



**APPENDIX:**  
**UNIVERSAL GIS METHODOLOGY FOR  
THE APPLICATION OF AMERICAN  
COMMUNITY SURVEY DATA TO  
WATER SYSTEM BOUNDARIES**

This appendix is related to the Drinking Water Needs Assessment. Learn more here: [Drinking Water Needs Assessment](#)

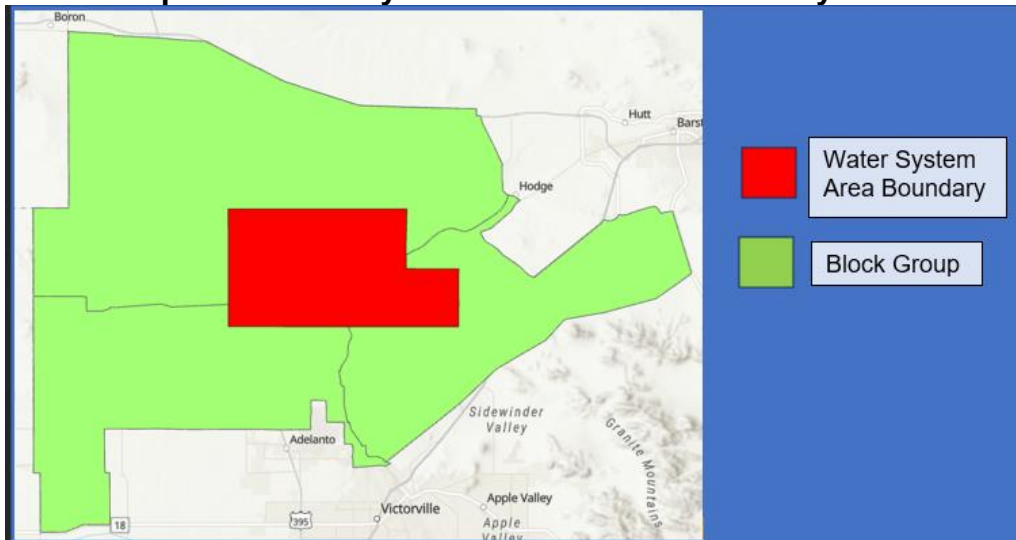
# CONTENTS

<b>APPENDIX: UNIVERSAL GIS METHODOLOGY FOR THE APPLICATION OF AMERICAN COMMUNITY SURVEY DATA TO WATER SYSTEM BOUNDARIES .....</b>	<b>I</b>
<b>INTRODUCTION.....</b>	<b>3</b>
Step 1: Data Sources .....	4
Step 2: Applying ACS to Block Groups .....	5
Step 3: Import service area boundary layer (SABI).....	10
Step 4: Excel Analysis.....	12

# INTRODUCTION

The American Community Survey (ACS) and the American Census are valuable sources of socio-economic data across California. This data does not align with communities served by water systems directly. However, there is overlap with the ACS data coverage and the water systems service area boundaries.

**Figure 1: Block Groups vs. Water System Service Area Boundary**



An approach was needed to determine the best metric to apply the data to water systems across California.

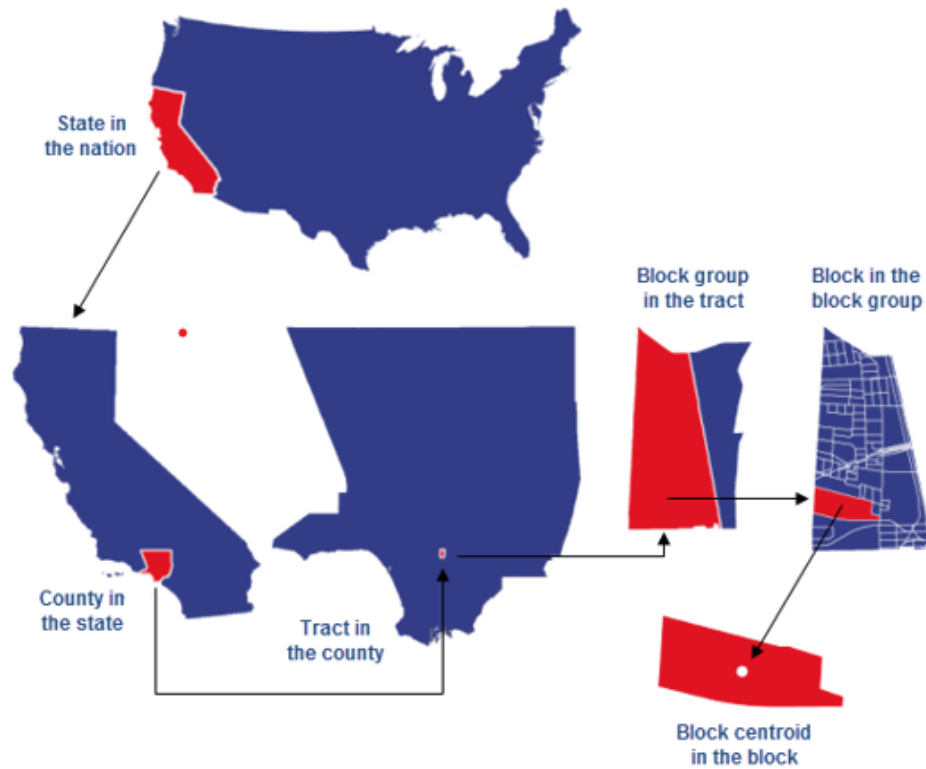
Census data is typically organized into distinct geographical boundaries that describe the size of the surveyed area. The smallest categories will be discussed for this analysis: Census Blocks, Block Groups, & Census Tracts.

