

# Data Visualization

## European Creativity Festival

Óscar Marín Miró - Outliers Collective  
@outliers\_es  
[www.outliers.es](http://www.outliers.es)

# Contents

## Visualization matters   Tool Overview

- ★What is data visualization?
- ★Why visualize?

- ★Coding tools
- ★Interactive tools

## Digging deeper

- ★Brief history
- ★Editorial approximations
- ★Data dimensions
- ★Relevant people
- ★A dataviz team

## Maps: CartoDB

## Networks: Gephi

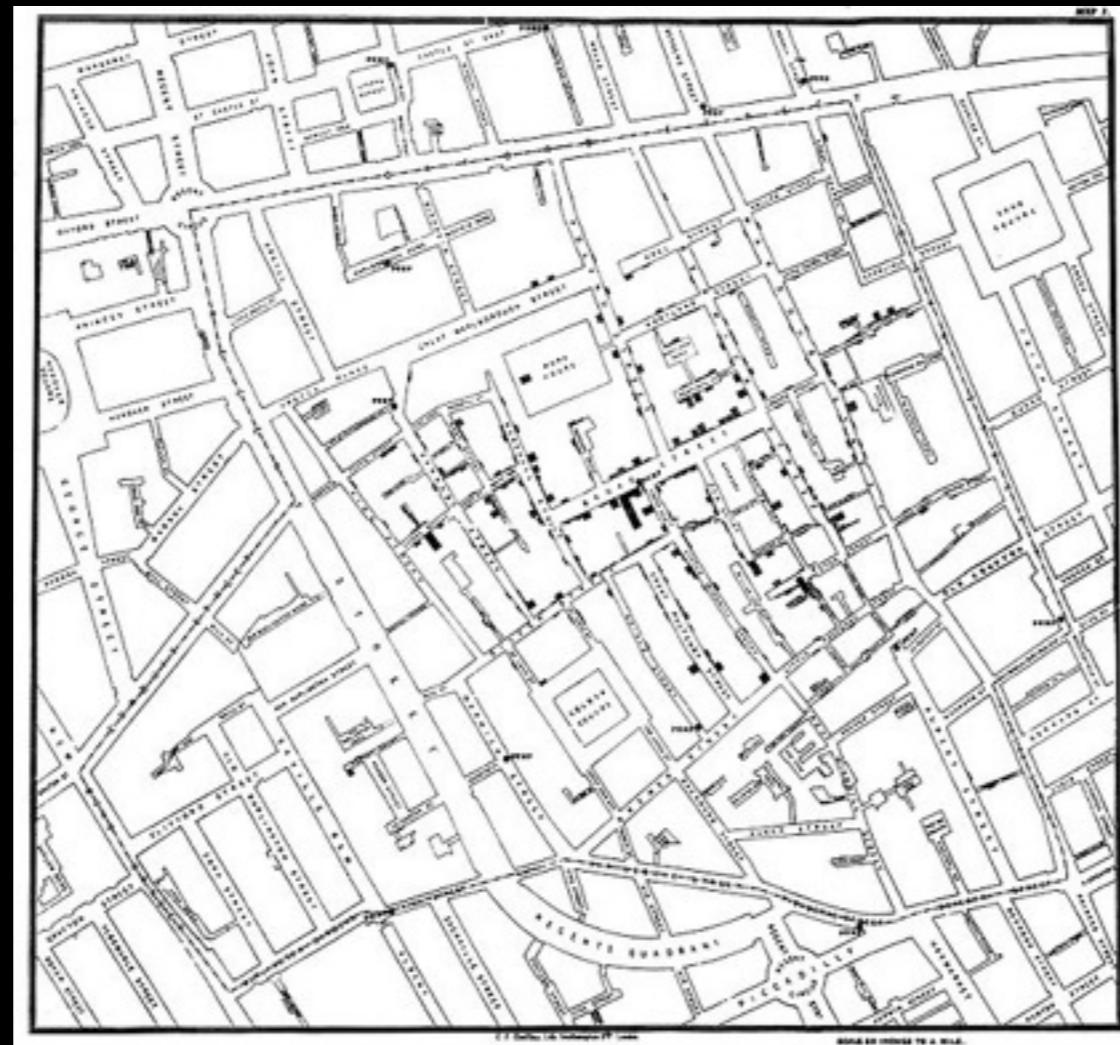
## All-purpose: Tableau

## References

# Links for the workshop

- ★Gephi download [link](#)
- ★CartoDB [sign up](#)
- ★‘Tableau public’ [download](#)
- ★Datasets + pdf on [github](#)

# Visualization matters

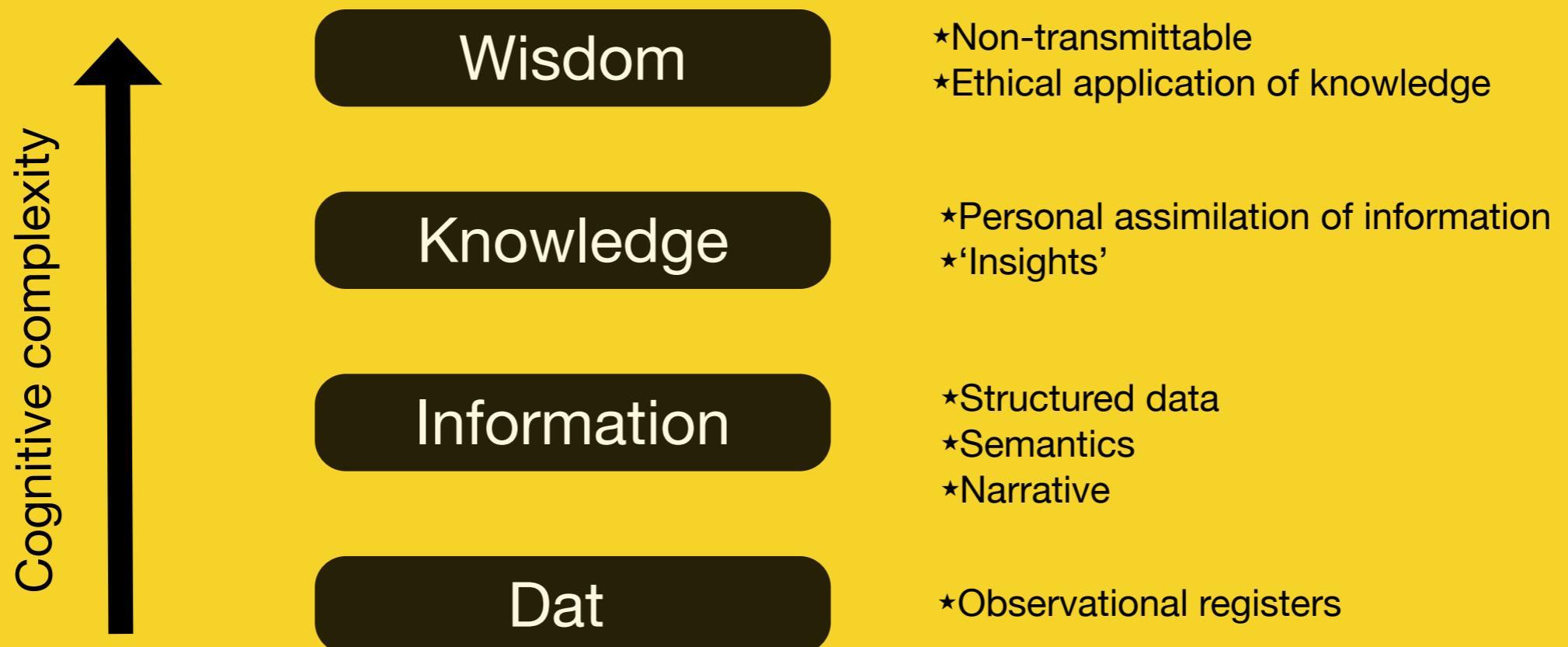


John Snow map

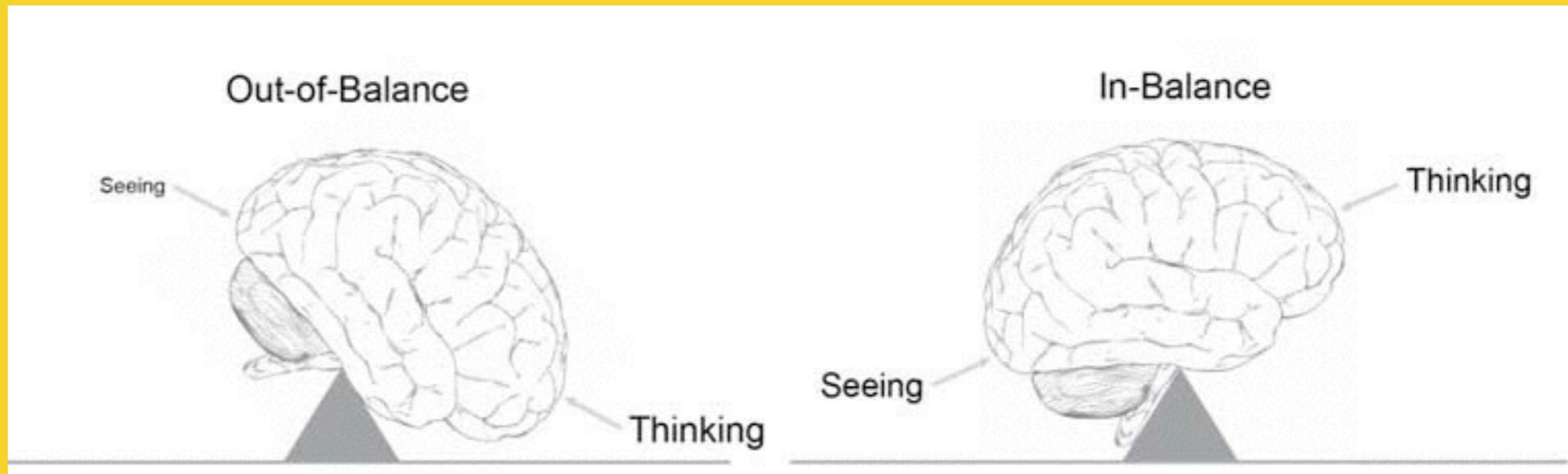
# What is Data Visualization?

“Technologies that transform data into information through visual elements”

# What is Data Visualization?



# Why do Data Visualization matter?



[link](#)

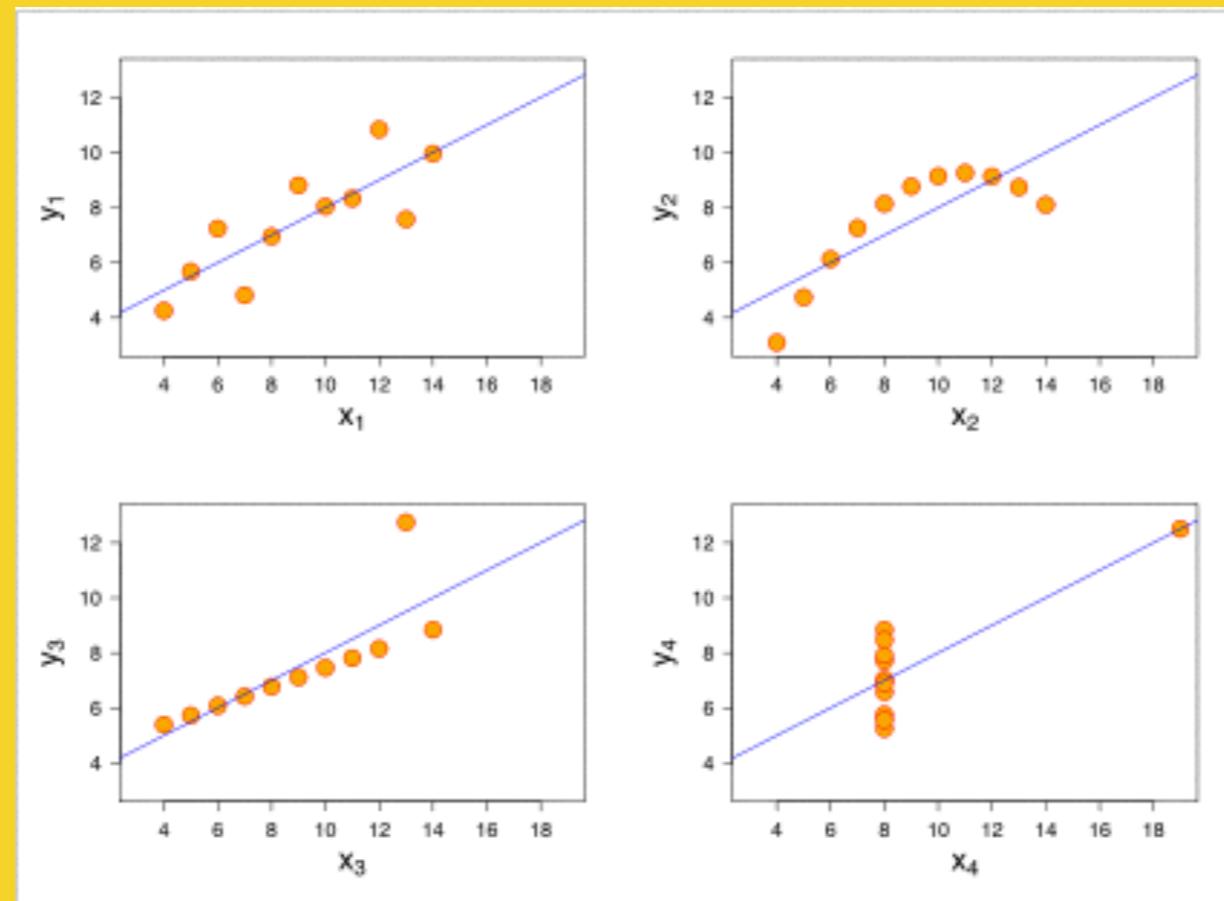
# Why do Data Visualization matter?

