Uncertainty Visualization

Michael Correll Tableau Research

Questions To Answer

What Does Uncertainty Mean?

How Should I Visualize It?

What Can Go Wrong?

Definitions and Bookkeeping

WHAT DOES UNCERTAINTY MEAN, ANYWAY?

Things "Uncertainty" Can Mean

Doubt

Risk

Variability

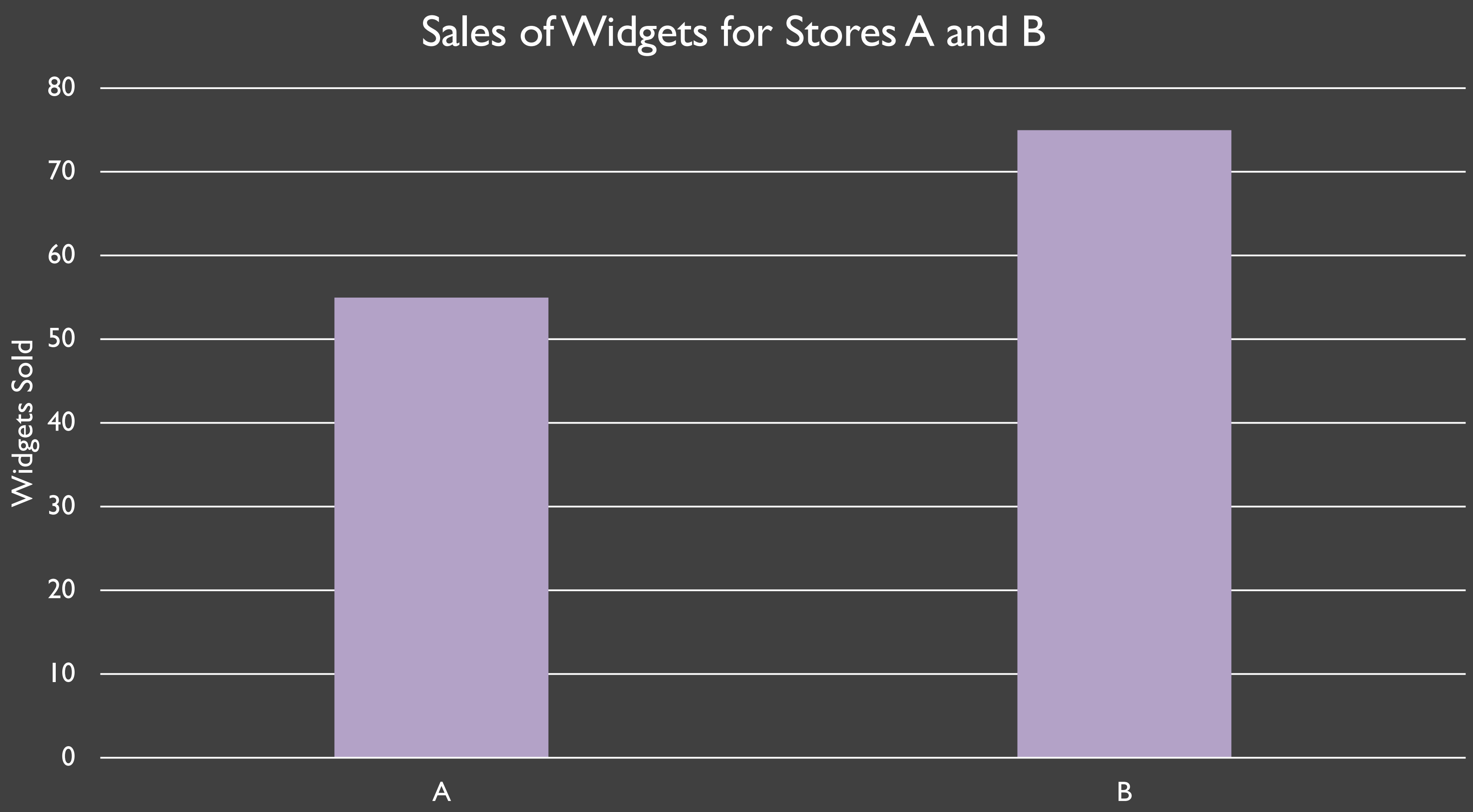
Error

Lack of Knowledge

Hedging

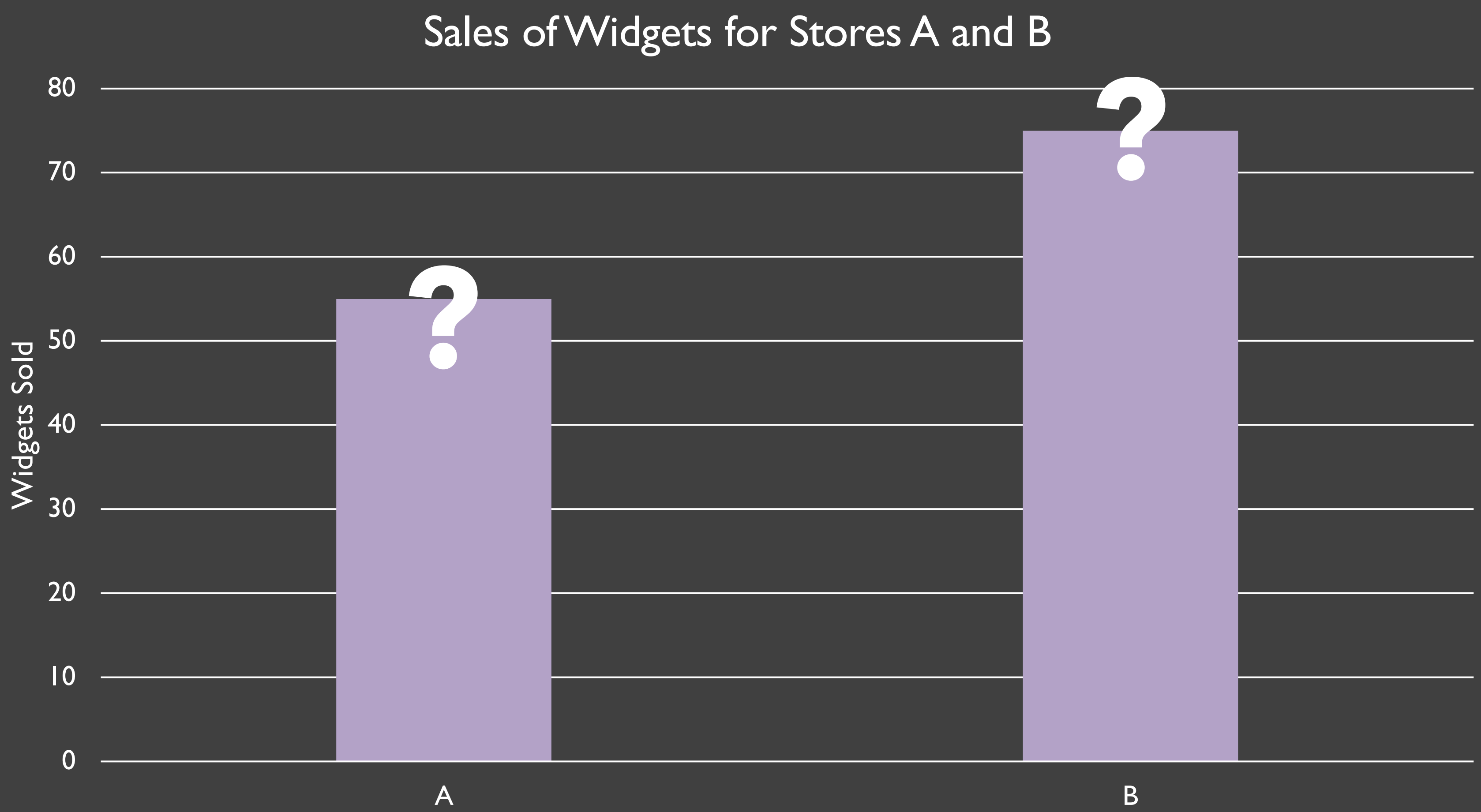
…

A Bar Chart

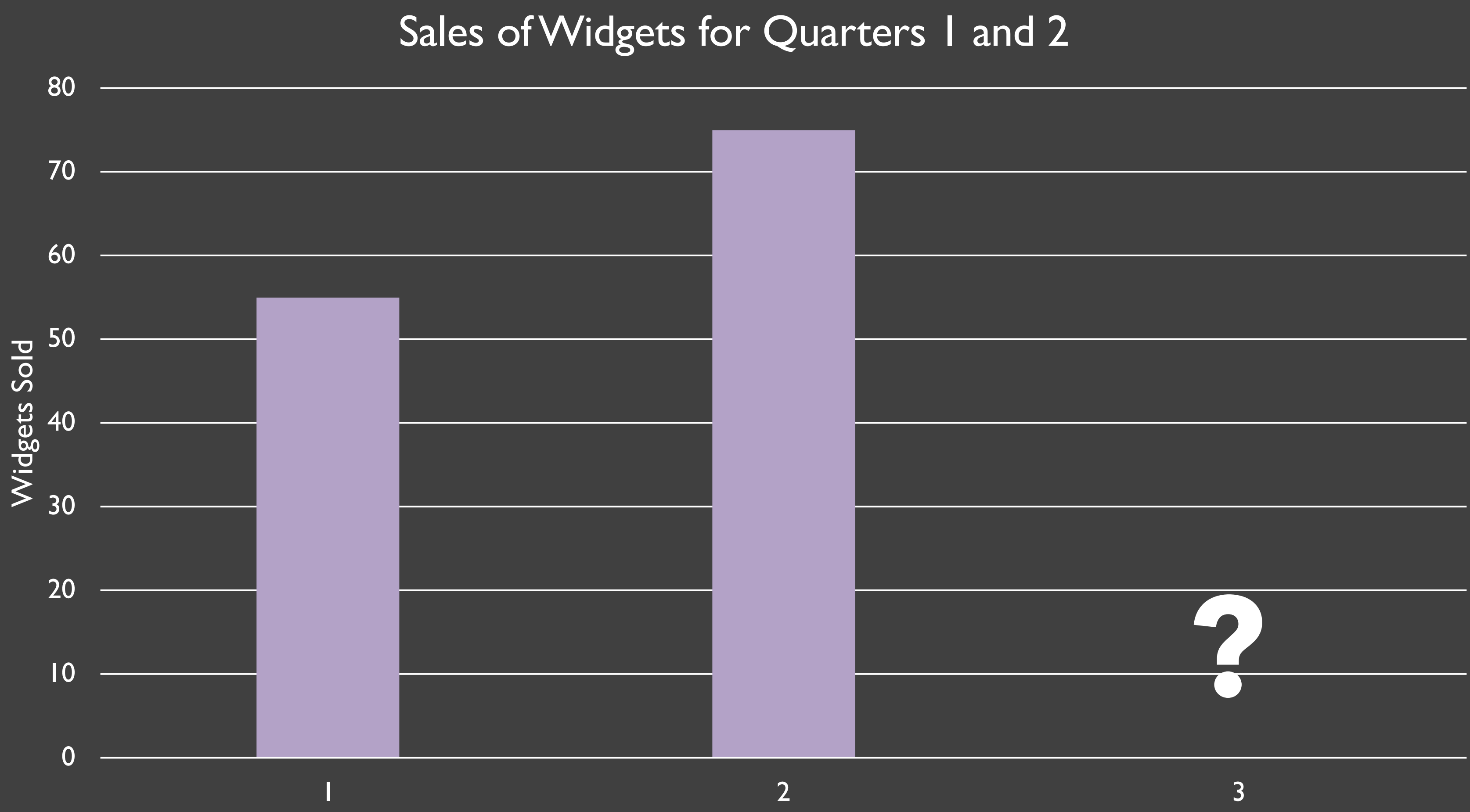


Sales of Widgets for Stores A and B
Widgets Sold
80
70
60
50
40
30
20
10
0
A
B

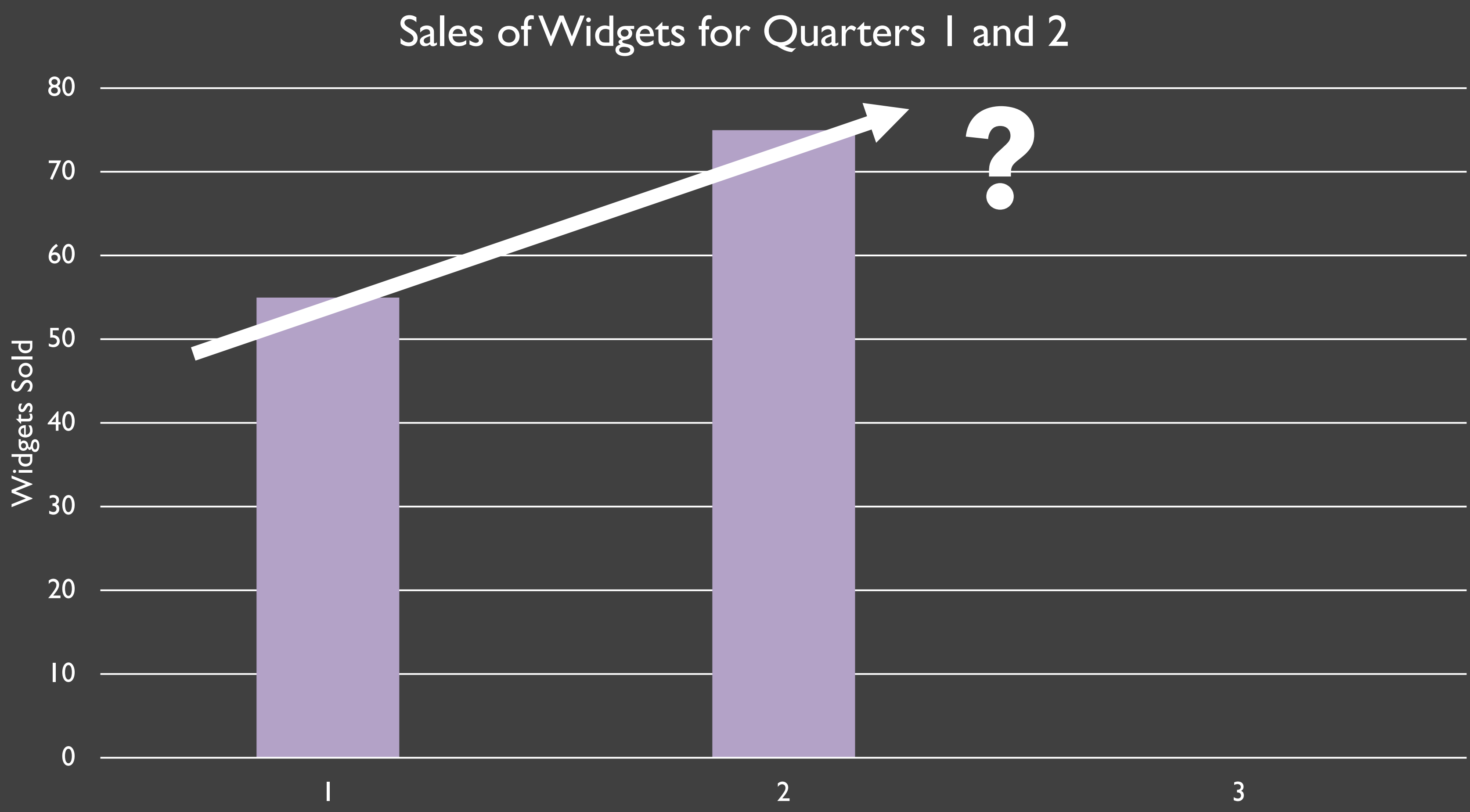
Measurement Uncertainty



Forecast Uncertainty

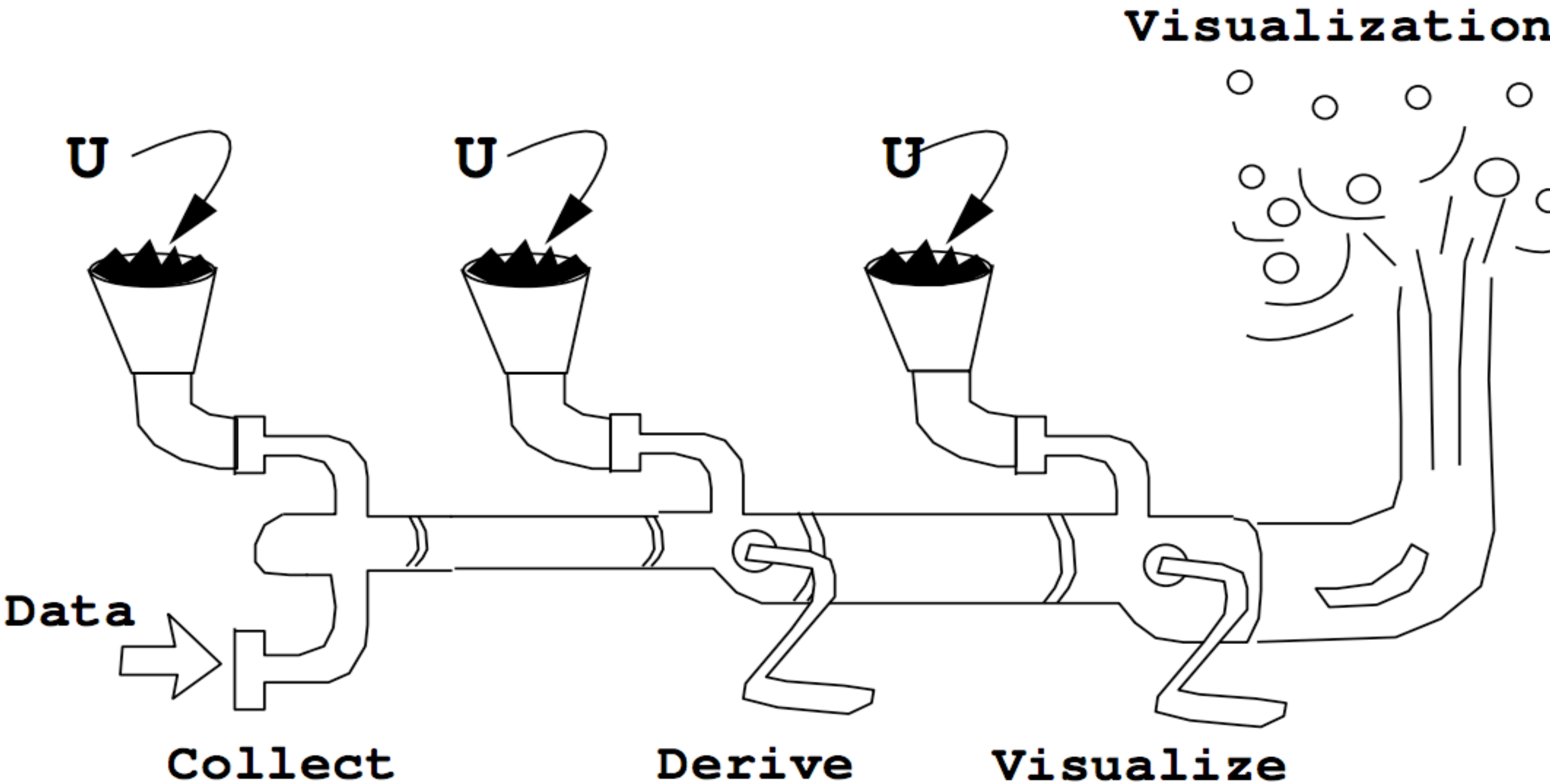


Model Uncertainty



Decision Uncertainty

Uncertainty Vis Pipeline



Pang et al. Approaches to Uncertainty Visualization. The Visual Computer, 1997.

Uncertainty Sources

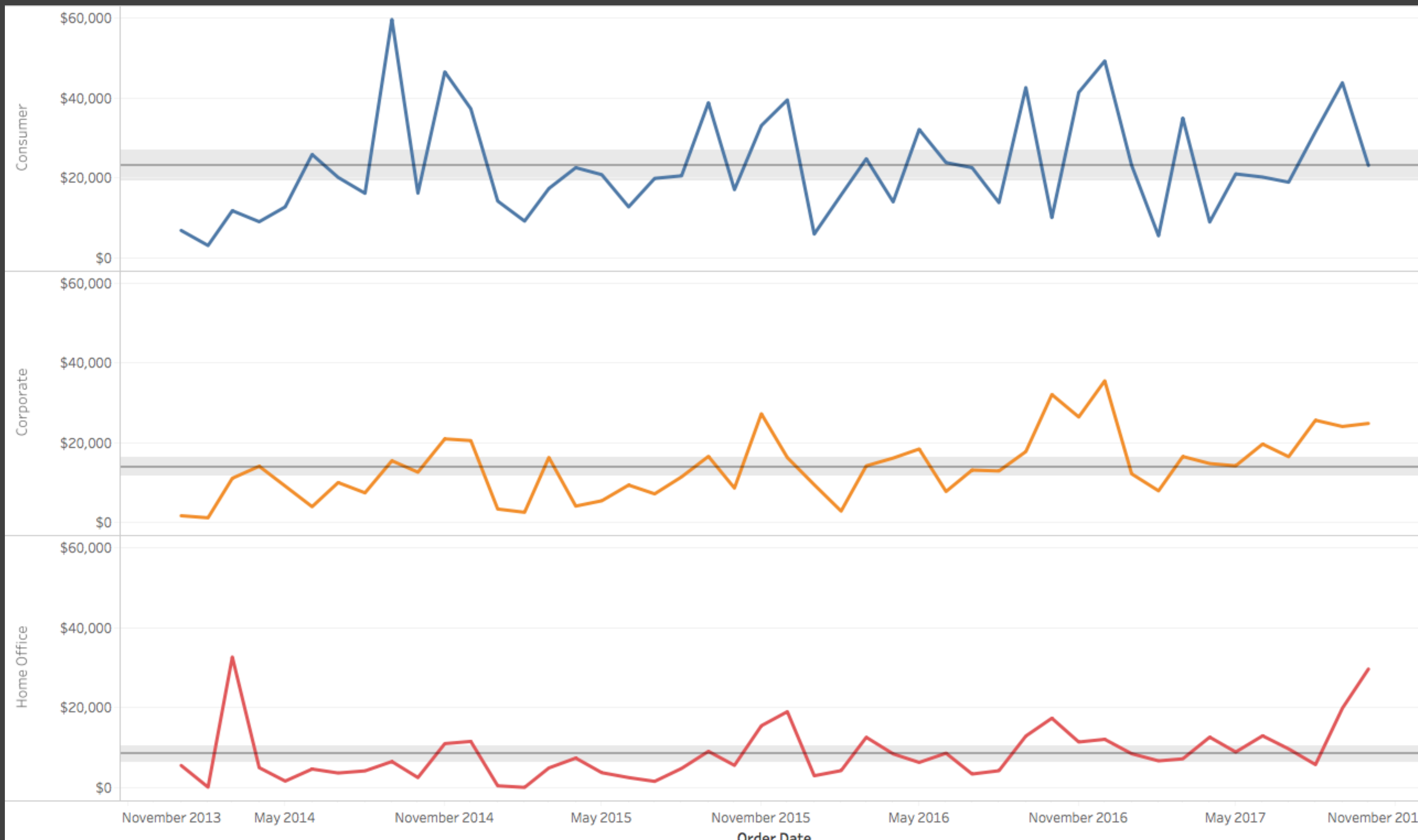
Measurement Uncertainty: "We're not sure what the data are"

Forecast Uncertainty: "We're not sure what will happen to the data next"

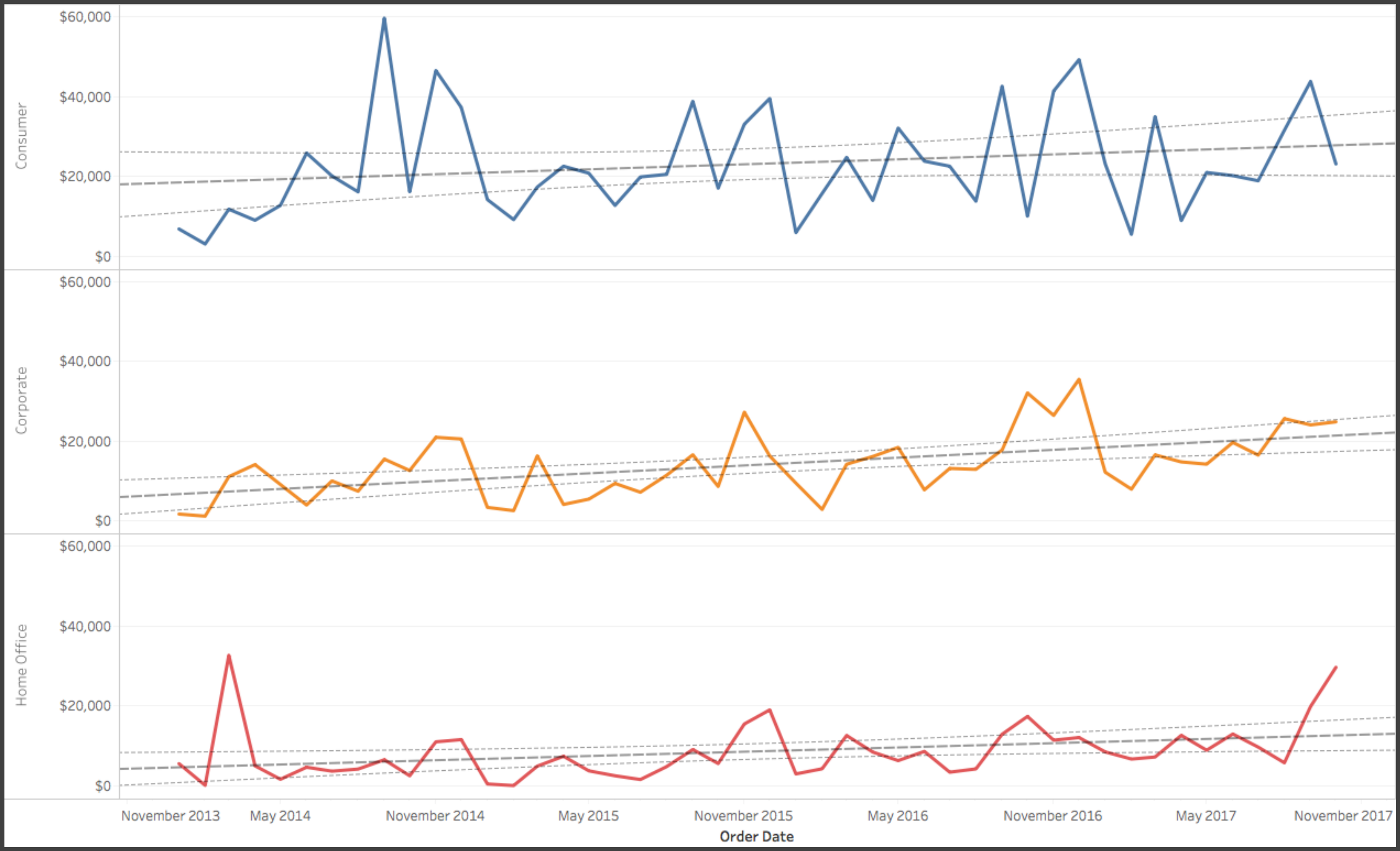
Model Uncertainty: "We're not sure how the data fit together"

Decision Uncertainty: "We're not sure what to do with the data"

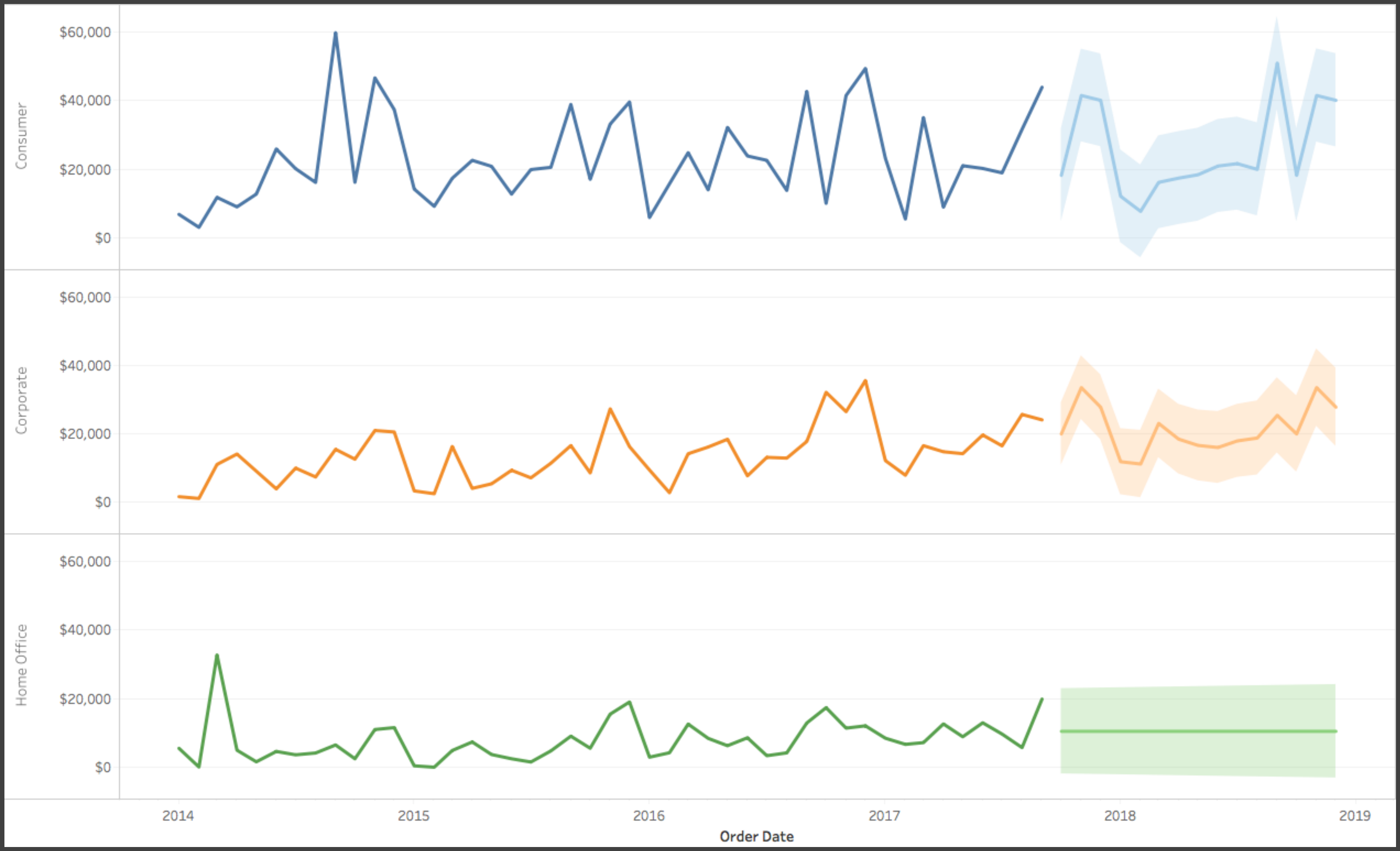
Measurement Uncertainty



Model Uncertainty



Forecast Uncertainty



Uncertainty Visualization

There are different **types** and **sources** of uncertainty.

We can **quantify** or **model** our uncertainty.

The visual presentation of uncertainty can **clash** with cognitive and perceptual biases.

Terminology

Aleatory Uncertainty

Epistemic Uncertainty

Type I error

Type II error

Precision

Bias

What Will Happen When I Flip This Coin?

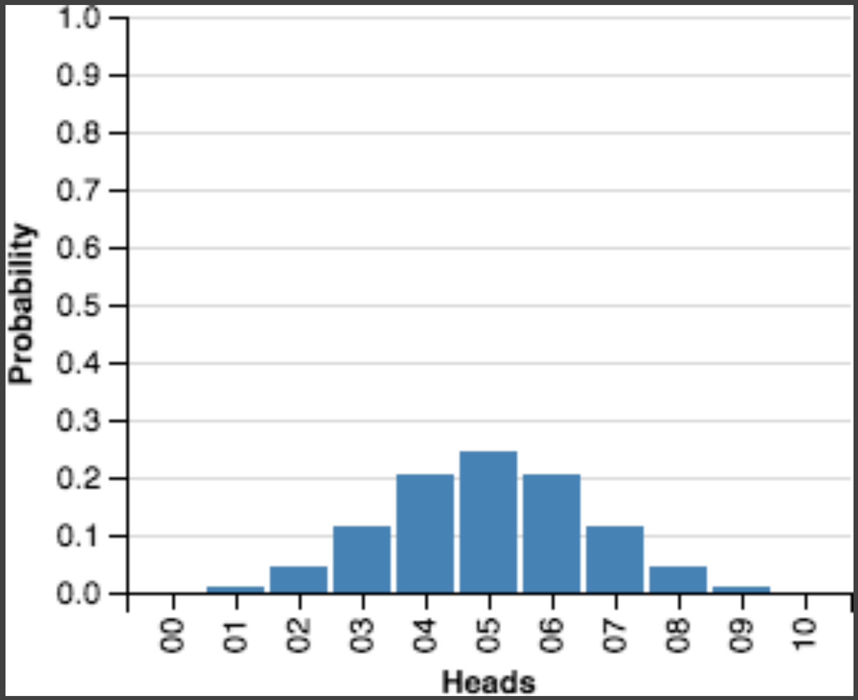
What Will Happen When I Flip This Coin?

