LISS2108

Statistical Inference for Social Networks Analysis

Session 1 - Introduction

Santiago Quintero

santiago.quintero_suarez@kcl.ac.uk

March 6th 2025

This class

4 × 3-ish hours sessions:

1. Introduction to inference in SNA
2. Randomisation, permutation and autoregressive models
3. Intro to ERGMs
4. ERGMs selection and applications

Each class will be a mixture of:

- Lecture content (45min – 1h)
- Coding walkthrough (45min – 1h)
- Completing coding exercises (30min)

This class

SNA is a big and expanding area of research!

We won't cover (although will mention):

- Ego networks
- Community detection algorithms
- Dynamic models

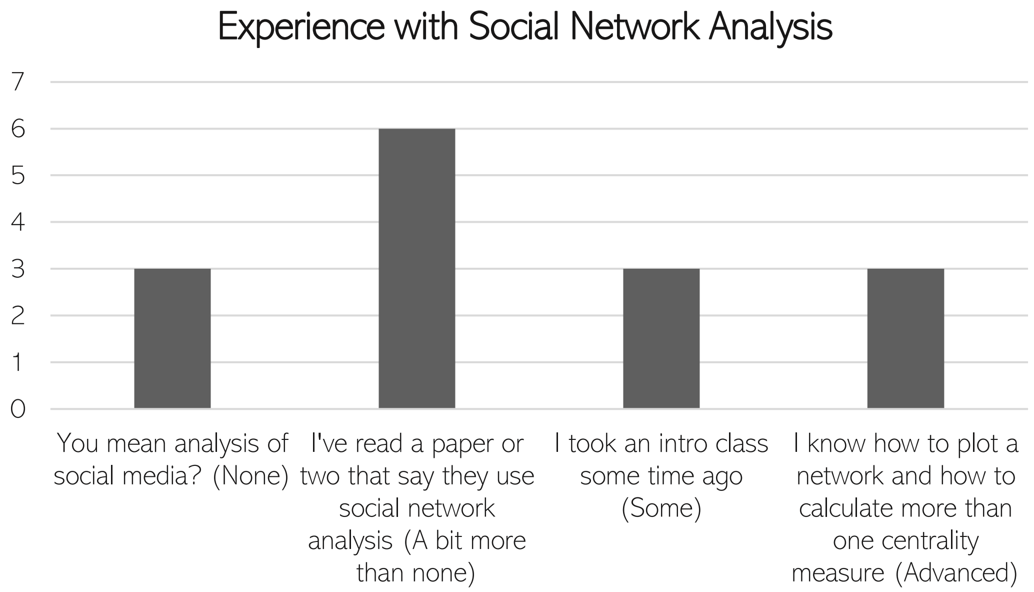
The key objectives

- Understanding the challenges and alternatives of **hypothesis testing** using **whole** network data
- Learning how to design a research question that can be tackled using inferential SNA
- Getting and intuitive understanding of the principles of some of the most popular models for conducting inference on whole networks—and how to implement them in R.

This class…

…is not an introduction to SNA!

- We will review some key concepts—but a basic idea of what networks are is very useful.



LISS-DTP has offered an [intro class to SNA before]!

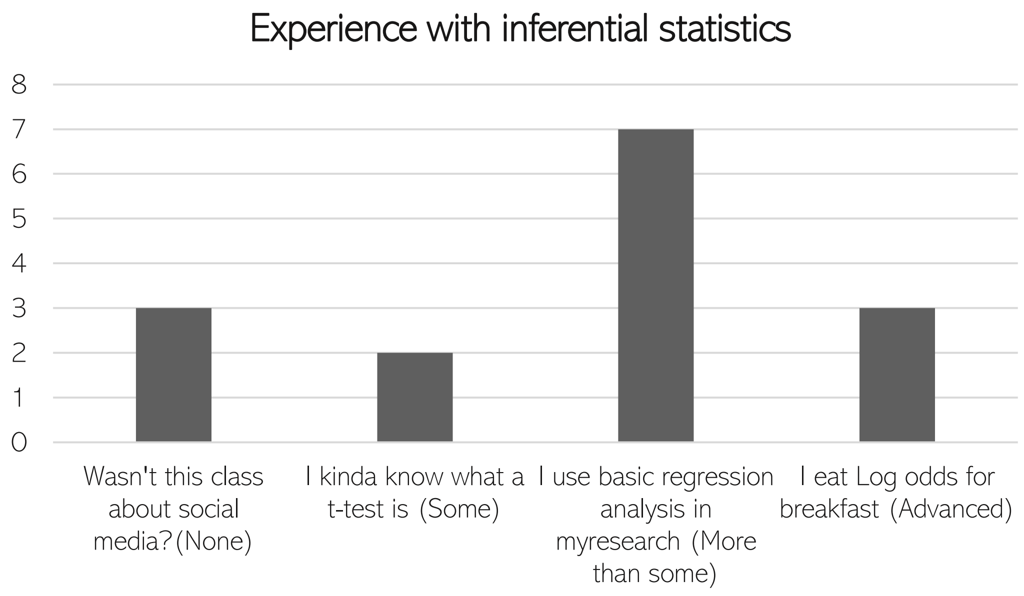
This class…

…is not an introduction to SNA!

- We will review some key concepts—but a basic idea of what networks are is very useful.

…is not an introduction to inferential stats!

- Some familiarity with statistical tests and regression is assumed, particularly with logistic regression.



There is also [statistical inference](#)with LISS-DTP!

This class…

…is not an introduction to SNA!

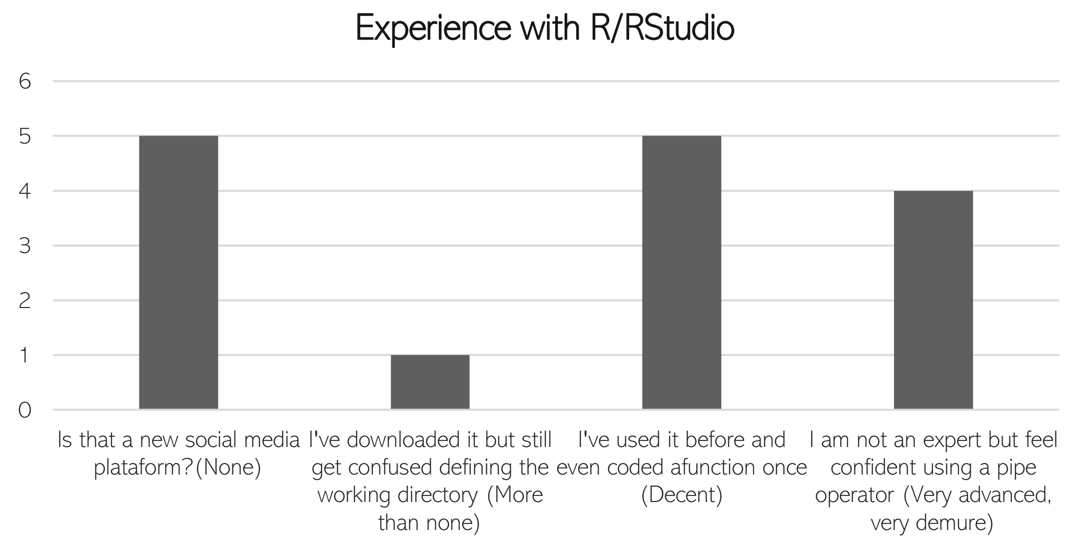
- We will review some key concepts—but a basic idea of what networks are is very useful.

…is not an introduction to inferential stats!

- Some familiarity with statistical tests and regression is assumed, particularly with logistic regression.

… is not an introduction to coding!

- We will work in R. While advanced coding skills are not necessary, familiarity with a coding environment is very useful. I will provide most of the code we will work with, though.



There are many, many, MANY resources out there to teach yourself!

This class

Some extra housekeeping:

- Classes will be on Zoom, 13:00 – 15:00, March 6th, 13th, 20th and 27th
- Classes will be recorded
- All materials, slides, readings and code will be available at the class GitHub page!
- You can reach me by email at any point ☺
- Questions?