Anscombe's quartet									
I		II		III		IV			
x	y	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		

Property	Value
Mean of $x$ in each case	9 (exact)
Variance of $x$ in each case	11 (exact)
Mean of $y$ in each case	7.50 (to 2 decimal places)
Variance of $y$ in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between $x$ and $y$ in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)

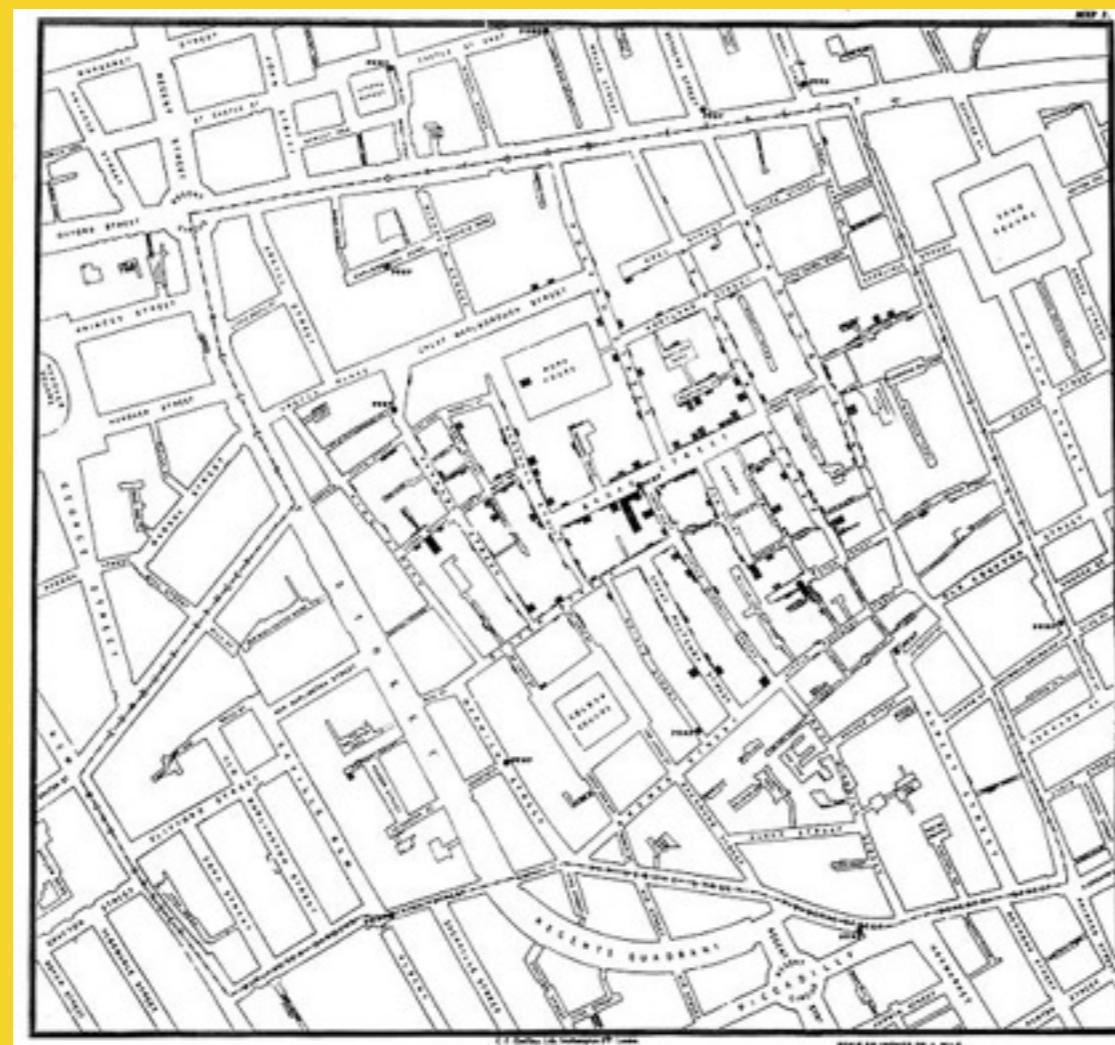
Anscombe's quartet

# Why do Data Visualization matter?



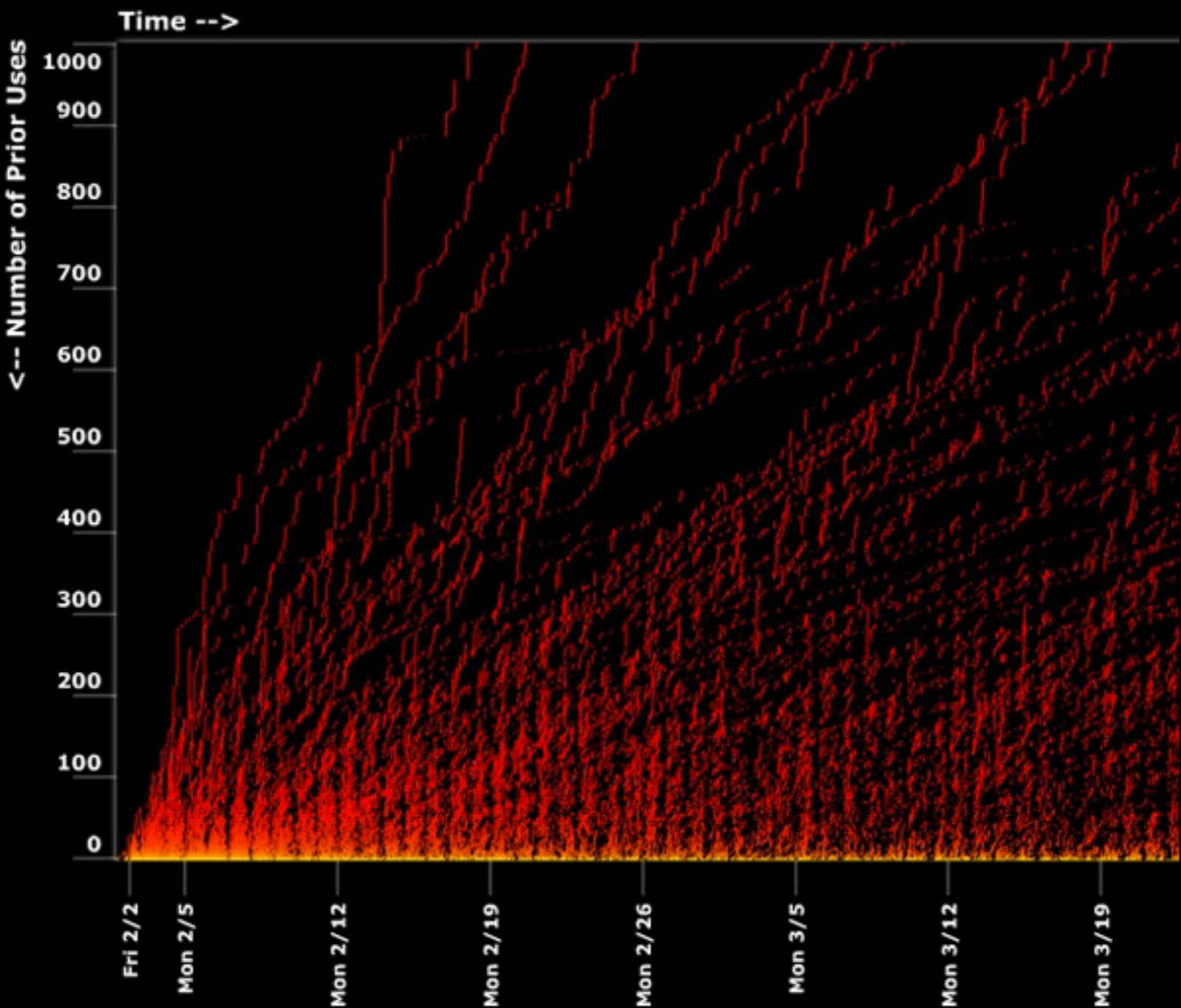
Anscombe's quartet

# Why do Data Visualization matter?



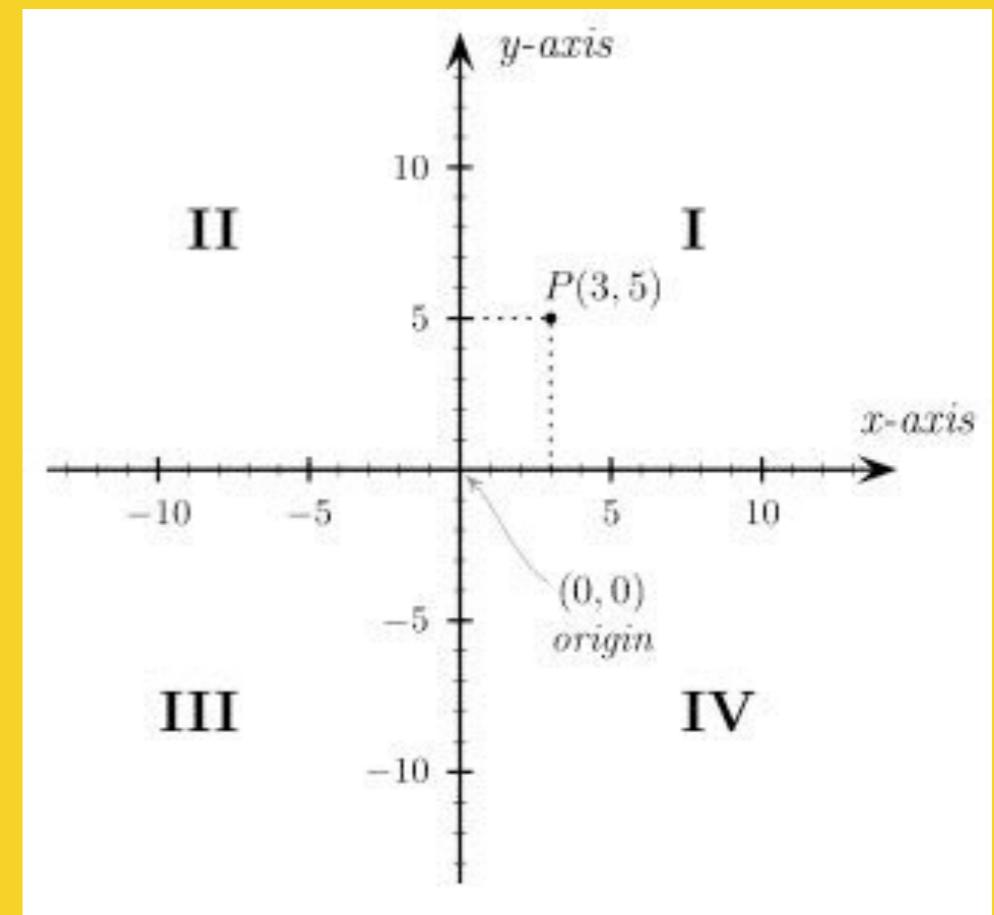
John Snow map

# Digging deeper



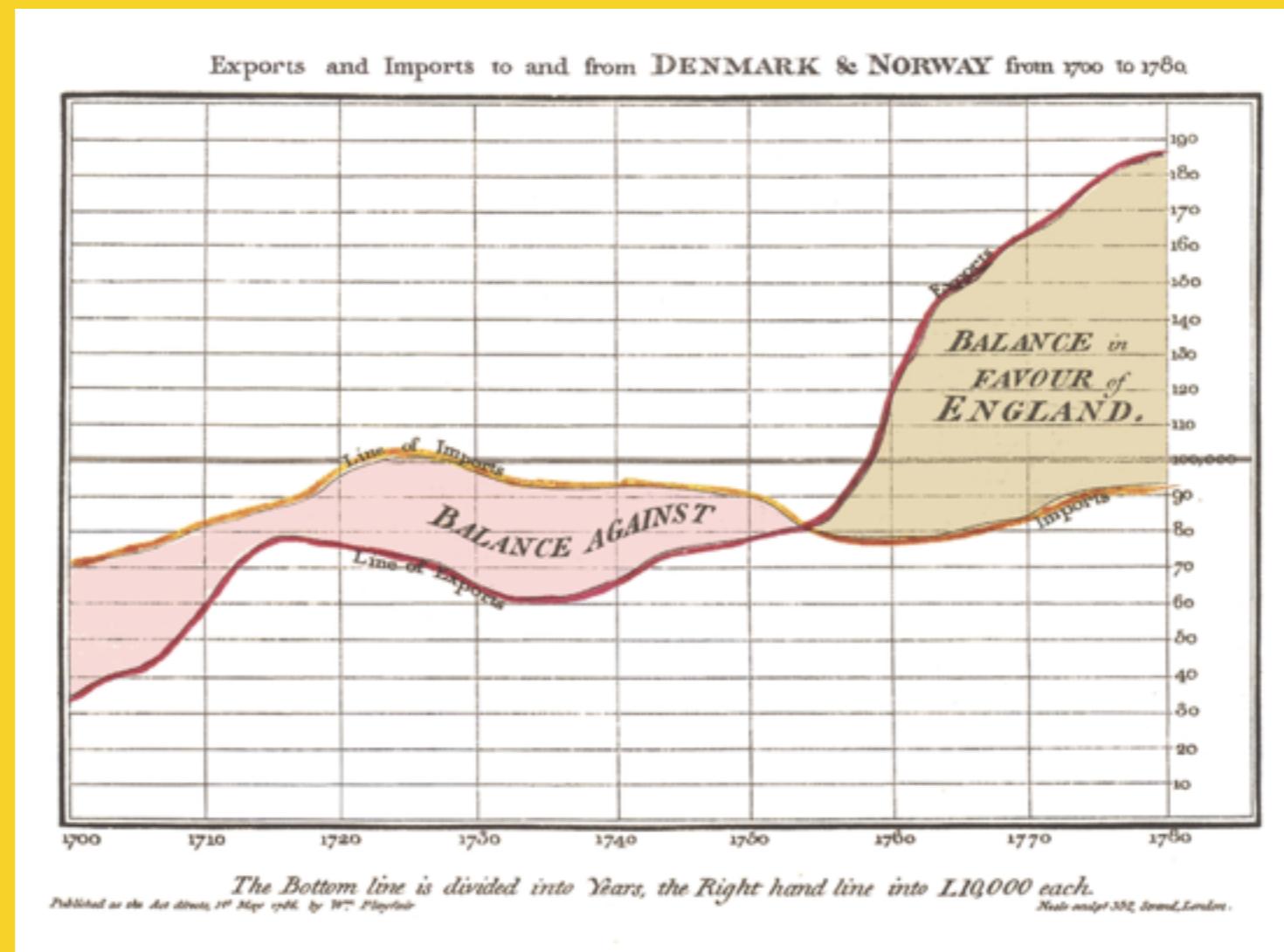
[link](#)

