(quiz)

(laptops)

Design

SI 649 W20: Information visualization

Matthew Kay, Assistant Professor, School of Information
& Computer Science and Engineering
University of Michigan
Portions of slides adapted from Eytan Adar

This week

Lecture

Design

Lab

Group project proposal critiques

Reminder: Last Altair lab due next week

The extra week is because web stuff is annoying

Use the time!

Remaining labs…

…are design labs, largely in-class.

From here on out, sit with your project group in lab

Design!

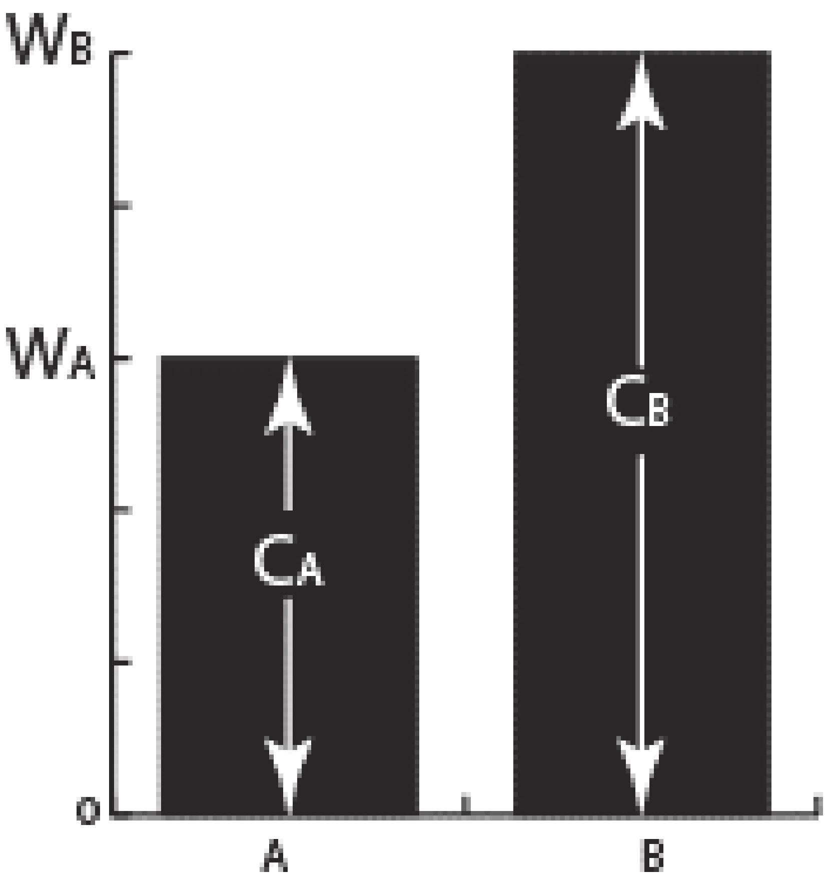
Let's talk Tufte for a bit:
lie factors and data-ink ratios

Lie factor =

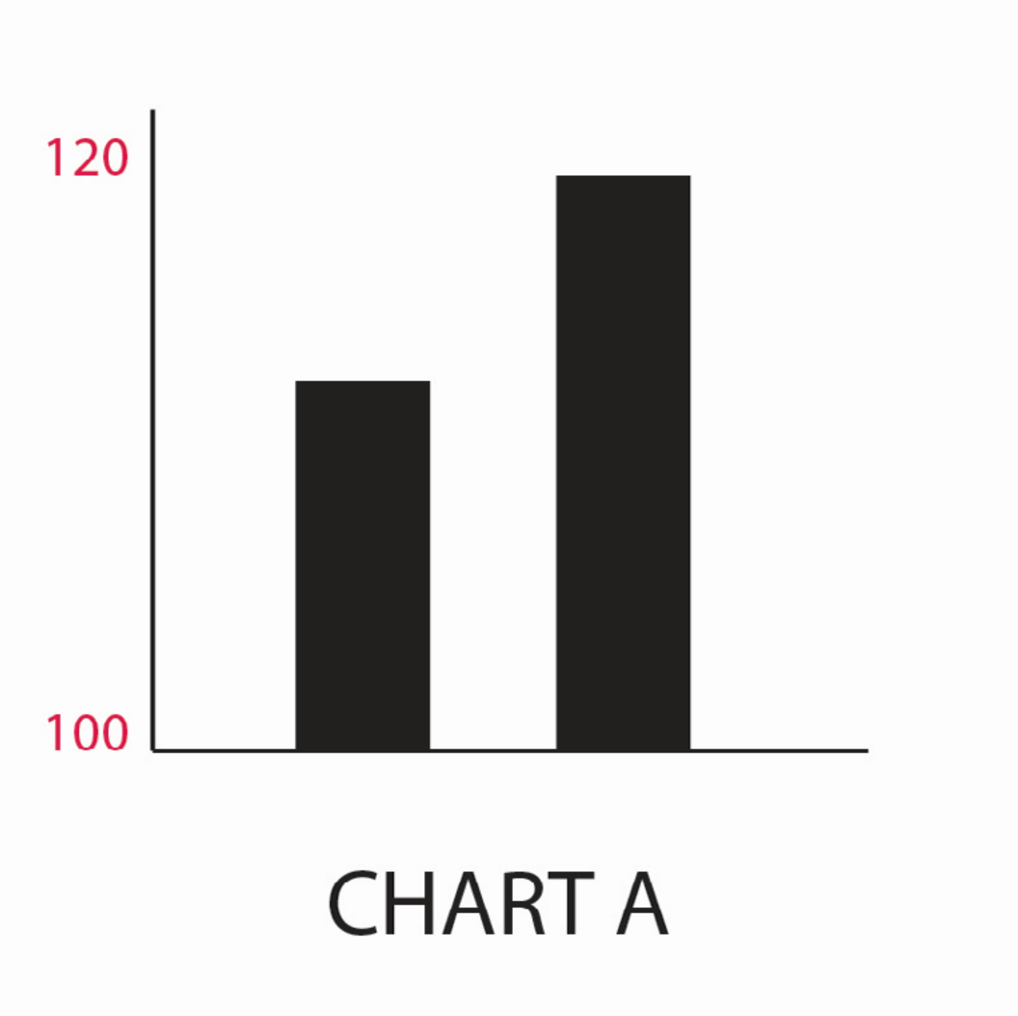
$$\frac{\text{size of effect in graphic}}{\text{size of effect in data}}$$

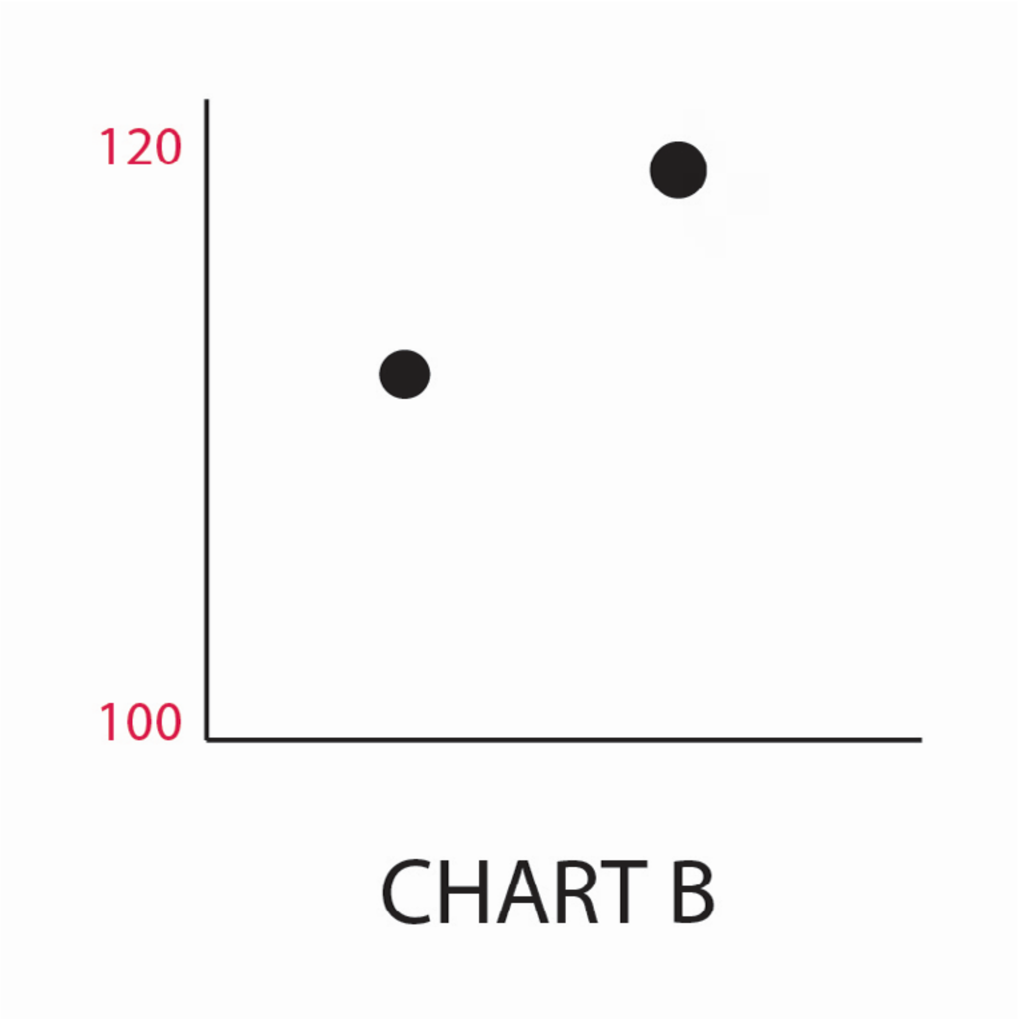
Don’t overthink this:
usually you want it to be ≈ 1

Connects back to effectiveness



Lie factor (quiz)

