Exercise 3 - Geodatabase Design for Legacy Data

Session Objectives: At the conclusion of this session, you will be able to:

- ❑ Create a file geodatabase with a thematic structure.
- ❑ Convert a mess of shapefiles and miscellaneous datasets into a clean and smart file geodatabase.

Material Created By: Eric Kelchlin (Oct 2011)

Revision: November 2013

Software: ArcGIS 10, SP 5

Data Path: C:\ or D:\CSP7200\Exercise3

Databases: NCTC.gdb (student creates)

Imagery: Imagery\Aerial_2009.sid

MXD:

SECTION 1 – Managing Legacy Data with a File Geodatabase

This section demonstrates how to use tools in ArcCatalog to convert a mess of shapefiles and other datasets into a clean and smart geodatabase. The first step is to examine your data layers and group similar items (i.e. shapefile layers) into file folders. These folders can become your feature datasets and thereby structure your geodatabase. The structure can be thematic (i.e. theme based), geographical, project specific or a combination of all three. You can have one geodatabase or many; this will depend on how much data you have and also on your project objectives and office organizational requirements.

When you create the feature datasets you will also have to decide on a datum and projection. You can assign different spatial references for the feature datasets within the same geodatabase, but this is best done for geographically based designs. Thematic designs will typically have the same datum and projection for the data layers in the geodatabase.

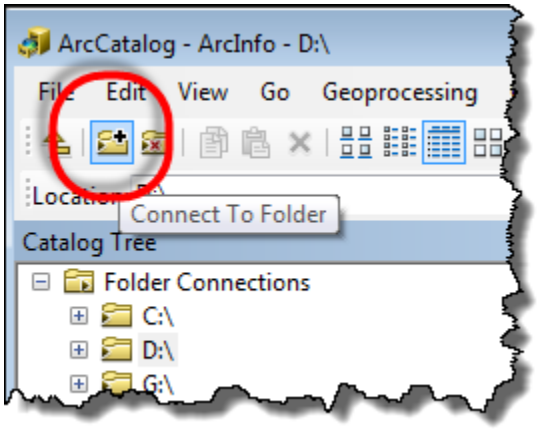
Once you have a geodatabase with a structure of empty feature datasets, the next step is to create feature classes from your legacy data. This is your opportunity to delete duplicate layers and messy data without metadata. The data clean-up and compilation efforts can be done as part of the data import/loading process while creating new feature classes.

All data are located in the **CSP7200\Exercise3** folder, unless otherwise stated.

TASK 1 – Create a File Geodatabase with a Thematic Structure

1. Open ArcCatalog.

2. Connect to the **CSP7200** folder by clicking on the Connect to Folder tool and navigating the folder. All of your class exercises reside in this folder.



3. Create a new file geodatabase in the **Exercise3** folder by simply right-clicking on the folder, choose **New > File Geodatabase**.

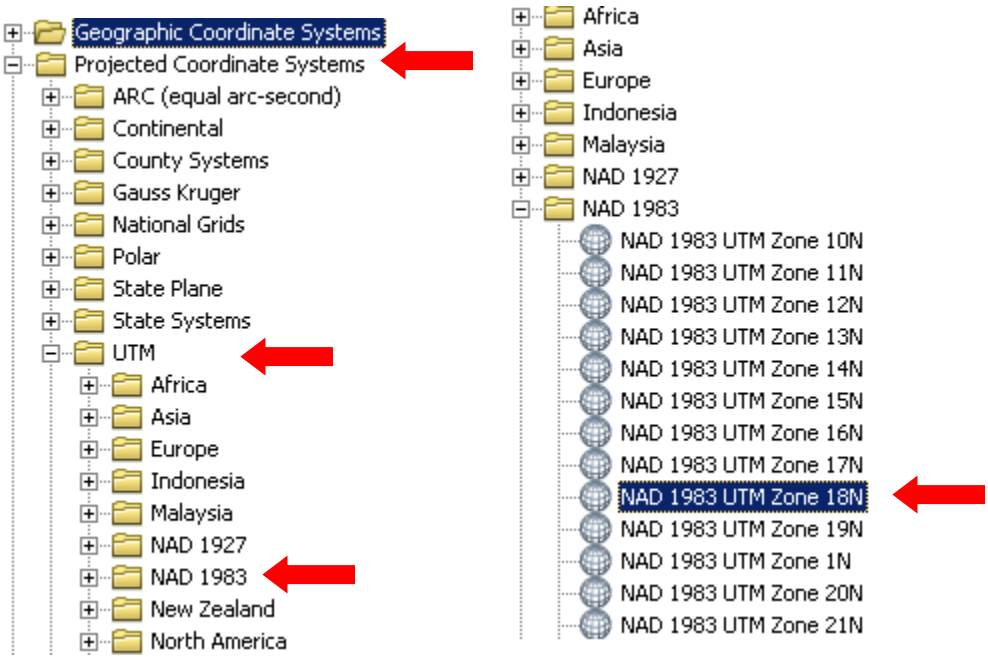
4. Name the geodatabase "**NCTC**" and hit enter on your keyboard. A new file geodatabase has been created.

5. Next, create a new feature dataset that will store our feature classes. Right-click the **NCTC** geodatabase and select **New > Feature Dataset**.

6. Name the feature dataset **Boundary**. Click **Next**

.

7. We need to assign a coordinate system for the feature dataset. Select **Projected Coordinate System > UTM > NAD 1983 > UTM Zone 18N** for the coordinate system. This is the standard projection used at NCTC. Click **Next**.



