# Scaling—a Plenitude of Power Laws

Last updated: 2024/09/10, 07:30:42 EDT

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number, 2024–2025

Prof. Peter Sheridan Dodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

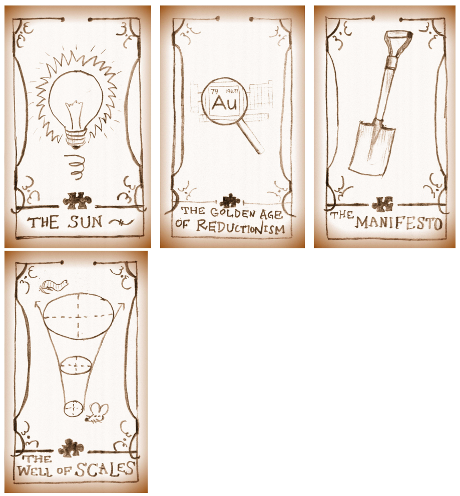




## Outline

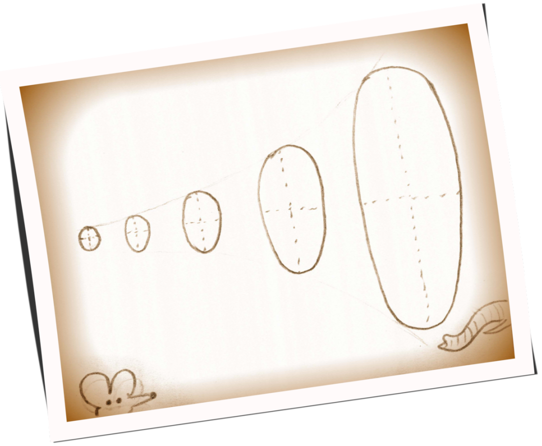









## Archival object:





## Scalingarama

### General observation:

Systems (complex or not) that cross many spatial and temporal scales often exhibit some form of scaling.

### Outline—All about scaling:

- Basic definitions.
- Examples.

### Possibly later:

- Advances in measuring your power-law relationships.
- Scaling in blood and river networks.
- The Unsolved Allometry Theoricides.



## Definitions

A power law relates two variables $x$ and $y$ as follows:

$$y = cx^{\alpha}$$

- $\alpha$ is the scaling exponent (or just exponent)
- $\alpha$ can be any number in principle but we will find various restrictions.
- $c$ is the prefactor (which can be important!)



## Definitions

- The prefactor $c$ must balance dimensions.
- Imagine the height $\ell$ and volume $v$ of a family of shapes are related as:
$$\ell = cv^{1/4}$$
- Using $[\cdot]$ to indicate dimension, then
$$[c] = [\ell]/[v^{1/4}] = L/L^{3/4} = L^{1/4}.$$
- More on this later with the Buckingham $\pi$ theorem.



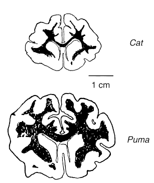
## Looking at data

- Power-law relationships are linear in log-log space:
$$y = cx^{\alpha}$$
$$\Rightarrow \log_b y = \alpha \log_b x + \log_b c$$
with slope equal to $\alpha$, the scaling exponent.
- Much searching for straight lines on log-log or double-logarithmic plots.
- Good practice: Always, always, always use base 10.
- Yes, the Dozenalists are right, 12 would be better.
- But: hands.[1] And social pressure.
- Talk only about orders of magnitude (powers of 10).

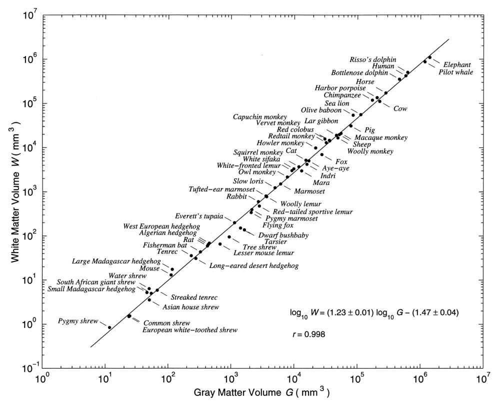
[1] Probably an accident of evolution—debated.



## A beautiful, heart-warming example:




- $G$ = volume of gray matter: 'computing elements'
- $W$ = volume of white matter: 'wiring'
- $W \sim cG^{1.23}$




- from Zhang & Sejnowski, PNAS (2000) [40]

# Why is $\alpha \simeq 1.23$?

## Quantities (following Zhang and Sejnowski):

- $G$ = Volume of gray matter (cortex/processors)
- $W$ = Volume of white matter (wiring)
- $T$ = Cortical thickness (wiring)
- $S$ = Cortical surface area
- $L$ = Average length of white matter fibers
- $p$ = density of axons on white matter/cortex interface

## A rough understanding:

- $G \sim ST$ (convolutions are okay)
- $W \sim \frac{1}{2}pSL$
- $G \sim L^3$
- Eliminate $S$ and $L$ to find $W \propto G^{4/3}/T$



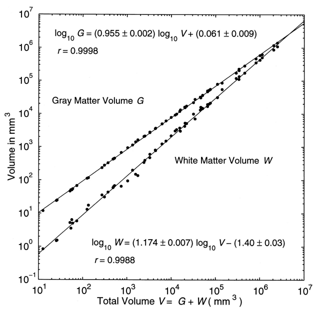
# Why is $\alpha \simeq 1.23$?

## A rough understanding:

- We are here: $W \propto G^{4/3}/T$
- Observe weak scaling $T \propto G^{0.10 \pm 0.02}$.
- Implies $S \propto G^{0.9} \rightarrow$ convolutions fill space.
- $\Rightarrow W \propto G^{4/3}/T \propto G^{1.23 \pm 0.02}$



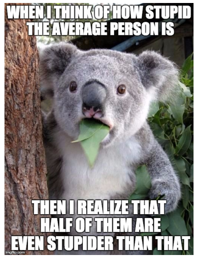
# Tricksiness:




- With $V = G + W$, some power laws must be approximations.
- Measuring exponents is a hairy business...



# Disappointing deviations from scaling:




- Per George Carlin
- Yes, should be the median. #painful

Image from here

The koala, a few roos short in the top paddock:

- Very small brains relative to body size.
- Wrinkle-free, smooth.
- Not many algorithms needed:
  - Only eat eucalyptus leaves (no water) (Will not eat leaves picked and presented to them)
  - Move to the next tree.
  - Sleep.
  - Defend themselves if needed (tree-climbing crocodiles, humans).
  - Occasionally make more koalas.



# Good scaling:

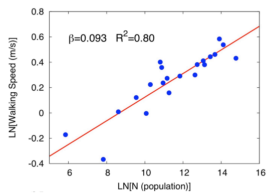
## General rules of thumb:

- *High quality:* scaling persists over three or more orders of magnitude for each variable.
- *Medium quality:* scaling persists over three or more orders of magnitude for only one variable and at least one for the other.
- *Very dubious:* scaling 'persists' over less than an order of magnitude for both variables.



# Unconvincing scaling:

## Average walking speed as a function of city population:




Two problems:

1. use of natural log, and
2. minute varation in dependent variable.

- from Bettencourt et al. (2007) [4]; otherwise totally great—more later.



# Definitions

Power laws are the signature of scale invariance:

Scale invariant 'objects'
look the 'same'
when they are appropriately rescaled.

- Objects = geometric shapes, time series, functions, relationships, distributions,...
- 'Same' might be 'statistically the same'
- To rescale means to change the units of measurement for the relevant variables



# Scale invariance

## Our friend $y = cx^\alpha$:

- If we rescale $x$ as $x = rx'$ and $y$ as $y = r^\alpha y'$,
- then
$$r^\alpha y' = c(rx')^\alpha$$
-
$$\Rightarrow y' = cr^\alpha x'^\alpha r^{-\alpha}$$
-
$$\Rightarrow y' = cx'^\alpha$$



# Scale invariance

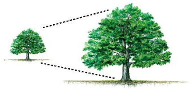
## Compare with $y = ce^{-\lambda x}$:

- If we rescale $x$ as $x = rx'$, then
$$y = ce^{-\lambda r x'}$$
- Original form cannot be recovered.
- Scale matters for the exponential.

## More on $y = ce^{-\lambda x}$:

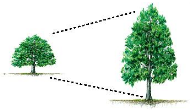
- Say $x_0 = 1/\lambda$ is the characteristic scale.
- For $x \gg x_0$, $y$ is small, while for $x \ll x_0$, $y$ is large.

## Isometry:



- Dimensions scale linearly with each other.

## Allometry:



- Dimensions scale nonlinearly.

## Allometry:

- Refers to differential growth rates of the parts of a living organism's body part or process.
- First proposed by Huxley and Teissier, Nature, 1936 "Terminology of relative growth" [15, 35]



# Definitions

## Isometry versus Allometry:

- Iso-metry = 'same measure'
- Allo-metry = 'other measure'

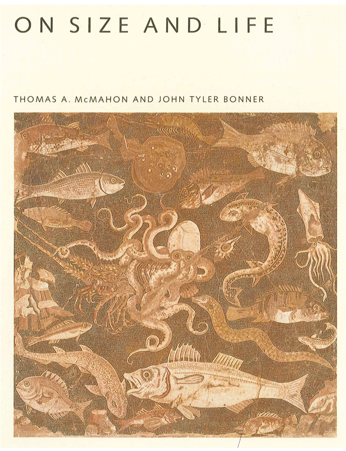
## We use allometric scaling to refer to both:

1. Nonlinear scaling of a dependent variable on an independent one (e.g., $y \propto x^{1/3}$)
2. The relative scaling of correlated measures (e.g., white and gray matter).