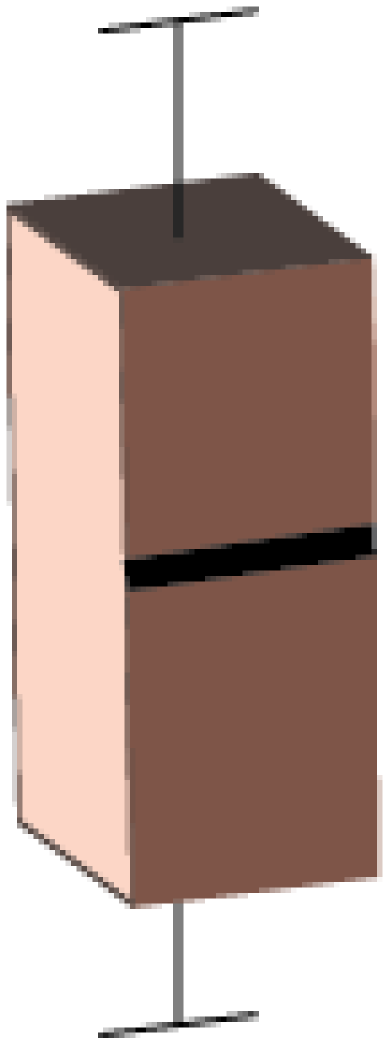




Data-ink ratio

Group activity: maximize the data-ink ratio of this 3d boxplot. E.g.:

1. Remove redundant / non-data ink
2. Refine the design
3. Repeat

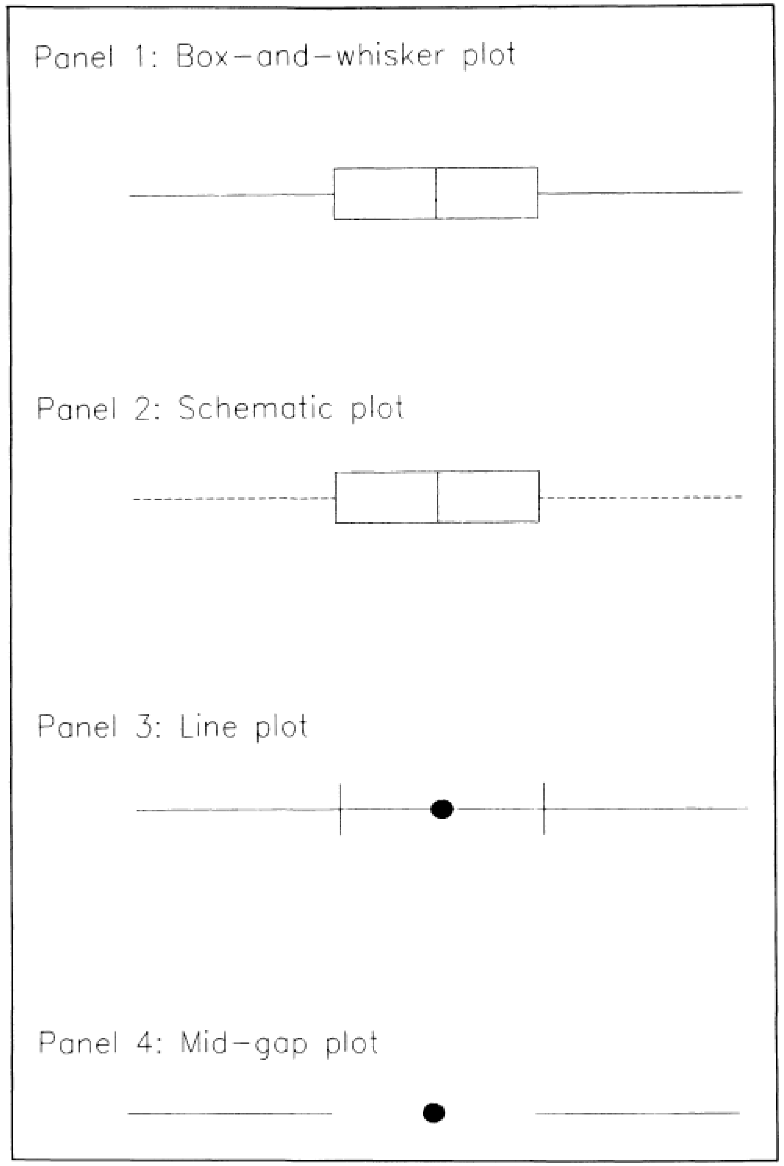


Data-ink ratio: boxplots and mid-gap plots

These have been tested:

The boxplot and lineplot led to more accurate estimates

[http://dx.doi.org/10.3102/10769986016001001]

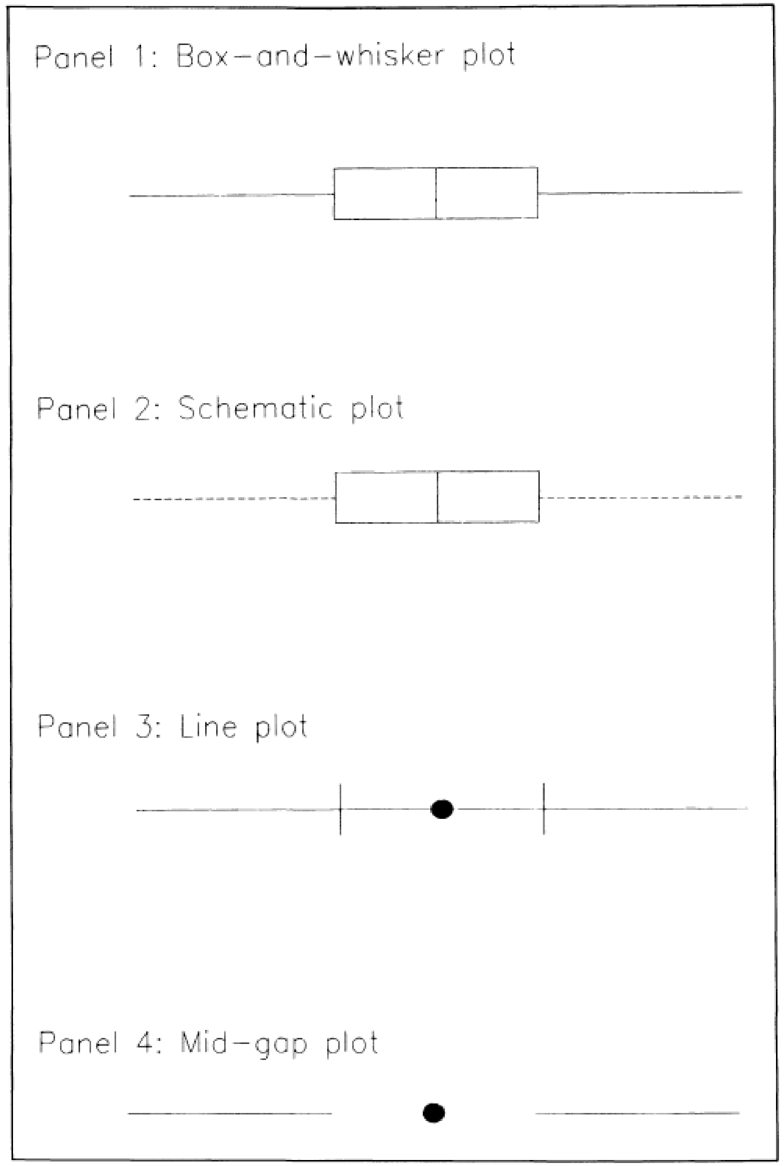


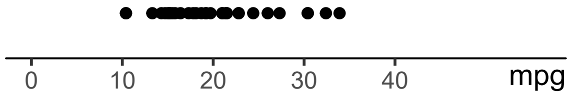
Data-ink ratio: boxplots and mid-gap plots

Limited expressiveness, Tufte-esque or otherwise

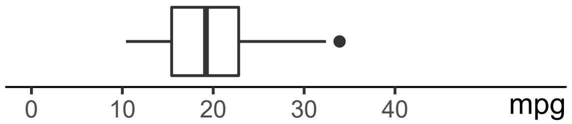
Many other, usually better, options for visualizing distributions...

(Hint: may be helpful for Individual Assignment)

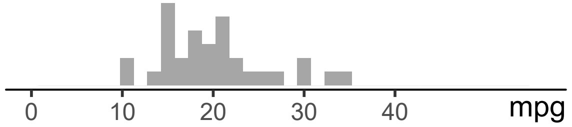




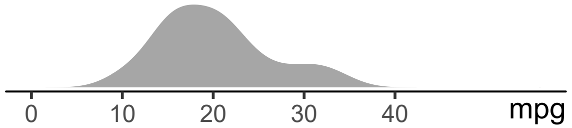
Raw data



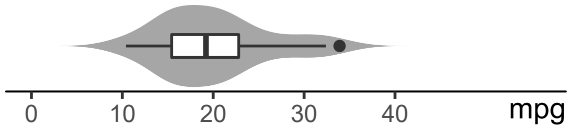
Boxplot



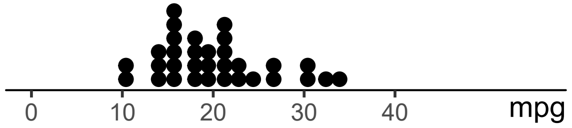
Histogram



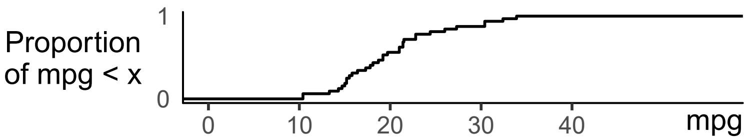
Density plot



Violin plot (or similar: “bean plot”)



(Wilkinson) dotplot (or similar: “beeswarm”)



Cumulative distribution function (CDF)

Is the data-ink ratio a good design guideline?

Is the data-ink ratio a good design guideline?

It is sensitive but not specific
(lots of false positives, hard to tell when to stop)

Better to rely on principles like effectiveness and
avoiding ambiguity

The data-ink ratio is better thought of as a minimalist aesthetic guideline that you should feel free to disagree with (I often do!)

