



“Introducing Our Nation’s Rural-Serving Postsecondary Institutions”

Rural-Serving Institutions Technical Documentation

Prepared By:

Andrew Koricich
Alisa Hicklin Fryar
Vanessa A. Sansone
Corey S. Sparks
Cecilia M. Orphan
Kevin R. McClure

Alliance for Research on Regional Colleges

January 2022



OUR PATH TO AN RSI METRIC

Rural communities are often overlooked, but they have been receiving more attention recently, leading to more interest in investing in postsecondary institutions that serve rural areas. Many policymakers, foundations, nonprofits, and national organizations are wanting to dedicate more time and attention to higher education institutions in rural communities, but when it's time to do the work, allocate the resources, or bring people to the table, decision-makers often get bogged down with the first and most fundamental question: how do we know what "counts" as rural?

This question is neither new nor unique to higher education. It is a long-standing debate within multiple literatures and policy domains. There are many reasons why defining "rural" is so complicated, but higher education is somewhat unique in the relationships between students, institutions, and communities. Some people are interested only in students who grew up in rural communities. Some are focused on colleges that are located in very small communities that are far from metropolitan areas. Others may be interested in universities whose service area includes rural communities.

In addition to differences in focus, we also know that perceptions of what "counts" as rural can vary widely. Most people would agree that a place like Baruch College in New York City would not "count" as rural-serving, while a place like The University of West Alabama in Livingston would be considered rural-serving by most people. The outliers are fairly easy to identify. But the middle space is far more complicated, often because many Americans think of rural only as "less urban," and even their understanding of urban is often deeply connected to their own life experiences. Most Americans, regardless of upbringing, know something about the largest cities in the country which offers a baseline level of knowledge about urban communities (even if ill-informed) that does not exist for rural communities.

Beyond the question of what "counts" as rural, we had to determine how to operationalize the *servicing* aspect of "rural-serving institution" (RSI). For Hispanic-Serving Institutions, a key criterion is an undergraduate enrollment of 25% or higher who identify as Hispanic, and for Native American-Serving Nontribal Institutions, the enrollment criterion is 10% or more of undergraduates who identify as Native American. These types of criteria are possible because institutions report the racial composition of their student bodies as part of an annual mandate to submit data to the Integrated Postsecondary Education Data System (IPEDS) hosted by the National Center for Education Statistics (NCES). The same is not true for rural students. Some institutions may specifically identify which students are from rural areas, but not all do. Furthermore, in order to have similar reporting in IPEDS, there would have to be one, universal definition of "rural" that all institutions used. So, in the absence of rural student enrollment data in IPEDS, we had to take a different approach, as well as consider other criteria that go beyond simply enrolling rural students.

Most of all, our research and exploration led us to one clear conclusion: definitions of rural and rural-serving are neither static nor universal. Any effort to understand and appreciate rural communities, rural populations, or rural institutions must be supported by a metric that is tailored to the focus and goal of that research. For this project, one way we chose to mitigate this complexity was to employ multiple

measures of place. The value in this approach is that it allows for a more-nuanced consideration of rurality—even with national-level, quantitative data—and this prevents institutions from being seen as rural-serving (or not) based solely on one measure of where they are located.

Our goal has been to create a measure that best supports studying, supporting, and discussing rural-serving postsecondary institutions. Our metric is inextricably tied to that purpose. The purpose of this document is to be fully transparent about our process and our metric, so that others are clear about the advantages and limitations of using the RSI metric for their work.

EXISTING METRICS FOR IDENTIFYING RURAL INSTITUTIONS

There are many ways that rurality has been defined in policy documents and academic literature (not including the work that discusses rural without defining or measuring it at all). When rural is defined/measured, there are several metrics that are most commonly used. The most common definitions rely heavily on population size and proximity to urban communities, and there are meaningful differences across metrics in the geographic/organizational unit of analysis (county, city, urban area, urban cluster, etc).

