

Visualization. Introduction

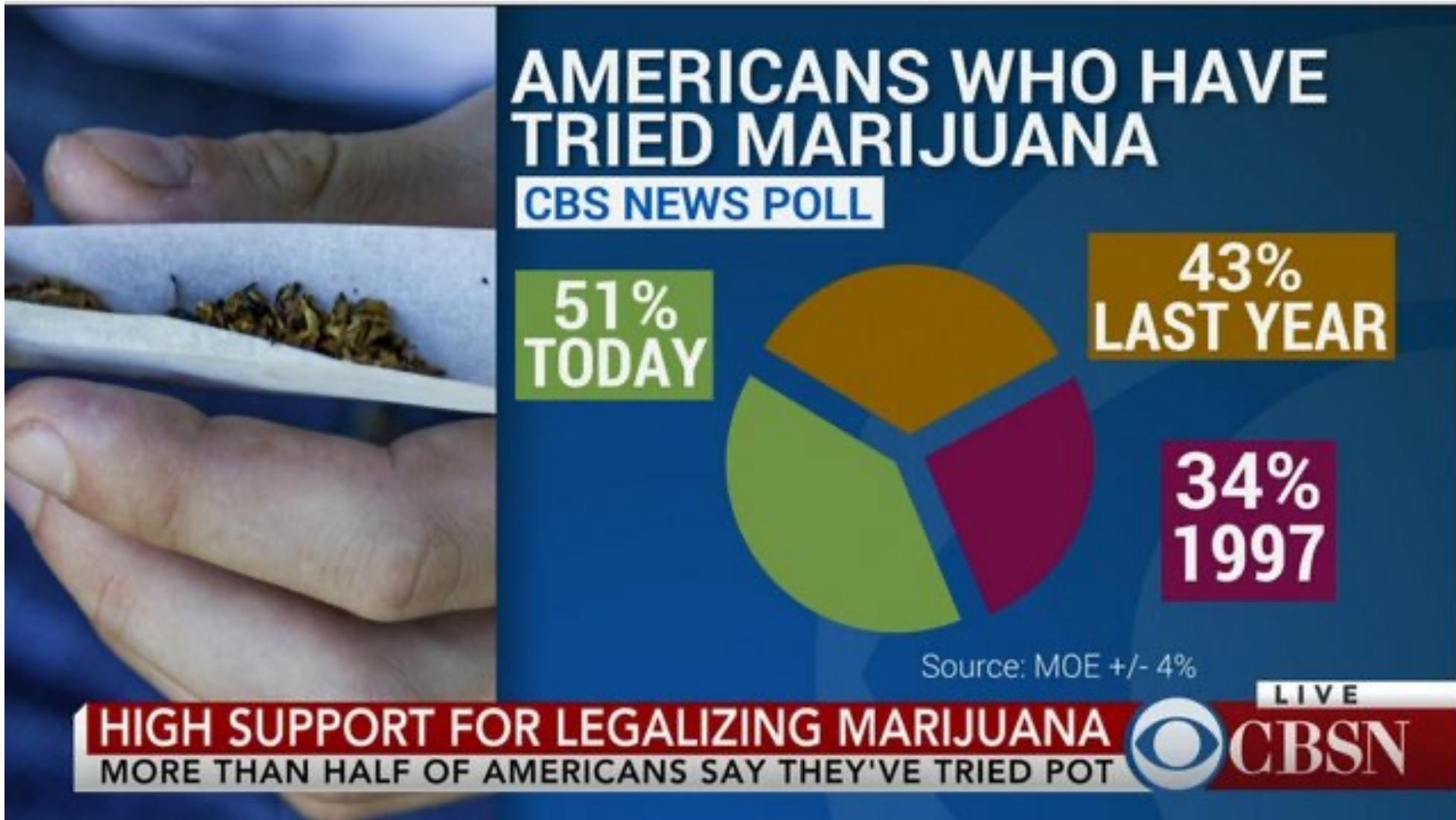
Pere-Pau Vázquez

Dept. Computer Science – UPC

“A graphic is not ‘drawn’ once and for all; it is ‘constructed’ and reconstructed until it reveals all the relationships constituted by the interplay of the data. The best graphic operations are those carried out by the decision-maker themselves.”

Jacques Bertin

Introduction



11 ABRIL 2021 Domingo

ABC.es

ABC

BARÓMETRO DE GAD 3 PARA EL 4-M

AYUSO DOBLEGARÁ AL SANCHISMO CON EL APOYO DE VOX

Los partidos de Casado y Abascal superarían en 5 diputados la mayoría absoluta en las elecciones madrileñas

[Editorial y páginas 18-19]

Isabel Díaz
Ayuso
PP

Rocío
Monasterio
Vox

Mónica
García
Más Madrid

Pablo
Iglesias
UP

Ángel
Gabilondo
PSOE
69
Mayoría absoluta en 2021

Vox
12

PP
62-63

PSOE
33-34

Más Madrid
17-19

Unidas
Podemos
10

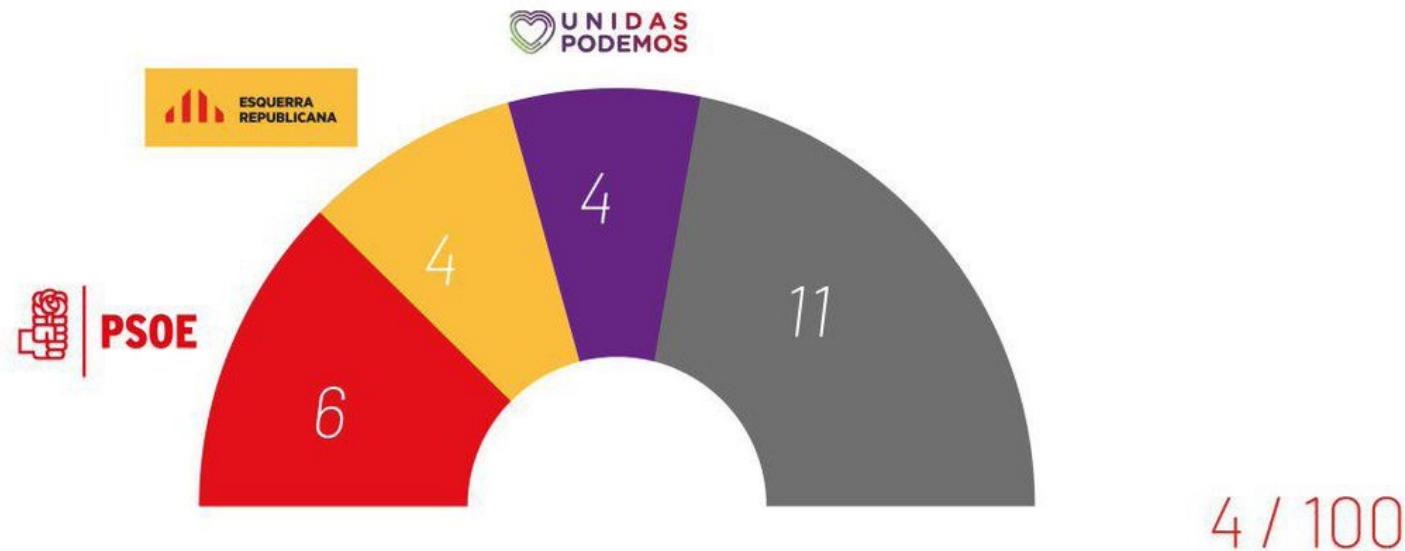
Número de escaños

Introduction

#Pactosdelainfamia

CASTELLDEFELS

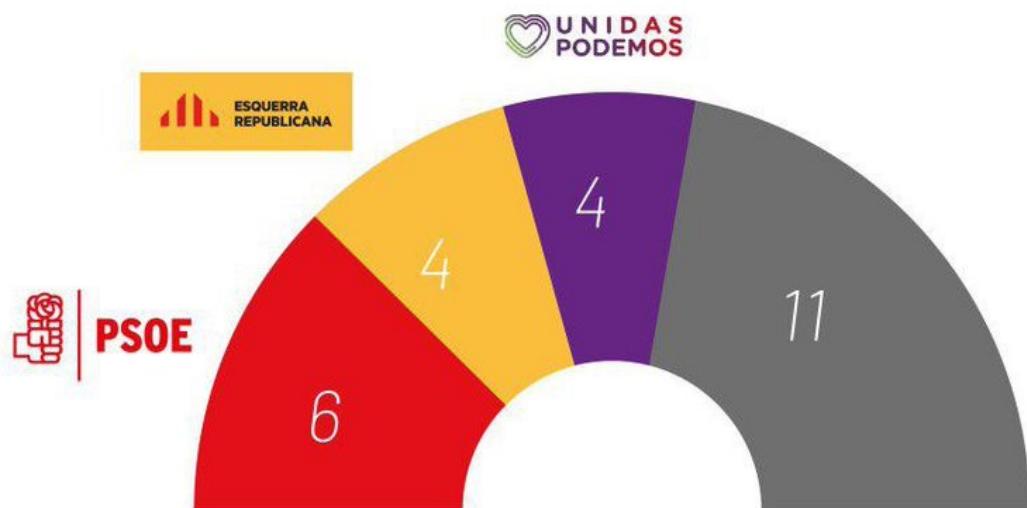
Cs acusa al PSOE de evitar que “gobierne el constitucionalismo, y eso que ha ganado las elecciones”.



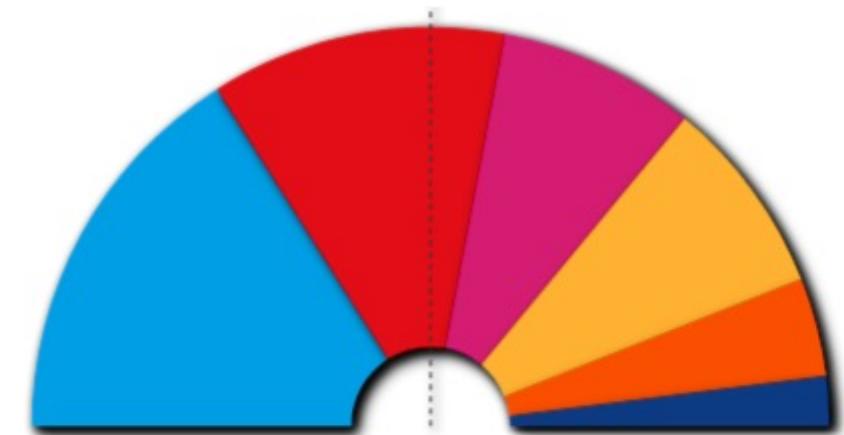
Introduction

#Pactosdelainfamia

CASTELLDEFELS



4 / 100



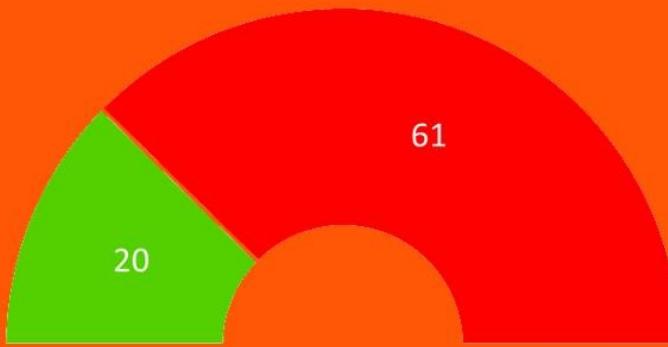
Partit	Vots	%	Regidors
PP	7.904	29,28%	8
PSC-CP	5.430	20,12%	6
MOVEM-ECP-ECG	4.330	16,04%	4
ERC - AM	4.290	15,89%	4
Cs	1.777	6,58%	2
JUNTS	1.525	5,65%	1
PRIMÀRIES	680	2,52%	0
SOM - AMUNT	299	1,11%	0
VOX	271	1,00%	0
CCSTF	144	0,53%	0
UCLL	123	0,46%	0

Introduction

GRÁFICOS DE LA INFAMIA

@PEDRODANIELPG

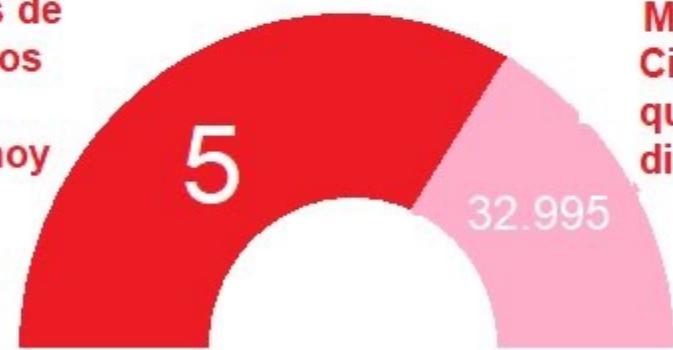
CORRECTOS: 20 (24.69%)
INCORRECTOS: 61 (75.31%)

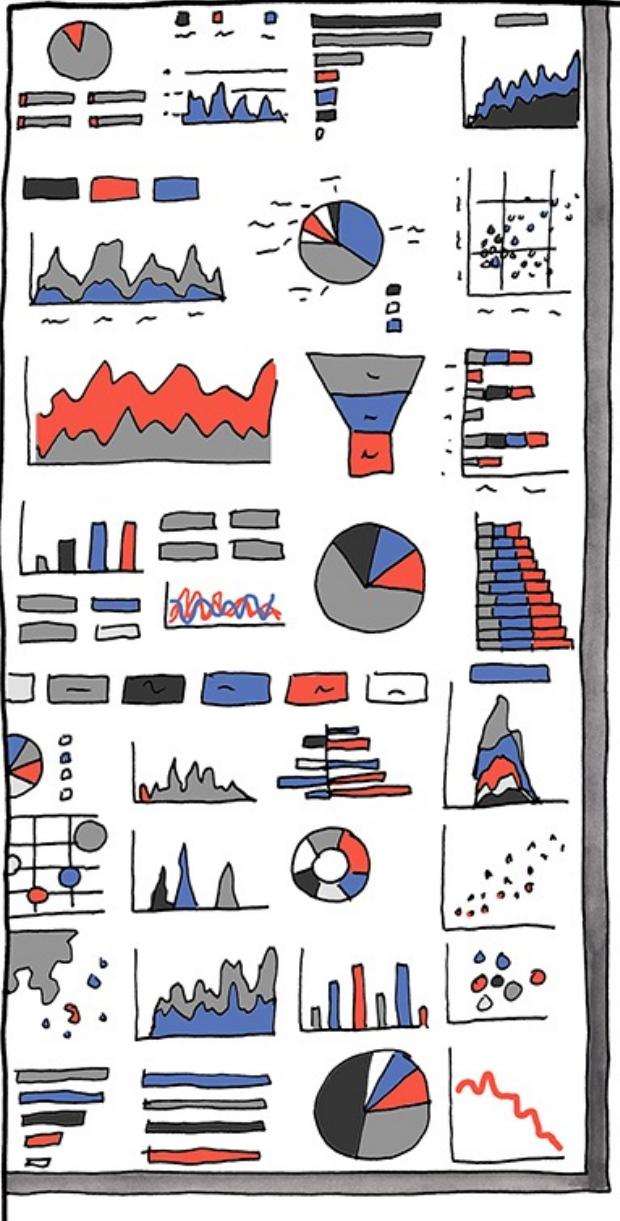


Introduction

La crisis de Ciudadanos en gráficos

Miembros de Ciudadanos que han dimitido hoy





OUR NEW DASHBOARD HAS ALL OF THE DIFFERENT KPI'S WE CAN TRACK NOW.



WHAT'S THAT KPI TRENDING TO ZERO?

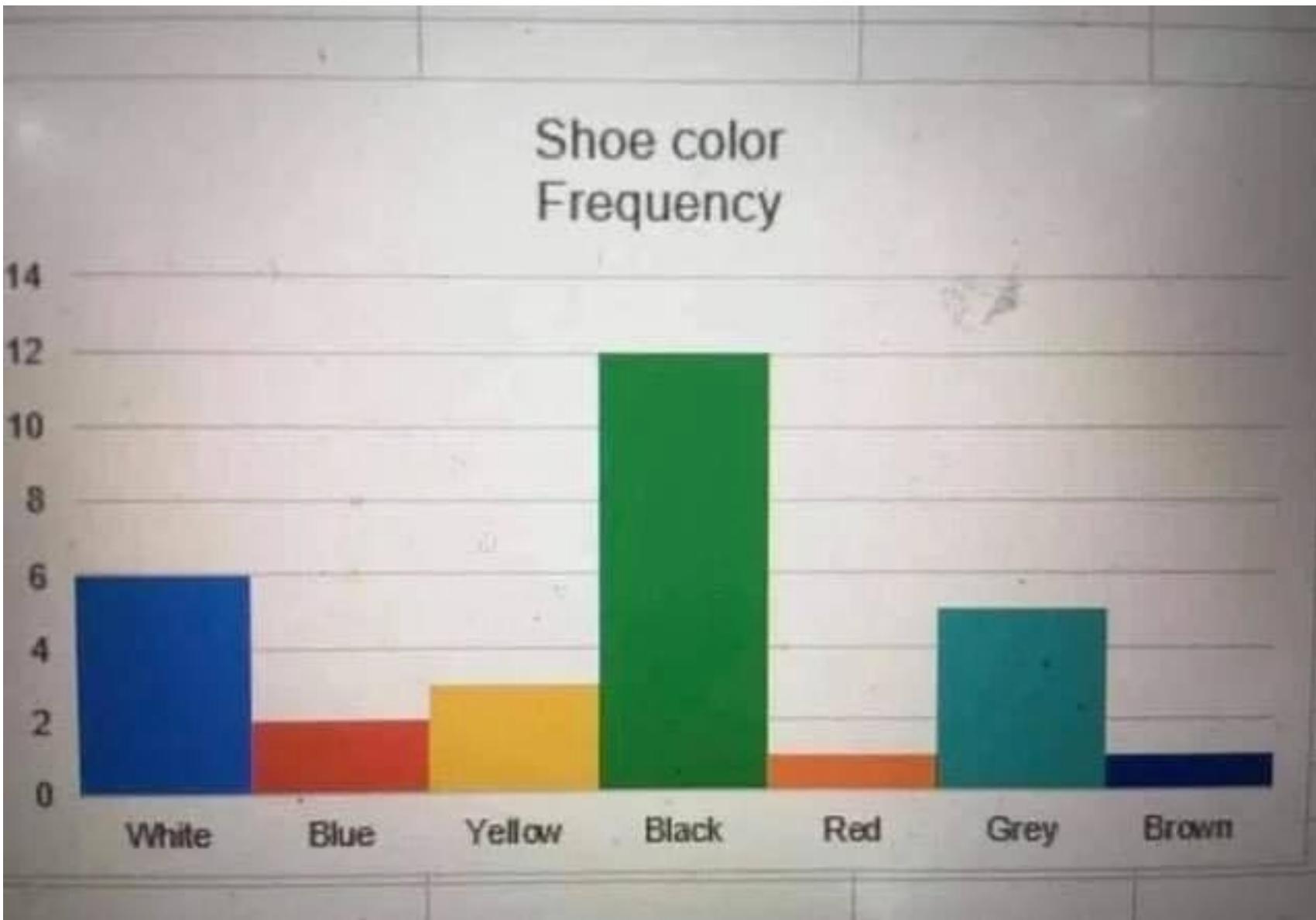


IT MEASURES HOW WELL WE UNDERSTAND THEM ALL.



