# Introduction to GIS: Lab 3

# University of Massachusetts – Amherst

This week’s lab focuses on working with and querying tables associated with spatial data. These tables hold the all-important data which allows you to make decisions with your analysis. Everything you will learn this week is primarily relevant for shapefiles (vector data) which have attribute tables associated with them. These tables are similar to information you might find in Excel[[1]](#footnote-1), with the added advantage that each row is also linked to a specific spatial location.

The data for this lab is available on the course website.

While the lab production activity is what you will turn in to Moodle for a grade, please note that practice questions throughout to test your own understanding of the concepts are usefully noted in **bold**.

This lab is divided into six parts, which you should complete in order:

**Part A:** Exploring the Attribute Table

**Objective:** Become familiar with the attribute table, how it functions, and the information it stores.

**Part B:** Table Operations: Calculate Geometry and Calculate Field

**Objective:** Use calculation capabilities to determine areas and create new attributes based on existing data.

**Part C:** Joining Tables to Spatial Data

**Objective:** Connect spatial data with attribute data that is not explicitly spatial.

**Part D:** Selecting by Attributes

**Objective:** Use the selection tools to select items from your datasets, while learning some simple SQL functions.

**Part E:** Summarizing and Table Statistics

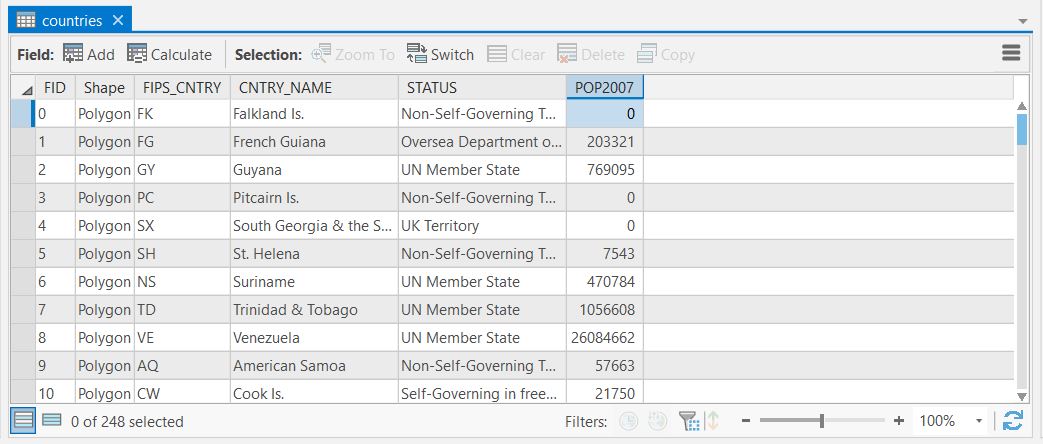
**Objective:** Use simple table manipulations to summarize information in your table and discover simple statistics about the data.

**Part F:** Lab Production: Analyzing Grizzly Bear Habitat

**Objective:** Apply what you have learned and analyze some bear conservation data.

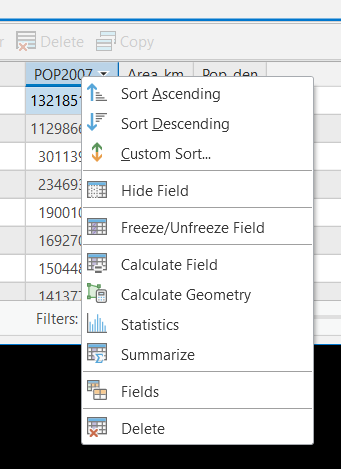
# Part A: Exploring the Attribute Table

* 1. Create a new folder (with an appropriate name) for this lab[[2]](#footnote-2).
  2. Unzip the lab data into your lab folder. Open ArcGIS Pro, add a folder connection to your lab data folder, and add the shapefile ‘countries’. This is a polygon shapefile of world countries as of 2010.
  3. Right click on countries in the Contents pane and select ‘Attribute Table’.
  4. The attribute table associated with this shapefile will pop up. All of the information associated with your shapefile is stored in this table. Everything you can symbolize, label, analyze, etc. is all here.
  5. This table currently contains information about the name of each country, the diplomatic status of the country, and the population of the country in 2007[[3]](#footnote-3) (although there are some zeros representing unknown population counts).

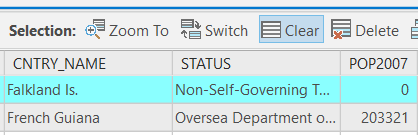


* 1. You can sort each column (as you might also do in excel) by right clicking on the column name and choosing ‘sort ascending’, ‘sort descending’, or ‘custom sort’[[4]](#footnote-4). For now, try sorting the table by population in descending order.