Tufte is a minimalist…

Okay okay,
Lie factors should be about 1,
Don't take data-ink ratios too seriously,
What should I do?

Some rough design guidelines*

1. (Match effectiveness with importance)
2. Avoid ambiguity
3. Locality is king / eyes beat memory
4. Establish viewing order
5. Layer, layer, layer
6. When in doubt, grid
7. Treat visual attributes like adjectives

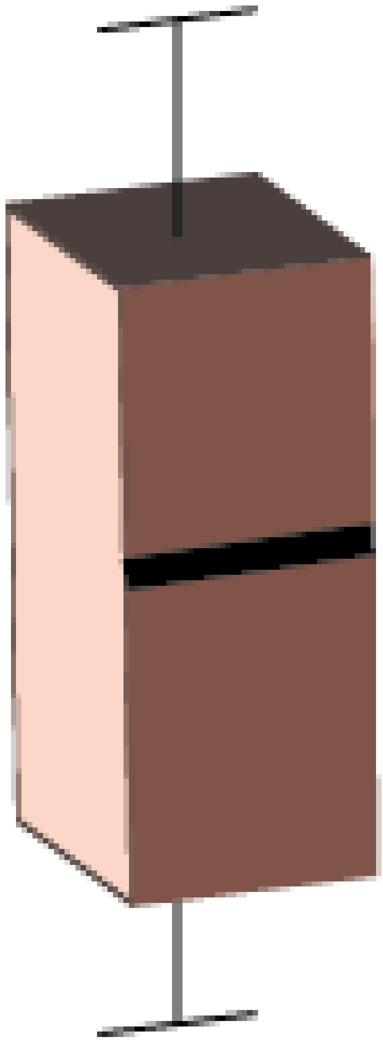
* These guidelines are drawn largely from my experience + personal preferences + the literature. Design is messy, these are not perfect, others will disagree with me, etc. *Caveat emptor.*

1. (Match effectiveness with importance)

(see also, Perception lecture)

2. Avoid ambiguity

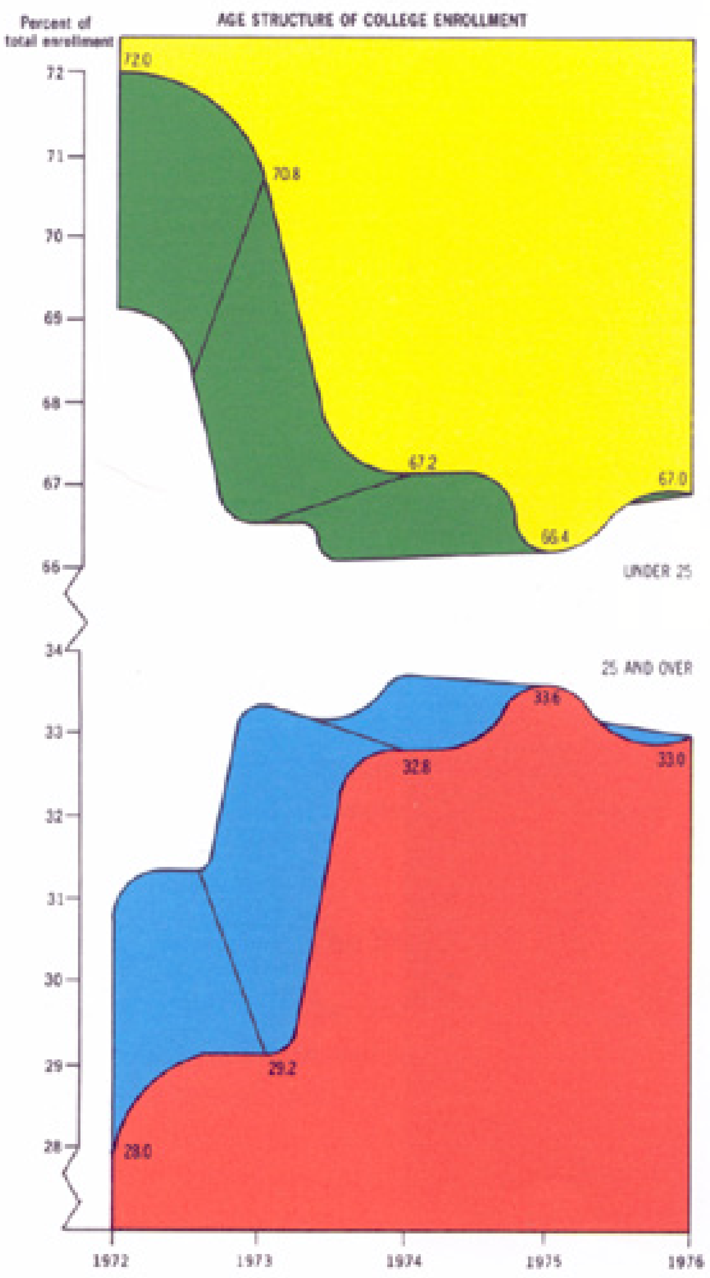
Most of the correct changes you make due to the data-ink ratio amount to this



2. Avoid ambiguity

Marks should not have multiple reasonable interpretations

If it looks like it could come from data, it should come from data



3. Locality is king / eyes beat memory

No:

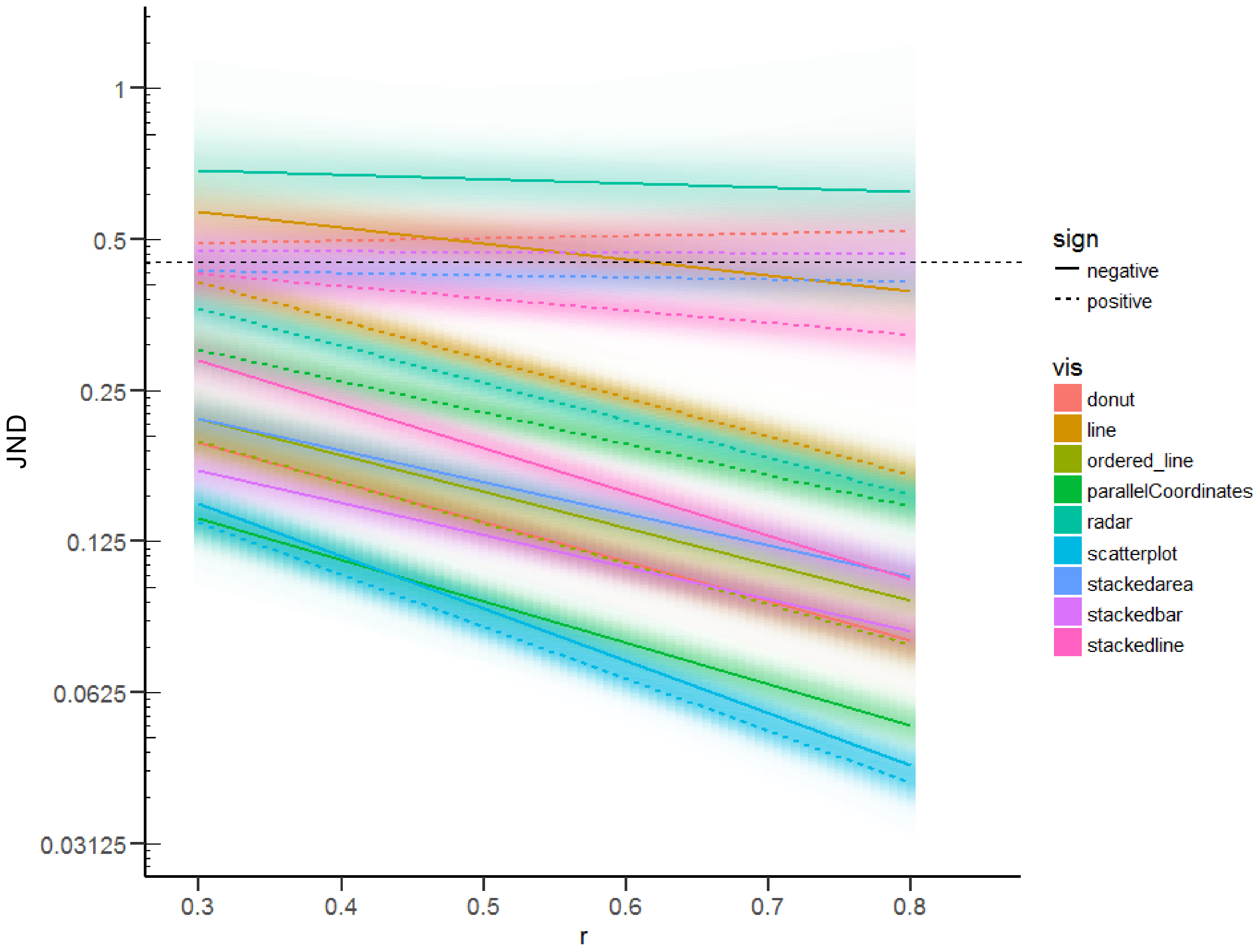
Thing

Information I need
to understand thing

Yes:

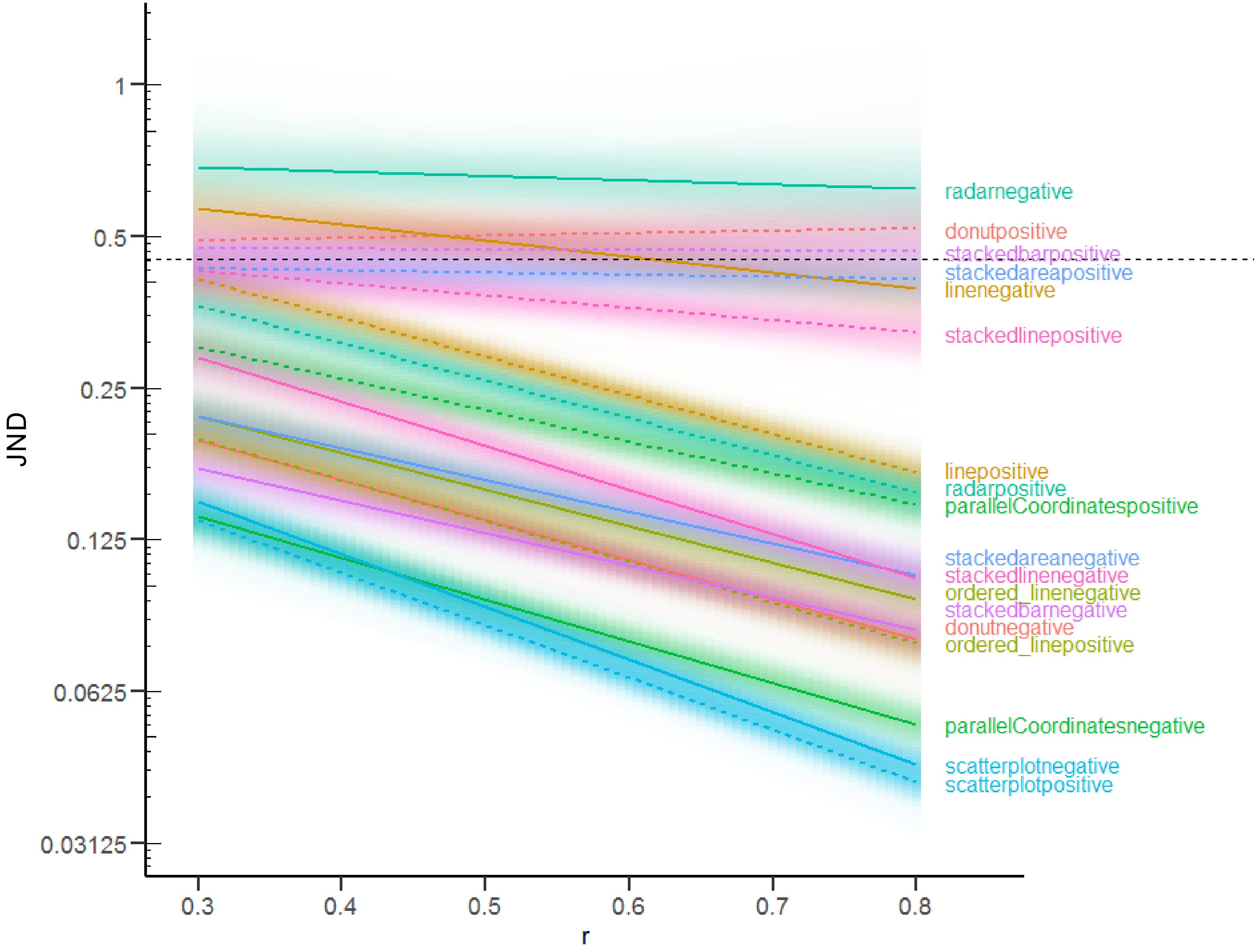
Thing ↔ Information I need
to understand thing

No



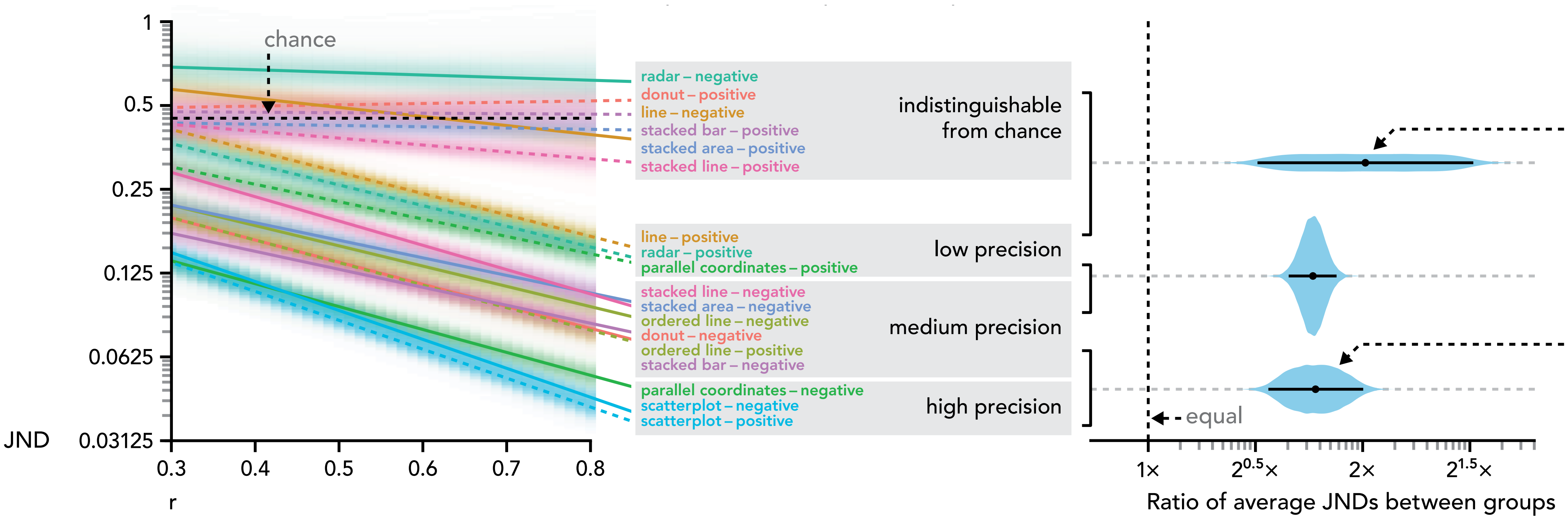
JND
1
0.5
0.25
0.125
0.0625
0.03125
0.3
0.4
0.5
0.6
0.7
0.8
r
sign
negative
positive
vis
donut
line
ordered_line
parallelCoordinates
radar
scatterplot
stackedarea
stackedbar
stackedline

Yes



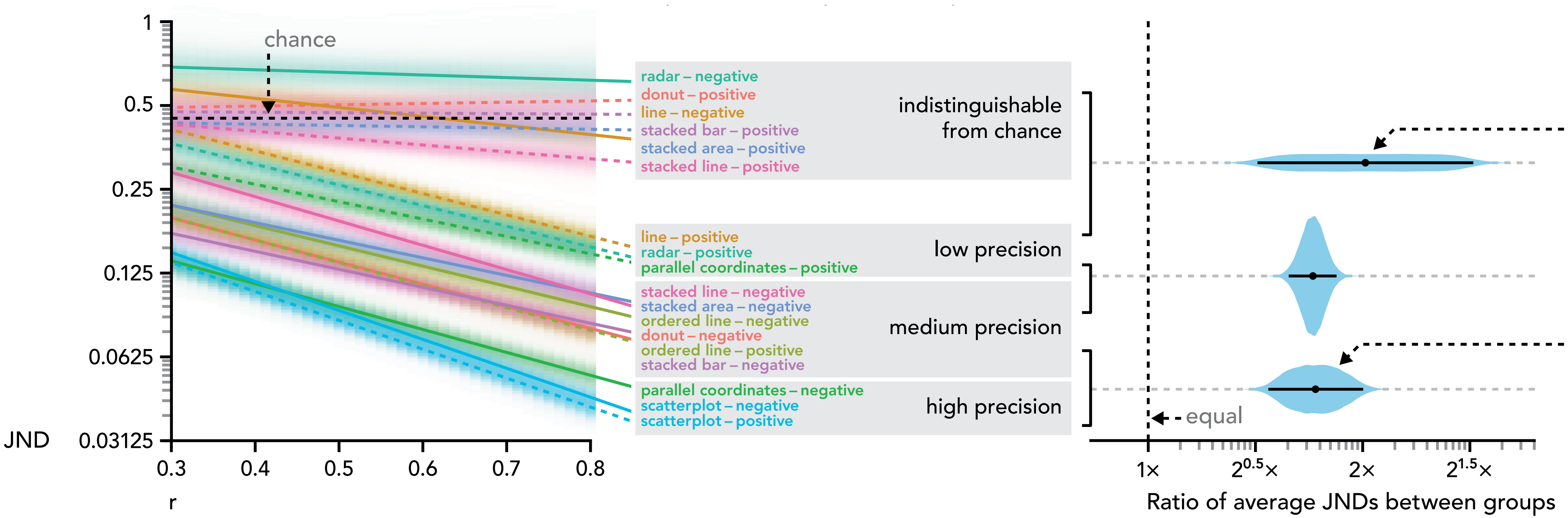
JND
r
1
0.5
0.25
0.125
0.0625
0.03125
0.3
0.4
0.5
0.6
0.7
0.8
radarnegative
donutpositive
stackedbarpositive
stackedareapositive
linenegative
stackedlinepositive
linepositive
radarpositive
parallelCoordinatespositive
stackedareanegative
stackedlinenegative
ordered_linenegative
stackedbarnegative
donutnegative
ordered_linepositive
parallelCoordinatesnegative
scatterplotnegative
scatterplotpositive

