

Lab 3: Working with Relational Database Management Systems

Purpose of the lab: To introduce the principles of interacting with attribute data stored in a GIS-based relational database management system.

1. Get Your Data

Using The Catalog Pan or Catalog View, copy the following data layers and files from the Cities201_Data folder over to your workspace:

Tiger\tracts.shp (for the lab assignment, you will need this file in UTM projection. If your tracts_utm.shp file is in good shape from Lab 2, you can re-use that file. If you are not sure, start fresh and reproject to UTM)

Tiger\sف1_tracts.dbf

Tiger\sف3_tracts.dbf

Using your computer's file management application, copy the following files from the Cities201_Data folder over to your Lab folder workspace (these are just normal PDF files, so you don't need to use The Catalog Pan or Catalog View for this):

Tiger sf1.pdf

Tiger sf3.pdf

Note: these might show up in your file management system as Google Chrome files.

2. Explore Your Data

You are by now familiar with the tracts shapefile. The other two data files, sf1_tracts and sf3_tracts, encode data from the **2000 U.S. Census**. These data are stored as tables; they are attribute data without the attached spatial data.

The sf1_tracts file contains data from *summary file 1*, which contains data gathered from questions asked of all people. The sf3_tracts file contains data from *summary file 3*, which contains data gathered from questions given to a sample of the population.

Open the sf1.pdf and sf3.pdf files. These are documentation files in Adobe Acrobat format for each of the respective summary file 1 and 3 data files. They describe the data and tell you about each of the variables contained in the data files. To learn more about the data in general, click on 'Chapter 1: Abstract' in the Bookmarks frame on the left side of the Acrobat Reader window and read 'Subject Content'.

Preview sf1_tracts and sf3_tracts in The Catalog Pan or Catalog View. Note that each **record** (meaning row) represents an individual tract. The field STFID is a unique identifier for each tract. Note that the tracts shapefile has the same field. Each field beginning with a 'P' is a variable, such as total population (sf1_tracts: P001001).

You can find out what each variable in the data file is by searching for it in the appropriate PDF document. For example, preview the sf1_tracts file and find the variable P015001. To find out what this variable is, go to the sf1.pdf document and press ctrl and F at the same time to bring up the search. Enter the name of the field and hit enter. The entry will be highlighted in the document, indicating that this field encodes the total number of households.

Search for a few other variables to get comfortable with the method. Remember to use the appropriate .pdf document for the appropriate data file (i.e. use the sf1.pdf to look up variables in the sf1_tracts table).

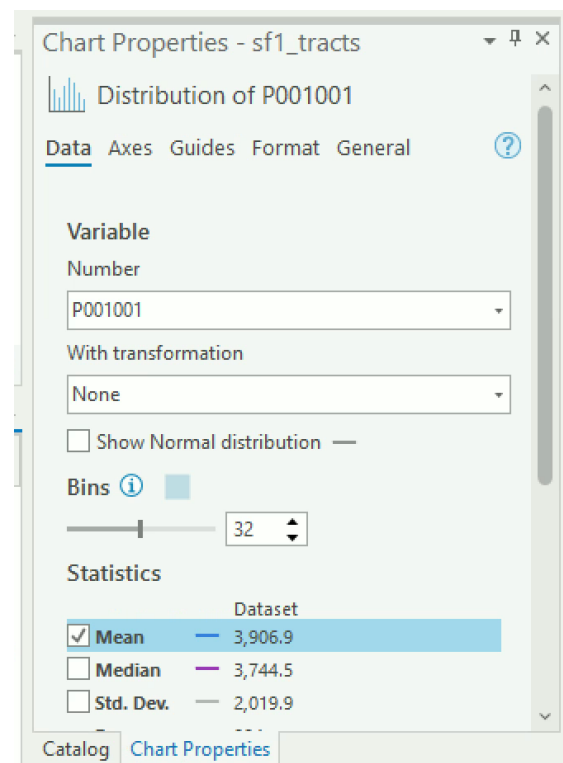
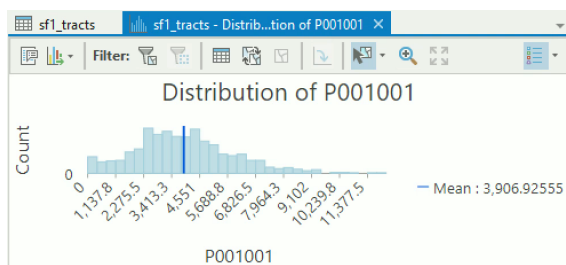
3. Manipulating the Display of Tables

