

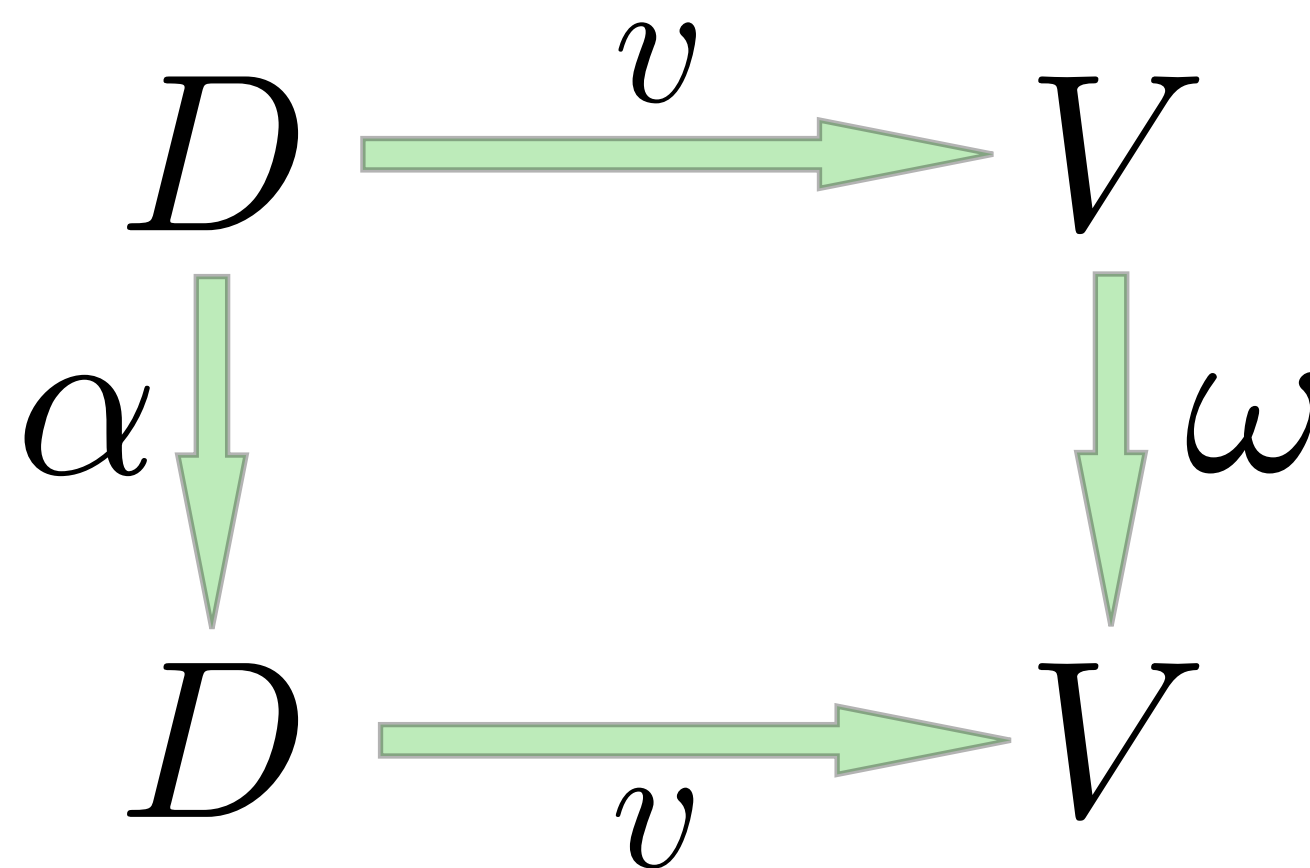
Algebraic Visualization Design for Pedagogy

Gordon Kindlmann

Computer Science & Computation Institute
University of Chicago

Carlos Scheidegger

Computer Science
University of Arizona



Our Basic Goal

Empower students with future-proof
critical thinking about reading and making
visualizations

(not: training on popular toolkits)

U of Chicago T-shirts: “That’s all very well in
practice, but how does it work in theory?”

Our Strategy

Structure presentation of vis methods and guidelines around a common underlying theory:

Algebraic Visualization Design

G Kindlmann and C Scheidegger. *An Algebraic Process for Visualization Design*. IEEE TVCG (Proceedings VIS 2014), 20(12):2181–2190, November 2014.

So first, a quick primer on Category Theory,
Functors, and Fiber Products

A joke! actually ...

Don't need math to ask focused questions

Two questions to ask of any vis:

1) If world was different in some interesting way, how would the data be different (α)?

2) How then would the visualization look different (ω)? Does it make visual sense?