P(Heads) = 0.5

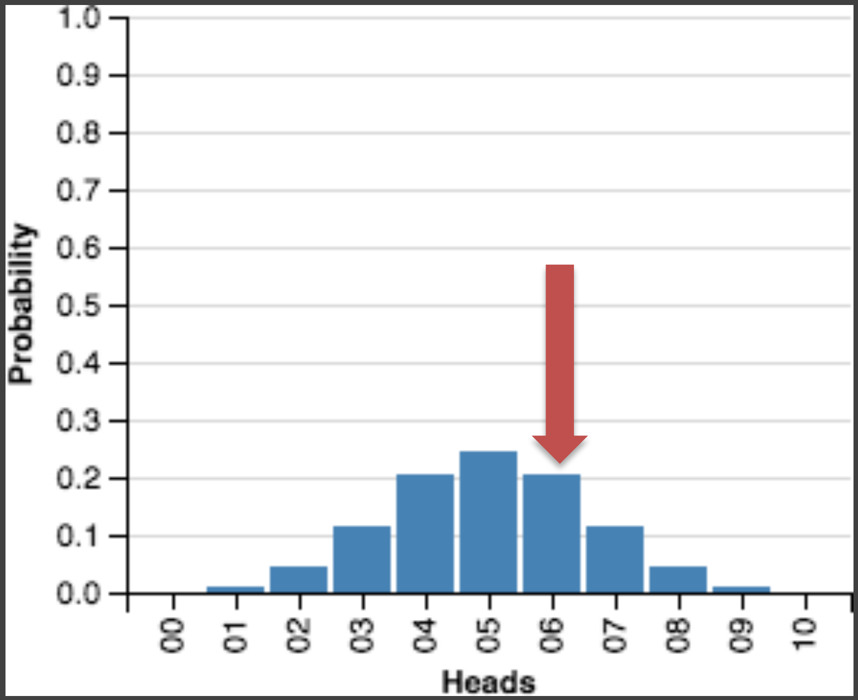
P(Tails) = 0.5

Aleatory Uncertainty

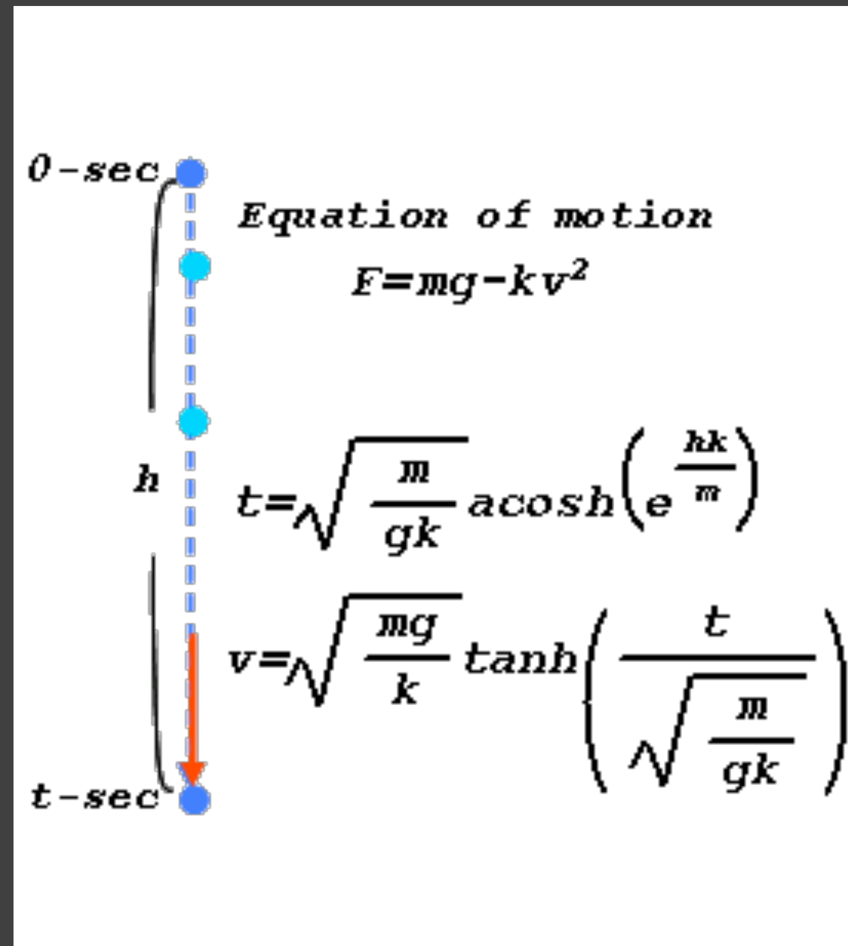
Aleatory Uncertainty



Aleatory Uncertainty



Epistemic Uncertainty



Uncertainty Types

Aleatory

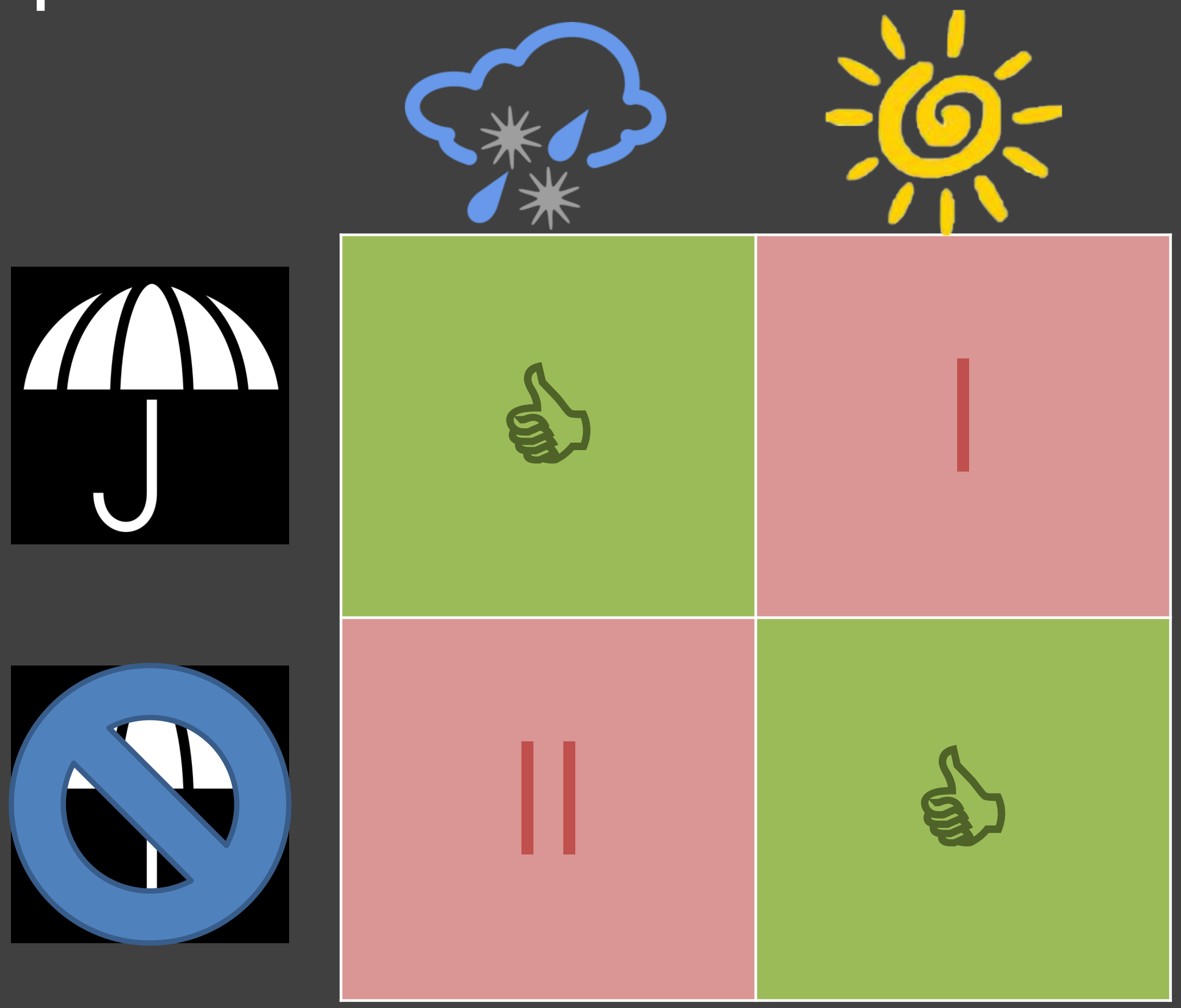
Variability: things that we don’t know (but can reason about the likelihood of).

Epistemic

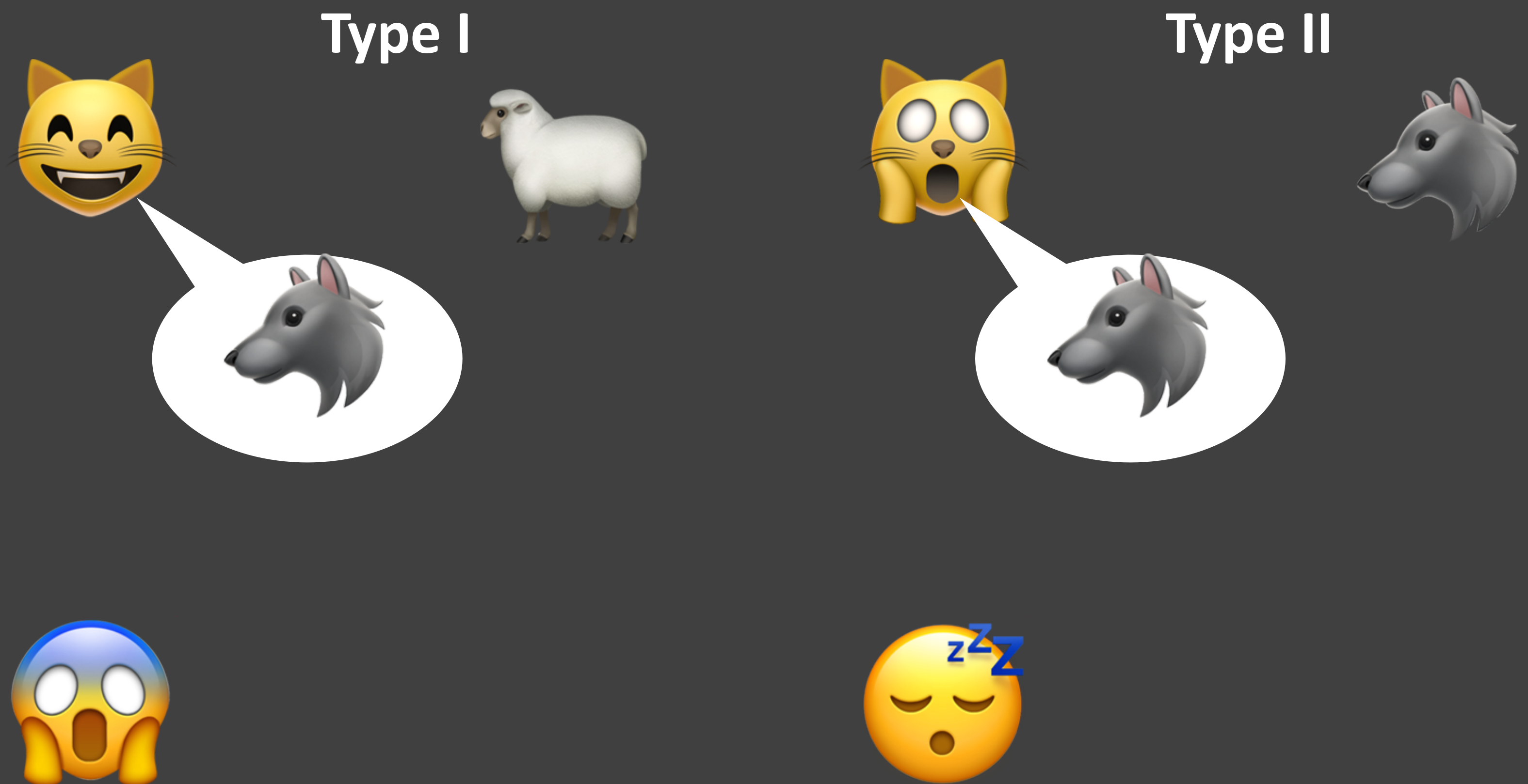
Things we could in principle know for certain, but have not measured.

Should I Bring an Umbrella?

Type I and II Errors



The Boy Who Cried Wolf



Did My Arrows Hit the Target?

Precision & Bias

Precision

Precision & Bias

Precision

Precision & Bias

Precision

Precision & Bias

Precision

Accuracy

Precision & Bias

Precision

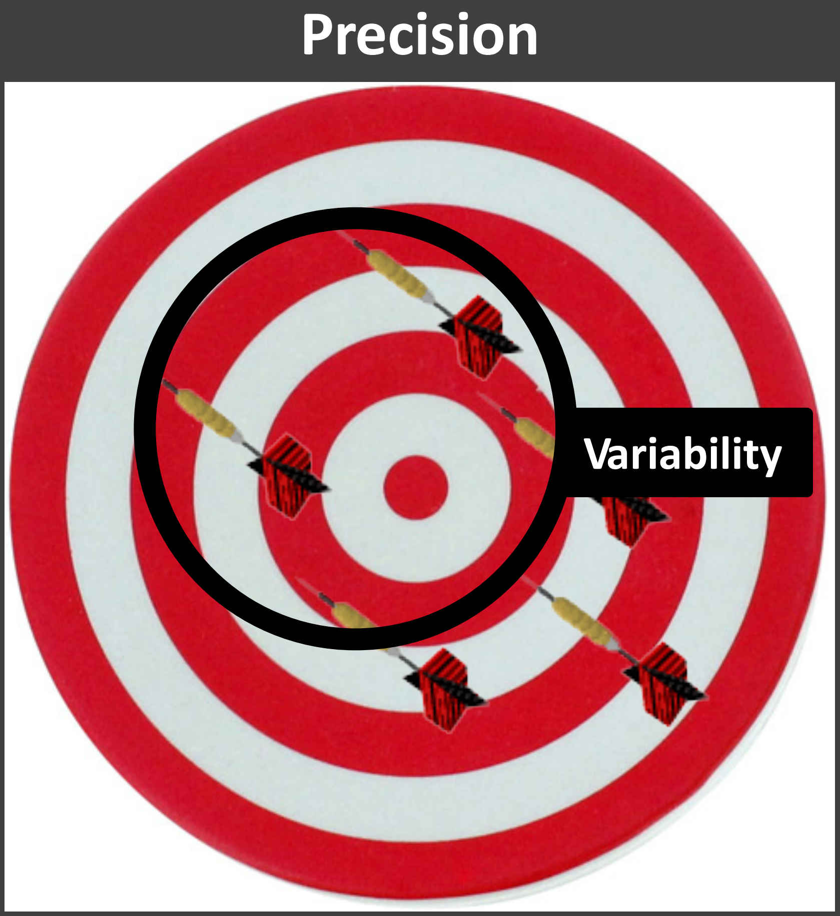
Accuracy

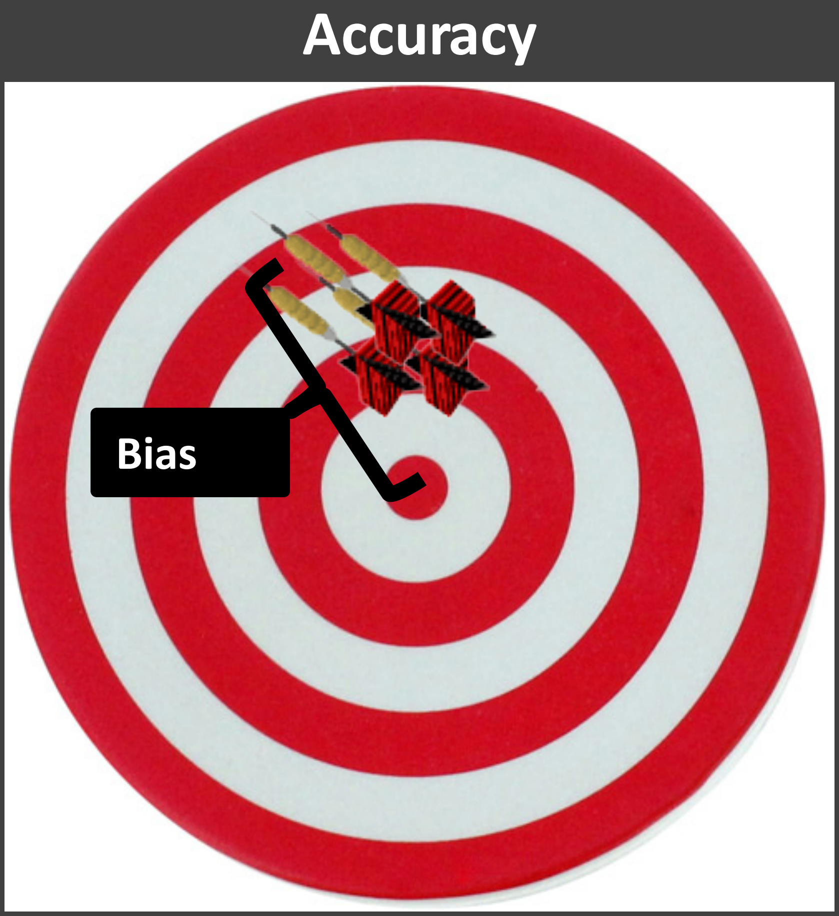
Precision & Bias

Precision

Accuracy

Precision & Bias





What Does Uncertainty Mean?

Any one of a number of potentially interconnected quantitative, qualitative, or factors that affect the quality, reliability, or utility of your data or data-driven decisions. Anything that can cause you to be unsure about your data or how to use it.

What Does Uncertainty Mean?

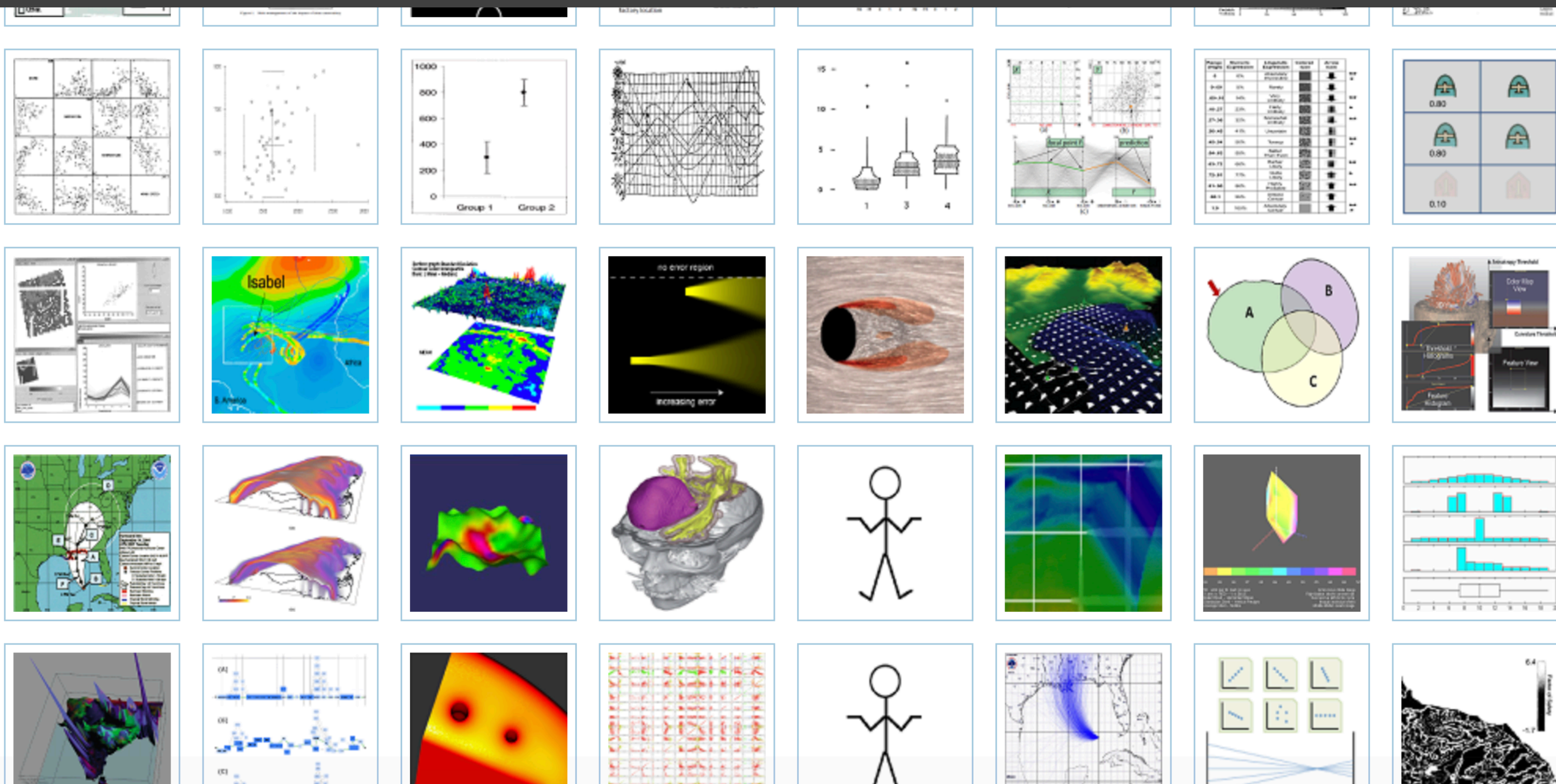
Any one of a number of potentially interconnected quantitative, qualitative, or factors that affect the quality, reliability, or utility of your data or data-driven decisions. Anything that can cause you to be unsure about your data or how to use it.

LOTS OF THINGS

Uncertainty Maps and Model Visualization

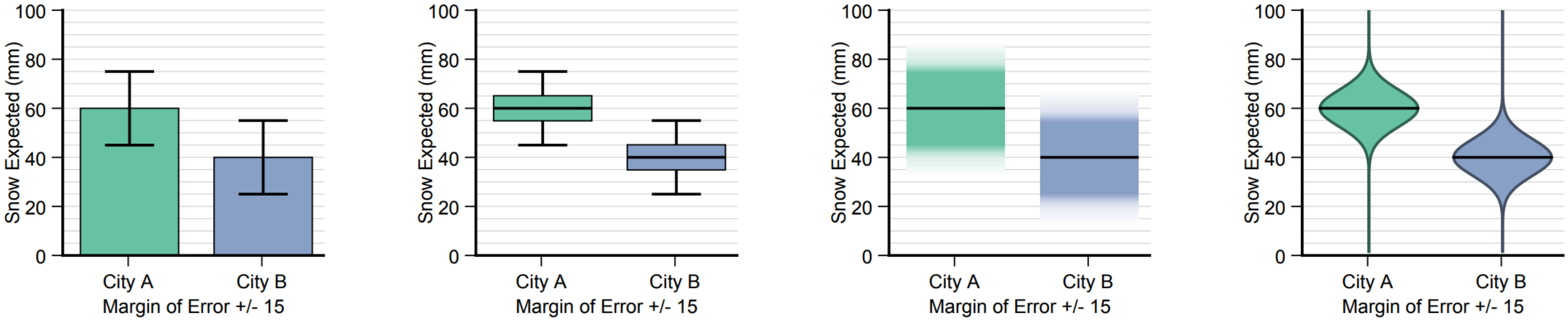
HOW SHOULD I VISUALIZE UNCERTAINTY?

Uncertainty Visualization Zoo

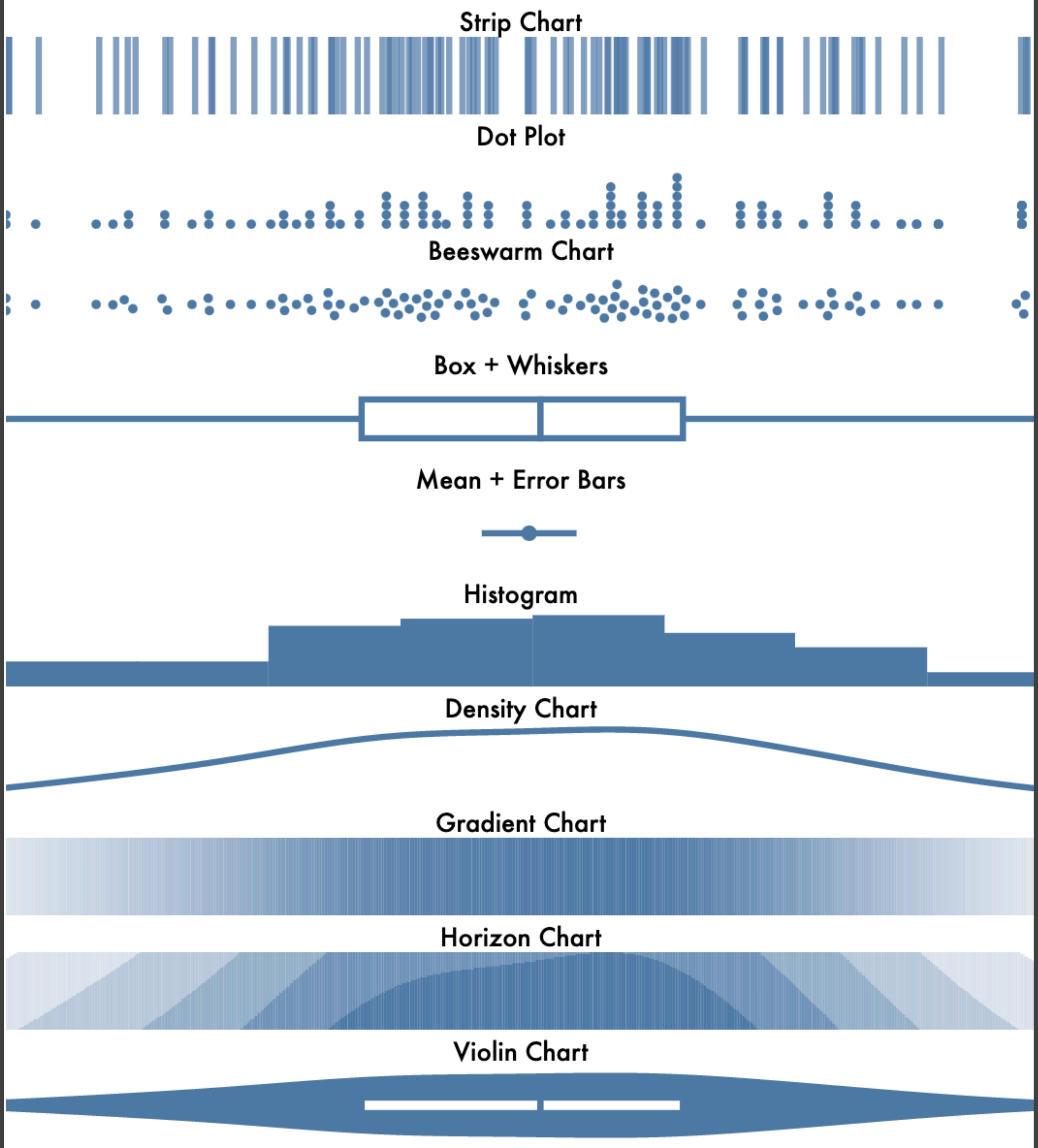


Jena et al. Uncertainty Visualisation: An Interactive Visual Survey. PACVIS, 2020.

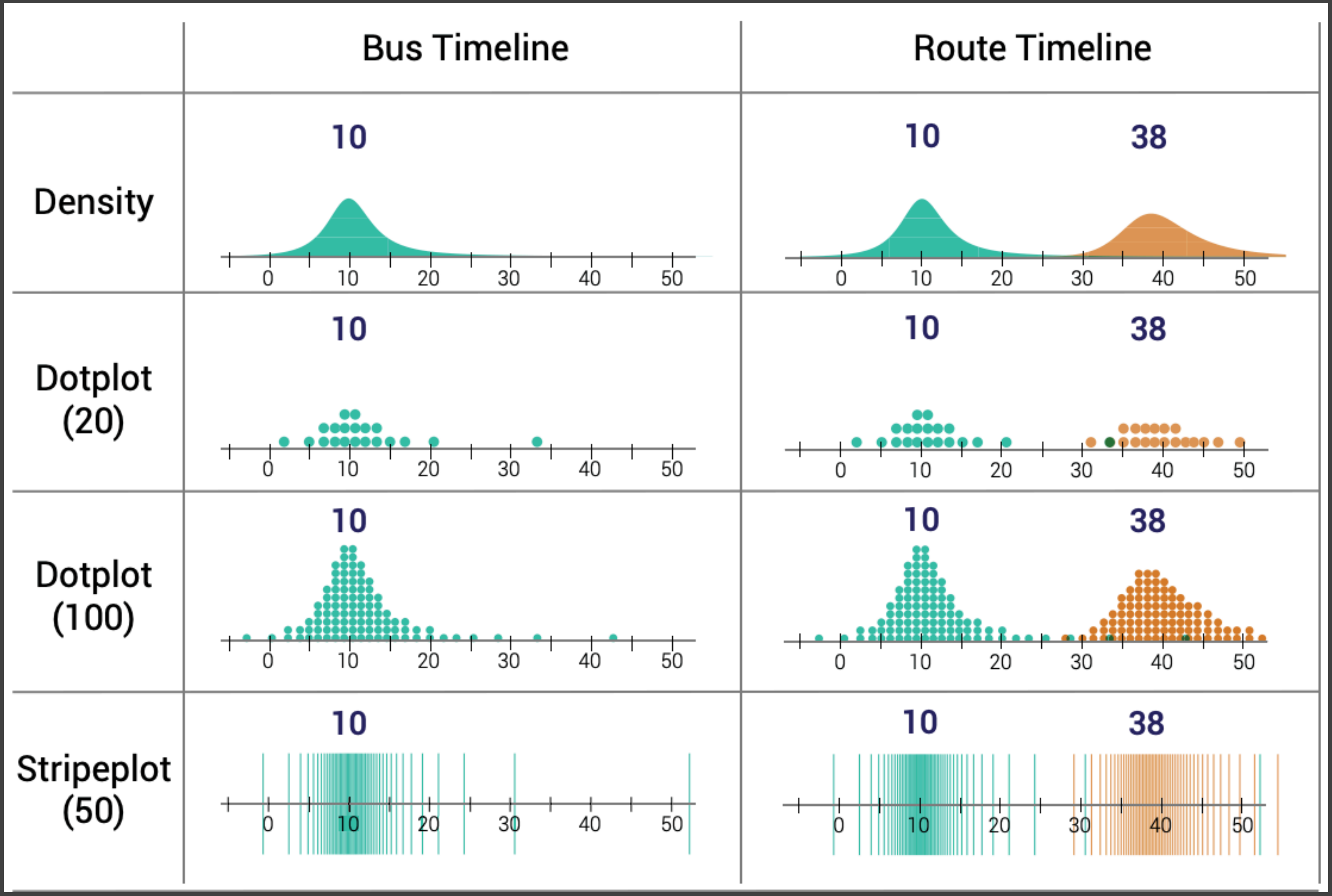
Intervals



Correll and Gleicher. Error Bars Considered Harmful:
Exploring Alternate Encodings for Mean and Error. VIS, 2014.

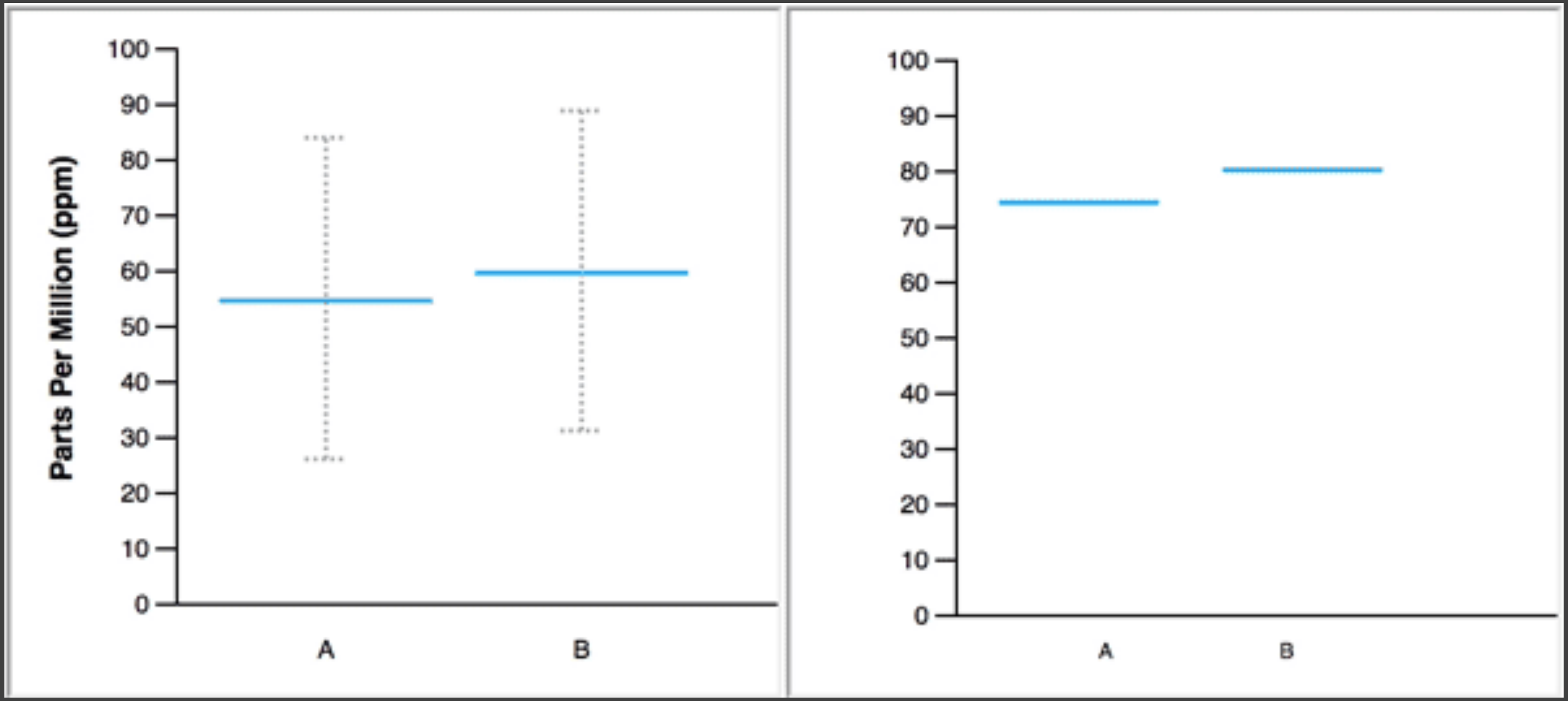

Strip Chart
Dot Plot
Beeswarm Chart
Box + Whiskers
Mean + Error Bars
Histogram
Density Chart
Gradient Chart
Horizon Chart
Violin Chart

Intervals

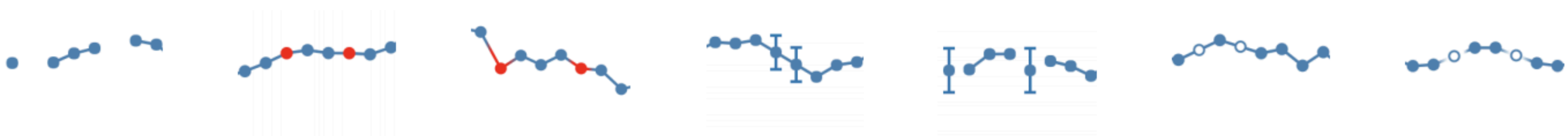


Kay et al. When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. CHI, 2016.

