

“The human understanding, on account of its own nature, readily supposes a greater order and uniformity in things than it finds. And ... it devises parallels and correspondences and relations which are not there.”

—Francis Bacon, 1620

Is what we see really there?

“The human understanding, on account of its own nature, readily supposes a greater order and uniformity in things than it finds. And ... it devises parallels and correspondences and relations which are not there.”

—Francis Bacon, 1620

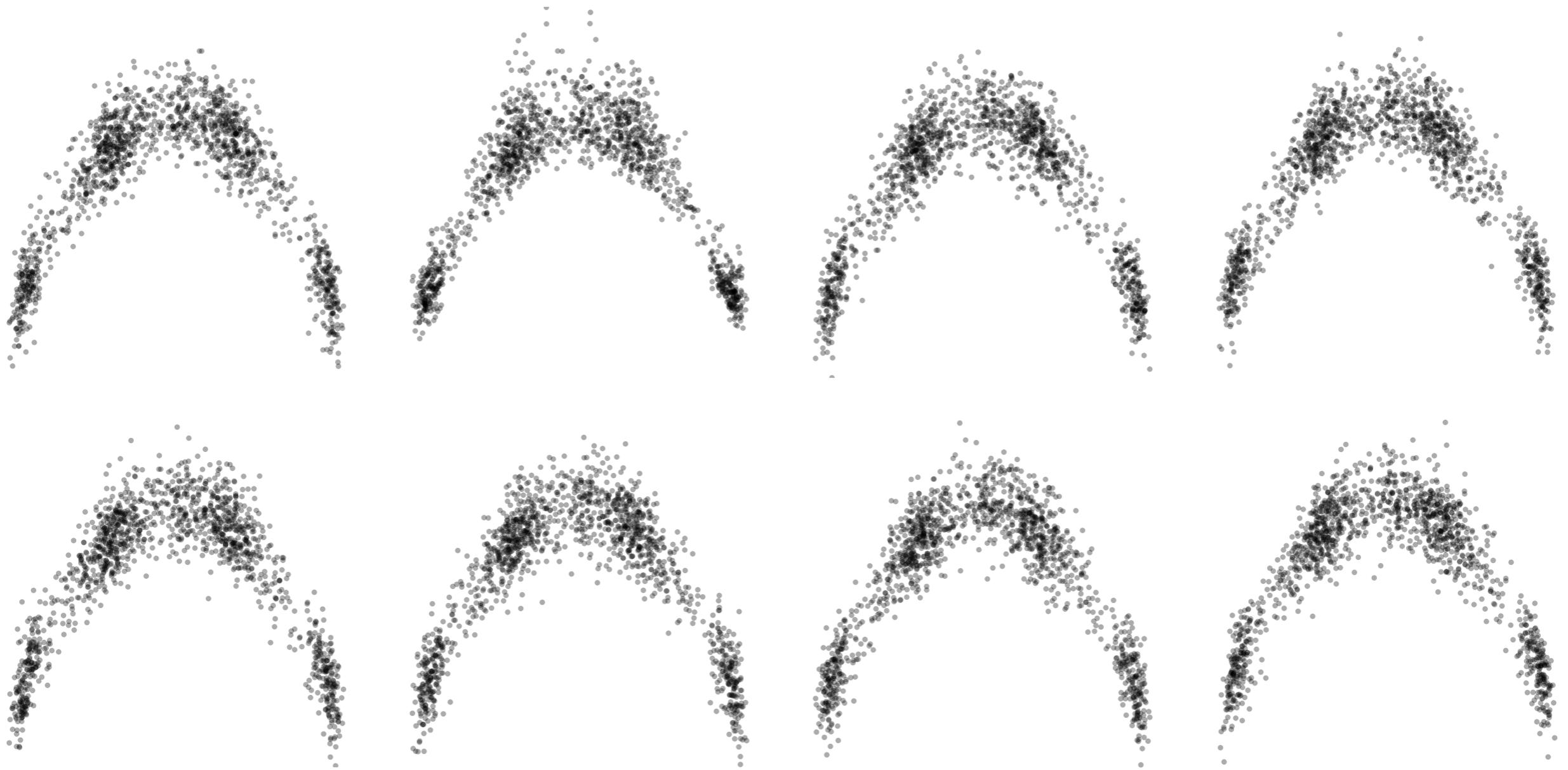
Graphical inference for infovis

Hadley Wickham, Dianne Cook,
Heike Hofmann, Andreas Buja

October 2010



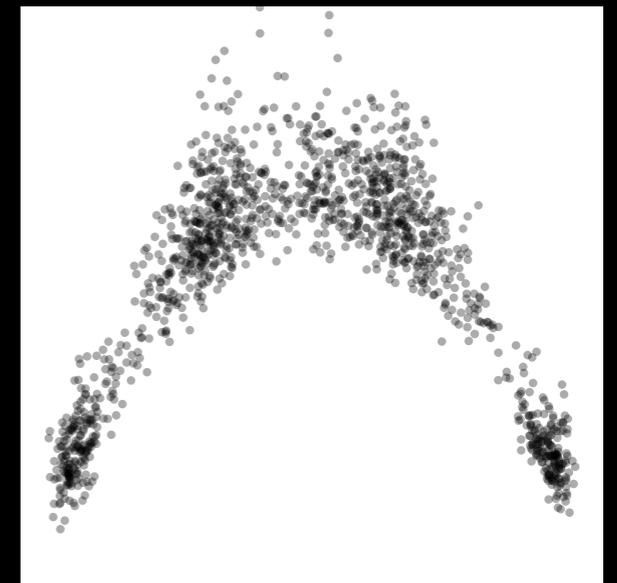
Wednesday, September 26, 12



Which one of these plots is not like the others?
Which of these plots just doesn't belong?

7 of those plots were plots of random (null) data. 1 plot was the real data.

If you correctly picked the true plot from the null plots then we have evidence that it really is different.



In fact, we have rigorous statistical evidence that there is a difference, just using Sesame Street skills!

