

Network Analysis using ArcGIS: Location-Allocation

Dr. Jing Yao, Urban Big Data Centre, University of Glasgow

This is a 2-hour lab designed for location-allocation analysis (a type of network analysis) using ArcGIS. Some basic knowledge of ArcGIS is required, but experience of network analysis is not necessary. The main content of this course includes:

- Build road network
- Implement location-allocation analysis
- Interpret the analysis results

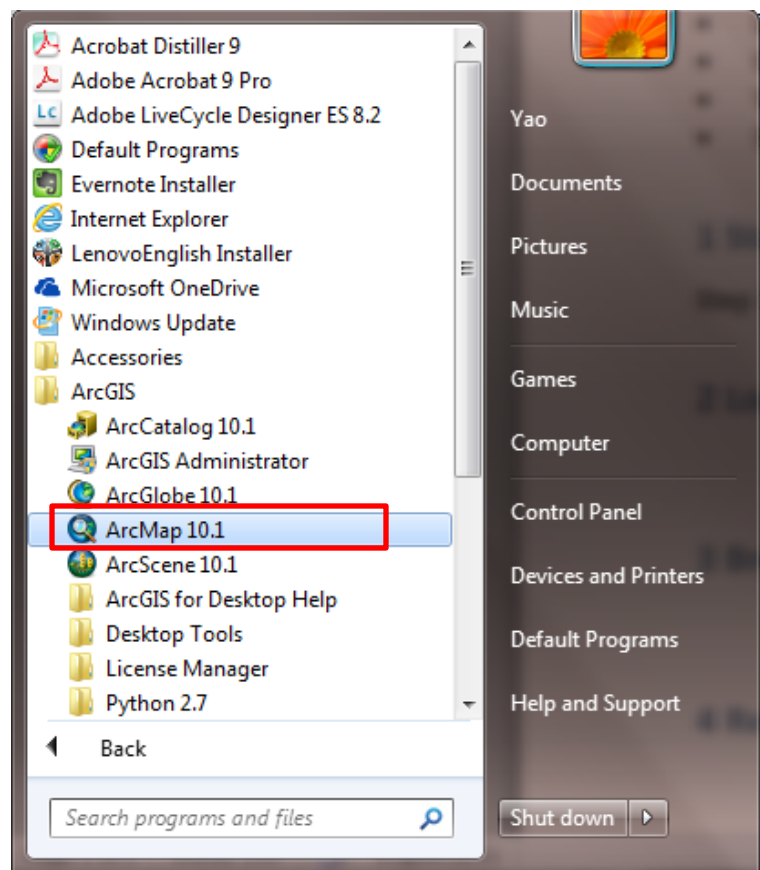
1 Getting Started

Network Analyst is an extension to ESRI's ArcGIS ArcMap application. As the name suggests, it provides a set of tools which allows us to analyze networks. In this session, we will particularly look at the location-allocation analysis. The aim is to show you what location-allocation analysis is, what kind of things can be done with it and how to go about doing them.

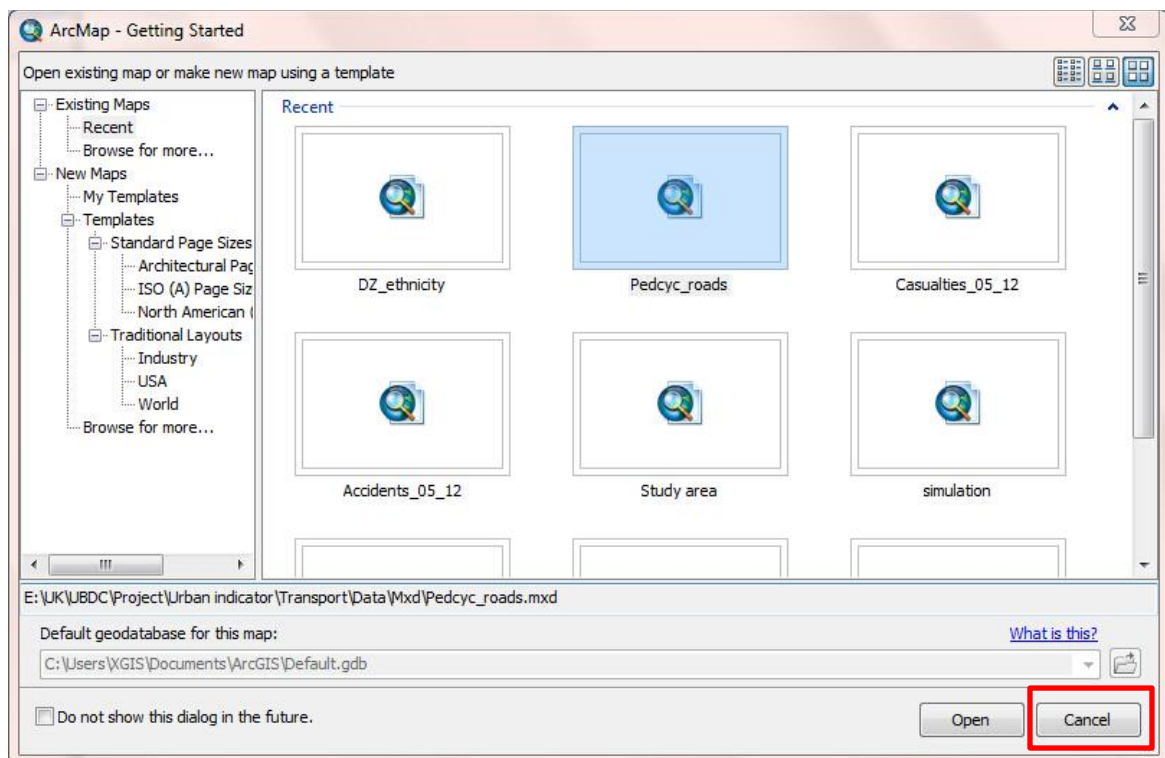
We will begin by discussing the network before moving on to the analysis. Network is used in many different ways. In this course, we will be thinking of the transportation network. The network (which may be multi-modal) connects different locations allowing people to travel between them. Moving between nodes in the network incurs some costs e.g. time, fuel, and vehicle depreciation. Network analysis in this context can be seen as an optimization problem where we attempt to minimize the cost of achieving some objective.

2 Start ArcGIS

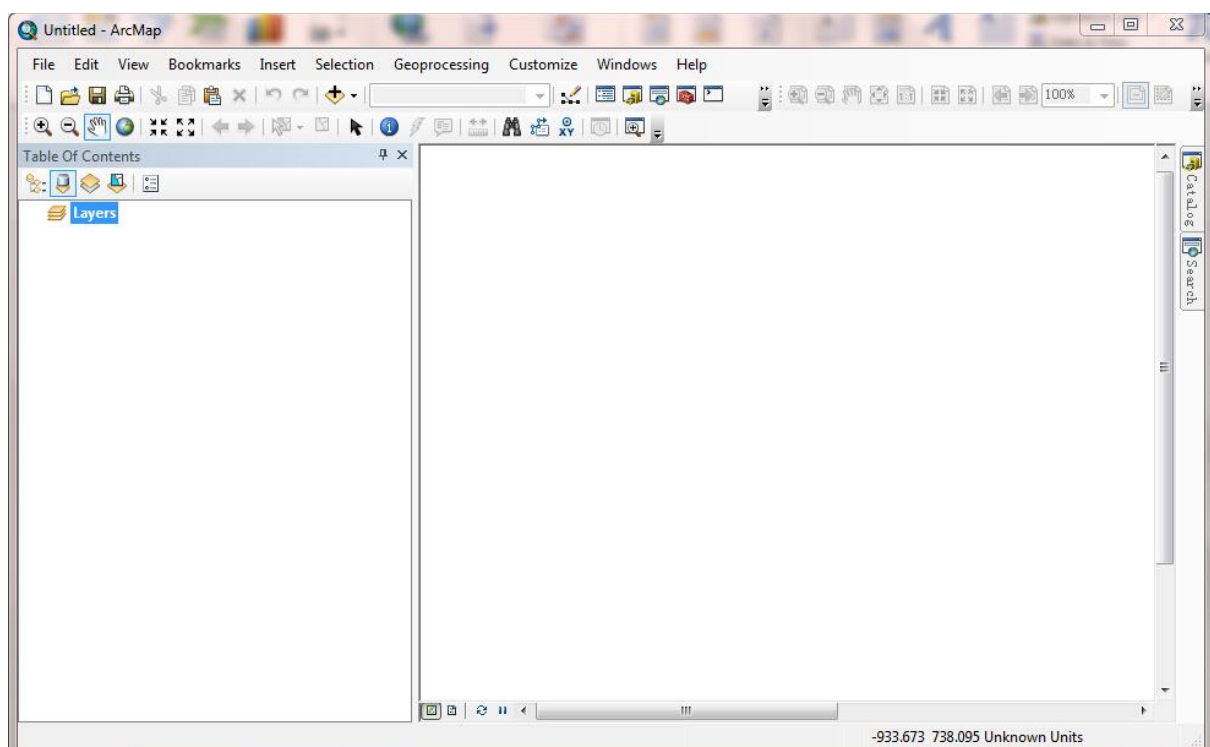
Step 1: Start **ArcGIS** from the **Start** menu. Choose the application **ArcMap**. Then, wait ...



