Uncertainty

DSC 106: Data Visualization

Sam Lau

UC San Diego

Announcements

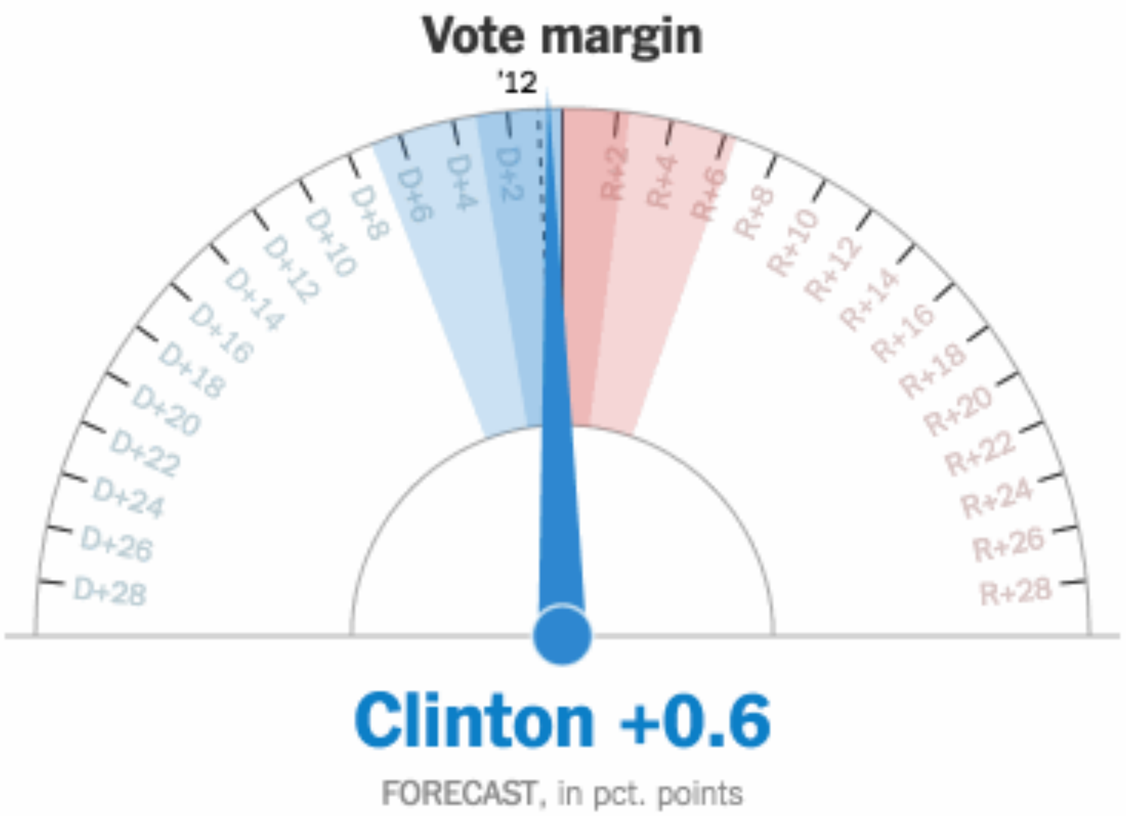
Lab 7 due **next week Tuesday** (not Friday anymore!).

Final project proposal (and teams) due next week Tuesday.

No lectures next week since Sam is traveling (only need to attend discussion for attendance).

FAQs:

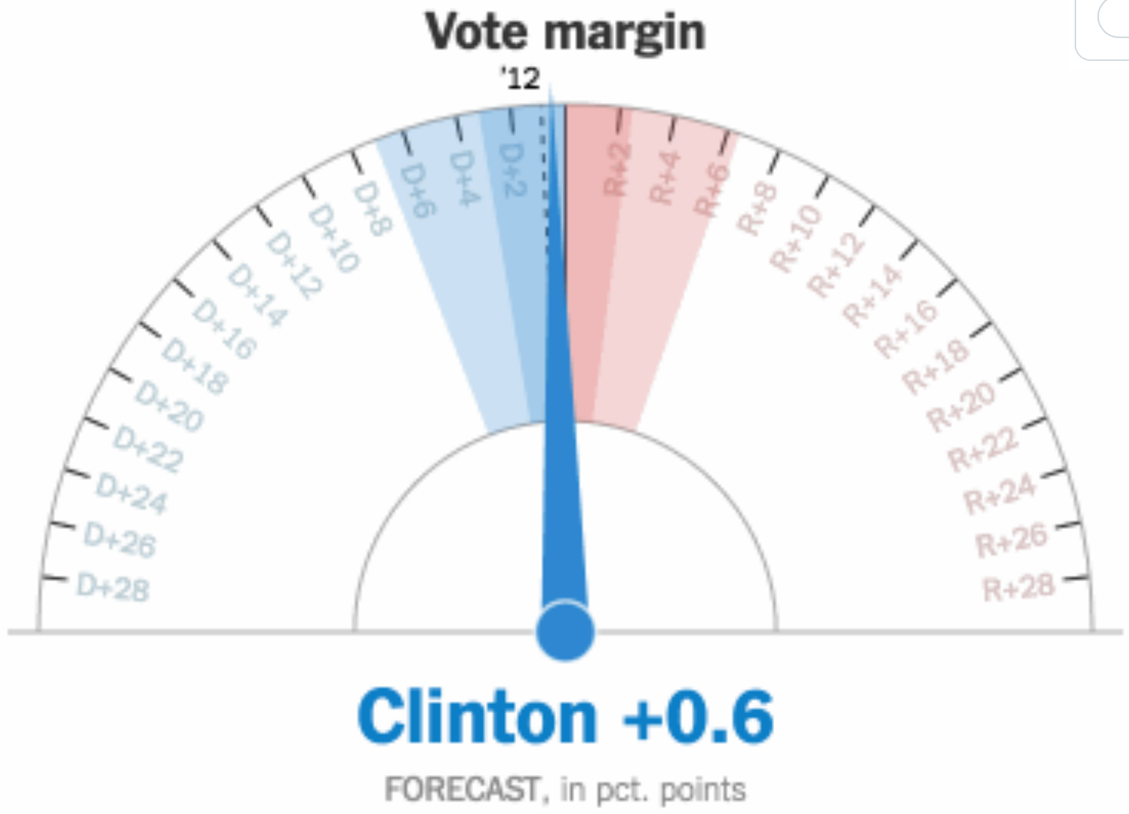
1.



What is being visualized?

What are the strengths and weaknesses of this visualization?

tryclassbuzz.com
Code: **needle**



(gobsmacked)
@nv... Follow

I'm taking
damn NY

10:08 PM ·

2

Danielle Rosen
@Daniel_Rosen · **Follow**

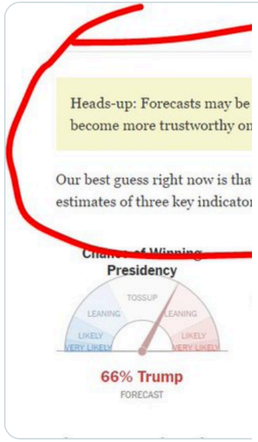
The NYT needle jittering is going to be in my nightmares no matter what happens tonight.

10:00 PM · Nov 8, 2016

2

Paul D. Shinkman
@PDShinkman · **Follow**

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

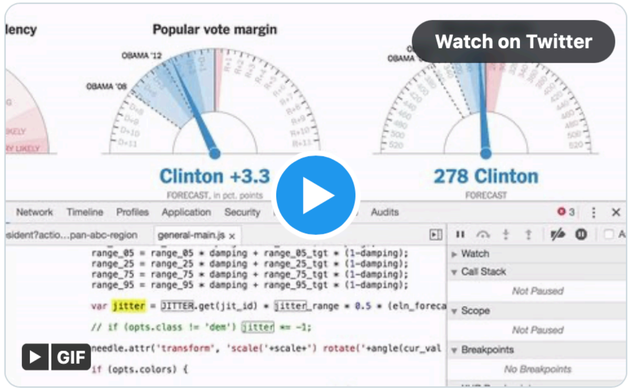


10:03 PM · Nov 8, 2016

Reply 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



@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

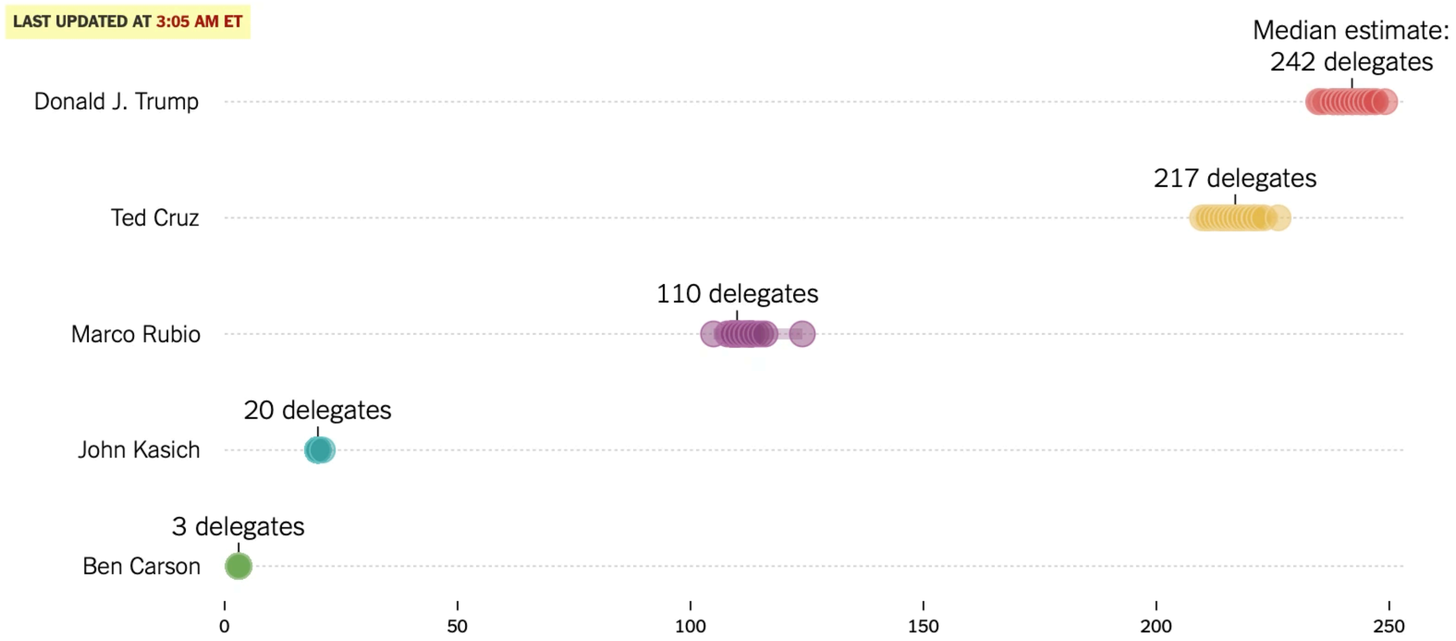
33 Reply Share

Read 5 replies

TheUpshot Live Models

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



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?

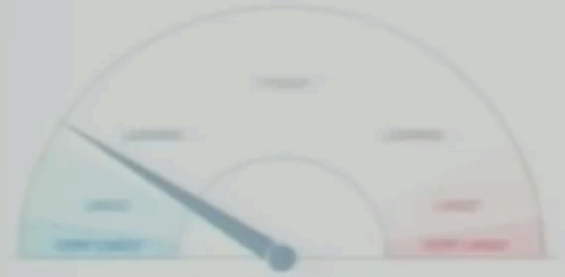
What are the strengths and weaknesses of this visualization?

How does it compare to the needle?

Live Presidential Forecast

Updated 7:55:14 PM ET

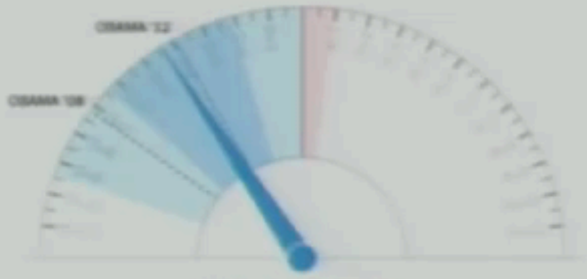
Chance of Winning Presidency



82% Clinton

FORECAST

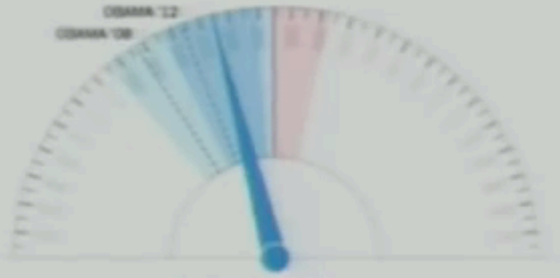
Popular vote margin



Clinton +4.4

FORECAST

Electoral votes



307 Clinton

FORECAST

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 votes	Reported margin	NYT projection	NYT win prob.
Michigan	0%	-	Clinton +6.5	75% Dem.
New Mexico	0%	-	Clinton +6.4	76% 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?

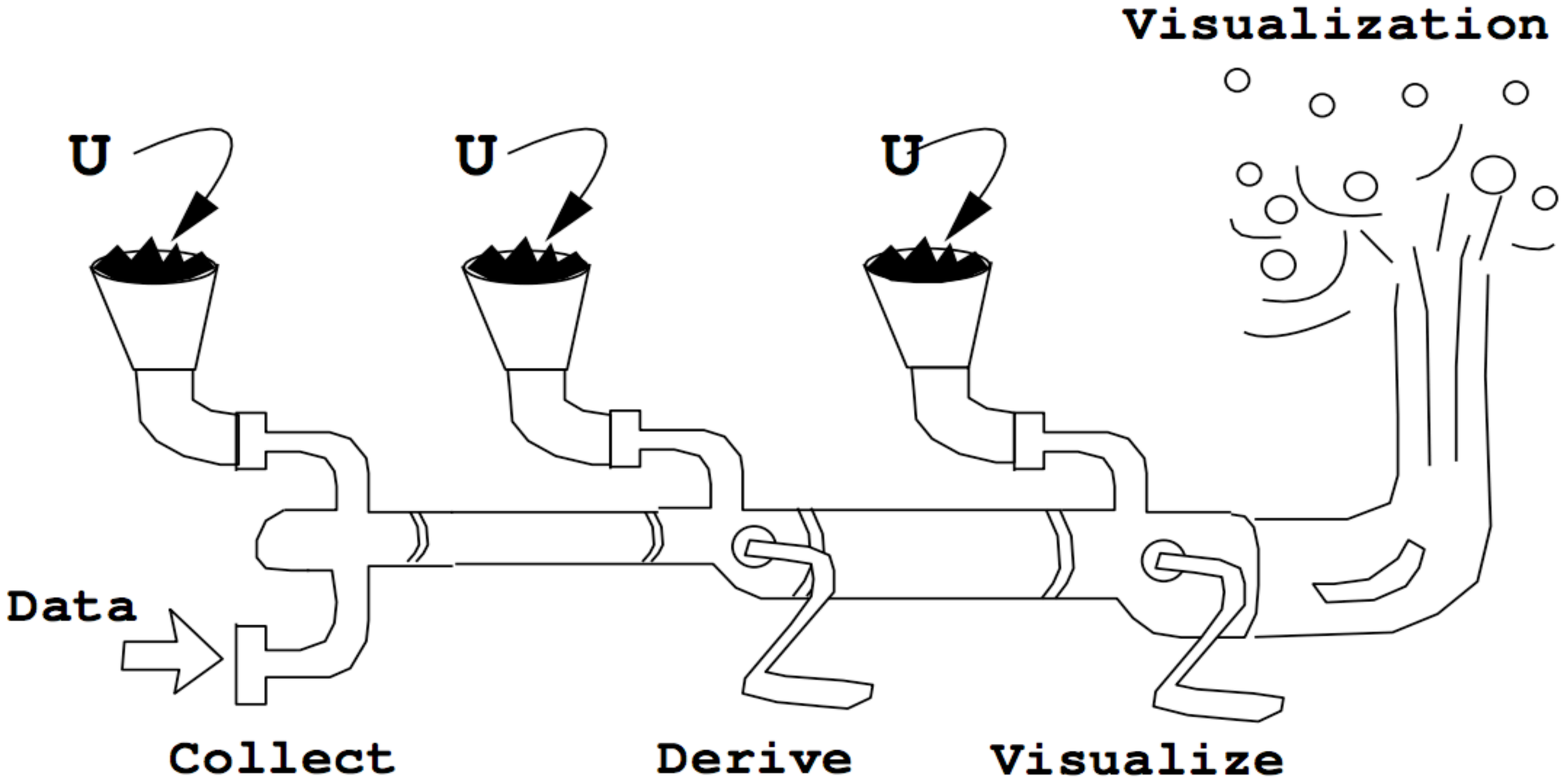
Uncertainty

What does it mean?

