

# Analysis and Visualisation with NetworkX and Altair

## Week 1: NetworkX

30 November 2020



THE UNIVERSITY of EDINBURGH  
Centre for Data, Culture & Society



# WELCOME

**REMEMBER TO MUTE YOURSELF IF YOU ARE NOT SPEAKING**

Points to remember:

- Use the chat function on the right to ask questions.
- If you would prefer to ask a question aloud, click the 'raise hand' icon on the right before un-muting your mic.
- Headphones can also help to reduce background noise.
- If you are having difficulties with participating, you can email [cdcs@ed.ac.uk](mailto:cdcs@ed.ac.uk) to let us know.

# Course Structure

Anticipate about ~7 hours/week

- 2 course meetings per week
  - 10:00 - 11:00 AM BST Mondays
  - 10:00 - 11:00 AM BST Fridays
- 1 assignment per week ~2 hours
- Office hours on Wednesdays for 30 minutes per participant
- Independent learning ~2 hours

**Teams** for introductions, meetings, office hours, questions, files

# Course Topics

Week 1: NetworkX for network analysis and visualisation

Week 2: Altair for data visualisation

# Instructor Introduction

- Pursuing a PhD in the School of Informatics ILCC
- AMSc Design Informatics, B.S. Information Systems
- Taught myself programming and data science skills outside courses using online resources
- Please share feedback on the course!

# For Participants

- Introduce material for you to review in greater depth on your own
- I'll direct you to further resources if you'd like to go beyond material covered in each week's assignment
- Course meetings won't be recorded
  - Three strike policy
  - Please let me know in advance if you cannot attend!
- Office hours: questions about assignments, your own projects
  - Chat with me on Teams to schedule

# For Participants

- During class meetings:
  - Take notes
  - Comment and ask questions
  - *Don't worry about writing the code I demo!*
- On your own:
  - Type your own code
  - *Avoid copying and pasting!*



# Getting Set Up

You will need:

- A. Python 3
- B. Pip/pip3 or conda
- C. NetworkX (along with its complementary packages)
- D. Jupyter Notebooks



# Getting Set Up: Python 3

If you use Anaconda or Miniconda to run Jupyter Notebooks, you already

*Note: You can have both Python 2.x and Python 3.x installed!*

# Getting Set Up: Pip or Conda

We'll be using Jupyter Notebooks

If you haven't already, you'll need to install:

- A. Python 3
- B. Pip or pip3
- C. NetworkX (along with its complementary packages)
- D. [Optional] Jupyter Notebooks <https://jupyter.org/install>

# Getting Set Up: NetworkX

If you're using Anaconda for Jupyter Notebooks:  
you already have NetworkX!

If you're using Miniconda for Jupyter Notebooks, in a Terminal/Shell run:

```
conda install networkx
```

Otherwise, in a Terminal/Shell run one of the following:

```
pip install networkx[all]  
Pip3 install networkx[all]
```

Reference: <https://networkx.org/documentation/stable/install.html>

# Getting Set Up: Jupyter Notebooks

We'll be using Jupyter Notebooks

3 options:

- A. MyBinder
- B. Using Miniconda or Anaconda
- C. Install to your computer: <https://jupyter.org/install>



# Network Analysis

Social relationships

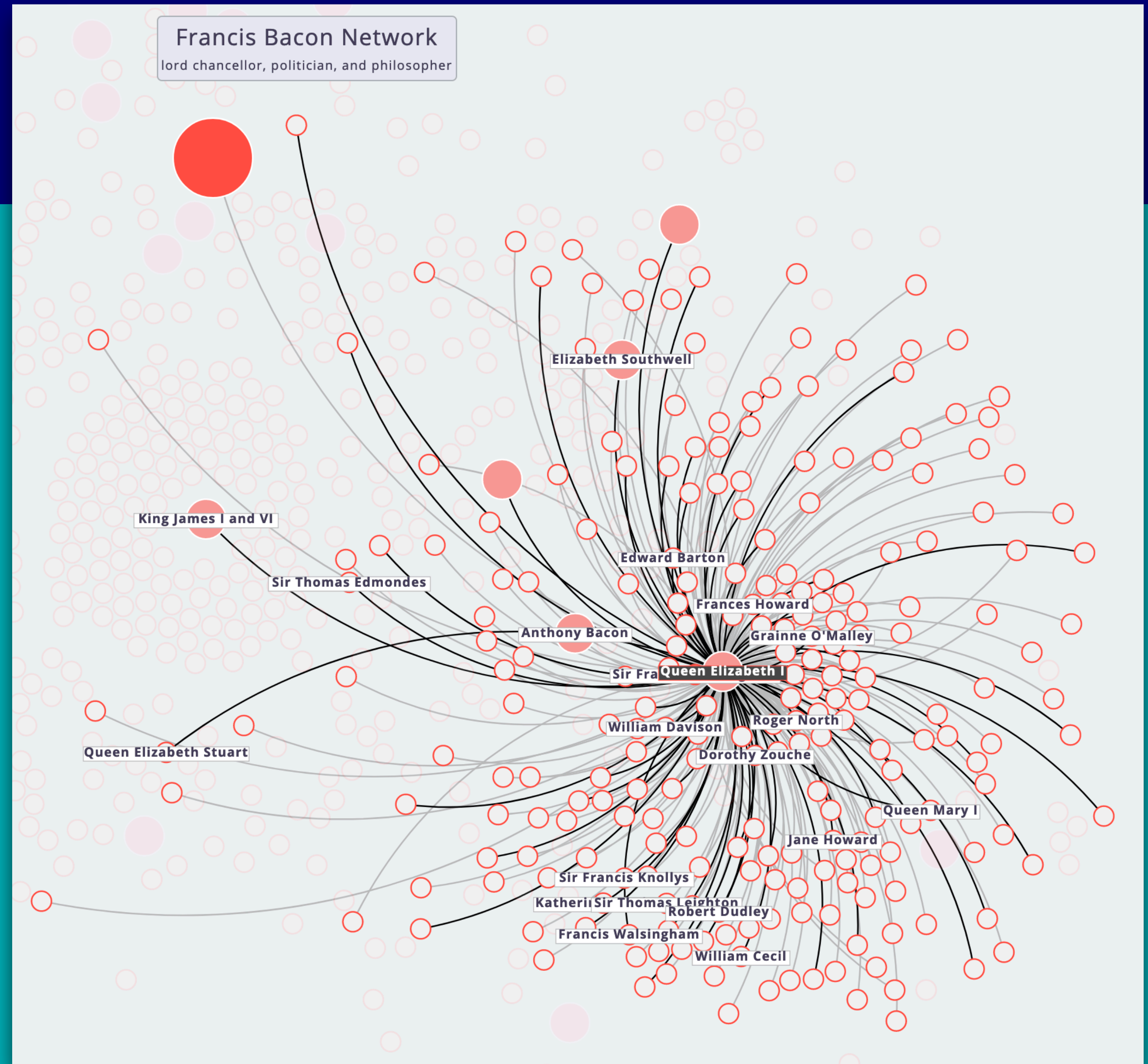
Who knows who?