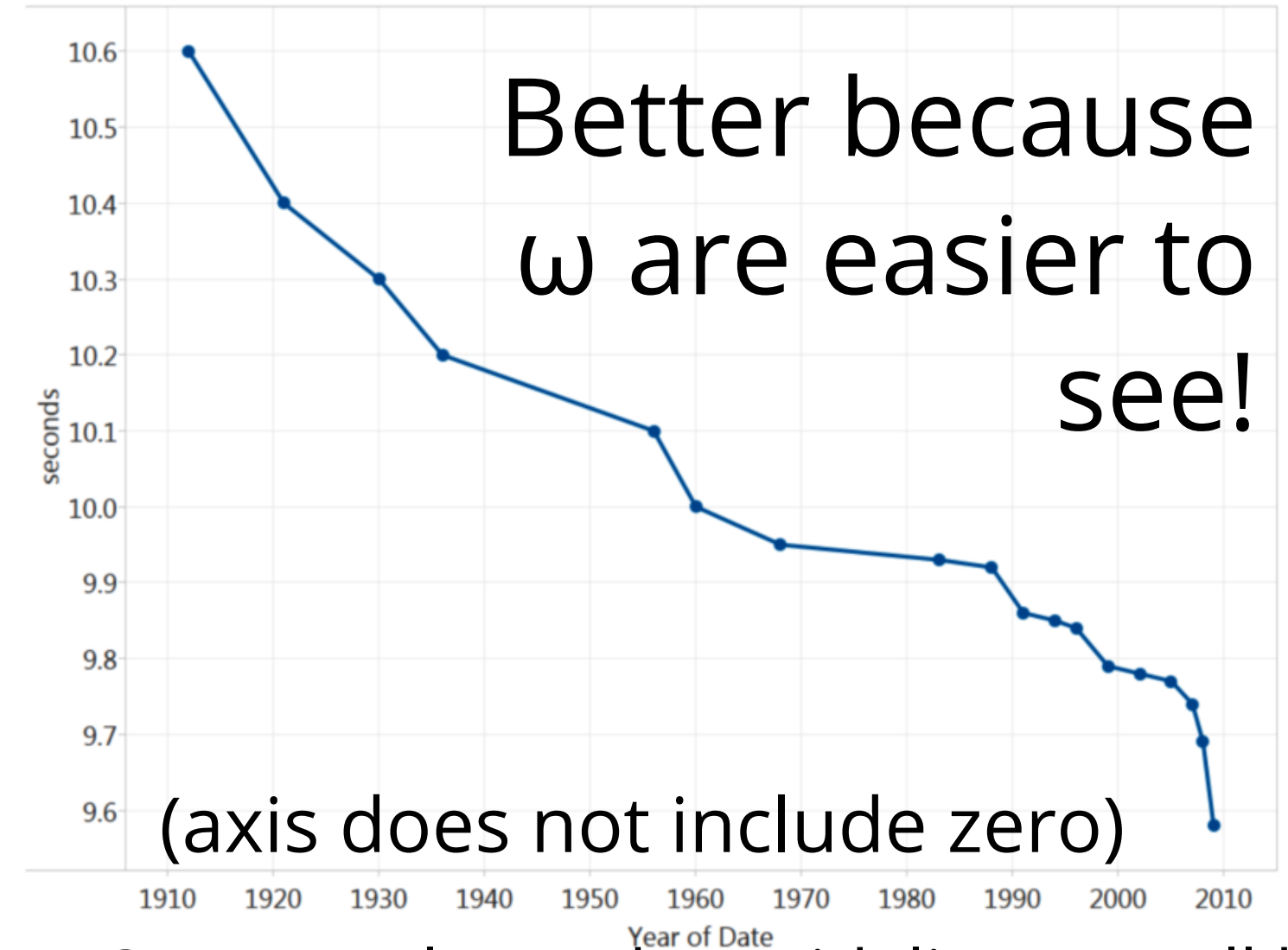
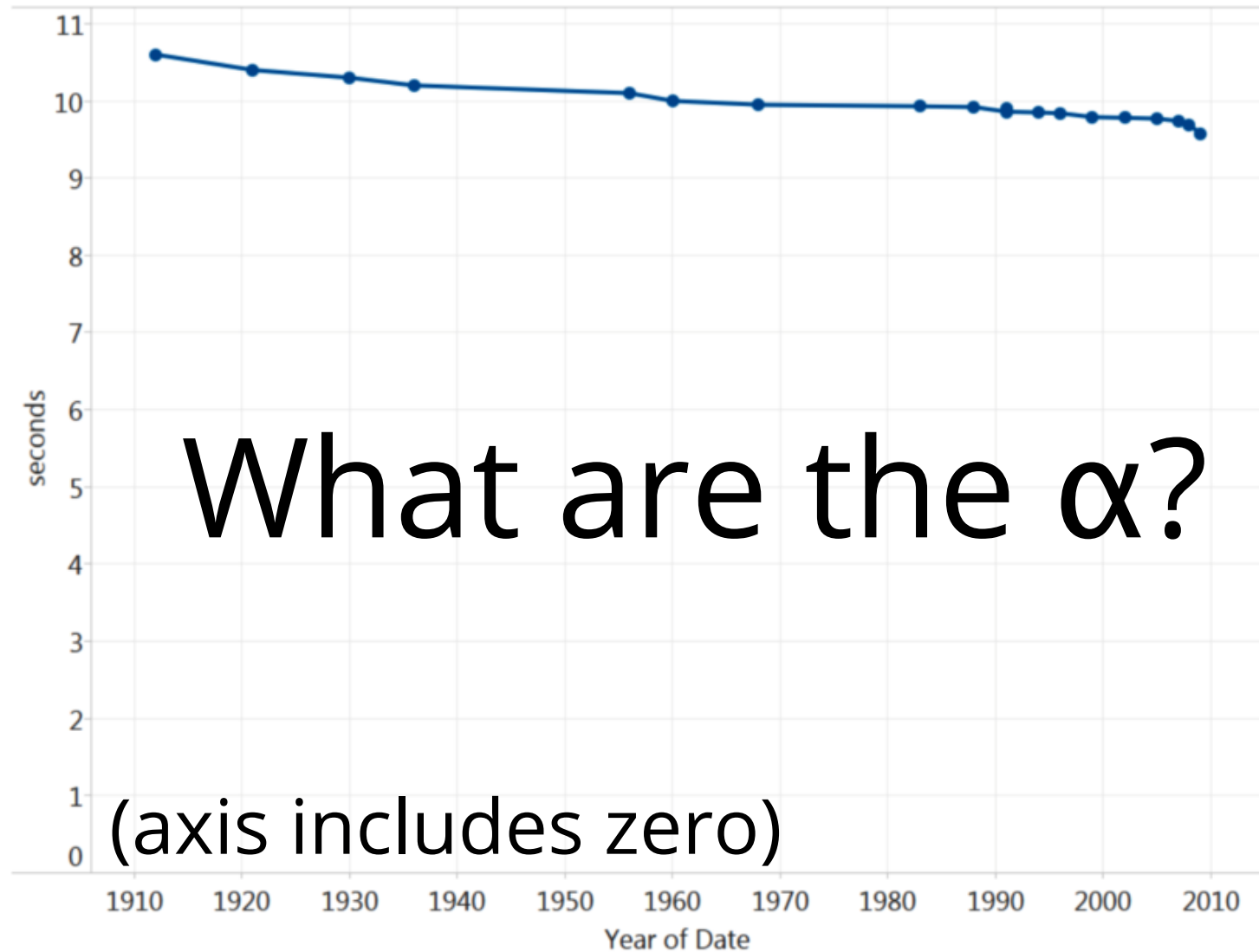
Connections: 1) world-data, 2) vis-user

Knowing Taxonomies \neq Answering questions

Include zero or not on vertical axis?

<http://gravyanecdote.com/uncategorized/mythbusters-should-you-start-your-axes-at-zero>

History of men's 100m dash world records

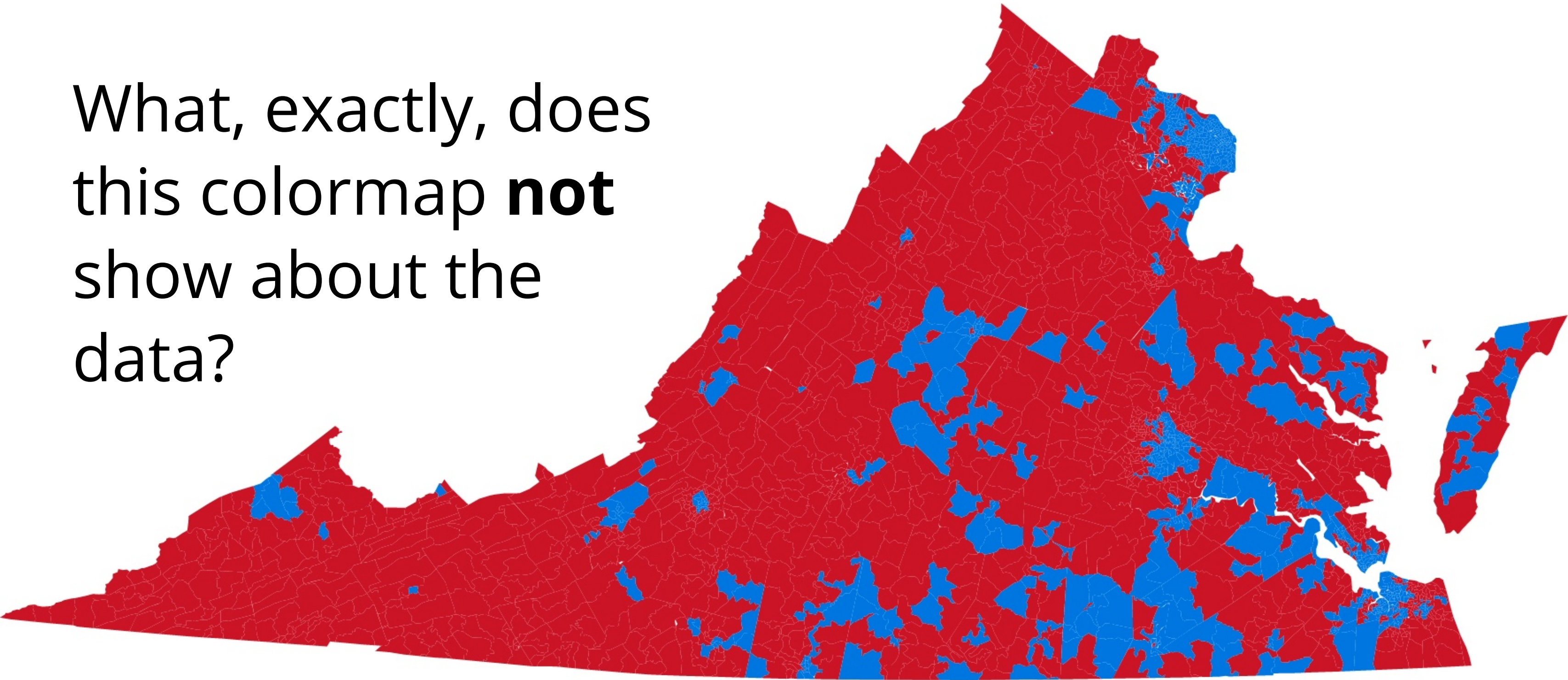


Cotgreave: "... doesn't really expose the change of the record over time ... or highlight impact Usain Bolt had on record"

