

Visualization without guesswork

Aneesh Karve, Quilt Data

Oracles

examples maxims models perceptual code

APIs

[KS16]

An algebraic process for visualization design
Kindlmann & Scheidegger

[KS2014]

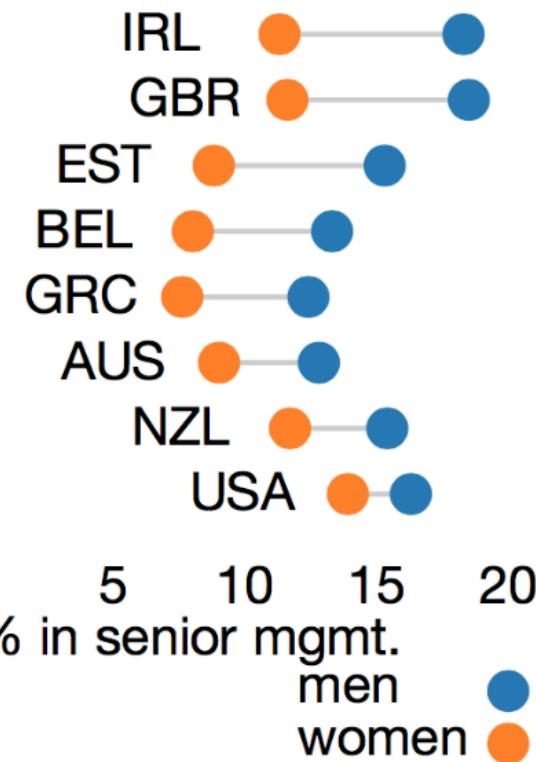
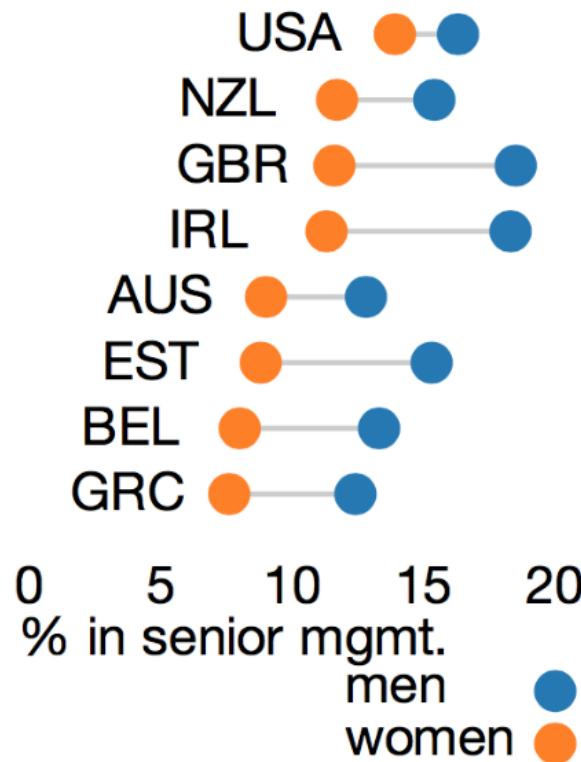
- Invariance
- Unambiguity
- Correspondence

Invariance

The same data
always produces
the same
visualization

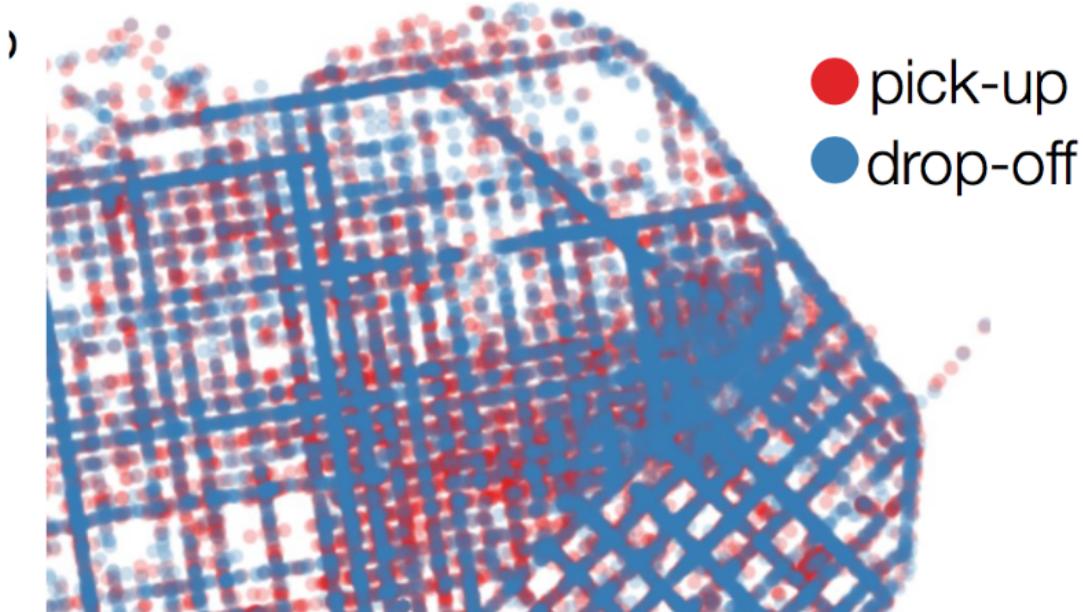
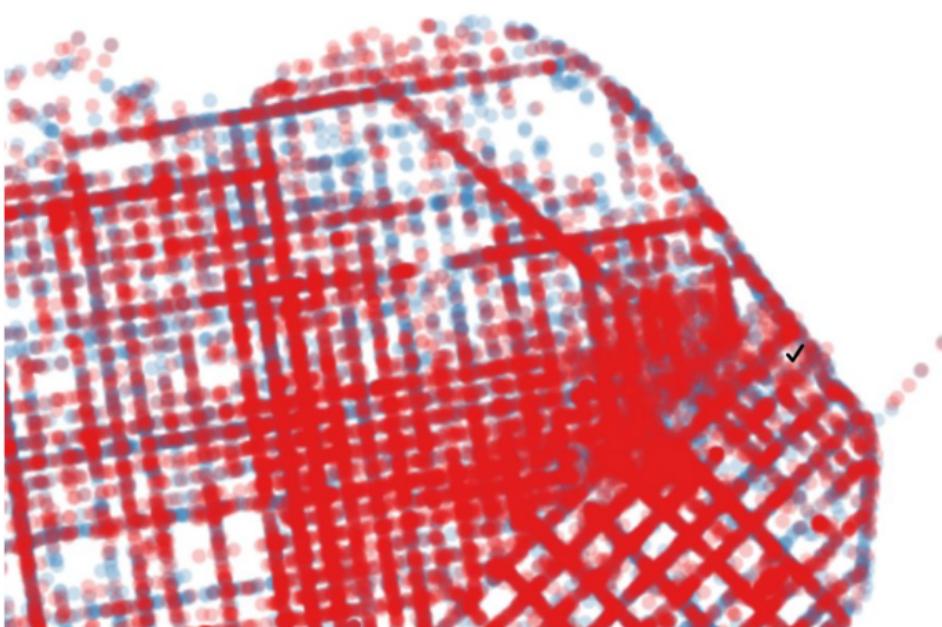
$$D \rightarrow V$$

Same data, different impression



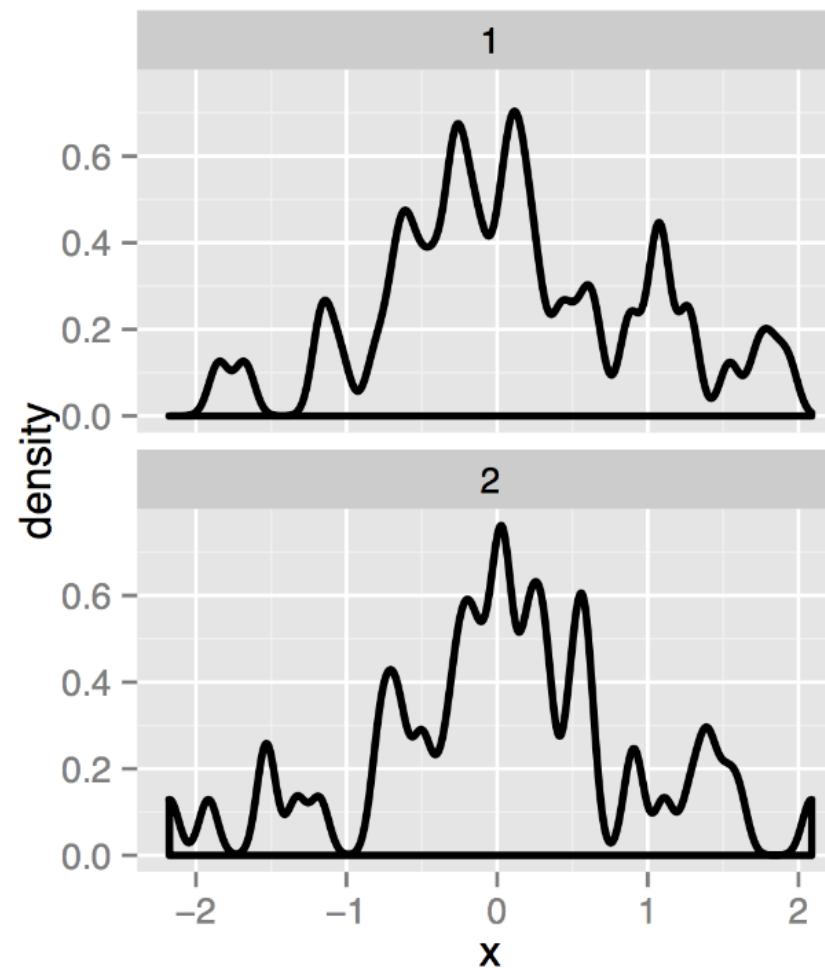
[KS14]

Same data, opposite viz



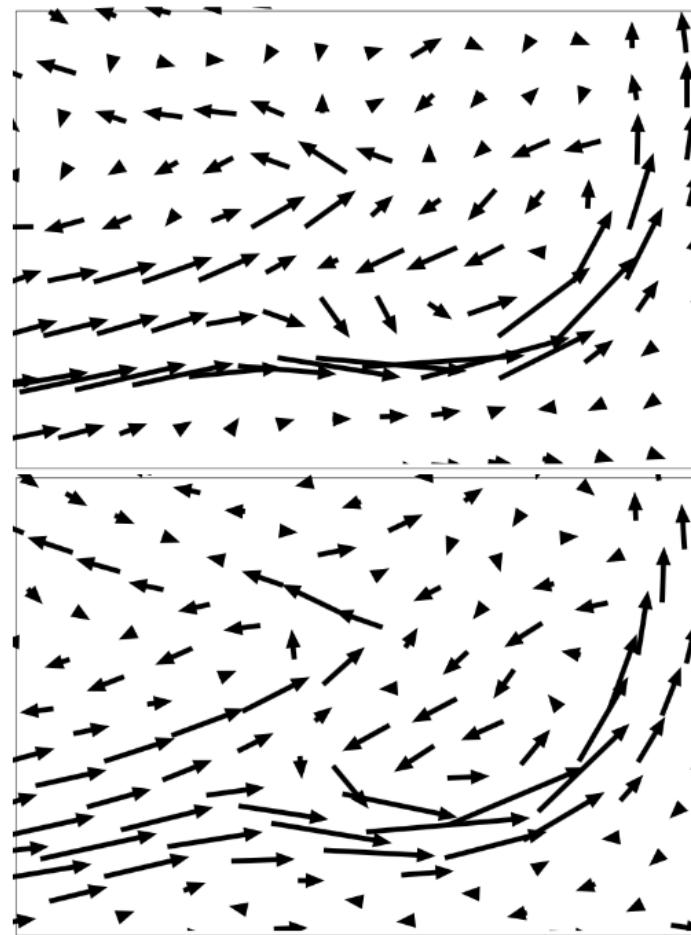
[KS14]

Same population, different samples

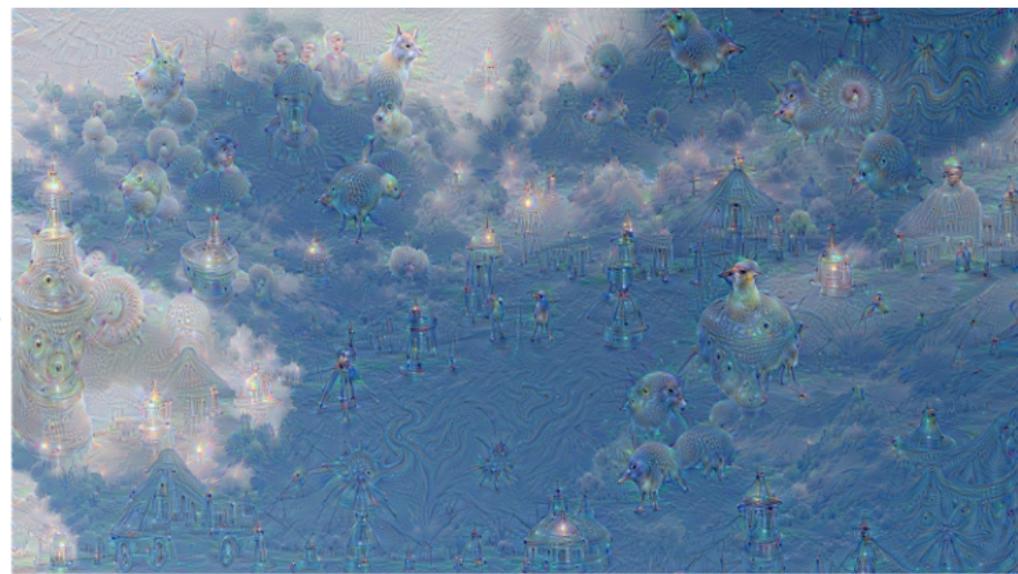
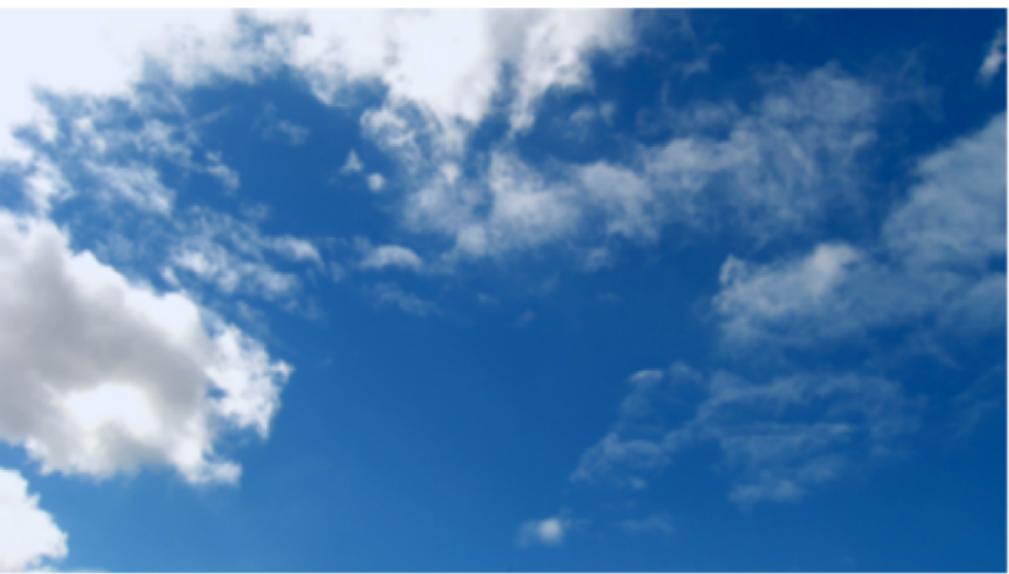


[KS14]

Same population, different samples



[KS14]





"Admiral Dog!"



