Network Analysis

CS102
## Spring 2020

Data Tools and Techniques

- **Basic Data Manipulation and Analysis**
  Performing well-defined computations or asking well-defined questions ("queries")

- **Data Mining**
  Looking for patterns in data

- **Machine Learning**
  Using data to build models and make predictions

- **Data Visualization**
  Graphical depiction of data

- **Data Collection and Preparation**

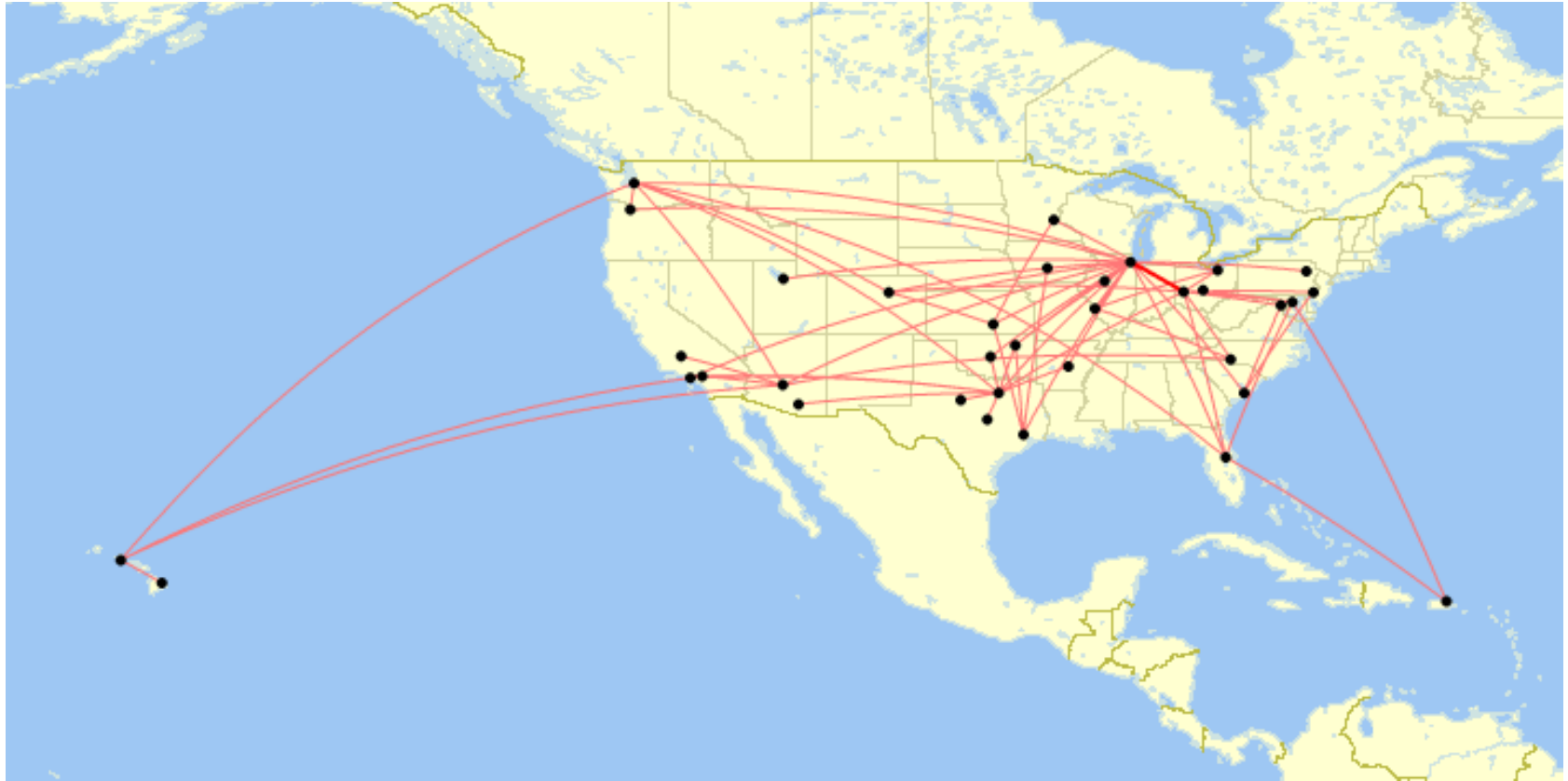Over a specific type of data

Networks

A real-world network is modeled in the computer as a graph:

- A set of nodes (or vertices, singular vertex)
- Some nodes are connected by edges (or links)
- Edges can be undirected or directed