How do ideas travel?

SNA: social network analysis

Example:

Six Degrees of Francis  
Bacon



Reference: [http://www.sixdegreesoffrancisbacon.com/?ids=10000473&min\\_confidence=60&type=network](http://www.sixdegreesoffrancisbacon.com/?ids=10000473&min_confidence=60&type=network)

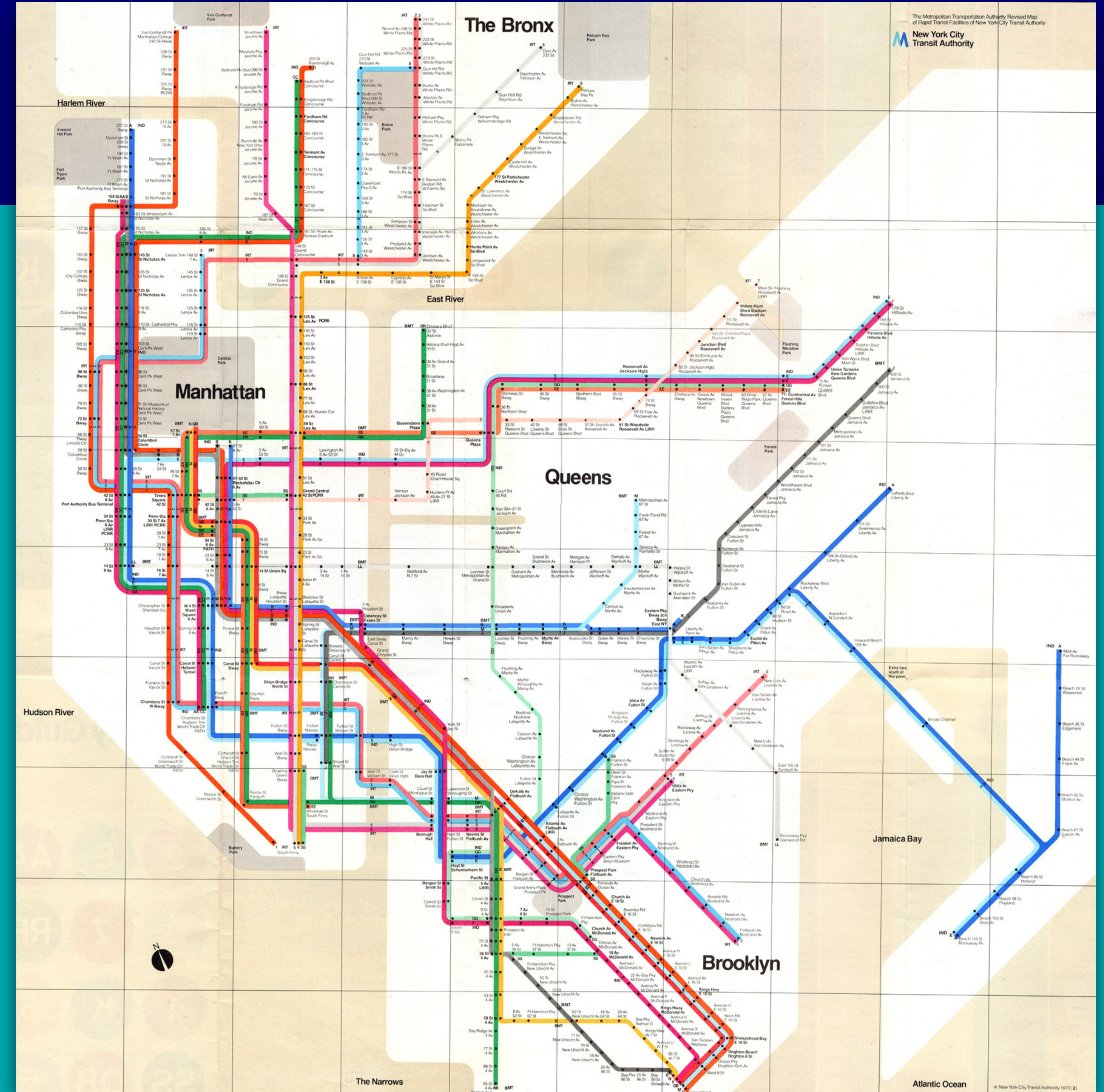


# Network Analysis

## Transportation

How can I get from  
point A to point B?