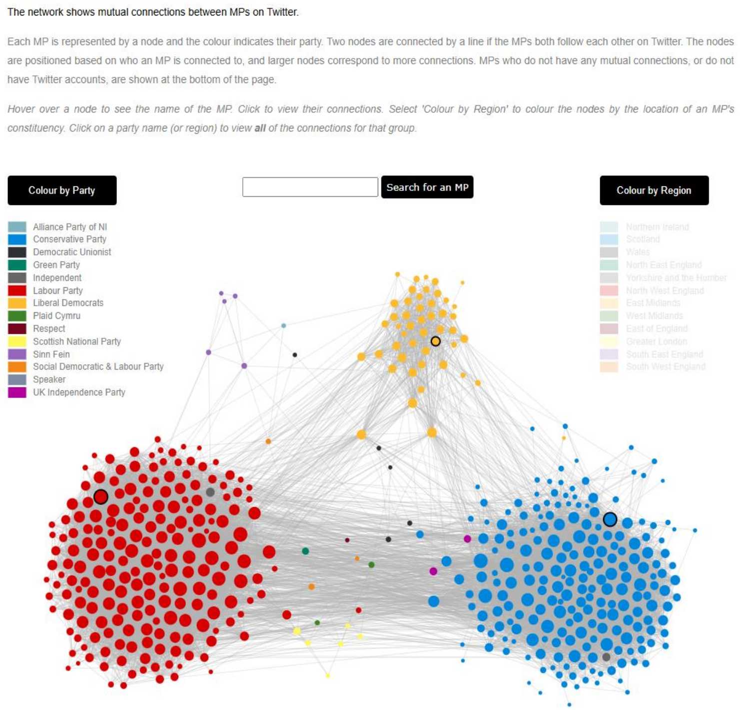
Session outline

1. SNA refresher
   - Networks as research objects
   - Representing and describing a network

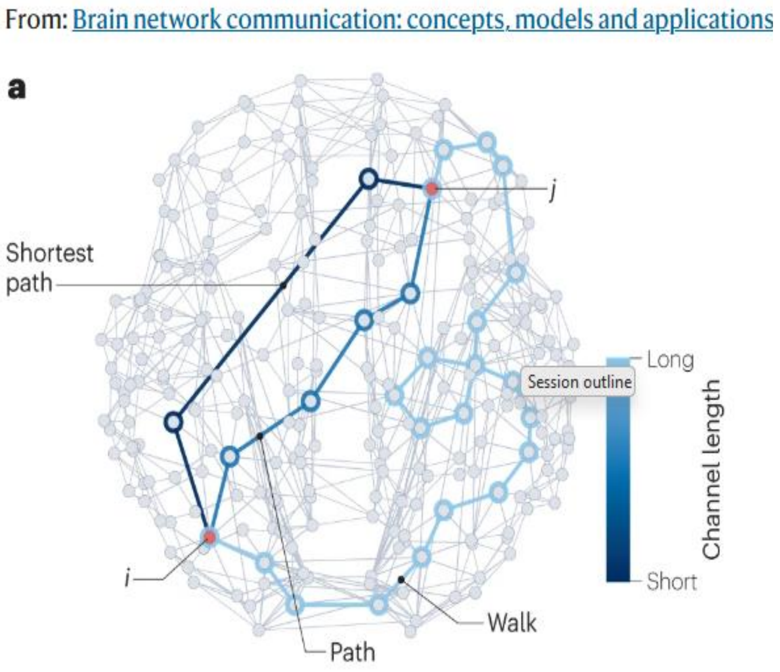
2. Research questions in SNA
   - Units and levels of analysis
   - Networks as independent and dependent variables

1. A (brief) review of Social Network Analysis

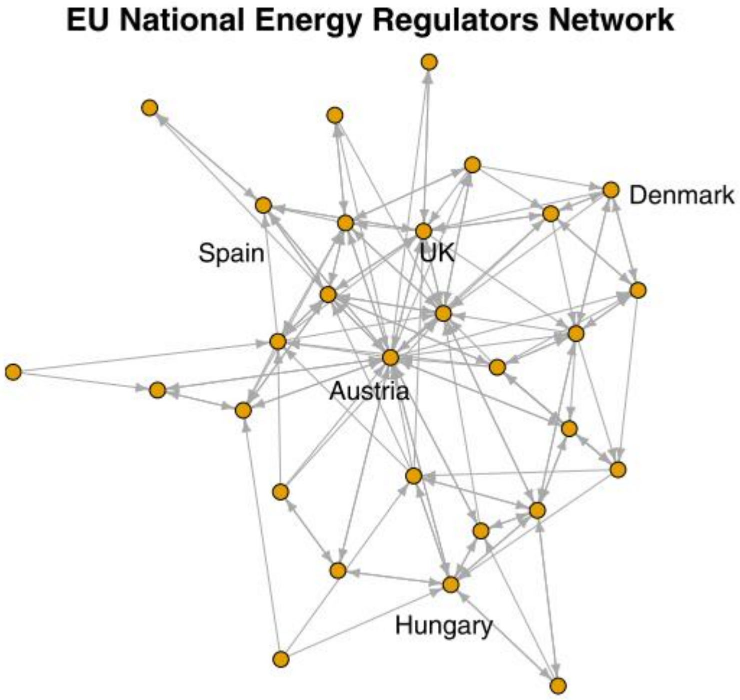
What are networks?



http://data.nesta.org.uk/policy/



Seguin, C., Sporns, O. & Zalesky, A. (2023). Brain network communication: concepts, models and applications. *Nat. Rev. Neurosci.* **24**, 557–574



Vantaggiato, F. P. (2019). The drivers of regulatory networking: policy learning between homophily and convergence. *Journal of Public Policy*, *39*(3), 443–464

What are networks?

People use the “networks” in slightly different ways, i.e.:

- Neural networks
- Policy networks
- Traffic networks
- Networks of cities

What are networks?

People use the “networks” in slightly different ways, i.e.:

- Neural networks
- Policy networks
- Traffic networks
- Networks of cities

In the context of this class, we’ll use a rather general approach: *A network is a **mathematical abstraction** of a set of relationships between entities in a (social) system* (cf. Borgatti et al., 2022).

- It implies a relational approach to understanding and conceptualising social systems
- We want to observe and analyse how entities (e.g., individuals, organisations, words–what we will call ***nodes***) interact, the reasons and results of their interactions, the structure of those interactions and their implications.

What are networks?

People use the "networks" in slightly different ways, i.e.:

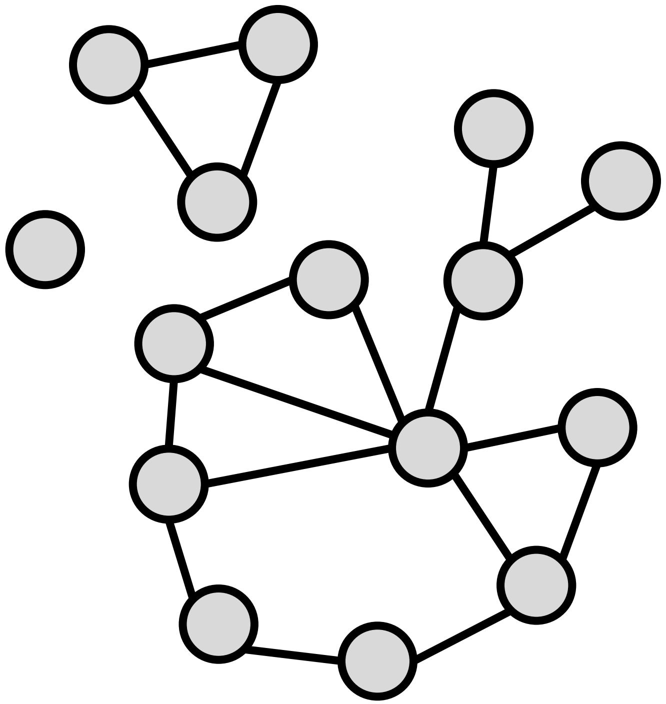
- Neural networks
- Policy networks
- Traffic networks
- Networks of cities

In the context of this class, we'll use a rather general approach: *A network is a* ***mathematical abstraction*** *of a set of relationships between entities in a (social) system* (cf. Borgatti et al., 2022).

- It implies a relational approach to understanding and conceptualising social systems
- We want to observe and analyse how entities (e.g., individuals, organisations, words–what we will call ***nodes***) interact, the reasons and results of their interactions, the structure of those interactions and their implications.

