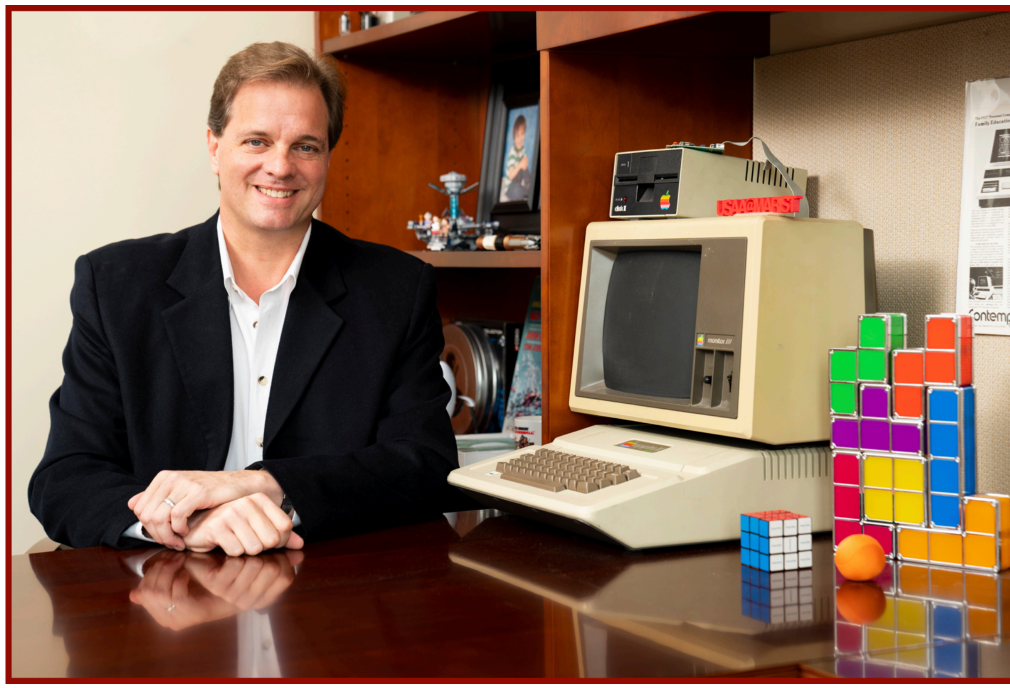
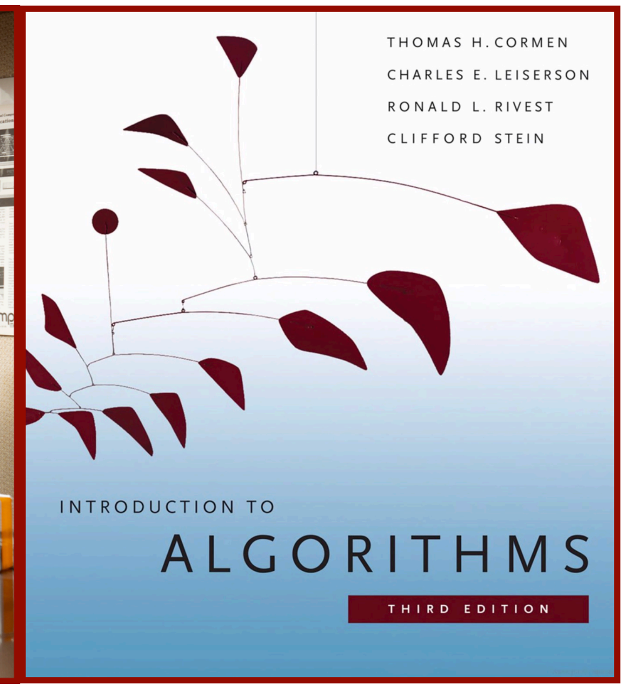
Graphs

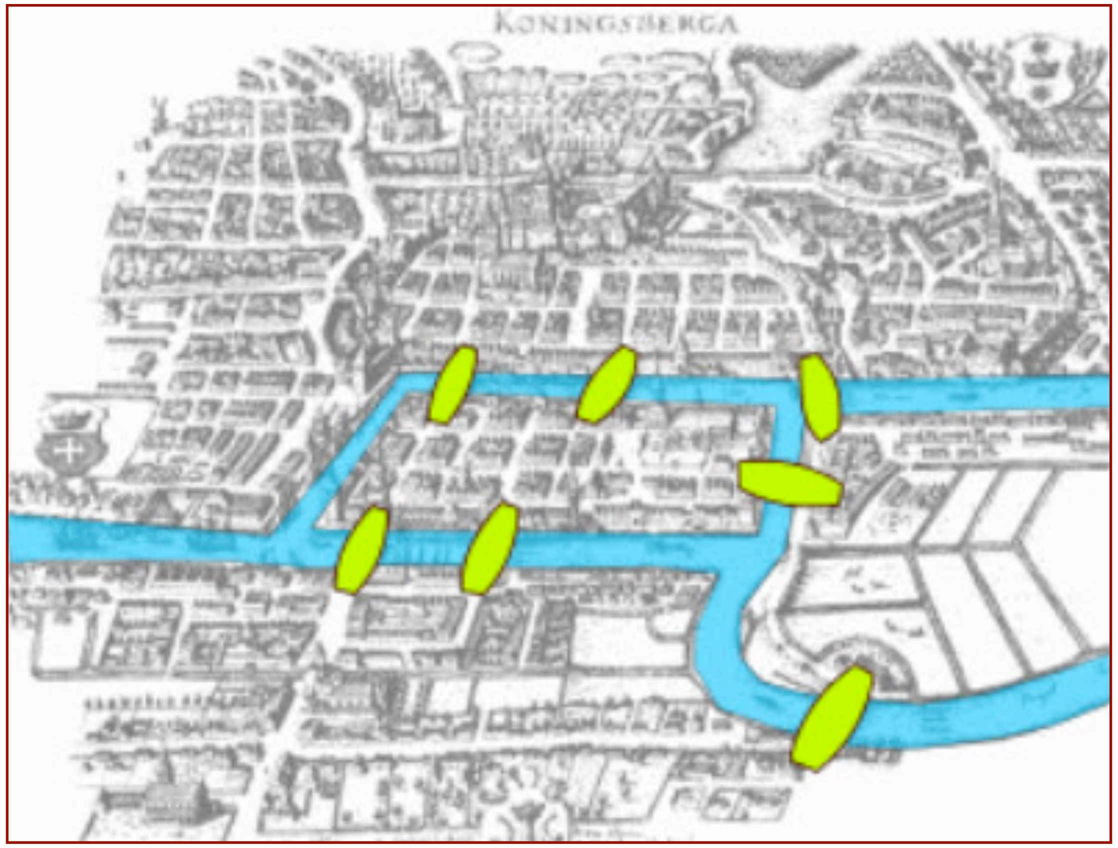




Alan G. Labouseur, Ph.D.
Alan.Labouseur@Marist.edu

A long time ago....

The city of Königsberg in Prussia (now Kaliningrad, Russia) was set on both sides of the Pregel River, and included two large islands which were connected to each other and the mainland by seven bridges.

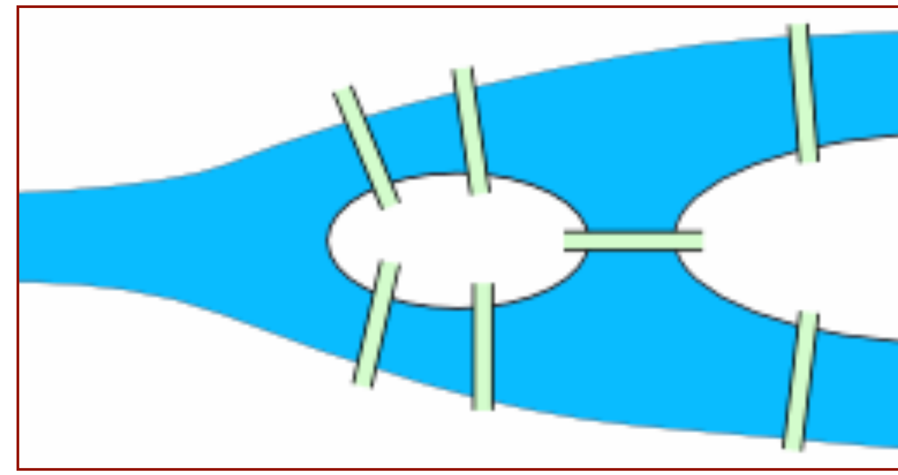


In 1752, Leonhard Euler wondered if he could wander through the entire city crossing each bridge only once.

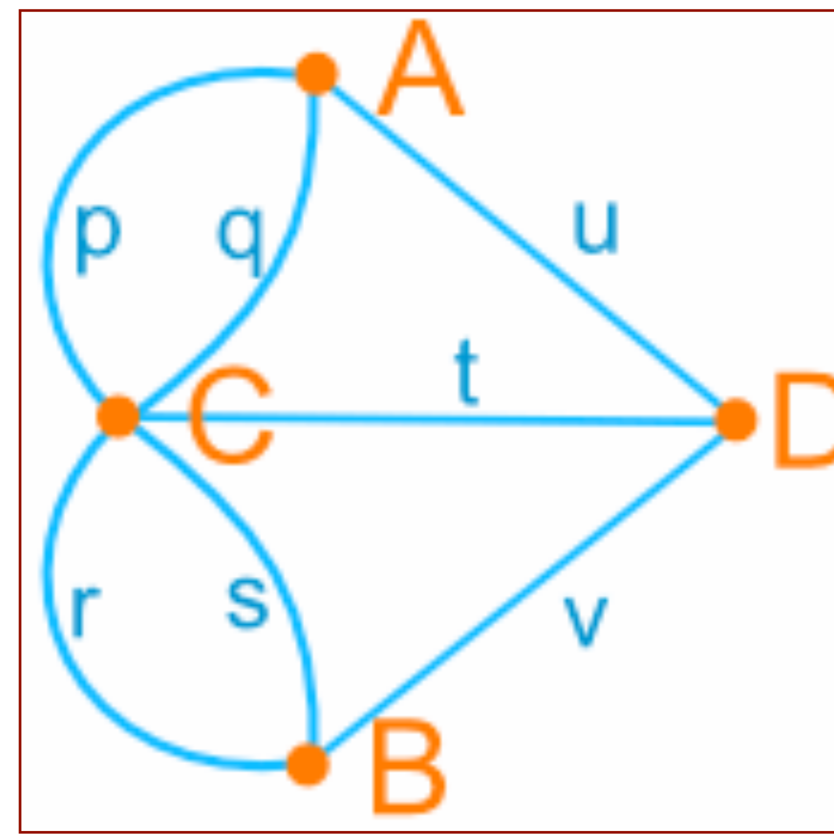
He ended up inventing Graph Theory.

A long time ago....

We can simplify the topography a little to make the problem more clear.



We can abstract it to nothing more than vertices and edges.

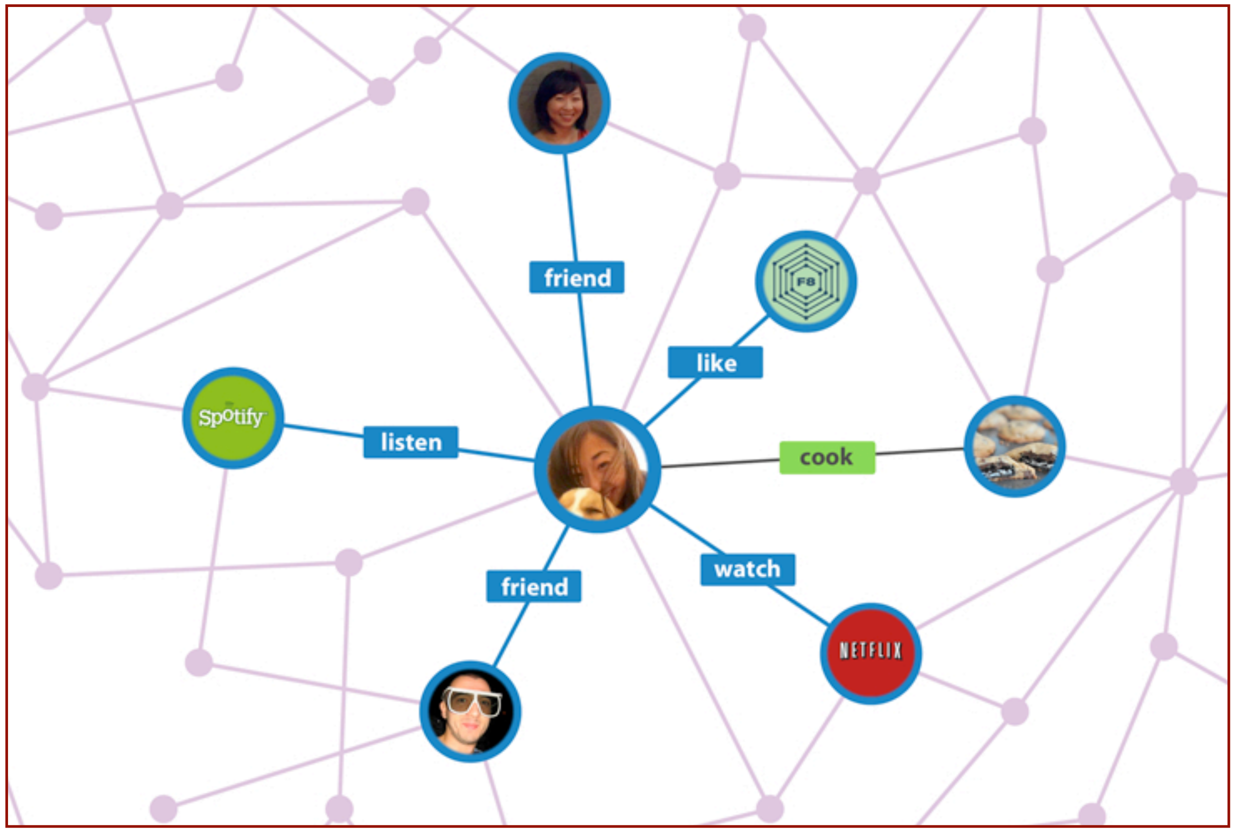


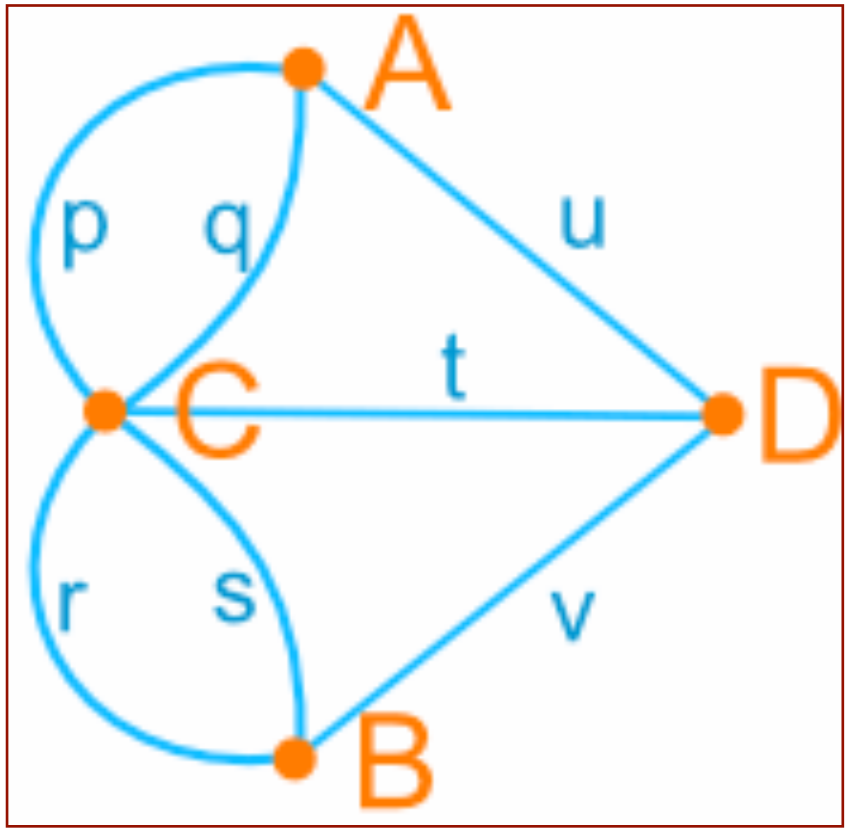
Now we have a graph, "G", which is comprised of a set of vertices ("V") and a set of edges among them ("E"), or G(V,E) for short.

