

Tutorial 7 – Vector Data Analysis

The objectives of this tutorial include:

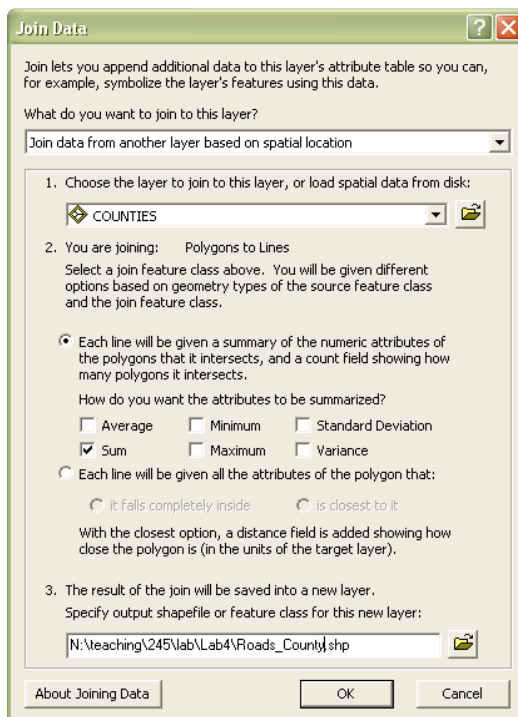
1. Selecting features by location (spatial joins and location queries)
2. Vector data analysis (buffering and unioning)
3. Feature manipulation for vector analysis (clipping and erasing)

Before beginning the tutorial, please COPY the lab7 zip archive to your server folder and unpack it.

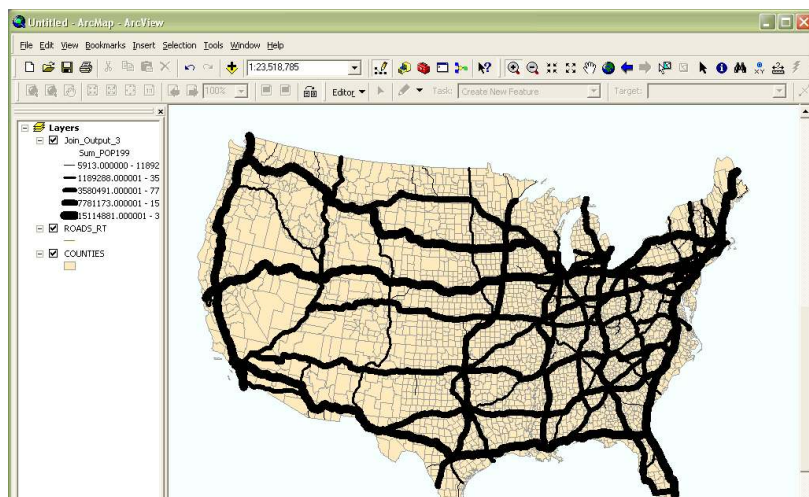
1. Spatial joins: Learning how to connect polygons to lines

Often times it is useful to join attribute information from a polygon map to a line map. We performed a similar operation in Lab 6 when we joined polygons to points for the Colgate land cover assessment exercise. In this case, let's say we want to determine the total number of people that live in counties that border each of the Interstate Highways in the United States.

- Launch ArcMap and open a new empty map. Add Counties and Roads_RT layers from the C:\ESRI\ESRIDATA\USA folder. Place the roads on top of the counties to see both.
- Make sure you change the projection of your data frame to a USA Contiguous Albers Equal Area Conic Projection. Do not pick the North American Albers Equal Area Conic projection. What's the difference between the USA Contiguous Conic USGS projection vs the USA Contiguous Conic projection?



- Right click on the Roads_RT layer and select 'Joins and Relates' → 'Join'. In the window that appears you select "Join data from another layer based on spatial location". In this dialogue box you can select the kind of spatial join you want to perform. Since we are interested in knowing the total number of people in each county that each road intersects you should select the first join option and 'sum'.
- Provide a name and location of the new shapefile that this tool will create. Once your window looks similar to the one above click 'ok' to run the command.
- Once the analysis is complete ArcGIS adds the new shapefile to your map. Open the new shapefile's attribute table to determine which interstate serves the greatest number of people based on the 1999 values. (Hint: sort descending on the 'Sum_Pop199' attribute).
- If you symbolize based on this new attribute you could very easily make a map like the one below (I used graduated symbols; see below).

