

# Visualization Design & Memorable Chart Junk

# Announcements

- Lecture videos will be posted to the YouTube playlist listed on the course page
- Slides will be available before lecture
- Paper discussions will require A/V if possible.
- My contact info will soon be updated on the course site

# Today

- “Good” Design (30 min)
  - Photography tips
  - Principles of Effective Website Design
  - Principles of Good User Interface Design
  - Examples of Good (Bad) Visualization Design
- Today’s Readings
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Kodak

## Top Ten Tips

1 Get down on their level



2 Use a plain background



3 Use flash outdoors



4 Move in close



5 Take some vertical pictures



6 Lock the focus



7 Move it from the middle



8 Know your flash's range



9 Watch the light



10 Be a picture director



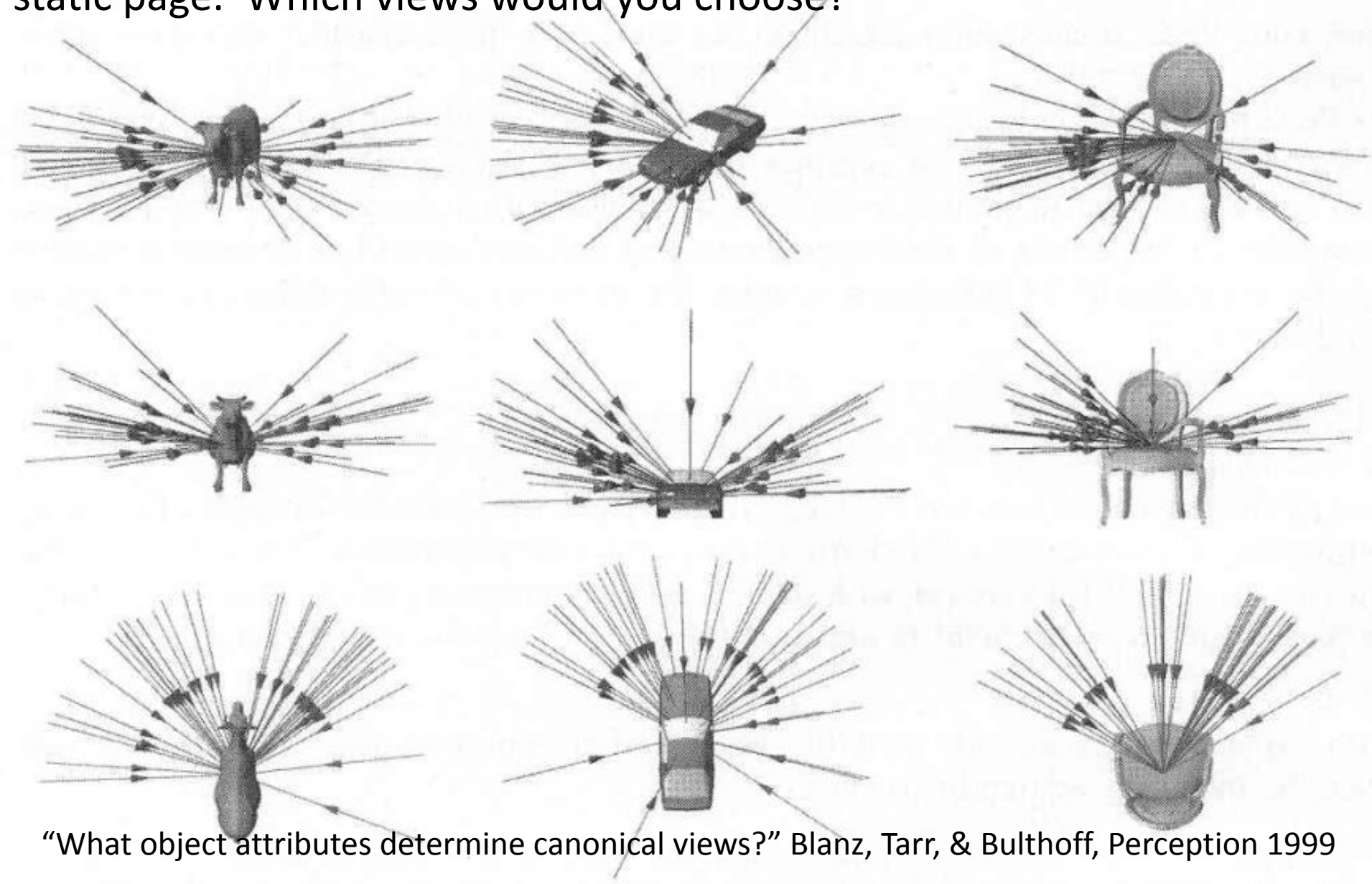
# “Canonical” Viewpoints

- From Dictionary.com:

- authorized; recognized; accepted
- the body of rules, principles, or standards accepted as axiomatic and universally binding in a field of study or art: the neoclassical canon
- a fundamental principle or general rule: the canons of good behavior
- a standard; criterion: the canons of taste

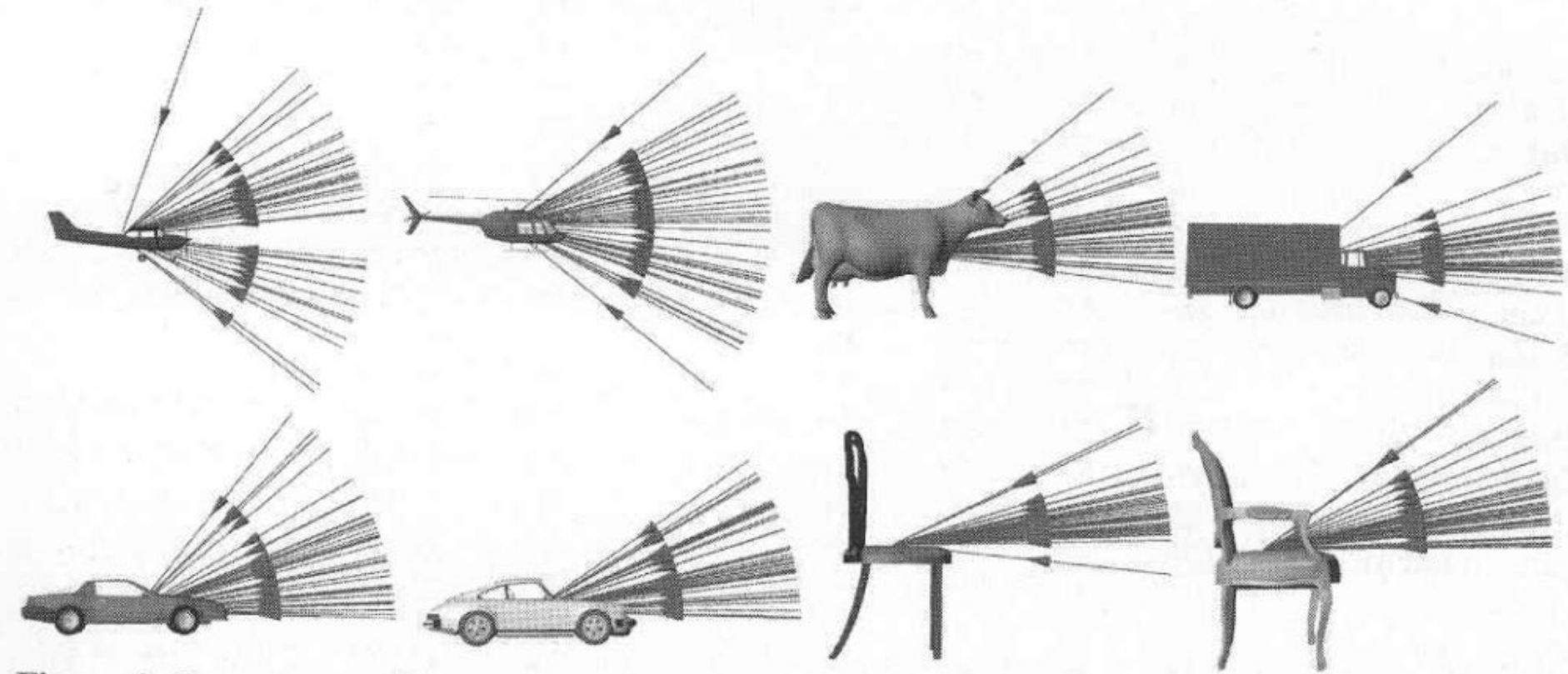
“What object attributes determine canonical views?” Blanz, Tarr, & Bulthoff, Perception 1999

Suppose you were making a brochure and you tried to give your customers the best possible impression of the objects shown on the static page. Which views would you choose?



“What object attributes determine canonical views?” Blanz, Tarr, & Bulthoff, Perception 1999

- Salience and significance of the features
- Stability of viewpoint to small transformations
- Minimize number of occluded features
- Familiarity, Functionality, Aesthetic criteria



“What object attributes determine canonical views?” Blanz, Tarr, & Bulthoff, Perception 1999



1a, 2



1a



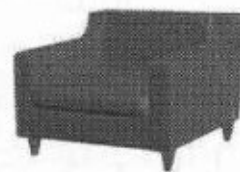
1a



1b



1b



2



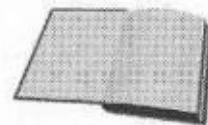
1a, 2



1a, 2



1b



1b



2



1b



1a, 2



1b



1a, 2



1b



1b



1b



1b



1a, 2



1a



1b



1b



1a, 2



1b



1b



1a, 2



1b



1a, 2



1a



1b



1a, 1b



# Rule of Thirds

[http://en.wikipedia.org/wiki/Rule\\_of\\_thirds](http://en.wikipedia.org/wiki/Rule_of_thirds)

- align subject with guide lines and intersection points, discourage placement of the subject at the center
- placing the horizon on the top or bottom line, avoid dividing picture in half



# Today

- “Good” Design (30 min)
  - Photography tips
  - Principles of Effective Website Design
  - Principles of Good User Interface Design
  - Examples of Good (Bad) Visualization Design
- Today’s Readings (30 min)
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Principles of Effective Website Design

- Guiding the eye (position, color, contrast, size, design elements)
- Spacing, padding, white/empty space, reduce cognitive load
- Navigation/orientation
- Typography (font, size, color, paragraphs)
- Usability/standards/conventions be obvious,  
“Don’t make users think”
- Consistency
- Alignment, polished, simplicity
- Effective writing
- Clarity, sharpness, contrast, exaggeration

# Principles of Good User Interface Design

- Consistency and standards
  - Match real world: words, phrases and concepts familiar to the user, real-world conventions, natural and logical order, coherency
- Flexibility and efficiency of use: cater/tailor to both inexperienced and experienced users
  - Know your user, user testing, listen to the user
- User control and freedom: a clearly marked "emergency exit" to leave the unwanted, support undo and redo
- Aesthetic and minimalist design: every extra unit of information competes with and diminishes visibility of relevant information
  - System status: keep users informed
- Recognize, diagnose, and recover from errors
  - Error prevention: good error messages, eliminate error-prone conditions, confirmation option
- Help and documentation
  - Recognition rather than recall: information/instructions should be visible or easily retrievable

[http://www.sylvantech.com/~talin/projects/ui\\_design.html](http://www.sylvantech.com/~talin/projects/ui_design.html)

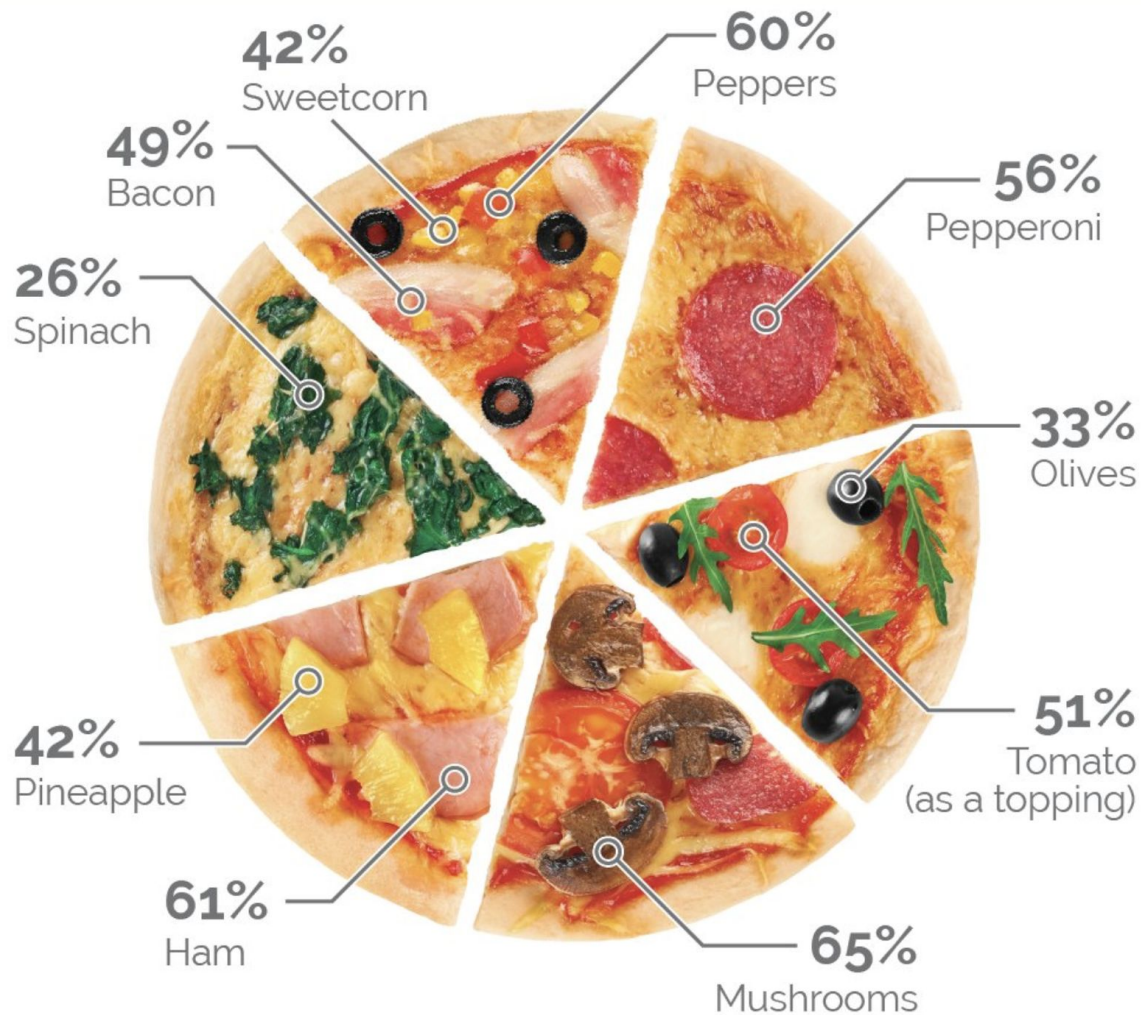
[http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html)

