

Uncertainty

DSC 106: Data Visualization

Sam Lau

UC San Diego

Join at
slido.com
#7283



Announcements

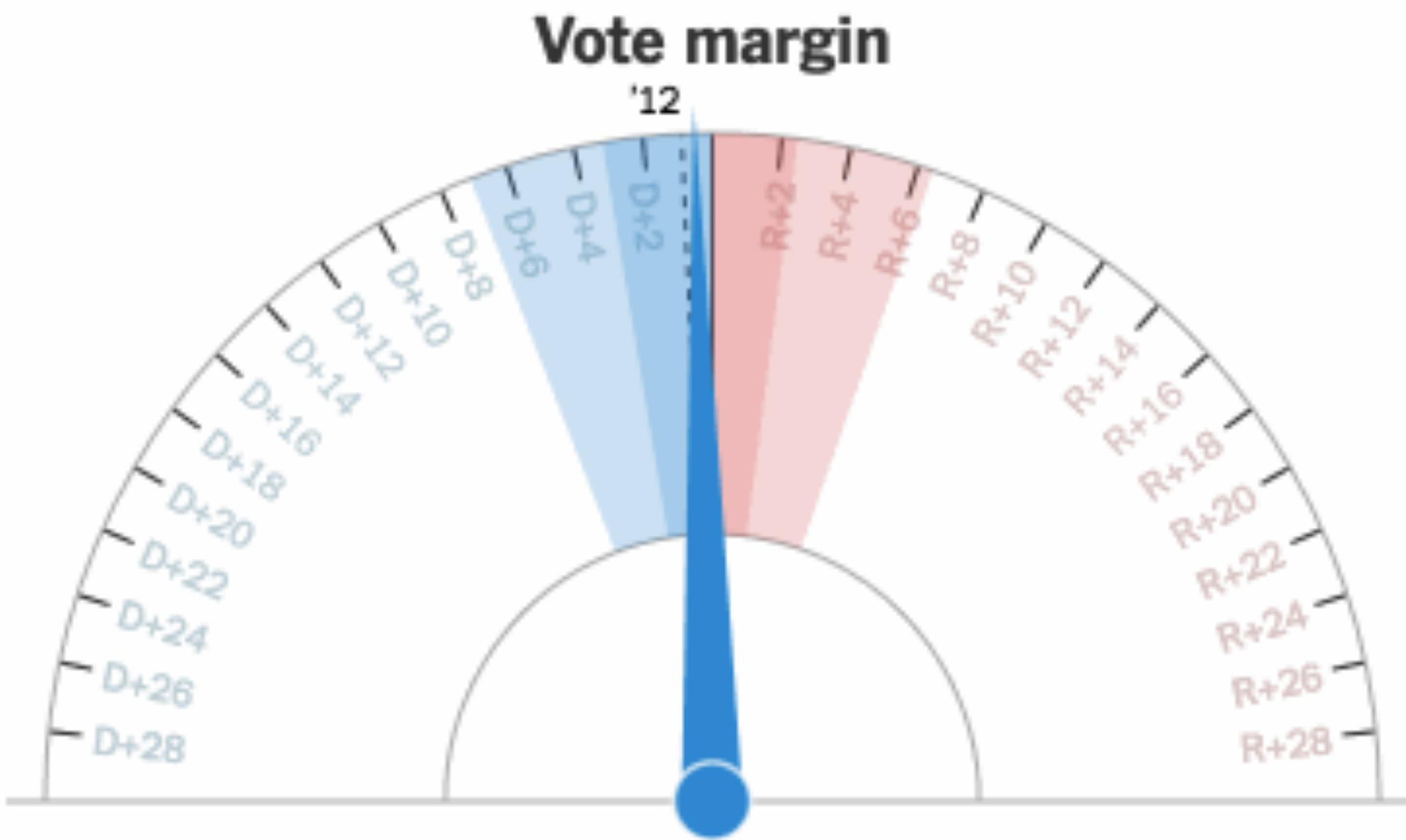
Final Project prototype due tomorrow.

Final Project video due next week Friday.

Lecture next Thurs. is Final Project peer feedback.

FAQs:

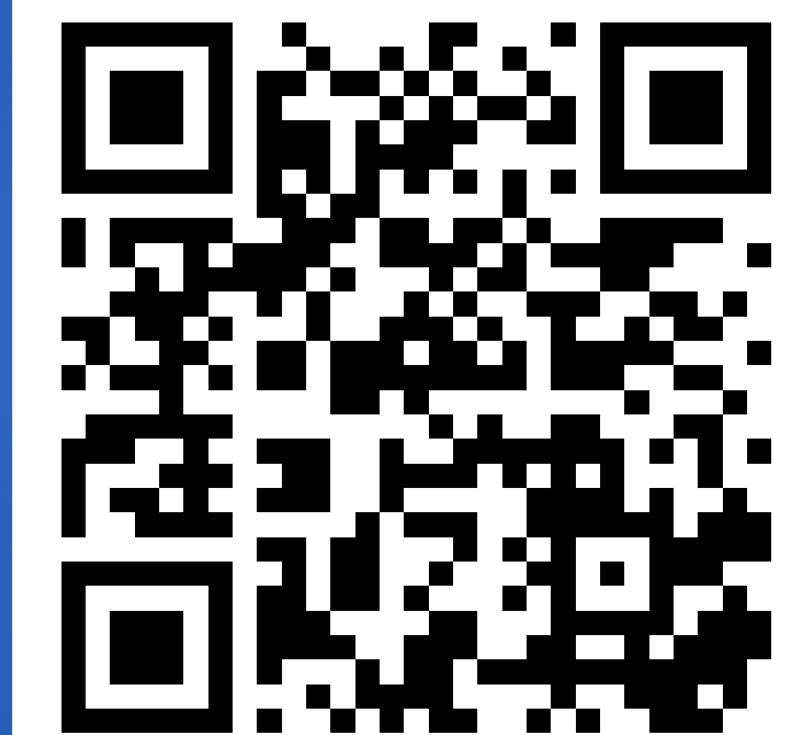
1. What is required for the final project prototype? See the website for full rubric.

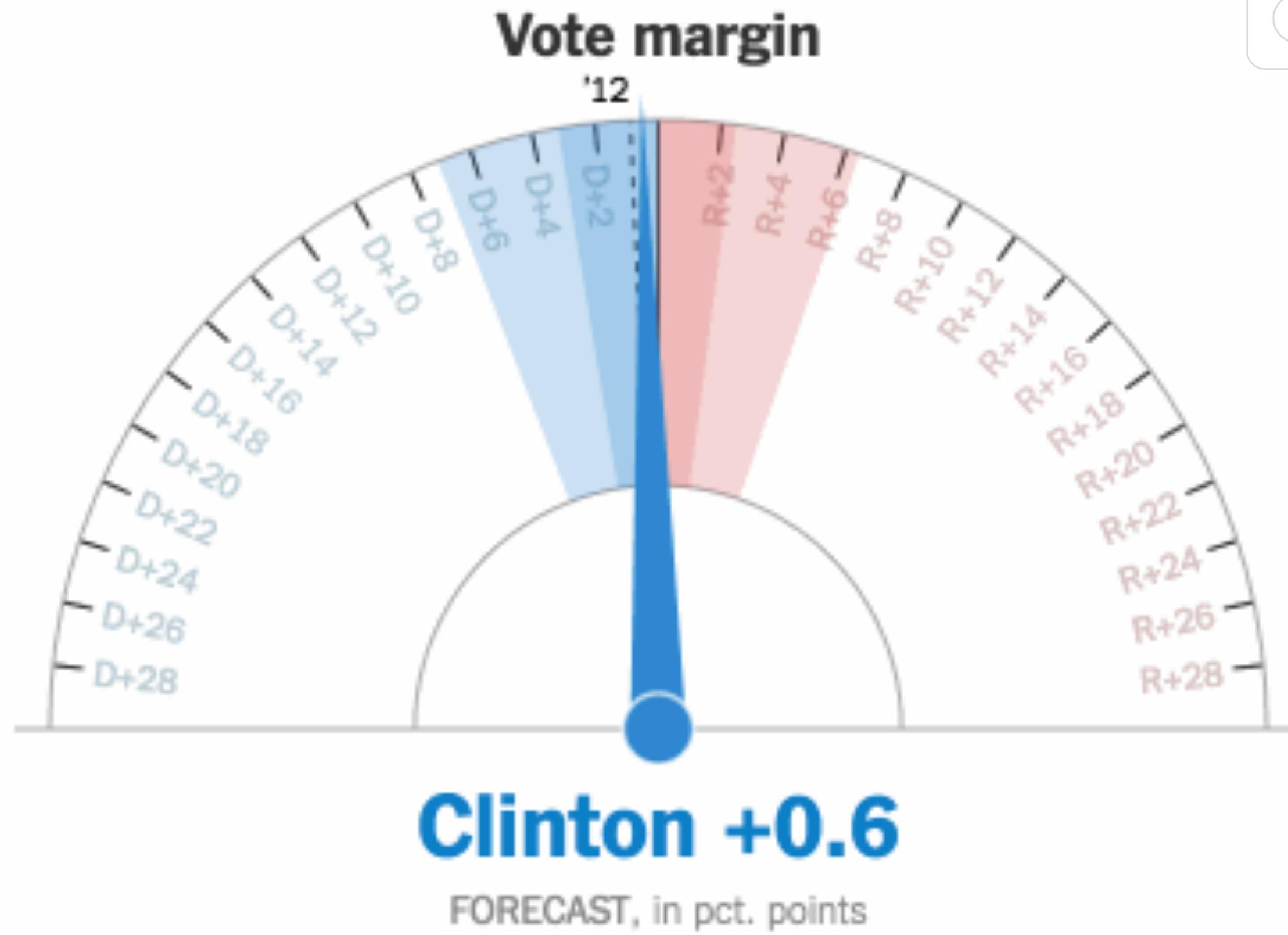


What is being visualized?

What are the strengths and weaknesses of this visualization?

Join at
slido.com
#7283





(gobsmacked) @nv... · 10:08 PM · Nov 8, 2016

I'm taking damn NY

10:08 PM · 2 1

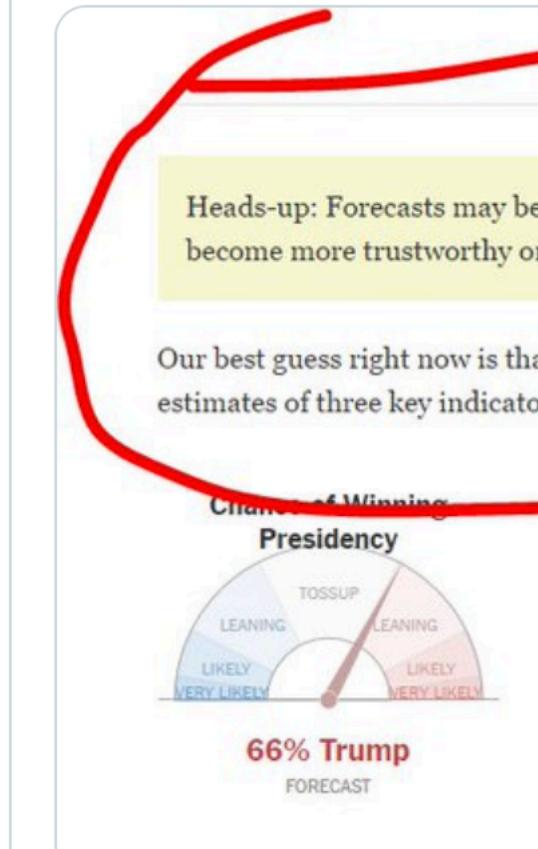
Danielle Rosen @Daniel_Rosen · Follow

The NYT needle jittering is going to be in my nightmares no matter what happens tonight.

10:00 PM · Nov 8, 2016

Paul D. Shinkman @PDShinkman · Follow

Has @nytimes caught on to Dems' election-night hair-greying freakout?



10:03 PM · Nov 8, 2016

1 1 S

Alp Toker Nov 8, 2016 · @atoker · Follow

Looking for trends in @nytimes's presidential forecast needle? Don't look too hard - the bounce is random jitter from your PC, not live data

Popular vote margin

Clinton +3.3

Watch on Twitter

FORECAST, in pct. points

Network Timeline Profiles Application Security Help

general-main.js x

```
range_85 = range_85 * damping + range_85_tgt * (1-damping);
range_25 = range_25 * damping + range_25_tgt * (1-damping);
range_75 = range_75 * damping + range_75_tgt * (1-damping);
range_95 = range_95 * damping + range_95_tgt * (1-damping);

var jitter = DITTER.get(jit_id) * jitter_range * 0.5 * (eln_forecast - 500);
if (opts.class != 'dem') jitter *= -1;

needle.attr('transform', 'scale('+scale+') rotate('+angle(cur_val)+')');
if (opts.colors) {
```

G

LIKELY

VERY LIKELY

TOSSUP

LEANING

LIKELY

VERY LIKELY

278 Clinton

FORECAST

Audits

③ : A

Watch Call Stack Not Paused Scope Not Paused Breakpoints No Breakpoints

@gka@vis.social @driven_by_data · Follow

not just random. this noise is conveying the uncertainty in our forecast (jitter range is from 25th–75th pctl in sims.)

10:01 PM · Nov 8, 2016

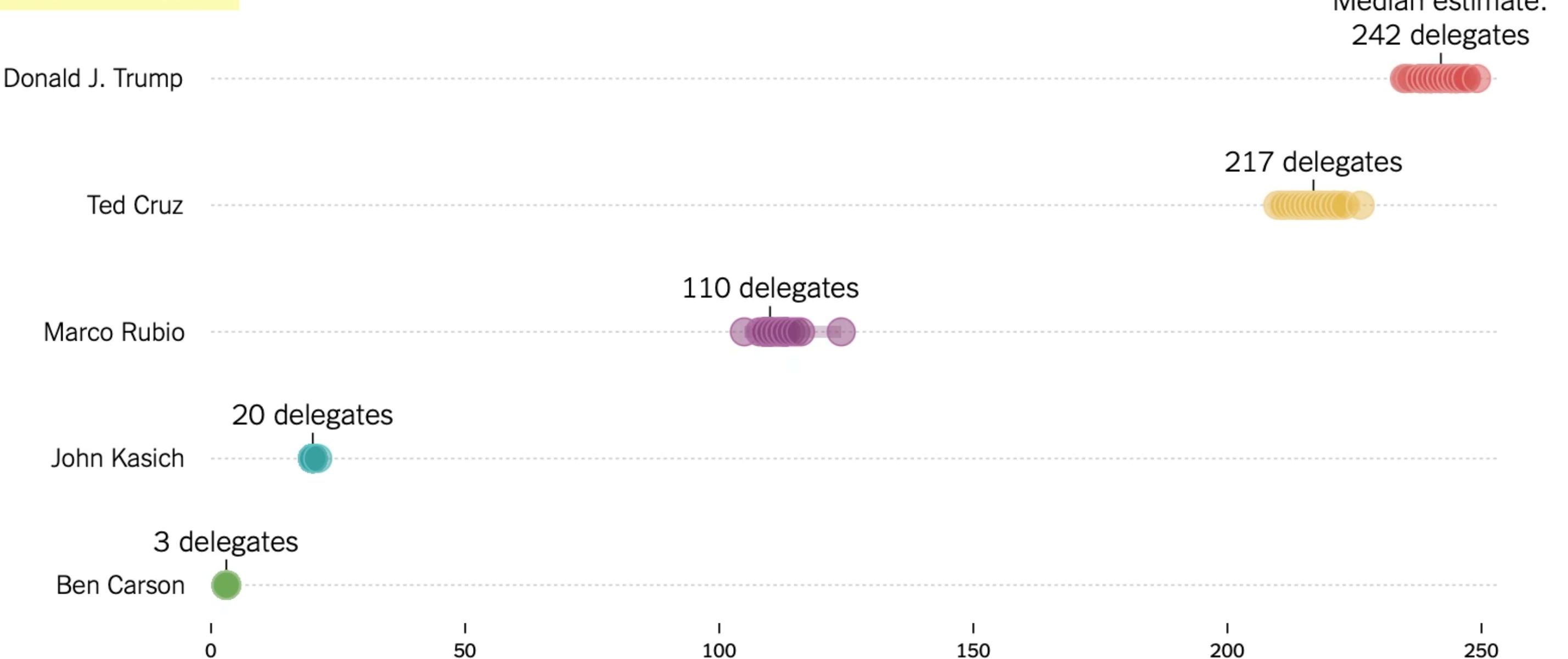
33 1 Share

Read 5 replies

Who Will Win Super Tuesday? Live Estimates of Tonight's Final Republican Delegate Count

By AMANDA COX, JOSH KATZ and KEVIN QUEALY MARCH 1, 2016

LAST UPDATED AT 3:05 AM ET



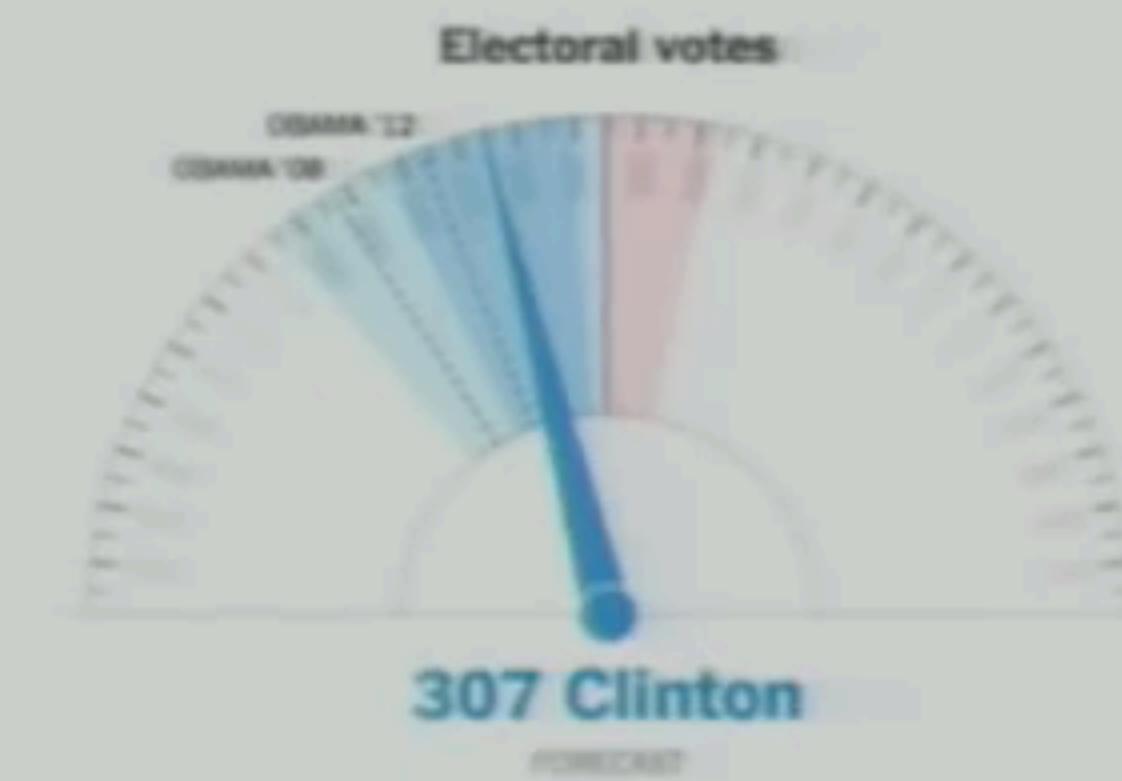
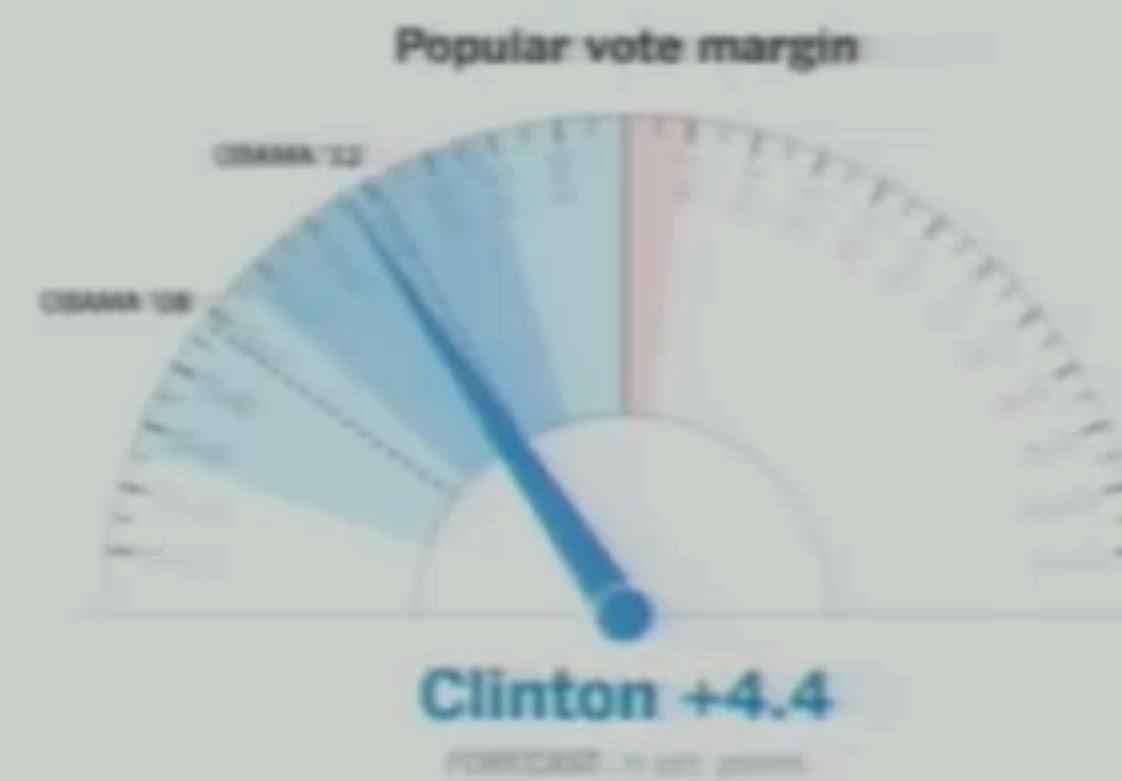
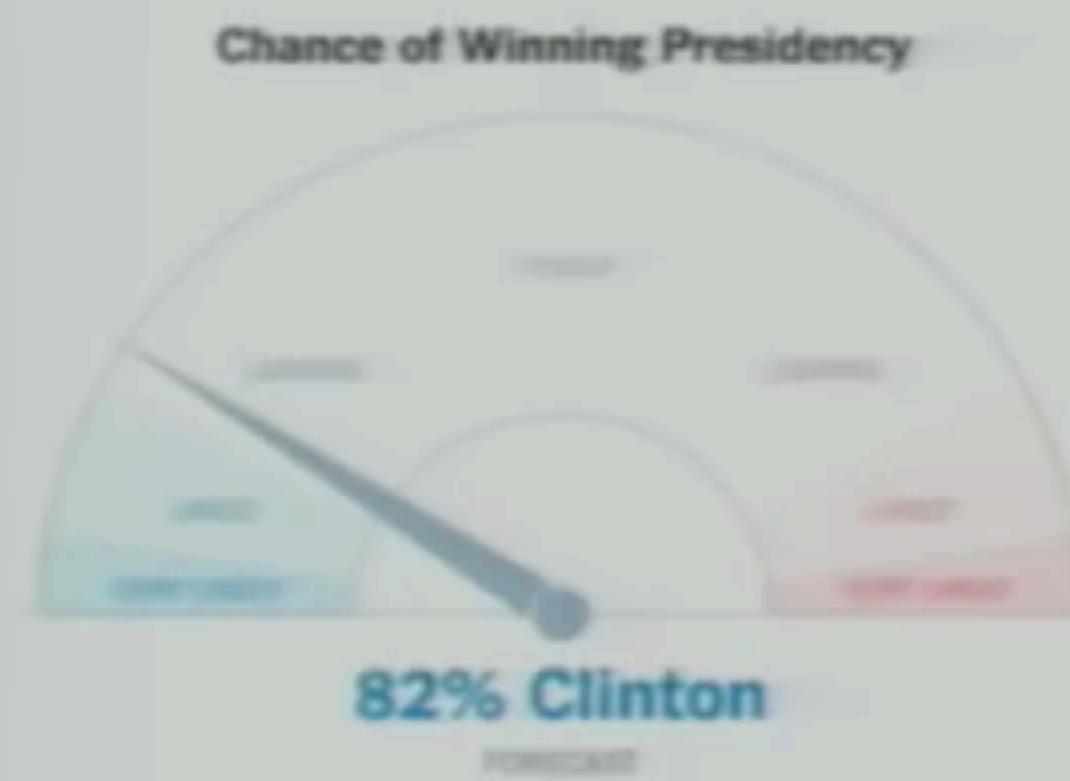
We're simulating the number of delegates each candidate will pick up on Super Tuesday. The dots above represent a range of possible outcomes.

What is being visualized?
What are the strengths
and weaknesses of this
visualization?

How does it compare to
the needle?

Live Presidential Forecast

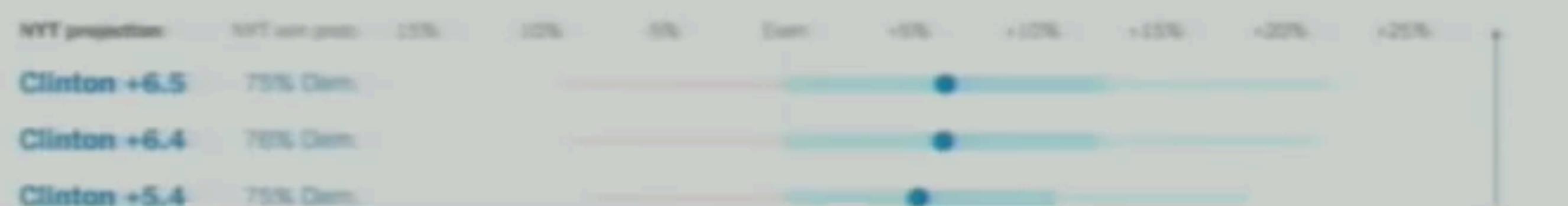
Updated 7:58:14 PM ET



The projections for each state are based on the votes reported so far and how those places have voted in previous elections.

