

CALIFORNIA DEPARTMENT OF WATER RESOURCES

# Estimating Population Across Different Boundaries



James Common & Melika Mani

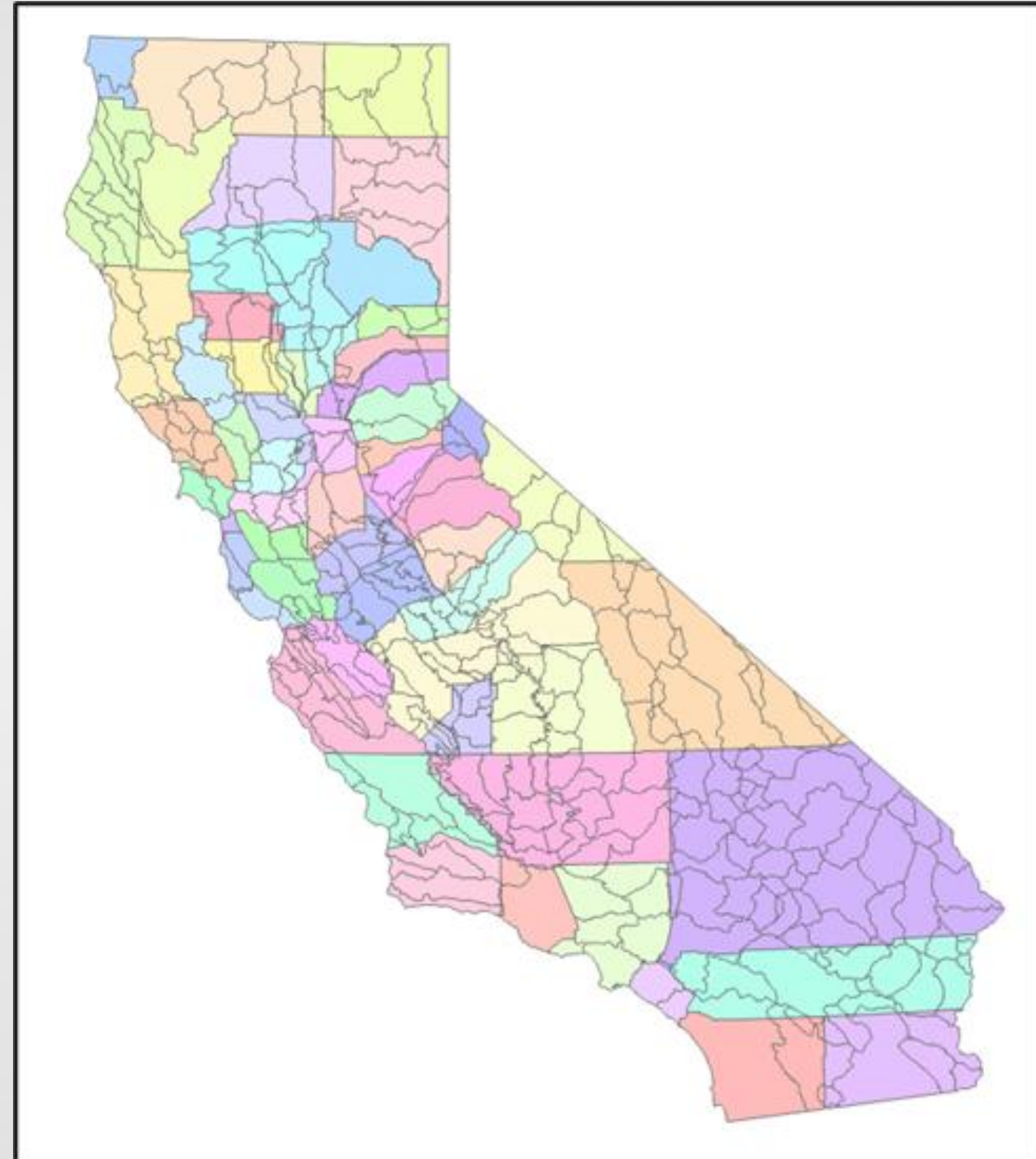
# Outline

- Past Efforts
- Existing Building Footprint Datasets
- Remotely Sensing Building Footprints
- Parcel Data and Description of Population Aggregation Process
- Tools to Automate Process
- Next Steps