"The Pig-Snail"



"The Camel-Bird"

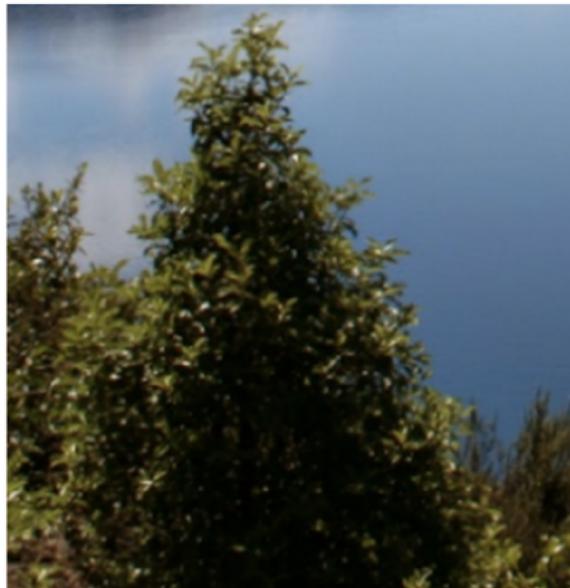


"The Dog-Fish"

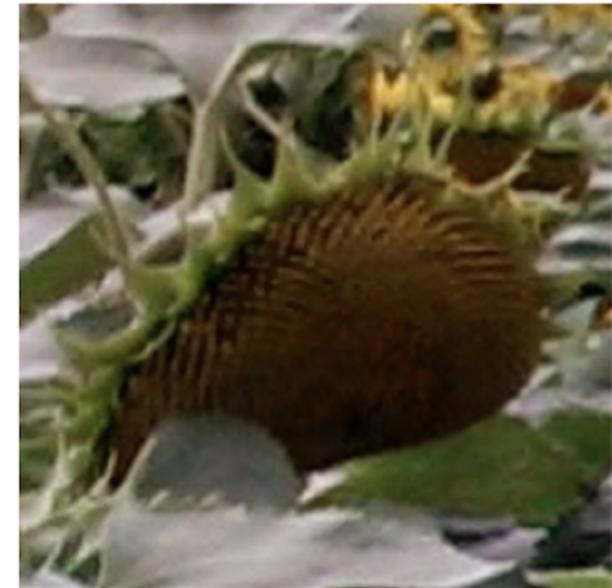




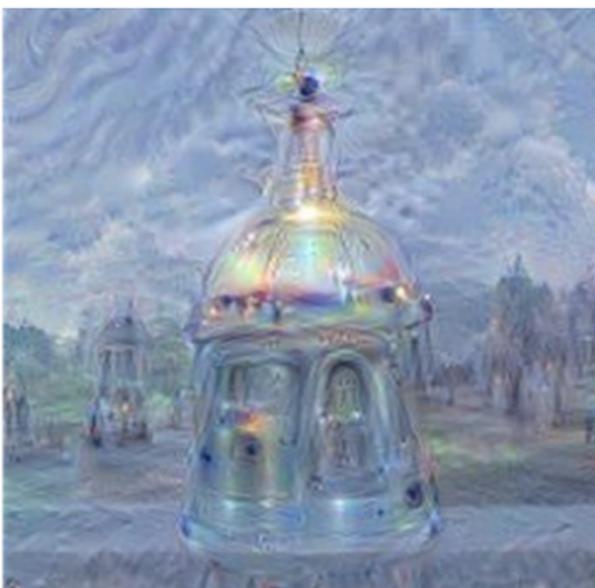
Horizon



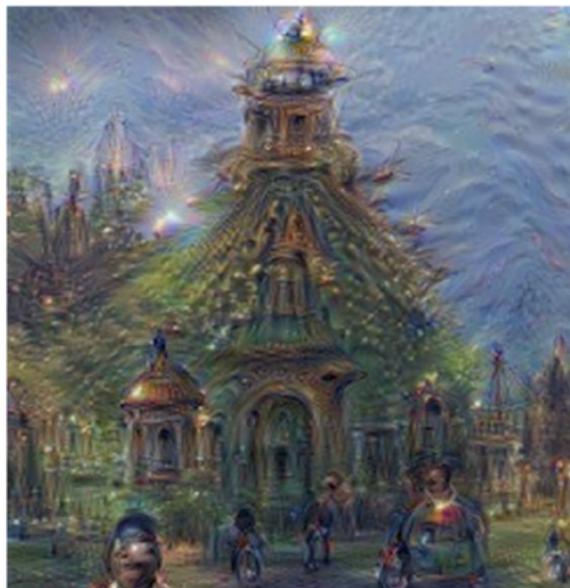
Trees



Leaves



Towers & Pagodas



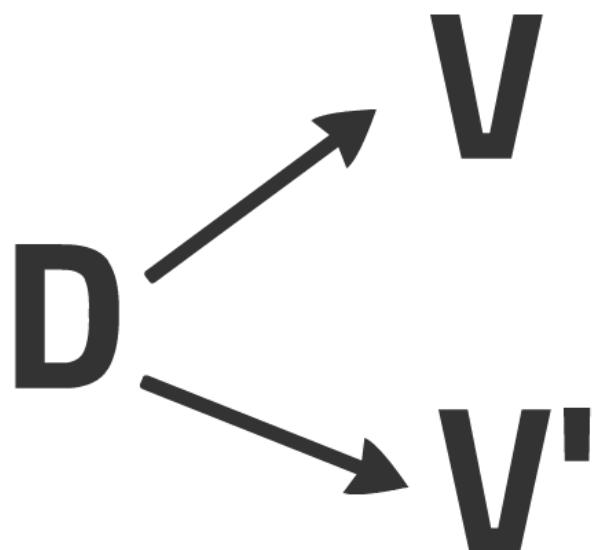
Buildings



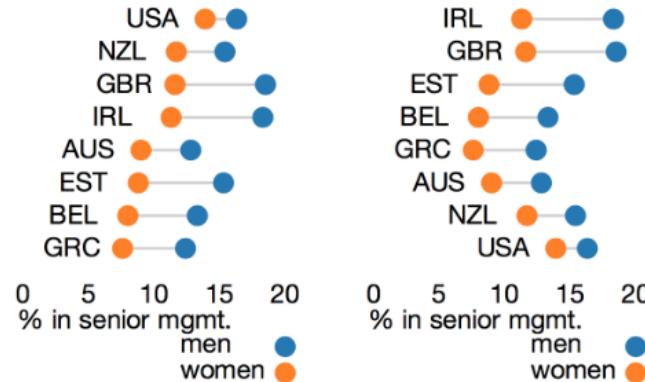
Birds & Insects

Hallucination

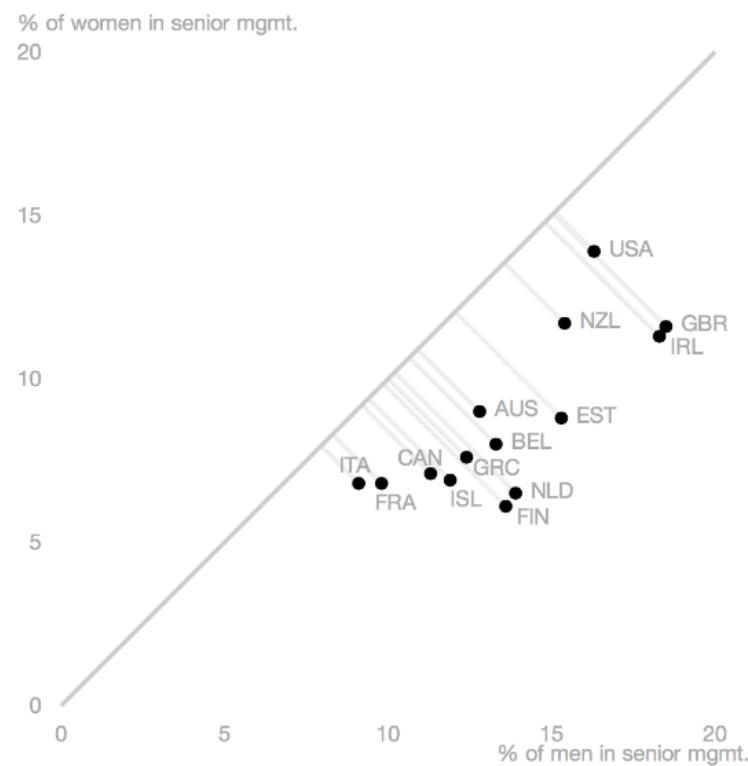
The same data
produces different
visualizations



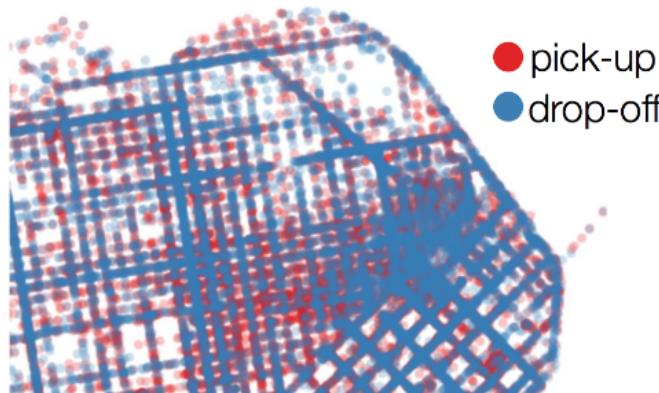
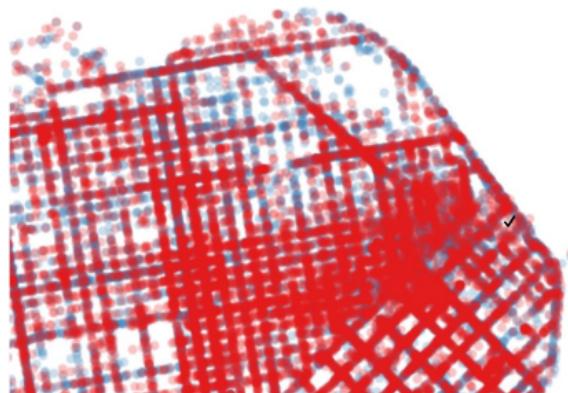
Same data, different impression



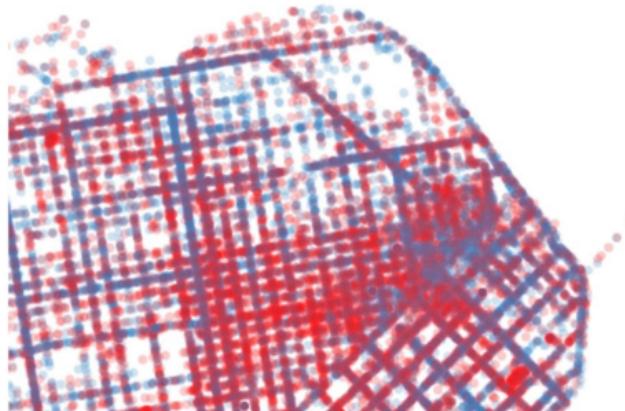
[KS14]



Same data, opposite viz



[KS14]

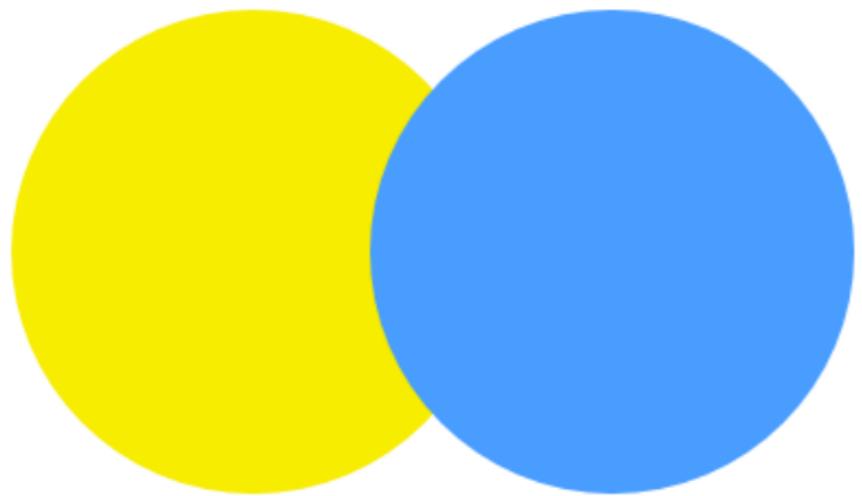
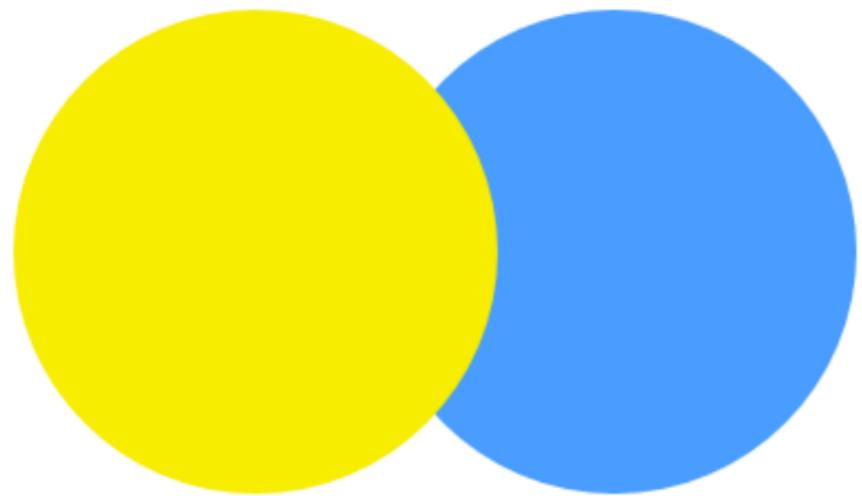


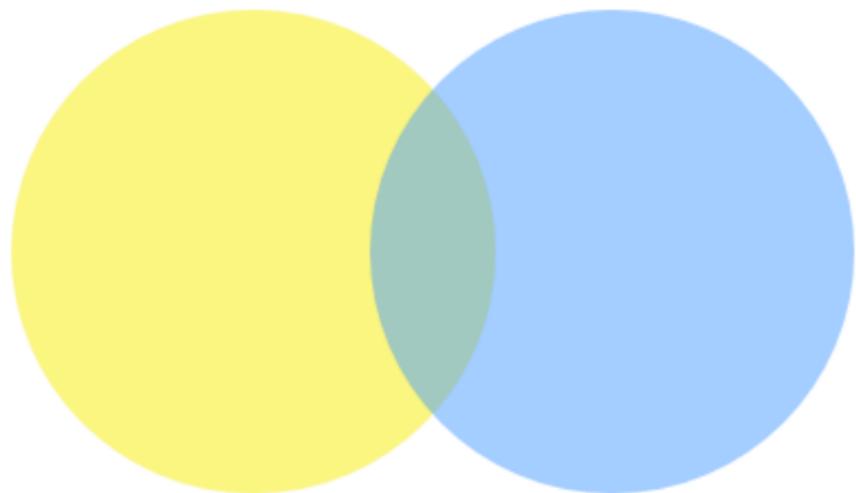
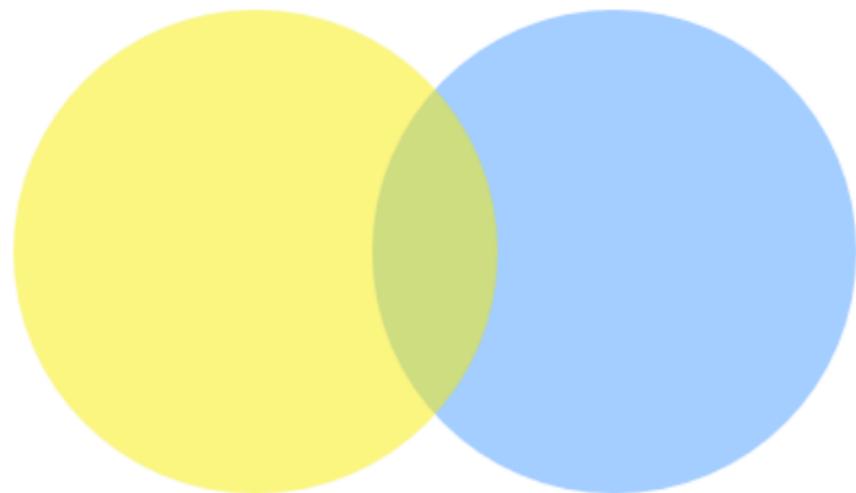
Data semantics differ from vis semantics