# Today

- “Good” Design (30 min)
  - Photography tips
  - Principles of Effective Website Design
  - Principles of Good User Interface Design
  - Examples of Good (Bad) Visualization Design
- Today’s Reading (30 min): “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Mushroom is the UK's most liked pizza topping

Generally speaking, which of the following toppings do you like on a pizza? Select as many as you like



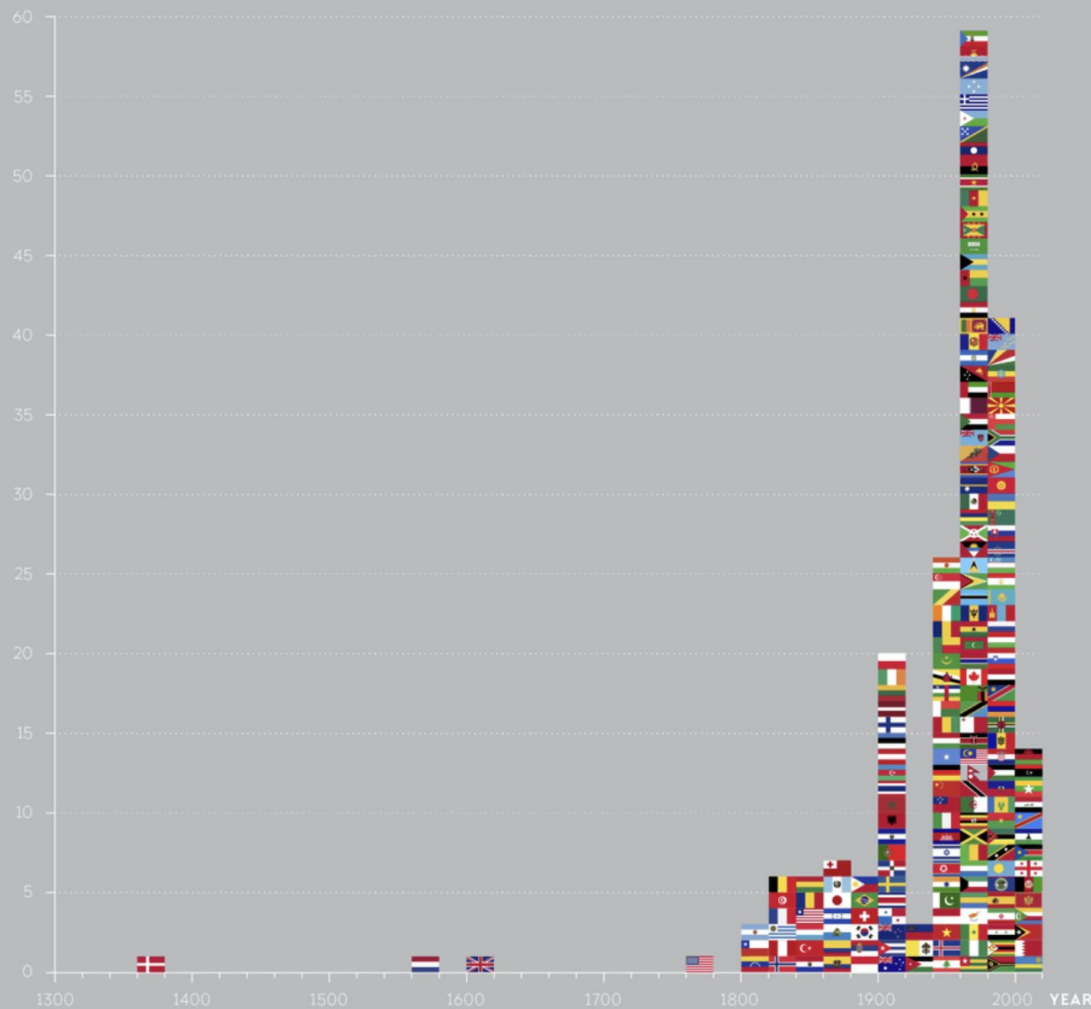
Other items not depicted include: onions (62%), chicken (56%), beef (36%), chillies (31%), jalapeños (30%), pork (25%), tuna (22%), anchovies (18%). 2% of people say they only like Margherita pizzas

*Worst pie chart ever?*

# How old are national flags?

This timeline is based on the date in which each nation adopted their current national flag and is sorted by a 20 year interval. As you can see only four of the current national flags are dated before the 19th century.

NUMBER OF FLAGS

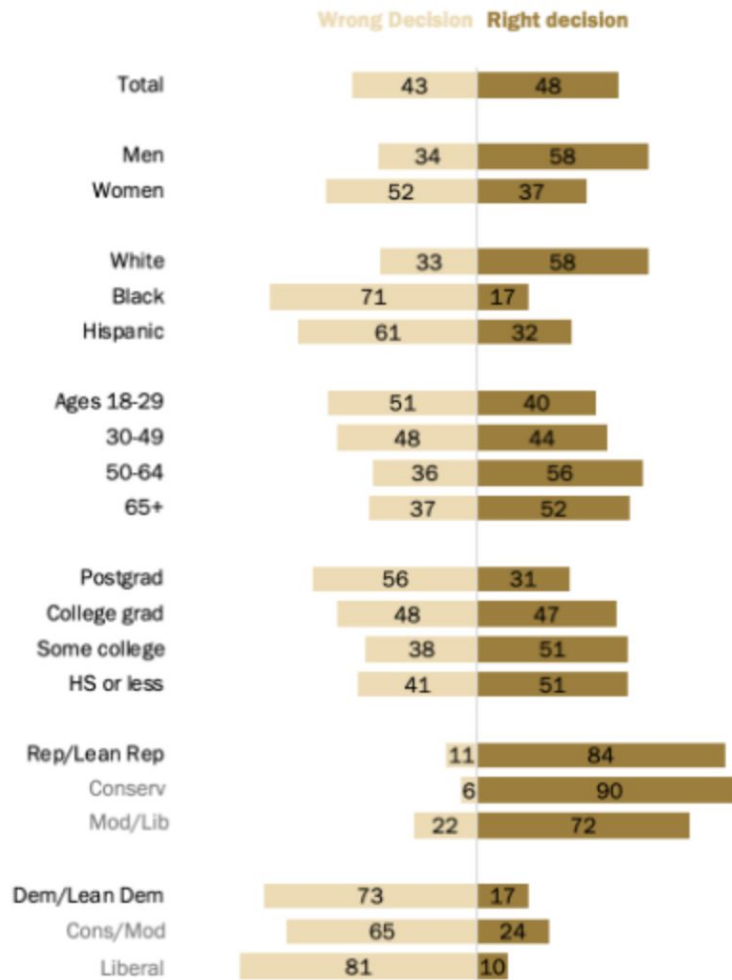


*Excellent use-case  
For on-chart imagery*



## Wide partisan, demographic differences in views of U.S. decision to conduct airstrike that killed Soleimani

% who say U.S. decision to conduct the airstrike that killed Iranian Gen. Soleimani was the ...



*Unbiased reporting  
Good data-ink ratio*

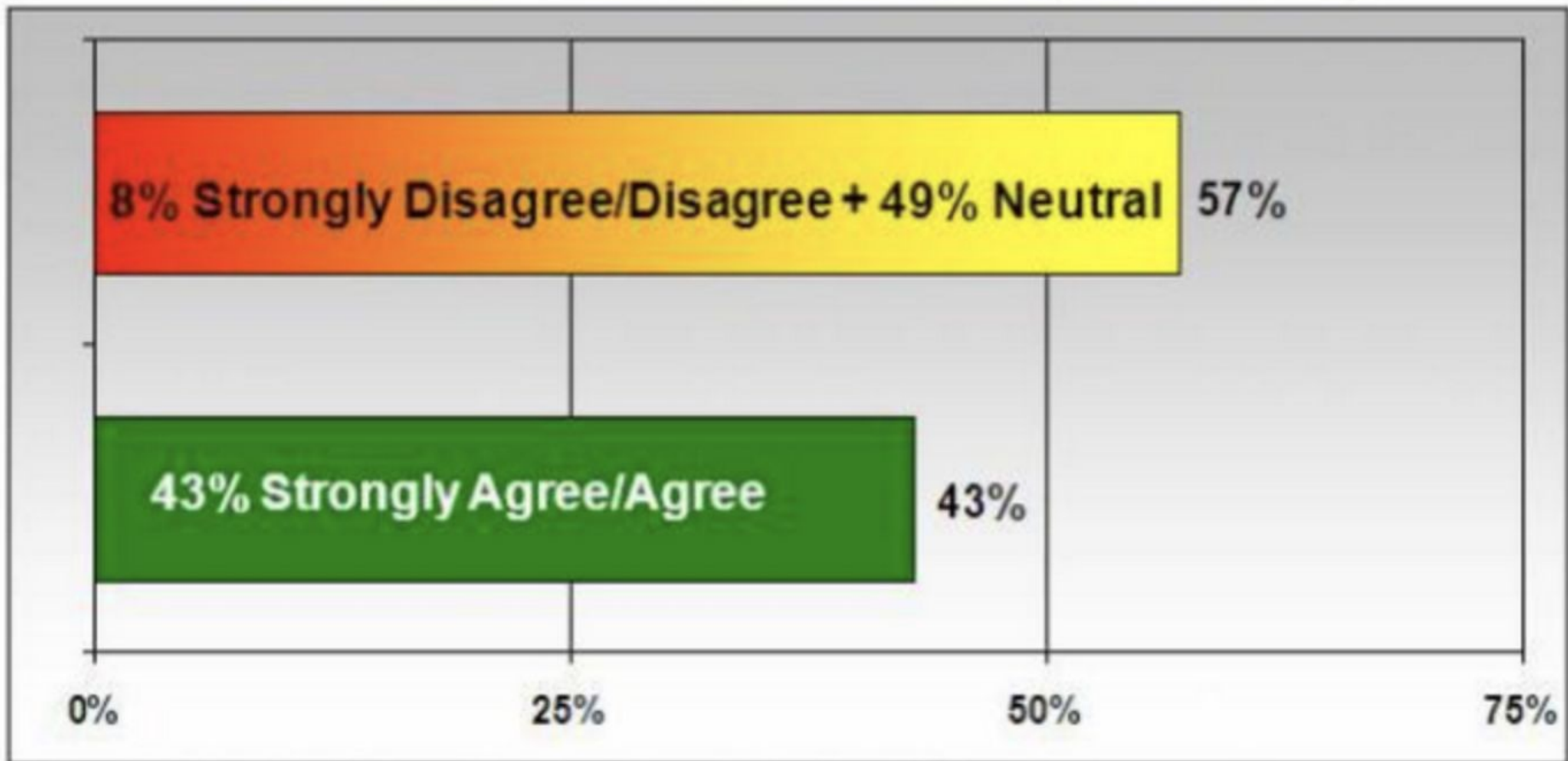
Notes: Don't know responses not shown. Whites and blacks include only those who are not Hispanic; Hispanics are of any race.

Source: Survey of U.S. adults conducted Jan. 8-13, 2020.

PEW RESEARCH CENTER

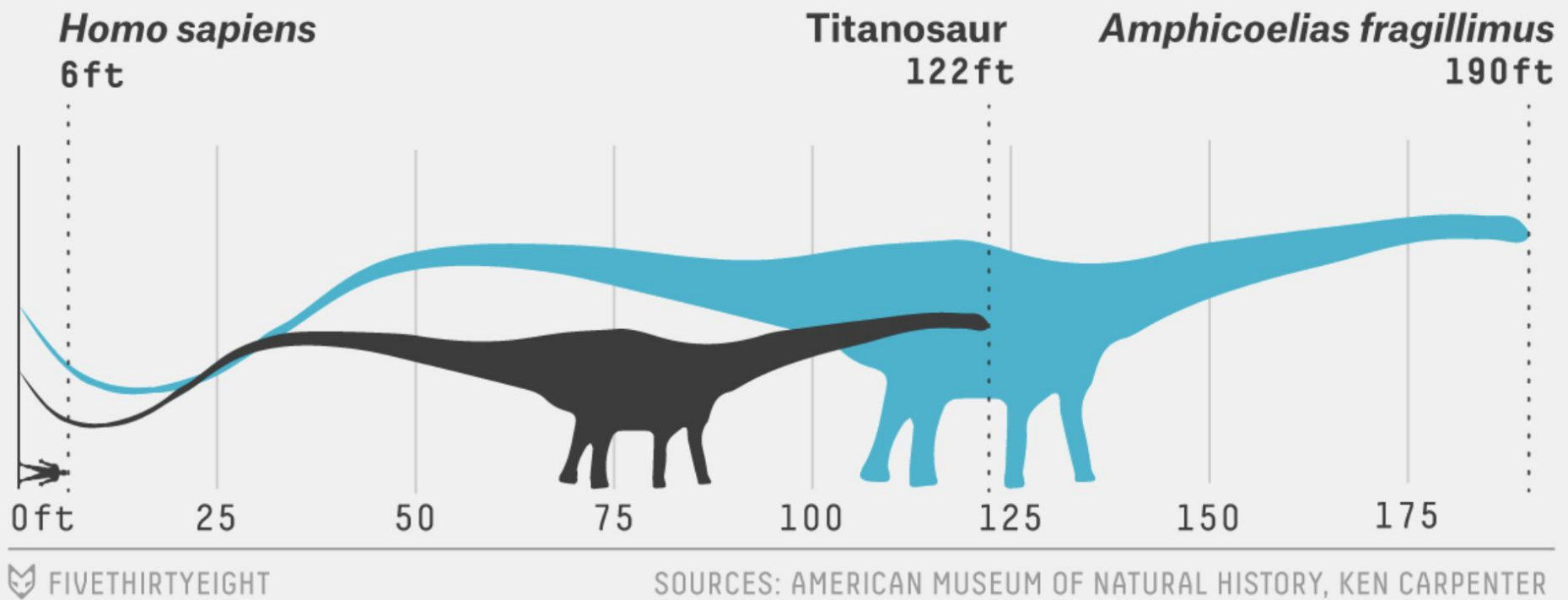


**Figure 2. Parent Response to  
“New vaccines are safe for my child(ren).”**



Source: C.S. Mott Children's Hospital National Poll on Children's Health, March 2007

*Over-simplification?  
Manipulative visualization?  
Confusing / unnecessary colors/gradient*



