

Uncertainty

DSC 106: Data Visualization

Jared Wilber

UC San Diego

Announcements

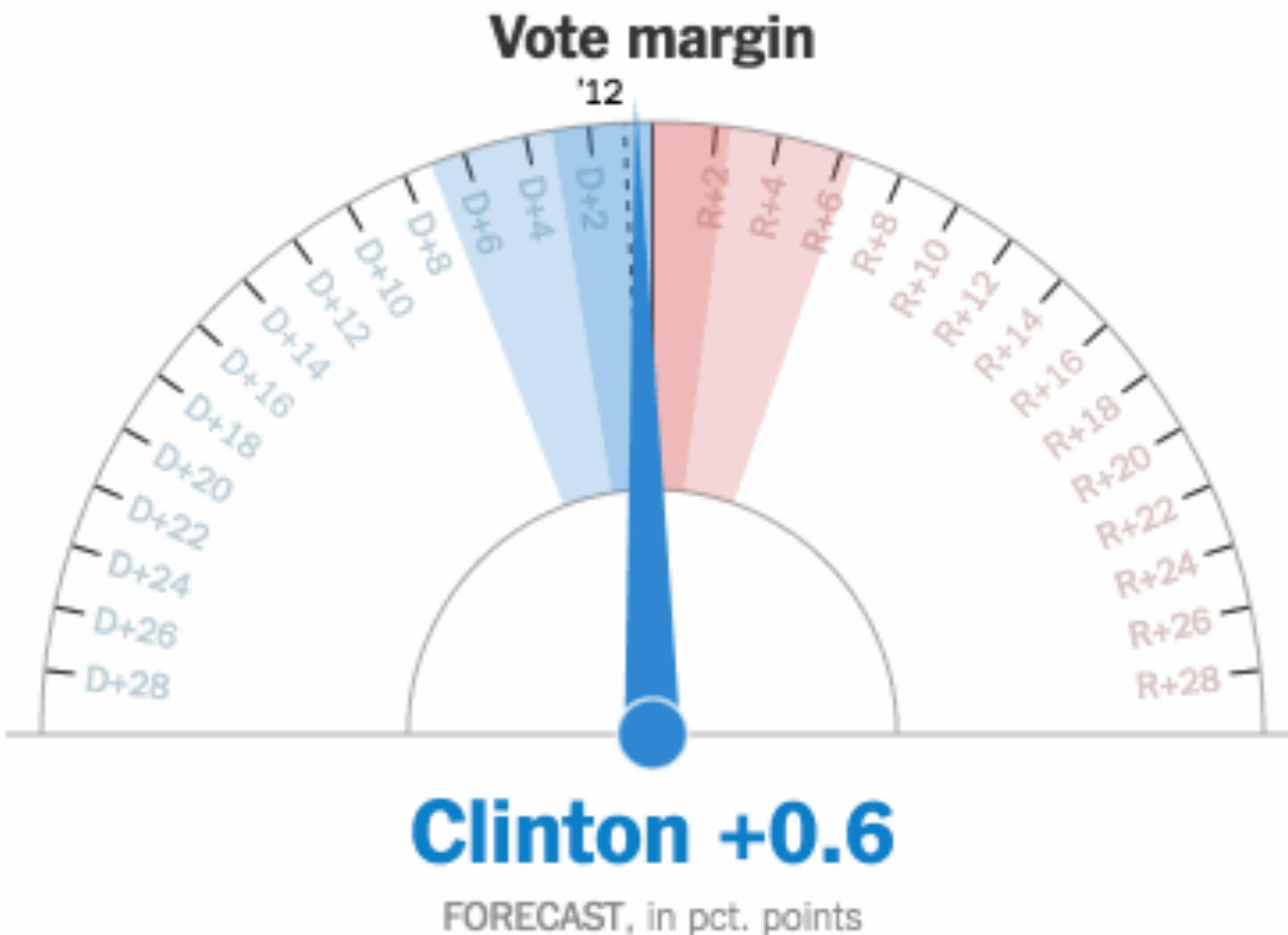
Final Project prototype due Monday.

Final Project video due the Monday thereafter.

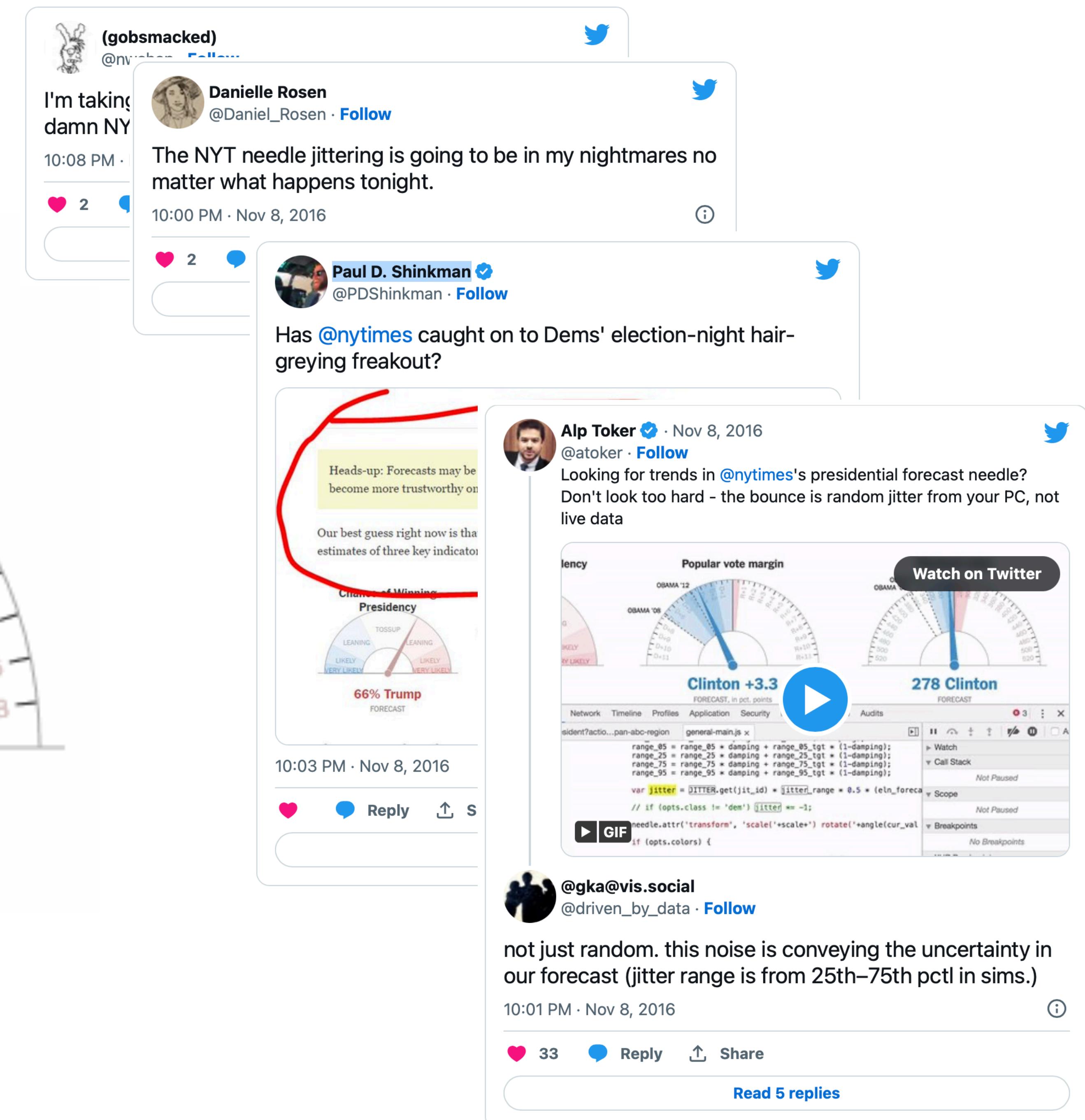
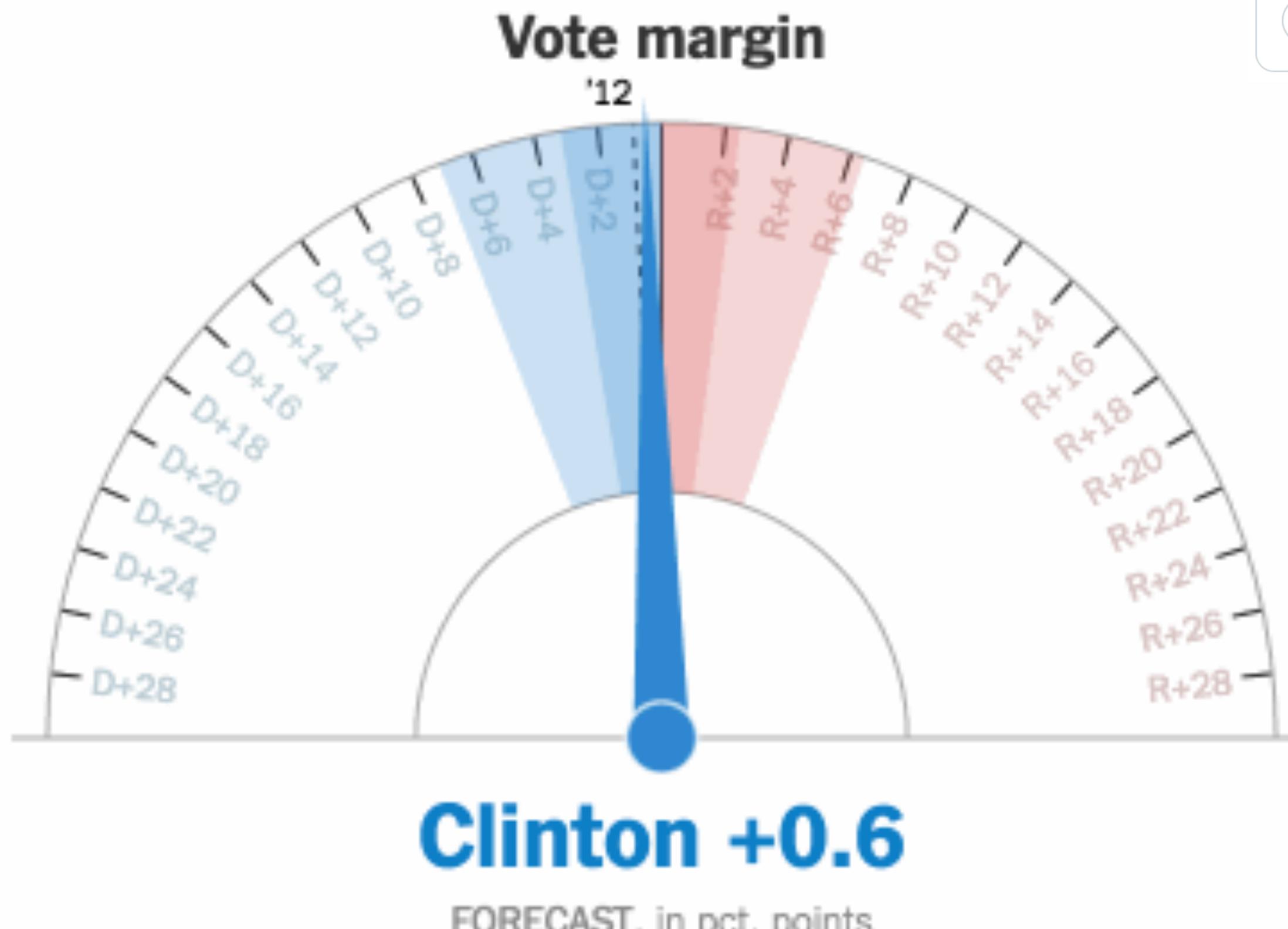
Lecture next Thurs. is Final Project peer feedback.

FAQs:

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



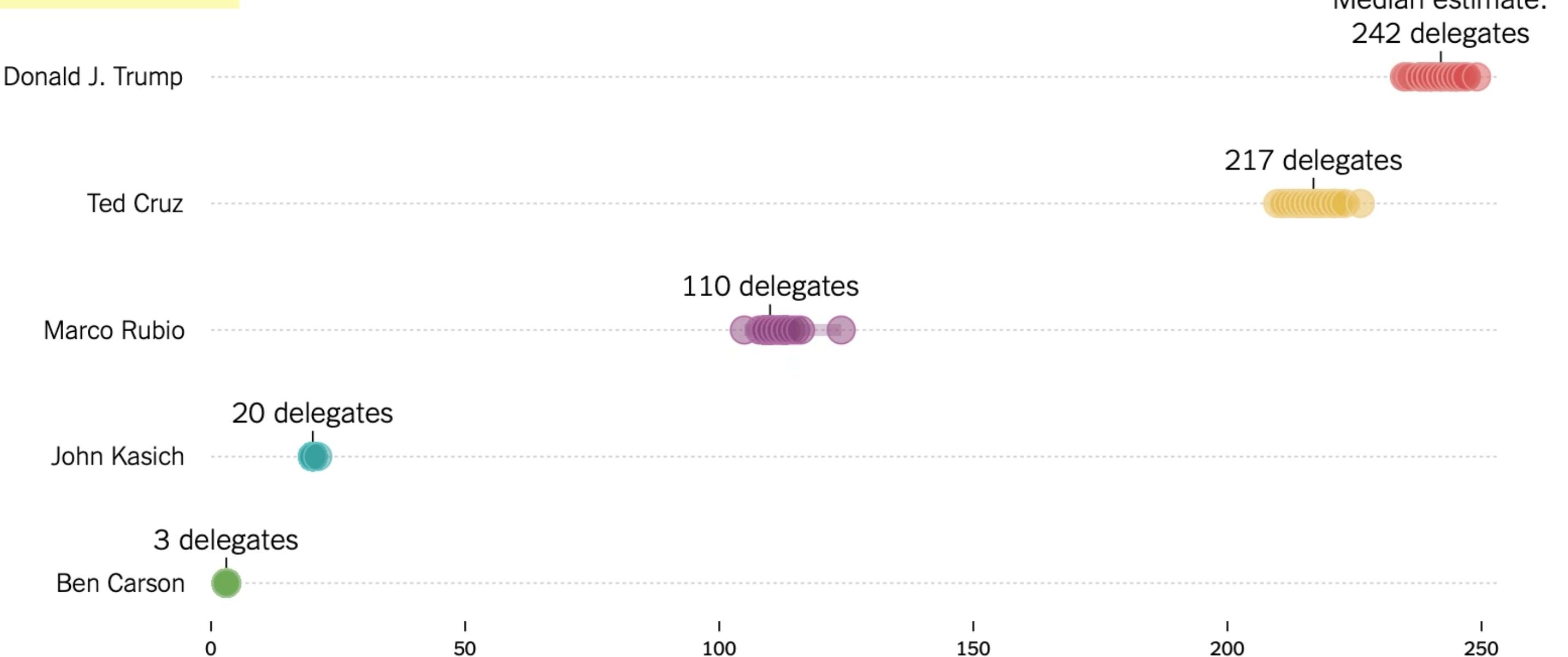
What is being visualized?
What are the strengths
and weaknesses of this
visualization?



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.

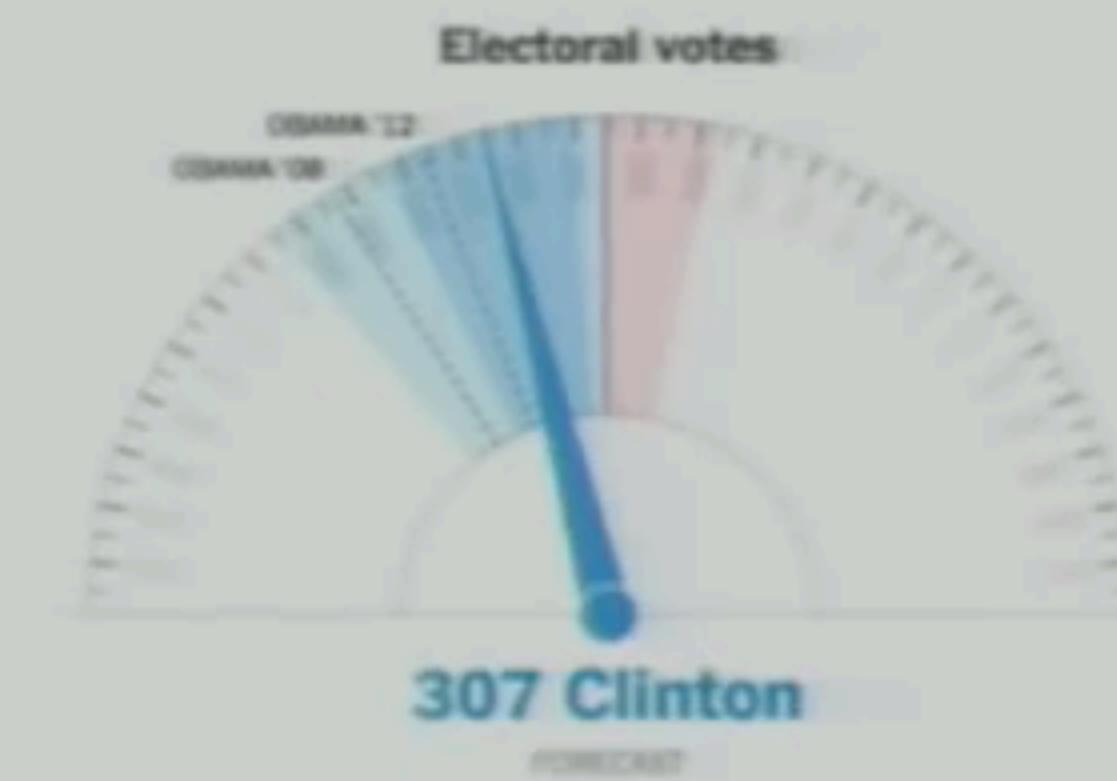
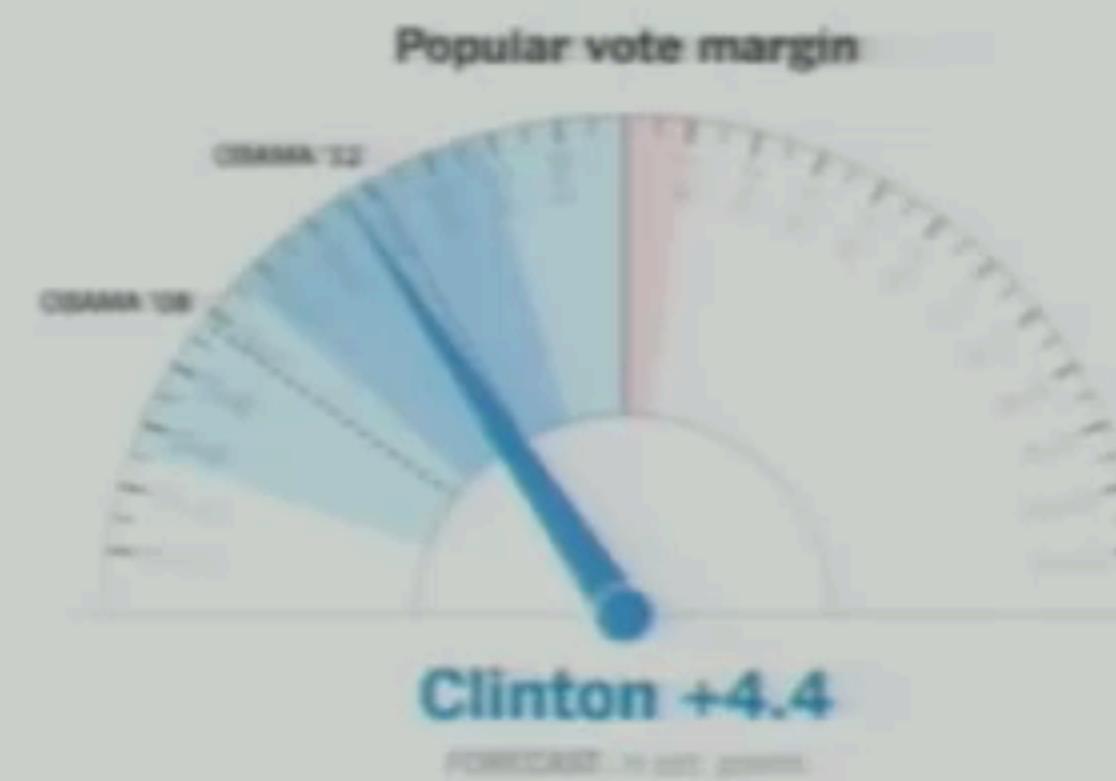
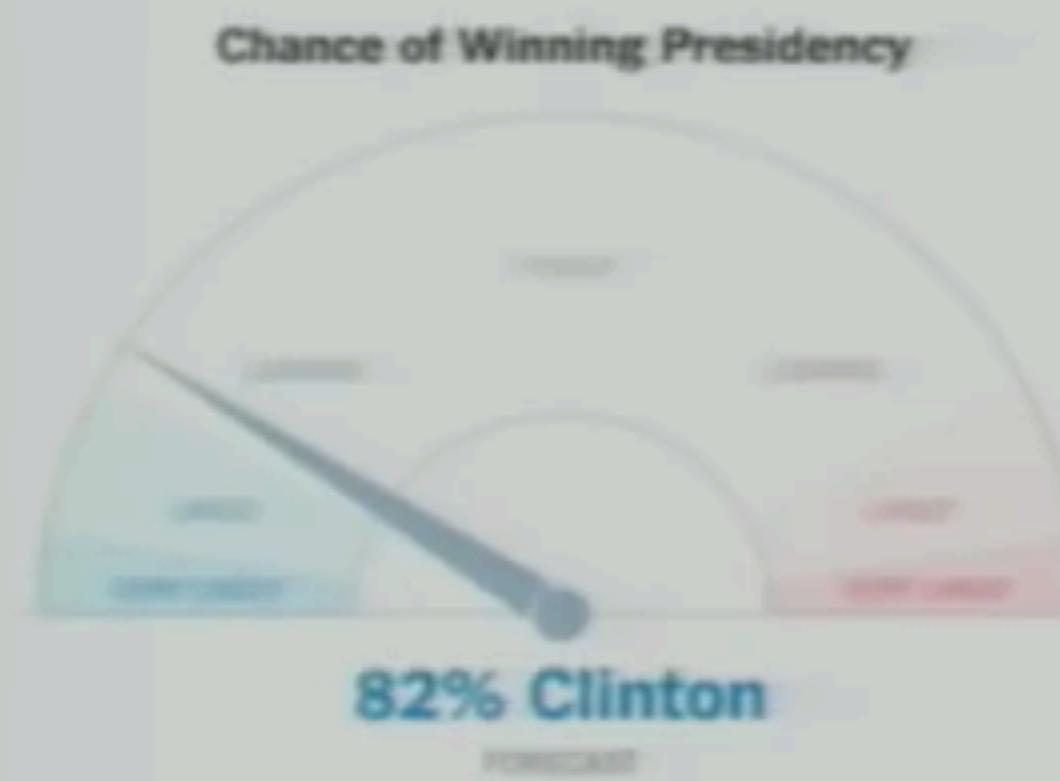
What is being visualized?

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. per cent of votes	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?

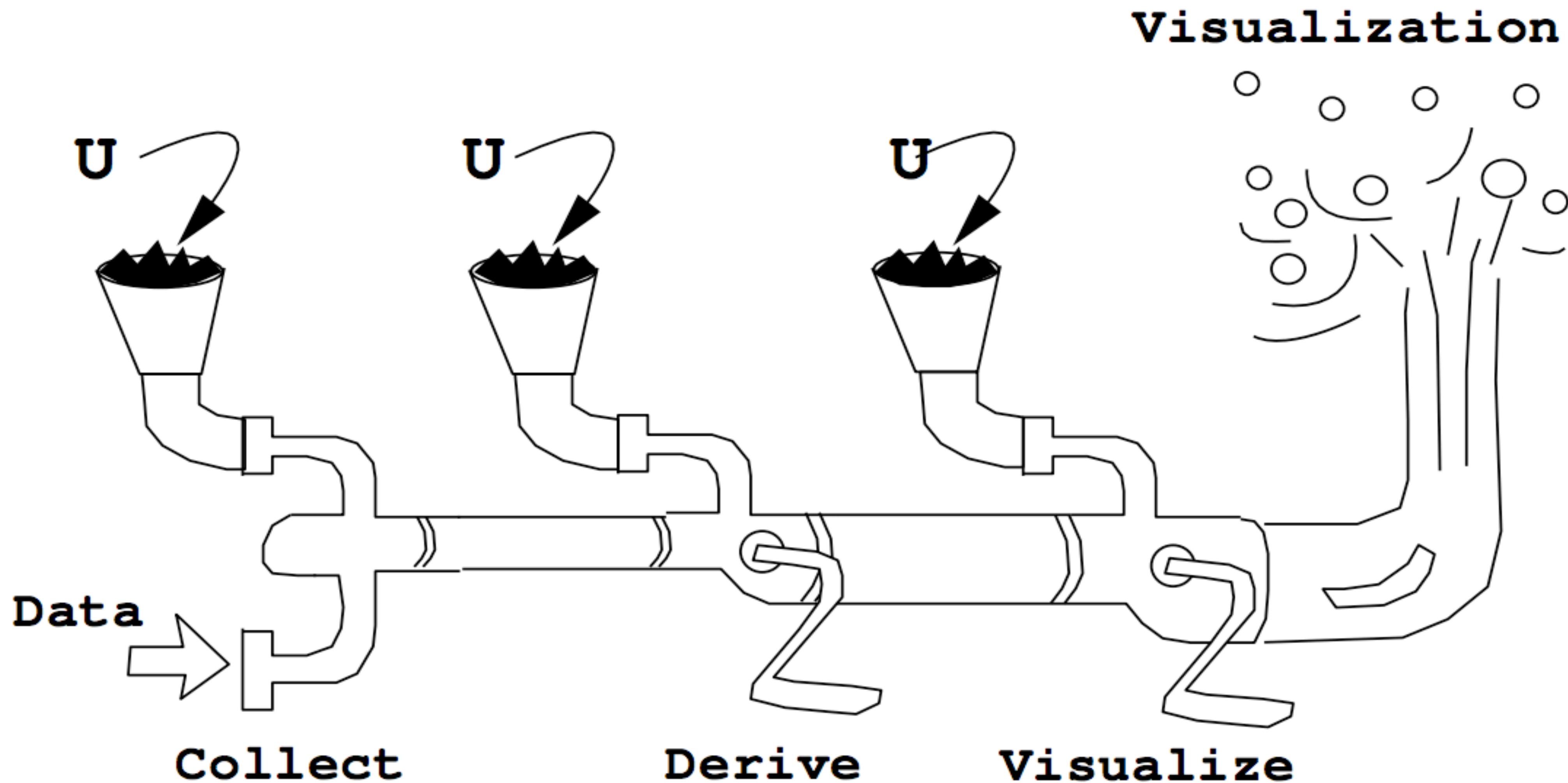
Uncertainty

What does it mean?

