

Introduction to Network Analysis

Mapping connections – Year 12 STEM Experience

Welcome to all of you!

- Where are we?
- Who am I?
- Who are you?
- Why are we here?
 - To learn about Network analysis and its applications to Geography and Spatial Analysis
 - through a mini-lecture
 - and a hands-on workshop

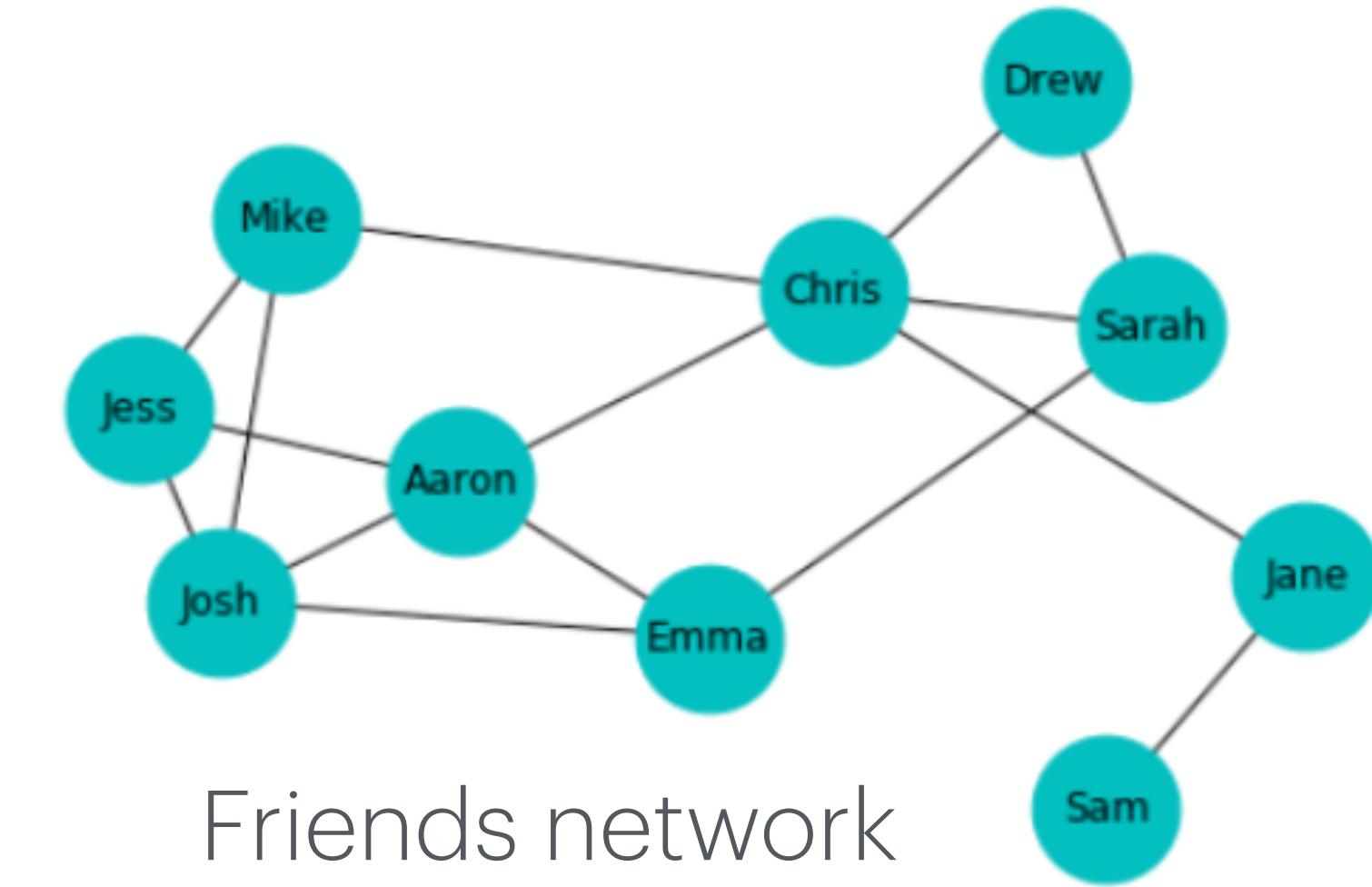
Let's get started with the mini-lecture

- But first, a **disclaimer**: do not get discouraged if you find some of these materials too challenging
- **THE MATERIALS OF THE MINI-LECTURE ARE ADVANCED**
 - These **slides are based on** those used by Prof. Jennifer Widom for a lecture in the **course titled “Working with data”**
 - The course is **addressed to** undergraduate and graduate **students from Stanford University**
 - Link to Stamford University course [here](#)
- AND **SO ARE THOSE OF THE HANDS-ON WORKSHOP**
 - **Based on online materials by Dr Martin Grandjean**, a Senior Researcher from the University of Lausanne

Grandjean, Martin (2015). "GEPHI – Introduction to Network Analysis and Visualization", <https://www.martingrandjean.ch/gephi-introduction>

What are networks?

- A **network** or a **graph** is an abstract concept, used to represent connections between elements
- A network is formed by
 - a set of **nodes** (also known as **vertices** – singular **vertex**)
 - and a set of **edges** (also known as links)
 - Edges can be **undirected** or **directed**



Friends network
(Undirected)

