# Statistical Analysis of Networks

## Lecture 2 – Basic concepts

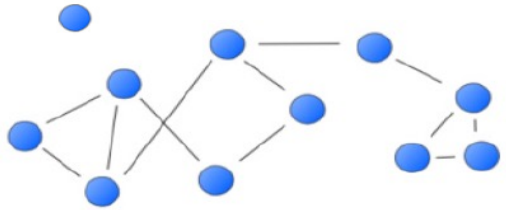
# NETWORK ANALYSIS: BASIC CONCEPTS

Network Analysis is based on:

- Relational data: relations connecting some entities to each other
  - Implication: conventional data focus on units or actors and their attributes, relational data focus on actors and relations

A (full one-mode) network is made up of:

- A set of entities $N := \{1, 2, 3 \ldots, n\}$ and at least one relation among them



entities = ***n*** nodes/***v*** vertices (in graph theory) or actors (in SNA) in ***N***

relations = ***L***/***E*** edges (in graph theory) or ties/links/connections (in SNA)

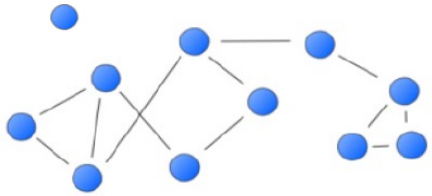
*graph and SNA terminology (and letters) will be used interchangeably!*

# NETWORK ANALYSIS: FULL VS EGO-NETWORKS

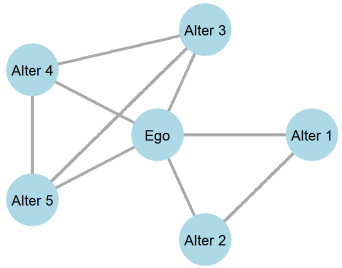
Important distinction, especially in SNA:

- Full (whole) network: relations among a given set of nodes/actors (socio-centric)

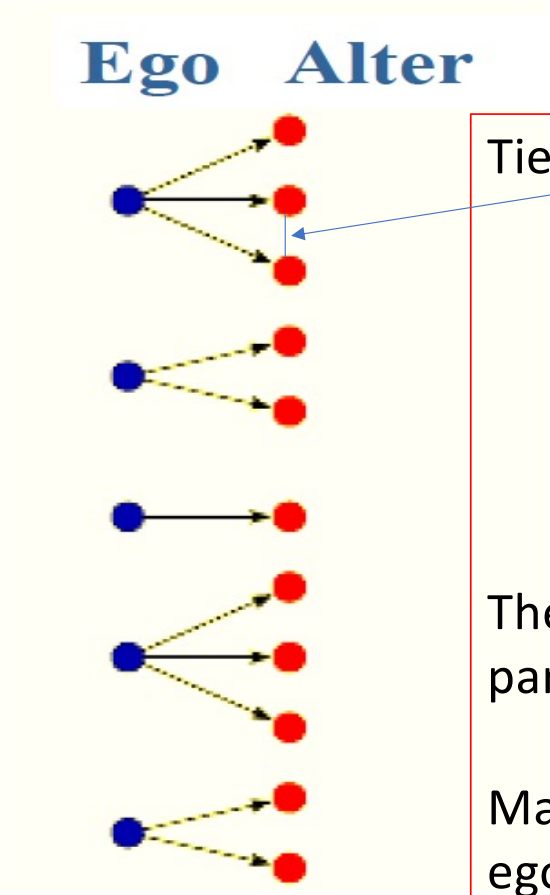
also called complete network, although complete has also a different meaning in (full) network analysis



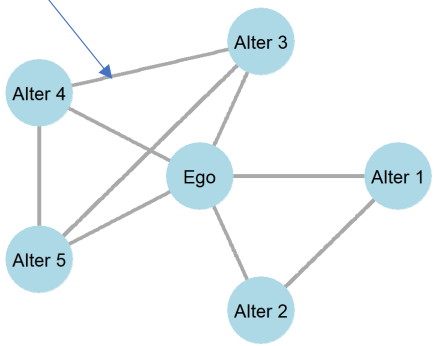
- Ego-centric network (ego-network): attention ONLY to the ties that a focal actor (Ego) has to network partners (Alters)