We're showing the closest states by default. [Show all states](#)

Name	Est. perf. of states	Estimated margin
Michigan	0%	Clinton +6.5
New Mexico	0%	Clinton +6.4
Wisconsin	0%	Clinton +5.4



OPEN
VIS 2017
CONF

Uncertainty

What does it mean?

How should I visualize it?

Uncertainty

What does it mean?

How should I visualize it?

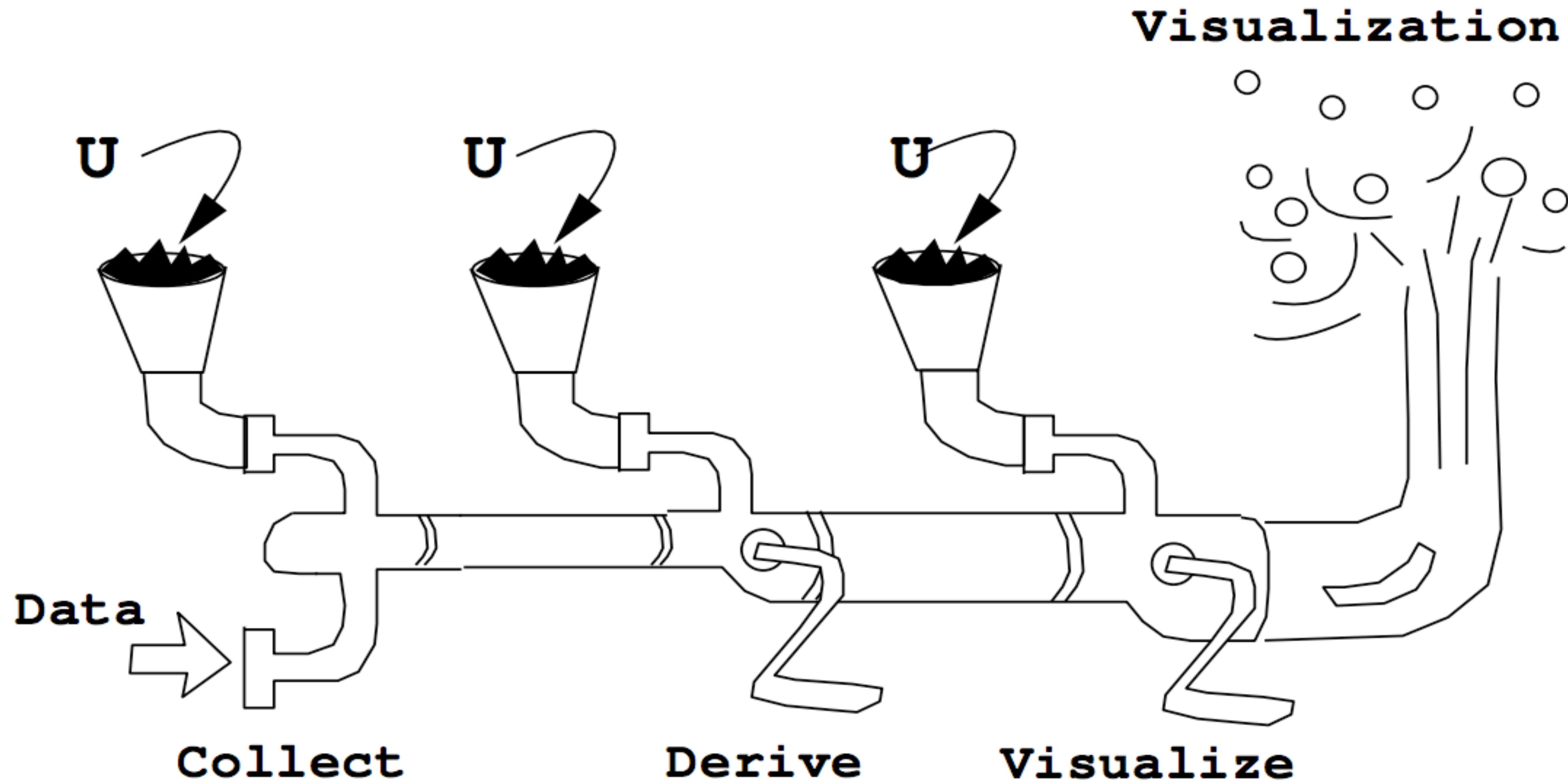
Uncertainty

What does it mean?

How should I visualize it?

Doubt
Risk
Variability
Error
Lack of knowledge
Hedging
etc...

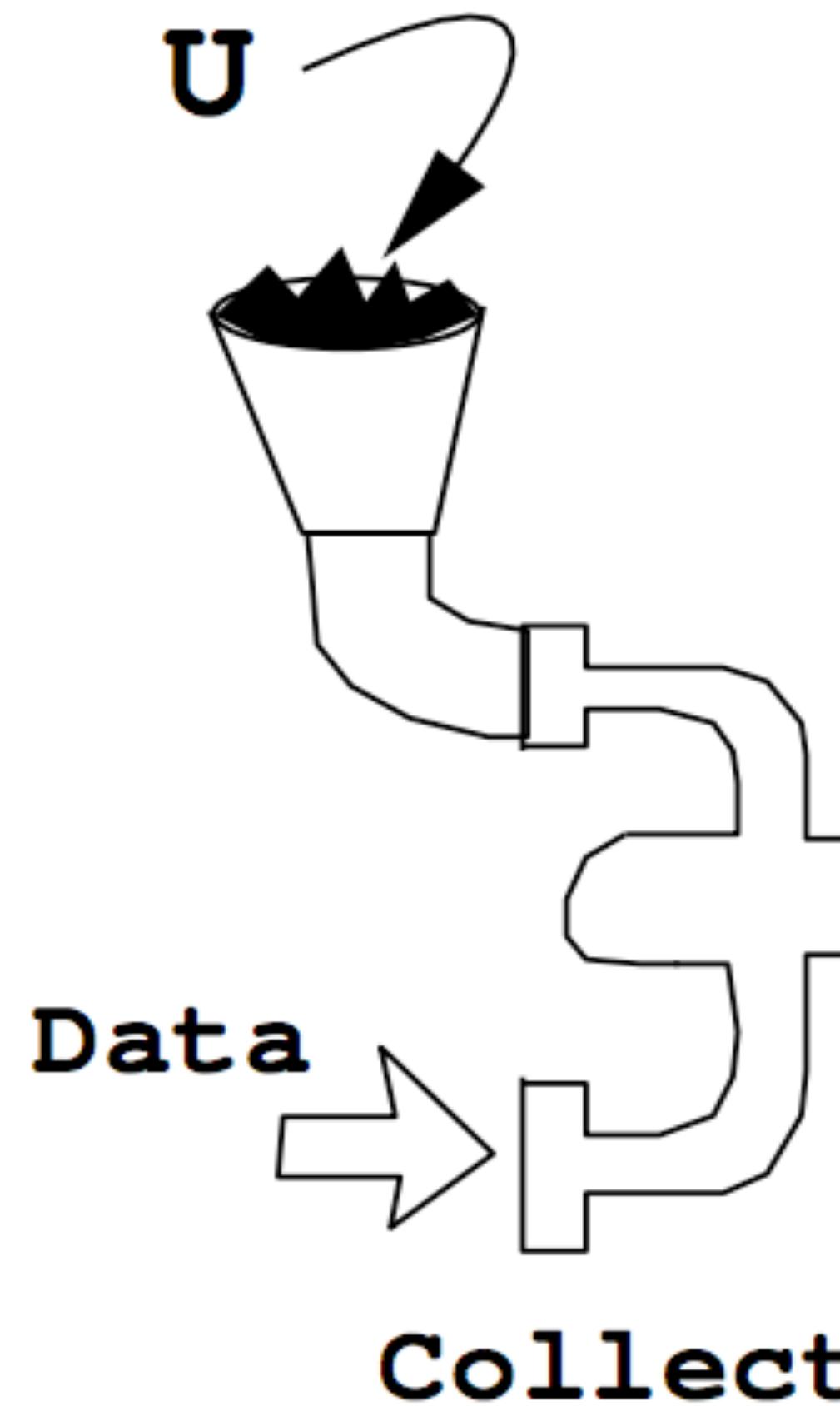
Sources and Types of Uncertainty



Sources and Types of Uncertainty

Measurement Uncertainty

How and how much should we sample the data?



Precision

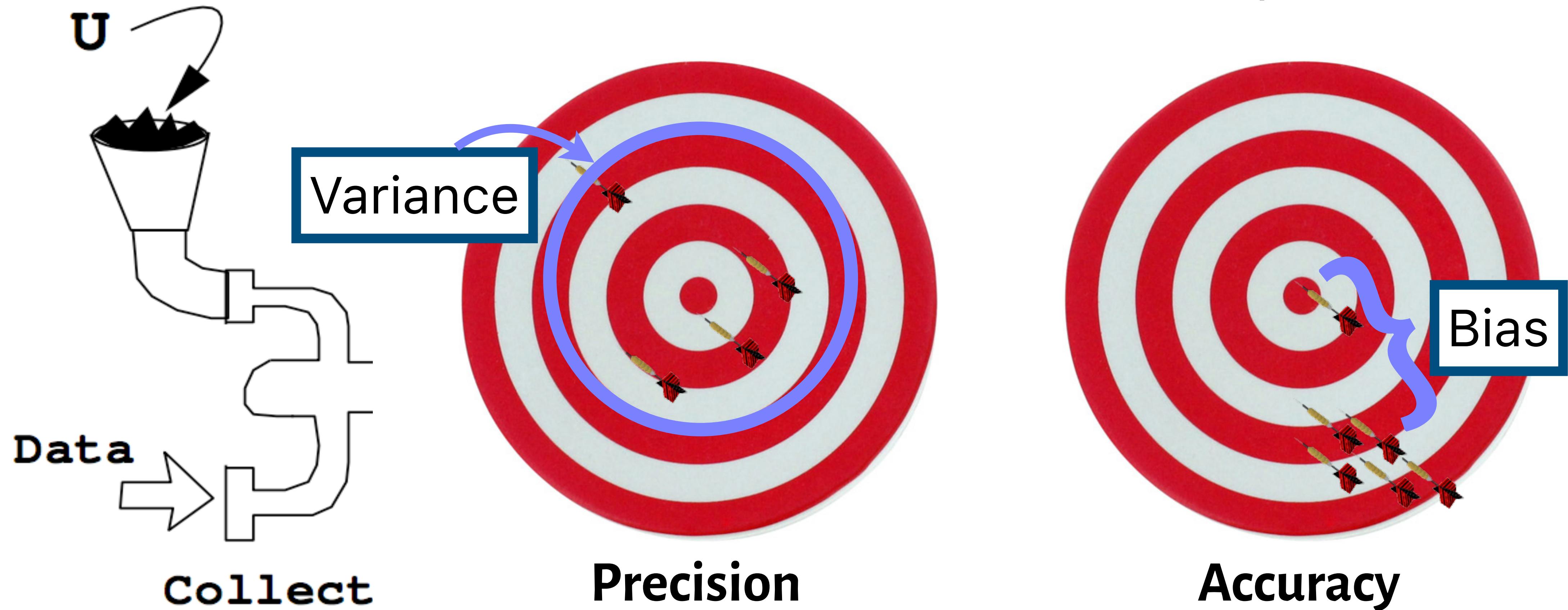


Accuracy

Sources and Types of Uncertainty

Measurement Uncertainty

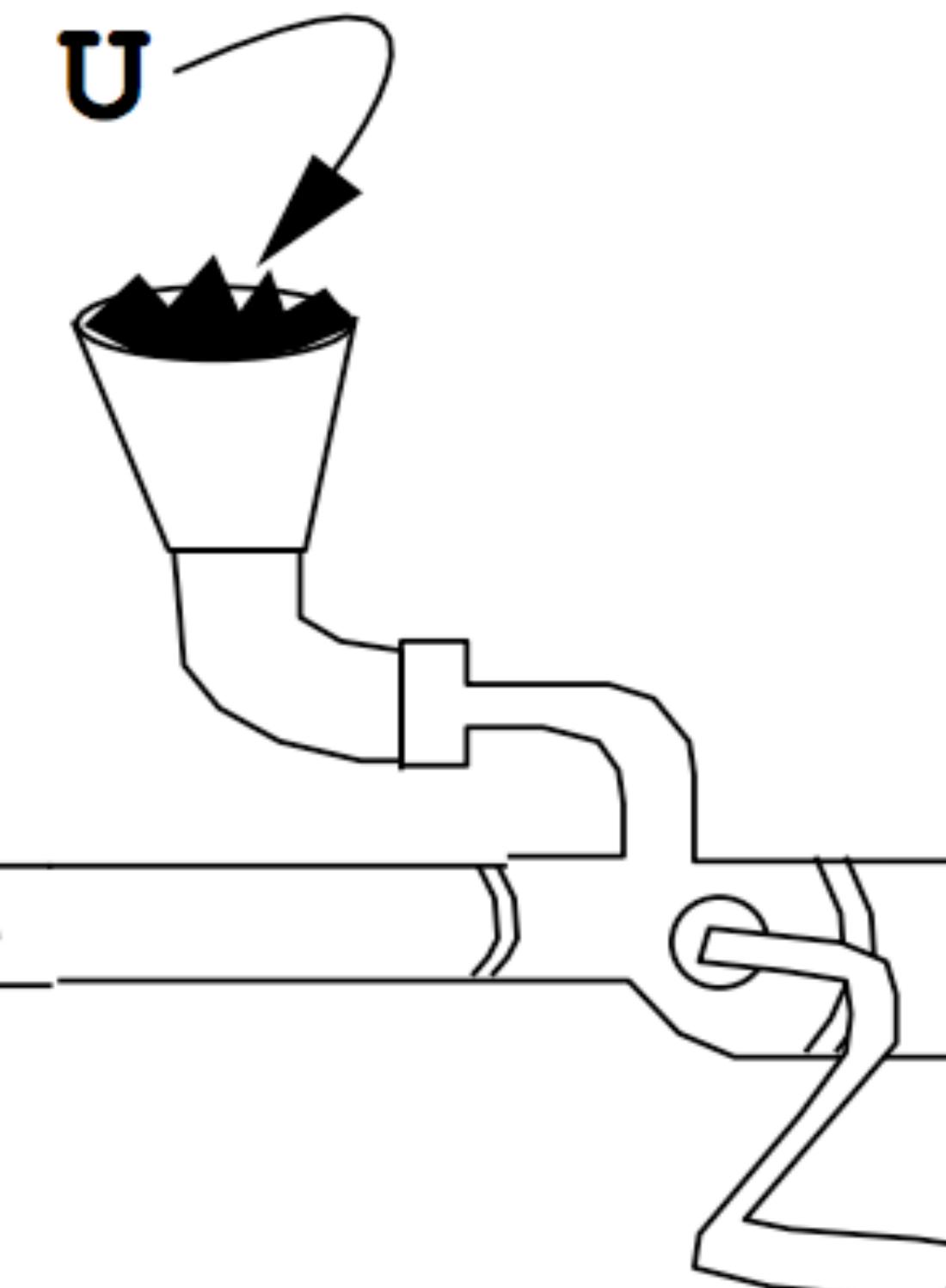
How and how much should we sample the data?



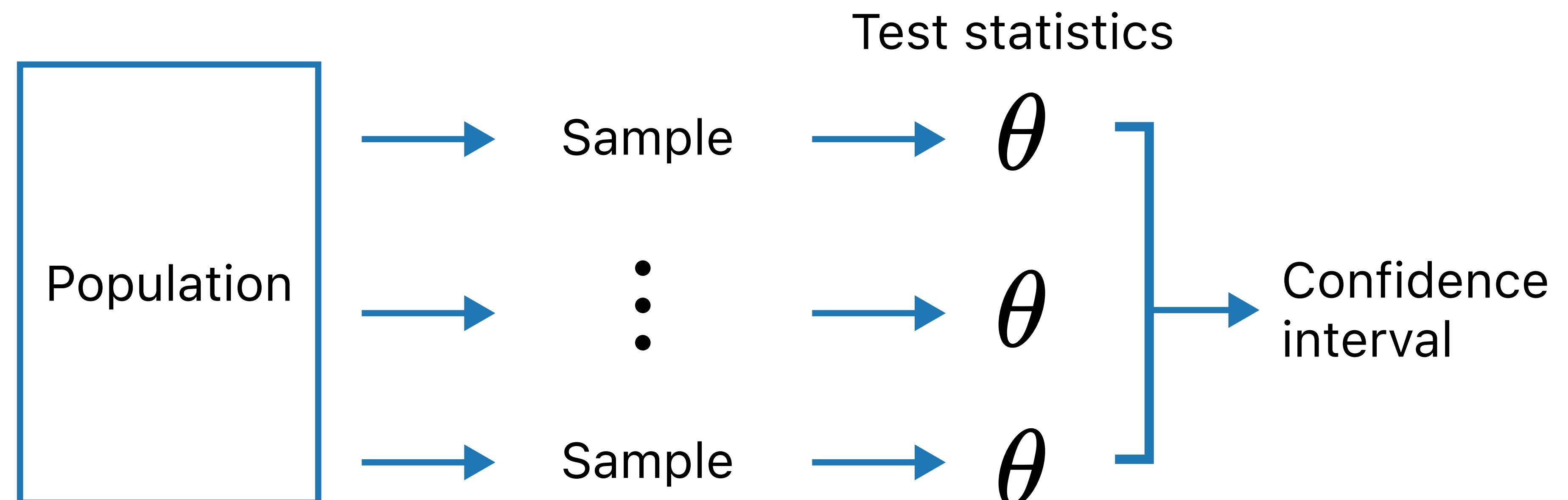
Sources and Types of Uncertainty

Model Uncertainty

How does the data fit together?

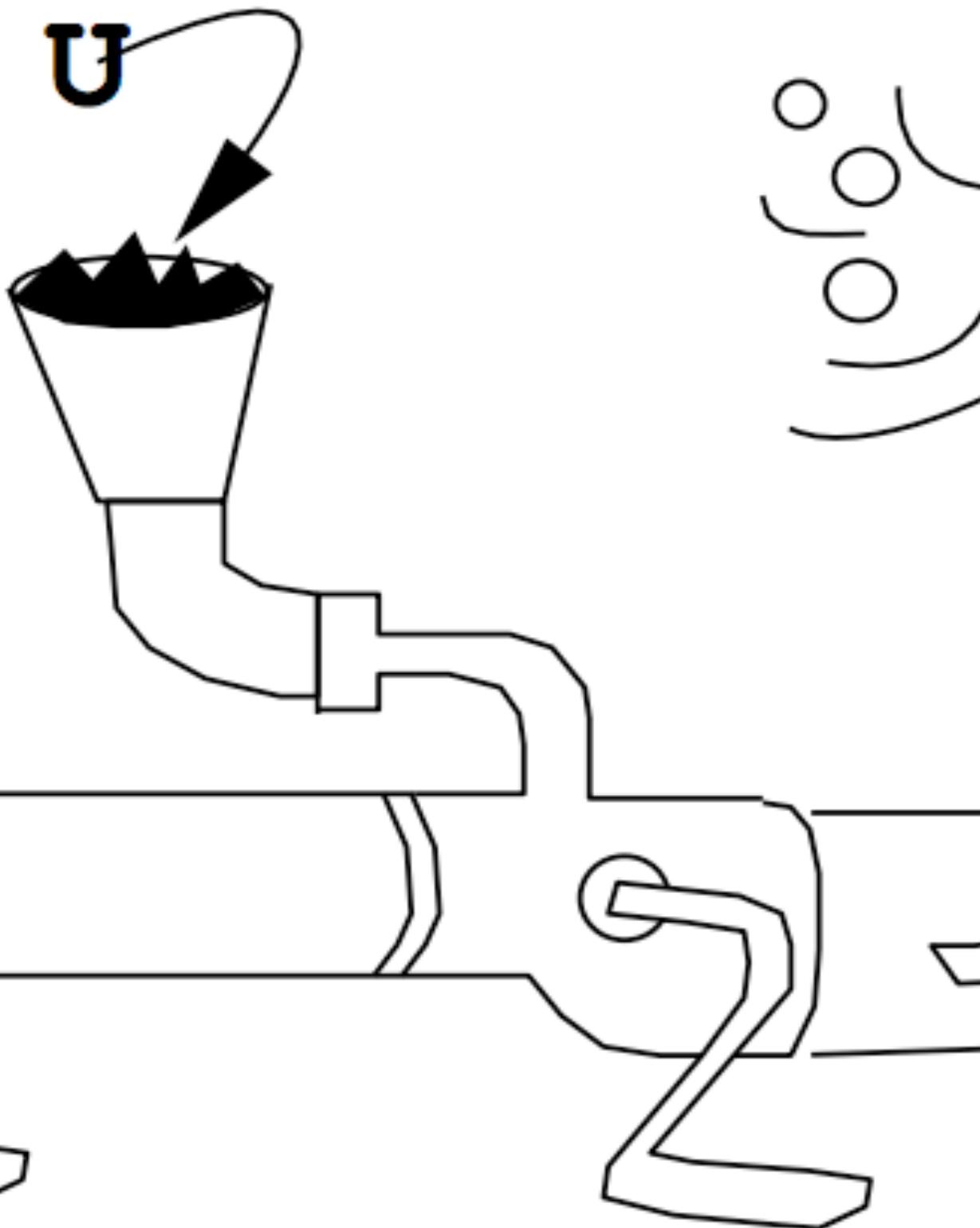


Derive



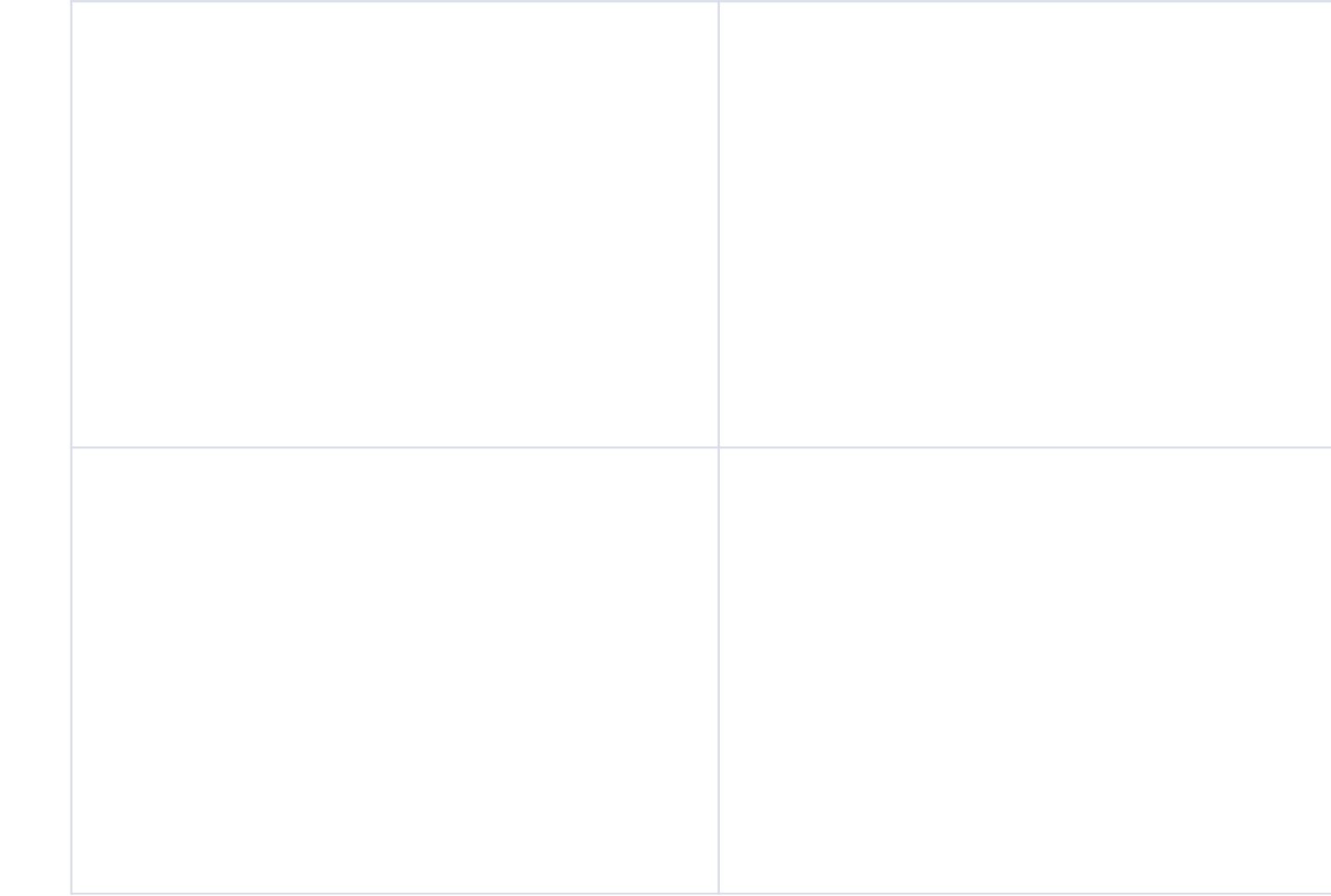
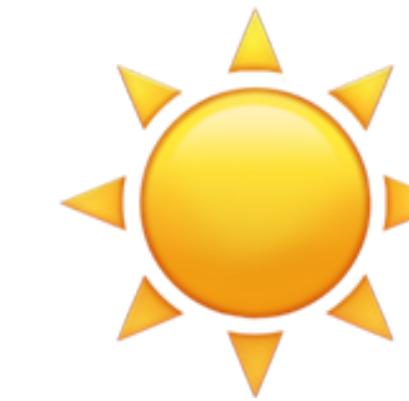
Sources and Types of Uncertainty

visualize



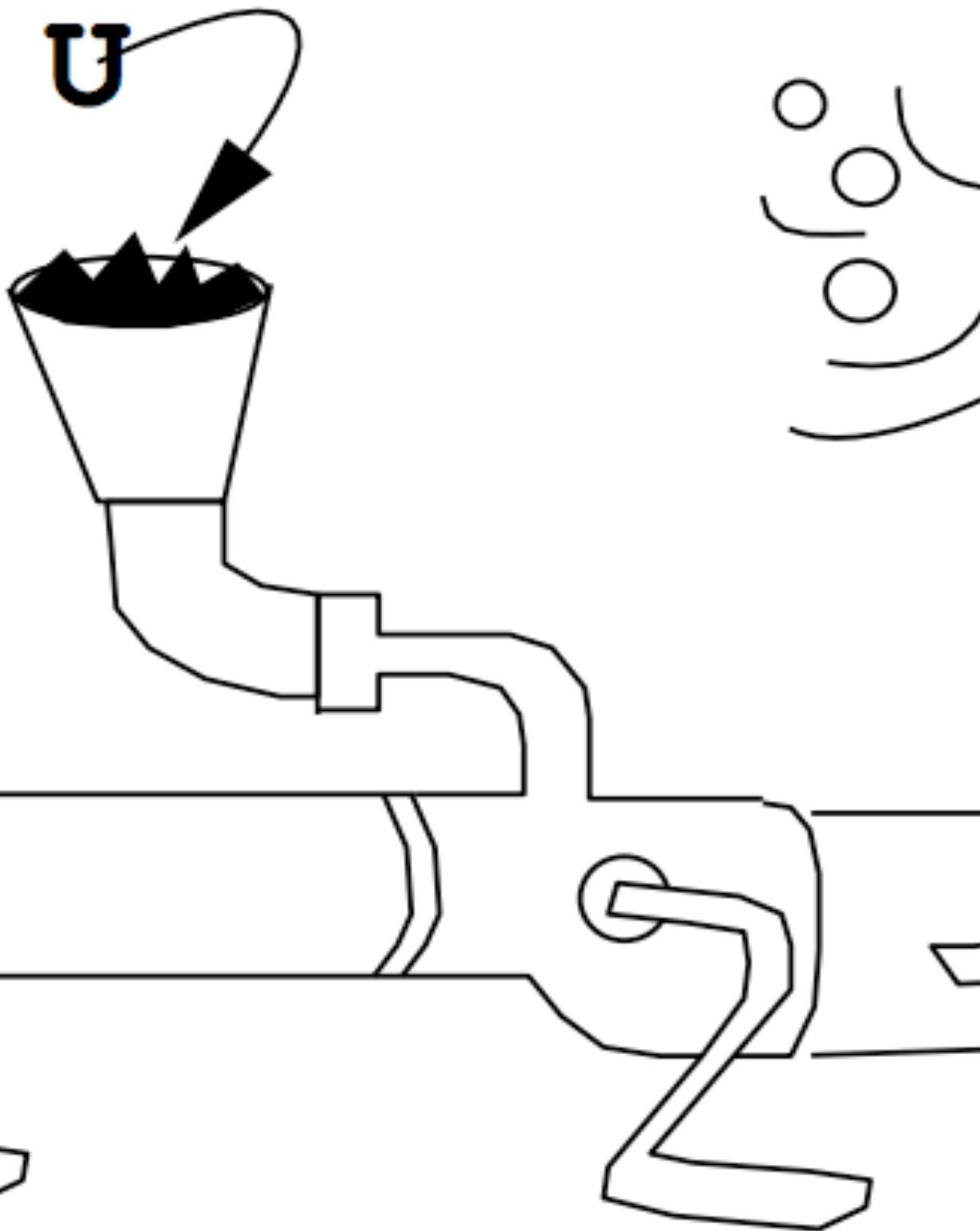
visualize

Decision/Forecast Uncertainty
How do I assess the risk or error?