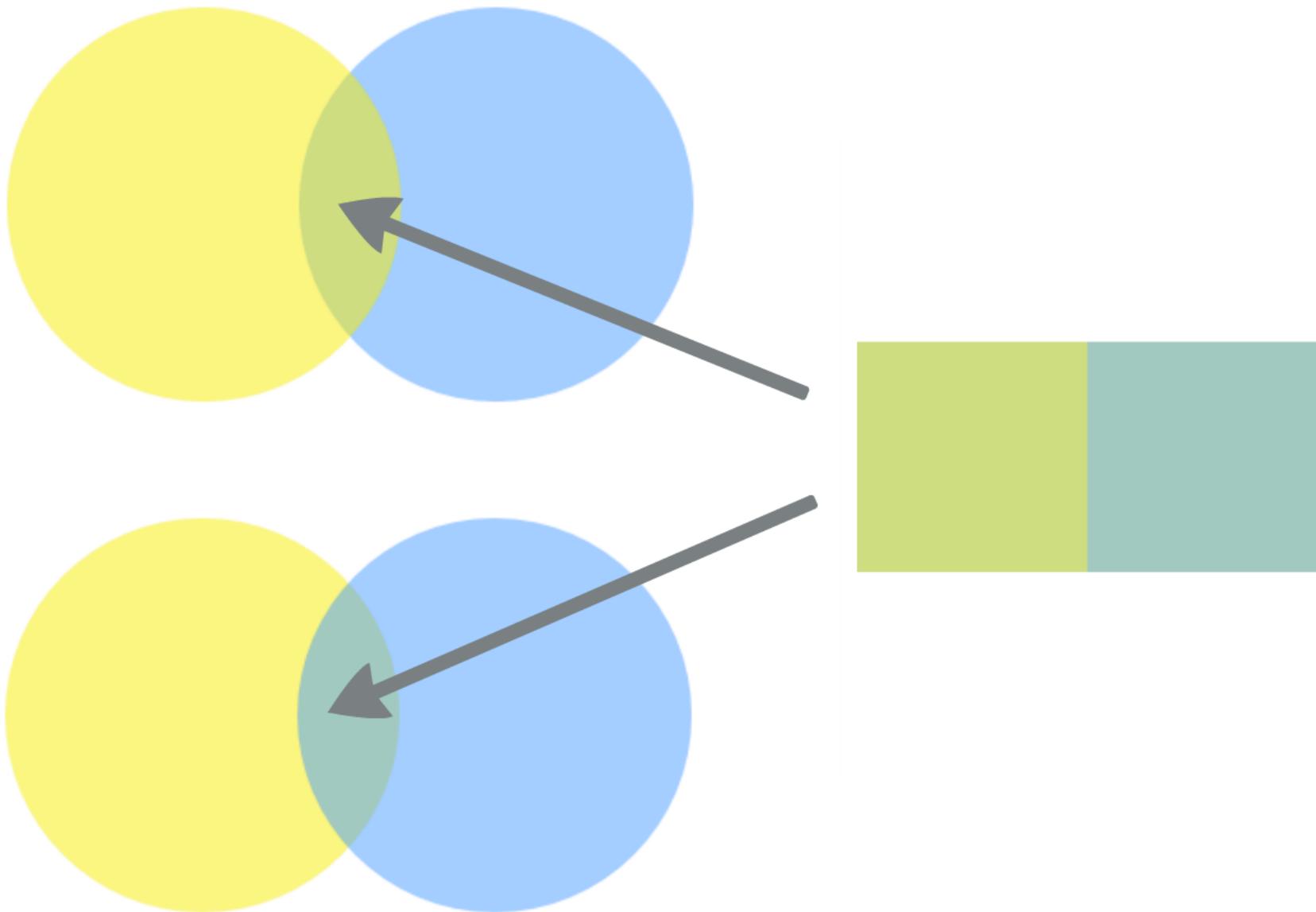
$$A \cap B = B \cap A$$

but

$$A \text{ over } B \neq B \text{ over } A$$

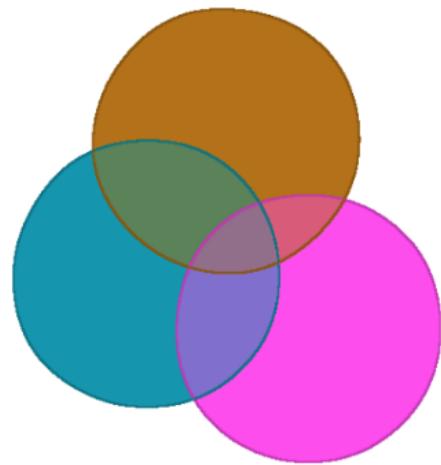




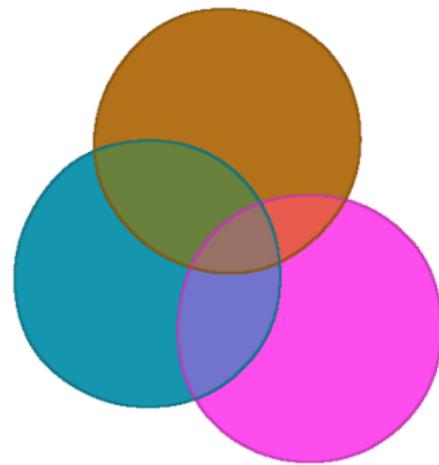


- Commutative blending
- Custom blending
- Sampling

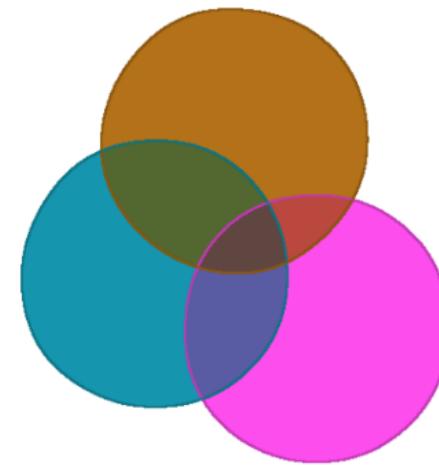
Splatterplots



(a) RGB
blending



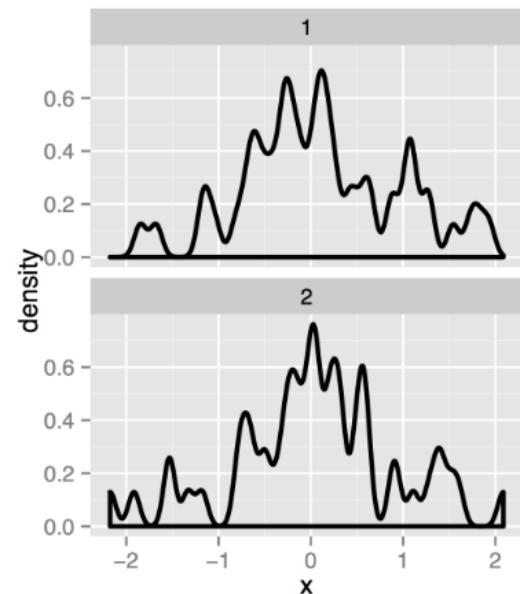
(b) LAB
blending



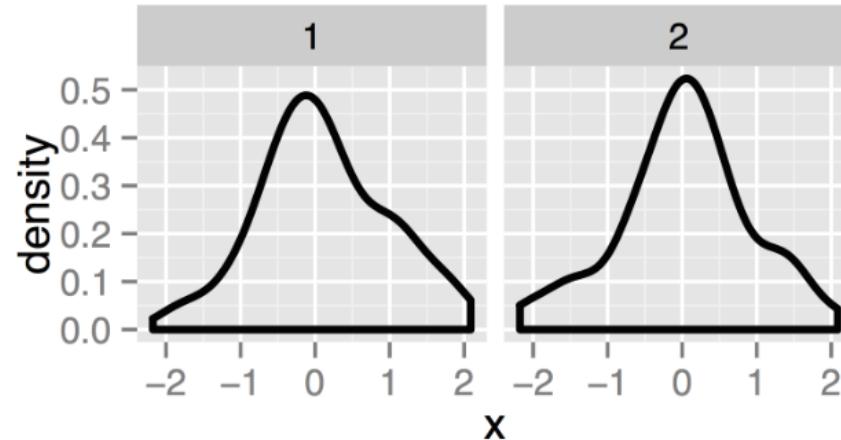
(c) With
attenuation

[MG13]

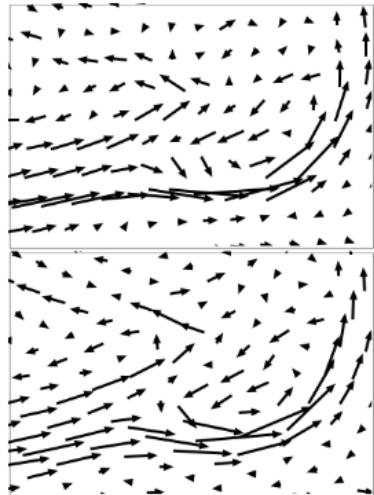
Same population, different samples



[KS14]



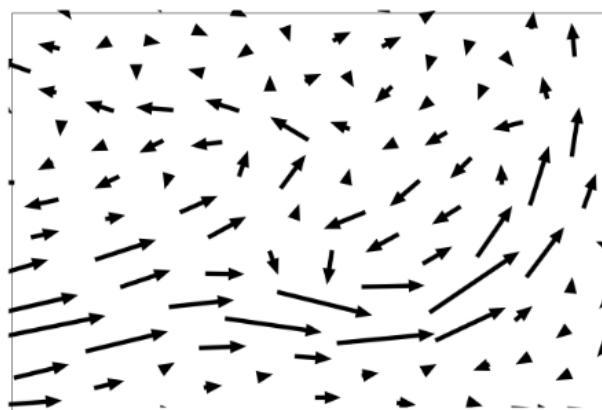
Same population, different samples



[KS14]



"Admiral Dog!"



G. Turk and D. Banks. Image-guided streamline placement.
In Proc. ACM SIGGRAPH, pages 453–460, 1996.

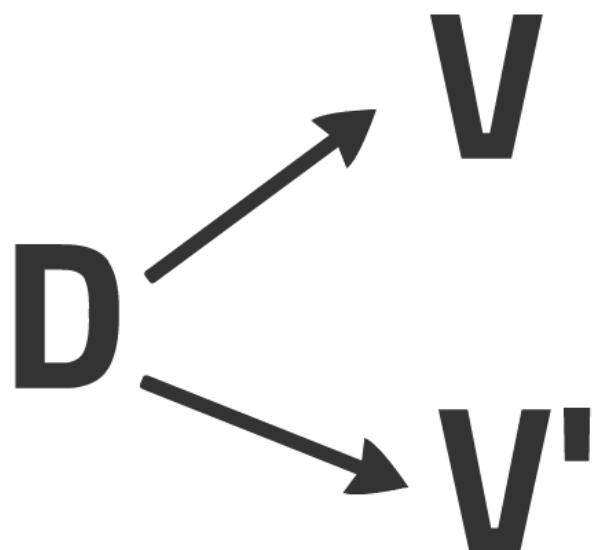


Horiz.



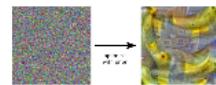
Hallucination

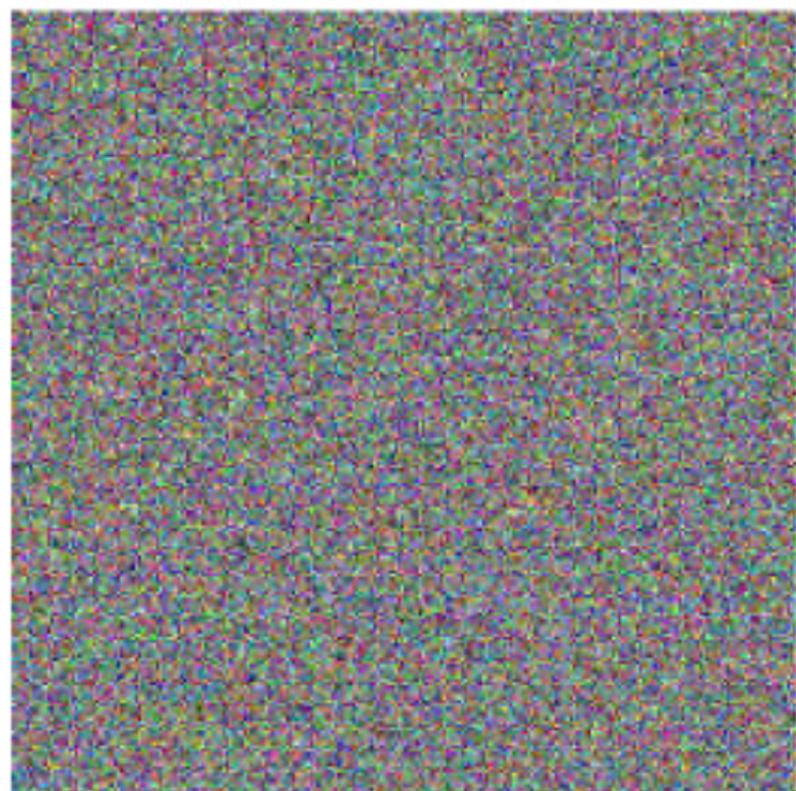
The same data
produces different
visualizations



Finding hallucinators

- Sampling
- Permutation
- Projection
- Inception





optimize
with prior



- Invariance
- Unambiguity
- Correspondence

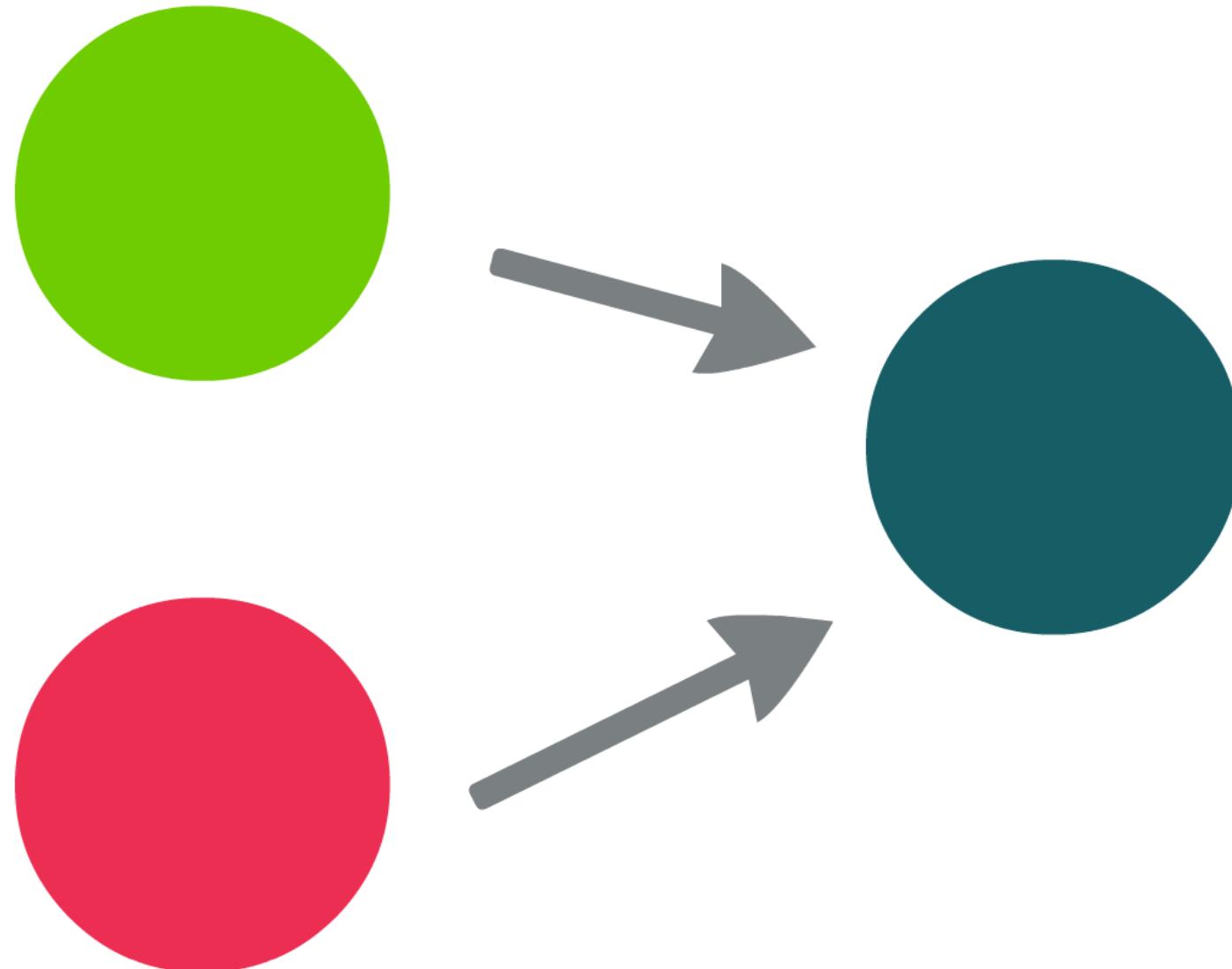
Unambiguity

Different data
always produces
different
visualizations

$$D \rightarrow V$$

$$D' \rightarrow V'$$

Color blindness



Confusion



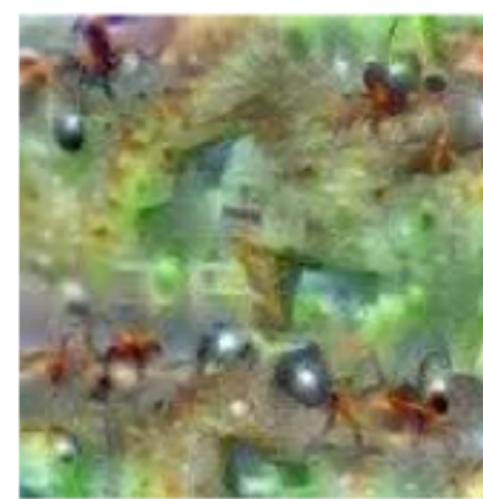
[In15]