## An interesting, earlier treatise on scaling:

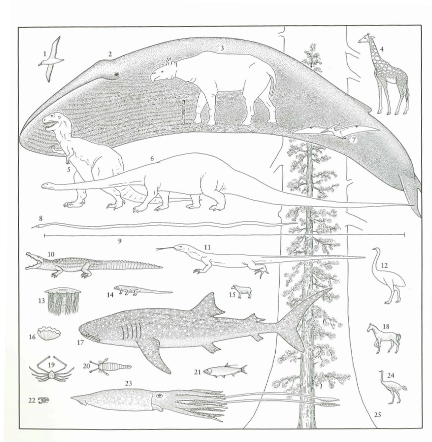
McMahon and Bonner, 1983 [26]




## The many scales of life:

The biggest living things (*left*). All the organisms are drawn to the same scale. *1*, The largest flying bird (albatross); *2*, the largest known animal (the blue whale), *3*, the largest extinct land mammal (*Baluchitherium*) with a human figure shown for scale; *4*, the tallest living land animal (giraffe); *5*, *Tyrannosaurus*; *6*, *Diplodocus*; *7*, one of the largest flying reptiles (*Pteranodon*); *8*, the largest extinct snake; *9*, the length of the largest tapeworm found in man; *10*, the largest living reptile (West African crocodile); *11*, the largest extinct lizard; *12*, the largest extinct bird (*Aepyornis*); *13*, the largest jellyfish (*Cyanea*); *14*, the largest living lizard (Komodo dragon); *15*, sheep; *16*, the largest bivalve mollusc (*Tridacna*); *17*; the largest fish (whale shark); *18*, horse; *19*, the largest crustacean (Japanese spider crab); *20*, the largest sea scorpion (Eurypterid); *21*, large tarpon; *22*, the largest lobster; *23*, the largest mollusc (deep-water squid, *Architeuthis*); *24*, ostrich; *25*, the lower 105 feet of the largest organism (giant sequoia), with a 100-foot larch superposed.

p. 2, McMahon and Bonner [26]

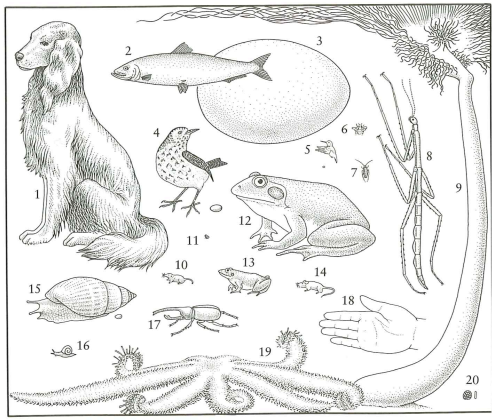


## The many scales of life:

Medium-sized creatures (*above*). *1*, Dog; *2*, common herring; *3*, the largest egg (*Aepyornis*); *4*, song thrush with egg; *5*, the smallest bird (hummingbird) with egg; *6*, queen bee; *7*, common cockroach; *8*, the largest stick insect; *9*, the largest polyp (*Branchiocerianthus*); *10*, the smallest mammal (flying shrew); *11*, the smallest vertebrate (a tropical frog); *12*, the largest frog (goliath frog); *13*, common grass frog; *14*, house mouse; *15*, the largest land snail (*Achatina*) with egg; *16*, common snail; *17*, the largest beetle (goliath beetle); *18*, human hand; *19*, the largest starfish (*Luidia*); *20*, the largest free-moving protozoan (an extinct nummulite).

p. 3, McMahon and Bonner [26]

More on the Elephant Bird here.

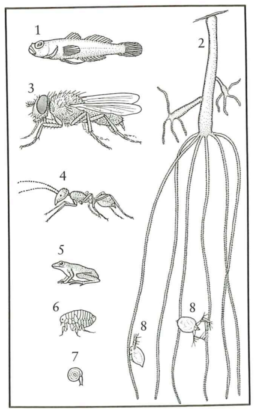


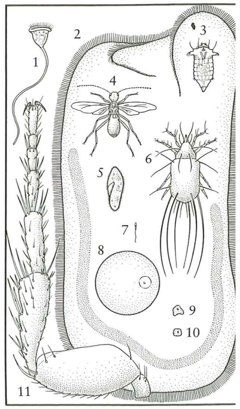
## The many scales of life:

Small, "naked-eye" creatures (*lower left*). *1*, One of the smallest fishes (*Trimmatom nanus*); *2*, common brown hydra, expanded; *3*, housefly; *4*, medium-sized ant; *5*, the smallest vertebrate (a tropical frog, the same as the one numbered *11* in the figure *above*); *6*, flea (*Xenopsylla cheopis*); *7*, the smallest land snail; *8*, common water flea (*Daphnia*).

The smallest "naked-eye" creatures and some large microscopic animals and cells (*below right*). *1*, *Vorticella*, a ciliate; *2*, the largest ciliate protozoan (*Bursaria*); *3*, the smallest many-celled animal (a rotifer); *4*, smallest flying insect (*Elaphis*); *5*, another ciliate (*Paramecium*); *6*, cheese mite; *7*, human sperm; *8*, human ovum; *9*, dysentery amoeba; *10*, human liver cell; *11*, the foreleg of the flea (numbered 6 in the figure to the *left*).

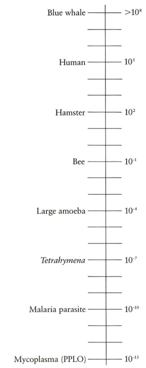
p. 3, McMahon and Bonner [26]



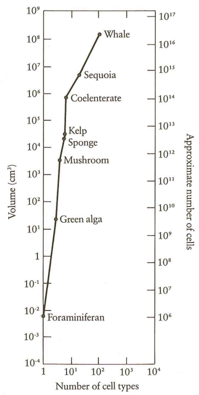




## Size range (in grams) and cell differentiation:

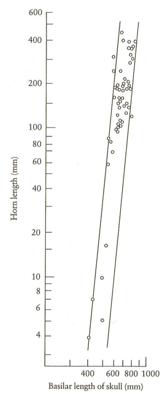


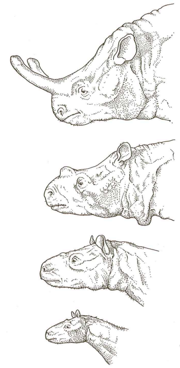

$10^{-13}$ to $10^{8}$ g, p. 3, McMahon and Bonner [26]






## Titanothere horns: $L_{\text{horn}} \sim L_{\text{skull}}{}^{4}$

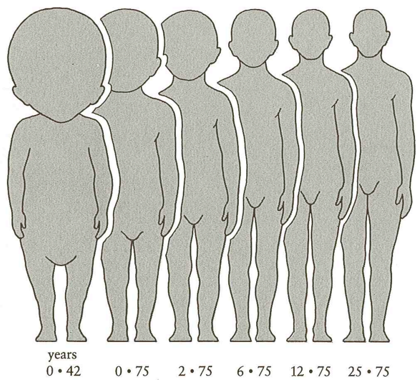


p. 36, McMahon and Bonner [26]; a bit dubious.



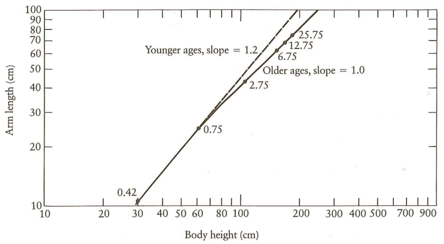
## Non-uniform growth:




p. 32, McMahon and Bonner [26]

# Non-uniform growth—arm length versus height:

## Good example of a break in scaling:

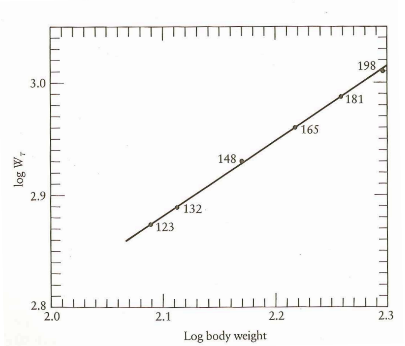



A crossover in scaling occurs around a height of 1 metre.

p. 32, McMahon and Bonner [26]



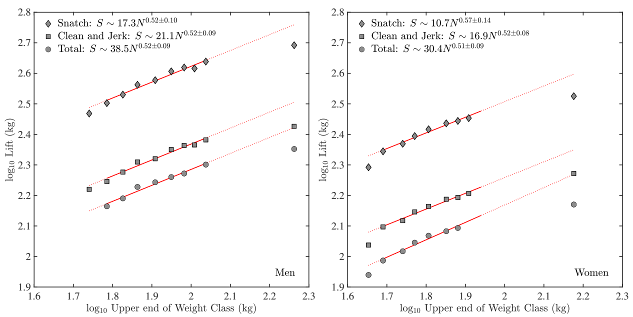
## Weightlifting: $M_{\text{world record}} \propto M_{\text{lifter}}^{2/3}$




Idea: Power ~ cross-sectional area of isometric lifters.
But modern data suggests an exponent of 1/2.

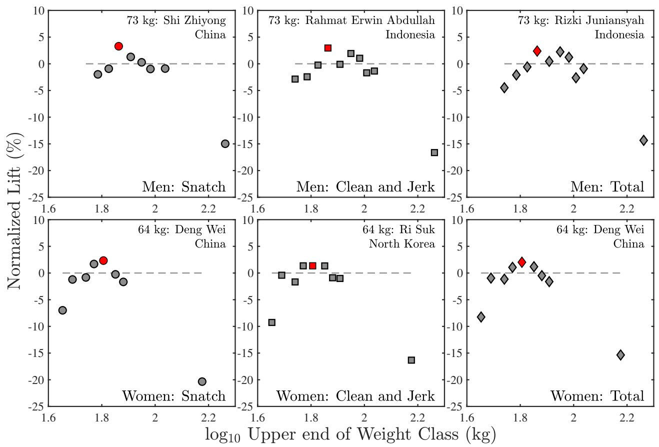
p. 53, McMahon and Bonner [26]

# Evidence for a 1/2 scaling exponent for weightlifting:




Li Wenwen's gold medal joy in Paris: Enjoy (at 2:25 with bonus Australian commentary).

# The "best" overall lifters:






# Stories—The Fraction Assassin:[2]

[1] *bonk bonk*



# Animal power

## Fundamental biological and ecological constraint:

$$P = c\,M^{\alpha}$$

$P$ = basal metabolic rate

$M$ = organismal body mass



# $P = c\,M^{\alpha}$

Prefactor $c$ depends on body plan and body temperature:

| | |
|---|---|
| Birds | 39–41 °$C$ |
| Eutherian Mammals | 36–38 °$C$ |
| Marsupials | 34–36 °$C$ |
| Monotremes | 30–31 °$C$ |



# What one might expect:

$\alpha = 2/3$ because ...

- Dimensional analysis suggests an energy balance surface law:

$$P \propto S \propto V^{2/3} \propto M^{2/3}$$

- Assumes isometric scaling (not quite the spherical cow).
- Lognormal fluctuations:
  Gaussian fluctuations in $\log P$ around $\log cM^{\alpha}$.
- Stefan-Boltzmann law for radiated energy:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \sigma\varepsilon S T^4 \propto S$$



# The prevailing belief of the Church of Quarterology:

$$\alpha = 3/4$$

$$P \propto M^{3/4}$$

Huh?

# The prevailing belief of the Church of Quarterology:



## Most obvious concern:

$$3/4 - 2/3 = 1/12$$

- An exponent higher than 2/3 points suggests a fundamental inefficiency in biology.
- Organisms must somehow be running 'hotter' than they need to balance heat loss.

# Related putative scalings:



## Wait! There's more!:

- number of capillaries $\propto M^{3/4}$
- time to reproductive maturity $\propto M^{1/4}$
- heart rate $\propto M^{-1/4}$
- cross-sectional area of aorta $\propto M^{3/4}$
- population density $\propto M^{-3/4}$

# The great 'law' of heartbeats:



## Assuming:

- Average lifespan $\propto M^{\beta}$
- Average heart rate $\propto M^{-\beta}$
- Irrelevant but perhaps $\beta = 1/4$.

