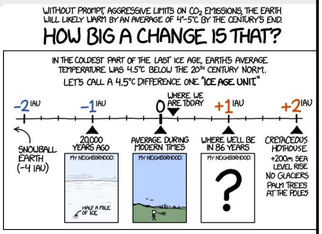
The role of visualization in the hypothetico-deductive method

Jean-Daniel Fekete, Inria
http://www.aviz.fr/~fekete

Two visualizations - two roles


WITHOUT PROMPT, AGGRESSIVE LIMITS ON CO₂ EMISSIONS, THE EARTH WILL LIKELY WARM BY AN AVERAGE OF 4°-5°C BY THE CENTURY'S END.
HOW BIG A CHANGE IS THAT?
IN THE COLDEST PART OF THE LAST ICE AGE, EARTH'S AVERAGE TEMPERATURE WAS 4.5°C BELOW THE 20TH CENTURY NORM.
LET'S CALL A 4.5°C DIFFERENCE ONE "ICE AGE UNIT."
-2 IAU
-1 IAU
0
WHERE WE ARE TODAY
+1 IAU
+2 IAU
SNOWBALL EARTH (-4 IAU)
20,000 YEARS AGO
AVERAGE DURING MODERN TIMES
WHERE WE'LL BE IN 86 YEARS
CRETACEOUS HOTHOUSE
+200m SEA LEVEL RISE
NO GLACIERS
PALM TREES AT THE POLES
MY NEIGHBORHOOD:
HALF A MILE OF ICE
MY NEIGHBORHOOD:
HI!
MY NEIGHBORHOOD:
?

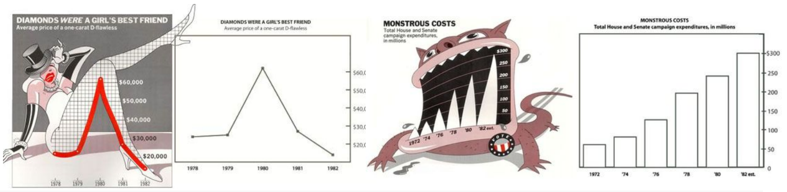
http://aviz.fr/~bbach/datacomics/

Visualization for exploration

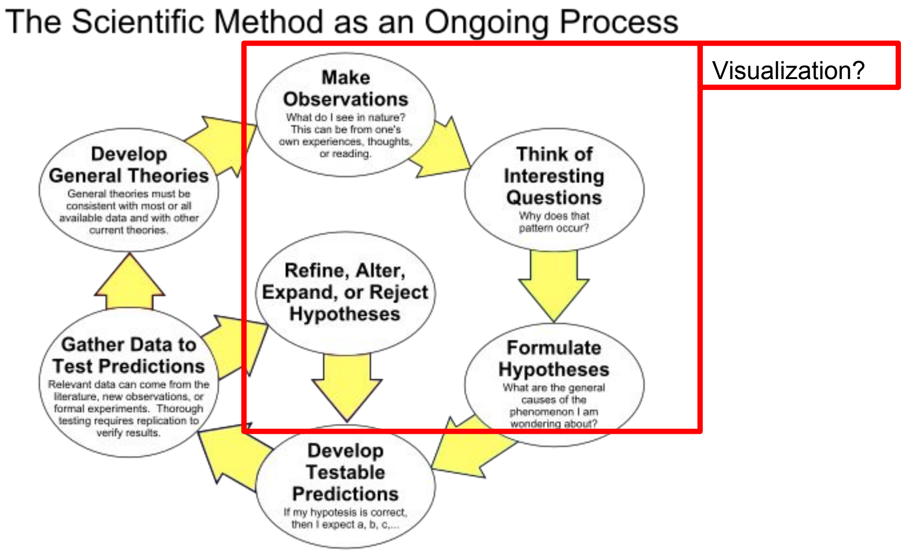
- Topic of this talk

Visualization for communication

- Communication is a rhetorical device
- Everything that helps communicate is fine
- But it should be understandable by a large audience



Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, Christopher Brooks, Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. *ACM* CHI, 2010.


The Scientific Method as an Ongoing Process
Visualization?
Make Observations
What do I see in nature? This can be from one's own experiences, thoughts, or reading.
Think of Interesting Questions
Why does that pattern occur?
Formulate Hypotheses
What are the general causes of the phenomenon I am wondering about?
Develop Testable Predictions
If my hypotesis is correct, then I expect a, b, c,...
Gather Data to Test Predictions
Relevant data can come from the literature, new observations, or formal experiments. Thorough testing requires replication to verify results.
Develop General Theories
General theories must be consistent with most or all available data and with other current theories.
Refine, Alter, Expand, or Reject Hypotheses

By ArchonMagnus (Own work) [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons

The role of visualization

What is visualization good at?

- Perceiving patterns in data

How (well) does it work?

- Let's see

Why is it important?