Hartebeest



Measuring Cup



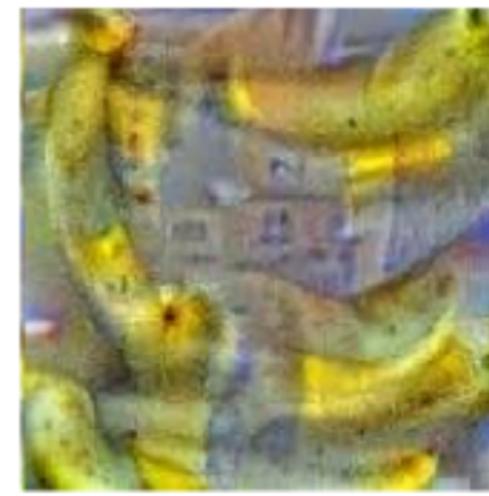
Ant



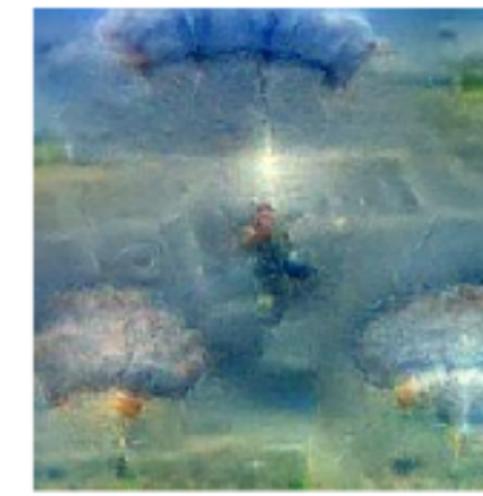
Starfish



Anemone Fish



Banana

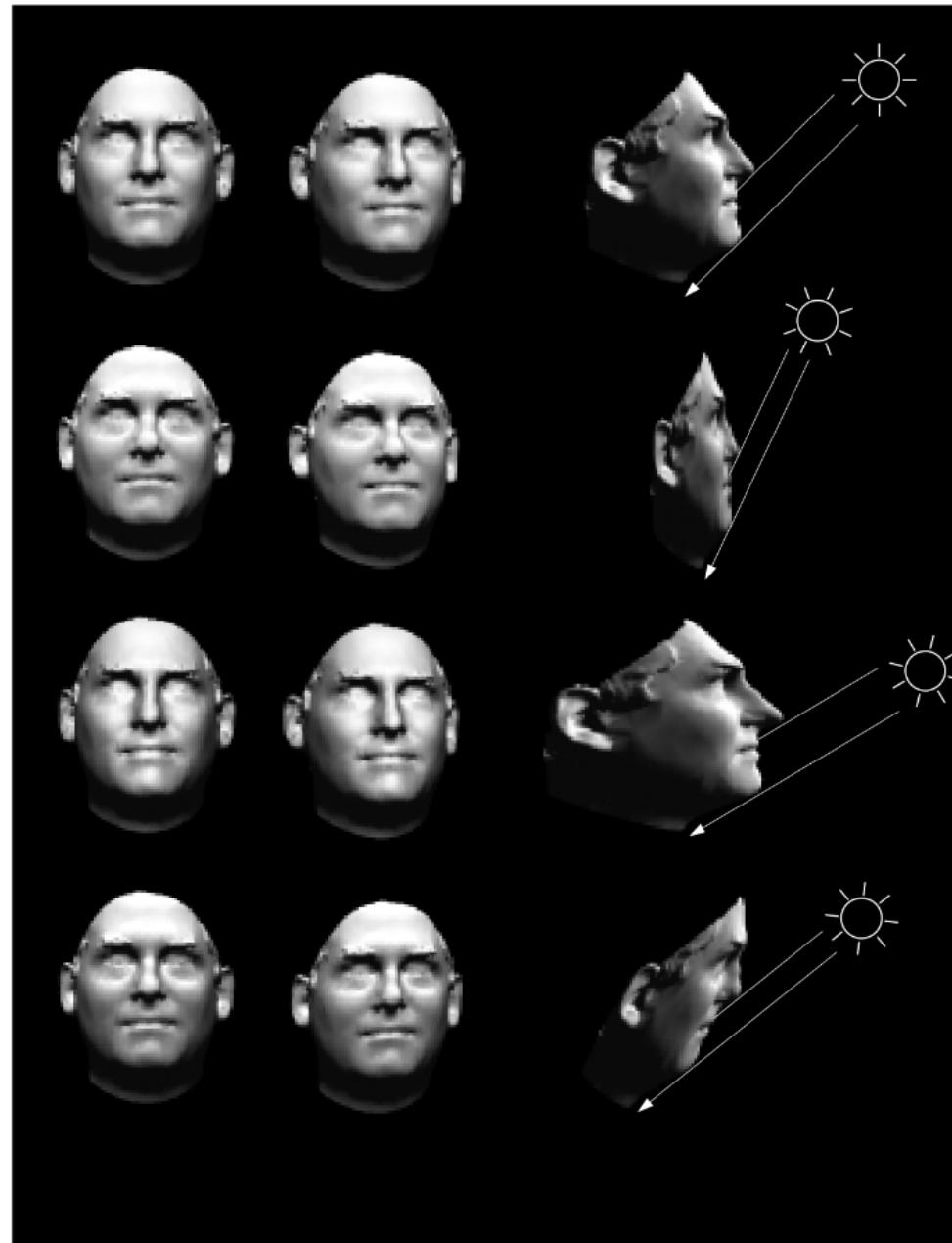


Parachute



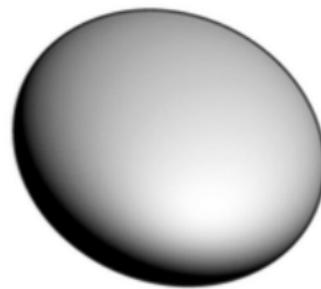
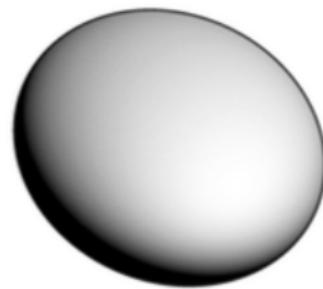
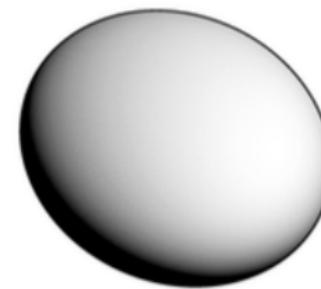
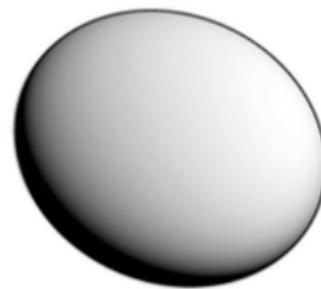
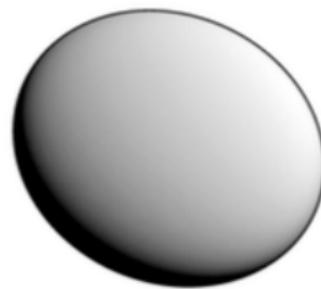
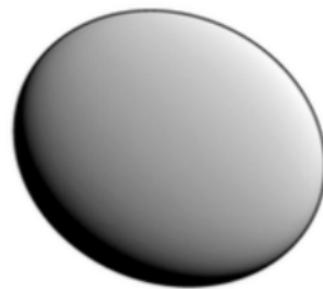
Screw

Bas-relief ambiguity

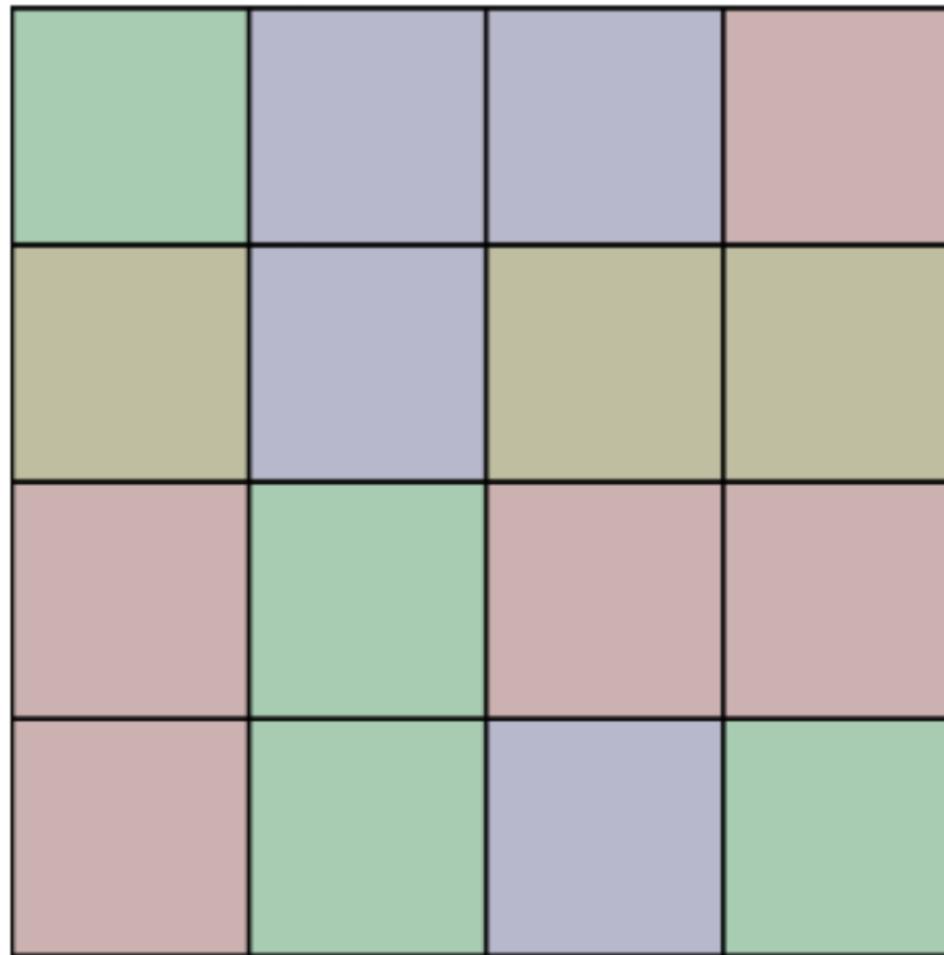


[BKY97]

Very different tensors look similar

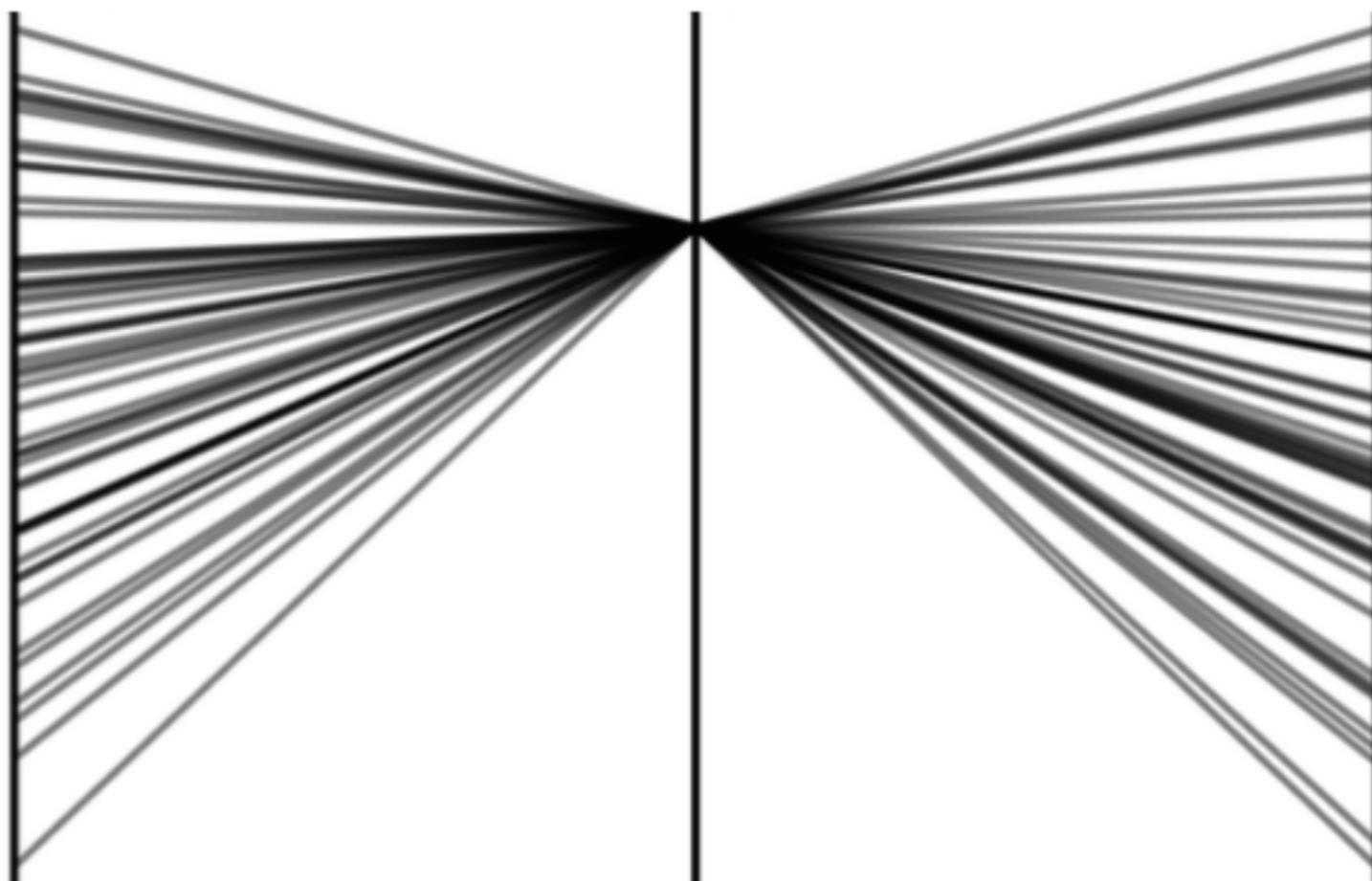


Different trees, same treemap



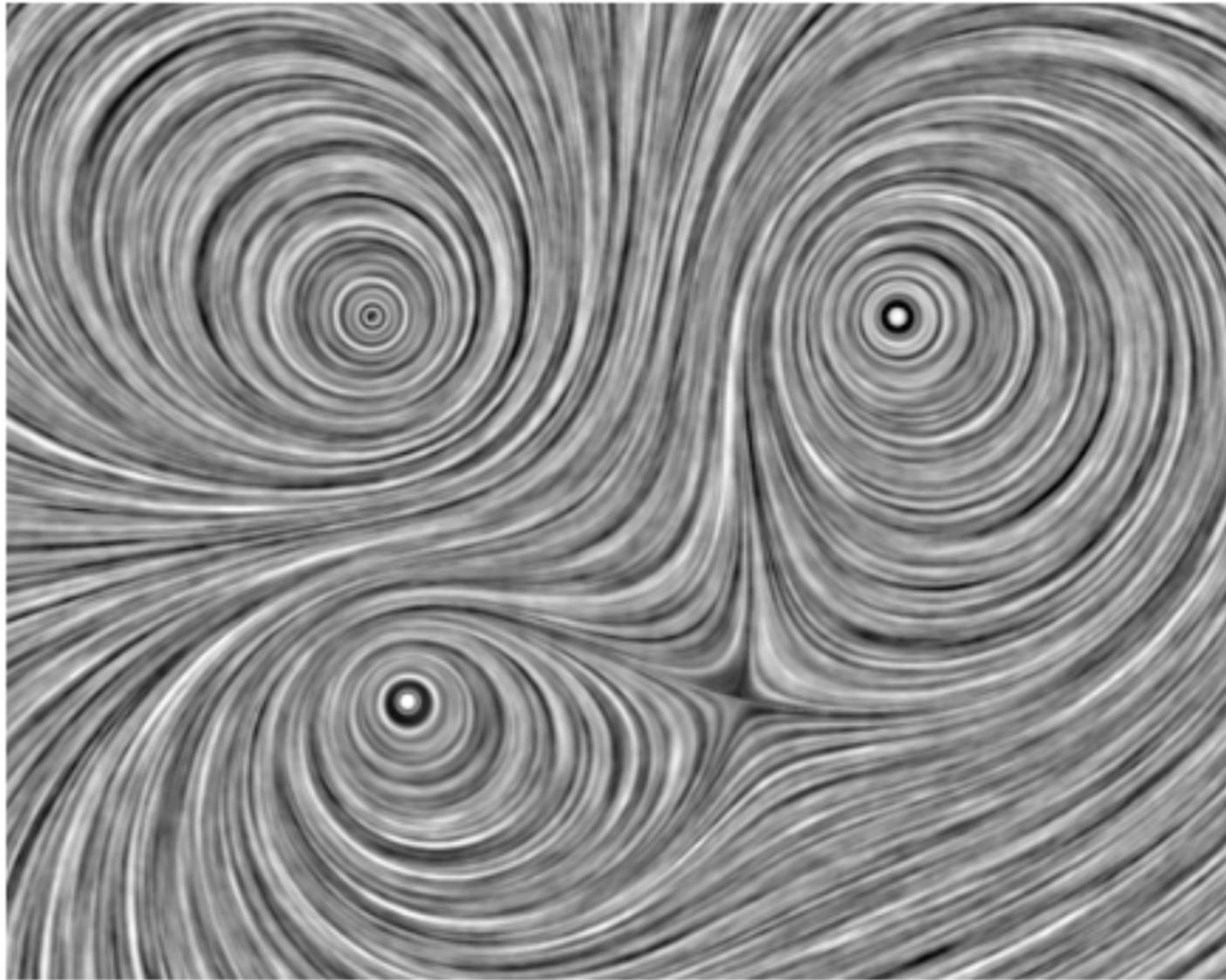
[KS14]

Different data sets, same plot

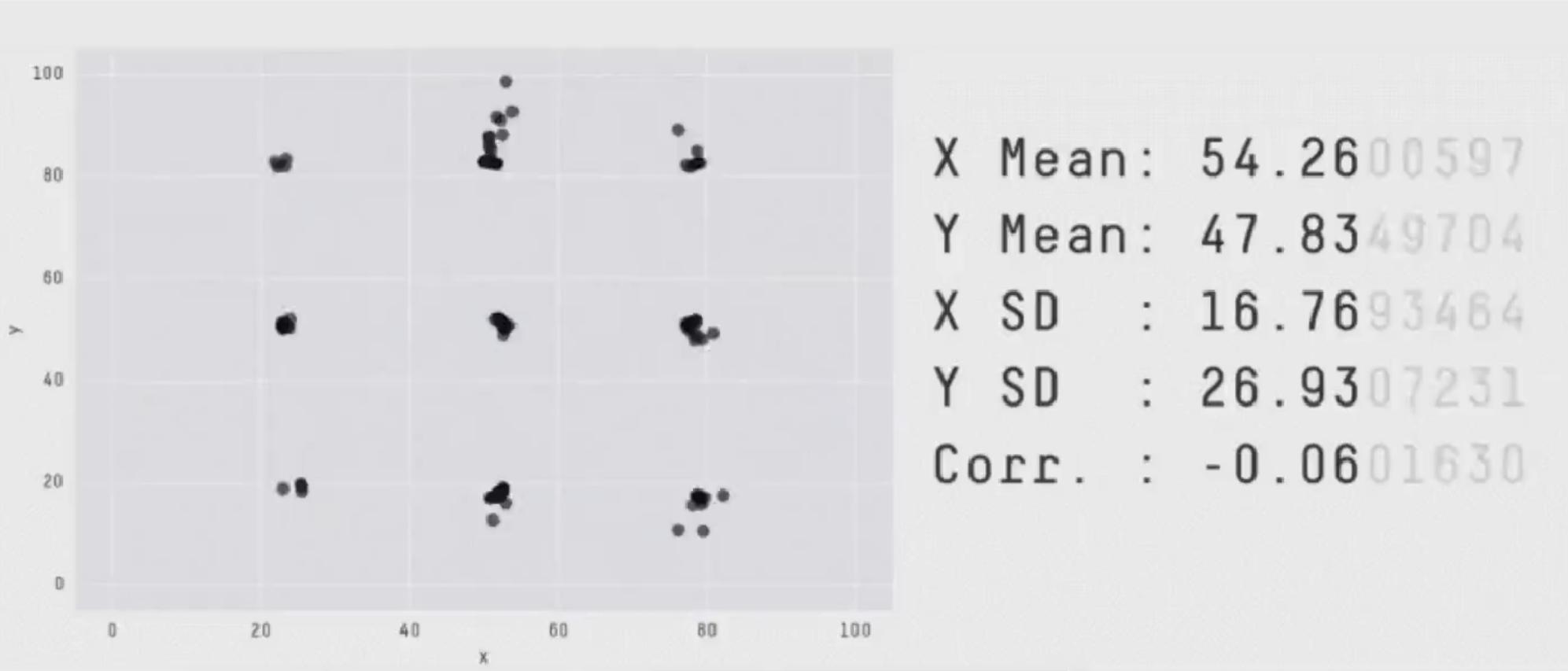


[KS14]