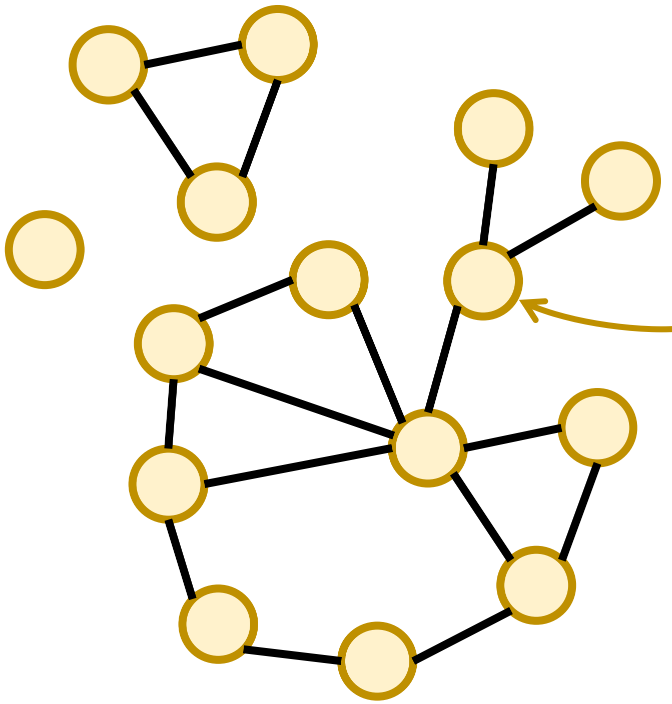
We try to connect observed patterns of interactions to theories and models of individual or social behaviour.

Representing a network



(Agneessens, 2023)

Representing a network



Nodes (vertices)

Students / Academics

Politicians

Countries

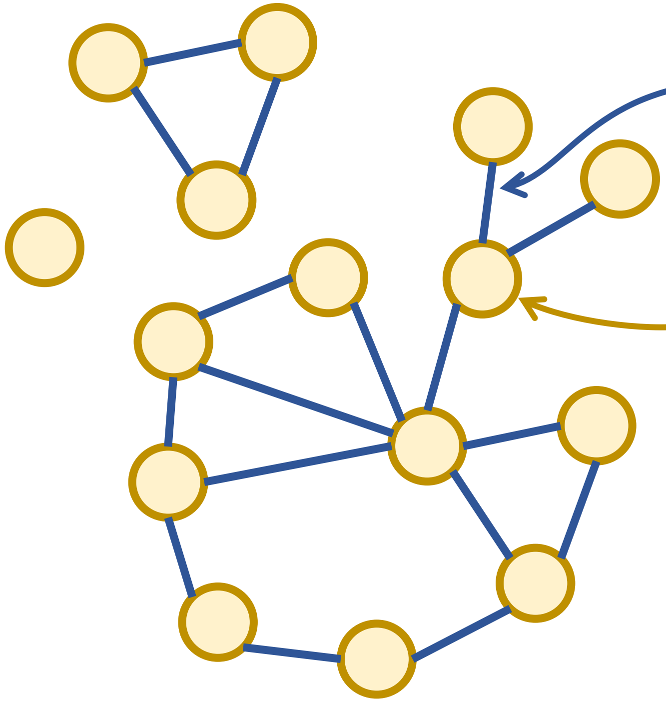
Businesses

Words

Organisations

(Agneessens, 2023)

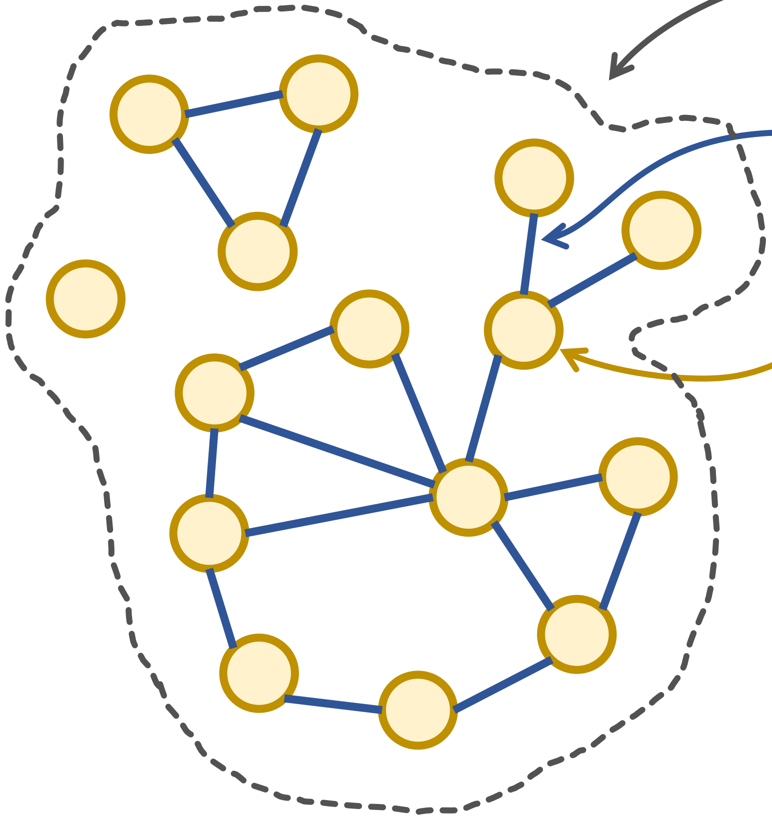
Representing a network



Nodes (vertices)	Ties (edges)
Students / Academics	Advice / co-authorship
Politicians	Partisanship
Countries	Alliances
Businesses	Trade
Words	Used together
Organisations	Employee exchange

(Agneessens, 2023)

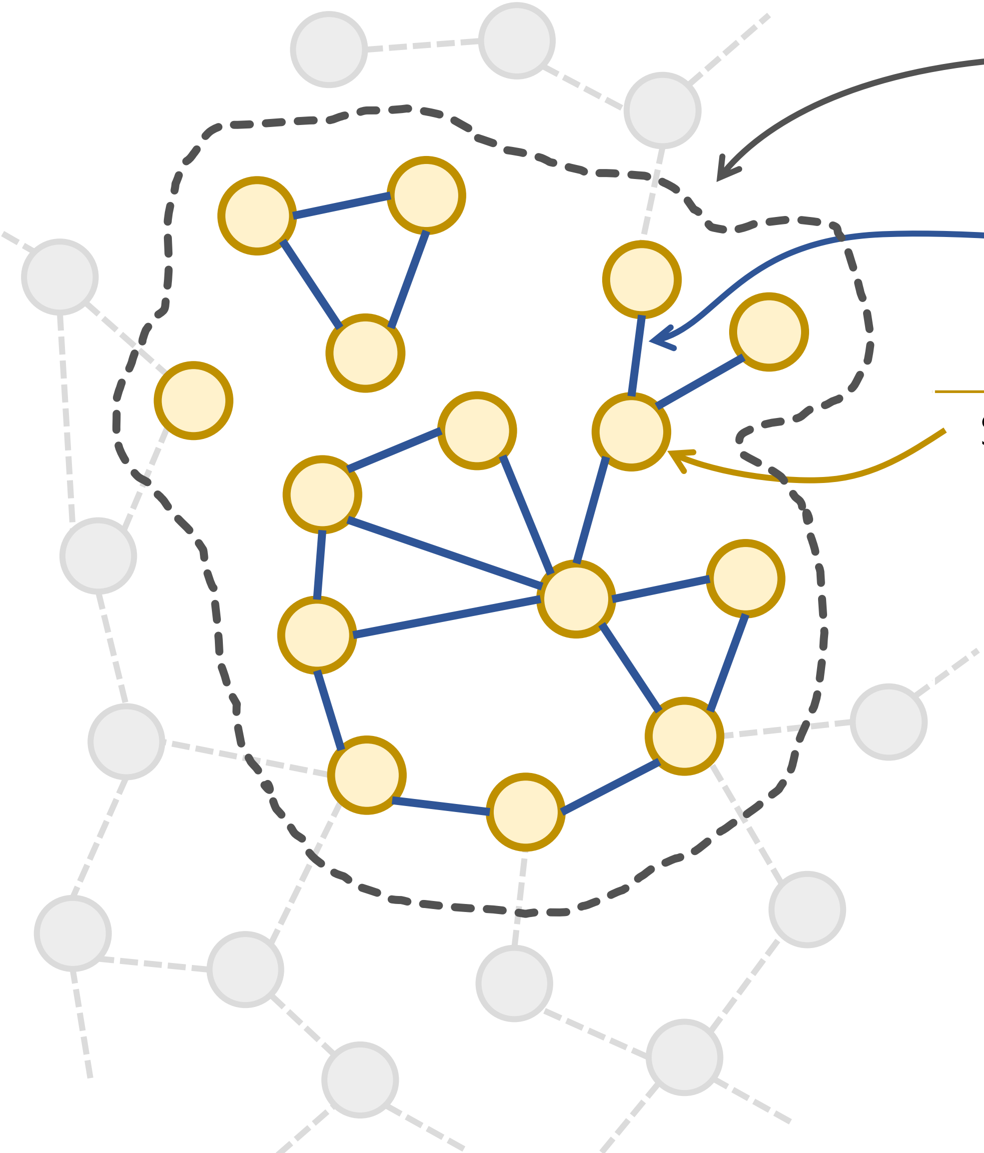
Representing a network



Nodes (vertices)	Ties (edges)	Boundary
Students / Academics	Advice / co-authorship	Classroom / Discipline
Politicians	Partisanship	Parliament
Countries	Alliances	A region / UN
Businesses	Trade	An industrial sector
Words	Used together	All tweets on a topic
Organisations	Employee exchange	A country

