



Geography

Steve Graves
Forensic Geography
Lab: Joining Data from Census to GIS

Background

Now that you have some experience downloading Census Data from the internet, you must now learn to make it appear on a map so that it can be analyzed.

Remember, many of the popular criminology theories assume that neighborhood variables are critical in our ability to understand or predict criminal activity.

Skills

You will demonstrate two software techniques this week, both of which are crucial to master and understand if you are to work independently as a GIS or crime analyst.

1. You will demonstrate an ability to recognize problems with US Census data.
 2. You will demonstrate an ability to correct a specific, common data problem with US Census Data, using the concatenate tool in Microsoft Excel
 3. You will “join” your corrected data file to a map file (.shp format) using ArcMap10 GIS program.
 4. You will use the joined files to make a map in which you identify a specific neighborhood in Los Angeles with a specific set of housing characteristics.
-

Your Task

This week’s assignment is reasonably simple.

1. You will be provided a data file in .csv (comma separated values) format. It was downloaded directly from the US Census Bureau. The file contains “general housing characteristics” from the 2010 US Census at the tract level for Los Angeles County.
 - a. File Name: LA_CNTY_TRACTS_HOUSING_2010.CSV
(Y:/COURSES/COURSES_SGRAVES/FORENSIC)
2. You will be provided a series of map files also downloaded directly from the US Census Bureau. This set of files (5 in total) are “controlled” by a master file that controls the subordinate files.
 - a. File Name: LA_COUNTY_TRACTS_2010.SHP
(Y:/COURSES/COURSES_SGRAVES/LA_COUNTY)
3. You should open the .csv file in MS Excel
4. Widen the data in columns A and B to see the entire cell contents.
5. Note that the data in column A is left justified indicating that data in these cells are functioning as text or string data. This is correct for this sort of “nominal” data.
6. Note that the data in column B is right justified. This indicates the data in these cells are function as numeric data. This is incorrect for this sort of “nominal” data.
 - a. You should highlight a few cells in column B and note that the software will begin to display sums, averages, etc. in the information ribbon across the bottom of the window.

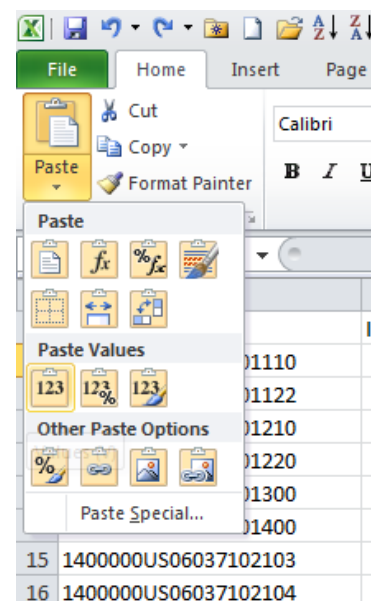
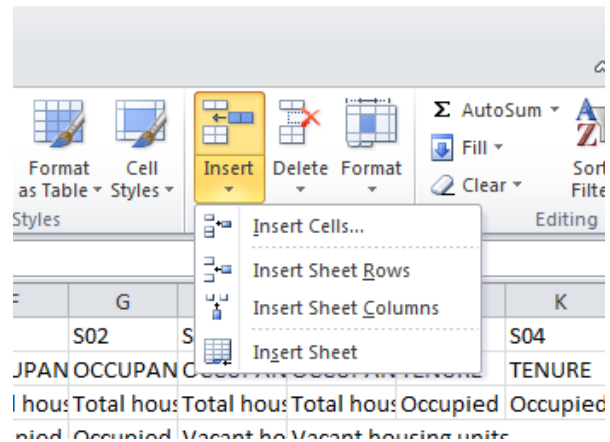
- b. Highlight a few cells in column A. Note that the software does not display anything but a count in the information ribbon across the bottom of the screen.
7. Note also that the first data cell in column B, just under “GEO.id/ID2” is “6037101110”. This is incorrectly displaying. Excel, unable to understand a leading ZERO in the data, inconveniently removes a leading zero from the cell contents. The cell should read “06037101110”
8. There are two methods for fixing this problem; both techniques are exceptionally useful for those who use Excel for managing data files. They are explained below.

Technique 1: Concatenation

Concatenate is a command in Excel (and other similar programs) that allow you to “paste” together information into a single cell. You can paste data from other cells, if you want to vary what is concatenated together, or you can simply insert a letter, number or symbol to any desired cell.

