



Overview on Information Visualization

BAIS 6140 – Information Visualization

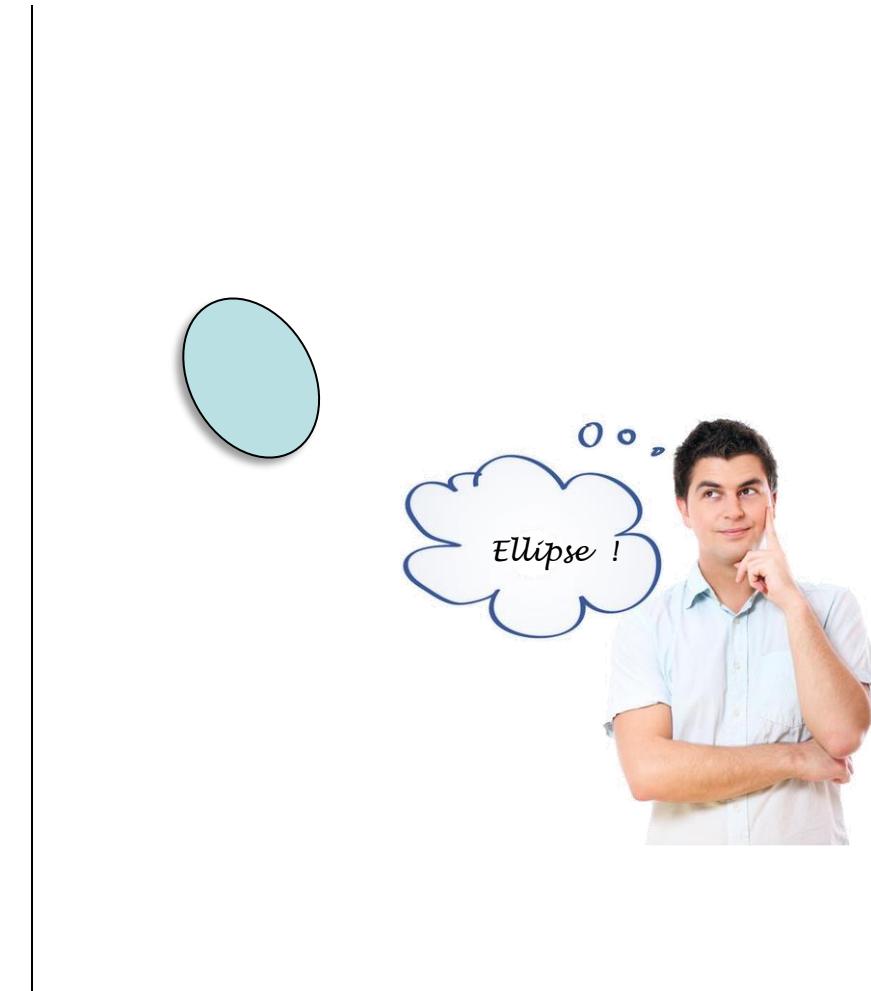
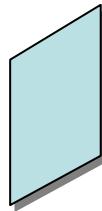
L. Miguel Encarnação

This class is about Truth!

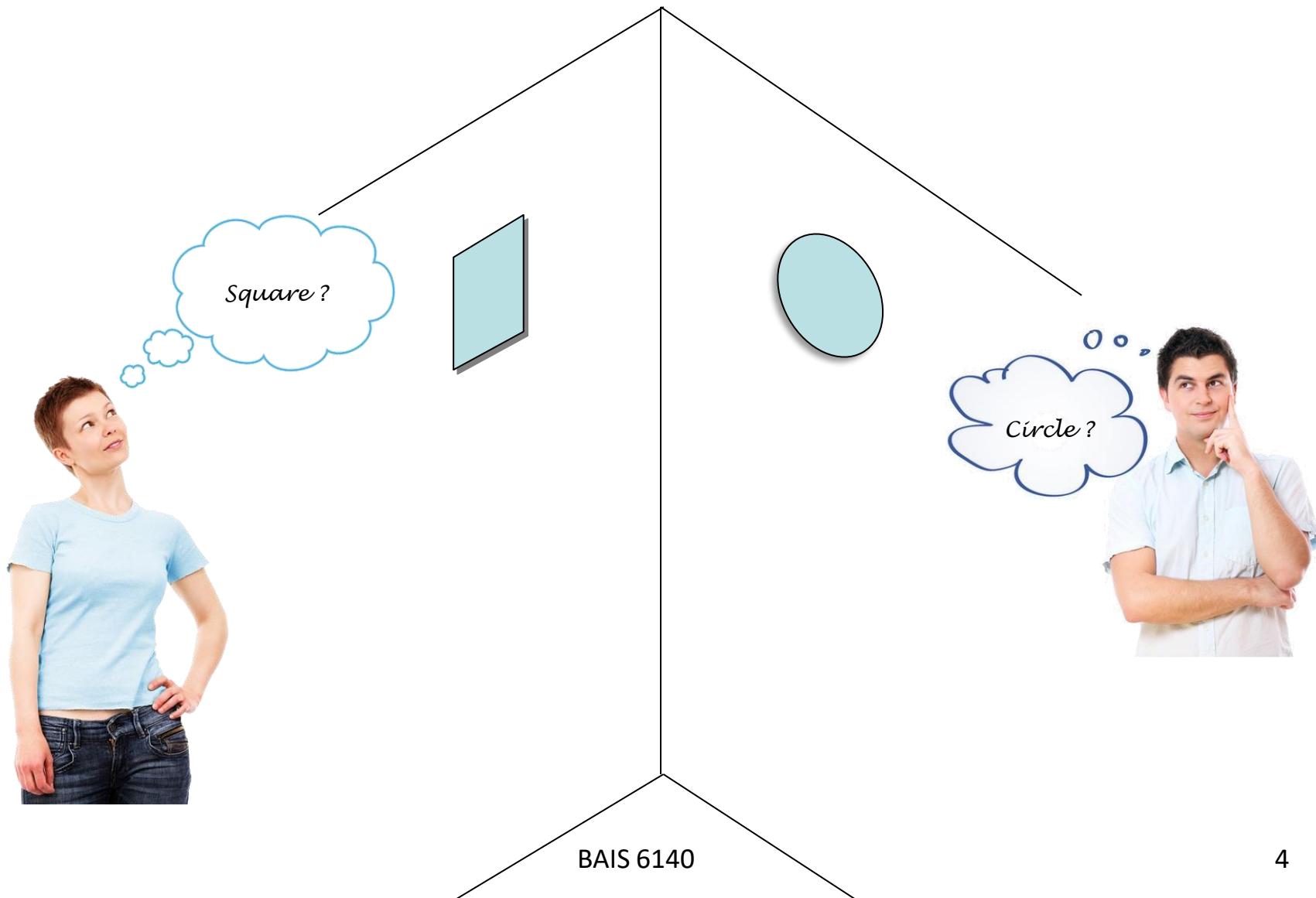


Rudy Giuliani, August 19, 2018

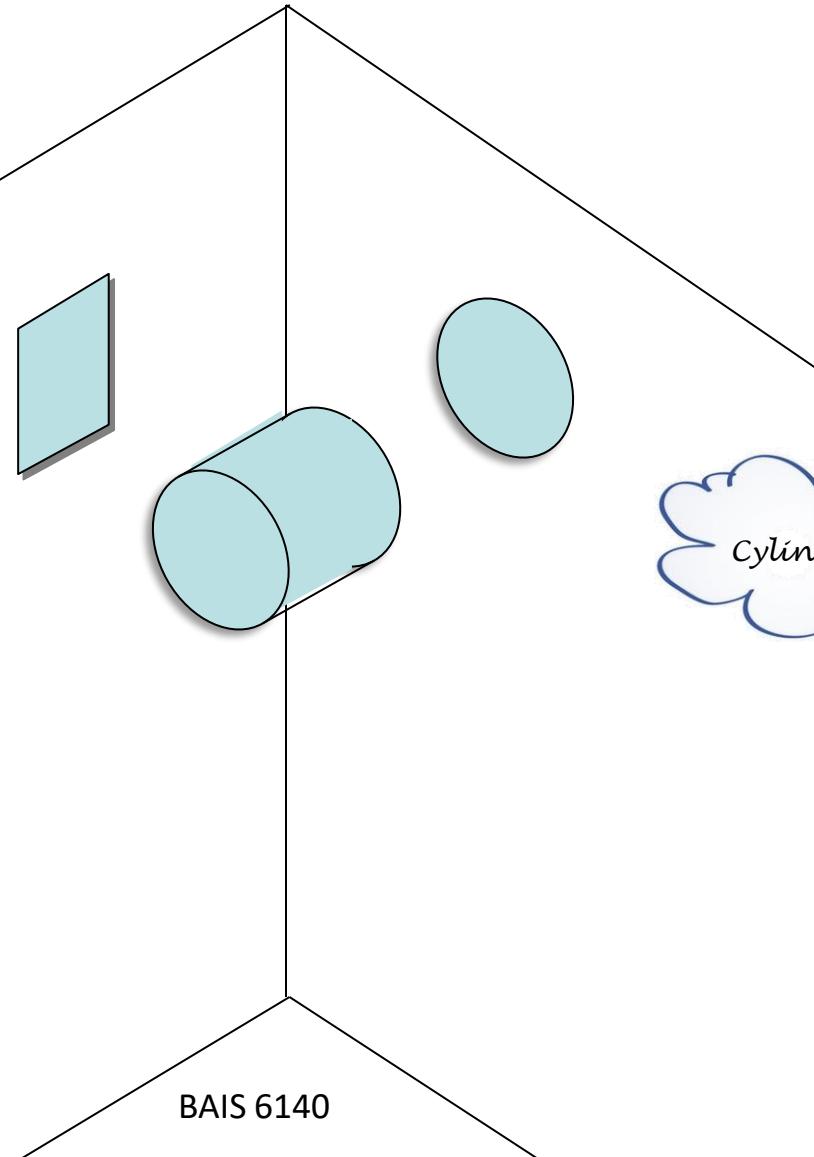
What IS true?



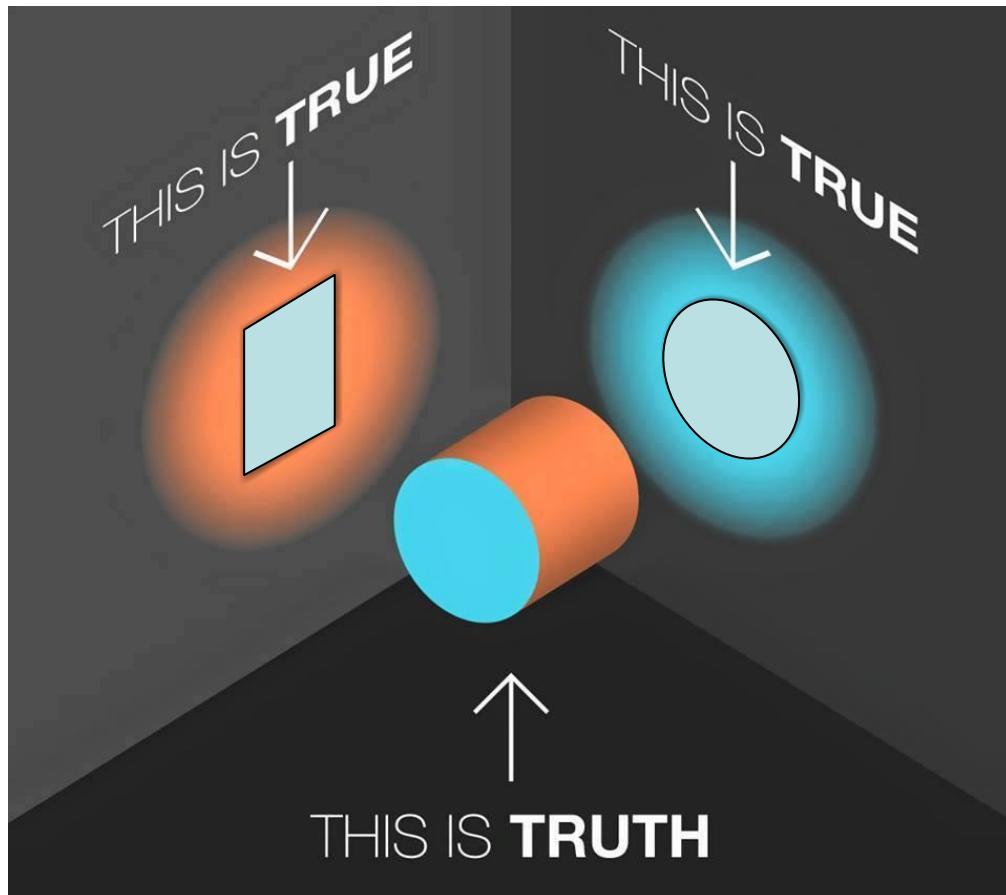
What IS true?



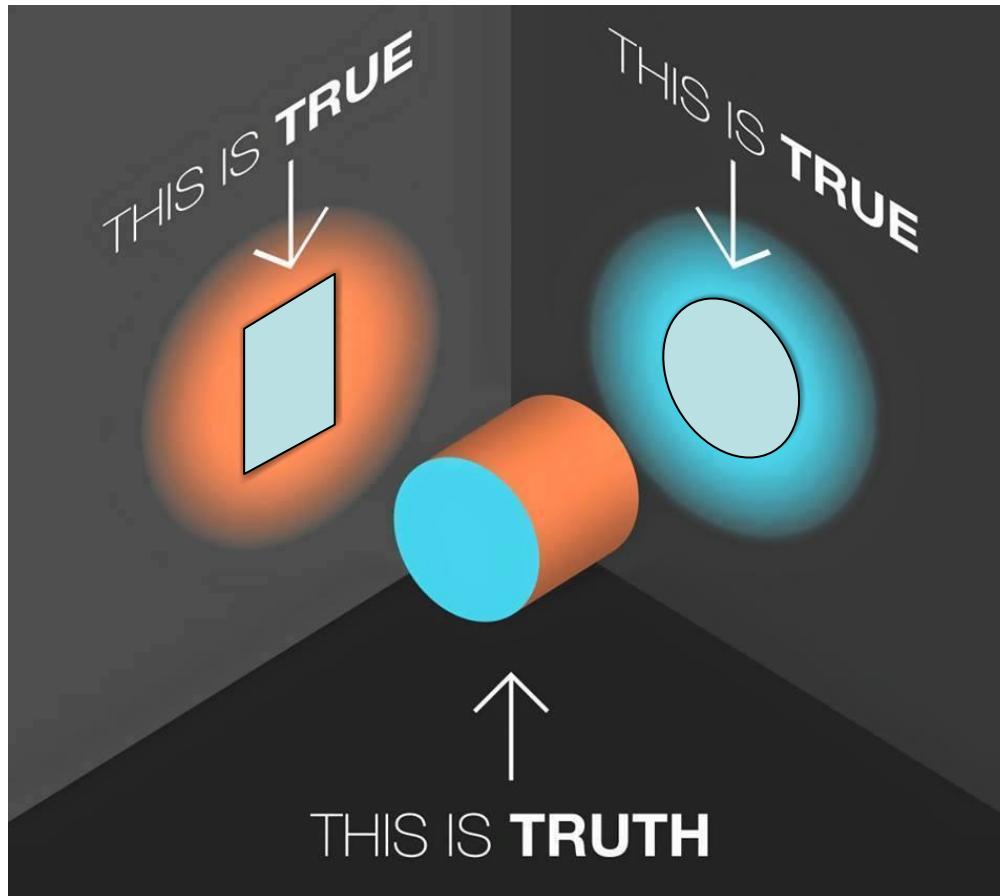
What IS true?