(Agneessens, 2023)

Representing a network



Nodes (vertices)	Ties (edges)	Boundary
Students / Academics	Advice / co-authorship	Classroom / Discipline
Politicians	Partisanship	Parliament
Countries	Alliances	A region / UN
Businesses	Trade	An industrial sector
Words	Used together	All tweets on a topic
Organisations	Employee exchange	A country

(Agneessens, 2023)

Representing a network

Graphical representations of networks are ~~cute~~ sometimes useful:

- To get a general idea of the data we are working with.
- In small samples, sometimes we can get interesting intuitions of what might be “happening”.

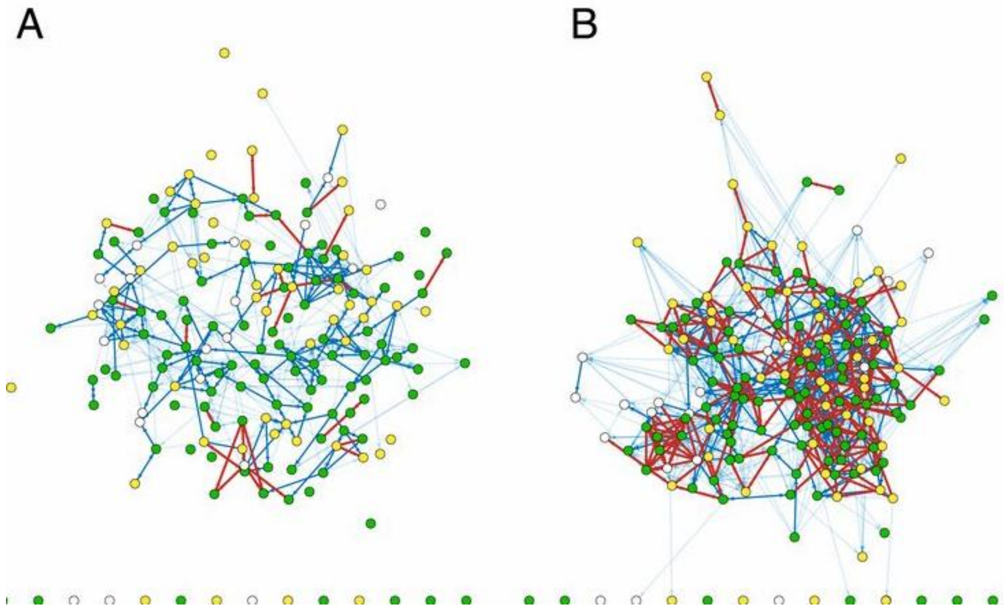
Representing a network

Graphical representations of networks are ~~cute~~ sometimes useful:

- To get a general idea of the data we are working with.
- In small samples, sometimes we can get interesting intuitions of what might be "happening".

However, in large datasets, mere graphical representations are not that useful.

Emergence of social networks between students



Network of environmental management collaborations in Colombia

C. Stadtfeld, A. Vörös, T. Elmer, Z. Boda, & I.J. Raabe. (2019). Integration in emerging social networks explains academic failure and success, *Proc. Natl. Acad. Sci*, 116 (3) 792-797

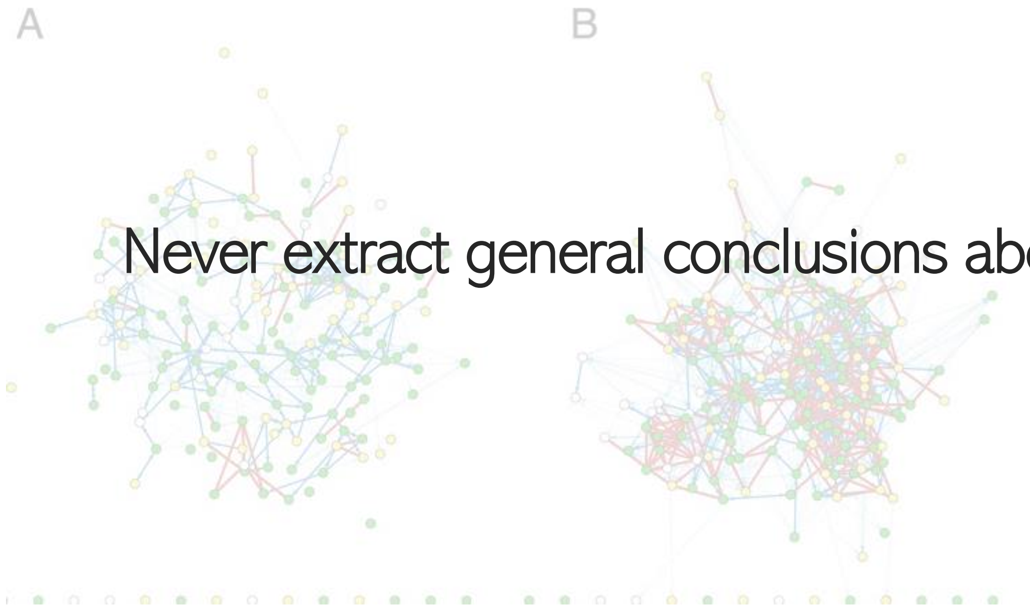
Representing a network

Graphical representations of networks are ~~cute~~ sometimes useful:

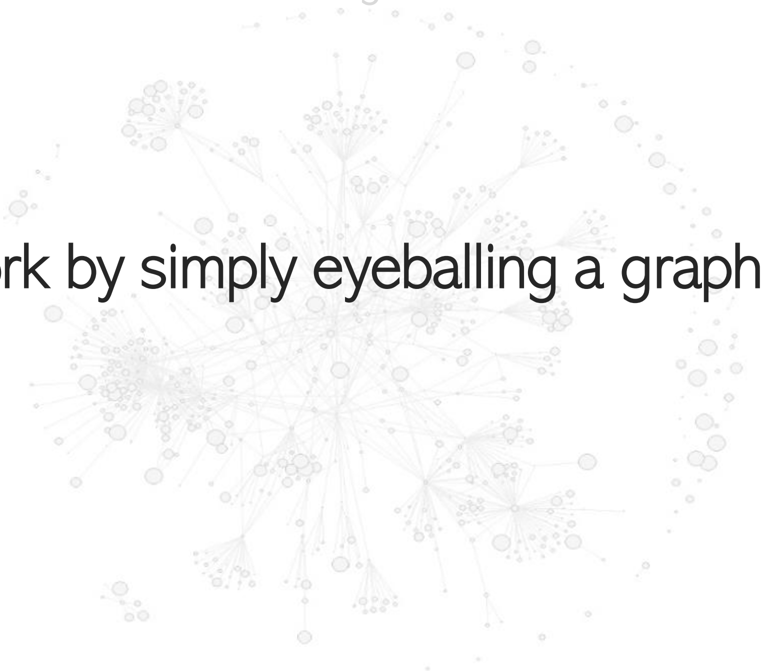
- to get a general idea of the data we are working with.
- In small samples, sometimes we can get interesting intuitions of what might be “happening”.

However, in large datasets, mere graphical representations are not that useful.

Emergence of social networks between students



Network of environmental management collaborations in Colombia



Never extract general conclusions about a network by simply eyeballing a graph!

