Introduction to GIS: Lab 4

Select by Location & Vector Data Analysis[[1]](#footnote-1)

This week, we’ll begin editing and manipulating data using our spatial tools and capabilities. Thinking back to our selection steps from Lab Two, and considering the vector data concepts discussed in lecture, today we will make spatial decisions to solve a theoretical problem concerning the West Nile Virus.

First, we’ll practice a few components from last week and go through some new data manipulation skills you’ll use throughout this lab and the course.

While the lab production activity is what you will turn in to Moodle for a grade, please note that practice questions throughout to test your own understanding of the concepts are usefully noted in **bold**. These questions are very useful for testing your capacities, and will be helpful to complete the lab production as well!

This lab is divided into seven parts, which you should complete in order:

**Part A:** Joins and SQL Practice

**Objective:** Practice more with joining a table to a spatial layer and writing query statements to explore the data.

**Part B:** Select by Location

**Objective:** Learn another selection tool by selecting based on spatial location.

**Part C:** Exporting a New Shapefile

**Objective:** Create new shapefiles from data you’ve selected.

**Part D:** Analysis Tools: Clip

 **Objective:** Use the clip tool and begin your first geospatial analysis tasks.

**Part E:** Analysis Tools: Buffer

**Objective:** Explore another analysis tool, the Buffer, to extend your analytical capabilities.

**Part F:** Other Analysis Tools

**Objective:** See the remainder of the simple analysis toolbox, while taking a brief dive into the Erase tool.

**Part G:** Lab Production: To Spray or Not to Spray: Solving a Public Health Problem with Spatial Analysis

**Objective:** Use your analysis skills to determine where to spray mosquitos for West Nile Virus after a dead bird has been found in a wetlands area.

# Part A: Joins and SQL Practice

*Summary: Reflect on some core skills from lab two. Please note there is no video walkthrough for this section.*

[*Geospatial Technology Competency Model*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 5.3.21.4, 5.3.20.1, 5.3.13, 5.2.21, 3.3.2, 2.8.6, 2.8.1, 2.3.2, 2.3.1, 2.1*

1. First, open the \hazmat folder and add all the shapefiles:
* Hazmat
* Schools
* Towns
* Counties
* Blackstone\_watershed
* Blackstone\_streams

These spatial layers surround the Blackstone River watershed, which provides drinking water to Rhode Island.

1. The hazmat folder also contains an excel spreadsheet with town population in 2000.

**Practice:** Add the population table to your Project and write a ‘select by’ query to answer the following question:

How many towns had a population of greater than 10,000 and an area of less than 50 km2?[[2]](#footnote-2)

1. Clear your selection and we’ll move on to the next part of the lab.

# Part B: Select by Location

*Summary: Explore another mode of selecting data using the spatial nature of data and layers.*

[*GTCM*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 5.3.21.4, 5.3.20.1, 5.3.13, 5.2.21, 3.3.2, 2.8.6, 2.8.1, 2.3.2, 2.3.1, 2.1*

[Link to Video Walkthrough](https://youtu.be/MV6-VGPfOpo)

1. In addition to Select by Attributes, which you recall from last week, another option in the Selection Group is Select by Location. This tool allows you to select features from one shapefile that are spatially related to another feature[[3]](#footnote-3).
2. For example, let’s say we want a count of all the Massachusetts towns that intersect the Blackstone River Watershed. Go to the Map tab > Selection Group and choose Select by Location.
3. In the Select by Location geoprocessing pane, we’ll choose a layer to apply our selection to, and a layer to do the selecting. The layer that will do the selecting is called a “Selecting Feature,” and the layer that the selection is applied to is called the “Input Feature.”

1. Since our goal is to select a certain number of towns, we’ll choose ‘towns’ as our Input Feature. The selection will be applied to this layer.
2. Next, we’ll choose our Selecting Feature. The layer in the Input Features parameter (towns) will be selected based on its relationship to the feature we choose as our Selecting Feature. For this example, we want to choose ‘Blackstone\_watershed.’
3. Notice also that you have several options under Relationship. This is where you choose the method by which the spatial relationship will be evaluated between the Input Feature and the Selecting Feature. The default is Intersect, but you can also select features within a specified distance of other features or select features that only fall completely within other features, for example. Keep the Relationship as Intersect for this example.
4. Finally, notice that you have several options under Selection type ­– similar to Select by Attributes, you can create a new selection or add/remove from an existing selection. Leave the Selection type as New Selection and click Apply.