8. Click **Next** in the vertical coordinate system dialog, you won't need to define this in most situations.

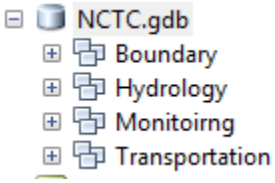
9. In the next dialog, keep the default tolerances and click **Finish**. Notice how the feature dataset is added below the **NCTC** geodatabase.

10. Create another feature dataset, however, since we are using the same coordinate system for all feature datasets, simply right-click on the **Boundary** feature dataset and select **Copy**.

11. Next, right-click on the **NCTC** geodatabase and select **Paste.** Click **OK** at the data transfer dialog.

12. Rename the feature class by simply right-clicking on the **Boundary_1** feature class and selecting **Rename**. Rename it to **Hydrology**. This technique can be quite the time saver.

13. Create two more feature datasets using the same methodology. Name the first "**Transportation"** and the other "**Monitoring"**.



14. Can you think of any other feature datasets that would help organize your GIS data at your office?

_____

_____

_____

_____

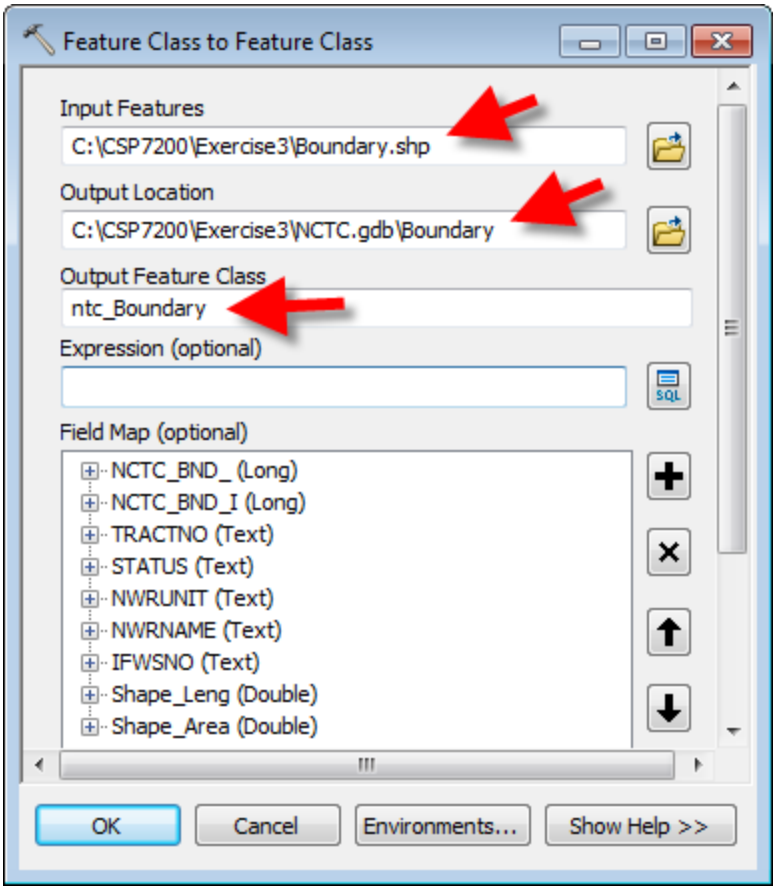
_____

TASK 2 – Creating a New Feature Class from an Old Shapefile

1. Right-click on the **Boundary** feature dataset. Select **Import > Feature Class (single…)**.

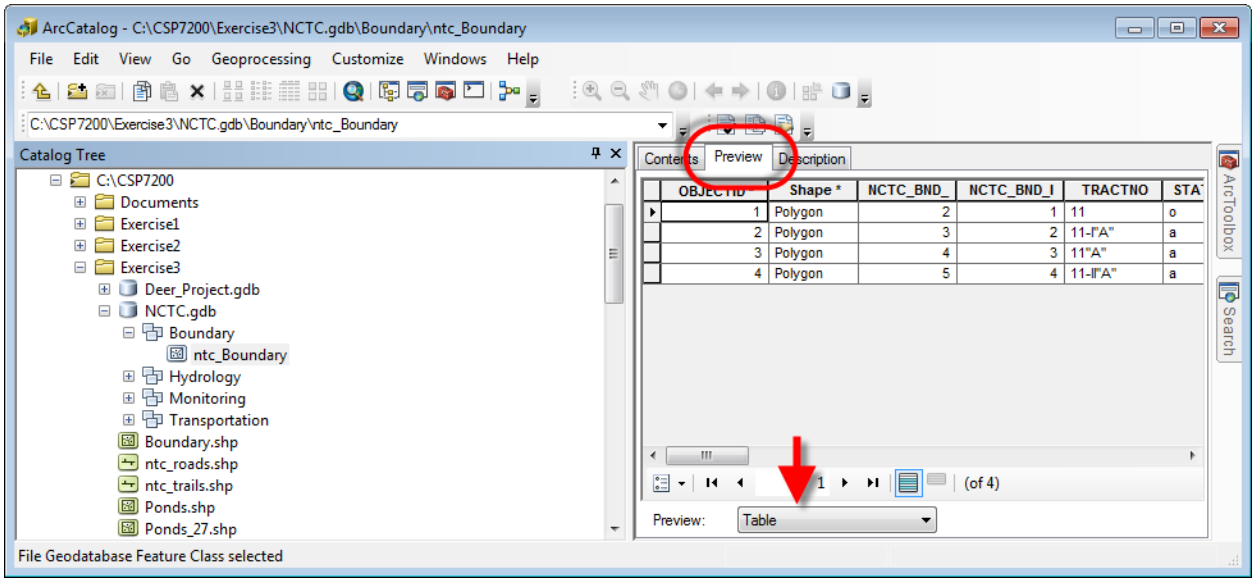
2. For **Input Features**, click the browse folder button and navigate to the **Exercise3** folder. Choose the **Boundary.shp** file and click the **Add** button.