TOM FISH BURNE

Introduction



Introduction

Moving a picture in Microsoft Word

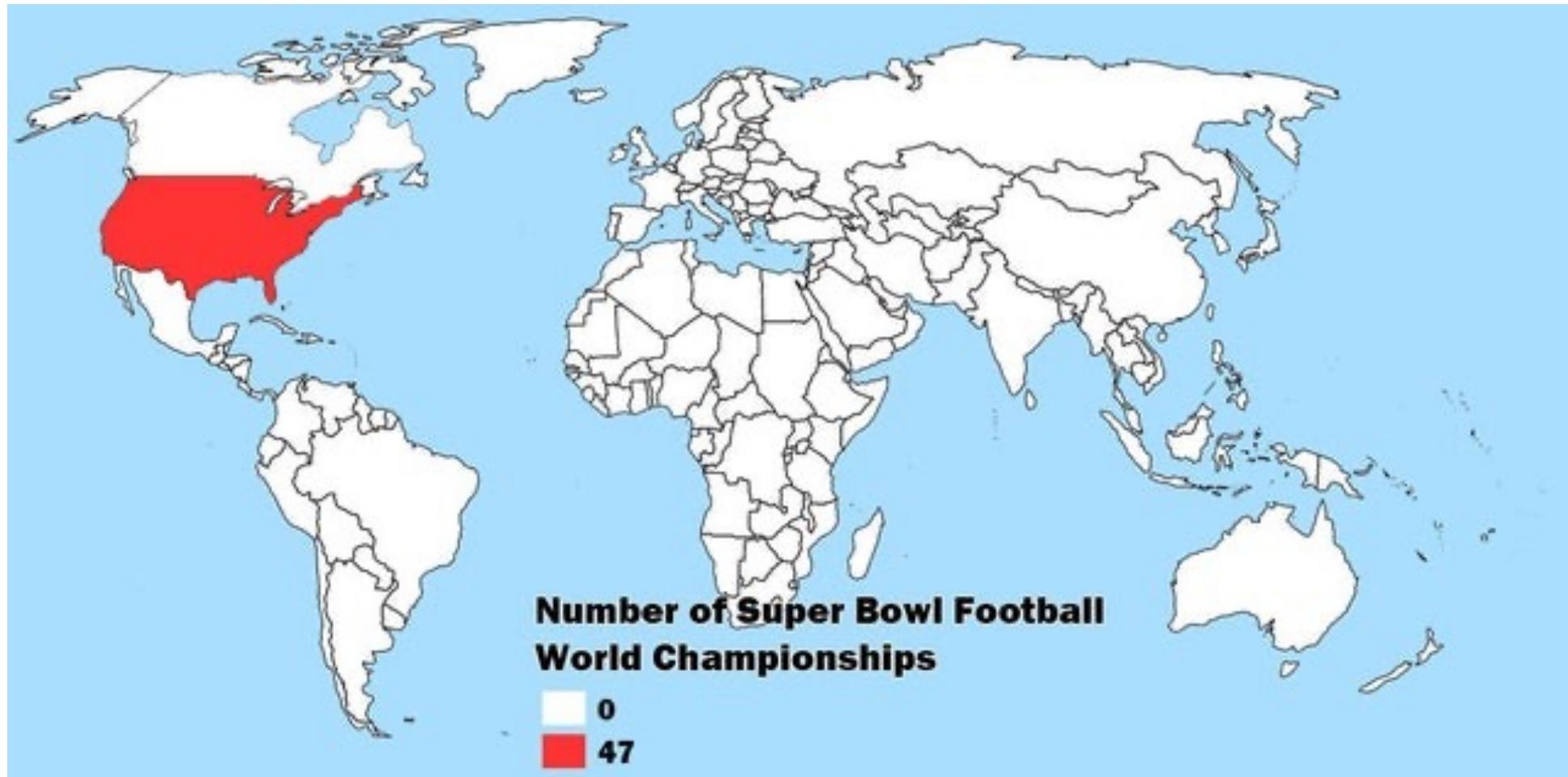
■ You
the



**mess up
whole document**

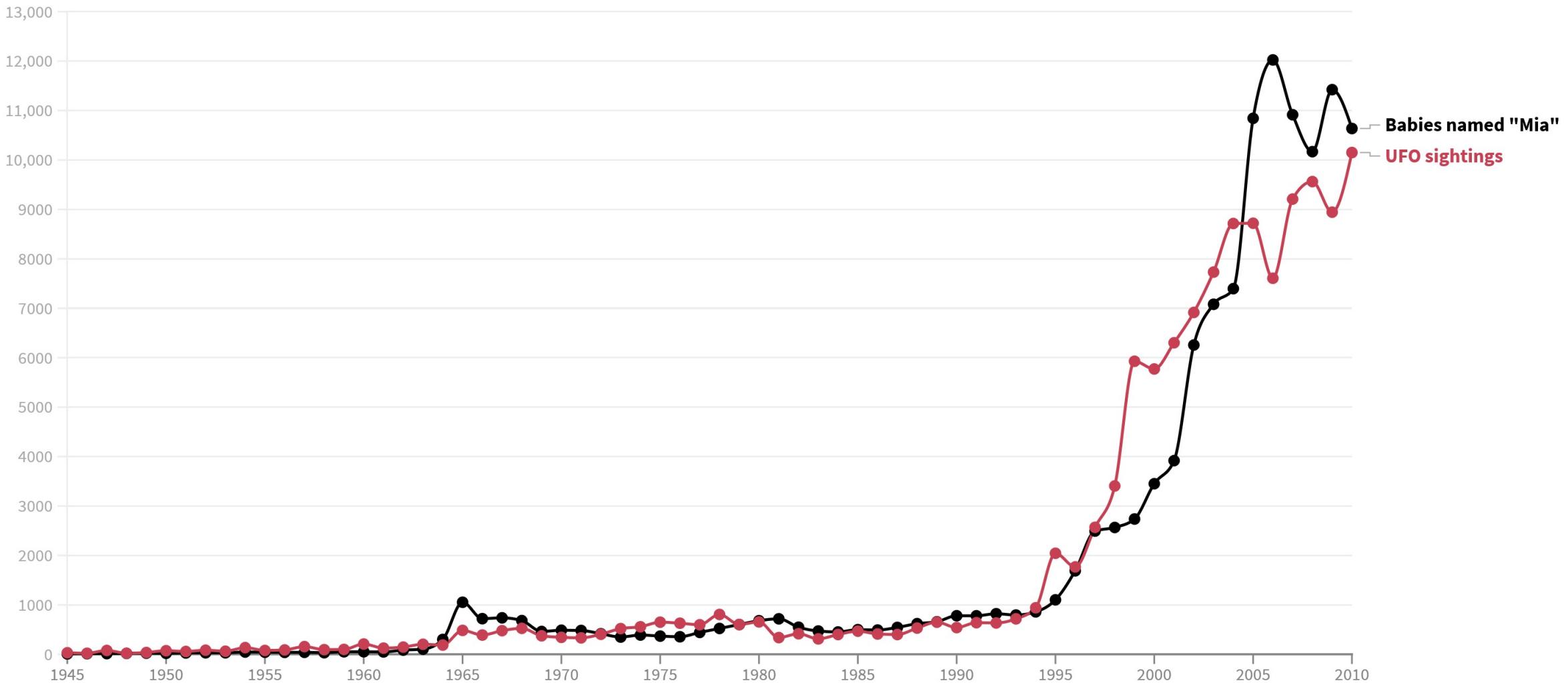
■ It
actually
does
what
you
want

Introduction



Correlation does not imply Causation

Number of **Babies named "Mia"** in the U.S. correlates with Number of **UFO Sightings** Around the World



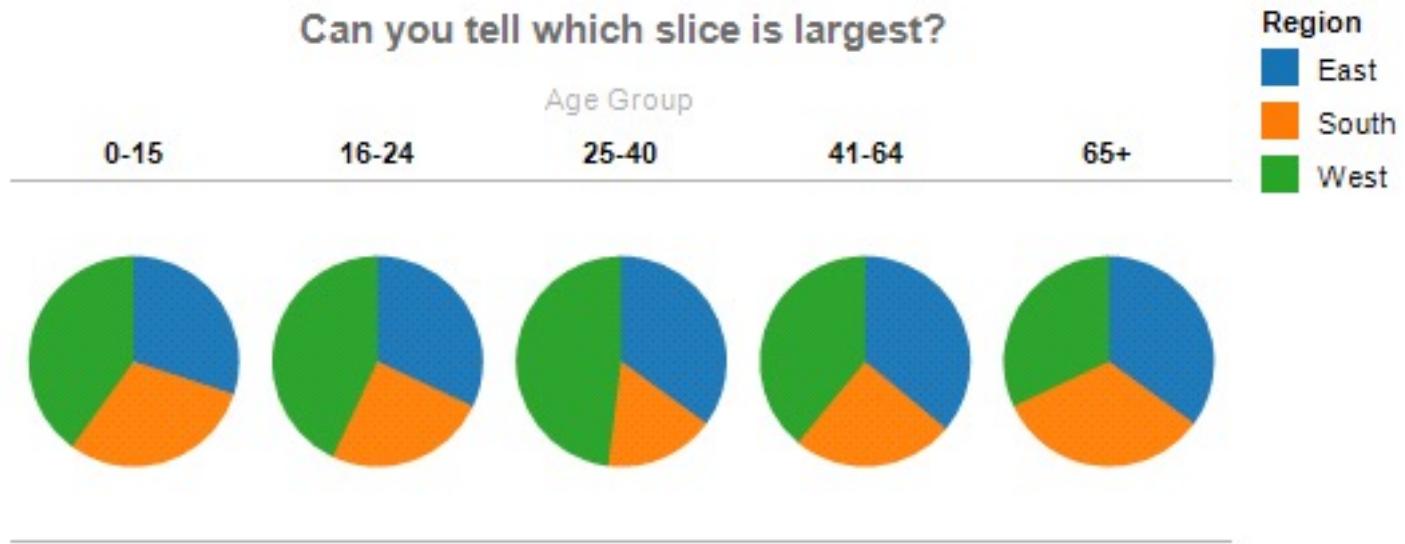
Data: NUFORC, SSA • Number of UFO sightings is shown by factor 5.2
Visualization by Cédric Scherer | #30DayChartChallenge 2021 | Day 13: Correlation

Takeaways

- Through the course, we will learn...

Takeaways

Can you tell which slice is largest?



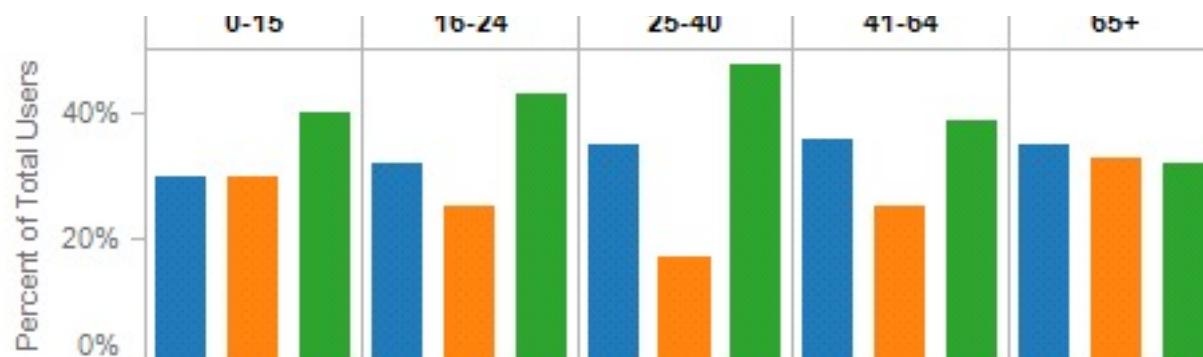
This is the same information represented as bars.
Comparisons are much easier.



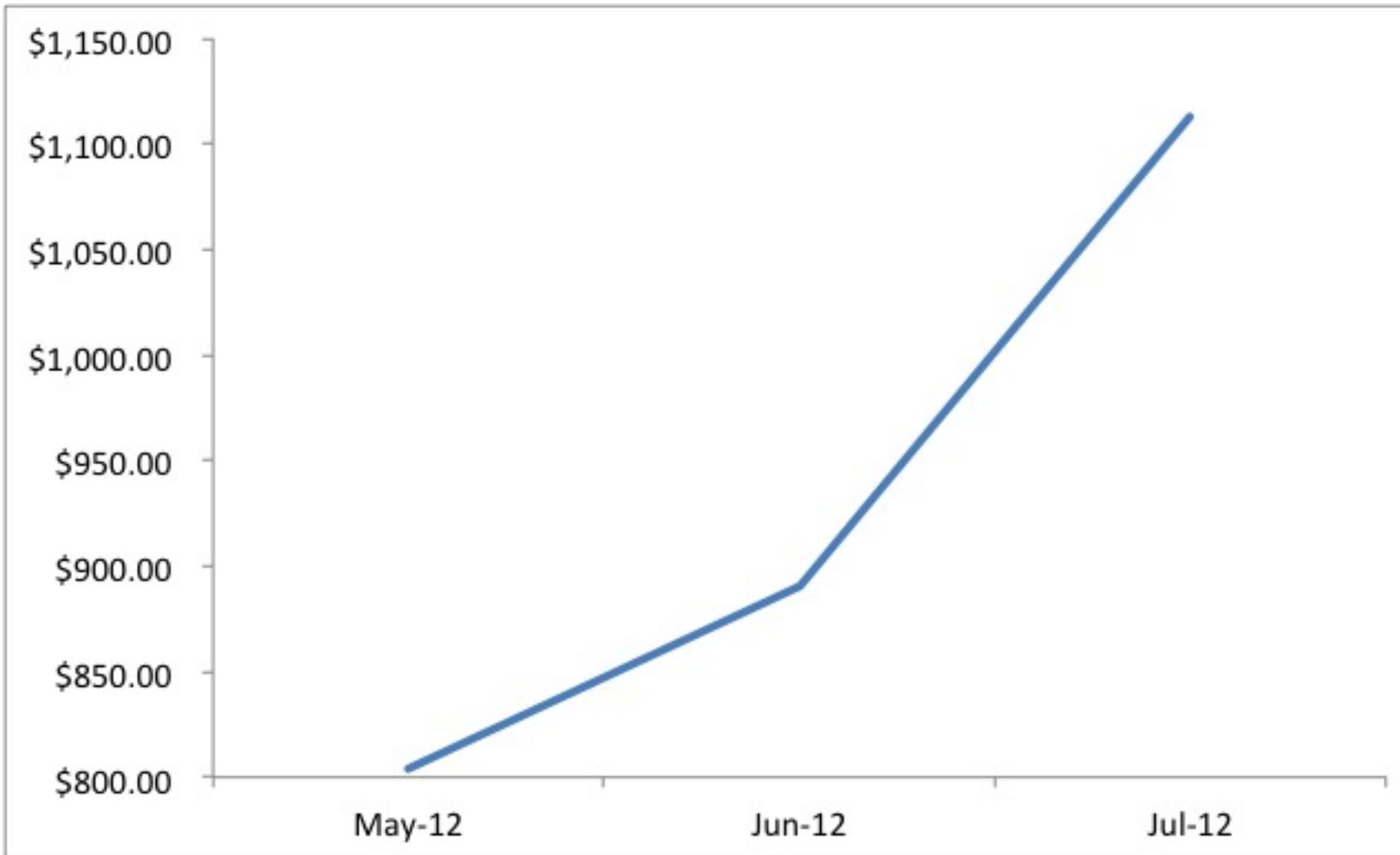
Takeaways



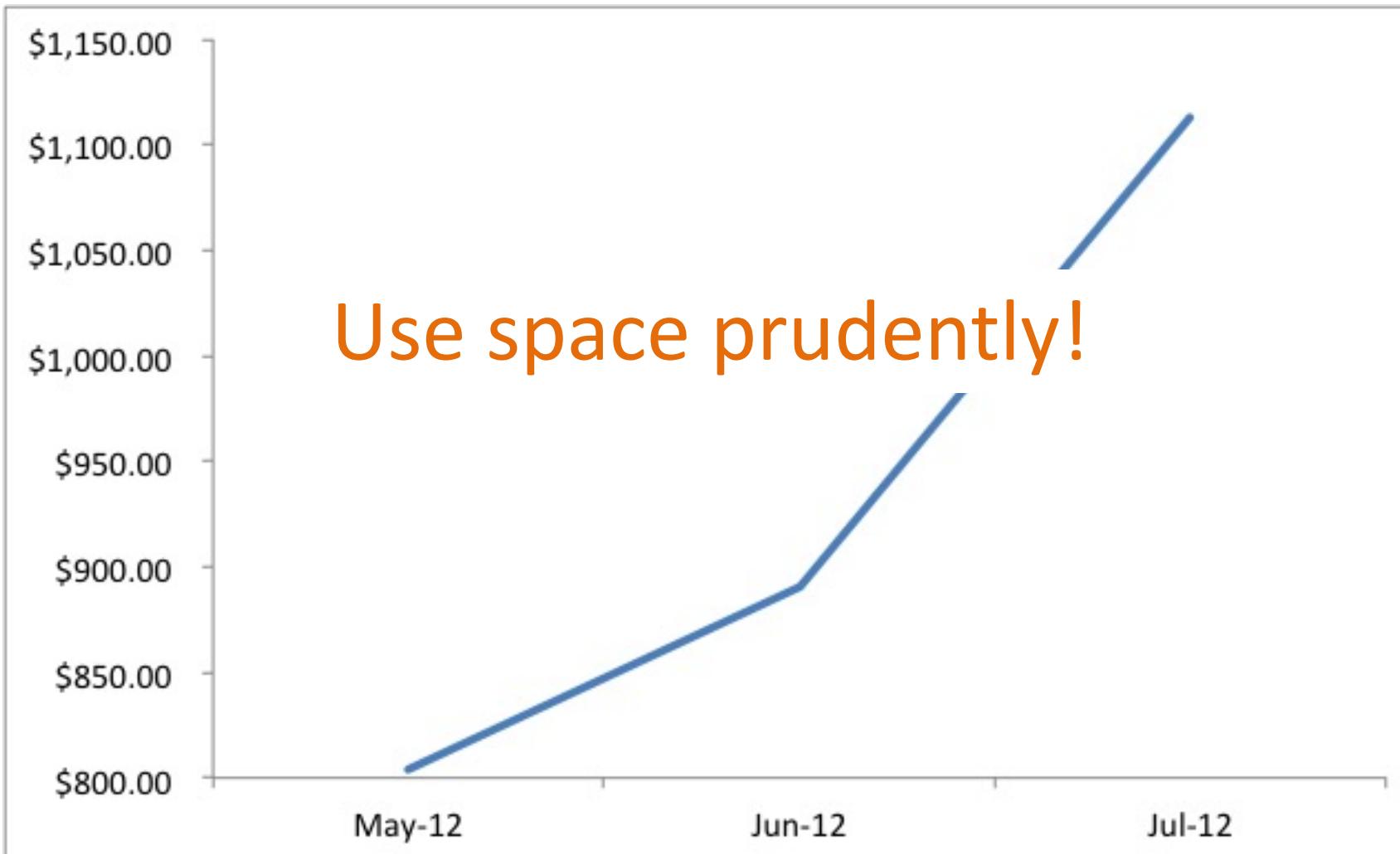
Not all visualization techniques
are equally [or even properly]
perceived!



