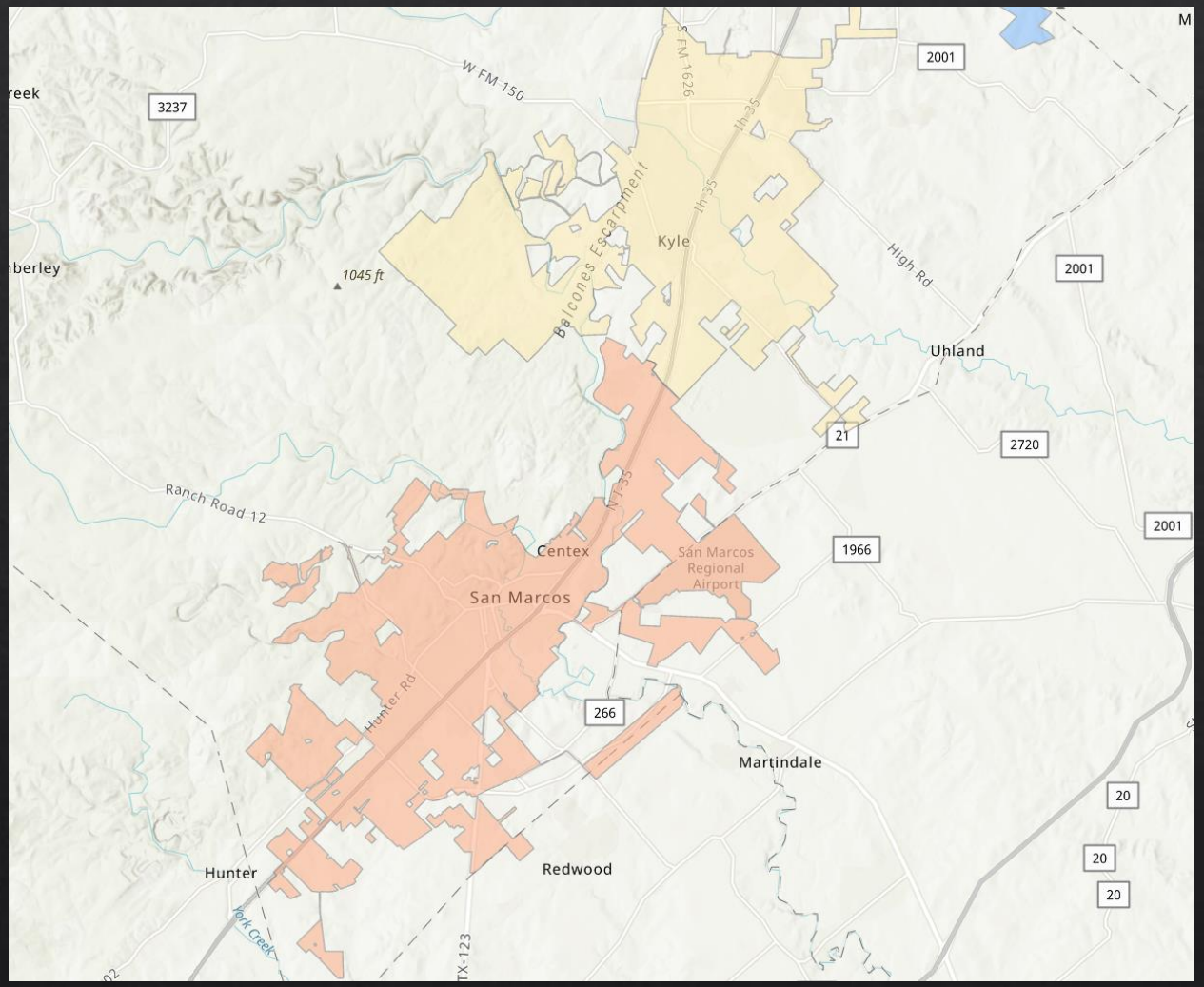

Municipal Area Population Estimation Through Areal Interpolation of Census Data

Chad Ramos | Kennedy Jackson
GEO 5419
5/3/23

TEXAS STATE
GEOGRAPHY AND ENVIRONMENTAL STUDIES

Introduction

- Population estimates of municipalities play a critical role in a broad range of planning decisions (Deng, Wu, and Wang 2010; Hoque 2012; Smith and Cody 2013).