**Practice:** What were the three countries with the highest population in 2007?[[5]](#footnote-5)



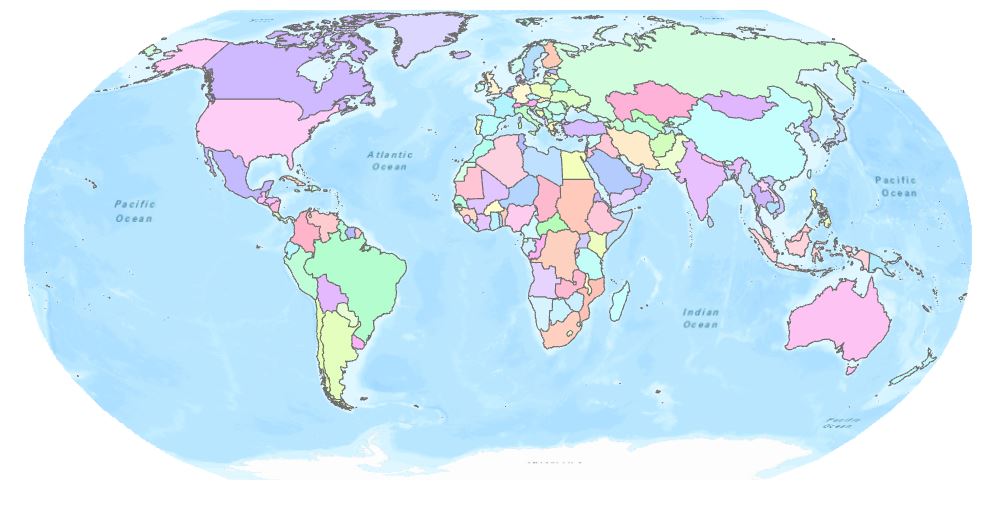
* 1. You can also use the table to select countries of interest. In the table, select the top 10 most populated countries by clicking on the grey squares at the left of the table. You can select multiples by <control>[[6]](#footnote-6) click, or by <shift>[[7]](#footnote-7) click. You’ll see that the selected countries have teal-colored outlines.
  2. You can clear a selection by clicking the ‘clear selection’ button in either the attribute table or in the display under the Map tab > Selection group > clear[[8]](#footnote-8)[[9]](#footnote-9).



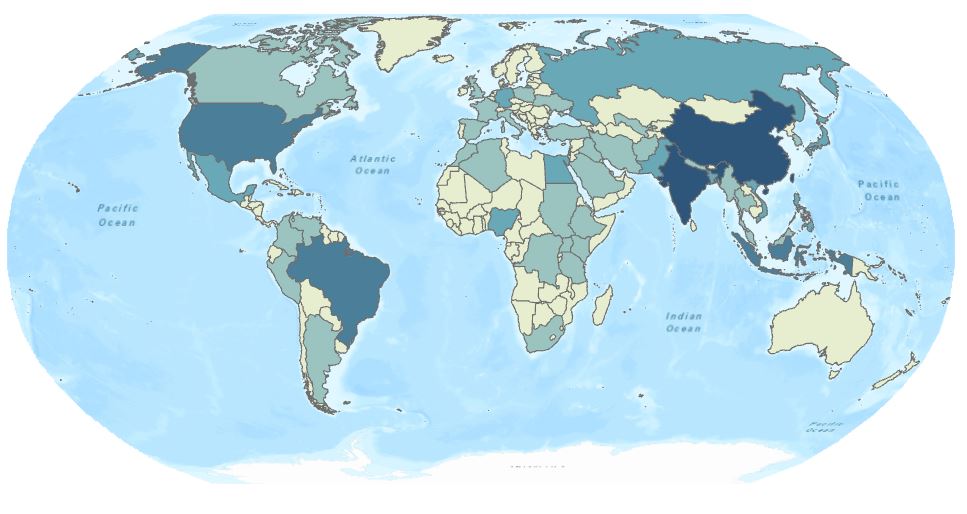
* 1. You can switch the selection between selected and unselected features with the switch selection button in the attribute table.[[10]](#footnote-10)



* 1. Practice visualizing by some different characteristics. Right-click on the layer in the Contents pane and choose Symbology to change the color display of the shapefile. Last week you used a display by category. Try that for these data by:
     1. Selecting Unique Values for the Primary symbology, then
     2. Choosing country name for Field 1. Your map may look something like below:[[11]](#footnote-11)



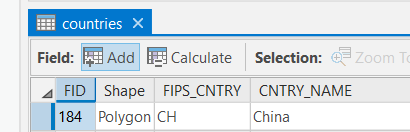
* 1. This map does a better job of emphasizing the different countries than just the borders alone.
  2. Now try displaying population[[12]](#footnote-12) using graduated colors. Your new map could look something like this:



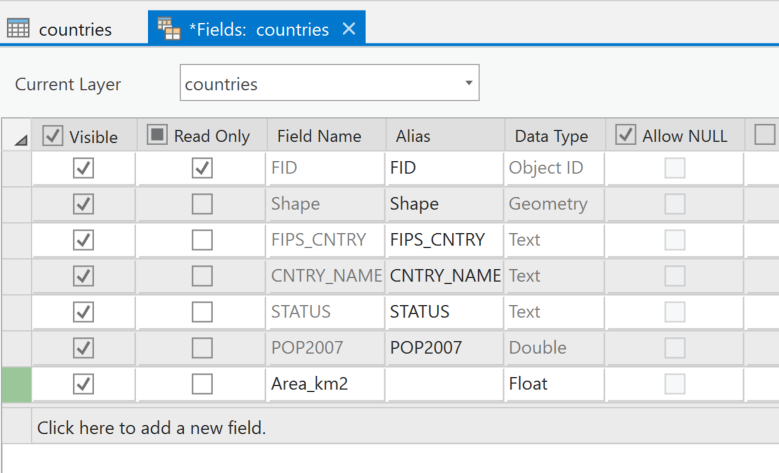
* 1. This new map still conveys the country borders, but also provides additional information about population.

