# Presenting Results

### 1. Visualization

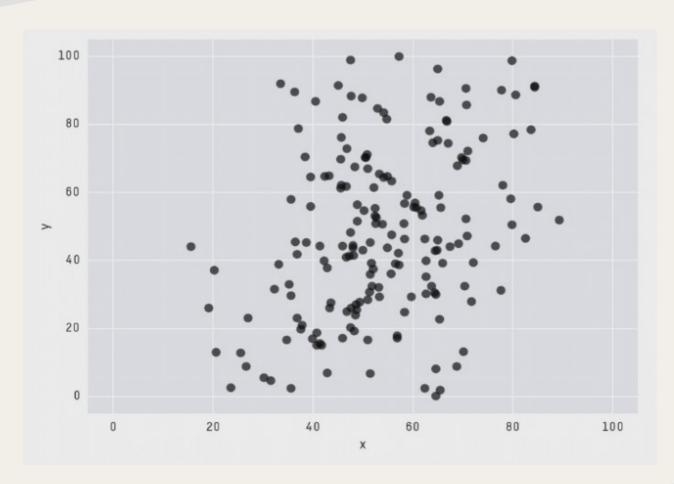
- "At their best, graphics are instruments for reasoning."
  - Edward Tufte
- "Numerical quantities focus on expected values, graphical summaries on unexpected values."
  - John Tukey
- "The greatest value of a picture is when it forces us to notice what we never expected to see."
  - John Tukey
- "If I can't picture it, I can't understand it."
  - Albert Einstein

Given two variables where:

$$\bar{X} = 54.26 \ (SD = 16.76)$$
 $\bar{Y} = 47.83 \ (SD = 26.93)$ 
 $r_{X,Y} = -.06$ 

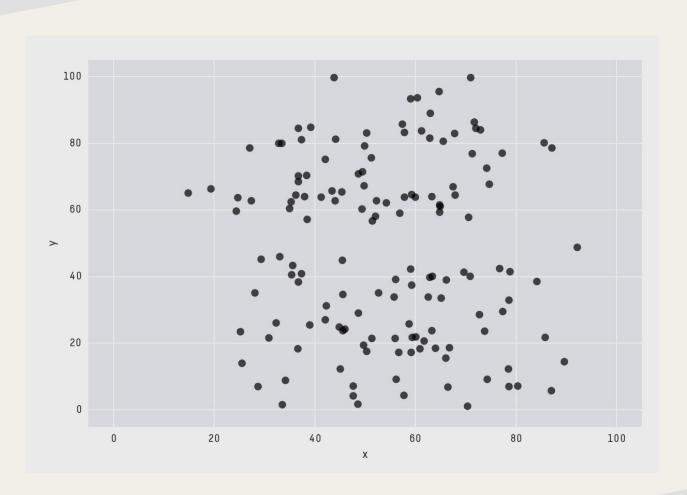
What do these data look like?

### This?



No,  $\overline{X} = 54.02$ ,  $\overline{Y} = 48.09$ ,  $r_{X,Y} = .32$ 

### This?



Yes,  $\overline{X} = 54.26$ ,  $\overline{Y} = 47.83$ ,  $r_{X,Y} = -.06$ 

## Or this?



Yes,  $\overline{X}$  = 54.26,  $\overline{Y}$  = 47.83,  $r_{X,Y}$  = -.06

### Elements of Good Data Visualization

- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

## Bad Graphs

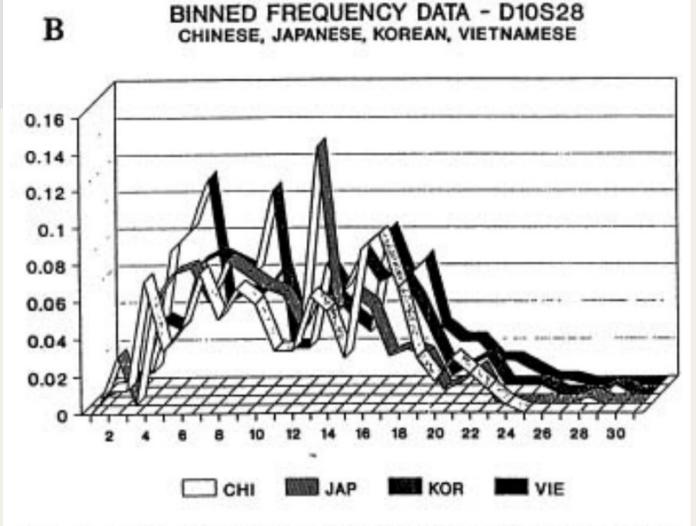


Fig. 4. Fixed bin distribution (histogram) for two loci and four Asian subpopulations (used with permission from John Hartmann): the boundaries of the 30 bins (vertical axis) are determined by the FBI; these bins are not of equal length. Sample sizes (numbers of individuals) for Chinese, Japanese, Korean and Vietnamese are 103, 125, 93 and 215 for D4S139 and 120, 137, 100 and 193 for D10S28. The horizontal axis is the bin number; bins are not of equal length.

#### Distribution of All TFBS Regions

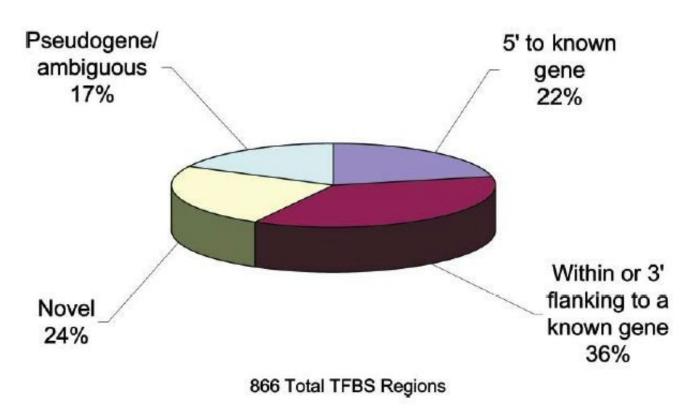


Figure 1. Classification of TFBS Regions

TFBS regions for Sp1, cMyc, and p53 were classified based upon proximity to annotations (RefSeq, Sanger hand-curated annotations, GenBank full-length mRNAs, and Ensembl predicted genes). The proximity was calculated from the center of each TFBS region. TFBS regions were classified as follows: within 5 kb of the 5' most exon of a gene, within 5 kb of the 3' terminal exon, or within a gene, novel or outside of any annotation, and pseudogene/ambiguous (TFBS overlapping or flanking pseudogene annotations, limited to chromosome 22, or TFBS regions falling into more than one of the above categories).

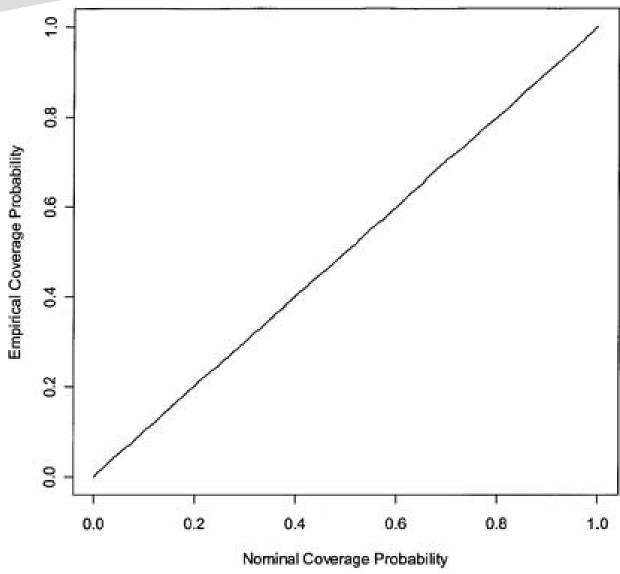


Figure 1 Empirical coverage of CIs for the relative-risk parameter  $\beta$  of haplotype 01100. Results are based on 10,000 simulated data sets with the same haplotype frequencies as the FUSION data. Haplotype 01100 has a multiplicative effect on disease risk, with  $\beta = 0.35$ .

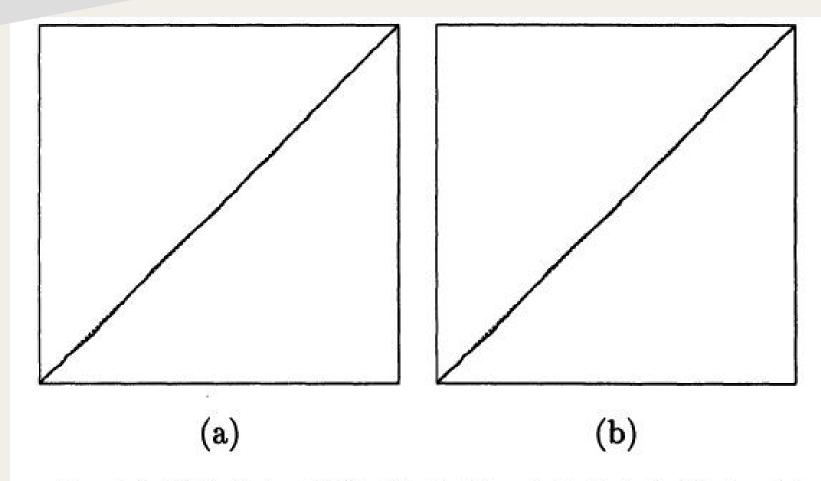
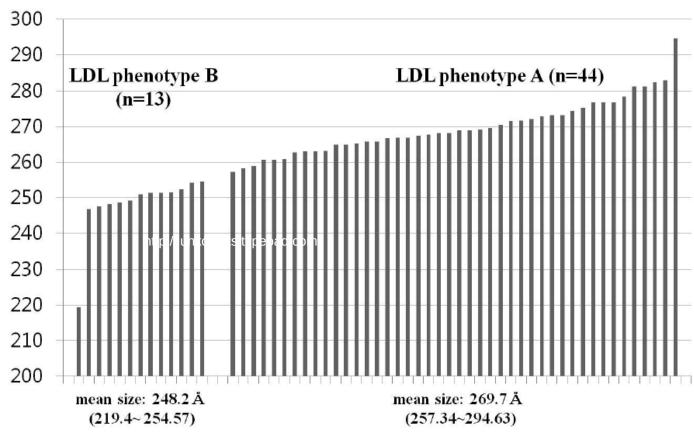


Figure 1. SRQ Plots of T<sub>i</sub>/T<sub>n</sub> (Vertical Axes) Against i/n (Horizontal Axes) for the Gibbs Sampler (a) and an Alternating Gibbs/Independence Sampler (b) for the Pump Failure Data Based on Runs of Length 5,000. Lines through the origin with unit slope are shown dashed; axis ranges are from 0 to 1 for all axes.



**Fig. 1.** Distribution of low-density lipoprotein (LDL) particle size in all study subjects (LDL phenotypes A and B). LDL phenotype A group (mean size: 269.7 Å, n = 44), subjects with buoyant-mode profiles [peak LDL particle diameter  $\geq 264$  Å] including intermediate LDL subclass pattern [256 Å  $\leq$  peak LDL particle diameter  $\leq 263$  Å]; LDL phenotype B group (mean size: 248.2 Å, n = 13), subjects with dense-mode profiles [peak LDL particle diameter  $\leq 255$  Å]

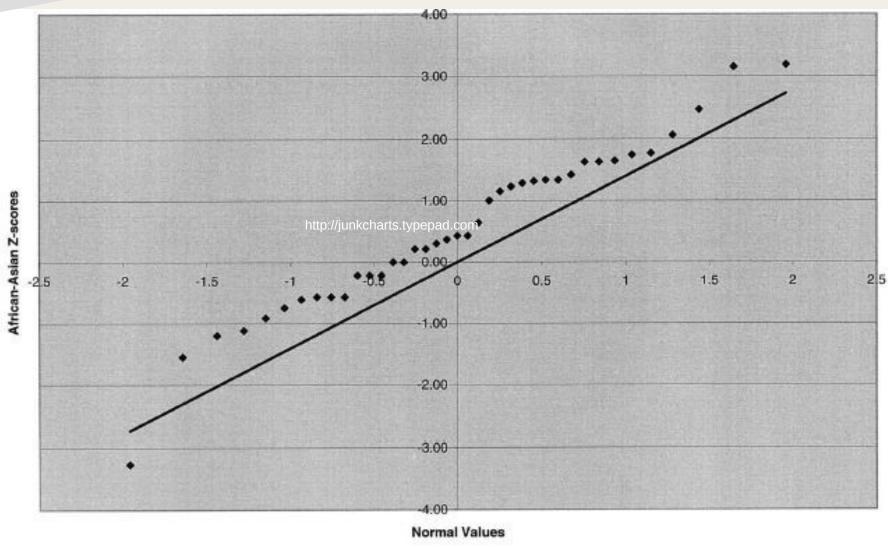
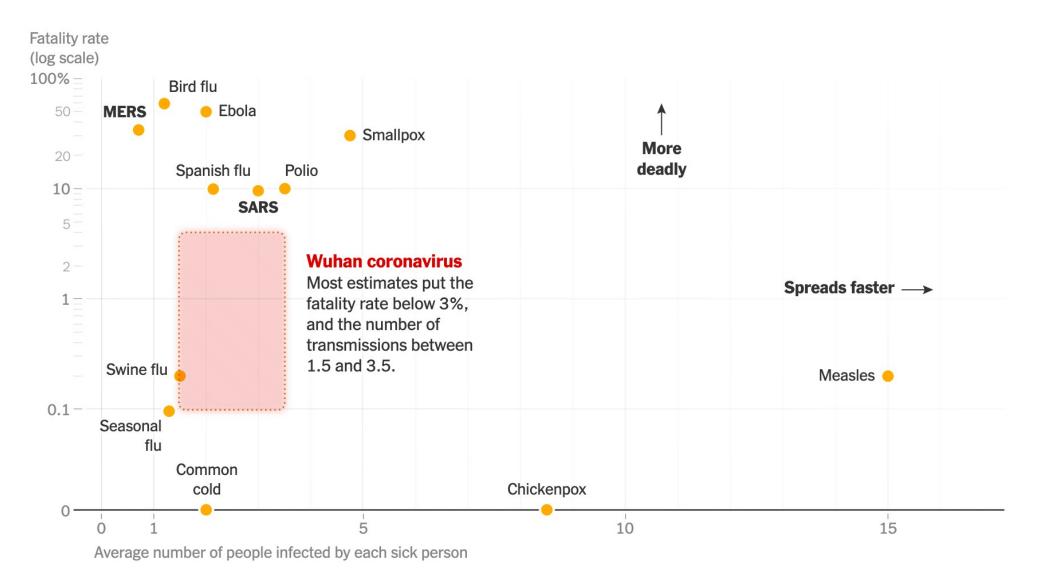


Figure 2 Q-Q plots of Z scores for telomeric interval-length differences. a, African Americans versus Asians. b, Whites versus Asians.



Note: Average case-fatality rates and transmission numbers are shown. Estimates of case-fatality rates can vary, and numbers for the Wuhan coronavirus are preliminary estimates.

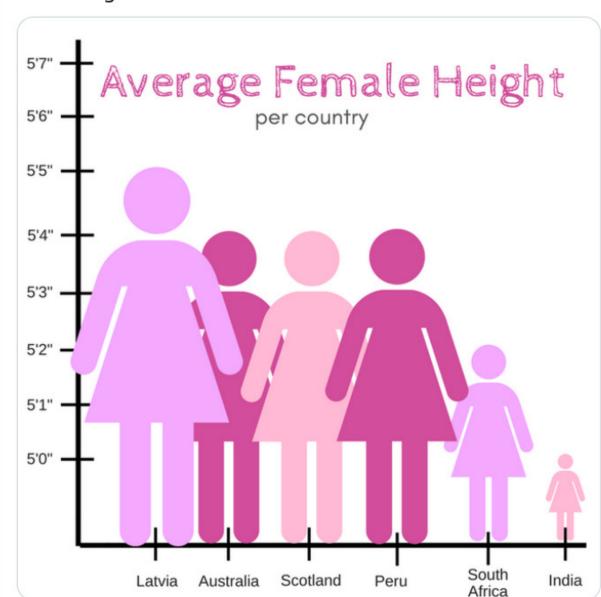






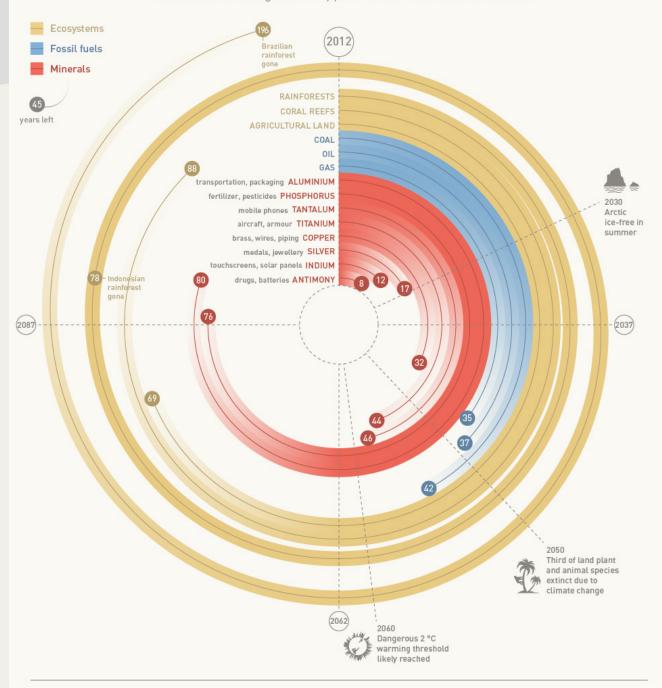


As an Indian woman, I can confirm that too much of my time is spent hiding behind a rock praying the terrifying gang of international giant ladies and their Latvian general don't find me



#### Stock Check

Estimated remaining world supplies of non-renewable resources



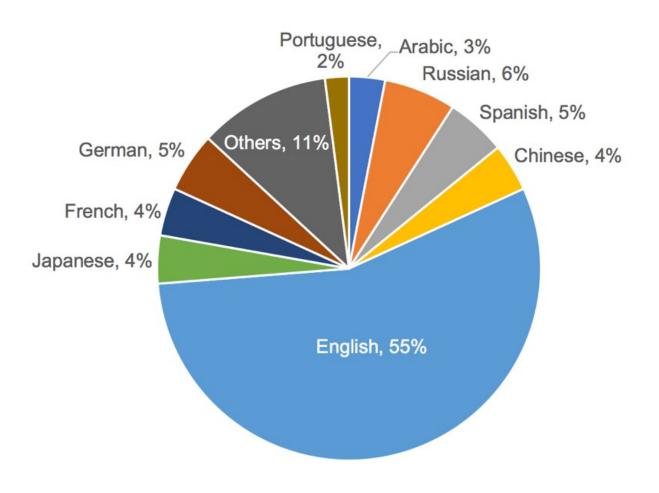
Sources: UN TEEB, US Geological Survey, BP, Worm et al (2006), London Metal Exchange. Figures are wortdwide. Living natural resources dates are worst-case based on published estimates. Minerals and fossil fuel data based on known reserves currently economical to extract, assuming fixed % increase in usage per year. No provision made for changes in demand caused by new technologies, discoveries of new reserves or market forces. Agricultural land means land suitable for rainfed cultivation net of other land usage. Thirty year historic agricultural expansion rates are applied.



