

Why do we need a visualization system?

Yu-Shuen Wang, CS, NCTU

Definition

- Computer-based **visualization** systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Tamara Munzner

- Why?
- Trade-off

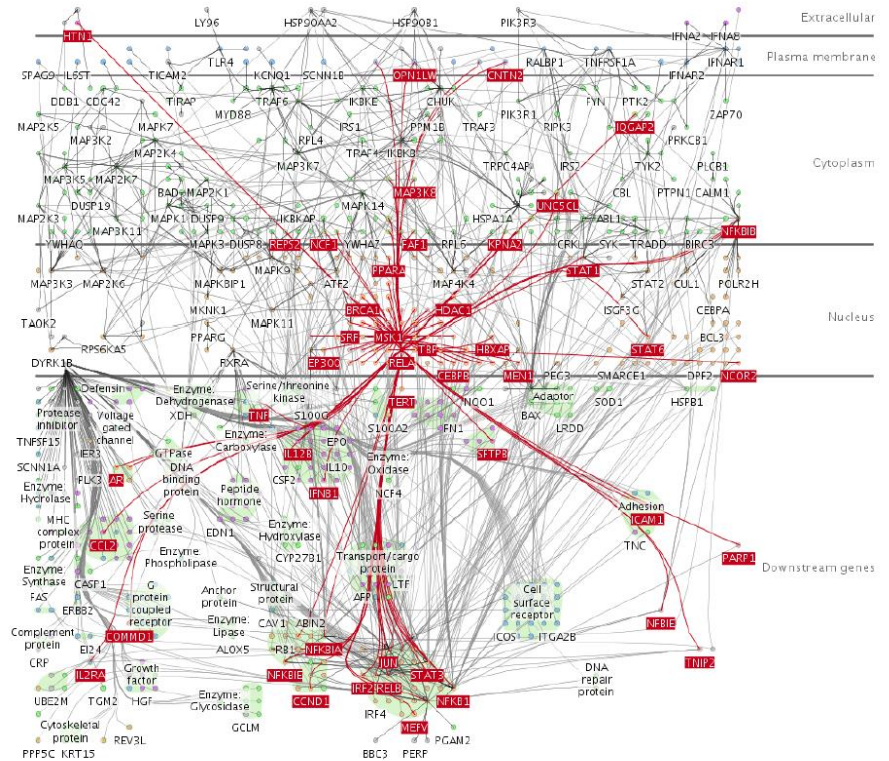
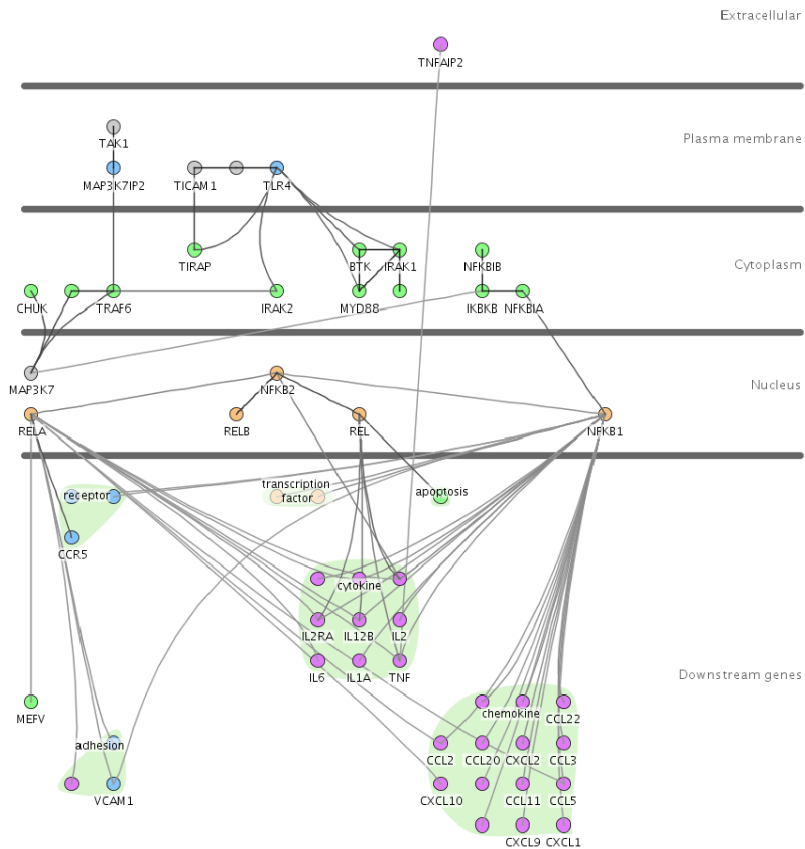
Why have a human in the loop?