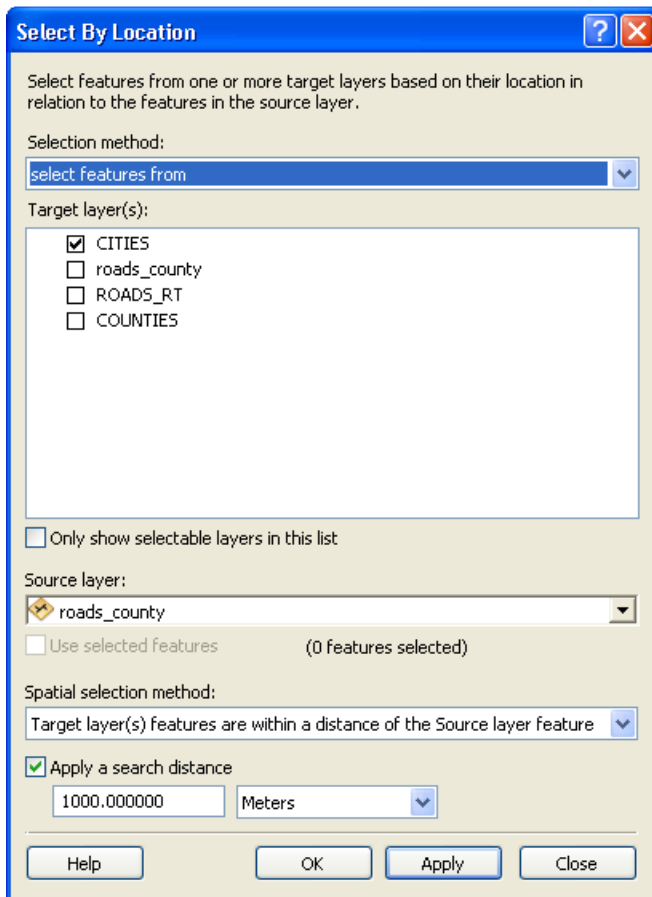


Note: some attribute names end up getting truncated during analysis because by definition they can only be 10 characters long and 'Sum_' was added to all attributes. If you have trouble figuring out the variables go back and open the original attribute table.

2. Select features by location

In previous exercises we have selected individual features according to attribute data values (e.g., selecting all the counties in the state of NY from the map of US counties). You can also select a subset of features based on their location. For example, let's imagine we are long-haul truckers and while we could live anywhere in the US, we'd prefer to live in cities within 1 kilometer of Interstate Highways. How many cities are available for us to choose from?

- Add the cities shapefile from the C:\ESRI\ESRIDATA\USA folder to your existing map of counties and roads. By default cities is placed on top of the other three maps.
- If you started a new ArcMap map document make sure you once again change the projection of your data frame to United States Albers Equal Area Projection.
- Under the selection pull-down menu choose 'Select by location'. The dialogue box that pops up enables you to perform a variety of different spatial queries (be sure to browse the options available in the 'spatial selection method' pulldown menu.
- At the bottom you can specify that we are interested in those cities that are within 1 km of the roads. Once you select 'apply' your selected cities should appear highlighted on your map. Don't select 'ok' just yet.

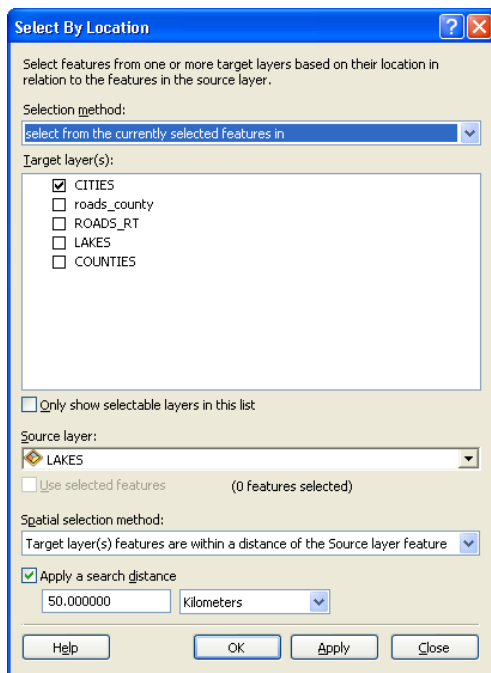


- You should now be able to open the attribute table of the Cities shapefile and view the selected cities. If you have performed the operation correctly there should be 316 cities selected from 3,149 possible.

Note: if you want to create a shapefile on only those cities within 1 km of an interstate you can right click on the cities shapefile and select 'Data→ Export data' and then choose to only export the selected features.

Let's further imagine that after visiting these cities you decide that you want to live within 50 km of a lake. How many of the 316 cities are within 50 km of a lake?

- Add the lakes shapefile from the C:\ESRI\ESRIDATA\USA folder to your existing map of counties and roads.
- In the "Selection Method" pulldown menu at the top of the 'Select By Location' dialogue box indicate that you want to 'select from the currently selected features in' option. This will enable you to select from the 316 cities that are within 1 km of an interstate.
- This time you 'want to select from the currently selected features in cities that are within a distance of the features in lakes.' Once again select the option to buffer, but this time select lakes and 50 km. Your window should look like the one below.