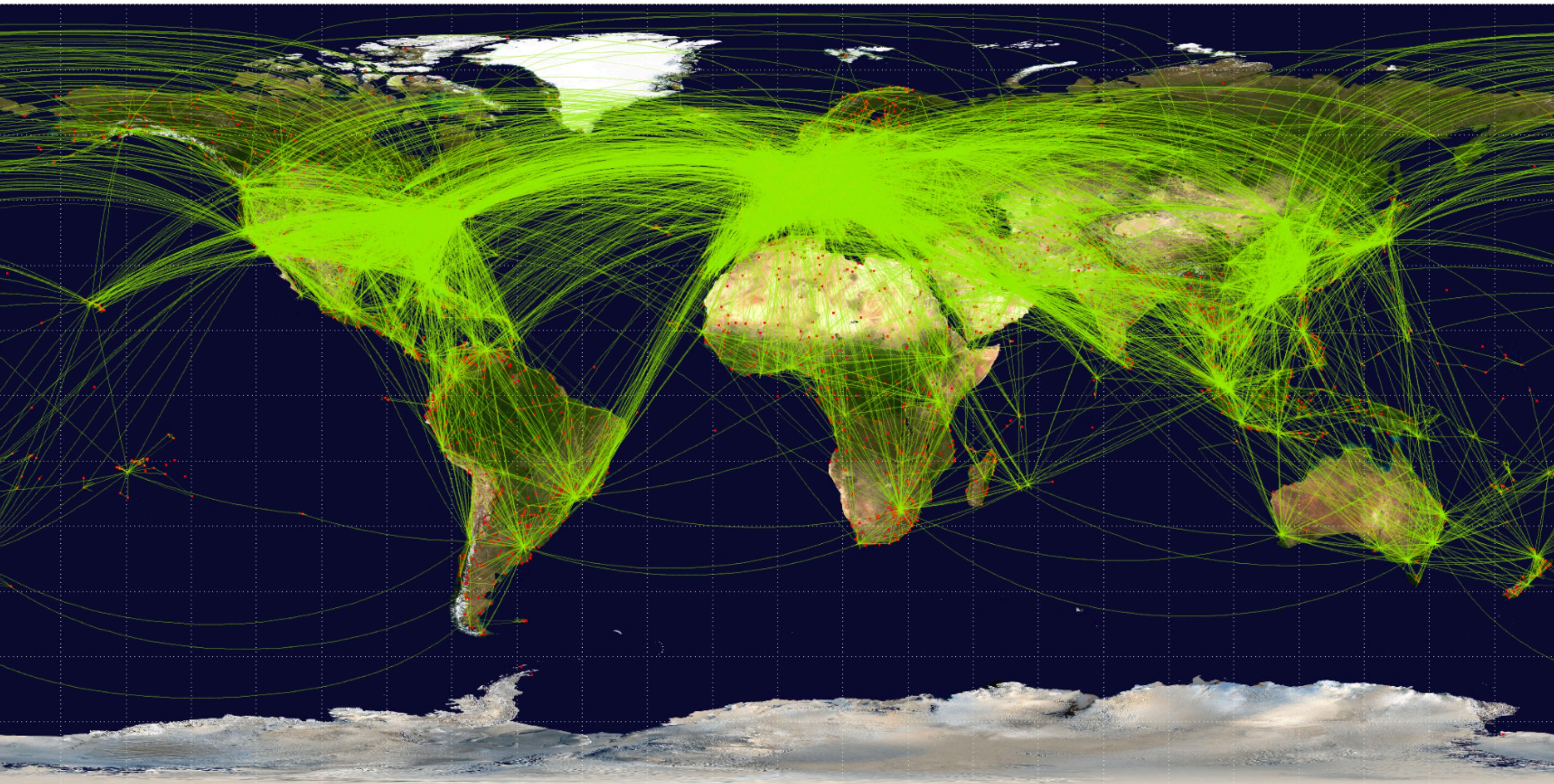
Example: Flight routes



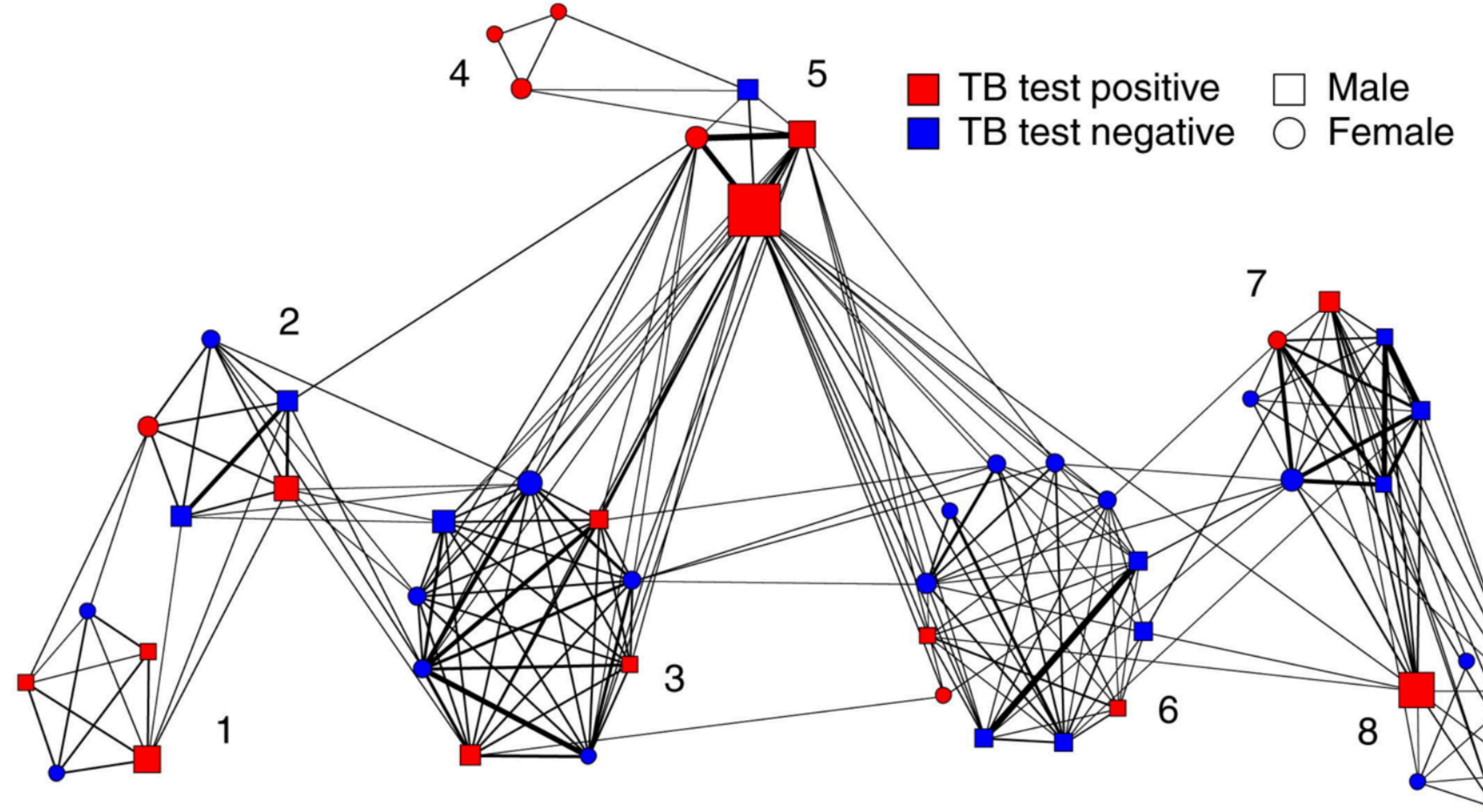
Example: Flight routes



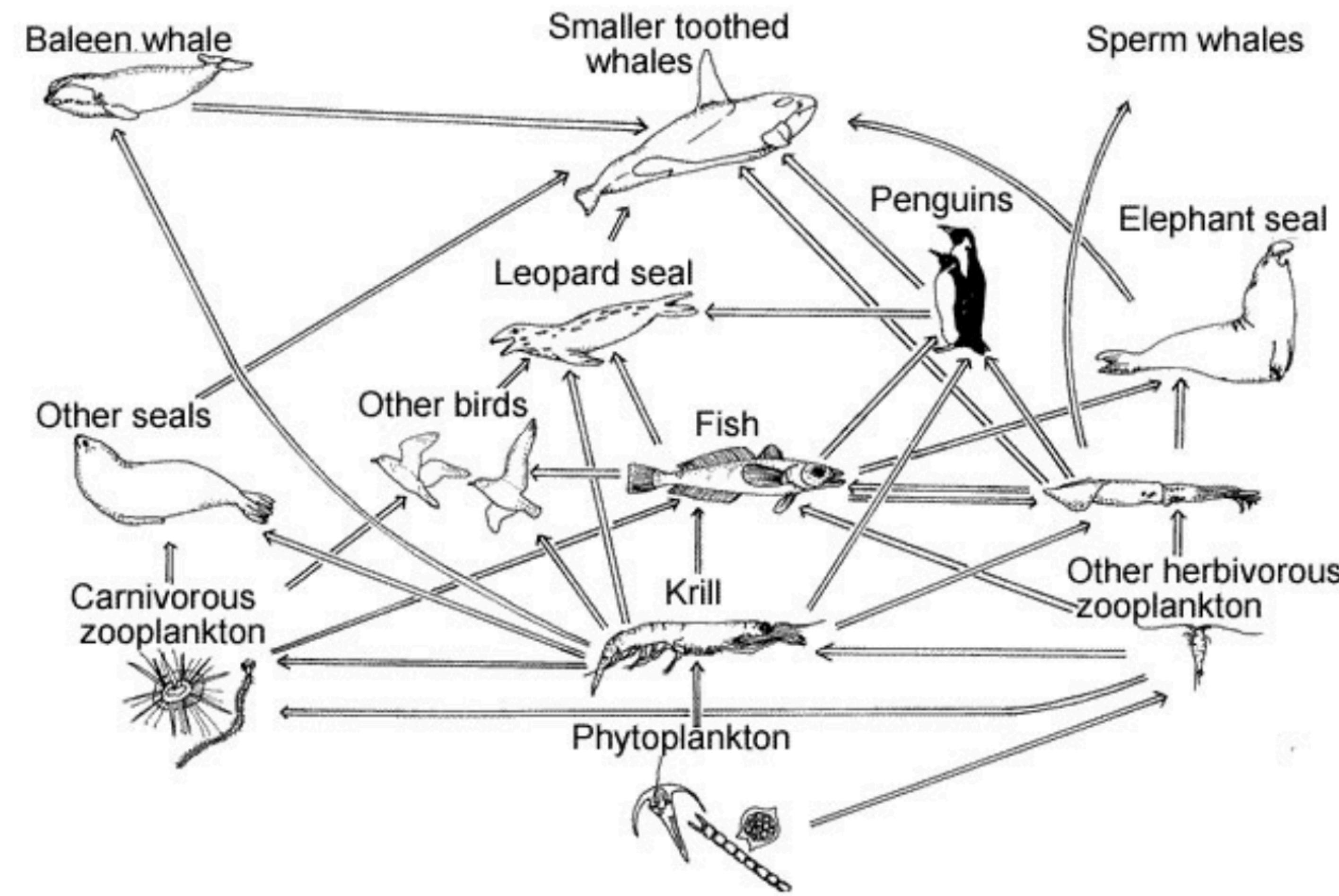
Example: Flight routes



Example: Disease transmission



Example: Food chain



Example: Instagram followers



And many more examples!

- Electricity grid
- Rail network
- Street network
- Migration network
- Organisation of staff in a firm
- **Can you think of more?!**