## Bad vs. Better

#### Bad

#### **Top Languages Used on the Internet**



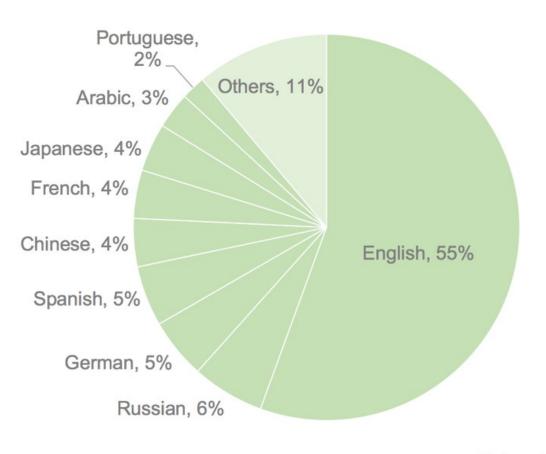
Data from Wikipedia Kaiser Fung | Junk Charts

## Bad vs. Better

#### Better

- Follows common conventions:
  - Places the first slice at the top vertical
  - Sorts slices from largest to smallest
    - (Except "Others")
  - Introduces multiple colors only to encode data differences.

#### **Top Languages Used on the Internet**

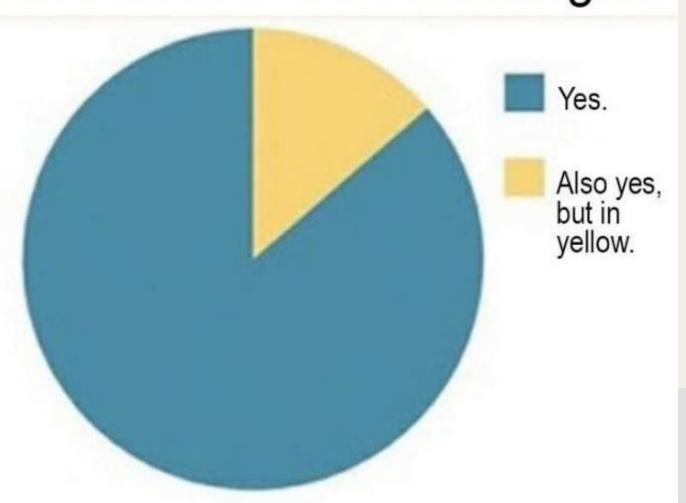


Data from Wikipedia

Kaiser Fung | Junk Charts

## Pie Charts

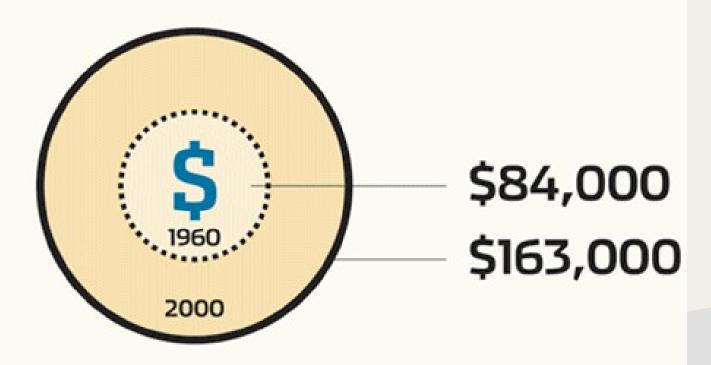
### Should I have a beer tonight?



## Pie Charts (cont.)

#### **MEDIAN FAMILY INCOME**

In the 14 most elite ZIP Codes (as of 1960, in today's purchasing power)



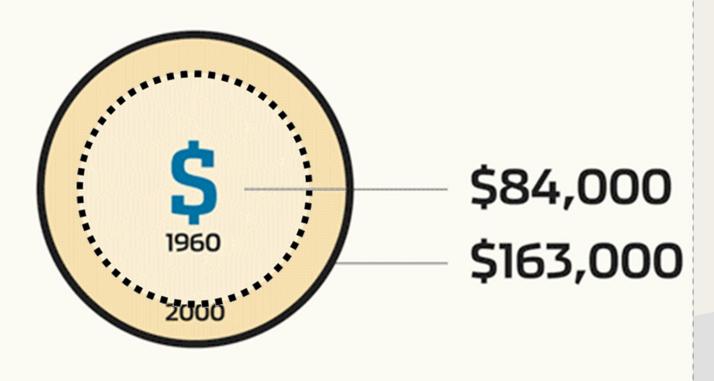
SOURCE: 'COMING APART'

22

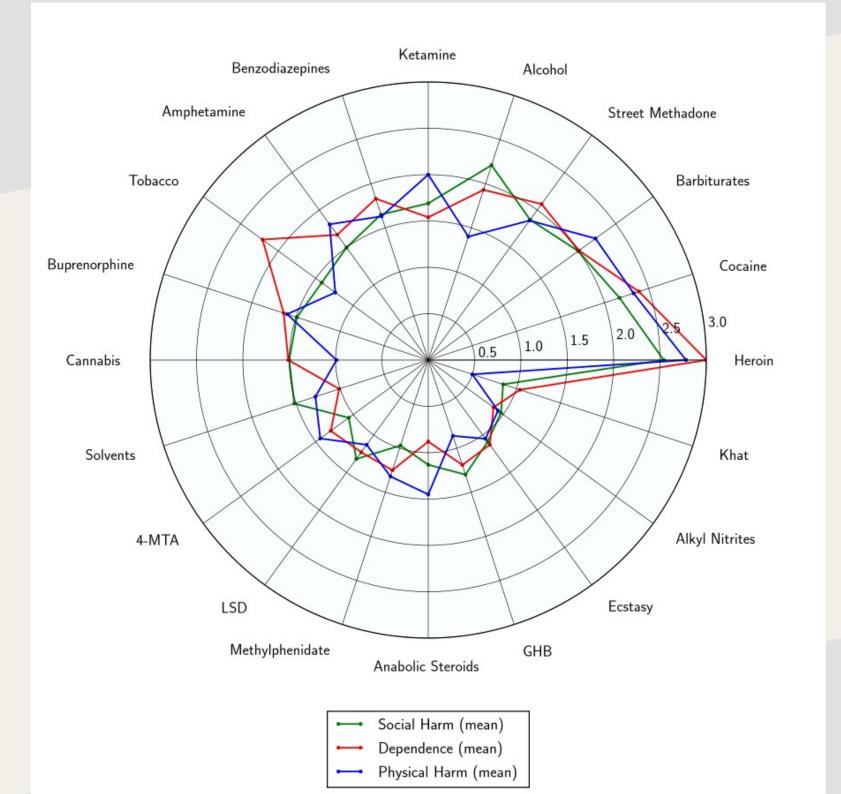
## Pie Charts (cont.)

#### **MEDIAN FAMILY INCOME**

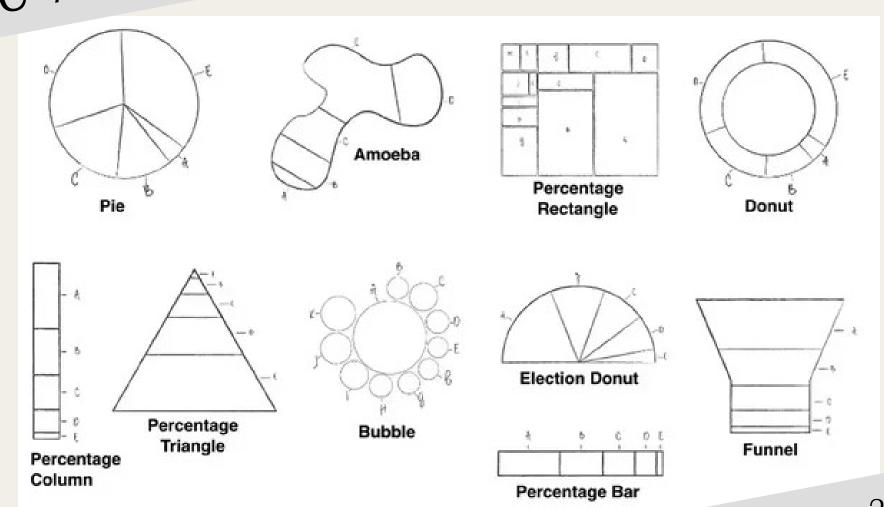
In the 14 most elite ZIP Codes (as of 1960, in today's purchasing power)



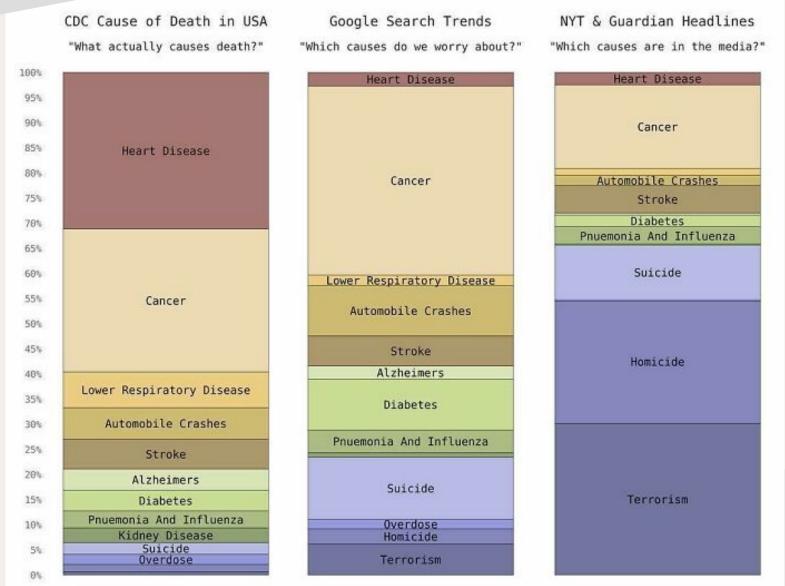
SOURCE: 'COMING APART'



## Pie Charts: Some Alternatives



### Pie Charts: Some Alternatives (cont.)

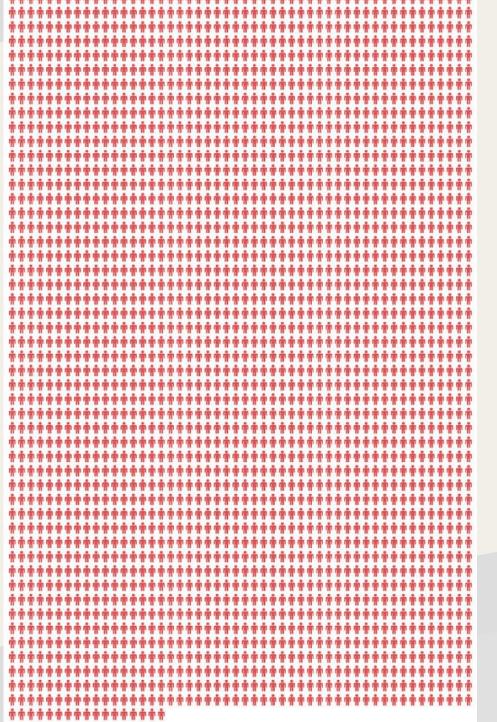


### Pie Charts: Some Alternatives

People killed in terrorist attacks <u>2015 - 2020</u>

**£** Europe

**?** Africa

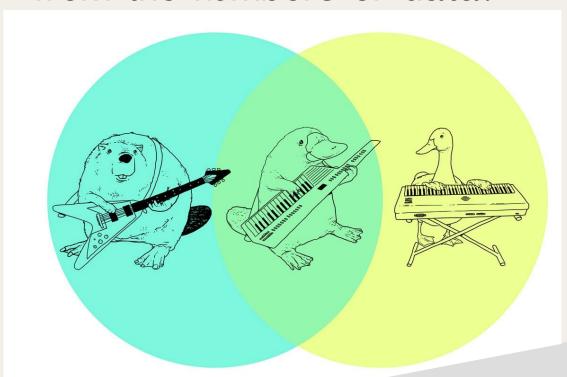


### Elements of Good Data Visualization

- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

### Self Sufficiency

- Can the figure be understood on its own
  - E.g., can the figure be understood divorced from the numbers or data?

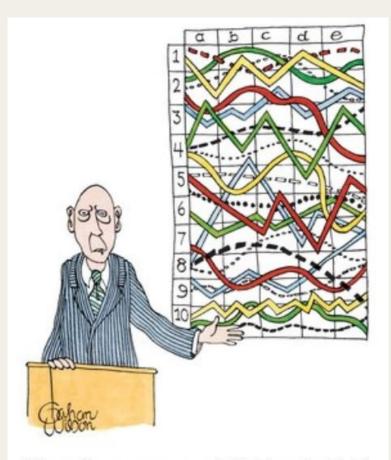


### Elements of Good Data Visualization

- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

## Efficient information transfer

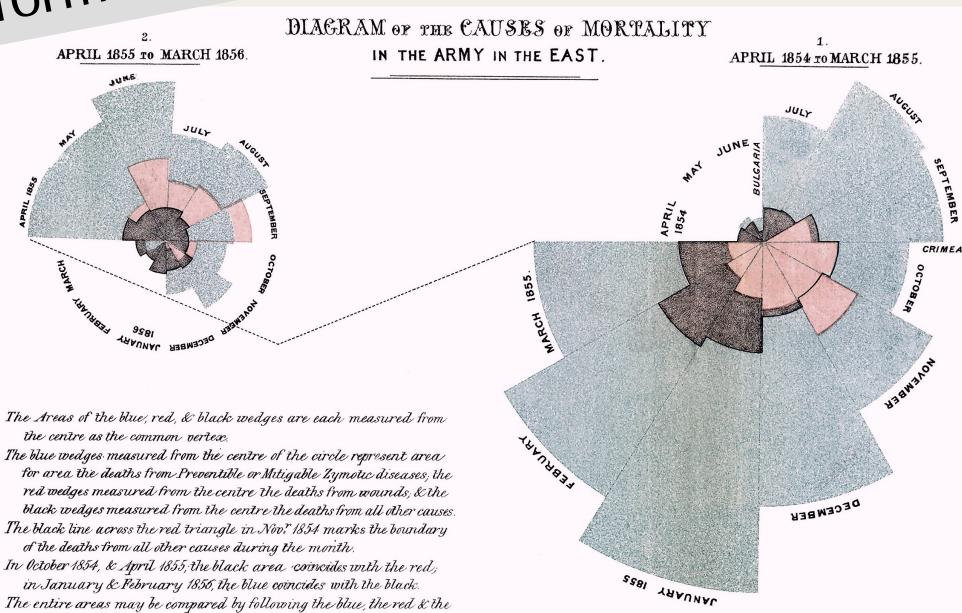
- How much information is conveyed
- How quickly & easily is it to understand
- N.b., many argue to convey a few, important pieces of information in a given figure



"I'll pause for a moment so you can let this information sink in."

# Efficient information transfer (cont.)

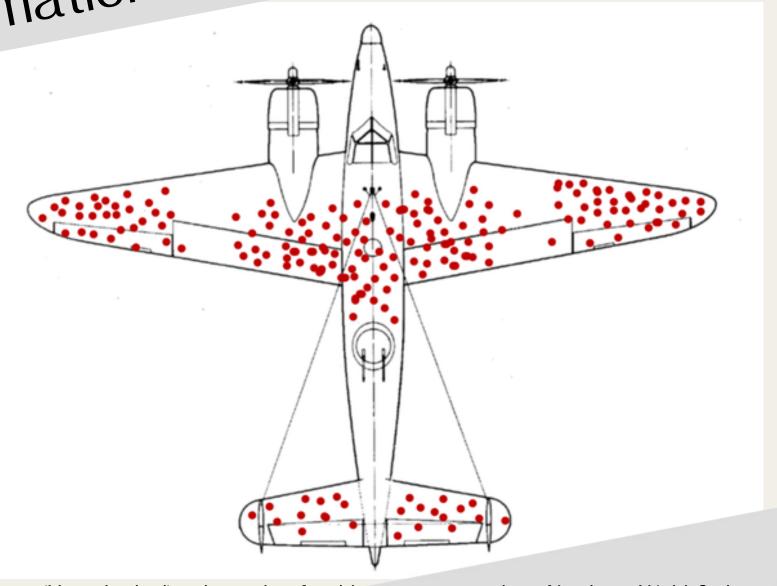
black lines enclosing them.



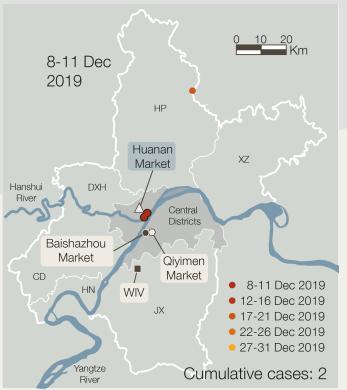


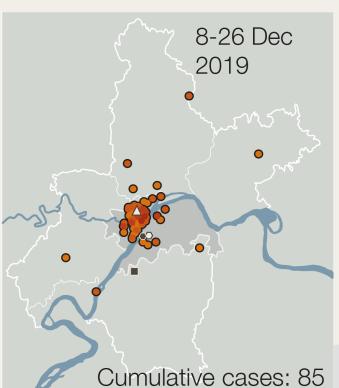
Snow, John (1849). On the Mode of Communication of Cholera. London: John Churchill.

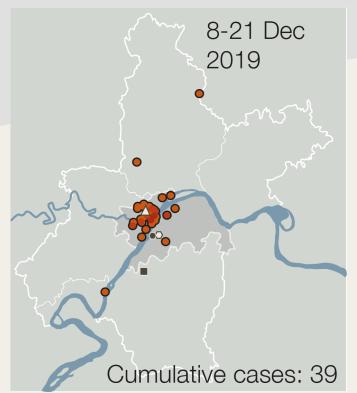
# Efficient information transfer (cont.)