## Then:

- Average number of heart beats in a lifespan
  $\simeq$ (Average lifespan) $\times$ (Average heart rate)
  $\propto M^{\beta-\beta}$
  $\propto M^{0}$
- Number of heartbeats per life time is independent of organism size!
- $\approx$ 1.5 billion....




# Ecology—Species-area law:



## Allegedly (data is messy): [21, 19]

"An equilibrium theory of insular zoogeography"
MacArthur and Wilson,
Evolution, **17**, 373–387, 1963. [21]

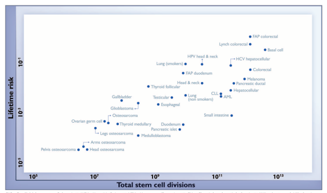
$$N_{\text{species}} \propto A^{\beta}$$

- According to physicists—on islands: $\beta \approx 1/4$.
- Also—on continuous land: $\beta \approx 1/8$.

# Cancer:



"Variation in cancer risk among tissues can be explained by the number of stem cell divisions"
Tomasetti and Vogelstein,
Science, **347**, 78–81, 2015. [37]




FAP = Familial Adenomatous Polyposis ◆ HCV = Hepatitis C virus ◆ HPV = Human papillomavirus ◆ CLL = Chronic lymphocytic leukemia ◆ AML = Acute myeloid leukemia

**Fig. 1. The relationship between the number of stem cell divisions in the lifetime of a given tissue and the lifetime risk of cancer in that tissue.** Values are from table S1, the derivation of which is discussed in the supplementary materials.

Roughly: $p \sim r^{2/3}$ where $p$ = life time probability and $r$ = rate of stem cell replication.



"How fast do living organisms move: Maximum speeds from bacteria to elephants and whales"
Meyer-Vernet and Rospars,
American Journal of Physics, **83**, 719–722, 2015. [28]

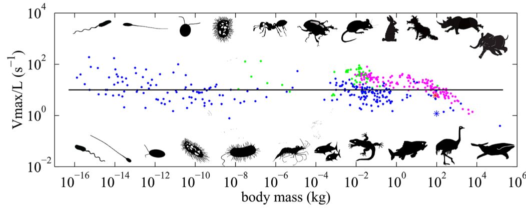



Fig. 1. Maximum relative speed versus body mass for 202 running species (157 mammals plotted in magenta and 45 non-mammals plotted in green), 127 swimming species and 91 micro-organisms (plotted in blue). The sources of the data are given in Ref. 16. The solid line is the maximum relative speed [Eq. (13)] estimated in Sec. III. The human world records are plotted as asterisks (upper for running and lower for swimming). Some examples of organisms of various masses are sketched in black (drawings by François Meyer).

Insert assignment question



"A general scaling law reveals why the largest animals are not the fastest"
Hirt et al.,
Nature Ecology & Evolution, **1**, 1116, 2017. [12]

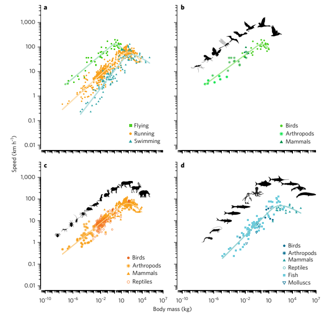



**Figure 2 | Empirical data and time-dependent model fit for the allometric scaling of maximum speed. a**, Comparison of scaling for the different locomotion modes (flying, running, swimming). **b–d**, Taxonomic differences are illustrated separately for flying (**b**; $n=55$), running (**c**; $n=458$) and swimming (**d**; $n=109$) animals. Overall model fit: $R^2=0.893$. The residual variation does not exhibit a signature of taxonomy (only a weak effect of thermoregulation; see Methods).



"A general scaling law reveals why the largest animals are not the fastest"
Hirt et al.,
Nature Ecology & Evolution, **1**, 1116, 2017. [12]

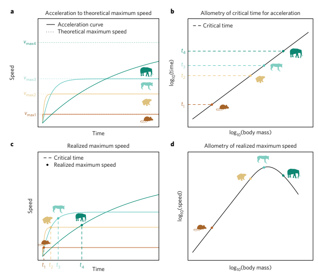



**Figure 1 | Concept of time-dependent and mass-dependent realized maximum speed of animals. a**, Acceleration of animals follows a saturation curve (solid lines) approaching the theoretical maximum speed (dotted lines) depending on body mass (colour code). **b**, The time available for acceleration increases with body mass following a power law. **c,d**, This critical time determines the realized maximum speed (**c**), yielding a hump-shaped increase of maximum speed with body mass (**d**).

## Theoretical story:

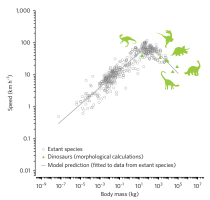


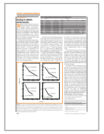

**Figure 4 | Predicting the maximum speed of extinct species with the time-dependent model.** The model prediction (grey line) is fitted to data of extant species (grey circles) and extended to higher body masses. Speed data for dinosaurs (green triangles) come from detailed morphological model calculations (values in Table 1) and were not used to obtain model parameters.

- Maximum speed increases with size: $v_{\max} = aM^b$
- Takes a while to get going: $v(t) = v_{\max}(1 - e^{-kt})$
- $k \sim F_{\max}/M \sim cM^{d-1}$
  Literature: $0.75 \lesssim d \lesssim 0.94$
- Acceleration time = depletion time for anaerobic energy: $\tau \sim fM^g$
  Literature: $0.76 \lesssim g \lesssim 1.27$
- $v_{\max} = aM^b\left(1 - e^{-hM^i}\right)$
- $i = d - 1 + g$ and $h = cf$

- Literature search for for maximum speeds of running, flying and swimming animals.
- Search terms: “maximum speed”, “escape speed”, and “sprint speed”.

Note: [28] not cited.





“Scaling in athletic world records”
Savaglio and Carbone,
Nature, **404**, 244, 2000. [34]

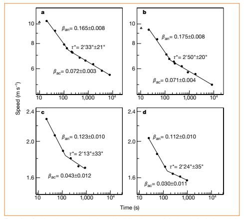


**Figure 1** Plots of world-record mean speeds against the record time (at November 1999). **a,b,** Running, and **c,d,** swimming records: for men (**a,c**), we consider 11 races (200 m, 400 m, 800 m, 1,000 m, 1,500 m, the mile, 3,000 m, 5,000 m, 10,000 m, 1 hour, and marathon); the same races are considered for women (**b,d**), apart from the 1 hour race. Lines represent the best fits. The scaling exponents $\beta$ and characteristic times $\tau^*$ of the breakpoints are shown; characteristic times have been determined by using a $\chi^2$ minimization on a broken power law. Triangles in **a,b** represent the 100 m race, which is excluded from the analysis because the mean speed is strongly affected by the standing start of athletes.

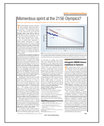
- Mean speed $\langle s \rangle$ decays with race time $\tau$:

$$\langle s \rangle \sim \tau^{-\beta}$$

- Break in scaling at around $\tau \simeq$ 150–170 seconds
- Anaerobic–aerobic transition
- Roughly 1 km running race
- Running decays faster than swimming

- Eek: Small scaling regimes





“Athletics: Momentous sprint at the 2156 Olympics?”
Tatem et al.,
Nature, **431**, 525–525, 2004. [36]

## Linear extrapolation for the 100 metres:

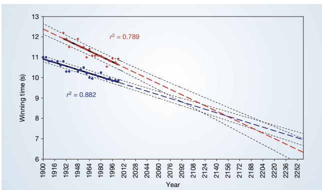
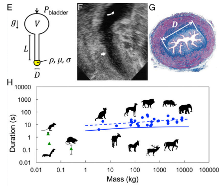



**Figure 1** The winning Olympic 100-metre sprint times for men (blue points) and women (red points), with superimposed best-fit linear regression lines (solid black lines) and coefficients of determination. The regression lines are extrapolated (broken blue and red lines for men and women, respectively) and 95% confidence intervals (dotted black lines) based on the available points are superimposed. The projections intersect just before the 2156 Olympics, when the winning women's 100-metre sprint time of 8.079 s will be faster than the men's at 8.098 s.

Tatem: “If I’m wrong anyone is welcome to come and question me about the result after the 2156 Olympics.”



“Duration of urination does not change with body size”,
Yang et al., Proceedings of the National Academy of Sciences, **111**, 11932–11937, 2014. [39]




- 32 mammals at Zoo Atlanta
- Figs. 1 and 2 are NSFTCR[3]
- $M = 3 \times 10^1$ g to $8 \times 10^6$ g
- For $\geq 3 \times 10^3$ g, $T \sim M^{1/6}$
- Duration $\sim 21 \pm 13$ seconds
- Smaller mammals: $T \sim M^0$
- Duration $\sim$ 0.02 to 2 seconds

**Table 1. Measured allometric relationships for the urinary system of animals**

| | Variable | Unit | Best fit | $R^2$ | $N$ |
|---|---|---|---|---|---|
| Duration of urination | $T$ | s | $8.2\ M^{0.13}$ | 0.2 | 32 |
| Urethral length | $L$ | mm | $35\ M^{0.43}$ | 0.9 | 47 |
| Urethral diameter | $D$ | mm | $2.0\ M^{0.39}$ | 0.9 | 22 |
| Shape factor | $\alpha$ | — | $0.2\ M^{-0.05}$ | 0.5 | 5 |
| Bladder capacity | $V$ | mL | $4.6\ M^{0.97}$ | 0.9 | 9 |
| Bladder pressure | $P_{bladder}$ | kPa | $5.2\ M^{-0.01}$ | 0.02 | 8 |
| Flow rate for females | $Q_F$ | mL/s | $1.8\ M^{0.66}$ | 0.9 | 16 |
| Flow rate for males | $Q_M$ | mL/s | $0.3\ M^{0.92}$ | 0.9 | 15 |

Body mass $M$ given in kilograms. Duration of urination corresponds to animals heavier than 3 kg. Urethral length and diameter, shape factor, bladder capacity, bladder pressure, and flow rates correspond to animals heavier than 0.02 kg.

[3] Not Safe For The Class Room



## Where this was always going:[4]

- Ig Nobel in Physics in 2015
- And again in 2019 for a paper on a peculiarity of wombats [?]

[4] David Hu’s papers on the fluid mechanics of interesting things



From How do wombats poop cubes? Scientists get to the bottom of the mystery, Science, 2021/01/27:

‘That just leaves one mystery: why wombats evolved cubic poop in the first place.

Hu speculates that because the animals climb up on rocks and logs to mark their territory, the flat-sided [poops] aren’t as likely to roll off from these high perches.

