

# Uncertainty

---

**DSC 106: Data Visualization**

Jared Wilber

UC San Diego

# Announcements

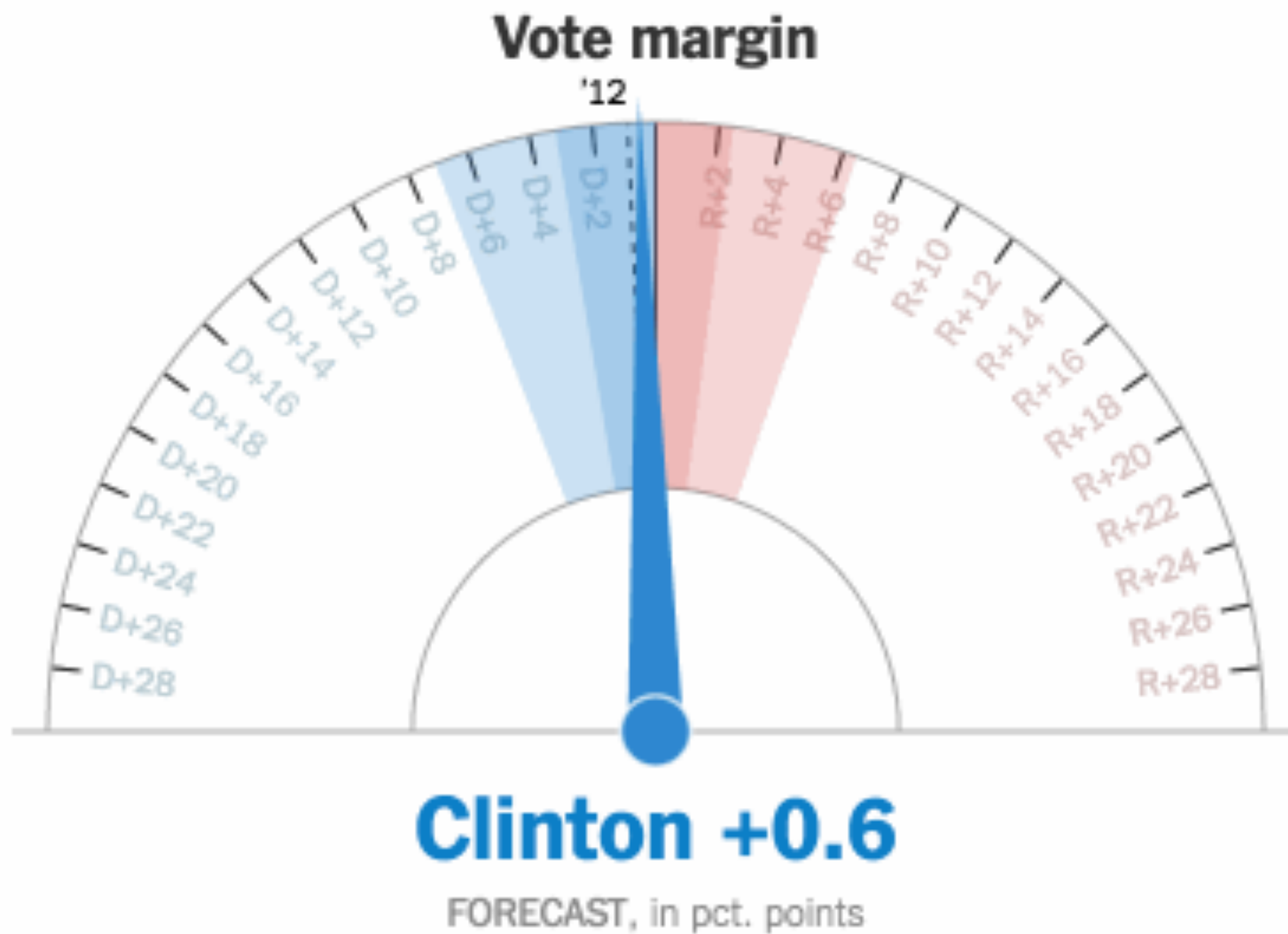
Final Project prototype due Monday.

Final Project video due the Monday thereafter.

Lecture next Thurs. is Final Project peer feedback.

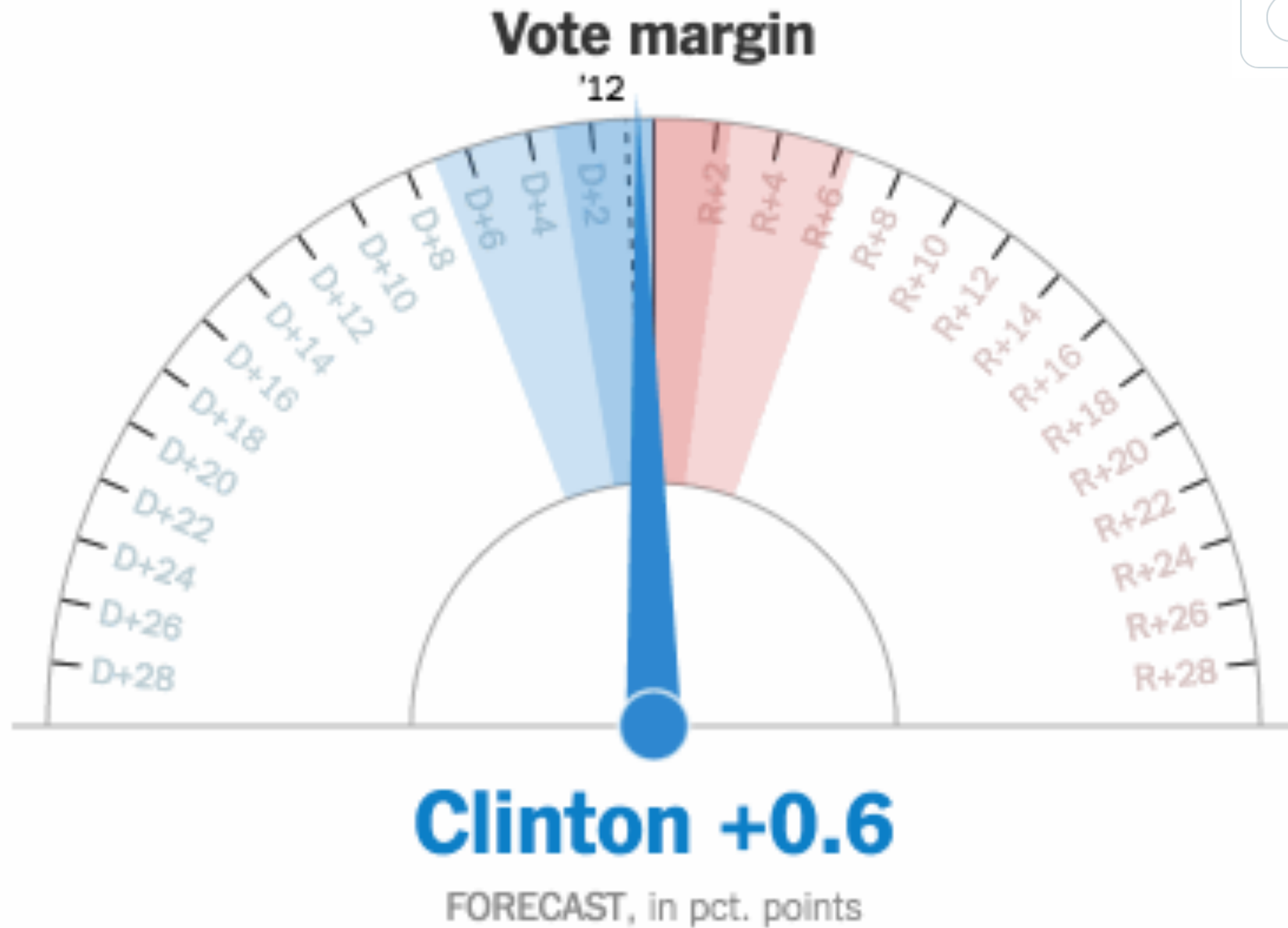
## **FAQs:**

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?



**(gobsmacked)**  
@nytimes

I'm taking damn NY

10:08 PM ·

2

---

**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

2

**Paul D. Shinkman** ✓  
@PDSrinkman · Follow

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

Heads-up: Forecasts may become more trustworthy on

Our best guess right now is the estimates of three key indicators

66% Trump FORECAST

10:03 PM · Nov 8, 2016

Reply

**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

```

range_05 = range_05 * damping + range_05_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 - range_75);
needle.attr('transform', 'scale('+scale+') rotate('+angle(cur_val) + jitter)');
if (opts.colors) {

```

**@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

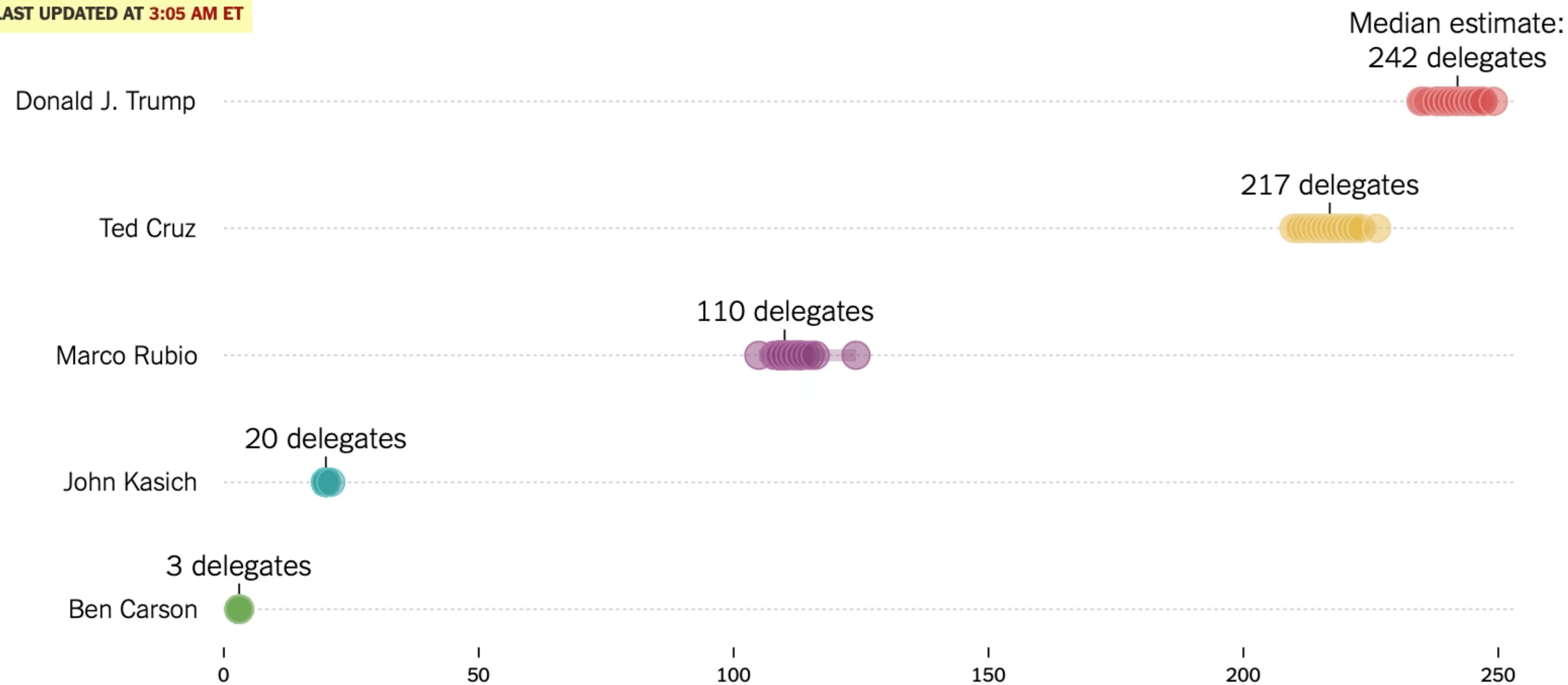
Reply Share

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?

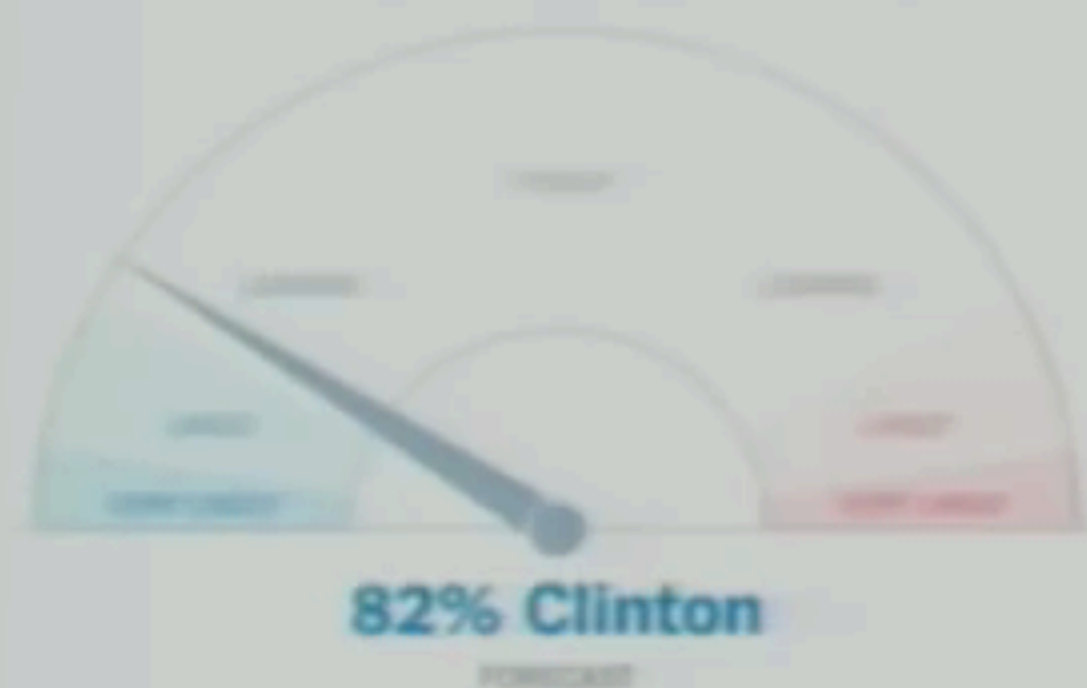
What are the strengths and weaknesses of this visualization?

How does it compare to the needle?

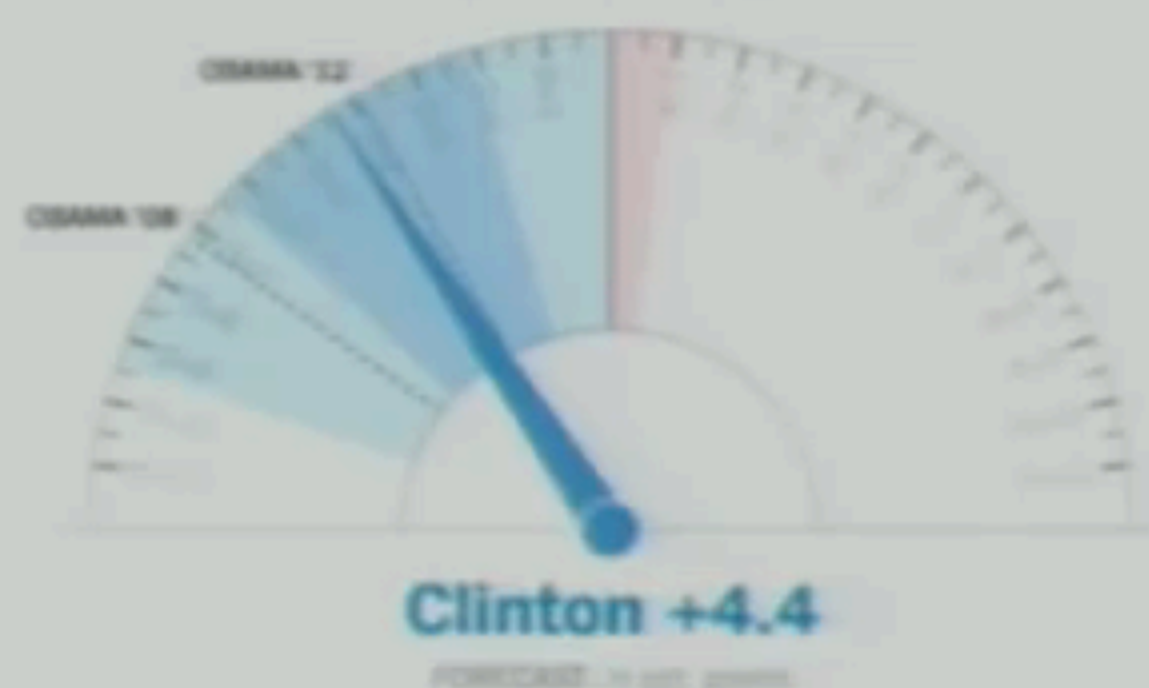
# Live Presidential Forecast

Updated 7:55:14 PM ET

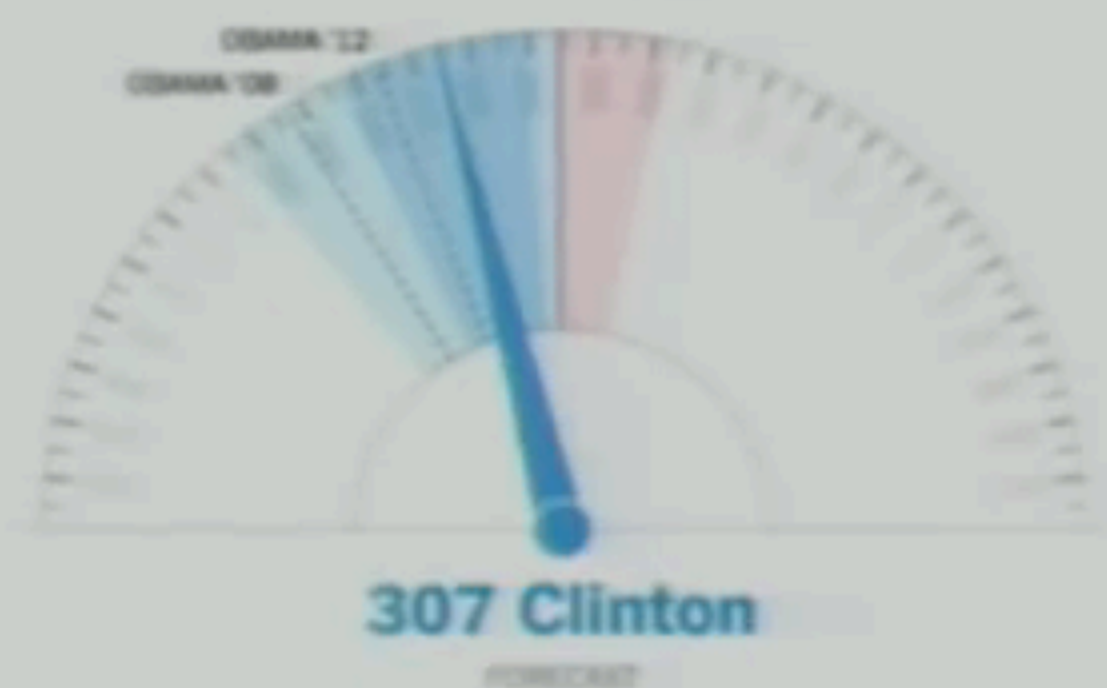
Chance of Winning Presidency



Popular vote margin



Electoral votes



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](#)

State	Est. pct. of voters	Reported margin	NYT projection	NYT win prob.
Michigan	0%		Clinton +6.5	75% Dem.
New Mexico	0%		Clinton +6.4	75% Dem.
Wisconsin	0%		Clinton +5.4	75% Dem.



OPEN  
VIS 2017  
CONF

# Uncertainty

What does it mean?

How should I visualize it?

# Uncertainty

What does it mean?

How should I visualize it?



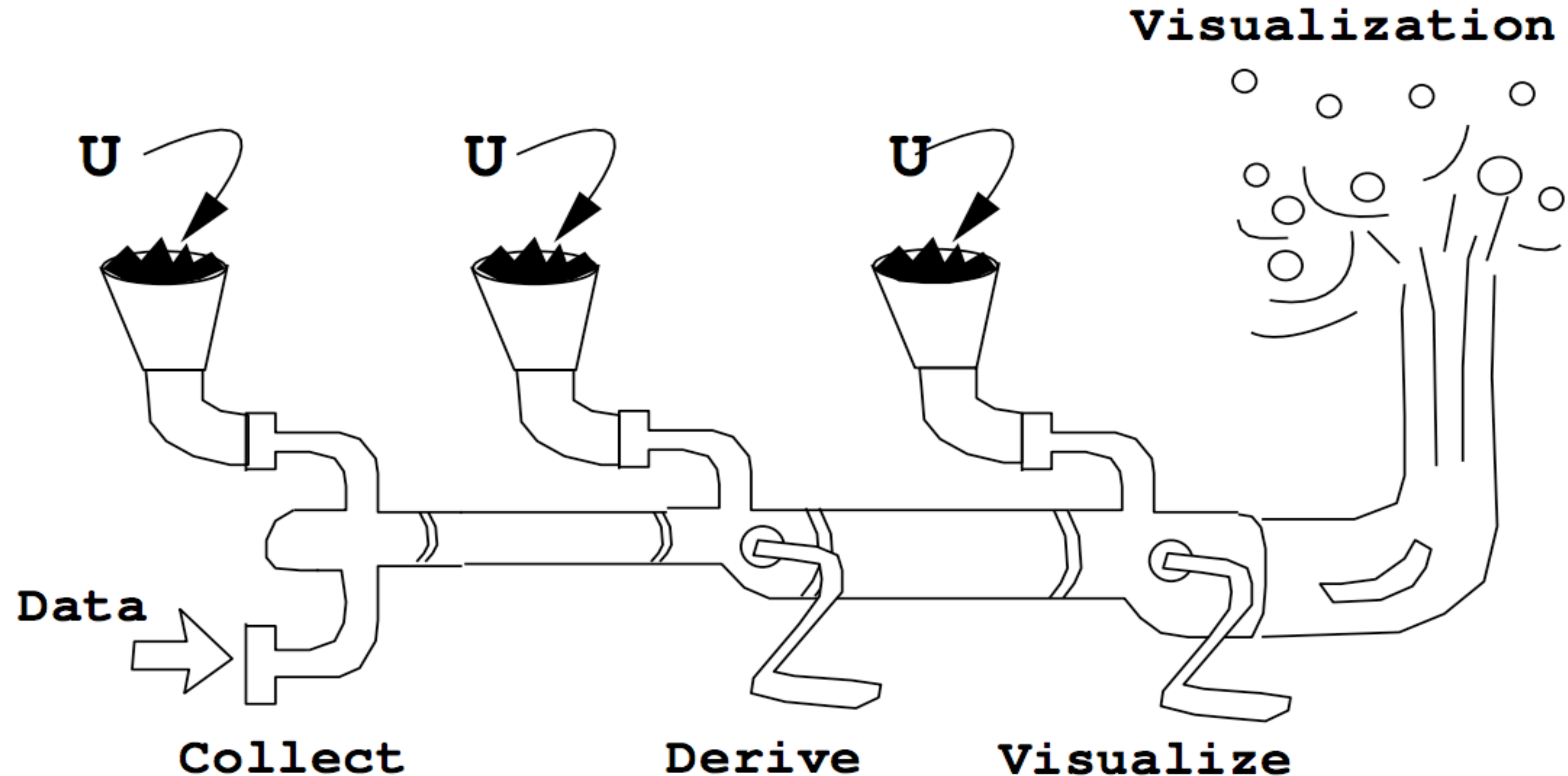
# Uncertainty

What does it mean?