How should I visualize it?

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

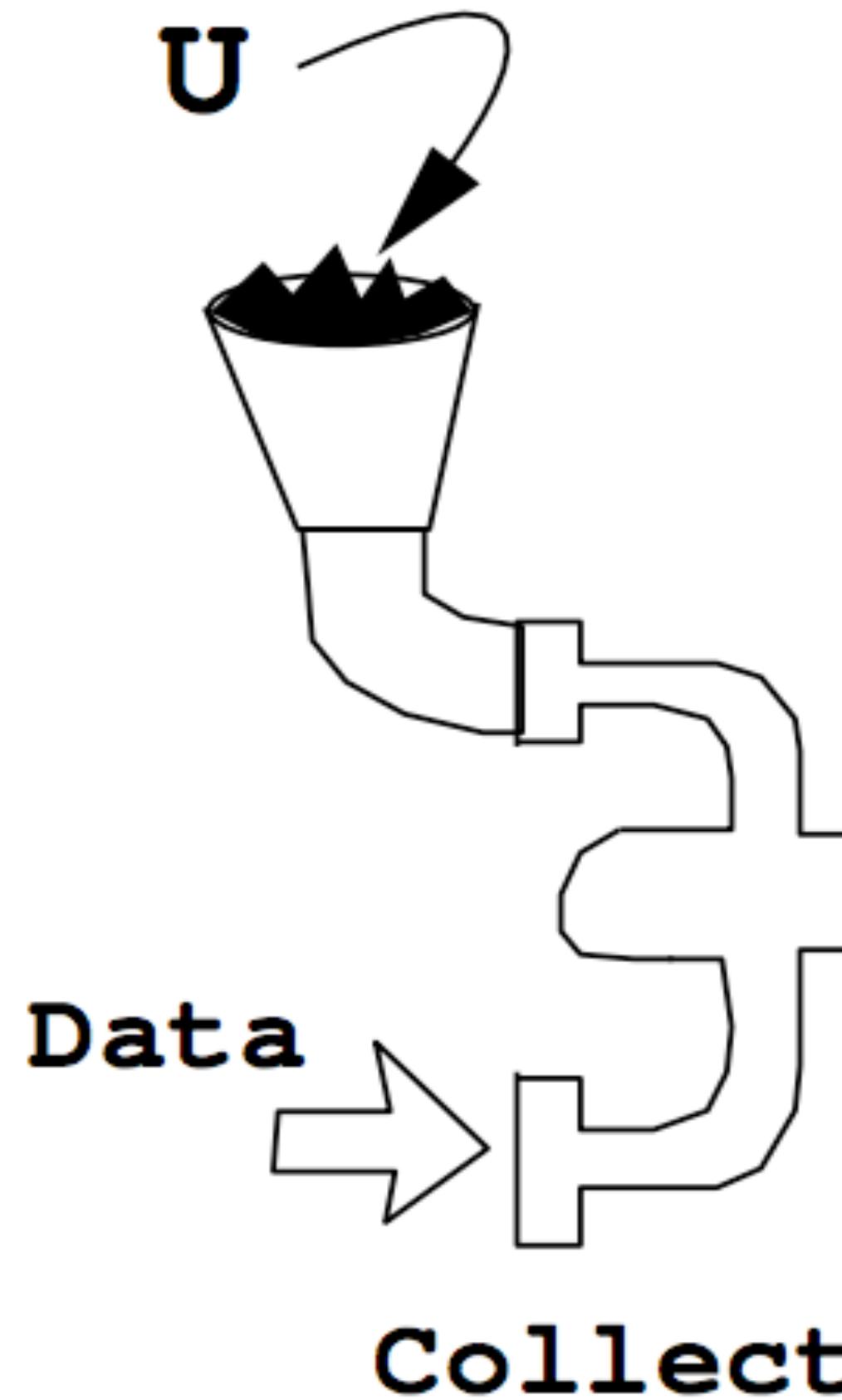
Sources and Types of Uncertainty



Sources and Types of Uncertainty

Measurement Uncertainty

How and how much should we sample the data?



Precision

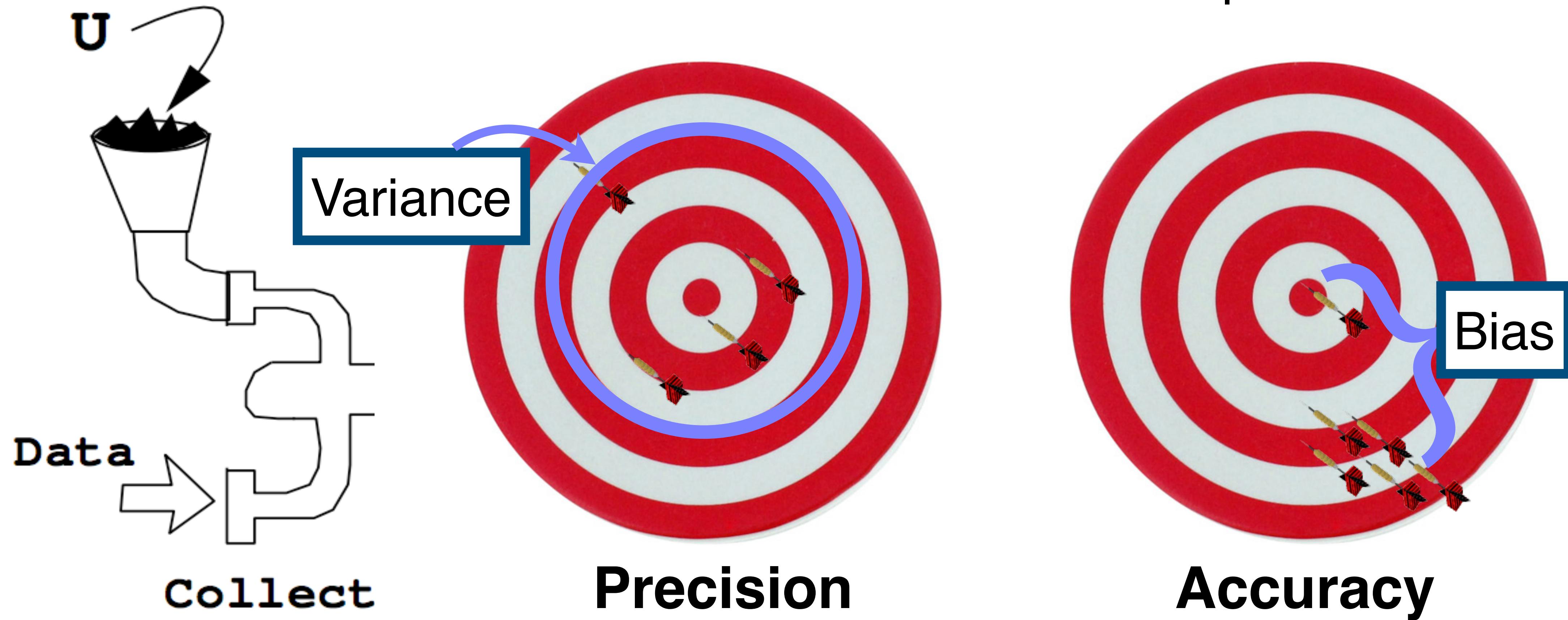


Accuracy

Sources and Types of Uncertainty

Measurement Uncertainty

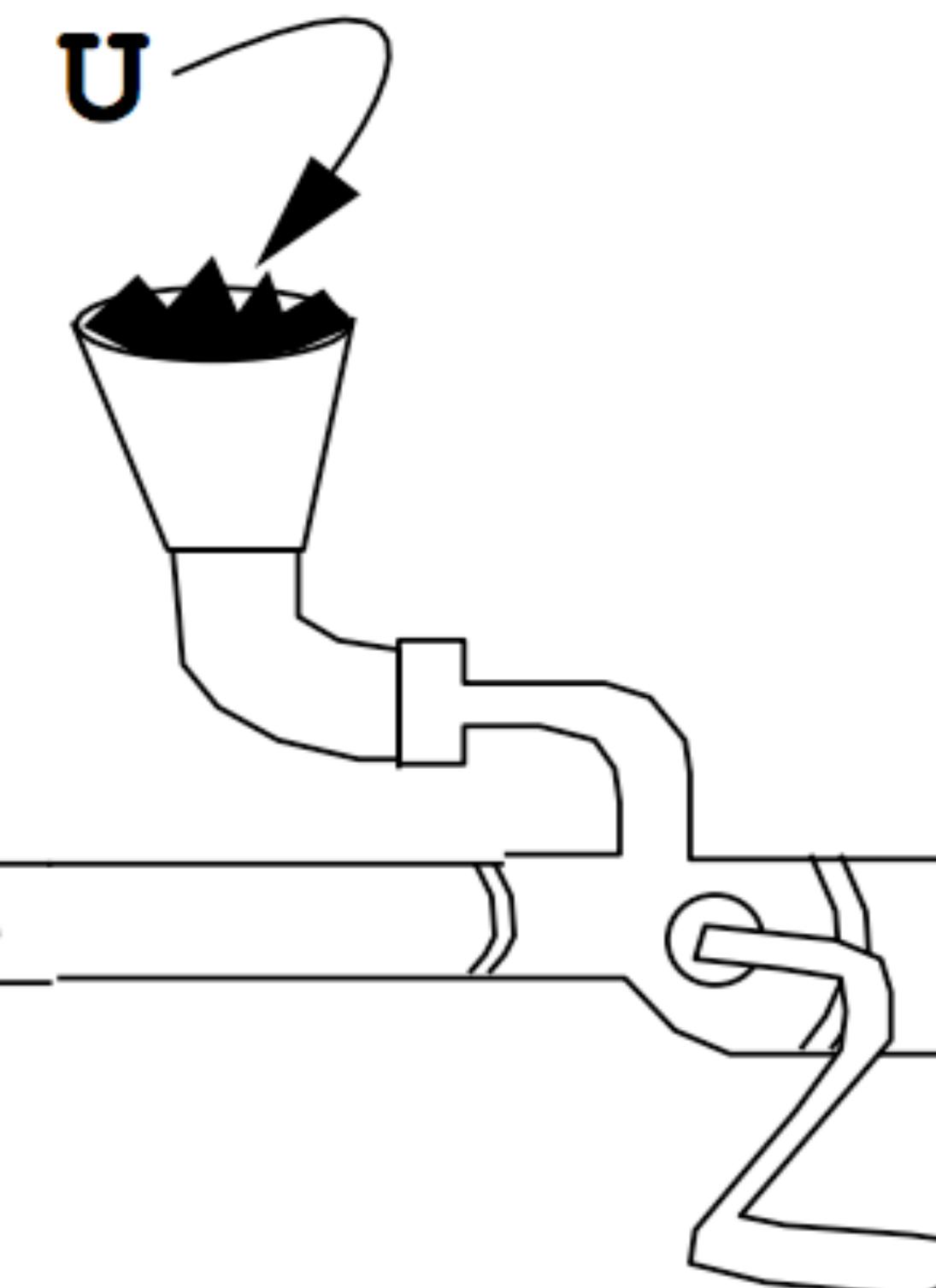
How and how much should we sample the data?



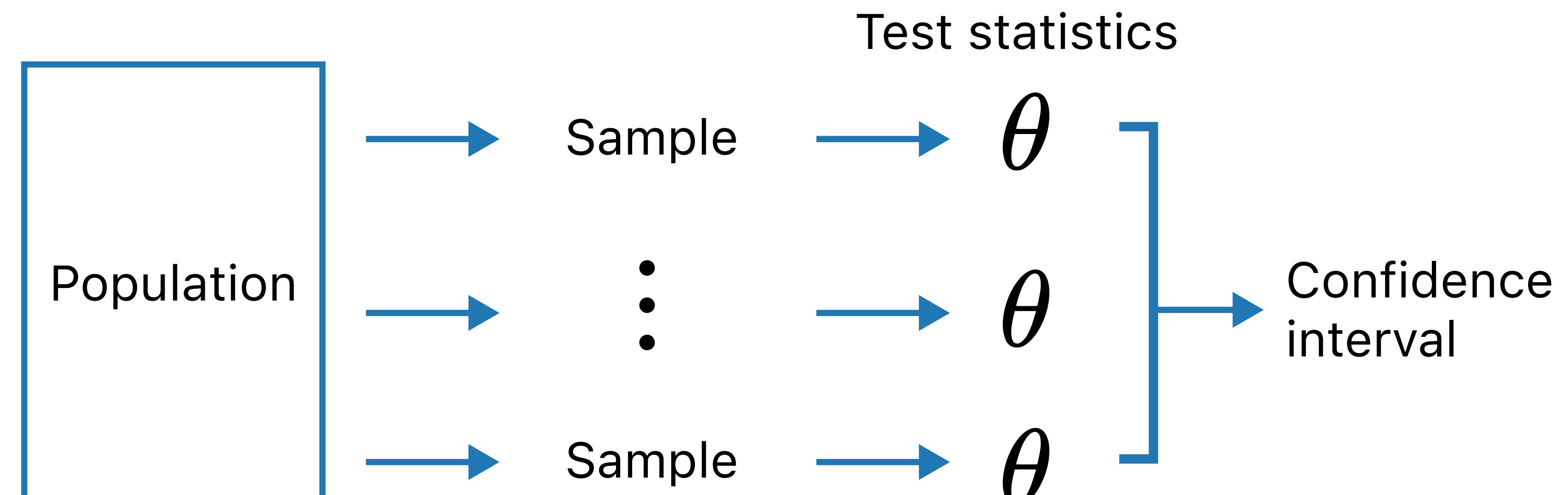
Sources and Types of Uncertainty

Model Uncertainty

How does the data fit together?

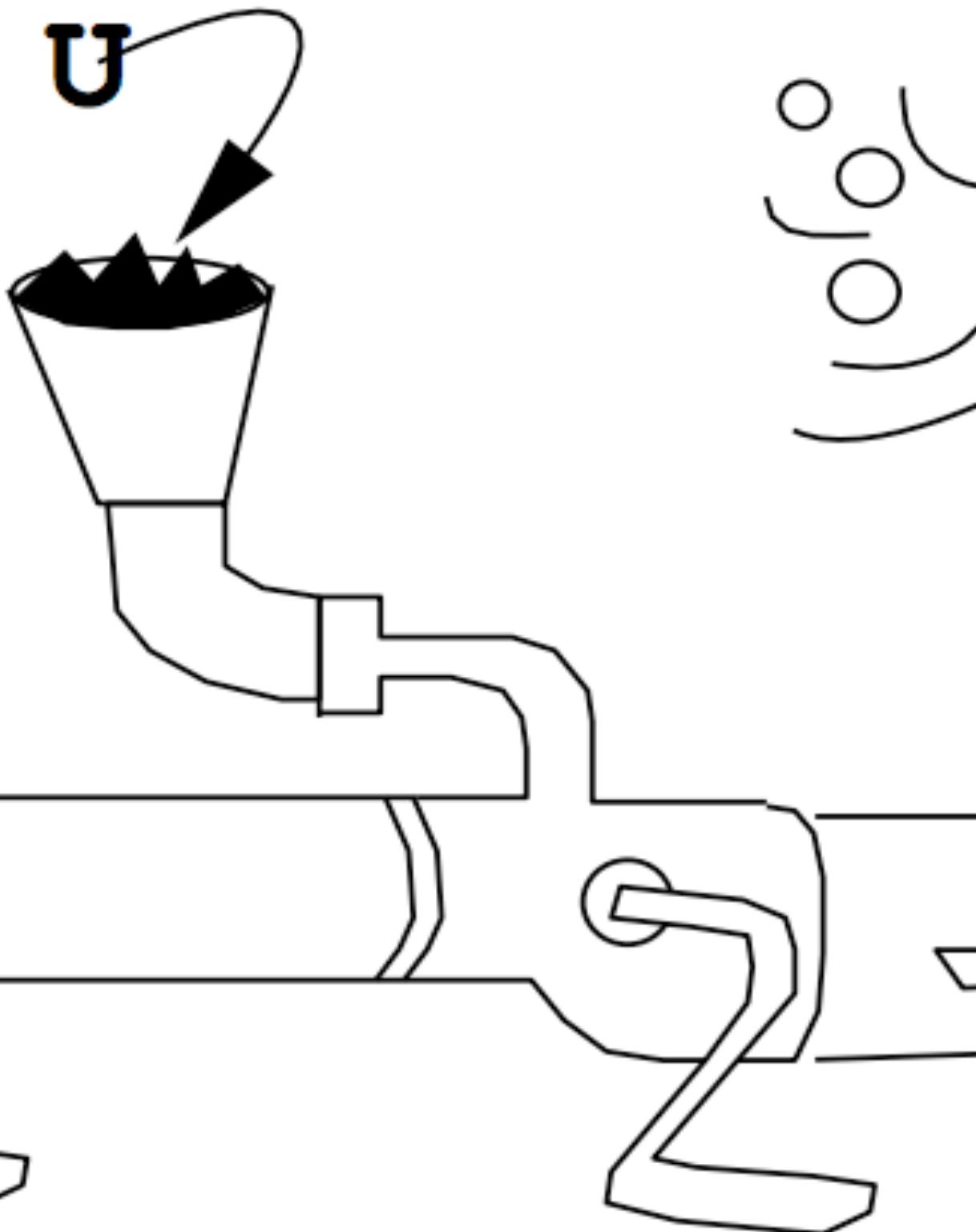


Derive



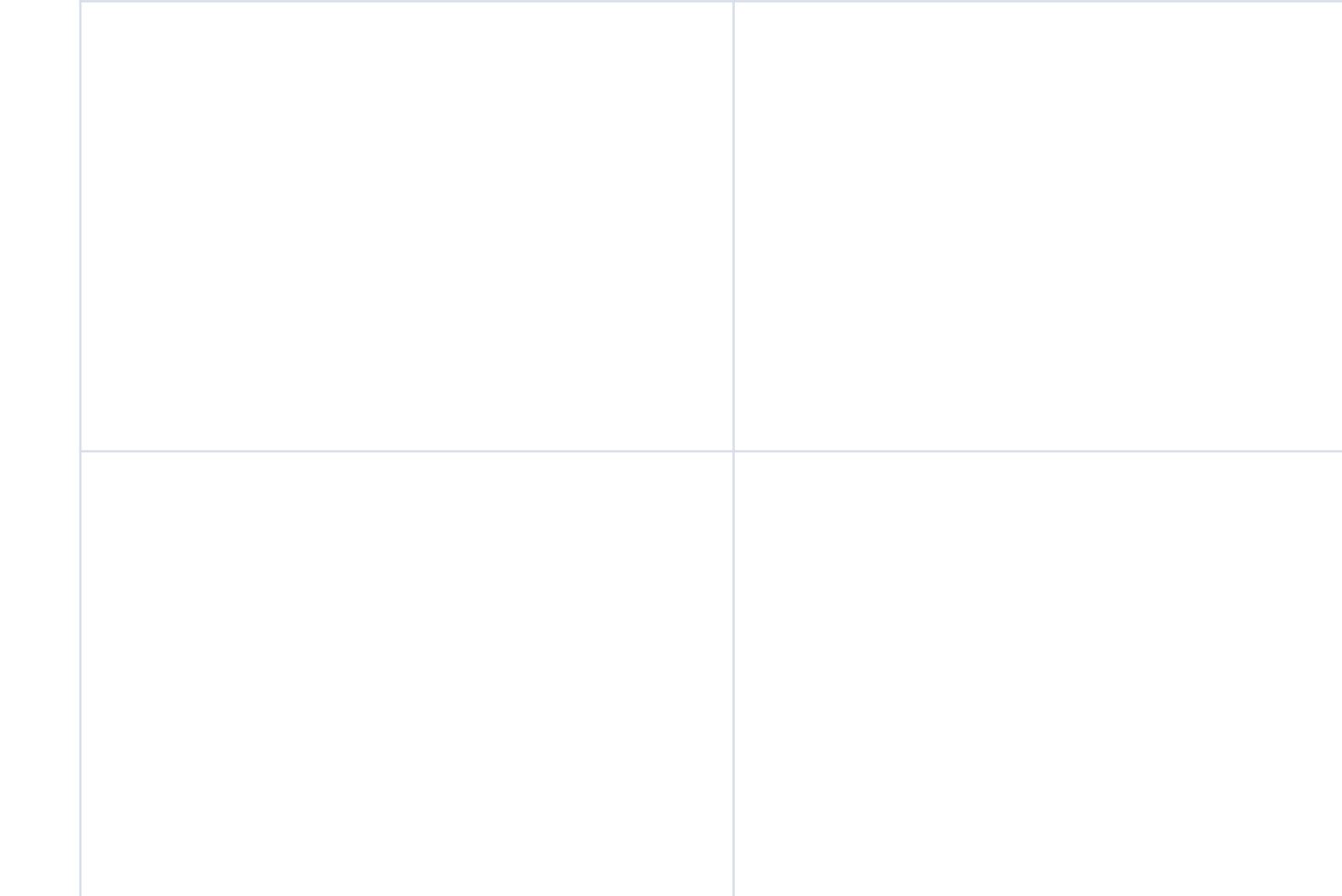
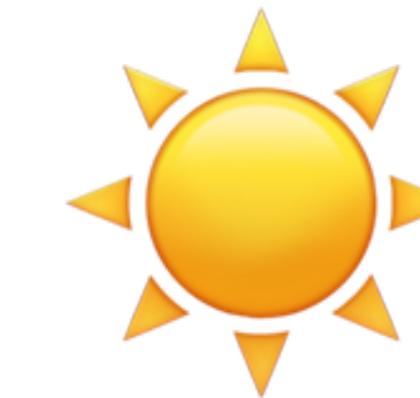
Sources and Types of Uncertainty