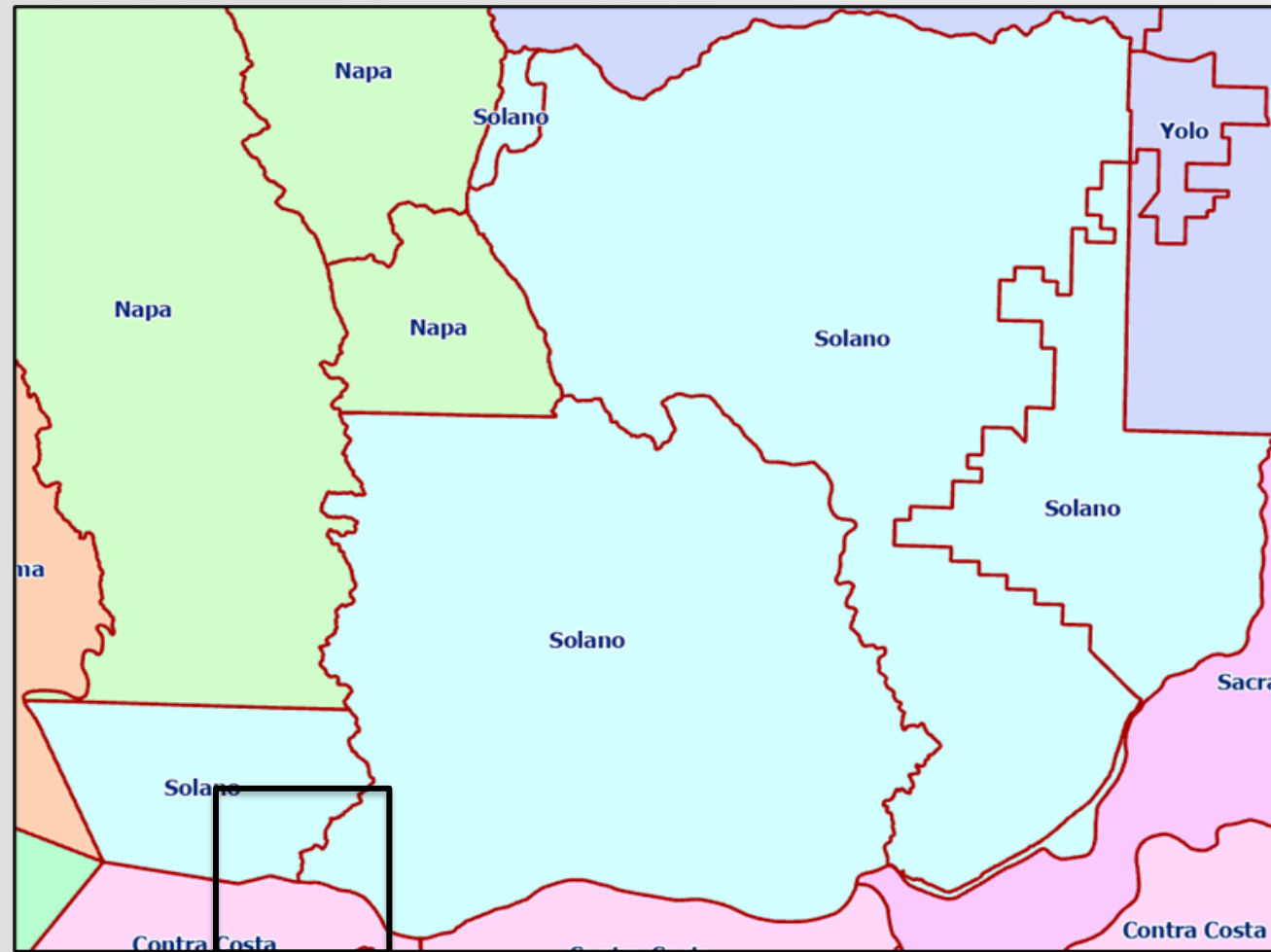
# Past Efforts

- DWR uses Detailed Analysis Units intersected with county boundaries (DAUCO) for the Water Plan water balance data
- Over 500 DAUCO's in the state
- Population estimates required for each DAUCO