Computer-based visualization systems provide visual representations of **datasets** designed to help **people** carry out tasks more effectively.

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- don't need vis when fully automatic solution exists and is trusted
- many analysis problems ill-specified
 - don't know exactly what questions to ask in advance
- possibilities
 - long-term use for end users (e.g. exploratory analysis of scientific data)
 - presentation of known results
 - stepping stone to better understanding of requirements before developing models
 - help developers of automatic solution refine/debug, determine parameters
 - help end users of automatic solutions verify, build trust

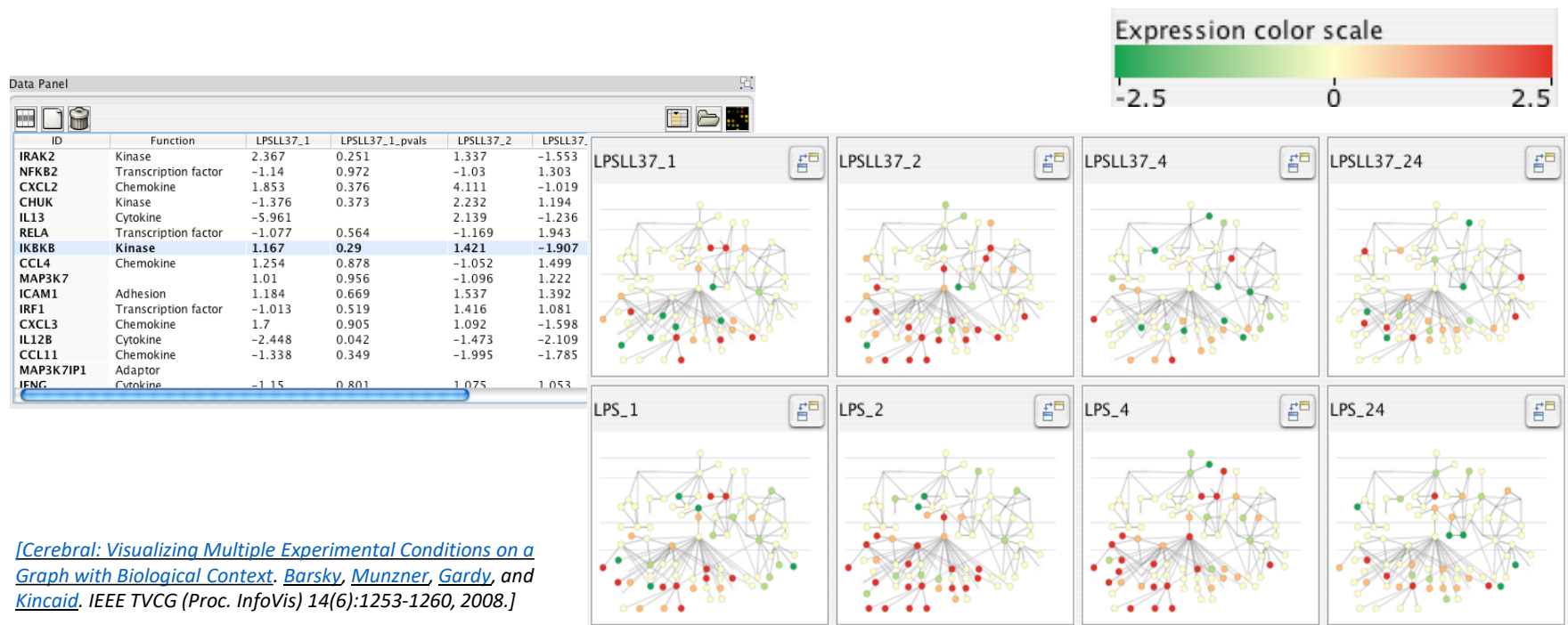
Why have a Computer in the loop?



Why use an external representation?