Get to a Map view in ArcGIS and drag and drop the sf1_tracts file into it. Note that the file is not displayed in the view window as there is no spatial data, but it shows up in the Contents Pane.

Open the attribute table for sf1_tracts.

Find the field 'P001001'. This field encodes the total population of each tract. Right click on name of the field. A drop down menu appears. Choose 'Sort Ascending' to reorder the table from lowest total population to highest. Choose 'Sort Descending' to reorder the table from highest to lowest total population.

Right click and choose Statistics for a summary of the field, including a histogram of the distribution and summary statistics such as the mean and standard deviation.



4. Selection and Export

The selection operation is a fundamental relational operation that you are already familiar with. You can select a single record manually by simply clicking on that record on the far left of the table window. To manually select multiple records, hold the Ctrl key down while selecting.

Select a few records manually. Note that at the bottom of the table the number of selected and total records is reported. Next to this number is the 'Show selected records' button – press it now. Now only the selected records are displayed. Press the 'Show all records' button to the left to view all the records.

Other ways of selecting are available by pressing the Table Options button at the top left of the table. Choose Select All to select all the records. To remove your selection, choose Clear Selection.

To do an attribute selection, choose Select by Attributes (from the Table Options or from the button at the top of the table). In the window, choose "P001001" <4000. 544 out of 994 records should be selected. Choose **Table Options > Switch Selection** to reverse the selection. Now those 450 other records with total population equal to or greater than 4001 should be selected.



You can save just the selected records to a new table. Go to Table Options > Export and for Output Location, send the Output Table to your workspace. For output name, use the extension .dbf (that's the extension for Dbase tables). Press OK. You should see the new table in Contents Pane. Open the new table. It should contain only the 450 records representing tracts with population over 4000.

Close your Select by Attributes window and the sf1_tracts table, leaving only the new exported table open.

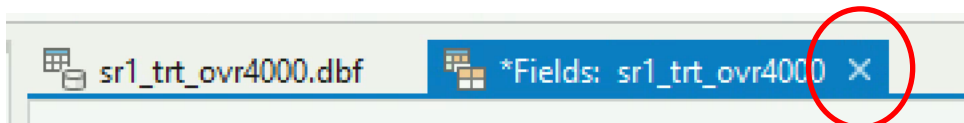
5. Creating and Editing Fields

It is possible to add a new field to the table and calculate new values for that field. Here we will create a new field encoding population density by dividing the total population of each tract (encoded in the P001001 field) by the land area of each tract (encoded in the Arealand field).

Go to Table Options and choose Add Field.



For the Name text box enter 'PopDensity'. For the Type dropdown menu, choose 'Float'. Right click and then save. After the process has run, close the table editing window.



Scroll to the far right of the table and see that a new field named 'PopDensity' has been created. Right click on the PopDensity field name and choose 'Calculate Field' (press the Yes button if prompted with a warning).

In the text entry area (where it says PopDensity =), enter the expression '[P001001] / [AREALAND]' by clicking on the appropriate field names and operator button. Press the OK button.

The PopDensity field should now encode the population density in people per square meter for each tract.

Delete the PopDensity field by right clicking on the name of the field and choosing 'Delete Field'.

