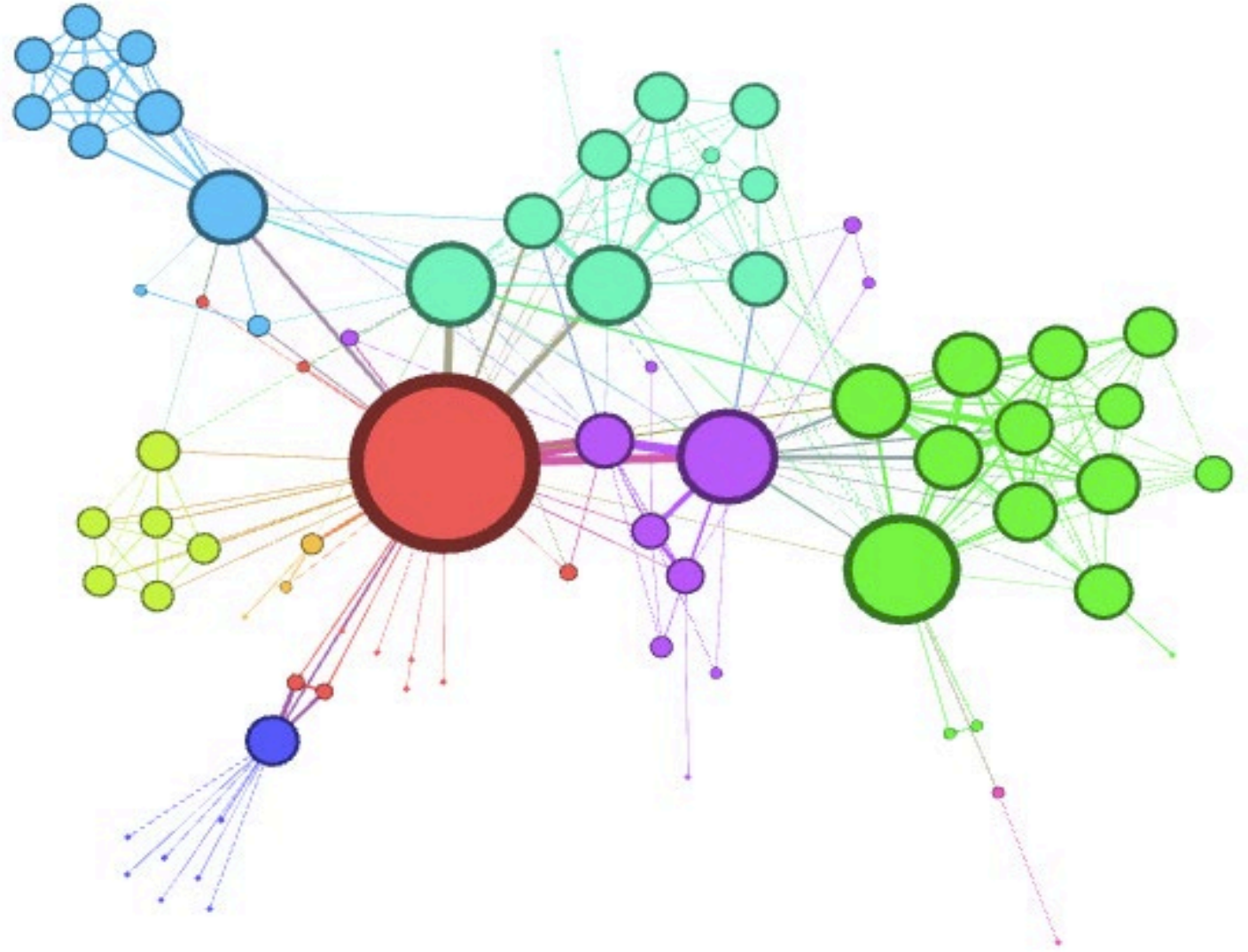
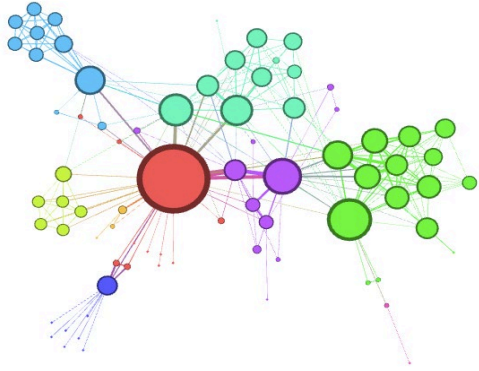


