

Statistical Analysis of Networks

Network Description

Learning Goals

- ❖ At the end of the lecture, you should be able to answer this question:
 - ❖ How can I describe some simple features of a network?

Simple Network Summary

- ❖ How big is it? (Size)
- ❖ How tightly connected is it? (Density)
- ❖ How separated is it? (Components / Subgraphs)
- ❖ How compact is it? (Diameter)
- ❖ How much clustering is there? (Transitivity)

Size

- ❖ How big is it?
 - ❖ This is just the number of nodes in the network, g .

Density

- ❖ Of the ties that could exist, what fraction are observed?
- ❖ This is the *density* of a graph.
- ❖ That is, how many ties there are compared to how many ties there could be.

Density

- ❖ The density of an **undirected** graph is given by:

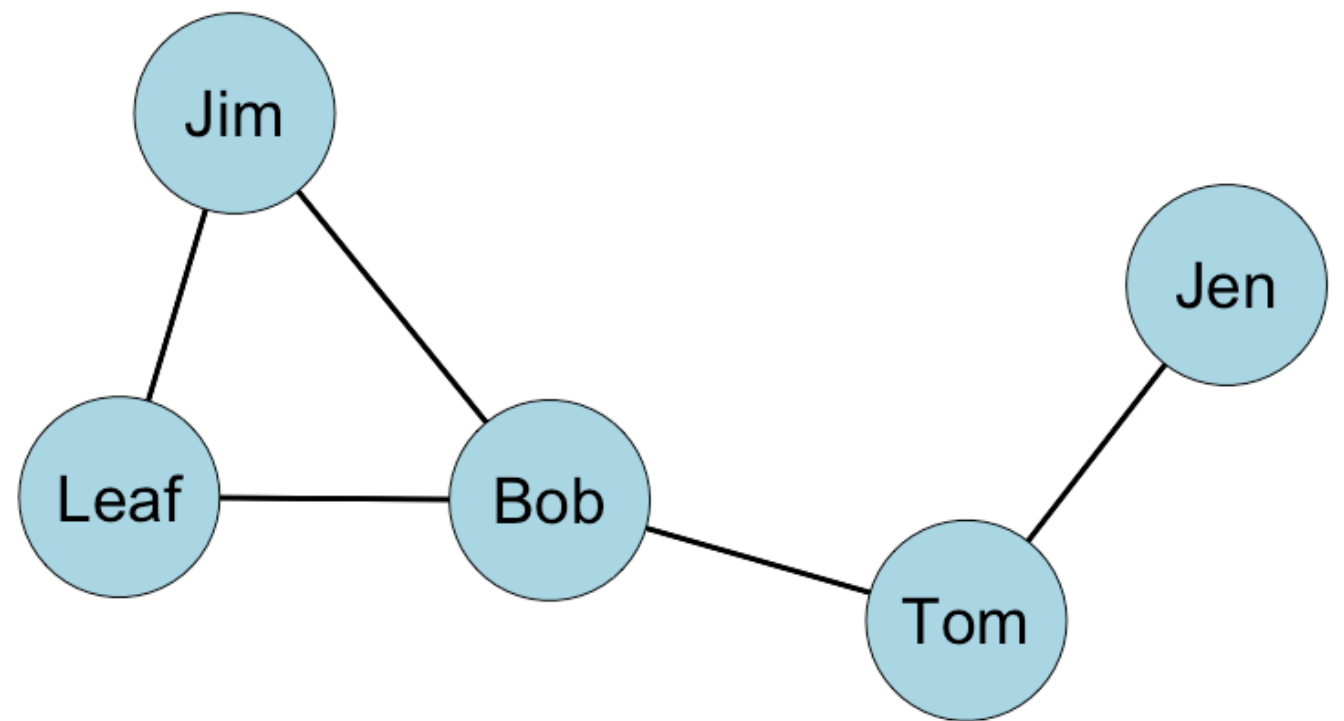
$$\frac{2L}{g(g-1)}$$

- ❖ Where L is the number of edges and g is the number of nodes.

Example: Undirected, Binary Network

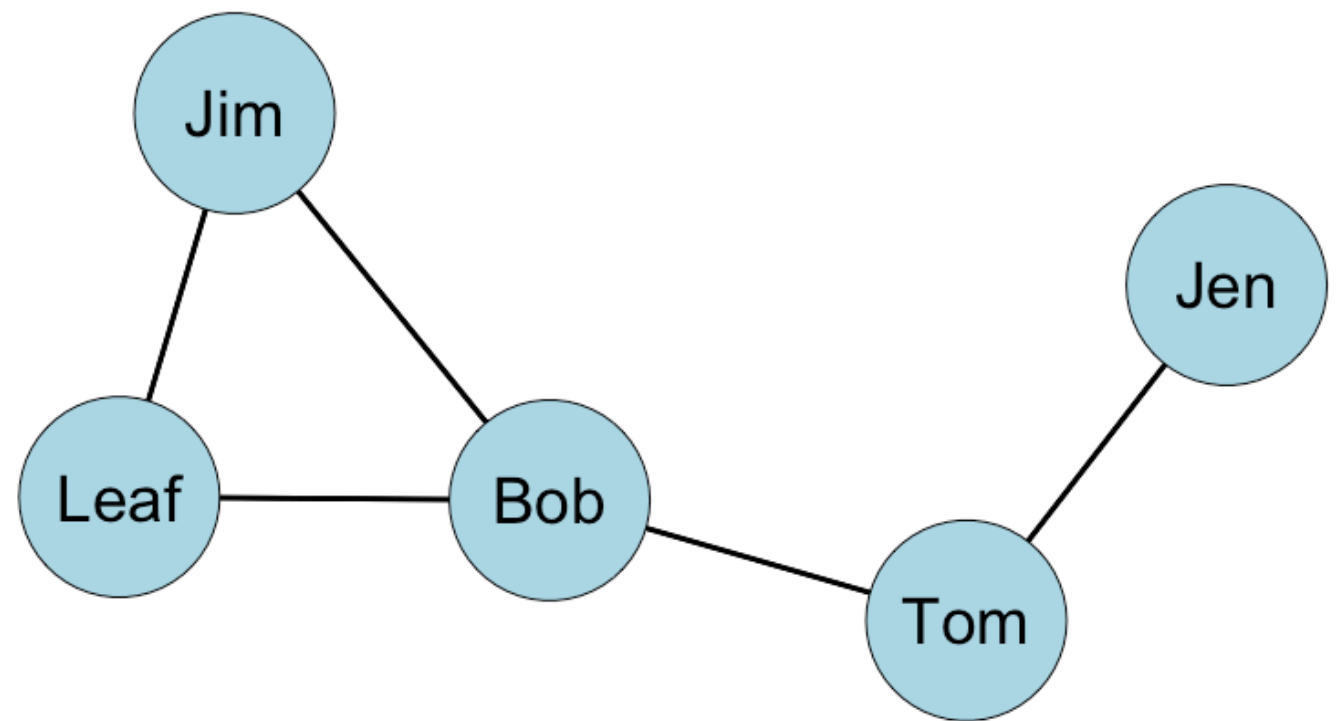
What is the density of this graph?

$$\frac{2L}{g(g-1)}$$



Example: Undirected, Binary Network

$$\frac{2L}{g(g-1)} = \frac{2 \times 5}{5(5-1)} = 0.5$$

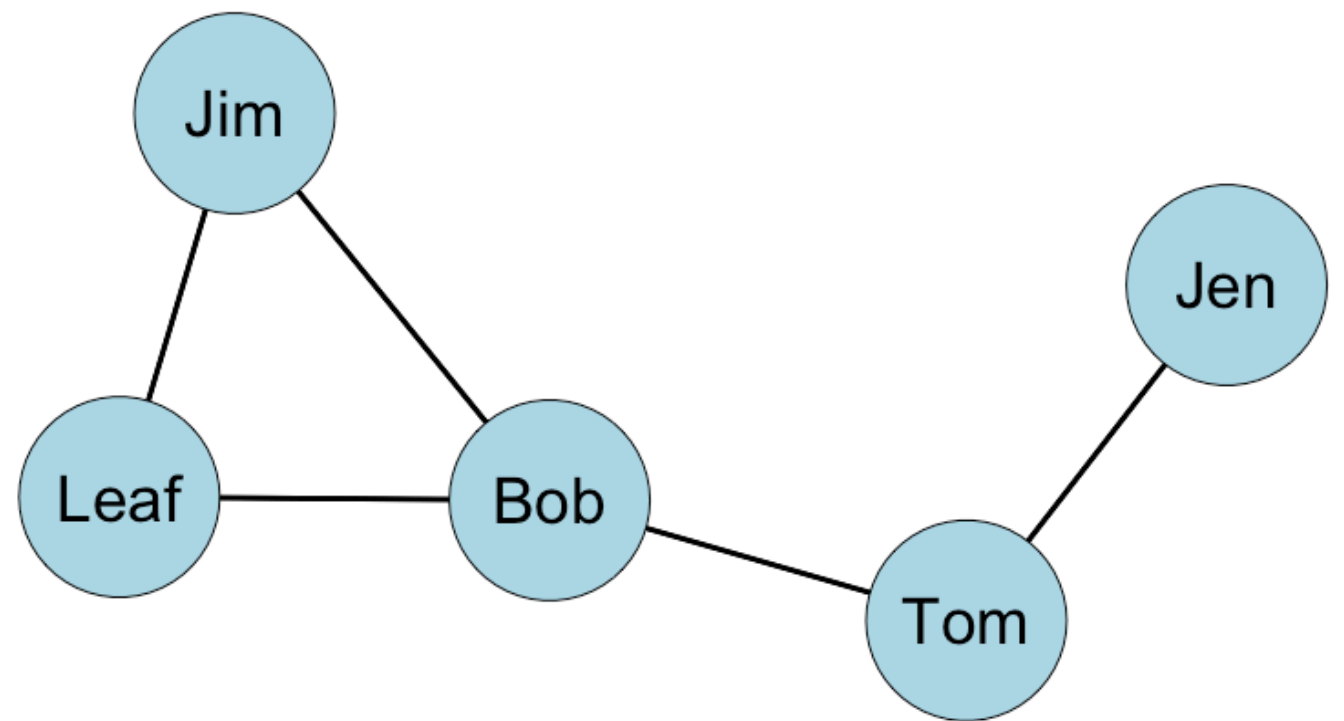


Example: Undirected, Binary Network

L is 5 because there are 5 edges

$$\frac{2L}{g(g-1)} = \frac{2 \times 5}{5(5-1)} = 0.5$$

g is 5 because there are 5 nodes



Density

- ❖ The density of a **directed** graph is given by:

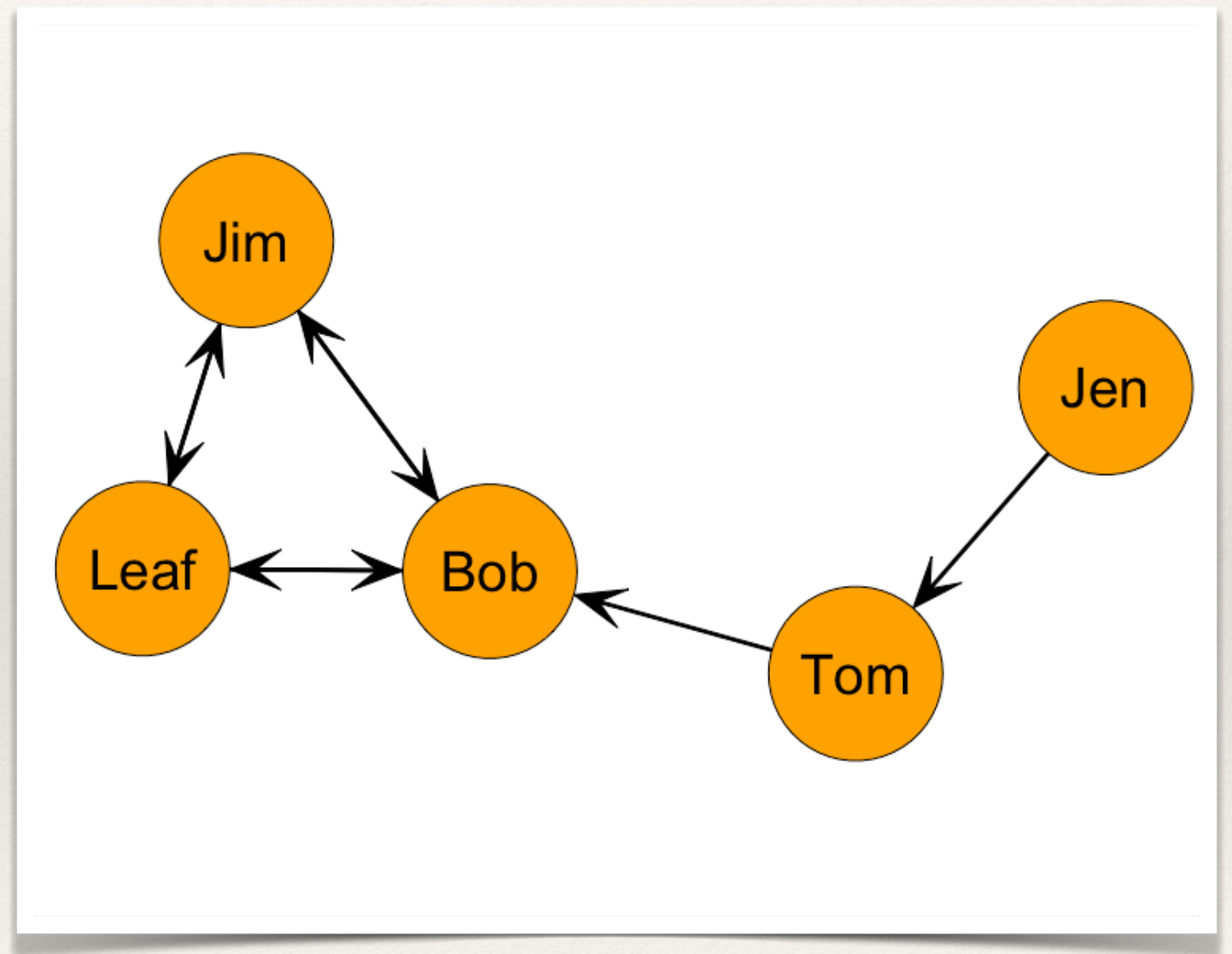
$$\frac{L}{g(g-1)}$$

- ❖ Where L is the number of edges and g is the number of nodes.

Example: Undirected, Binary Network

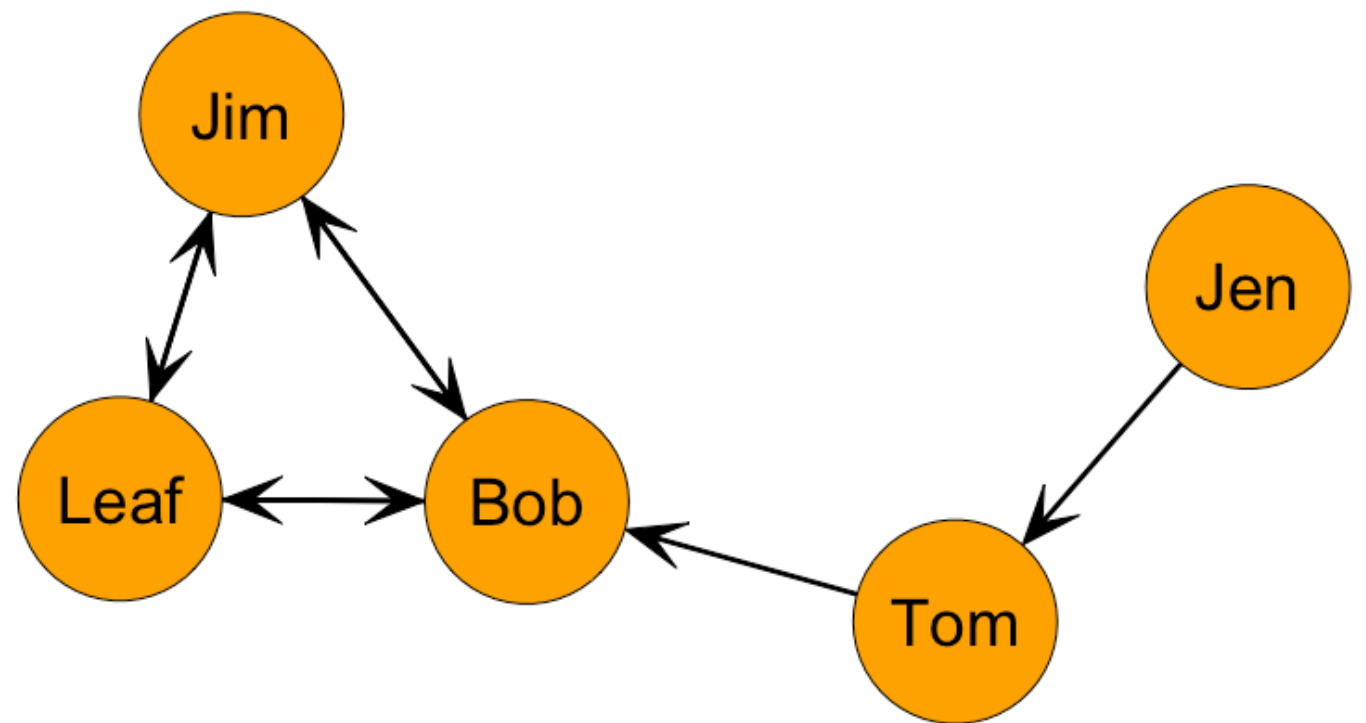
What is the density of this graph?

