

Uncertainty visualization

CS424: Visualization & Visual Analytics

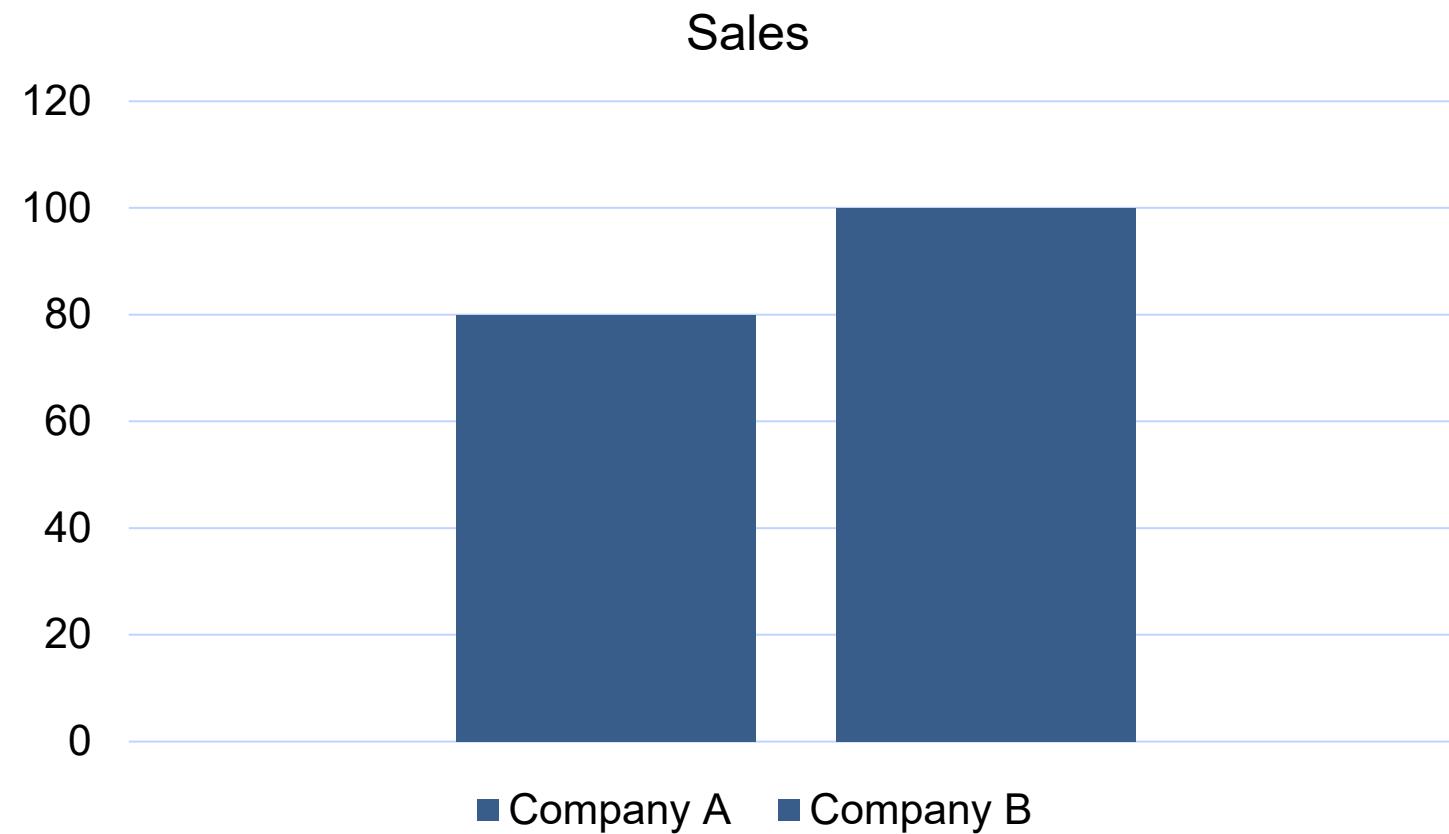
Fabio Miranda

<https://fmiranda.me>

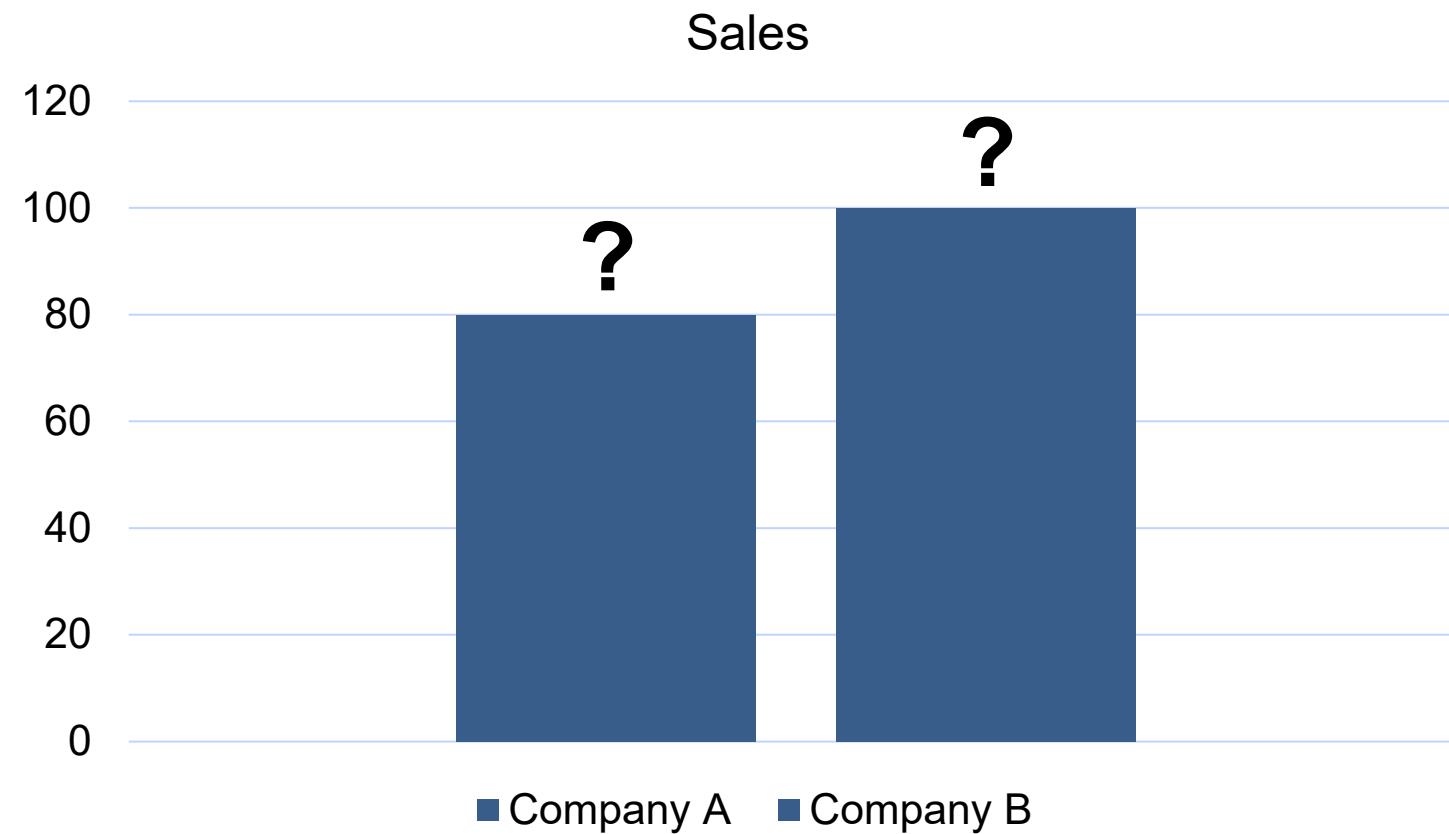
Uncertainty

- What is uncertainty?
 - Doubt
 - Risk
 - Variability
 - Error
 - Lack of knowledge
- Uncertainty is inherent to most data, and can enter the analysis pipeline during measurement, modeling and forecasting phases.
- It is important to effectively communicate uncertainty to establish transparency.