Graphs

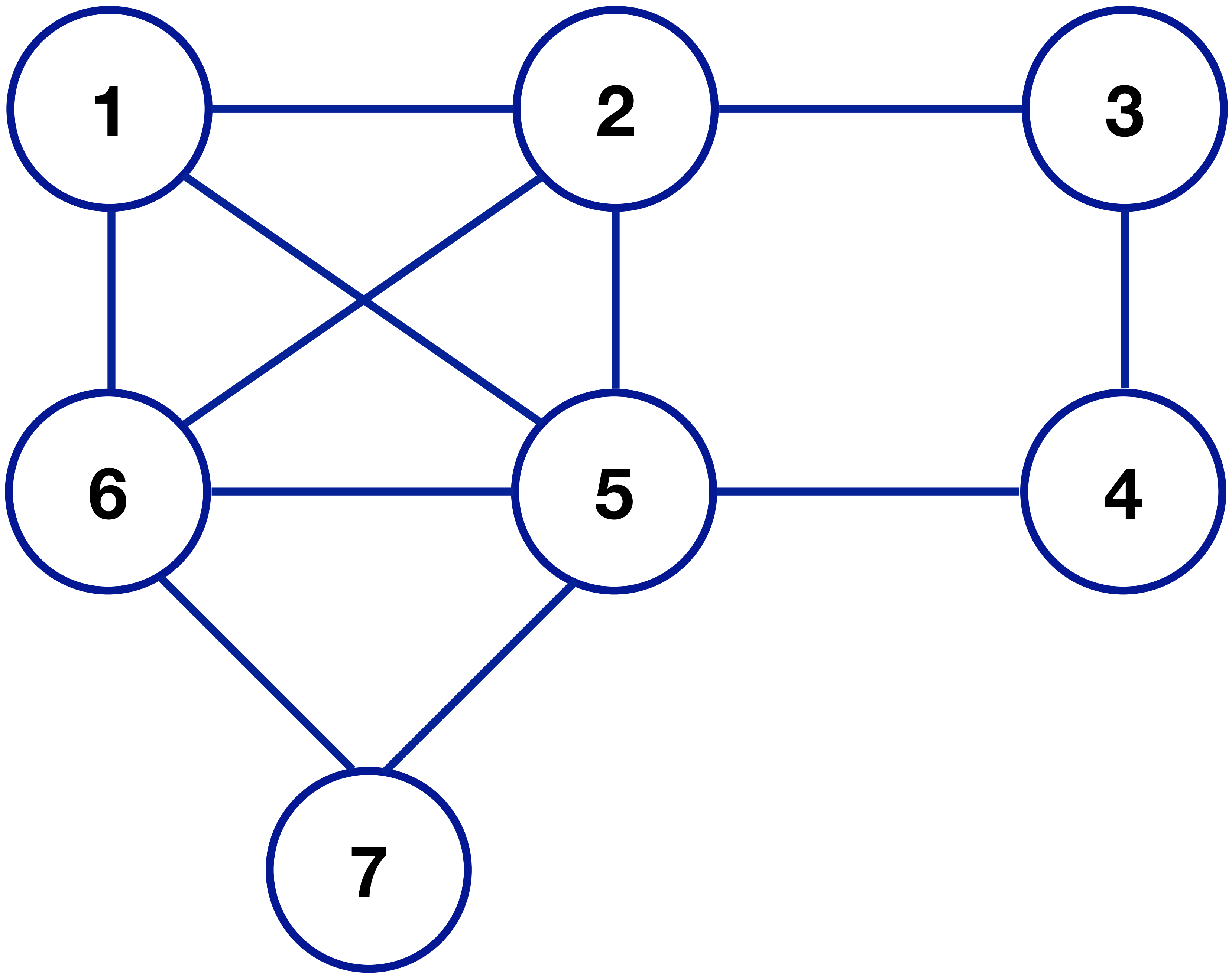
It turns out that graphs are ridiculously useful.

Euler's "tour" of bridges over the Pregel river kicked off a branch of mathematics without which we would not have social media or navigation systems all these years later.



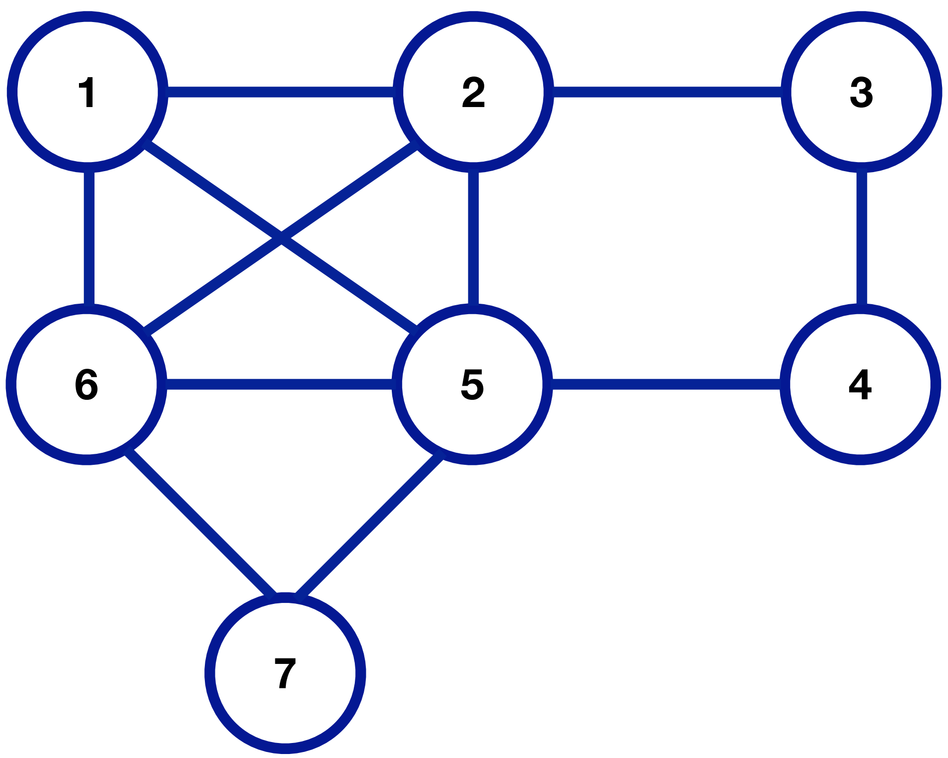


Graphs



Graphs

Graph . . .

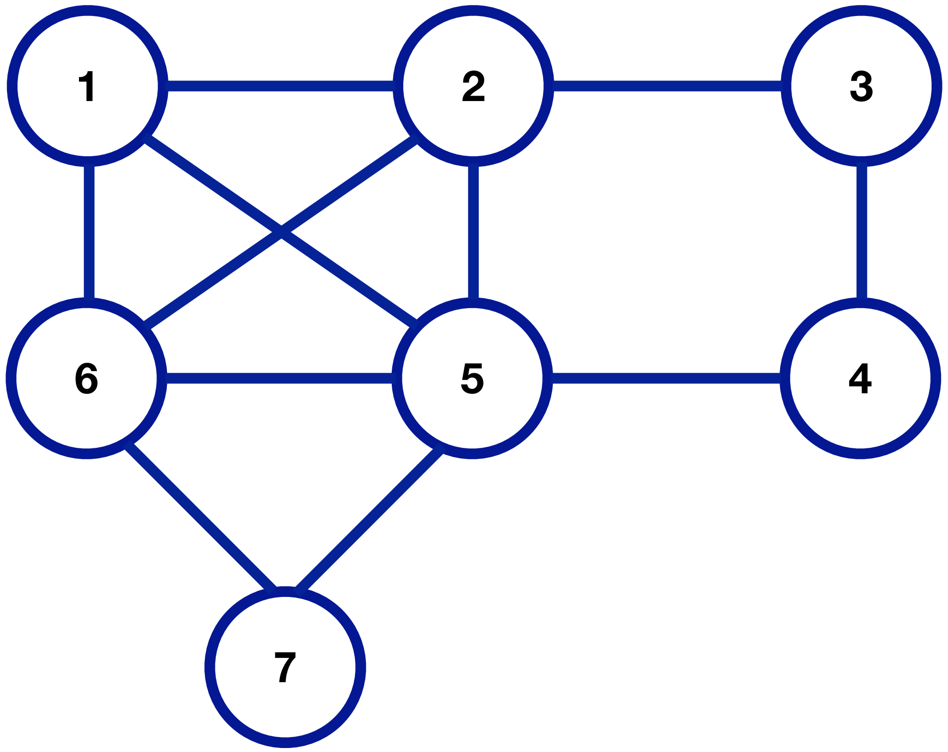


as Matrix

	1	2	3	4	5	6	7
1	.	1	.	.	1	1	.
2	1	.	1	.	1	1	.
3	.	1	.	1	.	.	.
4	.	.	1	.	1	.	.
5	1	1	.	1	.	1	1
6	1	1	.	.	1	.	1
7	.	.	.	.	1	1	.

Graphs

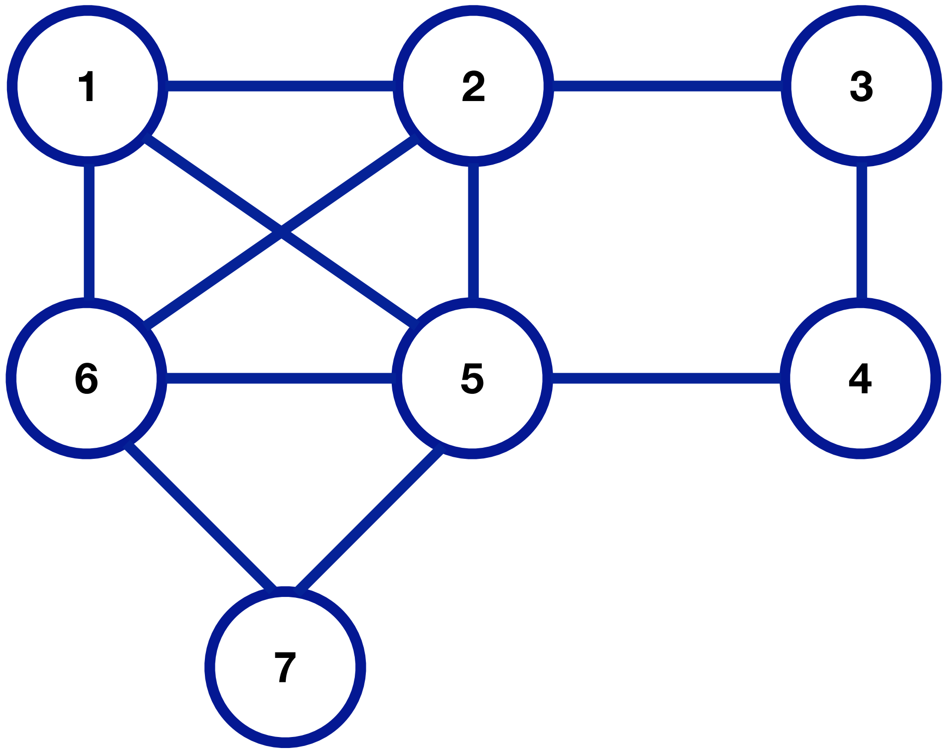
Graph . . .



as Adjacency List

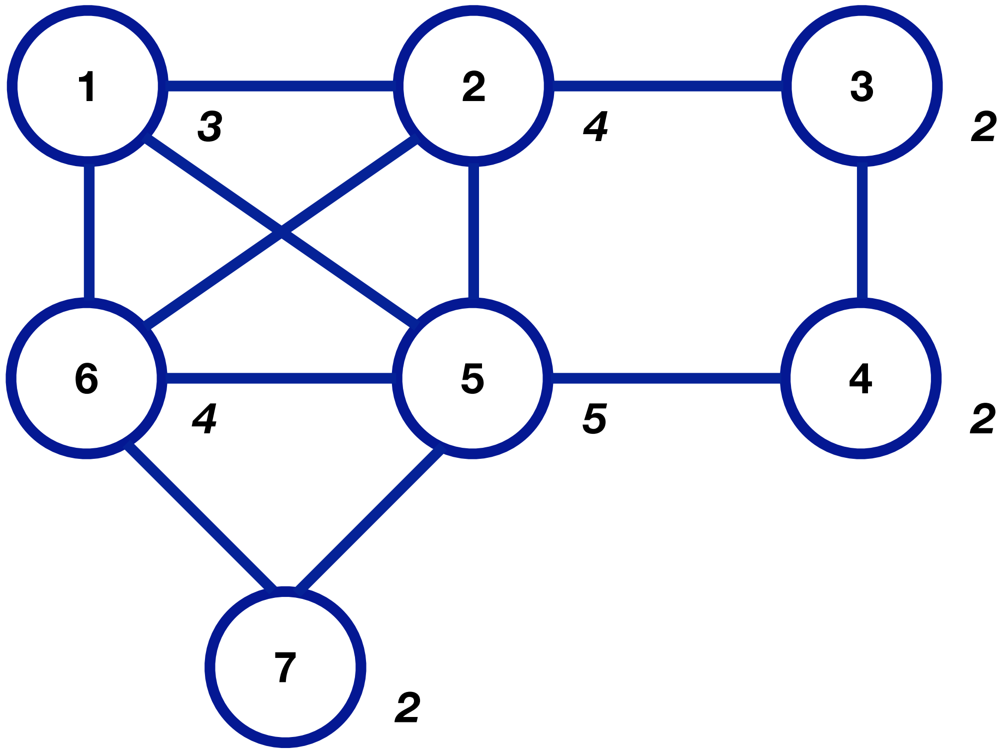
```
[1]  2  5  6
[2]  1  3  5  6
[3]  2  4
[4]  3  5
[5]  1  2  4  6  7
[6]  1  2  5  7
[7]  5  6
```

Vertex Degree



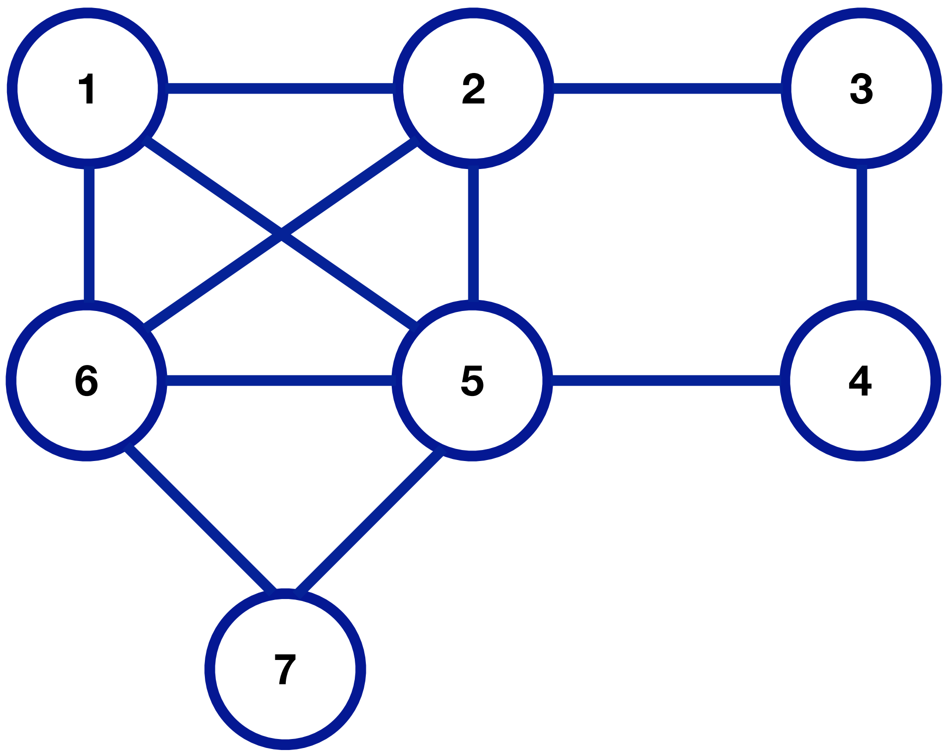
The **degree** of a vertex **v** in a graph $G = (V, E)$ is the number of edges incident on **v**.)