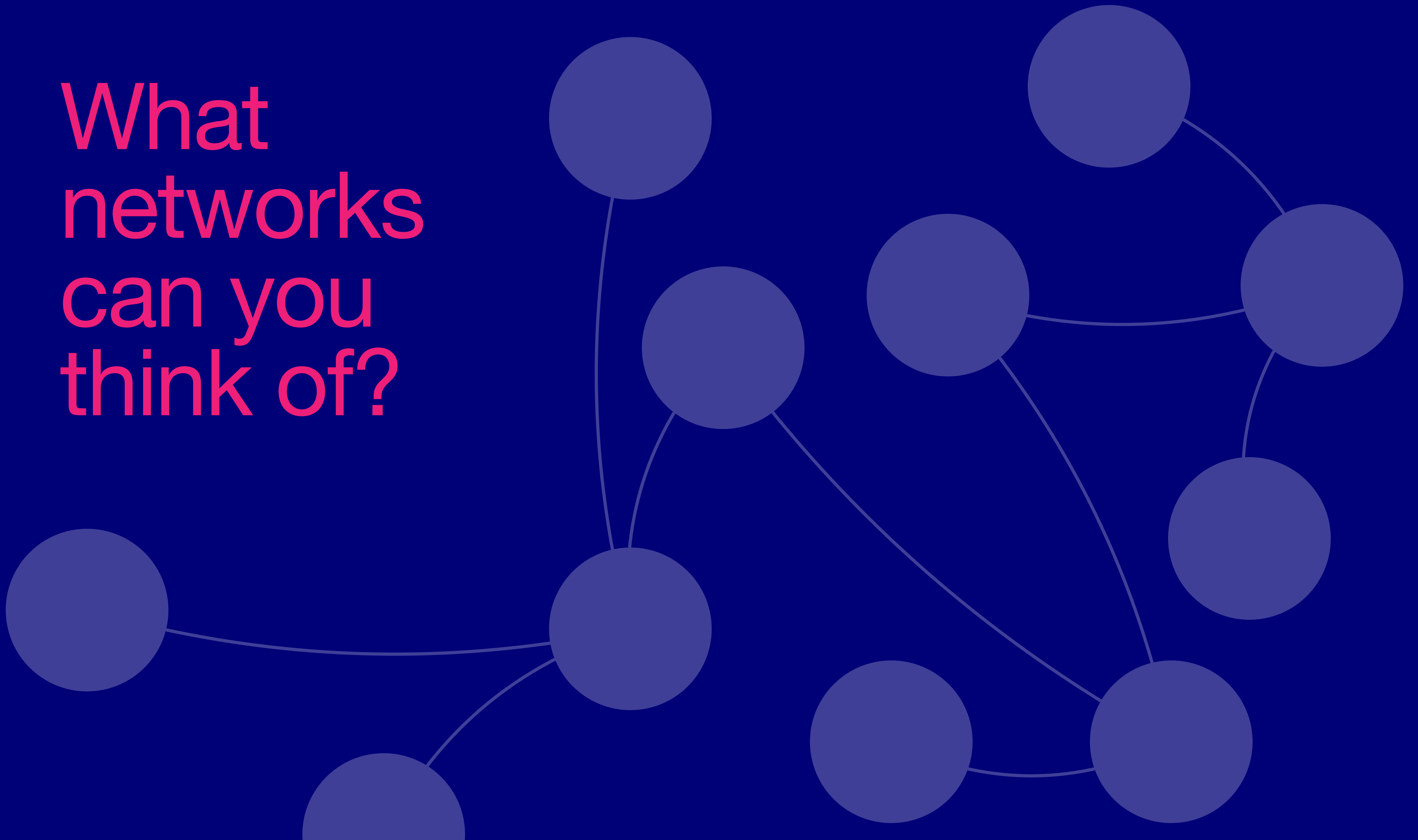
Example:  
New York City subway map



Reference: <http://www.visualcomplexity.com/vc/project.cfm?id=266>



What  
networks  
can you  
think of?





# Network Analysis

Networks, or **graphs**, contain:

1. **Nodes** (vertices, actors): entities, such as people or places
2. **Edges** (ties, relations): relationships, such as friendships or roads

Networks represent data that are interdependent

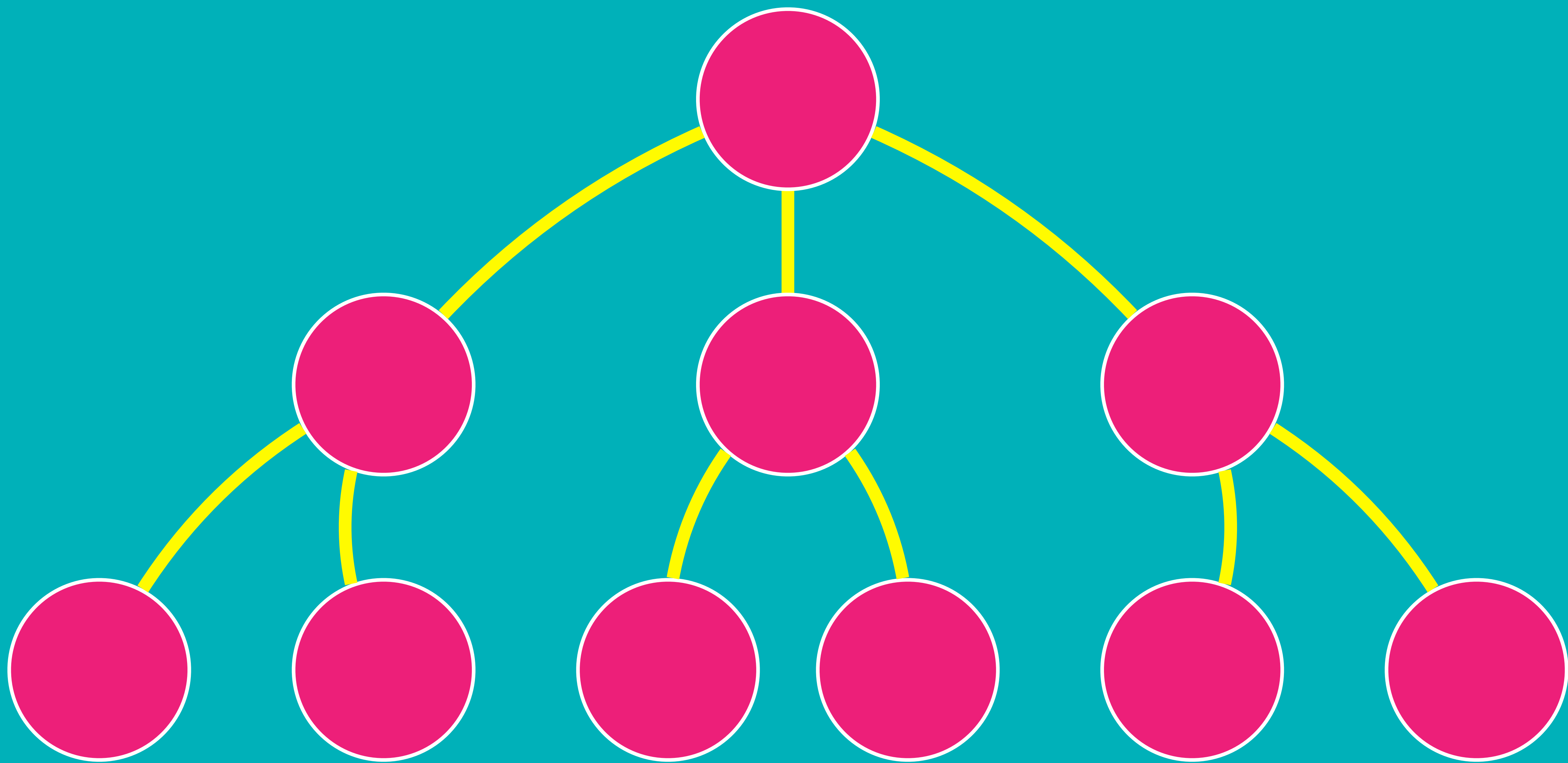
Reference: <http://www.scottbot.net/HIAL/index.html@p=6279.html>