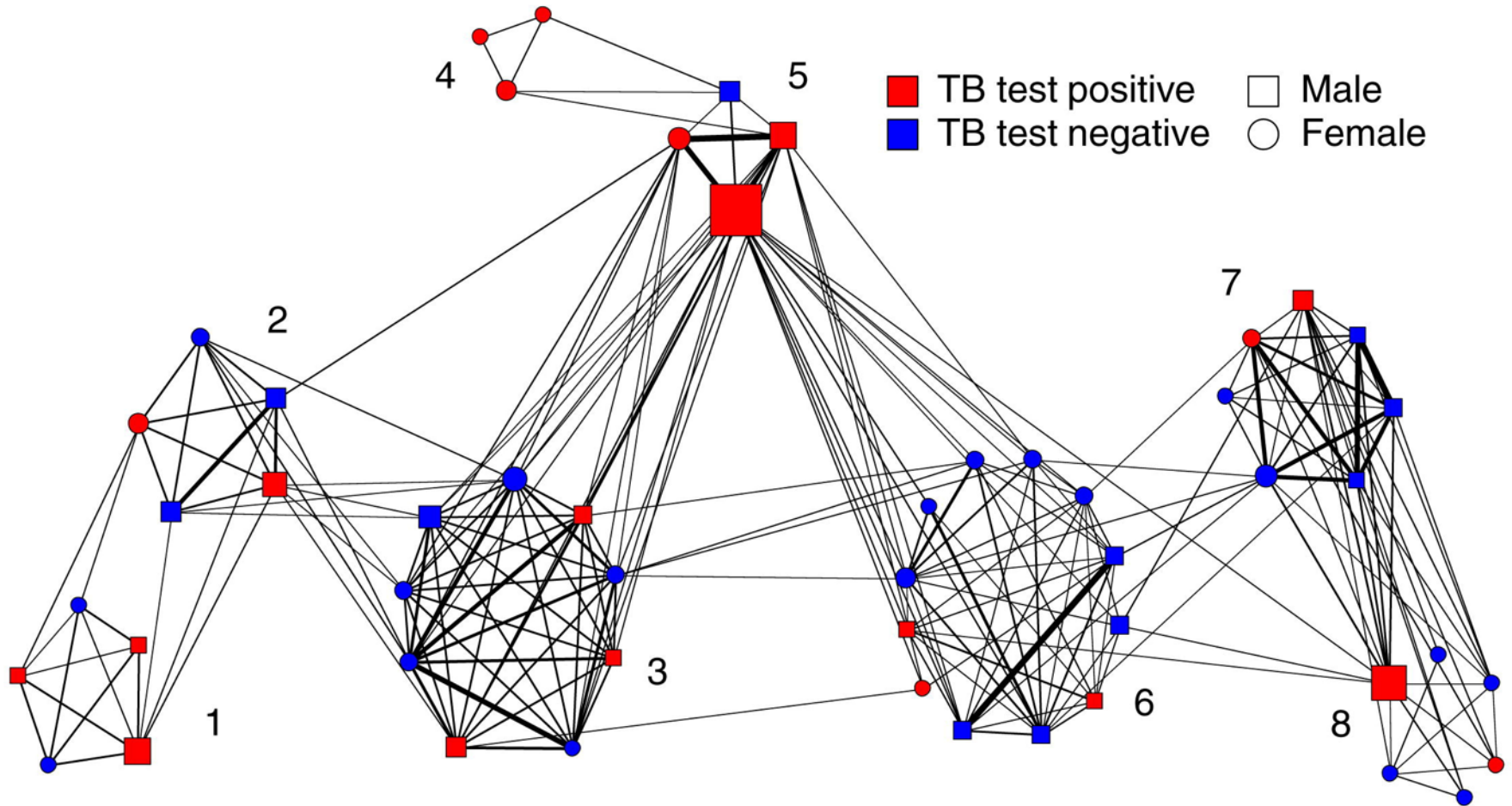Friends network (undirected)
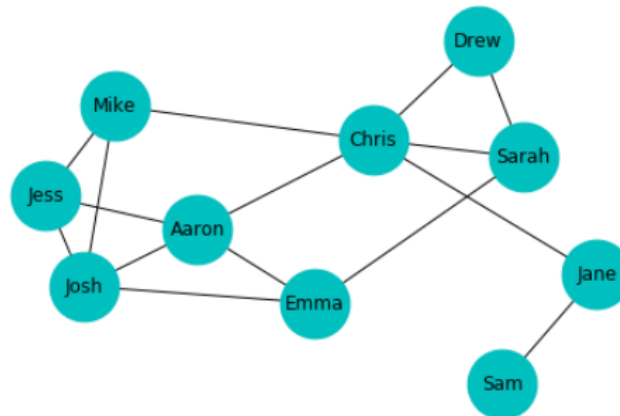