NCES Locale. In higher education policy, efforts to use quantitative data or objective classifications to discuss rural institutions often rely heavily on the “locale” variable in the NCES / IPEDS federal dataset (NCES Locale, 2016). The locale variable is based on Census data and includes four broad categories (city, suburb, town, and rural), and each of the four categories has three subcategories. For city and suburb, the subcategories include large, medium, and small, distinctions that are connected to population size (over 250K, 100K-250K, less than 100K). In these data, a town is defined as a place that is outside of an urbanized area but inside an urban cluster. Rural is outside of an urbanized area and an urban cluster. The subcategories for town and rural are fringe, distant, and remote, distinctions based on distance from an urbanized area (for town) and distance from an urban cluster (for remote).

This categorization is quite useful when studying urban areas, because it is a very urban-centric definition. In our early examination, we found great value in using the locale variable when trying to differentiate among the most-urban areas, but we found less utility in using locale to understand rural communities and the institutions that serve them. Neither the town/rural distinction, nor the fringe/distant/remote distinctions added much clarity to our understanding of rural serving institutions, nor did it aid us in better understanding the diversity among RSIs. In short, NCES locale helps us understand “not urban” far better than it helps us understand rural.

USDA-ERS County-Level Metro/Non-Metro. Much of the work on rural communities relies heavily on county-level data, and one of the most common data sources is the County Typology Codes dataset by the U.S. Department of Agriculture’s Economic Research Service (USDA-ERS). These data include a dichotomous, county-level “metropolitan/non-metropolitan” variable. For this variable, a county is considered a metro county if it includes one or more urbanized areas and has a population of at least

50,000 or is an adjacent county with a high level of integration with the urbanized area, based on commuting (USDA-ERS, 2015; Census 2010).

This measure is a straightforward categorization and a commonly used one, but for the purpose of our research, there were two key limitations. First, as a dichotomous variable, it limits our ability to understand the diversity among non-metro counties. Second, it's a single-item, county-level measure, which can be quite complicated when trying to examine institutions across the country, because county size and county dynamics are very different across the U.S. (See map at: https://www2.census.gov/geo/maps/general_ref/us_base/stco2003/stco2003.pdf). The variation in county size is problematic, because when we use county-level data to understand institutional characteristics, especially when it's the only measure used, we have to assume that the relationships between the institutional characteristics and the county characteristics are fairly consistent. And we know from empirical work, including our own research and analysis, that the characteristics between institutions and counties can be inconsistent.

For example, Oregon Institute of Technology and Lander University in South Carolina are both four-year institutions with comparable institutional missions. However, Lander University is located in Greenwood County, a county that covers 463 square miles, while Oregon Institute of Technology is located in Klamath County, which covers 6,136 square miles. The population size is comparable, but Klamath County is roughly 13 times the area of Greenwood County. The limitations that come from these differences in county size do not invalidate the use of county-level data, as other options also have their own limitations, but it must be considered.

RUCC Codes. The Rural-Urban Continuum Code (RUCC) data have been used in a wide range of work across multiple policy areas. RUCC is an ordinal, nine-point, county-level measure, developed by the USDA-ERS. The categories are based on population size and proximity to a metropolitan area and are updated every 10 years.

- 1 - Counties in metro areas of 1 million population or more
- 2 - Counties in metro areas of 250,000 to 1 million population
- 3 - Counties in metro areas of fewer than 250,000
- 4 - Urban population of 20,000 or more, adjacent to a metro area
- 5 - Urban population of 20,000 or more, not adjacent to a metro area
- 6 - Urban population of 2,500 to 19,999, adjacent to a metro area
- 7 - Urban population of 2,500 to 19,999, not adjacent to a metro area
- 8 - Completely rural or less than 2,500 urban population, adjacent to a metro area
- 9 - Completely rural or less than 2,500 urban population, not adjacent to a metro area

RUCC offers many benefits to rural scholars, especially in its ability to differentiate among various types of rural communities, but it has its limitations as well. We soon realized that if we were to use this metric as our only indicator, we would have similar problems that we face with other county-level data, but we also discovered a different limitation related to the ordinal nature of the scale. In our own analysis, we found that other characteristics most often connected to rural communities did not trend with the RUCC scale as closely as we had expected. Upon further review, we learned that “adjacent to a metro area” is a

separate phenomenon from population size, meaning that the evidence suggested that a “6” on the scale (urban population of 2,500 to 19,999, adjacent to a metro area) was not necessarily more rural, as understood by other measures, as a “5” (urban population of 20,000 or more, not adjacent to a metro area). Because RUCC is so common for rural scholars, we were reluctant to exclude it completely, but we realized that we need to look at size and metro adjacency separately.