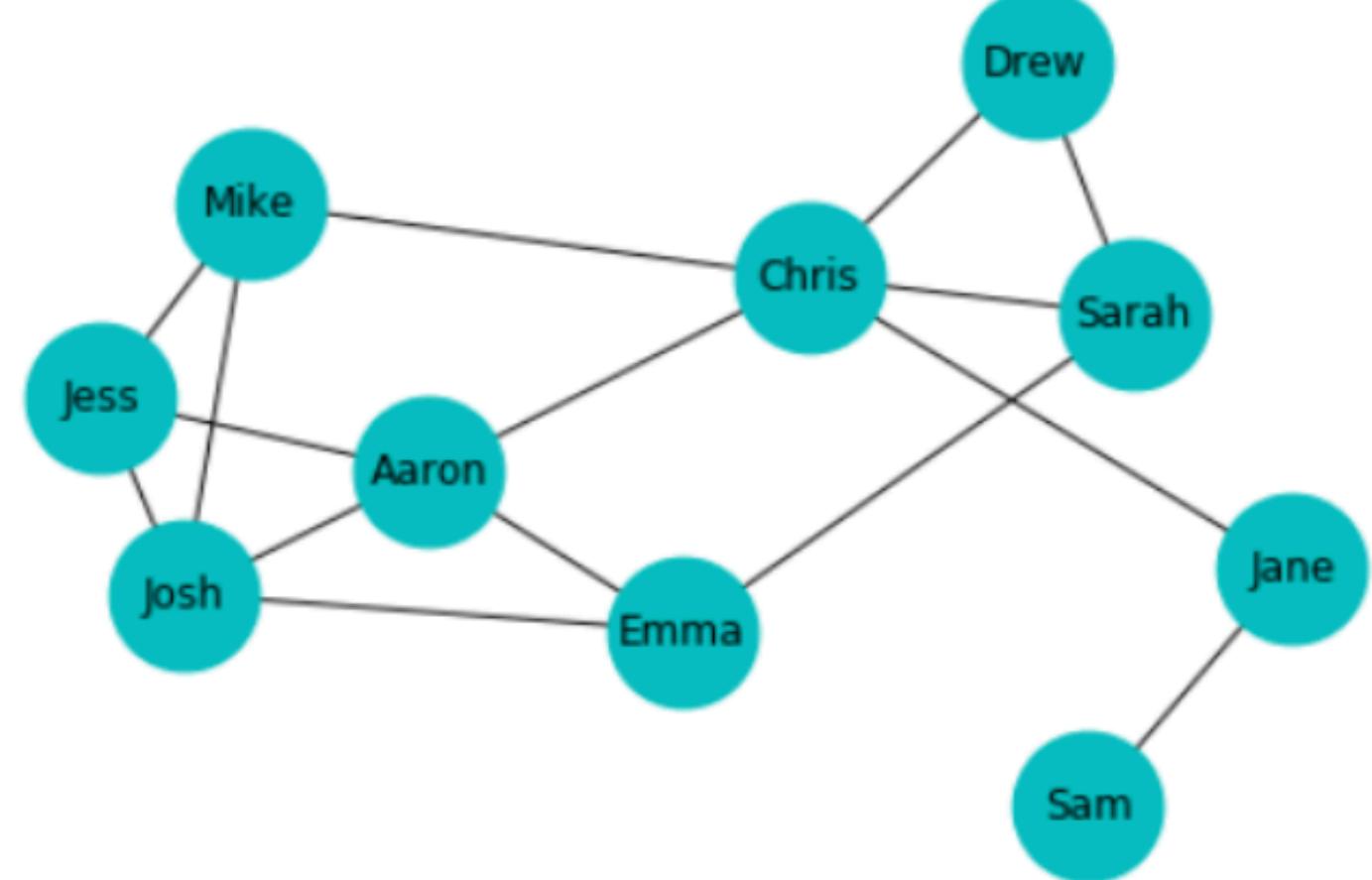
Becoming a Network Analysis pro requires:

- Learning enough mathematics and statistics
- Becoming good at computational methods and coding
- Being good with data manipulation
- Being able to do some nice data visualisation
- Mastering the art of interpreting (sometimes confusing) results