- You should find that 42 of the cities are both within 1 km of an interstate and 50 km of a lake. If you do not get this answer ask your lab instructor for help.

3. Analyzing vector data: buffering

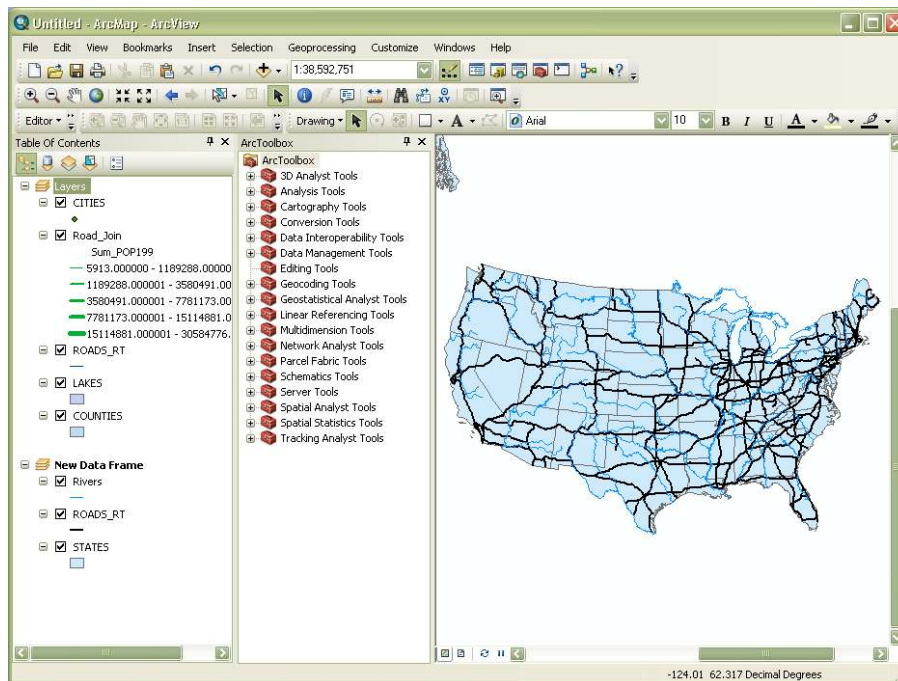
Each of the routines described above are ultimately designed to select a unique subset of objects from different map layers. However, suppose we want to analyze spatial data layers in a way that creates entirely new objects. For example, rather than wanting to know how many cities are within 1 km of an interstate (selecting a unique subset of cities) we want to know what land areas in the US are within 1 km of an interstate. These land areas (i.e., buffers) do not exist yet. They need to be created from existing data. These kinds of questions require a different set of tools.

Do you understand the difference between the two types of vector analyses described above? This is a critical distinction in the context of the exercise.

(((Parts 3 and 4.... ArcToolbox project (reproject) of roads_rt and rivers to USA Contiguous Albers Equal Area helps solve problem of inability to select out the rivers buffer portion, but results in 98 selected unioned areas, not 99 -) Might be useful to remind difference between projection of data frame and projection of data layers. Also, distance is 25000m and 50000m, so if changed, buff_dist must reflect that.) Or else fiddling with Environments...Output Coordinates, but haven't tried.)))

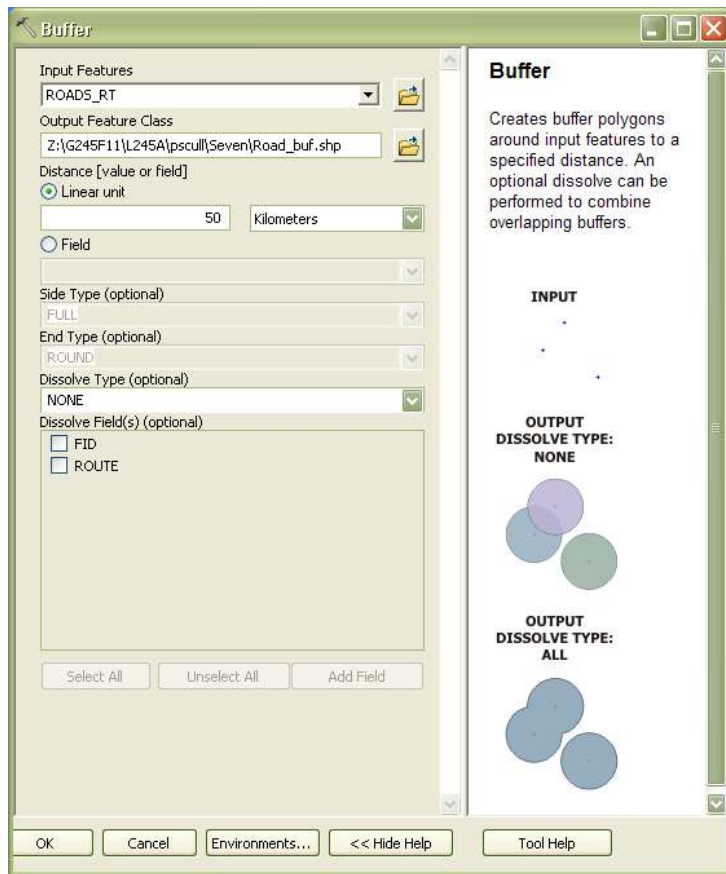
Let's imagine that we want to determine all those areas in the US that are not within 50 kilometers of an interstate, but are within 25 km of a river. We can do this by "buffering" the interstates at 50 km and overlaying the area outside the buffer with the area inside a 25 km buffer of streams.