OUR GOALS

Early in the process, we wanted to be clear that our efforts to identify RSIs were not just about identifying rural-located institutions. We wanted to find a better way to identify which institutions are well-situated to serve rural communities, and we wanted to do that in a way that was inclusive of the diversity of RSIs, but we didn’t want rural-serving to be so inclusive that it fails to offer new insight. That balance is quite difficult to achieve. Our review of existing metrics for rural helped us appreciate the limitations of a single metric (like using RUCC or Metro/Non-Metro) and the limitations of calling certain places “rural” when places that were labeled otherwise (like “town” in the NCES locale data) often had elements of rurality as well. In the end, we settled on a goal of creating a metric (1) based on multiple measures and (2) structured to be relative in nature, so that others could decide how strict they wanted to be in their view of what “counts” as rural-serving.

We also realized that we had a number of constraints to consider. Some of these constraints were obstacles that were unavoidable, but others offered opportunities for us to further refine our goals and our contributions through this work.

CONSTRAINTS AND OPPORTUNITIES

Rural as an Identity. Ask someone from a major city to name a rural town, and it’s very likely that they will name a town that is actually not that rural. And for many people in the rural communities, any town with a Walmart would be considered a decent-sized city. At the same time, there are people who identify as rural in ways that are not connected to where they reside. Someone who was raised in a rural community who continues to have some of the social markers that they connect to their hometown may think of themselves as rural, even if they’ve lived in Washington, D.C., for the past ten years. At the institutional level, the disconnect between rural-serving identity and measurable factors can be even more pronounced. If you have a research institution that’s located in a city of over 250,000 people, with a student population that is mostly coming from urban areas, many people would consider that an urban university. But a place like Texas Tech University in Lubbock, which is a city considered to be a rural place that also has a rural identity, is not usually what people think of when they think of an urban university.

As researchers seeking to use quantitative data to add clarity, this disconnect is complicated. For many, being connected to rural communities, either personally or institutionally, is very important to their identity, and, at the same time, people disagree strongly over what it means for a place to be rural or an institution to be rural-serving. We quickly realized that, once again, our goal is not actually to decide

which institutions are rural. Our goal has always been to create a national, inclusive metric, using quantitative data, to identify the institutions that are best-positioned to be rural-serving. Our metric does not invalidate the work of institutions that are not as closely connected to rural communities, nor should it give credit to institutions who are in rural areas but not doing the work. Our metric helps us have a better conversation about institutions that serve rural communities.

Variation in County Size. As mentioned in the discussion of existing, commonly-used metrics, using a single measure of county characteristics is complicated due to the variations in county size across the country. As we continued our work, we realized that variations in county size created two challenges: for geographically larger counties, we obscure diversity within the county, and for geographically smaller counties, we lose the ability to pick up on characteristics of surrounding communities that are meaningfully connected to the institution.

After much consideration, we realized that we could utilize different county-level metrics in ways that could help us address these shortcomings. We incorporate the Census county-level data on rural populations, combined with the county adjacency data, to try to offer a more nuanced view of county and regional rurality. As discussed in greater detail below, for each university, we incorporate a measure of the percentage of the county's population that is rural, and the average of the percent rural in adjacent counties. Utilizing these metrics is not a "fix" for the uneven county size, but it allows us to incorporate different ways of measuring rurality while still working to make sure our metric and the associated data are transparent and easily accessible.

Lack of Student Level Data. The most common question we are asked is whether we include the number or percent of rural students in the metric, similar to how the Hispanic-Serving Institution federal designation is structured, in part, on the percentage of students at an institution who identify as Hispanic. We do not include student data in the metric, because these data are not available at the national level. However, we are confident that, surprising to some, the lack of rural student numbers is not a meaningful limitation for our purposes.

First, as mentioned earlier, our goal is to identify institutions that are well-positioned to serve rural communities, which institutions do through student enrollment and other functions like community engagement, concurrent enrollment with high schools, business incubation, fine arts programming, and many, many other functions. Student counts would focus only on one aspect of that work.

