06 – Spatial Interpolation with Grids

J.S. Onésimo Sandoval
Sociology
Saint Louis University

Outline

- What is Spatial Interpolation?
- Modifiable Areal Unit Problem (MAUP)
- How to make a grid
- Assigning discrete or count variables to the grid cell
- Assigning ratio scales variables to the grid cell

What is Spatial Interpolation?

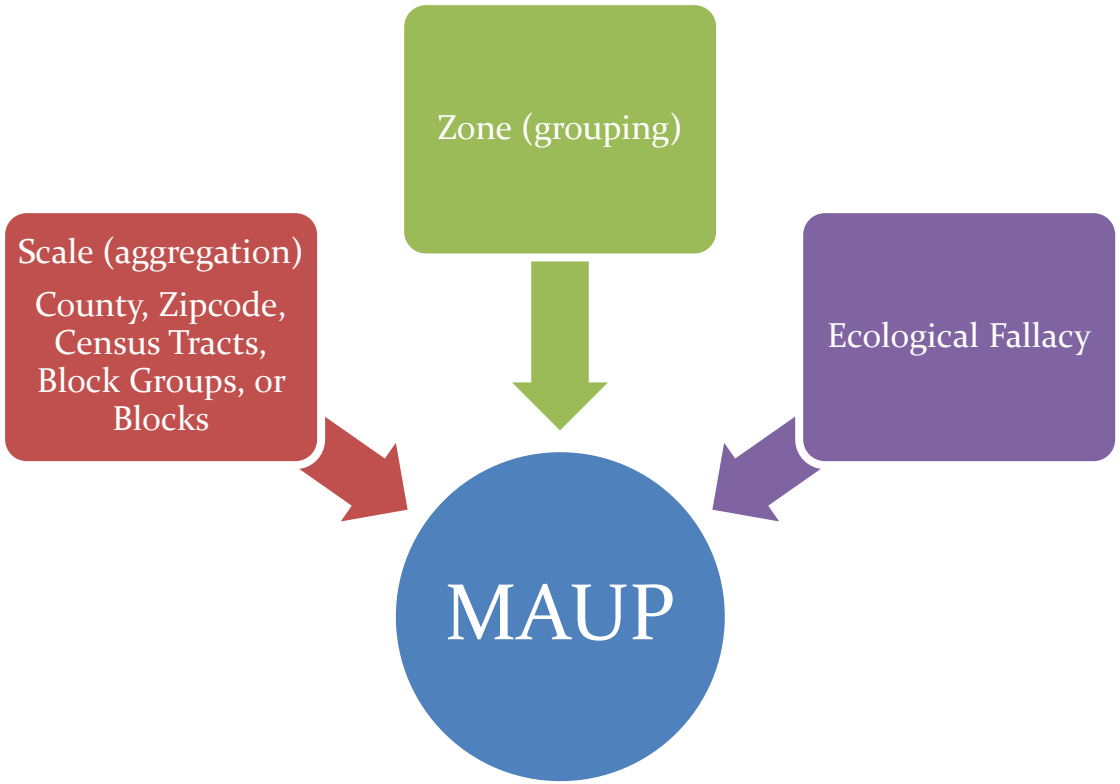
Spatial Interpolation

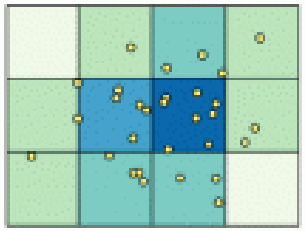
- Interpolation=Estimate
  - Estimate the unknown data values for specific locations using the known data values
- In a perfect world we would work with point data, however, most of our data is in the form of a shapefile
- Most ecological, environmental, economic, and social data represents continuous and dynamic values
- Time (data is fluid)
  - Spatial-temporal interaction effects
- Standardize variables

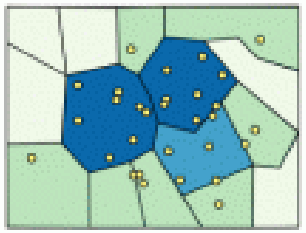
Modifiable Areal Unit Problem (MAUP)

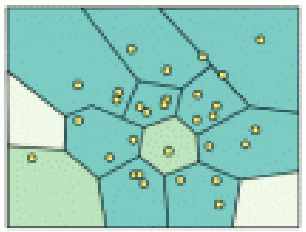
Modifiable Areal Unit Problem (MAUP)

- The polygons we use are artificial spatial representation of continuous geographical phenomena
- Results may depend on the specific geographic unit



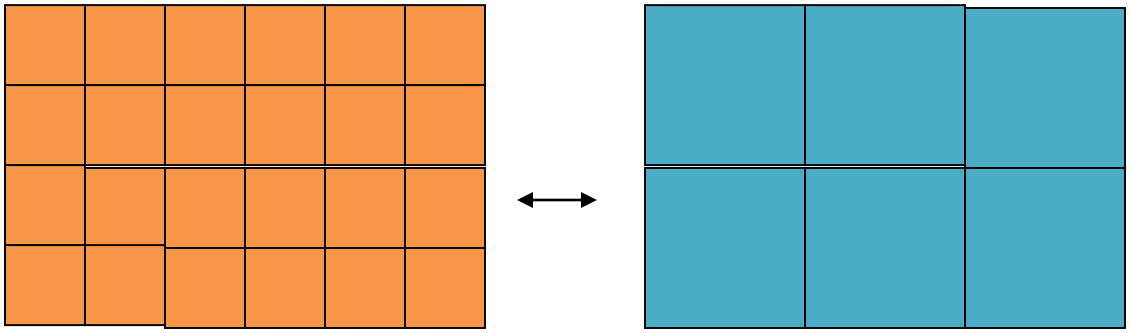






Modifiable Areal Unit Problem (MAUP)

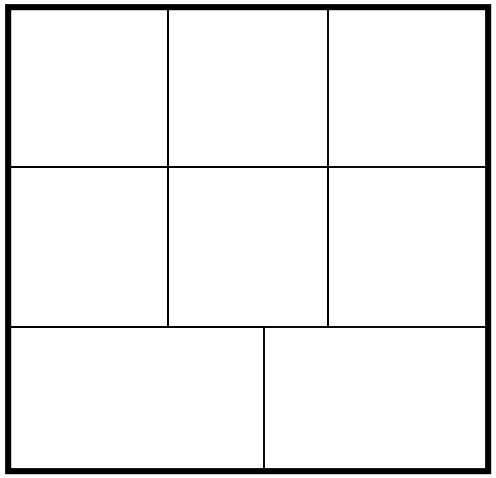
- Scale Issue: involves the aggregation of smaller units into larger ones. Generally speaking, the larger the spatial units, the stronger the relationship among variables.

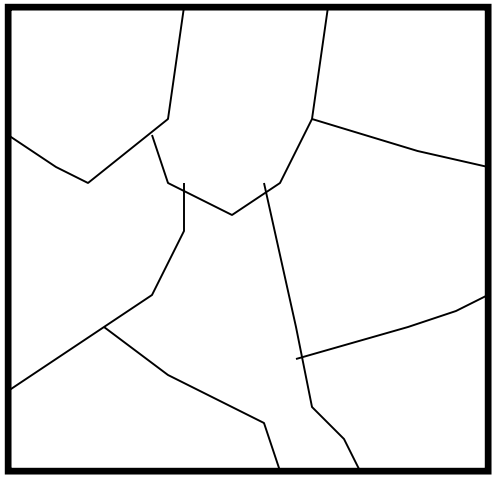


Aggregation (smoothed)

Modifiable Areal Unit Problem (MAUP)

- Modifiable Area or Group: Units are arbitrary defined and different organization of the units may create different analytical results.