# Past Efforts

- Census boundaries do not align with DAUCO
- Some census blocks are split



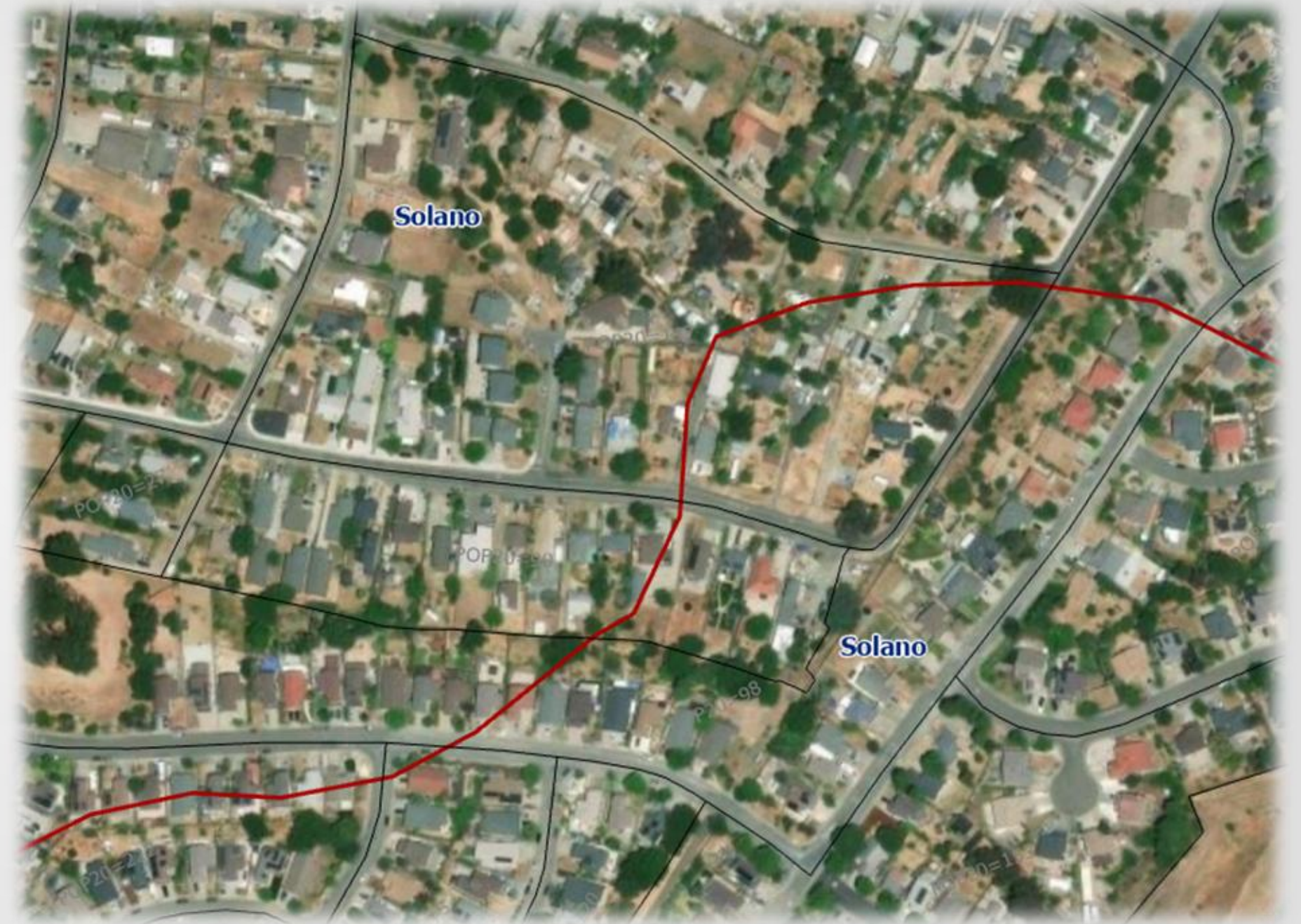
# Past Efforts

- Dark red lines represent the DAUCO boundaries
- Thin black lines are the census blocks
- Can assume the population of a census block completely inside would be 100%
- Population of split census block based on a manual count of houses (visual inspection)



# Past Efforts

- Visual inspection of split census blocks can be tedious



# Existing Building Footprint Datasets

- Available for free and covers all 58 counties
- Used by other state agencies
- Two datasets: Microsoft Global and US
  - Global version updated more frequently but contained unique errors
  - US version had building footprints in those areas but had less overall coverage for the state



# Existing Building Footprint Datasets

- Other Concerns
  - Does not contain attributes about building type or total housing units within buildings such as apartments
  - Large building footprint area could be apartments or commercial type
  - Missing building identifications in areas of heavy tree canopy



# Automate Building Footprint Extraction using ESRI Deep Learning

## Esri Building Footprint Extraction – USA

- Pretrained deep learning model for detecting buildings from high-resolution imagery
- Performs best with orthorectified 8-bit RGB (10–40 cm resolution)
- Reported average precision: **~0.72** (Esri)

## Key advantage

- Can be rerun on updated imagery

## Our Evaluation

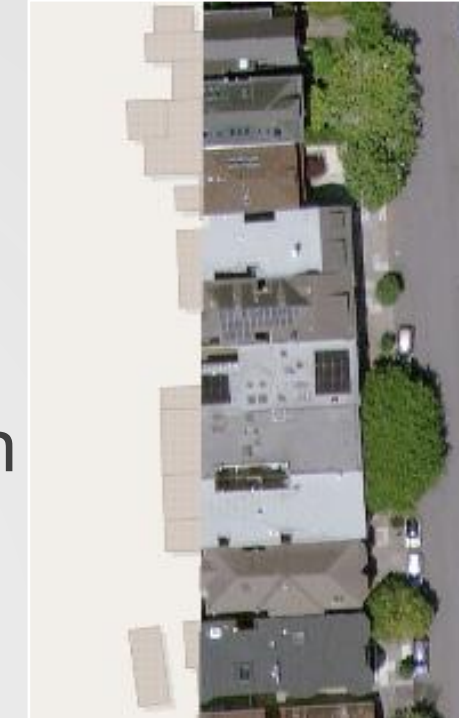
- Used NAIP 2022 imagery (60 cm) → lower resolution than model requirement
- Tested across different conditions:  
Lake Tahoe, Los Angeles, San Jose, San Francisco, Sacramento
- Captured about **60–74% of buildings** in tested areas
- Microsoft datasets captured more true buildings across all samples