1. The statistical justice system

2. Line up protocol

3. Rorschach protocol

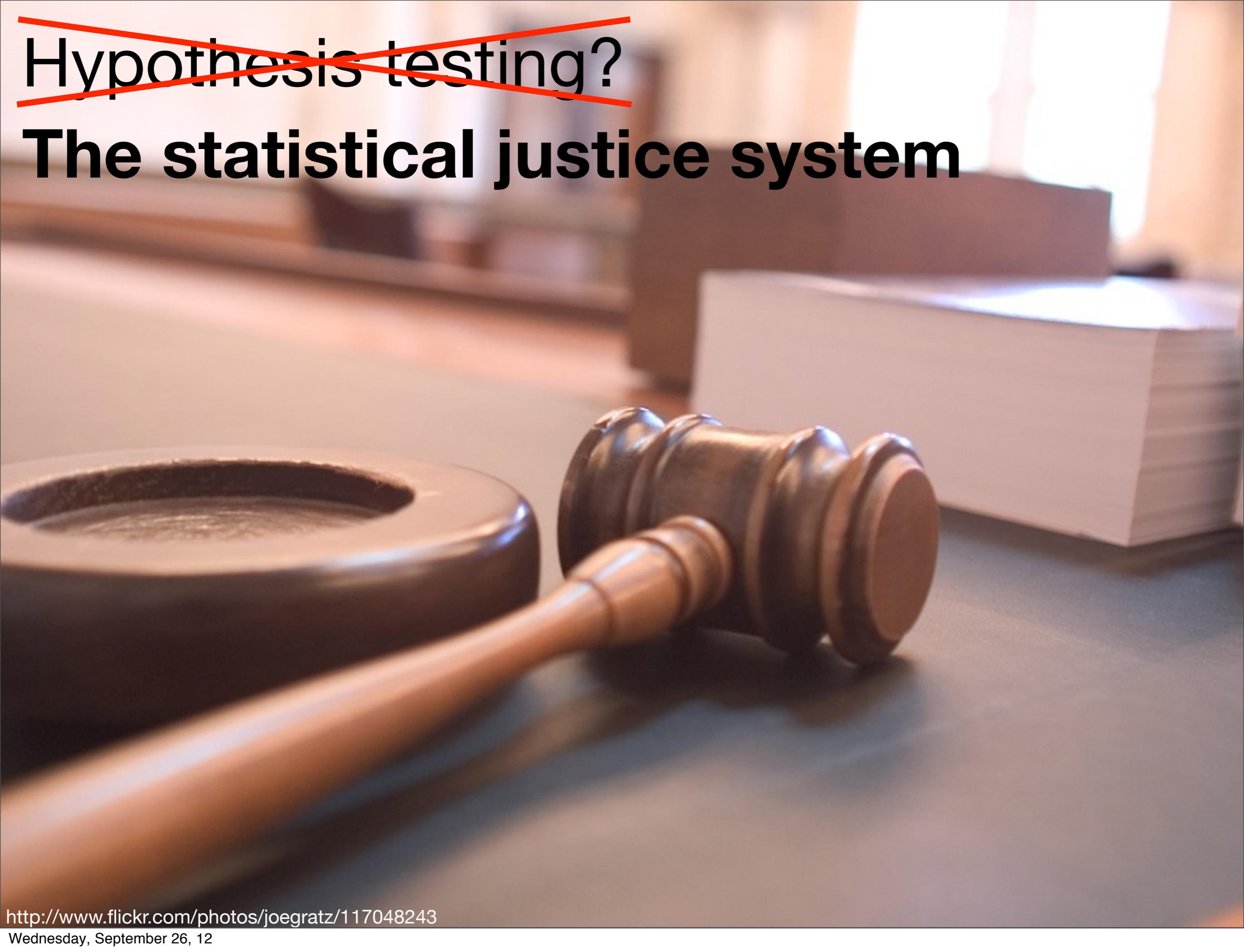
4. Future work

Hypothesis testing?



~~Hypothesis testing?~~

The statistical justice system



Ho: null hypothesis

Defence

Ha: alternative hypothesis

Prosecution

Ho: null hypothesis

Defence

Ha: alternative hypothesis

Prosecution

Null distribution

Innocents

Ho: null hypothesis

Defence

Ha: alternative hypothesis

Prosecution

Null distribution

Innocents

Reject the null

Guilty

Fail to reject the null

Not guilty

Ho: null hypothesis

Ha: alternative hypothesis

Null distribution

Reject the null

Fail to reject the null

p-value

Defence

Prosecution

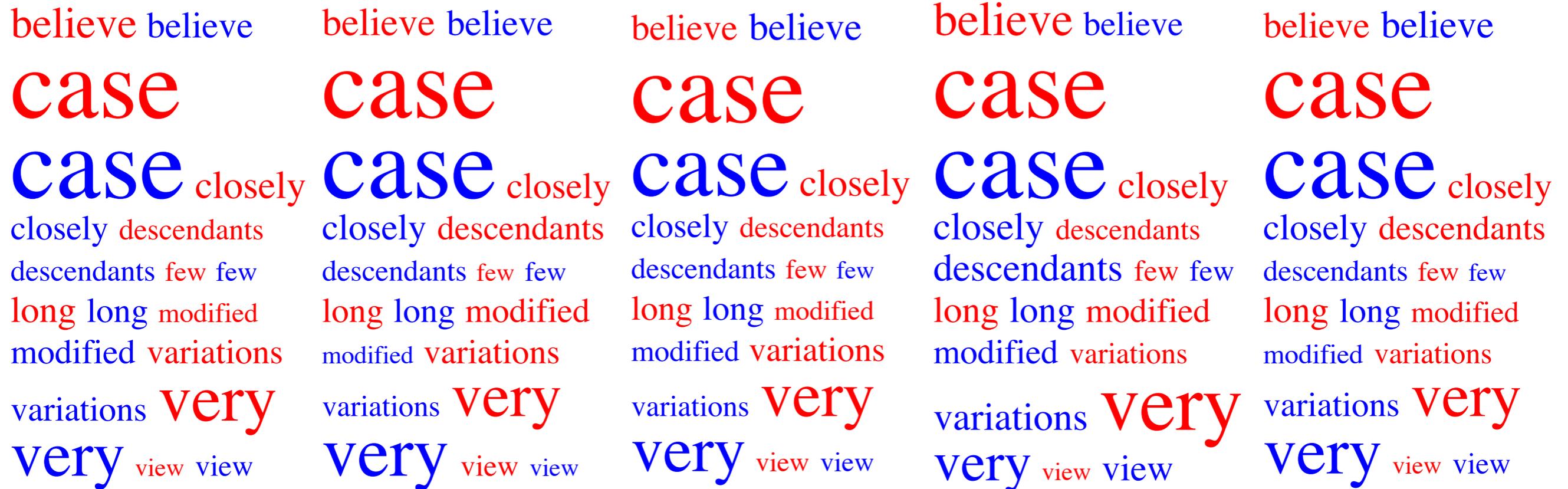
Innocents

Guilty

Not guilty

Probability that a truly innocent dataset would look as guilty as the suspect

Line up



Five tag clouds of selected words from the 1st (red) and 6th (blue) editions of Darwin’s “Origin of Species”. Four of the tag clouds were generated under the null hypothesis of no difference between editions, and one is the true data. Can you spot it?

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

Five tag clouds of selected words from the 1st (red) and 6th (blue) editions of Darwin's "Origin of Species". Four of the tag clouds were generated under the null hypothesis of no difference between editions, and one is the true data. Can you spot it?