**Census Blocks:** The smallest geographic area for which the Bureau of the Census collects and tabulates decennial census data.

**Block Groups:** A combination of census blocks that is a subdivision of a census tract.

**Census Tracts:** Small, relatively permanent statistical subdivisions of a county that are uniquely numbered within each county with a numeric code. On average they include approximately 4,000 inhabitants.

**Figure 2: Geographical Census Boundaries<sup>1</sup>**



Block Groups are typically used due to the wider availability of census data compared to the smaller Census Blocks and provide finer coverage than Census Tracts.

The methodology joins ACS data with block group shapefiles and then intersects the block group shapefiles with water system boundaries. An excel sheet is then produced and an area weighted average of all block groups within a system boundary is calculated to yield the data point result per water system.

The State Water Board commonly utilizes ArcPro<sup>2</sup> as one of its primary software for geospatial analysis and the following guidelines are tailored to this application.

## STEP 1: DATA SOURCES

- ACS derived or other data tied to physical boundaries such as block groups or census tracts.<sup>3</sup>

<sup>1</sup> [United States census geography—Related Concepts | Documentation \(arcgis.com\)](https://learn.arcgis.com/en/related-concepts/united-states-census-geography.htm)  
<https://learn.arcgis.com/en/related-concepts/united-states-census-geography.htm>

<sup>2</sup> Current version used in this analysis: ArcPro 3.2.1; California State Water Board GIS portal: [Gallery \(ca.gov\)](https://gispublic.waterboards.ca.gov/portal/home/gallery.html?view=grid&sortOrder=asc&sortField=title)  
<https://gispublic.waterboards.ca.gov/portal/home/gallery.html?view=grid&sortOrder=asc&sortField=title>

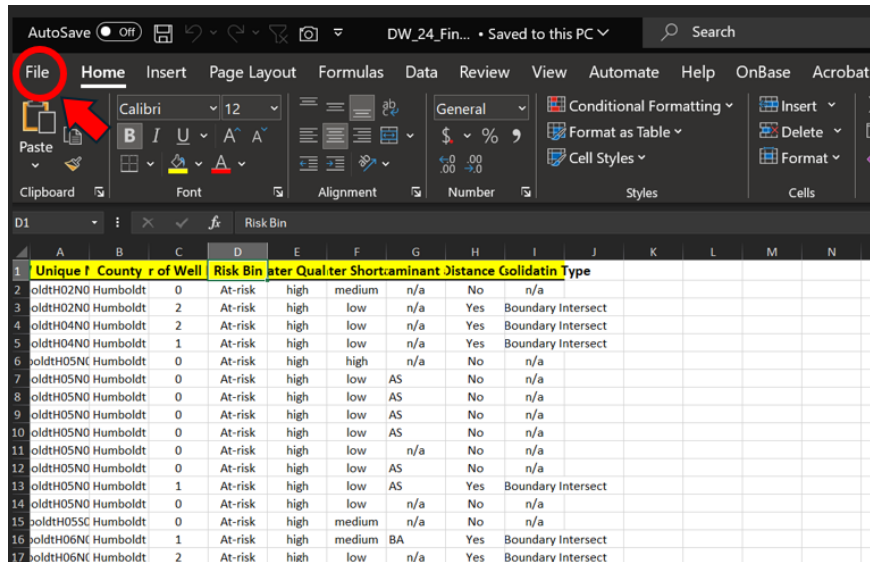
<sup>3</sup> [Census Bureau Tables](https://data.census.gov/table?g=040XX00US06$1500000) [https://data.census.gov/table?g=040XX00US06\\$1500000](https://data.census.gov/table?g=040XX00US06$1500000)

- Shapefile of either block groups or census tracts.<sup>4</sup>
- Service area boundary layer (SABL) for all targeted Water Systems.<sup>5</sup>

## STEP 2: APPLYING ACS TO BLOCK GROUPS

1. ACS data can be downloaded as an excel file, however in ArcPro a comma separated value (.csv) file type must be used.
2. Convert the ACS Excel sheet into a .csv by selecting “File” then select “Save As”.

