

Joining Data Tables to SABINS Data

September 2011

This guide shows users how to join the School Attendance Boundary Information System (SABINS) data to Census 2010 tabular data, Common Core of Data (CCD), and individual schools. In addition, this guide shows users how to store data in a file geodatabase.

This guide assumes the user has already downloaded and unzipped SABINS data from www.sabinsdata.org. For assistance in this, please see the *Using the SABINS Data Finder* tutorial.

Four files are necessary to follow all steps outlined in this tutorial document.

1. A shapefile of school attendance boundary polygons. For this example, we will use the kindergarten shapefile called **PY_SABINS_0910_00_US.shp**
2. The second file contains Census 2010 data aggregated by school attendance boundary. In this tutorial, we will work with race data. For kindergarten, **nhgisXXXX_ds171_sab_00.csv** (where *XXXX* is a number which represents how many extracts the user has previously completed).
3. The third file is a SABINS-to-CCD crosswalk table that links unique school IDs (*NCESSCH*) alongside their corresponding school boundary IDs (*SABINSID*). There is a unique table for each grade. For kindergarten, the table is **NS_SABINS_CCD_0910_00.dbf**
4. The fourth file contains CCD public school information by school ID. The one file, which contains all grades is **CCD0910.dbf**

Files 1 and 2 are downloaded using the SABINS Data Finder while files 3 and 4 are downloaded from www.sabinsdata.org/user-resources/school-data



Funding provided by the National Science Foundation.
Project support provided by the Minnesota Population Center and The College of William and Mary.

Background Information

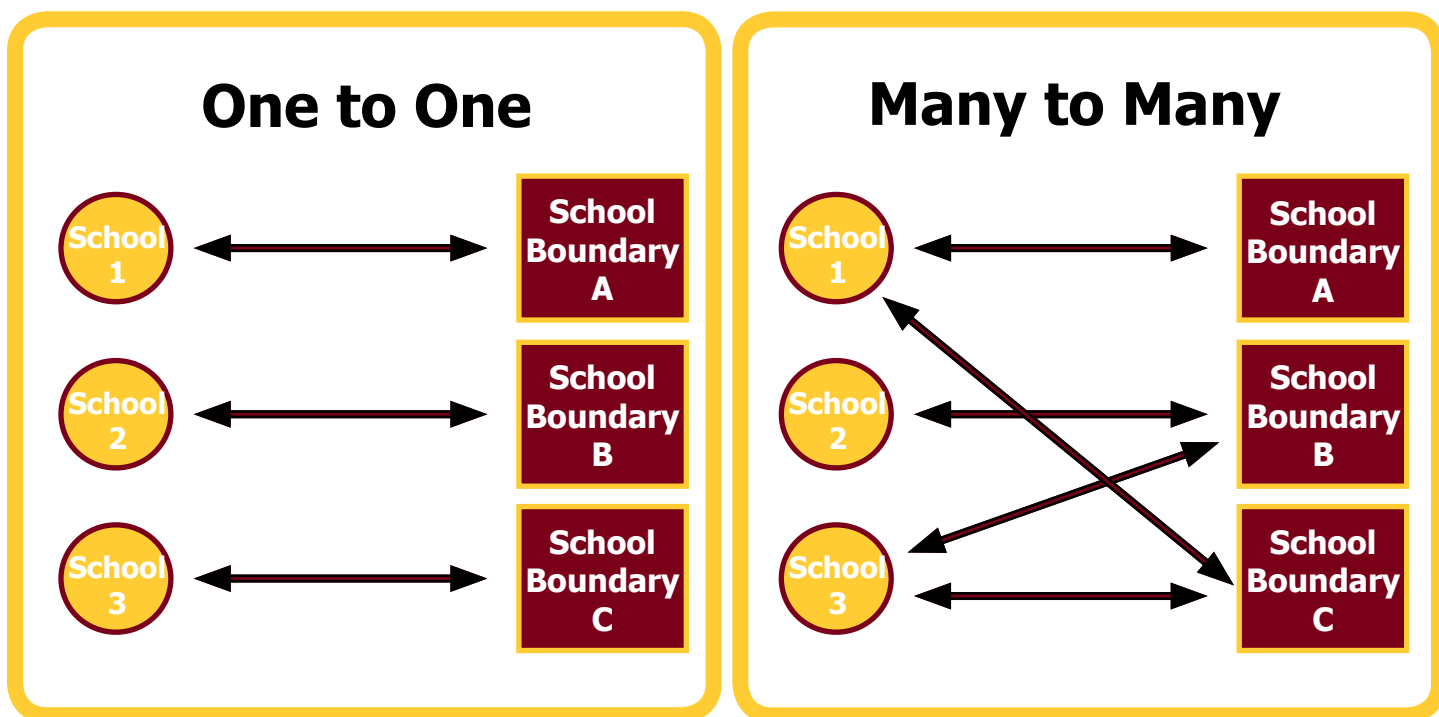
Schools and school attendance boundaries have a somewhat counter-intuitive relationship. In most cases, one school serves a single attendance boundary. But schools and school attendance boundaries do not always have a one-to-one relationship. Some schools serve multiple, separate attendance boundaries, and sometimes more than one school serves the same attendance boundary. This makes the relationship between schools and school attendance boundaries a more complicated many-to-many relationship.

To accommodate schools' and attendance boundaries' many-to-many relationship, the SABINS project stores information about either entity in separate tables. In the attendance boundary shapefile (**PY_SABINS_0910_00_US.shp**), each school attendance boundary is listed once and identified by a unique 27-digit SABINSID. In the public school information data table (**CCD0910.dbf**), each school is listed once and identified by a unique 12-digit NCESSCH. A third table lists the NCESSCH's and SABINSID's side by side; schools and school attendance boundaries may be listed multiple times depending on their corresponding relationships.

We accommodate this many-to-many relationship by providing two options for joining schools and attendance boundaries.

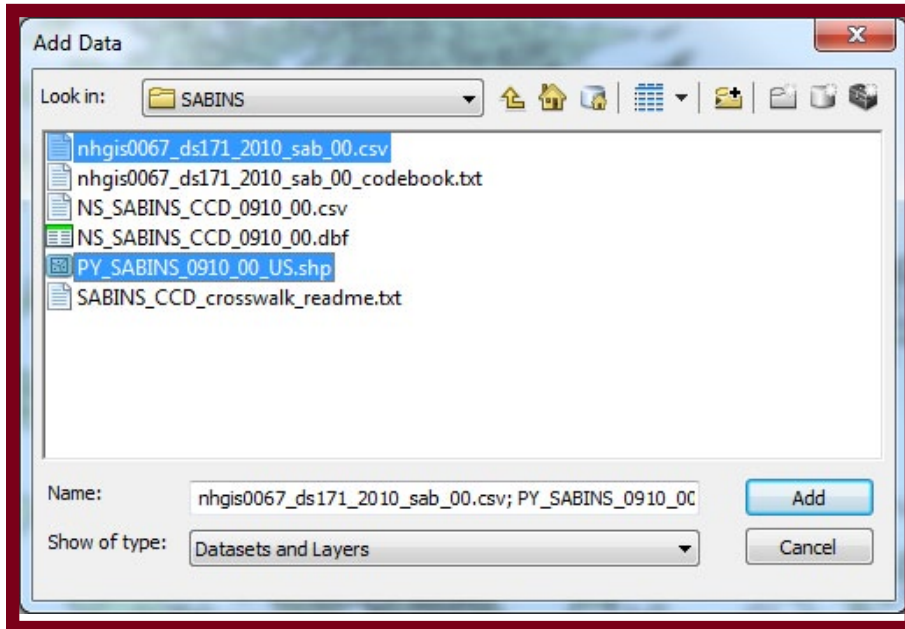
- Option A** joins school data to individual SABINS polygons. In this option, each school attendance boundary will be accounted for. But polygons associated with multiple schools will only have one school listed with them, i.e., not all schools will be listed.

- Option B** joins school boundary data to individual schools. In this option, each school will be accounted for. Duplicate SABINSID's will be listed where multiple schools share a single attendance boundary.