Hypothetical Outcome Plots



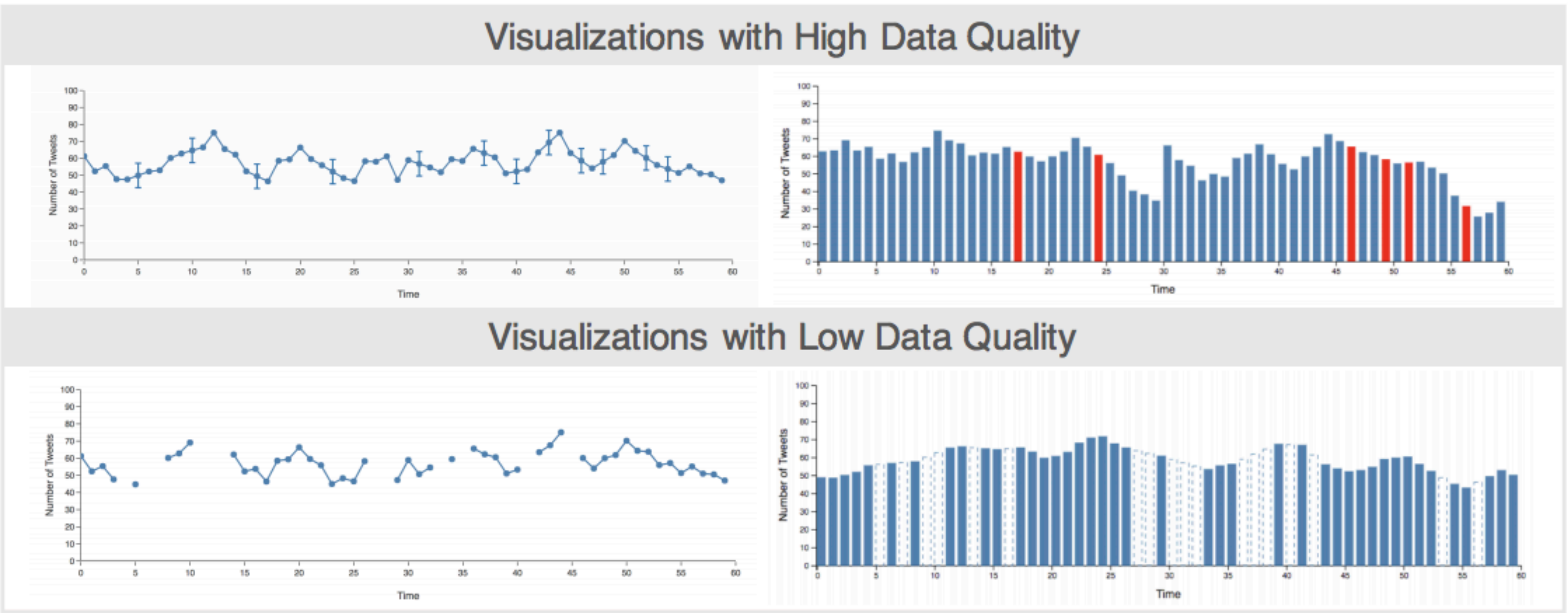
Missing Values



(a) Data Absent (b) Color Points (c) Color Points & Line Gradients (d) Connected Error Bars (e) Disconnected Error Bars (f) Unfilled Points (g) Unfilled Points & Line Gradients

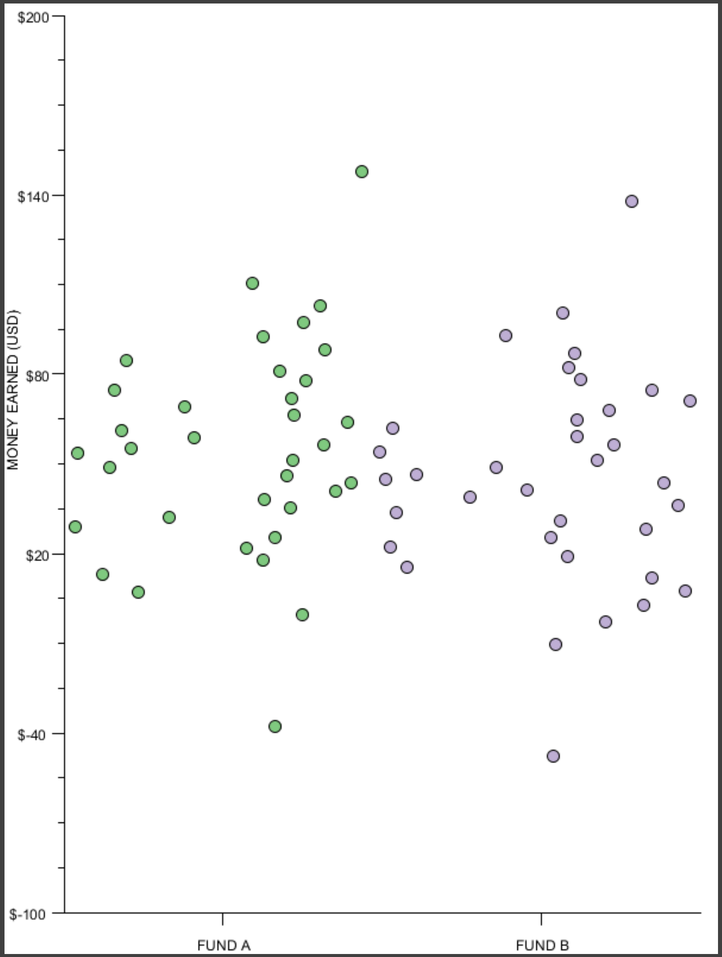
Song, Hayeon and Szafir, Danielle. Where's My Data? Evaluating Visualizations with Missing Data. IEEE VIS, 2018.

Missing Values

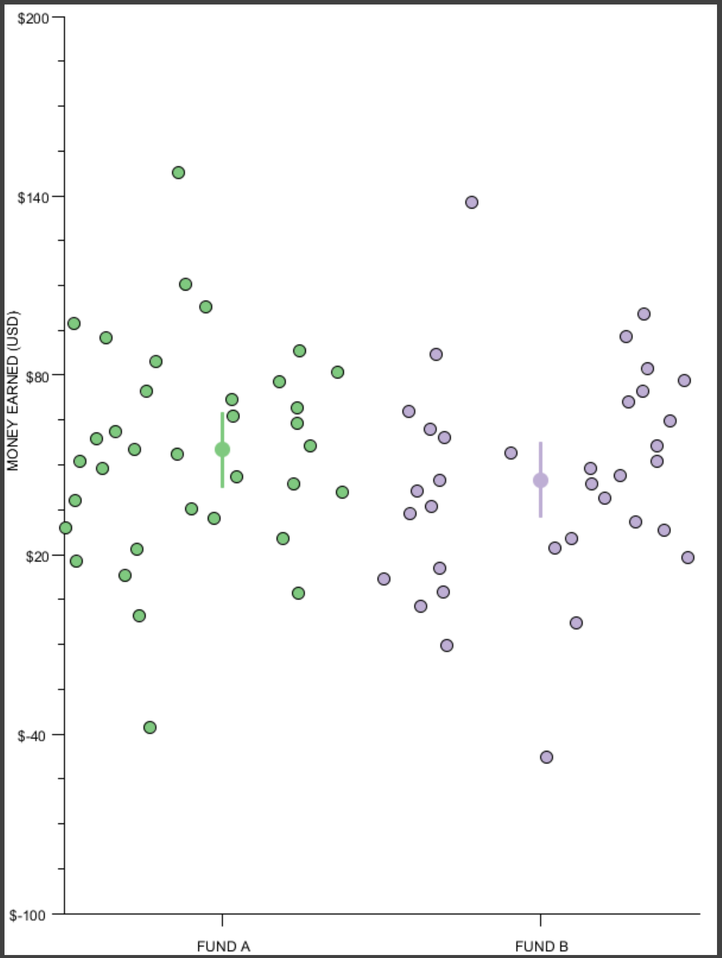


Song, Hayeon and Szafir, Danielle. Where's My Data? Evaluating Visualizations with Missing Data. IEEE VIS, 2018.

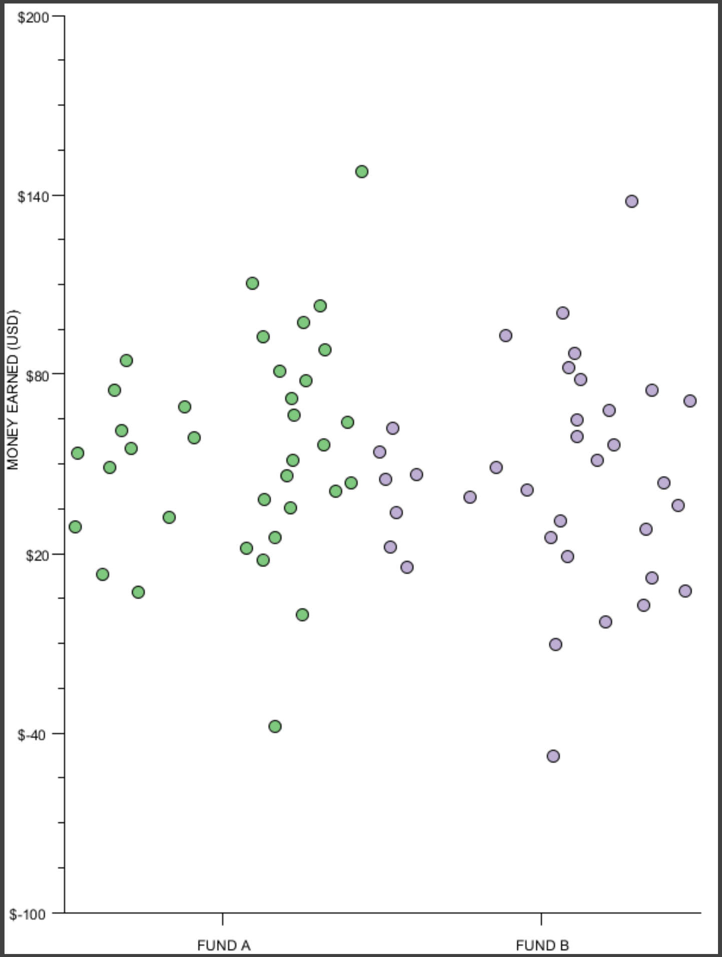
Special Case: Implicit Uncertainty

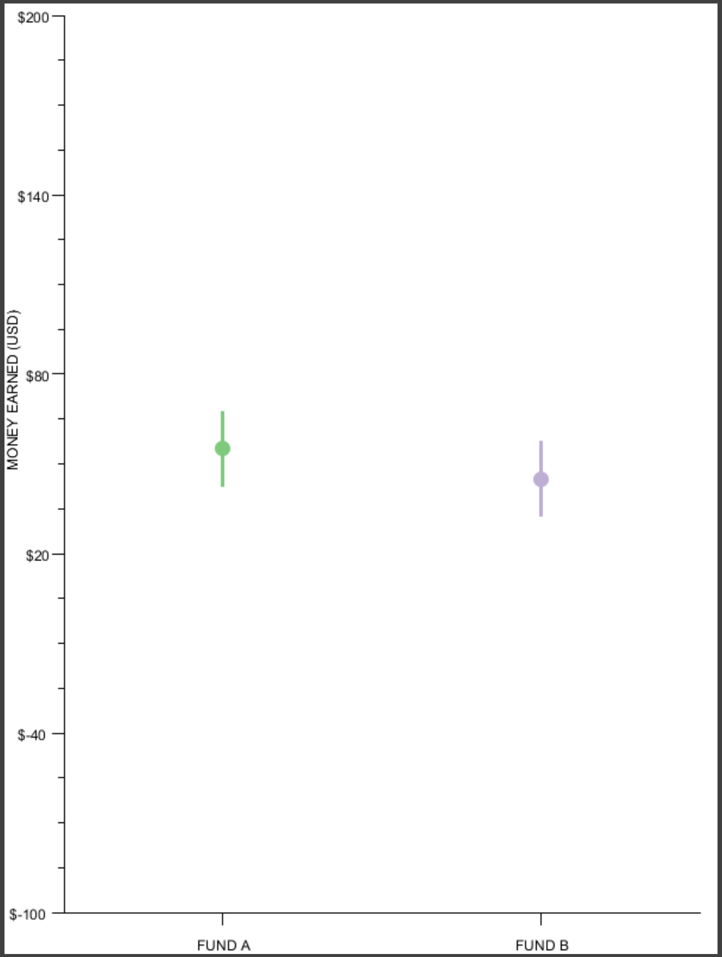


Special Case: Implicit Uncertainty



Special Case: Implicit Uncertainty



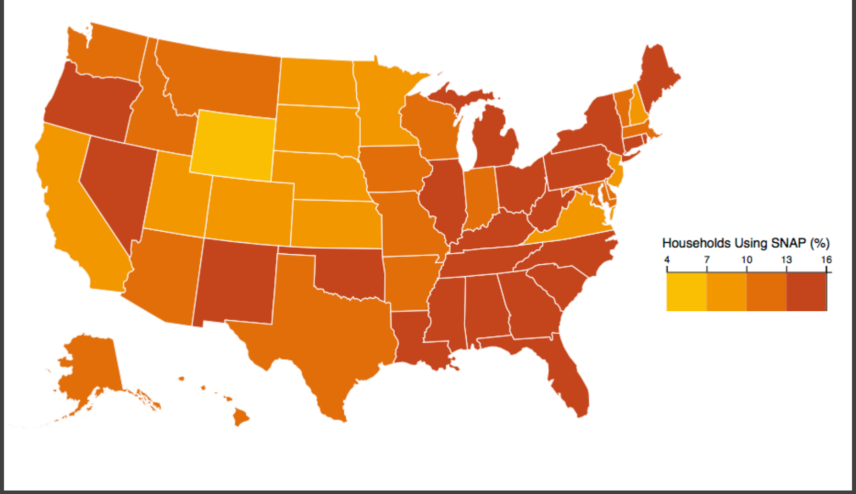


Uncertainty Vis Pipeline

1) Quantify Uncertainty
2) Choose a free visual variable
3) Encode uncertainty with the variable

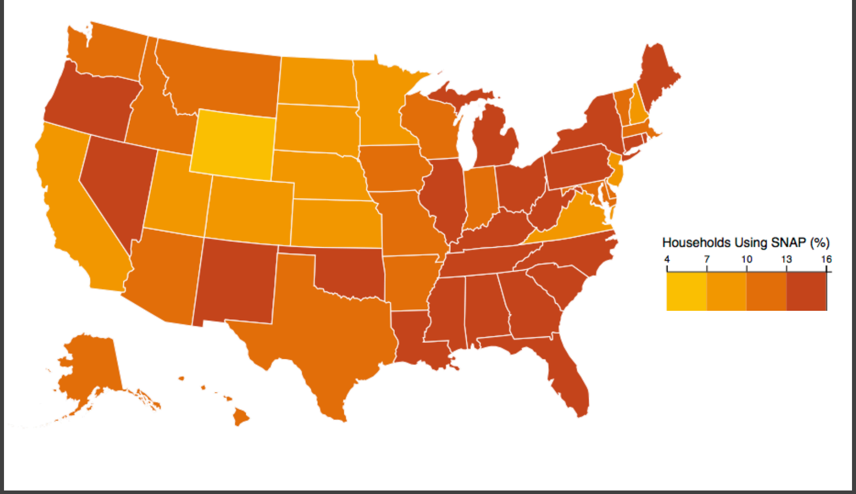
SNAP

Data Map

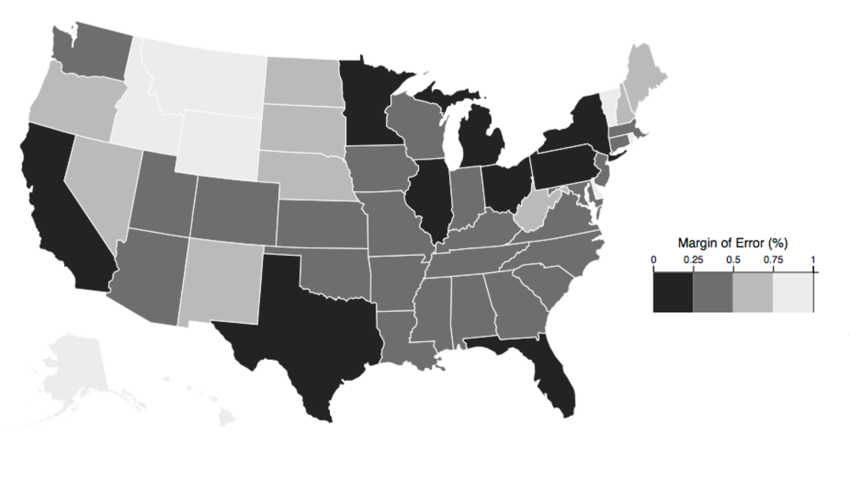


SNAP

Data Map



Uncertainty Map



Uncertainty Vis Pipeline

1) Quantify Uncertainty
2) Choose a free visual variable
3) Encode uncertainty with the variable

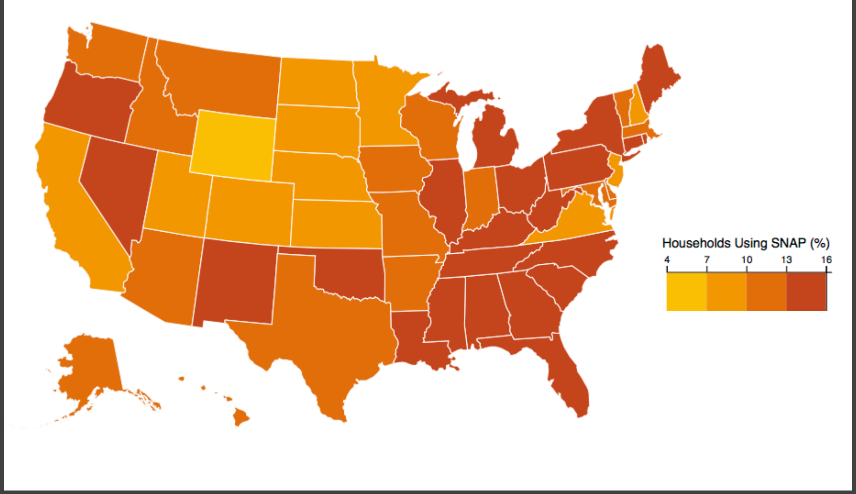
Uncertainty Vis Pipeline

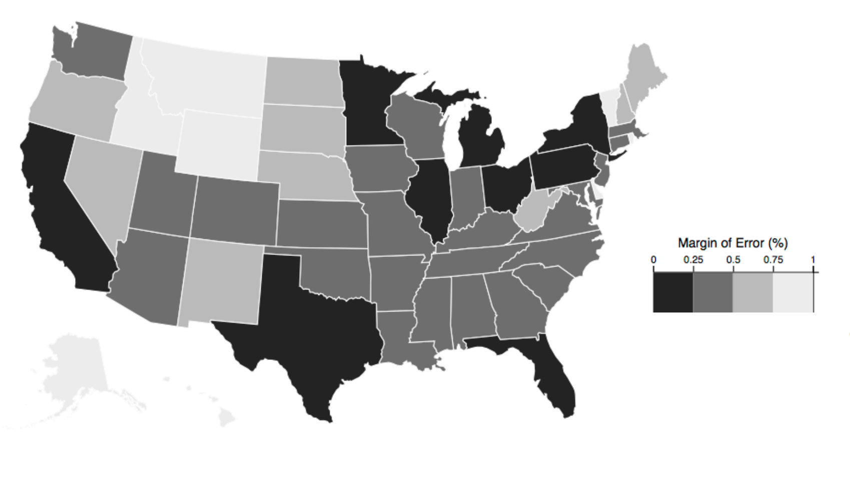
1) Quantify Uncertainty
2) Choose a free visual variable
3) Encode uncertainty with the variable
4) Unify the Data Map and Uncertainty Map

How to Unify?

Data Map

Uncertainty Map

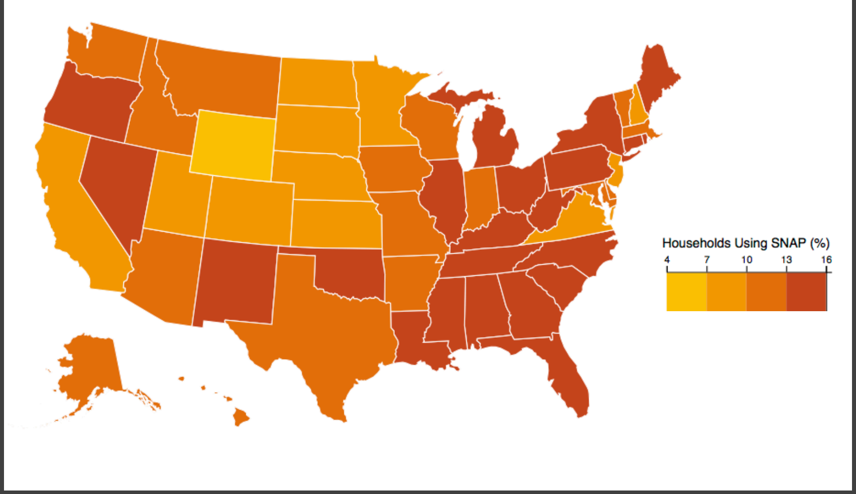


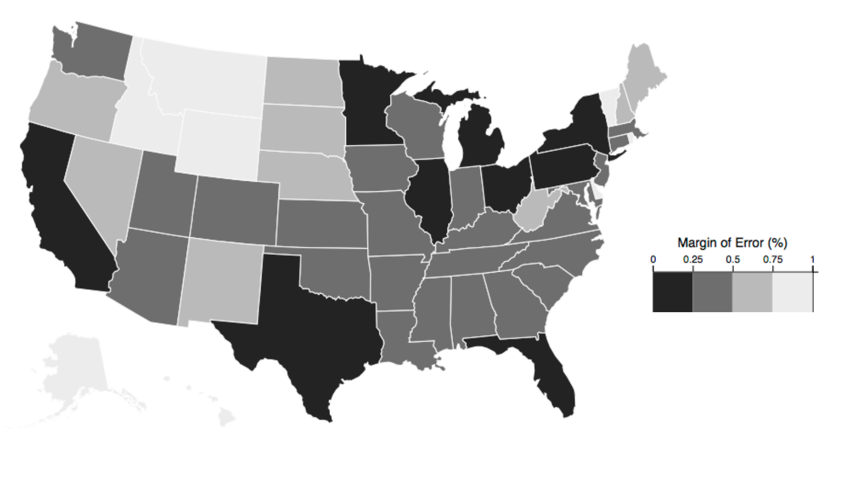


Juxtaposition

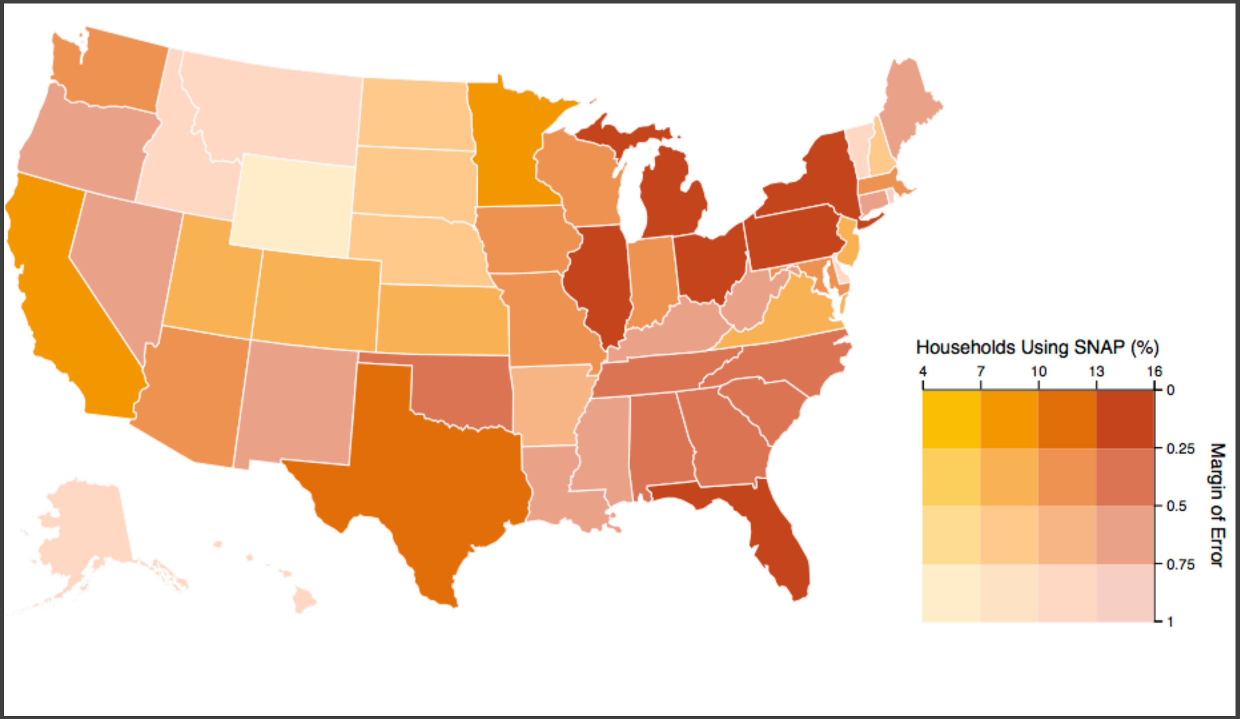
Data Map

Uncertainty Map

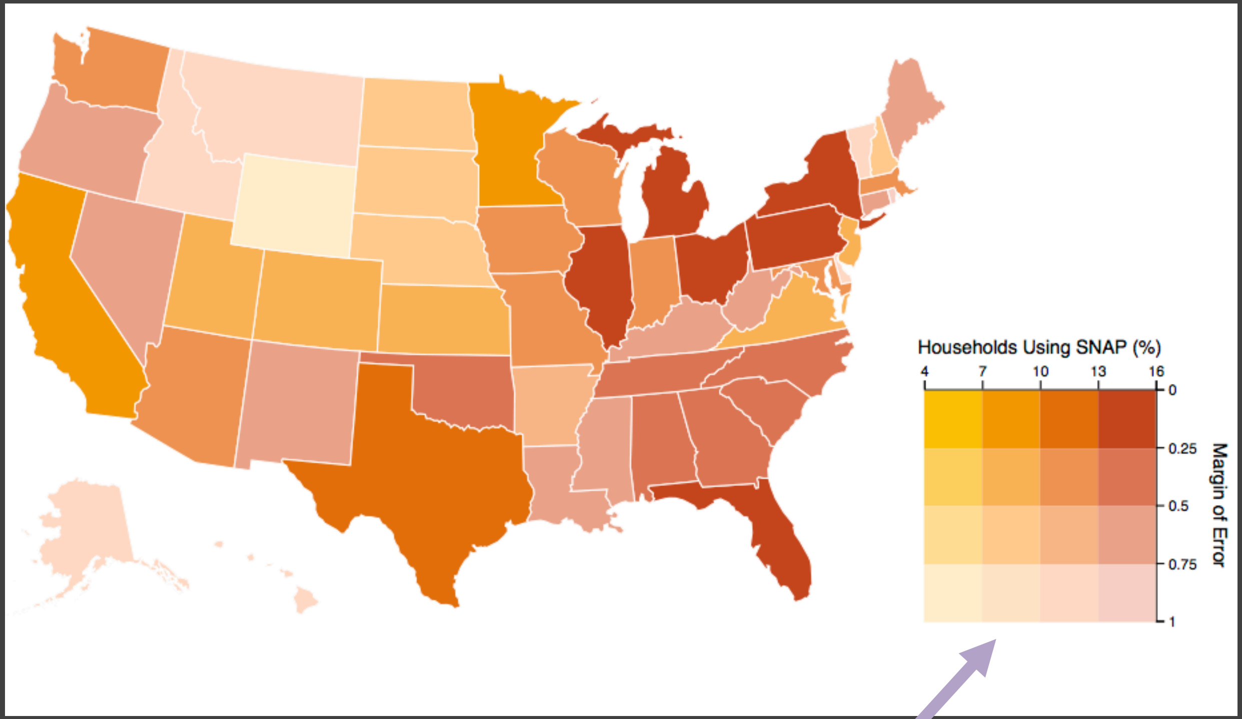




Superposition

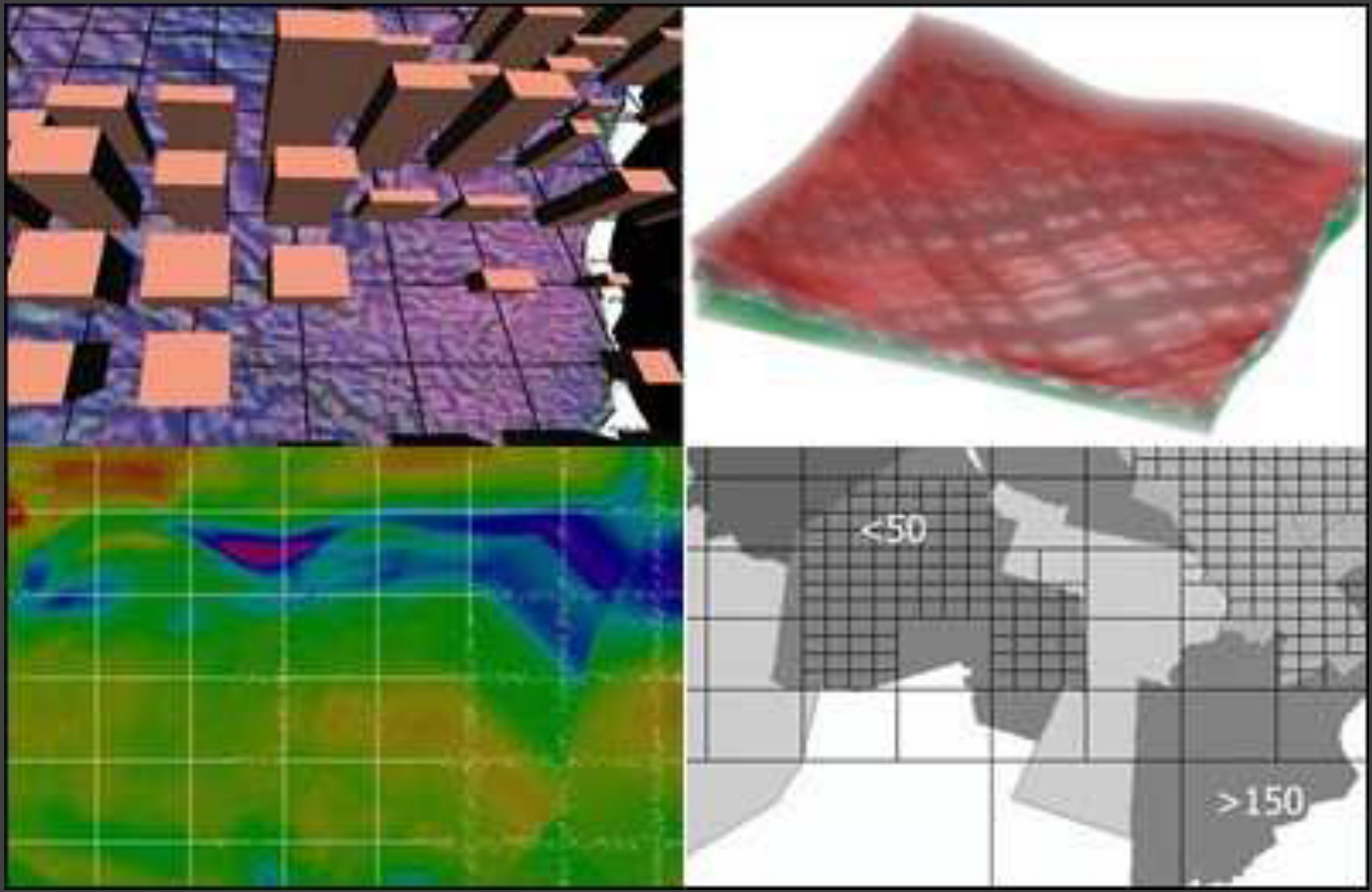


Superposition



Bivariate Map

Superposition



Griethe, Henning and Schumann, Heidrun. The Visualization of Uncertain Data: Methods and Problems. SimVis, 2006.