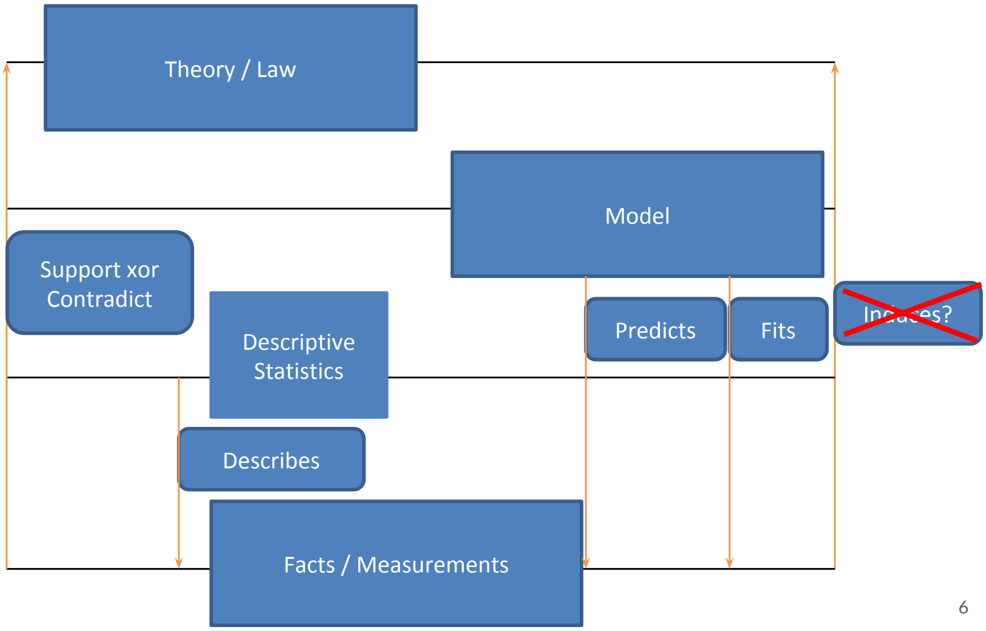
- Key question today: **how (well) does it fit in the scientific method?**

Visualization for exploration

Compact graphical presentation and user interface
for manipulating large numbers of items,
possibly extracted from far larger datasets.
Enables users to make
discoveries, decisions, or explanations
about
patterns (trend, cluster, gap, outlier...), groups of items, or individual items.
[Plaisant, 2001]

The study of (interactive) visual representations of [abstract] data to reinforce human cognition
[Wikipedia, information visualization, 2018]

Science 1.0



Theory / Law

Model

Support xor Contradict

Descriptive Statistics

Predicts

Fits

~~Induces?~~

Describes

Facts / Measurements

Stats vs. visualization: the Anscombe's Quartet

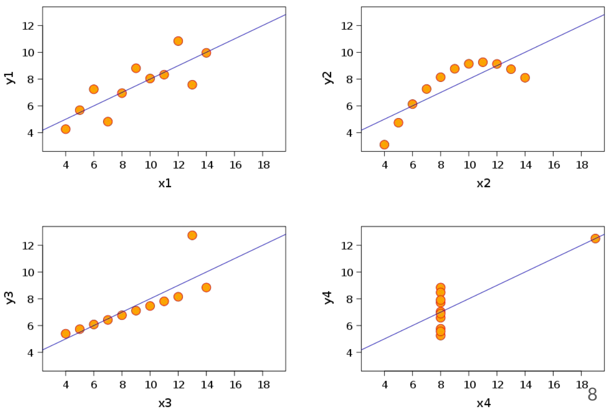
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	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

Mean of *x*	9.0
Variance of *x*	11.0
Mean of *y*	7.5
Variance of *y*	4.12
Correlation between *x* and *y*	0.816
Linear regression line	*y* = 3 + 0.5*x*

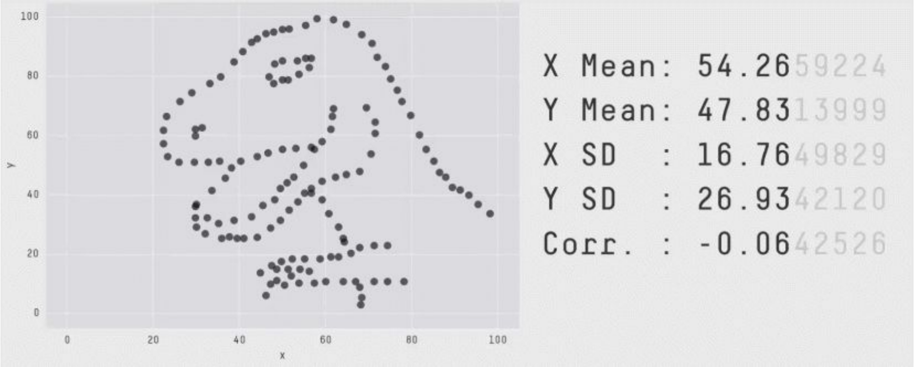
[Source: Anscombe's quartet, Wikipedia]