visualize



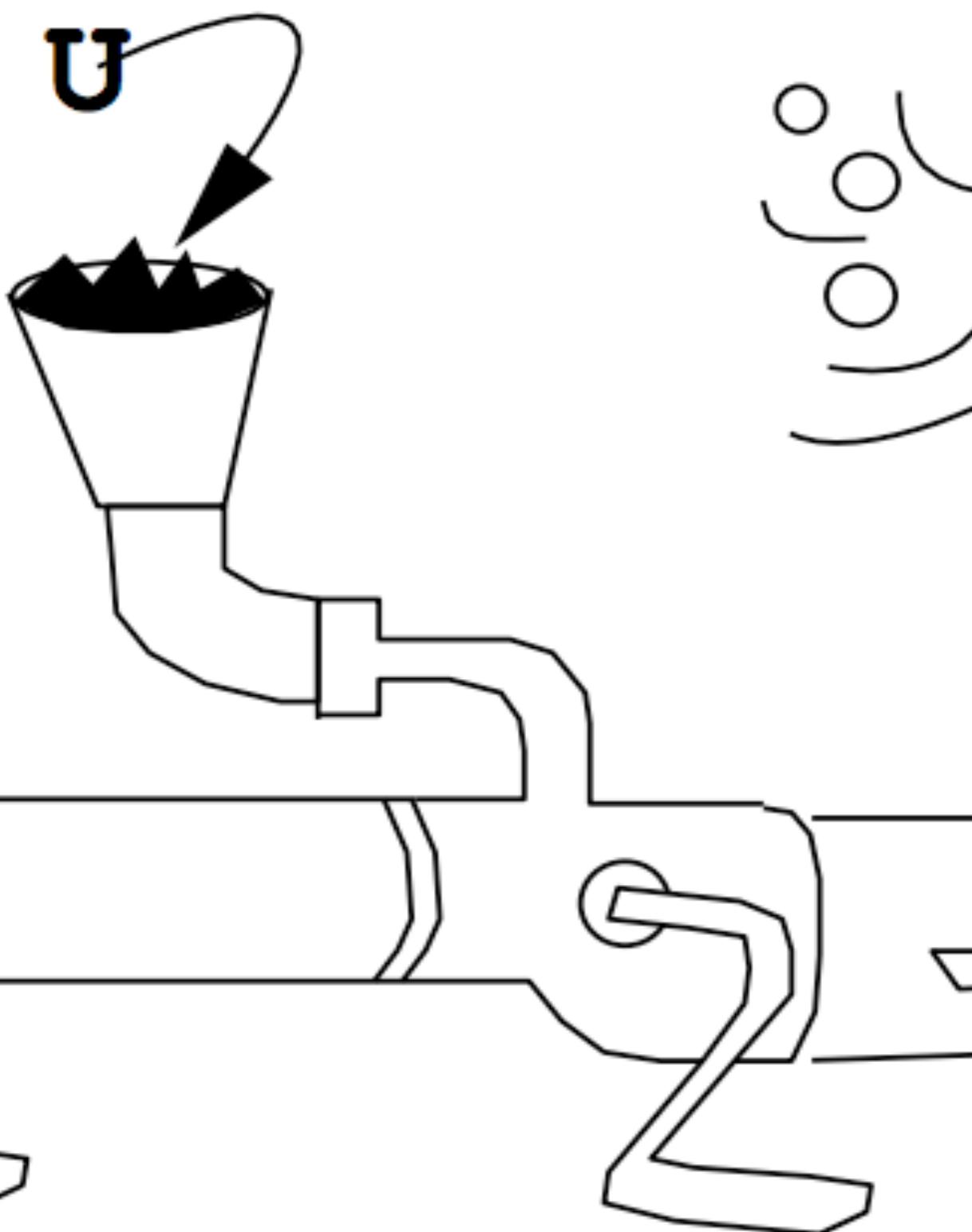
visualize

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



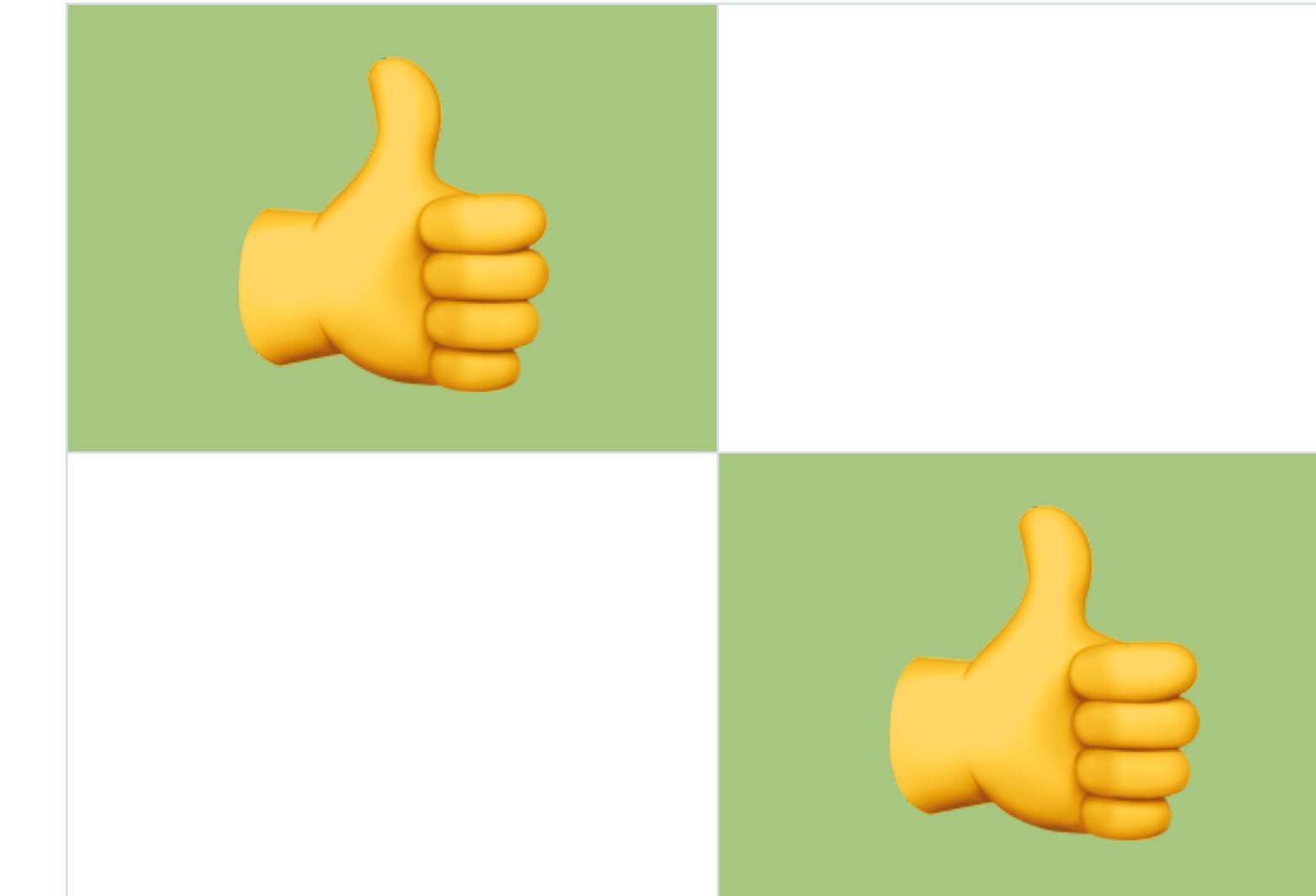
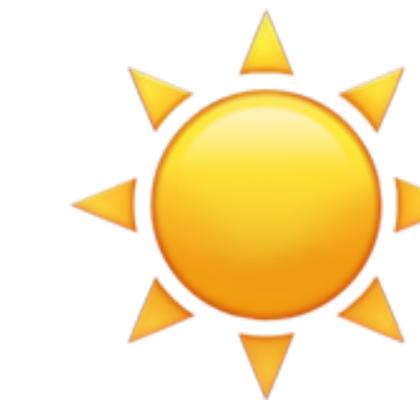
Sources and Types of Uncertainty

visualize



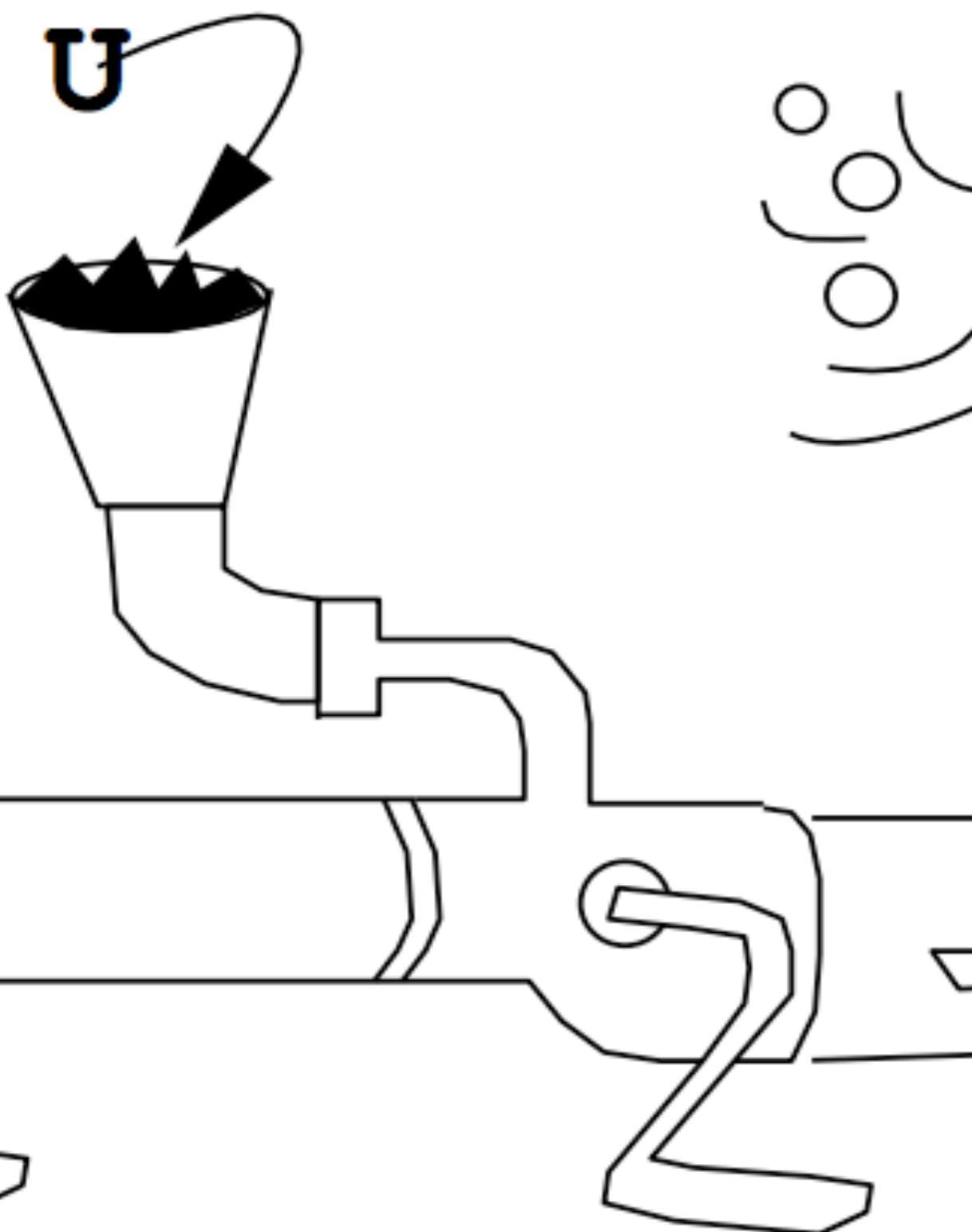
visualize

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



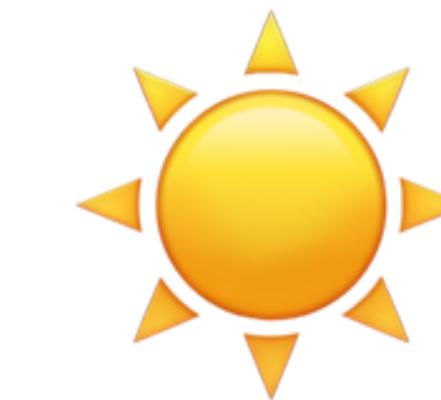
Sources and Types of Uncertainty

visualize



visualize

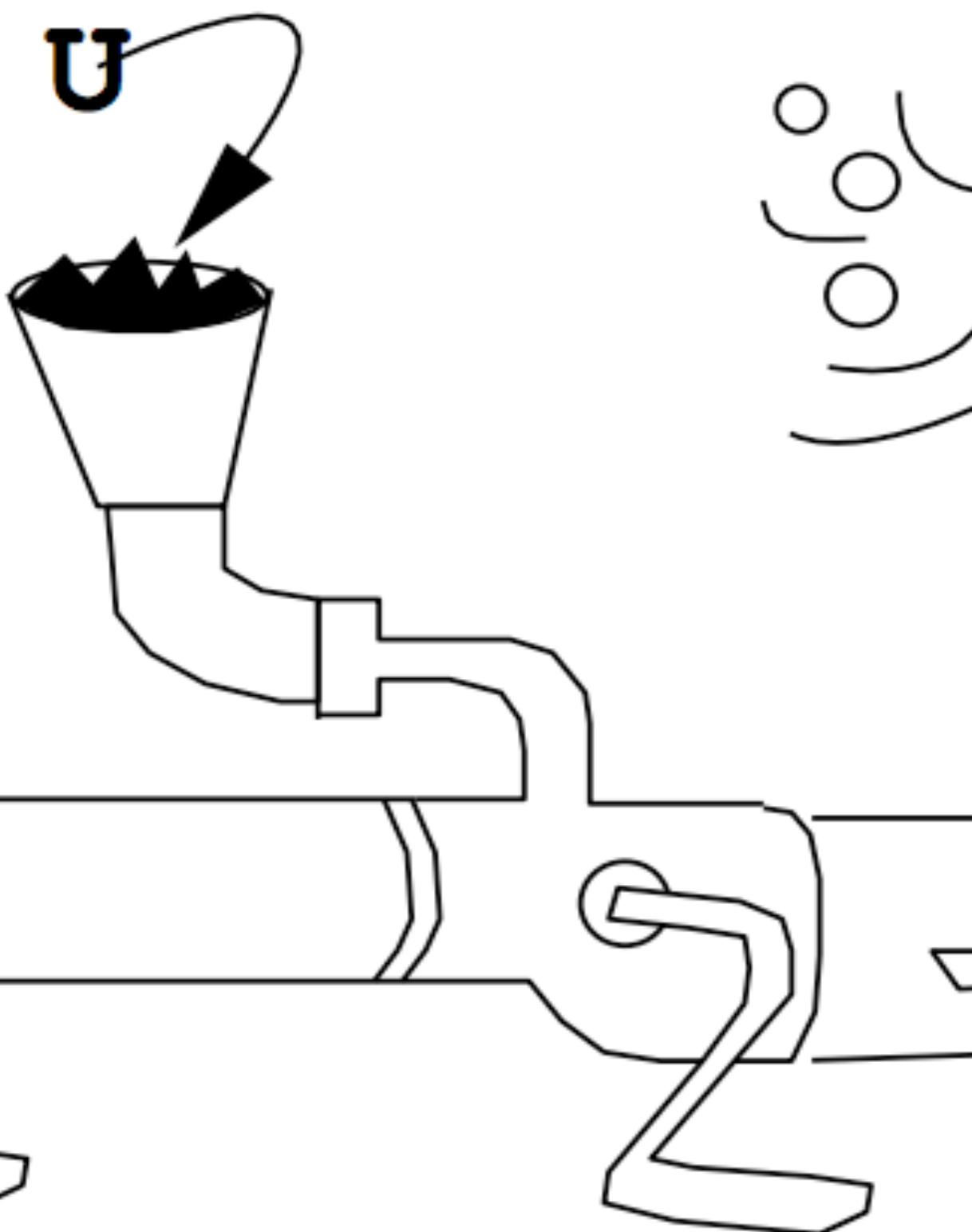
Decision/Forecast Uncertainty
How do I assess the risk or error?



	Type I False Positive

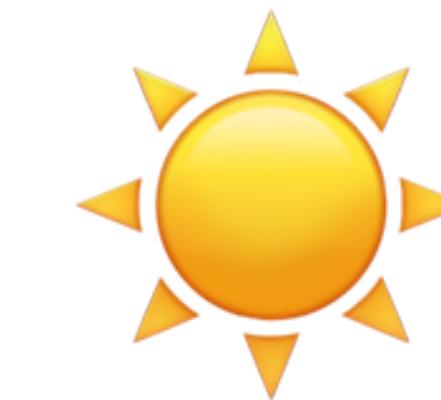
Sources and Types of Uncertainty

visualize



visualize

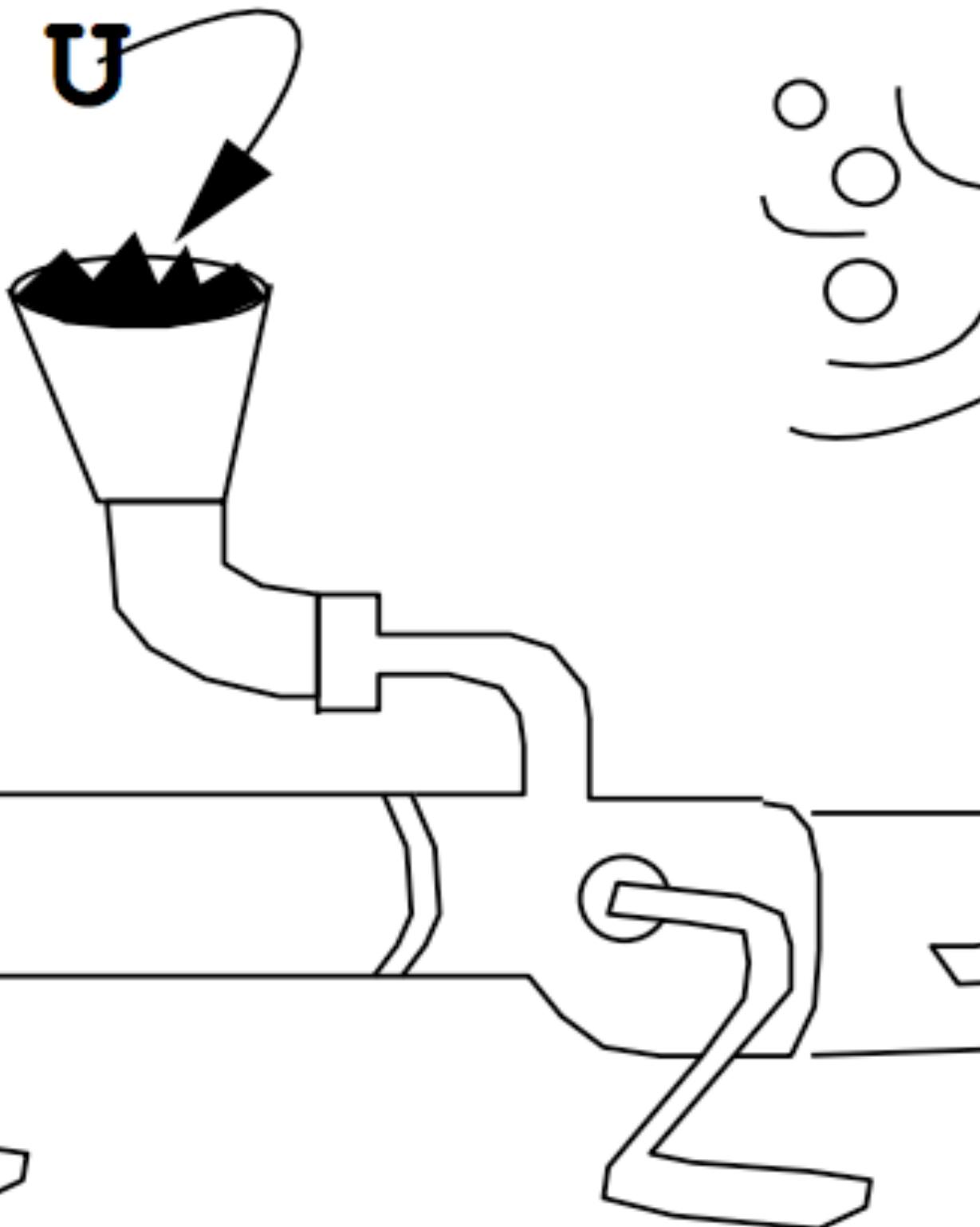
Decision/Forecast Uncertainty
How do I assess the risk or error?



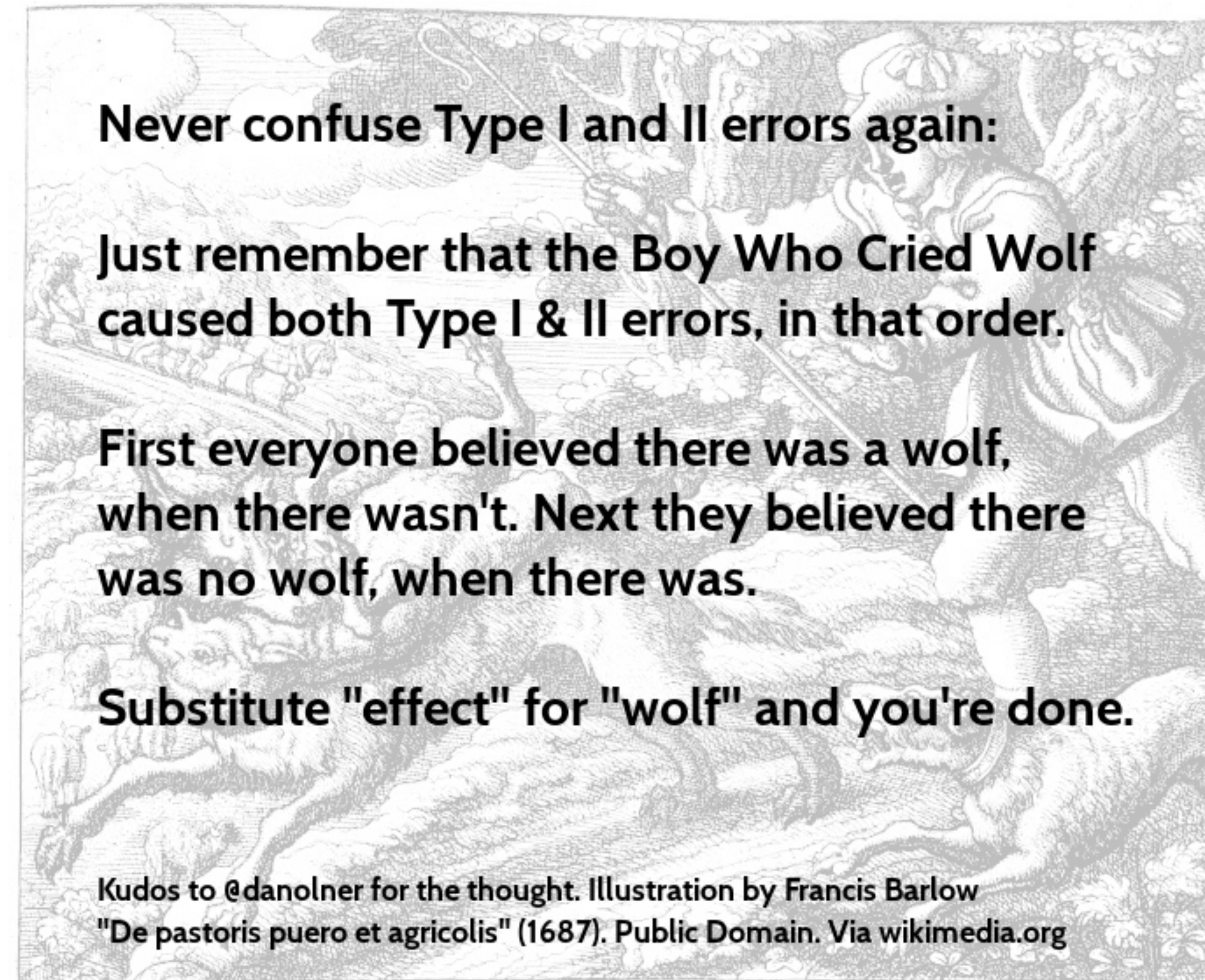
	Type I False Positive
Type II False Negative	

Sources and Types of Uncertainty

Visualize



Visualize



Uncertainty

What does it mean?

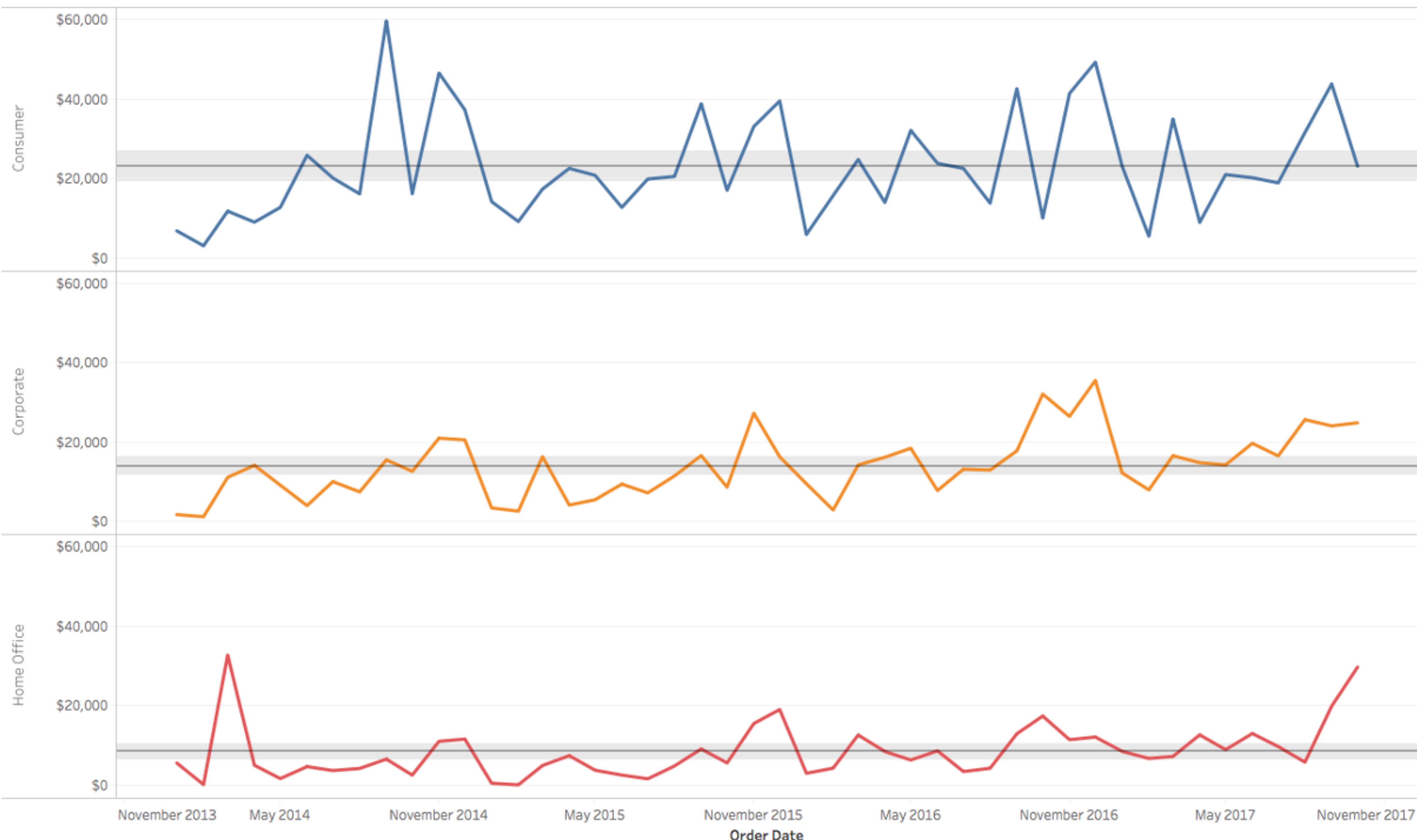
Lots of things!