...

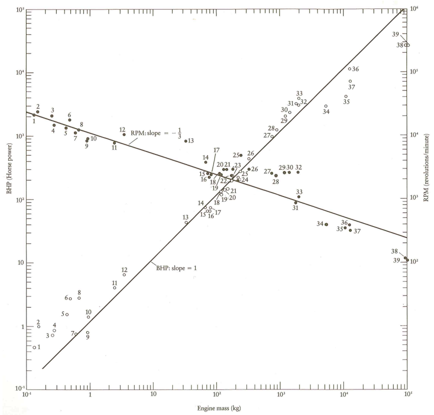
In the meantime, Hu also thinks this knowledge could help researchers raising wombats in captivity.

“Sometimes their [poops] aren’t as cubic as the [wild] ones,” he says.

The squarer the poop, the healthier the wombat.’



## Engines:

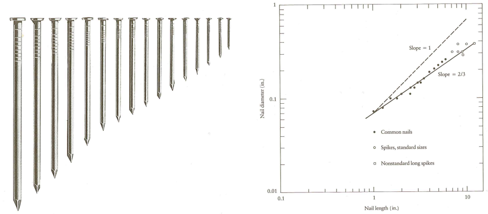



BHP = brake horse power



## The allometry of nails:

### Observed: Diameter $\propto$ Length$^{2/3}$ or $d \propto \ell^{2/3}$.




### Since $\ell d^2 \propto$ Volume $v$:

- Diameter $\propto$ Mass$^{2/7}$ or $d \propto v^{2/7}$.
- Length $\propto$ Mass$^{3/7}$ or $\ell \propto v^{3/7}$.
- Nails lengthen faster than they broaden (c.f. trees).

p. 58–59, McMahon and Bonner [26]



## The allometry of nails:

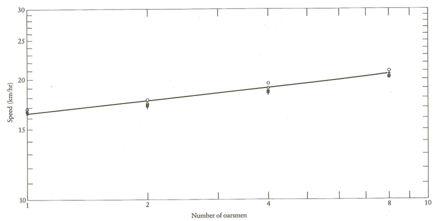
### A buckling instability?:

- Physics/Engineering result: Columns buckle under a load which depends on $d^4/\ell^2$.
- To drive nails in, posit resistive force $\propto$ nail circumference = $\pi d$.
- Match forces independent of nail size: $d^4/\ell^2 \propto d$.
- Leads to $d \propto \ell^{2/3}$.
- Argument made by Galileo [11] in 1638 in “Discourses on Two New Sciences.” Also, see here.
- Another smart person’s contribution: Euler, 1757
- Also see McMahon, “Size and Shape in Biology,” Science, 1973. [25]

## Rowing: Speed $\propto$ (number of rowers)$^{1/9}$

Shell dimensions and performances.

| No. of oarsmen | Modifying description | Length, $l$ (m) | Beam, $b$ (m) | $l/b$ | Boat mass per oarsman (kg) | Time for 2000 m (min) I | II | III | IV |
|---|---|---|---|---|---|---|---|---|---|
| 8 | Heavyweight | 18.28 | 0.610 | 30.0 | 14.7 | 5.87 | 5.92 | 5.82 | 5.73 |
| 8 | Lightweight | 18.28 | 0.598 | 30.6 | 14.7 | | | | |
| 4 | With coxswain | 12.80 | 0.574 | 22.3 | 18.1 | | | | |
| 4 | Without coxswain | 11.75 | 0.574 | 21.0 | 18.1 | 6.33 | 6.42 | 6.48 | 6.13 |
| 2 | Double scull | 9.76 | 0.381 | 25.6 | 13.6 | | | | |
| 2 | Pair-oared shell | 9.76 | 0.356 | 27.4 | 13.6 | 6.87 | 6.92 | 6.95 | 6.77 |
| 1 | Single scull | 7.93 | 0.293 | 27.0 | 16.3 | 7.16 | 7.25 | 7.28 | 7.17 |




- Very weak scaling and size variation but it's theoretically explainable ...



## Physics:

### Scaling in elementary laws of physics:

- Inverse-square law of gravity and Coulomb's law:

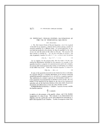
$$F \propto \frac{m_1 m_2}{r^2} \quad \text{and} \quad F \propto \frac{q_1 q_2}{r^2}.$$

- Force is diminished by expansion of space away from source.
- The square is $d - 1 = 3 - 1 = 2$, the dimension of a sphere's surface.
- We'll see a gravity law applies for a range of human phenomena.



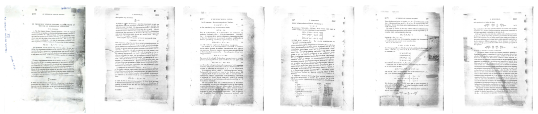
## Dimensional Analysis:

The Buckingham $\pi$ theorem:[5]



"On Physically Similar Systems: Illustrations of the Use of Dimensional Equations"
E. Buckingham,
Phys. Rev., **4**, 345–376, 1914. [7]

As captured in the 1990s in the MIT physics library:



[5] Stigler's Law of Eponymy applies yet again. See here. More later.



## Dimensional Analysis:[6]

### Fundamental equations cannot depend on units:

- System involves $n$ related quantities with some unknown equation $f(q_1, q_2, \ldots, q_n) = 0$.
- Geometric ex.: area of a square, side length $\ell$: $A = \ell^2$ where $[A] = L^2$ and $[\ell] = L$.
- Rewrite as a relation of $p \leq n$ independent dimensionless parameters where $p$ is the number of independent dimensions (mass, length, time, luminous intensity ...):

$$F(\pi_1, \pi_2, \ldots, \pi_p) = 0$$

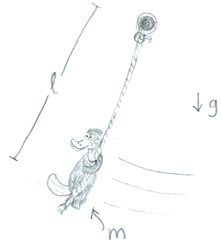
- e.g., $A/\ell^2 - 1 = 0$ where $\pi_1 = A/\ell^2$.
- Another example: $F = ma \Rightarrow F/ma - 1 = 0$.
- Plan: solve problems using only backs of envelopes.

[6] Length is a dimension, furlongs and smoots are units



## Example:

### Simple pendulum:




- Idealized mass/platypus swinging forever.
- Four quantities:
  1. Length $\ell$,
  2. mass $m$,
  3. gravitational acceleration $g$, and
  4. pendulum's period $\tau$.
- Variable dimensions: $[\ell] = L$, $[m] = M$, $[g] = LT^{-2}$, and $[\tau] = T$.
- Turn over your envelopes and find some $\pi$'s.



### A little formalism:

- Game: find all possible independent combinations of the $\{q_1, q_2, \ldots, q_n\}$, that form dimensionless quantities $\{\pi_1, \pi_2, \ldots, \pi_p\}$, where we need to figure out $p$ (which must be $\leq n$).
- Consider $\pi_i = q_1^{x_1} q_2^{x_2} \cdots q_n^{x_n}$.
- We (desperately) want to find all sets of powers $x_j$ that create dimensionless quantities.
- Dimensions: want $[\pi_i] = [q_1]^{x_1}[q_2]^{x_2} \cdots [q_n]^{x_n} = 1$.
- For the platypus pendulum we have $[q_1] = L$, $[q_2] = M$, $[q_3] = LT^{-2}$, and $[q_4] = T$,

  with dimensions $d_1 = L$, $d_2 = M$, and $d_3 = T$.
- So: $[\pi_i] = L^{x_1} M^{x_2} (LT^{-2})^{x_3} T^{x_4}$.
- We regroup: $[\pi_i] = L^{x_1 + x_3} M^{x_2} T^{-2x_3 + x_4}$.
- We now need: $x_1 + x_3 = 0$, $x_2 = 0$, and $-2x_3 + x_4 = 0$.
- Time for matrixology ...



### Well, of course there are matrices:

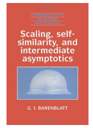
- Thrillingly, we have:

$$\mathbf{A}\vec{x} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- A nullspace equation: $\mathbf{A}\vec{x} = \vec{0}$.
- Number of dimensionless parameters = Dimension of null space = $n - r$ where $n$ is the number of columns of $\mathbf{A}$ and $r$ is the rank of $\mathbf{A}$.
- Here: $n = 4$ and $r = 3 \rightarrow F(\pi_1) = 0 \rightarrow \pi_1$ = const.
- In general: Create a matrix $\mathbf{A}$ where $ij$th entry is the power of dimension $i$ in the $j$th variable, and solve by row reduction to find basis null vectors.
- We (you) find: $\pi_1 = \ell/g\tau^2 = \text{const.}$ Upshot: $\tau \propto \sqrt{\ell}$.

  Insert assignment question

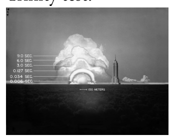





"Scaling, self-similarity, and intermediate asymptotics"
by G. I. Barenblatt (1996). [2]

### G. I. Taylor, magazines, and classified secrets:

1945
New Mexico
Trinity test:



Self-similar blast wave:

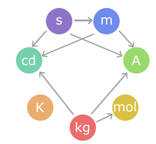
- Radius: $[R] = L$,
  Time: $[t] = T$,
  Density of air: $[\rho] = M/L^3$,
  Energy: $[E] = ML^2/T^2$.
- Four variables, three dimensions.
- One dimensionless variable: $E = \text{constant} \times \rho R^5/t^2$.
- Scaling: Speed decays as $1/R^{3/2}$.

Related: Radiolab's Elements on the Cold War, the Bomb Pulse, and the dating of cell age (33:30).

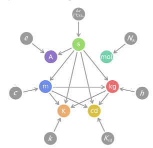


## Sorting out base units of fundamental measurement:

SI base units were redefined in 2019:




by Dono/Wikipedia

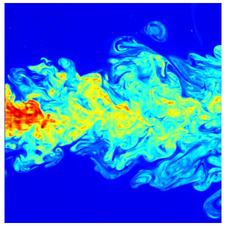


by Wikipetzi/Wikipedia

- Now: kilogram is an artifact in Sèvres, France.
- Defined by fixing Planck's constant as $6.62607015 \times 10^{-34}$ s$^{-1}$·m$^2$·kg.[3]
- Metre chosen to fix speed of light at 299,792,458 m·s$^{-1}$.
- Radiolab piece: $\leq$ kg

[3] Not without some arguing ...

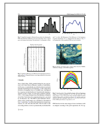
# Turbulence:



Big whirls have little whirls
That heed on their velocity,
And little whirls have littler whirls
And so on to viscosity.
— Lewis Fry Richardson

- Image from here.
- Jonathan Swift (1733): "Big fleas have little fleas upon their backs to bite 'em, And little fleas have lesser fleas, and so, ad infinitum." The Siphonaptera.