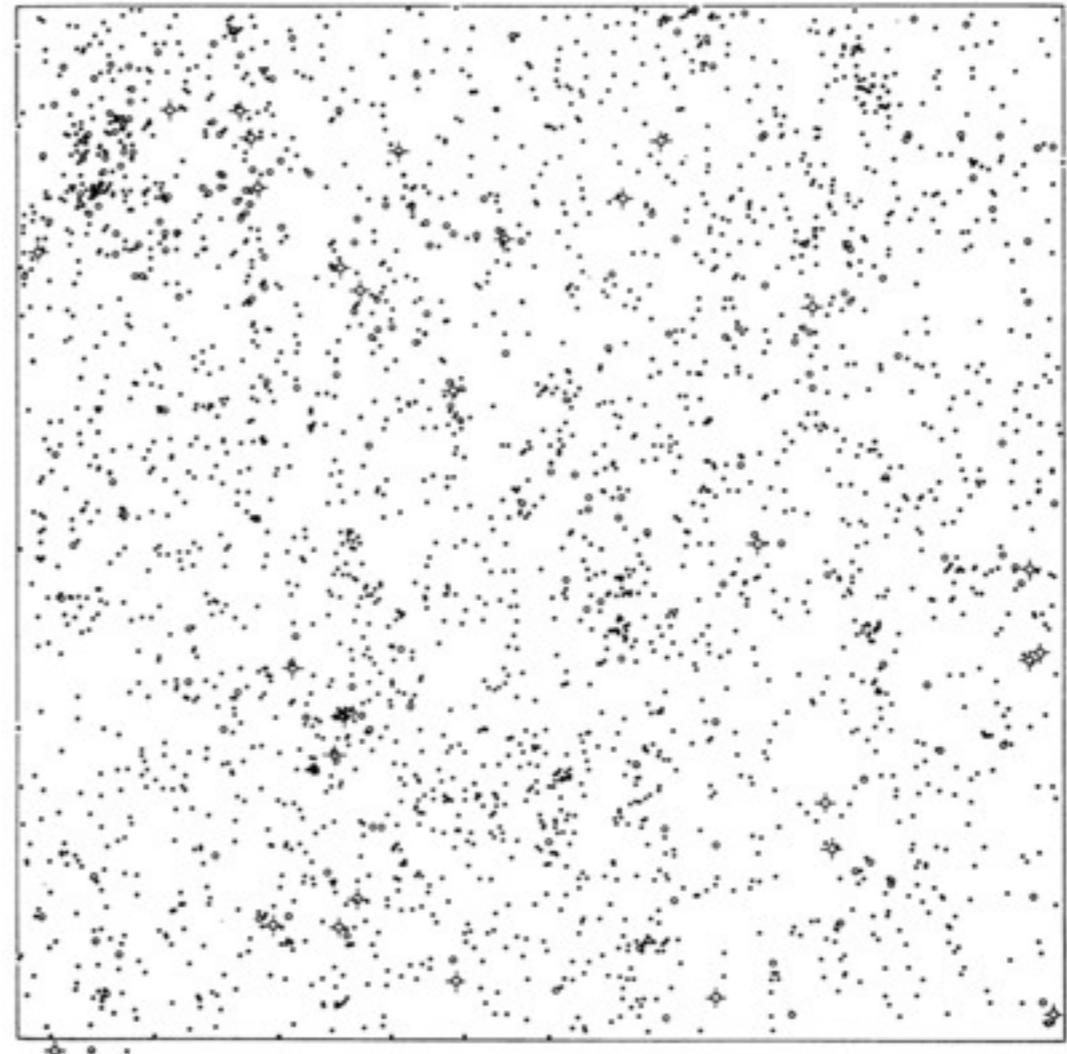
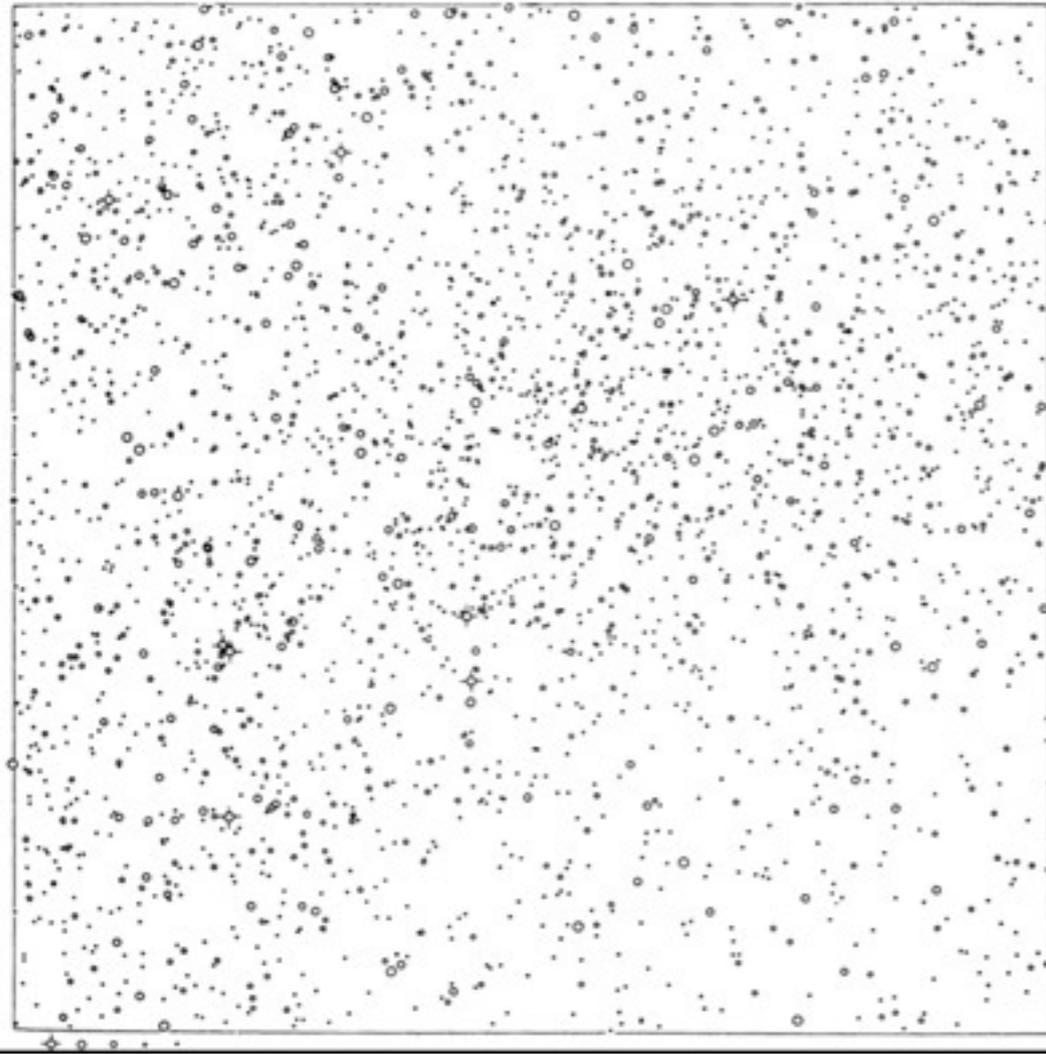
Protocol