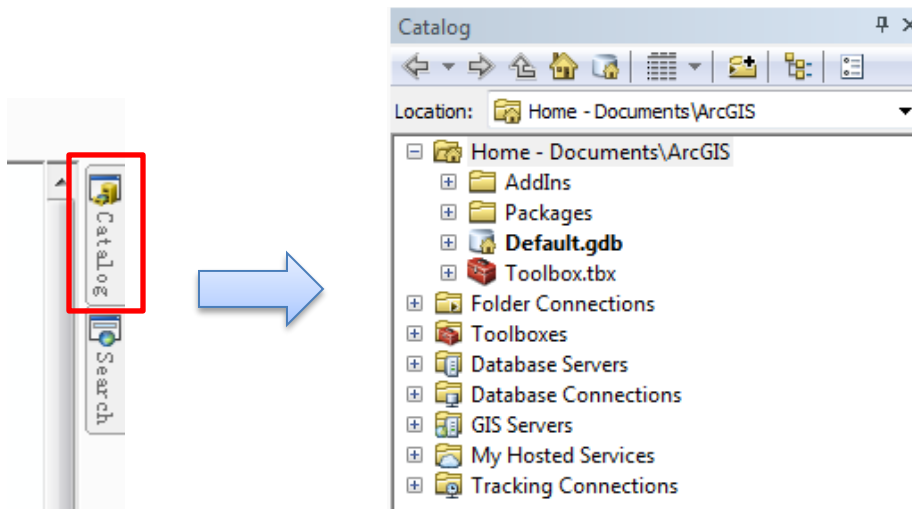
Step 2: In the *Getting Started* window, click on **Cancel**.




Step 3: Now the ArcGIS GUI should look as follows.

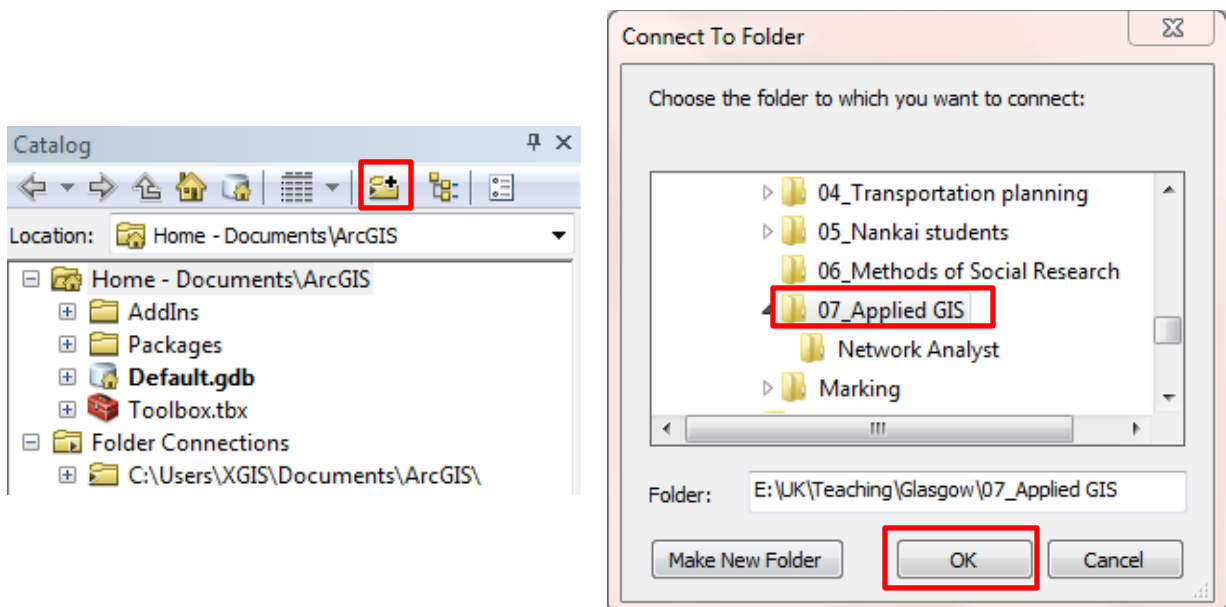


Step 4: Click on the **Catalog** tab on the right to open the **Catalog** window.




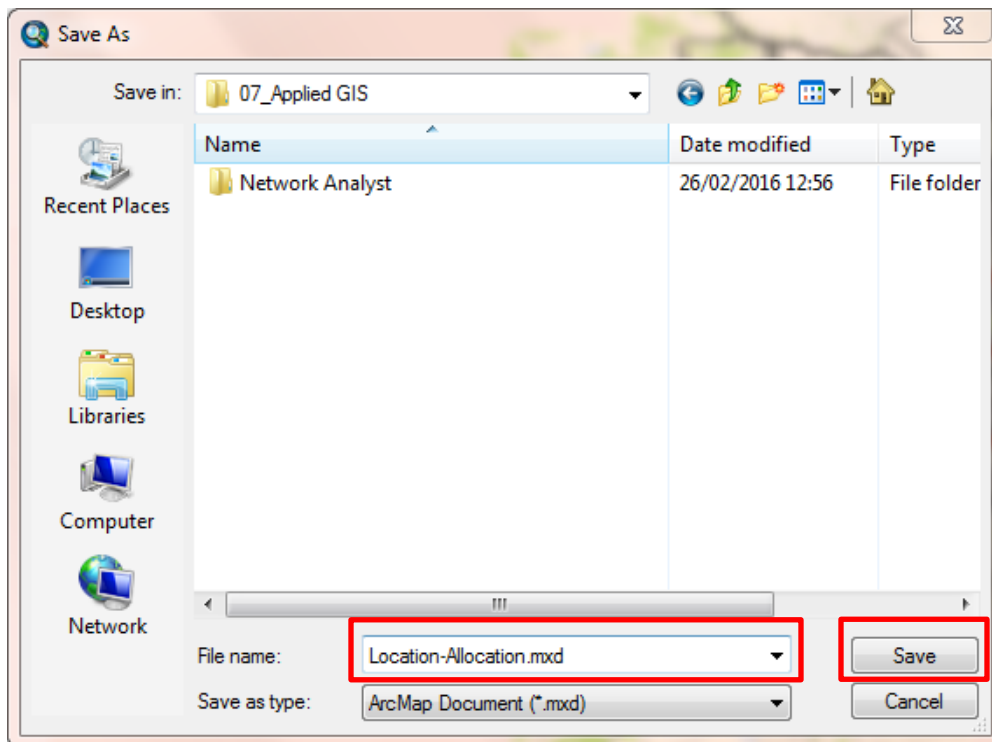
Step 5: Connect to your data using **Catalog**.

Click on  and choose the file directory where you save the data for this session (your directory might be different from mine. I suggest you save your data and all the later output to your USB disk because you don't have the permission to save files on the university's computer). Click on **OK**.



Step 6: Save the map document.

We'll save all our work in a map document, which contains the data source, colors, legend and symbols we'll use for the map. Click on . In the **Save as** window, name your map document as *Location-Allocation* (of course you can give it a different name) and choose a directory for it. Click on **Save**.




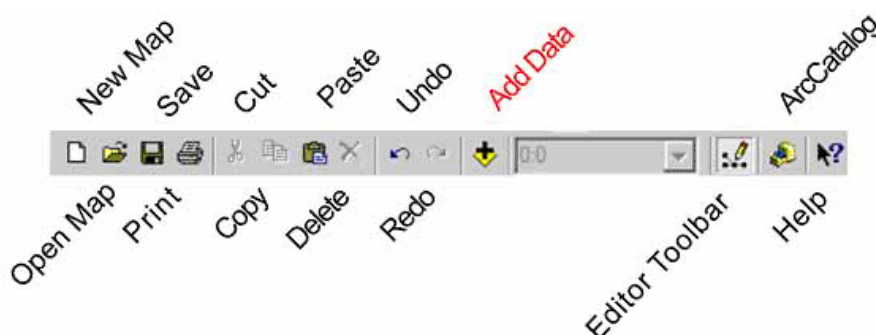
3 Create a Road Network

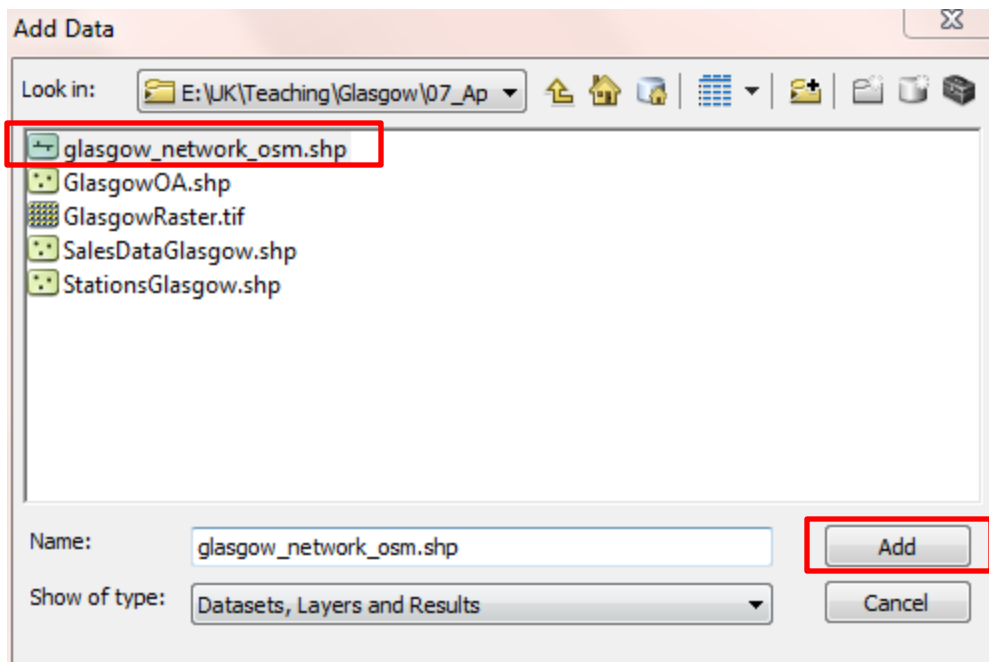
Before we can conduct any analysis, we have to define our network. The network may be provided in a variety of forms and formats. Usually we will receive one or more shapefiles. The network will consist of a set of links connecting a set of nodes. Our GIS software will use the spatial coincidence of nodes and links to define the connectivity of the network. We may also have additional data to help us in our analysis. For example, information on speed limits, average speeds at different times of day, turn restrictions, one-way streets, height/weight restrictions etc.