A Brief Introduction to Social Network Analysis

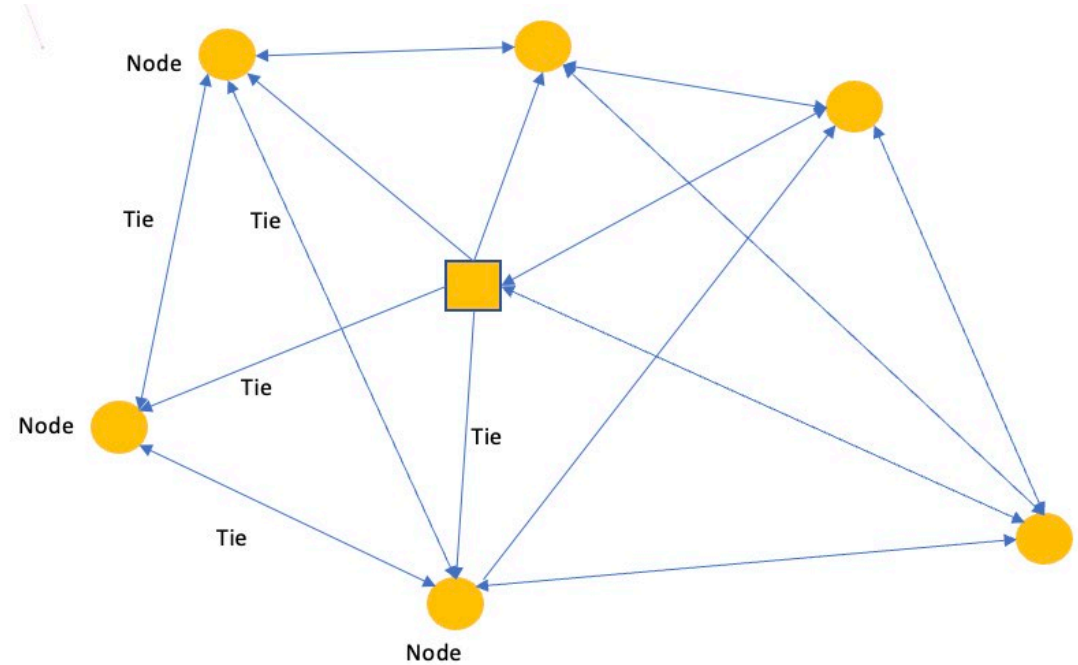
Amy Steinbugler
Data Analytics Hangout
February 8, 2023

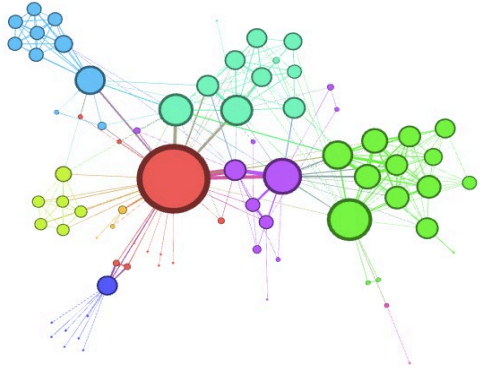




What is social network analysis?

The social world is comprised of interconnected actors (both humans and organizations). Social network analysis explores the structure, composition, and function of these webs.



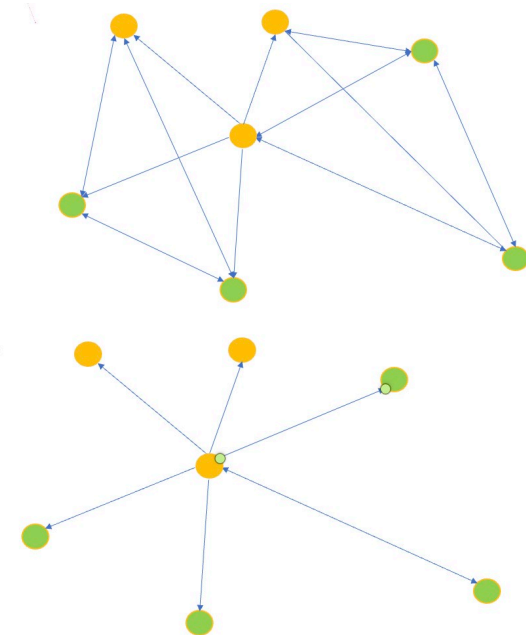
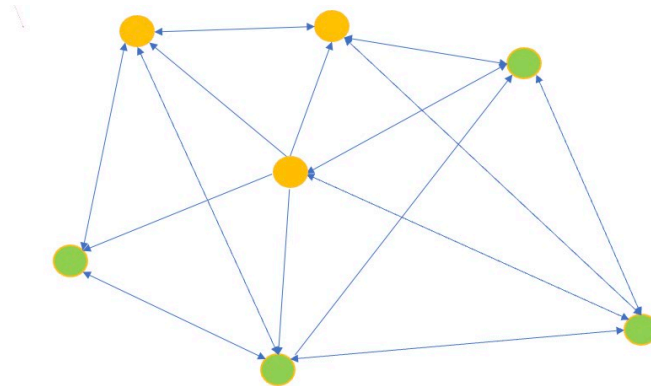


Network Metrics: Why Does Structure Matter?

