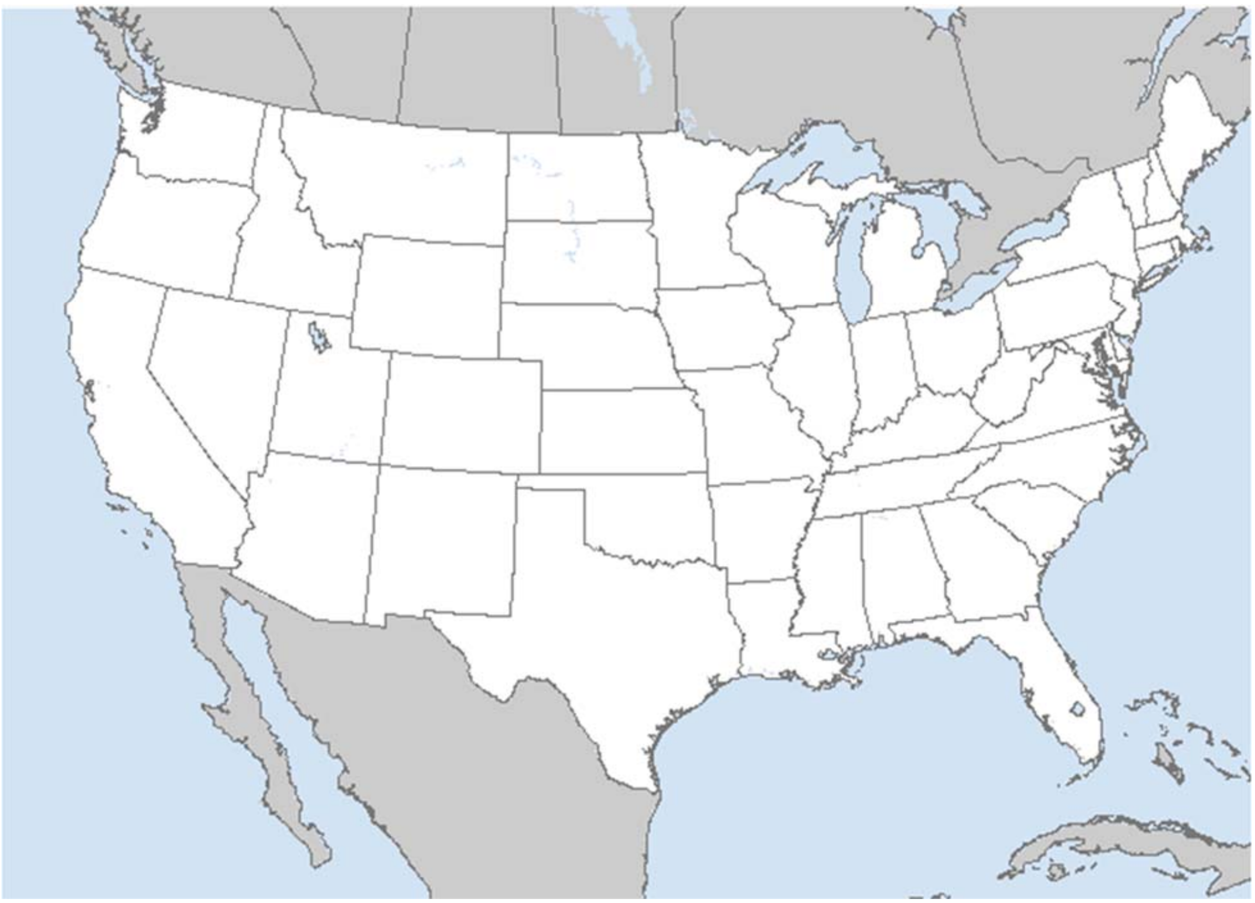
Discrete Math
Ogrodnik
Graph Theory Intro

Amazon, as you know, is one of the world's largest retailers. In 2017, the company made over 5 billion worldwide Prime shipments alone. An organized and efficient distribution network is needed to manage such a large quantity of shipments. While Amazon does not publicly release exact numbers, one logistics consulting company's research reports that Amazon has 122 fulfillment centers in the United States.