# Part B: Table Operations: Calculate Geometry and Calculate Field

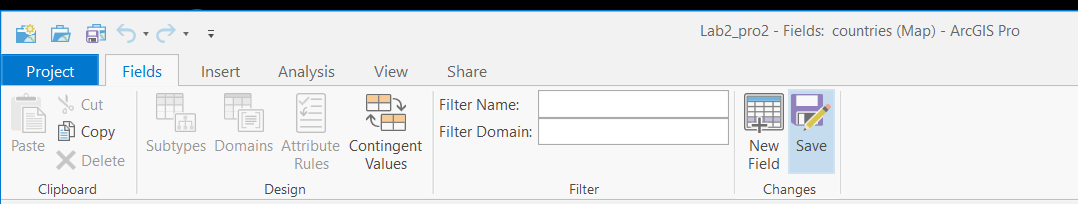
1. Now, let’s say that you want to create additional information in your shapefile. For example, we know population of each country. Let’s use GIS to calculate population density of each country in persons/km².[[13]](#footnote-13)
2. The first thing we need to calculate population density is country area[[14]](#footnote-14). Because features in a GIS are spatial, it is easy for ArcGIS Pro to calculate any geographic attributes.
3. First, you need to add a new column to your attribute table. To do this, open the countries attribute table. In the Field group at the top of the table, click the Add symbol. A table showing all the fields and their specifications pops up. This is called the Fields view.



1. Your new field has been added at the bottom. Name your new field ‘Area\_km2’.[[15]](#footnote-15)

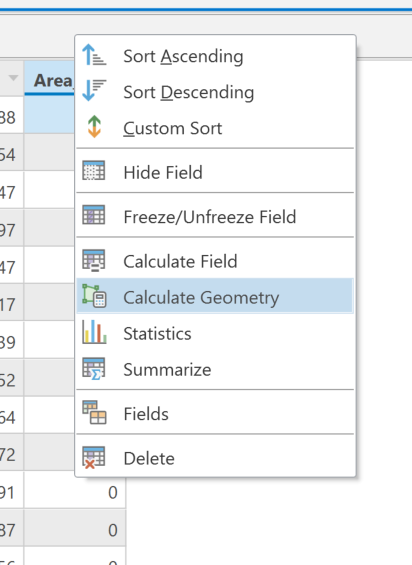
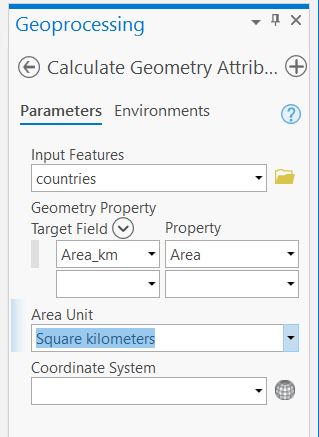


1. An extended footnote[[16]](#footnote-16): ArcGIS limits you to ten (10) characters in column titles. Use them wisely! It also restricts the use of spaces or other ‘special’ characters[[17]](#footnote-17). To help keep your units organized, use units in your column title! For example, if you just titled the column ‘Area’, you wouldn’t immediately know if it was square kilometers, square miles, acres, hectares etc.
2. Double-click in the Data Type column of your new field to bring up a drop-down menu. Chose **float**.[[18]](#footnote-18)
3. In order to save this new field, you have to click Save! The Save button is at the top of the document, in the ‘Fields’ tab that has appeared. Once you have saved, close out the Fields view.



|  |
| --- |
| What does Data Type mean?[[19]](#footnote-19)  “Type” refers to how the computer and software store information. Every column in an ArcGIS attribute table must have a type associated with it, which sets restrictions on what can be stored in that field, what calculations you can do, and so on. Whenever you add a field, ArcGIS Pro defaults to calling the data values “Long,” which stands for “long integer.” An integer is a whole number (e.g., 1, 2, 3589). A “short” integer in computer terms ranges from -32768 to 32,768. Those weird numbers encompass 216 [[20]](#footnote-20) numbers, which the computer stores as 16 bit (or 2 byte). If you select “long” integer, you get 224 integers to choose from (out to over 8 million).  Integers are all well and good, but your data analysis will usually require more precision, and for that we need decimals. For that, we want “float”, which you just used for your new area column above. “Float” refers to floating point, which includes both integers non-integer numbers (e.g., 3.14, 3589.01). If you ever calculate a fraction, make sure that you have assigned a column value of floating point – otherwise ArcMap will round to whole numbers.  Another important “Type” category to know for now is “text”, which is probably self-evident. If you want the values of the column to be letters instead of numbers, make it text. As you might guess, you can’t do any numeric calculations on text (try dividing Croatia/2 – doesn’t work)[[21]](#footnote-21). Likewise, if you name a column type as text and put numeric values into it, ArcGIS Pro won’t let you do any numeric operations because it understands the column as text[[22]](#footnote-22).  Once a type is set YOU CANNOT EVER CHANGE IT so type carefully. (Or, just be ready to delete the field and add it again with the correct choice of type). |

