

Perceptual model

Perceptual model (black diamond) produces flatter aspect ratios than other techniques





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

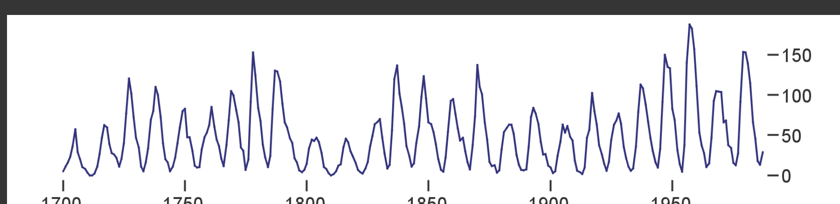
Multi-Scale Banking to 45°

Idea: Use Spectral Analysis to identify trends

Find strong frequency components

Lowpass filter to create trend lines

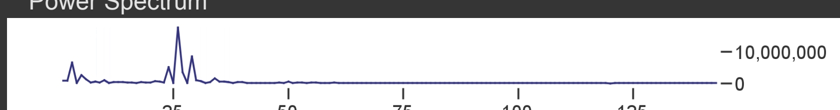
Compute Power Spectrum



Take Discrete Fourier Transform

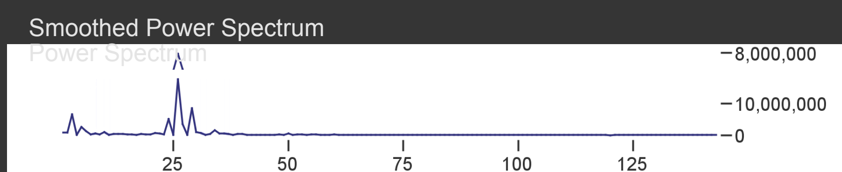
Compute squared magnitudes

Power Spectrum



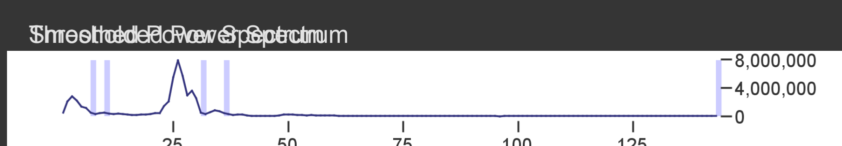
Smooth the Spectrum

Convolve with Gaussian filter
window size = 3, σ = 1

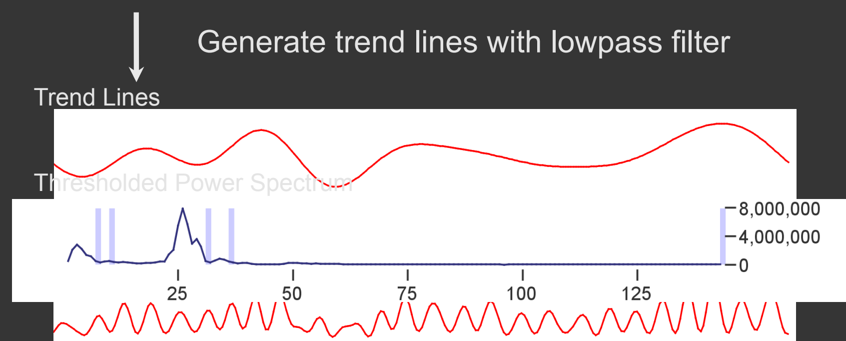


Threshold the Spectrum

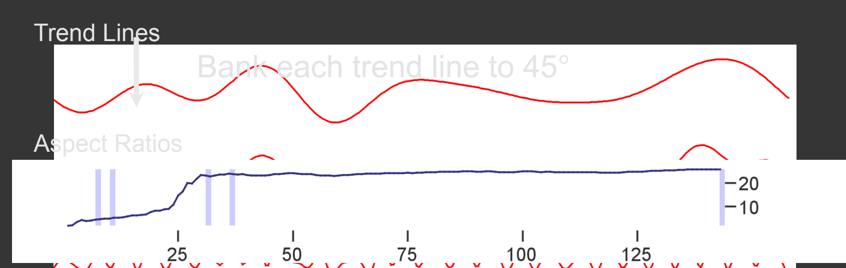
Threshold at mean of power spectrum
Retain last values of contiguous runs