# Common Issues in Building Detection

- False detections — roads, parking areas, and open land were sometimes identified as buildings
- Merged buildings — nearby or attached buildings, especially in dense urban areas, were combined into a single footprint
- Shape errors — large or irregular buildings were not captured accurately

*These issues reduced confidence in using the results for population allocation.*



# Parcel Data

- Dataset (Q2 2025 Lightbox)
- Attributes for “use code description” and for count of “housing units”
- Use codes allow for selection of only residential types

Parcels\_SOLANO

Field: Add Calculate Selection: Select By Attributes Zoom To Switch Clear Delete Copy

|       | AL_SEC_TWP_RNG | ASSR_MAP_REF | USE_CODE_MUNI_DESC     | USE_CODE_MUNI | USE_CODE_STD_LPS | USE_CODE_STD_DESC...    | Zoning | LOT_SIZE_AREA | LOT_SIZE_AREA_L ^ |
|-------|----------------|--------------|------------------------|---------------|------------------|-------------------------|--------|---------------|-------------------|
| 19371 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 5.01          | AC                |
| 19372 | all>           | <Null>       | MANUFACTURED HOME      | 1013          | 1006             | MOBILE/MANUFACTUR...    | <Null> | 5.01          | AC                |
| 19373 | all>           | <Null>       | TRANSFER BASE YR VA... | 1191          | 1001             | SINGLE FAMILY RESIDE... | <Null> | 4.89          | AC                |
| 19374 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 5.04          | AC                |
| 19375 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 5.03          | AC                |
| 19376 | all>           | <Null>       | ROW CROP               | 5052          | 7004             | CROP LAND, FIELD CR...  | <Null> | 20            | AC                |
| 19377 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 6.25          | AC                |
| 19378 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 6.25          | AC                |
| 19379 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 5             | AC                |
| 19380 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 2.49          | AC                |
| 19381 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 2.49          | AC                |
| 19382 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 2.5           | AC                |
| 19383 | all>           | <Null>       | RURAL RANCHETTE > 1... | 1012          | 1008             | RURAL/AGRICULTURAL...   | <Null> | 9             | AC                |

0 of 152,627 selected

Filters: 100%



# Parcel Data

- Intersect residential-only parcel types with census blocks
- This assigns each parcel with a census block GEOID to help relate other data
- Census blocks might cut through a parcel, but we had a requirement to have a parcel assigned to only one census block
- Used parcel centroid during the process



# Parcel Data

- Intersect the result with DAUCO boundary
- If census blocks are split by DAUCO, determine the count of parcel housing units in each DAUCO piece
- Census block population distributed to each DAUCO piece based on proportion of total housing units