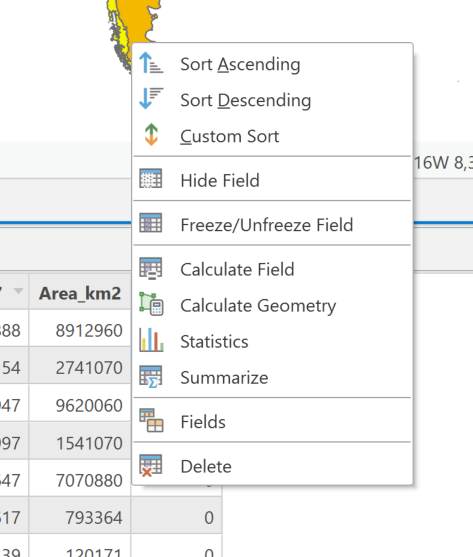
1. In the countries attribute table, right click on the new column title Area\_km2 and go to calculate geometry. A Geoprocessing[[23]](#footnote-23) pane will open at the right of the document.[[24]](#footnote-24)

1. By changing the Property of calculate geometry, you can calculate area, perimeter, X centroid or Y centroid of a polygon shapefile[[25]](#footnote-25). If this shapefile were a line file, you’d be able to calculate length (but not area)[[26]](#footnote-26).
2. Select Area from the Property drop-down, and choose Square kilometers for the Area Unit. Click Run (or OK depending on your Arc version). Now you’ve got a new column in your table of country land area in km2.

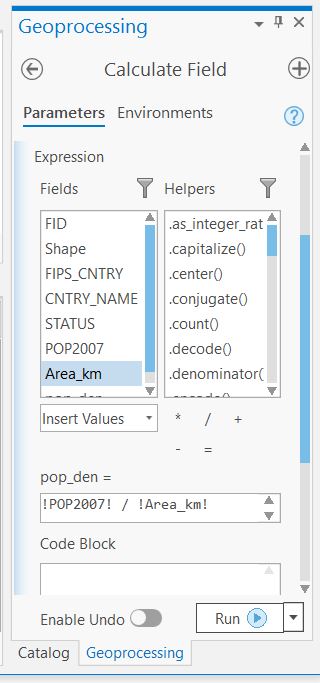
**Practice: What’s the largest country in terms of Area? How did you find out?**

1. Now you have the information you need to calculate population density. We just need a field to put it an. Add another new field. Call it pop\_den and make it float[[27]](#footnote-27). Right click on pop\_den and choose Calculate Field. Calculate Field allows you to do any numeric manipulations on columns of data (similar to Excel, although without as much functionality).



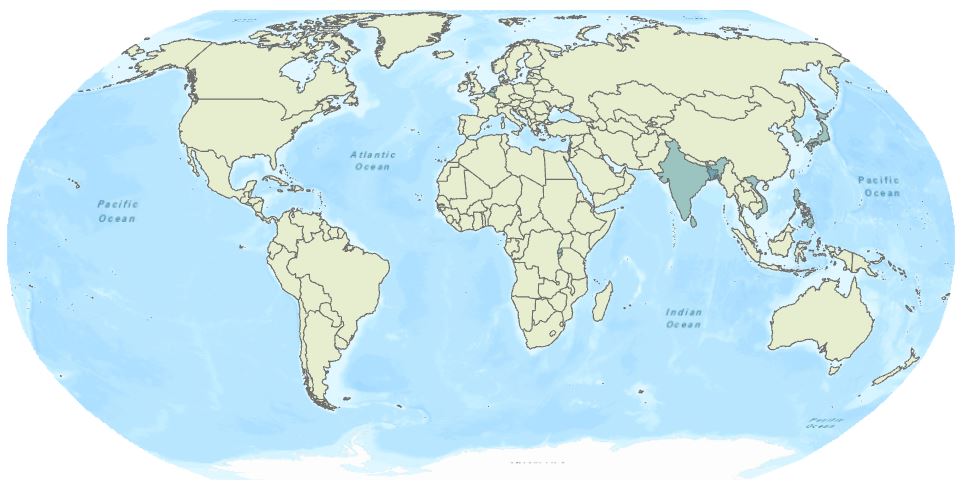
1. In the Geoprocessing pane that appears, divide POP2007 by Area\_km. The easiest way to do this in Calculate Field[[28]](#footnote-28) is to double click on the fields and the operations that you want. Operators are located below the ‘Helpers’ column. Once you’ve set up the equation, click Run.

Graphical user interface, application

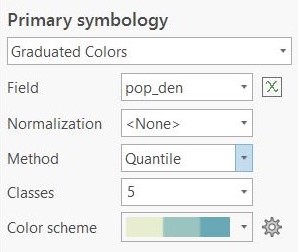
Description automatically generated

OR

1. Now, make a map of graduated colors based on population density. The default map might look something like[[29]](#footnote-29):