How should I visualize it?

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

Sources and Types of Uncertainty

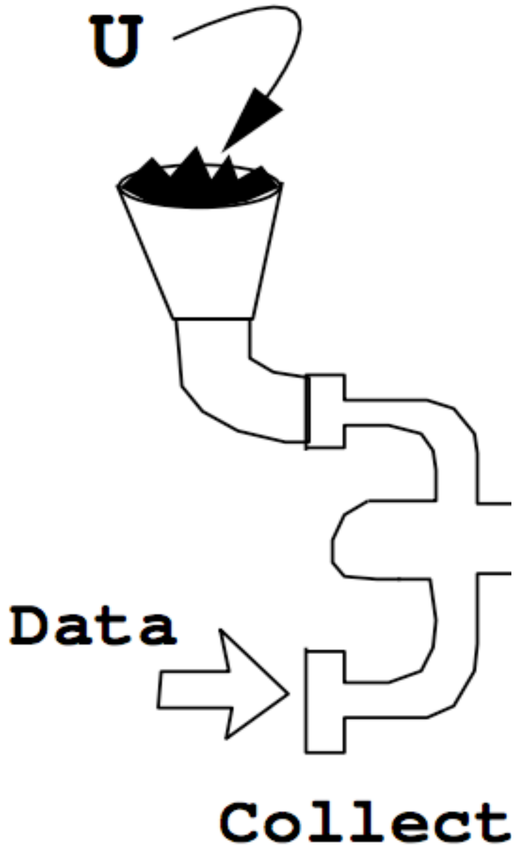


Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Sources and Types of Uncertainty

Measurement Uncertainty

How and how much should we sample the data?



Precision

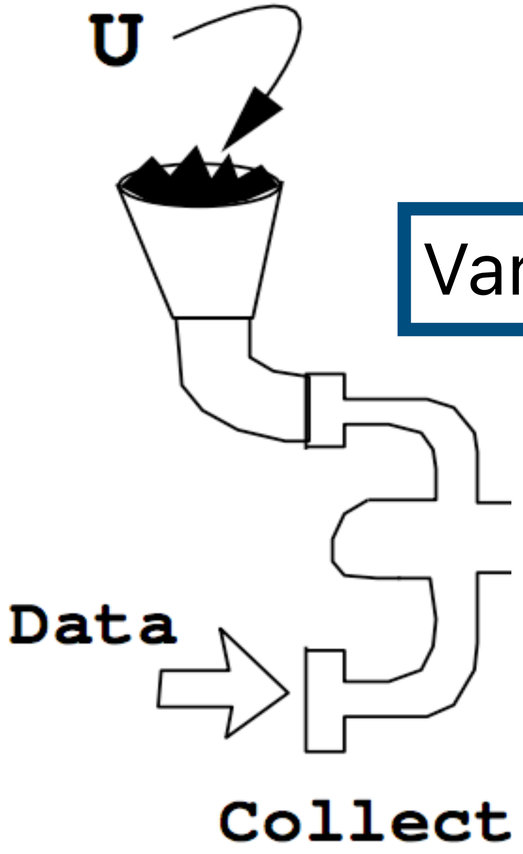
Accuracy

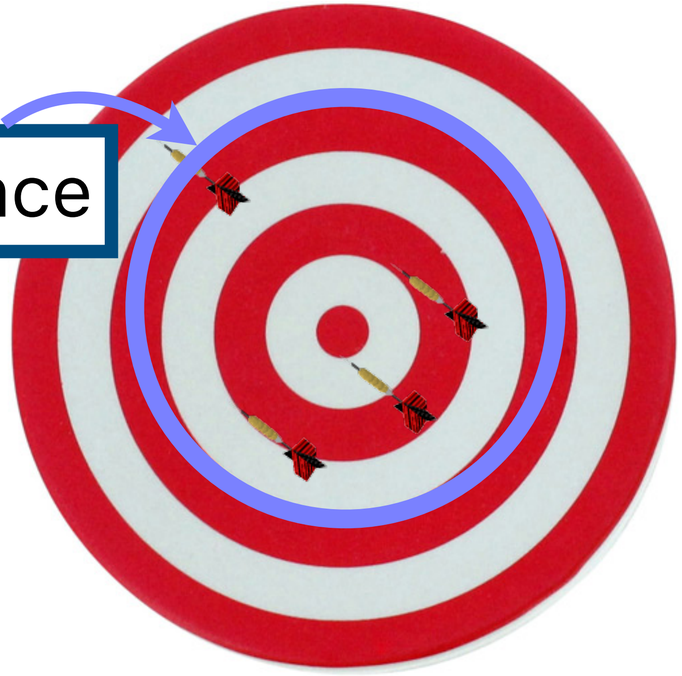
Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Sources and Types of Uncertainty

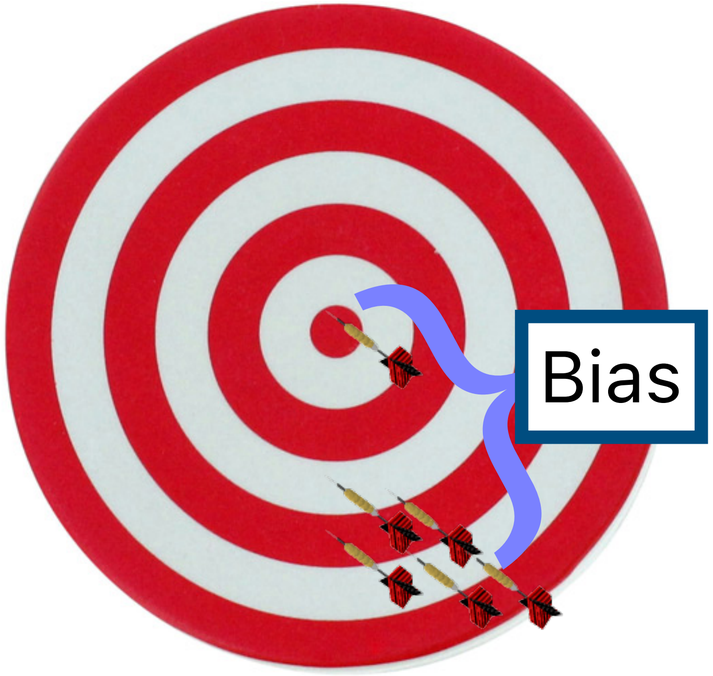
Measurement Uncertainty

How and how much should we sample the data?





Precision



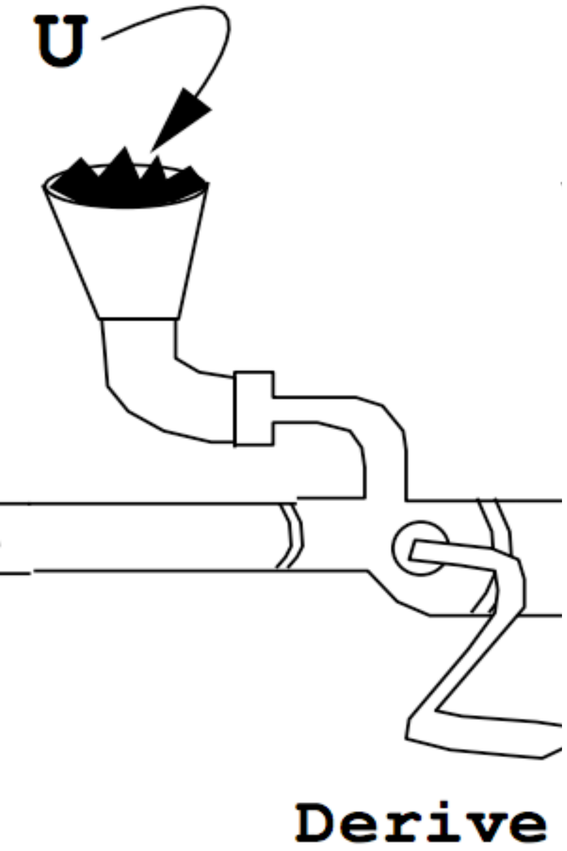
Accuracy

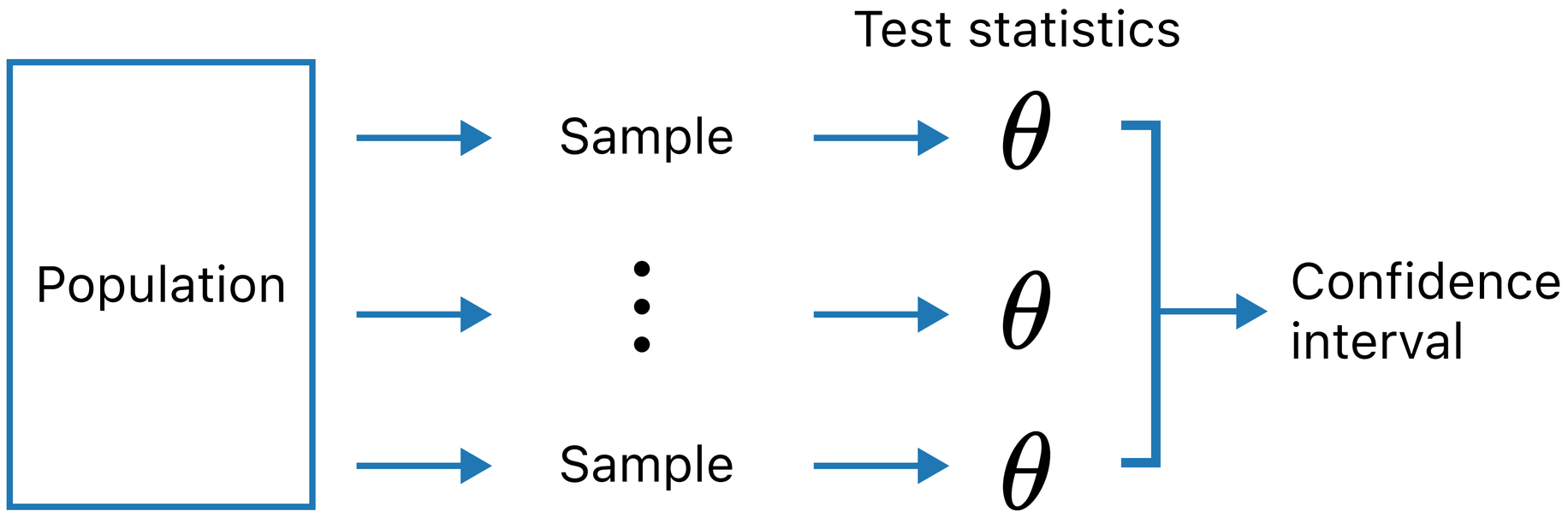
Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Sources and Types of Uncertainty

Model Uncertainty

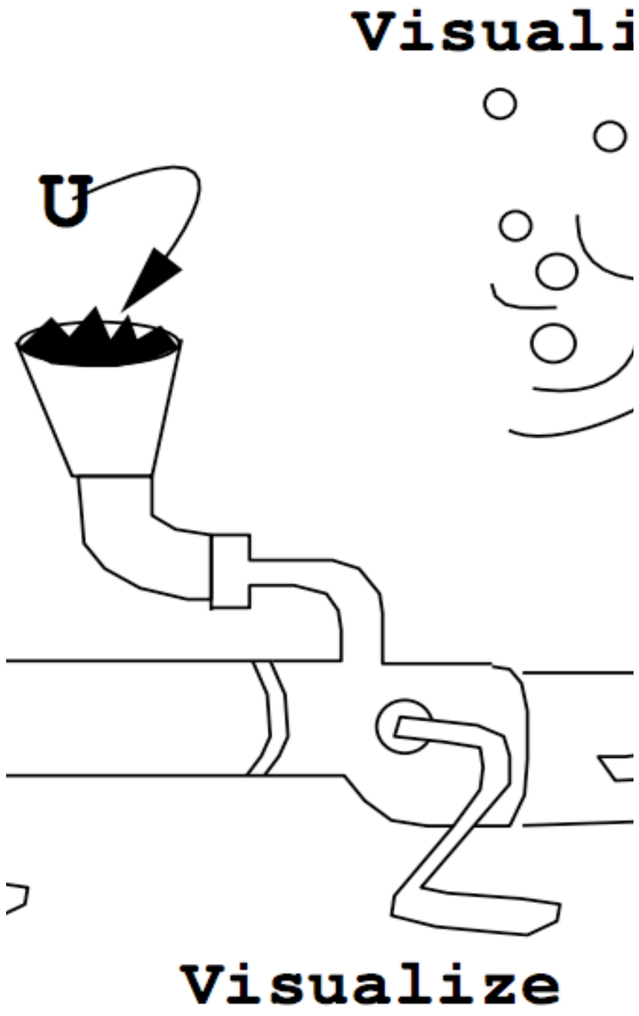
How does the data fit together?





Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

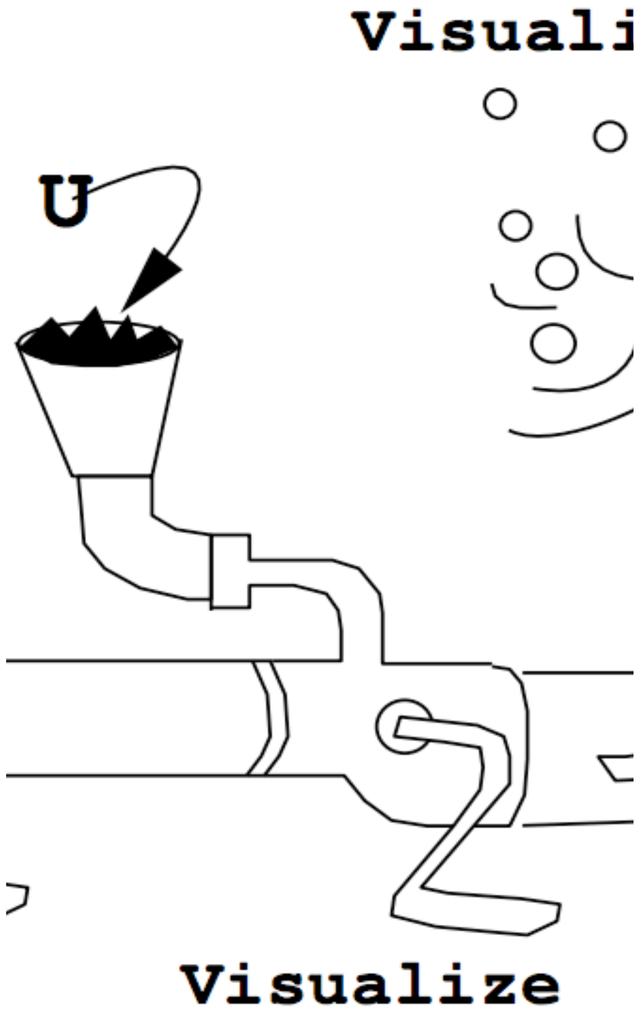
Sources and Types of Uncertainty



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

Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

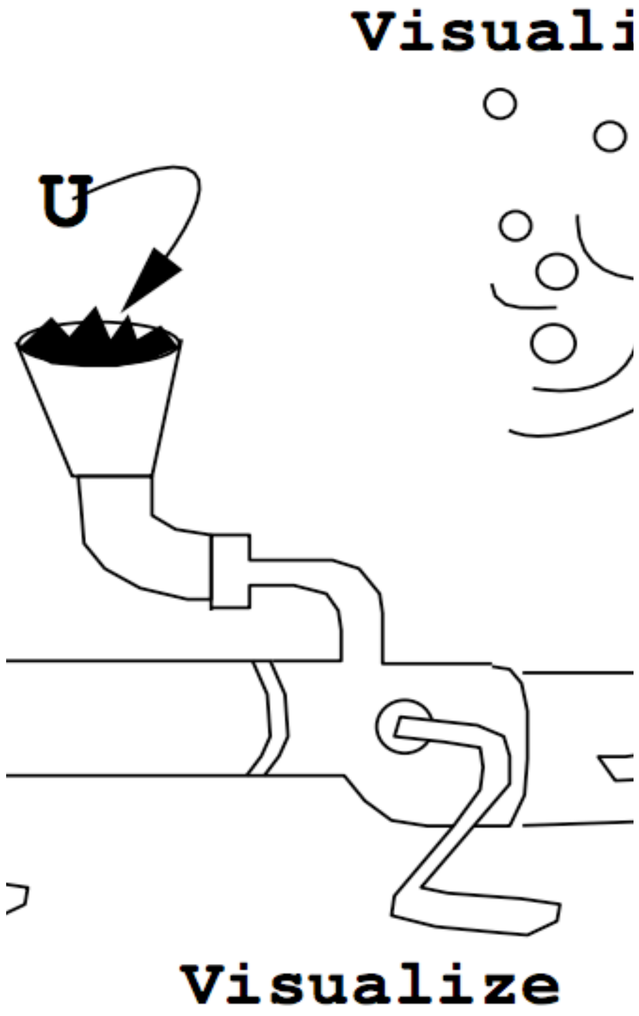
Sources and Types of Uncertainty



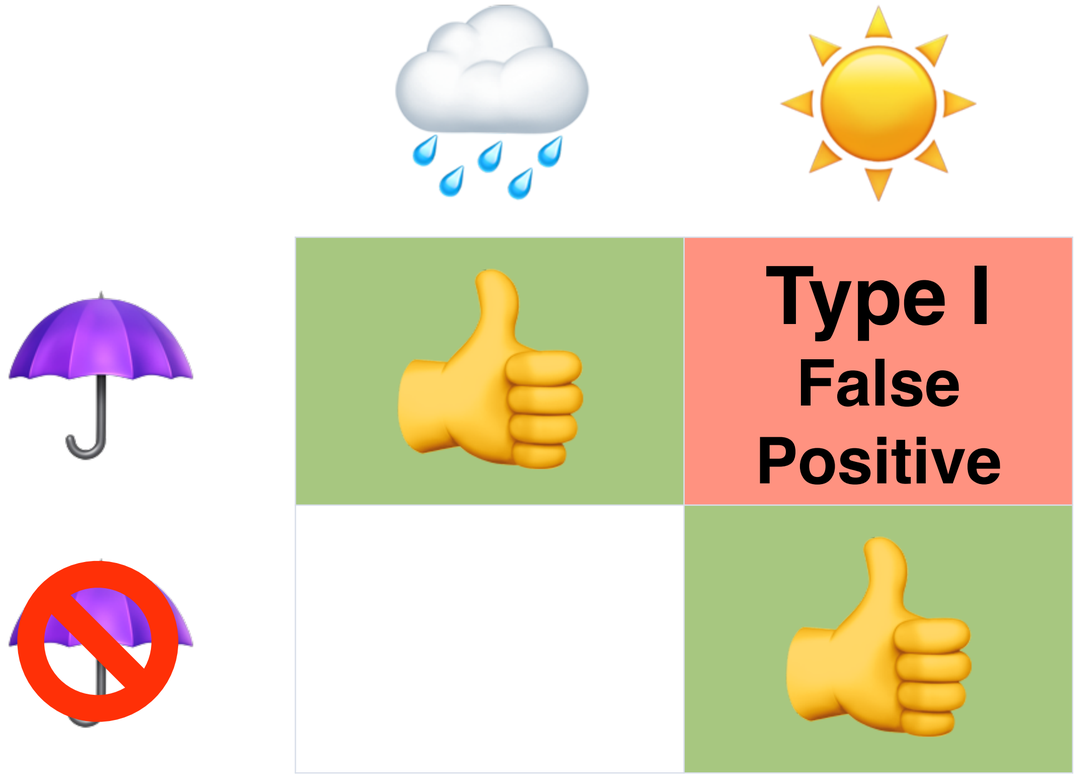
Decision/Forecast Uncertainty
How do I assess the risk or error?

Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Sources and Types of Uncertainty

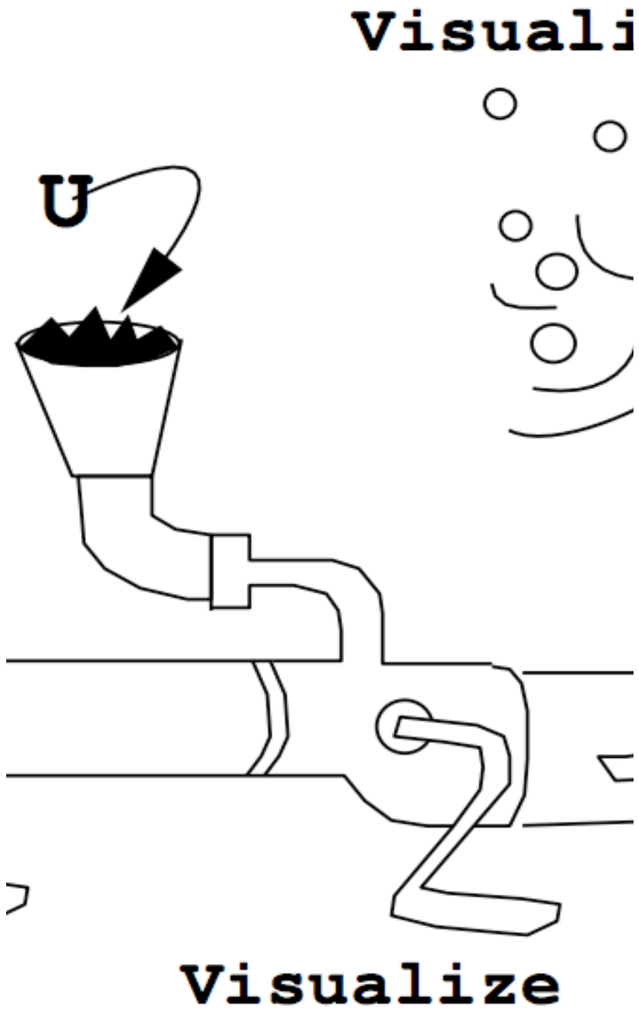


Decision/Forecast Uncertainty
How do I assess the risk or error?

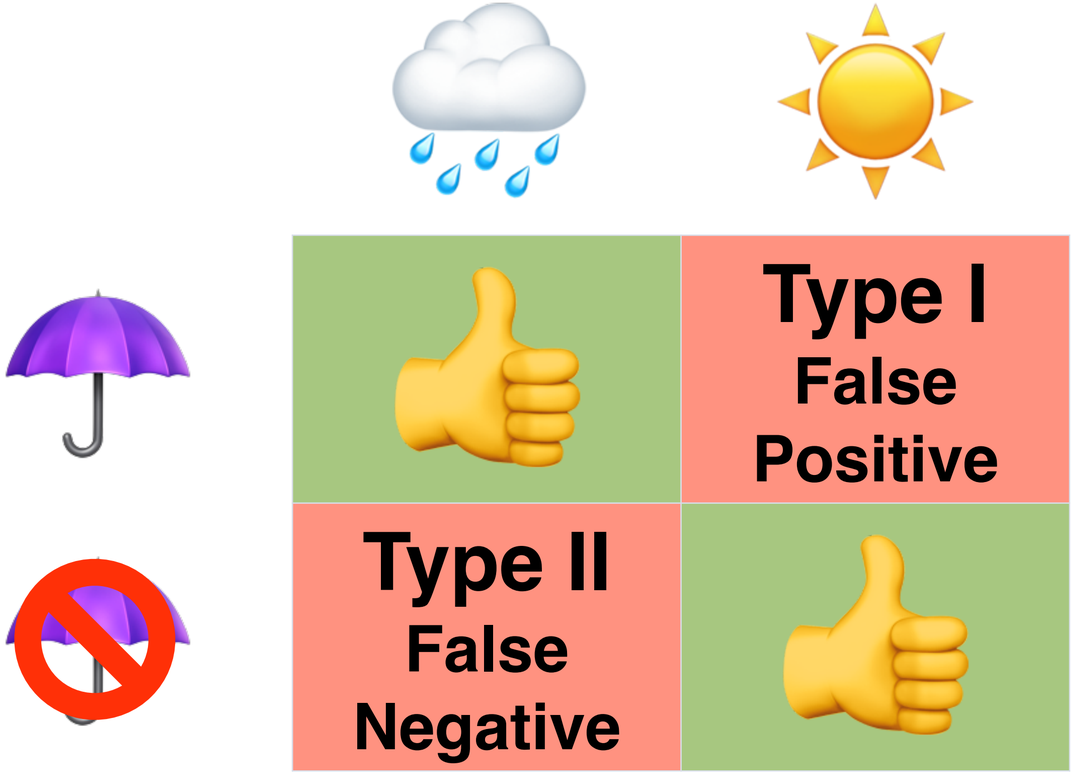


Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Sources and Types of Uncertainty



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



Pang, Wittenbrink, Lodha. Approaches to Uncertainty Visualization, 1997.

Uncertainty

What does it mean?

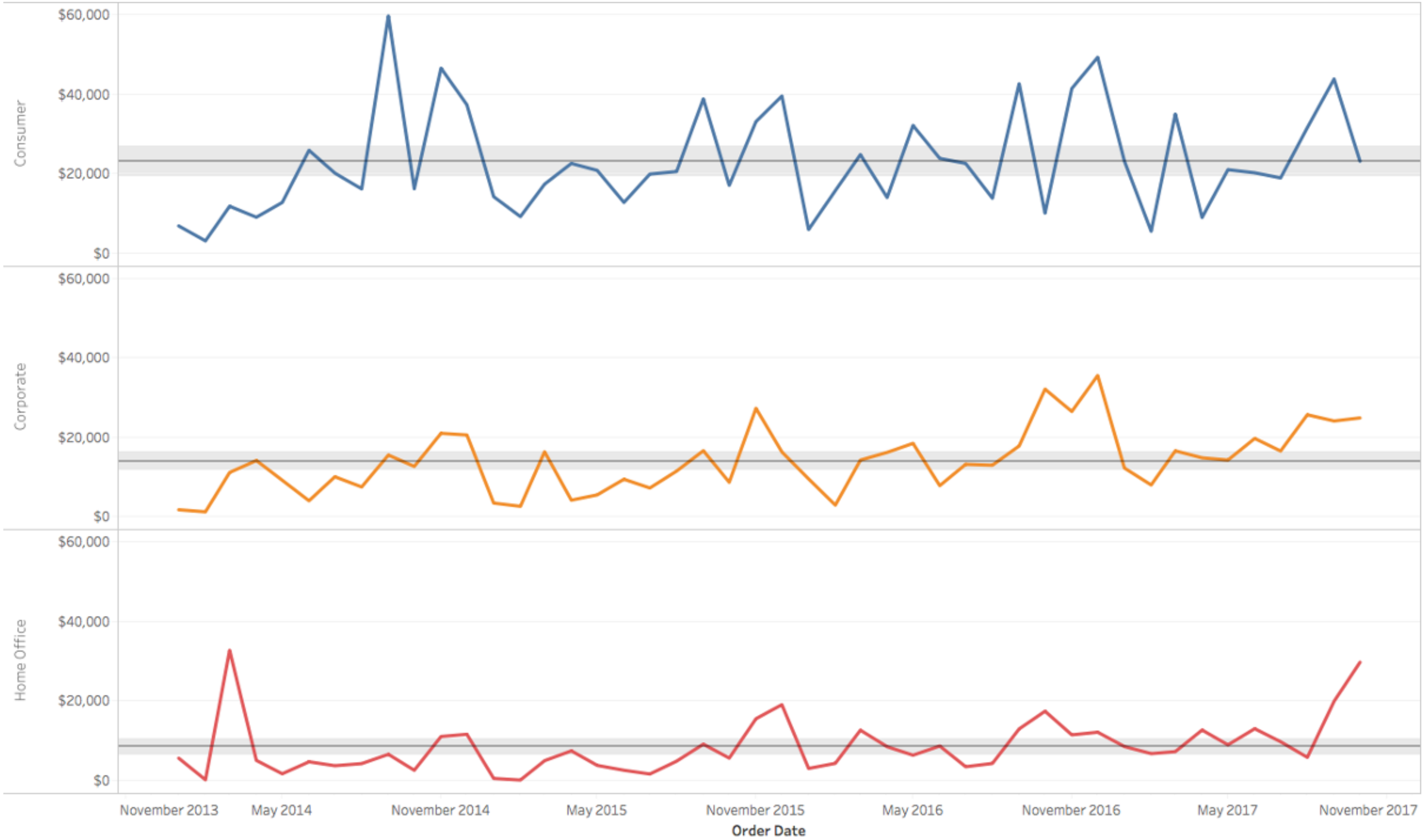
Lots of things!

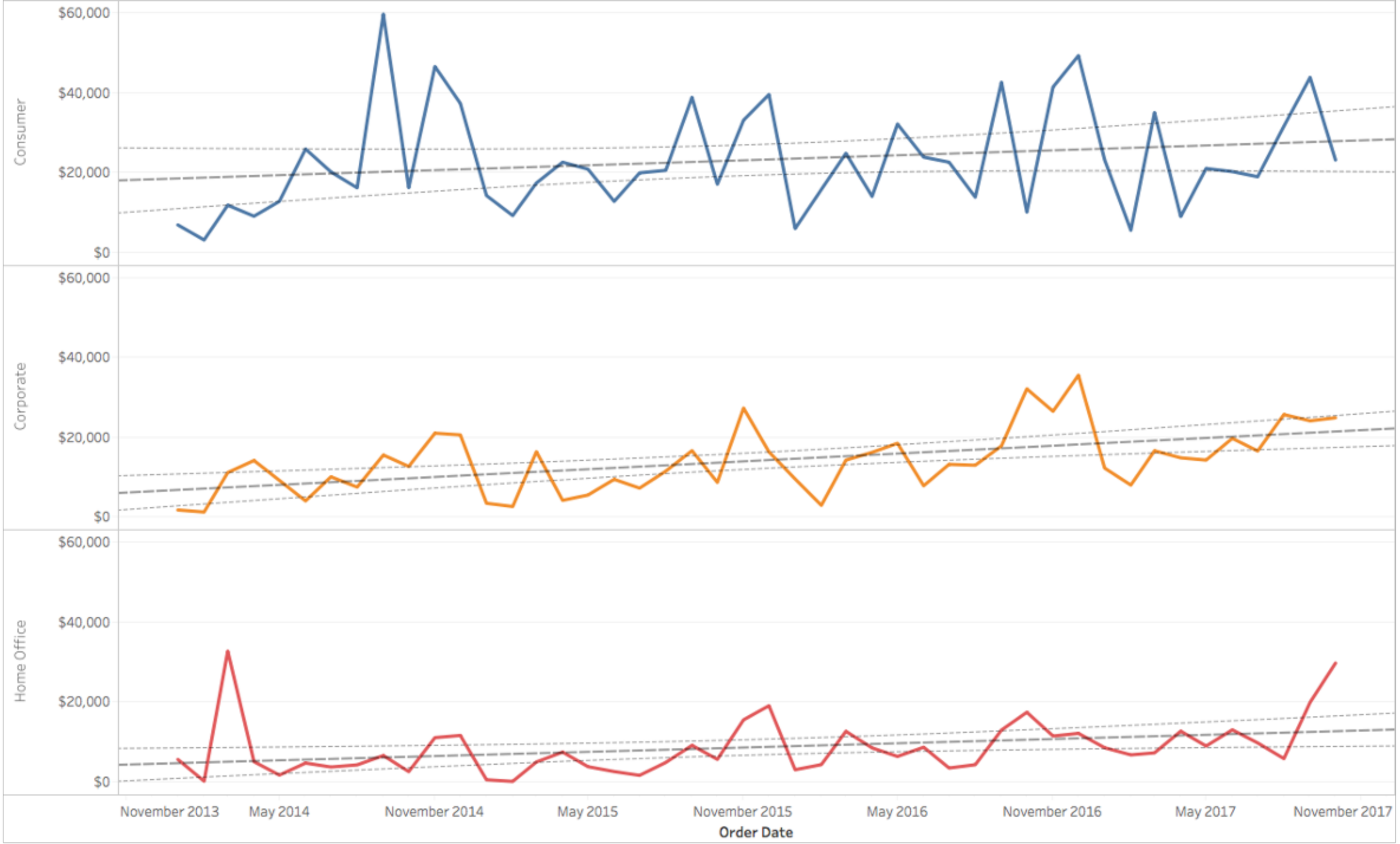
How should I visualize it?