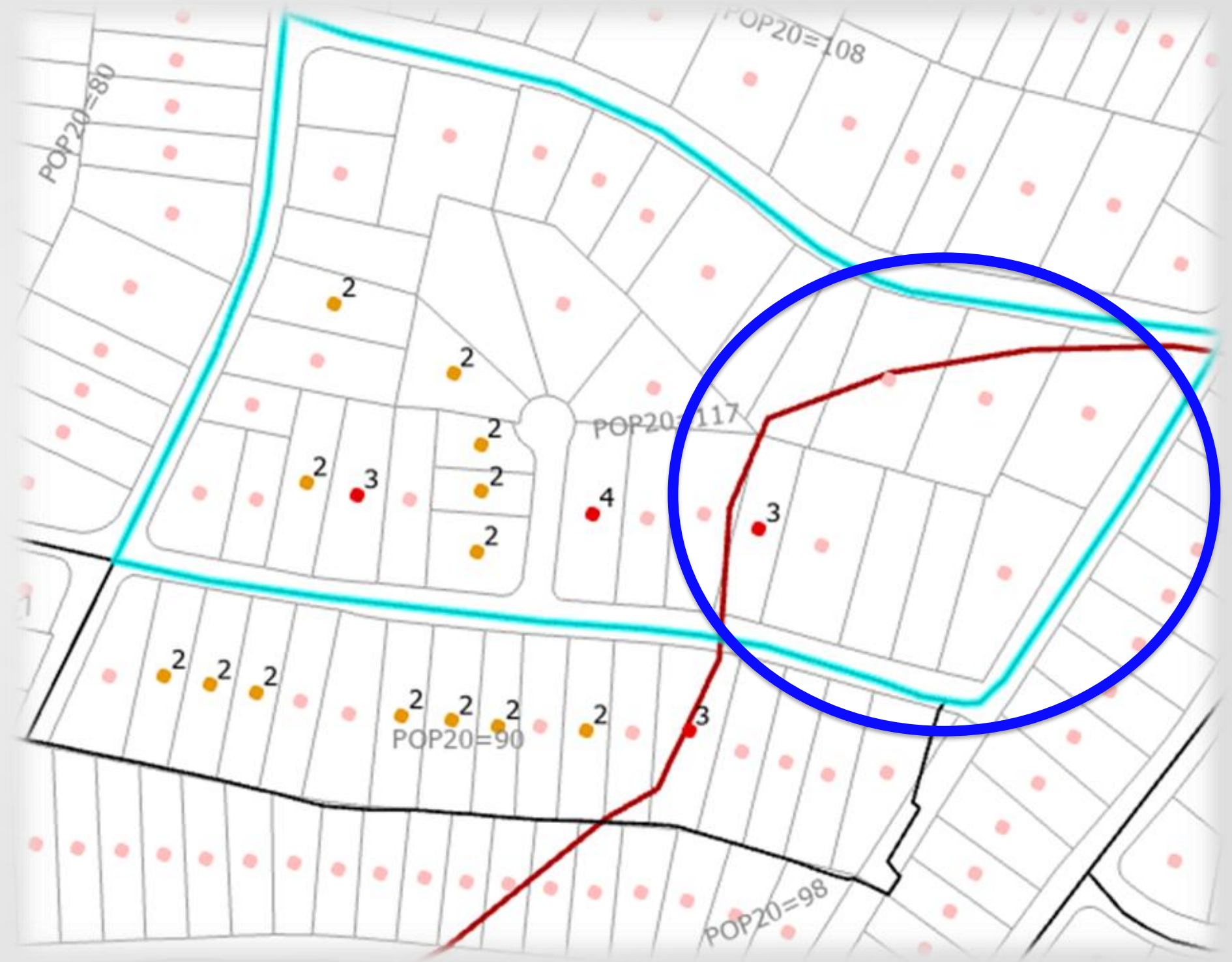
# Parcel Data

- For example, see bright blue highlighted census block



# Parcel Data

- $8/(8+34) = 19\%$
- So it was assumed 19% of the census block population is located in the smaller sized DAUCO piece on the right
- The census block population was 117 people, so 19% would be about 22 people in the smaller piece

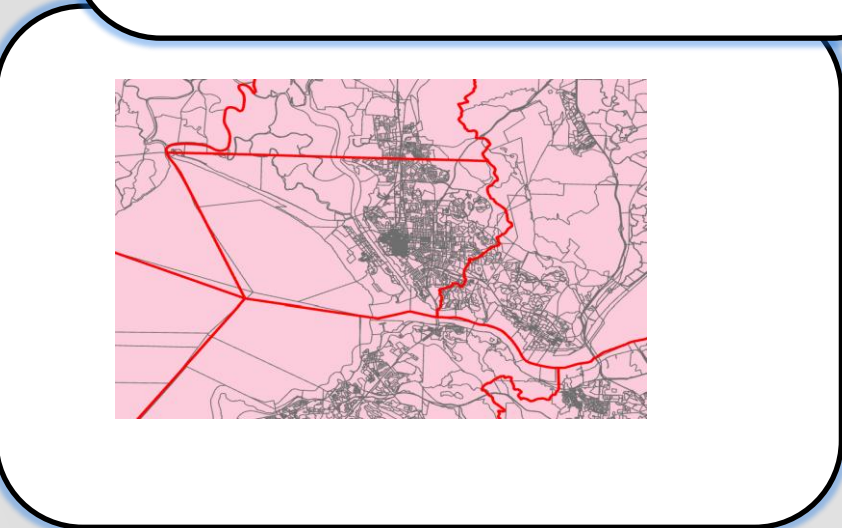


# Tool to Automate Process

- Convert manual steps into automated process using Python and ready to use ESRI geoprocessing tools