The **structure** of a social network (and the position of key individuals within it) is the architecture that facilitates or constrains the flow of resources through it.

Key Structural Metrics:

- * Degree
- * Density
- * Reciprocity
- * Betweenness



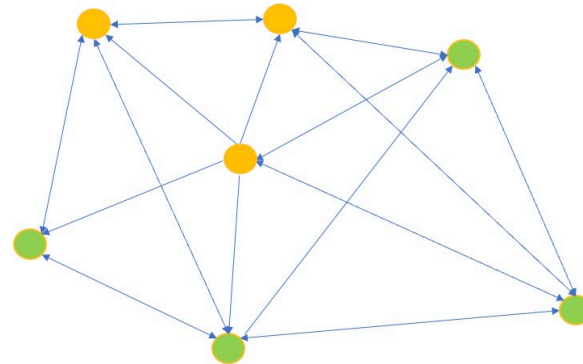


Network Metrics: Why Does Composition Matter?

The **composition** of a social network reveals the capacity for actors to access certain resources or status attributes. It also offers a way to explore patterns of homophily and heterophily.

Key Composition Metrics:

- * Homophily Index

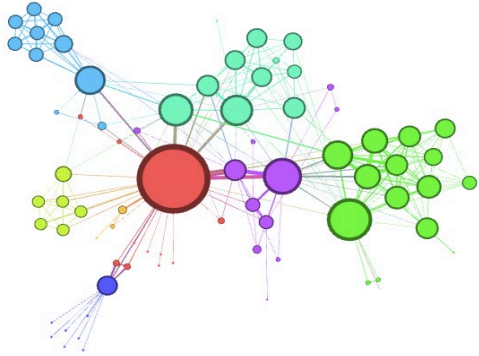




What about function?

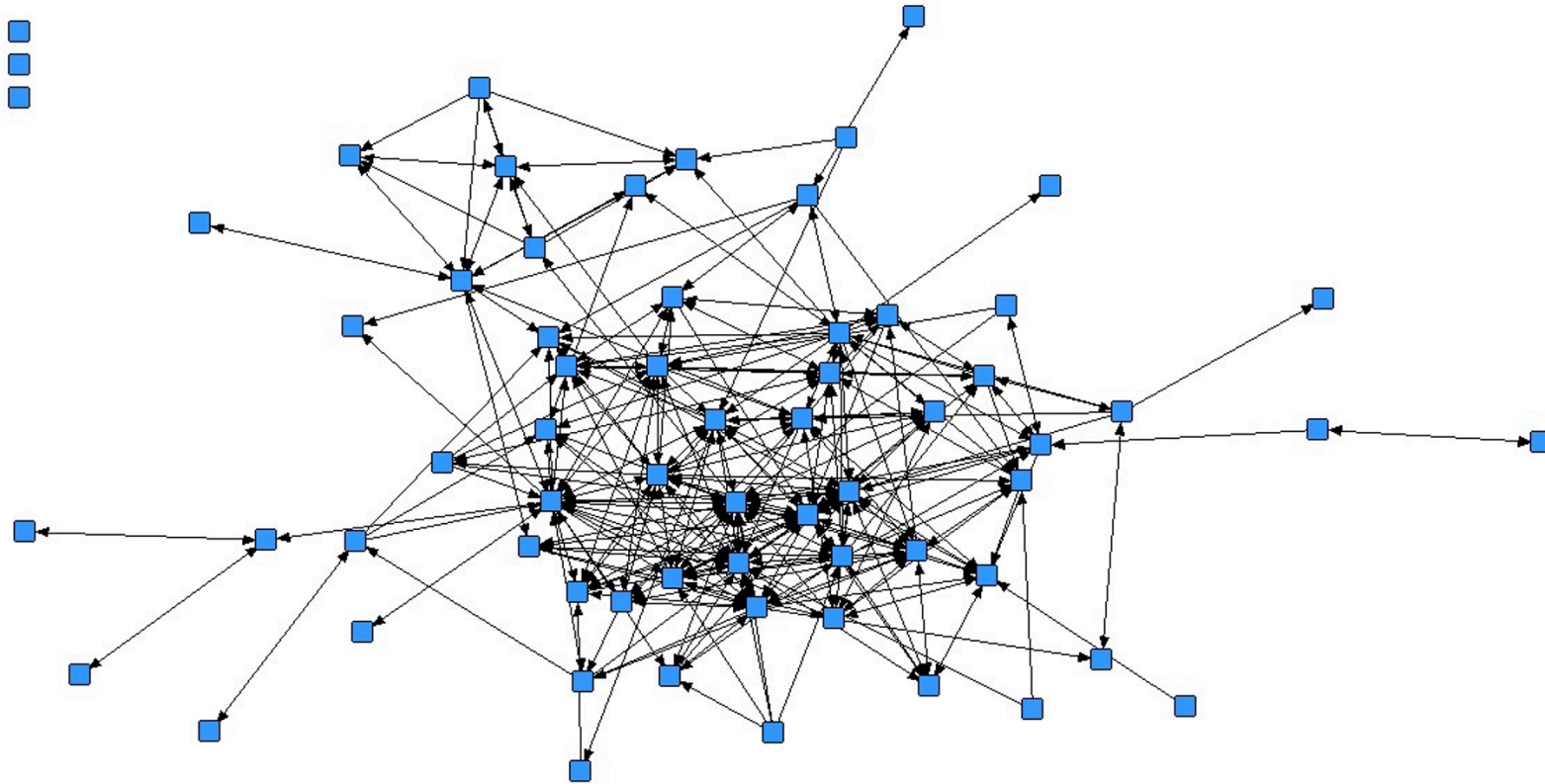
Researchers' focus on network function reflects the research question and theoretical perspective that propels the study.