Visualization reveals a different story

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

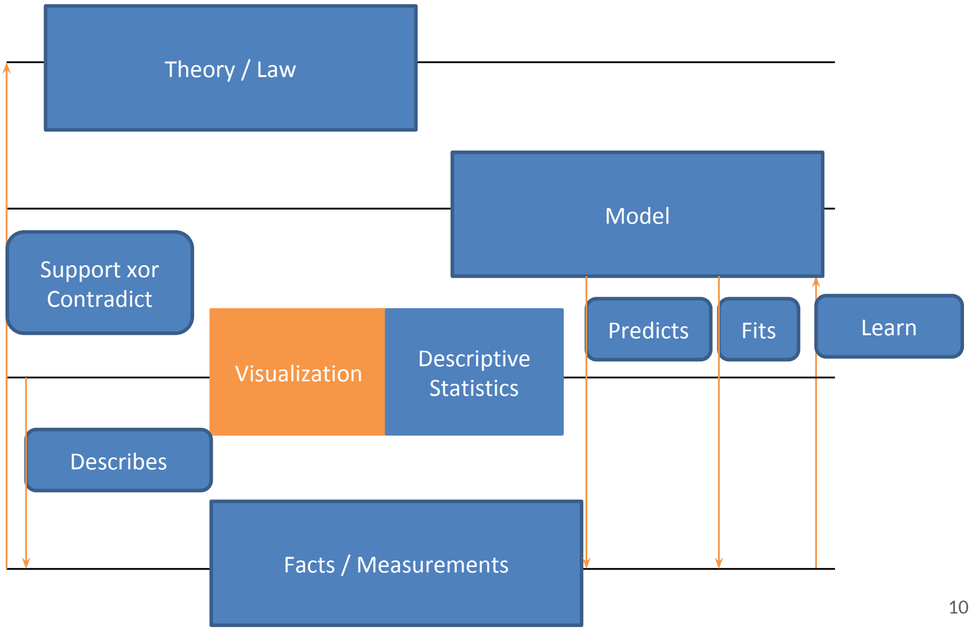


The same stats can take ~ any shape



J. Matejka and G. Fitzmaurice. 2017. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. ACM CHI'17
https://www.autodeskresearch.com/publications/samestats

Science 2.0



Theory / Law

Model

Support xor Contradict

Predicts

Fits

Learn

Visualization

Descriptive Statistics

Describes

Facts / Measurements

Science 2.0 relies more heavily on data

- With the increase of storage, more data is available, more opportunities for exploration
- More data is gathered from sensors
- More data is available online
- Exploration without a-priori hypotheses is possible
  - E.g. Metagenomics
- Data Driven Science (a.k.a Data Science) is now becoming popular
- Visualization allows rapid exploration of data
  - Fast feedback loop to express hypotheses, confirm or disprove
- Is this new method valid?

Visualization for exploration

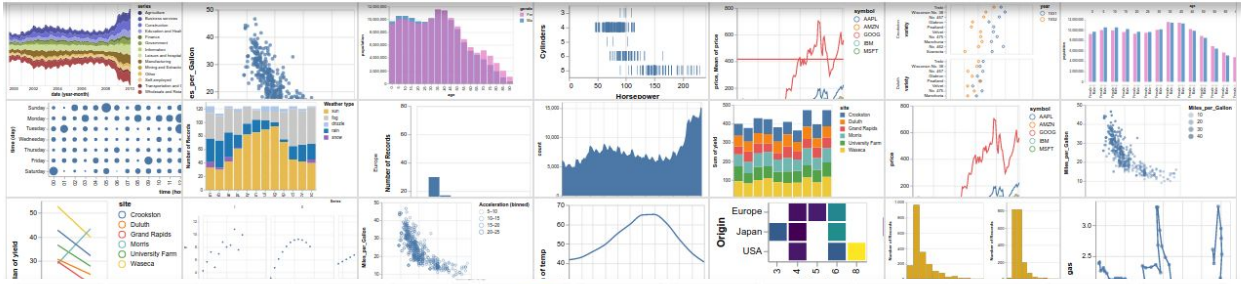
Tremendous progress in the last 15 years