One option for creating a network is to use the data from **OpenStreetMap** (<http://www.openstreetmap.org/>). You can find information about how to download the data at <http://wiki.openstreetmap.org/>.

Step 1: Add data.

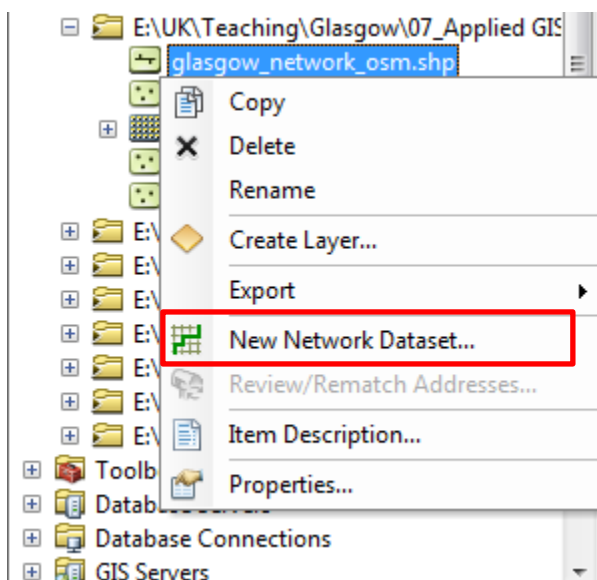
We will begin by adding the shapefile of Glasgow's road network into our layout. Click on the command . This will open the **Add Data** window. Add the shapefile *glasgow_network_osm*.





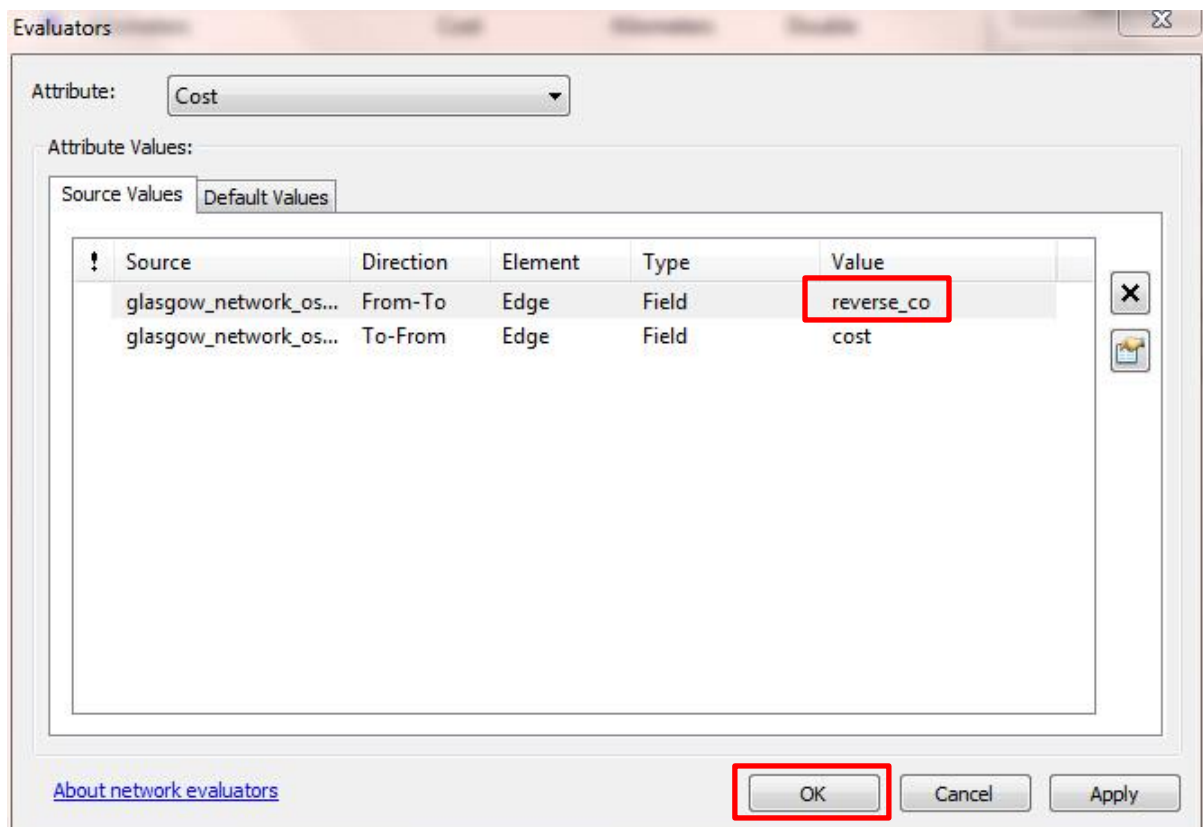
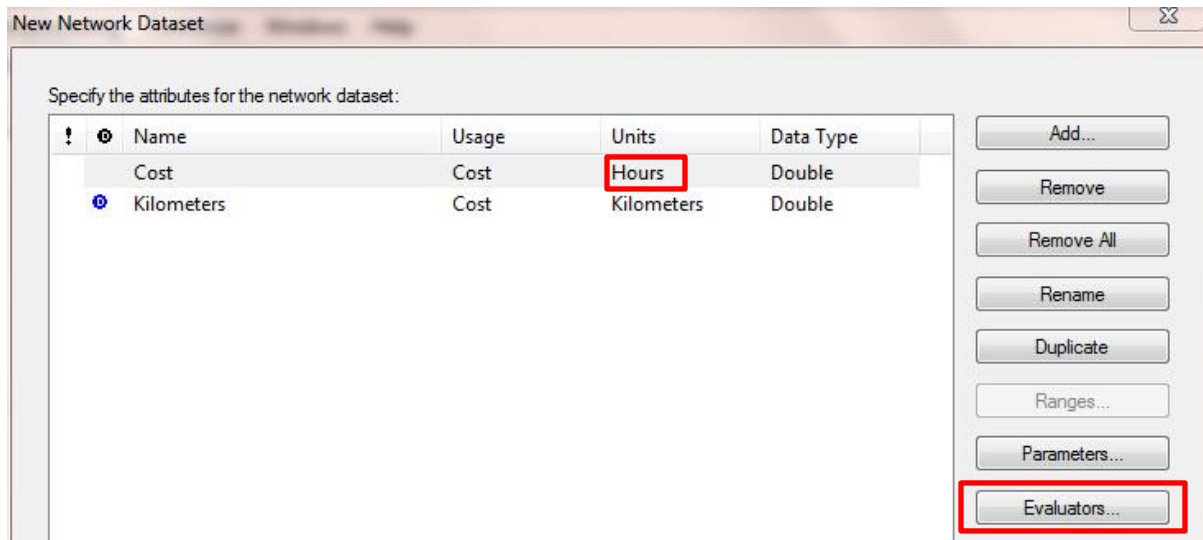
Step 2: Create a new network dataset.

- I. Using the Catalog window, navigate to the folder where the shapefile *glasgow_network_osm* is stored. Right-click on the file name and then click on **New Network Dataset**. Give the network a name such as *Glasgow_network*. Select **No** when asked if you want to model turns in the network. Accept the default specification of connectivity. Select **None** when asked if you want to model elevation.



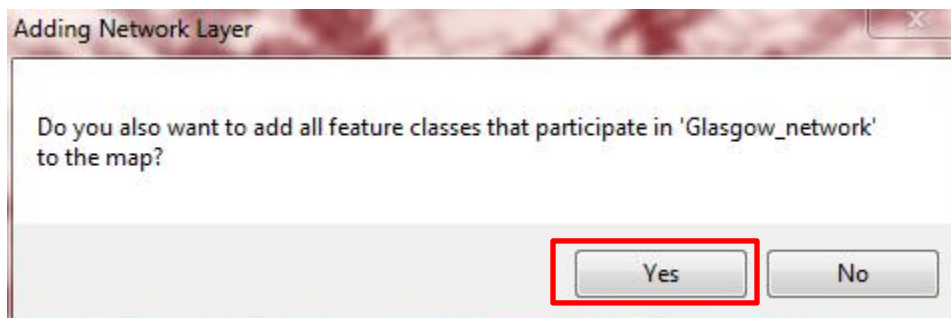
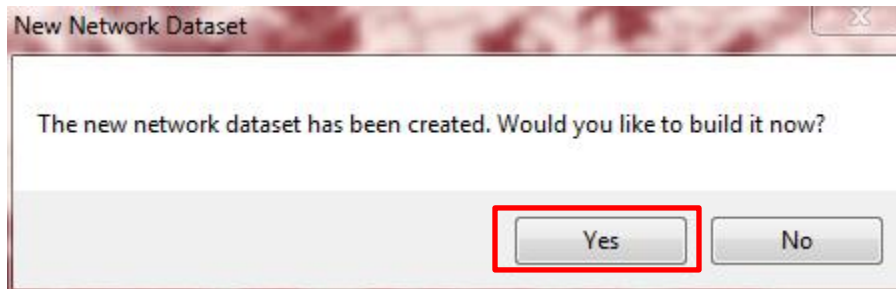
- II. When we get to the window asking us about attributes of the dataset, we can see a *Cost* and *Kilometers* attribute. The *Kilometers* attribute is measured directly from the shapefile. The *Cost* attribute which we are presented with is based on data attached to the shapefile about travel time by car. **Network Analyst** has tried to guess what the *Cost* attribute is. It has correctly identified it as a cost, but it doesn't know the units. In the units field, select *Hours* (since the data in the shapefile is time measured in hours). There is one more setting which we have to

change. Make sure that the *Cost* attribute is selected and then click on the **Evaluators...** button. *Network Analyst* models the cost by the direction of travel, so the speed of going from *X* to *Y* may not be the same as going from *Y* to *X*. Notice that the same field (*cost*) is being used for both directions at the moment. We can do better than this using the data. In the *From-To* line change the **Value** from *cost* to *reverse_co* (this should read *reverse_cost* but the field name has been shortened). Once you have done this, click on **OK**. Then click on **Next**.



