Disclosure Avoidance and the 2020 Census: How the TopDown Algorithm Works

2020 Census Briefs

By the Population Reference Bureau and
the U.S. Census Bureau's 2020 Census Data Products and Dissemination Team

C2020BR-04

March 2023

This is the third in a series of briefs describing how disclosure avoidance methods are being applied to 2020 Census data products and the implications of those methods for data users. This brief describes how differential privacy works and how it is applied to the 2020 Census Redistricting Data and Demographic and Housing Characteristics File (DHC). Additional information is available in the U.S. Census Bureau's handbook, "Disclosure Avoidance for the 2020 Census: An Introduction" and in "Disclosure Avoidance and the 2020 Census Redistricting Data."[1]

WHAT IS DISCLOSURE AVOIDANCE AND WHY IS IT IMPORTANT?

At the Census Bureau, **disclosure avoidance** is defined as a process used to protect the confidentiality of respondents' personal information. The Census Bureau has applied disclosure avoidance methods for decades to keep respondents' information confidential and maintain public trust in the data.

Differential Privacy

Differential privacy is the scientific term for a disclosure avoidance framework used to protect the privacy of respondents and the confidentiality of their data in the Census Bureau's published data products. The Census Bureau uses differential privacy on some of our published data products. It forms the foundation of the Disclosure Avoidance System (DAS) used to protect 2020 Census respondent confidentiality, and it is part of a broader family of disclosure avoidance approaches known as "formal privacy" that quantify the disclosure risk associated with each and every statistic published.

Differentially private disclosure avoidance methods infuse *statistical noise*—small random additions or subtractions—into the data to reduce the risk that someone could reidentify any person or household by combining the data from multiple statistics either through record linkage or other types of attacks.

The idea of using noise to protect confidentiality is not new; the Census Bureau has used similar techniques for decades. For a data product as large and detailed as the decennial census, and with threats posed by advances in computing technology and the rapid growth in the number of commercially available databases on people and households, differential privacy is the best science available at present to protect 2020 Census respondent confidentiality while minimizing the impact on statistical validity.

For more on why the Census Bureau chose differential privacy as a tool to protect the 2020 Census, refer to "Why the Census Bureau Chose Differential Privacy."[2]

[1] U.S. Census Bureau, "Disclosure Avoidance for the 2020 Census: An Introduction," <www.census.gov/library/publications/2021/decennial/2020-census-disclosure-avoidance-handbook.html>; and U.S. Census Bureau, "Disclosure Avoidance and the 2020 Census Redistricting Data," 2023, <www.census.gov/library/publications/2023/decennial/c2020br-02.html>.

[2] U.S. Census Bureau, "Why the Census Bureau Chose Differential Privacy," 2023, <www.census.gov/library/publications/2023/decennial/c2020br-03.html>.

HOW DOES DIFFERENTIAL PRIVACY WORK?

Imagine the image on a television screen: What appears to be a clear, crisp picture is actually composed of millions of pixels—tiny dots of color. If you were to zoom in, you could identify individual pixels. Applying differential privacy to the census data is like introducing small changes in color or hue to those individual pixels. The changes affect the appearance of individual pixels, but the overall picture remains when you zoom back out (Figure 1).[3]

Differential privacy works by adding statistical noise to data—increasing or decreasing some values in the published data, just like changing the hue in some of the pixels on a screen. The values associated with a particular data point are changed like the individual pixels in the picture, but when you look at the aggregation of those data points in a published statistic, the picture remains.

Adding less noise improves accuracy but increases the potential for disclosure. That risk of disclosure is referred to as privacy loss. Tracking that balance between accuracy and protection occurs through a privacy-loss budget. So, if the system is structured to add less noise to a certain result (e.g., single year of age at the tract level), it will balance the budget by adding more noise to certain other results that may not require as much accuracy.

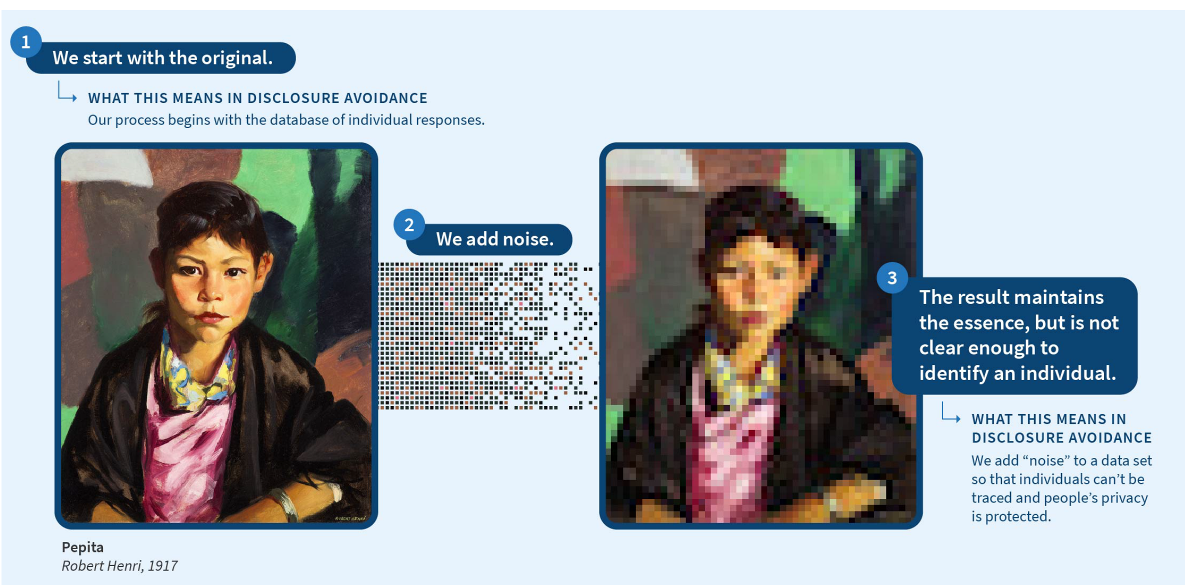
The total amount of noise can be set on a spectrum from high accuracy but no protection to no accuracy but high protection. High accuracy but no protection means that most of the noise added is at or close to zero—similar to making almost no changes to the pixels so the original is still recognizable. No accuracy but high protection means that large values are added or subtracted—similar to greatly changing most pixels and completely distorting the picture.