3. Type "**ntc_boundary** " in the **Output Feature Class** field. Click **OK**.

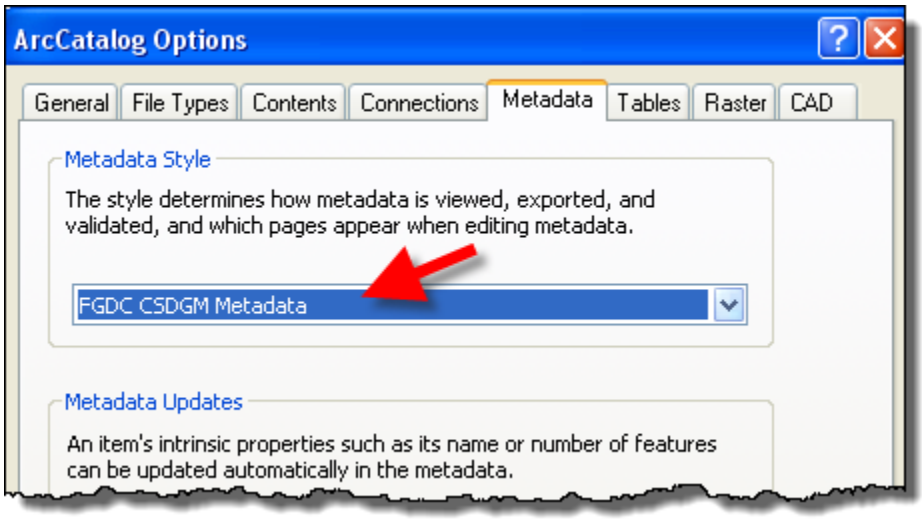


4. Expand the **Boundary** feature dataset and select the new feature class. Click on the **Preview** tab to view the new feature class you just imported.

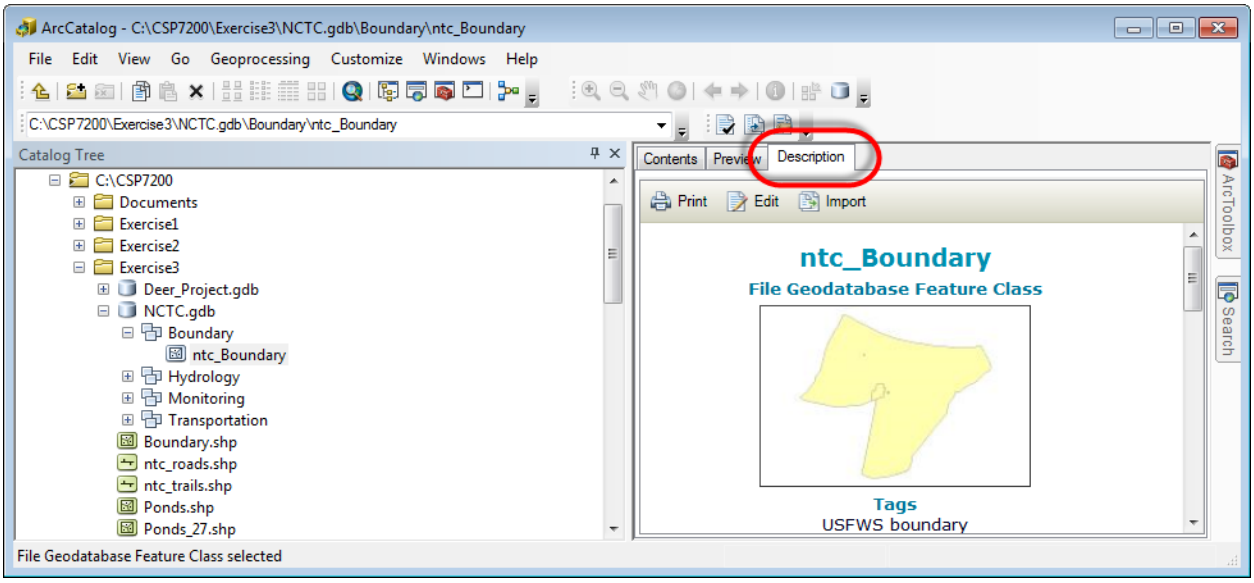
5. Change the Preview field from Geography to Table. What two land status attributes are listed in the table? _____ and _____.



6. On the main toolbar, click **Customize > ArcCatalog Options** and make sure your metadata style is **FGDC CSDGM Metadata**. This is short for the Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata.



7. Click on the **Description** Tab to view the metadata. Search through the metadata and find what these two attributes indicate.