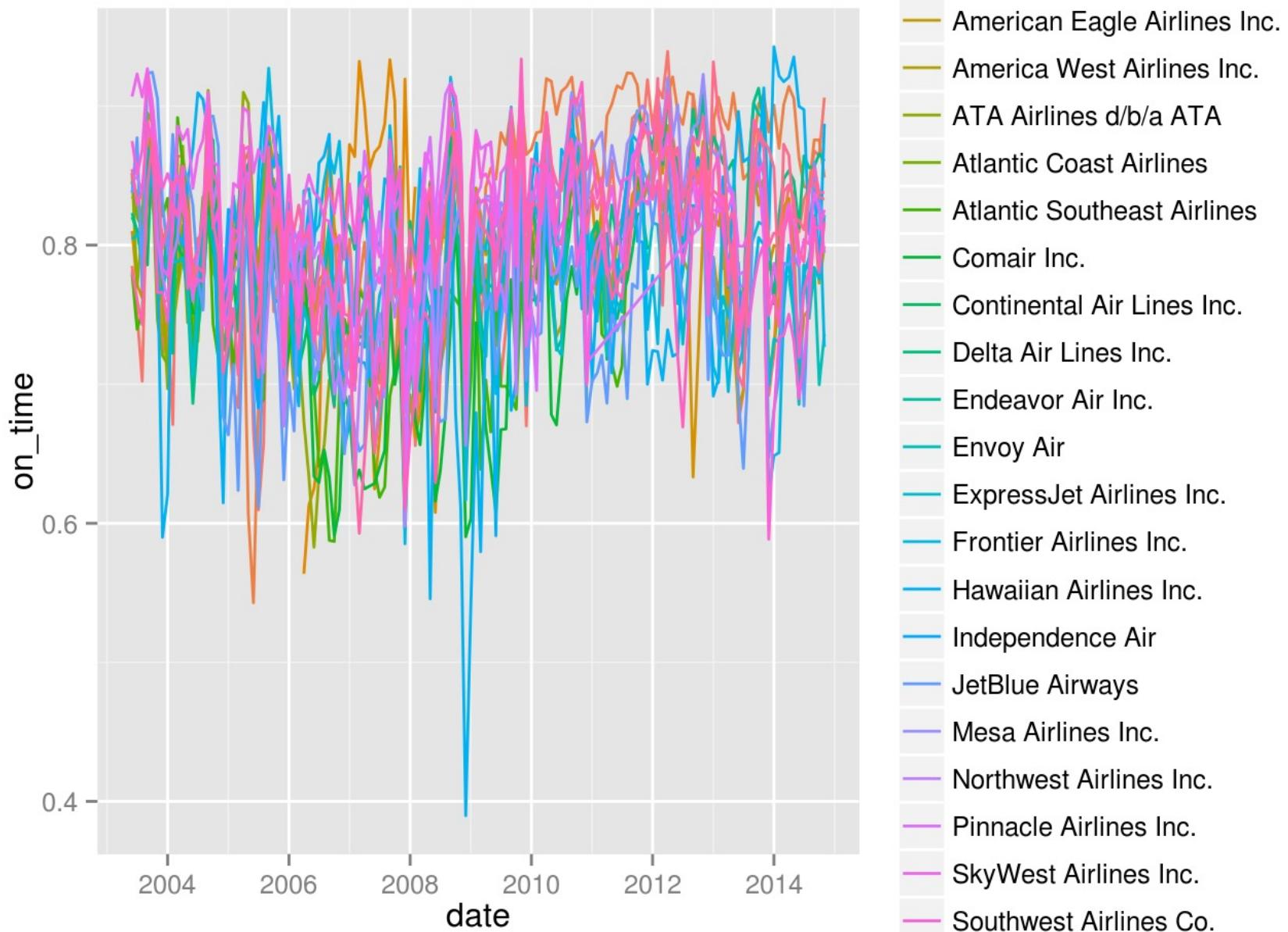
Takeaways



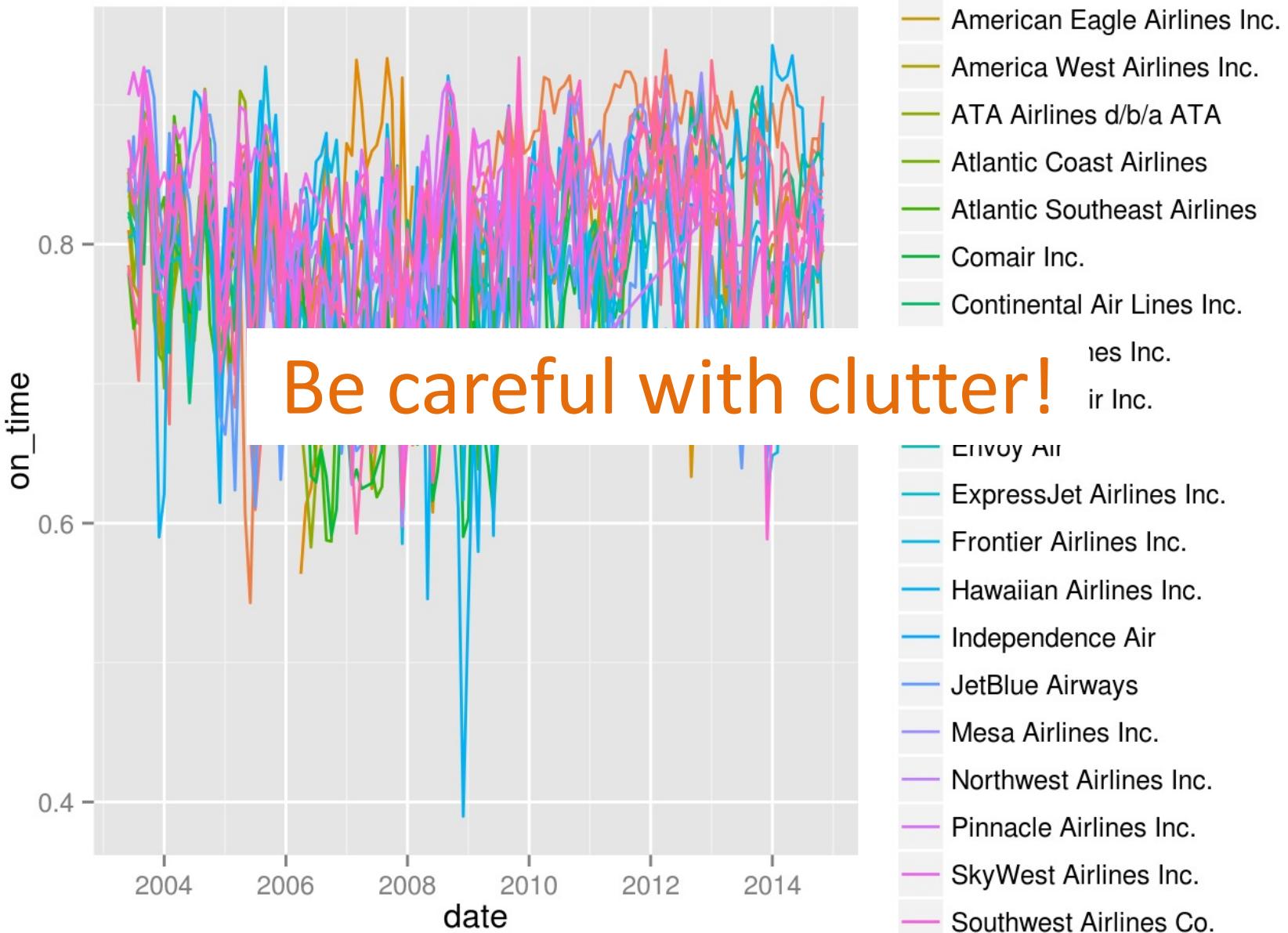
Takeaways



Takeaways



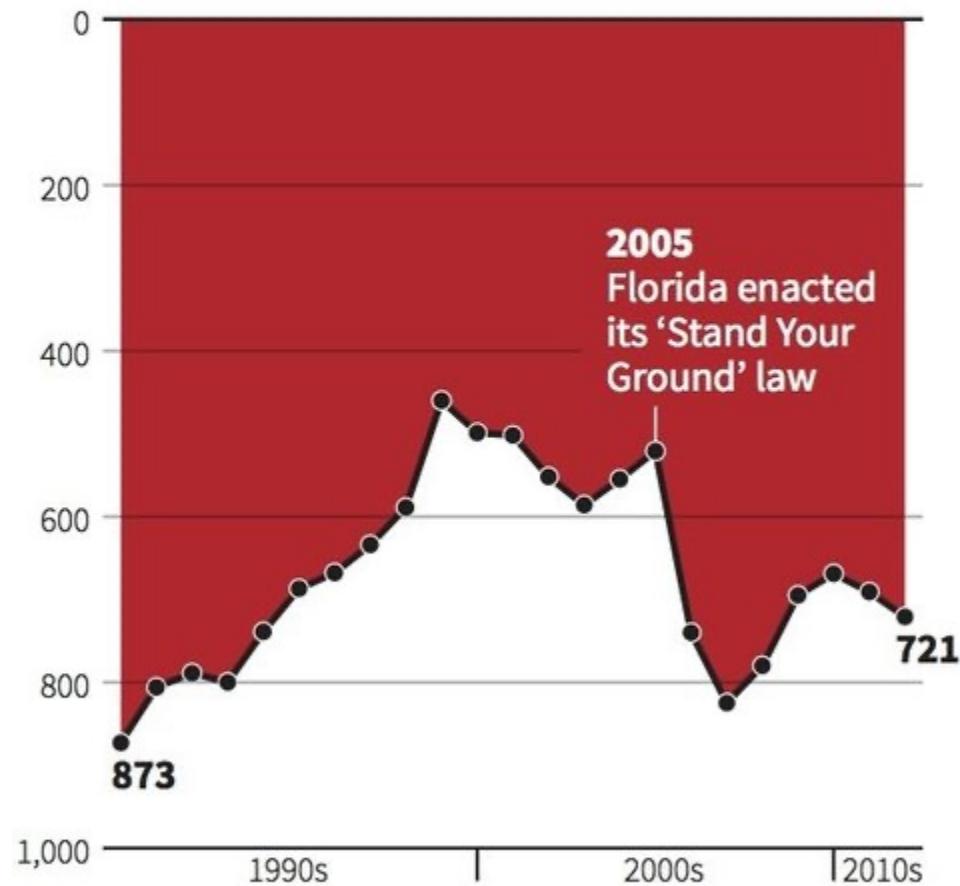
Takeaways



Takeaways

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

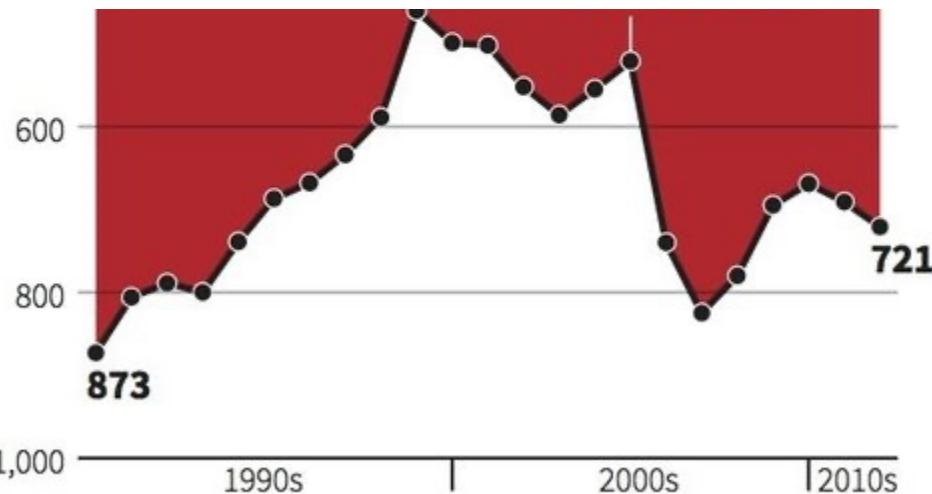
Takeaways

Gun deaths in Florida

Number of murders committed using firearms



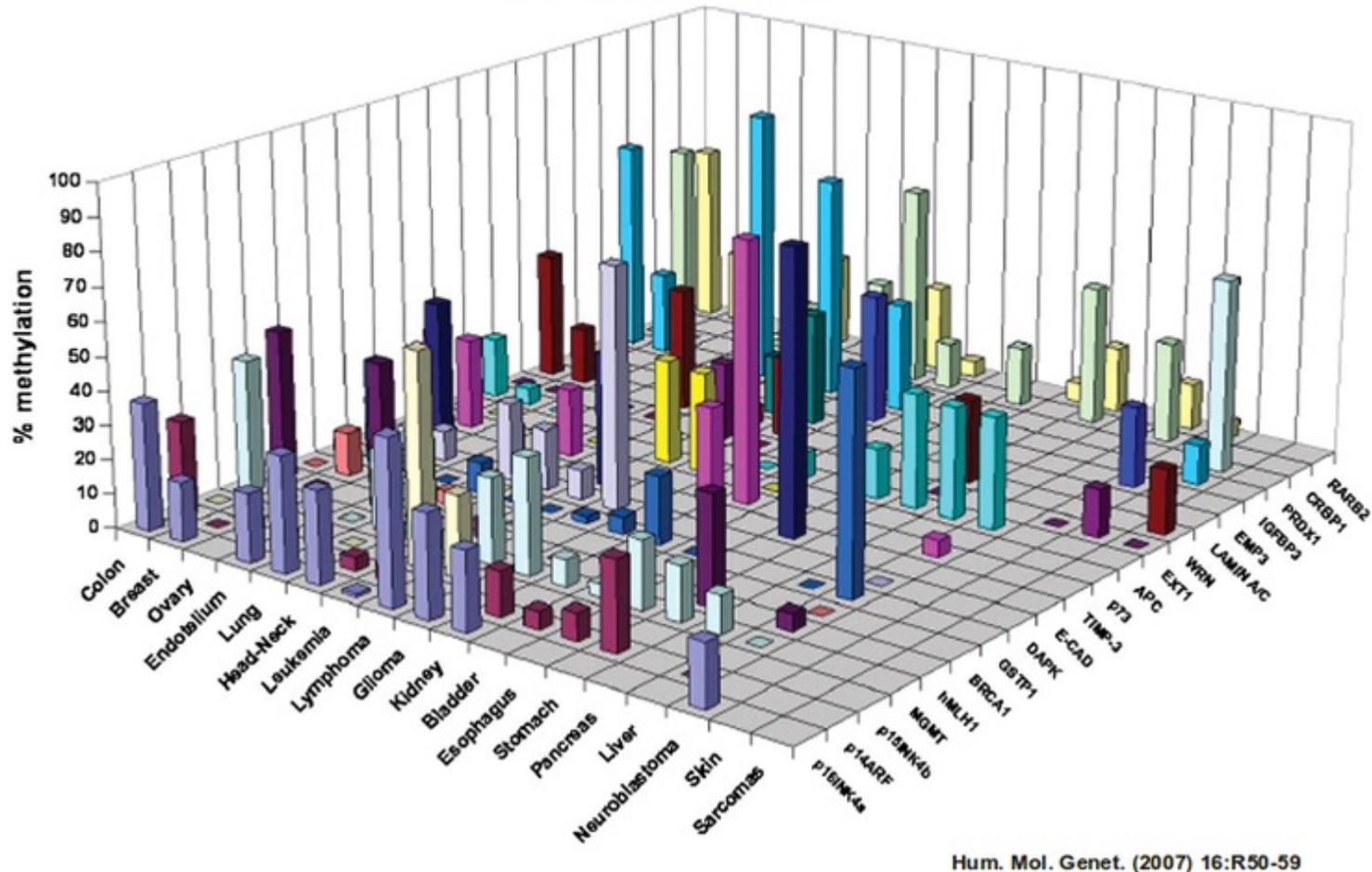
Use standard axes!



Source: Florida Department of Law Enforcement

Takeaways

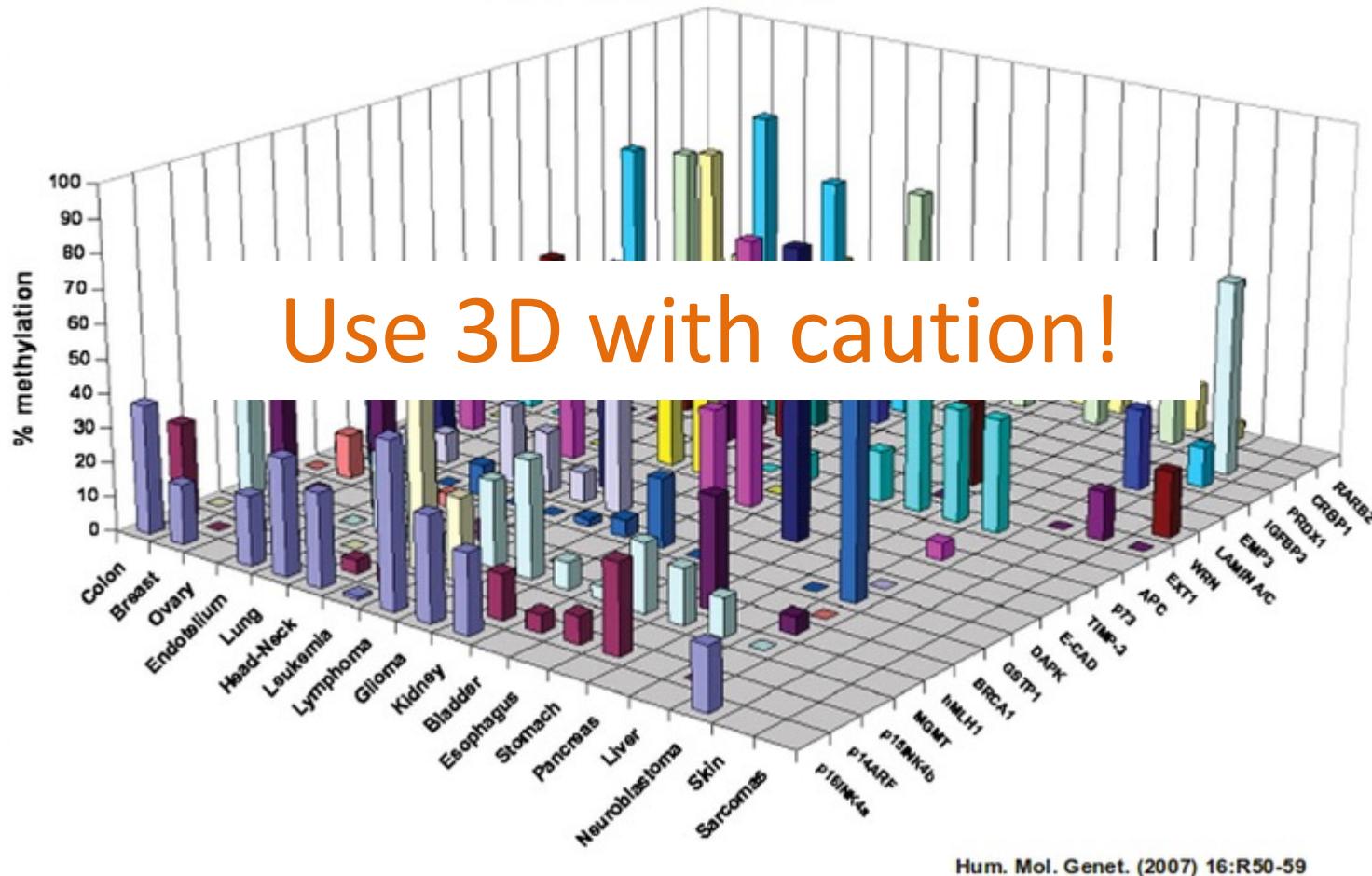
A CpG Island Hypermethylation Profile of Human Cancer