Generate Trend Lines



Bank Trend Lines to 45°



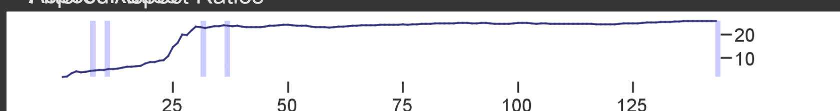
Filter Aspect Ratios

Filter similar aspect ratios

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

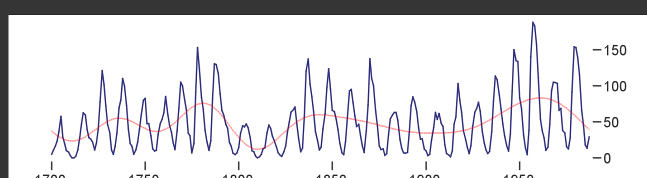
Aspect Ratios

Filtered Aspect Ratios

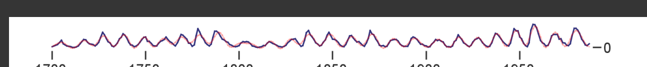


Sunspot Cycles

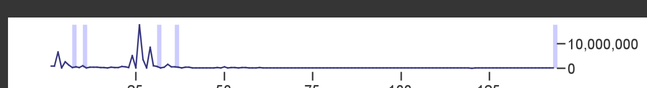
Yearly values 1700-1987



Aspect Ratio = 4.23



Aspect Ratio = 14.55



Power Spectrum

Aspect Ratios

CO_2

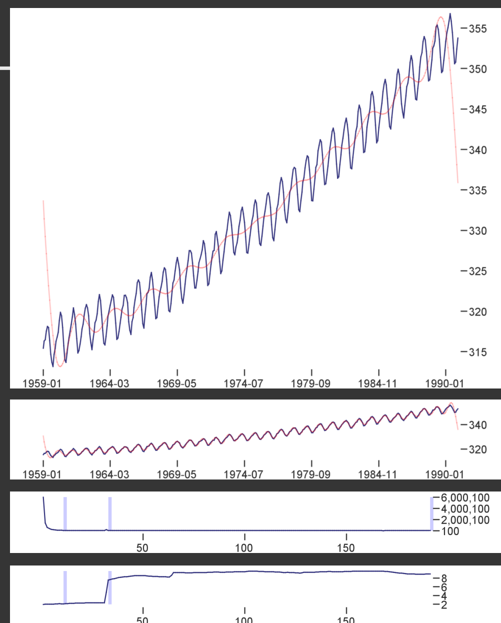
Monthly concentrations from the Mauna Loa Observatory, 1950-1990