**Figure 3: 'File' Button Location in Excel**



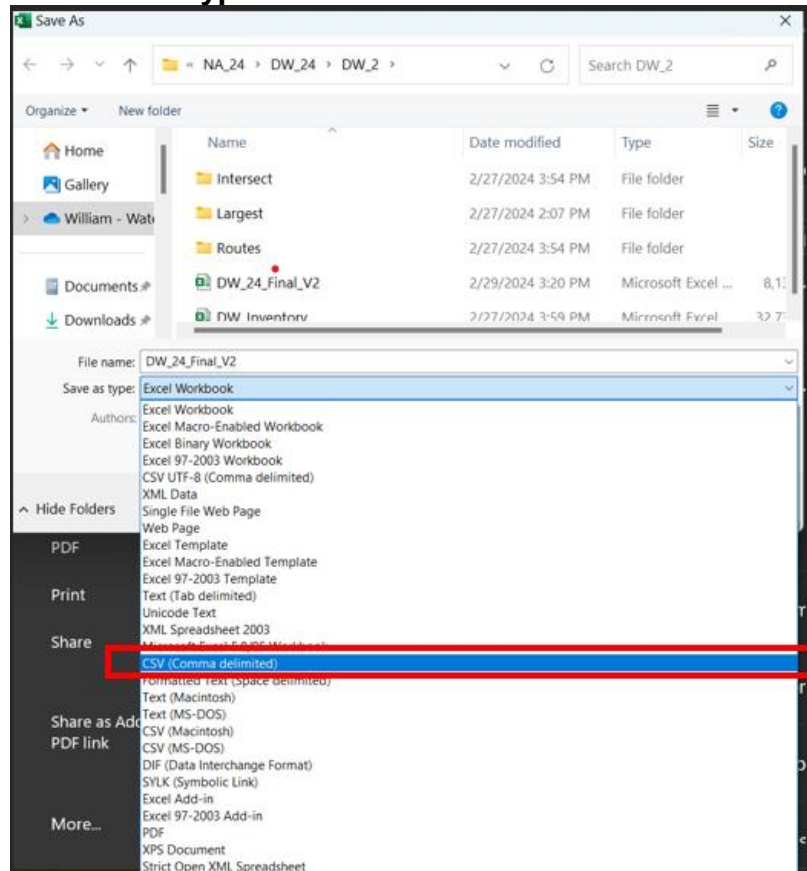
3. Under the options to save select the “CSV (Comma delineated)”.

<sup>4</sup> [TIGER/Line@ Shapefiles \(census.gov\)](https://www.census.gov/cgi-bin/geo/shapefiles/index.php) https://www.census.gov/cgi-bin/geo/shapefiles/index.php

<sup>5</sup> [System Area Boundary Layer \(SABL\) Look-up Tool \(ca.gov\)](https://gispublic.waterboards.ca.gov/portal/apps/webappviewer/index.html?id=272351aa7db14435989647a86e6d3ad8)