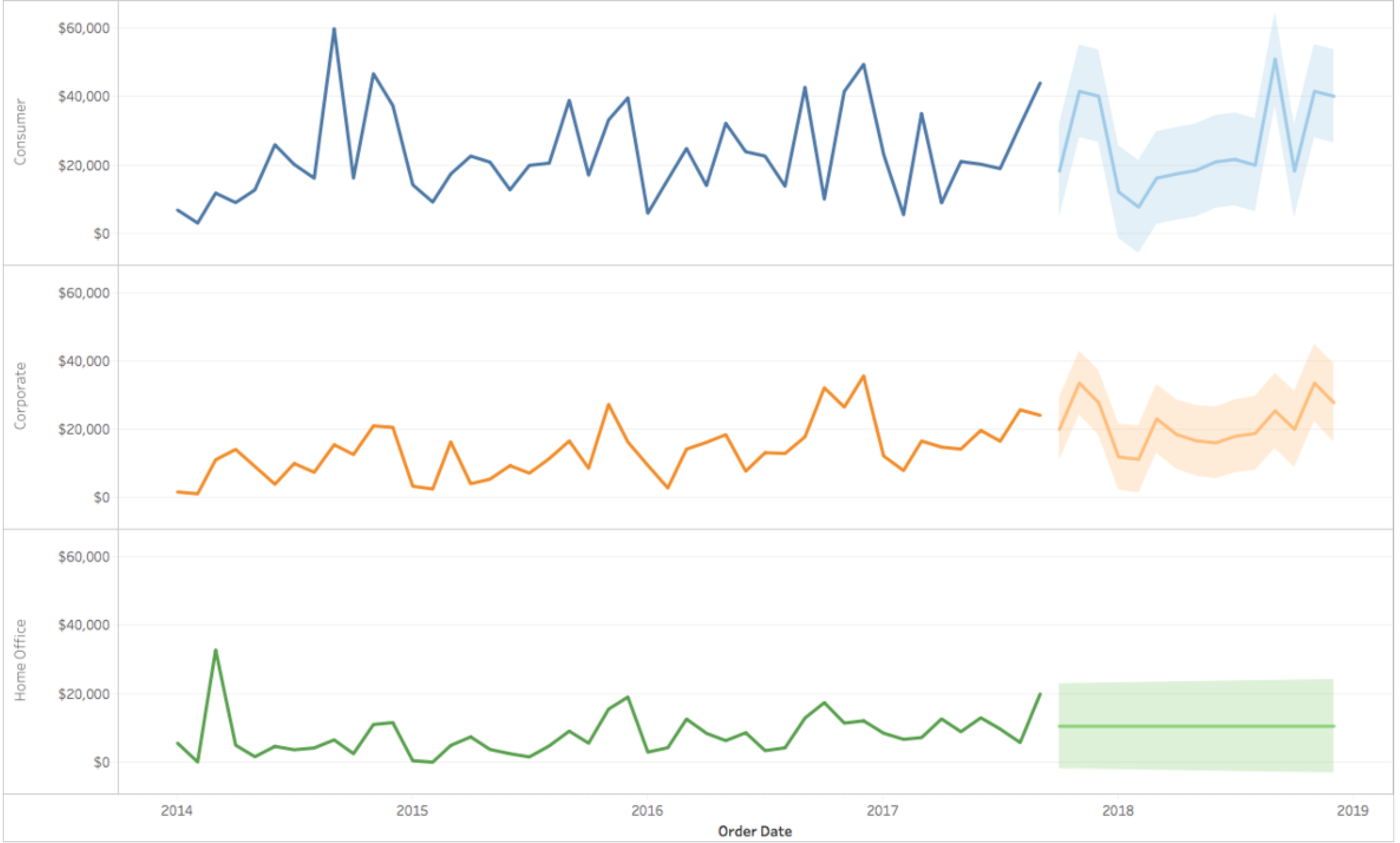
Uncertainty

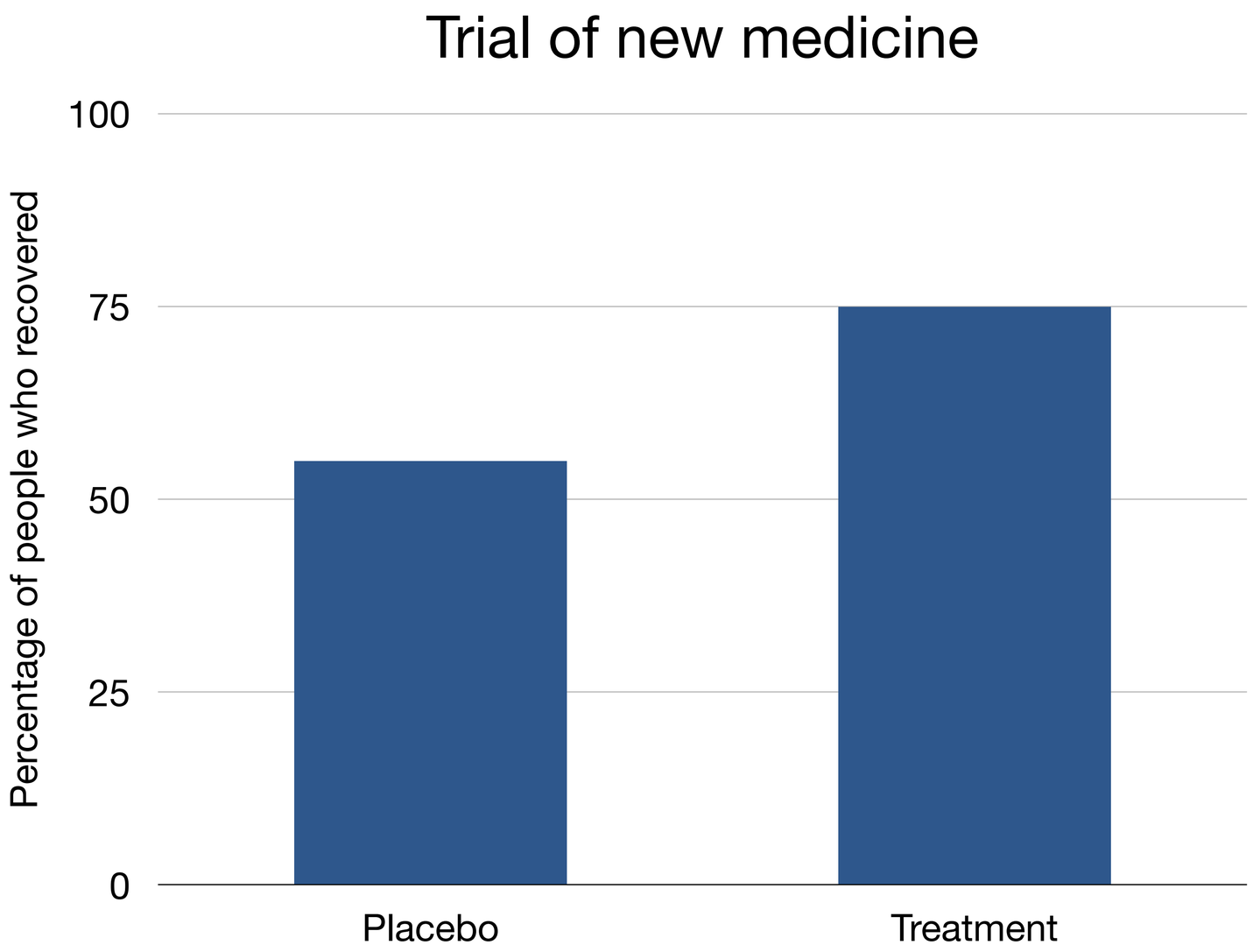
What does it mean?

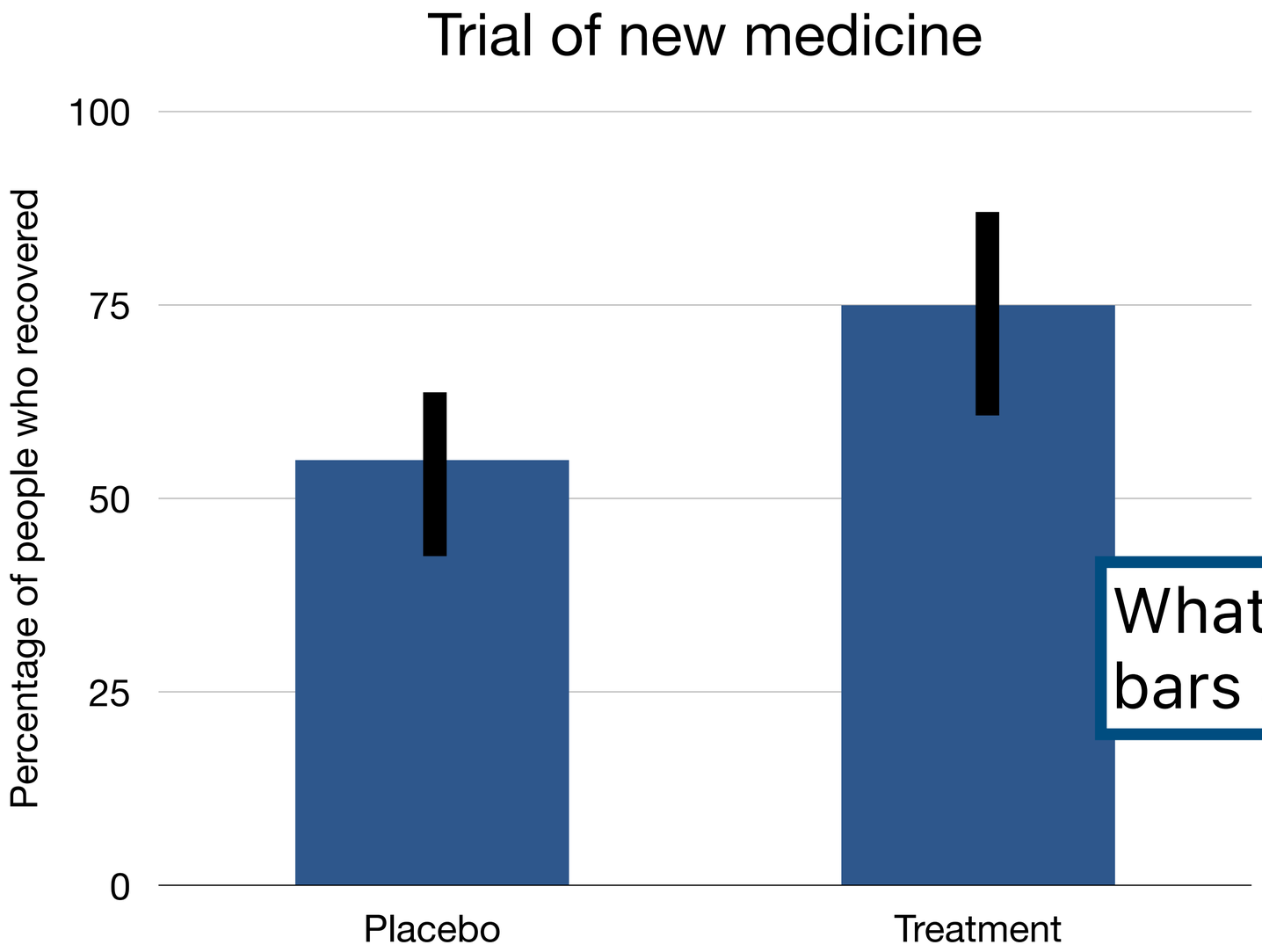
How should I visualize it?


Consumer
$60,000
$40,000
$20,000
$0
Corporate
$60,000
$40,000
$20,000
$0
Home Office
$60,000
$40,000
$20,000
$0
November 2013
May 2014
November 2014
May 2015
November 2015
May 2016
November 2016
May 2017
November 2017
Order Date


Consumer
$60,000
$40,000
$20,000
$0
Corporate
$60,000
$40,000
$20,000
$0
Home Office
$60,000
$40,000
$20,000
$0
November 2013
May 2014
November 2014
May 2015
November 2015
May 2016
November 2016
May 2017
November 2017
Order Date

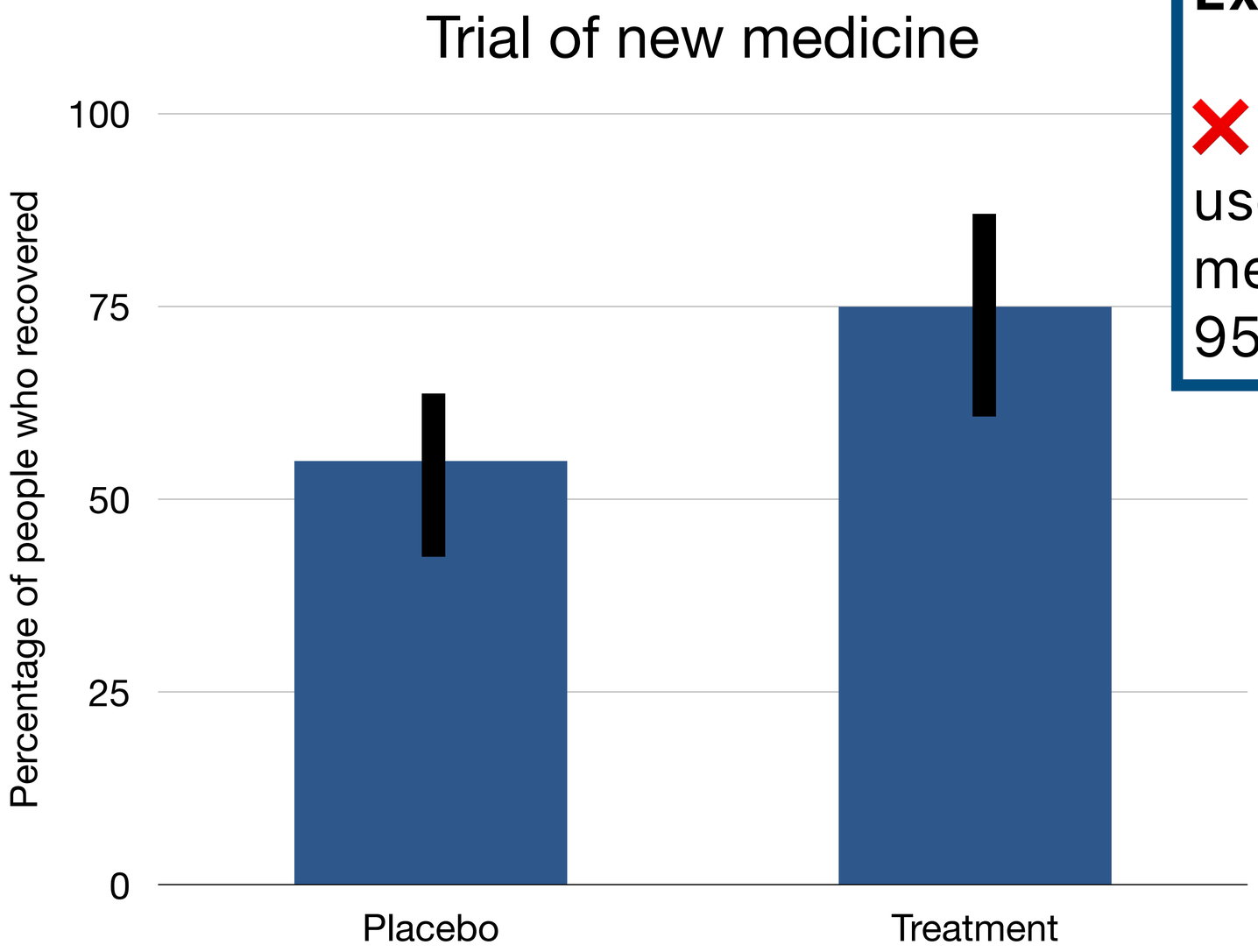

Consumer
$60,000
$40,000
$20,000
$0
Corporate
$60,000
$40,000
$20,000
$0
Home Office
$60,000
$40,000
$20,000
$0
2014
2015
2016
2017
2018
2019
Order Date


Trial of new medicine
Percentage of people who recovered
100
75
50
25
0
Placebo
Treatment



