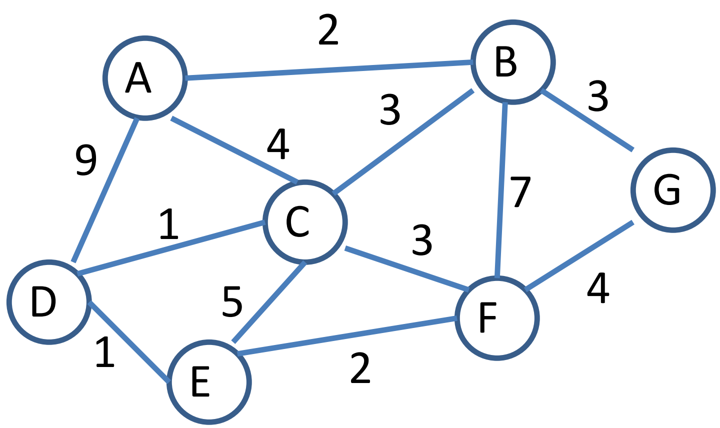
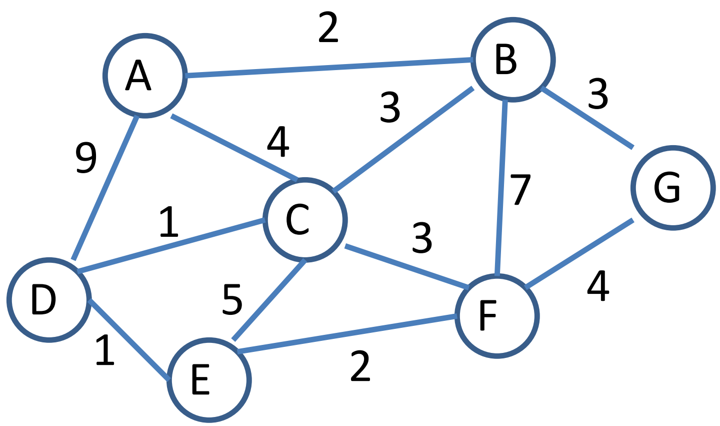
Minimum Spanning Trees

Question: You have 7 data points. Shown below are how "far apart" some points are from others. You need to divide the points into two clusters so that nearby points are in the same cluster. How would you divide these points?



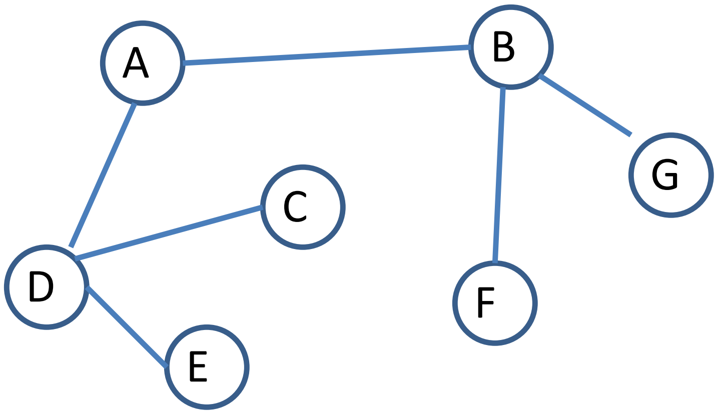
A. {A,B,C} {D,E,F,G}
B. {A,C,D,E} {B,F,G}
C. {C,D,E,F} {A,B,G}
D. {C,E,F} {A,B,D,G}

Suppose you are laying out electrical lines for a new housing subdivision with 7 houses. Naturally, every house needs electricity. Here are costs for connecting a line pairs of houses. What is on connection you are sure to use?



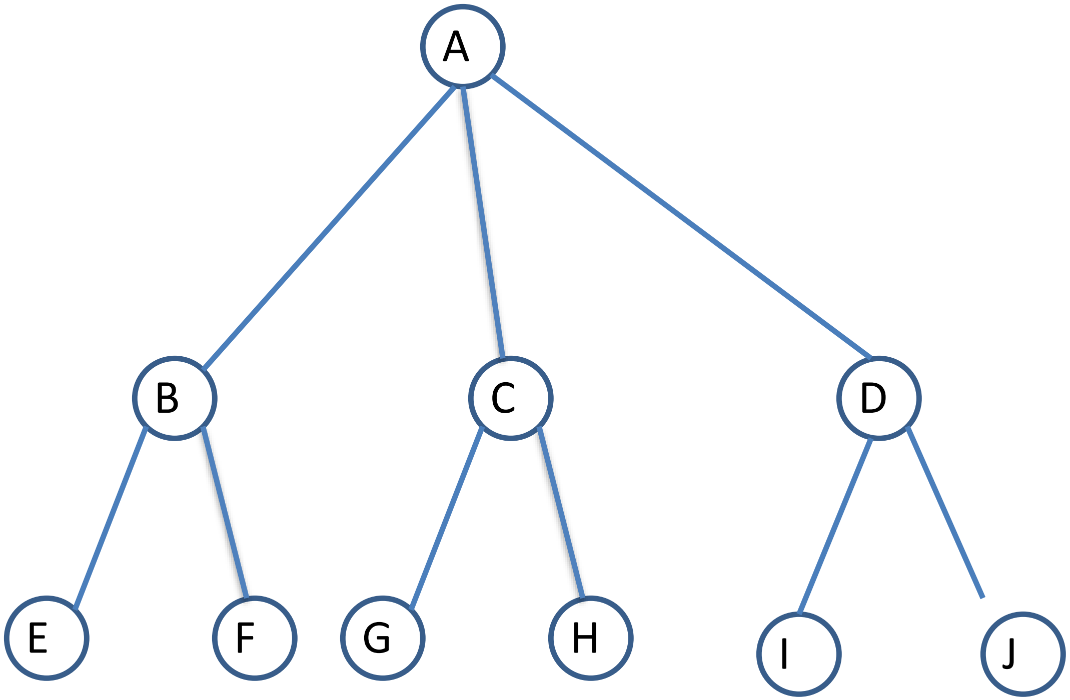
A. D to C
B. D to A
C. C to B
D. C to E

We have mostly focused on directed graphs. We call an undirected graph a *tree* if it is connected (you can go from any node to any other node) and has no cycles. For example:



Note that if we picked up this graph by any of its nodes and let gravity pull the others down, this would look like a directed tree.