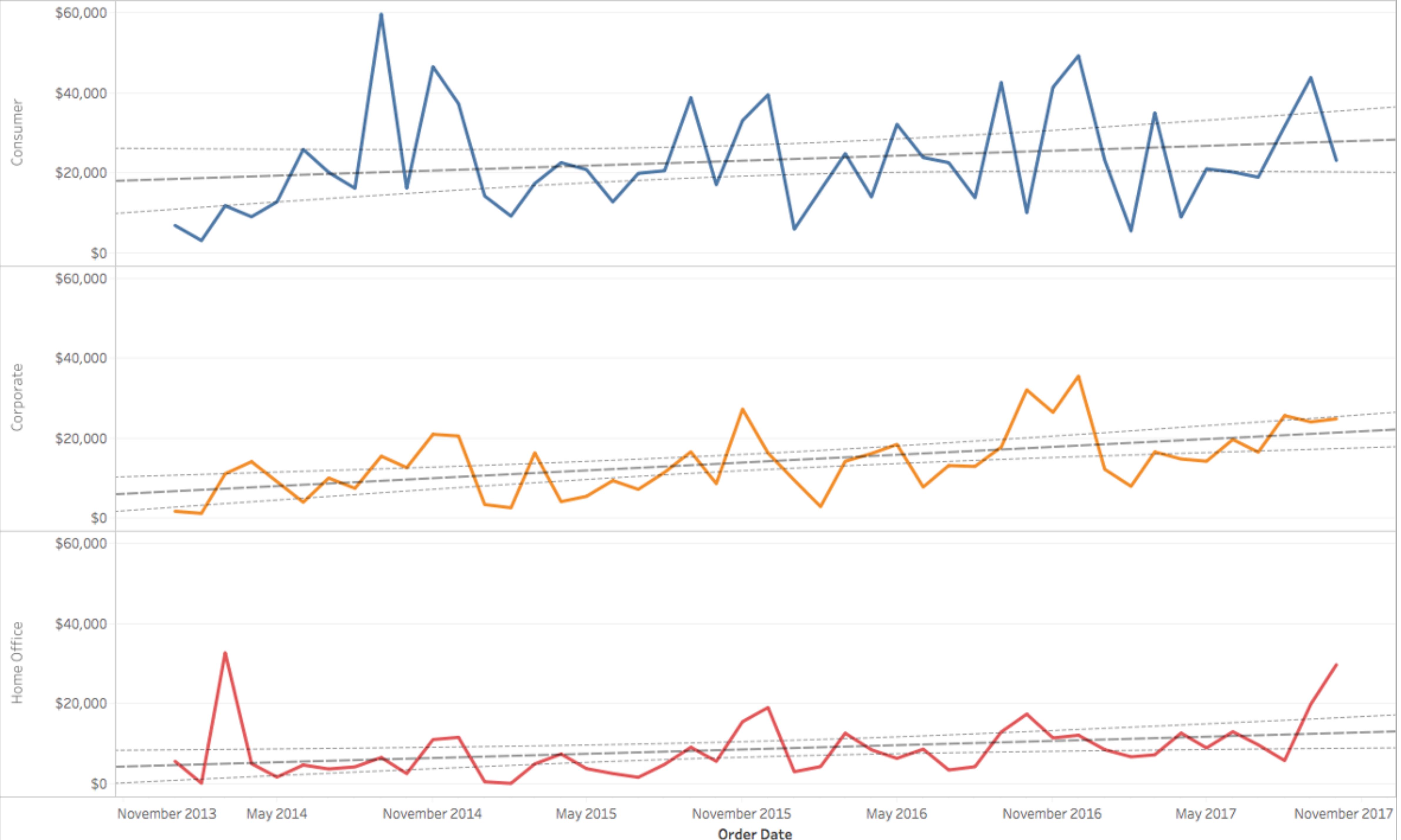
How should I visualize it?

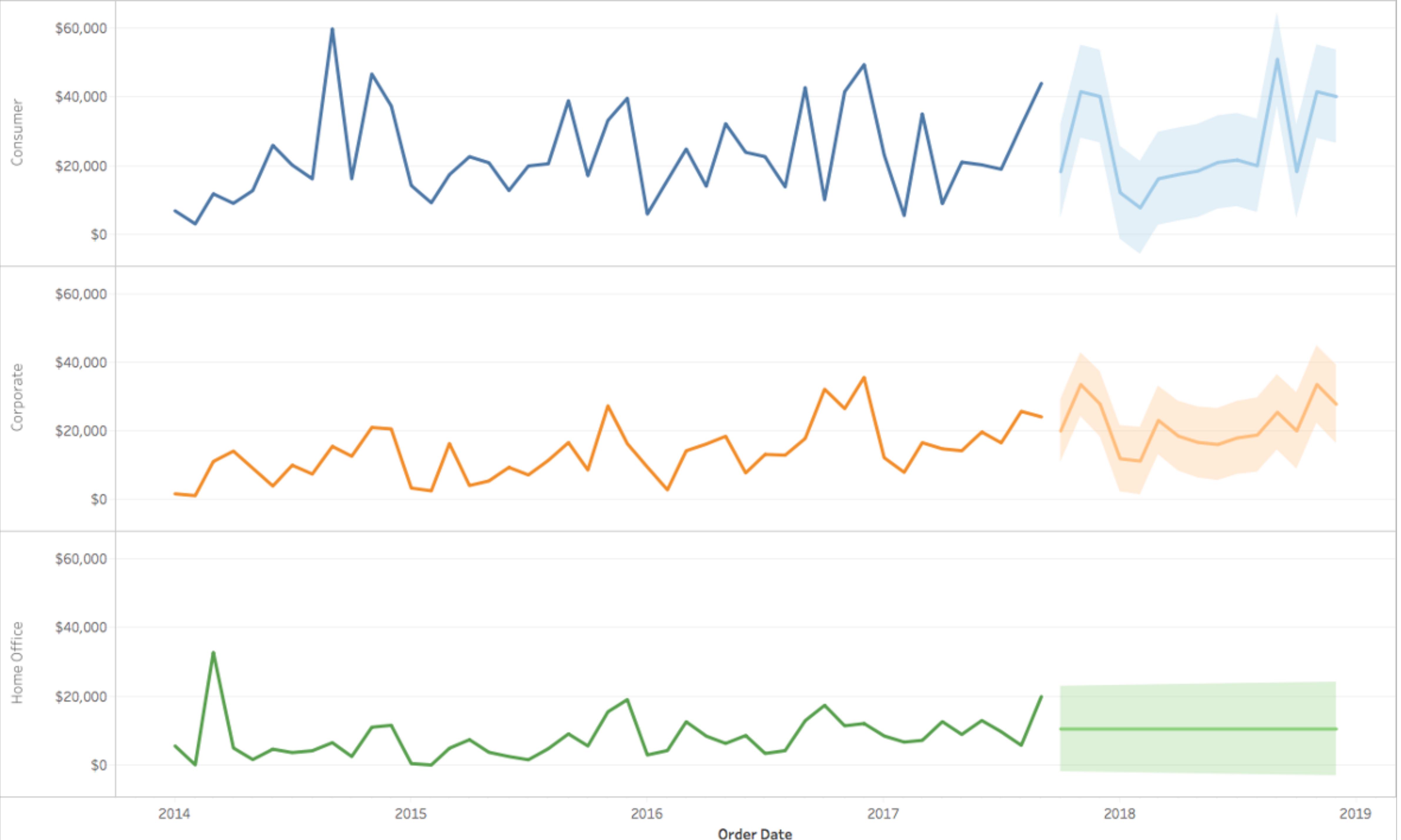
Uncertainty

What does it mean?

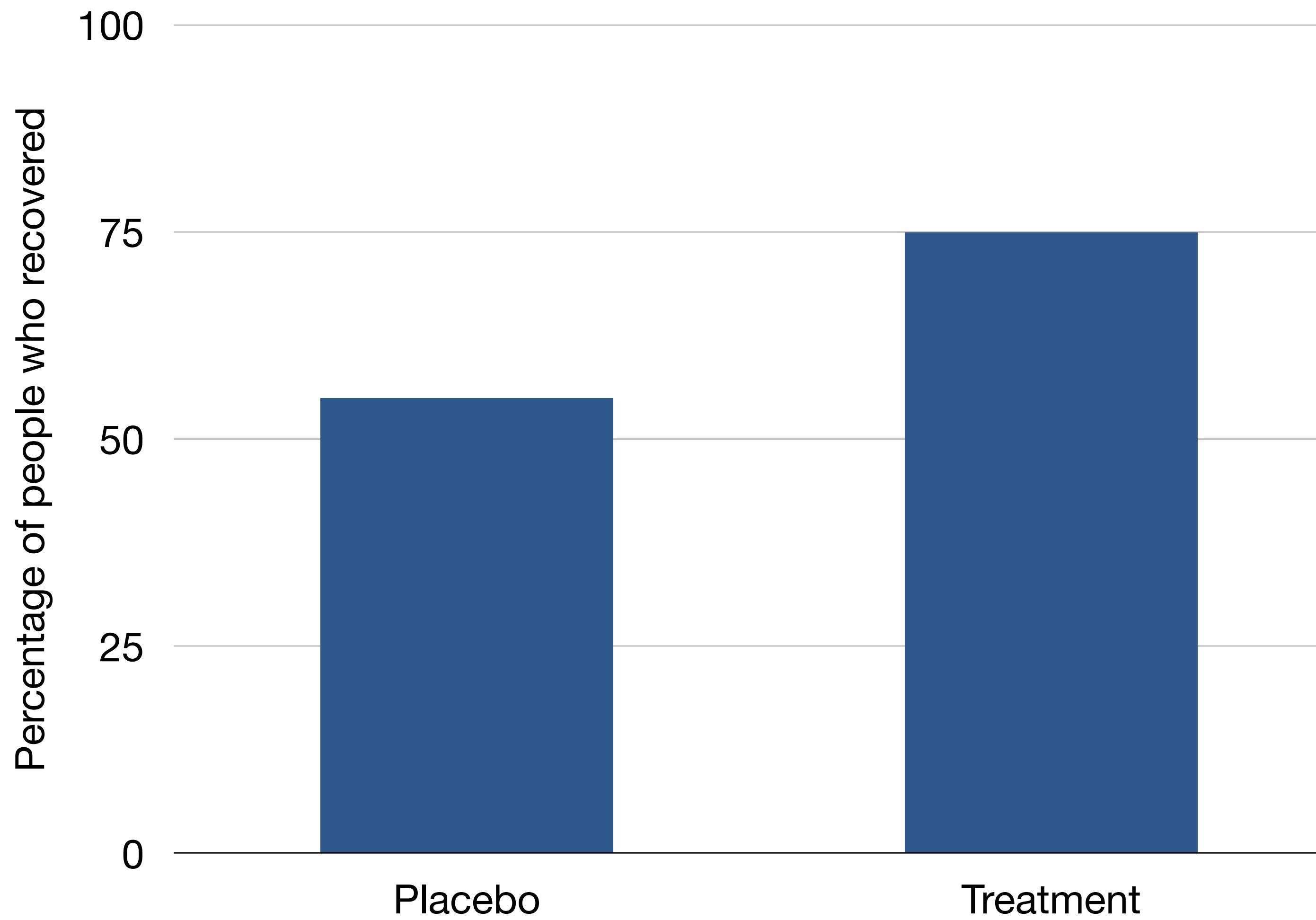
How should I visualize it?



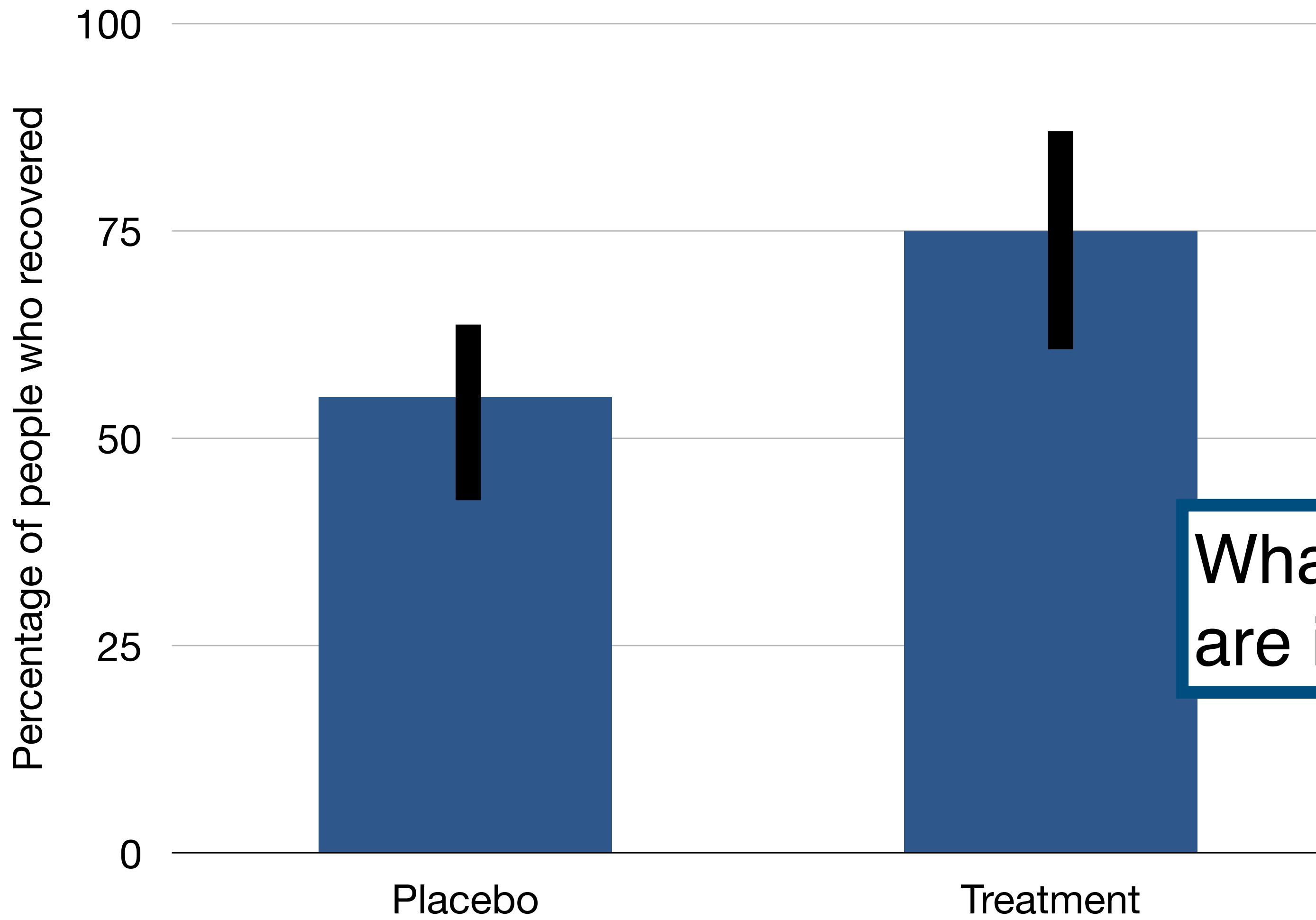




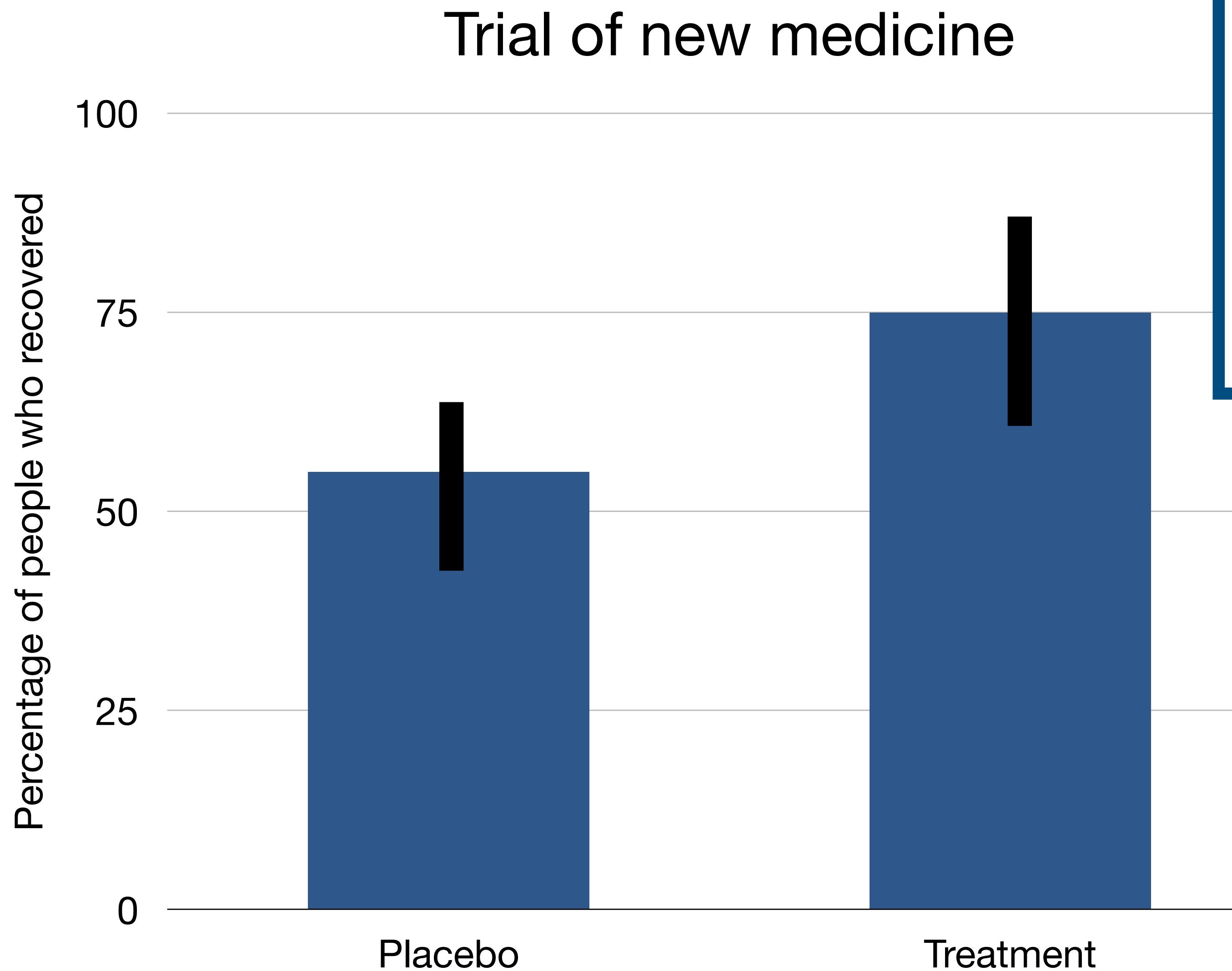
Trial of new medicine



Trial of new medicine

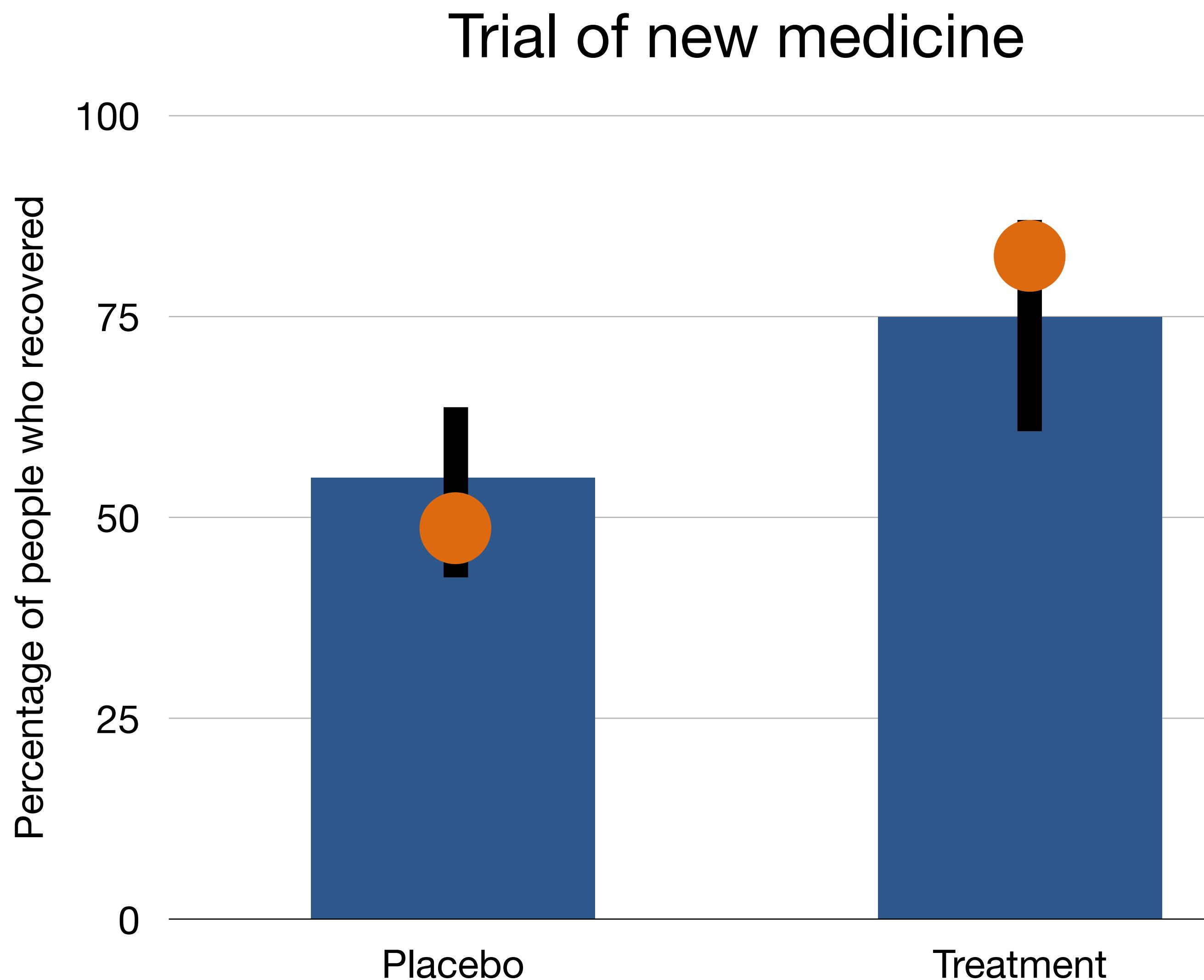


What do you think these error bars are implying?



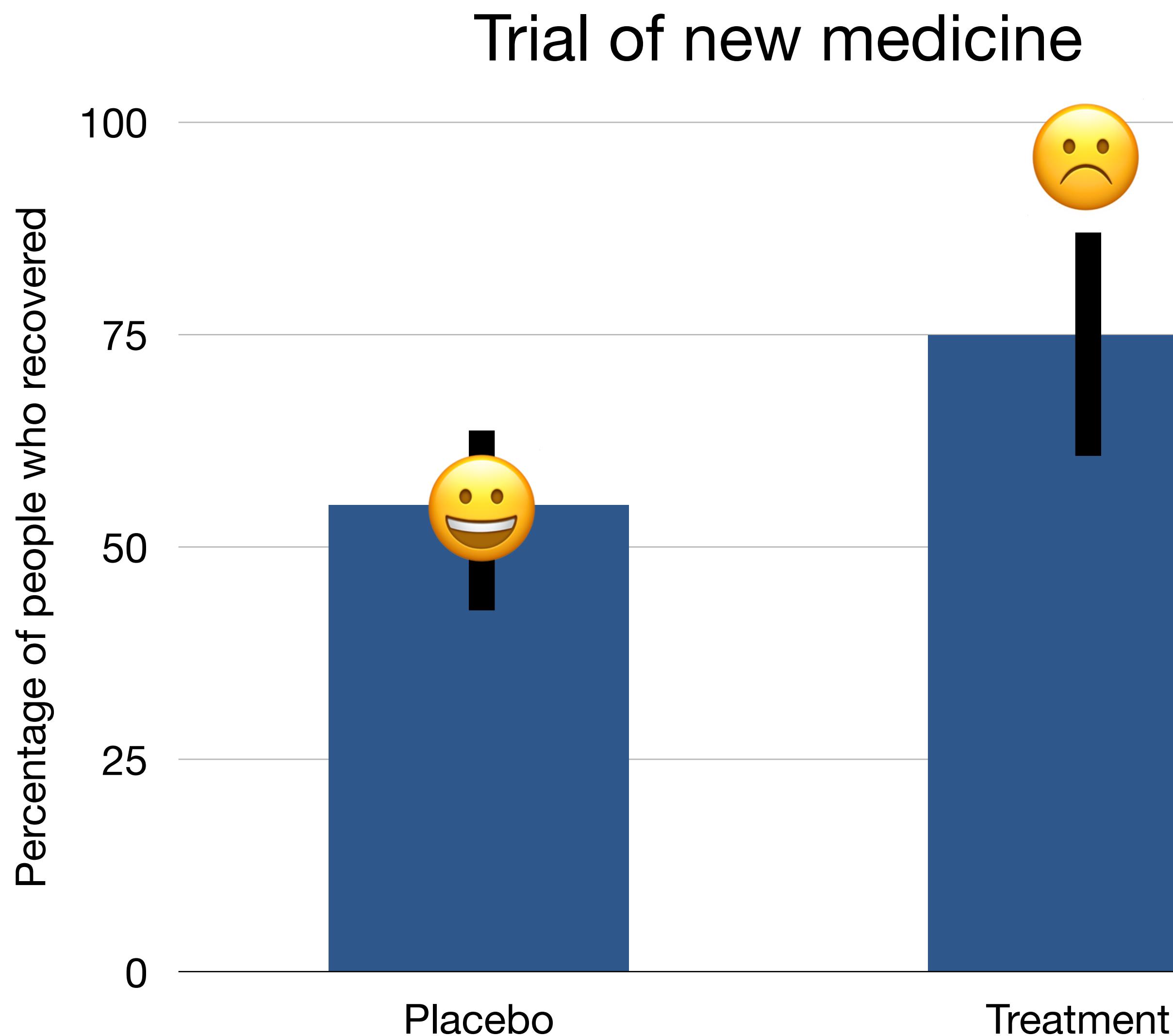
Expressiveness?