How should I visualize it?

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

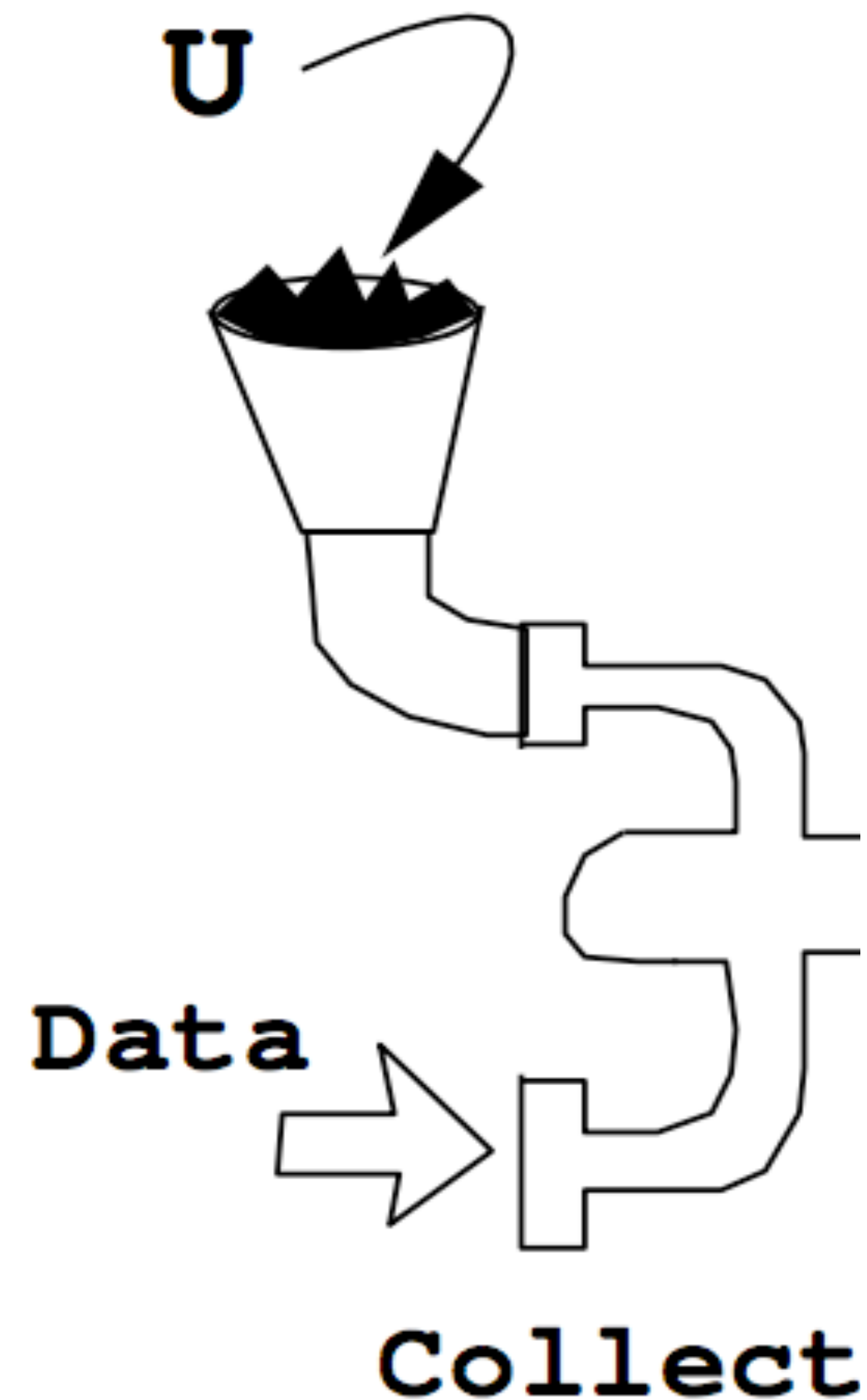
# Sources and Types of Uncertainty



# Sources and Types of Uncertainty

## Measurement Uncertainty

How and how much should we sample the data?



**Precision**

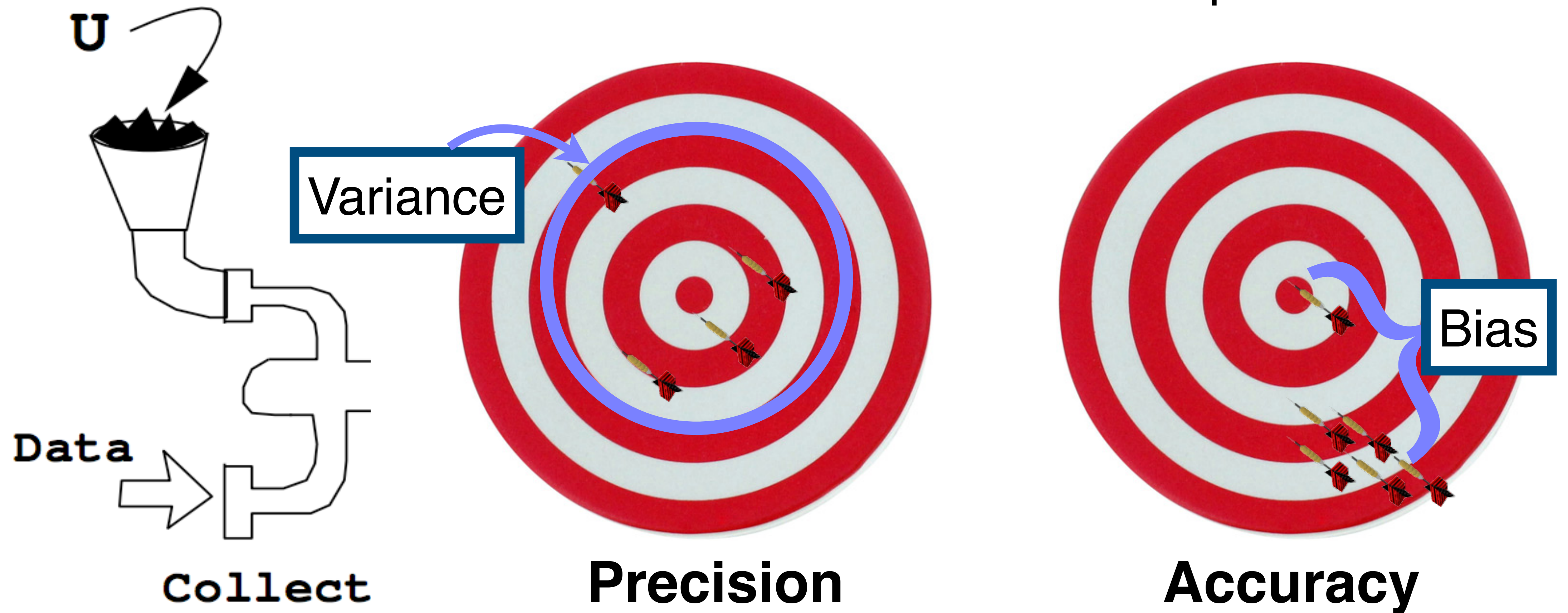


**Accuracy**

# Sources and Types of Uncertainty

## Measurement Uncertainty

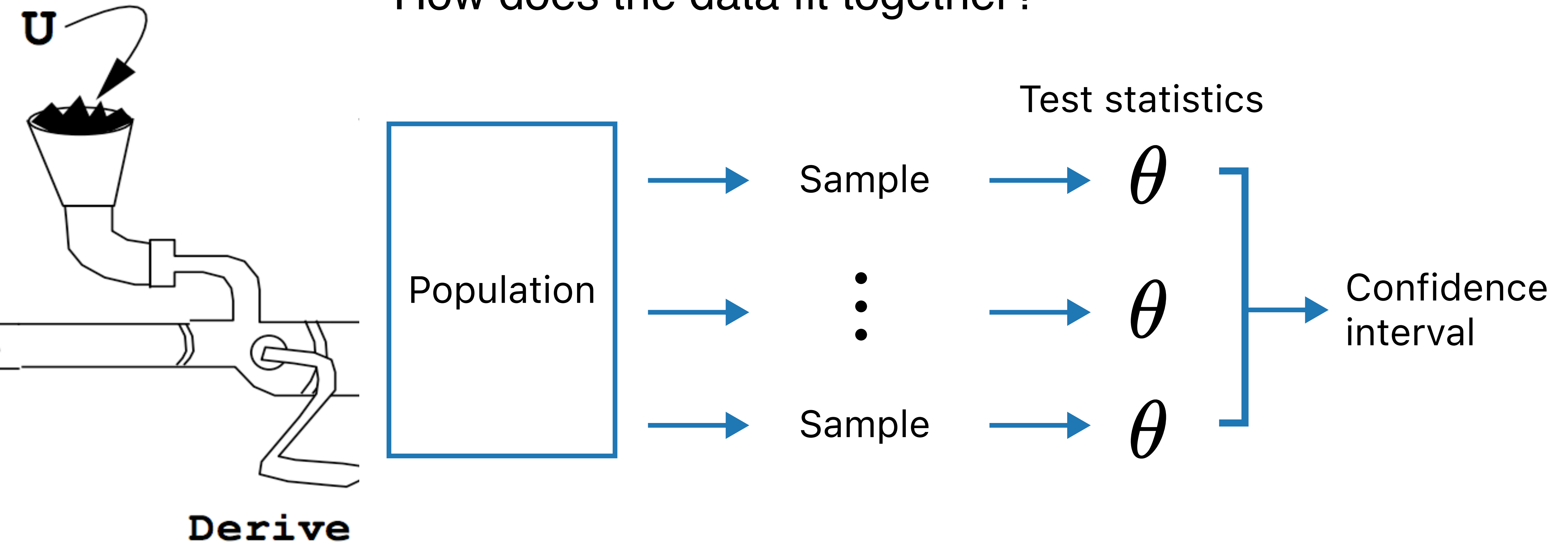
How and how much should we sample the data?



# Sources and Types of Uncertainty

## Model Uncertainty

How does the data fit together?

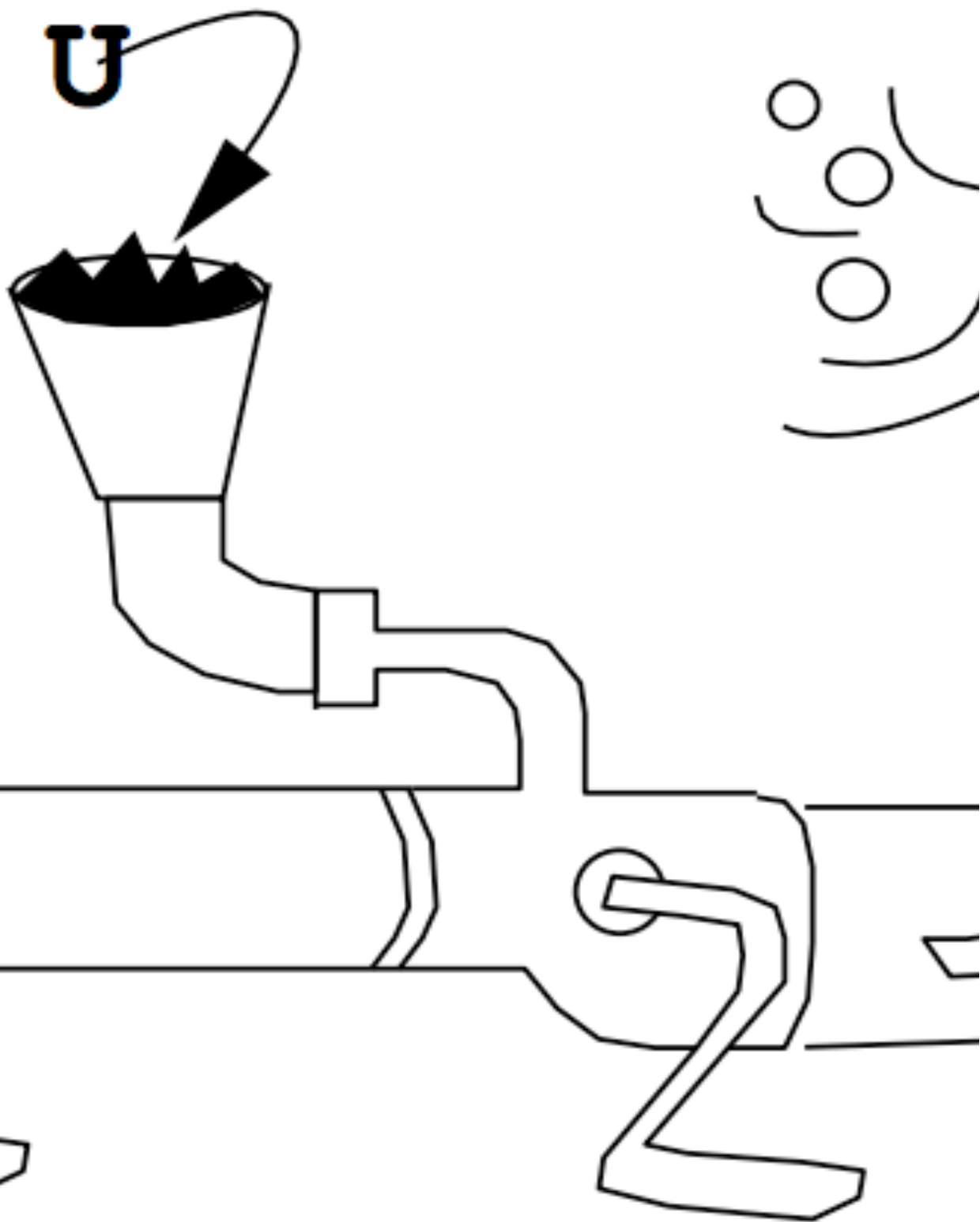


# Sources and Types of Uncertainty

Visuali



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

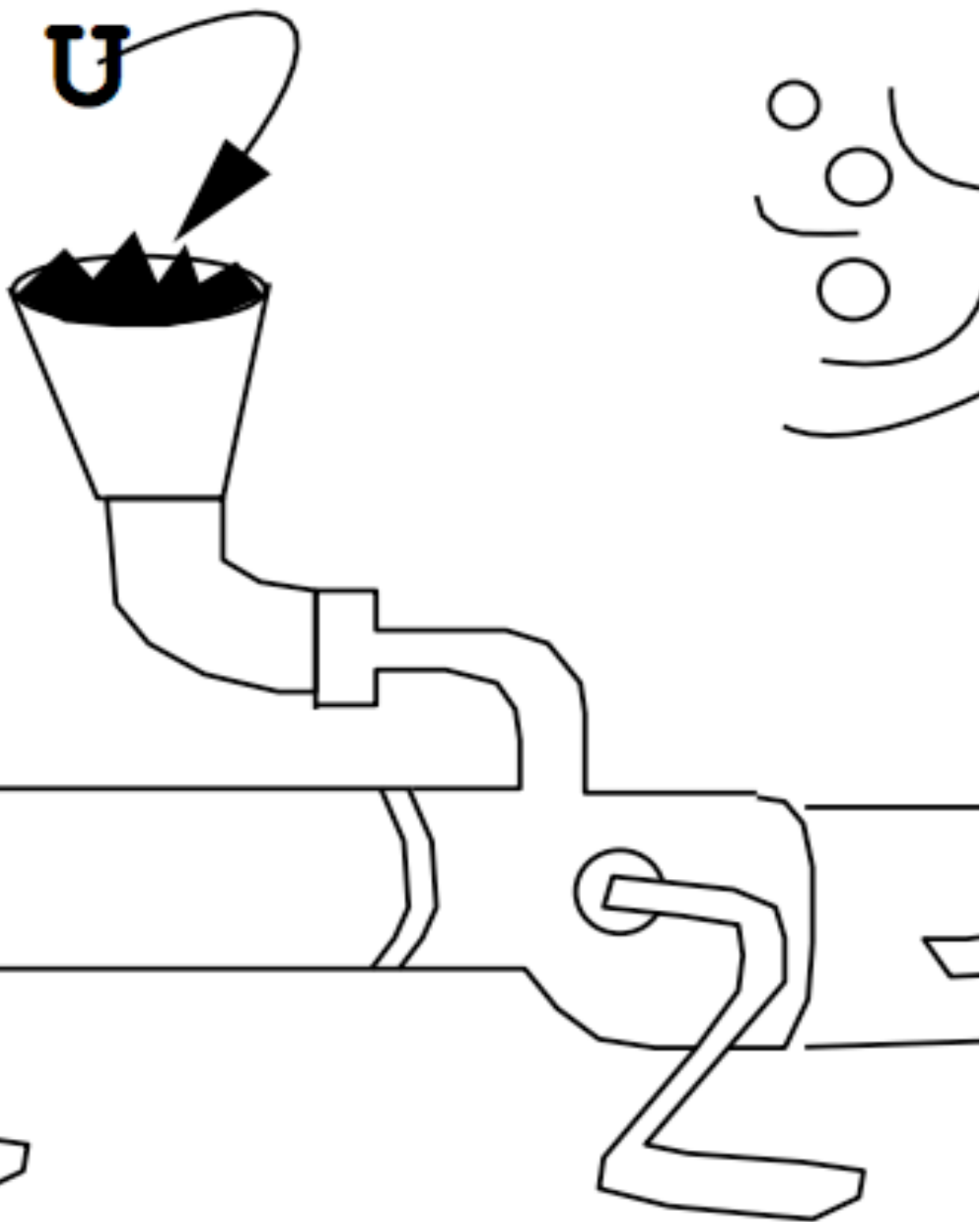


Visualize




# Sources and Types of Uncertainty

Visuali



Visualize

**Decision/Forecast Uncertainty**

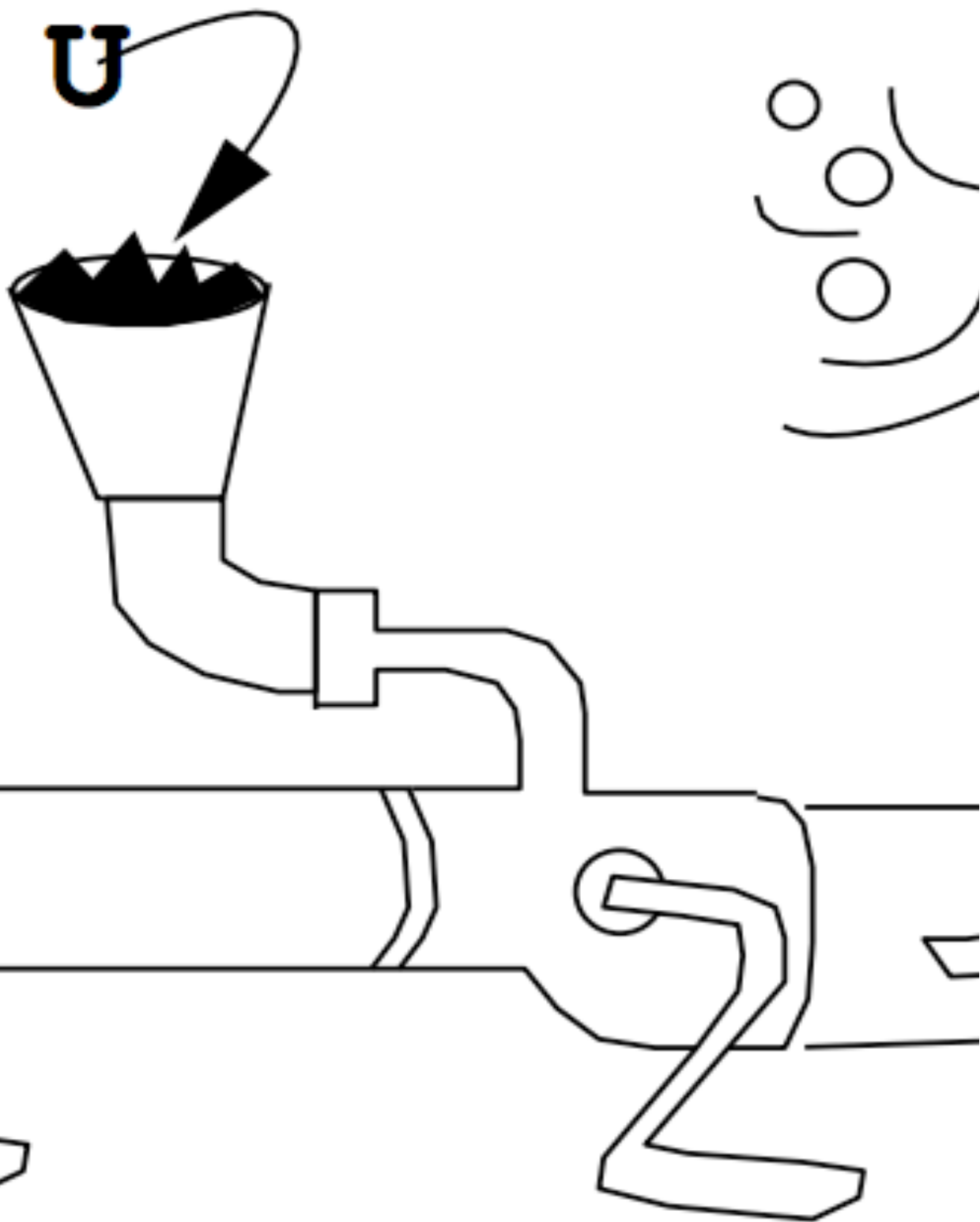
How do I assess the risk or error?