Here’s how:

1. Insert a new column to the right of column B (click on a cell in column C and click on Insert/Insert Sheet Columns)
2. A new column will appear to the right of column B.
3. Type “FIPS” in the cell to the right of cell B8 (has “Id2” in it)
4. In cell C9 (one below FIPS), type the following formula: `=concatenate(“0”, B9)`
5. Press Enter
6. Note that the formula result is now LEFT justified.
7. Copy and paste your result to the rest of the column’s blank cells.
8. Copy and paste the entire column again, but this time when you paste, paste VALUES ONLY.

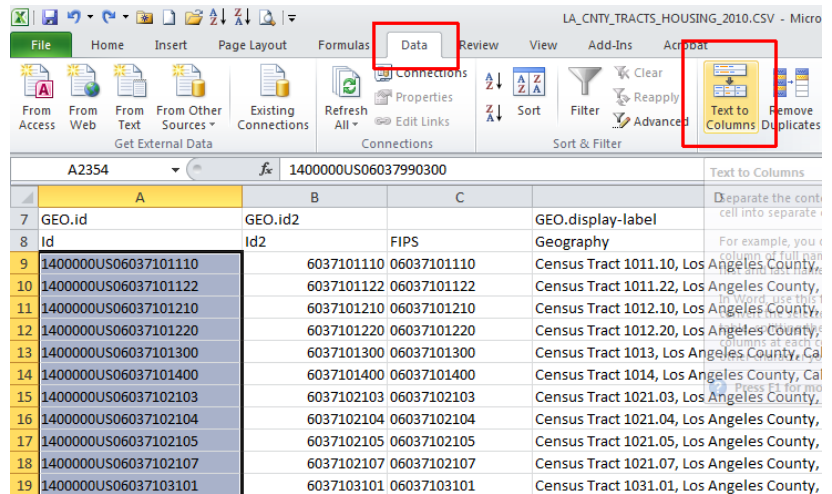


Technique 2: Text to Columns

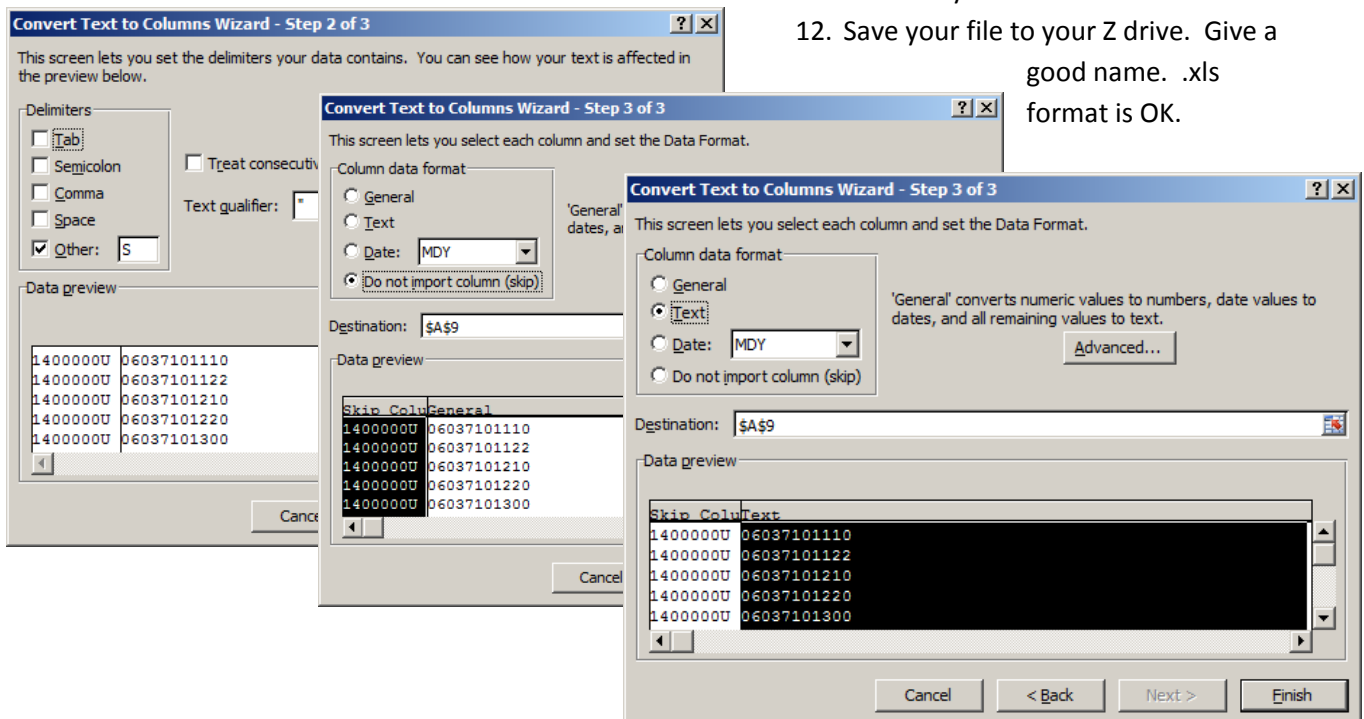
The second technique is used to break a single column of data in to multiple columns of data, or to truncate and discard a column of data. The REAL value of this technique is that it is the only foolproof way to get Excel to treat numeric data as text/string data.