Different velocities, same streamlines



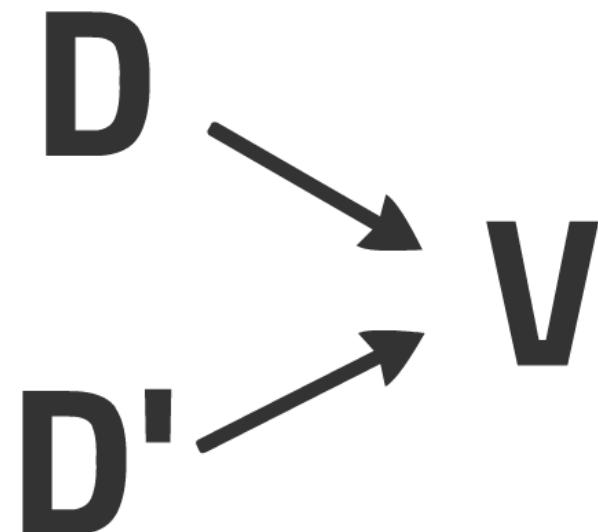
[KS14]



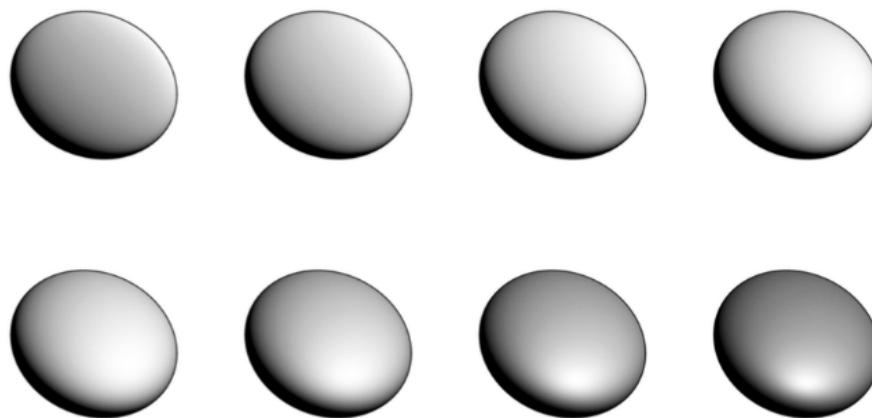
[MF17]

Confusion

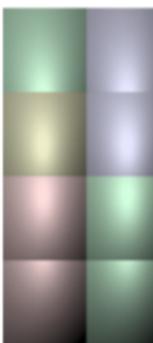
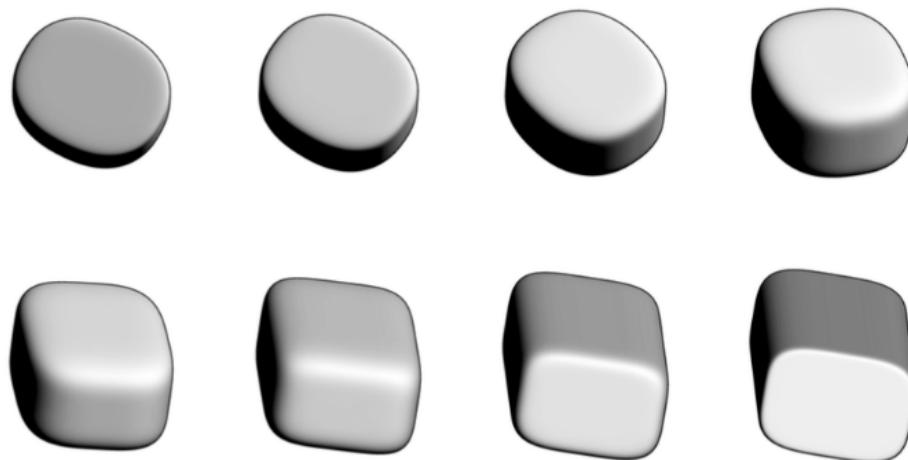
Different data
produces the same
visualization



Very different tensors look similar

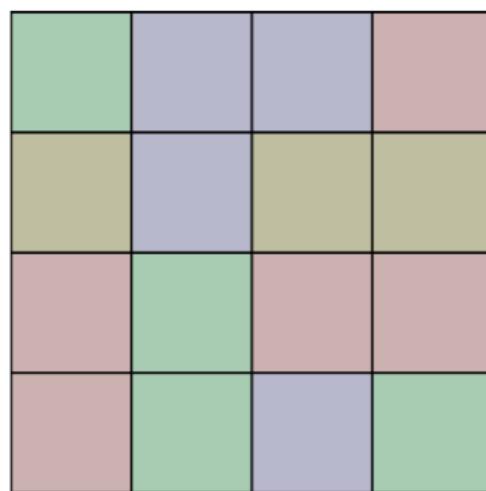


[KS14]



[KS14]

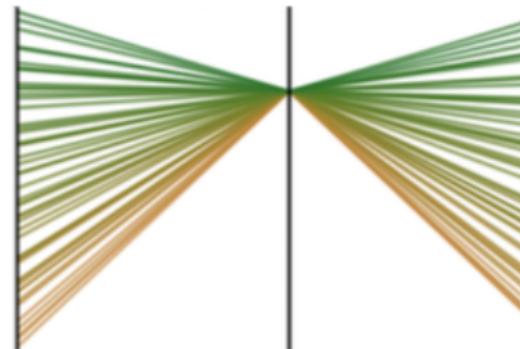
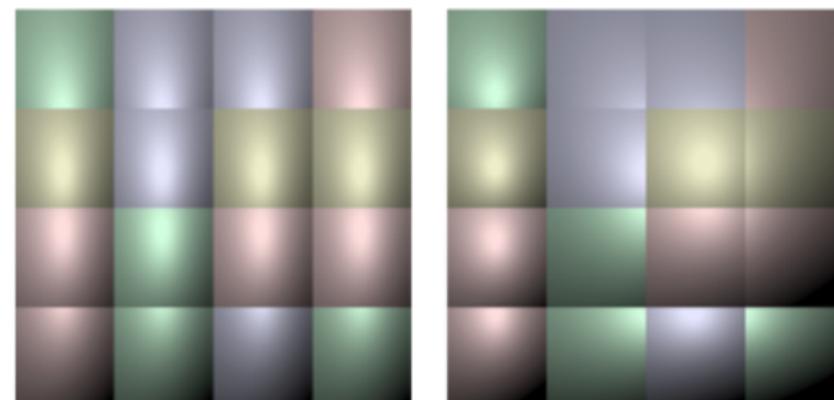
Different trees, same treemap



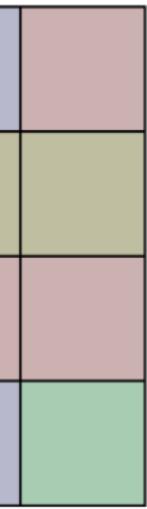
[KS14]

[KS14]

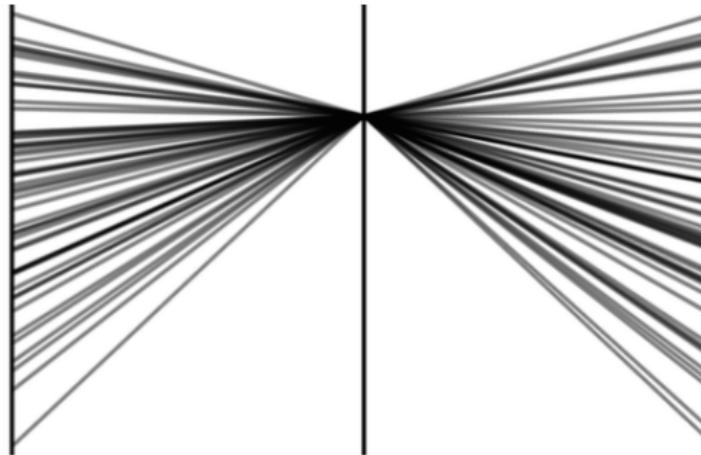
Different c



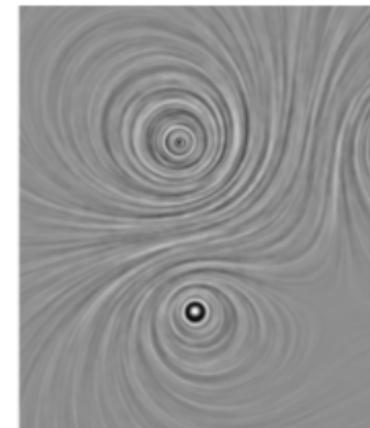
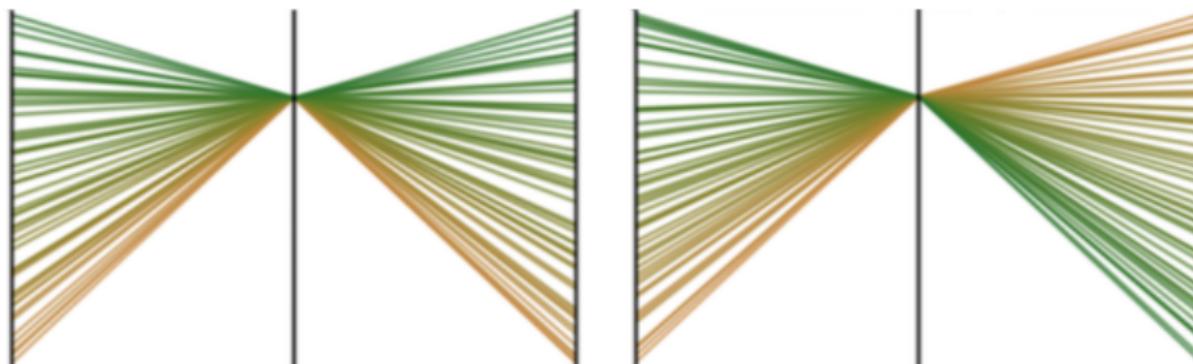
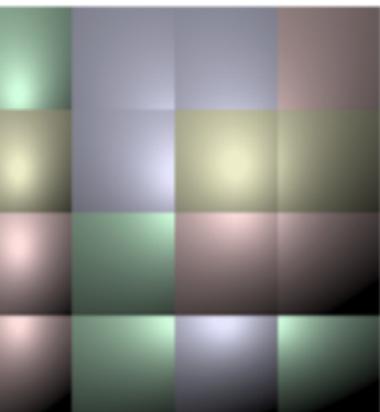
treemap



Different data sets, same plot

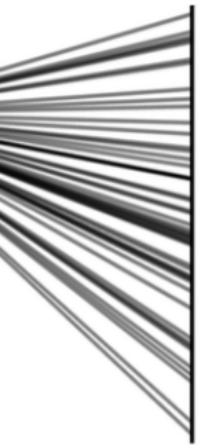


[KS14]

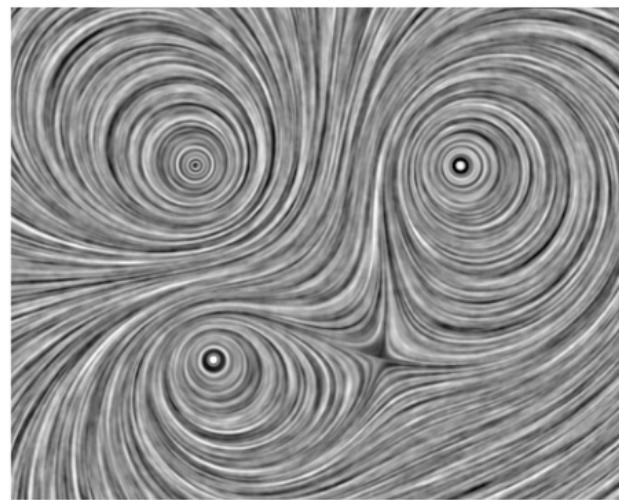


DI

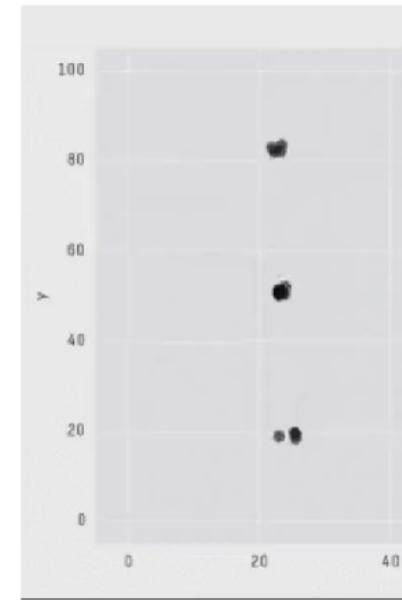
plot



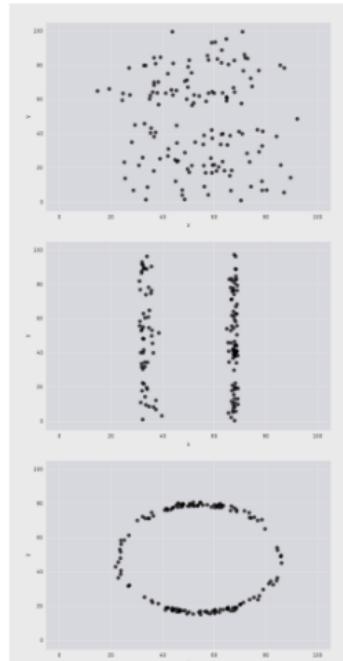
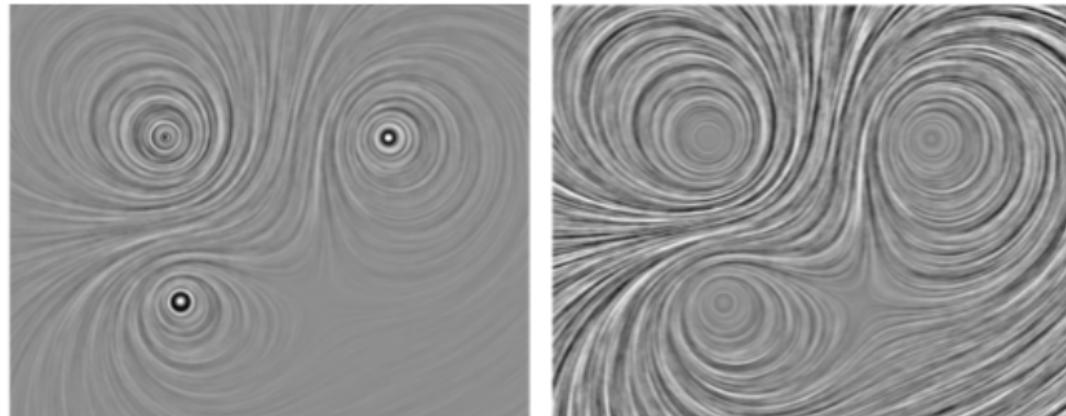
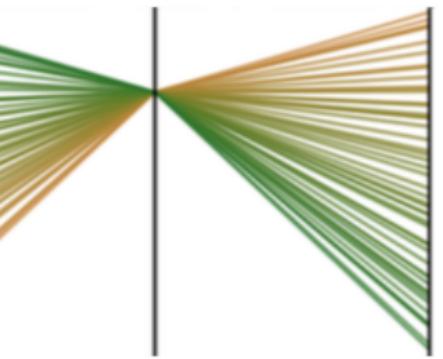
Different velocities, same streamlines



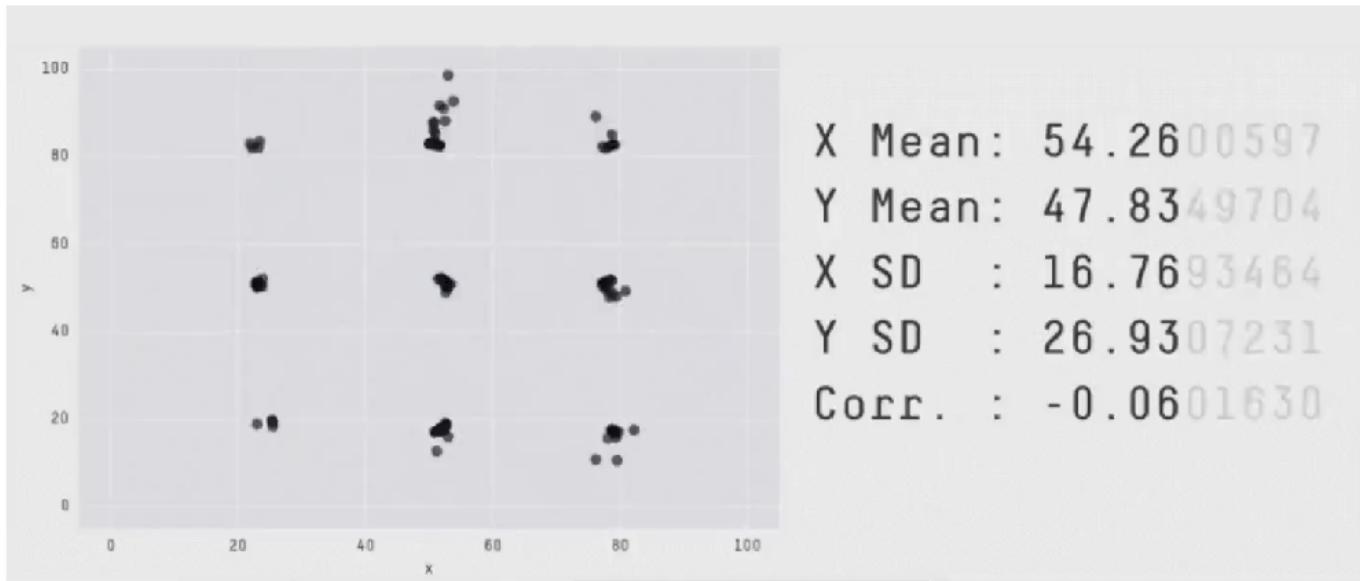
[KS14]