8. What does the coded “o” and “a” status attributes refer to?

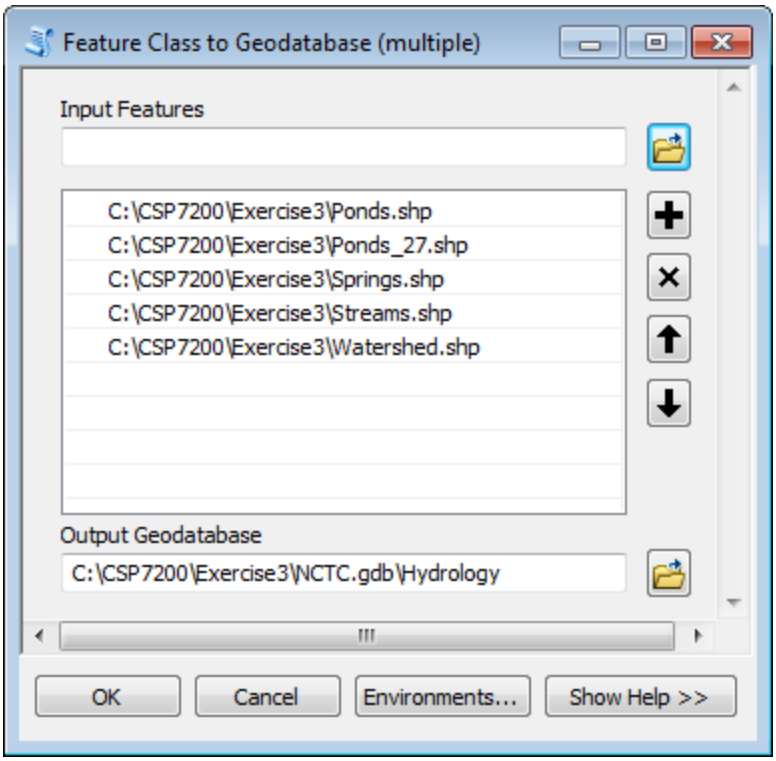
_____

_____

TASK 3 – Creating Multiple Feature Classes from Multiple Shapefiles

1. Next, let's bring in a batch of shapefiles all at once to save time. Right-click on the **Hydrology** feature dataset and select **Import > Feature Class (multiple…).**

2. Click the browse folder button and navigate to the **Exercise3** folder. Hold down the **Ctrl Key** and select the **Ponds, Ponds_27, Springs, Streams** and **Watershed** shapefiles. Click **Add** and **OK**.



3. Examine the spatial reference for the NEW **Ponds_27** feature classes by right-clicking on each feature class and selecting **Properties**, then clicking on the **XY Coordinates** Tab. What are the XY Coordinates for these layers?

   Ponds feature class:_____

   Ponds_27 feature class:_____

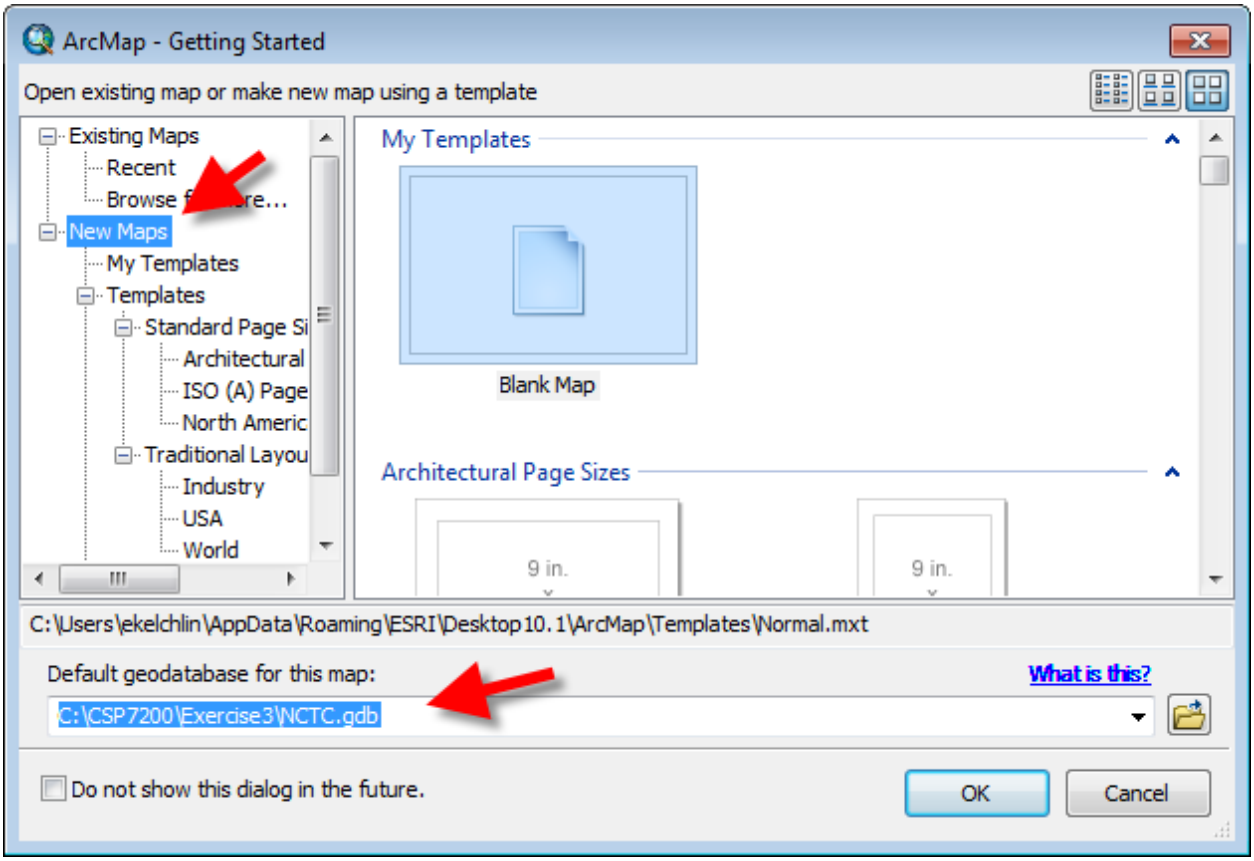
4. Do the same for the Hydrology feature dataset. The XY Coordinates should be the same as your feature classes.