# Population Split Toolbox



Population Split Tool

box

| Target_Pop_Summary     |            |          | CB_Target_Split_All                                                    |                 |                     |                       |                           |                 |
|------------------------|------------|----------|------------------------------------------------------------------------|-----------------|---------------------|-----------------------|---------------------------|-----------------|
| Field:  Add  Calculate |            |          | Selection:  Select By Attributes  Zoom To  Switch  Clear  Delete  Copy |                 |                     |                       |                           | Rows:  Insert ▾ |
|                        | OBJECTID * | TargetID | FREQUENCY                                                              | SUM_POP20_Alloc | SUM_HOUSING20_Alloc | SUM_POP20_Alloc_Whole | SUM_HOUSING20_Alloc_Whole |                 |
| 1                      | 1          | 00125    | 201                                                                    | 821.31646       | 409.152046          | 821                   | 409                       |                 |
| 2                      | 2          | 00147    | 133                                                                    | 1282.376966     | 551.836729          | 1283                  | 552                       |                 |
| 3                      | 3          | 00247    | 199                                                                    | 1684.782669     | 868.221383          | 1684                  | 868                       |                 |
| 4                      | 4          | 00347    | 295                                                                    | 5043.524478     | 2464.555218         | 5044                  | 2465                      |                 |
| 5                      | 5          | 00447    | 929                                                                    | 21547.485631    | 10245.930866        | 21547                 | 10244                     |                 |
| 6                      | 6          | 00712    | 276                                                                    | 4934.538666     | 2045.268766         | 4934                  | 2046                      |                 |
| 7                      | 7          | 00753    | 691                                                                    | 15293.45763     | 7420.856803         | 15294                 | 7421                      |                 |
| 8                      | 8          | 01008    | 51                                                                     | 1188.902115     | 578.642631          | 1189                  | 579                       |                 |
| 9                      | 9          | 01012    | 196                                                                    | 1127.827048     | 550.925644          | 1127                  | 550                       |                 |
| 10                     | 10         | 01047    | 574                                                                    | 4123.651054     | 2407.905093         | 4124                  | 2409                      |                 |
| 11                     | 11         | 01408    | 585                                                                    | 26554.097885    | 10511.357369        | 26554                 | 10511                     |                 |
| 12                     | 12         | 01723    | 633                                                                    | 20104.533302    | 10480.718204        | 20103                 | 10481                     |                 |
| 13                     | 13         | 01823    | 374                                                                    | 7627.742538     | 4144.617139         | 7628                  | 4145                      |                 |
| 14                     | 14         | 01923    | 29                                                                     | 642.036507      | 307.882304          | 643                   | 308                       |                 |
| 15                     | 15         | 01949    | 129                                                                    | 2184.31217      | 2384.74329          | 2184                  | 2384                      |                 |
| 16                     | 16         | 02012    | 598                                                                    | 22304.823937    | 9676.009685         | 22305                 | 9676                      |                 |
| 17                     | 17         | 02112    | 221                                                                    | 3059.776262     | 1472.916521         | 3059                  | 1473                      |                 |

Target\_Pop\_Summary

CB\_Target\_Split\_All

134

Field: Add Calculate Selection: Select By Attributes Zoom To Switch Clear Delete Copy Rows: Insert

2908

581

|    | OBJECTID * | GEOID20         | TargetID | Pct   | Method   | POP20 | POP20_Alloc | POP20_Alloc_Whole | HOUSING20 | HOUSING20_Alloc | HOUSING20_Alloc_Whole |     |
|----|------------|-----------------|----------|-------|----------|-------|-------------|-------------------|-----------|-----------------|-----------------------|-----|
| 21 | 21         | 060014001001030 | 04701    | 1     | CENTROID | 204   | 204         | 204               | 87        | 87              | 87                    | 1   |
| 22 | 22         | 060014001001031 | 04701    | 1     | CENTROID | 15    | 15          | 15                | 3         | 3               | 3                     | 543 |
| 23 | 23         | 060014001001032 | 04701    | 1     | CENTROID | 11    | 11          | 11                | 3         | 3               | 3                     | 235 |
| 24 | 24         | 060014001001033 | 04701    | 1     | CENTROID | 0     | 0           | 0                 | 0         | 0               | 0                     |     |
| 25 | 25         | 060014001002000 | 04701    | 0.375 | CENTROID | 43    | 16.125      | 16                | 9         | 3.375           | 3                     |     |
| 26 | 26         | 060014001002000 | 04707    | 0.625 | CENTROID | 43    | 26.875      | 27                | 9         | 5.625           | 6                     |     |
| 27 | 27         | 060014001002001 | 04701    | 1     | CENTROID | 99    | 99          | 99                | 39        | 39              | 39                    |     |
| 28 | 28         | 060014001002002 | 04701    | 1     | CENTROID | 25    | 25          | 25                | 6         | 6               | 6                     |     |
| 29 | 29         | 060014001002003 | 04701    | 1     | CENTROID | 128   | 128         | 128               | 63        | 63              | 63                    |     |
| 30 | 30         | 060014001002004 | 04701    | 1     | CENTROID | 55    | 55          | 55                | 27        | 27              | 27                    |     |
| 31 | 31         | 060014001002005 | 04701    | 1     | CENTROID | 66    | 66          | 66                | 28        | 28              | 28                    |     |
| 32 | 32         | 060014001002006 | 04701    | 1     | CENTROID | 51    | 51          | 51                | 23        | 23              | 23                    |     |
| 33 | 33         | 060014001002007 | 04701    | 1     | CENTROID | 108   | 108         | 108               | 46        | 46              | 46                    |     |
| 34 | 34         | 060014001002008 | 04701    | 1     | CENTROID | 88    | 88          | 88                | 46        | 46              | 46                    |     |
| 35 | 35         | 060014001002009 | 04701    | 1     | CENTROID | 77    | 77          | 77                | 43        | 43              | 43                    |     |
| 36 | 36         | 060014001002010 | 04701    | 1     | CENTROID | 69    | 69          | 69                | 39        | 39              | 39                    |     |
| 37 | 37         | 060014001002011 | 04701    | 1     | CENTROID | 16    | 16          | 16                | 8         | 8               | 8                     |     |
| 38 | 38         | 060014001002012 | 04701    | 1     | CENTROID | 52    | 52          | 52                | 24        | 24              | 24                    |     |
| 39 | 39         | 060014001002013 | 04701    | 1     | CENTROID | 123   | 123         | 123               | 71        | 71              | 71                    |     |
| 40 | 40         | 060014001002014 | 04701    | 1     | CENTROID | 50    | 50          | 50                | 29        | 29              | 29                    |     |