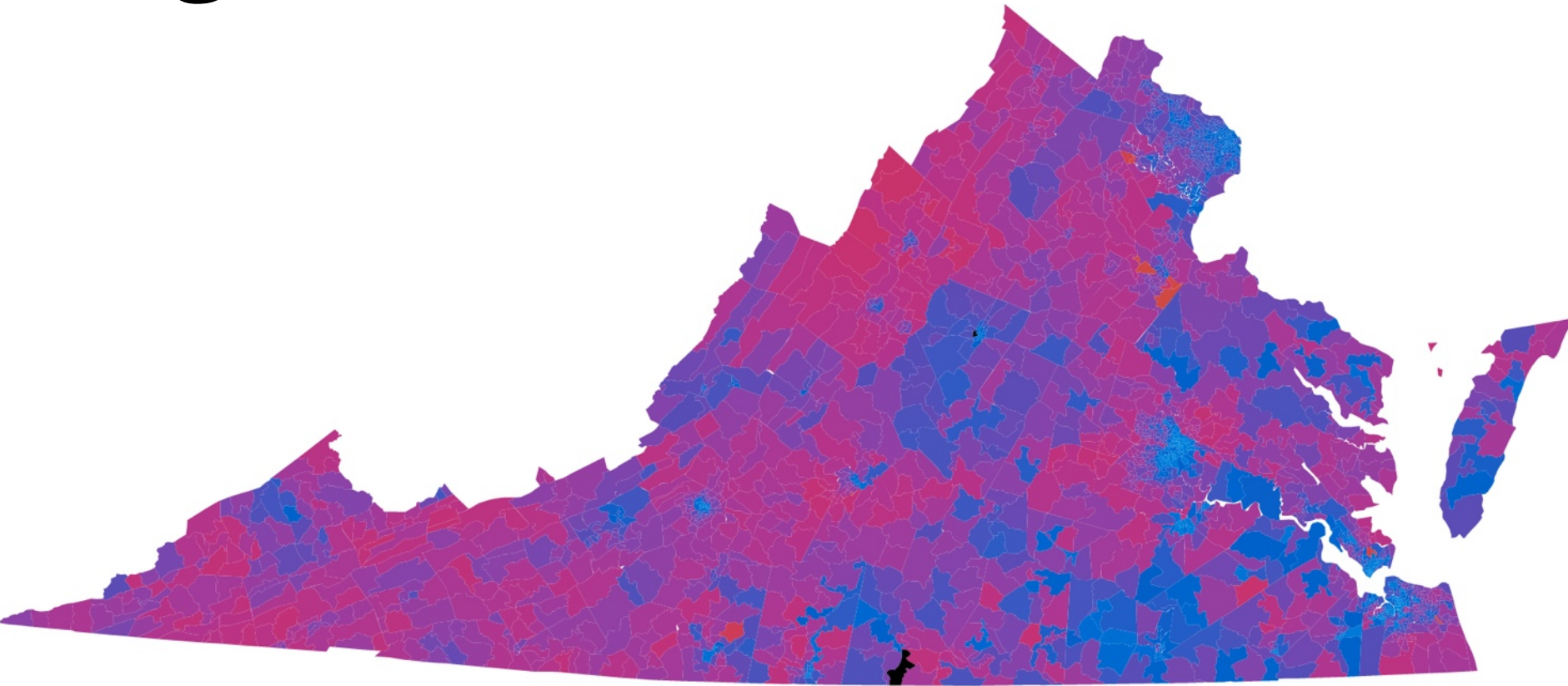
"Once you learn the guidelines, you'll be able to fine tune your charts by **bending or breaking them** according to your use case and objective." (not actionable)

Virginia (US) 2014 Senate Election results, per precinct

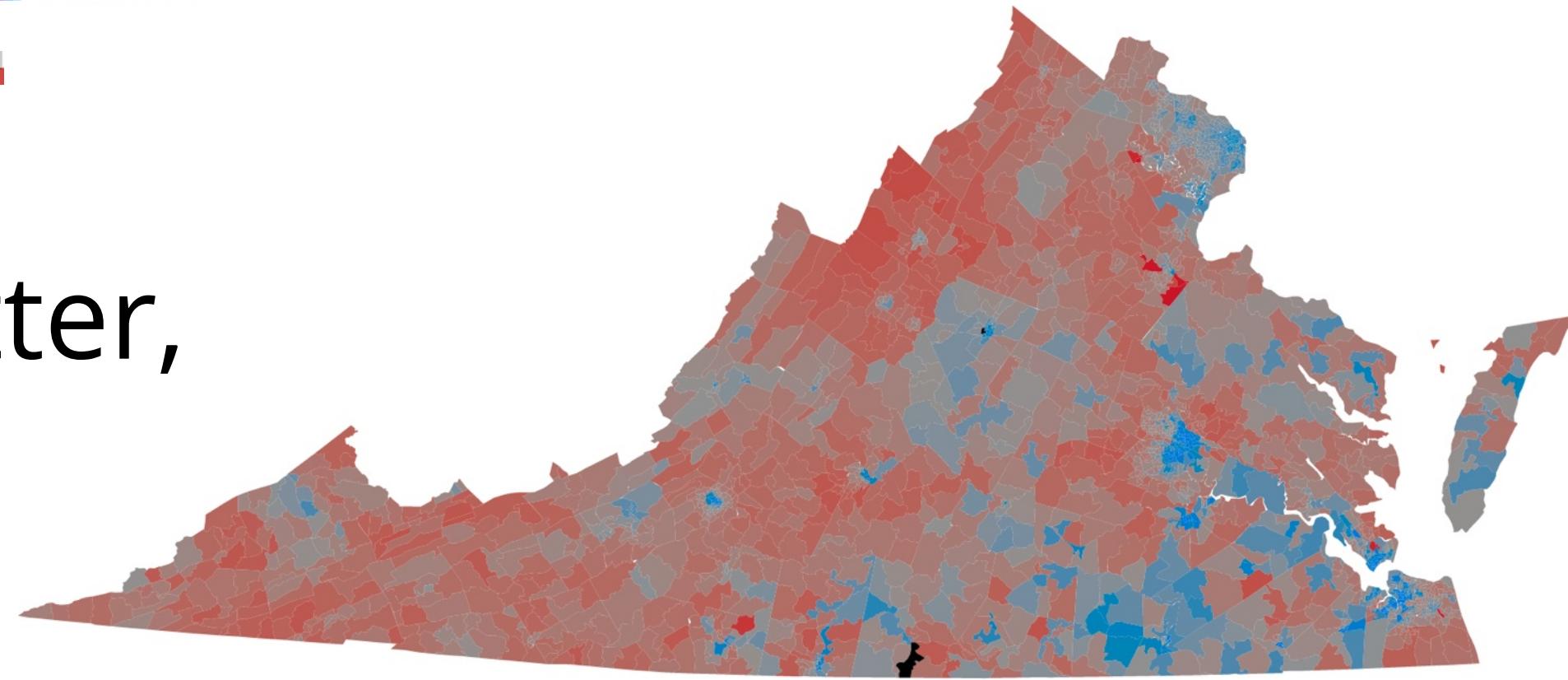
What, exactly, does
this colormap **not**
show about the
data?