The privacy-loss budget is set by the Census Bureau's Data Stewardship Executive Policy Committee, to ensure fitness for use of the 2020 Census data, while effectively protecting confidentiality.

There are many ways to implement differential privacy. For 2020 Census Redistricting and DHC data, the Census Bureau implemented differential privacy through a series of formulas and steps called the TopDown Algorithm (TDA).

[3] The analogy is useful but incomplete. The Census Bureau adds noise to many cropped versions of the photo, not the pixels (response data) themselves, but the understanding the analogy conveys is useful and correct—controlled noise infusion is consistent with accuracy.

Figure 1.
Adding Noise Blurs Information So That Individuals Are Less Recognizable



Source: Population Reference Bureau.

HOW DOES THE TDA WORK?

The Census Bureau's DAS for 2020 Census Redistricting Data has two parts: differentially private algorithms and postprocessing. Both take place within a framework known as TDA. The differentially private algorithms add noise to the data consistent with the privacy-loss budget. Then, postprocessing imposes certain consistencies (for example, ensuring that the population totals for counties within a state sum to the state's total population). Even with post-processing, however, some inconsistencies remain in the data by design (for example, more households than people in a given geography). Steps in the TDA process are described in more detail below.

What Are the Steps in the TDA?

Working with input from stakeholders, the Census Bureau first compiled a list of tables for the 2020 Census Redistricting Data files.[4]

Next, the Census Bureau consolidated all the Redistricting Data tables into one detailed cross-tabulation that reflected all the variables (i.e., race, Hispanic origin, voting age, major group quarters [GQ] type, and housing occupancy status) for each geographic level (from the nation, to states, down to census blocks), all categories for each variable (for example, Hispanic origin includes Hispanic and not Hispanic), and combinations of those categories.

The TDA reads data from the Census Edited File (CEF) microdata—the database of original, individual-level census responses that have been edited for quality. CEF processing includes processes such as filling in missing or erroneous demographic and housing characteristic data. Starting with the list of tabulations described above, the TDA queries the 2020 CEF to produce certain tabulations such as counts of the voting-age population for the nation (the process is repeated for other geographies).

The TDA adds noise to cells in those tabulations using a differential privacy mechanism—adding or subtracting small amounts at random, as described above.

After noise is added, the TDA runs postprocessing routines to maintain certain characteristics in the data (described in the "Postprocessing the Noisy Statistics to Produce Tables" section of this brief). Examples include keeping state population totals unchanged and removing negative counts in the final tables.

Finally, the TDA uses the noise-infused, postprocessed data to generate privacy-protected microdata records for the entire nation. The microdata include rows for each individual person or household and their characteristics for each geography. These individual records contain every level of geography on the Census Bureau's geographic hierarchy. The records are exported into the Census Bureau's tabulation system to generate published data products.

The Census Bureau repeated a similar process for DHC—working with stakeholders to identify a list of tables to be published, creating one detailed cross-tabulation that reflected all geographic, variable, and category combinations based on planned published tables, reading in and querying the CEF, adding noise, and postprocessing. For DHC, postprocessing also includes steps to maintain consistency with the previously-published Redistricting Data.

The result is a privacy-protected dataset. Census blocks contain the most noise and the most protection where disclosure risk is usually greatest. A higher level of geography, such as census tracts, will have increased relative accuracy (noise as a share of population size) and will frequently have greater absolute accuracy (overall noise) as well.