Aspect Ratio = 1.17

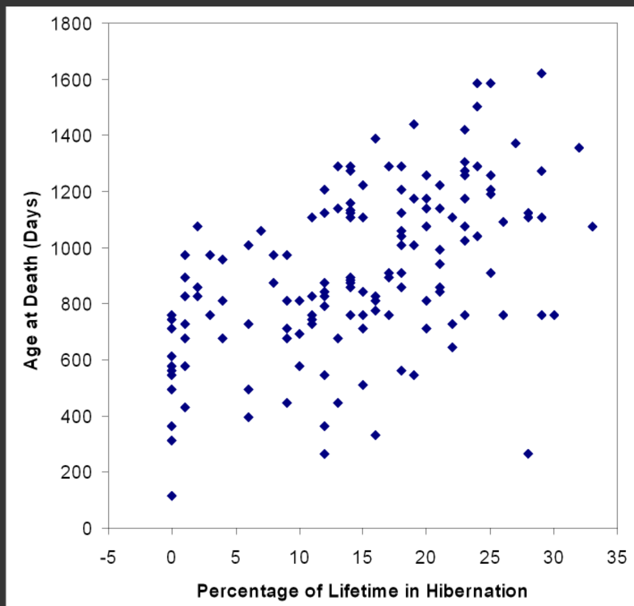
Aspect Ratio = 7.87

Power Spectrum

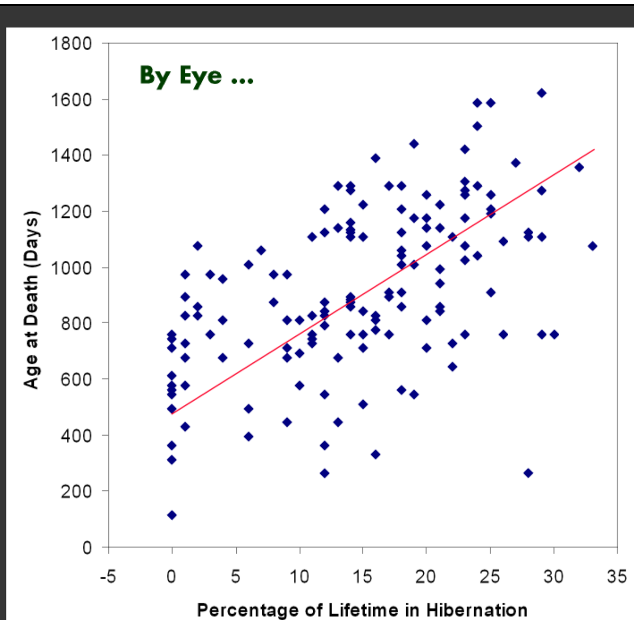
Aspect Ratios



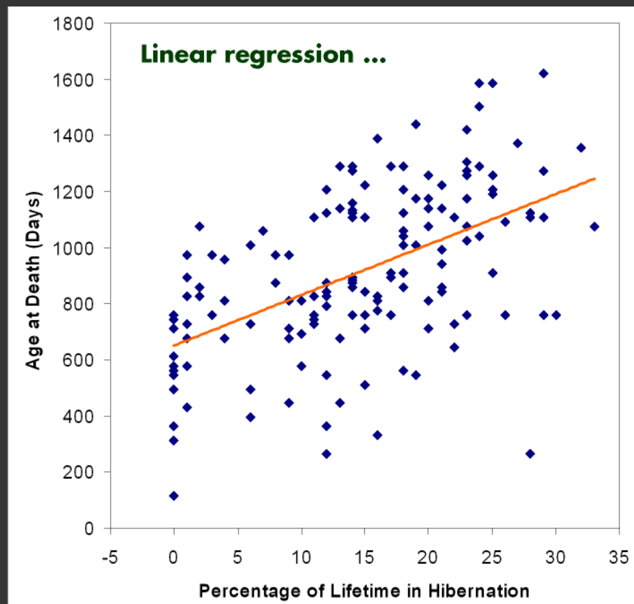
Fitting the Data



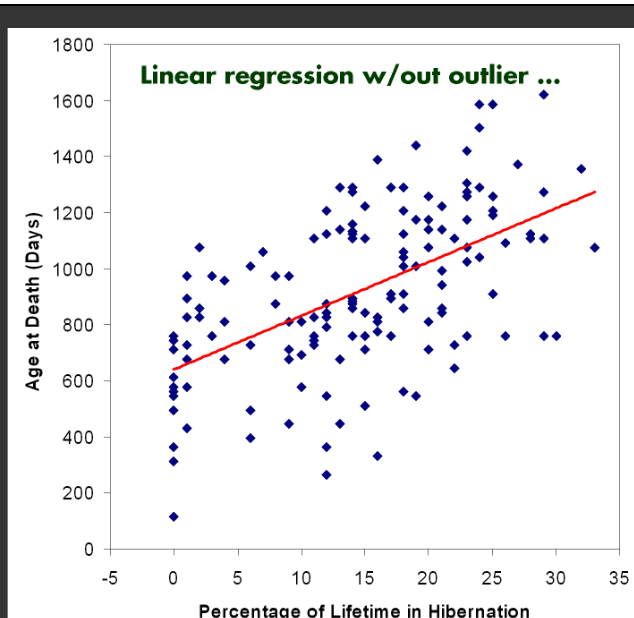
[The Elements of Graphing Data. Cleveland 94]



[The Elements of Graphing Data. Cleveland 94]



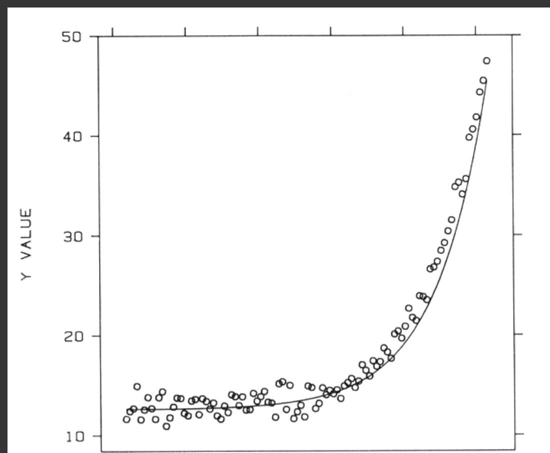
[The Elements of Graphing Data. Cleveland 94]



[The Elements of Graphing Data. Cleveland 94]

Transforming data

How well does curve fit data?

