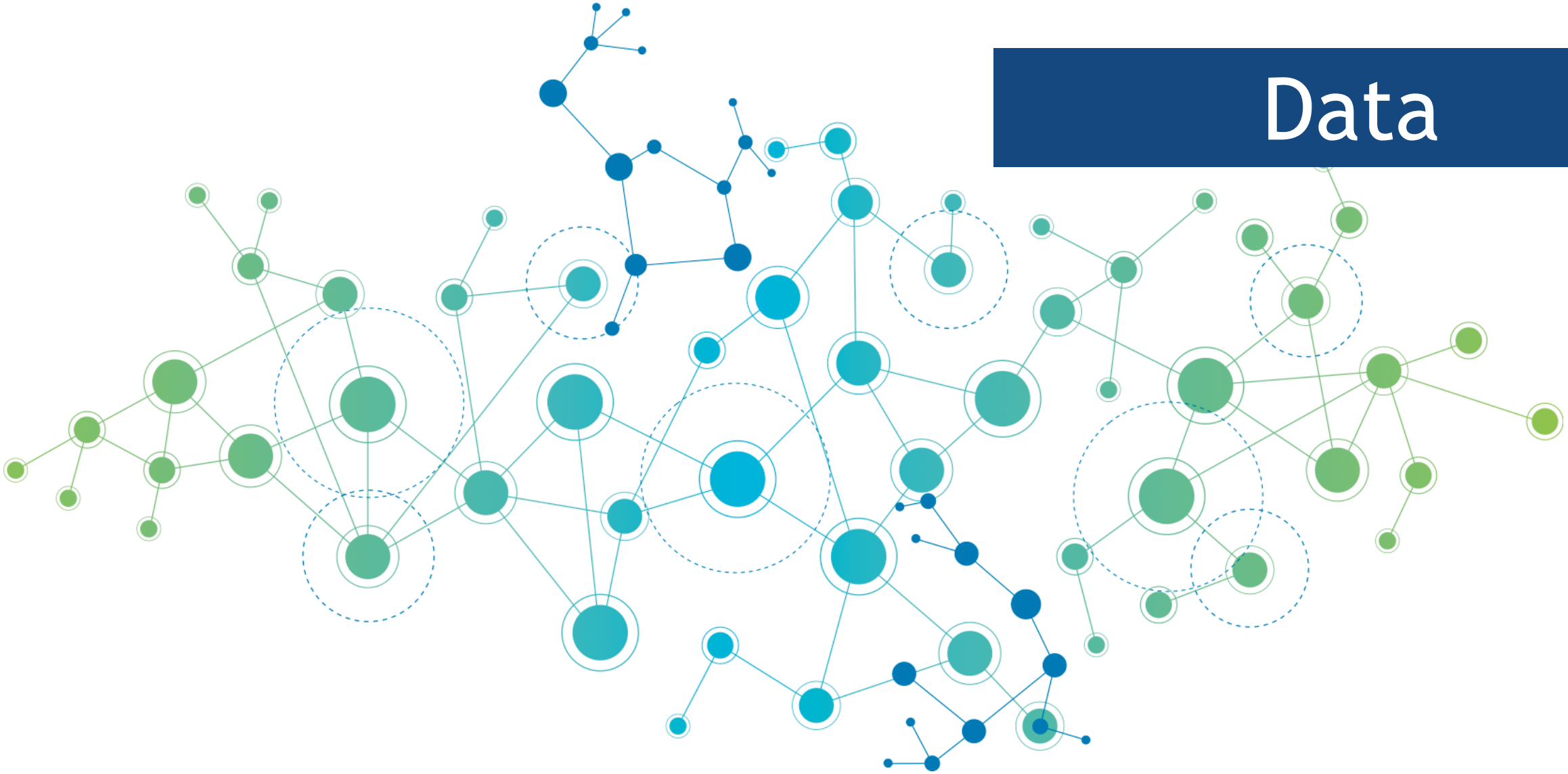
Data



Goals

- Understand the kind of information network analysis handles and how it can be displayed
- Learn how to decide who is in the network
- Learn how to collect network data
- Exploring various network data tasks in R

