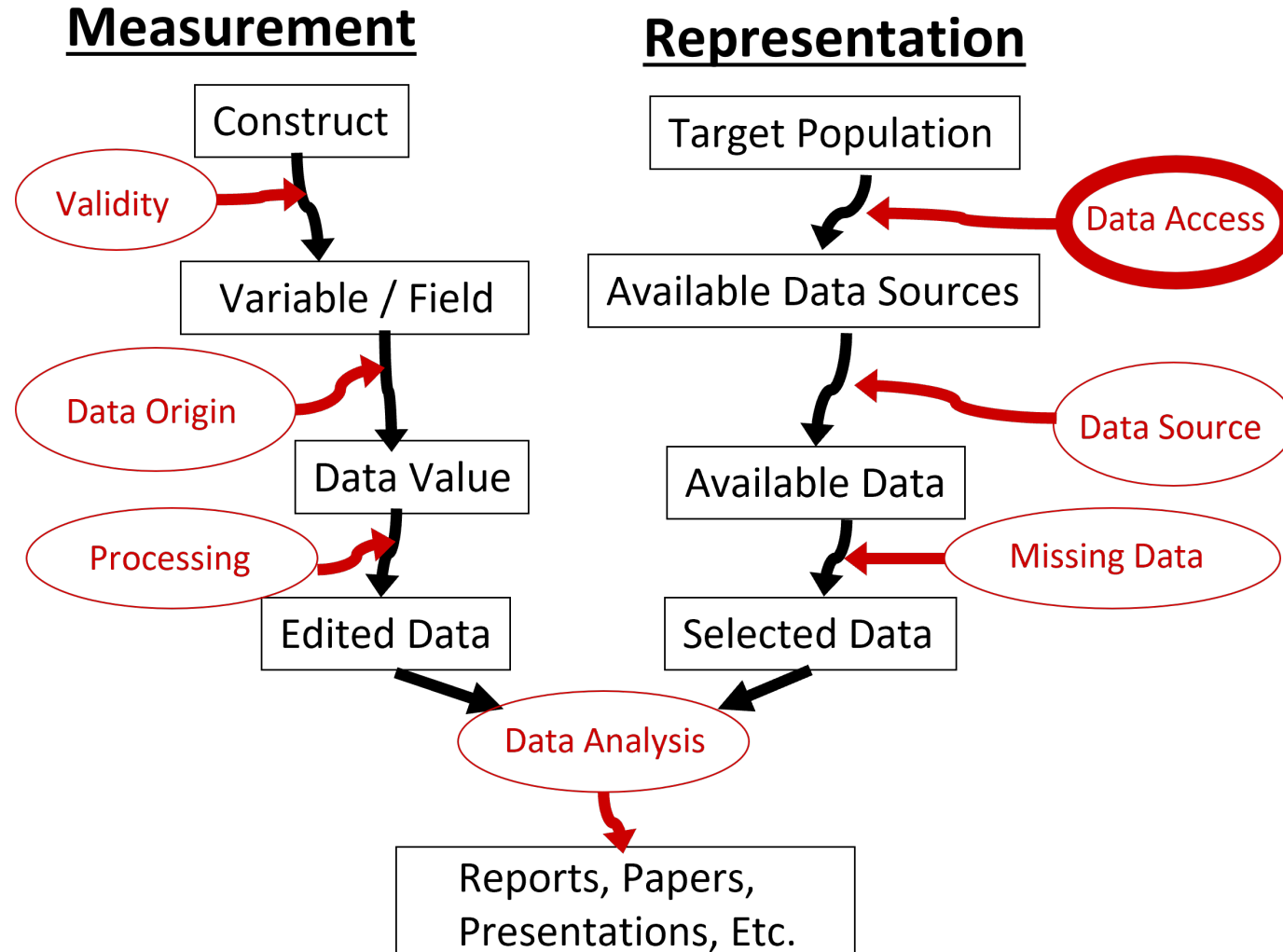


Data Access Threats: Designed Data By James Wagner

Dimensions of TDQ: The Big Picture!



Data Access Threats: Designed Data

- Researchers collecting designed data access the population through a **sampling frame**
- A sampling frame is a **set of materials or processes** through which the population is accessed
 - Example: List of persons
 - Example: Random digit dialing

Sampling Frame Problems

- Sampling frames are often imperfect
- Four main issues
 - Overcoverage
 - Undercoverage
 - Duplication
 - Clustering

Sampling Frame Problems: Overcoverage (1)

- Frame includes ineligible units
- Cost issue – need to identify eligibles

Sampling Frame Problems: Overcoverage (2)

- Example: Studying a rare group
 - Persons with diabetes, 10% of US population
 - No list of persons with diabetes exists
 - Sampling frame: A list of all residential addresses in the US
 - Many sampled units will not contain a person with diabetes

Sampling Frame Problems: Undercoverage (1)

- Frame does not include all eligible units
- Possible bias – are undercovered units different?

Sampling Frame Problems: Undercoverage (2)

- Example: Random Digit Dialing (Cellular)
 - Target population is adults 18+
 - According to Blumberg and Lukes (2019), 96.4% of households have at least one cell phone
 - The 3.6% without a cell phone are different: older, more rural, less educated
 - Will this difference lead to bias in your measurement?



© Faculty Presenter

**Except where otherwise noted, this
work is licensed under CC BY-NC 4.0**