Example: Flight Routes
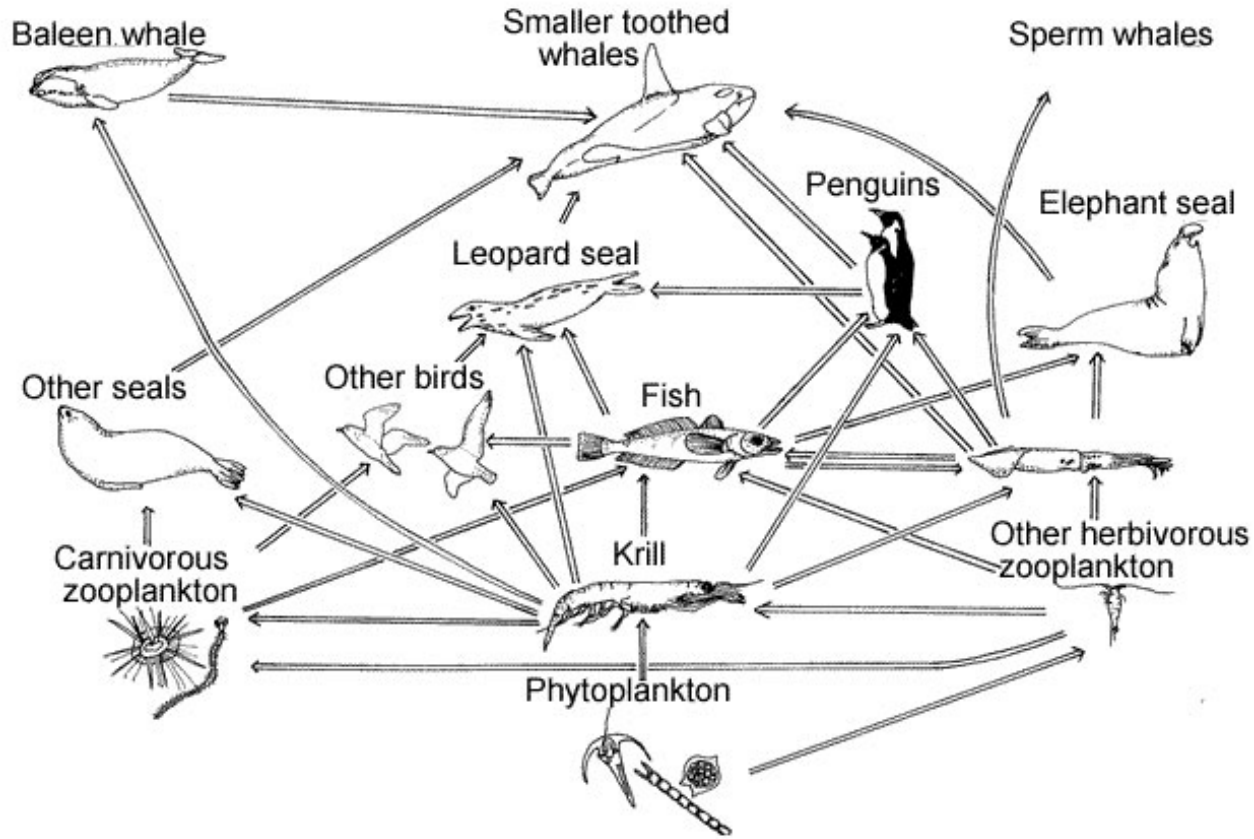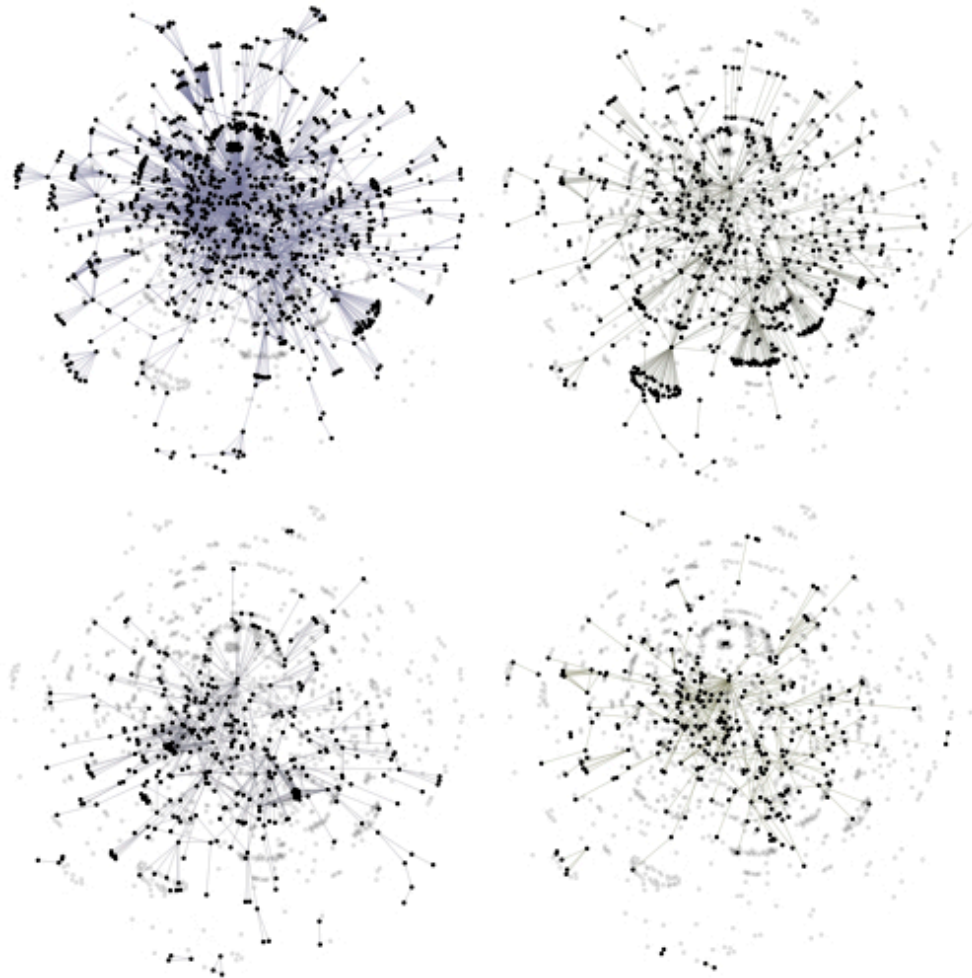
Example: Flight Routes