True vs. Truth



Information Visualization



Maximizing the likelihood to find consensus on truth for better decision-making.

Truth more important than ever

- Data Explosion
 - Society is more complex
 - There simply is more “stuff”
 - Computers, internet and web give people access to an incredible amount of data
 - news, sports, financial, purchases, etc...

So much data – per minute

2018 *This Is What Happens In An Internet Minute*



<https://www.weforum.org/agenda/2018/05/what-happens-in-an-internet-minute-in-2018/>, 2018

Created By:
@LoriLewis
@OfficiallyChadd

THINGS THAT HAPPEN EVERY **60** SECONDS

https://youtu.be/WWoLF_TiWcs?t=6

Data Overload

- Challenge: How to make use of the data
 - How do we make sense of the data?
 - How do we harness this data in decision-making processes?
 - How do we avoid being overwhelmed?

Why else you should care

- Data Analytics as 21st century skill

Why you should care – even more

- Data Analytics skills industry shortage

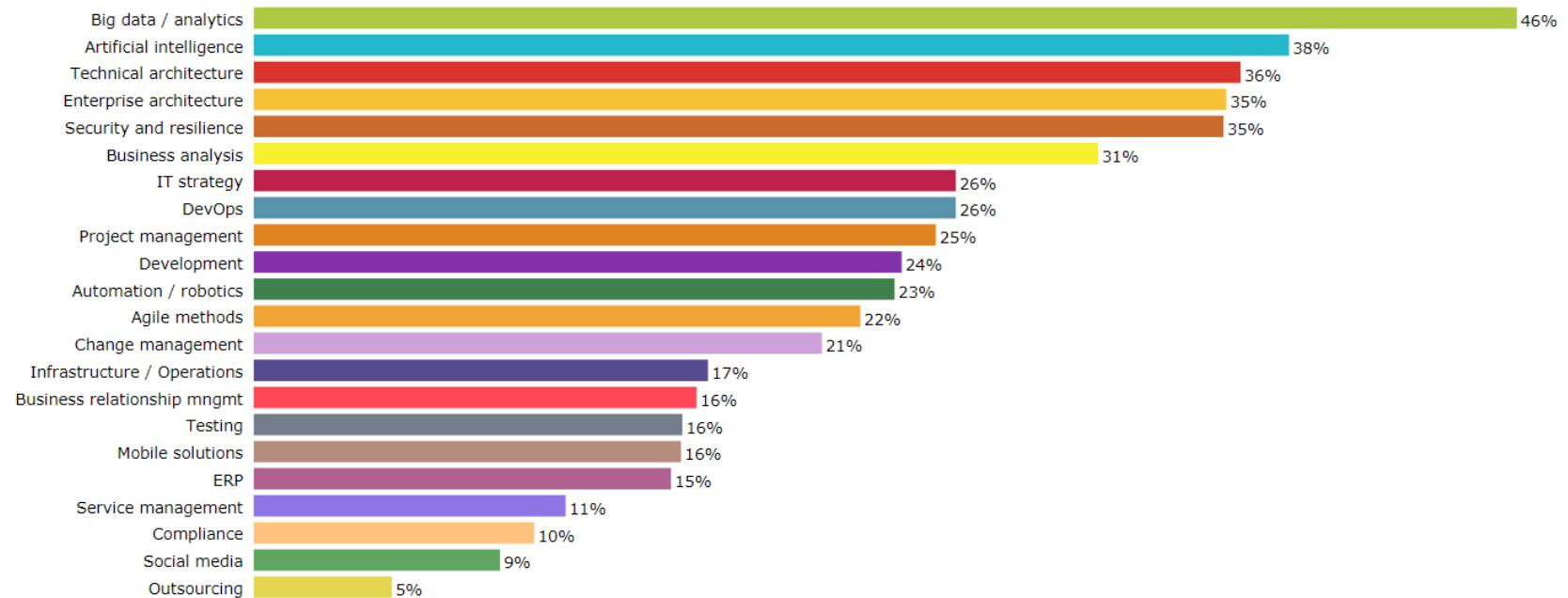
“As the use of analytics grows quickly, companies will need employees who understand the data. A May study from McKinsey & Co. found that **by 2018, the U.S. will face a shortage of 1.5 million managers who can use data to shape business decisions.**”

Why you should care – even more

- Data Science skills industry shortage

Which functions do you feel suffer from a skills shortage?

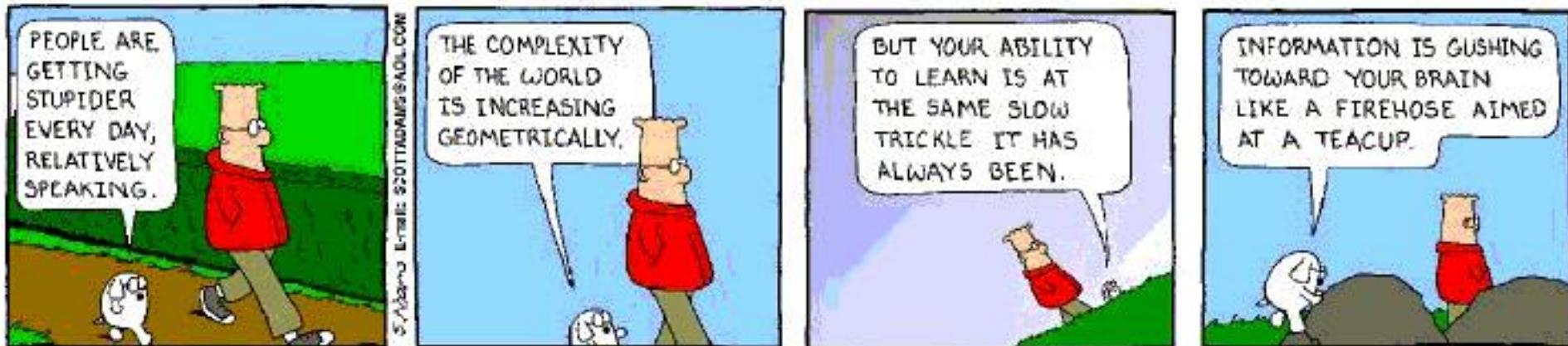
Select all that apply



©2019 KPMG CIO report

First step: Deal with all the Data

- Challenge: How to make use of the data
 - How do we make sense of the data?
 - How do we harness this data in decision-making processes?
 - How do we avoid being overwhelmed?

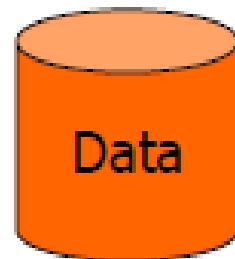


The Strategy

- Transform the data into information (understanding, insight) thus making it useful to people

The Problem

Web,
Books,
Papers,
Game scores,
Scientific data,
Biotech,
Shopping
People
Stock/finance
News



Data Transfer →



How?

Vision: 100 MB/s

Ears: <100 b/s

Telepathy

Haptic/tactile

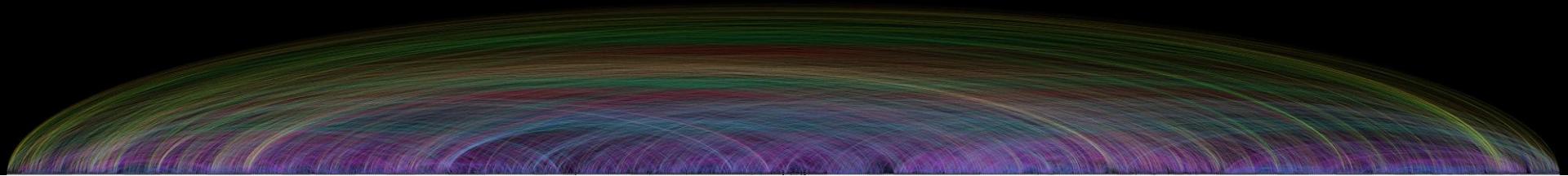
Smell

Taste

Two slides courtesy
of Chris North

Human Vision to the rescue

- Highest bandwidth sense
- Fast, parallel
- Pattern recognition
- Pre-attentive
- Extends memory and cognitive capacity
 - (Multiplication test)
- People think visually
- Cool: LET'S USE IT!



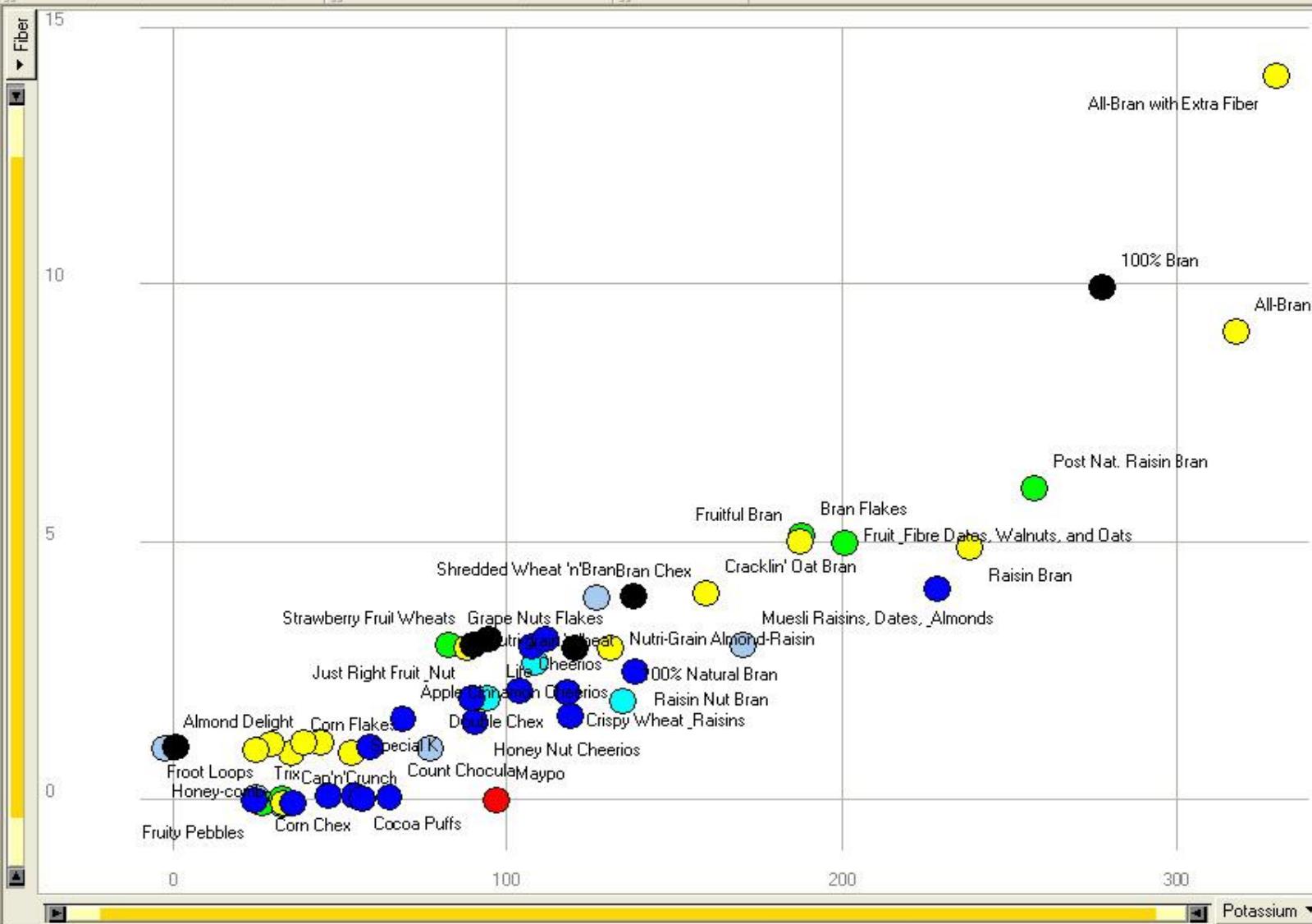
EXAMPLES

Questions

- Which cereal has the most/least potassium?
- Is there a relationship between potassium and fiber?
- If so, are there any outliers?
- Which manufacturer makes the healthiest cereals?

	A	B	C	D	
1	Cereal	Manufacturer	Fiber	Potassium	
2	100% Bran	N	10	280	
3	100% Natural Bran	Q	2	135	
4	All-Bran	K	9	320	
5	All-Bran with Extra Fiber	K	14	330	
6	Almond Delight	R	1	0	
7	Apple Cinnamon Cheerios	G	1.5	70	
8	Bran Chex	R	4	125	
9	Bran Flakes	P	5	190	
10	Cap'n Crunch	Q	0	35	
11	Cheerios	G	2	105	
12	Cocoa Puffs	G	0	55	
13	Corn Chex	R	0	25	
14	Corn Flakes	K	1	35	
15	Count Chocula	G	0	65	
16	Cracklin' Oat Bran	K	4	160	
17	Cream of Wheat (Quick)	N	1	0	
18	Crispy Wheat & Raisins	G	2	120	
19	Double Chex	R	1	80	
20	Froot Loops	K	1	30	
21	Frosted Flakes	K	1	25	
22	Fruit & Fibre Dates, Wal	P	5	200	
23	Fruitful Bran	K	5	190	
24	Fruity Pebbles	P	0	25	
25	Golden Grahams	G	0	45	
26	Grape Nuts Flakes	P	3	85	
27	Honey Nut Cheerios	G	1.5	90	
28	Honey-comb		P	0	35
29	Just Right Fruit & Nut		K	2	95
30	Life		Q	2	95
31	Lucky Charms		G	0	55
32	Maypo		A	0	95
33	Muesli Raisins, Dates, &		R	3	170
34	Multi-Grain Cheerios		G	2	90
35	Nutri-Grain Almond-Rais		K	3	130
36	Nutri-grain Wheat		K	3	90
37	Oatmeal Raisin Crisp		G	1.5	120
38	Post Nat. Raisin Bran		P	6	260
39	Product 19		K	1	45
40	Quaker Oatmeal		Q	2.7	110
41	Raisin Bran		K	5	240
42	Raisin Nut Bran		G	2.5	140
43	Rice Krispies		K	0	35
44	Shredded Wheat		N	3	95
45	Shredded Wheat 'n'Bran		N	4	140
46	Shredded Wheat spoon		N	3	120
47	Smacks		K	1	40
48	Special K		K	1	55
49	Strawberry Fruit Wheats		N	3	90
50	Total Corn Flakes		G	0	35
51	Total Raisin Bran		G	4	230
52	Total Whole Grain		G	3	110
53	Trix		G	0	25
54	Wheaties		G	3	110
55	Wheaties Honey Gold		G	1	60

Courtesy of John Stasko



Cereal (All)



Manufacturer

- A
- G
- K
- N
- P
- Q
- R

Fiber 0 14



Pot... 0 330



Even tougher ...