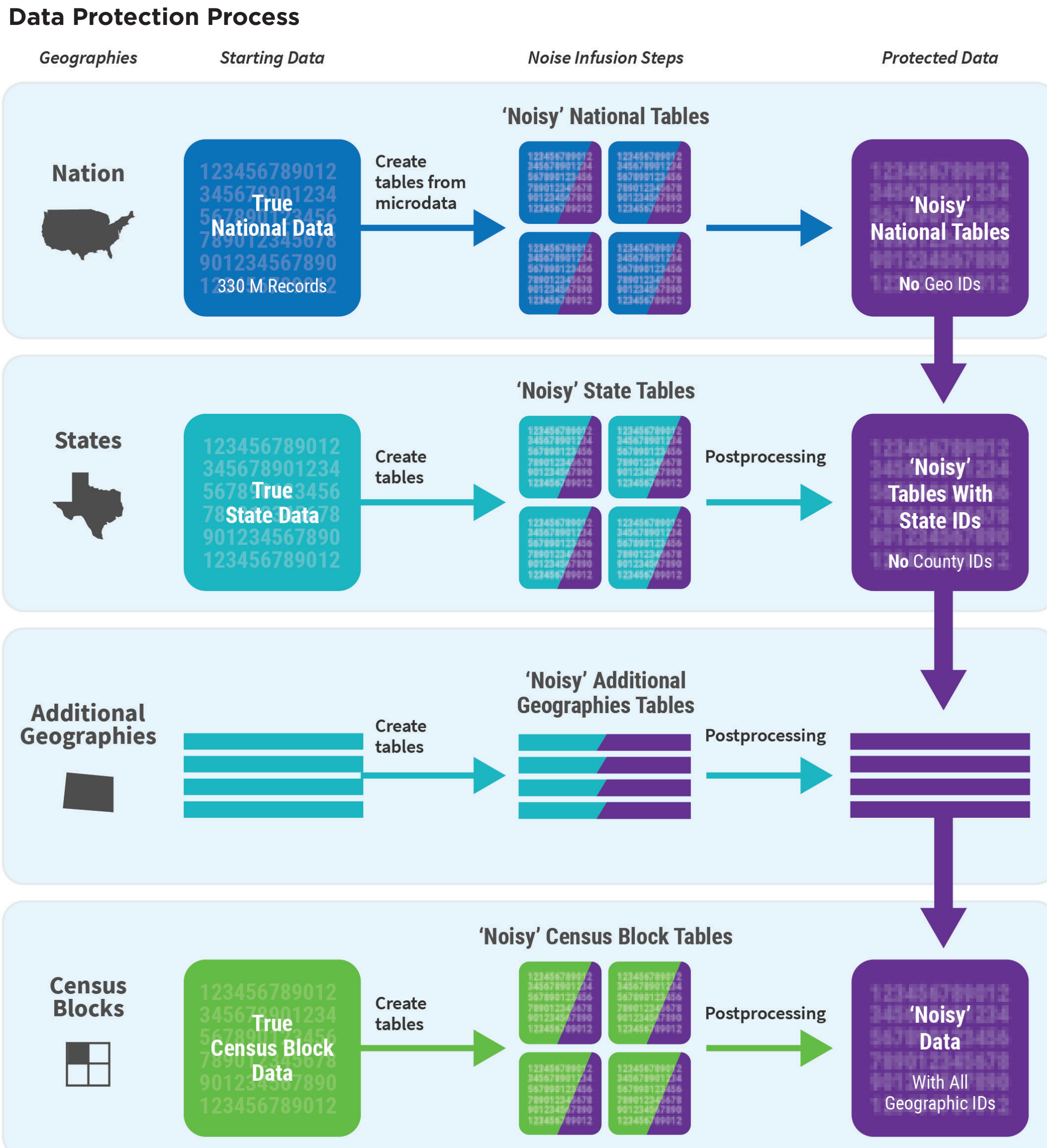
Starting With the Nation and Working Down to Blocks

The Census Bureau considers geographic nesting, such as counties within states, as it applies noise at different geographic levels.

- Starting at the national level, noise is infused into tables that represent all the combinations of characteristics across all the data. The result is a nation-wide, noise-infused dataset. These data include a noisy record representing every person in the United States but do not yet include geographic information (Figure 2).

[4] A detailed list of tables is available in the Census Bureau's "2020 Census State Redistricting Data (Public Law 94-171) Summary File Technical Documentation," <https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf>.

Figure 2.
Creating Differentially Private Data for the 2020 Census Redistricting Data and Demographic and Housing Characteristics Data

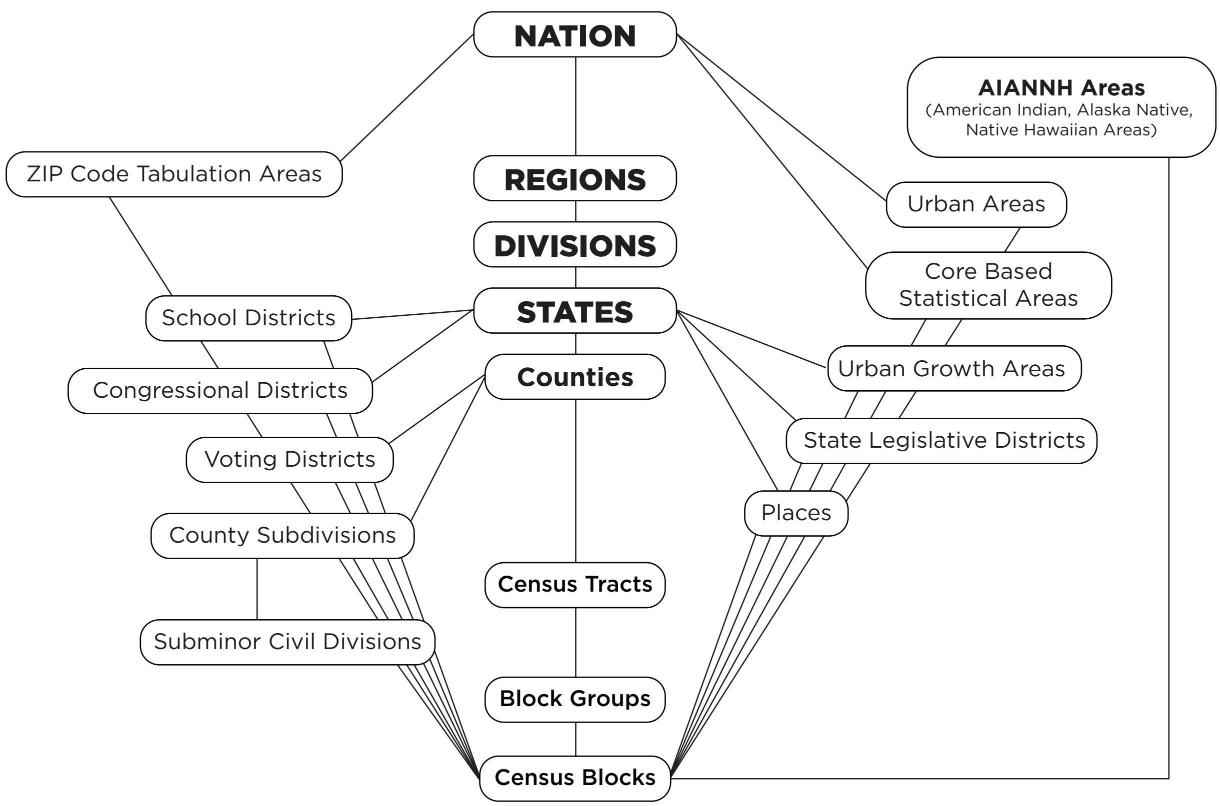


Source: Population Reference Bureau.