- Grants, locating public facilities, schools, housing developments, transportation routes, hospitals, etc. all require population estimates (Wu et al. 2005; Smith and Cody 2013).
- Population estimates are used by local businesses, large corporations, and federal, state, and local governments (Deng, Wu, and Wang 2010; Hoque 2012; Smith and Cody 2013).

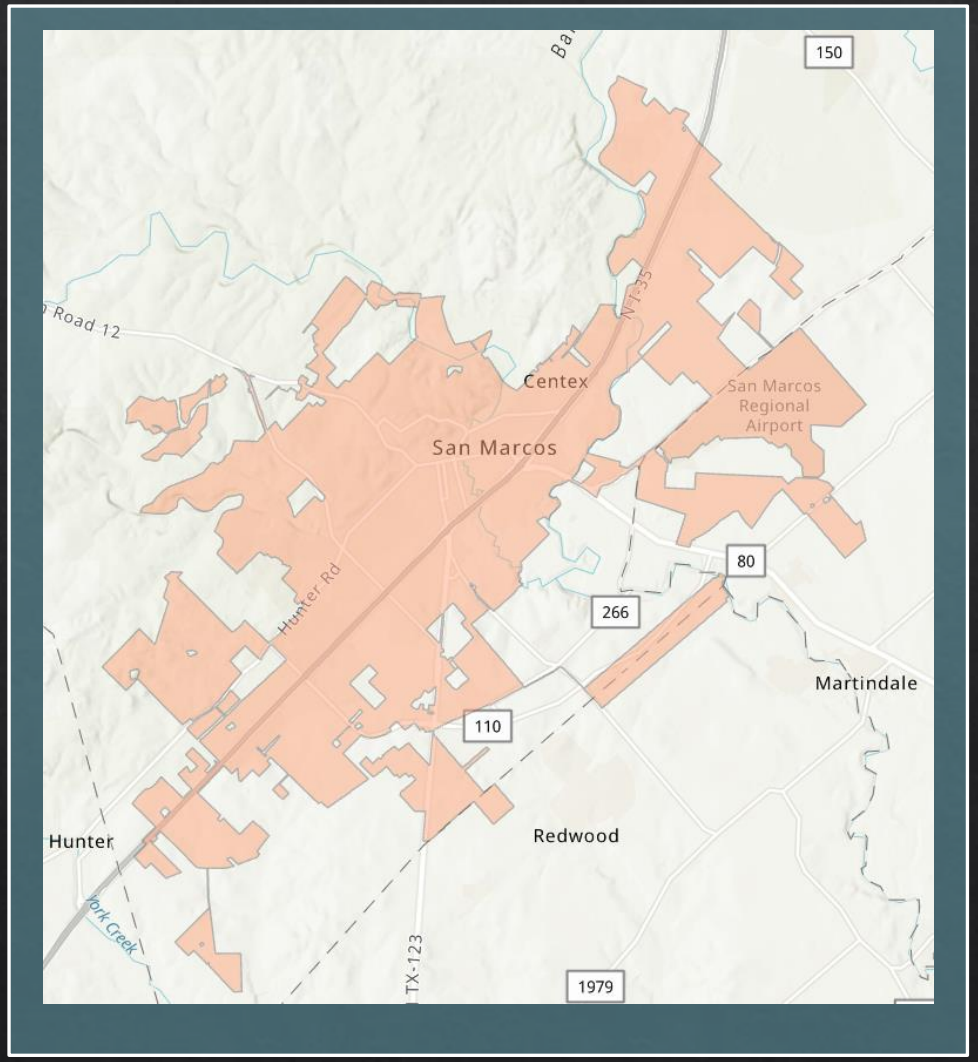

Figure 1. San Marcos city limits illustrating complexity of municipal boundaries. Map created by authors.