# Sources and Types of Uncertainty



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

Visuali



Visualize



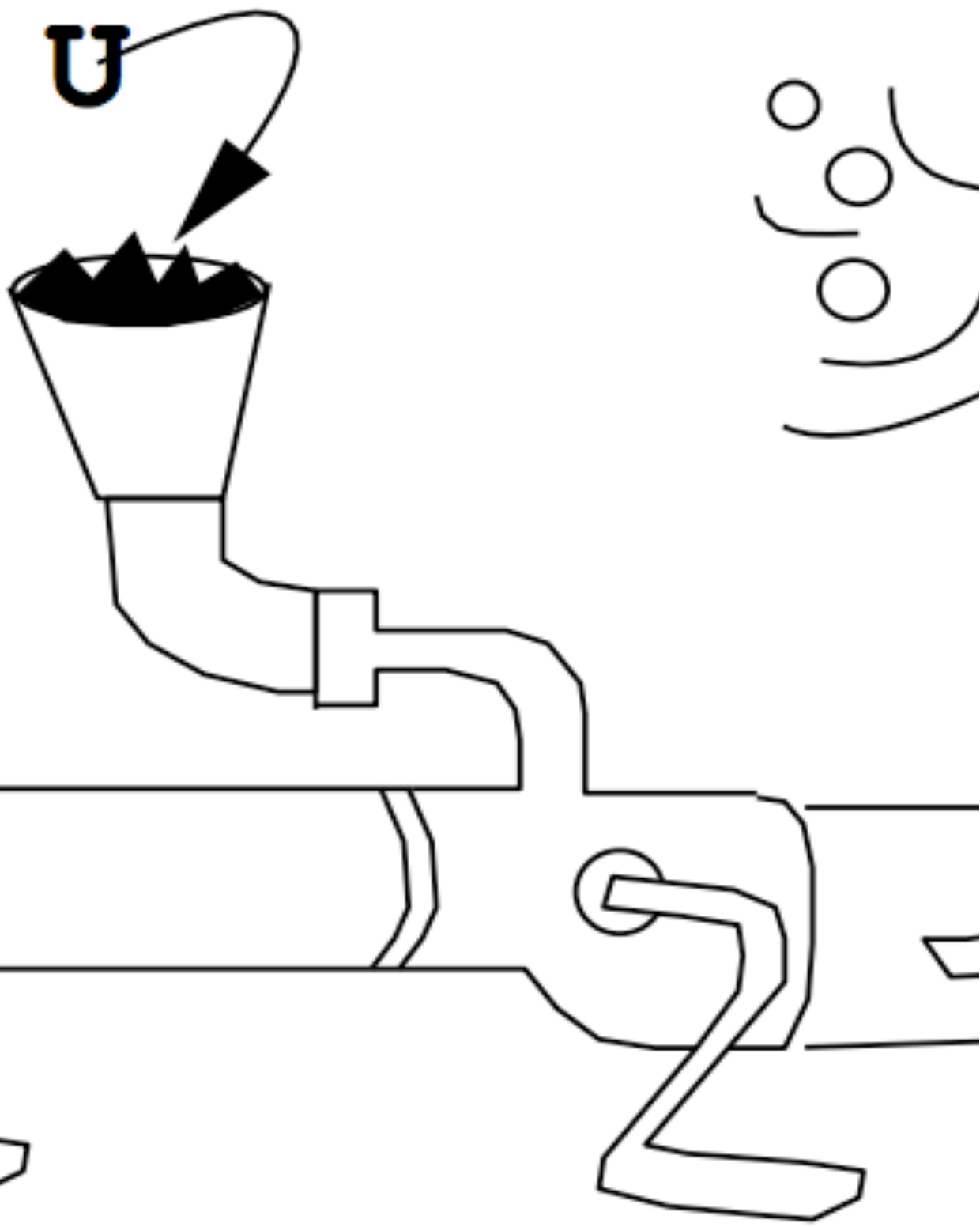
	<b>Type I False Positive</b>
	



# Sources and Types of Uncertainty



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

Visuali



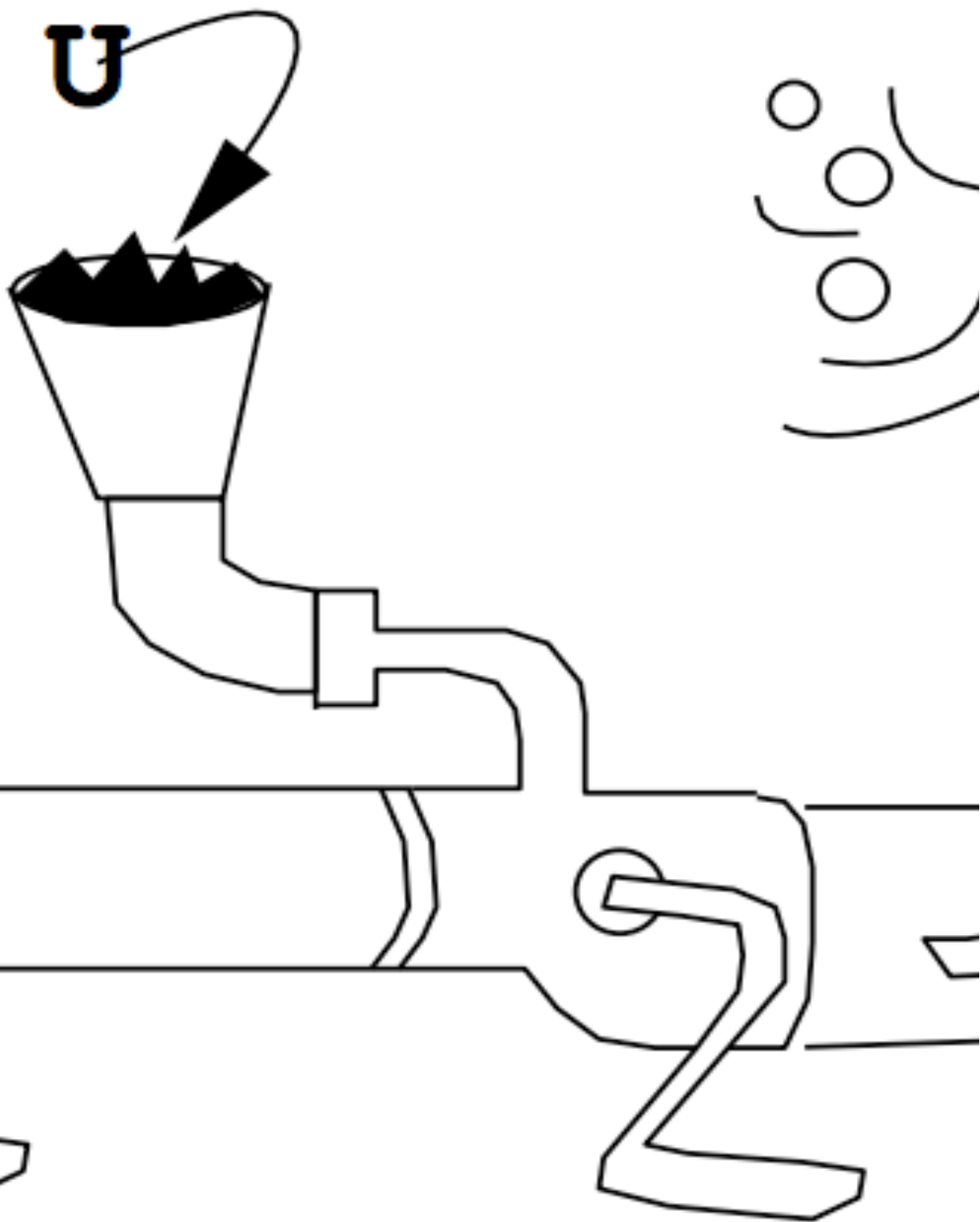
Visualize



	<b>Type I False Positive</b>
<b>Type II False Negative</b>	

# Sources and Types of Uncertainty

**Visuali**



**Visualize**

**Never confuse Type I and II errors again:**

**Just remember that the Boy Who Cried Wolf caused both Type I & II errors, in that order.**

**First everyone believed there was a wolf, when there wasn't. Next they believed there was no wolf, when there was.**

**Substitute "effect" for "wolf" and you're done.**

Kudos to @danolner for the thought. Illustration by Francis Barlow  
"De pastoris puero et agricolis" (1687). Public Domain. Via [wikimedia.org](https://commons.wikimedia.org/wiki/File:De_pastoris_puero_et_agricolis.jpg)

# Uncertainty

What does it mean?

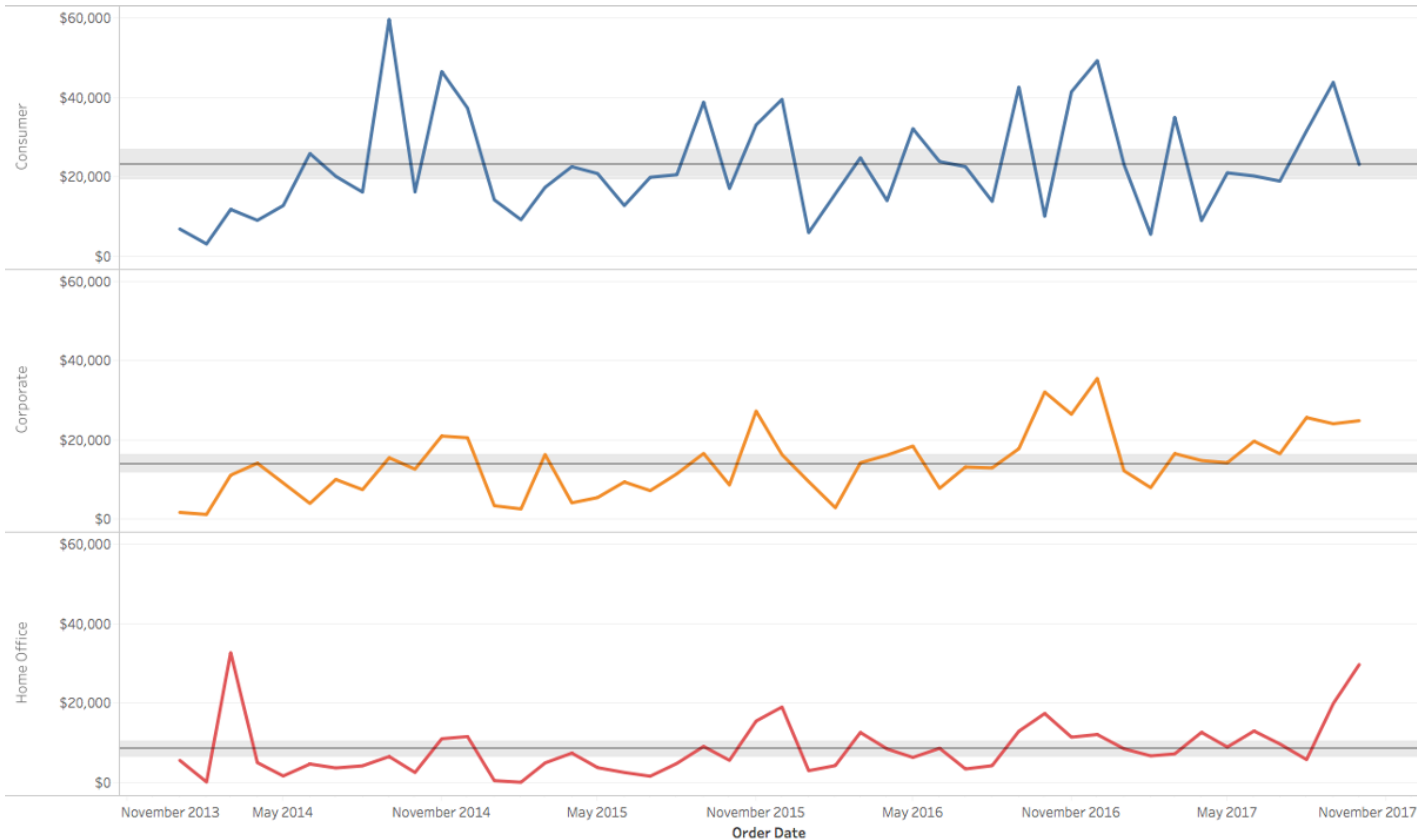
Lots of things!

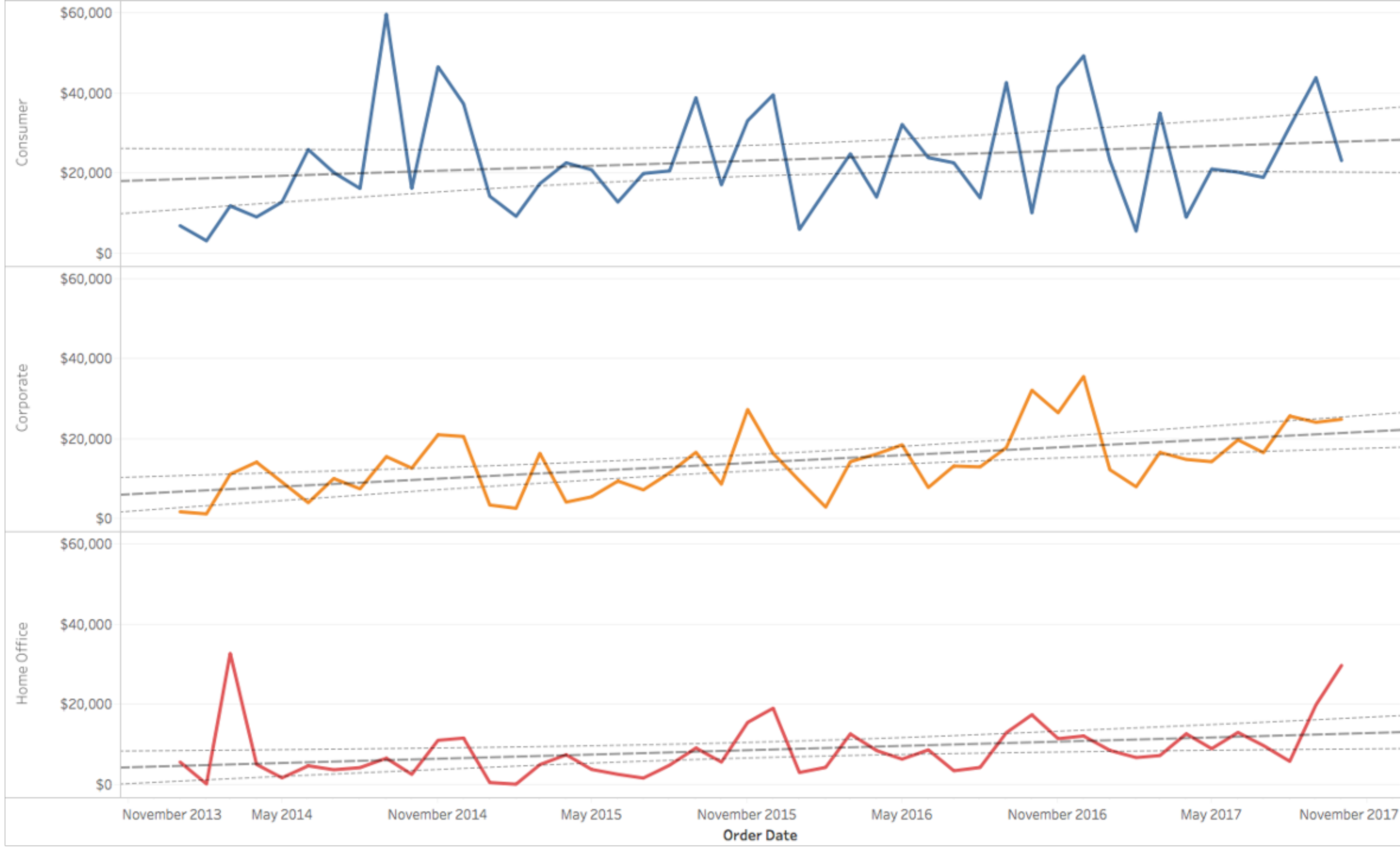
How should I visualize it?

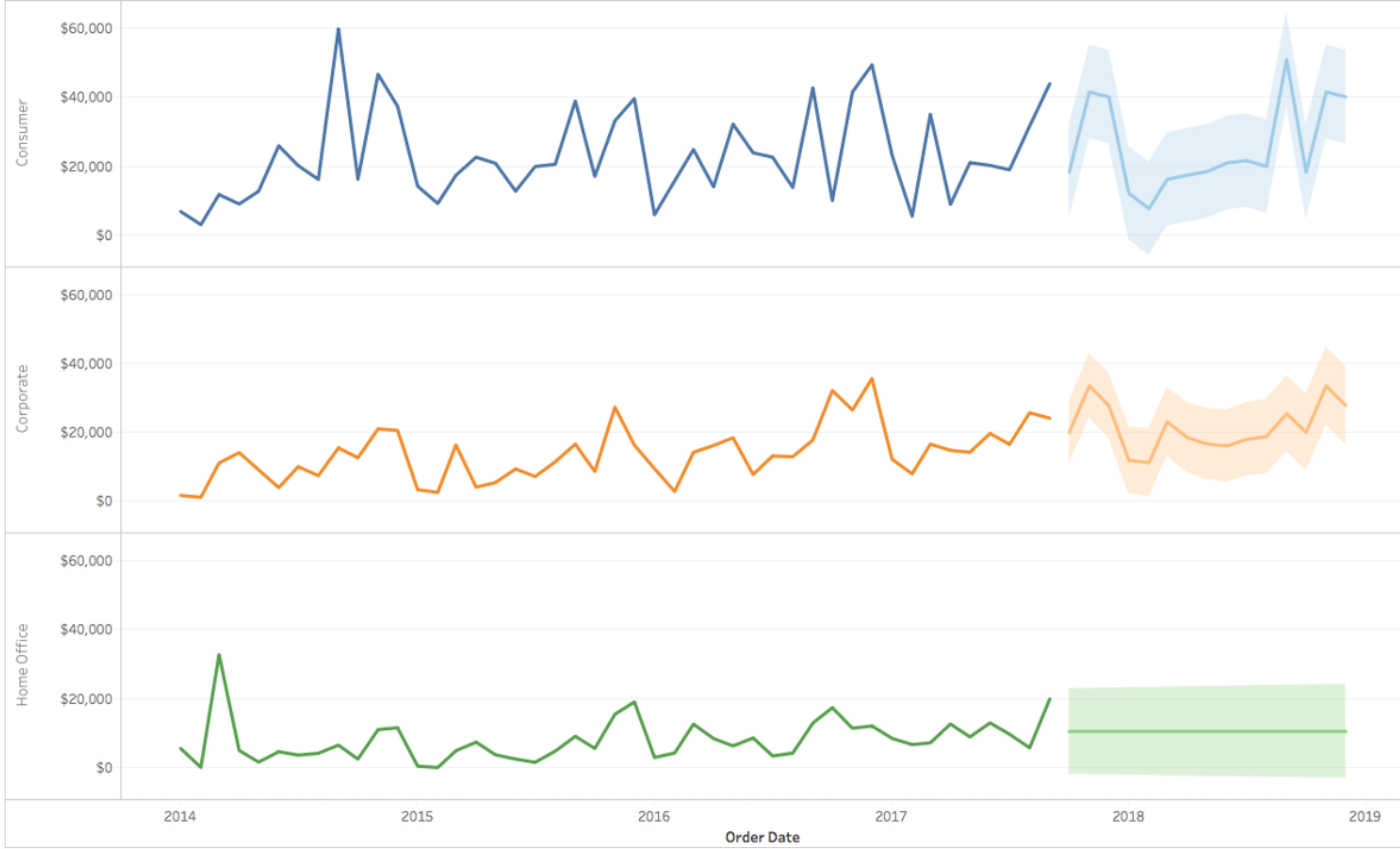
# Uncertainty

What does it mean?

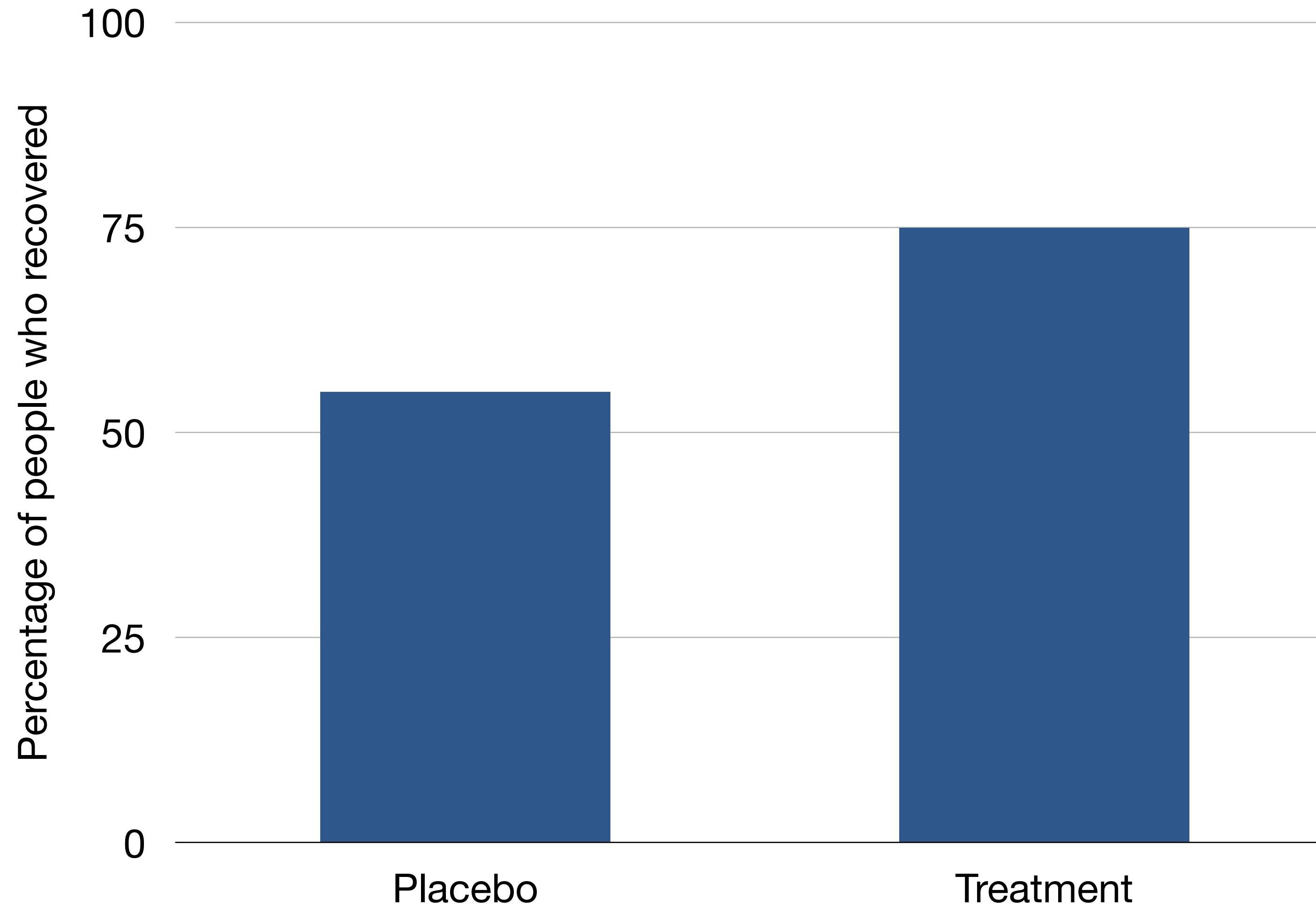
How should I visualize it?