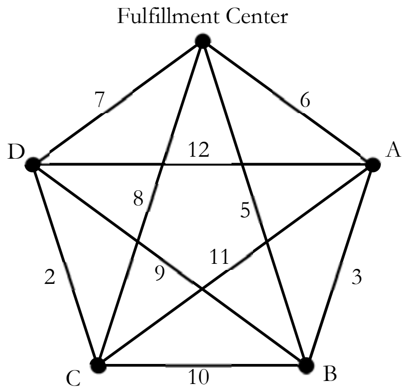
Imagine you work for Amazon and are in charge of planning the locations of these fulfillment centers. Where would you place these 122 centers? What factors would you need to consider.

Work in your group to determine a location for every one of the 122 Amazon fulfillment centers. On the map below you must display the location of each center. You must also include an explanation for how you determined these locations. Identify any resources you used in making your decisions as well as the reasoning behind your decisions.



Discrete Math
Ogrodnik
Graph Theory Intro

1. Suppose you are work for Amazon and are in charge of managing the delivery schedules. There are four packages that will be delivered by the same driver. The **graph** below represents the relative distance (but not position) between any two delivery locations. Delivery locations are represented by the vertices. Find the absolute shortest route this driver can take in delivering these four packages. The driver must leave from and return to the fulfillment center. Identify the route and its total distance.



$FC\overset{6}{-}A\overset{3}{-}B\overset{10}{—}C\overset{2}{-}D\overset{7}{-}FC=28$

$FC\overset{6}{-}A\overset{3}{-}B\overset{9}{-}D\overset{2}{-}C\overset{8}{-}FC=28$

$FC\overset{6}{-}A\overset{11}{—}C\overset{10}{—}B\overset{9}{-}D\overset{7}{-}FC=43$

$FC\overset{6}{-}A\overset{11}{—}C\overset{2}{-}D\overset{9}{-}B\overset{5}{-}FC=33$

$FC\overset{6}{-}A\overset{12}{—}D\overset{9}{-}B\overset{10}{—}C\overset{8}{-}FC=45$

$FC\overset{6}{-}A\overset{12}{—}D\overset{2}{-}C\overset{10}{—}B\overset{5}{-}FC=35$

$FC\overset{5}{-}B\overset{3}{-}A\overset{11}{—}C\overset{2}{-}D\overset{7}{-}FC=28$

$FC\overset{5}{-}B\overset{3}{-}A\overset{12}{—}D\overset{2}{-}C\overset{8}{-}FC=30$

$FC\overset{5}{-}B\overset{10}{—}C\overset{11}{—}A\overset{12}{—}D\overset{7}{-}FC=45$

$FC\overset{5}{-}B\overset{10}{—}C\overset{2}{-}D\overset{12}{—}A\overset{6}{-}FC=35$

$FC\overset{5}{-}B\overset{9}{-}D\overset{12}{—}A\overset{11}{—}C\overset{8}{-}FC=45$

$FC\overset{5}{-}B\overset{9}{-}D\overset{2}{-}C\overset{11}{—}A\overset{6}{-}FC=33$

Calculate the length of every possible cycle, then pick the shortest. Half of the work is shown above, though some cycles are repeats of others (in reverse).

2. Imagine a variation of the above problem with 10 deliveries that must be made. How many total **cycles** are possible that allow the driver to start and end at the fulfillment center and make every delivery?

Assuming each cycle has the same distance regardless of the direction it is travelled, then:

$\frac{10!}{2}=1{,}814{,}400$ different cycles

Discrete Math
Ogrodnik
Graph Theory Intro

There are different methods for solving a traveling salesperson problem, but only one that always produces an optimal solution – the **brute force method**. With this method, you simply calculate the length of every possible cycle. In these calculations the optimal solution will be evident.

1. Consider a traveling salesperson problem with 8 locations that must be visited. How many cycles would need to be examined to use the brute force method?

$$\frac{8!}{2} = 20{,}160$$

2. According to a Wired article, the average UPS driver makes 120 deliveries per day. How many cycles would need to be examined to use the brute force method to find the optimal route for these deliveries?