In the social science, network research often explores social capital or social support.



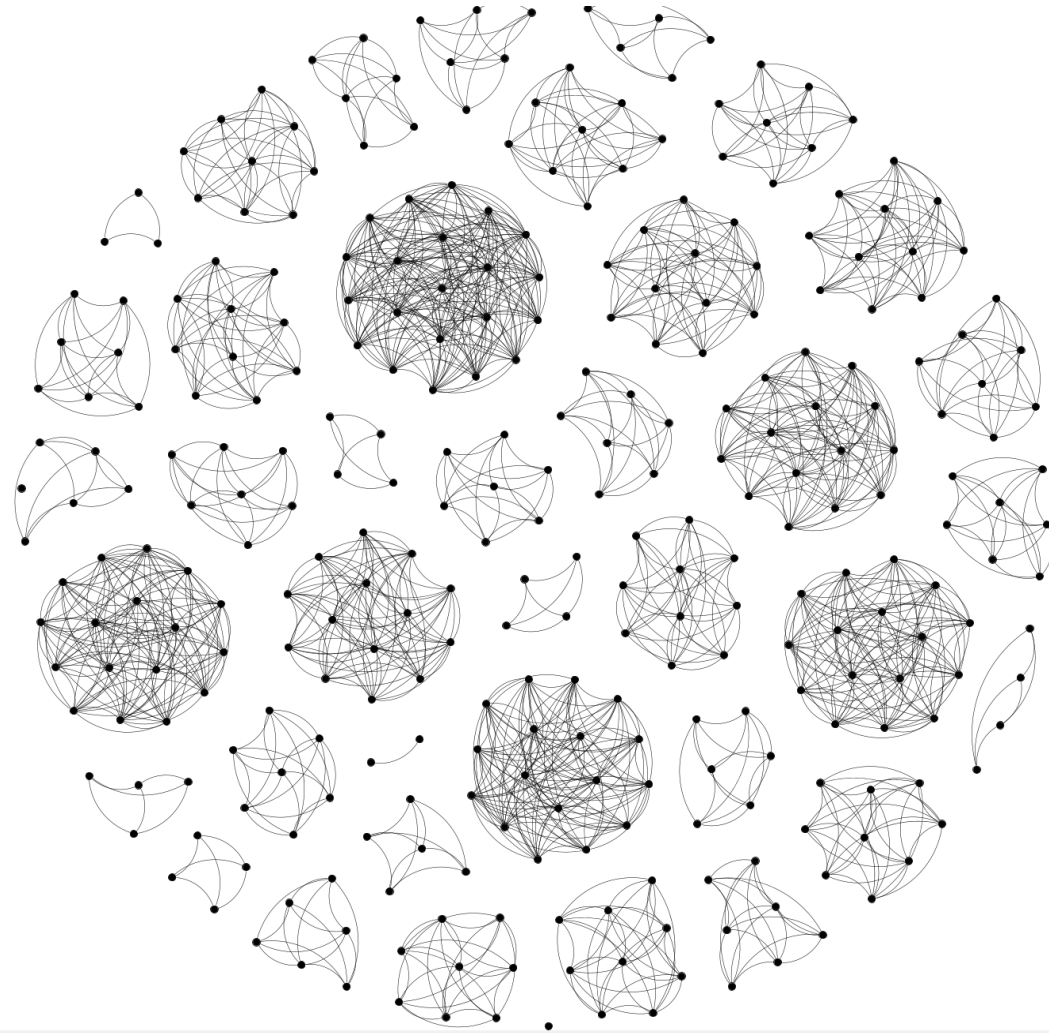
Whole Networks

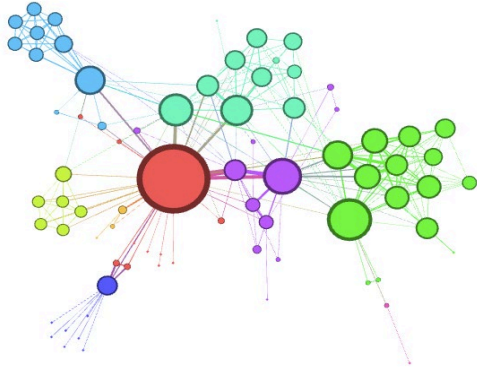
8th Grade Parent Ties—Who Knows Whom by Name





Whole Networks vs. Ego Networks

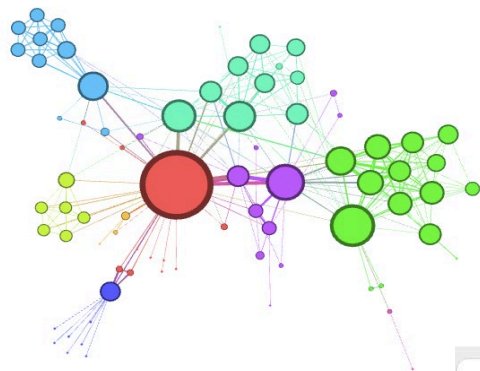




SNA Methodologies: Data Collection

Data Scraping—Social media platforms (Twitter, Facebooks, etc); Co-sponsorship of legislation; Linkages between Advisory Boards

Network Surveys—collect network data through social surveys (administered by Qualtrics or similar platform)

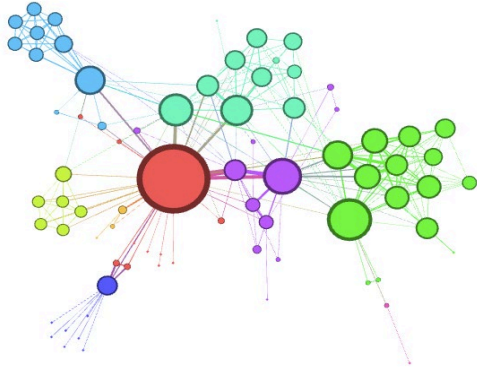


SNA Methodologies: Data Management

● ● ● Data Editor (Edit) — GSS2018.dta

Social network data is not stored in single flat file where rows represent case numbers and columns represent variables....

abany[1]		2								