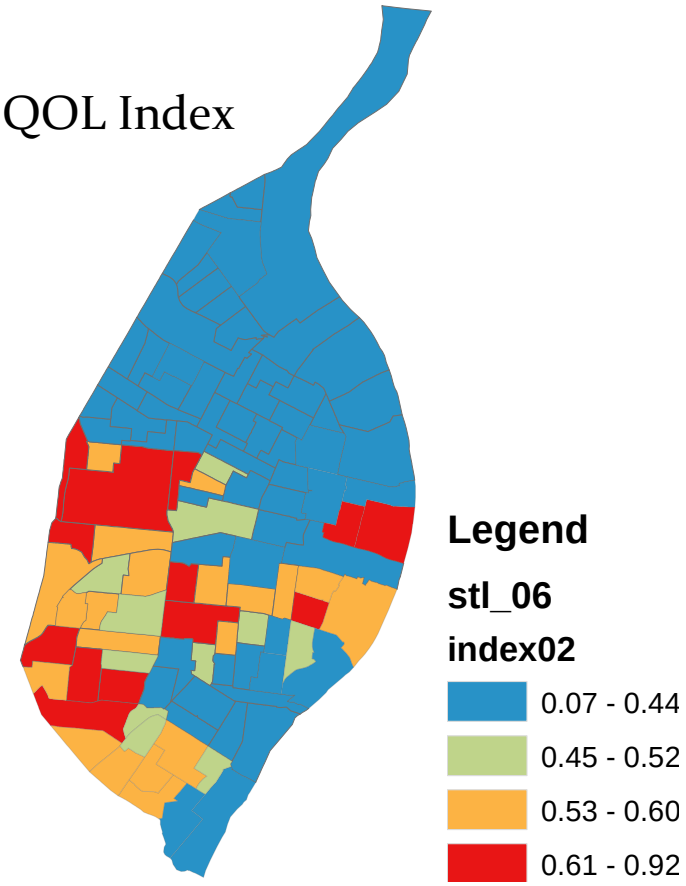




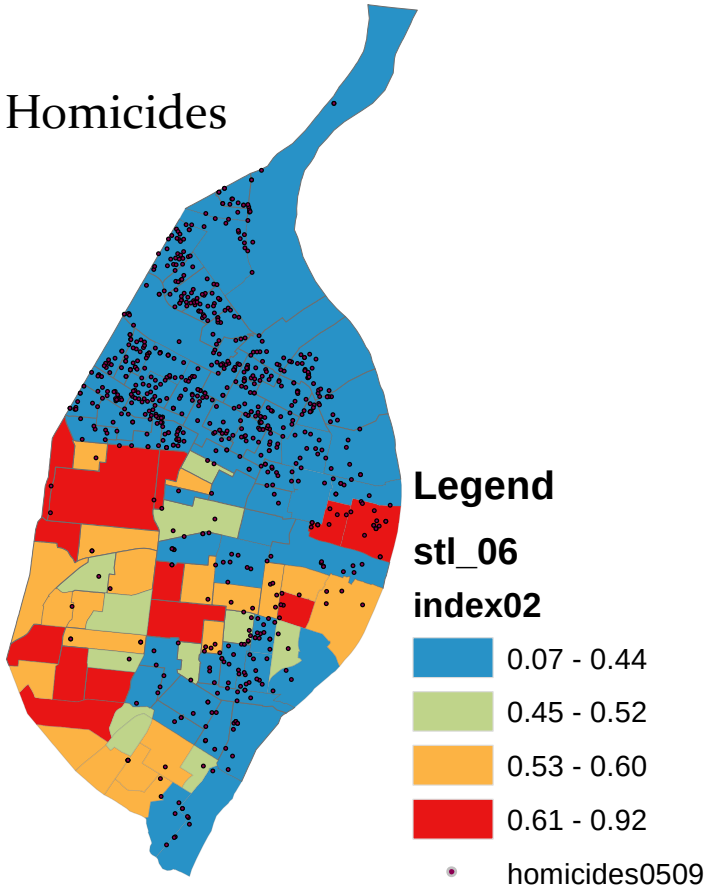
Modifiable Areal Unit Problem (MAUP)

- Ecological Fallacy: Results from aggregated data (e.g. census tracts) cannot be applied to individual people
- Cannot assume the people in blue areas commit crimes

QOL Index



Homicides

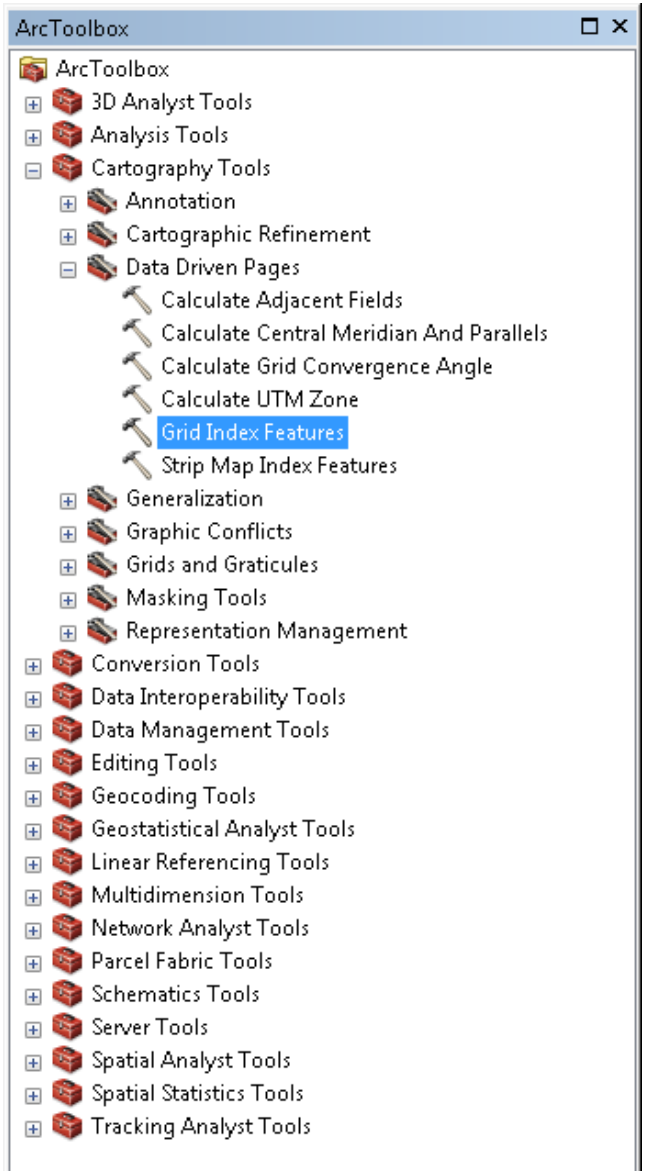


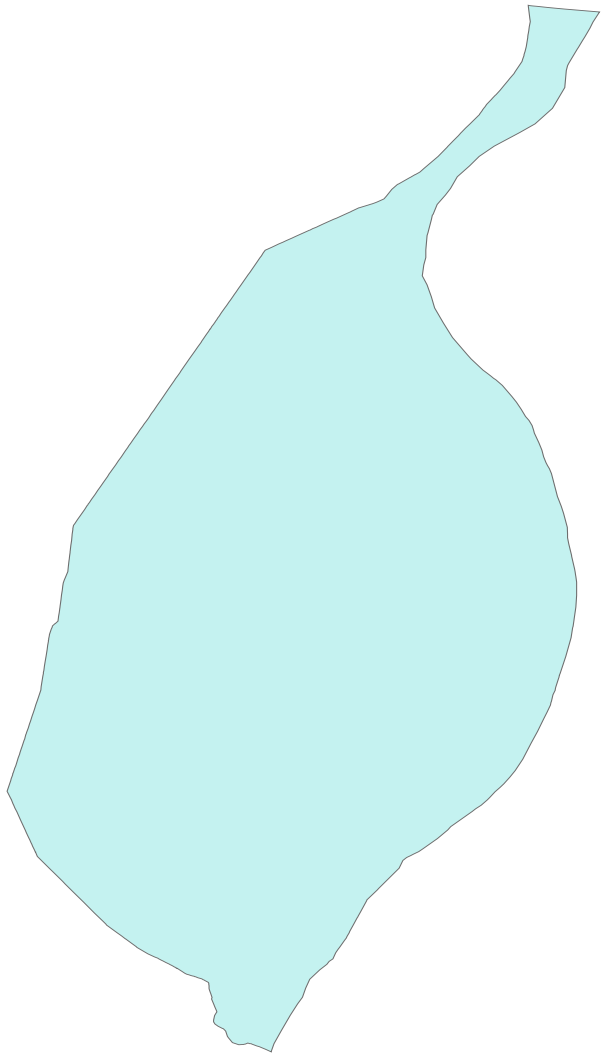
Solutions

- Standardized space over time – we need to compare apples to apples
  - One way to do is to normalize the 2000 boundaries to 2010.
  - A second way to do to this normalize the boundaries to standard space
- Advantages of the grid
  - We can synthesis social and economic data with raster images
  - We can avoid some potential pitfalls in bias with our statistics tests related to sample size and neighbors