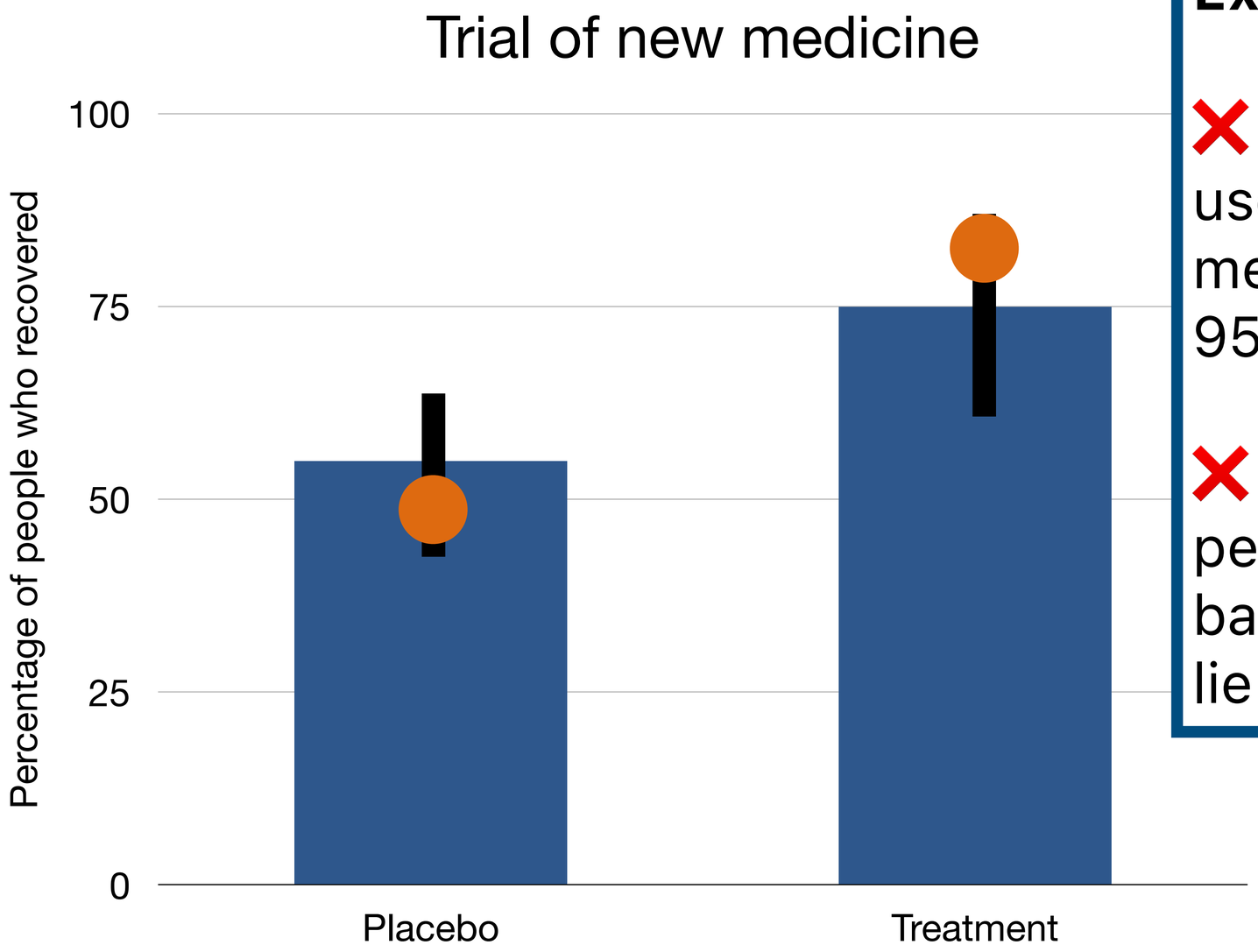
What do you think these error bars are implying?

tryclassbuzz.com
Code: **errorbars**



Expressiveness?

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

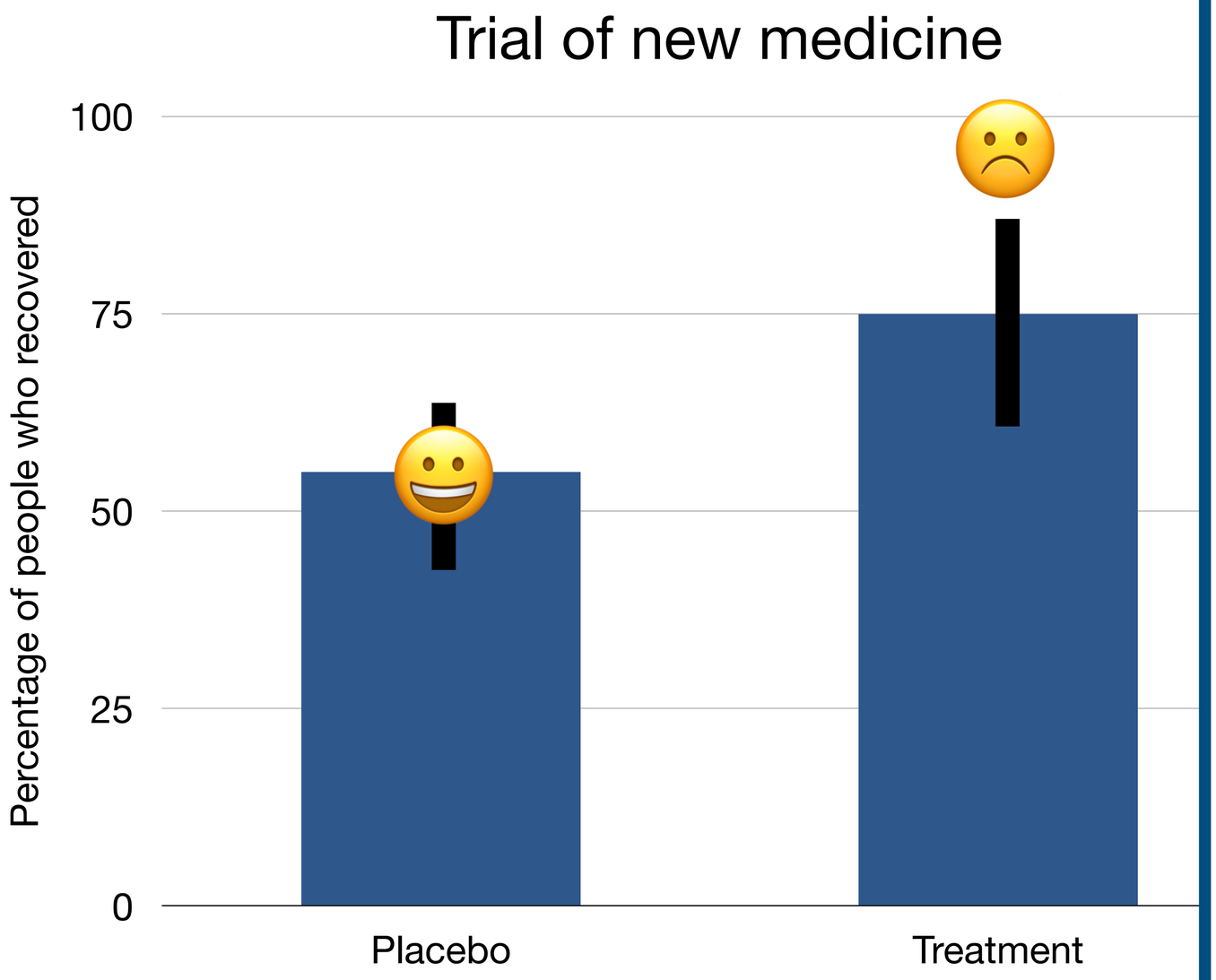


Expressiveness?

❌ 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.

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



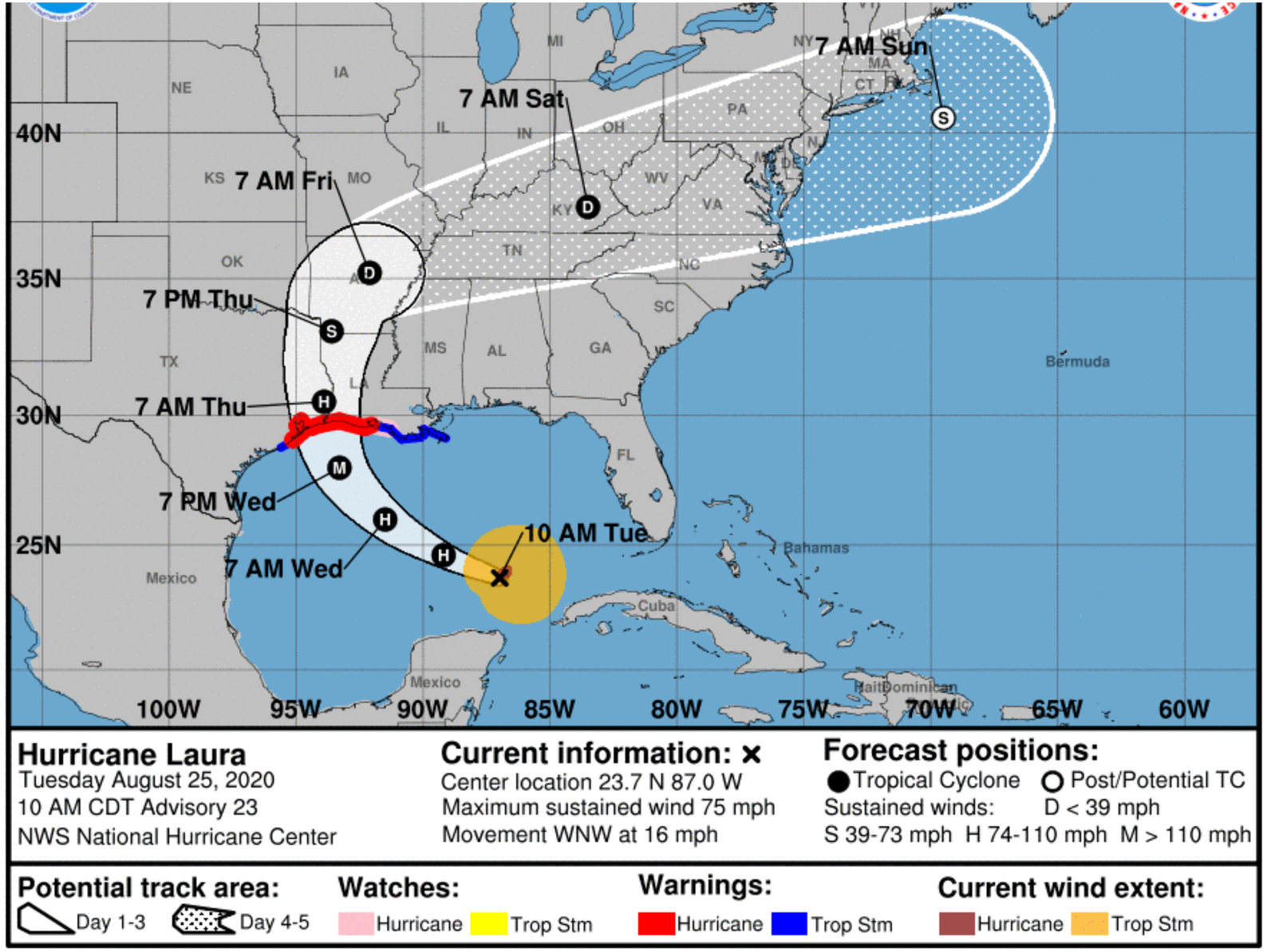
[Newman & Scholl, 2012]
[Correll & Gleicher, 2014]

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.

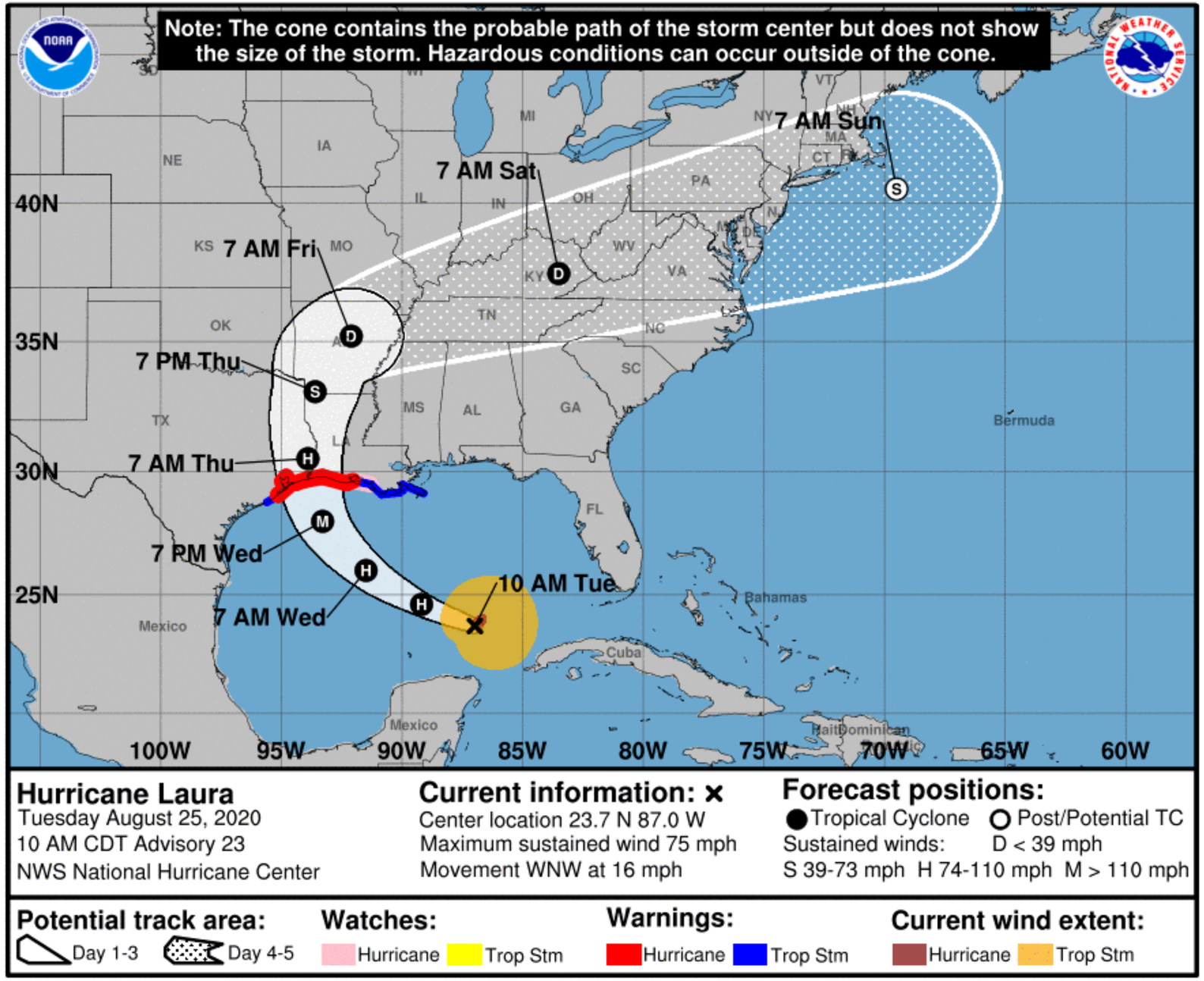
❌ Binary bias: people perceive values to either be in or out of the margins of error.



What is being visualized?