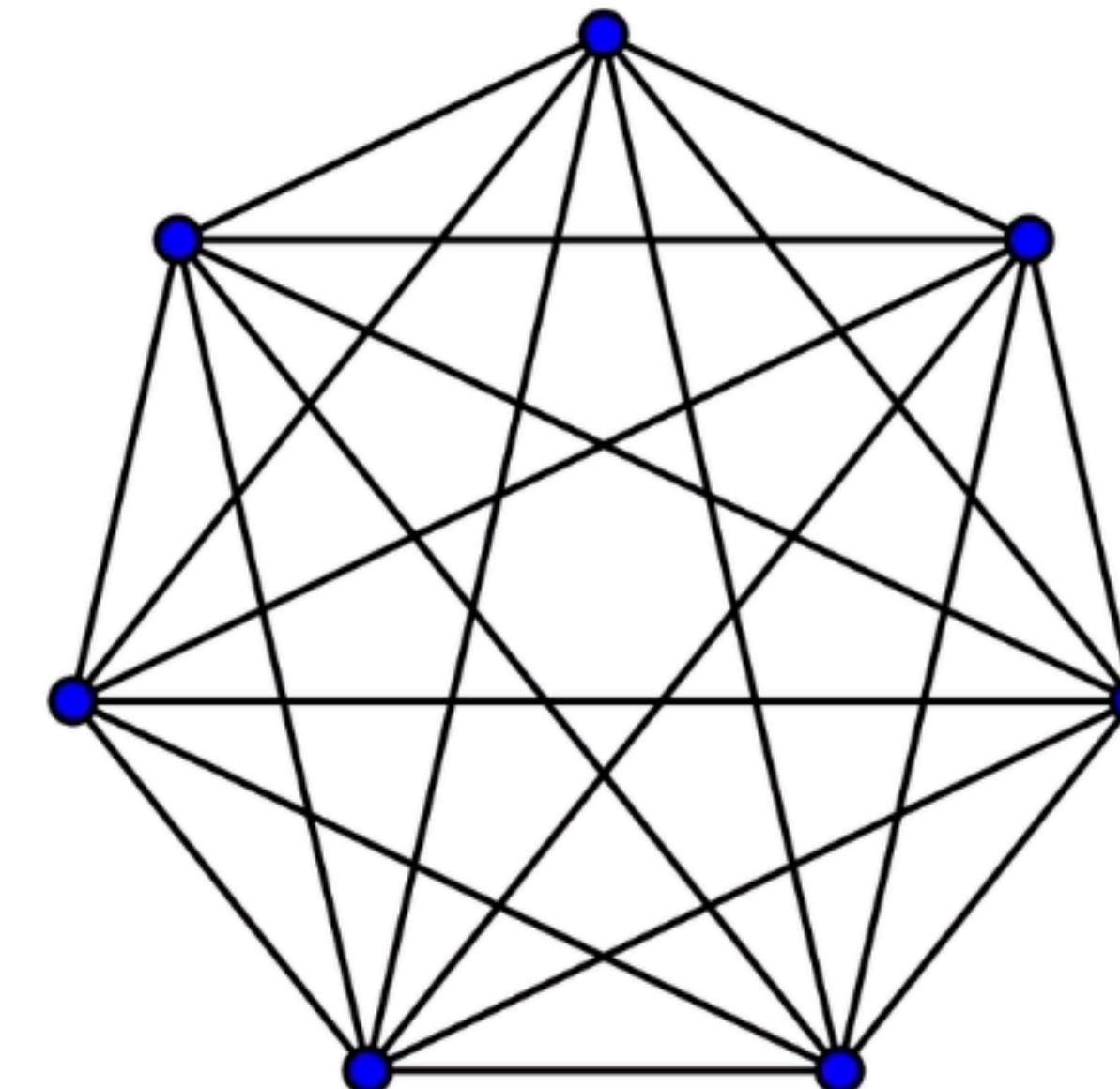
Properties of a network

Density

$$\frac{\text{Number of observed edges}}{\text{Number of possible edges}}$$



Network with density < 1

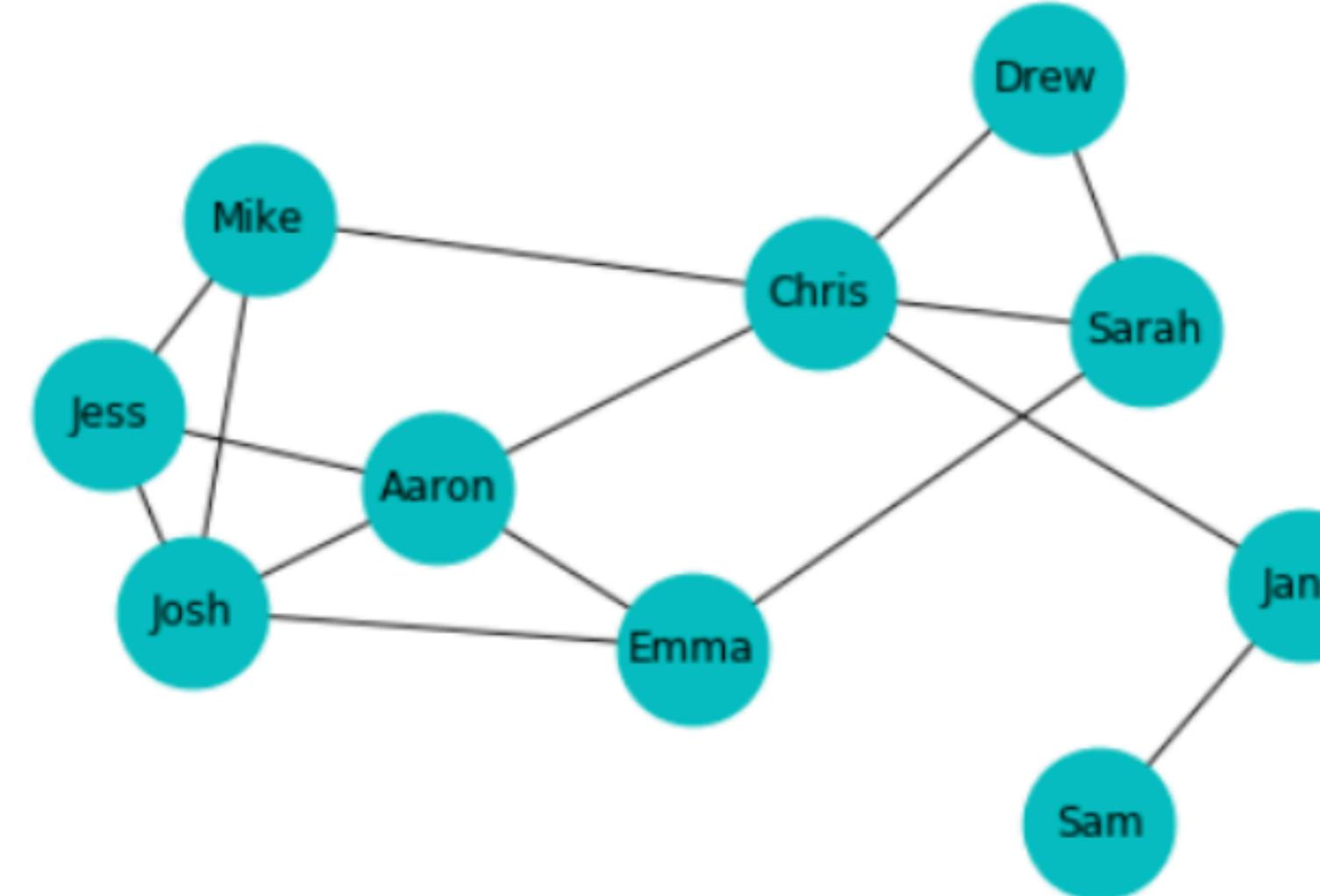


Network with density = 1

Properties of a network

Shortest paths between a pair of nodes

Minimum number of edges that need to be traversed to go from node A to node B