How to make a grid

1. Define your study area
2. ArcToolbox->Cartography Tools->Data Driven Pages -> Grid Index Features



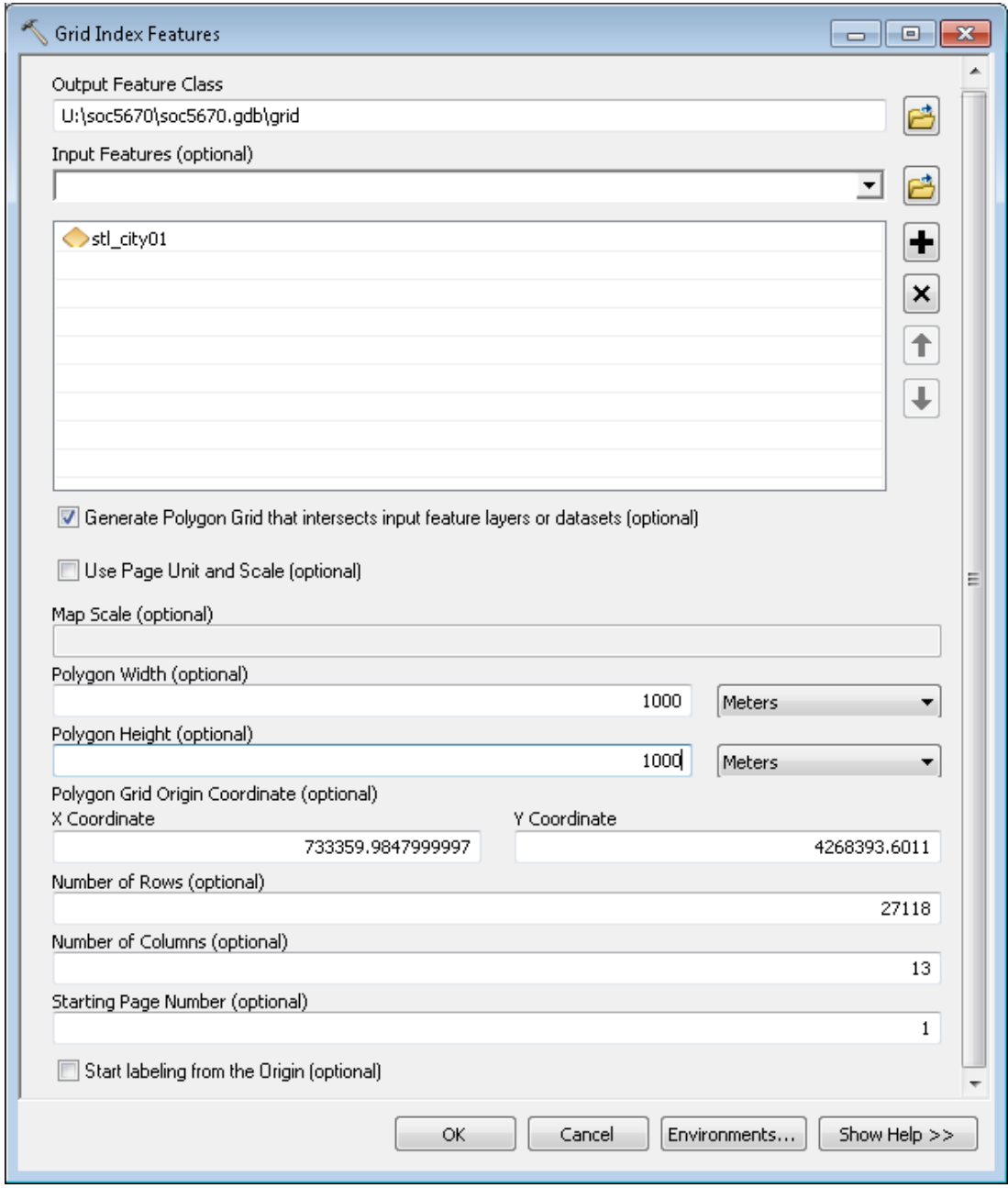


1. Once you select your study you need to name the new feature file

2. I typically use the “Generate Polygon Grid” option

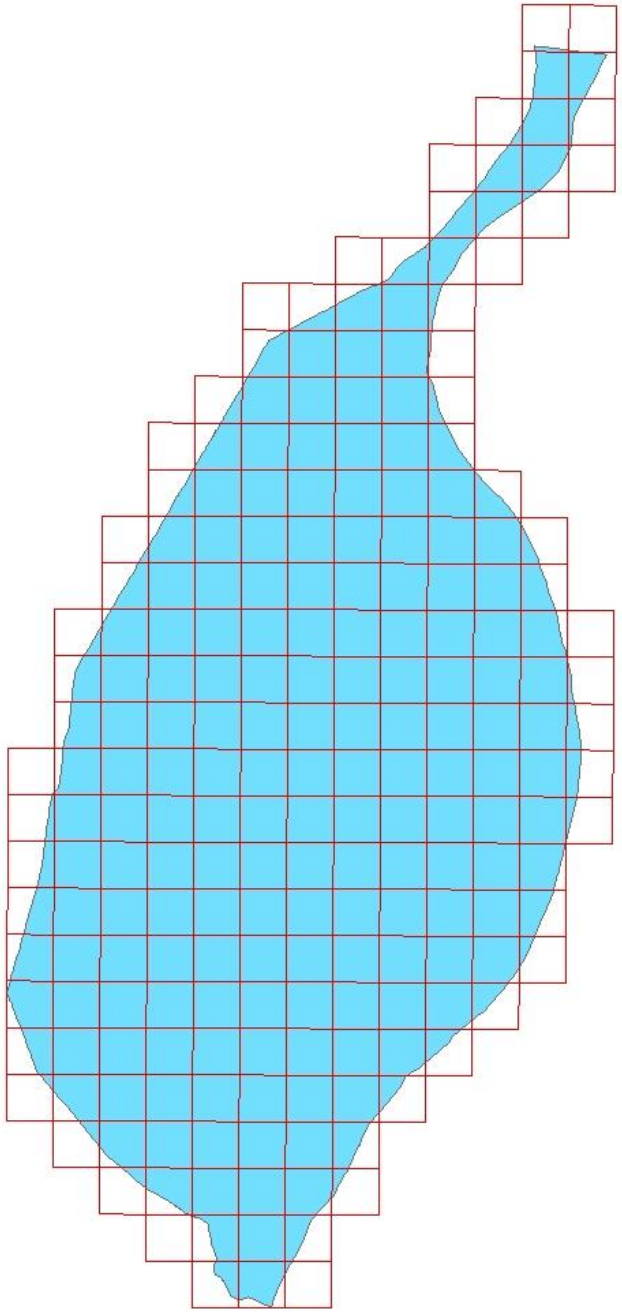
3. Now you have the flexibility to experience with different sizes for the grids. Remember we are working in meters

4. Everything also should populate as a default



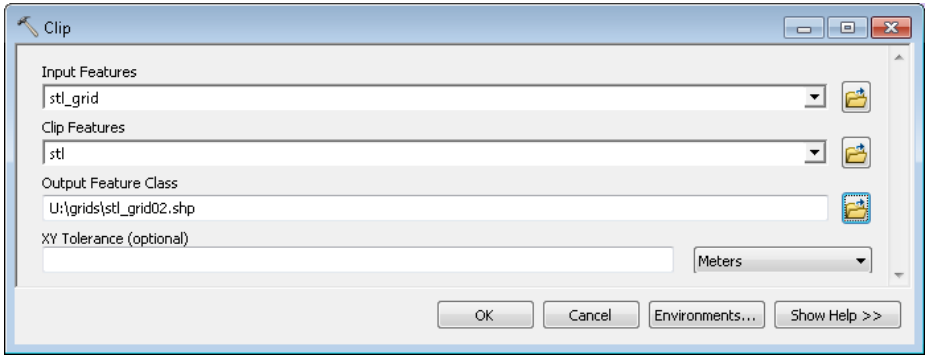
1. You will have something like this image to the right

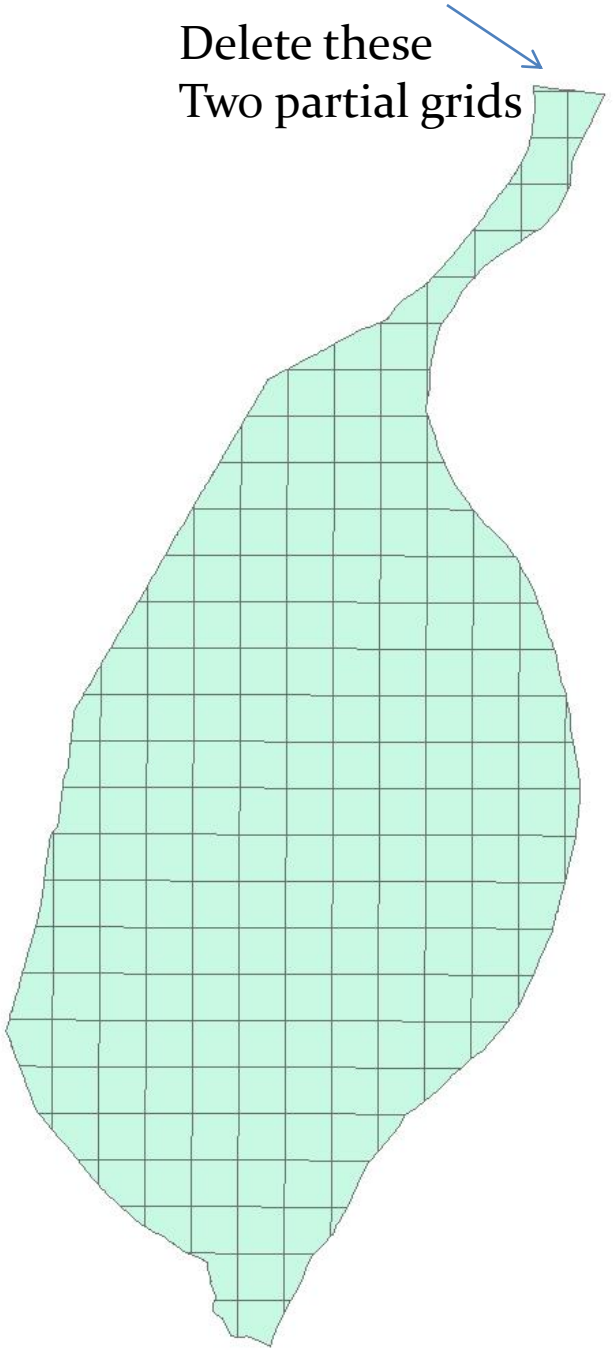
2. At this point you should inspect the partial grids and determine if you want to keep them or delete them