Here's how:

1. Highlight the column of data in column A (Ctrl+shift+↓)
2. Activate the Data tab, to bring up the data toolbar.
3. Click on the Text to Columns tool
4. Select "Delimited" from the step 1 window that appears. Click next
5. De-select Tab and select "Other" from the list of delimiters.
6. Enter the letter S in the small window to the left of the "other" option. Click Next
7. In the step 3 window, click first on "Do not import" to tell Excel not to import the left-most column.
8. Click and highlight the right most or second column (see image)
9. Select "Text" from the data format options
10. Click Finish.

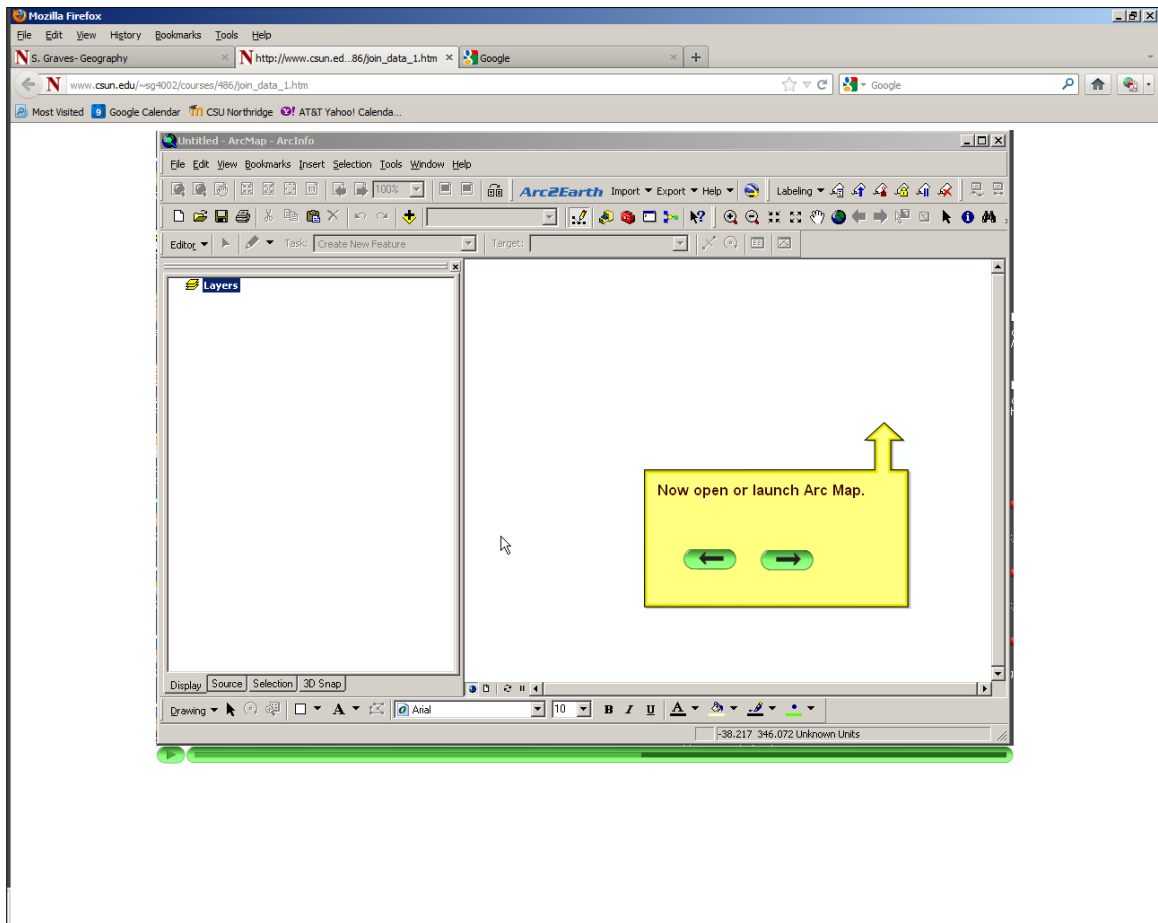


11. Examine your results.
12. Save your file to your Z drive. Give a good name. .xls format is OK.



Part 2. Joining your data to a file in ArcMap GIS.