Network Data Structures

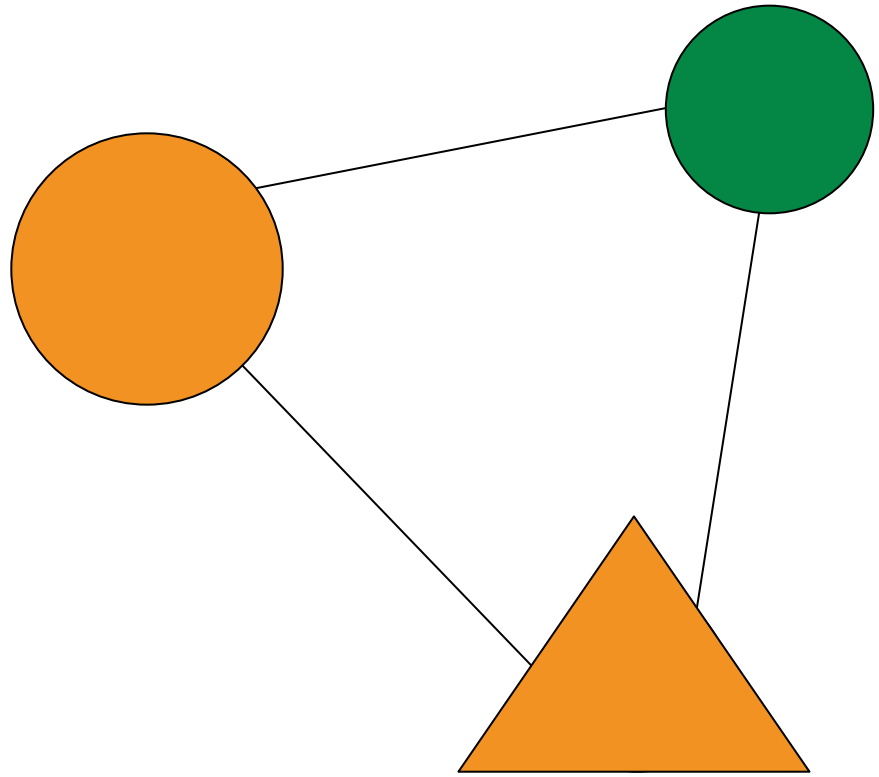
What kinds of data need to be in network objects?

Types of information in network data

Type	Description	Required?
Nodes	List of nodes in network, along with node names	Required
Ties	List of ties in the network	Required
Node attributes	Node characteristics	Optional
Tie attributes	Tie characteristics	Optional
Meta-data	Other information about the entire network	Depends