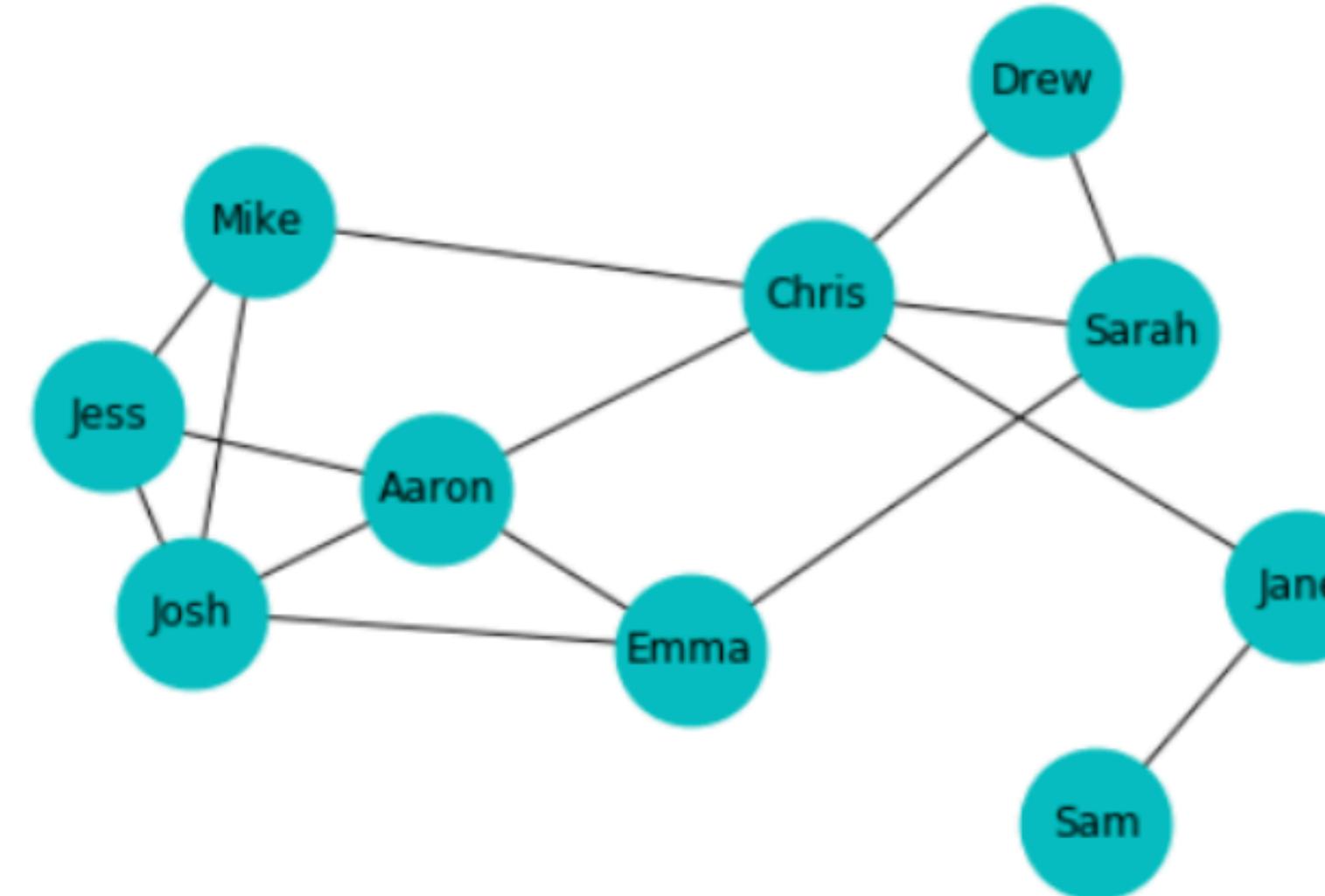
Idea of **small world** and “**Six degrees of separation**”

(Or even less with social media, which makes our world smaller)

Properties of a network

Diameter of a network

Maximum shortest path in a network

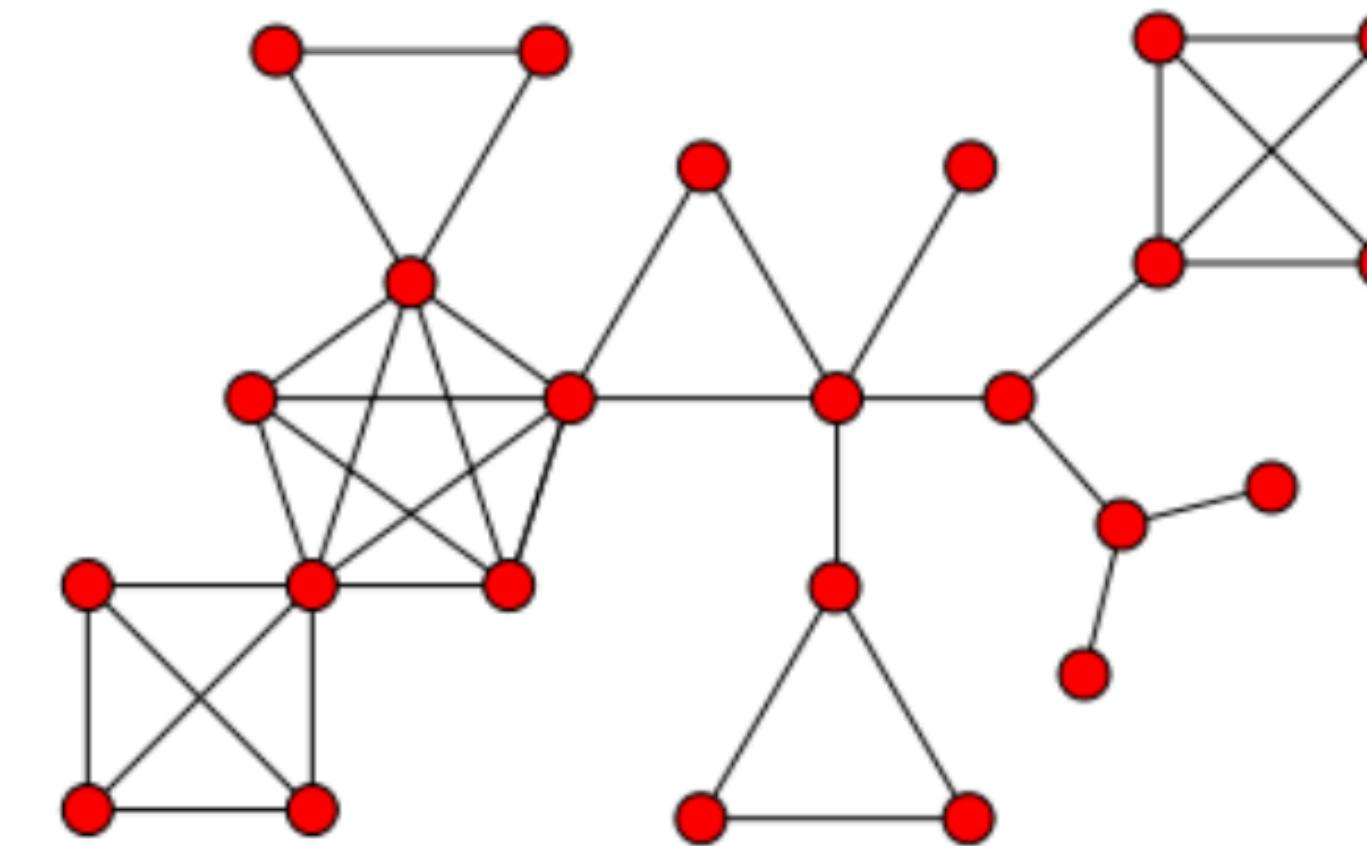
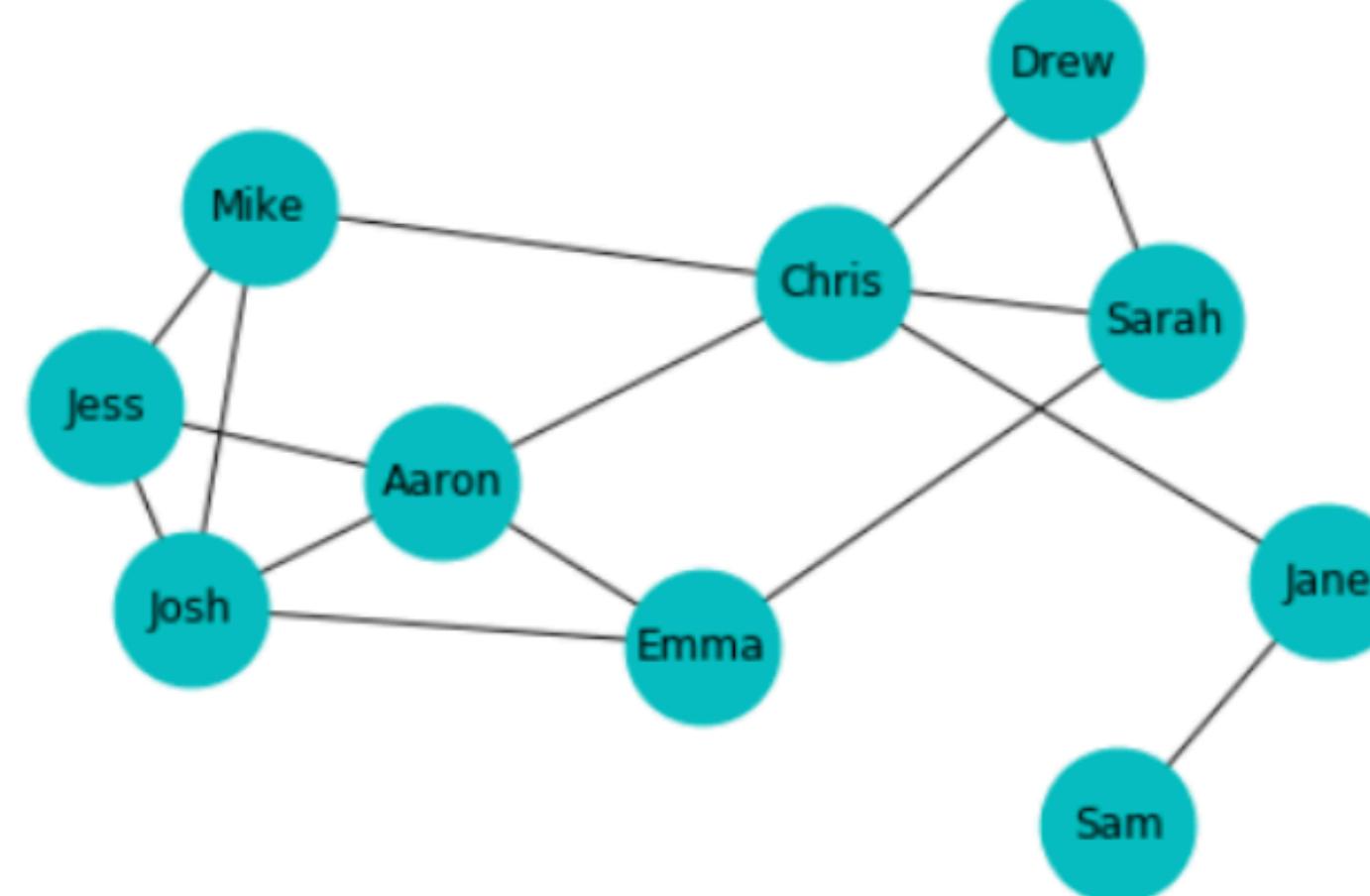