Generate $n-1$ decoys
(null datasets)

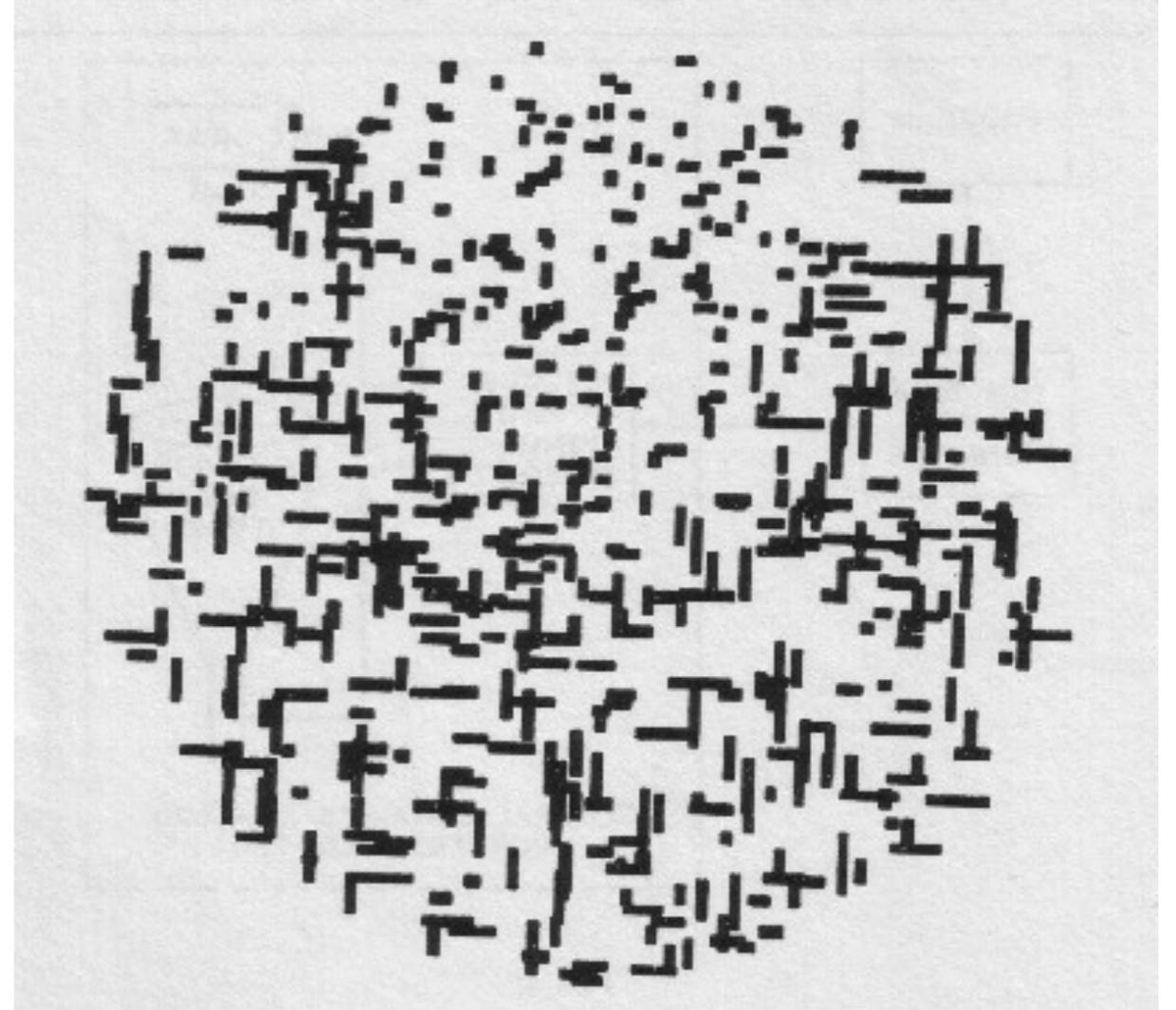
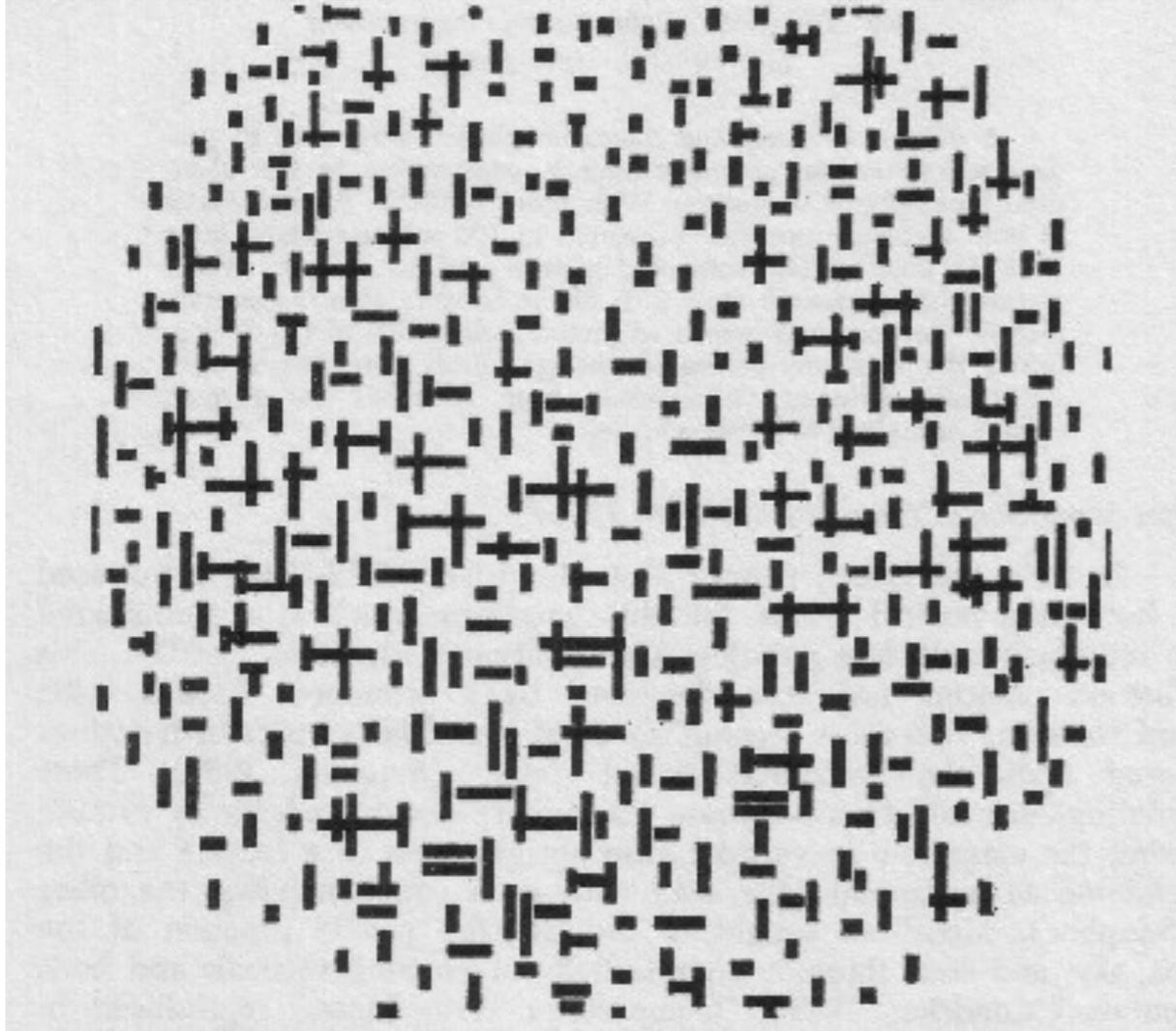
Plot the decoys + the real data
(randomly positioned)