# Network Analysis

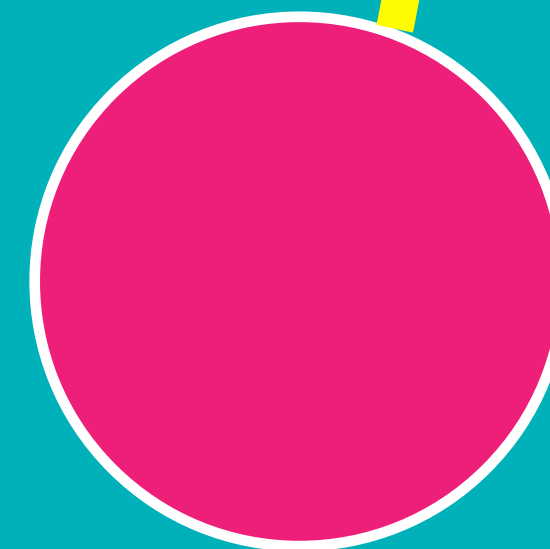
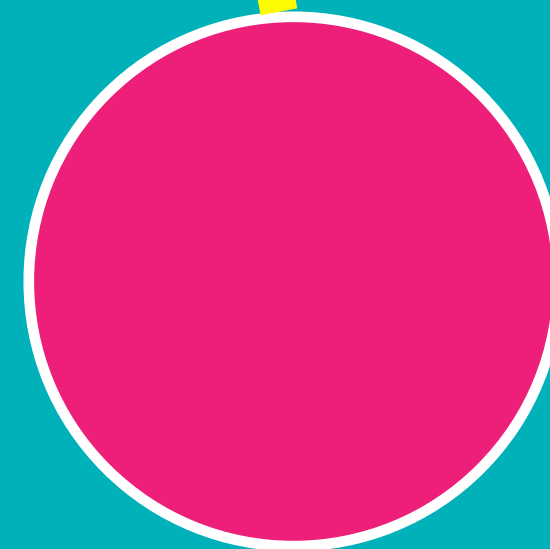
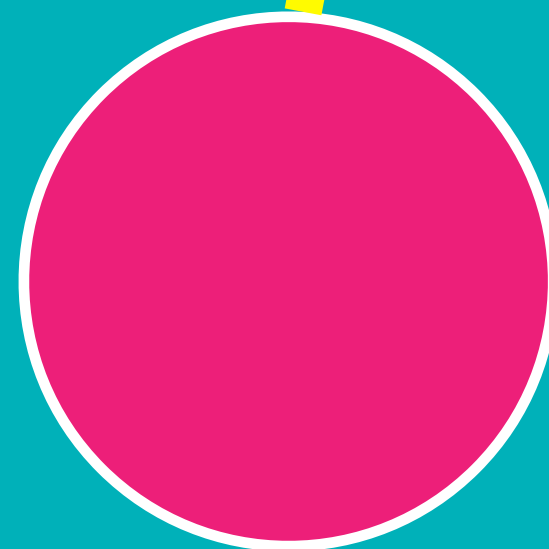
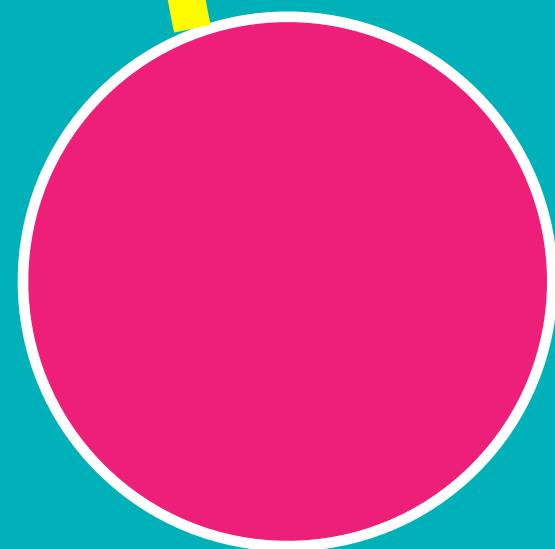
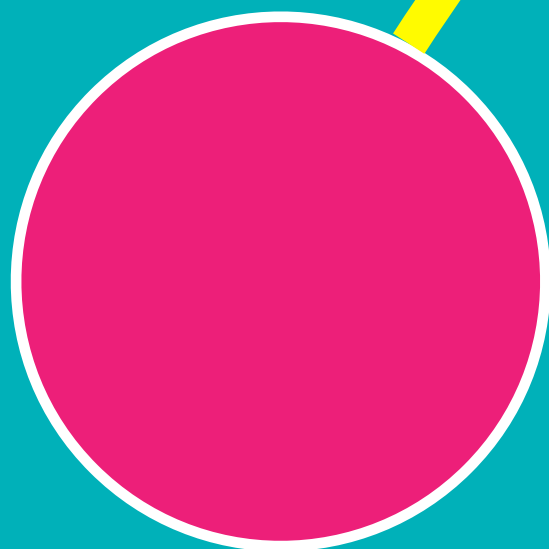
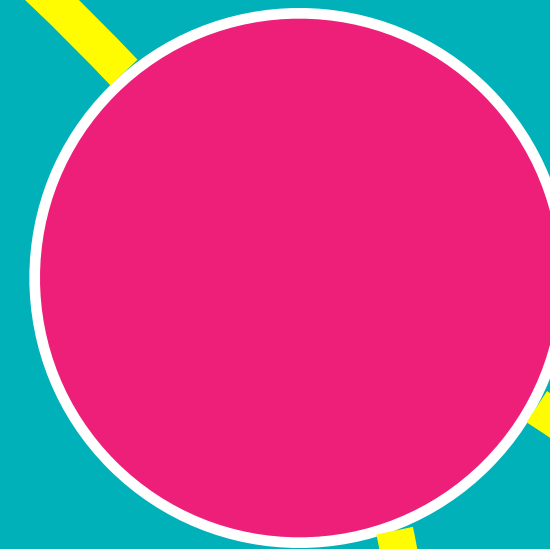
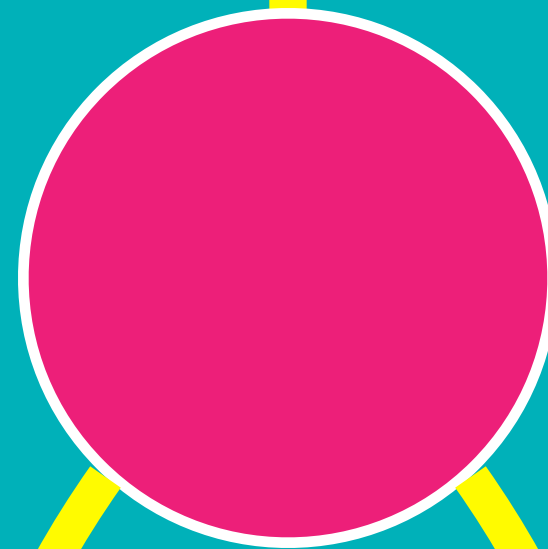
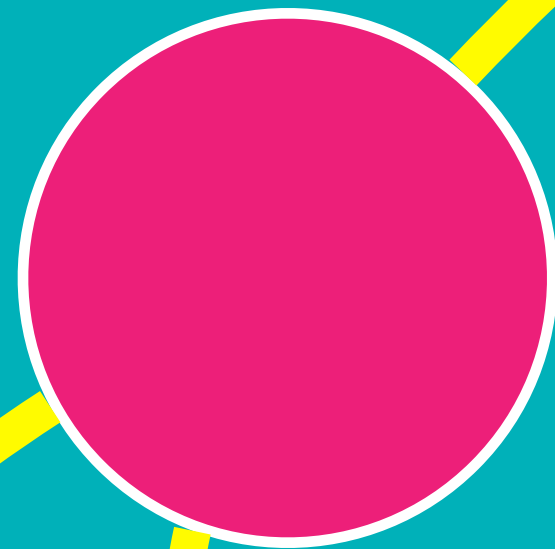
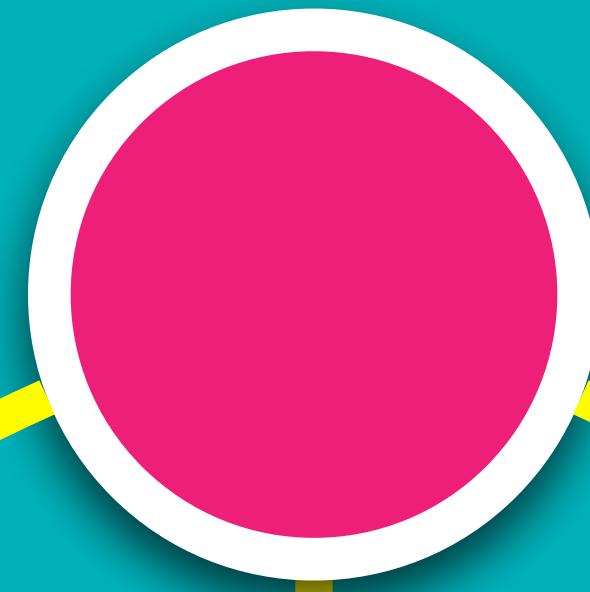
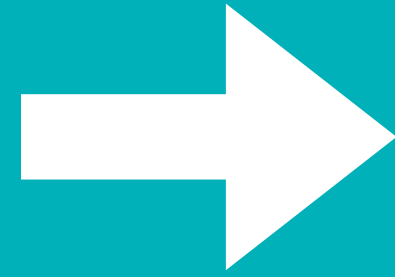
A graph's **edges** may have several characteristics:

1. Directed or undirected
2. Attributes
3. Weights

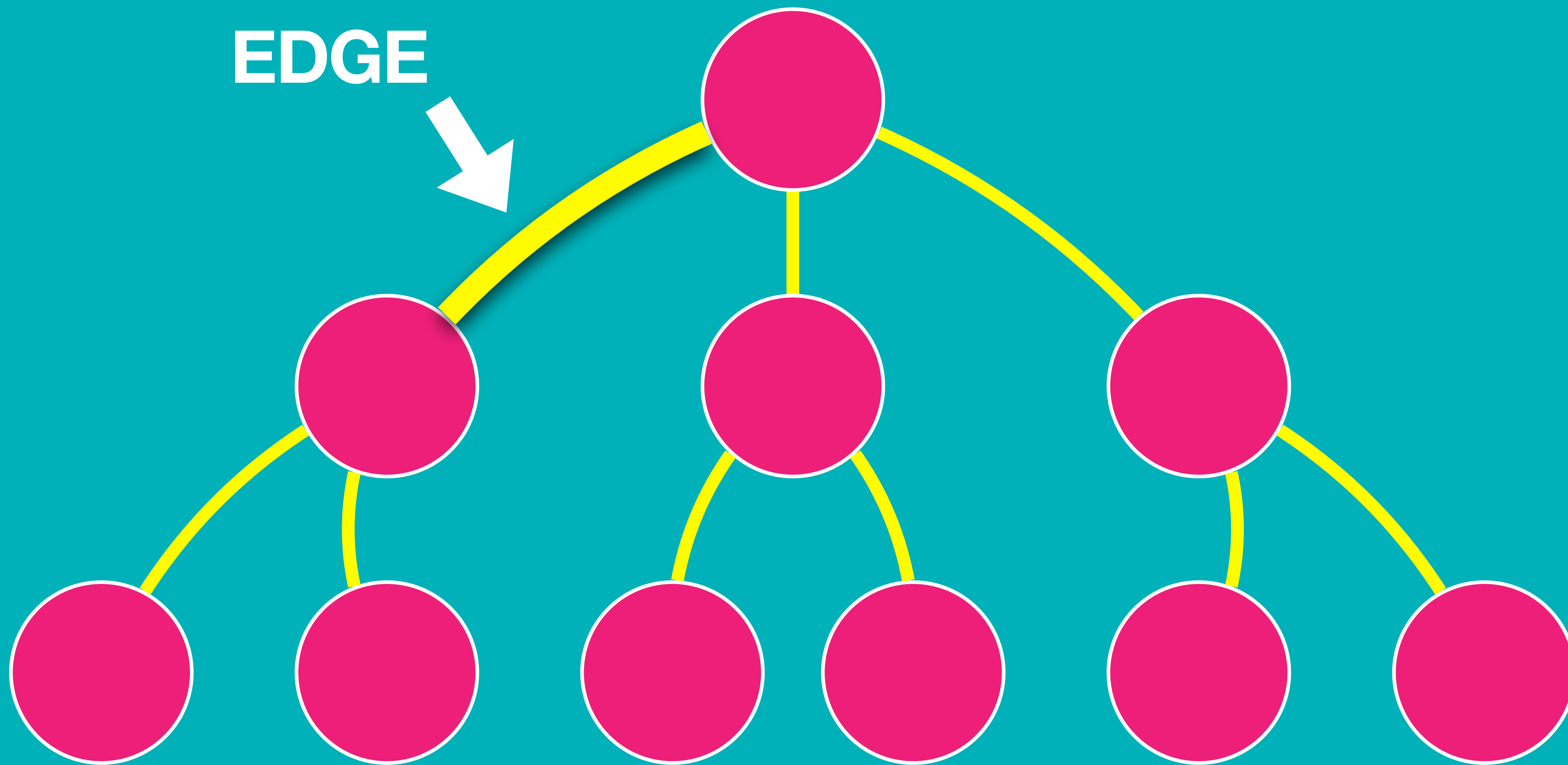
Reference: <http://www.scottbot.net/HIAL/index.html@p=6279.html>