*Simple, intuitive, effective*  
*Educational for all ages*

'18-'19

'19-'20

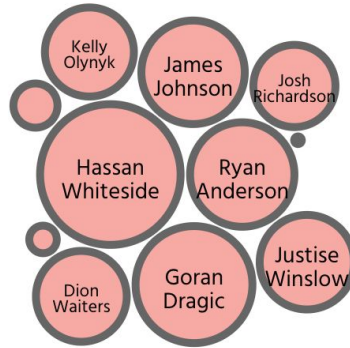
'20-'21

'21-'22

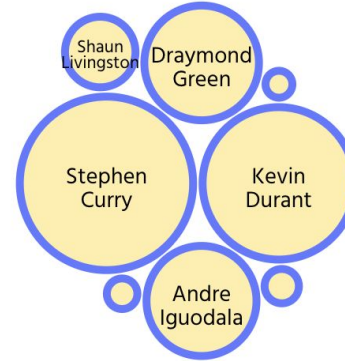
'22-'23

'23-'24

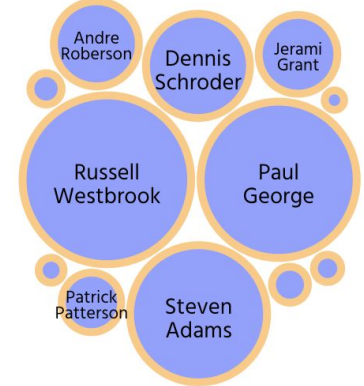
Miami Heat - \$129.6M



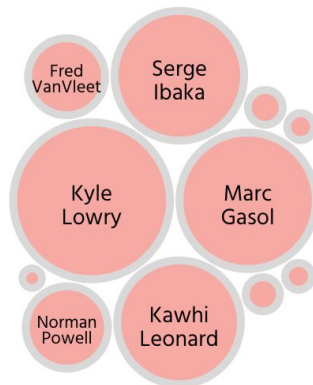
Golden State Warriors - \$121.0M



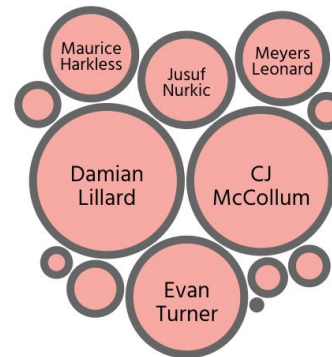
Oklahoma City Thunder - \$146.8M



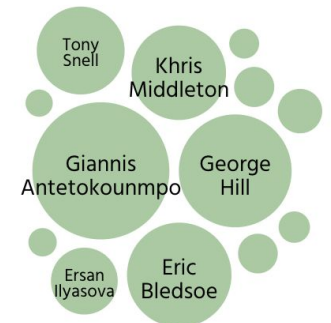
Toronto Raptors - \$131.4M



Portland Trail Blazers - \$126.0M

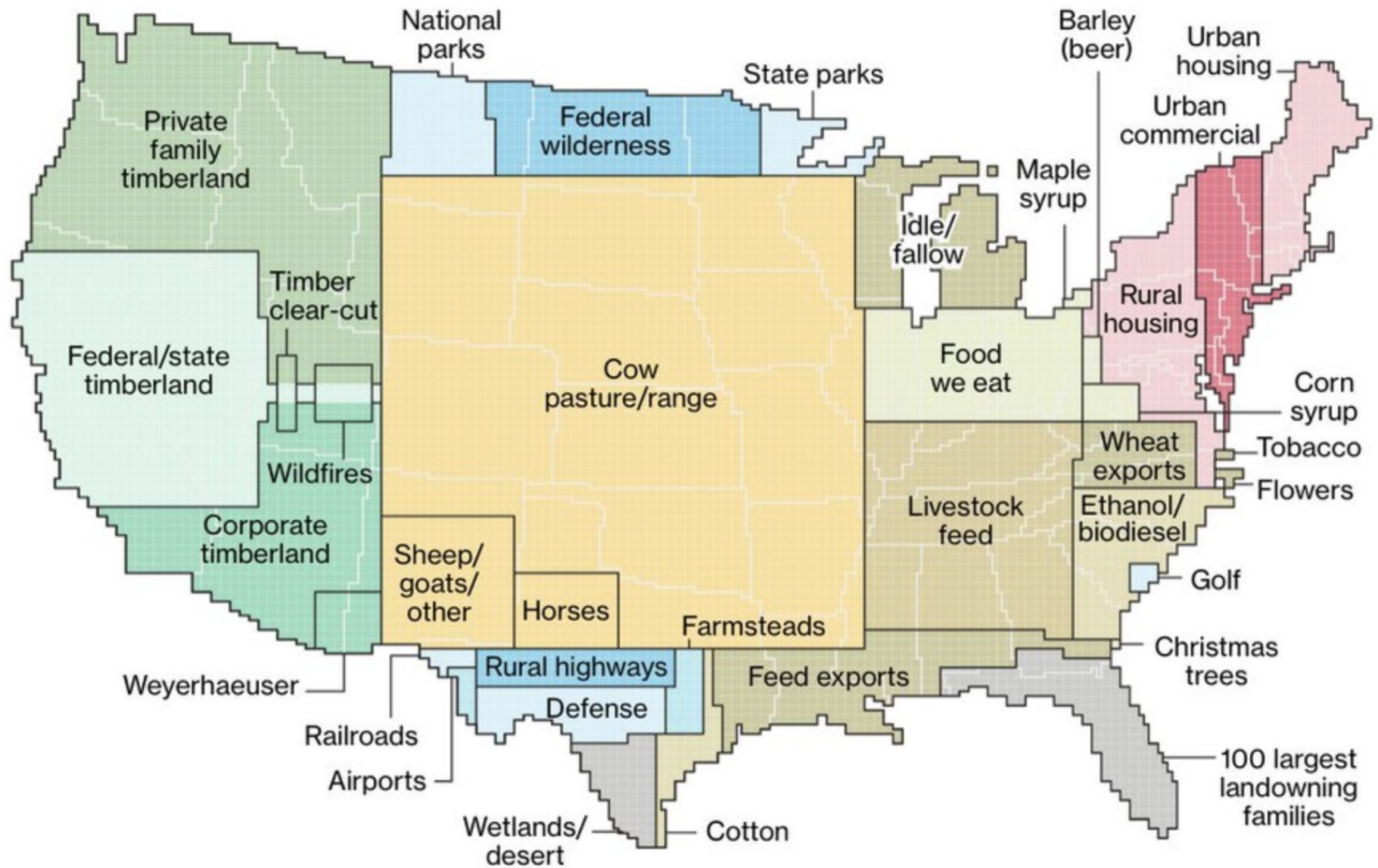


Milwaukee Bucks - \$107.4M



<http://cahaber.me/#/basketball>

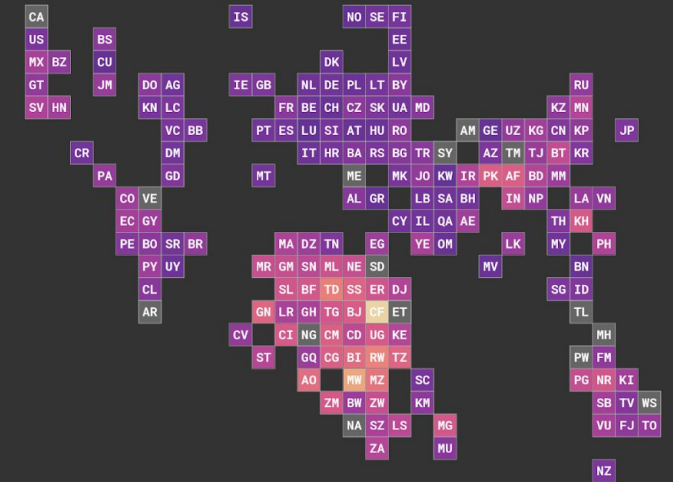
*Compare players within team*  
*Compare players on different teams*  
*Different team choices/strategies*



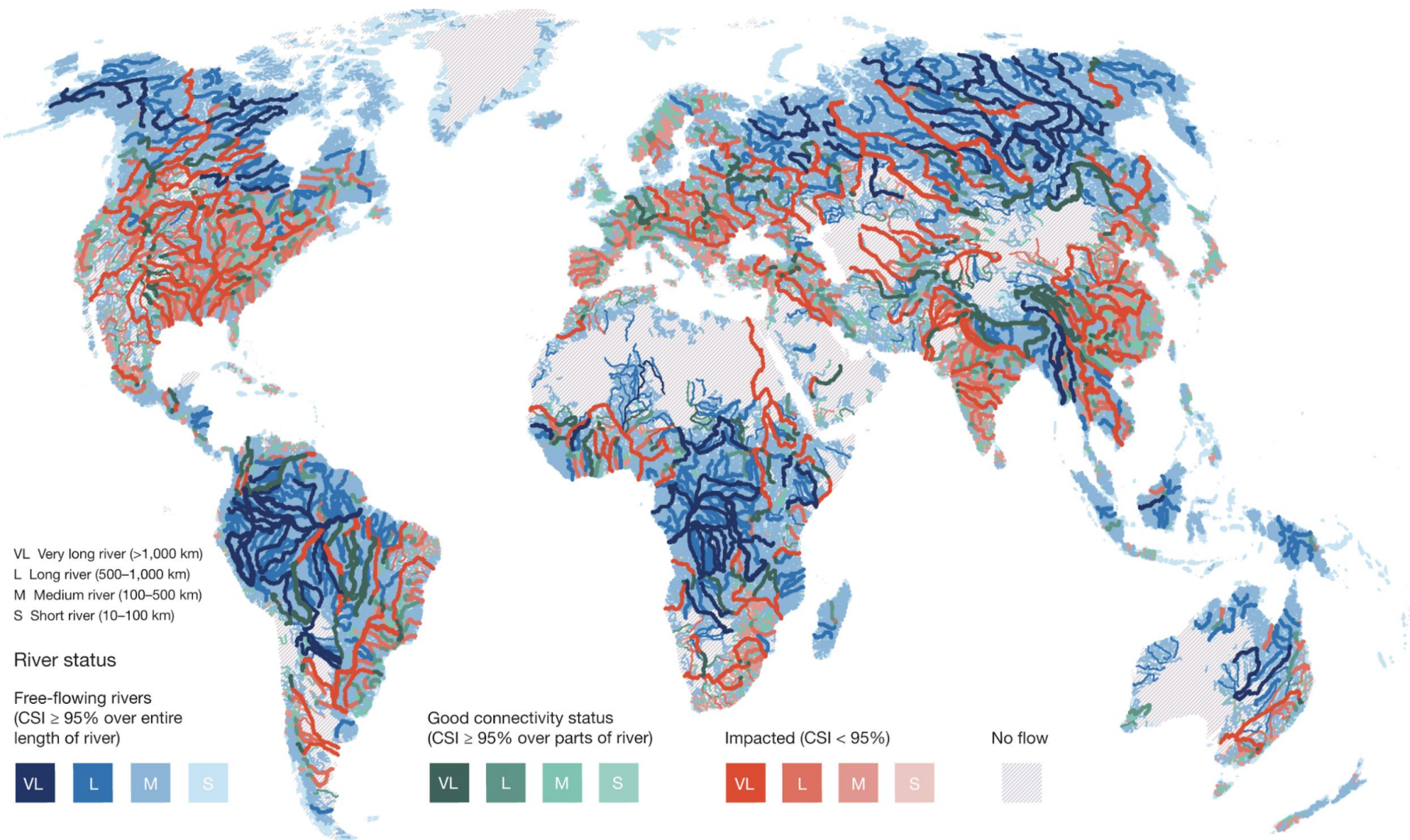
*Effective non-pie shaped  
pie chart!*



Latest reported student to teacher ratio per country and continent (2012-2018)



*Are all datasets relevant to plot on a world map?*



*Allows us to clearly see significant variations in distributions of rivers  
AND use (overuse?) of technology to control nature*



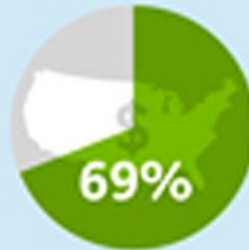
**1.7 million** people per year get an infection during a hospital stay

**98,987** people in the U.S. die annually from HAIs

System  
**\$35**  
Billion/yr



**9.4%**  
of total  
inpatient costs  
are HAI-related

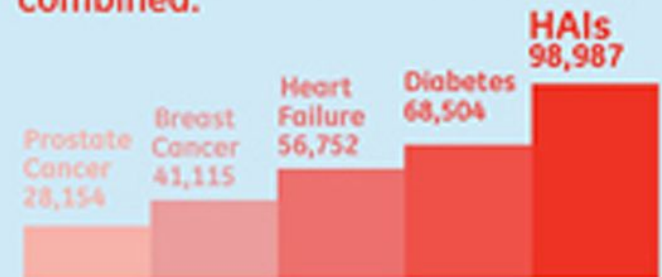


More than  $\frac{2}{3}$  of HAIs affect people with Medicare or Medicaid

Patient  
**\$1,100**  
per admission



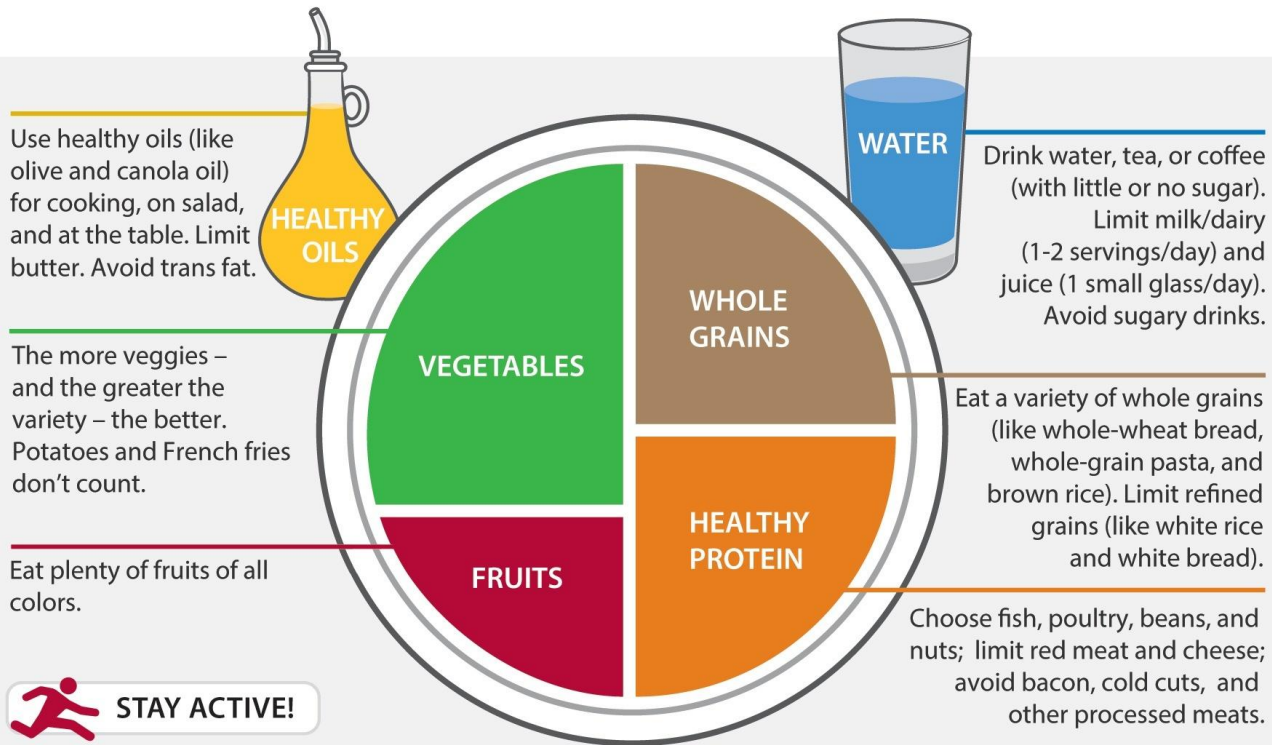
HAIs kill more people each year than Breast Cancer and Prostate Cancer combined.