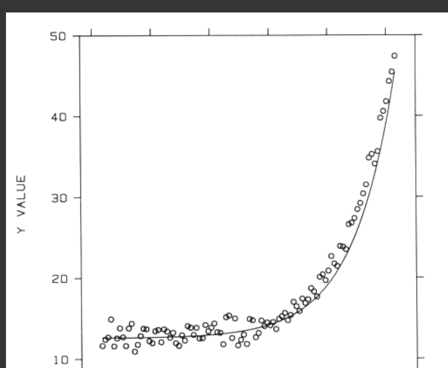


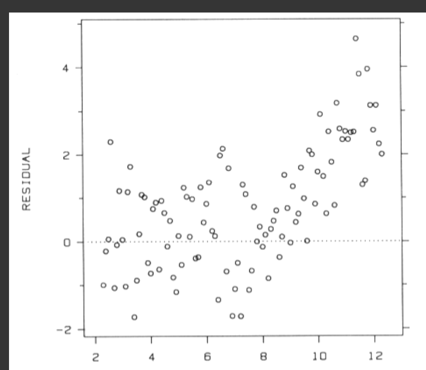
[Cleveland 85]

Transforming data

Residual graph

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

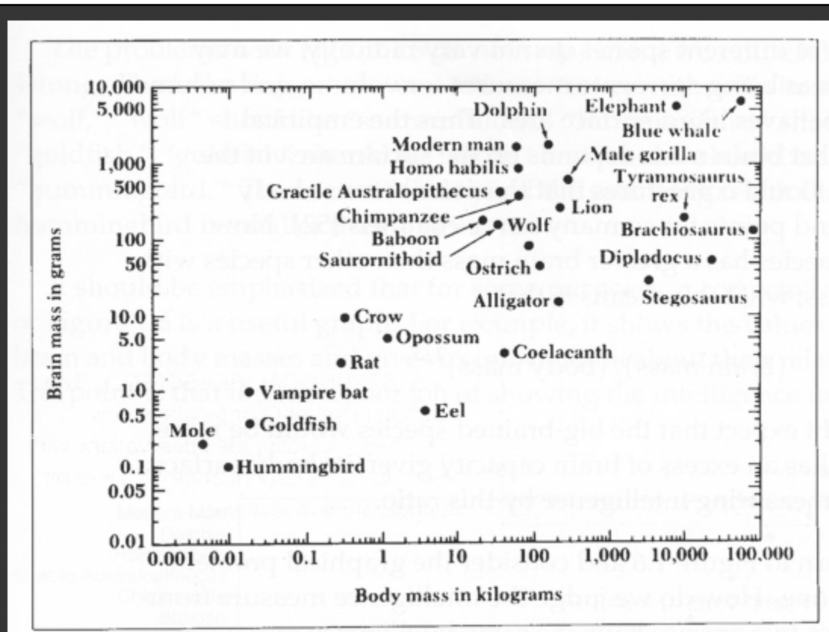




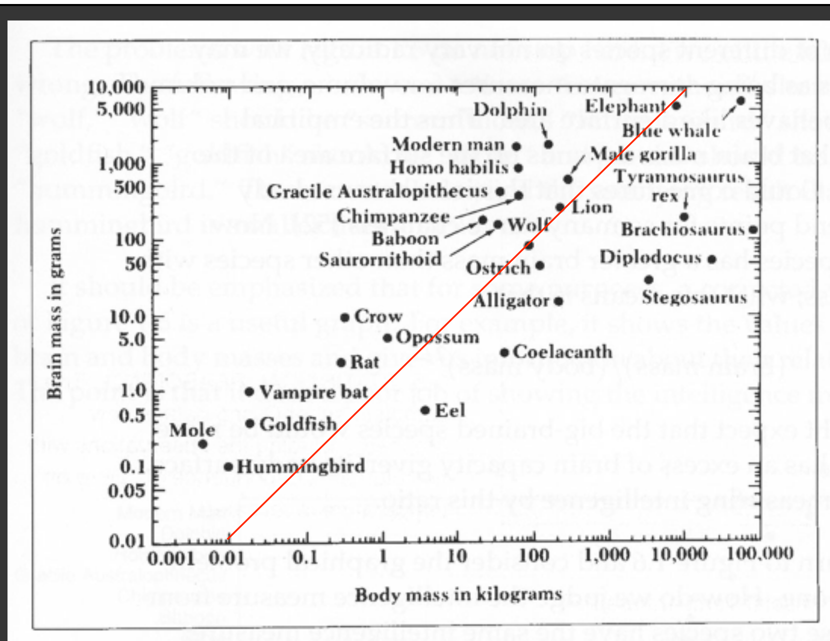
[Cleveland 85]

Most powerful brain?