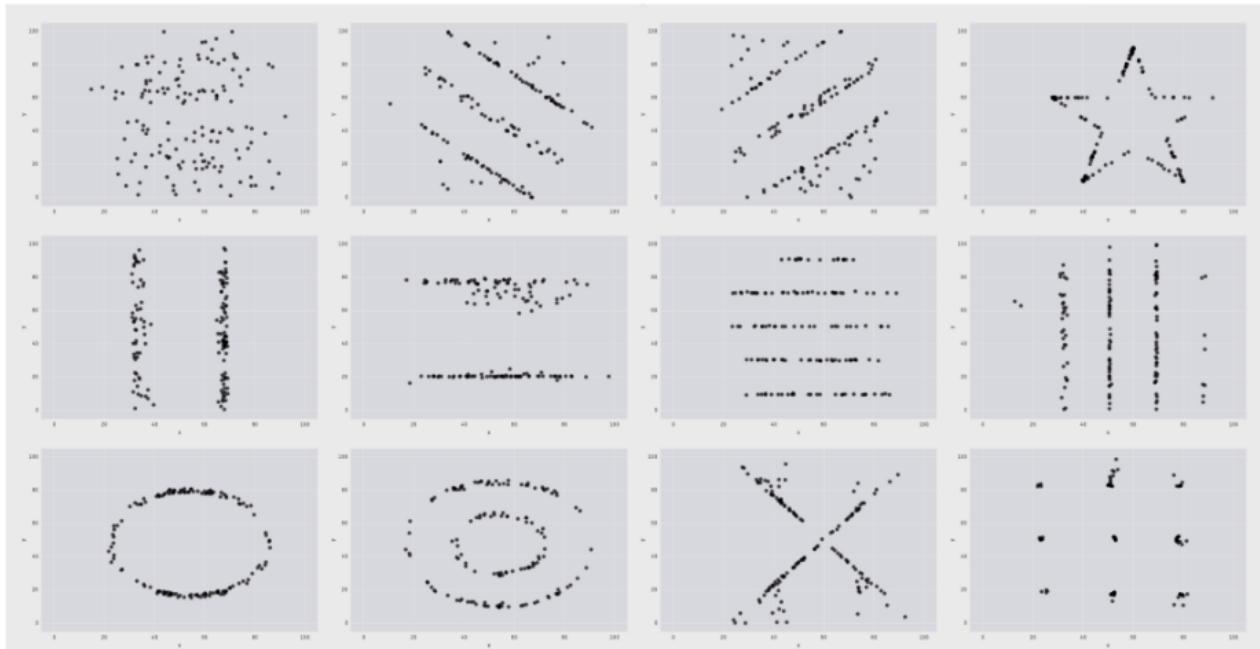
[MF17]



lines

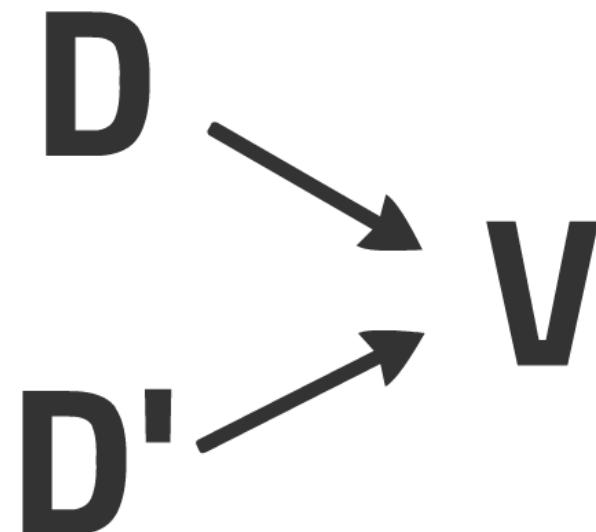


[MF17]



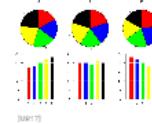
Confusion

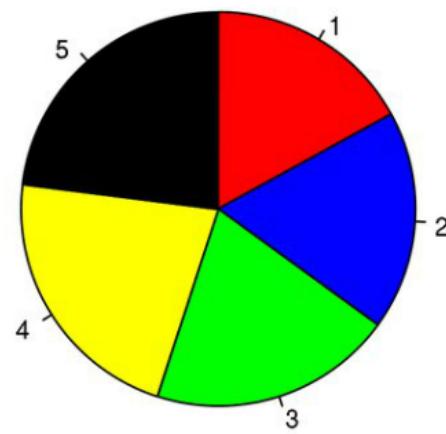
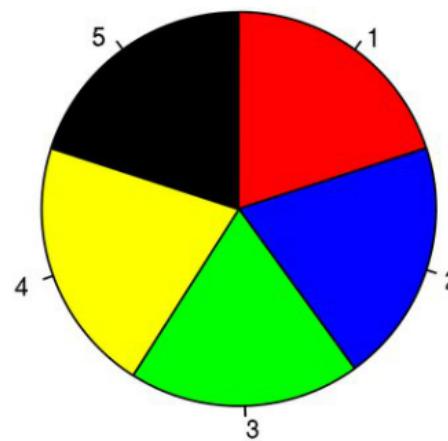
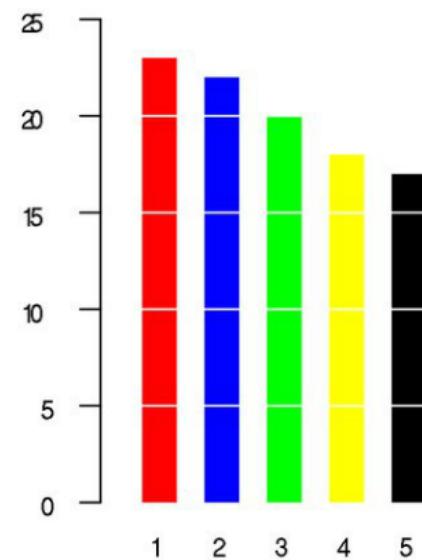
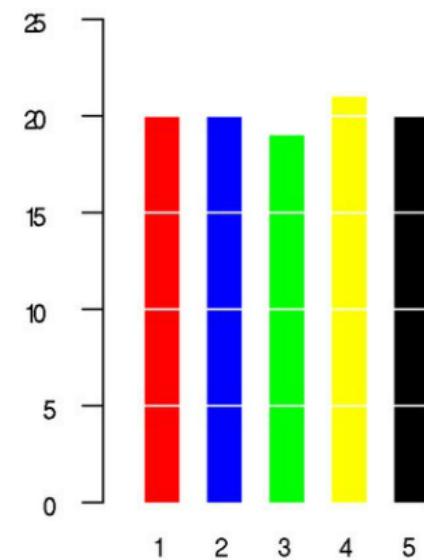
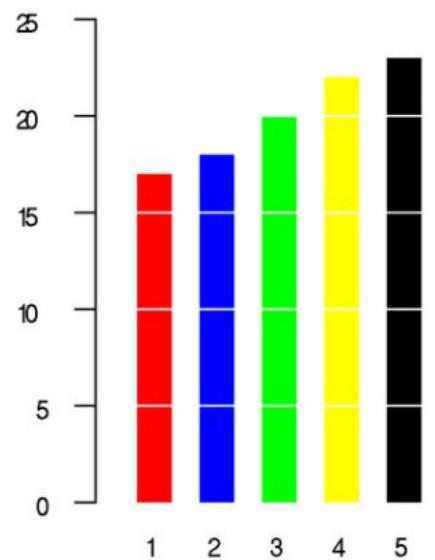
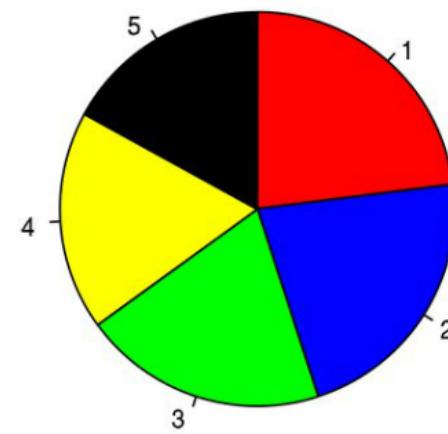
Different data
produces the same
visualization



Finding confusion

- Loss/gain of dimension
- Aliasing
- Human or computer

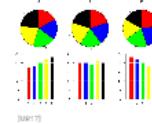


A**B****C**

[MR17]

Finding confusion

- Loss/gain of dimension
- Aliasing
- Human or computer

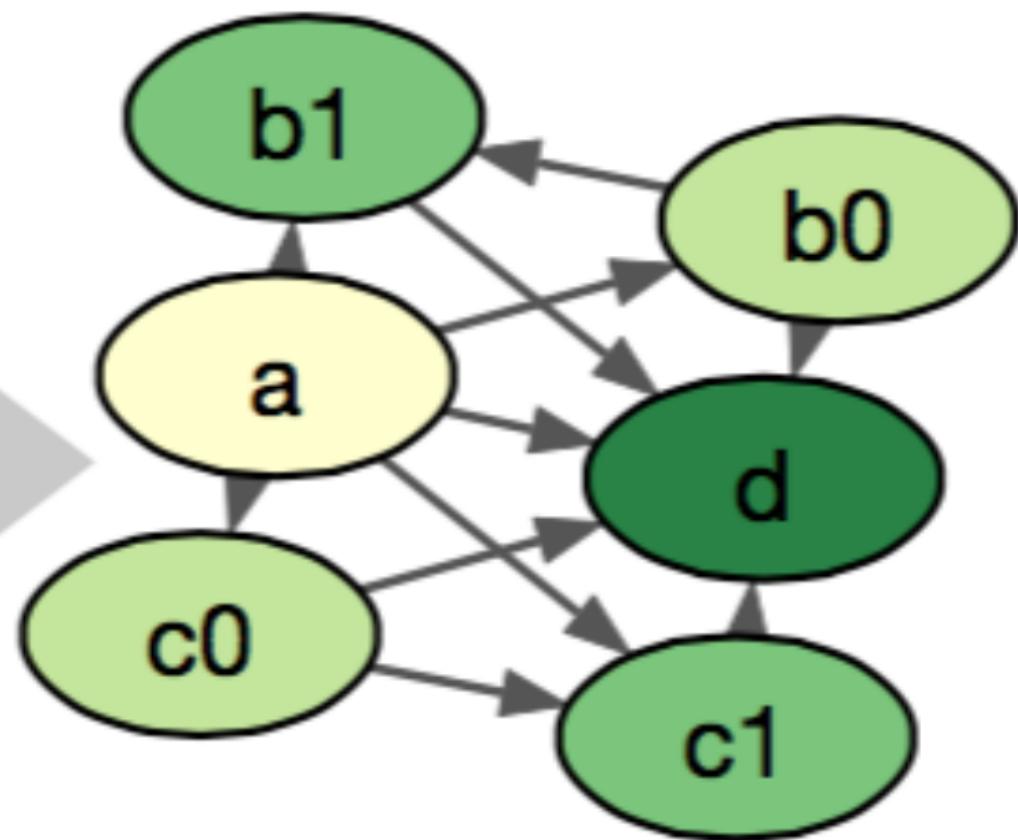
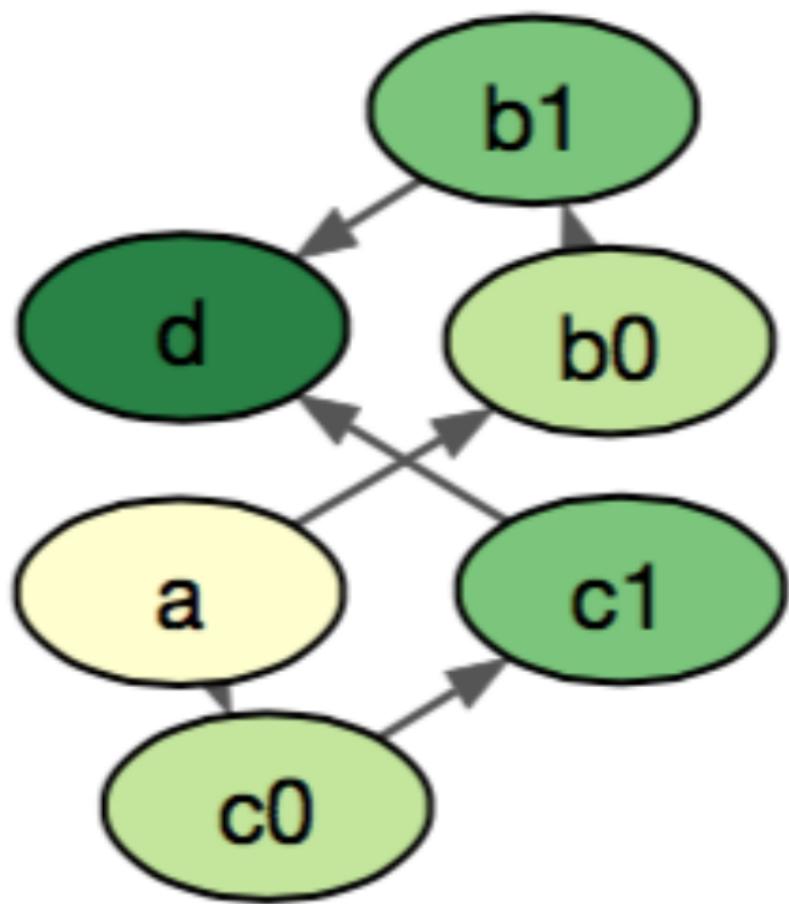


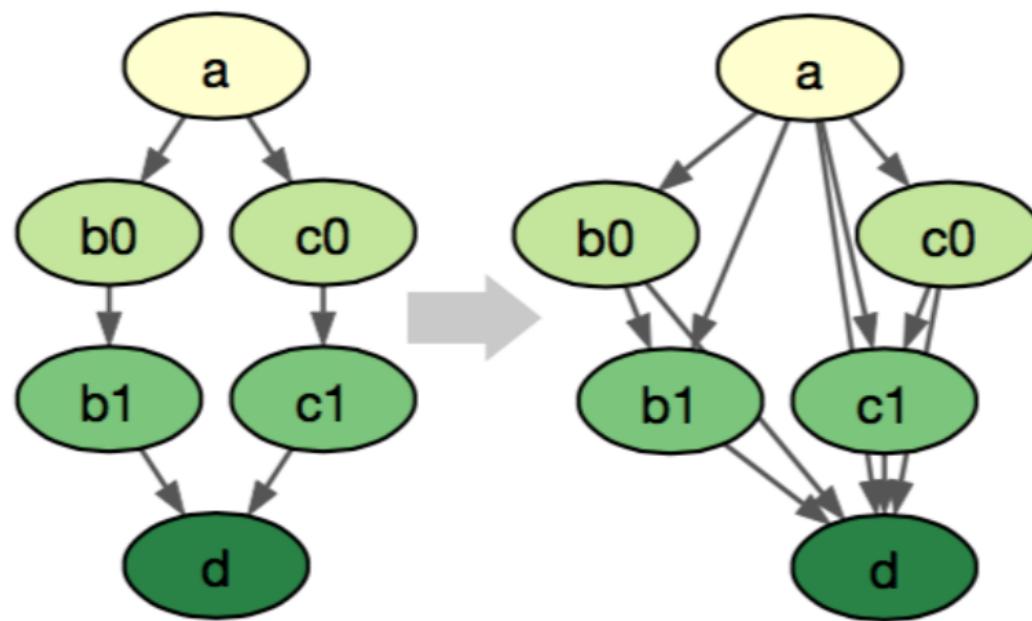
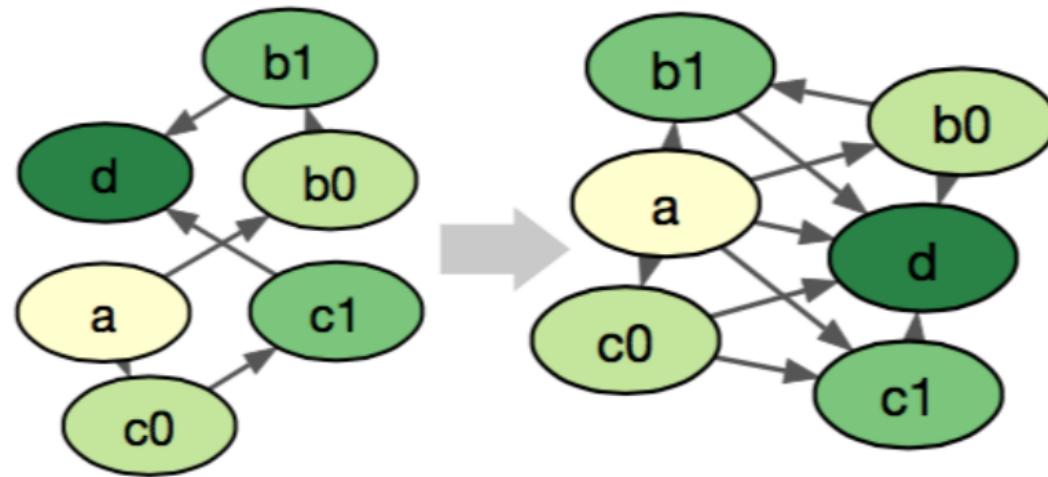
- Invariance
- Unambiguity
- Correspondence

Correspondence

- Invariant and Unambiguous
- Meaningful changes to the data lead to meaningful changes in the visualization (and vice versa)

$$D \cong V$$





[KS14]

- Invariance
- Unambiguity
- Correspondence

- **Data**
- **Representation**
- **Visualization**

D → R → V

- Sampling error
- Floating point error
- Algorithmic bias
- Editorial bias
- Tractability bias
- Modeling error
- ...