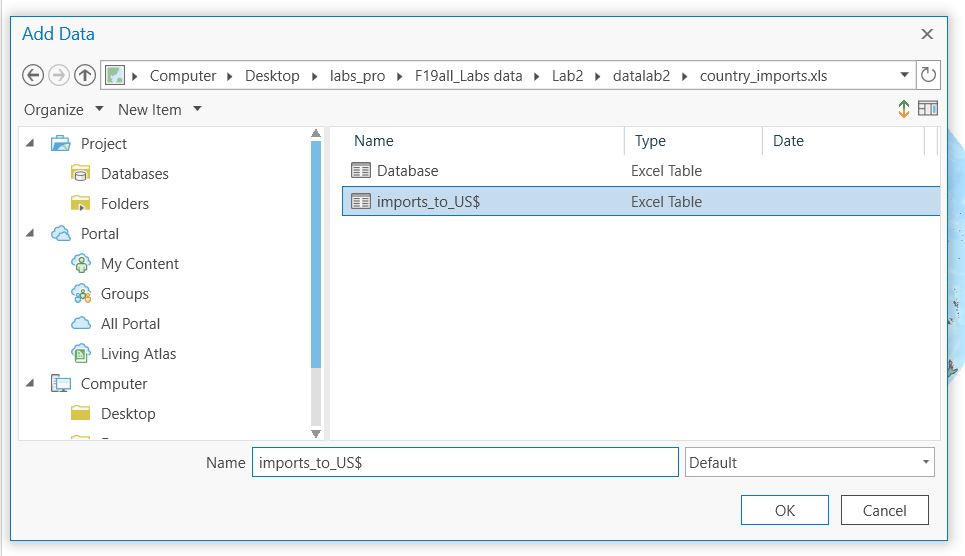
This color scheme is not terribly useful because the default display divides the colors in ‘Natural Breaks’ between 0 and 13377. The densest countries are messing up the scale! Make a better color stretch by choosing a new way to divide the classes from the Method drop-down menu.



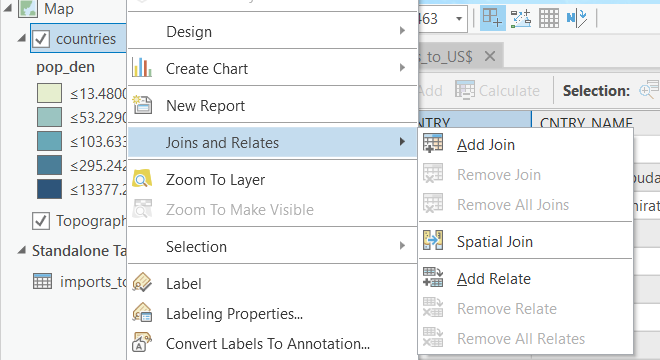
1. There are several methods of dividing the classes. Try using Quantile to see how that looks. If you have five quantiles, then the top 20% will be one color, second 20% another etc. If you choose ten quantiles, then the top 10% will be one color, second 10% another etc.[[30]](#footnote-30)
2. We’ll play with this more later on. For now, you’ve got a pretty new map that effectively displays global countries by population density.

# Part C: Joining Tables to Spatial Data

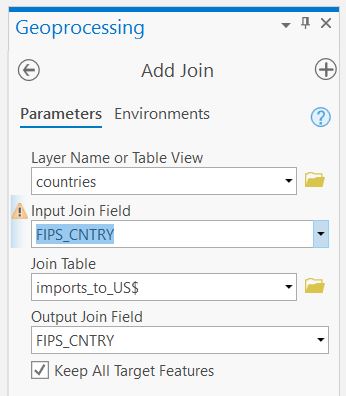
1. Sometimes spatial data already contain all the information you need. But, often you’ll have some other information that needs to be connected to the spatial data. In this example, you have a shapefile of countries of the world and information about their names and populations (and now population densities). Any additional information about countries you might want (e.g., native language, GDP, percent forest cover) will need to be joined to the shapefile.
2. In your lab folder you will also find an Excel spreadsheet named country\_imports.xls. Open the spreadsheet in Excel and check it out. This spreadsheet contains country codes, country names, and two new categories called ‘all\_imprt’ (which is the value (in $US) of all imports to the U.S. coming from that country), and Plnt\_imprt (the value (in $US) of all live plant imports to the U.S. coming from that country)[[31]](#footnote-31).
3. Once you are done peeking at the data, close Excel. (if you try to open the same file in Excel and ArcGIS at the same time, bad things happen[[32]](#footnote-32)). Let’s put this into our Project.
4. Unlike a shapefile, you can’t simply drag an Excel file from your Catalog to your Contents to add it. Instead, click the Add Data button on the Map tab and choose ‘Data.’ Navigate to the Excel file and select it. Choose imports\_to\_US$[[33]](#footnote-33)
   1. If you have weird errors adding the file to your map, try the country\_imports.csv version instead!



1. The Excel sheet is added to your Contents as a ‘Standalone Table’. Right-click on the file name and click ‘Open’ to open its attribute table. You should see the same table that you just saw in Excel. Notice that the table values FIPS\_CNTRY and CNTRY\_NAME are identical in this table and in the countries shapefile[[34]](#footnote-34).
2. Let’s put these together. Right click on countries, go to ‘Joins and Relates’, and click ‘Add Join’.



1. Join the imports\_to\_US$ table to the shapefile based on FIPS\_CNTRY and click Run.



**Practice: What do you expect to see when you open the newly joined attribute table of countries?**

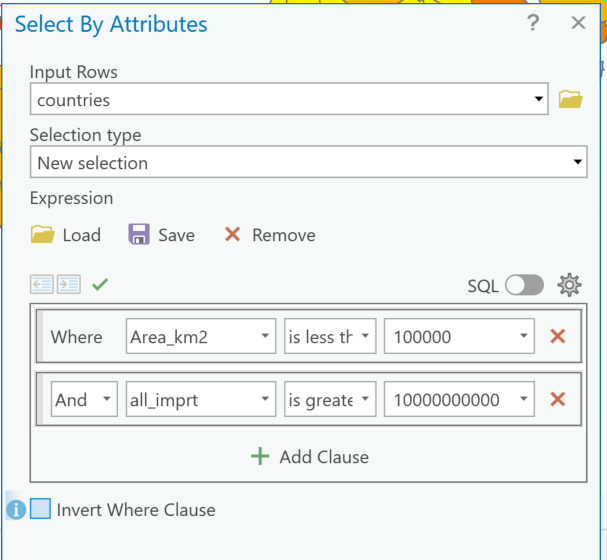
1. Let’s see what you just did! Open the attribute table of your countries shapefile. If you scroll over to the right, you’ll see that it now contains all of the columns from the imports\_to\_US table.

