Help

Explore frequently asked questions and learn about the tools.

What is an SSA?

SSA stands for Statistical Small Area. An SSA is a group of census tracts clustered based on key indicators like poverty, housing, and transportation. Much more information about how the SSA boundaries were developed appears later in this FAQ.

Why did you develop SSAs?

The San Antonio metro area is consistently identified as one of the most highly economically-segregated in the country, but racial/ethnic segregation is much less pronounced. Many disparities and inequities remain hidden unless data can be sliced by both race/ethnicity and place. Unfortunately, data thus disaggregated at the census tract level or lower falls apart due to wide margins of error in survey estimates, data suppression to protect privacy, and volatile rates. ZIP codes, which are quite a bit larger than census tracts, are commonly used but are meaningless in terms of anything except efficient mail delivery. To help solve that problem, we created SSAs: clusters of census tracts that better reflect real neighborhoods than do ZIP codes.

How did you develop the SSA boundaries?

CINow staff researched and tested several statistical methods of census tract aggregation, including meeting several times with multiple other investigators who have done similar work in the U.S. A recurring method in the literature on spatially aggregating geographies is using Principal Component Analysis (PCA) as a way to reduce dimensions (number of variables used). Staff ran a PCA with demographic variables such as race, income, and education along with a spatial K Means analysis to group like census tracts together based on the number of principal components that explain at least 90% of the original variance in the predictors. The Spatial 'K'luster Analysis by Tree Edge Removal (SKATER) algorithm is used to keep the clusters spatially contiguous. CINow staff created a function to automatically run the PCA, K-means, and SKATER algorithms for different combinations of minimum census tract groupings and number of regions (clusters) and calculate the coefficient of variation (CV) for the African American population for the newly created regions. The goal is to create new regions where the CV for the African American population is 0.5 or less, because although it is the third-largest race/ethnicity group in the county, at just over 7%of total population the African American population is the one for which super-neighborhoods should make the biggest improvement in margin of error and resulting usefulness of small-area estimates.

CINow developed four separate models where census tracts are grouped into 75 super-neighborhoods statistically clustered on different key criteria: poverty, housing, transportation, and general, which incorporates all three of the other key criteria. These four models were shared with core project partners to gather their input on which of the four should be moved forward for public input on the boundaries. The "General" model was put forward for public input, with feedback solicited through CINow's website, newsletter, and via Facebook. Public input led to two fundamental changes in the boundaries:

- The original term for the census tract clusters was "super-neighborhood". That term was abandoned in favor of "statistical small area" or SSA. The term "super-neighborhood" led to confusion that these boundaries would somehow compete with or replace the City of San Antonio's formal neighborhood association, historic district, and conservation district boundaries – all of which are quite politicized. "Statistical small area" was chosen to communicate that these are geographies that are smaller than county level and were created using statistical methods. If SSA does not perform well when the data platform undergoes beta testing with community users, "census tract cluster" or some other option may be chosen.

- To the extent possible, each small municipality in Bexar County (e.g., Leon Valley, Kirby) was kept intact in a single SSA, rather than its boundaries straddling two or more SSAs. The final boundaries divide the county into 92 SSAs. (In comparison, the county has 76 ZIP codes.)

Each Decennial Census changes how data are collected, measured, or defined. One of those changes that creates a challenge every ten years is the adaptation of geographic boundaries (Census tracts) to new population estimates. Census tracts must have between 2,500 and 8,000 residents, so every 10 years, tracts with too few or too many residents are combined or split to result in a resident count within the required range. Those changes in tract boundaries complicate analysis over time. The SSA boundaries were created to fit both 2010 and 2020 Census tract geographies. This allows us to accurately trend between 2010-2020 decennial Census and 2010-2029 American Community Survey datasets.