Housing Unit Method

- The most popular method of population estimation is the **Housing Unit Method (HU)**, used by municipal governments, state agencies, and the Census Bureau (Deng, Wu, and Wang 2010, Smith and Cody 2013).
- The HU method is based on the number of housing units in a location and the average number of people living in each unit. This is a popular technique because it is simple and can adapt to various data sources (Deng, Wu, and Wang 2010 , Hoque, 2012, Smith and Cody 2013).

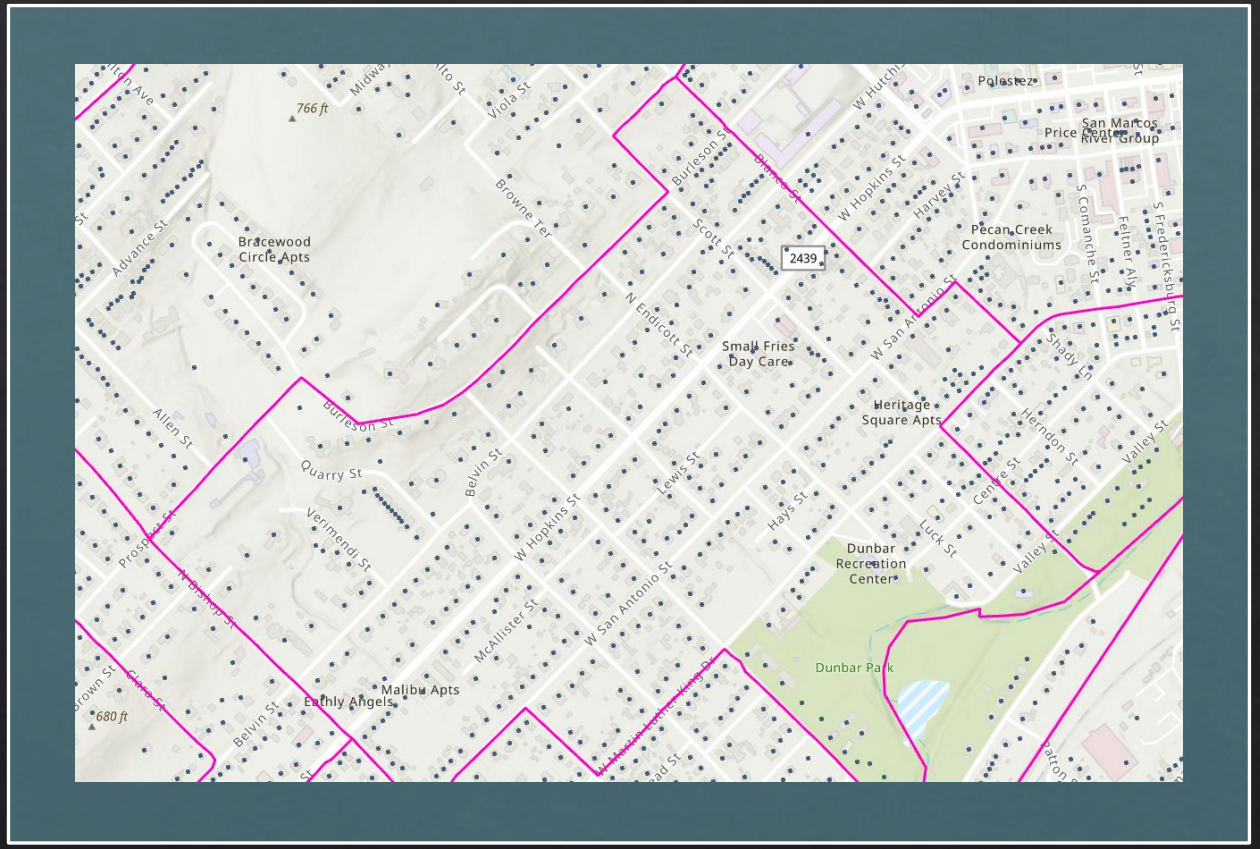

Figure 2. Address points in Census Block, San Marcos, Tx. Map created by authors.