Vertex Degree



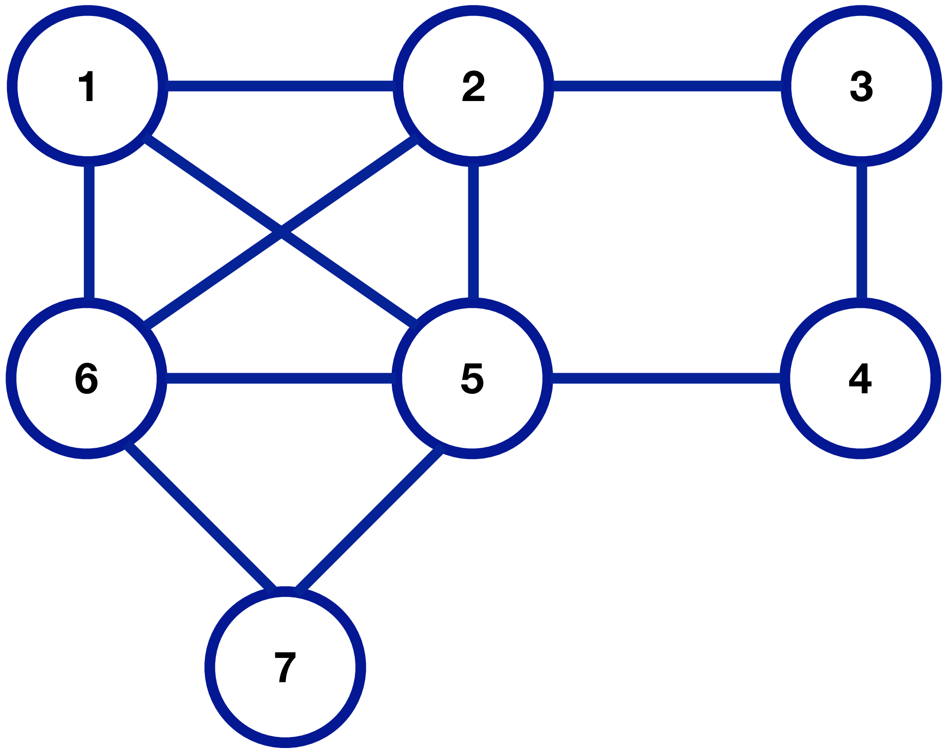
The **degree** of a vertex **v** in a graph *G* = (*V*, *E*) is the number of edges incident on **v**.)

Independent Sets



An **independent set** in a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that, for all *u*,*v* in *V'*, the edge {*u*,*v*} is **not** in E. (I.e., no two vertices in *V'* are adjacent.)

Independent Sets

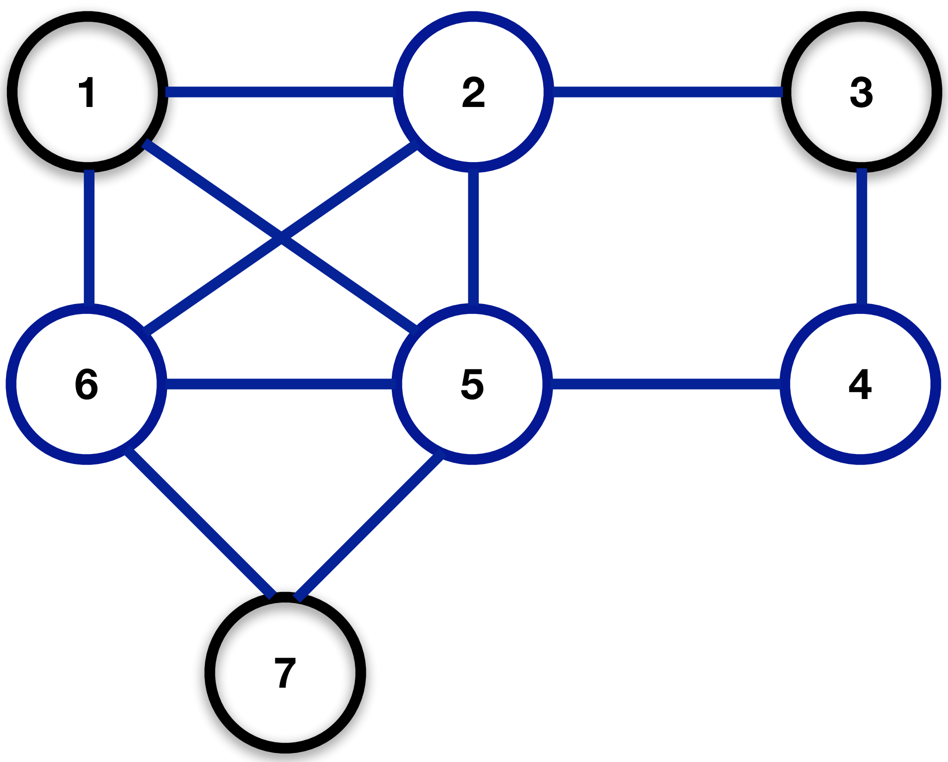


An **independent set** in a graph $G = (V, E)$ is a subset $V' \subseteq V$ such that, for all u,v in V', the edge $\{u,v\}$ is **not** in E. (I.e., no two vertices in V' are adjacent.)

A **maximum independent set** is an independent set of the largest possible cardinality.

Can you find one?

Independent Sets



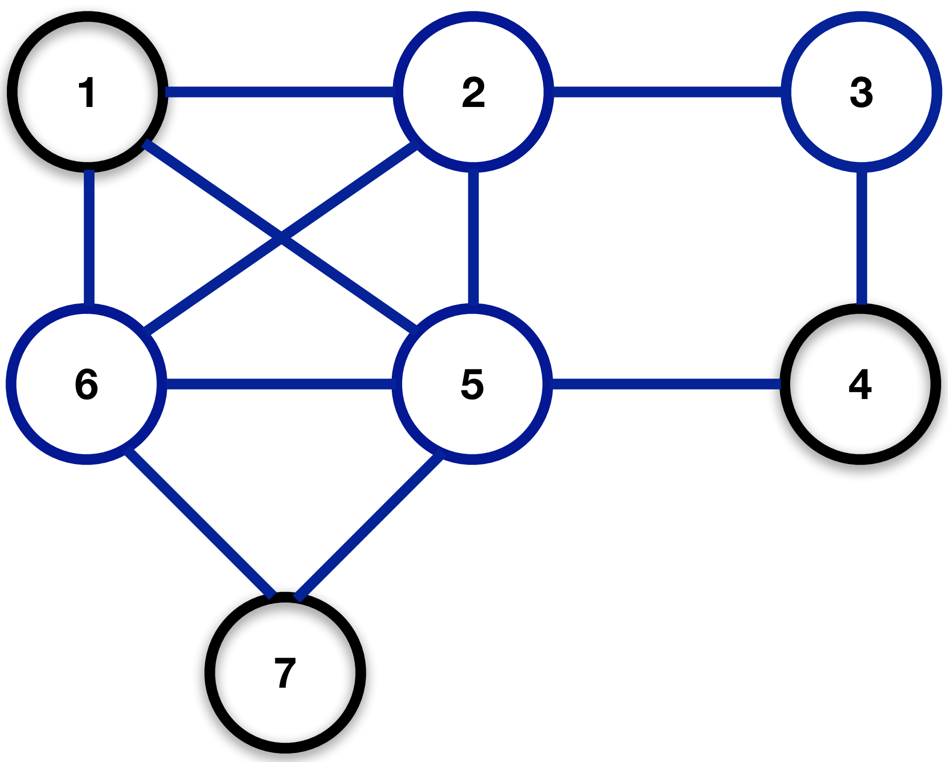
An **independent set** in a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that, for all *u*,*v* in *V'*, the edge {*u*,*v*} is **not** in E. (I.e., no two vertices in *V'* are adjacent.)

A **maximum independent set** is an independent set of the largest possible cardinality.

Can you find one? 1 3 7

Can you find two?

Independent Sets



An **independent set** in a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that, for all *u*,*v* in *V'*, the edge {*u*,*v*} is **not** in E. (I.e., no two vertices in *V'* are adjacent.)

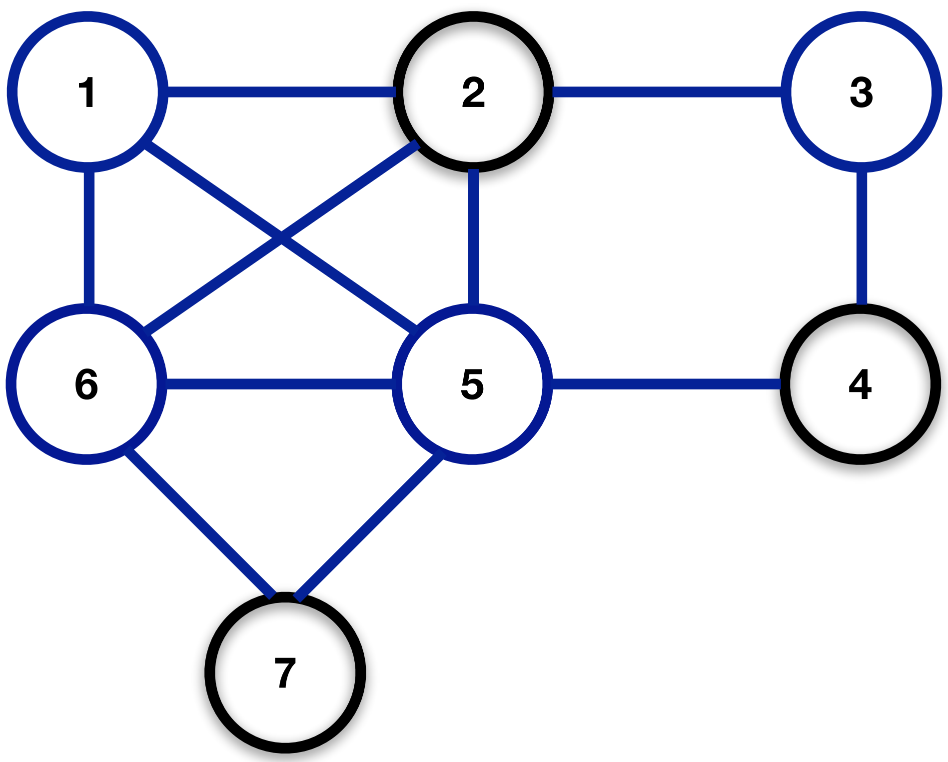
A **maximum independent set** is an independent set of the largest possible cardinality.

Can you find one? 1 3 7

Can you find two? 1 4 7

Can you find three?

Independent Sets



An **independent set** in a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that, for all *u*,*v* in *V'*, the edge {*u*,*v*} is **not** in E. (I.e., no two vertices in *V'* are adjacent.)

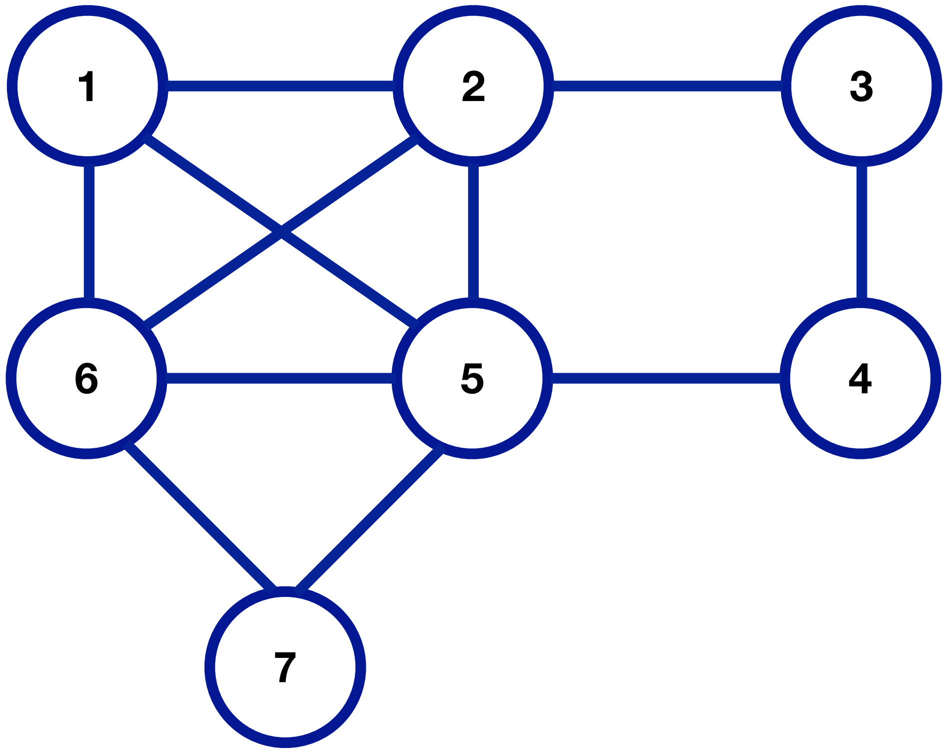
A **maximum independent set** is an independent set of the largest possible cardinality.

Can you find one? 1 3 7

Can you find two? 1 4 7

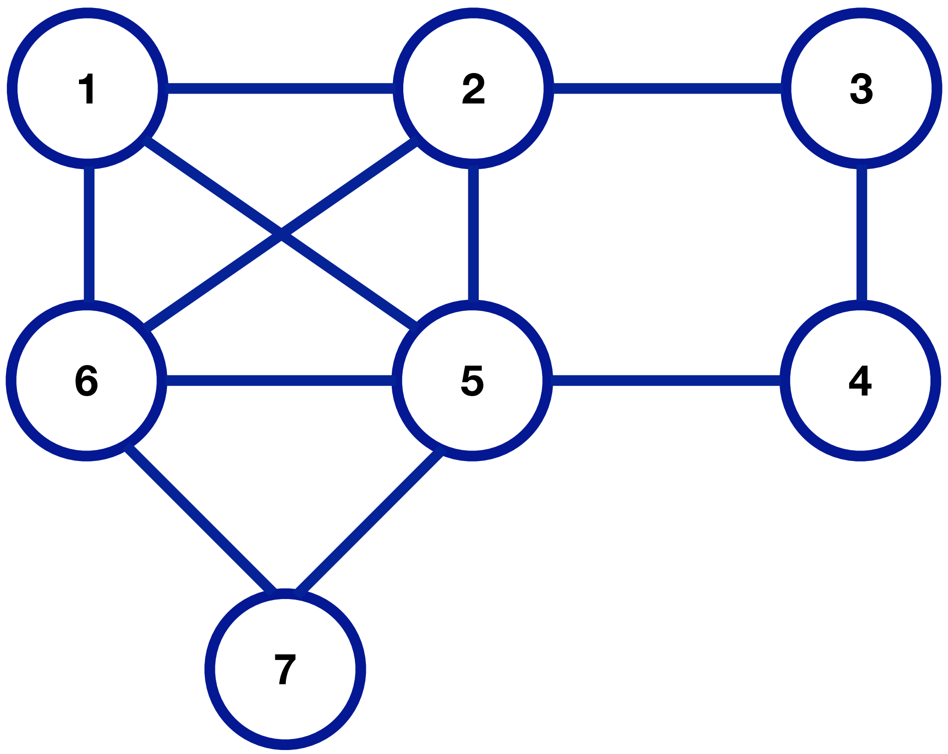
Can you find three? 2 4 7

Vertex Cover



A **vertex cover** of a graph $G = (V, E)$ is a subset $V' \subseteq V$ such that if $\{u,v\}$ is an edge in G, then either u is in V' or v is in V' or both are. (I.e., it's a set of vertices V' such that each edge of graph G has at least one member of V' as an endpoint.)