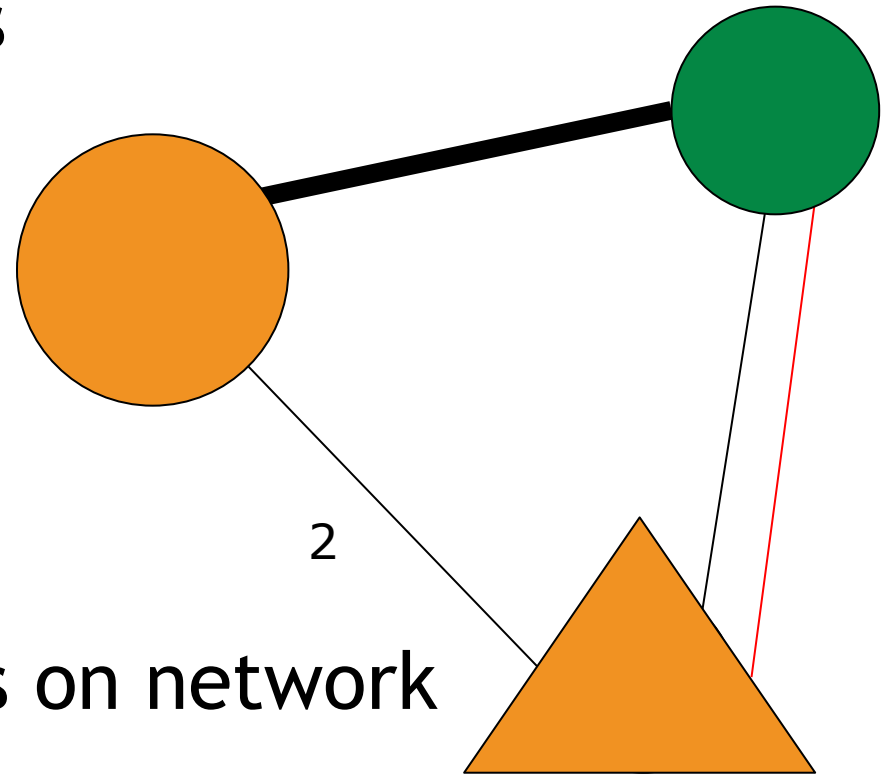
Node Attributes

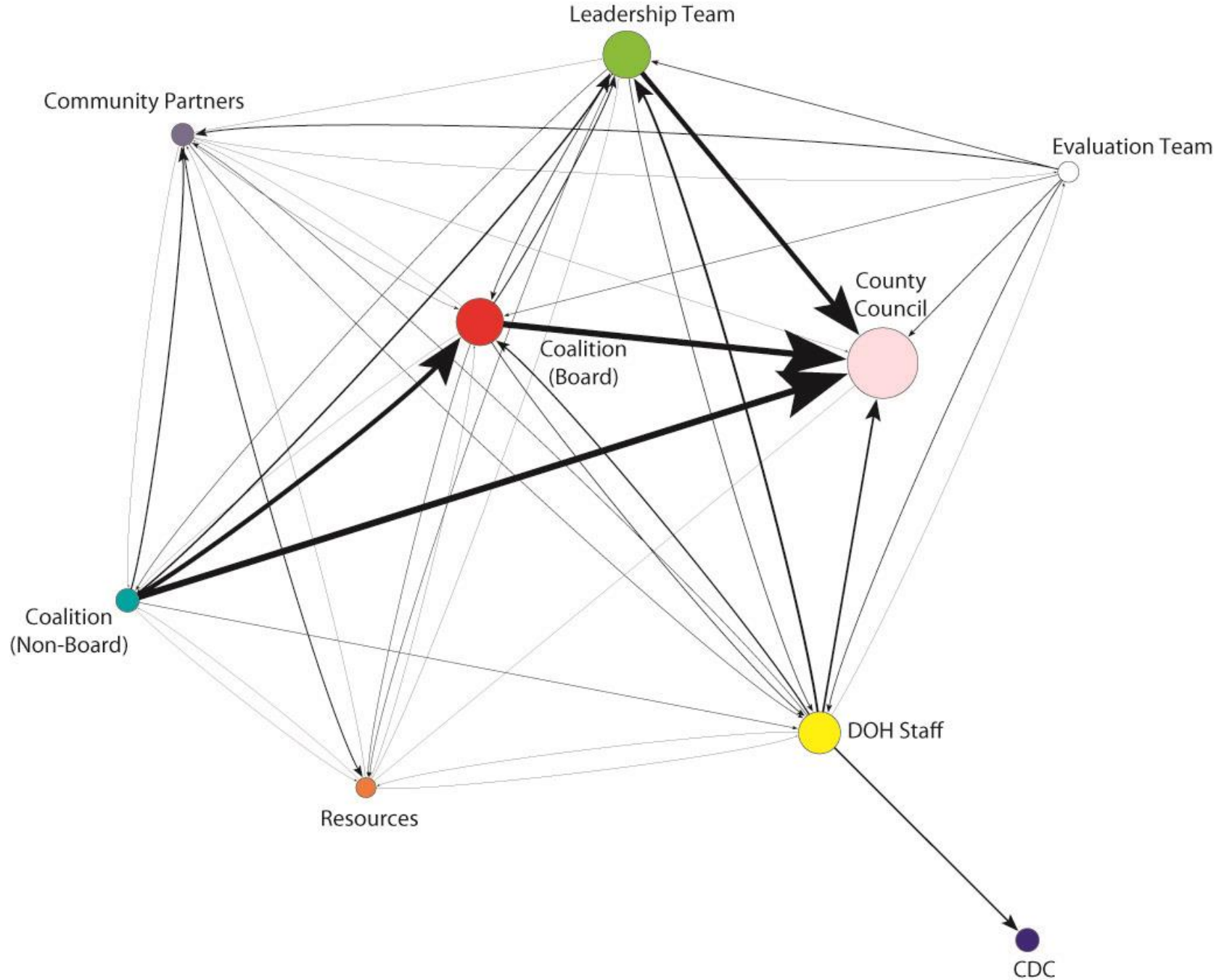
- Categorical - color & shape
- Continuous - size