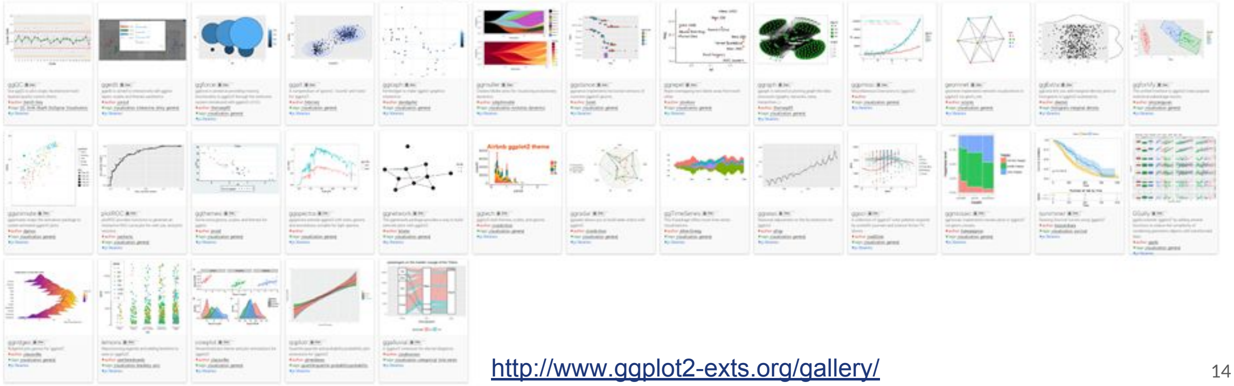
- Almost complete model for building visualizations
- Some understanding of how interaction works (although much more is needed)
- Understanding of fundamental visual perception capabilities
- Understanding of color perception from low-level perception up to some level of cognition
- Understanding of ensemble perception (correlation) and limitations (crowding, contrast)
- Better understanding of perception biases (change blindness, limitations of animations)
- Better understanding of cognitive biases
- Deeper understanding of specific visualization techniques (scatterplots, graphs)

Building visualizations in a principled way

- The Grammar of Graphics (Wilkinson, 2005) introduces the idea of describing visualizations through a few constructs
- Visualization is not limited to a series of baked components, it can be described and constructed consistently *(if not completely)*
  - gglot2 with R (Wickham, 2010) implements it with some variations
    - Hadley Wickham. A layered grammar of graphics. Journal of Computational and Graphical Statistics, vol. 19, no. 1, pp. 3–28, 2010.
  - Vega in JavaScript (...), implements it with extensions for specifying interactions
    - A. Satyanarayan, D. Moritz, K. Wongsuphasawat, J. Heer, Vega-Lite: A Grammar of Interactive Graphics, IEEE Trans. Visualization & Comp. Graphics 2017
    - Vega is accessible from other languages such as Python (Altair), Juli, R, all the ones compiled in JS, etc.

https://vega.github.io/vega-lite/examples/





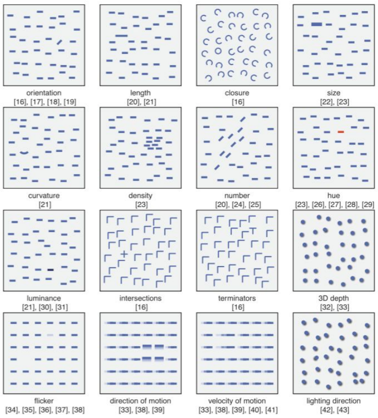
http://www.ggplot2-exts.org/gallery/

Fast perception with Preattentive Processing

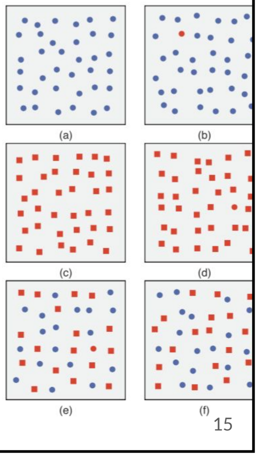
“tasks that can be performed on large multi-element displays in less than 200-250 milliseconds”: preattentive processing is done quickly, effortlessly and in parallel without any attention being focused on the display.

[Treisman, 1985] A. Treisman, Preattentive Processing in Vision, *Computer Vision, Graphics, and Image Processing*, 31(2):156-177, August 1985.

[Treisman, 1986] A. Treisman, Features and Objects in Visual Processing, *Scientific American*, 255(5):114-125, 1986.



orientation [16], [17], [18], [19]
length [20], [21]
closure [16]
size [22], [23]
curvature [21]
density [23]
number [20], [24], [25]
hue [23], [26], [27], [28], [29]
luminance [21], [30], [31]
intersections [16]
terminators [16]
3D depth [32], [33]
flicker [34], [35], [36], [37], [38]
direction of motion [33], [38], [39]
velocity of motion [33], [38], [39], [40], [41]
lighting direction [42], [43]



Rationales for charts

Wealth of information: https://github.com/widged/data-for-good/wiki/Visualisation-::-Choosing-a-chart

- A Tour through the Visualization Zoo https://queue.acm.org/detail.cfm?id=1805128
- Web sites surveying techniques
  - treevis.net
  - textvis.lnu.se
  - www.timeviz.net
  -