1. Clip the grids to the STL City boundaries

2. N=215 - You will get partial grids. Note I will delete two very small partial grids at the top of the city

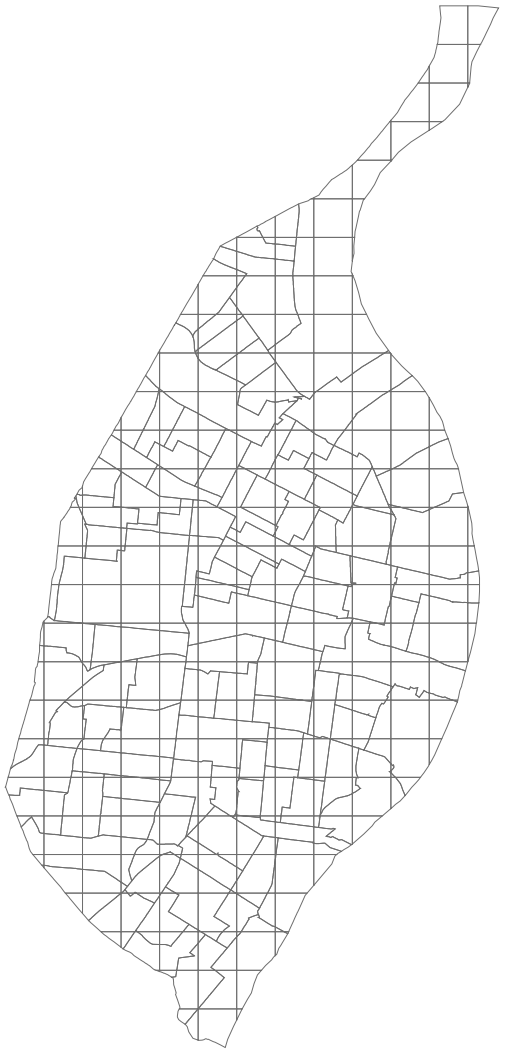


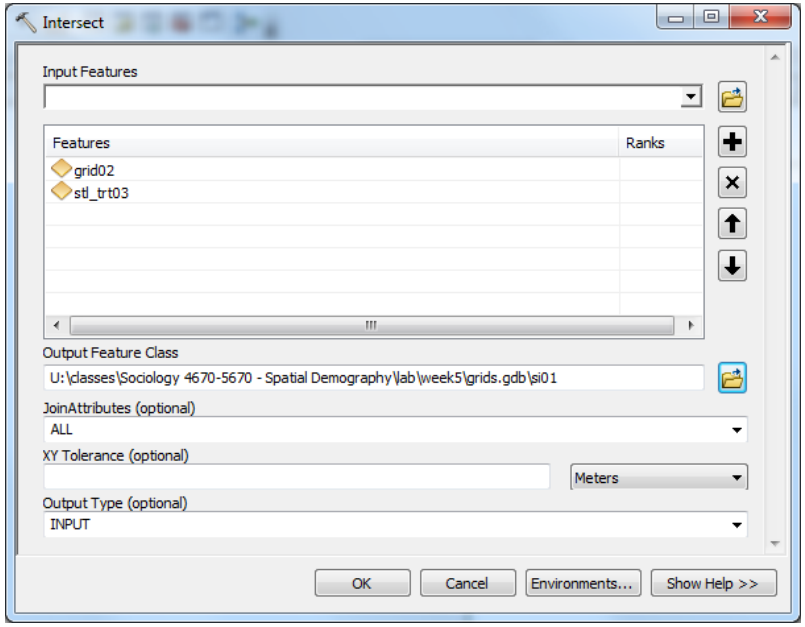


Assigning discrete or count variables to the grid cell

Step 1 – use the intersect function to intersect the grids with the census tracts

You have 634 unique polygons





Table

si01

OBJECTID *	Shape *	FID_grid02	PageName	PageNumber	FID_stl_trt03	STATEFP	COUNTYFP	TRACTCE	GEOID	NAME	NAMELSAD	M
1	Polygon	3	B12	3	69	29	510	127000	29510127000	1270	Census Tract 1270	G
2	Polygon	4	B13	4	69	29	510	127000	29510127000	1270	Census Tract 1270	G
3	Polygon	5	C11	5	69	29	510	127000	29510127000	1270	Census Tract 1270	G
4	Polygon	6	C12	6	69	29	510	127000	29510127000	1270	Census Tract 1270	G
5	Polygon	7	C13	7	69	29	510	127000	29510127000	1270	Census Tract 1270	G
6	Polygon	8	D10	8	69	29	510	127000	29510127000	1270	Census Tract 1270	G
7	Polygon	9	D11	9	69	29	510	127000	29510127000	1270	Census Tract 1270	G
8	Polygon	10	D12	10	69	29	510	127000	29510127000	1270	Census Tract 1270	G
9	Polygon	11	D13	11	69	29	510	127000	29510127000	1270	Census Tract 1270	G
10	Polygon	12	E10	12	69	29	510	127000	29510127000	1270	Census Tract 1270	G
11	Polygon	13	E11	13	69	29	510	127000	29510127000	1270	Census Tract 1270	G
12	Polygon	14	E12	14	69	29	510	127000	29510127000	1270	Census Tract 1270	G
13	Polygon	15	F8	15	69	29	510	127000	29510127000	1270	Census Tract 1270	G
14	Polygon	16	F9	16	69	29	510	127000	29510127000	1270	Census Tract 1270	G
15	Polygon	17	F10	17	69	29	510	127000	29510127000	1270	Census Tract 1270	G
16	Polygon	18	F11	18	69	29	510	127000	29510127000	1270	Census Tract 1270	G
17	Polygon	19	G6	19	55	29	510	108200	29510108200	1082	Census Tract 1082	G
18	Polygon	20	G7	20	55	29	510	108200	29510108200	1082	Census Tract 1082	G
19	Polygon	20	G7	20	56	29	510	108300	29510108300	1083	Census Tract 1083	G
20	Polygon	21	G8	21	56	29	510	108300	29510108300	1083	Census Tract 1083	G
21	Polygon	21	G8	21	69	29	510	127000	29510127000	1270	Census Tract 1270	G
22	Polygon	22	G9	22	69	29	510	127000	29510127000	1270	Census Tract 1270	G
23	Polygon	23	G10	23	69	29	510	127000	29510127000	1270	Census Tract 1270	G
24	Polygon	24	H6	24	46	29	510	107300	29510107300	1073	Census Tract 1073	G
25	Polygon	24	H6	24	55	29	510	108200	29510108200	1082	Census Tract 1082	G
26	Polygon	24	H6	24	103	29	510	108100	29510108100	1081	Census Tract 1081	G
27	Polygon	25	H7	25	46	29	510	107300	29510107300	1073	Census Tract 1073	G
28	Polygon	25	H7	25	55	29	510	108200	29510108200	1082	Census Tract 1082	G

1 (0 out of 634 Selected)

si01

$$A_w = \frac{A_i}{A_t}$$

A_w = Partial census tract area weight

A_i = Individual area of each census tract

A_t = Total area of the census tract parts

$$T_v = \sum_{i=1}^{n} C_t * A_w$$

T_v = Census tract population

C_t = Census tract total population

A_w = Partial census tract area weight

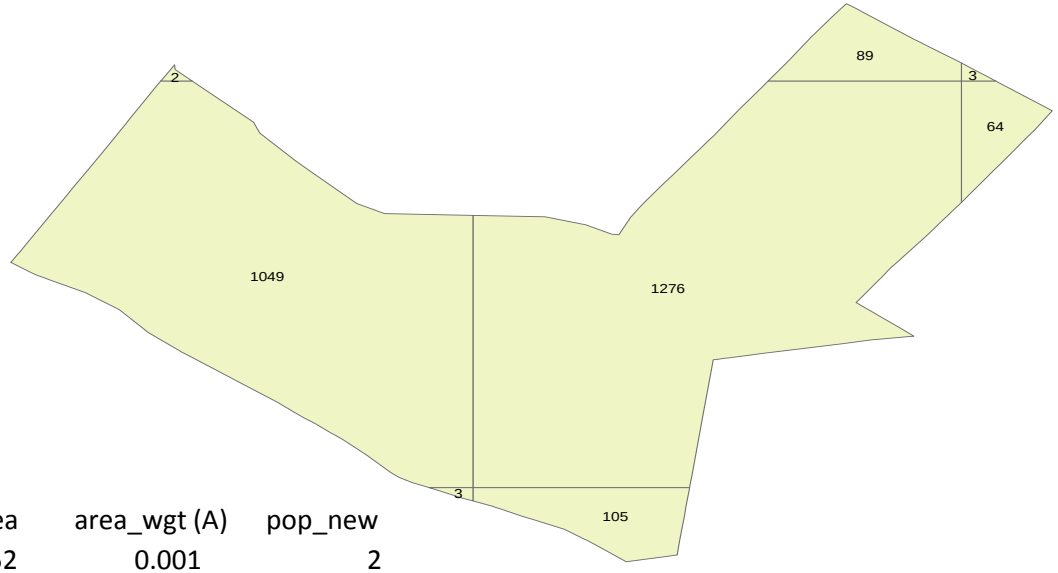
Table

si02

SE_T001_0012	area	OBJECTID_1	GEOID	Count_GEOID	Sum_area	Shape_Length	Shape_Area	area_wgt	pop_new
2591	0.001113	1	29510101100	8	1.251657	171.761061	1112.672362	0.000889	2.303294
2591	0.043112	1	29510101100	8	1.251657	967.052956	43111.50779	0.034444	89.243239
2591	0.00159	1	29510101100	8	1.251657	199.675425	1589.862493	0.00127	3.291105
2591	0.50658	1	29510101100	8	1.251657	3090.13167	506580.41584	0.404728	1048.649872
2591	0.61657	1	29510101100	8	1.251657	3804.781168	616569.78395	0.492603	1276.334033
2591	0.030697	1	29510101100	8	1.251657	798.027909	30696.927695	0.024525	63.544362
2591	0.001478	1	29510101100	8	1.251657	217.533358	1478.062647	0.001181	3.059673
2591	0.050518	1	29510101100	8	1.251657	1097.927209	50517.676122	0.040361	104.574423