**Practice: Change the display to show each country by value of goods they export to the U.S[[35]](#footnote-35). What country do we import the most from in terms of dollar value?[[36]](#footnote-36)**

1. If you wanted to, you could now perform new Calculate Field calculations with your new fields. Yay math!
2. Joins do not permanently add the table data to the shapefile. If you want one of the new imports columns permanently within your shapefile, you’d need to add a new field and set it equal to one of the imports columns using field calculator. Alternatively, you could save a new copy of the shapefile that contains all of the joined data[[37]](#footnote-37).

# Part D: Selecting by Attributes

1. Finding the highest or lowest value of something is easy to do by sorting the attribute table. But, sometimes you want to select features using more complicated queries. For example, what if you wanted to select all the ‘small’ countries that do ‘a lot’ of exporting to the U.S.?
2. Under the Map tab at the top of the document, click Select by Attributes in the Selection group. A Geoprocessing pane will open. (You’ll explore Select by Location in next week’s lab).
3. Select by Attributes defaults to ‘New selection’ for Selection type. Since we haven’t made a selection yet, the other options aren’t very useful!
4. This time, create a new selection of countries that have areas less than 100,000 km2 (roughly the size of Nicaragua)[[38]](#footnote-38) AND total exports to the US greater than $10,000,000,000[[39]](#footnote-39). We’ll pull our Input Rows from ‘countries’ – that’s where all of our data is, right?
   1. Lots of zeroes…. Lots of opportunities for typos! Scientific notation can help.
5. To start, click on ‘New expression’ next to the green plus sign to bring up the equation dialogue. You’ll need to click the Add Clause button to add more than one clause. Your equation should look something like[[40]](#footnote-40):



1. How many countries fit these criteria?[[41]](#footnote-41)
2. If your calculation didn’t come out right, check the values in the selection clauses, Arc understands scientific notation:

Graphical user interface, text, application, email

Description automatically generated

1. Now, see if you can find the small countries with areas of less than 100,000 km2 that export **between** $10,000,000 and $30,000,000 of plants to the U.S[[42]](#footnote-42). Which countries fit these criteria?[[43]](#footnote-43)
2. Selecting is an art and a science[[44]](#footnote-44)! If you’ve successfully completed a complicated selection operation, you should save your selection expression so you don’t have to re-do it later (and possibly make a mistake). Choose a descriptive filename.
   1. You can hit the Load button to load a previously saved selection expression

Graphical user interface, application

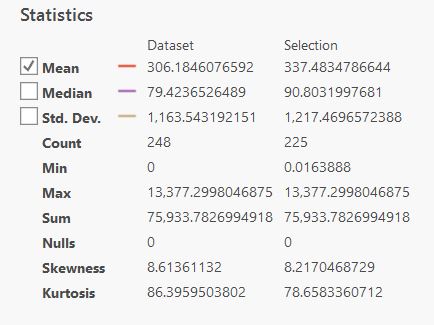
Description automatically generated

# Part E: Summarizing and Table Statistics

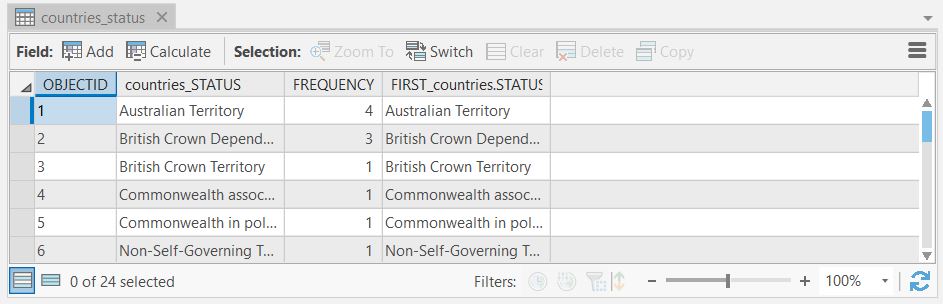
1. **Clear any previous selections.** Open the attribute table for countries again. What if you want some basic statistics about population density for all countries in the world[[45]](#footnote-45)? Right click on the pop\_den column heading and select Statistics.



1. Statistics gives you the min, max, mean of the column. But, there’s a problem with the minimum and the mean here – recall that some countries didn’t have data for population, so our calculated population values of 0 are not actually 0 people/km2, instead they refer to “No Data”. It would be better if our statistics didn’t include the zeros[[46]](#footnote-46).
2. Fortunately, you can also perform statistics on a selection within your table! In fact, statistics will always show you stats for the current selection unless you have nothing selected. Use select by attributes again to select all countries with pop\_den greater than 0 and recalculate statistics.