What are the strengths and weaknesses of this visualization?

tryclassbuzz.com
Code: **hurricane**



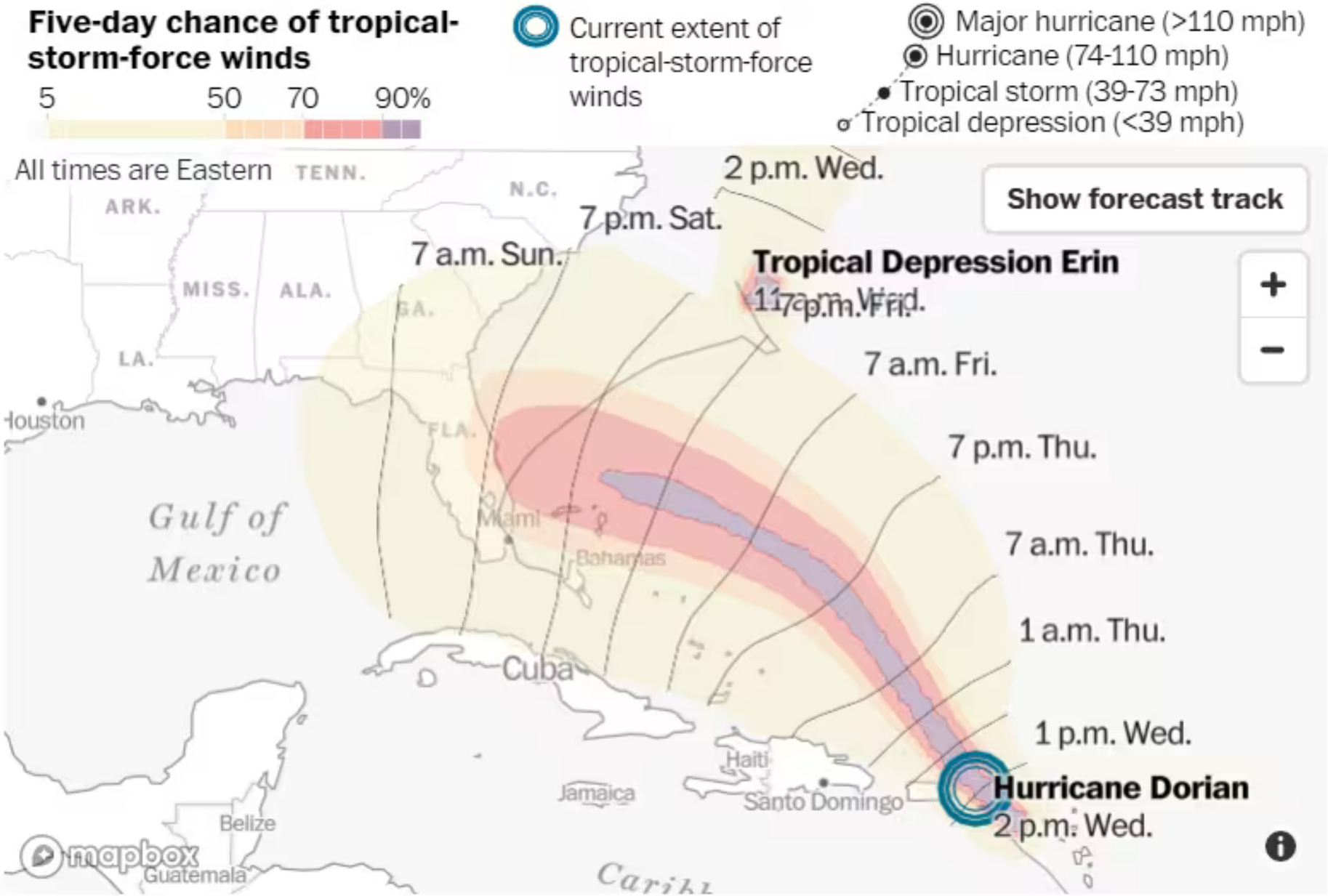
What is being visualized?

What are the strengths and weaknesses of this visualization?


FOX4 HURRICANE IAN
5:00 AM ADVISORY
SAT 2:00 AM
40 MPH
FRI 2:00 AM
90 MPH
THU 2:00 AM
115 MPH
WED 2:00 PM
130 MPH
WED 2:00 AM
140 MPH
FORT MYERS
NAPLES
CURRENT LOCATION: 90 MI SW OF G

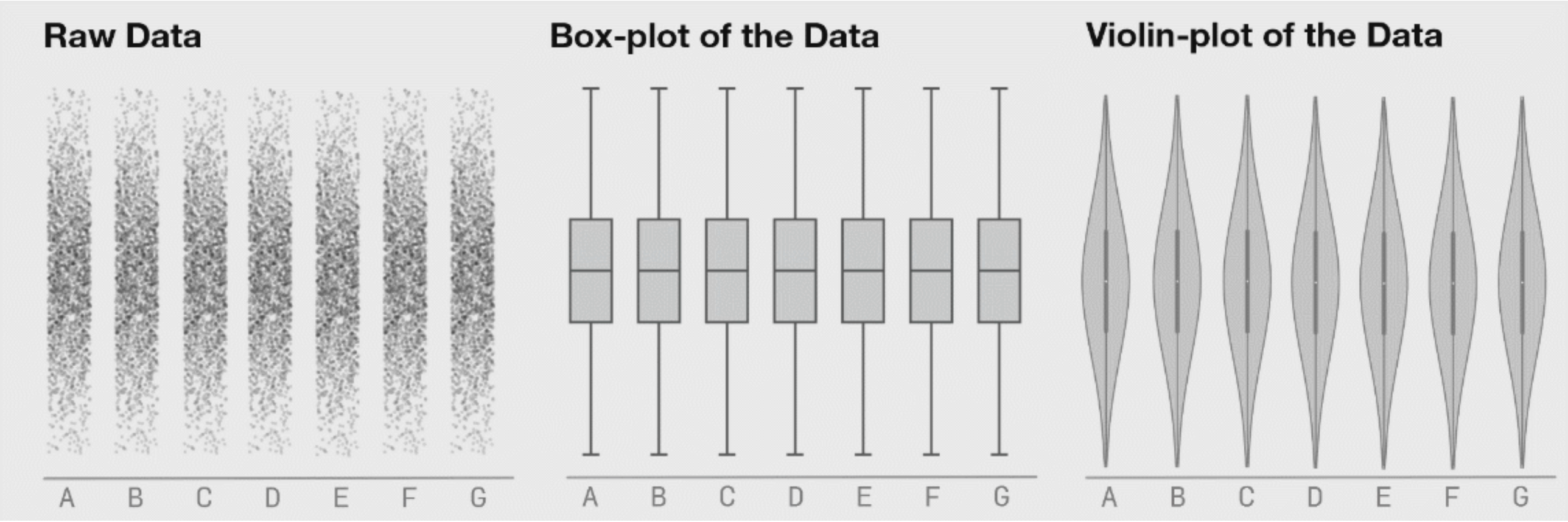

Rapid City
Omaha
St. Louis
Cincinnati
Washington
New York
FIRST WARNING WEATHER
HURRICANE DORIAN
Atlanta
Jacksonville
Wednesday 8:00 AM EDT
75 mph
Tuesday 8:00 AM EDT
125 mph
Monday 8:00 AM EDT
140 mph
Sunday 8:00 AM EDT
130 mph
Saturday 8:00 AM EDT
125 mph
Location: 24.5°N, 69.8°W
Wind: 110 mph
Gusts: 130 mph
Movement: NW at 10 mph
Pressure: 28.70 in


Hurricane Dorian Forecast Track and Intensity
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.
8 AM Tue
8 AM Sun
8 AM Sat
8 PM Fri
8 AM Fri
8 AM Mon
11 AM Thu
30N
25N
20N
90W
85W
80W
75W
70W
65W
60W
Hurricane Dorian
Thursday August 29, 2019
11 AM AST Advisory 21
NWS National Hurricane Center
Current information:
Center location 21.4 N 67.2 W
Maximum sustained wind 85 mph
Movement NW at 13 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

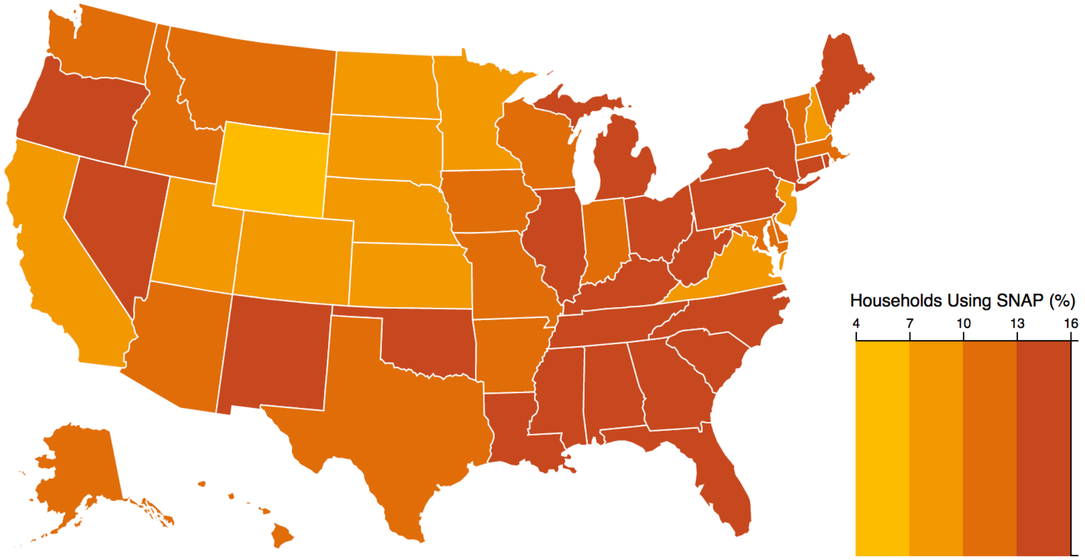


Source: National Weather Service. Note: Impact lines represent the earliest reasonable arrival time of tropical-storm-force winds.

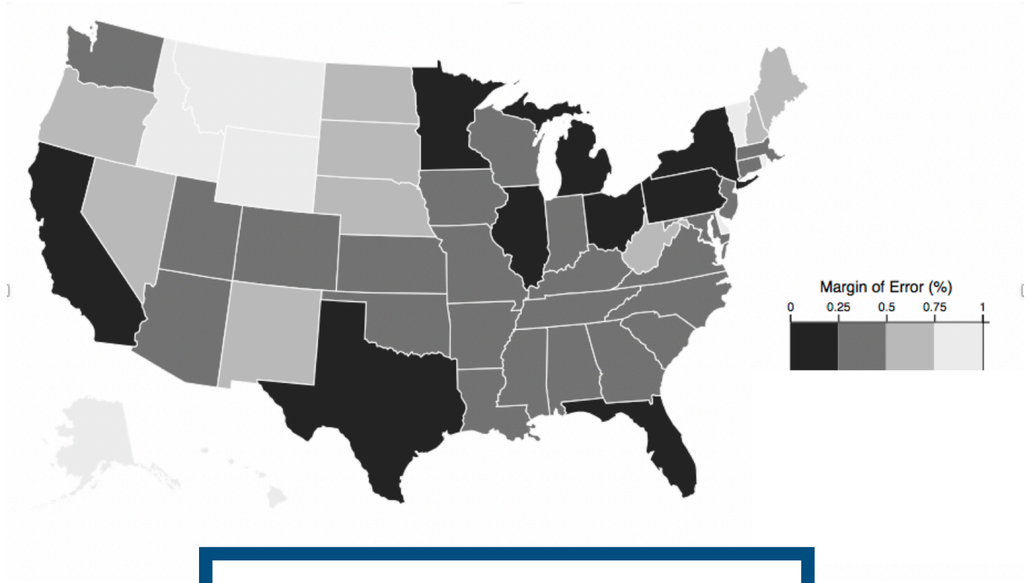
For uncertainty, use **visual variables** instead of visualizing point estimates



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



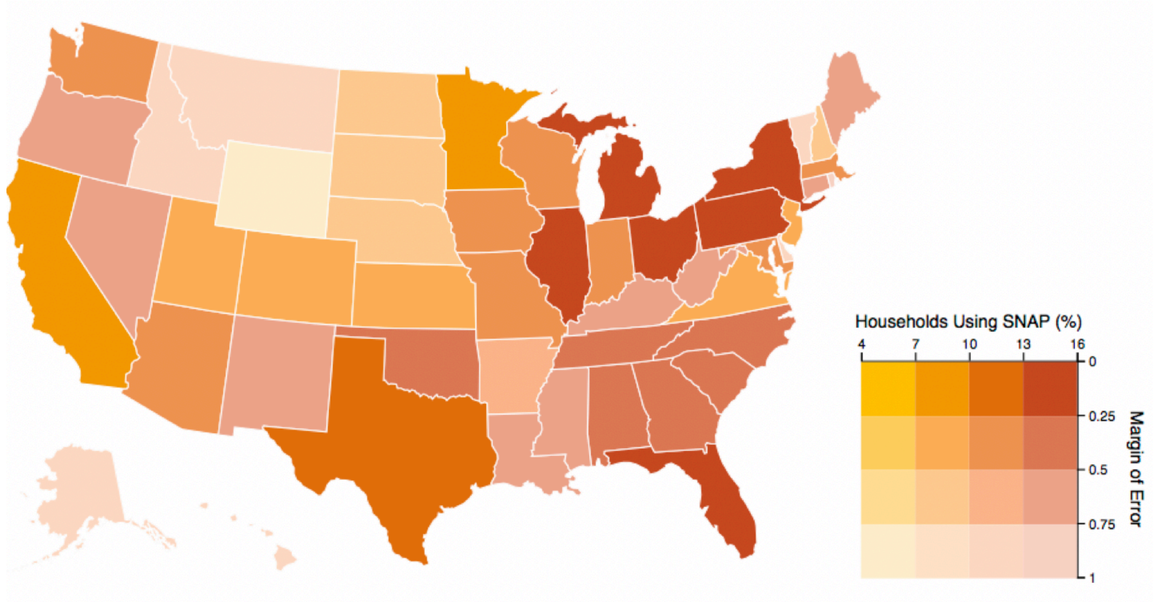
Data Map

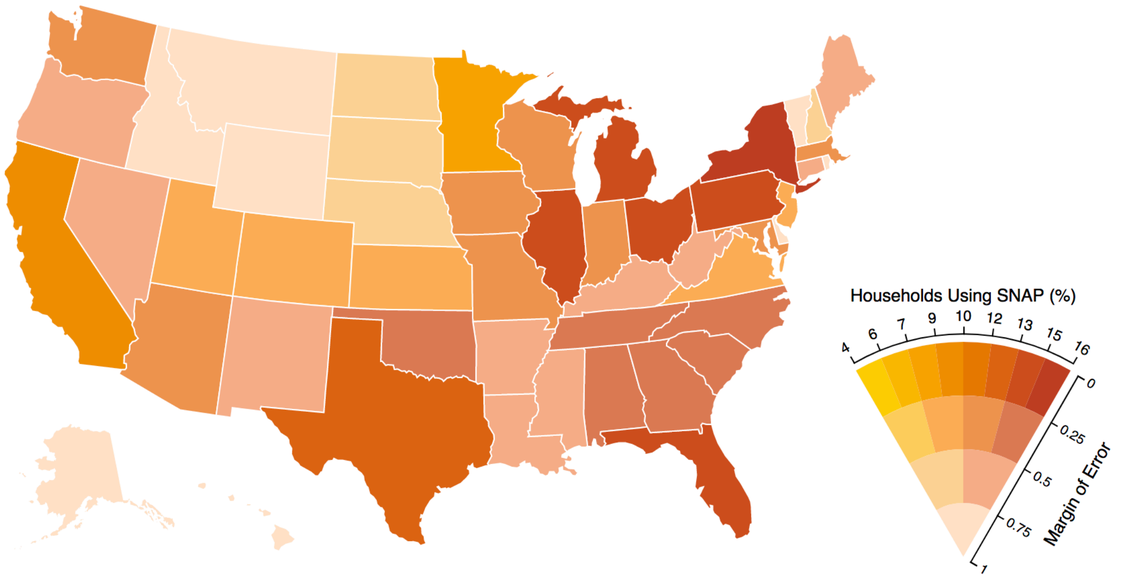


Uncertainty Map

[Correll, Moritz, & Heer, 2018]

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



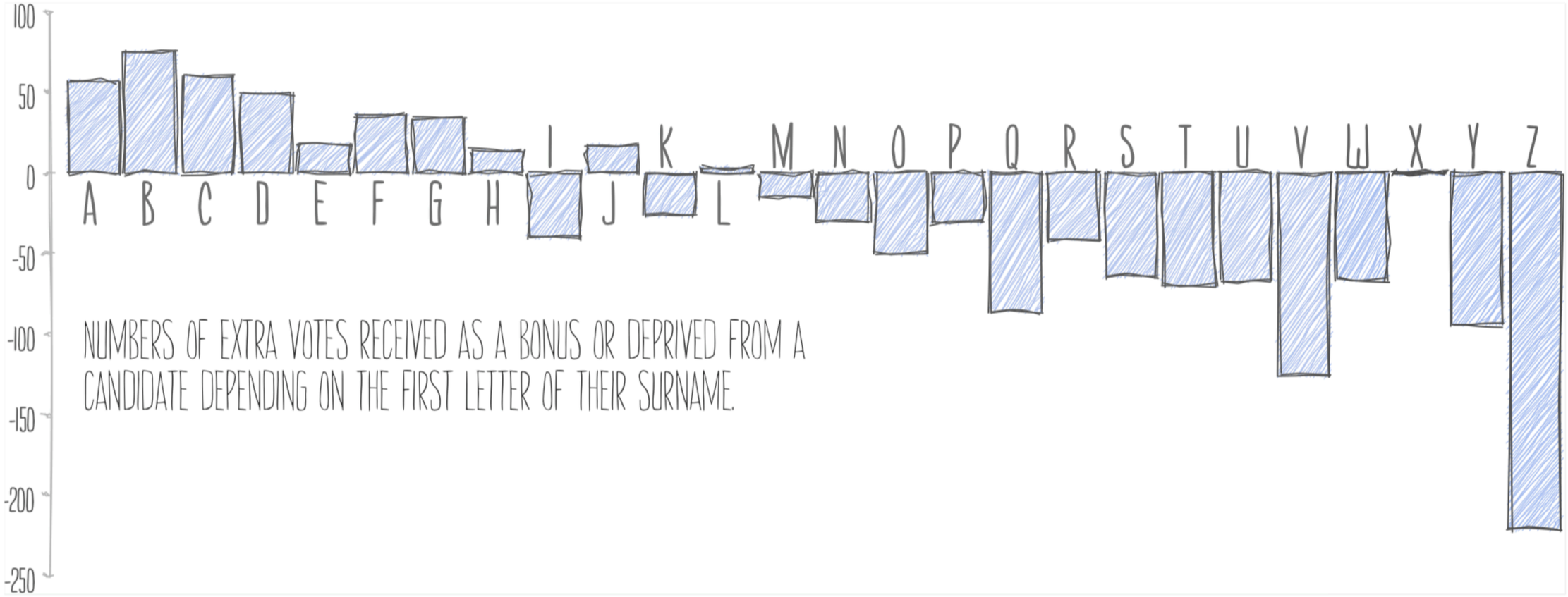


Bivariate Map (Data + Uncertainty)