$$\frac{L}{g(g-1)}$$



Example: Undirected, Binary Network

$$\frac{L}{g(g-1)} = \frac{8}{5(5-1)} = 0.4$$

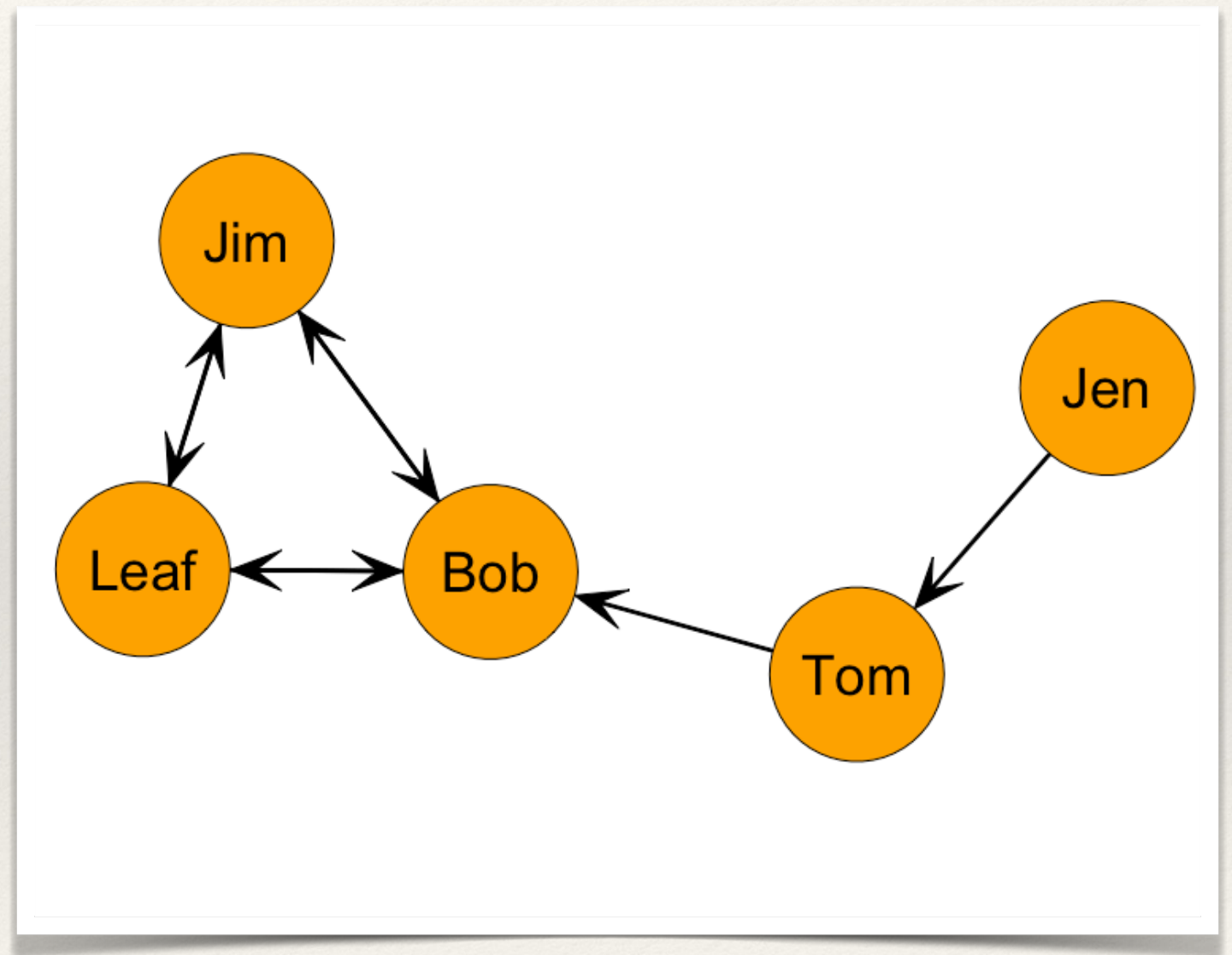


Example: Undirected, Binary Network

L is 8 because there are 8 edges

$$\frac{L}{g(g-1)} = \frac{8}{5(5-1)} = 0.4$$

g is 5 because there are 5 nodes

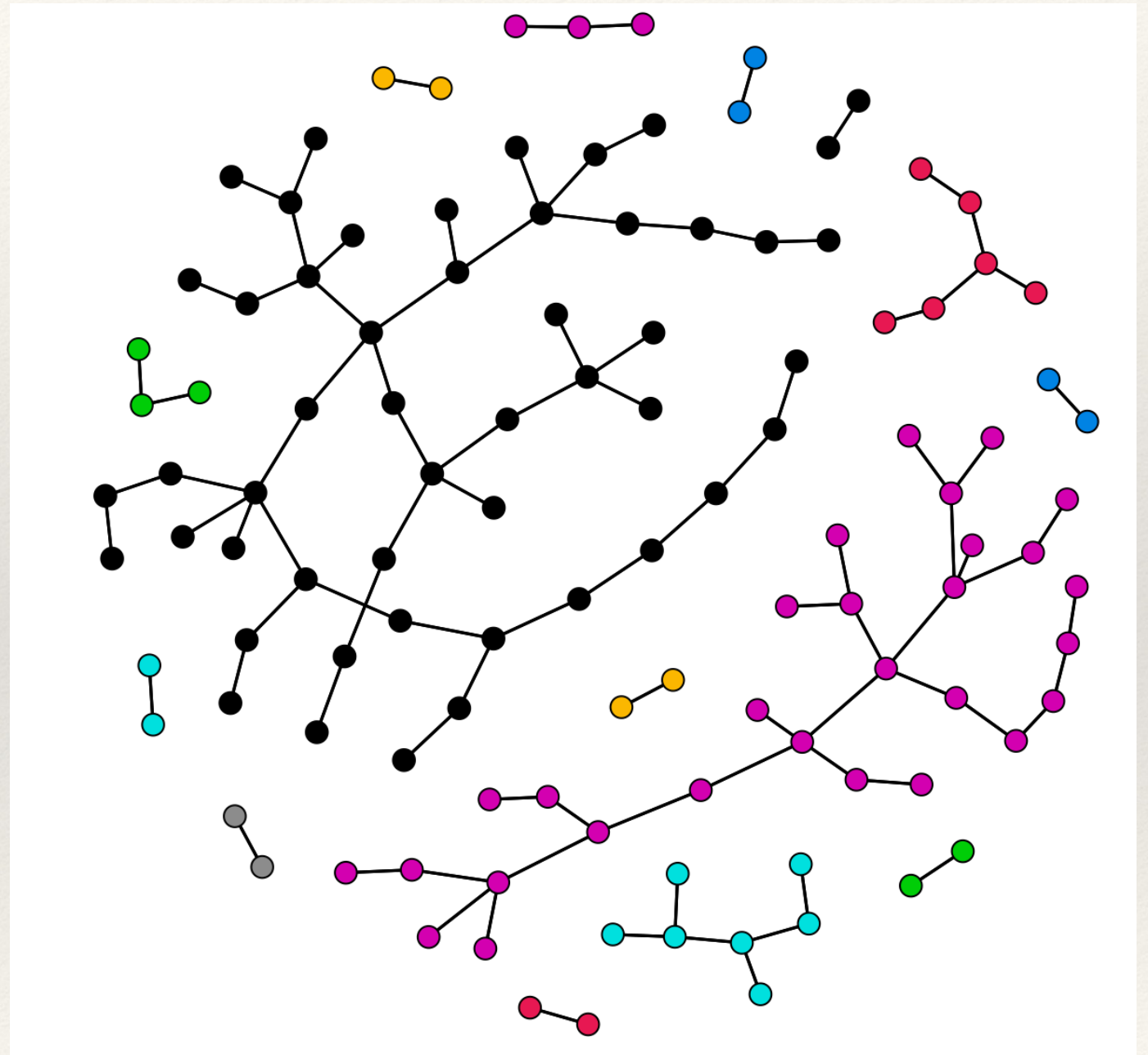


Components

- ❖ How many subgroups are in a network?
- ❖ Put differently, how many *disconnected* subgraphs are there in the graph?
- ❖ These are called **components**

Components

This graph has 15 components

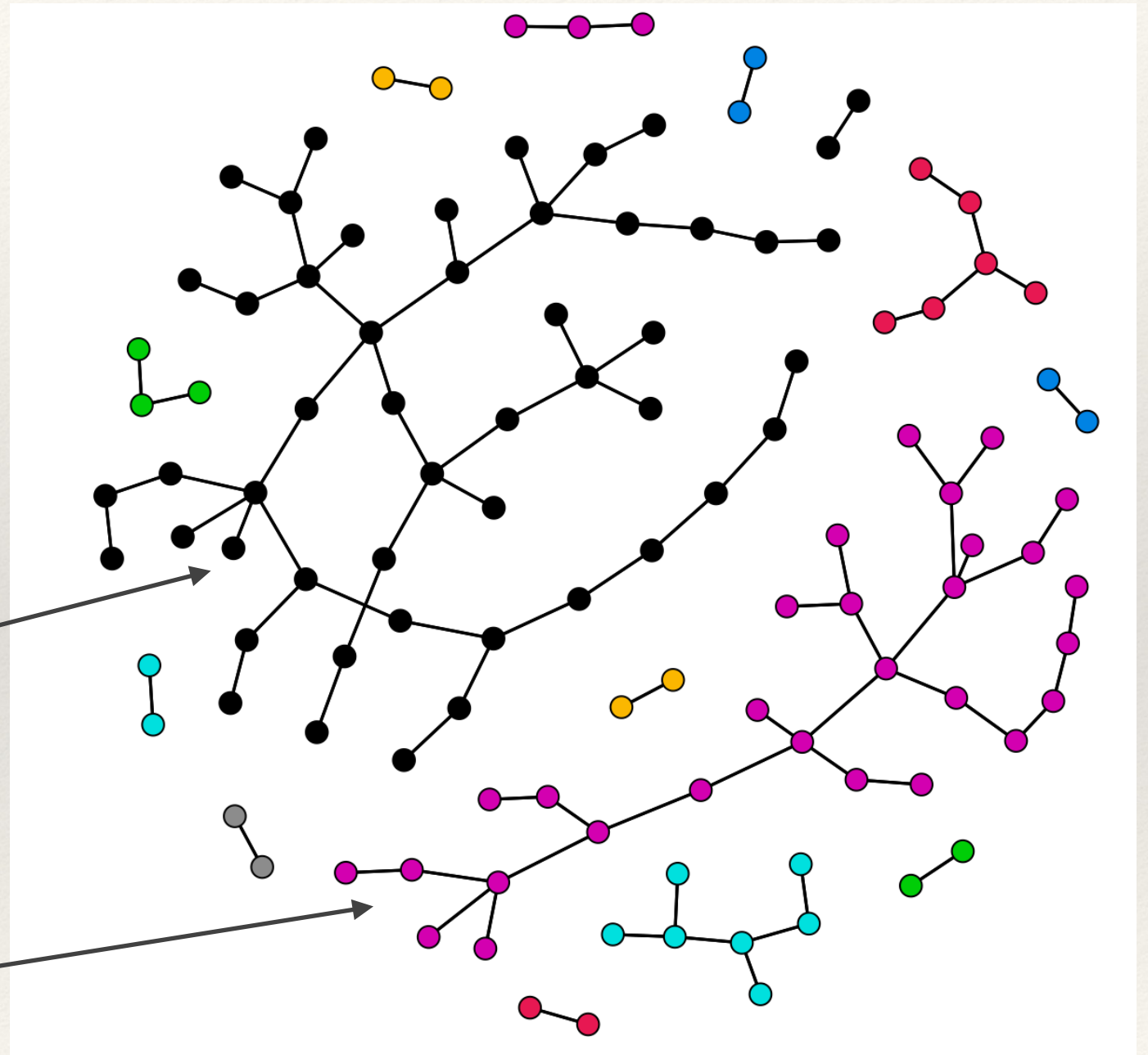


Components

Components are often
ordered by size

First component

Second component

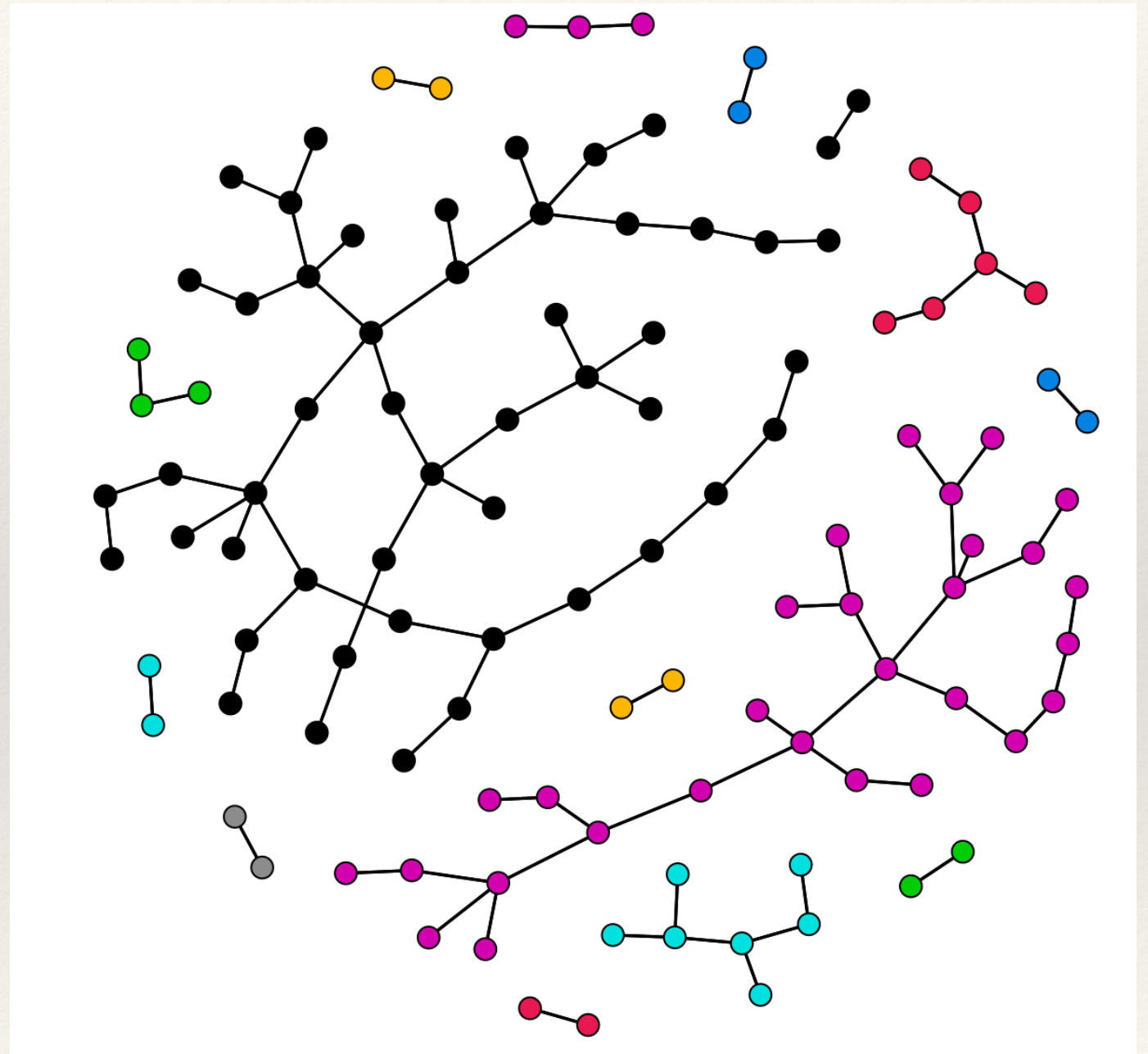


Components

Conceptually, what does it mean if
the graph has:

A lot of components?

No components?

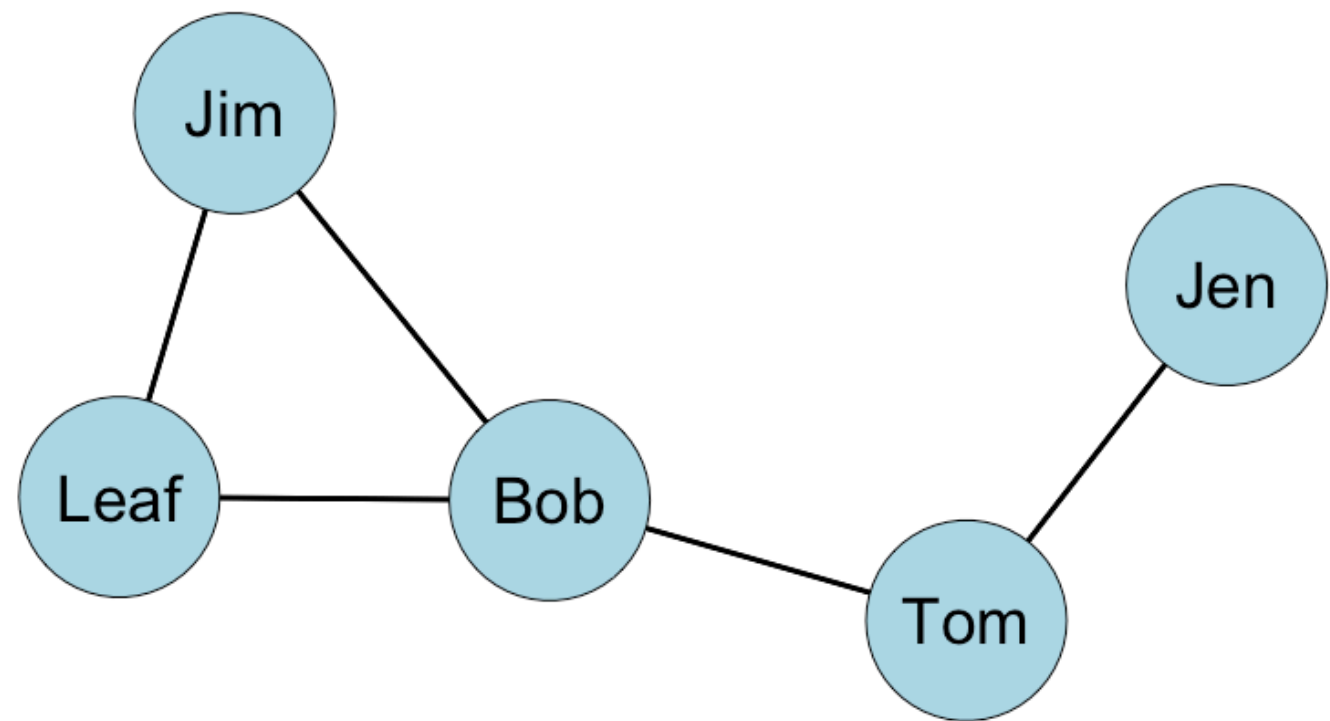


Diameter

- ❖ How compact is the network?
 - ❖ Take two nodes, A and B, on average, how many steps does A have to take to reach B?
 - ❖ A **step** is movement along an edge.
 - ❖ A **shortest path** is the fewest number of steps A has to take to reach B.
 - ❖ The **diameter** is the longest of all the shortest paths in the network.