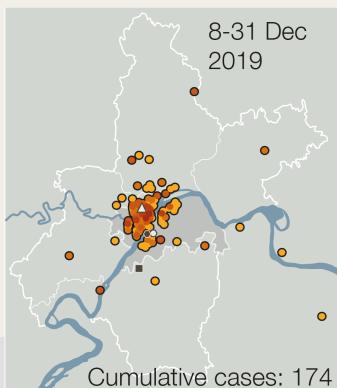


(Hypothetical) schematic of evidence presented to <u>Abraham Wald</u> & the rest of the Statistical Research Group, WWII







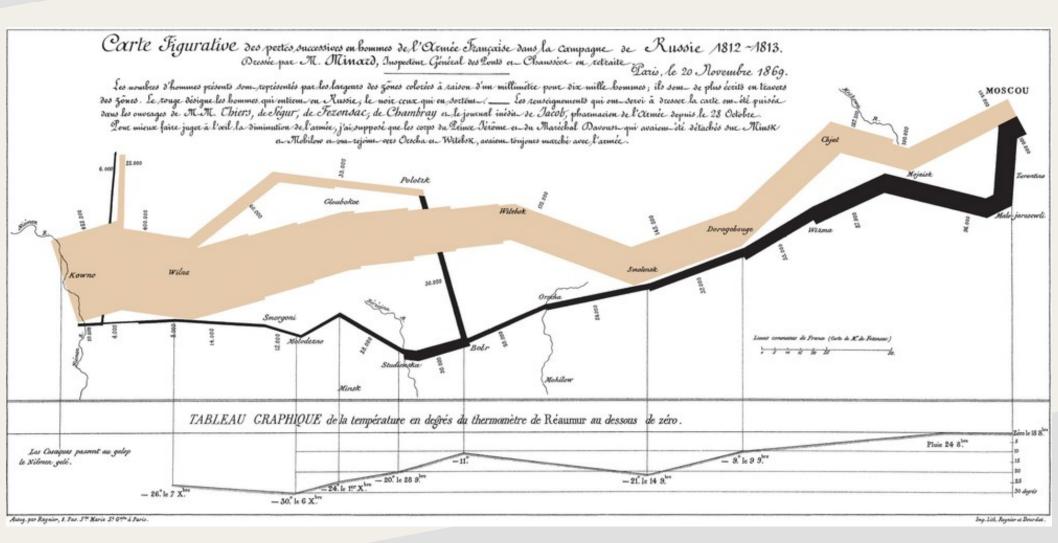


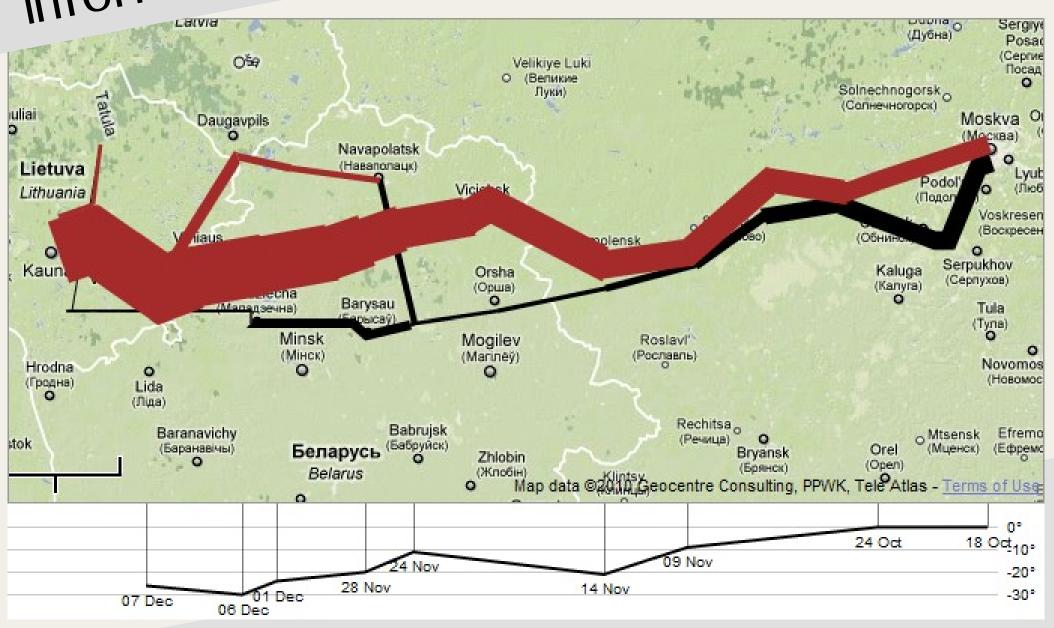
Huanan Market: A "wet" market where wild animals were illegally sold.

**WIV:** Wuhan Institute of Virology

Cited in a <u>Goats & Soda</u> article from: Holmes, E. C et al. (<u>2021, July 7</u>). The Origins of SARS-CoV-2: A Critical Review (Version 1.0.2). *Zenodo*. http://doi.org/10.5281/zenodo.5112546

# Efficient information transfer (cont.)

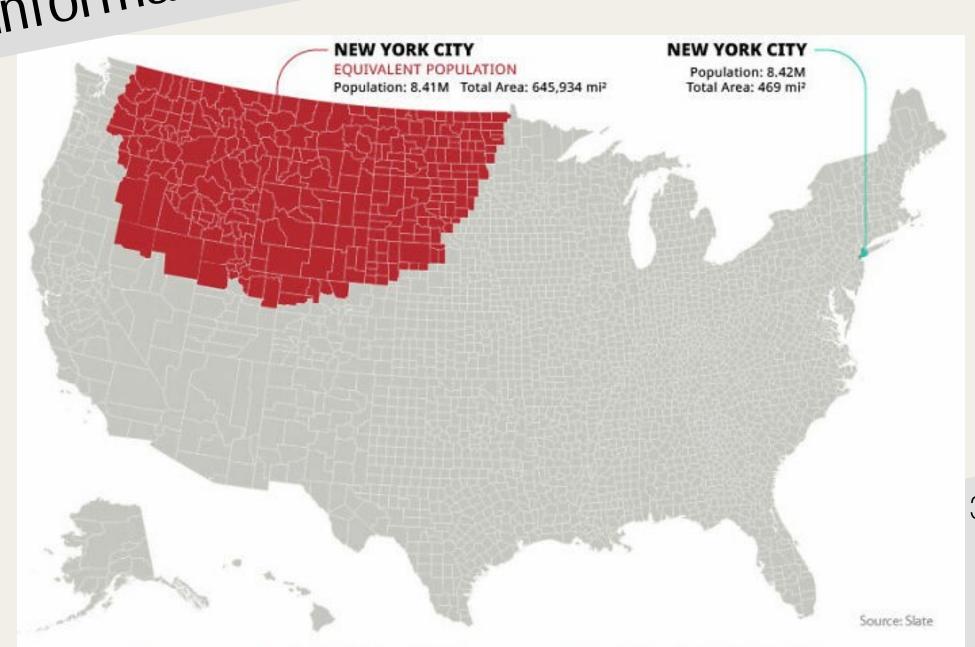




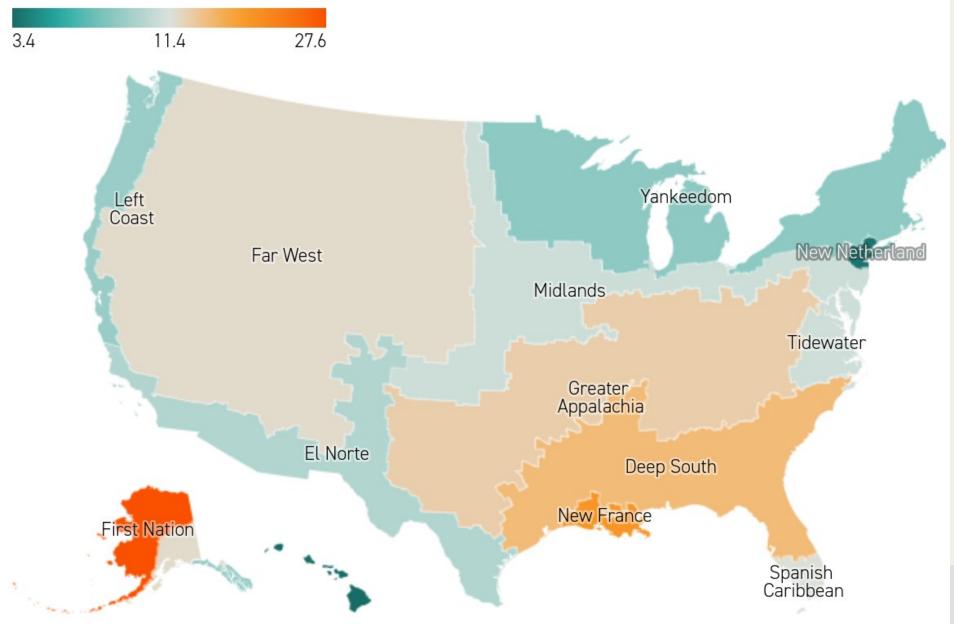
#### POPULATION OF THE UNITED STATES



50% LIVE HERE 50% LIVE HERE



Average of annual gun deaths per 100,000 people in each region from 2010-2020 Nationwide rate: 11.4 gun deaths per 100,000 people

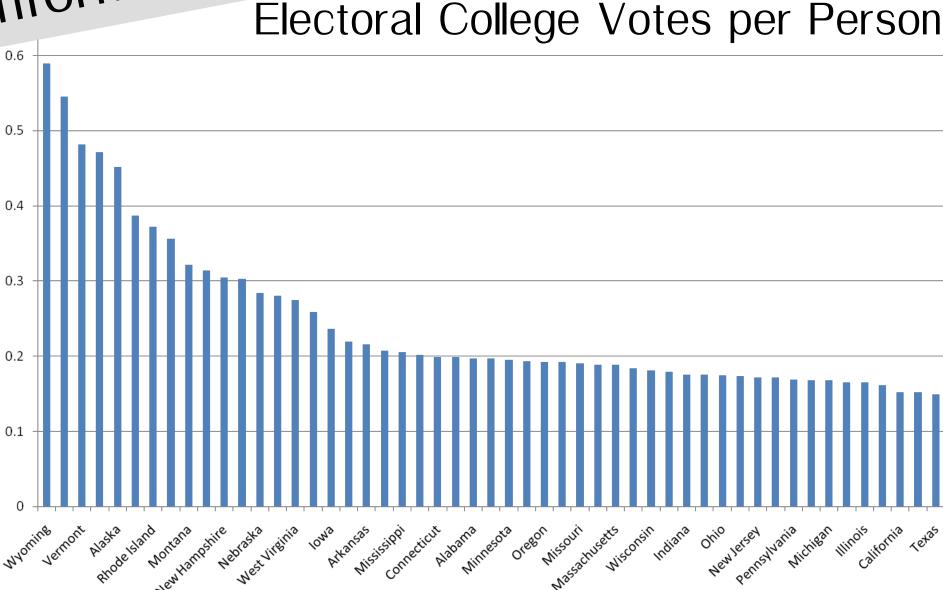


Notes: First Nation, Greater Polynesia, New France and Spanish Caribbean are not included among major regions. Rates are smoothed across counties due to CDC data suppression.

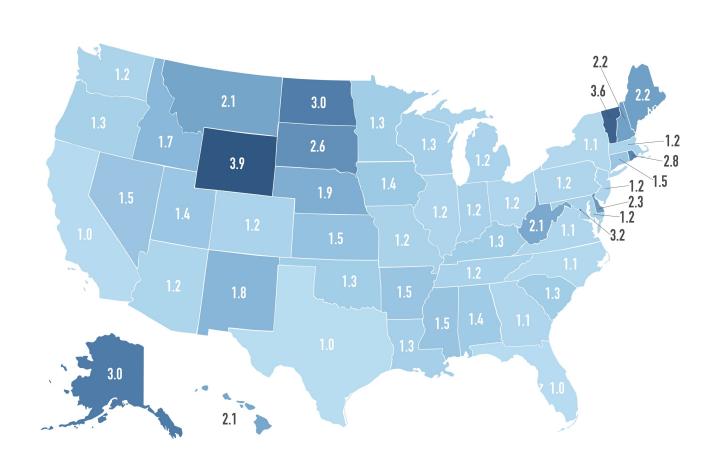
Source: Nationhood Lab at Salve Regina University

Grant Schwab/POLITICO

#### Efficient transfer (cont.) information transfer (cont.)

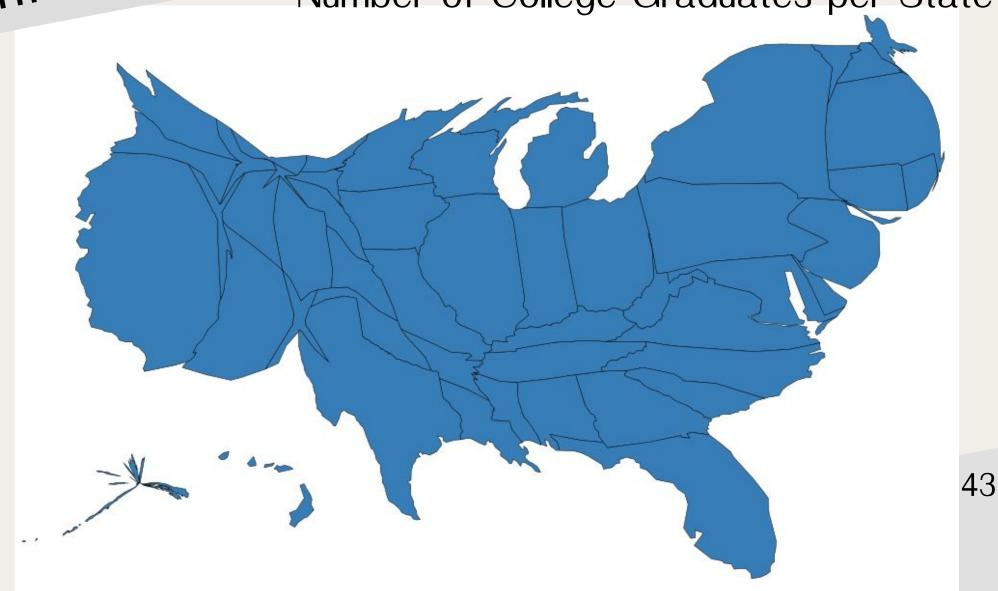


#### Votes per Person in US Electoral College System

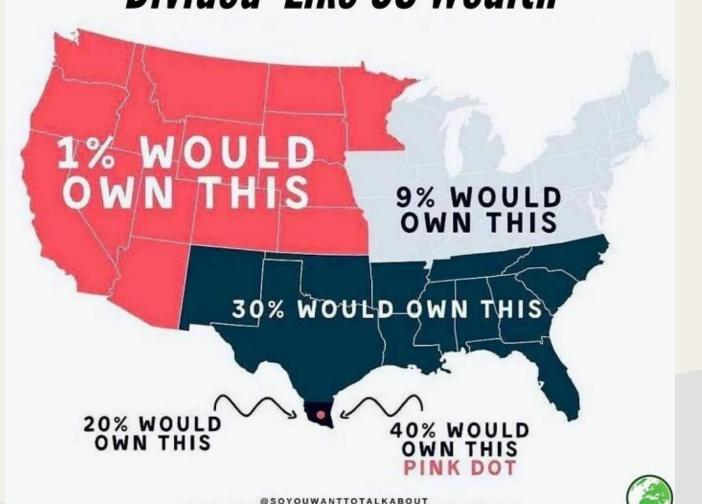


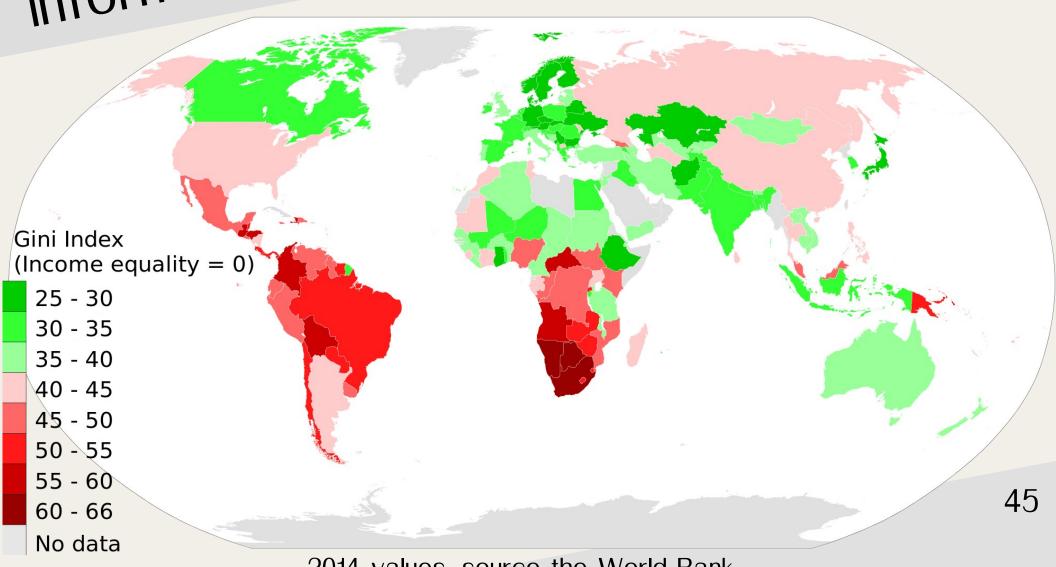
= 4,000 Graduates

Number of College Graduates per State

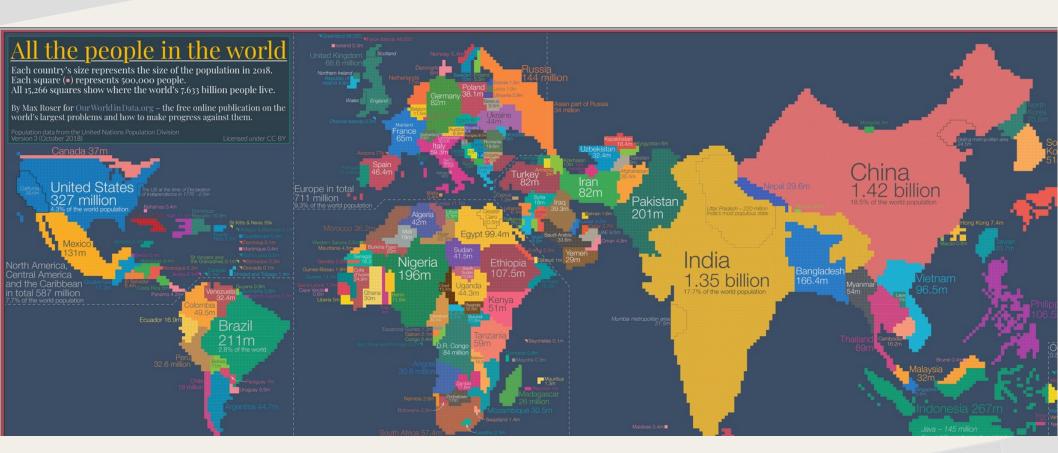


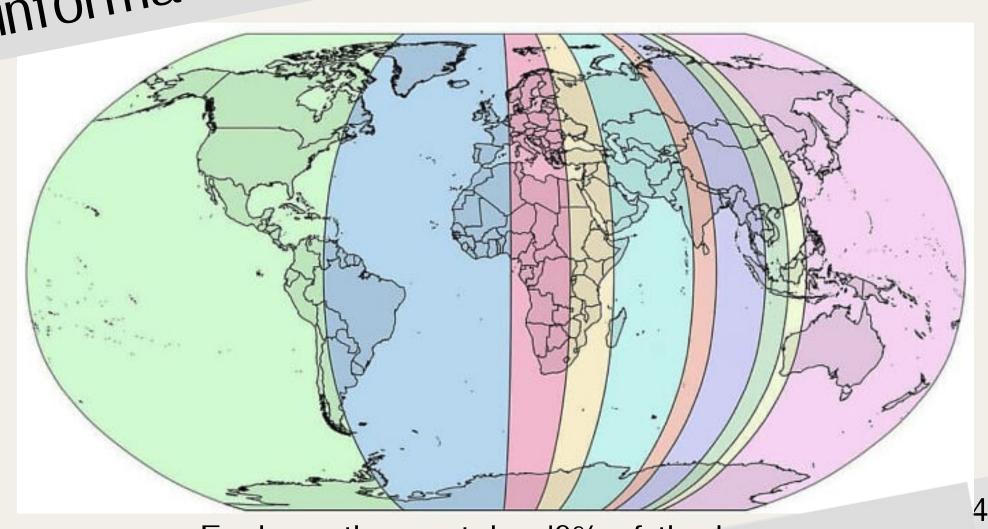
#### If The United States Was Divided Like US Wealth





