



1

Network Study Designs & Data

Study designs determine what you can observe

Which determines what you can model statistically

... And what you can't



Network data

Two key aspects to consider

- Network properties
 - **Domain**: what type of network data is this?
 - **Temporality**: cross-sectional or longitudinal measurement?
- Measurement properties
 - **Sample**: what is the population of interest, and how was it sampled?
 - **Instrument**: how was the information collected?

1. Domain

(Network properties)

- Human social networks
 - Links can be contact/exchange/affect/role-based/genetic
 - Multi-level designs can include persons and places
- Animal networks
 - Links can be contact/movement/genetic
 - Multi-level designs can include animals and places
- Institutional networks
 - Links can be persons! (e.g., hospital transfers)
 - Or goods/money/etc.

2. Temporality

(Network properties)

- Cross-sectional designs collect data at one time point
 - This does not prevent you from collecting retrospective data
 - For example, on the start and end of a previous partnership
 - So this can be used this for dynamic modeling
 - If durational information is collected
- Longitudinal designs collect data at more than one time point
 - Panel designs vs. continuous measurement
 - Open vs. closed cohort

3. Sample

(Measurement properties)

- A network census is data on every node and every link
- Network sampling designs include:
 - Adaptively sampled networks
 - Link tracing designs (e.g., snowball, random walk, RDS)
 - Egocentrically sampled networks
 - Enroll population sample (“egos”)
 - Ask them about their partners (“alters”)
 - Optional: ask ego to report on alter-alter ties
 - Convenience samples
 - For example, online social networks

4. Instrument

(Measurement properties)

How are your data collected?

- Traditional designs
 - Interview (for humans)
 - Interviewer administered (face-to-face or T/CAPI)
 - Self administered
 - Passive observation and recording (for other types of nodes)
- Electronic passive capture
 - Scraping (web data)
 - Sensor data

What network statistics can you observe?

- Degree
 - Mean degree
 - Degree distributions
- Nodal attributes
 - Types of nodes
 - Heterogeneity in degree by nodal attributes
 - Mixing by nodal attributes
- Triads
 - And all of their possible configurations
- Larger configurations (which ones are of interest?)
- Timing of ties
 - Start/End
 - Duration of active and completed partnerships
- Other attributes of ties
 - Type
 - Intensity
 - Direction
- Multiplexity
 - Multiple tie types joining nodes

Next module

Statistical Models for Networks – Theory

- Principles of statistical inference with network data
- Static models: ERGMs
- Temporal/dynamic models: TERGMs