1. How many Massachusetts towns intersect the Blackstone watershed?[[4]](#footnote-4)
2. You can think of the Select by Location tool as a way to construct a sentence that will answer you query. In our example, we are asking, “How many towns (Input Feature) intersect (Relationship) the Blackstone watershed (Selecting Feature)?”

**Practice:** Using all of your selection skills – answer the following question: How many public schools (Type = ‘PUB’) are in the Blackstone watershed?

**Hint 1:** Think about what selections you need to do (there are two separate ones). What are the processes to do them?[[5]](#footnote-5)

**Hint 2:** Check out the options in the Selection type dropdown. Which is the most appropriate yere?

**Practice:** How many hazmat sites are located within 100 m of one of the Blackstone streams?[[6]](#footnote-6)

# Part C: Exporting a New Shapefile

*Summary: Explore how to manage your data by creating a new shapefile from a selection.*

[*GTCM*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 2.8.1, 2.4.2.1*

[Link to Video Walkthrough](https://youtu.be/IFgX0zSD8HE)

1. Let’s say we’re only interested in towns that intersect the Blackstone watershed. Why bother to keep all these extra towns? They’re just wasting space!
2. Use ‘Select by Location’ to select all the towns that intersect the Blackstone watershed (if they aren’t already selected).
3. Right click on towns and go to Data > Export Feature. A Geoprocessing pane appears. For Input Features, choose towns. Select where you want to save your new shapefile under Output Location and name your shapefile under Output Feature Class.[[7]](#footnote-7)



1. **When you have data selected, ArcGIS Pro will default to perform the chosen operation on ONLY the selected data**. So, export will default to exporting the selected features. The same will happen with analytical tools – so watch out for what you have selected!
2. Click Run and voila – a new shapefile is created and added to your map that includes only the towns in the Blackstone watershed.[[8]](#footnote-8)

**Practice:** What is the area of all the towns that intersect the Blackstone watershed?[[9]](#footnote-9)[[10]](#footnote-10)

# Part D: Analysis Tools: Clip

*Summary: Learn your first geoprocessing tool, Clip, which allows you to create a new file of only those features within another area of interest.*

[*GTCM*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 5.2.19.1, 2.4.2.1*

[Link to Video Walkthrough](https://youtu.be/LGdnq1DUCXI)

1. You just calculated the area of all the towns that intersect the Blackstone watershed. But, what if we don’t want all the area that is outside of the watershed to be included? Take, for example, Shrewsbury. It is about half in the watershed and half out. How can we calculate the area of Shrewsbury that is inside the Blackstone watershed?



|  |
| --- |
| Note on finding toolsThe easiest way to find tools is via the Tools search bar: go to the Analysis tab and click on the red toolbox. This brings up a geoprocessing pane with a search bar (“Find Tools”). In order to search for tools in this search bar, you must either know what the tool is called or what operations it completes. If you don’t know the name of the tool, perhaps try using Google rather than ArcGIS help to try to find it. Google has more flexibility in search terms – just make sure you include ArcGIS Pro[[11]](#footnote-11) in addition to whatever you’re searching for. Or, better yet, ask one of us! The more you use ArcGIS, the more familiar you’ll become with the different tool names. But, for now, consider keeping a notebook with some tool names that you can refer back to later. |

1. One tool we can use to cut off features based on another feature is the ‘clip’ tool. Clip functions like a cookie cutter and takes from an input layer only the data that fits the clip layer (see below).



1. Open the Tools geoprocessing pane (Analysis tab > tools). Find the Clip tool either by searching or by going to the Toolboxes tab > Analysis Tools > Extract > Clip.



1. Selecting the Clip tool brings up a new geoprocessing pane. The ‘input features’ at the top is the layer you want to clip, the ‘clip features’ below it is the shape you want to use (‘clip features’ is like the cookie cutter). Specify a new[[12]](#footnote-12) output shapefile and save it where you want it to go.



1. Click Run to execute the command. You’ll know it’s working if you see a blue progress bar at the bottom of the geoprocessing pane.[[13]](#footnote-13)



1. You’ll know it’s finished when a green checkmark appears.[[14]](#footnote-14) A details window will also pop up briefly, showing time elapsed, messages, and any errors or warnings.

 

1. Find Shrewsbury and see what it looks like now – it should only include the portion of the town that is inside the watershed.