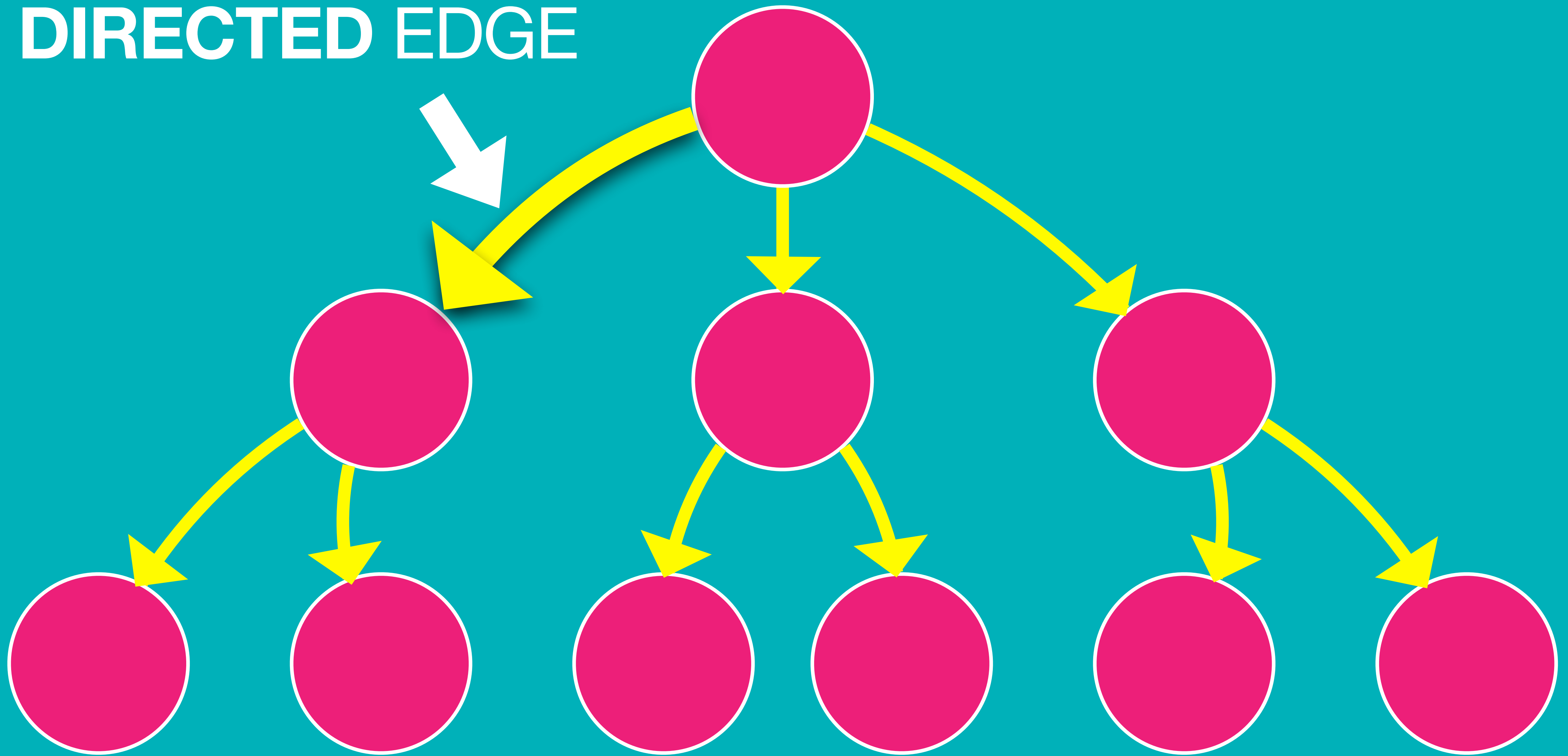
NODE



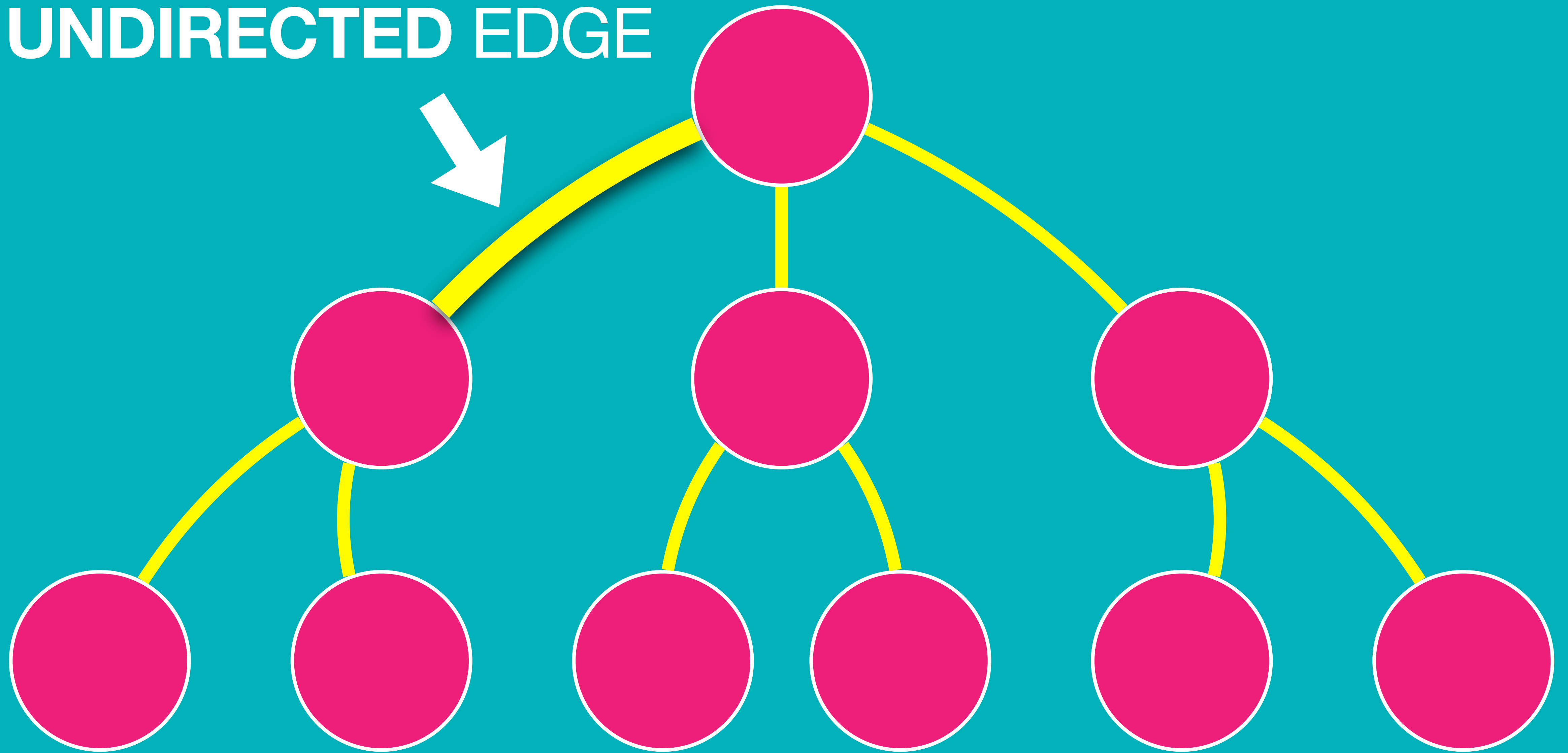
EDGE



**DIRECTED EDGE**

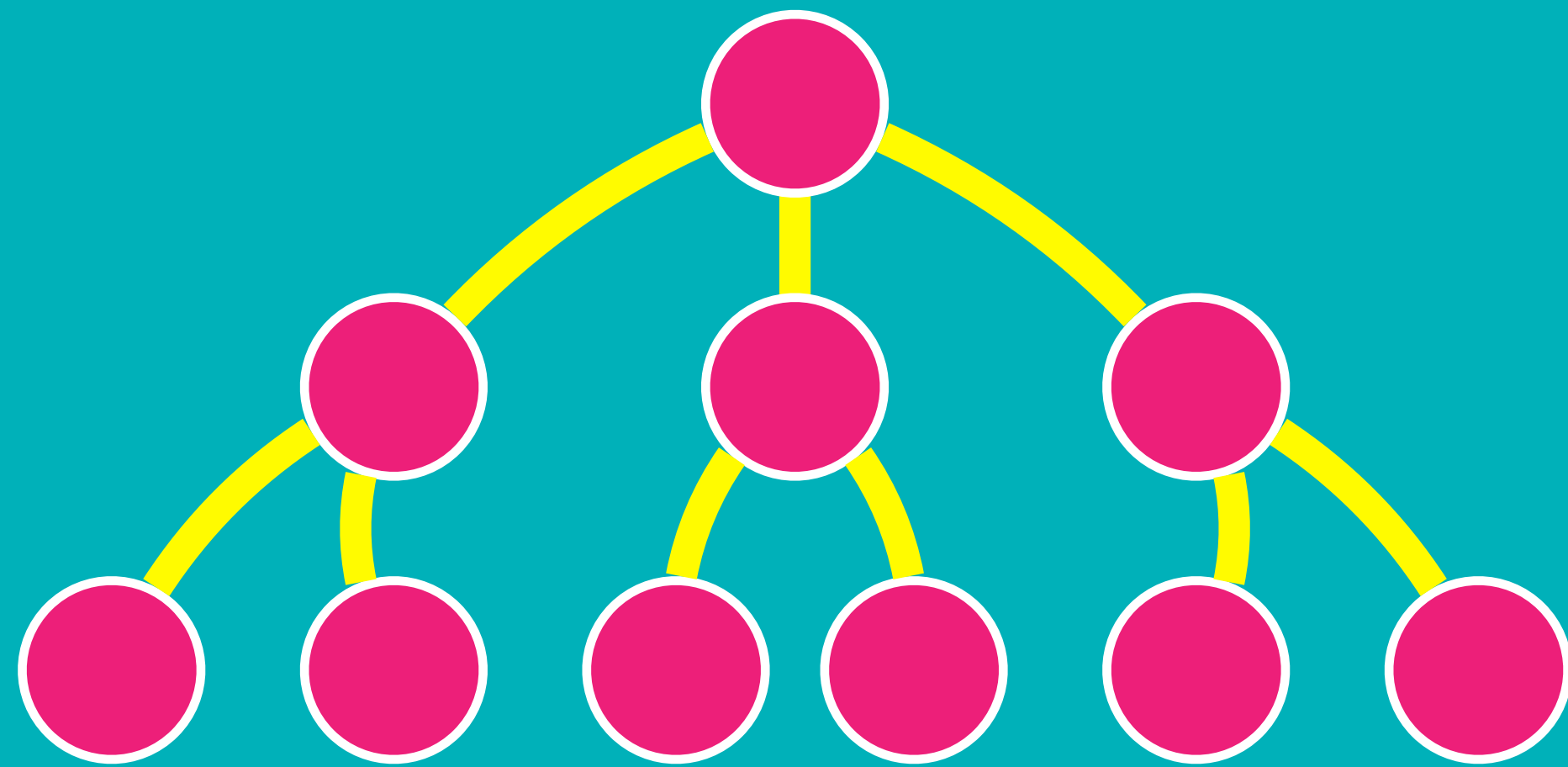


UNDIRECTED EDGE



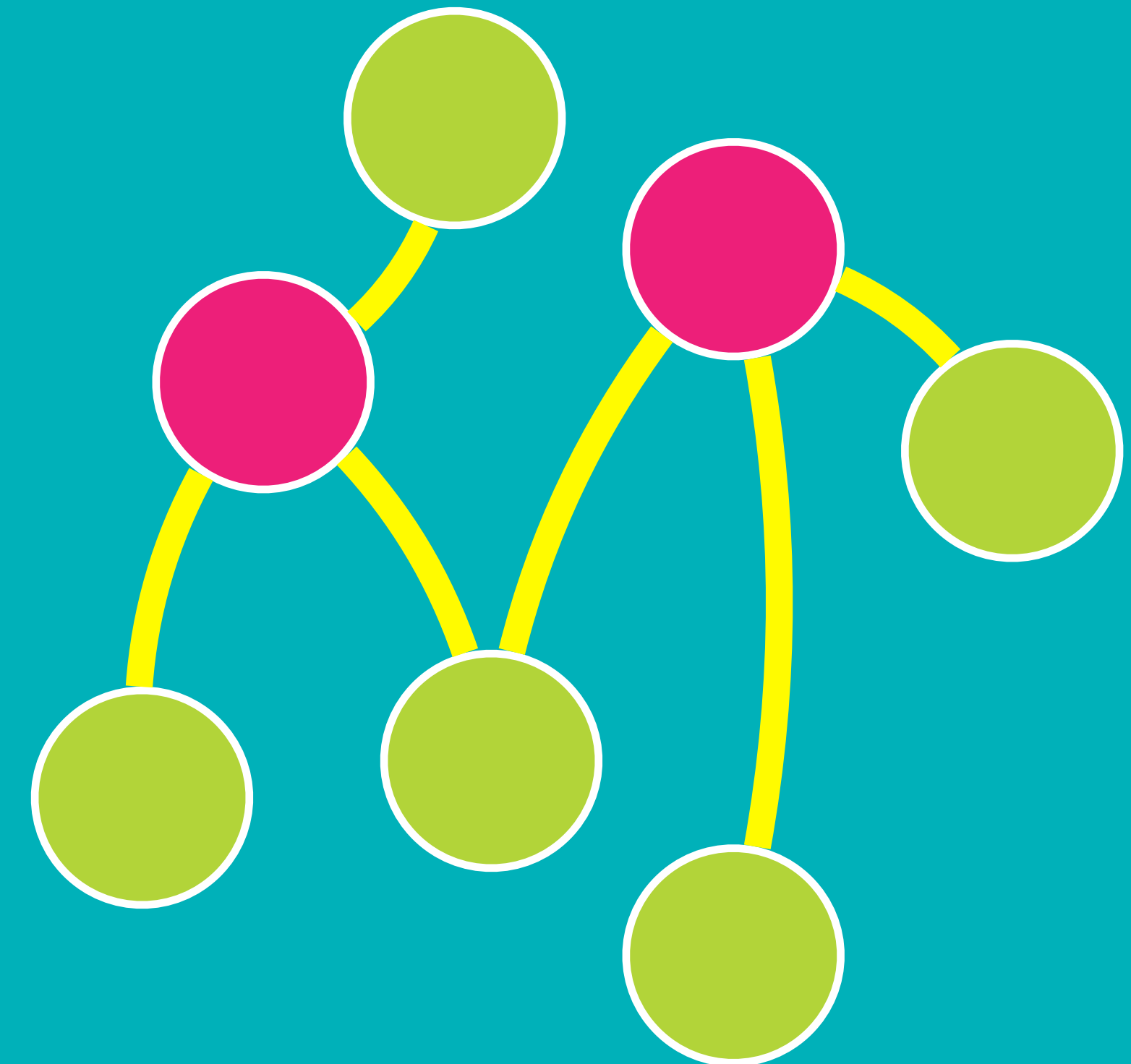


## hierarchical



*It's best to keep it simple, with 1,  
or if you must, 2, types of nodes*

## bimodal or bipartite



# NetworkX

**NetworkX** is a Python library for creating and analyzing networks

To create a network, you need two data files:

1. Edges file with 2 columns for source and target nodes
2. Nodes file with columns node names and any attributes of those nodes