No



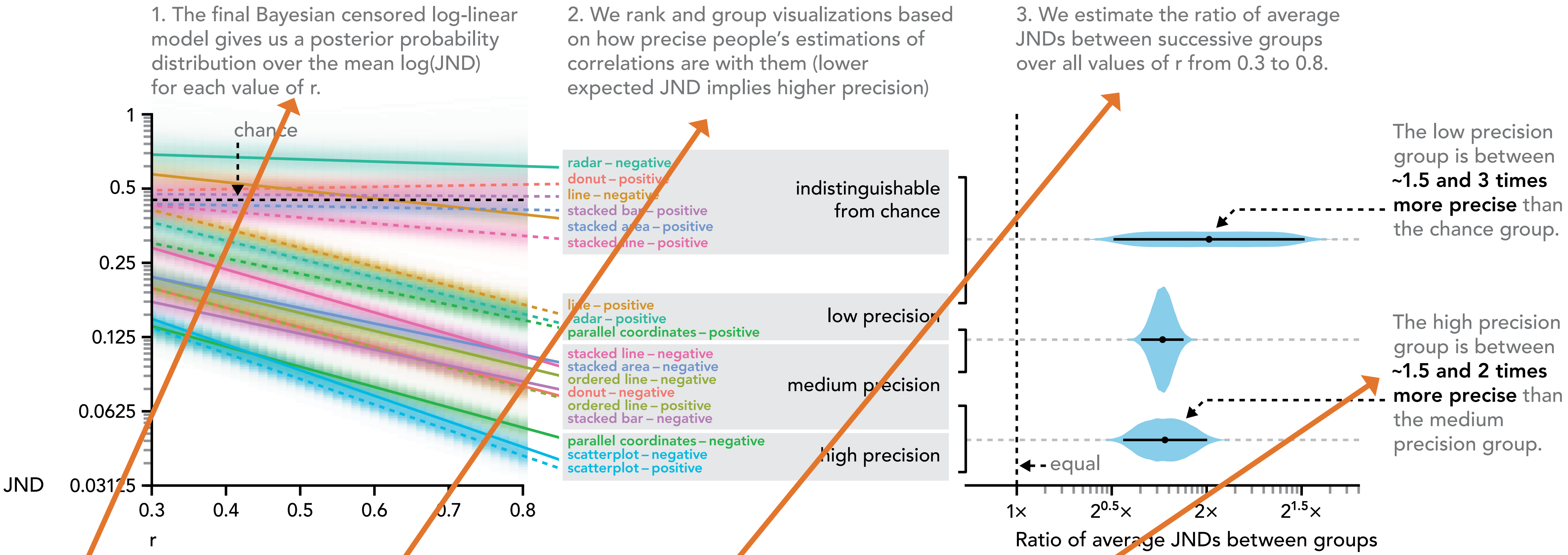
The left panel shows the Bayesian censored log-linear model, which gives us a posterior probability distribution over the mean log(JND) for each value of r. In the center panel we rank and group visualizations based on how precise estimations of correlations are with them (lower expected JND implies higher precision). In the right panel we estimate the ratio of average JNDs between succesive groups over all values of r from 0.3 to 0.8. The low precision group is between ~1.5 and 3 times more precise than the chance group. The high precision group is between ~1.5 and 2 times more precise than the medium precision group.

No



The left panel shows the Bayesian censored log-linear model, which gives us a posterior probability distribution over the mean log(JND) for each value of r. In the center panel we rank and group visualizations based on how precise estimations of correlations are with them (lower expected JND implies higher precision). In the right panel we estimate the ratio of average JNDs between succesive groups over all values of r from 0.3 to 0.8. The low precision group is between ~1.5 and 3 times more precise than the chance group. The high precision group is between ~1.5 and 2 times more precise than the medium precision group.

Yes



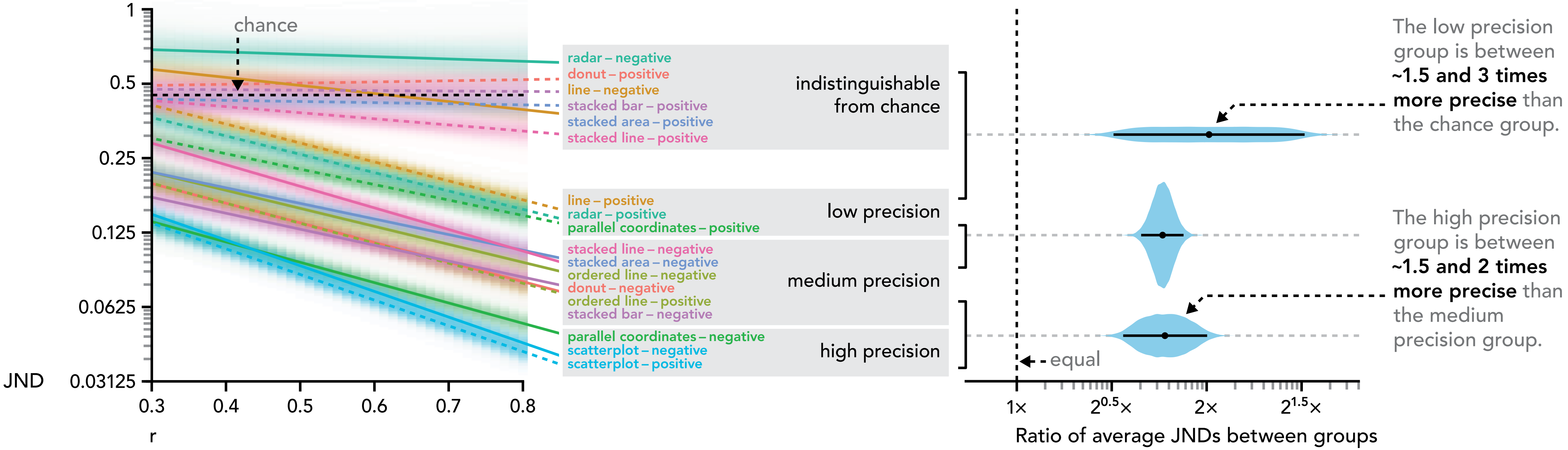
The left panel shows the Bayesian censored log-linear model, which gives us a posterior probability distribution over the mean log(JND) for each value of r. In the center panel we rank and group visualizations based on how precise estimations of correlations are with them (lower expected JND implies higher precision). In the right panel we estimate the ratio of average JNDs between succesive groups over all values of r from 0.3 to 0.8. The low precision group is between ~1.5 and 3 times more precise than the chance group. The high precision group is between ~1.5 and 2 times more precise than the medium precision group.

Yes

1. The final Bayesian censored log-linear model gives us a posterior probability distribution over the mean log(JND) for each value of r.

2. We rank and group visualizations based on how precise people's estimations of correlations are with them (lower expected JND implies higher precision)

3. We estimate the ratio of average JNDs between successive groups over all values of r from 0.3 to 0.8.



3. Locality is king / eyes beat memory

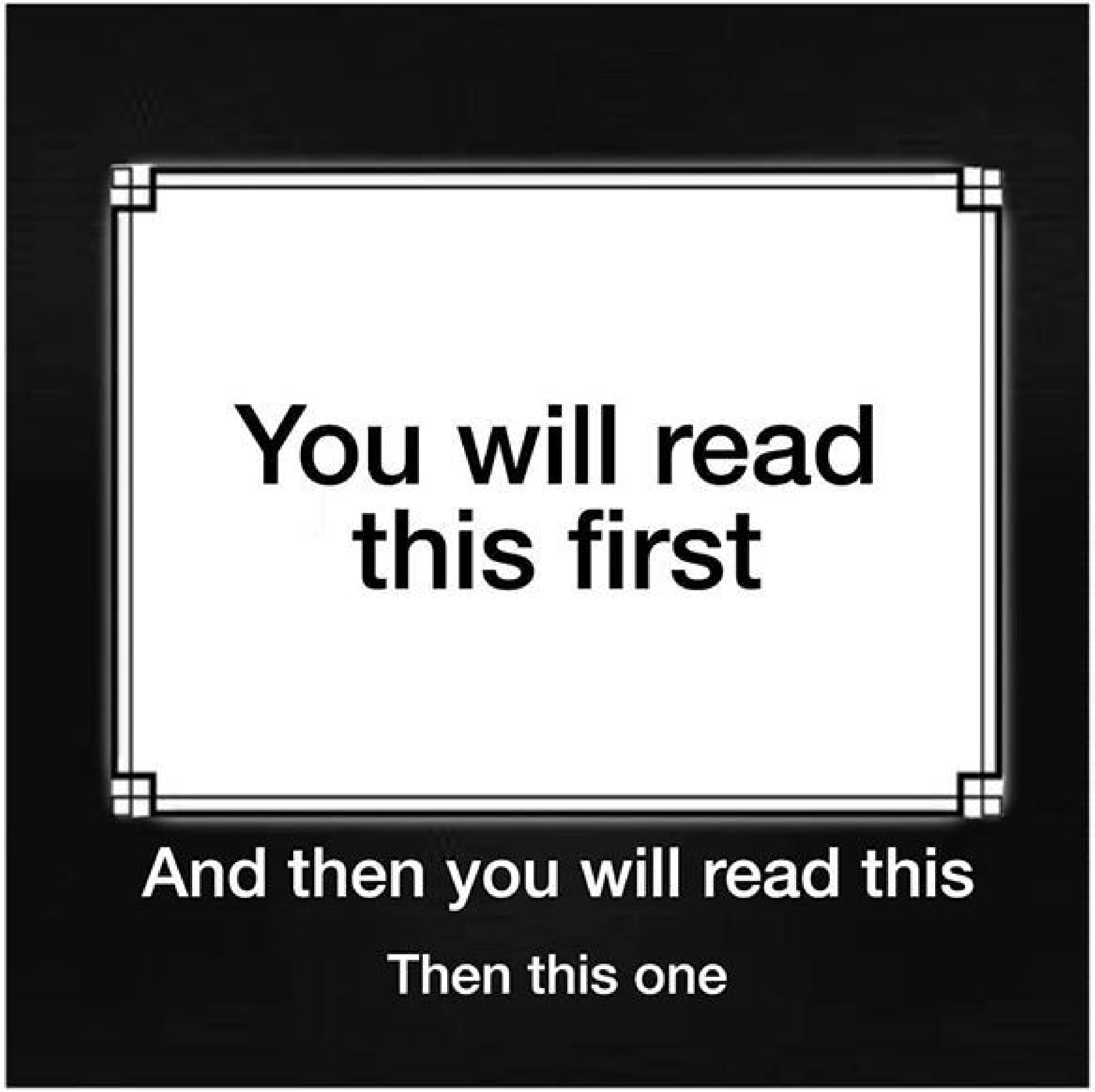
Implies animation / interaction can make things worse (though they are often useful)

E.g.: If the comparison you want is in a tooltip, you have to remember values to make comparisons

Count lookups!