Uncertainty: a simple example



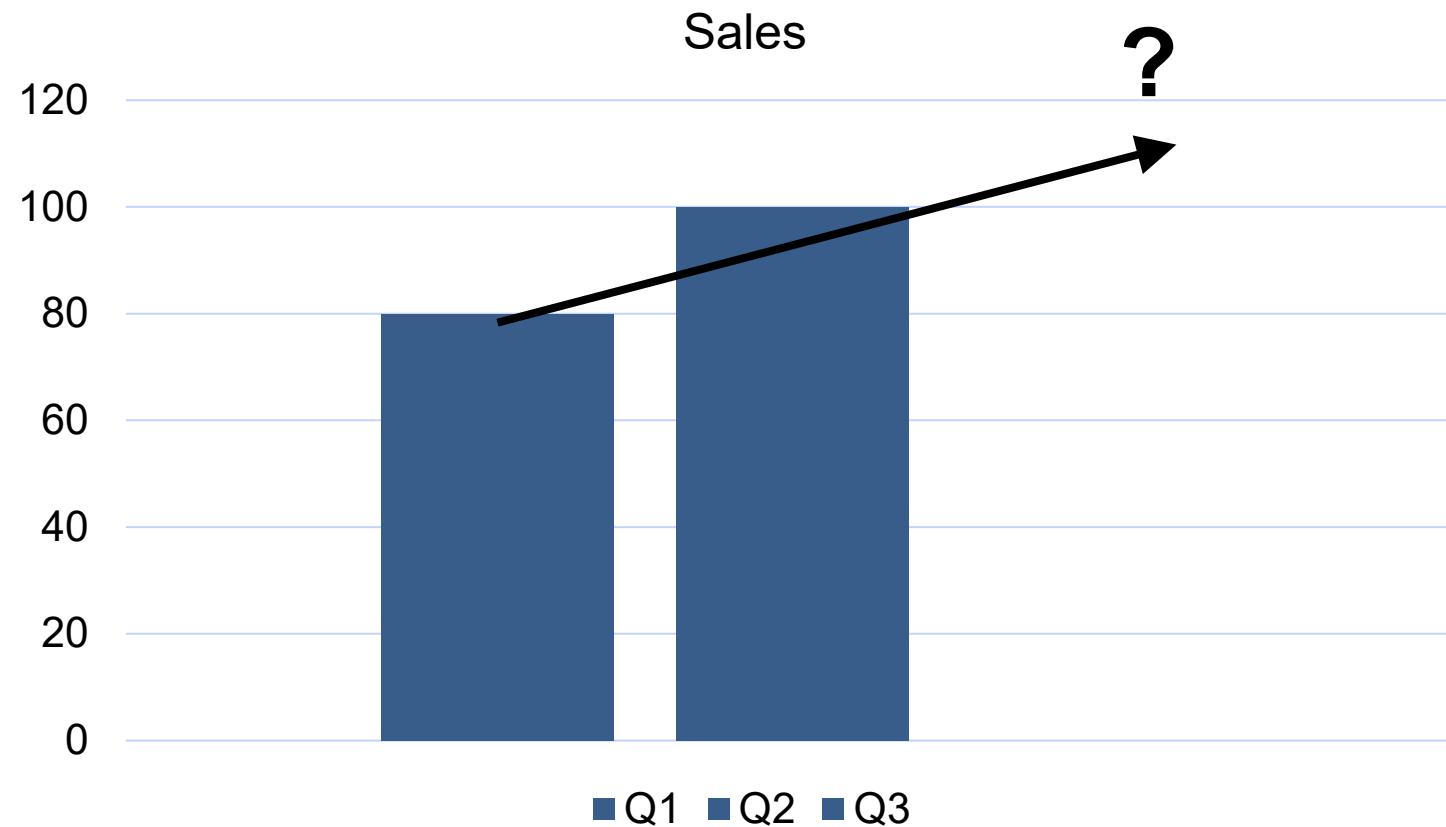
Measurement uncertainty



Forecast uncertainty



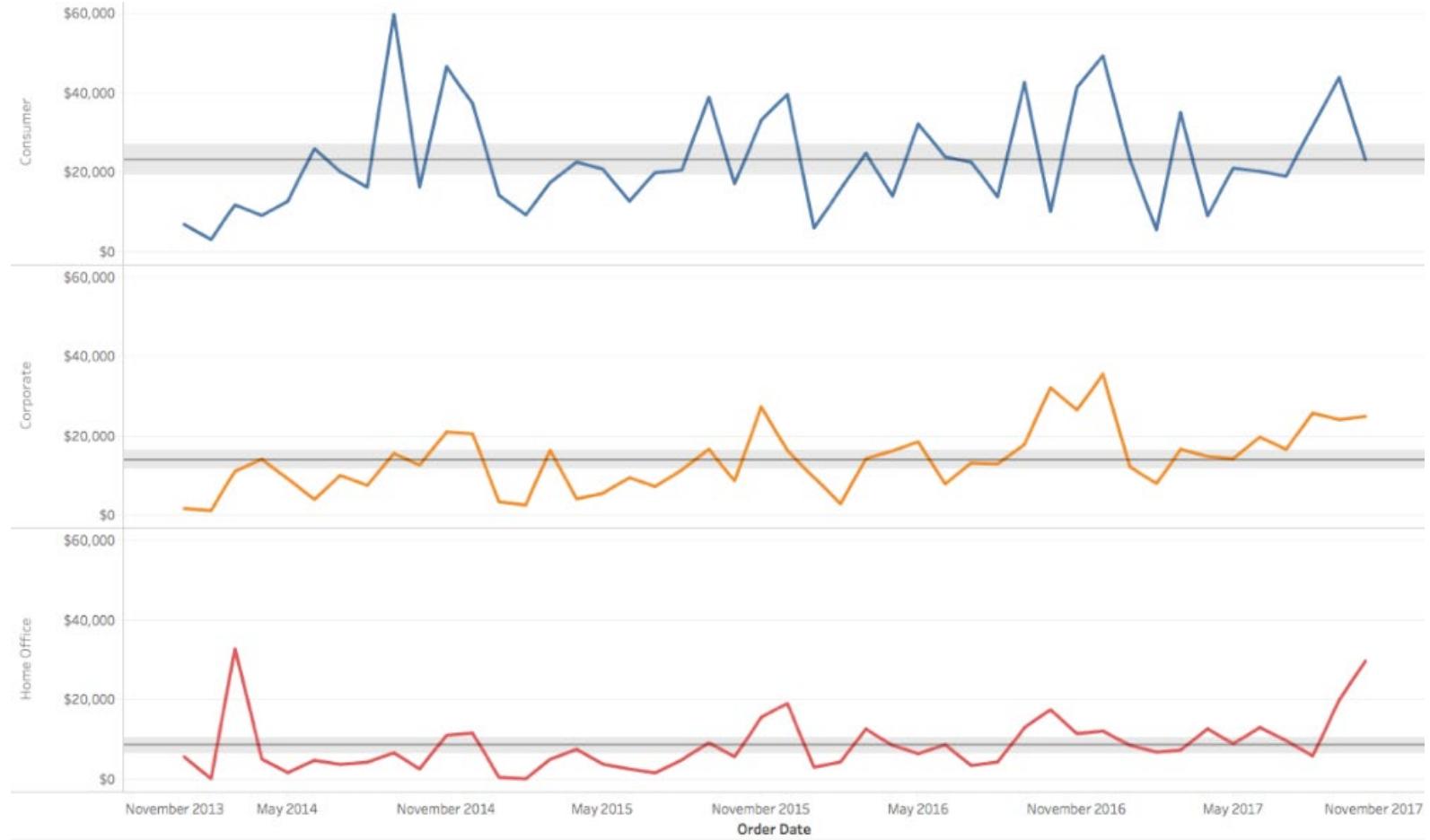
Model uncertainty