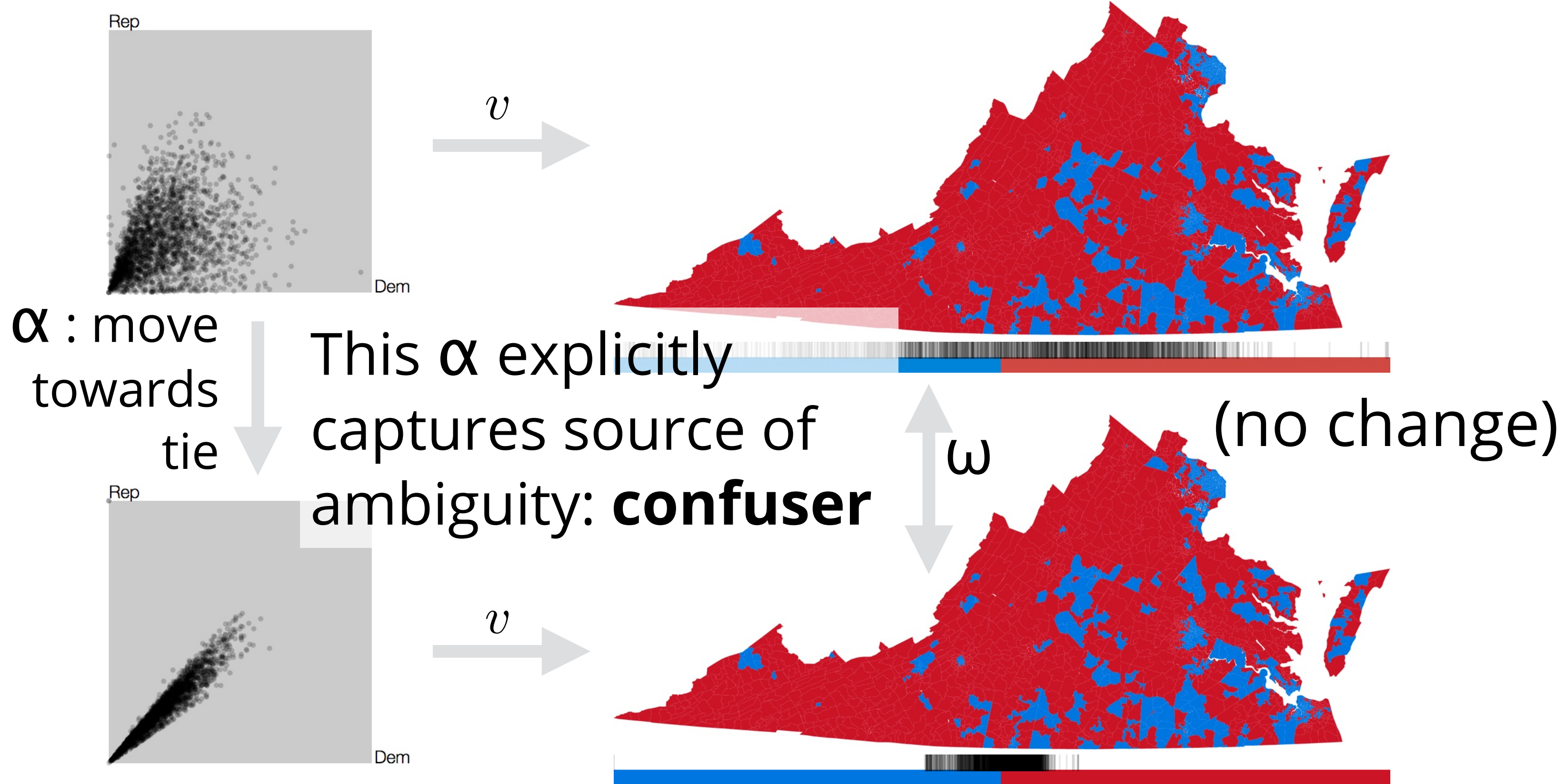
Virginia (US) 2014 Senate Election results, per precinct



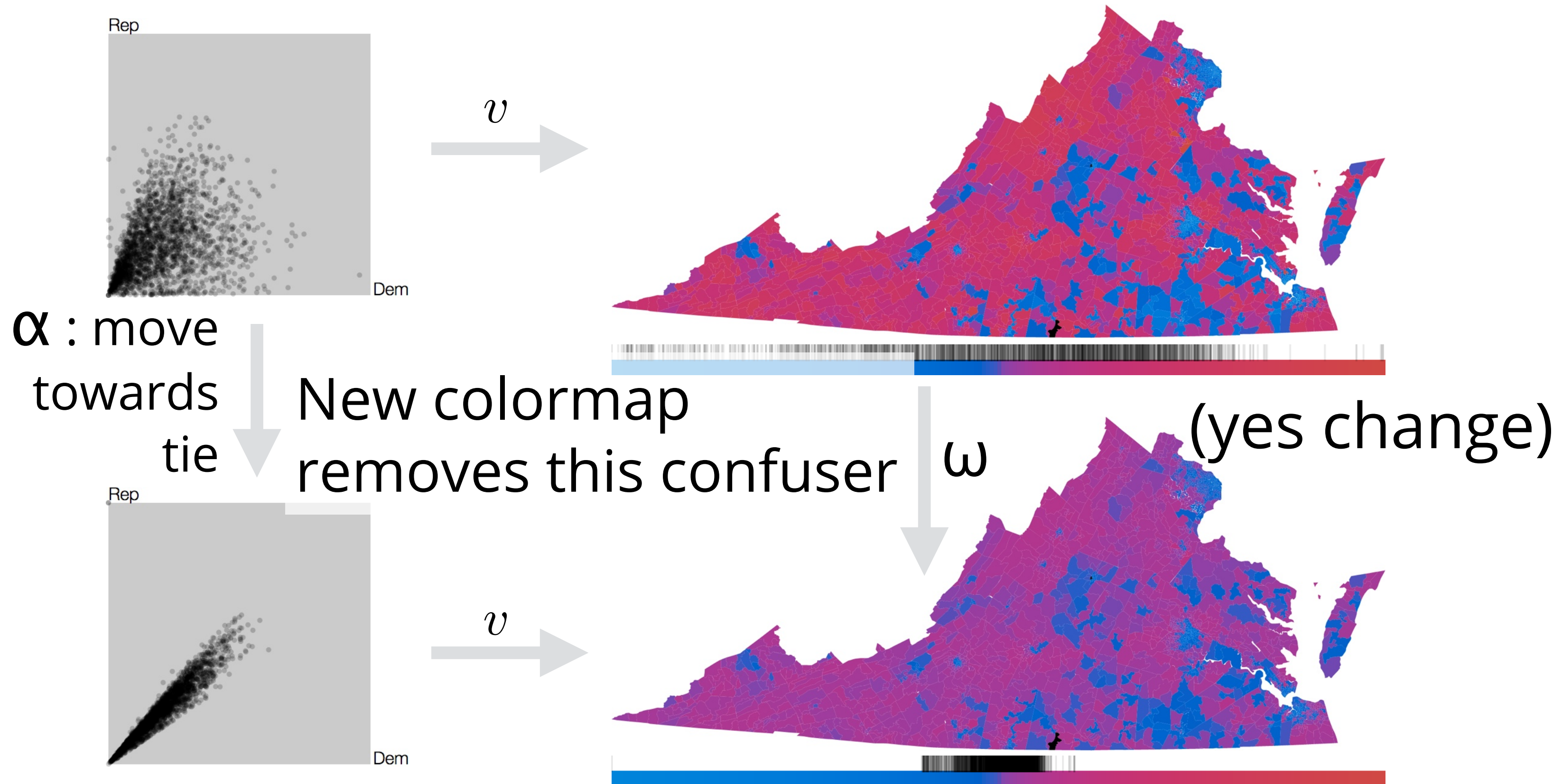
So which one is better,
for what purpose?