[http://www3.gehealthcare.com/en/Products/Categories/Healthcare\\_IT/Quality\\_Management#tabs/tab1900328377C74CAC8AD7E8D4A2072591](http://www3.gehealthcare.com/en/Products/Categories/Healthcare_IT/Quality_Management#tabs/tab1900328377C74CAC8AD7E8D4A2072591)



# HEALTHY EATING PLATE



© Harvard University



Harvard T.H. Chan School of Public Health  
The Nutrition Source  
[www.hsph.harvard.edu/nutritionsource](http://www.hsph.harvard.edu/nutritionsource)

Harvard Medical School  
Harvard Health Publications  
[www.health.harvard.edu](http://www.health.harvard.edu)



<https://cdn1.sph.harvard.edu/wp-content/uploads/sites/30/2012/09/HEPJan2015.jpg>

*Vast improvement on original food pyramid*



- **Scientific Visualization (SciVis)**
  - really large quantities of data
  - data usually has inherent structure
  - often has a spatial and/or temporal component (coordinate system)
  - often appropriate to use of 3D visualization techniques
  - such as medical, hurricane, computational fluid dynamics (CFD) data
- **Information Visualization (InfoVis)**
  - smaller datasets
  - data that does not have an inherent structure (may not have coordinate system)
  - financial stock market data, demographic census data, genetic data, etc.
- **Visual Analytics**
  - involves a cycle of rapidly creating visualizations to answer questions and generate new questions about a dataset
- **Infographics** are typically in the realm of InfoVis, and often they show the results of the visual analytics process, but SciVis is not really a part of most infographics.
- Annual IEEE Visualization and IEEE InfoVis (Information Visualization) conferences are two separate entities. The set of people organizing, attending and involved one conference is almost disjoint from the other set.

<http://blog.visual.ly/the-beautiful-world-of-scivis/>

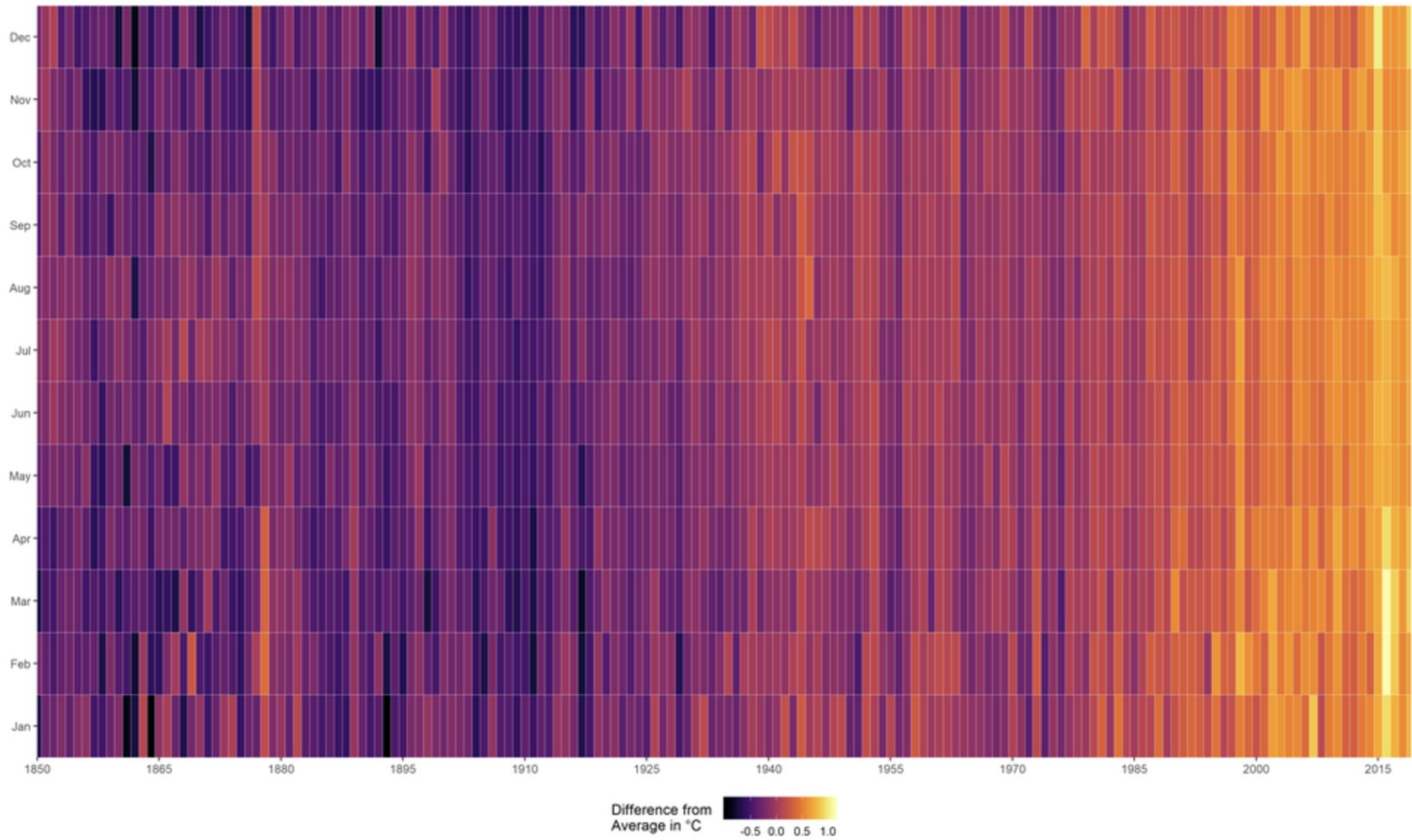
<https://visualizeit.wordpress.com/2007/06/07/the-great-infovis-and-scivis-divide/>



*An Interactive Visualization!  
(interactive data collection)*

# Average World Temperature Since 1850

20th century average 13.7°C, Data HadCRUT4

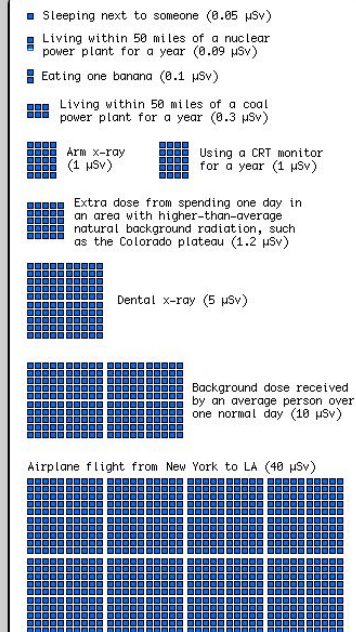


*Visualization Challenge: very large datasets*



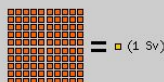
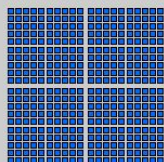
# Radiation Dose Chart

This is a chart of the ionizing radiation dose a person can absorb from various sources. The unit for absorbed dose is "sievert" (Sv), and measures the effect a dose of radiation will have on the cells of the body. One sievert (all at once) will make you sick, and too many more will kill you, but we safely absorb small amounts of natural radiation daily. Note: The same number of sieverts absorbed in a shorter time will generally cause more damage, but your cumulative long-term dose plays a big role in things like cancer risk.



Using a cell phone (0  $\mu$ Sv)—a cell phone's transmitter does not produce ionizing radiation\* and does not cause cancer.  
\* Unless it's a bananaphone.

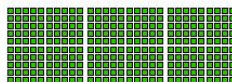
■ = (0.05  $\mu$ Sv)



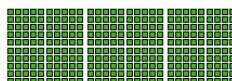
Ten minutes next to the Chernobyl reactor core after explosion and meltdown (50 Sv)



EPA yearly release target for a nuclear power plant (30  $\mu$ Sv)

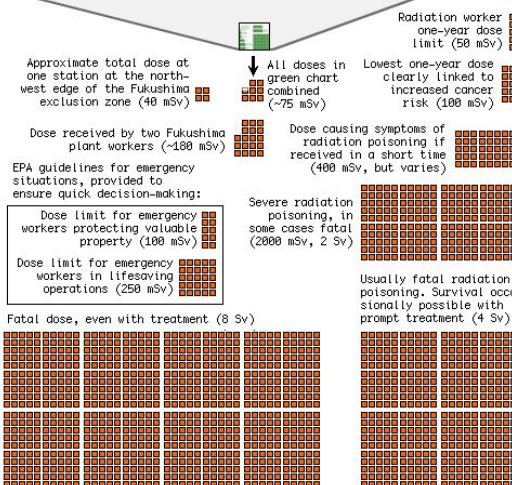
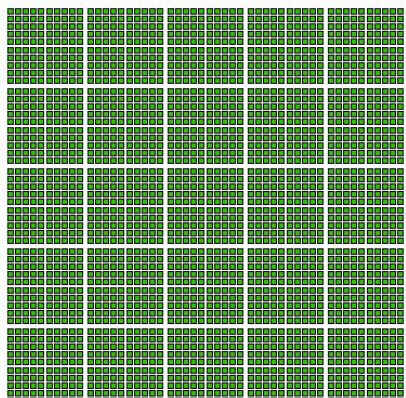


Dose from spending an hour on the grounds at the Chernobyl plant in 2010 (6 mSv in one spot, but varies wildly)



Chest CT scan (7 mSv)

Maximum yearly dose permitted for US radiation workers (50 mSv)



SOURCES:  
<http://www.nrc.gov/reading-rm/doc-collections/cfr/part020/>  
<http://www.nema.ne.gov/technological/dose-limits.html>  
<http://www.deq.idaho.gov/nl/oversight/radiation/dose-calculator.cfm>  
<http://www.deq.idaho.gov/nl/oversight/radiation/radiation-guide.cfm>  
[http://www.bnl.gov/bnlweb/DOE/03580/Chapter\\_8.pdf](http://www.bnl.gov/bnlweb/DOE/03580/Chapter_8.pdf)  
[http://dels-old.nas.edu/dels/rpt\\_briefs/verr\\_final.pdf](http://dels-old.nas.edu/dels/rpt_briefs/verr_final.pdf)  
<http://people.reed.edu/~emcanalis/radiation.html>  
<http://en.wikipedia.org/wiki/Sievert>  
<http://blog.vornaskotti.com/2010/07/18/into-the-zone-chernobyl-prigpat/>  
<http://www.nrc.gov/reading-rm/doc-collections/r2act-sheets/tritium-radiation-fa.html>  
[http://www.merit.go.jp/component/\\_dMenu/other/detail/\\_icsFiles/aidefile/2011/03/16/1303727-1716.pdf](http://www.merit.go.jp/component/_dMenu/other/detail/_icsFiles/aidefile/2011/03/16/1303727-1716.pdf)  
<http://radiology.rsna.org/content/248/1/254>

Chart by Randall Munroe, with help from Ellen, Senior Reactor Operator at the Reed Research Reactor, who suggested the idea and provided a lot of the sources. I'm sure I've added in lots of mistakes; it's for general education only. If you're basing radiation safety procedures on an internet PNG image and things go wrong, you have no one to blame but yourself.

<https://xkcd.com/radiation/>

Visualization Challenge:  
representing different scales

# Periodic Table of Elements

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18									
1 <b>H</b> Hydrogen 1.00794	Atomic # Symbol Name Atomic Mass																2 <b>He</b> Helium 4.002602									
<div><div><div>C Solid</div><div>Hg Liquid</div><div>H Gas</div><div>Rf Unknown</div></div><div><div>Metals</div><div>Alkali metals</div><div>Alkaline earth metals</div><div>Lanthanoids</div><div>Actinoids</div><div>Transition metals</div><div>Poor metals</div><div>Other nonmetals</div><div>Noble gases</div></div><div>Nonmetals</div></div>																										
3 <b>Li</b> Lithium 6.941	4 <b>Be</b> Beryllium 9.012182											5 <b>B</b> Boron 10.811	6 <b>C</b> Carbon 12.0107	7 <b>N</b> Nitrogen 14.0067	8 <b>O</b> Oxygen 15.9994	9 <b>F</b> Fluorine 18.9984032	10 <b>Ne</b> Neon 20.1797									
11 <b>Na</b> Sodium 22.98976928	12 <b>Mg</b> Magnesium 24.3050											13 <b>Al</b> Aluminium 26.9815386	14 <b>Si</b> Silicon 28.0855	15 <b>P</b> Phosphorus 30.973762	16 <b>S</b> Sulfur 32.065	17 <b>Cl</b> Chlorine 35.453	18 <b>Ar</b> Argon 39.948									
19 <b>K</b> Potassium 39.0983	20 <b>Ca</b> Calcium 40.078	21 <b>Sc</b> Scandium 44.955912	22 <b>Ti</b> Titanium 47.867	23 <b>V</b> Vanadium 50.9415	24 <b>Cr</b> Chromium 51.9961	25 <b>Mn</b> Manganese 54.938045	26 <b>Fe</b> Iron 55.845	27 <b>Co</b> Cobalt 58.933195	28 <b>Ni</b> Nickel 58.6934	29 <b>Cu</b> Copper 63.546	30 <b>Zn</b> Zinc 65.38	31 <b>Ga</b> Gallium 69.723	32 <b>Ge</b> Germanium 72.64	33 <b>As</b> Arsenic 74.92160	34 <b>Se</b> Selenium 78.96	35 <b>Br</b> Bromine 79.904	36 <b>Kr</b> Krypton 83.798									
37 <b>Rb</b> Rubidium 85.4678	38 <b>Sr</b> Strontium 87.62	39 <b>Y</b> Yttrium 88.90585	40 <b>Zr</b> Zirconium 91.224	41 <b>Nb</b> Niobium 92.90638	42 <b>Mo</b> Molybdenum 95.96	43 <b>Tc</b> Technetium (97.9072)	44 <b>Ru</b> Ruthenium 101.07	45 <b>Rh</b> Rhodium 102.90550	46 <b>Pd</b> Palladium 106.42	47 <b>Ag</b> Silver 107.8682	48 <b>Cd</b> Cadmium 112.411	49 <b>In</b> Indium 114.818	50 <b>Sn</b> Tin 118.710	51 <b>Sb</b> Antimony 121.760	52 <b>Te</b> Tellurium 127.60	53 <b>I</b> Iodine 126.90447	54 <b>Xe</b> Xenon 131.293									
55 <b>Cs</b> Caesium 132.9054519	56 <b>Ba</b> Barium 137.327	57–71										72 <b>Hf</b> Hafnium 178.49	73 <b>Ta</b> Tantalum 180.94788	74 <b>W</b> Tungsten 183.84	75 <b>Re</b> Rhenium 186.207	76 <b>Os</b> Osmium 190.23	77 <b>Ir</b> Iridium 192.217	78 <b>Pt</b> Platinum 195.084	79 <b>Au</b> Gold 196.966569	80 <b>Hg</b> Mercury 200.59	81 <b>Tl</b> Thallium 204.3833	82 <b>Pb</b> Lead 207.2	83 <b>Bi</b> Bismuth 208.98040	84 <b>Po</b> Polonium (209.9824)	85 <b>At</b> Astatine (209.9871)	86 <b>Rn</b> Radon (222.0176)
87 <b>Fr</b> Francium (223)	88 <b>Ra</b> Radium (226)	89–103										104 <b>Rf</b> Rutherfordium (261)	105 <b>Db</b> Dubnium (262)	106 <b>Sg</b> Seaborgium (266)	107 <b>Bh</b> Bohrium (264)	108 <b>Hs</b> Hassium (277)	109 <b>Mt</b> Meitnerium (268)	110 <b>Ds</b> Darmstadtium (271)	111 <b>Rg</b> Roentgenium (272)	112 <b>Uub</b> Ununbium (285)	113 <b>Uut</b> Ununtrium (284)	114 <b>Uuq</b> Ununquadium (289)	115 <b>Uup</b> Ununpentium (288)	116 <b>Uuh</b> Ununhexium (292)	117 <b>Uus</b> Ununseptium (294)	118 <b>Uuo</b> Ununoctium (294)

For elements with no stable isotopes, the mass number of the isotope with the longest half-life is in parentheses.