[[47]](#footnote-47)

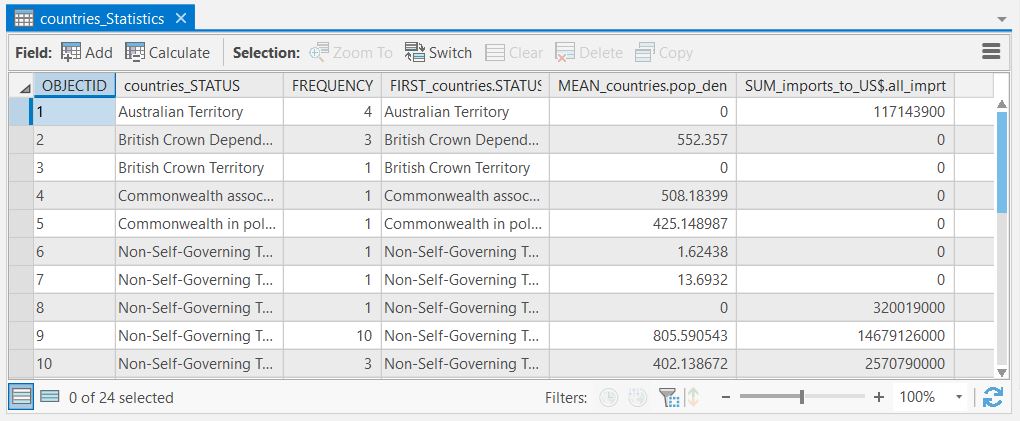
1. If you expand the Chart Properties pane, you can see that Statistics now has two columns: one for the Dataset (as a whole) and one for the Selection. In the Selection column, 0 is no longer the smallest number. Can you figure out the mean value of total exports to the U.S. coming from each country?[[48]](#footnote-48)
2. **Clear all selections** **again** so we can look at a final data analysis tool.
3. Lastly, let’s Summarize. This tool is extremely useful when you have multiple features that fall under the same category. For example, in last week’s lab, Bellevue had sewer pipes with different diameters. You could use calculate geometry to figure out the length of each sewer pipeline type, and then use summarize to calculate the average length for each pipe diameter!
4. Open the countries attribute table and notice that several countries have the same “status” (e.g., UN Member Country). Right click on status and select Summarize. In the popup pane that appears, you can rename the output table and specify where you want to keep it. You also need to fill in the Statistics Fields. Numeric attribute fields can be summarized using any statistic. Text attribute fields (like STATUS) can be summarized using types MIN, MAX, COUNT, FIRST, LAST, and UNIQUE. Choose ‘countries.STATUS’ for Field and select Count for the Statistic Type.



1. Open the table. This summary gives you a count of the number of features (countries) that have each status. How many countries are UN Member States?

**Practice: Can you think of another way that you could have found this information using Select by Attributes?**

1. Let’s try Summarize again, but this time we’ll get a little more information out of it. In the Summary Statistics pane, leave the ‘Status’ field as is (statistic type Count), and add a field for pop\_den with statistic type Mean, and all\_imprt with statistic type Sum.



1. With this new summary, you can now figure out (for example): What is the average population density of UN Member States? How does that compare to the Occupied territory? What is the total value of annual U.S. imports from Non-Self-Governing Territories of the UK?[[49]](#footnote-49)

# Part F, Lab Production: Analyzing Grizzly Bear Habitat

**Problem Statement:**

A century and a half ago, nearly one hundred thousand grizzly bears roamed the mountains and plains of the western United States. Today, only a few thousand remain. The grizzly bear, *Ursus arctos*, is listed as a threatened species under the U.S. Endangered Species Act, requiring that the federal government restore the bear's population to a level that removes them from the threat of extinction. To do this, the U.S. Fish and Wildlife Service designated five recovery zones in the Northern Rockies of Wyoming, Montana, Idaho, and northeastern Washington. A sixth zone, called private inholdings within one of the recovery zones, is included in this dataset but is not yet undergoing data collection. For purposes of analysis and conservation, this sixth zone is ***not*** considered a recovery zone.

**Data:**

The data you need to answer these questions is in the \Grizzlies folder

* A shapefile called region.shp which is the study area
* A shapefile called grizone.shp which are the grizzly bear recovery zones
* A .csv file called griznum.csv which contains a count of the number of bears observed in each recovery zone[[50]](#footnote-50)
* An image called Grizzly.jpg[[51]](#footnote-51)

**Questions to be answered:**

* 1. What are the names of the grizzly bear recovery zones?
  2. What is the total area of all recovery zones (*in square km*)? Briefly explain the steps you used to determine this information.
  3. What is the observed count of bears for each recovery zone? Briefly explain the steps you used to determine this information.
  4. What is the density of bears[[52]](#footnote-52) in each recovery zone (*in bears per km2*)? Briefly explain the steps you used to determine this information.

# Map requirements:

1. Include two maps[[53]](#footnote-53) in the Layout.

*Map Frame #1*

1. Includes the three regional states, outlined and labeled, for context. You can use any basemap with this outline.
2. Each of the grizzly bear recovery zones is symbolized with a unique color[[54]](#footnote-54).
3. A legend is included showing the name of the recovery regions related to each color (Remove any unnecessary words or symbols from the legend).
4. The map contains a scale bar (in km) and north arrow.