- Once the national data are set, the process is repeated for states. In the state step, mathematical optimization routines ensure that the state totals for different population or housing characteristics are as close as possible to the noisy measurements and these state totals, when added together, are consistent with the national data from the prior step. The result is an updated dataset that now includes state identifiers (but no county IDs).
- This optimization process is repeated for a series of ever smaller geographic units ending with census blocks.
- In the very last step, the tabular census block data are converted back into microdata that includes identifiers for all the geographic levels.

The TDA originally used the conventional hierarchy, or nesting scheme, of geographies for census data

Figure 3.
Standard Hierarchy of Geographic Areas



Source: U.S. Census Bureau.

products, sometimes called the geographic spine (Figure 3). Along the spine—the path down the middle from nation to block—each *child* geography perfectly nests within its *parent* geography. For example, all counties nest within one (and only one) state. Starting with the smallest unit along the geographic spine and working upward: blocks nest within block groups, block groups within census tracts, census tracts within counties, counties within states, states within divisions, divisions within regions, and regions within the nation.

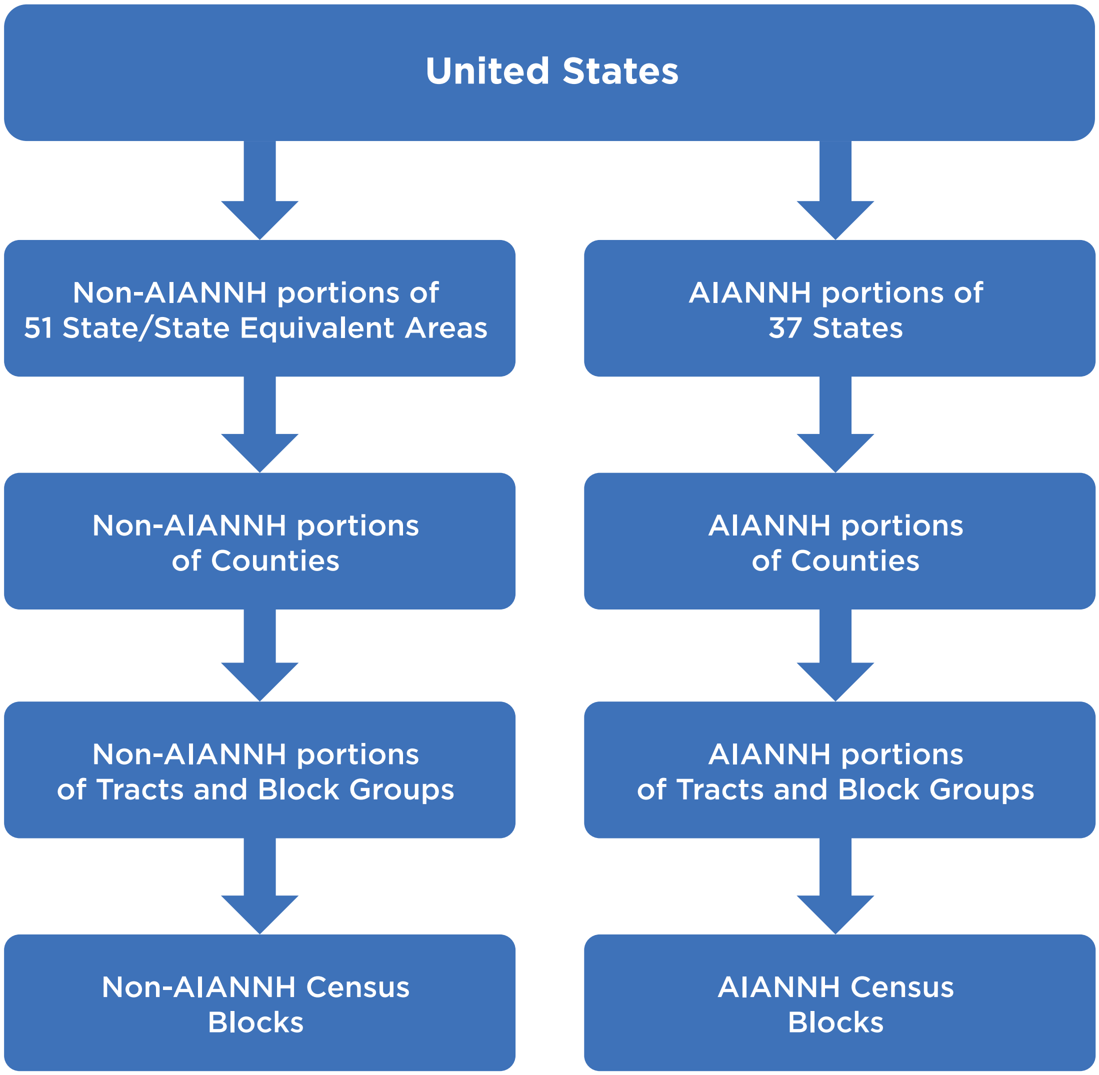
Some geographies, however, do not fit within the nesting scheme. School districts, for example, can be summed up from blocks and fit within states but do not necessarily follow block group or tract boundaries. To address feedback from data users about the importance of accurate data for some of the off-spine geographies, the Census Bureau made changes to the geographic hierarchy used for the TDA.

These nonnested, or *off-spine*, geographies are not part of the TDA processing routine. As a consequence, the noise-infused data for these areas may be noisier than those for the on-spine geographies.[5]

By making changes to the TDA's geographic hierarchy, the Census Bureau was able to improve data accuracy for targeted off-spine geographic areas. For example, the Census Bureau responded to feedback on the accuracy of counts for American Indian/Alaska Native/Native Hawaiian (AIANNH) areas. To address this, the Census Bureau changed how it processes AIANNH areas. For states with AIANNH areas, the AIANNH and non-AIANNH portions of the state are split and processed separately (Figure 4).