Example: Flight Routes
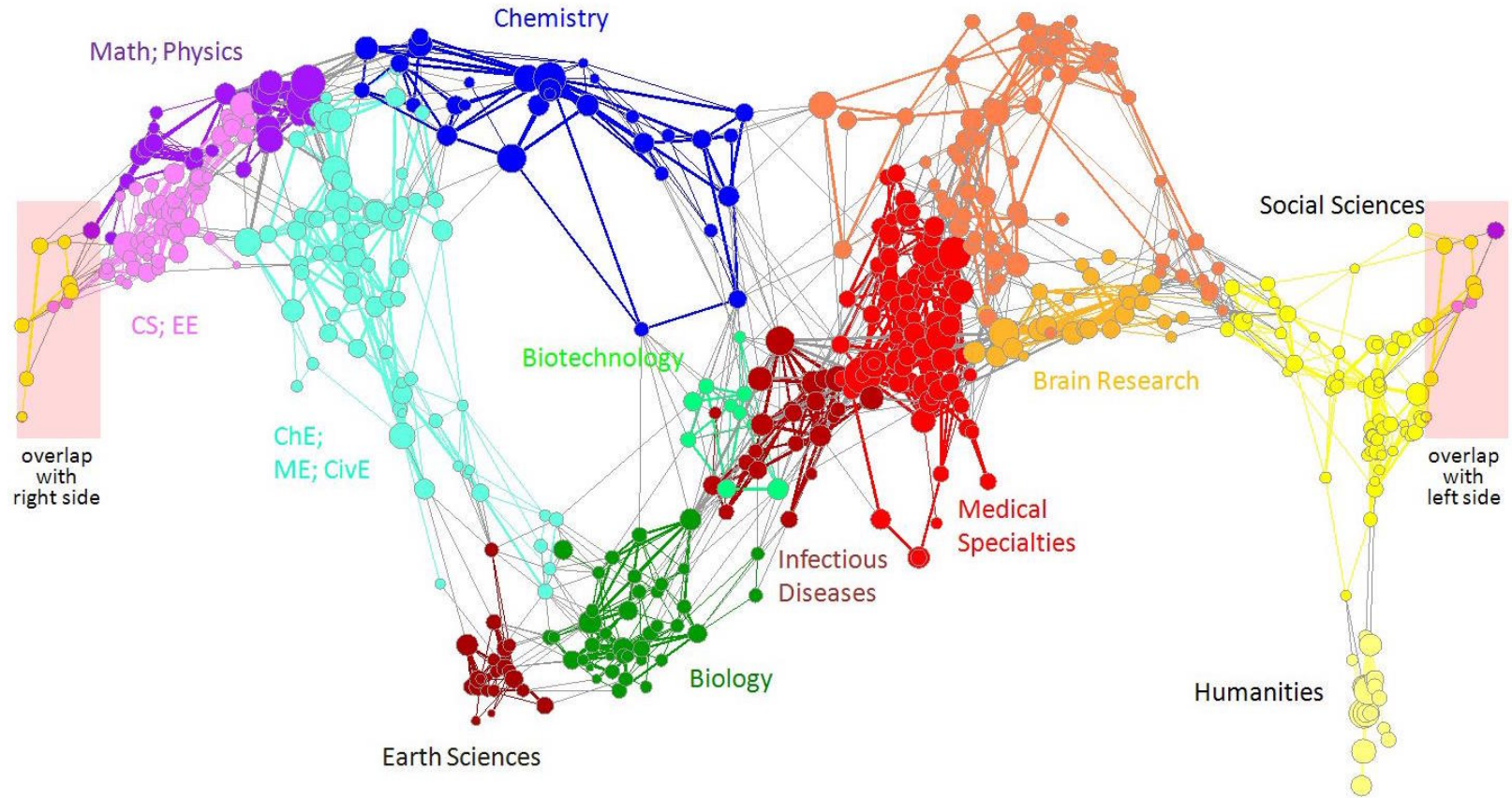
Example: Disease Transmission
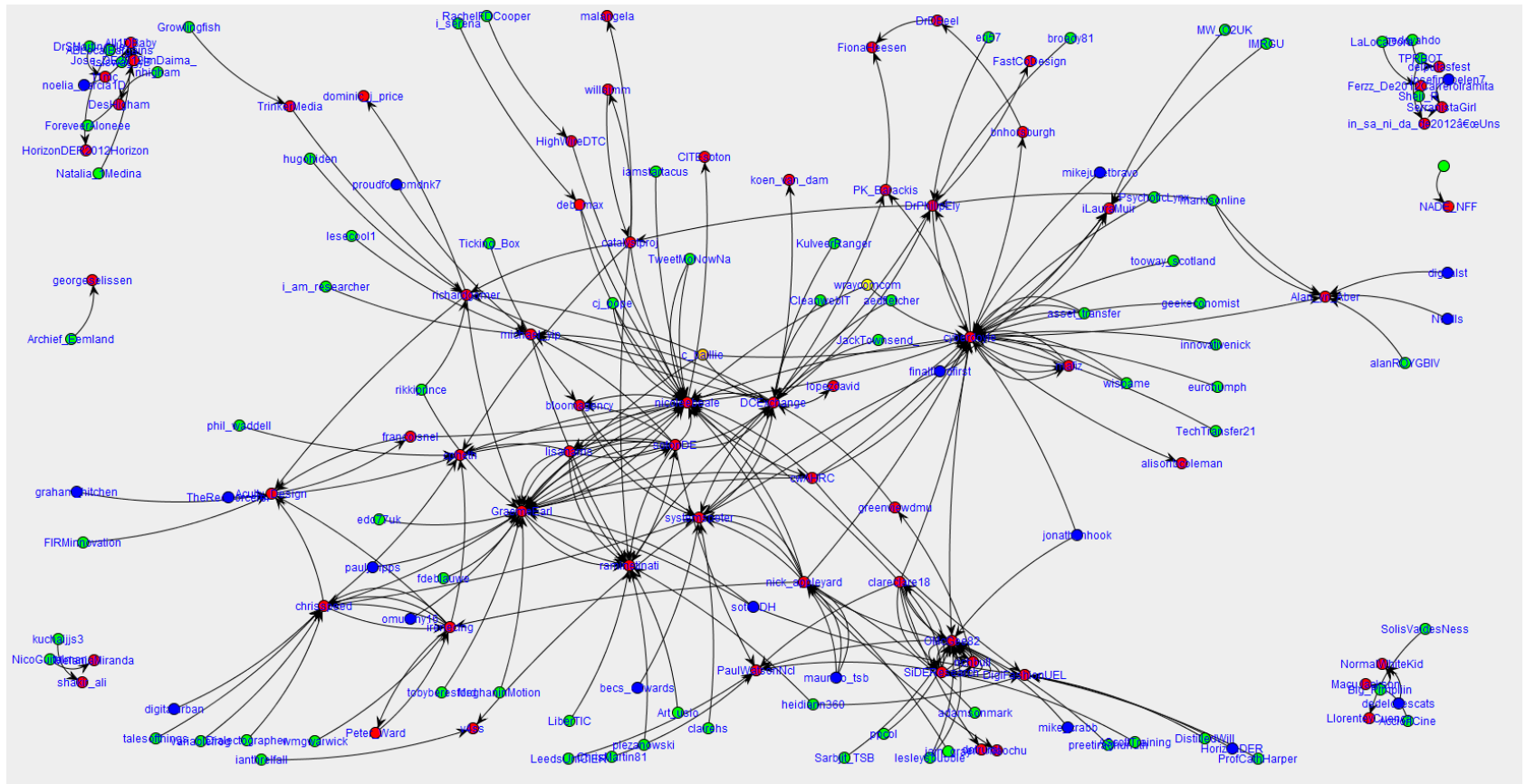
Example: Food Chain