	abany	abdefect	abfelegl	abhhelp1	abhhelp2	abhhelp3	abhhelp4	abhlth	abinspay	abmedgov1
1	no	yes	IAP	yes	yes	yes	yes	yes	people ...	the gove...
2	yes	yes	it depe...	no	no	no	no	yes	people ...	IAP
3	IAP	IAP	IAP	yes	no	yes	yes	IAP	people ...	a woman ...
4	IAP	IAP	should	yes	yes	yes	yes	IAP	people ...	IAP
5	no	yes	IAP	no	no	no	yes	yes	people ...	IAP
6	yes	yes	should	yes	yes	yes	yes	yes	people ...	IAP
7	yes	yes	it depe...	yes	no	yes	yes	yes	people ...	IAP
8	no	yes	IAP	yes	no	yes	yes	yes	No answ...	a woman ...
9	IAP	IAP	it depe...	yes	yes	yes	yes	IAP	people ...	a woman ...
10	IAP	IAP	IAP	yes	no	no	yes	IAP	people ...	IAP
11	yes	yes	IAP	yes	yes	yes	yes	yes	people ...	IAP
12	IAP	IAP	it depe...	yes	no	no	yes	IAP	Don't k...	IAP
13	no	yes	it depe...	no	no	no	no	yes	Don't k...	a woman ...
14	yes	yes	IAP	yes	no	yes	yes	yes	people ...	IAP
15	IAP	IAP	should	yes	yes	yes	yes	IAP	people ...	a woman ...