- Add a new dataframe to your map and add the Rivers, Roads_RT, and States layers from the ESRI map files.
- Before proceeding any further make sure you change the projection of your dataframe to United States Albers Equal Area Projection.
- Add the ArcToolbox to your map and open the buffering tool (located within Analysis Tools → Proximity; see below).

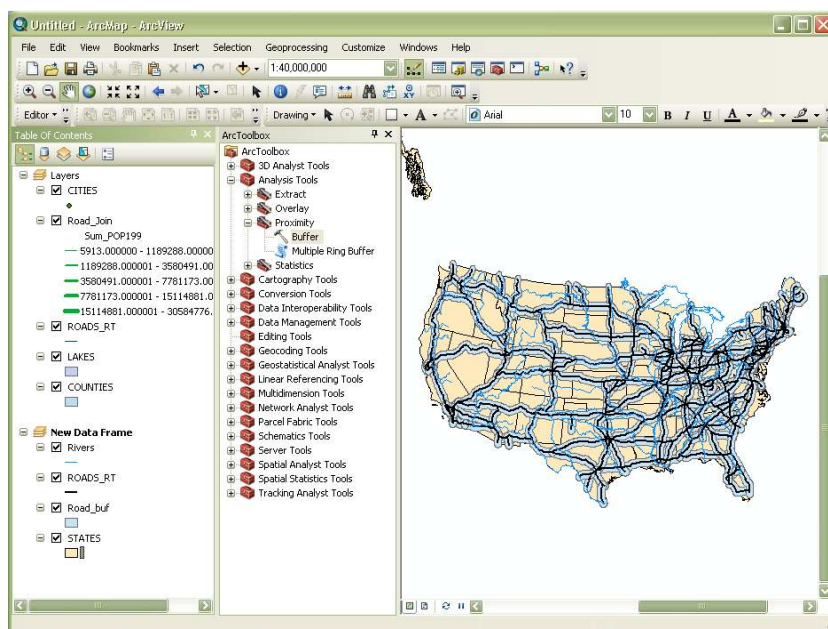


- Double left click on the buffer tool. Specify 'Roads_RT' as your input feature and define an appropriate output file name and path.
- Next, select 'Linear Unit' as 50 km. Here we could buffer each road differently based on information in their attribute table (e.g. size), but we will treat all roads as the same in our analysis.
- Choose the remainder of the defaults and click ok once your window looks like the one below. Can you see the difference between the two different dissolve options?

GEOG 245: Geographic Information Systems
Lab 07

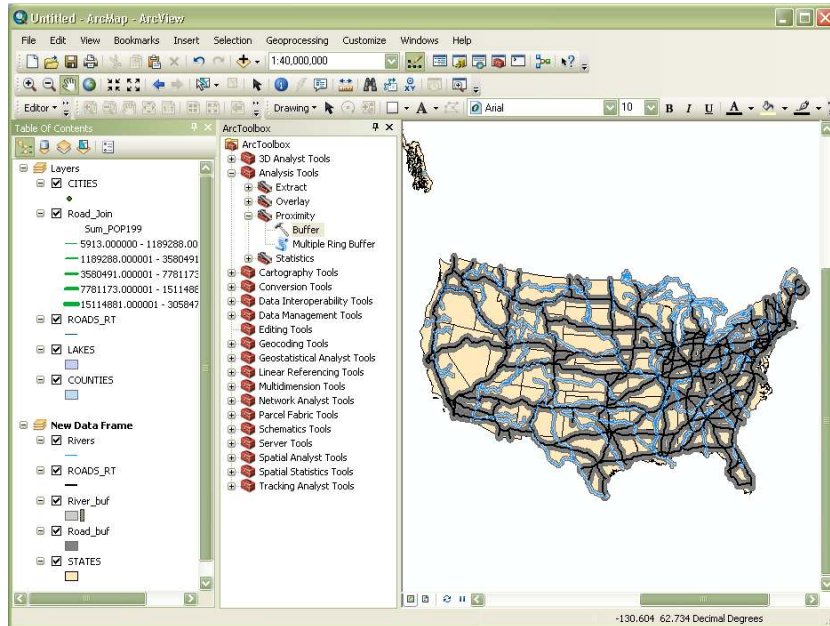


- By default the new layer is added to your map. It should look something like the one below.