Show to an **impartial** observer.
Can they spot the real data?

If so, you have evidence for true difference
(p-value = $1/n$)



E. L. Scott, C. D. Shane, and M. D. Swanson. Comparison of the synthetic and actual distribution of galaxies on a photographic plate. *Astrophysical Journal*, 119:91–112, Jan. 1954.



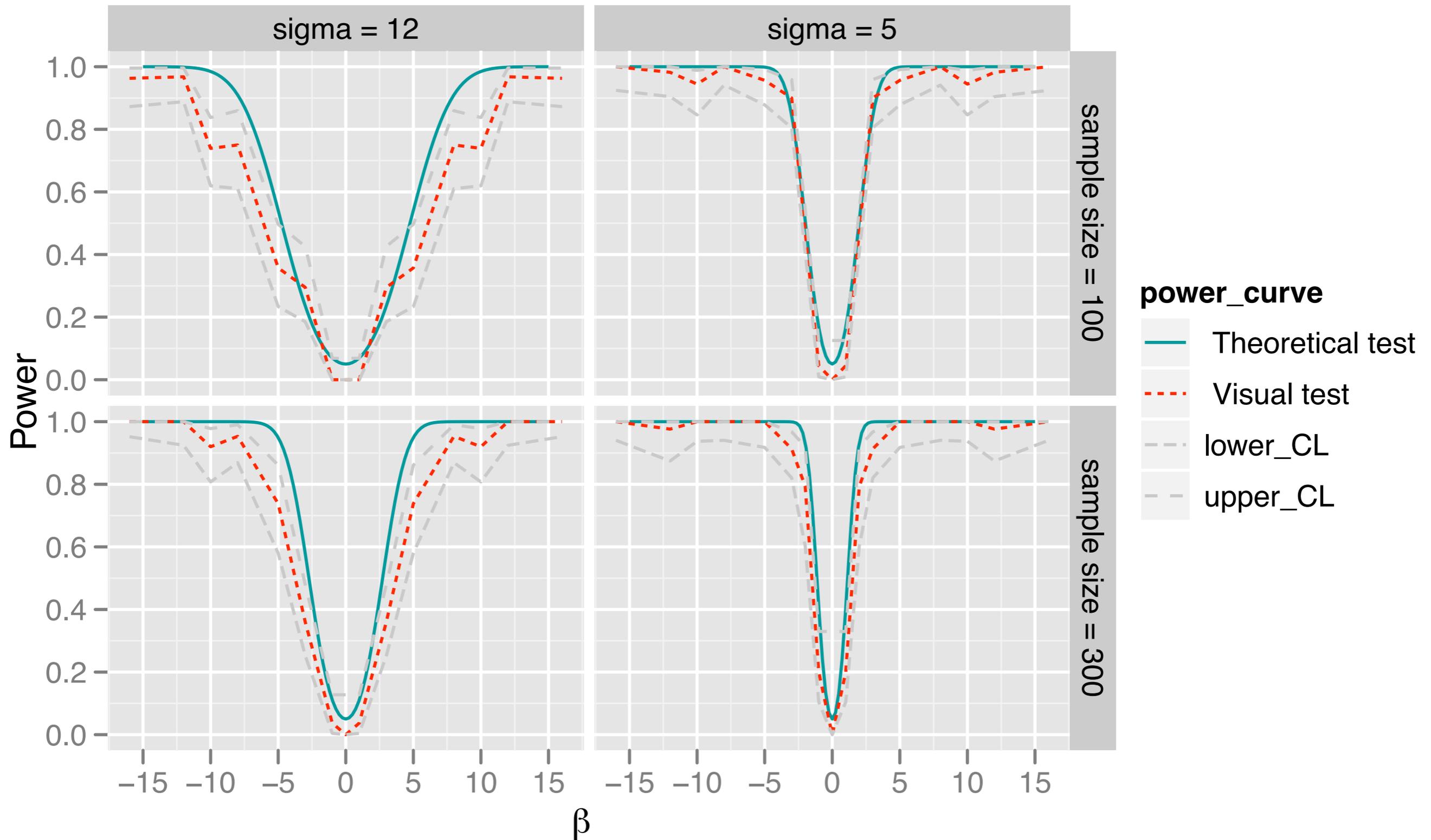
A. M. Noll. Human or machine: A subjective comparison of Piet Mondrian's "composition with lines" (1917) and a computer-generated picture. *The Psychological Record*, 16:1-10, 1966.