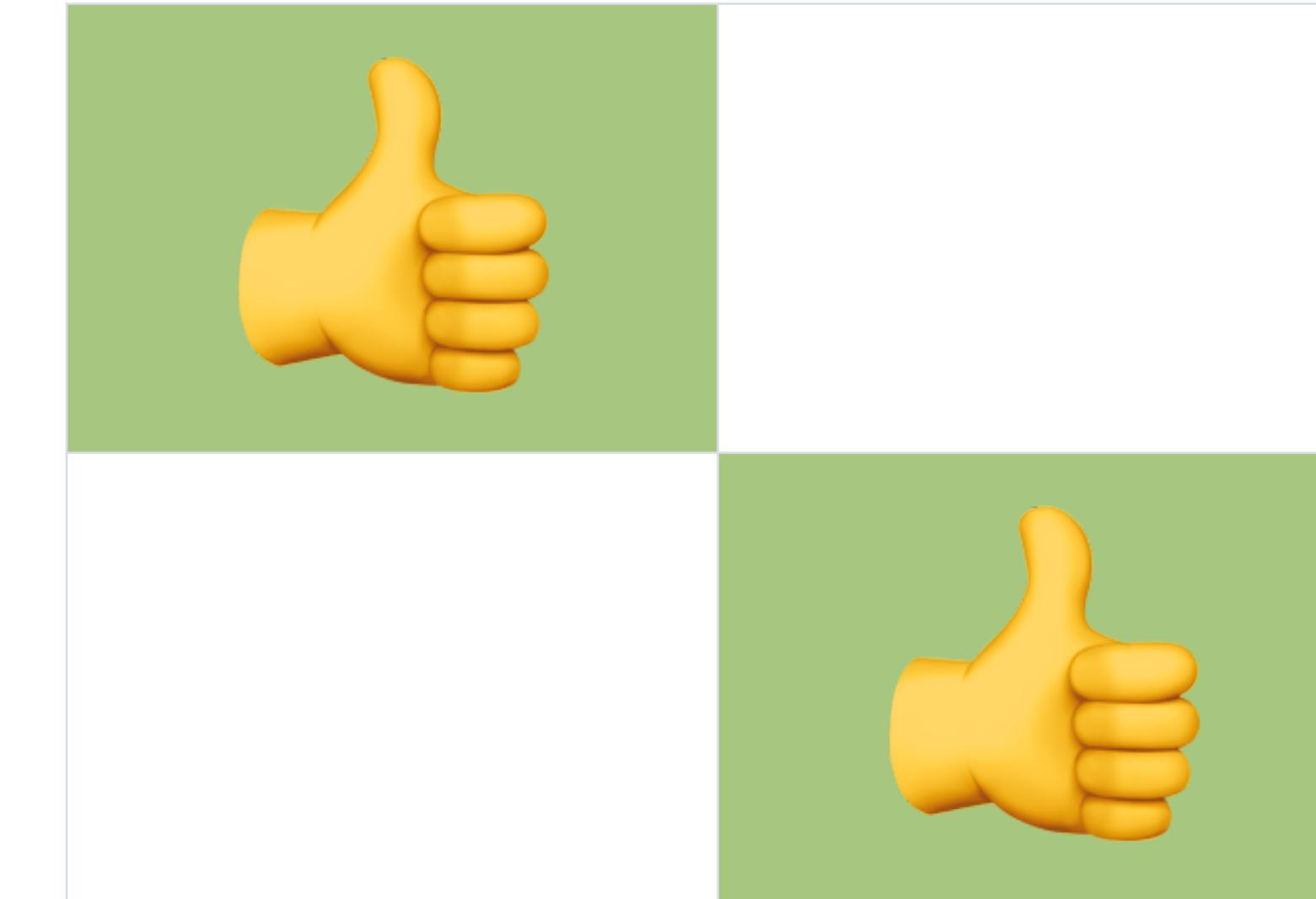
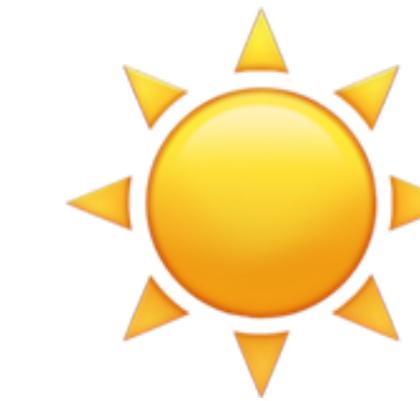
Sources and Types of Uncertainty

visualize



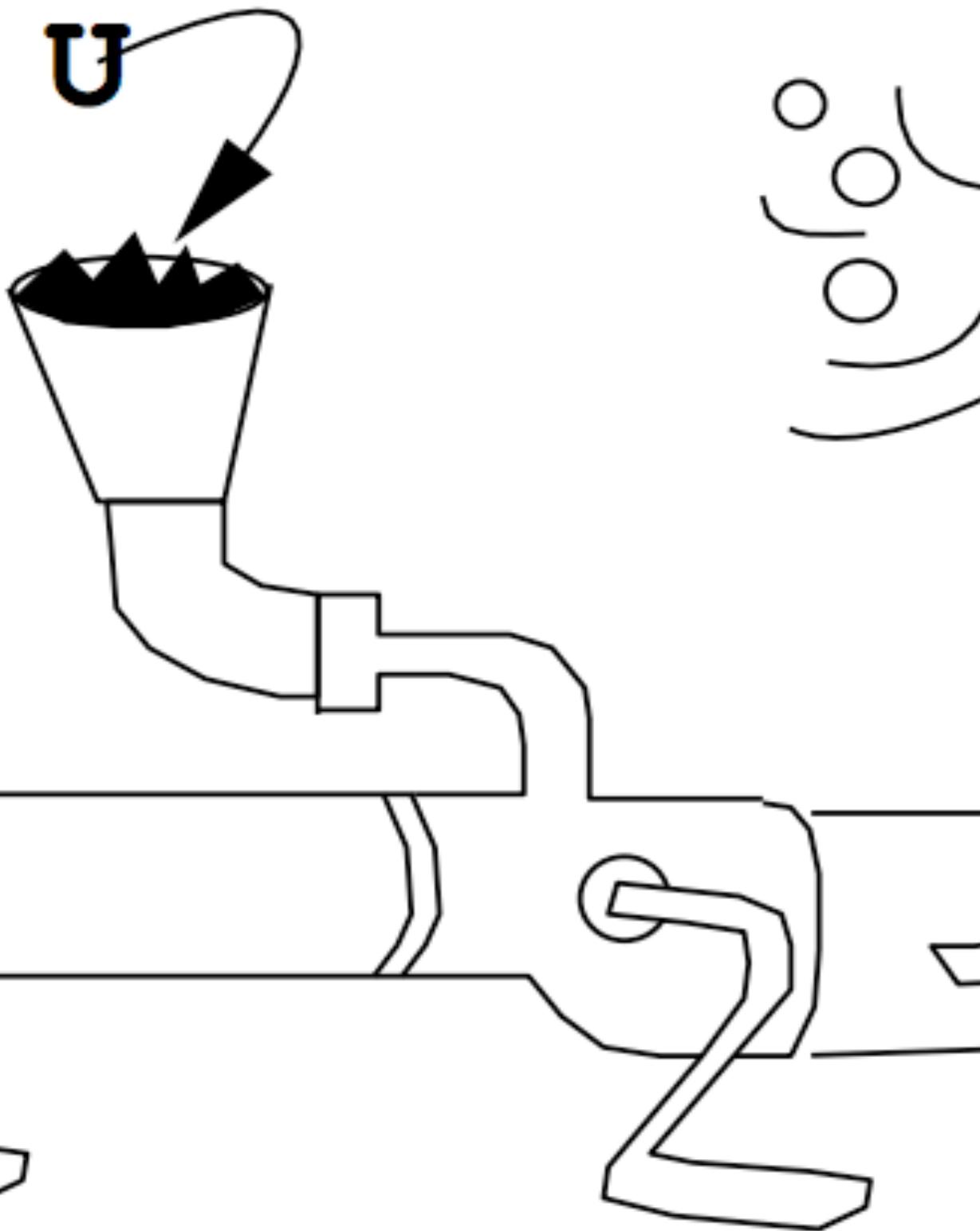
visualize

Decision/Forecast Uncertainty
How do I assess the risk or error?



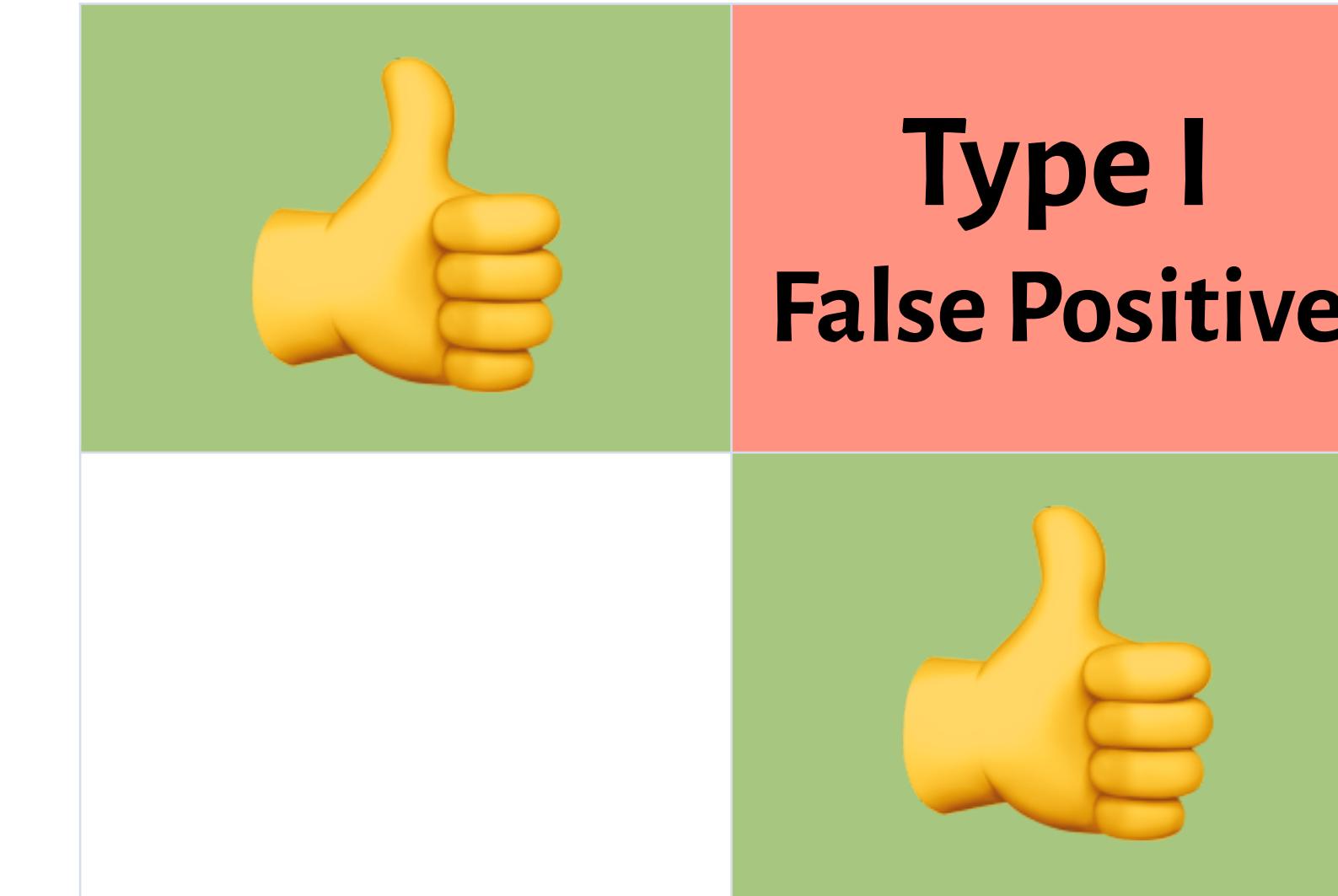
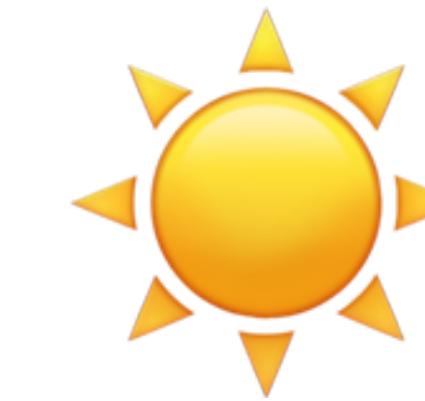
Sources and Types of Uncertainty

visualize



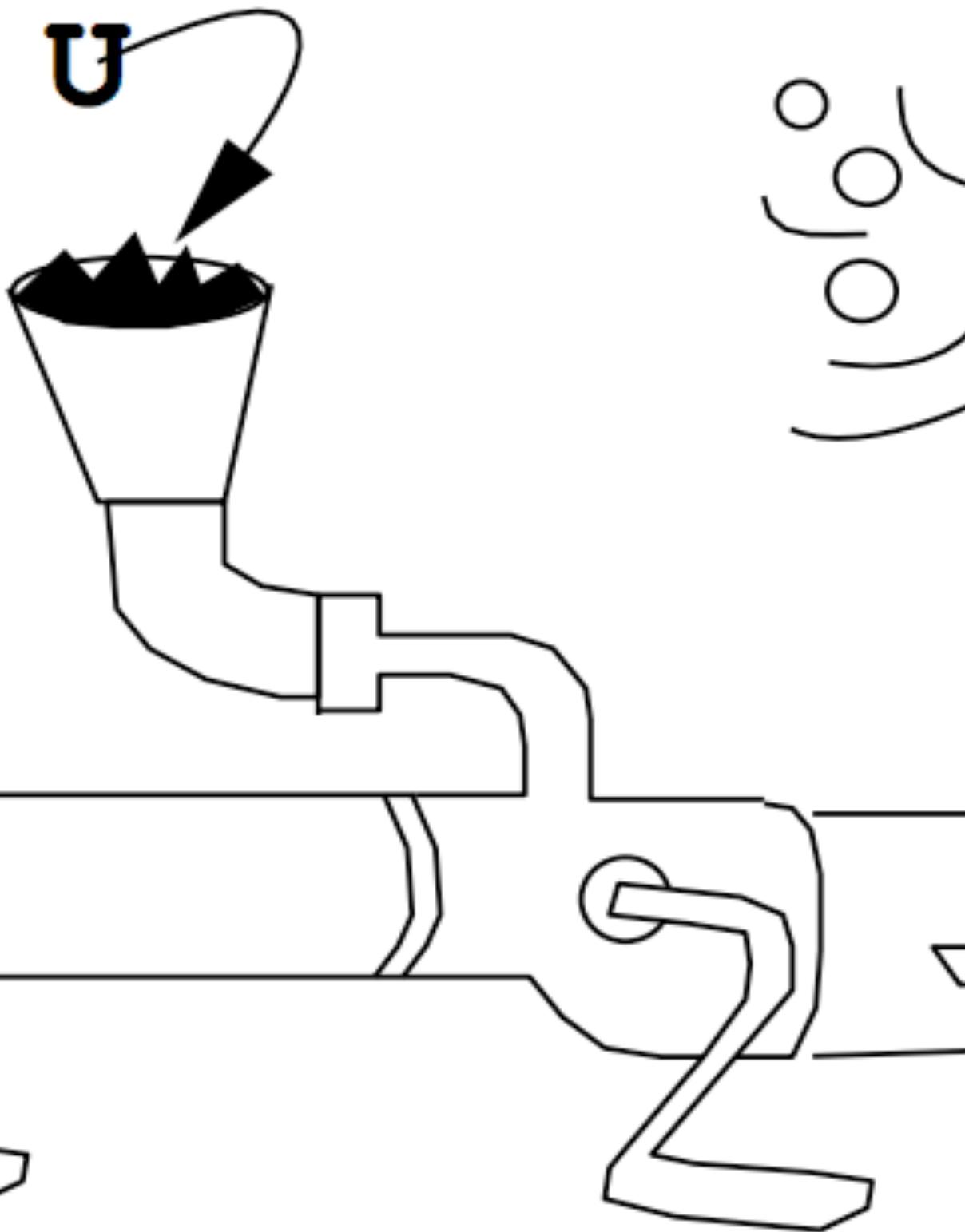
visualize

Decision/Forecast Uncertainty
How do I assess the risk or error?



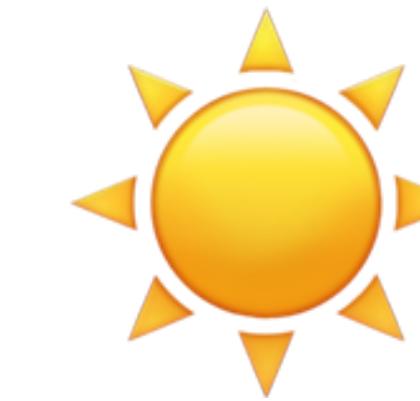
Sources and Types of Uncertainty

visualize



visualize

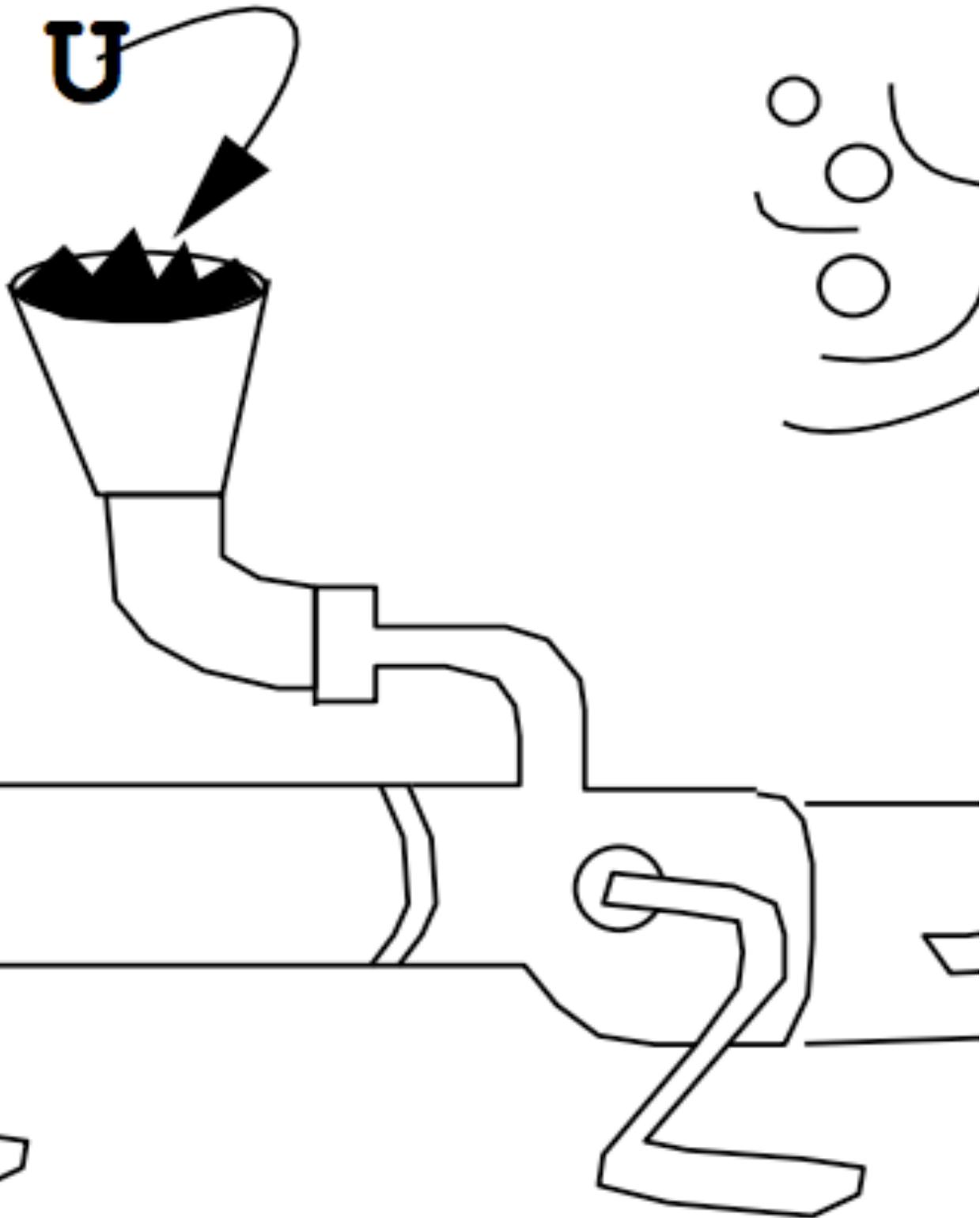
Decision/Forecast Uncertainty
How do I assess the risk or error?



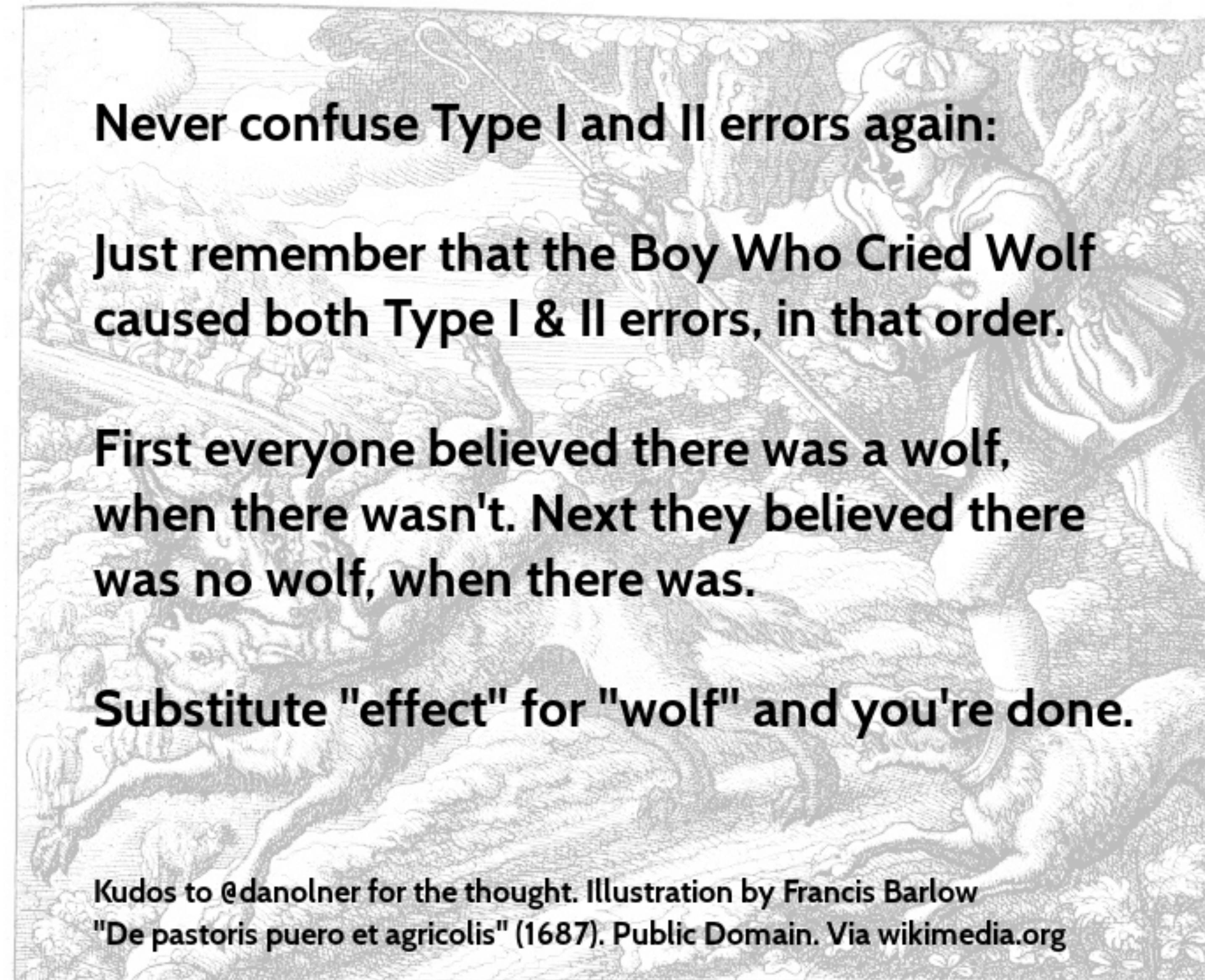
	Type I False Positive
Type II False Negative	

Sources and Types of Uncertainty

Visualize



Visualize



Uncertainty

What does it mean?