Housing Unit Method

- Municipalities have data on the number of HUs and the number of group housing units-apartments etc. (GH) (Smith and Cody 2013).
- They rely on surveys like the Census or ACS for the number of persons per household (PPH), and the vacancy rate (VR) (Smith and Cody 2013).
- The calculation can become tedious as each Census designated area can have a different PPH and VR (Smith and Cody 2013).

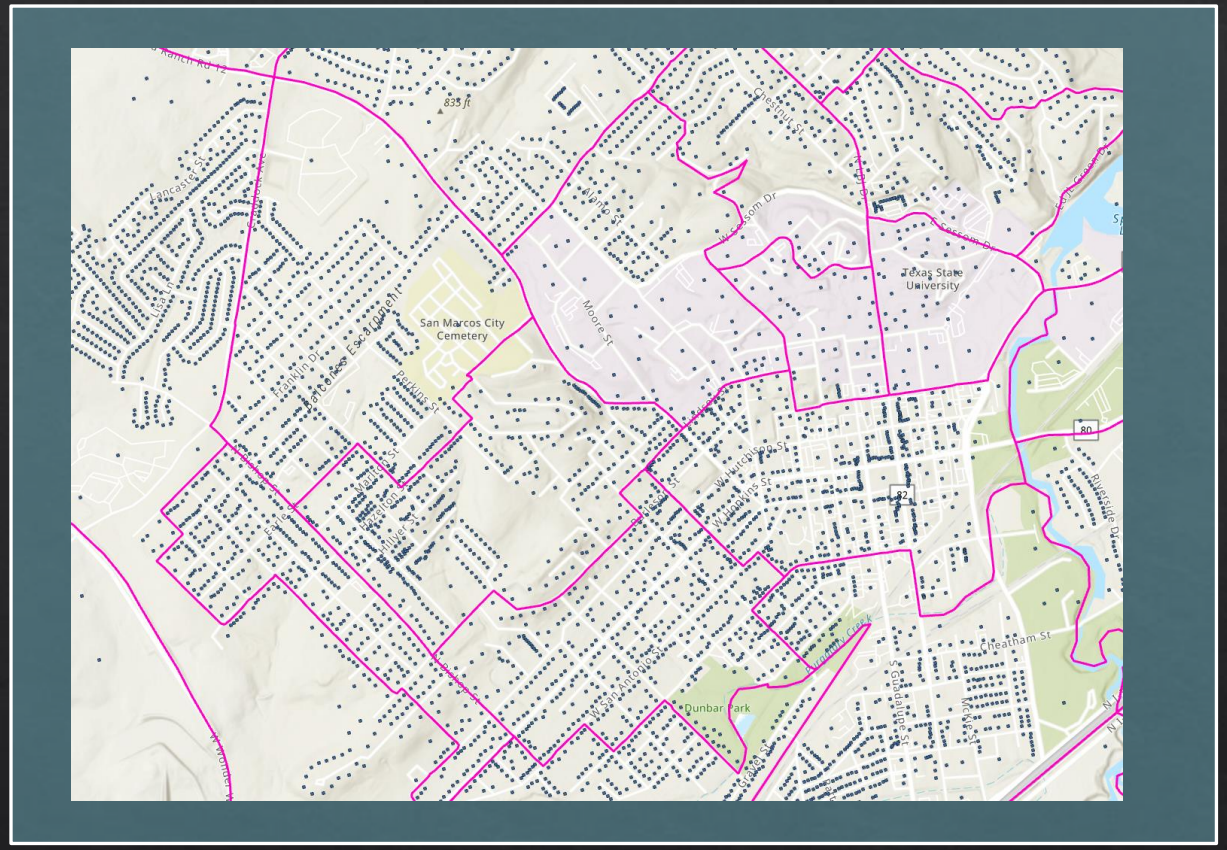

Figure 3. Address points in Census Block illustrating difficulty in calculating population by HU method. Map created by authors.