# EGO-CENTRIC NETWORK



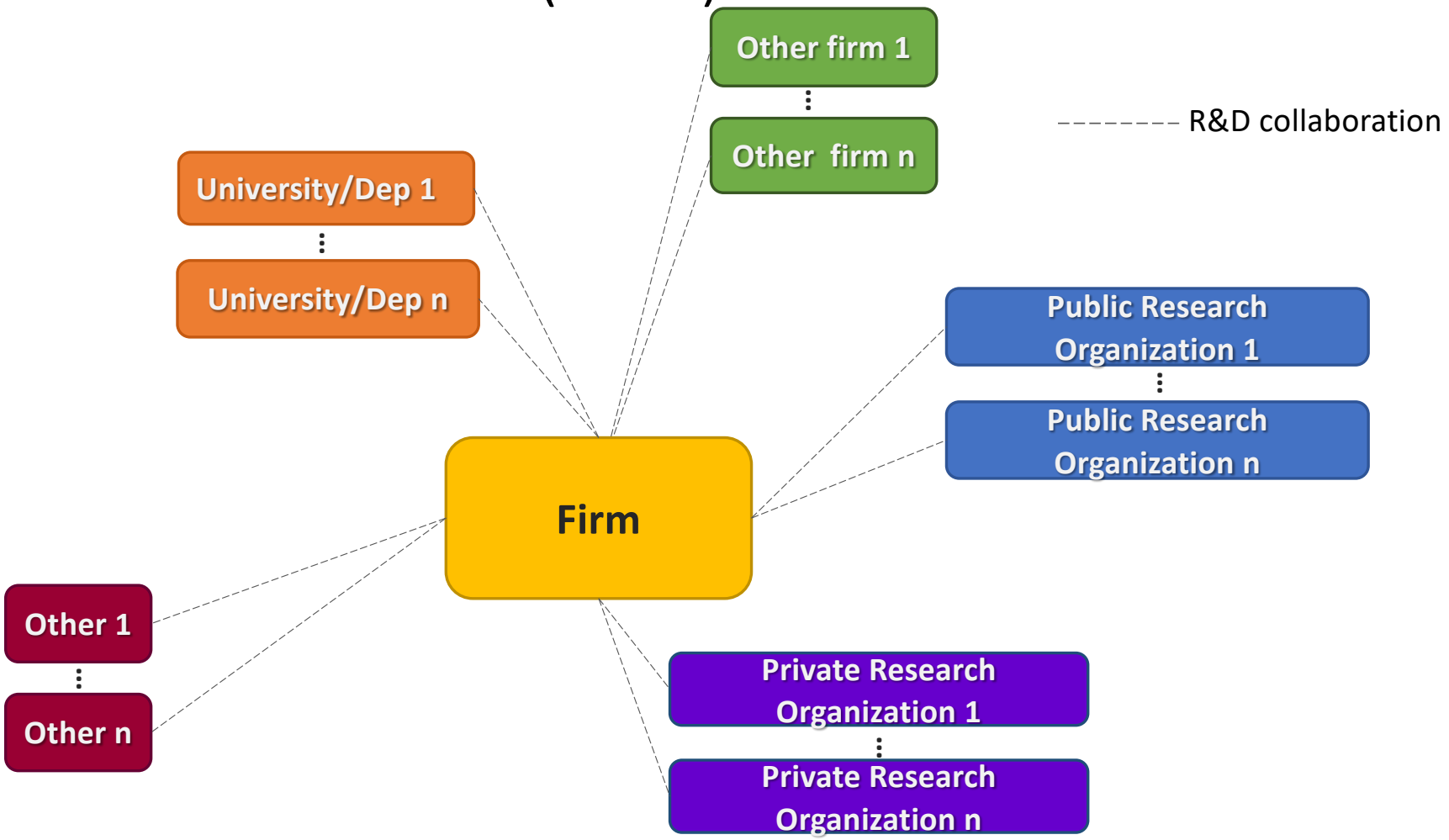

Ties among alters can also be observed (ego's evaluation)