Vertex Cover

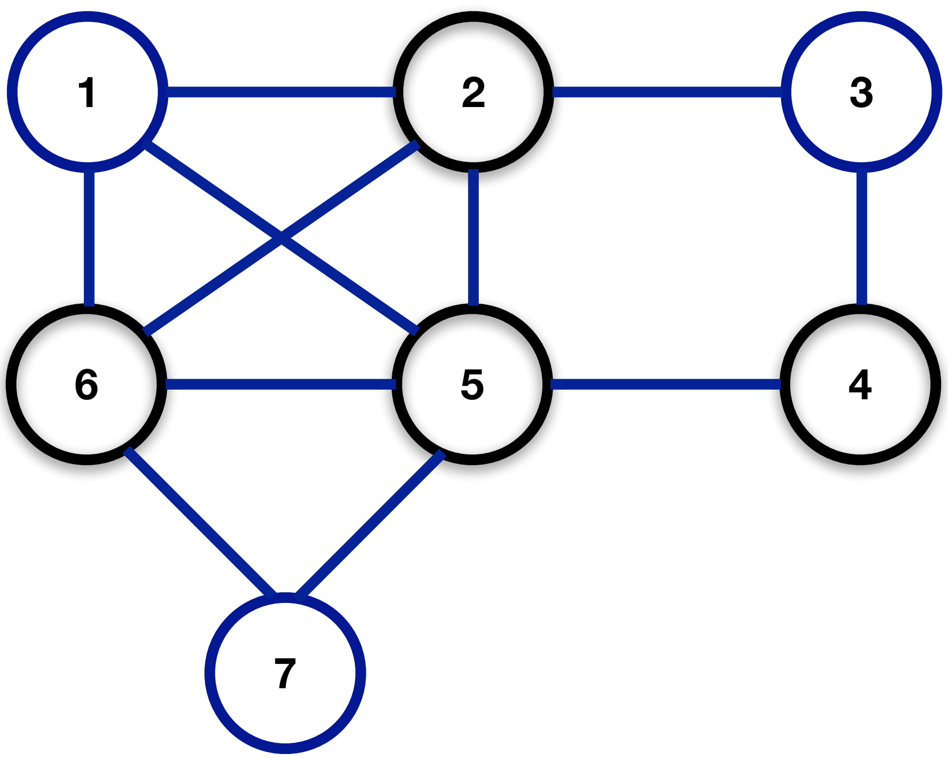


A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one?

Vertex Cover



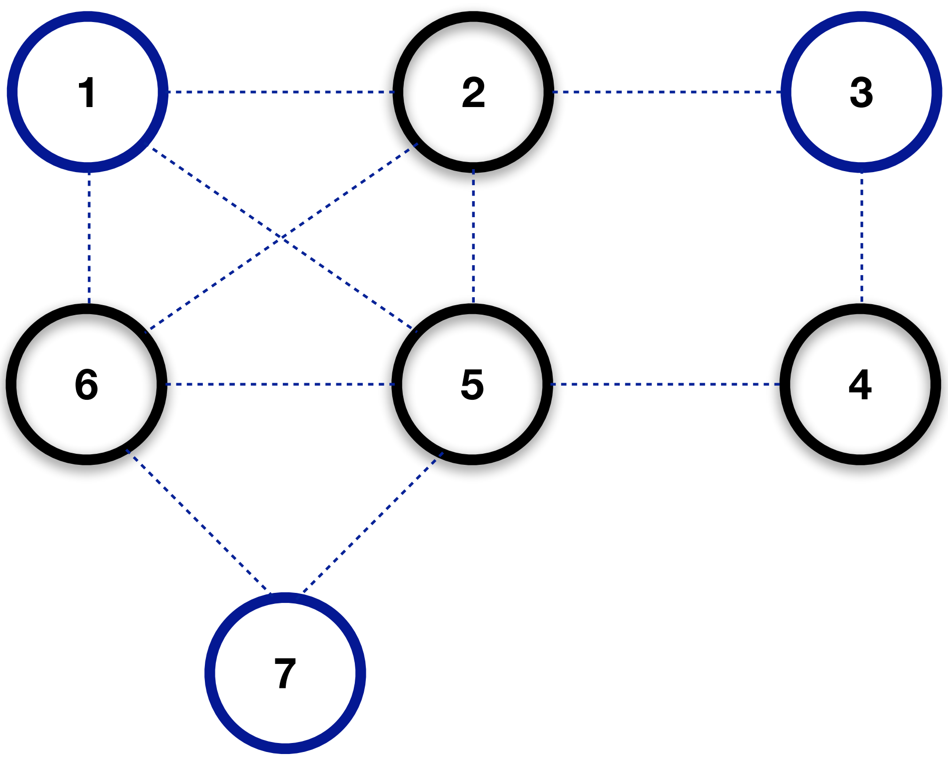
A **vertex cover** of a graph $G = (V, E)$ is a subset $V' \subseteq V$ such that if $\{u,v\}$ is an edge in G, then either u is in V' or v is in V' or both are. (I.e., it's a set of vertices V' such that each edge of graph G has at least one member of V' as an endpoint.)

A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 6

Really?

Vertex Cover



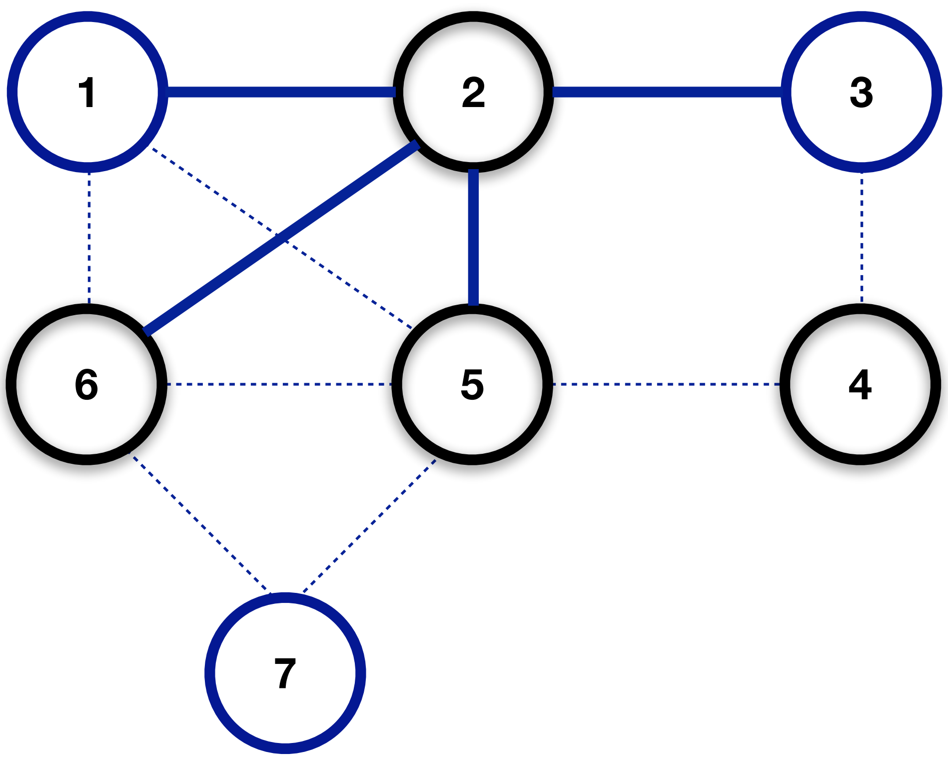
A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 6

Really? Let's test it.

Vertex Cover



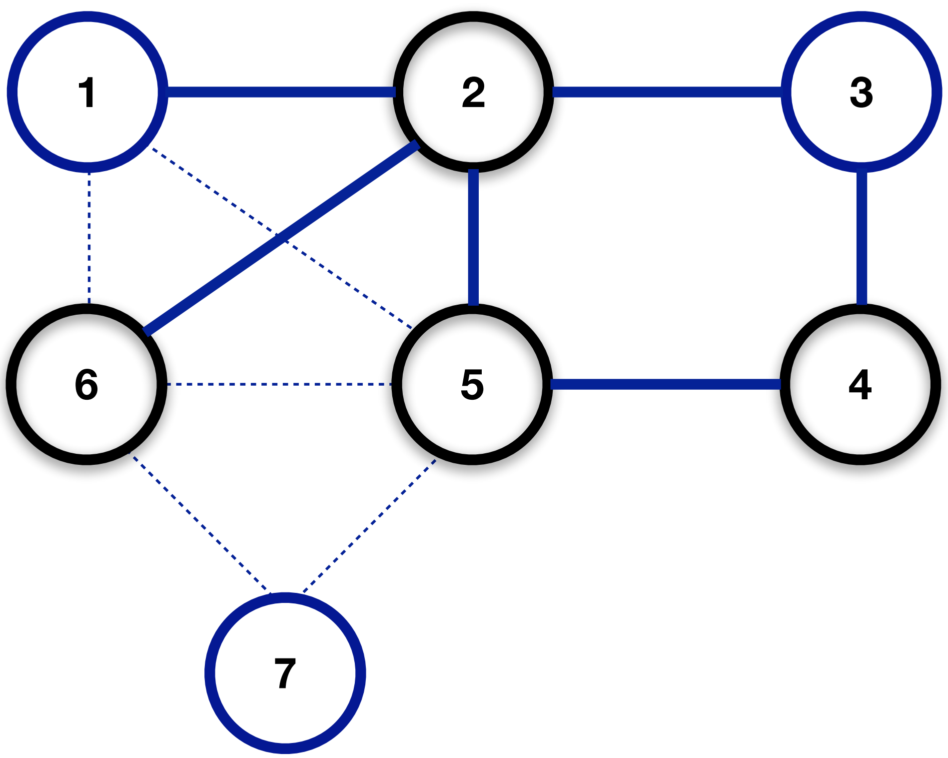
A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? **2** 4 5 6

Really? Let's test it.
Vertex 2 covers 4 edges.

Vertex Cover



A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

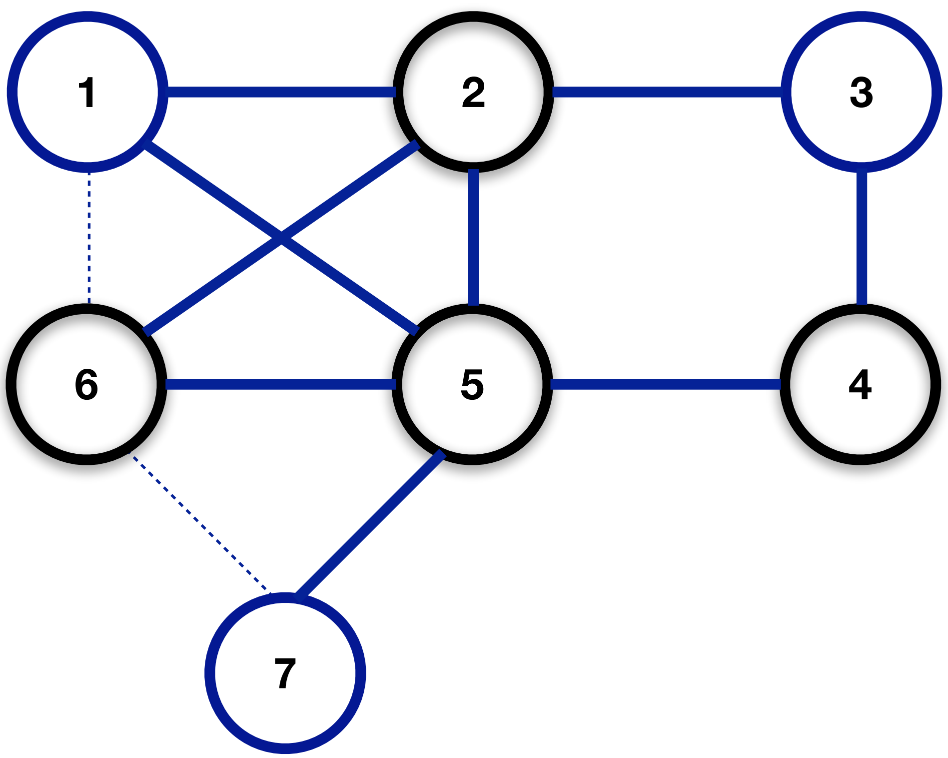
Can you find one? 2 **4** 5 6

Really? Let's test it.

Vertex 2 covers 4 edges.

Vertex 4 covers 2 more edges.

Vertex Cover



A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

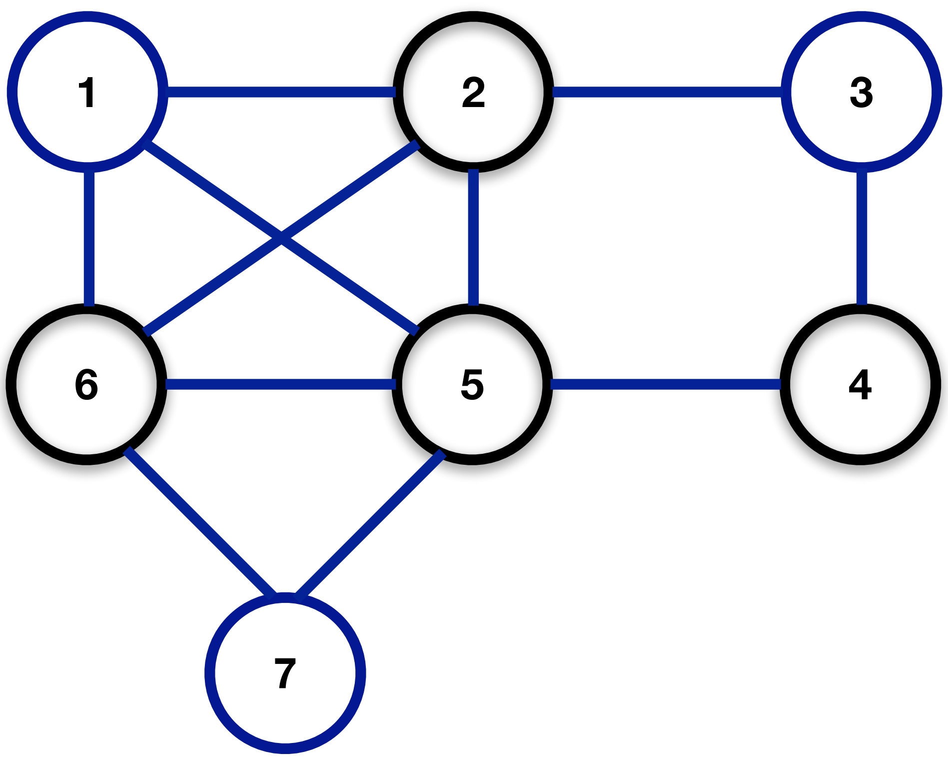
A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 **5** 6

Really? Let's test it.

Vertex 2 covers 4 edges.
Vertex 4 covers 2 more edges.
Vertex 5 covers 3 more edges.

Vertex Cover



A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

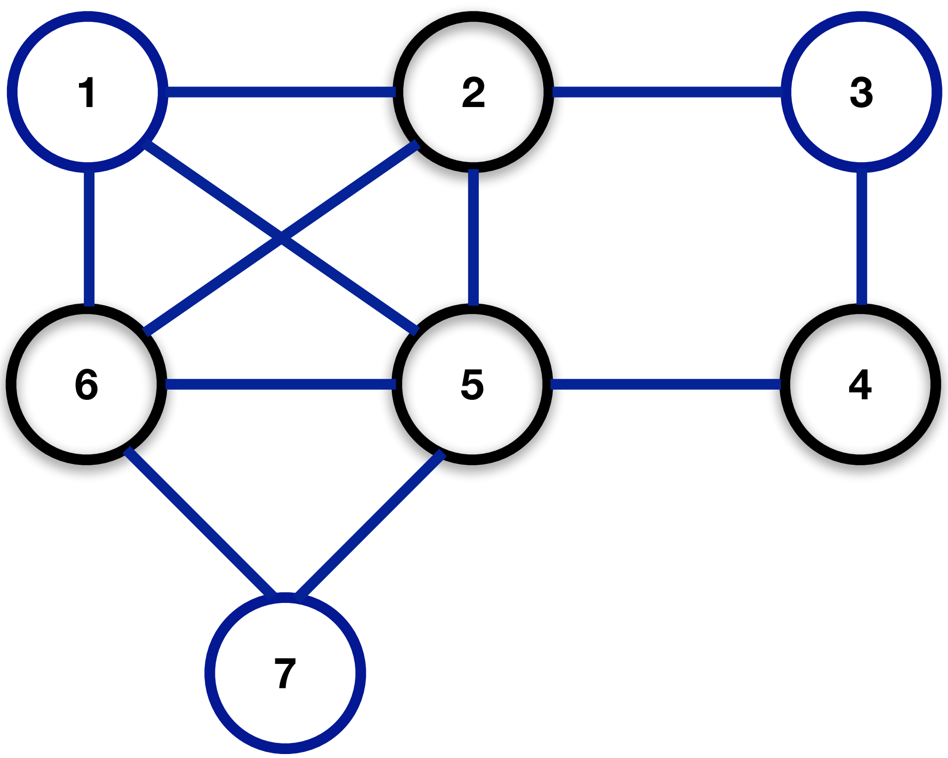
A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 **6**

Really? Let's test it.

