

Relationships Count when Mapping?

Database “joins” and “relates” form the basis of many GIS analyses where tabular data need to be mapped. Tabular data can take many forms including Excel Spreadsheets, Comma and/or Tab-delimited files, and even simple text files. Of course, to map tabular data, the data need to be organized into fields (columns), each with a header, and one field, a “key field”, must be present in both the tabular data file and a GIS file which contains some mappable polygon, line, or point geometry (ex. a shapefile of states, counties, census blocks, ZIPcodes, wildlife management areas, etc.) To make a map graphic, the tabular data need to be somehow appended to the geometry file. This brings us into the world of databases and spatial relationships, more specifically, joins and relates.

When the relationship between the geometry (point, line, or polygon) and the tabular data is a one-to-one relationship, i.e. a row in the tabular data will only match one geometric feature, the tabular “join” is the way to go. In these cases, all you need to do is construct the join using the two key fields, export the joined data to a new file (joins are only temporary) and continue on with your mapping. You can symbolize your geometry using any attribute in the tabular data... easy peasy!

BUT... what if there is a many-to-one relationship, i.e. multiple entries in the tabular data match to a single geometric feature. Think of the case where there is a list of 100s of members of a group with a data table containing the City, State, and ZIP code for each entry and your goal is to construct a map at the ZIP code-level. You have a high expectation that more than one person lives in any particular ZIP code, hence a one-to-many relationship. Now you are into the world of “relates” which brings new questions and options because the tabular data are not immediately “mappable”.

As with joins, when relating tabular data to GIS geometry, the tables are related through a common “key field”, but because the geometry can only be mapped once, the related data appear only when an “identify” or “query” of the mapped feature is executed. Then, the tabular data are usually displayed in a “pop-up”. This is great for a web-map, but... what to do when you want to make a paper, .PDF, or graphic figure version of the map for publication or other dissemination? A typical example of this case is when you may have the membership of some group (no addresses or personal data, please) as a list of county names (or ZIPcodes)

and you want to map their locations AND indicate the number of individuals per county (or ZIPcode) for publication in a journal or magazine article.

As with most GIS analyses and constructions, there is always more than one way to get there, you just need to be creative. For this month’s Tips & Tricks, I’ll show you two ways to join tabular data to a GIS geometry file when the underlying relationship is a one-to-many and produce a map.

TIP #1: USING STATISTICS IN ARCGIS PRO

Assuming that you are starting with some GIS-compatible geometry file (a shapefile or feature class of ZIP code polygons) and a Microsoft Excel tabular file of attributes which includes a five-digit ZIP code for each record in the Excel table,

Step 1 — IMPORT the Excel tabular file into your geodatabase.

Tip 1A: Remember to check and adjust, as needed, the column headers in the Excel file. The headers need to start with letters, most special characters and/or spaces are not permitted. Underscores are permitted characters in the column header. Column header length may be truncated if too many characters.

Tip 1B: Remember to check the compatibility between the key fields in the geometry and tabular files. The data in the key fields MUST be identical. It is easy to think that numbers are always numbers, but numerical fields are right-justified; character fields (numbers as characters) are left-justified. Numerical numbers and character numbers will not join or relate.

Tip 1C: NOTE: Depending on your ArcGIS Pro (or Desktop) installation, if you have not installed the Microsoft drivers, you will get an Error Message and the Excel file will not be imported into the geodatabase. Do not panic. Simply open the file in Excel, export the file (or single worksheet) as either a Tab- or Comma-delimited DOS file, and then import that DOS file into the project’s geodatabase. After importing,

Photogrammetric Engineering & Remote Sensing
Vol. 89, No. 11, November 2023, pp. 661-664.
0099-1112/22/661-664

© 2023 American Society for Photogrammetry
and Remote Sensing
doi: 10.14358/PERS.89.11.661

I like to open the file just to make sure that you are getting what you think you are getting. If when opening the file, you are not seeing the file imported correctly, try a different DOS format.