$$\text{Population} = (\text{HUs} * \text{PPH} * \text{VR}) + (\text{GH} * \text{PPH} * \text{VR})$$

For each Census Area

Objective

- Use Areal Interpolation of Census Data to create an accurate and user-friendly method of small area population estimation, targeted at Municipal Government use.

Why

- More efficient than Housing Unit Method
- Uses Authoritative Data Source by default
- Versatile – calculate populations for subset of city limits

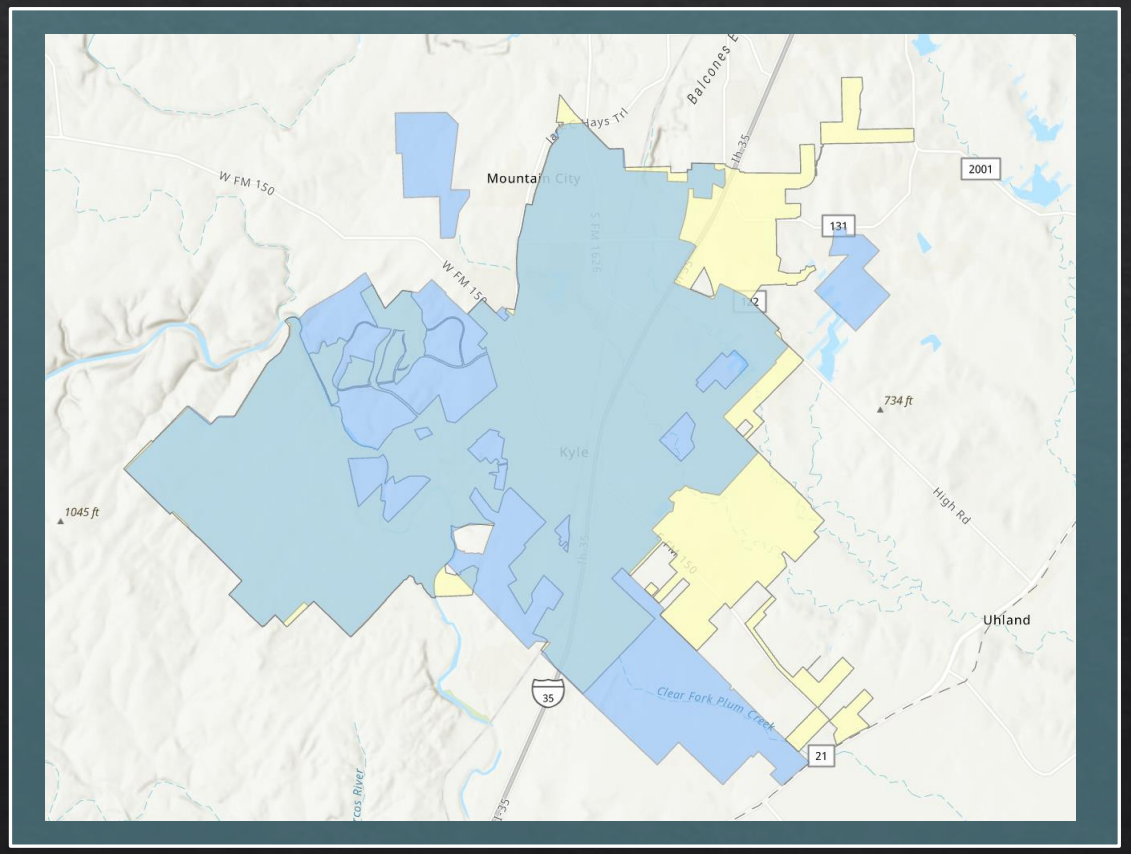

Figure 4. City of Kyle Water CCN and city limits. Map created by authors.