vs. classical tests

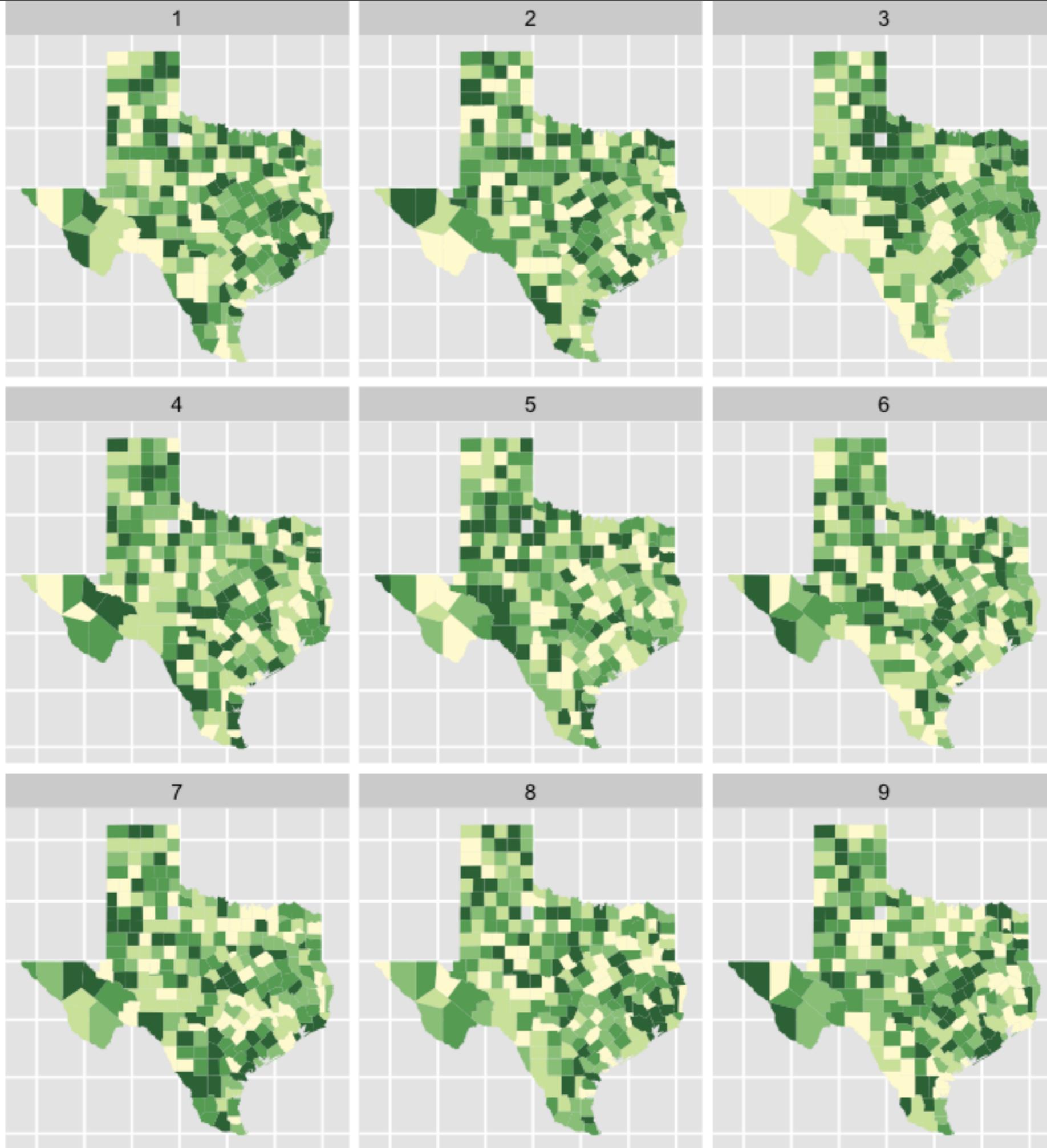
Of course, if we know what we're looking for, we can always develop an algorithm or numerical test.

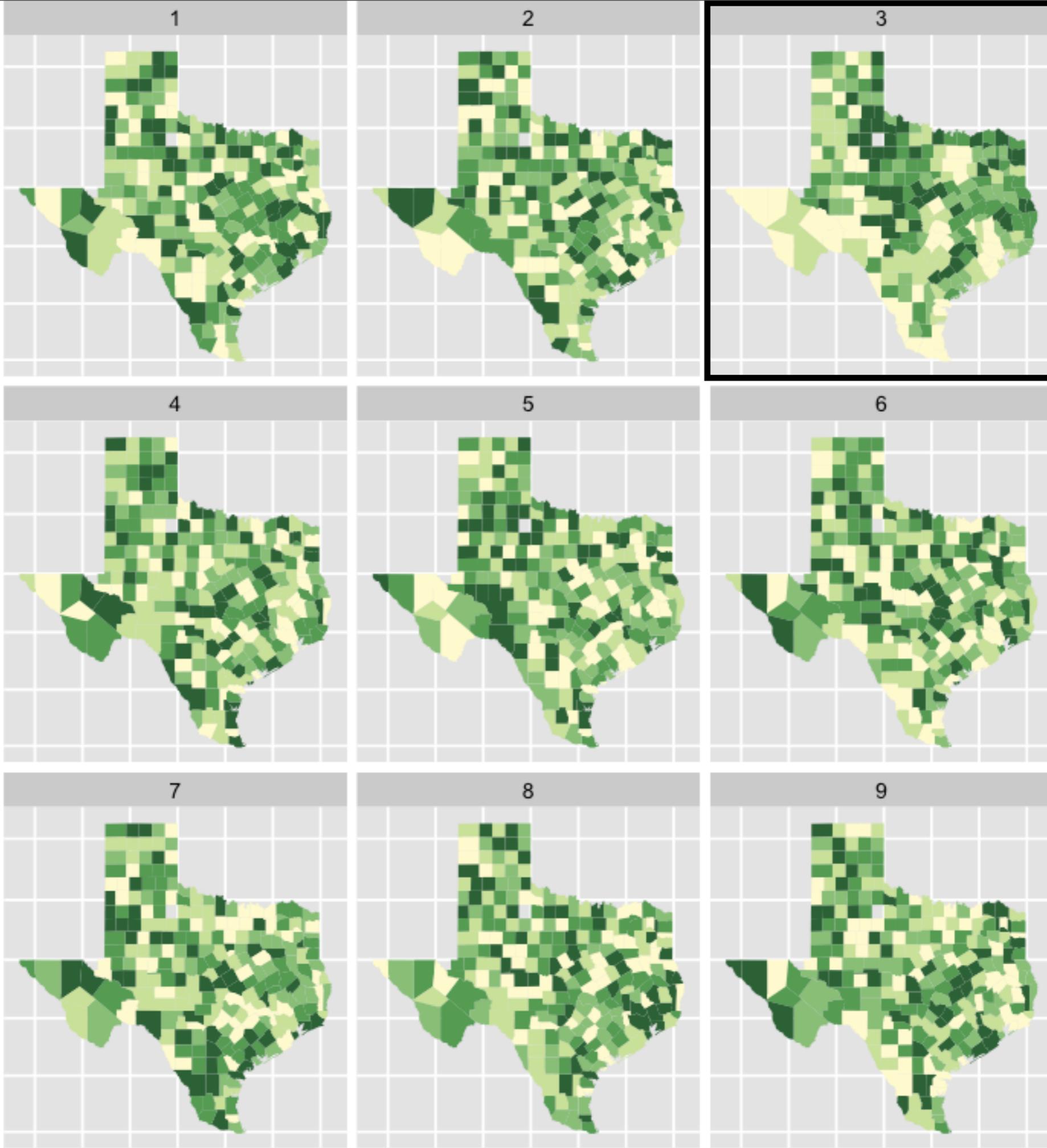
The **advantage** of visual inference is that works for very general tasks, including when you don't know exactly what you're looking for.