Gives us an indication of “**the size**” of the network

Properties of a network

Cliques in a network

Sets of fully-connected nodes

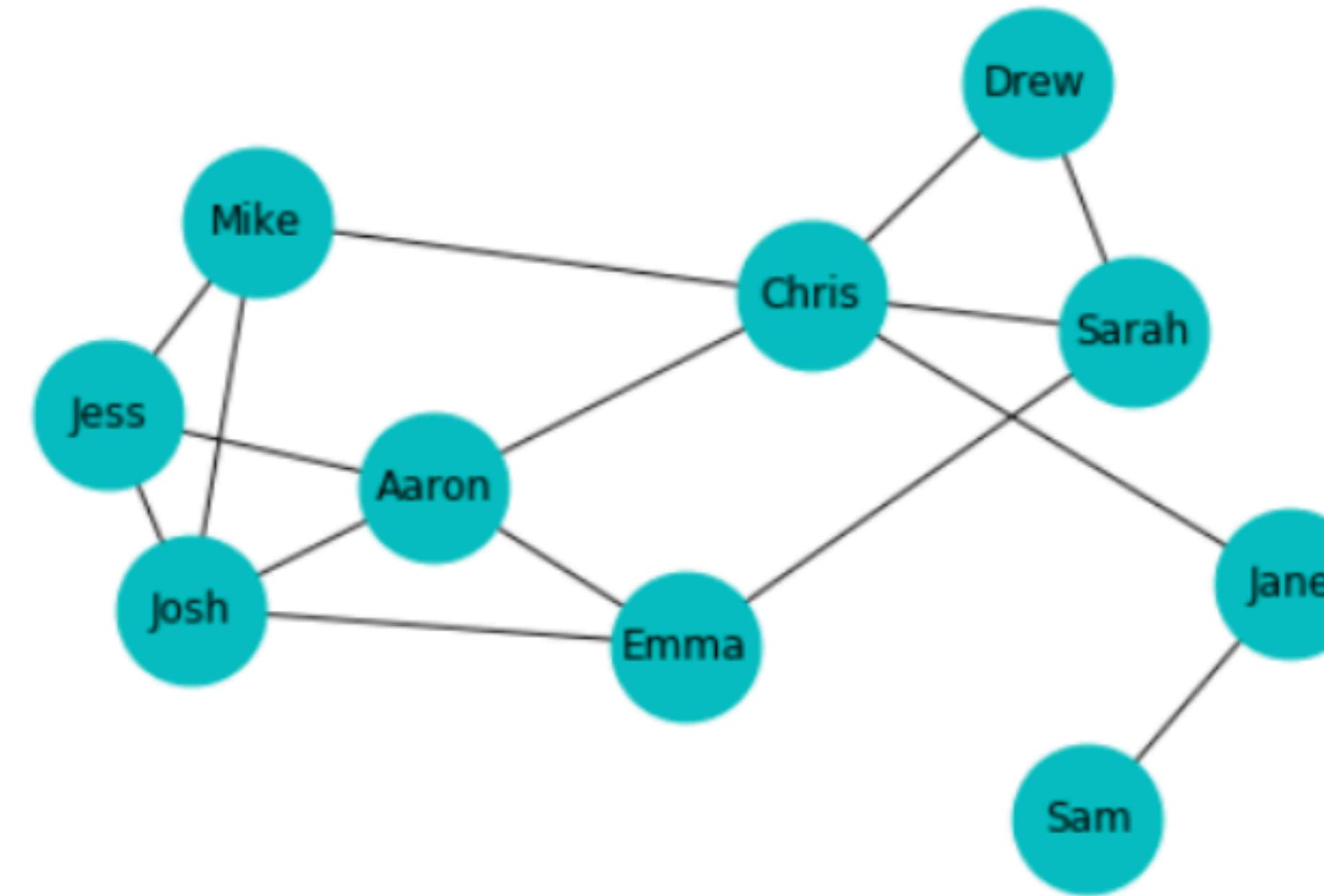


Commonly found in social networks where edges represent friendships

Properties of a network

Closeness centrality

Gives an indication of the average shortest distance to all other nodes

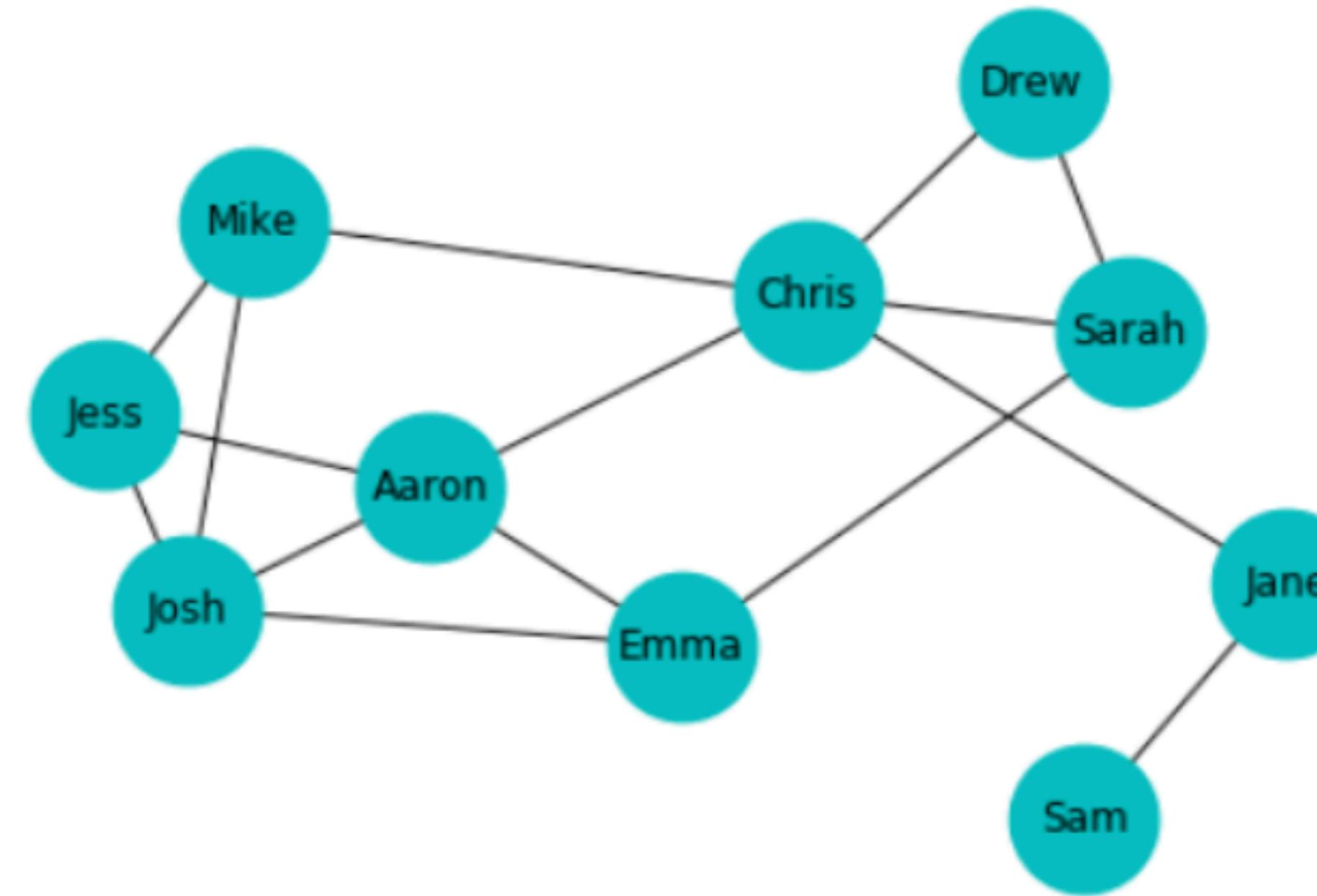


If a node has high closeness centrality, it means that the shortest path between itself and any other node in the network tends to be relatively small

Properties of a network

Betweenness centrality

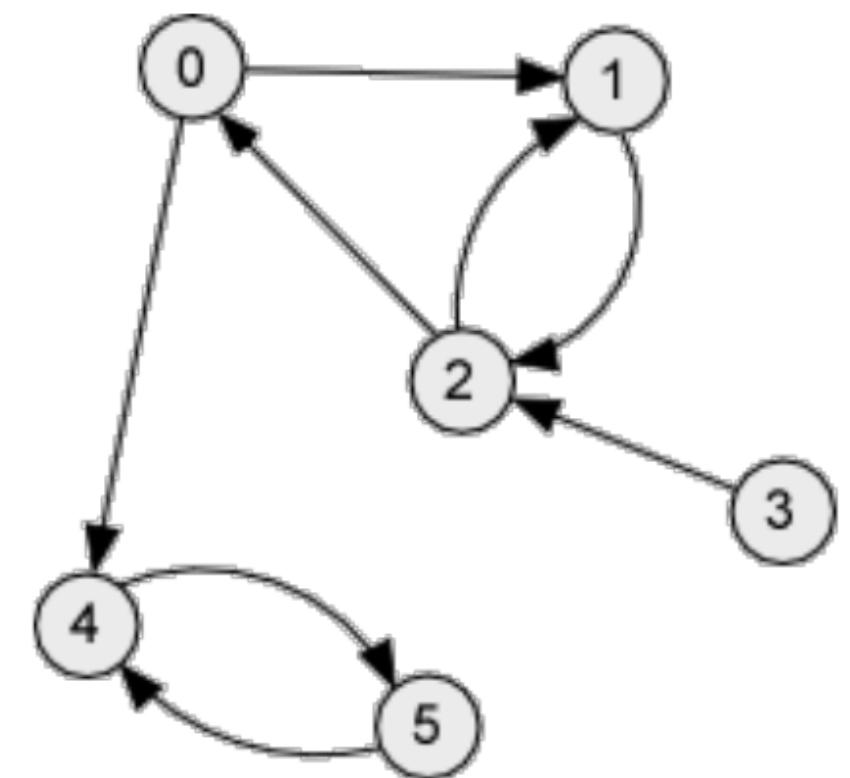
Number of shortest paths that a node lies in



If a node has high betweenness centrality, it means that it is very important to ensure the connectivity of the network

Properties of a directed network

- **In-degree** – How many edges are pointing towards a node? **How many followers?**
- **Out-degree** – How many edges are emerging from a node? **How many following?**
- **Reciprocity** – How often are links bidirectional?
- **Cycles**



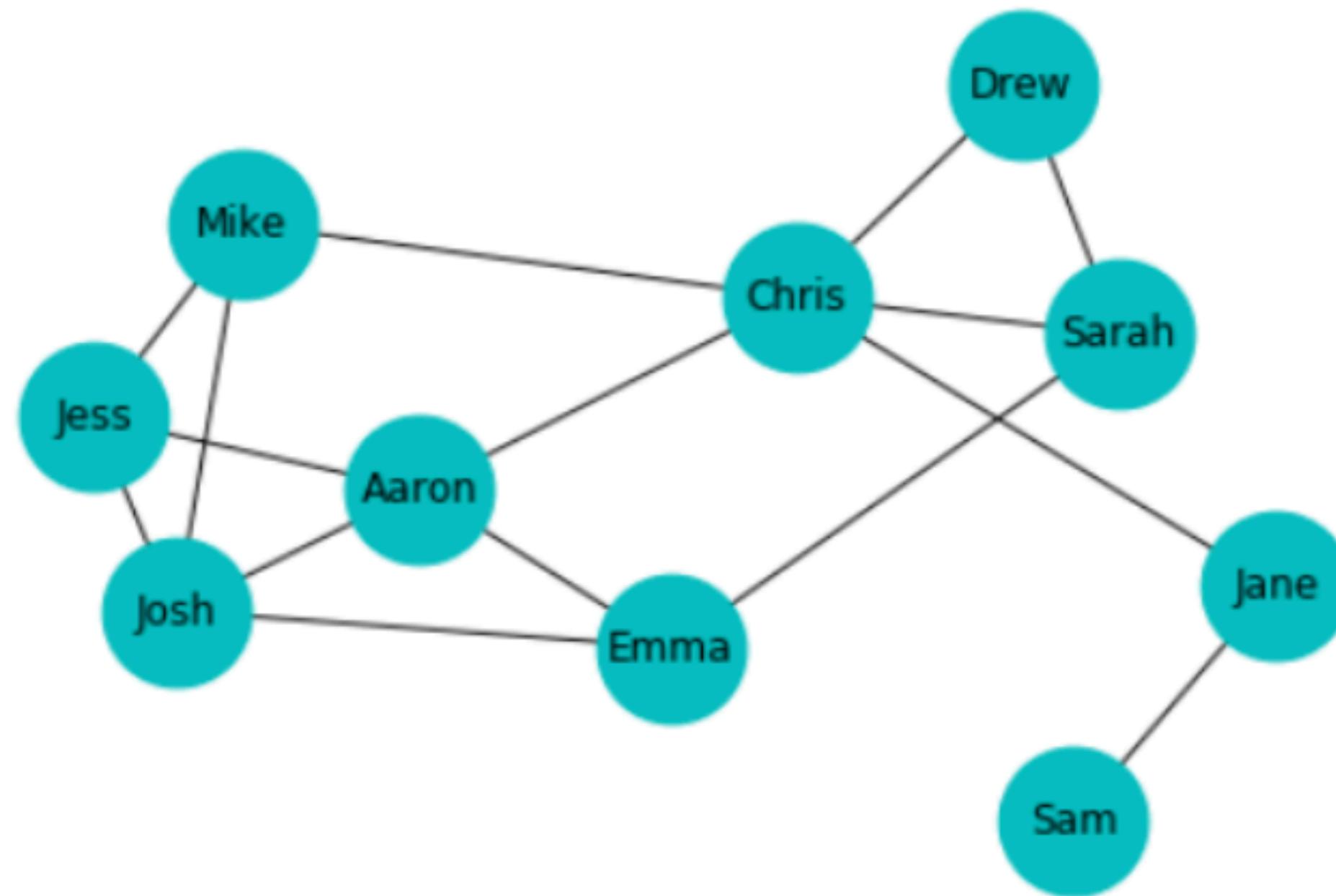
Properties of a directed network

- Nodes and edges can have **attributes**
- Examples:
 - **Name** of a node
 - **If node is a city**, node attribute: **population** of the city
 - If an edge represents a migratory flow between cities, edge attribute: **number of people migrating**