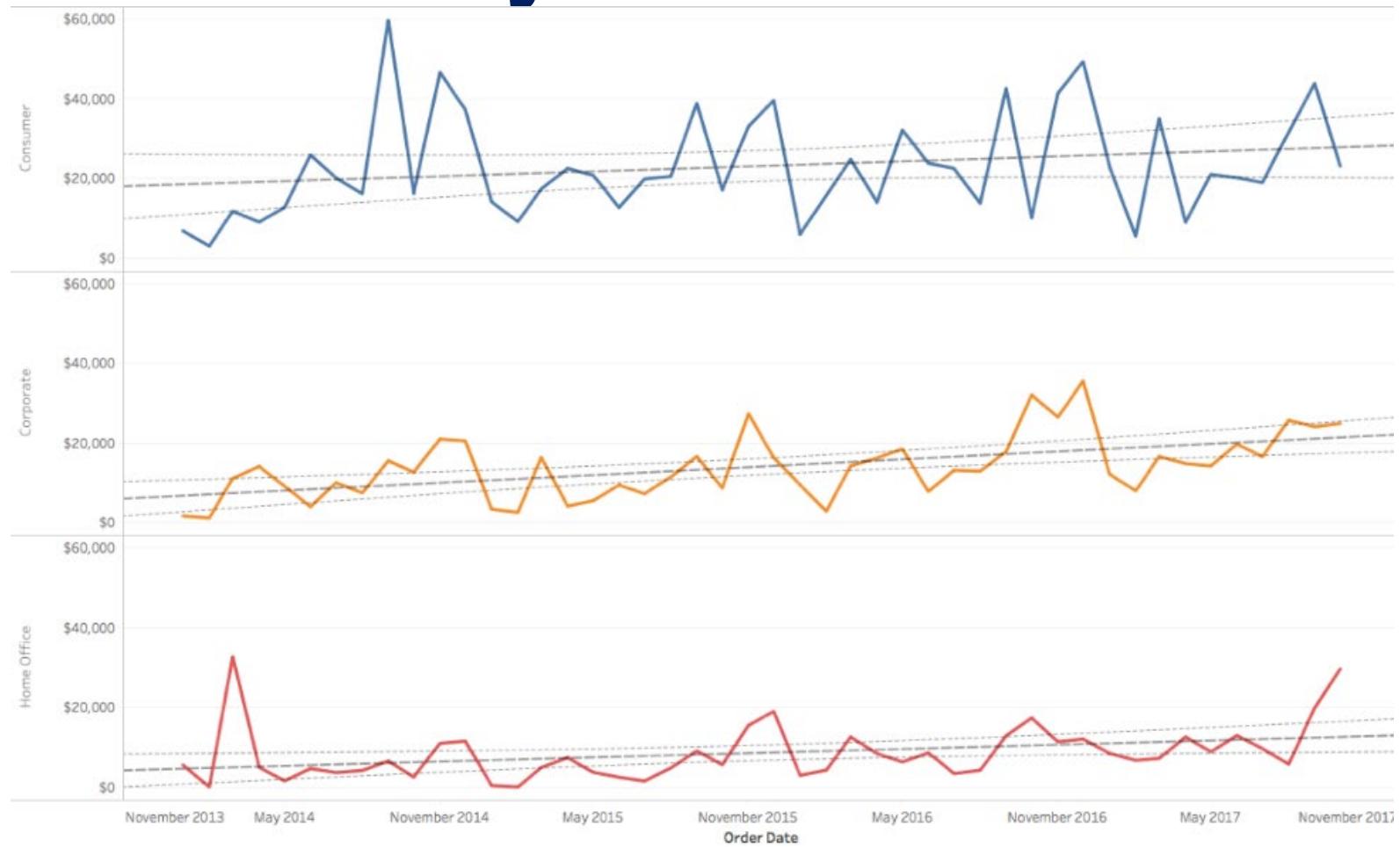
Sources of uncertainty in data

- Measurement: are we sure about the data?
- Forecast uncertainty: what will happen with the data next?
- Model uncertainty: how the data fit together?