Vertex 2 covers 4 edges.
Vertex 4 covers 2 more edges.
Vertex 5 covers 3 more edges.
Vertex 6 covers 2 more edges.

Vertex Cover



A **vertex cover** of a graph $G = (V, E)$ is a subset $V' \subseteq V$ such that if $\{u,v\}$ is an edge in G, then either u is in V' or v is in V' or both are. (I.e., it's a set of vertices V' such that each edge of graph G has at least one member of V' as an endpoint.)

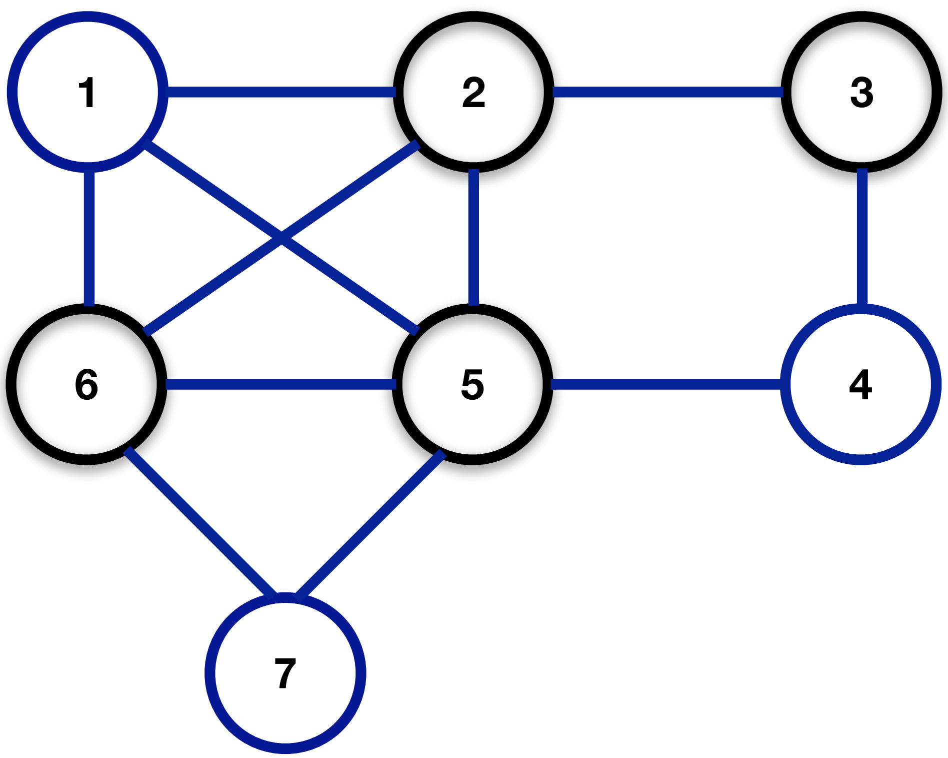
A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 6

Really? Let's test it.
Vertex 2 covers 4 edges.
Vertex 4 covers 2 more edges.
Vertex 5 covers 3 more edges.
Vertex 6 covers 2 more edges.
Yes, really.

Are there others?

Vertex Cover



A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

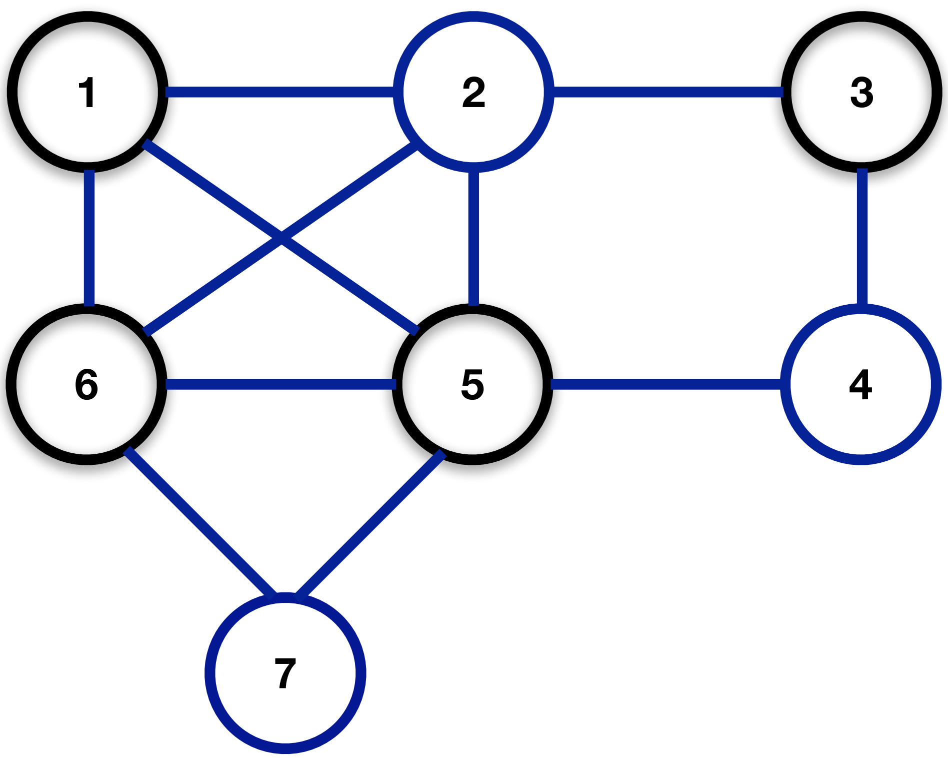
A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 6

Can you find two? 2 3 5 6

Can you find three?

Vertex Cover



A **vertex cover** of a graph *G* = (*V*, *E*) is a subset *V'* ⊆ *V* such that if {*u*,*v*} is an edge in *G*, then either *u* is in *V'* or *v* is in *V'* or both are. (I.e., it's a set of vertices *V'* such that each edge of graph *G* has at least one member of *V'* as an endpoint.)

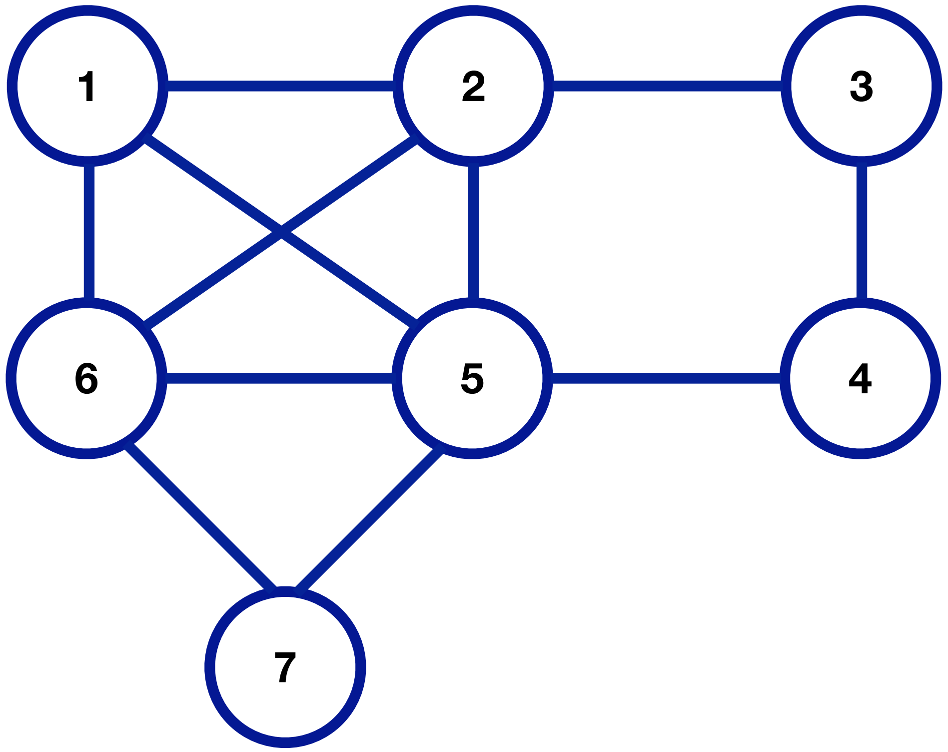
A **optimal vertex cover** is a vertex cover of minimum size for a given graph.

Can you find one? 2 4 5 6

Can you find two? 2 3 5 6

Can you find three? 1 3 5 6

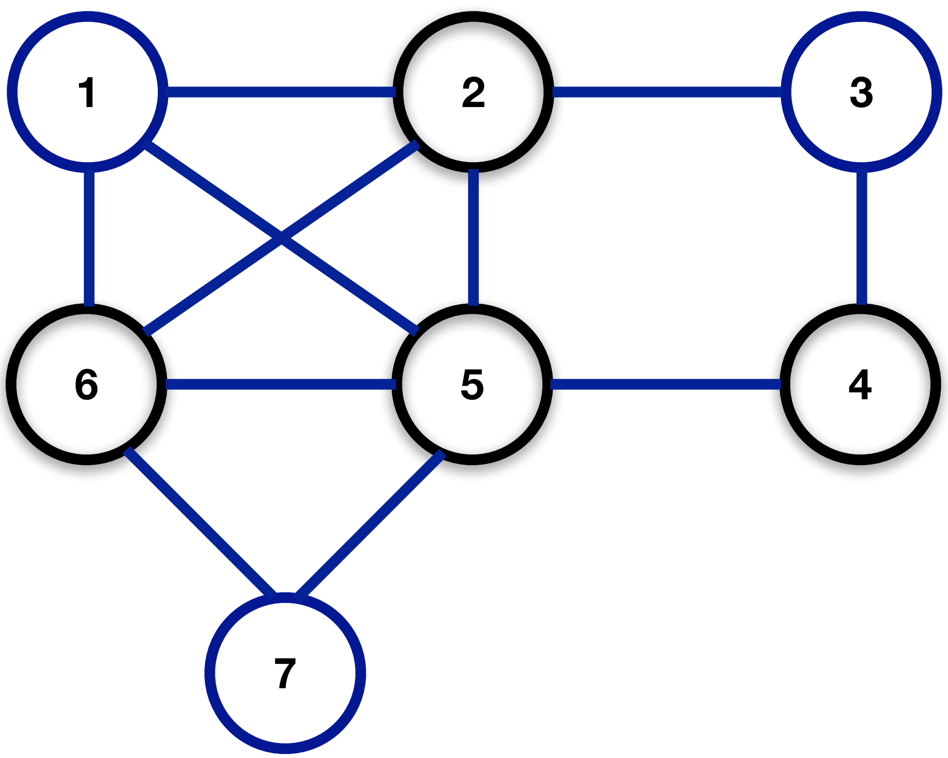
Vertex Cover and Independent Sets



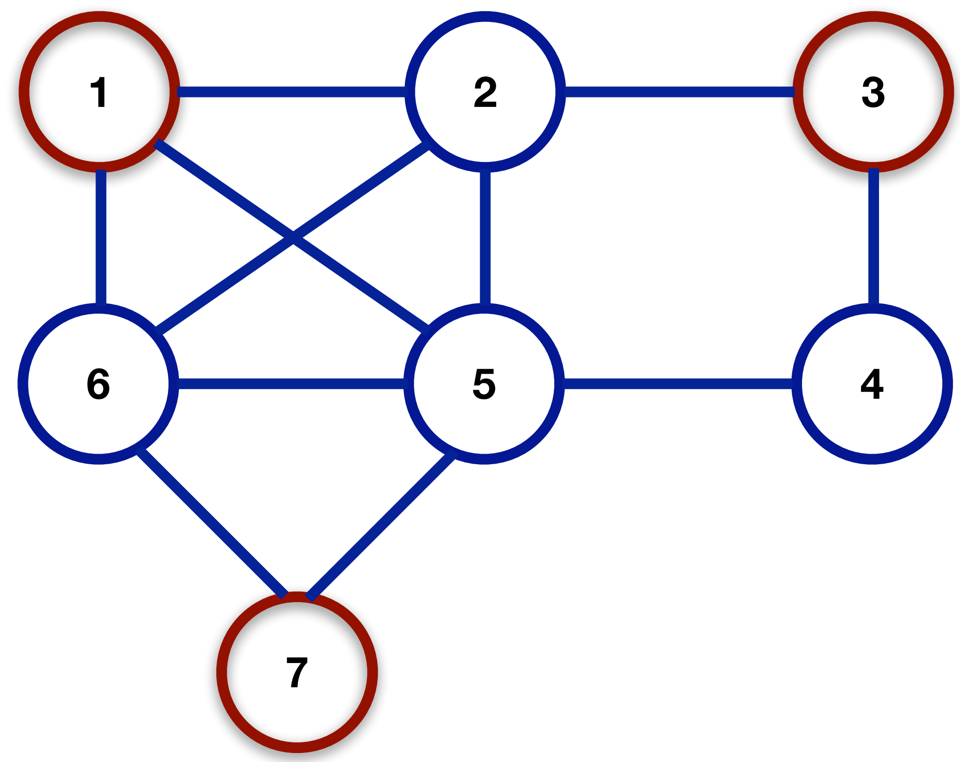
Did you notice a relationship between
vertex cover and **independent set**?

Graphs

Vertex Cover and Independent Sets

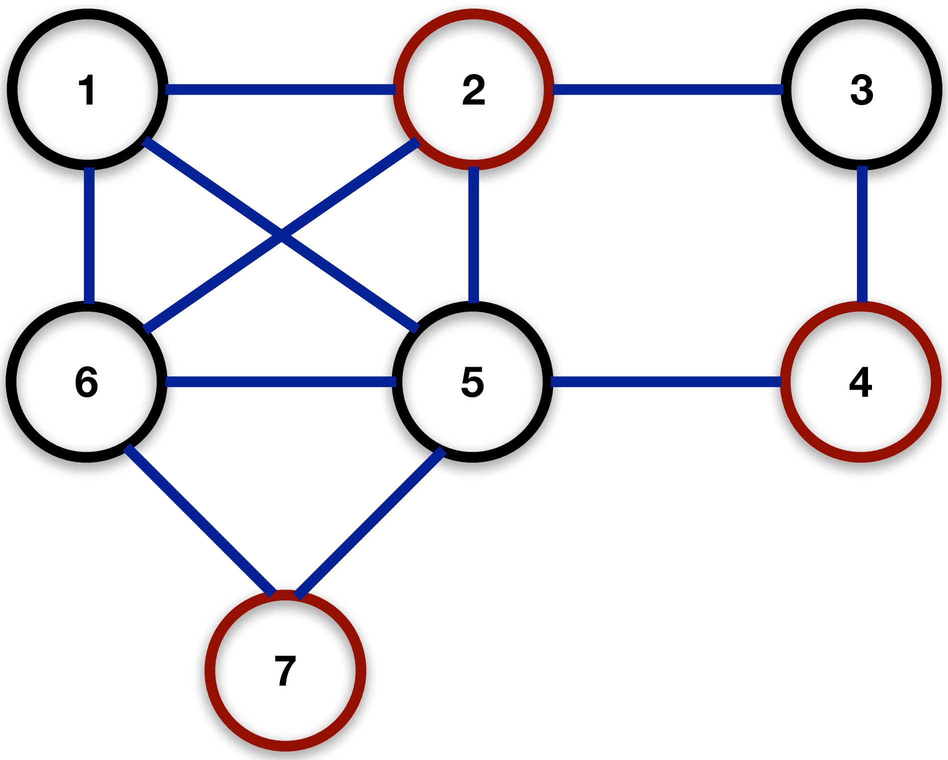


Vertex Cover



Independent Set

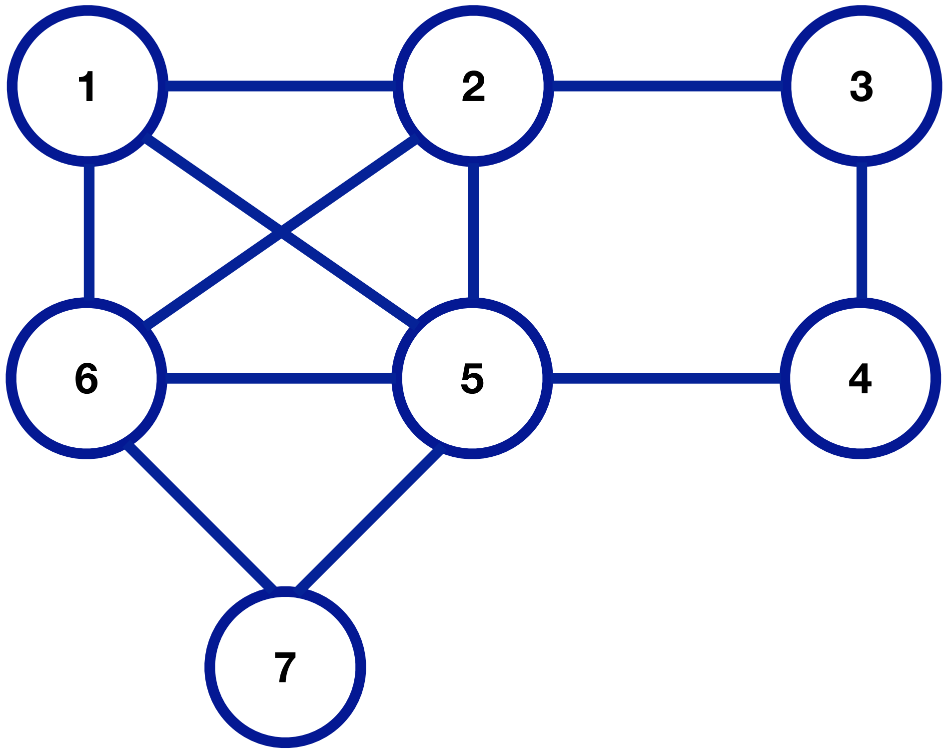
Vertex Cover and Independent Sets



Did you notice a relationship between **vertex cover** and **independent set**?