$$\frac{120!}{2} = 3.34 \times 10^{198}$$

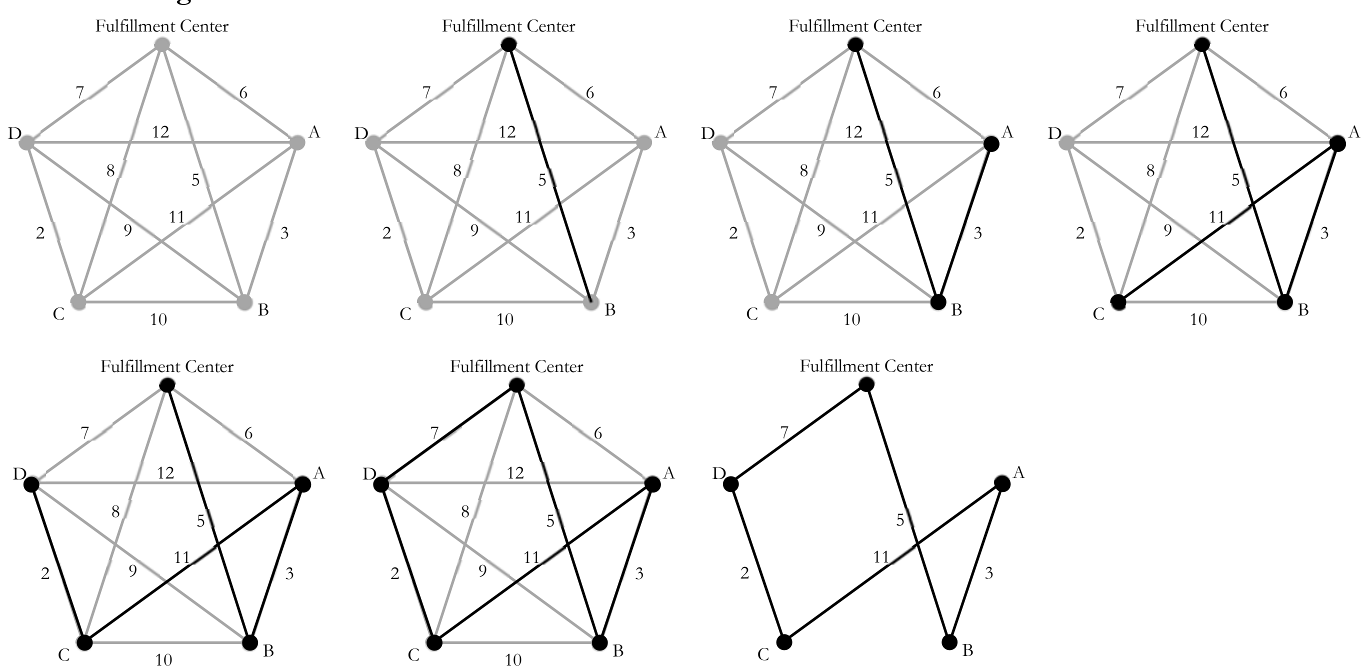
3. Suppose a computer can analyze a single cycle from #2 in one nanosecond. How long would it take for a calculator to analyze every cycle from the problem in #2?

$$\frac{3.34 \times 10^{198}}{1{,}000{,}000{,}000} = 3.34 \times 10^{189} \text{ seconds} = 1.06 \times 10^{182} \text{ years}$$

Because of the severe inefficiency of the brute force solution we need methods that have a chance of producing the optimal solution but require far less work.

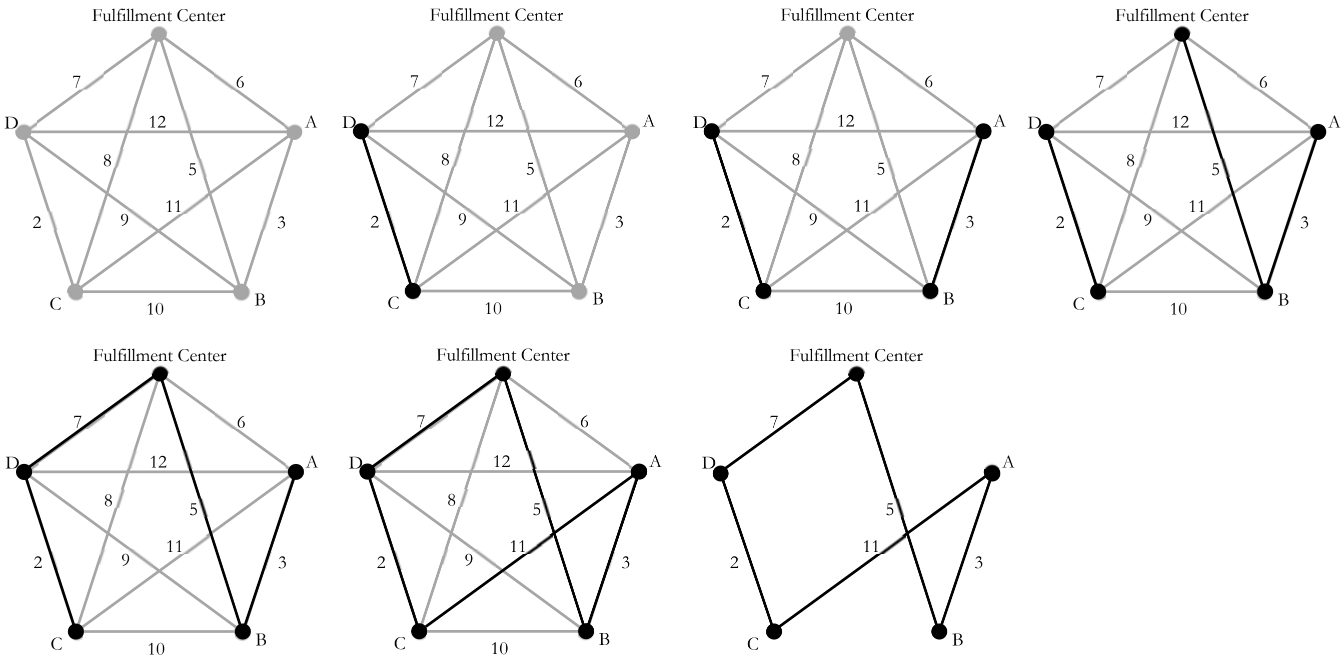
The diagrams below illustrate two **suboptimal solutions** to our initial problem from day 1.