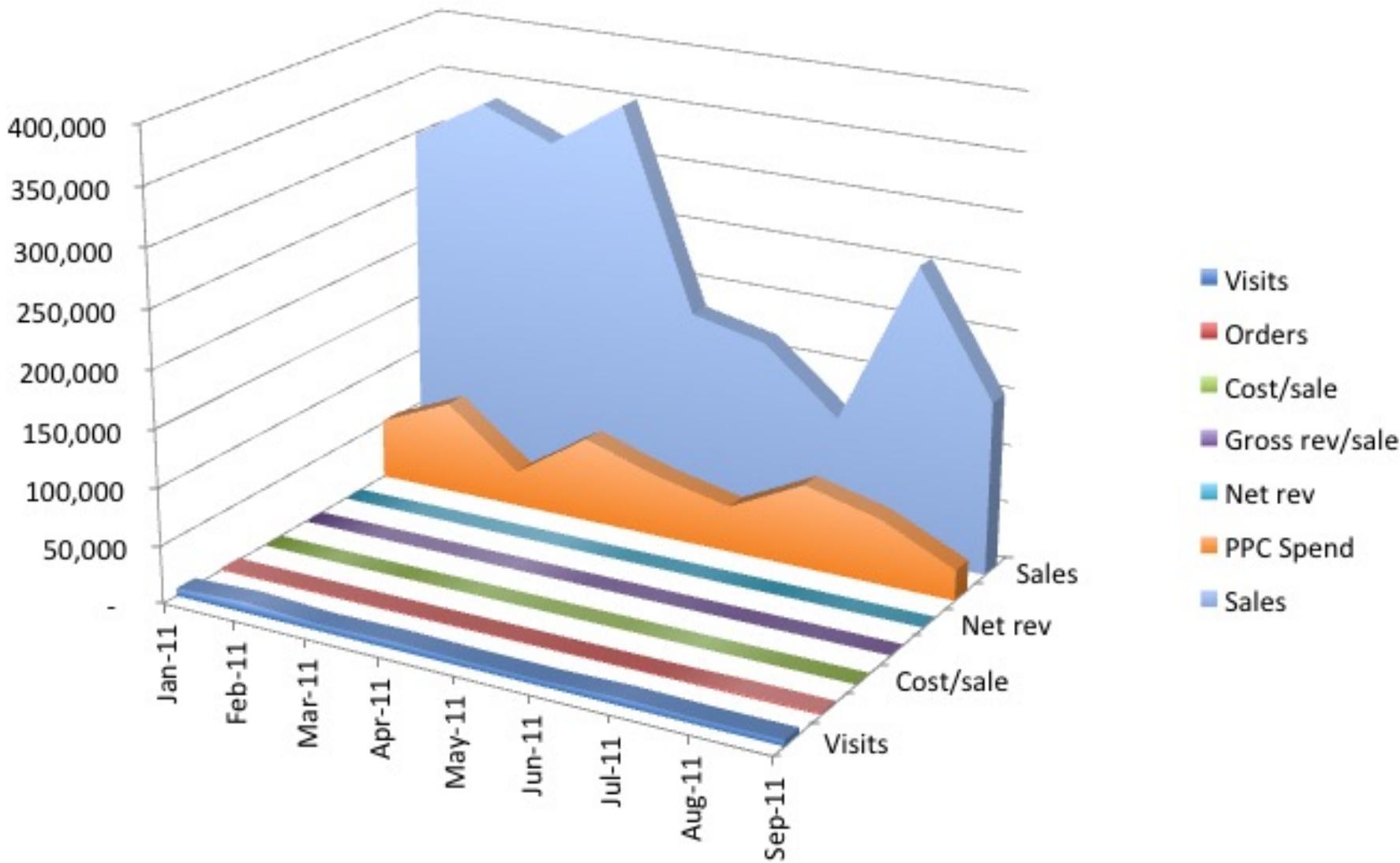
Hum. Mol. Genet. (2007) 16:R50-59

Takeaways

A CpG Island Hypermethylation Profile of Human Cancer



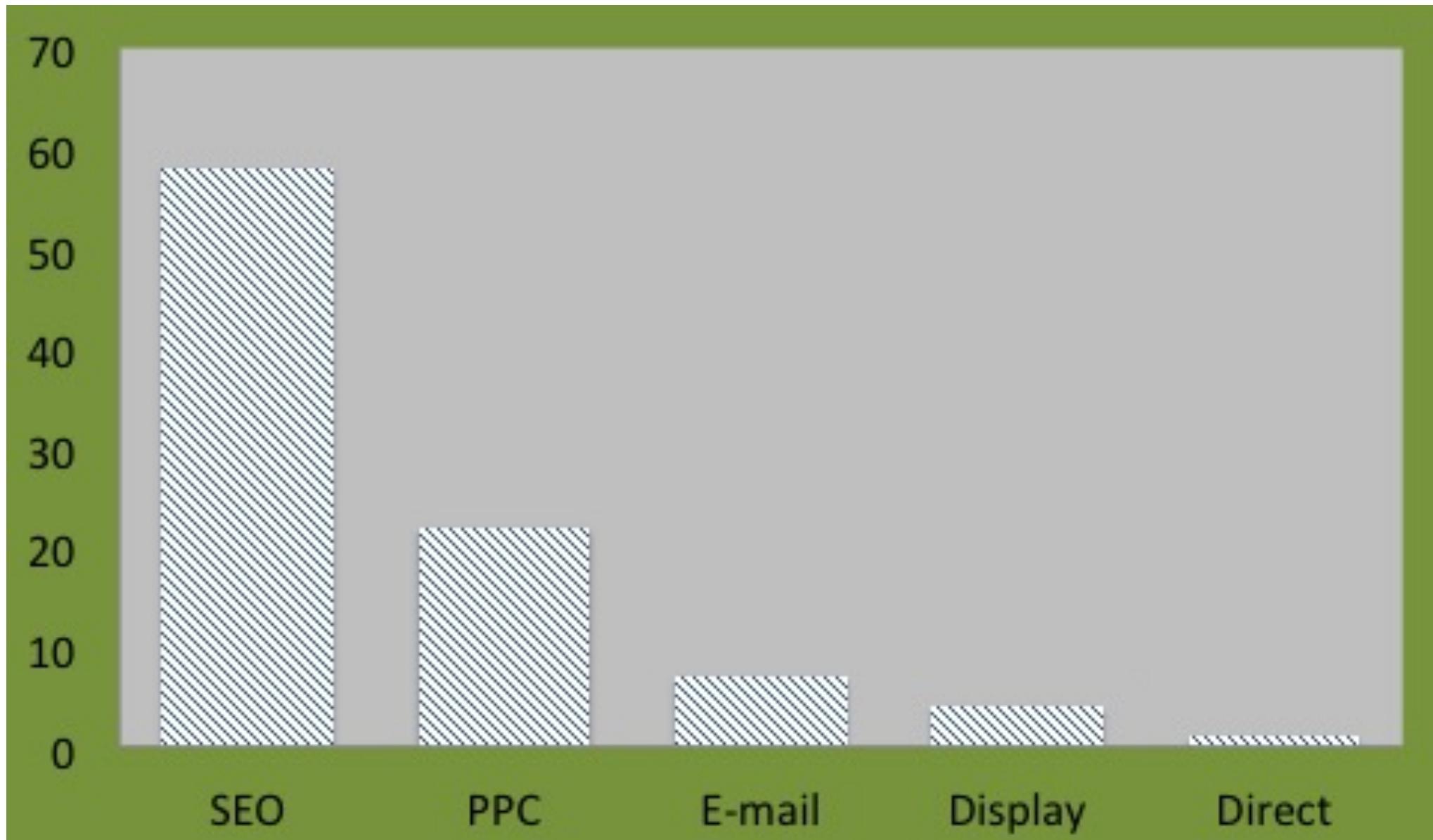
Takeaways



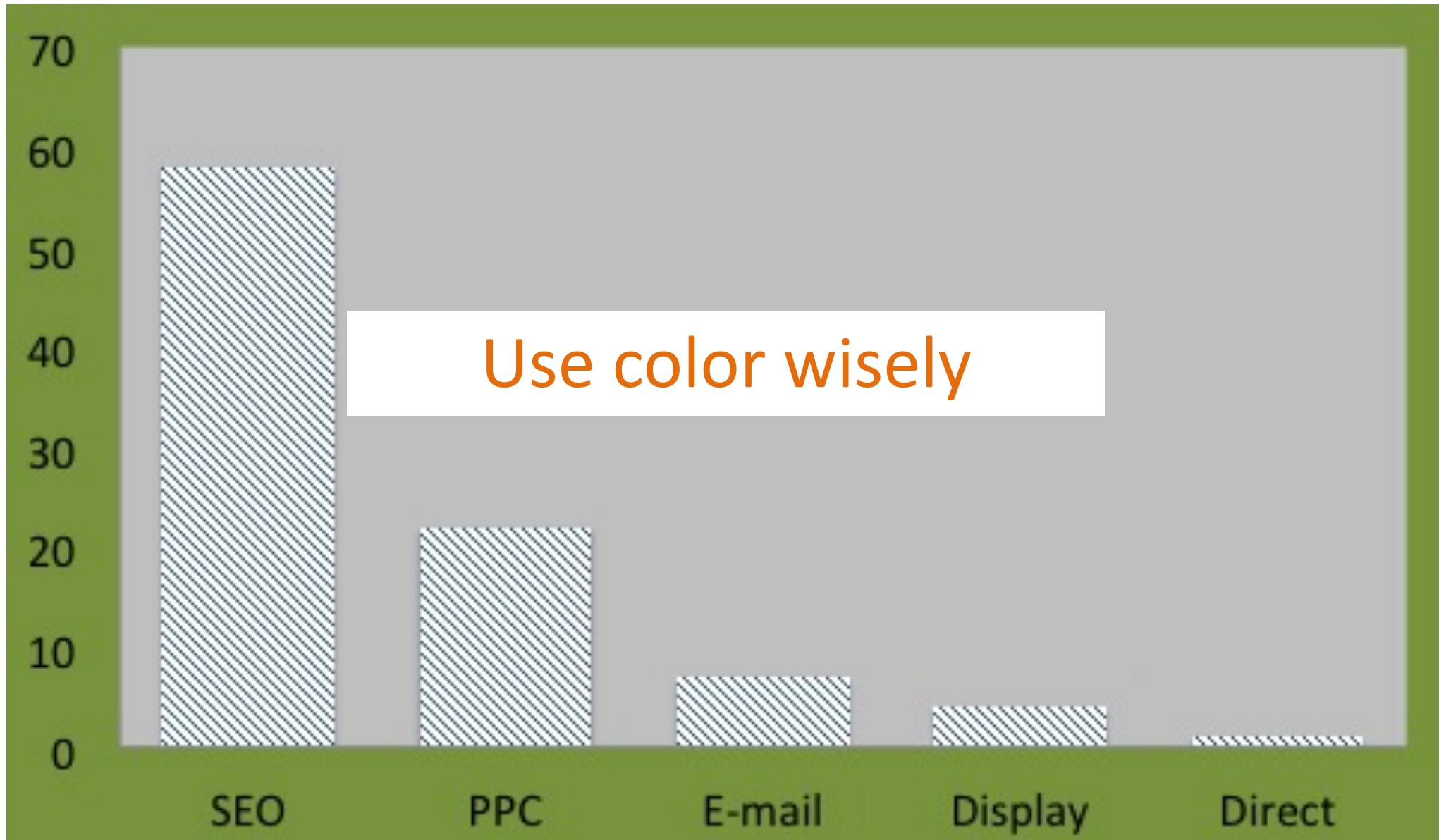
Takeaways



Takeaways



Takeaways



Outline

- Visualization. The basics
- History
- Visualize: What, why, and how
- The value of visualization
- The visualization mantra
- Visualization Areas
- Visual variables

Outline

- **Visualization. The basics**
- History
- Visualize: What, why, and how
- The value of visualization
- The visualization mantra
- Visualization Areas
- Visual variables

Visualization. The basics

Yeah! We are finally doing visualization...



Visualization. The basics

... errmmm... no. This is not visualization (this is CG)



Visualization. The basics

- Definition (by T. Munzner)
 - Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively
 - People and tasks
 - Augment the capabilities of the human rather than simply replace them by some decision making method that is computational
 - Some human in the loop that needs the details
 - When you do not know that the problem is in advance

Visualization. The basics

- Some human in the loop that needs the details
 - When you do not know that the problem is in detail in advance
- Exploratory data analysis
- Presentation of results
- Sometimes helping the developers build an automatic model
- Understand the situation
- Refine an algorithm
- Build trust among users
- Issues of vis regarding to use something that is external to users in order to swap cognitive tasks for perceptual ones
- Intended task, measurable definitions

Visualization. The basics

- Visualization is related to **understanding the underlying data**
 - Cognitive process
 - Gain understanding on data
 - Helping the user (human) to understand data

Visualization. The basics

- Visualization is related to **understanding the underlying data**
 - Cognitive process
 - Gain understanding on data
 - Helping the user (human) to understand data
- Visual representations help us understand
 - How?

Visualization. The basics

- Visual representations help us understand. How?
 - By augmenting human capabilities
 - Instead of replacing people with computational decision-making methods
 - Computer-based visualization systems provide visual representations of datasets designed to help **people** carry out tasks more effectively
 - Putting **human in the loop** is fundamental

Visualization. The basics

- Human in the loop needs the details [& no trusted automatic solution exists]
 - Doesn't know exactly what questions to ask in advance
 - Exploratory data analysis
 - speed up through human-in-the-loop visual data analysis
 - Present known results to others
 - Stepping stone towards automation
 - Before model creation to provide understanding
 - During algorithm creation to refine, debug, set parameters
 - Before or during deployment to build trust and monitor

Visualization. The basics

- If the result is a calculation, then you should probably not be using visualization at all.
 - If you know what you are looking for, you probably do not need the visualization
- “*The purpose of computing is insight, not numbers*”

Richard Hamming, 1971

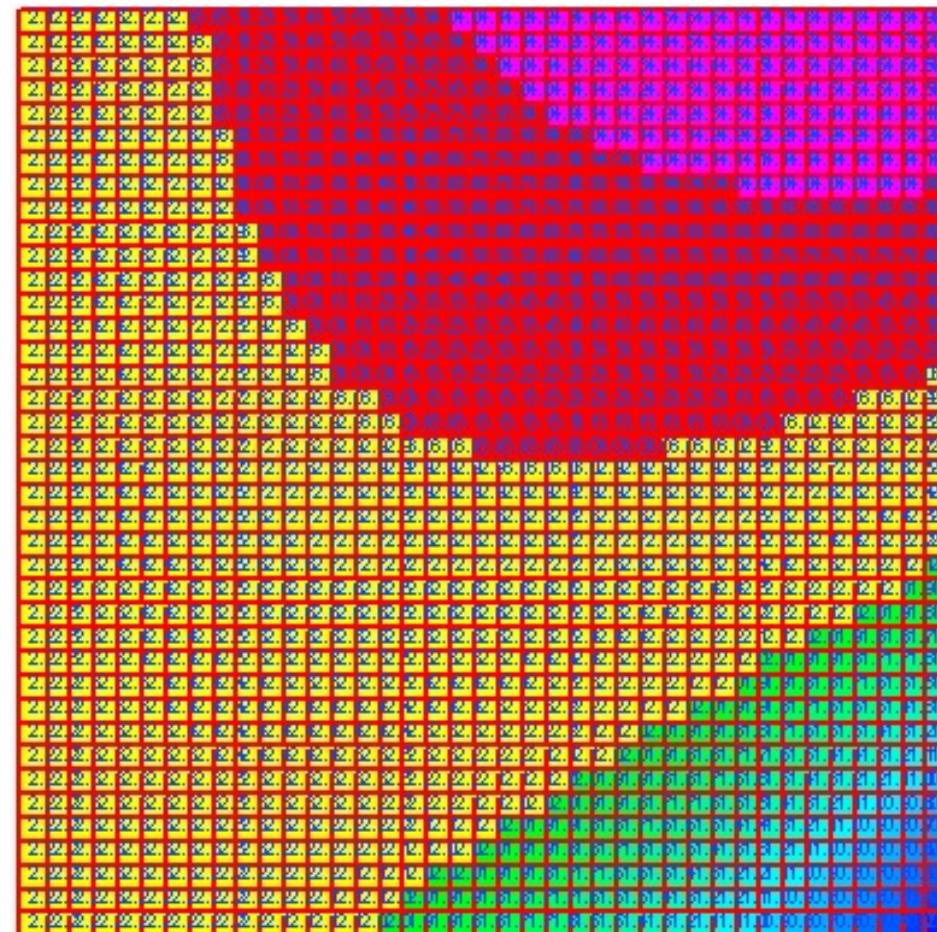
Visualization. The basics

- Try to get the domain from this:

1.3	1.8	2.2	2.6	3.0	3.3	3.6	3.9	4.1	4.3	4.4	4.5
1.3	1.8	2.2	2.5	2.9	3.2	3.4	3.6	3.8	4.0	4.1	4.1
1.4	1.8	2.1	2.5	2.7	3.0	3.2	3.4	3.5	3.7	3.7	3.8
1.4	1.7	2.1	2.4	2.6	2.8	3.0	3.2	3.3	3.3	3.4	3.4
1.4	1.7	2.0	2.3	2.5	2.7	2.8	2.9	3.0	3.0	3.0	3.0
1.4	1.7	2.0	2.2	2.4	2.5	2.6	2.7	2.7	2.7	2.7	2.6
1.5	1.7	1.9	2.1	2.3	2.4	2.4	2.5	2.5	2.4	2.3	2.2
1.5	1.7	1.9	2.0	2.1	2.2	2.2	2.2	2.2	2.1	2.0	1.9
1.5	1.7	1.8	1.9	2.0	2.1	2.0	2.0	1.9	1.8	1.7	1.5
1.5	1.7	1.8	1.9	1.9	1.9	1.9	1.8	1.7	1.5	1.3	1.1
1.6	1.7	1.7	1.8	1.8	1.7	1.7	1.5	1.4	1.2	1.0	0.7
1.6	1.7	1.7	1.7	1.7	1.7	1.6	1.5	1.3	1.1	0.9	0.6

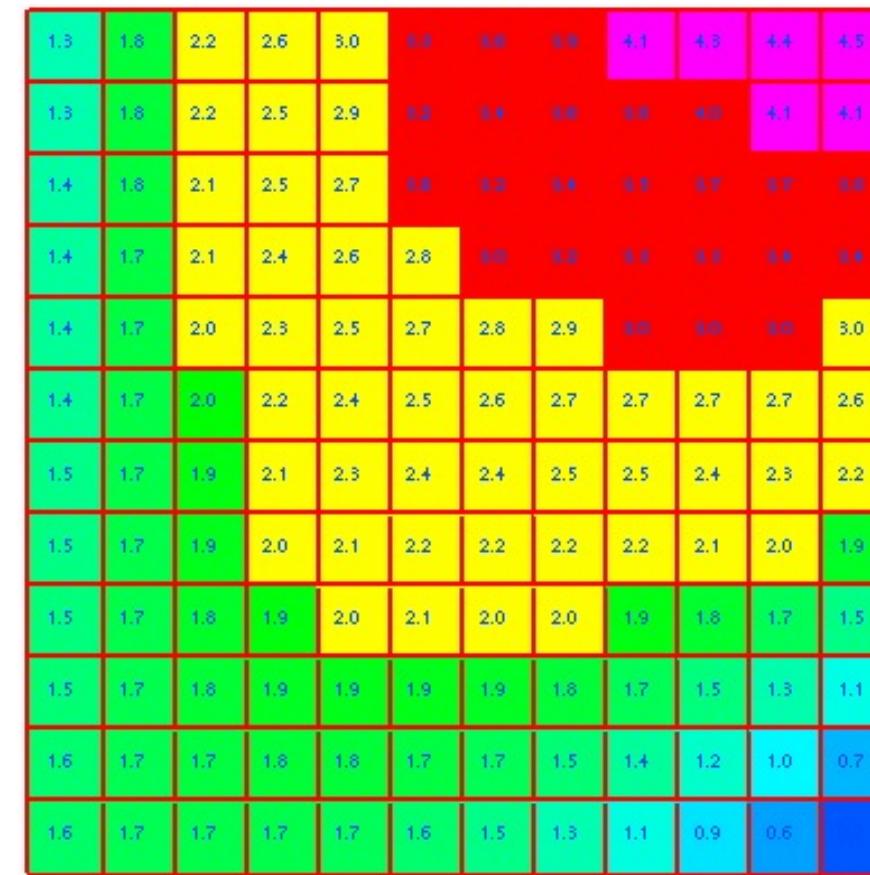
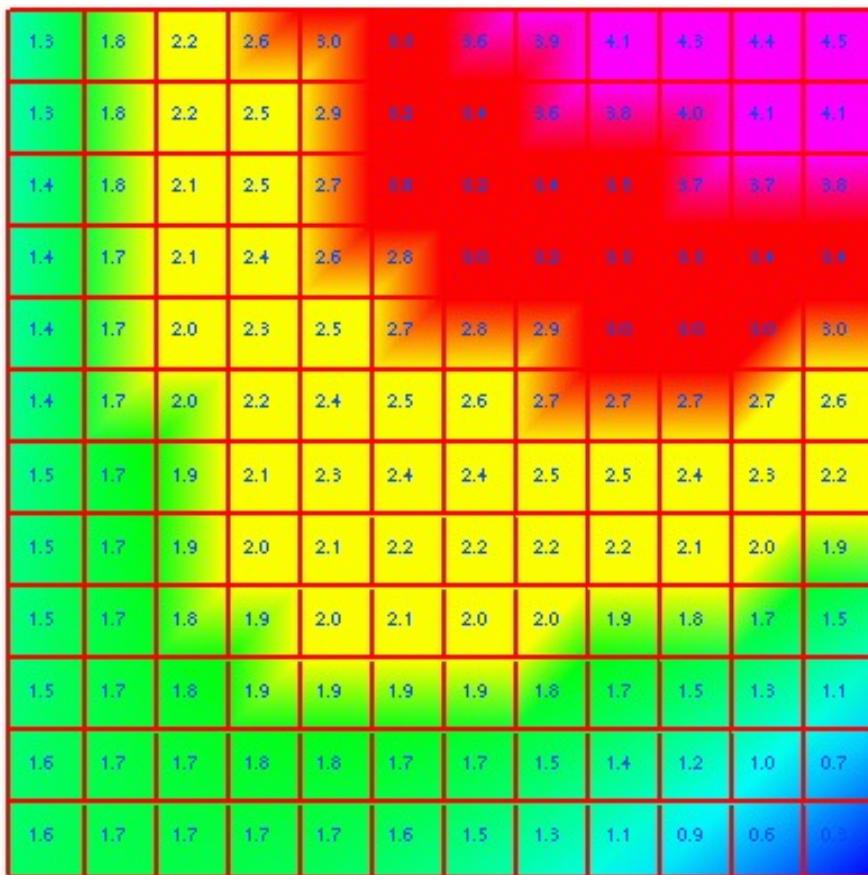
Visualization. The basics

- What if we use images?



Visualization. The basics

- And even simplify?