- What if you could only see one cereal's data at a time?
 - (e.g. some websites)
- What if I read the data to you?

Another example

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89



$$N = 11$$

$$\text{mean of } X\text{'s} = 9.0$$

$$\text{mean of } Y\text{'s} = 7.5$$

$$\text{equation of regression line: } Y = 3 + 0.5X$$

$$\text{standard error of estimate of slope} = 0.118$$

$$t = 4.24$$

$$\text{sum of squares } X - \bar{X} = 110.0$$

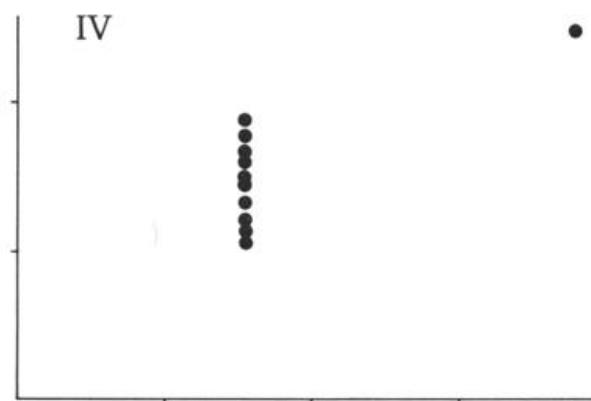
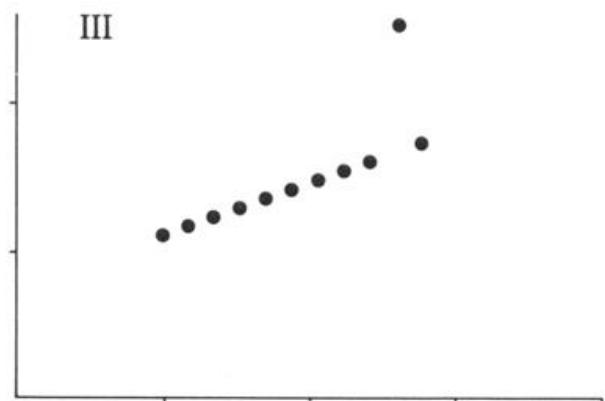
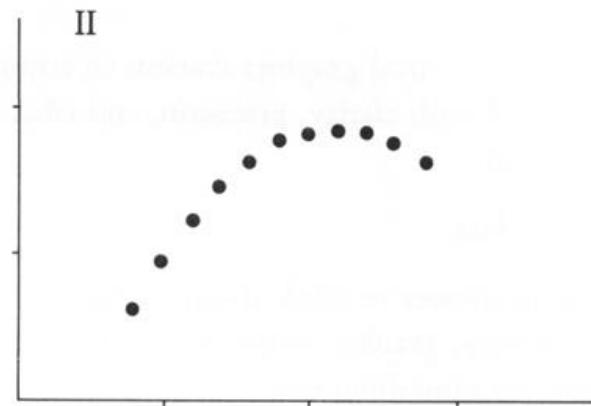
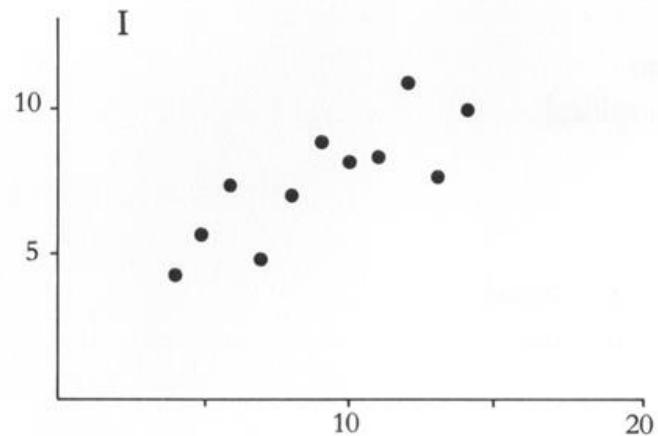
$$\text{regression sum of squares} = 27.50$$

$$\text{residual sum of squares of } Y = 13.75$$

$$\text{correlation coefficient} = .82$$

$$r^2 = .67$$

Aha !



Visualization

- Definition

“The use of computer-supported, interactive visual representations of data to amplify cognition.”

from [Card, Mackinlay Shneiderman ‘98]

Visualization

- Often thought of as process of making a graphic or an image
- Really is a cognitive process
 - Form a mental image of something
 - Internalize an understanding
- “The purpose of visualization is insight, not pictures”
 - Insight: discovery, decision making, explanation

Main Idea

- Visuals help us think
 - Provide a frame of reference, a temporary storage area
- Cognition → Perception
- Pattern matching
- External cognition aid
 - Role of external world in thinking and reason

[Larkin & Simon '87]

[Card, Mackinlay, Shneiderman '98]

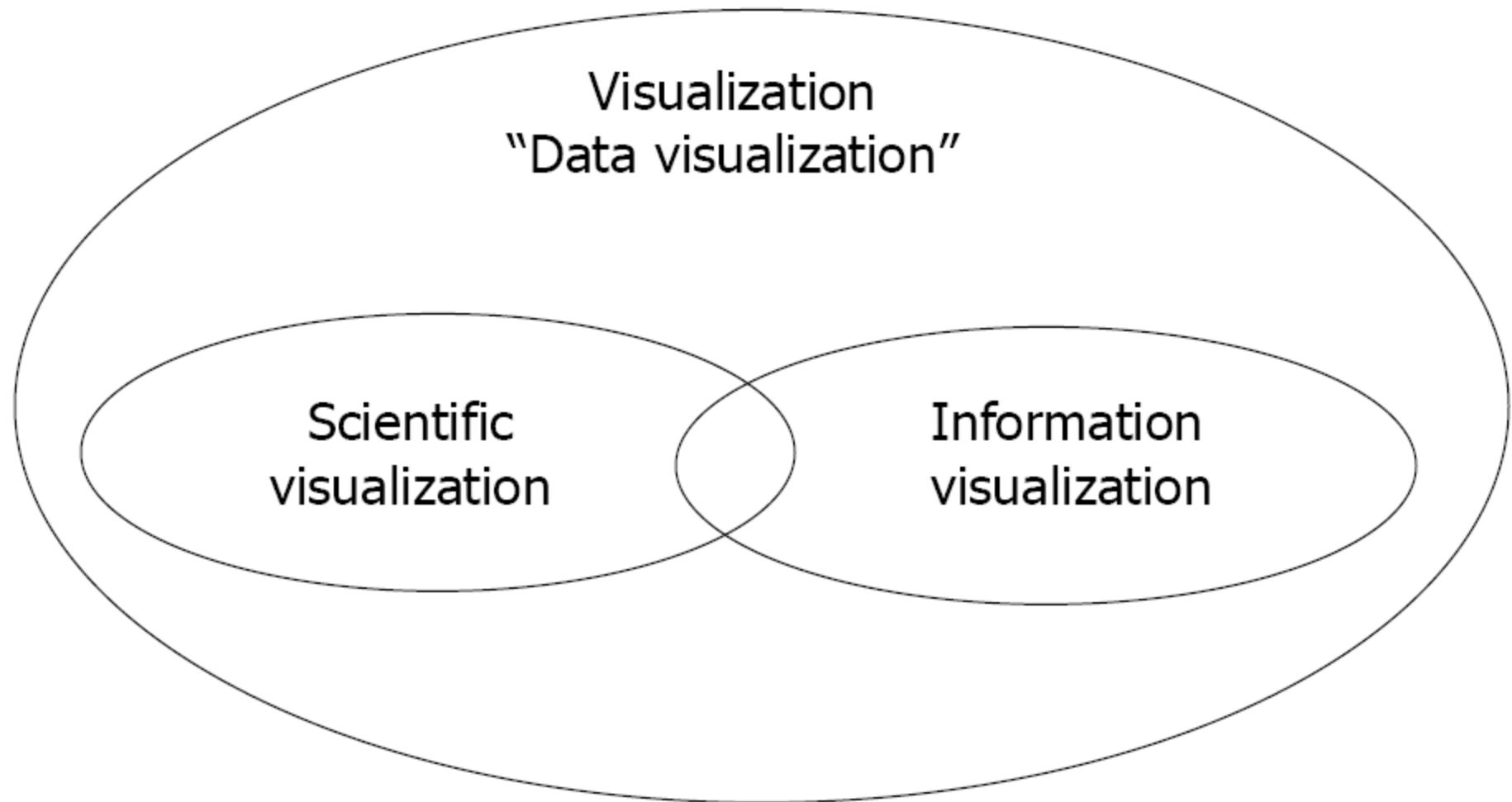
When to apply?

- Many other techniques for data analysis
 - Data mining, DB queries, machine learning...
- Visualization most useful in **exploratory data analysis**
 - Don't know what you're looking for
 - Don't have a priori question
 - Want to know what questions to ask

Part of our Culture

- “I see what you’re saying”
- “Seeing is believing”
- “A picture is worth a thousand words”

Overview



Scientific Visualization

- Primarily relates to and represents something physical or geometric
 - Often 3-D
 - Examples
 - Air flow over a wing
 - Stresses on a girder
 - Torrents inside a tornado
 - Organs in the human body
 - Molecular bonding

Information Visualization

- What is “information”?
 - Items, entities, things which do not have a direct physical correspondence
 - Notion of abstractness of the entities is important too
 - Examples
 - baseball statistics
 - stock trends
 - connections between criminals
 - car attributes...

Information Visualization

- What is “visualization”?

“The use of computer-supported, interactive visual representations of data to amplify cognition.”

from [Card, Mackinlay Shneiderman ‘98]

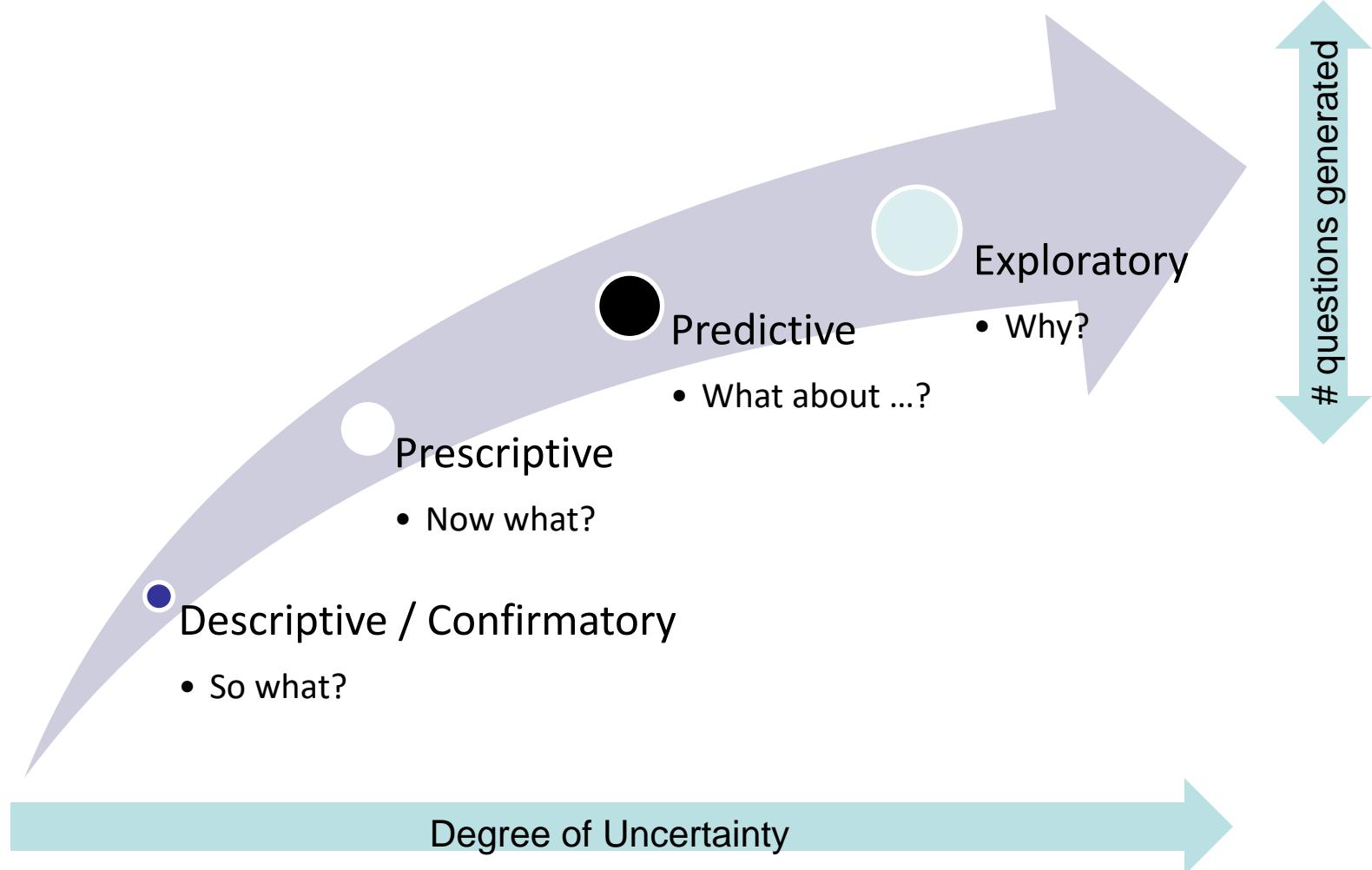
Information Visualization

- Components
 - Taking items without a direct physical correspondence and mapping them to a 2-D or 3-D physical space
 - Giving information a visual representation that is useful for analysis and decision-making