Step 2 — OPEN the tabular file and make any needed selections. In the case below, I was only interested in those members who live in the State of Florida. Then select (left-click) and right-click on the field of interest, in this case “ZipCode5” and choose “Summarize” (Figure 1):

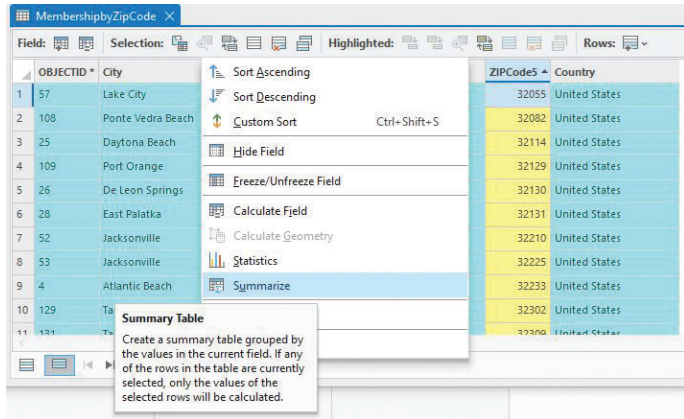


Figure 1. Selecting a field in the data table to summarize in ArcGIS Pro.

Step 3 — The Summary Statistics dialog box will allow you to; (1) specify an output table to your geodatabase, (2) choose the field to summarize, and (3) specify the statistic type for the new table as in Figure 2. In this case, because I want to make a choropleth map, ZIP code polygons colored by the number of members in each ZIP code, I chose “Count”. The resulting table (Figure 3), contains the ZIPcode5 and the FREQUENCY (=Count) for each ZIPcode5.

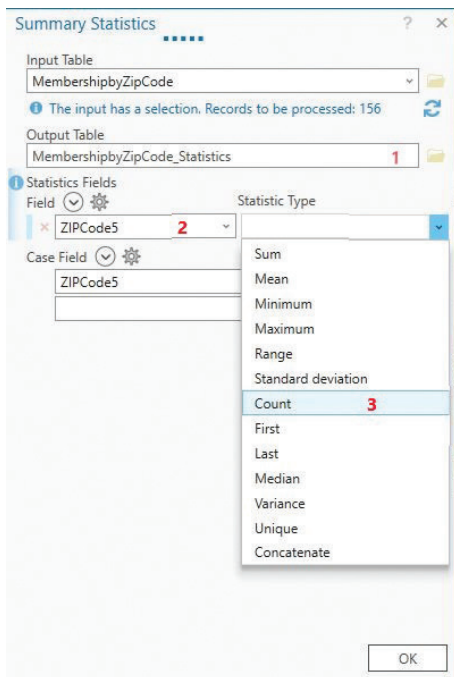


Figure 2. Summary Statistics menu showing the “Count” as the Statistic Type.

OBJECTID *	ZIPCode5	FREQUENCY	SUM_ZIPCode5
12	32311	1	32311
13	32312	2	64624
14	32579	1	32579
15	32603	1	32603
16	32605	1	32605
17	32607	1	32607
18	32609	1	32609
19	32610	1	32610
20	32611	4	130444
21	32615	1	32615
22	32653	2	65306
23	32708	1	32708

Figure 3. The table that results from the running the summary statistics tool contains the FREQUENCY field for each ZIPCode5.

Step 4 — Join this new table to your ZIP Code geometry file for mapping (Figure 4). In this case, I also labeled the ZIP code polygons with the number (=FREQUENCY) of members.