Other analysis

Link prediction

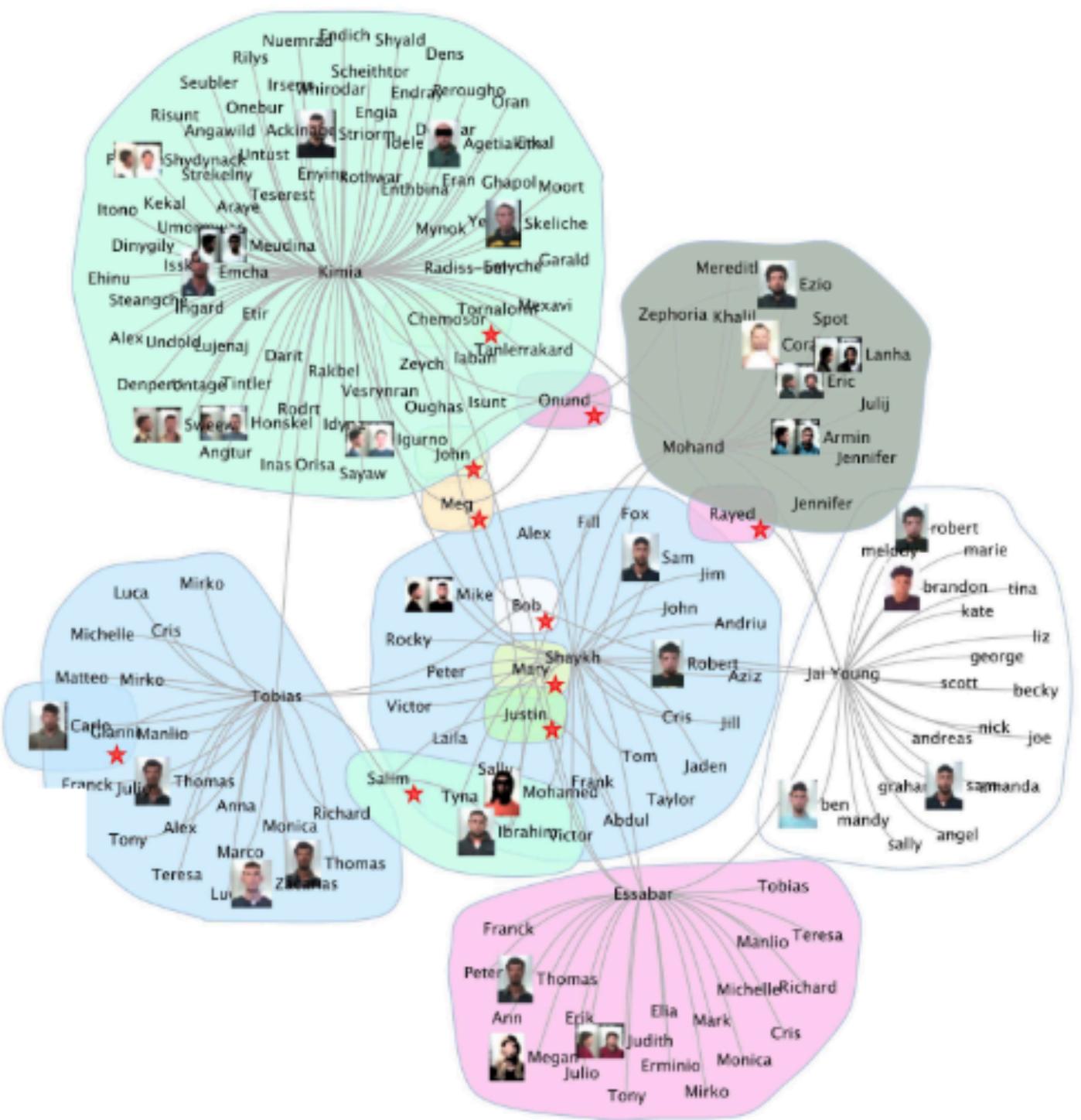
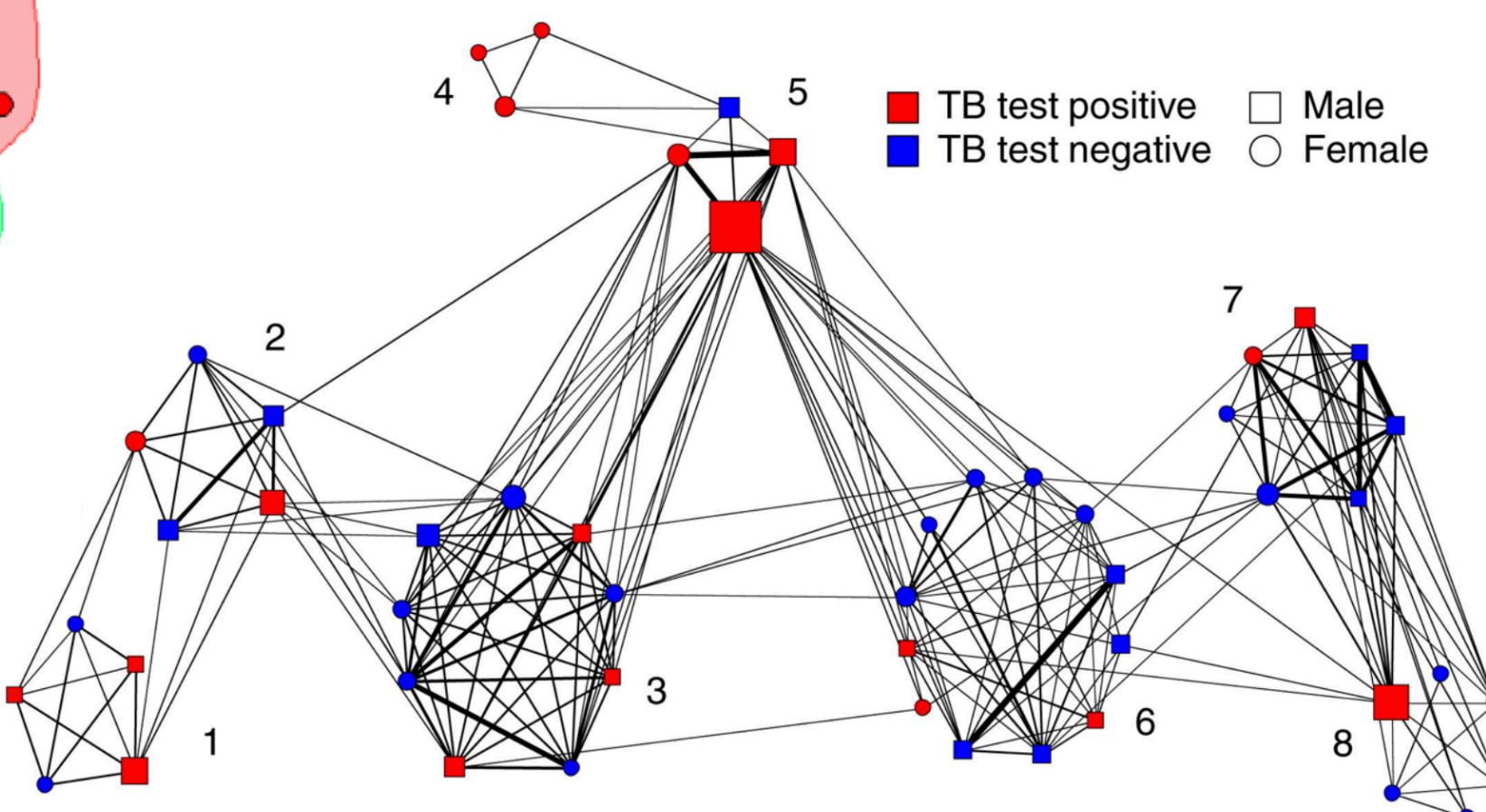
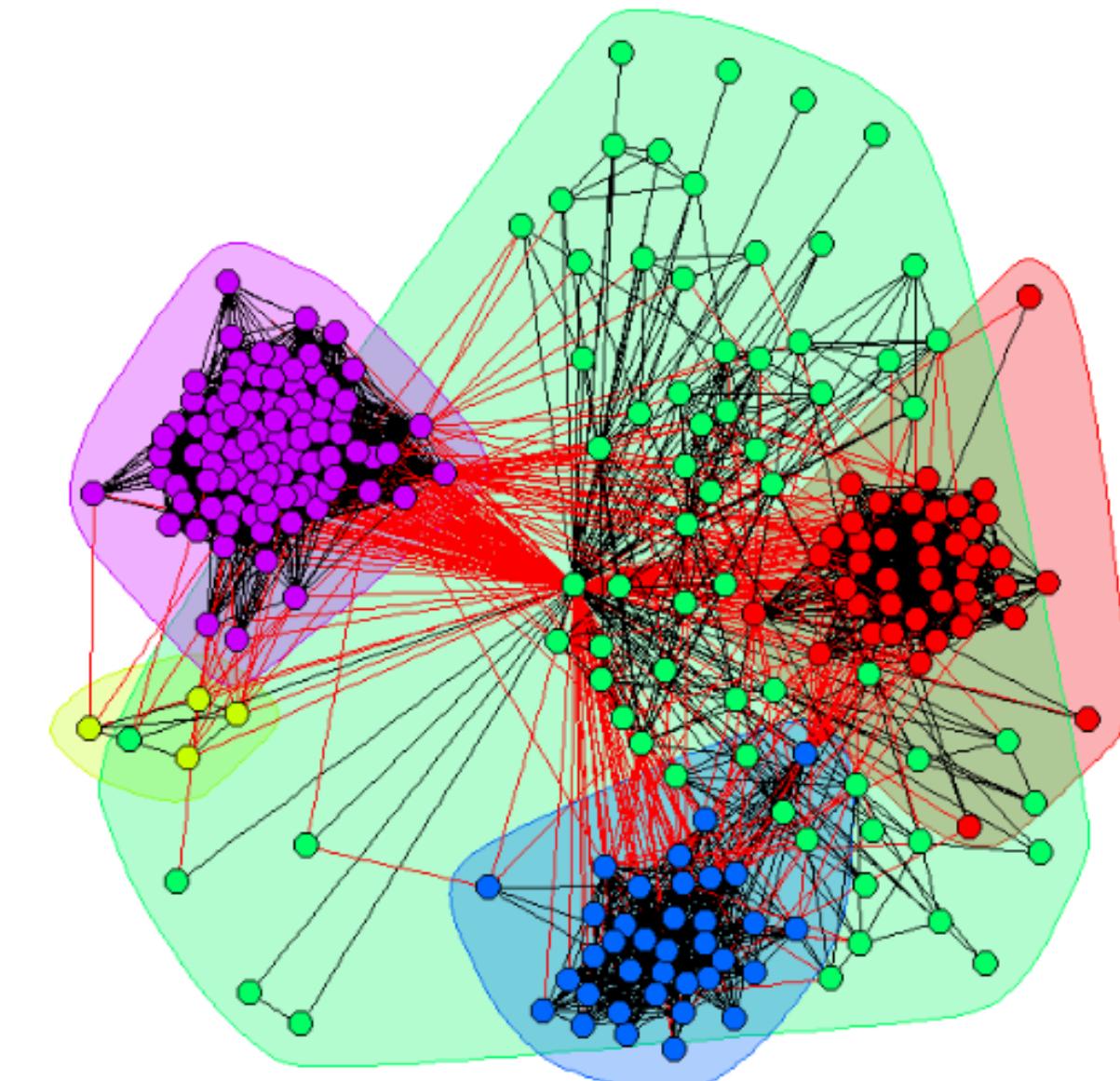
Predict future edges added to the network