Value-Suppressing Uncertainty Map

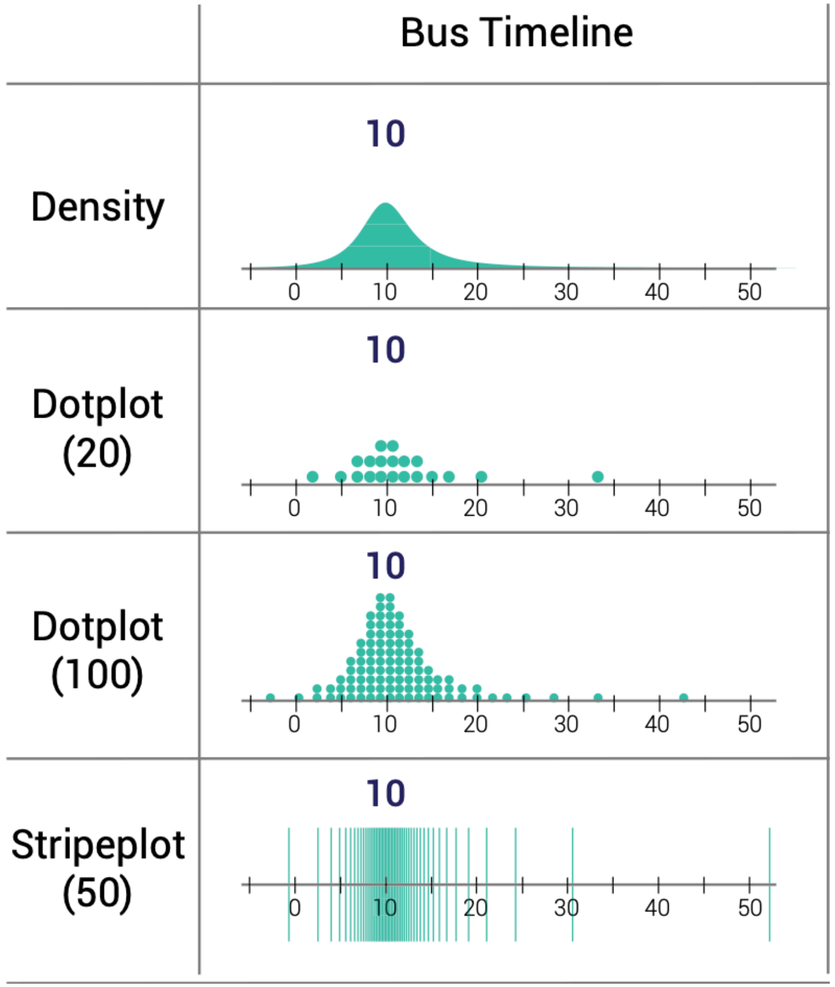
[Correll, Moritz, & Heer, 2018]

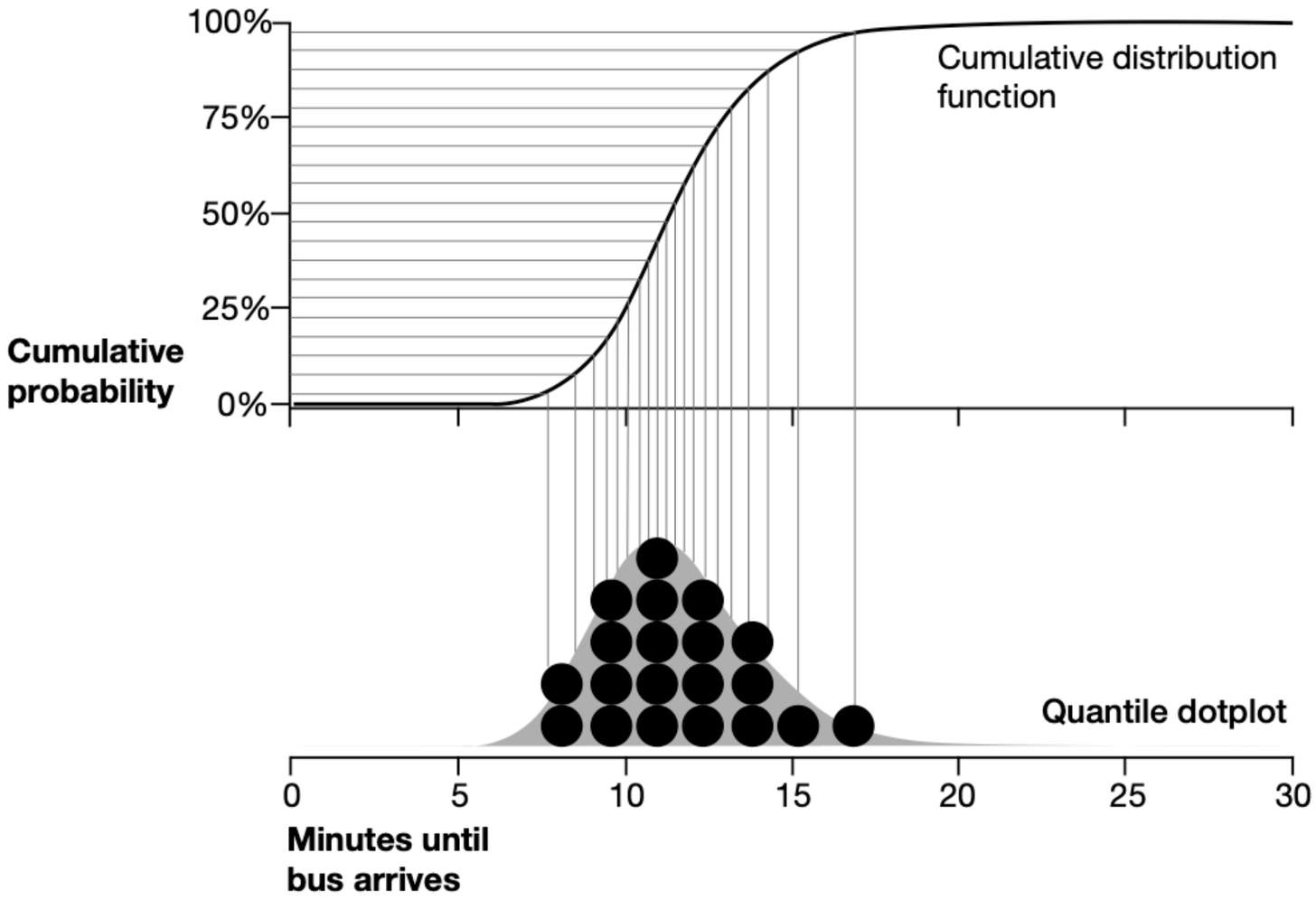
For uncertainty, use **visual variables** instead of visualizing point estimates



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

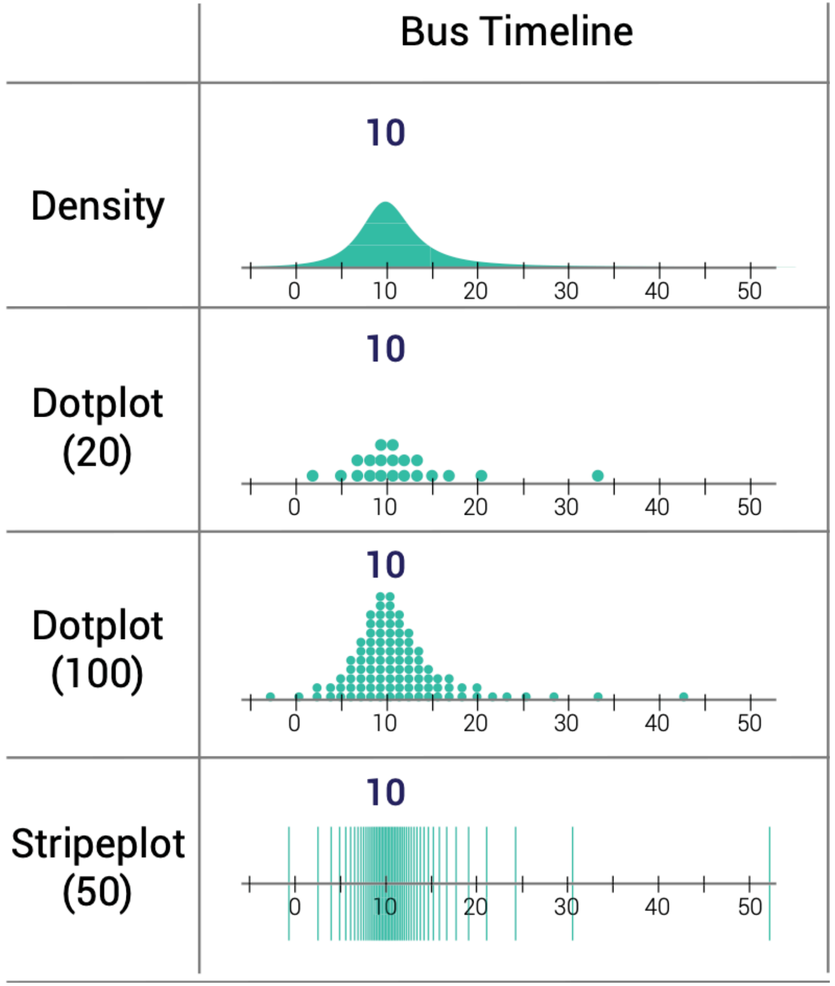
"Set of draws" technique

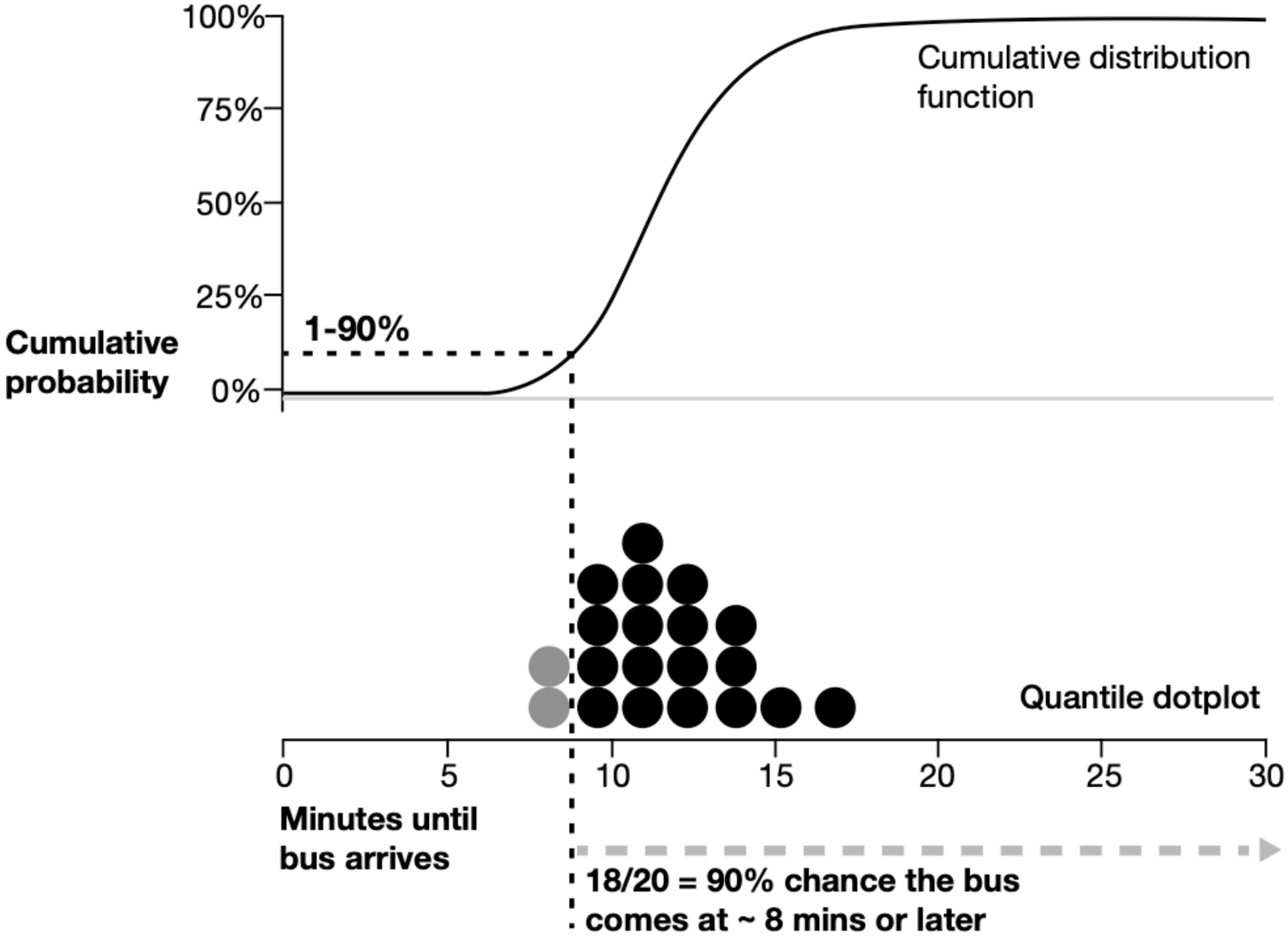




[Kay et al., 2016]

"Set of draws" technique

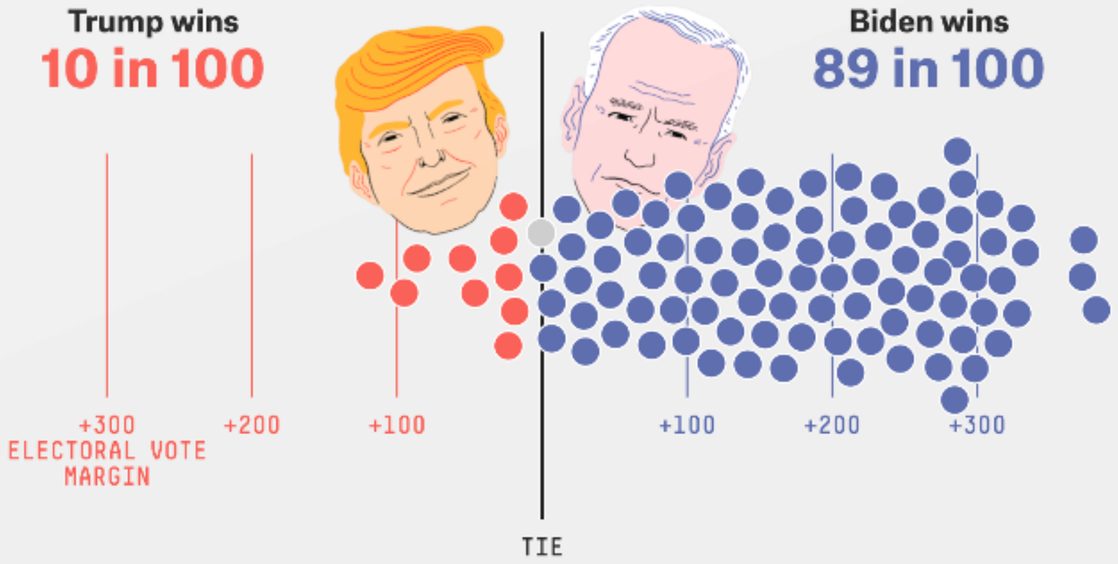




[Kay et al., 2016]