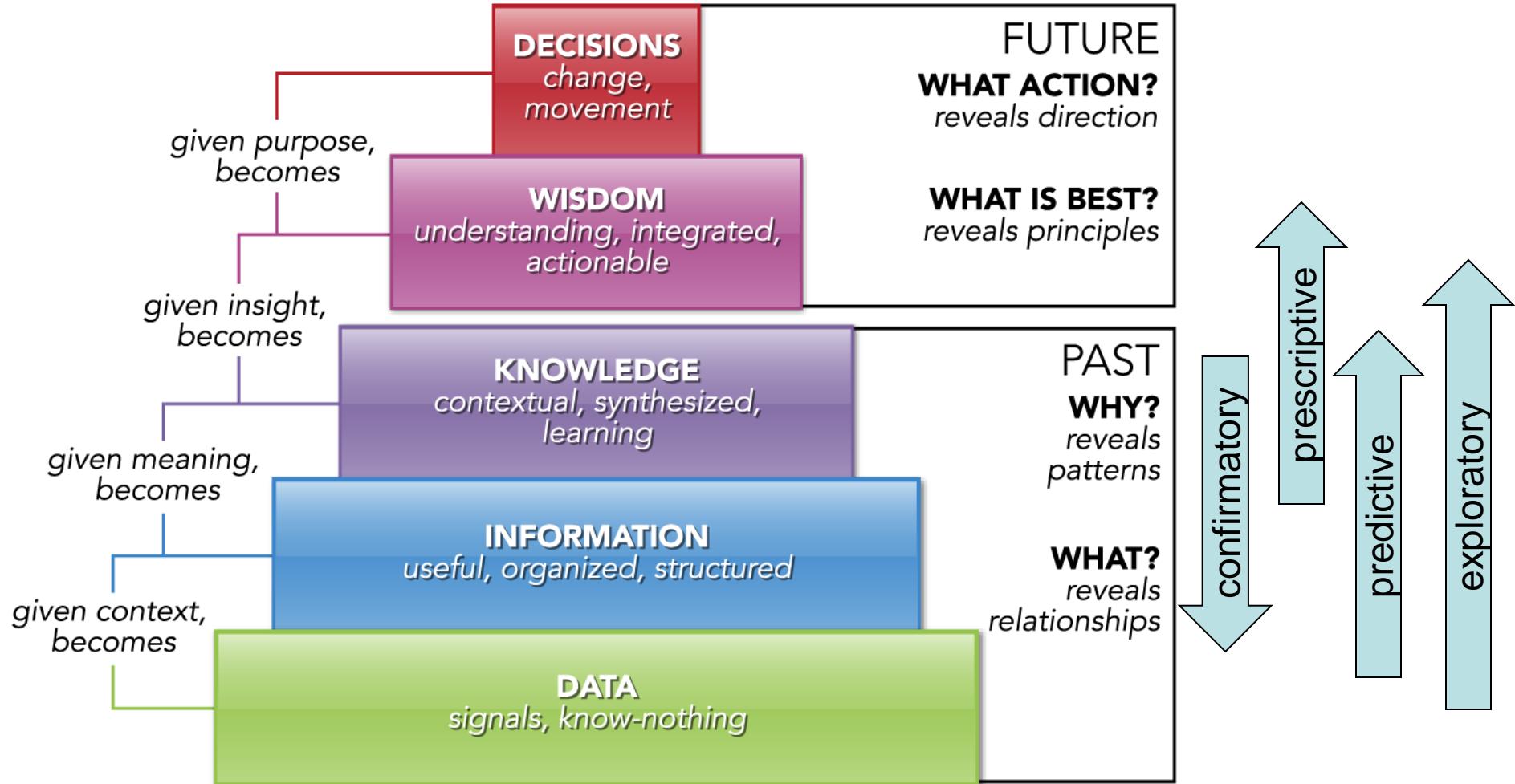
Information Visualization

- Two main uses
 - Analysis
 - Understand your data better and act upon that understanding
 - Presentation
 - Communicate and inform others more effectively

Types of Analysis Purposes



Analytic purposes & Decision Making



DIKW pyramid

Information Visualization

- Two main uses
 - Analysis
 - Understand your data better and act upon that understanding
 - **Presentation**
 - **Communicate and inform others more effectively**

InfoVis for Presentation Purposes



<https://www.youtube.com/watch?v=mkEXx7sDXAI>

Information Visualization

- Two key attributes
 - Scale
 - Challenge often arises when data sets become very large
 - Interactivity
 - Want to show multiple different perspectives on the data

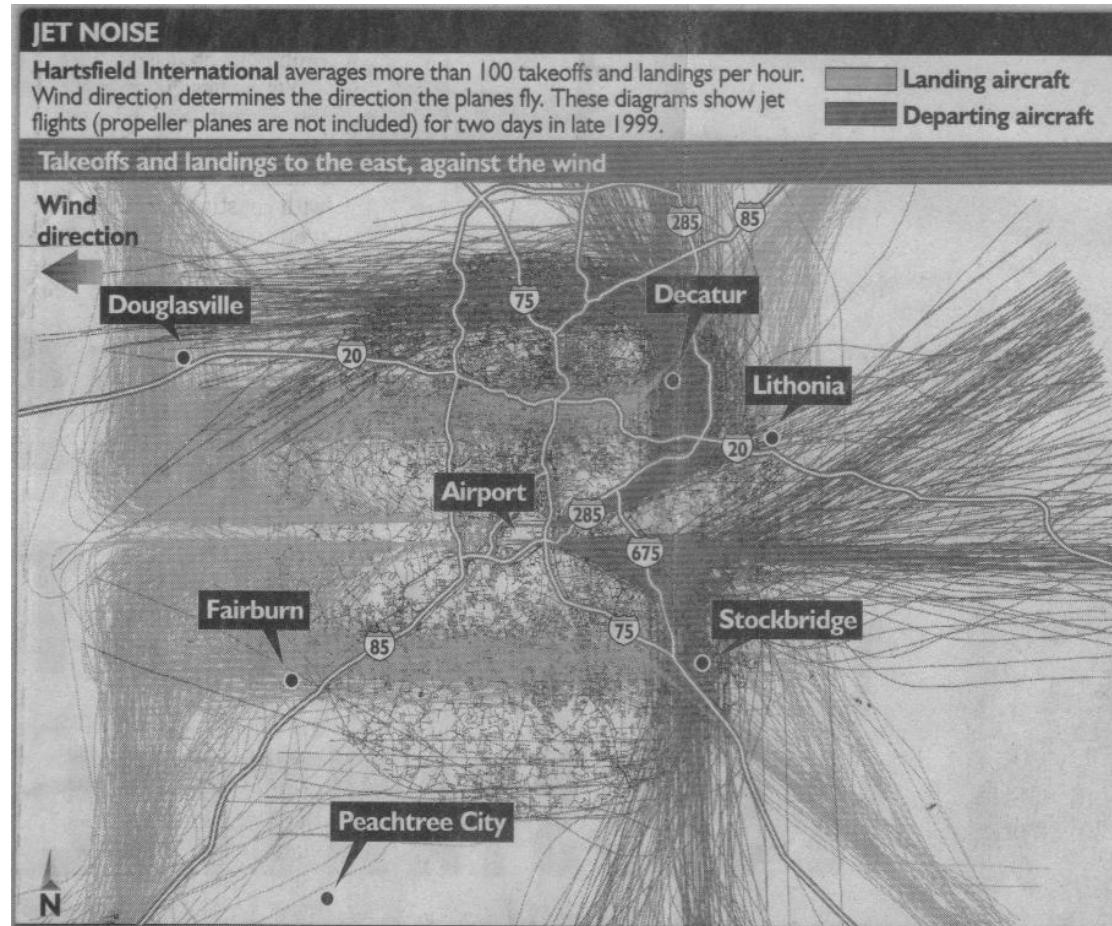
Information Visualization

- Example domains
 - Text
 - Statistics
 - Financial/business data
 - Internet information
 - Software
 - ...

Information Visualization

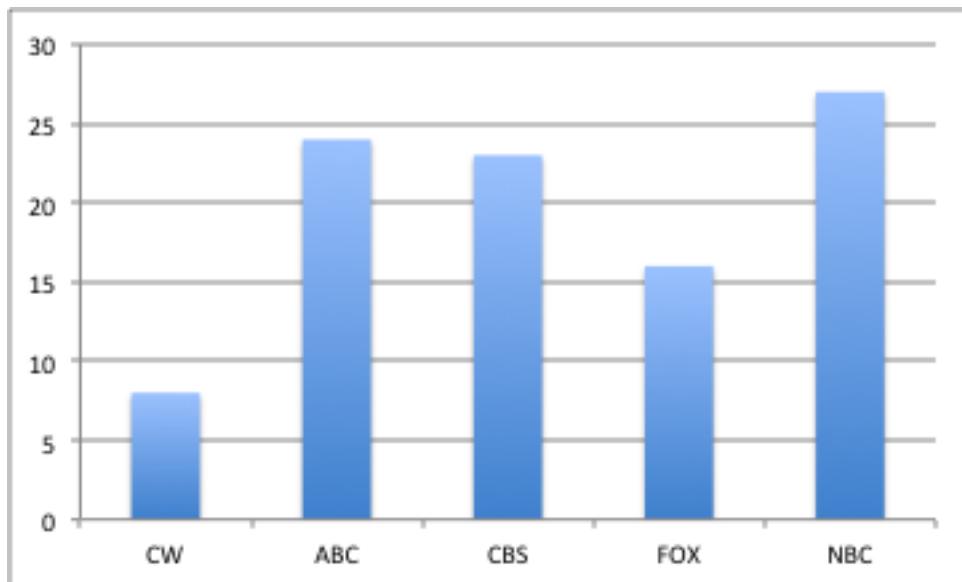
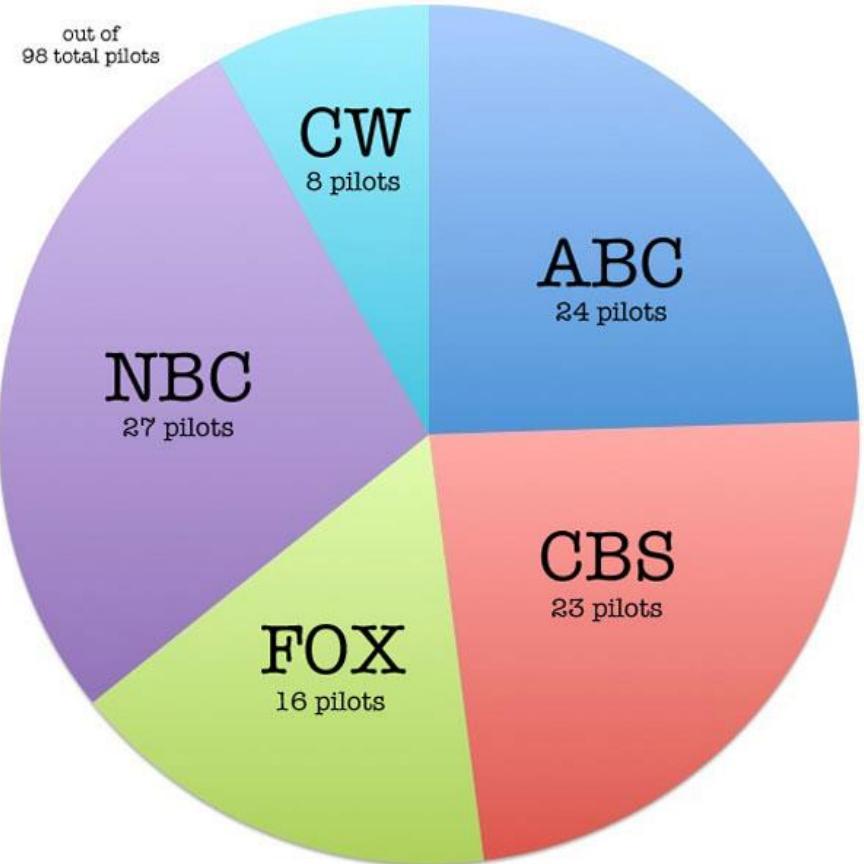
- Examples
 - Start with static picture
 - Are they information visualization according to our definition?

Atlanta Flight Traffic

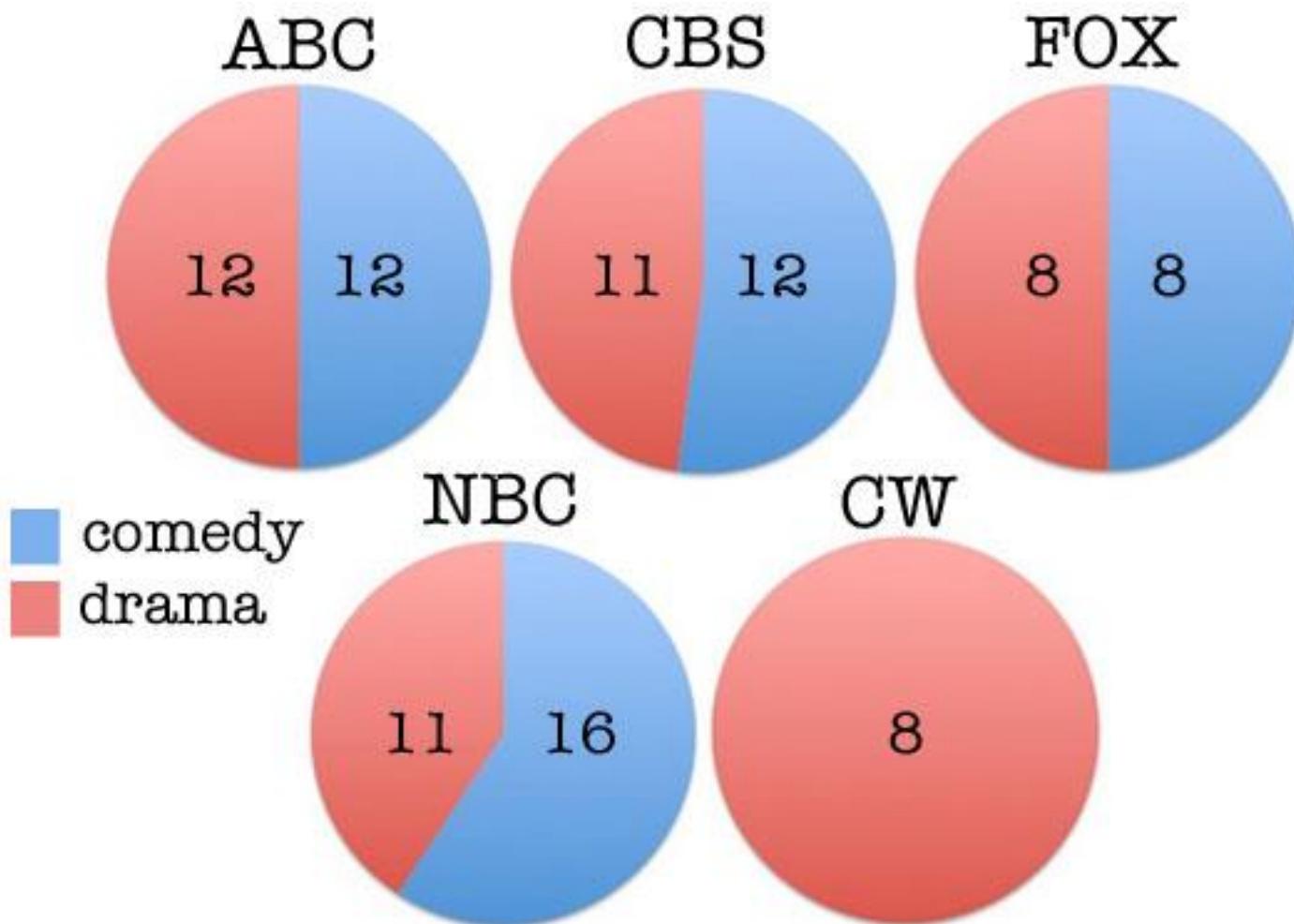


Atlanta Journal
April 30, 2000

2013 Pilot Season

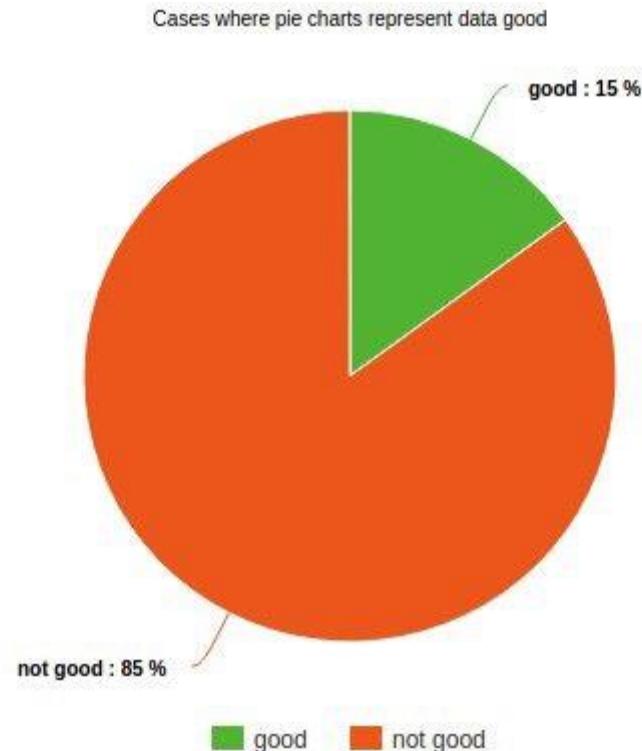


2013 Pilot Season



Pie charts are useful ...