Design and Interface Copyright © 1997 Michael Dayah (michael@dayah.com). <http://www.ptable.com/>

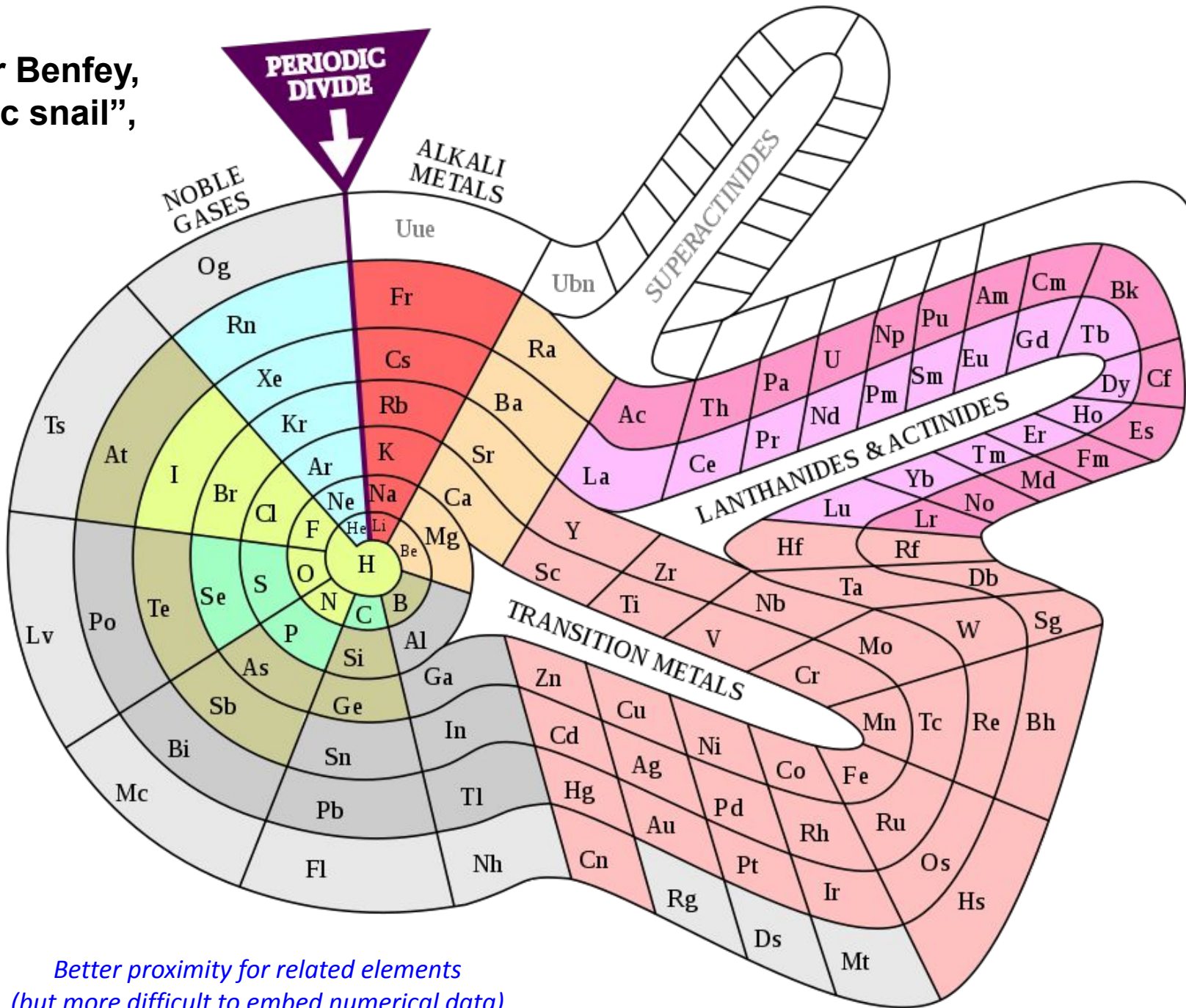


57 <b>La</b> Lanthanum 138.90547	58 <b>Ce</b> Cerium 140.116	59 <b>Pr</b> Praseodymium 140.90768	60 <b>Nd</b> Neodymium 144.242	61 <b>Pm</b> Promethium (145)	62 <b>Sm</b> Samarium 150.36	63 <b>Eu</b> Europium 151.964	64 <b>Gd</b> Gadolinium 157.25	65 <b>Tb</b> Terbium 158.92535	66 <b>Dy</b> Dysprosium 162.500	67 <b>Ho</b> Holmium 164.93032	68 <b>Er</b> Erbium 167.259	69 <b>Tm</b> Thulium 168.93421	70 <b>Yb</b> Ytterbium 173.054	71 <b>Lu</b> Lutetium 174.9668
89 <b>Ac</b> Actinium (227)	90 <b>Th</b> Thorium 232.03806	91 <b>Pa</b> Protactinium 231.03688	92 <b>U</b> Uranium 238.02891	93 <b>Np</b> Neptunium (237)	94 <b>Pu</b> Plutonium (244)	95 <b>Am</b> Americium (243)	96 <b>Cm</b> Curium (247)	97 <b>Bk</b> Berkelium (247)	98 <b>Cf</b> Californium (251)	99 <b>Es</b> Einsteinium (252)	100 <b>Fm</b> Fermium (257)	101 <b>Md</b> Mendelevium (258)	102 <b>No</b> Nobelium (259)	103 <b>Lr</b> Lawrencium (262)

<http://www.ptable.com/>



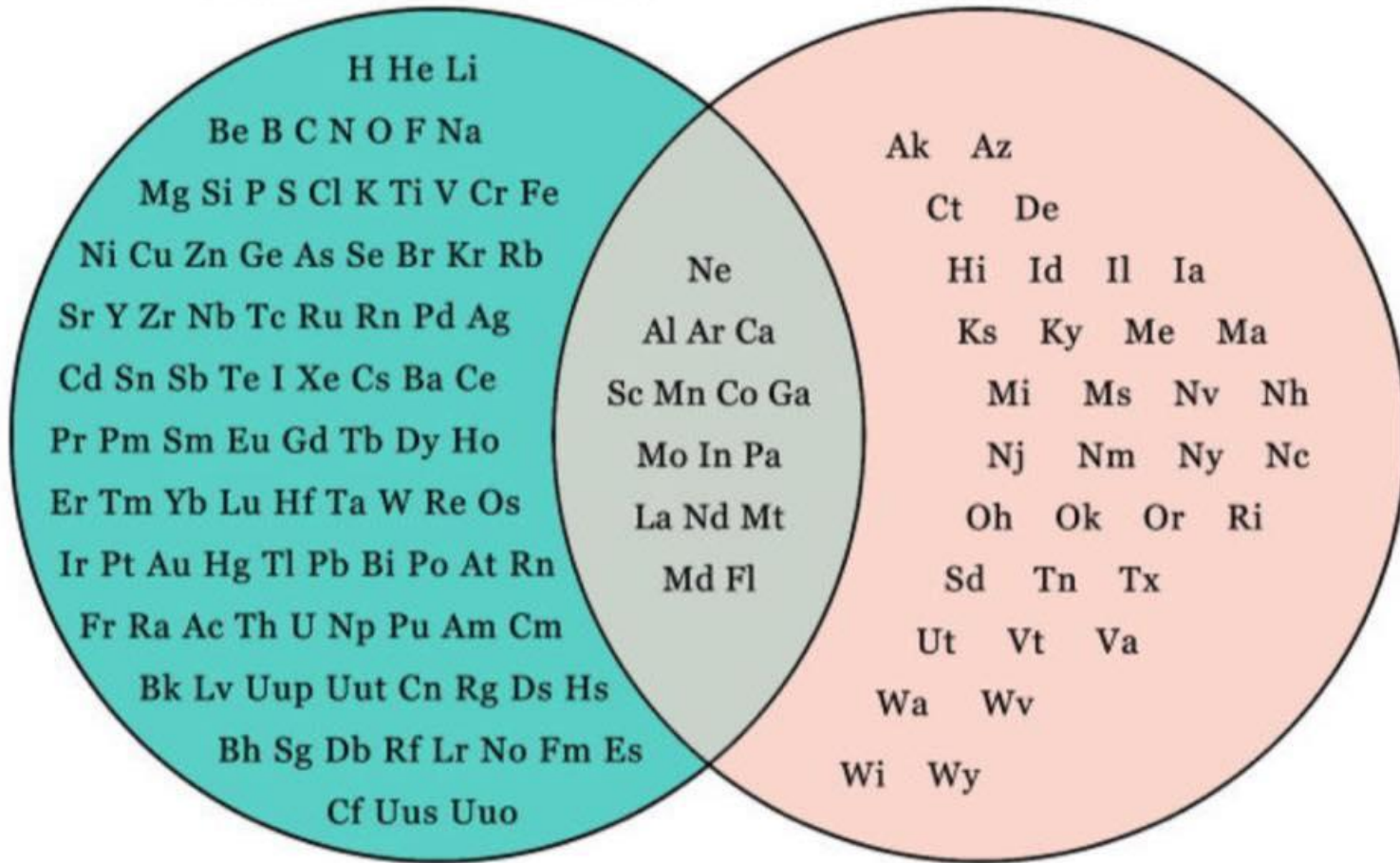
**Theodor Benfey,  
“periodic snail”,  
1964**



*Better proximity for related elements  
(but more difficult to embed numerical data)*

chemical elements

us states



chicanaspice

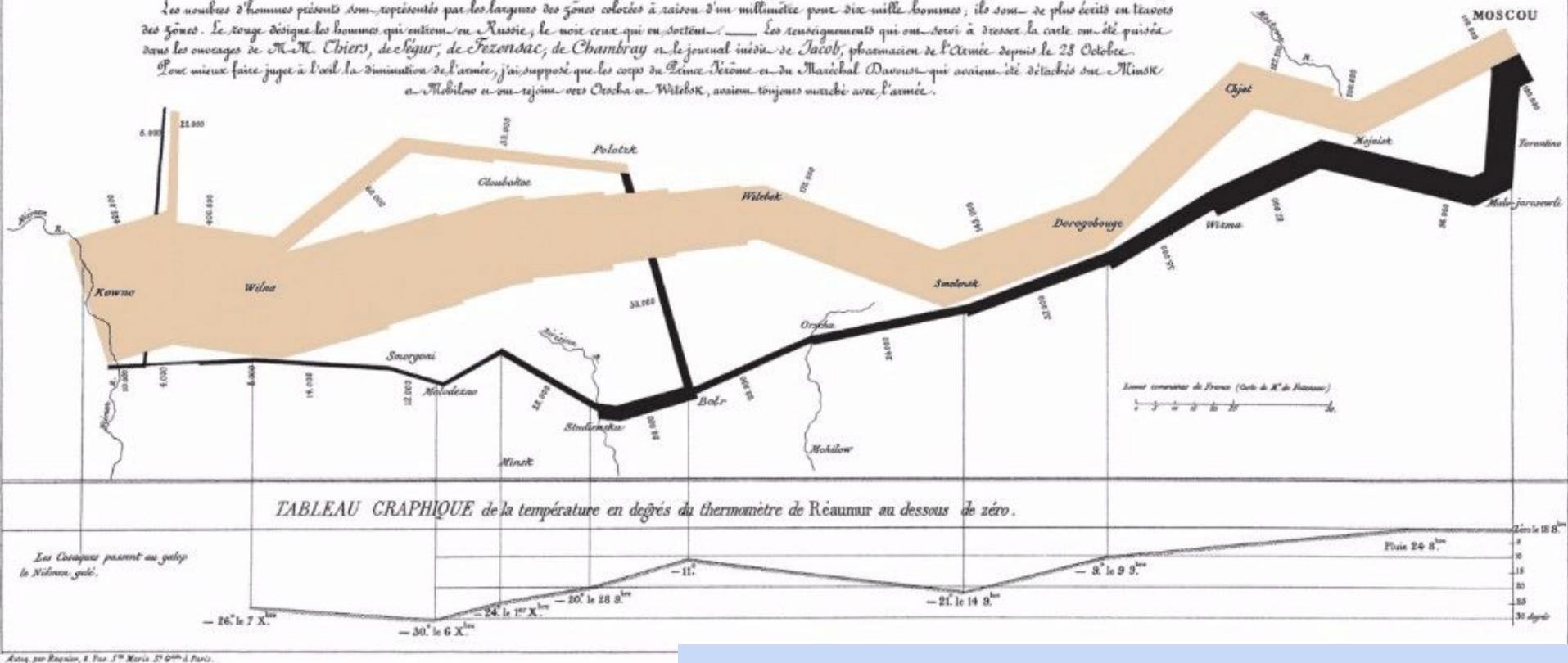
this information is so satisfying but idk what to do with it

# Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées et extraite de Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les longueurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps de Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilew et qui rejoins vers Orscha et Vitebsk, avaient toujours marché avec l'armée.