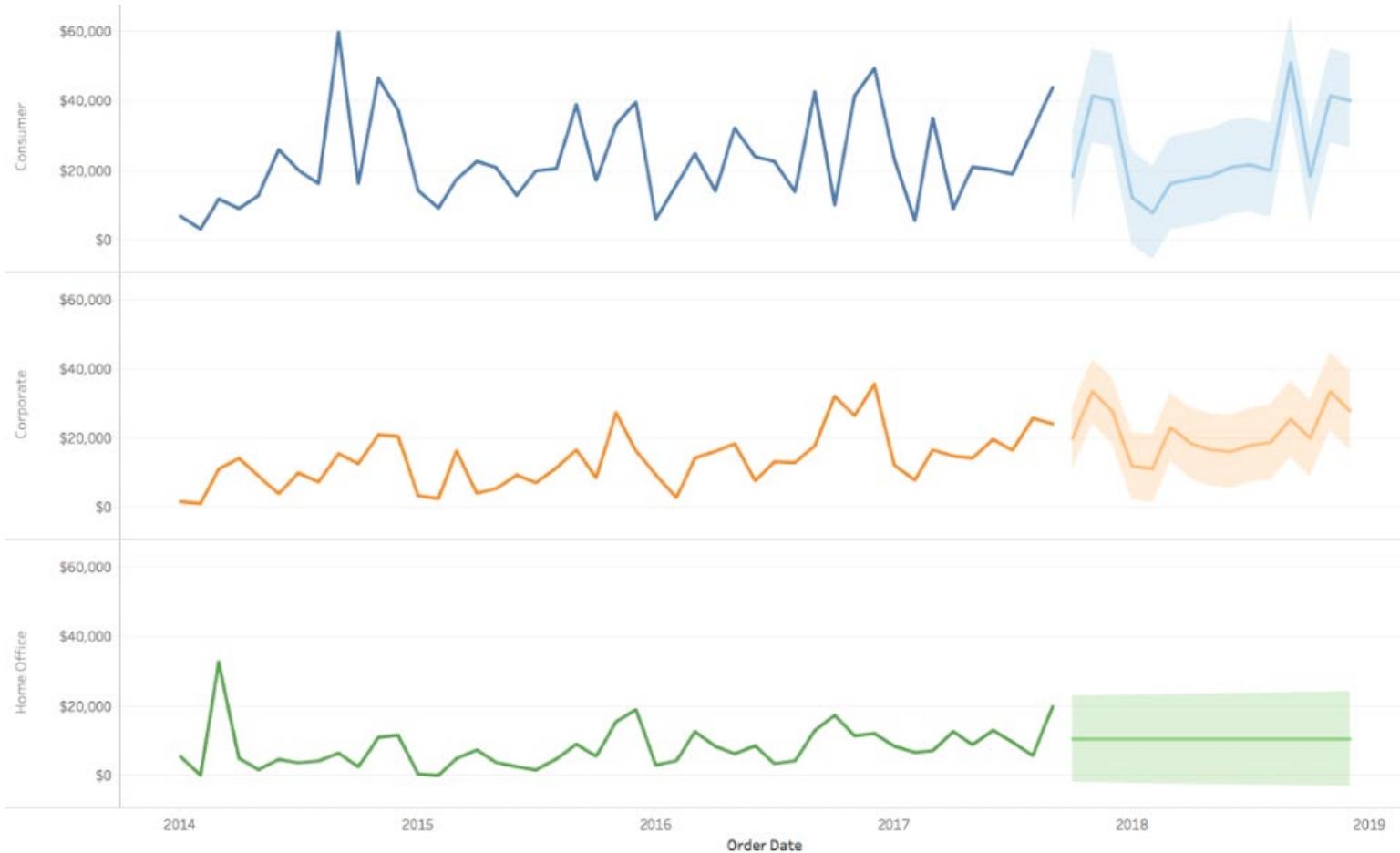
Measurement uncertainty



Model uncertainty



Forecast uncertainty



Approval rating

FiveThirtyEight

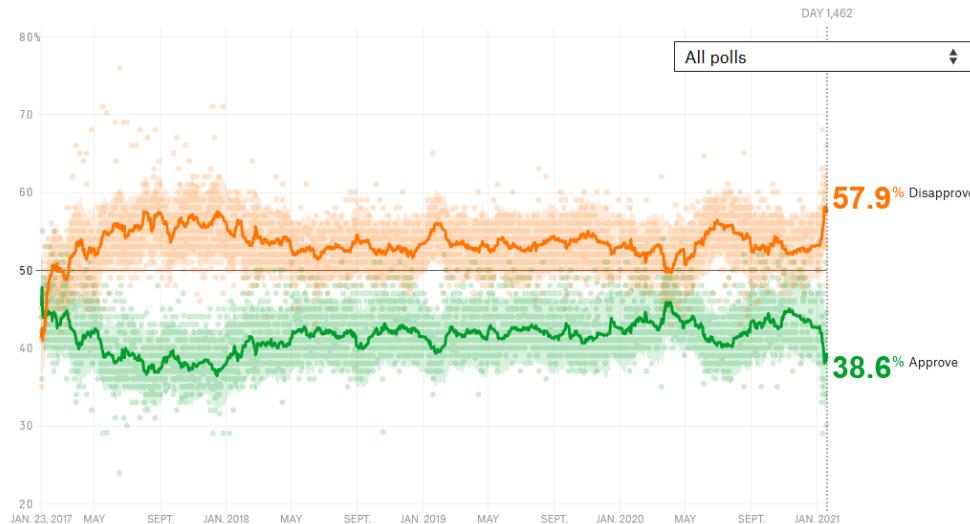


Our Trump approval rating polling averages are final and no longer updating.
[How popular is Joe Biden? »](#)

UPDATED JAN. 20, 2021 AT 11:57 AM

How unpopular is Donald Trump?

An updating calculation of the president's approval rating, accounting for each poll's quality, recency, sample size and partisan lean. [How this works »](#)



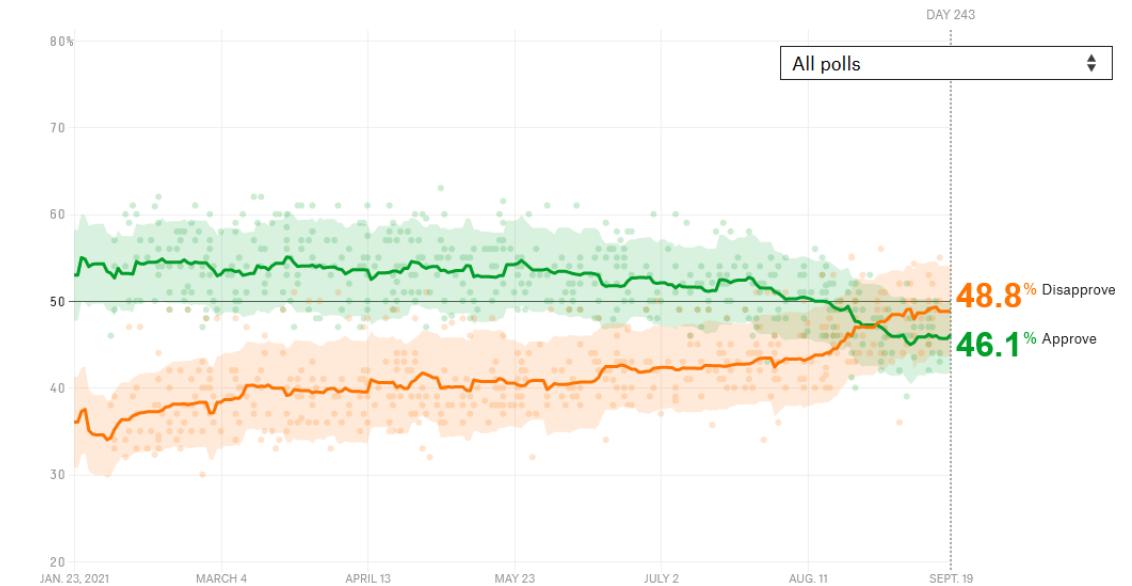
FiveThirtyEight



UPDATED SEP. 10, 2021, AT 1:37 PM

How unpopular is Joe Biden?

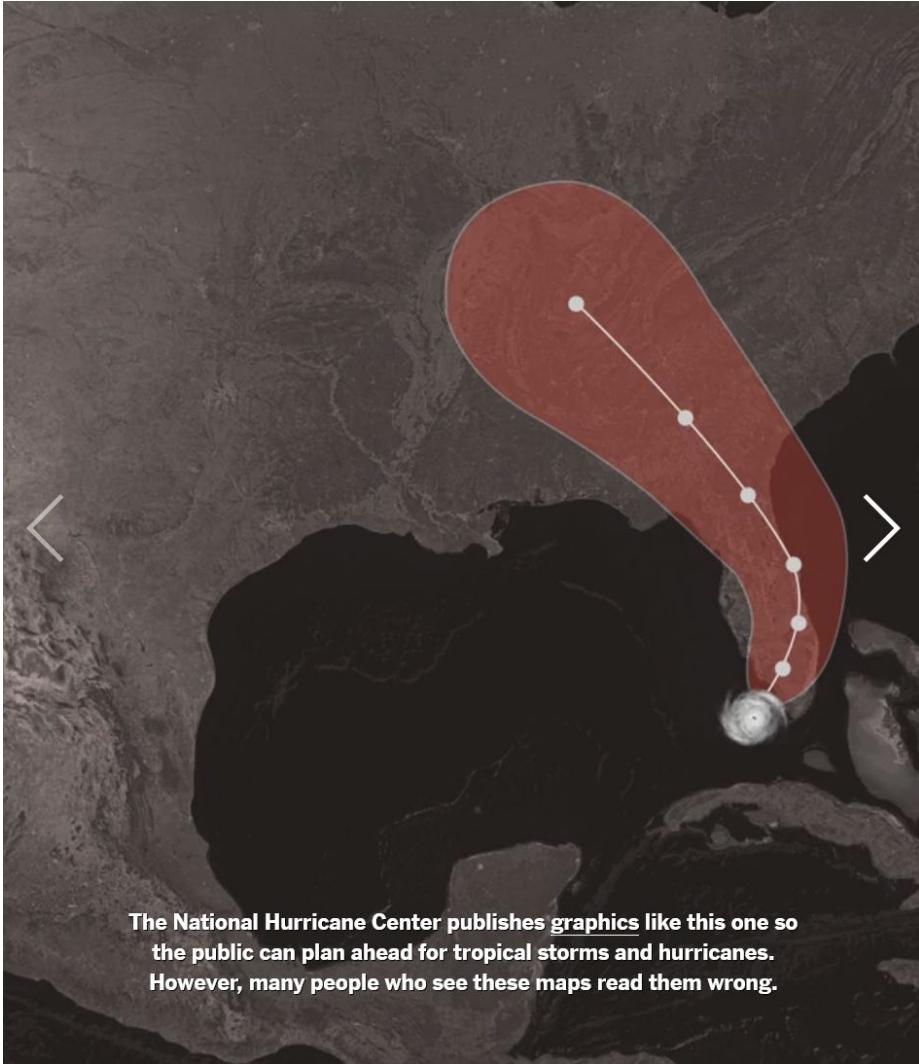
An updating calculation of the president's approval rating, accounting for each poll's quality, recency, sample size and partisan lean. [How this works »](#)