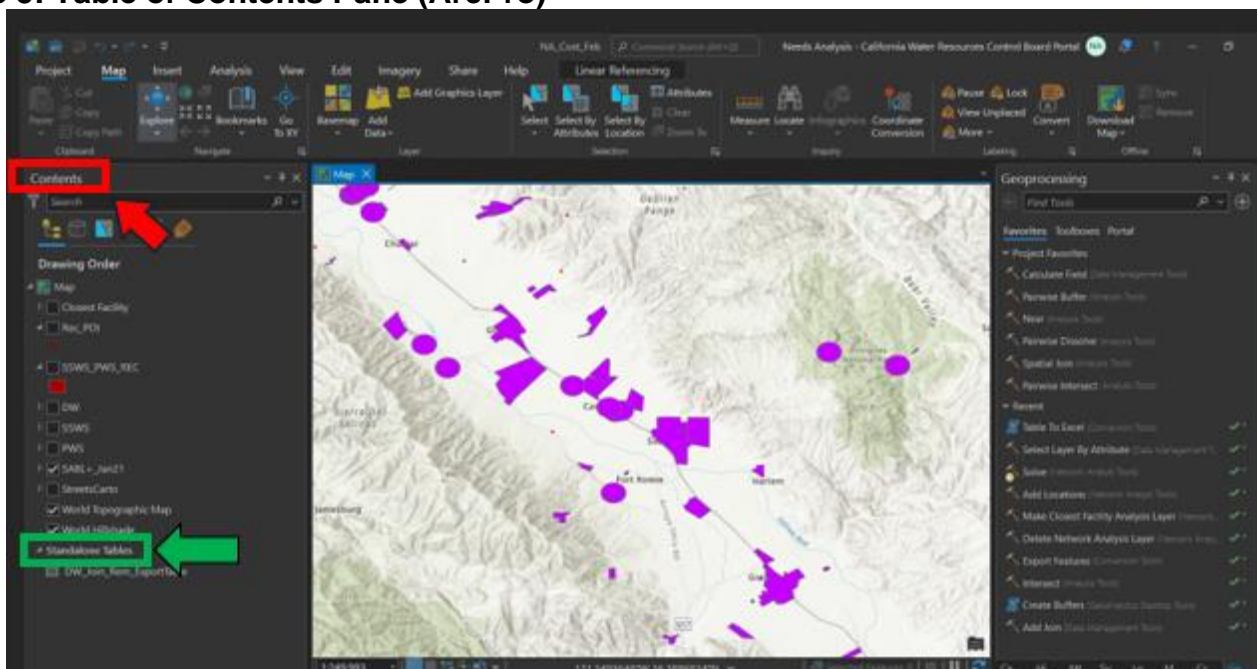
https://gispublic.waterboards.ca.gov/portal/apps/webappviewer/index.html?id=272351aa7db14435989647a86e6d3ad8

Figure 4: Comma Delimited File Type



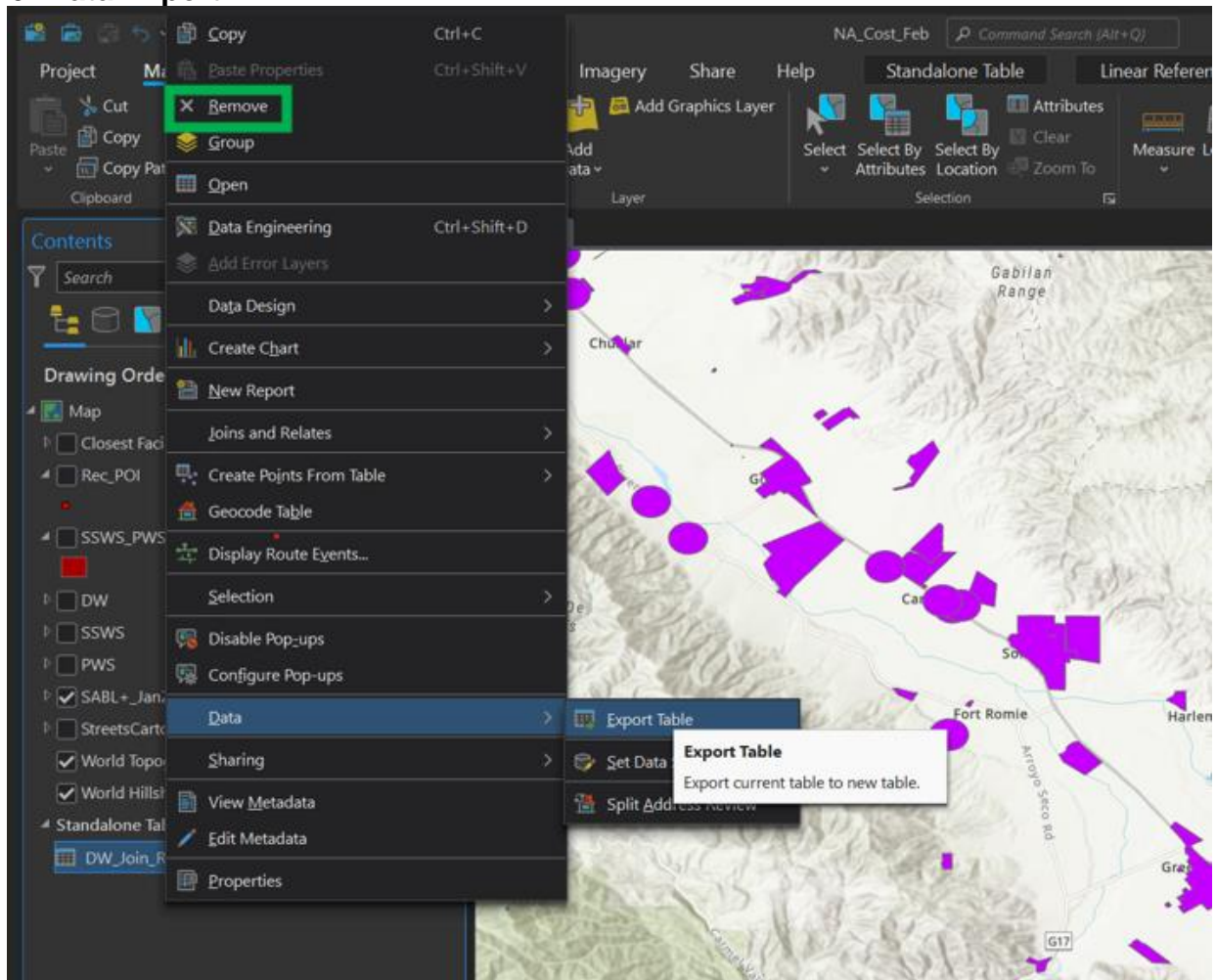
4. Once the file is saved, open ArcPro and drop the .csv into the “Table of Contents” pane.