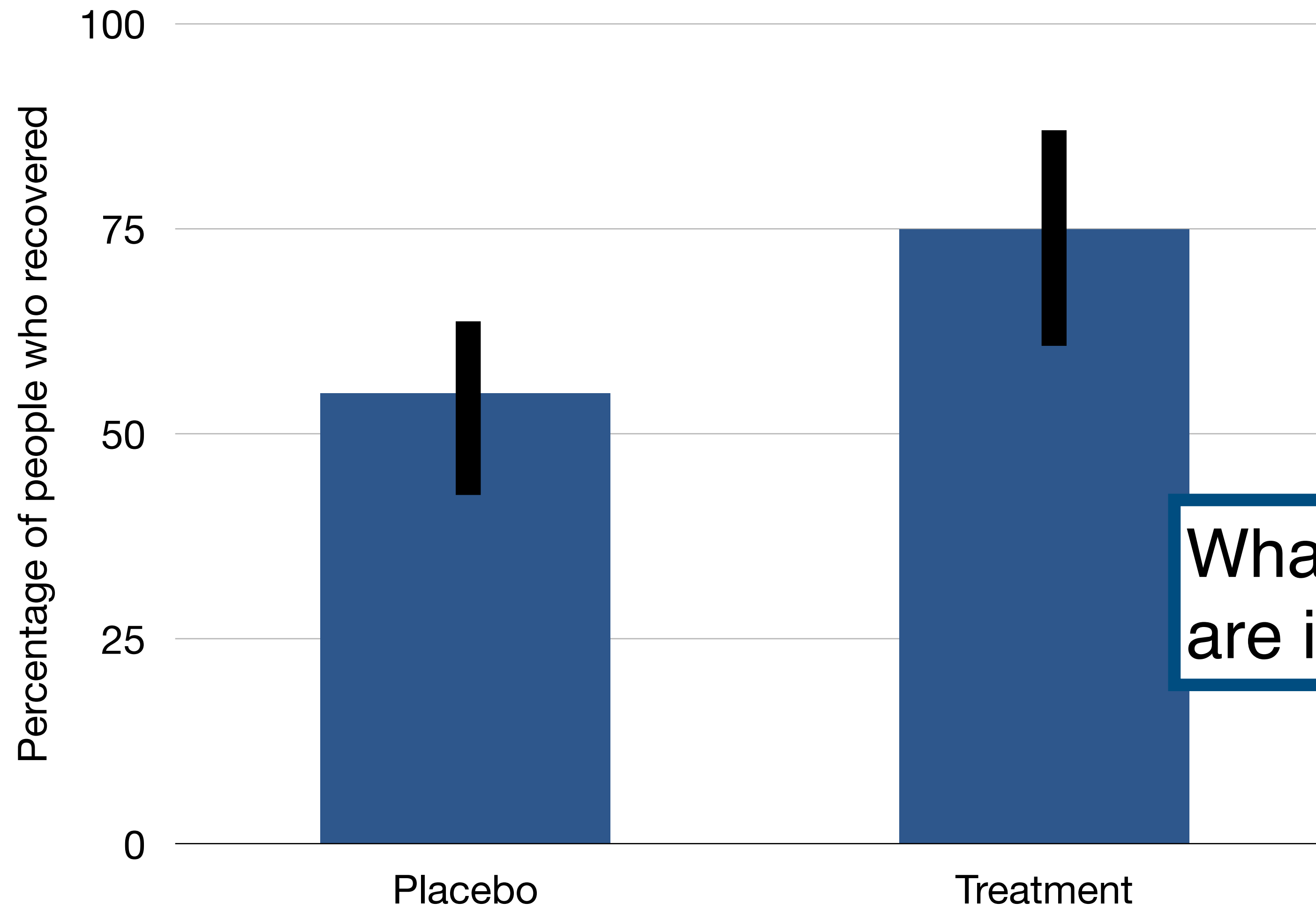


# Trial of new medicine



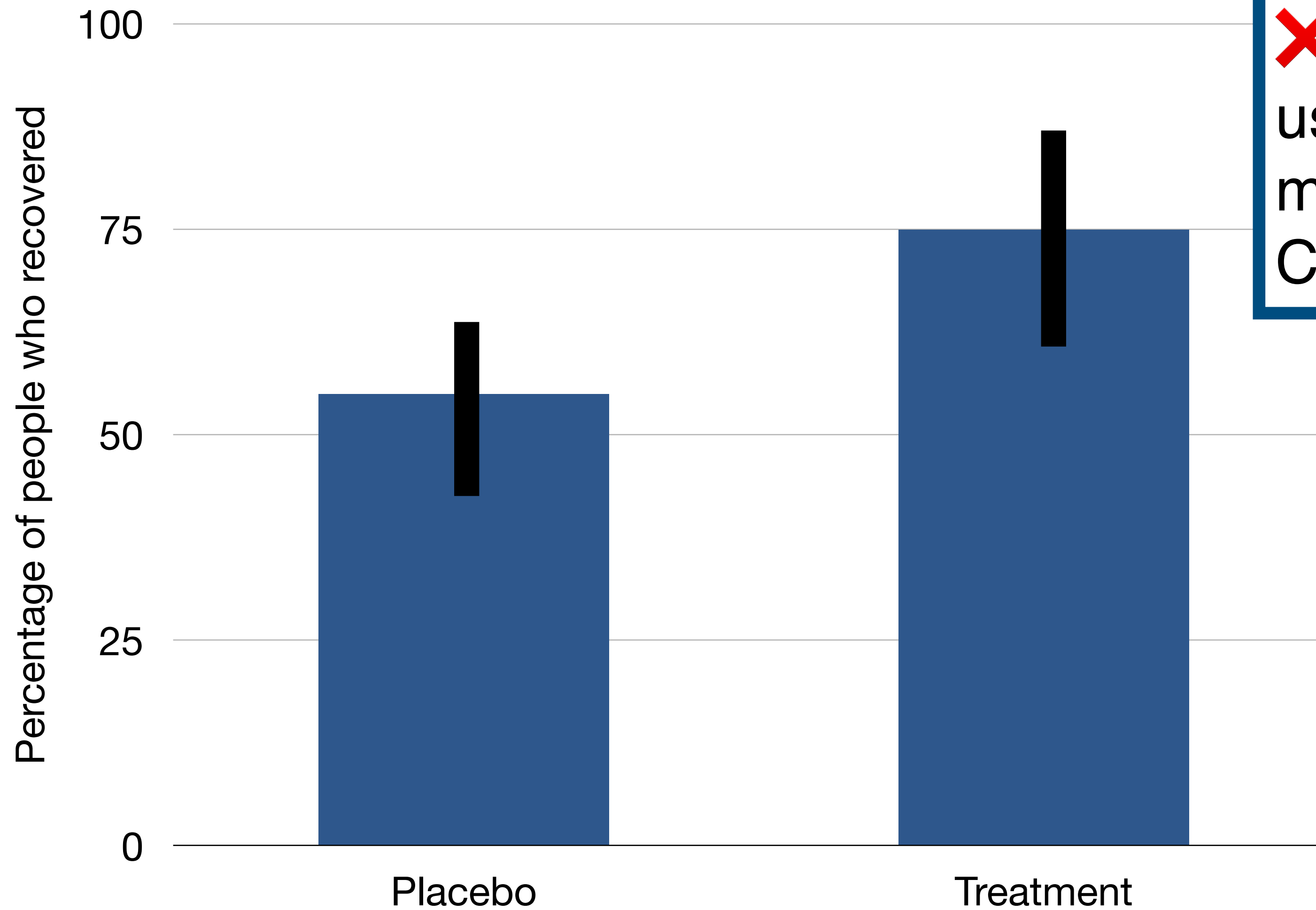


# Trial of new medicine



What do you think these error bars are implying?

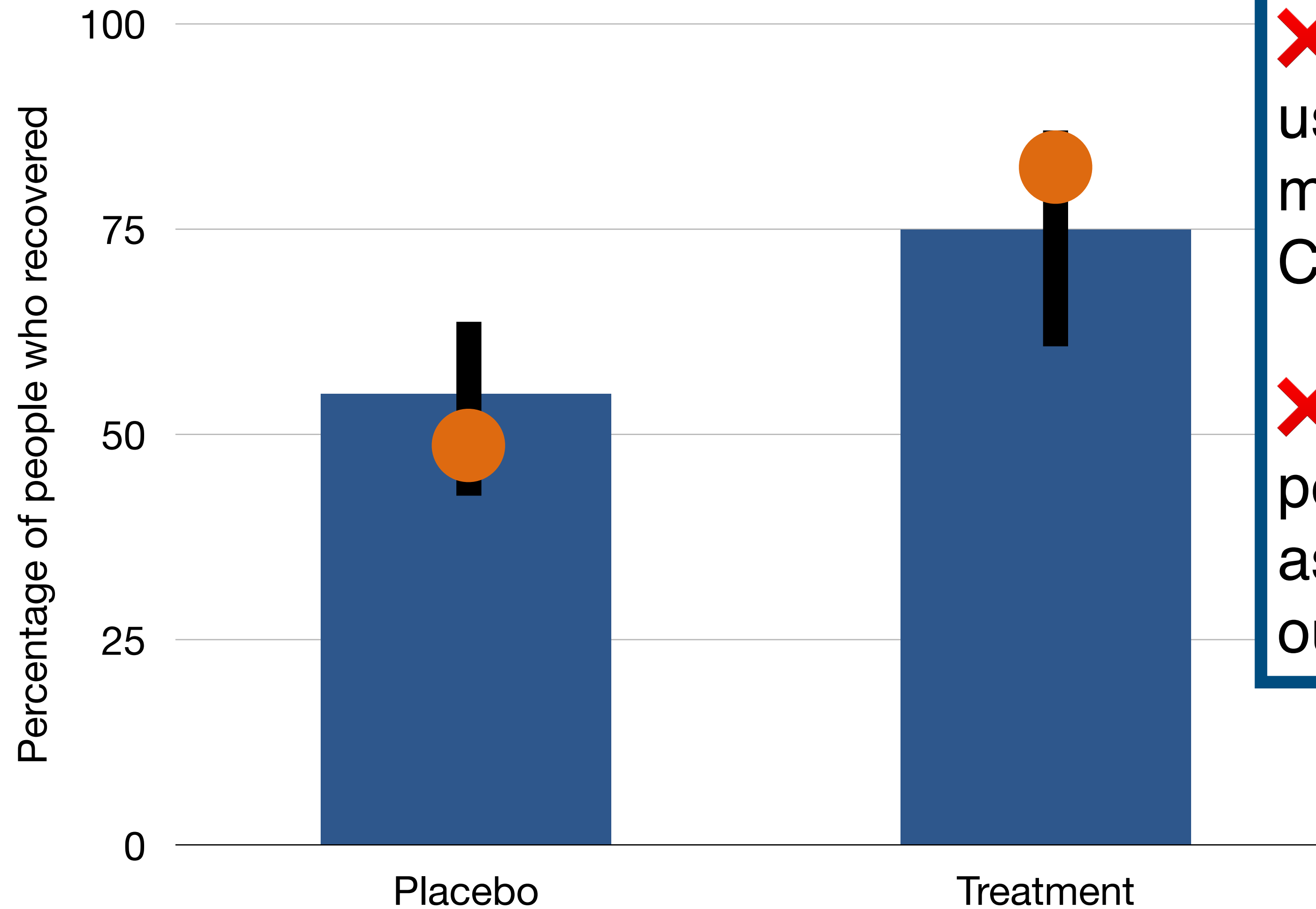
## Trial of new medicine



### Expressiveness?

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

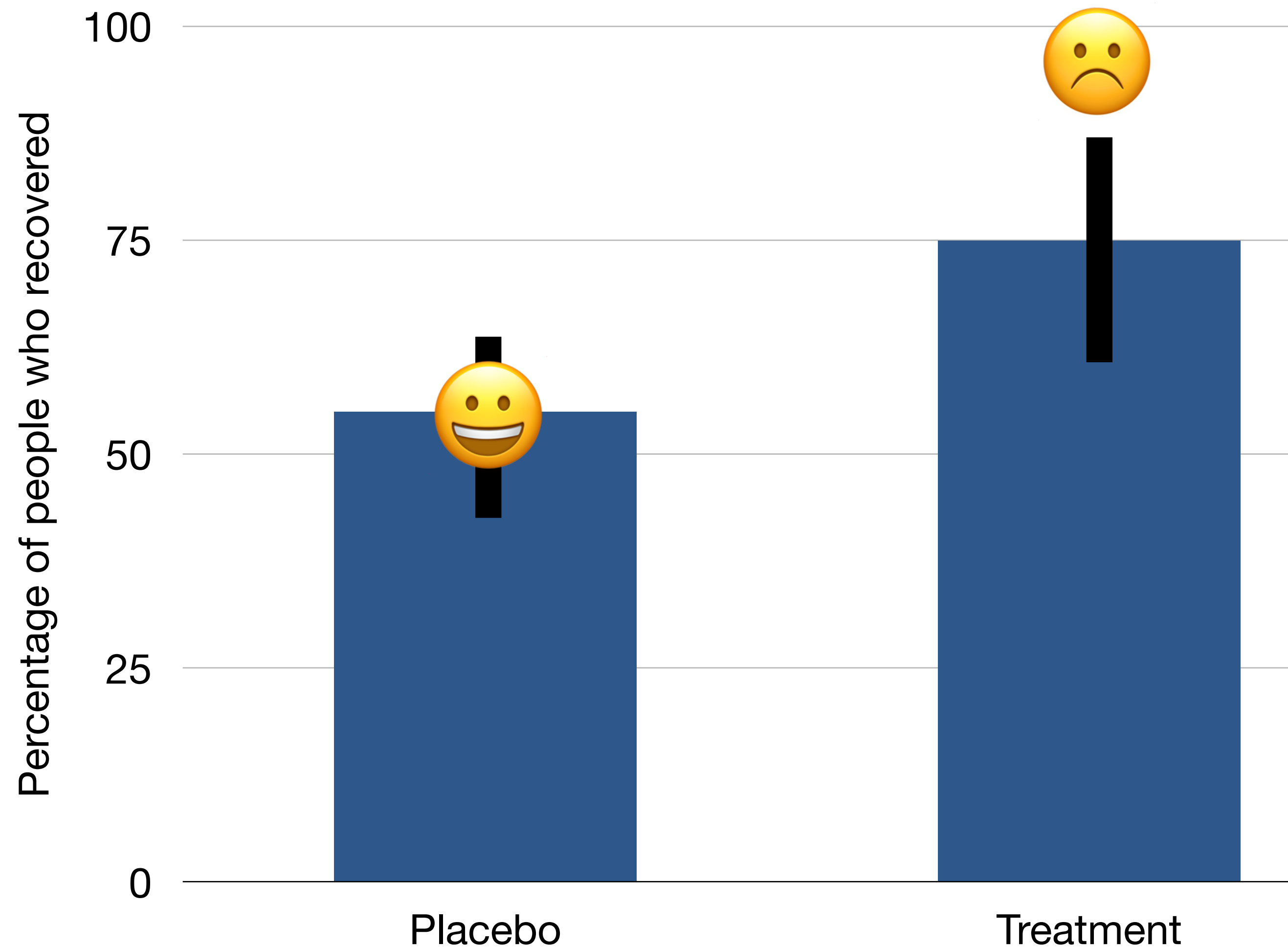
## Trial of new medicine



### Expressiveness?

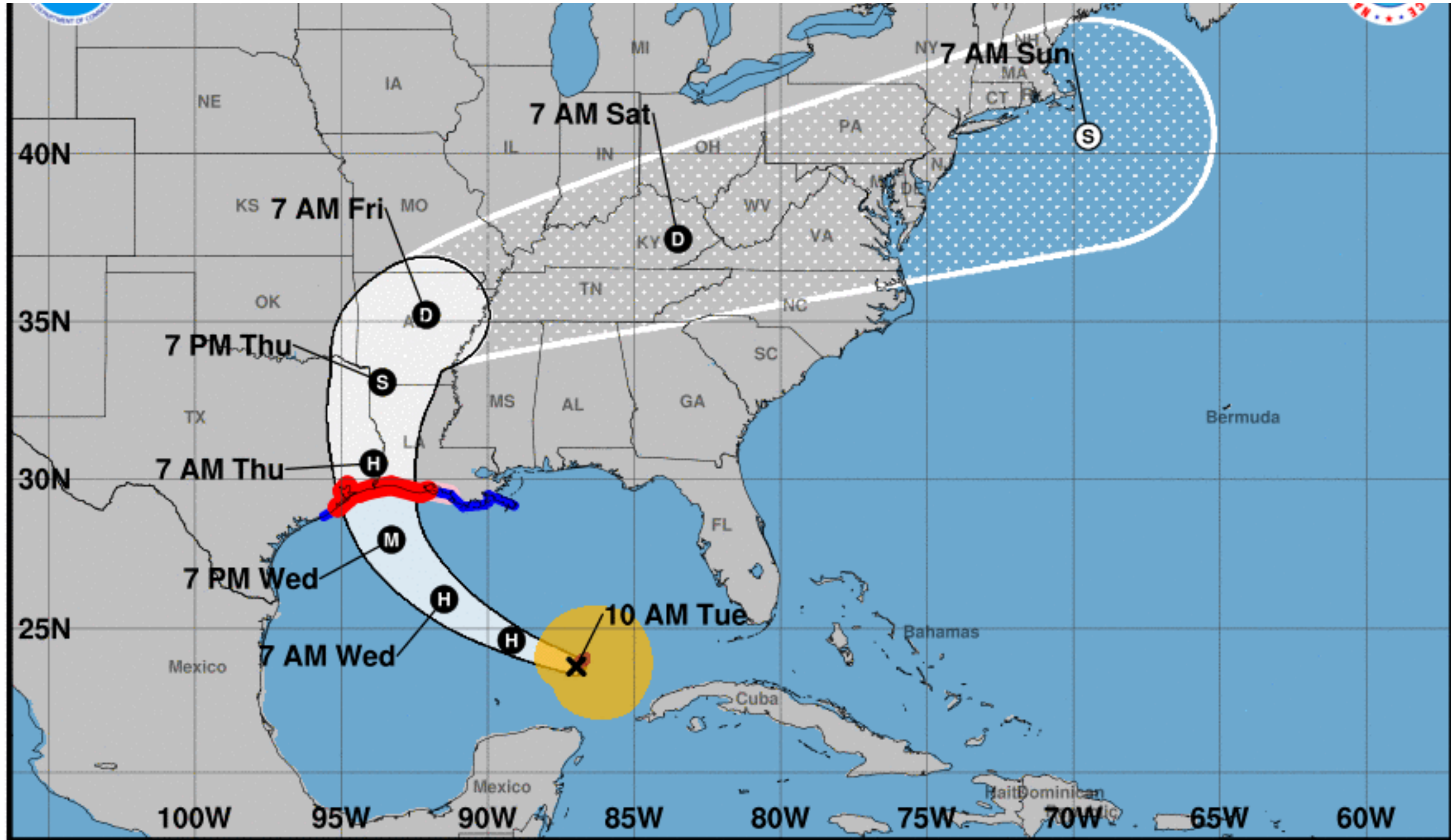
- ✗ Error bars aren't consistently used to visualize the same measure (standard error, IQR, 95% CI, etc.).
- ✗ Within-the-bar bias: people perceive points falling within the bar as more likely than those that lie outside.

## Trial of new medicine



### 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.



**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

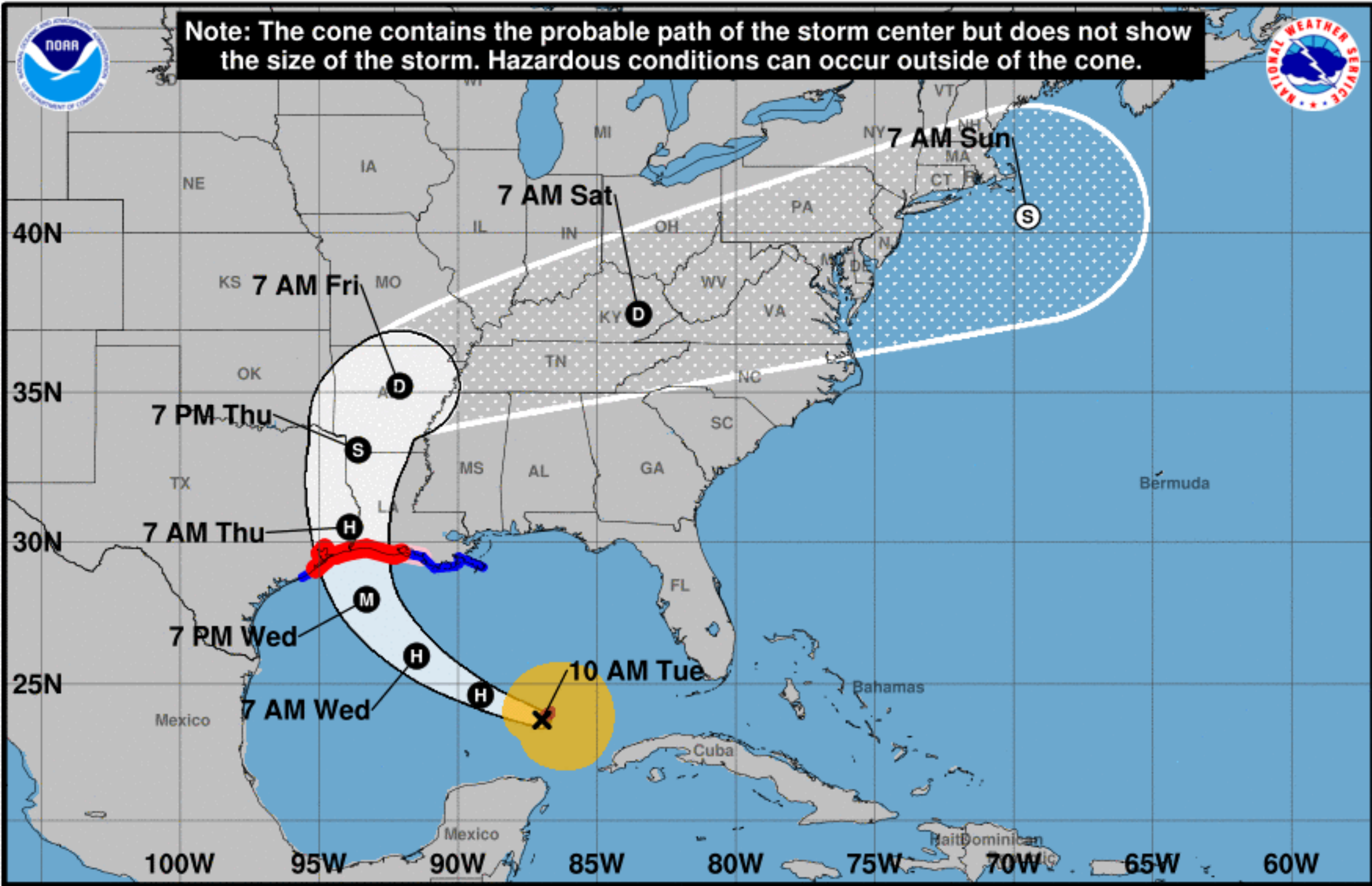
**Potential track area:**  
 Day 1-3 Day 4-5

**Watches:**  
 Hurricane Trop Stm

**Warnings:**  
 Hurricane Trop Stm

**Current wind extent:**  
 Hurricane Trop Stm

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



Note: The cone contains the probable path of the storm center but does not show the size of the storm. Hazardous conditions can occur outside of the cone.