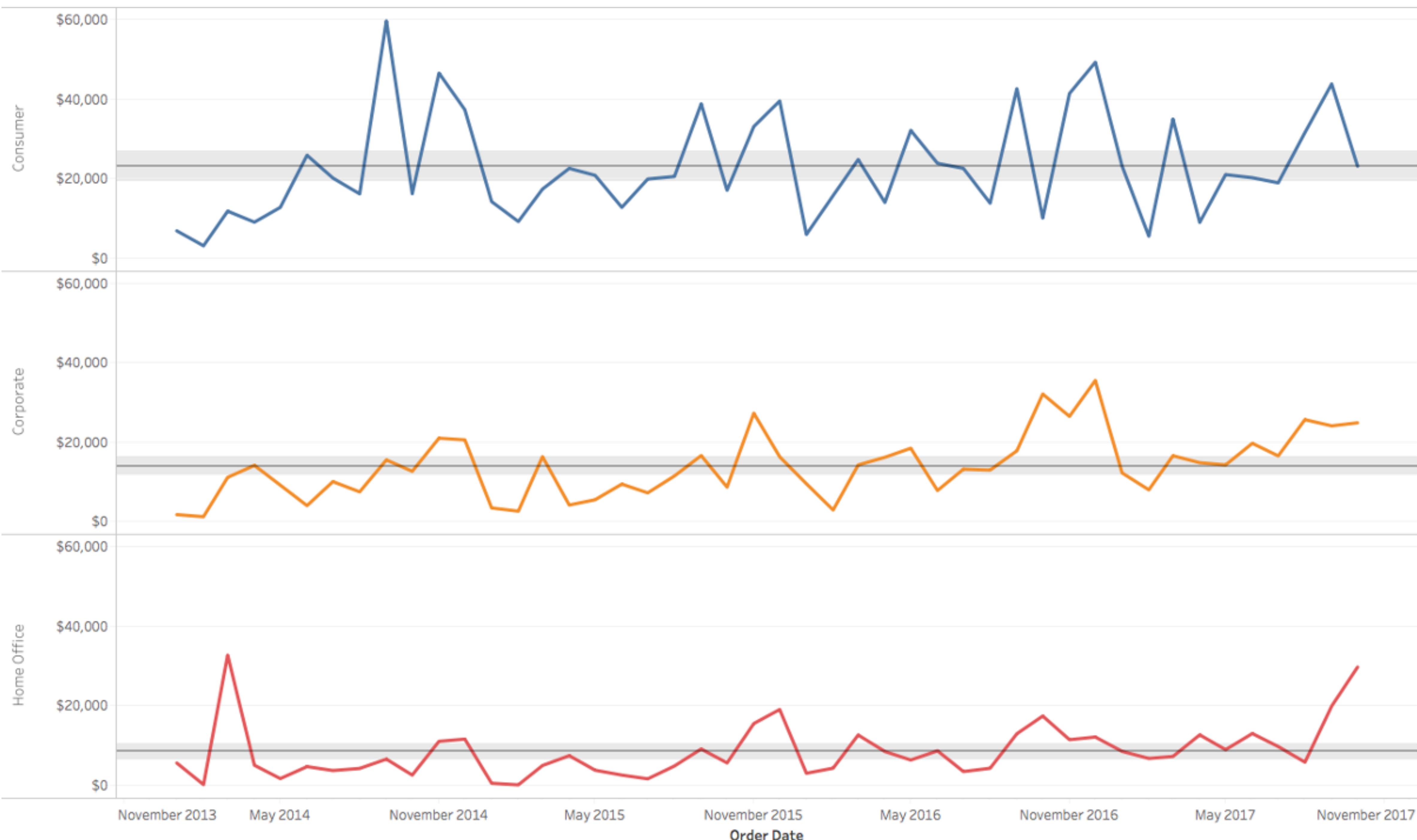
Lots of things!

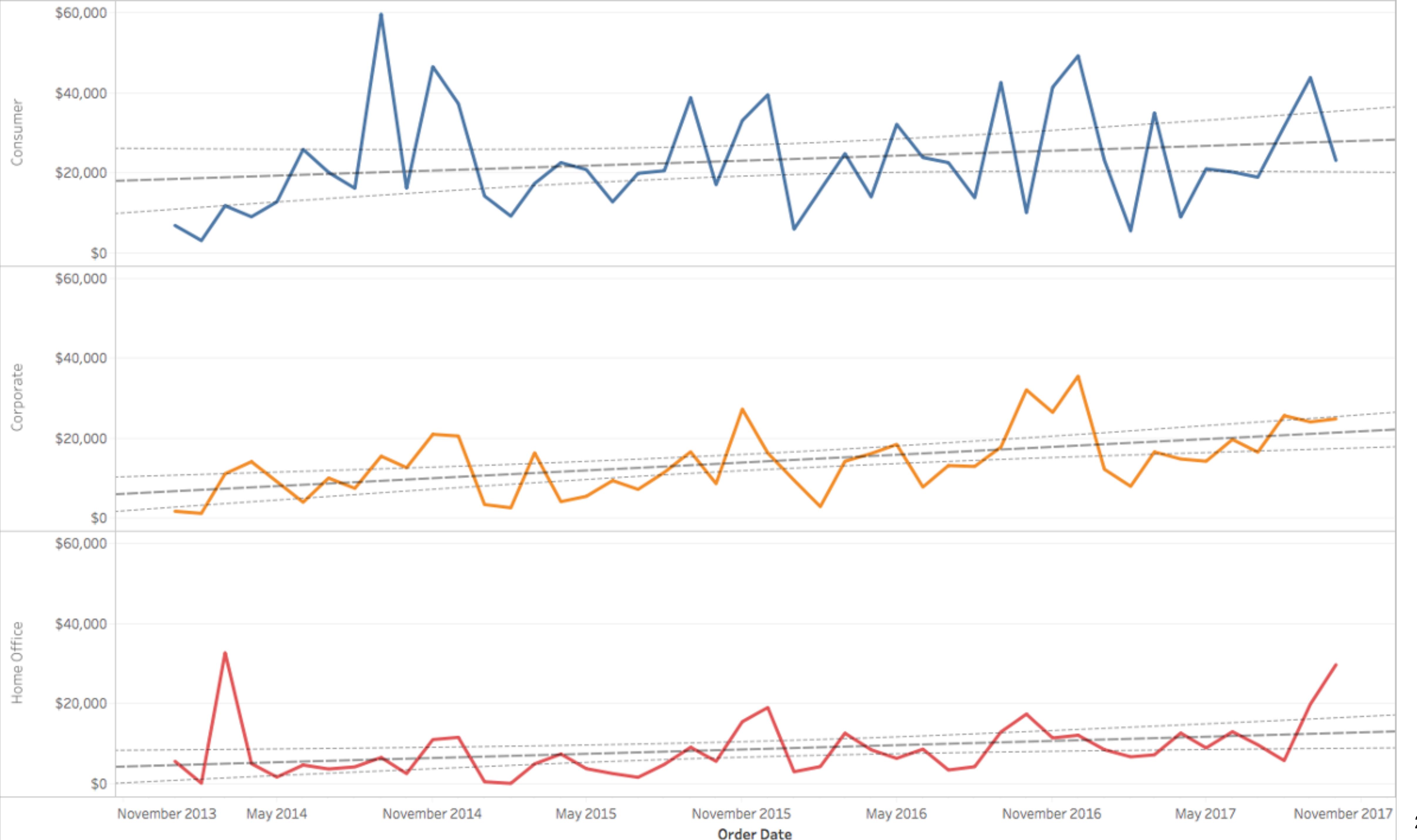
How should I visualize it?

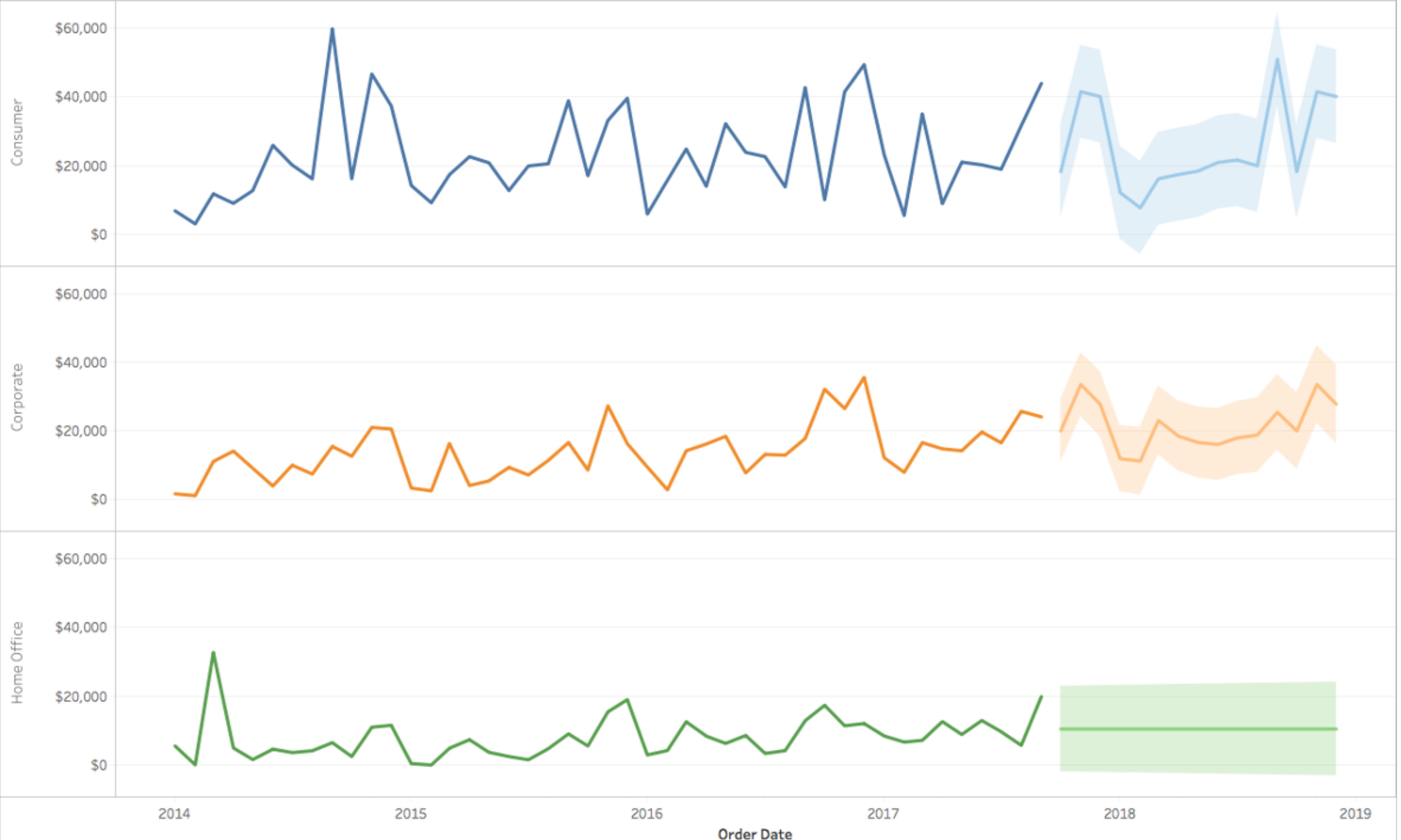
Uncertainty

What does it mean?

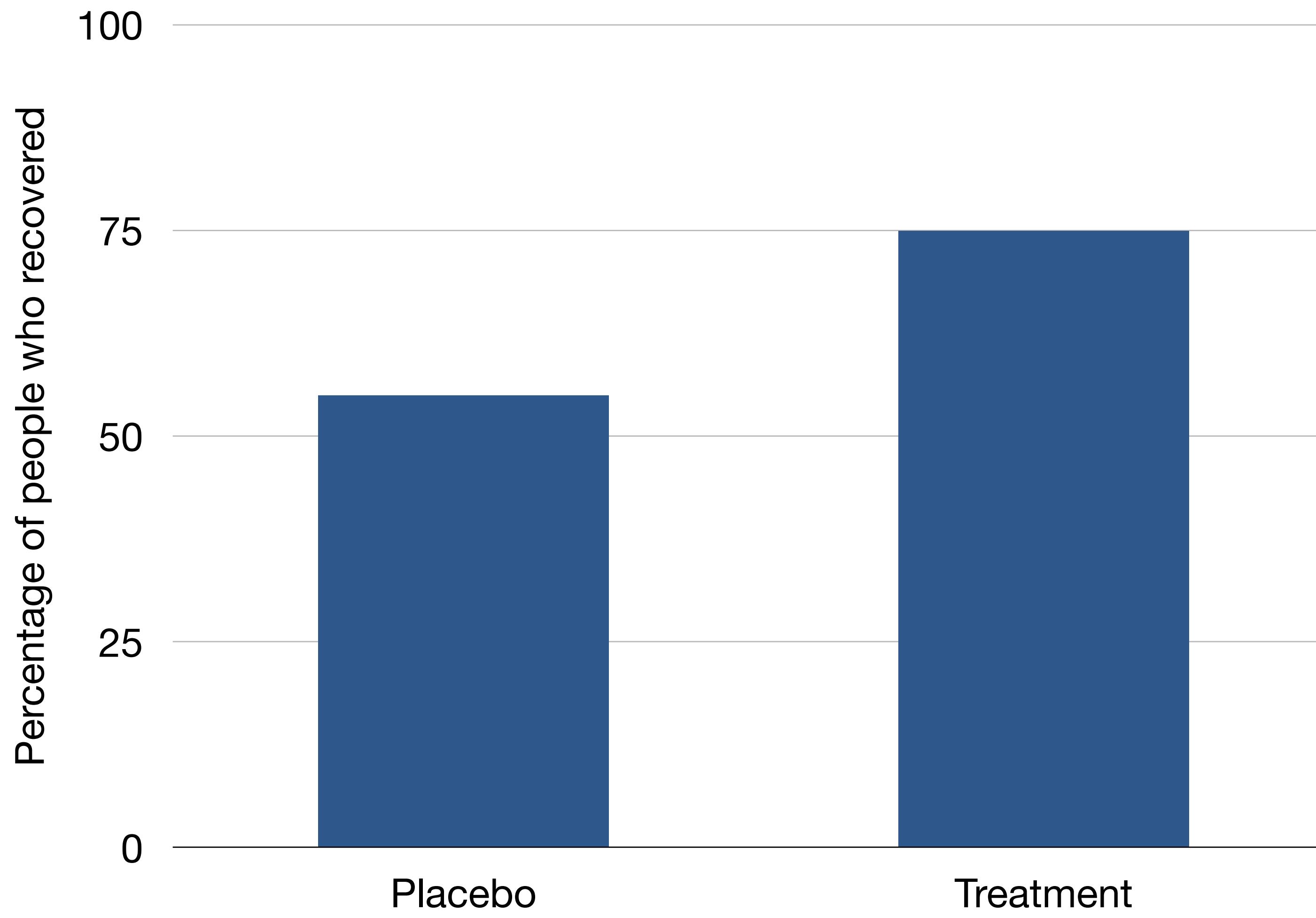
How should I visualize it?



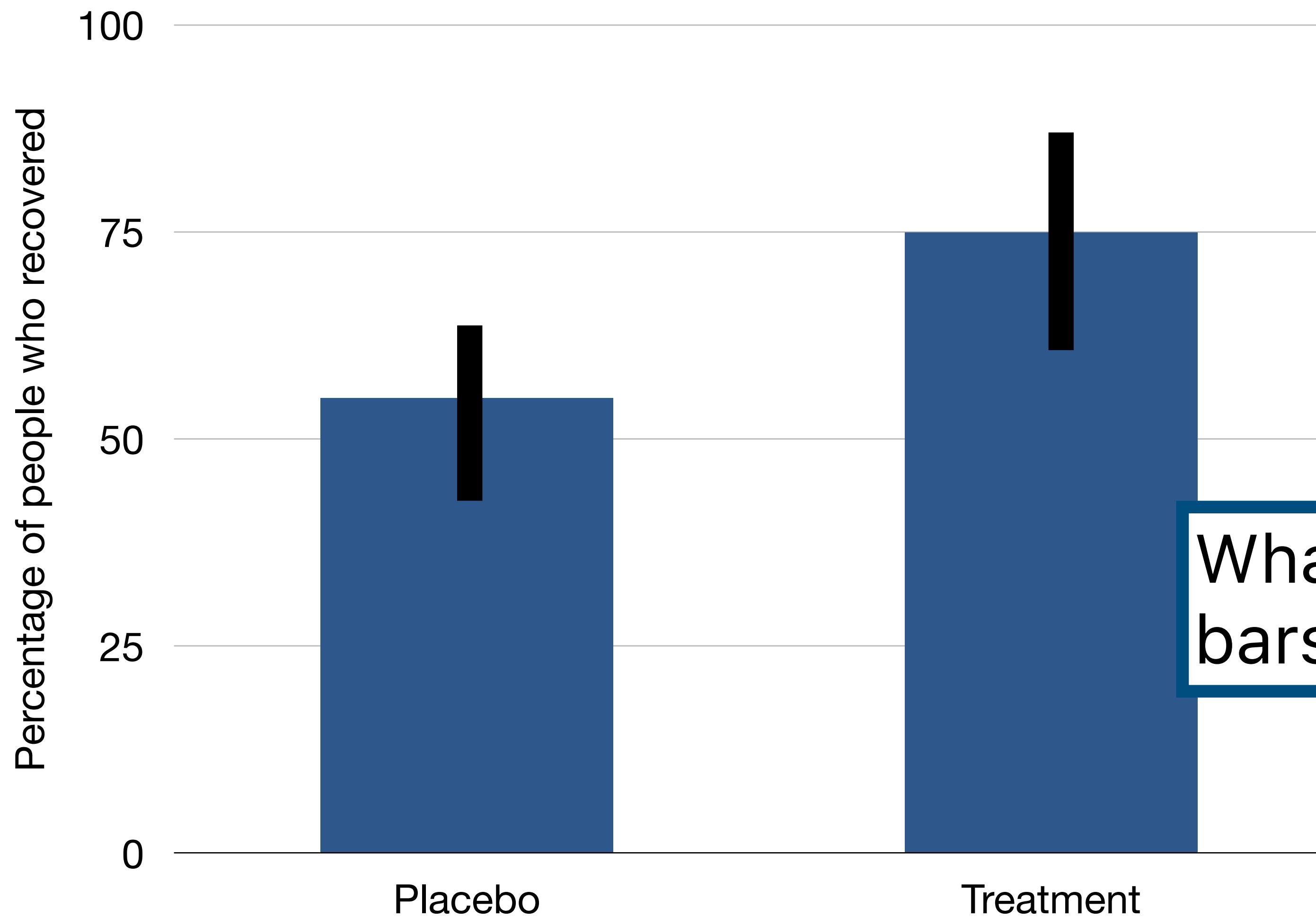




Trial of new medicine

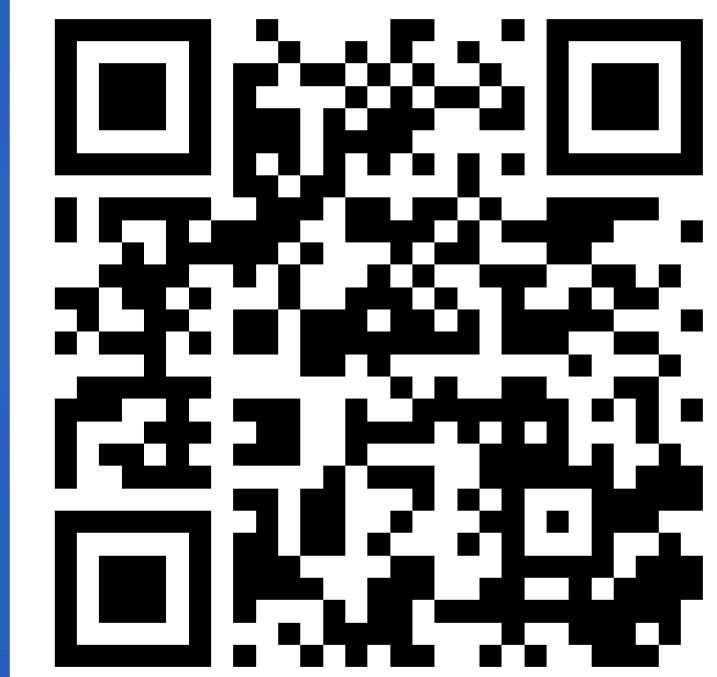


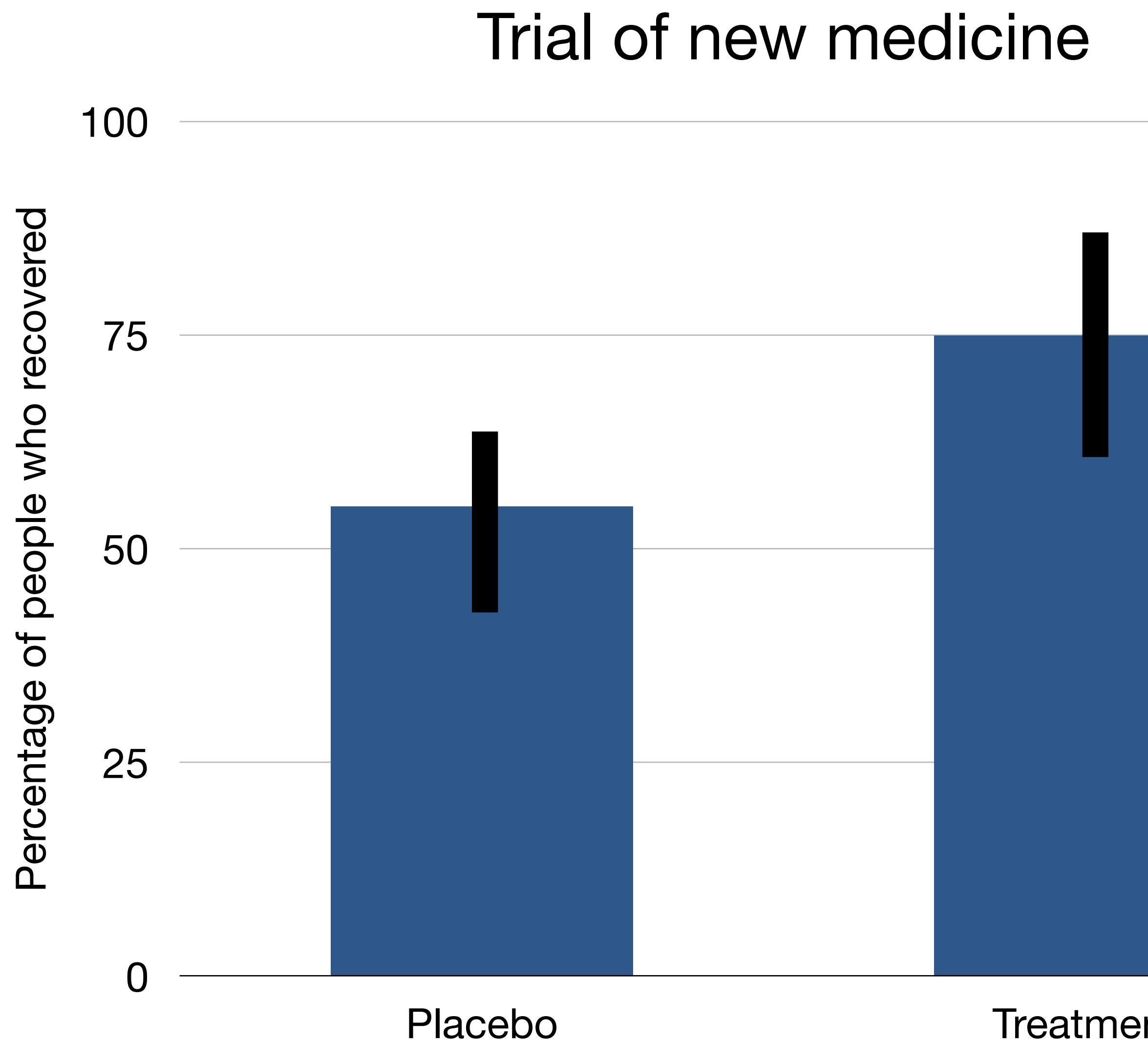
Trial of new medicine



What do you think these error bars are implying?