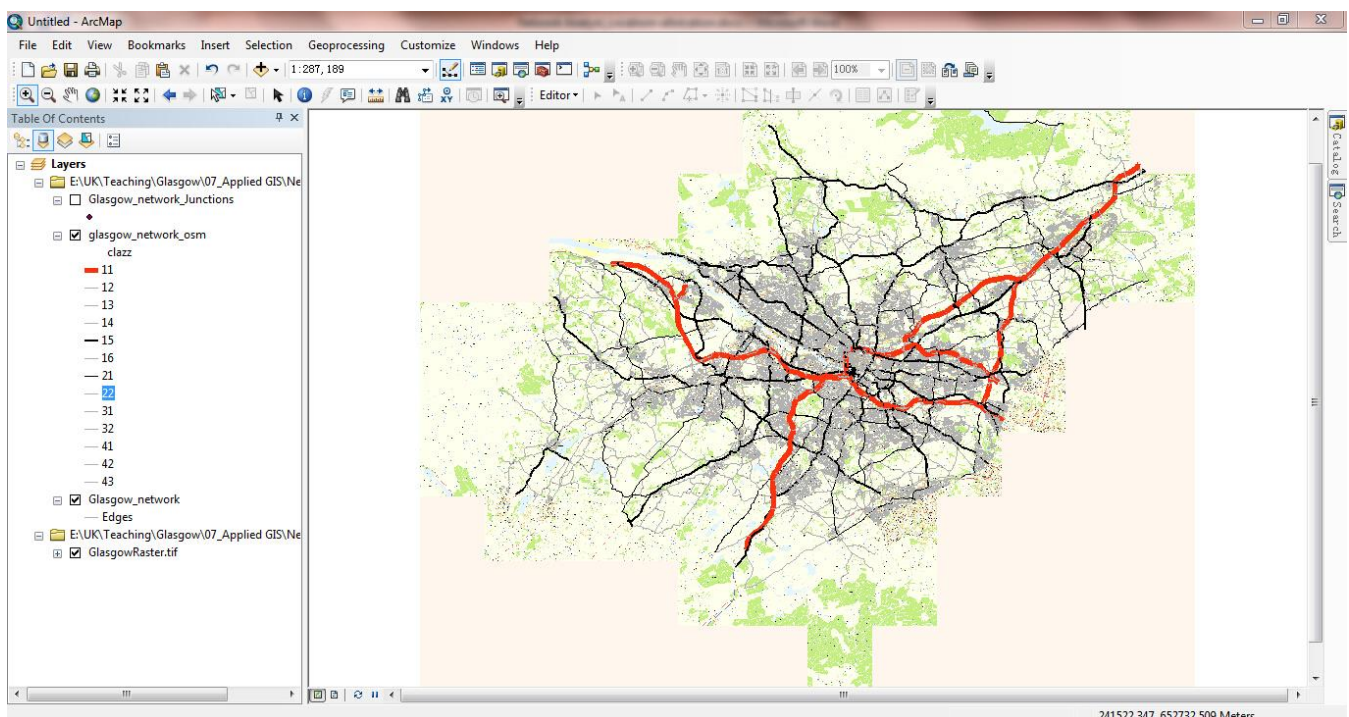
- III. We do not want to establish driving directions so select **No** when asked. Accept the summary by clicking on **Finish** and then build the network when asked and add it to the layout. Hide all of the features except the roads. Change the symbology on the roads to improve the display. Although not perfect, the *clazz*

attribute gives some details about the type of road. A value of 11 is a motorway, 15 is an **A** road and 21 is a **B** road. Choose an appropriate way to display these. Note that these symbology changes have no impact on the analysis, they are purely cosmetic. They may help us get out bearings.



4 Location-allocation Analysis

Before beginning with our analysis, make sure that you have our built OSM network added to the layout. Remove any others which you may have added previously. I suggest displaying only the network and hiding the junctions. In order to provide some context, we will add a raster map of Glasgow to the background of our layout. Add the *GlasgowRaster.tif* file to the layout.



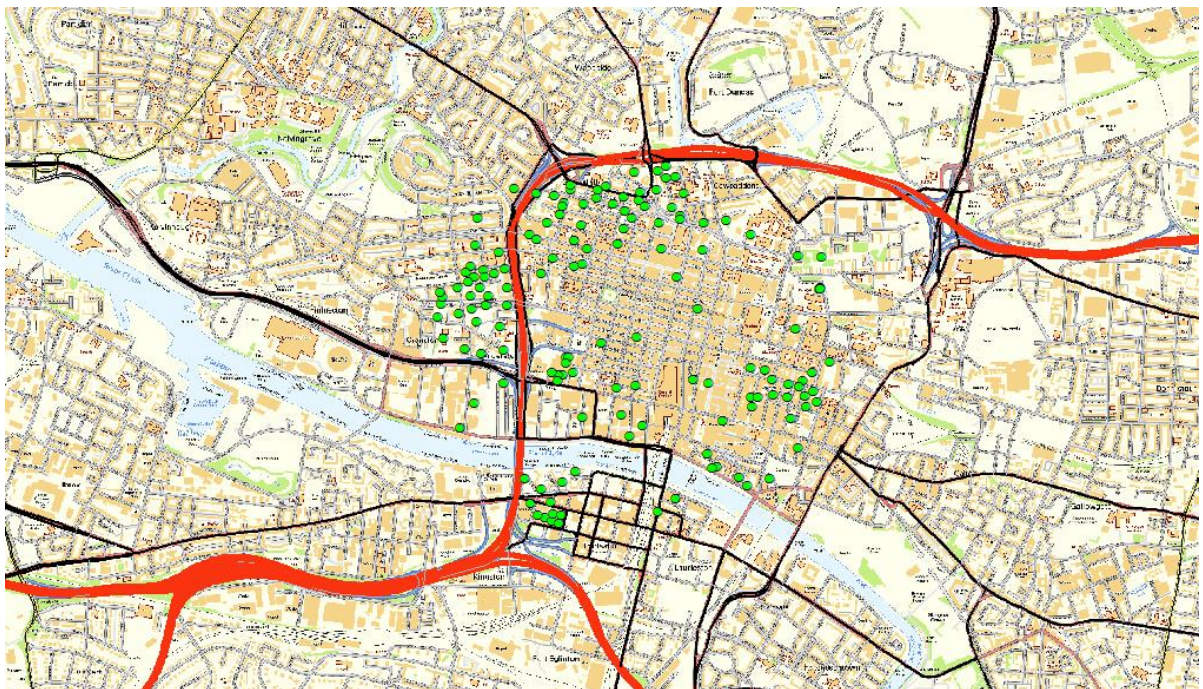
The **Location-Allocation** tool can be used to solve several related but distinct kinds of problems. There are two stages to solving these problems. Firstly, a number of facilities must be located from a set of feasible locations. Secondly, demand must be allocated to these facilities. Depending on what sort of problem we are trying to solve, the rules for locating the facilities will vary. We will look at a couple of different problems we could solve. We will work with small example since solving these kinds of problems on a large scale can be computationally demanding.

Example 1

For our first example, suppose that we want to locate some recycling facilities (glass recycling bins for example) in Glasgow city centre. Suppose further that we have decided that everyone should be within a 500m distance of a recycling bin and we want to know a) what is the smallest number of bins we could use to do this, and b) where should we place them.

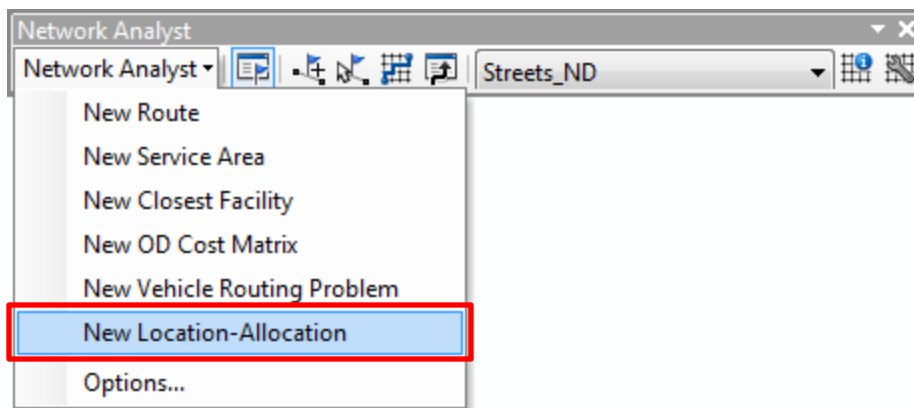
Step 1: Add data.

Add *GlasgowOA* to the layout. This shapefile has centroids for the output areas in Glasgow. The shapefile has data on the number of people living in each zone (actually this is based on the number of workers living in the area).

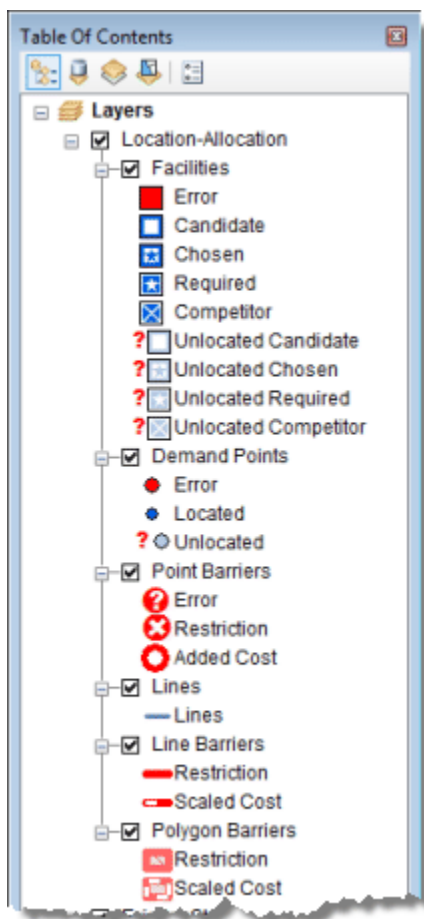


Step 2: Add a new location-allocation layer.