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



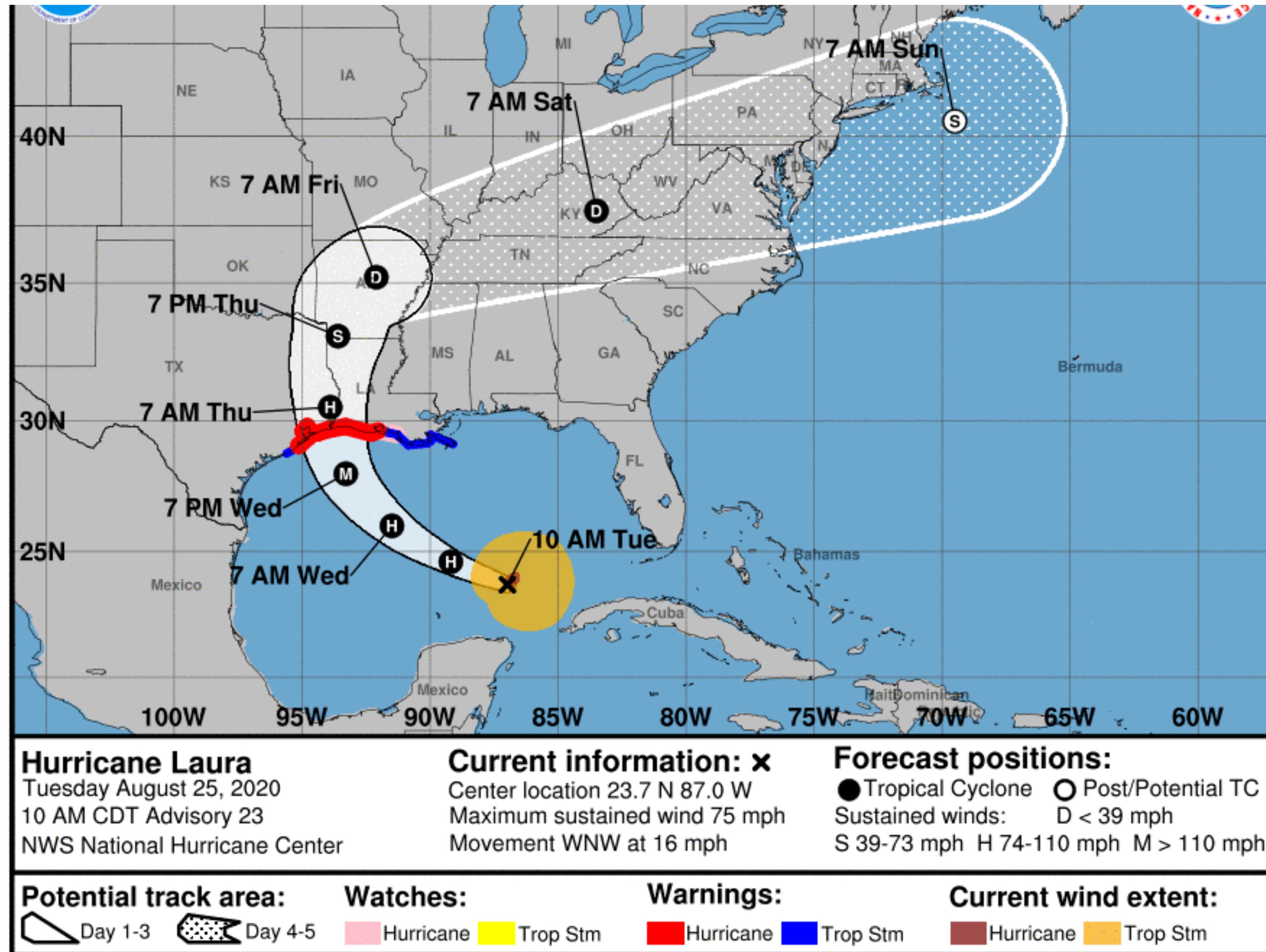
Expressiveness?

- ✖ 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.

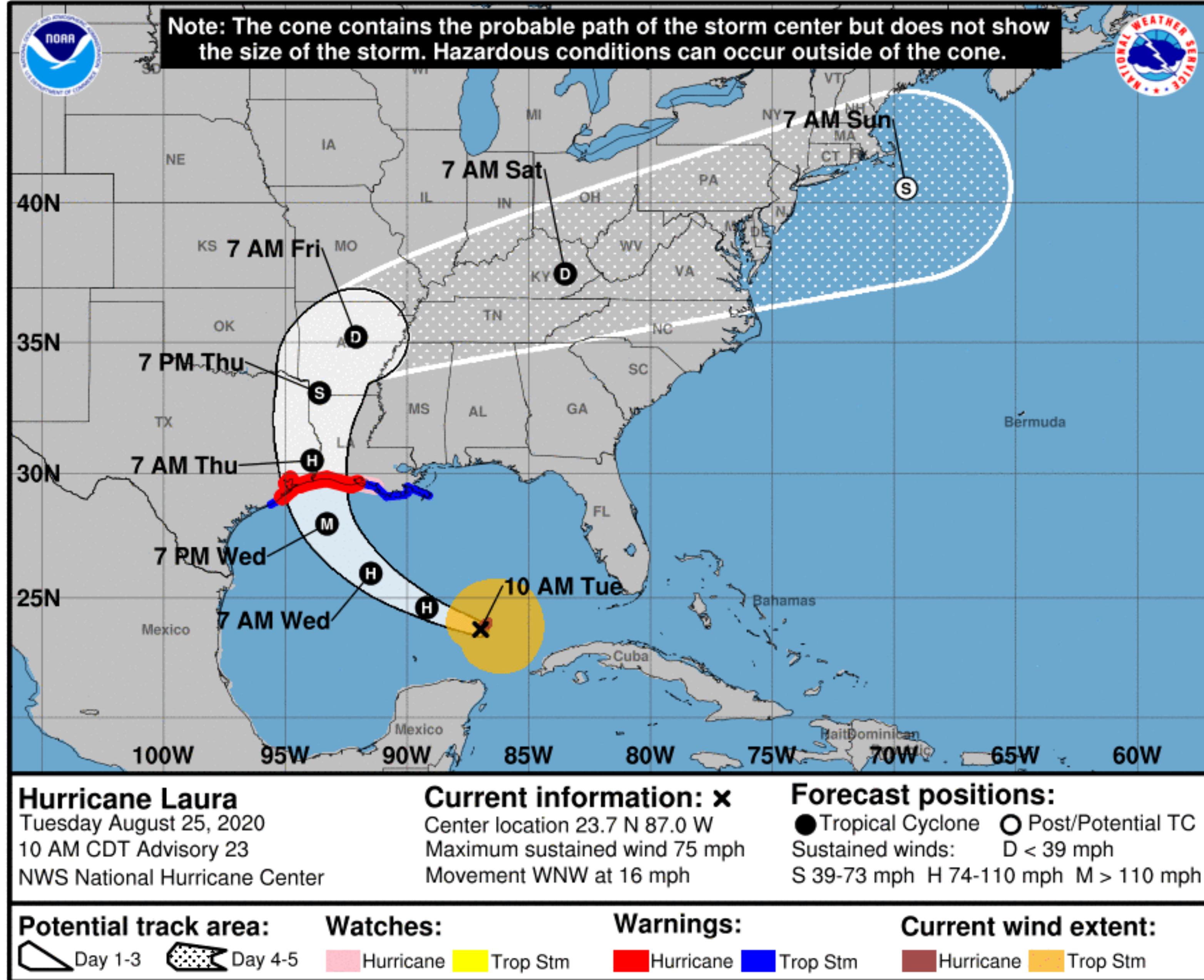


Expressiveness?

- ✖ 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.



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

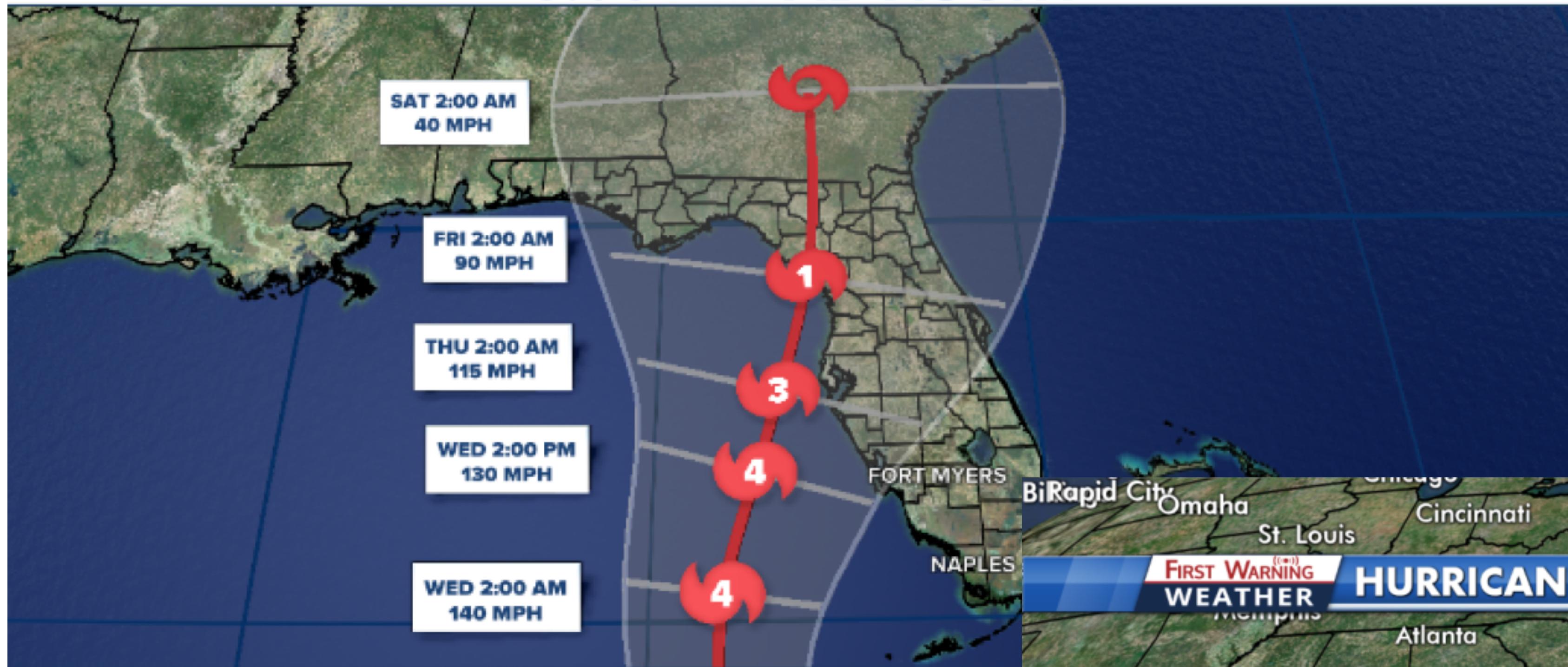


What is being visualized?
What are the strengths and weaknesses of this visualization?

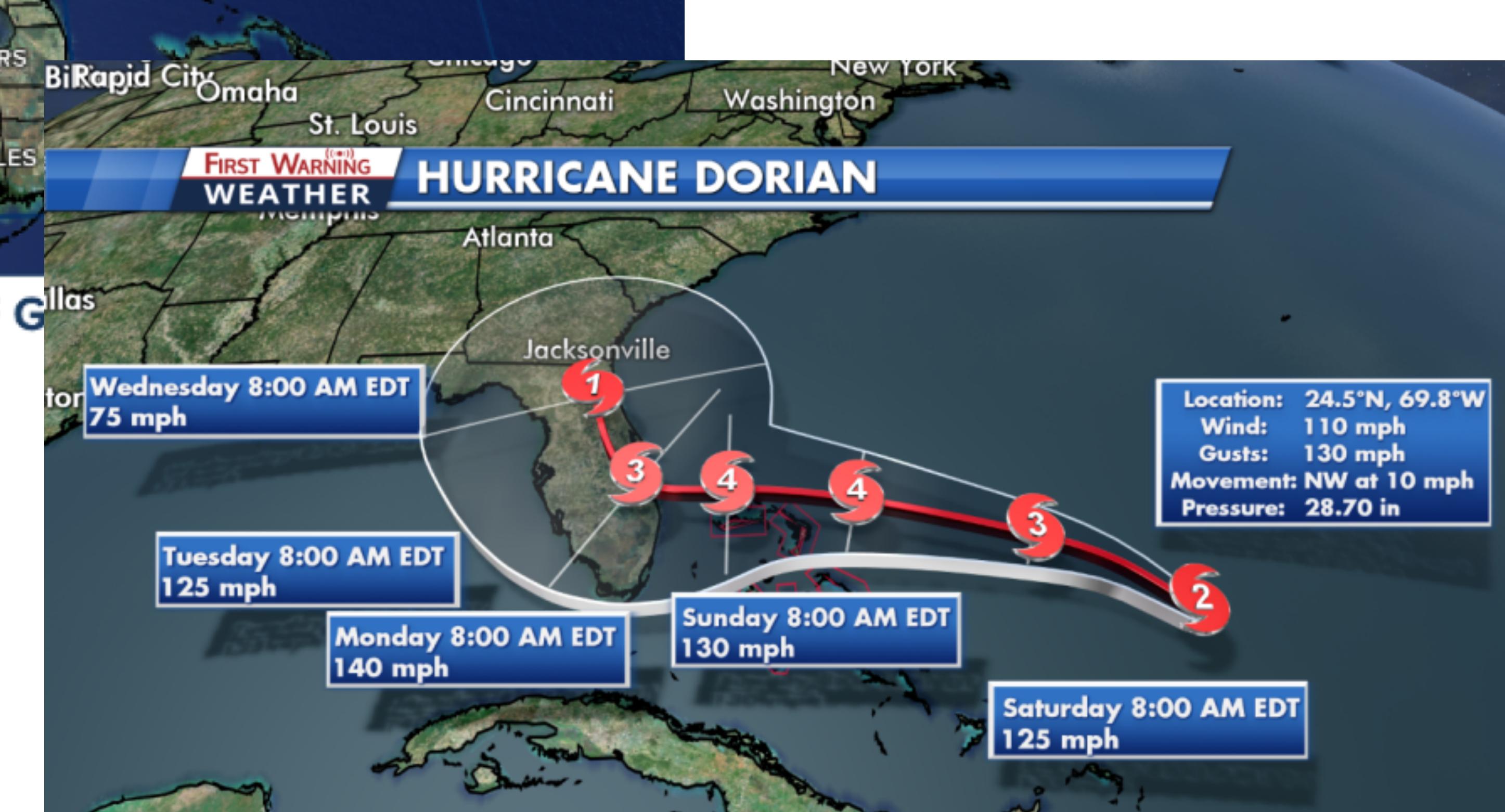
FOX 4

HURRICANE IAN

5:00 AM ADVISORY

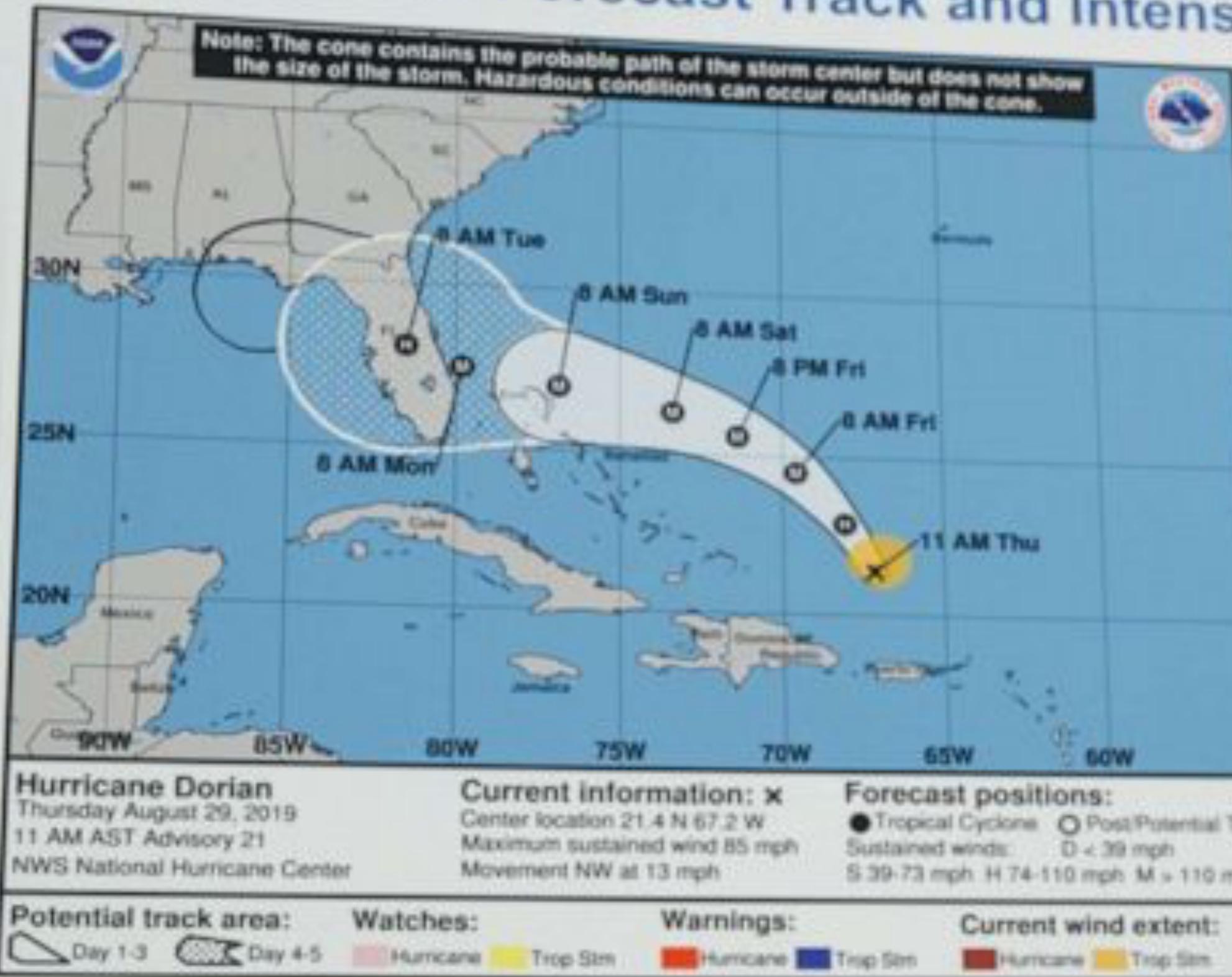


CURRENT LOCATION: 90 MI SW OF GULFPORT, MS

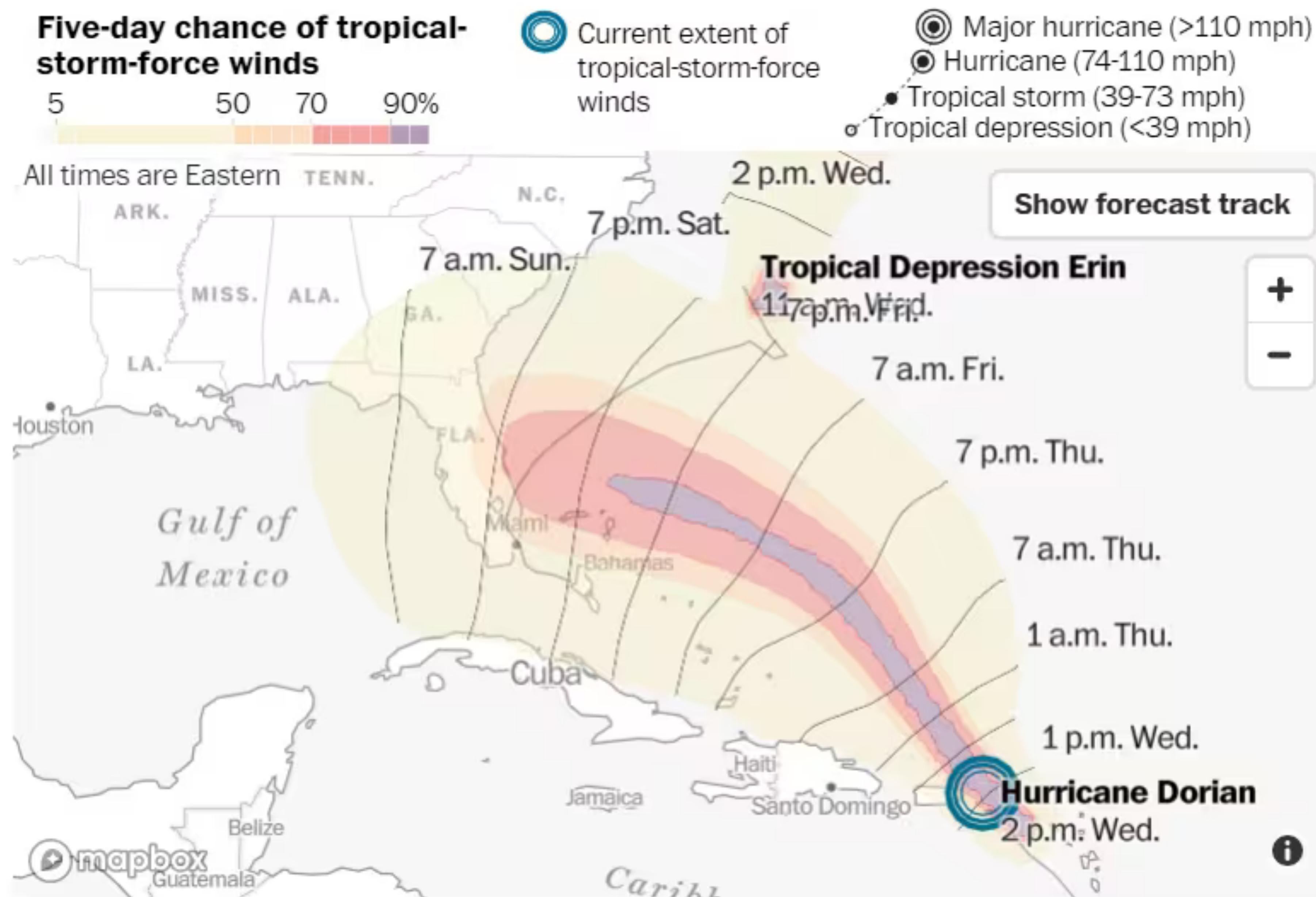




Hurricane Dorian Forecast Track and Intensity



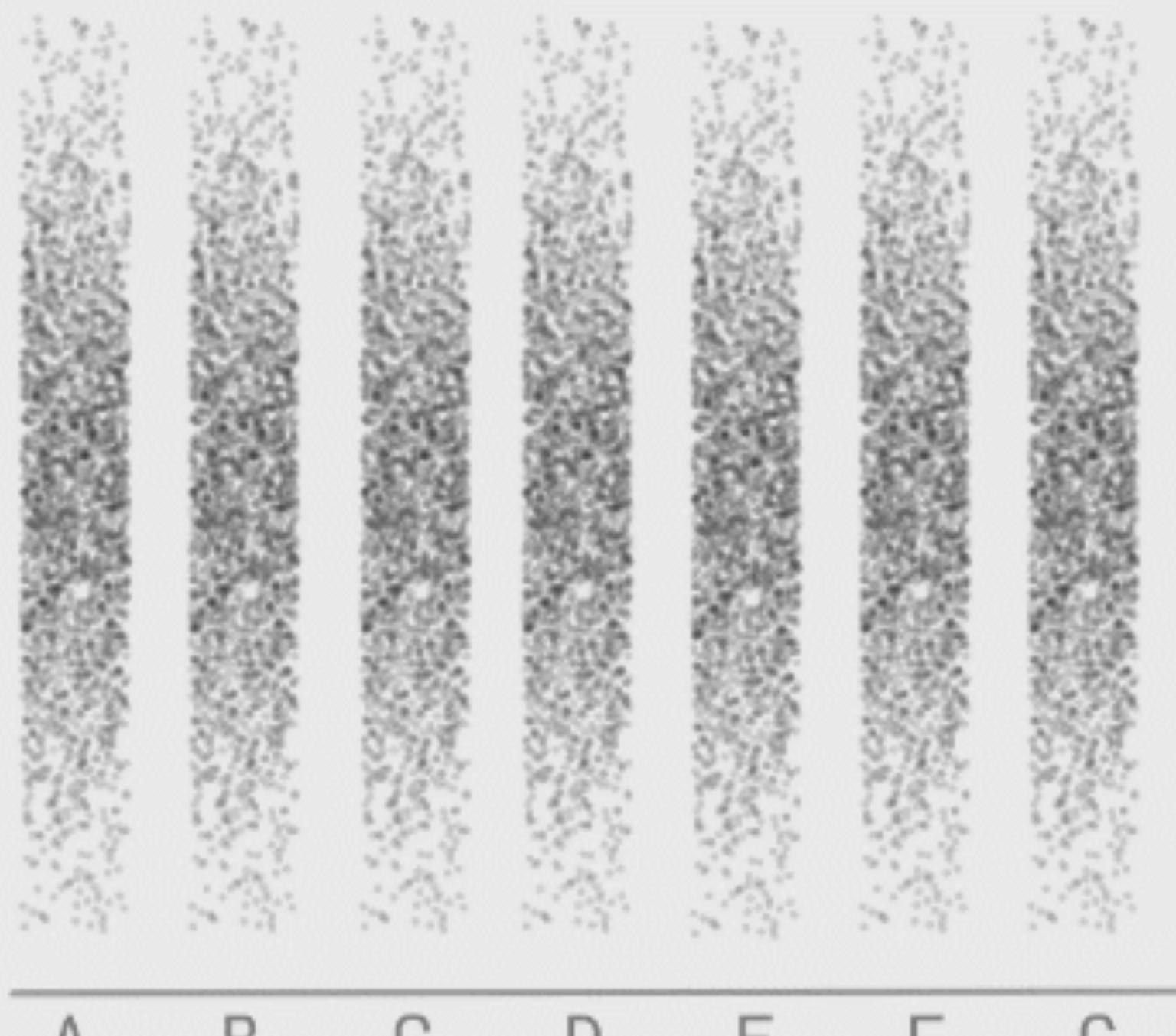
Five-day chance of tropical storm-force winds