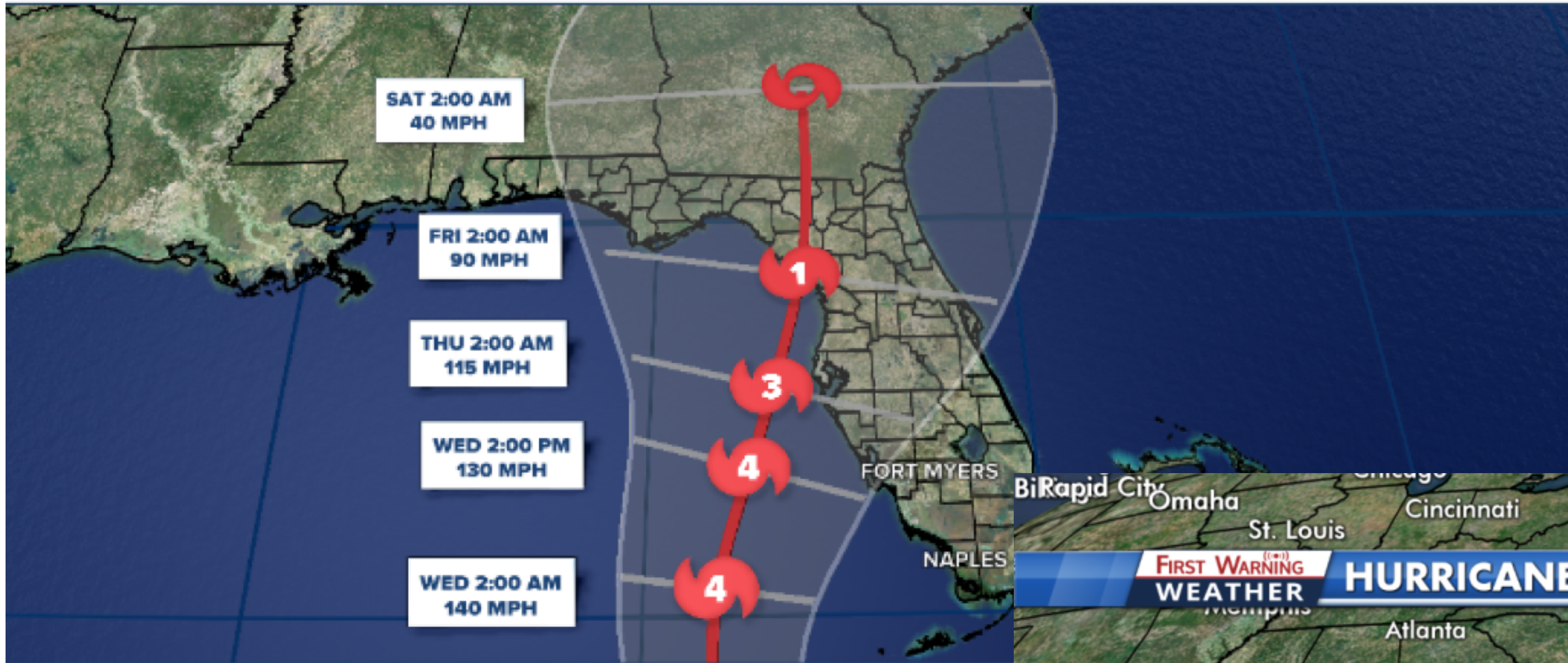
<b>Hurricane Laura</b> Tuesday August 25, 2020 10 AM CDT Advisory 23 NWS National Hurricane Center		<b>Current information: x</b> Center location 23.7 N 87.0 W Maximum sustained wind 75 mph Movement WNW at 16 mph		<b>Forecast positions:</b> ● Tropical Cyclone    ○ Post/Potential TC Sustained winds:    D < 39 mph S 39-73 mph    H 74-110 mph    M > 110 mph	
<b>Potential track area:</b> □ Day 1-3    ▨ Day 4-5	<b>Watches:</b> ■ Hurricane    ■ Trop Stm	<b>Warnings:</b> ■ Hurricane    ■ Trop Stm	<b>Current wind extent:</b> ■ Hurricane    ■ Trop Stm		

What is being visualized?

What are the strengths and weaknesses of this visualization?

# FOX 4 HURRICANE IAN

## 5:00 AM ADVISORY



CURRENT LOCATION: 90 MI SW OF G



### FIRST WARNING WEATHER HURRICANE DORIAN



# Hurricane Dorian Forecast Track and Intensity



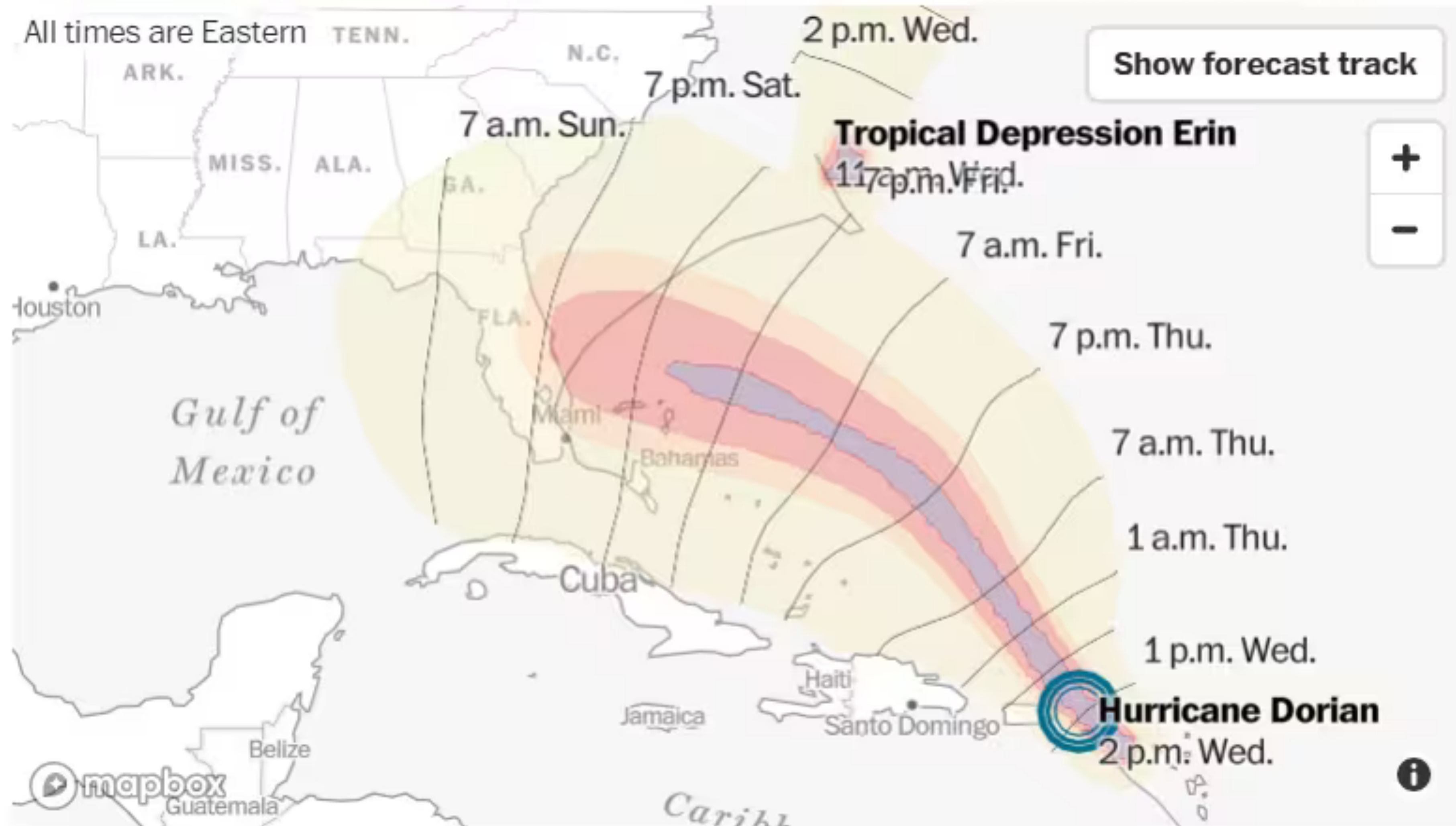


### Five-day chance of tropical-storm-force winds



Current extent of tropical-storm-force winds

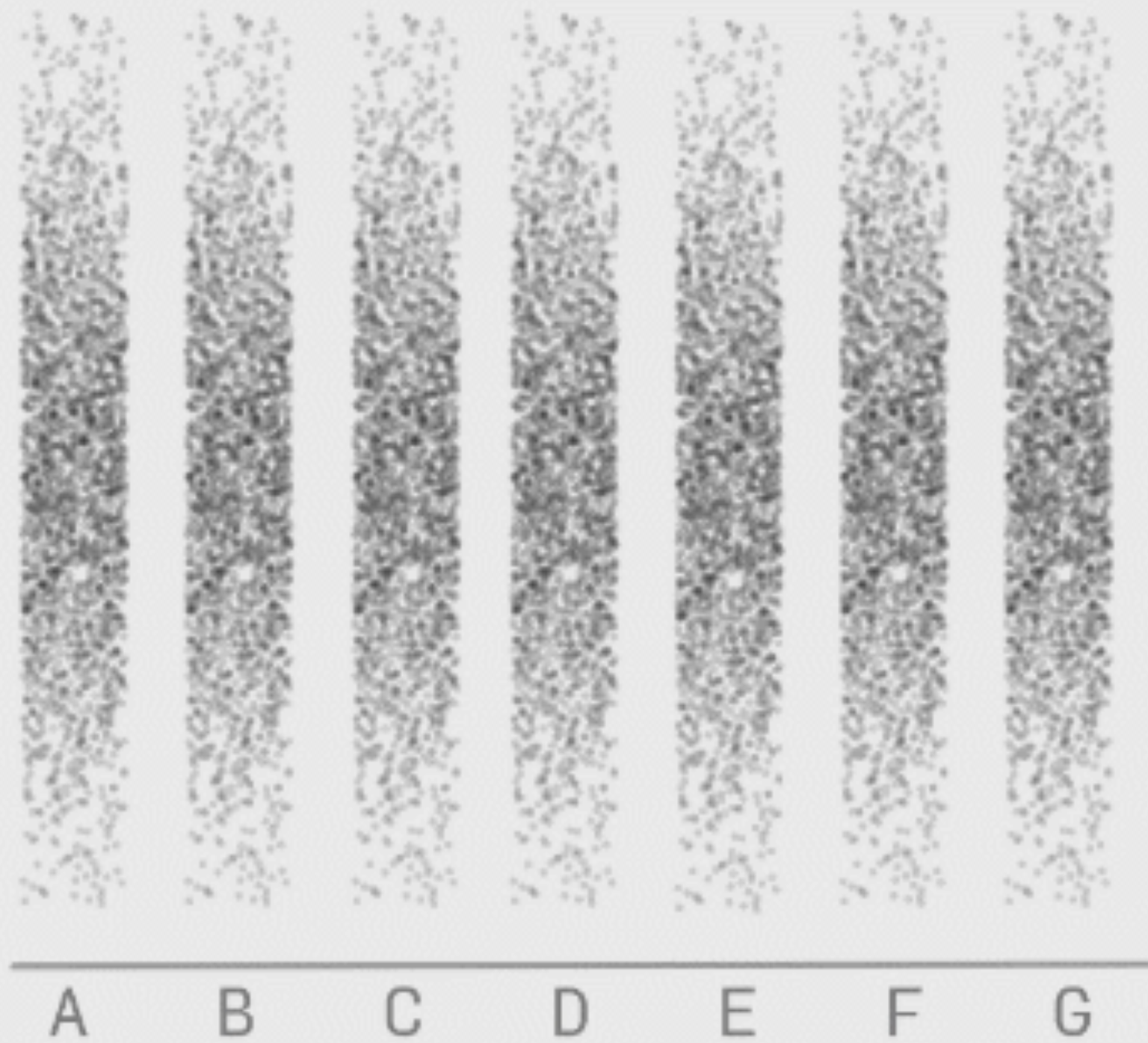
Major hurricane (>110 mph)  
 Hurricane (74-110 mph)  
 Tropical storm (39-73 mph)  
 Tropical depression (<39 mph)



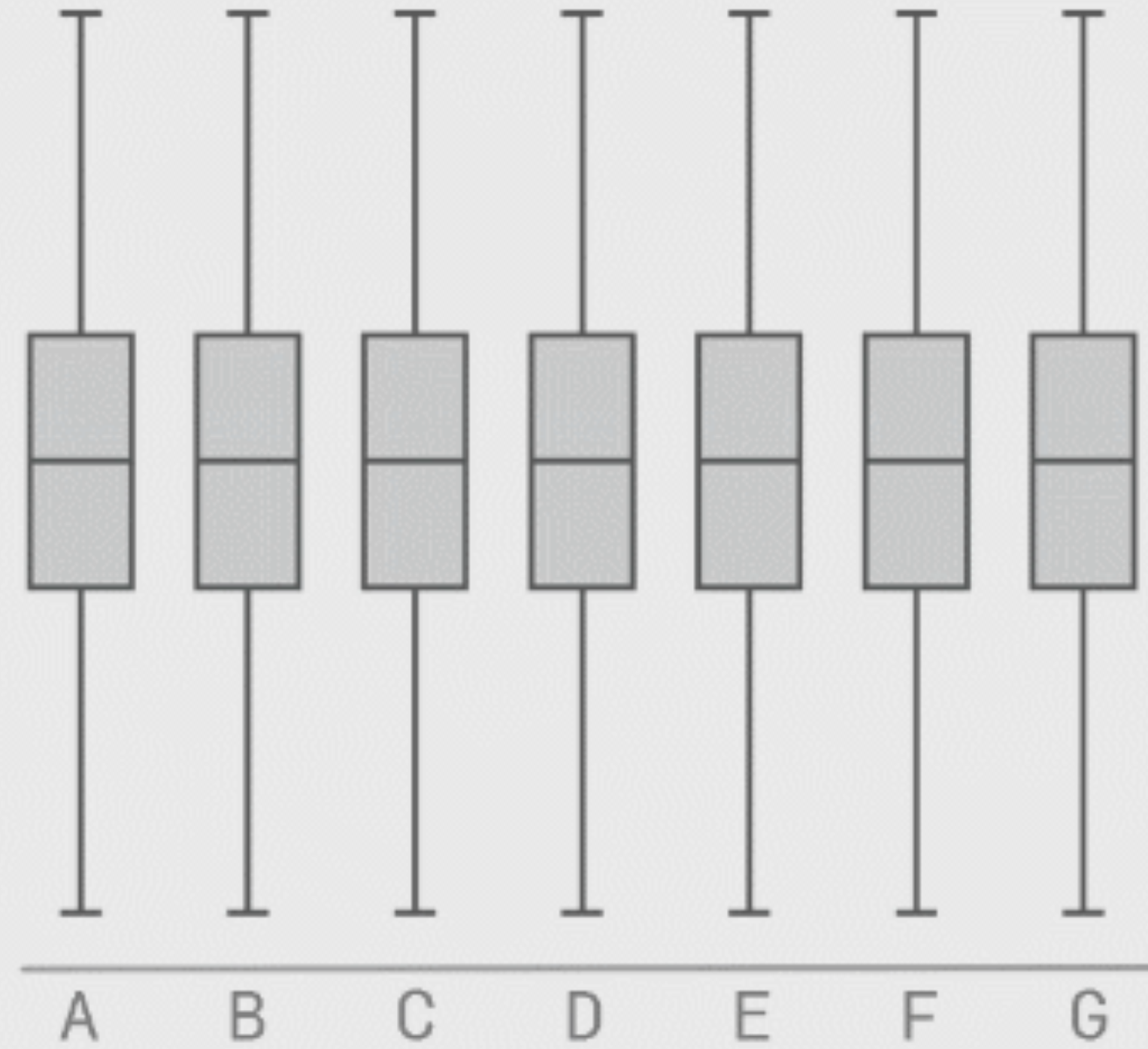
Source: National Weather Service. Note: Impact lines represent the earliest reasonable arrival time of tropical-storm-force winds.