5. Now, examine XY Coordinates on the OLD **Ponds** and **Ponds_27** shapefiles. What are the XY Coordinates for these layers?

   Ponds shapefile:_____Ponds_27 shapefile:_____

6. Click on the button on the main toolbar in ArcCatalog to launch ArcMap and examine your new feature classes.

7. Click on **New Maps** in the ArcMap Getting Started dialog. Change your default geodatabase to NCTC and click **OK**.



8. Add the **Aerial_2009.sid** file from the **Imagery** folder.

9. Add the **Ponds** and **Ponds_27** feature classes.

10. Do the ponds overlay correctly or is there a shift?_____.

11. If the shapefile has an undefined coordinate system (i.e., does not have an associated projection .prj file) the shapefile will be imported into the feature dataset and assume what?_____.

The point here is that you need to have a projection defined for all your shapefiles before you perform a multiple/batch import. The **Ponds_27** shapefile was actually digitized in the NAD27 datum, not NAD83.

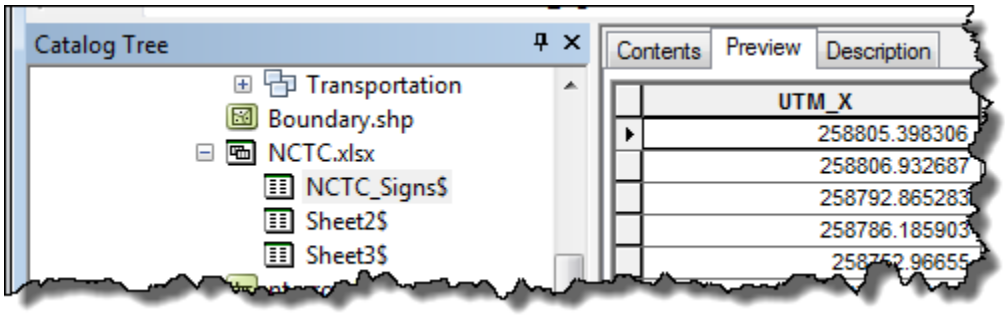
The geodatabase will assign each imported layer whatever you have defined in the feature dataset, which can lead to large spatial errors. It does not matter what projection you have for the shapefile, just as long as you have one. The geodatabase will automatically transform the data into the spatial reference you have in the feature dataset.

TASK 4 – Importing XY Tabular Data

This example shows you how easy it is to take an Excel spreadsheet containing XY coordinate data and create a new feature class.

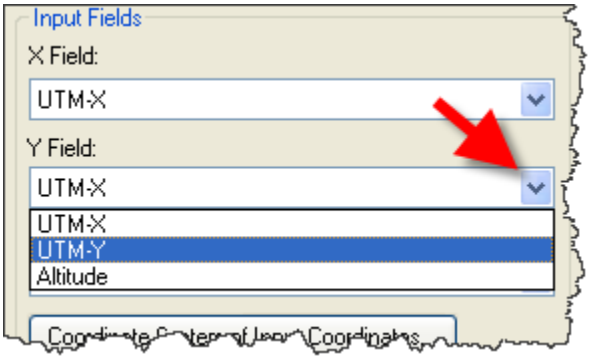
1. Maximize the **NCTC.xlsx** file in the Catalog Tree and highlight the **NCTC_Signs$** worksheet. The **$** symbol indicates that it's an excel worksheet.

2. Look at the table in the **Preview tab** and notice that we already have a UTM_X and UTM_Y coordinate fields and we also know the datum and coordinate system.



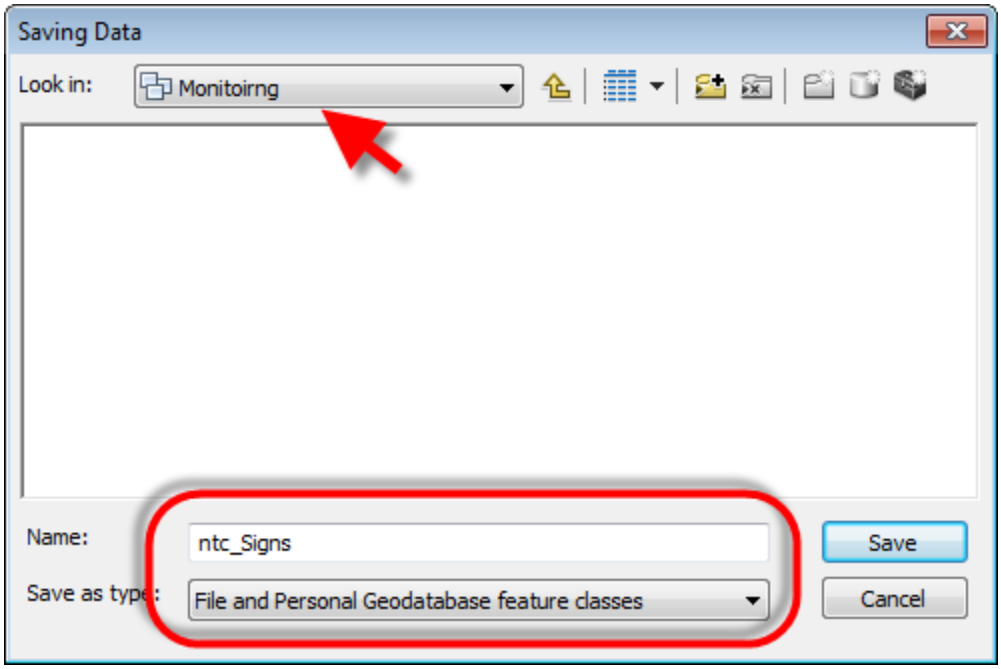
3. Right-click on **NCTC_Signs$** worksheet and select **Create Feature Class > From XY Table…**

4. Change the **Y Field** to **UTM-Y**.



5. You can chose to leave the Z Field blank or change it to the Altitude field. If you change it to Altitude you'll have a Point Z data type which is used for 3 dimensional data viewing.