Uncertainty Vis Pipeline

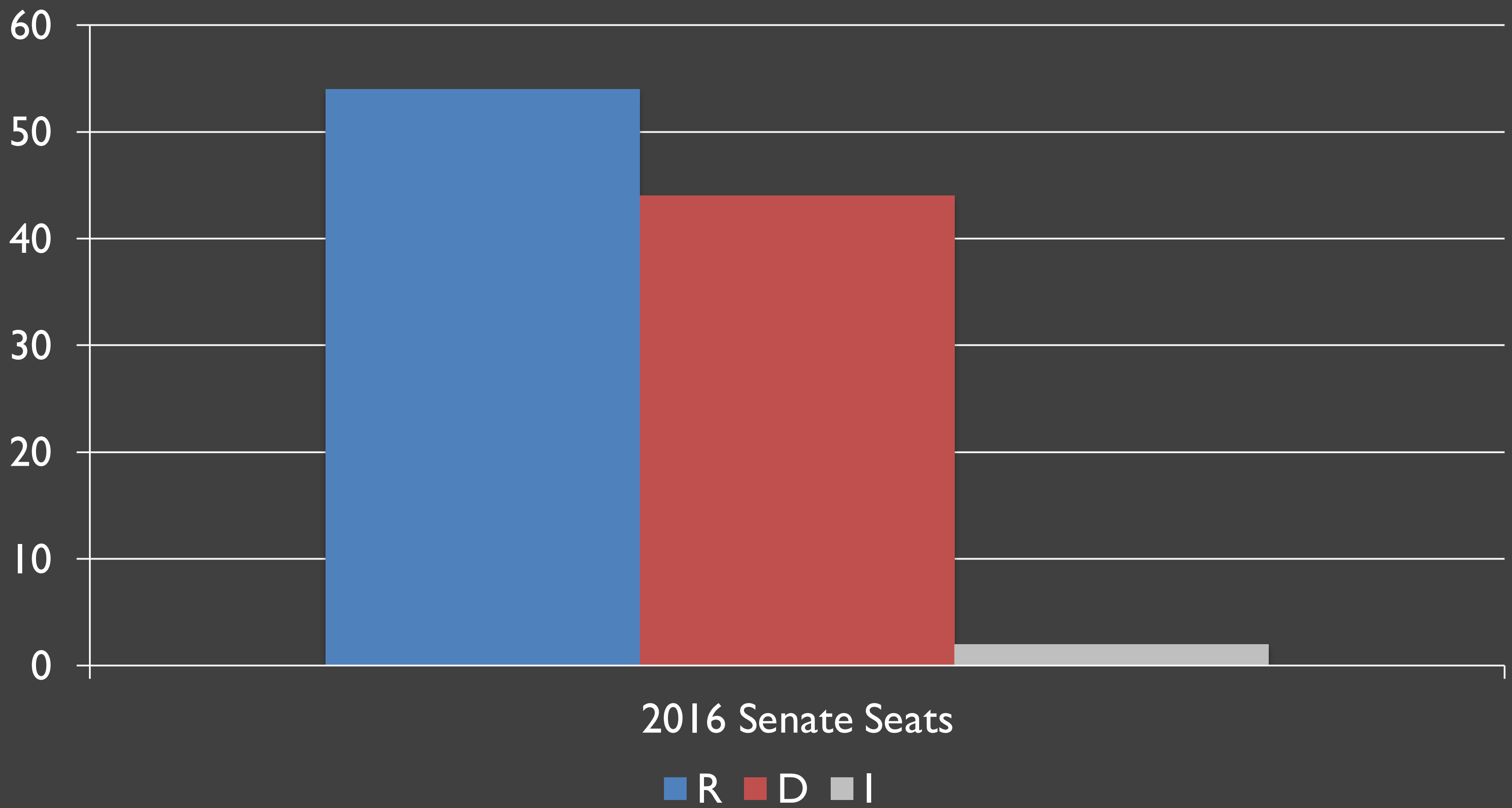
1) Quantify Uncertainty
2) Choose a free visual variable
3) Encode uncertainty with the variable
4) Unify the Data Map and Uncertainty Map

Uncertainty Vis Pipeline

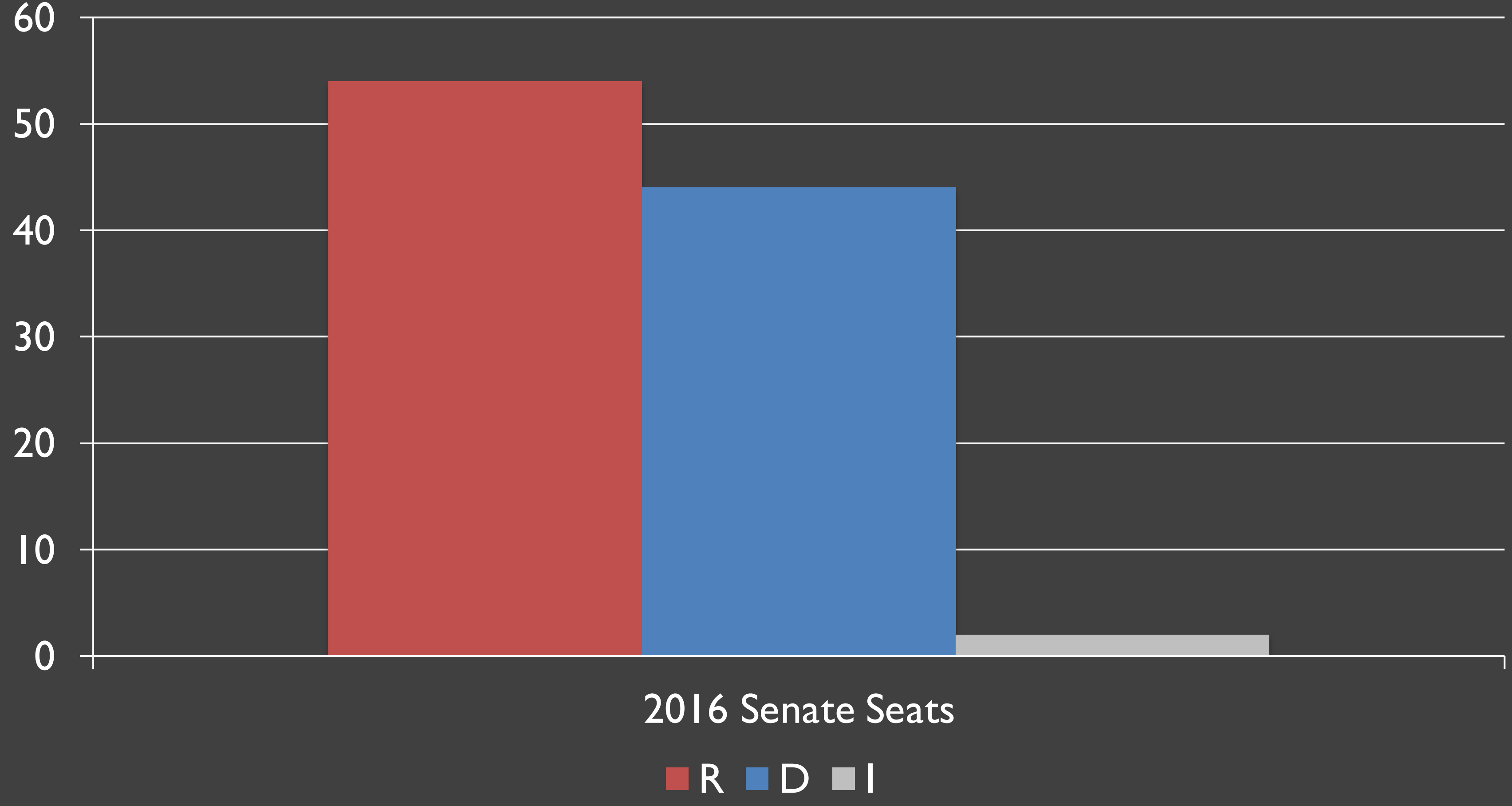
1) Quantify Uncertainty
2) Choose a free **visual variable**
3) Encode uncertainty with the variable
4) Unify the Data Map and Uncertainty Map

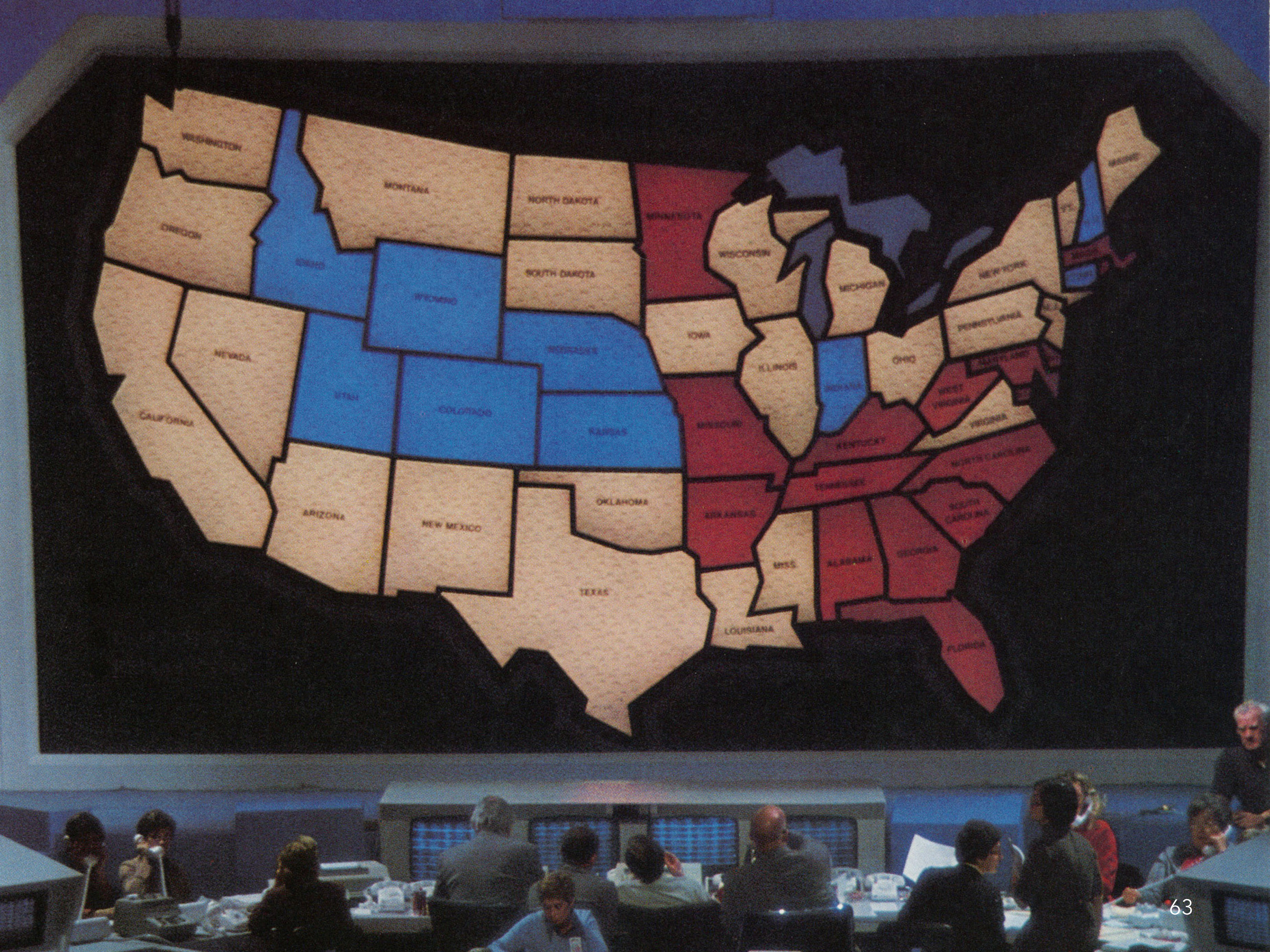
Semiotics of Uncertainty

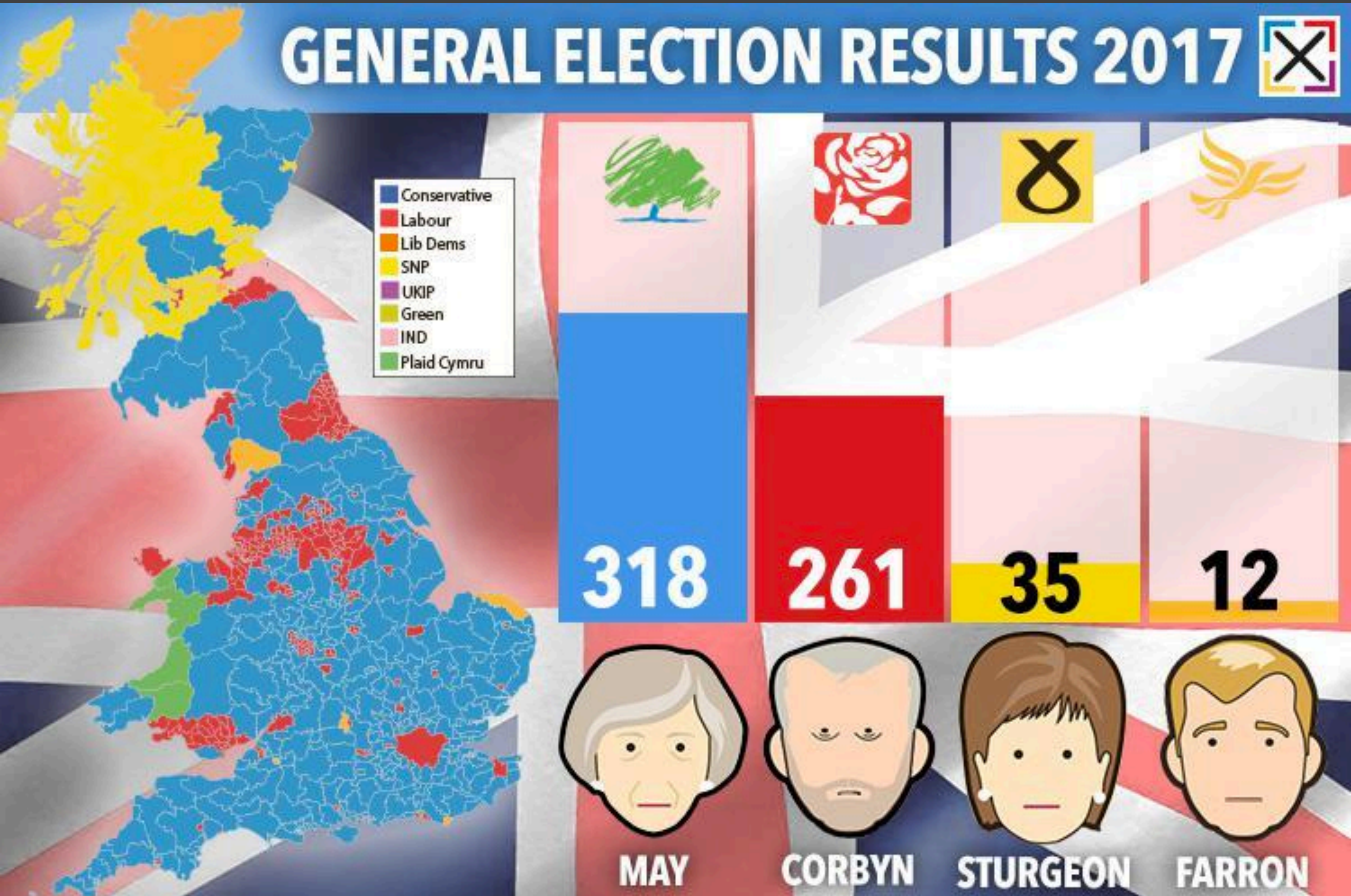
The Variable Matters!

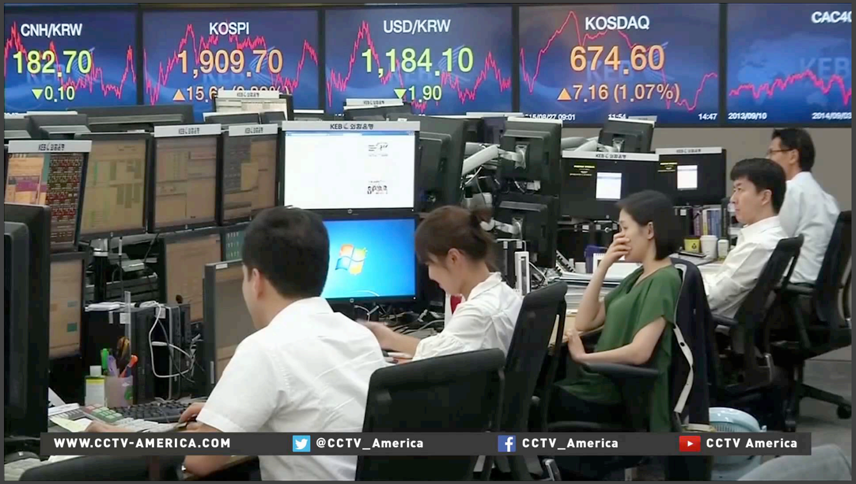


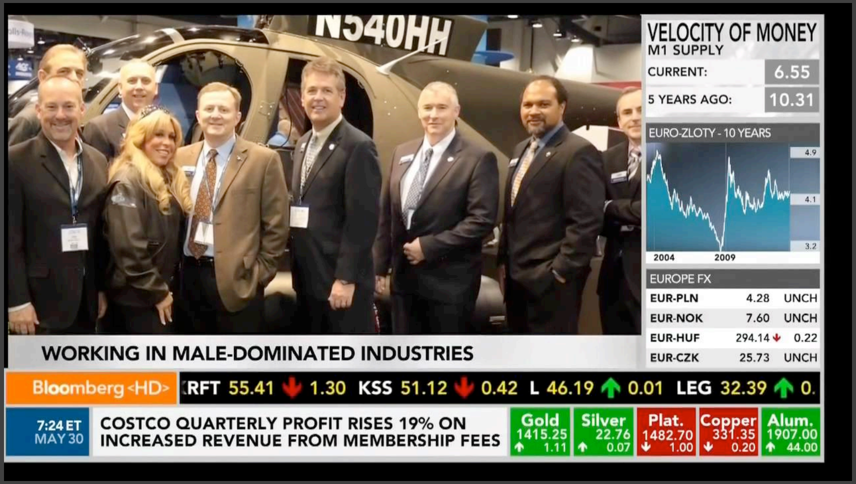
The Variable Matters!




WASHINGTON
OREGON
MONTANA
NORTH DAKOTA
SOUTH DAKOTA
MINNESOTA
WISCONSIN
MICHIGAN
NEW YORK
PENNSYLVANIA
OHIO
ILLINOIS
IOWA
NEVADA
UTAH
COLORADO
KANSAS
CALIFORNIA
ARIZONA
NEW MEXICO
TEXAS
OKLAHOMA
MISSOURI
ARKANSAS
KENTUCKY
TENNESSEE
WEST VIRGINIA
VIRGINIA
NORTH CAROLINA
SOUTH CAROLINA
GEORGIA
ALABAMA
MISS.
LOUISIANA
FLORIDA

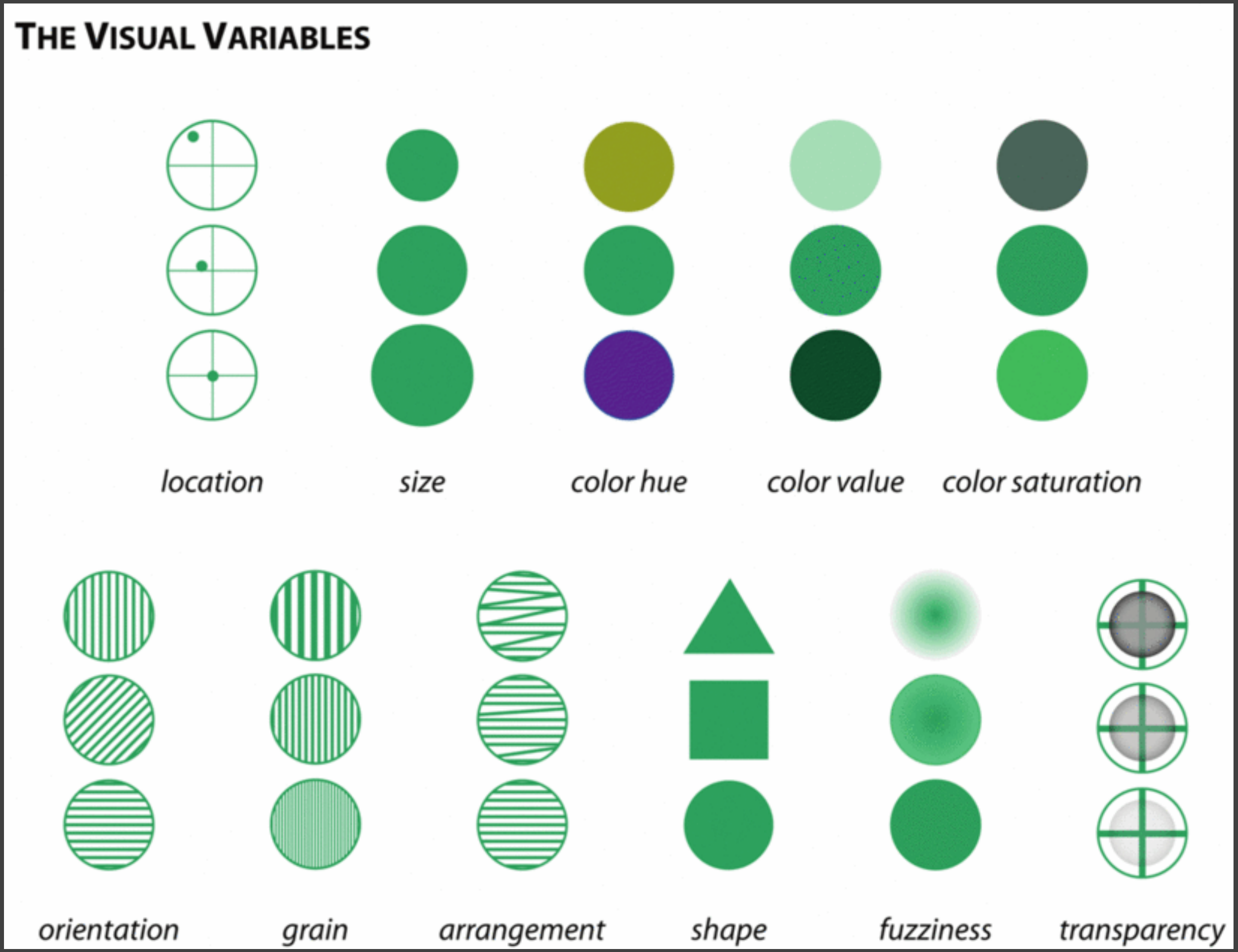

GENERAL ELECTION RESULTS 2017
Conservative
Labour
Lib Dems
SNP
UKIP
Green
IND
Plaid Cymru
318
261
35
12
MAY
CORBYN
STURGEON
FARRON


CNH/KRW
182.70
▼0.10
KOSPI
1,909.70
USD/KRW
1,184.10
▼1.90
KOSDAQ
674.60
▲7.16 (1.07%)
WWW.CCTV-AMERICA.COM
@CCTV_America
CCTV_America
CCTV America

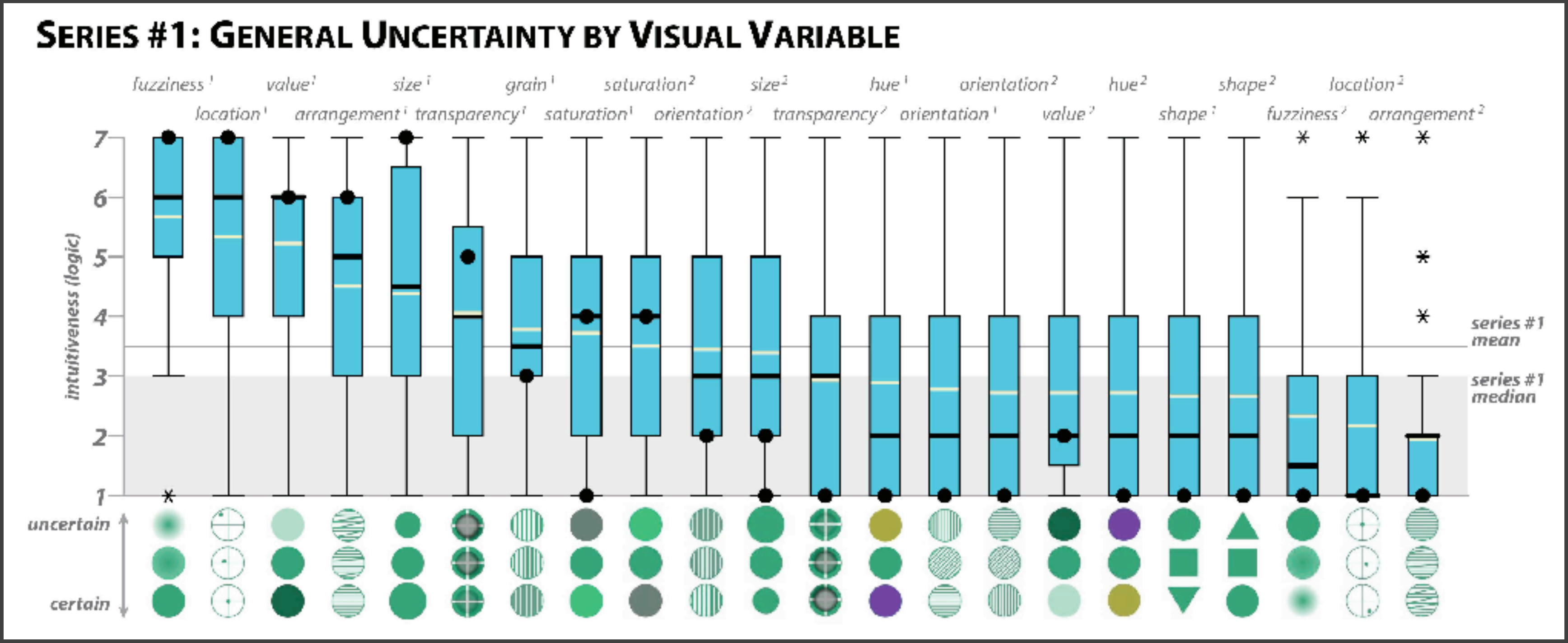

N540HH
VELOCITY OF MONEY
M1 SUPPLY
CURRENT: 6.55
5 YEARS AGO: 10.31
EURO-ZLOTY - 10 YEARS
2004
2009
EUROPE FX
EUR-PLN 4.28 UNCH
EUR-NOK 7.60 UNCH
EUR-HUF 294.14 0.22
EUR-CZK 25.73 UNCH
WORKING IN MALE-DOMINATED INDUSTRIES
Bloomberg <HD>
RFT 55.41 1.30 KSS 51.12 0.42 L 46.19 0.01 LEG 32.39
7:24 ET MAY 30
COSTCO QUARTERLY PROFIT RISES 19% ON INCREASED REVENUE FROM MEMBERSHIP FEES
Gold 1415.25 1.11
Silver 22.76 0.07
Plat. 1482.70 1.00
Copper 331.35 0.20
Alum. 1907.00 44.00

Semiotics of Uncertainty

Semiotics of Uncertainty

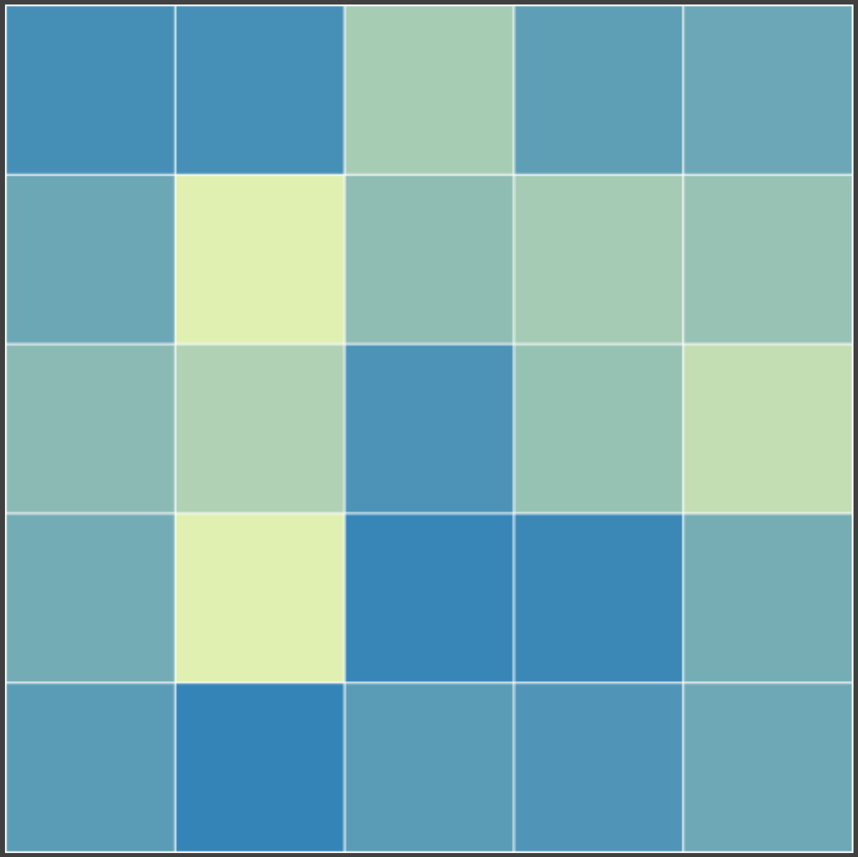


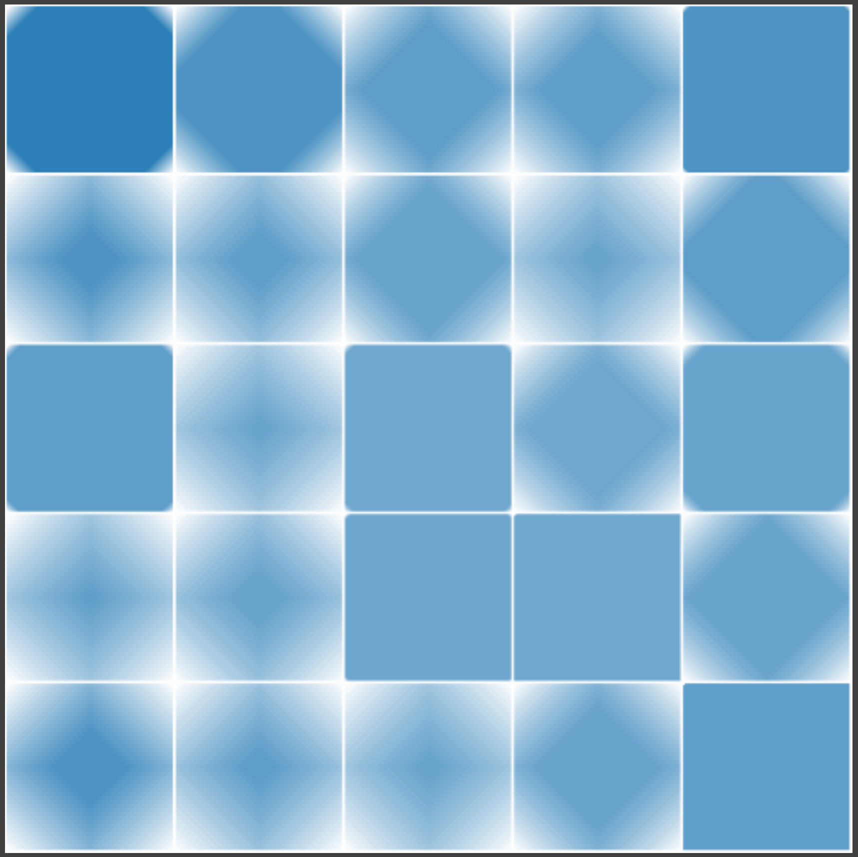
MacEachren, Alan et al. Visual Semiotics & Uncertainty Visualization: An empirical study. IEEE VIS, 2012.



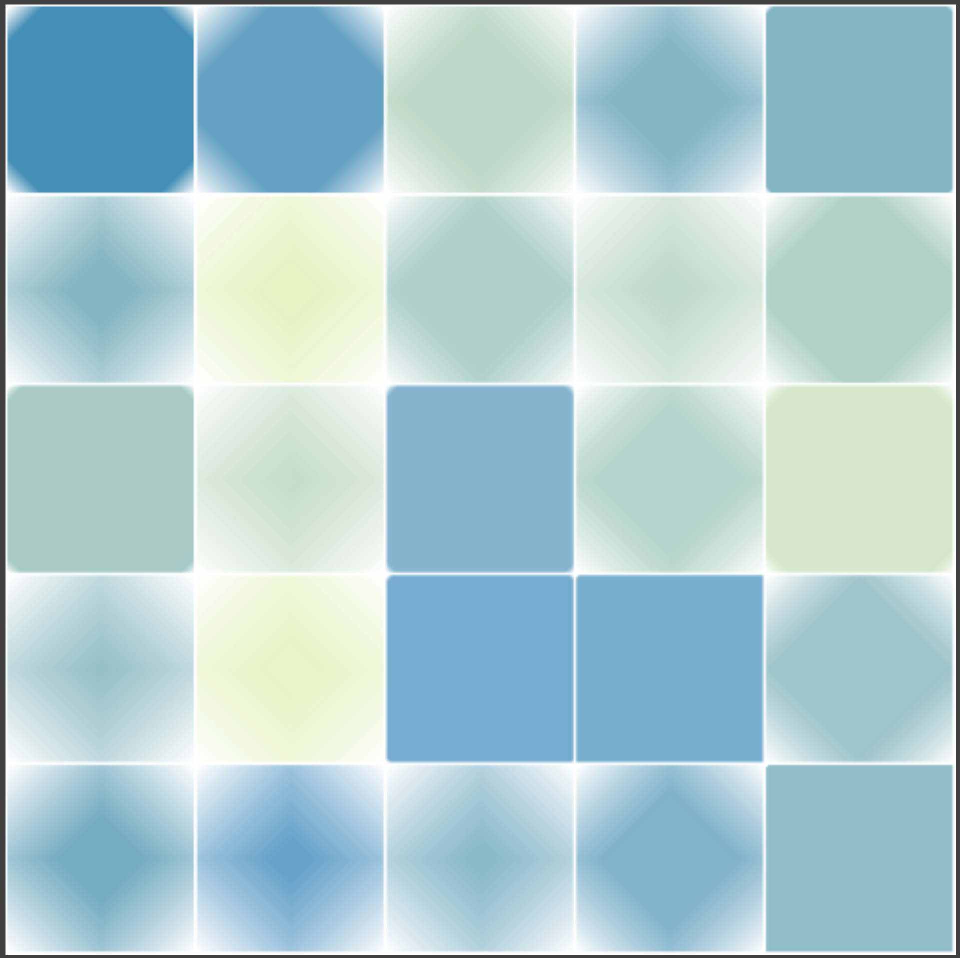
SERIES #1: GENERAL UNCERTAINTY BY VISUAL VARIABLE
fuzziness 1
location 1
value 1
arrangement 1
size 1
transparency 1
grain 1
saturation 1
saturation 2
orientation 2
size 2
transparency 2
hue 1
orientation 1
orientation 2
value 2
hue 2
shape 1
shape 2
fuzziness 2
location 2
arrangement 2
intuitiveness (logic)
7
6
5
4
3
2
1
uncertain
certain
series #1 mean
series #1 median