1 (0 out of 8 Selected)

si02



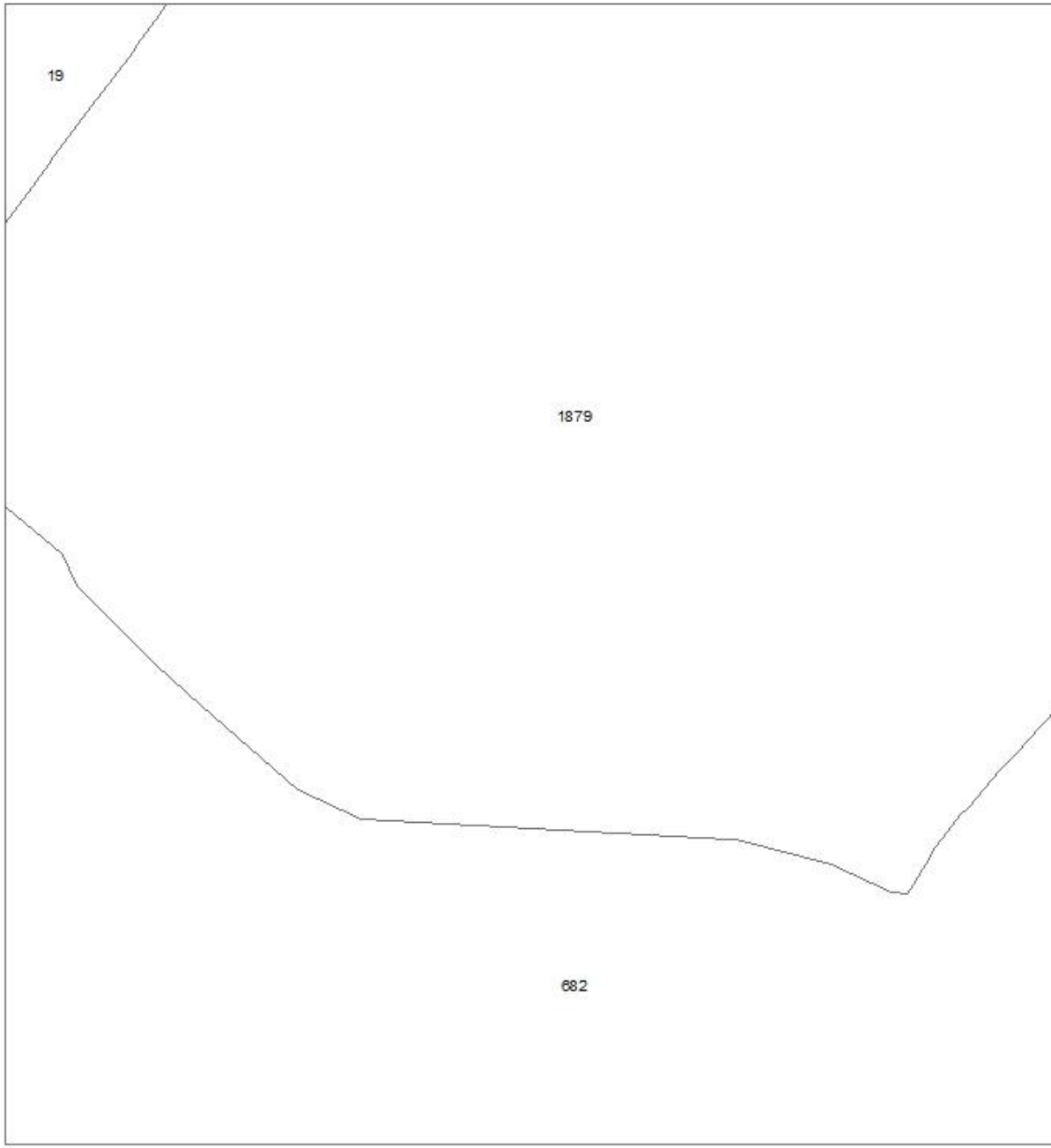
GEOID	Pop (C)	area	Sum_area	area_wgt (A)	pop_new
29510101100	2591	0.001	1.252	0.001	2
29510101100	2591	0.043	1.252	0.034	89
29510101100	2591	0.002	1.252	0.001	3
29510101100	2591	0.507	1.252	0.405	1049
29510101100	2591	0.617	1.252	0.493	1276
29510101100	2591	0.031	1.252	0.025	64
29510101100	2591	0.001	1.252	0.001	3
29510101100	2591	0.051	1.252	0.040	105
Total		1.252		1.000	2591

$$G_v = \sum_{i=1}^{n} C_t * A_w$$

G_v = Grid population

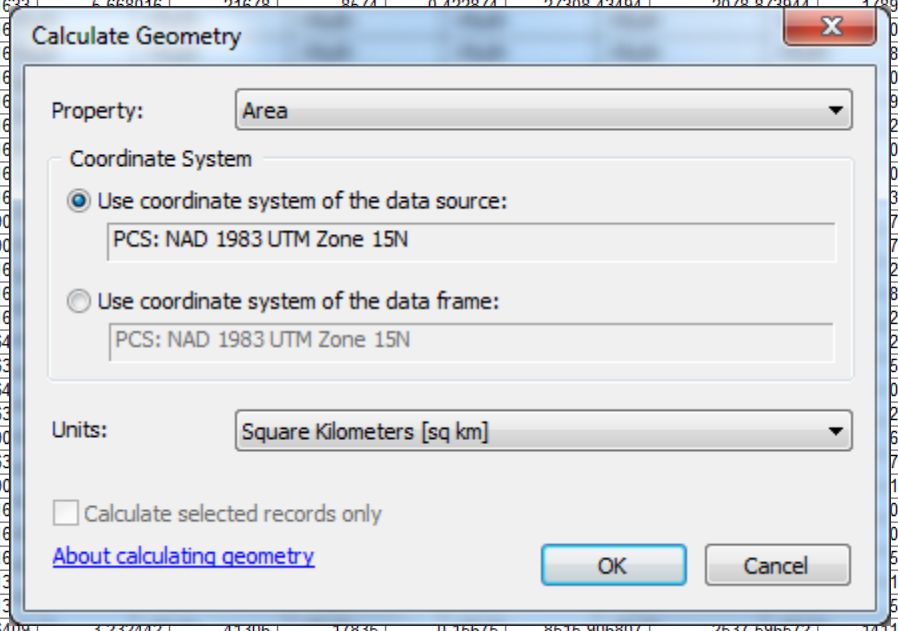
C_t = Census tract total population

A_w = Partial census tract area weight



Step 2 – Create a new variable called “area”. Note you can use the shape area as well. Divide by 1 million

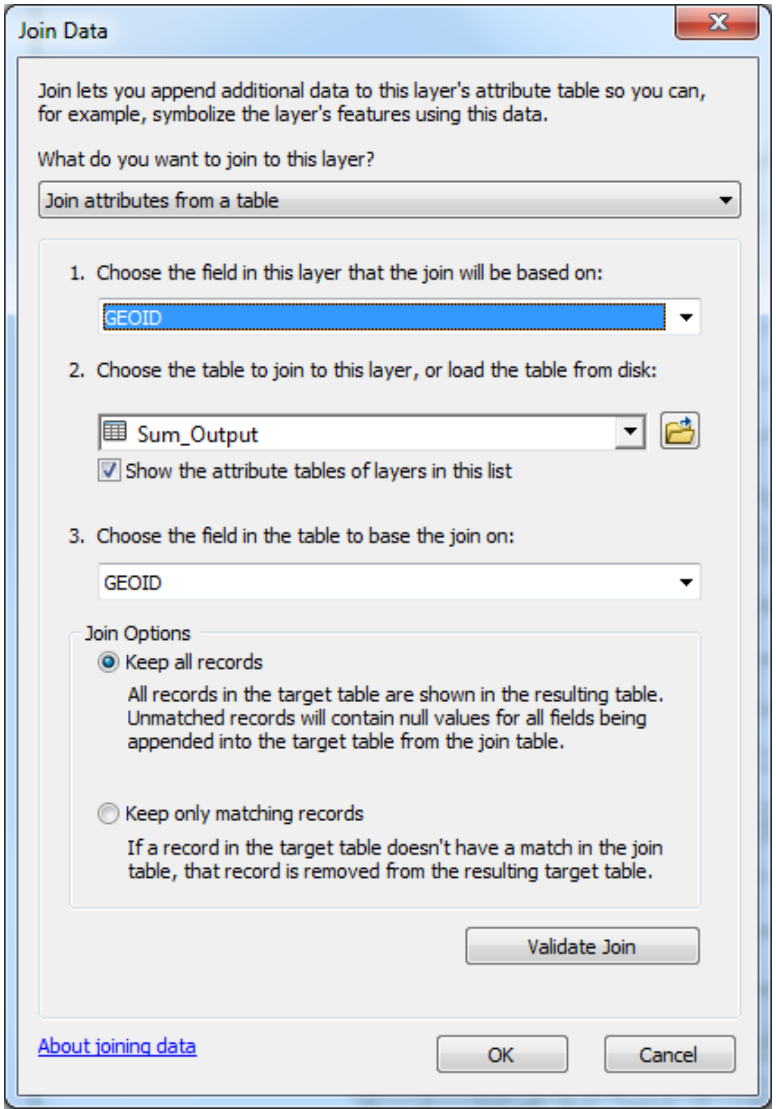
pct_wh	pct_blk	pct_oth	diversity	gini	crime_rat	med_hhi	per_cai	pov_rate	Shape_Leng	Shape_Length	Shape_Area	area
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	2115.325412	262464.667673	0.262465
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	2532.176998	325240.624478	0.325241
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	3418.036676	669538.002102	0.669538
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	3169.148839	525413.460137	0.525413
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	2078.873044	178946.810165	0.178947
0.140081	0.793522	0.066397	0.581579	0.416							099.043426	0.966099
0.140081	0.793522	0.066397	0.581579	0.416							801.342798	0.136801
0.140081	0.793522	0.066397	0.581579	0.416							0517.71588	0.060518
0.140081	0.793522	0.066397	0.581579	0.416							966.416708	0.826966
0.140081	0.793522	0.066397	0.581579	0.416							2198.39015	0.492198
0.140081	0.793522	0.066397	0.581579	0.416							0510.34918	0.06051
0.140081	0.793522	0.066397	0.581579	0.416							0514.13387	0.690514
0.140081	0.793522	0.066397	0.581579	0.416							316.241833	0.221316
0.053375	0.9206	0.026026	0.29813	0.490							790.794155	0.002791
0.053375	0.9206	0.026026	0.29813	0.490							798.675538	0.130799
0.140081	0.793522	0.066397	0.581579	0.416							251.727134	0.159252
0.140081	0.793522	0.066397	0.581579	0.416							863.322616	0.856863
0.140081	0.793522	0.066397	0.581579	0.416							2744.29175	0.402744
0.039671	0.937114	0.023215	0.251449	0.3664							254.102946	0.138254
0.035953	0.953669	0.010378	0.193162	0.5063							596.951945	0.089597
0.039671	0.937114	0.023215	0.251449	0.3664							037.062323	0.042037
0.035953	0.953669	0.010378	0.193162	0.5063							233.054623	0.534233
0.053375	0.9206	0.026026	0.29813	0.490							653.214456	0.223653
0.035953	0.953669	0.010378	0.193162	0.5063							775.381116	0.149775
0.053375	0.9206	0.026026	0.29813	0.490							144.584717	0.393145
0.140081	0.793522	0.066397	0.581579	0.416							080.034167	0.45708
0.140081	0.793522	0.066397	0.581579	0.416							000.000003	1
0.140081	0.793522	0.066397	0.581579	0.416							574.211006	0.068574
0.010034	0.971704	0.018262	0.133959	0.3813							1098.49885	0.061098
0.010034	0.971704	0.018262	0.133959	0.3813							537.641532	0.752538
0.039671	0.937114	0.023215	0.251449	0.366409	3.232442	41306	17835	0.15675	8515.906807	2537.596572	141109.569327	0.14111
0.010034	0.971704	0.018262	0.133959	0.381393	5.21774	30671	15364	0.273075	5608.172018	1008.573885	44568.165902	0.044568
0.039671	0.937114	0.023215	0.251449	0.366409	3.232442	41306	17835	0.15675	8515.906807	4175.787147	943240.11191	0.94324
0.035953	0.953669	0.010378	0.193162	0.506303	0.74129	34781	22485	0.205585	4775.556532	948.315047	12191.722186	0.012192
0.039671	0.937114	0.023215	0.251449	0.366409	3.232442	41306	17835	0.15675	8515.906807	2747.194528	396349.33057	0.396349
0.035953	0.953669	0.010378	0.193162	0.506303	0.74129	34781	22485	0.205585	4775.556532	1117.366009	41404.703205	0.041405
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	3088.458353	562245.966225	0.562246
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	4018.825954	980417.987881	0.980418
0.140081	0.793522	0.066397	0.581579	0.41633	5.668016	21678	8574	0.422874	27308.43494	1065.314151	26428.538916	0.026429
0.014903	0.967958	0.017139	0.149189	0.464102	8.196721	32708	17188	0.320776	3957.156052	794.272027	32102.773088	0.032103
0.010034	0.971704	0.018262	0.133959	0.381393	5.21774	30671	15364	0.273075	5608.172018	2667.655197	398605.198281	0.398605
0.008971	0.969896	0.021132	0.139668	0.466597	4.186603	28780	15007	0.294731	11983.090545	1676.790321	142643.777575	0.142644
0.014903	0.967958	0.017139	0.149189	0.464102	8.196721	32708	17188	0.320776	3957.156052	3082.37378	494916.724987	0.494917
0.010034	0.971704	0.018262	0.133959	0.381393	5.21774	30671	15364	0.273075	5608.172018	2619.797431	330694.482969	0.330694
0.011478	0.96198	0.026542	0.16829	0.464404	7.890961	27028	15117	0.373361	4397.149963	1777.618672	130270.741688	0.130271
0.039671	0.937114	0.023215	0.251449	0.366409	3.232442	41306	17835	0.15675	8515.906807	1021.634188	44118.050356	0.044118
0.014903	0.967958	0.017139	0.149189	0.464102	8.196721	32708	17188	0.320776	3957.156052	514.128113	10899.967624	0.0109