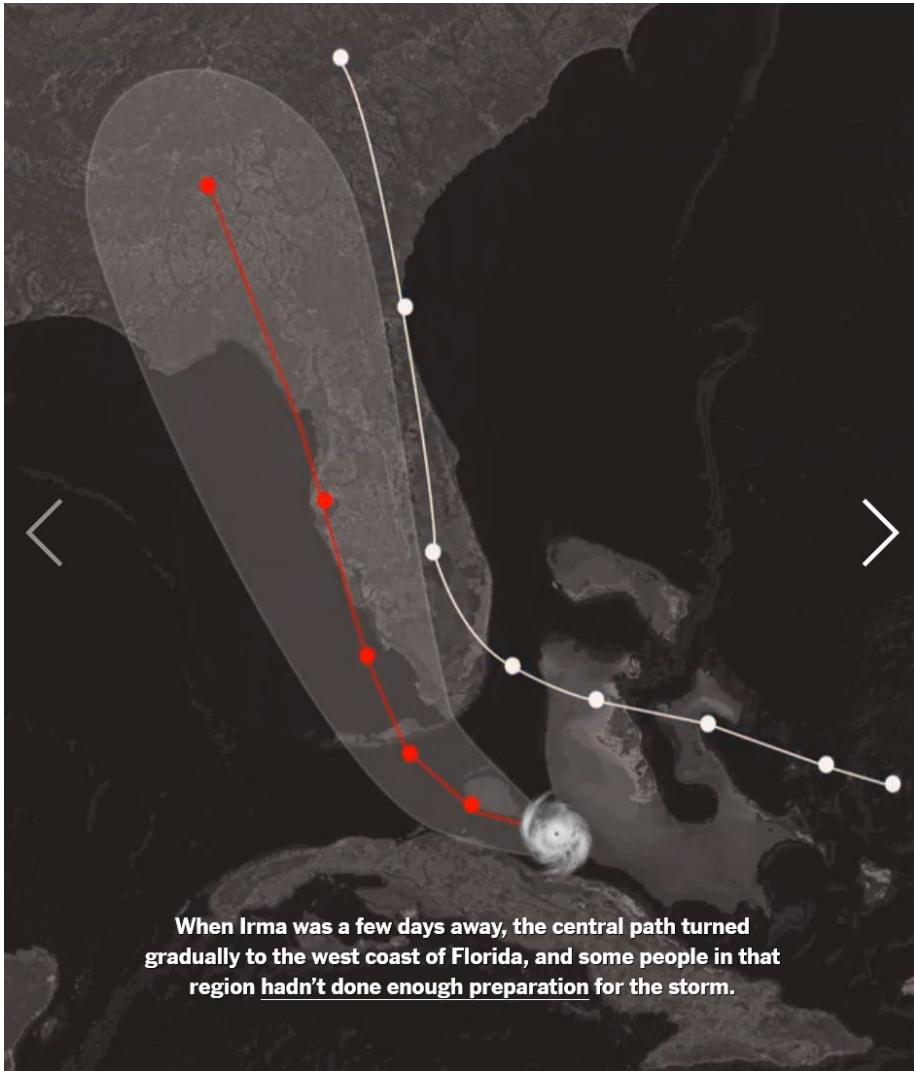
Hurricane Irma



Hurricane Irma



Hurricane Irma



Hurricane Irma



Chance of rain



100% chance in 1/3 of the city

0% chance in 2/3 of the city

33% chance for the city

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.”

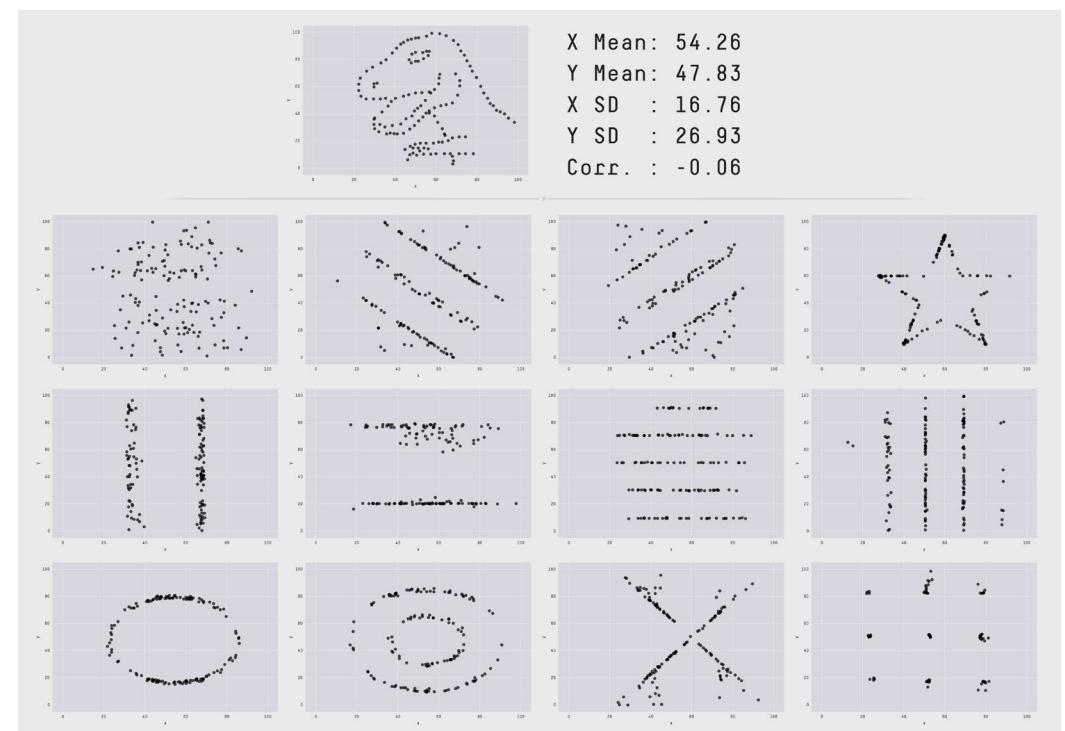
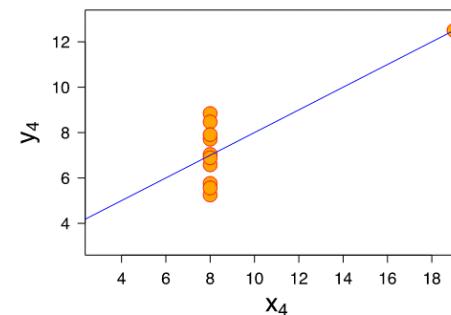
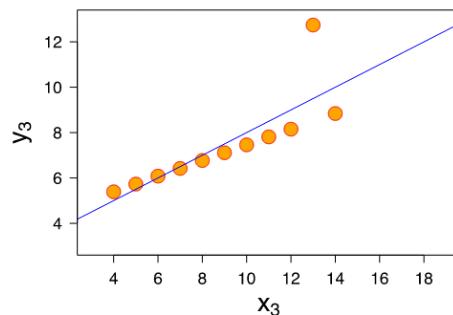
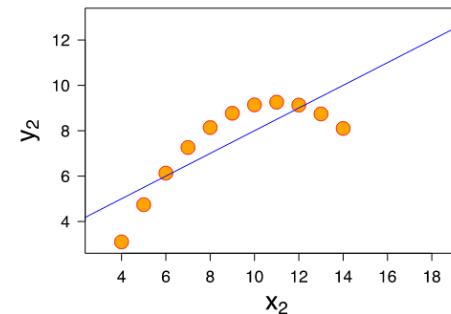
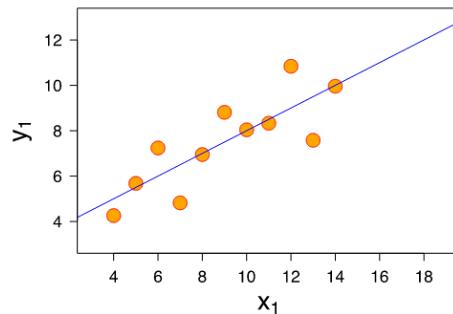
Michael Correll
Tableau Research