# Brief history



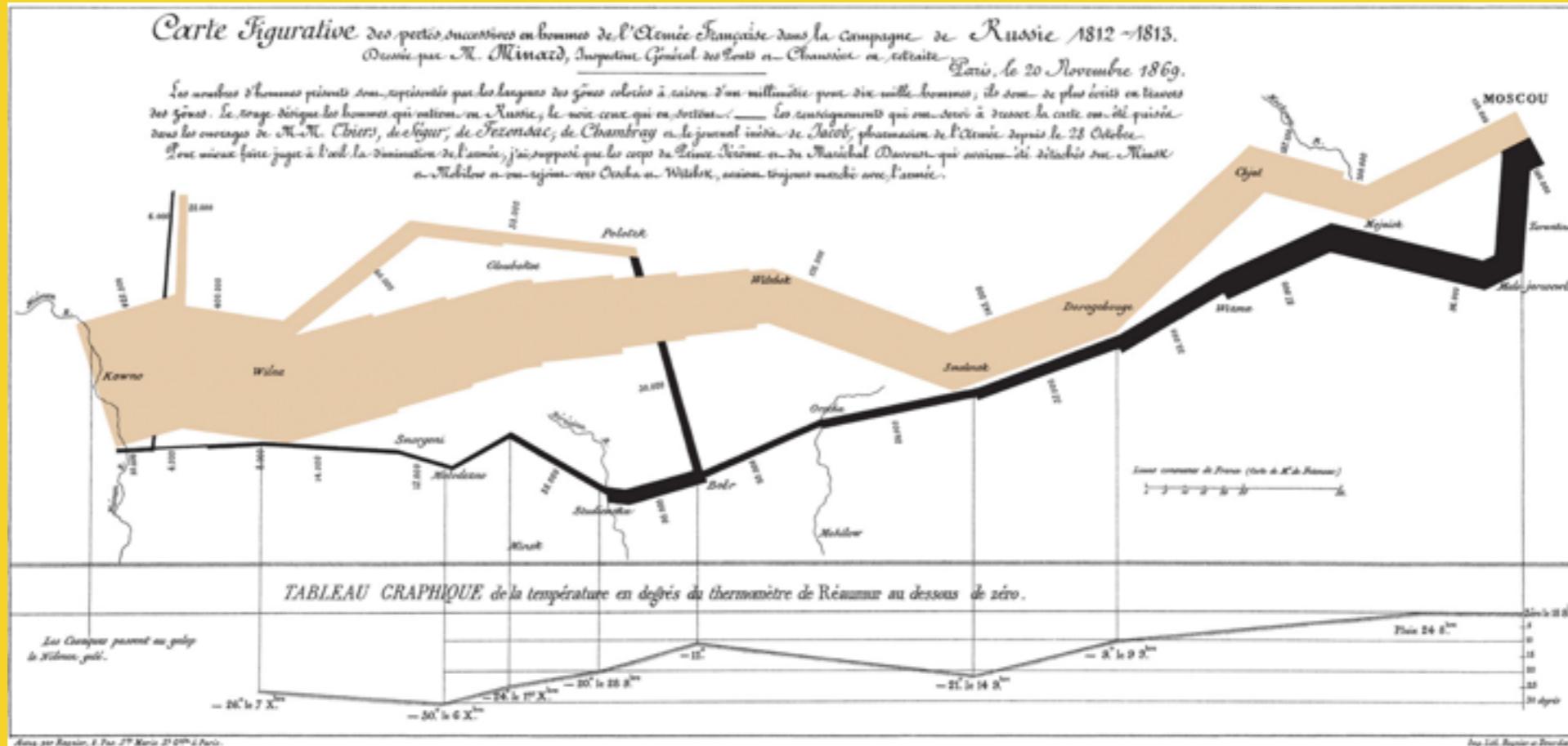
XVII century: René Descartes. Descriptive geometry

# Brief history



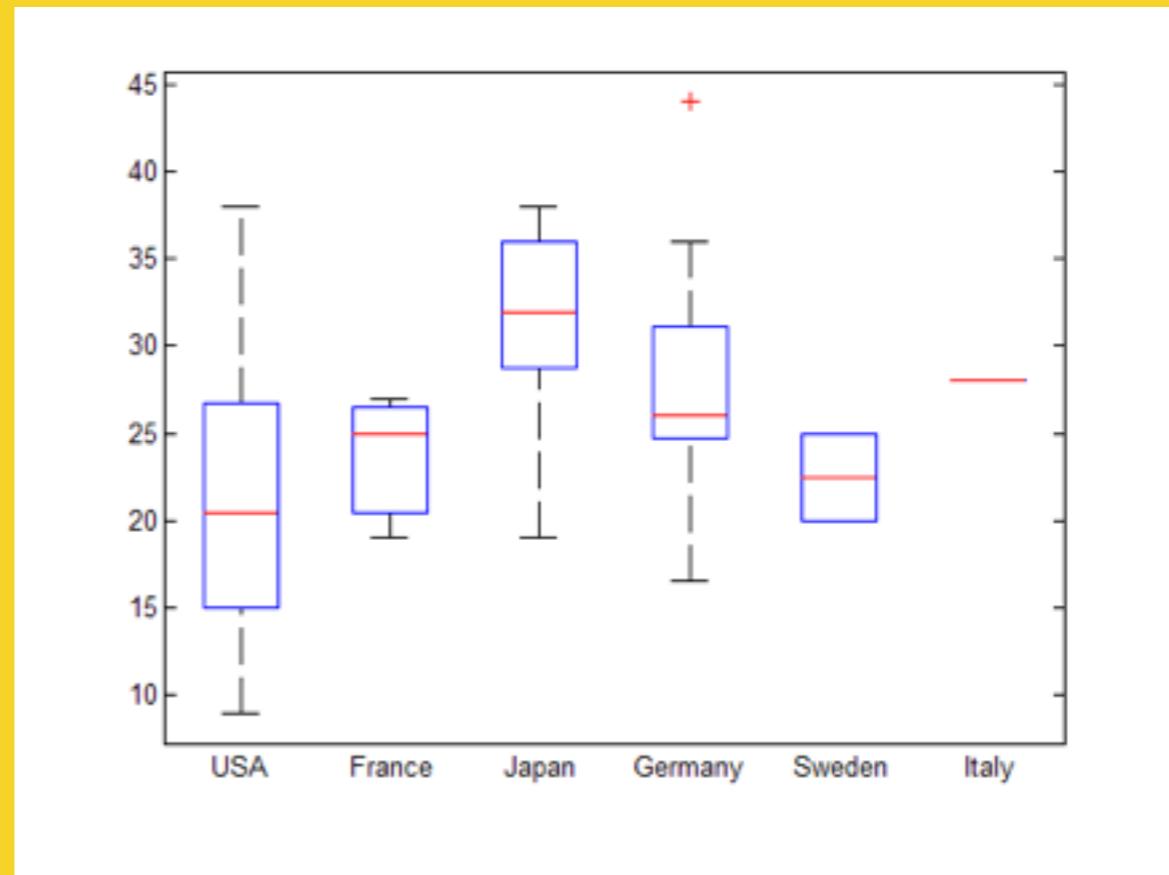
XVIII century: William Playfair. Statistical graphics

# Brief history



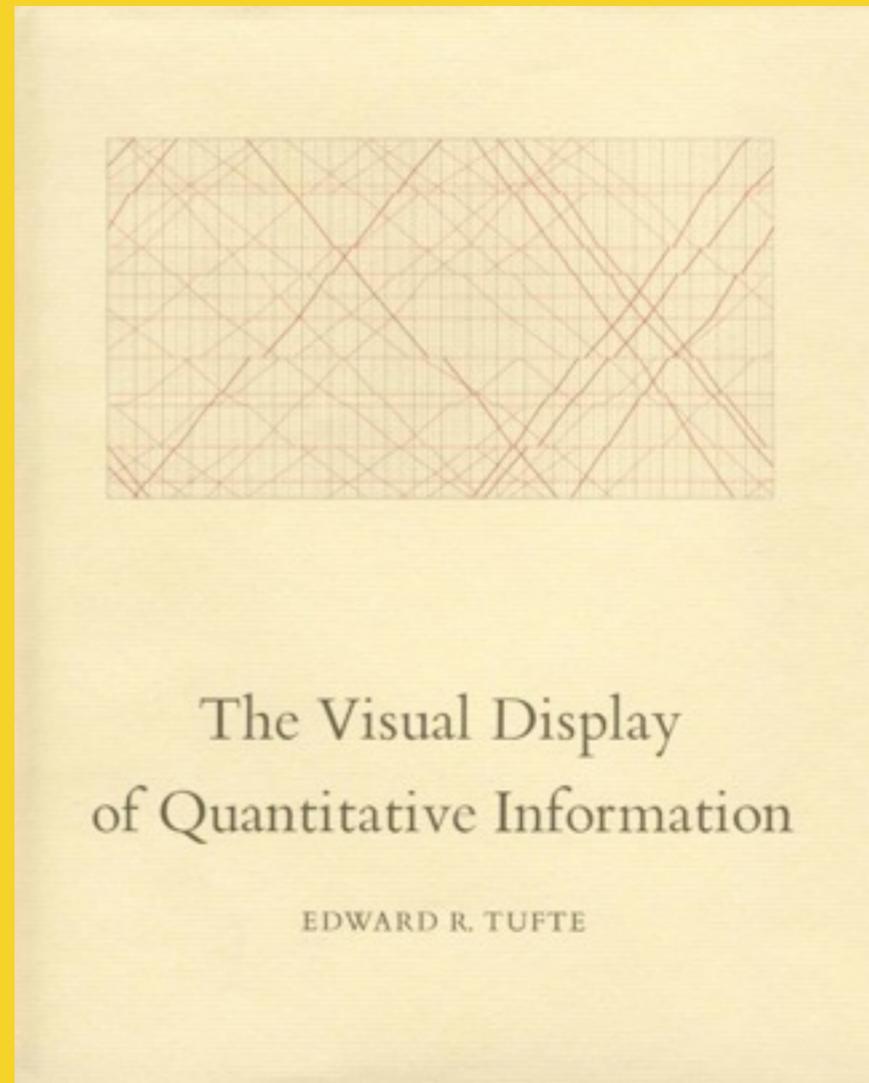
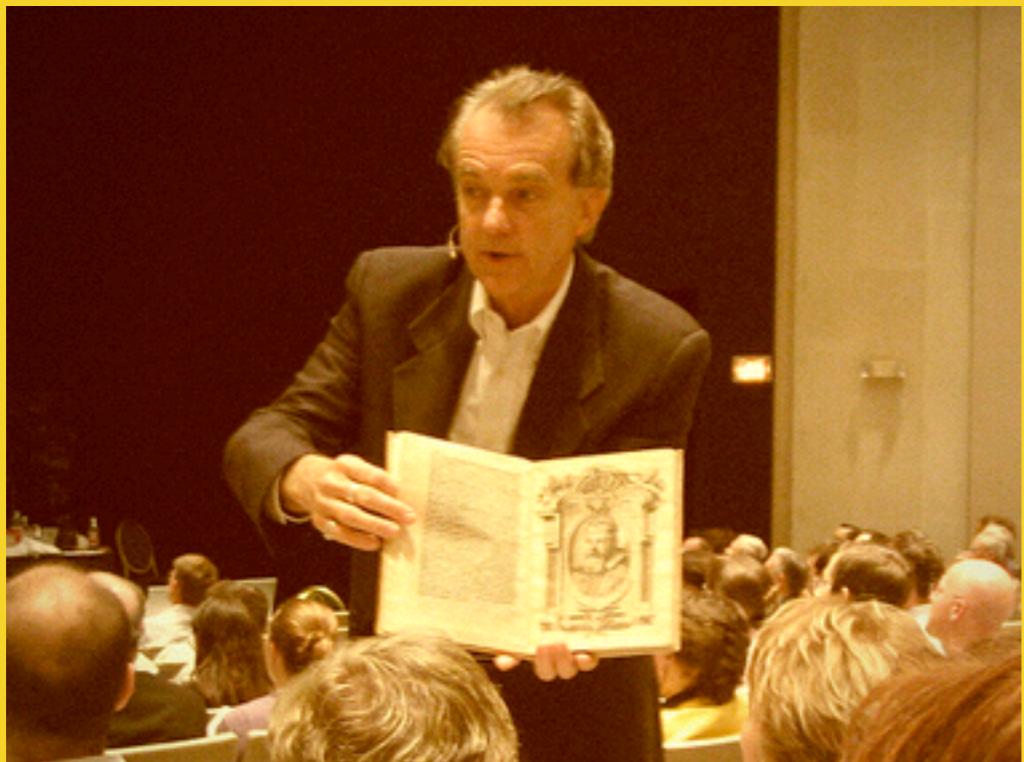
XIX century: Charles Joseph Minard. “Infographic”

# Brief history



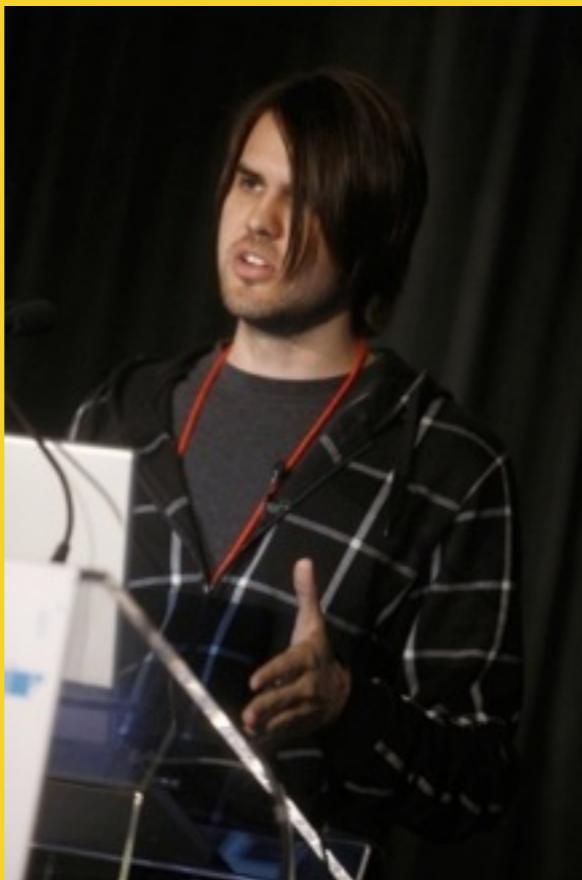
Siglo XX: John Tukey. FFT. Box Plot. Exploratory Data Analysis