Figure 5: Table of Contents Pane (ArcPro)



- The csv. will be automatically located under the “Standalone Table” group at the bottom (see green arrow). This table will need to be exported to the project’s geodatabase. A geodatabase is where all the data within the ArcPro project is saved and stored. For certain tools to function properly within the app it is highly recommended to import all data being used into the same database.
- To do this, simply right click the recently imported csv and select “Data” -> “Export Table”.

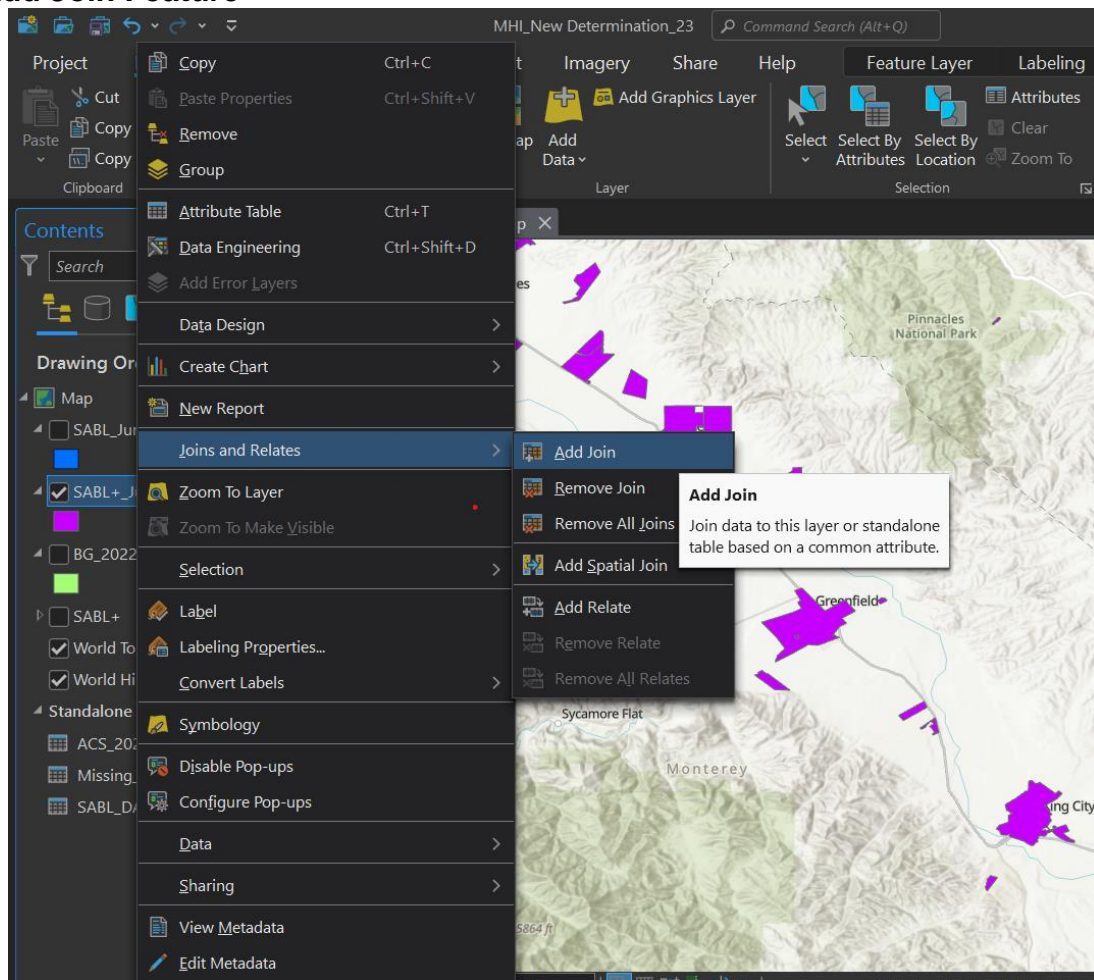
**Figure 6: Data Export**



- Save the file under the geodatabase this is being used for the project. Once saved a copy of the same .csv will be produced under the first one. Delete the original file by right clicking the file then selecting “Remove” (see green box).