Joining Census 2010 Data to a SABINS Shapefile

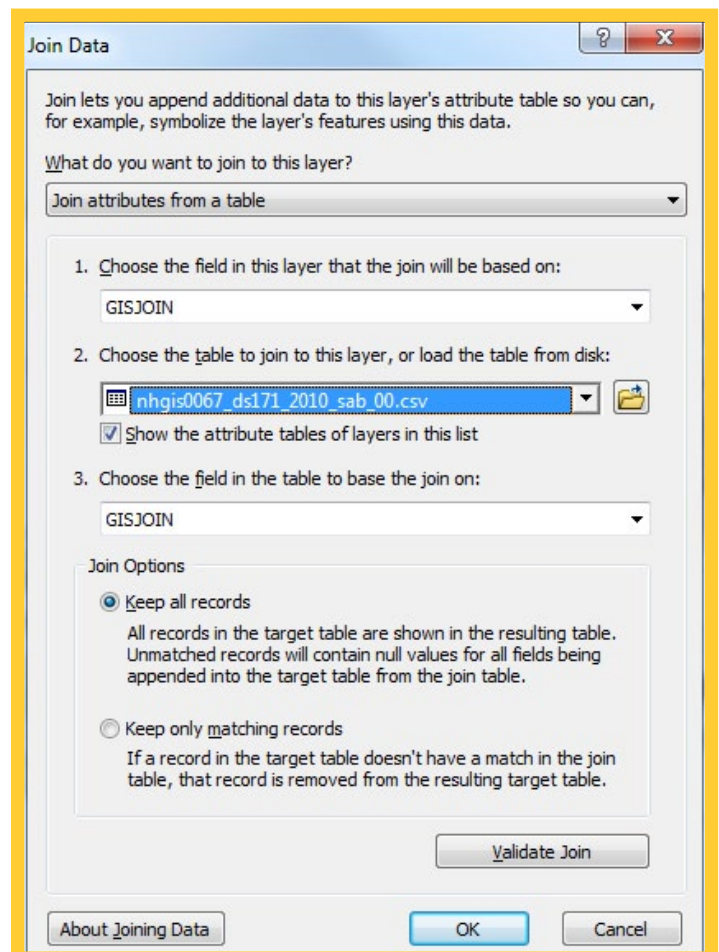
Start Esri ArcMap and open a new document. Click the add data button, and browse to the location of the Kindergarten shapefile and the race data file. For this example, the files names are as follows: **PY_SABINS_0910_00_US.shp** and **nhgisXXXX_ds171_2010_sab_00.csv**



Join the two files by right clicking **PY_SABINS_0910_00_US.shp** in the table of contents. Then select Joins and Relates>Join... In the dialogue box that opens:

- Under "What do you want to join to this layer?" select **Join attributes from a table**.
- For both the numbers 1 and 3 (the field the join will be based on), select **GISJOIN**. Select **nhgisXXXX_ds171_2010_sab_00.csv** under number 2, "Choose the table to join to this layer, or load the table from disk:"
- Make sure the radio button next to "Keep all records" is selected.
- Press OK.

You may view the joined table by right clicking **PY_SABINS_0910_00_US.shp** in the table of contents and selecting Open.



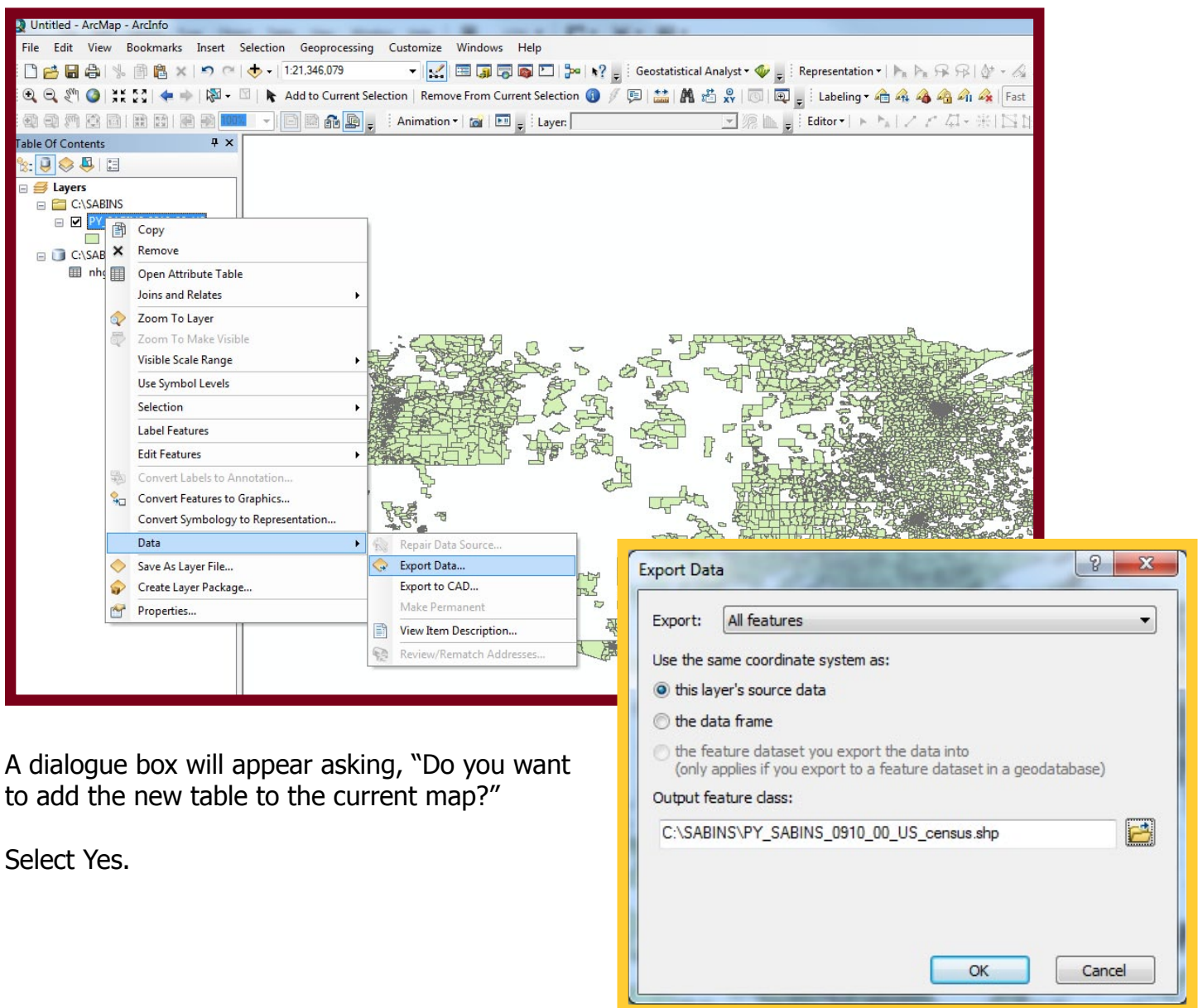
Joining Census 2010 Data to a SABINS Shapefile

After we verify the join worked correctly by inspecting the results in the shapefile's attribute table, it is best to export the joined shapefile. It is important to realize that a table join does not result in a permanent change to the files involved. Rather, it is considered virtual because the join disappears if the map document is not saved. Exporting the shapefile while a join is intact will create a new file that permanently contains all of the attributes of the original shapefile and the joined data.

To export, right click **PY_SABINS_0910_00_US.shp** in the table of contents, and select Data>Export Data...

Choose to export the joined data as a shapefile.

For the purposes of this tutorial, the exported file is called **PY_SABINS_0910_00_US_census.shp**.



A dialogue box will appear asking, "Do you want to add the new table to the current map?"

Select Yes.

Option A: Joining CCD School-level Data to School Attendance Boundaries

Link SABINS Shapefile to CCD Association Table

The SABINS shapefiles can be joined to the CCD School-level data through the use of an association table, as the shapefiles do not contain the unique key used to identify schools in CCD.

Add the following files to your map document:

- PY_SABINS_0910_00_US_census.shp** – Kindergarten school attendance boundaries with census data attached.
- NS_SABINS_CCD_0910_00.dbf** – Crosswalk table that connects school attendance boundaries to school in the CCD.
- CCD0910.dbf** – 2009-2010 Common Core of Data file.