# Brief history



Edward Tufte. 1983

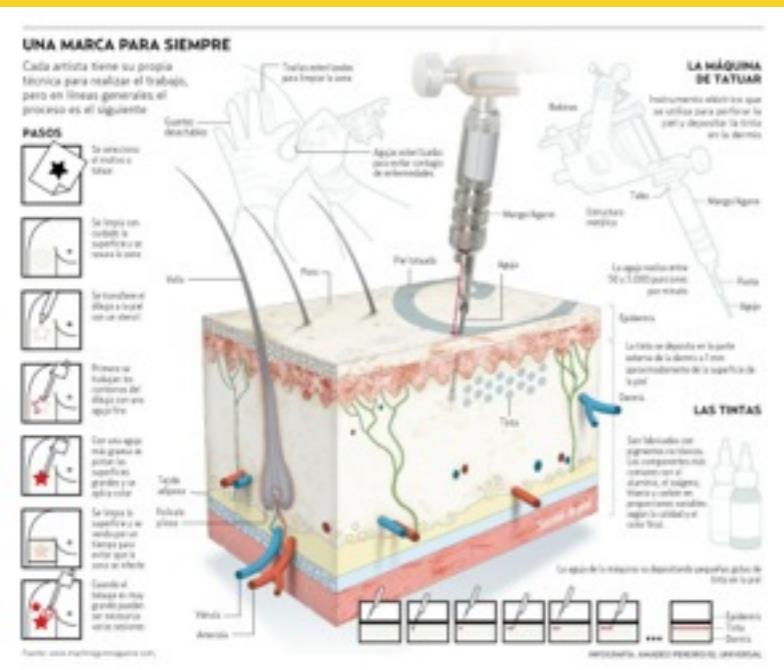
# Brief history



XXI century: Alberto Cairo, Aaron Koblin, Mike Bostock

# Brief history

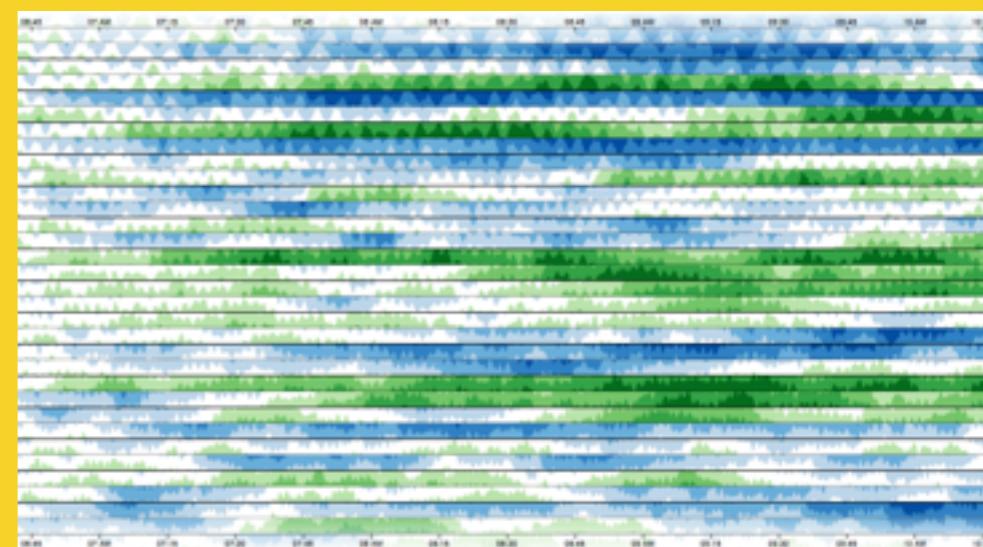
Cairo



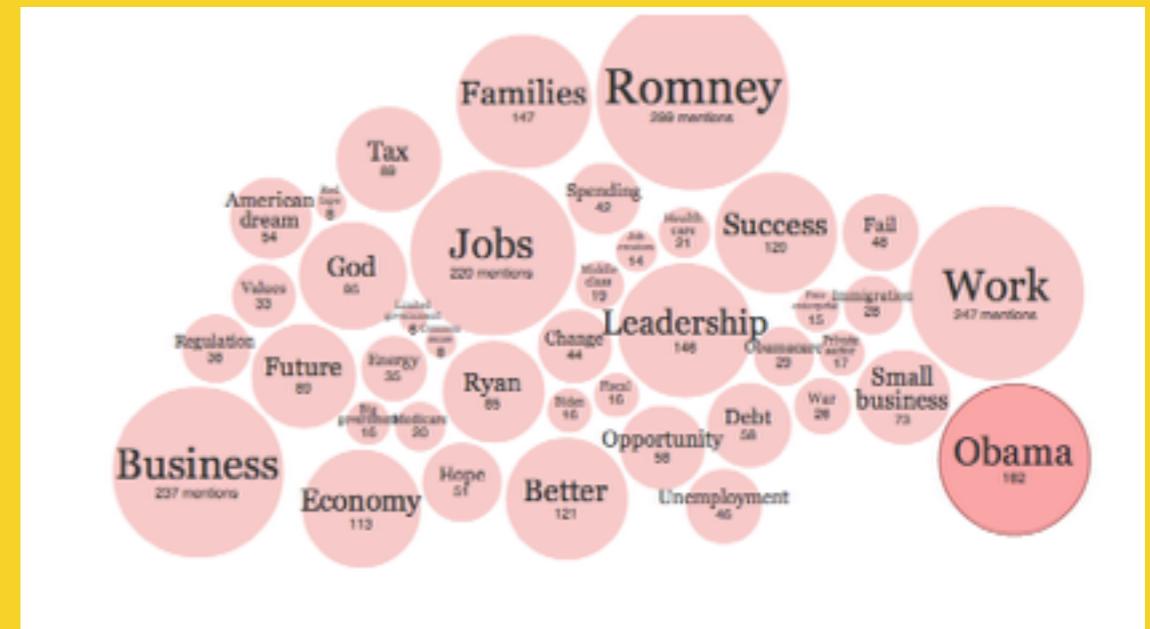
Koblin



Bostock



# Narrative matters, too



[Link](#)

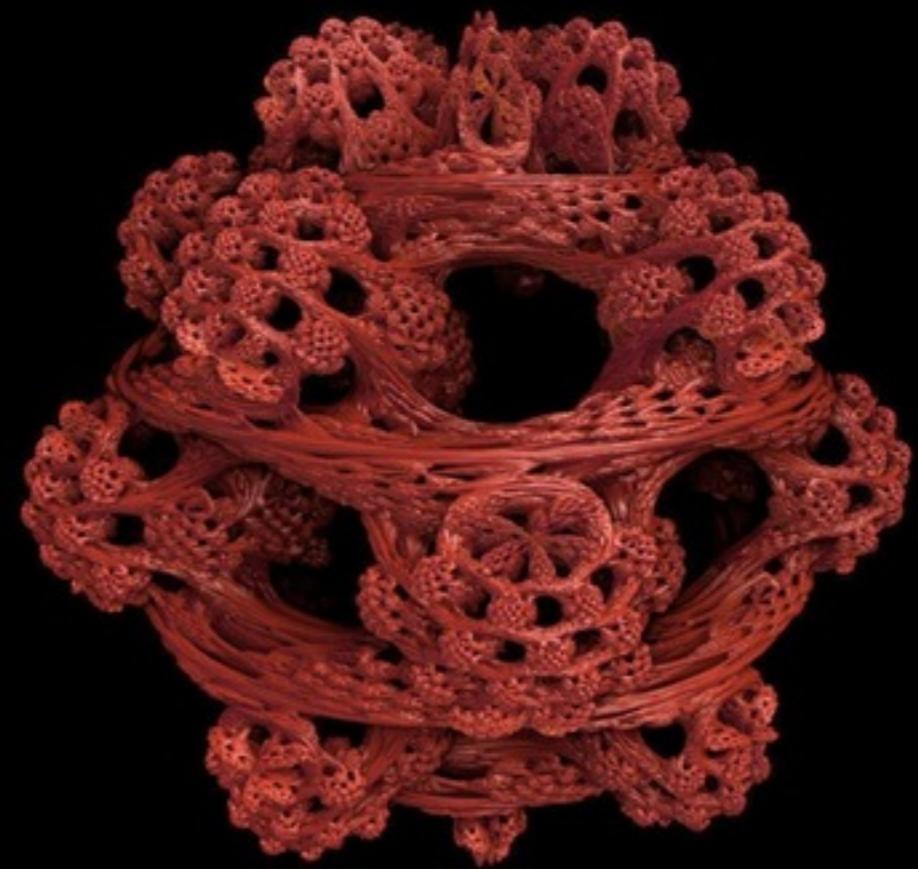
[Link](#)

# Narrative matters, too



[Link](#)

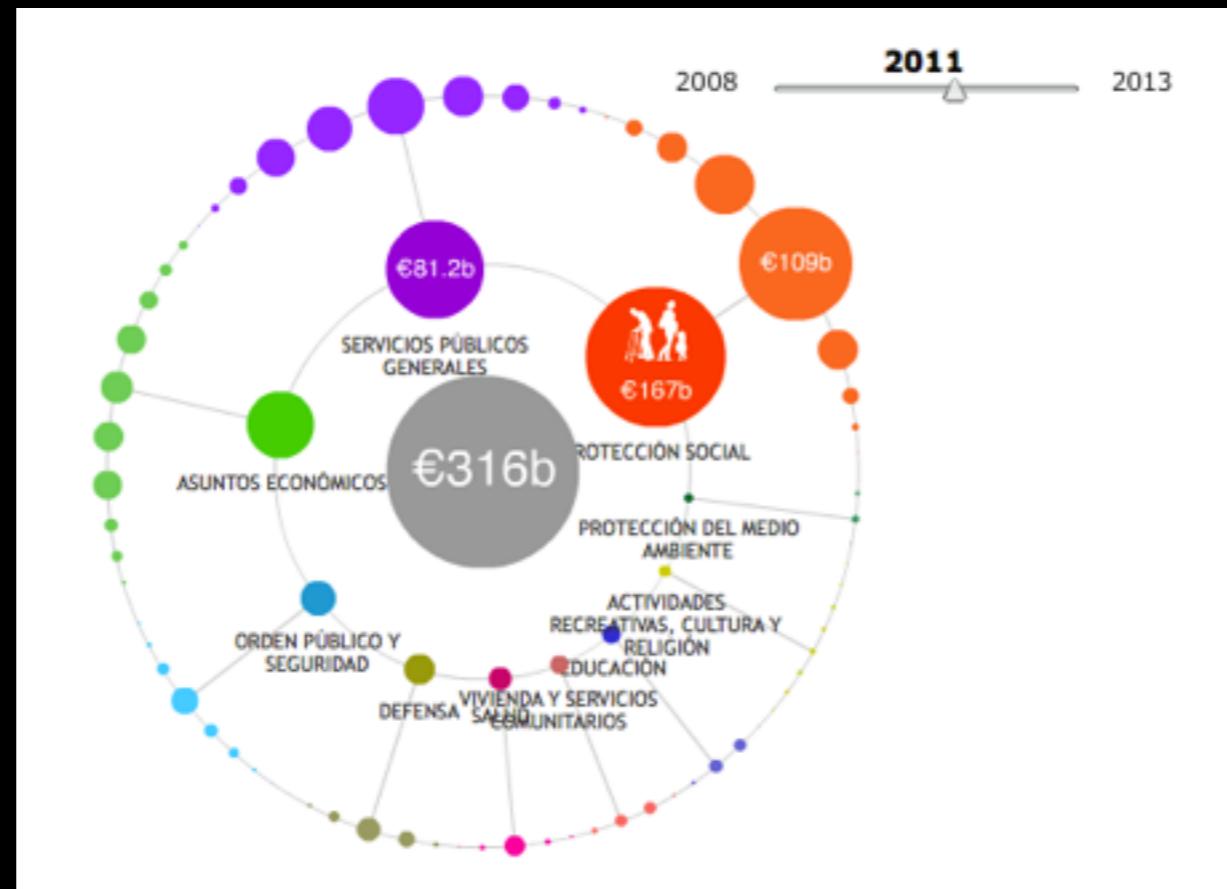
# Art vs function



(generative) Art

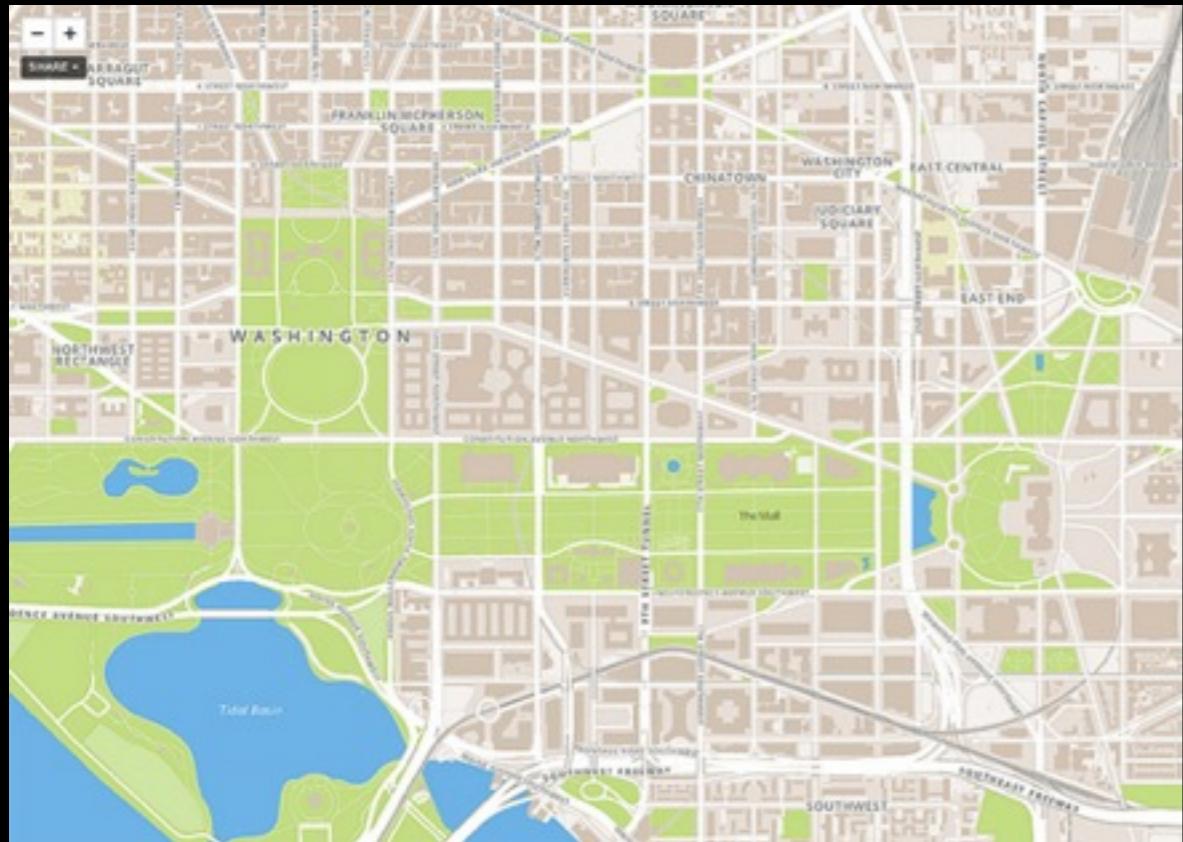
[Link](#)