Edward Tufte says it "may well be the best statistical graphic ever drawn"

## Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812-1813.

Charles Joseph Minard, 1869.

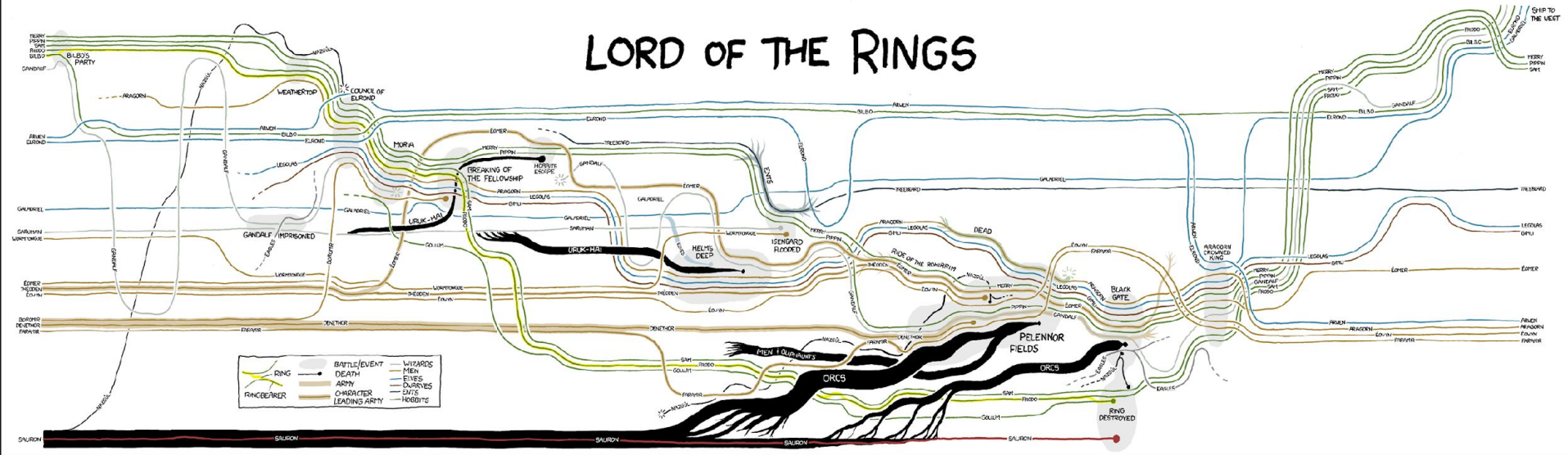
The numbers of men present are represented by the widths of the colored zones at a rate of one millimeter for every ten thousand men; they are further written across the zones. The red designates the men who enter Russia, the black those who leave it. — The information which has served to draw up the map has been extracted from the works of M.M. Thiers, de Ségur, de Fezensac, de Chambray and the unpublished diary of Jacob, the pharmacist of the Army since October 28th.

In order to better judge with the eye the diminution of the army, I have assumed that the troops of Prince Jérôme and of Marshal Davoust, who had been detached at Minsk and Mogilev and have rejoined near Orsha and Vitebsk, had always marched with the army.



THESE CHARTS SHOW MOVIE CHARACTER INTERACTIONS.  
THE HORIZONTAL AXIS IS TIME. THE VERTICAL GROUPING OF THE  
LINES INDICATES WHICH CHARACTERS ARE TOGETHER AT A GIVEN TIME.

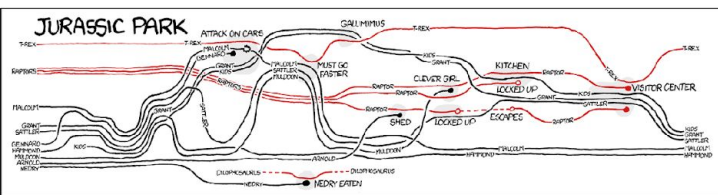
## LORD OF THE RINGS



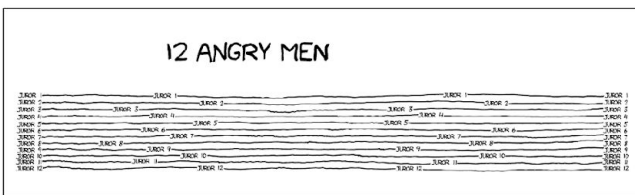
## STAR WARS (ORIGINAL TRILOGY)



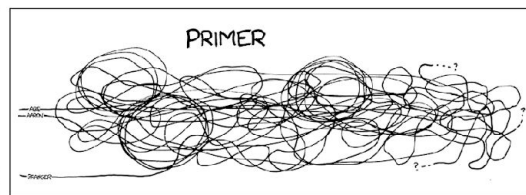
## JURASSIC PARK



## 12 ANGRY MEN



## PRIMER



# Today

- “Good” Design (30 min)
- Today’s Readings:
  - How To
    - “Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message”, Stephen Few, Intelligent Enterprise, 2004”
    - “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Today

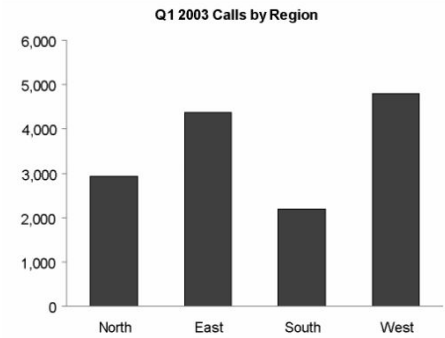
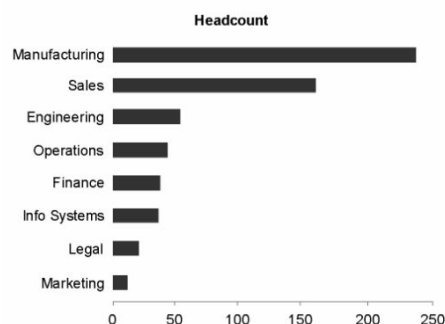
- “Good” Design (30 min)
- Today’s Readings:
  - How To
  - “Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message”, Stephen Few, Intelligent Enterprise, 2004”
  - “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Today

- “Good” Design (30 min)
- Today’s Readings:
  - How To
  - “Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message”, Stephen Few, Intelligent Enterprise, 2004”
  - “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

"Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message", Stephen Few, Intelligent Enterprise, 2004

- Learn conventions/patterns, applies to almost all business data (not necessarily all scientific data)
- Avoid viewer confusion / mis-information
- Definition: Categorical vs. quantitative
- Definition: Nominal (order free), ordinal (ordered), interval (e.g. histogram)
- Available visual attributes for encoding data: location, size, shape, orientation, color  
*choose wisely!!*

Type/Description	Encoding Methods	Example																																							
<b>Nominal Comparison</b> A simple comparison of the categorical subdivisions of one or more measures in no particular order	<ul style="list-style-type: none"><li>Bars only (horizontal or vertical)</li></ul>	 <p>Q1 2003 Calls by Region</p> <table><thead><tr><th>Region</th><th>Calls</th></tr></thead><tbody><tr><td>North</td><td>3,000</td></tr><tr><td>East</td><td>4,500</td></tr><tr><td>South</td><td>2,200</td></tr><tr><td>West</td><td>4,800</td></tr></tbody></table>	Region	Calls	North	3,000	East	4,500	South	2,200	West	4,800																													
Region	Calls																																								
North	3,000																																								
East	4,500																																								
South	2,200																																								
West	4,800																																								
<b>Time Series</b> Multiple instances of one or more measures taken at equidistant points in time	<ul style="list-style-type: none"><li>Lines to emphasize overall pattern</li><li>Bars to emphasize individual values</li><li>Points connected by lines to slightly emphasize individual values while still highlighting the overall pattern</li><li>Always place time on the horizontal axis</li></ul>	 <p>2003 Sales</p> <table><thead><tr><th>Month</th><th>Category 1 Sales</th><th>Category 2 Sales</th></tr></thead><tbody><tr><td>Jan</td><td>2,200</td><td>1,000</td></tr><tr><td>Feb</td><td>2,500</td><td>1,000</td></tr><tr><td>Mar</td><td>2,800</td><td>1,000</td></tr><tr><td>Apr</td><td>2,500</td><td>1,000</td></tr><tr><td>May</td><td>3,000</td><td>1,000</td></tr><tr><td>Jun</td><td>3,000</td><td>1,000</td></tr><tr><td>Jul</td><td>2,500</td><td>1,000</td></tr><tr><td>Aug</td><td>2,500</td><td>500</td></tr><tr><td>Sep</td><td>3,000</td><td>1,000</td></tr><tr><td>Oct</td><td>2,800</td><td>1,000</td></tr><tr><td>Nov</td><td>3,200</td><td>1,000</td></tr><tr><td>Dec</td><td>3,800</td><td>1,000</td></tr></tbody></table>	Month	Category 1 Sales	Category 2 Sales	Jan	2,200	1,000	Feb	2,500	1,000	Mar	2,800	1,000	Apr	2,500	1,000	May	3,000	1,000	Jun	3,000	1,000	Jul	2,500	1,000	Aug	2,500	500	Sep	3,000	1,000	Oct	2,800	1,000	Nov	3,200	1,000	Dec	3,800	1,000
Month	Category 1 Sales	Category 2 Sales																																							
Jan	2,200	1,000																																							
Feb	2,500	1,000																																							
Mar	2,800	1,000																																							
Apr	2,500	1,000																																							
May	3,000	1,000																																							
Jun	3,000	1,000																																							
Jul	2,500	1,000																																							
Aug	2,500	500																																							
Sep	3,000	1,000																																							
Oct	2,800	1,000																																							
Nov	3,200	1,000																																							
Dec	3,800	1,000																																							
<b>Ranking</b> Categorical subdivisions of a measure ordered by size (either descending or ascending)	<ul style="list-style-type: none"><li>Bars only (horizontal or vertical)</li><li>To highlight high values, sort in descending order</li><li>To highlight low values, sort in ascending order</li></ul>	 <p>Headcount</p> <table><thead><tr><th>Department</th><th>Headcount</th></tr></thead><tbody><tr><td>Manufacturing</td><td>230</td></tr><tr><td>Sales</td><td>160</td></tr><tr><td>Engineering</td><td>50</td></tr><tr><td>Operations</td><td>45</td></tr><tr><td>Finance</td><td>40</td></tr><tr><td>Info Systems</td><td>35</td></tr><tr><td>Legal</td><td>15</td></tr><tr><td>Marketing</td><td>10</td></tr></tbody></table>	Department	Headcount	Manufacturing	230	Sales	160	Engineering	50	Operations	45	Finance	40	Info Systems	35	Legal	15	Marketing	10																					
Department	Headcount																																								
Manufacturing	230																																								
Sales	160																																								
Engineering	50																																								
Operations	45																																								
Finance	40																																								
Info Systems	35																																								
Legal	15																																								
Marketing	10																																								
<b>Part-to-Whole</b> Measures of individual categorical subdivisions as ratios to the whole	<ul style="list-style-type: none"><li>Bars only (horizontal or vertical)</li><li>Use stacked bars only when you must display measures of the whole as well as the parts</li></ul>	 <p>Regional % of Total Expenses</p> <table><thead><tr><th>Region</th><th>% of Total Expenses</th></tr></thead><tbody><tr><td>West</td><td>34%</td></tr><tr><td>East</td><td>31%</td></tr><tr><td>North</td><td>21%</td></tr><tr><td>South</td><td>16%</td></tr></tbody></table>	Region	% of Total Expenses	West	34%	East	31%	North	21%	South	16%																													
Region	% of Total Expenses																																								
West	34%																																								
East	31%																																								
North	21%																																								
South	16%																																								

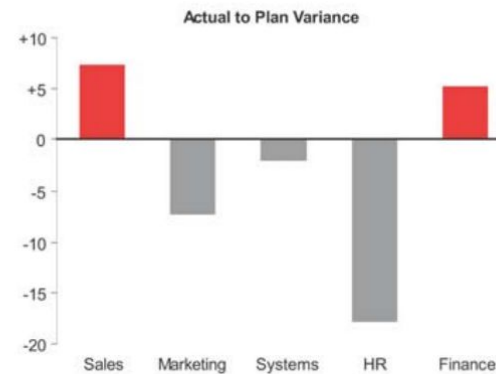
"Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message",  
Stephen Few,  
Intelligent Enterprise,  
2004



## Deviation

Categorical subdivisions of a measure compared to a reference measure, expressed as the differences between them

- Lines to emphasize the overall pattern only when displaying deviation and time-series relationships together
- Points connected by lines to slightly emphasize individual data points while also highlighting the overall pattern when displaying deviation and time-series relationships together
- Bars to emphasize individual values, but limit to vertical bars when a time-series relationship is included
- Always include a reference line to compare the measures of deviation against

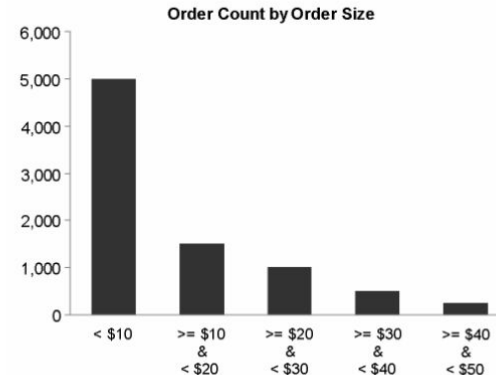


"Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message", Stephen Few, Intelligent Enterprise, 2004

## Frequency Distribution

Counts of something per categorical subdivisions (intervals) of a quantitative range

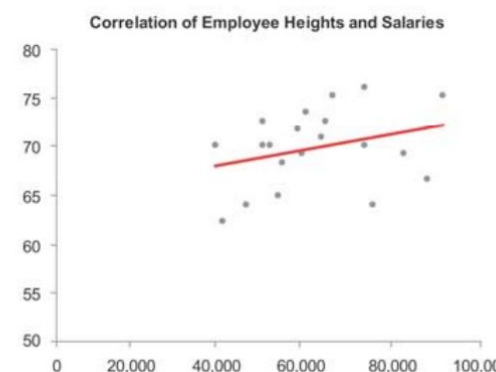
- Vertical bars to emphasize individual values (called a *histogram*)
- Lines to emphasize the overall pattern (called a *frequency polygon*)



## Correlation

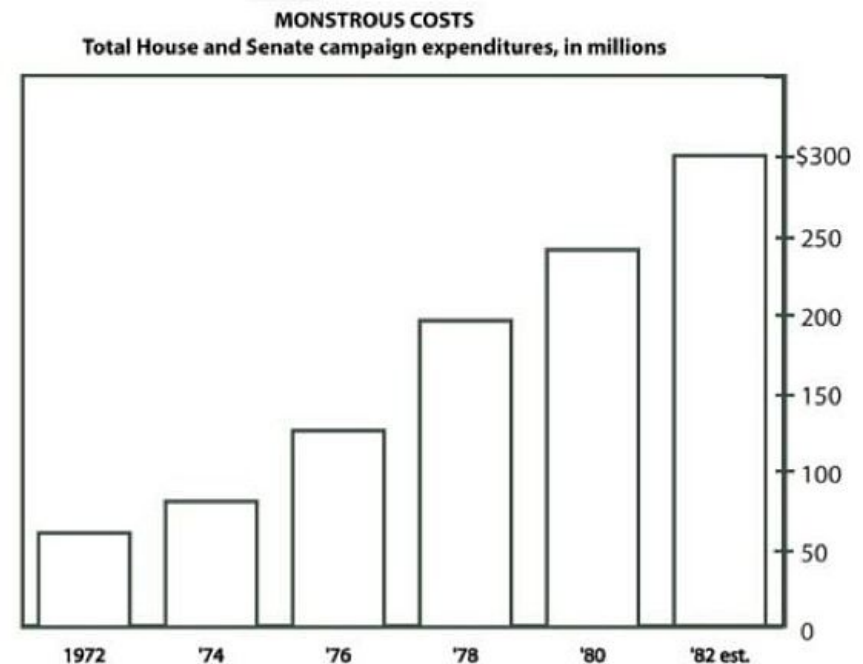
Comparisons of two paired sets of measures to determine if as one set goes up the other set goes either up or down in a corresponding manner, and if so, how strongly

- Points and a trend line in the form of a scatter plot
- Bars may be used, arranged as a *paired bar graph* or a *correlation bar graph*, if scatter plots are unfamiliar
- (Note: For descriptions of these graphs, see my book *Show Me the Numbers*.)