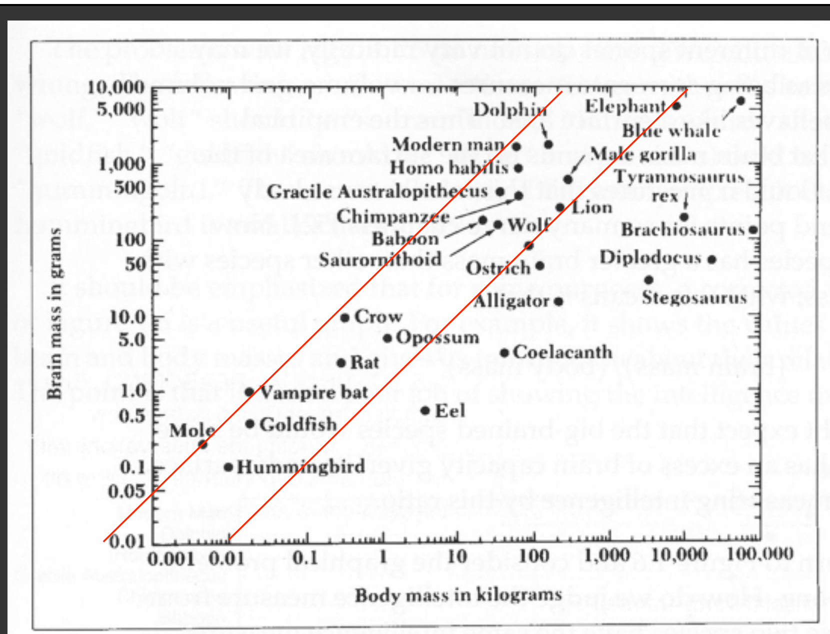
ID	Name	Body Weight	Brain Weight
1	Lesser Short-tailed Shrew	5	0.14
2	Little Brown Bat	10	0.25
3	Mouse	23	0.3
4	Big Brown Bat	23	0.4
5	Musk Shrew	48	0.33
6	Star Nosed Mole	60	1
7	Eastern American Mole	75	1.2
8	Ground Squirrel	101	4
9	Tree Shrew	104	2.5
10	Golden Hamster	120	1
11	Mole Rate	122	3
12	Galago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



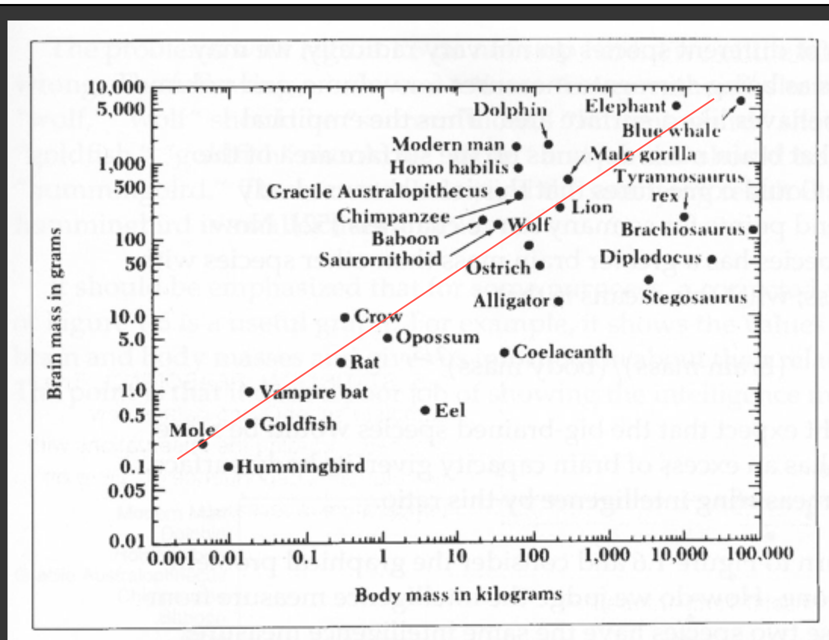
The Dragons of Eden [Carl Sagan]