Computer-based visualization systems provide **visual representations** of datasets designed to help people carry out tasks more effectively.

- external representation: replace cognition with perception



[\[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Incaid. IEEE TVCG \(Proc. InfoVis\) 14\(6\):1253-1260, 2008.\]](#)

Why depend on vision?

Computer-based visualization systems provide **visual** representations of datasets designed to help people carry out tasks more effectively.

- human visual system is high-bandwidth channel to brain
 - overview possible due to background processing
 - subjective experience of seeing everything simultaneously
 - significant processing occurs in parallel and pre-attentively
- sound: lower bandwidth and different semantics
 - overview not supported
 - subjective experience of sequential stream
- touch/haptics: impoverished record/replay capacity
 - only very low-bandwidth communication thus far
- taste, smell: no viable record/replay devices

Why represent all the data?

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	

Why represent all the data?

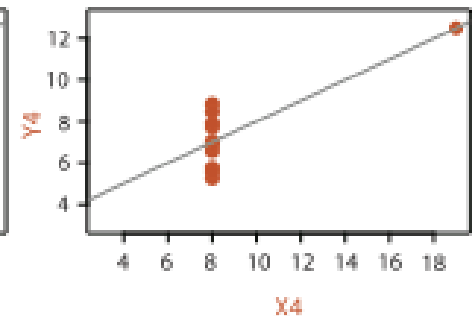
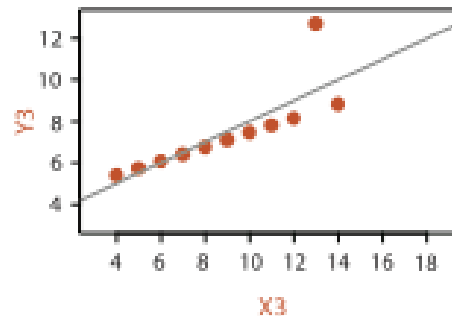
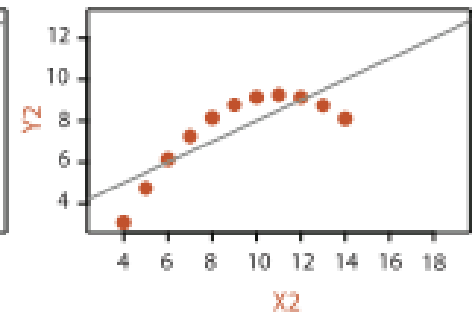
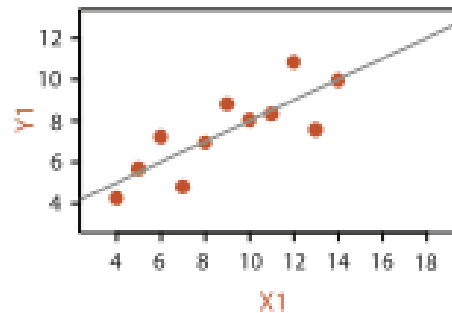
Computer-based visualization systems provide visual **representations of datasets** designed to help people carry out tasks more effectively.

- summaries lose information, details matter
 - confirm expected and find unexpected patterns
 - assess validity of statistical model