- Download the shapefile for the California block groups.

9. Drag and drop the shapefile into the “Table of Contents” pane.
10. Right click on the block group shapefile, then go to “Joins and Relates” then go to “Add Join”. Select the ACS .csv file as the layer to join.

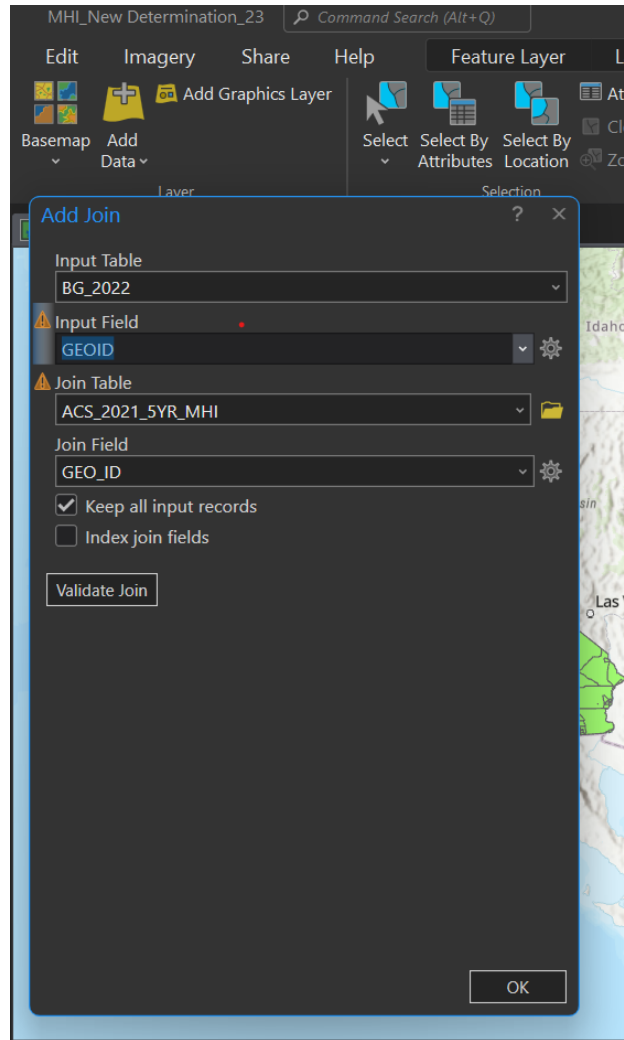
**Figure 7: Add Join Feature**



11. The “Add Join” tool window will pop up:
  - a. Select the block group shapefile as the “Input Table”.
  - b. Use the “GEOID” as the “Input Field”.
  - c. Set the ACS .csv file as the “Join Table”.
  - d. Use “GEOID” for “Join Field”.

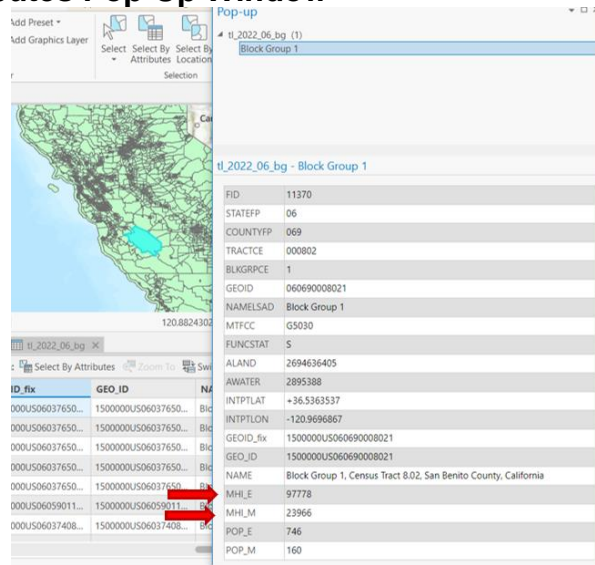


**Figure 8: Add Join Tool Window**



12. Select "OK".
13. Check the join by selecting any random block group shapefile on the map. If done successfully there will be data from the ACS .csv file within the table.