Data Collection

Data	Source
Test City Limits Shapefile-Kyle	https://city-of-kyle-maps-giskyle.hub.arcgis.com/datasets/GISKyle::jurisdiction-1/explore?location=29.986848%2C-97.776775%2C11.69
Test City Limits Shapefile-San Marcos	https://data-cosm.hub.arcgis.com/datasets/city-limits/explore?location=29.868371%2C-97.930650%2C11.91
Test City Limits Shapefile-Austin	https://data.austintexas.gov/dataset/BOUNDARIES_jurisdictions/3pzb-6mbr
Default population data	https://www.arcgis.com/home/item.html?id=b3642e91b49548f5af772394b0537681#overview
Authoritative Population Estimates to compare accuracy-Texas Demographic Center.	https://demographics.texas.gov/data/tpepp/estimates/

Methodology

- Sum Block Level (or user input) Population data for all Blocks within a user-specified area, with special consideration for edge cases.

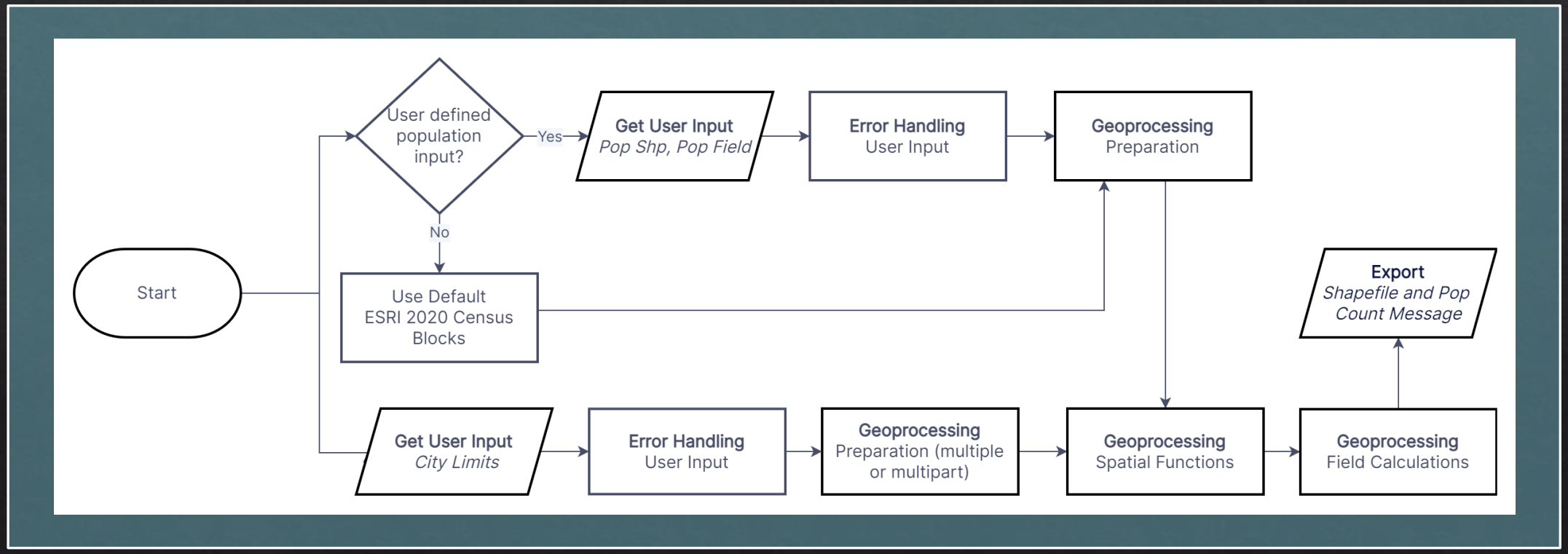

Figure 5. Conceptual flowchart of research script. Created by authors.