Link Attributes

- Multiple relationships (multiplexity)
  - Multiple lines
  - Color optional
- Valued relationships
  - Number
  - Line thickness
- Visualization depends on network size




Leadership Team
Community Partners
Evaluation Team
County Council
Coalition (Board)
Coalition (Non-Board)
DOH Staff
Resources
CDC

Network data collection

Network delineation, boundary specification, data collection modes

Boundary Specification Goals

- Determine who is in the network = population of interest
  - “2012 ISSH attendees”
  - “St. Louis tobacco control partners”
- Try to measure everyone
  - High response rate more critical than in other kinds of research
- Including key actors is more important than the size of the network

Approaches

- Realist (internal) – network members could identify themselves and each other (i.e. club membership)
- Nominalist (external) – determined by external criteria, possibly researcher-defined (citation networks)
- Laumann criteria
  - Positional – formal membership
  - Reputational – knowledgeable person names members
  - Event – participation in certain activities
  - Relational – contact with others in the network

Example

National Networks Identification

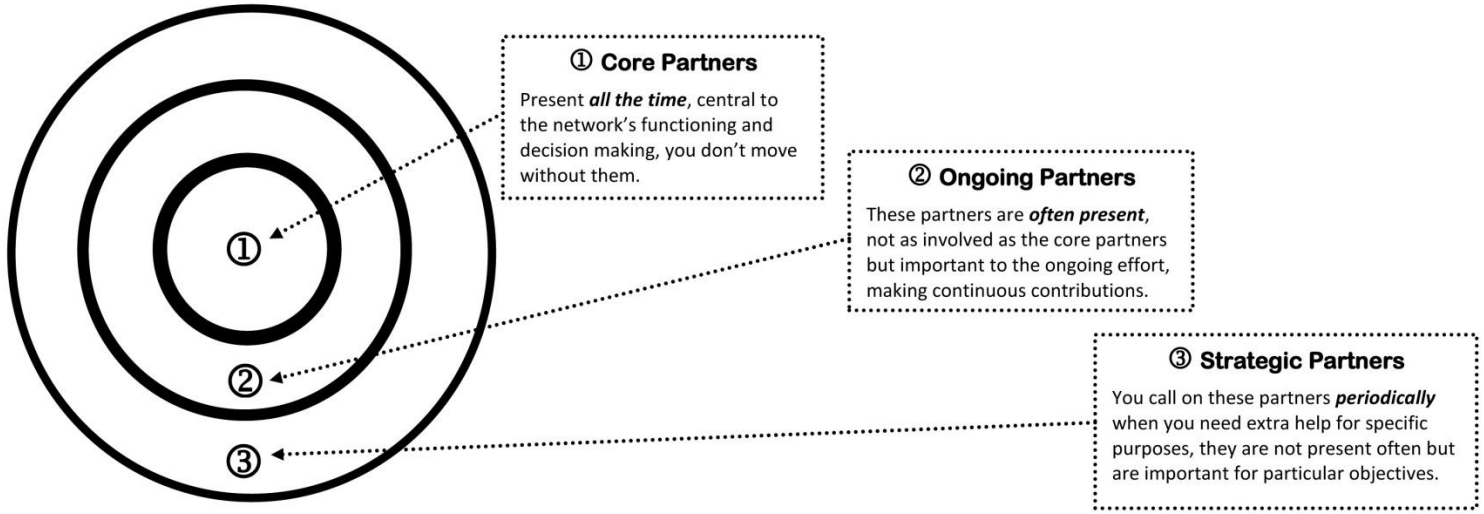
Goal: To collect an initial list of people who will participate in our web survey evaluation of partners working together to achieve common goals in tobacco control for the [specify population].

Definition: The *core, ongoing, and strategic network* is a partnership of critical agencies that are leading efforts in commercial tobacco control among the [specify network].