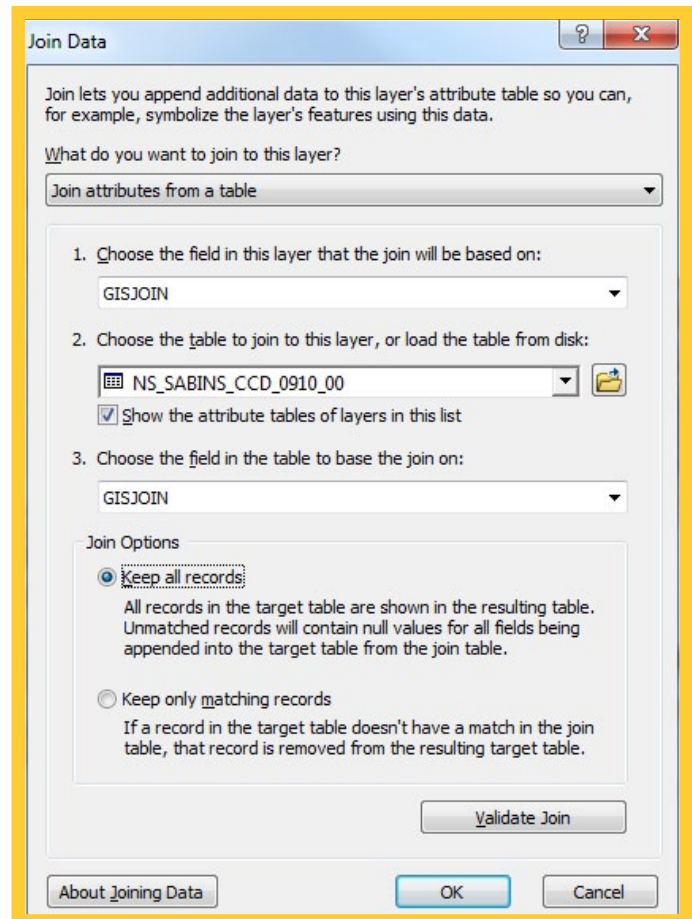
Join the association table **NS_SABINS_CCD_0910_00.dbf** to the shapefile by right clicking **PY_SABINS_0910_00_US_census.shp** in the table of contents. Then select Joins and Relates >Join...

a. Under “What do you want to join to this layer?” select **Join attributes from a table**.

b. For both the numbers 1 and 3 (the field the join will be based on), select **GISJOIN**. Select **NS_SABINS_CCD_0910_00** under number 2, “Choose the table to join to this layer, or load the table from disk:”

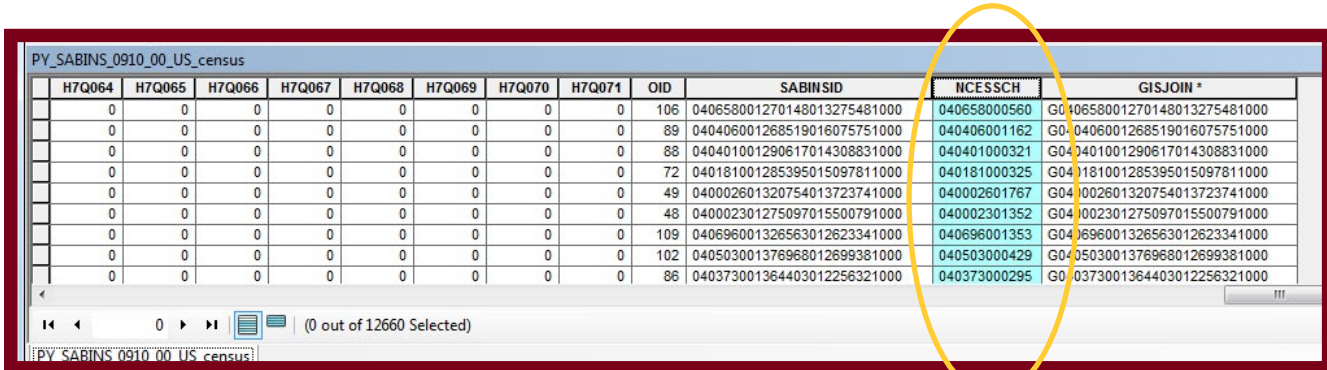
c. Make sure the radio button next to “Keep all records” is selected.

d. Press OK.



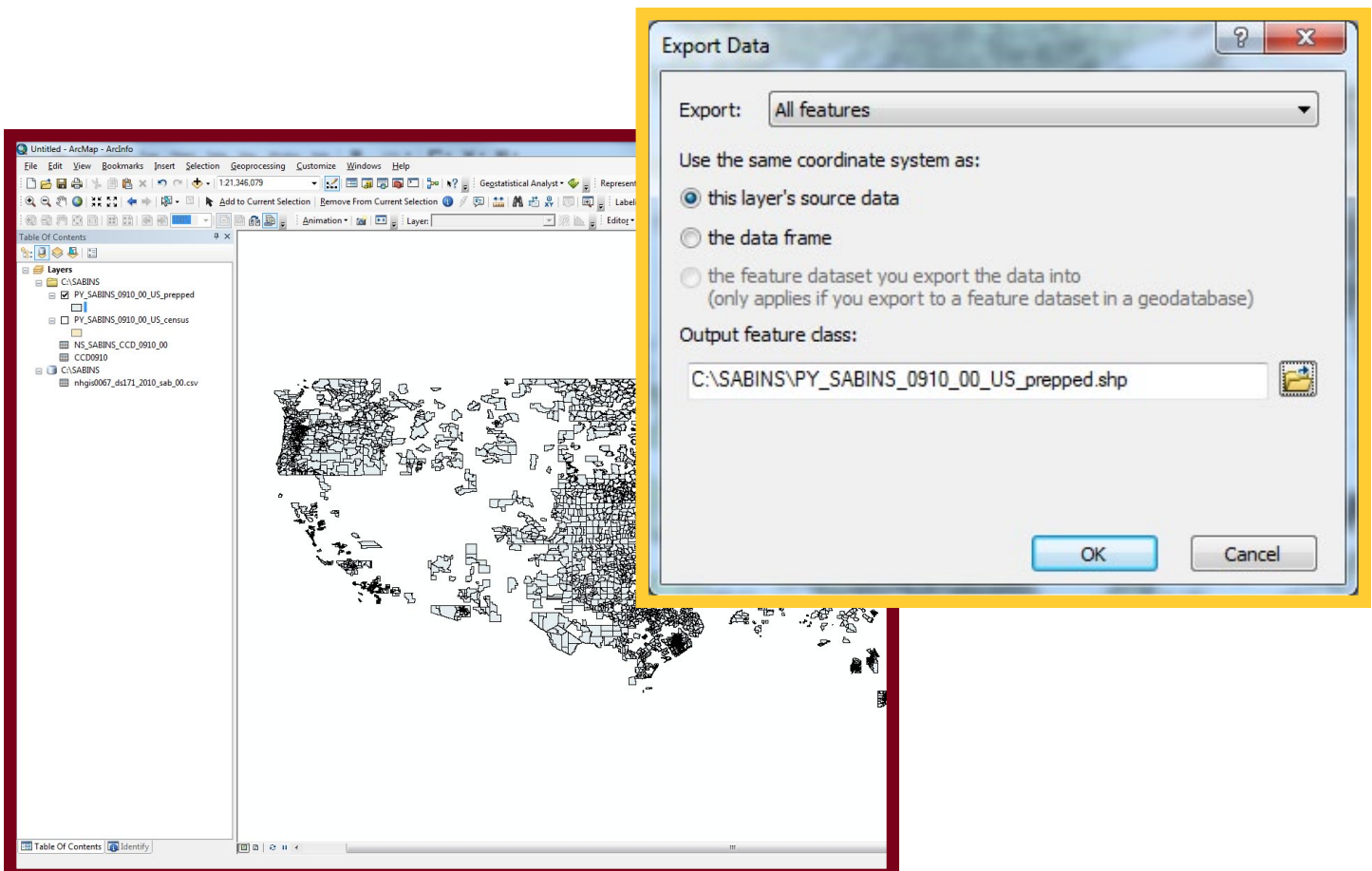
Link SABINS Shapefile to CCD Association Table

You may view the joined table by right clicking **PY_SABINS_0910_00_US_census.shp** in the table of contents and selecting Open. The figure below shows the tables that have been joined, with the critical **NCESSCH** field included.



H7Q064	H7Q065	H7Q066	H7Q067	H7Q068	H7Q069	H7Q070	H7Q071	OID	SABINSID	NCESSCH	GISJOIN *
0	0	0	0	0	0	0	0	106	040658001270148013275481000	040658000560	G040658001270148013275481000
0	0	0	0	0	0	0	0	89	040406001268519016075751000	040406001162	G040406001268519016075751000
0	0	0	0	0	0	0	0	88	040401001290617014308831000	040401000321	G040401001290617014308831000
0	0	0	0	0	0	0	0	72	040181001285395015097811000	040181000325	G040181001285395015097811000
0	0	0	0	0	0	0	0	49	040002601320754013723741000	040002601767	G040002601320754013723741000
0	0	0	0	0	0	0	0	48	040002301275097015500791000	040002301352	G040002301275097015500791000
0	0	0	0	0	0	0	0	109	040696001326563012623341000	040696001353	G040696001326563012623341000
0	0	0	0	0	0	0	0	102	040503001376968012699381000	040503000429	G040503001376968012699381000
0	0	0	0	0	0	0	0	86	040373001364403012256321000	040373000295	G040373001364403012256321000

Next, export the joined shapefile. To export, right click **PY_SABINS_0910_00_US_census.shp** in the table of contents, and select Data>Export Data... Choose to export the joined data as a shapefile. For the purposes of this tutorial, the exported file is called **PY_SABINS_0910_00_US_prepped.shp**. Add the new shapefile to the map when prompted.