C. Stadtfeld, A. Vörös, T. Elmer, Z. Boda, & I.J. Raabe. (2019). Integration in emerging social networks explains academic failure and success, *Proc. Natl. Acad. Sci*, 116 (3) 792-797

Representing a network

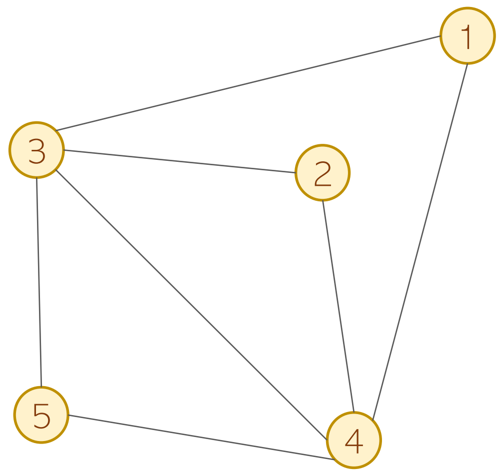
We are better off by getting used to thinking of networks as (adjacency) matrices!

- Total number of vertices: n
- Total number of edges: m
- Generic vertices: i and j
- Generic edge: (i, j)
- Adjacency matrix: $\mathbf{A}$ of size n by n
- Particular edge in matrix: A_{ij}

Representing a network

We are better off by getting used to thinking of networks as (adjacency) matrices!

- Total number of vertices: n
- Total number of edges: m
- Generic vertices: i and j
- Generic edge: (i, j)
- Adjacency matrix: $\mathbf{A}$ of size n by n
- Particular edge in matrix: A_{ij}

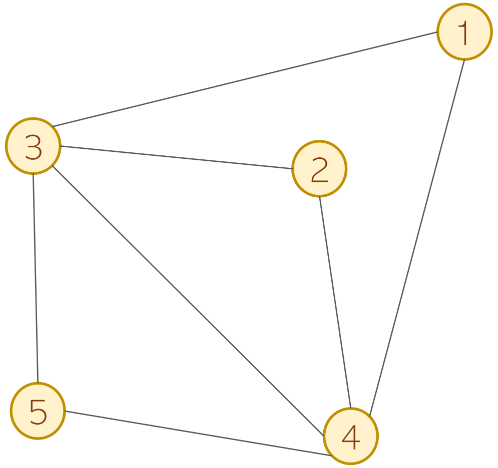


$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{bmatrix}$$

Representing a network

We are better off by getting used to thinking of networks as (adjacency) matrices!

- Total number of vertices: n
- Total number of edges: m
- Generic vertices: i and j
- Generic edge: (i, j)
- Adjacency matrix: $\mathbf{A}$ of size n by n
- Particular edge in matrix: A_{ij}

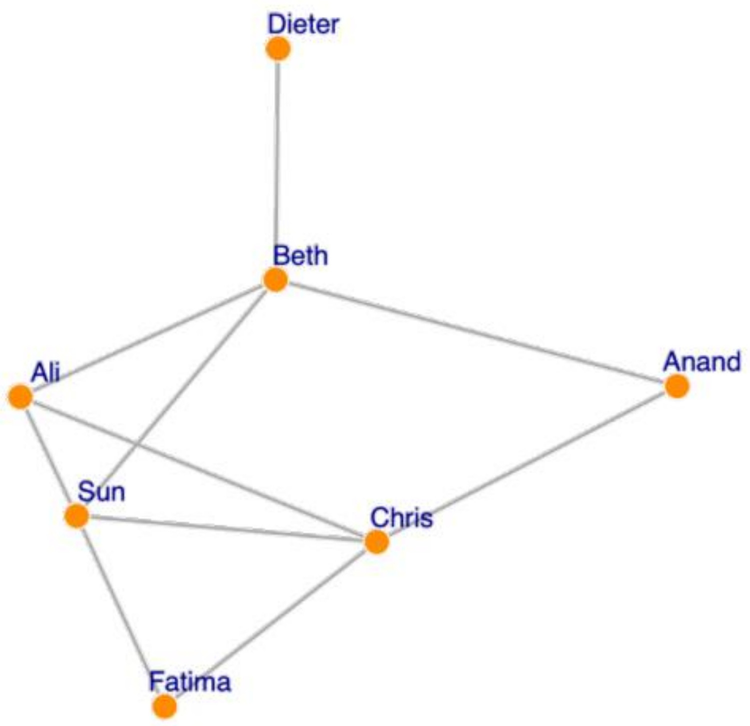


$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Representing a network

We are better off by getting used to thinking of networks as (adjacency) matrices!

	Anand	Beth	Chris	Dieter	Sun	Ali	Fatima
Anand	0	1	1	0	0	0	0
Beth	1	0	0	1	1	1	0
Chris	1	0	0	0	1	1	1
Dieter	0	1	0	0	0	0	0
Sun	0	1	1	0	0	1	1
Ali	0	1	1	0	1	0	0
Fatima	0	0	1	0	1	0	0

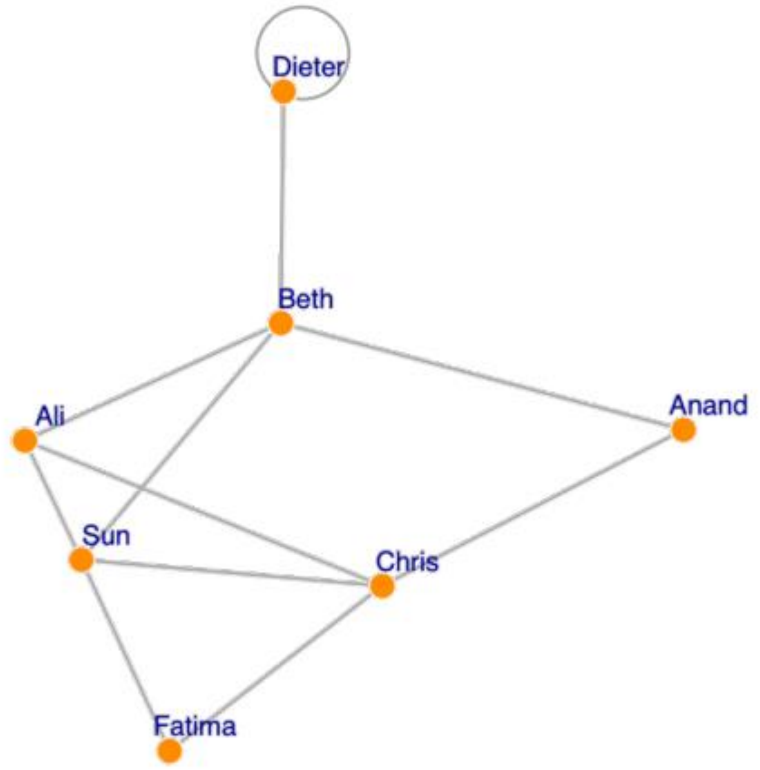


(Power, 2021)

Representing a network

We are better off by getting used to thinking of networks as (adjacency) matrices!

	Anand	Beth	Chris	Dieter	Sun	Ali	Fatima
Anand	0	1	1	0	0	0	0
Beth	1	0	0	1	1	1	0
Chris	1	0	0	0	1	1	1
Dieter	0	1	0	2	0	0	0
Sun	0	1	1	0	0	1	1
Ali	0	1	1	0	1	0	0
Fatima	0	0	1	0	1	0	0

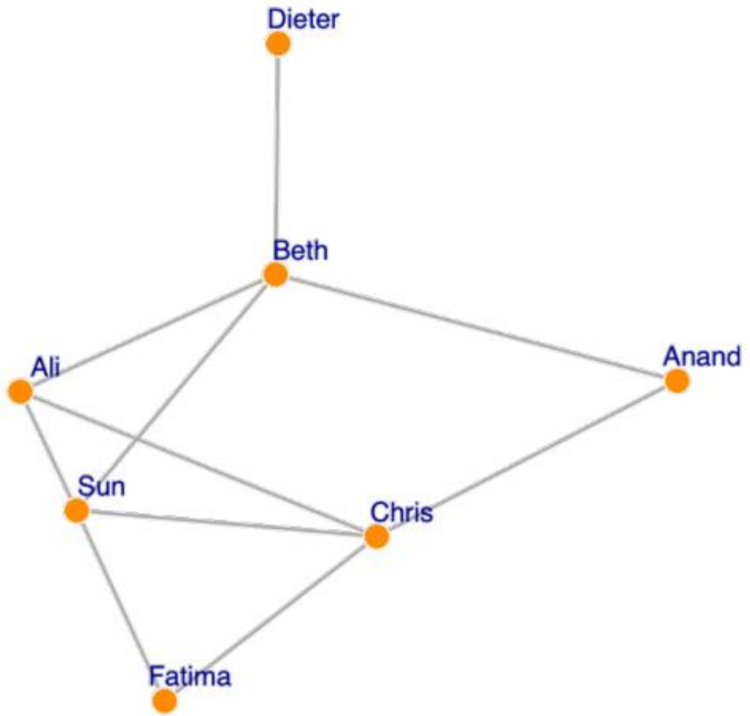


(Power, 2021)

Representing a *undirected* network

We are better off by getting used to thinking of networks as (adjacency) matrices!