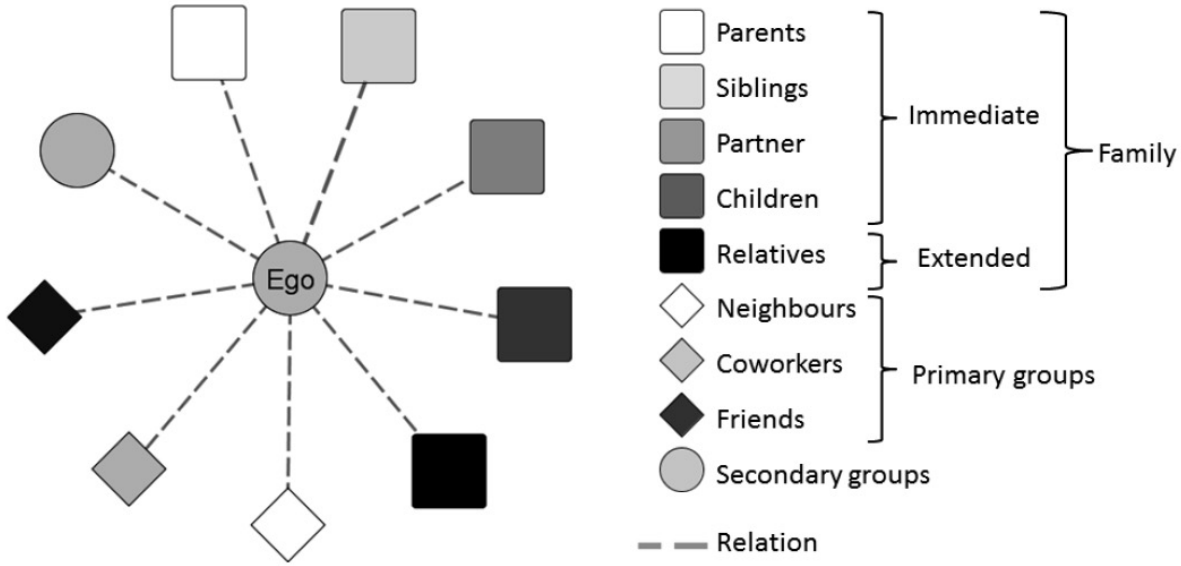

The ego-centered network design can be easily embedded as part of a representative survey of a large population

Many of the concepts of full network analysis apply to egocentric network as well

# EGO-CENTRIC NETWORK (FIRM)



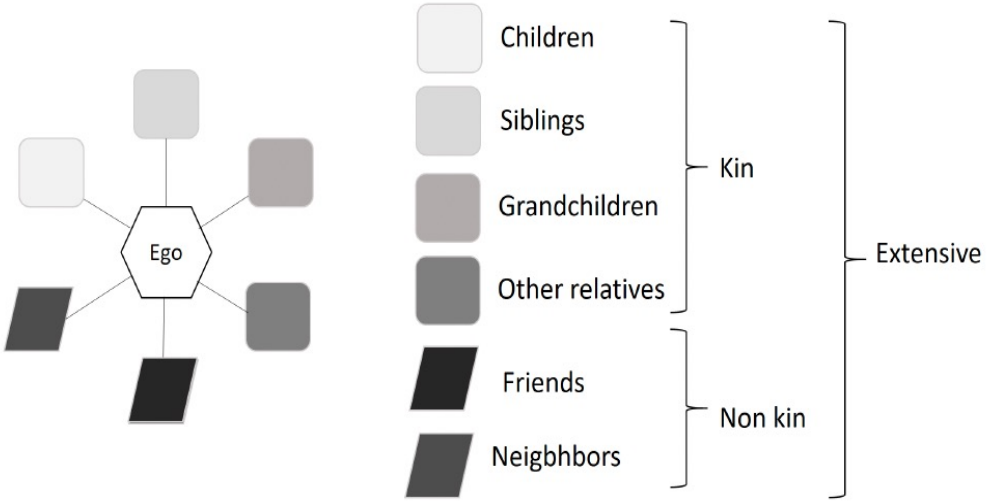
# EGOCENTRIC NETWORK – SUPPORT NETWORK (INDIVIDUAL)