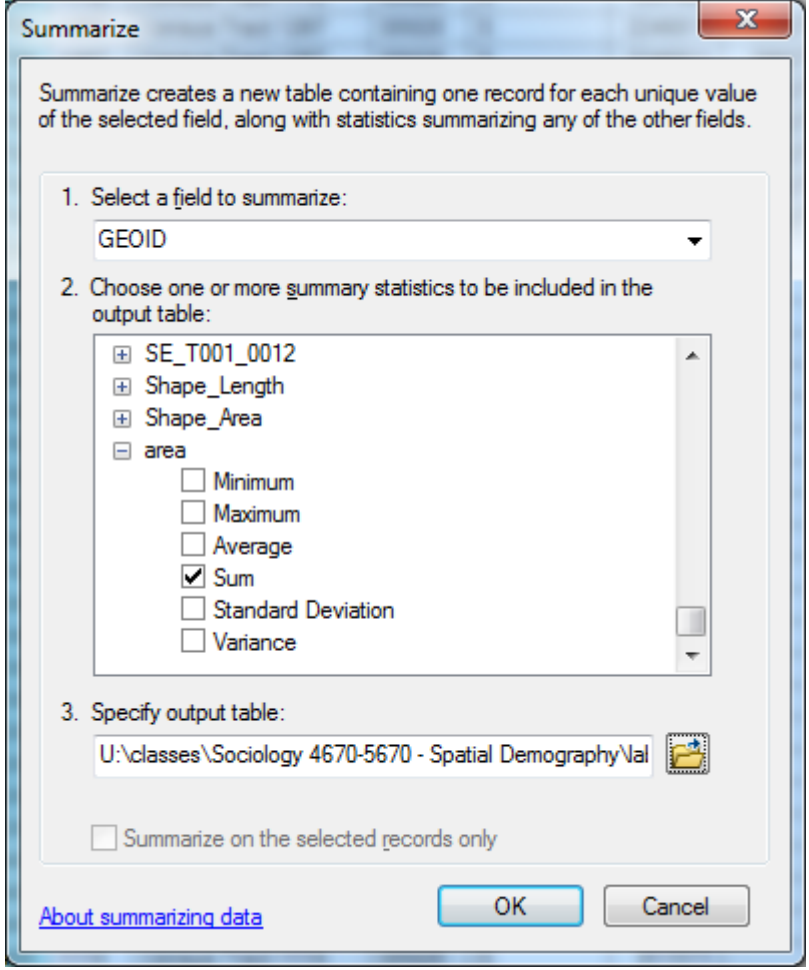


Step 3 – Summarize GEOID by area

Step 4 - Join Summarized Table to “si01”

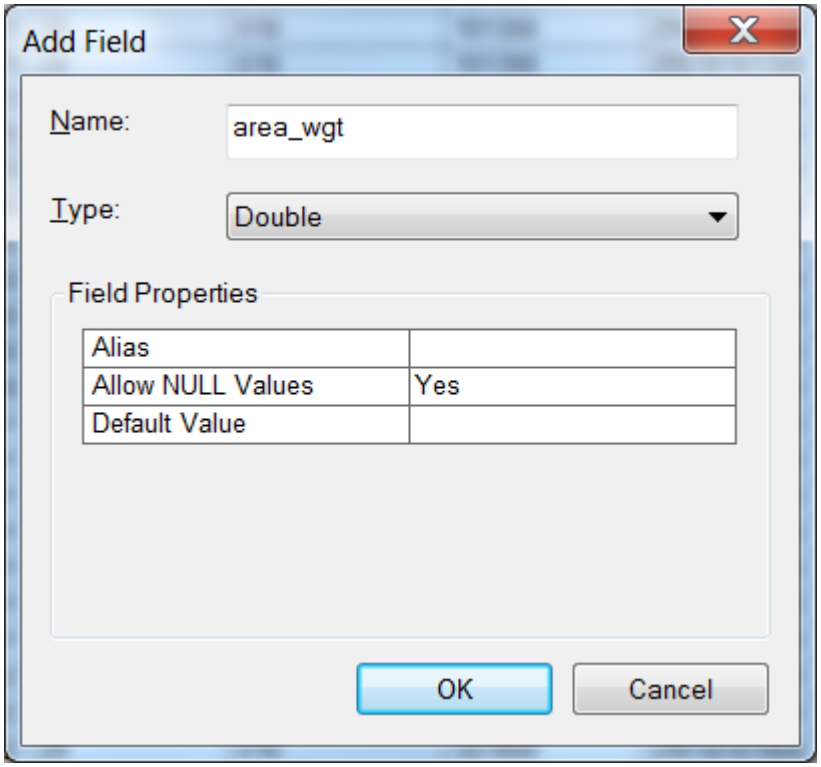
Step 5 – Make it a permanent join (e.g., new file)

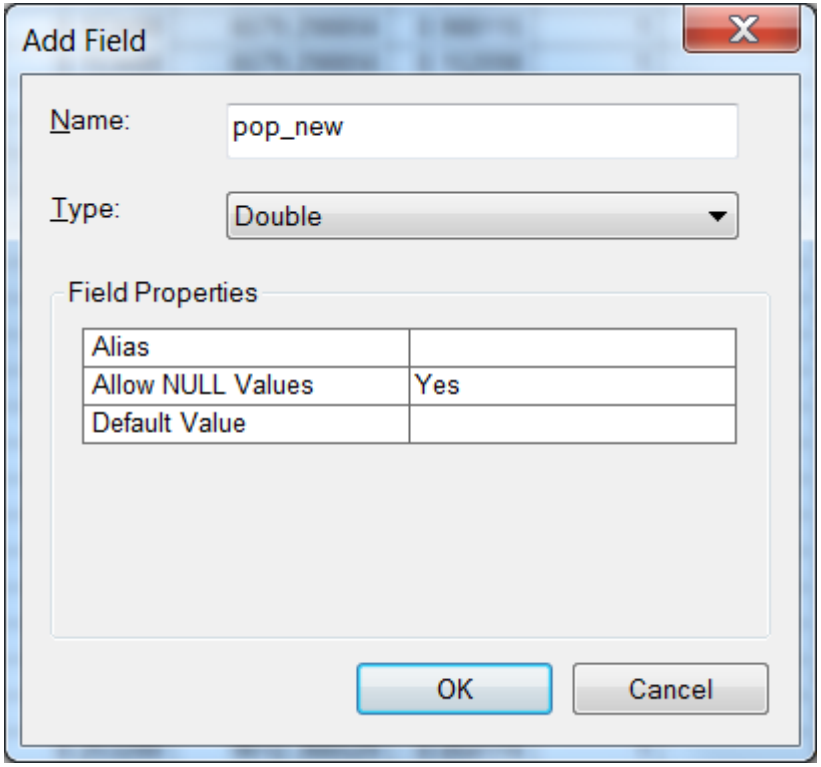




Step 6 – Add a field called “area_wgt”