# Art vs function

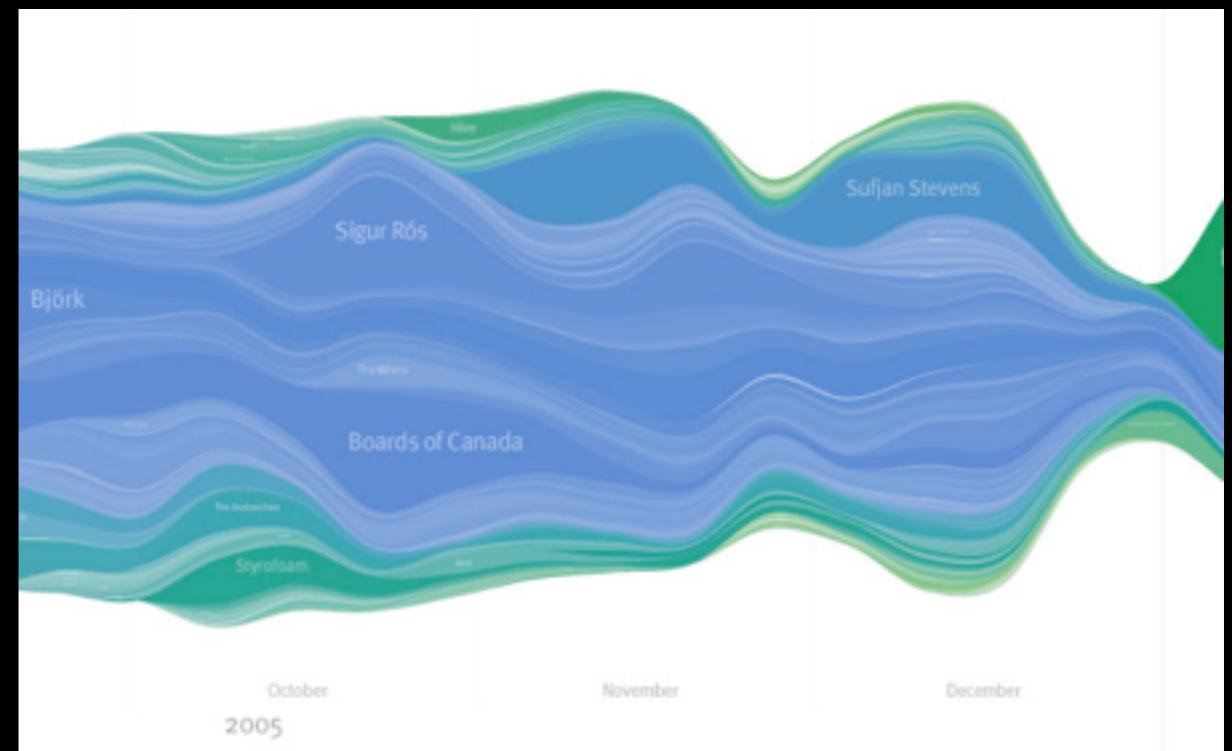


Function  
Link

# Figurative vs non-figurative

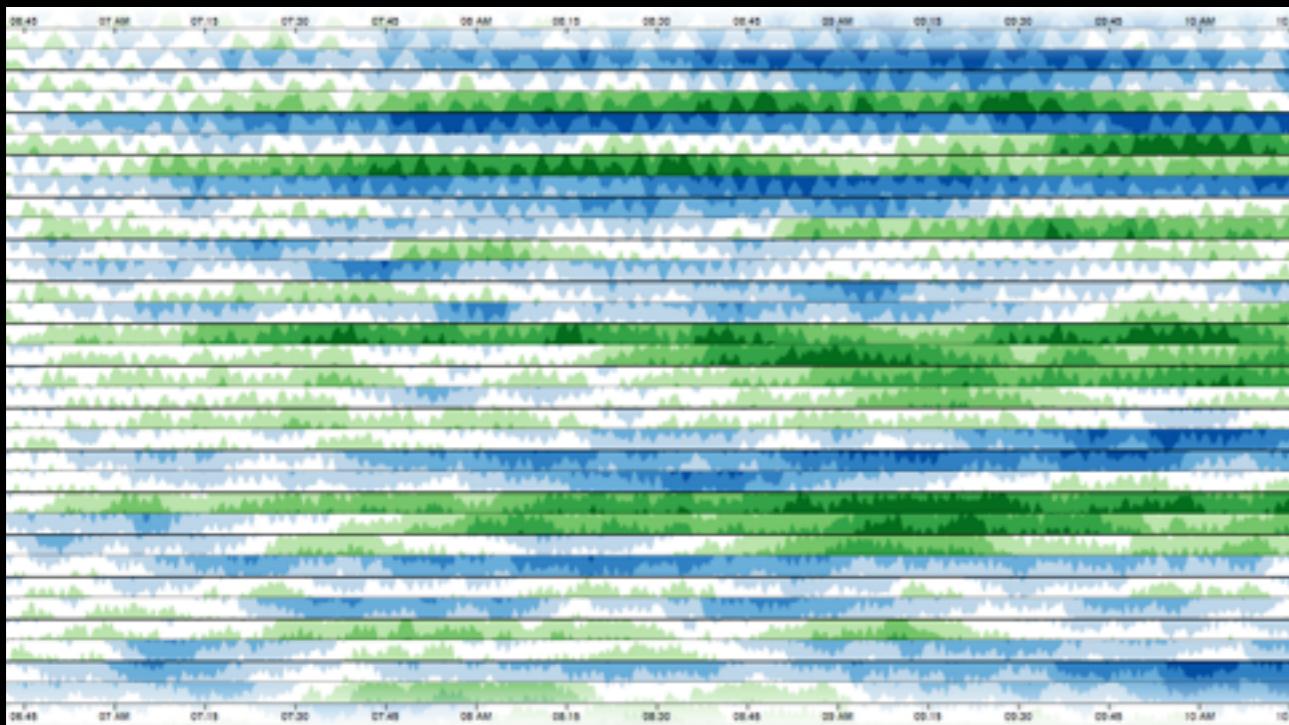


Figurative

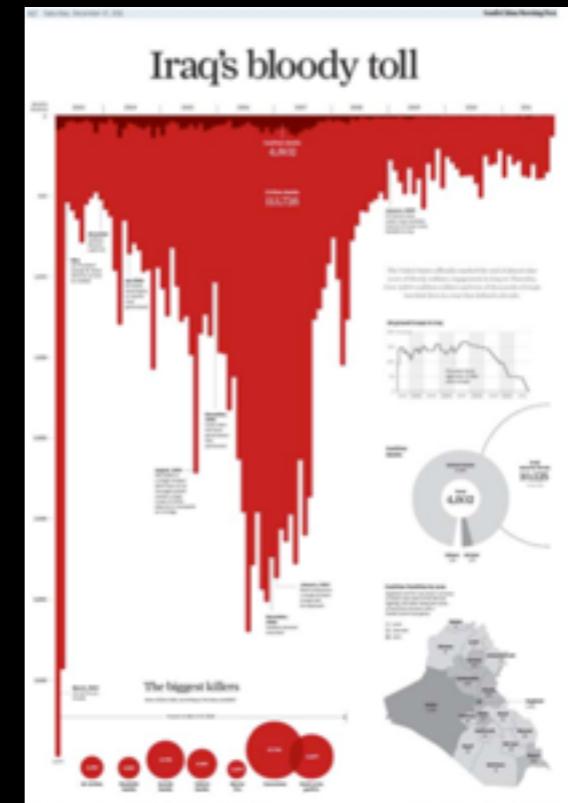


Non-figurative

# Tone



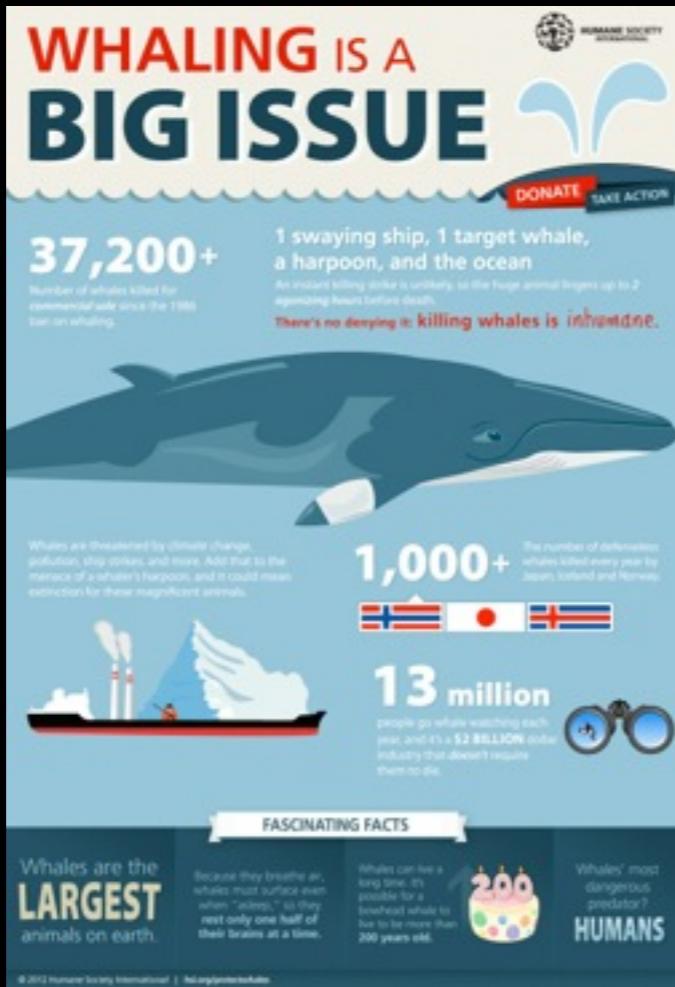
Analysis



Emotion

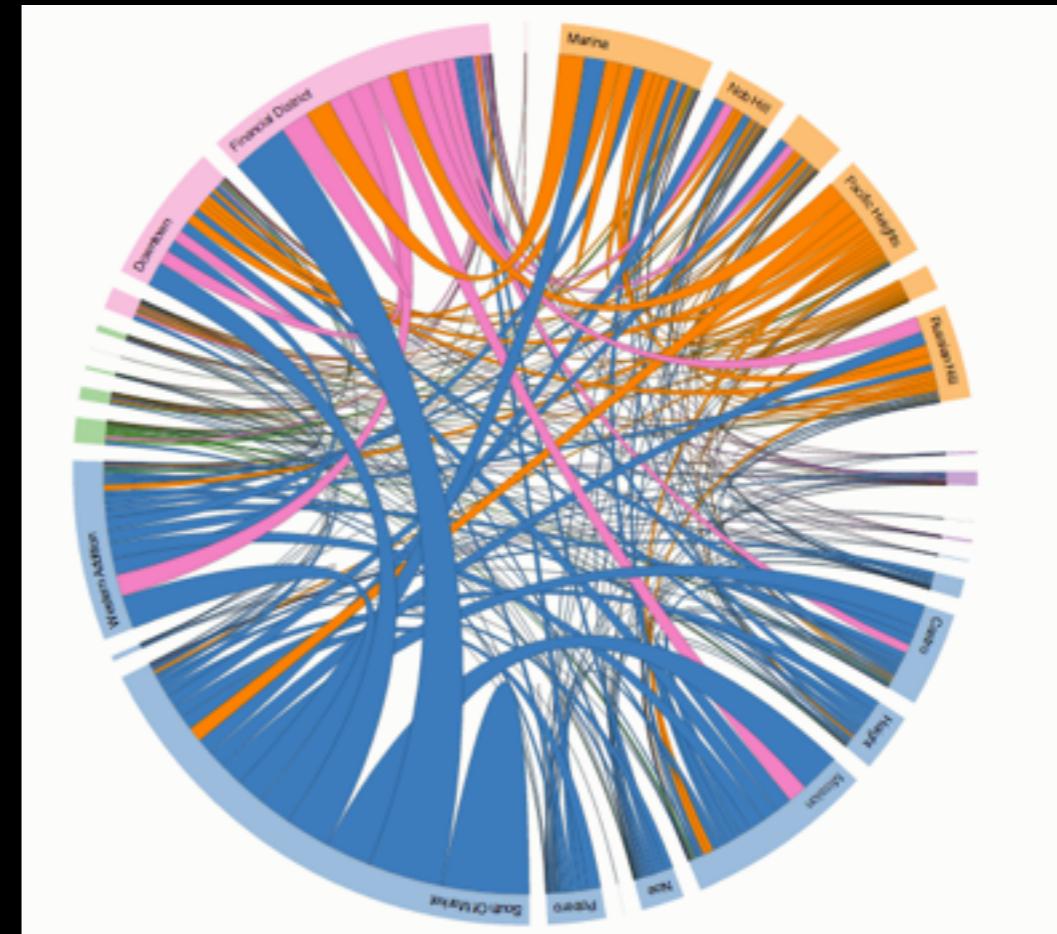
[Link](#)

# Intention: Expose vs Explore



Expose

[Link](#)



Explore

[Link](#)