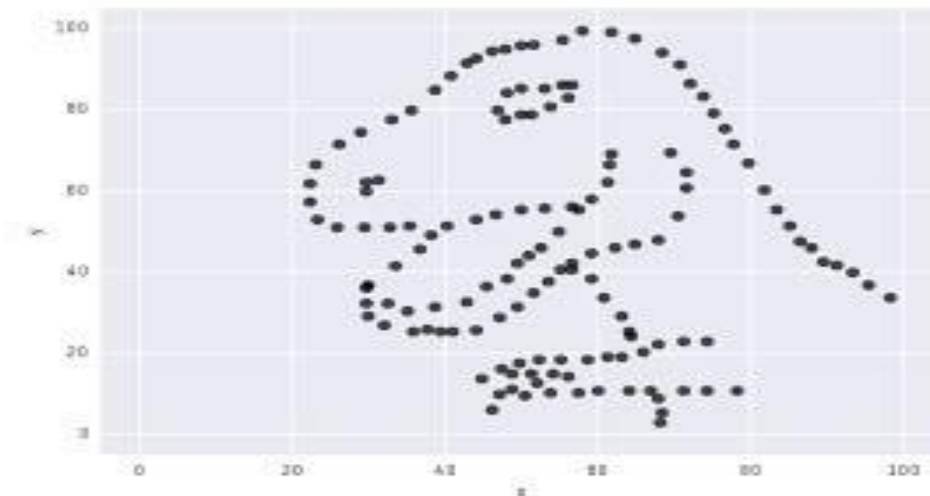
Anscombe's Quartet

Identical statistics

x mean	9
x variance	10
y mean	7.5
y variance	3.75
x/y correlation	0.816



Same Stats, Different Graphs - CHI 2017

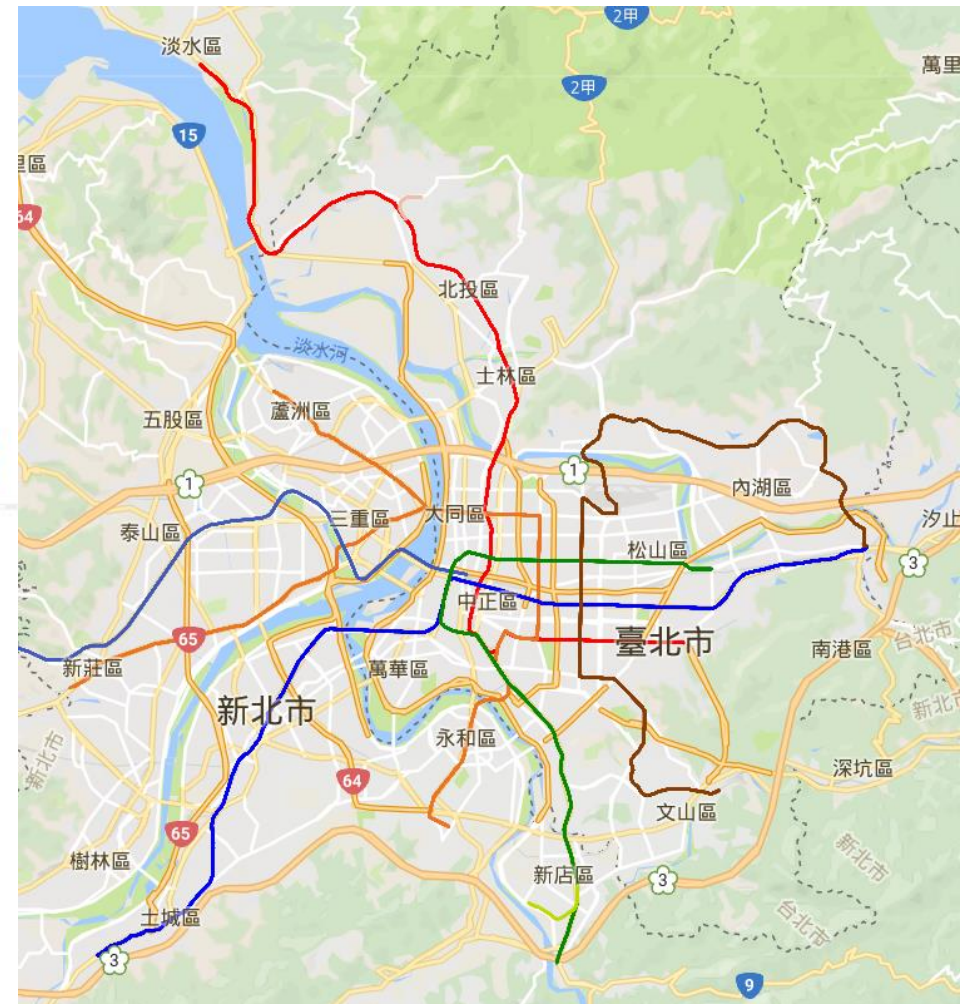


X Mean: 54.2632025
Y Mean: 47.8315781
X SD : 16.7650109
Y SD : 26.9353144
Corr. : -0.0645195