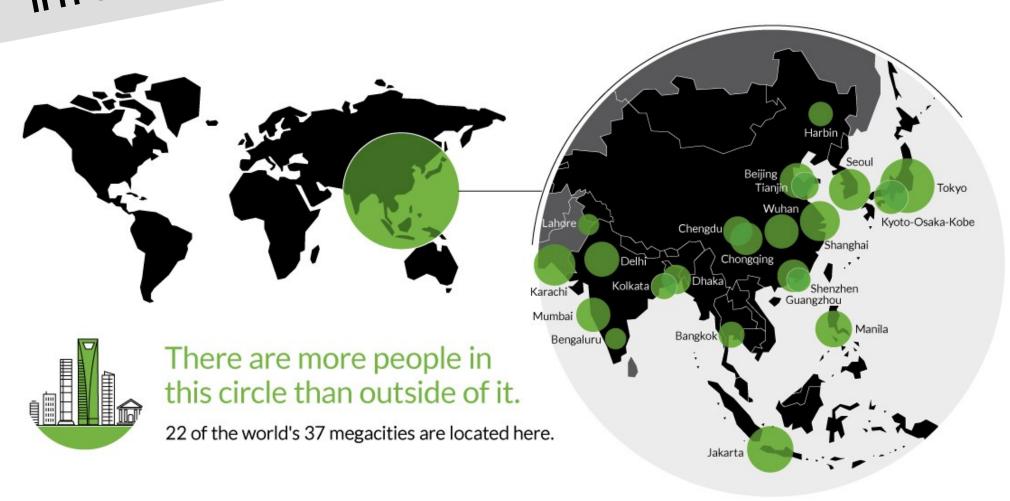
2014 values, source the World Bank

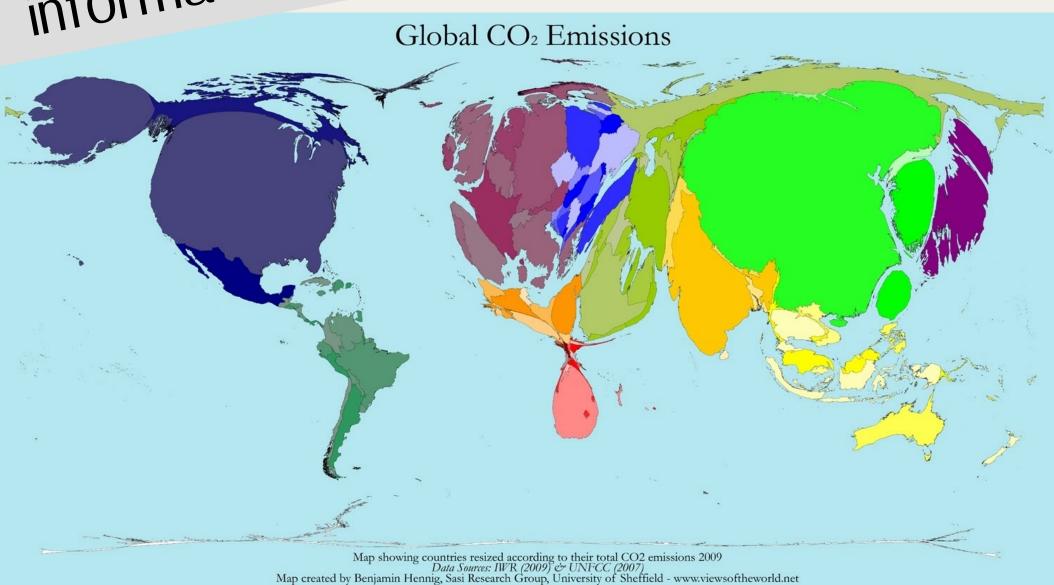




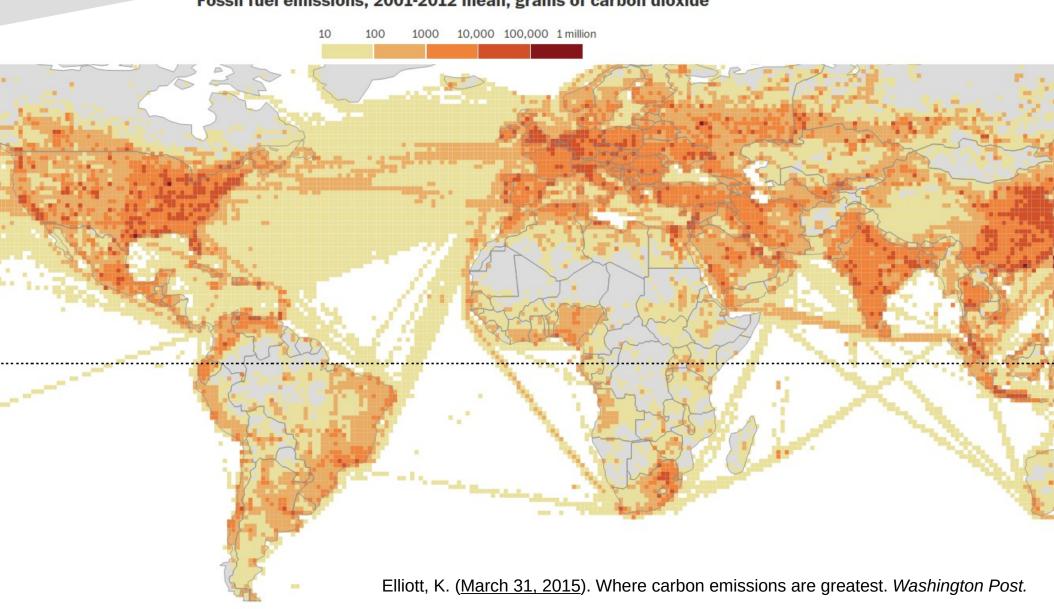
Each section contains 10% of the human population (c. 2020)

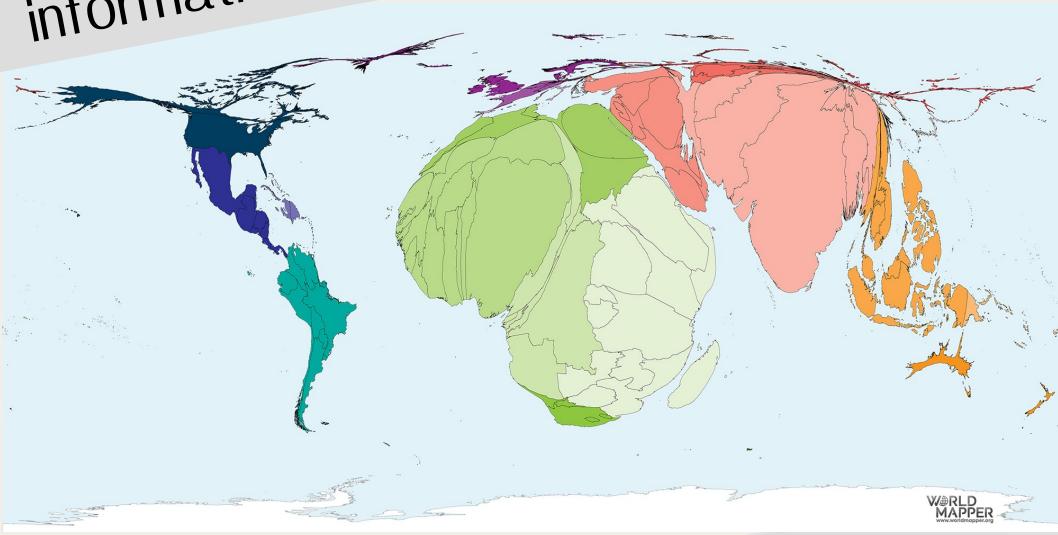
47



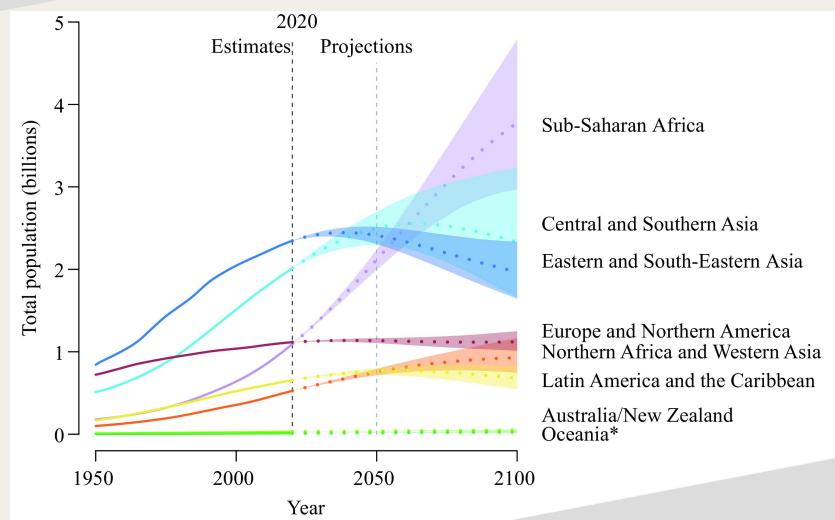


Fossil fuel emissions, 2001-2012 mean, grams of carbon dioxide



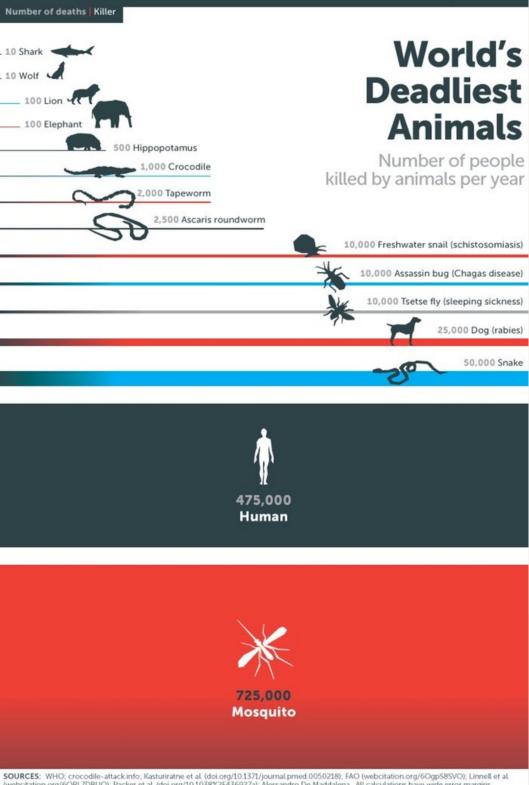


World population growth rates 2015 - 2050

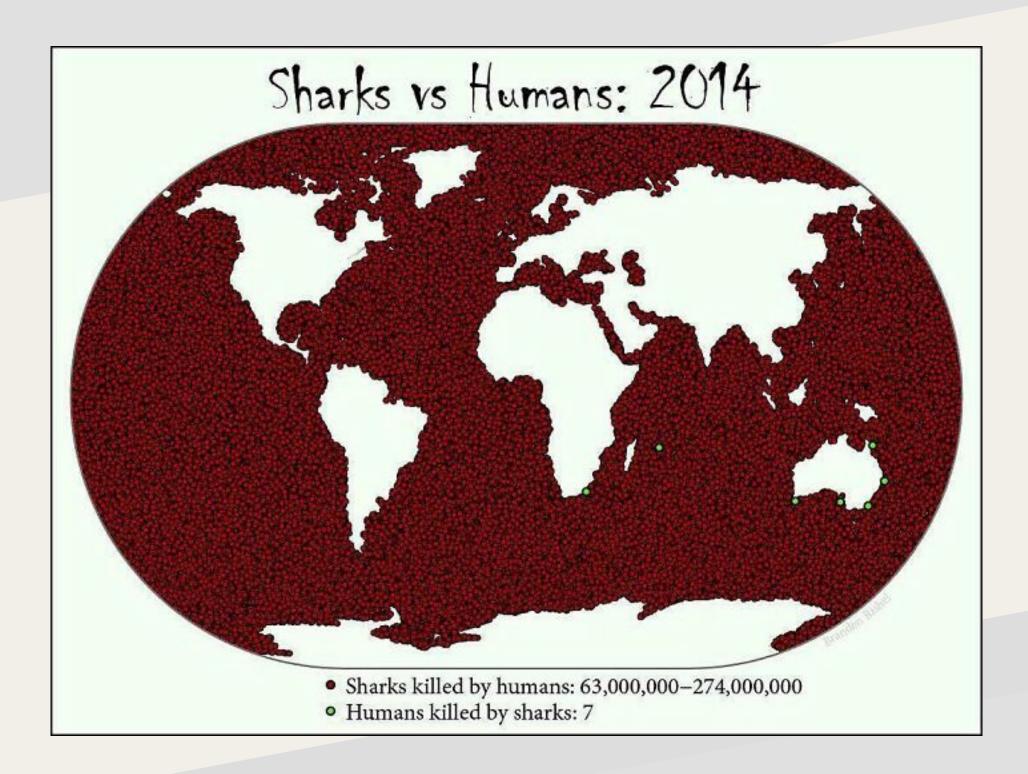


Source: Gu, D., Andreev, K., & Dupre, M. E. (2021). Major trends in population growth around the world. *China CDC weekly*, *3*(28), 604–613. <a href="https://doi.org/10.46234/ccdcw2021.160">https://doi.org/10.46234/ccdcw2021.160</a>

# Efficient information transfer

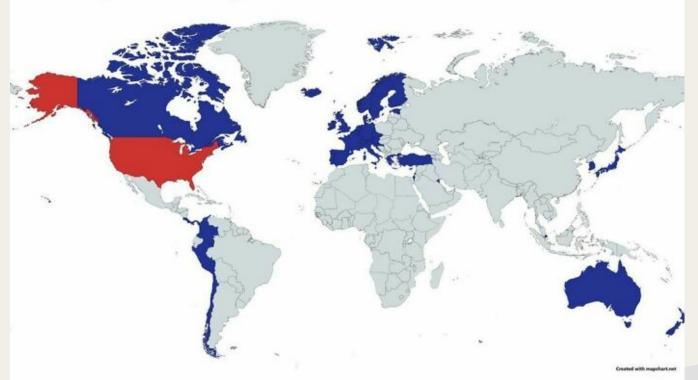


(webcitation.org/6ORL7DBUO); Packer et al. (doi.org/10.1038%2F436927a); Alessandro De Maddalena. All calculations have wide error margins.



## Efficient transfer (cont.) information transfer (with a little with a li

#### Countries With a Higher Life Expectancy than the United States



United States

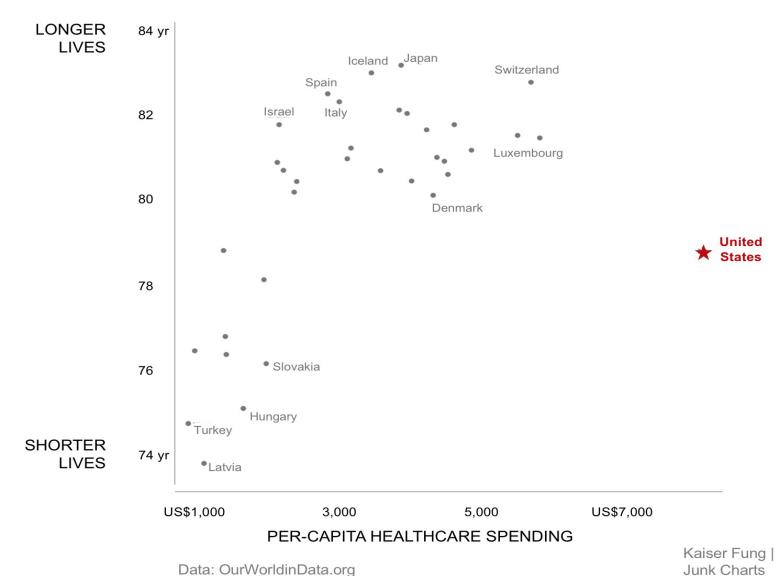
Countries With A Higher Life Expectancy than The United States (WHO 2019)



55

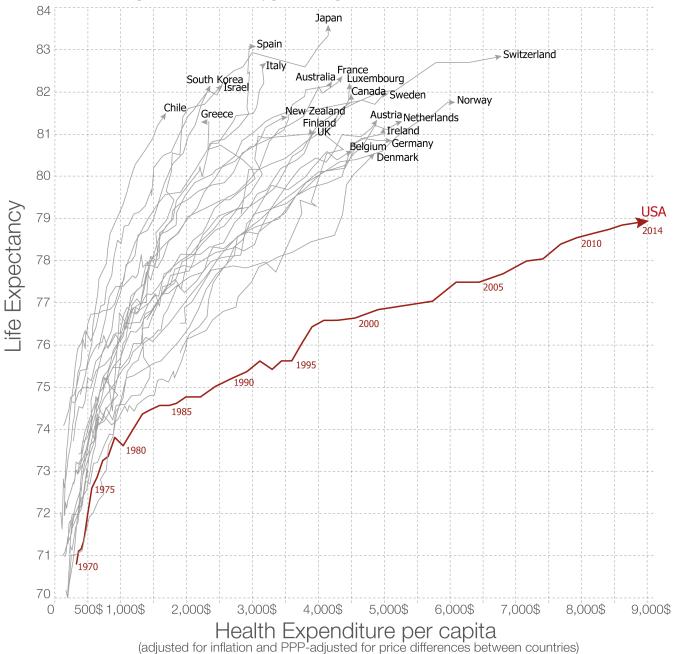
## Efficient transfer (cont.) information transfer (cont.)

The United States spent at least 50% more per capita on health care, yet its life expectancy in 2012 lagged most developed nations.



#### Life expectancy vs. health expenditure over time (1970-2014) Our World

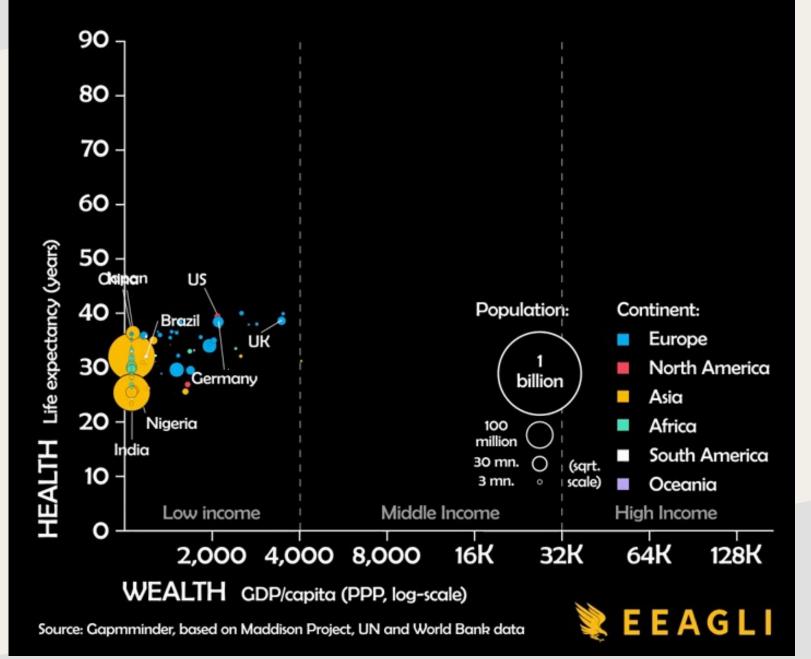
in Data Health spending measures the consumption of health care goods and services, including personal health care (curative care, rehabilitative care, long-term care, ancillary services and medical goods) and collective services (prevention and public health services as well as health administration), but excluding spending on investments. Shown is total health expenditure (financed by public and private sources).