4. Establish viewing order

And you will read this at the end



4. Establish viewing order

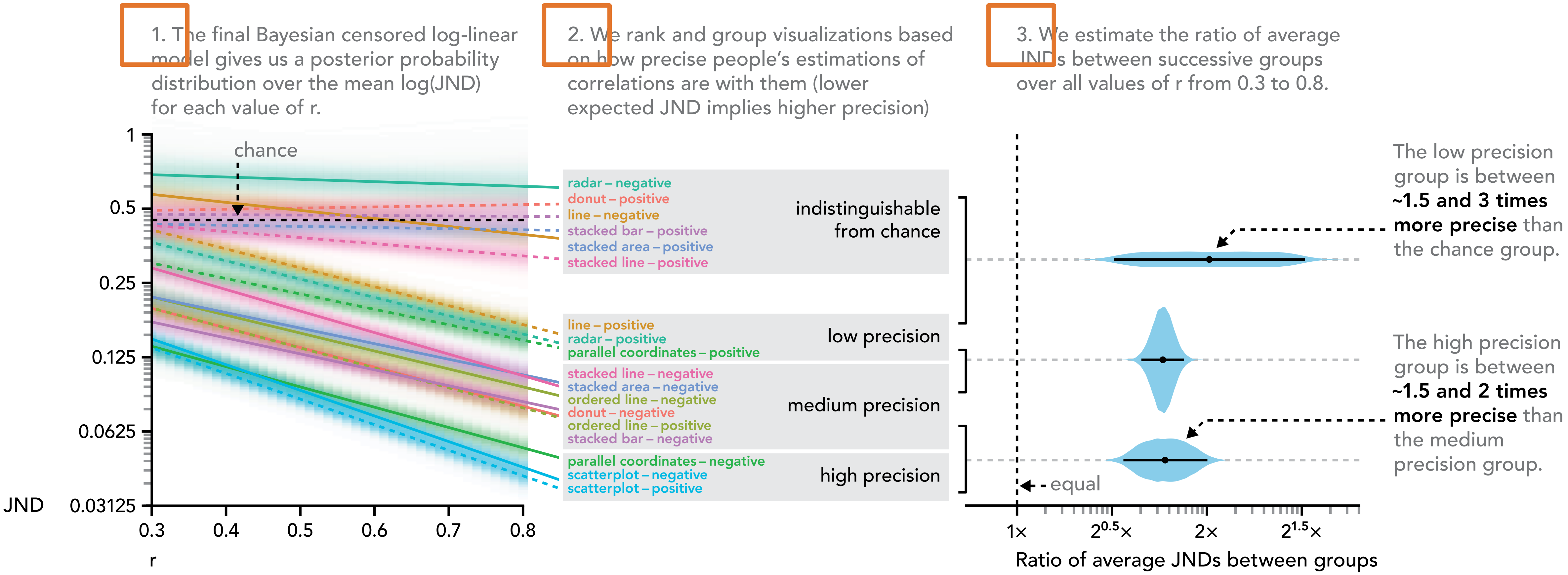
Know where your audience will look first, second.

Remember salience from perception.

Think like a movie director. Are you telling a story?

4. Establish viewing order

Can be as simple as some numbers…

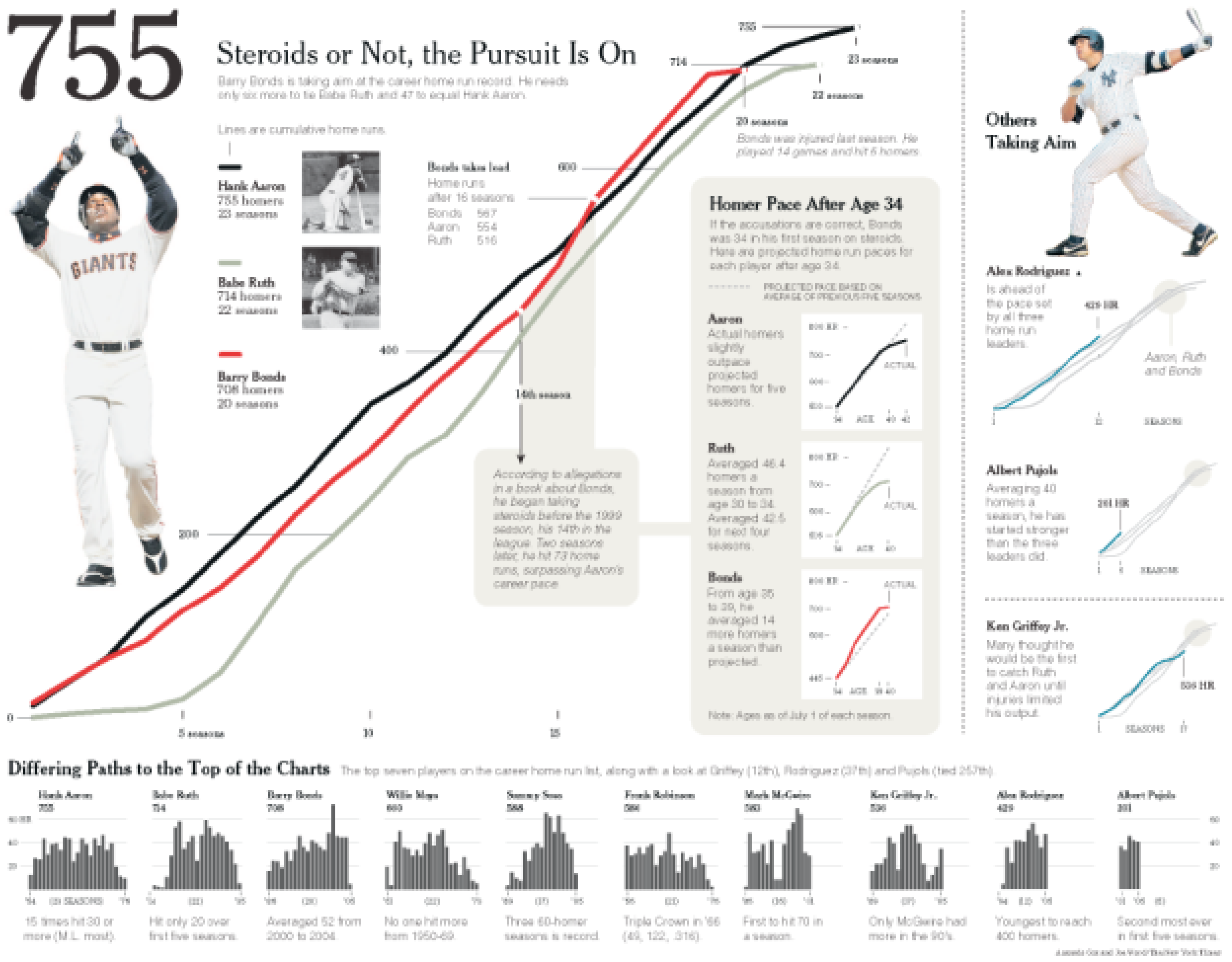


4. Establish viewing order

Or more complex, relying on salience, other visual cues, viewer expectations (maybe)

4. Establish viewing order

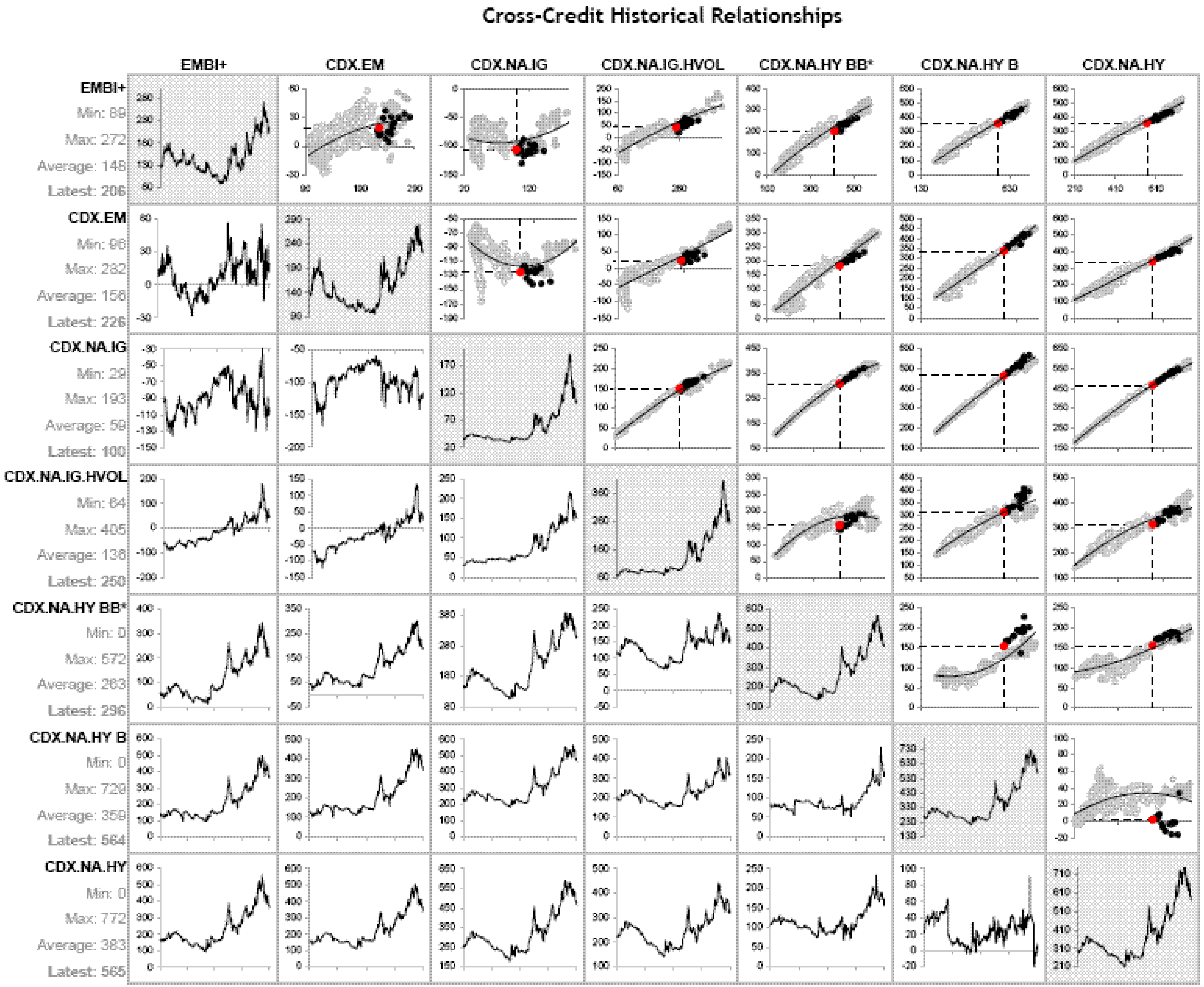
Or more complex, relying on salience, other visual cues, viewer expectations (maybe)