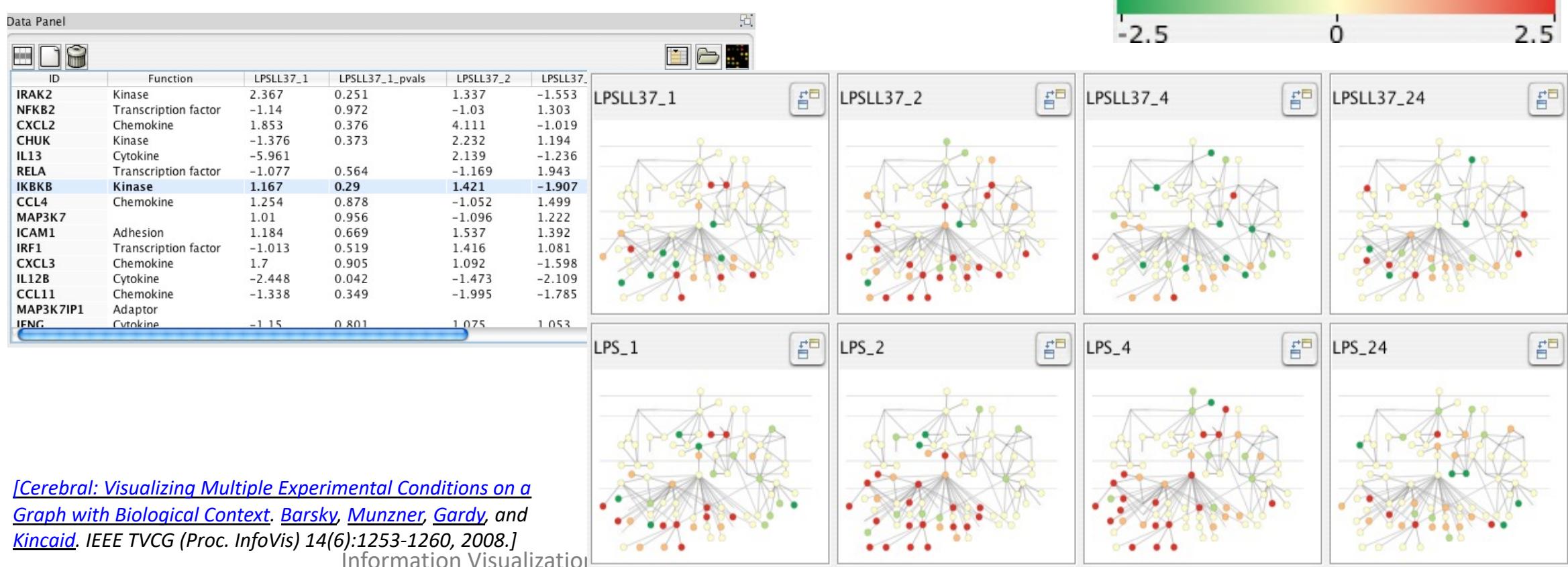


Visualization. The basics

- Computer-based visualization systems provide **visual representations of datasets** designed to **help people** carry out tasks more **effectively**
 - Visual representations → human visual system
 - Datasets
 - People
 - Effective

Visualization. The basics

- Visual [external] representations. Why depending on vision?
 - Replace cognition with perception



Visualization. The basics

- **Representations of datasets**
 - Summaries lose information, details matter
 - Confirm expected and find unexpected patterns
 - Assess validity of statistical model
 - Famous example, the Anscombe's quartet

Visualization. The basics

Anscombe's Quartet: Raw Data

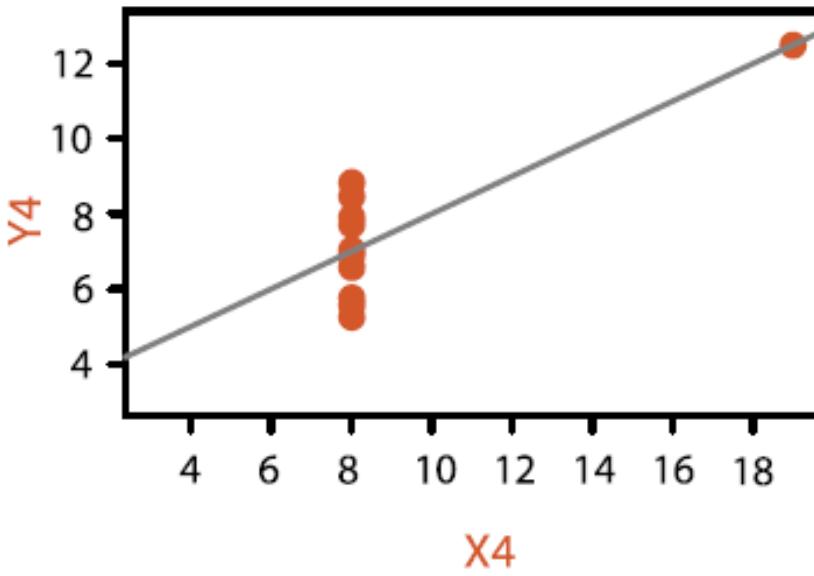
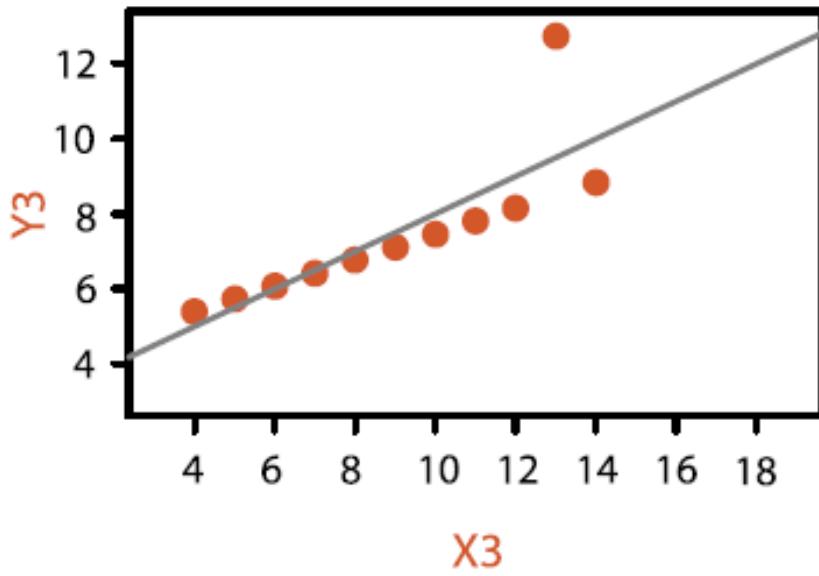
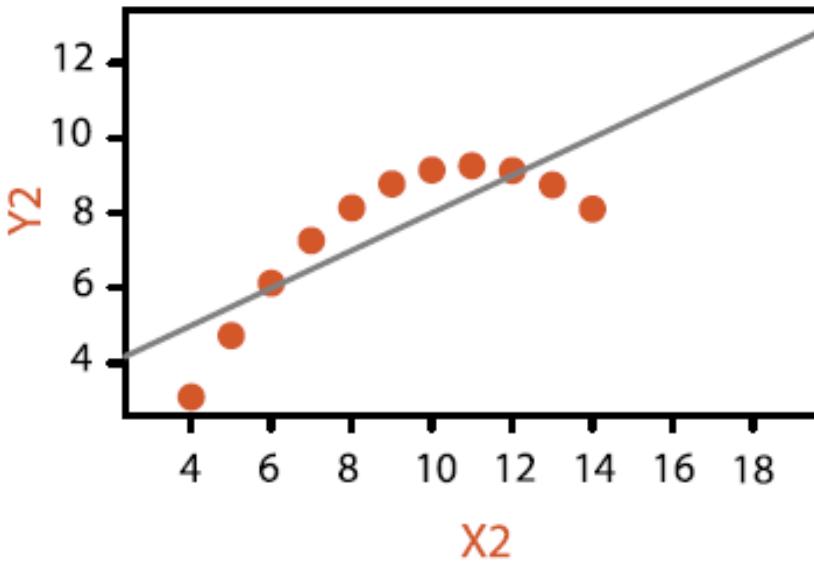
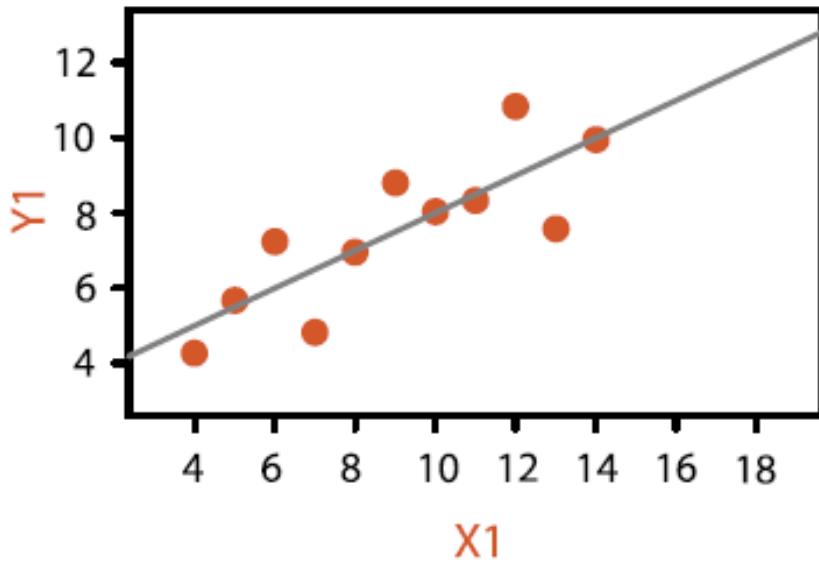
	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	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

Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

Visualization. The basics

- Same statistical properties:
 - Number of observations (n): 11
 - Mean x: 9.0
 - Mean y: 7.5
 - Equation of regression line: $y = 3 + 0.5x$
 - Sums of squares of $x - \text{mean}_x$: 110.0
 - Regression sums of squares: 27.50 (1 d.f.)
 - Residual sums of squares of y: 13.75 (9 d.f.)
 - Multiple R^2 : 0.667

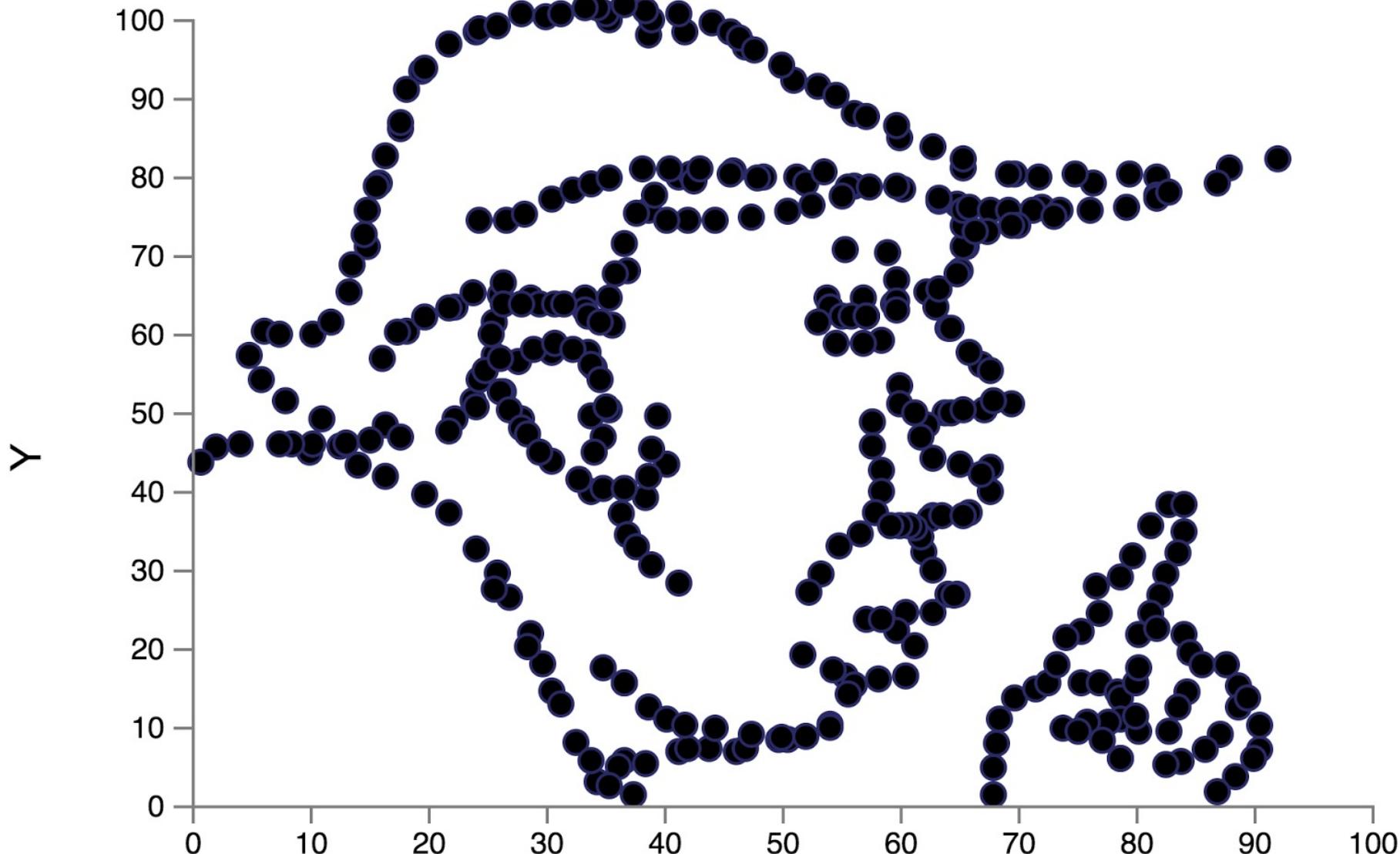
Visualization. The basics



Visualization. The basics

- **Representations of datasets**
 - We can go even further
 - <https://www.youtube.com/watch?v=DbJyPELmhJc>

N = 383 ; X mean = 49.4905 ; X SD = 22.0677 ; Y mean = 50.4382 ; Y SD = 27.7443 ; Pearson correlation = -0.284



Visualization. The basics

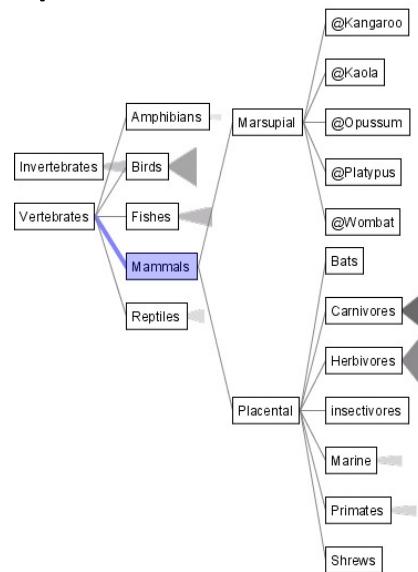
- Carry out tasks more **effectively**
 - Effectiveness requires match between data/task and representation
 - Set of representations is huge
 - Many are ineffective mismatch for specific data/task combo
 - Increases chance of finding good solutions if you understand full space of possibilities
 - What counts as effective?
 - Novel: enable entirely new kinds of analysis
 - Faster: speed up existing workflows
 - How to validate effectiveness
 - Many methods, must pick appropriate one for your context

Visualization. The basics

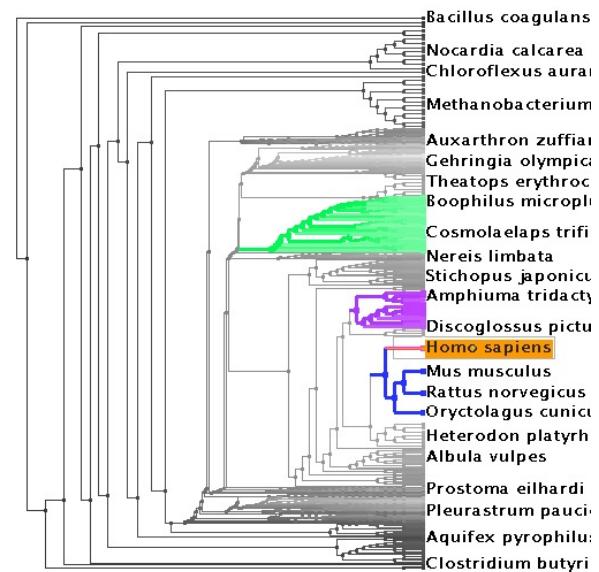
- Analyzing visualizations. Why?
 - Imposes structure on huge design space
 - Scaffold to help you think systematically about choices
 - Analyzing existing as stepping stone to designing new
 - Most possibilities ineffective for particular task/data combination

• Analyzing visualizations

SpaceTree



TreeJuxtaposer



What?

⊕ Tree



Why?

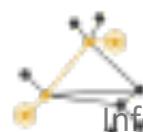
⊕ Actions

- Present
- Locate
- Identify



⊕ Targets

- Path between two nodes



How?

⊕ SpaceTree

- Encode
- Navigate
- Select
- Filter
- Aggregate



[*SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation.* Grosjean, Plaisant, and Bederson. Proc. InfoVis 2002, p 57–64.]

⊕ TreeJuxtaposer

- Encode
- Navigate
- Select
- Arrange



Visualization. The basics

- Purpose: Better understanding of mortality rate, life expectancy, growth ... using publicly available data
 - Hans Rosling, TED Talk 2006