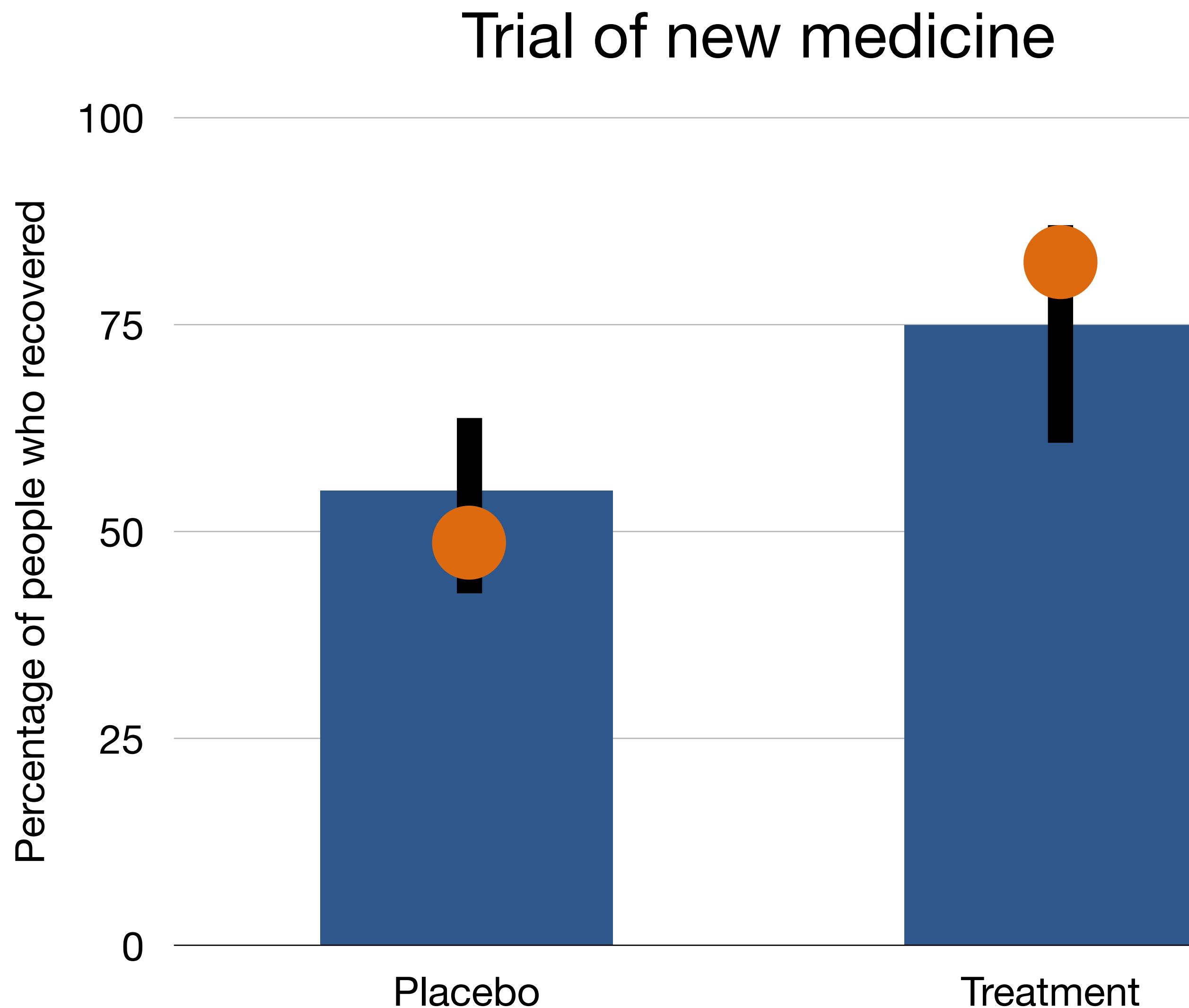
Join at
slido.com
#7283





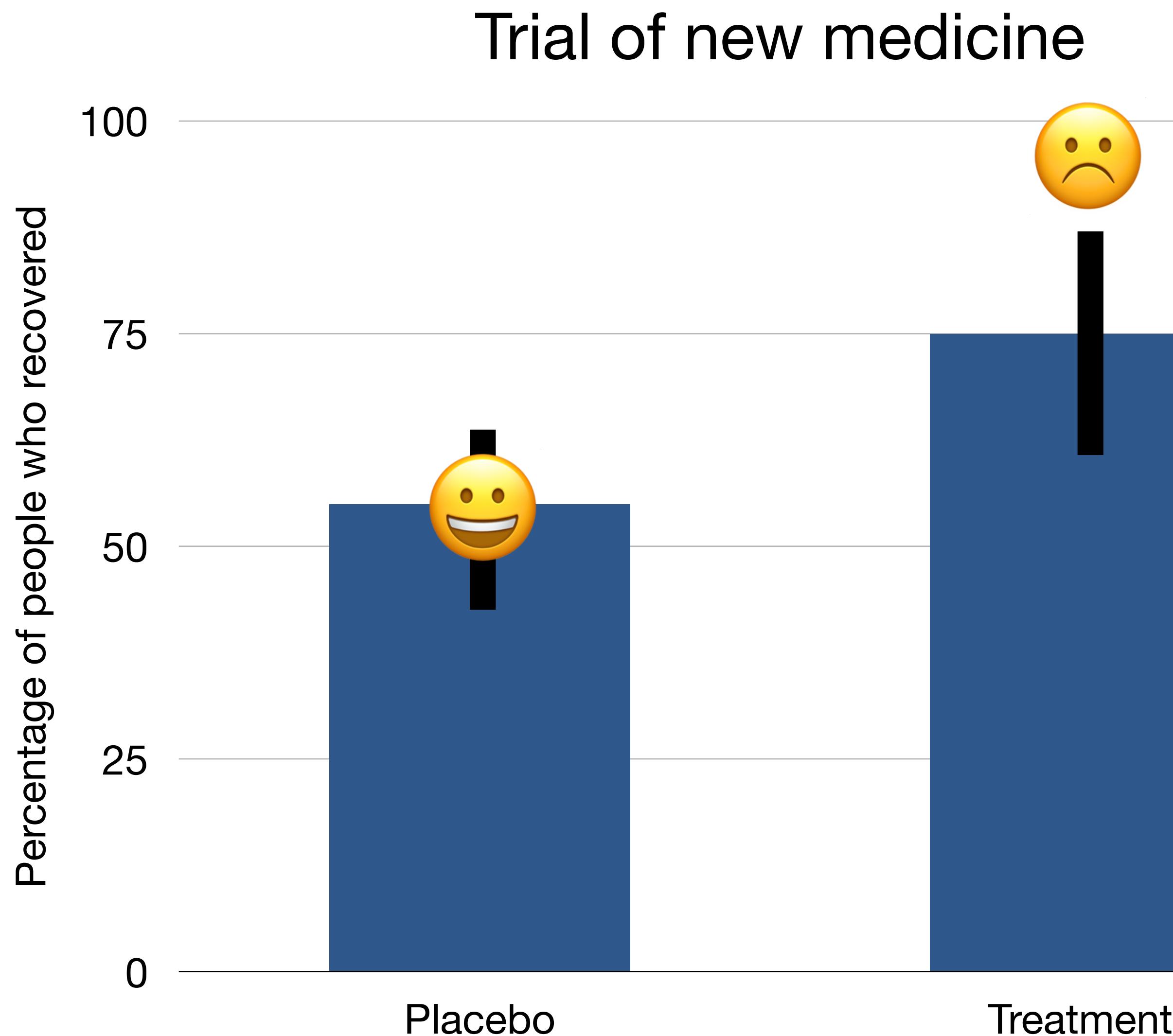
Expressiveness?

✗ Error bars aren't consistently used to visualize the same measure (standard error, IQR, 95% CI, etc.).



Expressiveness?

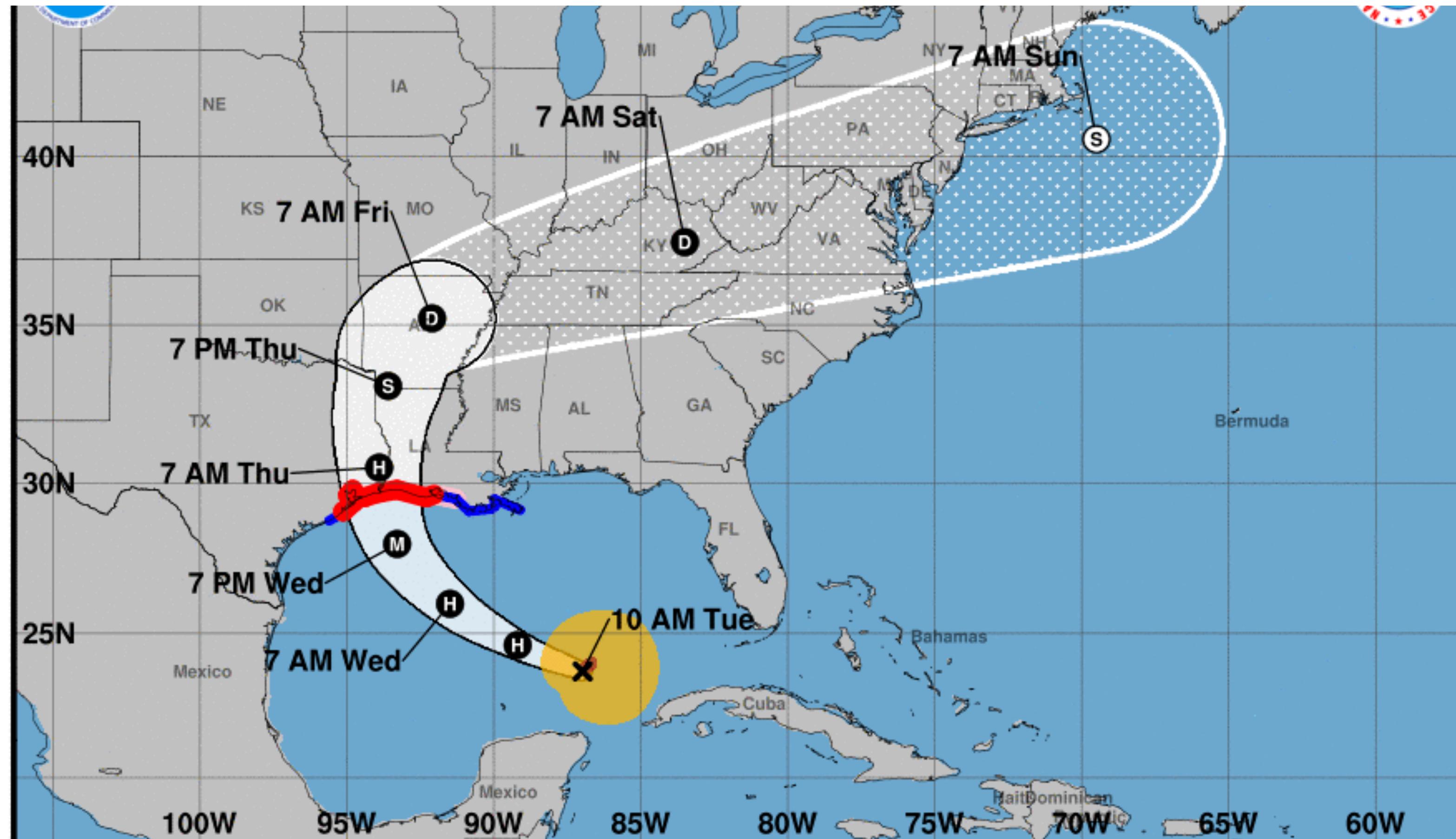
- ✖ Error bars aren't consistently used to visualize the same measure (standard error, IQR, 95% CI, etc.).
- ✖ Within-the-bar bias: people perceive points falling within the bar as more likely than those that lie outside.



Expressiveness?

- ✖ Error bars aren't consistently used to visualize the same measure (standard error, IQR, 95% CI, etc.).
- ✖ Within-the-bar bias: people perceive points falling within the bar as more likely than those that lie outside.
- ✖ Binary bias: people perceive values to either be in or out of the margins of error.

[Newman & Scholl, 2012]
[Correll & Gleicher, 2014]



Hurricane Laura

Tuesday August 25, 2020

10 AM CDT Advisory 23

NWS National Hurricane Center

Current information: x

Center location 23.7 N 87.0 W
Maximum sustained wind 75 mph
Movement WNW at 16 mph

Forecast positions:

● Tropical Cyclone ○ Post/Potential TC
Sustained winds: D < 39 mph
S 39-73 mph H 74-110 mph M > 110 mph

Potential track area:



Watches:



Warnings:



Current wind extent:



What is being visualized?

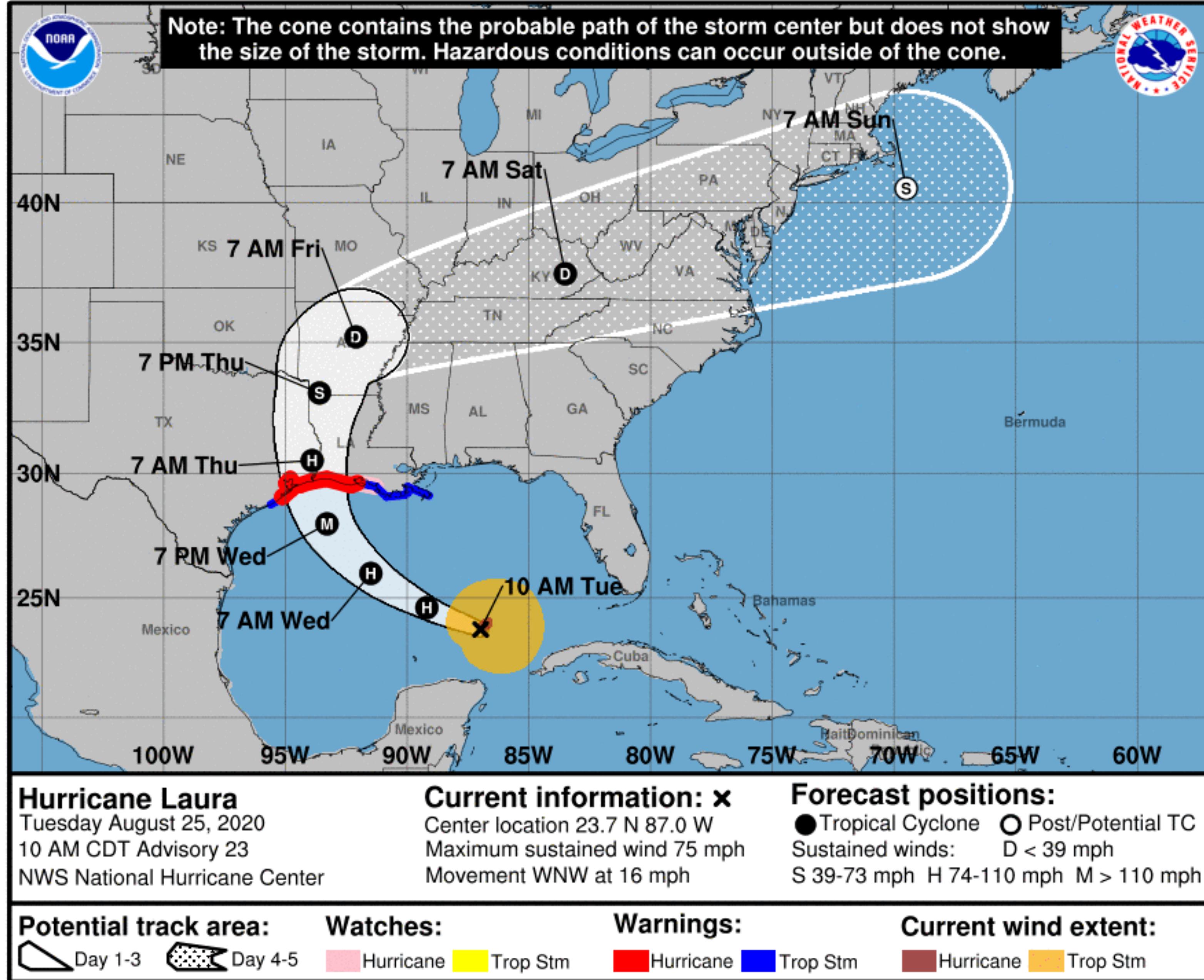
What are the strengths and weaknesses of this visualization?

Join at

slido.com

#7283





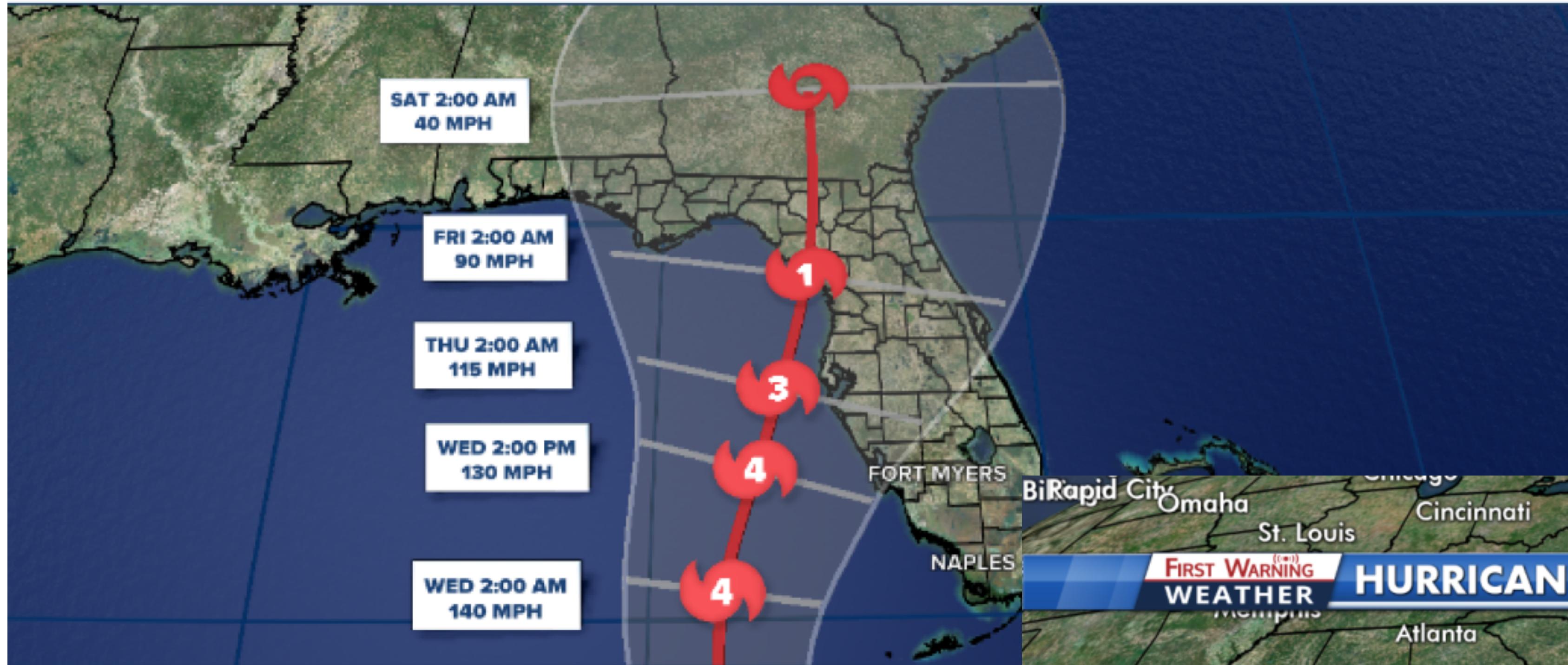
What is being visualized?

What are the strengths and weaknesses of this visualization?