**Practice:** What is the area of Shrewsbury inside of the Blackstone watershed?[[15]](#footnote-15)

|  |
| --- |
| Warning: Calculated Area does not automatically update!If you found the answers to the reminder question above (good job!), then you probably opened your attribute table to find the area of Shrewsbury and saw the same value in the Area column that you calculated earlier. And, while the shape area column has updated, it’s still in some mystery unit that’s hard to understand[[16]](#footnote-16). Any time you clip a shapefile (or perform any other analysis tools), calculations you’ve made in new columns like area or perimeter (for polygons) and length (for lines) will not automatically update. You will need to redo Calculate Geometry to find the area of Shrewsbury inside the Blackstone watershed. |

# Part E: Analysis Tools: Buffer

*Summary: Perform more spatial analysis with the Buffer tool, allowing the creation of areas around items of interest.*

[*GTCM*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 5.2.19.1, 5.2.1, 2.4.2.1*

[Link to Video Walkthrough](https://youtu.be/kP0JCnBUmss)

1. On to a new problem. Now, let’s say that we want to find the area of land within 100m of a Blackstone stream in the town of Shrewsbury.
2. Blackstone streams is a line shapefile. These lines have no width or size, so we’ll need to create a new polygon to tell us the information we want to know about proximity to the streams. We can use the ‘Buffer’ tool to create a file showing the area ‘around’ a point, line, or polygon feature. Neat!
3. Open the buffer tool by finding it in Toolboxes or using search.



1. Input features should be blackstone\_streams[[17]](#footnote-17). Save a new output shapefile somewhere in your Lab folder. In this case, we want the buffer to be 100m. You should also change the dissolve type to ‘Dissolve all output features into a single feature.’
	1. This means that the output will be one single buffer feature surrounding all the streams.
	2. If you use the default dissolve type (No Dissolve), you’ll get a separate buffer around every individual stream feature.
	3. Buffers around each individual stream will include some overlapping areas, so the area calculation would be incorrect.
2. Click Run to execute the buffer.
3. Now, clip your new stream buffer with the town of Shrewsbury.
	1. Remember how to Export Data? Remember how to clip? Combine!
	2. Select just the town of Shrewsbury in your watershed-only clipped towns file.
	3. Export that Shrewsbury as a new layer and use it as the clip feature. Data management score!

**Practice:** How much land area falls within 100 m of a Blackstone watershed stream in Shrewsbury?[[18]](#footnote-18)

# Part F: Other Analysis Tools

*Summary: Learn about some additional analysis tools and practice one – Erase! Erase will be very useful for your future lab work. Remember, erase will be helpful for the production activity.*

[*GTCM*](https://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx)*: 5.2.19.1, 2.4.2.1*

[Link to Video Walkthrough](https://youtu.be/X9bzsRA5pZc)

1. There are lots of other analysis tools that you can explore in Toolboxes. Below are a few examples. Chances are, if you can think of it, there’s probably a tool to do it…it’s just a matter of finding it!
2. Hovering over a tool’s name in the Toolbox will give you a brief description of what the tool does. You can read more about what tools do by clicking the blue question mark in the upper right-hand corner once you’ve opened the tool.



1. Example analysis tools[[19]](#footnote-19):

    

1. The ‘Union’ tool combines features into one file. If those features overlap, it retains their separate geometries, forming new shapes. Useful for when you have a lot of data of the same type in different files you need to combine, but want to see the separate components that make up those data.
2. Erase works like a big spatial subtraction tool. Taking some input, you can erase any feature from that input. Good when trying to remove features based on some exclusion principle.
3. Split carves up your data based on some pre-existing feature dataset. Using split can help manage your data into more manageable pieces, or separate data across some characteristic to run separate analyses.
4. Dissolve removes the ‘lines’ or the separations between components of your data. If you don’t need separate shapes in your layer file, dissolve allows you to unify components based on some commonality.
5. Let’s try one for fun. We’ll use Erase because it has some useful applications for the lab production activity.[[20]](#footnote-20)
6. Let’s use our last two layers as an erase example. If we wanted to remove the river buffer from our clipped Shrewsbury file completely, we could use erase to ‘delete’ that portion of our file. Open up the Erase tool and you’ll see how it works.

[[21]](#footnote-21)

1. In our case, we have an input file of our clipped Shrewsbury town. We want to erase, or remove, the 100m river buffer so it doesn’t even appear in our town file.
2. The output will be a new file showing the area of Shrewsbury in the watershed (our first clip) with the area within 100m of the river removed (our erase). Try it out!

[[22]](#footnote-22)

1. Learning how and when to clip, erase, buffer, intersect, union, and so forth is a key spatial skill. We’ll practice some of these in lab today, and all of them moving forward throughout the semester.