Visualization. The basics

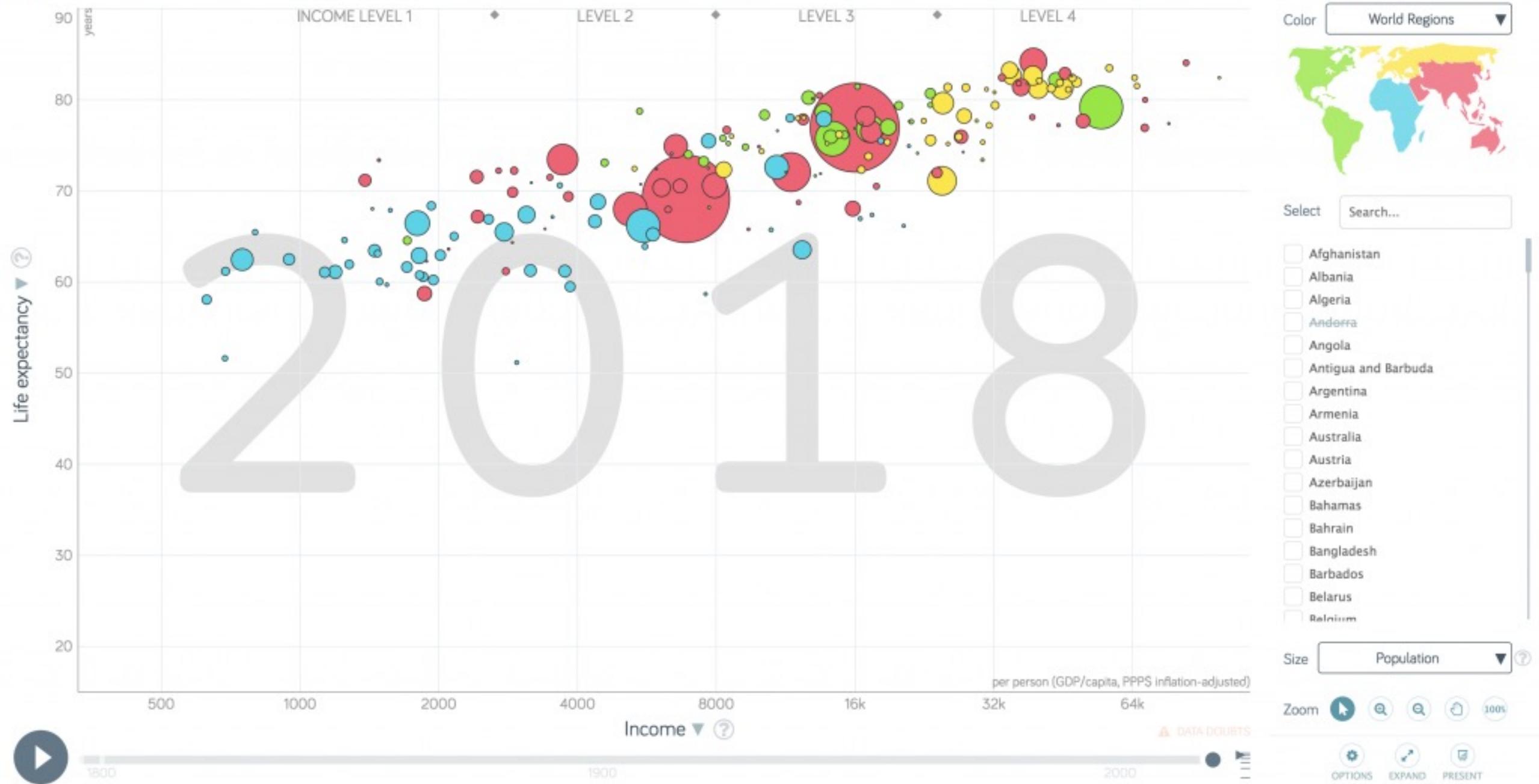


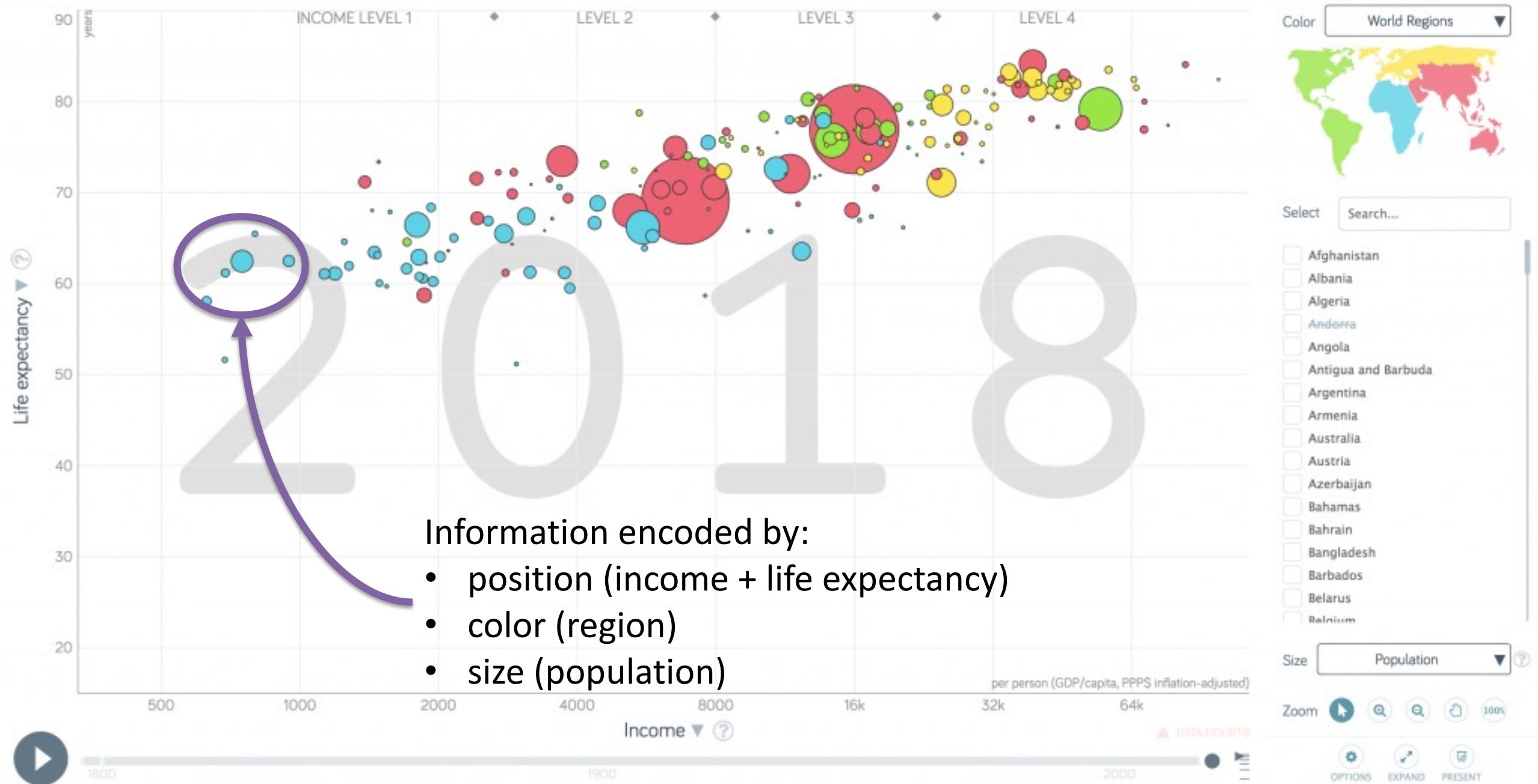
Visualization. Motivation

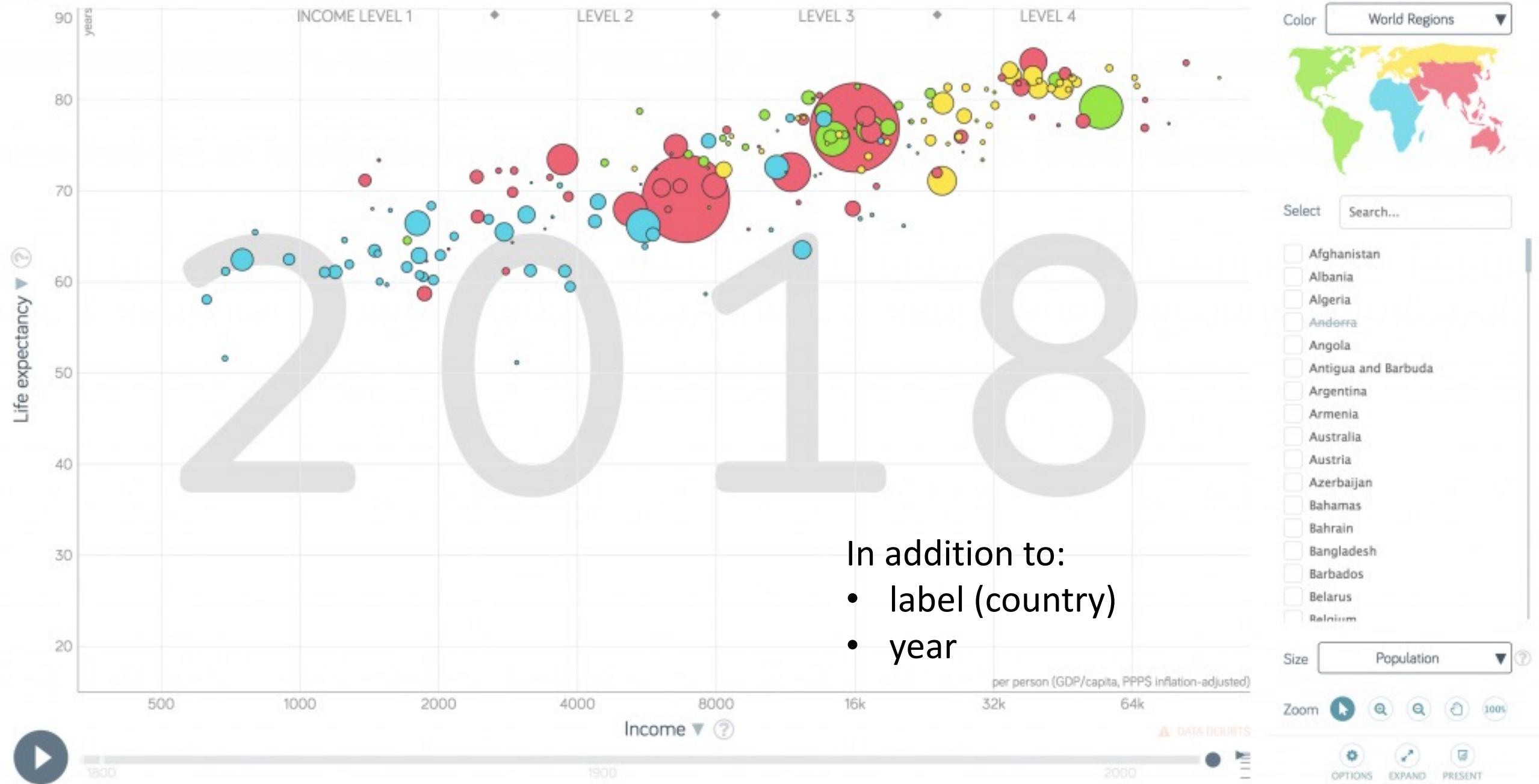
- Purpose: Better understanding of mortality rate, life expectancy, growth ... using publicly available data
 - Hans Rosling, TED Talk 2006
https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen
 - Other talks, just Google “Hans Rosling TED”
 - Also by Ola Rosling: <https://youtu.be/ooovn9PS-KY>

Visualization. Homework

- Answer those questions:
 - How many variables (and which) are displayed in the following charts:
 - The one shown at 2:26
 - The one shown at 13:07
 - Find an example of visual representation that does not effectively communicate the message (minute and second, and reason why)
 - What happens in terms of variables when he “splits South Africa”?
 - Bonus: Who is the “ghost”?







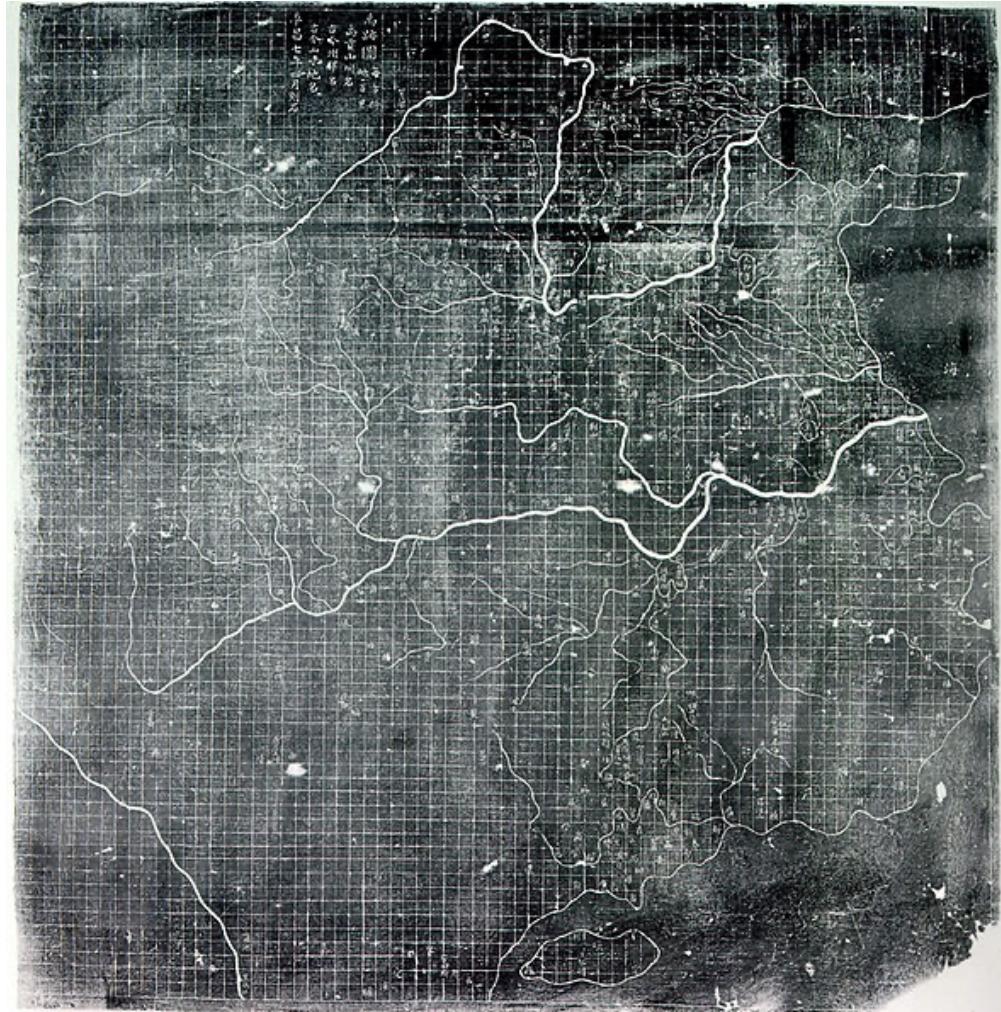
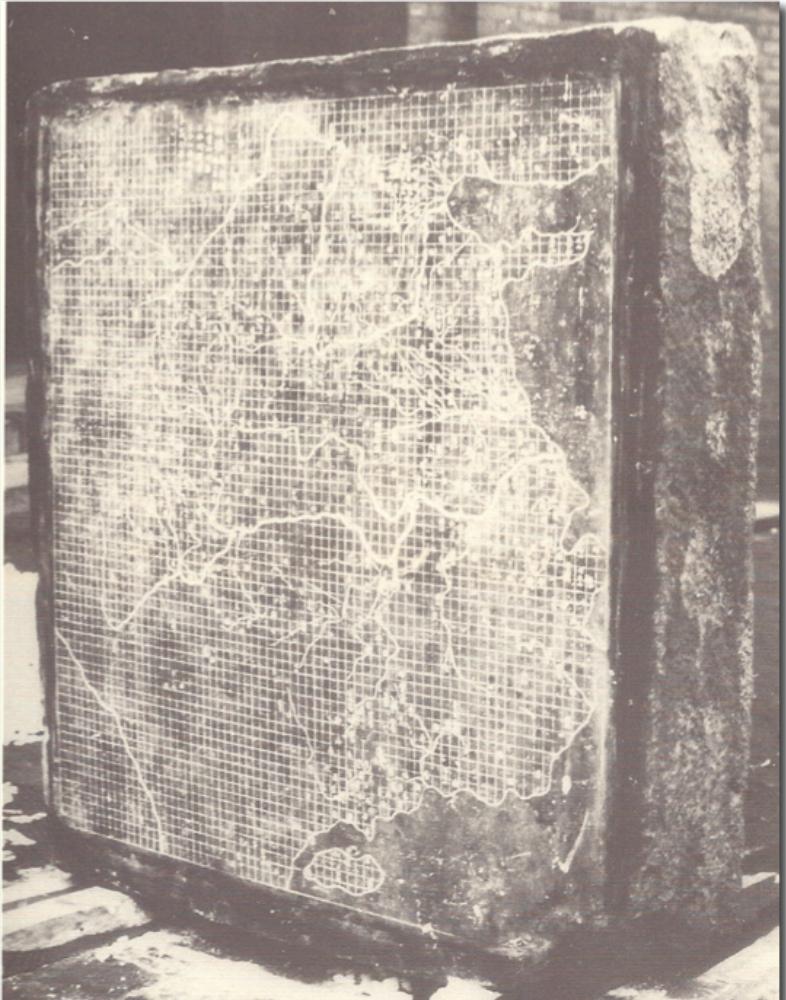
Outline

- *Visualization. The basics*
- **History**
- Visualize: What, why, and how
- The value of visualization
- The visualization mantra
- Visualization Areas
- Visual variables

History

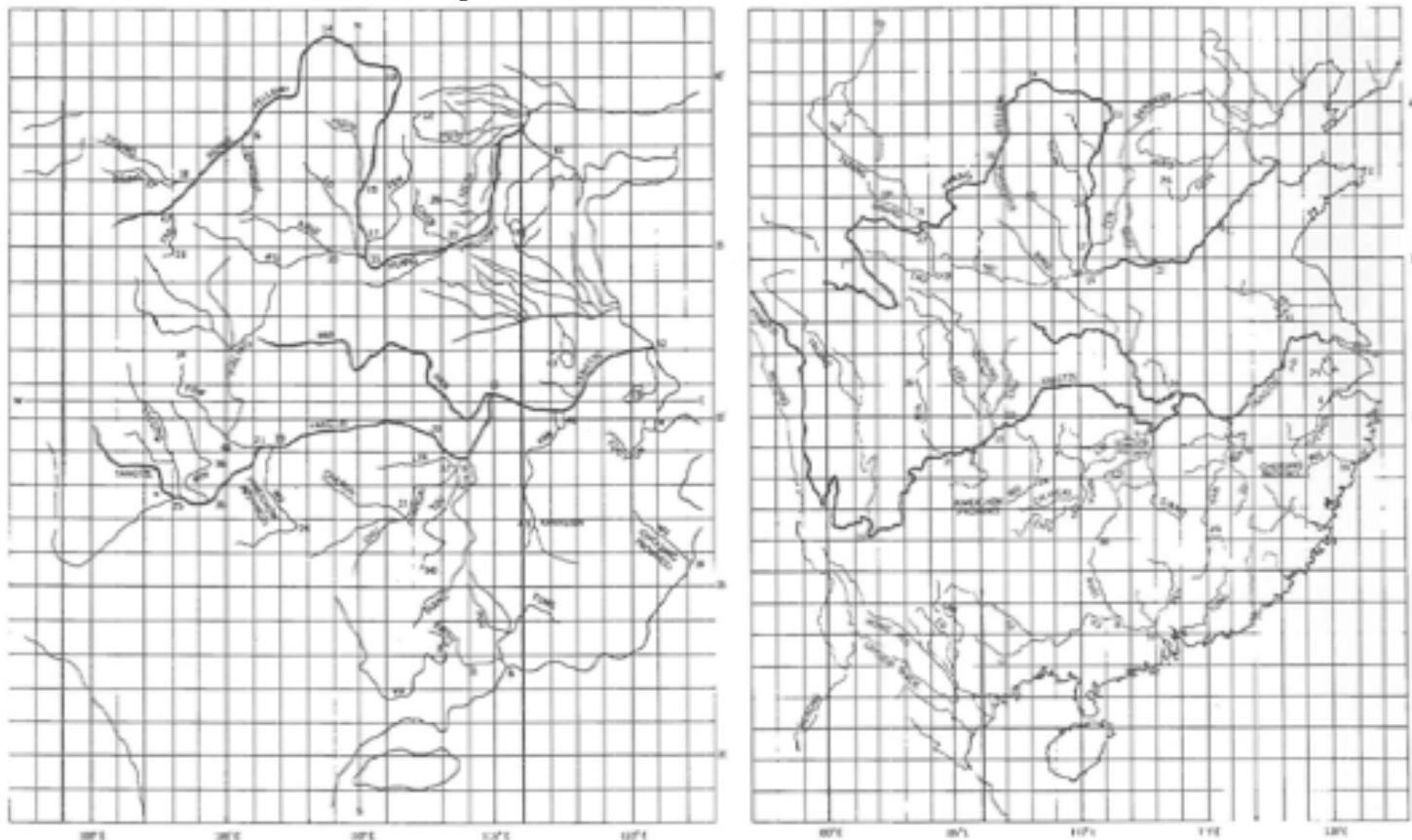
- But when did this start?
 - Quite long ago...
 - E. g. we may track cartographic examples for thousands of years
 - E.g.: The Yu Ji Tu (Map of the Tracks of Yu the Great), a map carved into stone in the year 1137 during the Song Dynasty
 - Uses Cartesian coordinates
 - » Grid with longitudinal and latitudinal lines

History



History

- ... vs current China map



A redrawing of the Yü-chi T'u with an oblong grid constructed empirically from the geography, the numbers refer to the accompanying table.
(from Hapgood)

A modern map of China, on a comparable scale to the Yü-chi T'u and corresponding numbers

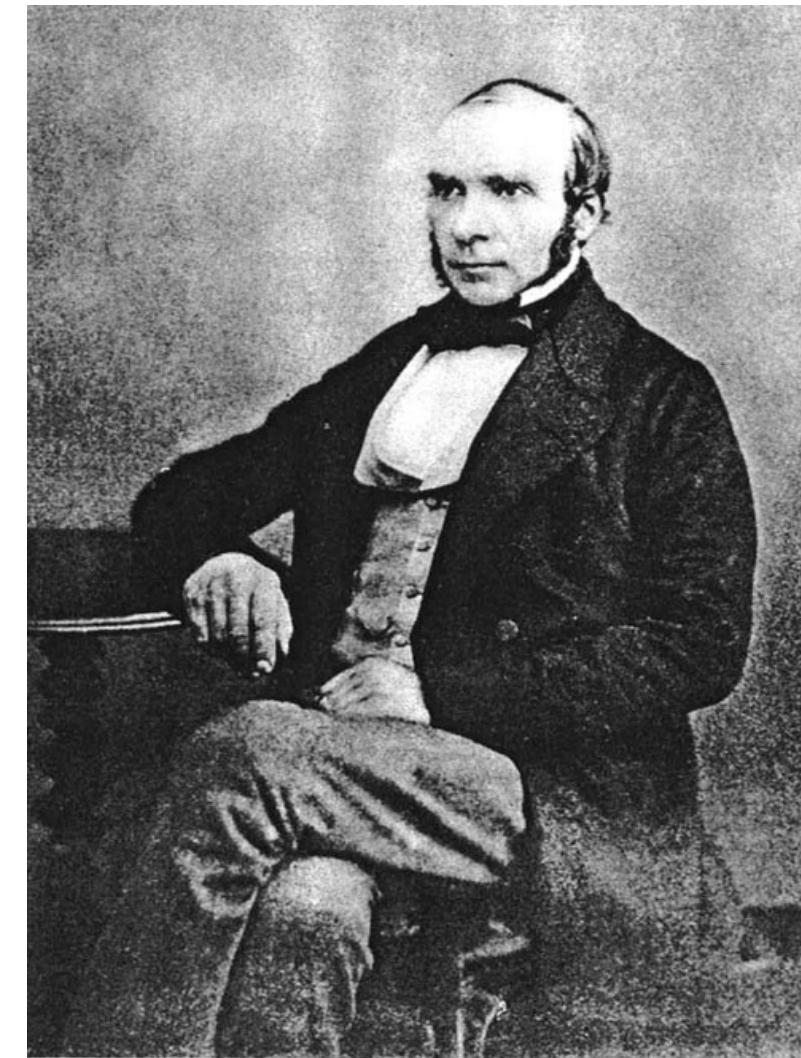
History

- Cholera epidemic in London 1854



History

- Cholera epidemic in London 1854
 - Dr. John Snow was able to trace the source of the cholera outbreak in Soho
 - His findings inspired fundamental changes in the water and waste systems in London
 - He is considered one of the fathers of modern epidemiology