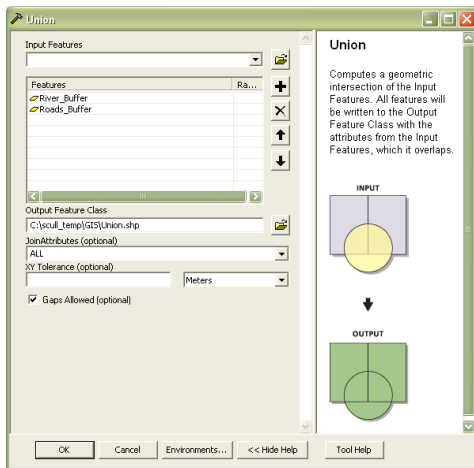
GEOG 245: Geographic Information Systems
Lab 07

- Open up the attribute table for your new layer. You should see a new attribute called 'Buff_Dist'. This attribute keeps track of the fact that you buffered the roads at 50 km, but the distance is tabulated in the original units of the data which is in degrees. (Why doesn't the attribute table have the distance in km?) Since 50 km is approximately equal to 0.449661 degrees on the Earth sphere, this is the number present in the table.
- Next, buffer the rivers layer at 25 km using a similar routine. Your river buffer should end up looking like the one below. The BUFF_DIST should be 0.224831 degrees in the attribute table.

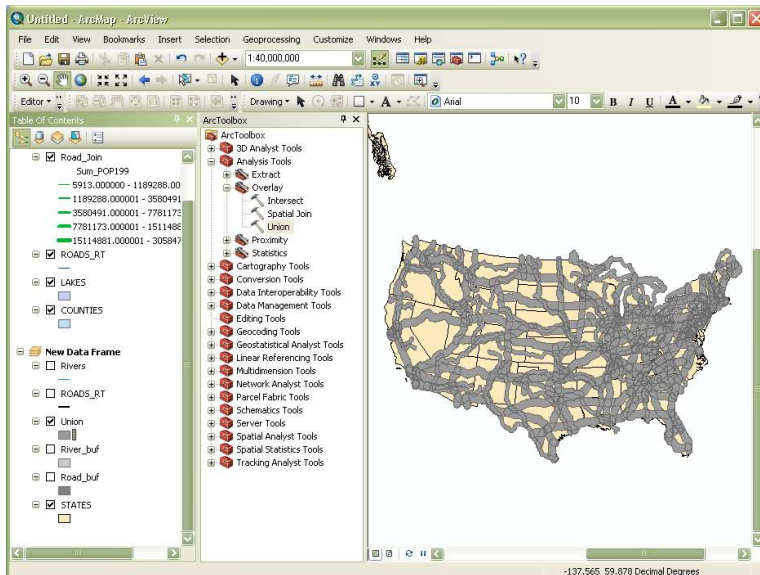


4. Analyzing vector data: Unioning

- Next we need to overlay the two buffers to identify the area outside the roads buffer and inside the river buffer. In the ArcToolbox open the union tool (Analysis tools → Overlay → Union).
- The input features are the layers you want to union. You can also see in the right panel of the dialogue box a schematic that illustrates what a union actually does. You should enter both of the buffer maps, define an appropriate filename, and make sure you are combining 'all' attributes (see below). Click ok.



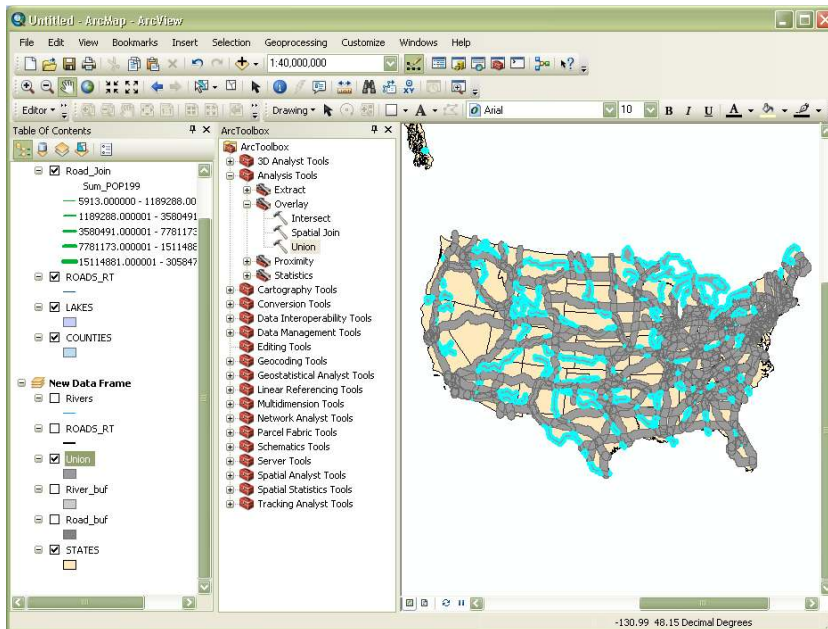
- After the analysis is complete the new layer ('Union' in my example) is added to your map. It should look something like the one below.