Example: Criminal Networks

Example: Science Citations

Example: Retweets

Example: Facebook Friends

Other Examples

- Electricity grid + other civil infrastructure

- The brain + other biological structures

- Organizations and organizational behavior

- Spread of memes, other social phenomena

- And many, many more …

Network Analysis

Properties specific to graph structure

- **Basic Data Manipulation and Analysis**
  Asking well-defined questions

- **Data Mining**
  Looking for patterns

  Today:
  a few examples

- **Machine Learning**
  Building models, making predictions
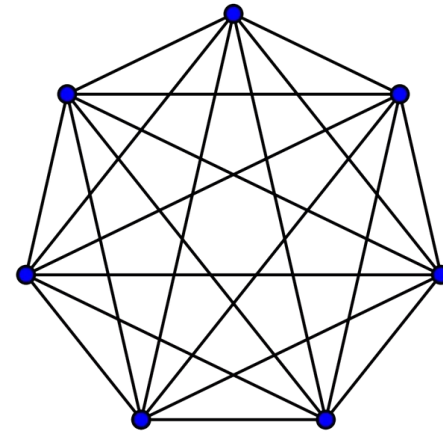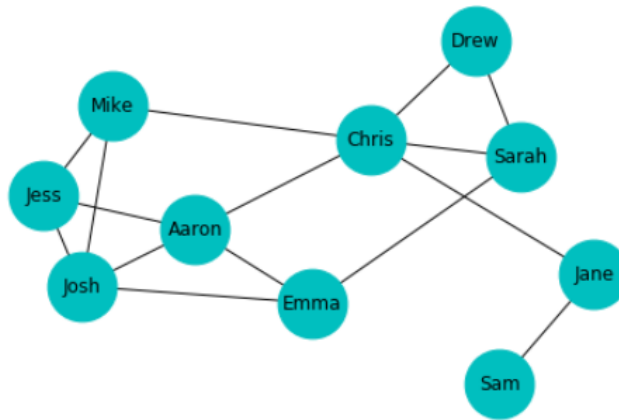
- **Data Visualization**
  Graphical depiction

- **Data Collection and Preparation**

Properties of Undirected Graphs

Density of graph

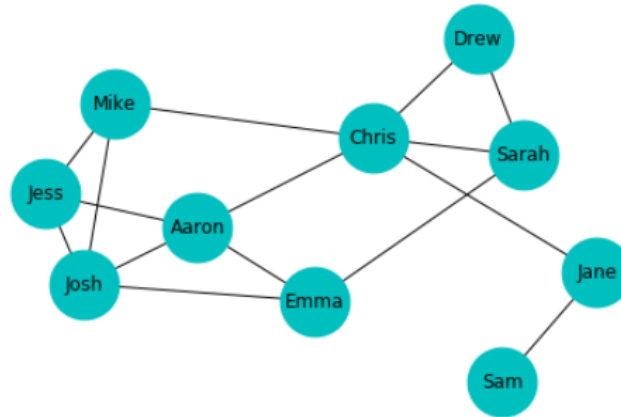$$\frac{\text{\# of edges}}{\text{\# of possible edges}}$$

Properties of Undirected Graphs

Shortest paths in graph

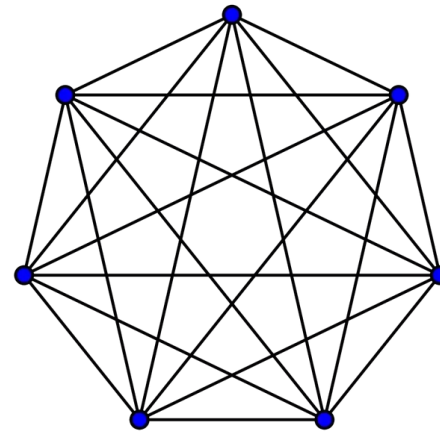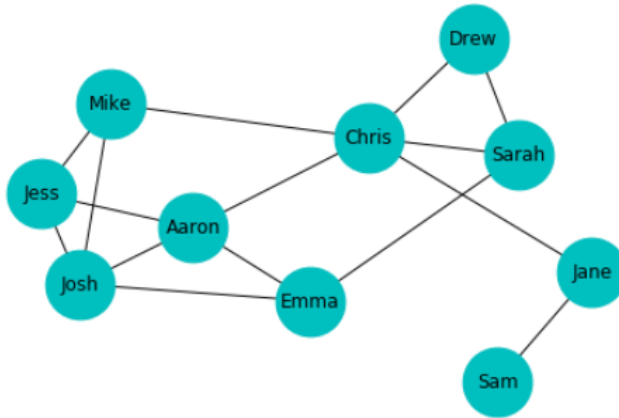*Shortest distance between given pair of nodes*