FOX 4

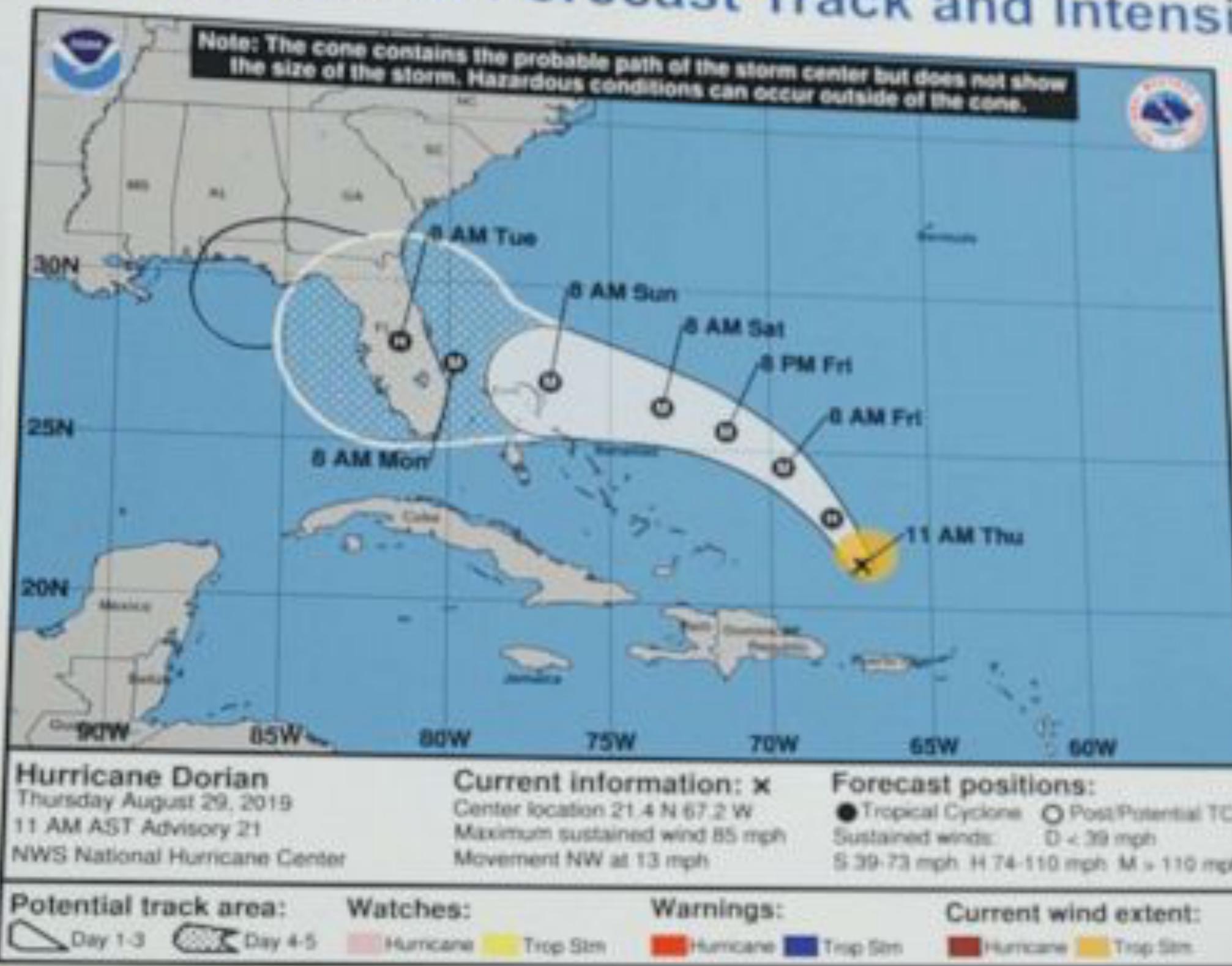
HURRICANE IAN

5:00 AM ADVISORY





Hurricane Dorian Forecast Track and Intensity

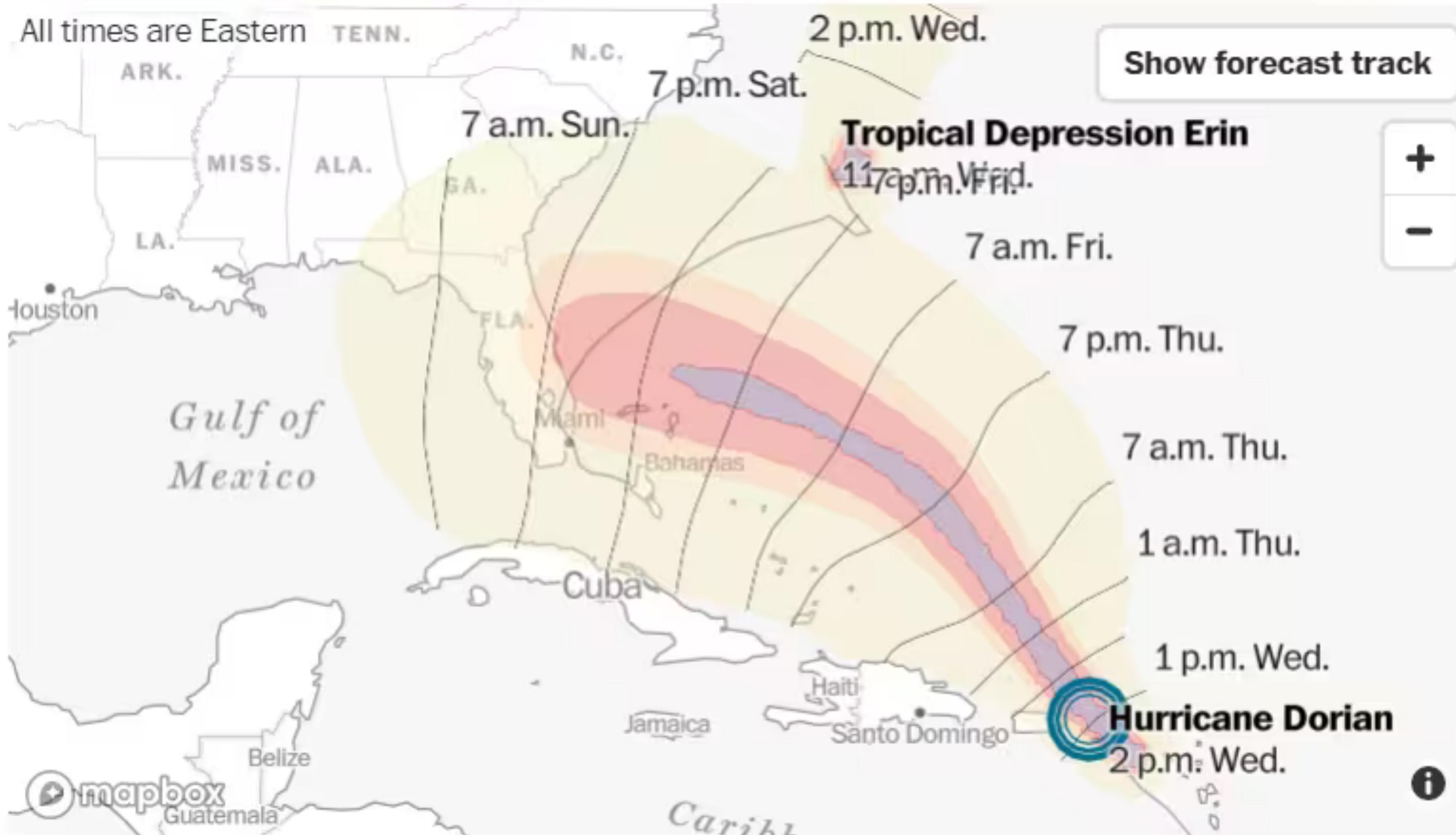


Five-day chance of tropical-storm-force winds

5 50 70 90%

Current extent of tropical-storm-force winds

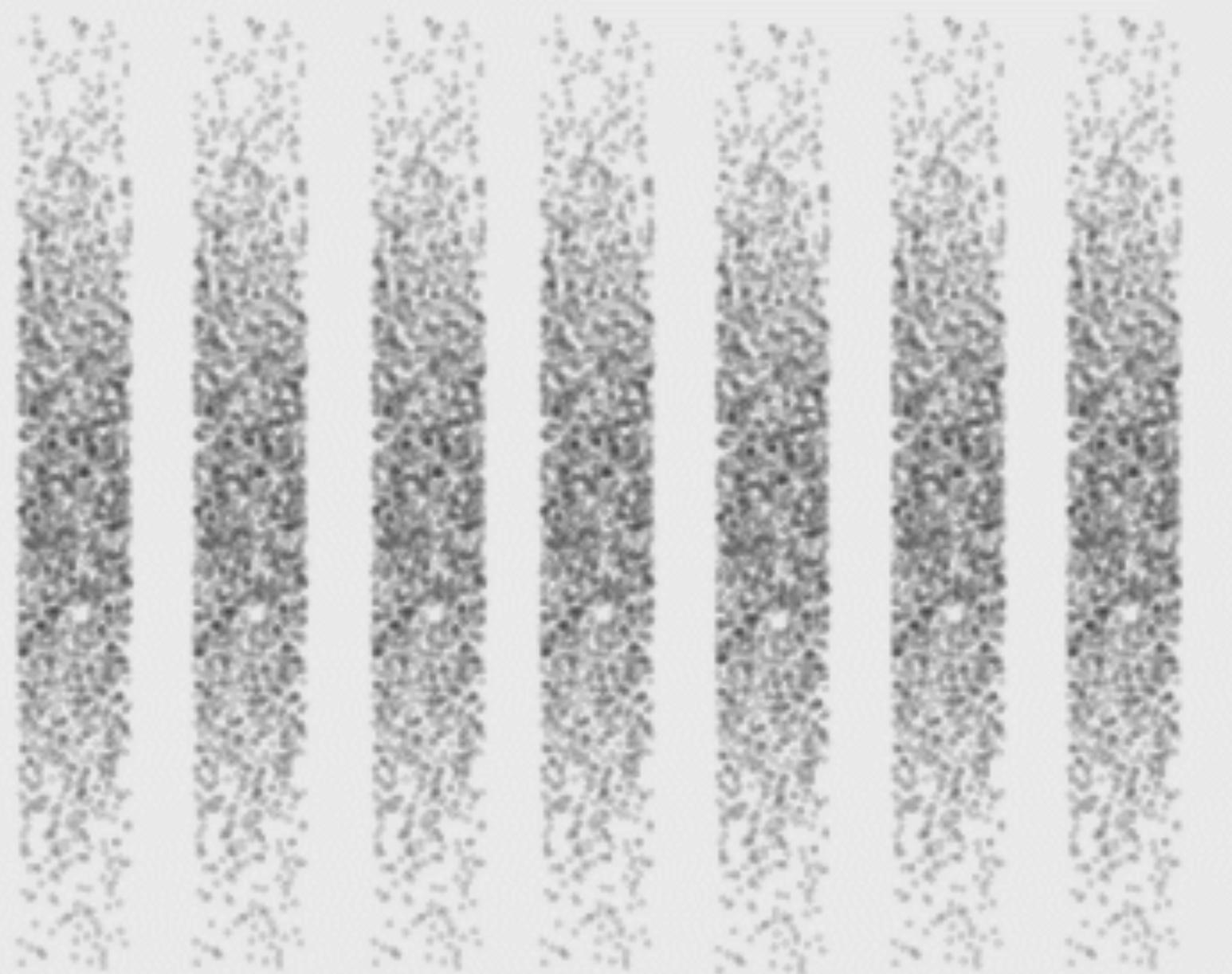
- Major hurricane (>110 mph)
- Hurricane (74-110 mph)
- Tropical storm (39-73 mph)
- Tropical depression (<39 mph)



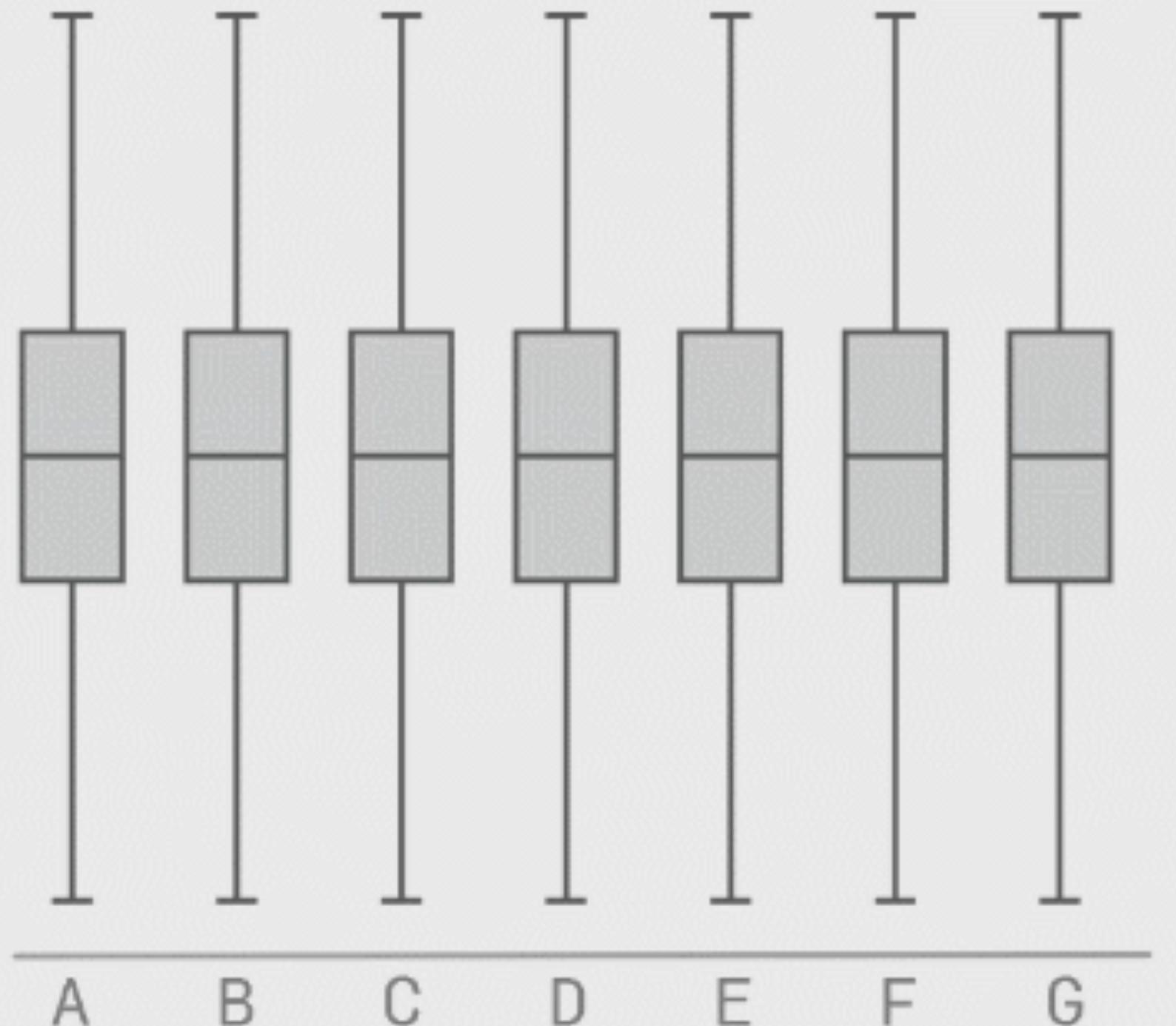
Source: National Weather Service. Note: Impact lines represent the earliest reasonable arrival time of tropical-storm-force winds.