Reference: <https://programminghistorian.org/en/lessons/creating-network-diagrams-from-historical-sources>

**DEMO**

# NetworkX: Creating a Network

1. Create a **list** of node names
2. Create a **list** of edges, where each list item is a tuple of two related nodes
3. Create **dictionaries** of the attributes of the nodes (one dictionary per attribute type)
4. Add those attributes to the Graph

Reference: <https://programminghistorian.org/en/lessons/creating-network-diagrams-from-historical-sources>

**DEMO**

# Assignment

Watch the videos in *Use Case 2: Discovering Collaboration*, in this LinkedIn Learning course:

<https://www.linkedin.com/learning/applied-ai-for-human-resources/organization-design?u=50251009>

**Follow along in your own Jupyter Notebook!**

Analyze the graph we created together in class, starting from the section *Metrics available in NetworkX* in this tutorial:

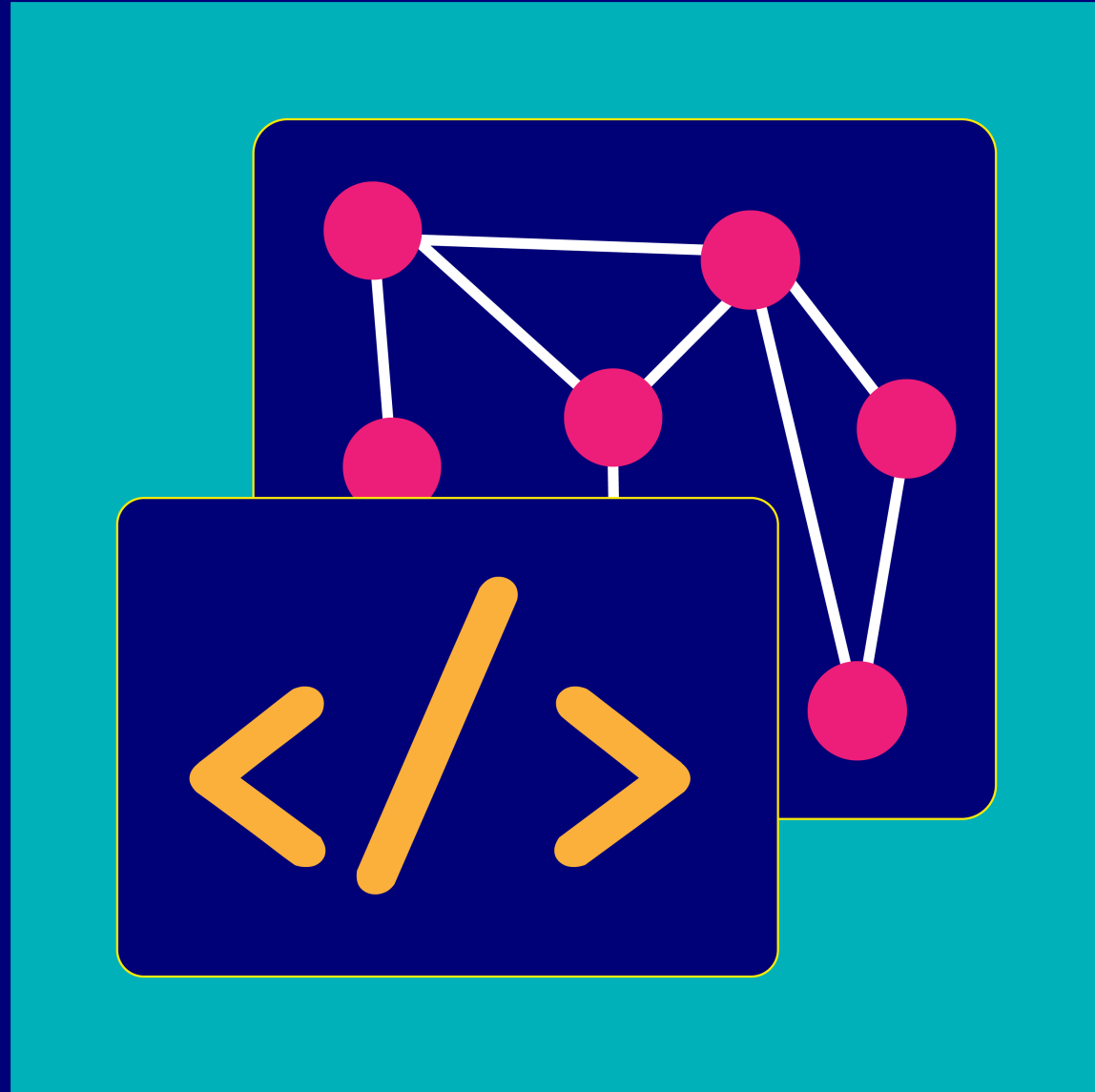
<https://programminghistorian.org/en/lessons/exploring-and-analyzing-network-data-with-python#fn:pipinstall>

**You can use the Notebook we demoed today!**

Volunteers  
to share your  
network analysis  
experience?







# Thanks everyone!

Next course meeting: Friday, 10:00-11:00 AM BST

Office hours available on Wednesday (30 minutes)

*To schedule, please message me on Teams!*