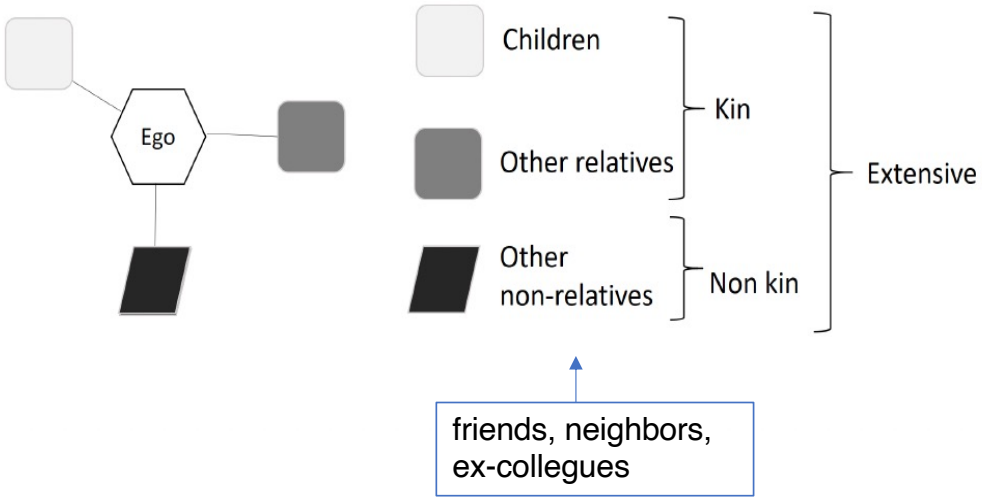



- Social contacts with others that entail a certain level of qualitative exchange such as “socioemotional aid, instrumental aid and informational aid” (Thoits, 1985; van der Poel, 1993)

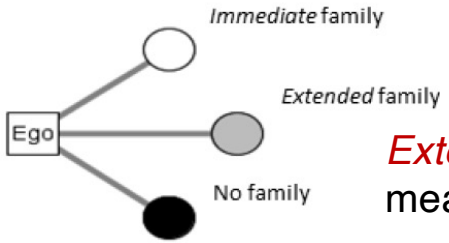
# CONTACT AND SUPPORT NETWORK OF ELDERLY PEOPLE IN ITALY (65+)

data on f-t-f contacts from large national surveys




ISTAT Family and Social Subjects Survey

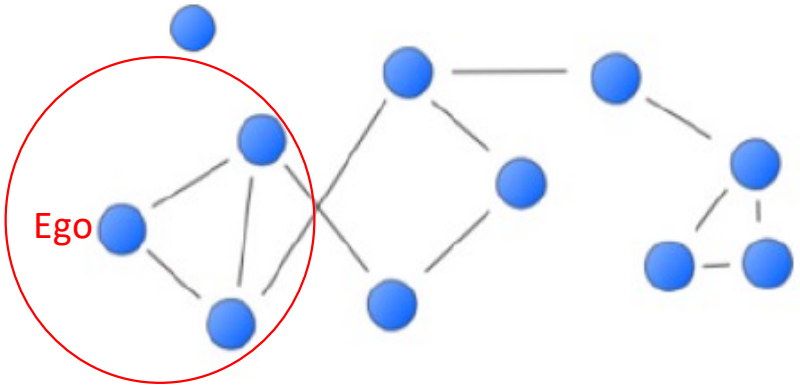



Survey of Health, Aging and Retirement in Europe – SHARE Corona Survey

*Extensive* ego-network: 30% of elderly living as singles or partners in a couple in 2016
mean size: 9 persons