"Six degrees of separation"
(Four in Facebook)
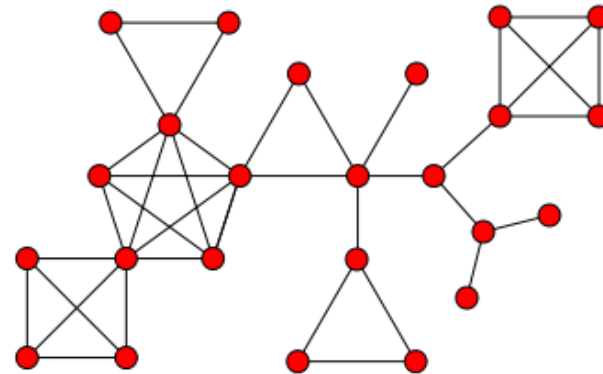
Properties of Undirected Graphs
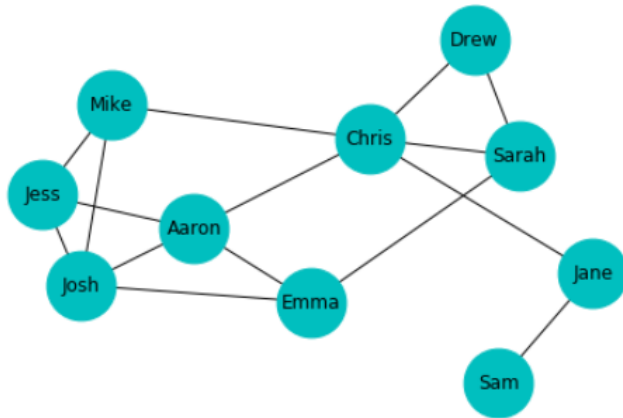
Diameter of graph

Maximum shortest path in graph

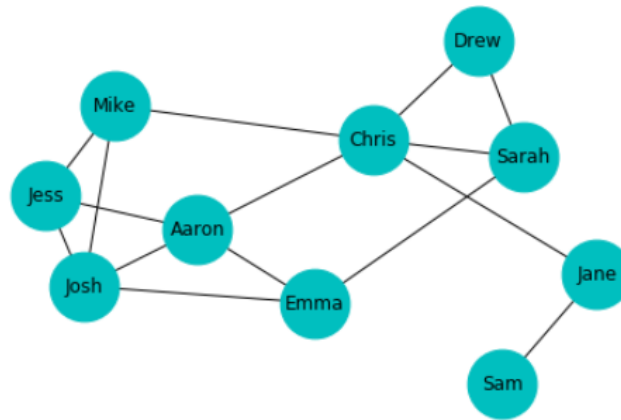# Properties of Undirected Graphs

Cliques in graph

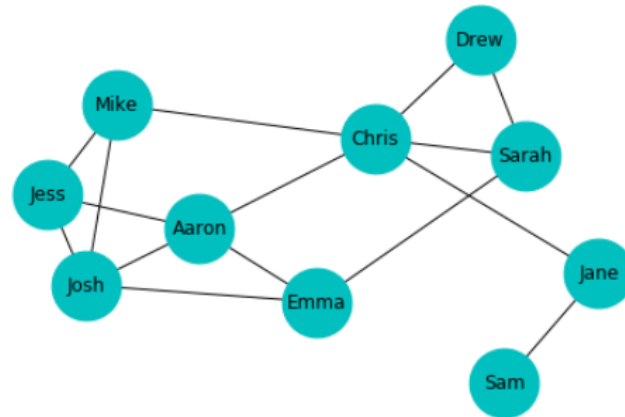*Sets of fully-connected nodes*

Properties of Undirected Graphs

Closeness centrality of a node in a graph

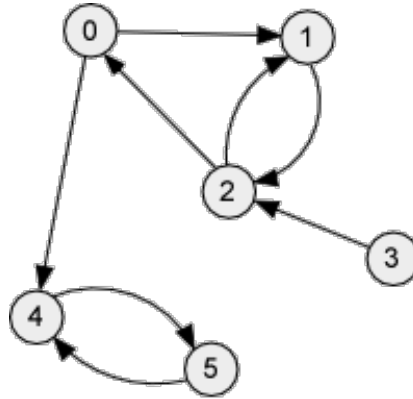*Average shortest distance to all other nodes (inverted so higher is "better")*