Step 7 – Add a field called “pop_new”



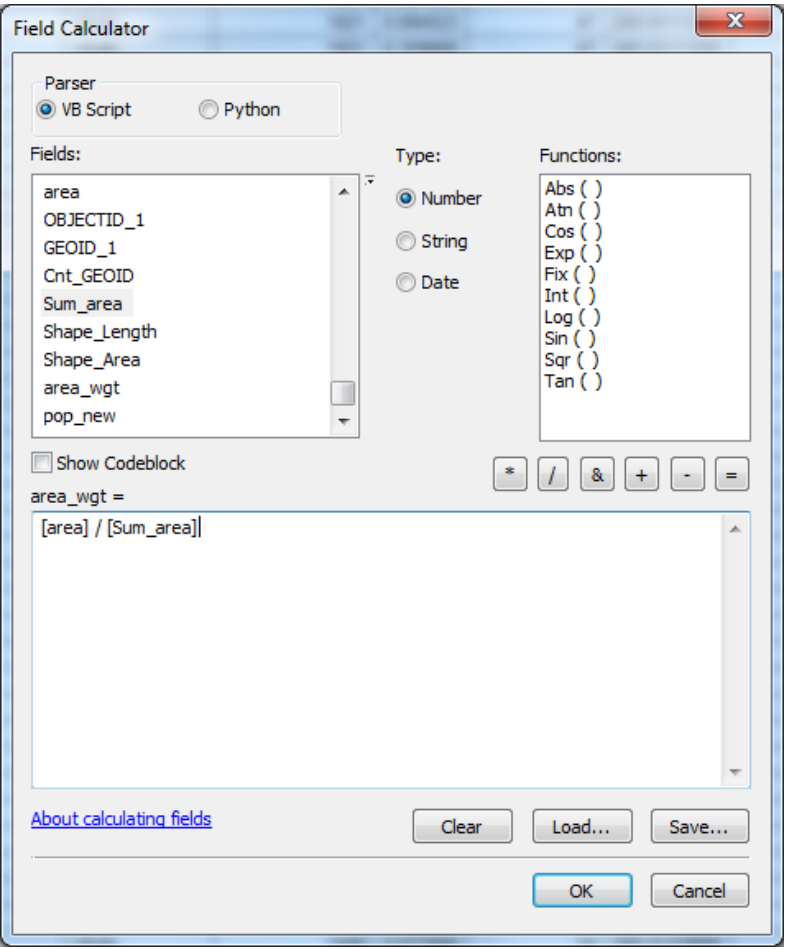


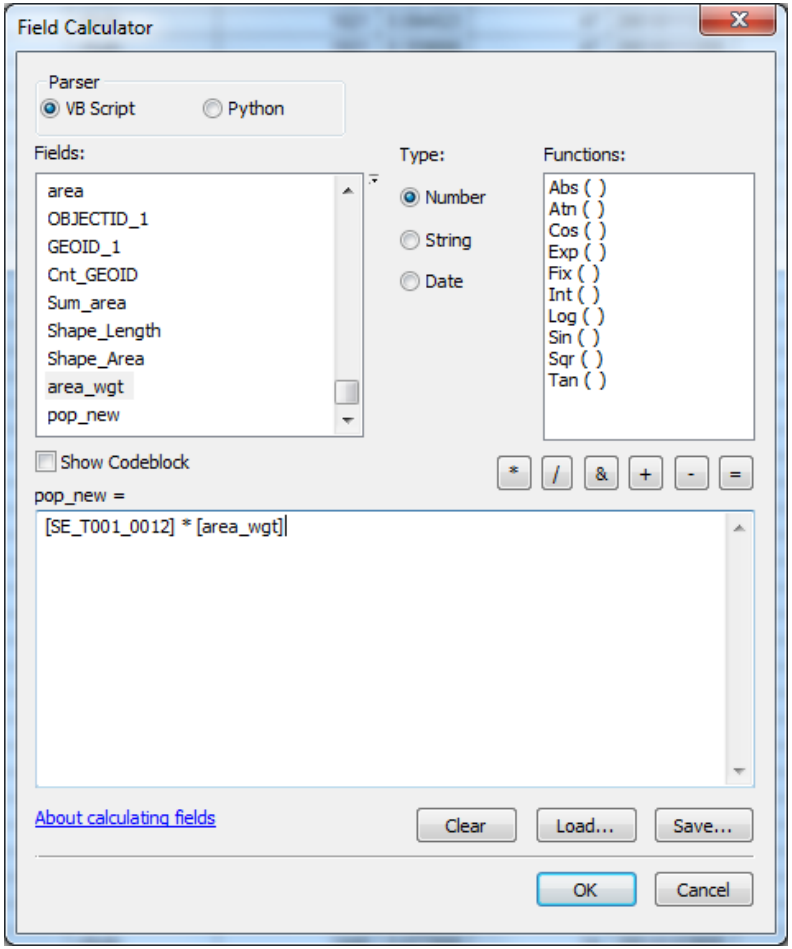
Step 8 – compute the formula for area_wgt

area_wgt=area/Sum_area

Step 9 – compute the formula for pop_new

pop_new=tot_pop*area_wgt (in our case tot_pop=SE_T001_0012)

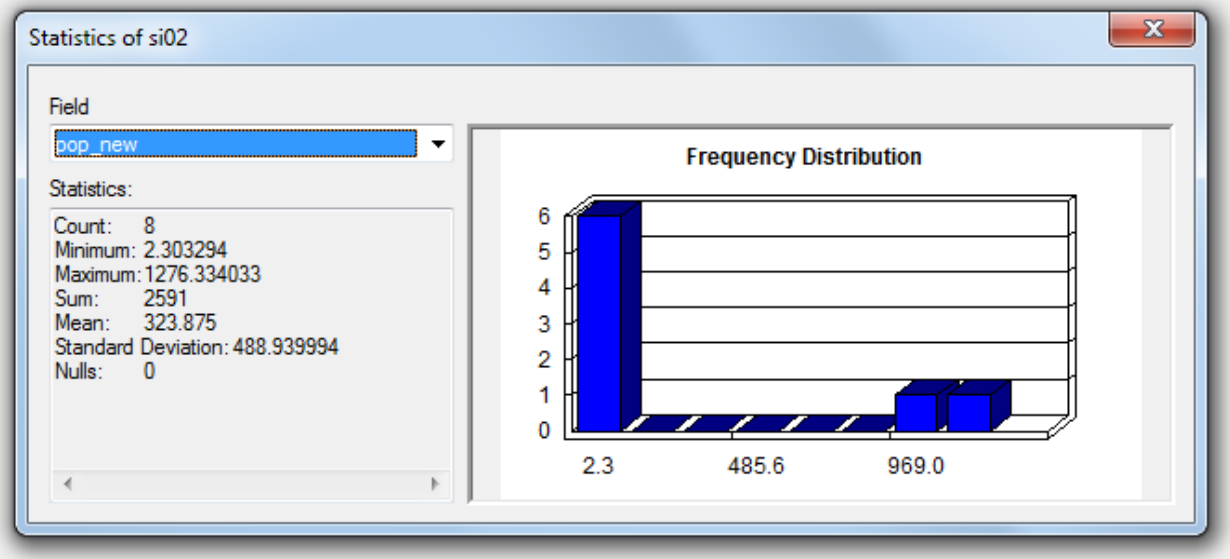




Step 10 – check your work - the new population must equal the tract population (n=2591)

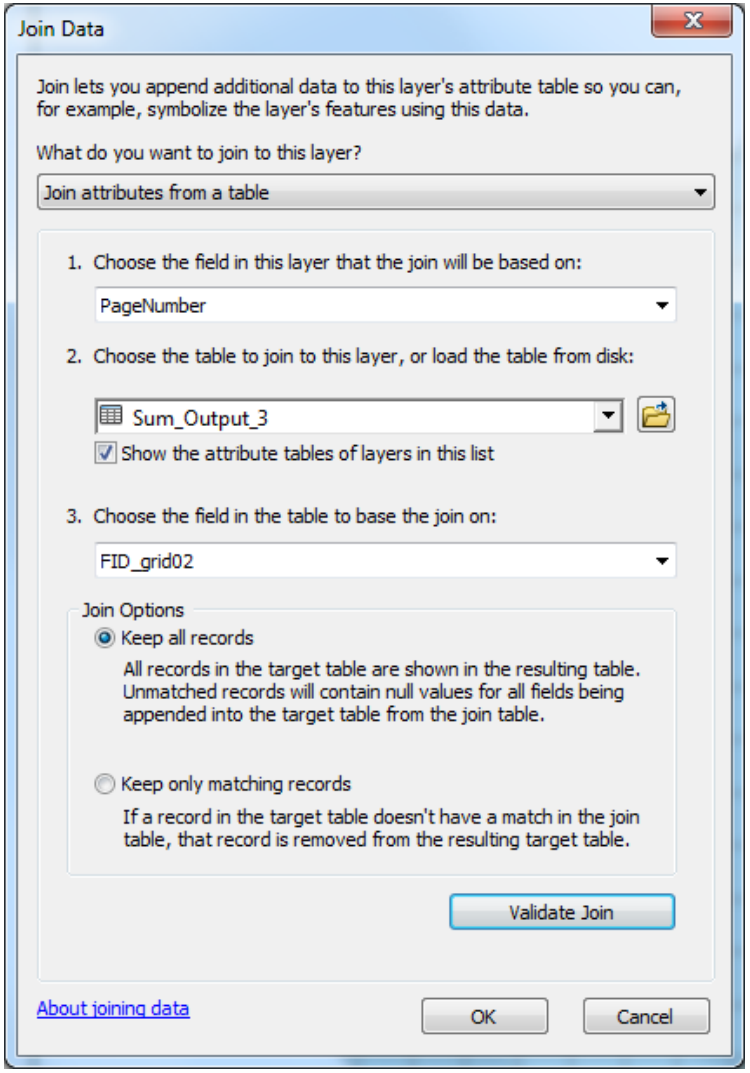
Everything looks right. The 2591 residents have been distributed over the 8 grids