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

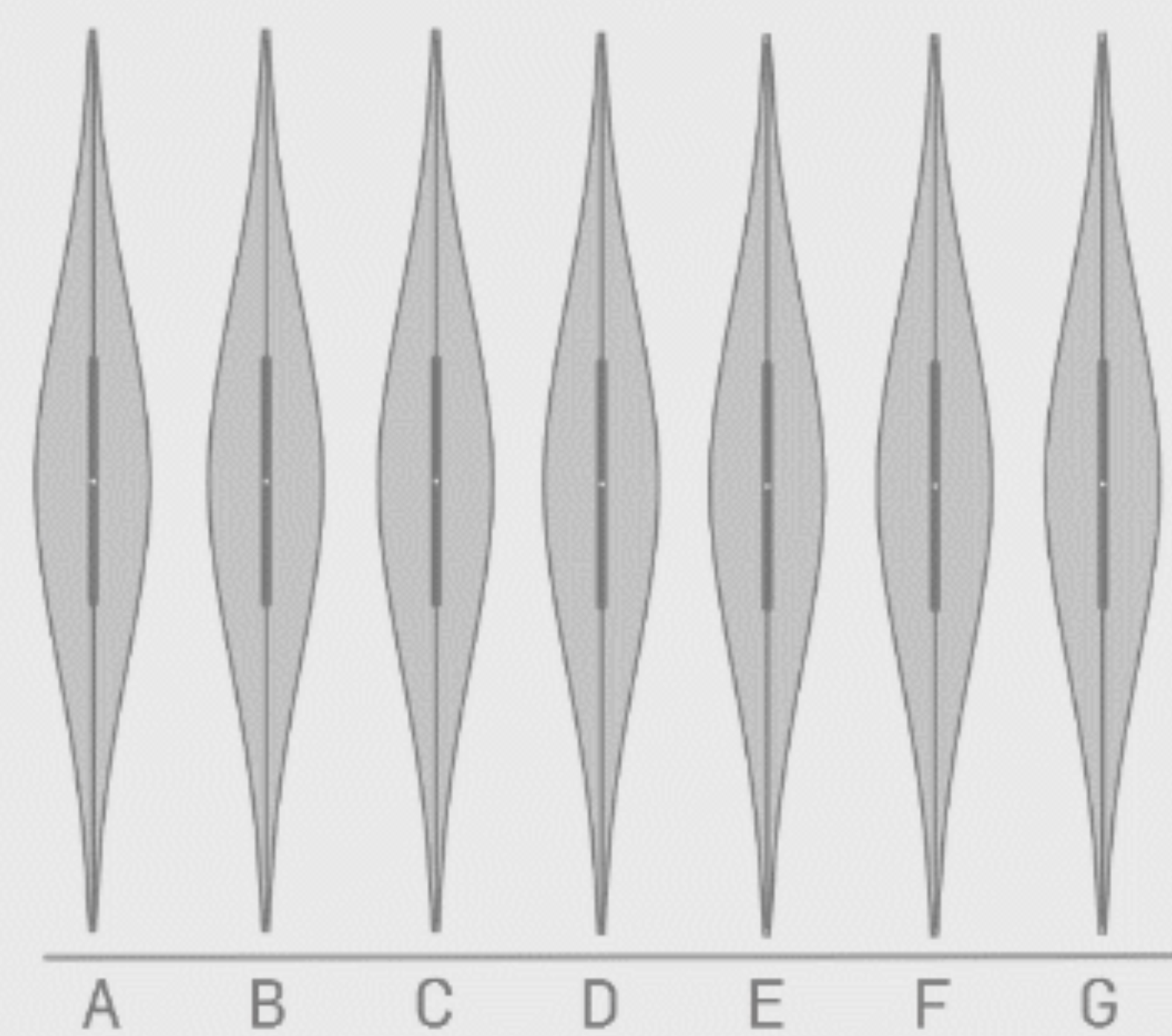
Raw Data



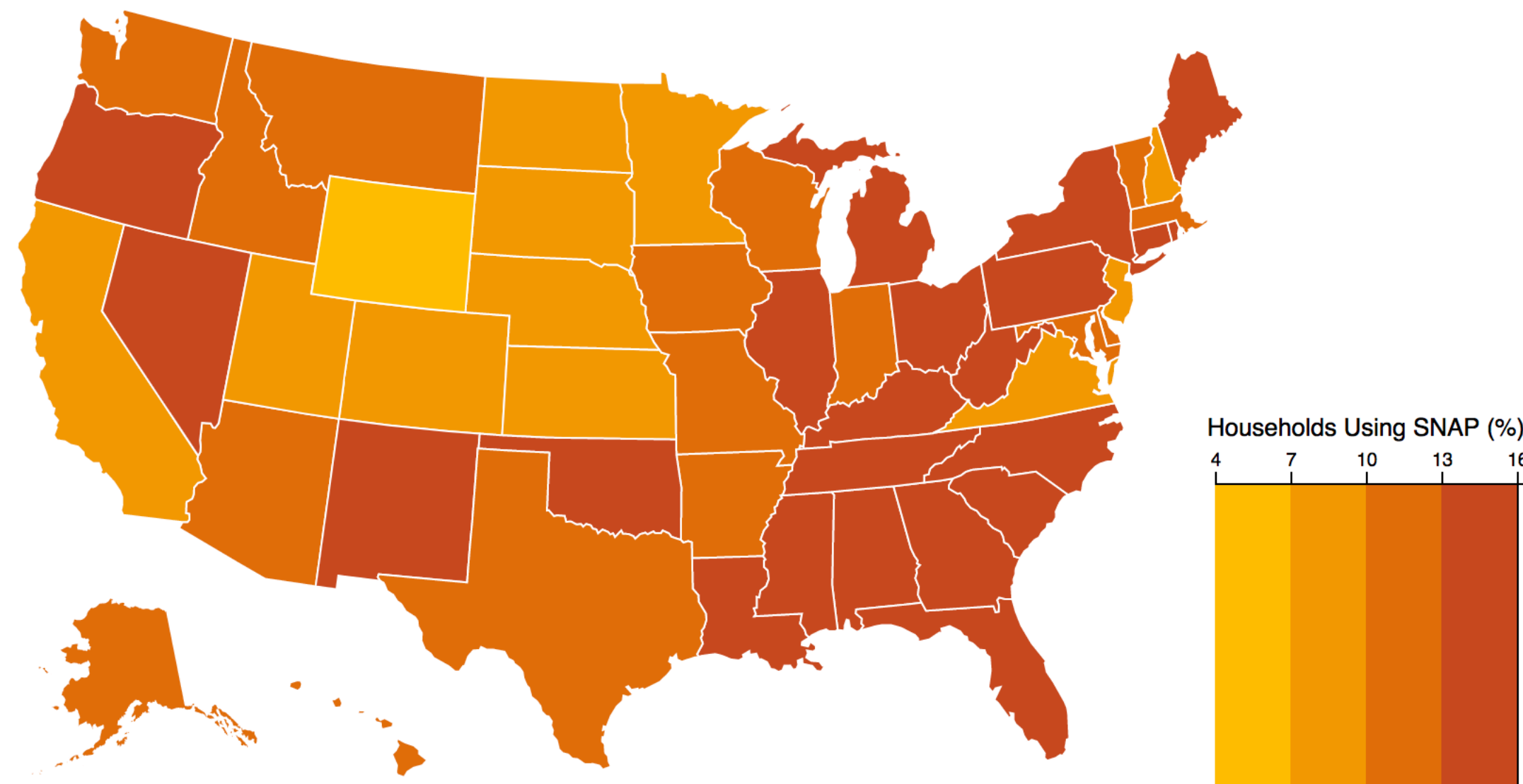
Box-plot of the Data



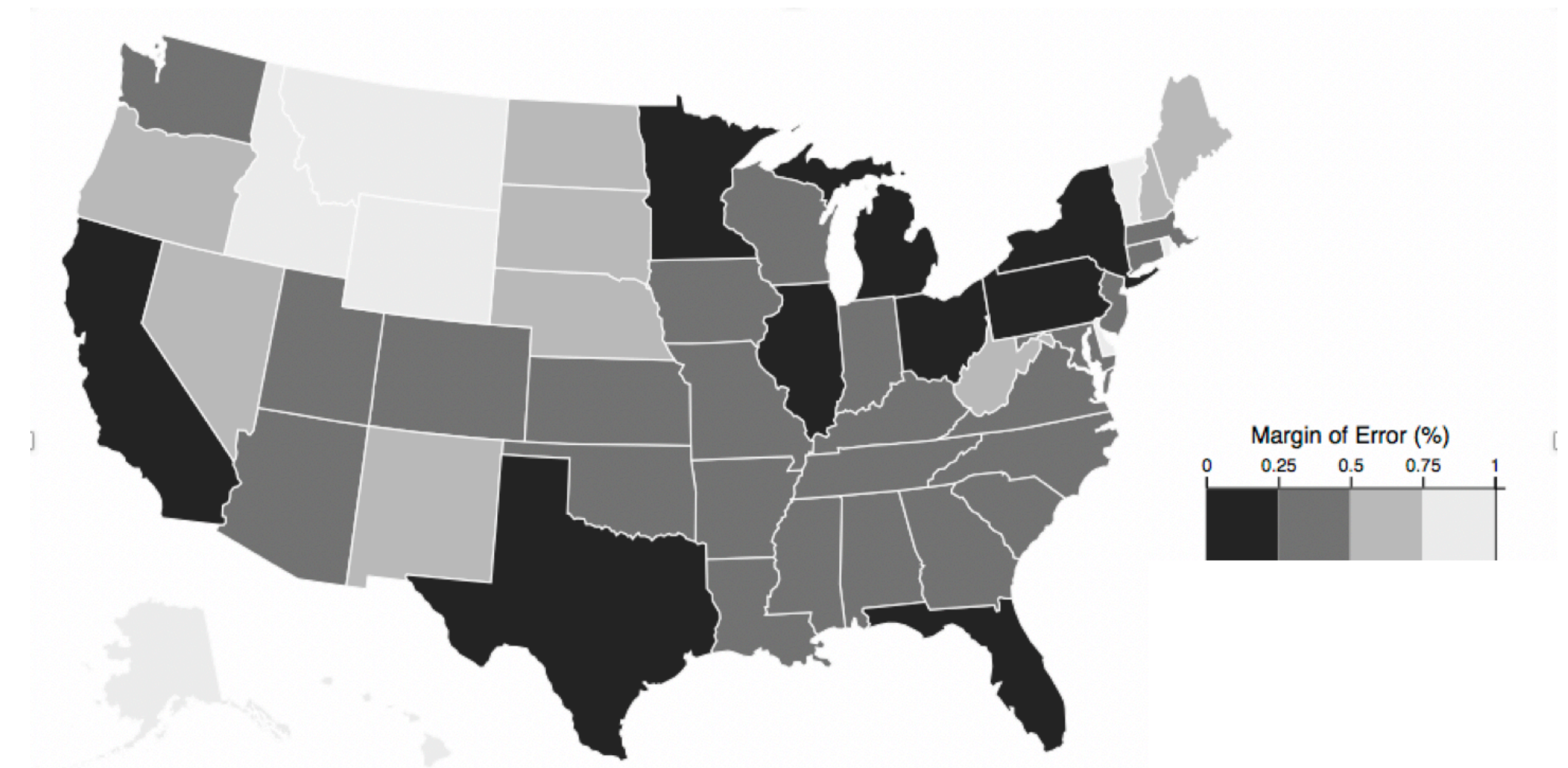
Violin-plot of the Data



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

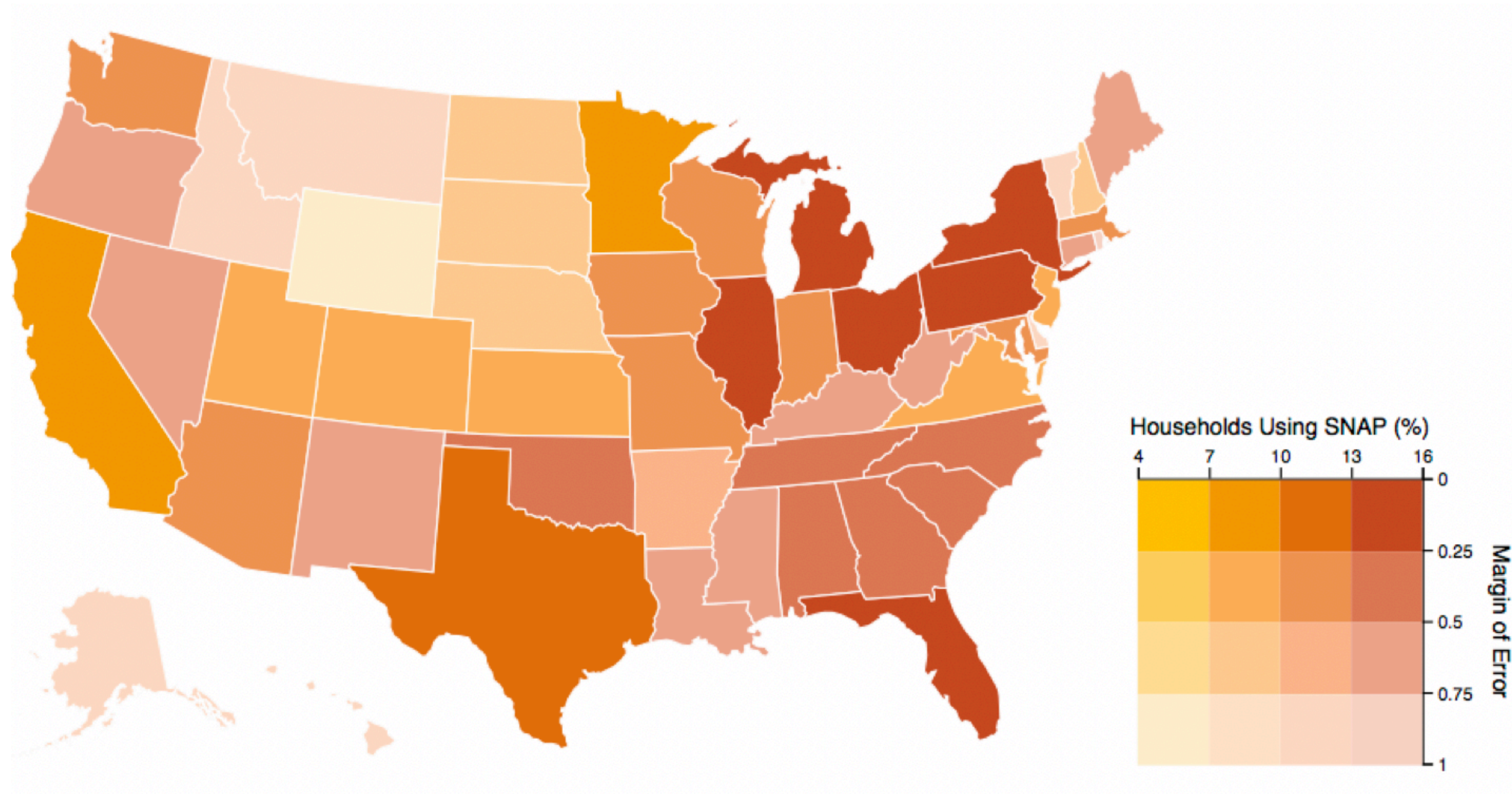


Data Map

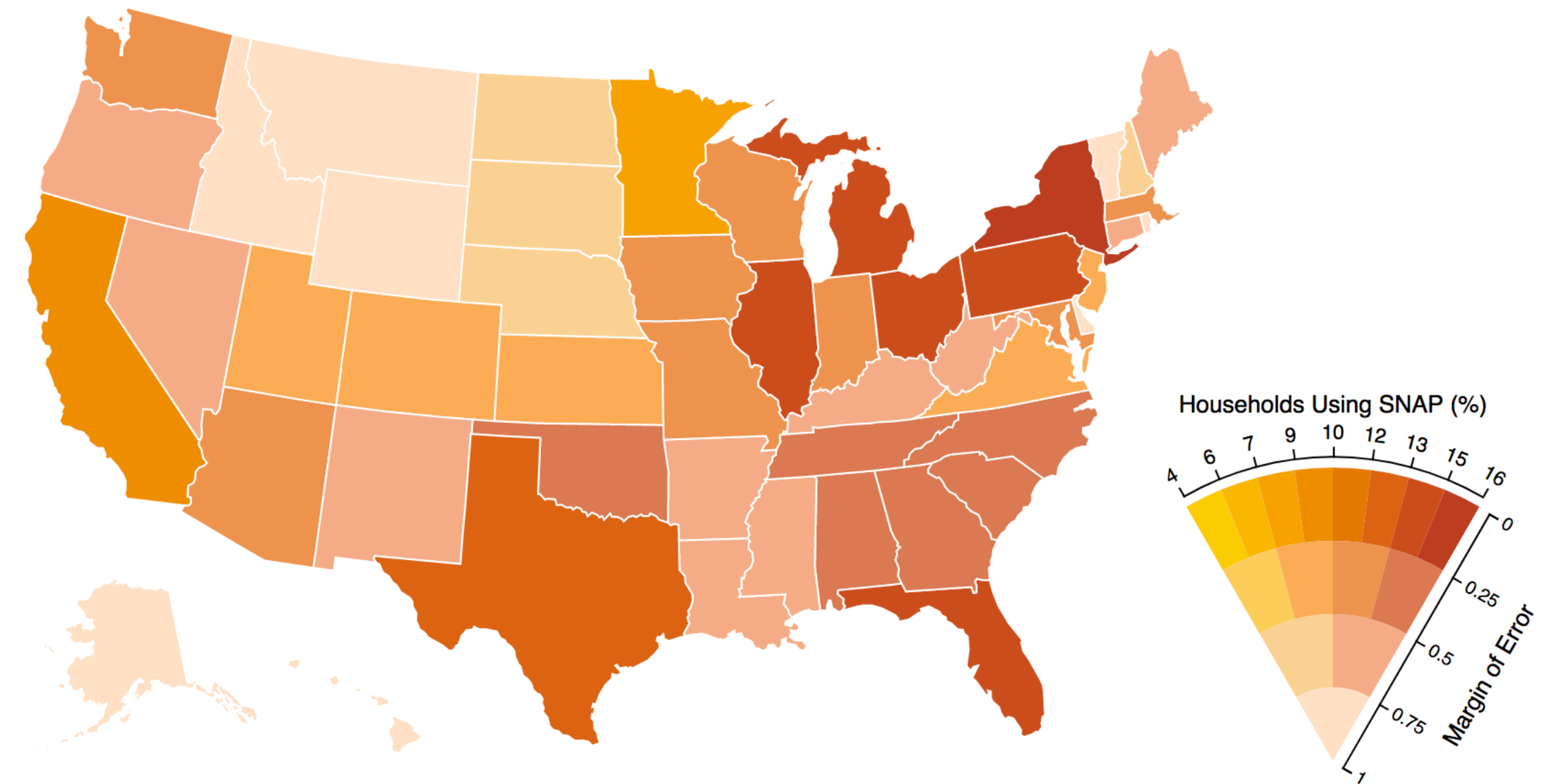


Uncertainty Map

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

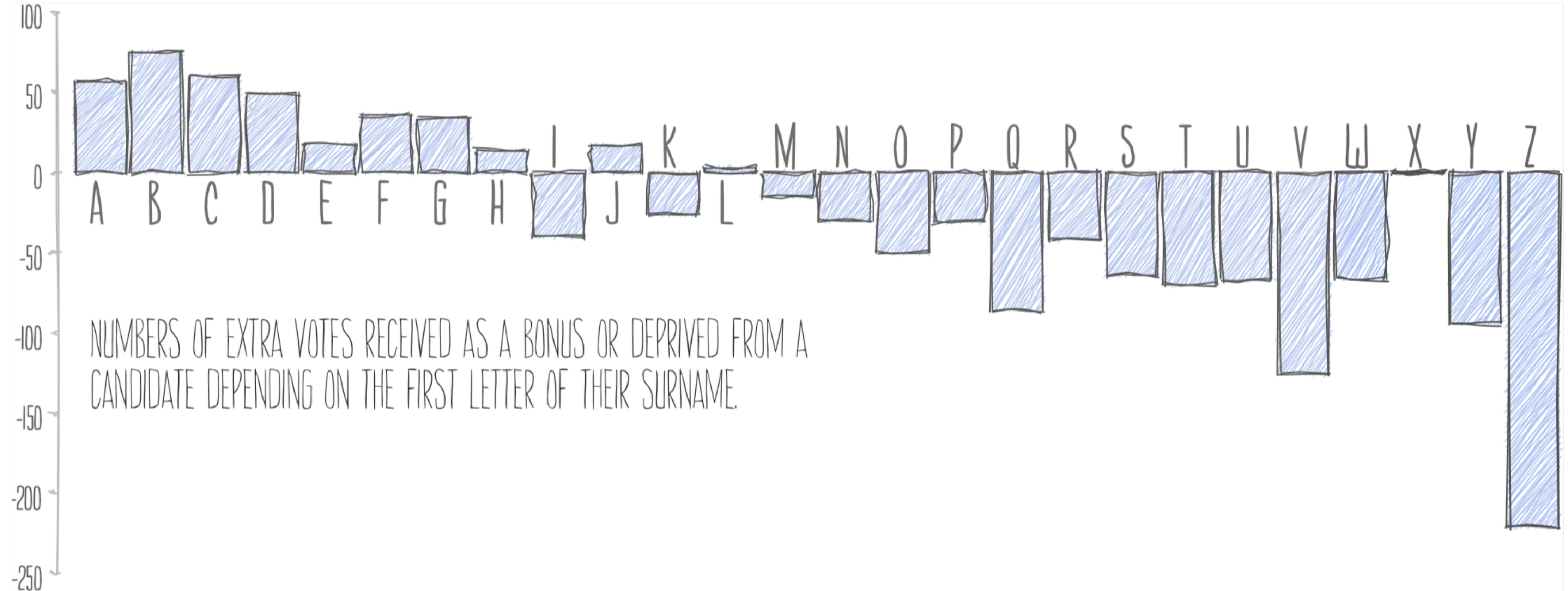


Bivariate Map (Data + Uncertainty)



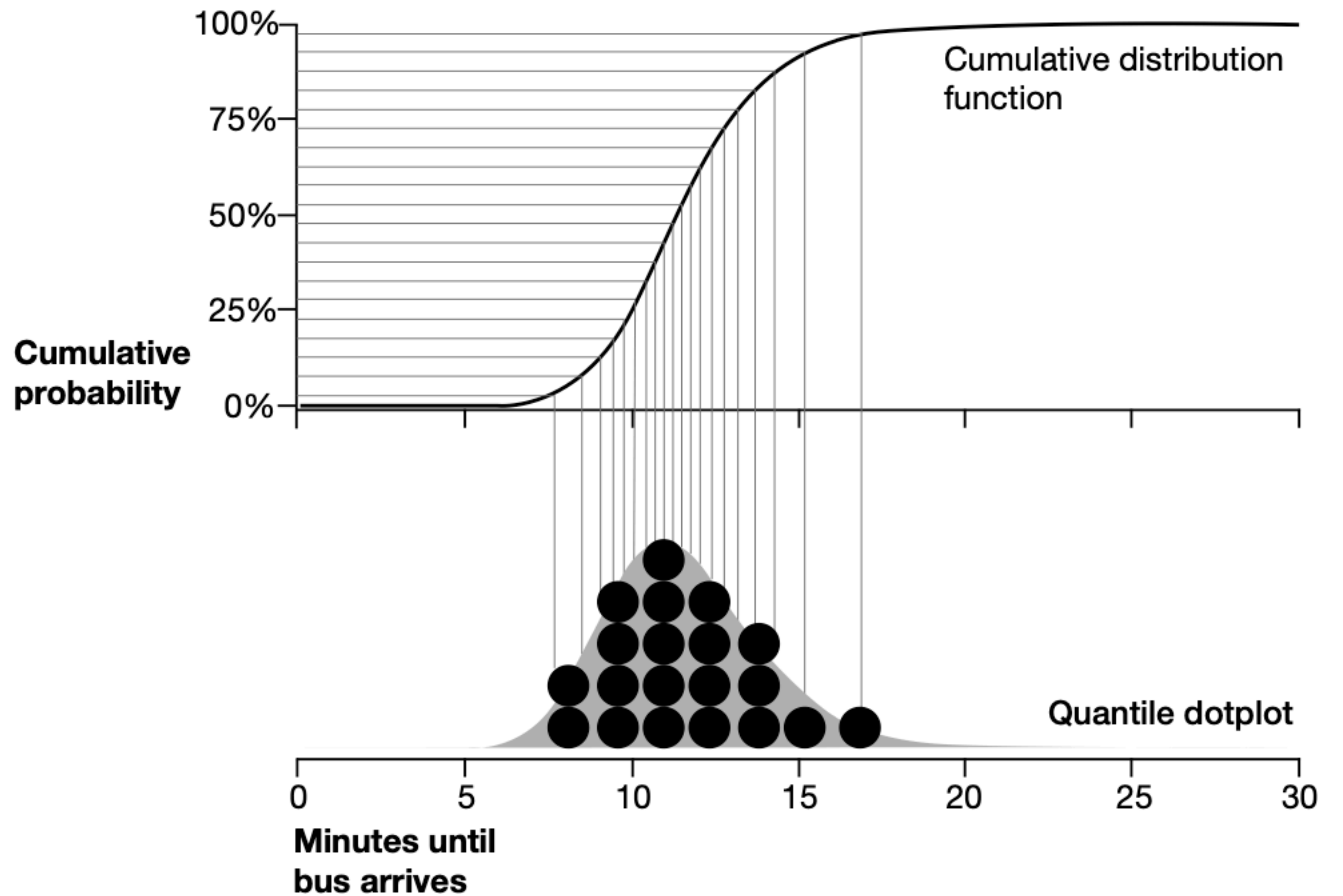
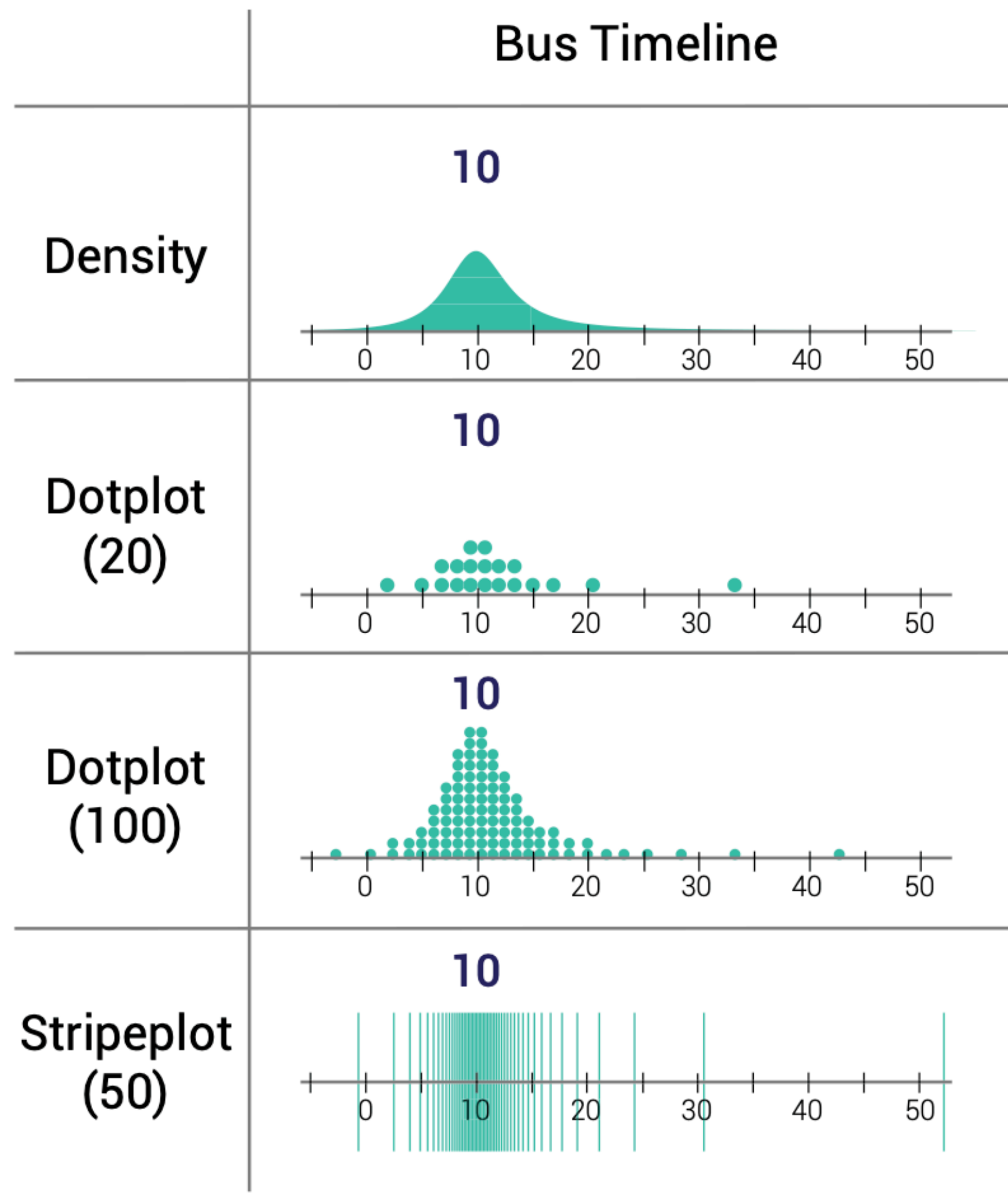
Value-Suppressing Uncertainty Map