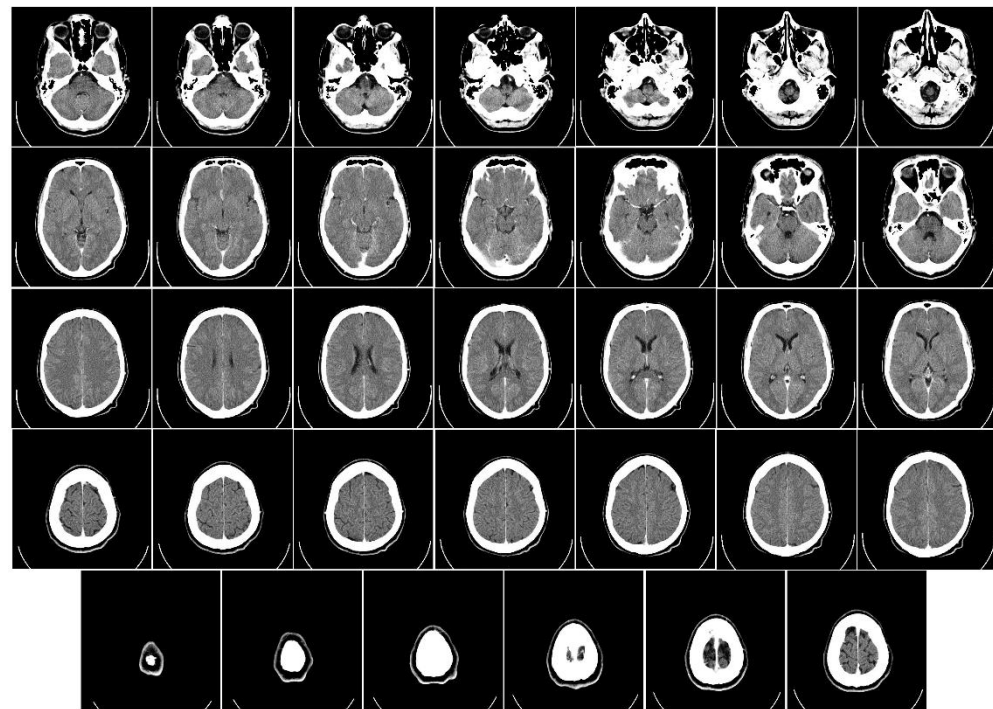
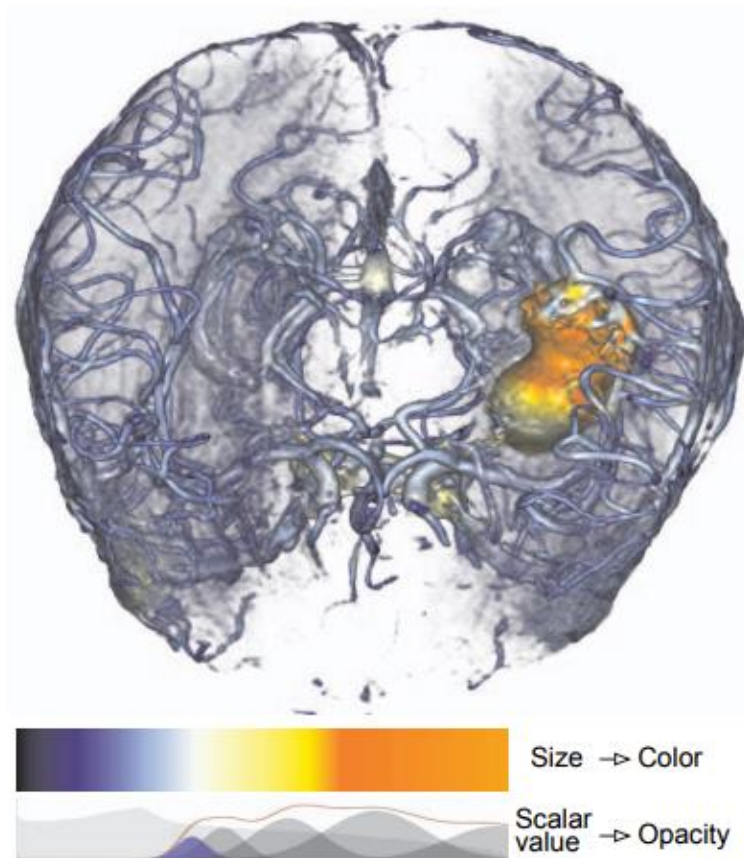
Why focus on tasks and effectiveness?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out **tasks** more **effectively**.