Fuzziness Juxtaposition

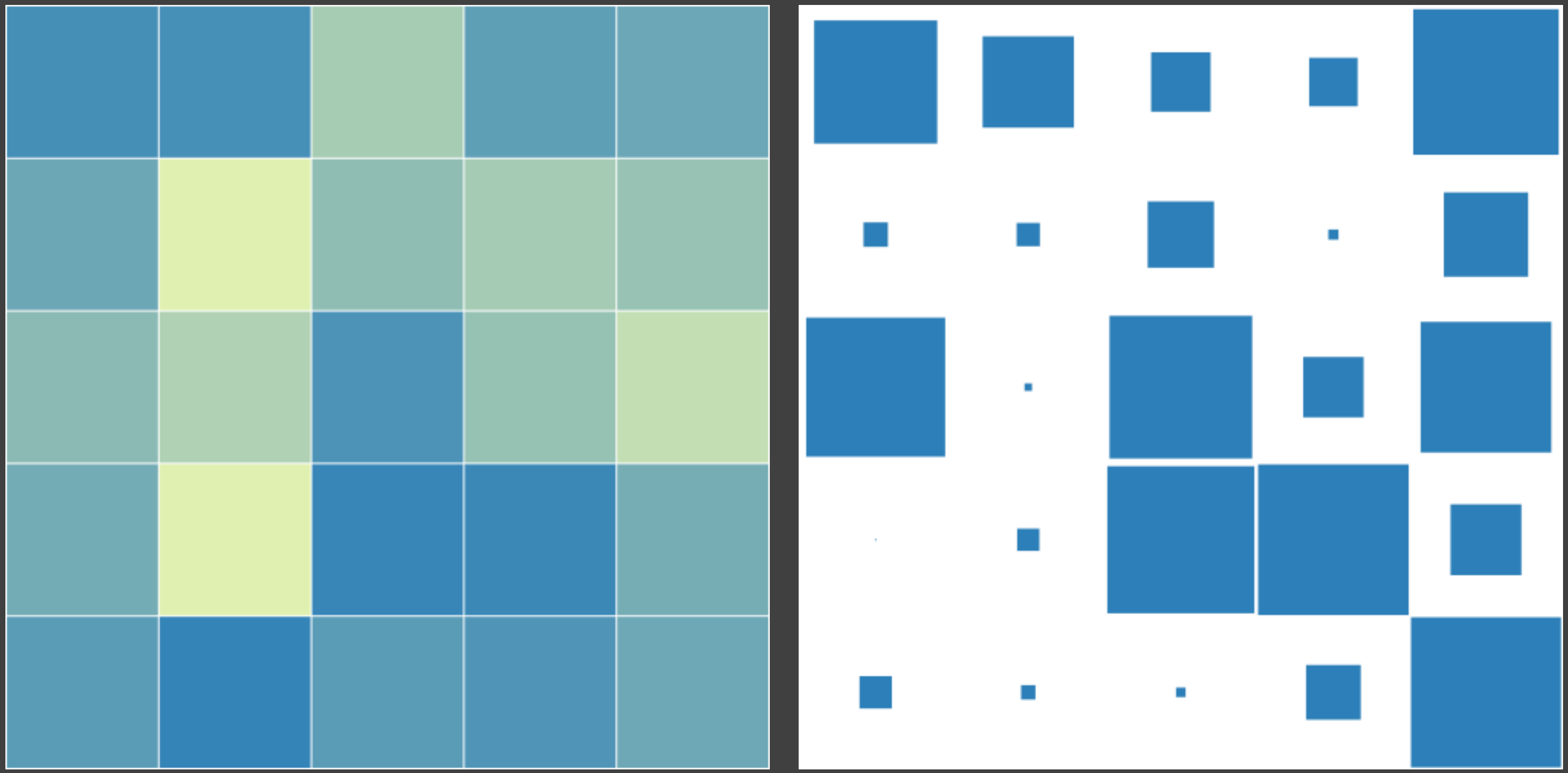




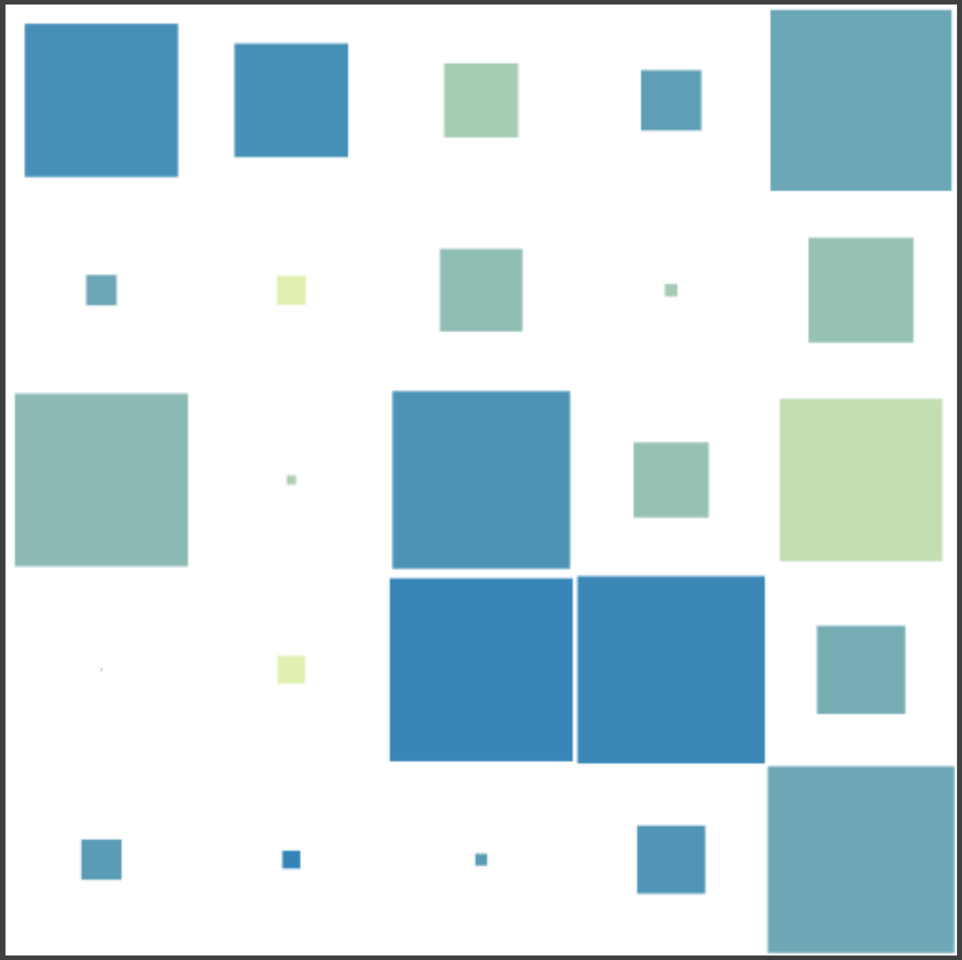
Fuzziness Superposition



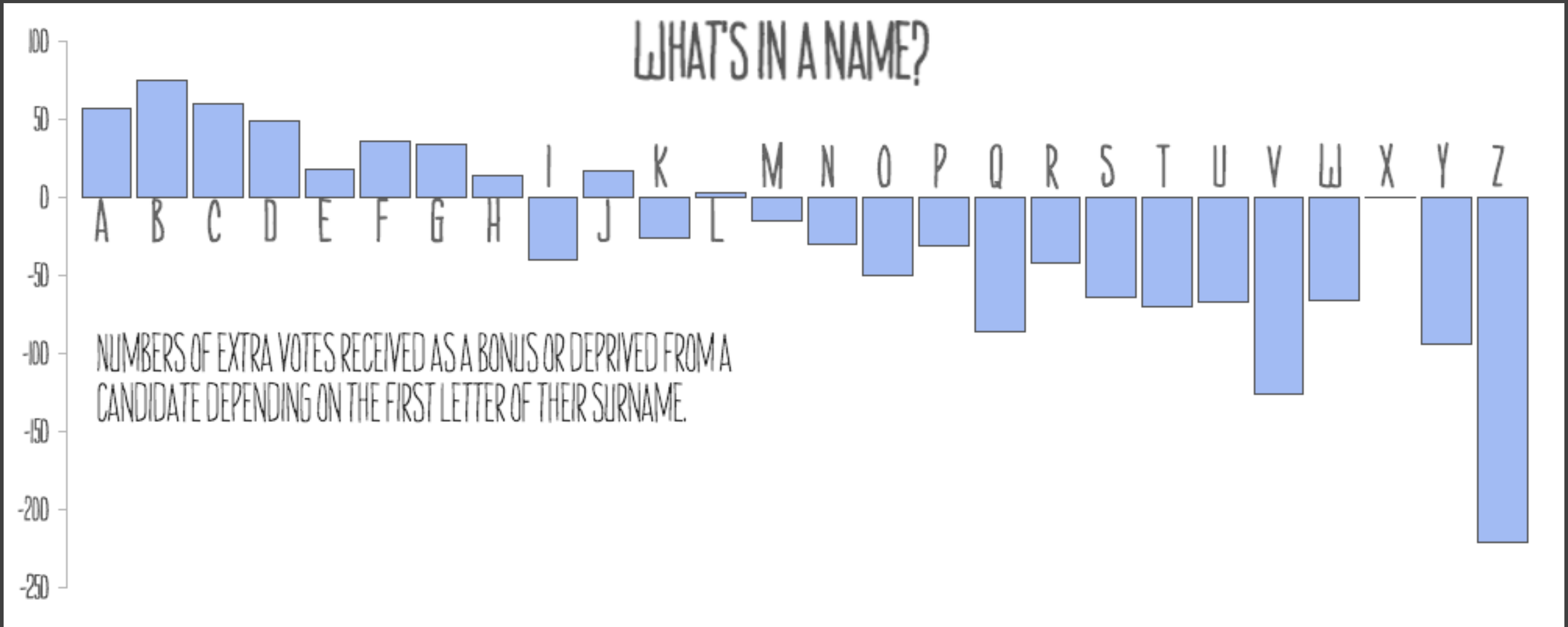
Size Juxtaposition



Size Superposition



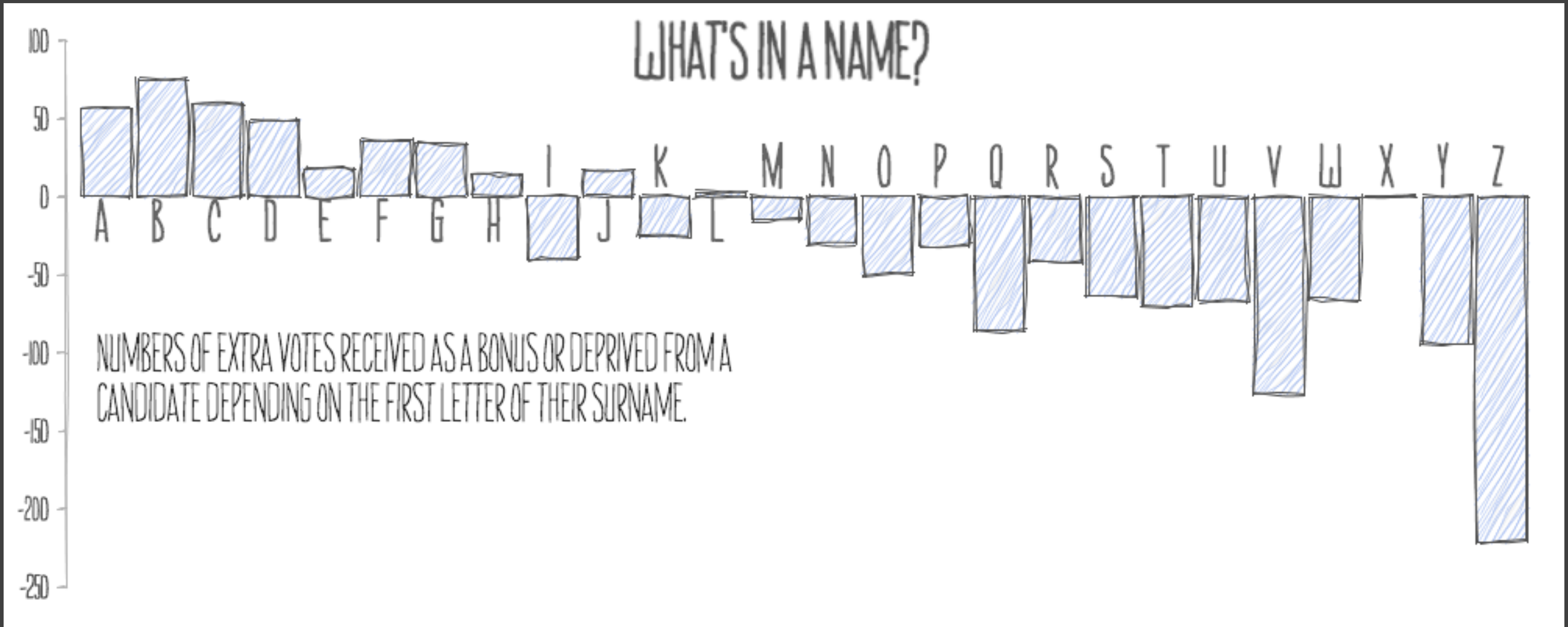
"Sketchiness"



Wood, Jo et al. Sketchy rendering for information visualization. IEEE VIS, 2012.

Boukhelifa, Nadia et al. Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. IEEE VIS, 2012.

“Sketchiness”



Wood, Jo et al. Sketchy rendering for information visualization. IEEE VIS, 2012.

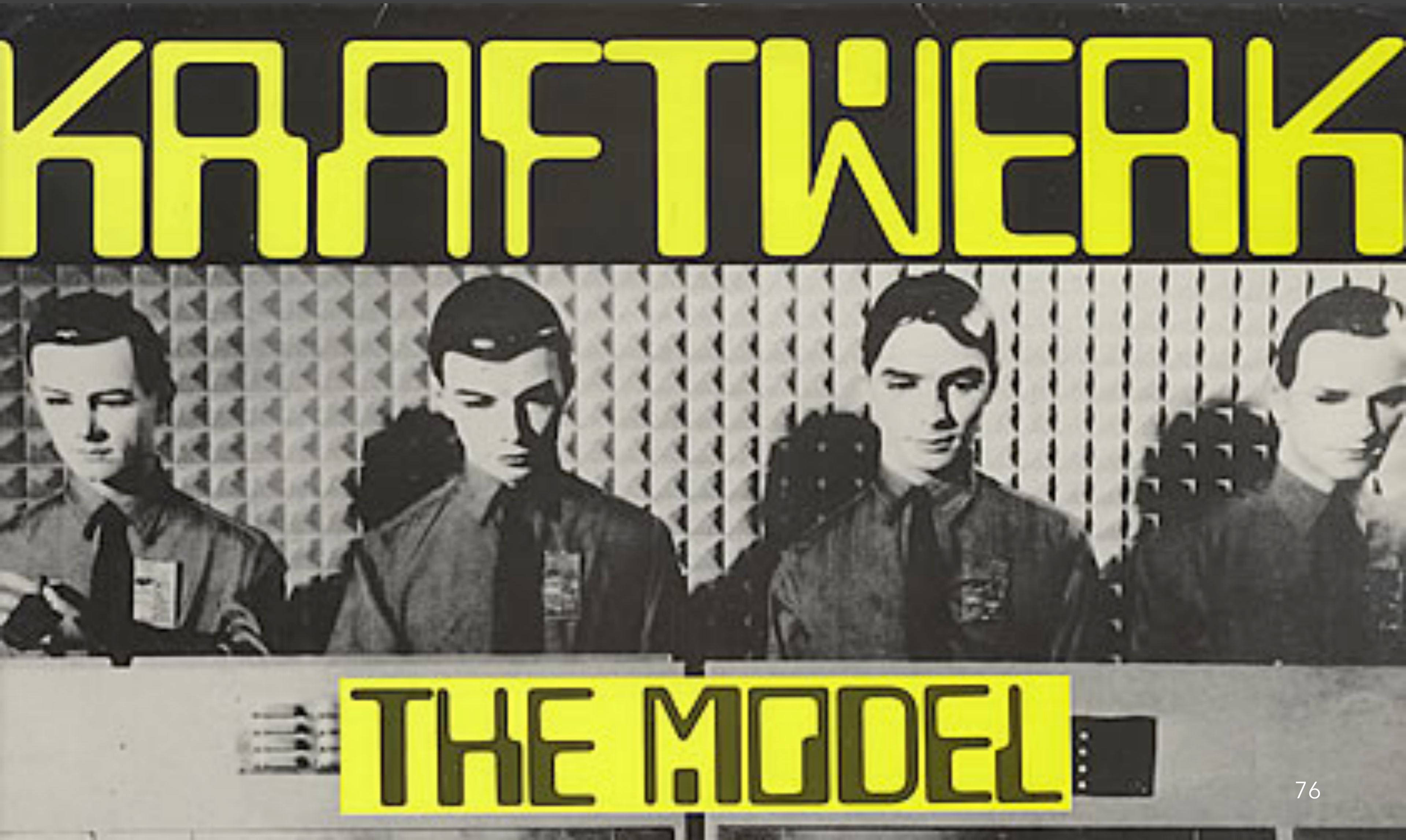
Boukhelifa, Nadia et al. Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. IEEE VIS, 2012.

Encoding Uncertainty

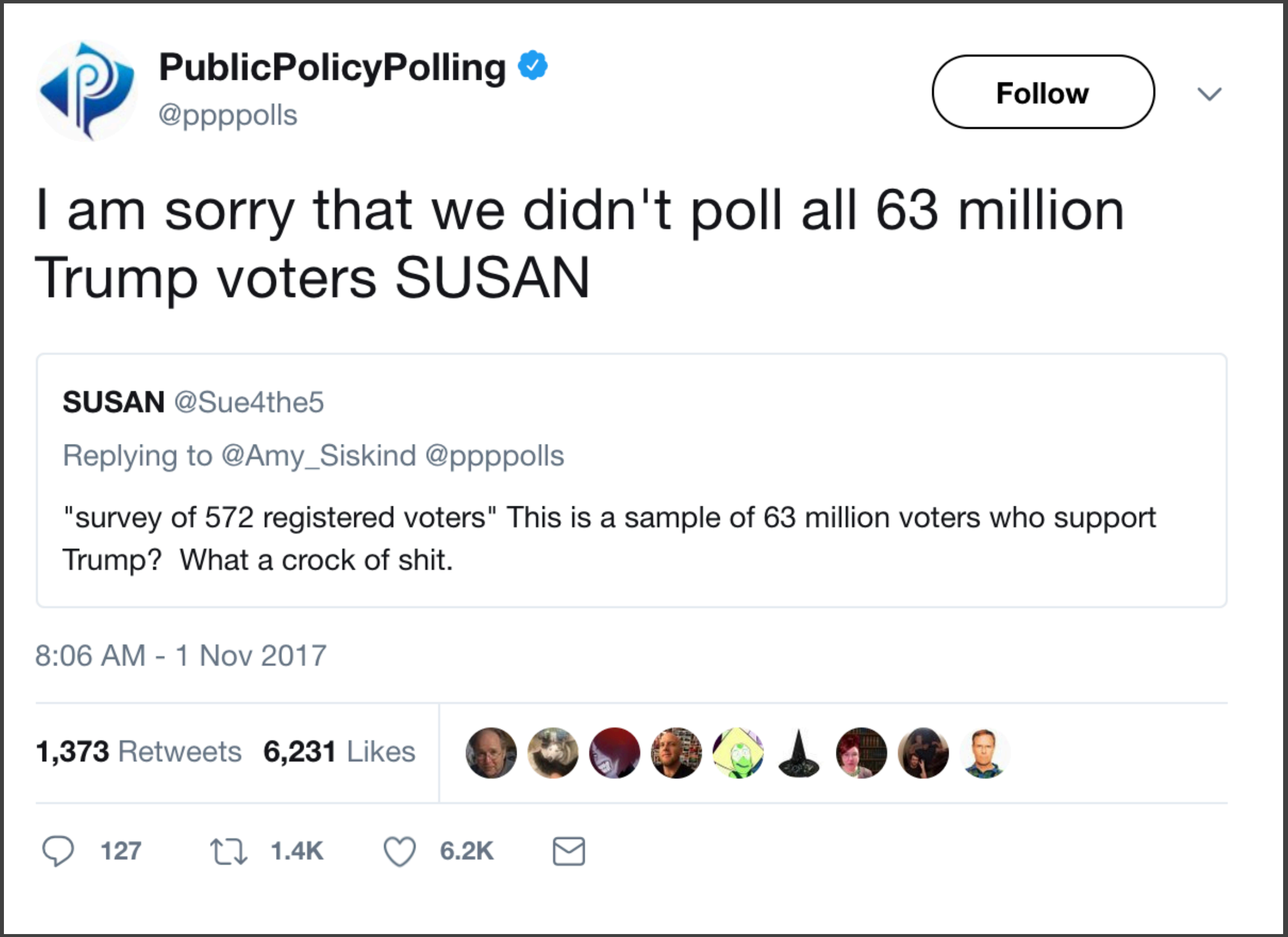
Some visual variables (like fuzziness and value) have a **semiotic connection** to uncertainty.

However, intuitive variables may not always be accurately interpreted!

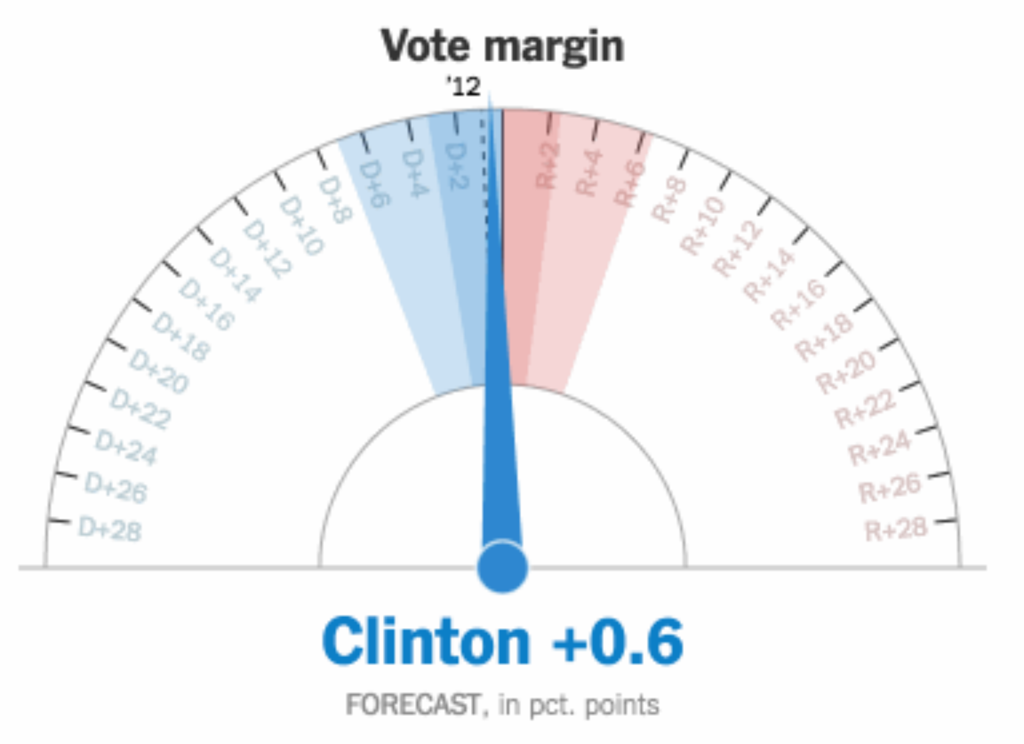
Model Visualization



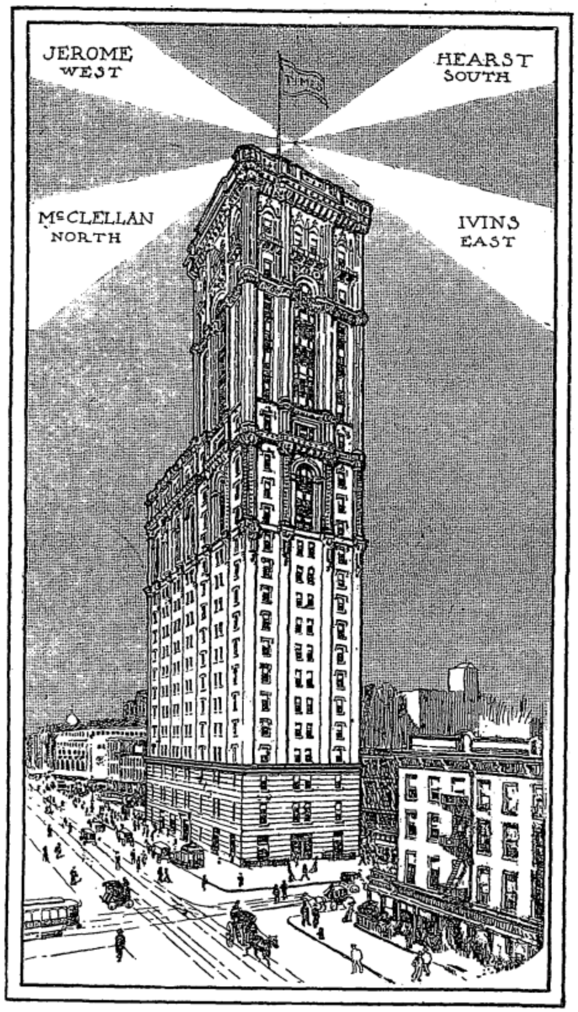
Polling Data



The NYT Needle



ELECTION RESULTS BY SEARCHLIGHT.



The Times Election Searchlight Code.

News Will Be Flashed from the Tower of The Times Building on Tuesday Night.

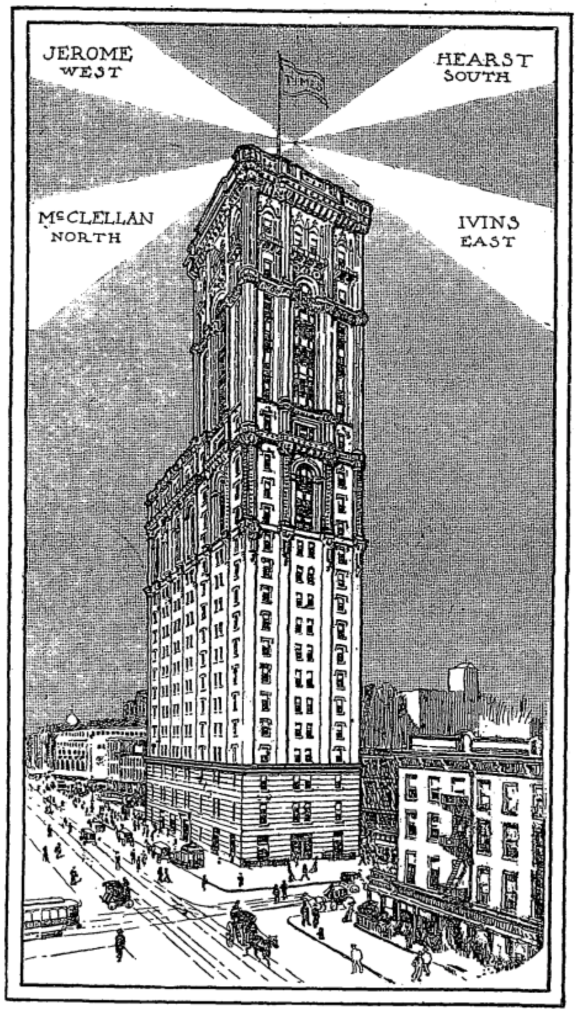
The results of the election next Tuesday night will be flashed by electric light from the tower of the Times Building, so that for miles around people will be able to tell which of the candidates has won.

This will be entirely separate and distinct from the elaborate bulletin service which THE TIMES will also maintain. To display the detailed bulletins so that the crowds can see them easily and comfortably, a stereopticon machine will be set up in the triangle north of the Times Building and the bulletins displayed on canvas stretched from the north side of the building. There will be a similar service at the Harlem office of THE TIMES, 129 West 125th Street.

The electric signals from the tower of the Times Building will be flashed from a point 365 feet above the street level. A steady light to the north will show that McClellan has been elected; a steady light to the east will indicate Ivins's election, and a steady light to the south will indicate that Hearst has won.

Jerome's election will be indicated by a steady light to the west. A light to the north, waving from east to west, will indicate Osborne's election. A light to the south, waving from east to west, will indicate Shearn's election.

ELECTION RESULTS BY SEARCHLIGHT.



The Times Election Searchlight Code.

News Will Be Flashed from the Tower of The Times Building on Tuesday Night.

The results of the election next Tuesday night will be flashed by electric light from the tower of the Times Building, so that for miles around people will be able to tell which of the candidates has won.

This will be entirely separate and distinct from the elaborate bulletin service which THE TIMES will also maintain. To display the detailed bulletins so that the crowds can see them easily and comfortably, a stereopticon machine will be set up in the triangle north of the Times Building and the bulletins displayed on canvas stretched from the north side of the building. There will be a similar service at the Harlem office of THE TIMES, 129 West 125th Street.

The electric signals from the tower of the Times Building will be flashed from a point 365 feet above the street level. A steady light to the north will show that McClellan has been elected; a steady light to the east will indicate Ivins's election, and a steady light to the south will indicate that Hearst has won.

Jerome's election will be indicated by a steady light to the west. A light to the north, waving from east to west, will indicate Osborne's election. A light to the south, waving from east to west, will indicate Shearn's election.

Election Bulletins

BY BOMBS.

TUESDAY NIGHT

THE TRIBUNE

will send up from the roof of the

GREAT NORTHERN HOTEL

hourly, shells containing blue and red stars—exactly on the hour—at 7, 8, 9, 10, 11 p. m. 12 midnight, 1 and 2 a. m. Wednesday morning, unless election is decided earlier, in which case twelve bombs will be sent up in rapid succession. Blue to indicate McKinley's election. Red to indicate Bryan's election.

SIX BOMBS EVERY HOUR.

The first bomb sent up, if blue, indicates the returns in **COOK COUNTY** at that hour are favorable to McKinley; if red, favorable to Bryan.

After sixty seconds two bombs will be sent up in rapid succession, and will indicate, if blue, that returns from **ILLINOIS** favor McKinley; if red, Bryan.

After sixty seconds more three bombs will be sent up in rapid succession, and if blue will indicate that at that hour returns from the **entire country** favor McKinley; if red, Bryan. Each bomb bursts high in the air, scattering a shower of stars.

Polling Data

Candidate A is ahead of Candidate B in the polls, with 55% of the likely voters*

Polling Data

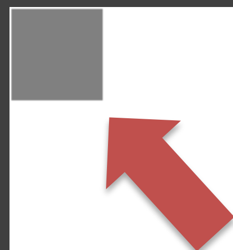
Candidate A is ahead of Candidate B in the polls, with 55% of the likely voters*

*poll of 100 people, margin of error +/-5

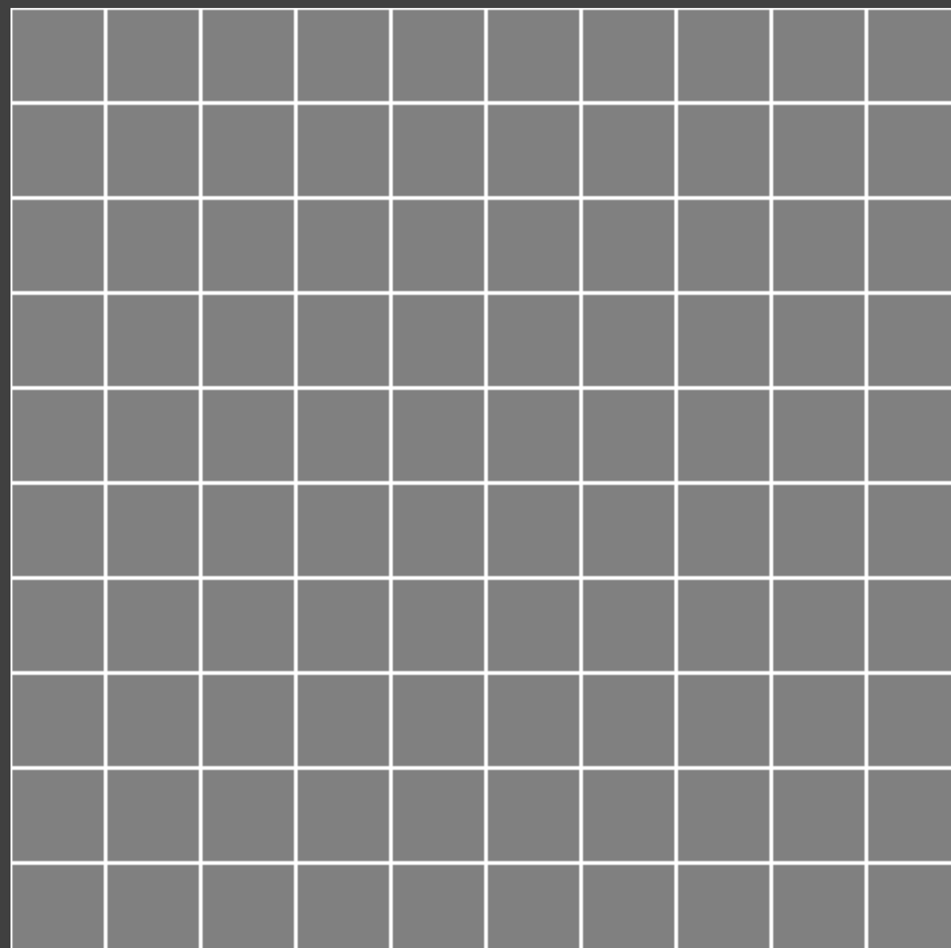
Monte Carlo Approach

Candidate A is ahead of Candidate B in the polls, with 55% of the likely voters*

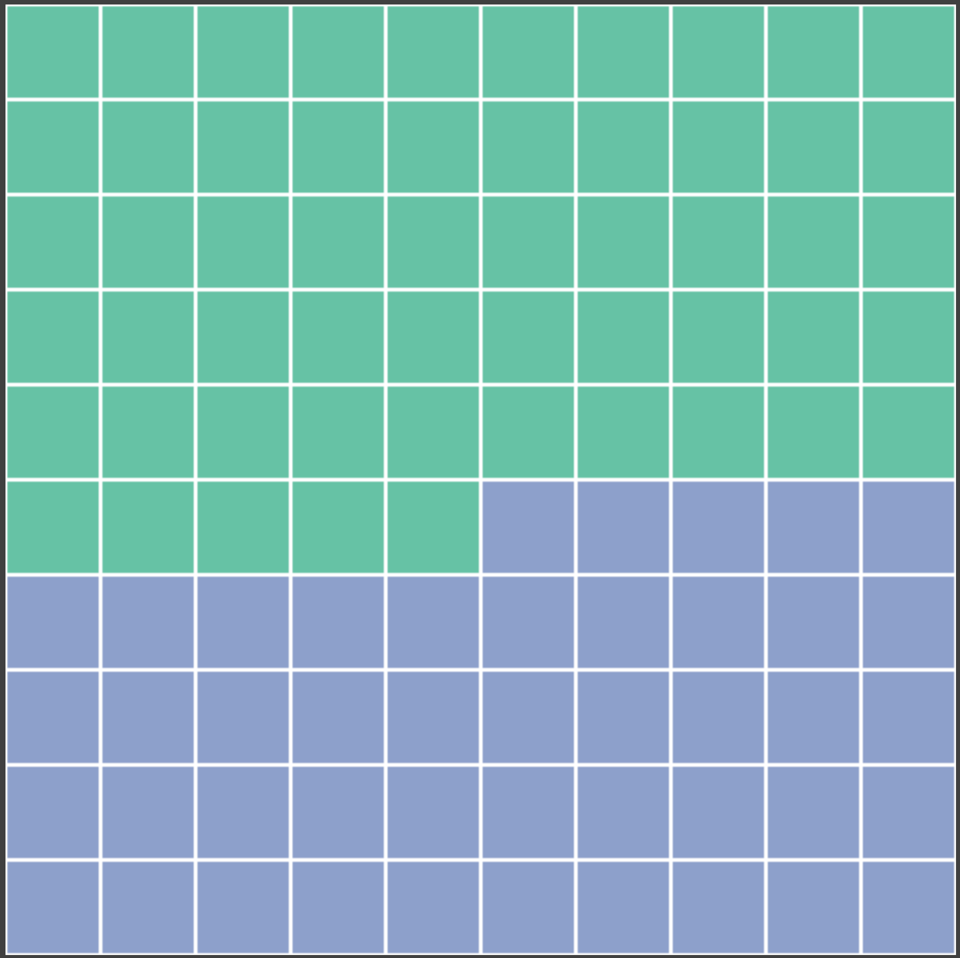
*poll of 100 people, margin of error +/-5