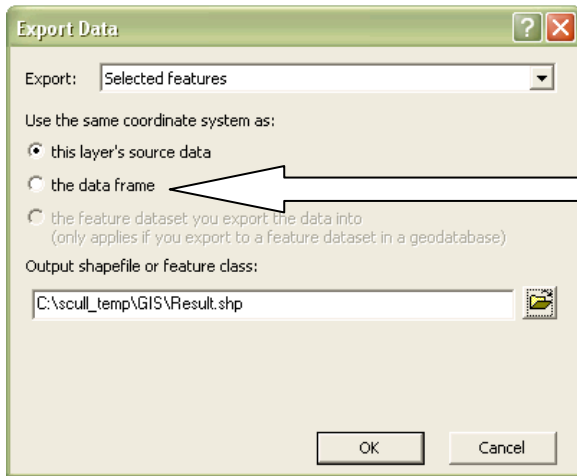
- Next, we need to select those polygons outside the roads buffer and inside the rivers buffer. Open up the attribute table for the union shapefile (see below). The 'BUFF_DIST' is the attribute with our river information and the 'BUFF_DIS_1' is the attribute with our Roads information. Even though the original two attributes had the same name (Buff_dist) we can tell them apart based on their proximity to other attributes in the table.

FID	Shape	FID_River	NAME	SYSTEM	BUFF_DIST	FID_Road_b	ROUTE	BUFF_DIS_1
0	Polygon	0	Alabama		0.224831	-1		0
1	Polygon	1	Arkansas	Mississippi	0.224831	-1		0
2	Polygon	2	Bear		0.224831	-1		0
3	Polygon	3	Brazos		0.224831	-1		0
4	Polygon	4	Canadian	Mississippi	0.224831	-1		0
5	Polygon	5	Chattahoochee		0.224831	-1		0
6	Polygon	6	Clark Fork	Columbia	0.224831	-1		0
7	Polygon	7	Colorado	Colorado	0.224831	-1		0
8	Polygon	8	Columbia	Columbia	0.224831	-1		0
9	Polygon	9	Connecticut		0.224831	-1		0
10	Polygon	11	Flathead	Columbia	0.224831	-1		0
11	Polygon	12	Gila	Colorado	0.224831	-1		0
12	Polygon	13	Green	Colorado	0.224831	-1		0
13	Polygon	14	Illinois	Mississippi	0.224831	-1		0

- Remember, we are interested in those polygons that have a river buffer but are outside the roads buffer. These are the polygons with BUFF_DIST = 0.224831 and BUFF_DIS_1 = 0.
- Select these polygons using the 'Select by Attributes' tool. Do you remember how to do this? Once selected your map should look like the one below (99 records should be selected).

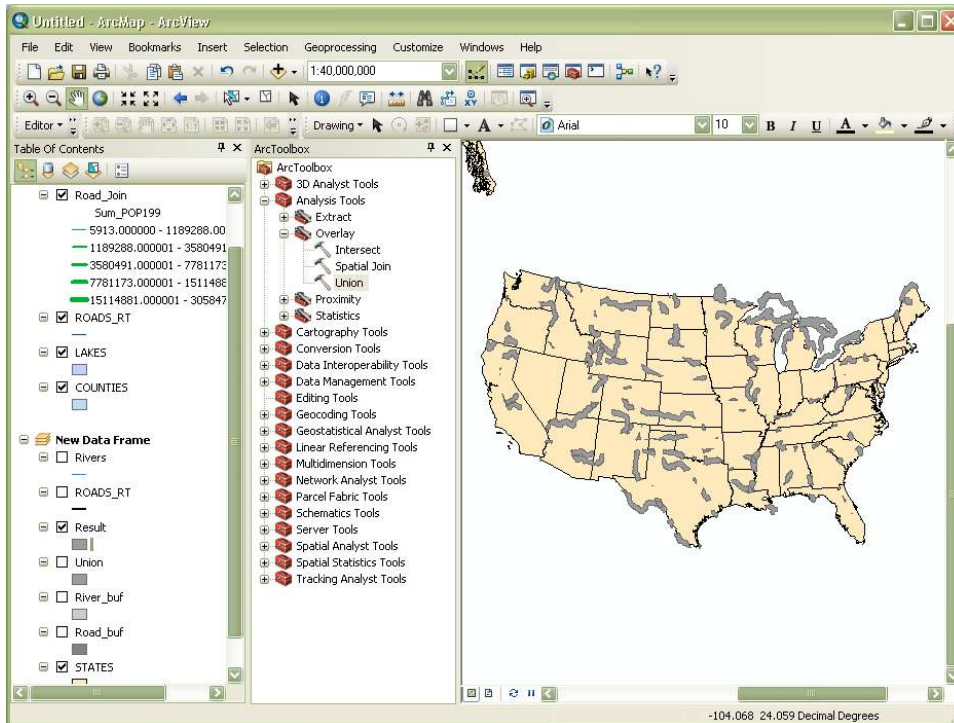


- Save these selected polygons to a new shapefile by right clicking on the layer in the table of contents and selecting 'Data → Export Data' using the Selected features option (see below).



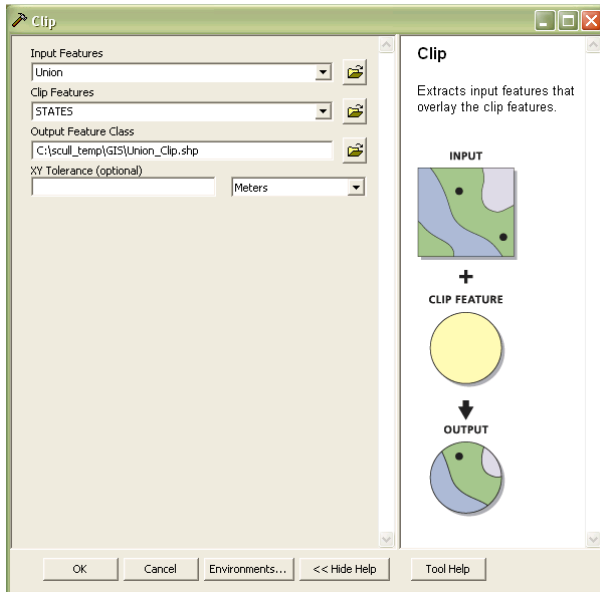
What would selecting this option do to your output shapefile/feature class?

- Your new layer should look like the one below. Notice that some of the polygons are located over the border in Mexico and actually within the great lakes. These are mistakes that resulted from the buffering portion of our analysis. In essence, we have selected areas outside of our study area (US).

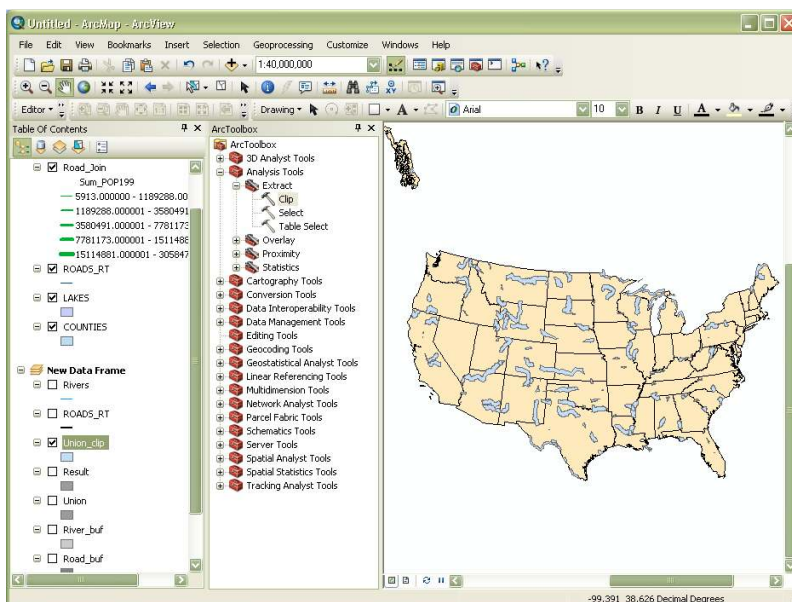


5. Feature manipulation: Clipping between layers

- We can remove these non-US areas by cleaning up our analysis. In this particular context this is called a “clip”. The tool is located in the ArcToolbox (Analysis tools → Extract → Clip). Double left click the tool.
- Your input feature, as shown in the schematic to the right once the tool is open, is your union shapefile. The clip feature is your states shapefile. Define an appropriate output file name, accept the rest of the defaults and click OK (see below).

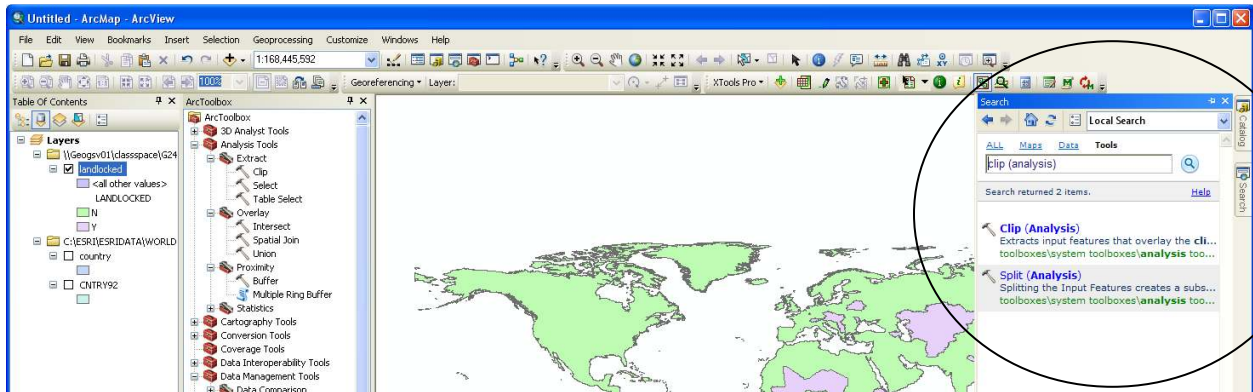


- If all goes as planned your map should look like the figure below, which shows all the areas in the US that are within 25 km of a river and outside of 50 km of an interstate.



6. Tool tip: Locating and using an analytical tool

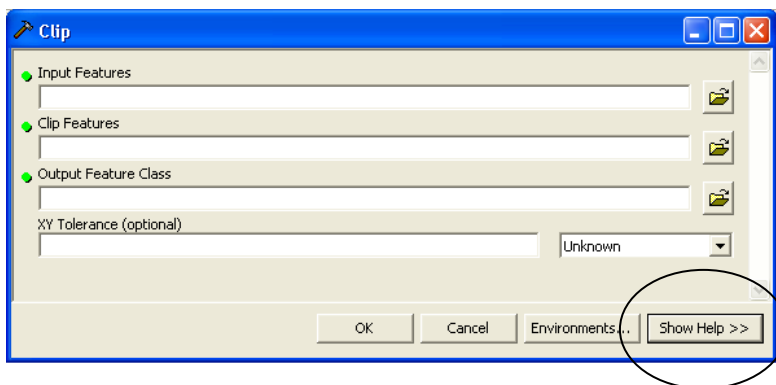
Let's imagine that you want to perform a Clip routine but can't remember where to find the tool in the ArcToolbox and how to use it. Go to the GUI, Windows, Search. You'll find a tab on the upper right hand part of your map document called "Search."



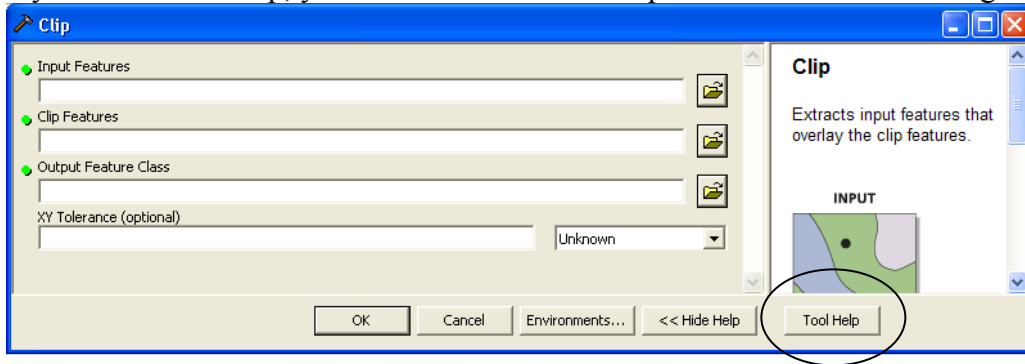
Click on the "Tools" above the empty box. Then start typing the name of your command (or what you can remember of it) in the empty field.

Then you can either click on the tool name (the parentheses indicate where in the ArcToolbox hierarchy you can find it), to start using it right away, or else click the text below to get a full description of the tool.

But how do I use this tool? you ask. Well, as the tool menu comes up, you can click "Show Help" button on the bottom and a general informational panel comes up next to the menu.



If you need more help, you can click on “Tool Help” on the bottom and it’ll give you details.



If you’re not certain what layers to put in what place, click in the space where you’d type in the input name and Help will give you an idea of what goes where.

Extra Tool: Erase

Erase: for when you want that area outside the clip. (An operation that might be defined as the inverse of clip.)