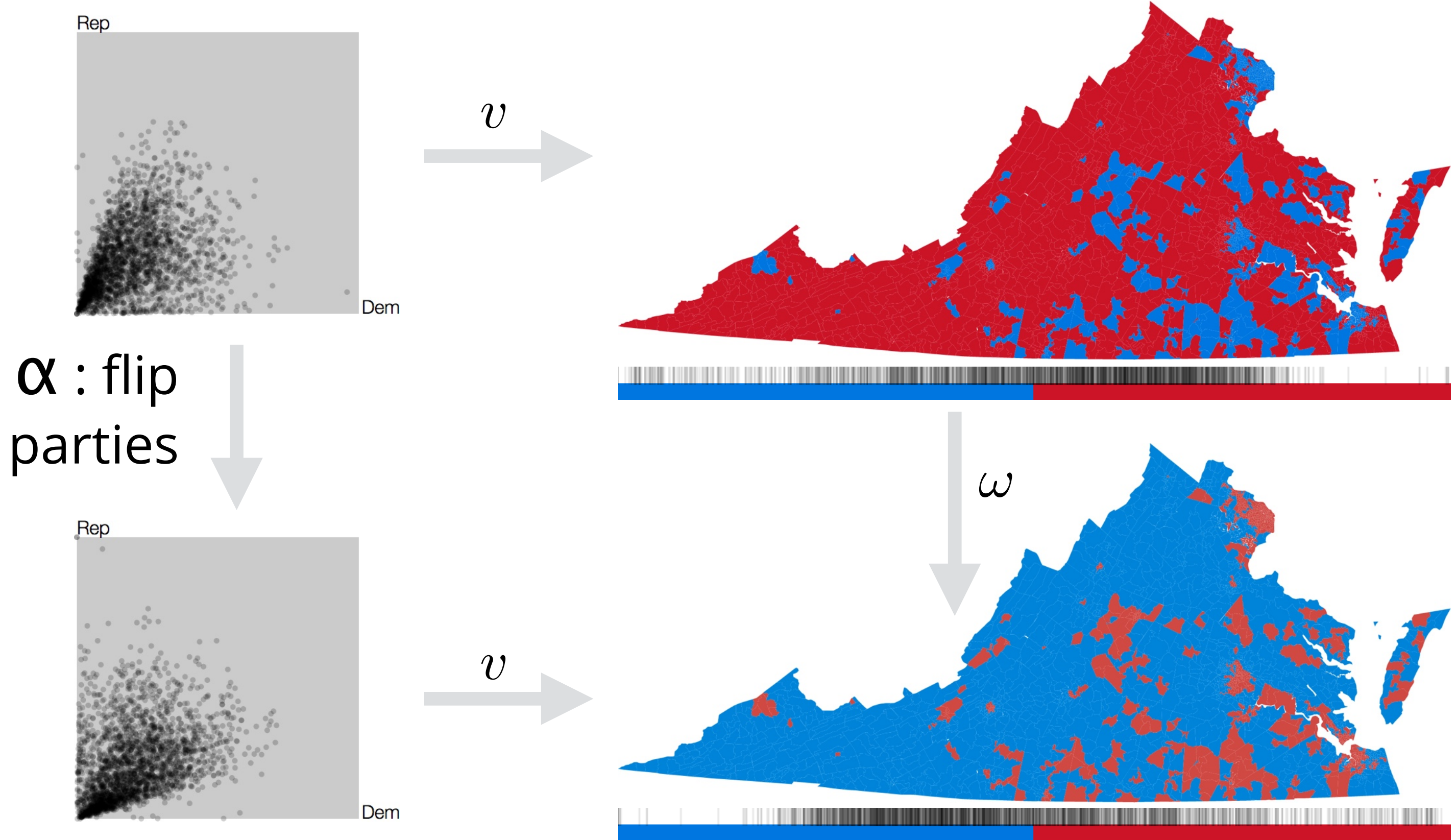
Teach using α s to probe vis properties



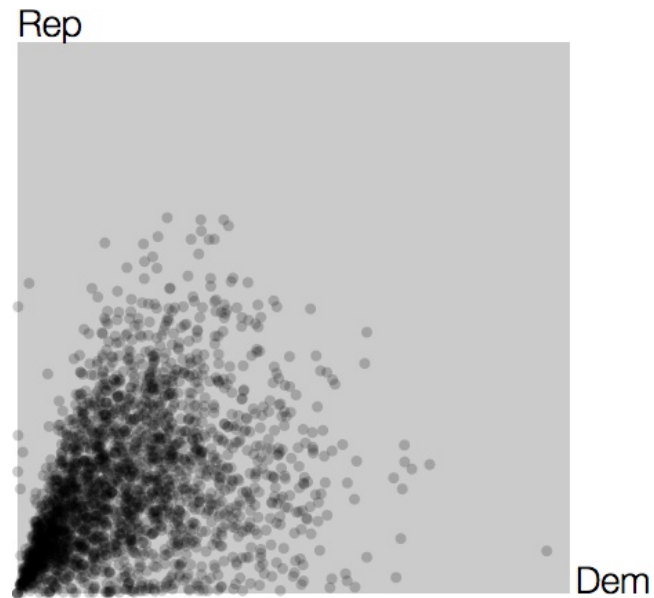
... now with a different visualization



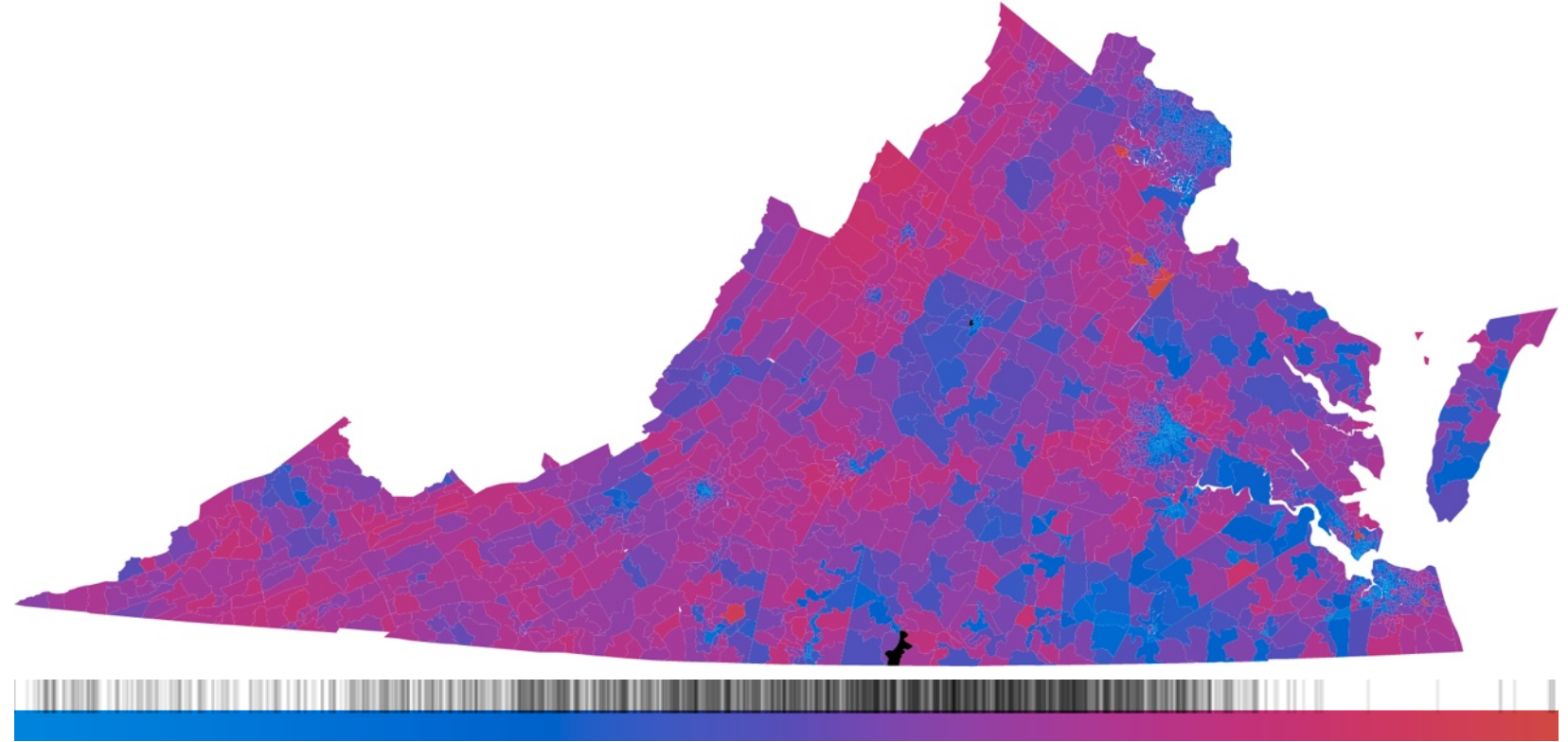
Where was vote evenly split?