*Map Frame #2*

1. Includes the regional states for context.
2. Each of the grizzly bear recovery zones should be symbolized with a dot density map showing bears per square kilometer.[[55]](#footnote-55)
3. A legend is included telling us dots/unit area.
4. The map contains a scale bar (in km) and north arrow.
5. Include on your layout a labeled image of a grizzly bear[[56]](#footnote-56).

Export your layout as an image and insert the image into a document with the answers to questions 1-4. Remember to convert your lab report file to .pdf before submitting on Moodle!

1. Except more fun [↑](#footnote-ref-1)
2. Eventually this should just be your automatic reaction [↑](#footnote-ref-2)
3. This really should be updated [↑](#footnote-ref-3)
4. See screenshot next page [↑](#footnote-ref-4)
5. China, India, USA [↑](#footnote-ref-5)
6. ctrl [↑](#footnote-ref-6)
7. …shift [↑](#footnote-ref-7)
8. This action is very important to remember. You can also clear a selection in the attributes table by pressing CTRL + SHIFT + A. [↑](#footnote-ref-8)
9. See screenshot next page. [↑](#footnote-ref-9)
10. You are going to use these a lot, forever. [↑](#footnote-ref-10)
11. Depending on which crazy color scheme you pick [↑](#footnote-ref-11)
12. Think: What do you need to change to make the symbology different? [↑](#footnote-ref-12)
13. Not a footnote, but an exponent. It’s…going to get weird. [↑](#footnote-ref-13)
14. Because density is things per unit. [↑](#footnote-ref-14)
15. Without the quotations, with the underscore ( \_ ) [↑](#footnote-ref-15)
16. People don’t always read these, so the footnote is the text. Unicorns. [↑](#footnote-ref-16)
17. @&\*, for example. [↑](#footnote-ref-17)
18. Short for floating point. [↑](#footnote-ref-18)
19. Font color changes are another way to draw attention to components of a document [↑](#footnote-ref-19)
20. First is an exponent, second is this footnote [↑](#footnote-ref-20)
21. Yugoslavia is divisible into many separate pieces, however [↑](#footnote-ref-21)
22. ArcGIS only does what you tell it to [↑](#footnote-ref-22)
23. Depending on your version of Arc, this popup may be called ‘Calculate Geometry’, and may look slightly different than the screenshot. [↑](#footnote-ref-23)
24. We want area in this column – so let’s put it there! [↑](#footnote-ref-24)
25. Note the projection information we’re ignoring for now. Soon. [↑](#footnote-ref-25)
26. Two-dimensions and all [↑](#footnote-ref-26)
27. ArcGIS rarely makes it rain [↑](#footnote-ref-27)
28. And the way that Arc complains the least about your calculations. Best practice is to click the buttons. [↑](#footnote-ref-28)
29. Dramatic pause [↑](#footnote-ref-29)
30. Data visualization will be a theme in the course. [↑](#footnote-ref-30)
31. Consider this your metadata dictionary for this file [↑](#footnote-ref-31)
32. ArcGIS and Excel fight over who has ‘write’ privileges and no one, specially you, dear analyst, is happy. [↑](#footnote-ref-32)
33. You might not have ever noticed, but each sheet within Excel is its own file: ArcGIS recognizes this, and makes sure you select the one you want to import [↑](#footnote-ref-33)
34. This is important. We join like to like - wondering which field is the same? Open the table and look! [↑](#footnote-ref-34)
35. Column name is all\_import… weird! [↑](#footnote-ref-35)
36. Surprise, it’s Japan! [↑](#footnote-ref-36)
37. This process can be a bit tedious… We’ll wait until we’re a bit more comfortable before trying. [↑](#footnote-ref-37)
38. ‘small’ [↑](#footnote-ref-38)
39. ‘a lot’ [↑](#footnote-ref-39)
40. The next page, not a blank space. Focus on the lab, Swift [↑](#footnote-ref-40)
41. The answer should be 26 – make sure you can figure out how to find this information. [↑](#footnote-ref-41)
42. There’s a lot going on here! Think about the different processes you’re asking the query to make, the field you’re making the query on… lots of moving parts. [↑](#footnote-ref-42)
43. Ask the TAs for the answer on this one. [↑](#footnote-ref-43)
44. For better or worse, you’ll get lots of practice. [↑](#footnote-ref-44)
45. Power! [↑](#footnote-ref-45)
46. Remember that ‘No data’ and ‘0’ are two different concepts! [↑](#footnote-ref-46)
47. So many decimals…. [↑](#footnote-ref-47)
48. ~$32.5 million [↑](#footnote-ref-48)
49. Mostly rhetorical, but practice data manipulation with these questions (or others!) in mind to gain experience with the capabilities. [↑](#footnote-ref-49)
50. This file is also available as an xls file, for fun. [↑](#footnote-ref-50)
51. Not a spatial data layer, just a picture of a bear [↑](#footnote-ref-51)
52. How could you calculate this? [↑](#footnote-ref-52)
53. Insert tab -> New Map. Work on one map at a time, though. [↑](#footnote-ref-53)
54. Beware the dreaded <all other values>… Click the More dropdown! [↑](#footnote-ref-54)
55. You will have to adjust this yourself by typing in the appropriate value. What do you have more of, bears or square kilometers? Your dots will have to represent fractional bears – poor bears! [↑](#footnote-ref-55)
56. Beets and Battlestar Galactica references optional [↑](#footnote-ref-56)