# Reading for Today

- “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”  
Bateman et al., CHI 2010.

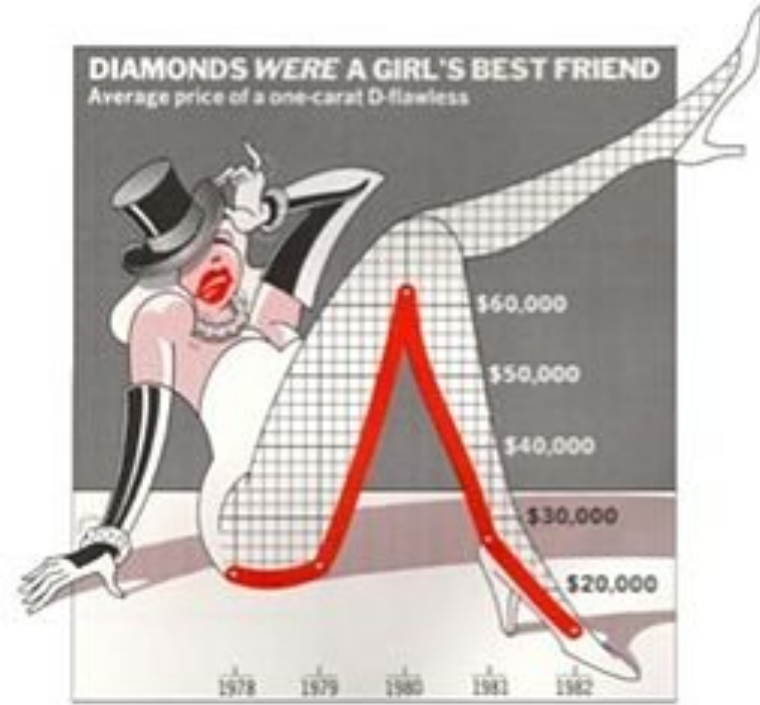


- Article discussed here:  
<http://eagereyes.org/criticism/chart-junk-considered-useful-after-all>



# What is “Chart Junk”?

- Extraneous elements in a chart or visualization
- Does not represent data
- Data-to-ink ratio (aim to convey more data with less ink)
- According to Edward Tufte:  
It's not just unnecessary, it's harmful (distracting)
- According to Nigel Holmes:  
Visualization should engage the reader's interest



Nigel Holmes

# Study Design

“Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts” Bateman, Mandryk, Gutwin, Genest, McDine, & Brooks, CHI 2010

- Compare embellished charts to plain ones
- Measured:
  - interpretation accuracy  
*was no worse for embellished charts*
  - long-term recall (2-3 weeks later)  
*was better for embellished charts,  
topic & details of the chart were more memorable*
- Prior work:
  - Higher data-to-ink → faster response & greater accuracy [Gilan & Richman]
  - Other work shows a somewhat weak correlation between data-to-ink and interpretability or aesthetics
- Author’s caution:
  - Not an endorsement of chart junk
  - Embellishments can lead to bias!

- Relatively small sample pool
  - 10 tested with ~5 min recall
  - 10 tested with 2-3 week recall
- Found no difference in time to read & describe embellished vs. non-embellished
- Participants preferred the embellished charts and found them more attractive

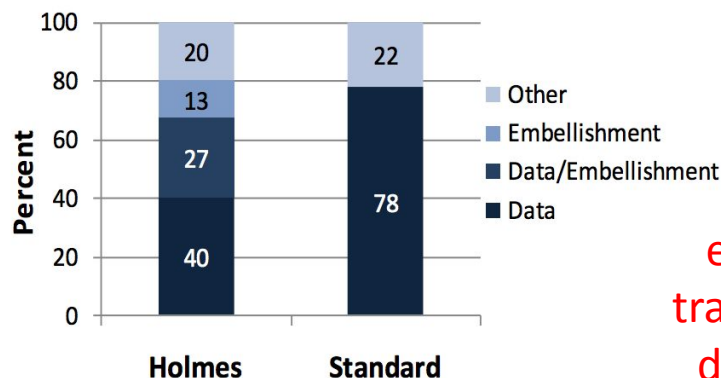


Figure 9. Percentage of on-screen time spent looking at different chart elements for Holmes and Plain charts.

eye  
tracker  
data

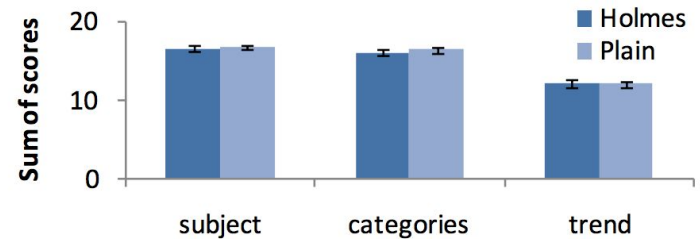


Figure 4. Means  $\pm$  SE for description scores.

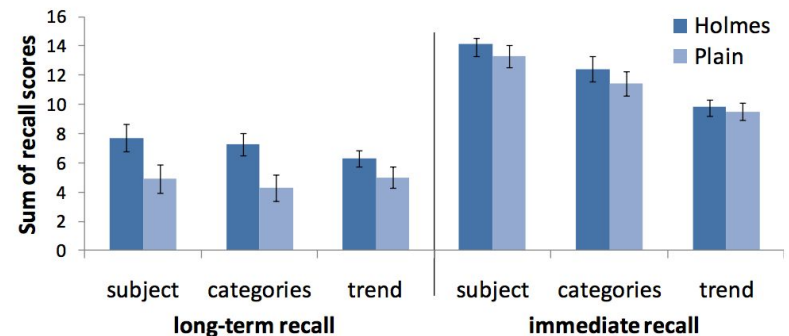


Figure 5. Means  $\pm$  SE for recall scores for long-term and immediate recall.

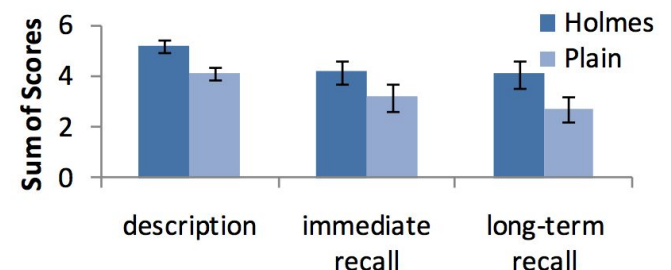


Figure 6. Means  $\pm$  SE for sum of value message scores.

- Viewing time was unlimited for this study
  - Participants ended up spending the ~same amount of time on embellished vs. non-embellished
  - Effect of limiting time *not measured*
- Chart junk for these examples was tightly coupled with subject & details of chart
  - Quote from Holmes: “I think [Tufte] missed the point of much that I was trying to do: TIME magazine charts were aimed at lay readers, not unintelligent ones, but busy ones. I knew they’d get the point quicker if they were somehow attracted to the graphic.”
- What about charts from paper on last slide? What was their point? How good is your recall? Will your recall them in 2-3 weeks? Why didn’t the authors use embellishment?

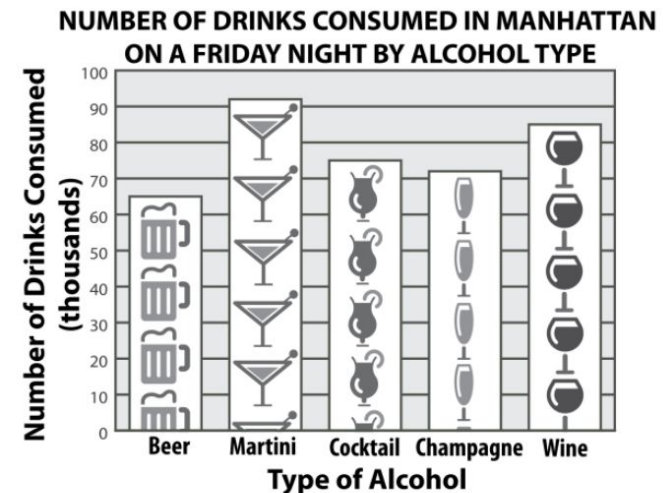


Figure 10. Less extreme visual imagery in charts.





[http://nigelholmes.com/  
graphic/data-dump/](http://nigelholmes.com/graphic/data-dump/)

# Homework Assignment 2:

## Time-Based Datasets

- Team of 2 or 3
  - Obtain an interesting time-based dataset
    - Should be collectable\* from online sources, and
    - Require a modest effort to prepare\*
- \* = you'll submit your scripts/code to document*
- Use Excel or Google Sheets
    - Create a variety (one of each?!) of the charts following the guidelines from "Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message"
    - Excellent labels and captions for each.
  - Upload your assignment to Submitty by **11:59pm on Thursday**. And post one of the images on the forum...

# Pair Worksheet (~20 minutes)

- Meet new people
  - *Work with someone you did not know before this class*
  - *Work with a different partner every time*
- 1 worksheet per team of 3
  - *Sketch & brainstorm on the page*
  - *Submit on Submittity*
- Use color! Be creative!

# Tools for scraping data from the web

- copy-paste
- wget
- grep / sed / awk / sort / uniq
- Favorite programming language to parse/strip out unnecessary html formatting
- Save as .csv (comma separated value) files to upload to Excel / Google Sheets
- Python has lots of packages for parsing (e.g., json format)
- Selenium for automated browsing of websites



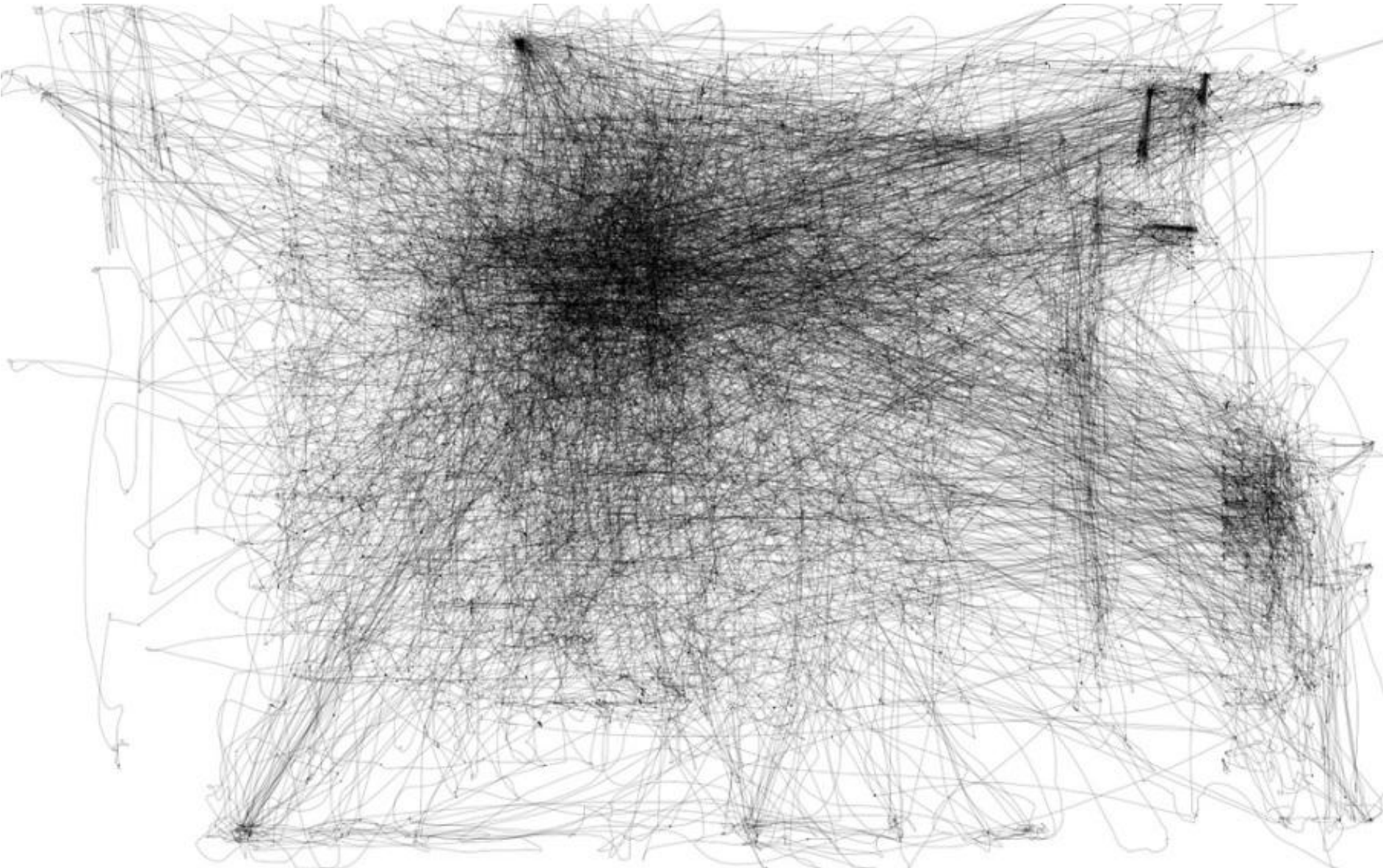
# Today

- “Good” Design (30 min)
  - Photography tips
  - Principles of Effective Website Design
  - Principles of Good User Interface Design
  - Examples of Good (Bad) Visualization Design
- Today’s Reading (30 min): “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing



# Graph Terminology I

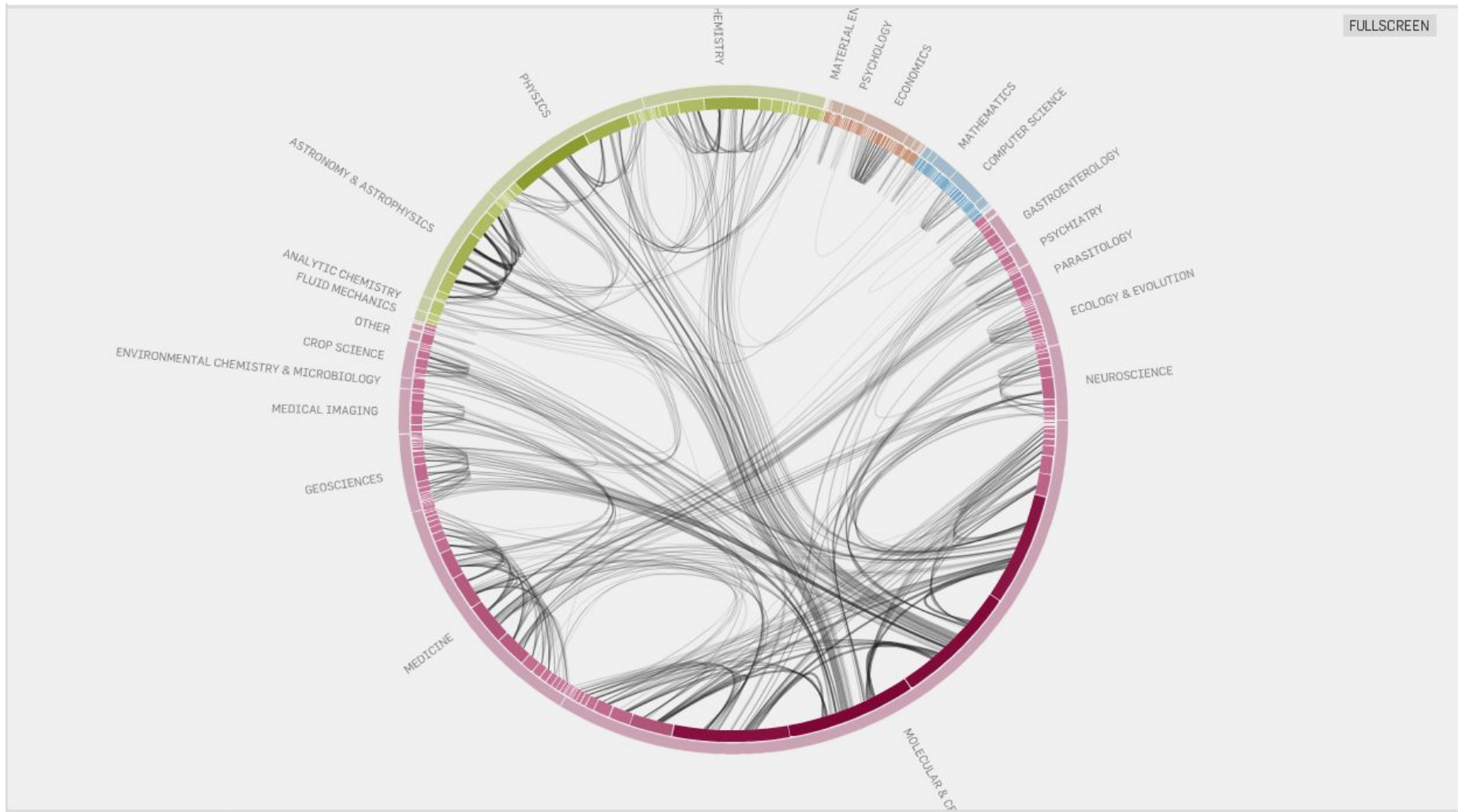
- Directed / Undirected Edges
- Tree (no cycles) vs. Graph (cycles allowed)
  - Cycle: A path along edges through the graph where only starting & ending vertices are repeated.
  - Walk: A sequence of vertices and edges
    - closed walk, simple cycle, directed cycle, ...
- Valence (a.k.a. Degree) of a Vertex:  
# of edges incident on the vertex
- Regular: Each vertex has same valence,  
a 3-regular graph is also called cubic





# Graph Terminology II

- Polygon: 2D flat or on a sphere, with straight or great circle edges
- Polyhedron: 3D solid formed by flat faces
- Polytope: flat sides in any dimension
- Bipartite: vertices can be split into two groups, A & B. No edge connects a vertex in A to another vertex in A. Same for B.
- Clique – subset of vertices in an undirected graph with an edge connecting every pair of vertices in the subset.



Stefaner, Moritz. "Citation Patterns." *Wellformed Eigenfactor*. N.p., 2008. Web.

# Graph Terminology III

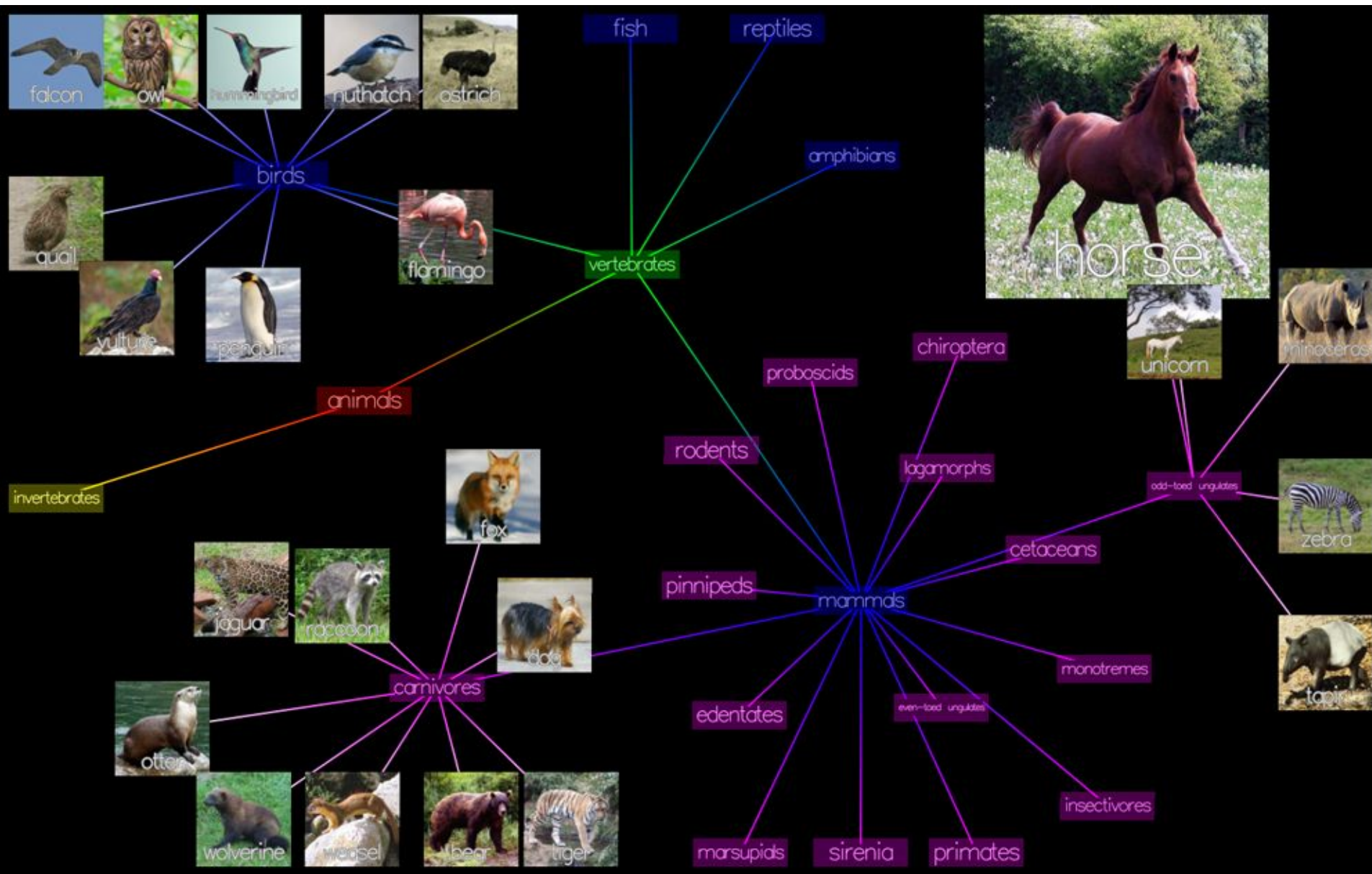
- Upward Drawing (of a tree) – no child is drawn with vertically above (higher  $y$  value than) its parent.
- Plane Graph – A 2D drawing of the graph where no edges cross (touching at the endpoint vertices they share is ok)
- Planar Graph – A graph for which a Plane Graph exists.
- Euler's Theorem for planar graphs:  
For a plane graph with  $n$  vertices,  $m$  edges and  $f$  faces, we have  $n - m + f = 2$ .





# Graph Drawing Goals

- Automated!
- Can read all of the labels
- Can follow the line and see exactly which 2 vertices it connects
- Aesthetically pleasing
- Layout should display as much symmetry as possible
- Crossing free or minimal-crossing layout
- All edge lengths are approximately equal
- Even vertex distribution
- Distance between nodes in final layout should be as close as possible to “graph distance” (# of edges on shortest path between those nodes)



# Graph Drawing Questions

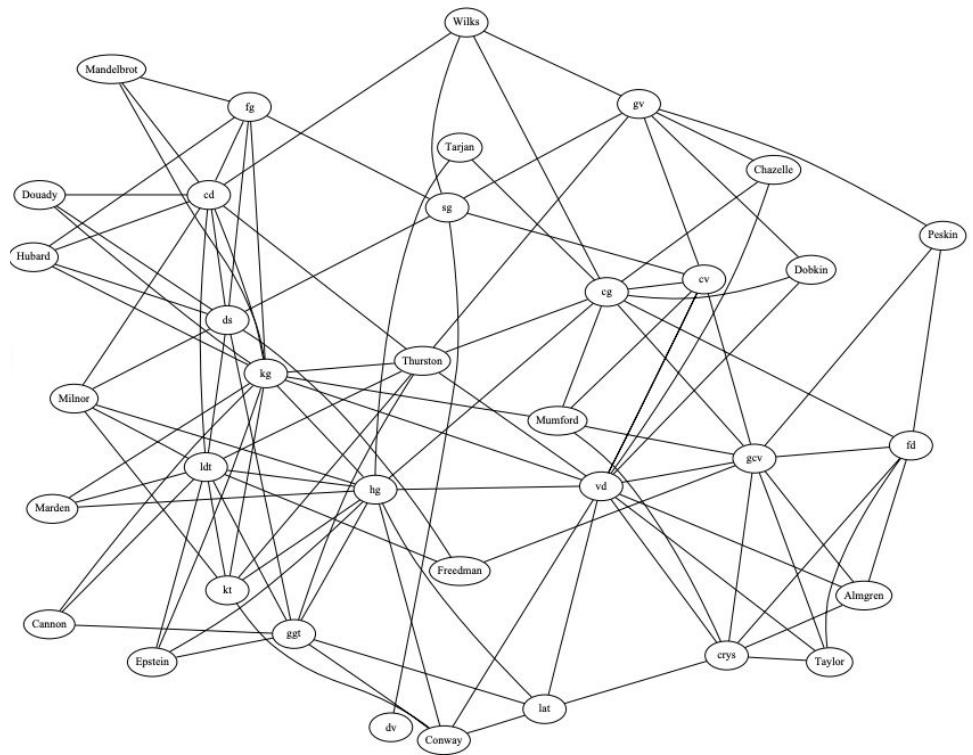
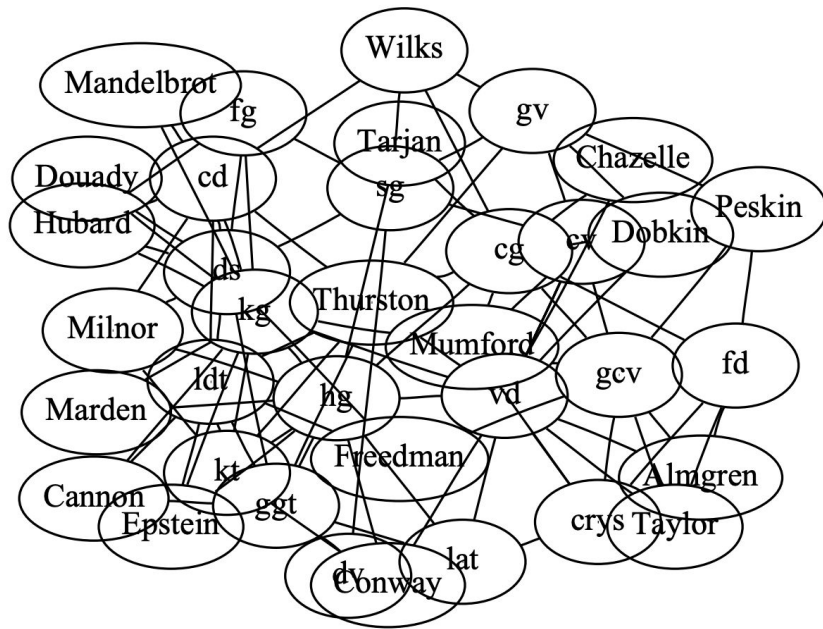
- What is the metric of success for each of our goals?
- Can we guarantee to find a solution? The optimal or best solution?
- Can we use randomness? Does it help?
- How expensive/slow are the different algorithms to draw graphs?
- How does it scale with more nodes/edges?
  - Does it lose effectiveness in meeting our goals?
  - How is the running time affected?
- How do we label the nodes/edges with color/words/images?

# Today

- “Good” Design (30 min)
  - Photography tips
  - Principles of Effective Website Design
  - Principles of Good User Interface Design
  - Examples of Good (Bad) Visualization Design
- Today’s Reading (30 min): “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts”
- Worksheet: Time-based Data & Simple Charts (20 min)
- Graph Drawing Preview: Terminology & Goals (15 min)
- Readings for Tuesday: Graph Drawing

# Reading for Tuesday *(pick one)*

“Improved force-directed layouts”, Gansner and North, Graph Drawing, 1999.





# Reading for Tuesday *(pick one)*

“A Technique for  
Drawing Directed  
Graphs” Gansner,  
Koutsofios, North, &  
Vo, IEEE Trans. on  
Software Engineering,  
1993.