Properties of Undirected Graphs

Betweenness centrality of a node in a graph
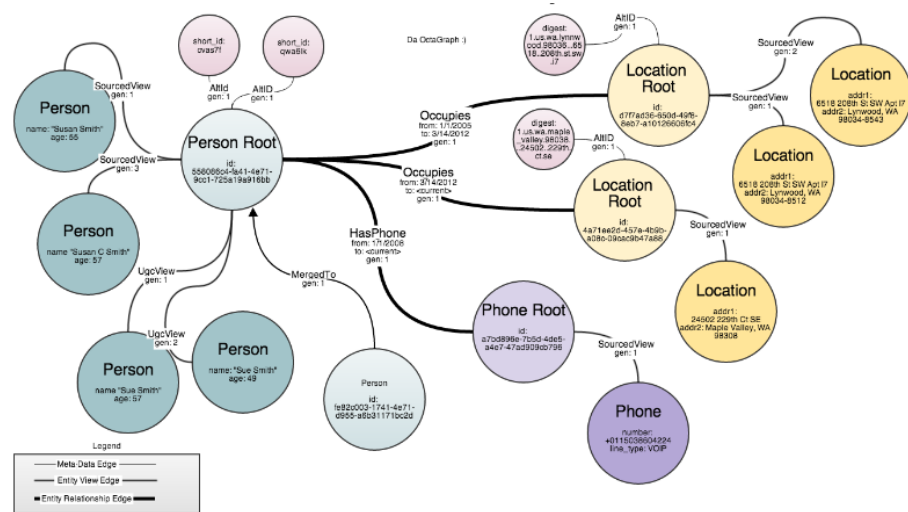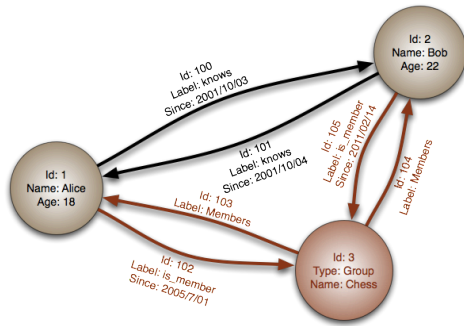
Number of shortest paths the node lies on

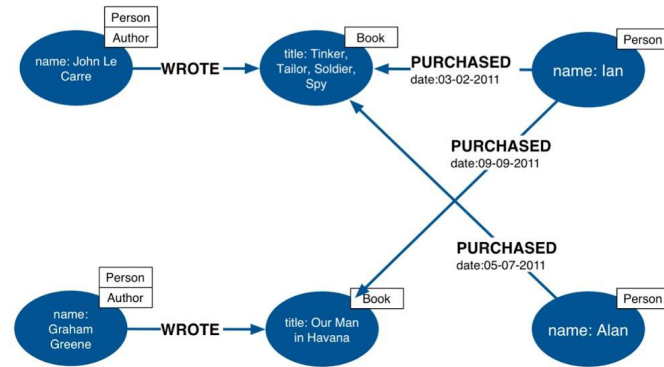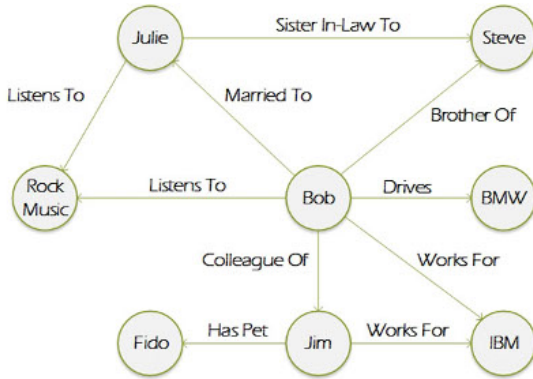# Directed Graphs

- In-degree - How many "followers"

- Out-degree - How many "following"