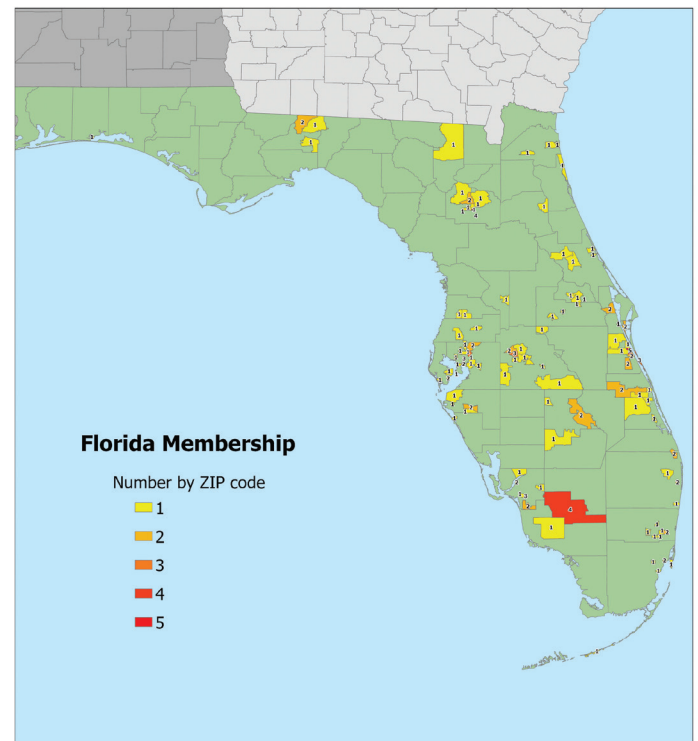


Figure 4. Map showing membership summarized by the ZIPcode5 attribute joined to the ZIP code polygons.

TIP #2 — MAKING A HEATMAP IN ARCGIS PRO

As an alternative to the simple choropleth map (Figure 4) a “Heat Map” shows the same data but tends to draw the viewer’s attention to areas with higher concentrations of the mapping variable and is more visually appealing to many observers. However, to make a heat map, the geometry must be points rather than polygons. To make a heat map using the same data:

Step 1: Use the Data Management | Feature to Point tool to convert the ZIP code polygons into Centroids,

Step 2: Join the MembershipbyZIPcode_Statistics table to the ZIP code centroids,

Step 3: Open the Symbology pane and choose Heat Map from the dropdown list of Primary symbology (Figure 5),

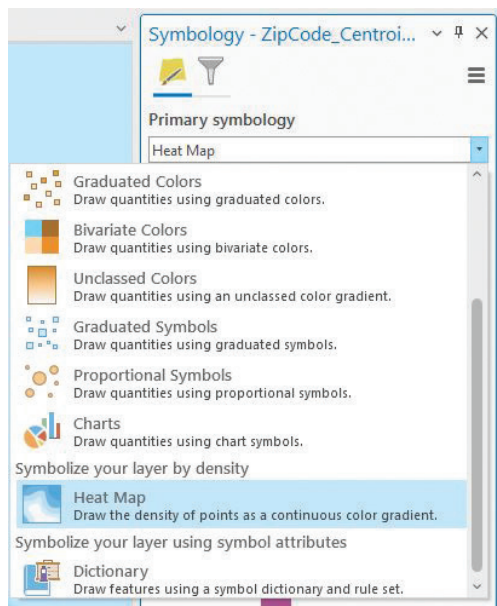


Figure 5. Choosing the Heat Map option from the Symbology Menu for a point feature class.

Step 4: Adjust the Radius, Color scheme, and Method (Figure 6) to suit your needs (remember... never accept the defaults), and

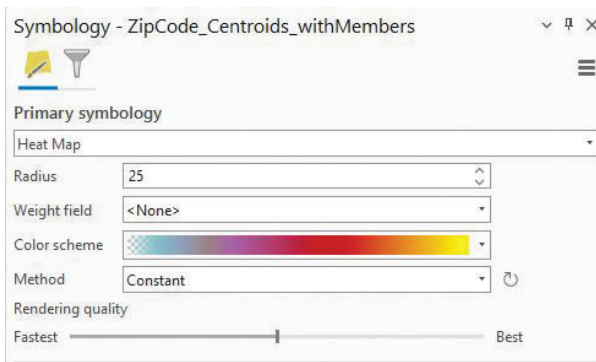


Figure 6. User-specified options for Heat Map Symbology

Step 5: Label you map as needed (Figure 7.)