Data source: Health expenditure from the OECD; Life expectancy from the World Bank Licensed under CC-BY-SA by the author Max Roser. The data visualization is available at OurWorldinData.org and there you find more research and visualizations on this topic.

#### 221 Years of Health and Wealth

Year: 1800



#### **Deaths Due to Complications from Abortions in the USA**

Since the passage of Roe v Wade, in the early 70s, there has been a dramatic drop in abortion deaths.

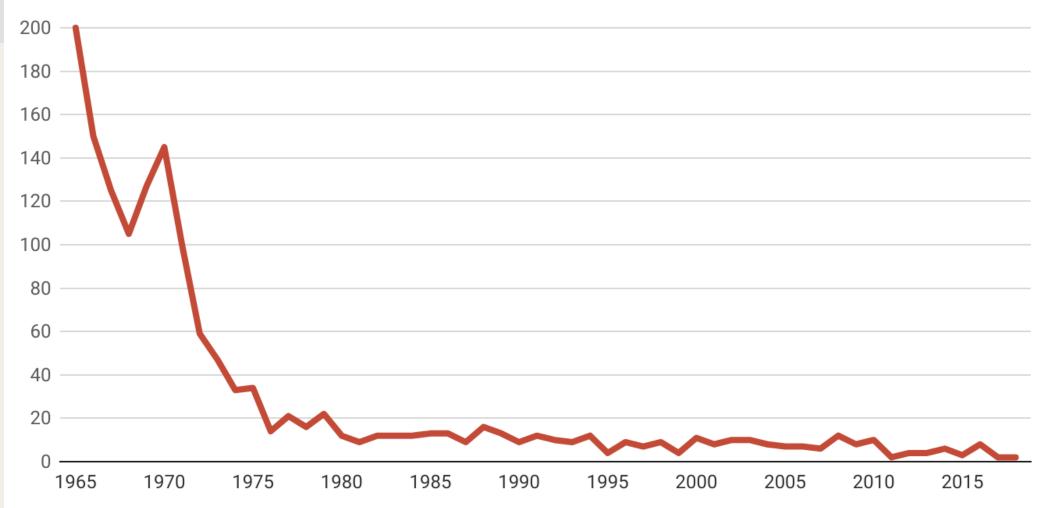
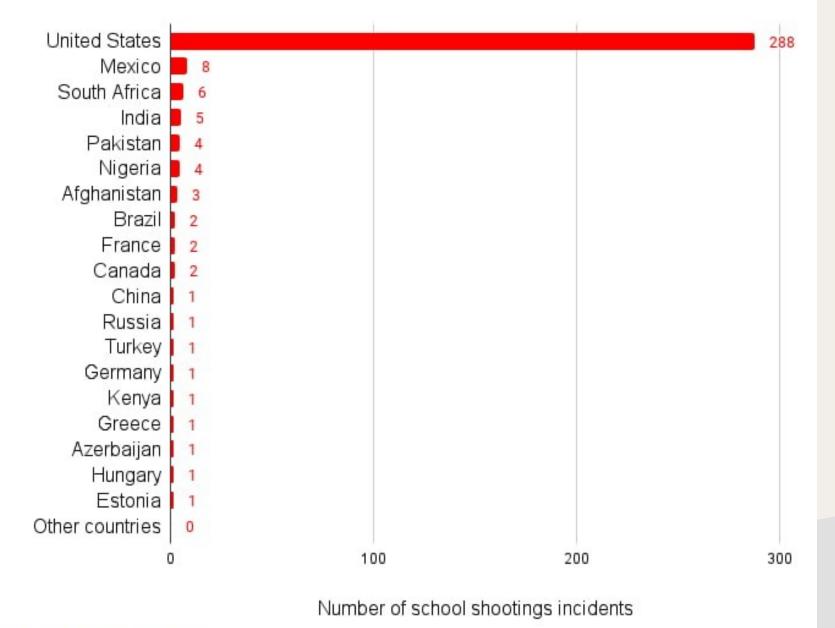


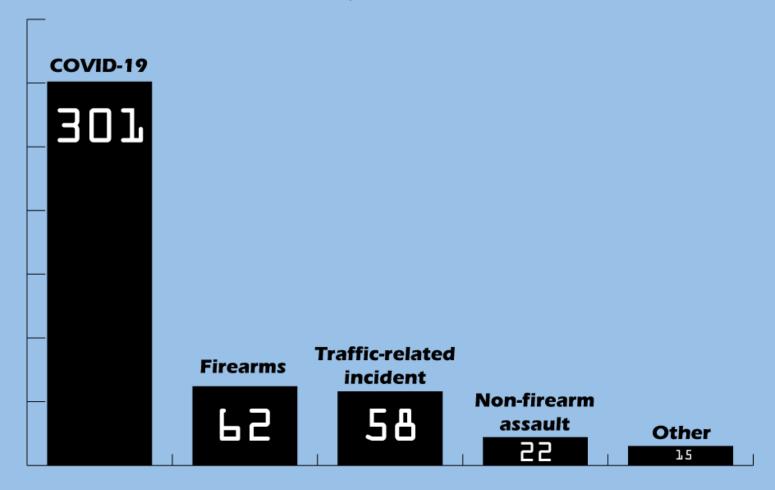
Chart: u/thedataracer • Created with Datawrapper

#### Countries with school shootings incidents (Jan 2009-May 2018)

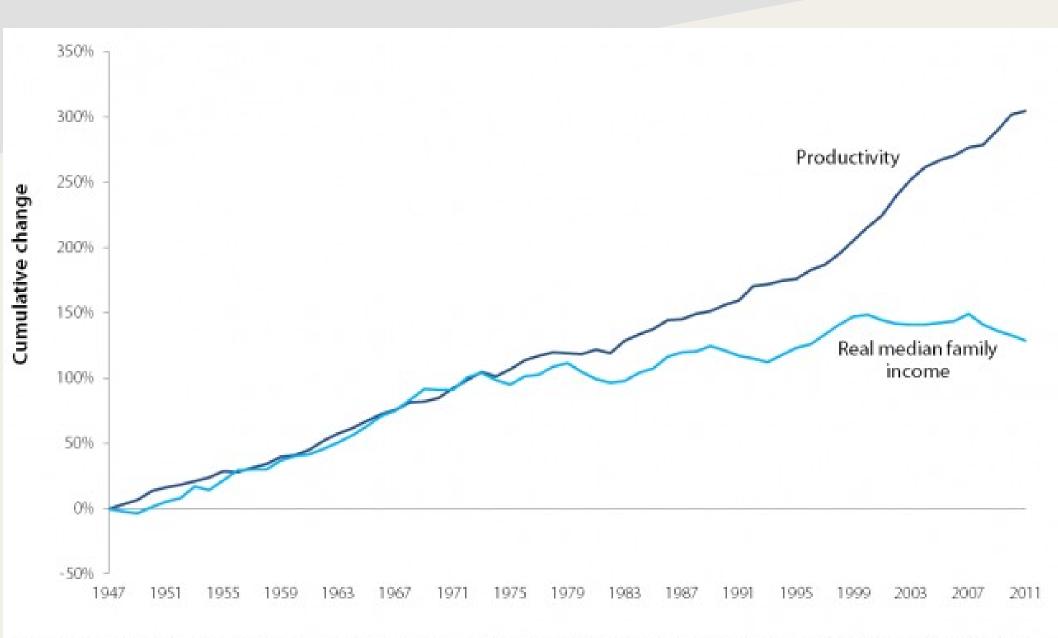
Source: World Population Review



#### 458 law enforcement officers died in the line of duty in 2021



Created by TrueBirch using data from the National Law Enforcement Officers Memorial Fund. COVID-19 fatalities are only counted when an officer "died due to direct exposure to the virus during the commission of their official duties." Officers who died from a case of COVID-19 that did not have a proven link to "official duties" are not counted.



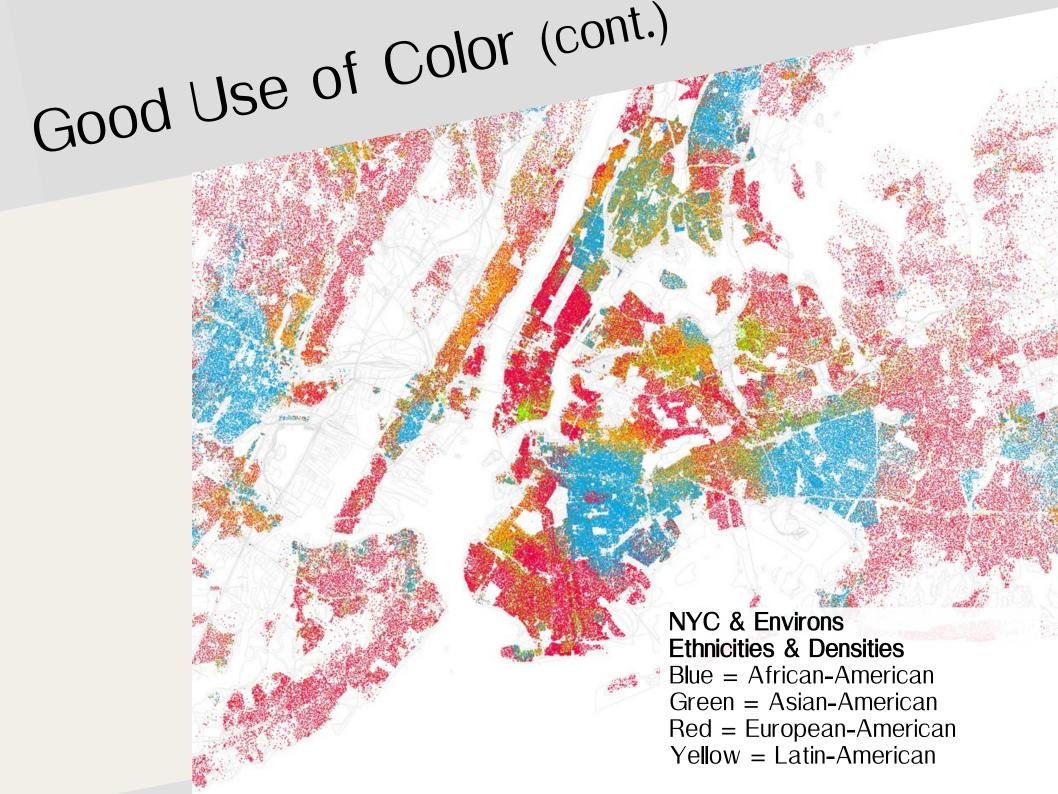
Source: Authors' analysis of Current Population Survey Annual Social and Economic Supplement Historical Income Tables, (Table F-5) and Bureau of Labor Statistics, Productivity – Major Sector Productivity and Costs Database (2012)

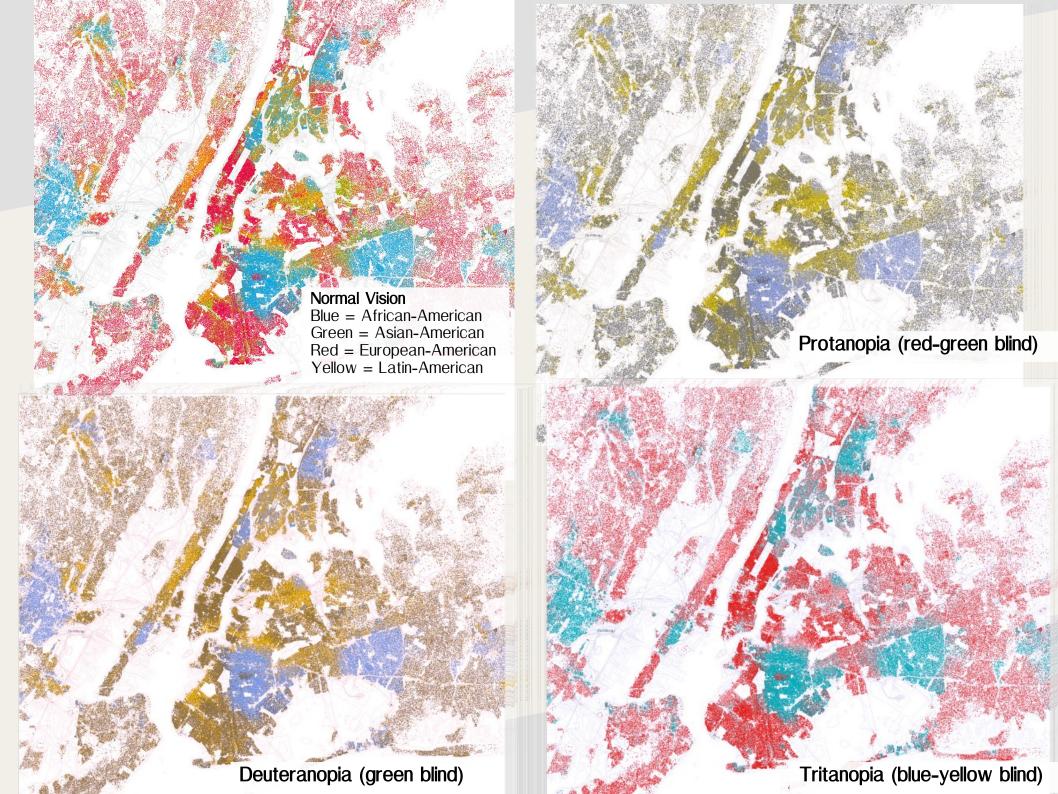
#### Good Use of Color

- Use a few contrasting colors
  - Black & grey scales are safe & easily reproduced
  - Good combinations to accommodate color blindness:
    - Indigo & orange (+ green + vermillion)
    - Deep purple & mint (+ orange + teal)
- Sites to get examples & check accessibility:
  - J\*Fly Site about Color Blindness & Figures
  - Coblis: Simulates how images appear to different users

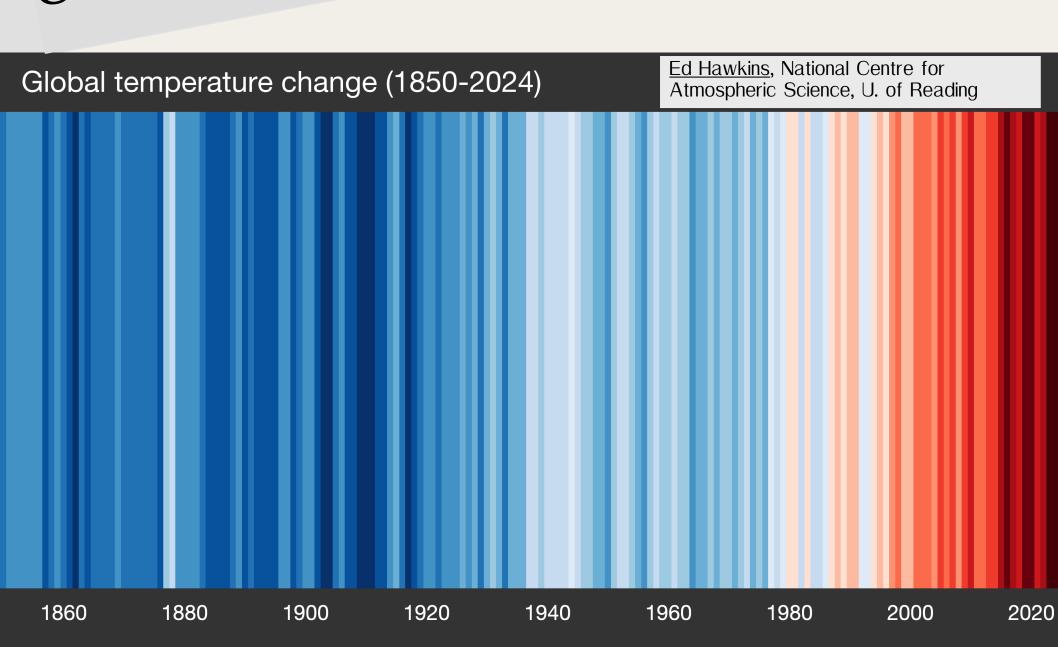
63

- Colorgorical: Generate perceptually distinct images
- ColorBrewer2: Has a "colorblind-safe" filter





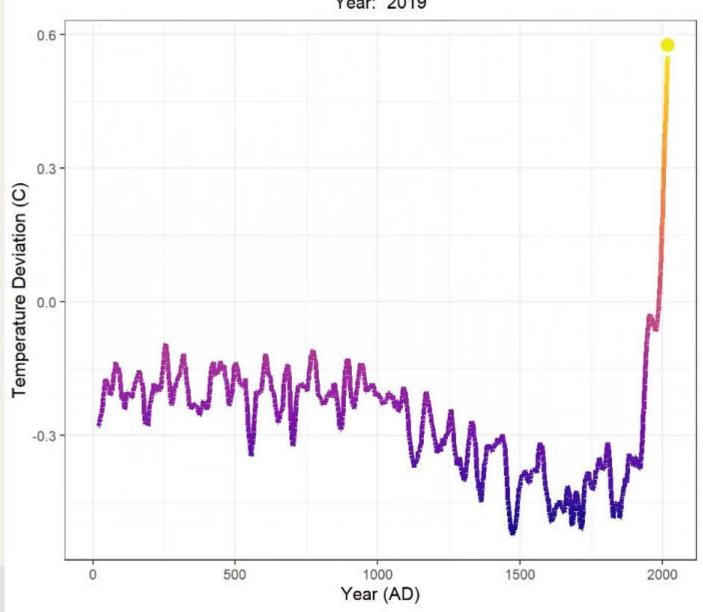
#### Good Use of Color (cont.)