Sources of uncertainty in data

- Data ensembles: collections of simulations that explore a parameter space or realize a collection of instances.
- Model inadequacy: lack of knowledge of true phenomena that is being approximated.
- Algorithmic approximations: caused when translating models to computational settings.

Data aggregation generates uncertainty

- Same stats, different graphs [Matejka and Fitzmaurice, 2017]



Summary statistics

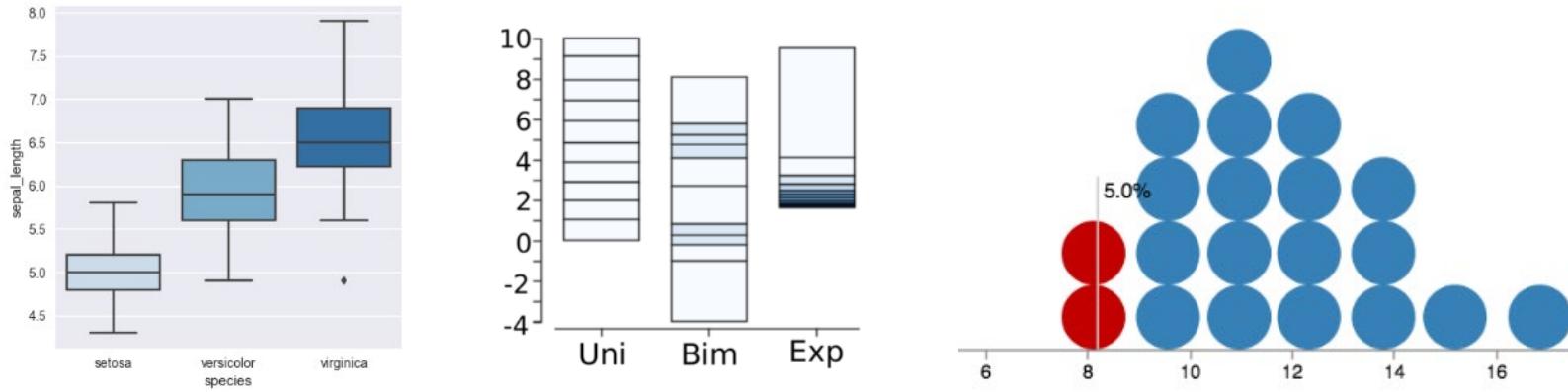
- Summarizes statistical information
- Measures of locality:
 - Mean
 - Median
 - Interquartile mean
- Measure of spread:
 - Standard deviation
 - Range
 - Variance
 - Absolute deviation
 - Interquartile range

Uncertainty visualization

- When you have a whole range or distribution of numbers, avoid visualizing single numbers.
- Graphical annotations of distributional properties
 - Histograms and density plots
 - Intervals and ratios
 - Distributions
- Visual encodings of uncertainty
- Hybrid approaches

Possible solution: quantiles

- Robust statistics with clear interpretation



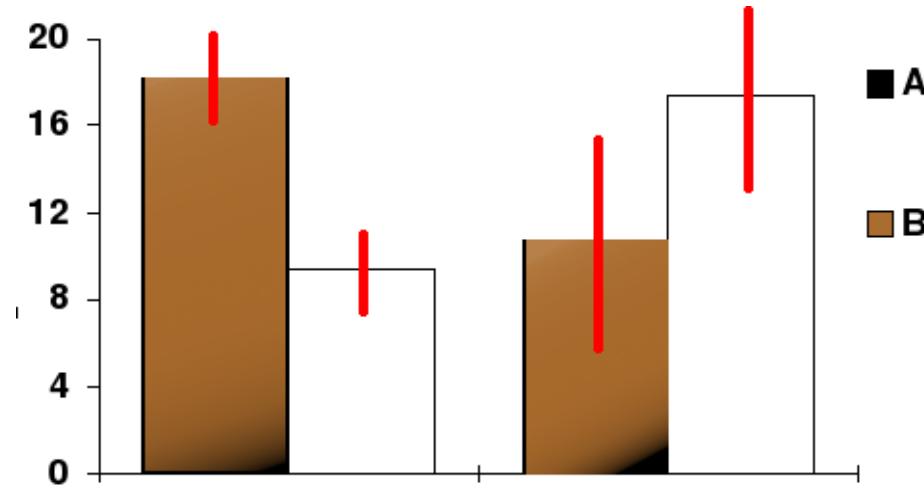
- Important: quantiles cannot be merged

$$\text{Mean}(A \cup B) = \frac{(\sum A + \sum B)}{(|A| + |B|)}$$

$\text{Median}(A \cup B)$ is not a function of $\text{Median}(A)$ and $\text{Median}(B)$

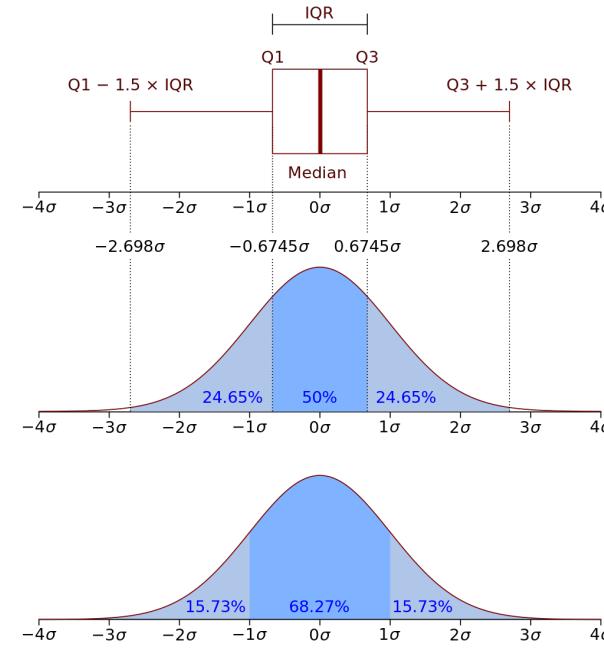
Intervals and ratios

Error bars: show variability of data, indicating error or uncertainty.