Join CCD Data to SABINS Shapefile

With the SABINS shapefile now containing the NCESSCH field, it can be successfully joined to the CCD which also contains the NCESSCH field.

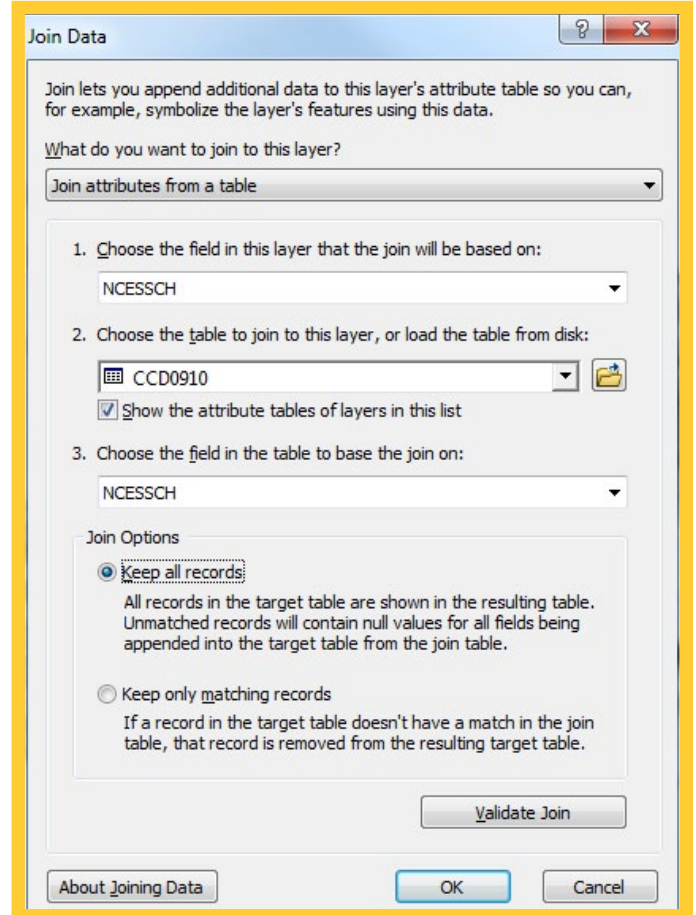
Join the school data (**CCD0910.dbf**) to the school boundary shapefile created in last step. Right click on **PY_SABINS_0910_00_US_prepped.shp** in the table of contents. Then select Joins and Relates>Join...

a. Under "What do you want to join to this layer?" select **Join attributes from a table**.

b. For both the numbers 1 and 3 (the field the join will be based on), select **NCESSCH**. Select **CCD0910.dbf** under number 2, "Choose the table to join to this layer, or load the table from disk:"

c. Make sure the radio button next to "Keep all records" is selected.

d. Press OK.



You may view the joined table by right clicking **PY_SABINS_0910_00_US_prepped.shp** in the table of contents and selecting Open. The figure below shows the tables that have been joined, with the CCD fields now attached to each polygon record in the SABINS shapefile. Notice that the school names are now visible, whereas before this join, they were not included in the attribute table of this shapefile.

	NCESSCH	GISJOIN_12	OID	NCESSCH *	FIPST	LEAID	SCHNO	STID09	SEASCH09	LEANM09	SCHNAM09
040658000560	G040658001270148013275481000	3915	040658000560	04	0406580	00560	4397	5645		BLUE RIDGE UNIFIED DISTRICT	BLUE RIDGE ELEMENTARY SCHOOL
040406001162	G040406001268519016075751000	3491	040406001162	04	0404060	01162	4396	5643		KAYENTA UNIFIED DISTRICT	KAYENTA PRIMARY SCHOOL
040401000321	G040401001290617014308831000	3485	040401000321	04	0404010	00321	4388	5605		JOSEPH CITY UNIFIED DISTRICT	JOSEPH CITY ELEMENTARY SCHOOL
040181000325	G040181001285395015097811000	3125	040181000325	04	0401810	00325	4395	5638		CEDAR UNIFIED DISTRICT	JEDDITO SCHOOL
040002601767	G040002601320754013723741000	2177	040002601767	04	0400026	01767	4392	6053		HEBER-OVERGAARD UNIFIED DISTRICT	MOUNTAIN MEADOWS PRIMARY
040002301352	G040002301275097015500791000	2171	040002301352	04	0400023	01352	4390	5611		PINON UNIFIED DISTRICT	PINON ELEMENTARY SCHOOL
040696001353	G040696001326563012623341000	3953	040696001353	04	0406960	01353	4210	4862		SAN CARLOS UNIFIED DISTRICT	RICE SCHOOL
040503000429	G040503001376968012699381000	3702	040503000429	04	0405030	00429	4211	4867		MIAMI UNIFIED DISTRICT	LAS LOMAS ELEMENTARY SCHOOL
040373000295	G040373001364403012256321000	3421	040373000295	04	0403730	00295	4212	4870		HAYDEN-WINKELMAN UNIFIED DISTRICT	WINKELMAN ELEMENTARY SCHOOL

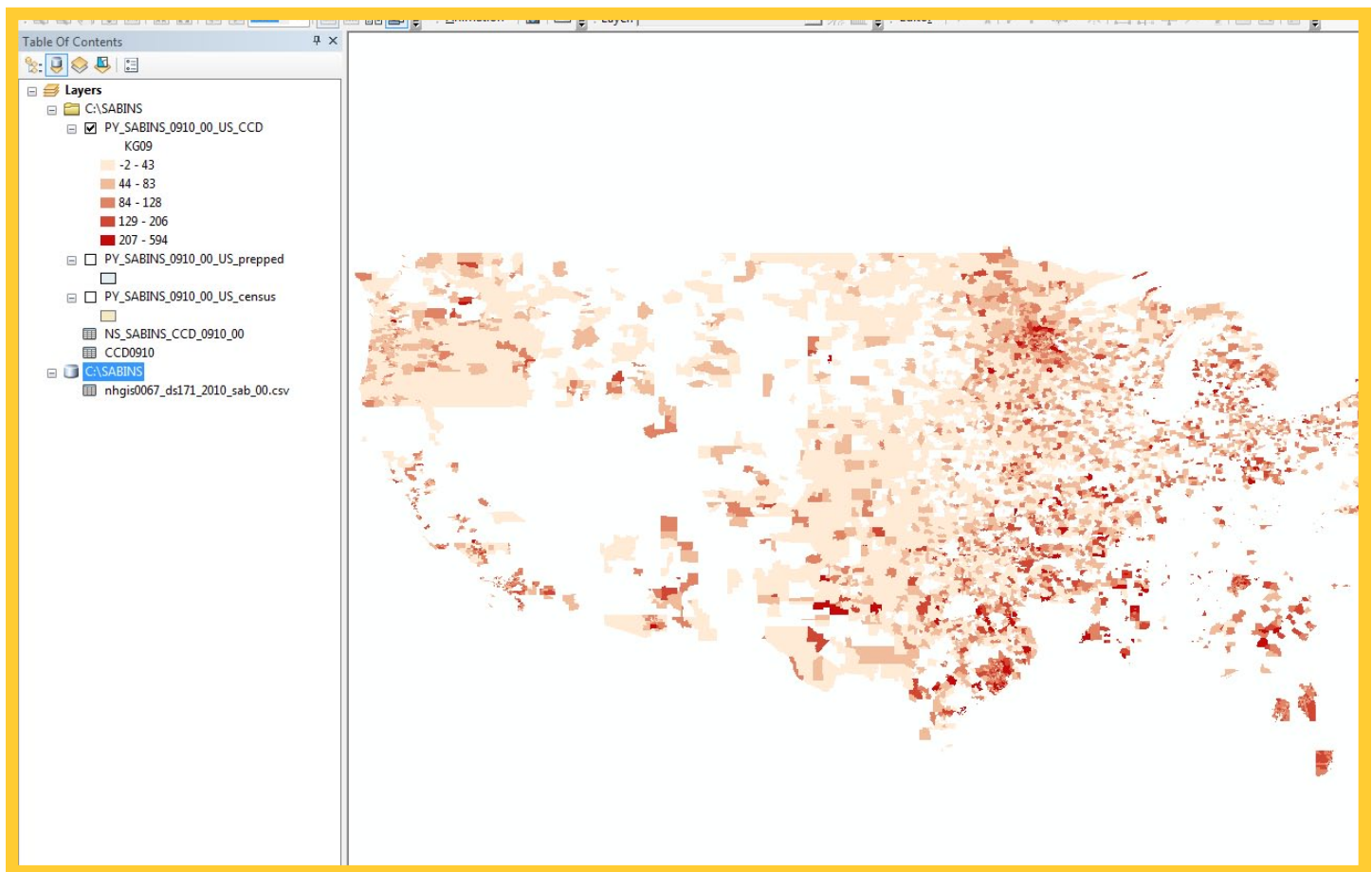
Join CCD Data to SABINS Shapefile

Finally, export the joined shapefile. To export, right click **PY_SABINS_0910_00_US_census.shp** in the table of contents, and select Data>Export Data...