1. Save your data to your Z drive on the server or a flash drive. You may leave the file as a .csv file, or convert it to an .xls file.
2. Open http://www.csun.edu/~sg4002/courses/486/join_data_1.htm It's a "video" file. You can skip the first part if you'd like and move to just beyond the halfway point to see a video of how to join an Excel file to a Eris shape file. (see image below for start point) Exact instruction for this assignment follow.



3. Open ArcMap 10.
4. Click on the open folder icon on the toolbar. Select blank map or "cancel" if you are prompted.
5. If you're working from within the Sierra Hall lab, look for the Y drive. The Y:/courses/courses_sgraves/ folder has many of the files you'll need. For this lab find the ArcMap "document" file, "LA_2010_tracts_housing" (it has an extension of .mxd that you can't see) = (Y:/COURSES/COURSES_SGRAVES/LA_COUNTY/LA_2010_tracts_housing.mxd)
6. When it opens, there are only 2 map layers in the table of contents:
LA_COUNTY_TRACTS_2010.SHP and LA_AREA_ZIPS.shp
7. Both files are relatively empty of significant data, so you must join the Tracts map to the Excel data that you just saved on your Z drive. Right click on the layer name (LA_COUNTY_TRACTS_2010) and select, "Open Attribute Table" to see what's included in the

tabular data...not much, but the GEOID10 column should be noted, because it is the “join field” to which you will match your Excel data.

8. Click on the little plus sign tool on the tool bar to “add data”.
9. An “add data” window will appear.
10. Find your Z drive folder where you saved your Excel file. If you have not done so, and for some reason the Z drive is not visible, click the “connect to folder” icon and find your Z drive among the possible drives.
11. When you find your housing data file, click to add it. You may be prompted to identify the “sheet” you want to add. It’s probably “sheet1\$” unless you renamed sheet 1 in the spreadsheet. If you saved the file as .csv or .dbf, then you won’t have sheets. Click Add.
12. Now you should see two layers in your Table of Contents window. One map and one datasheet. Not that the Table of Contents has automatically changed to display all layers by SOURCE. It’s not the default....which would hide your xls/database.
13. Let’s join them. Right click on the “LA_COUNTY_TRACTS_2010” layer and select “Joins and Relates” from the drop down menu.
14. By default, the join type should be “Join attributes from a table”.
15. Step 1 “Chose the field..”, select GEOID10 because that is the field (column) that contains unique string/text identifier for each census tract in LA county.
16. Step 2 should be filled out for you and show something like “sheet1\$”.
17. Step 3 may already have some suggested field (column) from your Excel file to base the join on...if you’ve followed the instructions carefully, your FIPS column or your “GEOID” column should work. Chose FIPS.
18. Click OK. It may look like nothing happened, but open the attribute table for the tract map and see how many additional new columns you have!
19. All that data is mappable now, but you have one last recommended step...you must EXPORT this map layer (tracts map) in order to “cement” or make permanent the join. Otherwise, if you opened the file again in another project, the join would be broken.
20. Here’s how (not mandatory): Right click on the LA_COUNTY_TRACTS_2010 layer name.
21. Chose Data/Export Data from the drop down list.
22. Find a folder you want to save these joined files together (as 1), probably on your Z drive.
23. Click OK.
24. You’ll be prompted to add this new “joined” layer to your project. Accept this and then delete the two ‘unjoined’ layers (the tract map and the excel file) by right click/remove of each layer.
25. Make a choropleth / thematic map depicting black, non-Hispanic homeownership in Los Angeles County by 2010 census tracts.
26. Find the tract with the highest rate of black homeownership in LA County. Note the name of the city/ZIP code this tract lies in.
27. Zoom in and get a screen capture “Control+Alt+Print Screen” of this location.
28. Paste your screen capture in an MS Word document along with a caption that includes the name and ZIP code of the specific neighborhood you’ve identified to your instructor. Make sure your last name is in the file name you send to your instructor.