Second, our metric is more directly tied to the rurality of the institution's location, but from our analysis there is strong evidence that the rurality of the location of the institution and the enrollment of rural students is meaningfully connected. The 2016 National Postsecondary Student Aid Study data allows us to examine these trends by comparing the locale of the institution to the locale of the student's home address. These data show that the majority of students at rural institutions come from rural communities, while less than 10% of students from urban communities attend rural institutions. So even though we cannot connect all rural students to individual campuses at the national level, we can safely say that there is strong evidence that students from rural communities are far more likely to attend rural institutions than students from urban areas and that institutions located in rural areas are largely enrolling rural students.

Multi-Campus Institutions. A number of institutions have more than one physical location, and, in some cases, those physical locations can be very different. Due to data limitations, we rely exclusively on the county of residence attached to an individual institution, as listed in IPEDS. However, our measure of rural-serving includes data on more than just the single county in which the institution is listed. We also include data on contiguous counties, as well as data on degrees awarded, which encompasses all of the units within that institutional record.

Online Education. As the percentage of students enrolled in exclusively online education increases, some may argue that the physical location of the institution is less connected to the work they do and the students they educate. At the extremes, with all students exclusively online, it is reasonable to suspect that the institution is functioning differently than we usually observe, and as such, we chose to exclude those institutions from our analysis. However, we chose to keep the other institutions in the dataset, as evidence suggests that many RSIs are supporting online education as a means for reaching a greater number of students located in regions with low population density and longer commute times to a physical campus.

Rural Mission. In the early stages of the project, we were asked if we would be capturing evidence of whether a commitment to serving rural communities was explicitly included in an institution's mission statement or marketing materials. It is clear that some institutions are cognizant of, and invested in, their rural-serving work, and that commitment is important. However, that focus is most expressed with respect to the region they serve, not an all-encompassing commitment to rural communities. This is important to mention because we believe not having an explicit mission to serve rural communities does not mean that an institution is *not* rural-serving. Only one institution within the most recent data year of IPEDS data on mission statements explicitly mentions rural at all. There was clearly no feasible way to collect national data on the role of rurality in institutional mission statements and/or branding.

CREATING THE METRIC

As we started our work to create the metric, we surveyed our options and landed on three main data sources that we found most useful: USDA-ERS RUCC Scores, Census population data, and degrees awarded in rural-relevant fields. Upon further examination, we found that these three sources offered a number of options for how to incorporate them into the analysis. For our final metric, we include two data elements from USDA-ERS RUCC scores, two elements from the Census data, and one variable from IPEDS on degrees awarded.

RUCC in Two Parts. As discussed in the earlier section, RUCC scores are widely used in the scholarship on rurality and offer tremendous benefit for speaking across literatures and substantive areas. However, RUCC is a nine-point ordinal scale constructed by incorporating two pieces of information: population size and adjacency to a metro area. In the original RUCC scores, metro adjacency is functionally used in a way that moderates the relationship between population size and rurality. Our analysis suggests that, while connected, county population size and metro adjacency function differently. For our measure, we break up RUCC scores into two parts: a six-point ordinal variable for population size and a dummy variable for metro adjacency.

Rural Population in the County and Region. Our measure also includes two variables that are drawn from Census data. The first is the percent of the institution's home county population that is coded as rural. Unlike the county-based rurality codes, this measure captures some of the variation within counties. We also use these data to include a measure of regional population rurality, which we measure as the average percent rural population for all adjacent counties. Including these two variables in our RSI scores allows us to take a broader, more problem-based approach to measuring rurality using geographic data.