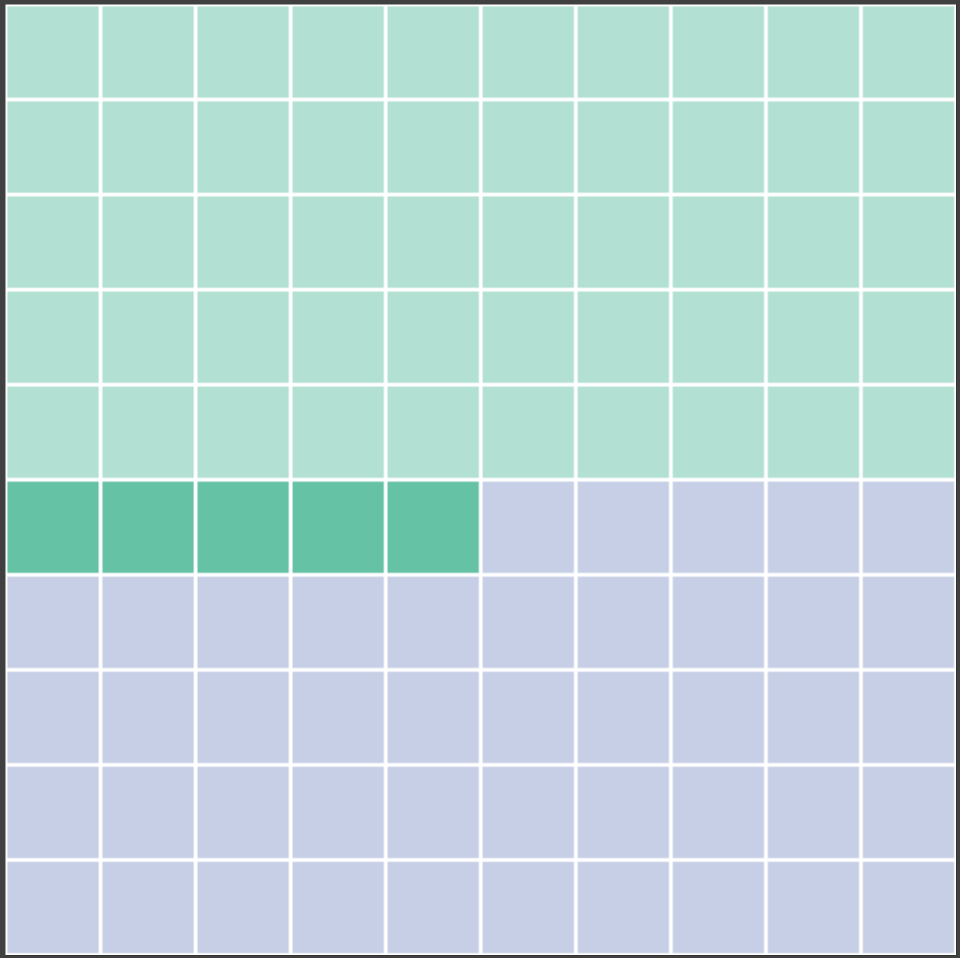
A Likely Voter



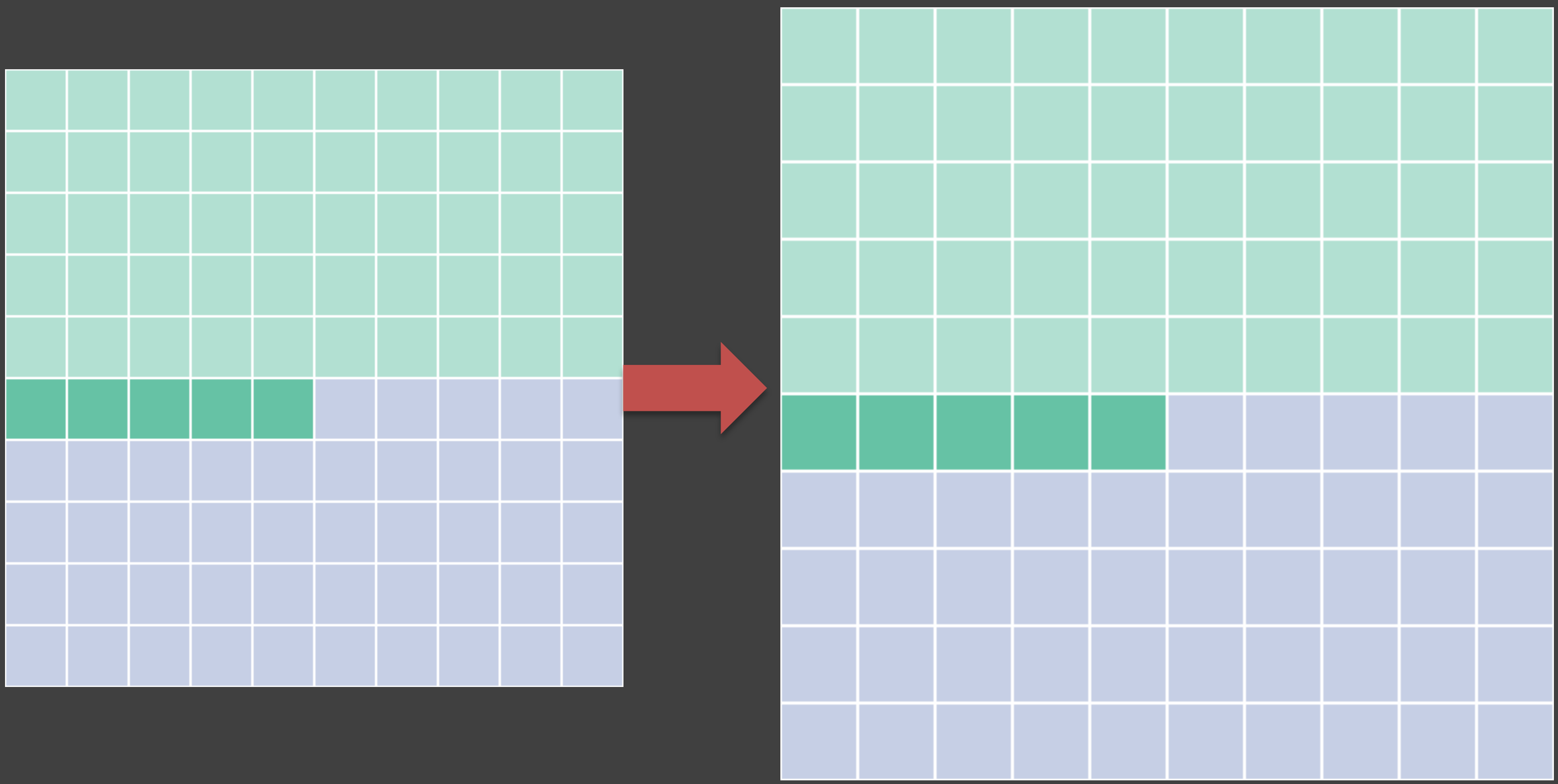
Poll



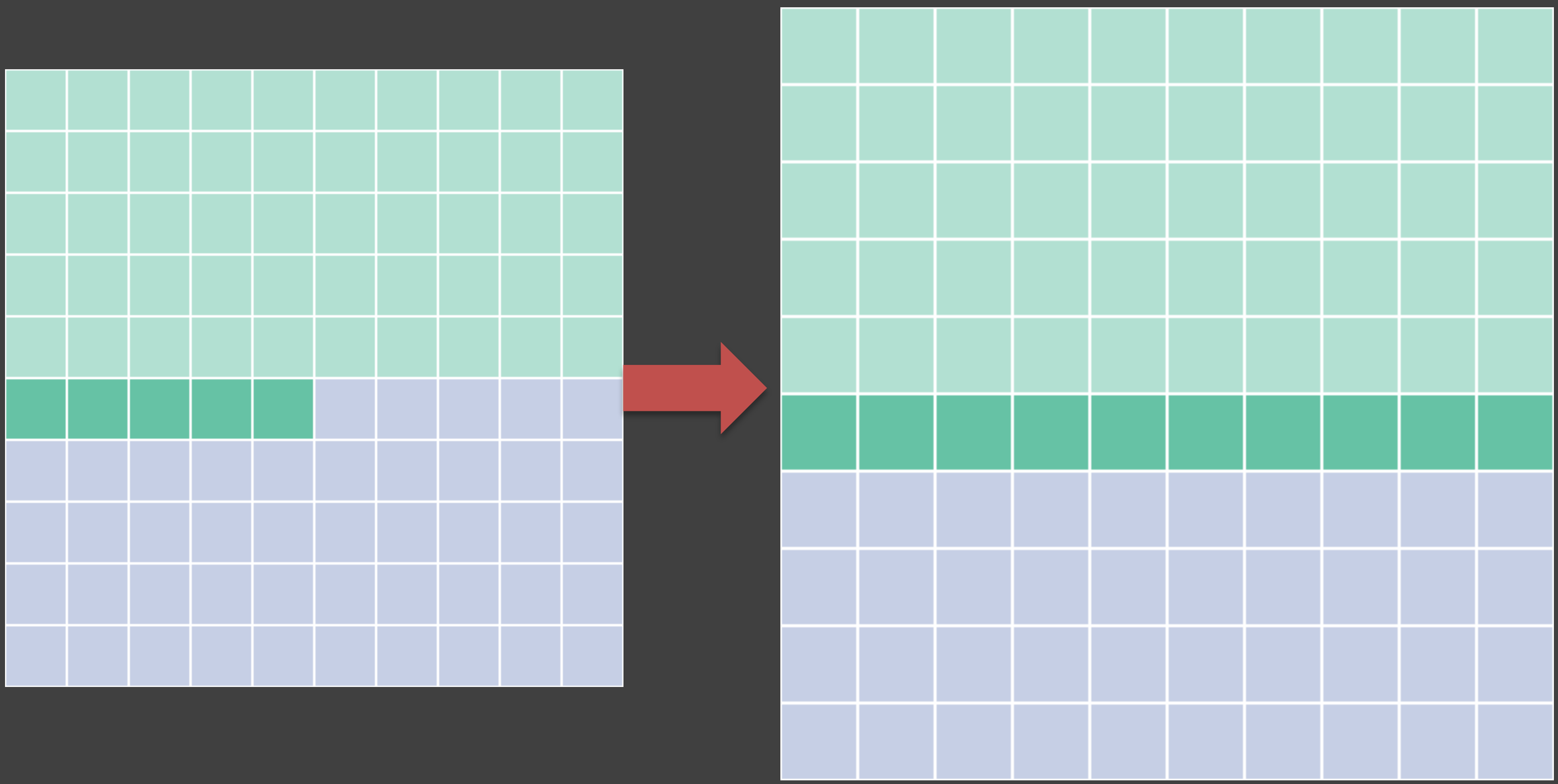
Poll



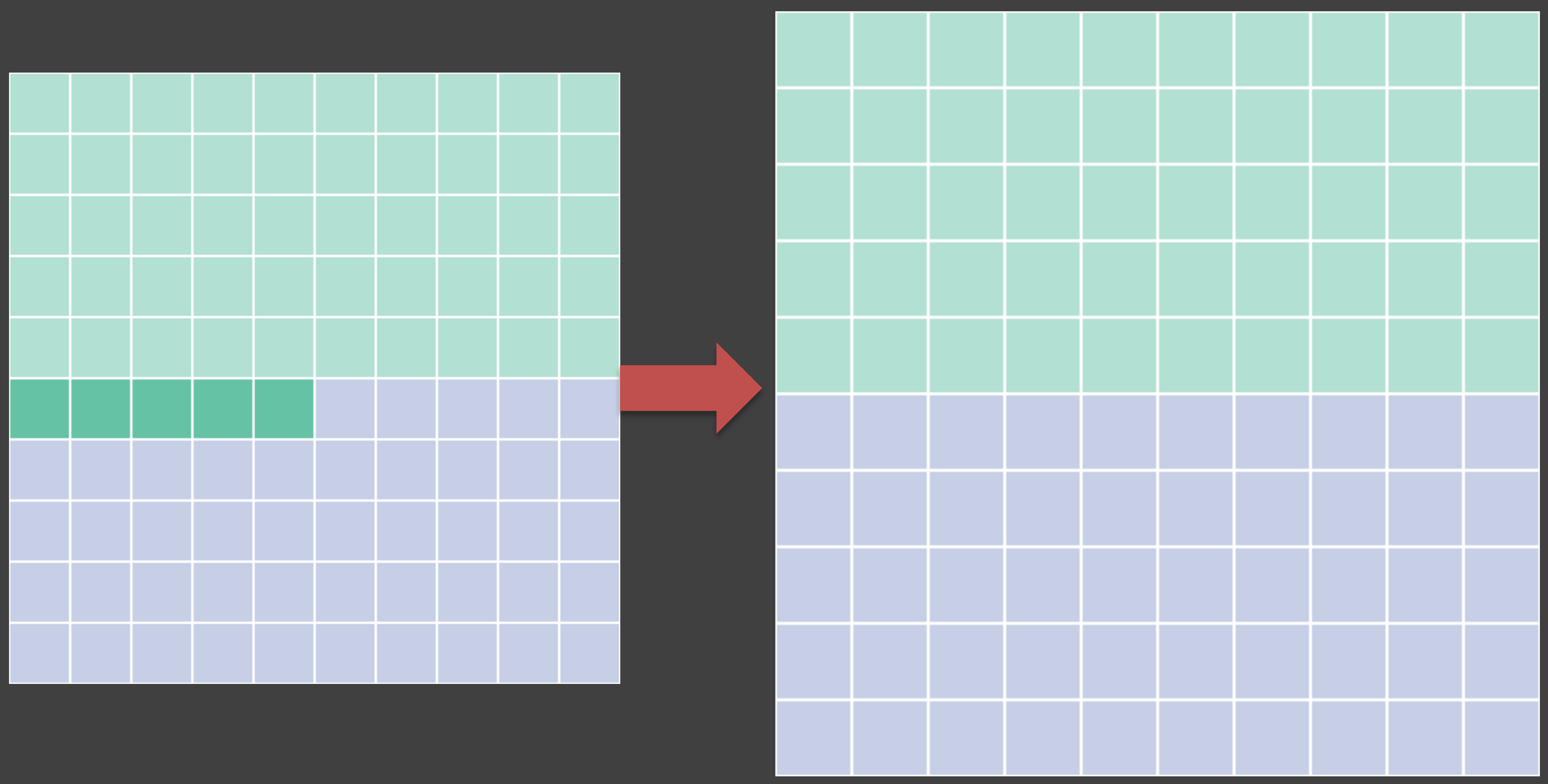
Actual Election?



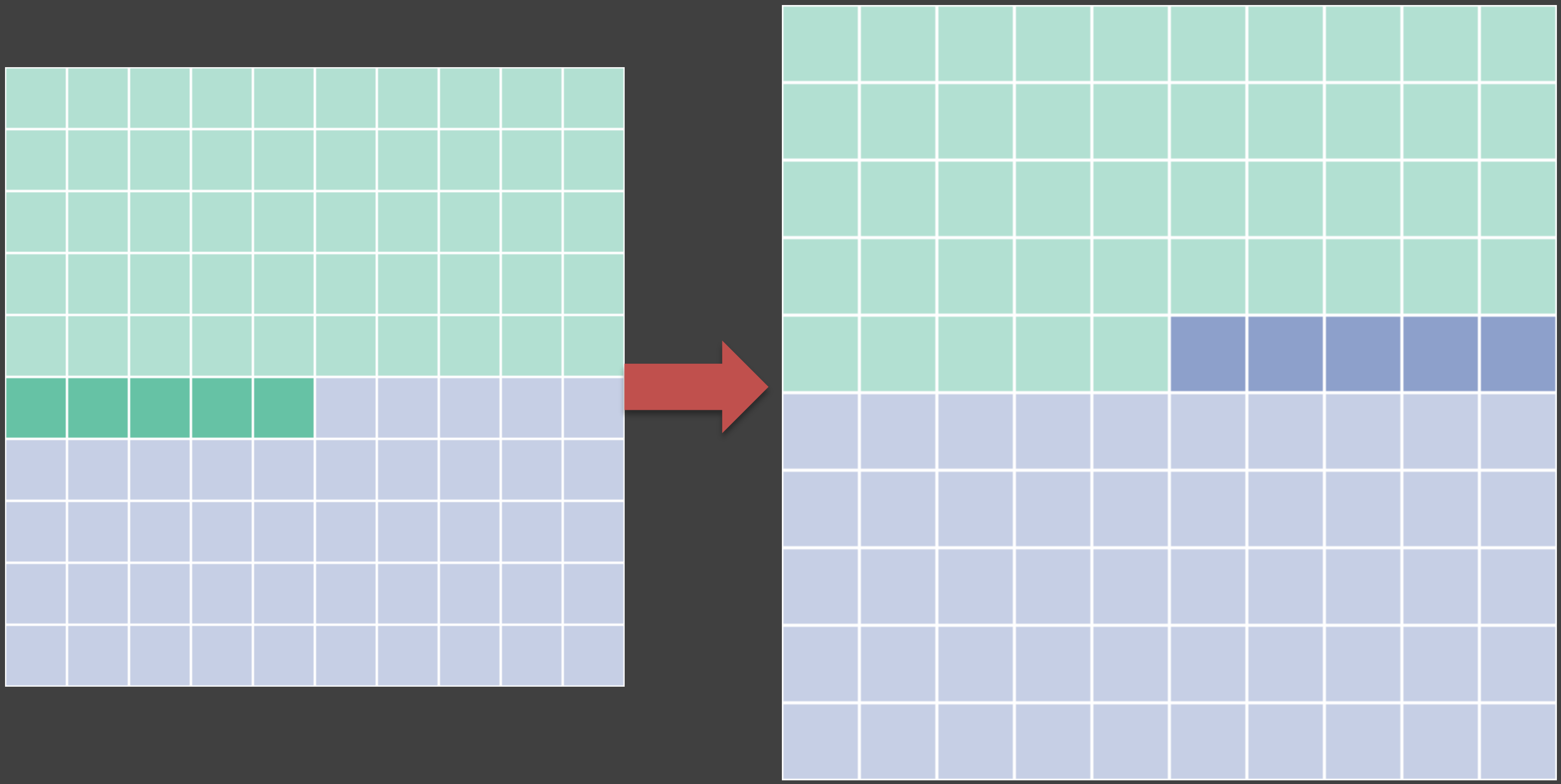
Actual Election?

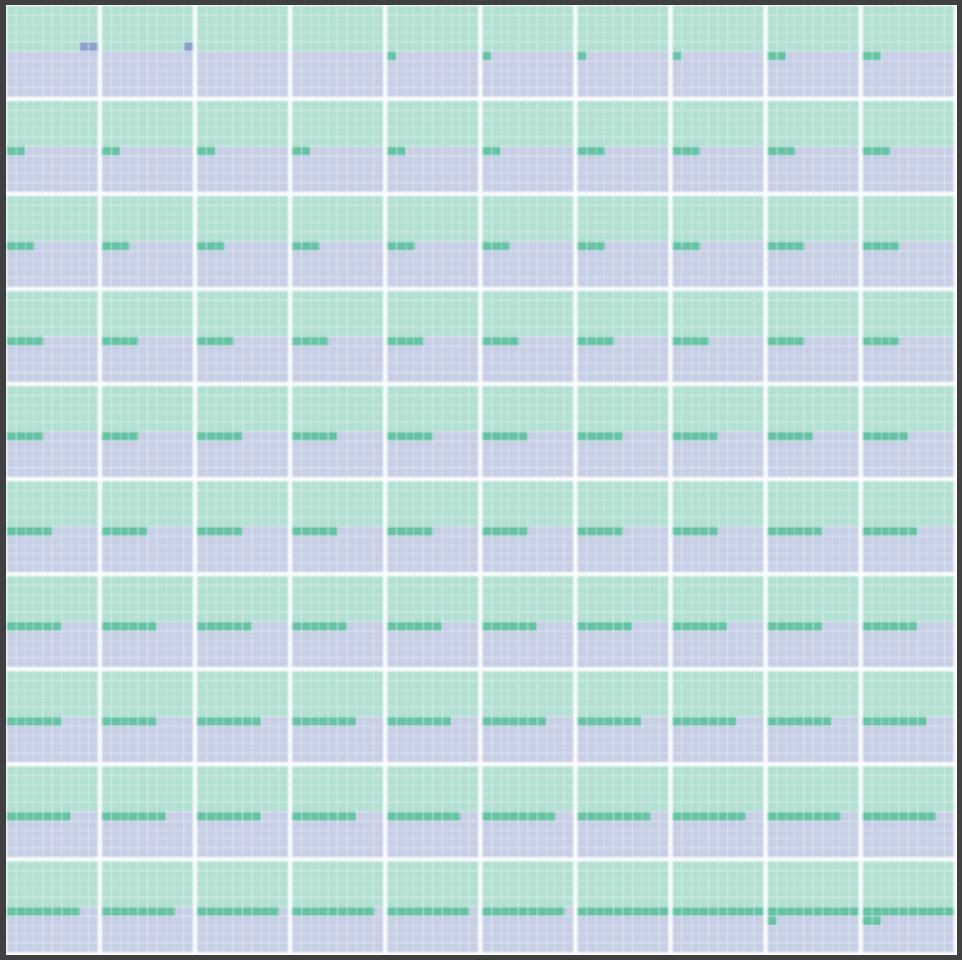


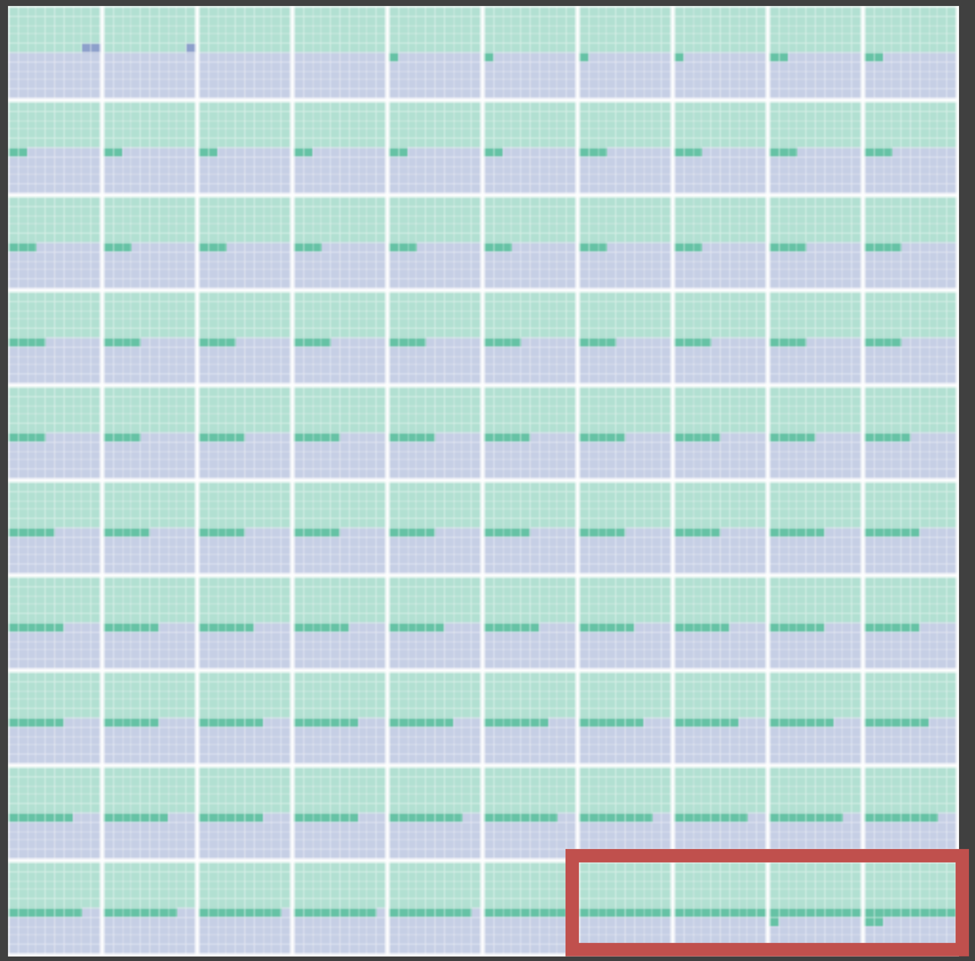
Actual Election?

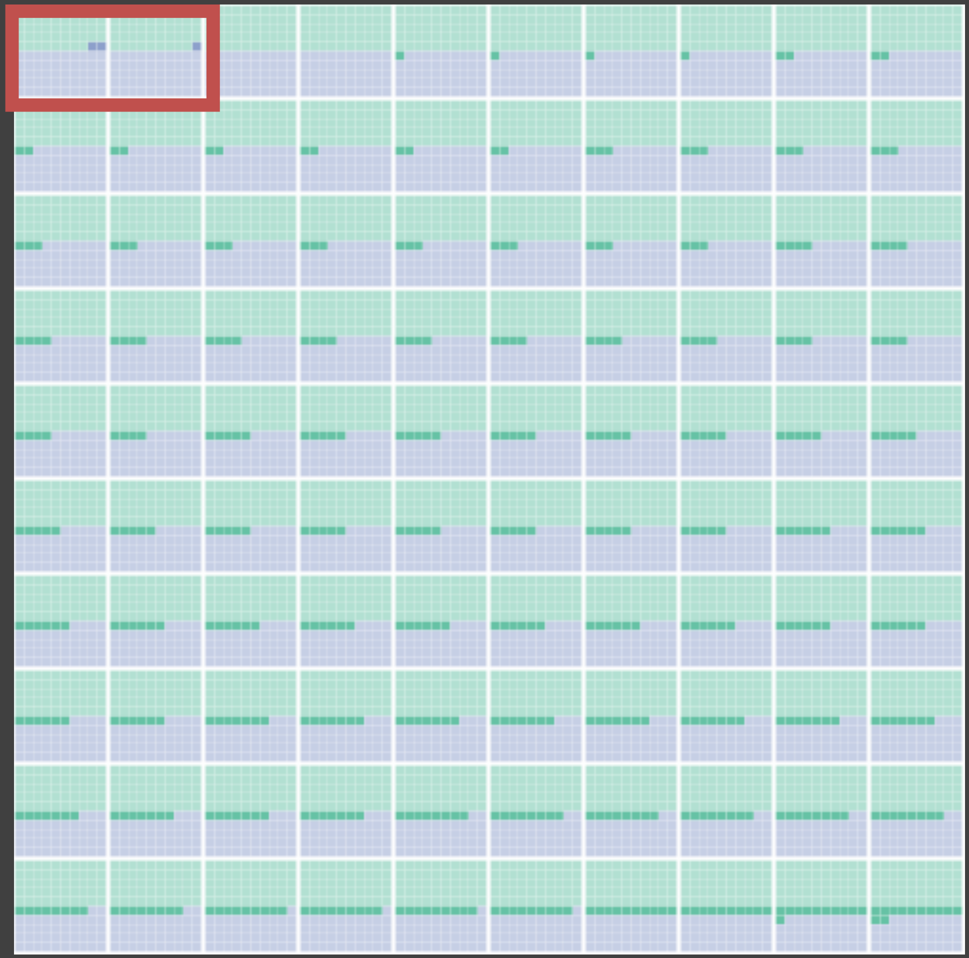


Actual Election?



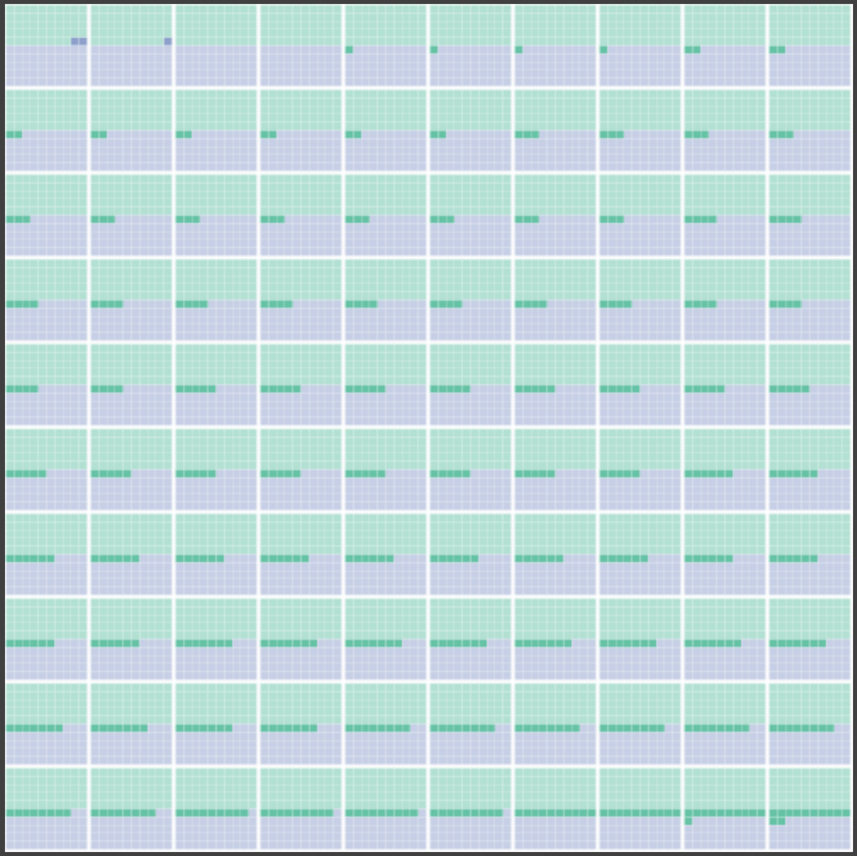






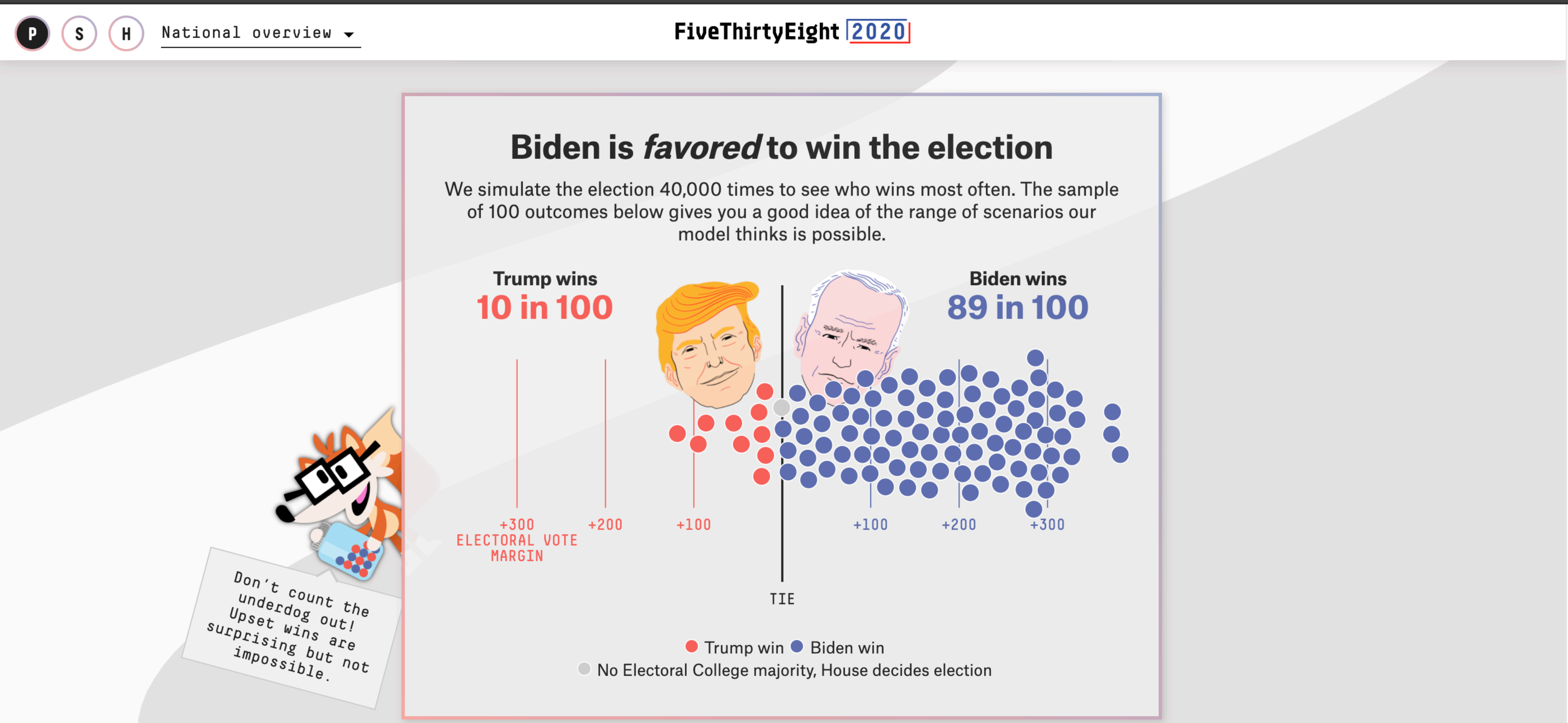
Pangloss Plot

Candidate A is ahead of Candidate B in the polls, with 55% of the likely voters*

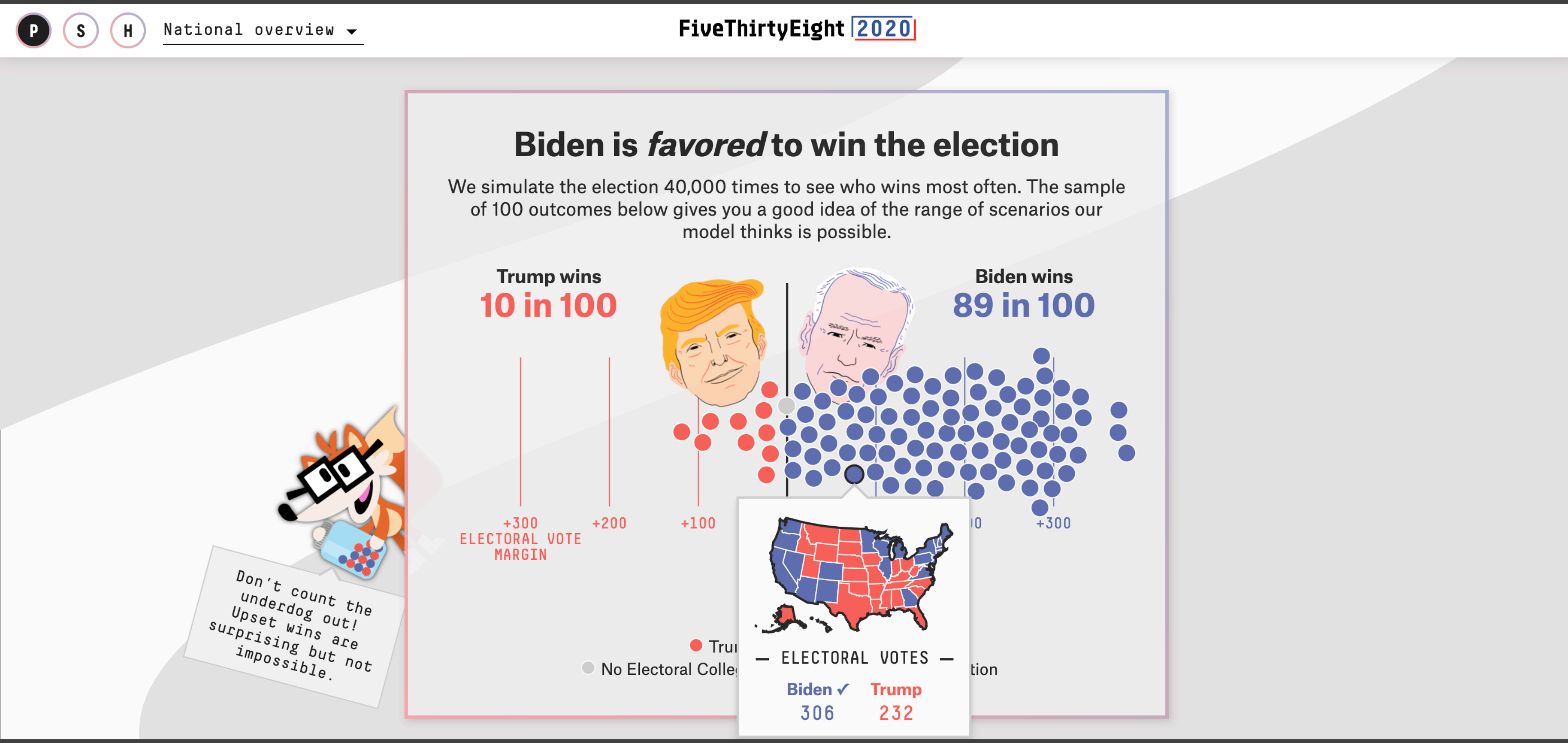


*poll of 100 people, margin of error +/-5

Bubble Swarm?



