

## Annex. Frequently Asked Questions Concerning Social Network Analysis

### What software programs are popular for SNA?

Many standalone software programs for SNA have a graphical user interface and do not require programming experience. Some popular options of this type include Gephi and Ucinet. Additionally, most statistical programming languages include packages, libraries, or modules for SNA. Currently, among open-source options, R and Python have particularly good SNA support. A popular choice for Microsoft Excel users is NodeXL, a feature-rich Excel add-in for SNA.

### How computationally intensive is SNA?

SNA tends to be more computationally intensive than traditional statistical methods, as many of the core algorithms do not scale linearly with network size.\* For example, betweenness centrality requires computing the shortest paths between all pairs of actors, which exhibits cubic growth with network size. Creating efficient network algorithms and ways of calculating SNA metrics in parallel are active areas of research. However, for most network studies using survey responses, the networks are generally not large enough to create excessive computational burden. For example, we analyzed the case studies presented in this paper using a consumer-grade laptop. If researchers know they will be collecting a large amount of network data from administrative, bibliographic, or social media records, additional computational resources or careful choice of algorithms may be required to efficiently perform SNA.

### Is there a minimum sample size required for SNA?

Data collection considerations are highly dependent on the type of SNA analysis planned. The most popular SNA design is the whole network or “network census” approach that assumes all members of a population of interest are represented in your data. In practice, defining the population and relevant network boundaries can be challenging, and the approach requires high response rates, unless additional assumptions can be made about the missing data. Other SNA designs have different sample considerations that reflect the types of data collected and research questions of interest. For example, egocentric network analysis has the actor as the unit of analysis and therefore has the same sample size considerations as traditional statistical analyses.\*\* Multiple network designs, which compare network metrics across many mutually exclusive networks (e.g., classrooms), have sample size considerations similar to multilevel regression modeling,\*\*\* where both individual- and group-level sample sizes are important. Conversations between stakeholders and social network researchers can often help clarify these requirements before data collection.

### Where can I take a class on SNA?

The sociology departments of large universities typically offer SNA classes. Faculty in schools of social sciences, medicine, business, communication, or public health in dedicated centers, such as Duke University’s Network Analysis Center, may also offer such classes. A good place to start is by looking at the course catalog of your local university or a directory of graduate programs from a professional society such as the International Network for Social Network Analysis. If in-class instruction is not available in your location, platforms such as Coursera or EdX offer several SNA massive online open courses (MOOCs).