"Turbulent luminance in impassioned van Gogh paintings"
Aragón et al.,
J. Math. Imaging Vis., **30**, 275–283, 2008. [1]

- Examined the probability pixels a distance $R$ apart share the same luminance.
- "Van Gogh painted perfect turbulence" by Phillip Ball, July 2006.
- Apparently not observed in other famous painter's works or when van Gogh was stable.
- Oops: Small ranges and natural log used.



# Advances in turbulence:

## In 1941, Kolmogorov, armed only with dimensional analysis and an envelope figures this out: [18]

$$E(k) = C\epsilon^{2/3}k^{-5/3}$$

- $E(k)$ = energy spectrum function.
- $\epsilon$ = rate of energy dissipation.
- $k = 2\pi/\lambda$ = wavenumber.

- Energy is distributed across all modes, decaying with wave number.
- No internal characteristic scale to turbulence.
- Stands up well experimentally and there has been no other advance of similar magnitude.



## "The Geometry of Nature": Fractals

- "Anomalous" scaling of lengths, areas, volumes relative to each other.
- The enduring question: how do self-similar geometries form?[7]
- Robert E. Horton: Self-similarity of river (branching) networks (1945). [13]
- Harold Hurst—Roughness of time series (1951). [14]
- Lewis Fry Richardson—Coastlines (1961).
- Benoît B. Mandelbrot—Introduced the term "Fractals" and explored them everywhere, 1960s on. [22, 23, 24]

[7] Note to self: Make millions with the "Fractal Diet"



# Scaling in Cities:

"Growth, innovation, scaling, and the pace of life in cities"
Bettencourt et al.,
Proc. Natl. Acad. Sci., **104**, 7301–7306, 2007. [4]

- Quantified levels of
  - Infrastructure
  - Wealth
  - Crime levels
  - Disease
  - Energy consumption

  as a function of city size $N$ (population).



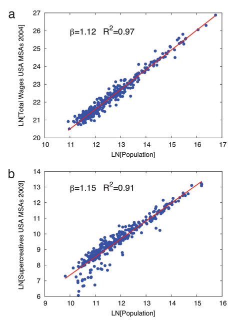



**Fig. 1.** Examples of scaling relationships. (a) Total wages per MSA in 2004 for the U.S. (blue points) vs. metropolitan population. (b) Supercreative employment per MSA in 2003, for the U.S. (blue points) vs. metropolitan population. Best-fit scaling relations are shown as solid lines.

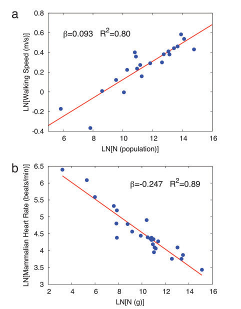



**Fig. 2.** The pace of urban life increases with city size in contrast to the pace of biological life, which decreases with organism size. (a) Scaling of walking speed vs. population for cities around the world. (b) Heart rate vs. the size (mass) of organisms.



# Scaling in Cities:

**Table 1. Scaling exponents for urban indicators vs. city size**

| $Y$ | $\beta$ | 95% CI | Adj-$R^2$ | Observations | Country–year |
|---|---|---|---|---|---|
| New patents | 1.27 | [1.25,1.29] | 0.72 | 331 | U.S. 2001 |
| Inventors | 1.25 | [1.22,1.27] | 0.76 | 331 | U.S. 2001 |
| Private R&D employment | 1.34 | [1.29,1.39] | 0.92 | 266 | U.S. 2002 |
| "Supercreative" employment | 1.15 | [1.11,1.18] | 0.89 | 287 | U.S. 2003 |
| R&D establishments | 1.19 | [1.14,1.22] | 0.77 | 287 | U.S. 1997 |
| R&D employment | 1.26 | [1.18,1.43] | 0.93 | 295 | China 2002 |
| Total wages | 1.12 | [1.09,1.13] | 0.96 | 361 | U.S. 2002 |
| Total bank deposits | 1.08 | [1.03,1.11] | 0.91 | 267 | U.S. 1996 |
| GDP | 1.15 | [1.06,1.23] | 0.96 | 295 | China 2002 |
| GDP | 1.26 | [1.09,1.46] | 0.64 | 196 | EU 1999–2003 |
| GDP | 1.13 | [1.03,1.23] | 0.94 | 37 | Germany 2003 |
| Total electrical consumption | 1.07 | [1.03,1.11] | 0.88 | 392 | Germany 2002 |
| New AIDS cases | 1.23 | [1.18,1.29] | 0.76 | 93 | U.S. 2002–2003 |
| Serious crimes | 1.16 | [1.11, 1.18] | 0.89 | 287 | U.S. 2003 |
| Total housing | 1.00 | [0.99,1.01] | 0.99 | 316 | U.S. 1990 |
| Total employment | 1.01 | [0.99,1.02] | 0.98 | 331 | U.S. 2001 |
| Household electrical consumption | 1.00 | [0.94,1.06] | 0.88 | 377 | Germany 2002 |
| Household electrical consumption | 1.05 | [0.89,1.22] | 0.91 | 295 | China 2002 |
| Household water consumption | 1.01 | [0.89,1.11] | 0.96 | 295 | China 2002 |
| Gasoline stations | 0.77 | [0.74,0.81] | 0.93 | 318 | U.S. 2001 |
| Gasoline sales | 0.79 | [0.73,0.80] | 0.94 | 318 | U.S. 2001 |
| Length of electrical cables | 0.87 | [0.82,0.92] | 0.75 | 380 | Germany 2002 |
| Road surface | 0.83 | [0.74,0.92] | 0.87 | 29 | Germany 2002 |

Data sources are shown in *SI Text*. CI, confidence interval; Adj-$R^2$, adjusted $R^2$; GDP, gross domestic product.



# Scaling in Cities:

## Intriguing findings:

- Global supply costs scale sublinearly with $N$ ($\beta < 1$).
  - Returns to scale for infrastructure.
- Total individual costs scale linearly with $N$ ($\beta = 1$)
  - Individuals consume similar amounts independent of city size.
- Social quantities scale superlinearly with $N$ ($\beta > 1$)
  - Creativity (# patents), wealth, disease, crime, ...

## Density doesn't seem to matter...

- Surprising given that across the world, we observe two orders of magnitude variation in area covered by agglomerations of fixed populations.



"Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities"
Bettencourt et al.,
PLoS ONE, **5**, e13541, 2010. [5]

## Comparing city features across populations:

- Cities = Metropolitan Statistical Areas (MSAs)
- Story: Fit scaling law and examine residuals
- Does a city have more or less crime than expected when normalized for population?
- Same idea as Encephalization Quotient (EQ).

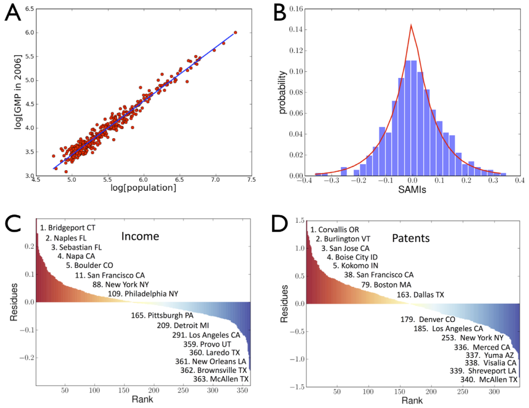
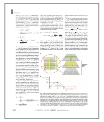



**Figure 1. Urban Agglomeration effects result in per capita nonlinear scaling of urban metrics.** Subtracting these effects produces a truly local measure of urban dynamics and a reference scale for ranking cities. a) A typical superlinear scaling law (solid line): Gross Metropolitan Product of US MSAs in 2006 (red dots) vs. population; the slope of the solid line has exponent, $\beta = 1.126$ (95% CI [1.101,1.149]). b) Histogram showing frequency of residuals, (SAMIs, see Eq. (2)); the statistics of residuals is well described by a Laplace distribution (red line). Scale independent ranking (SAMIs) for US MSAs by c) personal income and d) patenting (red denotes above average performance, blue below). For more details see Text S1, Table S1 and Figure S1.
doi:10.1371/journal.pone.0013541.g001



## A possible theoretical explanation?



"The origins of scaling in cities"
Luís M. A. Bettencourt,
Science, **340**, 1438–1441, 2013. [3]

#sixthology



## Non-simple scaling for death:

"Statistical signs of social influence on suicides"
Melo et al.,
Scientific Reports, **4**, 6239, 2014. [27]

- Bettencourt *et al.*'s initial work suggested social phenomena would follow superlinear scaling (wealth, crime, disease)
- Homicide, traffic, and suicide [10] all tied to social context in complex, different ways.
- For cities in Brazil, Melo *et al.* show:
  - Homicide appears to follow superlinear scaling ($\beta = 1.24 \pm 0.01$)
  - Traffic accident deaths appear to follow linear scaling ($\beta = 0.99 \pm 0.02$)
  - Suicide appears to follow sublinear scaling. ($\beta = 0.84 \pm 0.02$)

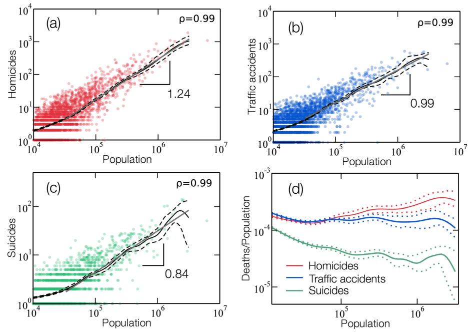



**Figure 1 | Scaling relations for homicides, traffic accidents, and suicides for the year of 2009 in Brazil.** The small circles show the total number of deaths by (a) homicides (red), (b) traffic accidents (blue), and (c) suicides (green) vs the population of each city. Each graph represents only one urban indicator, and the solid gray line indicate the best fit for a power-law relation, using OLS regression, between the average total number of deaths and the city size (population). To reduce the fluctuations we also performed a Nadaraya-Watson kernel regression[17,18]. The dashed lines show the 95% confidence band for the Nadaraya-Watson kernel regression. The ordinary least-squares (OLS)[19] fit to the Nadaraya-Watson kernel regression applied to the data on homicides in (a) reveals an allometric exponent $\beta = 1.24 \pm 0.01$, with a 95% confidence interval estimated by bootstrap. This is compatible with previous results obtained for U.S.[2] that also indicate a super-linear scaling relation with population and an exponent $\beta = 1.16$. Using the same procedure, we find $\beta = 0.99 \pm 0.02$ and $0.84 \pm 0.02$ for the numbers of deaths in traffic accidents (b) and suicides (c), respectively. The values of the Pearson correlation coefficients $\rho$ associated with these scaling relations are shown in each plot. This non-linear behavior observed for homicides and suicides certainly reflects the complexity of human social relations and strongly suggests that the the topology of the social network plays an important role on the rate of these events. (d) The solid lines show the Nadaraya-Watson kernel regression rate of deaths (total number of deaths divided by the population of a city) for each urban indicator, namely, homicides (red), traffic accidents (blue), and suicides (green). The dashed lines represent the 95% confidence bands. While the rate of fatal traffic accidents remains approximately invariant, the rate of homicides systematically increases, and the rate of suicides decreases with population.