Where was vote evenly split??



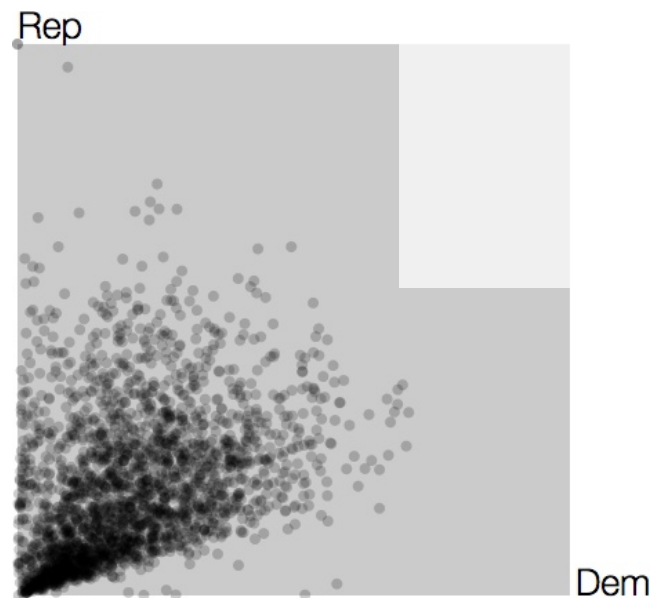
v



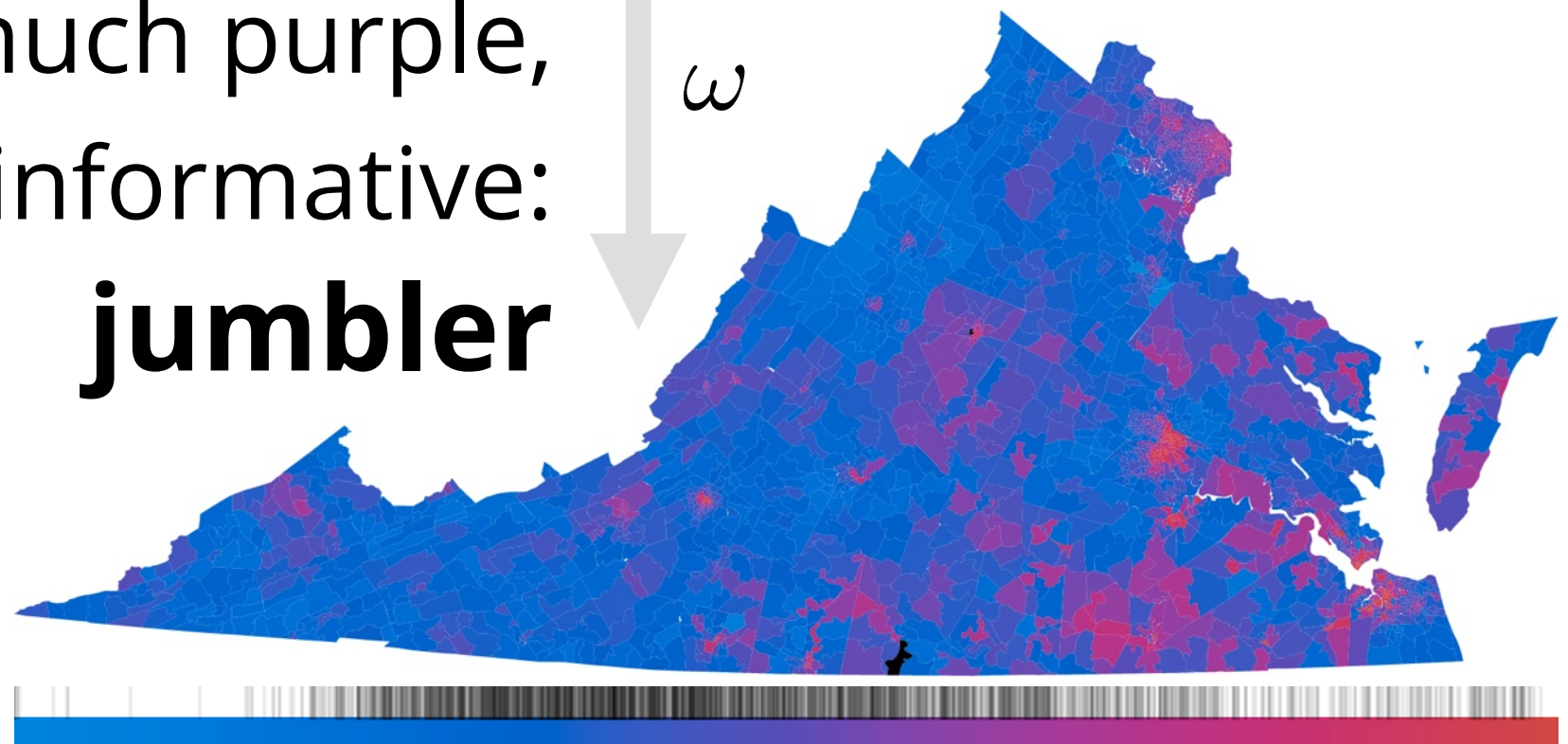
α : flip parties

Yikes - so much purple, not informative:

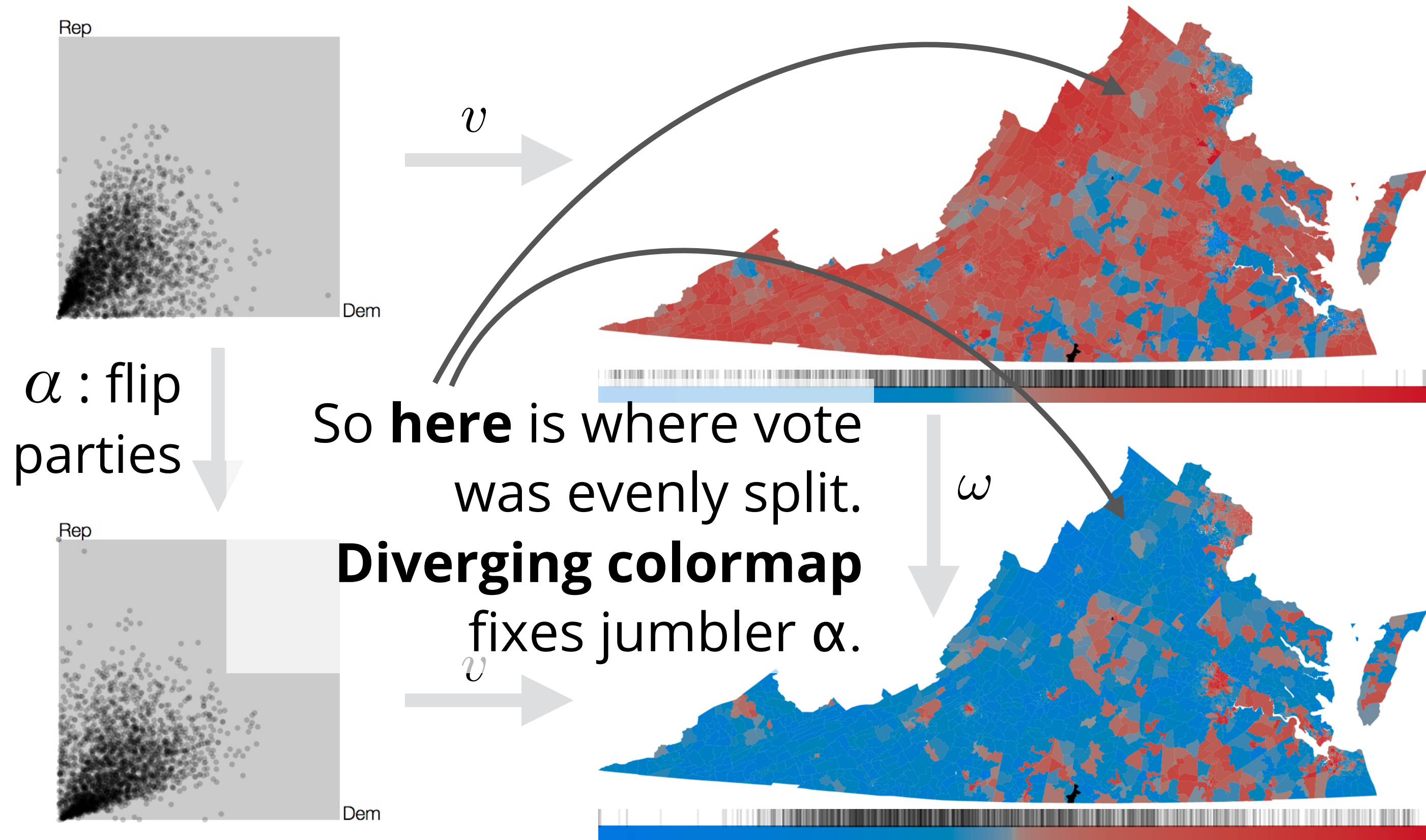
jumbler



v



... now with a different visualization

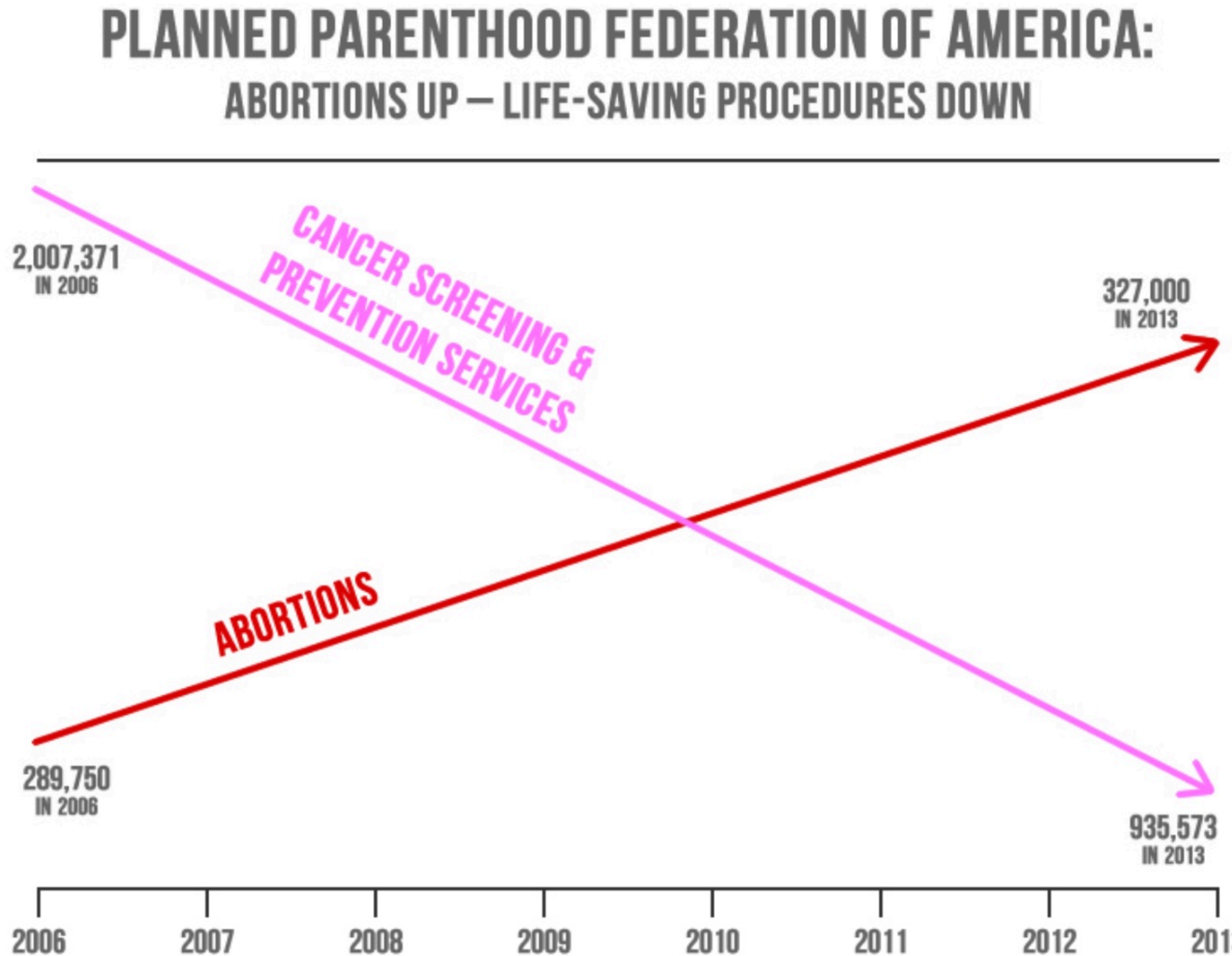


What is bad about this visualization?