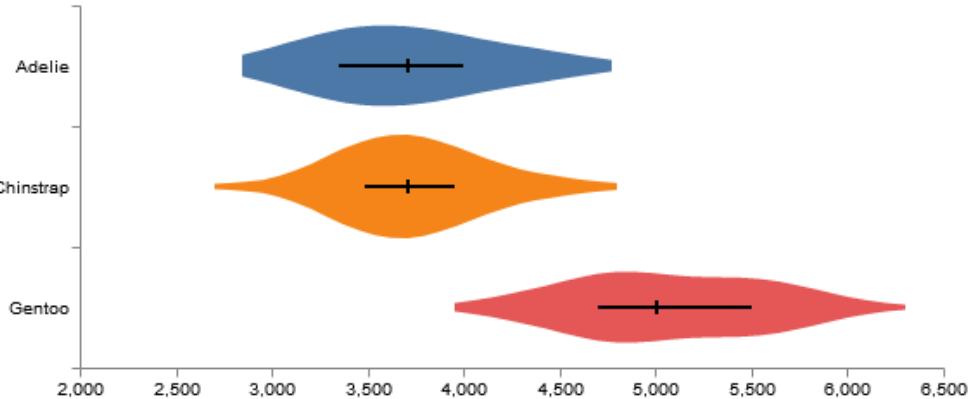
$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i + \mu)^2}$$

Error bars: show variability of data, indicating error or uncertainty.

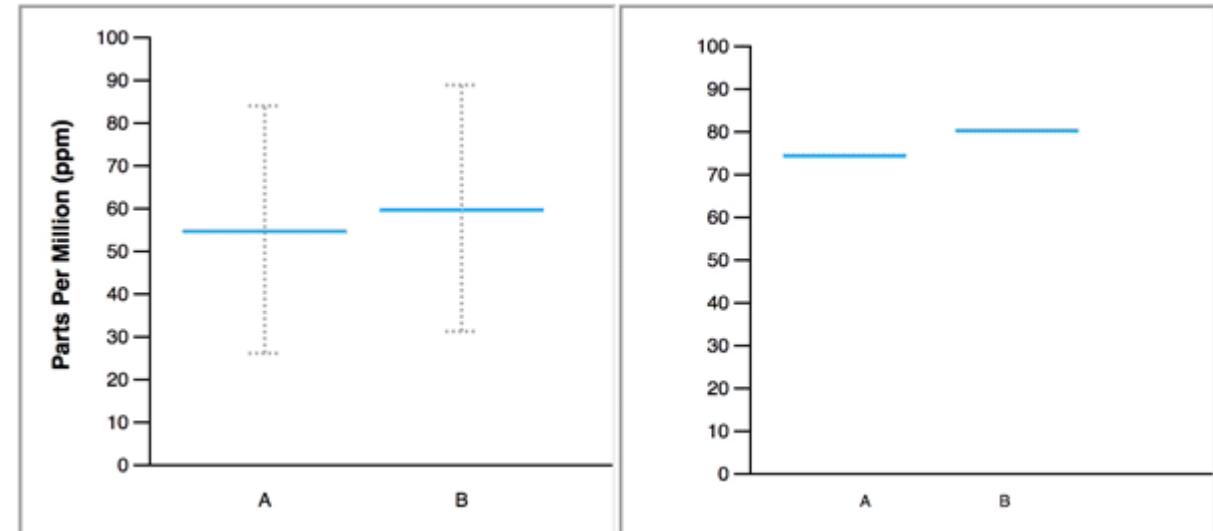


Distributions

Violin plot: similar to box plots, but show probability density at different values.



Hypothetical outcome plots: visualize a set of draws from a distribution.



Quantile dot plots



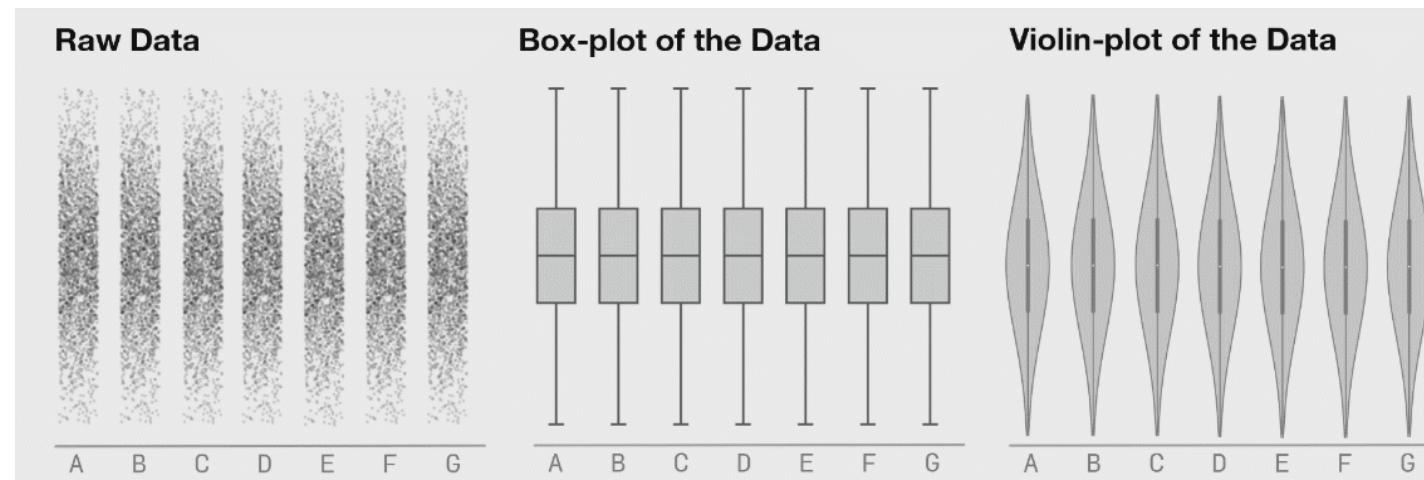
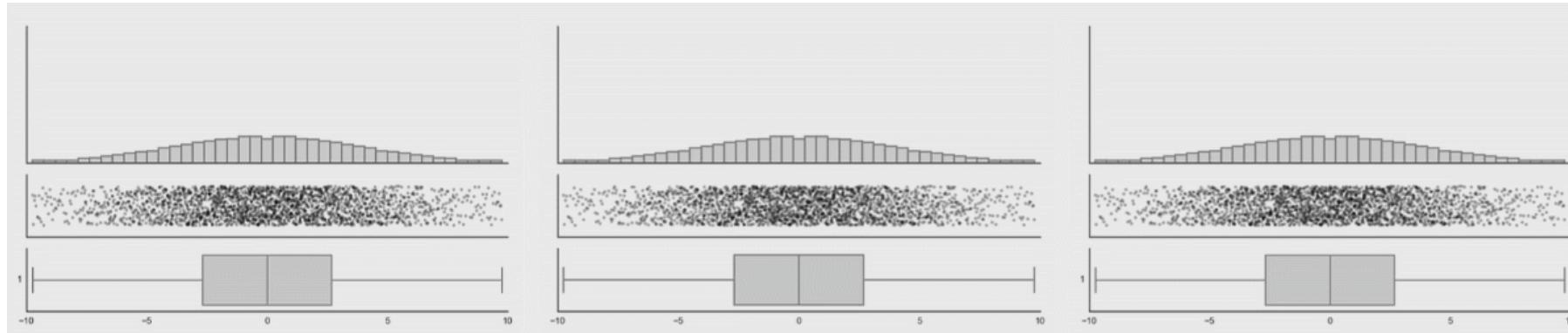
Hypothetical outcome plots