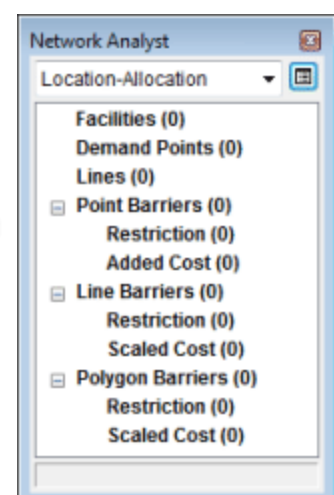
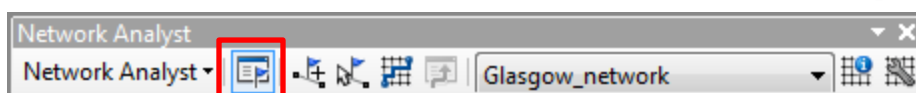
Using the **Network Analyst** toolbar (go to the menu **Customize** → **Toolbars** → **Network Analyst**) to add a **New Location-Allocation** layer to the layout.



The location-allocation analysis layer also appears in the **table of contents** as a composite layer containing six corresponding feature layers: *Facilities*, *Demand Points*, *Lines*, *Point Barriers*, *Line Barriers*, and *Polygon Barriers*. Each of the six feature layers has default symbology that can be modified in its **Layer Properties** window.



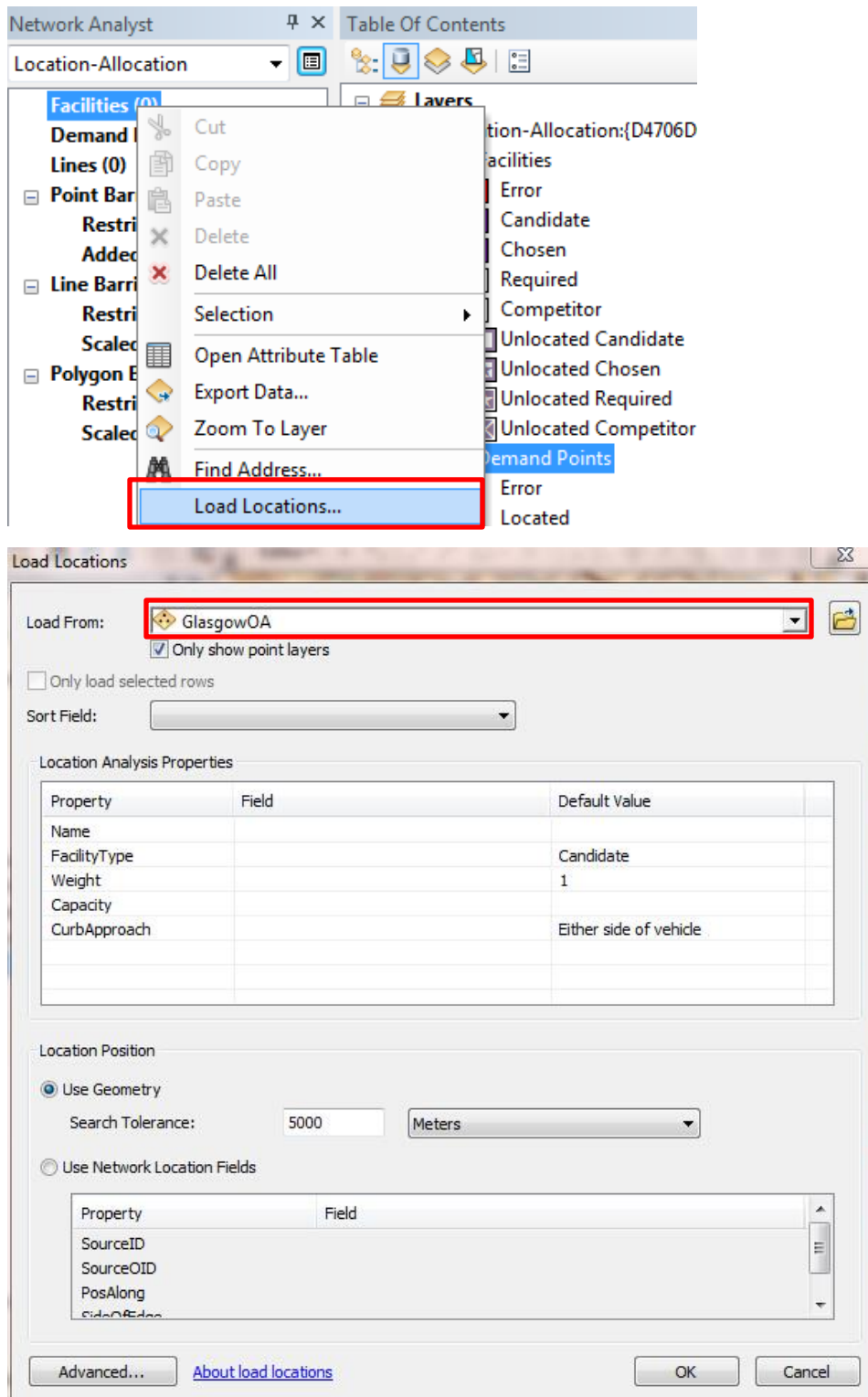
Click on  to open the *Network Analyst* window.



Step 3: Load facilities.

We will begin by loading some locations into the **Facilities**. These are not the locations of the facilities, but represent the possible/feasible locations. In this case, we will assume that any of the output areas are suitable locations.

Right click on **Facilities**→**Load Locations...** to load in the output area centroids as **Facilities**.



Step 4: Load demand points.

Next we will load in demand points. This is where our demand nodes are located. In this case, we assume that everyone living in an output area will generate demand. Open the **Load Locations** window like before and load the output area centroids. As not all of the output areas have the same number of people living in them, we will define a weight in the **Location Analysis Properties** box. Select *Population* in the **Weight** field. Click on **OK** when you have done this to load the locations.

Load Locations

Load From: GlasgowOA

☒ Only show point layers

☐ Only load selected rows

Sort Field:

Location Analysis Properties

Property	Field	Default Value
Name		
Weight	Population	1
GroupName		
ImpedanceTransformat...		
ImpedanceParameter		
CurbApproach		Either side of vehicle
Cutoff_Kilometers		
Cutoff_Cost		

Location Position


☒ Use Geometry

Search Tolerance: 5000 Meters

☐ Use Network Location Fields

Property	Field
SourceID	
SourceOID	
PosAlong	
SideOfEdge	

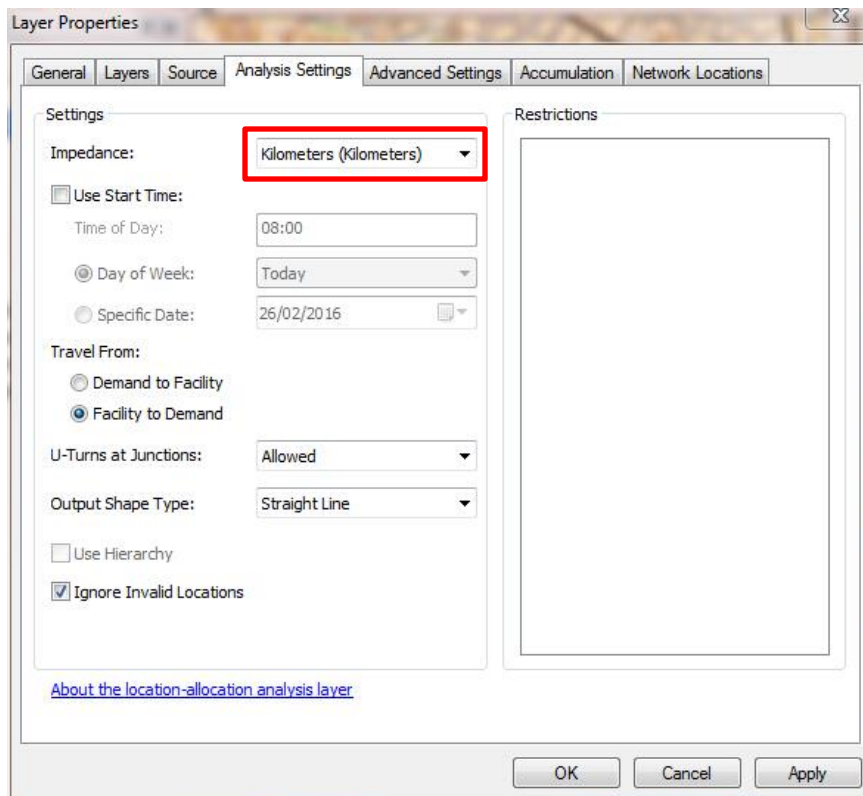
Advanced... [About load locations](#) OK Cancel

Step 5: Click on  to open the **Layer Properties** window. Make sure that **impedance** is set to *Kilometers*.