VISUALIZATION



- Sampling error
- Floating point error
- Algorithmic bias
- Editorial bias
- Tractability bias
- Modeling error
- ...

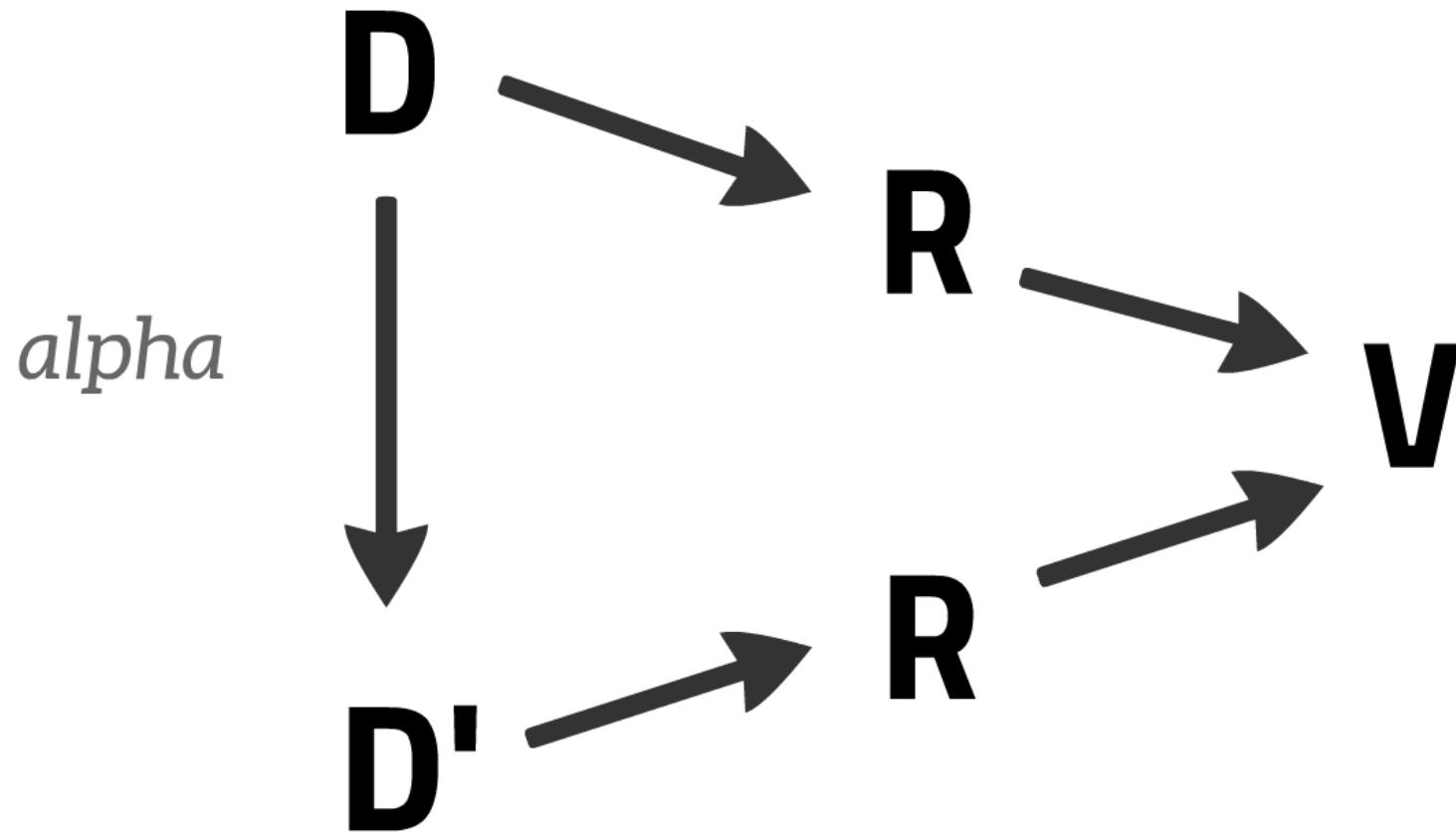
Visualization defined

Visualization defined

- ~~data => marks~~
- Δ data $<=>$ Δ marks
- meaningful changes to data $<=>$ meaningful changes to vis
- **accurate visualizations are isomorphic to data**

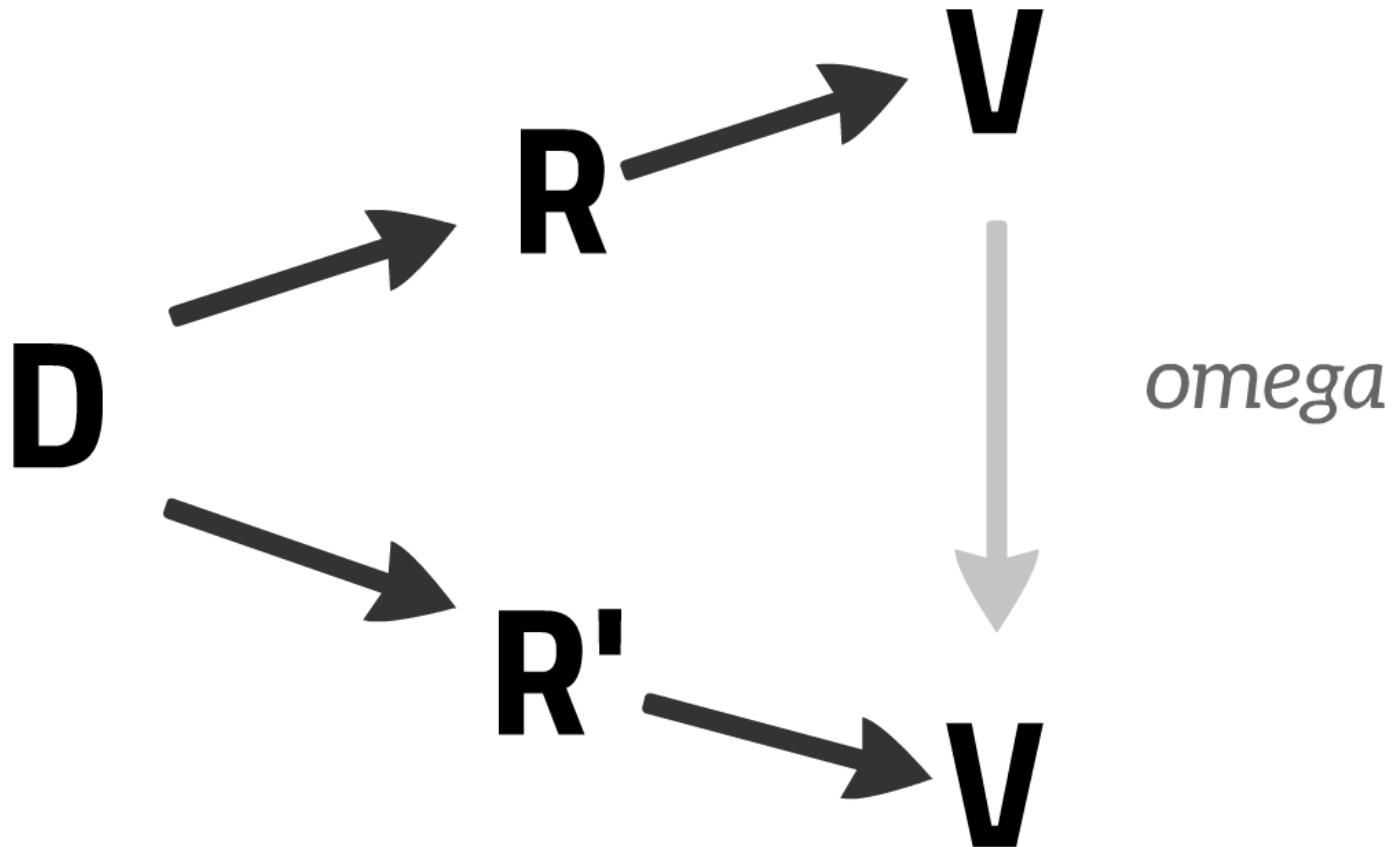
unambiguous depiction

confusion: something in the data that's not in the visualization



invariant representation

hallucination: something in the visualization that's not in the data



correspondence

	alpha	omega	property
invariance	1	1	well-defined
unambiguity	big	big	injective
correspondence	same	same	isomorphism



	alpha	omega	property
invariance	1	1	well-defined
unambiguity	big	big	injective
correspondence	same	same	isomorphism

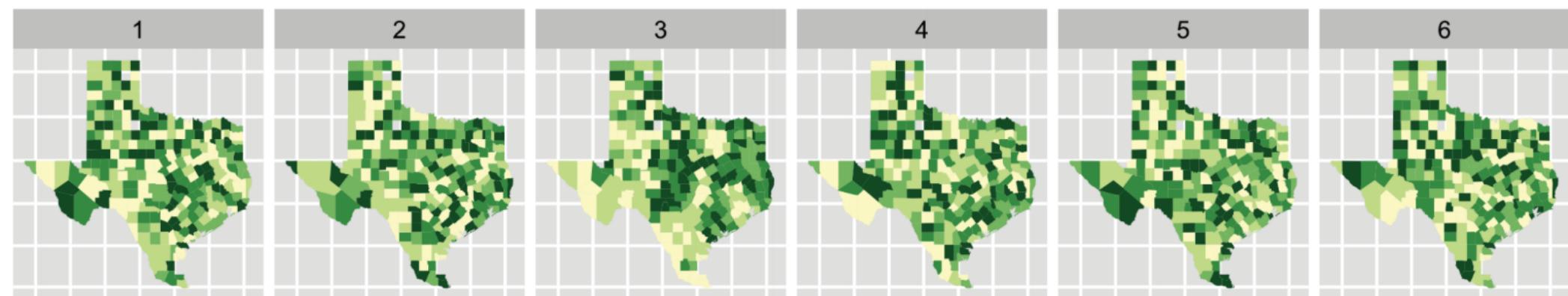
Three questions to improve visualization

- Is there anything in the visualization that's not in the data? (hallucination)
- Is there anything in the data that's not in the visualization? (confusion)
- Do meaningful changes in the data imply meaningful changes in the vis, and vice versa? (correspondence)

correspondence

	alpha	omega	property
invariance	1	1	well-defined
unambiguity	big	big	injective
correspondence	same	same	isomorphism





[WCHA10]

believe	believe	believe	believe	believe
case	case	case	case	case
case	closely	case	closely	case
closely descendants				
descendants few few				
long long modified				
modified variations				
variations very				

believe believe

case

case closely

closely descendants

descendants few few

long long modified

modified variations

variations **very**

very view view

believe believe

case

case closely

closely descendants

descendants few few

long long modified

modified variations

variations **very**

very view view

believe believe

case

case closely

closely descendants

descendants few few

long long modified

modified variations

variations **very**

very view view

believe believe

case

case closely

closely descendants

descendants few few

long long modified

modified variations

variations **very**

very view view

believe believe

case

case closely

closely descendants

descendants few few

long long modified

modified variations

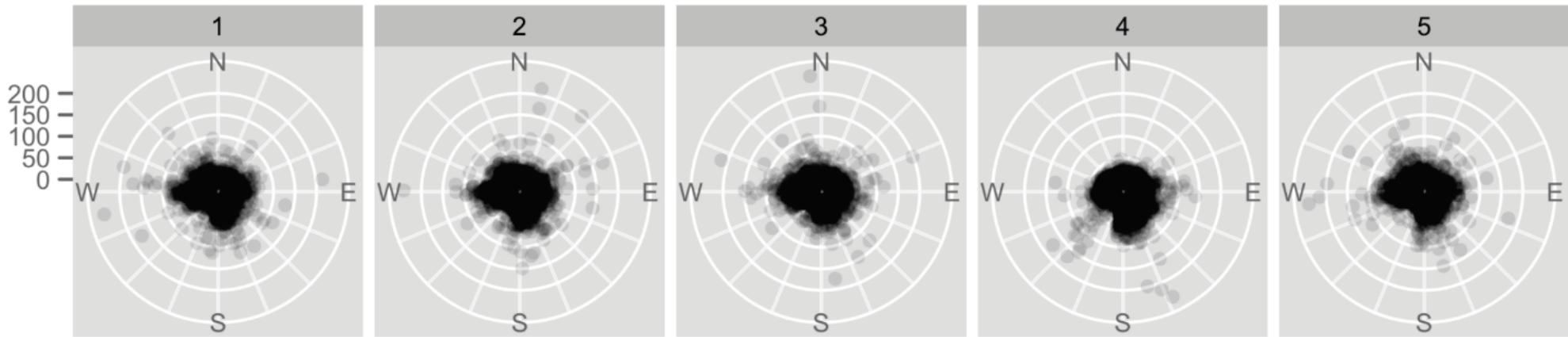
variations **very**

very view view

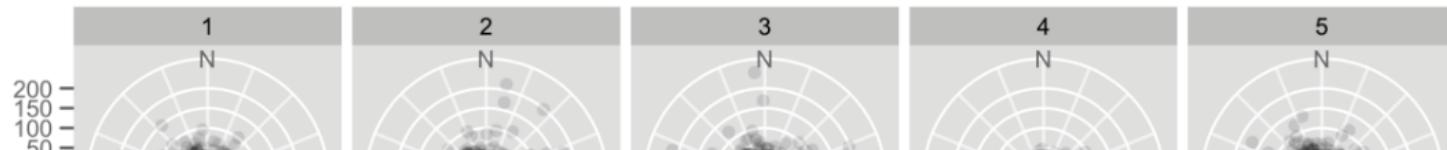
[WCHA10]

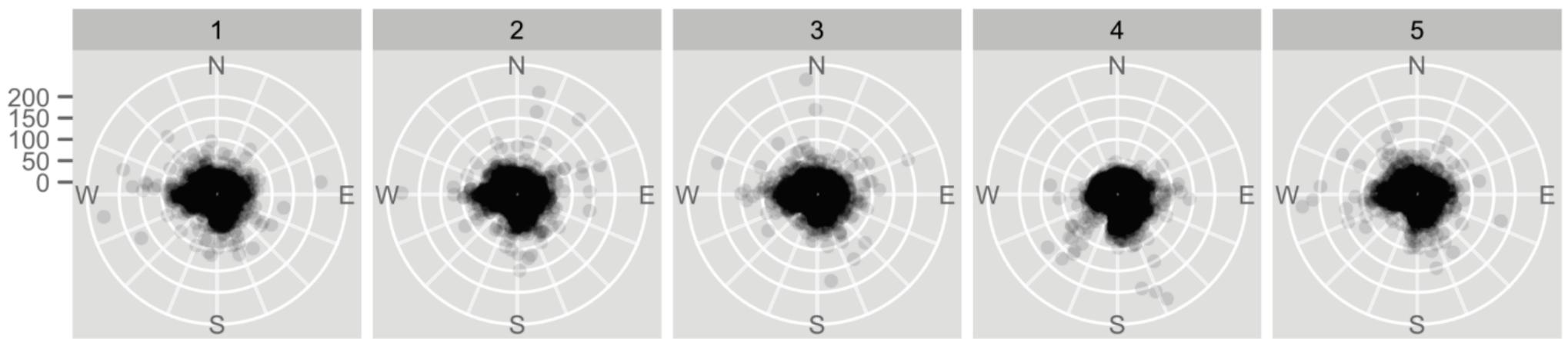
variations **very**
very view view

[WCHA10]