Pseudocode

1. Get input parameters
   a. Required user input study area
   b. Optional study area dissolve field
   c. Optional user submitted population data
   d. Optional field containing population data
   e. Required output table
2. Set path for ESRI Living Atlas population data
3. Set ArcGIS Pro Project as CURRENT and get map as Active Map
4. Perform error checks and provide messages on user input study and dissolve field
5. Perform error checks and provide messages on user input population data, if provided
6. Import ESRI data if no user submitted population data
7. Set variables and pass error-checked input parameters into tabulate intersections tool
8. Pass output from tabulate intersections into statistics tool to sum and group by
9. Add and drop fields to clean output table
10. Add output table to map

Results-Tool Function

- Functions as expected, provides population estimation for user-input study area
- Provides warning and error messages for bad input data
- Runtimes for multiple study areas in one shapefile can be excessively high
- Tabulate intersections tool will only output 5 features at a time, despite no indication of this limitation in the documentation

Results-Edge Case Methodology

- Edge cases are accounted for by apportioning the population of each edge-case census block to the study area based on the proportion of the overlap between the edge-case census block and the study area
- Could be argued that census blocks are small enough and aligned closely enough to real borders that edge cases can be ignored
- The results to the right show the population estimates for the two extreme cases of ignoring edge cases.
- Indicates that the methodology presented by the research makes a significant difference in population estimates.

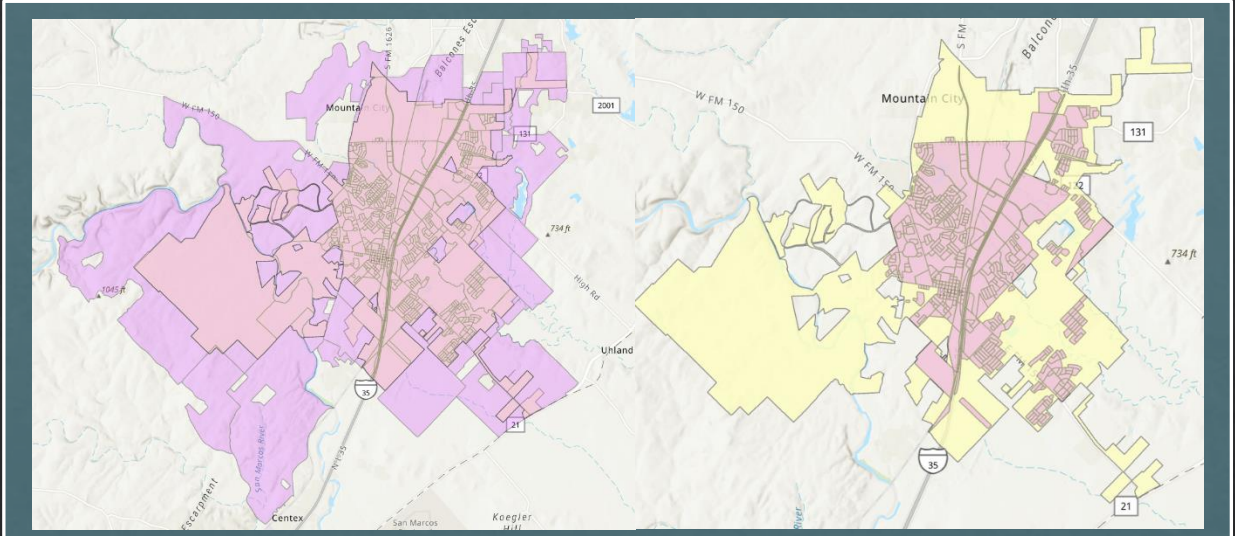

Figure 6. The two extremes of ignoring edge cases: selecting all census blocks intersecting the study area (left) and selecting only census blocks completely within the City of Kyle study area (right). Map created by authors.