SNA Methodologies: Data Organization/Storage

Here, tie data is stored as an adjacency matrix, where each tie is recorded.

Here, ties can be binary or valued.

	A	B	C	D	E	F	G	H	I	J	K	L
1	neighbor	101-1	101-2	101-3	101-4	101-5	101-6	101-7	101-8	101-9	101-10	101-11
2	101-1											
3	101-2	2										
4	101-3	1	1									
5	101-4	2	2	1								
6	101-5				1							
7	101-6		1		1	2						
8	101-7	2	2		2		1					
9	101-8	2	2		2	2	1	2				
10	101-9	1			1				1			
11	101-10				3			1	1	1		
12	101-11	2	1		1			1	1	1	1	

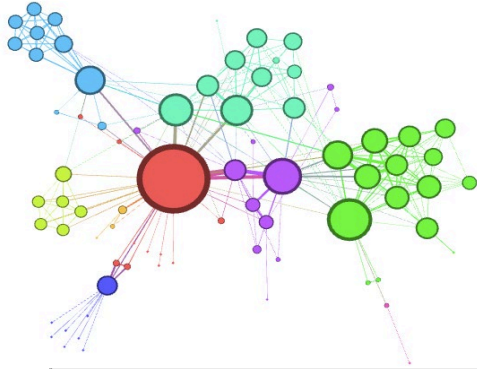
Values indicate type of relationship

blank = no relationship/don't know

1 = Acquaintance

2 = Friend

3 = Close Friend



SNA Methodologies: Data Organization/Storage

Same_worship_[1]		0										
	person1	person2	relation	Same_worsh~	Same_work_	Nkid_same_~	Also_relat~	Exchanges~n_	Exchanges~p_	Exchanges~e_	Discusses_~	Talks_abou~
1	101-1	101-2	1	0	0	1	0	1	1	1	1	0
2	101-1	101-3	0	0	0	1	0	1	1	1	1	0
3	101-1	101-4	0	0	0	1	0	1	1	1	1	0
4	101-1	101-5	0	0	0	1	0	1	1	1	1	0
5	101-1	101-6	0	0	0	1	0	1	1	1	1	0
6	101-1	101-7	0	0	0	1	0	1	1	1	1	0
7	101-1	101-8	0	0	0	1	0	1	1	1	1	0
8	101-2	101-3	2	0	0	0	0	1	1	0	1	0
9	101-2	101-4	1	0	0	0	0	1	1	0	1	0
10	101-2	101-5	2	0	0	0	0	1	1	0	1	0
11	101-2	101-6	2	0	0	0	0	1	1	0	1	0
12	101-2	101-7	1	0	0	0	0	1	1	0	1	0
13	101-2	101-8	0	0	0	0	0	1	1	0	1	0
14	101-3	101-4	0	0	0	0	0	1	1	0	0	0
15	101-3	101-5	2	0	0	0	0	1	1	0	0	0

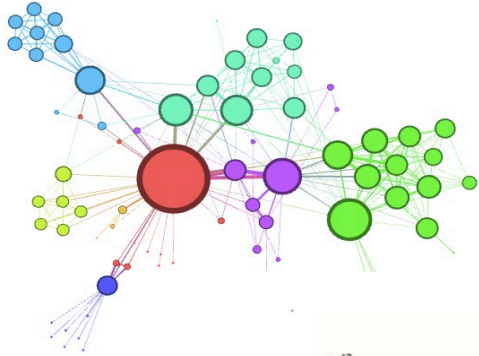
Here, tie data is stored as an edgelist, where each unit represents a tie (or dyad). Columns represent properties of the tie—attributes as well as resources exchanged.



SNA Methodologies: Data Organization/Storage

	A	B	C	D	E	F	G	H	I
1	InterviewNumber	RecipientLastName	RecipientFirstName	RecordedDate	UserLanguage	Variable 1	Variable 2	Variable 3	Variable 4
2	101	Last Name 1	First Name 1	10/18/19 12:10	EN	4	39	2	6725829
3	102	Last Name 2	First Name 2	10/23/19 6:10	EN	3	55	3	844423
4	103	Last Name 3	First Name 3	10/24/19 7:02	EN	1	91	4	492044
5	104	Last Name 4	First Name 4	11/1/19 11:46	EN	4	32	3	2196041
6	105	Last Name 5	First Name 5	11/7/19 3:01	EN	3	71	4	3221134
7	106	Last Name 6	First Name 6	11/15/19 12:20	EN	2	65	9	945734
8	109	Last Name 7	First Name 7	11/25/19 7:58	EN	3	88	2	9749584
9	110	Last Name 8	First Name 8	11/25/19 10:11	EN	4	46	9	2748932
10	111	Last Name 9	First Name 9	11/26/19 3:19	EN	3	74	9	1852410
11	112	Last Name 10	First Name 10	12/5/19 11:02	EN	1	12	4	99341
12	113	Last Name 11	First Name 11	12/16/19 8:09	EN	1	72	5	5763218
13	115	Last Name 12	First Name 12	12/17/19 7:58	EN	1	61	1	2375943
14	114	Last Name 13	First Name 13	12/17/19 8:41	EN	3	88	1	2158157
15	116	Last Name 14	First Name 14	12/18/19 7:09	EN	3	36	2	8018470