Dynamics (Brazil):

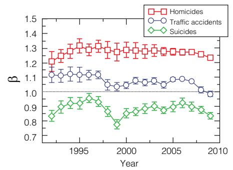
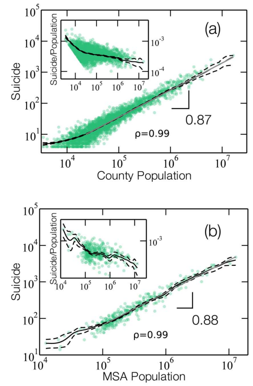


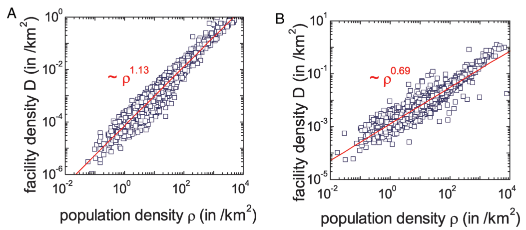

**Figure 2 | Temporal evolution of allometric exponent $\beta$ for homicides (red squares), deaths in traffic accidents (blue circles), and suicides (green diamonds).** Time evolution of the power-law exponent $\beta$ for each behavioral urban indicator in Brazil from 1992 to 2009. We can see that the non-linear behavior for homicides and suicides are robust for this 19 years period, and for the traffic accidents the exponent remain close to 1.0.

US data:






## Density of public and private facilities:

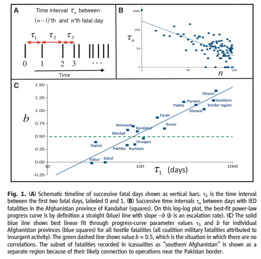



$$\rho_{\mathrm{fac}} \propto \rho_{\mathrm{pop}}^{\alpha}$$

- Left plot: ambulatory hospitals in the U.S.
- Right plot: public schools in the U.S.

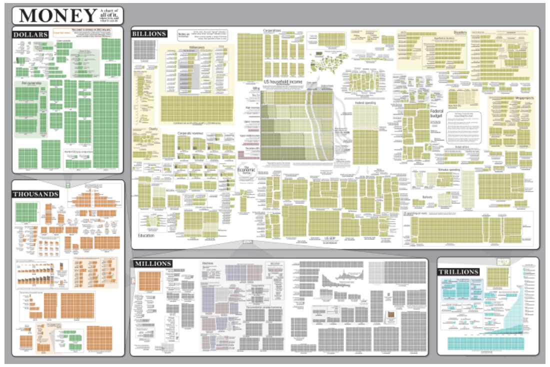


"Pattern in escalations in insurgent and terrorist activity"
Johnson et al.,
Science, **333**, 81–84, 2011. [16]



- Escalation: $\tau_n \sim \tau_1 n^{-b}$
- $b$ = scaling exponent (escalation rate)
- Interevent time $\tau_n$ between fatal attacks $n-1$ and $n$ (binned by days)
- Learning curves for organizations [38]
- More later on size distributions [9, 17, 6]






Explore the original zoomable and interactive version here: http://xkcd.com/980/.



# Irregular verbs

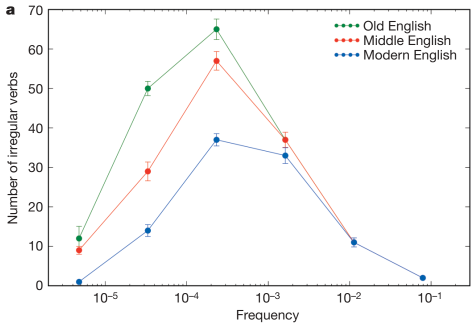
## Cleaning up the code that is English:

"Quantifying the evolutionary dynamics of language"
Lieberman et al.,
Nature, **449**, 713–716, 2007. [20]


- Exploration of how verbs with irregular conjugation gradually become regular over time.
- Comparison of verb behavior in Old, Middle, and Modern English.

# Irregular verbs

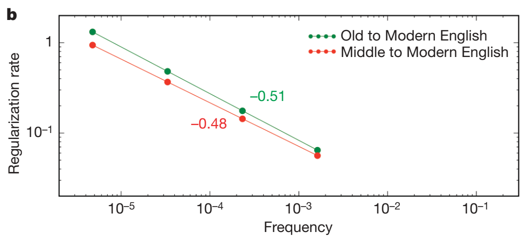





- Universal tendency towards regular conjugation
- Rare verbs tend to be regular in the first place

# Irregular verbs






- Rates are relative.
- The more common a verb is, the more resilient it is to change.

# Irregular verbs

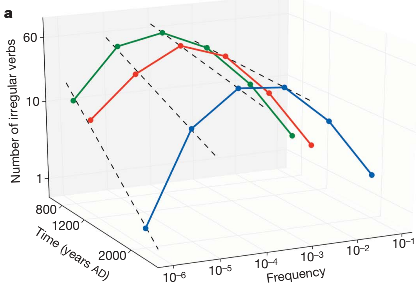
**Table 1 | The 177 irregular verbs studied**

| Frequency | Verbs | Regularization (%) | Half-life (yr) |
|---|---|---|---|
| $10^{-1}$–1 | be, have | 0 | 38,800 |
| $10^{-2}$–$10^{-1}$ | come, do, find, get, give, go, know, say, see, take, think | 0 | 14,400 |
| $10^{-3}$–$10^{-2}$ | begin, break, bring, buy, choose, draw, drink, drive, eat, fall, fight, forget, grow, hang, help, hold, leave, let, lie, lose, reach, rise, run, seek, set, shake, sit, sleep, speak, stand, teach, throw, understand, walk, win, work, write | 10 | 5,400 |
| $10^{-4}$–$10^{-3}$ | arise, bake, bear, beat, bind, bite, blow, bow, burn, burst, carve, chew, climb, cling, creep, dare, dig, drag, flee, float, flow, fly, fold, freeze, grind, leap, lend, lock, melt, reckon, ride, rush, shape, shine, shoot, shrink, sigh, sing, sink, slide, slip, smoke, spin, spring, starve, steal, step, stretch, strike, stroke, suck, swallow, swear, sweep, swim, swing, tear, wake, wash, weave, weep, weigh, wind, yell, yield | 43 | 2,000 |
| $10^{-5}$–$10^{-4}$ | bark, bellow, bid, blend, braid, brew, cleave, cringe, crow, dive, drip, fare, fret, glide, gnaw, grip, heave, knead, low, milk, mourn, mow, prescribe, redden, reek, row, scrape, seethe, shear, shed, shove, slay, slit, smite, sow, span, spurn, sting, stink, strew, stride, swell, tread, uproot, wade, warp, wax, wield, wring, writhe | 72 | 700 |
| $10^{-6}$–$10^{-5}$ | bide, chide, delve, flay, hew, rue, shrive, slink, snip, spew, sup, wreak | 91 | 300 |

177 Old English irregular verbs were compiled for this study. These are arranged according to frequency bin, and in alphabetical order within each bin. Also shown is the percentage of verbs in each bin that have regularized. The half-life is shown in years. Verbs that have regularized are indicated in red. As we move down the list, an increasingly large fraction of the verbs are red; the frequency-dependent regularization of irregular verbs becomes immediately apparent.

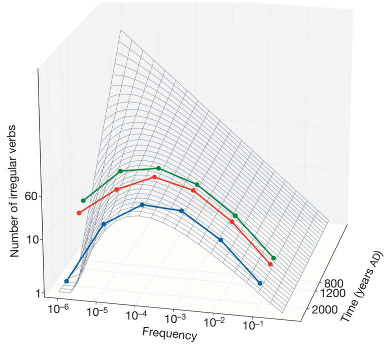
- Red = regularized
- Estimates of half-life for regularization ($\propto f^{1/2}$)






- 'Wed' is next to go.
- -ed is the winning rule...
- But 'snuck' is sneaking up on sneaked. [29]

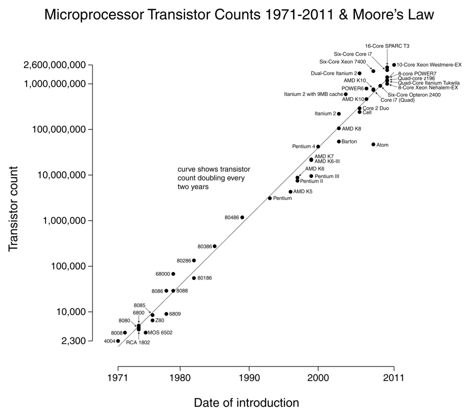


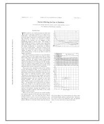



- Projecting back in time to proto-Zipf story of many tools.

# Moore's Law:

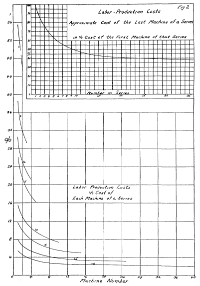


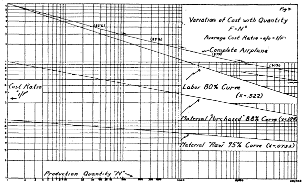


"Factors affecting the costs of airplanes"
T. P. Wright,
Journal of Aeronautical Sciences, **10**, 302–328, 1936. [38]







- Power law decay of cost with number of planes produced.
- "The present writer started his studies of the variation of cost with quantity in 1922."

# Scaling laws for technology production:



- "Statistical Basis for Predicting Technological Progress" Nagy et al., PLoS ONE, 2013. [31]
- $y_t$ = stuff unit cost; $x_t$ = total amount of stuff made.
- Wright's Law, cost decreases as a power of total stuff made: [38]

$$y_t \propto x_t^{-w}.$$

- Moore's Law, framed as cost decrease connected with doubling of transistor density every two years: [30]

$$y_t \propto e^{-mt}.$$

- Sahal's observation that Moore's law gives rise to Wright's law if stuff production grows exponentially: [33]

$$x_t \propto e^{gt}.$$

- Sahal + Moore gives Wright with $w = m/g$.