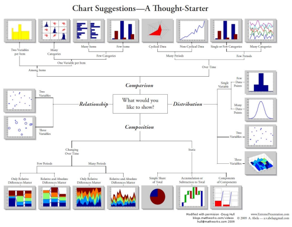
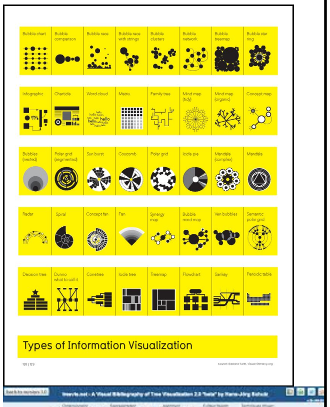
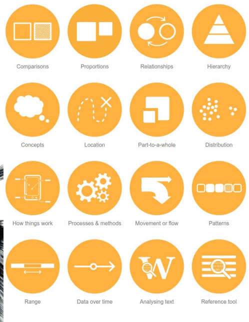
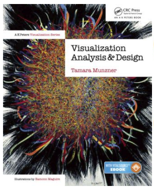


Chart Suggestions—A Thought-Starter
What would you like to show?
Comparison
Relationship
Distribution
Composition



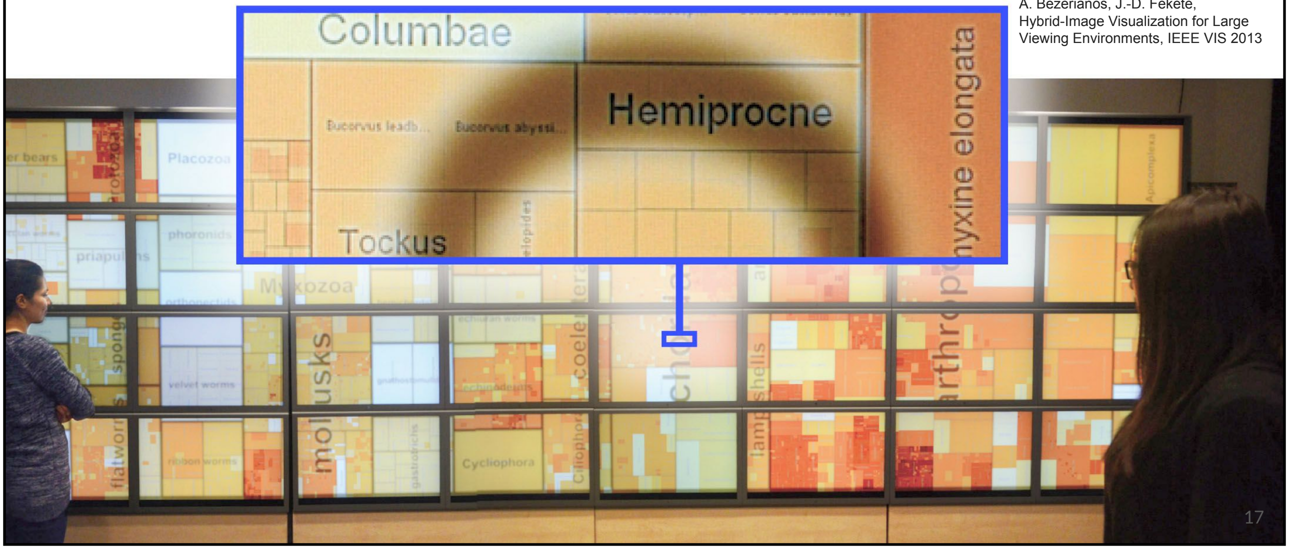




CRC Press
A K Peters Visualization Series
Visualization Analysis & Design
Tamara Munzner

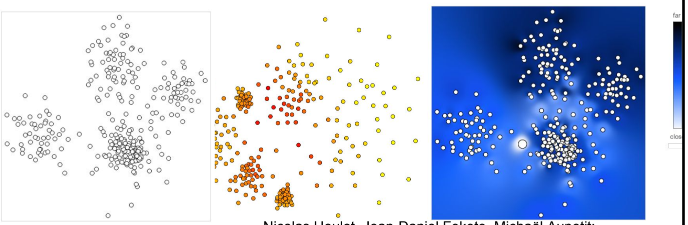
New surfaces for collaborative visualization

P. Isenberg, P. Dragicevic, W. Willett,
A. Bezerianos, J.-D. Fekete,
Hybrid-Image Visualization for Large
Viewing Environments, IEEE VIS 2013

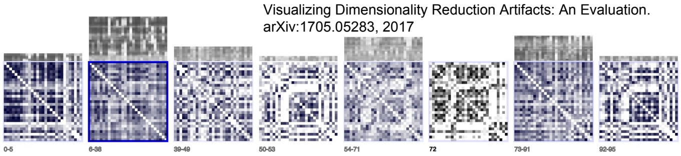


Beyond standard charts: deeper analyses

- When data gets bigger or more complex, visualization relies on computations
- These computation produce artifacts
  - **always**
- Practitioners need to
  - be aware of them
  - have methods to control them



Nicolas Heulot, Jean-Daniel Fekete, Michaël Aupetit:
Visualizing Dimensionality Reduction Artifacts: An Evaluation.
arXiv:1705.05283, 2017