In addition to storing information about ties (the architecture of the network), additional files must record the attributes of individual egos.



SNA Methodologies: Analysis/Visualizations

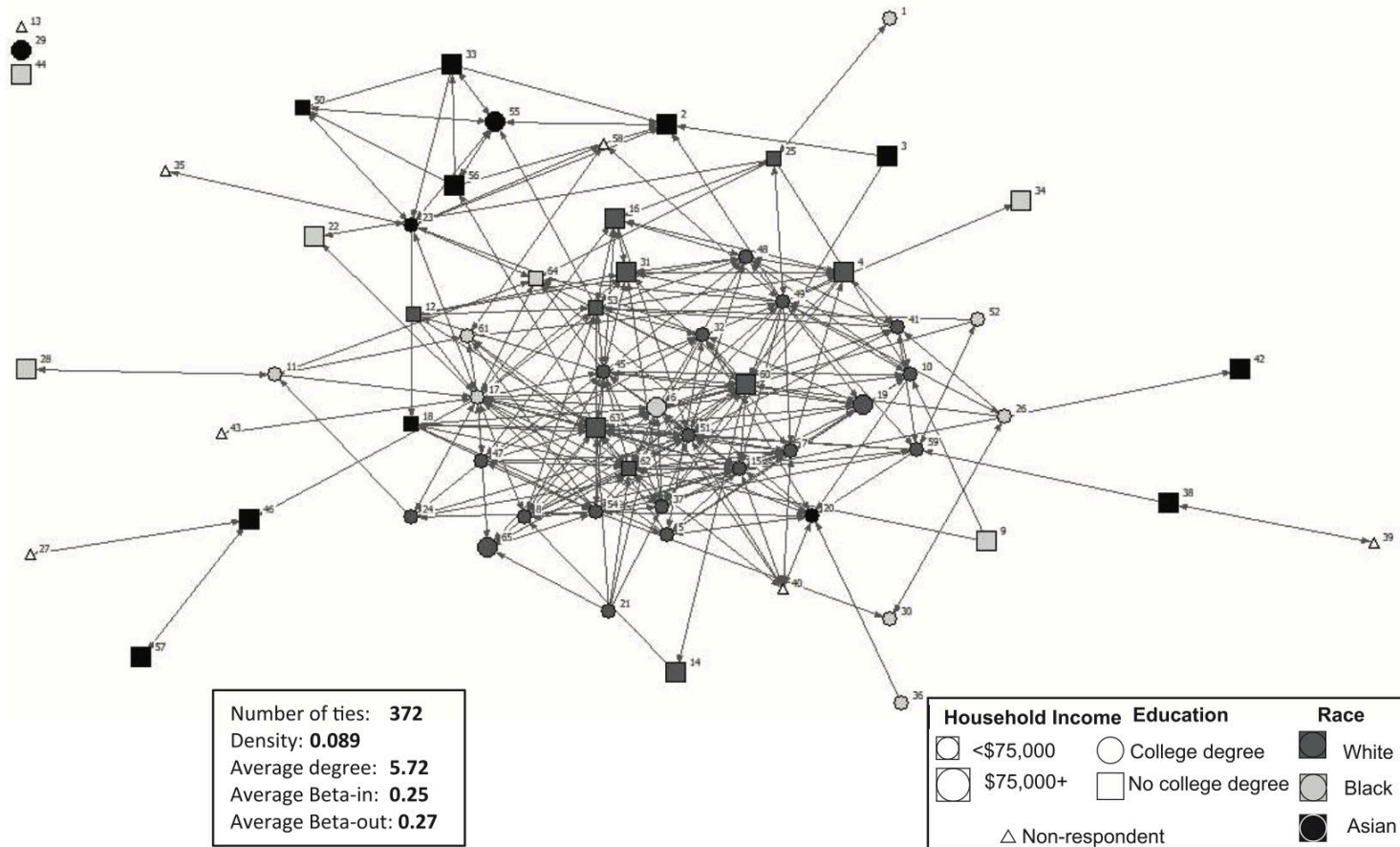


FIGURE 1. *Acquaintance network at City Magnet.*



SNA Methodologies: Analysis/Visualizations

Data analysis of network data is not limited to data visualizations (and the metrics that can be derived from these matrices).