6. Click on the browse button in the output section and save your new feature class as **ntc_signs** in the **Monitoring** feature dataset. Make sure to change the file type to "File and Personal Geodatabase feature classes". Click **Save**, then **OK**.



7. Peruse your geodatabase. Was **ntc_signs** added to the **Monitoring** feature dataset_____? This is a strange quirk. Use your mouse to drag and drop your new feature class into the **Monitoring** feature class.

8. Right-click on the **ntc_signs** feature class and select Properties. Click on the **XY Coordinate System tab** and notice that the proper datum and coordinate system was applied to the data.

9. Repeat the same table to feature class process, but this time save the feature class inside the geodatabase, not inside any feature dataset. Use a different file name. Look at the XY Coordinate System for this second feature class. Is it defined_____?

10. How would you define the XY Coordinate System of a table during table to feature class process if you're not saving it into a feature dataset?

TASK 5 – Performing SQL Queries while Importing Shapefiles

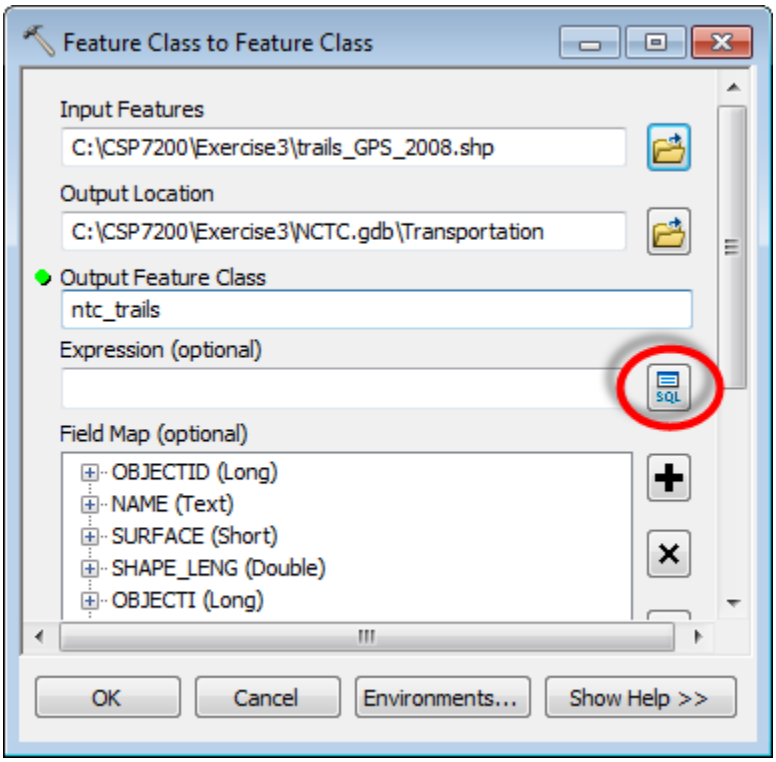
Let's import another shapefile into the NCTC geodatabase, but this time we're going to clean-up the data during the import with an SQL query.

1. Close ArcMap if you have not done so already.

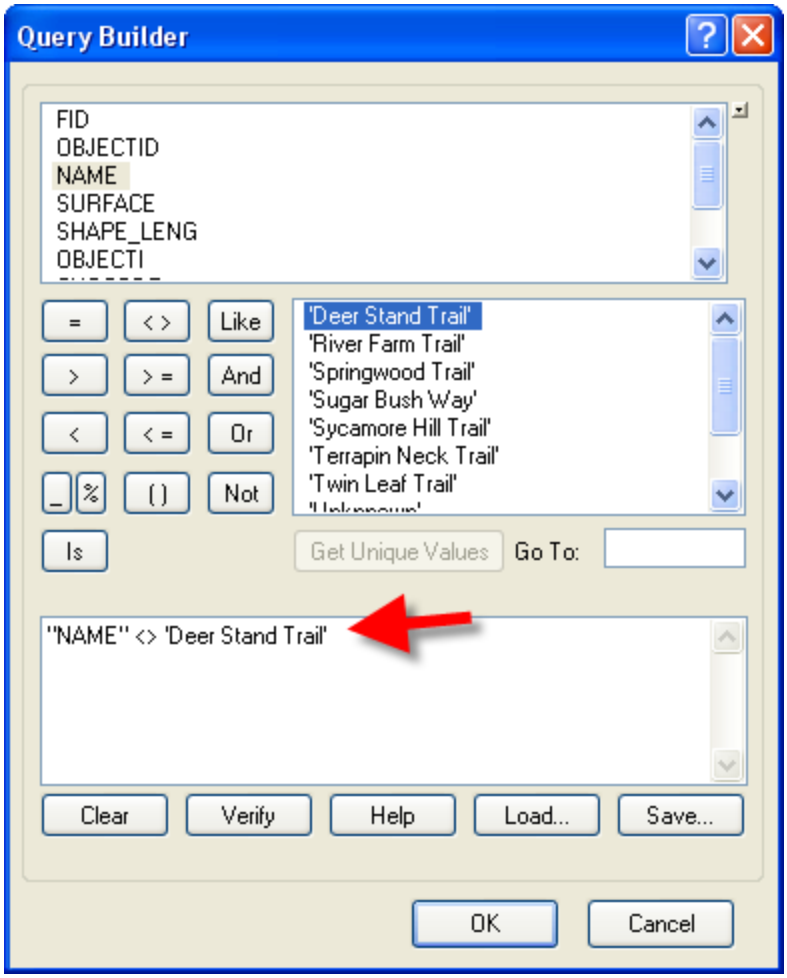
2. Right-click on the **Transportation** feature dataset and select **Import > Feature Class (single…).** Input the **trails_GPS_2008** shapefile.