Bubble Swarm?



Model Visualization

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

How Should I Visualize Uncertainty?

Choose an appropriate visual variable based on the domain, literacy, and expertise of your audience. Be mindful that any display of uncertainty inherently increases the complexity of your visualization, and that there is a preference/performance gap.

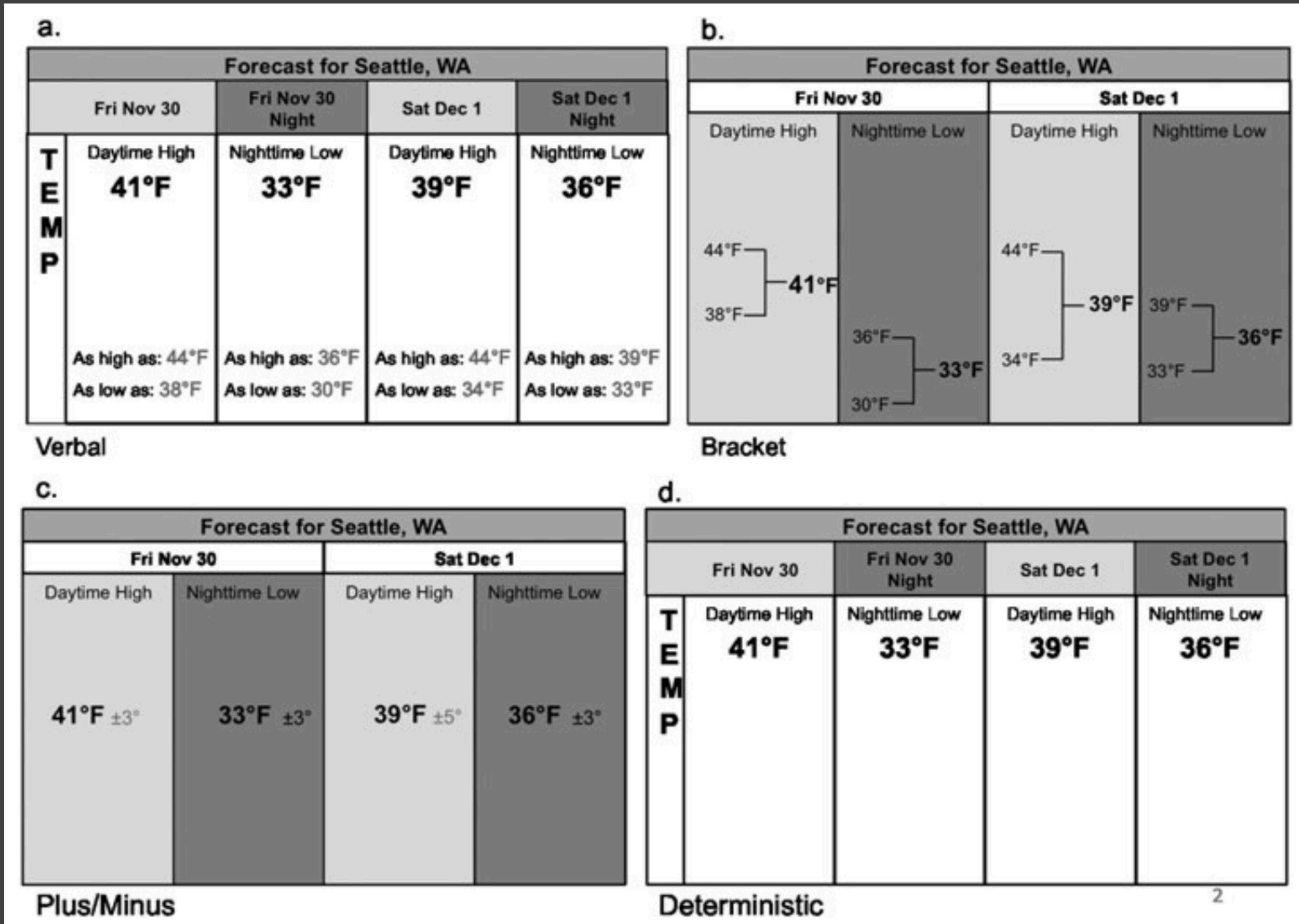
How Should I Visualize Uncertainty?

Choose an appropriate visual variable based on the domain, literacy, and expertise of your audience. Be mindful that any display of uncertainty inherently increases the complexity of your visualization, and that there is a preference/performance gap.

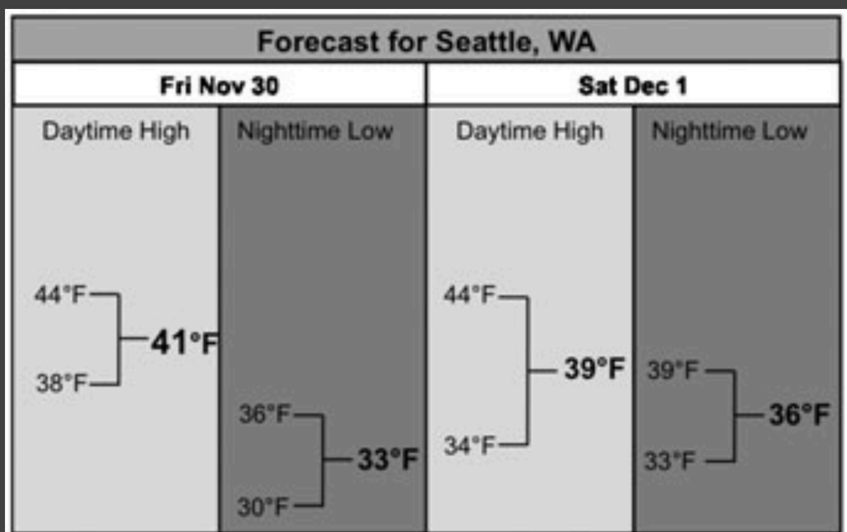
IT DEPENDS

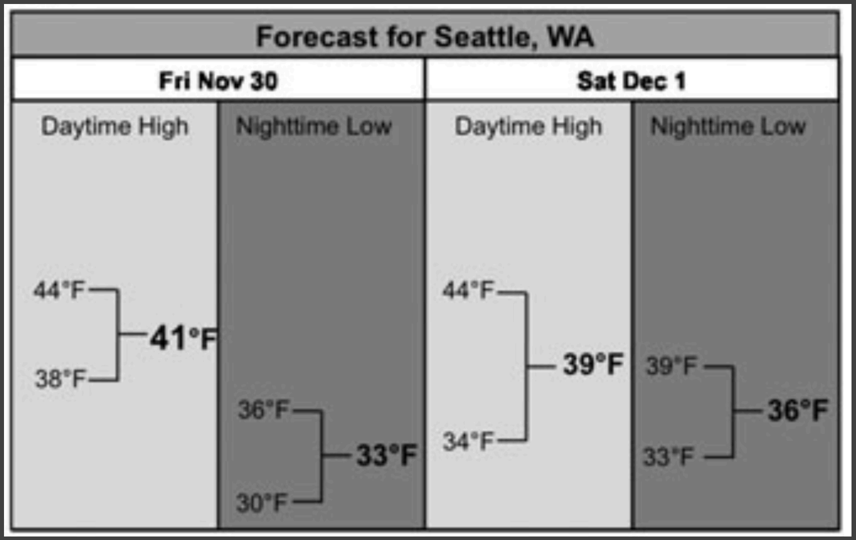
Cognitive and Perceptual Biases and Disfluencies

WHAT CAN GO WRONG WHEN VISUALIZING UNCERTAINTY?



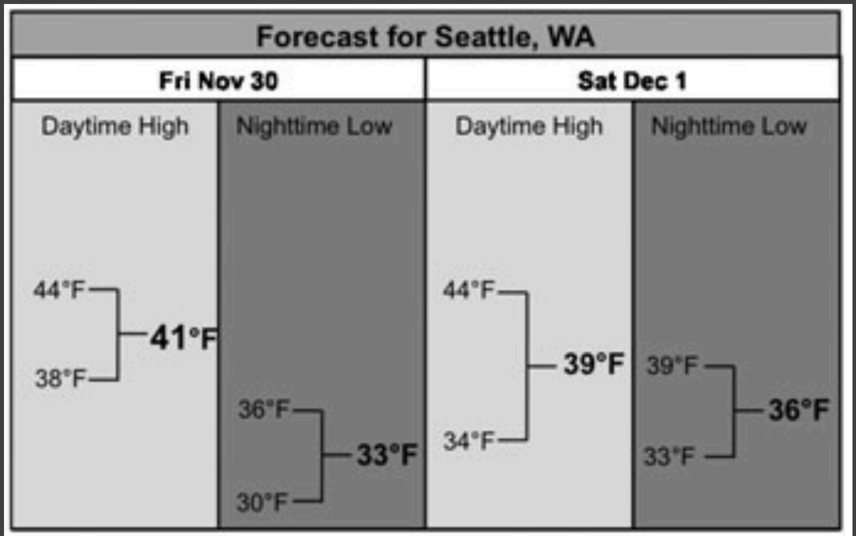
Savelli and Joslyn. The Advantages of Predictive Interval Forecasts for Non-Expert Users and the Impact of Visualizations. Applied Cog. Psych., 2013.


Forecast for Seattle, WA
Fri Nov 30
Sat Dec 1
Daytime High
Nighttime Low
Daytime High
Nighttime Low
44°F
41°F
38°F
36°F
33°F
30°F
44°F
39°F
34°F
39°F
36°F
33°F

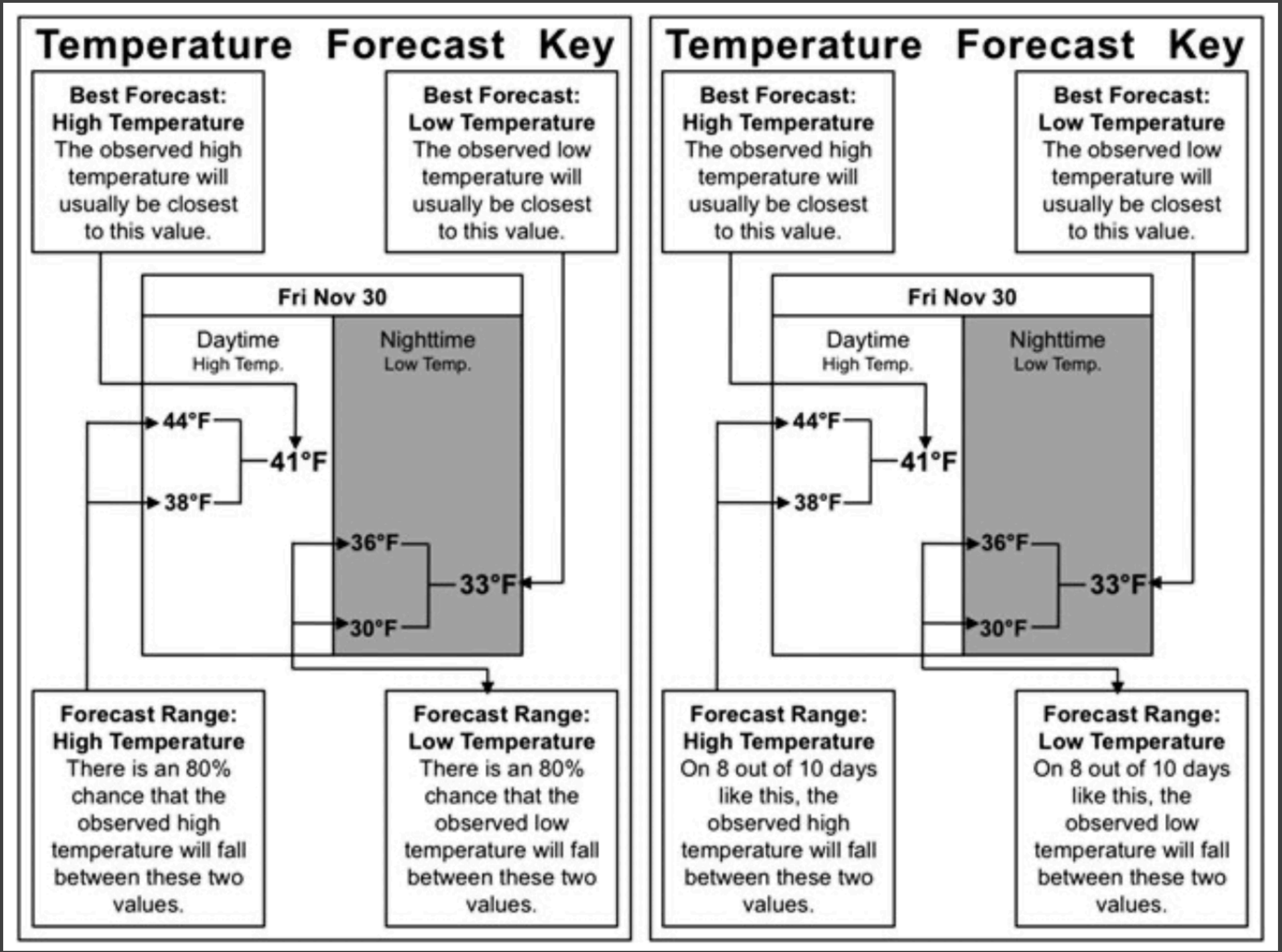


"The high tomorrow will be 44, and the low will be 38"

Deterministic Construal Error

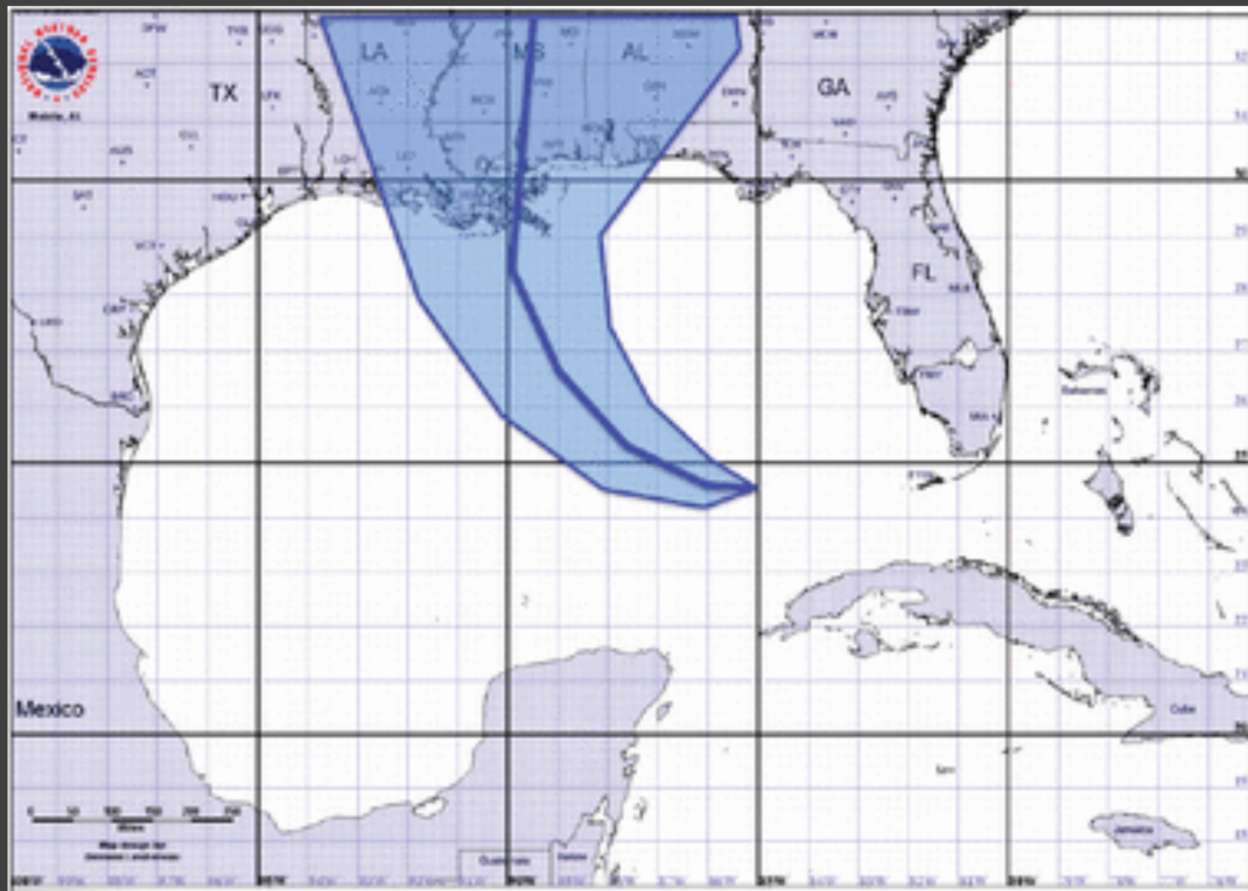


Probabilistic data is misinterpreted as being deterministic.



Savelli and Joslyn. The Advantages of Predictive Interval Forecasts for Non-Expert Users and the Impact of Visualizations. Applied Cog. Psych., 2013.

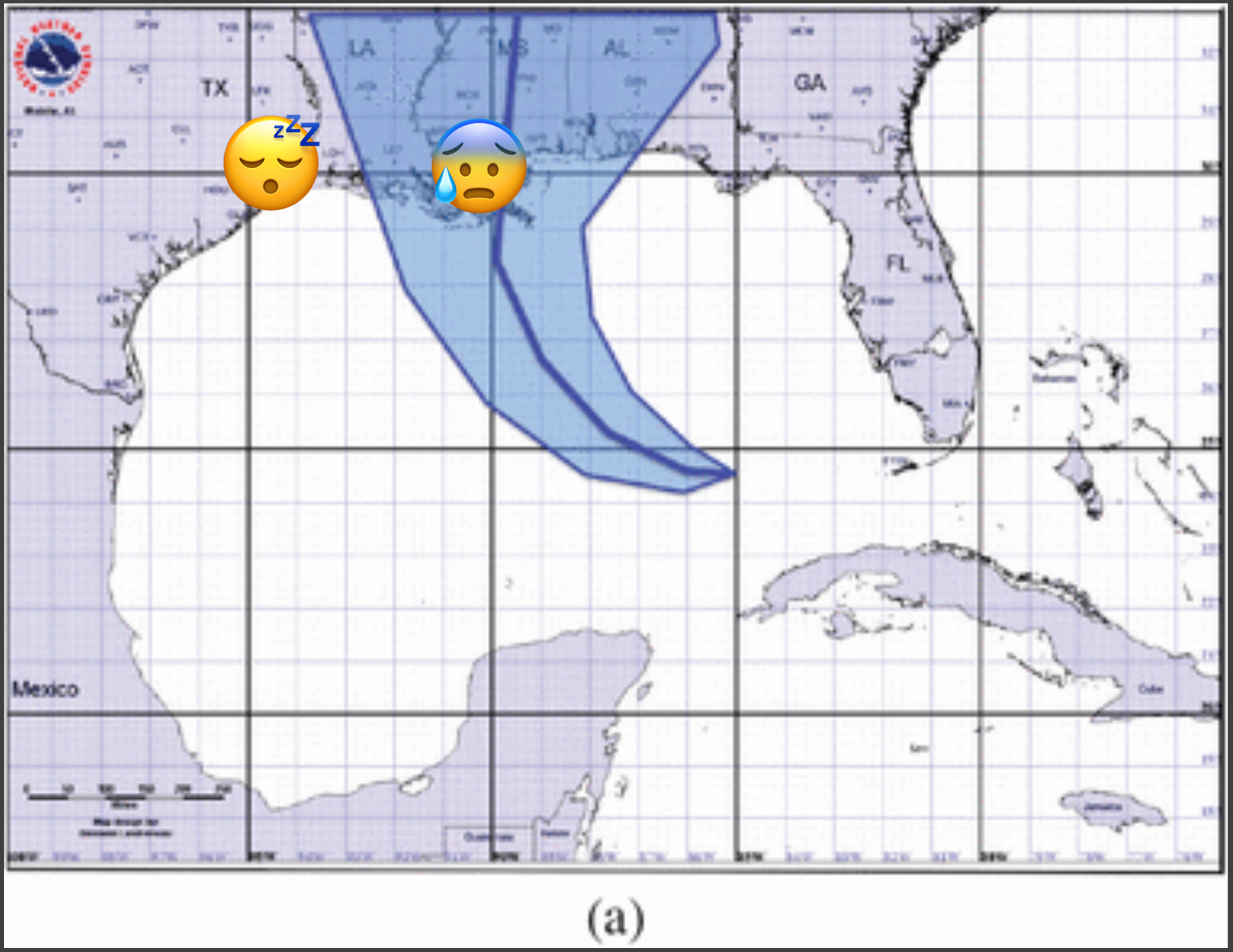
Cone of Doom



(a)

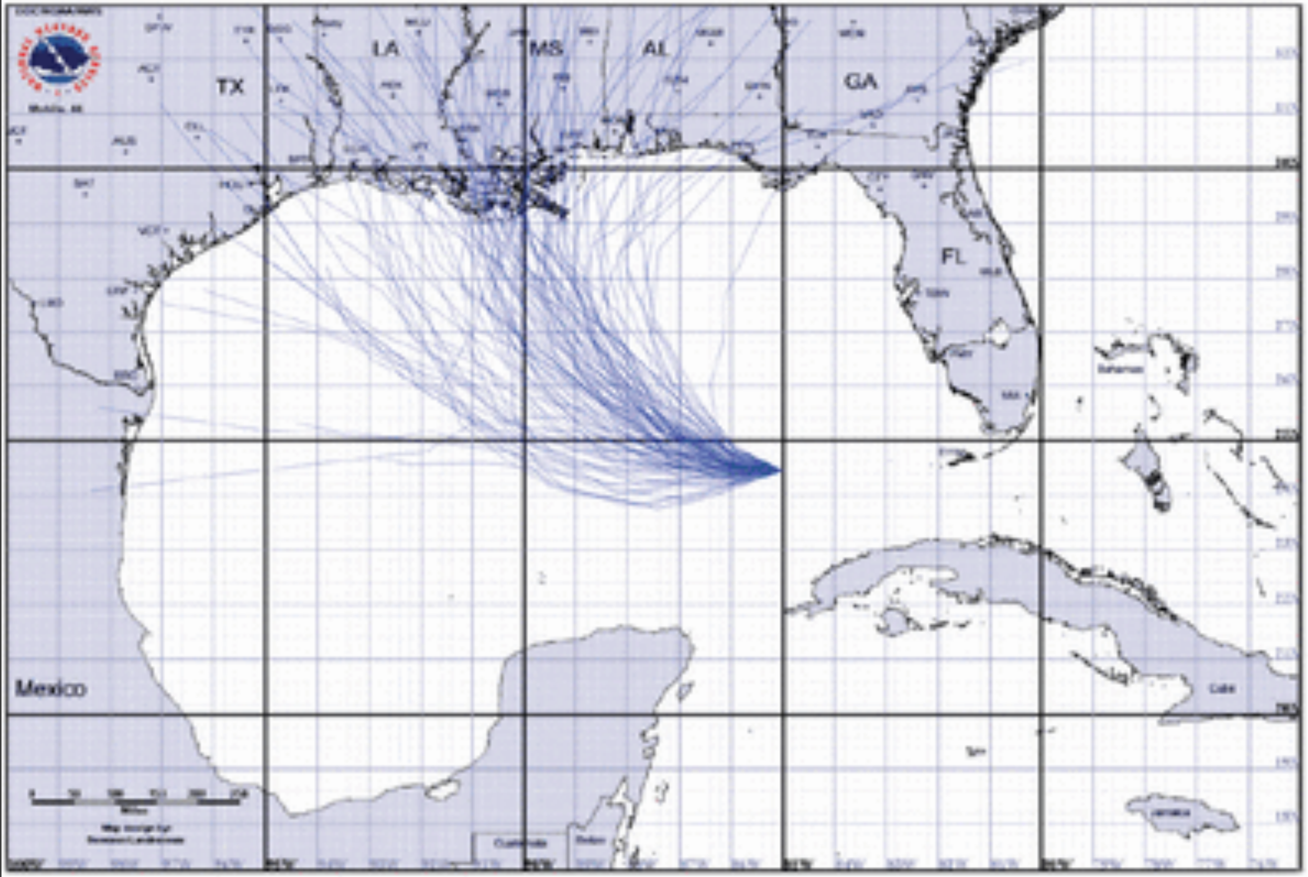
Cox, Jonathan and House, Donald and Lindell, Michael. Visuazlising uncertainty in predicted hurricane tracks. International Journal for Uncertainty Quantification, 2013.

Cone of Doom



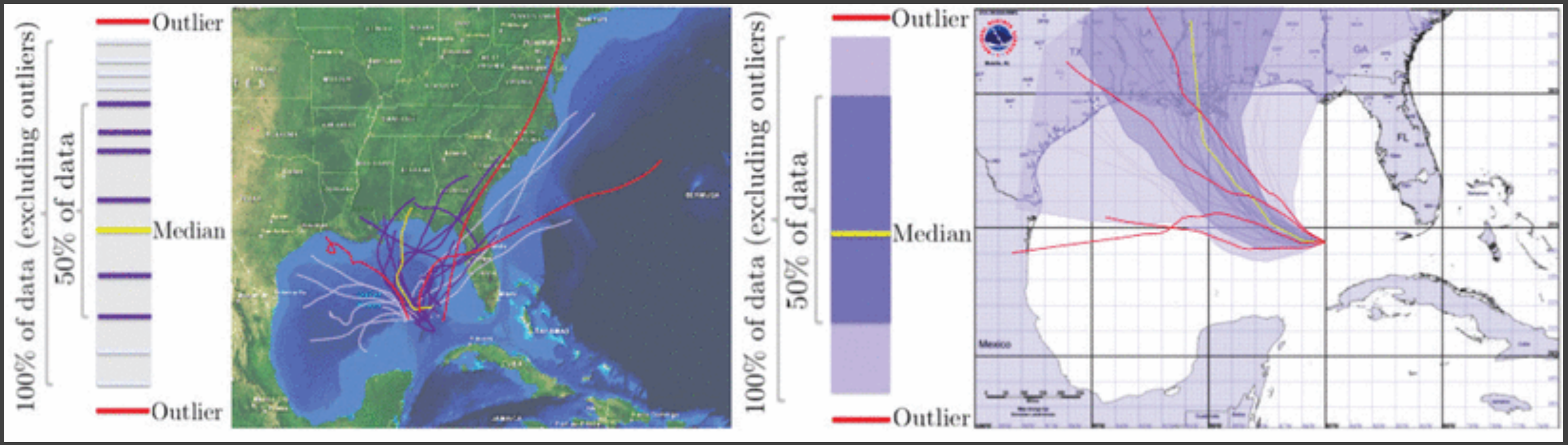
Cox, Jonathan and House, Donald and Lindell, Michael. Visuazlising uncertainty in predicted hurricane tracks. International Journal for Uncertainty Quantification, 2013.

Spaghetti/Ensemble Plots



(b)

Spaghetti/Ensemble Plots

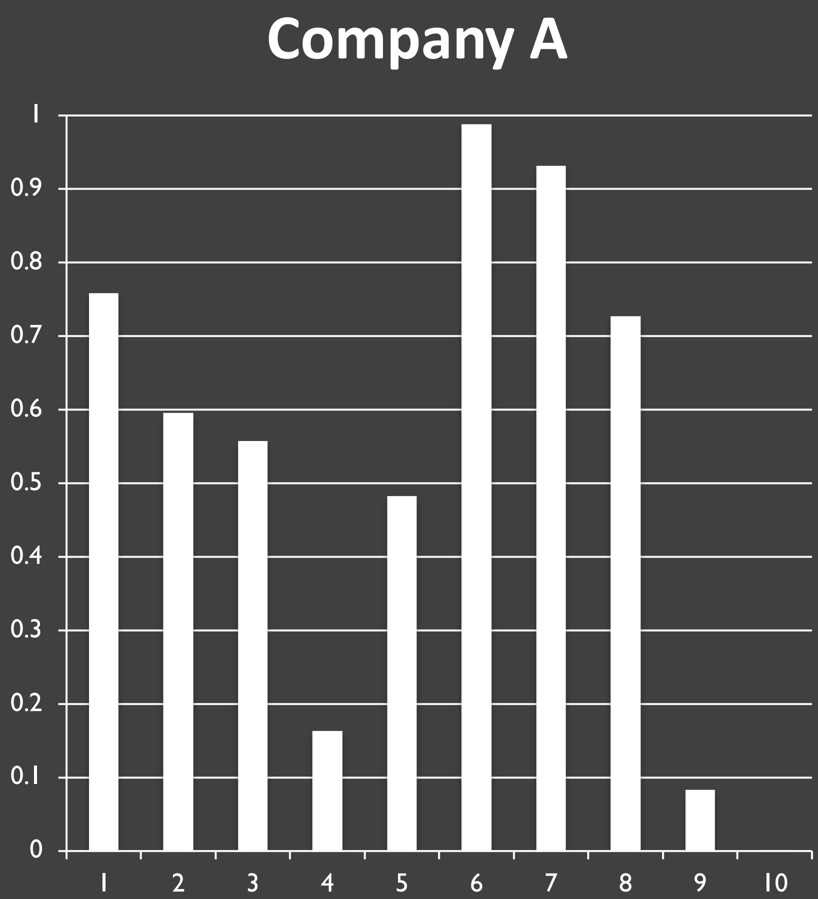


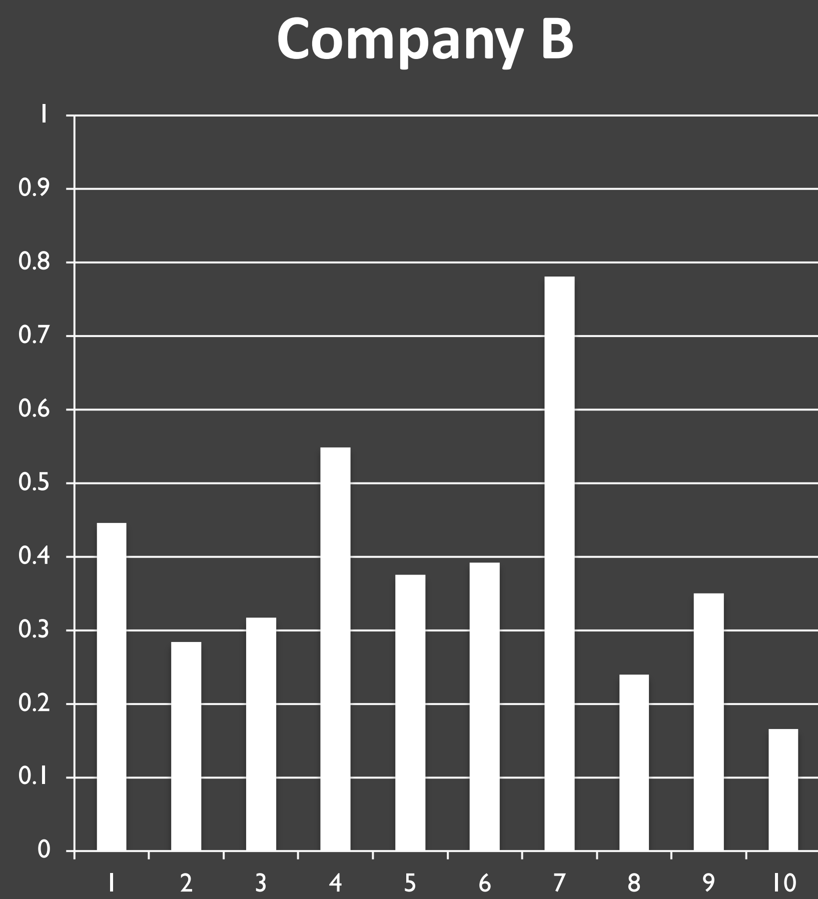
M. Mirzargar, R. Whitaker and R. Kirby. Curve Boxplot: Generalization of Boxplot for Ensembles of Curves. IEEE VIS 2014.

Things That Can Wrong

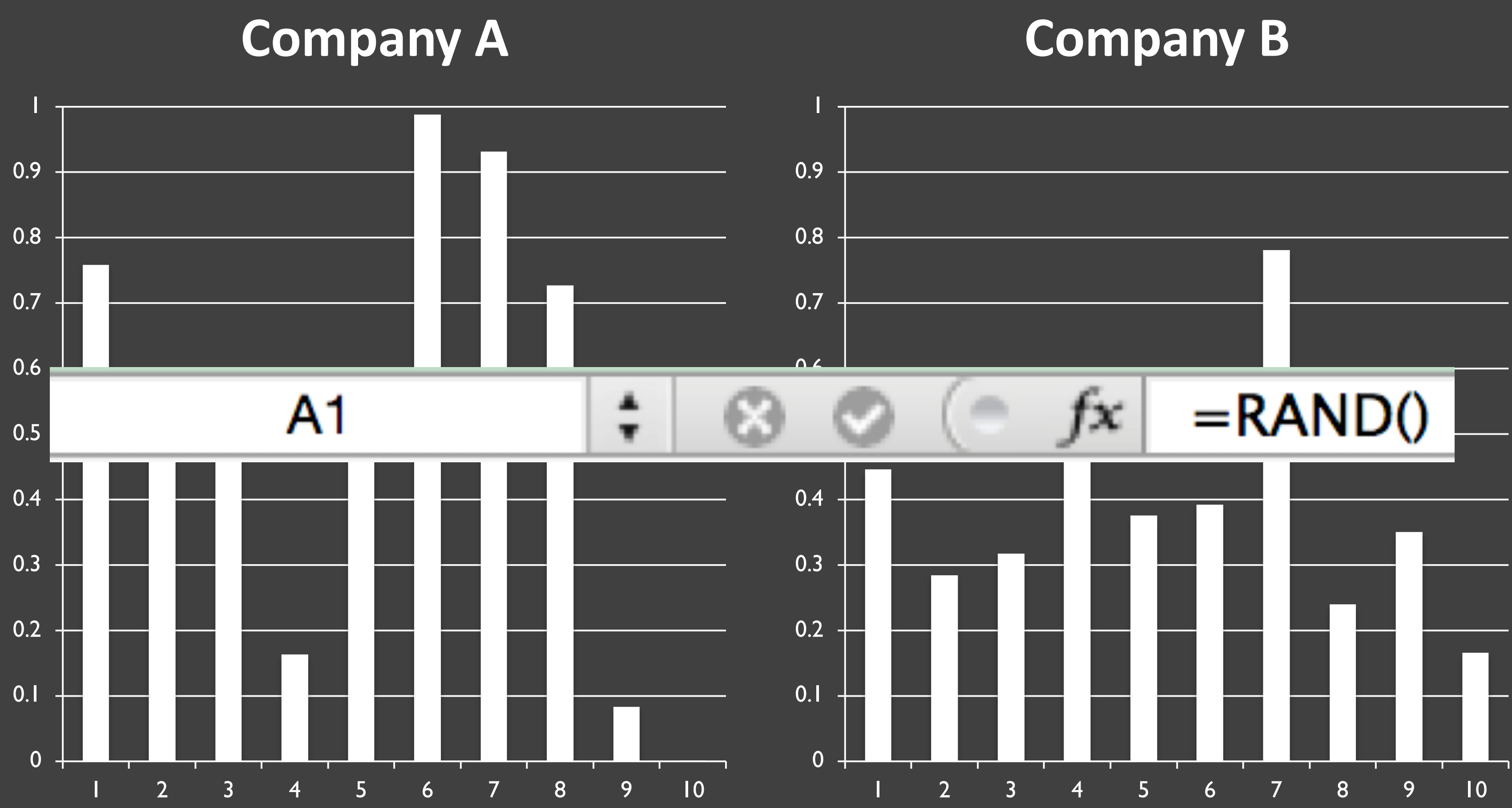
People Confuse Uncertainty with Certainty

Which Stock To Buy?