5. Layer, layer, layer, layer

Design for reading at different levels of detail

Pre-attentive attributes help

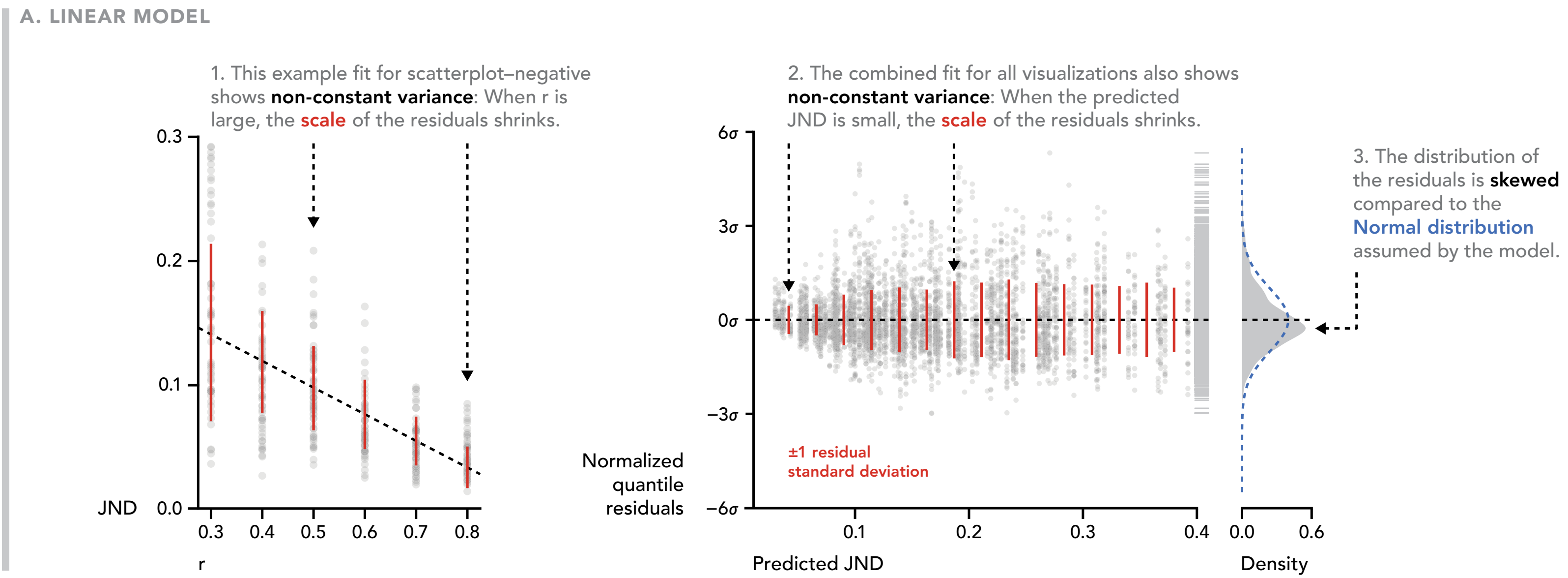


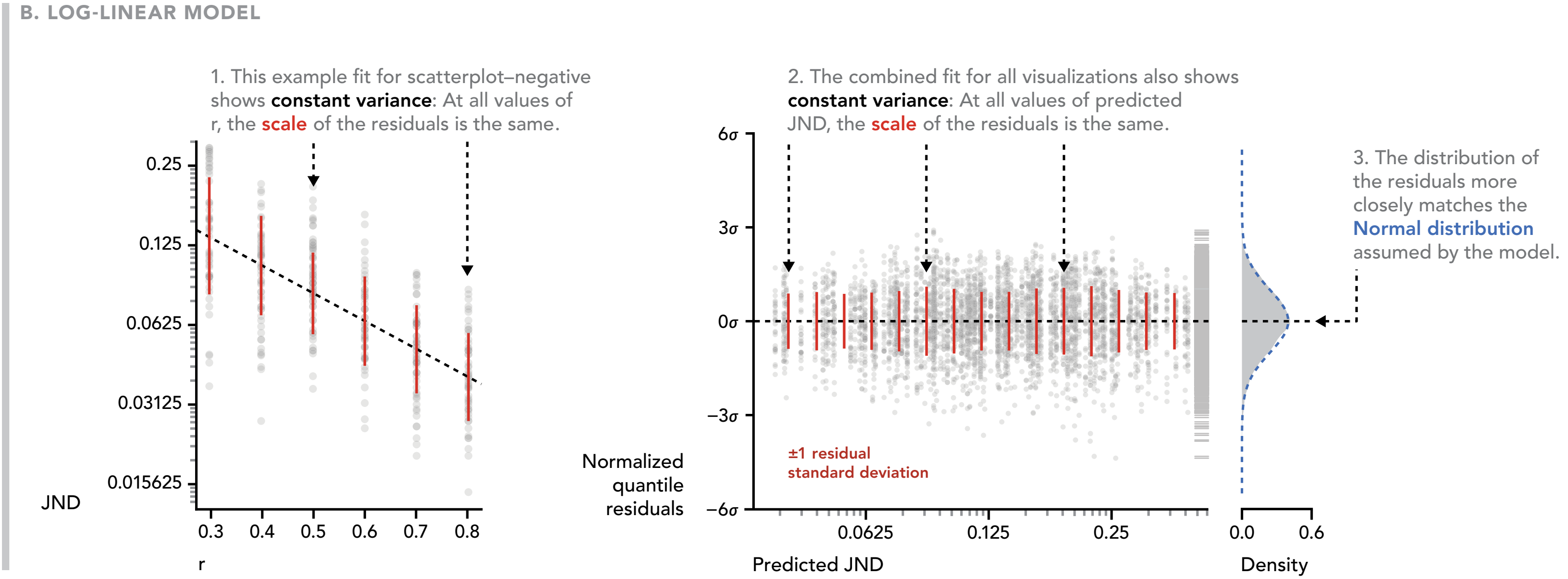
Note: * Spread data for this index does not include the roll to the latest contract due to significant roll effect