#### Good Use of Color (cont.)

#### Global Temperature Trend from 0 to 2019 AD

Temperature Deviation (C) from 1961-1990 Average Year: 2019



#### Good Use of Color (cont.)



#### Good Use of Color (end)

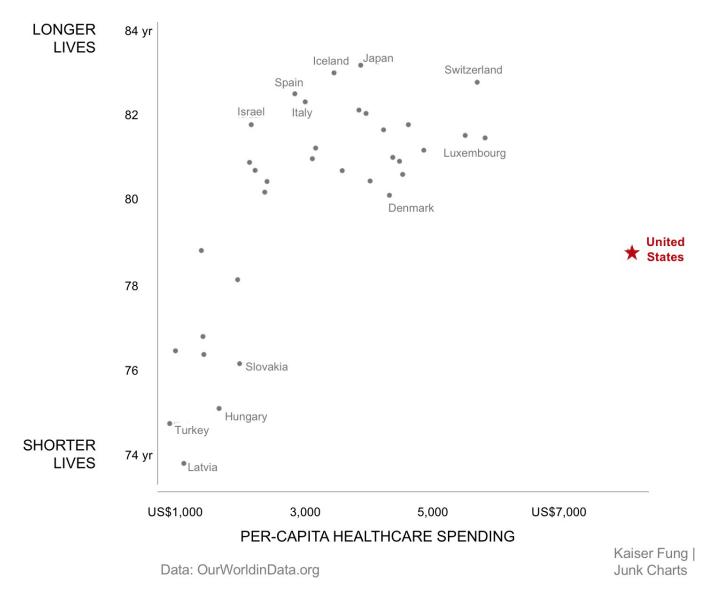


#### Elements of Good Data Visualization

- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

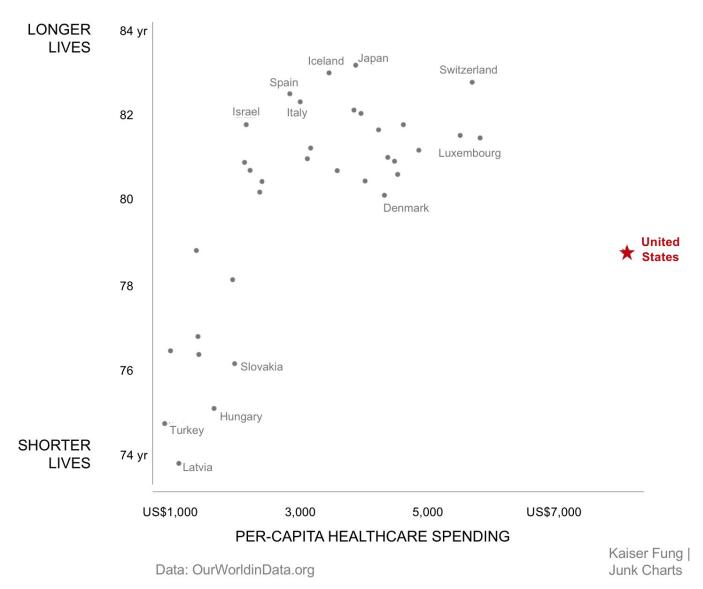
## Efficient transfer (cont.) information transfer (cont.)

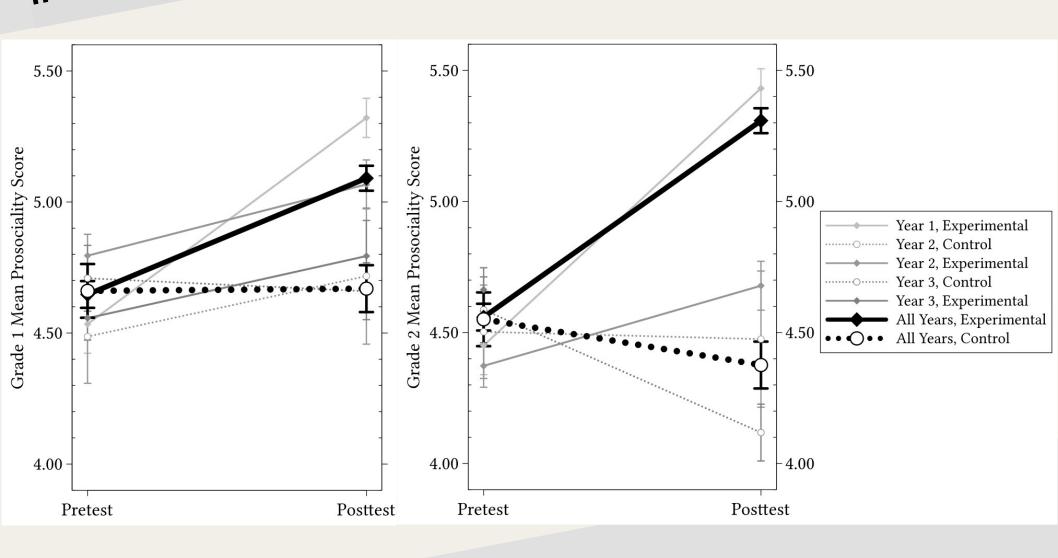
The United States spent at least 50% more per capita on health care, yet its life expectancy in 2012 lagged most developed nations.



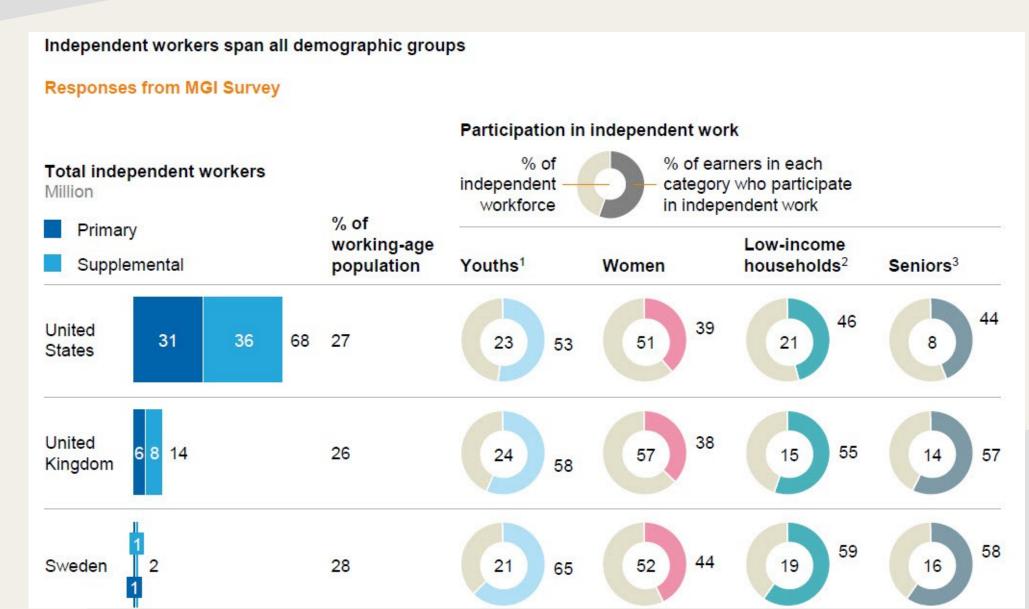
# Efficient transfer (cont.) information transfer

The United States spent at least 50% more per capita on health care, yet its life expectancy in 2012 lagged most developed nations.

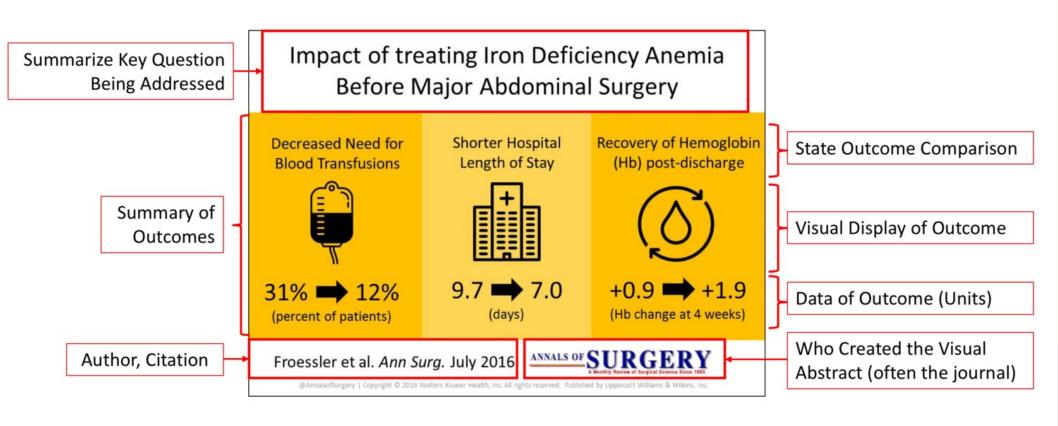




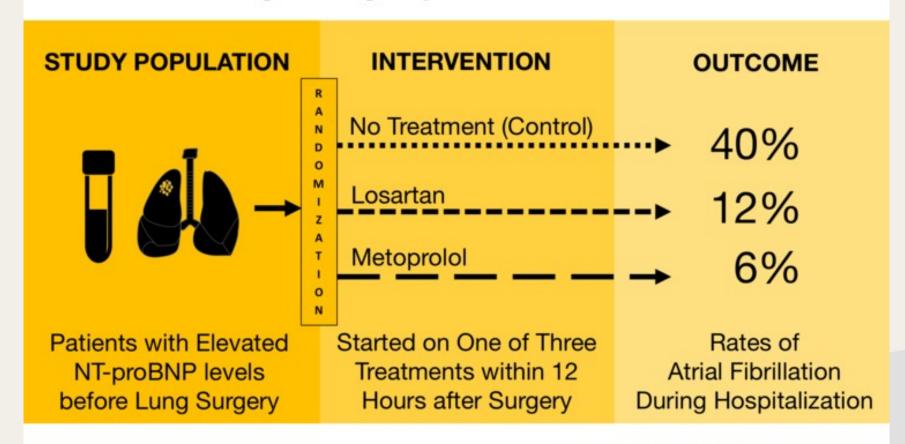
But too much is too much



#### **COMPONENTS OF AN EFFECTIVE VISUAL ABSTRACT**



Reducing the Rate of Atrial Fibrillation after Lung Surgery: the PRESAGE Trial



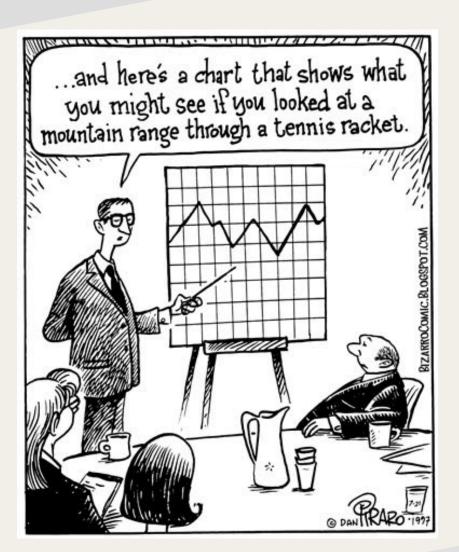
Cardinale et al. Ann Surg. Aug 2016.



## Elements of Good Data Visualization

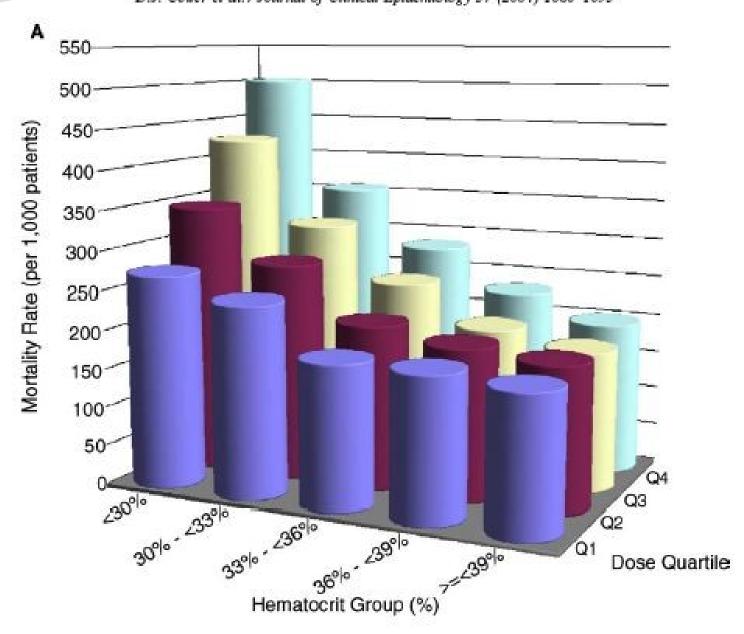
- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

## Data-to-Ink Ratio



## Data-to-Ink Ratio

D.J. Cotter et al. / Journal of Clinical Epidemiology 57 (2004) 1086-1095



## Elements of Good Data Visualization

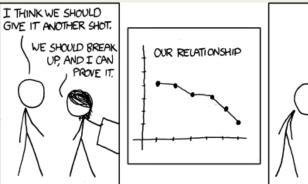
- Self sufficiency
- Efficient information transfer
- Data-to-ink ratio
- Follow conventions & readers' expectations

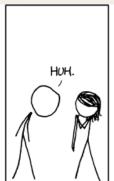
## Follow conventions & readers' expectations

- Label axes, give titles, consider a legend
  - Use title as the figure's message
  - Consider labeling elements within the figure
- Time progresses from left to right
  - Or clockwise

Changes in color track changes in

conditions

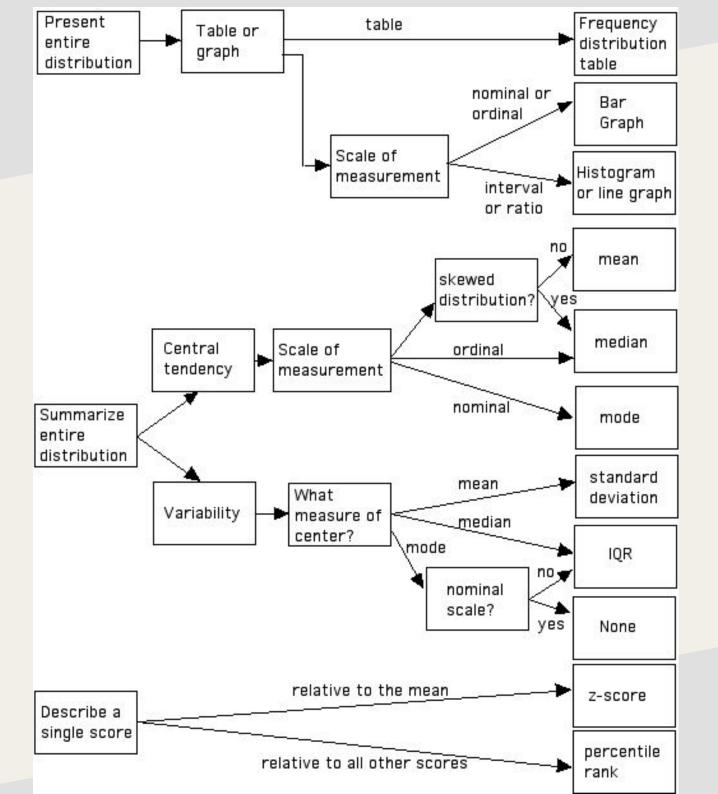


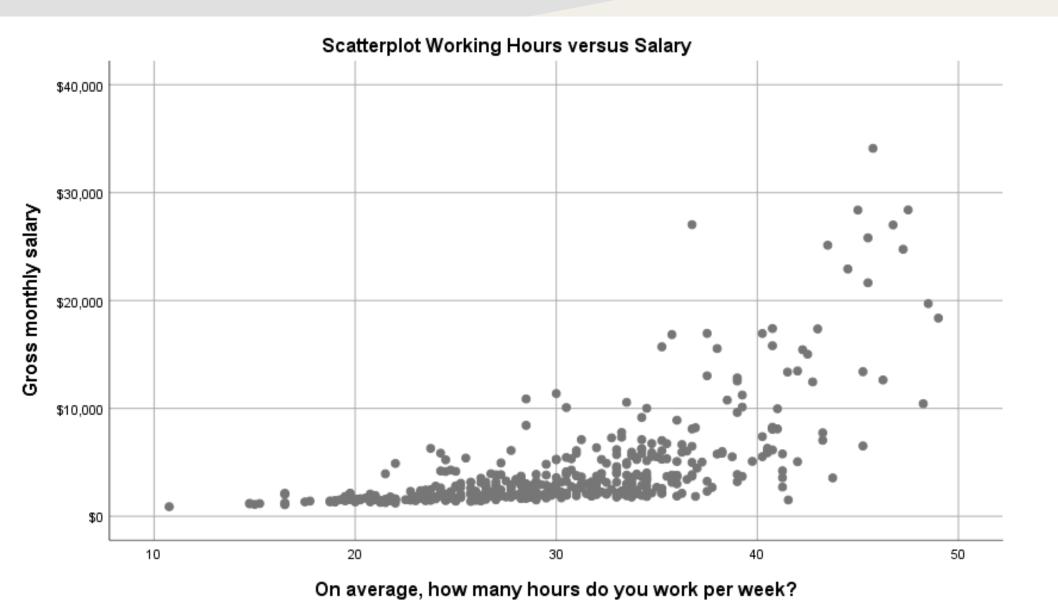


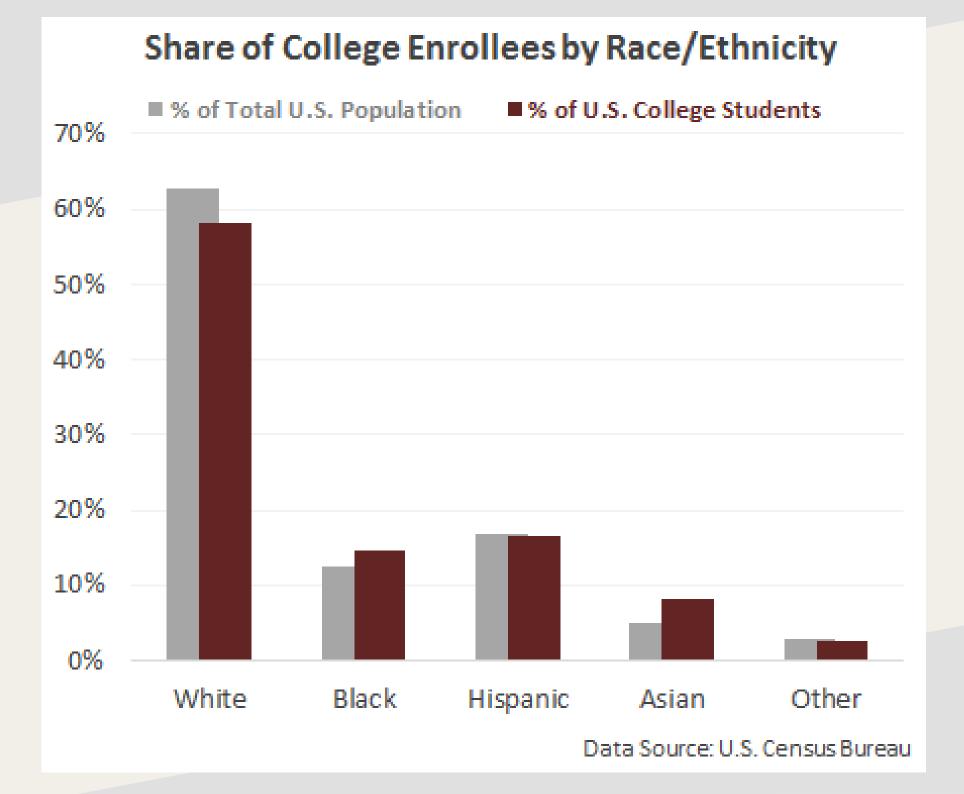


## Follow conventions 8 readers' expectations (cont.)

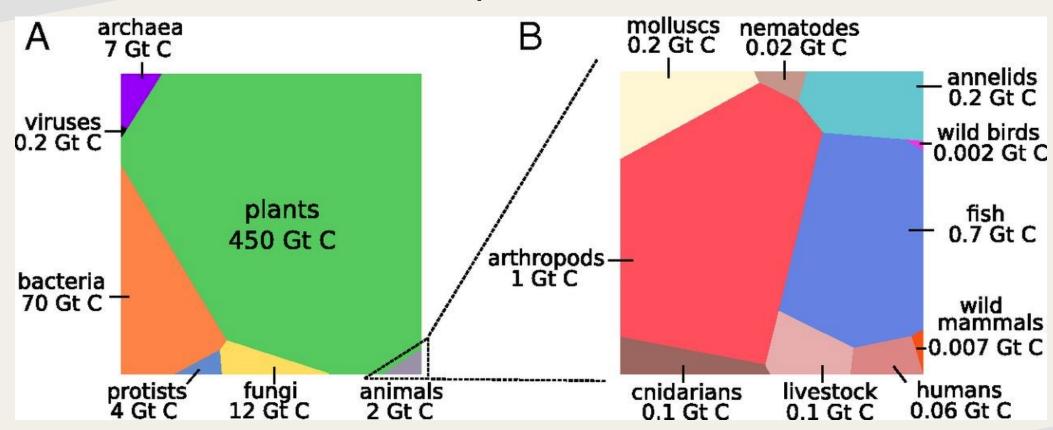
- Use same scale for all parts of a figure
  - And between figures if possible
- Area of objects scale magnitude
  - But humans are bad at comparing areas
  - Lengths are much easier
- x-Axis on histograms & bar charts is zero
  - Or clearly marked as not