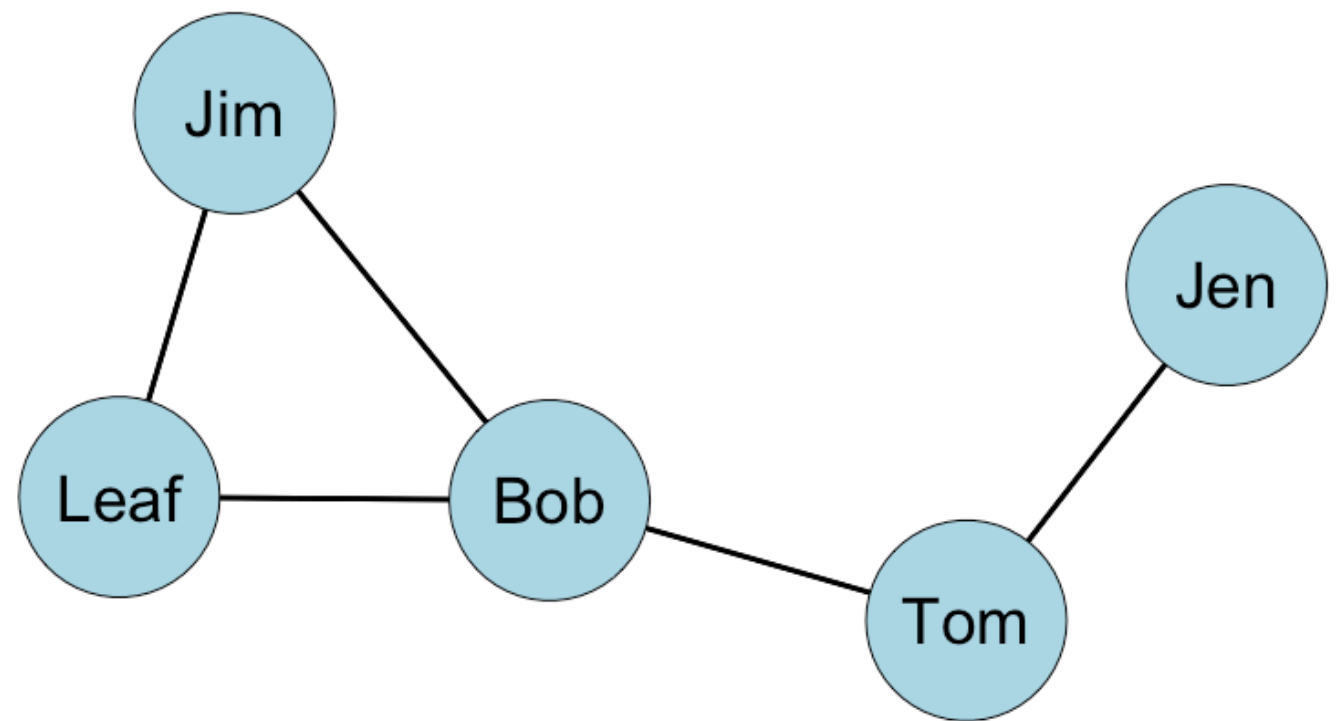
Example: Undirected, Binary Network

What is the shortest path from Leaf to Jen?



Example: Undirected, Binary Network

*What are the paths from Leaf to Jen?
And if there is more than one, which
is shortest?*

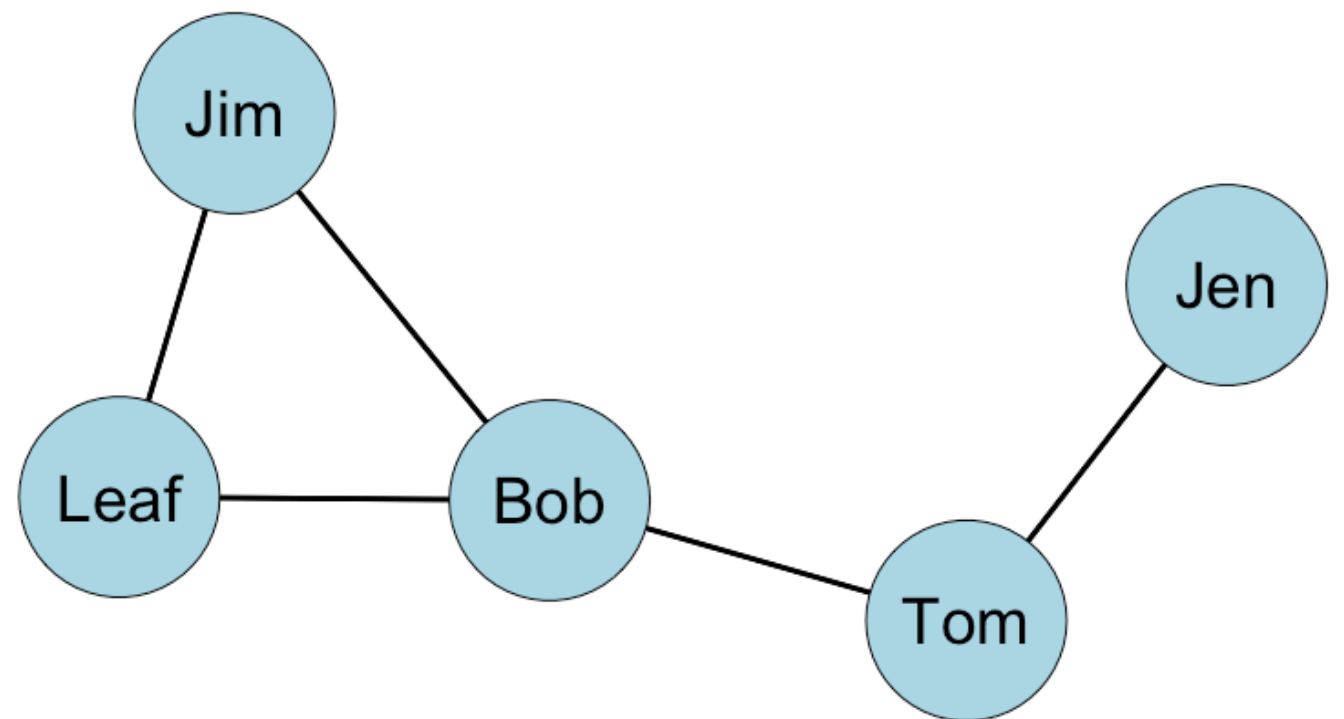


Example: Undirected, Binary Network

*What are the paths from Leaf to Jen?
And if there is more than one, which
is shortest?*

Leaf - Jim - Bob - Tom - Jen?

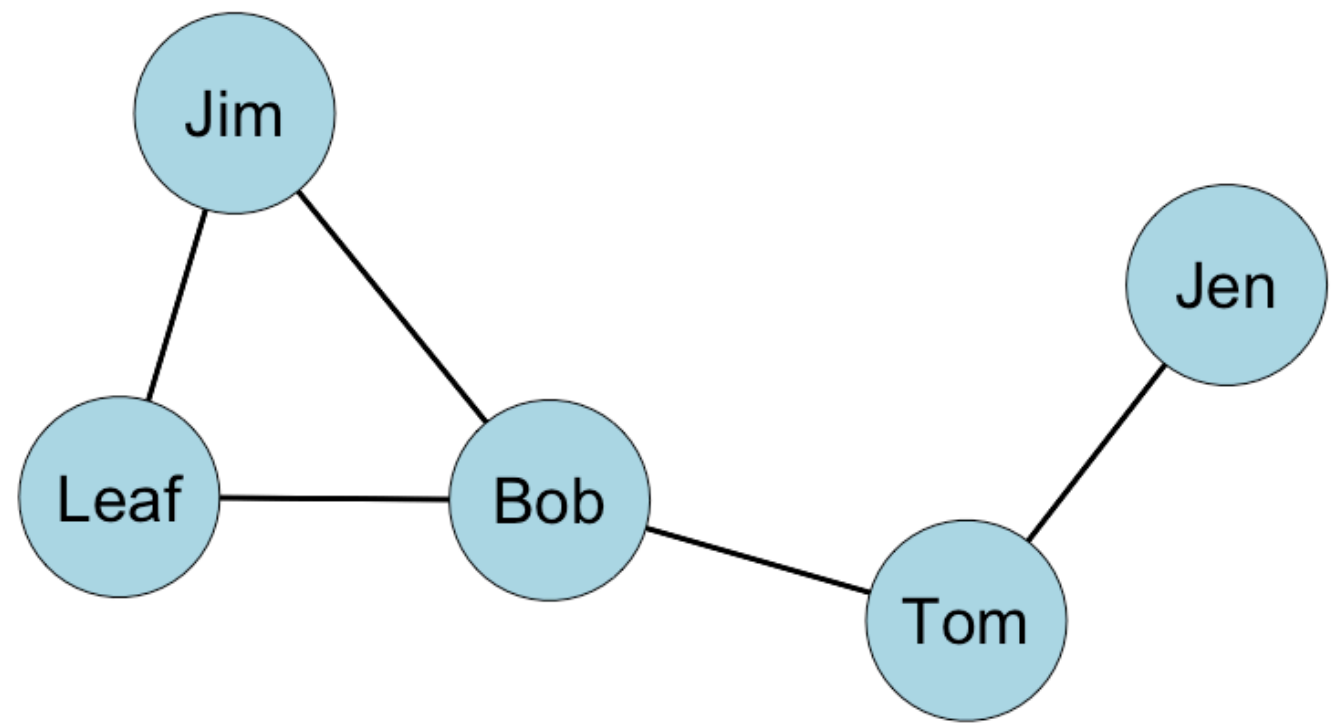
Leaf - Bob - Tom - Jen?



Example: Undirected, Binary Network

What is the diameter of the network?

Or, what two actors have the longest path?



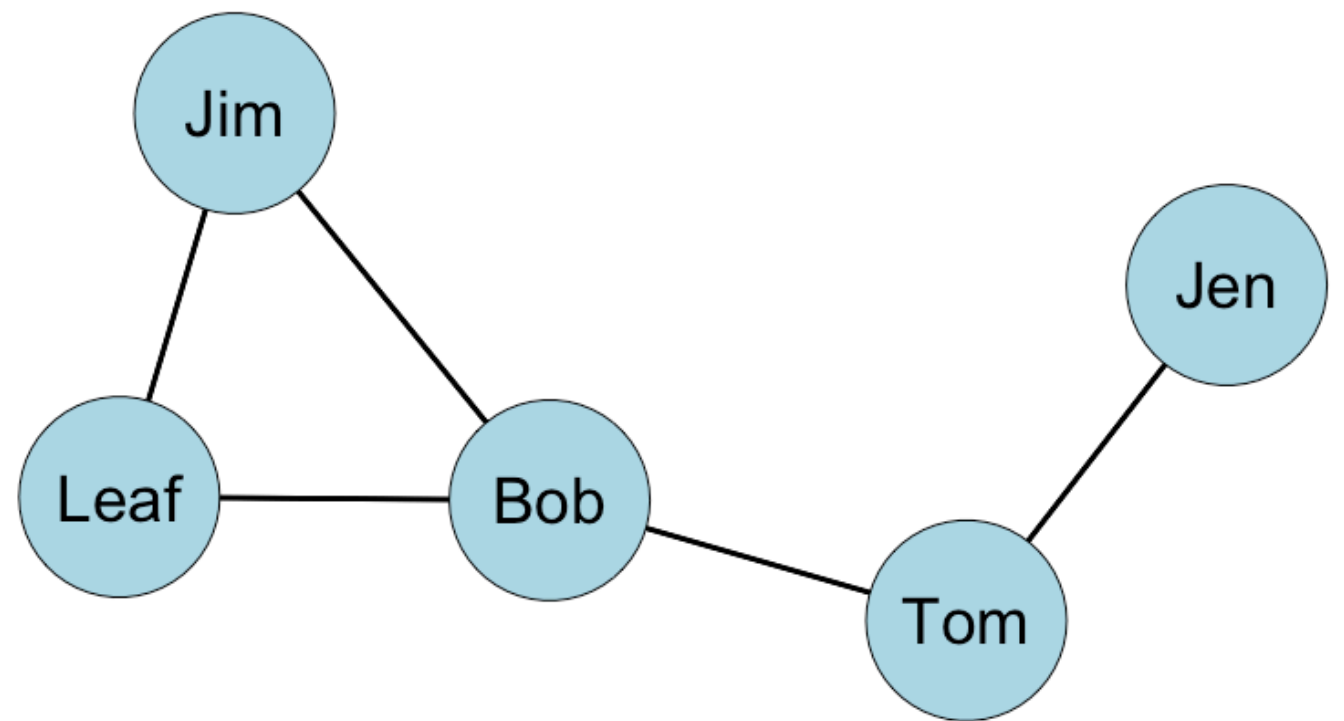
Example: Undirected, Binary Network

What is the diameter of the network?

Or, what two actors have the longest path?

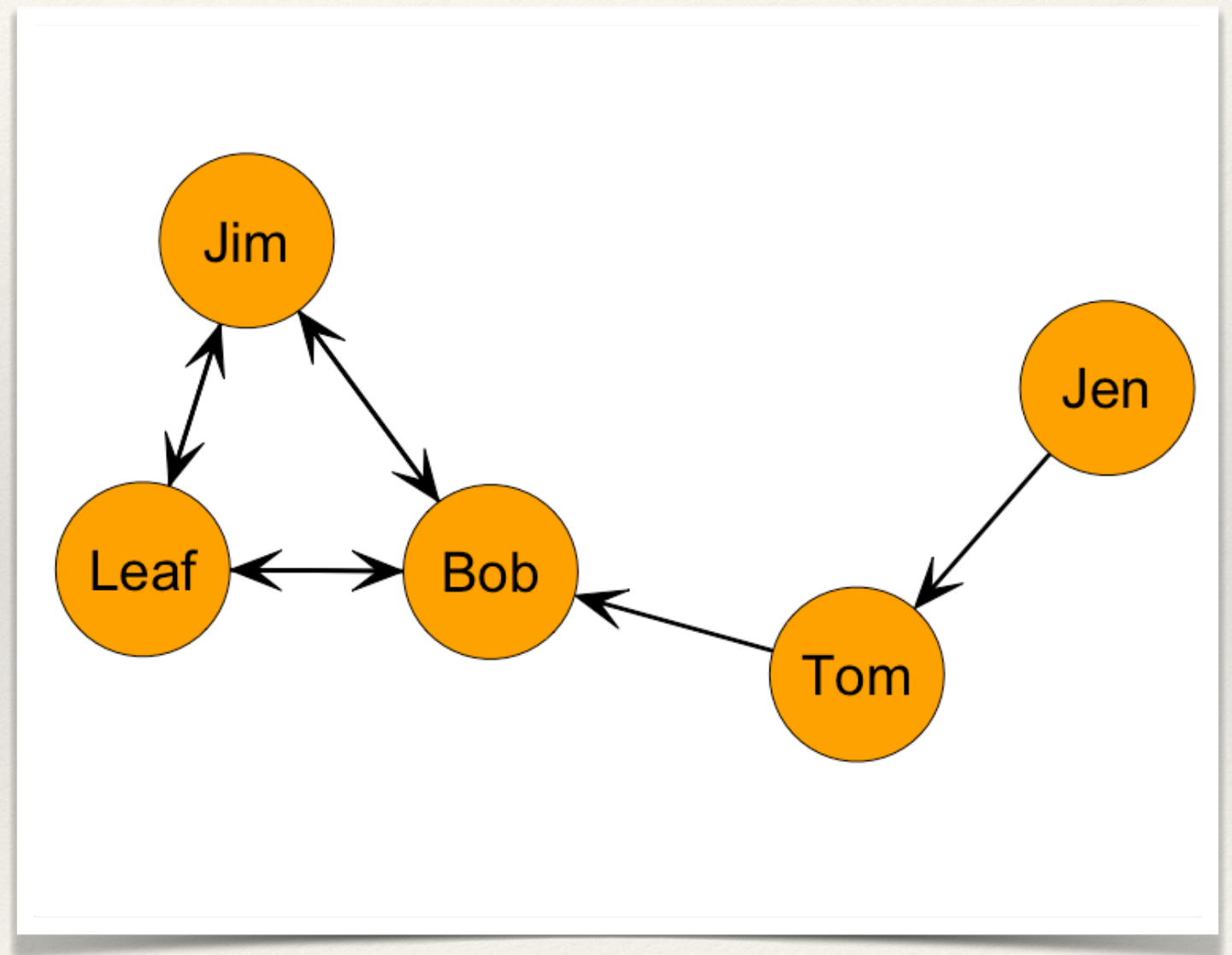
Jen and Leaf/Jim

So the diameter is 3.