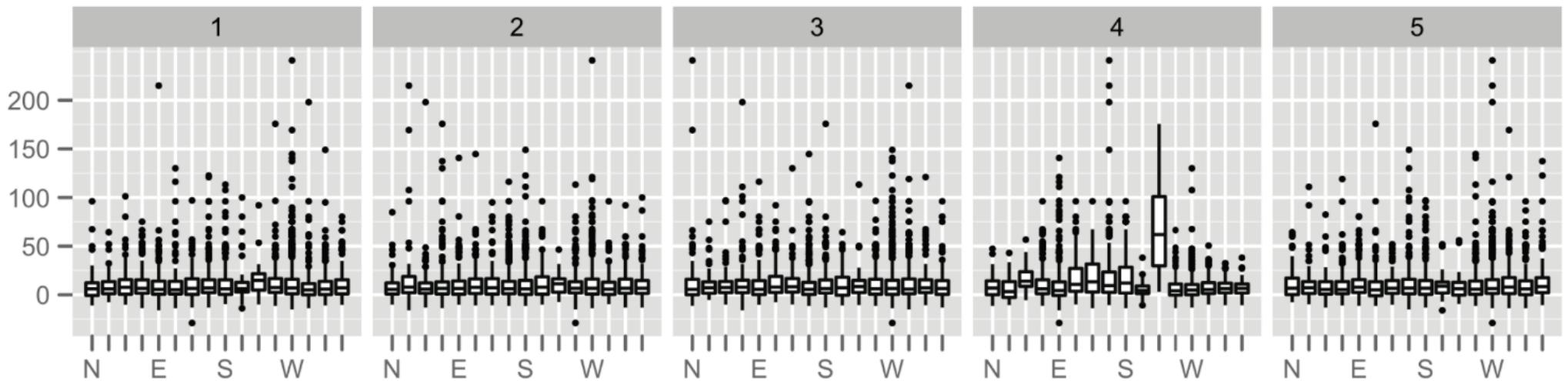


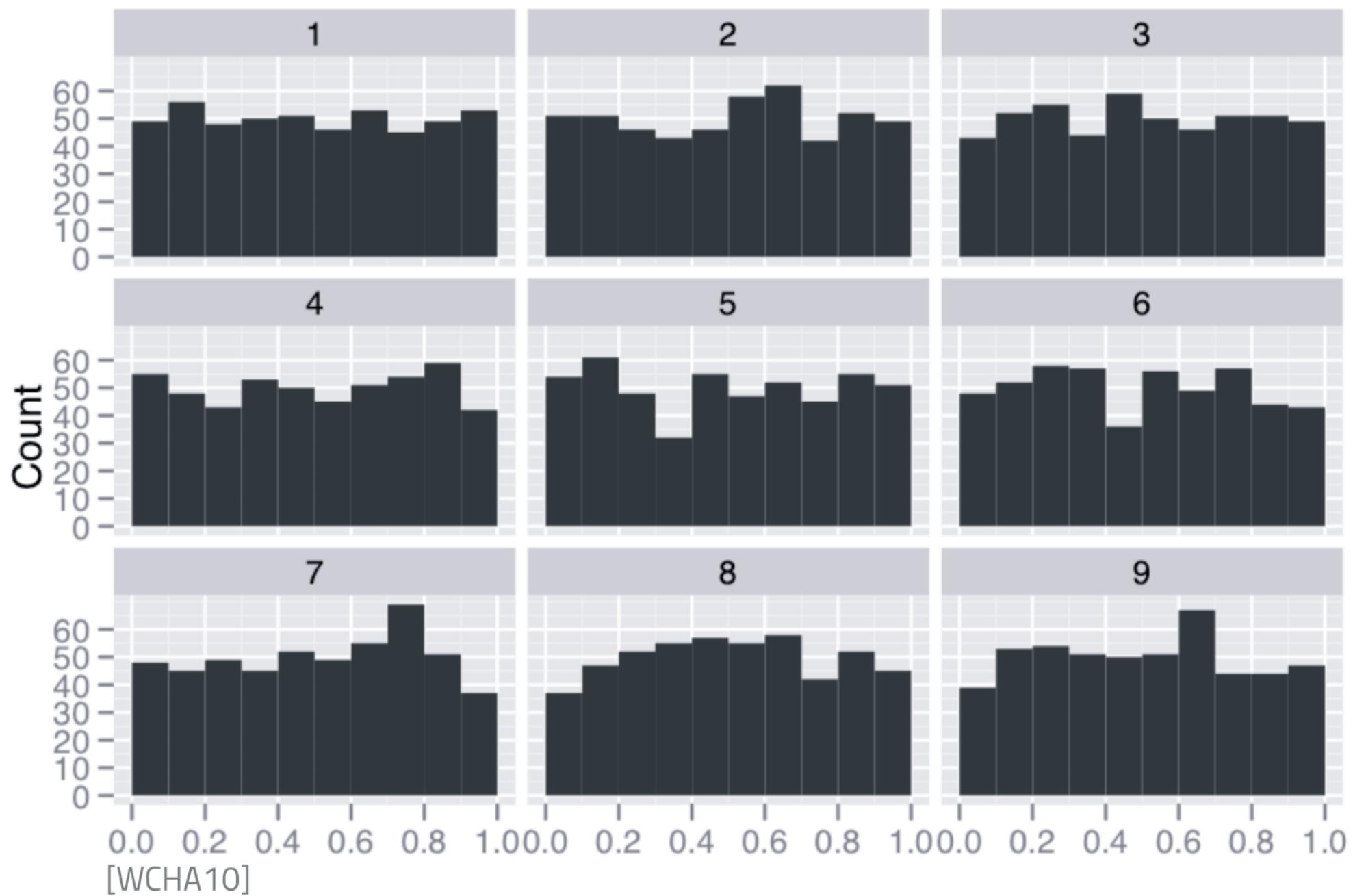
[WCHA10]





[WCHA10]





visualization? (confusion)

- Do meaningful changes in the data imply meaningful changes in the vis, and vice versa? (correspondence)

Computers can verify visualizations

Python <https://github.com/quiltdata/algebraic-vis>

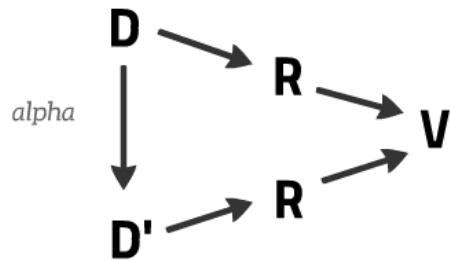
d3.js <https://github.com/hdc-arizona/d3-algebra>

Visualization defined

- ~~data => marks~~
- $\Delta \text{ data} \Leftrightarrow \Delta \text{ marks}$
- meaningful changes to data \Leftrightarrow meaningful changes to vis
- **accurate visualizations are isomorphic to data**

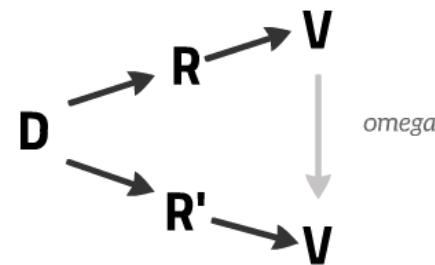
unambiguous depiction

confusion: something in the data that's not in the visualization

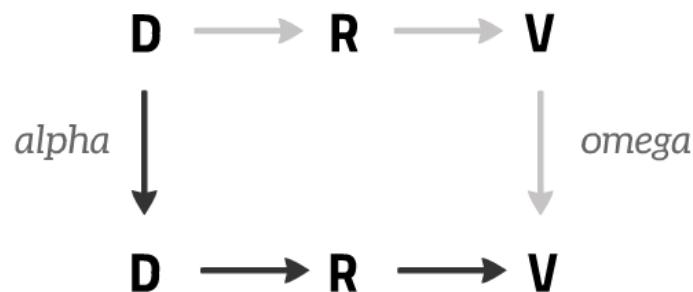


invariant representation

hallucination: something in the visualization that's not in the data

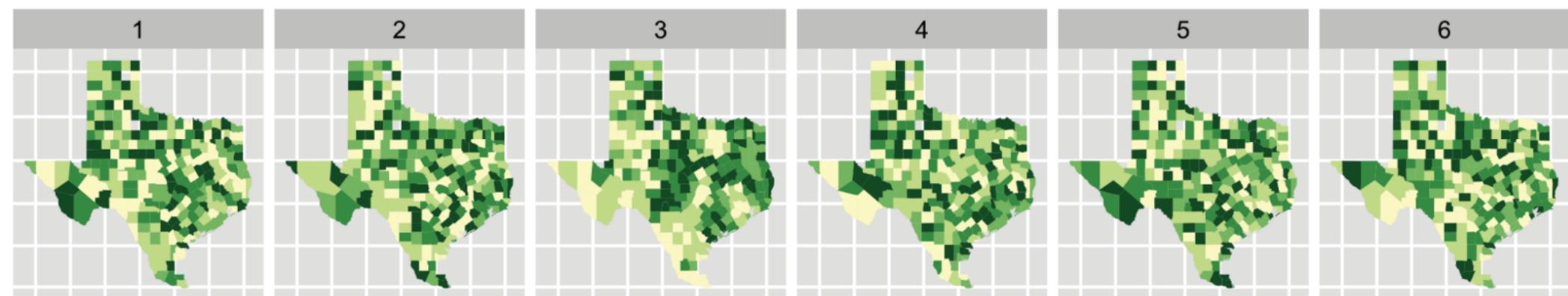


correspondence



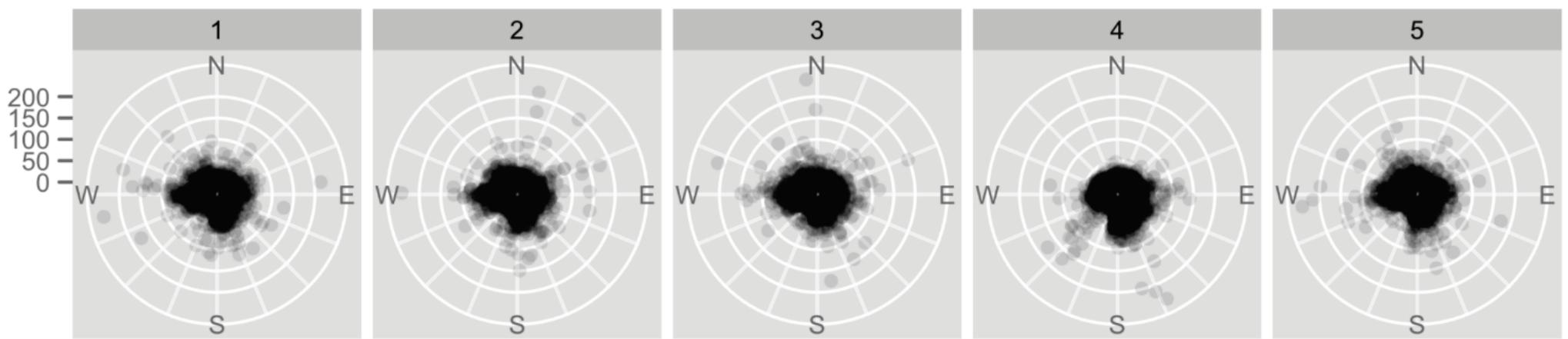
Three questions to improve visualization

- Is there anything in the visualization that's not in the data? (hallucination)
- Is there anything in the data that's not in the visualization? (confusion)
- Do meaningful changes in the data imply meaningful changes in the vis, and vice versa? (correspondence)

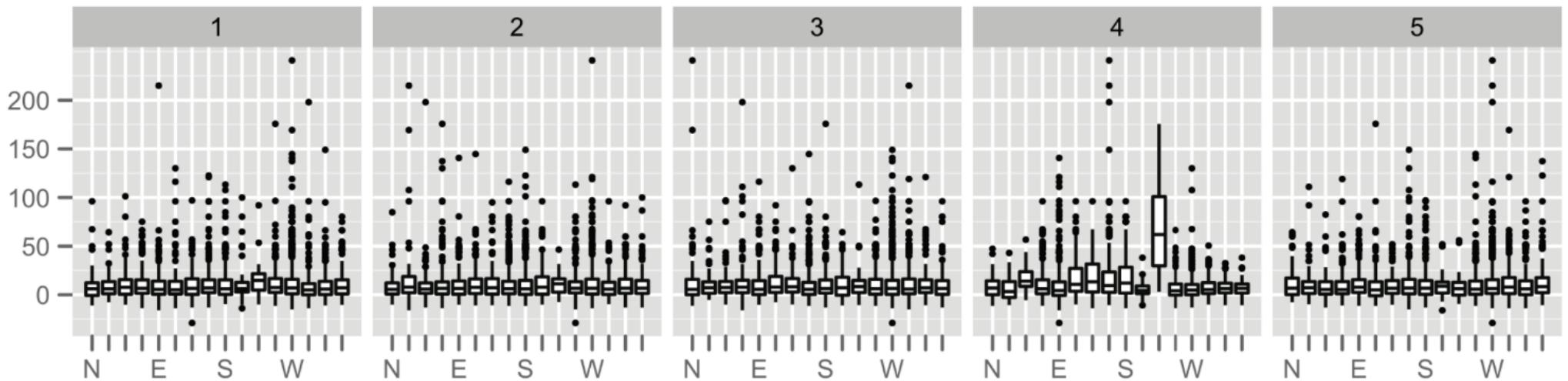


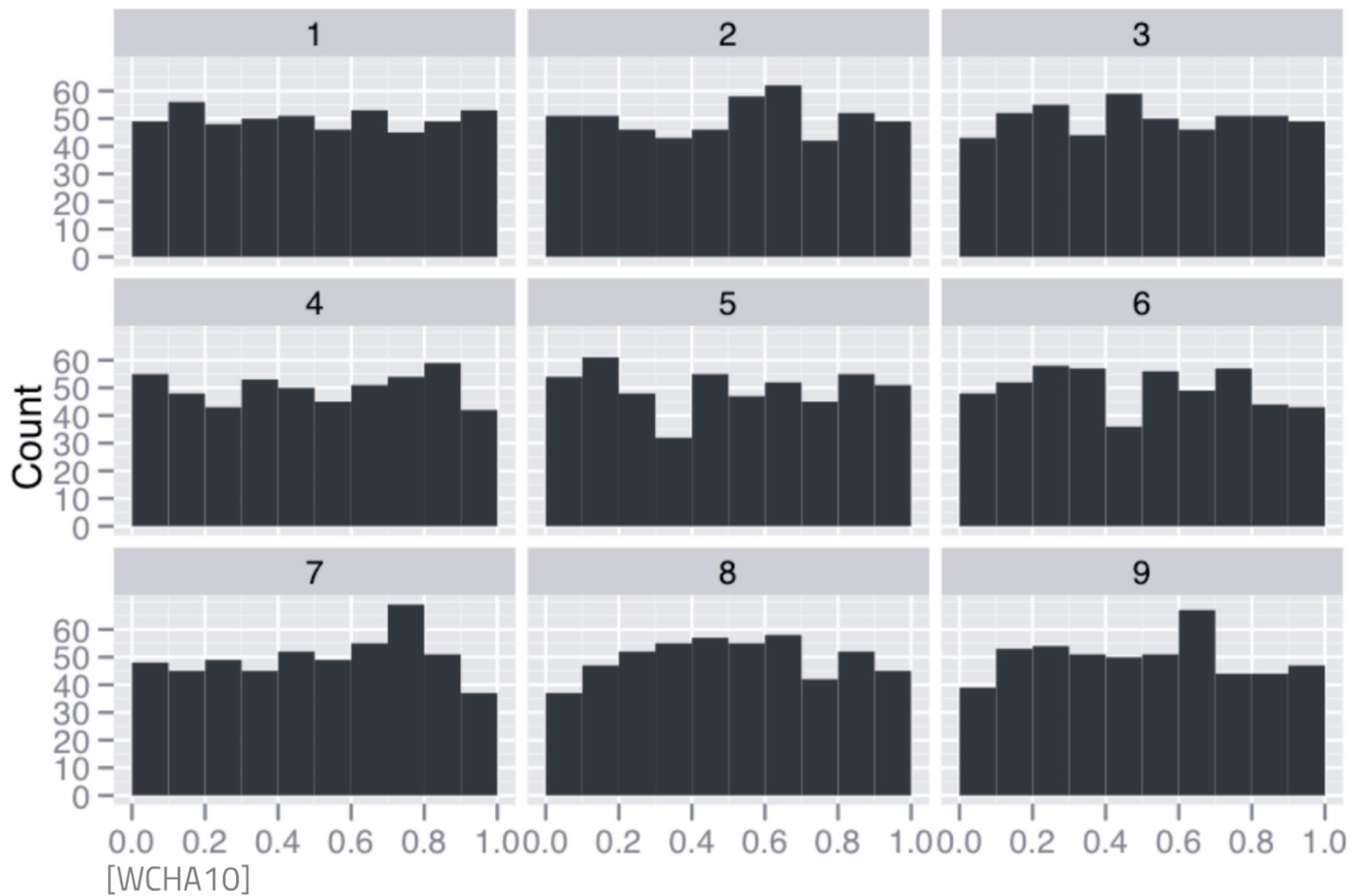
[WCHA10]

believe	believe	believe	believe	believe
case	case	case	case	case
case	closely	case	closely	case
closely descendants				
descendants few few				
long long modified				
modified variations				
variations very				



[WCHA10]





correspondence

	alpha	omega	property
invariance	1	1	well-defined
unambiguity	big	big	injective
correspondence	same	same	isomorphism



Bibliography

- *An Algebraic Process for Visualization Design*, G. Kindlmann and C. Scheidegger [**KS14 (KS15)**]
- *Algebraic Visualization Design for Pedagogy*, G. Kindlmann and C. Scheidegger [**KS16**]
- *Splatterplots: Overcoming Overdraw in Scatter Plots*, IEEE Transactions on Visualization and Computer Graphics, Volume 19, Number 9, page 1526--1538 — Sep 2013, Adrian Mayorga, Michael Gleicher [**MG13**]
- *Inceptionism: Going Deeper into Neural Networks*, Alexander Mordvintsev, Software Engineer, Christopher Olah, Software Engineering Intern and Mike Tyka, Software Engineer <http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html> [**In15**]
- *The bas-relief ambiguity*, Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on, P. N. Belhumeur, D. J. Kriegman, A. L. Yuille [**BKY97**]
- *Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing*, Justin Matejka, George Fitzmaurice [**MF17**]
- <https://twitter.com/MaxCRoser/status/857389434756505600> [**MR17**]
- *Graphical Inference for Infovis*, Hadley Wickham, Dianne Cook, Heike Hofmann, and Andreas Buja, IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 16, NO. 6, NOVEMBER/DECEMBER 2010 [**WCHA10**]

aneesh@quiltdata.io
quiltdata.com