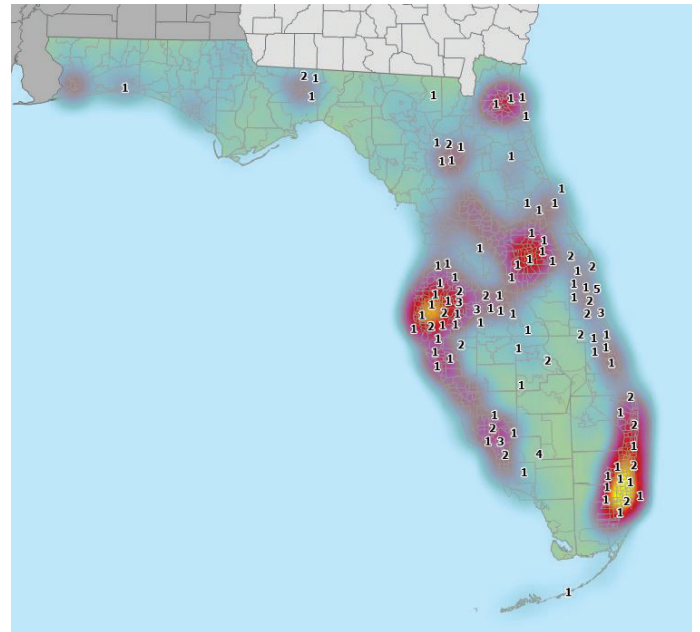


Figure 7. Heat map showing the same data as Figure 4 but using the ZIP code centroids joined to the data table.

TIP #3 — USE EXCEL TO CONSTRUCT THE STATISTICS TABLE

Excel can be used to count the number of occurrences of each ZIPcode5 in a two-step process. I recommend making a copy of the data or copying the data to a new worksheet as a precautionary measure.

Step 1: Label a new column “UNIQUE” to store the unique ZIPcode5 values. Then use the UNIQUE function to find the unique values from ZIPcode5 as in Figure 8. Enter the formula: =Unique(C2:C21) into the first cell in the UNIQUE field. (Of course, your data may have more than 21 records and may be in a different column.) For HELP with Excel formulas, use the Excel Help or visit <https://support.microsoft.com/en-au/office/overview-of-formulas-in-excel-ecfdc708-9162-49e8-b993-c311f47ca173>.

	City	State	ZIPCode5	UNIQUE
2	Tallahassee	FL	32312	=Unique(C2:C21)
3	Tallahassee	FL	32312	
4	Gainesville	FL	32610	
5	Gainesville	FL	32611	
6	Gainesville	FL	32611	
7	Gainesville	FL	32611	
8	Gainesville	FL	32611	
9	Alachua	FL	32615	
10	Sanford	FL	32772	
11	Titusville	FL	32780	
12	Titusville	FL	32780	
13	Orlando	FL	32806	
14	Orlando	FL	32816	
15	Orlando	FL	32819	
16	Winter Park	FL	32826	
17	Melbourne	FL	32901	
18	Melbourne	FL	32901	
19	Melbourne	FL	32901	
20	Melbourne	FL	32901	
21	Melbourne	FL	32901	

Figure 8. Excel spreadsheet with the formula for finding the unique values in the ZIPcode5 field.

Step 2: Label a new column “COUNT” to store the number of occurrences per ZIPcode5. Then use the COUNTIF function to count the number of occurrences of each ZIPcode5 (Figure 9) as =COUNTIF(\$C\$2:\$C\$21,E2). (Of course, your data may have more than records and you may have selected different columns for UNIQUE and COUNT.)

	A	B	C	D	E	F	G
1	City	State	ZIPcode5		UNIQUE	COUNT	
2	Tallahassee	FL	32312			=COUNTIF(\$C\$2:\$C\$21,E2)	
3	Tallahassee	FL	32312			COUNTIF(range, criteria)	
4	Gainesville	FL	32610			32610	
5	Gainesville	FL	32611			32611	
6	Gainesville	FL	32611			32615	
7	Gainesville	FL	32611			32772	
8	Gainesville	FL	32611			32780	
9	Alachua	FL	32615			32806	
10	Sanford	FL	32772			32816	
11	Titusville	FL	32780			32819	
12	Titusville	FL	32780			32826	
13	Orlando	FL	32806			32901	
14	Orlando	FL	32816				
15	Orlando	FL	32819				
16	Winter Park	FL	32826				
17	Melbourne	FL	32901				
18	Melbourne	FL	32901				
19	Melbourne	FL	32901				
20	Melbourne	FL	32901				
21	Melbourne	FL	32901				
22							

Figure 9. Excel spreadsheet using the COUNTIF function to count the number of occurrences of each UNIQUE value of ZIPcode5.