[5] Off-spine entities are formed by combining—adding or subtracting—combinations of on-spine entities. When on-spine geographic entities are combined like this, the DP noise accumulates in the off-spine entity. To control this cumulative error, TDA's optimized geographic spine minimizes the number of on-spine entities used to represent off-spine entities.

Figure 4.
Hierarchy for Disclosure Avoidance System Processing for Redistricting Data



Source: U.S. Census Bureau.

Within the TDA, all AIANNH areas in a state are grouped together for data processing. This reduces the tendency for the populations of AIANNH areas to be dispersed into surrounding areas and, thus, helps avoid undercounting those populations in the tribal areas. For example, at the state level, five American Indian areas in Kansas—the Iowa (KS-NE) Reservation and Off-Reservation Trust Land, the Kickapoo (KS) Reservation, the Prairie Band of Potawatomi Nation Reservation, the Sac and Fox Nation Reservation and Off-Reservation Trust Land, and the Kickapoo (KS) Reservation/Sac and Fox Nation Trust Land joint-use area—are processed together, separate from the rest of Kansas. At lower geographic levels, these individual tribal areas are then processed separately from each other.

Another important departure from the standard geographic hierarchy is how blocks—the smallest geographic unit the Census Bureau uses for tabulation—are grouped before being aggregated to tracts. Rather than using the Census Bureau's standard block groups, blocks are combined—sometimes in groups of nonbordering blocks—to improve the TDA's processing efficiency and reduce postprocessing error, especially for people living in GQ facilities such as correctional facilities, college/university student housing, or military quarters.

In most states, the District of Columbia, and Puerto Rico, these block combinations (called *optimized block groups* in the technical documentation) were redefined to more closely approximate places (such as cities). In 12 states, blocks were combined to more closely approximate minor civil divisions (cities, boroughs, and towns/townships). These are states where minor civil divisions, a type of county subdivision, serve as general purpose governments and provide the same government functions as incorporated places.

In the TDA, the Census Bureau processes all the smaller geographic units within a larger geographic area only after the larger area has been processed first, and uses TDA counts in the larger geographic units. This improves statistical accuracy by allowing TDA to identify large populations first and only later using those populations to try to identify smaller populations.

In addition, there is special processing for cases where a larger *parent* geography (such as a tract) has only one smaller *child* geography (such as a block group) within it. In these cases, the privacy-loss budget of the child geography is reallocated to the parent geography instead, and the algorithm only uses the noisy measurements for the parent geography.[6]

The Geographic Hierarchy for DHC Includes Additional Adjustments

For the DHC data, the Census Bureau made some additional modifications to the geographic processing hierarchy. The first change adds a geography to align with the annual Population Estimates (called *Population Estimates Primitive Geographies*). The second change is to introduce a new split-tract geography. These split tracts (called tract subsets) reflect the intersection of Population Estimates geographies, AIANNH areas, and tracts (Figure 5). These adjustments further improve data quality for key geographies.

Testing showed that the optimized geographic spine used for TDA processing yielded significant data quality improvements for the same level of protection.

[6] U.S. Census Bureau, Geographic Spines in the 2020 Census Disclosure Avoidance System TopDown Algorithm, <https://www2.census.gov/adrm/CED/Papers/CY21/2021-004-AbowdAshmeadCumingsMenonetal.pdf>.

Postprocessing the Noisy Statistics to Produce Tables

Invariants

The DAS departs from *textbook* differential privacy in one important way. The Redistricting and DHC data files include certain invariants—data that are kept exactly as enumerated with no noise added. Invariant statistics for the 2020 Census Redistricting Data are:

- Total number of people in each state, the District of Columbia, and Puerto Rico.
- Total number of housing units (but not population counts) in each census block, and all other geographic levels.
- Number of *occupied* GQ facilities (but not population counts) in each census block by the following types:
    - Correctional facilities for adults.
    - Juvenile facilities.
    - Nursing facilities/skilled-nursing facilities.
    - Other institutional facilities.
    - College/university student housing.
    - Military quarters.
    - Other noninstitutional facilities.

All other population and housing characteristics, including population counts for every geography below the state level, have had noise introduced. These invariants were added to meet statutory requirements (such as apportionment counts) and to retain as much publicly available information as possible (such as the number of housing units in a block and the presence or absence of occupied GQ facilities, which are important parts of the Local Update of Census Addresses Program).[7]