**Figure 9: Select Tool; Attributes Pop-Up Window**



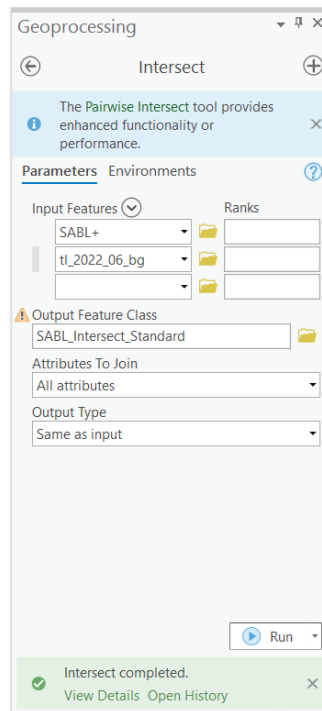
14. Use the “Select” tool located in the top ribbon under the “Map” pane to pick a random block group to check if the join was added successfully. The two arrows below show this block group has both an MHI and an MHI margin of error within its attribute table meaning that the join was applied correctly.

### STEP 3: IMPORT SERVICE AREA BOUNDARY LAYER (SABL)

15. Download the latest SABL layer either from the external waterboards GIS link or from the Water49 Database.<sup>6</sup>
16. Import the SABL layer into the “Table of Contents” pane.
17. Select the “Intersect” tool and import the block group shapefile and the SABL layers into the tool.

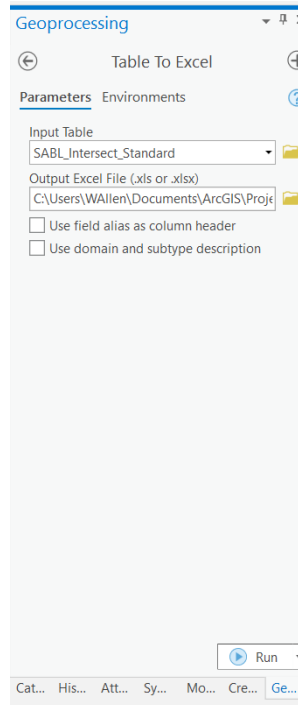
<sup>6</sup> [fetch.php](http://wiki/gis/lib/exe/fetch.php?media=connectingtowater49.pdf) http://wiki/gis/lib/exe/fetch.php?media=connectingtowater49.pdf

**Figure 10: Intersect Tool Window**



- a. Select “SABL+” and the “Block Group” shapefile as the “Input Features”.
  - b. Name the output as “Intersect” and specify where the intersect is being saved.
  - c. Select “All Attributes” for “Attributes to Join”.
  - d. “Same as Input” for “Output Type”.
18. Select “Run” a new layer representing the intersection of the two shapefiles will be produced.
  19. Search for the "Table to Excel" tool to export the data table from the intersection into an excel sheet format.
  20. Select the intersecting layer for “Input Table” and specify where the table will be saved on the computer.
  21. Select “Run”.
  22. Select the layer that is going to have its data table exported and converted into an excel sheet for the second half of the analysis. Specify where the new excel file will be saved and select “Run”.

**Figure 11: Table to Excel Tool Window**



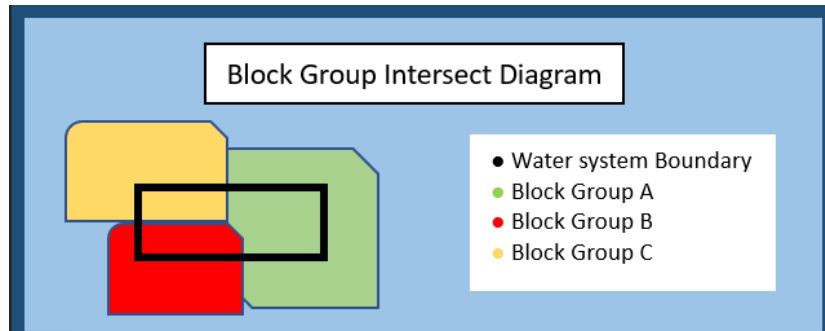
**STEP 4: EXCEL ANALYSIS**

23. For this example, the ACS data being used is Median Household Income (MHI) and MHI Margin of Error (MOE), however the same process can be applied to alternative ACS data points.

24. When SABL and the block group shapefiles are intersected they produce a series of data points where water system boundaries are split into various parcels.