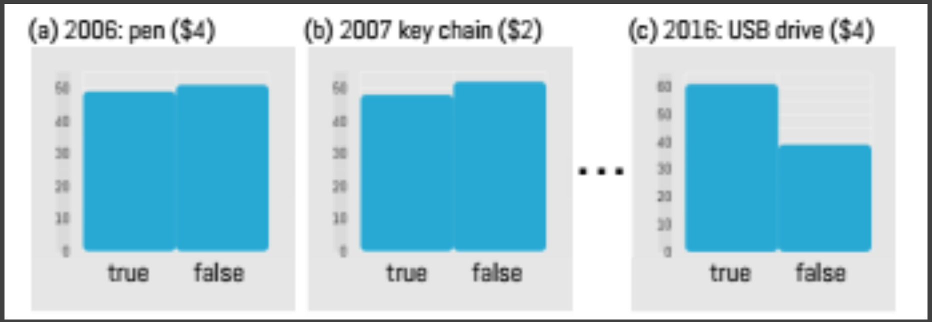




Neither!


Company A
Company B
A1
fx
=RAND()
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
10

What Swag Should We Send?



Zgraggen et al. "Investigating the Effect of the Multiple. Comparisons Problem in Visual Analysis. CHI 2018, to appear.

Fake Insights

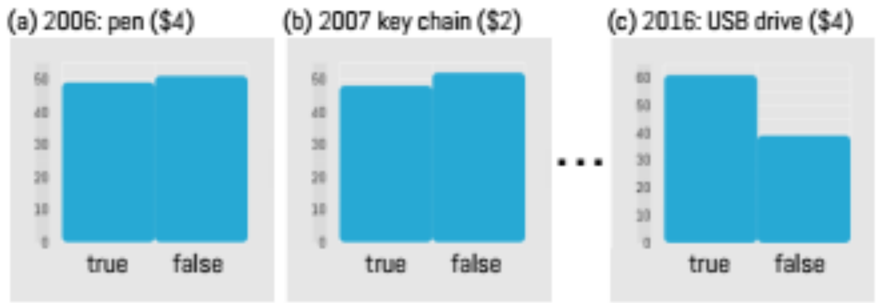


Figure 1. A user inspects several graphs and wrongly flags (c) as an insight because it looks different than (a) and (b). All were generated from the same uniform distribution and are the "same". By viewing lots of visualizations, the chances increase of seeing an apparent insight that is actually the product of random noise.

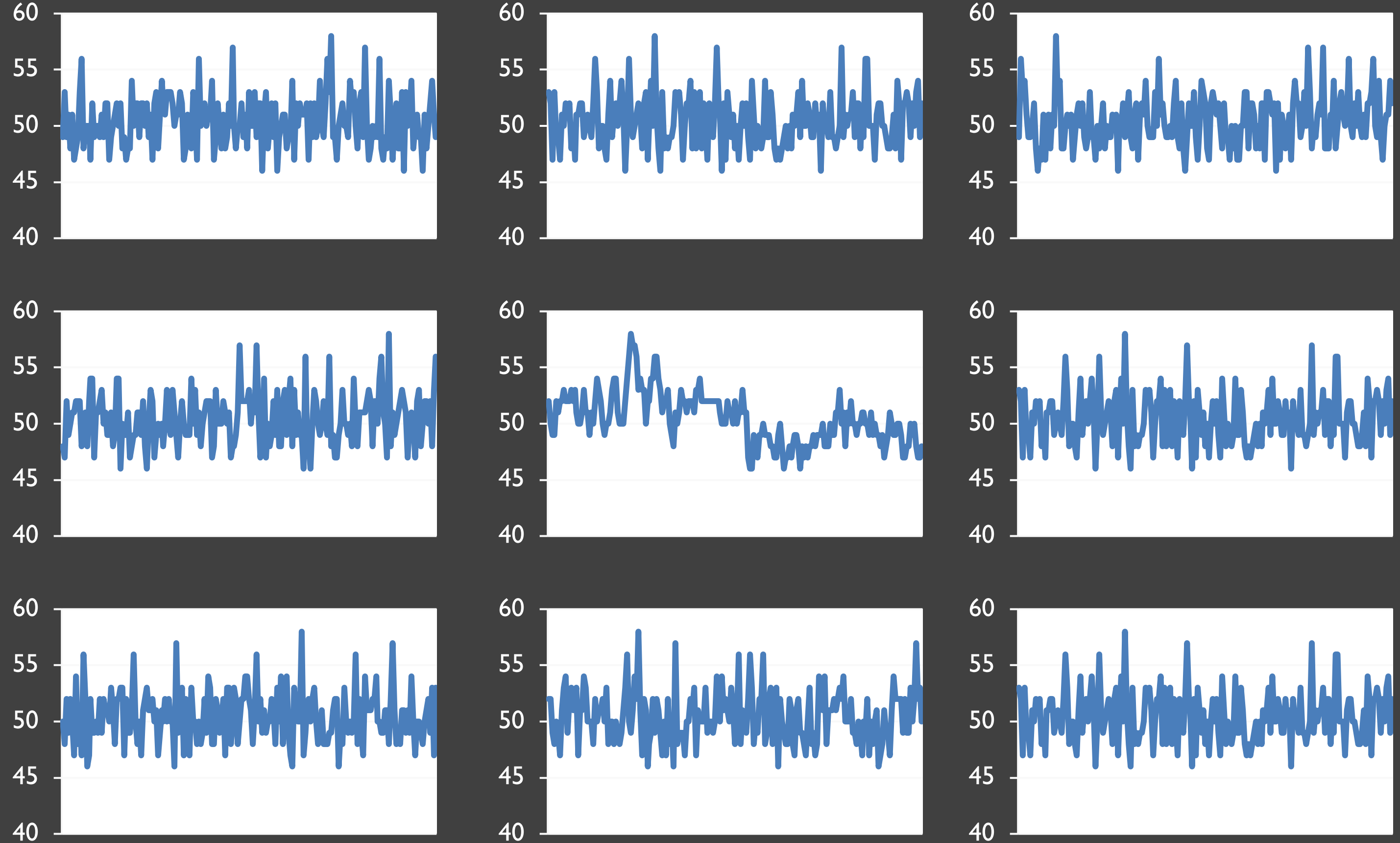
Zgraggen et al. "Investigating the Effect of the Multiple. Comparisons Problem in Visual Analysis. CHI 2018, to appear.

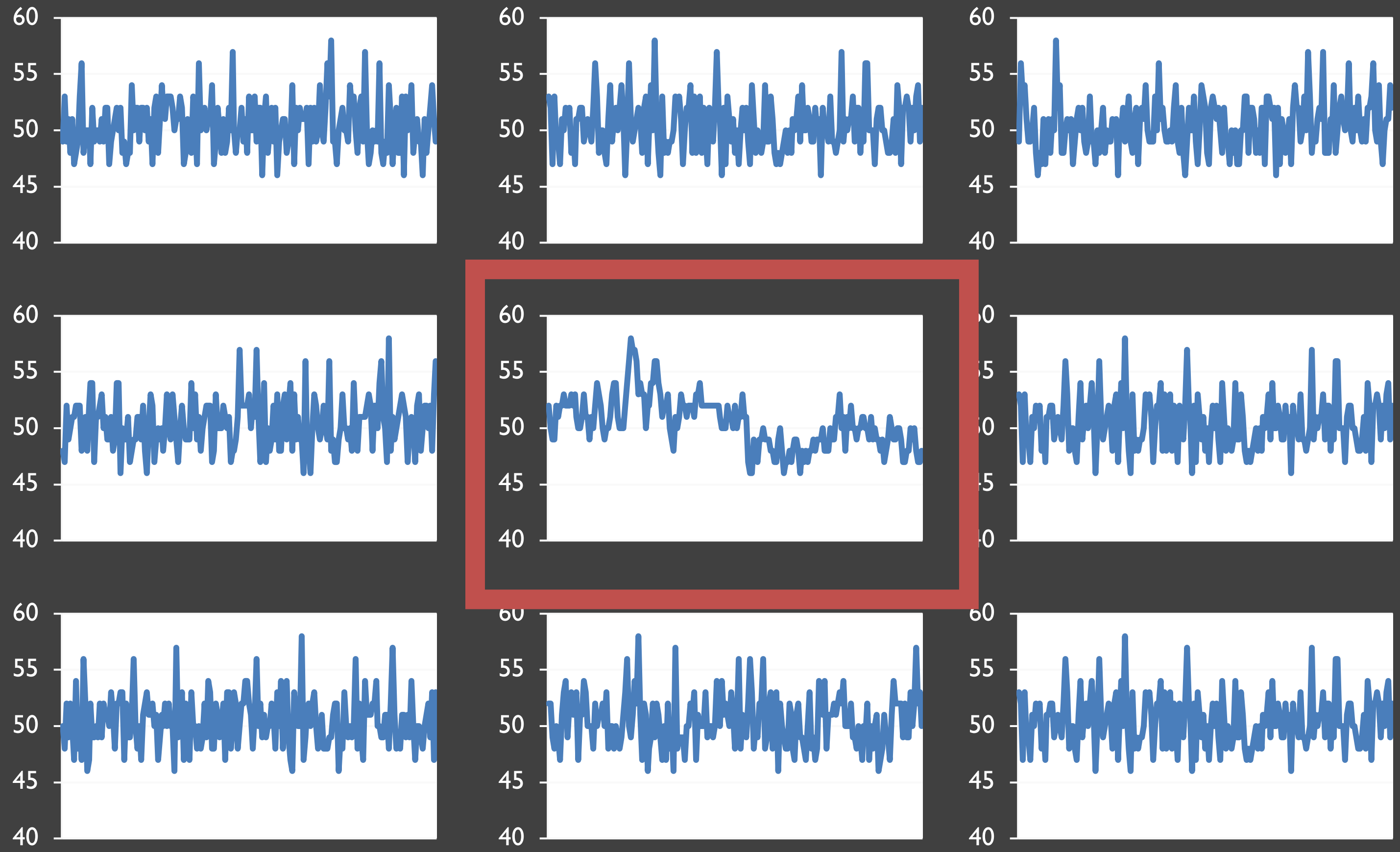
Wu Wei

無為

Pareidolia

Have People Made Up Their Mind About Obama?


60
55
50
45
40


60
55
50
45
40

Lineups Protocol

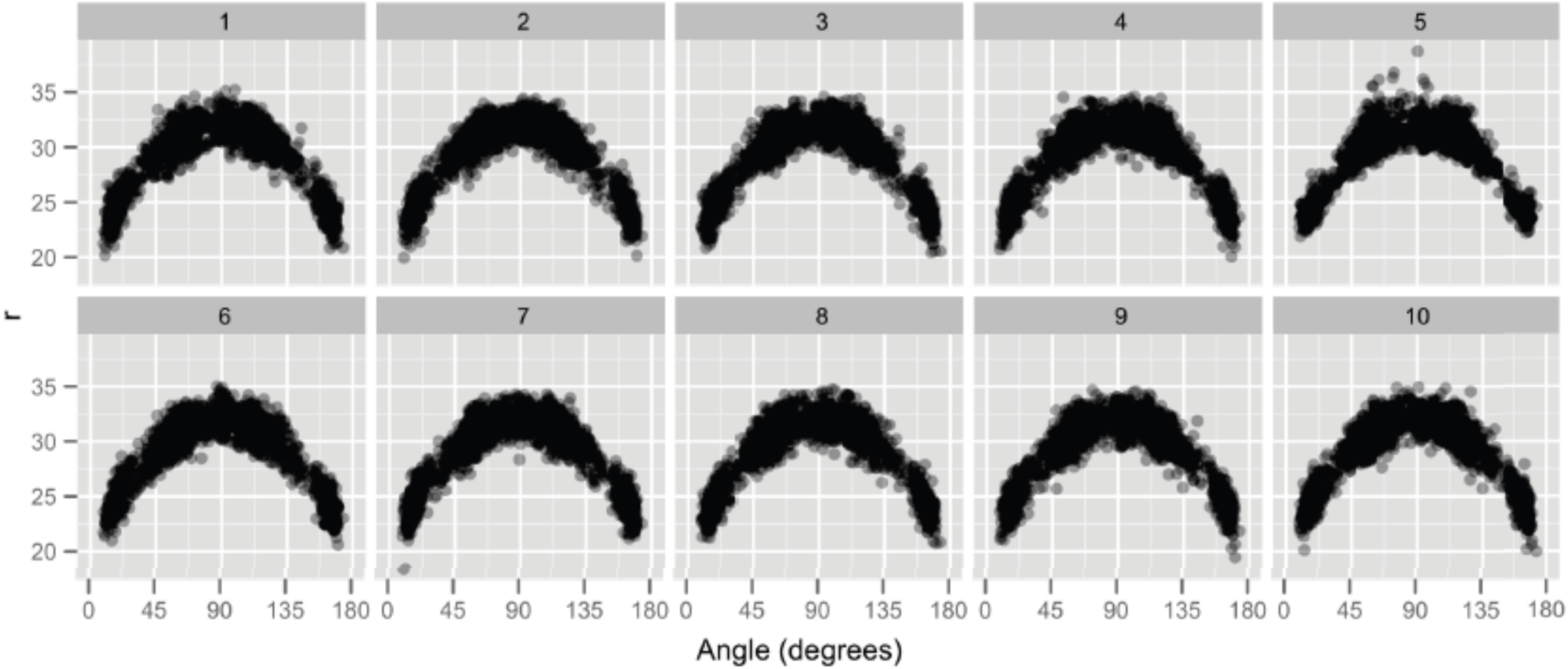
Buja et al. Statistical inference for exploratory data analysis and model diagnostics.
Royal Society, 2009.

Lineups Protocol

Buja et al. Statistical inference for exploratory data analysis and model diagnostics.
Royal Society, 2009.

Lineups Protocol!

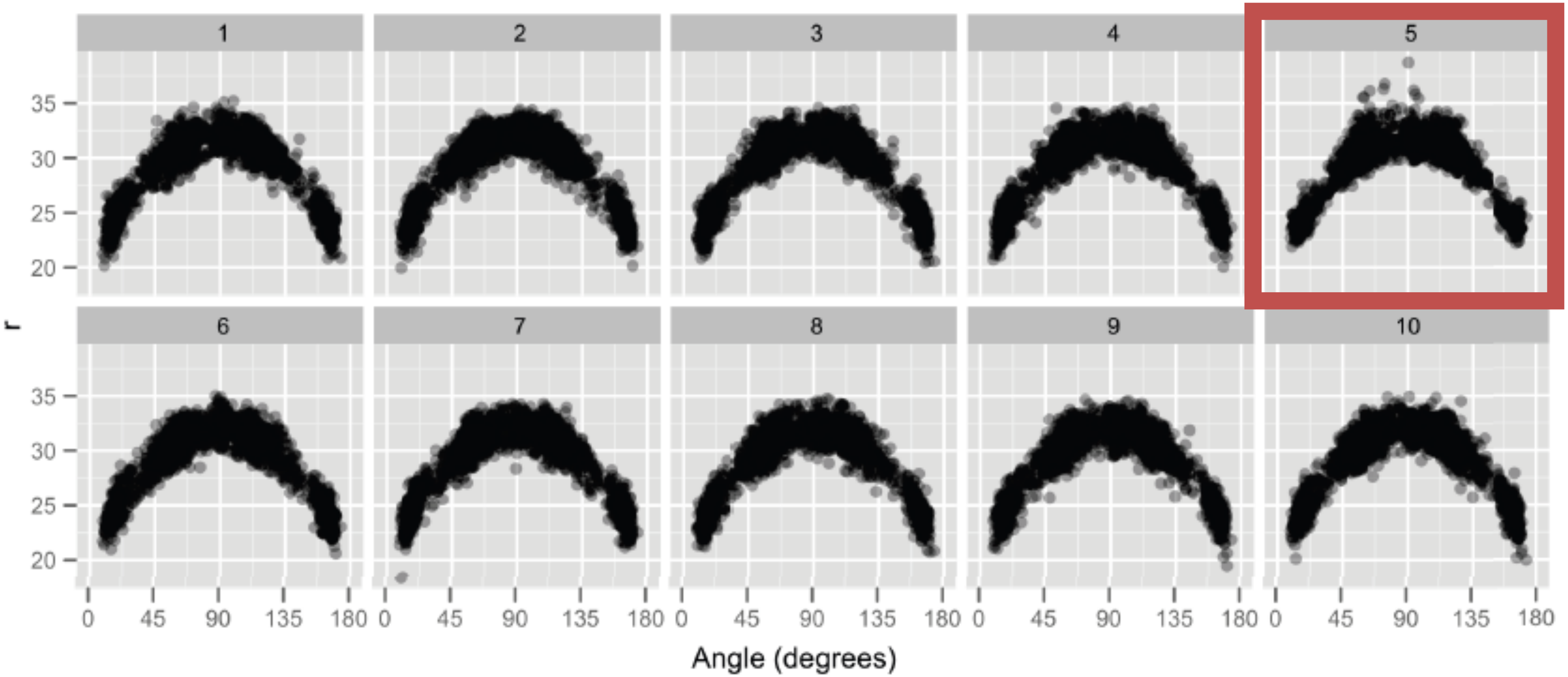
Buja et al. Statistical inference for exploratory data analysis and model diagnostics.
Royal Society, 2009.



Distance vs. angle for 3 point shots by the LA Lakers.

One plot is the real data. The others are generated according to a null hypothesis of quadratic relationship.

Hadley Wickham et al. "Graphical inference for Infovis." IEEE transactions on visualization and computer graphics 16.6 (2010): 973–9.



Distance vs. angle for 3 point shots by the LA Lakers.

One plot is the real data. The others are generated according to a null hypothesis of quadratic relationship.

Hadley Wickham et al. "Graphical inference for Infovis." IEEE transactions on visualization and computer graphics 16.6 (2010): 973–9.

Negative Results

People tend to analyze patterns and make decisions, even if there is “nothing to see.”

Negative or null results can correspond to weak and non-robust visual patterns across a model space.

Things That Can Wrong

People Confuse Uncertainty with Certainty

People Confuse Signal with Noise

Base Rate Fallacy

1% of the villagers are werewolves

80% of werewolves are allergic to silver.

10% of innocent villagers are allergic to silver.

If a villager is allergic to silver, what's the probability they are a werewolf?

Bayes' Law

$$P(A|B) = P(B|A)P(A) / P(B)$$

Bayes' Law

P(A|B) = P(B|A)P(A) / P(B)

P(🐺 | +Test) = P(+Test|🐺)P(🐺)/P(+Test)

Bayes' Law

P(A|B) = P(B|A)P(A) / P(B)

P(🐺 | +Test) = P(+Test|🐺)P(🐺)/P(+Test)

P(+) = P(+ ^ 🐺)P(🐺) + P(+^ ~🐺)P(~🐺)

Bayes' Law

$P(A|B) = P(B|A)P(A) / P(B)$

P(🐺 | +Test) = P(+Test|🐺)P(🐺)/P(+Test)

P(+) = P(+ ^ 🐺)P(🐺) + P(+^ ~🐺)P(~🐺)

$P(+) = 0.01*0.8 + 0.99*0.1$

$P(+) = 0.107$

P(🐺 | +) = 0.8 * 0.01 / 0.107 ≈ **0.075**

Problems

People are bad at this.

People who should be good at this are bad at it.

How you present the problem affects how bad people are at it.

How To Present Probabilities

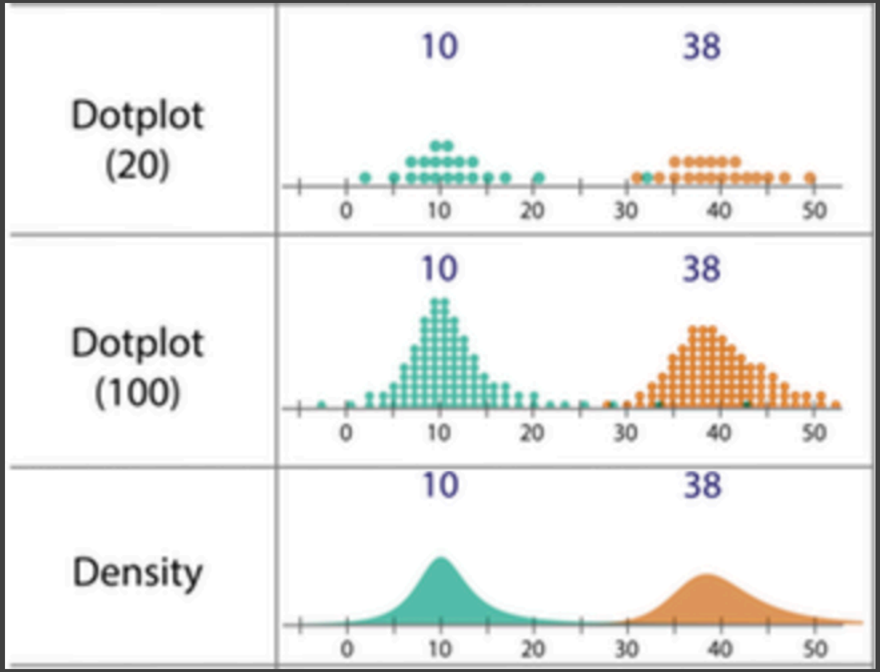
Less Intuitive	Probability	$P(A) = 0.6$
	Percentage	60% chance of A
More Intuitive	Natural Frequency	3 out of 5 times, A happens.

Ottley, et al. “Improving Bayesian reasoning: the effects of phrasing, visualization, and spatial ability.” VIS 2016.

Quantile Dot Plots

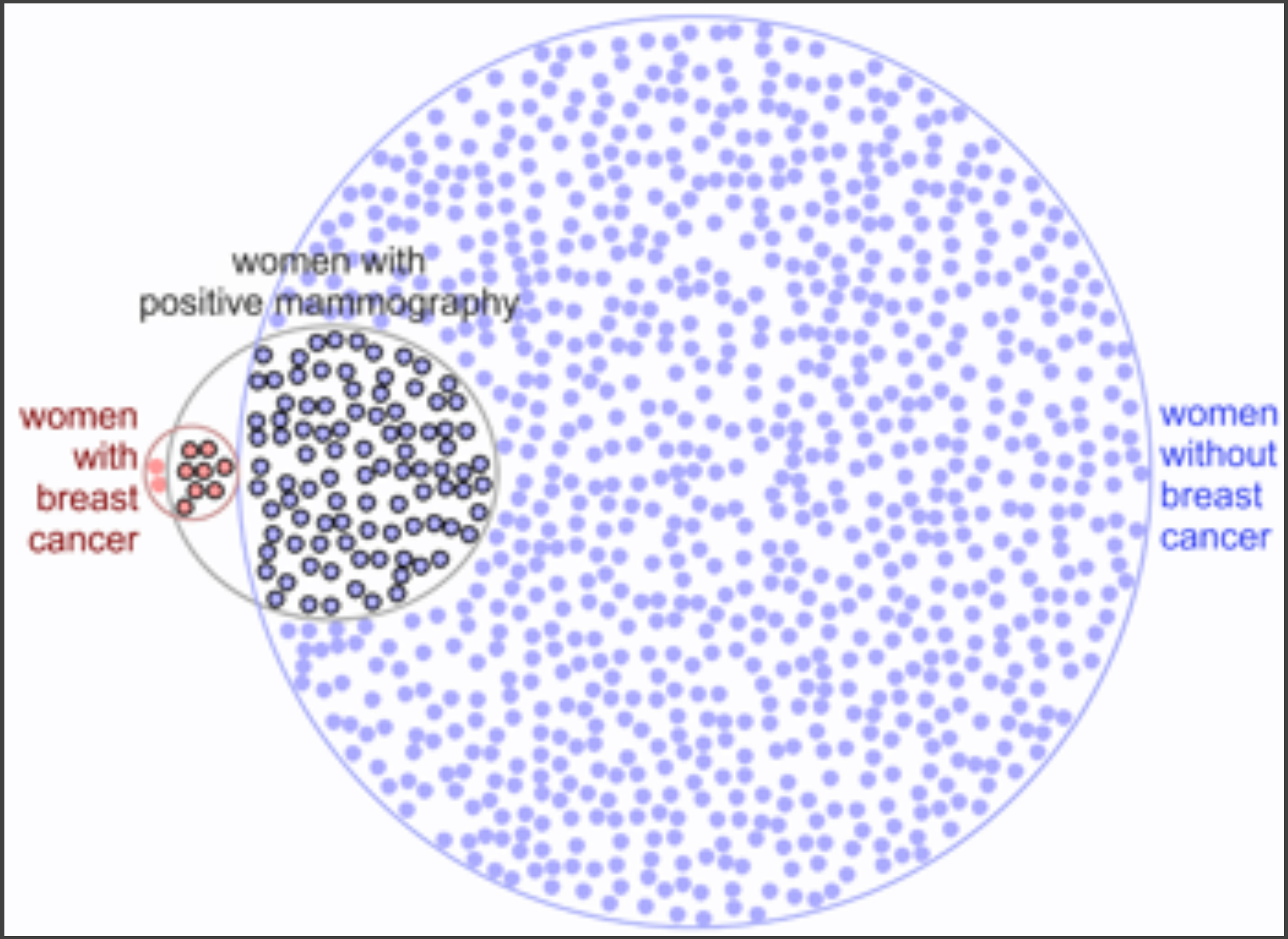
Less Error

More Error



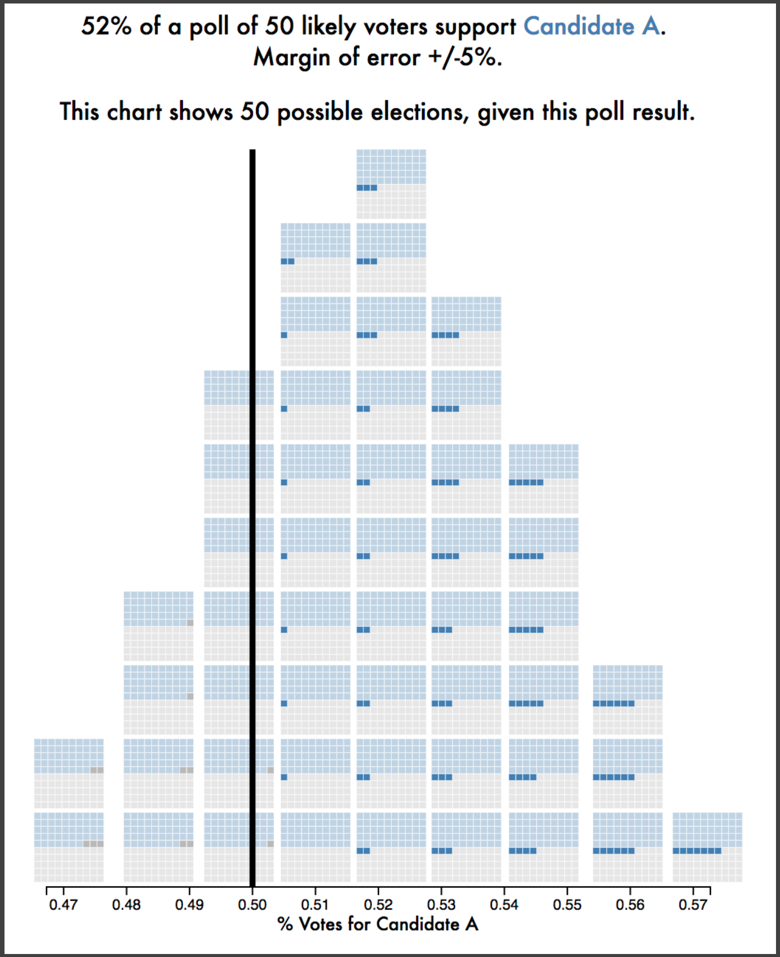
Kay et al. “When(ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems.” CHI 2016.

Base Rate Fallacy



Micallef et al. "Assessing the Effect of Visualizations on Bayesian Reasoning Through Crowdsourcing." VIS 2012.

Pangloss Dot Plot?



Things That Can Wrong

People Confuse Uncertainty with Certainty

People Confuse Signal with Noise

People Confuse Probabilities with ???

What Can Go Wrong?

Uncertainty can be difficult to understand and require a statistical background and high numeracy. Additionally, cognitive and perceptual biases can result in people making poor or error-prone decisions from uncertain data.

What Can Go Wrong?

Uncertainty can be difficult to understand and require a statistical background and high numeracy. Additionally, cognitive and perceptual biases can result in people making poor or error-prone decisions from uncertain data.

A LOT

Questions To Answer

What Does Uncertainty Mean?

How Should I Visualize It?

What Can Go Wrong?

Questions To Answer

What Does Uncertainty Mean?

LOTS OF THINGS

How Should I Visualize It?

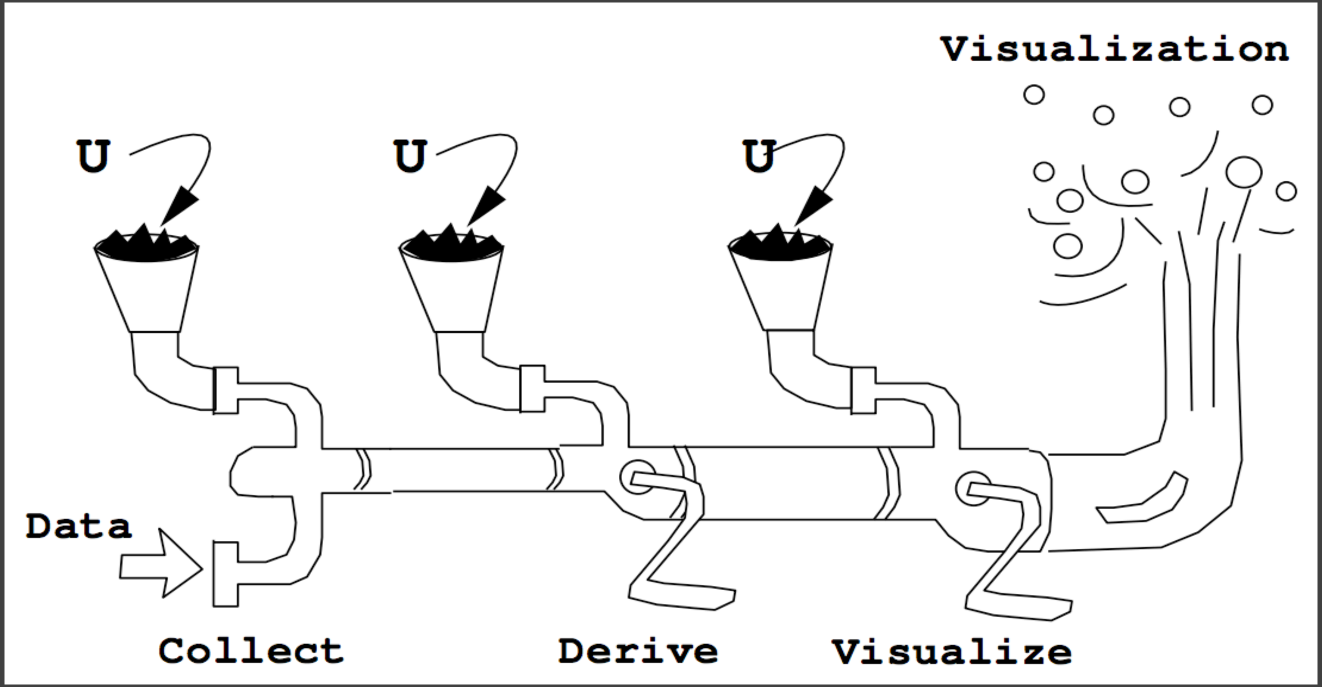
IT DEPENDS

What Can Go Wrong?

A LOT

Wrap Up

Uncertainty can happen at all stages of the analysis process, from data collection to final decision-making

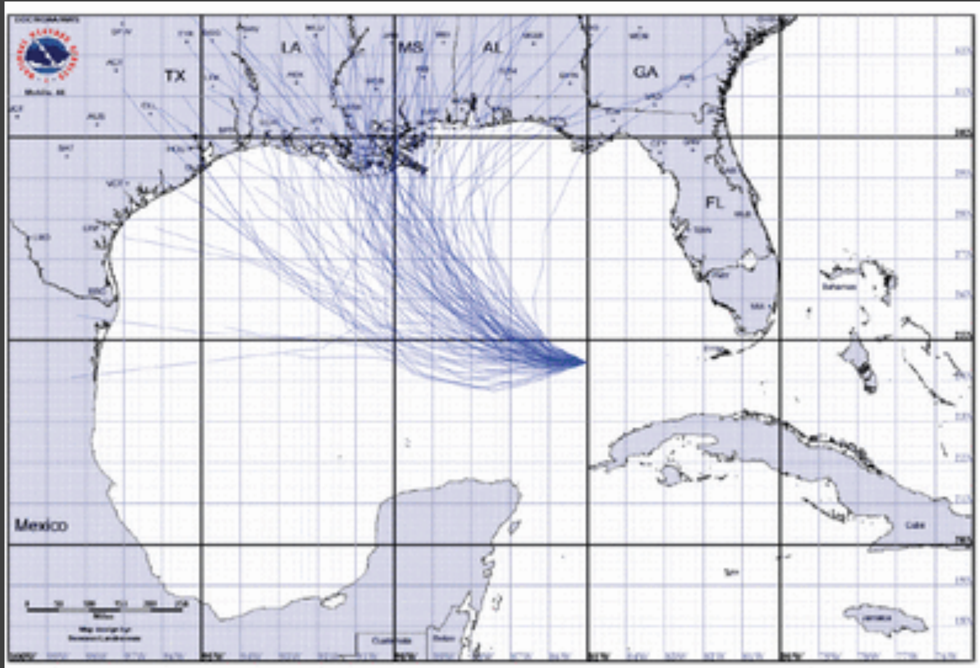


Wrap Up

Variables like blur and transparency can be intuitive for showing uncertainty, but hard to decode.

Wrap Up

Consider using discrete samples to show variation and uncertainty in a model



(b)

Wrap Up

Consider when uncertainty is high enough that doing *nothing* is the right thing to do.

Topics I Didn't Cover

Uncertainty Quantification

Uncertainty Visualization Evaluation

Visualization Verification

… lots more

Questions?

Michael Correll Tableau Research