Nearest Neighbor



Discrete Math
Ogrodnik
Graph Theory Intro

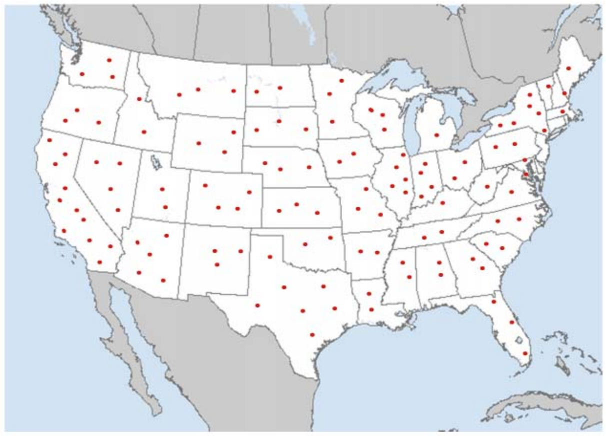
Sorted Edges

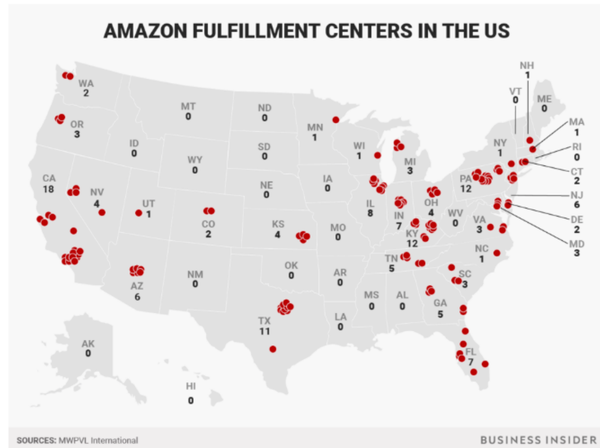


In this case, both the nearest neighbor method and the sorted edges method produce the optimal solution – a cycle with a length of 28. Neither method will always produce the optimal solution, but each method is likely to produce a near optimal solution and do so requiring very little work.

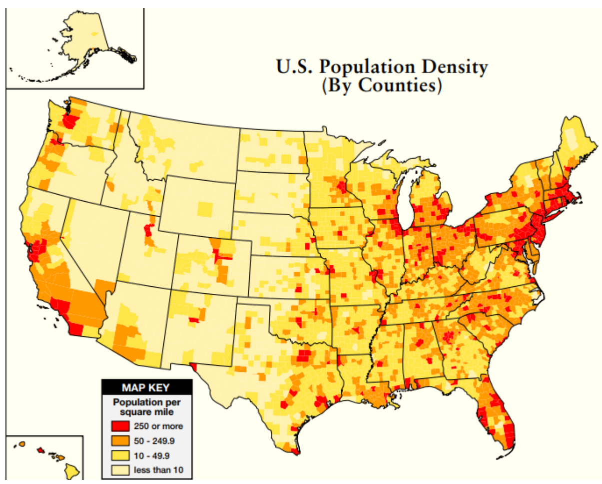
Discrete Math
Ogrodnik
Graph Theory Intro

Recall our first day of this unit, where you attempted to place the 122 Amazon fulfillment centers across the country. The map on the left is a student solution that is representative of the work many of you submitted. The map on the right is the actual placement of these fulfillment centers.