- a. These parcels need to be recombined in excel for each water system.
- b. A weighted average is used to accomplish this, where the weight is the area of intersection.

**Figure 12: Illustration of the intersection between the block groups and the water system boundaries.**

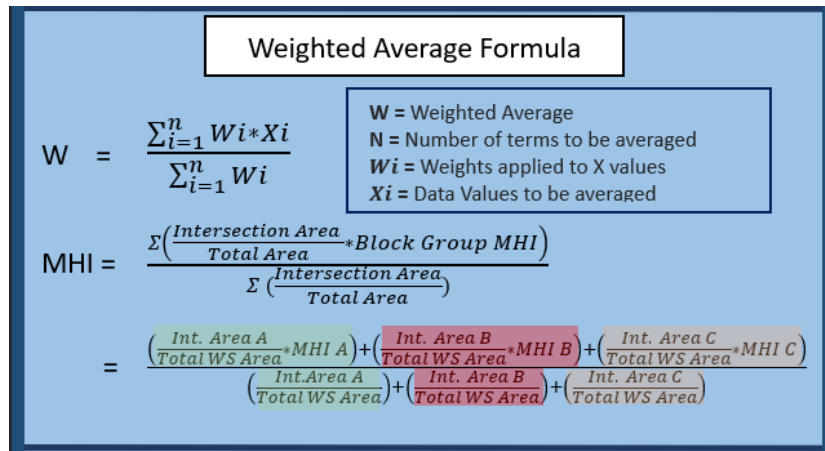


25. The weighted average is given by the following formula:

a. 
$$W = \frac{\sum_{i=1}^n W_i * X_i}{\sum_{i=1}^n W_i}$$

- b. Where:
- i. W = the weighted average
  - ii. n= the number of unique intersections within a water system
  - iii. Wi = weight (area of intersection/ total area of water system)
  - iv. Xi is the data value to be averaged (MHI in this example but other ACS data may be used here by each block group)
- c. This formula can be expanded to fit the data points here:
- i. 
$$W = \frac{\sum \left( \frac{\text{Intersection Area}}{\text{Total Area}} * \text{Block Group ACS} \right)}{\sum \left( \frac{\text{Intersection Area}}{\text{Total Area}} \right)}$$
- d. The following diagram breaks down how this formula further:

**Figure 13: Explanation of the weighted average formula used to determine MHI. The highlighted colors correspond to the block group intersecting areas seen above in Figure 12**



26. Open the excel file that was just created from the last step. Filter the following data-points from the first tab into a “Clean\_V1” new tab:
- a. PWSID
  - b. Population
  - c. Federal Classification
  - d. MHI\_E (ACS data point)
  - e. MHI\_MOE (ACS provided data point)
  - f. Parcel/part Area

**Figure 14: Snippet of the required data fields needed from the excel sheet produced by the intersection to be collected on a new tab**

PWSID	Population	Federal Classification	MHI_E	MHI_MOE	Part Area
CA3310001	270000	COMMUNITY	17269	8101	47397.81724
CA3310001	270000	COMMUNITY	17269	8101	24916784.66
CA3310007	45727	COMMUNITY	17269	8101	24916784.66
CA4910009	173648	COMMUNITY	17324	7949	232518.0242
CA1310001	26928	COMMUNITY	17328	17284	579557.6504

27. Add a new Column labelled, “MOE”.

- a. Subtract the values in MOE from the values in the ACS data column ("MHI"). This will be the set of ACS values that will be used in the final calculation.
  - b. Select for all records with an MHI estimate and copy them over to a new Tab labelled, "Clean\_V2".
28. Add a new column to the spreadsheet and label it "Total Area".
29. Select the entire sheet and insert it into a pivot table, label the new tab created as "Pivot 1".
  - a. Filter the Pivot by "PWSID" for rows and "Sum of Part Area" for values.
30. Go back to "Clean\_V2" and use the "vlookup" function to copy the values from "Pivot 1" into the column "Total Area" organized by PWSID.
31. Add new column labeled "Part/Total" and divide columns "Part Area" by "Total Area".
32. Add new Column labeled "MHI \* (Part/total)" and multiple "Part/Total" by the "MHI w/MOE".
33. Insert the whole tab into a new pivot table labeled "Pivot 2".
  - a. Organize the pivot table by selecting "PWSID" for rows and both "Sum of MHI \* Part/Total" and "Sum of Part/Total" for values.
34. Divide "Sum of MHI \* Part/Total" by "Sum of Part/Total" for each PWSID.
35. This will give you the weighted average for each water system.