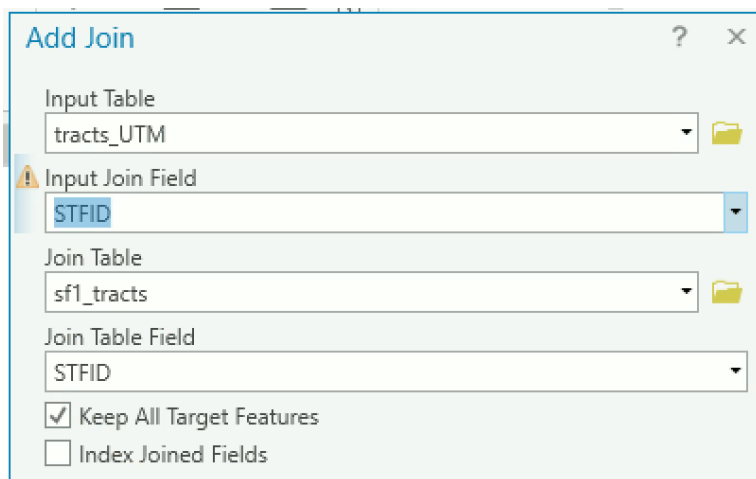
Close and remove the exported table from your map. You won't need this table again—it was just for practice.

6. Joining Tables

Here, we will **join** tables together. The sf1_tracts table will be joined onto the tracts shapefile table in order to associate the census data encoded in the sf1_tracts table with individual tract polygons encoded in the tracts shapefile.

Add the tracts shapefile to your map and open its attribute table. Note that there are only a few **fields** (meaning columns). Close the attribute table.

In the Contents Pane, right click on tracts, choose 'Joins and Relates', and then choose 'Add Join'.



For number Input Join Field, choose 'STFID'. This is the field in the tracts shapefile table used for the join.

For number Join Table, choose 'sf1_tracts'. This is the table to join onto the tracts shapefile table.

For number Join Table Field, choose 'STFID'. This is the field in the sf1_tracts table used for the join.

'Keep All Target Features' should be selected in the Join Options window.

Press the OK button (click Yes on the pop up window).

Open the tracts shapefile attribute table. Note that the entire sf1_tracts table has been appended onto the tracts shapefile table.

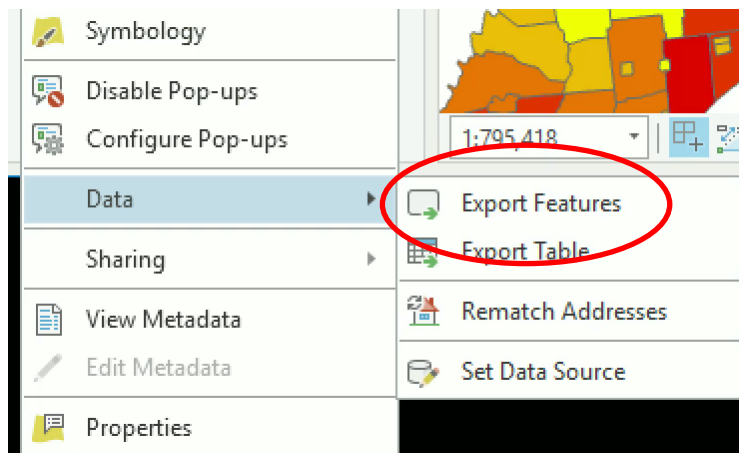
**** If your map disappears when the join is complete, change the symbology to graduated colors, and it should re-appear. You can now change the symbology back to single symbol if you want. ****

5. Create a New Shapefile from Selected Records

The US Census has designated census tracts for all territory of the United States, including parks and industrial land. This means that there are some tracts in Philadelphia that have zero population. A tract with zero population is very different than a tract with low population, so it is typical to designate zero population tracts on a map, often with a grey shade. One way to do that is to create a shapefile with **only** zero population tracts.

Using your tracts shapefile that you joined the sf1 table to, select all of the tracts with 0 population (either use the Select by Attributes operation or sort your attribute table by population in the attribute table and manually select the records with 0 population). 9 tracts should be selected.