Identification criteria: The *core, ongoing, and strategic network* includes any critical organizations or individuals within your network who you have had direct contact with *in the last 12 months*, and meet *any* of the following criteria:

- Has provided input for strategic planning on tobacco control for the Network,
- Has conducted program and/or policy efforts in tobacco control at the national, regional, or local levels,
- Has helped enhance the ability of the Network to perform its activities and achieve its goals for tobacco control

The figure below may help you think of the variety of partners in your network. Please think of the partners you have in each category, and include them on the form that follows if they meet the above criteria. Don't worry if you have trouble sorting a partner into one category or the other – it is more important to have a qualified partner included somewhere than it is to have them "correctly" sorted. Please list only one individual per organization. Please provide phone and email information for all partners. This is very important, as that will enable us to contact each individual. Feel free to include multiple phone numbers or email addresses for each.



*Adapted from the Center for Managing Chronic Disease at the University of Michigan

Roster vs. Free Recall

- Roster: respondent given entire list of network members to answer questions about

- Example roster screening question
  - All network actors appear
  - “No” responses do not appear in later questions

Are you aware of the following individuals' work regarding tobacco control?

	Yes	*No*
Marty Dannenfelser, ACF	☐	☐
Anne Linehan, ACF	☐	☐
Harriett Bennett, AHRQ	☐	☐
Elizabeth (Beth) Edgerton, AHRQ	☐	☐
Ernstene (Tina) Murray, AHRQ	☐	☐
Christine Williams, AHRQ	☐	☐
Rosemarie Henson, CDC	☐	☐
Corinne Husten, CDC	☐	☐
Laura Kann, CDC	☐	☐
Paula Kocher, CDC	☐	☐
Brick Lancaster, CDC	☐	☐
Ann Malarcher, CDC	☐	☐
Jeffrey (Jeff) McKenna, CDC	☐	☐
David E. Nelson, CDC	☐	☐
Terry Pechacek, CDC	☐	☐
Patricia Richter, CDC	☐	☐
Charlotte Schoenborn, CDC	☐	☐
Michael Schooley, CDC	☐	☐
Dana Shelton, CDC	☐	☐
James (Jim) Coan, CMS	☐	☐
Jyme Schafer, CMS	☐	☐
Steve Galson, FDA	☐	☐

Roster vs. Free Recall

- Roster: respondent given entire list of network members to answer questions about
- Free recall: respondent asked to identify network members
  - Name generator questions used to elicit a set of network members
    - “Who are the 10 people most important to the completion of grant activities?”
    - “Who are the 10 people you work with most on grant activities?”
    - “Who are your 5 closest friends?”
  - Common to use multiple name generators to ensure adequate coverage
  - Caution: be *very* clear in your instructions

Example free recall name generator question

In this section, we want to find out who you feel is most important to the completion of the CPPW activities. The major CPPW activities are:

- Youth media campaign,
- Strengthening the County smokefree workplace ordinance,
- Strengthening ordinances in targeted municipalities,
- Strengthening K-12 and/or higher education policies,
- Augmenting required signage to include a graphic warning,
- Auditing tobacco retailers for compliance with FDA and County regulations,
- Increasing use of the Missouri Quitline, and
- Increasing workplace and/or general cessation provision.

Please identify up to 15 people who you think are the *most important to the successful completion* of the above CPPW activities. (Do not include those receiving interventions. In order for your information to be useful, you *must* include the names of *individual people* in the spaces for First and Last Name. Please include only one name per space.)

Important: 1

First Name	John
Last Name	Smith
Organization	DOH

Snowball Sampling

- Use free recall name generators w/ starting set of actors
- New names get recruited and asked the same name generator questions
- Repeat the process until some natural or imposed limit is reached. Sometimes the same actors are nominated after a certain amount of time → “natural” network has been identified.
- May want to add new members only if they’ve been mentioned more than once.
- Be realistic with workload estimates

Network data sources

Where can you find or obtain network data?

Archival/ Secondary

- Web of Science/Elsevier
- Administrative sources
- Listserv posts
- Emails
  - Enron
  - (https://www.gregorybufithis.com/2019/04/04/text-analytics-the-enron-data-set-and-andrew-fastow-get-a-new-role/)
- Facebook, any social media

Collect your own

- Interviews
  - Telephone
  - In-person
- Surveys
  - Internet/web based
  - Paper
- Observation

Example node attribute questions

- Keep consistent with level of analysis
  - Individuals
  - Organizations

How long have you been employed at your agency?