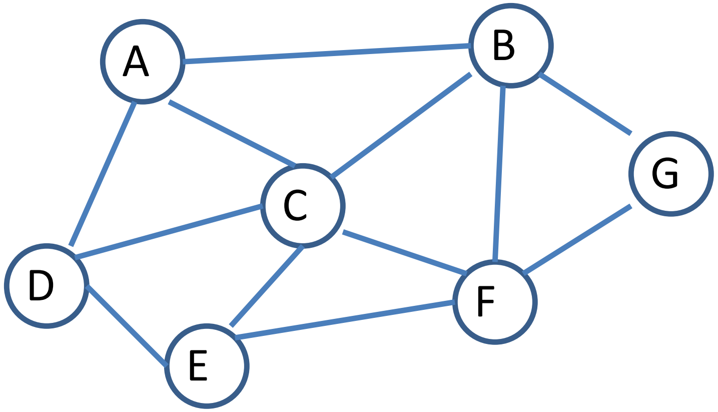
Note that a tree with n nodes has exactly n-1 edges, since each node except for the root is the destination of exactly one edge:



This has 10 nodes and 9 edges.

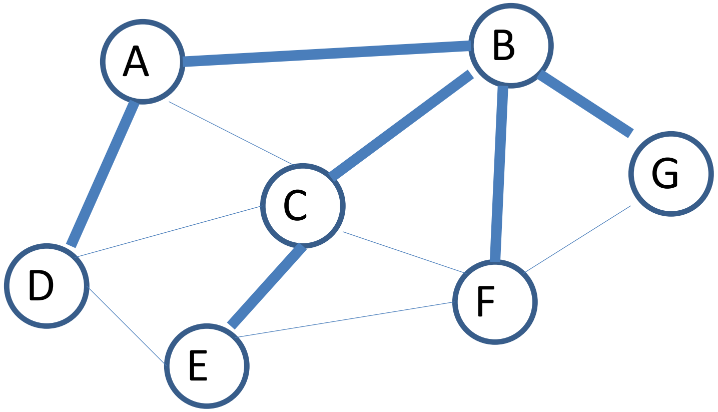
Similarly, any connected graph with n nodes and n-1 edges must be a tree, for if it contained a cycle there wouldn't be enough edges to connect all of the vertices.

Now consider a general undirected graph, such as



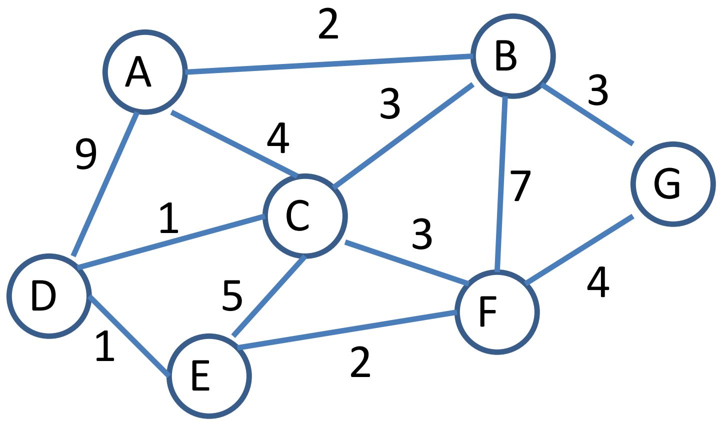
A *spanning* tree for such a graph is a graph with the same nodes and a subsect of the edges that is connected and acyclic. In other words, it is a subset of the edges that forms a tree and reaches all of the nodes of the graph.

Here the dark edges form a spanning tree



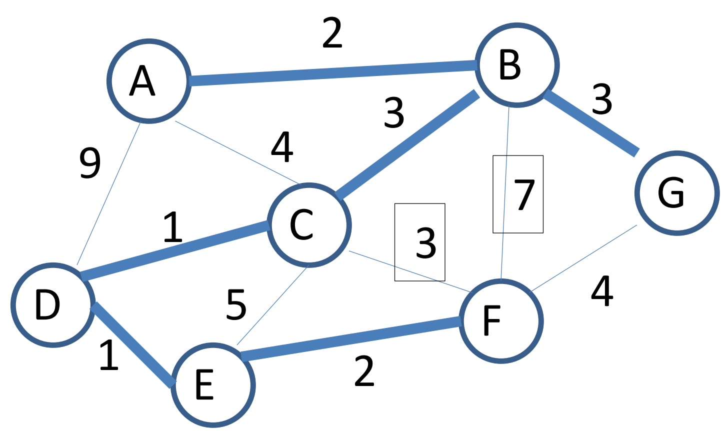
Spanning trees are useful in many situations. For example, if the graph represents possible connections between nodes on a network, a spanning tree represents the smallest set of connections needed to link all of the nodes.

Now add weights to the edges, such as



A *minimum spanning tree* is the spanning tree whose edge weights sum to the smallest amount. Minimum spanning trees have many applications, from network design to clustering algorithms, maze generation, optimal path selection and so forth.

Here is a minimum spanning tree for this graph:



We would like an algorithm to find the minimum spanning tree for a general weighted undirected graph.