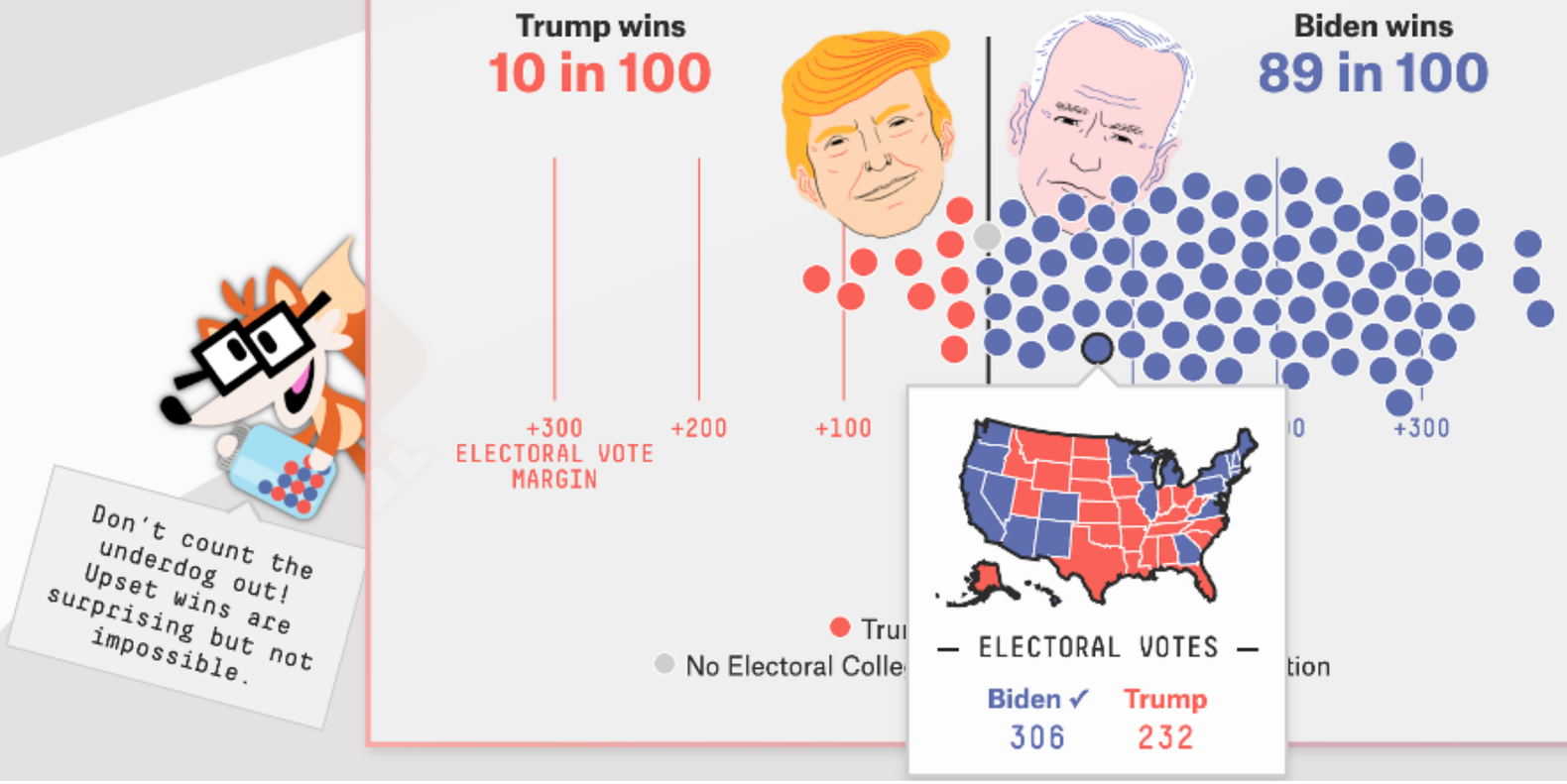
Biden is *favored* to win the election

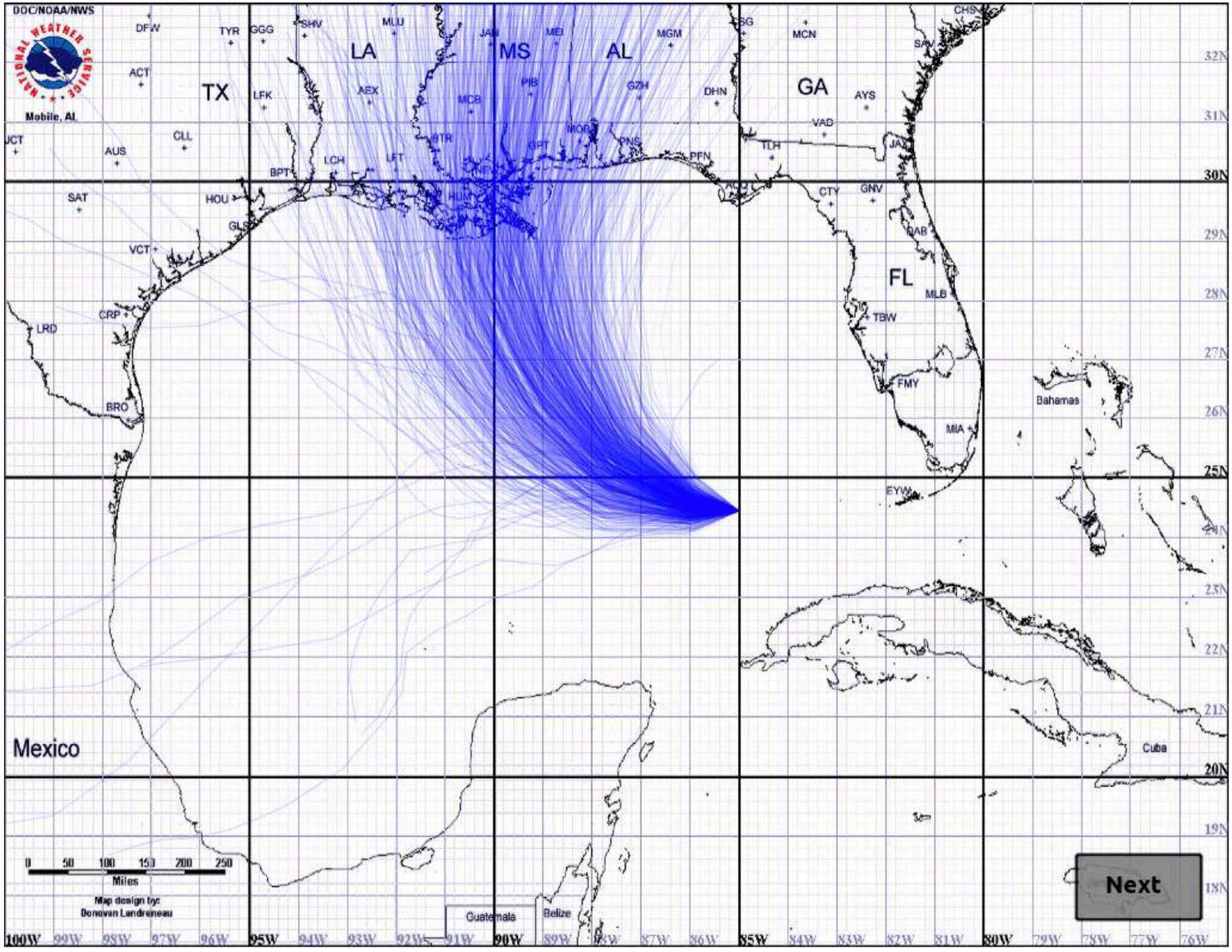
We simulate the election 40,000 times to see who wins most often. The sample of 100 outcomes below gives you a good idea of the range of scenarios our model thinks is possible.



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.




DOC/NOAA/NWS
NATIONAL WEATHER SERVICE
Mobile, AL
TX
LA
MS
AL
GA
FL
Mexico
Bahamas
Cuba
Guatemala
Belize
DFW
TYR
GGG
SHV
MLU
JAN
MEI
MGM
SG
MCN
CHS
SAV
ACT
LFK
AEX
MCB
PIB
GZH
DHN
AYS
VAD
JCT
AUS
CLL
BPT
LCH
LFT
HTR
GPT
MOB
PNS
PFN
TLH
JAX
SAT
HOU
GLS
CTY
GNV
VCT
DAB
MLB
CRP
LRD
TBW
FMY
BRO
MIA
EYW
0 50 100 150 200 250
Miles
Map design by:
Donovan Landreneau
Next
32N
31N
30N
29N
28N
27N
26N
25N
24N
23N
22N
21N
20N
19N
18N
100W
99W
98W
97W
96W
95W
94W
93W
92W
91W
90W
89W
88W
87W
86W
85W
84W
83W
82W
81W
80W
79W
78W
77W
76W

The New York Times

TheUpshot

STATISTICAL NOISE

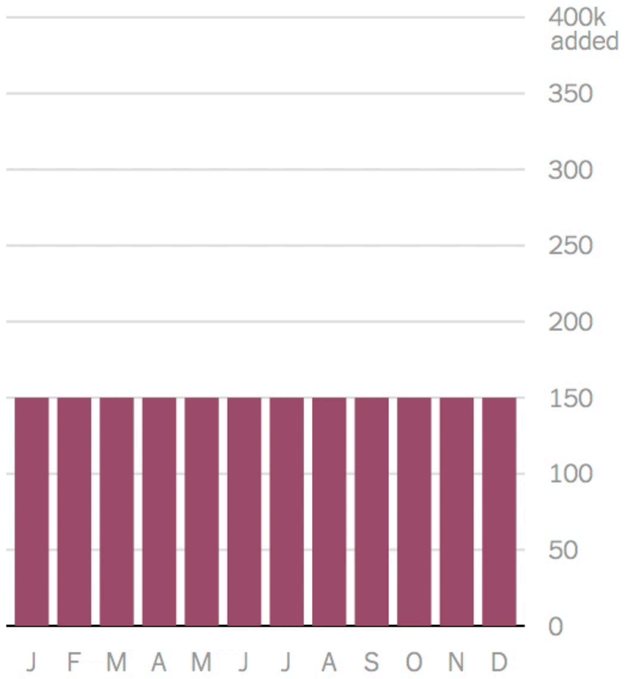
How Not to Be Misled by the Jobs Report

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.

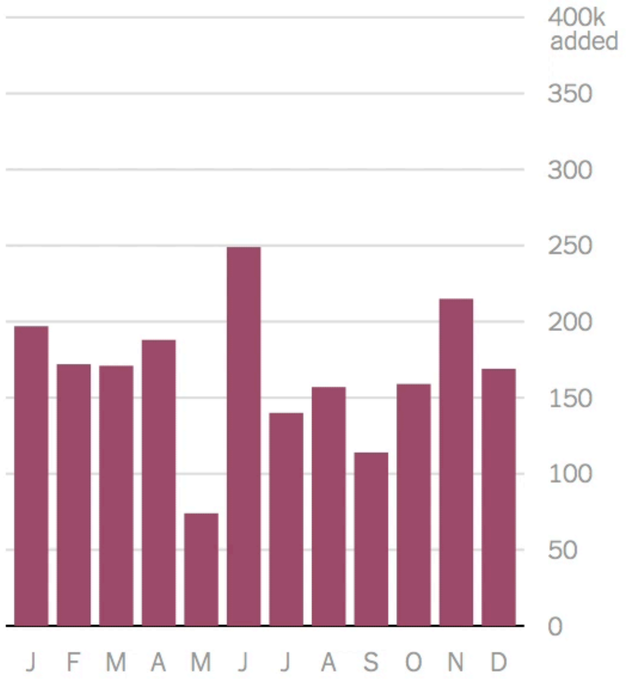
Under 55,000 jobs	55,000 to 110,000	110,000 to 140,000	160,000 to 190,000	190,000 to 245,000	245,000+
4% chance	19% chance	19% chance	19% chance	19% chance	4% chance

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

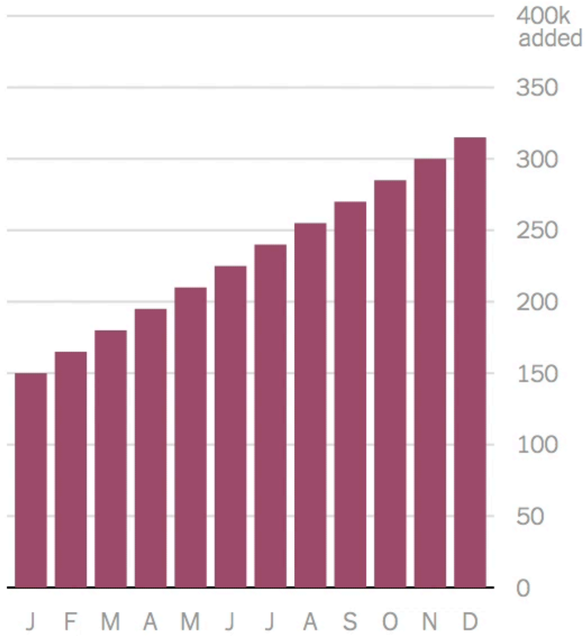


...the jobs report **could look like this**:

Play

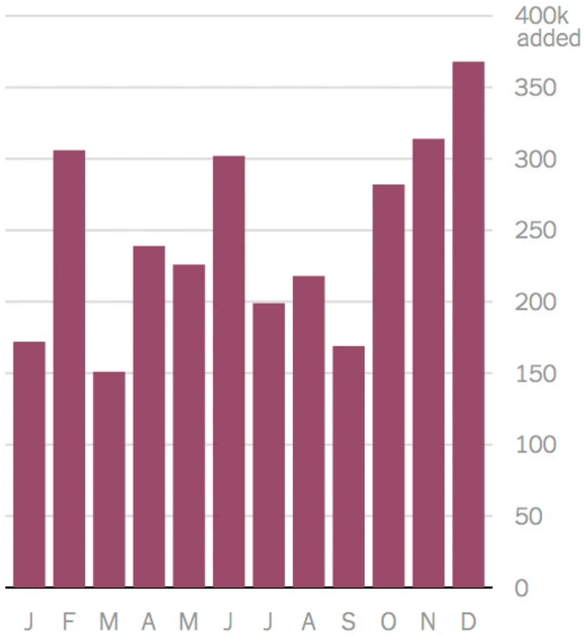


If job growth **had been accelerating**...



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

Play



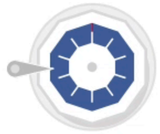
SPIN

Democrats: ?? **Republicans: ??**

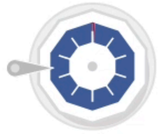
Likely Democratic

R.I.

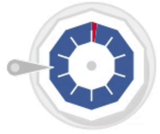
Va.



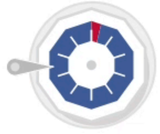
N.J.



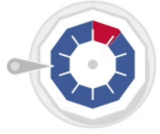
Ill.



Ore.

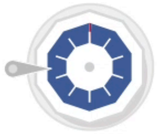


N.H.

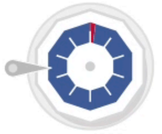


Hawaii

N.M.



Mass.



Del.

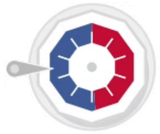
Minn.

Iowa

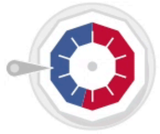
Competitive

Colo.

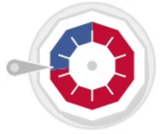
Mich.



Ark.



Ga.



N.C.

Alaska

La.

Ky.

Likely Republican

Ala.

Wyo.

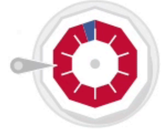
Okla.

Kan.

Tenn.

S.D.

S.C.



Miss.

Okla. (S)

Idaho

Tex.

Me.

S.C. (S)

Neb.

W.Va.

Mont.

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:

TRUMP	BUSH	WALKER	PAUL	RUBIO	CRUZ	HUCKABEE	CARSON	CHRISTIE	PERRY	KASICH	SANTORUM	JINDAL	FIORINA	PATAKI	GRAHAM
18.0%	**14.0%**	**11.0%**	**7.0%**	**6.0%**	**6.0%**	**5.0%**	**5.0%**	**3.0%**	**2.0%**	**2.0%**	**2.0%**	**1.0%**	**1.0%**	**1.0%**	**0.0%**

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.

Hullman, 2019

Uncertainty

What does it mean?

Lots of things!

How should I visualize it?

It depends!