# Data ‘dimensions’

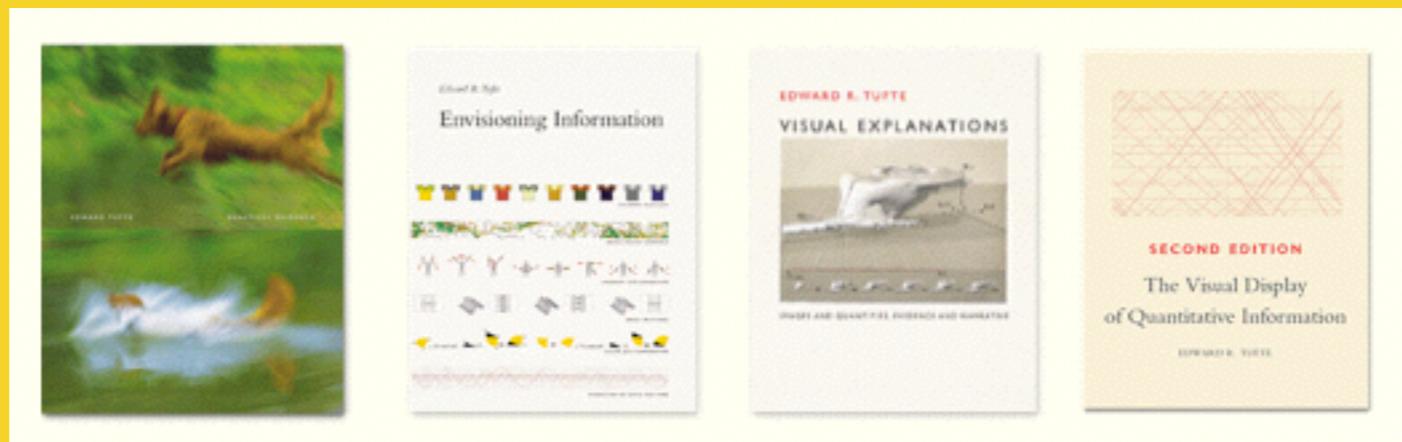
## LATCH



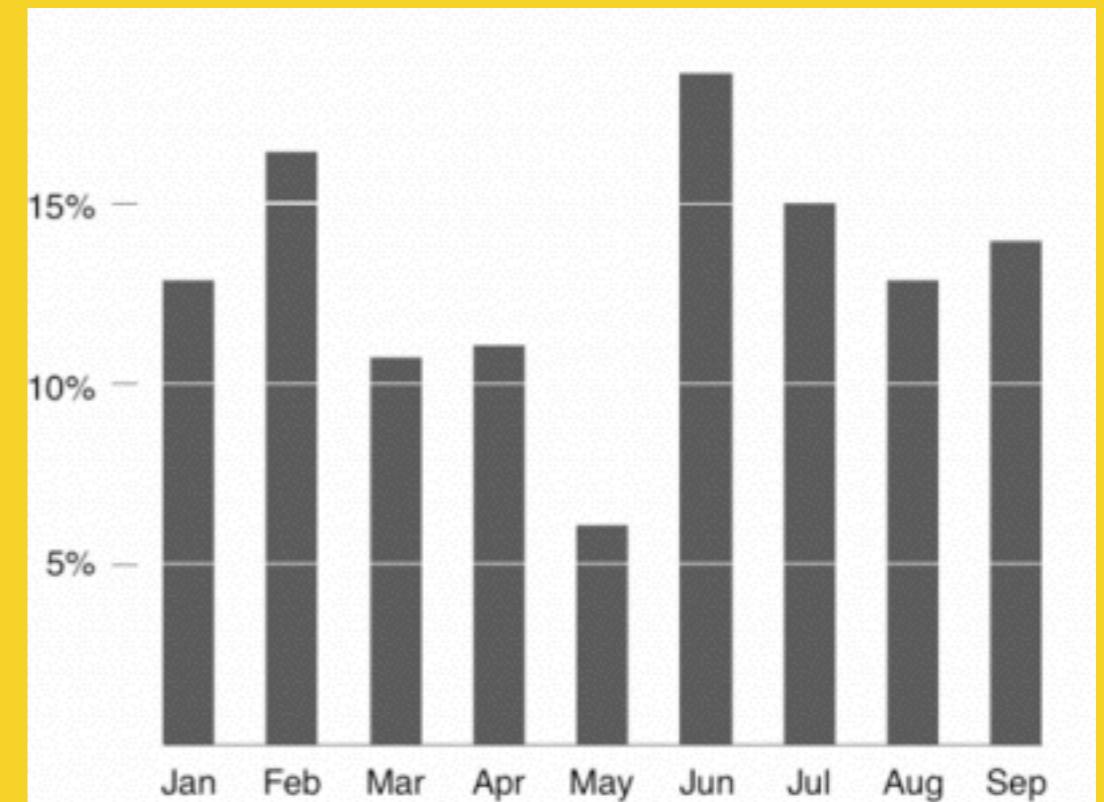
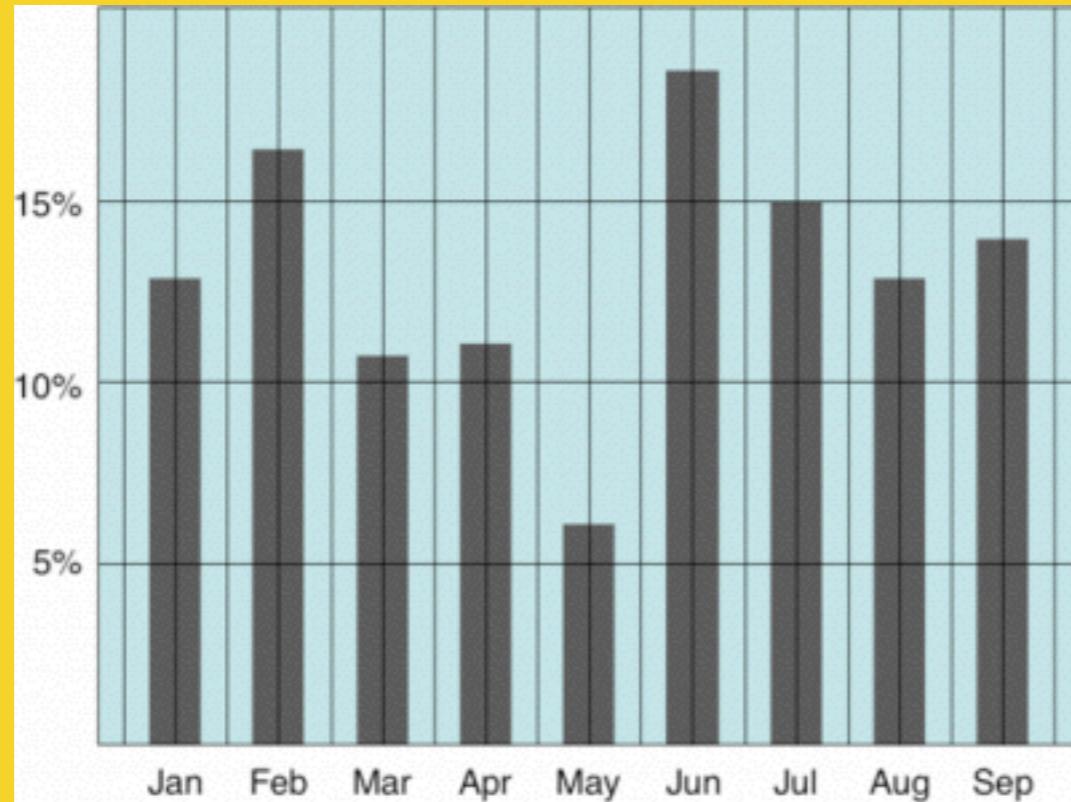
- ★ Location
- ★ Alphabet
- ★ Category
- ★ Hierarchy

# Relevant people: Tufte

- ★ Edward Rolf Tufte
- ★ Personal battle against ‘chartjunk’
- ★ Statistician and emeritus teacher of Political Sciences, Statistics and Computer Science at Yale University



# Relevant people: Tuft

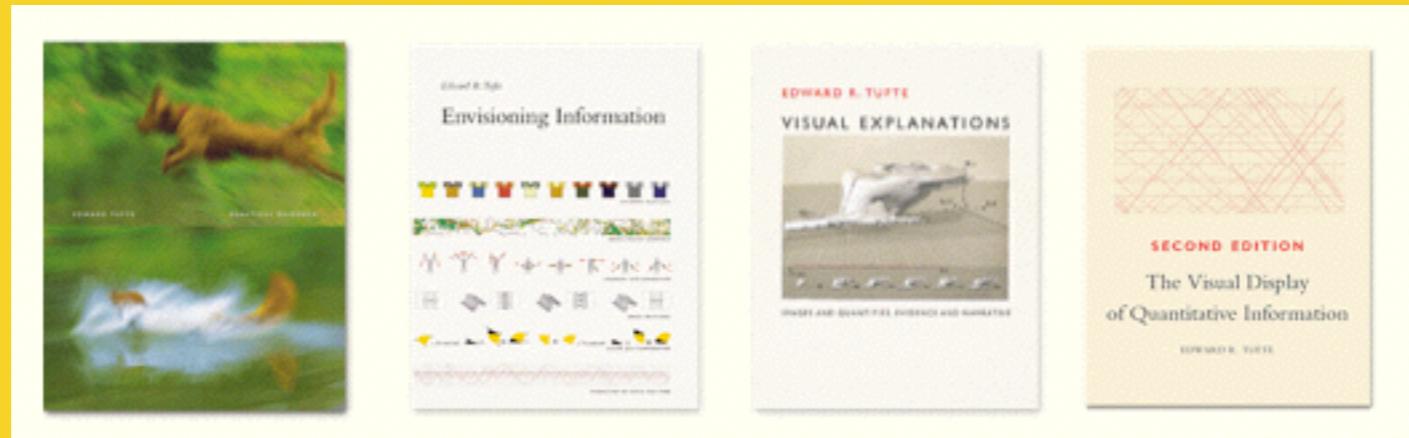


Data-Ink Ratio

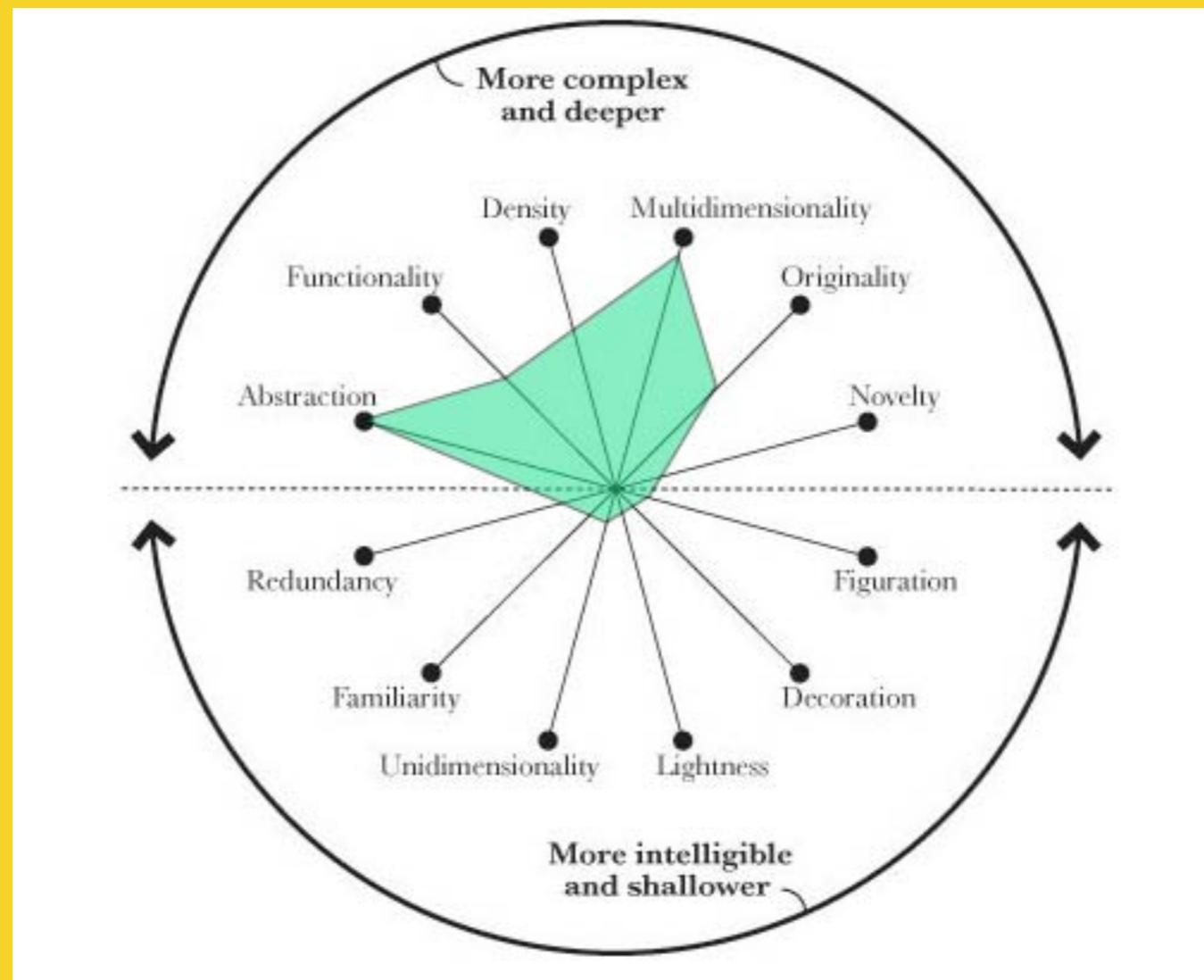
[Link](#)

# Relevant people: Cairo

- ★ Alberto Cairo
- ★ Infographics and information visualisation consultant
- ★ Teacher and journalist
- ★ Head of infographics at ‘El Mundo’ [2000-2005]
- ★ “Infographics 2.0” and “The Functional Art” author
- ★ Journalist approach to data visualisation



# Relevant people: Cairo



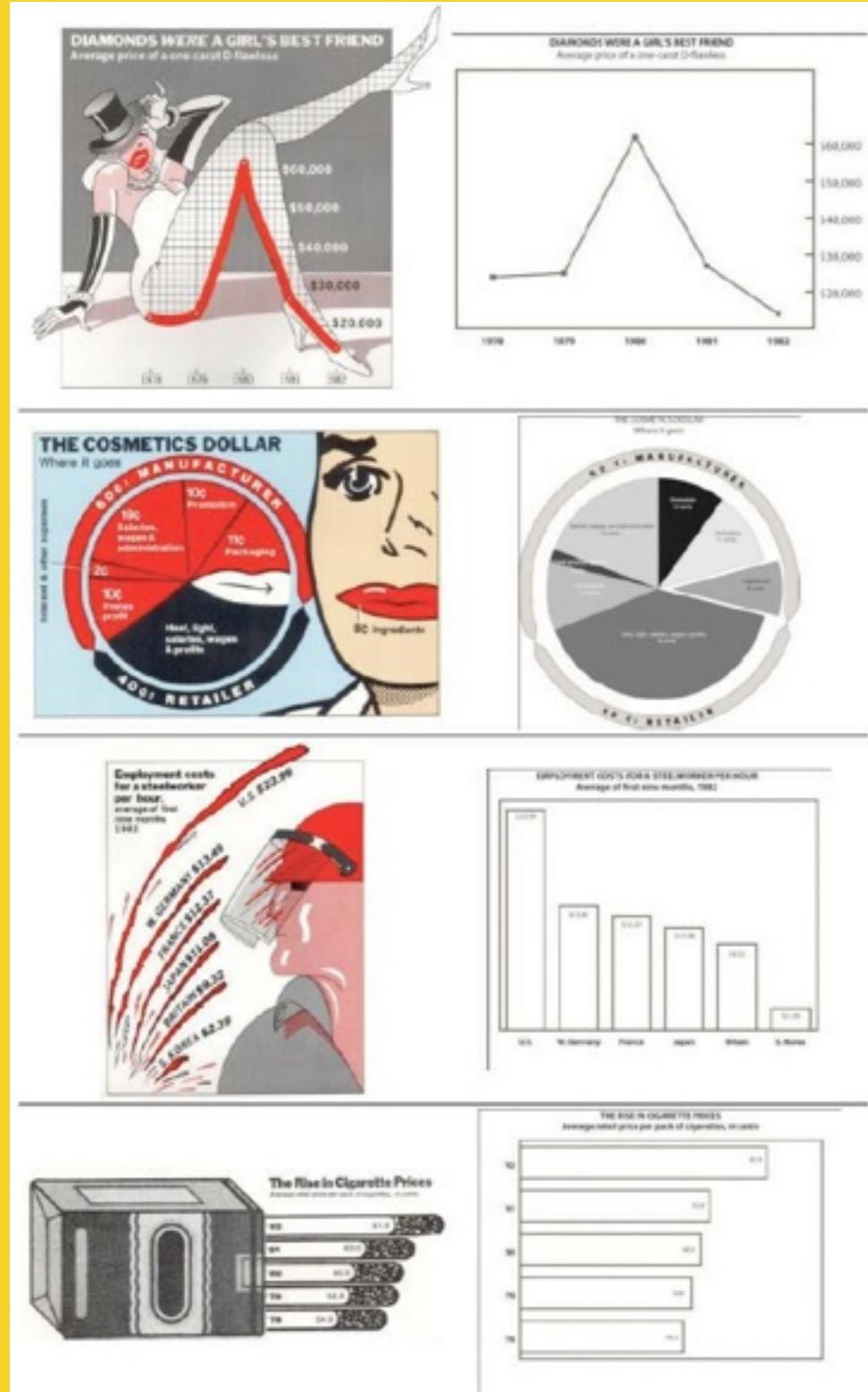
“Visual wheel”