# FULL NETWORK VS EGO-CENTRIC NETWORK

The set of relations of a single node in a full network represents the ego-centric network of that node




**Our focus:** full network

Many of the concepts of full network analysis apply to ego-centric network as well

# NETWORK ANALYSIS: BASIC CONCEPTS

- (full) one-mode network: a single set of entities is involved
- (full) two-mode network: two different set of entities (people and their organization, people and social events, papers and authors,…)
  - relations occur between different types of nodes
  - two-mode network: tracks the *affiliation* or *membership* of one set (affiliation network or *bipartite* network)
  - *mode*: number of sets on which a relation is observed (more than two-mode: multilevel network)

- multiplex network: multiple relations on the same set of nodes
- multilayer network: multiple kinds of interactions among different set of nodes

*Our attention: mainly on one-mode full network*

*(some extentions to two-mode networks)*

# AFFILIATION (BIPARTITE) NETWORKS: A FAMOUS DATA SET

Which women (18) attended which events (14) of the "social season" in a small southern town? Data provided by the society pages of the local newspaper.
(Davis A., Gardner B.B., Gardner M.R. *Deep South; a social anthropological study of caste and class,* University of Chicago Press, 1941)

| Names of Participants of Group I | Code Numbers and Dates of Social Events Reported in *Old City Herald* | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) 6/27 | (2) 3/2 | (3) 4/12 | (4) 9/26 | (5) 2/25 | (6) 5/19 | (7) 3/15 | (8) 9/16 | (9) 4/8 | (10) 6/10 | (11) 2/23 | (12) 4/7 | (13) 11/21 | (14) 8/3 |
| 1. Mrs. Evelyn Jefferson | X | X | X | X | X | X | . | X | X | . | . | . | . | . |
| 2. Miss Laura Mandeville | X | X | X | . | X | X | X | X | . | . | . | . | . | . |
| 3. Miss Theresa Anderson | . | X | X | X | X | X | X | X | X | . | . | . | . | . |
| 4. Miss Brenda Rogers | X | . | X | X | X | X | X | X | . | . | . | . | . | . |
| 5. Miss Charlotte McDowd | . | . | X | X | X | . | X | . | . | . | . | . | . | . |
| 6. Miss Frances Anderson | . | . | X | . | X | X | . | X | . | . | . | . | . | . |
| 7. Miss Eleanor Nye | . | . | . | . | X | X | X | X | . | . | . | . | . | . |
| 8. Miss Pearl Oglethorpe | . | . | . | . | . | X | . | X | X | . | . | . | . | . |
| 9. Miss Ruth DeSand | . | . | . | . | X | . | X | X | X | . | . | . | . | . |
| 10. Miss Verne Sanderson | . | . | . | . | . | . | X | X | X | . | . | X | . | . |
| 11. Miss Myra Liddell | . | . | . | . | . | . | . | X | X | X | . | X | . | . |
| 12. Miss Katherine Rogers | . | . | . | . | . | . | . | X | X | X | . | X | X | X |
| 13. Mrs. Sylvia Avondale | . | . | . | . | . | . | X | X | X | X | . | X | X | X |
| 14. Mrs. Nora Fayette | . | . | . | . | . | X | X | . | X | X | X | X | X | X |
| 15. Mrs. Helen Lloyd | . | . | . | . | . | . | X | X | . | X | X | X | . | . |
| 16. Mrs. Dorothy Murchison | . | . | . | . | . | . | . | X | X | . | . | . | . | . |
| 17. Mrs. Olivia Carleton | . | . | . | . | . | . | . | . | X | . | X | . | . | . |
| 18. Mrs. Flora Price | . | . | . | . | . | . | . | . | X | . | X | . | . | . |