Choose to export the joined data as a shapefile.

For the purposes of this tutorial, the exported file is called **PY_SABINS_0910_00_US_CCD.shp**. Add the new shapefile to the map when prompted.

With this final shapefile, we can now map school specific attributes through each school attendance boundary. In the example below, darker shades of red represent schools with a greater number of enrolled kindergarten students.



Option B: Joining SABINS Polygon and/or Census Data to CCD School-level Data

Link CCD to Association Table

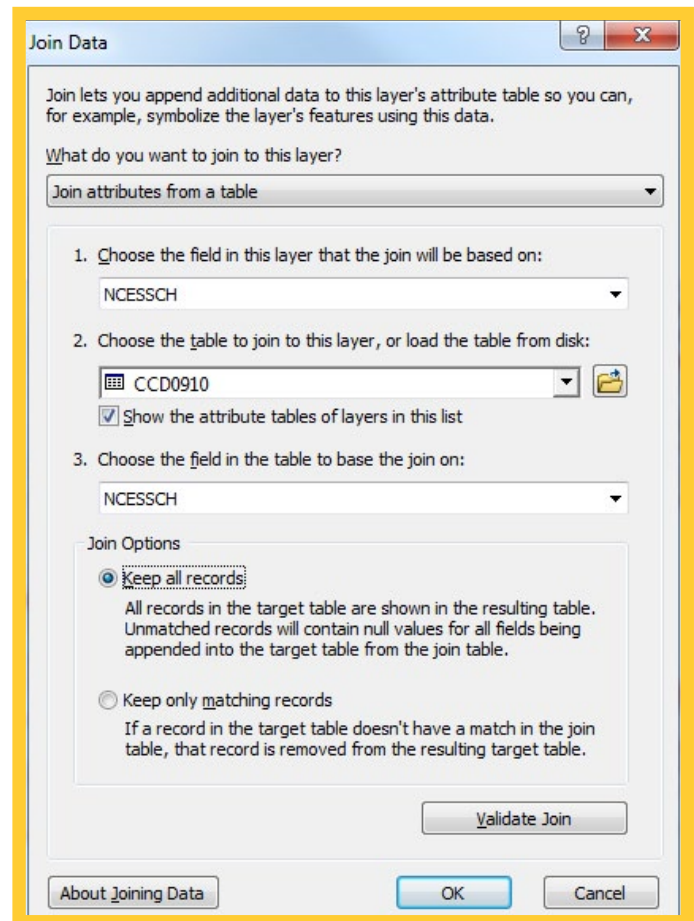
The SABINS shapefiles can be joined to the CCD School-level data through the use of an association table, as the shapefiles do not contain the unique key used to identify schools in CCD.

Add the following files to your map document:

- a. **PY_SABINS_0910_00_US_census.shp** – Kindergarten school attendance boundaries with census data attached.
- b. **NS_SABINS_CCD_0910_00.dbf** – Crosswalk table that connects school attendance boundaries to school in the CCD.
- c. **CCD0910.dbf** – 2009-2010 Common Core of Data file.

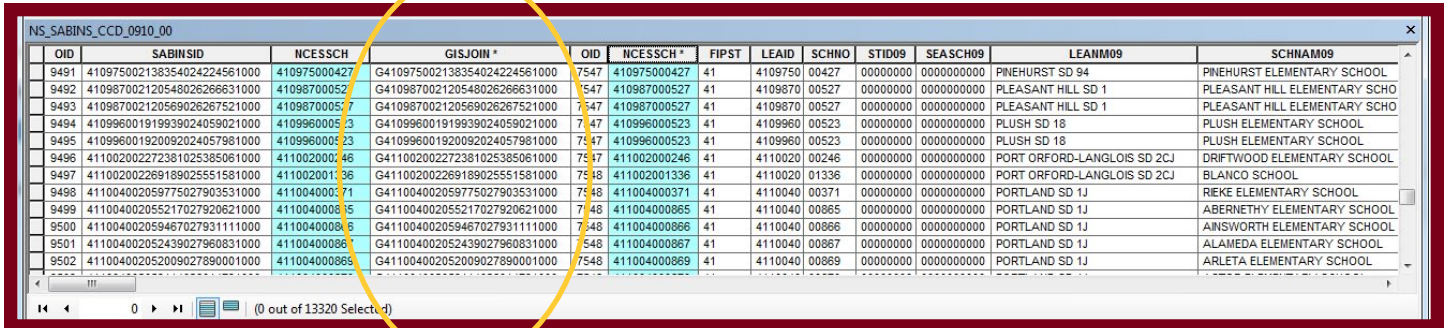
Join the CCD table **CCD0910.dbf** to the association table by right clicking **NS_SABINS_CCD_0910_00.dbf** in the table of contents. Then select Joins and Relates >Join...

- a. Under “What do you want to join to this layer?” select **Join attributes from a table**.
- b. For both the numbers 1 and 3 (the field the join will be based on), select **NCESSCH**. Select **CCD0910.dbf** under number 2, “Choose the table to join to this layer, or load the table from disk:”
- c. Make sure the radio button next to “Keep all records” is selected.
- d. Press OK.



Link CCD to Association Table

You may view the joined table by right clicking **NS_SABINS_CCD_0910_00.dbf** in the table of contents and selecting Open. The figure below shows the tables that have been joined, with the critical GISJOIN field now included.



OID	SABINSID	NCESSCH	GISJOIN *	OID	NCESSCH *	FIPST	LEAID	SCHNO	STID09	SEASCH09	LEANM09	SCHNAM09
9491	410975002138354024224561000	410975000427	G410975002138354024224561000	7547	410975000427	41	4109750	00427	00000000	0000000000	PINEHURST SD 94	PINEHURST ELEMENTARY SCHOOL
9492	410987002120548026266631000	410987000527	G410987002120548026266631000	547	410987000527	41	4109870	00527	00000000	0000000000	PLEASANT HILL SD 1	PLEASANT HILL ELEMENTARY SCHO
9493	410987002120569026267521000	410987000527	G410987002120569026267521000	547	410987000527	41	4109870	00527	00000000	0000000000	PLEASANT HILL SD 1	PLEASANT HILL ELEMENTARY SCHO
9494	410996001919939024059021000	410996000523	G410996001919939024059021000	747	410996000523	41	4109960	00523	00000000	0000000000	PLUSH SD 18	PLUSH ELEMENTARY SCHOOL
9495	410996001920092024057981000	410996000523	G410996001920092024057981000	747	410996000523	41	4109960	00523	00000000	0000000000	PLUSH SD 18	PLUSH ELEMENTARY SCHOOL
9496	411002002272381025385061000	411002000246	G411002002272381025385061000	7547	411002000246	41	4110020	00246	00000000	0000000000	PORT ORFORD-LANGLOIS SD 2CJ	DRIFTWOOD ELEMENTARY SCHOOL
9497	411002002269189025551581000	411002000136	G411002002269189025551581000	7548	411002000136	41	4110020	01336	00000000	0000000000	PORT ORFORD-LANGLOIS SD 2CJ	BLANCO SCHOOL
9498	411004002059775027903531000	411004000371	G411004002059775027903531000	7548	411004000371	41	4110040	00371	00000000	0000000000	PORTLAND SD 1J	RIEKE ELEMENTARY SCHOOL
9499	411004002055217027920621000	411004000865	G411004002055217027920621000	748	411004000865	41	4110040	00865	00000000	0000000000	PORTLAND SD 1J	ABERNETHY ELEMENTARY SCHOOL
9500	411004002059467027931111000	411004000866	G411004002059467027931111000	748	411004000866	41	4110040	00866	00000000	0000000000	PORTLAND SD 1J	AINSWORTH ELEMENTARY SCHOOL
9501	411004002052439027960831000	411004000867	G411004002052439027960831000	548	411004000867	41	4110040	00867	00000000	0000000000	PORTLAND SD 1J	ALAMEDA ELEMENTARY SCHOOL
9502	411004002052009027890001000	411004000869	G411004002052009027890001000	7548	411004000869	41	4110040	00869	00000000	0000000000	PORTLAND SD 1J	ARLETA ELEMENTARY SCHOOL

Next, export the joined .dbf file. To export, right click **NS_SABINS_CCD_0910_00.dbf** in the table of contents, and select Data>Export...