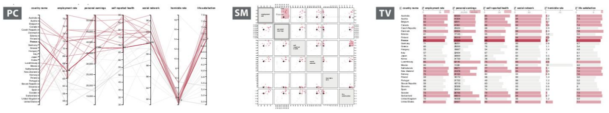


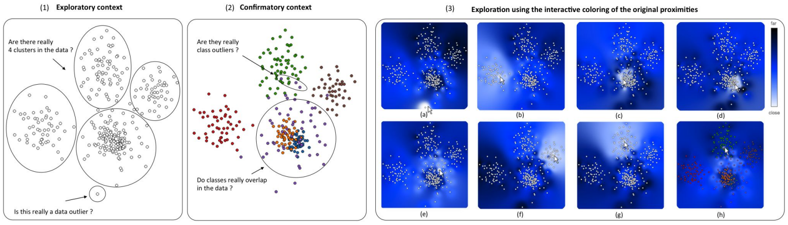
Bach, B., Henry-Riche, N., Dwyer, T., Madhyastha, T., Fekete, J.-D. and Grabowski, T. (2015), Small MultiPiles: Piling Time to Explore Temporal Patterns in Dynamic Networks. Computer Graphics Forum, 34: 31–40. http://aviz.fr/~bbach/multipiles/

Beyond simple charts: controlling artifacts

Visualizing high-dimensional Data

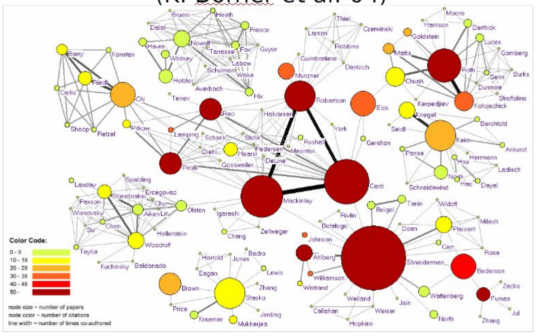
- Up to 10-30 dimensions, some simple visualization techniques work
    - Scatterplot Matrix
    - Parallel Coordinates
    - Data Matrix
- Above that
    - Multidimensions Projections
    - Data Matrix to some extent



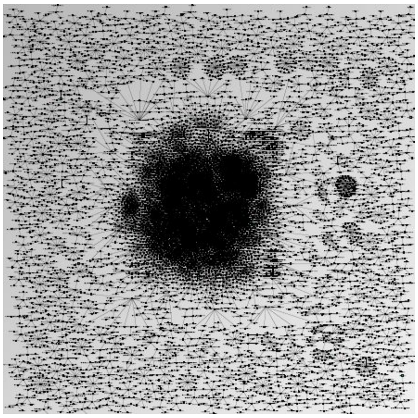


Deeper Analysis of Networks

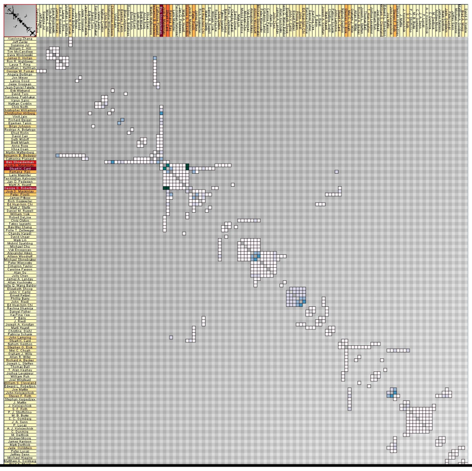
InfoVis Co-authoring
(K. Börner et al. 04)

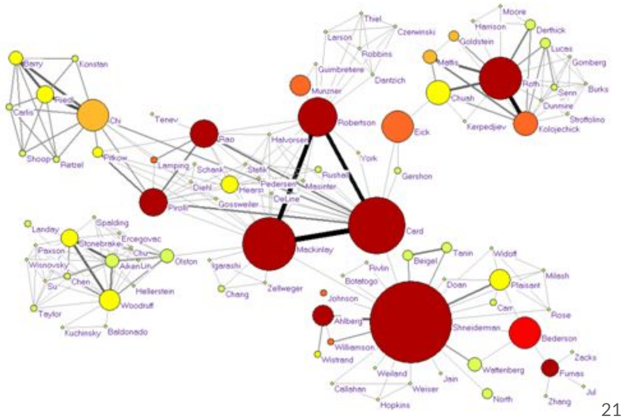

Color Code:
0 - 9
10 - 19
20 - 29
30 - 39
40 - 49
50 -
node size ~ number of papers
node color ~ number of citations
line width ~ number of times co-authored