For any graph $G = (V, E)$, if $V' \subseteq V$ is a vertex cover for G then $V - V'$ is an independent set in G.

Cliques

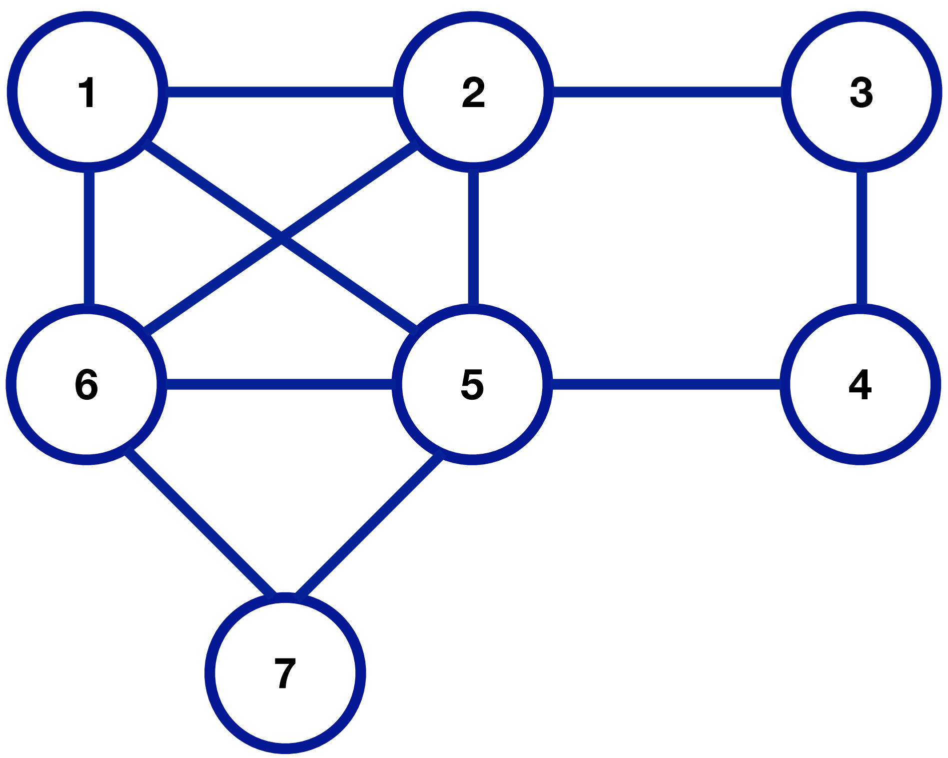


A **clique** in an undirected graph *G* = (*V*, *E*)

… is a subset of the vertex set *V’* ⊆ *V*, such that for every two vertices in *V’* there exists an edge connecting them.

… is the subgraph induced by *V’* as long as it is complete.

Cliques



A **clique** in an undirected graph *G* = (*V*, *E*)

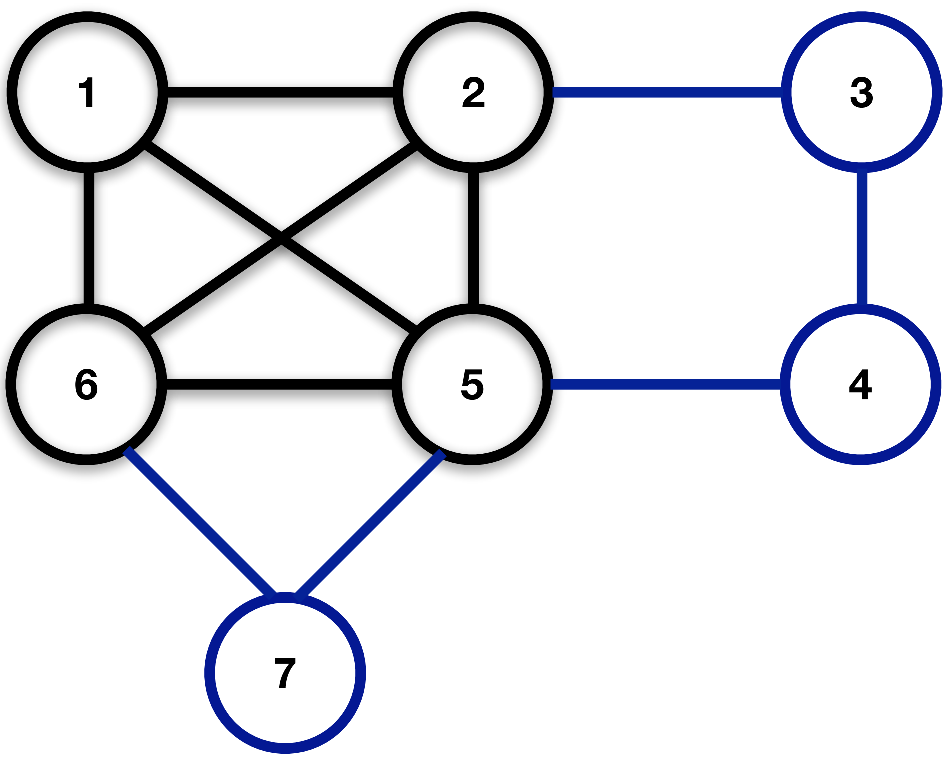
… is a subset of the vertex set *V'* ⊆ *V*, such that for every two vertices in *V'* there exists an edge connecting them.

… is the subgraph induced by *V'* as long as it is complete.

A **maximal clique** is a clique that cannot be extended by including even one more adjacent vertex.

Can you find a maximal clique?

Cliques



A **clique** in an undirected graph $G = (V, E)$

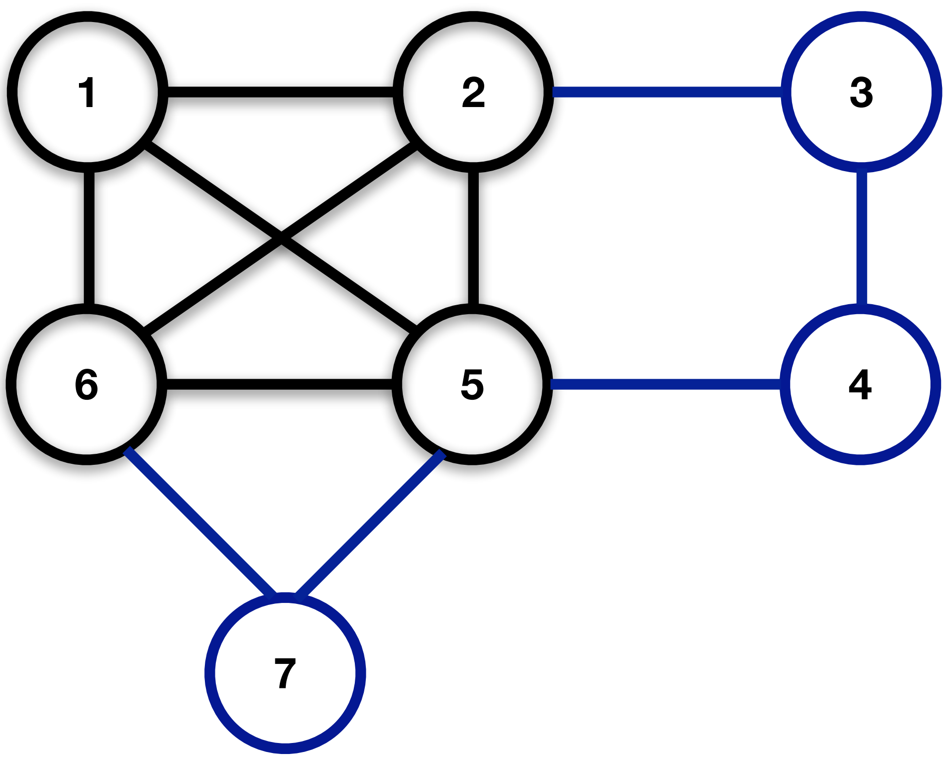
… is a subset of the vertex set $V' \subseteq V$, such that for every two vertices in V' there exists an edge connecting them.

… is the subgraph induced by V' as long as it is complete.

A **maximal clique** is a clique that cannot be extended by including even one more adjacent vertex.

Can you find a maximal clique? 1 2 5 6

Cliques



A **clique** in an undirected graph $G = (V, E)$

… is a subset of the vertex set $V' \subseteq V$, such that for every two vertices in V' there exists an edge connecting them.

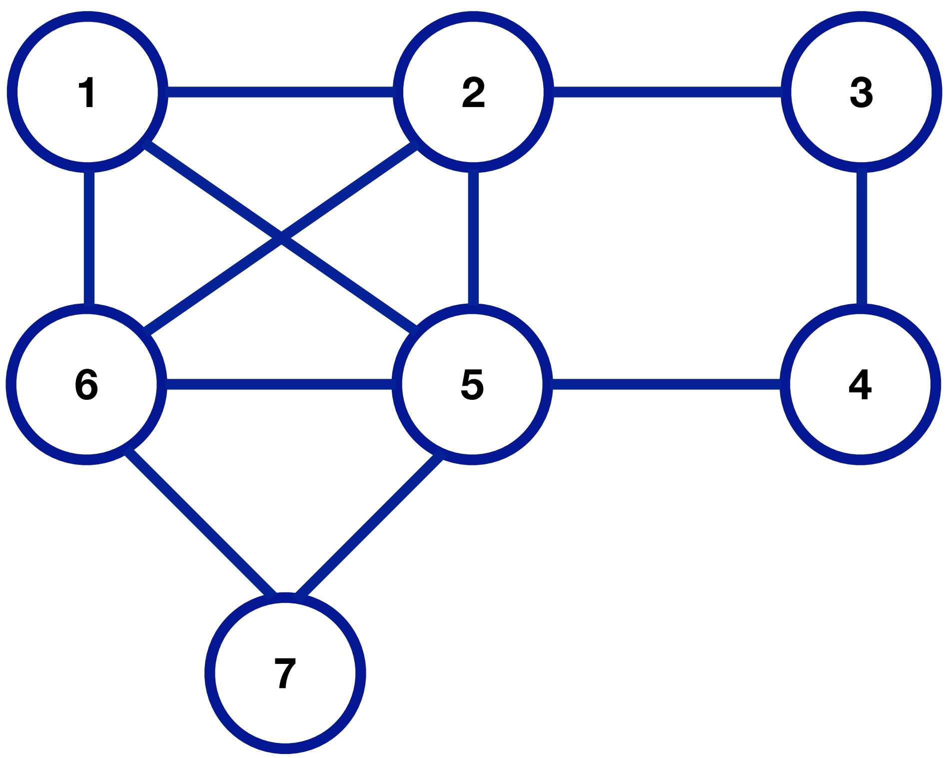
… is the subgraph induced by V' as long as it is complete.

A **maximal clique** is a clique that cannot be extended by including even one more adjacent vertex.

Maximal clique: 1 2 5 6

These are highly connected vertices.
How connected?

Clustering Coefficient

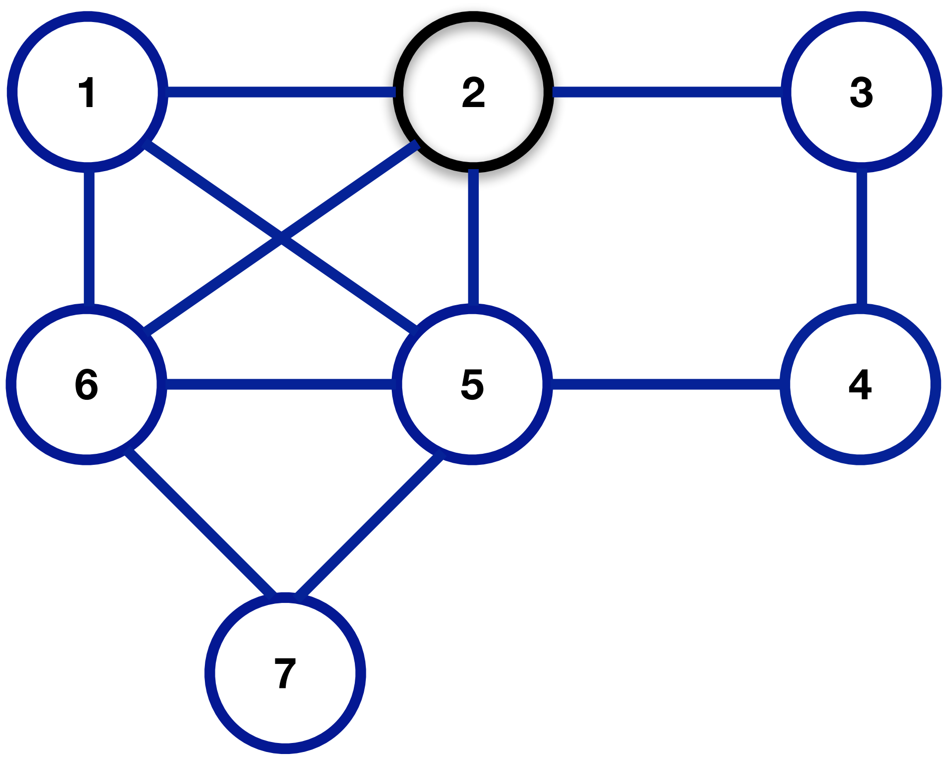


The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

Clustering Coefficient



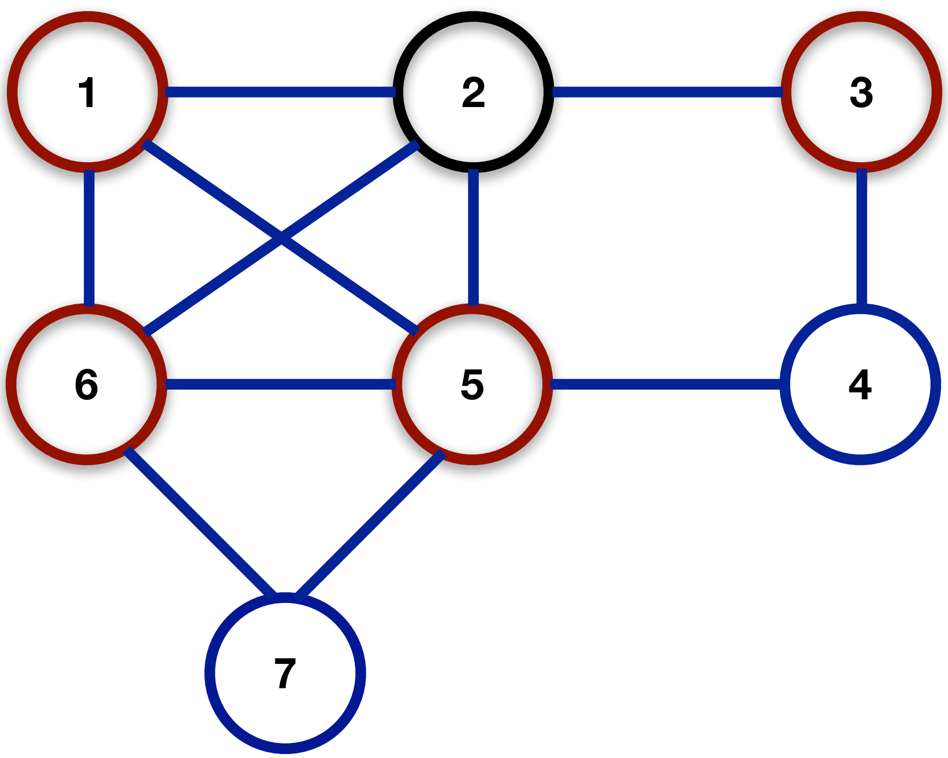
$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