	Anand	Beth	Chris	Dieter	Sun	Ali	Fatima
Anand	0	1	1	0	0	0	0
Beth	1	0	0	1	1	1	0
Chris	1	0	0	0	1	1	1
Dieter	0	1	0	0	0	0	0
Sun	0	1	1	0	0	1	1
Ali	0	1	1	0	1	0	0
Fatima	0	0	1	0	1	0	0

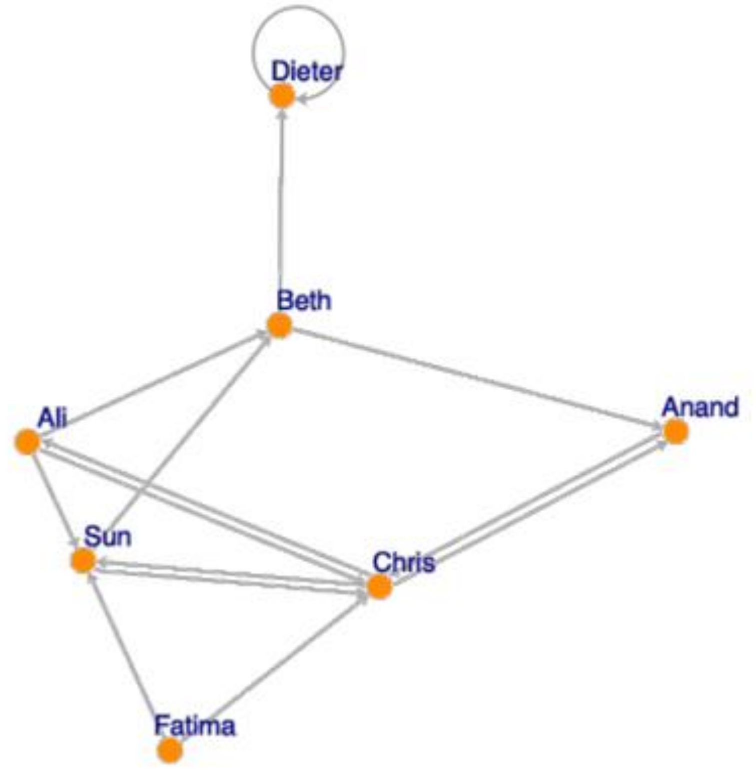


(Power, 2021)

Representing a *directed* network

We are better off by getting used to thinking of networks as (adjacency) matrices!

	Anand	Beth	Chris	Dieter	Sun	Ali	Fatima
Anand	0	0	1	0	0	0	0
Beth	1	0	0	1	0	0	0
Chris	1	0	0	0	1	1	0
Dieter	0	0	0	1	0	0	0
Sun	0	1	1	0	0	0	0
Ali	0	1	1	0	1	0	0
Fatima	0	0	1	0	1	0	0

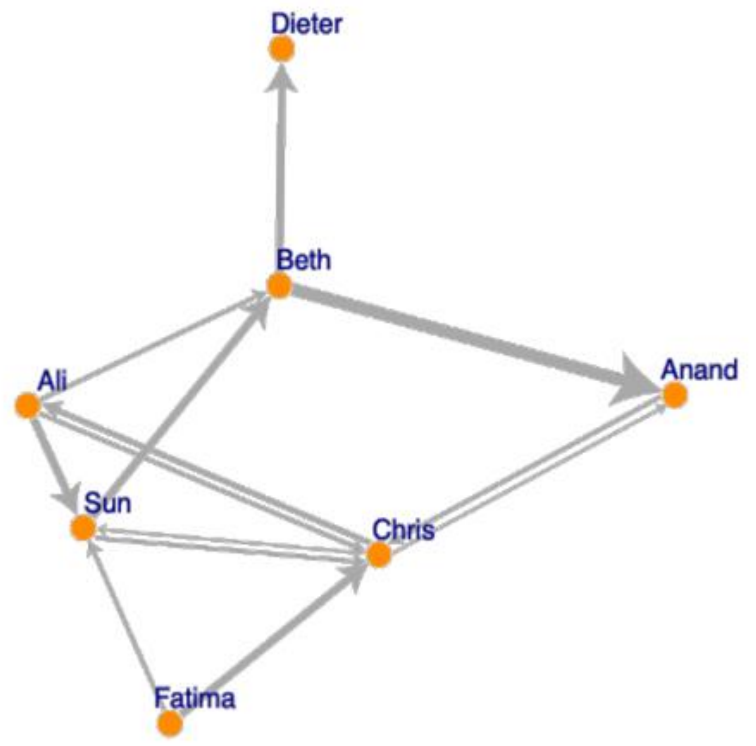


(Power, 2021)

Representing a network – Edge attributes (weighted edges)

We are better off by getting used to thinking of networks as (adjacency) matrices!

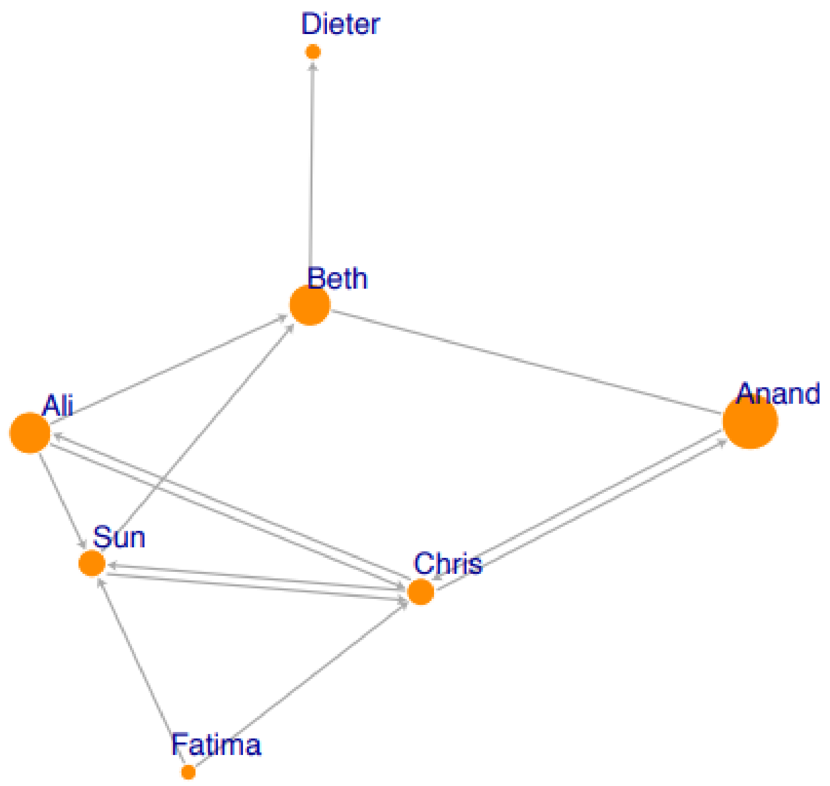
	Anand	Beth	Chris	Dieter	Sun	Ali	Fatima
Anand	0	0	2	0	0	0	0
Beth	4	0	0	3	0	0	0
Chris	1	0	0	0	1	2	0
Dieter	0	0	0	0	0	0	0
Sun	0	3	2	0	0	0	0
Ali	0	2	2	0	3	0	0
Fatima	0	0	3	0	2	0	0



(Power, 2021)