Example: Undirected, Binary Network

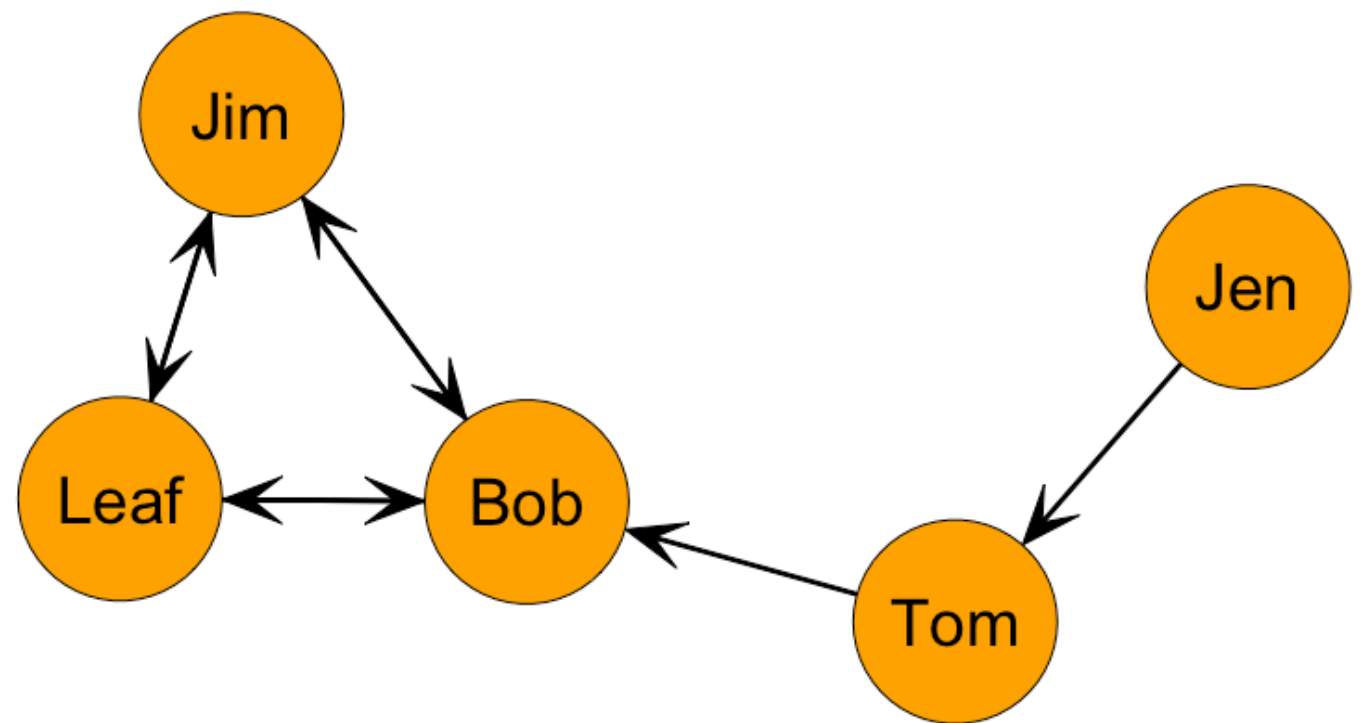
What is the shortest path from Leaf to Jen?



Example: Undirected, Binary Network

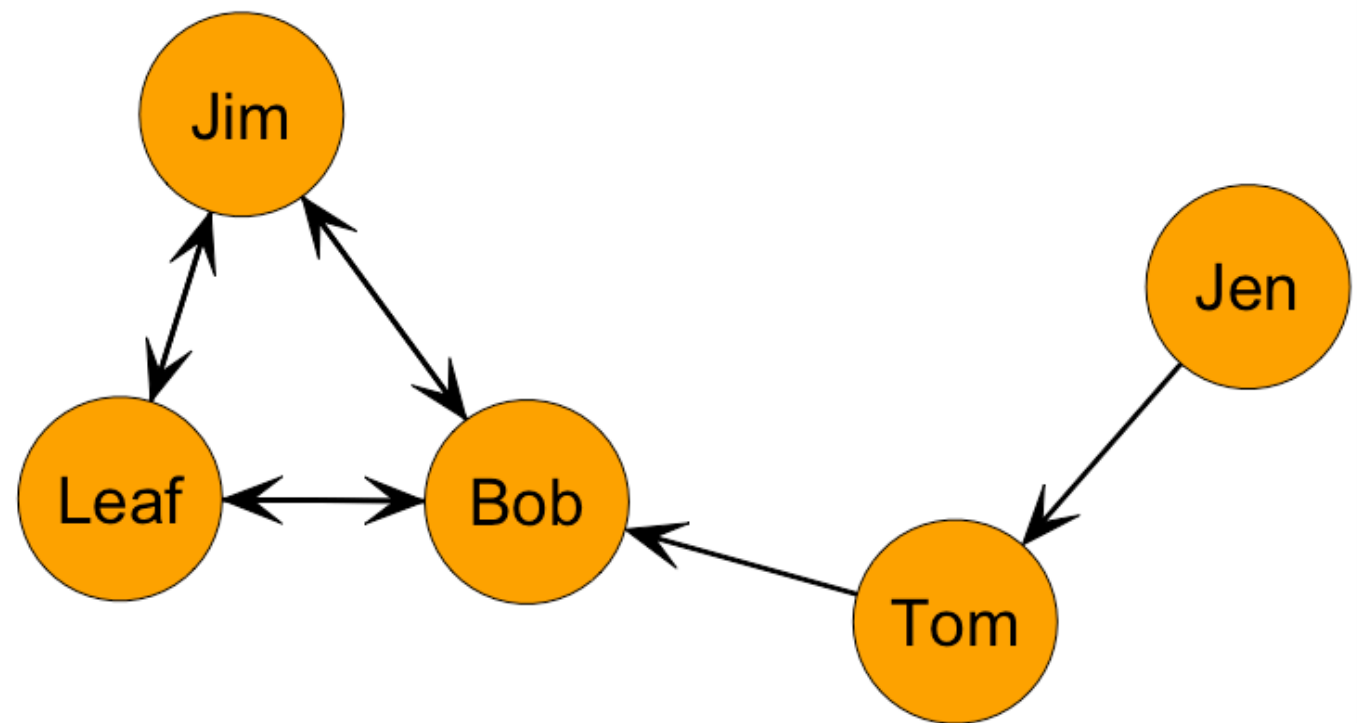
What is the shortest path from Leaf to Jen?

Leaf - Bob



Example: Undirected, Binary Network

What is the shortest path from Jen to Jim?

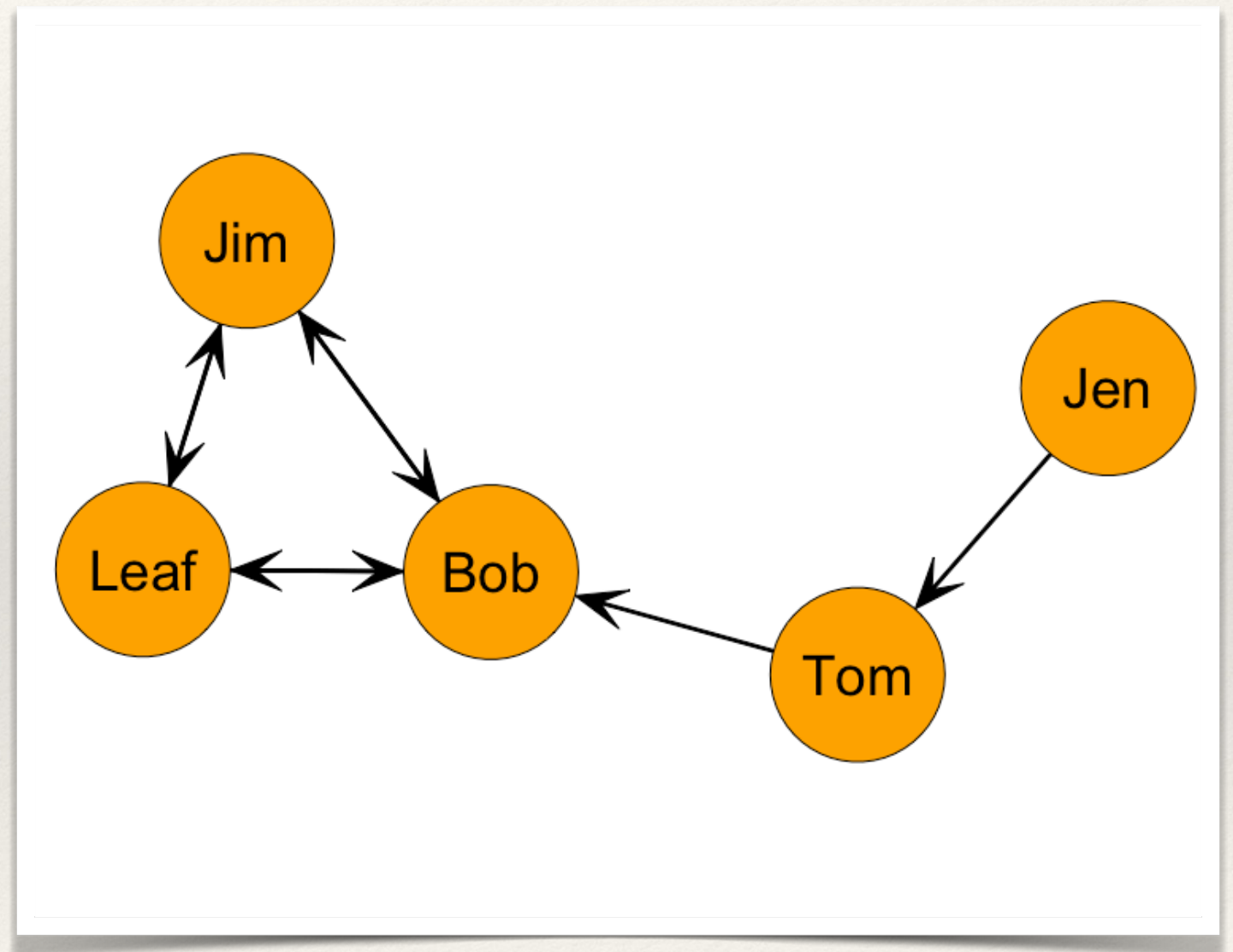


Example: Undirected, Binary Network

What is the shortest path from Jen to Jim?

Jen -> Tom -> Bob -> Leaf -> Jim ?

Jen -> Tom -> Bob -> Jim ?



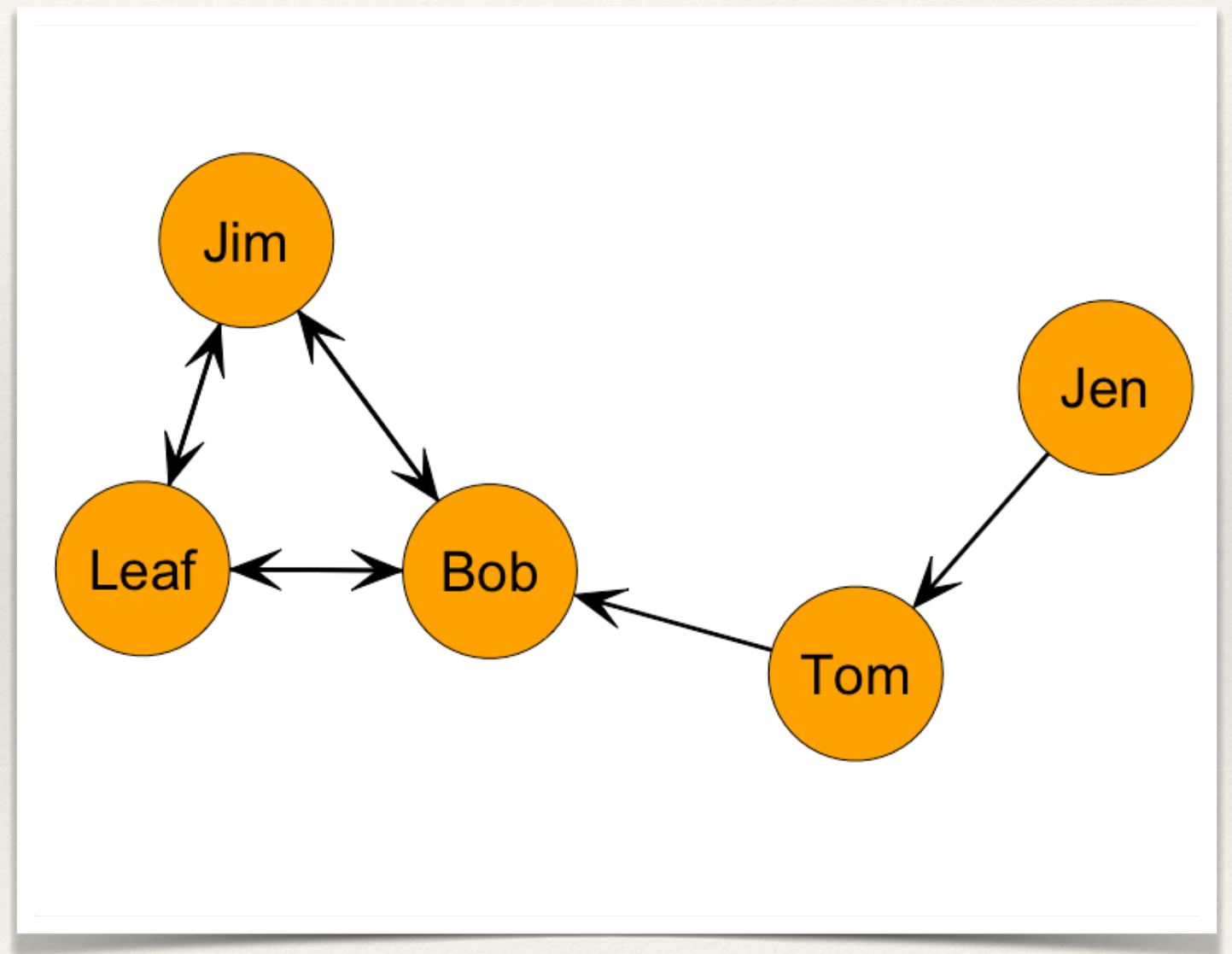
Example: Undirected, Binary Network

What is the diameter of the network?

Or, what two actors have the longest path?

Jen and Jim/Leaf

So the diameter is 3.

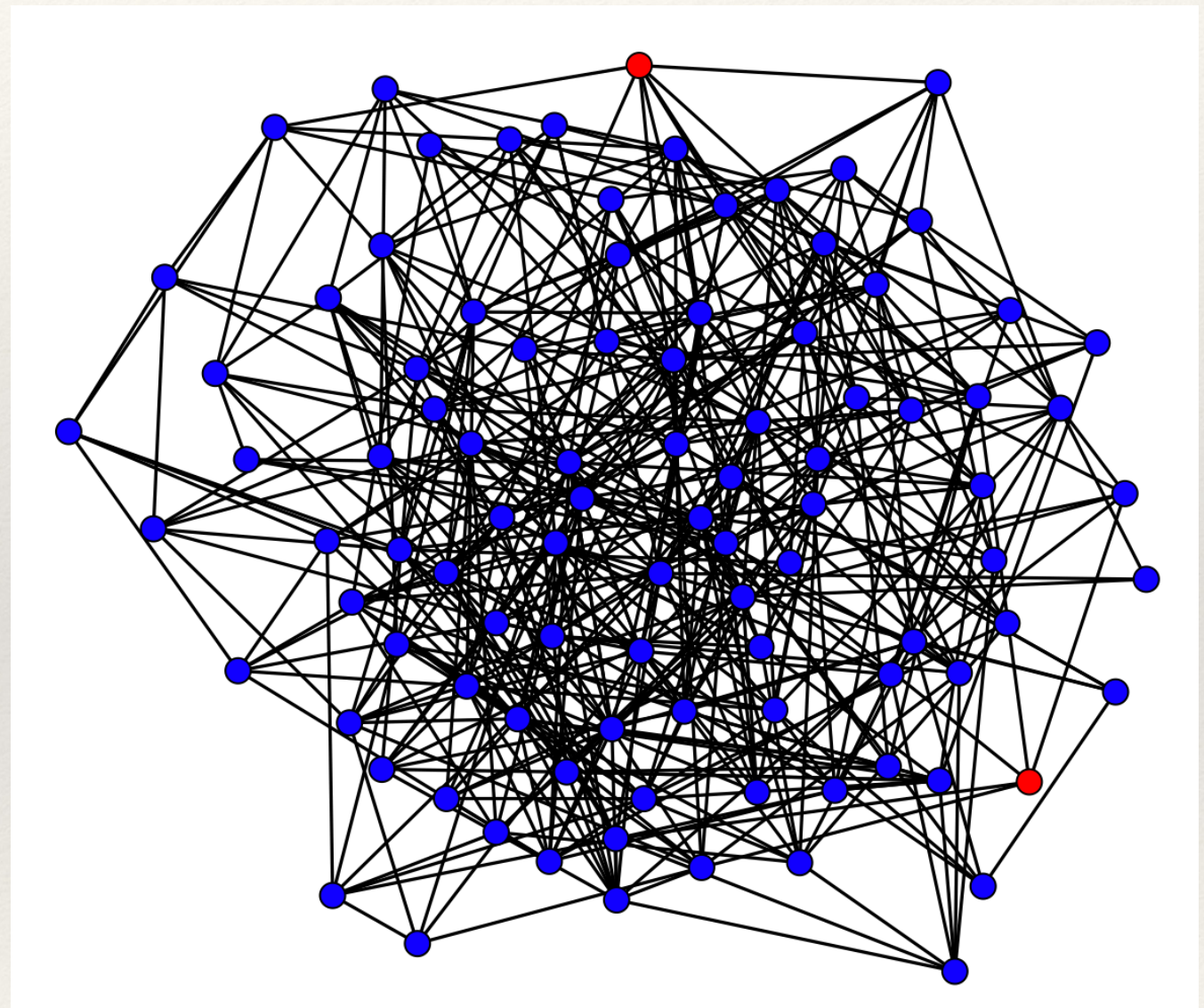


Example: Undirected, Binary Network

In larger graphs, it isn't as easy (possible?) to count the paths just by looking at the network

Diameter is 11: it takes 11 steps from the two red dots

As we will see later, these distances can be represented in a matrix



Transitivity

- ❖ How much clustering is in the network?
 - ❖ Take three nodes, A, B, C.
 - ❖ A **triplet** is three nodes connected by either two or three edges.
 - ❖ A *closed triplet* exists if A-B, B-C, and A-C.
 - ❖ A closed triplet is a **triangle**.
 - ❖ An *open triplet* exists if A-B, B-C, but not A-C.
 - ❖ **Transitivity** is the proportion of closed triplets to all triplets.