This is a map of the population density of the country, by county.



1. You should notice a closer correlation between population density and number of Amazon fulfillment centers. There are some locations, however, that are dense but not near a fulfillment center, as well as other locations that are not dense but are home to a fulfillment center. What other factors might Amazon be considering when placing these centers?

   Cost of building the center in certain locations
   Proximity to other facilities (warehouses, producers, etc.)
   Taxes

2. Finally, read the article at the following link: https://www.wired.com/2013/06/ups-astronomical-math/

Discrete Math
Ogrodnik
Graph Theory Intro

1. A city is planning a new subway system. The system must connect the seven locations listed below. This system does not need to form a cycle; the main goal is to minimize the cost of the project while still forming a connected system. Assume the cost per mile is the same between any two locations.

   The distances between any two locations are shown below, in miles.

	A	B	C	D	E	F	G
A		2.3	3.7	1.2	4.8	0.9	1.5
B	2.3		3.6	2.4	1.7	2.8	3.9
C	3.7	3.6		4.0	2.5	3.0	4.1
D	1.2	2.4	4.0		2.2	3.8	0.5
E	4.8	1.7	2.5	2.2		2.6	1.3
F	0.9	2.8	3.0	3.8	2.6		4.4
G	1.5	3.9	4.1	0.5	1.3	4.4	

   Find the optimal solution to this problem. Show the connections that will be made, and the total distance of this solution.

   - You are not necessarily forming a Hamiltonian path. You are definitely not forming a Hamiltonian cycle.
   - If you create a graph of your solution, it should be a **connected graph**.
   - The total "weight" of the graph is what you must minimize.
   - To minimize the weight, you shouldn't have any redundant edges – there should be only one path connecting any two vertices.

Discrete Math
Ogrodnik
Graph Theory Intro

Your solution to #1 should form a **tree** – a connected graph in which every edge is a bridge. A tree that includes every vertex of a given graph is called a **spanning tree**. When the edges have weights (such as distance, cost, time, etc.), the tree with the lowest overall weight is called the **minimum cost spanning tree**.

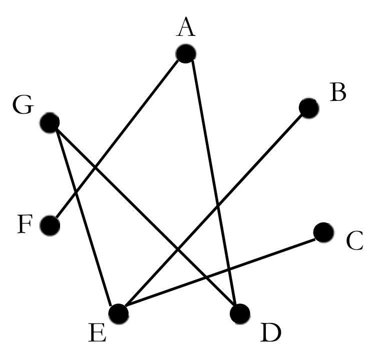
Prim's algorithm can be used to find the minimum cost spanning tree for any graph.

Prim's Algorithm
Define the sets Open and Connected, such that Open includes every vertex in your graph and Connected is an empty set.

1. Choose any vertex in Open and add it to Connected. Remove the vertex from Open.
2. Find the edge with the lowest weight that connects a vertex in Open to a vertex in Connected.
3. Add the vertex from Open used in #2 to Connected, and remove the vertex from Open.
4. Add the edge used in #2 to the tree.
5. If Open is now empty, terminate the algorithm. Else, return to step 2.