Step 3: Copy the COUNTIF formula into all of the UNIQUE value cells to get the count (Figure 10):

	A	B	C	D	E	F	G
1	City	State	ZIPcode5		UNIQUE	COUNT	
2	Tallahassee	FL	32312			2	
3	Tallahassee	FL	32312			1	
4	Gainesville	FL	32610			4	
5	Gainesville	FL	32611			1	
6	Gainesville	FL	32611			1	
7	Gainesville	FL	32611			2	
8	Gainesville	FL	32611			1	
9	Alachua	FL	32615			1	
10	Sanford	FL	32772			1	
11	Titusville	FL	32780			1	
12	Titusville	FL	32780			5	
13	Orlando	FL	32806				
14	Orlando	FL	32816				
15	Orlando	FL	32819				
16	Winter Park	FL	32826				
17	Melbourne	FL	32901				
18	Melbourne	FL	32901				
19	Melbourne	FL	32901				
20	Melbourne	FL	32901				
21	Melbourne	FL	32901				
22							

Figure 10. Results of copying the COUNTIF function for each UNIQUE ZIPcode5.

Now it is just a simple task of copy/paste special the values from the UNIQUE and COUNT columns to a separate worksheet and joining the data to the GIS Geometry file for mapping.

Send your questions, comments, and tips to GISTT@ASPRS.org.

Al Karlin, Ph.D., CMS-L, GISP is with Dewberry's Geospatial and Technology Services group in Tampa, FL. As a senior geospatial scientist, Al works with all aspects of Lidar, remote sensing, photogrammetry, and GIS-related projects.

JOURNAL STAFF

Editor-In-Chief

Alper Yilmaz, Ph.D., PERSEditor@asprs.org

Associate Editors

Ravishankar Dwivedi, Ph.D., rsdwivedi51@gmail.com
 Valérie Gouet-Brunet, Ph.D., valerie.gouet@ign.fr
 Petra Helmholz, Ph.D., Petra.Helmholz@curtin.edu.au
 Dorota Iwaszczuk, Ph.D., dorota.iwaszczuk@tum.de
 Desheng Liu, Ph.D., liu.738@osu.edu
 Clement Mallet, Ph.D., clemallet@gmail.com
 Sidike Paheding, Ph.D., spahedin@mtu.edu
 Norbert Pfeifer, np@ipf.tuwien.ac.at
 Rongjun Qin, Ph.D., qin.324@osu.edu
 Ribana Roscher, Ph.D., ribana.roscher@uni-bonn.de
 Zhenfeng Shao, Ph.D., shaozhenfeng@whu.edu.cn
 Filiz Sunar, Ph.D., fsunar@itu.edu.tr
 Prasad Thenkabail, Ph.D., pthenkabail@usgs.gov
 Dongdong Wang, Ph.D., ddwang@umd.edu
 Qunming Wang, Ph.D., wqm11111@126.com
 Ruisheng Wang, Ph.D., ruiswang@ucalgary.ca
 Jan Dirk Wegner, jan.wegner@geod.baug.ethz.ch
 Bo Wu, Ph.D., bo.wu@polyu.edu.hk
 Michael Yang, Ph.D., michael.yang@utwente.nl
 Hongyan Zhang, zhanghongyan@whu.edu.cn

Contributing Editors

Highlight Editor

Jie Shan, Ph.D., jshan@ecn.purdue.edu

Feature Articles

Michael Joos, CP, GISP, featureeditor@asprs.org

Grids & Datums Column

Clifford J. Mugnier, C.P., C.M.S., cjmce@lsu.edu

Book Reviews

Sagar Deshpande, Ph.D., bookreview@asprs.org

Mapping Matters Column

Qassim Abdullah, Ph.D., Mapping_Matters@asprs.org

GIS Tips & Tricks

Alvan Karlin, Ph.D., CMS-L, GISP akarlin@Dewberry.com

SectorInsight

Youssef Kaddoura, Ph.D., kaddoura@ufl.edu
 Bob Ryerson, Ph.D., FASPRS, bryerson@kimgeomatics.com
 Hamdy Elsayed, Hamdy.Elsayed@teledyne.com

ASPRS Staff

Assistant Director — Publications

Rae Kelley, rkelley@asprs.org

Electronic Publications Manager/Graphic Artist

Matthew Austin, maustin@asprs.org

Advertising Sales Representative

Bill Spilman, bill@innovativemediasolutions.com