



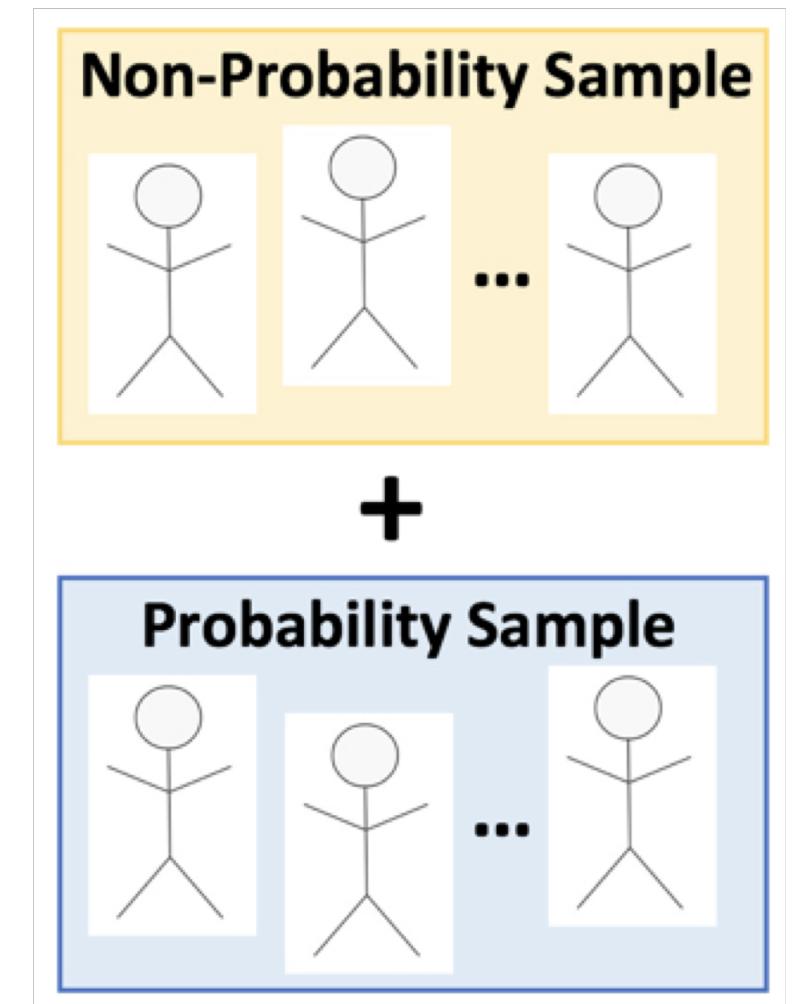
# Non-Probability Sampling, Part 2

*Brady T. West*

# Population Inference Approaches

## “Pseudo-Randomization Approach”

- **Combine non-probability sample with a probability sample**
- **Estimate probability of being included in non-probability sample** as a function of auxiliary information available in both samples
- **Treat estimated probabilities of selection as “known”** for non-probability sample, use probability sampling methods for analysis



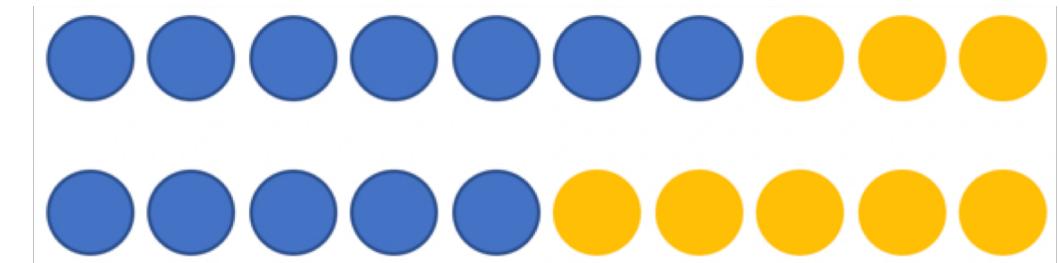
# Population Inference Approaches

## “Calibration” Approach

- **Compute weights for responding units** in non-probability sample that allow weighted sampled to mirror a known population

**Non-probability sample:** 70% female, 30% male

**Population:** 50% female, 50% male



→ **Down-weight females and Up-weight males**

# Population Inference Approaches

## “Calibration” Approach

- **Compute weights for responding units** in non-probability sample that allow weighted sampled to mirror a known population
- **Limitation:** if weighting factor not related to variable(s) of interest  
→ will not reduce possible sampling bias

# Twitter Example: Non-Probability Sample

API to extract info from several hundred thousand tweets  
and indicator of support for President Trump computed

- **Probability** of a tweet being selected **cannot be determined**
- **Twitter users not a random sample** of larger population
- **Lots of data, but** ,,,
  - high potential for sampling bias
  - lack of representation: may only capture people with strong opinions!



[Logo from Twitter](#)

# What's Next?

- **Sampling distributions and sampling variance** ~  
how to estimate features of these distributions  
*based on only one probability sample*
- **Examples of making population inferences**  
*based on type of sample selected*
- Introduce **model-based** approaches to analyzing data