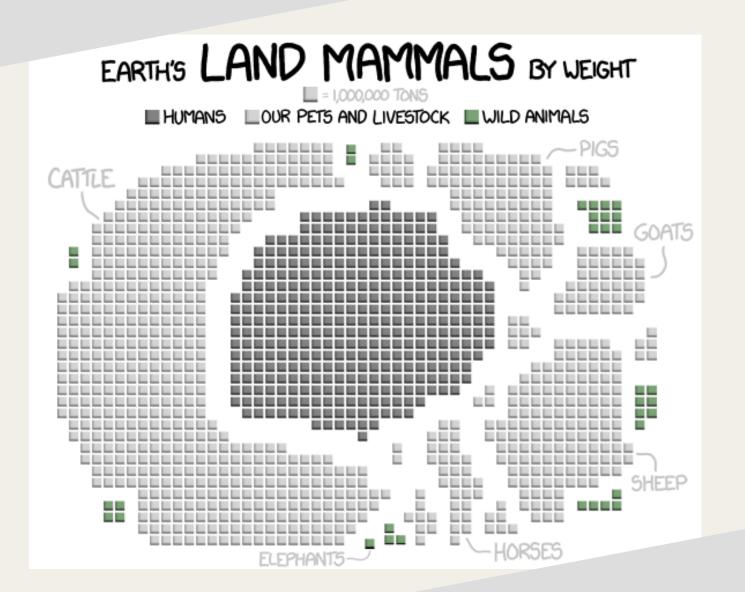


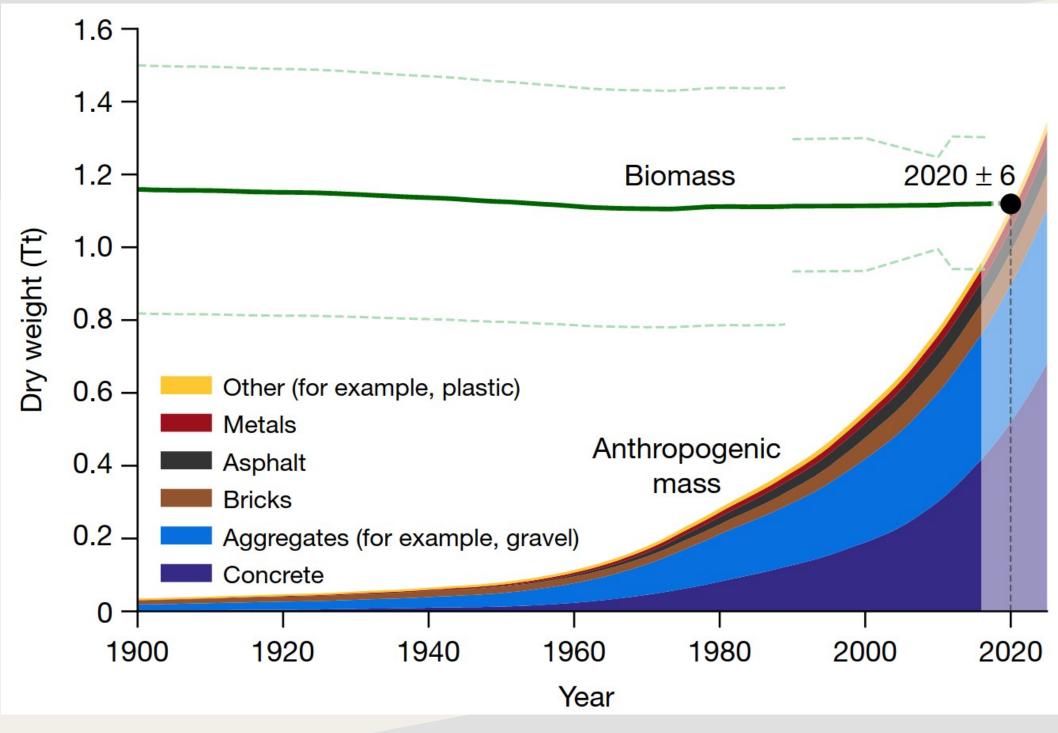


## Proportions of Earth's Biomass



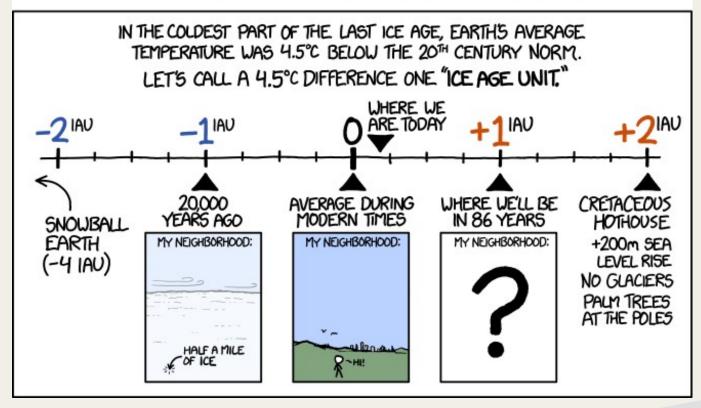
88





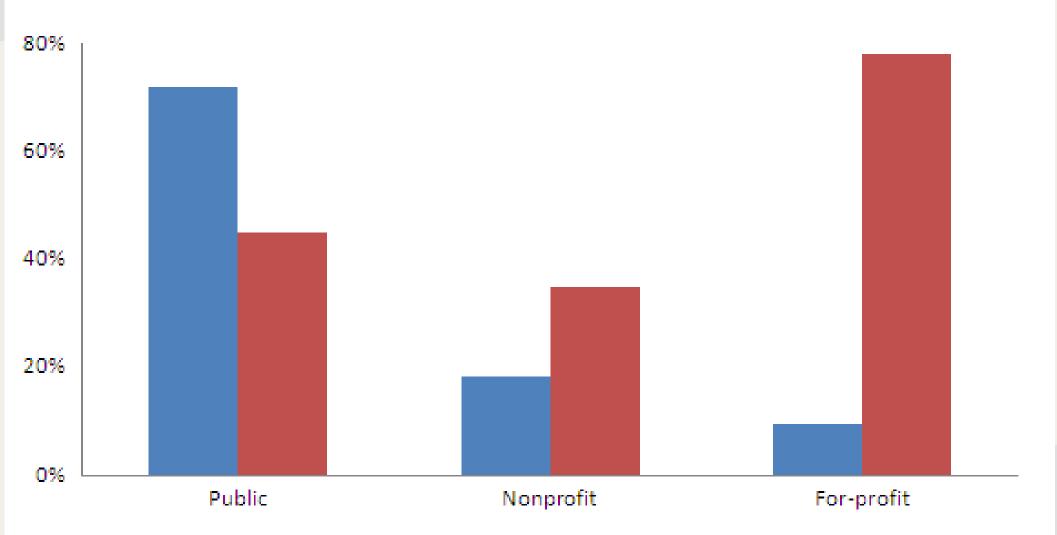
WITHOUT PROMPT, AGGRESSIVE LIMITS ON CO2 EMISSIONS, THE EARTH WILL LIKELY WARM BY AN AVERAGE OF 4°-5°C BY THE CENTURY'S END.

## HOW BIG A CHANGE IS THAT?



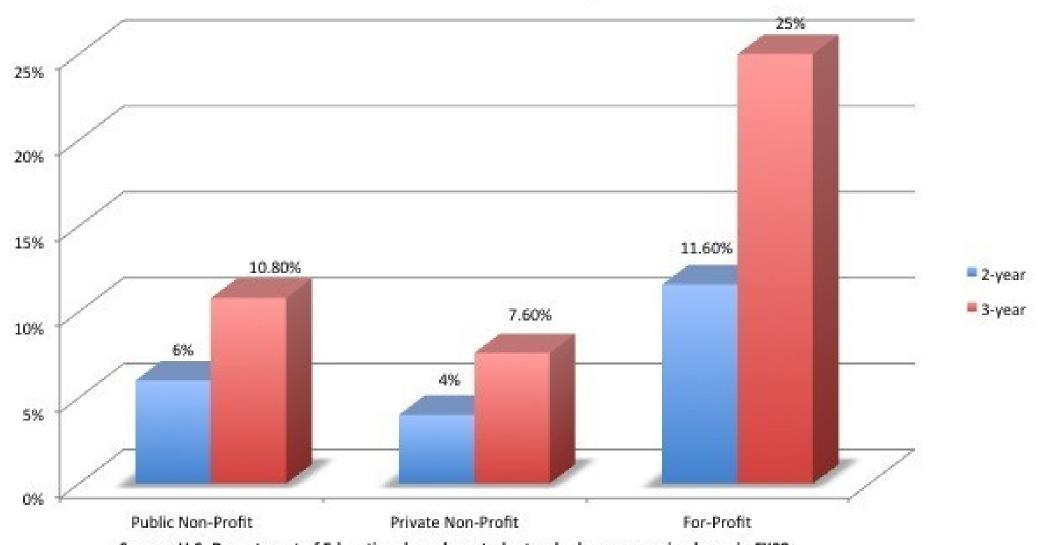
## College enrollment and drop-out rates at different four-year institutions

- Percentage of total national student enrollment
- Percentage who fail to get a degree within six years



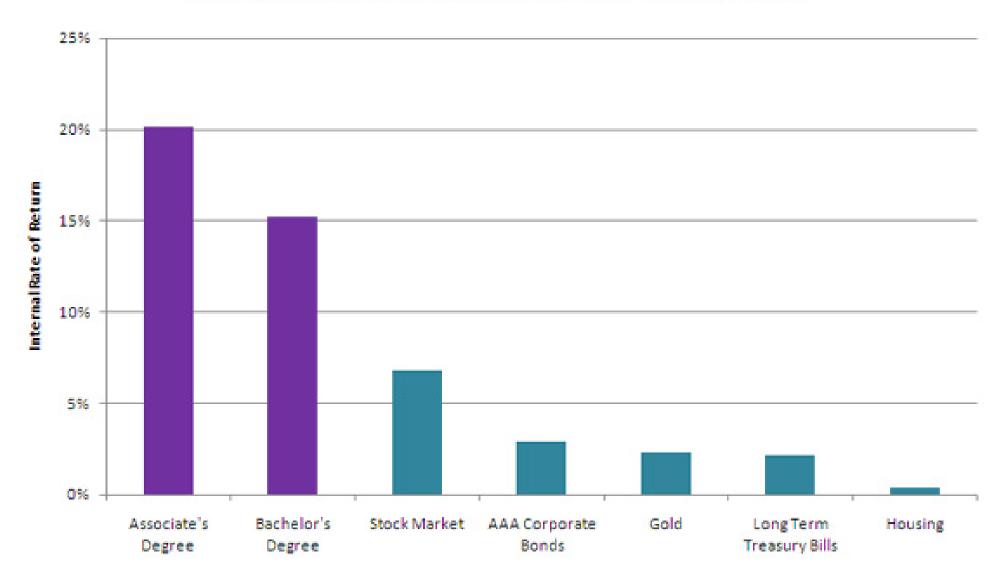
Source: NCES, Harvard Graduate School of Education

### **Student Loan Default Rates-By Education Sector**



Source: U.S. Department of Education, based on students who began repaying loans in FY08

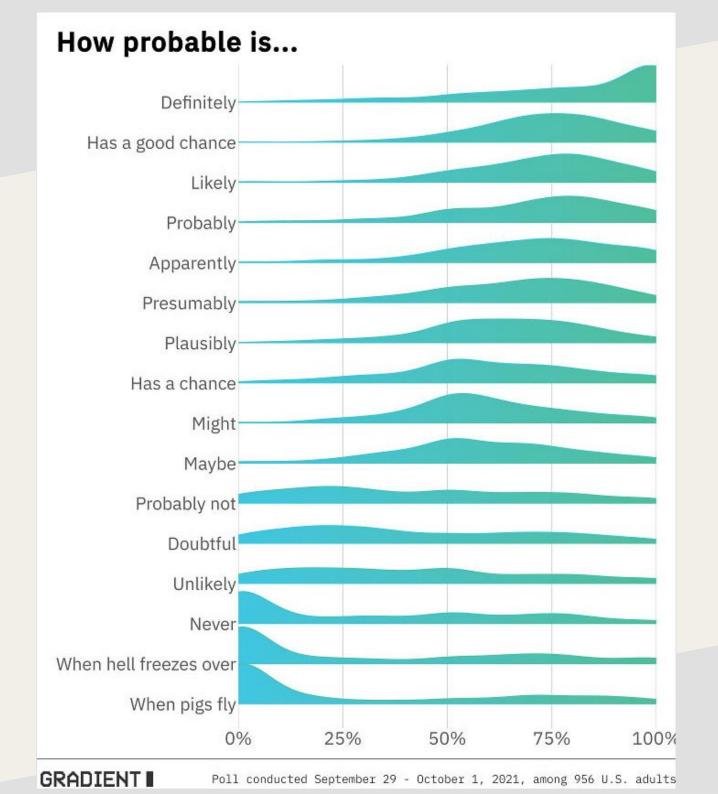
### College Compared to Alternative Investments



Note: Asset returns are geometric averages since 1950. Source: Author's calculations of IRR, values adjusted for inflation using the CPI-U; March Current Population Survey (CPS) 2007-2010 averages; National Mining Association; National Center for Education Statistics; Robert Shiller online data, Long Term Treasury Bills have 10 year maturities after 1953; Federal Reserve Bank of St. Louis.

#### THE ECONOMIC MOBILITY A COLLEGE DEGREE BRINGS Without a College Degree With a College Degree 50% FAMILY A CHILD BORN INTO A LOW-INCOME END UP IN EACH INCOME BRACKET 40 30 20 CHANCE THAT WILL 10 0 Lowest Second Middle Fourth Highest Quintile Quintile Quintile Quintile Quintile

People born into the lowest economic quintile have a much better chance of escaping poverty as adults if they earn a college degree.



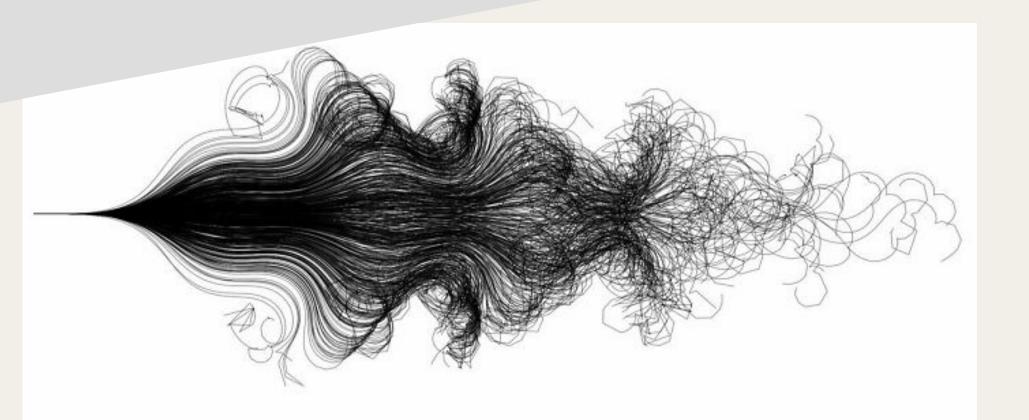
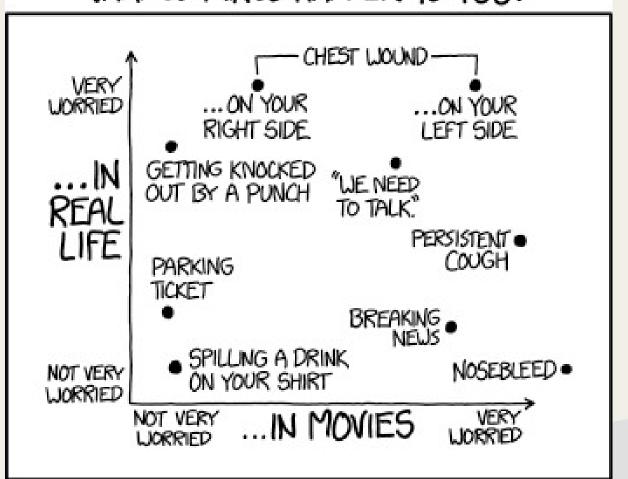


Figure 2: Instability of an unsteered bicycle. This shows 800 runs of a bicycle being pushed to the right. For each run, the path of the front wheel on the ground is shown until the bicycle has fallen over. The unstable oscillatory nature is due to the subcritical speed of the bicycle, which loses further speed with each oscillation.

## HOW WORRIED SHOULD YOU BE WHEN VARIOUS THINGS HAPPEN TO YOU:



HOW EASY IT IS TO GROW UP EASY WITHOUT LEARNING [THING]

- 100 DIGITS OF PI
  - LYRICS TO LIE DIDN'T START THE FIRE

HOW BAD IT IS IF YOU DON'T KNOW [THING]

NOT\_ BAD

HOW TO RIDE A BIKE .