2y history shown

6. When in doubt, grid

And get **synchronized axes** as a bonus



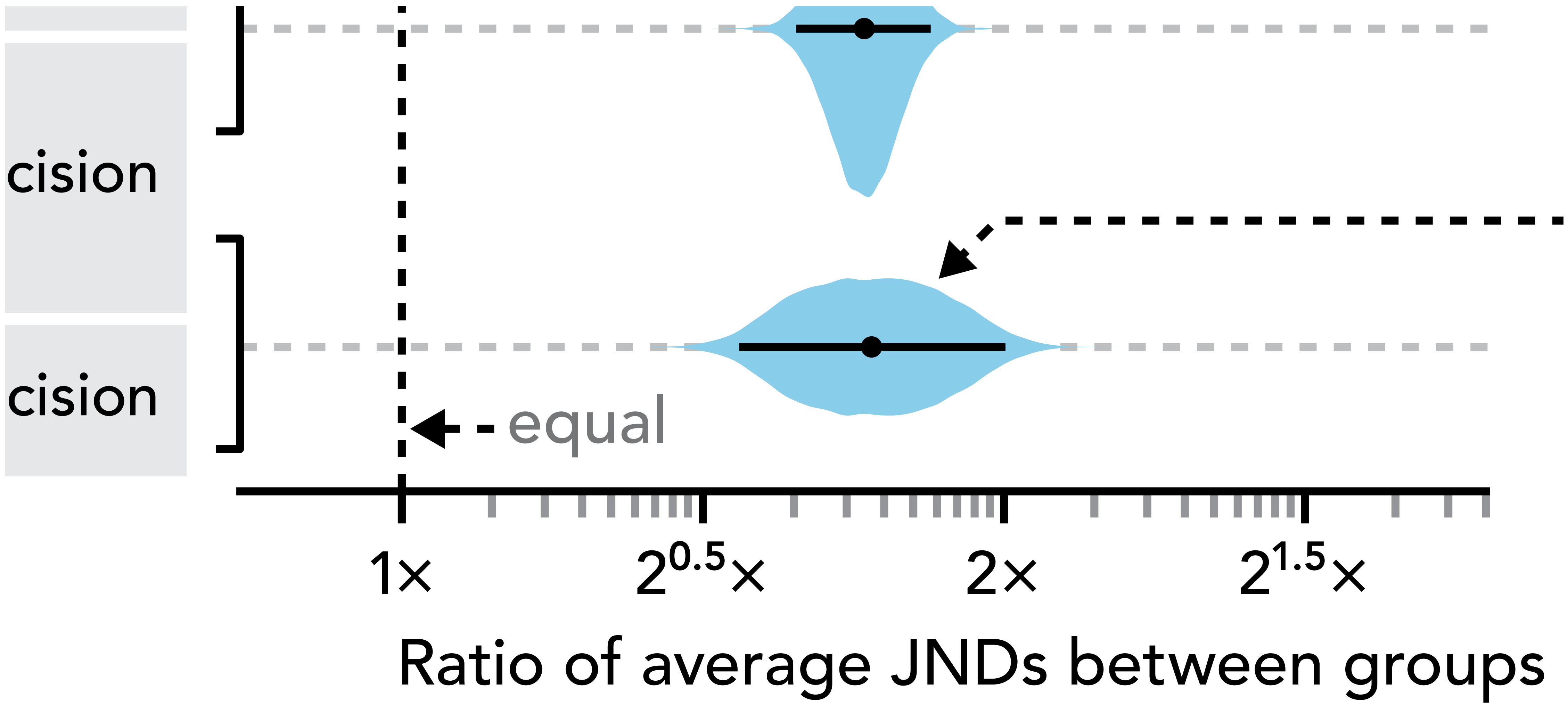


7. Treat visual attributes like adjectives

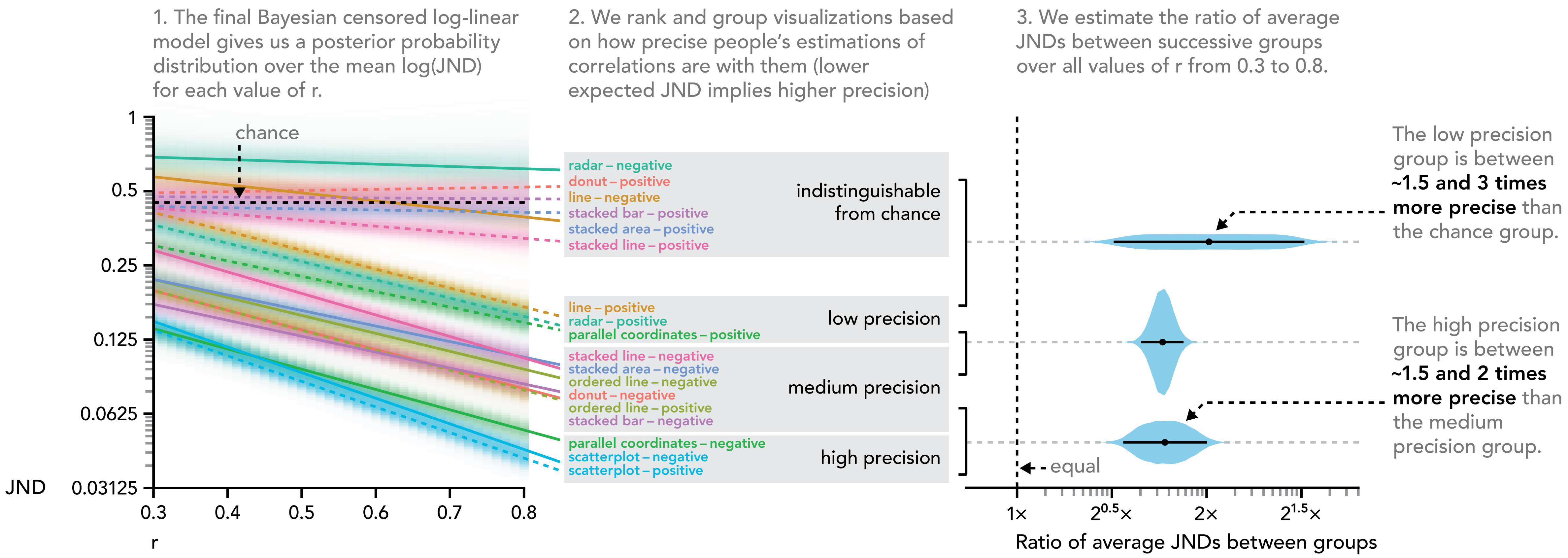
Don’t use three attributes (size, color, shape, ...) to create emphasis where one or two will do.

The very tall building is very extremely tall.

(7b. Obey the pen)



(7b. Obey the pen)



(7b. Obey the pen)

Even visual texture is pleasing

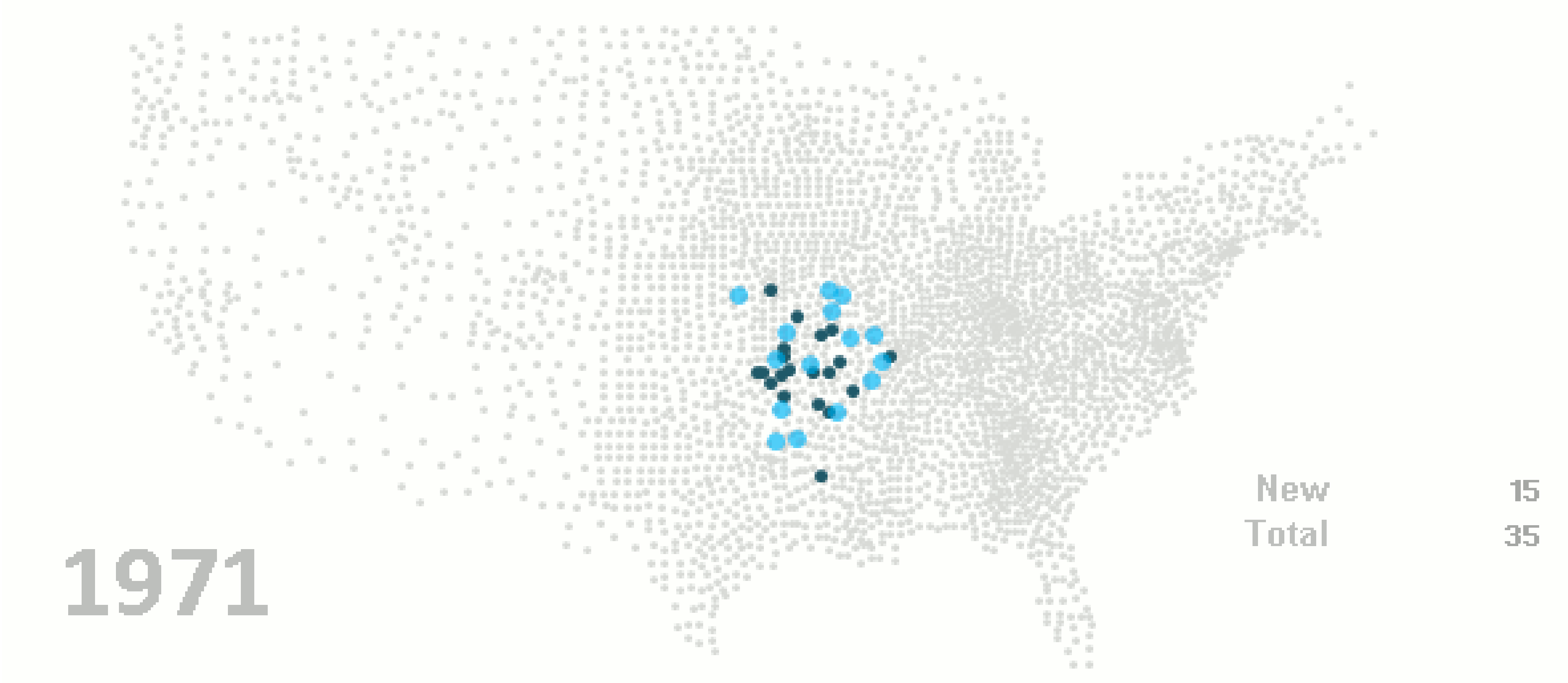
Also makes it easier to create visual hierarchy and call out something important when you need to

Some rough design guidelines*

1. (Match effectiveness with importance)
2. Avoid ambiguity
3. Locality is king / eyes beat memory
4. Establish viewing order
5. Layer, layer, layer
6. When in doubt, grid
7. Treat visual attributes like adjectives

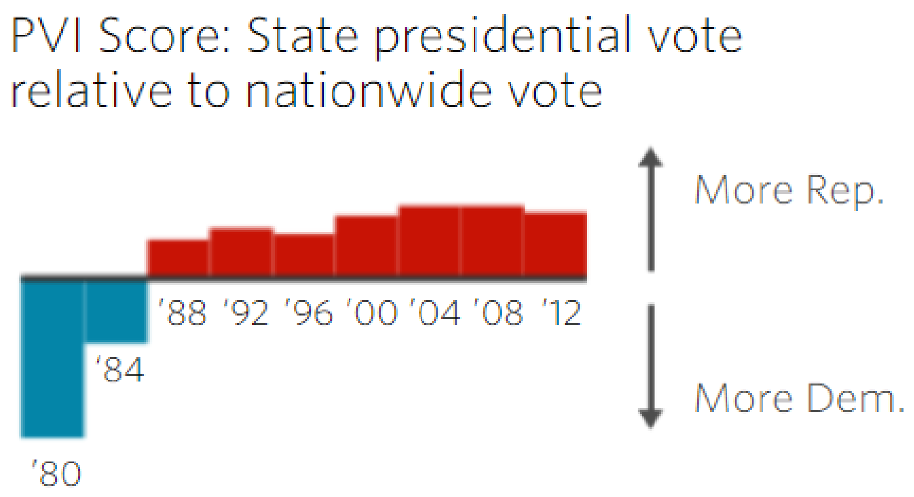
* These guidelines are drawn largely from my experience + personal preferences + the literature. Design is messy, these are not perfect, others will disagree with me, etc. *Caveat emptor.*

Small multiples



[https://excelcharts.com/animation-small-multiples-growth-walmart-excel-edition/]

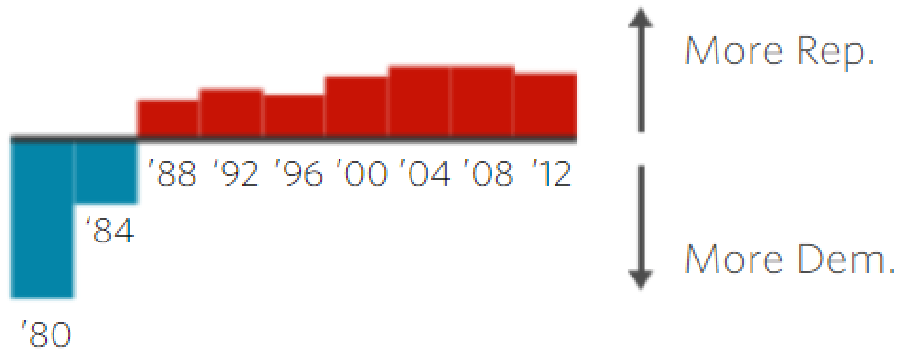
Small multiples



[http://graphics.wsj.com/elections/2016/field-guide-red-blue-america/]

Small multiples

PVI Score: State presidential vote relative to nationwide vote



[http://graphics.wsj.com/elections/2016/field-guide-red-blue-america/]