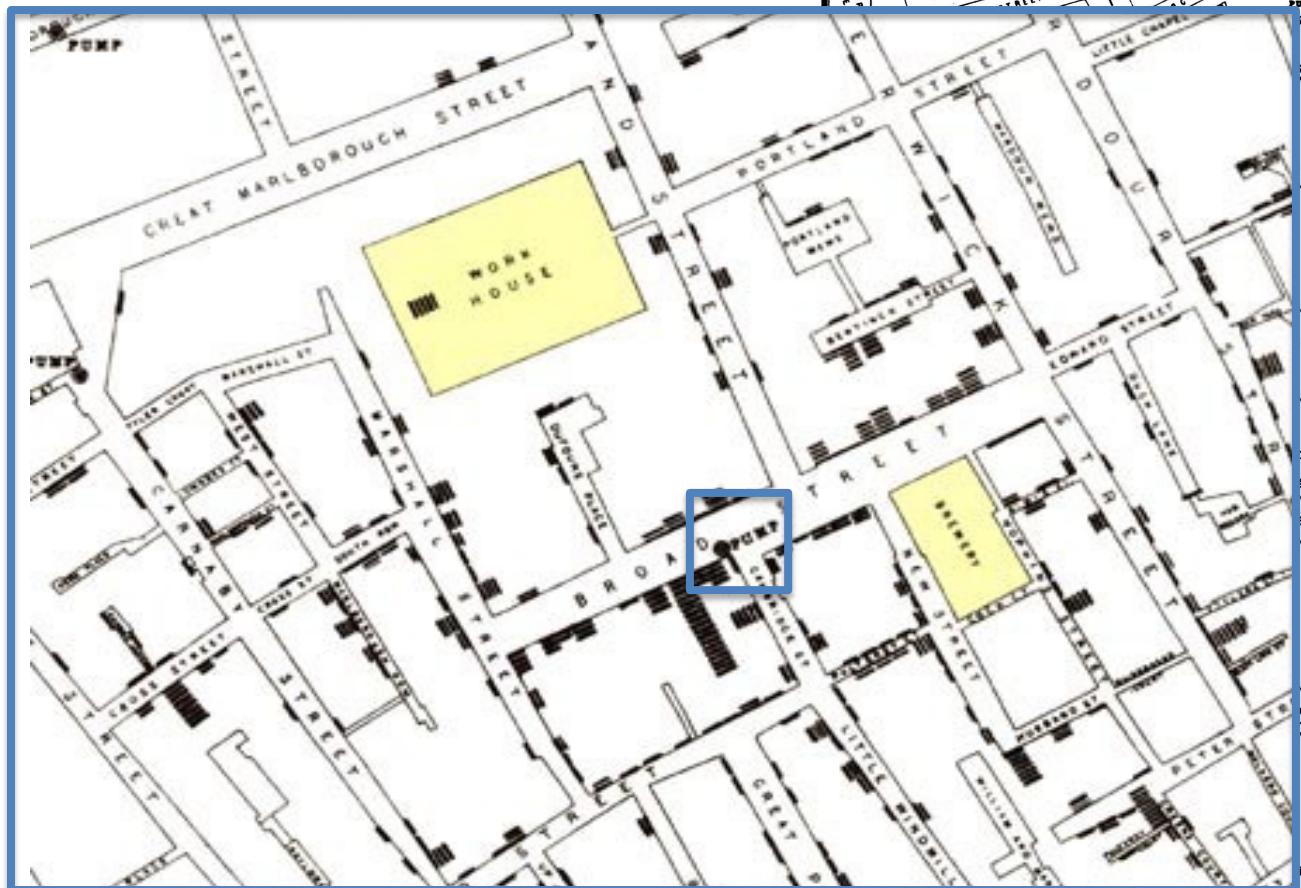


John Snow

History

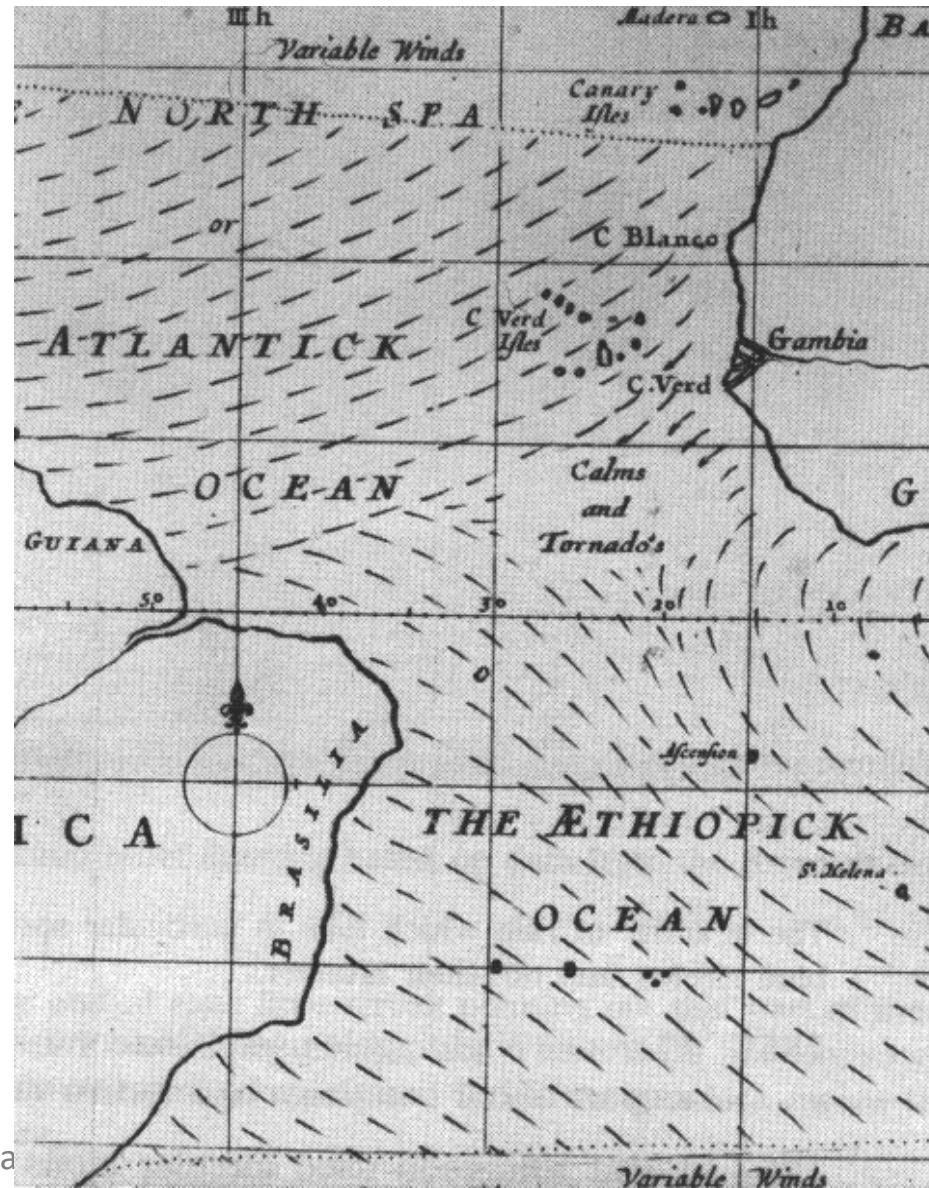
- Cholera epidemic in London 1854
 - “On proceeding to the spot, I found that **nearly all the deaths had taken place within a short distance of the [Broad Street] pump.** There were only ten deaths in houses situated decidedly nearer to another street-pump. In five of these cases the families of the deceased persons informed me that they always sent to the pump in Broad Street, as they preferred the water to that of the pumps which were nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street...”

- Drawing by John Snow



History

- Other examples
 - Wind flow visualization

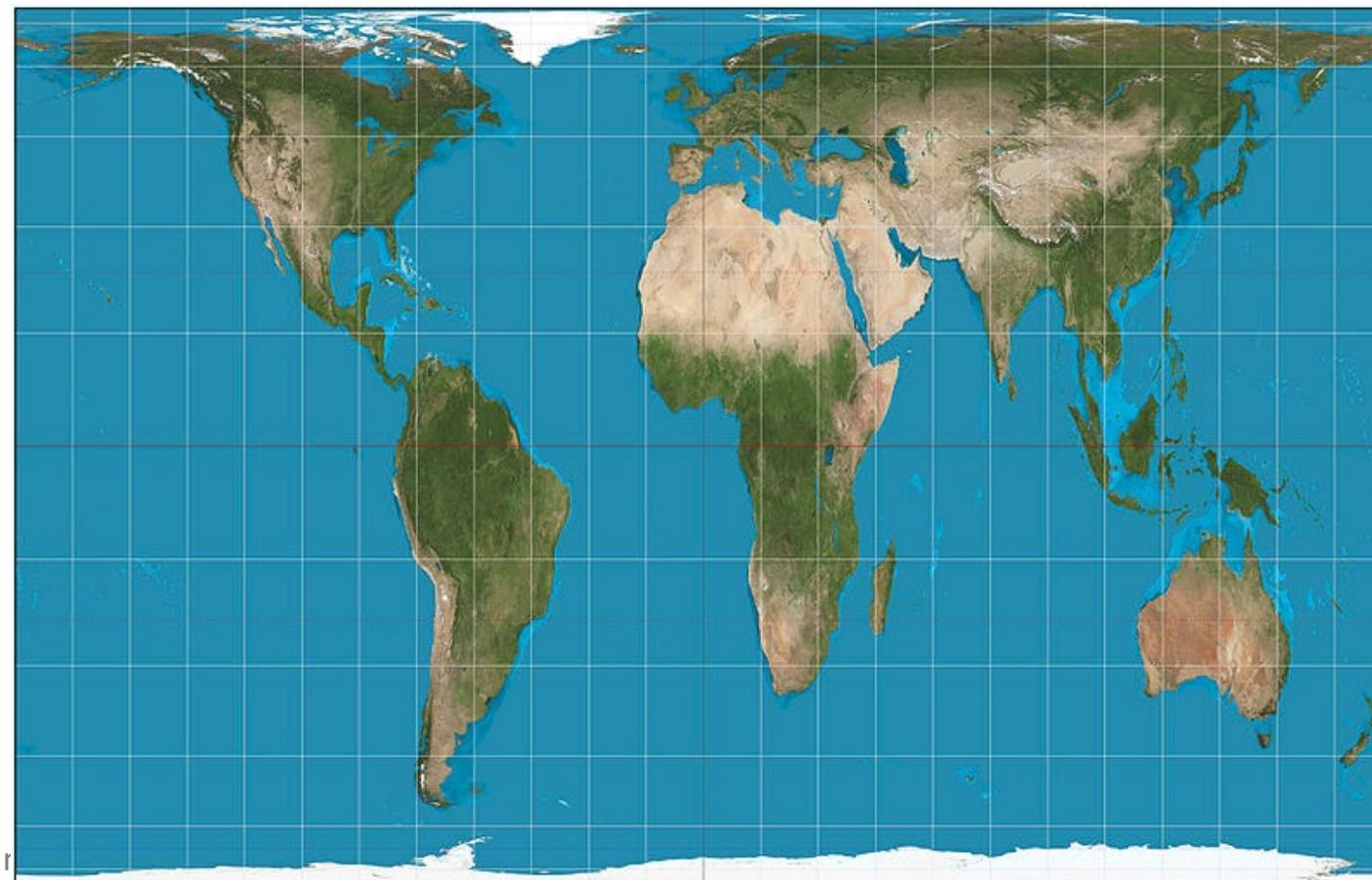


History of Visualization

- Map projections, a real problem:
 - <https://www.youtube.com/watch?v=vVX-PrBRtTY&t=105s>
 - <https://www.youtube.com/watch?v=LAoBLrLWoPE>

History

- Mercator vs Gall-Peters projection

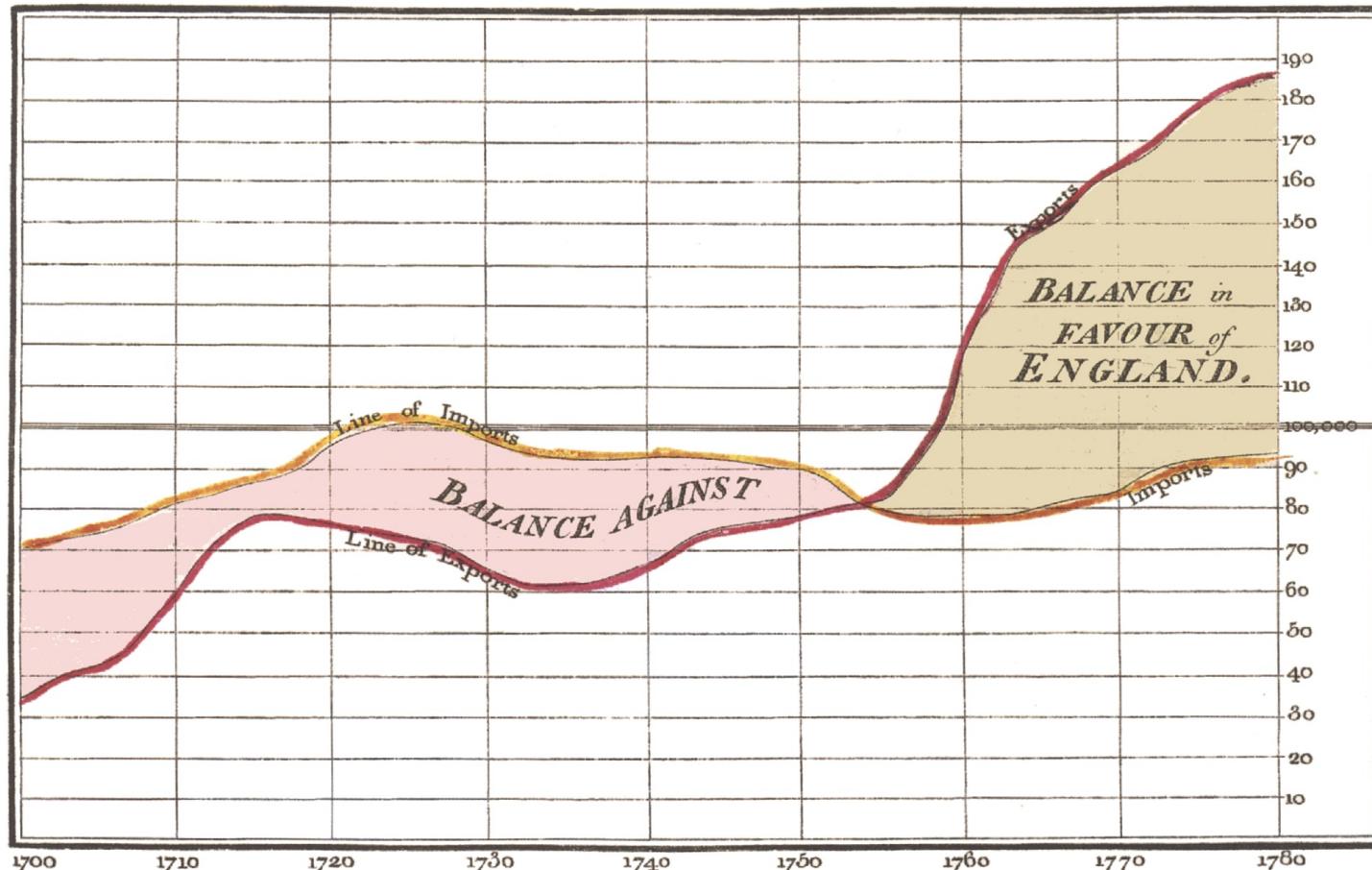


History

- The first visual representations of abstract, nonspatial datasets were created in the 18th century by William Playfair.

History

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



The Bottom line is divided into Years, the Right hand line into £10,000 each.

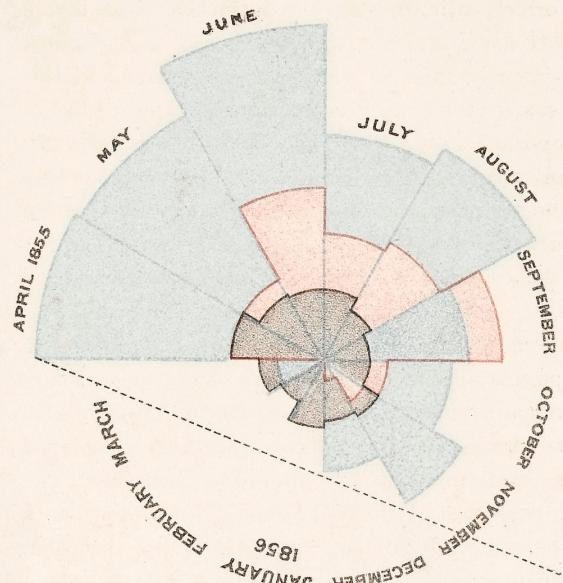
Published as the Act directs, 18th May 1786, by W^m Playfair
Neale sculpt 352, Strand, London.

Introduction

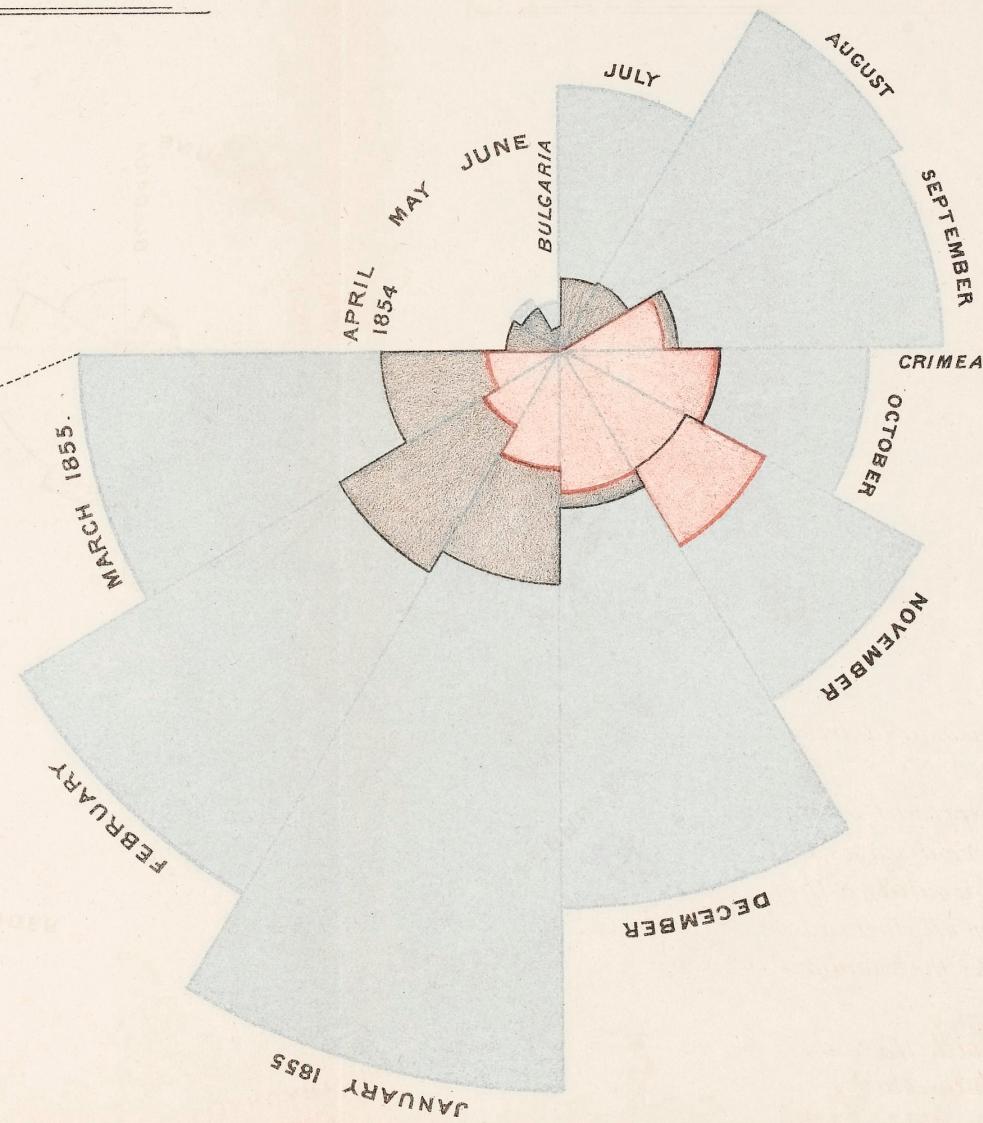
- Florence Nithingale, analyzing mortality in the Crimean War
 - Mortality due to zymotic disease was reduced by increasing sanitary improvements in military barracks

DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.