Network Analyst

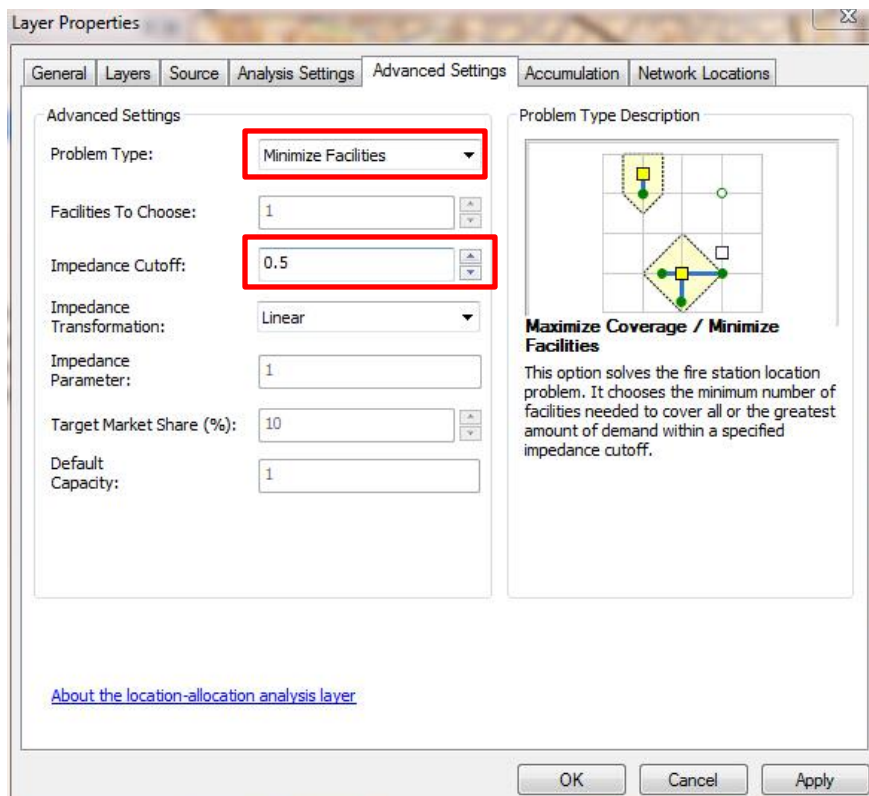
Location-Allocation

- Facilities (146)
- Demand Points (146)
- Lines (0)
- Point Barriers (0)
 - Restriction (0)
 - Added Cost (0)
- Line Barriers (0)
 - Restriction (0)
 - Scaled Cost (0)
- Polygon Barriers (0)
 - Restriction (0)
 - Scaled Cost (0)




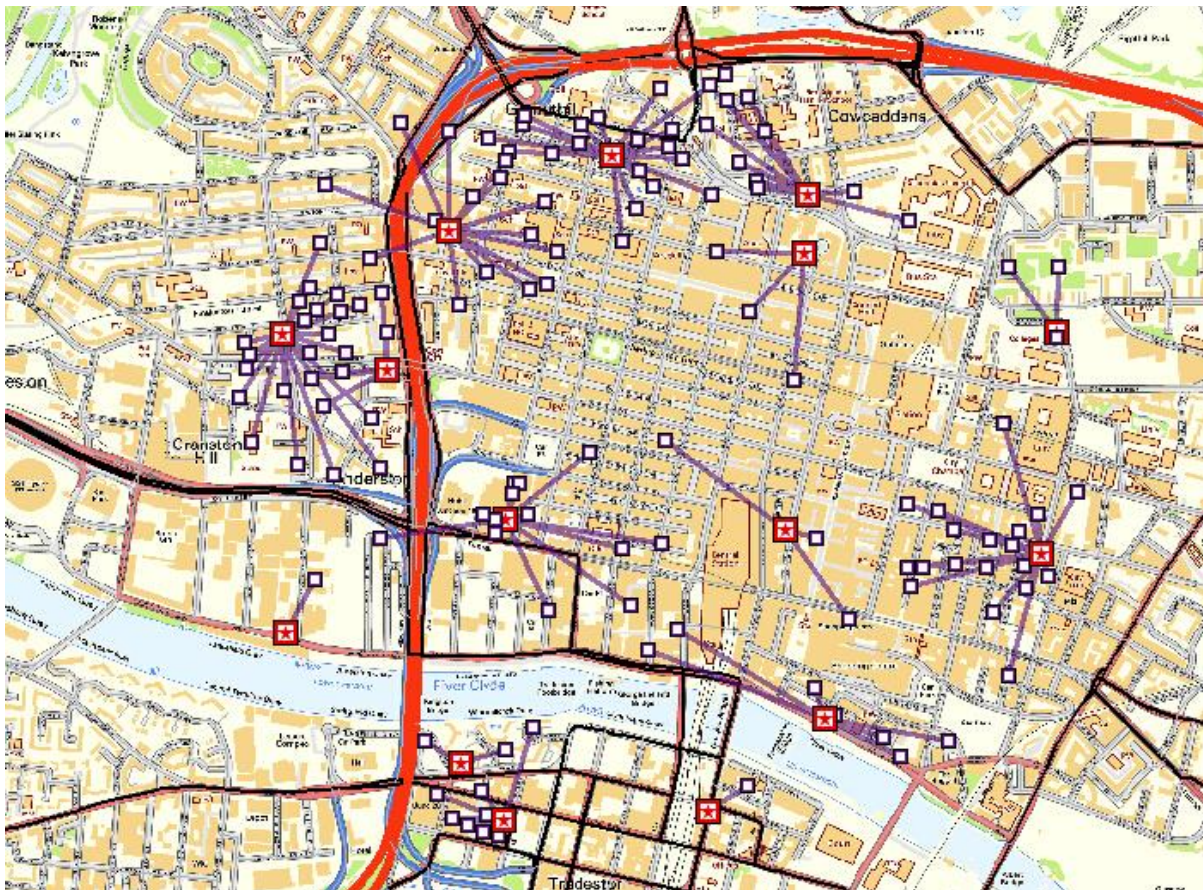
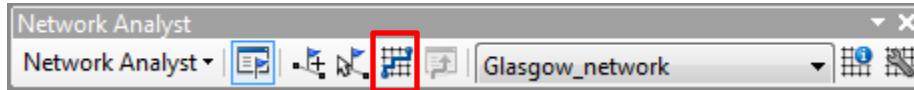
Step 6: Model settings.

Click on the **Advance Settings** tab. This is where we select the type of problems we would like to solve and specify the parameters. For **Problem Type** select *Minimize Facilities*. We want every node to be within 500m of a facility, so we set the **Impedance Cutoff** to 0.5 (since we have measures Km in the network). Click on **OK**.



Step 7: Solve the location-allocation problem.


Click on  to solve the problem. You will now see the minimum number of facilities you need to have all demand points with 500m of a facility. You will also see how the demand points should be assigned to facilities.




Example 2

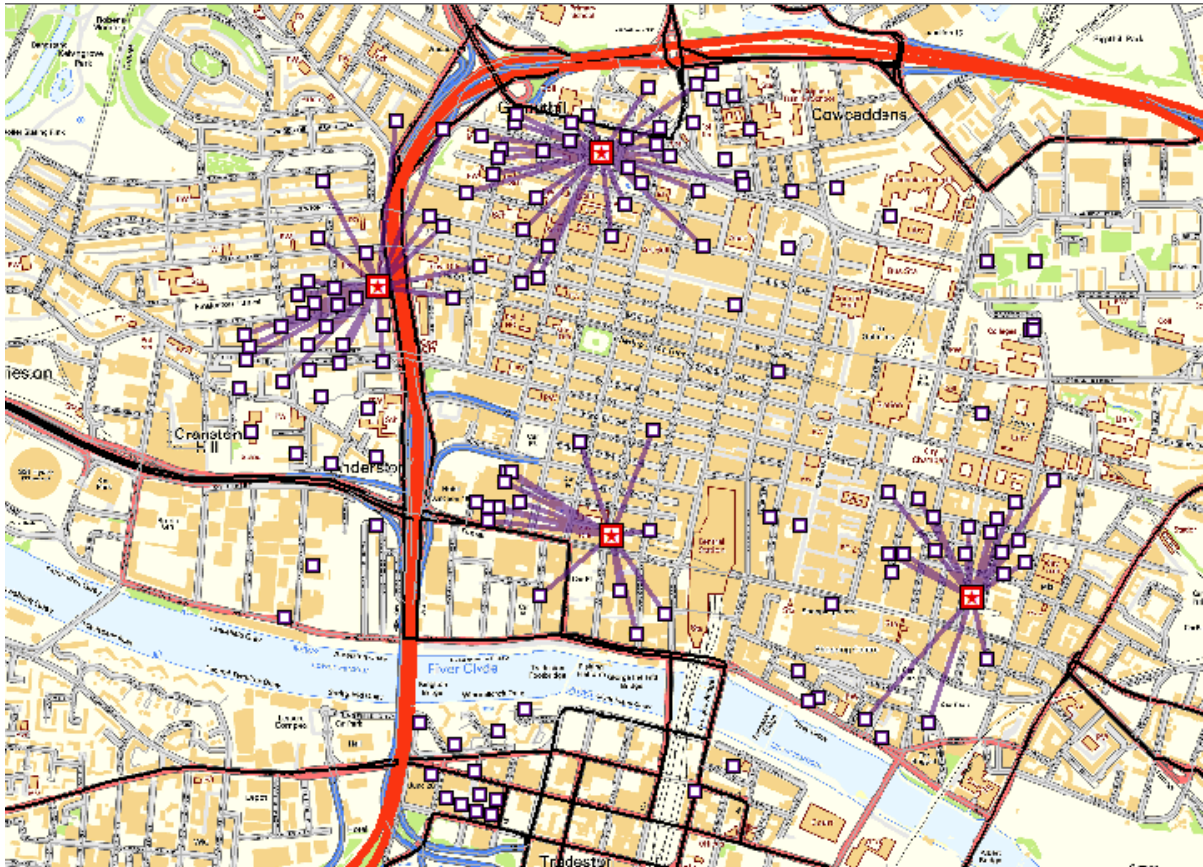
We will now consider a related problem. Imagine that we have a budget to have four recycling facilities and we want to know where to locate them.

Step 1: Change Model settings.

Click on  to open the **Layer Properties** window. Open the **Advanced Settings** tab. In the **Problem Type** box select *Maximize Coverage*. We want to locate four facilities, so in the **Facilities To Choose** box type 4. We want to maximize the number of people within 500m of a facility, so type 0.5 into the **Impedance Cutoff** box. Click on **OK**.

Step 2: Solve the location-allocation problem.


Click on  to generate the solution. You will be presented with a solution showing where the four facilities should be located to maximize the number of people with 500m of a facility.