2. Apply Prim's algorithm to the problem introduced in #1. Compare this solution to your original solution from #1.

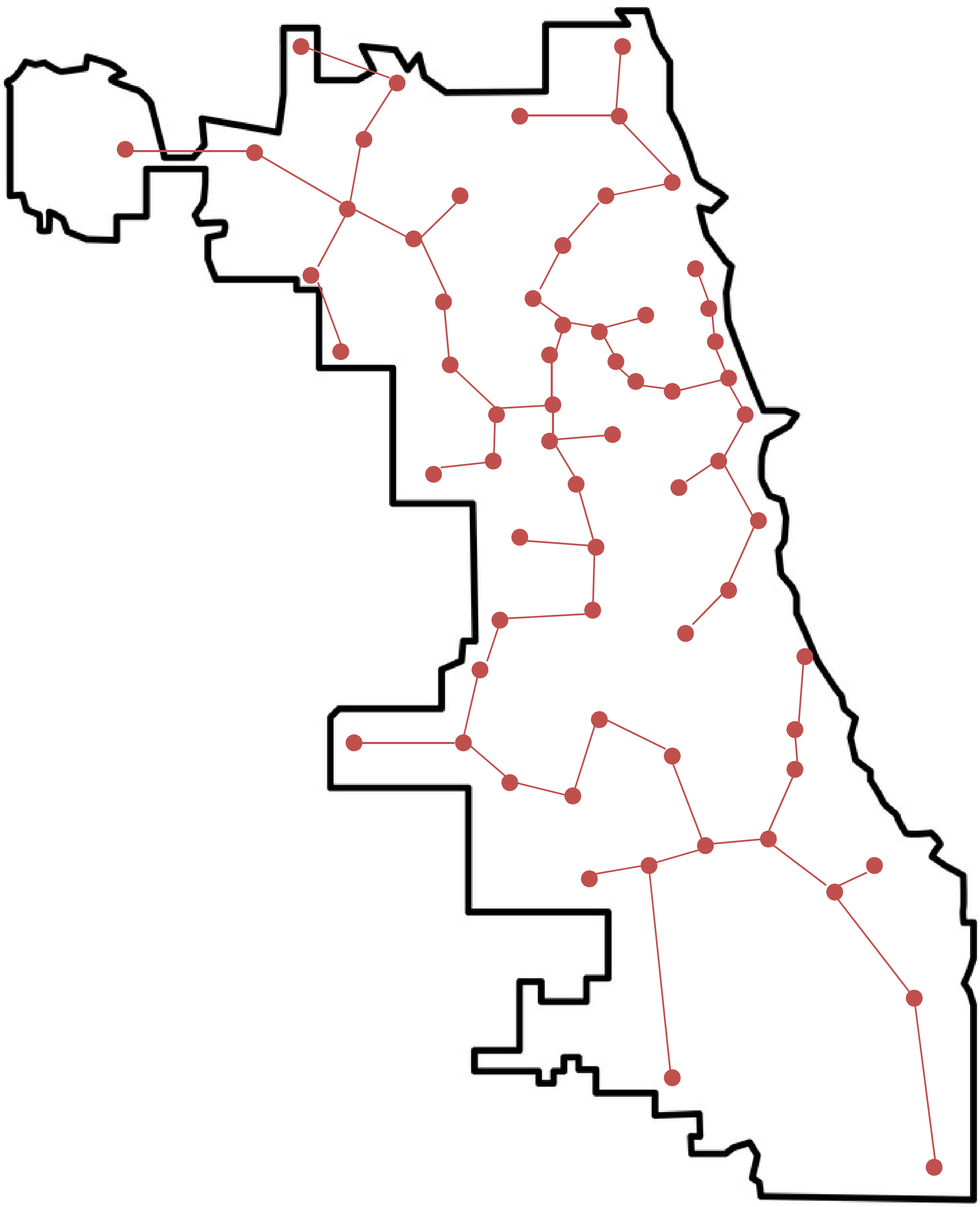
	A	B	C	D	E	F	G
A		2.3	3.7	1.2	4.8	0.9	1.5
B	2.3		3.6	2.4	1.7	2.8	3.9
C	3.7	3.6		4.0	2.5	3.0	4.1
D	1.2	2.4	4.0		2.2	3.8	0.5
E	4.8	1.7	2.5	2.2		2.6	1.3
F	0.9	2.8	3.0	3.8	2.6		4.4
G	1.5	3.9	4.1	0.5	1.3	4.4	



Open	Connected	Edge	Edge Weight	Total Weight
{B,C,D,E,F,G}	{A}	AF	0.9	0.9
{B,C,D,E,G}	{A,F}	AD	1.2	2.1
{B,C,E,G}	{A,D,F}	DG	0.5	2.6
{B,C,E}	{A,D,F,G}	GE	1.3	3.9
{B,C}	{A,D,E,F,G}	EB	1.7	5.6
{C}	{A,B,D,E,F,G}	EC	2.5	8.1

Discrete Math
Ogrodnik
Graph Theory Intro

3. Imagine Chicago is going to totally renovate their elevated train system. The map below shows the planned locations of the various stops (red dots) that must be included in the new system. You must determine the way in which the stops should be connected such that the result is the minimum cost spanning tree. Assume the map is drawn to scale.