Representing a network – Vertex attributes

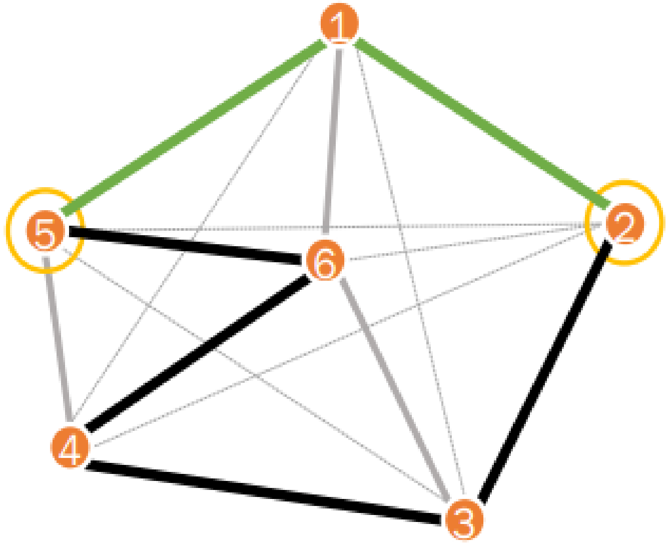
Name	Gender	Age
Anand	1	30
Beth	0	25
Chris	1	22
Sun	0	22
Ali	1	24
Fatima	0	18
Dieter	1	18



(Power, 2021)

Describing a network

- Total number of vertices: n — 6
- Total number of edges: m — 9
- Degree : $k_i = \sum_{j=1}^{n} A_{ij}$ * — $k_1=3$, $k_2=2$, $k_3=3$
- Mean degree: $c = \frac{2m}{n}$ * — 3
- Density: $\rho = \frac{2m}{n(n-1)} = \frac{c}{n-1} \approx \frac{c}{n}$ * — 0,6
- Diameter: longest geodesic path length — 2

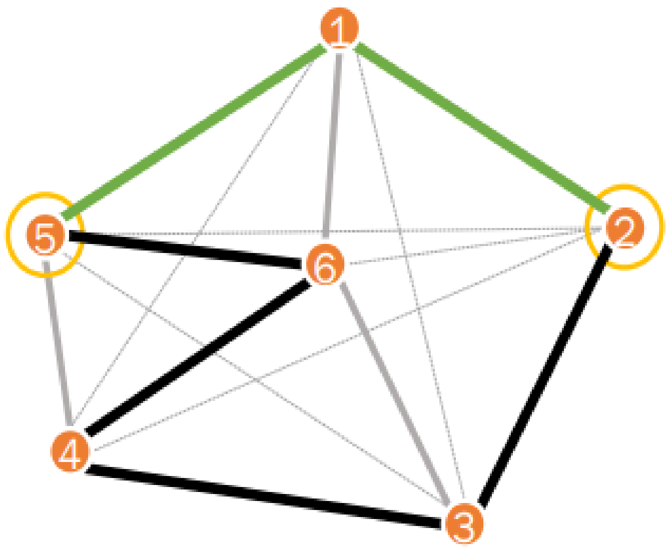


*Different in directed networks

(Power, 2021)

Describing a network

- Total number of vertices: n `vcount()`
- Total number of edges: m `ecount()`
- Degree : $k_i = \sum_{j=1}^{n} A_{ij}$ * `degree()`
- Mean degree: $c = \frac{2m}{n}$ * `mean(degree())`
- Density: $\rho = \frac{2m}{n(n-1)} = \frac{c}{n-1} \approx \frac{c}{n}$ * `graph.density()`
- Diameter: longest geodesic path length `diameter()`



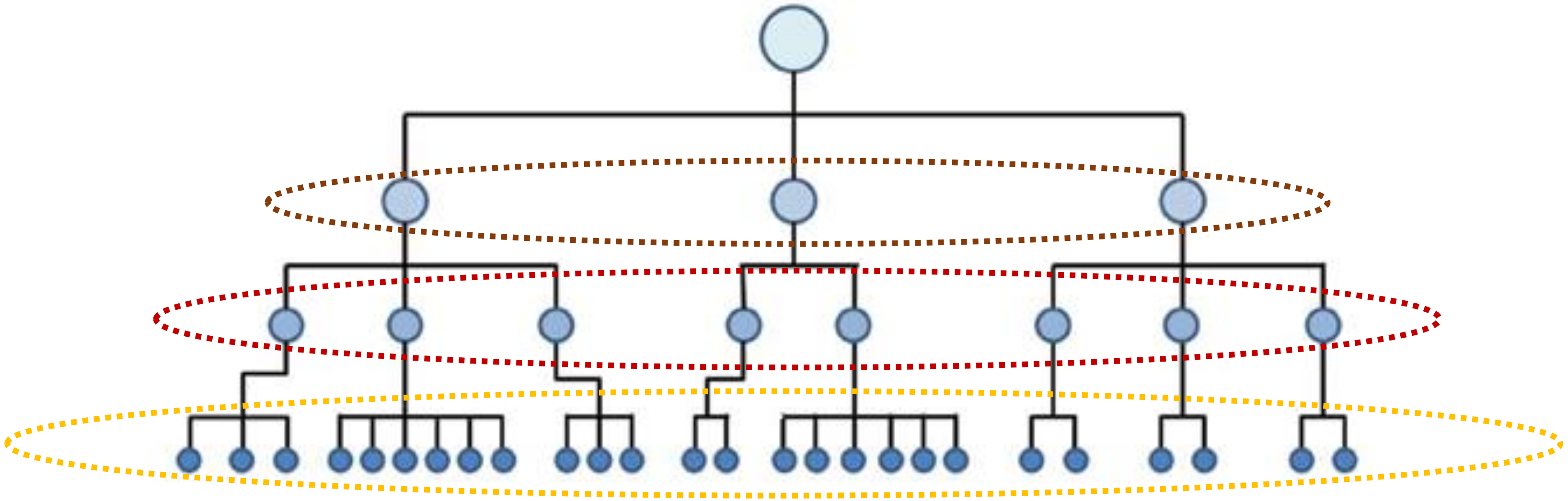
*Different in directed networks

(Power, 2021)

Describing a network

There are many ways to descriptively analyse a network

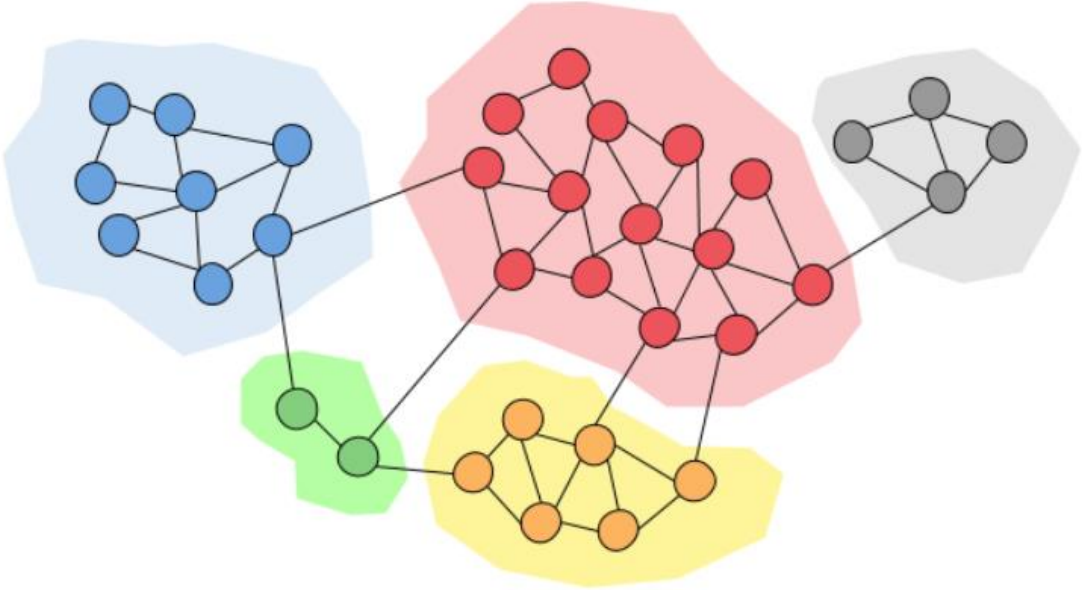
- Centrality measures (most important nodes?)
- Structural equivalence (are two nodes “structurally equivalent”?)



Describing a network

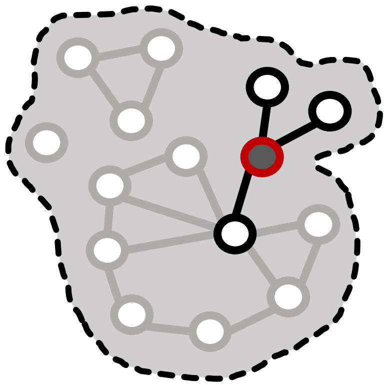
There are many ways to descriptively analyse a network

- Centrality measures (most important nodes?)
- Structural equivalence (are two nodes “structurally equivalent”?)
- Community detection (are there meaningful sub-partitions of the network?)