2.
APRIL 1855 TO MARCH 1856.



1.
APRIL 1854 TO MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

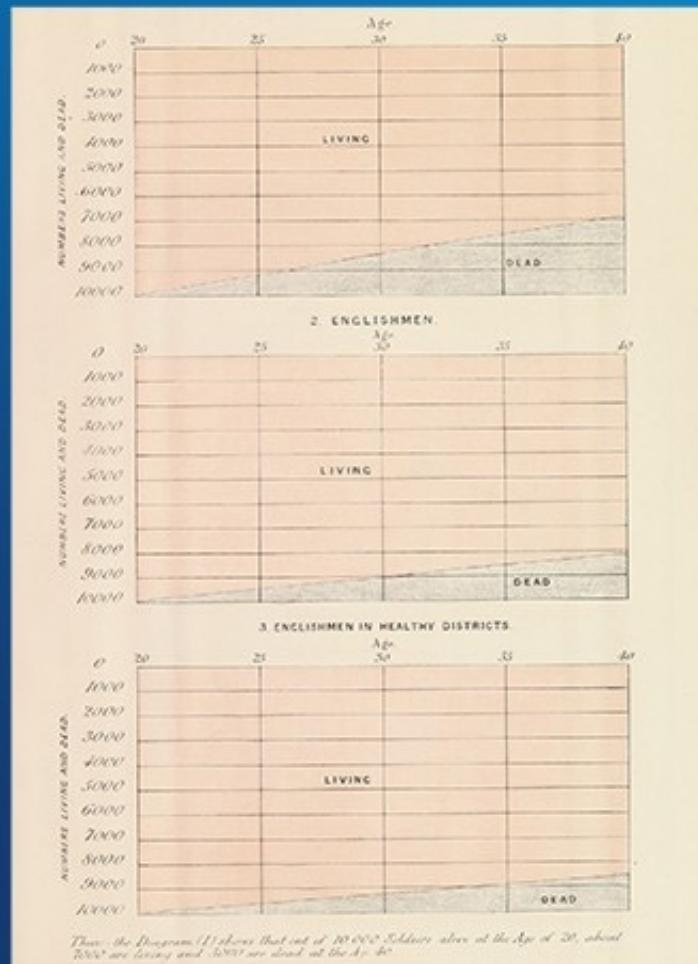
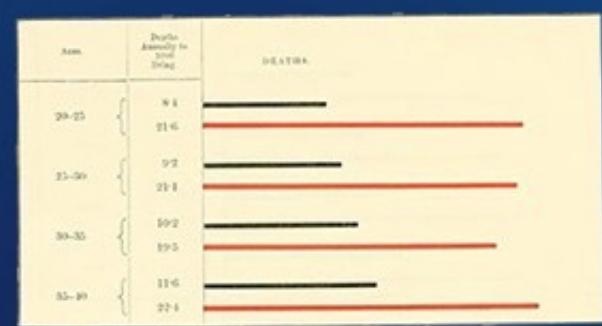
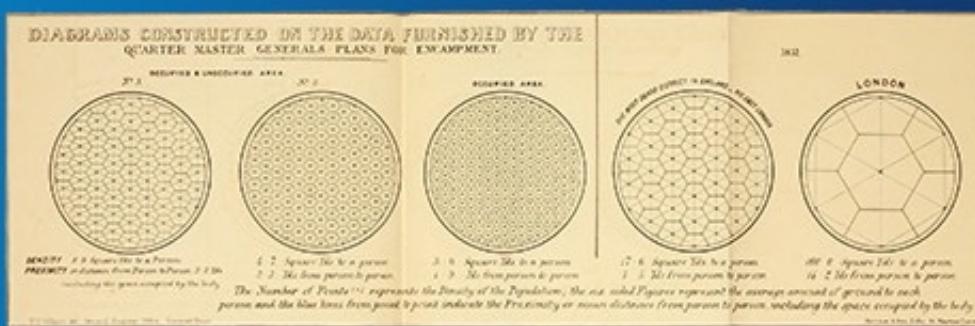
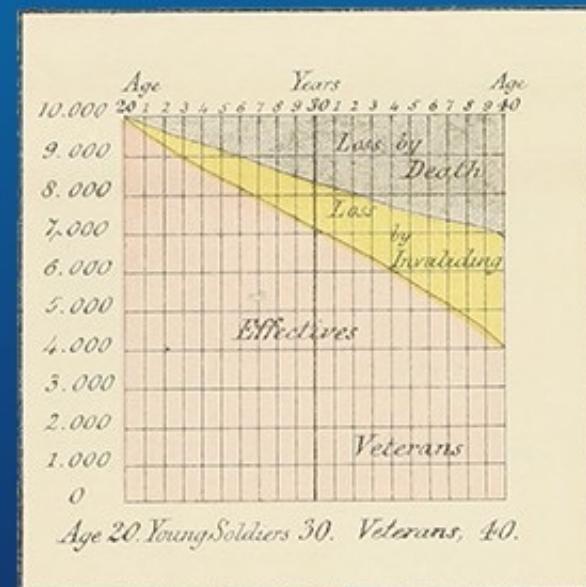
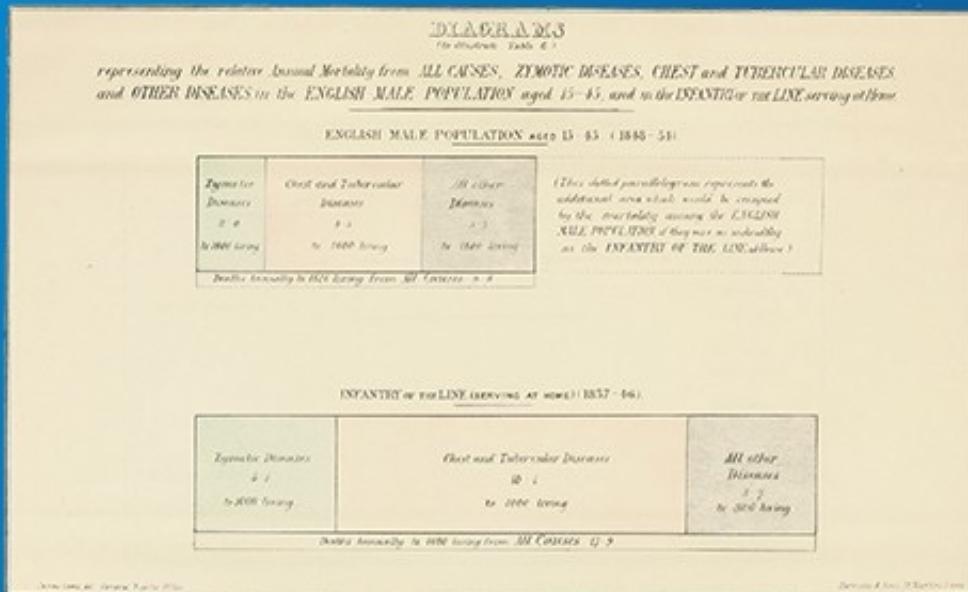
The black line across the red triangle in Nov. 1854 marks the boundary of the deaths from all other causes during the month.

In October 1854, & April 1855, the black area coincides with the red; in January & February 1855, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

Introduction

Diagrams from “Mortality of the British Army: at home and abroad” 1858
[\(<https://archive.org/details/mortalityofbriti00lond/page/n53/mode/2up>\).](https://archive.org/details/mortalityofbriti00lond/page/n53/mode/2up)



History

- Other examples... the Russian campaign by Napoleon

History

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 en retraite

Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs 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. Chiers, 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 du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Malibow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

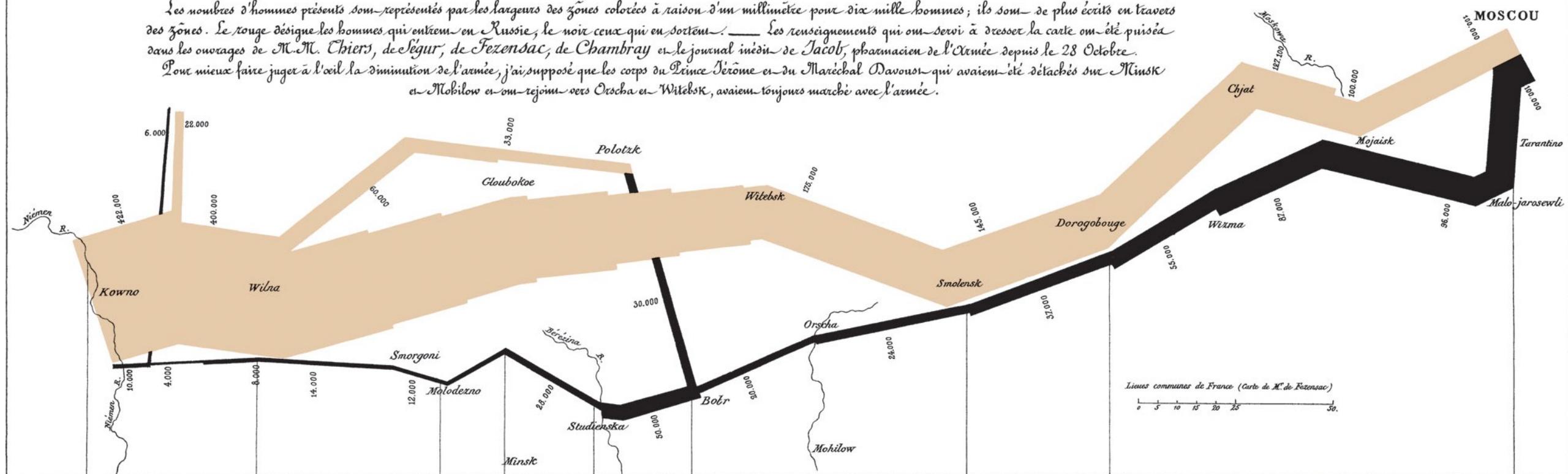


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Niemen gelé.

- 26° le 7 X.^{bre}

- 30° le 6 X.^{bre}

- 24. le 1^{er} X.^{bre}

- 20. le 28 9.^{bre}

- 11°

- 21. le 14 9.^{bre}

- 9° le 9 9.^{bre}

Pluie 24 8.^{bre}

Zéro le 18 8.^{bre}

5

10

15

20

25

30 degrés

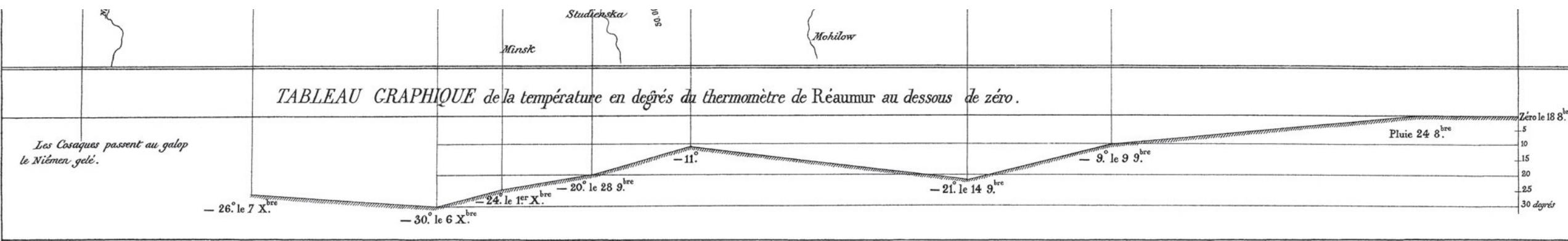
History

*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 en retraite.*

Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs 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. Chiers, 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 du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.



General rules

- Be honest (check & verify the data)
- Above all, show the data
- Explain encodings
- Label your axes
- Make the geometry change as the data
- Take into account your audience

Thanks

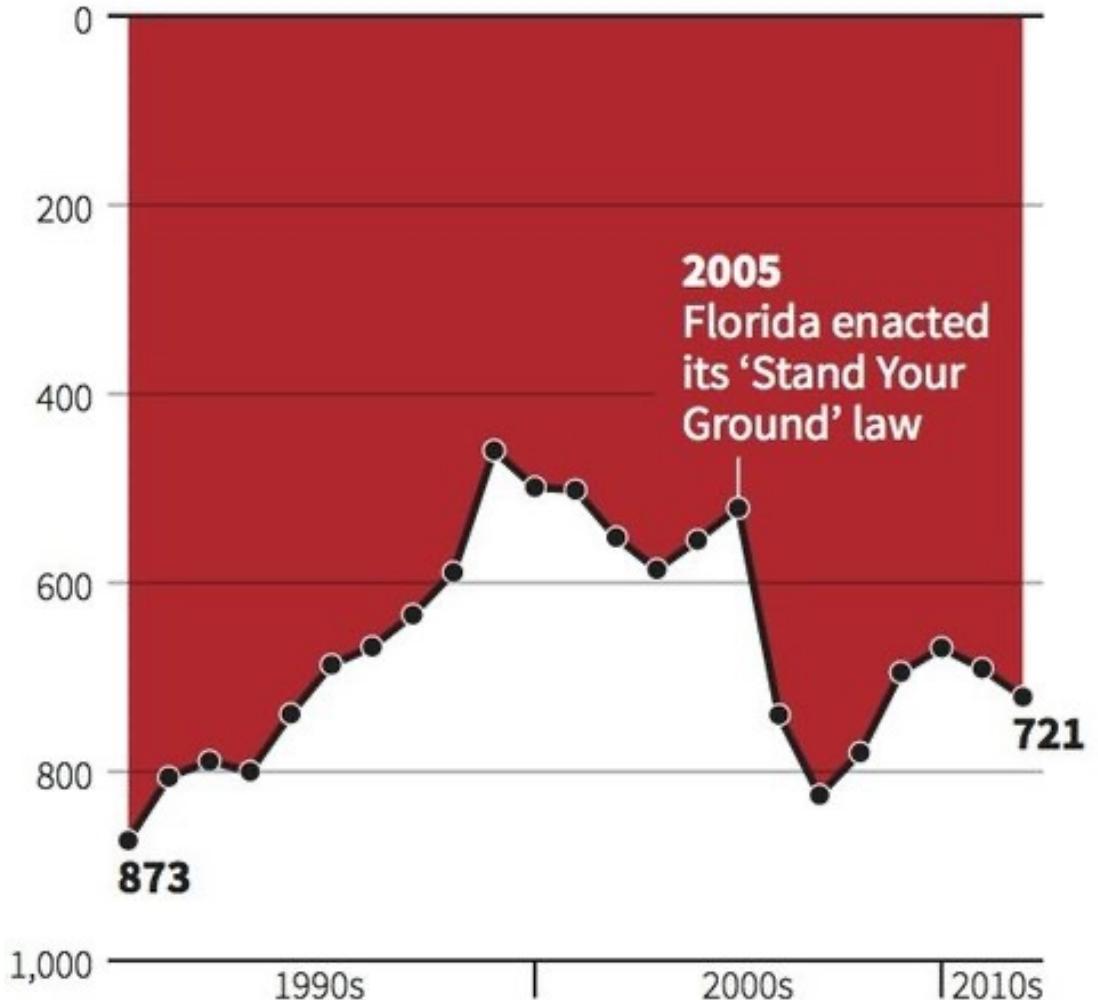
- The information and some of the images come from different presentations/papers/course notes/books from:
 - A. Vilanova, T. Munzer, M. E. Groeller, M. Hadwiger, D. Weiskopf, T. Möller, R. Peikert, P. Muigg, J. Stasko, S. Oeltze-Jafra, J. Bertin, E. Tufte... and/or other papers/blogs/webpages
 - Thanks to all of them!

Other resources

- T. Munzner's presentation (with the visualization definition):
<https://www.youtube.com/watch?v=Bc7Ajludois>
- Minard's Napoleon March analyzed by R. Kosara:
https://www.youtube.com/watch?v=hlb1uM_SOcE
- Florence Nightingale's Famous Rose Chart, by R. Kosara:
https://www.youtube.com/watch?v=JZh8tUy_bnM

Gun deaths in Florida

Number of murders committed using firearms



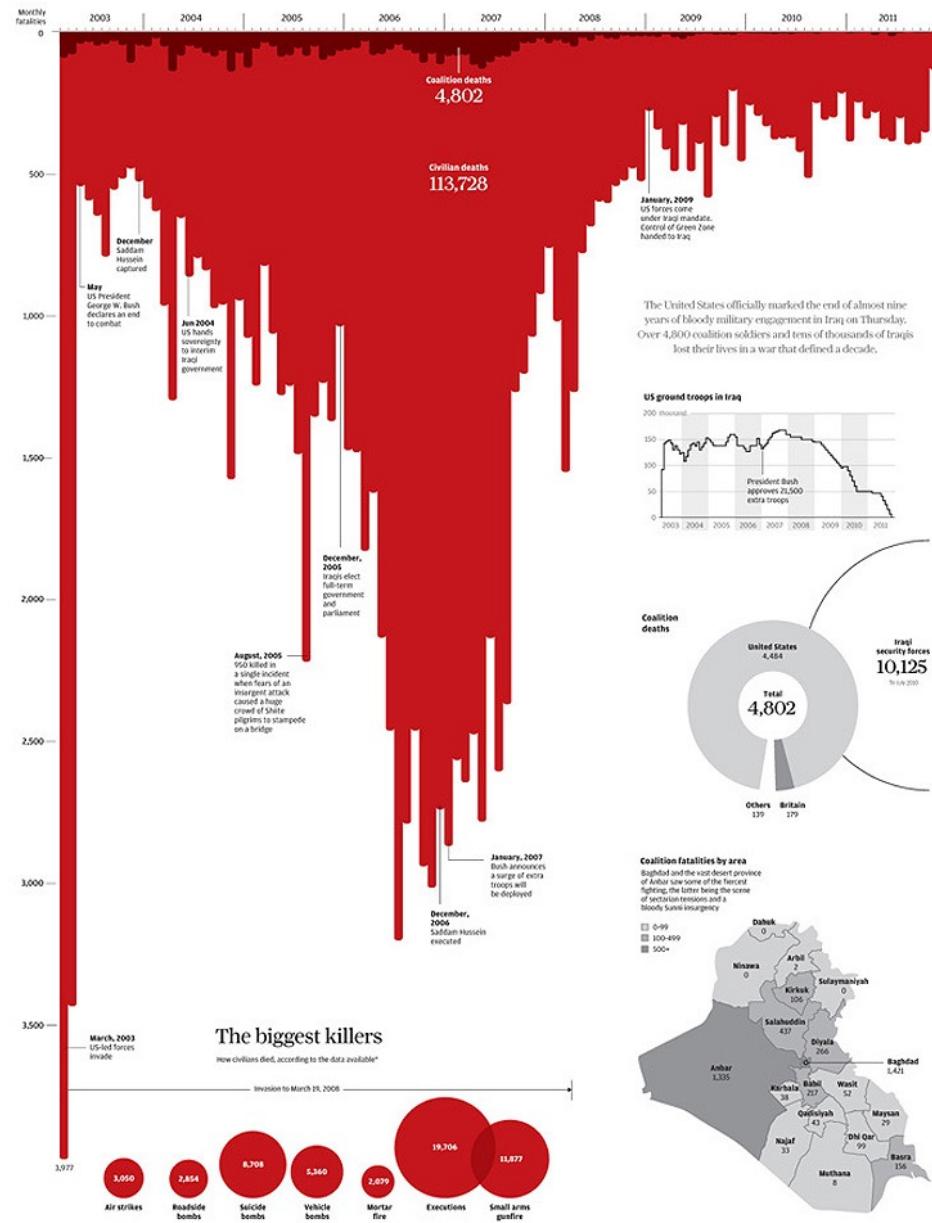
Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

Jáquez

Iraq's bloody toll

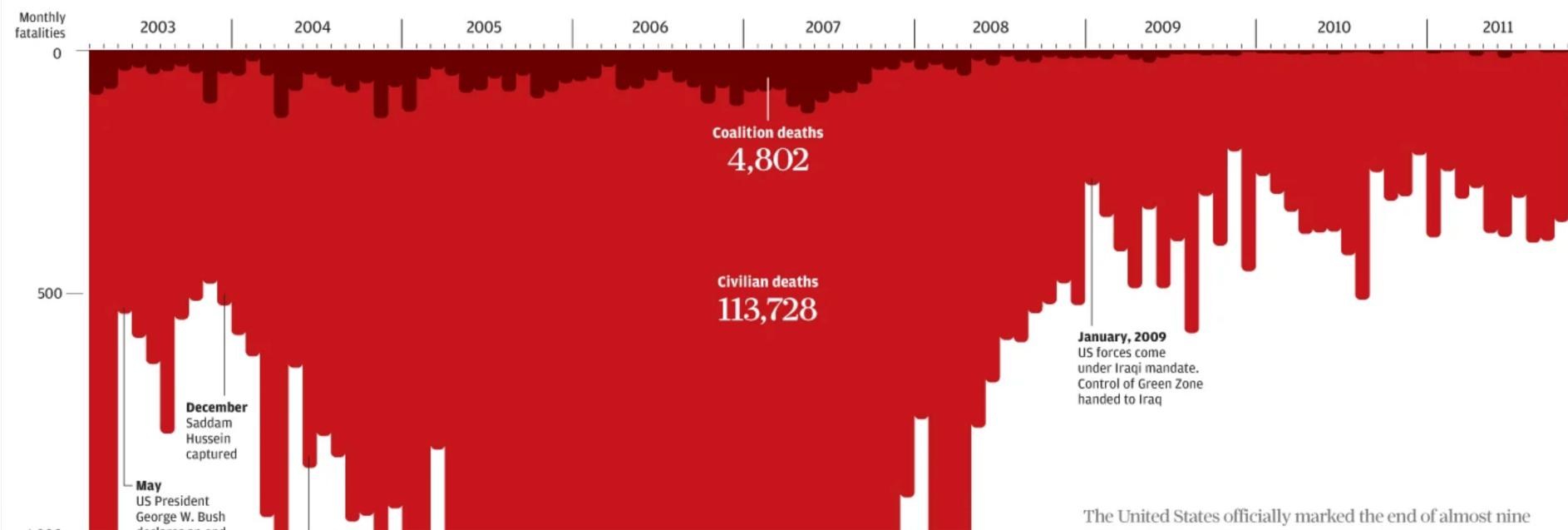


Infographics / Infographics

Iraq's bloody toll

[f](#) [t](#) [%](#) [✉](#)

Iraq's bloody toll



Visualization. Introduction

Pere-Pau Vázquez

Dept. Computer Science – UPC