1

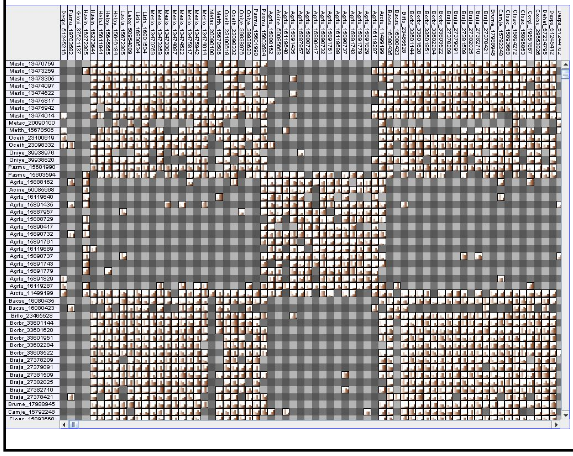


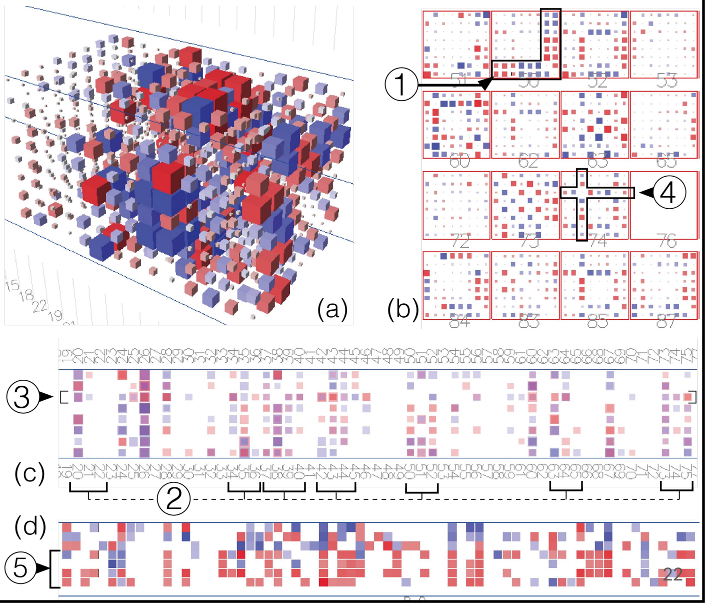
Adjacency matrix better at exploring





Analysis of complex networks





Visualization controversies

- Explorations through interactive visualizations is not replicable
  - Is it a problem? Is Einstein replicable?
- Visualization is done by humans who sometimes see patterns in noise
  - Type I error
- Visualization is done by humans who can miss patterns
  - Type II error

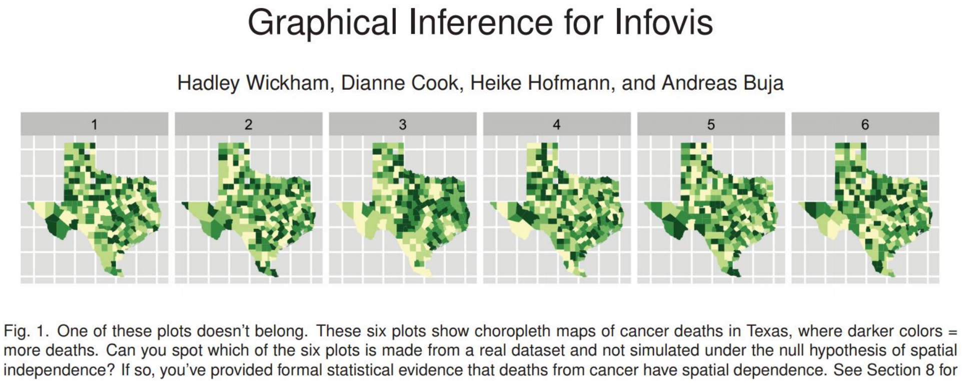
The hypothetico-deductive method

"Scientific inquiry proceeds by formulating a hypothesis in a form that could conceivably be falsified by a test on observable data."

Visualizations generate observations and hypotheses.

- Is the visualization sufficient as a test?
  - ... it depends
- Can more sophisticated tests improve the situation?
  - ... it depends

Visual null hypothesis testing



Graphical Inference for Infovis

Hadley Wickham, Dianne Cook, Heike Hofmann, and Andreas Buja

Fig. 1. One of these plots doesn't belong. These six plots show choropleth maps of cancer deaths in Texas, where darker colors = more deaths. Can you spot which of the six plots is made from a real dataset and not simulated under the null hypothesis of spatial independence? If so, you've provided formal statistical evidence that deaths from cancer have spatial dependence. See Section 8 for

Hadley Wickham, Dianne Cook, Heike Hofmann, and Andreas Buja. 2010. Graphical inference for infovis. IEEE Transactions on Visualization and Computer Graphics 16, 6 (November 2010), 973-979.

Visual null hypothesis testing

"The basic protocol of the line up is simple:

- Generate n−1 decoys (null data sets).
- Make plots of the decoys, and randomly position a plot of the true data.
- Show to an impartial observer.