Choose to export the joined data as a dBASE file.

For the purposes of this tutorial, the exported file is called **NS_SABINS_CCD_0910_00_schools.dbf**. Add the new file to the map document when prompted.

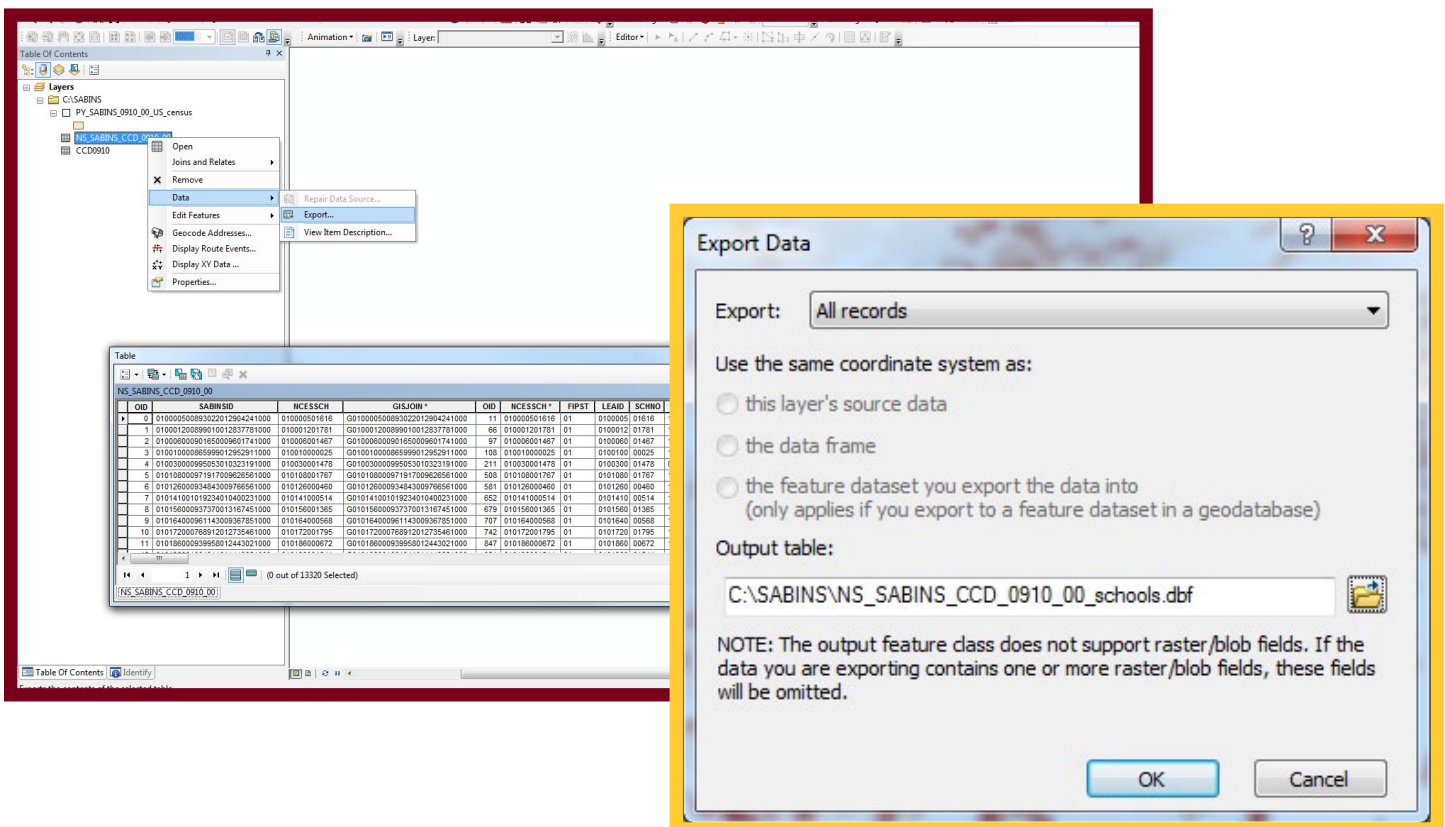


Table of Contents

- Layers
 - NS_SABINS_CCD_0910_00
 - NS_SABINS_CCD_0910_00_US_census
 - NS_SABINS_CCD_0910_00
 - CC0910

Export Data

Export: All records

Use the same coordinate system as:

- ☐ this layer's source data
- ☐ the data frame
- ☐ the feature dataset you export the data into (only applies if you export to a geodatabase)

Output table:

C:\SABINS\NS_SABINS_CCD_0910_00_schools.dbf

NOTE: The output feature class does not support raster/blob fields. If the data you are exporting contains one or more raster/blob fields, these fields will be omitted.

OK Cancel

Joining SABINS Data to Schools

With the school data now containing the GISJOIN field, it can be successfully joined to the SABINS shapefile which also contains the GISJOIN field.

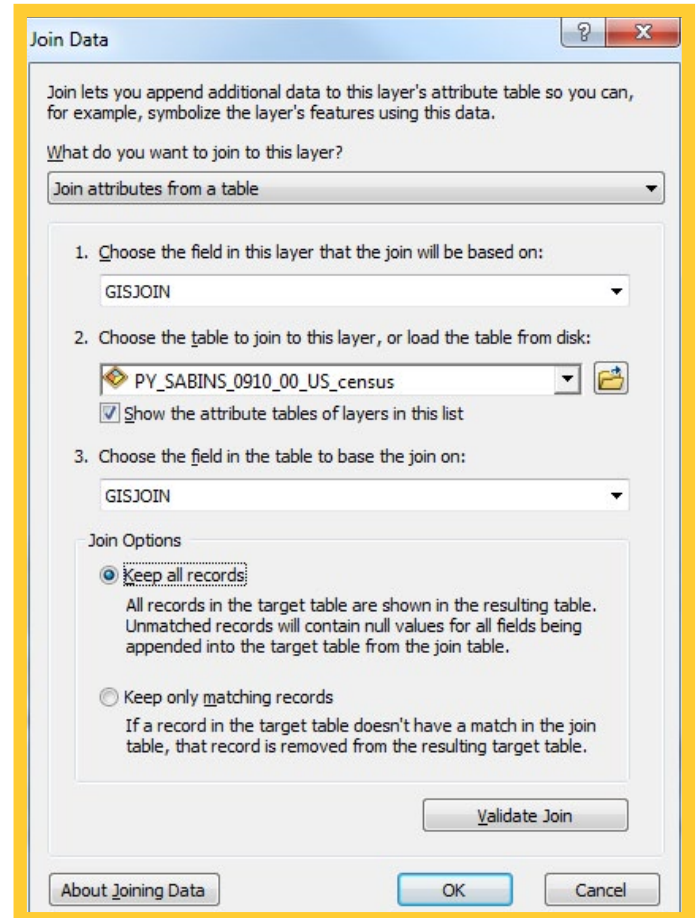
Join the SABINS shapefile (**PY_SABINS_0910_00_census.shp**) to the school data created in last step. Right click on **NS_SABINS_CCD_0910_00_schools.dbf** in the table of contents. Then select Joins and Relates>Join...

a. Under "What do you want to join to this layer?" select **Join attributes from a table**.

b. For both the numbers 1 and 3 (the field the join will be based on), select **GISJOIN**. Select **PY_SABINS_0910_00_census.shp** under number 2, "Choose the table to join to this layer, or load the table from disk:"

c. Make sure the radio button next to "Keep all records" is selected.

d. Press OK.



You may view the joined table by right clicking **NS_SABINS_CCD_0910_00_schools.dbf** in the table of contents and selecting Open. The figure below shows the tables that have been joined, with the SABINSID field now attached to each individual school record. Notice that the multiple schools that share a single SABINSID are now all visible.