3. Name the new feature class **ntc_trails**. Continue onto the next step before finishing.

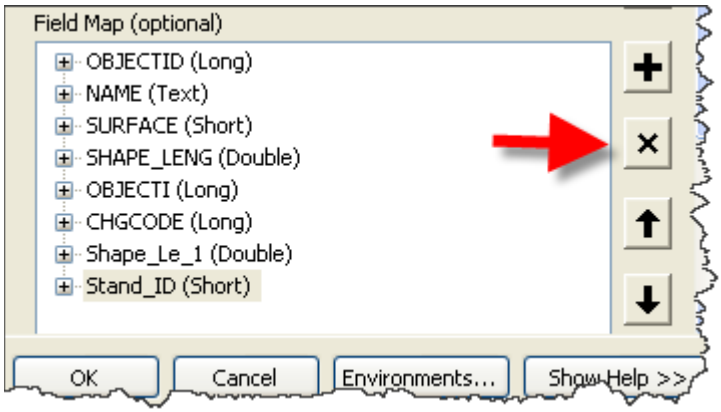
4. Click the **SQL** button to create a query.



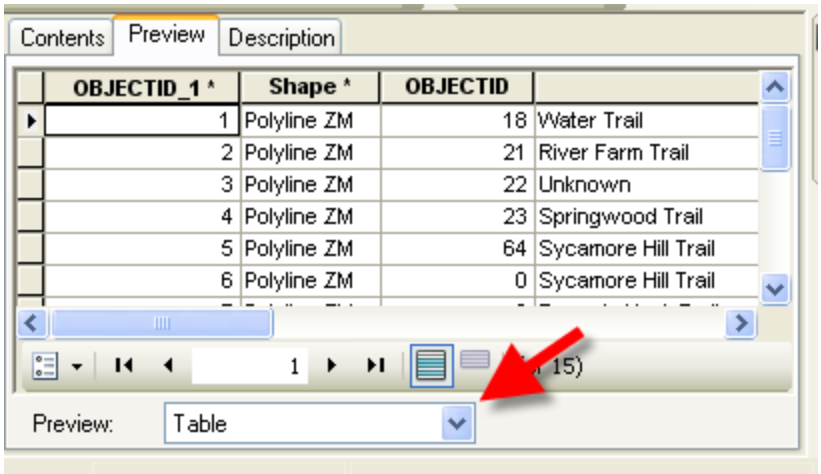
5. Create the query to include all the trails except Deer Stand Trails as shown below. The instructor will walk you through this part. Click **OK** when done to return to the previous screen.



6. Let's clean-up the data a bit more before we finish the import. Delete the last 5 fields in the **Field Map** section using the delete button. Click **OK**.



7. Look at the Table in the **Preview tab** to make sure the data was imported correctly.



8. Repeat the same process above, but this time only import the Deer Stand Trails. Name the new feature class **ntc_deer_stand_trails**.

TASK 6– Importing a Feature Class

We're almost done. Let's import one more data type; a feature class from a file geodatabase.

1. Right-click on the **Transportation** feature dataset and select **Import > Feature Class (single…)**.

2. For **Input Features**, click the browse button and navigate to the **Deer_Project** geodatabase.

3. Select the **ntc_hunter_parking** feature class from the **Deer_Hunt** feature dataset. Click the **Add** button.

4. Name the Output Feature Class **ntc_hunter_parking.** Click **OK** when done.

5. Click on the **Preview tab** and examine the data you just imported.

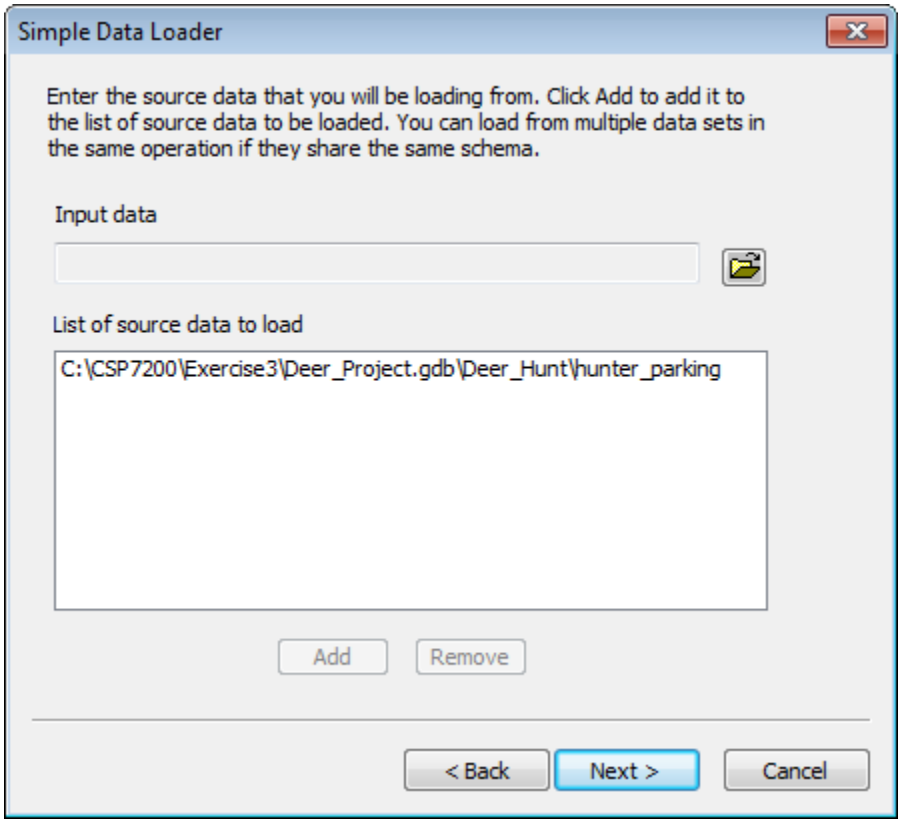
TASK 7– Appending Data using the Load Command