Example 3

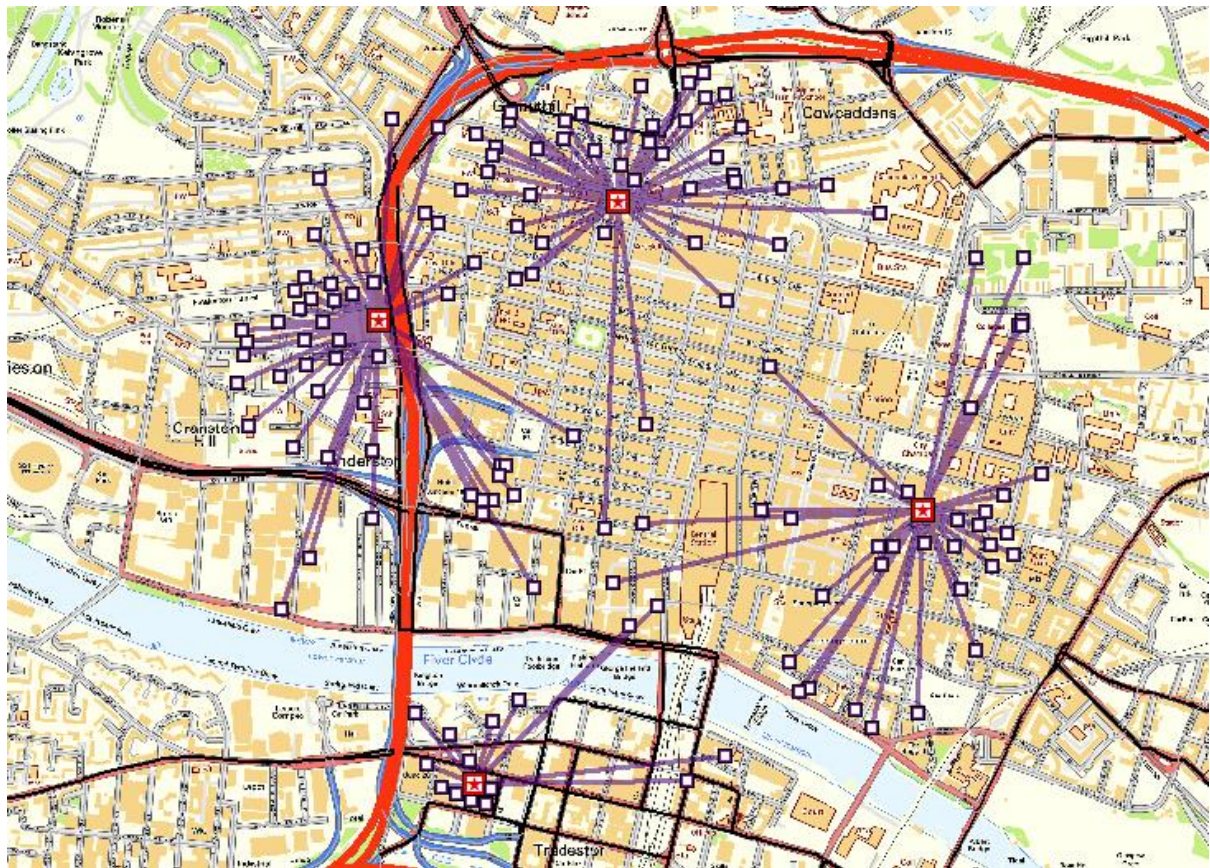
Rather than maximizing the number of people within 500m of a facility, suppose we wanted to minimize the weighted distance from each demand node to each facility.

Step 1: Change Model settings.

Click on  to open the **Layer Properties** window. Open the **Advanced Settings** tab. In the **Problem Type** box select *Minimize Impedance*. We will locate 4 facilities. This time we will not use an **Impedance Cutoff**, so delete anything in the box so that it reverts to <None>. Click on **OK**.

Step 2: Solve the location-allocation problem.

Click on . You will now see where the facilities should be located in order to minimize the total weighted distance. Each demand node will be allocated to its closest facility.



Appendix

Location-allocation

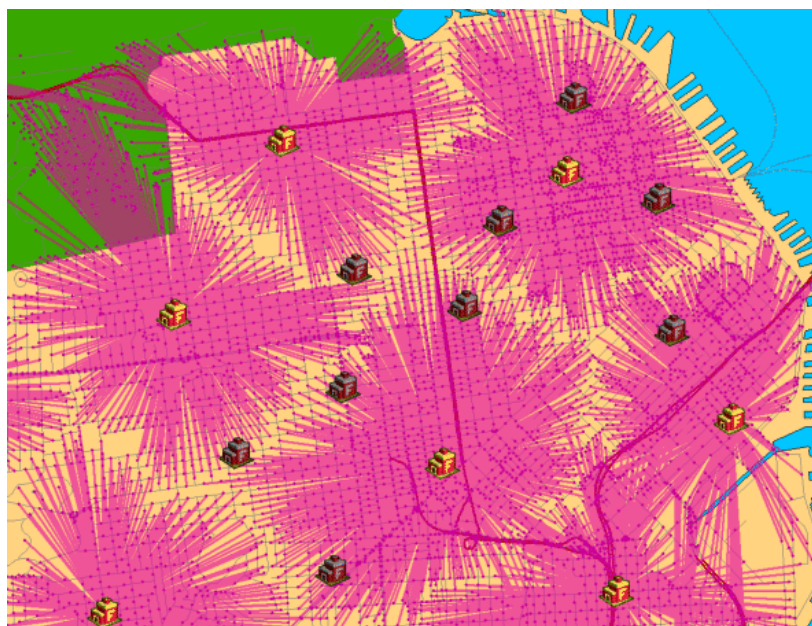
Location-allocation helps you choose which facilities from a set of facilities to operate based on their potential interaction with demand points. It can help you answer questions like the following:

- Given a set of existing fire stations, which site for a new fire station would provide the best response times for the community?
- If a retail company has to downsize, which stores should it close to maintain the most overall demand?
- Where should a factory be built to minimize the distance to distribution centers?

In these examples, facilities would represent the fire stations, retail stores, and factories; demand points would represent buildings, customers, and distribution centers.

The objective may be to minimize the overall distance between demand points and facilities, maximize the number of demand points covered within a certain distance of facilities, maximize an apportioned amount of demand that decays with increasing distance from a facility, or maximize the amount of demand captured in an environment of friendly and competing facilities.

The map below shows the results of a location-allocation analysis meant to determine which fire stations are redundant. The following information was provided to the solver: an array of fire stations (facilities), street midpoints (demand points), and a maximum allowable response time. The response time is the time it takes firefighters to reach a given location. The location-allocation solver determined that the fire department can close several fire stations and still maintain a three-minute response time.

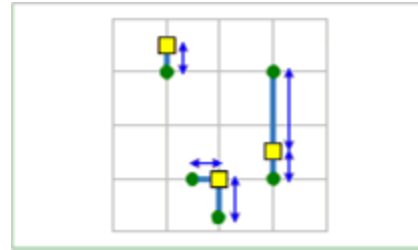


Out of the current set of fire stations, nine fire stations can close, and a minimum of seven are needed for the department to still be able to respond to emergencies within three minutes. (Source: ArcGIS Desktop Help 10.1)

Location-allocation problem types

Minimize Impedance (P-Median)

Facilities are located such that the sum of all weighted costs between demand points and solution facilities is minimized. The arrows in the graphic below highlight the fact that allocation is based on distance among all demand points.



Minimize Impedance chooses facilities such that the sum of weighted impedances (demand allocated to a facility multiplied by the impedance to the facility) is minimized.

This problem type is traditionally used to locate warehouses, because it can reduce the overall transportation costs of delivering goods to outlets. Since Minimize Impedance reduces the overall distance the public needs to travel to reach the chosen facilities, the minimize impedance problem without an impedance cutoff is ordinarily regarded as more equitable than other problem types for locating some public-sector facilities such as libraries, regional airports, museums, department of motor vehicles offices, and health clinics.

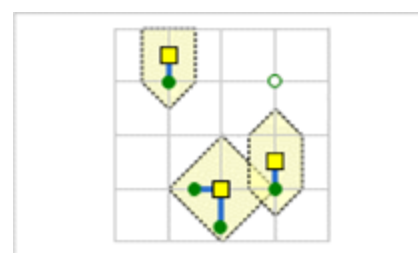
The following list describes how the minimize impedance problem type handles demand:

- If an impedance cutoff is set, any demand outside all the facilities' impedance cutoffs is not allocated.
- A demand point inside the impedance cutoff of one facility has all its demand weight allocated to that facility.
- A demand point inside the impedance cutoff of two or more facilities has all its demand weight allocated to the nearest facility only.

Maximize Coverage

Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff.

Maximize Coverage is frequently used to locate fire stations, police stations, and ERS centers, because emergency services are often required to arrive at all demand points within a specified response time. Note that it is important for all organizations, and critical for emergency services, to have accurate and precise data so that analysis results correctly model real-world results.



Maximize Coverage chooses facilities such that as much demand as possible is covered by the impedance cutoff of facilities. In this graphic, the solver was directed to choose three facilities.

Pizza delivery businesses, as opposed to eat-in pizzerias, try to locate stores where they can cover the most people within a certain drive time. People who order pizzas for delivery don't typically worry about how far away the pizzeria is; they are mainly

concerned with the pizza arriving within an advertised time window. Therefore, a pizza-delivery business would subtract pizza-preparation time from their advertised delivery time and solve a maximize coverage problem to choose the candidate facility that would capture the most potential customers in the coverage area. (Potential customers of eat-in pizzerias are more affected by distance, since they need to travel to the restaurant; thus, the attendance maximizing or market share problem types would better suit eat-in restaurants.)