- ☐ *less than 1 year*
- ☐ *1-5 years*
- ☐ *6-10 years*
- ☐ *11-15 years*
- ☐ *16 or more years*

How long have you been in your current position?

- ☐ *less than 1 year*
- ☐ *1-5 years*
- ☐ *6-10 years*
- ☐ *11-15 years*
- ☐ *16 or more years*

How long has your work included a focus on tobacco?

- ☐ *I do not work in tobacco*
- ☐ *less than 1 year*
- ☐ *1-5 years*
- ☐ *6-10 years*
- ☐ *11-15 years*
- ☐ *16 or more years*

Interview/Survey question formats

- Free vs. fixed choice
  - Free: no limit on nomination (“How much time do you spend with each of your best friends?”)
  - Fixed choice: specific or upper limit (“How much time do you spend with your 3 best friends?”)
- Ratings vs. rankings: name interpreter (item used to measure relational characteristics between actors)
  - Rating – score for each network member
  - Ranking – rank or sort each network member

Example name interpreter question

Which response best describes your current level of collaboration (within the last 12 months) between you and the following partners?

	Do not work together	Share information only	Work together informally	Work as a formal team	Work as a formal team on multiple projects
Person 1 - Organization 1	◉	○	○	○	○
Person 3 - Organization 1	◉	○	○	○	○
Person 4 - Organization 1	○	◉	○	○	○
Person 6 - Organization 2	○	◉	○	○	○
Person 8 - Organization 2	○	◉	○	○	○

Example multiple relations question

For the people you named earlier, what types of CPPW activities have you worked together on? (Check all that apply.)

	Youth Media Campaign	*Smokefree County Ordinance*	*Smokefree Municipality Ordinances*	*K-12/ Higher Ed Policies*	*Retailer Graphic Warning Ordinance*	*Promotional Item/ Discounted Product Ban Ordinance*	*Missouri Quitline*	*Workplace Policies/ Cessation*	*None of these*
John Smith -- Centers for Disease Control and Prevention	☐	☐	☐	☐	☐	☐	☐	☐	○
Jen Thompson -- St Louis County Department of Health	☐	☐	☐	☐	☐	☐	☐	☐	○
Sean Jones -- Clear Channel Radio	☐	☐	☐	☐	☐	☐	☐	☐	○
Janet Shelton -- St Louis Post Dispatch	☐	☐	☐	☐	☐	☐	☐	☐	○

Content types

Type	Examples
Filter/Recognition/Name Generator	Are you aware of this person's work in tobacco control?
	Name 15 people who you work with on your CDC grant.
Relationship Strength	Contact (Daily, Monthly, Weekly)
	Level of collaboration (Share Information, Informal Partnership, Formal Partnership)
Evaluative (good/bad)	How satisfied are you with the quality of communication you have with each of your partners?
	What barriers have you experienced in your work with each of your partners? (Time, Capacity, Politics)
Flow	Which of the following clinics do you refer patients to?
	Which of the following agencies do you receive funding from?

Participant burden

- Each question gets asked for each actor in the network
- True load = # of actors answered for X # of questions

Actor	Q1	Q2	Q3	Q4
Person 1	1	2	3	4
Person 2	2	4	6	8
Person 3	3	6	9	12
Person 4	4	8	12	16
Person 5	5	10	15	20
…				
Person 100	100	200	300	400

Measurement quality

- Accuracy higher for stable/long-term characteristics
- Individuals as representatives of organizations
- Reliability higher w/ closer relationships and when ratings/rankings done for all network members
- Roster format more reliable than free recall

Online (network specific)

Tool	URL
PARTNER Tool	http://www.partnertool.net/
C-IKNOW	http://ciknow.northwestern.edu/
ONA	https://www.s2.onasurveys.com/
FirmNet Online	http://www.firmnetonline.com/
Network Genie	https://secure.networkgenie.com

Online (non-network specific)

Tool	URL
Survey Monkey	http://www.surveymonkey.com/
REDCap	http://project-redcap.org/
Snap	http://www.snapsurveys.com/
Qualtrics	http://www.qualtrics.com/
Roster Example	http://sswwustl.qualtrics.com/SE/?SID=SV_3wm7g5nIWDOVIGg
Free Recall/Name Generator Example	http://sswwustl.qualtrics.com/SE/?SID=SV_5vin0B5WbqmegtK