Estimation Type	Population	% Difference
Texas Demographic Center (Authoritative)	45,697	n/a
All Intersecting Blocks	54,942	20.23%
Completely Within	36,604	-24.84%
Research Tool	46,204	1.10%

Table 2. Population estimates based on extremes of edge cases.

Results-Tool Accuracy

- Overall, results show that the tool can be highly accurate when compared to authoritative population estimates done by the Texas Demographic Center
- Most cities show accuracy within 3%
- Some cities above 10% error and some over 50% error.
- Of the 50 cities tested, all with error above 10% are border cities

City	Texas Demographic Center Estimate	Tool Estimate	Percent Difference
McAllen	142210	142207	0.01
Laredo	255205	255279	0.03
Buda	15108	15114	0.04
Lockhart	14379	14371	0.06
Pharr	79715	79658	0.07
Edinburg	100243	100148	0.09
San Juan	35294	35330	0.1
Mission	85778	85686	0.11
Benbrook	24520	24489	0.13
Taylor	16267	16295	0.17
Weslaco	40160	40089	0.18
Conroe	89956	90141	0.21
San Marcos	67553	67801	0.37
Alamo	19493	19413	0.41
Mercedes	16258	16330	0.44
La Villa	2804	2818	0.5
Kerrville	24278	24431	0.63
Palmview	15830	15716	0.72
Waco	138486	139733	0.9
Granjeno	283	286	1.06
Kyle	45697	46204	1.11
Austin	961855	943826	1.87
Sullivan City	3908	3982	1.89
Progreso	4807	4651	3.25
Elsa	5668	5470	3.49
Hidalgo	13964	13423	3.87
Rio Bravo	4450	4228	4.99
Penitas	6460	6054	6.28
El Cenizo	2540	2317	8.78
Pleasanton	10648	9679	9.1
La Joya	4457	4907	10.09
Edcouch	2732	3096	13.32
Roma	11561	9002	22.13
Rio Grande City	15317	11194	26.92
Escobares	2588	11 39	55.99
La Feria	6817	2344	65.62
Santa Rosa	2450	831	66.08

Conclusions

- Overall, results show that the tool can be highly accurate when compared to authoritative population estimates done by the Texas Demographic Center
  Most cities show accuracy within 3%

- Some cities above 10% error and some over 50% error.

- Of the 50 cities tested, all with error above 10% are border cities

Demo

Work Cited

Deng, C., C. Wu, and L. Wang. 2010. Improving the housing-unit method for small-area population estimation using remote-sensing and GIS information. *International Journal of Remote Sensing* 31 (21):5673–5688. https://www.tandfonline.com/doi/full/10.1080/01431161.2010.496806 (last accessed 20 March 2023).

Hoque, N. 2012. Evaluation of small area population estimates produced by Housing Unit, Ratio-correlation and Component Method II compared to 2000 Census counts. *Canadian Studies in Population* 39 (1–2):91. https://journals.library.ualberta.ca/csp/index.php/csp/article/view/17838 (last accessed 20 March 2023).

Smith, S. K., and S. Cody. 2013. Making the Housing Unit Method Work: An Evaluation of 2010 Population Estimates in Florida. *Population Research and Policy Review* 32 (2):221–242. http://link.springer.com/10.1007/s11113-012-9265-2 (last accessed 20 March 2023).

Wu, S., X. Qiu, and L. Wang. 2005. Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing* 42 (1):80–96. https://www.tandfonline.com/doi/full/10.2747/1548-1603.42.1.80 (last accessed 15 March 2023).