Interaction maps



Data aggregation generates uncertainty

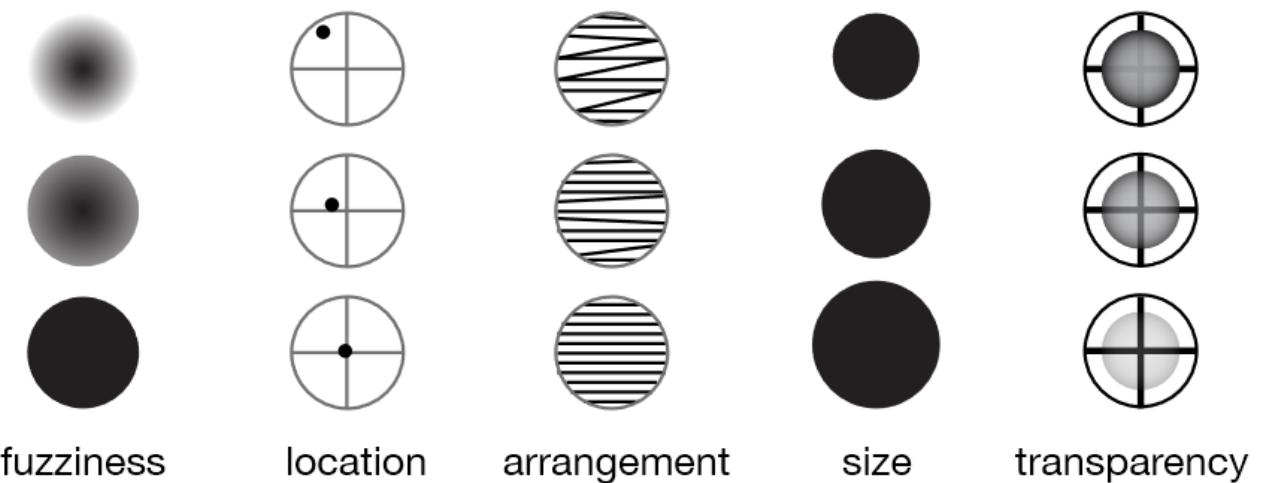


Distributions

- Histogram and density plots
- Violin plot
- Gradient plot
- Hypothetical outcome plot
- Quantile dot plot
- Ensemble plot

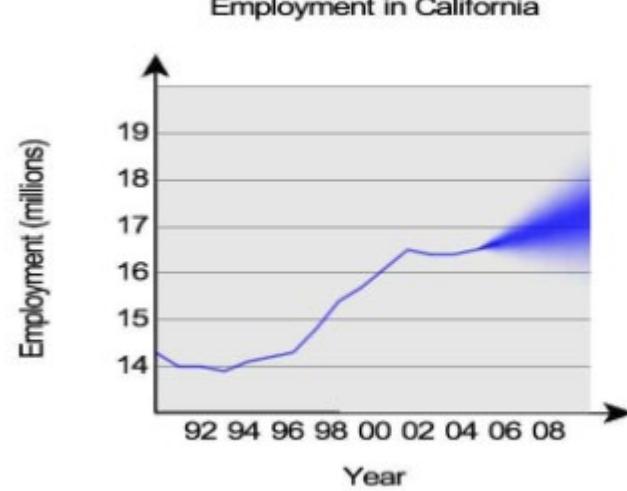
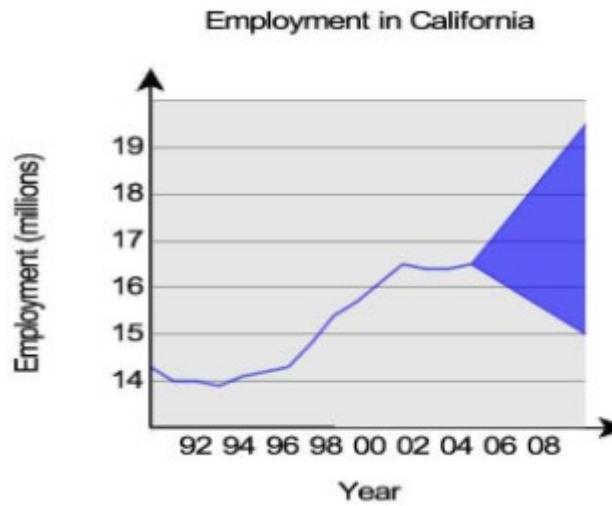
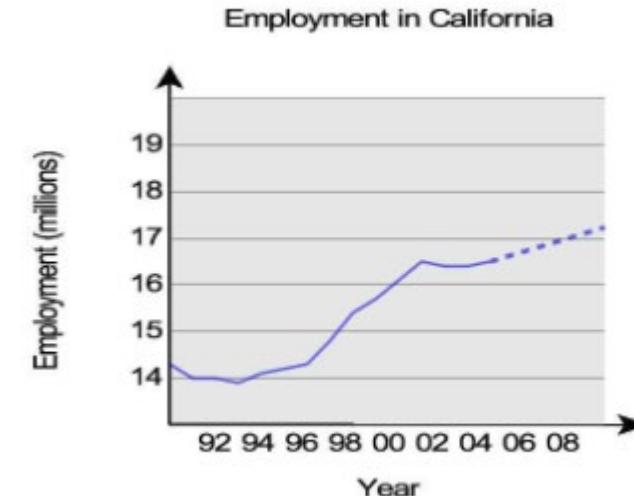
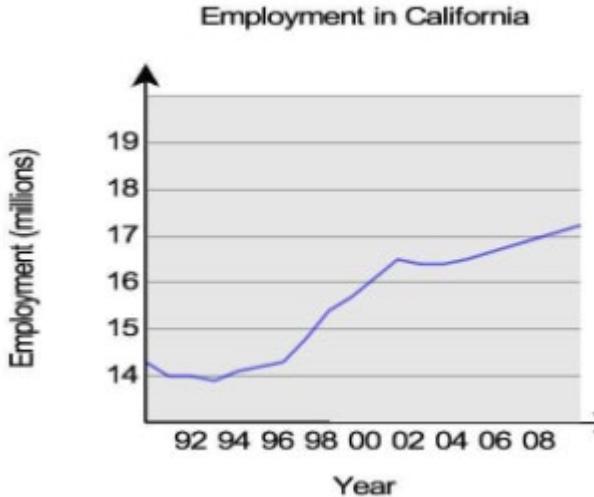
Visual encodings of uncertainty

- Additional visual channels can be used to characterize the uncertainty:
hue, texture, opacity
 - Fuzziness
 - Location
 - Arrangement
 - Size transparency



[Padilla et al., 2020]

Visual encodings of uncertainty



[Streit et al., 2007]

Hybrid approaches

Contour boxplot:

