GIS II**: Data Management: Creation, edition and maintenance of geographic data**
Module 1: Leveraging the where of your geographic data

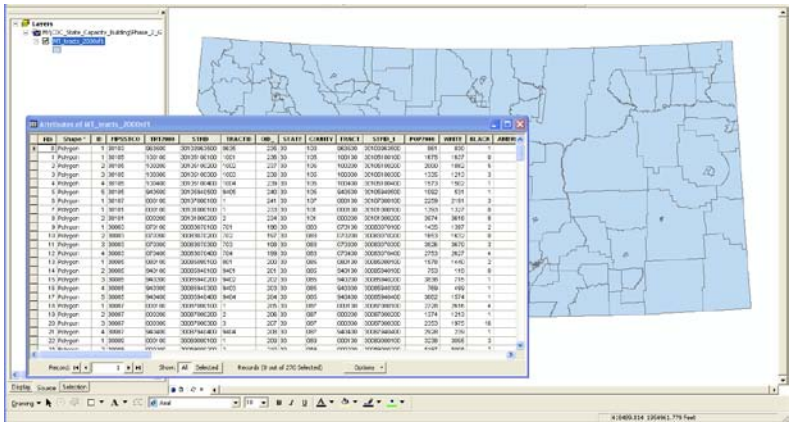
*** Files needed for exercise: *mt_sic4_5812.dbf*, and *MT_tracts_2000sf1.shp*

Goals: After completing this exercise, you will have experience in performing spatial joins and selections and a reviewed your understanding of spatial data projections.

Skills: After completing this exercise, you will be able to spatially join a point dataset (businesses of interest) to a polygon dataset (US Census 2000 Tracts). This is a useful method for the enumeration of points within polygons: i.e. what is the count for points of interest within geography of interest?

Creating Folder Connections:

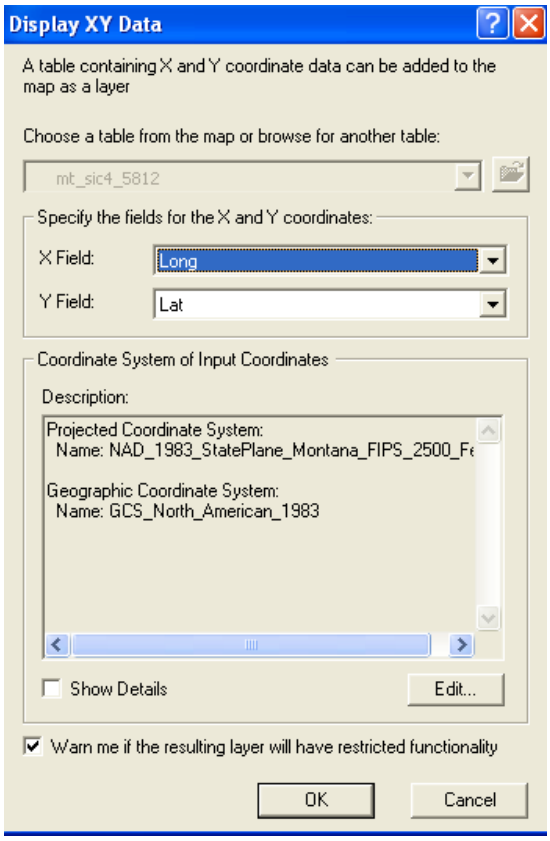
1. Open ArcMap. Choose to start a new empty map.
2. Click the Add Data button .
3. Click on the Connect to Folder button . Connect to the Exercise_01_data folder.
4. Browse to your folder and connect to it. We now have a permanent connection to that folder.
5. Double click on *MT_tracts_2000sft.shp* to add it to your project. This shapefile has Census 2000 summary file 1 demographic data appended to it.
6. Open the table for the shapefile and take a look at the attribute fields.



Add your tabular data and display x/y coordinates:

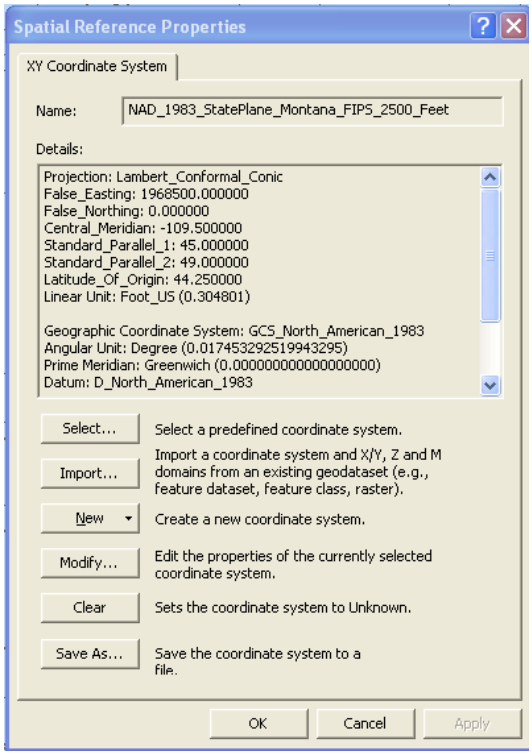
1. Add the file *mt_sic4_5812.dbf* to your project. You should now see the *mt_sic4_5812.dbf* table in the ArcMap table of contents on the left hand side. This table is a list of all businesses classified by 4 digit SIC (standard industry classification) as eating establishments in the state of MT as of 2008.
2. We want to turn this tabular data into a spatial dataset; right-click on the *mt_sic4_5812.dbf* table and choose Display XY Data.

3. On this screen, ArcMap will have already selected LONG for the X Field and LAT for the Y Field. Also note that ArcMap has automatically defaulted to the projected coordinate system information associated with the first shapefile you added to the project. You will need to edit this because your coordinate information is in decimal degrees and will not work in a projected state plane system without first being defined in an angular geographic coordinate system.

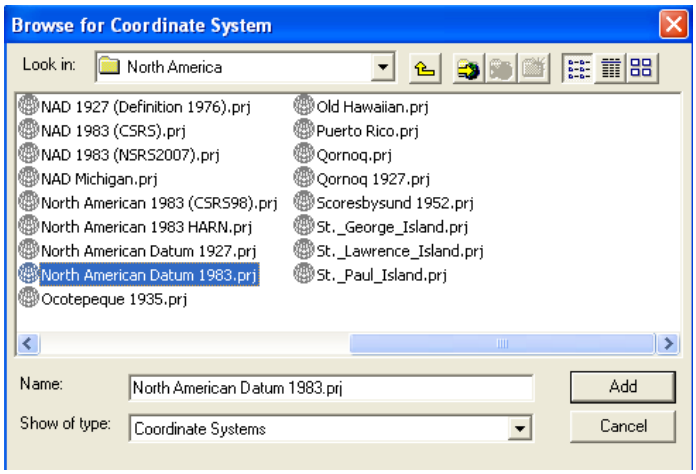


For kicks you can go with this default and see what the result will be by sticking with the default option.

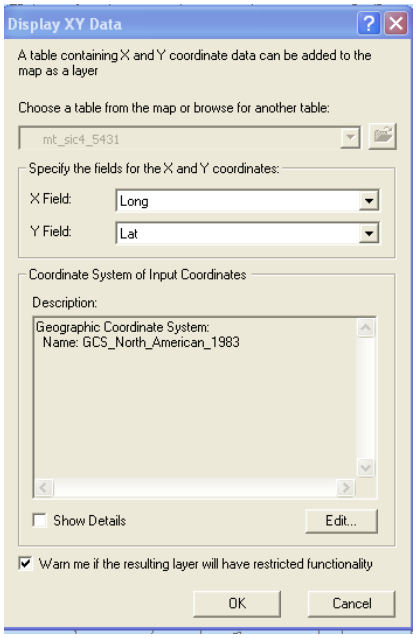
4. Let's edit the coordinate system of input coordinates so they will display properly click Edit and then Select a predefined coordinate system.



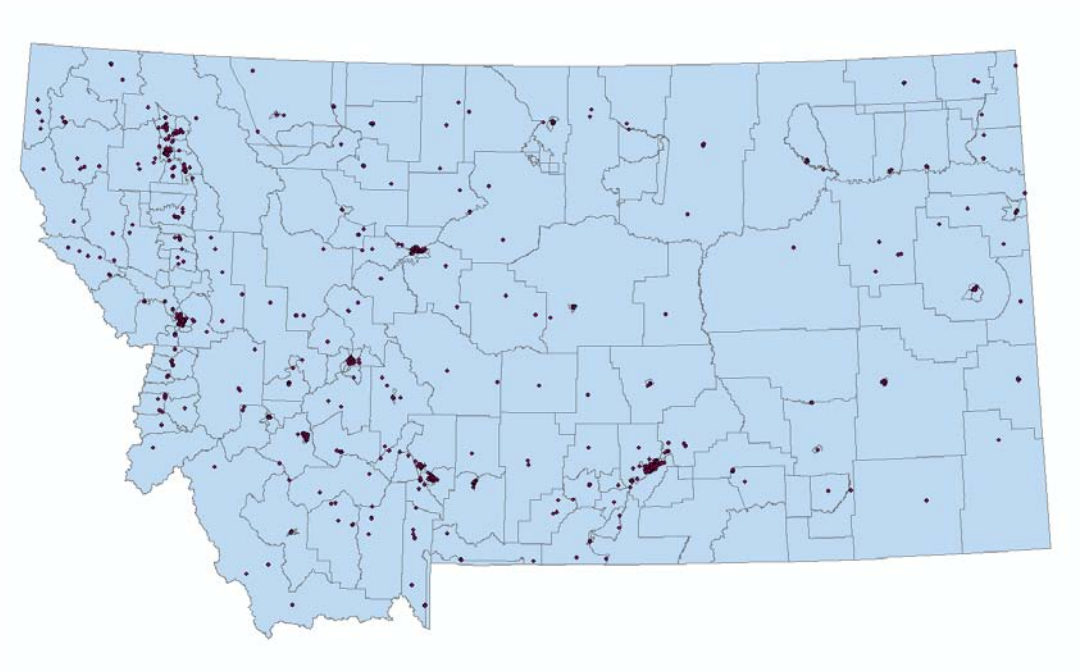
5. Choose Geographic Coordinate Systems > North America > *North American Datum 1983.prj*



6. Click OK. You have now changed the coordinate system for the input coordinates to an angular geographic coordinate system that can handle your decimal degree coordinates.



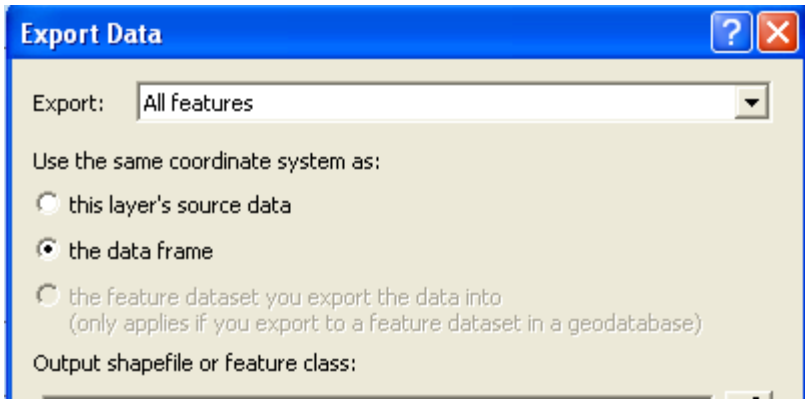
7. The eating establishments should now be displayed above your MT Census 2000 tracts.



Exporting and projecting your events to a new shapefile:

1. You now have a layer in your Table of Contents called *mt_sic4_5812_Events*. This is a temporary layer; we want to make it permanent.

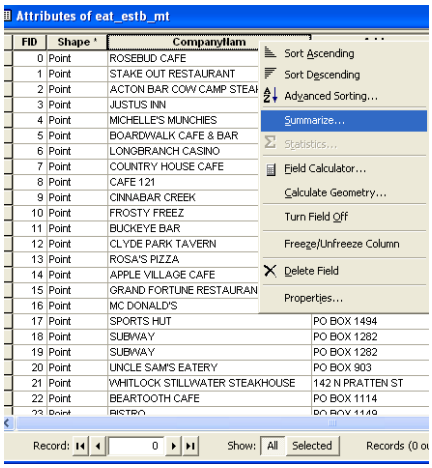
2. Right click on *mt_sic4_5812_Events* and go to Data > Export Data. You should choose to export All features, with this layer's source data, to your folder. This will save the data as a file on your computer, instead of a temporary file saved only in the computer's memory. Choose the radio button to indicate that you want to use the coordinate system from the data frame. Recall that you have defined the projection for *the mt_sic4_5812_Events*; with this information ArcMap can project the data to MT's projected coordinate system.



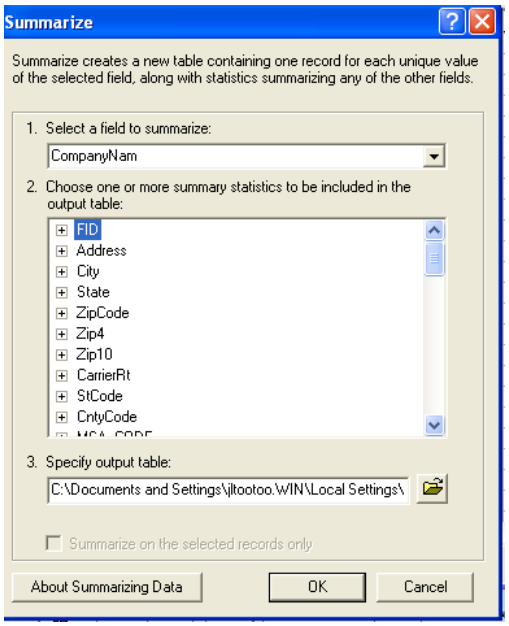
3. Name your file *eat_estb_mt.shp* and save it.

Spatial selection:

1. Now that both of datasets are correctly projected in the same system we can responsibly perform some spatial operations, let's start with a spatial selection. To make this interesting let's say we are looking for the most numerous franchise restaurants in the state of MT.
2. Open the table of your *eat_estb_mt.shp* and right click on the CompanyNam attribute field (should be 3rd from the left). Choose Summarize to get a count of the number of records in the table with names in common.

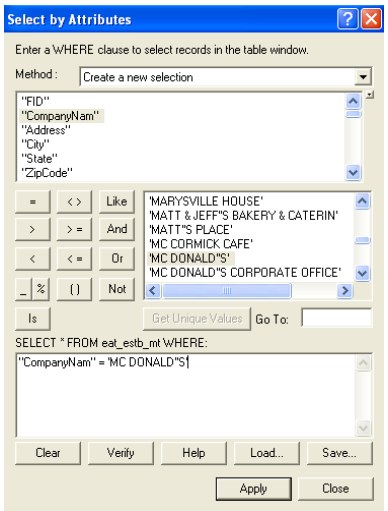


3. Stick with the defaults, but name and save your summary table appropriately.

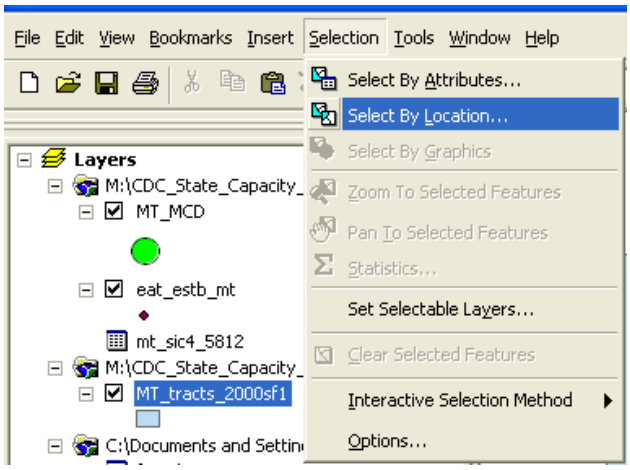


4. Once the operation is complete, add the result table to your map. Open the table and sort the Count_CompanyNam field by descending order. Which franchise is most common in the dataset?

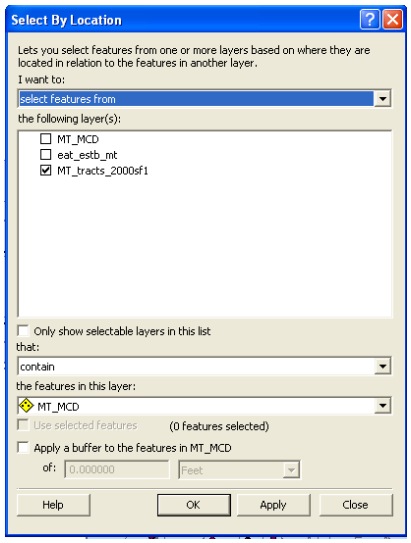
5. Let's say we are interested in knowing all Census 2000 tracts that contain a McDonald's Restaurant. To do this you need to first create a tabular selection in *eat_estb_mt.shp.* Open the attribute table and click Options > Select by Attribute. Select records using the query "CompanyNam"="MC DONALD'S".



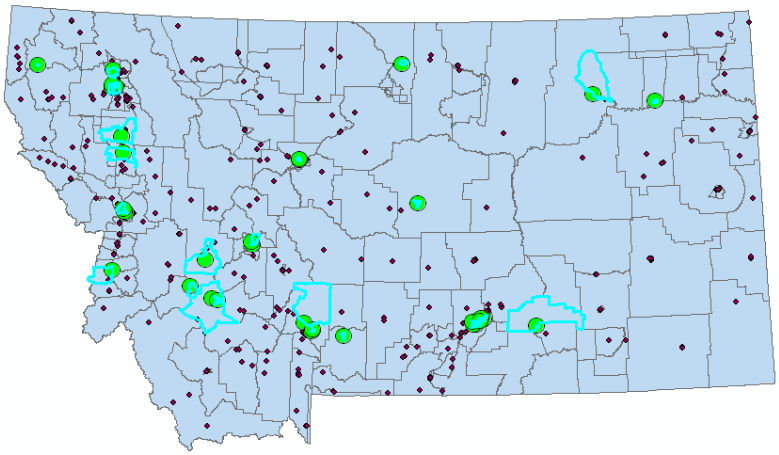
6. Apply the query and take a look at the table. How many records did you get for McDonald's? The number should match up with the count for the restaurant in your summary table.
7. Close your table and take a look at your map, all the eating establishments that met the requirements of your table query are selected. Export this selection as a new shapefile, name it *MT_MCD.shp* and choose to add it to your map.
8. Now we are ready to make a spatial selection. Click on the selection tab and choose Select by Location to answer the question: which census tracts contain a McDonald's Restaurant?



9. You will select features from *MT_tracts_2000sf1* that contain features from the *MT_MCD* (MT McDonald's) layer, click OK once you have done this.



10. Take a look at your map, these are the Census 2000 tracts that contain McDonald's.

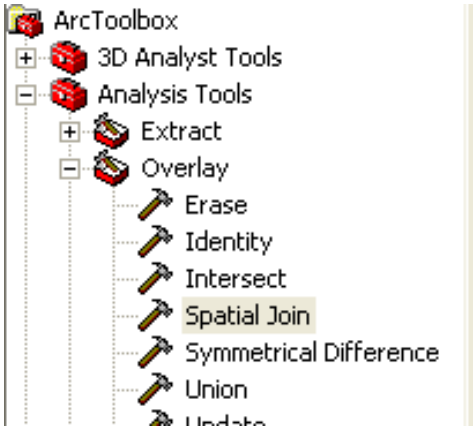


11. Now open the attribute table associated with *MT_tracts_2000_sf1* and toggle the buttons on the bottom of the table to Show only selected records. How many tracts contain McDonald's?
12. This is great information and a good way to understand your data with a query, but what about making a more formal connection between the data? How about determining the count of the McDonald's restaurants by 2000 Census tract? We can accomplish this with a spatial join.

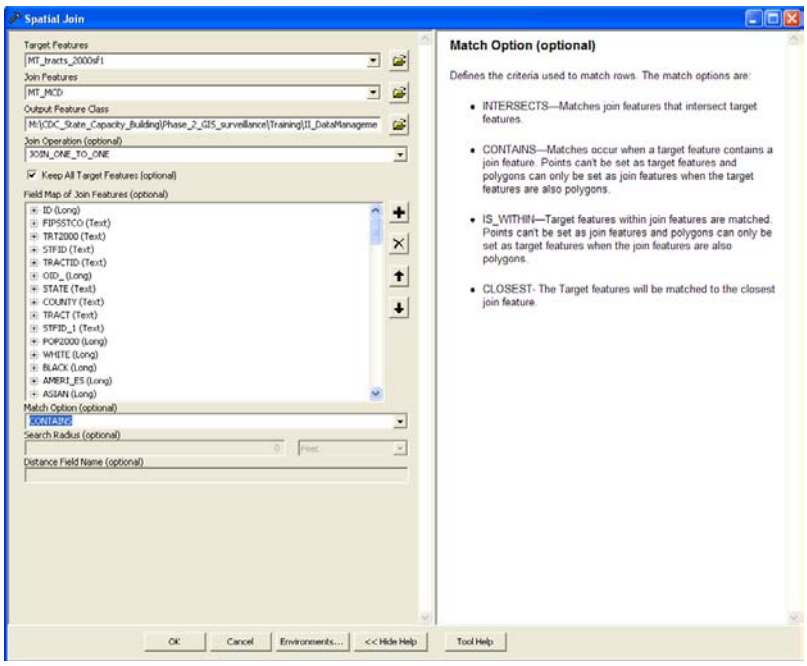
Spatial Join:

1. We are now ready to perform a spatial join between two related layers of data with the goal of answering a question that was initiated by a spatial selection: How many restaurants of interest are present with each host tract?

2. Open ArcToolbox > Analysis Tools > Overlay > Spatial Join.



3. Our Target Feature (the data to which we are appending) will be: *MT_tracts_2000sf1*. Our Join Feature (the data we are appending) will be: *MT_MCD.* Join Operation: we want to JOIN_ONE_TO_ONE. Keep default keep all target features checked. Match option: CONTAINS.



4. Notice that the spatial operation creates a new dataset that you will need to name and find a home for on your computer. Once you have saved it take a look at your table. For every record in your tract table (each record represents a closed polygon geometry) where a McDonalds is present; information for the restaurant has been appended to the tract. You can check this by looking at the Join_count field.

5. Are there any tracts with multiple McDonald's? If we included sales volume or number of employees (numeric values) in the original eating establishment information we would have had the option of summarizing these values in cases where more than one McDonalds is present.

If you have time…

Try repeating the processing above for another sic code (there are two additional .dbf files with other sic codes of interest: 5432 (fruit and vegetable markets) and 5921 (liquor stores).