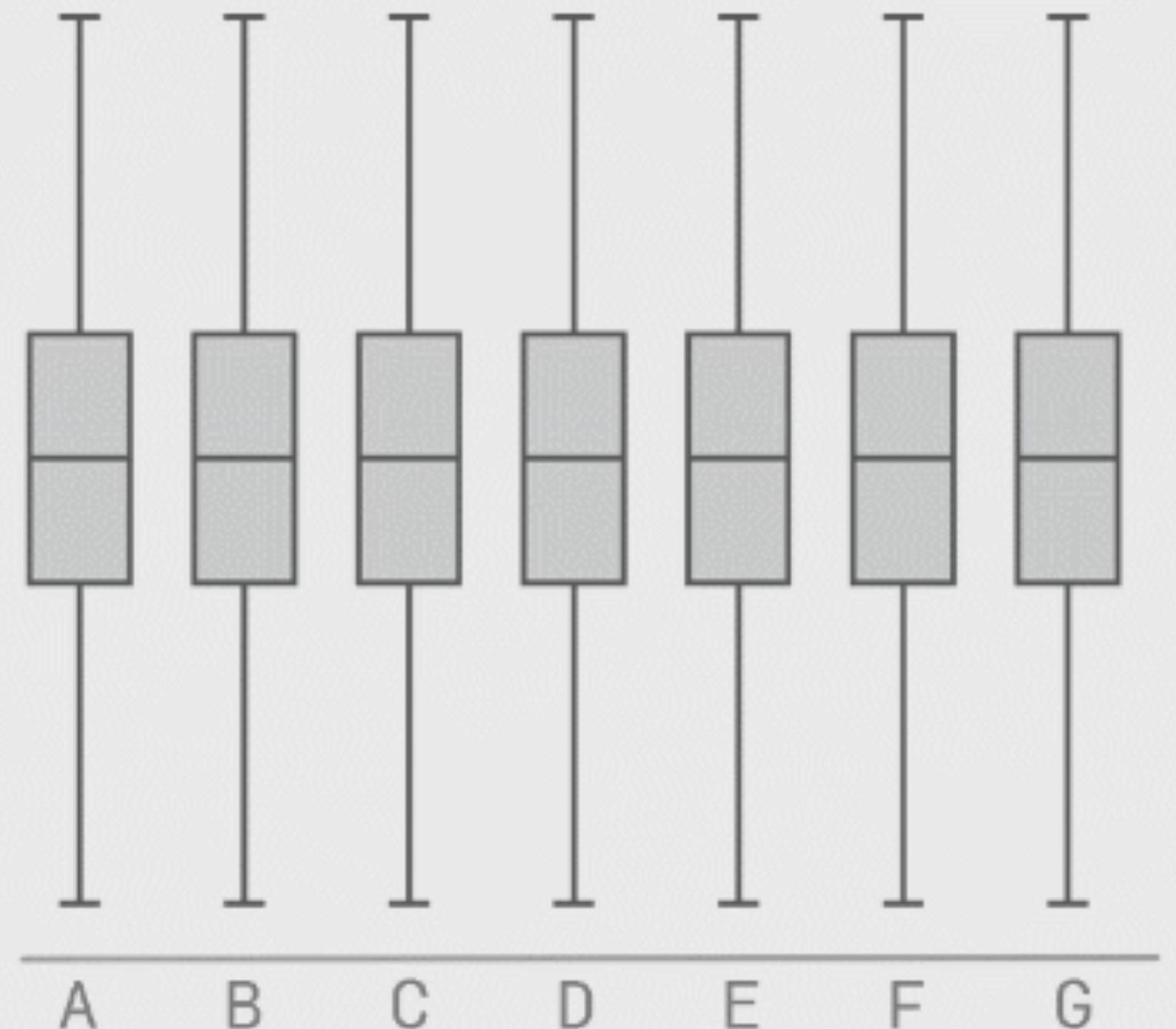
Source: National Weather Service. Note: Impact lines represent the earliest reasonable arrival time of tropical-storm-force winds.

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

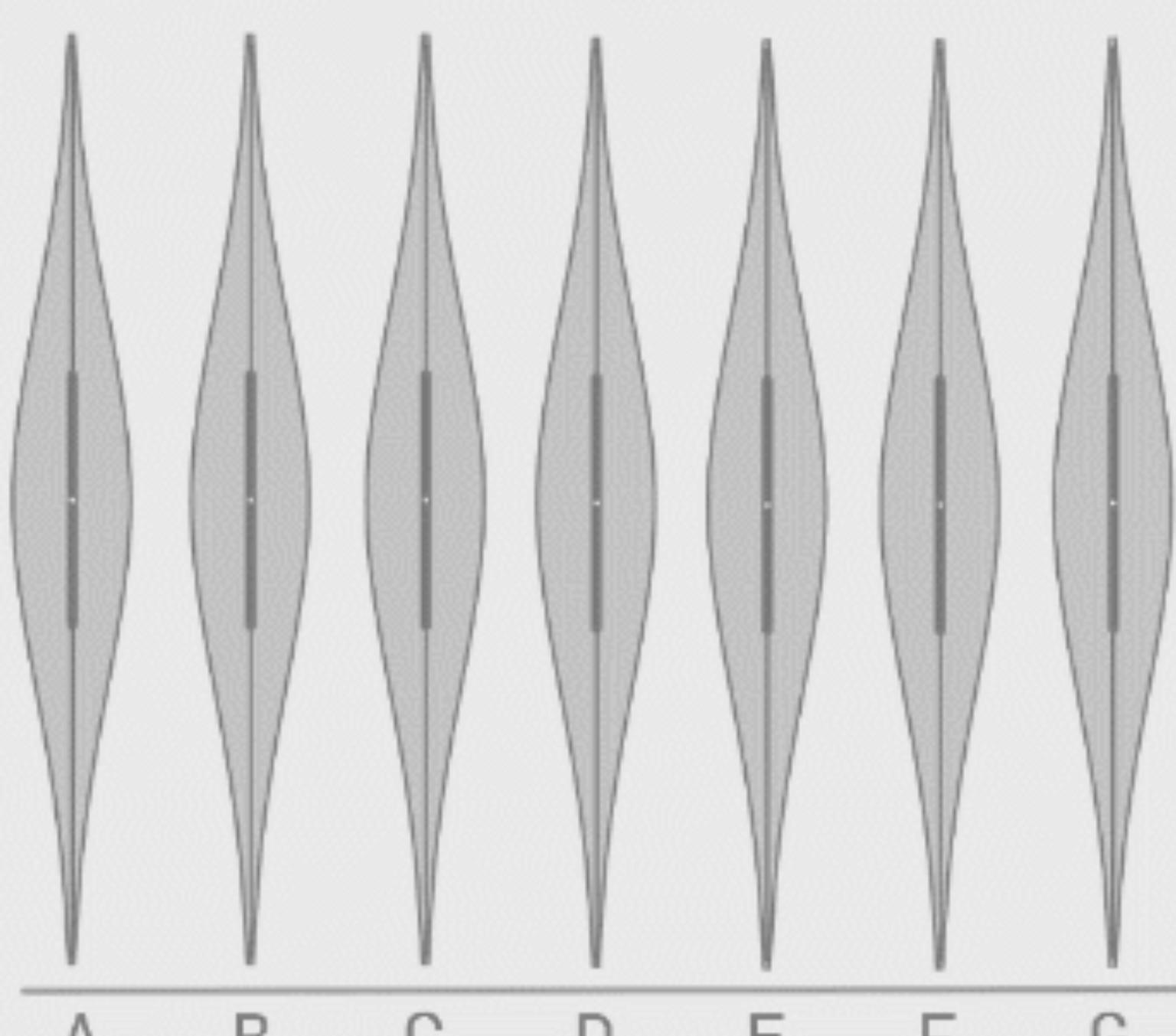
Raw Data



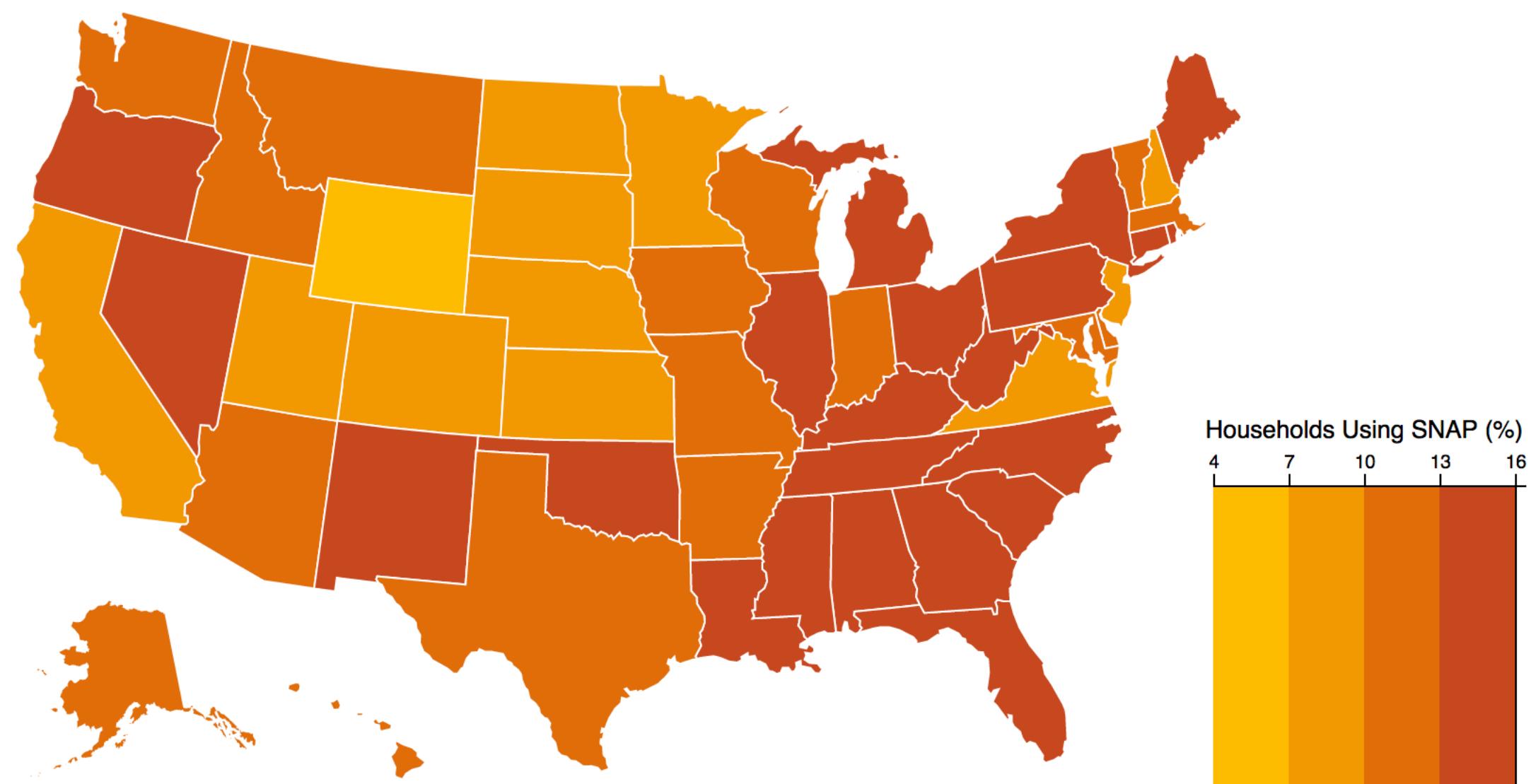
Box-plot of the Data



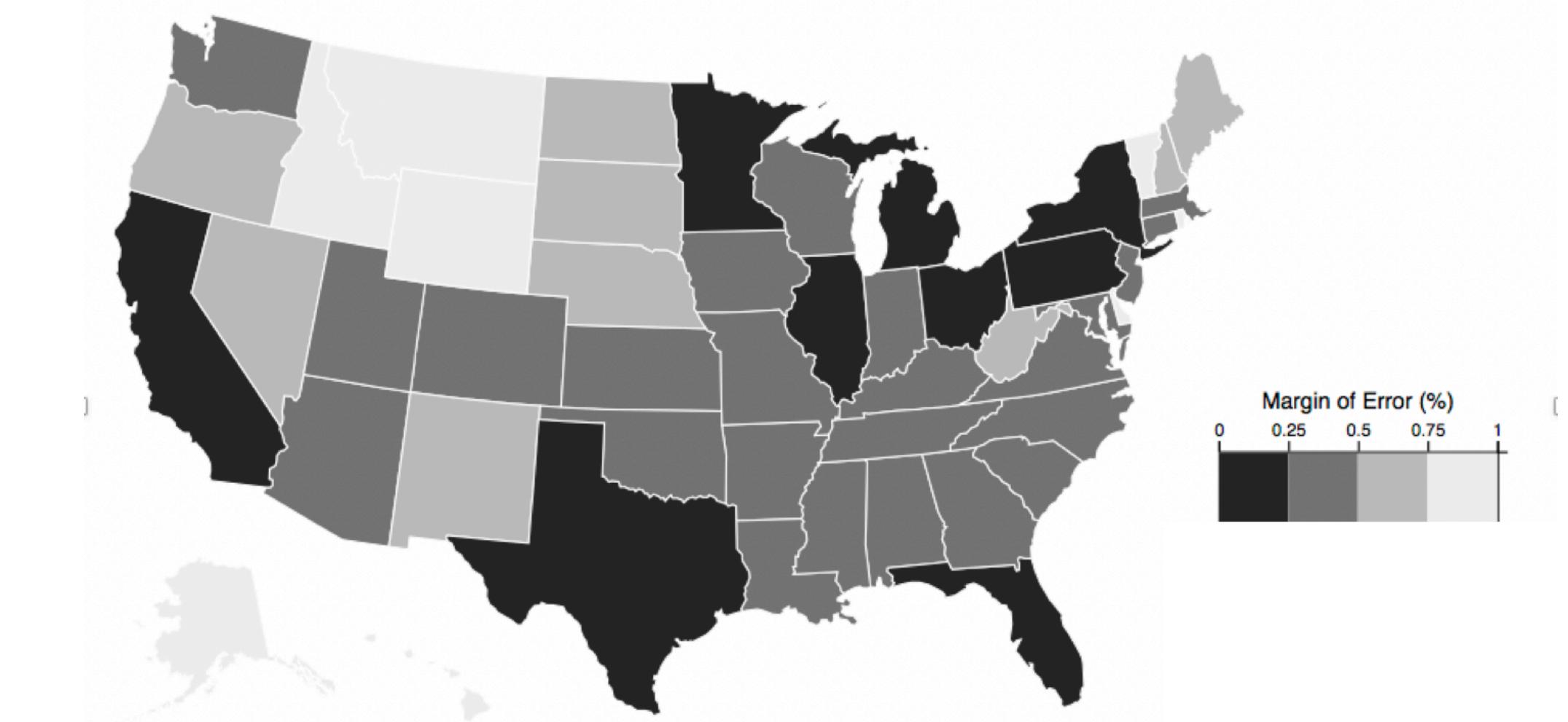
Violin-plot of the Data



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

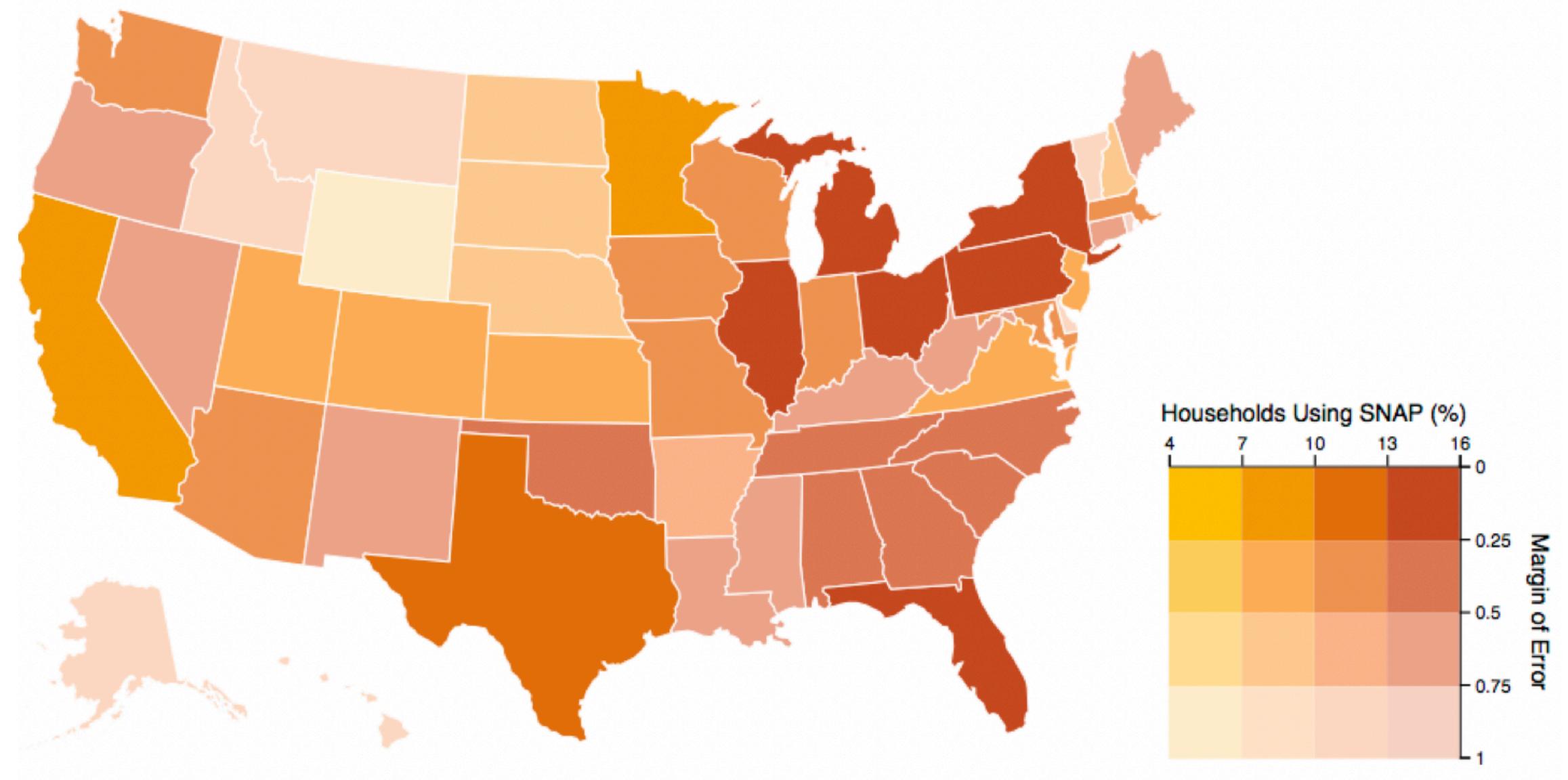


Data Map

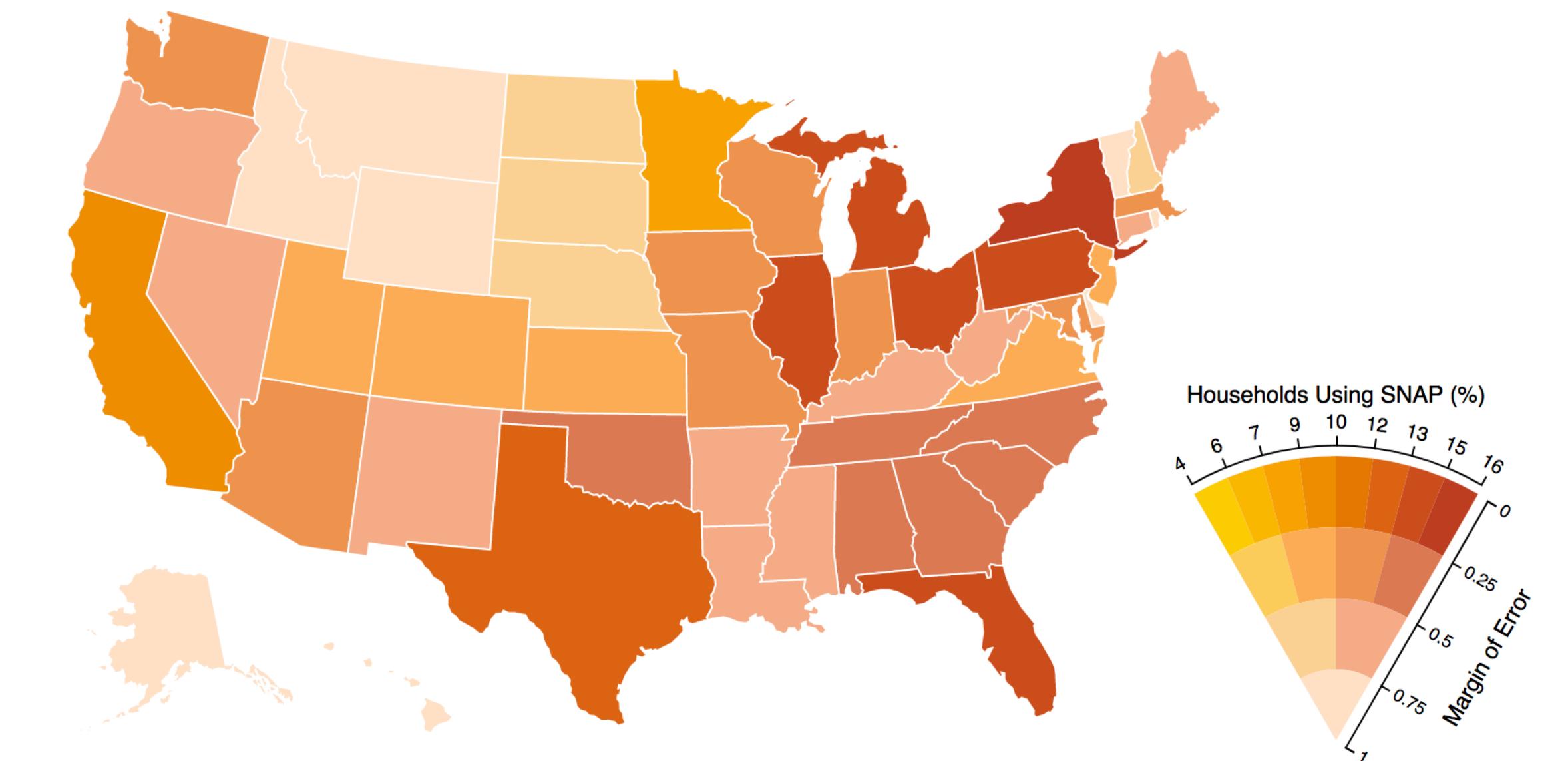


Uncertainty Map

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

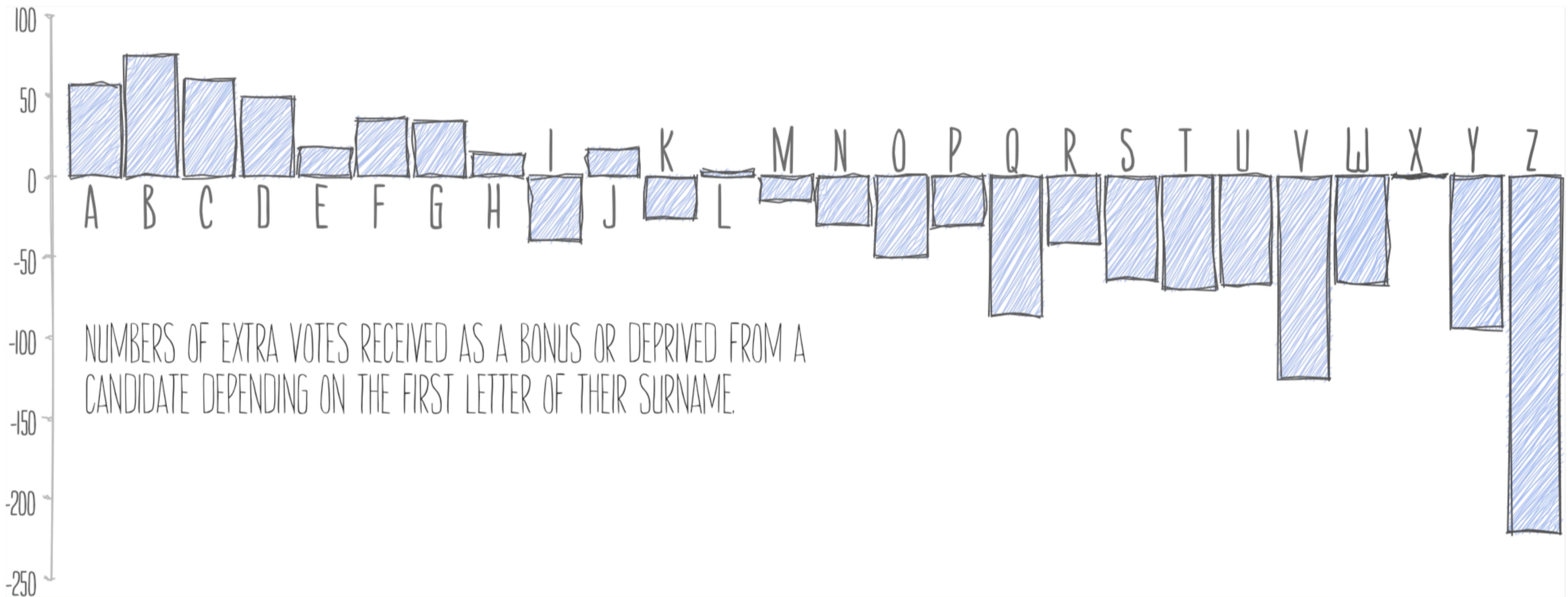


Bivariate Map (Data + Uncertainty)