<http://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading/>

29 Sept 2015 US
Congressional
hearing on Planned
Parenthood

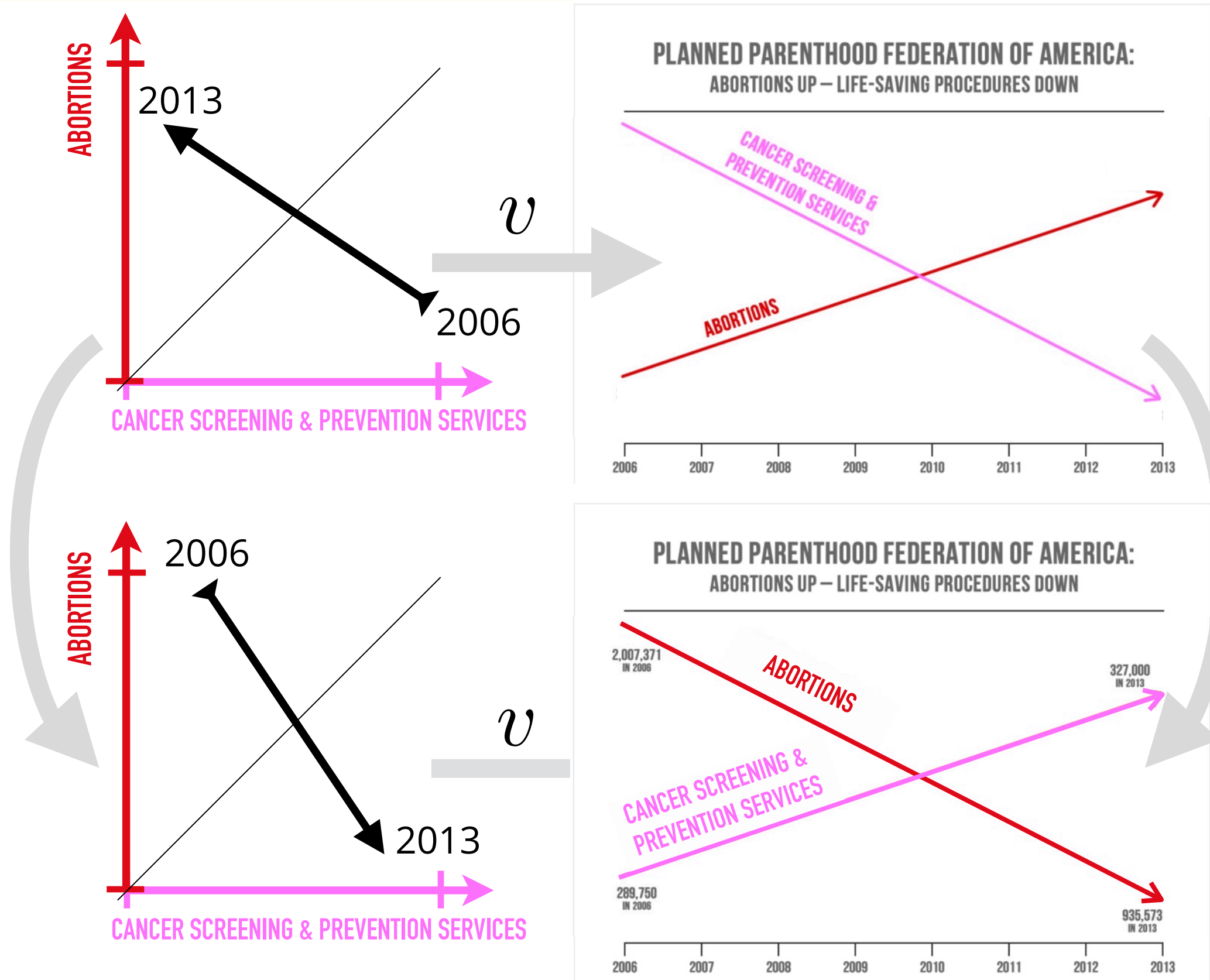
Visualization shown
by Rep. Jason
Chaffetz, (R Utah)



It is the visual symmetry that's misleading

(Strong) visual symmetry ω suggests meaningful data change α , i.e.

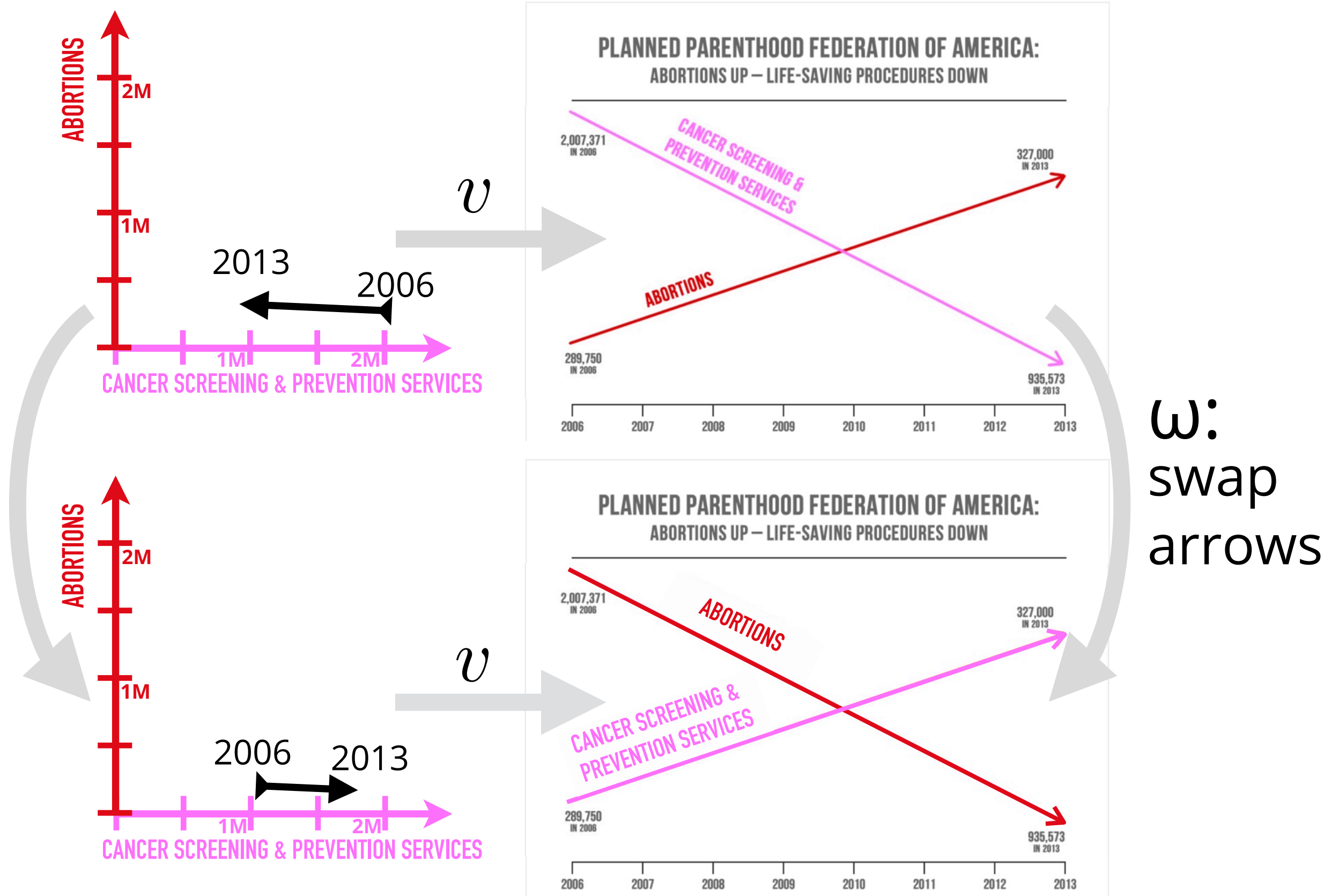
Preserving strong negative correlation, but flipping temporal trend



ω : swap arrows

Actual data change not as interesting

Dual axes
create a
misleader:
an apparent
visual
symmetry ω
not grounded
in interesting
data change
 α



Representation Invariance is old idea

SCIENCE

[Stevens 1946]

Vol. 103, No. 2684

Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Acoustic Laboratory, Harvard University

FOR SEVEN YEARS A COMMITTEE of the British Association for the Advancement of Science debated the problem of measurement. Appointed in 1932 to represent Section A (Mathematical and Physical Sciences) and Section J (Psychology), the committee was instructed to consider and report upon the possibility of "quantitative estimates of sensory events"—meaning simply: Is it possible to measure human sensation? Deliberation led only to disagreement, mainly about what is meant by the term measurement. An interim report in 1938 found one member complaining that his colleagues

by the formal (mathematical) properties of the scales. Furthermore—and this is of great concern to several of the sciences—the statistical manipulations that can legitimately be applied to empirical data depend upon the type of scale against which the data are ordered.

A CLASSIFICATION OF SCALES OF MEASUREMENT

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads

Representation Invariance is old idea

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	<i>Permutation group</i> $x' = f(x)$ <i>f(x) means any one-to-one substitution</i>	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	<i>Isotonic group</i> $x' = f(x)$ <i>f(x) means any monotonic increasing function</i>	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	<i>General linear group</i> $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	<i>Similarity group</i> $x' = ax$	Coefficient of variation

possible hallucinators

e.g. taking median commutes with applying a monotonic function, but taking the mean does not

Like others: pedagogy via critique

Algebraic vis design provides operational vocabulary for **describing a visualization** (confuser, jumbler/misleader, hallucinator)

Also: tensor glyphs and tensor field features

“Tools, not Rules” (George Gopen, The Sense of Structure)

In progress: library of α s that implements

Brehmer & Munzner’s task typology “A Multi-Level

Typology of Abstract Visualization Tasks” TVCG 9(12): 2376-2385 (2013)