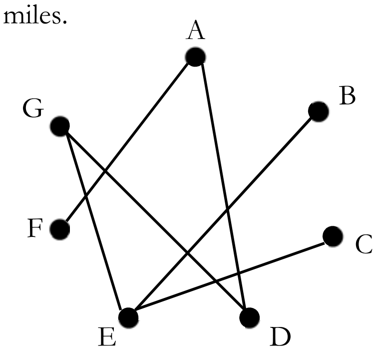


Discrete Math
Ogrodnik
Graph Theory Intro

Recall the problem below, from the Day 8 notes. The solution was found using Prim's algorithm. The result is a minimum cost spanning tree.

> A city is planning a new subway system. The system must connect the seven locations listed below. This system does not need to form a cycle; the main goal is to minimize the cost of the project while still forming a connected system. Assume the cost per mile is the same between any two locations.
>
> The distances between any two locations are shown below, in miles.

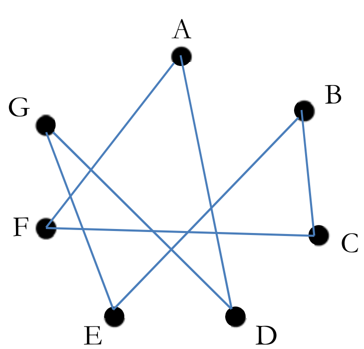
	A	B	C	D	E	F	G
A		2.3	3.7	1.2	4.8	0.9	1.5
B	2.3		3.6	2.4	1.7	2.8	3.9
C	3.7	3.6		4.0	2.5	3.0	4.1
D	1.2	2.4	4.0		2.2	3.8	0.5
E	4.8	1.7	2.5	2.2		2.6	1.3
F	0.9	2.8	3.0	3.8	2.6		4.4
G	1.5	3.9	4.1	0.5	1.3	4.4	



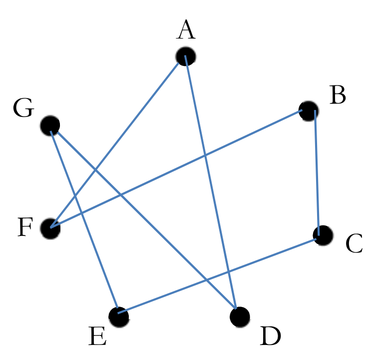
We can use a minimum cost spanning tree of a graph as the basis for finding a solution to the traveling salesperson problem for that same graph. Following the algorithm below to find a solution to the traveling salesperson problem for the locations in the above table.

Minimum Spanning Tree to Traveling Salesperson

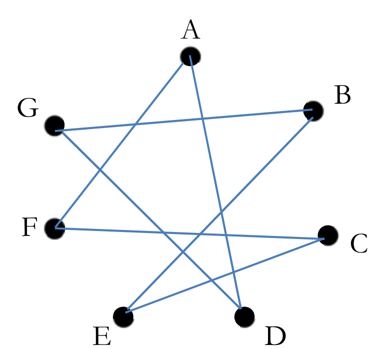
1. Complete the minimum cost spanning tree for the graph.
2. Consider an edge in the spanning tree. If the edge is connected to two vertices that each have a valence less than or equal to two, include the edge in the solution.
3. If every edge in the spanning tree has been considered, continue to step 4. Else, return to step 2.
4. Resolve every unvisited vertex that can be visited by edges from the spanning tree. When multiple variations are possible, consider each possible variation as a new potential solution.
5. For each potential solution, resolve every unvisited vertex that cannot be visited by edges from the spanning tree. When multiple variations are possible, consider each possible variation as a new potential solution.
6. Complete the cycle for every potential solution by adding any necessary edges.
7. Of all potential solutions, find the total length of the given cycle.
8. The shortest cycle of all potential solutions is the solution to the traveling salesperson cycle.



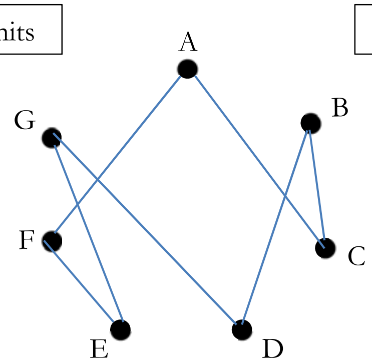
12.2 units



12.8 units



13.7 units



Nearest Neighbor: 15 units