Transitivity

❖ How much clustering is in the network?

❖ Take three

❖ A **triple**

❖ A *closed*

❖ A *closed*

❖ An *open*

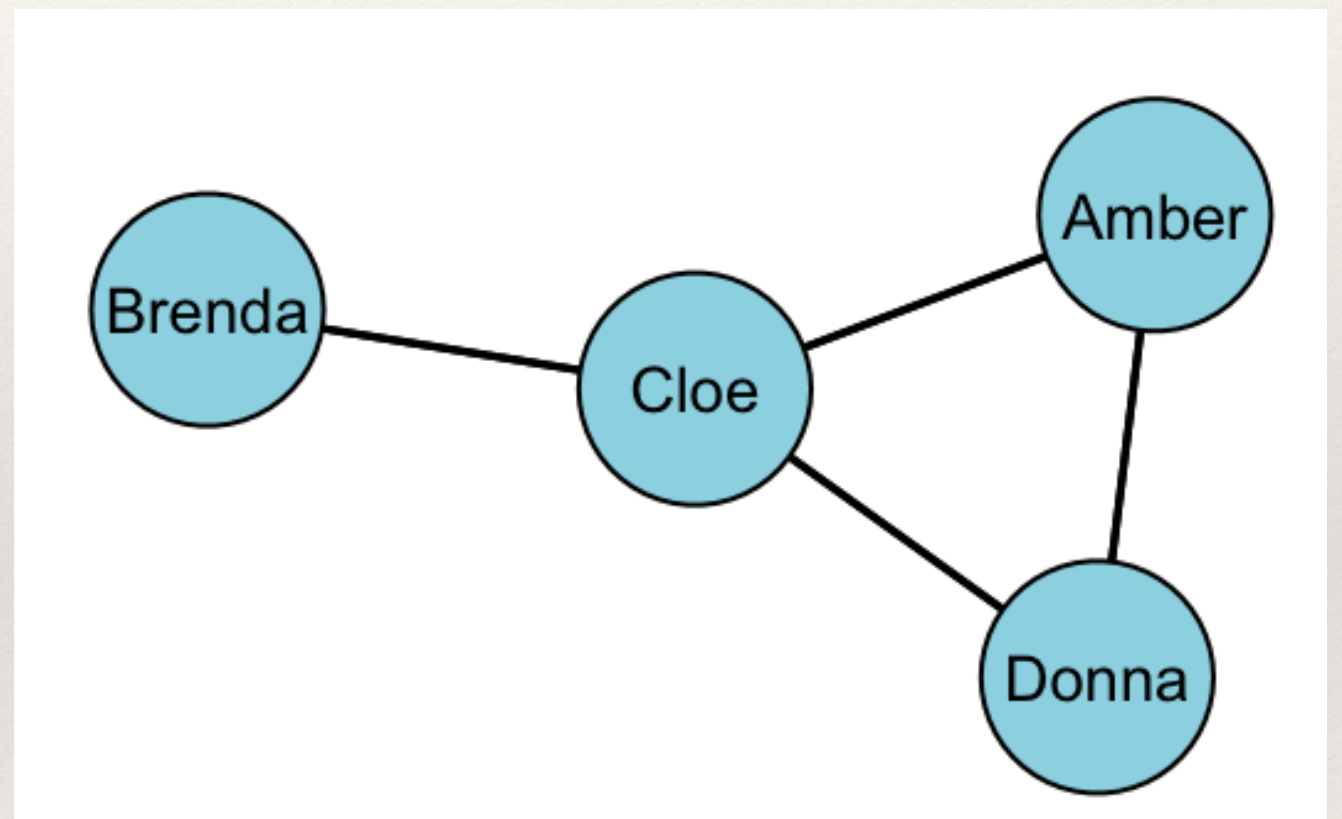
❖ **Transitivity** is the proportion of closed triplets to all triplets.



or three edges.

Example: Undirected, Binary Network

*How many **closed** triplets are there in the graph?*



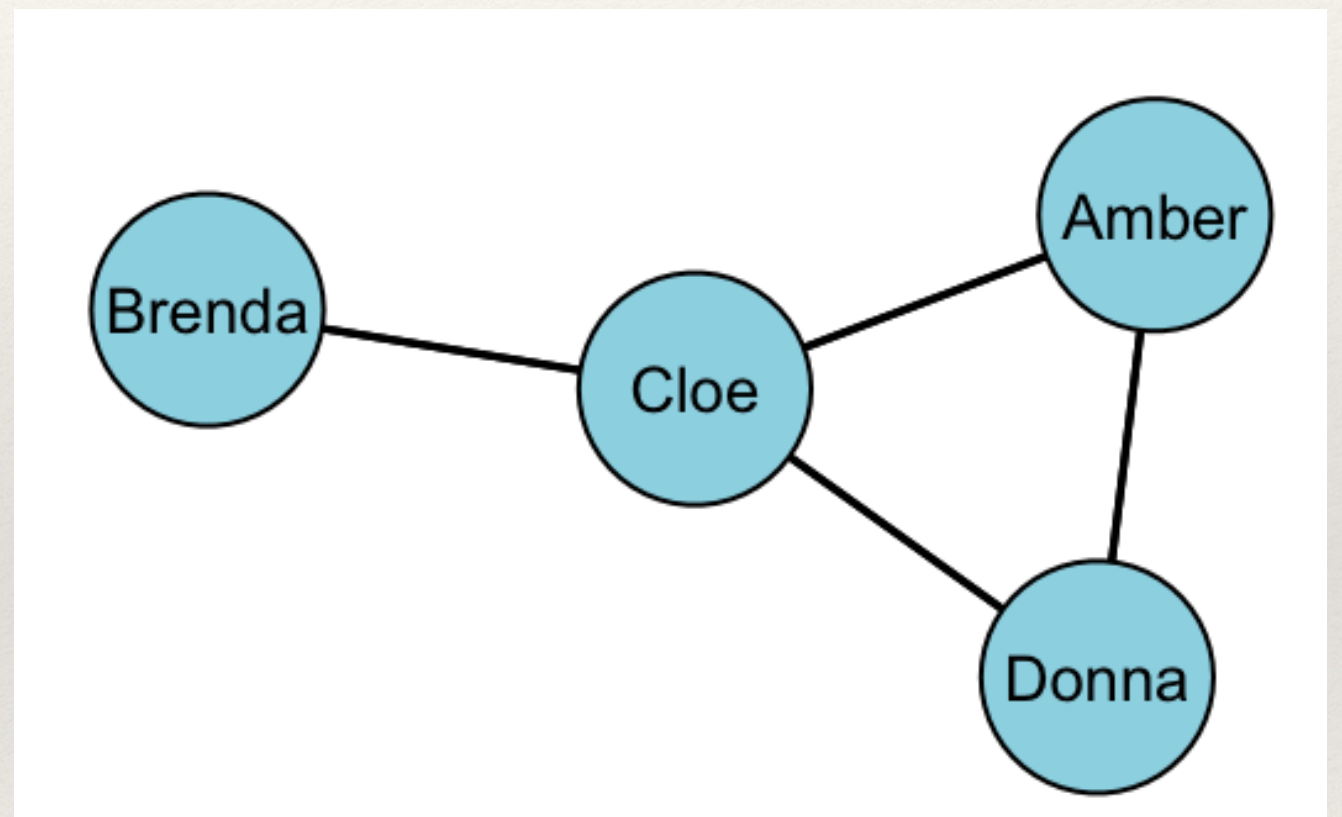
Example: Undirected, Binary Network

*How many **closed** triplets are there in the graph?*

Cloe, Amber, Donna

Amber, Cloe, Donna

Donna, Amber, Cloe

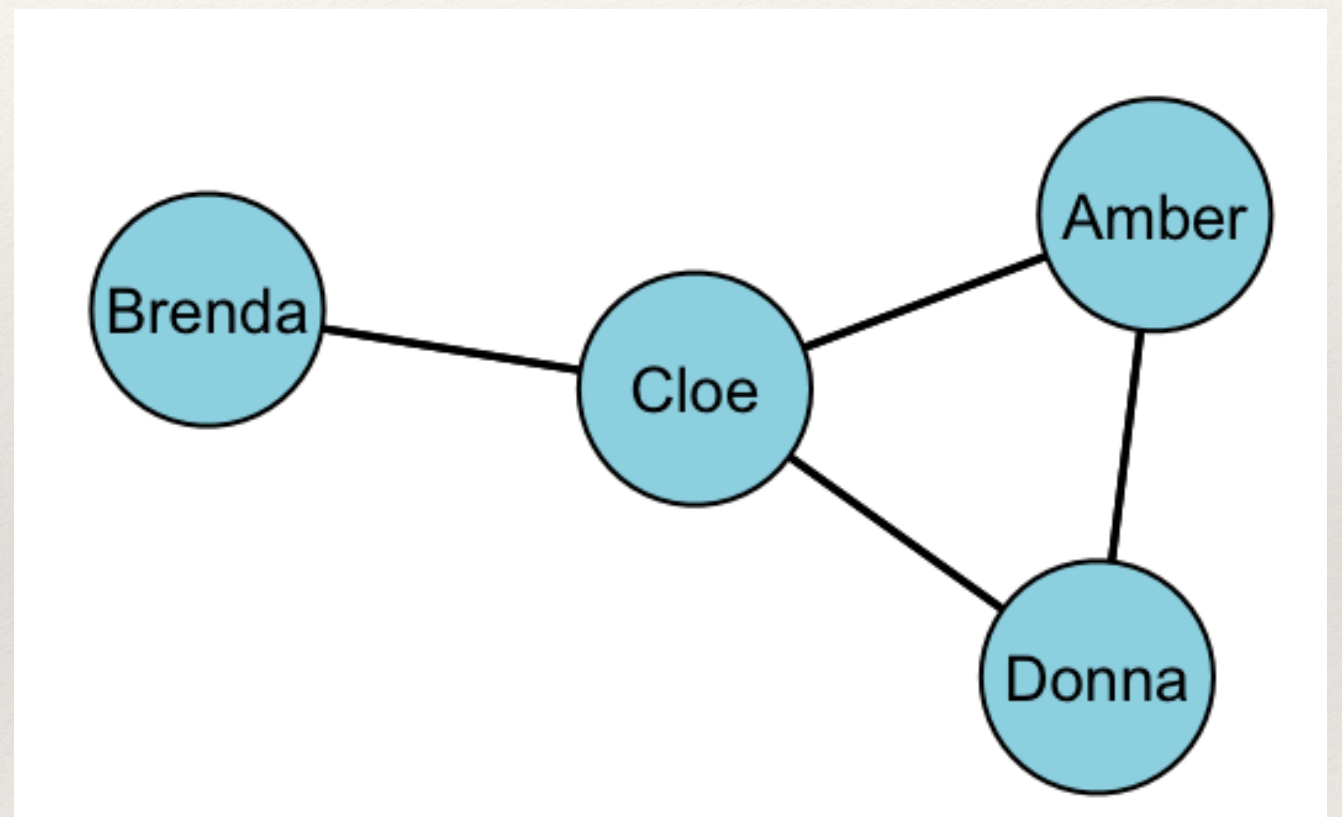


Example: Undirected, Binary Network

*How many **open** triplets are there in the graph?*

Brenda, Cloe, Amber

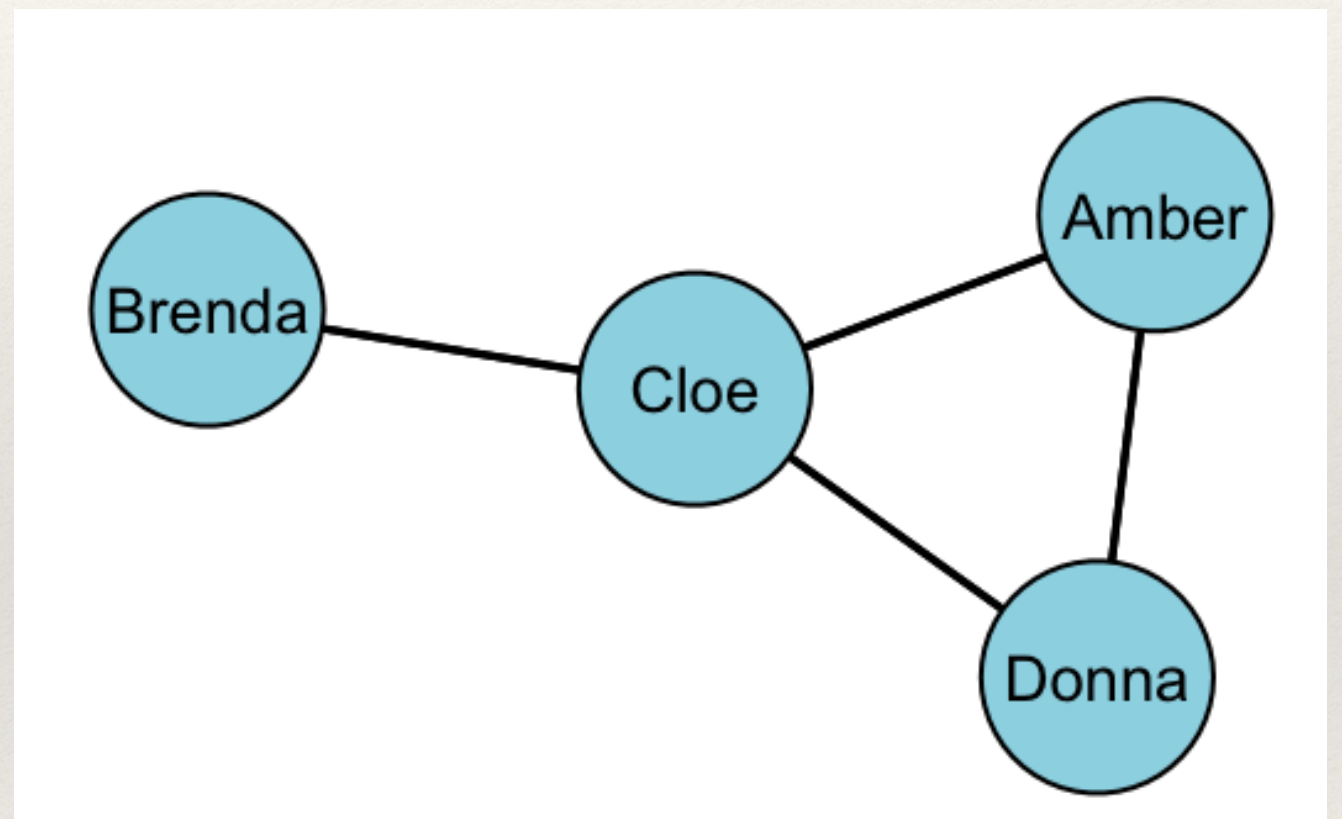
Brenda, Cloe, Donna



Example: Undirected, Binary Network

That gives us:

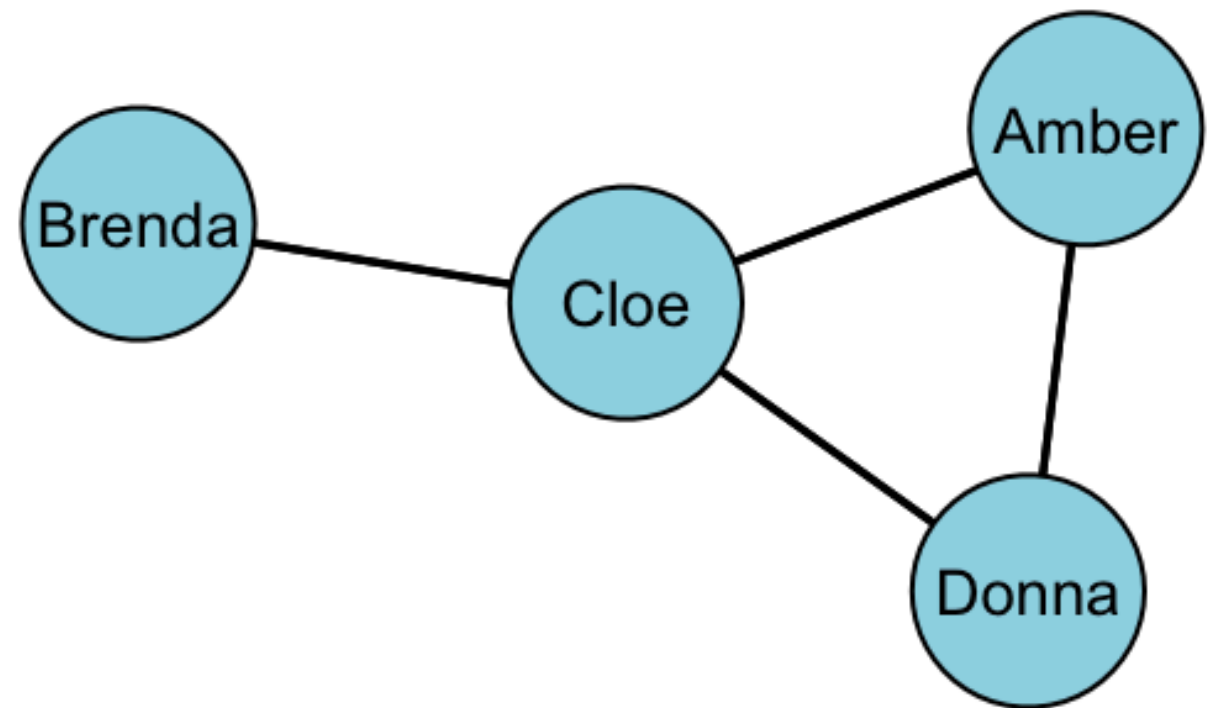
5 triplets total, and 3 are closed



Example: Undirected, Binary Network

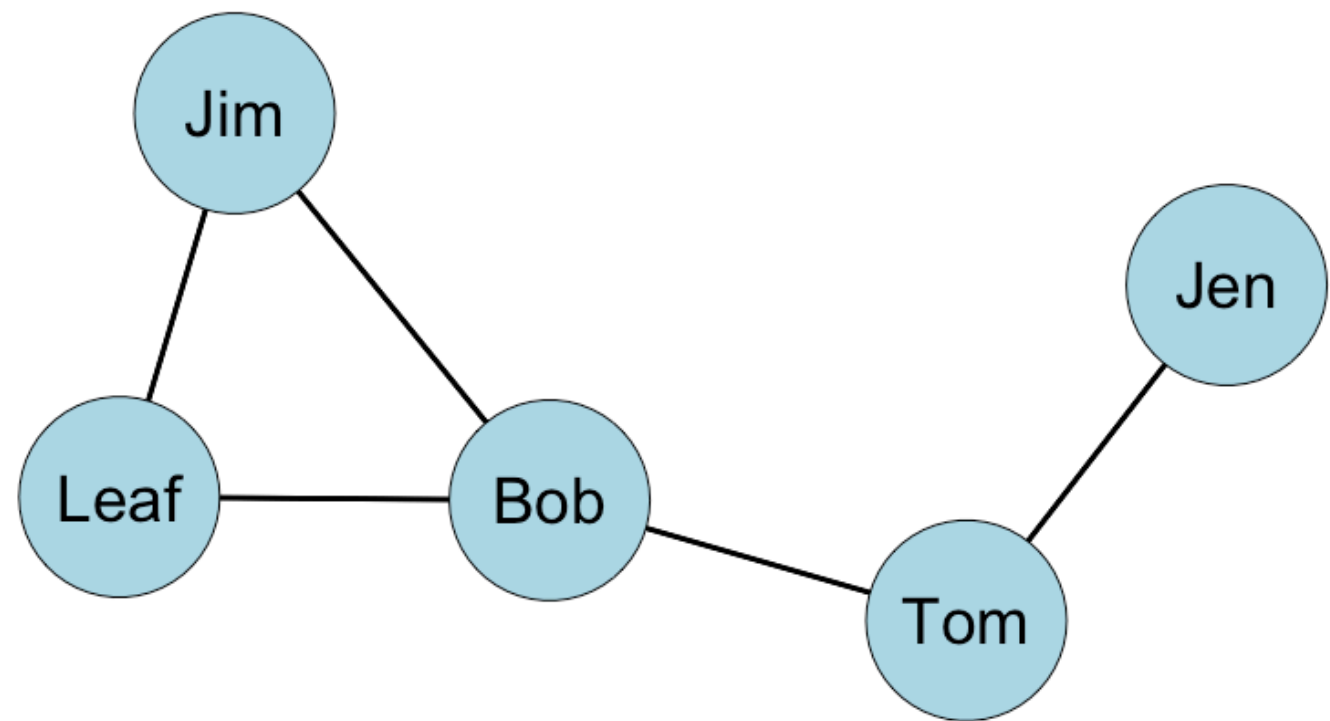
$$C = \frac{\text{number of closed triplets}}{\text{number of all triplets (open and closed)}}$$

$$C = \frac{3}{5} = 0.6$$



Example: Undirected, Binary Network

*How many **closed** triplets are there in the graph?*



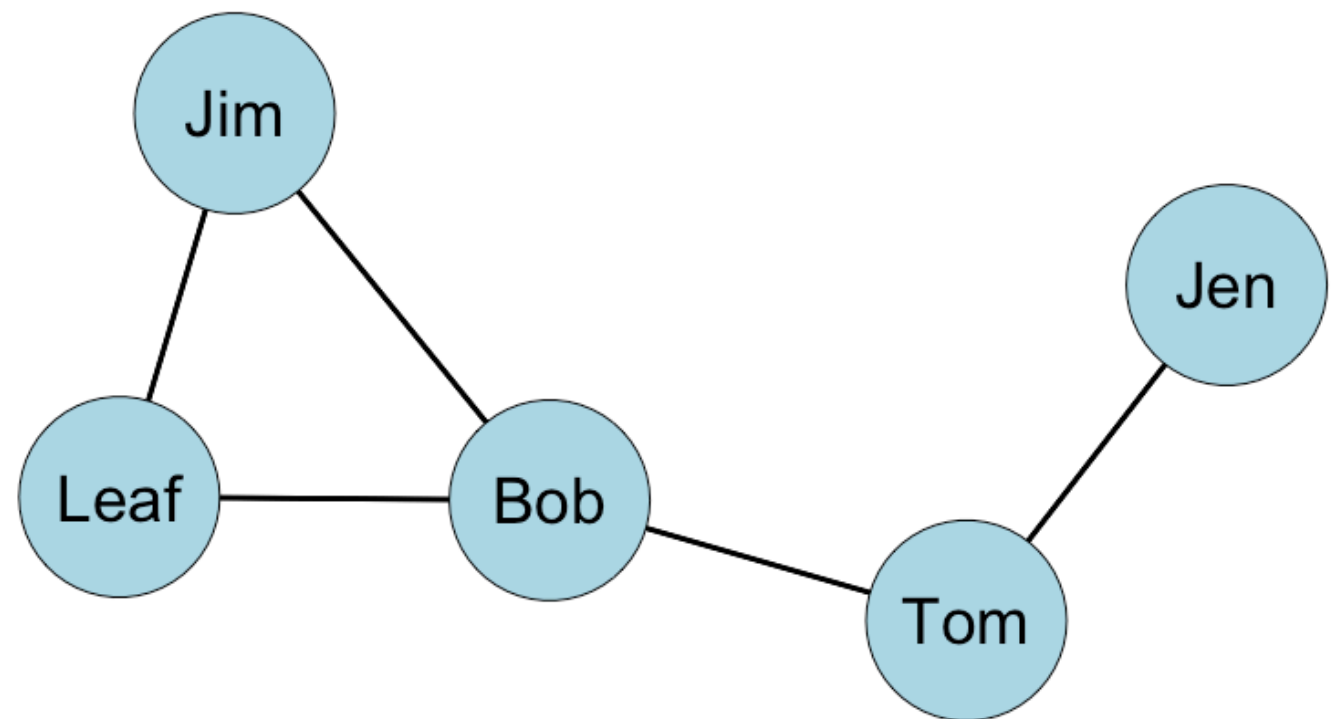
Example: Undirected, Binary Network

*How many **closed** triplets are there in the graph?*

Bob, Leaf, Jim

Leaf, Jim, Bob

Jim, Bob, Leaf



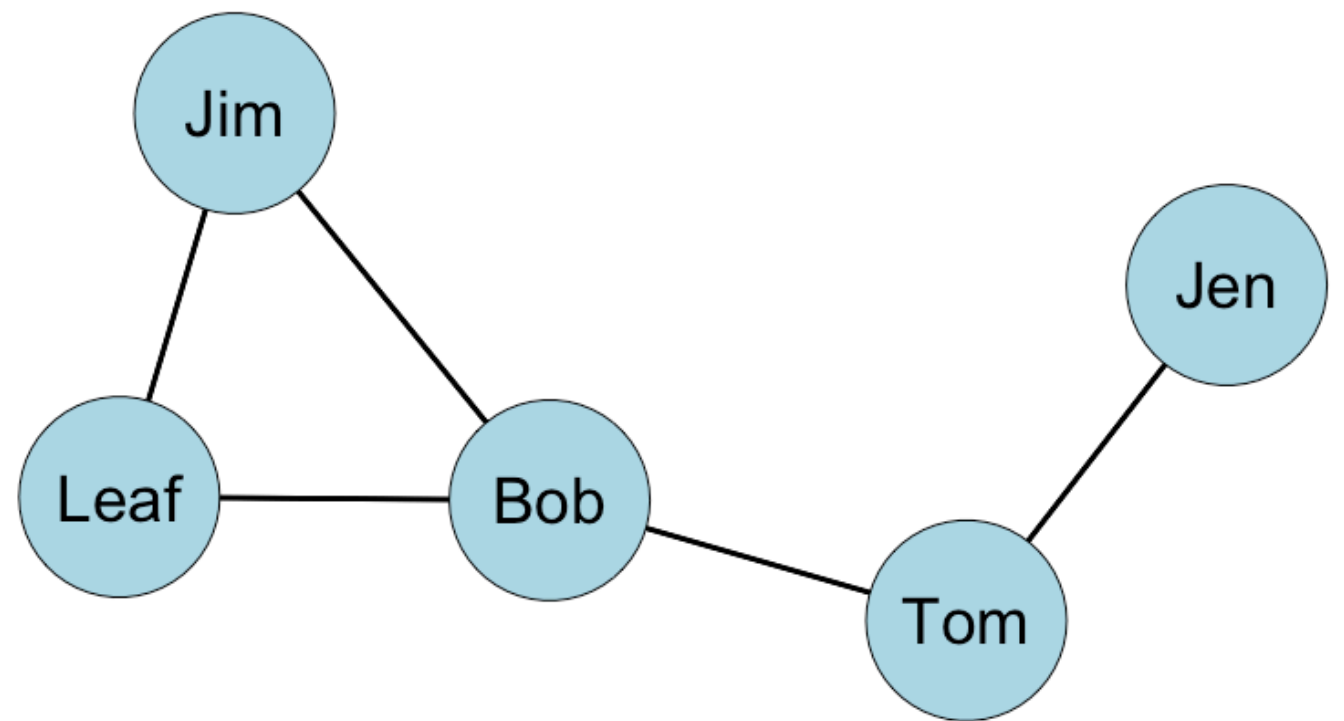
Example: Undirected, Binary Network

*How many **open** triplets are there in the graph?*

Jen, Tom, Bob

Leaf, Bob, Tom

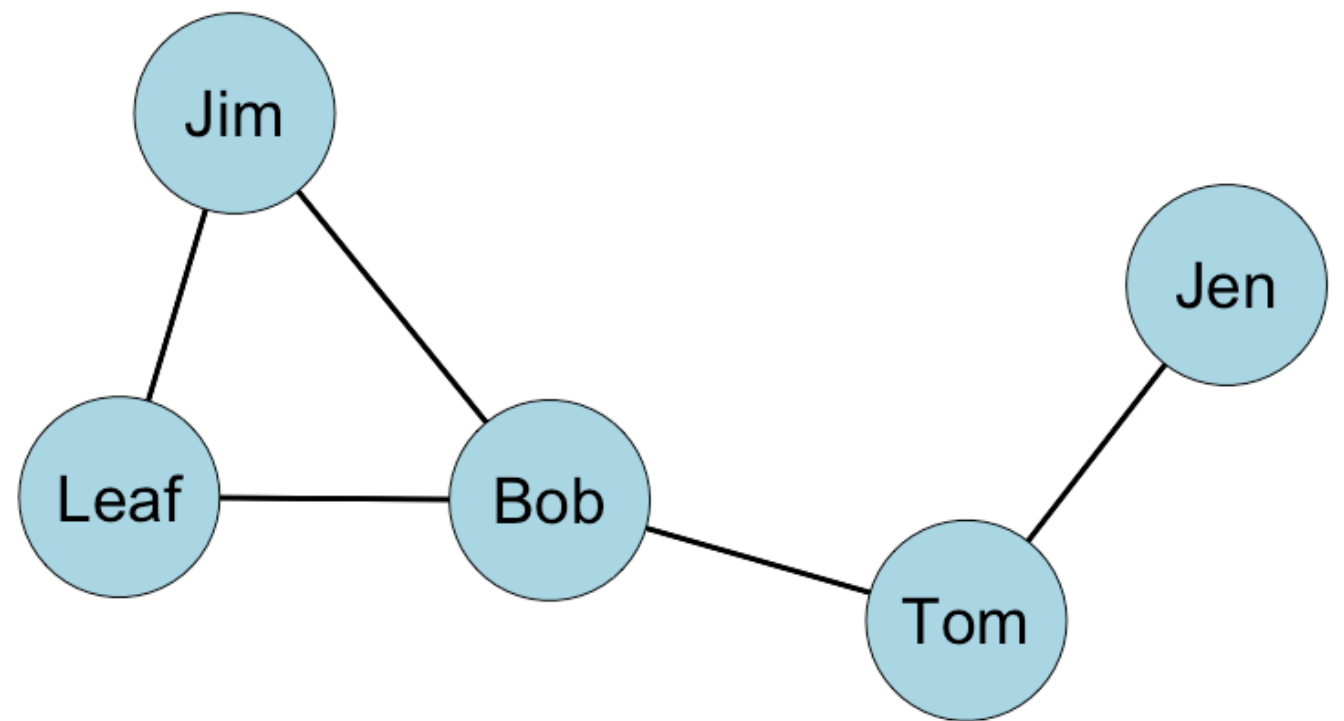
Tom, Bob, Jim



Example: Undirected, Binary Network

That gives us:

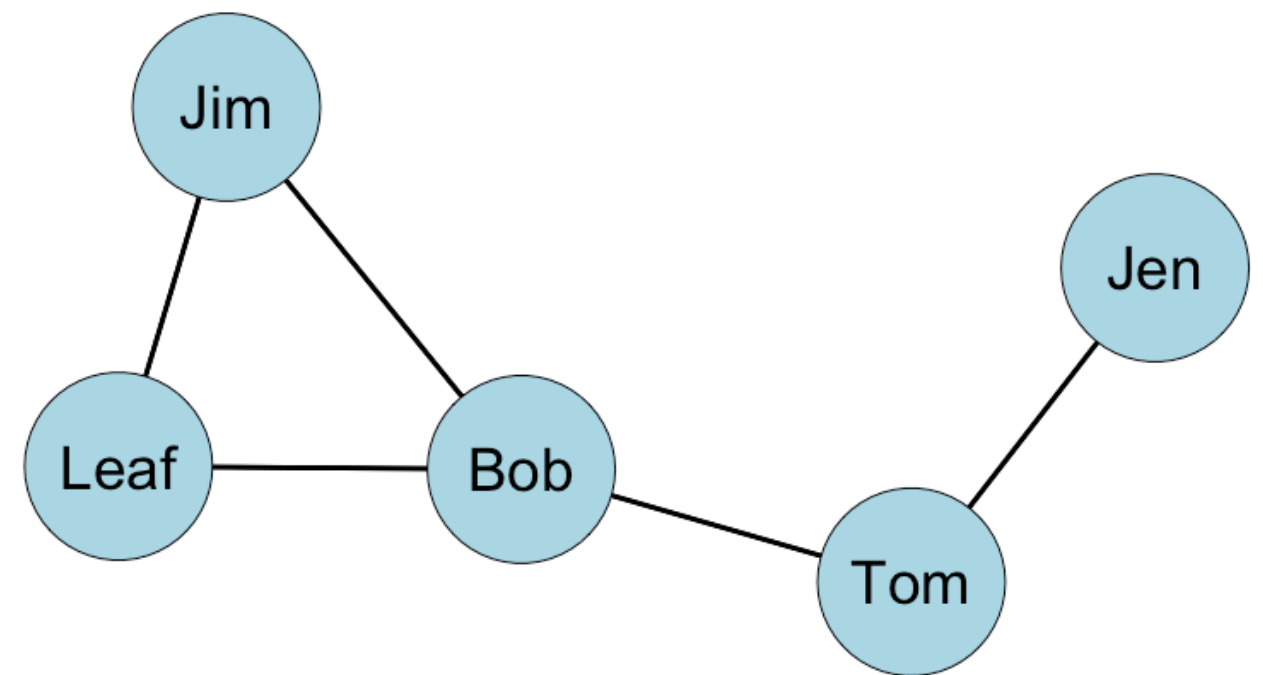
6 triplets total, and 3 are closed



Example: Undirected, Binary Network

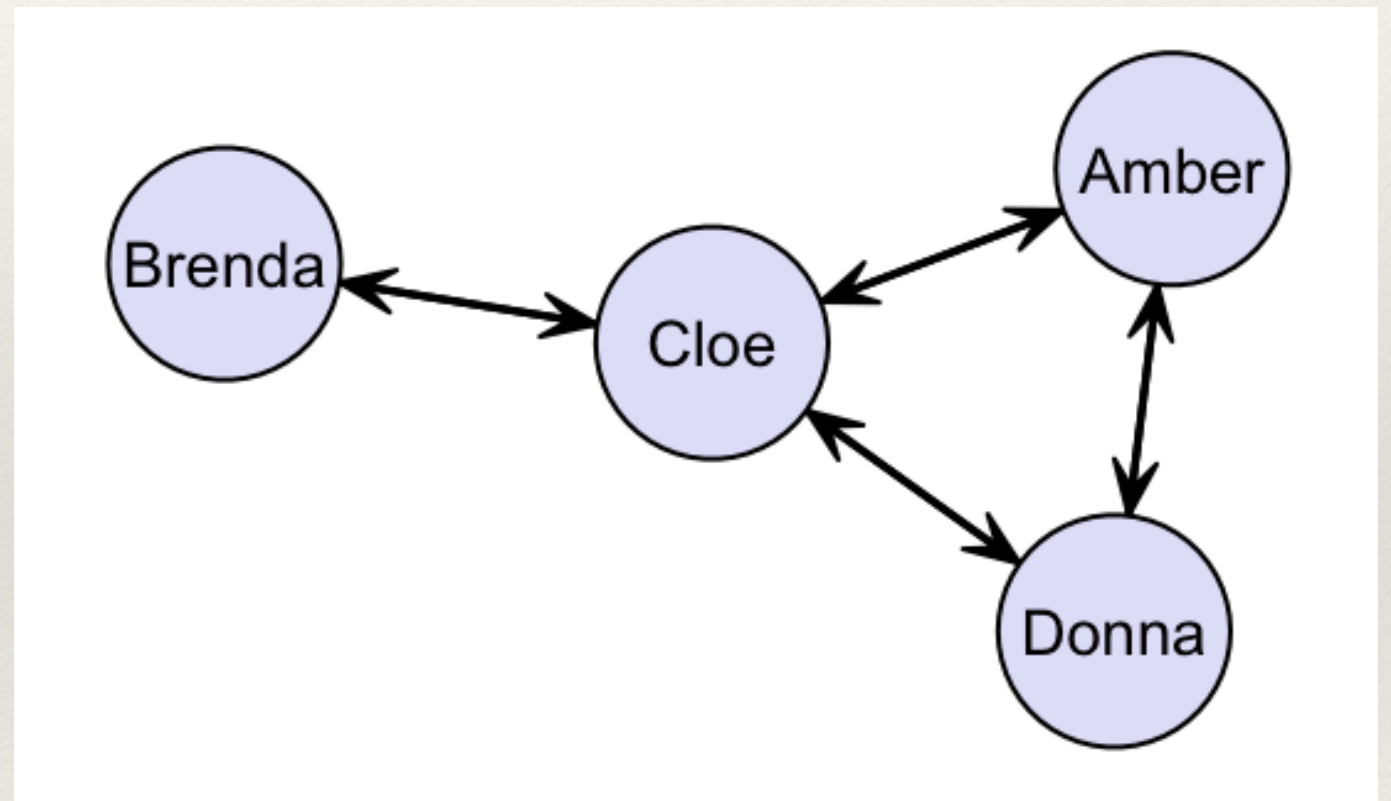
$$C = \frac{\text{number of closed triplets}}{\text{number of all triplets (open and closed)}}$$