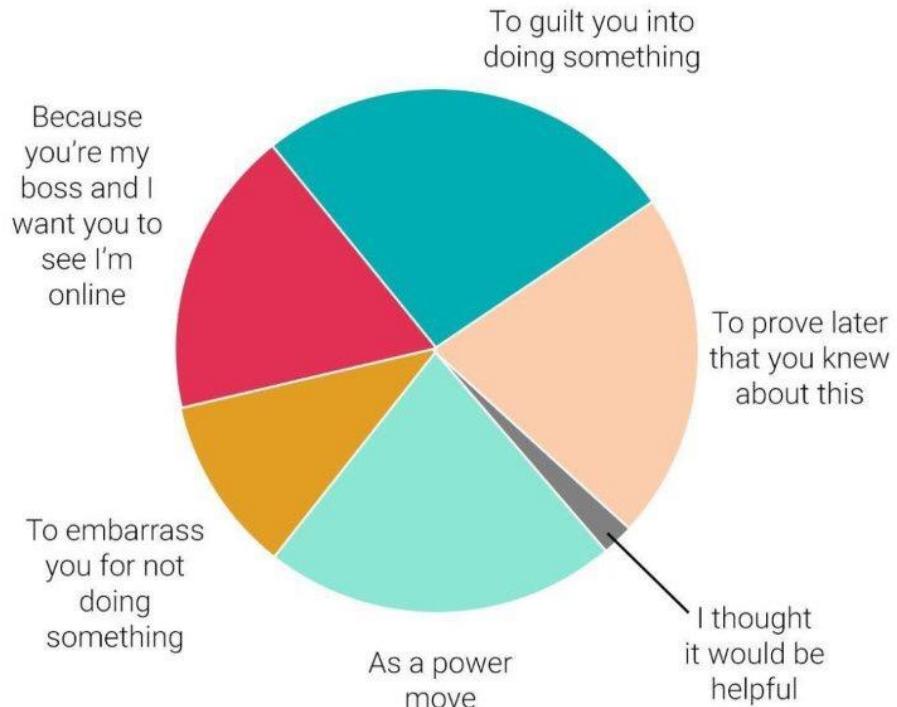
- for visualizing simple part-of-a whole relationship

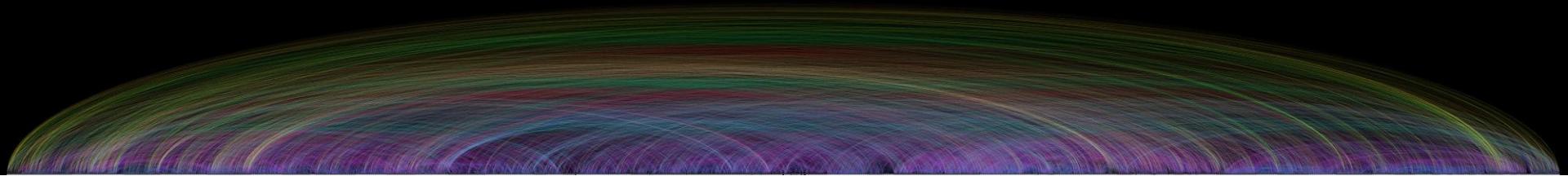


... which might be in the message

Why I CC'ed You on this E-mail

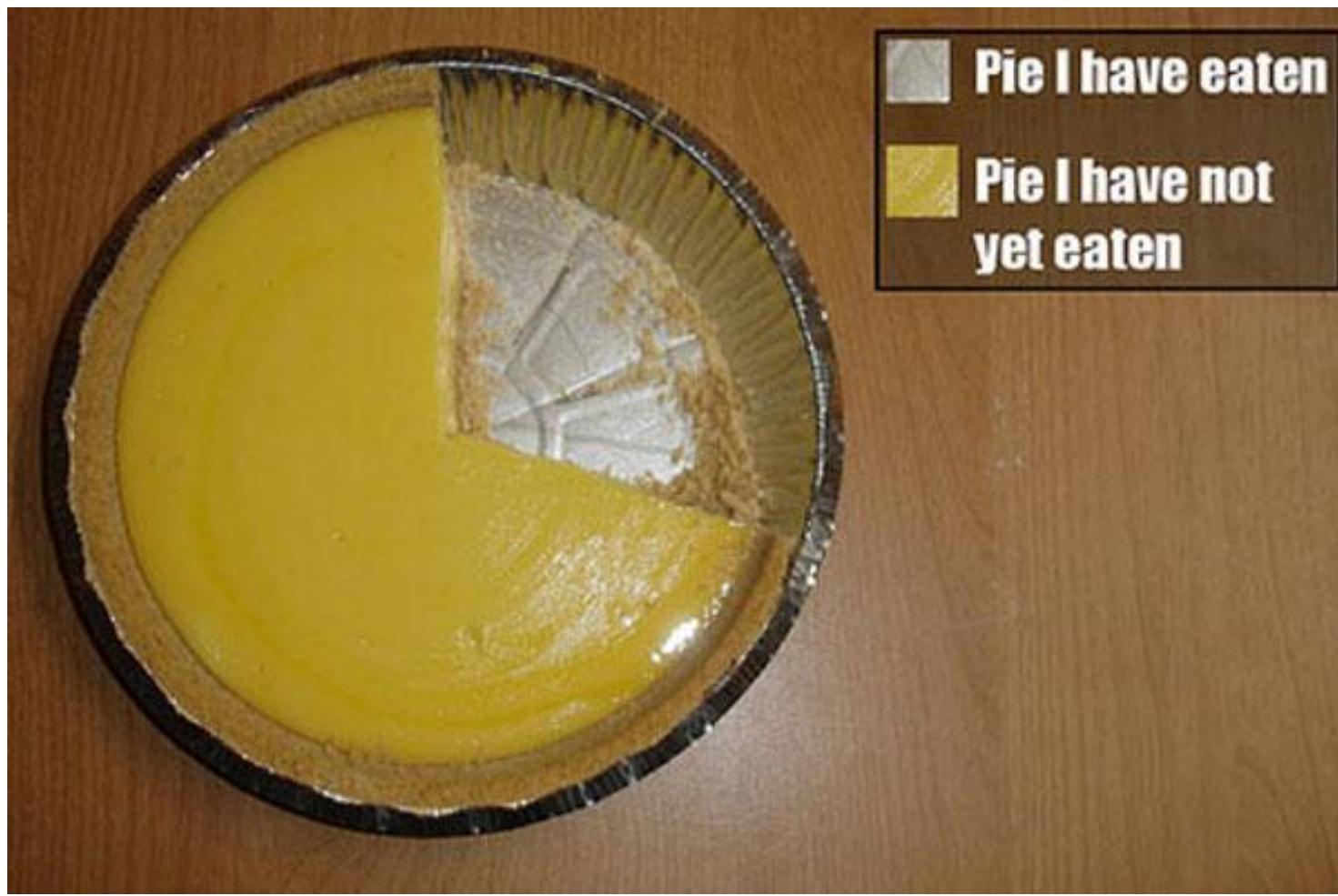
TheCooperReview.com





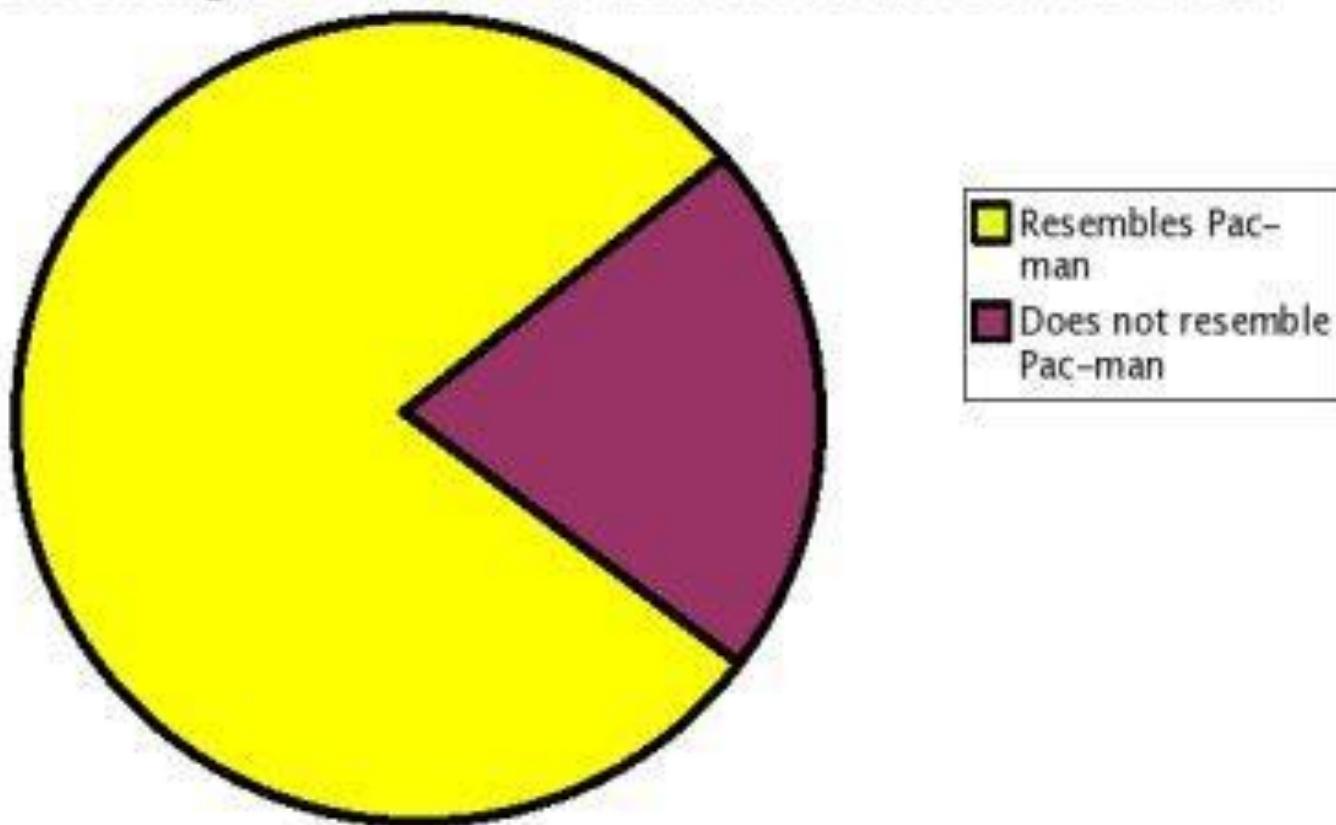
(and one of the few times pie charts are the best choice ...)

OR JUST FOR FUN

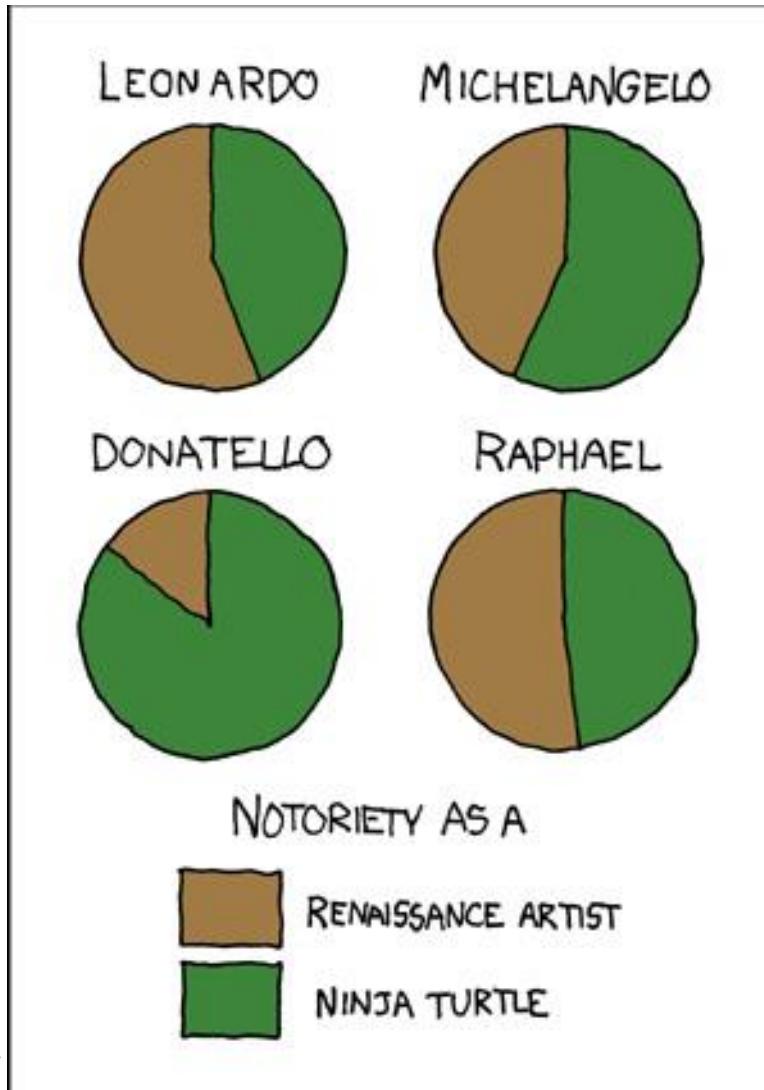


http://infosthetics.com/archives/2008/09/funniest_pie_chart_ever.html

Percentage of Chart Which Resembles Pac-man



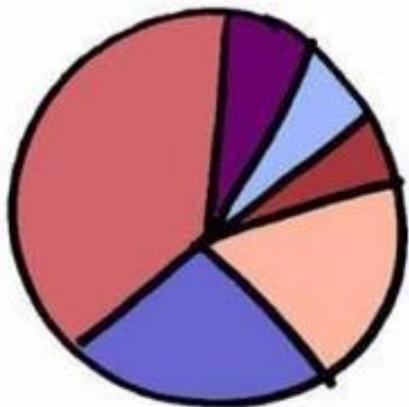
<http://www.boingboing.net/2006/11/02/hilarious-piechartvi.html>