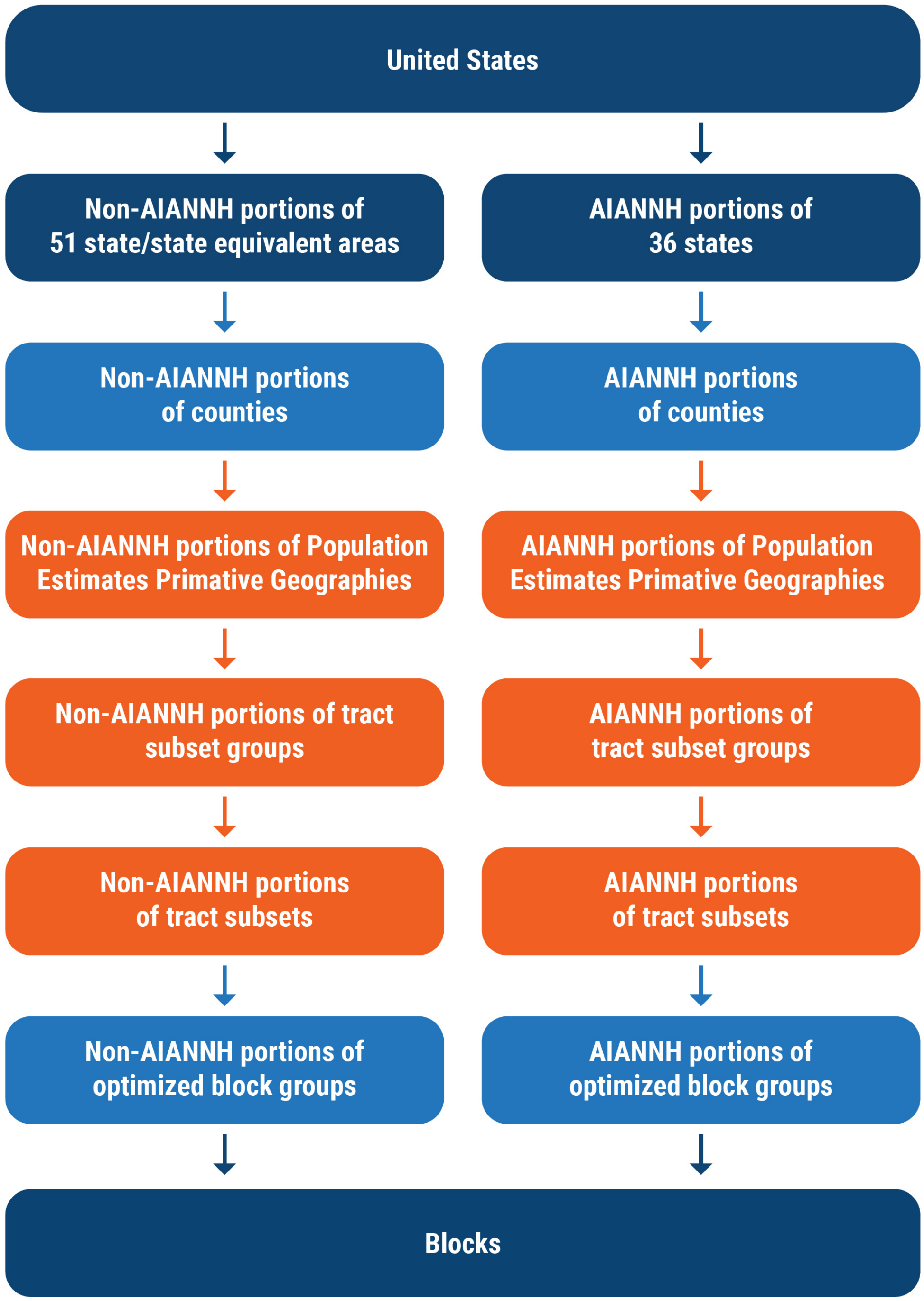
Additional Constraints

In addition to the invariants noted above, there are some constraints within the TDA that are applied at all geographic levels. For the Redistricting Data, these constraints include:

- Population and housing counts must be integers and may not be negative.
- The parts of a table must sum to subtotals and totals within the table.

[7] The Census Bureau's Master Address File has information regarding addresses that have been previously recorded as GQs, but if the final edited census response for that address indicates that it had no occupants on April 1, 2020, the address is removed from the census frame. The Census Edited File, therefore, contains no information about addresses that are unoccupied GQs.

Figure 5.
Geographic Hierarchy for the TopDown Algorithm



Source: U.S. Census Bureau.

- Population-related counts must be consistent within tables, across tables, and across geographies. For example, the population by race must sum to the total population, and the population in each county within a state must sum to the state's total population.
- Housing-related counts must be consistent within tables, across tables, and across geographies. For example, the number of occupied and vacant housing units must sum to the total number of housing units, and the number of vacant housing units in each county within a state must sum to the state's total number of vacant housing units.
- If there are zero housing units and zero GQ facilities in a geography, then no people may be assigned to that geography.
- The number of people per GQ facility must be greater than or equal to 1. This requirement translates the previously discussed requirement that the number of occupied GQ facilities (by major type) has no noise.
- The number of people per housing unit must be less than or equal to 99,999, and the number of people per GQ facility must be less than or equal to 99,999.
- There are zero people aged less than 18 in GQ type 301, *Nursing facilities/skilled nursing facilities*.

DHC includes additional constraints to preserve demographic reasonableness. To ensure published tables in the DHC data are consistent with previously published Redistricting Data, counts in the DHC are constrained to the published Redistricting Data totals. For example, the household population for a given location will be the same in Redistricting Data tables and in DHC tables. In addition:

- Housing units cannot have both an owner/renter tenure value and a vacancy value.
- The count of housing units has no noise, and the count of *occupied* GQ facilities (by major type) have no noise.
- Each relationship status must contain valid age ranges (e.g., householder must be between 15 and 115 years).
- Each GQ person must have a valid age range dependent on GQ type (e.g., people in Nursing facilities/Skilled-nursing facilities must be between 20 and 115 years).[8]

While several constraints have been applied in the TDA, some inconsistencies may remain in the Redistricting Data and the DHC data. These inconsistencies are described in detail in "Disclosure Avoidance and the 2020 Census Redistricting Data."[9]

Multipass Optimization

While adding noise to the data does not introduce bias (since the additions and subtractions are random, independent, and will be roughly equal), the postprocessing step introduces bias by removing negative values or imposing other constraints on the resulting data.

To reduce bias for small geographic areas and population subgroups, the Census Bureau implemented a postprocessing routine, called *multipass optimization*. Multipass optimization processes certain elements of the data first and then uses those results as an input to subsequent steps.

At the national level, the state level, and then for lower levels of geography, multipass first processes the population count for each unit within that geographic level (for example, the population for each county within a state or each census tract within a county). Next, the algorithm generates the remaining statistics, constraining those statistics to the population counts determined in the first pass.[10]

Processing total population first helps limit the tendency of small counts to get larger, large counts to get smaller, and limits the likelihood that TDA noise introduced into the total population counts could be inadvertently correlated with other demographic characteristics.