[Link](#)

# Relevant people: Holmes

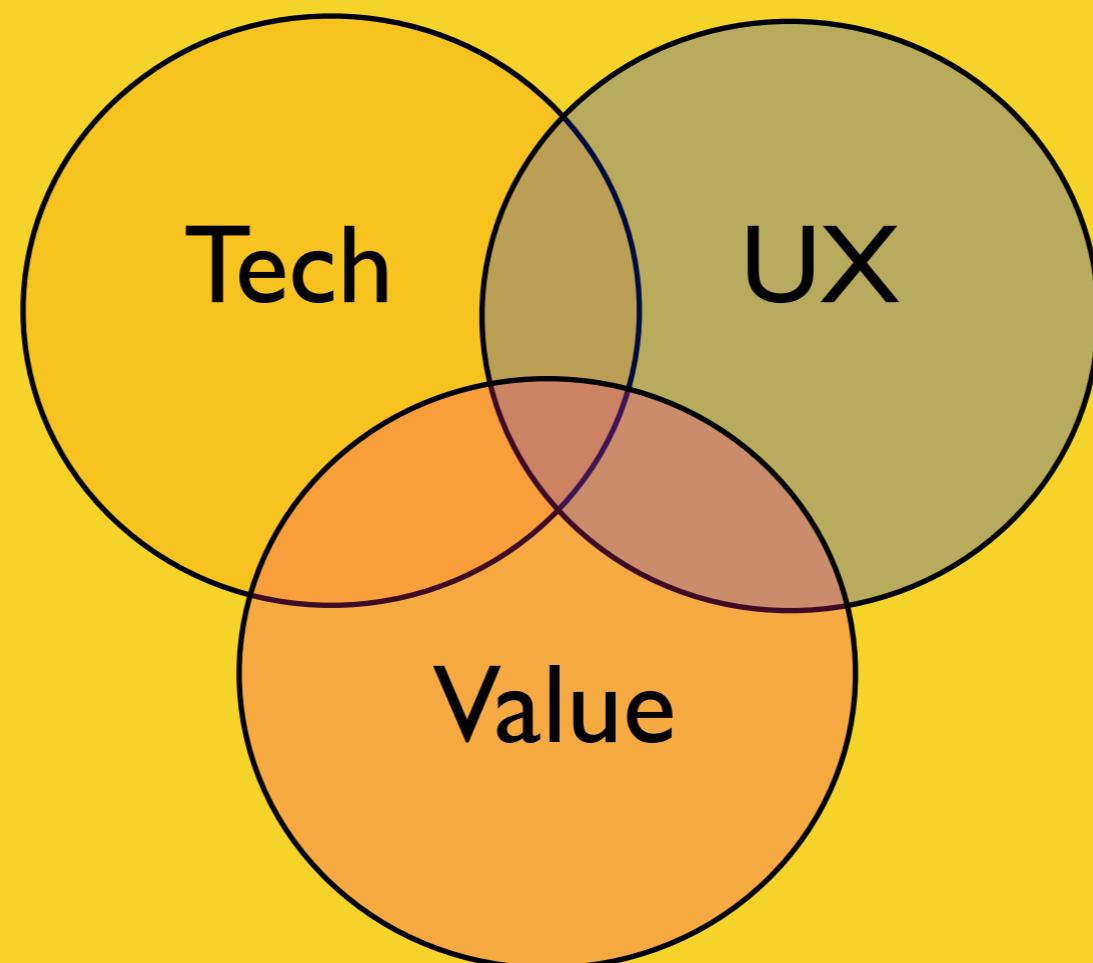
- ★ Nigel Holmes
- ★ Graphic Designer
- ★ Visually rich infographics
- ★ Has written eight books about graphic design and information visualization
- ★ Worked for ‘Times Magazine’ [1978-1994], designing ‘explanation graphics’

# Relevant people: Holmes

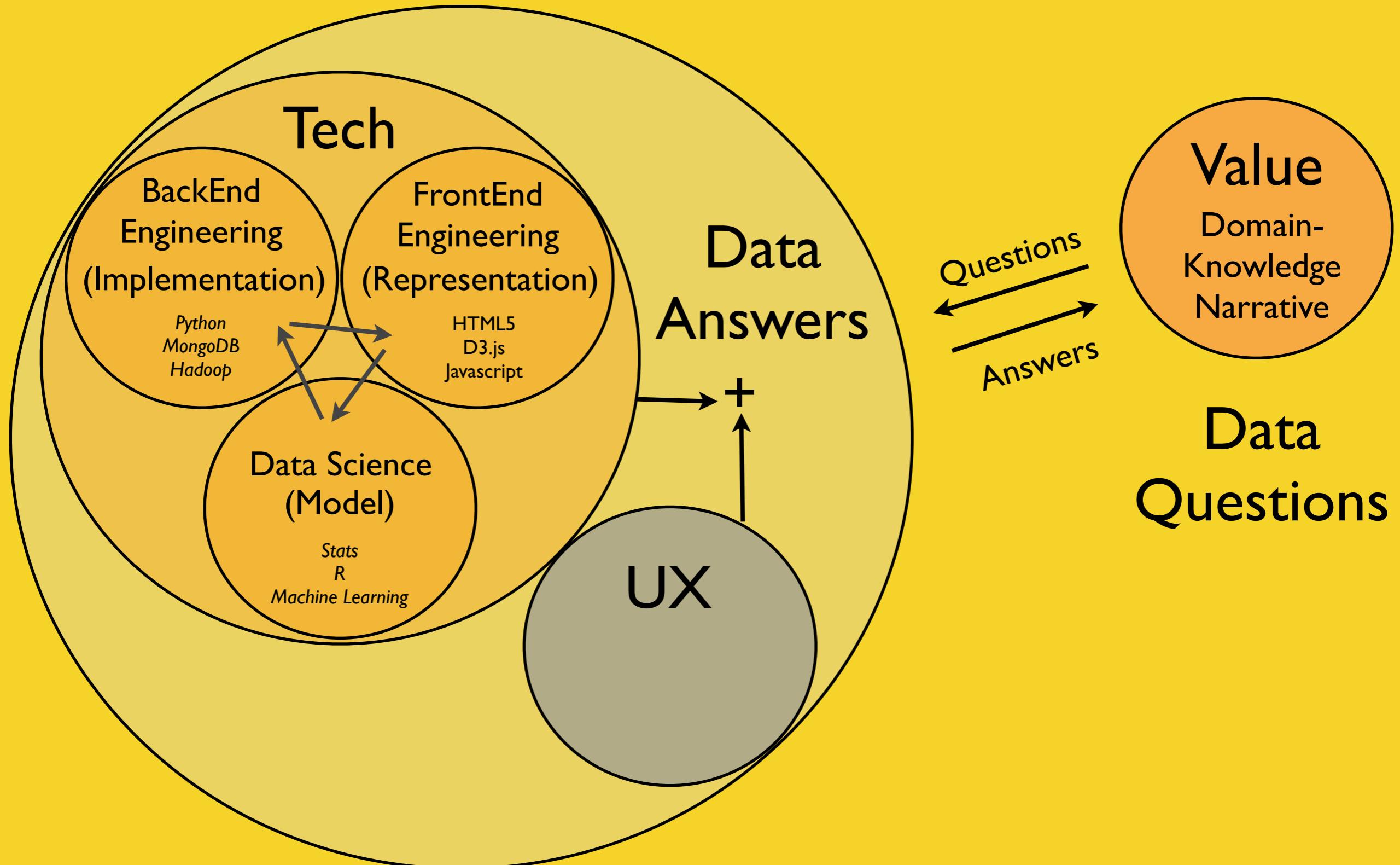


- ★ Saskatchewan University: 2010
- ★ Short+ long term recall was measured for:
  - ★ Central subject
  - ★ Variables
  - ★ Forms
  - ★ Subjectivity
- ★ In 'chartjunk' cases, at short term, both subject and form was much better recalled
- ★ In the long term, only subject was much better remembered

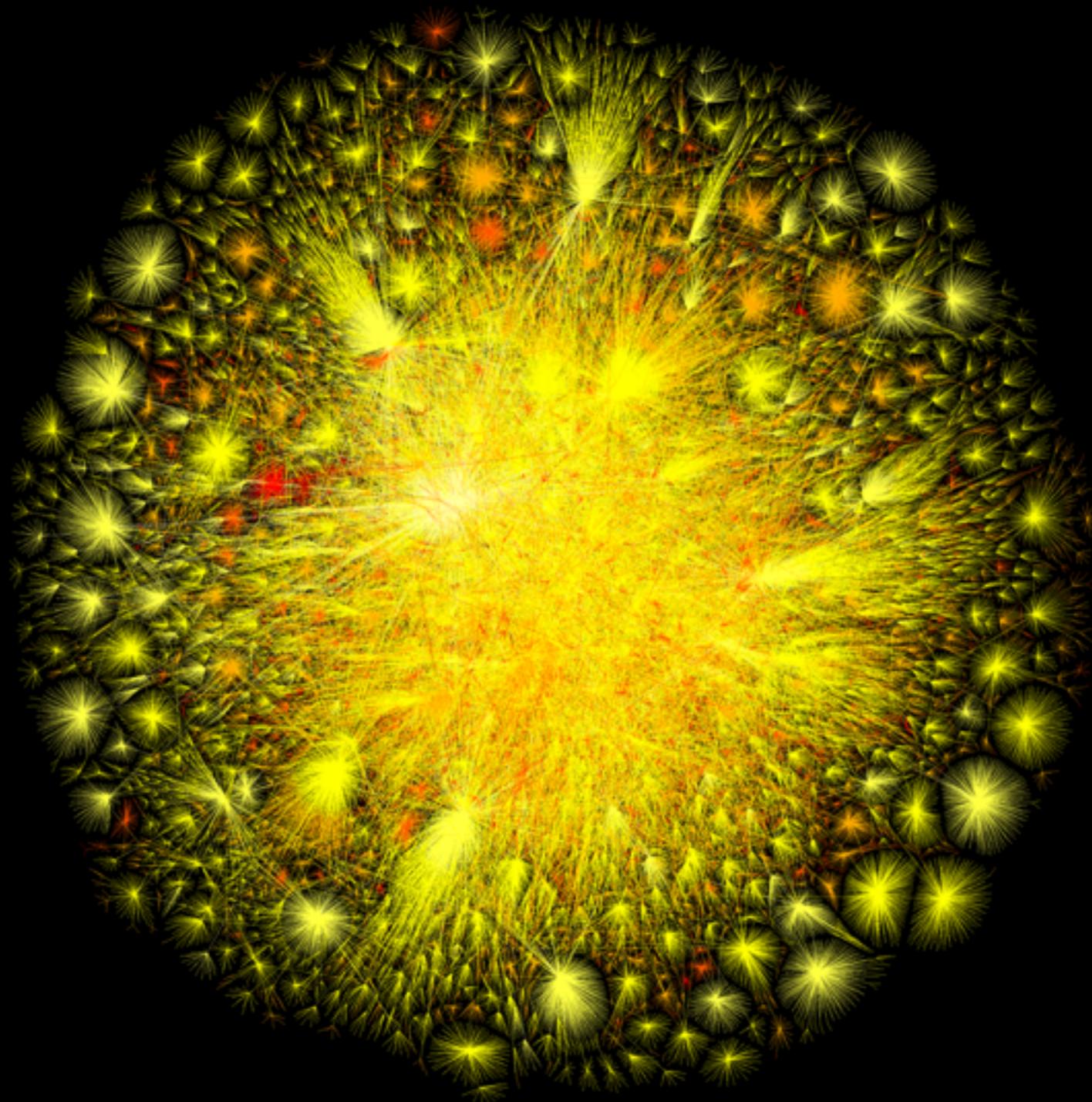
# Meet the Dataviz Team (I)