# Part G, Lab Production: To Spray or Not to Spray

# Solving a Public Health Problem with Spatial Analysis

[Link to Video Walkthrough](https://youtu.be/d6pAJmB1H6E)

**Problem Statement**: The citizens and government in Middlesex County, Massachusetts are concerned by the recent discovery of a dead bird carrying West Nile Virus. The bird is believed to have lived in the local habitat, and the Centers for Disease Control (CDC) is considering whether to spray the surrounding area with a pesticide aimed at eliminating mosquitoes that may have contracted the disease. A representative from the CDC has arrived at your GIS lab and is asking you to help her evaluate whether or not the CDC’s current criteria for spraying a pesticide to eliminate the mosquitoes that carry the virus are likely to be effective. She has provided the spraying criteria below.

**The Spraying Criteria**: According to the CDC’s current criteria, spraying a pesticide by trucks must occur within a 2-kilometer radius around the site of any animal that has been found dead of the virus. However, it is known that pesticides sprayed by truck will only reach to 50 m on either side of roads. Further, the Environmental Protection Agency (EPA) has regulated that spraying must not occur within 100 m of any wetland because they don’t want to harm wetland biota.

**Data:**

To undertake the evaluation, you are provided with three layers in the Lab3 zip. The layers are:

1. *deadbird* containing the location of a bird that has been found to have died of the virus
2. *wetlands* representing wetlands as polygons
3. *roads* representing roads as lines

**Questions to be turned in:**

Given this spatial data and the spraying criteria, answer the following questions:

1. What is the ideal spray area per the CDC guidelines (in square km) if there were no logistical or regulatory constraints?
2. How much land area (in km2) is available for pesticide spraying given the above wetland and road distance limitations?
3. What percentage of the recommended total spraying radius does this available area represent?
4. In your opinion, how effective is this treatment likely to be? Explain your reasoning.
5. Using a flow chart or an enumerated set of clearly written instructions, explain the steps you used to complete this analysis.

**Map to be turned in:**

1. Create an 8 ½ x 11 layout/poster (may be portrait or landscape orientation) showing the potential and realized spray areas, along with the location of the dead bird, roads, wetlands, and any other spatial data that would help illustrate your results. *Make sure your map has a legend, scale bar (in km), and north arrow as well as a title.* When you are finished, export the map as an image and insert it in a word document with your answers to the above questions. Submit your report as a .pdf on Moodle.
1. University of Massachusetts – Amherst, ArcGIS Pro Edition

Written by Bethany Bradley, Forrest J. Bowlick, Sophie Argetsinger, Steven Bittner, Brit Laginhas, Chloe Thompson, Connor Hughes, and many others [↑](#footnote-ref-1)
2. Should be 11. [↑](#footnote-ref-2)
3. Remember when you were in elementary (or equivalent) school, and you learned words like ‘around’ ‘on’ ‘near’, and so forth? Think like that with spatial decisions. [↑](#footnote-ref-3)
4. 25 [↑](#footnote-ref-4)
5. (Answer should be 96 – make sure you know how to get that info) [↑](#footnote-ref-5)
6. (Answer should be 33) [↑](#footnote-ref-6)
7. Accept no defaults! You are in control of the data – you control where that data goes. [↑](#footnote-ref-7)
8. Memorize this task order for future efficiency. [↑](#footnote-ref-8)
9. (should be 1382 km2) [↑](#footnote-ref-9)
10. Yes, you do already have an area field. Do you know what unit it is in? Could you find out easily? Probably not – so make a new field! Calculate Geometry! You know what to do. [↑](#footnote-ref-10)
11. You may see lots of results for ArcMap, ArcGIS PRO’s predecessor. [↑](#footnote-ref-11)
12. Clip and other analysis tools do not change your original data. Instead they make a new file based on the tool you use. Neat! [↑](#footnote-ref-12)
13. Get ready to stare at this bar longingly for the rest of the semester. [↑](#footnote-ref-13)
14. Other options include it not working. [↑](#footnote-ref-14)
15. (It should be 35.5 km2) [↑](#footnote-ref-15)
16. Square milimeters maybe? How is that helpful, Arc? [↑](#footnote-ref-16)
17. The file you want to buffer. [↑](#footnote-ref-17)
18. (Should be 14.4 km2) [↑](#footnote-ref-18)
19. Big screenshot next page [↑](#footnote-ref-19)
20. HINT HINT HINT [↑](#footnote-ref-20)
21. These great diagrams from the Esri help page. [↑](#footnote-ref-21)
22. Your resulting file should have holes! [↑](#footnote-ref-22)