Two-mode network:

women (first set of entities)
social events (second set of entities)

# NETWORK ANALYSIS: TYPE OF RELATIONS

Relations can be:

- undirected: there is no ordering in the vertices defining an edge $i \to j = j \to i$
  - if the relation is present between *i* and *j,* the relation is naturally mutual or reciprocated (collaboration, co-authorship, marriage, business)
  - symmetric relation

- directed: there is an ordering in the vertices defining an edge
  - a directed edge is also called arc
  - if only one arc is present, the relation is asymmetric
    - vertex *i(j)* is the *sender, j(i)* is the *receiver*
  - if both arcs are present with the vertices playing opposite roles of head and tail for the respective arcs, the two arcs are said to be mutual/reciprocated (the relation as a whole)

$i \to j \neq j \to i$

# NETWORK AND GRAPH: BASIC CONCEPTS

## (GRAPH IS A MATHEMATICAL STRUCTURE WITH A SET OF VERTICES AND A SET OF EDGES/LINKS)

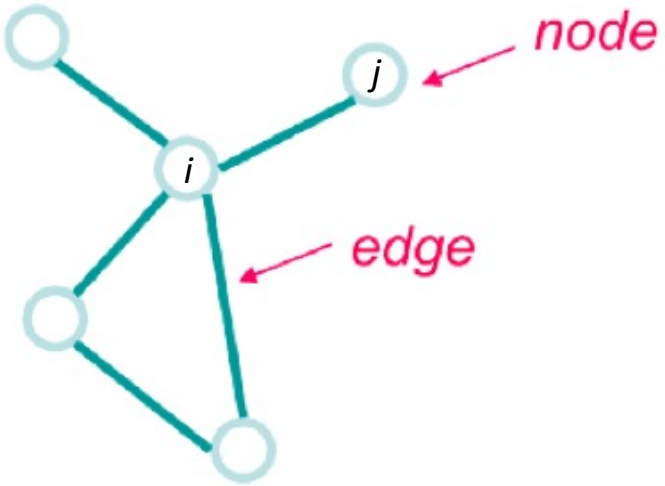
A network can be viewed as a collection of 'points' joined by 'lines'

Formally a network is a graph G(V,E) with *card(V) = n* and *card(E) = m*

*n = size or order of the network*

Distinction is not clear or universally accepted:

- *small networks* – some 10 units and lines
- *large networks* – some 1000 units and lines
- *"huge" networks* – millions of units and lines



# NETWORK AND GRAPH: BASIC CONCEPTS

## (GRAPH IS A MATHEMATICAL STRUCTURE WITH A SET OF VERTICES AND A SET OF EDGES/LINKS)

(Usually) A graph has no edges for which both endpoints connect to a single vertex (called loops) and no pairs of vertices with more than one edge between them (called multi-edges).

- A graph with either of these properties is called a multi-graph.

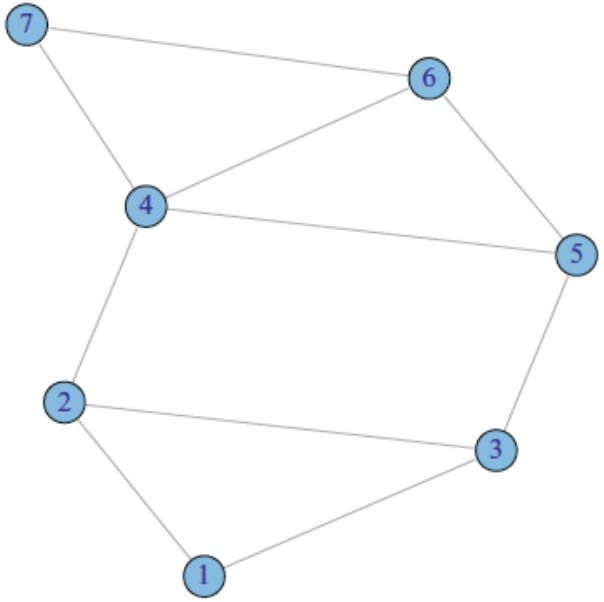
A graph G for which each edge in E has an ordering to its vertices (i.e., so that $\{i,j\}$ is distinct from $\{j,i\}$, for $i,j \in V$) is called a directed graph or digraph.