For uncertainty, use **visual variables**
instead of visualizing point estimates

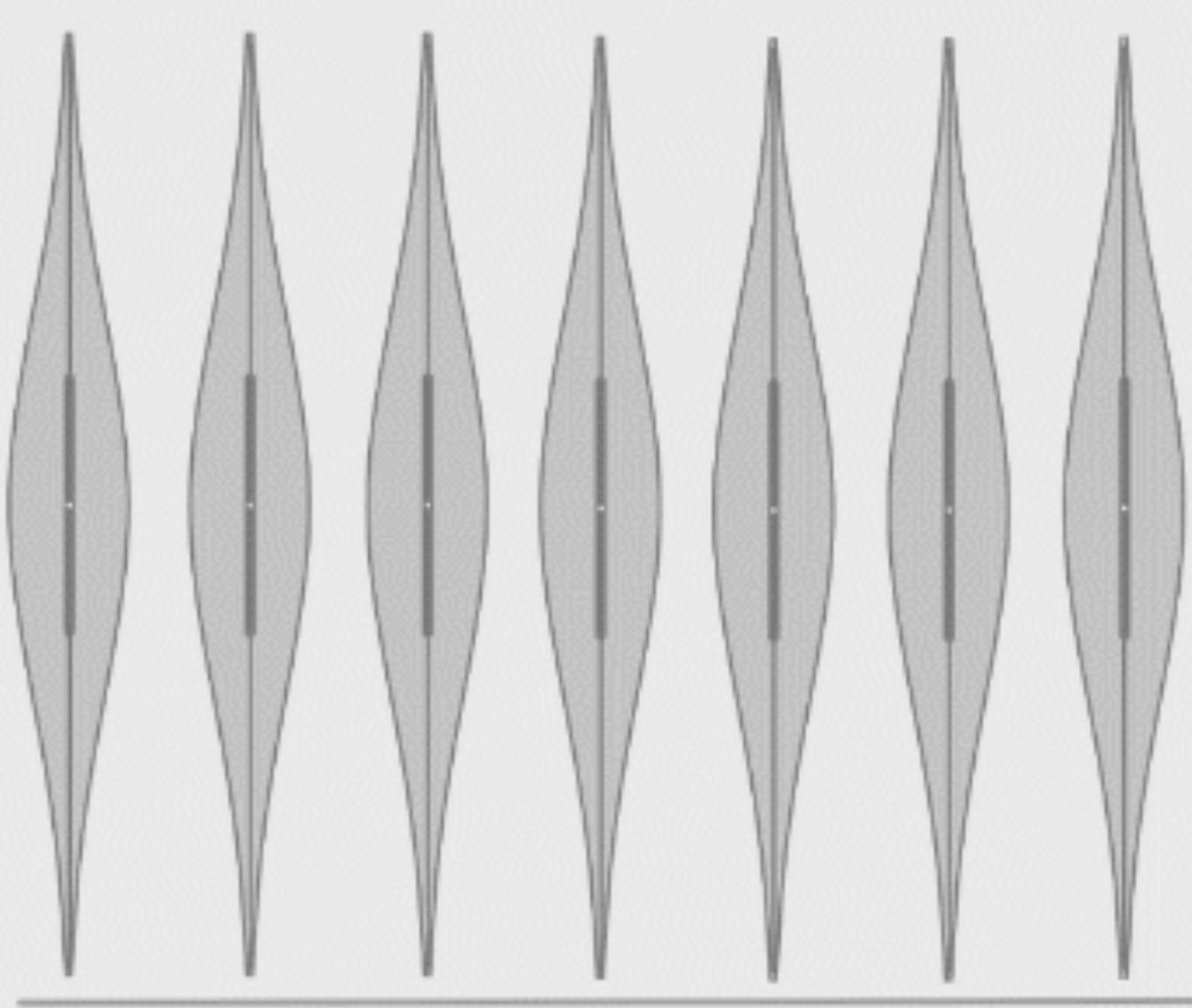
Raw Data



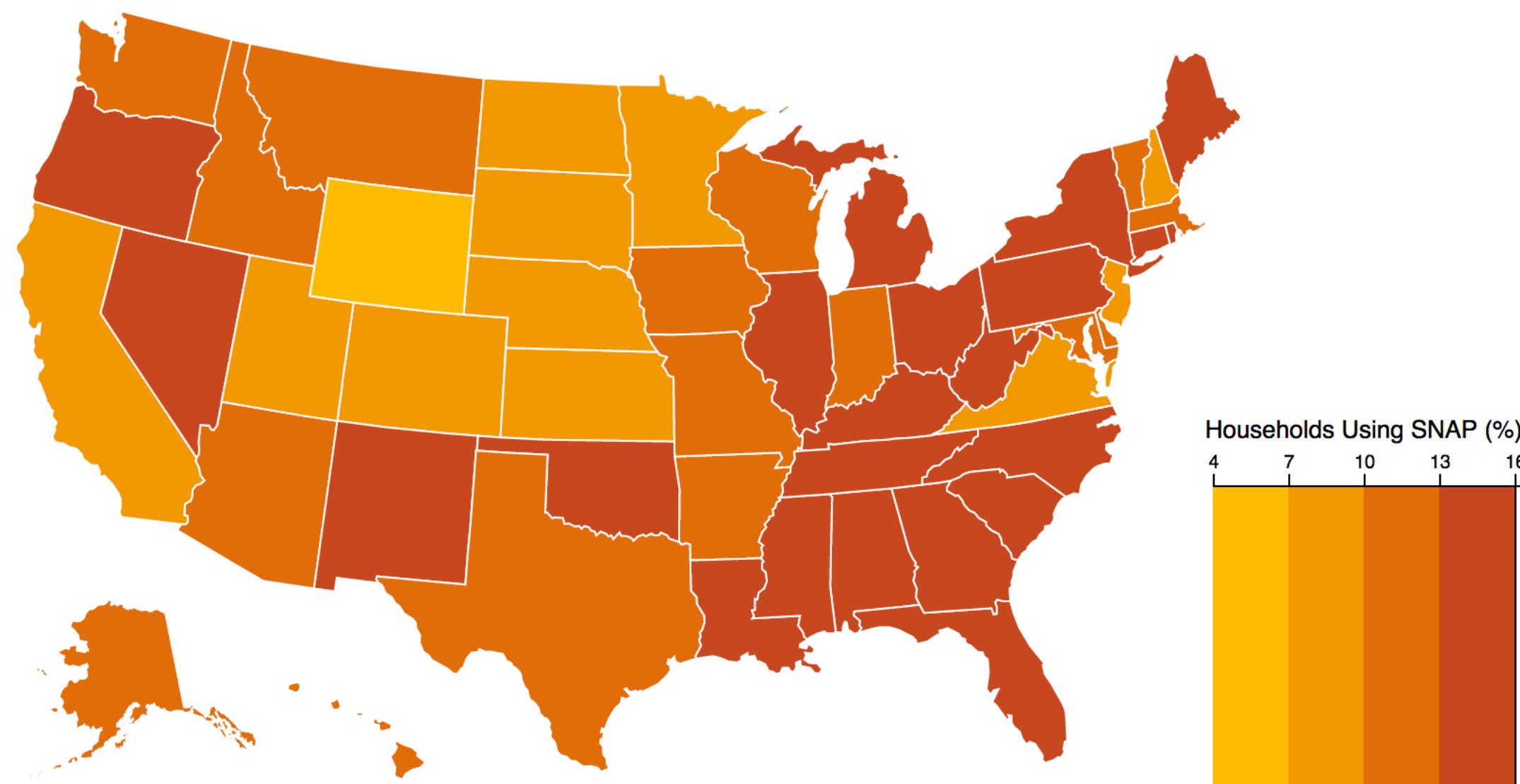
Box-plot of the Data



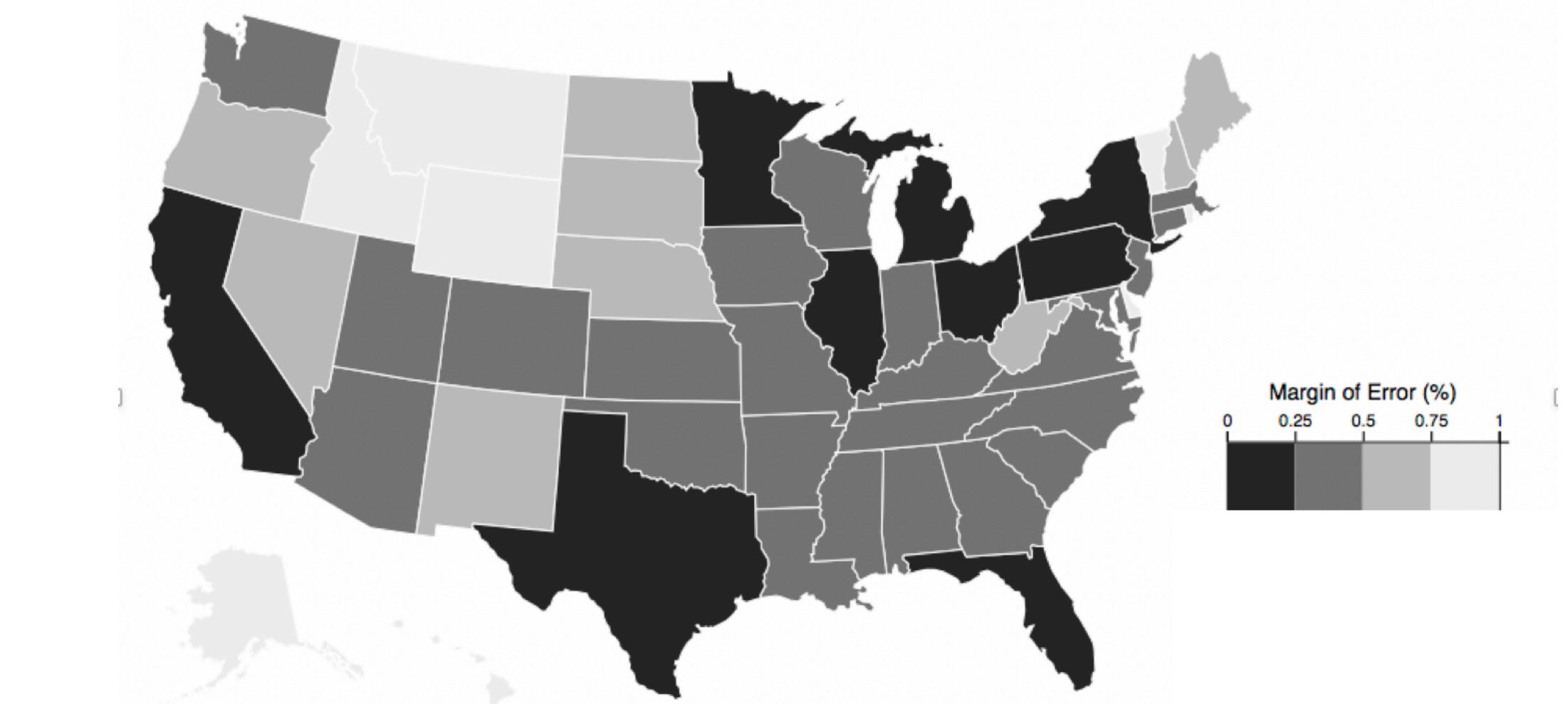
Violin-plot of the Data



For uncertainty, use **visual variables**
instead of visualizing point estimates

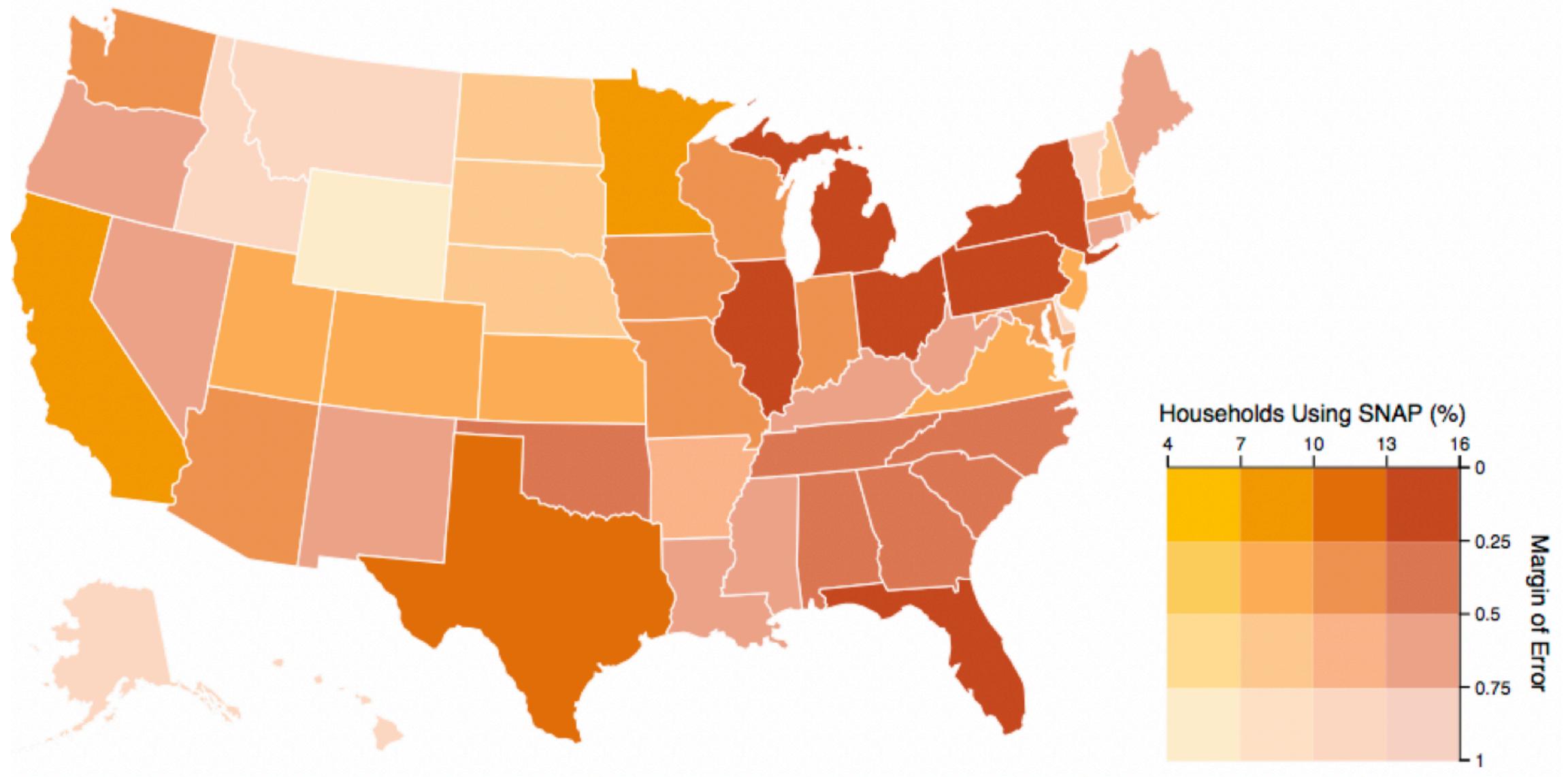


Data Map

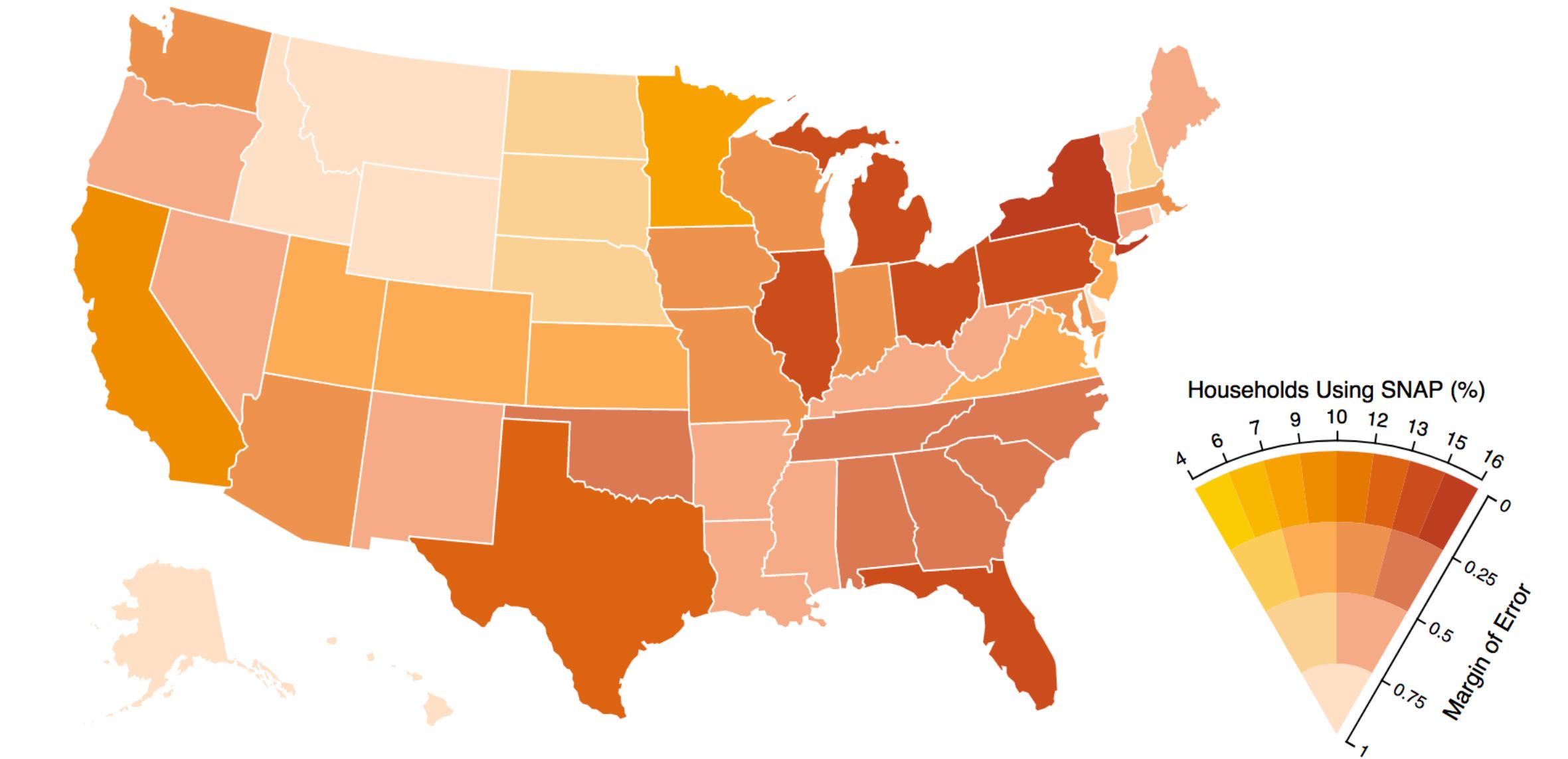


Uncertainty Map

For uncertainty, use **visual variables**
instead of visualizing point estimates

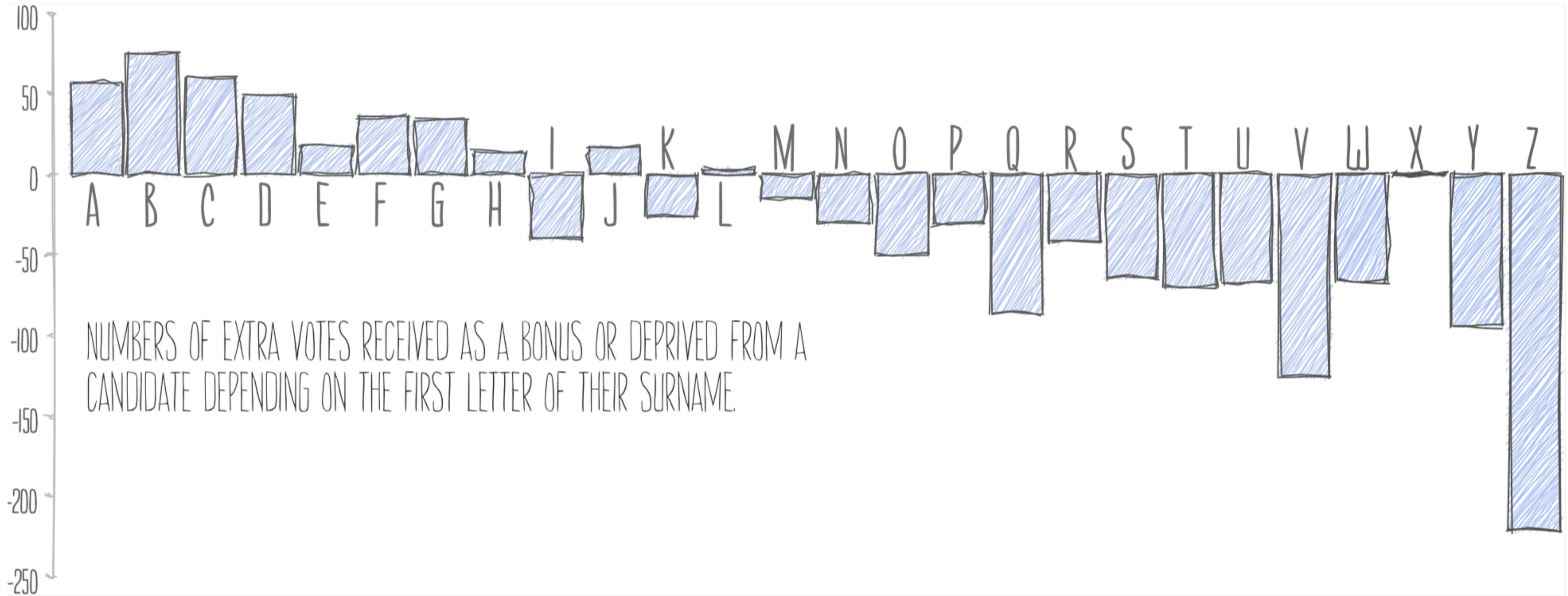


Bivariate Map (Data + Uncertainty)



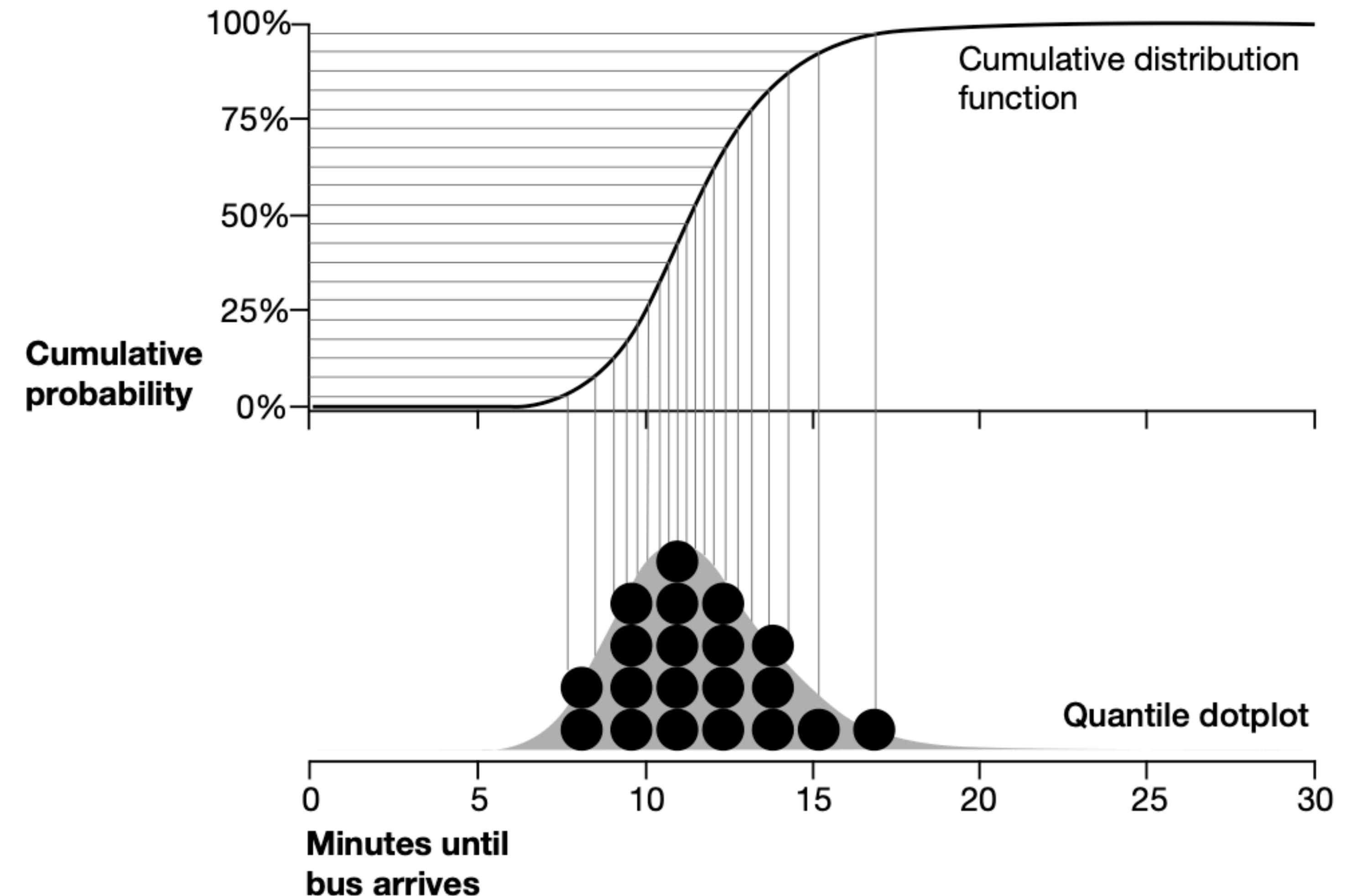
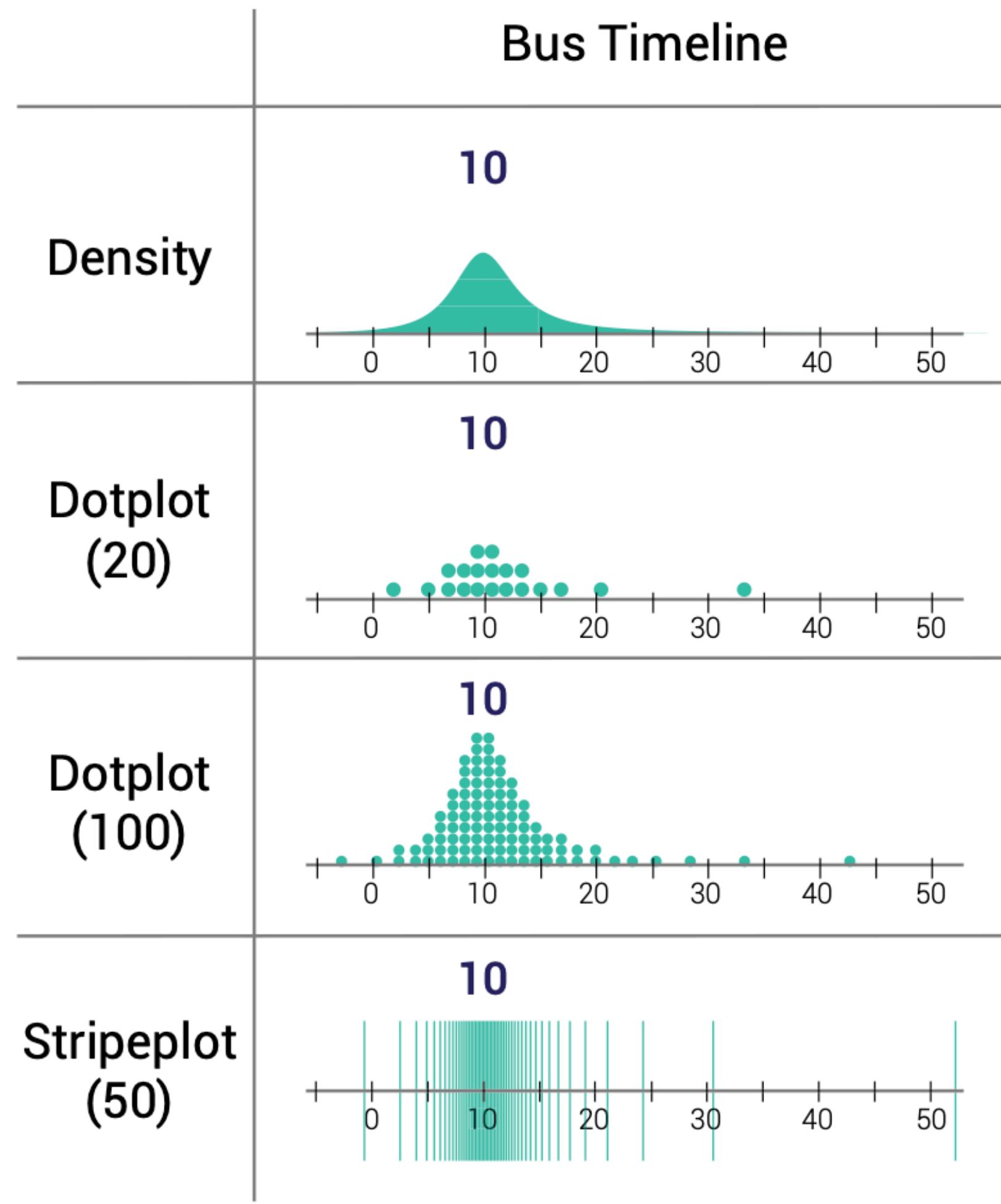
Value-Suppressing Uncertainty Map

For uncertainty, use **visual variables** instead of visualizing point estimates

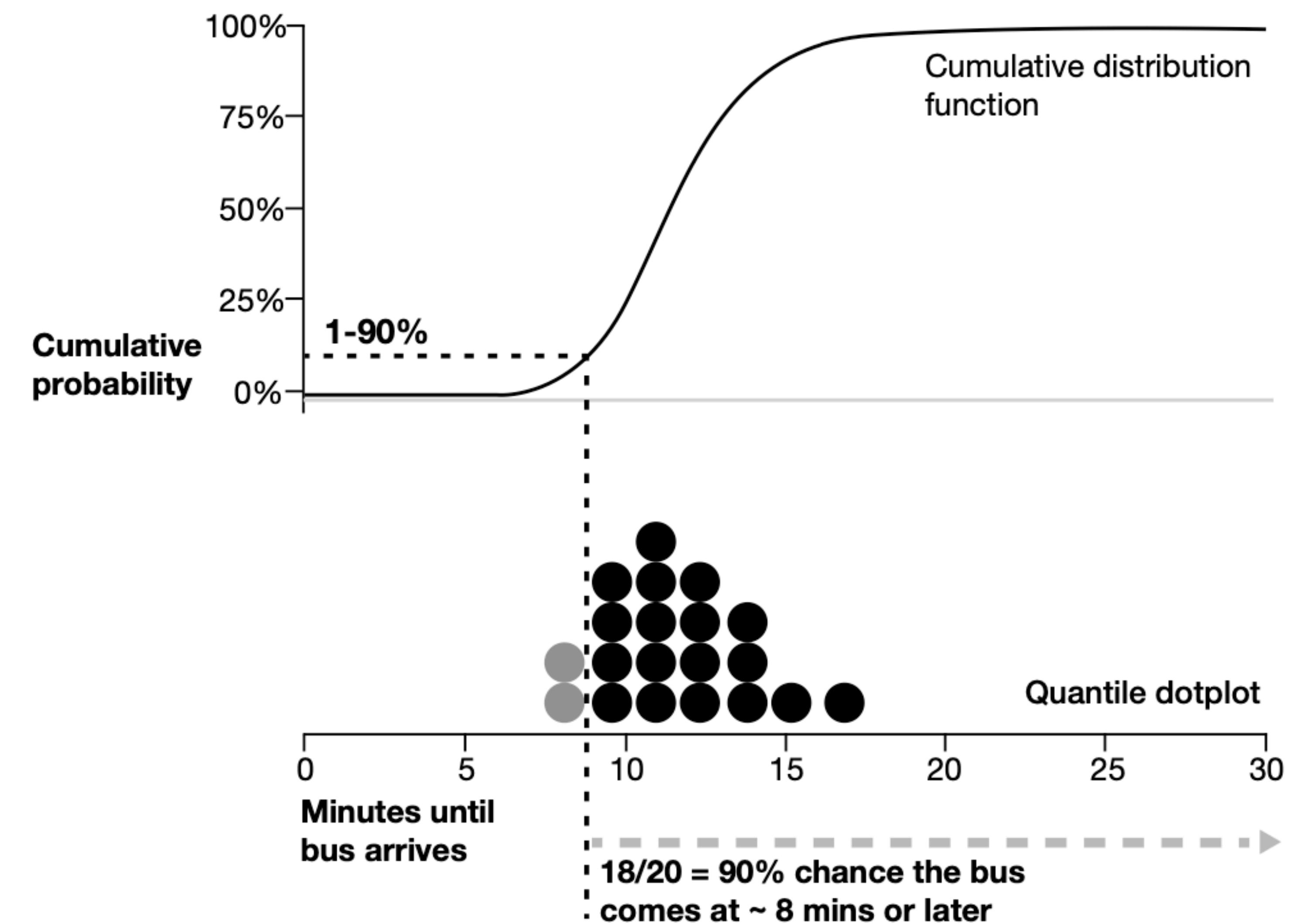
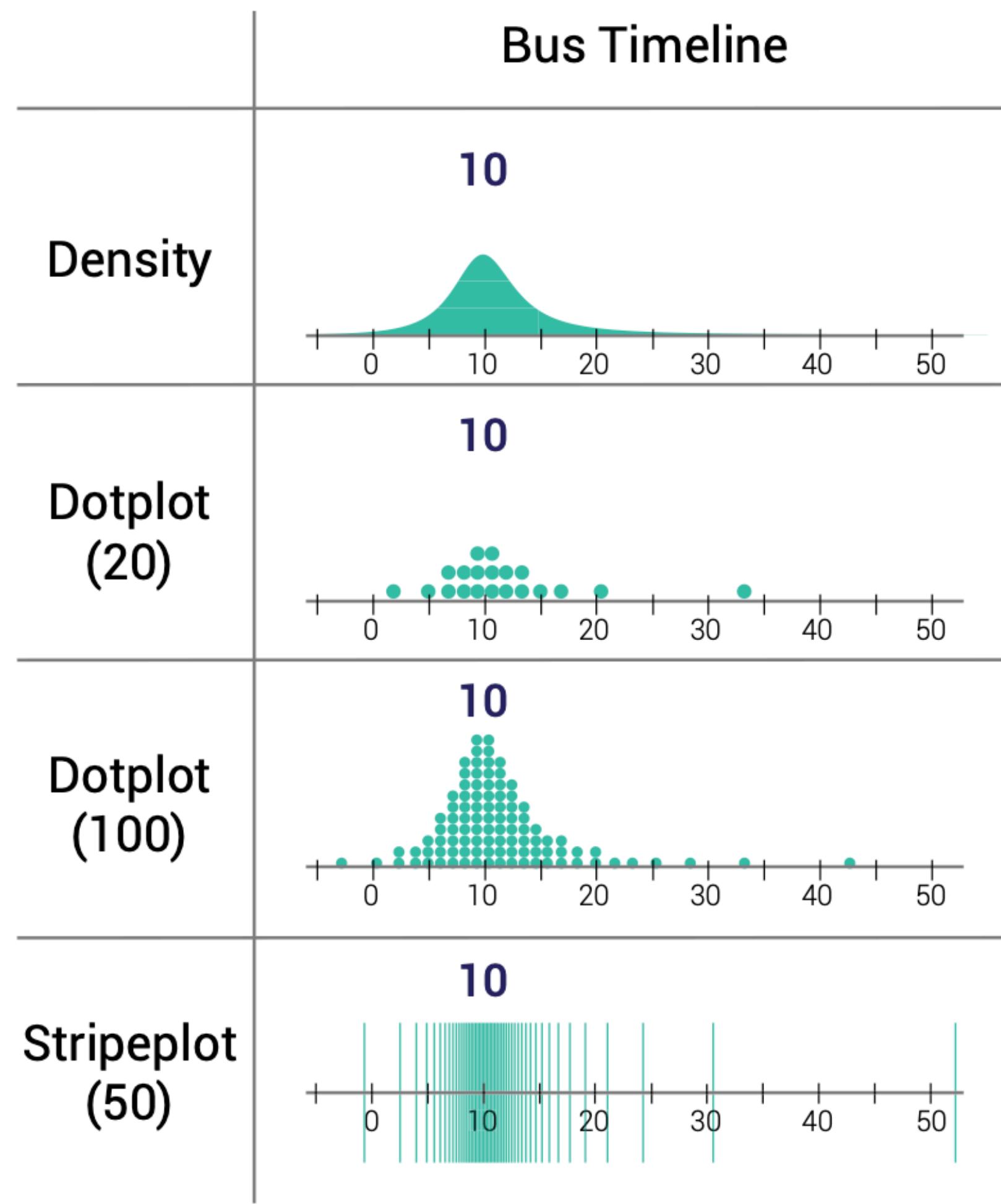


[Wood et al., 2012]
[Boukhelifa et al., 2012]

"Set of draws" technique

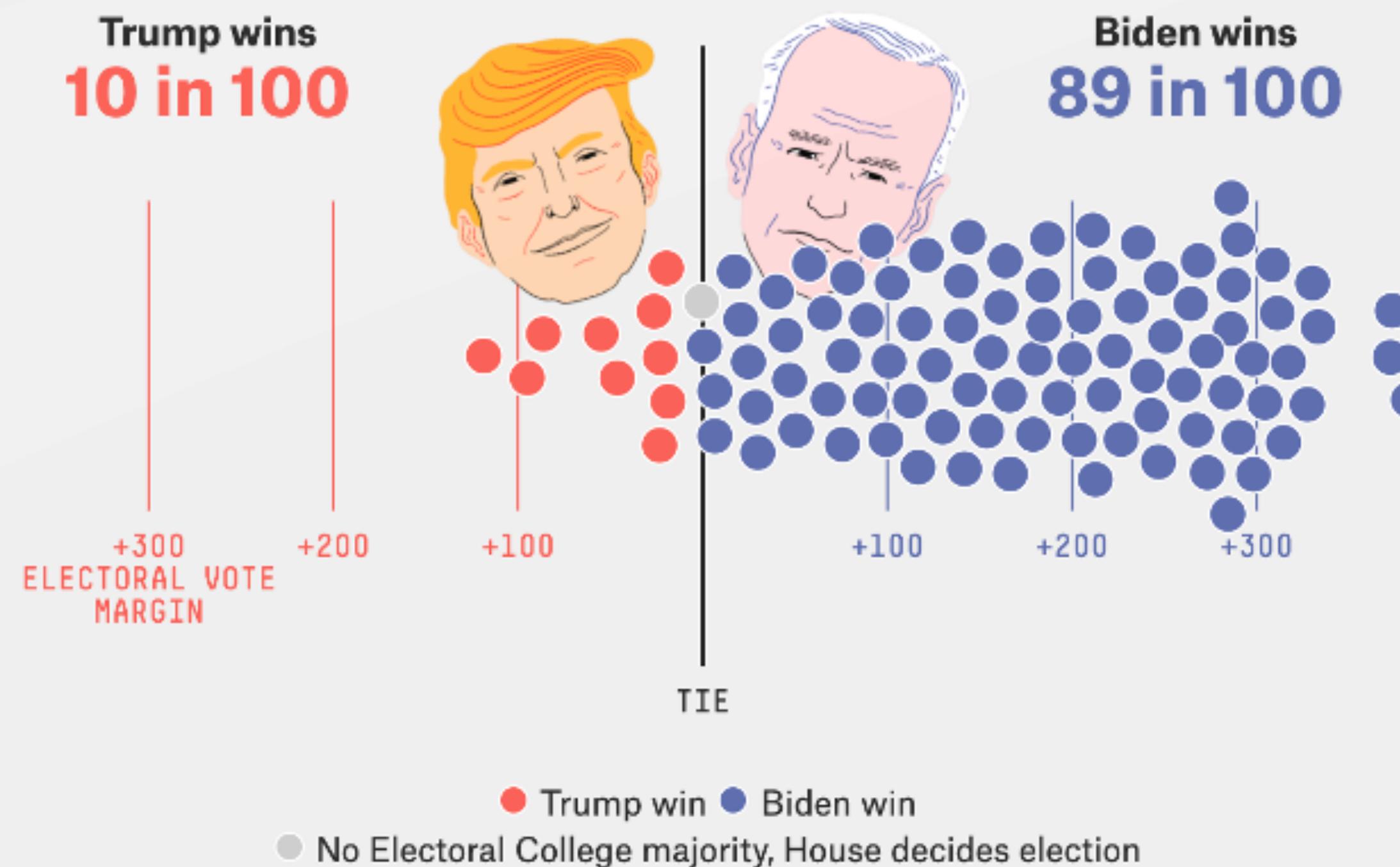


"Set of draws" technique



Biden is *favored* to win the election

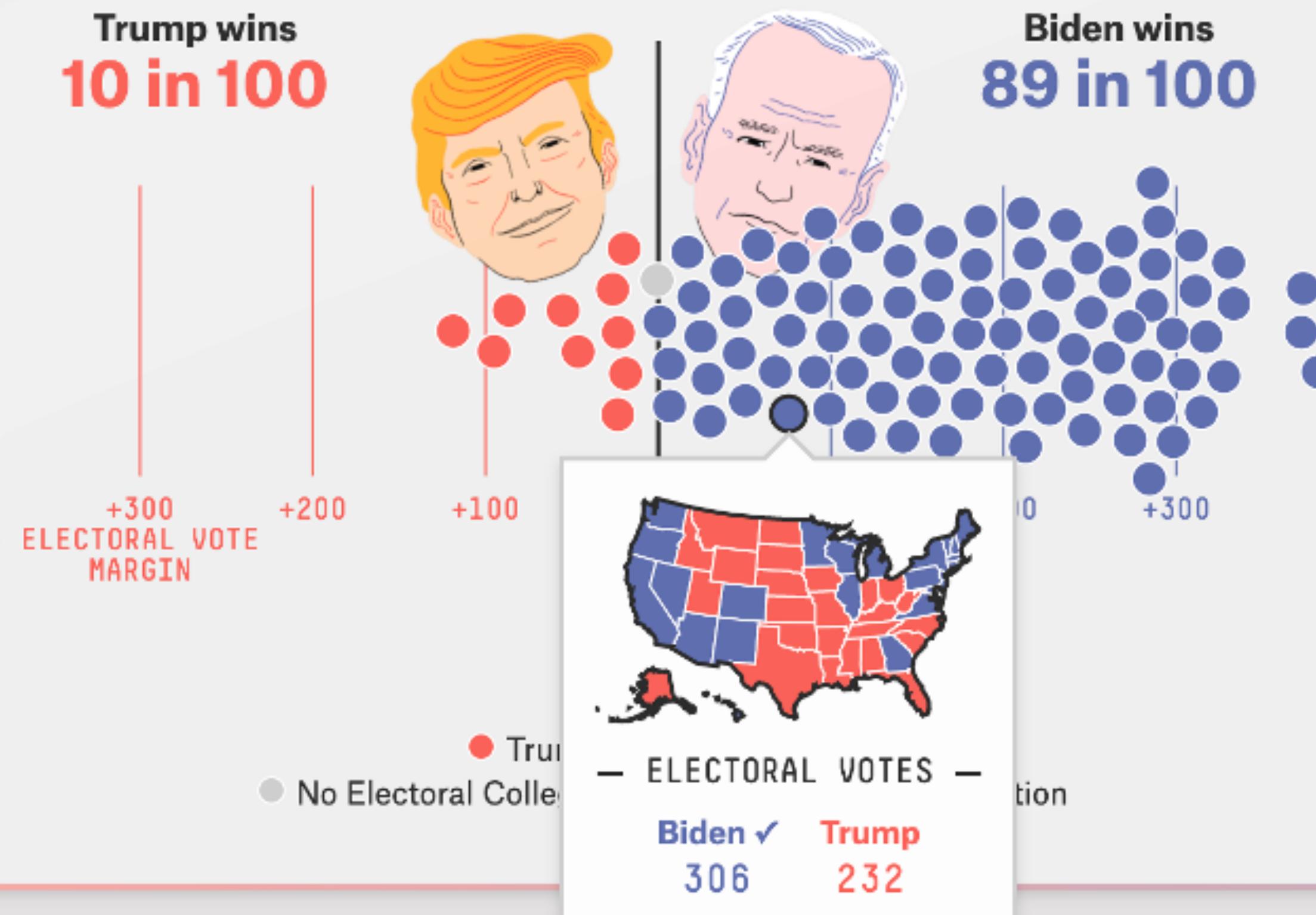
We simulate the election 40,000 times to see who wins most often. The sample of 100 outcomes below gives you a good idea of the range of scenarios our model thinks is possible.