- tasks serve as constraint on design (as does data)
 - idioms do not serve all tasks equally!
 - challenge: recast tasks from domain-specific vocabulary to abstract forms
- most possibilities ineffective
 - validation is necessary, but tricky
 - 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



Which vis design is effective?



https://en.wikipedia.org/wiki/CT_scan

C. Correa and K. L. Ma, "Size-based Transfer Functions: A New Volume Exploration Technique," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1380-1387, 2008.

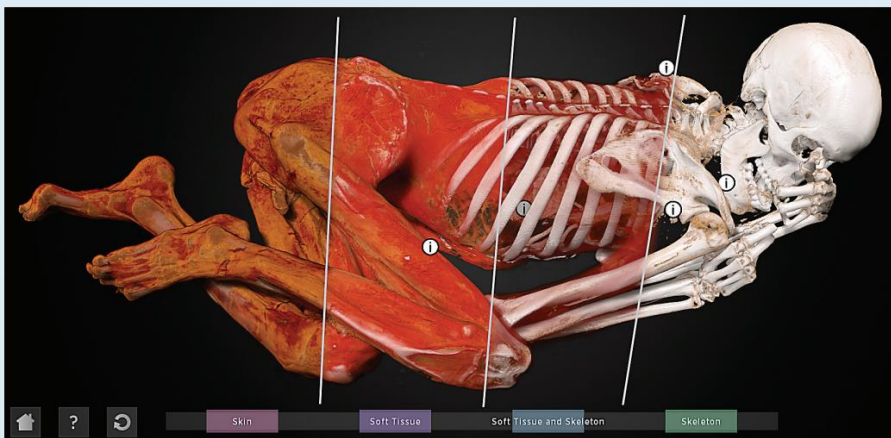
Point of Interest

Rotate

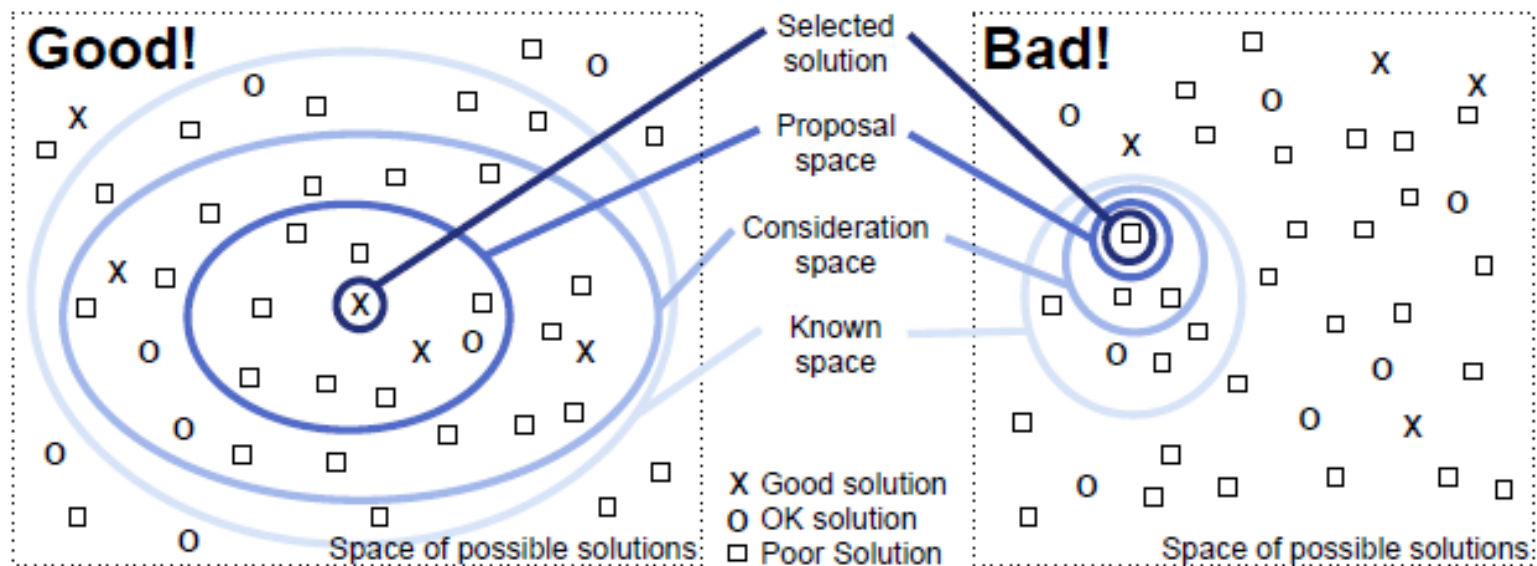
Zoom

Clipping

Layers

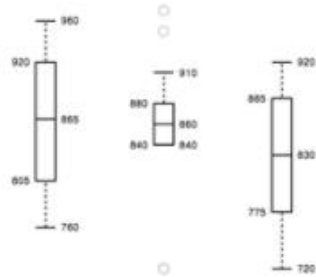


Why are most designs ineffective?