This is an example where we mapped hunter parking lots on two separate occasions using different GPS units. We now have to do some data management to append the new data to our existing **ntc_hunter_parking** feature class. The Load command allows you to append data if the source schema differs, which is the case here. The Append tool in ArcToolbox can be used if the data have the same schema.

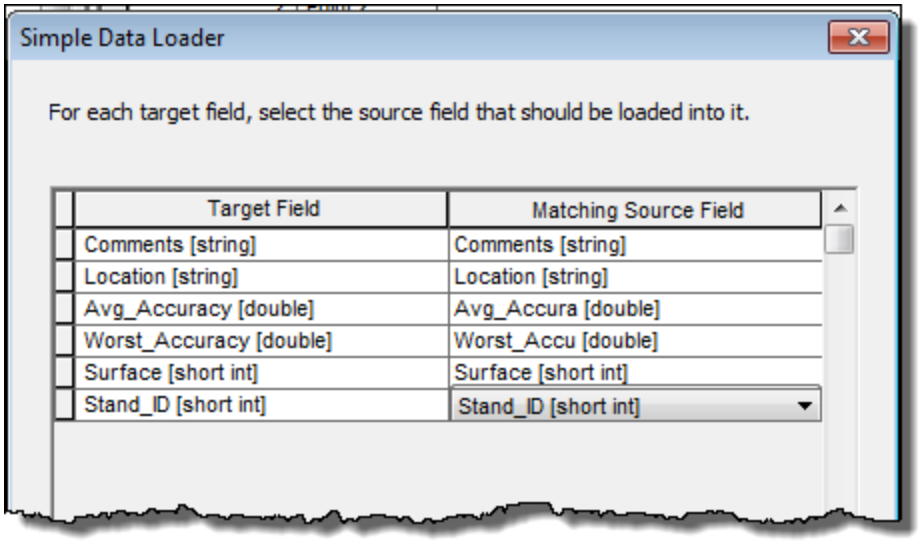
1. Right-click on the **ntc_hunter_parking** feature class you just imported. Choose **Load > Load Data**. Click **Next** on the first screen that appears.

2. For input data, click browse and navigate to the **Deer_Project** geodatabase in the **Project** folder. Select the **hunter_parking** feature class.

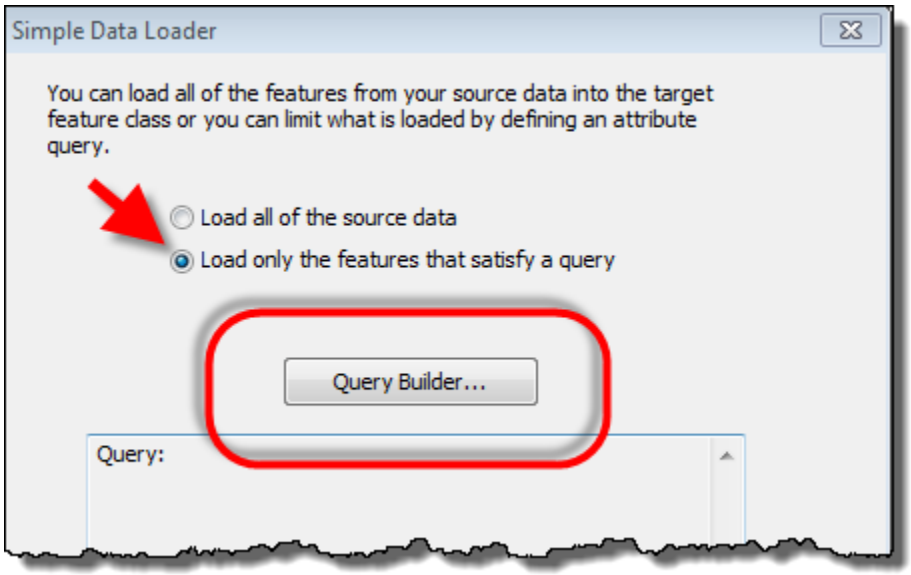
3. From the **Simple Data Loader** dialog, click **Add** to include the data in the **List of source data to load** field. Click **Next** twice.



4. Click on the drop-down arrows for the **<none>** values in the **Matching Source Field** column and match-up the Source with the Target fields. Note: any unmatched fields will be null in the appended records. Click **Next.**



5. Now, the important part. We need to create a simple query to select only the records we want appended to our existing feature class. Ensure that the "**Load only the features that satisfy a query**" is checked.



6. Click the **Query Builder** button and select only values that = **1** in the **chgcode** field. Do you remember how to do this?

7. Preview the **ntc_hunter_parking** feature class tabular data when done to make sure the records were appended.

Congratulations! You imported shapefiles, tables and future classes into your new file geodatabase.