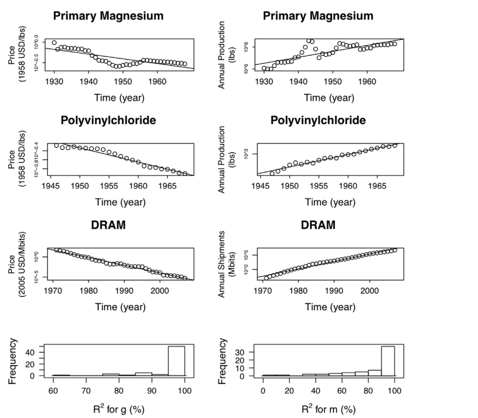



**Figure 3. Three examples showing the logarithm of price as a function of time in the left column and the logarithm of production as a function of time in the right column, based on industry-wide data.** We have chosen these examples to be representative: The top row contains an example with one of the worst fits, the second row an example with an intermediate goodness of fit, and the third row one of the best examples. The fourth row of the figure shows histograms of $R^2$ values for fitting $g$ and $m$ for the 62 datasets.
doi:10.1371/journal.pone.0052669.g003

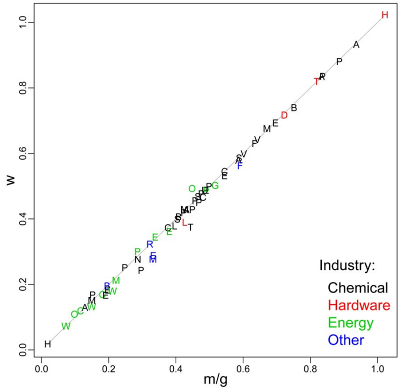
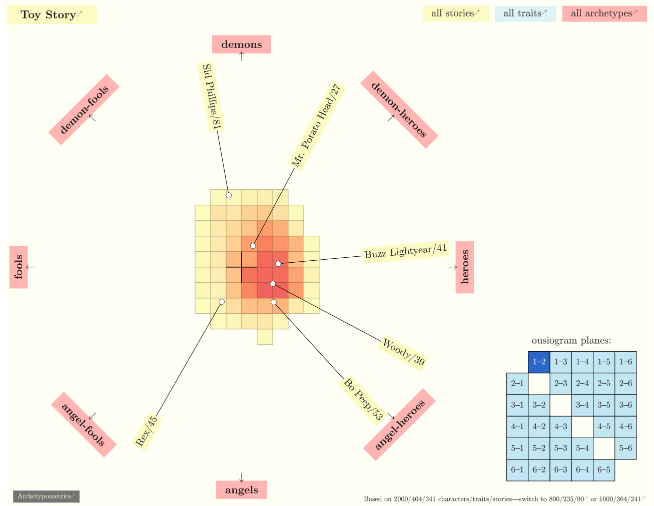



**Figure 4. An illustration that the combination of exponentially increasing production and exponentially decreasing cost are equivalent to Wright's law.** The value of the Wright parameter $w$ is plotted against the prediction $m/g$ based on the Sahal formula, where $m$ is the exponent of cost reduction and $g$ the exponent of the increase in cumulative production.
doi:10.1371/journal.pone.0052669.g004




https://compstorylab.org/archetypometrics/cards/Toy-Story-dimensions-1-vs-2-2000-464-341.pdf

## Toy Story and Moore's law:



'When the group moved to California to become part of Lucasfilm, we got close to making a computer-animated movie again in the mid-1980s – this time about a monkey with godlike powers but a missing prefrontal cortex. We had a sponsor, a story treatment, and a marketing survey. We were prepared to make a screen test: Our hot young animator John Lasseter had sketched numerous studies of the hero monkey and had the sponsor salivating over a glass-dragon protagonist.'

[7]"How Pixar Used Moore's Law to Predict the Future," Wired, 2013/04/17 https://www.wired.com/2013/04/how-pixar-used-moores-law-to-predict-the-future/

## Toy Story and Moore's law:



"But when it came time to harden the deal and run the numbers for the contracts, I discovered to my dismay that computers were still too slow: The projected production cost was too high and the computation time way too long. We had to back out of the deal. This time, we did know enough detail to correctly apply Moore's Law – and it told us that we had to wait another five years to start making the first movie. And sure enough, five years later Disney approached us to make Toy Story."

[7]"How Pixar Used Moore's Law to Predict the Future," Wired, 2013/04/17 https://www.wired.com/2013/04/how-pixar-used-moores-law-to-predict-the-future/

## Toy Story and Moore's law:



'We implement each step to see if it actually works, then gain the courage, the insight, and the engineering mastery to proceed to the next step.
Moore's Law told us that the new company we were starting, Pixar, had to bide its time—building hardware instead of making movies.'

[7]"How Pixar Used Moore's Law to Predict the Future," Wired, 2013/04/17 https://www.wired.com/2013/04/how-pixar-used-moores-law-to-predict-the-future/

## Toy Story and Moore's law:



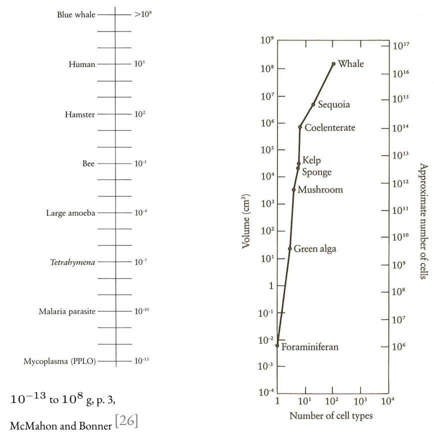
### Rhetoric of maybeness with hook to "More is different"

'That's the reason for expressing Moore's Law in orders of magnitude rather than factors of 10. The latter form is merely arithmetic, but the former implies an intellectual challenge. We use "order of magnitude" to imply a change so great that it requires new thought processes, new conceptualizations: It's not simply more, it's different.'

[7]"How Pixar Used Moore's Law to Predict the Future," Wired, 2013/04/17 https://www.wired.com/2013/04/how-pixar-used-moores-law-to-predict-the-future/

## Size range (in grams) and cell differentiation:






$10^{-13}$ to $10^{8}$ g, p. 3, McMahon and Bonner [26]

## Scaling of Specialization:



"Scaling of Differentiation in Networks: Nervous Systems, Organisms, Ant Colonies, Ecosystems, Businesses, Universities, Cities, Electronic Circuits, and Legos"
Changizi, McDannald, and Widders, J. Theor. Biol, **218**, 215–237, 2002. [8]

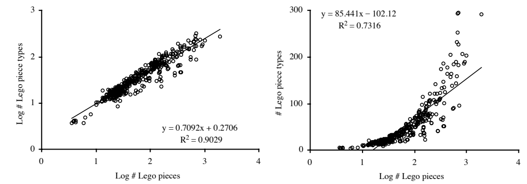



FIG. 3. Log–log (base 10) (left) and semi-log (right) plots of the number of Lego piece types vs. the total number of parts in Lego structures ($n = 391$). To help to distinguish the data points, logarithmic values were perturbed by adding a random number in the interval [−0.05, 0.05], and non-logarithmic values were perturbed by adding a random number in the interval [−1, 1].

- 2012 wired.com write-up

## $C \sim N^{1/d}, d \geq 1$:



- $C$ = network differentiation = # node types.
- $N$ = network size = # nodes.
- $d$ = combinatorial degree.
- Low $d$: strongly specialized parts.
- High $d$: strongly combinatorial in nature, parts are reused.
- Claim: Natural selection produces high $d$ systems.
- Claim: Engineering/brains produces low $d$ systems.
- For language: See the naturally-incorrectly-attributed[8] Heaps' Law
- Most generally: $N_{\textrm{types}} \sim N^{\beta}_{\textrm{things}}$ where $0 < \beta \leq 1$. More later.

[8]Plus one for Stigler's Law of Eponymy. More later.

TABLE 1
*Summary of results**

| Network | Node | No. data points | Range of log $N$ | Log–log $R^2$ | Semi-log $R^2$ | $p_{power}/p_{log}$ | Relationship between $C$ and $N$ | Comb. degree | Exponent $v$ for type-net scaling | Figure in text |
|---|---|---|---|---|---|---|---|---|---|---|
| *Selected networks* | | | | | | | | | | |
| Electronic circuits | Component | 373 | 2.12 | 0.747 | 0.602 | 0.05/4e−5 | Power law | 2.29 | 0.92 | 2 |
| Legos™ | Piece | 391 | 2.65 | 0.903 | 0.732 | 0.09/1e−7 | Power law | 1.41 | — | 3 |
| Businesses | | | | | | | | | | |
| military vessels | Employee | 13 | 1.88 | 0.971 | 0.832 | 0.05/3e−3 | Power law | 1.60 | — | 4 |
| military offices | Employee | 8 | 1.59 | 0.964 | 0.789 | 0.16/0.16 | Increasing | 1.13 | — | 4 |
| universities | Employee | 9 | 1.55 | 0.786 | 0.749 | 0.27/0.27 | Increasing | 1.37 | — | 4 |
| insurance co. | Employee | 52 | 2.30 | 0.748 | 0.685 | 0.11/0.10 | Increasing | 3.04 | — | 4 |
| Universities | | | | | | | | | | |
| across schools | Faculty | 112 | 2.72 | 0.695 | 0.549 | 0.09/0.01 | Power law | 1.81 | — | 5 |
| history of Duke | Faculty | 46 | 0.94 | 0.921 | 0.892 | 0.09/0.05 | Increasing | 2.07 | — | 5 |
| Ant colonies | | | | | | | | | | |
| caste = type | Ant | 46 | 6.00 | 0.481 | 0.454 | 0.11/0.04 | Power law | 8.16 | — | 6 |
| size range = type | Ant | 22 | 5.24 | 0.658 | 0.548 | 0.17/0.04 | Power law | 8.00 | — | 6 |
| Organisms | Cell | 134 | 12.40 | 0.249 | 0.165 | 0.08/0.02 | Power law | 17.73 | — | 7 |
| Neocortex | Neuron | 10 | 0.85 | 0.520 | 0.584 | 0.16/0.16 | Increasing | 4.56 | — | 9 |
| *Competitive networks* | | | | | | | | | | |
| Biotas | Organism | — | — | — | — | — | Power law | ≈3 | 0.3 to 1.0 | — |
| Cities | Business | 82 | 2.44 | 0.985 | 0.832 | 0.08/8e-8 | Power law | 1.56 | — | 10 |

*(1) The kind of network, (2) what the nodes are within that kind of network, (3) the number of data points, (4) the logarithmic range of network sizes $N$ (i.e. $\log(N_{max}/N_{min})$), (5) the log–log correlation, (6) the semi-log correlation, (7) the serial-dependence probabilities under, respectively, power-law and logarithmic models, (8) the empirically determined best-fit relationship between differentiation $C$ and organization size $N$ (if one of the two models can be refuted with $p<0.05$; otherwise we just write "increasing" to denote that neither model can be rejected), (9) the combinatorial degree (i.e. the inverse of the best-fit slope of a log–log plot of $C$ versus $N$), (10) the scaling exponent for how quickly the edge-degree $\delta$ scales with type-network size $C$ (in those places for which data exist), (11) figure in this text where the plots are presented. Values for biotas represent the broad trend from the literature.