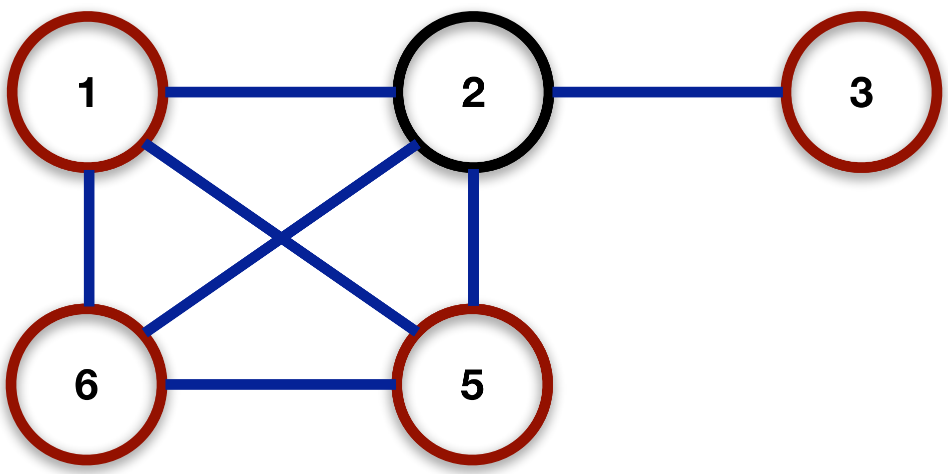
The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

Vertex 2 is adjacent to vertices 1, 3, 5, and 6.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

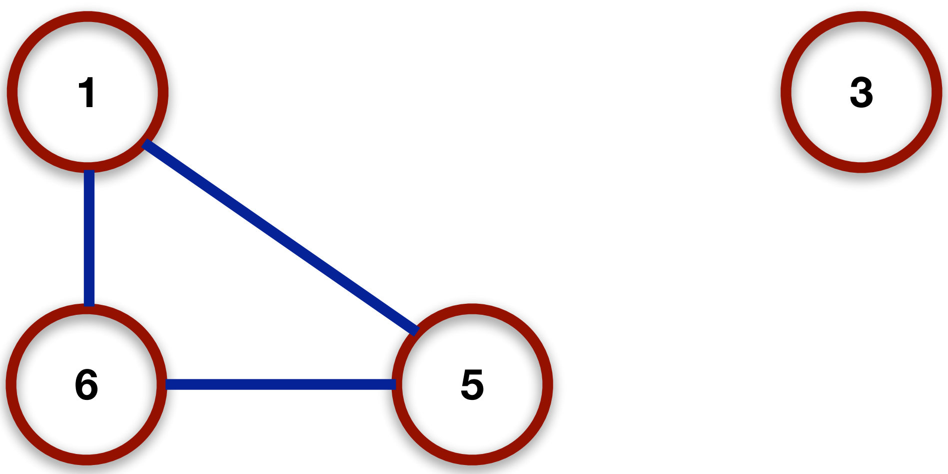
The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

Vertex 2 is adjacent to vertices 1, 3, 5, and 6. Ignore the rest.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

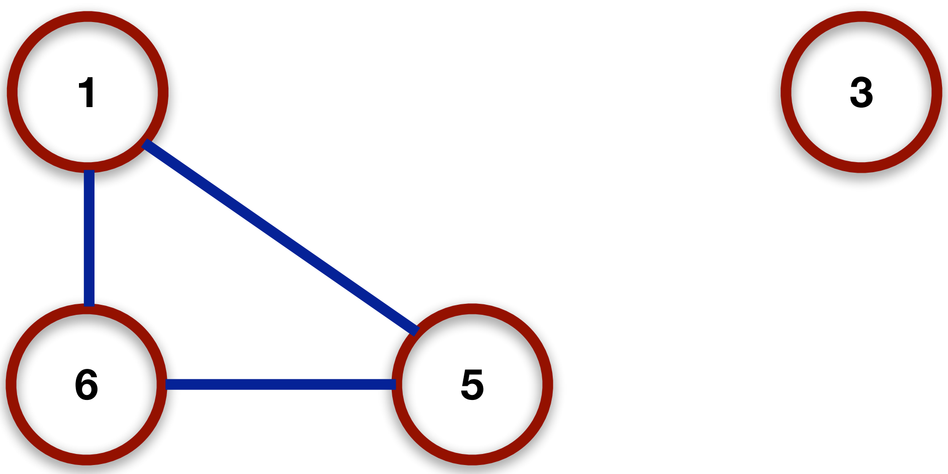
The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

Vertex 2 is adjacent to vertices 1, 3, 5, and 6. Ignore the rest. Ignore vertex 2 itself, since this is about the neighbors.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

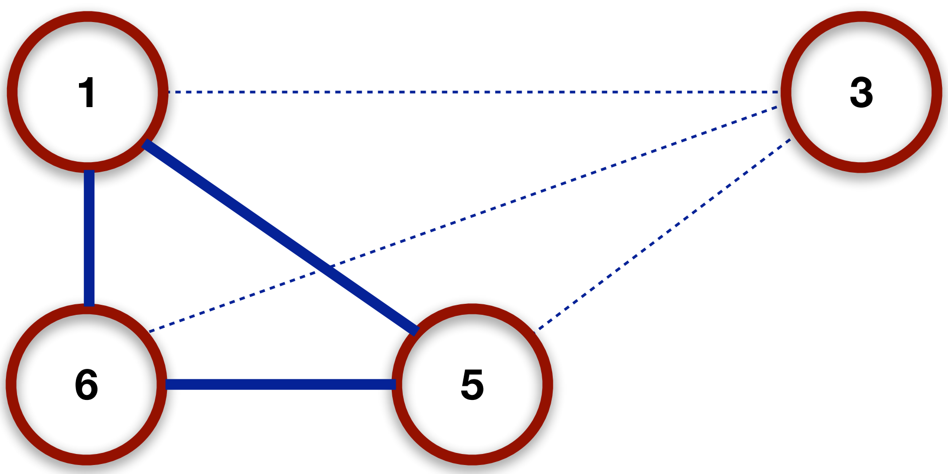
It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

Vertex 2 is adjacent to vertices 1, 3, 5, and 6. Ignore the rest. Ignore vertex 2 itself, since this is about the neighbors.

There are 3 actual edges among the remaining vertices 1, 3, 5, and 6.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

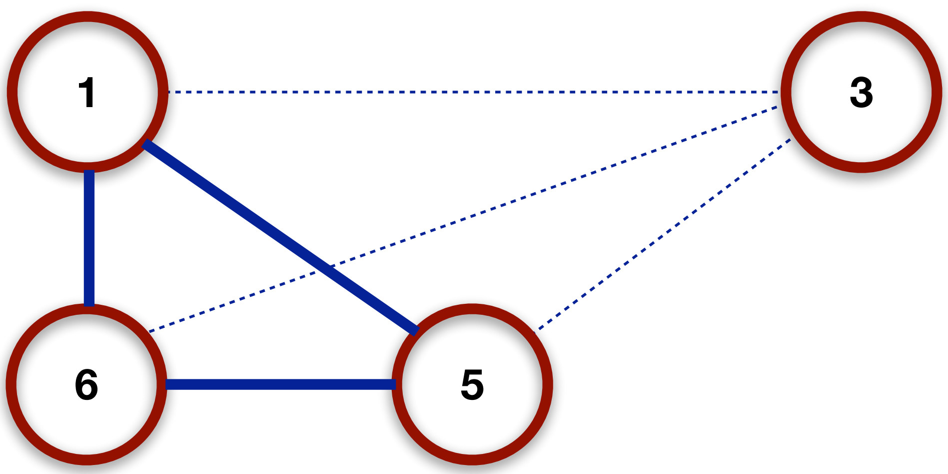
Let's calculate the clustering coefficient of vertex 2.

Vertex 2 is adjacent to vertices 1, 3, 5, and 6. Ignore the rest. Ignore vertex 2 itself, since this is about the neighbors.

There are 3 actual edges among the remaining vertices 1, 3, 5, and 6.

There are $n(n-1)/2$ possible edges among the remaining vertices 1, 3, 5, and 6 = 12/2 = 6.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

Let's calculate the clustering coefficient of vertex 2.

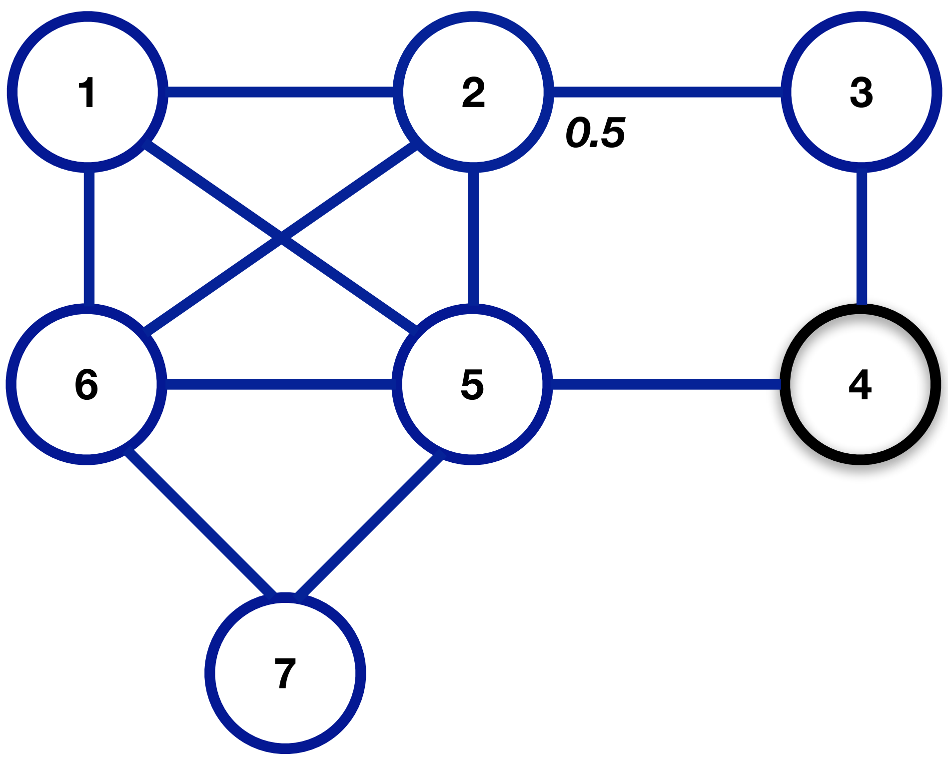
Vertex 2 is adjacent to vertices 1, 3, 5, and 6. Ignore the rest. Ignore vertex 2 itself, since this is about the neighbors.

There are 3 actual edges among the remaining vertices 1, 3, 5, and 6.

There are $n(n-1)/2$ possible edges among the remaining vertices 1, 3, 5, and 6 = 12/2 = 6.

ClusteringCoefficient(vertex 2) = 3/6 = 0.5

Clustering Coefficient



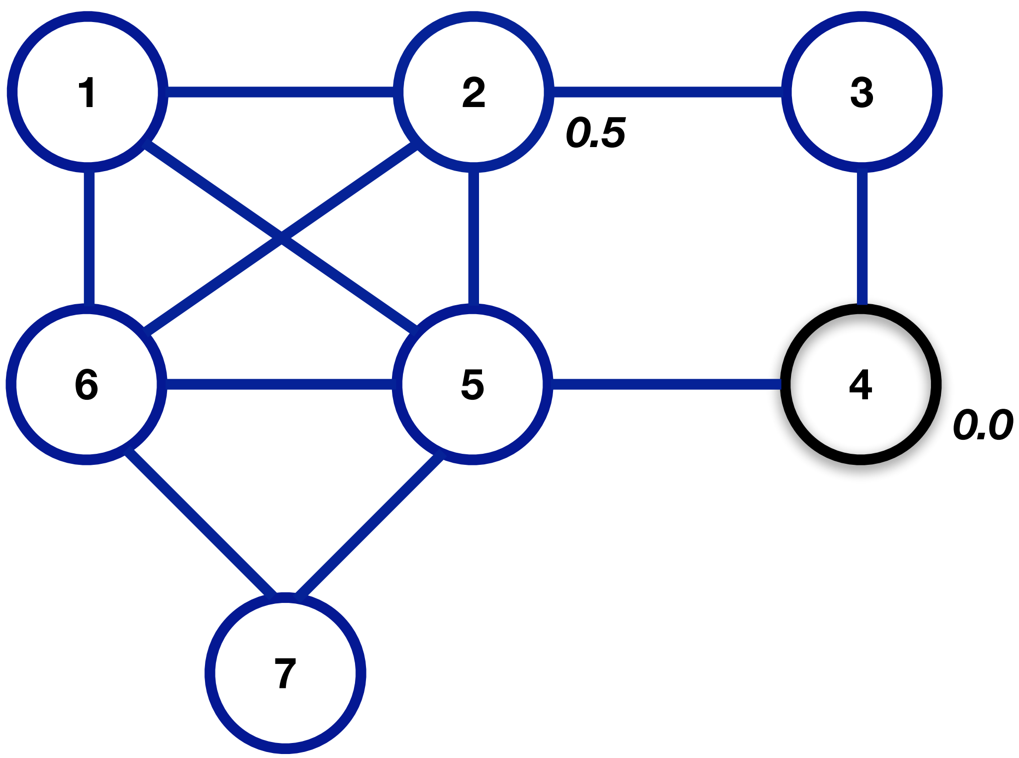
$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

What's the clustering coefficient of vertex 4?

Clustering Coefficient



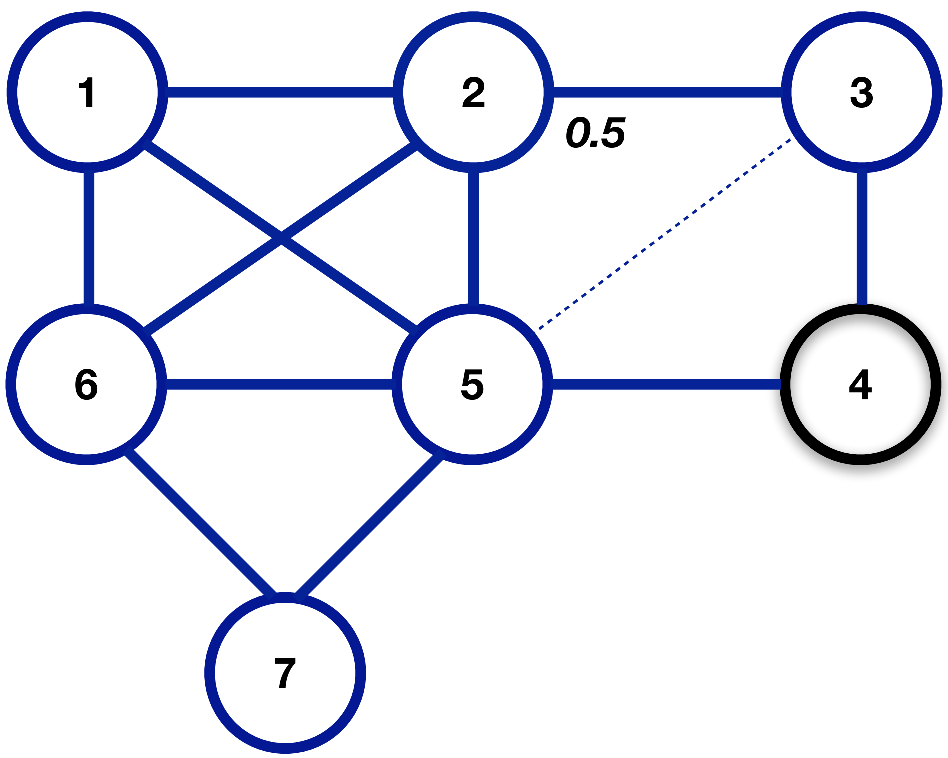
$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

What's the clustering coefficient of vertex 4? It's 0.