Can they spot the real data? In practice, we would typically set n = 19, so that if the accused is innocent, the probability of picking the accused by chance is 1/20 = 0.05, the traditional boundary for statistical significance. Comparing 20 plots is also reasonably feasible for a human observer. "

Well, maybe, but how many times?

The multiple comparisons problem (MCP)

"As more comparisons are made, the probability rapidly increases of encountering interesting-looking (e.g., data trend, unexpected distribution, etc.), but still random events. Treating such inevitable patterns as insights is a false discovery (Type I error) and the analyst `loses' if they act on such false insights." [Zgraggen, Zhao, Zeleznik and Kraska, Investigating the Effect of the Multiple Comparison Problem in Visual Analysis in CHI 2018.]

- Need of a hold-out dataset to confirm the findings made by exploration, or gathering more data
- Also, need to control the biases added in the data through multiple filters and aggregations

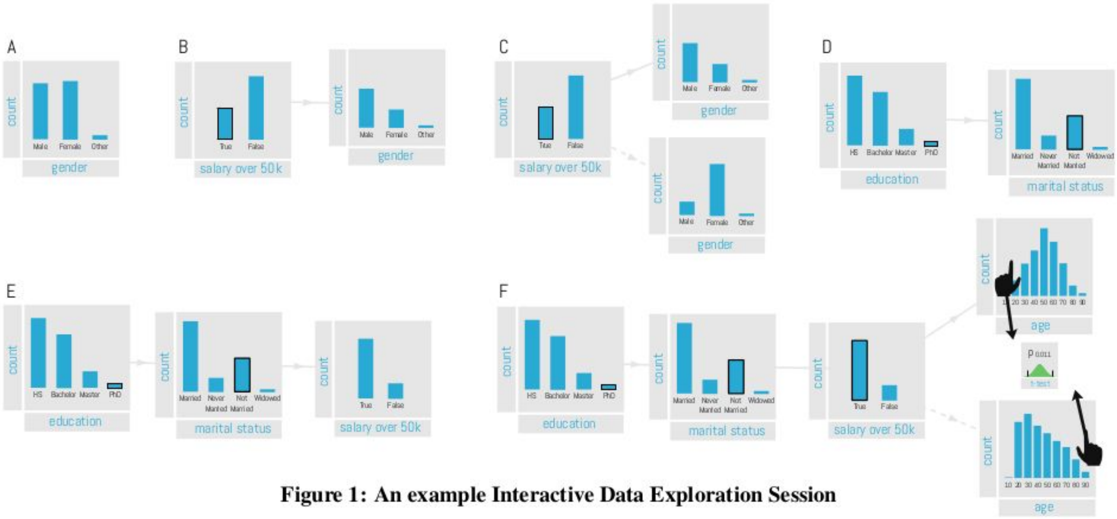


Figure 1: An example Interactive Data Exploration Session

The multiple comparisons problem

- The problem is identical when using any exploratory technique
  - visualization is neither immune, nor worse than others
- A few solutions, but users of exploratory systems should be aware of the problem
- Raging debate: **can expertise overcome spurious discoveries during exploration?**

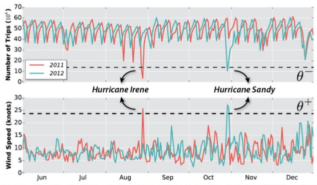


Figure 1: Variation of the number of taxi trips in NYC and its relationship with wind speed.

F. Chirigati, H. Doraiswamy, T. Damoulas, and J. Freire. **Data Polygamy: The Many-Many Relationships among Urban Spatio-Temporal Data Sets**.SIGMOD '16

Z. Zhao, L. De Stefani, E. Zgraggen, C. Binnig, E. Upfal, and T. Kraska. **Controlling False Discoveries During Interactive Data Exploration**.SIGMOD '17

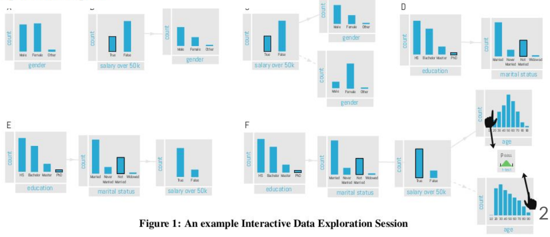


Figure 1: An example Interactive Data Exploration Session

Take away messages

- Visualization is an effective way to generate hypotheses, and sometimes validate them
- It comes with its own artifacts
  - Need to develop sufficient visualization literacy to overcome them
- It allows multiple comparisons to be made quickly
  - Need to control for the multiple comparison problem
- To test that patterns are not spurious, either
  - more data is available for uncorrelated tests
  - more powerful tests exist
  - human expertise can validate the patterns
- Explain the hypothesis discovery pipeline, not only the validation of the hypothesis
  - To help your colleagues find more hypotheses