$$C = \frac{3}{6} = 0.5$$



Example: Directed, Binary Network

*How many **closed** triplets are there in the graph?*



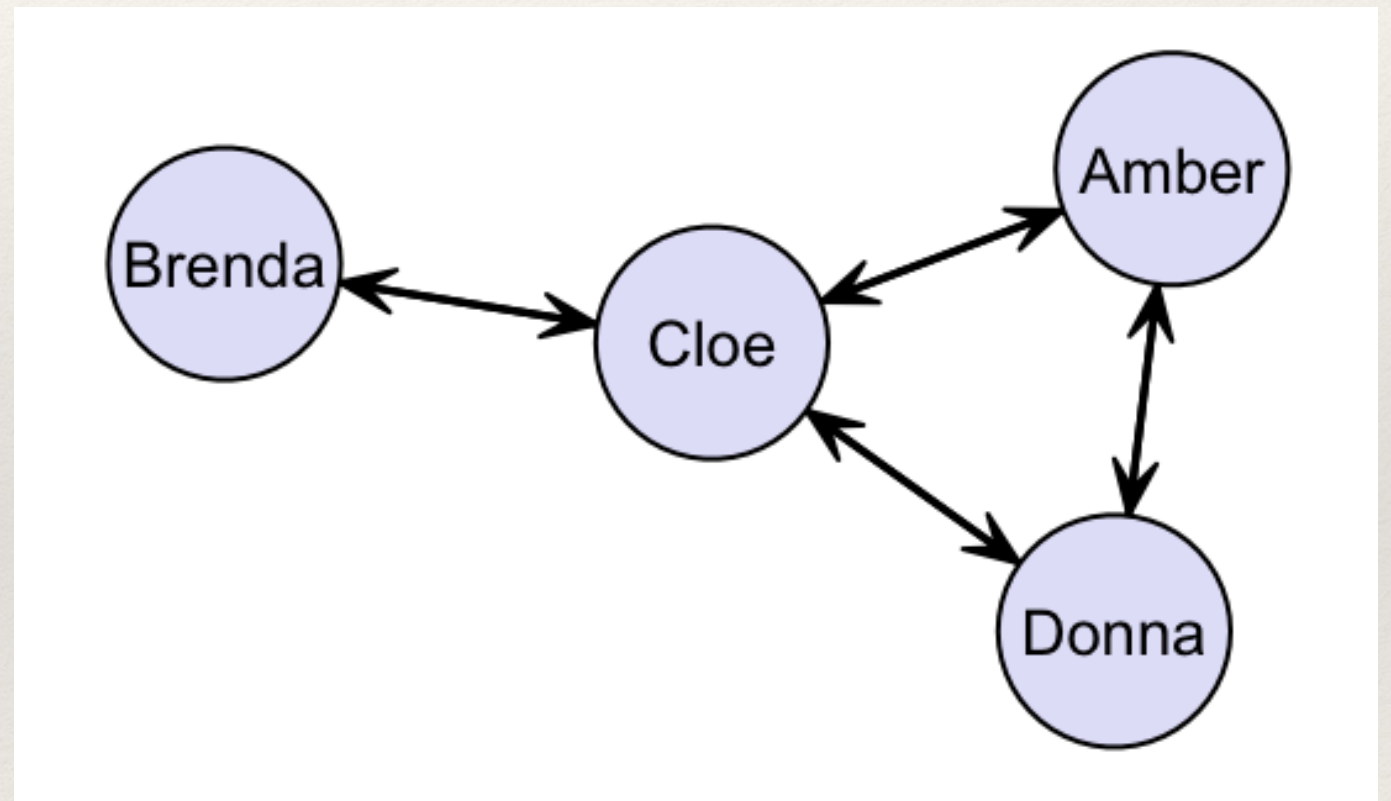
Example: Directed, Binary Network

*How many **closed** triplets are there in the graph?*

Amber, Cloe, Donna
Amber, Donna, Cloe

Cloe, Amber, Donna
Cloe, Donna, Amber

Donna, Cloe, Amber
Donna, Amber, Cloe

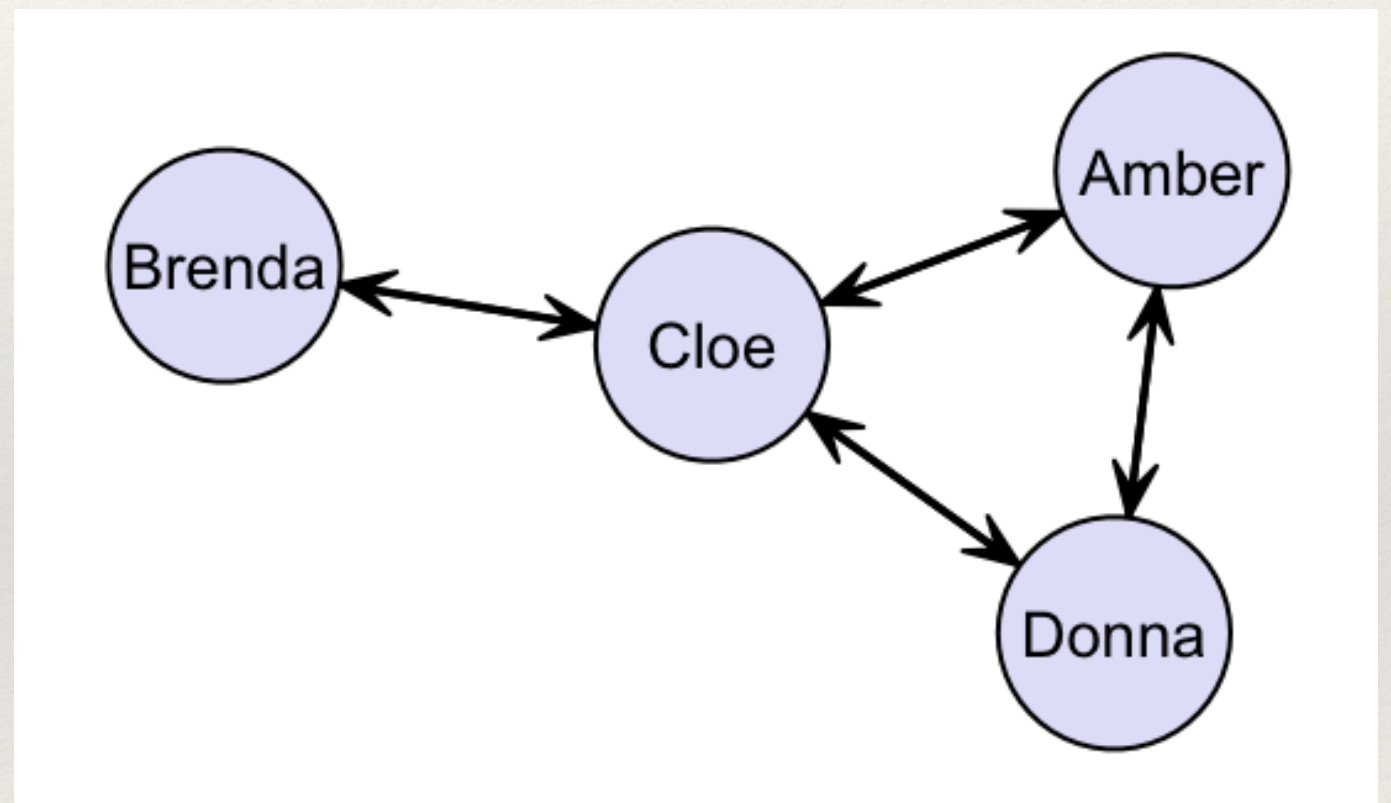


Example: Directed, Binary Network

*How many **open** triplets are there in the graph?*

Brenda, Cloe, Amber
Brenda, Cloe, Donna

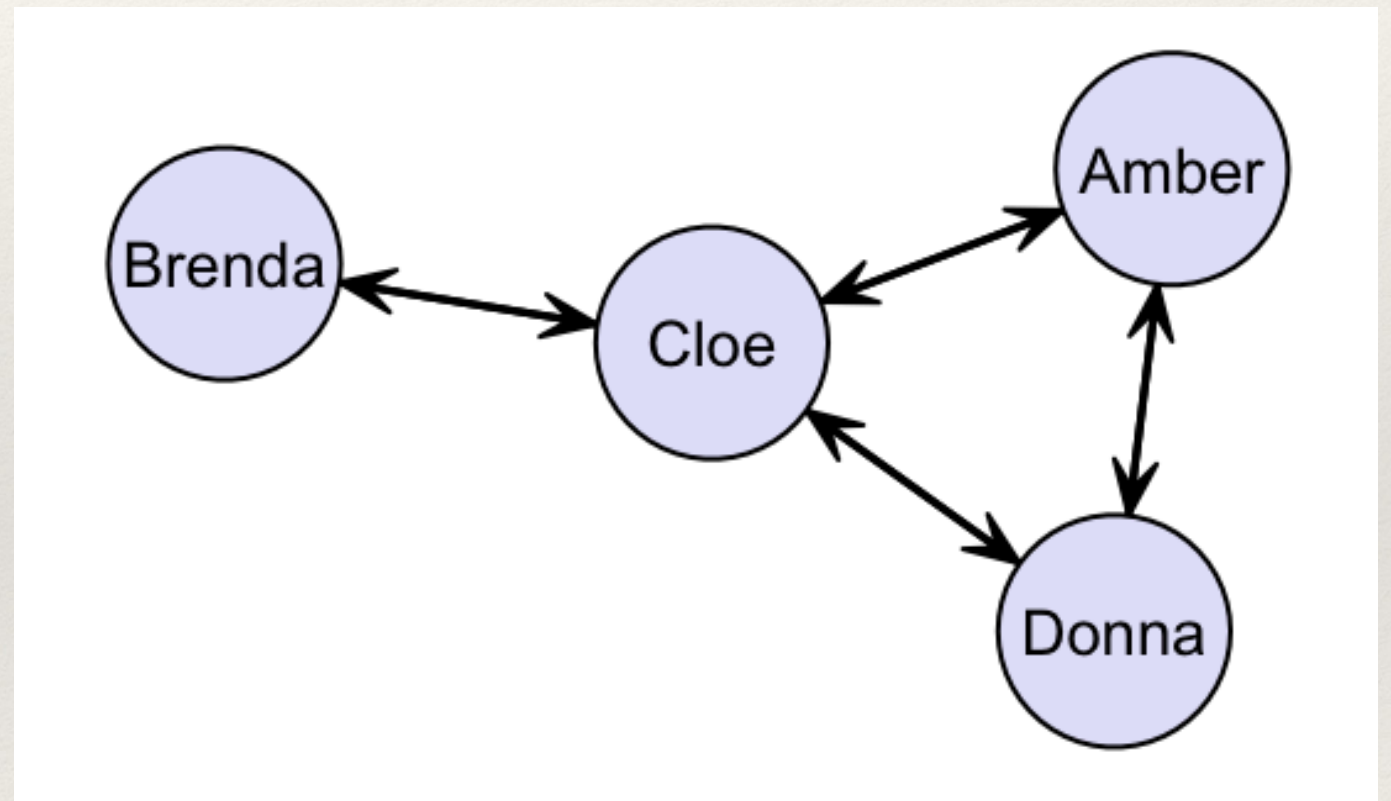
Amber, Cloe, Brenda
Donna, Cloe, Brenda



Example: Directed, Binary Network

That gives us:

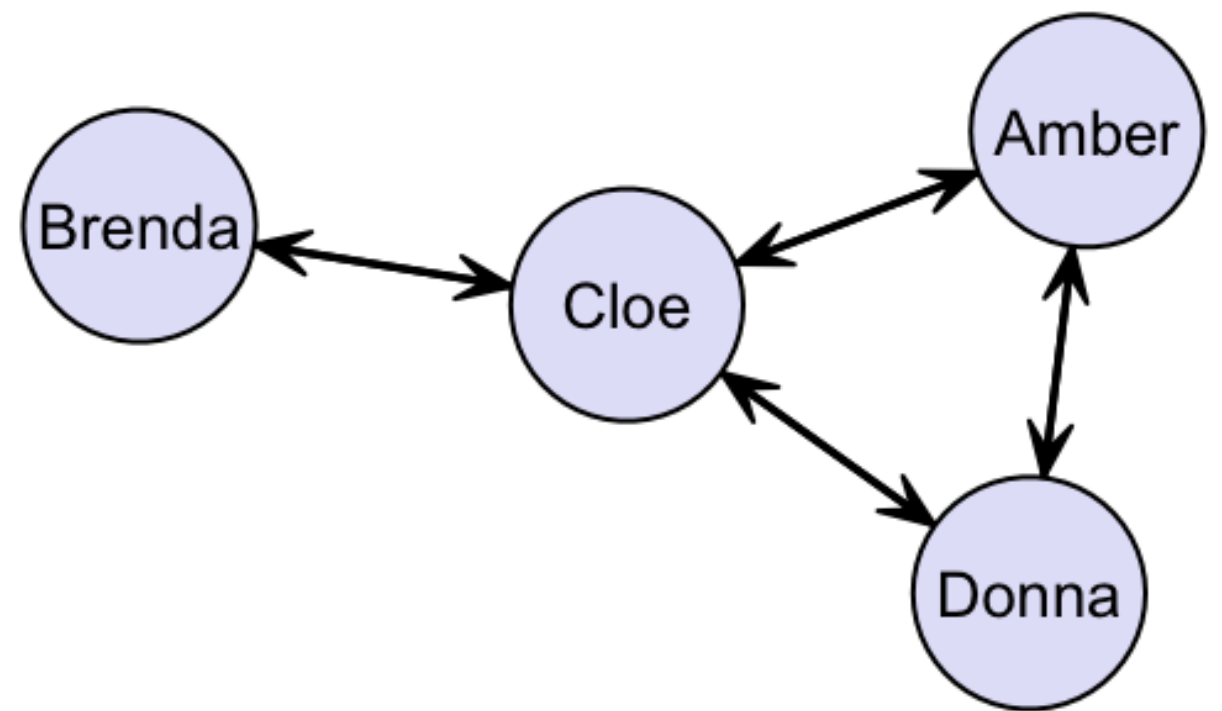
10 triplets total, and 6 are closed



Example: Directed, Binary Network

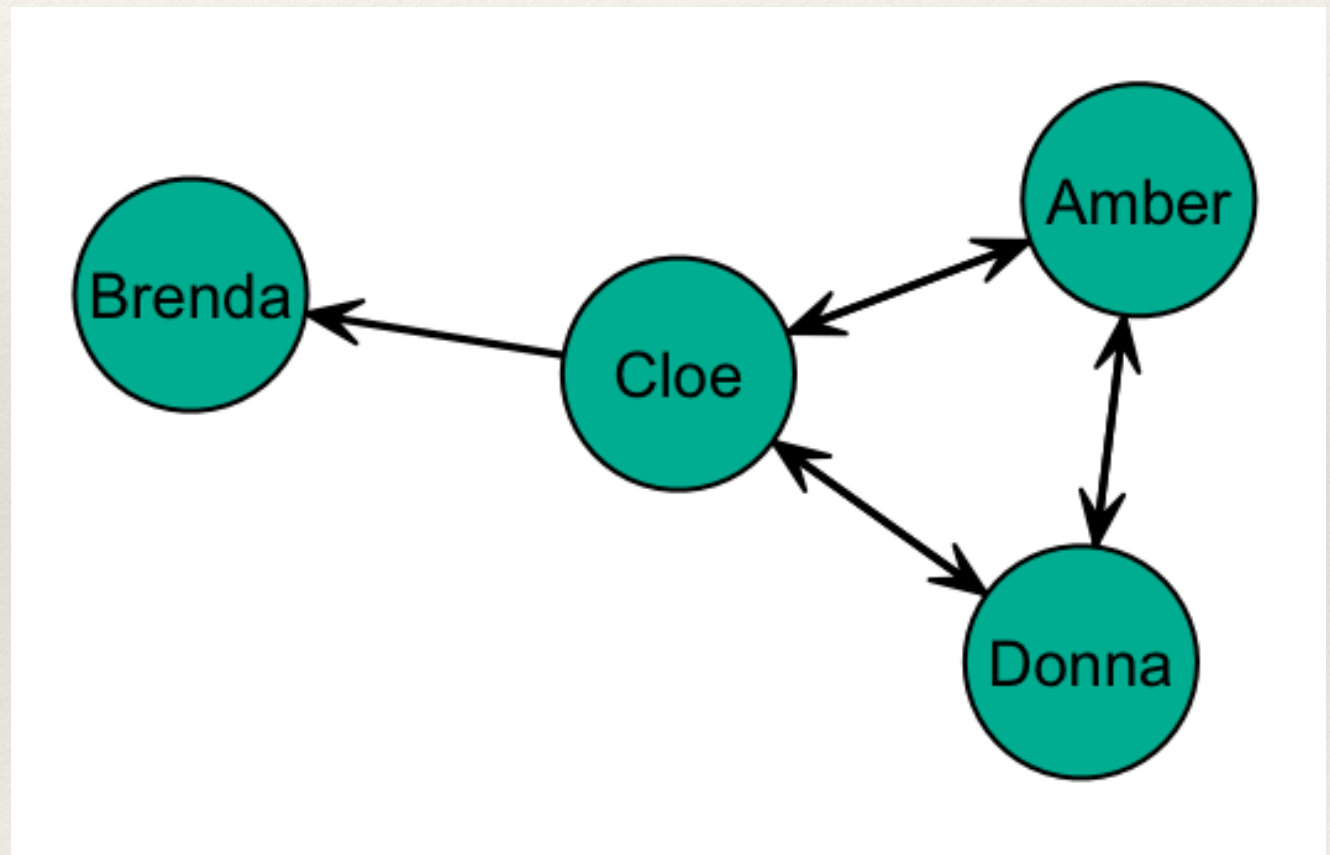
$$C = \frac{\text{number of closed triplets}}{\text{number of all triplets (open and closed)}}$$