How Much Noise Is Added to the Data? Understanding the Privacy-Loss Budget Allocation

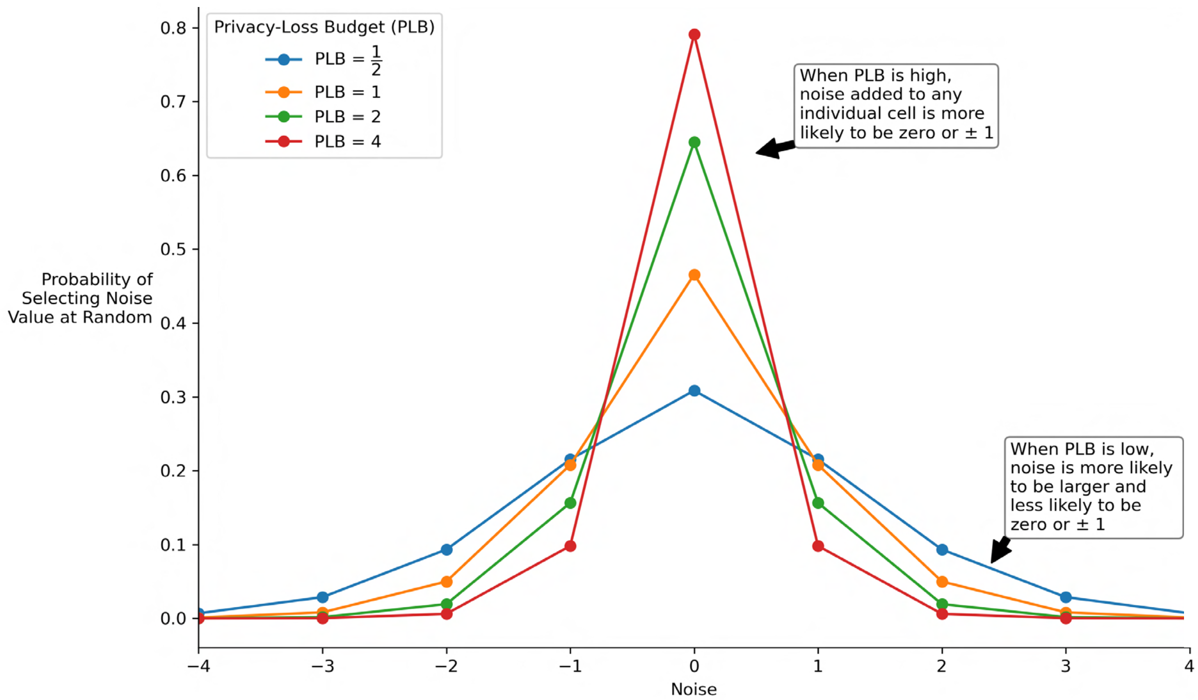
Noise must be added throughout the published tables to ensure the confidentiality of the data. That means the overall privacy-loss budget is distributed across published census products (tables and microdata), population and housing characteristics, and geographic levels. Allocating more budget to improve accuracy for one dimension of the data (such as more accurate total population counts for counties) means

[8] The Census Bureau's Master Address File has information regarding addresses that have been previously recorded as GQs, but if the final edited census response for that address indicates that it had no occupants on April 1, 2020, the address is removed from the census frame. The Census Edited File, therefore, contains no information about addresses that are unoccupied GQs.

[9] U.S. Census Bureau, "Disclosure Avoidance and the 2020 Census Redistricting Data," 2023, <www.census.gov/library/publications/2023/decennial/c2020br-02.html>.

[10] John M. Abowd and Victoria A. Velkoff, "Modernizing Disclosure Avoidance: A Multipass Solution to Post-Processing Error," U.S. Census Bureau, Washington, DC, 2020, <www.census.gov/newsroom/blogs/research-matters/2020/06/modernizing_disclosu.html>.

Figure 6.
The Privacy-Loss Budget Acts as a Dial on the Level of Noise



Source: U.S. Census Bureau.

that there is less budget for accuracy in another dimension (such as race detail).

As the privacy-loss budget (represented by the Greek letter *rho*) rises, the noise added to any given cell is increasingly likely to be zero (Figure 6). Lower budgets imply less accuracy and more protection, as the noise distribution spreads out away from zero, and larger amounts of noise added to a cell become increasingly likely. In the most extreme cases, a budget of zero would reflect complete noise with no accuracy. Conversely, a budget of infinity would reflect complete accuracy with no noise.

The privacy-loss budget is not the only factor that influences the shape of the distribution—and thus the likelihood of selecting random noise far from zero. The type of statistical distribution also matters. Some statistical distributions allow, for example, for sizeable outliers. But for decennial census data, confidentiality concerns need to be balanced with the accuracy of the data and adding large amounts of noise to some cells may harm the data's fitness for use.

For the 2020 Census, the Census Bureau chose a statistical distribution that yields substantially greater accuracy for comparable disclosure risk. The selected distribution is more tightly clustered around zero with a lower probability of choosing unusually large noise values at random.[11] Therefore, there are fewer outliers.

[11] This distribution is called the discrete Gaussian distribution in the differential privacy literature. Relative to other commonly used distributions in the differential privacy literature, such as the Laplace distribution—our two-sided geometric distribution, this distribution has thinner "tails," i.e., outlier observations are less common.

For more specifics on the amount of noise, bias, implausible results, and user guidance, refer to:

- Disclosure Avoidance for the 2020 Census: An Introduction <www.census.gov/library/publications/2021/decennial/2020-census-disclosure-avoidance-handbook.html>.
- Disclosure Avoidance and the 2020 Census Redistricting Data <www.census.gov/library/publications/2023/decennial/c2020br-02.html>.