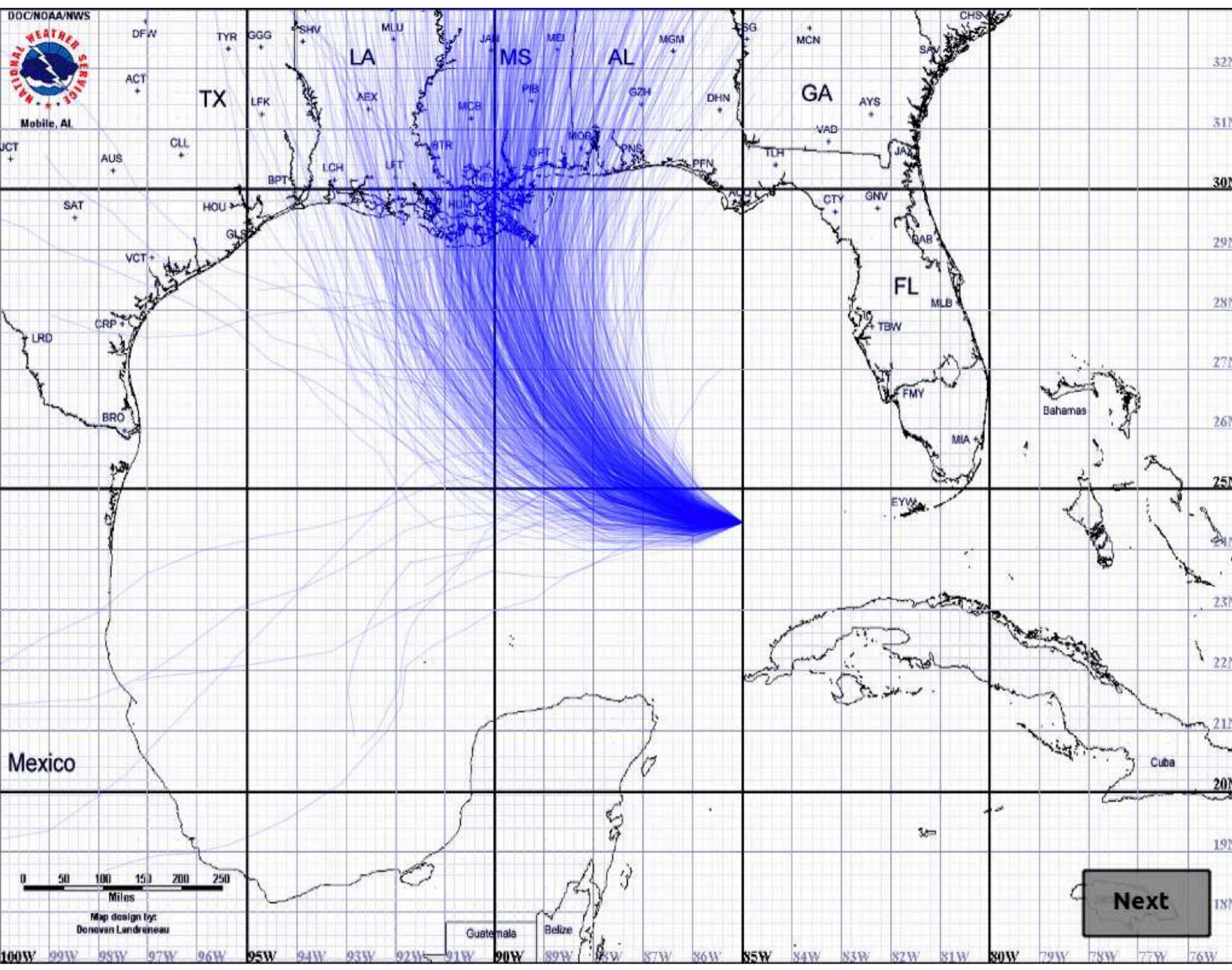


Don't count the underdog out! Upset wins are surprising but not impossible.

Biden is *favored* to win the election

We simulate the election 40,000 times to see who wins most often. The sample of 100 outcomes below gives you a good idea of the range of scenarios our model thinks is possible.





 **TheUpshot**

STATISTICAL NOISE

How Not to Be Misled by the Jobs Report

If the economy actually added 150,000 jobs last month, it would be possible to see any of these headlines:

The jobs number is just an estimate, and it comes with uncertainty.



Under 55,000 jobs
4% chance

55,000 to 110,000
19% chance

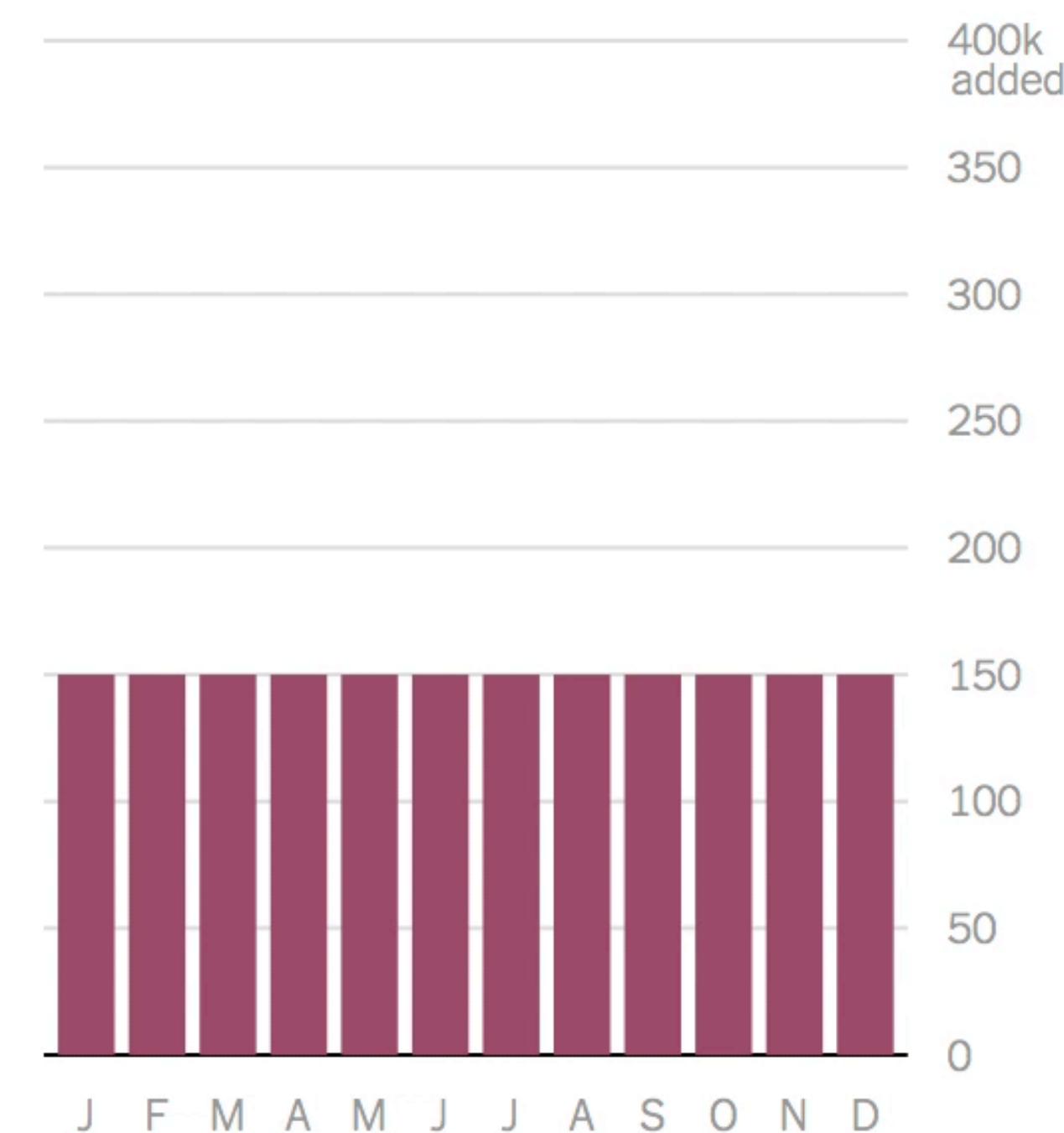
110,000 to 140,000
19% chance

160,000 to 190,000
19% chance

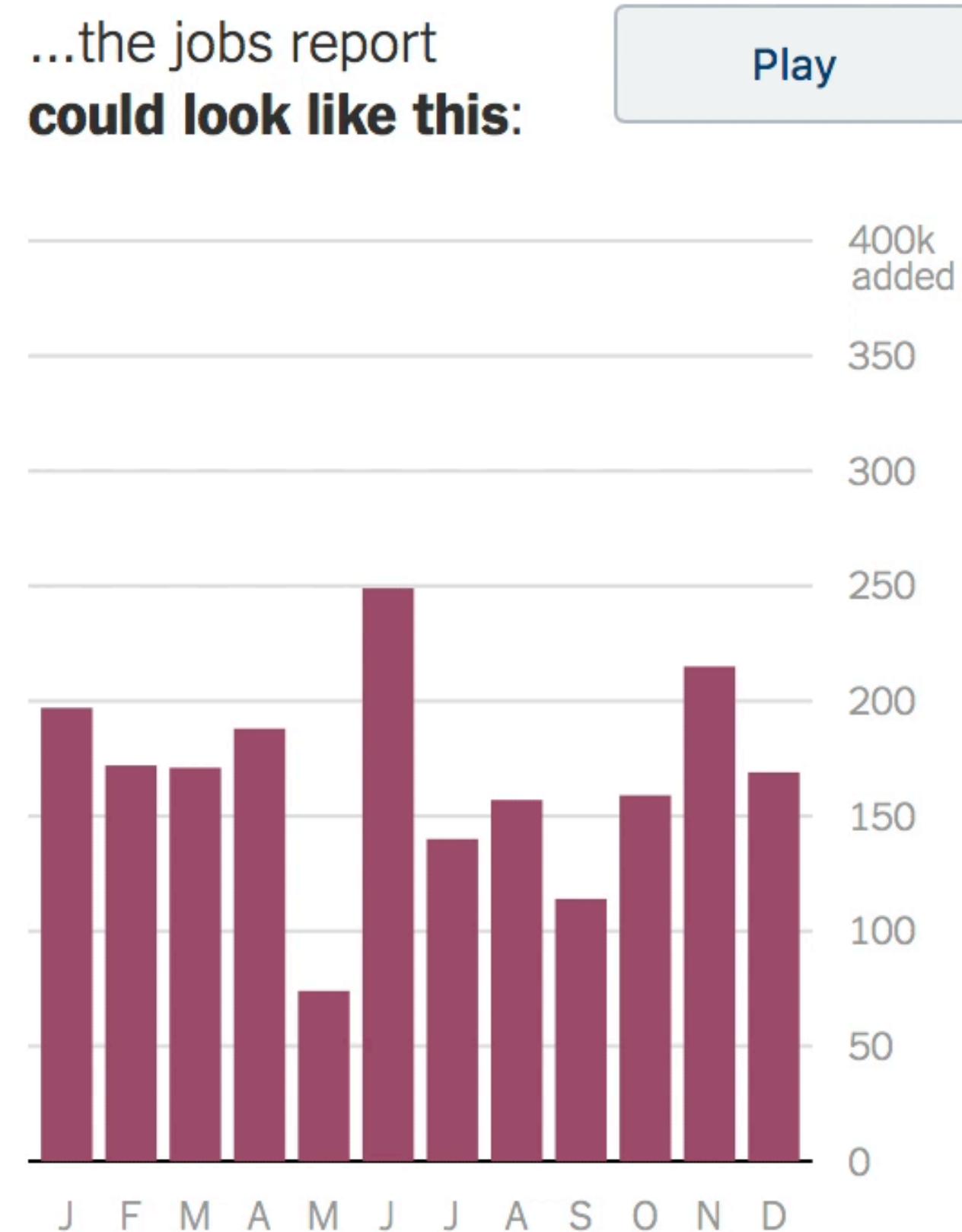
190,000 to 245,000
19% chance

245,000+
4% chance

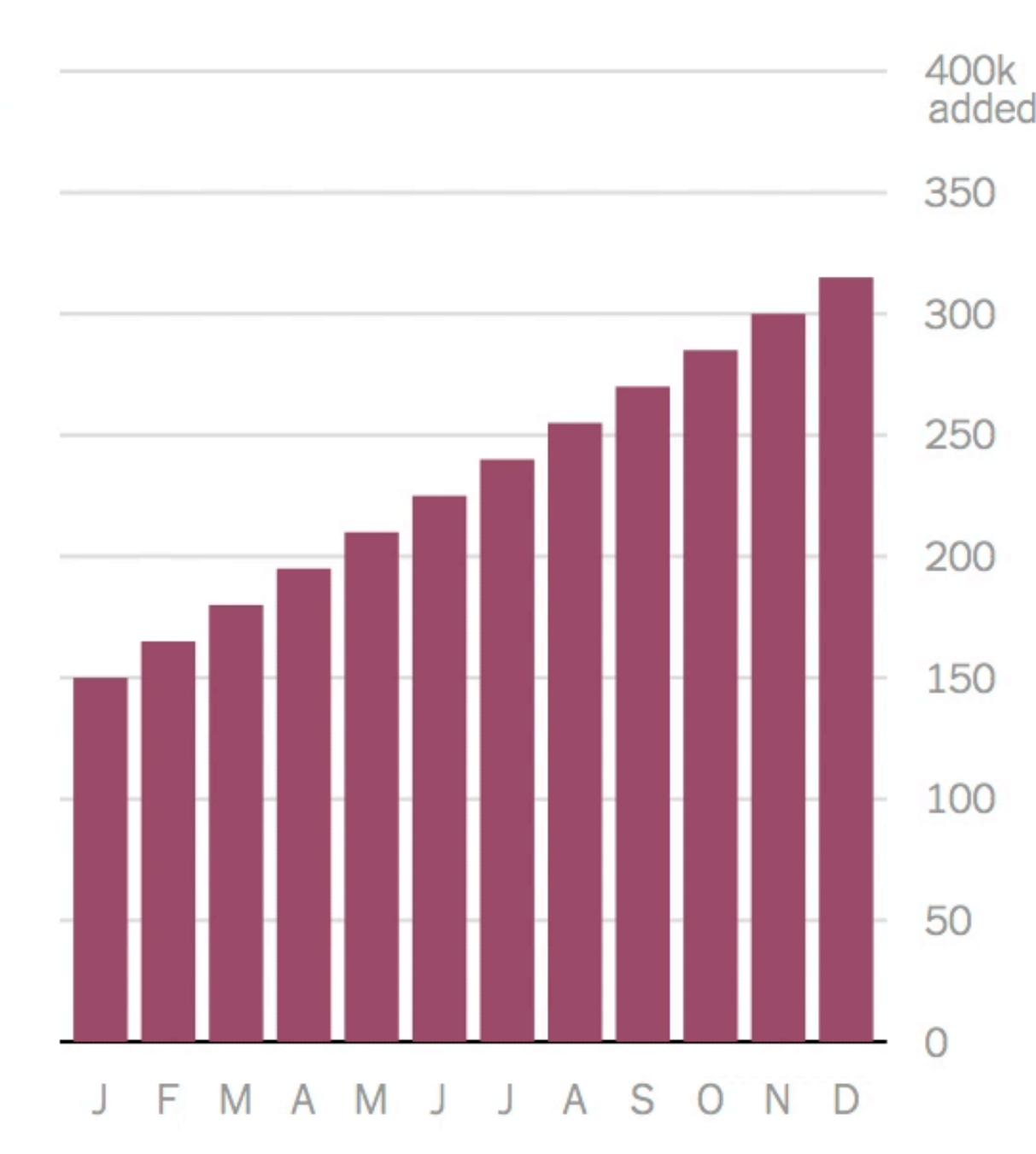
If job growth **were actually steady**
over the last 12 months...



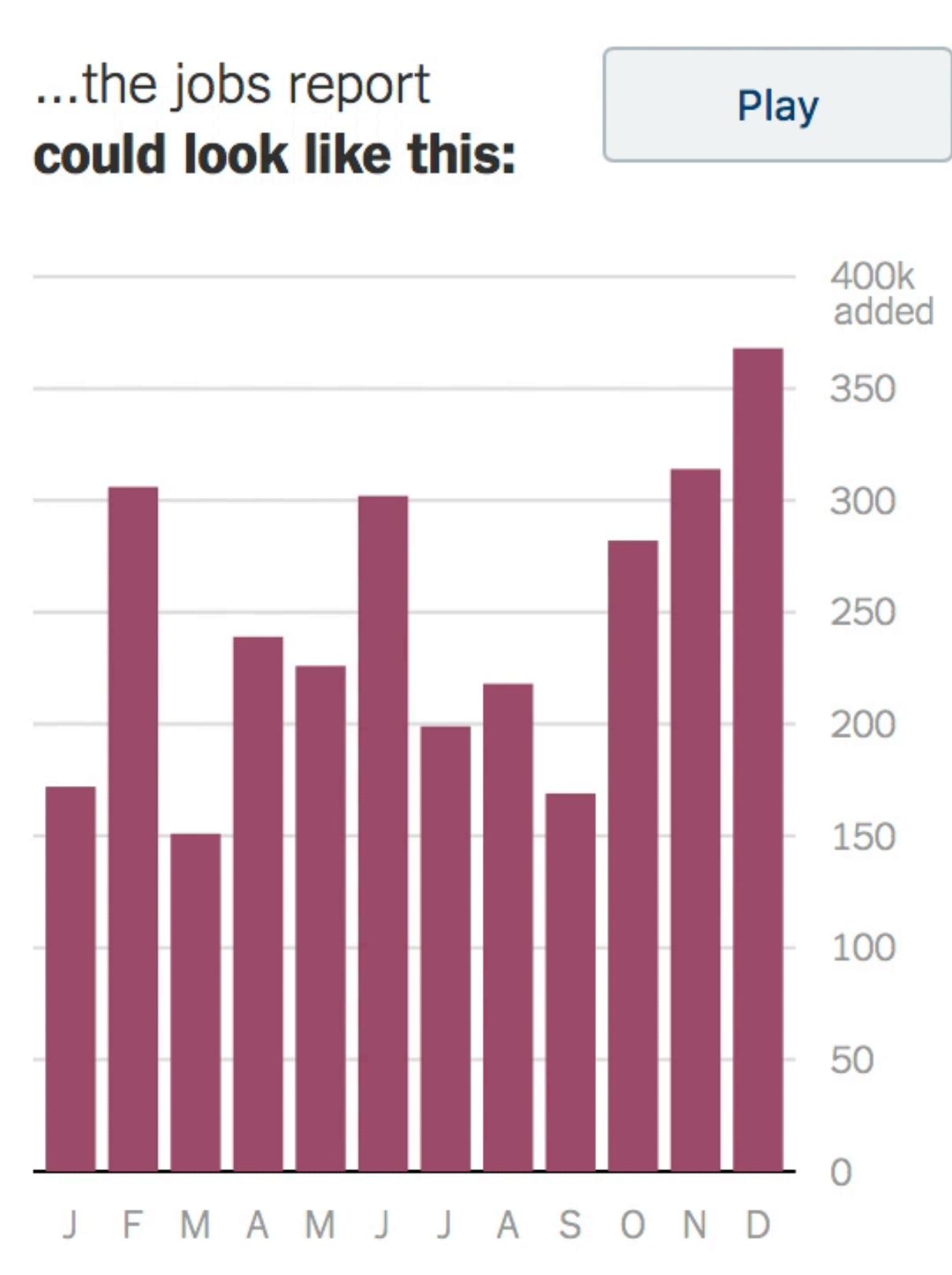
...the jobs report
could look like this:



If job growth **had been accelerating**...



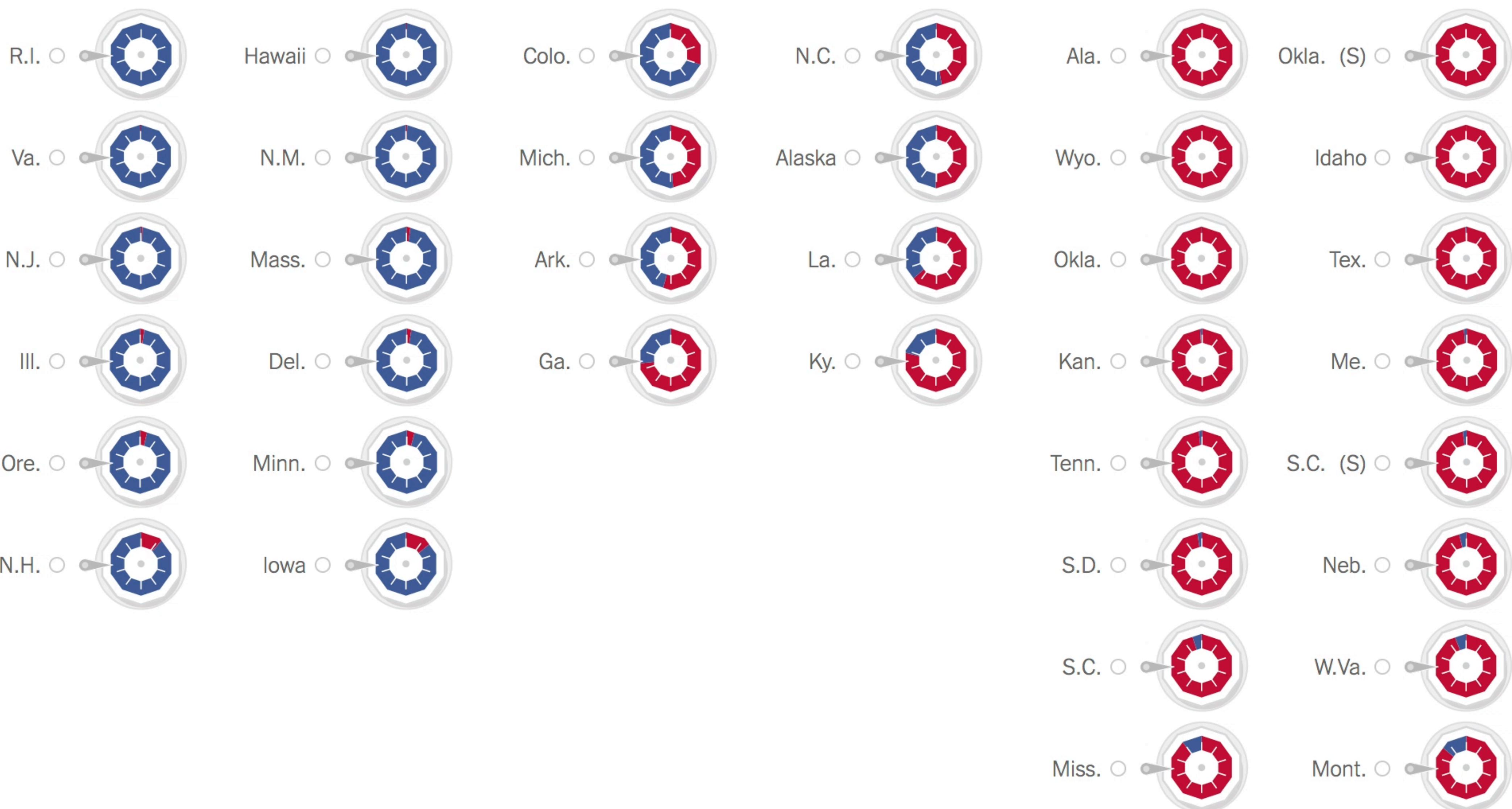
...the jobs report
could look like this:



Likely Democratic

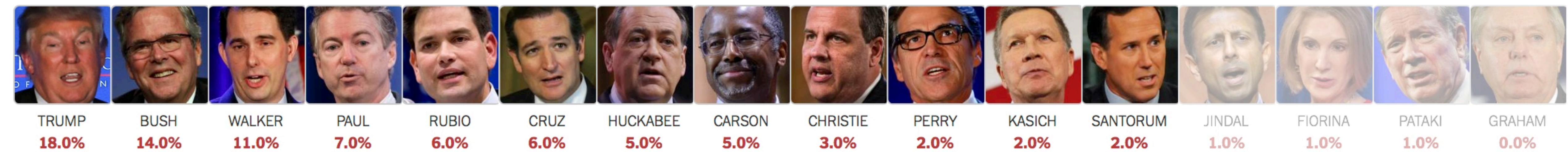
Competitive

Likely Republican



Here's a simulation of who could be in and who could be out if the candidates' averages were rounded to the nearest whole number.

If the averages are correct, but rounding is to the nearest whole number:



Rounding to fewer decimal places could be welcome news for candidates on the cusp like Mr. Santorum (who has already called the debate rules "a miscarriage"), Mr. Kasich or Mr. Jindal.

Uncertainty

What does it mean?

Building models is necessary to quantify uncertainty.

It is important to communicate the variability in model outcomes.

How should I visualize it?

Dynamic or ensemble displays can help communicate complex models.

Why Authors Sometimes Don't Visualize Uncertainty

A visualization expresses a signal

Authors simplify, crystallize, abstract the complexity of data.

Process validates signal

Authors decide whether process has "low enough" uncertainty.

Uncertainty obfuscates signal

Could distract, or require too much work from the reader.

Uncertainty

What does it mean?

Lots of things!

How should I visualize it?

It depends!