- Such edges are called directed edges or arcs, with the direction of an arc $\{i,j\}$ read from left to right, from the tail $i$ to the head $j$.

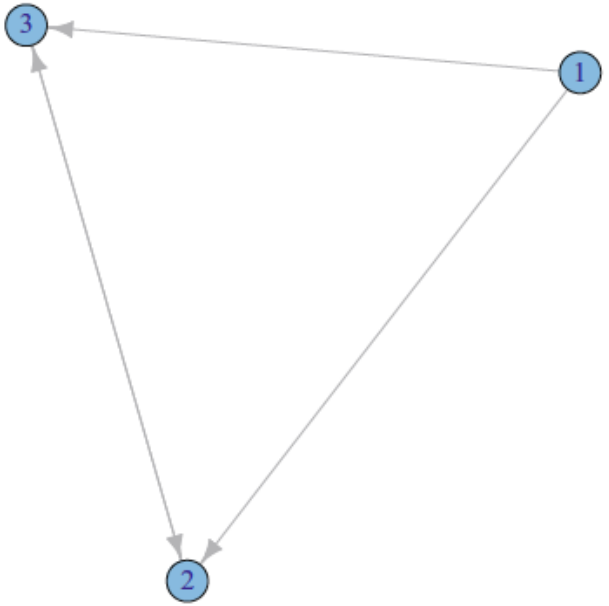
- If vertices play opposite roles of head and tail for the respective arcs the two arcs are said to be mutual (or reciprocated)

# UNDIRECTED AND DIRECTED NETWOKS/GRAPHS

undirected (*graph*)




directed (*digraph*)



# BIPARTITE GRAPH - DAVIS DATASET

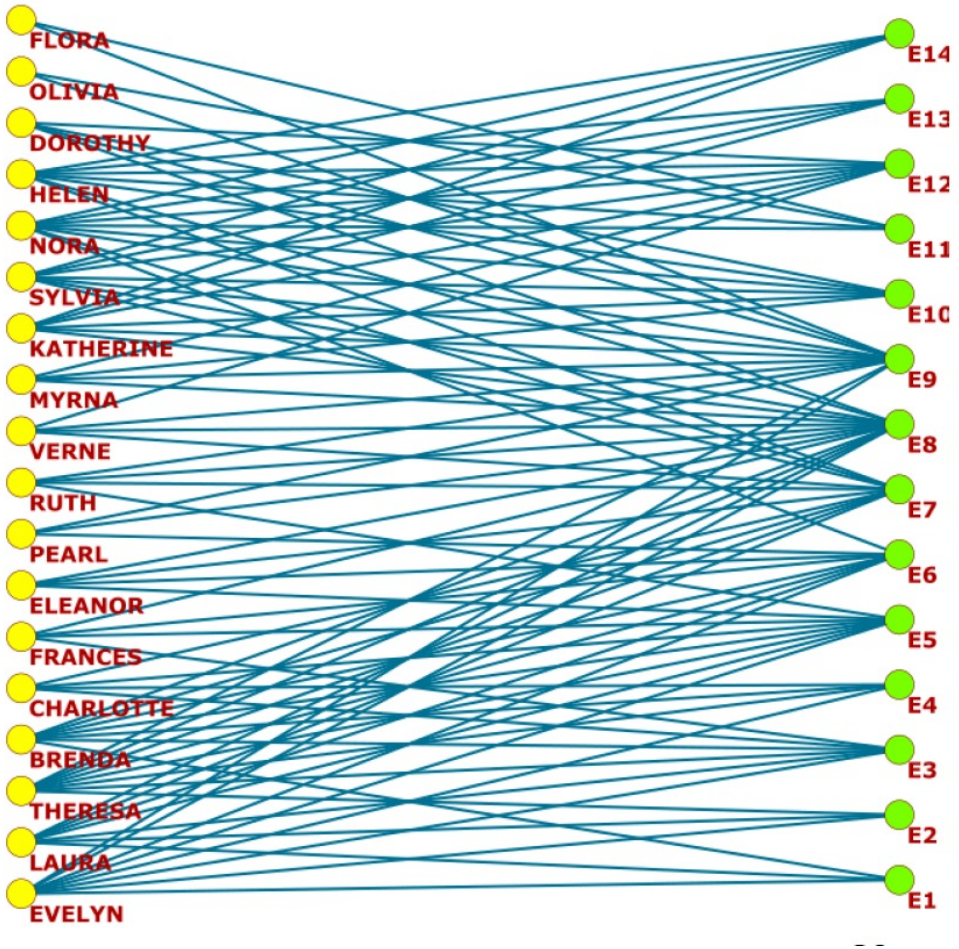
A bipartite graph is a graph G = (V,E)
such that the vertex set V may be partitioned
into two disjoint sets, say V1 and V2,
and each edge in E has one endpoint
in V1 and the other in V2.