Visual Index

Box Plots



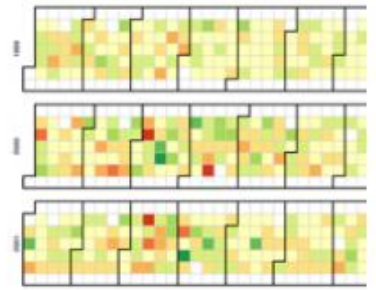
Bubble Chart



Bullet Charts



Calendar View



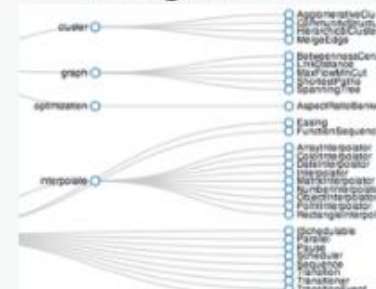
Non-contiguous Cartogram



Chord Diagram



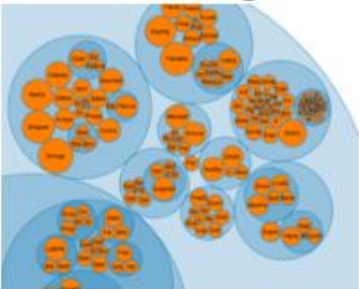
Dendrogram



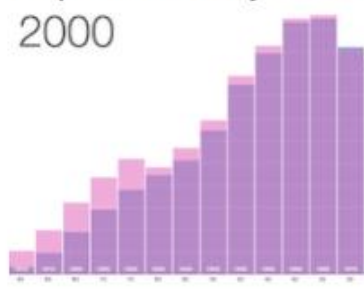
Force-Directed Graph



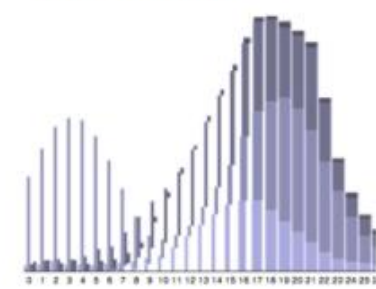
Circle Packing



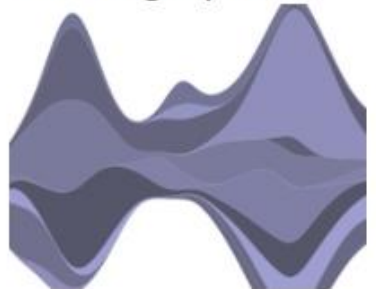
Population Pyramid



Stacked Bars



Streamgraph

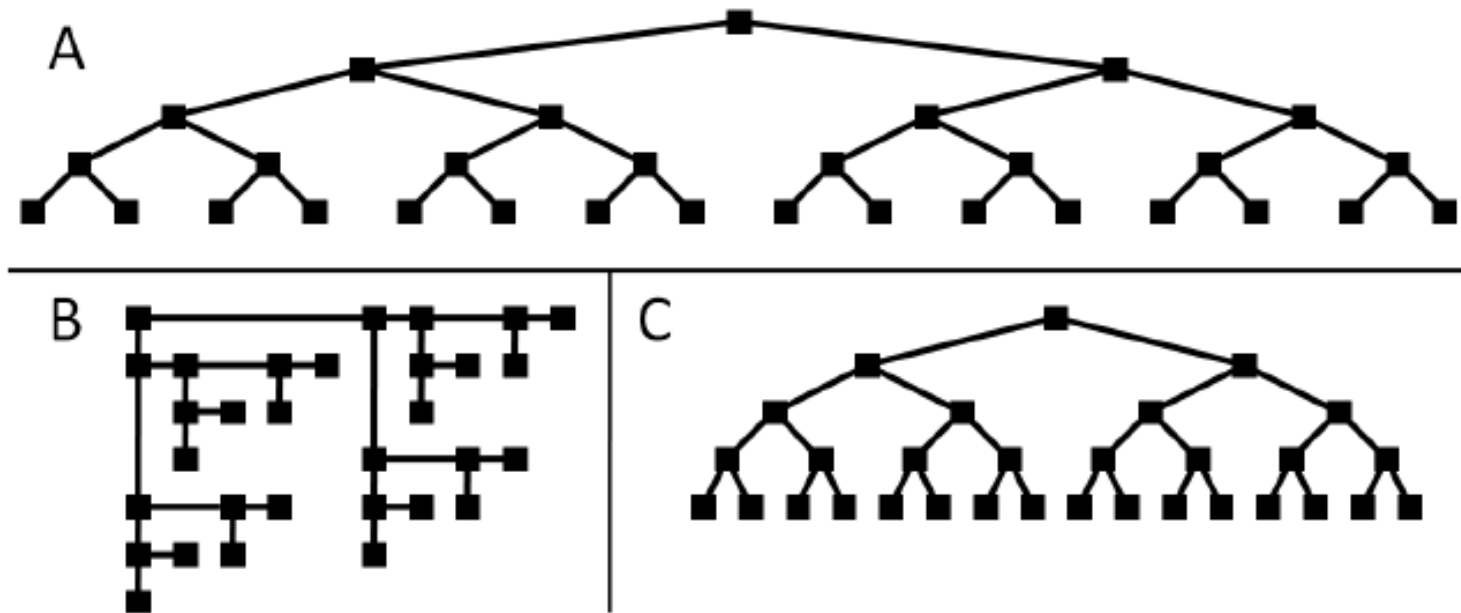


What resource limitations are we faced with?

Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

- Computational limits
 - processing time
 - system memory
- Human limits
 - human attention and memory
- Display limits
 - pixels are precious resource, the most constrained resource
 - information density: ratio of space used to encode info vs unused whitespace
 - tradeoff between clutter and wasting space, find sweet spot between dense and sparse