- How to escape Movie Quicksand
- LYRICS TO 12 DAYS
   OF CHRISTMAS
- TV THEME SONGS

THAT CAT BITES ARE REALLY SERIOUS AND IF BITTEN YOU NEED TO WASH THE BITE AND CALL A DOCTOR IMMEDIATELY

THE RED FLAGS FOR AN ABUSIVE RELATIONSHIP

THE SIGNS OF A STROKE

COUGH INTO YOUR ELBOW, NOT YOUR HAND

VERY

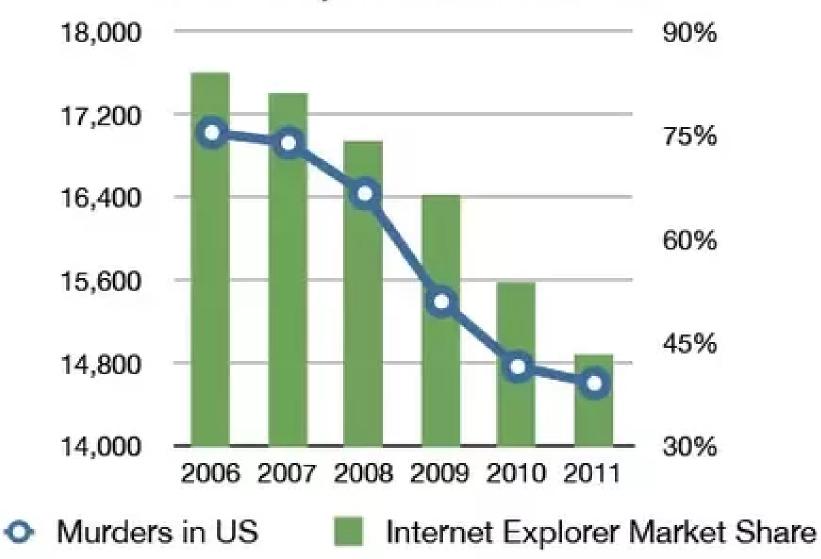
THAT YOU HAVE TO EMPTY THE DRYER LINT TRAP

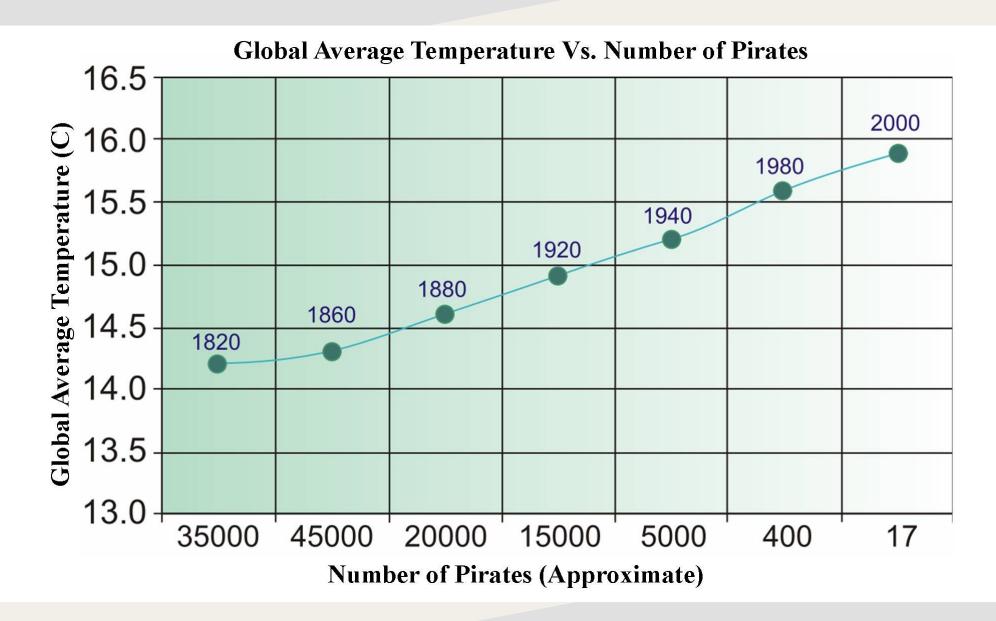
STOP, DROP, AND ROLL

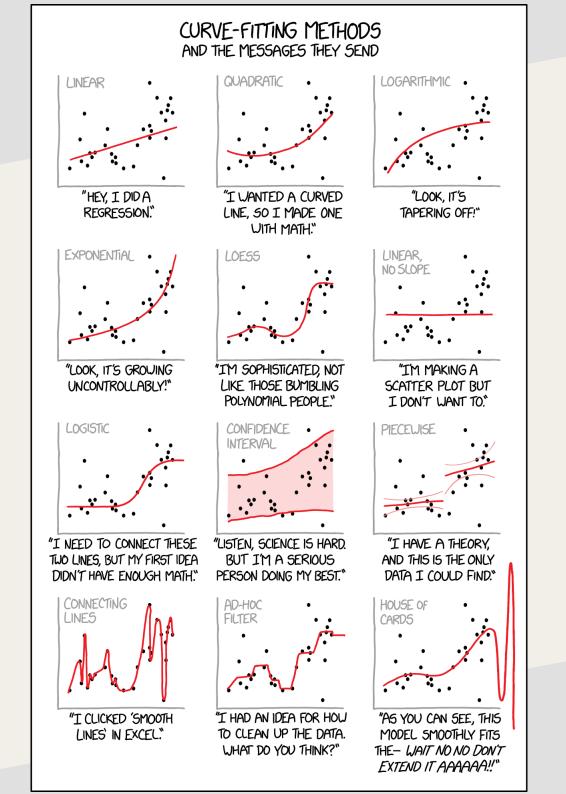
THAT YOU HAVE ... TO PAY TAXES ...

HARD

### Internet Explorer vs Murder Rate





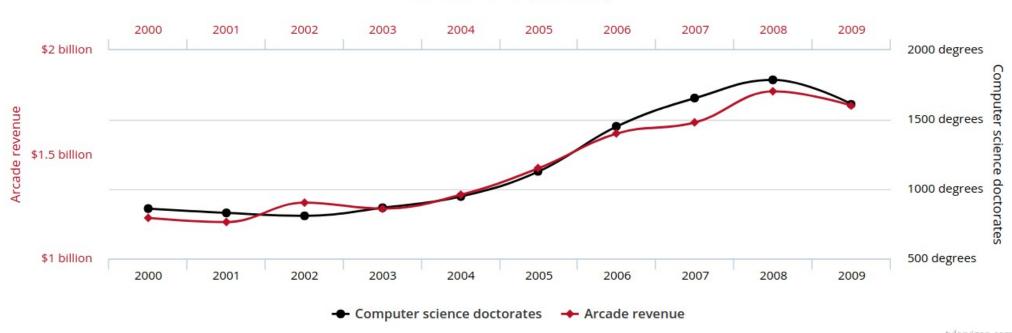


## Total revenue generated by arcades

correlates with

### Computer science doctorates awarded in the US

Correlation: 98.51% (r=0.985065)



tylervigen.com

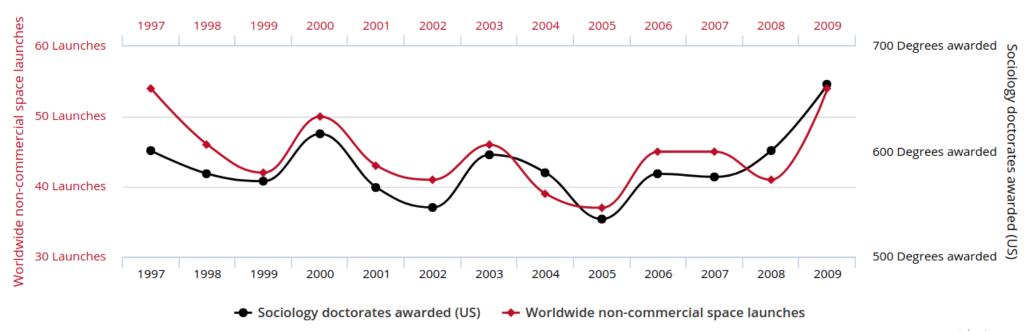
Data sources: U.S. Census Bureau and National Science Foundation

### Worldwide non-commercial space launches

correlates with

### Sociology doctorates awarded (US)

Correlation: 78.92% (r=0.78915)



tylervigen.com

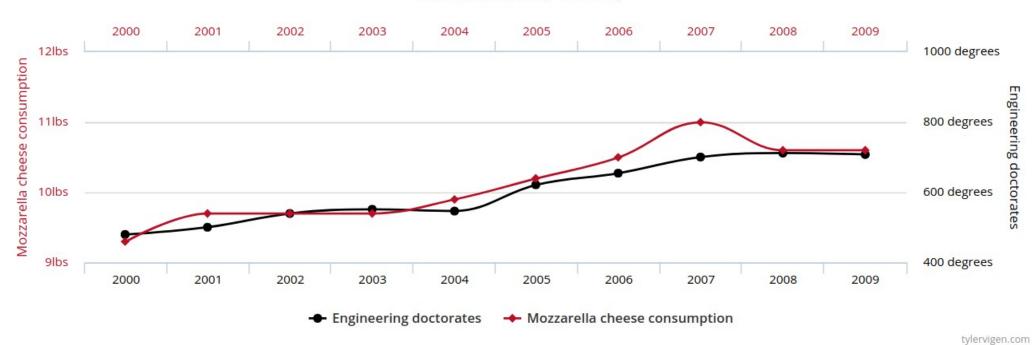
Data sources: Federal Aviation Administration and National Science Foundation

### Per capita consumption of mozzarella cheese

correlates with

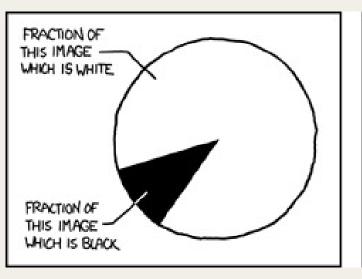
## Civil engineering doctorates awarded

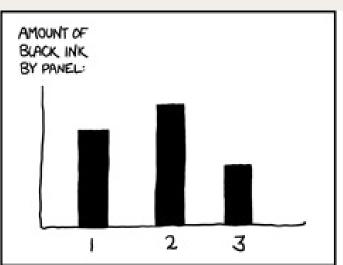
Correlation: 95.86% (r=0.958648)

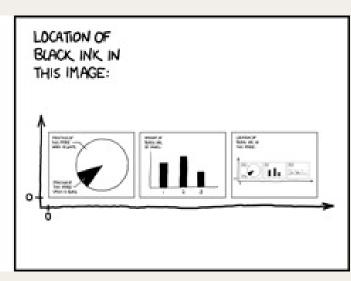


Data sources: U.S. Department of Agriculture and National Science Foundation



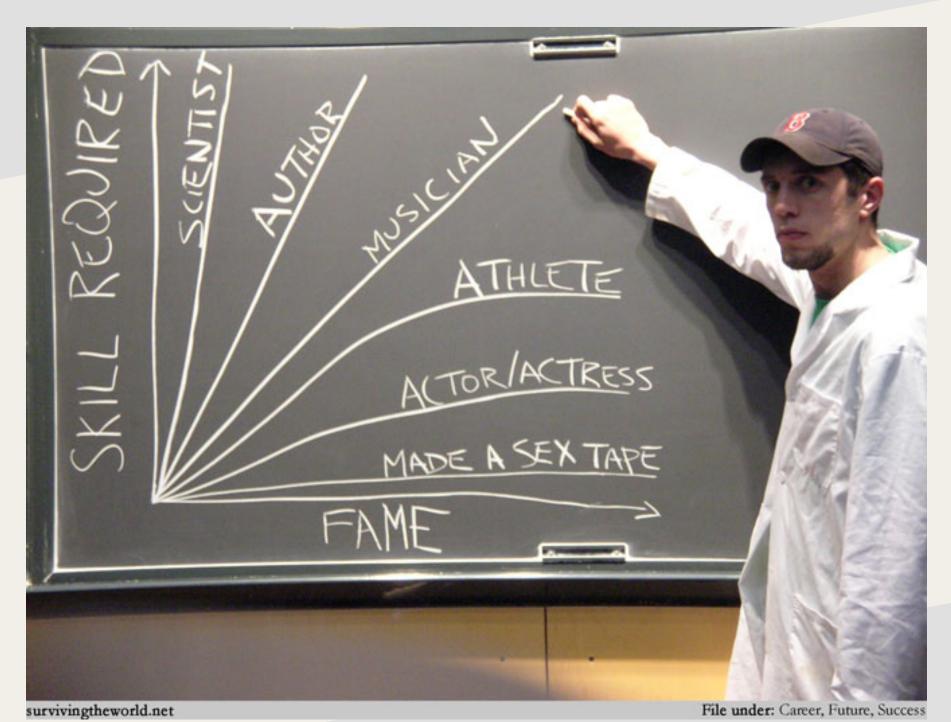




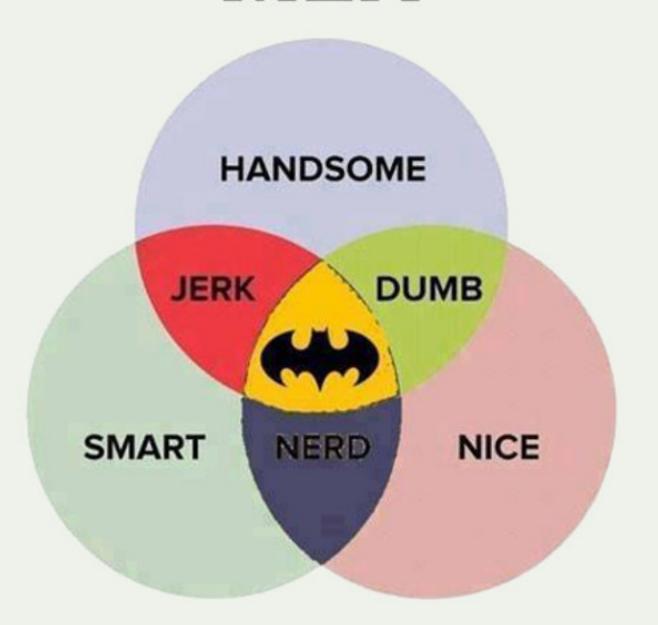


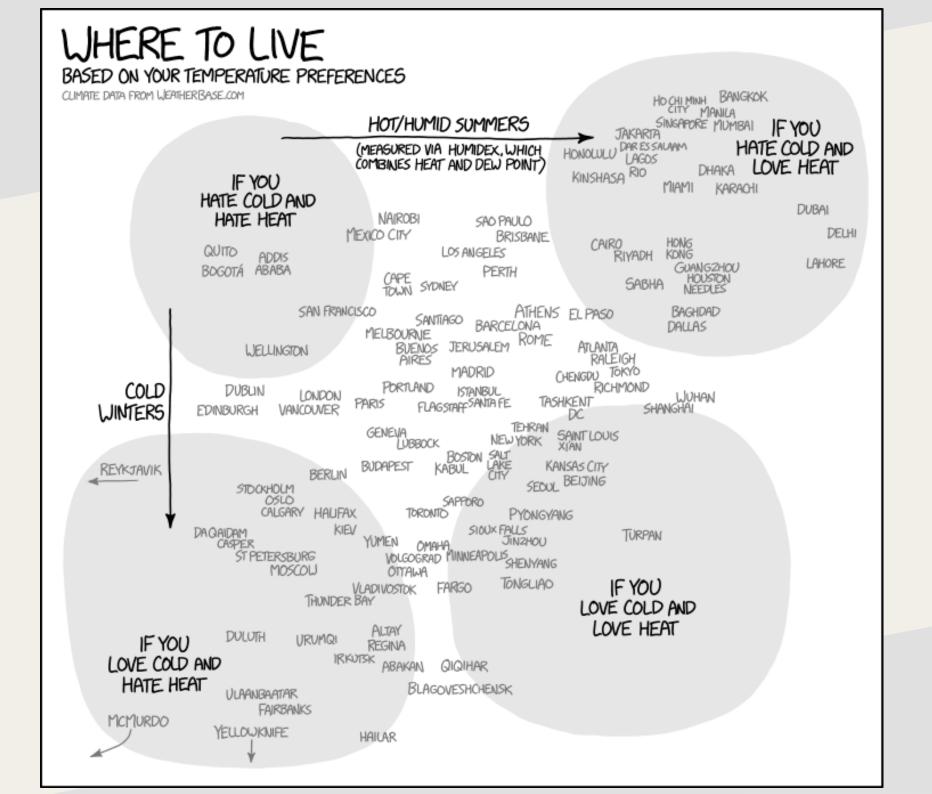
The contents of any one panel are dependent on the contents of every panel including itself. The graph of panel dependencies is complete and bidirectional, and each node has a loop.

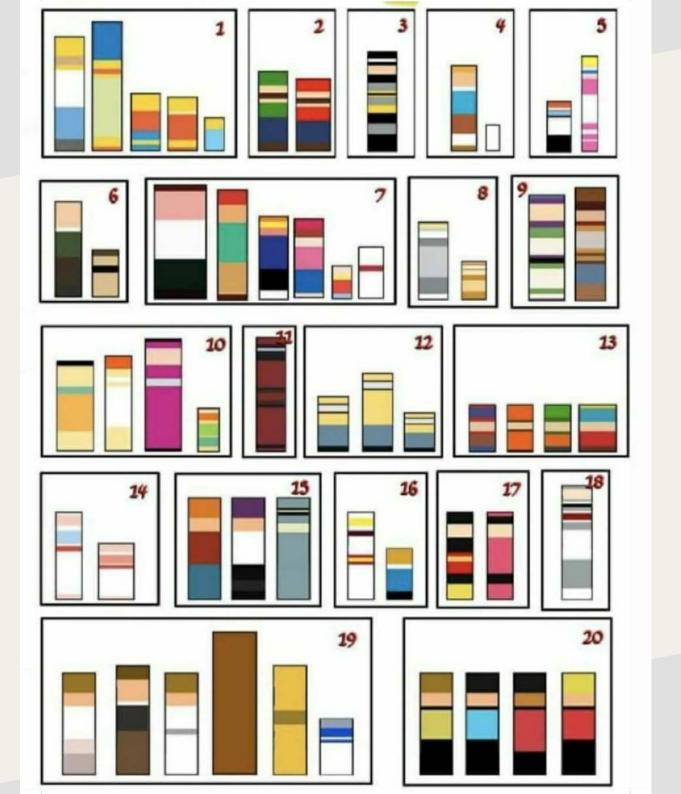
106



## MEN







## 2. Text



A bot wrote this obituary and I'm losing my mind

## RAI ARIIAVKA

Brenda Tent retired from living at the age of old, surrounded by family and natural causes. A librarian from birth, Brenda was an avid collector of dust. She had a sweet heart and married her high school. She loved having hobbies and helping her sons to be



disadvantaged youths. She had no horses but thought she did. The church gave her a choir because she sang like bird and looked like bird and Brenda was a bird. She owed us so many poems.

The funeral will be held in 1977 at heaven. In lieu of flowers, send Brenda more life.

113

# Writing Results Sections

- Tell a story
  - Hopefully about how your data address your research goals / hypotheses
- Use figures & tables as talking points
  - Describe the figure (nearly restating the title & legend)
  - Guide the reader through the main points made by the figure
- Use statistics as "citations" to support assertions

# Writing Results Sections (cont.)

- Text in Results can also clarify unclear points in figures
  - And qualifications to analyses
    - That may be further considered in Discussion & Limitations sections
- Results vs. Discussion sections:
  - Results: Just the facts (and how to understand them)
  - Discussion: Interpreting the facts, discussing their implications for research & practice

## References & Resources

- Junk Charts (https://junkcharts.typepad.com/)
- U. of Leicester: *Presenting numerical data*
- The unspoken rules of visualisation (and when to break them)
- Tableau: Data visualization beginner's guide:
   A definition, examples, and learning resources
- Avinash Kaushik: 7 Data Presentation Tips
- Lots more from United Nations Economic Commission for Europe
- Writing results for dissertations

## THE END

## More Resources

- XKCD's chart of the average temperature of the earth: <a href="https://xkcd.com/1732/">https://xkcd.com/1732/</a>. It's too long for here, but very eye-opening
- Information Is Beautiful, an website devoted to presenting interesting data (many fields & interests). Including:
  - The MicrobeScope: Deadliness, contagiousness, n sufferers, media coverage and more of various microbes
  - <u>Snake Oil Sports Supplements</u>? What can you take to actually improve performance, according to science?
  - The <u>longevity</u> of a married, happy-go-lucky, outdoors-loving, sex-mad, hippy, party girl in senior management with a cat
- The Fallen of World War II, a gorgeously presented video of those deaths and the value of the Long Peace we're now loosing
- 3D Map of Galaxies, a rotating map of surveyed galaxies around the Milky Way (our to ~13 billion light years)
- VizWorld: A website about visualizations with a concentration or economics & "innovative thinkers" (<a href="https://vizworld.com/">https://vizworld.com/</a>)