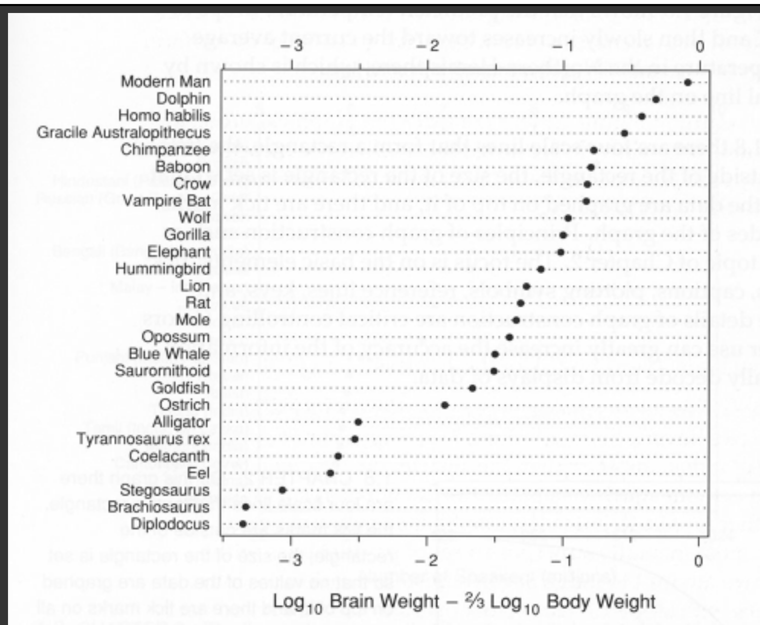
The Dragons of Eden [Carl Sagan]



The Dragons of Eden [Carl Sagan]

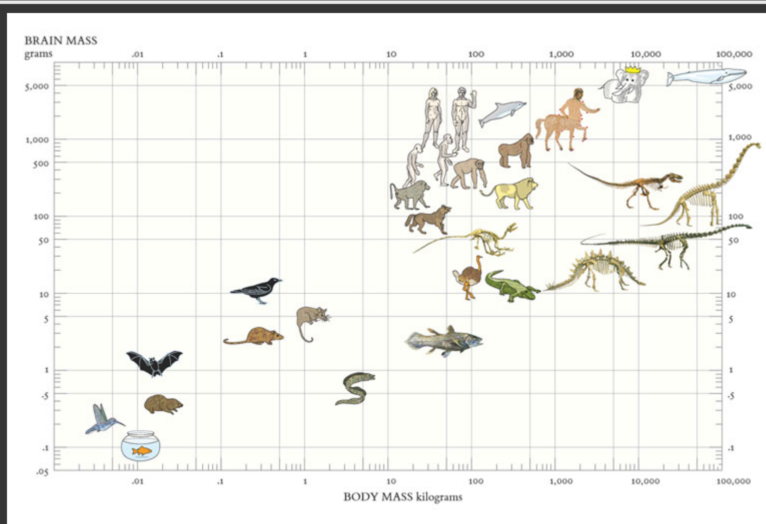


The Dragons of Eden [Carl Sagan]



The Elements of Graphing Data [Cleveland]

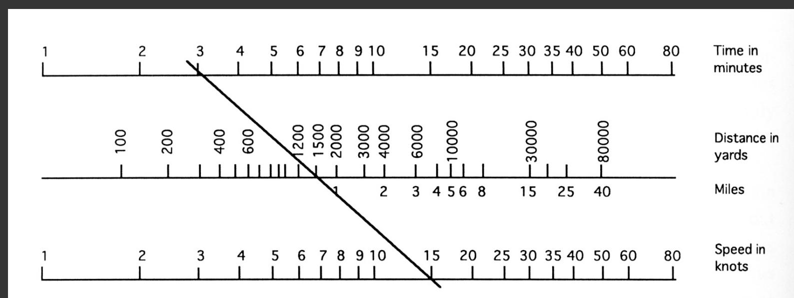
Most powerful brain



Beautiful Evidence [Tufte]