- Such graphs are typically used to represent

'membership' networks, for example,
with 'members' (e.g., people) in V1 and
the corresponding organizations in V2

- Or to represent 'affiliation' networks,

for example, with 'participants' (e.g., people)
in V1 and the events in V2



# HYPERGRAPH AND BIPARTITE GRAPH

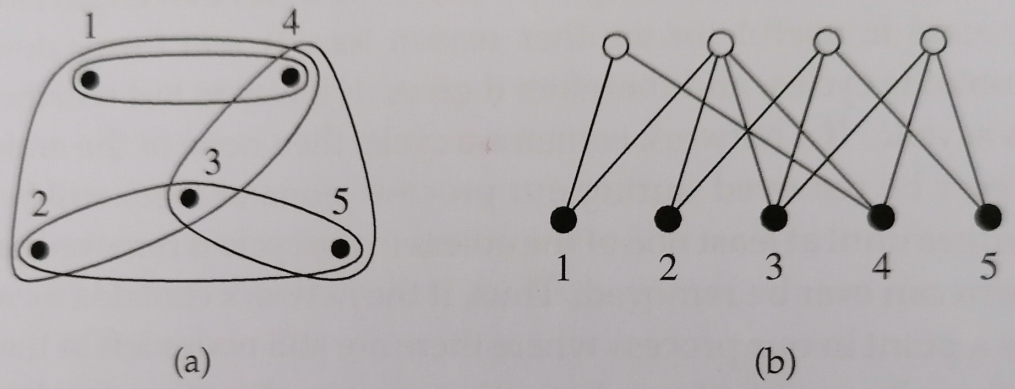



**Figure 6.4: A hypergraph and corresponding bipartite graph.** These two networks convey the same information—the membership of five nodes in four different groups. (a) The hypergraph representation in which the groups are represented as hyperedges, denoted by the loops circling sets of nodes. (b) The bipartite representation in which we introduce four new nodes (open circles at the top) representing the four groups, with edges connecting each of the original five nodes (bottom) to the groups to which it belongs.

Examples of networks that can be represented as hypergraphs or bipartite graphs

| Network | Node | Group (membership) |
|---|---|---|
| Movie actors | Actor | Cast of a movie |
| Co-authorship | Author | Authors of a paper |
| Board of directors | Director | Board of a company |
| Social events | People | Participants at social event |
| Rail connections | Station | Train routes |
| Metabolic reactions | Metabolites | Participants in a reaction |

Newman M. (2018) *Networks*, Oxford, University Press, 2nd Edition

# HYPERGRAPH AND BIPARTITE GRAPH: CO-AUTHORSHIP