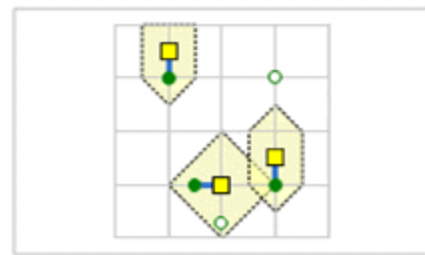
The following list describes how the Maximize Coverage problem handles demand:

- Any demand point outside all the facilities' impedance cutoffs is not allocated.
- A demand point inside the impedance cutoff of one facility has all its demand weight allocated to that facility.
- A demand point inside the impedance cutoff of two or more facilities has all its demand weight allocated to the nearest facility only.

Maximize Capacitated Coverage

Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff; additionally, the weighted demand allocated to a facility can't exceed the facility's capacity.

Maximize Capacitated Coverage behaves like either the Minimize Impedance or Maximize Coverage problem type but with the added constraint of capacity. (If **Impedance Cutoff** is set to <none>, it behaves like a capacitated version of Minimize Impedance.) You can specify a capacity for a facility by assigning a numeric value to its Capacity property. If the Capacity property is null, the facility is assigned a capacity from the **Default Capacity** property of the analysis layer.



Use-cases for Maximize Capacitated Coverage include creating territories that encompass a given number of people or businesses, locating hospitals or other medical facilities with a limited number of beds or patients who can be treated, or locating warehouses whose inventory isn't assumed to be unlimited.

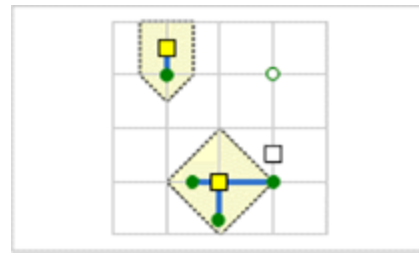
The following list describes how the Maximize Capacitated Coverage problem handles demand:

- Unlike Maximize Coverage, Maximize Capacitated Coverage doesn't require an impedance cutoff; however, when an impedance cutoff is specified, any demand point outside all the facilities' impedance cutoffs is not allocated.
- An allocated demand point has all or none of its demand weight assigned to a facility; that is, demand isn't apportioned with this problem type.
- If the total demand within the impedance cutoff of a facility is greater than the capacity of the facility, only the demand points that maximize total captured demand and minimize total weighted impedance are allocated.

Minimize Facilities

Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff; additionally, the number of facilities required to cover demand points is minimized.

Minimize Facilities is the same as Maximize Coverage but with the exception of the number of facilities to locate, which in this case is determined by the solver. When the cost of building facilities is not a limiting factor, the same kinds of organizations that use Maximize Coverage (emergency response, for instance) use Minimize Facilities so that all possible demand points will be covered. Minimize Facilities is also used to choose school bus stops when students are required to walk a certain distance before another school bus stop is provided closer to the student's residence.



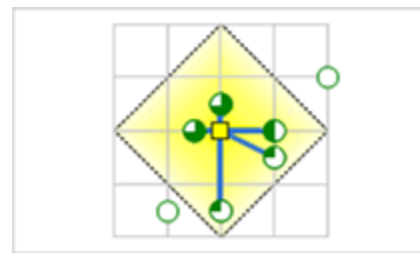
Minimize Facilities chooses facilities such that as many demand points as possible are within the impedance cutoff of facilities. Additionally, the number of facilities required to cover all demand points is minimized. In this graphic, the solver was able to cover all demand points with only two facilities.

The following list describes how the Minimize Facilities problem handles demand:

- Any demand point outside all the facilities' impedance cutoffs is not allocated.
- A demand point inside the impedance cutoff of one facility has all its demand weight allocated to that facility.
- A demand point inside the impedance cutoff of two or more facilities has all its demand weight allocated to the nearest facility only.

Maximize Attendance

Facilities are chosen such that as much demand weight as possible is allocated to facilities while assuming the demand weight decreases in relation to the distance between the facility and the demand point.



Specialty stores that have little or no competition benefit from this problem type, but it may also be beneficial to general retailers and restaurants that don't have the data on competitors necessary to perform market share problem types. Some businesses that might benefit from this problem type include coffee shops, fitness centers, dental and medical offices, bowling alleys, and electronics stores. Public transit bus stops are often chosen with the help of Maximize Attendance. Maximize Attendance assumes that the farther people have to travel to reach your facility, the less likely they are to use it. This is reflected in how the amount

Maximize Attendance chooses facilities such that as much demand weight as possible is allocated to facilities while assuming the demand weight decreases with distance. The demand points, represented by pie charts in this graphic, show how much of their total demand is captured by the facility.

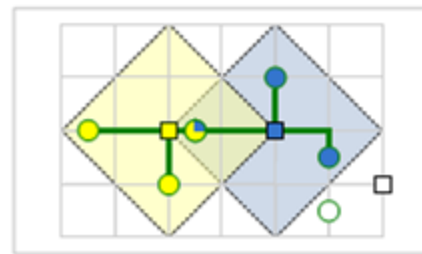
of demand allocated to facilities diminishes with distance. You specify the distance decay with the impedance transformation.

The following list describes how the Maximize Attendance problem handles demand:

- Demand outside the impedance cutoff of all facilities is not allocated to any facility.
- When a demand point is inside the impedance cutoff of one facility, its demand weight is partially allocated according to the cutoff and impedance transformation. The demand points in the graphic above have pie charts to represent the ratio of their total demand weight that was captured by the chosen facility.
- The weight of a demand point covered by more than one facility's impedance cutoff is allocated only to the nearest facility.

Maximize Market Share

A specific number of facilities are chosen such that the allocated demand is maximized in the presence of competitors. The goal is to capture as much of the total market share as possible with a given number of facilities, which you specify. The total market share is the sum of all demand weight for valid demand points.



Maximize Market Share chooses facilities such that the largest amount of allocated demand is captured in the presence of competitors. You specify the number of facilities you want it to choose.

The market share problem types require the most data because, along with knowing your own facilities' weight, you also need to know that of your competitors' facilities. The same types of facilities that use the Maximize Attendance problem type can also use market share problem types given that they have comprehensive information that includes competitor data. Large discount stores typically use Maximize Market Share to locate a finite set of new stores. The market share problem types use a Huff model, which is also known as a gravity model or spatial interaction.

The following list describes how the Maximize Market Share problem handles demand:

- Demand outside the impedance cutoff of all facilities is not allocated to any facility.
- A demand point inside the impedance cutoff of one facility has all its demand weight allocated to that facility.
- A demand point inside the impedance cutoff of two or more facilities has all its demand weight allocated to the facilities that cover it; furthermore, the weight is split among the facilities proportionally to the facilities' attractiveness (facility weight) and inversely proportional to the distance between the facility and demand point. Given equal facility weights, this means more demand weight is assigned to near facilities than far facilities. This behavior is demonstrated in the Maximize Market Share graphic above. Assume the three facilities (squares) have equivalent weights and note that one of the six demand points (circles) is within

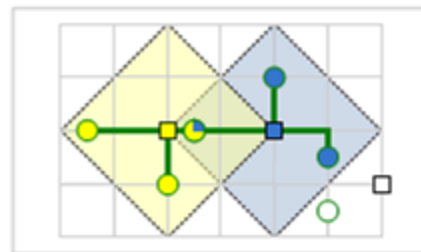
the impedance cutoff of two competing facilities and has its demand split between the facilities. (The amount of yellow or blue in the demand point correlates with the percent of demand weight captured by the facility of the same color.) The demand point near the center of the graphic is covered by both the yellow facility on the left and the blue facility in the center. Since the demand point is closer to the yellow facility, more of the demand is allocated to that facility.

The demand point on the bottom right did not have any of its demand allocated. The nearest facility to that demand point was not chosen to be part of the solution because the Facilities To Choose property was set to one.

- The total market share, which can be used to calculate the captured market share, is the sum of the weight of all demand points located on the network; unlocated demand points do not contribute to the total market share and should be relocated on the network if they are to be counted.

Target Market Share

Target Market Share chooses the minimum number of facilities necessary to capture a specific percentage of the total market share in the presence of competitors. The total market share is the sum of all demand weight for valid demand points. You set the percent of the market share you want to reach and let the solver choose the fewest number of facilities necessary to meet that threshold.



Target Market Share works in the presence of competitors and tries to choose the fewest facilities necessary to capture the market share that you specify.

The market share problem types require the most data because, along with knowing your own facilities' weight, you also need to know that of your competitors' facilities. The same types of facilities that use the Maximize Attendance problem type can also use market share problem types given that they have comprehensive information that includes competitor data.

Large discount stores typically use the Target Market Share problem type when they want to know how much expansion would be required to reach a certain level of the market share or see what strategy would be needed just to maintain their current market share given the introduction of new competing facilities. The results often represent what stores would like to do if budgets weren't a concern. In other cases where budget is a concern, stores revert to the Maximize Market Share problem and simply capture as much of the market share as possible with a limited number of facilities.

The following list describes how the target market share problem handles demand:

- The total market share, which is used in calculating the captured market share, is the sum of the weight of all demand points located on the network; unlocated demand points do not contribute to the total market share and should be relocated on the network if they are to be counted.

- Demand outside the impedance cutoff of all facilities is not allocated to any facility.
- A demand point inside the impedance cutoff of one facility has all its demand weight allocated to that facility.
- A demand point inside the impedance cutoff of two or more facilities has all its demand weight allocated to the facilities that cover it; furthermore, the weight is split among the facilities proportionally to the facilities' attractiveness (facility weight) and inversely proportional to the distance between the facility and demand point. Given equal facility weights, this means more demand weight is assigned to near facilities than far facilities. This behavior is demonstrated in the Target Market Share graphic above. Assume the three facilities (squares) have equivalent weights and note that two of the six demand points (circles) are within the impedance cutoffs of two different facilities and have their demand split between the facilities. (The amount of yellow or blue in the demand point correlates to the percent of demand weight captured by the facility of the same color.) The demand point near the center of the graphic is covered by both the yellow facility on the left and the blue facility in the center. Since the demand point is closer to the yellow facility, more of the demand is allocated to that facility.

Another demand point has its weight evenly split between the blue facility and the yellow facility on the right because it is equidistant to both facilities.