- Reciprocity - How often links are bidirectional

- Cycles

Labeled Graphs
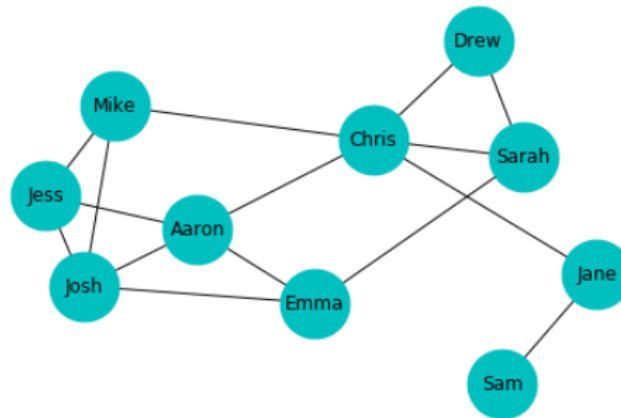
Other Analyses

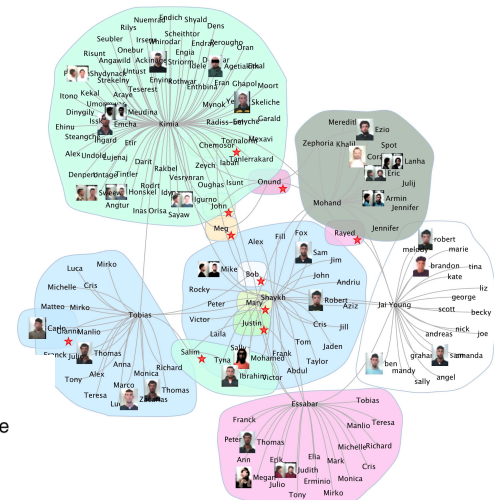## "Link Prediction"
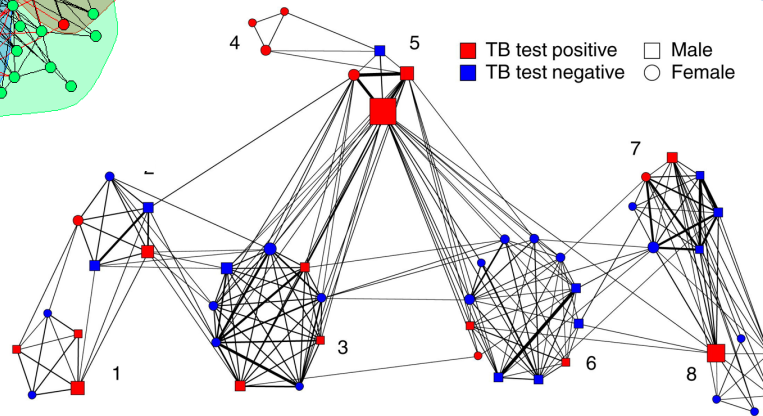
Predict future edges added to the graph
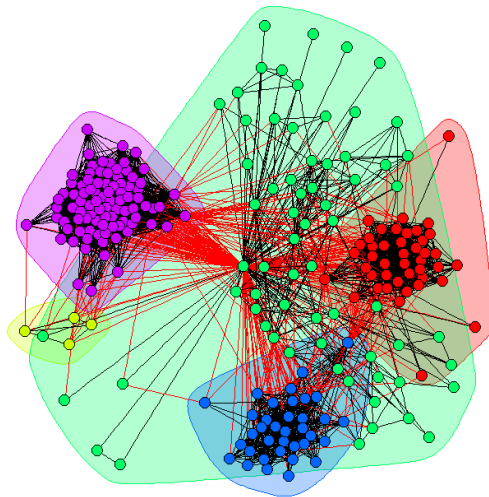

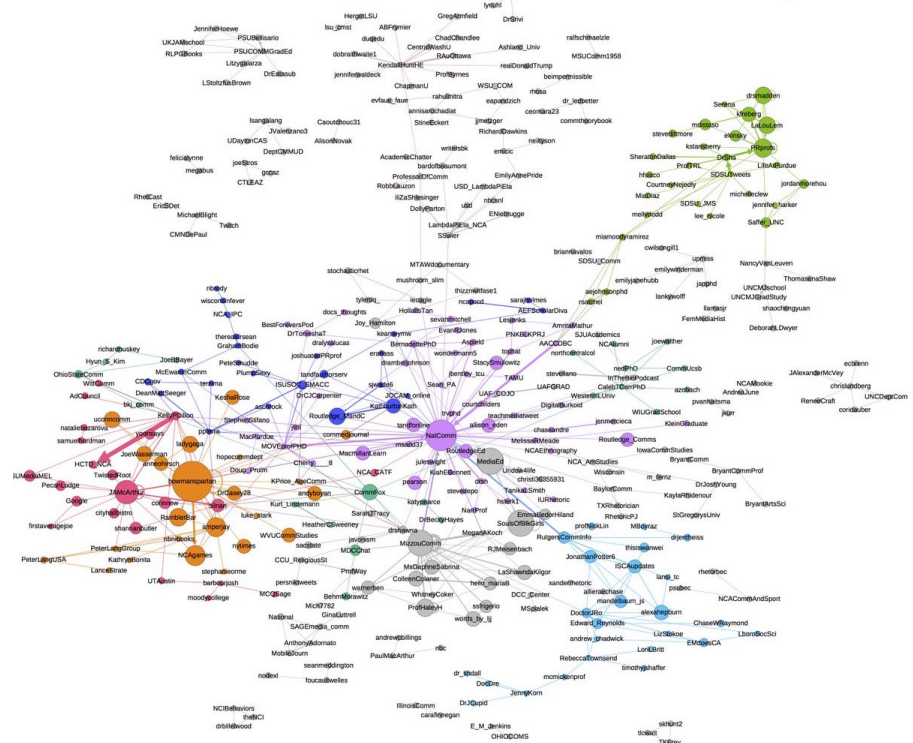
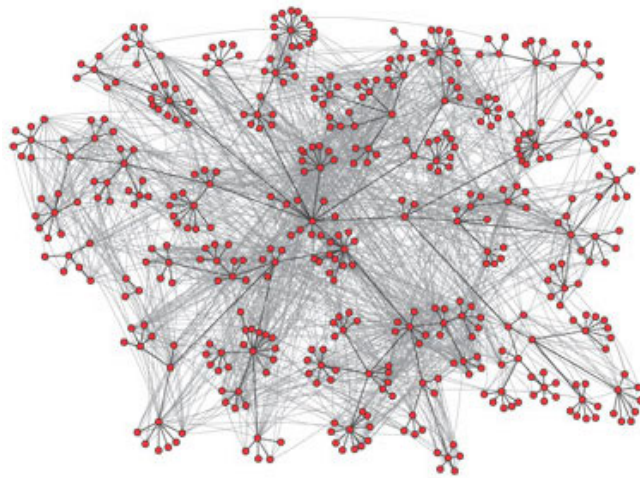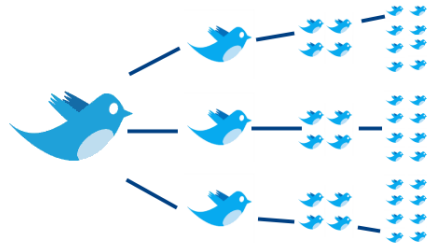Friends (or Follows) recommendations

"Community Detection"

Sets of interlinked/similar nodes

Other Analyses

"Cascades" - *Information propagation*

Hands-On Network Analysis

- ■ Datasets
  - • Tiny "friends" network (undirected)
  - • Tiny "follows" network (directed)
  - • Dolphin associations (assignment)

- ■ Python networkx package