<http://xkcd.com/197/>

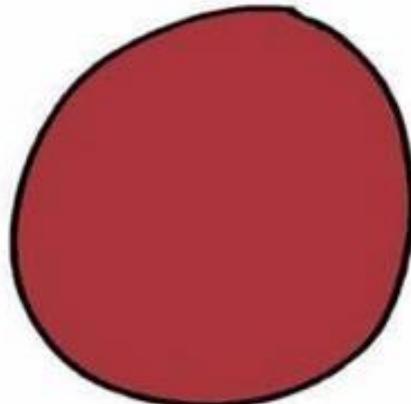
Looking at a photo of you

WHAT YOU SEE

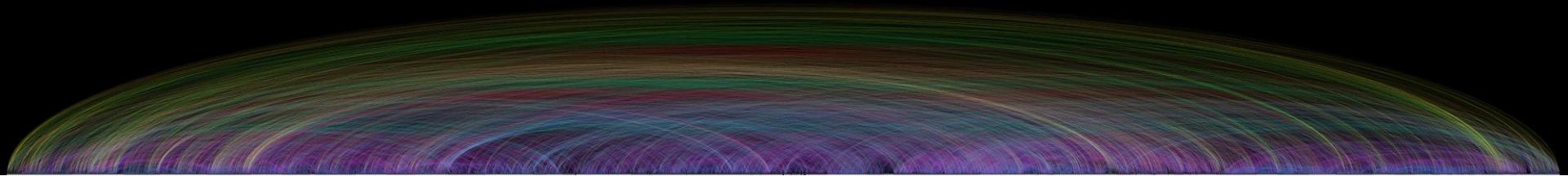


- nooo
- oof
- :)
- ugh
- why
- that's what i look like?

WHAT OTHER PEOPLE SEE



- Oh cool a
picture



OTHER EXAMPLES

Country Music



Figure 14. States Mentioned in Country-Music Lyrics

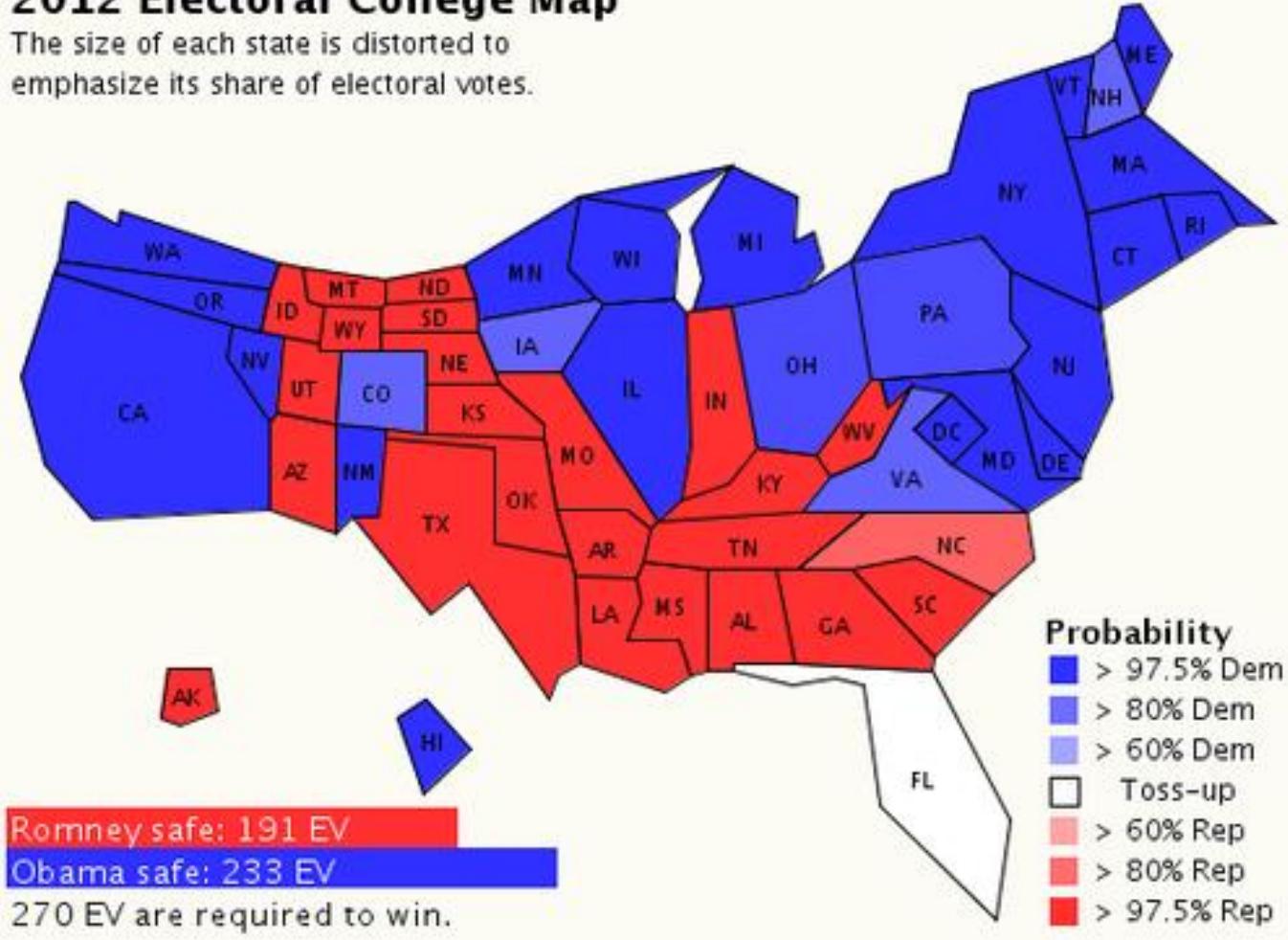
Source: Ben Marsh, "A Rose-Colored Map," *Harper's*, July 1977, 80. Used by permission.

Note: The size of each state is proportional to the number of times it is mentioned.

Electoral College

2012 Electoral College Map

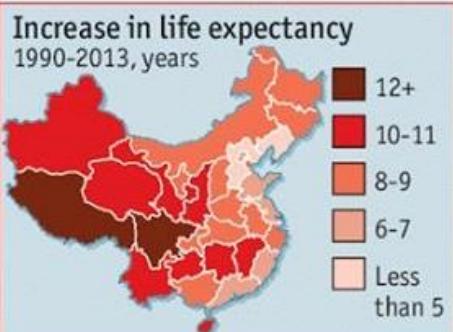
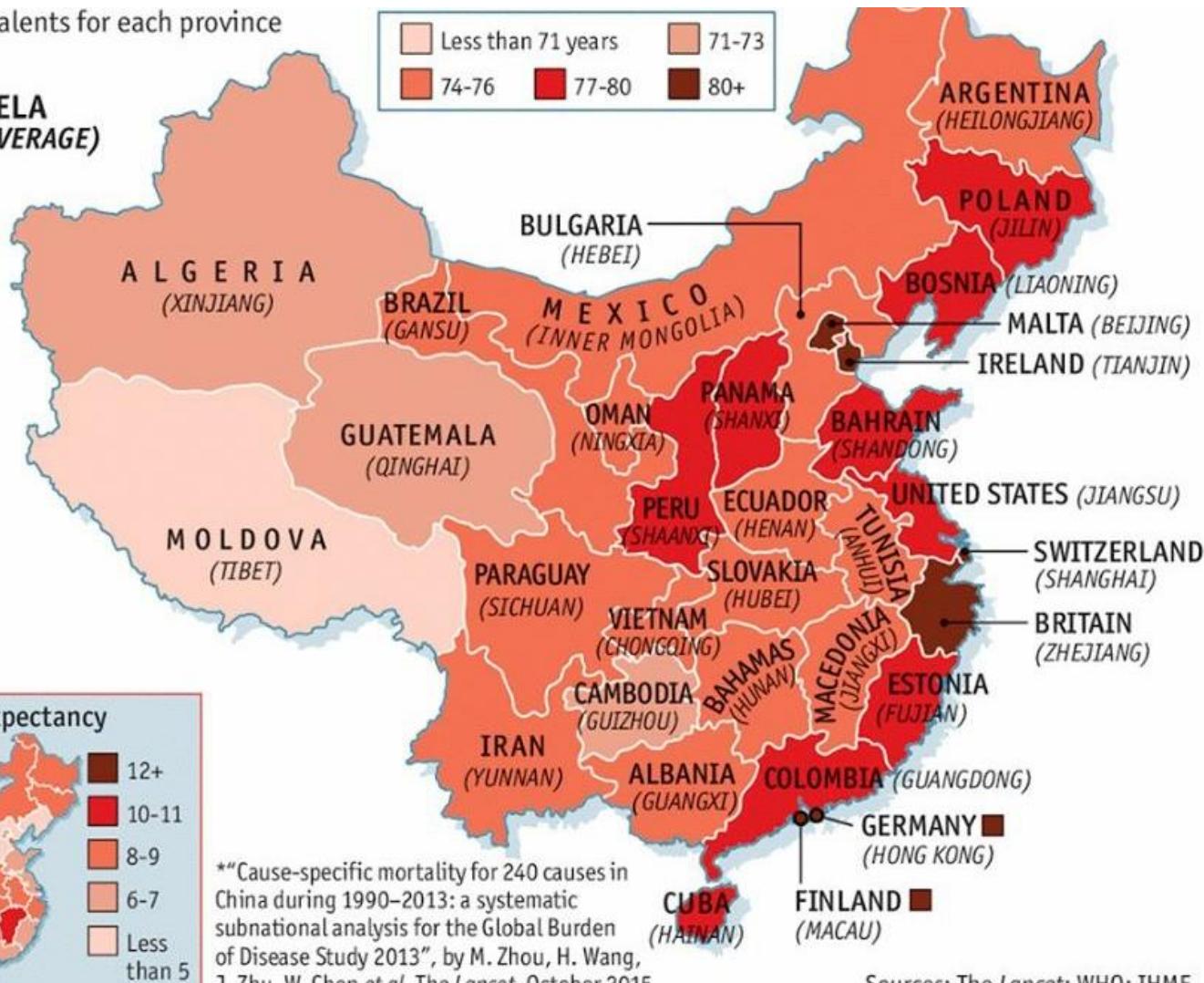
The size of each state is distorted to emphasize its share of electoral votes.



Life Expectancy in China

Country equivalents for each province
2013

VENEZUELA
(CHINA AVERAGE)