Degrees Awarded. Historically, many institutions have had strong reputations for their work in rural communities tied to their support and excellence in agricultural programs, co-ops, and extension offices. We wanted to be thoughtful about those ties, while also recognizing that there are other areas of emphasis, outside of agriculture, that are disproportionately applicable to rural communities. In order to identify these other fields, we relied on several sources of information: our expertise in rural issues, extant literature on the rural workforce, and ad-hoc analyses of our location metrics and degree fields to understand how fields of study are spread across institutional locales. Using the Classification of Institutional Programs (CIP) codes contained in IPEDS, we were able to test our assumptions and determine that, in addition to Agriculture, the two fields of particular rural relevance were Natural Resources and Parks and Recreation. For the purposes of this project, we totaled the degrees and certificates awarded in those three fields and divided that by the total number of degrees and certificates awarded by the institution that year to calculate the percentage of institutional awards conferred in Agriculture, Natural Resources, and Parks and Recreation.

Index Construction Methodology. Factor analysis is used to construct the RSI metric using the principal factor method. Factor analysis is commonly used in statistical analysis when researchers wish to examine patterns of correlation among several variables. The Principal Factor Method seeks to find the so-called principal components of a set of variables that identify separate, uncorrelated patterns among the variables. In this sense, it is often used as a variable reduction strategy to form indices that represent information measured on several variables at the same time. It does so by finding linear combinations of the observed variables that correspond to separate components of the original data. These components of the overall variation are often identified as latent, or unobserved, factors that represent a combination of observed characteristics. Factor analysis derives loadings of the original variables that describe how the original variables relate to the new latent variable. These loadings act as weights for the original variables, much in the way parameters in a regression model work to predict the value of an outcome variable, so that the latent variable values are a weighted additive sum of the loading coefficients and the original data. The RSI index score is calculated using the loadings from the factor analysis applied to the original data. These values are naturally scaled to have a mean of 0 and variance of 1. To aid in interpretation of the RSI index, the index is re-scaled to have a minimum value of 0.

WHAT'S NOT THERE/OUT OF SCOPE

Although we have made clear what *is* included in our creation of the RSI metric, we also want to add clarity to what *is not* included in the metric and out of its scope. First, the metric does not take into account the economic development of the rural community or the economic contributions of institutions to their communities. A primary reason economic development is not included in the RSI metric is

because there are data limitations and attribution issues with measuring the concept of economic development. We also considered that economic development is associated with the size and demographics of a geographic area, which does not accurately reflect the nuances of rural economies. Additionally, our RSI scores are calculated for all institutions, including urban institutions, creating challenges for specifying attribution. To best support our goal of creating an RSI metric that is both inclusive and focused, we excluded economic development from the RSI metric.

Second, the RSI does not include special focus institutions, specifically we excluded vocational and career technical colleges. Our justification for excluding these types of institutions is because the data reporting systems and structures were found to differ from those that two-year colleges and four-year universities reported to IPEDS. Because we have an interest in integrating the RSI metric with national data systems like IPEDS, we chose to exclude these institutions from our RSI metric. Lastly, we want to make explicit that the RSI metric is measuring institutions who are best positioned to be rural-serving, not necessarily who is doing it *well*. This is an important distinction to consider because we created the RSI metric with an asset-based approach. Using an asset-based lens meant we were intentional about creating an RSI metric that empowered the very institutions and rural communities the metric was meant to serve. We did this through our development of a metric that offers institutions the power to determine their own scale of rural-*servingsness*. In doing so, we considered the politics of data and our position as researchers/data analysts (Dixon-Román, 2017; Stone, 2020) and intentionally worked to equalize as best we could the power dynamics that are situated within the construction of an RSI metric. As such, we want to make clear that looking at institutions who are serving rural communities *well* goes beyond the scope of the RSI metric.

Throughout this project, we were guided by our commitment to create a metric that is designed primarily to support the work of others. We hope that our discussion of the process for creating the RSI metric will provide more clarity and transparency to those who share our interest in rural higher education.

REFERENCES

Dixon-Román, E. 2017. *Inheriting possibility: Social reproduction and quantification in education*. University of Minnesota Press.