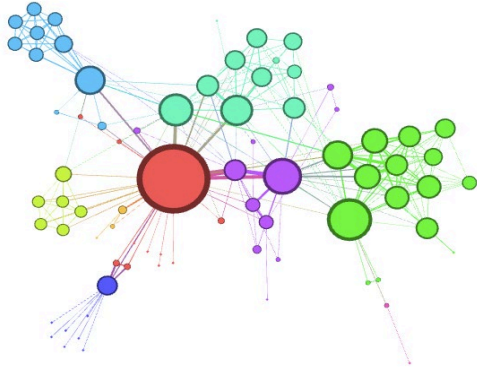
More complex models facilitate inferential analysis:

- * OLS regression (with ego-level dependent variable)
- * Logistic regression (multi-level models with tie-level data)
- * ERGM (exponential random graph model)
- * QAP (quadratic assignment procedure)



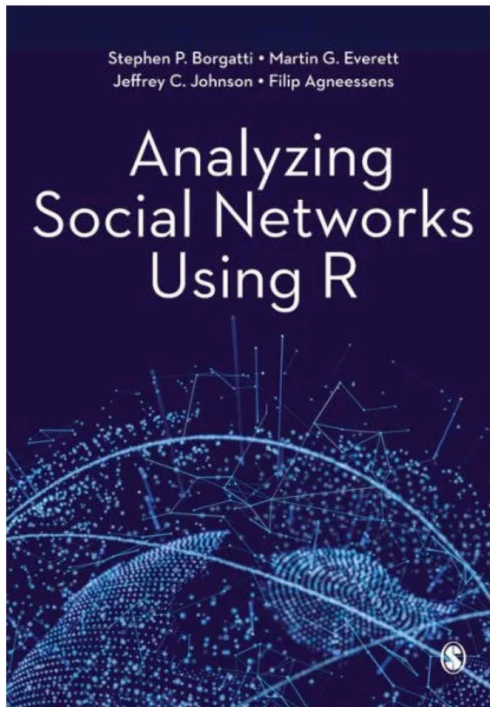
How might YOU use SNA?

	Type of Node	
Level of Analysis	Individuals	Collectives
Node Level	Are Data Analytics majors who are more central to their major's friendship network more likely to do well on course exams?	Are firms with more diverse technology partners more likely to introduce innovative products into the market?
Dyad Level	Are Data Analytics majors who take the same disciplinary (non-DA courses) more likely to develop friendships?	Are firms with similar organizational cultures more likely to form joint ventures with each other?
Network Level	Do academic majors (such as Data Analytics) with dense friendship networks grow faster than other academic majors?	When a network of firms is densely connected, does this place the network at greater risk of catastrophic failure (because of cascade effects)?

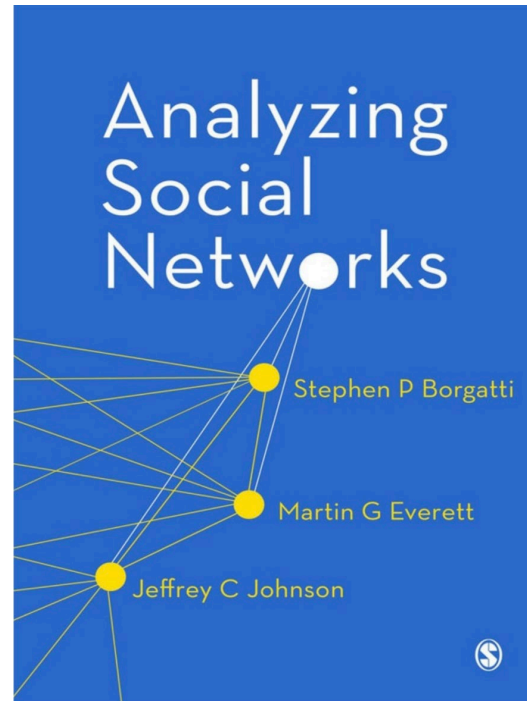


SNA Methodologies: Analysis Programs

R



UCInet/NetDraw



Pajek
Gephi
Graphstream
and many others...

Questions or Ideas?

Get in touch.

Amy Steinbugler

steinbua@dickinson.edu