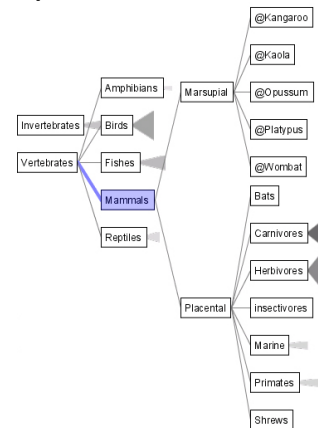
Display limits



Why analyze?

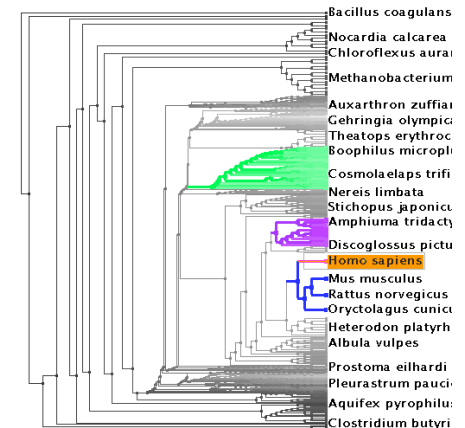
- imposes structure on huge design space
 - scaffold to help you think systematically about choices
 - analyzing existing as stepping stone to designing new
 - most possibilities ineffective for particular task/data combination

SpaceTree



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

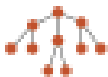
TreeJuxtaposer



[TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context With Guaranteed Visibility. ACM Trans. on Graphics (Proc. SIGGRAPH) 22:453–462, 2003.]

What?

Tree



Why?

Actions



Targets

→ Path between two nodes



How?

SpaceTree



TreeJuxtaposer



What?

Why?

How?

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
 - *Chap 1: What's Vis, and Why Do It?*
- [A Tour through the Visualization Zoo.](#) Communications of the ACM, 53(6), pp. 59-67, Jun 2010