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

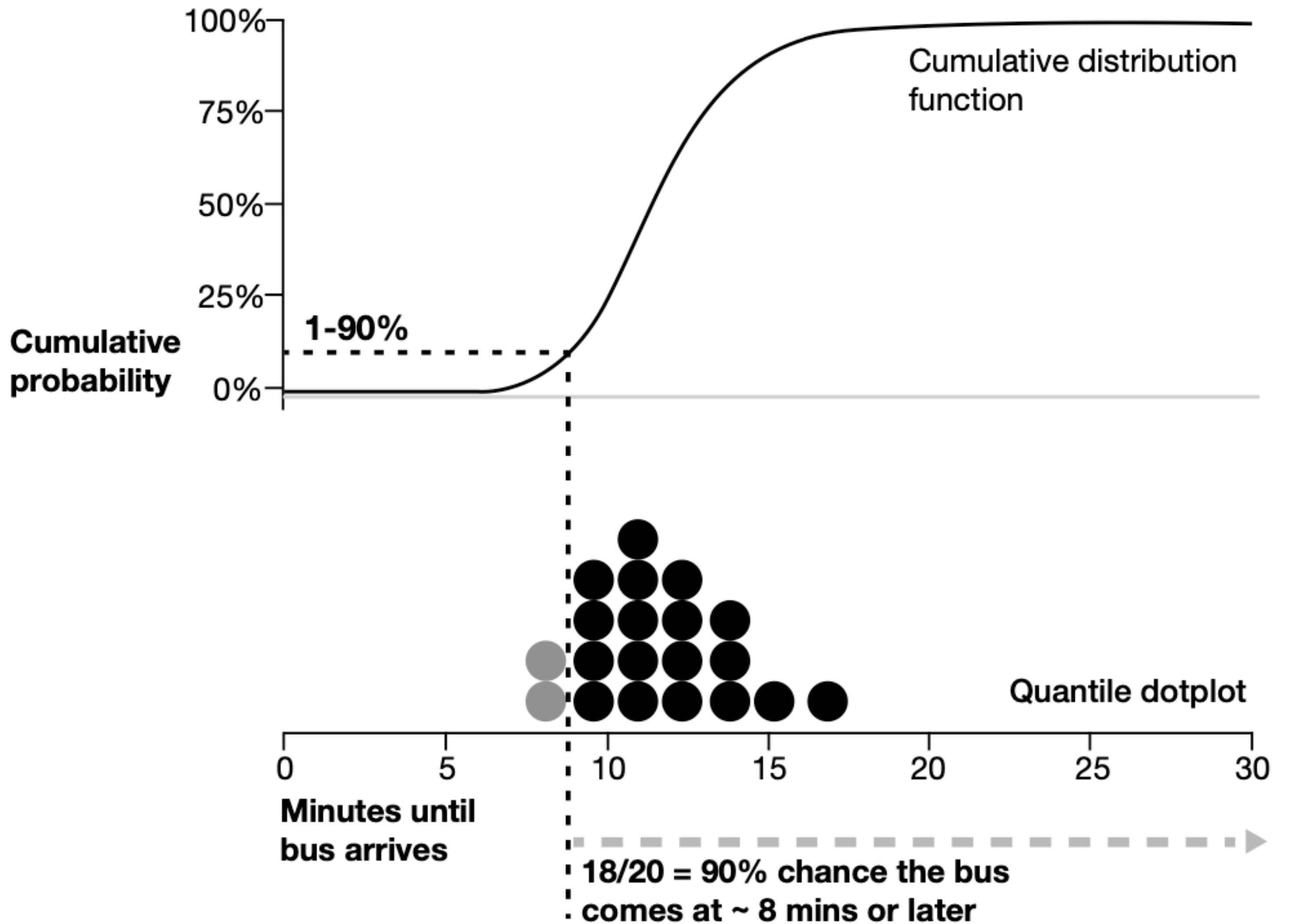
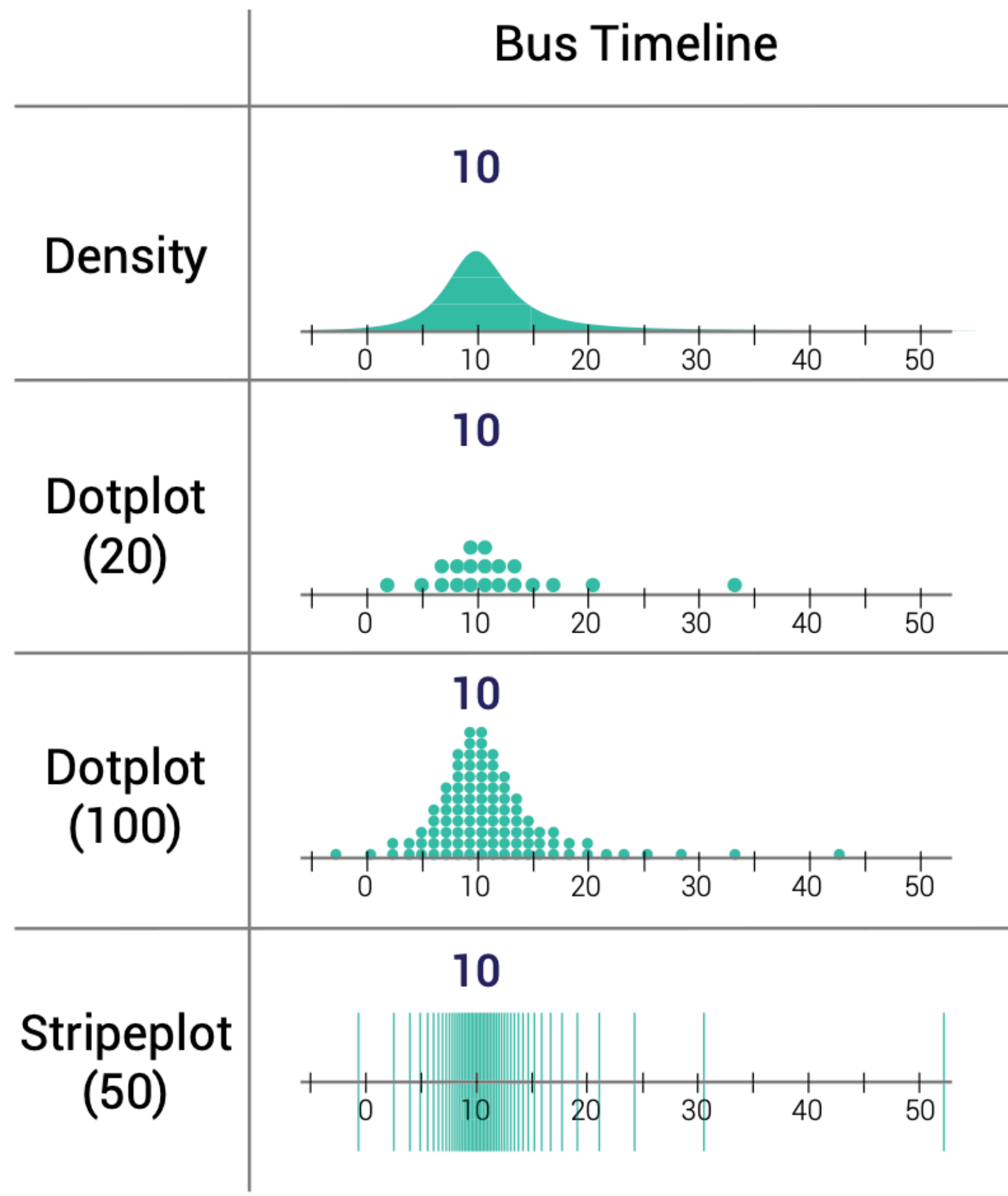


[Wood et al., 2012]  
[Boukhelifa et al., 2012]

# "Set of draws" technique

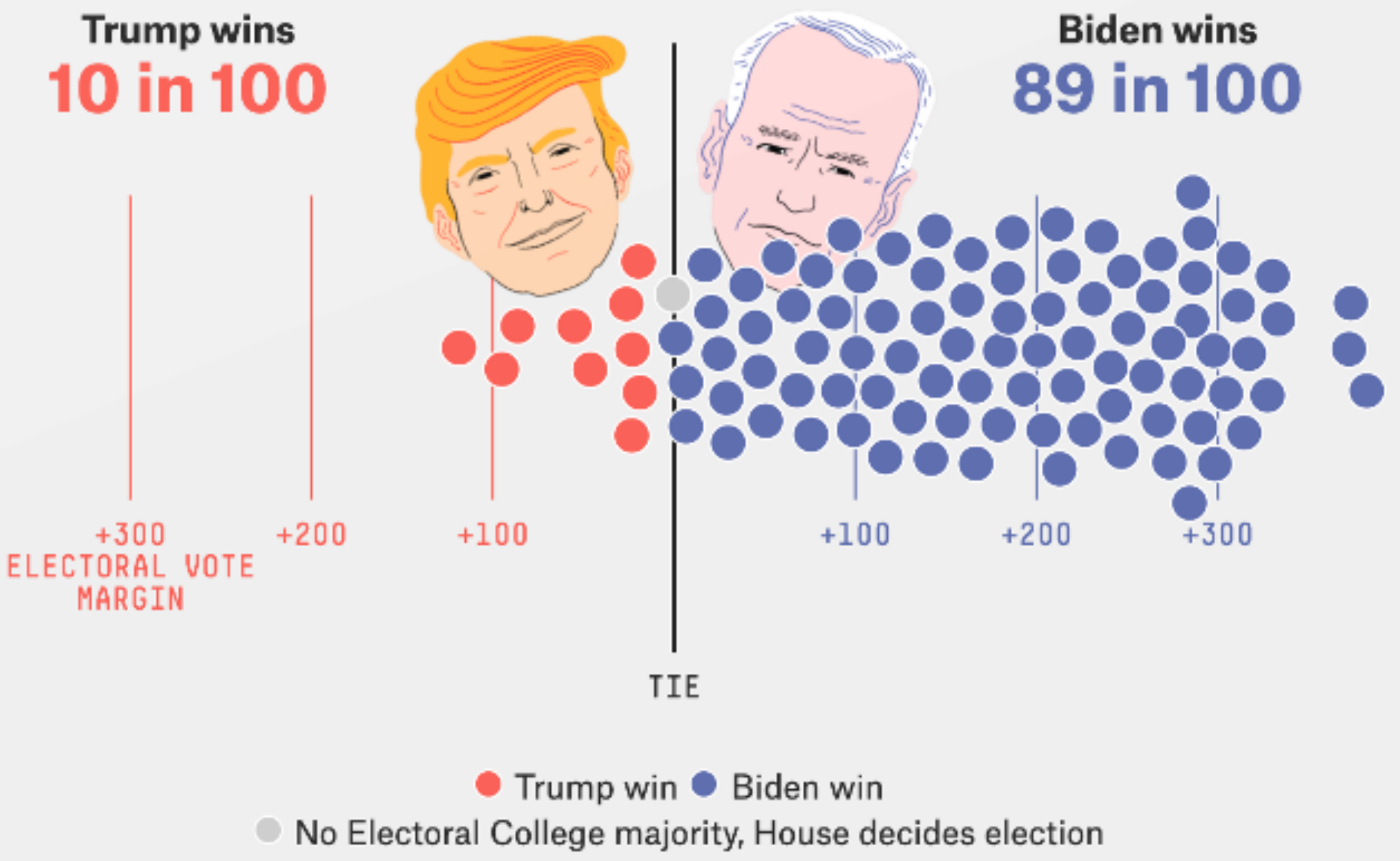


# "Set of draws" technique



### 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.

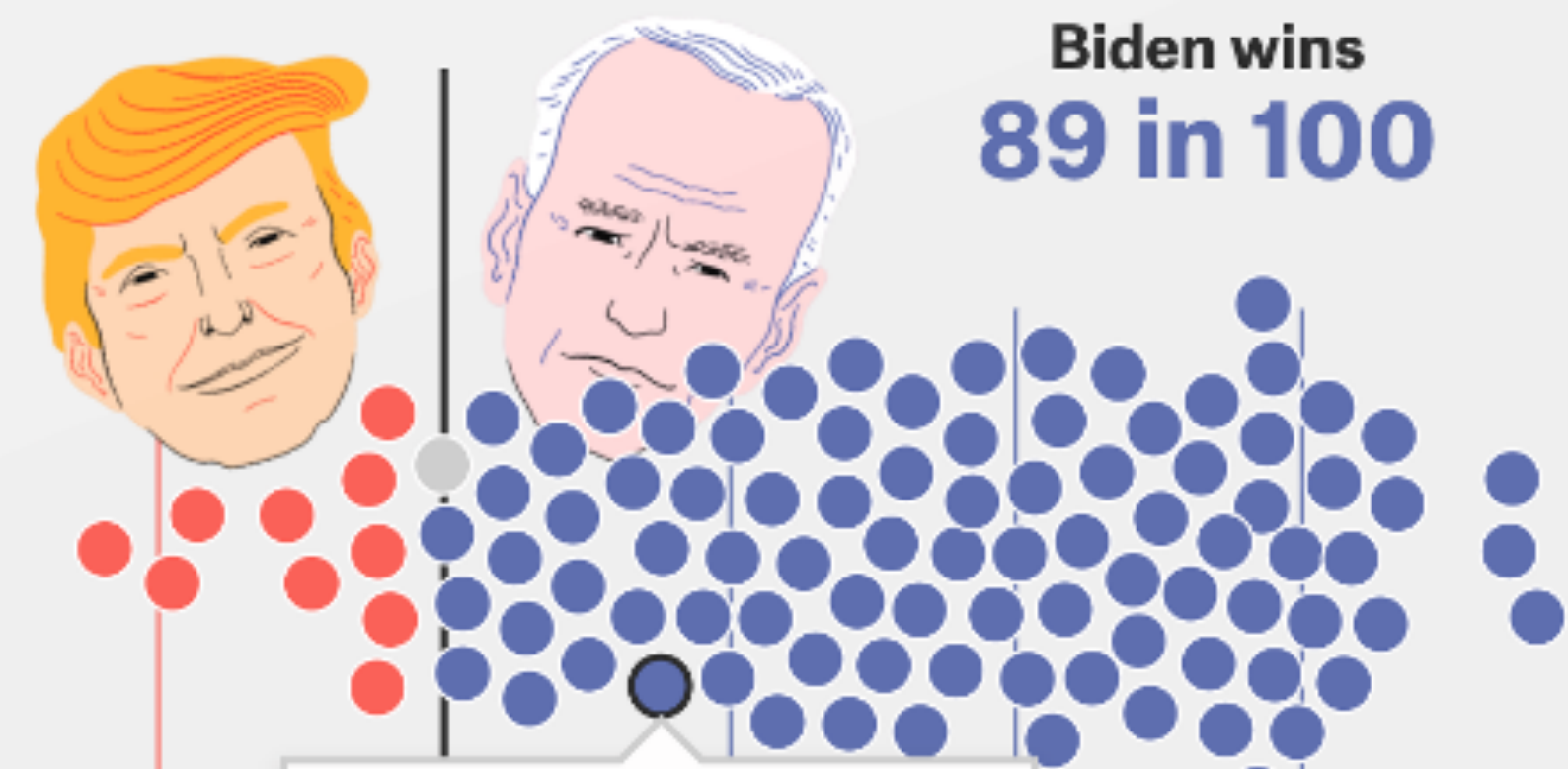


# Biden is *favored* to win the election

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.

Trump wins  
**10 in 100**

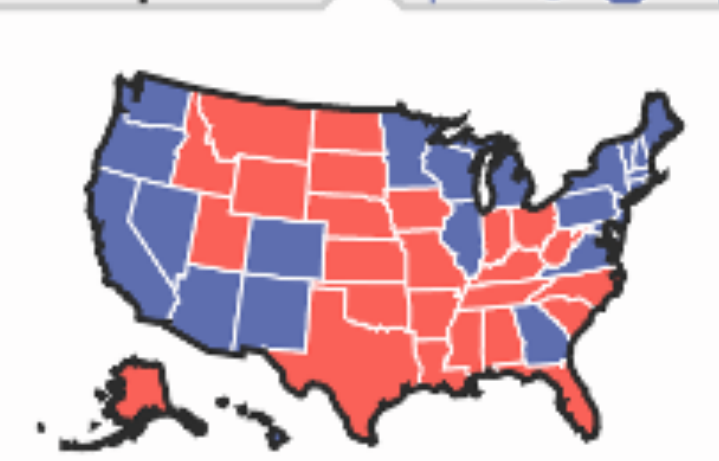
Biden wins  
**89 in 100**



+300  
+200  
+100  
ELECTORAL VOTE MARGIN

0  
+300

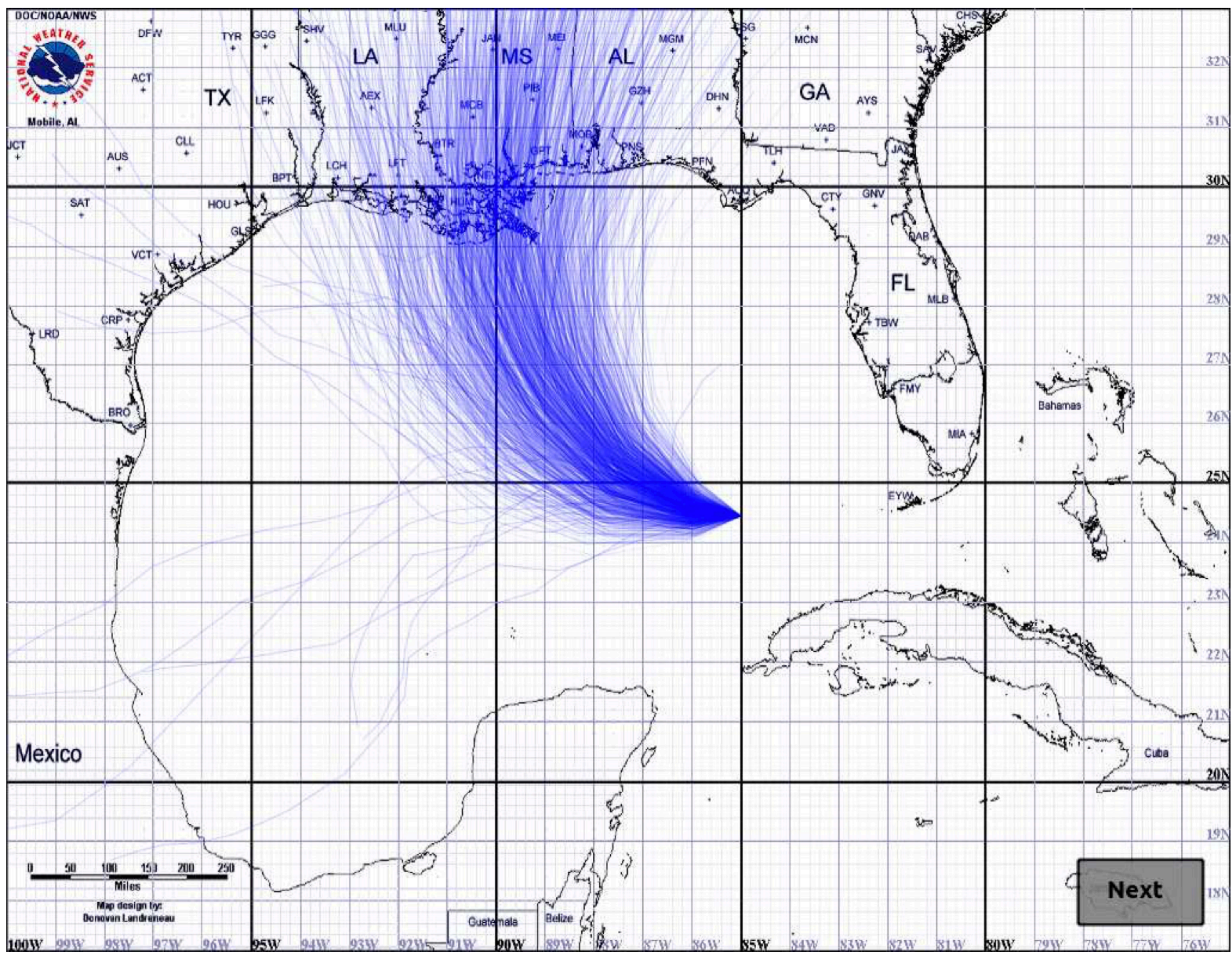
Don't count the underdog out! Upset wins are surprising but not impossible.



— ELECTORAL VOTES —

Biden ✓	Trump
306	232

● Trump  
● No Electoral College



**• TheUpshot**

**STATISTICAL NOISE**

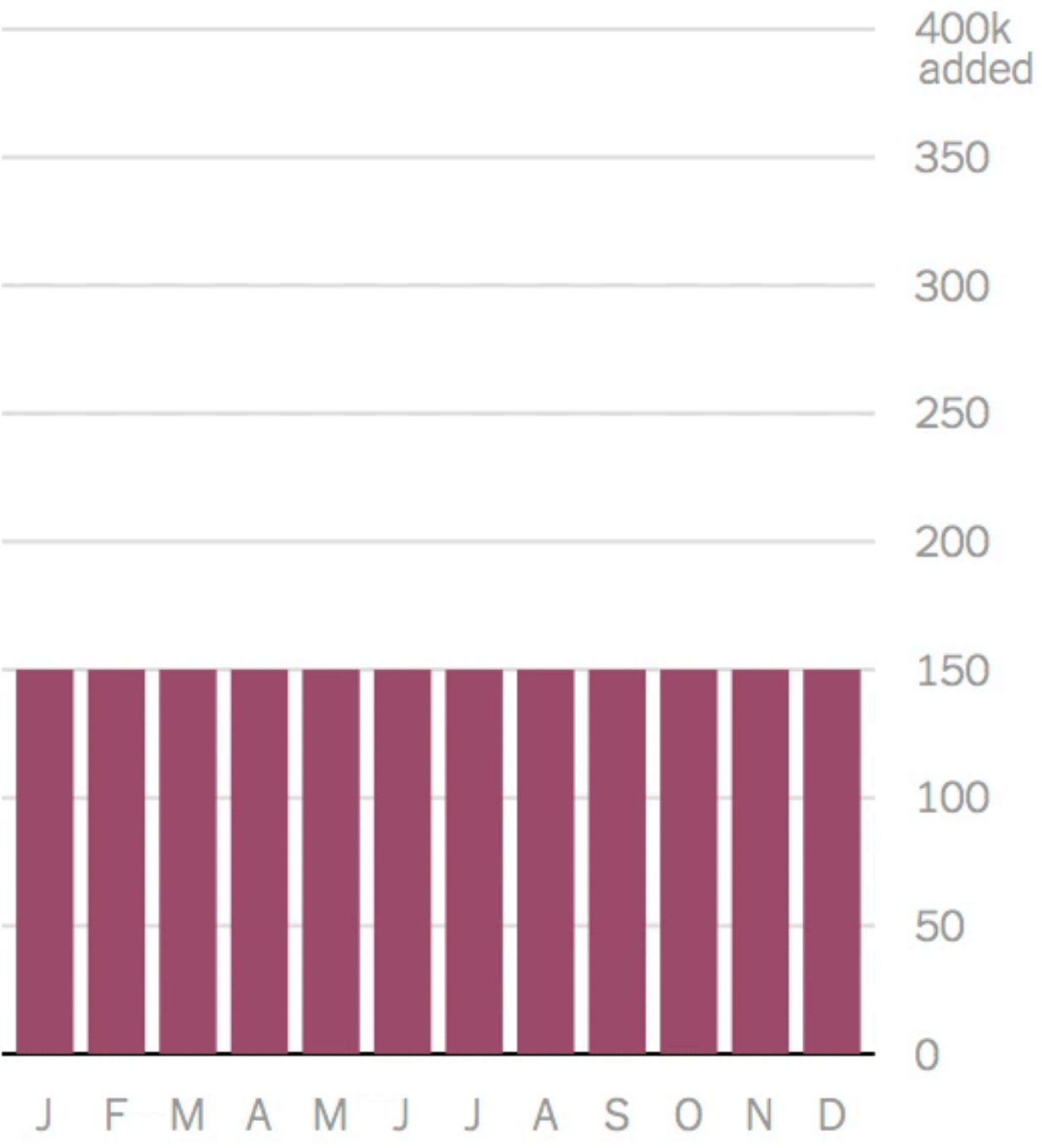
# How Not to Be Misled by the Jobs Report