Clustering Coefficient



$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

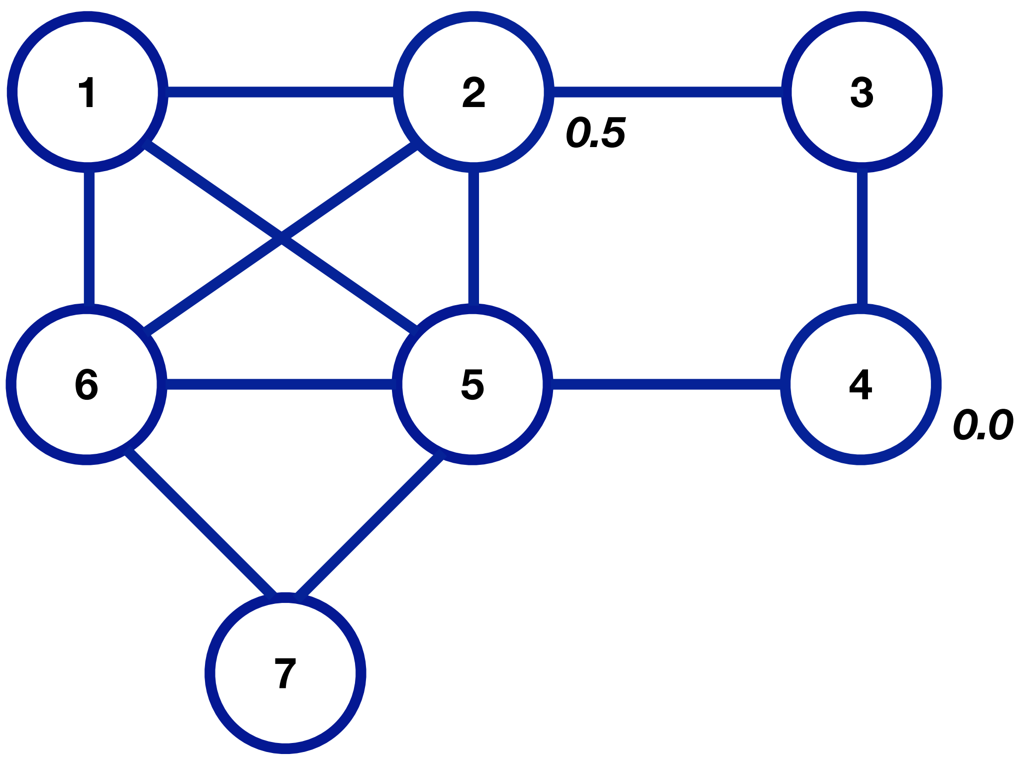
It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

What's the clustering coefficient of vertex 4? It's 0.

We could increase it — and this make the graph more "clique-ie" — by closing the 3-4-5 triangle (math pun intended).

This **Triadic Closure** is a common graph operation, and what social networks do to suggest people you may know.

Clustering Coefficient



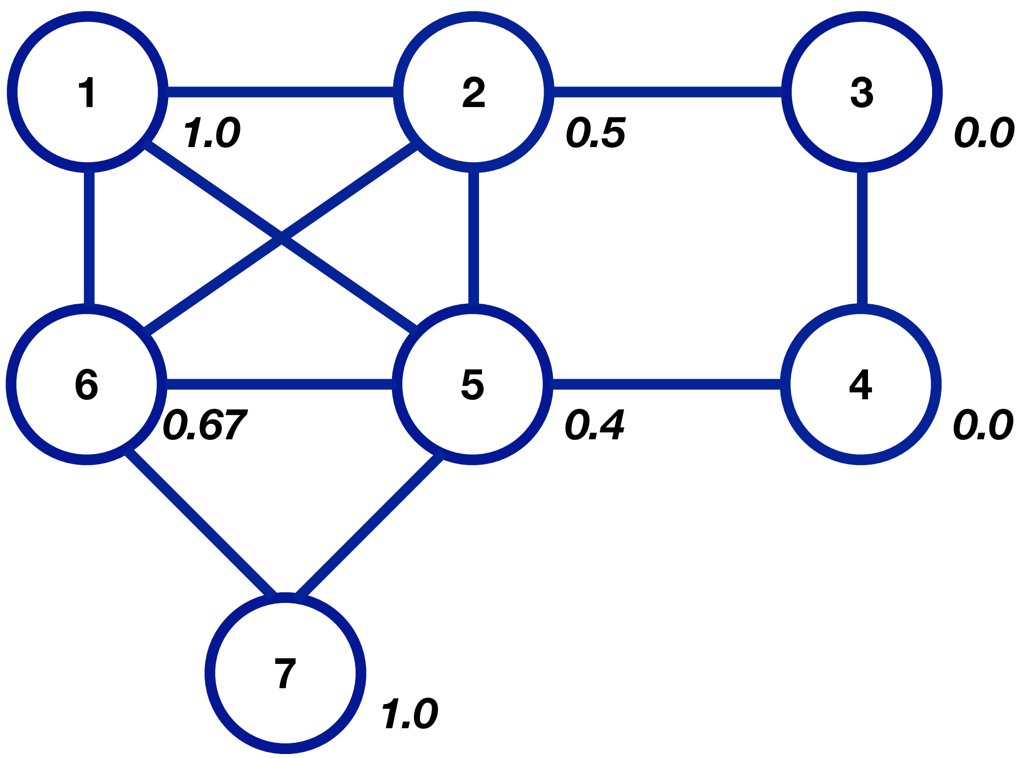
$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It's a measure of the local connectivity or "clique-iness" of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

How about the rest?

Clustering Coefficient



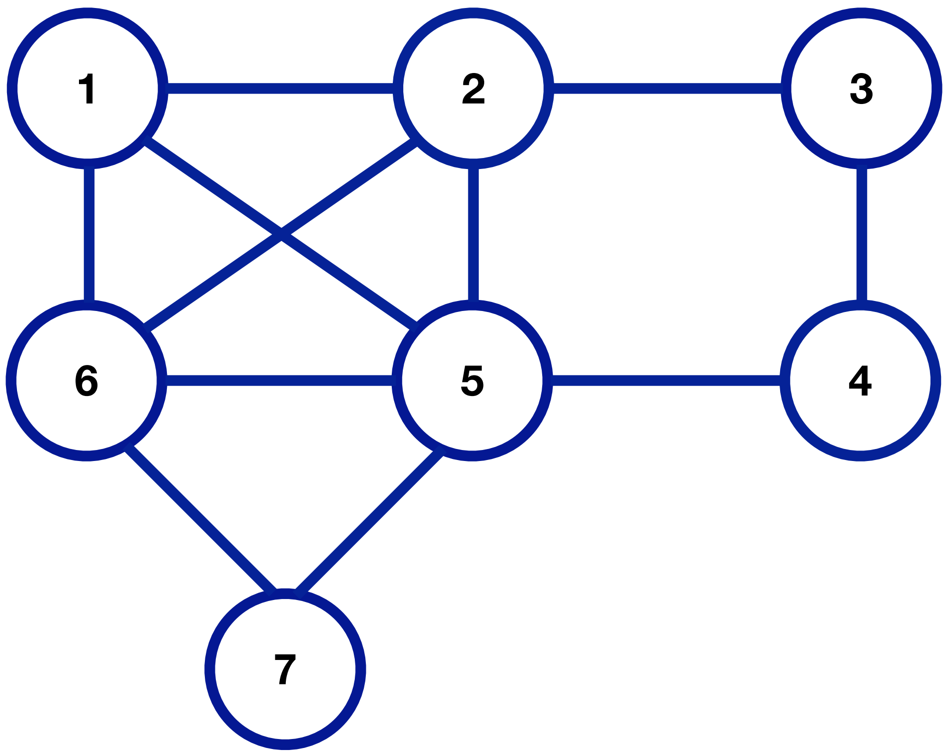
$$\frac{\text{number of actual edges among a vertex's neighbors}}{\text{number of possible edges among a vertex's neighbors}}$$

The **clustering coefficient** of a vertex in a graph is the probability that the neighbors of that vertex are also connected to each other.

It’s a measure of the local connectivity or “clique-iness” of a (all or part of a) graph. (I.e., the probability that the adjacent vertices of a given vertex are connected.)

How about the rest?

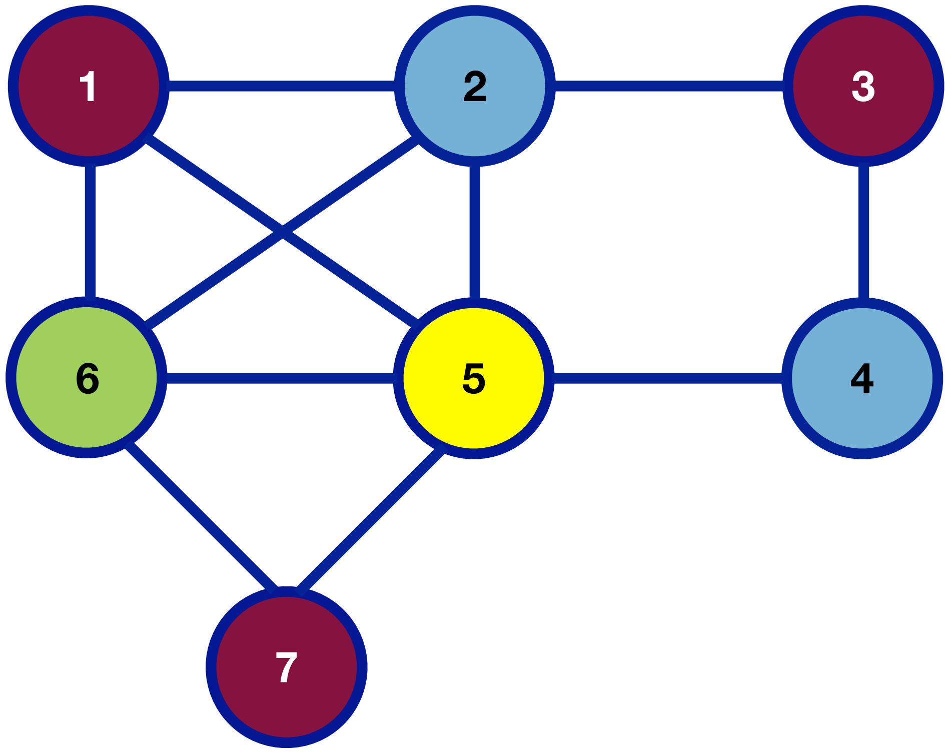
Graph Coloring



A **graph coloring** is a way of coloring the vertices of a graph such that no neighbors vertices share the same color.

The graph coloring problem is accomplishing this with a few colors as possible.

Graph Coloring



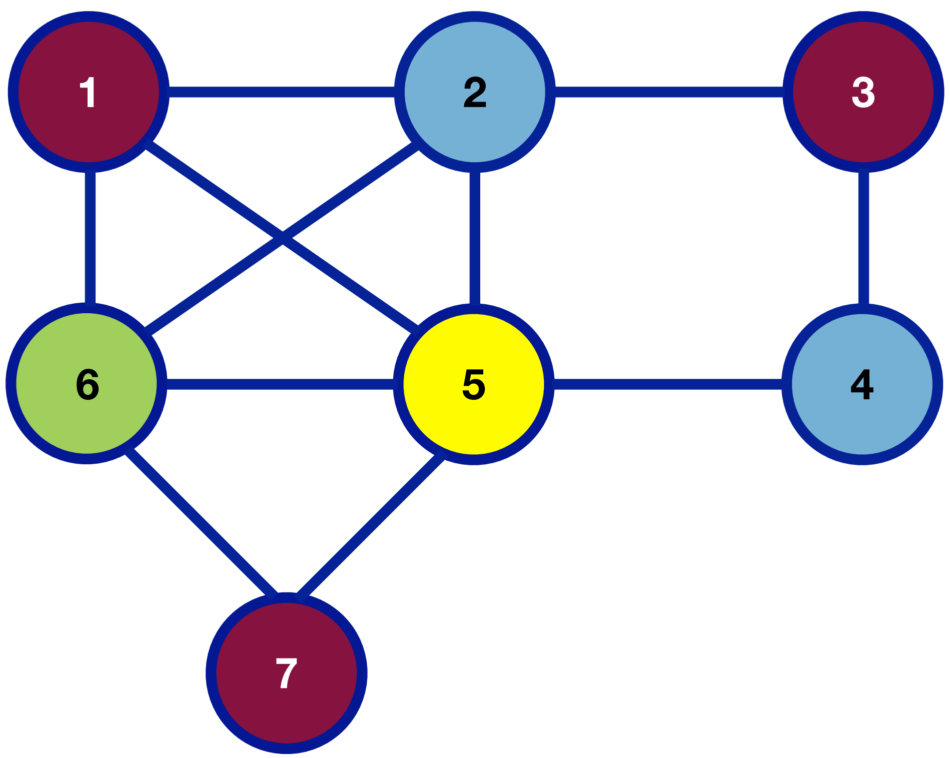
A **graph coloring** is a way of coloring the vertices of a graph such that no neighbors vertices share the same color.

The graph coloring problem is accomplishing this with a few colors as possible.

Here is a 4-color solution.

Is there a 3-color solution?

Graph Coloring



A **graph coloring** is a way of coloring the vertices of a graph such that no neighbors vertices share the same color.

The graph coloring problem is accomplishing this with a few colors as possible.

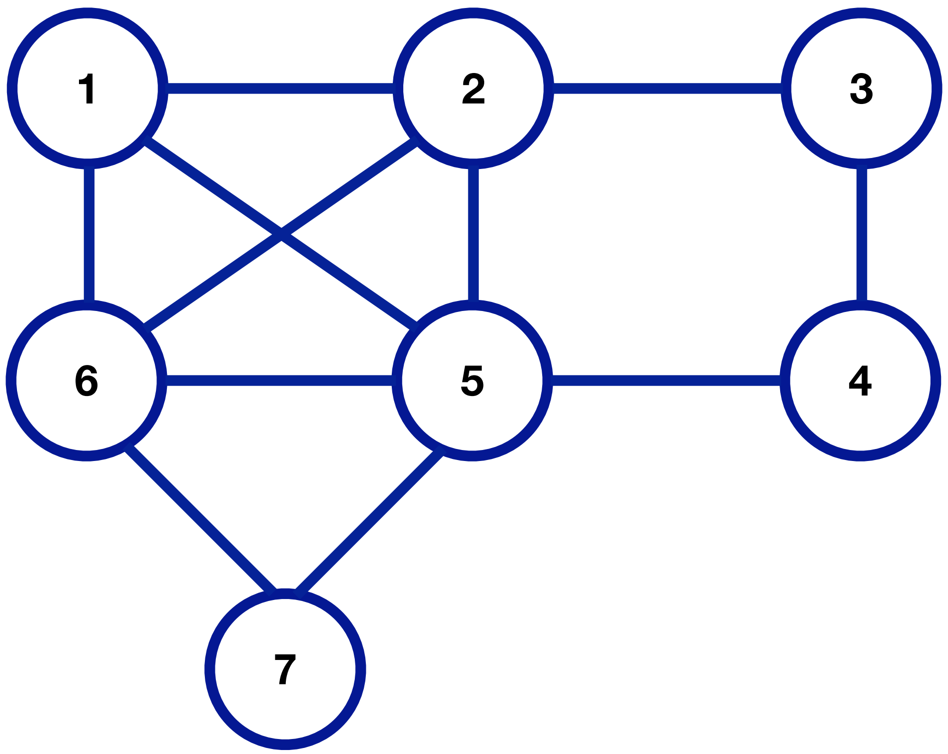
Here is a 4-color solution.

Is there a 3-color solution?

We can use graph coloring to model Sudoku.

Graphs

There's more . . .



- Distribution of vertex degrees
- Distribution of clustering coefficients and triadic closure
- Network density
- Size of connected components
- Shortest distance between pairs of vertices
- The centrality or eccentricity of vertices by various measures (PageRank and closeness centrality are of particular interest.)

… but it will have to wait.