Interpreting the Privacy-Loss Budget

- In the TDA, the balance between accuracy and degree of protection is managed using the parameter rho.
- *Rho* promises an upper bound on how much someone else can learn about you, relative to what they could have learned if your information were replaced with random values. Research shows that this interpretation is very robust; it is true no matter how much external information an attacker has, no matter how clever they are or what algorithms they know, and no matter how much computing power they have.
- The portion of *rho* spent only on the block level has a similar interpretation: it promises a bound on how much someone else can learn about you, relative to what they could have learned if your block had been replaced by another block in your block group (but your other information was left as is).
- *Rho* is not a measure of harm, and it does not reflect how we expect an attacker to use the information. Instead, rho limits how much information about a person or household could be leaked, and Title 13 requires the Census Bureau protect all respondent data.
- You can get different levels of accuracy given the same rho. For example, as improvements were made to the TDA design during Census Bureau internal experiments, there were sometimes significant increases in accuracy without corresponding increases in *rho*.

WHAT ARE THE IMPLICATIONS FOR DATA USERS?

The disclosure avoidance system results in some issues that data users should be aware of when using Redistricting Data. For guidance, refer to "Disclosure Avoidance and the 2020 Census Redistricting Data."[12] Guidance for the DHC is forthcoming.

HOW HAS DATA USER FEEDBACK INFORMED THE PLANNING PROCESS?

The Census Bureau received invaluable feedback on disclosure avoidance from external stakeholders that informed our efforts and decision-making. These came via the 2020 DAS email <2020DAS@census.gov>, advisory meetings, tribal consultations, and comments provided during presentations at conferences and the "Disclosure Avoidance Webinar Series."[13, 14] The Census Bureau and external data users identified several issues with preliminary versions of the DAS that needed additional attention before it could be applied to the 2020 Census data, including:

- Situations where small populations tended to gain population, whereas larger populations tended to lose population, such as rural counties with small populations gaining population and large urban counties losing population.
- Limitations of the noise-infused data for emergency planning operations such as the population of older adults living alone.

[12] U.S. Census Bureau, "Disclosure Avoidance and the 2020 Census Redistricting Data," 2023, <www.census.gov/library/publications/2023/decennial/c2020br-02.html>.

[13] To view any webinar in the series, visit <www.census.gov/data/academy/webinars/series/disclosure-avoidance.html>.

[14] Feedback received to date can be accessed by visiting Round 1 feedback, 2010 Demonstration Data Demographic and Housing Characteristics File (DHC) v. 2022-03-16 (6/23/2022) <https://www2.census.gov/programs-surveys/decennial/2020/program-management/round_1_feedback.pdf>, and Round 2 Feedback, 2010 Demonstration Data Demographic and Housing Characteristics File (DHC) v. 2022-08-25 (8/25/2022) <https://www2.census.gov/programs-surveys/decennial/2020/program-management/round_2_feedback.pdf>.

- Issues for populations living on American Indian reservations such as large changes in population counts.
- Problems with the accuracy of census data for certain geographic areas that do not follow the Census Bureau's standard geographic hierarchy such as school districts (Figure 5).[15]
- Identification of extreme outliers such as areas with children under the age of 18 but no adult population.
- Identification of demographic and housing characteristics and geographies that needed additional accuracy.

The Census Bureau used this feedback to make improvements to the DAS and adjust privacy settings to improve overall accuracy for geographic areas and other characteristics but never to favor a particular demographic group over another. As a result of this work, the Census Bureau was able to greatly reduce many of these limitations.

Data user feedback was also incorporated in a series of demonstration products to test whether the noise-infused data were fit for use.[16] Advanced data users may download demonstration data that was generated by applying the 2020 DAS to the 2010 Census data.

While not all data user concerns were, or could be addressed, the Census Bureau has continued gathering feedback to help inform future products. In addition, the Census Bureau is working on tools that will help data users interpret the accuracy of the data for their specific needs.

WHERE CAN I LEARN MORE?

- Disclosure Avoidance and the 2020 Census Redistricting Data <www.census.gov/library/publications/2023/decennial/c2020br-02.html>
- Disclosure Avoidance for the 2020 Census: An Introduction <www.census.gov/library/publications/2021/decennial/2020-census-disclosure-avoidance-handbook.html>
- Why the Census Bureau Chose Differential Privacy <www.census.gov/library/publications/2023/decennial/c2020br-03.html>
- 2020 Census Data Product Planning Crosswalk (v. 2022-08-25) <https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/02-Demographic_and_Housing_Characteristics/2022-08-25_Summary_File/2022-08-25_2020_Census_Data_Product_Planning_Crosswalk.xlsx>
- Disclosure Avoidance: Latest Frequently Asked Questions <www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-updates/2020-das-faqs.html>
- 2020 Decennial Census: Processing the Count: Disclosure Avoidance Modernization <www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html>
- Disclosure Avoidance Webinar Series <www.census.gov/data/academy/webinars/series/disclosure-avoidance.2021.List_882320526.html#list-tab-List_882320526>

You can also subscribe to the Census Bureau's "2020 Census Data Products Newsletter" for timely updates and contact us at <2020DAS@census.gov> if you have questions.[17]

[15] Committee on National Statistics, "2020 Census Data Products: Data Needs and Privacy Considerations" workshop <https://sites.nationalacademies.org/DBASSE/CNSTAT/DBASSE_196518>.

[16] U.S. Census Bureau, "Developing the DAS: Demonstration Data and Progress Metrics, Detailed Summary Metrics for Production Settings," June 8, 2021, <www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html>.

[17] U.S. Census Bureau, "Decennial Census: Data Products and Operational Updates," <https://public.govdelivery.com/accounts/USCENSUS/signup/15409>.