Right click on the tracts layer in your map table of contents, then got to Data→Export Features. In the next dialogue box, make sure you save as a Shapefile in the location you want with the name you want.



The new shapefile will be added to your map. Slide it to the top of the map layers, and color the tracts a shade of grey. **Every time you make a map with census data, you should either use the Exclusions feature in the Symbology Pane or you should add this shapefile to your map so that census tracts with zero population show differently than your graduated colors.**

8. More Exporting Shapefiles

Do another Select by Attributes on the tracts shapefile. For your selection criteria, select all those tracts with water cover, where the value of the AREAWATR field > 0 (AREAWATR is the square meters of water cover).

273 tracts should be selected.

Once the appropriate tracts have been selected, right click the tracts shapefile in the Table of Contents and choose Data, then choose Export Data. Next to Export choose 'Selected Features' and send the new shapefile you are creating to your workspace. Click the Yes button to add the exported data to ArcMap.

Turn off the tracts shapefile to see that the new shapefile encodes only those tracts with water in them.

ASSIGNMENT

Lab #3-6: Must do 1 Full Lab Reports & 3 Slim Lab Reports – your choice which is which

Objective

The objective of this assignment is to explore the distribution of poverty rate in southeast Pennsylvania as it relates to school districts.

Deliverables

Turn in a report addressing this objective. Refer back to your Lab Report Instructions for direction about the key sections of your report—introduction, methods, results, discussion, tables and figures, and summary of operations. Your report must include a tract-level choropleth map of the poverty rate (the percentage of persons living below the poverty line) in southeast Pennsylvania. This tract-level map should also display school district boundaries. Moreover, you should report the STFIDs of the tracts with the five highest poverty rates. Also, report the school districts that intersect these five tracts and include a map that identifies them.

Getting Started

First, think through the analysis steps for this task. Which data layers or tables will you need? [Make sure all of your files are in UTM projection.](#)

Do you have all of the variables you need, or will you have to create new variables?

Hint: The **sf3_tracts** file has information on poverty rate. The field P087001 encodes the number of people for whom poverty status was checked. The field P087002 encodes the number of people with income below the poverty level. Thus, the poverty rate for each tract can be calculated as:

$$\text{P087002 (i.e. number of people with income below poverty line) divided by P087001 (i.e. number of people for whom poverty status was checked) or (P087002/P087001).}$$

Another hint: Remember algebra - is it possible to divide by zero? In order to calculate a poverty rate, you must first select those tracts for which the poverty level has been checked for at least 1 person ($\text{P087001} > 0$)—you should get 980 records out of 994 total—then continue with your calculation.

Finally, when you consider the limitations of the analysis with this data, consider if there is anything puzzling about the data.

Lab 3 Summary of Operations

Operation	Briefly describe what this operation does	List the steps to complete this operation
Export data to new shapefile		
Table: add field		
Table: calculate field		
Table: export selected records		
Table: join to shapefile		

Lab 3 will be graded with the following criteria (see General Lab Report instructions for more detail):

FULL REPORT

	Pass	Good	Excellent
Objective			
Clear statement of GIS analysis objective			
Analysis Steps			
Correct steps & understanding of GIS operations			
Results			
Clear summary of results that describes the patterns seen			
Discussion			
Limitations			
Extensions			
Decisionmaker applications			
Tables & Figures			
Map has required analysis features			
Maps are in UTM projection			
Map layout displays qualities identified in Map Design Checklist			
Poverty rate info is reported as required			
GIS Operations Summary			
Table is completed			
Other Notes			
Notes (see tracked changes below for more)			
Grade			

SLIM REPORT

	Pass	Good	Excellent
Analysis Steps			
Correct steps & understanding of GIS operations			
Tables & Figures			
Map has required analysis features			
Maps are in UTM projection			
Map design: overall composition			
Map design: required elements			
Map design: data classes (for map 3)			
Map design: colors and line weights			
Poverty rate info is reported as required			
GIS Operations Summary			
Table is completed			
Other Notes			
Notes (see tracked changes below for more)			
Grade			