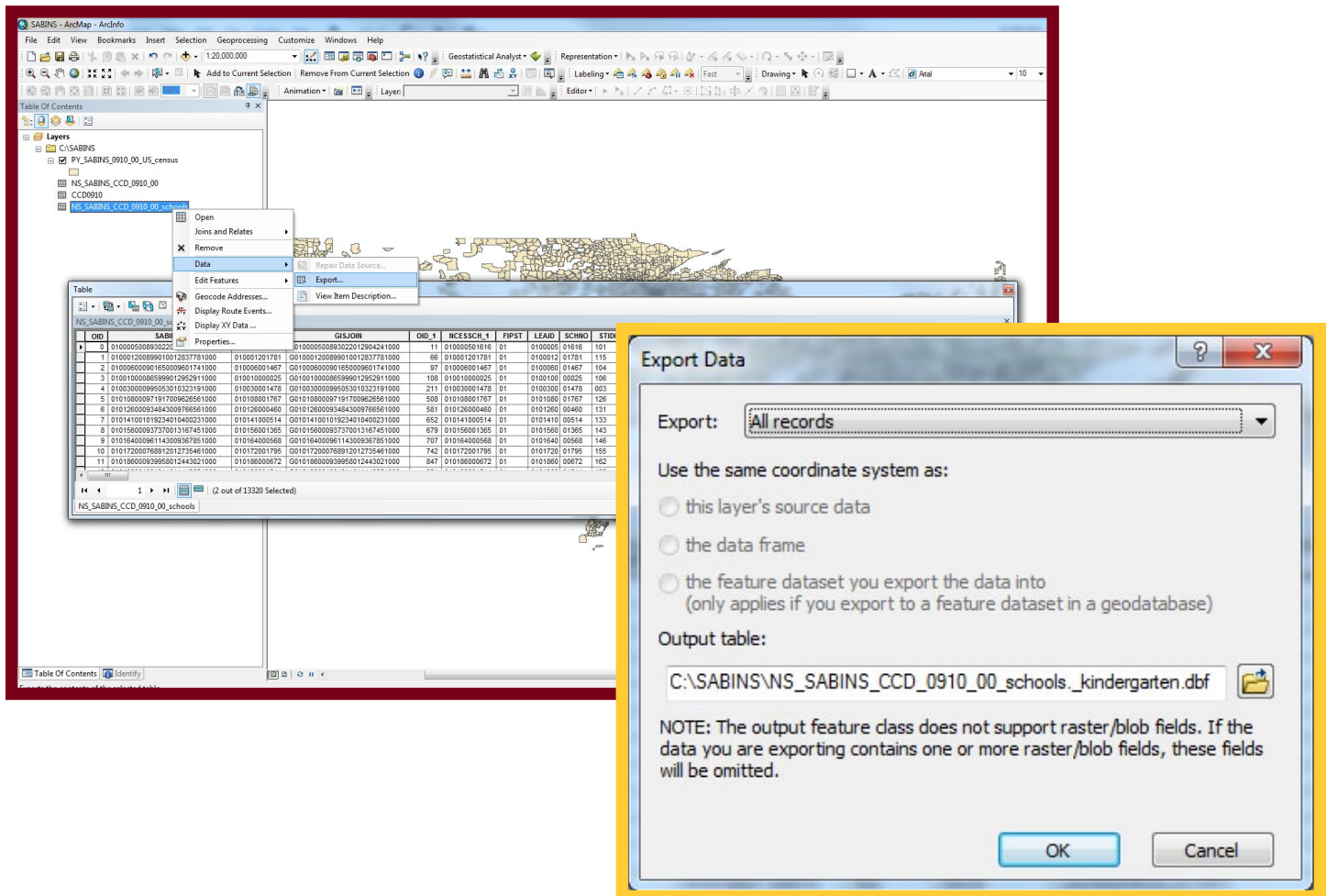
OID	SABINSID	NCESSCH	GISJOIN	OID_1	NCESSCH_1	FIPST	LEAID	SCHNO	STID09	SEASCH09	LEANN09	SCHNAM09	PHONE09
24	020000503626200056729851000	020000500582	G020000503626200056729851000	1670	020000500582	02	0200005	00582	55	550010	KASHUNAMUT SCHOOL DISTRICT	CHEVAK SCHOOL	9078587
25	0200003003593849051196211000	020000300626	G0200003003593849051196211000	1697	020000300626	02	0200030	00626	08	080020	BRISTOL BAY BOROUGH SCHOOL DISTRICT	NAKNEK ELEMENTARY	9072464
26	0200006002903928048839431000	020000600627	G0200006002903928048839431000	1707	020000600627	02	0200060	00627	12	120020	CORDOVA CITY SCHOOL DISTRICT	MT. ECCLES ELEMENTARY	9074243
27	0200009002414036040659361000	020000900160	G0200009002414036040659361000	1717	020000900160	02	0200090	00160	13	130010	CRAIG CITY SCHOOL DISTRICT	PAGE CORRESPONDENCE	9078263
28	0200009002414036040659361000	020000900160	G0200009002414036040659361000	1717	020000900160	02	0200090	00160	13	130020	CRAIG CITY SCHOOL DISTRICT	SHARPLEY ELEMENTARY	9078425
29	0200009002414036040659361000	020000900160	G0200009002414036040659361000	1717	020000900160	02	0200090	00160	13	130020	CRAIG CITY SCHOOL DISTRICT	SHARPLEY ELEMENTARY	9078425
30	0200024002477258042845411000	020002400135	G0200024002477258042845411000	1856	020002400135	02	0200240	00135	42	420010	SITKA SCHOOL DISTRICT	BARANOF ELEMENTARY	9077475
31	0200024002477258042845411000	020002400142	G0200024002477258042845411000	1858	020002400142	02	0200240	00142	42	420010	SITKA SCHOOL DISTRICT	SITKA CORRESPONDENCE	9077477
32	0200030002445730043763731000	020003000146	G0200030002445730043763731000	1865	020003000146	02	0200300	00146	19	190010	HOONAH CITY SCHOOL DISTRICT	HOONAH ELEMENTARY	9078452
33	0200030002445730043763731000	020003000148	G0200030002445730043763731000	1867	020003000148	02	0200300	00148	20	200010	HYDABURG CITY SCHOOL DISTRICT	HYDABURG SCHOOL	9072853
34	0200036002403509042297691000	020003600150	G0200036002403509042297691000	1868	020003600150	02	0200360	00150	23	230010	KAKE CITY SCHOOL DISTRICT	KAKE ELEMENTARY & HIGH SCHOOL	9077853
35	0200045002409428040715171000	020004500179	G0200045002409428040715171000	1913	020004500179	02	0200450	00179	27	270010	KLAWOCK CITY SCHOOL DISTRICT	KLAWOCK CITY SCHOOL	9077552

Joining SABINS Data to Schools

Finally, export the joined .dbf file. To export, right click **NS_SABINS_CCD_0910_00_schools.dbf** in the table of contents, and select Data>Export...

Choose to export the joined data as a dBASE file.

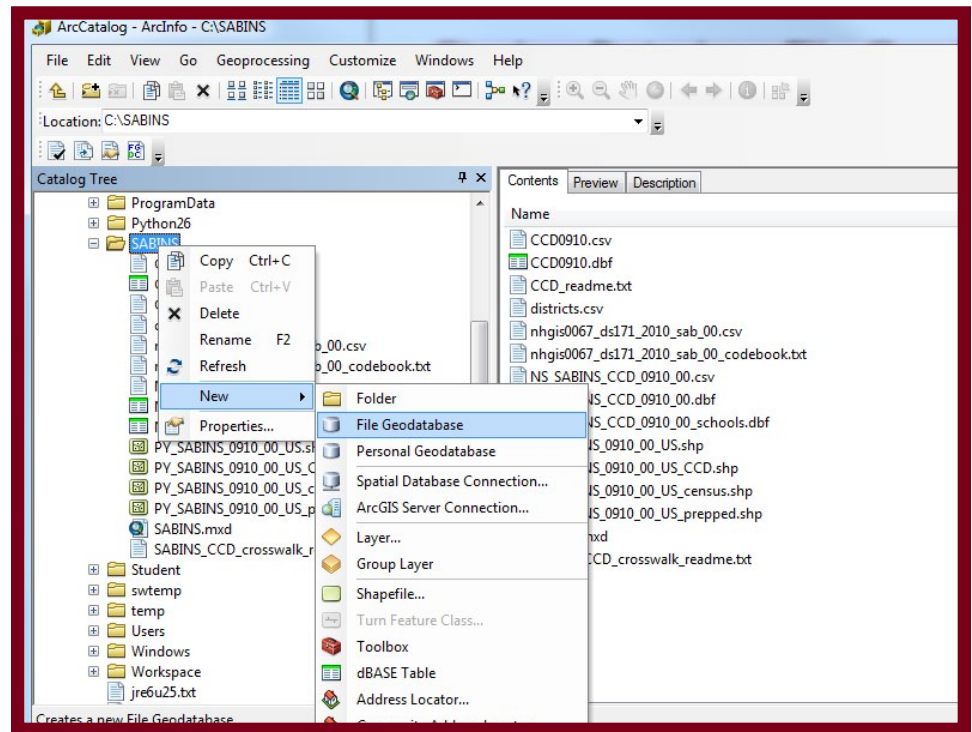
For the purposes of this tutorial, the exported file is called **NS_SABINS_CCD_0910_00_schools_kindergarten.dbf**. Add the new file to the map document when prompted.