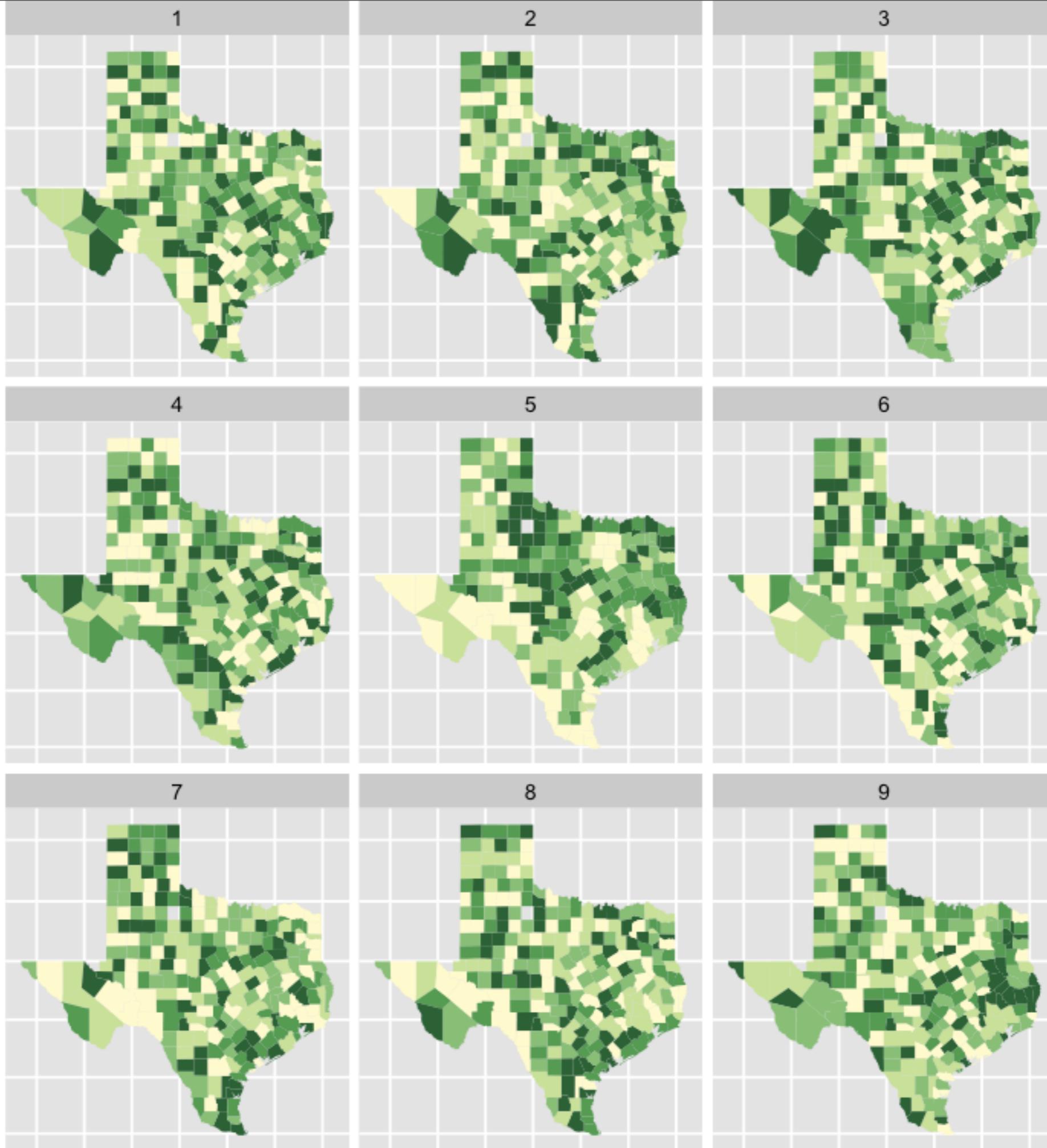
Recent work shows that power only a little worse than classical test