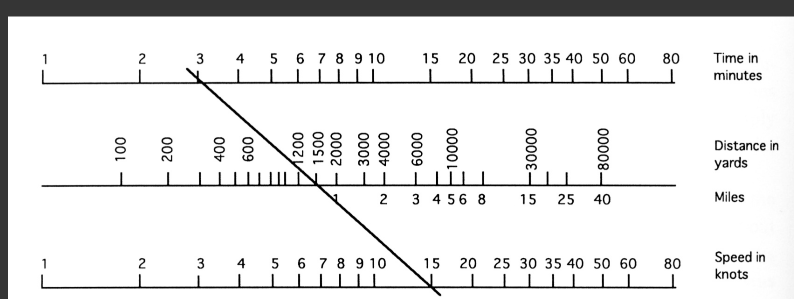
Graphical Calculations

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

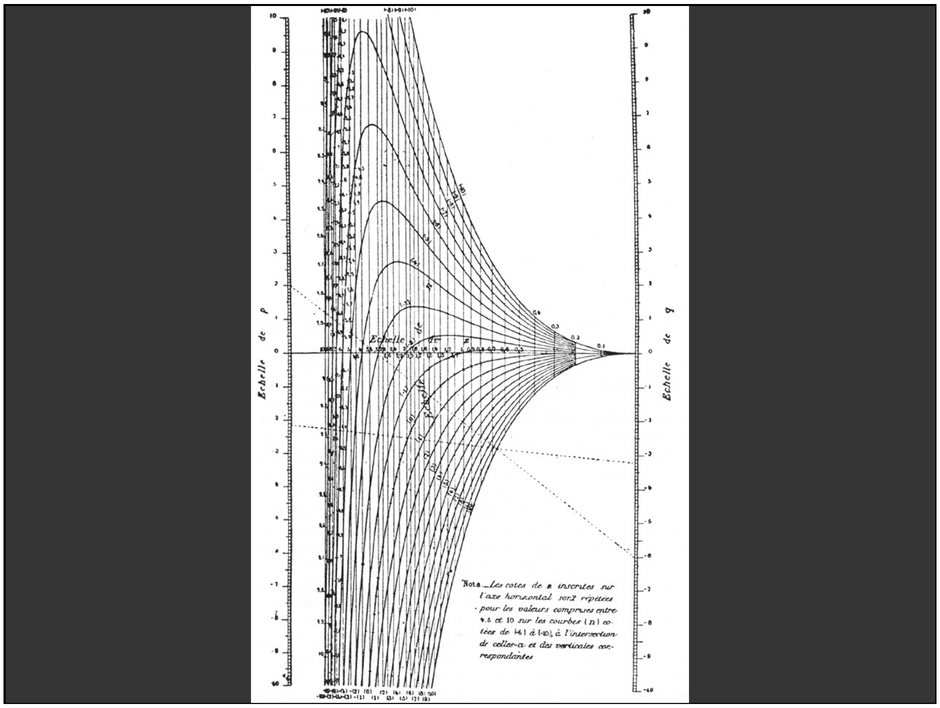
$$\begin{vmatrix} x_1(u) & y_1(u) & w_1(u) \\ x_2(v) & y_2(v) & w_2(v) \\ x_3(s,t) & y_3(s,t) & w_3(s,t) \end{vmatrix} = 0$$