2. Research questions in Social Network Analysis

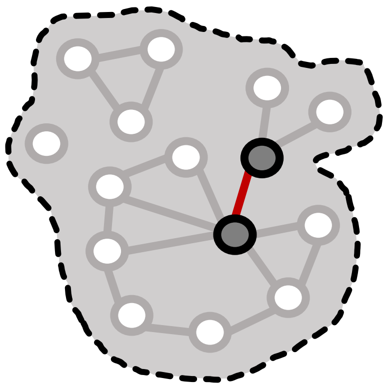
Levels of analysis



Node level
Degree centrality
Level at which a node is similar to its alters (direct connections)
How much a node is a structural hole (broker)
How do node characteristics determine its position in the network

(Agneessens, 2020; 2023)

Levels of analysis



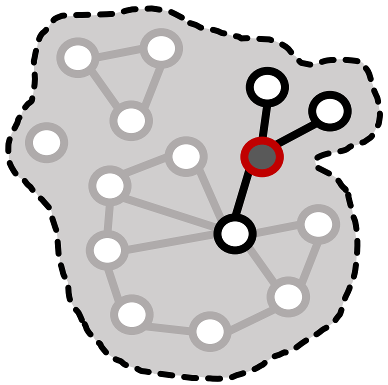
Dyad level

Presence or absence of a tie between two nodes

Distance between two nodes

Number of closed triads

Whether 2 connected nodes are similar on some characteristic



Node level

Degree centrality

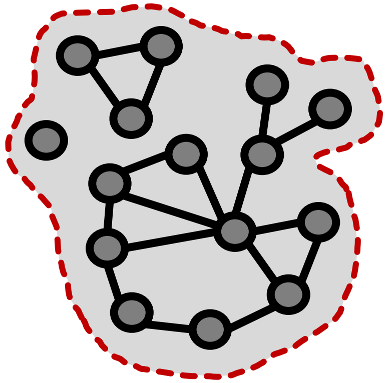
Level at which a node is similar to its alters (direct connections)

How much a node is a structural hole (broker)

How do node characteristics determine its position in the network

(Agneessens, 2020; 2023)

Levels of analysis



Group level

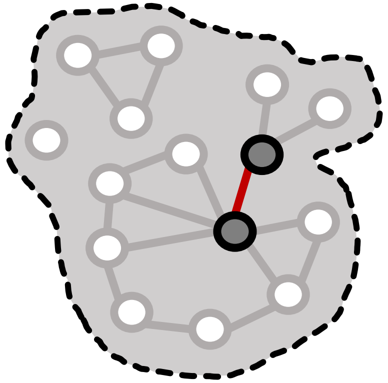
Density

Number of cliques

Level of “core-peripheriness”

Level of homophily in a group

Dyad level



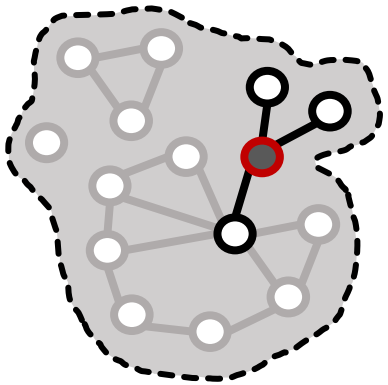
Presence or absence of a tie between two nodes

Distance between two nodes

Number of closed triads

Whether 2 connected nodes are similar on some characteristic

Node level



Degree centrality

Level at which a node is similar to its alters (direct connections)

How much a node is a structural hole (broker)

How do node characteristics determine its position in the network

(Agneessens, 2020; 2023)

Modelling networks

Antecedents of networks
(Networks as *Dependent variable*)

<u>Group characteristics</u>

- Group size
- Composition
- Formal structure

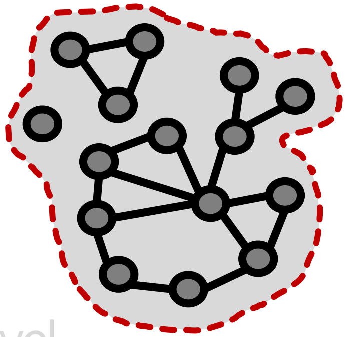
Group level

Density

Number of cliques

Level of "core-peripheriness"

Level of homophily in a group



Dyad level

Presence or absence of a tie between two nodes

Distance between two nodes

Number of closed triads

Whether 2 connected nodes are similar on some characteristic

Node level

Degree centrality

Level at which a node is similar to its alters (direct connections)

How much a node is a structural hole (broker)

How do node characteristics determine its position in the network

Consequences of networks
(Networks as *Inependent variable*)

<u>Group outcomes</u>

- Performance
- Group culture
- Intervention effect

(Agneessens, 2020; 2023)

Modelling networks

Antecedents of networks
(Networks as *Dependent variable*)

Consequences of networks
(Networks as *Inependent variable*)

Group level

Density

Number of cliques

Level of "core-peripheriness"

Level of homophily in a group

Dyad level

<u>Dyadic characteristics</u>

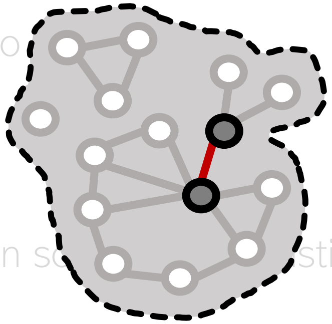
- Similarity (homophily)
- Peer-effects
- Sorrounding

Presence or absence of a tie between two

Distance between two nodes

Number of closed triads

Whether 2 connected nodes are similar on s… …stic



<u>Dyadic outcomes</u>

- Transfer

Node level

Degree centrality

Level at which a node is similar to its alters (direct connections)

How much a node is a structural hole (broker)

How do node characteristics determine its position in the network

(Agneessens, 2020; 2023)

Modelling networks

Antecedents of networks
(Networks as *Dependent variable*)

Consequences of networks
(Networks as *Inependent variable*)

Group level

Density

Number of cliques

Level of "core-peripheriness"

Level of homophily in a group

Dyad level

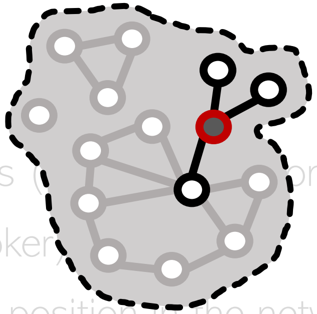
Presence or absence of a tie between two nodes

Distance between two nodes

Number of closed triads

Whether 2 connected nodes are similar on some characteristic

Node level



<u>Nodal (individual) characteristics</u>

- Age
- Values
- Reputation

→

Degree centrality

Level at which a node is similar to its alters (… ons)

How much a node is a structural hole (broker…

How do node characteristics determine its position in the network

→

<u>Nodal (individual) outcomes</u>

- Performance
- Beliefs
- Well-being

(Agneessens, 2020; 2023)

(Agneessens, 2023)

Bibliography

- Agneessens, F. (2020). Dyadic, nodal, and group-level approaches to study the antecedents and consequences of networks: Which social network models to use and when? In: Light, R., & Moody, J. (Eds.). (2020). *The Oxford Handbook of Social Networks*. Oxford University Press. https://academic.oup.com/edited-volume/34294/chapter-abstract/290740610?redirectedFrom=fulltext
- Agneessens, F. (2023). *Statistical Analysis for Cross-Sectional and Longitudinal Social Network Analysis*. Essex Summer School in Social Science Data Analysis. University of Essex
- Borgatti, S., Everett, M.G., Johnson, J.C. & Agneessens, F. (2022). Analyzing Social Networks Using R. SAGE Publications
- Newman, M.E.J. (2010). *Networks: An introduction*. Oxford, UK: Oxford University Press. Chapter 1.
- Power, E. (2021). *MY461. Social Network Analysis*. Department of Methodology, London School of Economics and Political Science
- Seguin, C., Sporns, O. & Zalesky, A. (2023). Brain network communication: concepts, models and applications. *Nat. Rev. Neurosci.* 24, 557–574
- Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z. & Raabe, I.J. (2019). Integration in emerging social networks explains academic failure and success, *Proc. Natl. Acad. Sci*, 116 (3) 792-797
- Vantaggiato, F. P. (2019). The drivers of regulatory networking: policy learning between homophily and convergence. *Journal of Public Policy*, *39*(3), 443–464