# A key framing from language:

## Types and Tokens:

- In linguistics, words are described on the two levels of types and tokens [32].
- In semiotics, signs can be thought of having two components of the signified and the signifier.

## Example:

- Types are 1-grams, e.g., '!', 'the', 'love', and 'spork'.
- Tokens are 1-grams as written down.
- In "Pride and Prejudice", for example, there are 498 '!'s, 4,058 'the's, 90 'love's, and 0 'spork's.



# Types and Things and Measures, Oh My!

## Beyond language:

Lift out and expand the type-token framing to complex systems in general.

## ~~Three~~ Four possible parts:

1. Type: A kind or class of category of individual things based on shared characteristics.
2. Thing: An individual manifestation of a type.
3. Measure: A quantification of the manifestation of things.
4. Experience: An interaction of any kind with a manifestation of a type.[9]

[9] Fame.



## Language:

1. Type: A defined word.
2. Thing (token): An instance of spoken or printed word.
3. Number or Frequency (counts of tokens).
4. Experience: Listening to others, reading a book.

## Atoms:

1. Type: Atom
2. Thing: Element (stuff made of a given atom; e.g., gold)
3. Measure: Mass; could be Number.
4. Experience: Atomic bonds.



## Water:

1. Type: Water molecule, $H^2O$.
2. Thing: Water.
3. Measure: Volume (liters, gallons); given pressure and temperature, equivalent to Number (counts of molecules) and then Mass.
4. Experience: Rain.

## Biology:

- Example type: The species Ornithorhynchus anatinus, the platypus.
- Thing: Any given platypus.
- Measure: The number of platypuses ('instances' of the species) living in Australia in the wild.
- Experience: Seeing a platypus in the wild; being hunted by a platypus.



## Moneyspace:

- Example type: Corporation.
- Things: The publicly traded companies of Apple and Microsoft.
- Measure: Market capitalization.
- Experience: Being sued by Microsoft.

- Apple and Microsoft may be viewed as components of the publicly-owned corporate world.
- The sizes of corporations may be broken down into many rankable dimensions such as annual revenue or number of employees worldwide.
- In principle, market capitalization represents a kind of current collective belief in terms of money.



## Sizes and Rankings:

- We will often consider systems where each component type $\tau$ has at least one measurable—and hence rankable—'size' $s_\tau$.
- Perceived size is a combination of Measure (what exists) and Experience (what is measured).
- Important: We may also have rankings where we do not know the underlying 'size' (e.g., book/thing sales on Amazon).



## Three examples which show some of the range of what 'size' can mean:

1. Size for a word in a corpus means the number of indistinguishable instances of that word (many identical entites—tokens);
2. Size for species means the number of 'biological replications' of an individual type (many genetically similar entities of varying ages); and
3. Size for a corporation might mean monetary value (market cap, one entity).
4. May have more than one measure of a system:
   - Total biomass of a species.[10]
   - Number of employees in a corporation.
   - Number of stars in a galaxy.[10]
5. Measure of size allows for rankings.
6. Again, sizes may be hidden.

[10] Somewhat hard to estimate.



## When tokens are fungible:

- Randomly permute all of the words (tokens) of the same type in Pride and Prejudice.
- Measure and Experience will be unchanged.
- NFTs: Non-fungible tokens.
- Tricking people into thinking tokens are types.
- "The Oxymoron for Morons."

## When tokens are funguses:

- NFF: Non-fungible fungus (from a sentient fungus's point of view).
- But in cooking, funguses are fungible.
- Lack of exposure leads to fungibility of "the other."[11]

[11] Universal: Identical twins look the same until they don't.

### Shell of the nut:

- Scaling is a fundamental feature of complex systems.
- Basic distinction between isometric and allometric scaling.
- Powerful envelope-based approach: Dimensional analysis.
- "Oh yeah, well that's just dimensional analysis" said the [insert your own adjective] physicist.
- Tricksiness: A wide variety of mechanisms give rise to scalings.[12]
- Some mechanisms are common, some are rare.[13]

[12] It's not your great-great-great-grandparents' normal distribution
[13] To be understood: The scaling story of scaling-making mechanisms



## References I

[1] J. L. Aragón, G. G. Naumis, M. Bai, M. Torres, and P. K. Maini.
Turbulent luminance in impassioned van Gogh paintings.
J. Math. Imaging Vis., 30:275–283, 2008. pdf

[2] G. I. Barenblatt.
Scaling, self-similarity, and intermediate asymptotics, volume 14 of Cambridge Texts in Applied Mathematics.
Cambridge University Press, 1996.

[3] L. M. A. Bettencourt.
The origins of scaling in cities.
Science, 340:1438–1441, 2013. pdf

[4] L. M. A. Bettencourt, J. Lobo, D. Helbing, Kühnhert, and G. B. West.
Growth, innovation, scaling, and the pace of life in cities.
Proc. Natl. Acad. Sci., 104(17):7301–7306, 2007. pdf



## References II

[5] L. M. A. Bettencourt, J. Lobo, D. Strumsky, and G. B. West.
Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities.
PLoS ONE, 5:e13541, 2010. pdf

[6] J. C. Bohorquez, S. Gourley, A. R. Dixon, M. Spagat, and N. F. Johnson.
Common ecology quantifies human insurgency.
Nature, 462:911–914, 2009. pdf

[7] E. Buckingham.
On physically similar systems: Illustrations of the use of dimensional equations.
Phys. Rev., 4:345–376, 1914. pdf



## References III

[8] M. A. Changizi, M. A. McDannald, and D. Widders.
Scaling of differentiation in networks: Nervous systems, organisms, ant colonies, ecosystems, businesses, universities, cities, electronic circuits, and Legos.
J. Theor. Biol, 218:215–237, 2002. pdf

[9] A. Clauset, M. Young, and K. S. Gleditsch.
On the Frequency of Severe Terrorist Events.
Journal of Conflict Resolution, 51(1):58–87, 2007. pdf

[10] E. Durkheim.
Suicide: A study in sociology.
Free Press, 2005.
Reissue edition (February 1, 1997).



## References IV

[11] G. Galilei.
Dialogues Concerning Two New Sciences.
Kessinger Publishing, 2010.
Translated by Henry Crew and Alfonso De Salvio.

[12] M. R. Hirt, W. Jetz, B. C. Rall, and U. Brose.
A general scaling law reveals why the largest animals are not the fastest.
Nature Ecology & Evolution, 1:1116, 2017. pdf

[13] R. E. Horton.
Erosional development of streams and their drainage basins; hydrophysical approach to quatitative morphology.
Bulletin of the Geological Society of America, 56(3):275–370, 1945. pdf



## References V

[14] H. E. Hurst.
Long term storage capacity of reservoirs.
Transactions of the American Society of Civil Engineers, 116:770–808, 1951.

[15] J. S. Huxley and G. Teissier.
Terminology of relative growth.
Nature, 137:780–781, 1936. pdf

[16] N. Johnson, S. Carran, J. Botner, K. Fontaine, N. Laxague, P. Nuetzel, J. Turnley, and B. Tivnan.
Pattern in escalations in insurgent and terrorist activity.
Science, 333:81–84, 2011. pdf

[17] N. F. Johnson, M. Spagat, J. A. Restrepo, O. Becerra, J. C. Bohorquez, N. Suarez, E. M. Restrepo, and R. Zarama.
Universal patterns underlying ongoing wars and terrorism, 2006. pdf



## References VI

[18] A. N. Kolmogorov.
The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers.
Proceedings of the USSR Academy of Sciences, 30:299–303, 1941.

[19] S. Levin.
The problem of pattern and scale in ecology.
Ecology, 73(6):1943–1967, 1992.
. pdf

[20] E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, and M. A. Nowak.
Quantifying the evolutionary dynamics of language.
Nature, 449:713–716, 2007. pdf



## References VII

[21] R. H. MacArthur and E. O. Wilson.
An equilibrium theory of insular zoogeography.
Evolution, 17:373–387, 1963. pdf

[22] B. B. Mandelbrot.
How long is the coast of britain? statistical self-similarity and fractional dimension.
Science, 156(3775):636–638, 1967. pdf

[23] B. B. Mandelbrot.
Fractals: Form, Chance, and Dimension.
Freeman, San Francisco, 1977.

[24] B. B. Mandelbrot.
The Fractal Geometry of Nature.
Freeman, San Francisco, 1983.



## References VIII

[25] T. McMahon.
Size and shape in biology.
Science, 179:1201–1204, 1973. pdf

[26] T. A. McMahon and J. T. Bonner.
On Size and Life.
Scientific American Library, New York, 1983.

[27] H. P. M. Melo, A. A. Moreira, É. Batista, H. A. Makse, and J. S. Andrade.
Statistical signs of social influence on suicides.
Scientific Reports, 4:6239, 2014. pdf

[28] N. Meyer-Vernet and J.-P. Rospars.
How fast do living organisms move: Maximum speeds from bacteria to elephants and whales.
American Journal of Physics, pages 719–722, 2015. pdf

# References IX



[29] J.-B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, T. G. B. Team, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M. A. Nowak, and E. A. Lieberman.
Quantitative analysis of culture using millions of digitized books.
Science, 2010. pdf

[30] G. E. Moore.
Cramming more components onto integrated circuits.
Electronics Magazine, 38:114–117, 1965.

[31] B. Nagy, J. D. Farmer, Q. M. Bui, and J. E. Trancik.
Statistical basis for predicting technological progress.
PloS one, 8(2):e52669, 2013. pdf

[32] C. S. S. Peirce.
Prolegomena to an apology for pragmaticism.
The Monist, 16(4):492–546, 1906. pdf

# References X



[33] D. Sahal.
A theory of progress functions.
AIIE Transactions, 11:23–29, 1979.

[34] S. Savaglio and V. Carbone.
Scaling in athletic world records.
Nature, 404:244, 2000. pdf

[35] A. Shingleton.
Allometry: The study of biological scaling.
Nature Education Knowledge, 1:2, 2010.

[36] A. J. Tatem, C. A. Guerra, P. M. Atkinson, and S. I. Hay.
Athletics: Momentous sprint at the 2156 Olympics?
Nature, 431(7008):525–525, 2004. pdf

# References XI



[37] C. Tomasetti and B. Vogelstein.
Variation in cancer risk among tissues can be explained by the number of stem cell divisions.
Science, 347:78–81, 2015. pdf

[38] T. P. Wright.
Factors affecting the costs of airplanes.
Journal of Aeronautical Sciences, 10:302–328, 1936. pdf

[39] P. J. Yang, J. Pham, J. Choo, and D. L. Hu.
Duration of urination does not change with body size.
Proceedings of the National Academy of Sciences, 111:11932–11937, 2014. pdf

# References XII



[40] K. Zhang and T. J. Sejnowski.
A universal scaling law between gray matter and white matter of cerebral cortex.
Proceedings of the National Academy of Sciences, 97:5621–5626, 2000. pdf