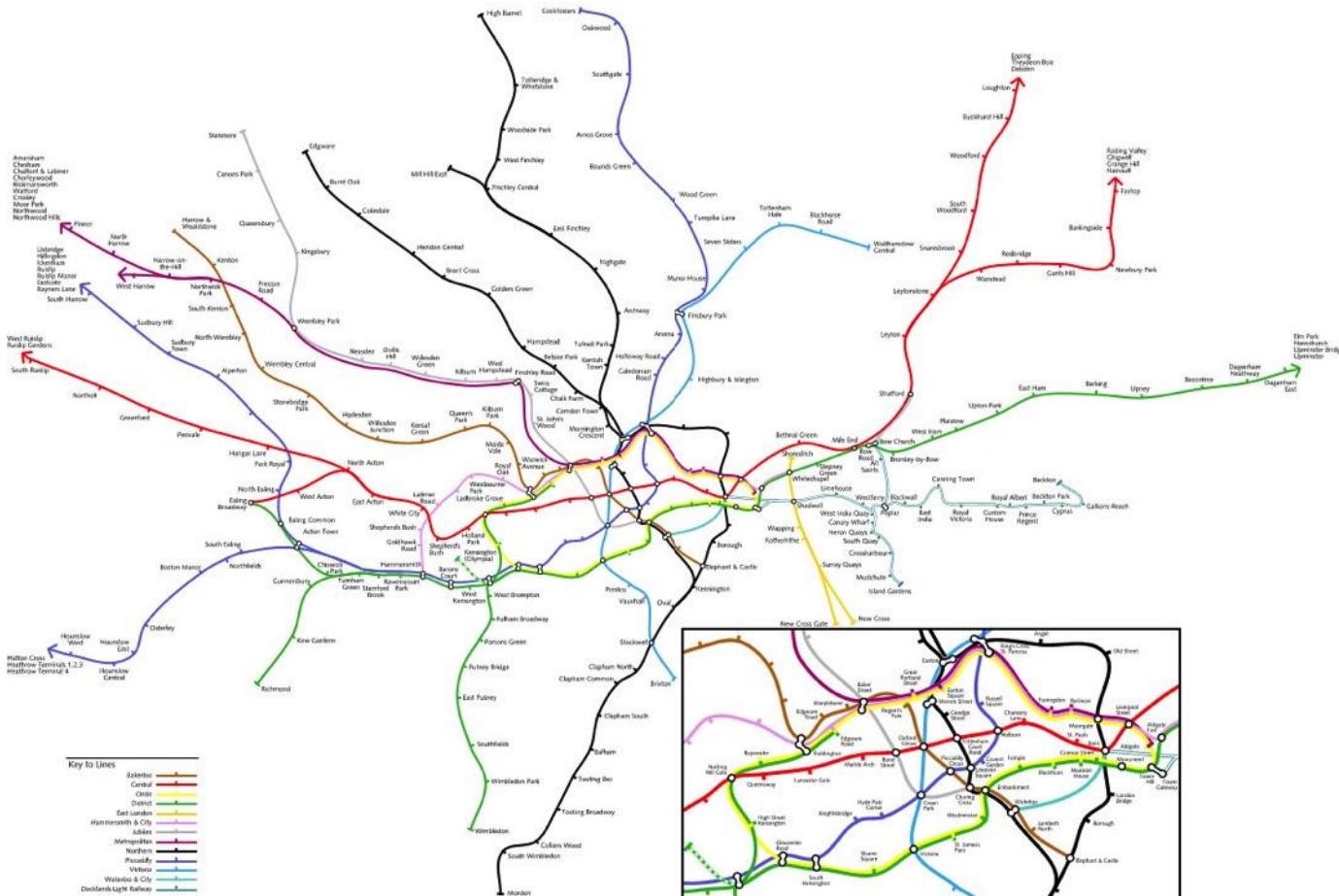


London Subway

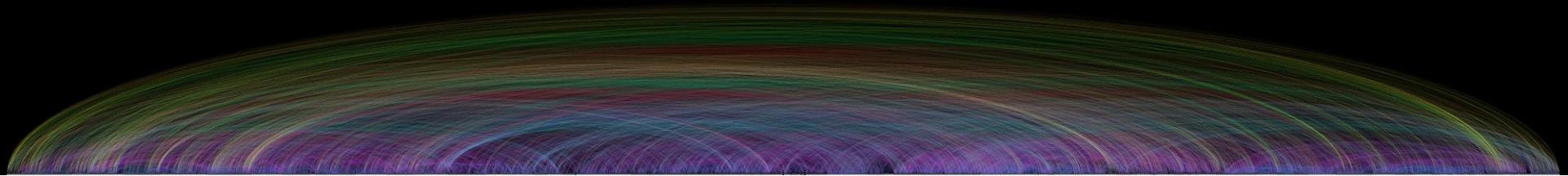


www.thetube.com

London Subway – True Geography

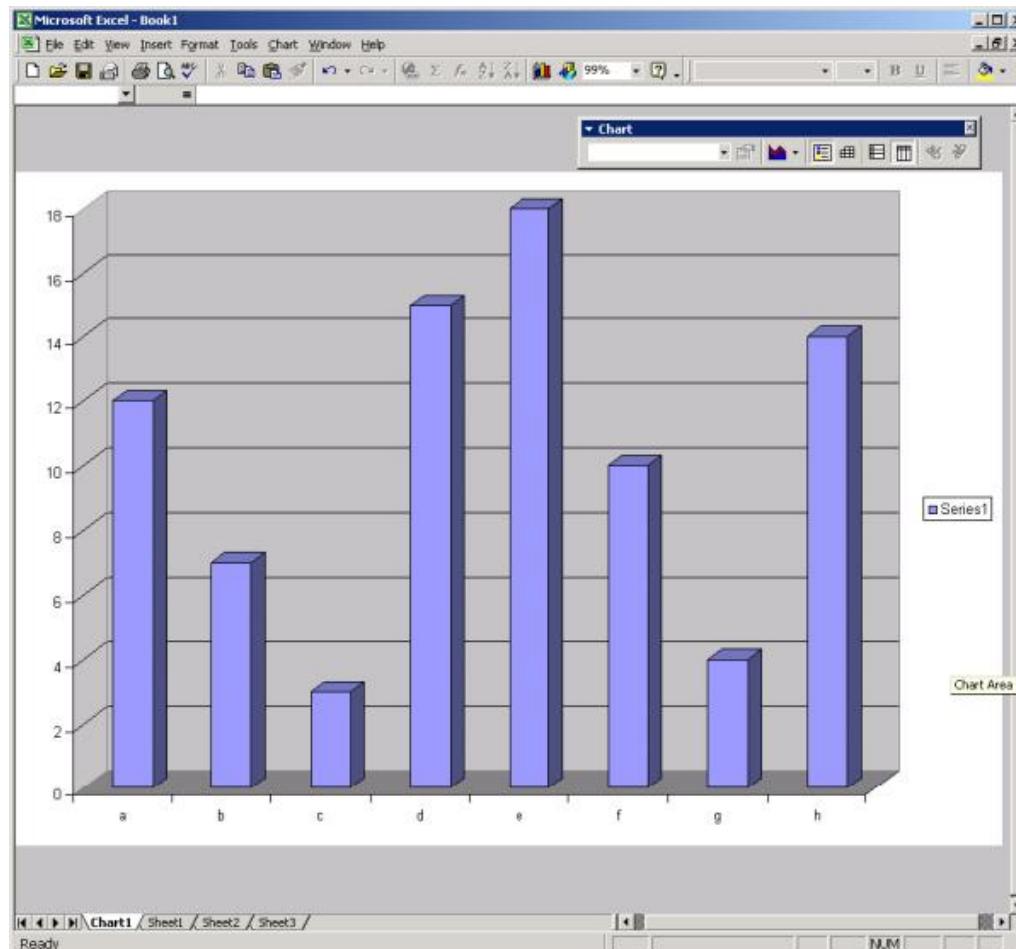


www.kottke.org/plus/misc/images/tubegeo.gif



BUT DON'T DO THIS

Excel



3D bars - Why ????

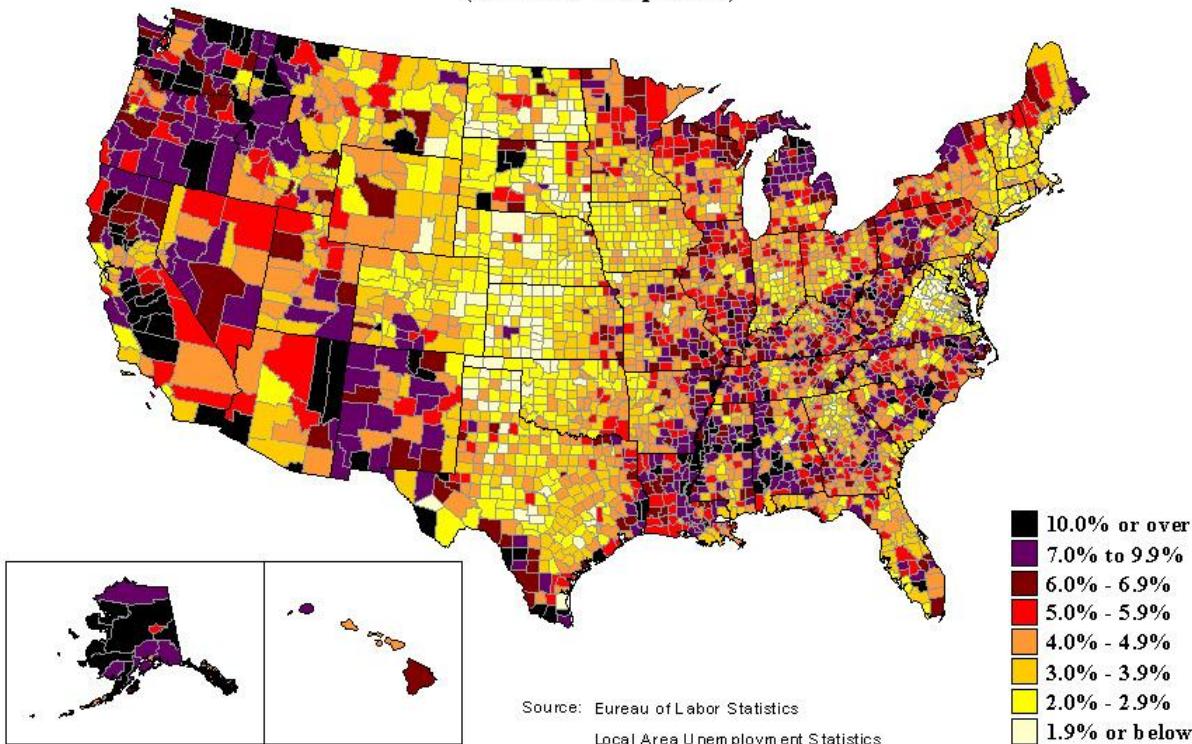
USA Today Graphics

Even worse ...

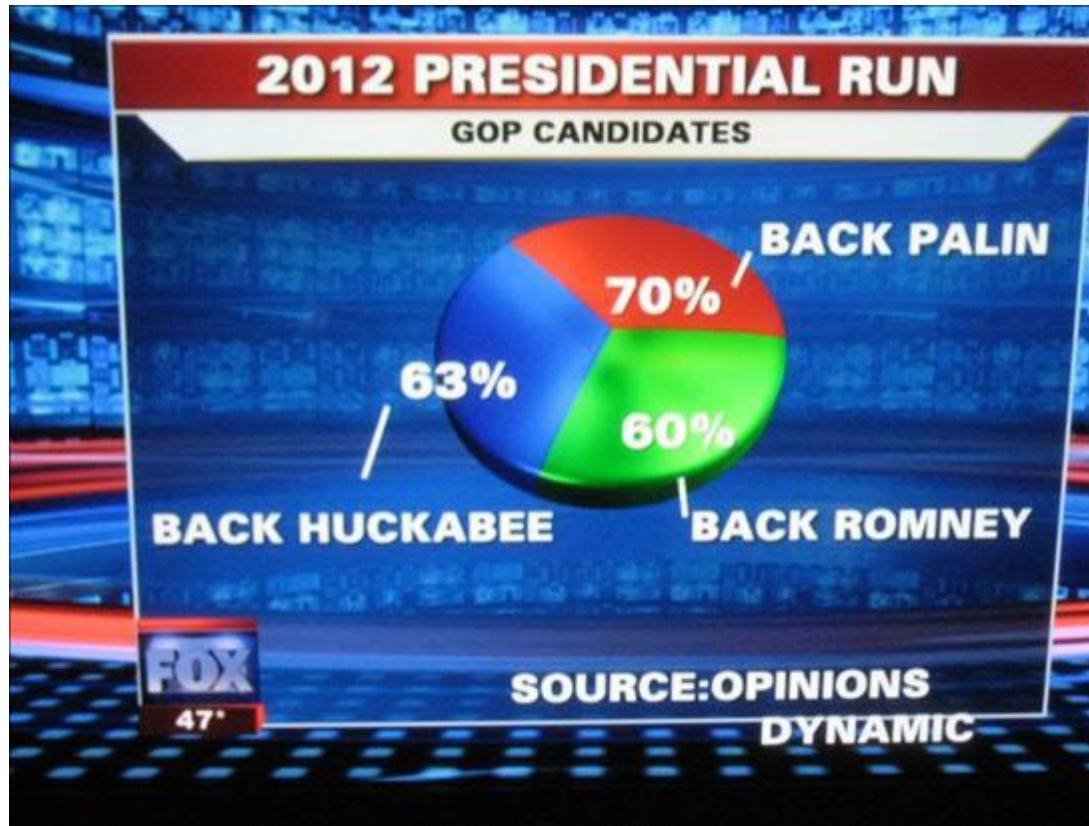


Unemployment Rates

Unemployment rates by county,
December 2000 - November 2001 averages
(U.S. rate = 4.6 percent)



Fox “News”



<http://wonkette.com/412361/all-193-of-republicans-support-palin-romney-and-huckabee>

Fan Loyalty

WHAT WOULD YOU DO IF YOUR
FAVORITE COLLEGE FOOTBALL TEAM
WAS PLAYING DURING TRICK-OR-TREATING
IN YOUR NEIGHBORHOOD?

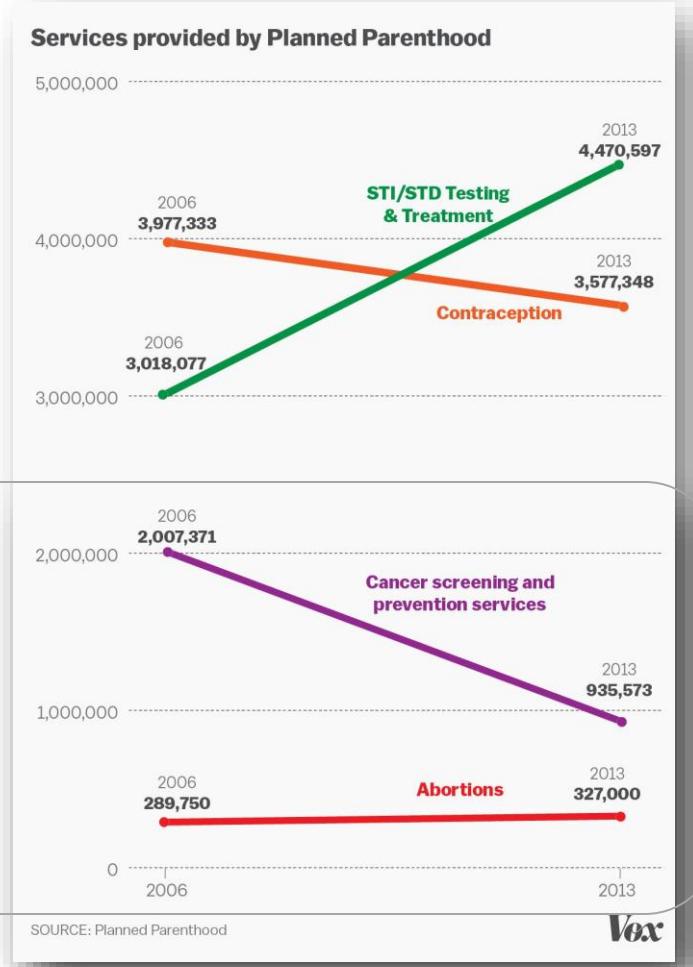
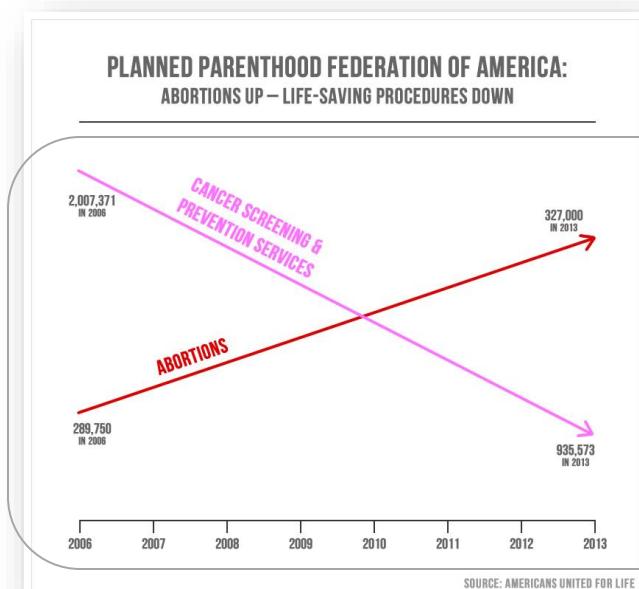


Source: Rasmussen Reports
and Discover Financial Services

FOLLOW US ON  www.Twitter.com/Discover

First and foremost: DO NO HARM

i.e. Don't Lie!



So far ...

- Discussion of Information Visualization in static images

Tools/Systems

- Now Interaction becomes important
- Examples ...

Map of the Market



ⓘ Use mouse wheel to zoom in and out. Drag zoomed map to pan it. Double-click a ticker to display detailed information in a new window.
Hover mouse cursor over a ticker to see its main competitors in a stacked view with a 3-month history graph.

-3% -2% -1% 0% +1% +2% +3%

Demo

<https://finviz.com/map.ashx>

Some days it looks like this



Use mouse wheel to zoom in and out. Drag zoomed map to pan it. Double-click a ticker to display detailed information in a new window.
Hover mouse cursor over a ticker to see its main competitors in a stacked view with a 3-month history graph.

-30% -20% -10% 0% +10% +20% +30%



By DEREK WATKINS, K.K. REBECCA LAI, LARRY
BUCHANAN and KAREN YOURISH NOV. 7, 2018

<https://www.nytimes.com/interactive/2018/11/07/us/politics/how-democrats-took-the-house.html>

Baby Names Popularity - NameVoyager: Baby Name Wizard Graph of Most Popular Baby Names

NameVoyager: Explore baby names and name trends letter by letter

Looking for the perfect baby name? [Sign up for free](#) to receive access to our expert tools!