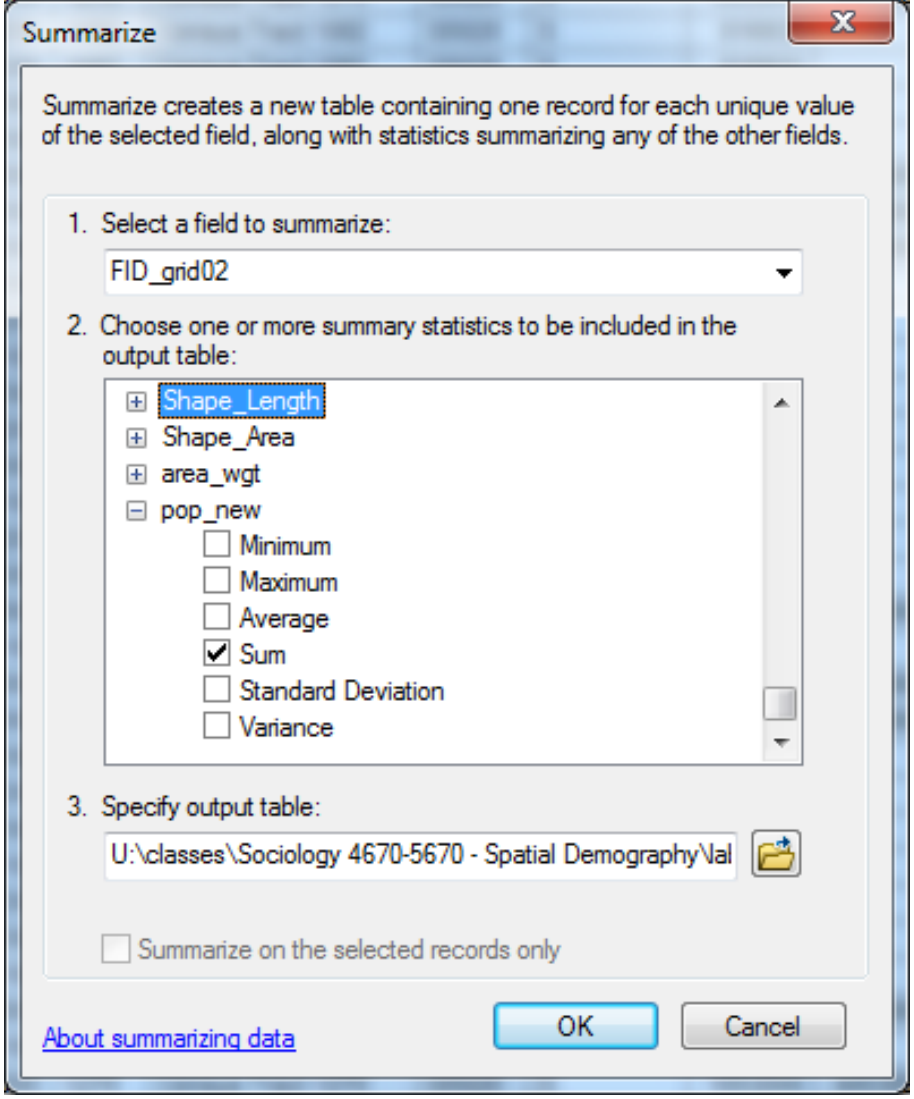
SE_T001_0012	area	OBJECTID_1	GEOID	Count_GEOID	Sum_area	Shape_Length	Shape_Area	area_wgt	pop_new
2591	0.001113	1	29510101100	8	1.251657	171.761061	1112.672362	0.000889	2.303294
2591	0.043112	1	29510101100	8	1.251657	967.052956	43111.50779	0.034444	89.243239
2591	0.00159	1	29510101100	8	1.251657	199.675425	1589.862493	0.00127	3.291105
2591	0.50658	1	29510101100	8	1.251657	3090.13167	506580.41584	0.404728	1048.649872
2591	0.61657	1	29510101100	8	1.251657	3804.781168	616569.78395	0.492603	1276.334033
2591	0.030697	1	29510101100	8	1.251657	798.027909	30696.927695	0.024525	63.544362
2591	0.001478	1	29510101100	8	1.251657	217.533358	1478.062647	0.001181	3.059673
2591	0.050518	1	29510101100	8	1.251657	1097.927209	50517.676122	0.040361	104.574423

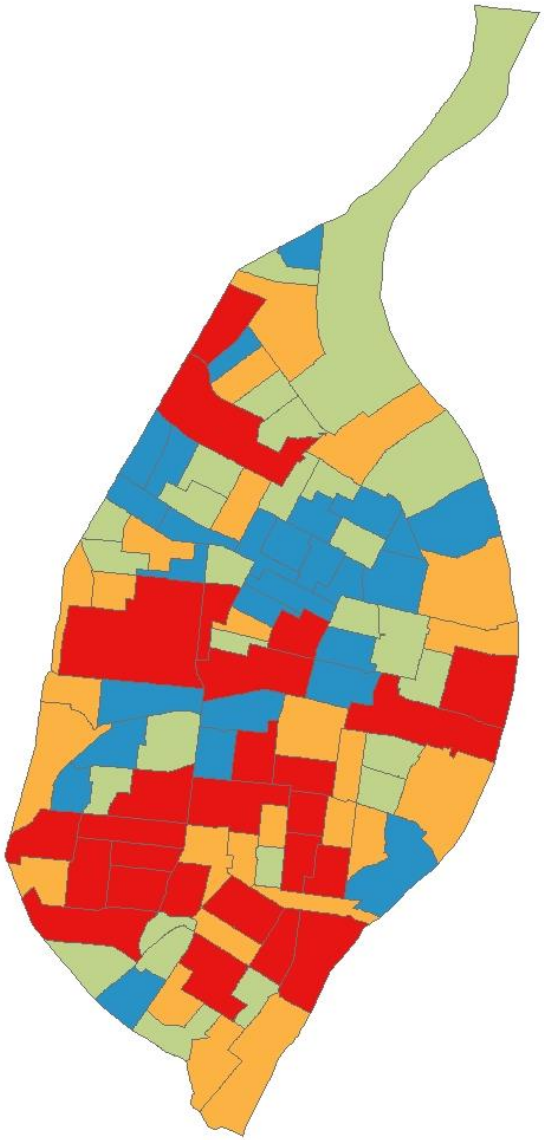


Step 11 – Summarize the GRIDCODE by the "pop_new" variable

Step 12 – Join the summarized table to the unpopulated grid shapefile.



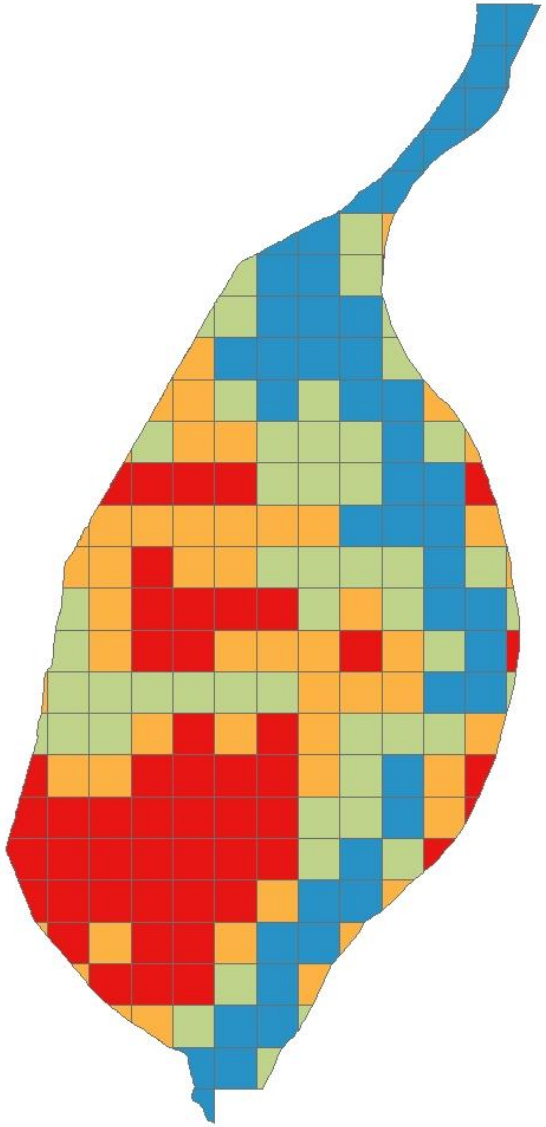




Legend

SE_T001_0012

- 620 - 2024
- 2025 - 2920
- 2921 - 3789
- 3790 - 7069



Legend

grid02

Sum_pop_new

- 0 - 468
- 469 - 1359
- 1360 - 2280
- 2281 - 5207