# Population Split Toolbox

## Four Step Workflow in ArcGIS Pro

### 1 Parcel Data Preparation

Geoprocessing

Step 1 Parcel Preparing (A) — Copy Parcels + Audit Use Codes

Parameters Environments

\* Folder containing ZIP files of parcels

\* Output FileGDB (.gdb)

USE\_CODE\_FIELD (exact match)

USE\_CODE\_STD\_DESC (exact match)

Baseline use code list (C:\Users\mmuniz\Documents\Baseline use code list.txt)

☐ Rebuild output geodatabase

☒ Overwrite output

File parallel processing factor (e.g., "50%" or "4")

50%

Geoprocessing

Step 1 Parcel Preparing (B) — Build Residential Centroids

Parameters Environments

\* Parcel FileGDB (output from Tool 1, contains Parcels <COUNTY> polygon PCs)

USE\_CODE\_FIELD (exact field name, case insensitive match)

USE\_CODE\_STD\_DESC\_LPS

Use code query list (new per line, edit to add/remove)

APARTMENT HOUSE (100+ UNITS)

APARTMENT HOUSE (5+ UNITS)

APARTMENTS (GENERIC)

BOARDING HOUSE, ROOMING HOUSE, APT HOTEL, TRANSIENT LODGINGS, HOSTEL

CHILDREN'S HOME, ORPHANAGE

CLUSTER HOME (RESIDENTIAL)

COMMERCIAL/OFFICE/RESIDENTIAL (MIXED USE)

CONDOMINIUM UNIT (RESIDENTIAL)

COOPERATIVE UNIT (RESIDENTIAL)

CORRECTIONAL FACILITY, JAIL, PRISON, INSANE ASYLUM

CULTURAL, HISTORICAL (MONUMENTS, HOMES, MUSEUMS, OTHER)

DORMITORY, GROUP QUARTERS (RESIDENTIAL)

DUPLEX (2 UNITS, ANY COMBINATION)

FRATERNITY HOUSE, SORORITY HOUSE

☒ Create statewide merged centroids in the parcels GDB

Statewide centroids feature class name (if creating statewide merge)

parcels\_centroids

☒ Overwrite output

File parallel processing factor (e.g., "50%" or "4")

50%

### 2 Centroid, Census & Boundary Integration

Geoprocessing

Step 2 — Intersect Centroids

Parameters Environments

\* Centroids (points)

\* Census Blocks (polygons)

\* Target Boundary (polygons)

\* Output Geodatabase

Output name for Centroids-in-Blocks

Cen\_CB

Output name for Centroids-in-Blocks-in-Target

Cen\_CB\_TB

### 3 Compute Population Allocation Weights

Geoprocessing

Step 3 — Dissolve & Percent Split

Parameters Environments

\* Centroids intersected with CensusBlocks (Cen\_CB)

\* Centroids intersected with CensusBlocks and Target Boundary (Cen\_CB\_TB)

\* Target Boundary Unique ID Field (Dissolve Field)

Keep target fields as FIRST (Statistics Fields)

\* Output Geodatabase

Output name for Block Totals

Centroids inside CensusBlocks

Output name for Block-Target Split

CensusBlocks\_TargetBoundary\_Split\_Centroids

Mode

AUTO

### 4 Gap Filling & Final Calculations

Geoprocessing

Step 4 — Area fallback + Combine + Export

Parameters Environments

\* Centroids (points)

\* Census Blocks (polygons)

\* Target Boundary (polygons)

\* Target unique ID field (in Target Boundary)

\* Centroid\_census\_target layer dissolve (CB\_Target\_Split\_Centroid from Step 3)

\* Output Geodatabase

Output name for Area split feature class

CB\_Target\_Split\_Area

Output name for Combined table

CB\_Target\_Split\_All

Output name for Target summary table

Target Pop Summary

Census Block numeric fields to allocate (default POP20; examples: POP20, HOUSING20, HH)

☒ Calculate integer allocation fields (use for count data like POP, HH; avoid for continuous values)

Output folder for CSV exports (optional)



# Population Split Toolbox

## Step 1A–1B: audit use codes and build centroids

### Step 1 — Parcel Preparation

- Load county parcel datasets into a statewide geodatabase
- Compare parcel use codes with the baseline residential list
- Flag new or unmatched codes for review
- Update the residential list as needed

→ **Generate residential parcel centroids for next steps of processing.**

The screenshot shows the 'Geoprocessing' window for 'Step 1 Parcel Preparing (B) — Build Residential Centroids'. The 'Parameters' tab is active. The tool requires a 'Parcels FileGDB (output from Tool 1; contains Parcels\_<COUNTY> polygon FCs)' and a 'USE\_CODE\_FIELD (exact field name; case-insensitive match)' set to 'USE\_CODE\_STD\_DESC\_LPS'. A 'Use-code query list' is provided with various residential categories. Checkboxes for 'Create statewide merged centroids in the parcels GDB' and 'Env: overwriteOutput' are checked. The 'Statewide centroids feature class name' is 'parcels\_centroids' and the 'Env: parallelProcessingFactor' is '50%'.