Baby Name >



Both Boys Girls

Current rank: boys

1000

500

100

25

1

girls

1000

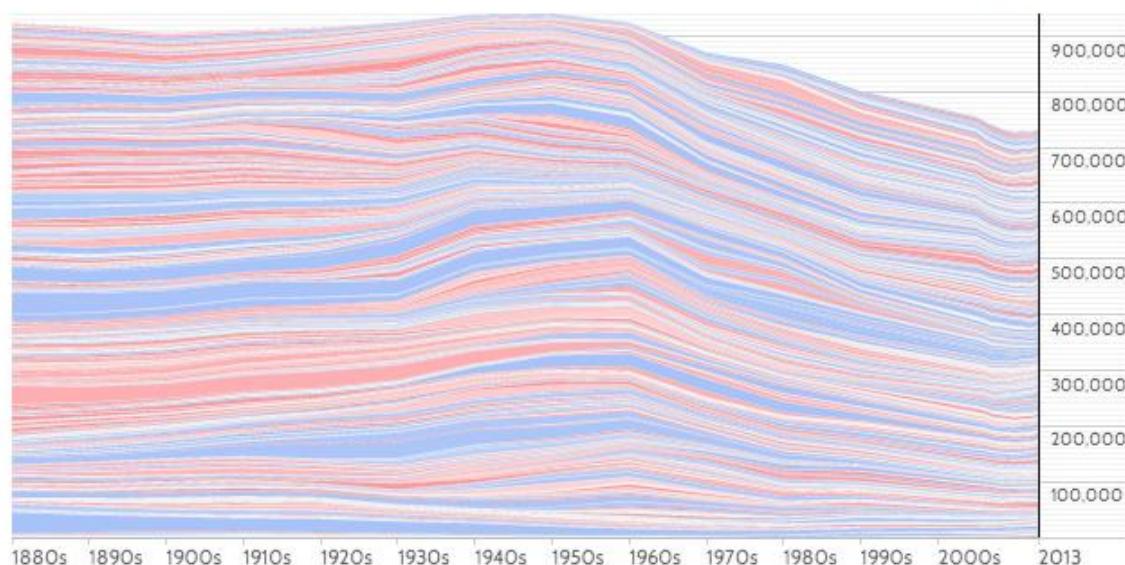
500

100

25

1

per million births



Click a name graph to view that name. Double-click to read more about it.

[enlarge](#)

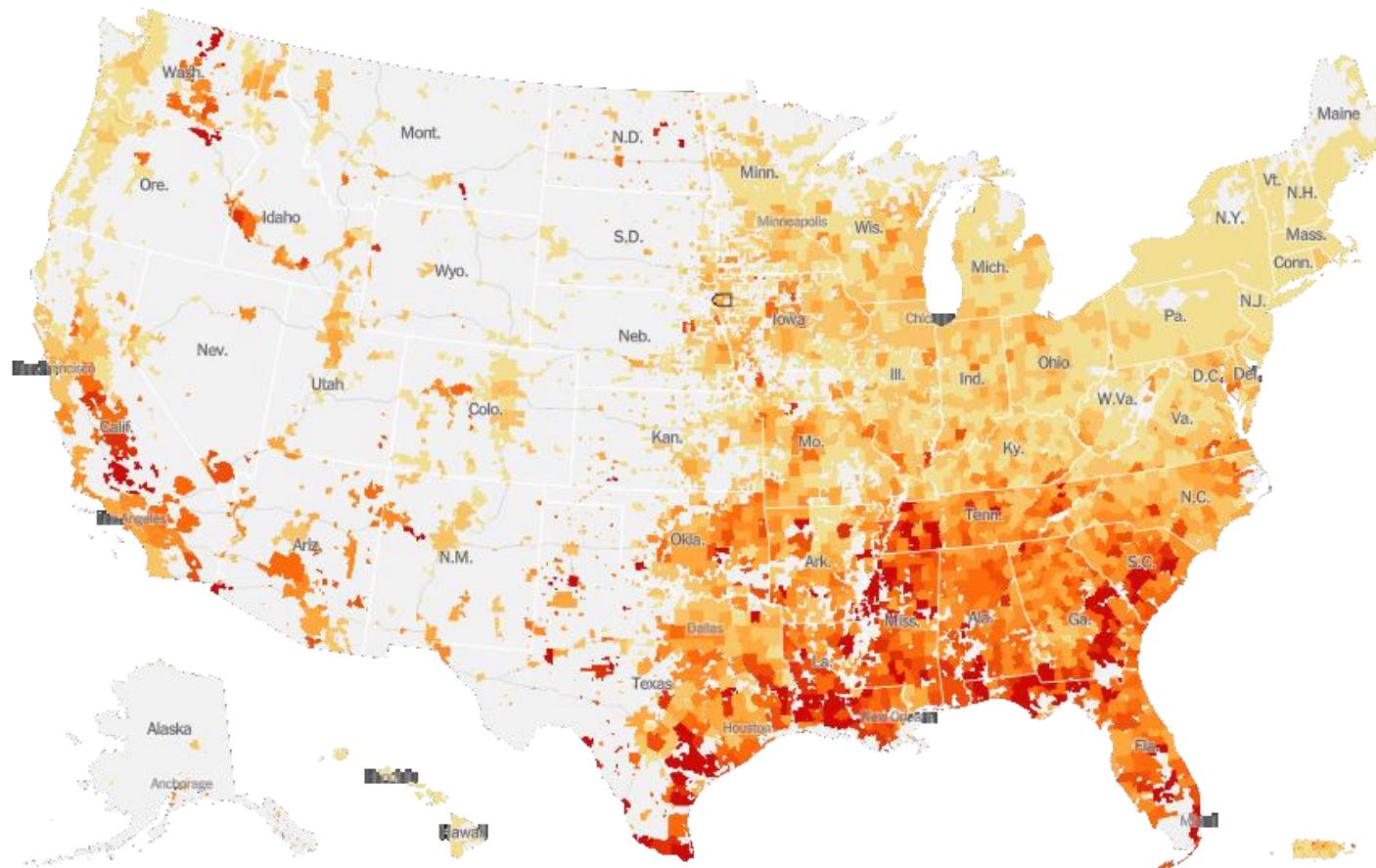
See top names by decade, trends by letters, and more with Expert Name Voyager! [Sign up for free!](#)



15,612 people like this. Be the first of your friends.

www.babynamewizard.com/namevoyager/

Coronavirus in the U.S. (NY Times)



<https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html?action=click&module=Top%20Stories&pgtype=Homepage#states>

RIP David Bowie

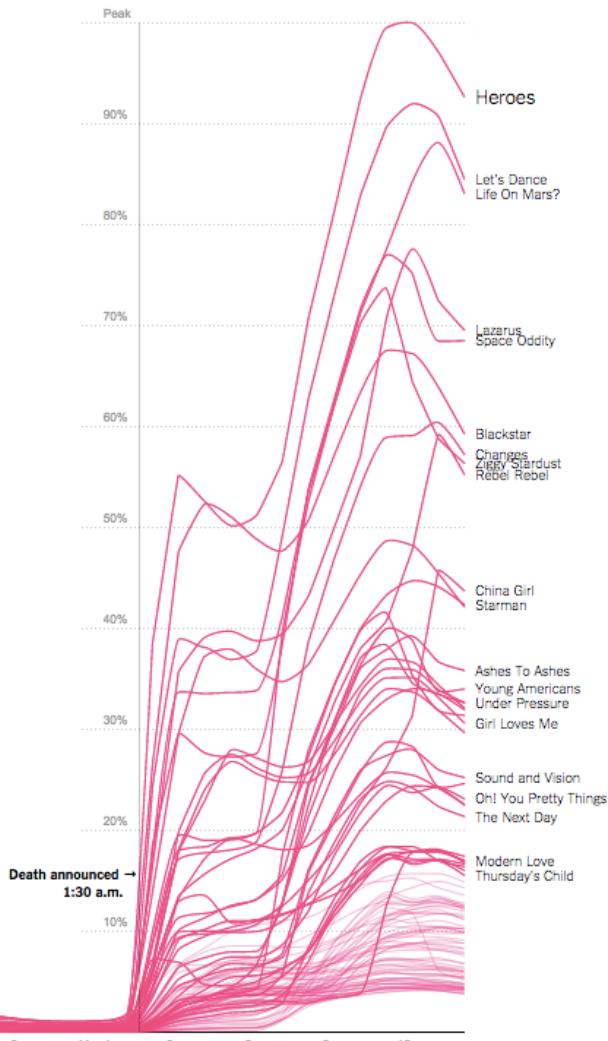
The David Bowie
Song That Fans Are
Listening to Most:
'Heroes'

By QUOCTRUNG BUI, JOSH KATZ and JASMINE C. LEE JAN. 12, 2016

Popularity of David Bowie songs on Spotify

Plays per hour

Sunday 3am 6am 9am 12pm 3pm 6pm 9pm Monday



Plays are expressed as a percentage of the maximum hourly plays for the leading song.

<http://www.nytimes.com/interactive/2016/01/12/upshot/david-bowie-songs-that-fans-are-listening-most-heroes-starman-major-tom.html>

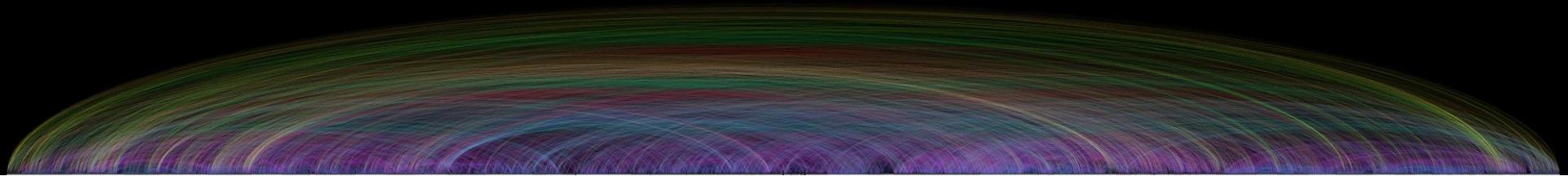
Tasks for InfoVis?

- Search (OK)
 - Finding a specific piece of information
 - How many games did the Braves win in 1995?
 - What novels did Ian Fleming author?
- Browsing (Better)
 - Look over or inspect something in a more casual manner, seek interesting information
 - Learn about crystallography
 - What has Jane been up to lately?

Tasks for Info Vis

- Analysis
 - Comparison / Difference
 - Outliers, Extremes
 - Patterns
- Assimilation
- Monitoring
- Awareness

More in a future class ...



Good Resources ...

InfoVis Wiki

Log in / create account

Page Discussion Read Edit View history Go Search

Share +

Visualization Design Patterns

Contents [hide]

- 1 Overview
- 2 Foundational Patterns
 - 2.1 Visualization Architecture
 - 2.2 Interaction Patterns
 - 2.3 Display Rendering Patterns
- 3 Programming Patterns
 - 3.1 Data Modelling and Management
 - 3.2 View
 - 3.3 Controller
- 4 Development Patterns
 - 4.1 Diagnosis
 - 4.2 Evaluation and Testing
- 5 References

Overview [edit]

For any design discipline, including Information and Scientific Visualization, there are design problems within the domain that tend to arise time and again. Although effective solutions may be known, they are not helpful unless that knowledge is coherently structured and communicated to colleagues. Without good structure, design knowledge may be disparate and disjointed, and relationships among solutions may not be apparent. Without good communication, designers who are unaware of solutions may be forced to 'reinvent the wheel', wasting unnecessary effort on a problem that has already been solved by others. Some means of capturing and codifying solutions to Information Visualisation design problems would be useful for those engaged in practice. The Visualization Design Patterns described here have been proposed for this purpose.

Foundational Patterns [edit]

These patterns are the building blocks of Visualizations. They are more "reference patterns" than true "design patterns". Visualization designers will assume they are a given component of visualization design. The patterns are included to provide a common technical lexicon, to build a *lingua franca* used throughout the language, to encourage communications between novices and experts, and to facilitate a discussion of pros and cons.

<https://www.infovis-wiki.net>

Infosthetics Blog

information aesthetics. Where form follows data.

SUGGEST ARCHIVES ABOUT [f](#) [e](#) [n](#)

Visualizing Publicly Available US Government Data On! »
19 September 2014 • no comments • infographic / open data • by infosthetics

FROGS & TOADS
47% at risk

SALAMANDERS
67% at risk

Bolitoglossa magnifica
Magnificent Web-footed Salamander

The Disappearing Planet: Comparing The Extinction Rate! »
18 September 2014 • no comments • aesthetic / collection / infographic • by infosthetics

Githut: The Universe Of Programming Languages Across! »
12 September 2014 • no comments • collection / open data • by infosthetics

Slice and Dice Data

5 Most Influential Visualizations of All Time

CHECK THEM OUT

advertisers here »

SEARCH

A LONG TIME AGO...

No images found in XML-file.

CONNECT

[Like](#) [Share](#) 3,349 people like this. Be the first of your friends.

[Follow @infosthetics](#) 18.2K followers

THE CLASSICS

Brad Paley - TextArc »
Hans Rosling - TED Talk (video) »
Jonathan Harris - We Feel Fine »
Marcos Weskamp - Newsmap »
Martin Wattenberg - Baby Name Voyager »
Michael Wesch - Web 2.0 (video) »

<http://infosthetics.com/cgi-bin/mt/mt-search.cgi?search=election&IncludeBlogs=1&limit=20>

NY Times (@NYTgraphics)



NYT Graphics • 5,134 tweets

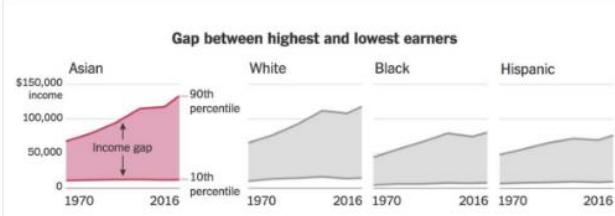
Tweets 5,134 Following 66 Followers 137K Likes 69 Lists 1 Follow

Who to follow · Refresh · View all

- Post Graphics** • Follow
- WSJ Graphics** • Follow
- Graham Roberts** • Follow

Tweets **Tweets & replies** **Media**

NYT Graphics • 20h
How "Crazy Rich" Asians have led to the largest income gap in the U.S.
nytimes.com/interactive/20...



Gap between highest and lowest earners

Year	Asian	White	Black	Hispanic
1970	\$50,000	\$50,000	\$50,000	\$50,000
2016	\$150,000	\$100,000	\$70,000	\$60,000

12 Followers you know



Trends for you · Change

- #iahsfb
- #Hawkeyes

<http://getdolphins.com/blog/interactive-data-visualizations-new-york-times/>
<https://www.nytimes.com/interactive/2018/us/2018-year-in-graphics.html>

Additional Books