Storing Data in a File Geodatabase

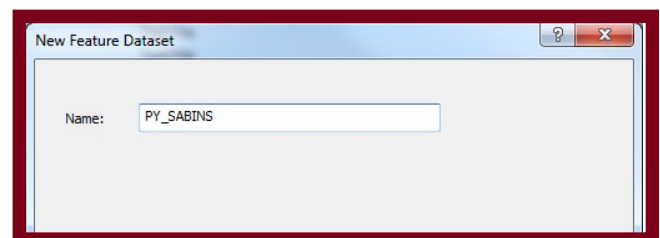
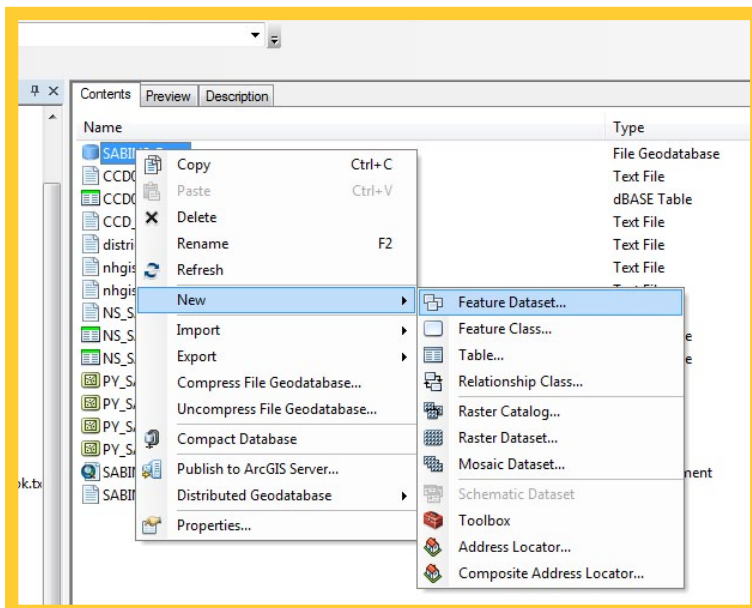
Migrating your data to a file geodatabase is beneficial for a number of reasons. It is easier to keep track of the various files, and you may experience faster processing speed with the exceptionally large files that SABINS data users encounter.

Start ArcCatalog. In the Catalog Tree, right click the folder where you plan to store SABINS data, and select New>File Geodatabase. Name the geodatabase as you see fit, in this example it is called **SABINS_Data.gdb**



Create a new feature dataset by right clicking the new geodatabase (**SABINS_Data.gdb**) and select New>Feature Dataset...

Then, in the first dialogue box labeled "New Feature Dataset," name your feature dataset **PY_SABINS**. Click Next>.



Storing Data in a File Geodatabase

In the next dialogue box, select a coordinate system by clicking **Projected Coordinate Systems>Continental>North America>USA Contiguous Albers Equal Area Conic**.

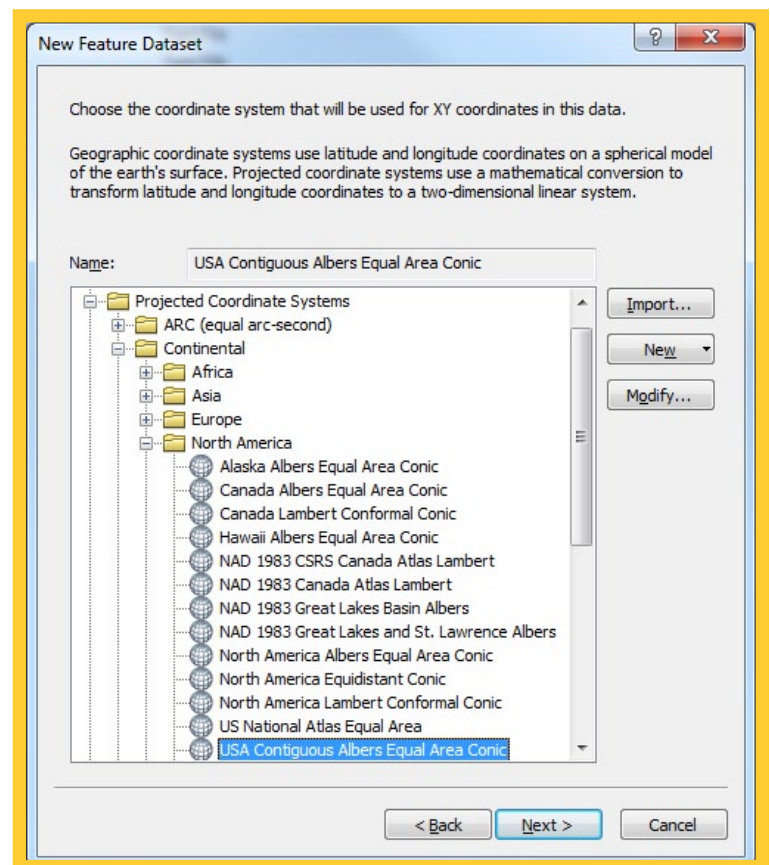
Press Next >.

The next dialogue box will ask you to "Choose the coordinate system that will be used for Z coordinates in this data." Leave this blank by selecting nothing.

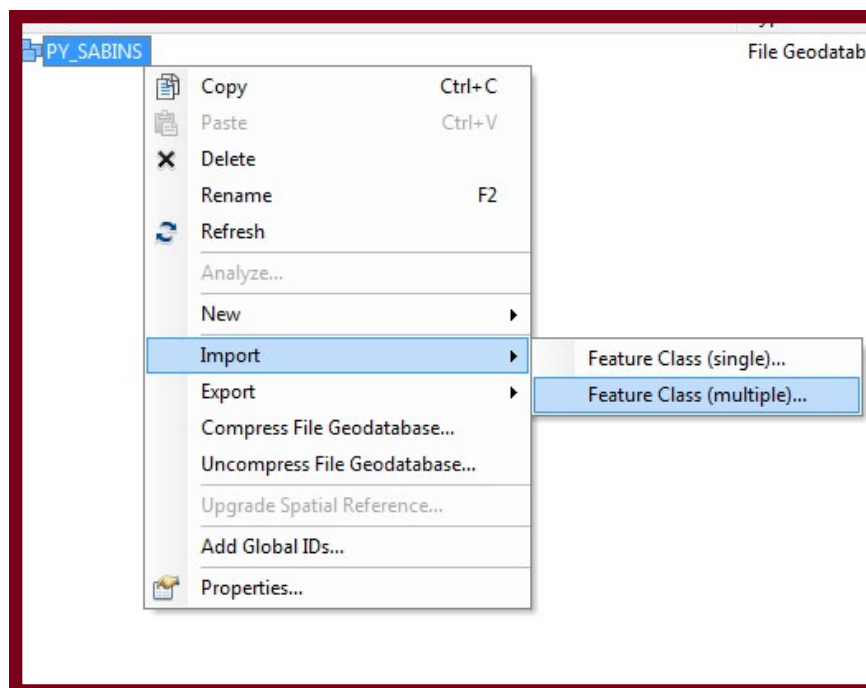
Press Next >.

The next dialogue box will ask you to set the XY Tolerance. Be sure the box is checked next to "Accept default resolution and domain extent (recommend)."

Press Finish.



In order to import data, right click your new feature dataset (**PY_SABINS**) and select Import>Feature Class (multiple)...



A new dialogue box will appear labeled "Feature Class to Geodatabase (multiple)".

Press the folder icon next to the box labeled "Input Features," and navigate to the location of your shapefiles and data tables. Select the ones you would like to add to the geodatabase, and click Add. When importing multiple feature classes at once, ArcCatalog will automatically assign each feature class the same name as the input features.

Finish by pressing OK. The actual import may take a minute or so to complete.