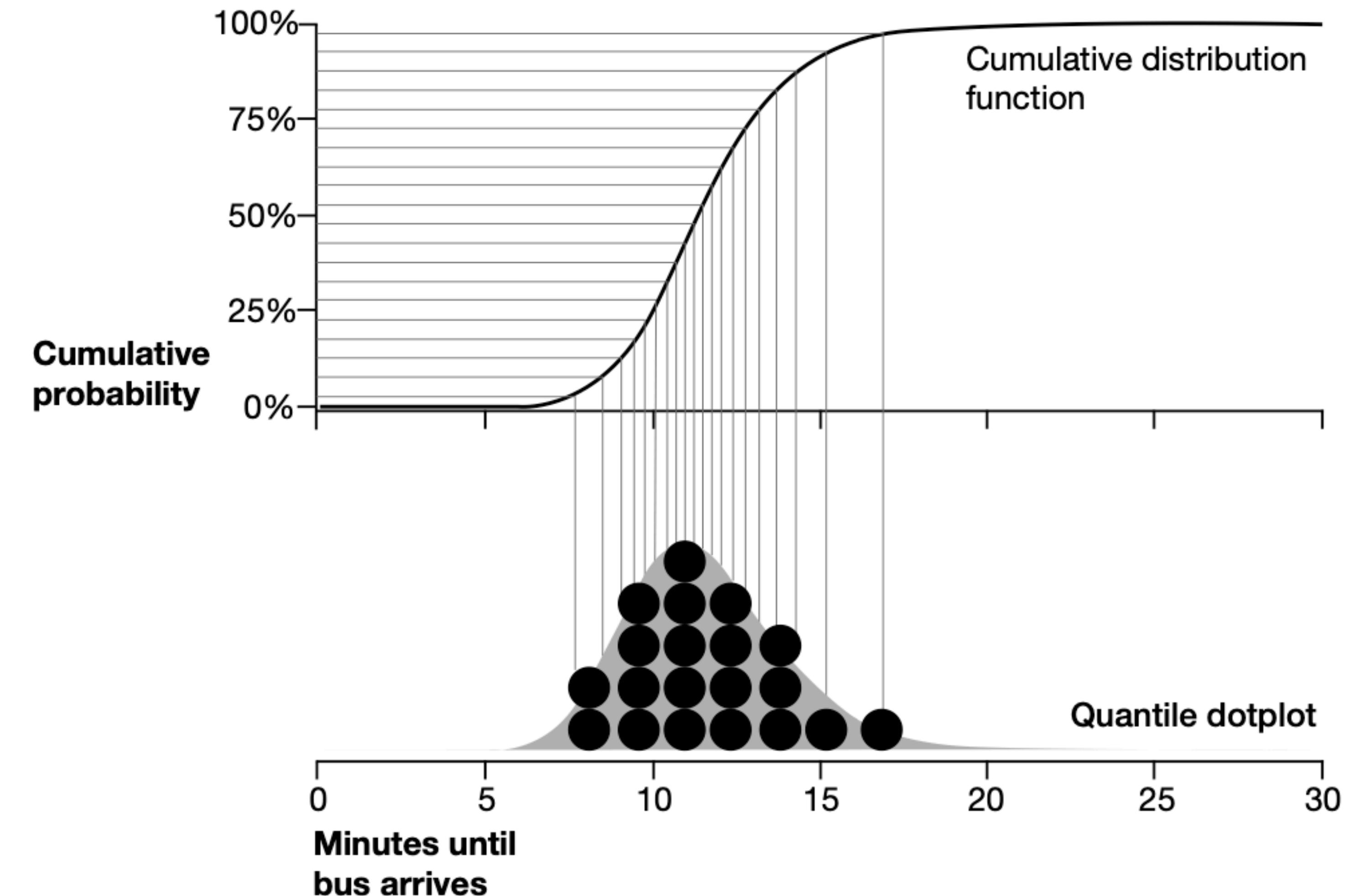
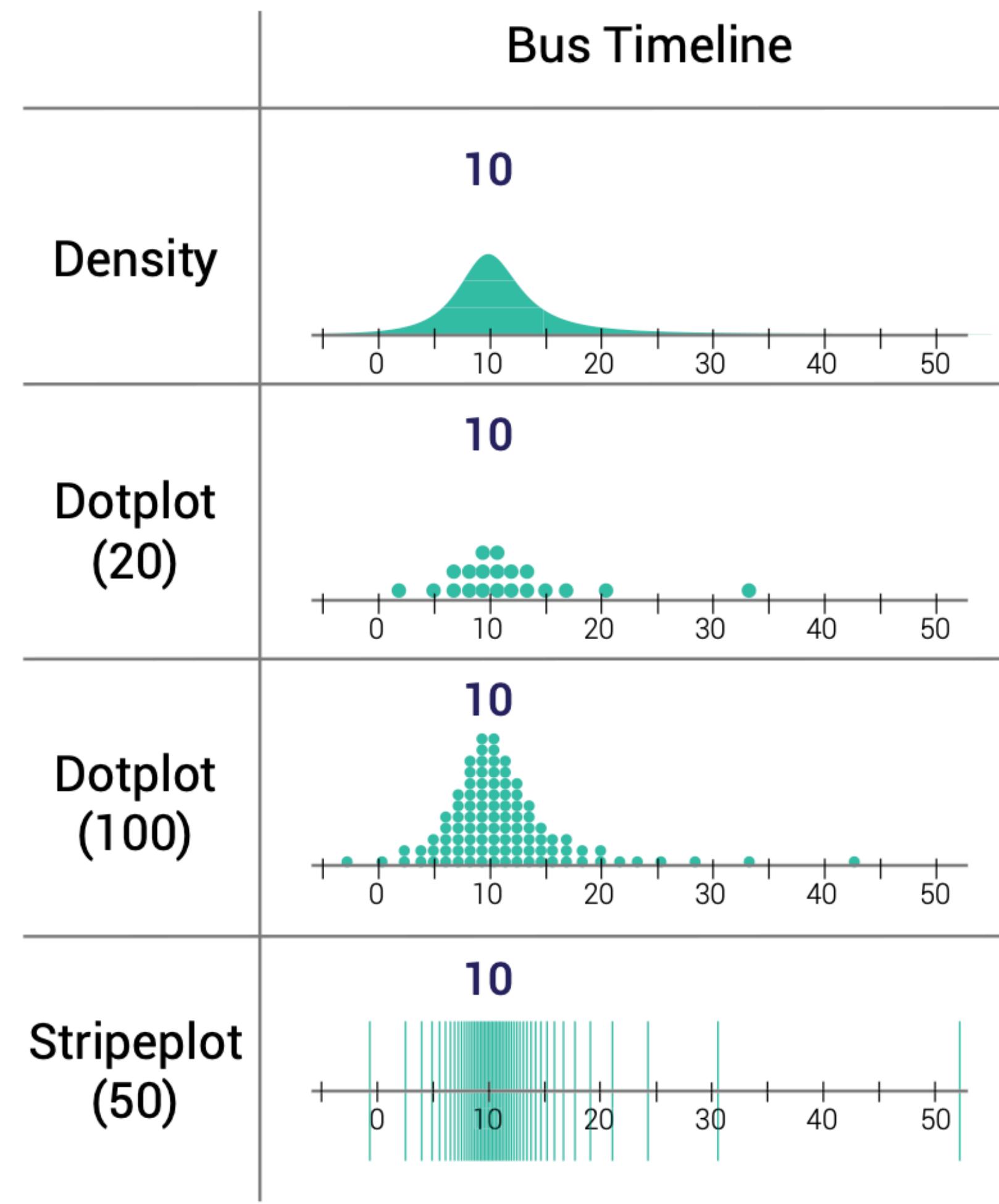
Value-Suppressing Uncertainty Map

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

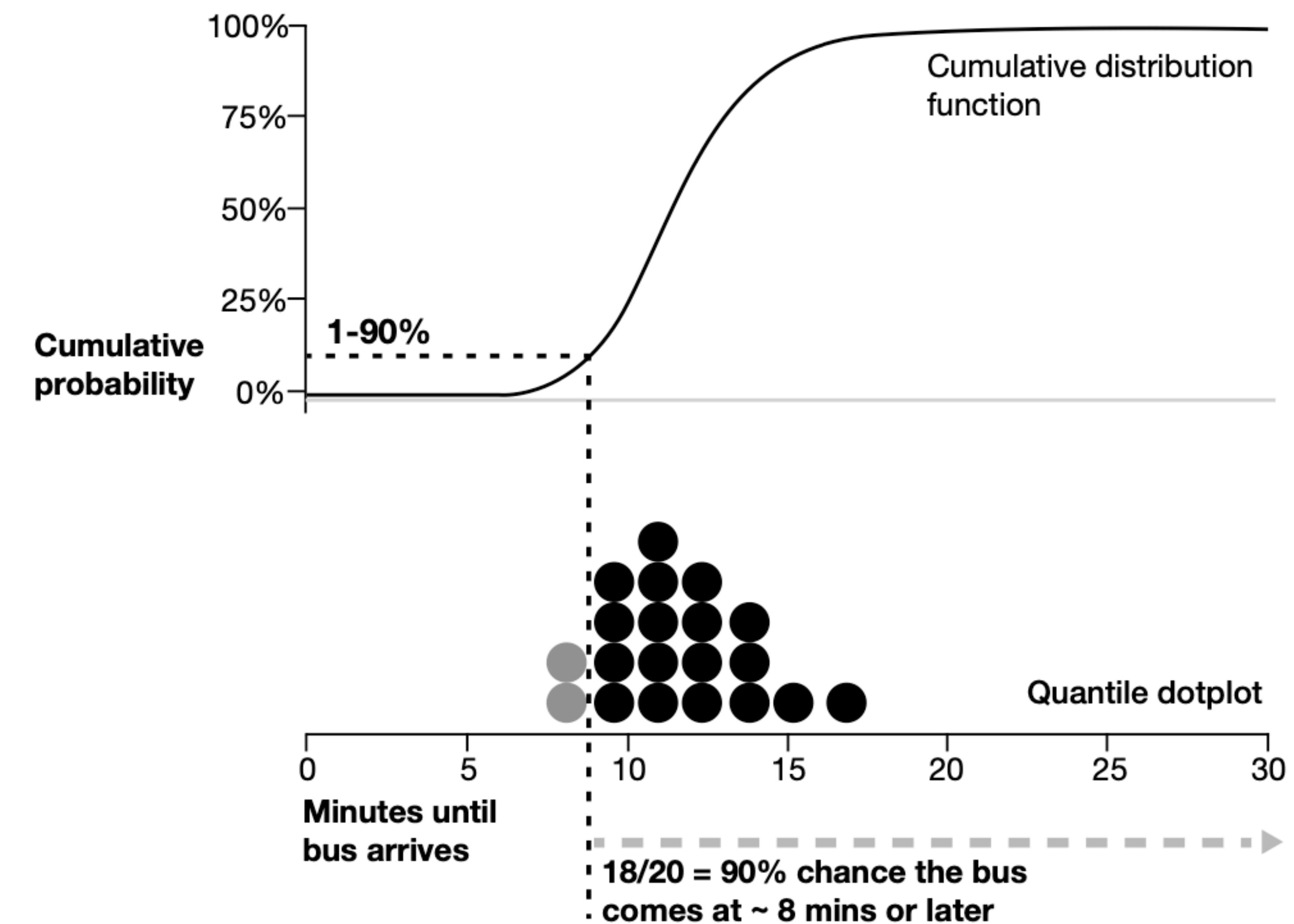
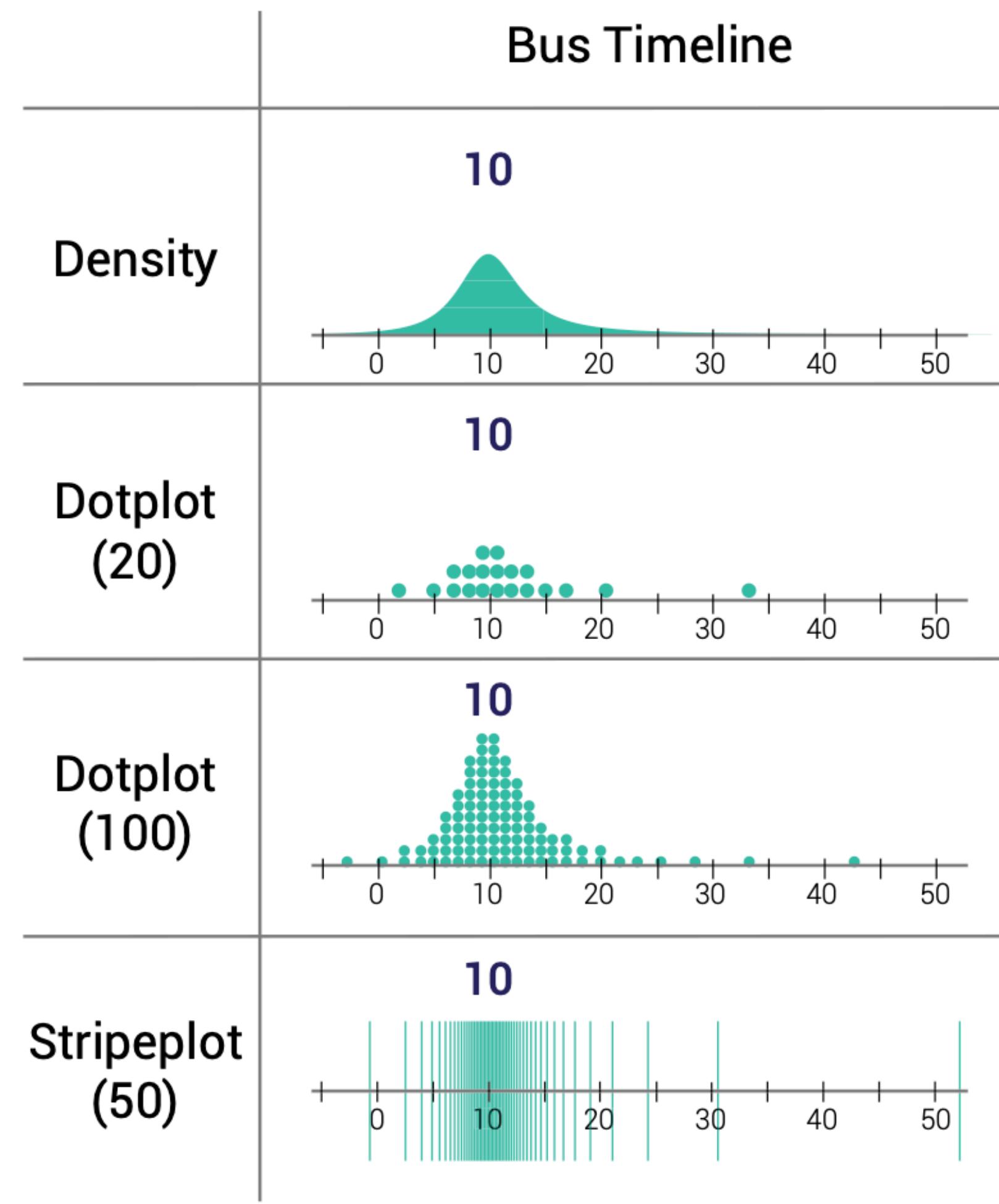


[Wood et al., 2012]
[Boukhelifa et al., 2012]

"Set of draws" technique

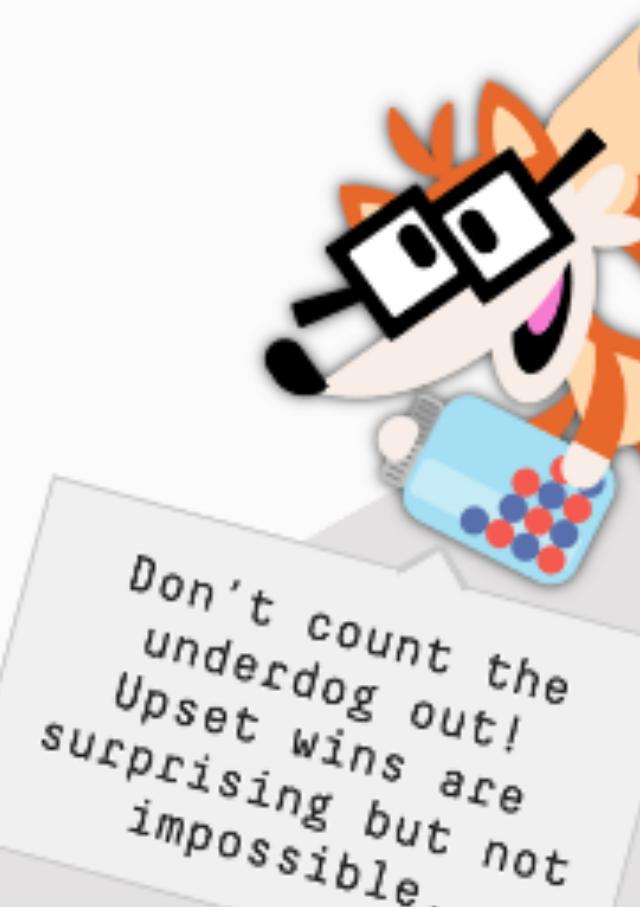
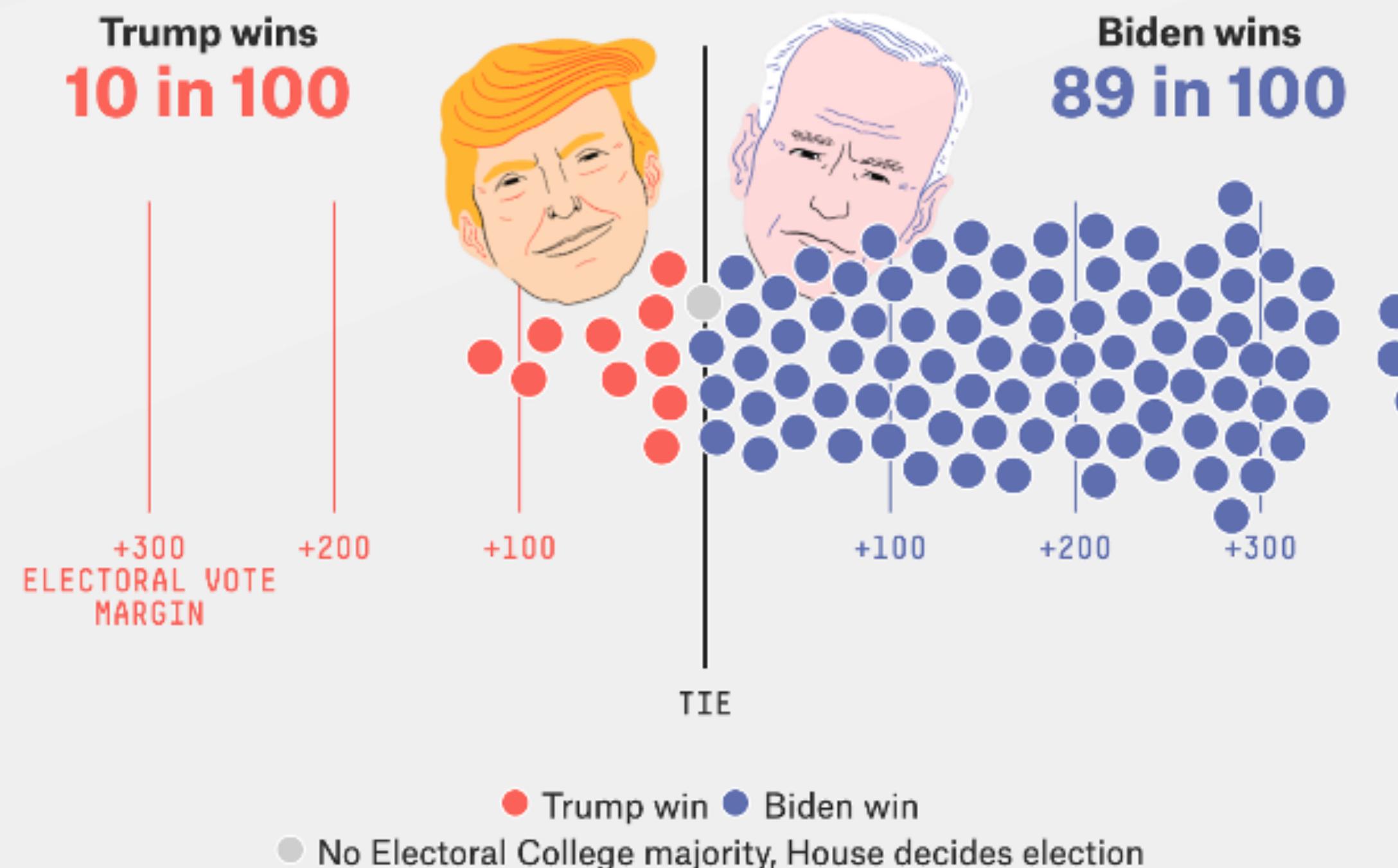