$$C = \frac{6}{10} = 0.6$$



Example: Directed, Binary Network

*How many **closed** triplets are there in the graph?*



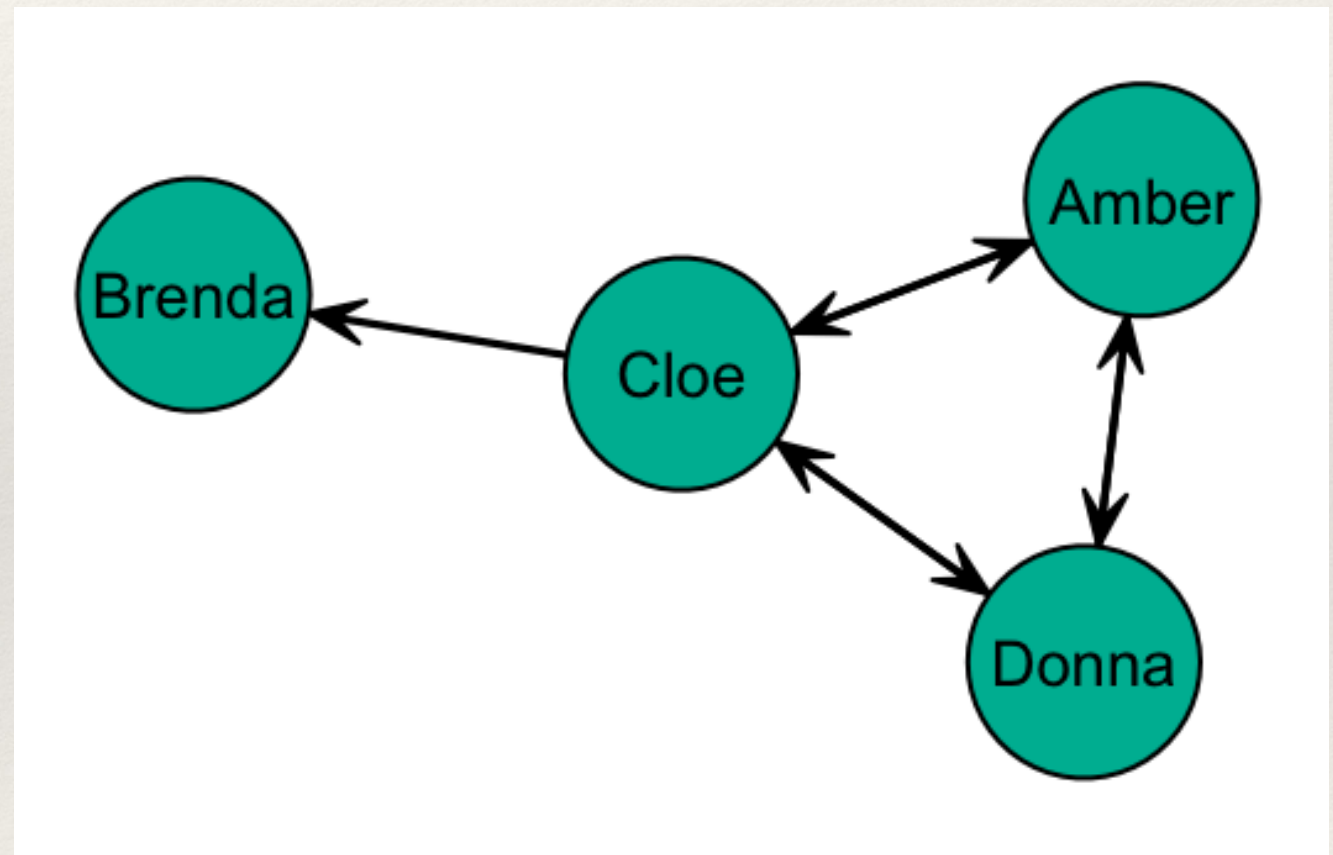
Example: Directed, Binary Network

*How many **closed** triplets are there in the graph?*

Amber, Cloe, Donna
Amber, Donna, Cloe

Cloe, Amber, Donna
Cloe, Donna, Amber

Donna, Cloe, Amber
Donna, Amber, Cloe

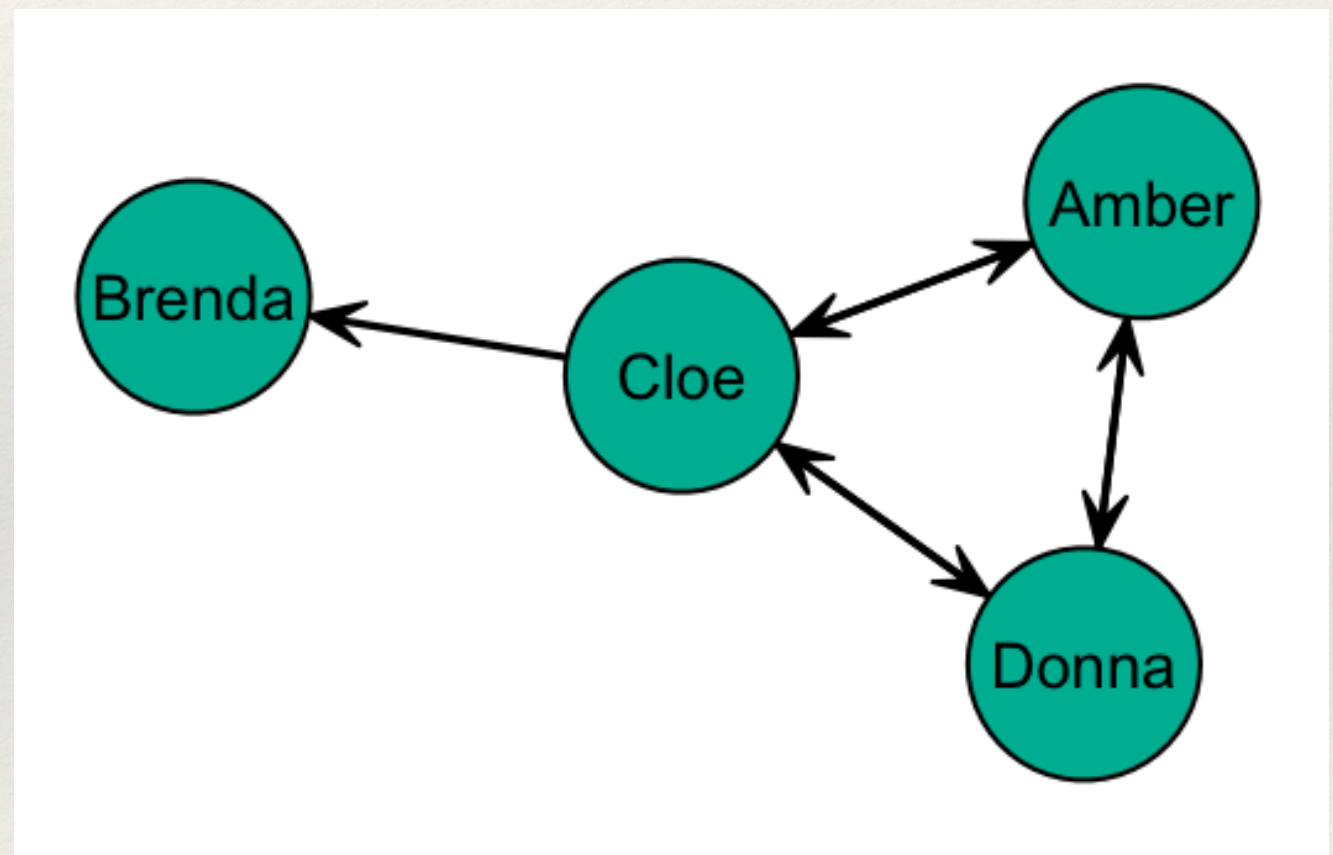


Example: Directed, Binary Network

How many *open* triplets are there in the graph?

Why?

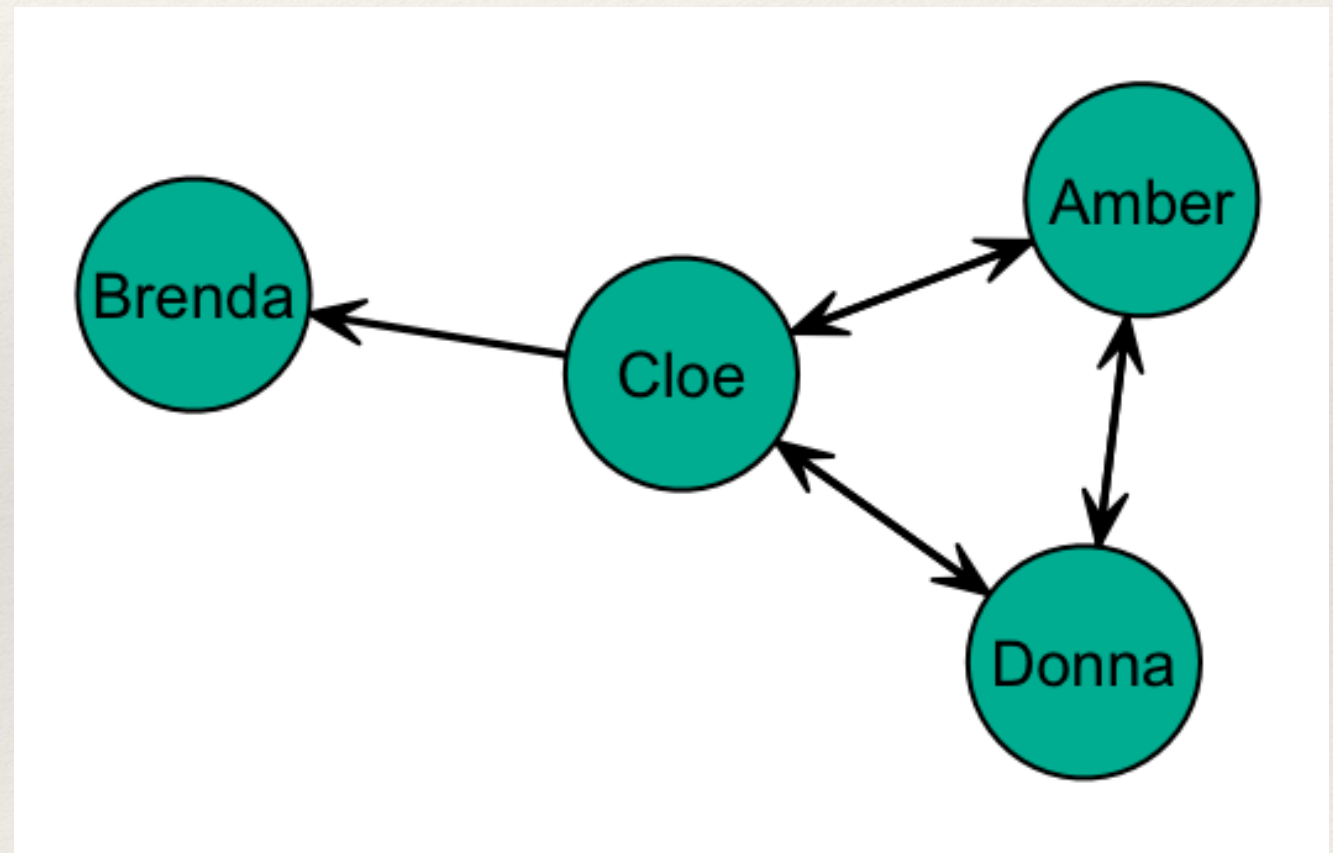
- Brenda, Cloe, Amber
- ~~Brenda, Cloe, Donna~~
- Amber, Cloe, Brenda
- ~~Donna, Cloe, Brenda~~



Example: Directed, Binary Network

That gives us:

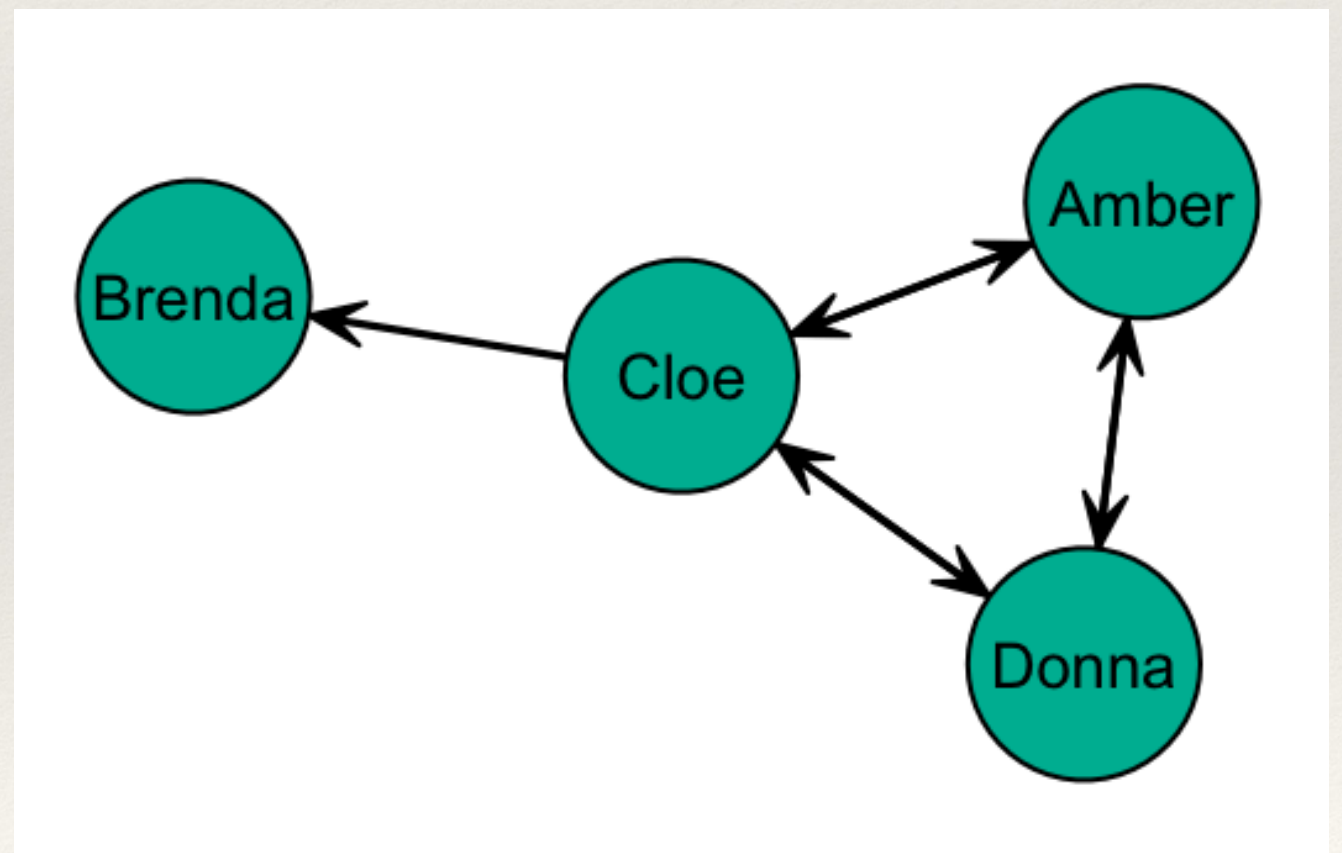
8 triplets total, and 6 are closed



Example: Directed, Binary Network

$$C = \frac{\text{number of closed triplets}}{\text{number of all triplets (open and closed)}}$$

$$C = \frac{6}{8} = 0.75$$



Summary

- ❖ So what?
 - ❖ We have covered five ways to quickly describe a network:
 - ❖ Size, density, components, diameter, transitivity
 - ❖ These measures help us understand the structure of a network and allow us to make comparisons across networks.

Learning Goals

- ❖ Now, you should be able to answer this question:
 - ❖ How can I describe some simple features of a network?

Questions?