Other analysis

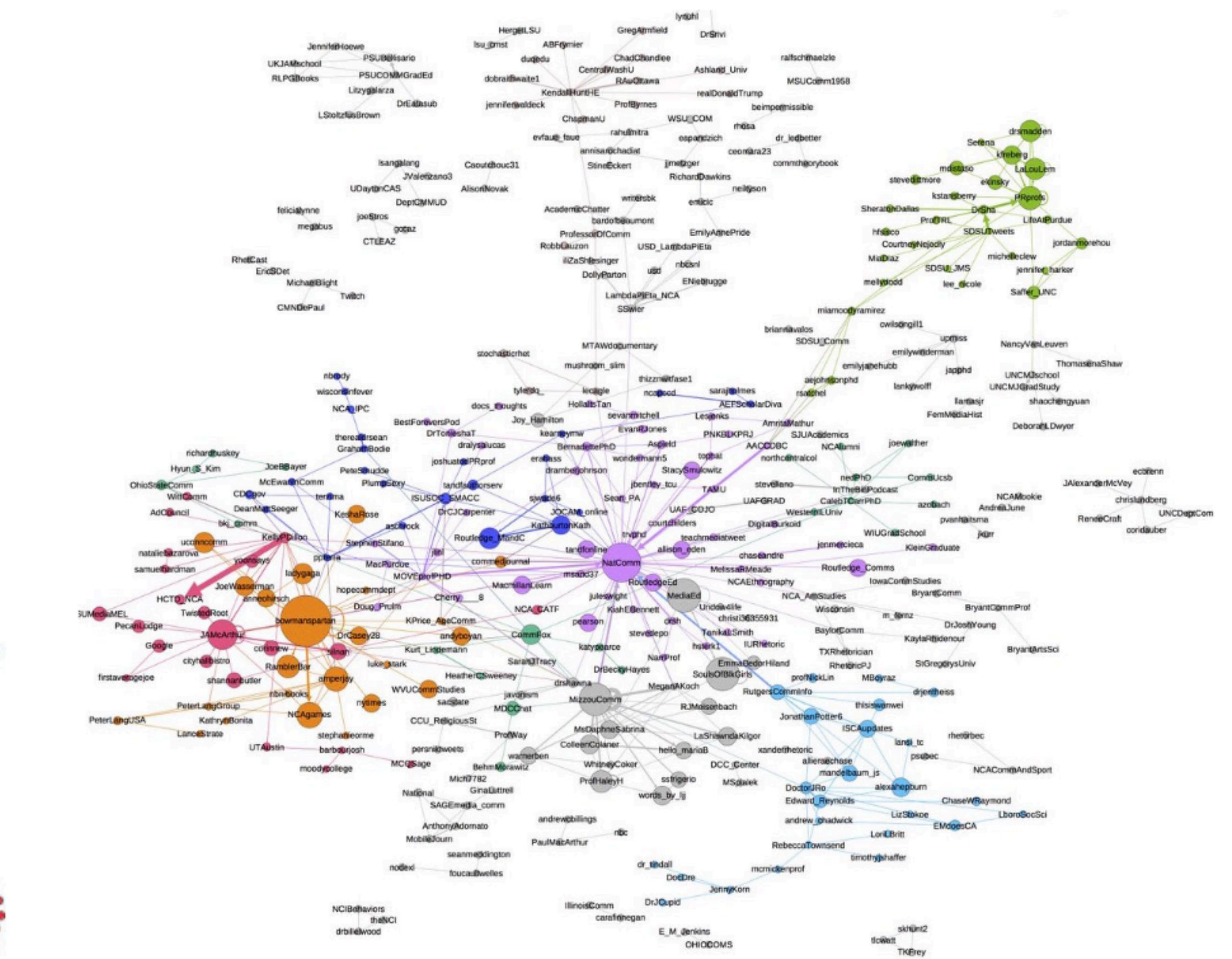
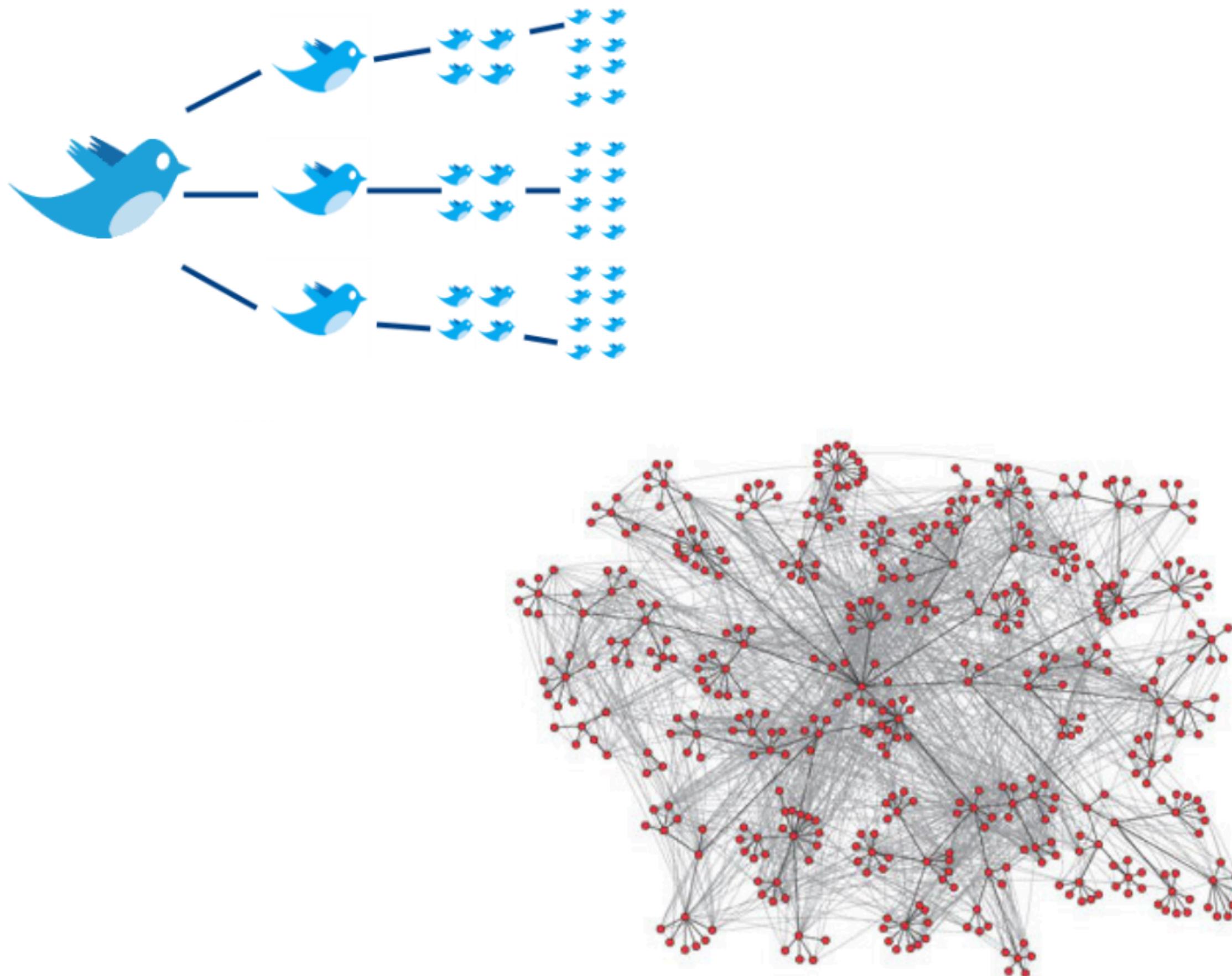
Community detection

Sets of interlinked or similar nodes



Other analysis

Modelling of “cascades” or propagation of information



Next on, a hands-on workshop on network analysis

- **Datasets**

- Mail data for number of letters sent between pairs of European cities

- **Software** for network analysis:

- Programming languages such as:

- Python (**NetworkX**)

- R (**iGraph**)

- Applications for network visualisation:

- **Gephi**