"Set of draws" technique



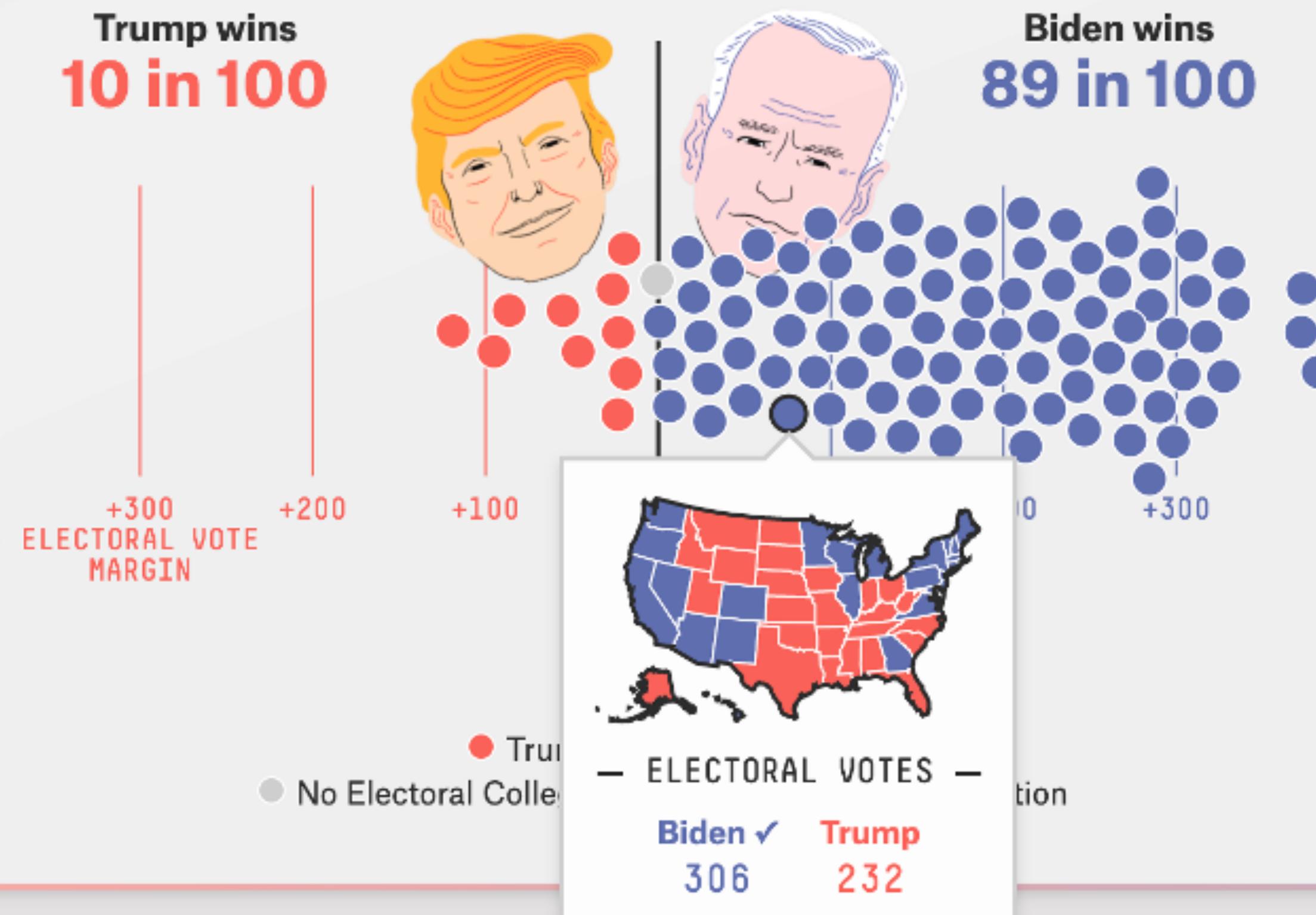
Biden is *favored* to win the election

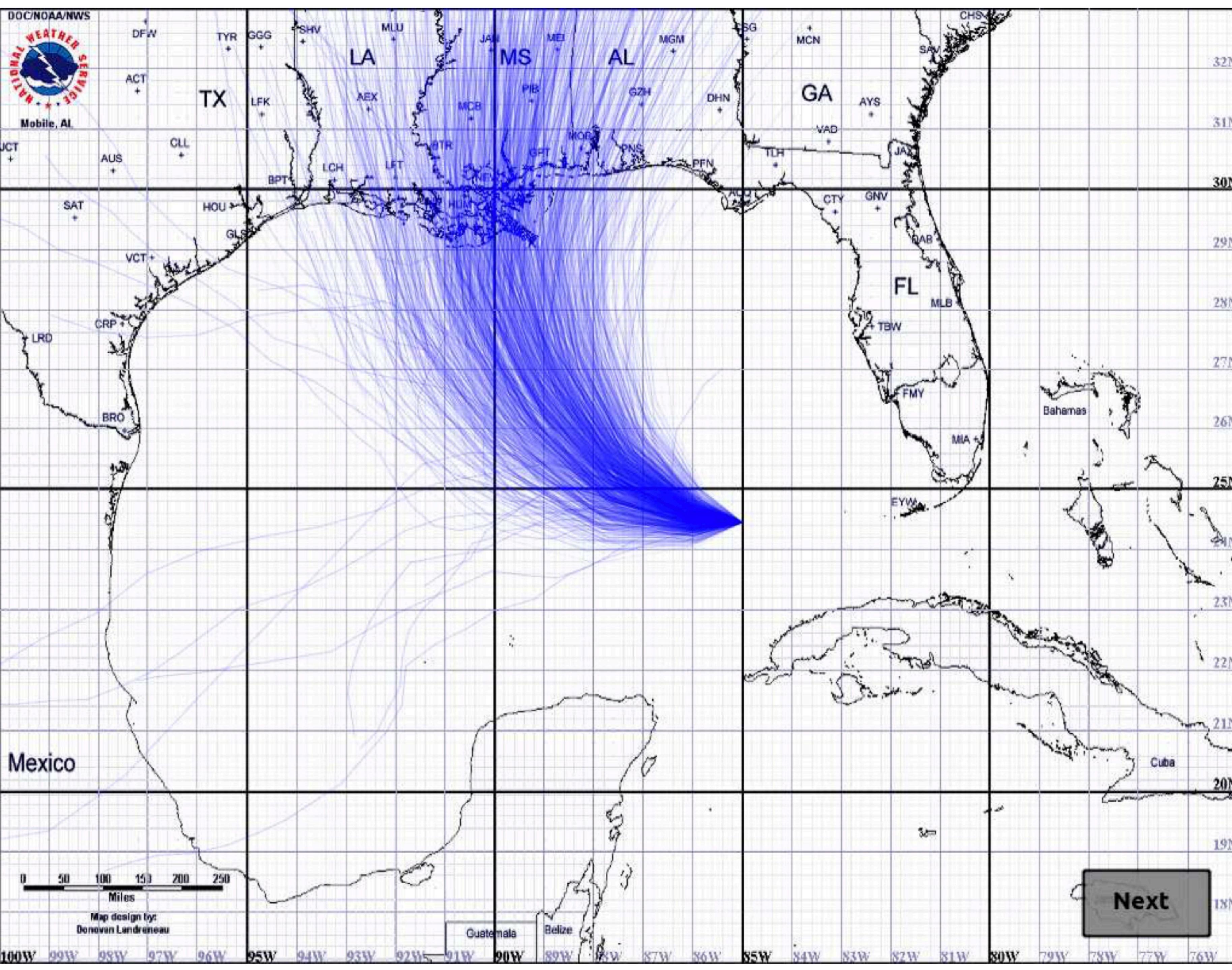
We simulate the election 40,000 times to see who wins most often. The sample of 100 outcomes below gives you a good idea of the range of scenarios our model thinks is possible.



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.





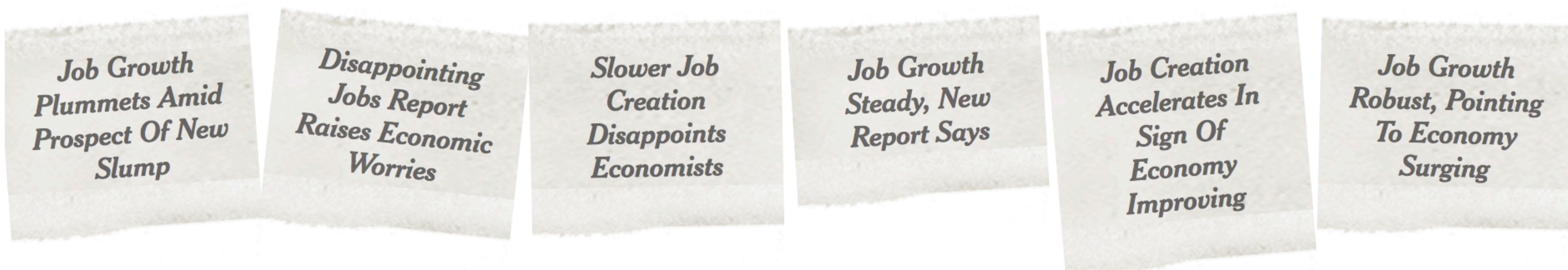
 **TheUpshot**

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

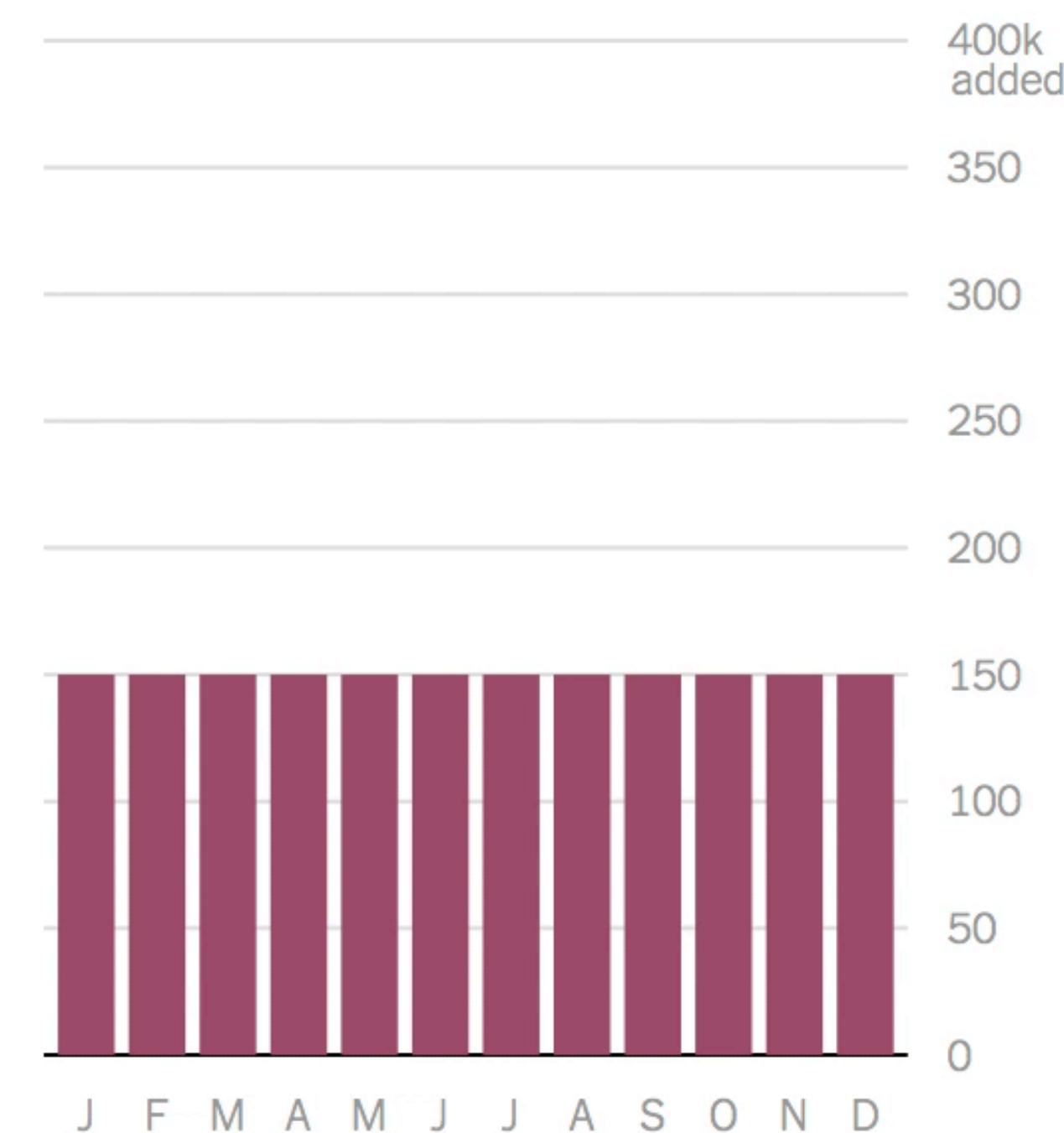
110,000 to 140,000
19% chance

160,000 to 190,000
19% chance

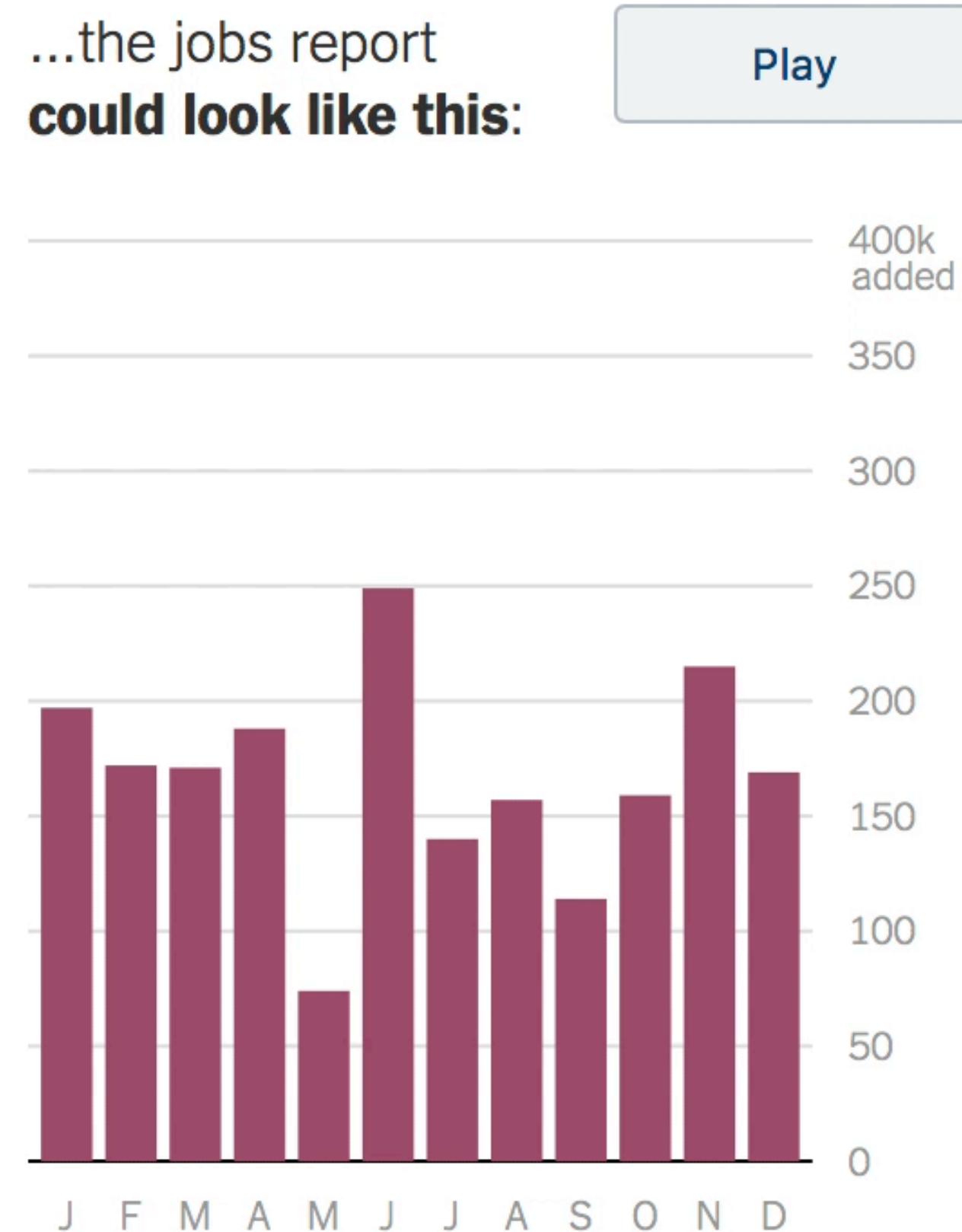
190,000 to 245,000
19% chance

245,000+
4% chance

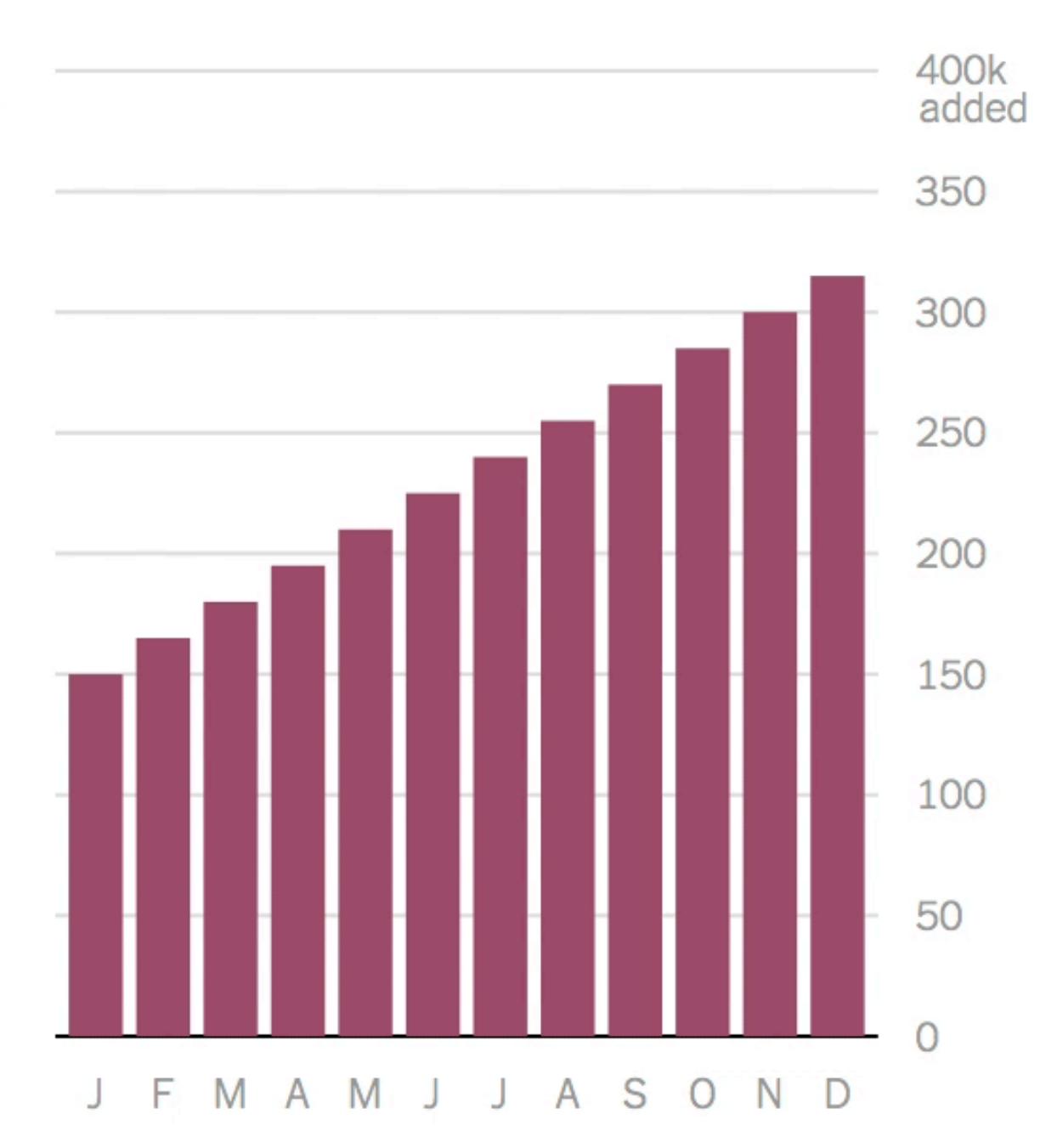
If job growth **were actually steady**
over the last 12 months...



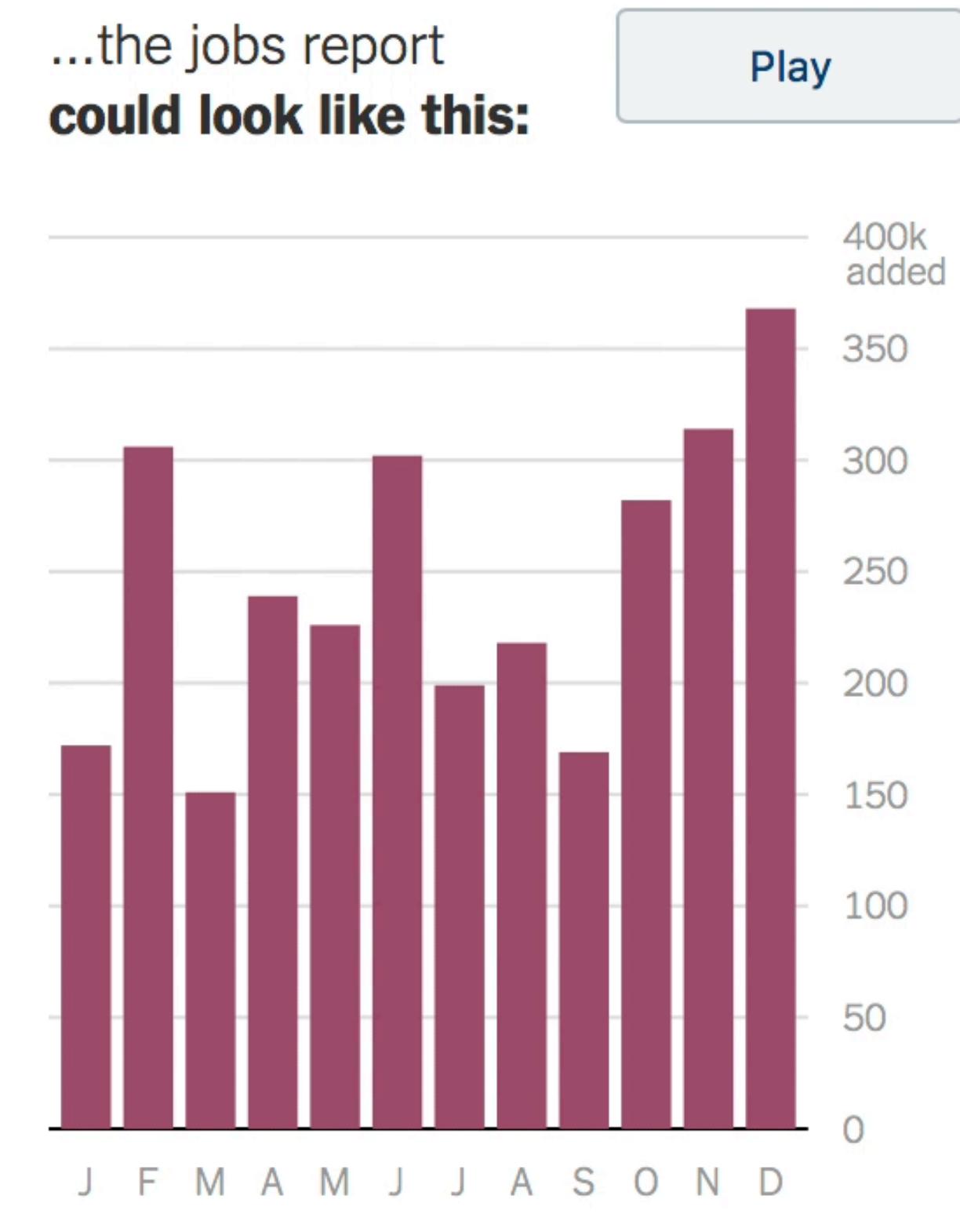
...the jobs report
could look like this:



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



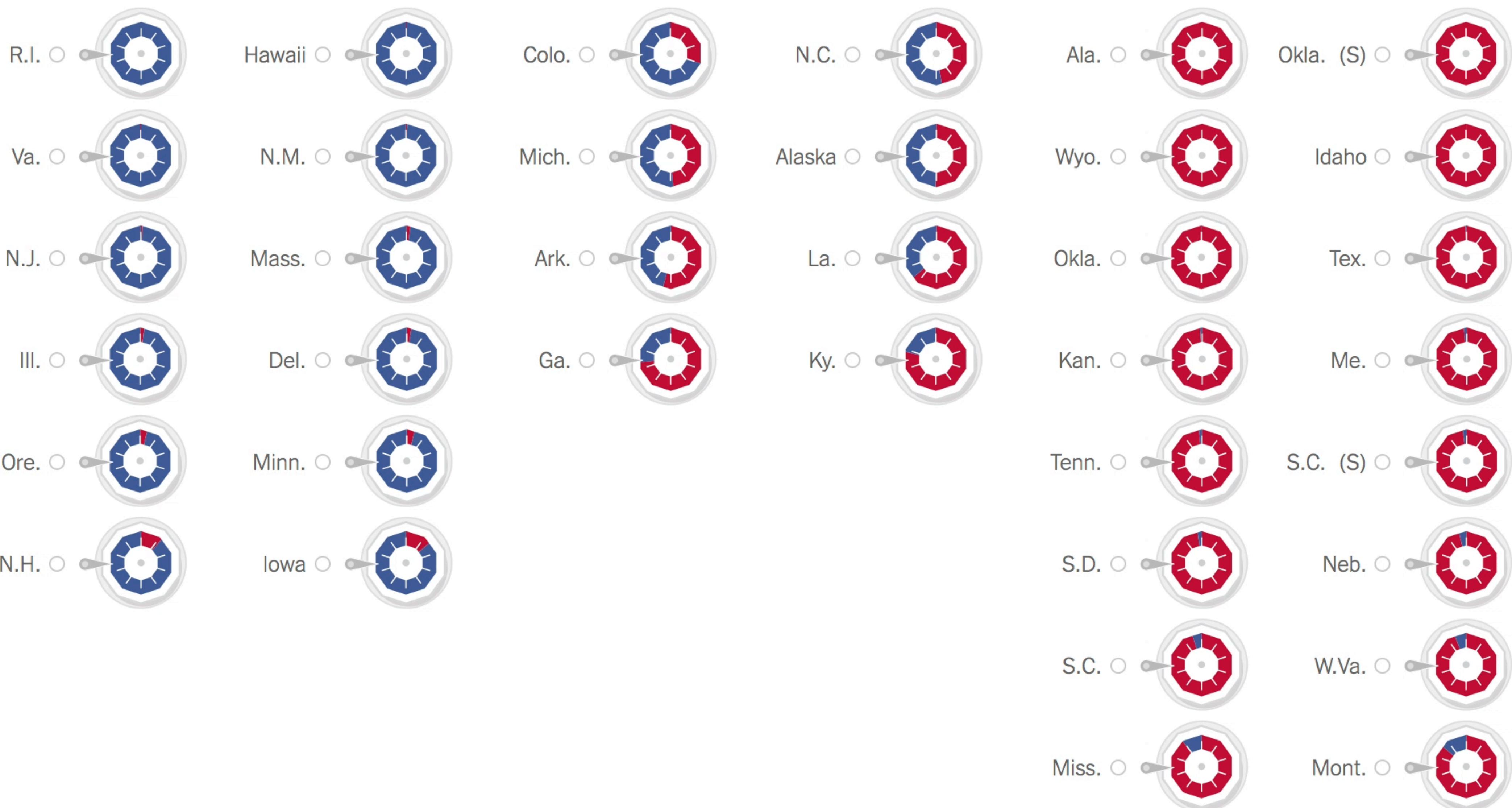
...the jobs report
could look like this:



Likely Democratic

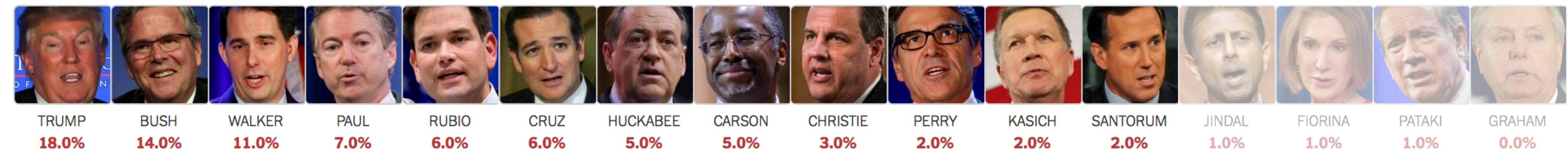
Competitive

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!