Geoprocessing

Step 1 Parcel Preparing (A) — Copy Parcels + Audit Use Codes

Parameters Environments

\* Folder containing ZIP files of parcels

Geoprocessing

Step 1 Parcel Preparing (B) — Build Residential Centroids

Parameters Environments

\* Parcels FileGDB (output from Tool 1; contains Parcels\_<COUNTY> polygon FCs)

USE\_CODE\_FIELD (exact field name; case-insensitive match)

USE\_CODE\_STD\_DESC\_LPS

Use-code query list (one per line; edit to add/remove)

APARTMENT HOUSE (100+ UNITS)  
APARTMENT HOUSE (5+ UNITS)  
APARTMENTS (GENERIC)  
BOARDING HOUSE, ROOMING HOUSE, APT HOTEL, TRANSIENT LODGINGS, HOSTEL  
CHILDREN'S HOME, ORPHANAGE  
CLUSTER HOME (RESIDENTIAL)  
COMMERCIAL/OFFICE/RESIDENTIAL (MIXED USE)  
CONDOMINIUM UNIT (RESIDENTIAL)  
COOPERATIVE UNIT (RESIDENTIAL)  
CORRECTIONAL FACILITY, JAILS, PRISONS, INSANE ASYLUM  
CULTURAL, HISTORICAL (MONUMENTS; HOMES; MUSEUMS; OTHER)  
DORMITORY, GROUP QUARTERS (RESIDENTIAL)  
DUPLEX (2 UNITS, ANY COMBINATION)  
FRATERNITY HOUSE, SORORITY HOUSE

☒ Create statewide merged centroids in the parcels GDB

Statewide centroids feature class name (if creating statewide merge)

parcels\_centroids

☒ Env: overwriteOutput

Env: parallelProcessingFactor (e.g., "50%" or "4")

50%



# Population Split Toolbox

- **Step 2:** Assign centroids to census blocks and target areas
- **Step 3:** Calculate split percentages from centroid counts
- **Step 4:** Use area-based fallback for populated blocks with no centroids
- **Step 4:** Merge results and summarize by target area

**Step 2**

Geoprocessing

Step 2 – Intersect Centroids

Parameters Environments

\* Centroids (points)

\* Census Blocks (polygons)

\* Target Boundary (polygons)

\* Output Geodatabase

Output name for Centroids-in-Blocks  
Cen\_CB

Output name for Centroids-in-Blocks-in-Target  
Cen\_CB\_TB

**Step 3**

Geoprocessing

Step 3 – Dissolve & Percent Split

Parameters Environments

\* Centroids intersected with CensusBlocks (Cen\_CB)

\* Centroids intersected with CensusBlocks and Target Boundary (Cen\_CB\_TB)

\* Target Boundary Unique ID Field (Dissolve Field)

Keep target fields as FIRST (Statistics Fields)

\* Output Geodatabase

Output name for Block Totals  
Centroids inside CensusBlocks

Output name for Block-Target Split  
CensusBlocks\_TargetBoundary\_Split\_Centroids

Mode  
AUTO

**Step 4**

Geoprocessing

Step 4 – Area fallback + Combine + Export

Parameters Environments

\* Centroids (points)

\* Census Blocks (polygons)

\* Target Boundary (polygons)

\* Target unique ID field (in Target Boundary)

\* Centroid\_census\_target layer dissolve (CB\_Target\_Split\_Centroid from Step 3)

\* Output Geodatabase

Output name for Area split feature class  
CB\_Target\_Split\_Area

Output name for Combined table  
CB\_Target\_Split\_All

Output name for Target summary table  
Target\_Pop\_Summary

Census Block numeric fields to allocate (default POP20; examples: POP20, HOUSING20, HH)

☐ Calculate integer allocation fields (use for count data like POP, HU, HH; avoid for continuous values)

Output folder for CSV exports (optional)



# Next Steps

- As population is the most consistent indicator of urban water usage, operationalizing these tools to refine urban water use estimation would improve water accounting accuracy
- Funding permitting, utilize tools developed by this pilot to assist in Water Plan transition from DAUCO to HUC8
- Funding permitting, explore data related to Single Family and Multi Family water use to refine the estimation of those datasets in the Water Plan data