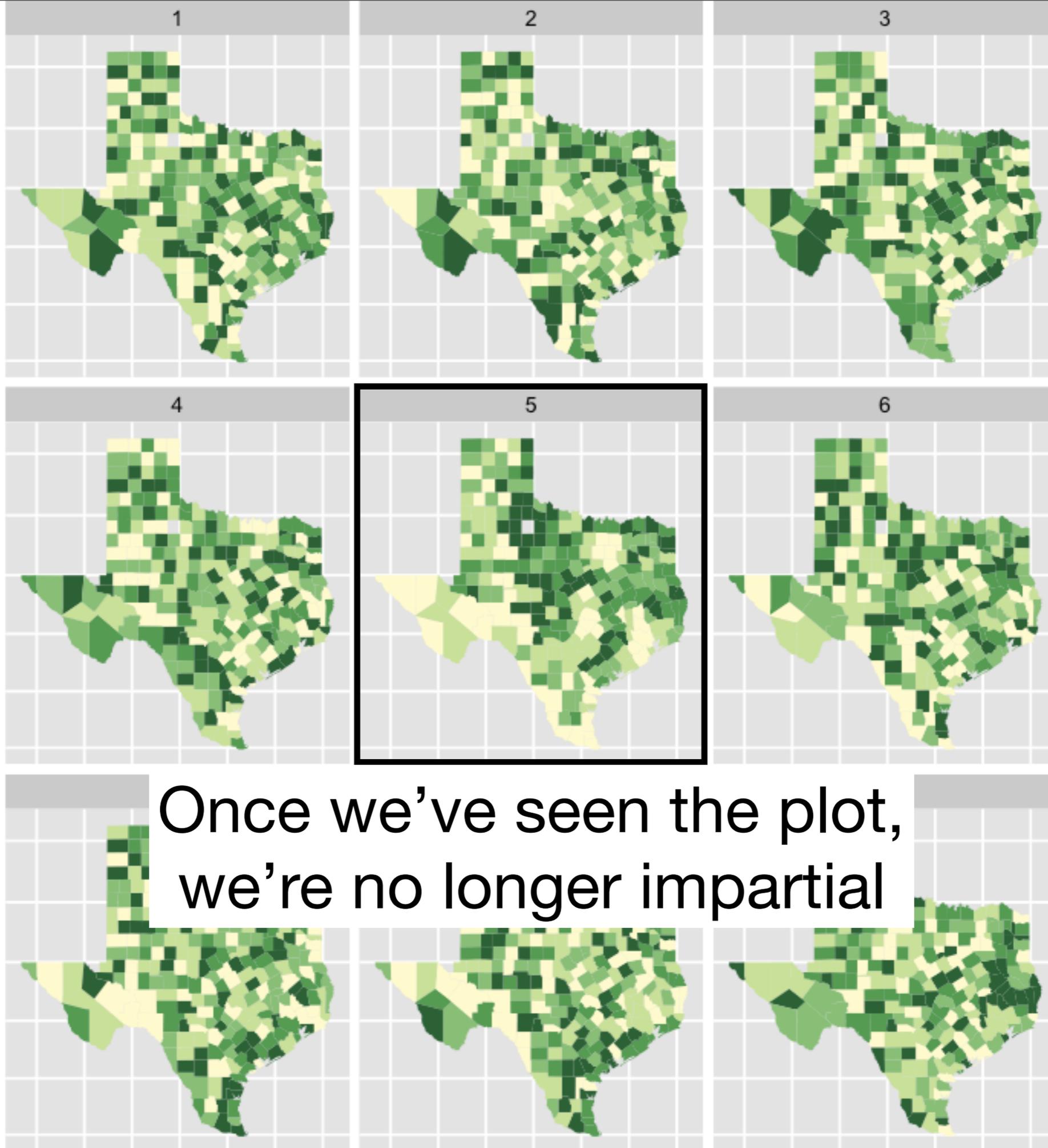


Plot	Task
Choropleth map	Is there a spatial trend?
Treemap	Is the distribution in higher level categories the same?
Scatterplot	Are the two variables independent?
Time series	Is there a trend in mean or variability?









Code

```
# Support package written in R
# http://github.com/ggobi/nulllabor
# Provides reference implementation of ideas

library(nulllabor)
library(ggplot2)

qplot(angle * 180 / pi, r, data = threept) %+%
  lineup(null_model(r ~ poly(angle, 2)), n = 10) +
  facet_wrap(~ .sample, ncol = 5)
```

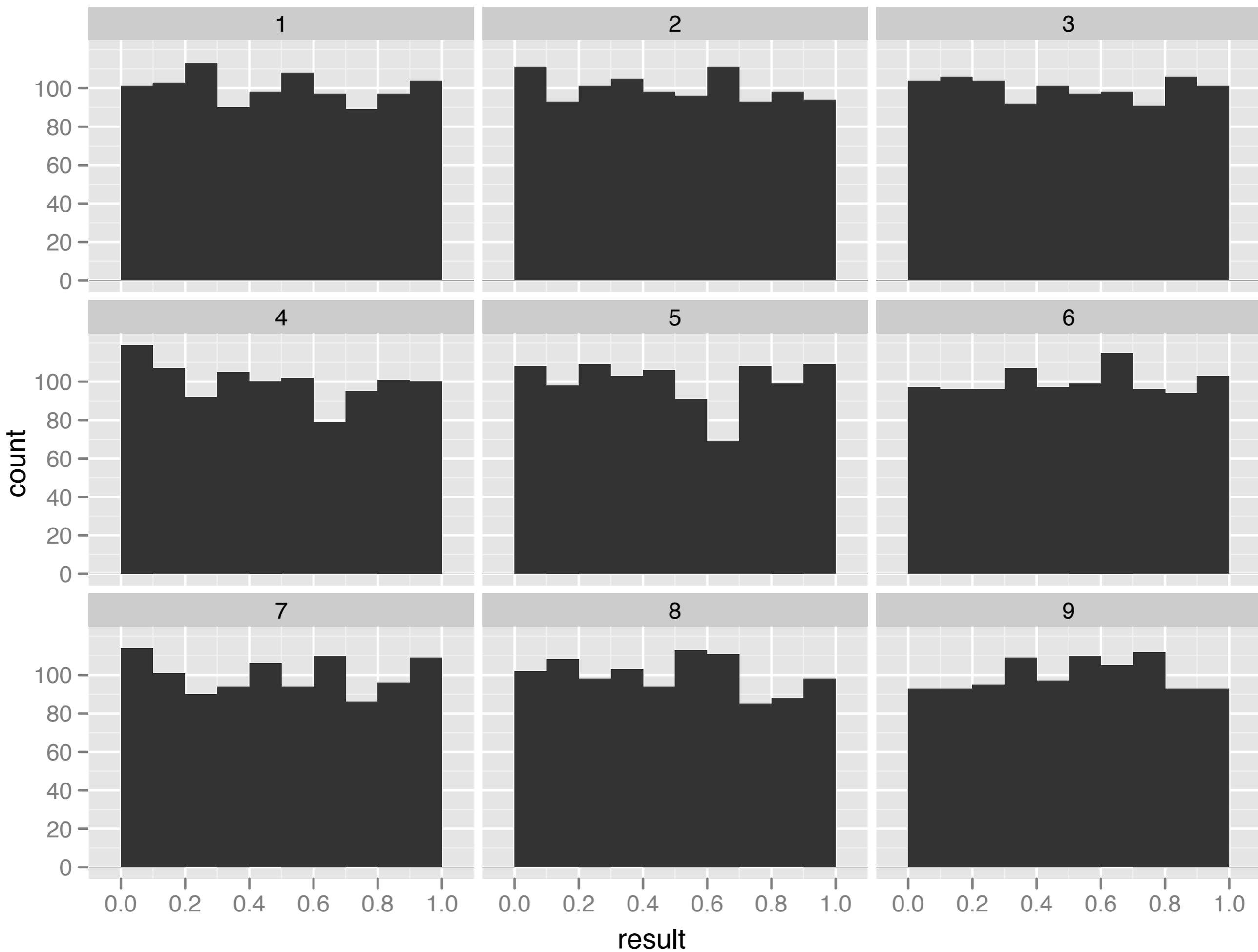
Rorschach

Rorschach

We're surprisingly bad at appreciating the amount of variation in random data.

Showing only null plots is a good way to calibrate our intuition.

We also plan on using these plots as an empirical tool to understand what features people pick up on. Anecdotally, undergrads focus too much on outliers



Future work

Future work

How can visual inference be integrated into visualisation software at a fundamental level?

How does training impact results? How do novices vs. experts differ?

What patterns do people pick up on?
What are the alternatives that people respond to?

Questions?

This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.