# Meet the Dataviz Team (I)



# Visualization tools



OPTE Project

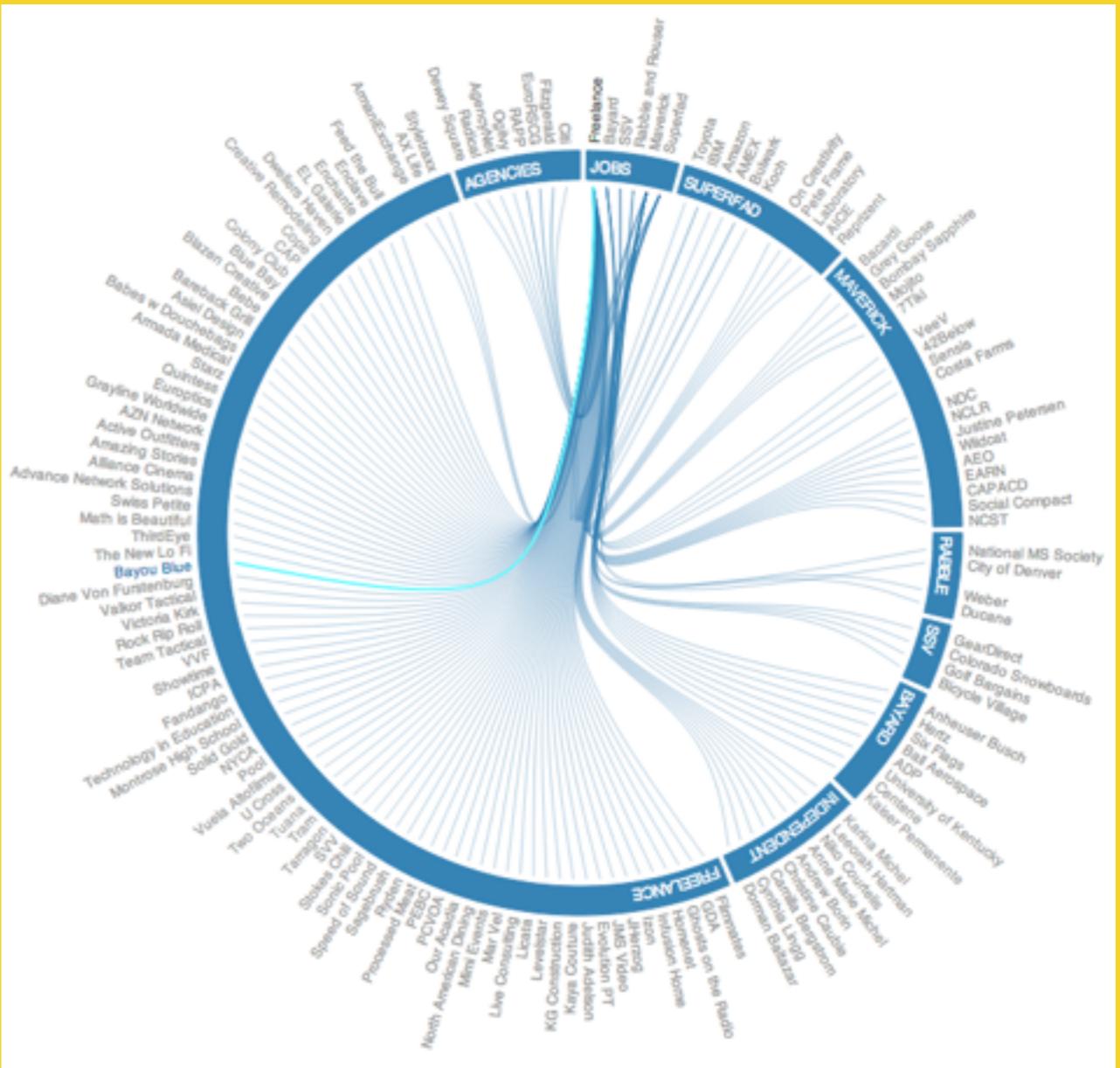
# Coding tools: Processing

- ★ Desktop oriented
- ★ Java-based
- ★ Good for multimedia artists

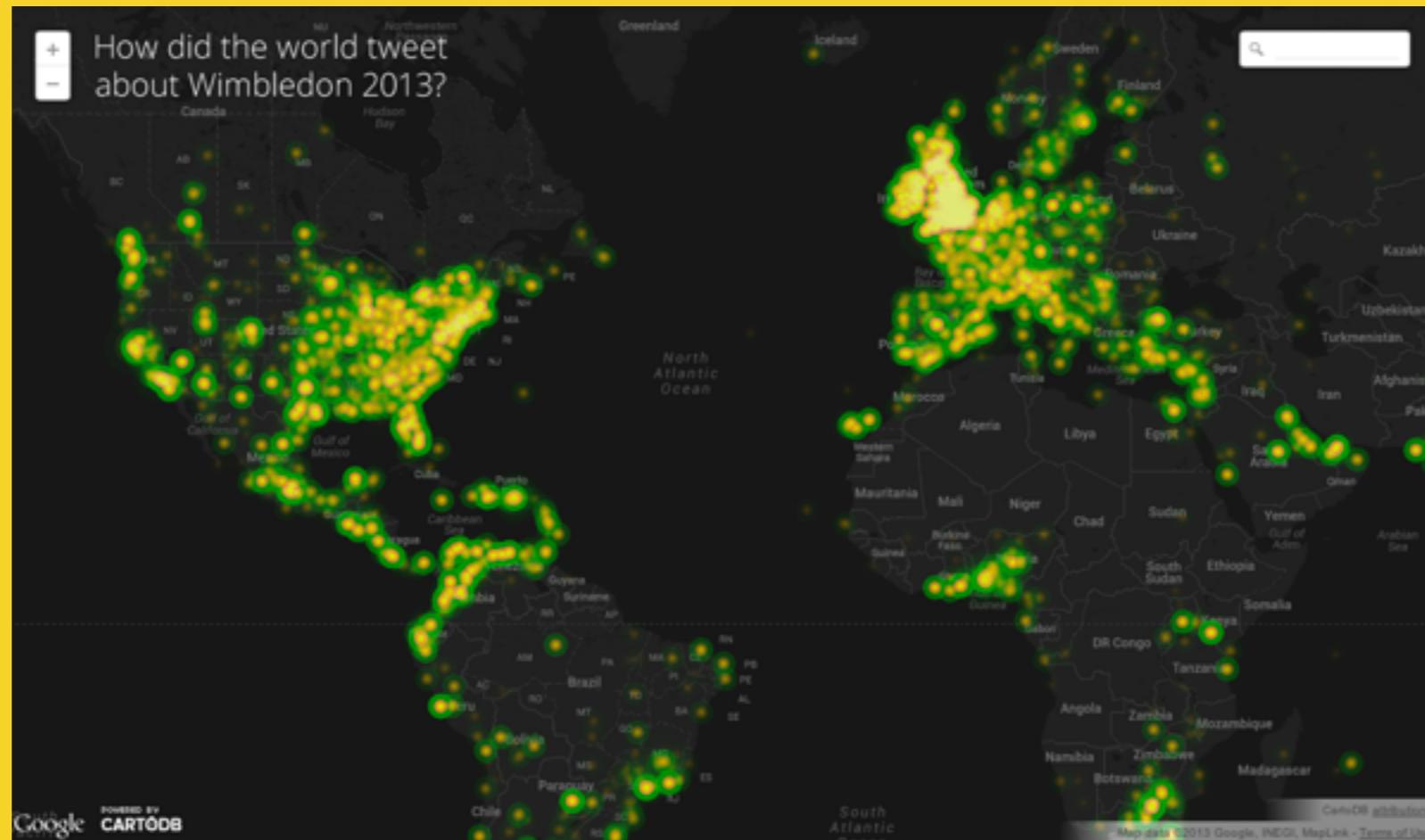


# Coding tools: D3.js

- ★ Web oriented
- ★ HTML5 based
- ★ Good for frontend developers



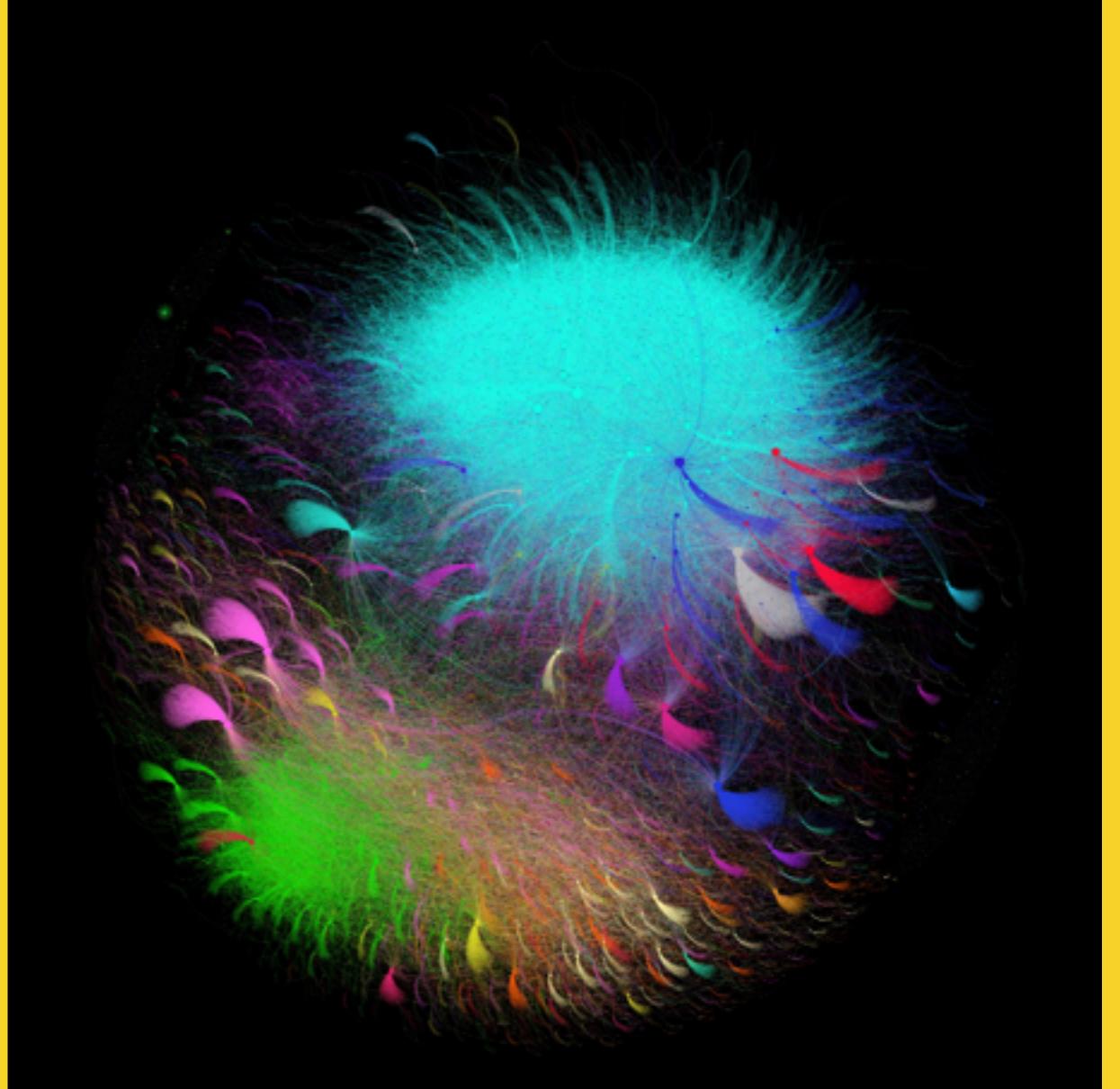
# CartoDB



- ★ ‘Excel’ for maps
- ★ Free (limited) version
- ★ Continuous improvement

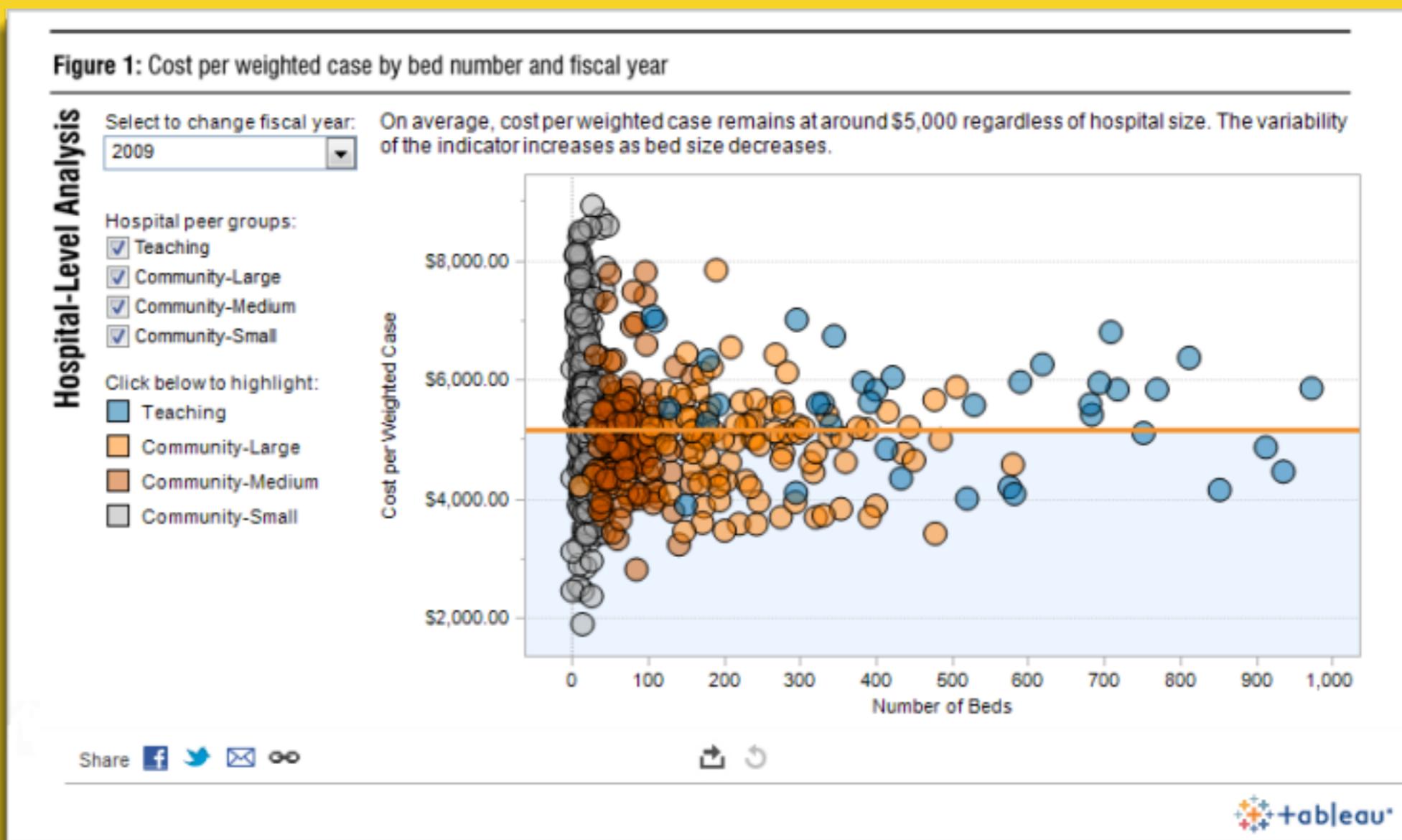
# Gephi

- ★ ‘Excel’ for networks
- ★ Network analysis also
- ★ 100% free



[Link](#)

# Tableau



- ★ ‘Excel’ for everything else
- ★ Interactive dataviz+dashboards
- ★ Free ‘Tableau Public’ version

# References



# Further reading

- ★ “The functional art” - Alberto Cairo
- ★ “The visual display of quantitative information” - Edward R. Tufte
- ★ “Information Graphics” - Several authors
- ★ “Dataclysm” - Christian Rudder

Questions?  
Thanks for your attention!