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

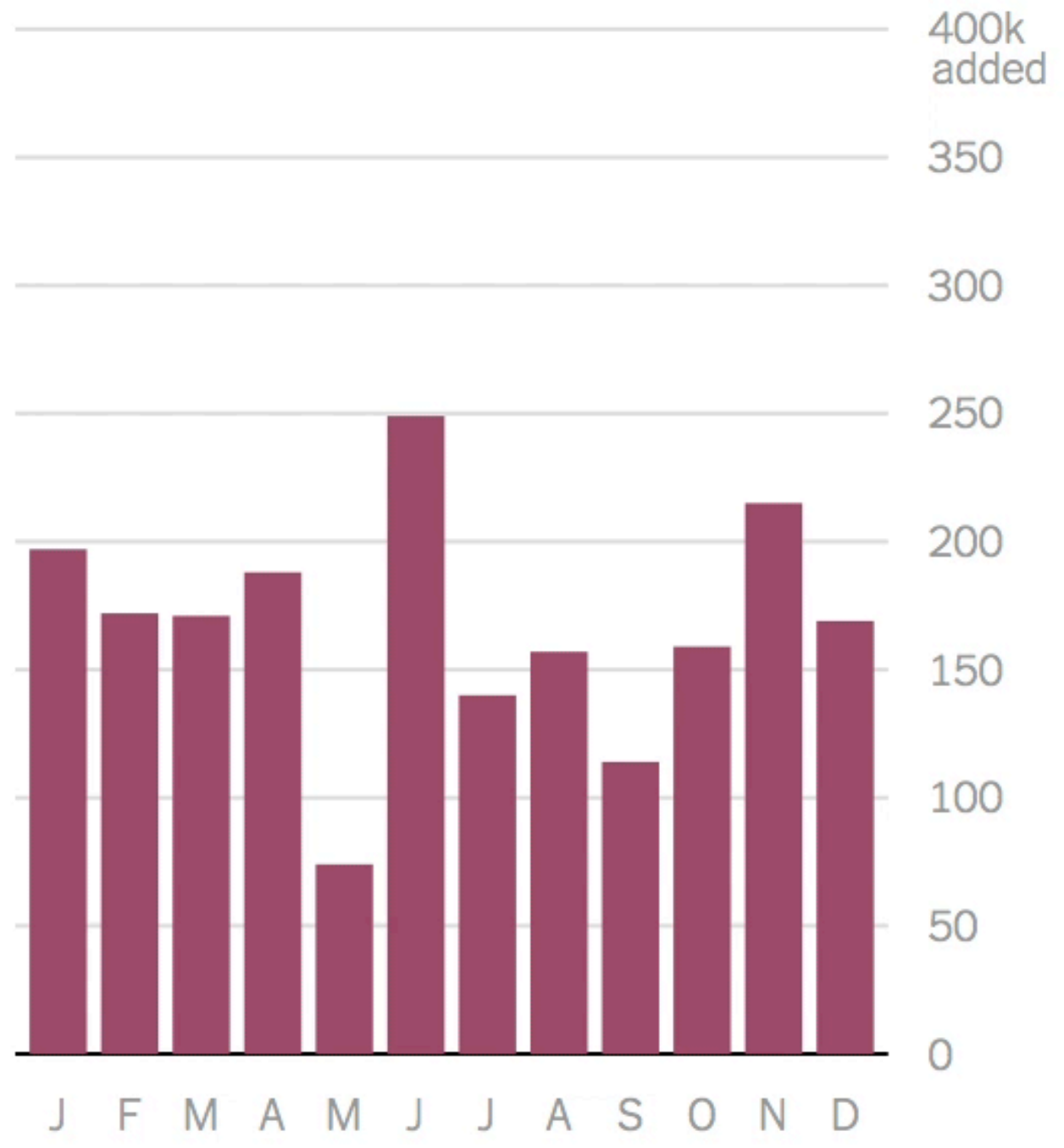
The jobs number is just an estimate, and it comes with uncertainty.



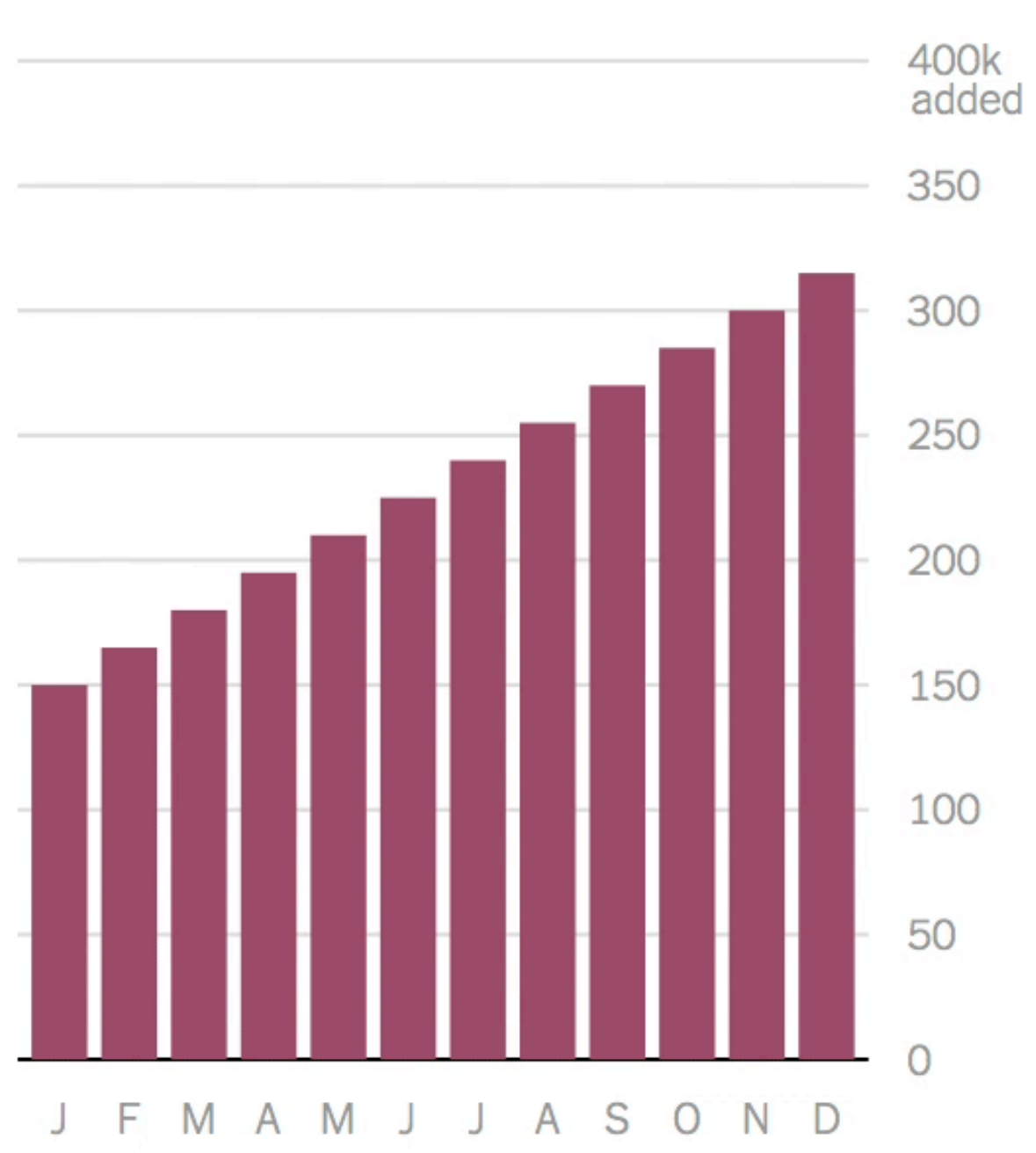
If job growth **were actually steady** over the last 12 months...



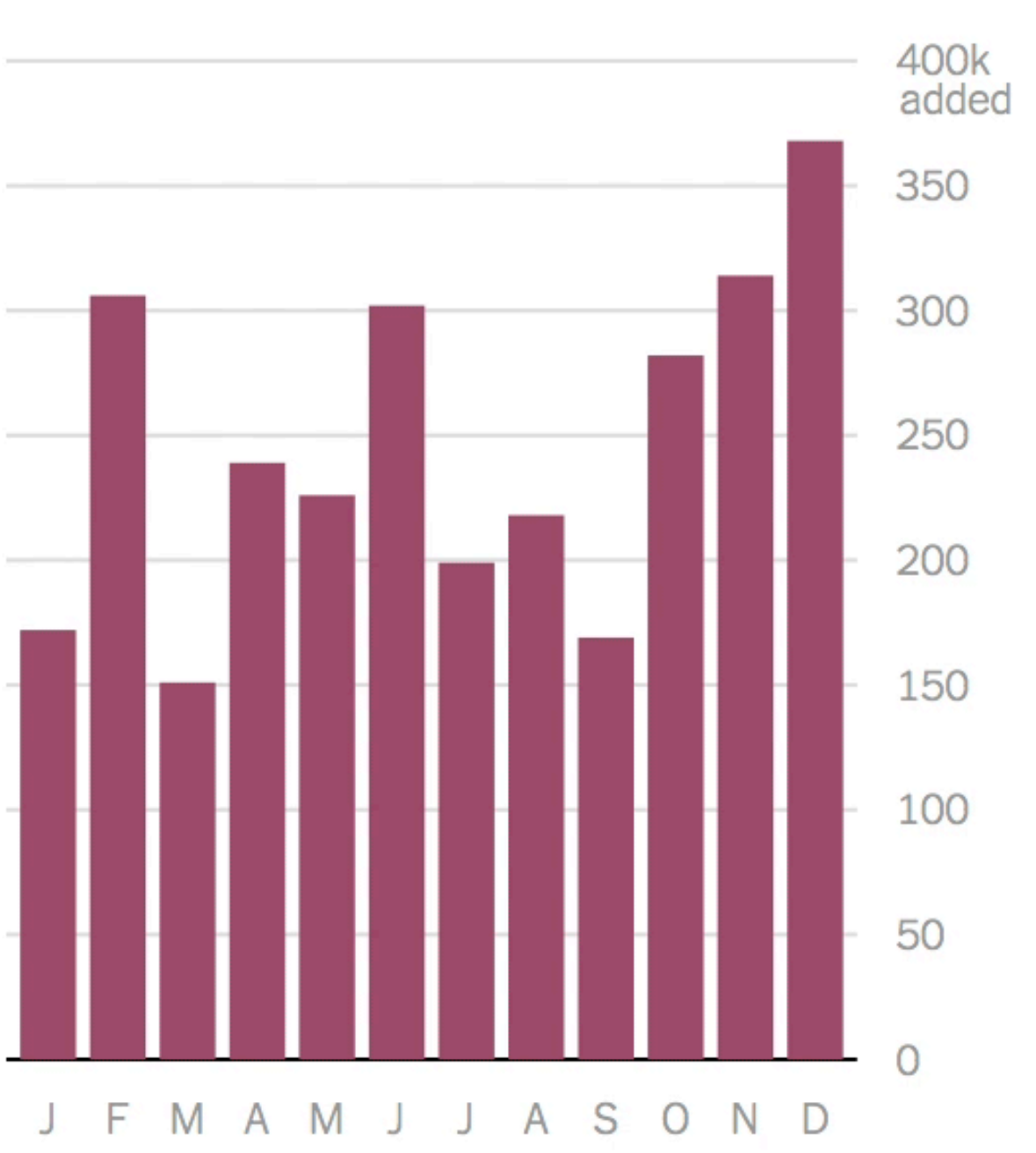
...the jobs report **could look like this:**



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



...the jobs report **could look like this:**



SPIN

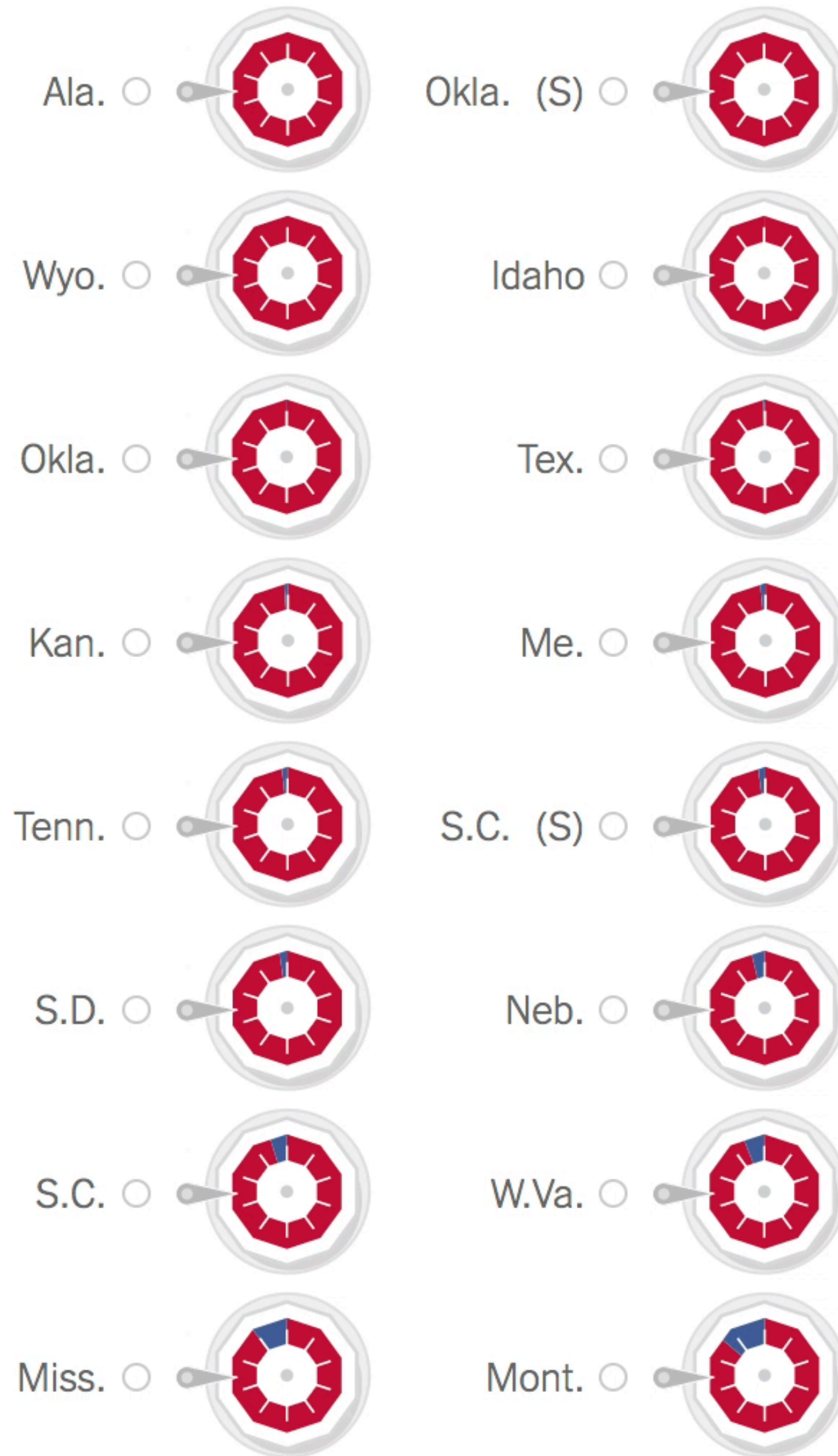
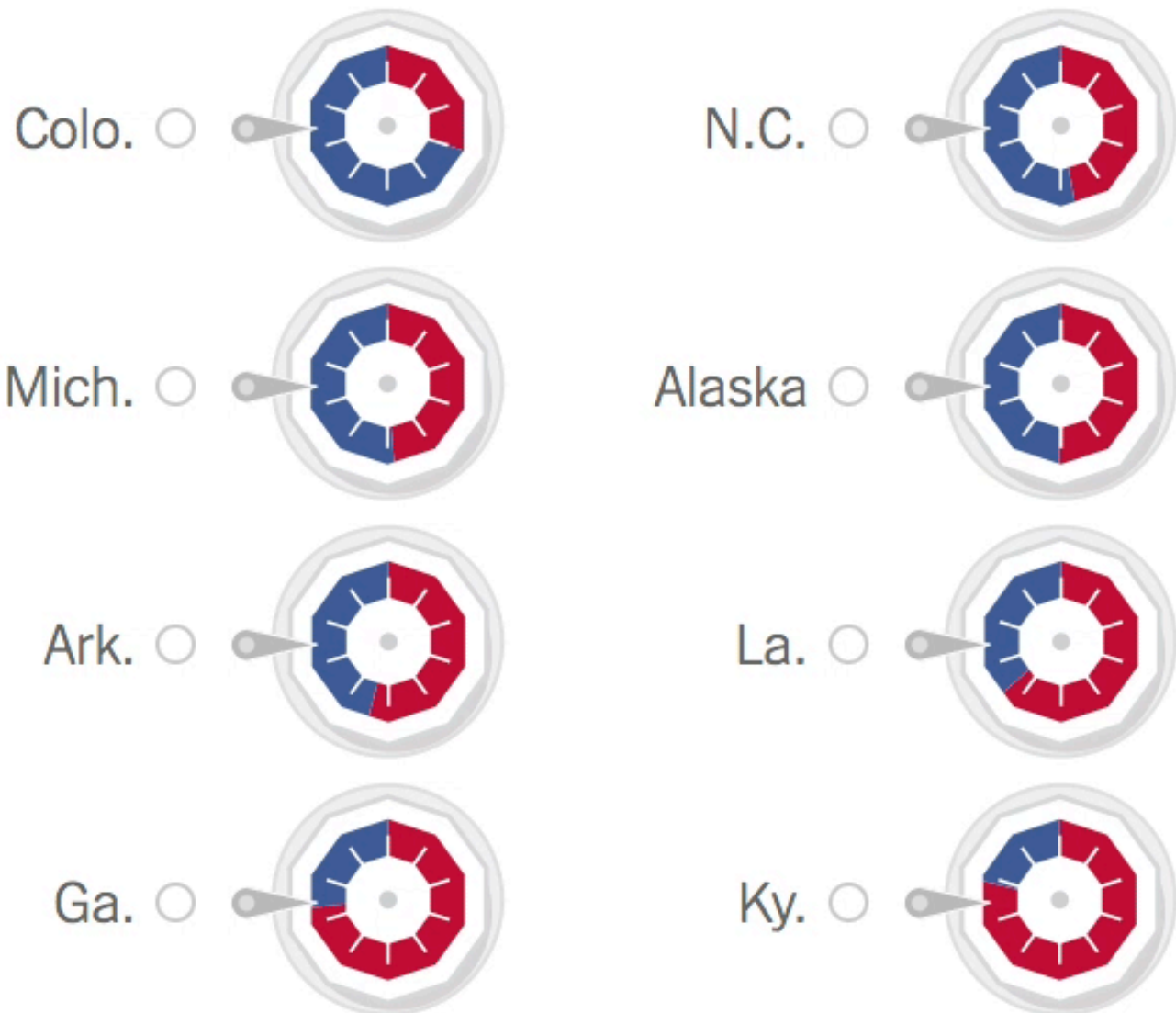
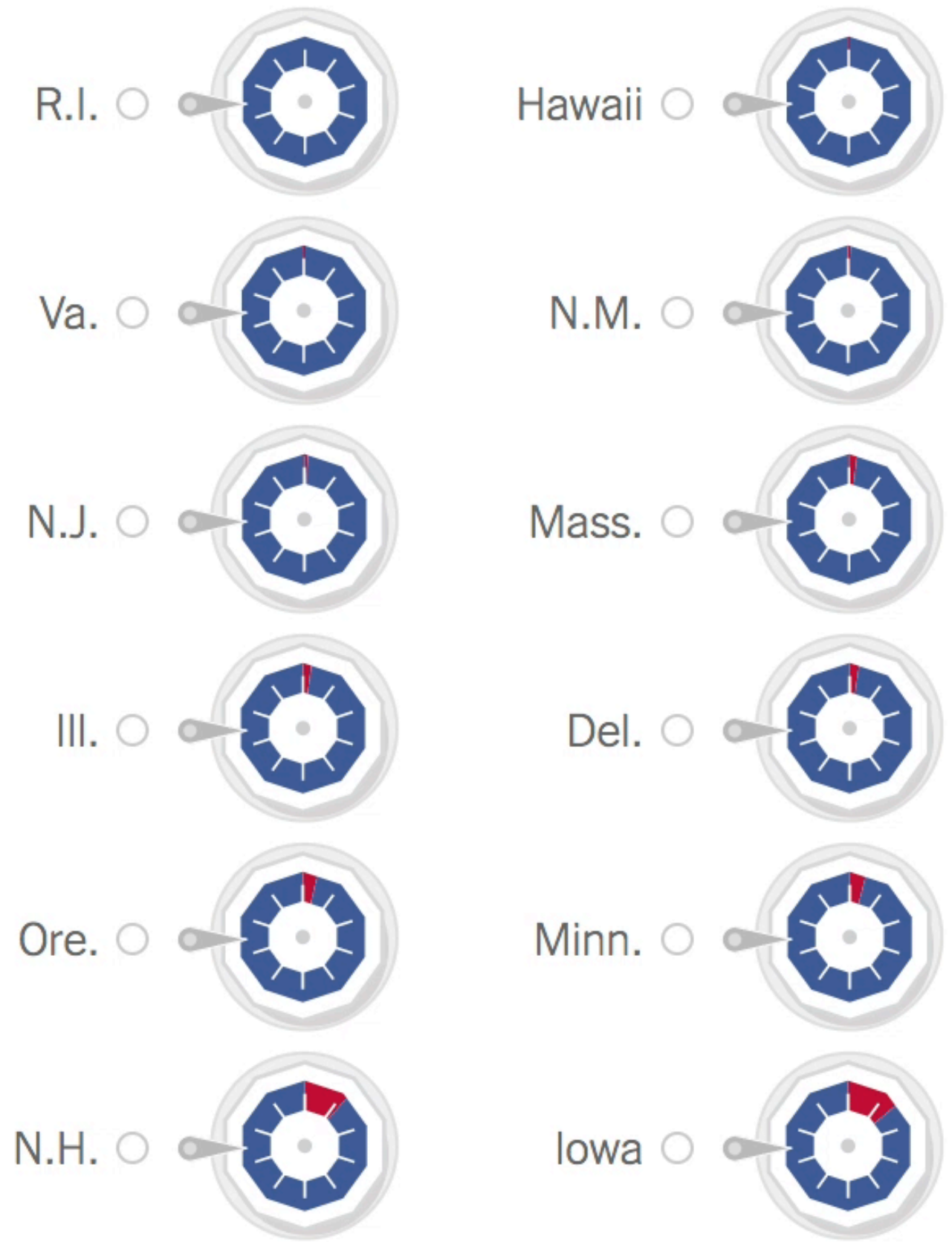
Democrats: ??

Republicans: ??

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?

How should I visualize it?

Building models is necessary to quantify uncertainty.

It is important to communicate the variability in model outcomes.

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!**