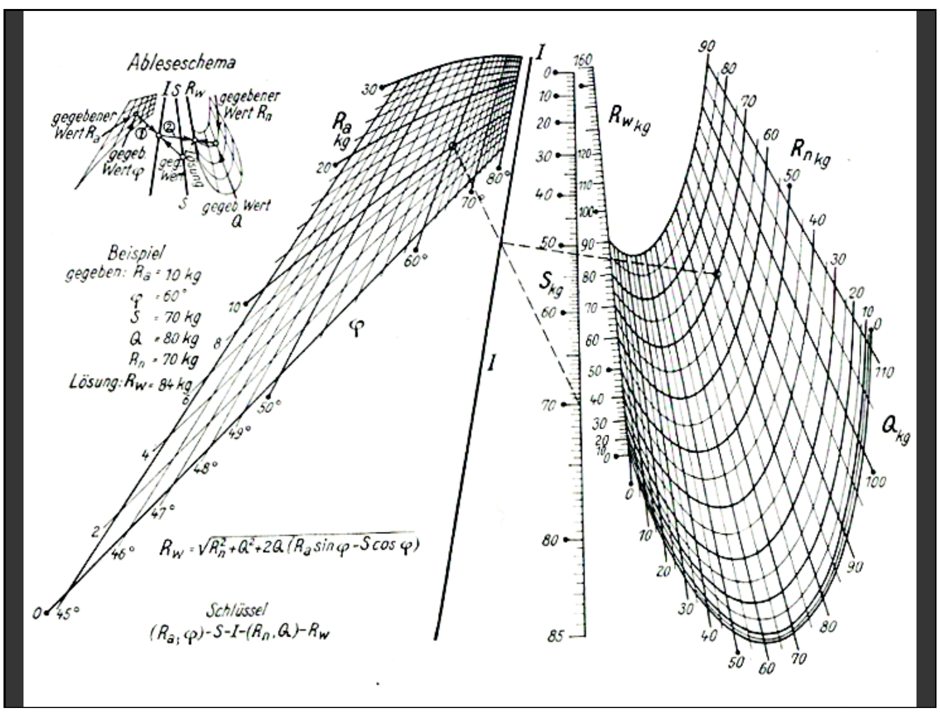
Slide rule

Model 1474-66 Electrotechnica 18 Scales

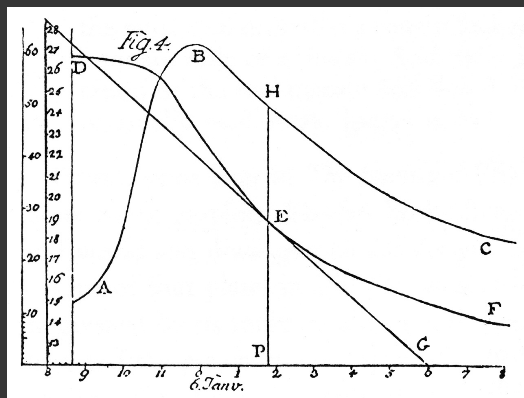
Tehnolemn Timisoara Slide Rule Archive

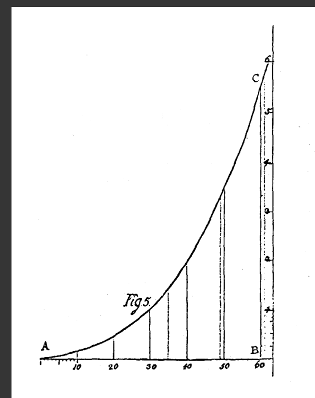
http://pubpages.unh.edu/~jwc/tehnolemn/


Echelle de p
Echelle de q
Echelle de n
Nota _ Les cotes de n inscrites sur l'axe horizontal sont répétées pour les valeurs comprises entre 4.5 et 10 sur les courbes (n) cotées de (4) à (10), à l'intersection de celles-ci et des verticales correspondantes

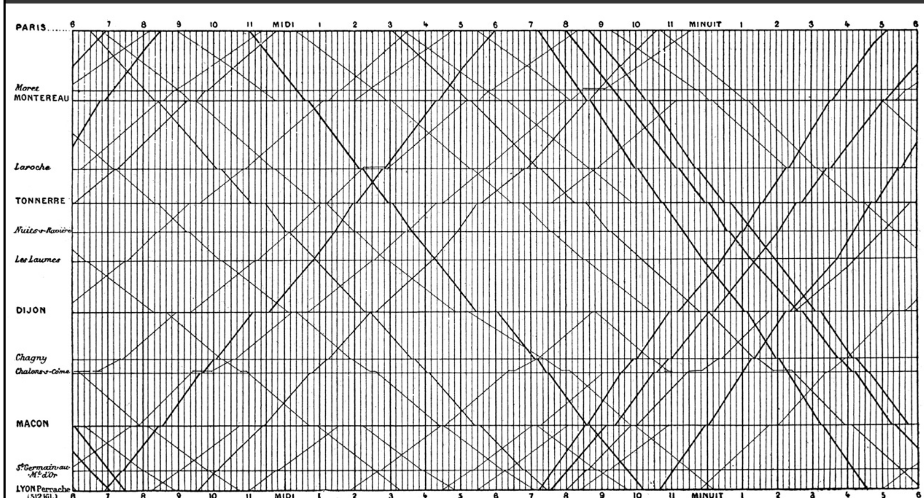

Ableseschema
gegebener Wert R_a
gegeb Wert φ
gegebener Wert R_n
geg. Wert
gegeb Wert Q
Lösung
Beispiel
gegeben: R_a = 10 kg
φ = 60°
S = 70 kg
Q = 80 kg
R_n = 70 kg
Lösung: R_w = 84 kg
$R_w = \sqrt{R_n^2 + Q^2 + 2Q(R_a \sin\varphi - S\cos\varphi)}$
Schlüssel
$(R_a; \varphi)$-S-I-(R_n, Q)-R_w
R_a kg
φ
I
R_w kg
S_{kg}
R_n kg
Q_{kg}

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.eamesoffice.com/the-work/powers-of-ten/]