NCES Locale. 2016.

https://nces.ed.gov/programs/edge/docs/NCES_LOCALE_USERSMANUAL_2016012.pdf

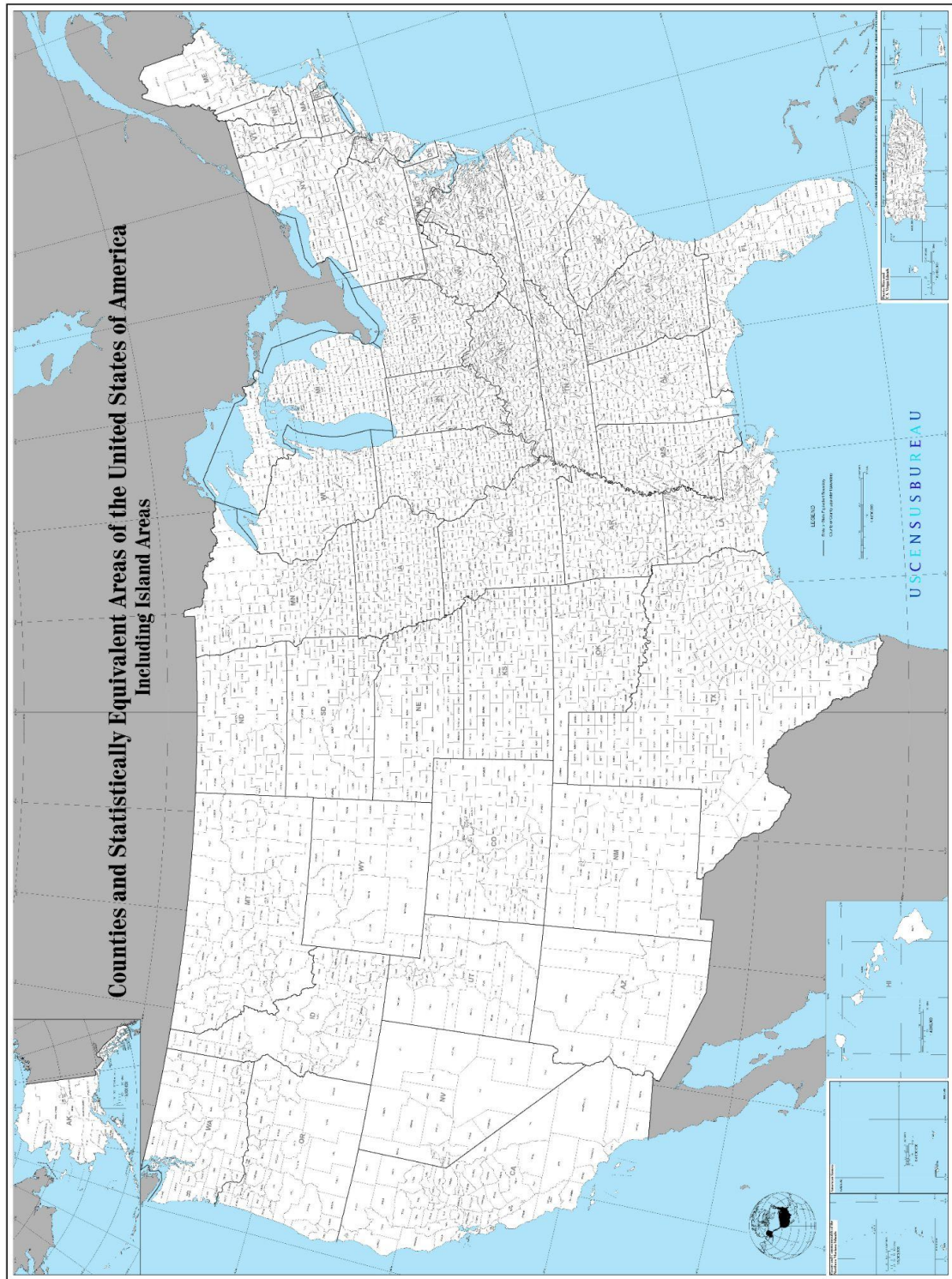
Stone, D. (2020). *Counting: How we use numbers to decide what matters*. Liveright Publishing.

USDA-ERS 2015.

<https://www.ers.usda.gov/webdocs/DataFiles/48652/ERSCountyTypology2015Edition.xls?v=7503.8>

Census 2010. <https://www.govinfo.gov/content/pkg/FR-2010-06-28/pdf/2010-15605.pdf>

Appendix: Map of US Counties



Source: US Census Bureau

https://www2.census.gov/geo/maps/general_ref/us_base/stco2003/stco2003.pdf