METHODOLOGY FOR THE SUBCOUNTY TOTAL RESIDENT POPULATION ESTIMATES (VINTAGE 2023): APRIL 1, 2020 TO JULY 1, 2023[1]

OVERVIEW

The Population Estimates Program (PEP) at the U.S. Census Bureau produces estimates of total resident population for all areas of general-purpose government on an annual basis. The subcounty areas consist of both incorporated places (such as cities, boroughs, and villages) and minor civil divisions (such as towns and townships). We produce county-level population estimates with a cohort-component method.[2] This method measures change since the date of the last census using administrative data on births, deaths, and domestic and international migration. Then we use updated subcounty housing unit estimates to distribute county household population to subcounty areas.[3] Finally, subcounty group quarter (GQ) population estimates are added to estimate the resident population.

METHOD

The 2020 Census base counts of housing units and associated population of each governmental unit are geographically updated each year to reflect legal boundary changes reported in the Boundary and Annexation Survey (BAS), from other geographic program revisions, and from 2020 Census corrections when available. These updates are applied at the primitive level of geography. *Primitive geographies* are a complete, mutually exclusive partition of the United States which can thus be aggregated into all geographic units for which PEP produces estimates. This enables the processing of estimates despite the fact that subcounty geographic units are not always mutually exclusive, nor do they always exist within a single county.

The subcounty population estimates may include revisions from accepted challenges to the estimates and special censuses for full jurisdictions.[4,5] Both types of revisions are incorporated through the population components. The Vintage 2023 population estimates include accepted challenges for seven areas but no special censuses.

The Census Bureau develops the subcounty population estimates for the household and group quarters populations separately, and then combines them to calculate the resident population. We estimate the household population by applying the "Distributive Housing Unit Method" to the county-level household population to distribute it to each subcounty area as illustrated by the following steps:

---

[1] Vintage 2023 data products are associated with Data Management System projects P-6000042, P-7501659, and P-7527355. The U.S. Census Bureau reviewed these data products for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release (CBDRB-FY23-0063).

[2] For a more detailed description of the methodology used to produce the Vintage 2023 estimates of state and county population, please refer to: https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2020-2023/methods-statement-v2023.pdf.

[3] For a more detailed description of the methodology used to produce the Vintage 2023 estimates of housing units, please refer to: https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2020-2023/2023-hu-meth.pdf.

[4] Population Estimates Challenge Program results are available here:
 https://www.census.gov/programs-surveys/popest/about/challenge-program/results.html.

[5] Special Census Program results are available at https://www.census.gov/programs-surveys/specialcensus/data_products/official_counts.html and special census areas included in previous series of estimates are available here: https://www.census.gov/programs-surveys/popest/about/special-census.html.

Step 1 - Producing an Uncontrolled Subcounty Household Population Estimate

The Vintage 2023 uncontrolled subcounty household population estimates for each date begin with the Vintage 2023 housing unit estimates. We multiply each year's estimate by the Household Population Per Housing Unit ratio, which is the quotient of the 2020 Census household population[6] and the 2020 Census housing units. This produces the final uncontrolled population estimate for each year.[7]

Step 2 - Producing a Controlled Subcounty Household Population Estimate

We then control the subcounty estimates of household population from Step 1 so that they sum to the published county totals for each date. To do this, we divide the Vintage 2023 county-level household population estimate by the sum of the uncontrolled subcounty household population estimates within the county. We multiply this adjustment or "rake factor" by the uncontrolled subcounty household population estimates calculated in Step 1. This calculation produces the controlled subcounty household population estimate. We round this controlled estimate using a variation of the greatest mantissa methodology.[8]

Step 3 – Estimating the Group Quarters Population

The GQ component of the total estimate is a combination of persons residing in institutional facilities and non-institutional facilities.

The institutional facilities include four types:

- Correctional facilities for adults
- Juvenile facilities
- Nursing facilities/Skilled-nursing facilities
- Other institutional facilities

The non-institutional facilities include three types:

- College/university student housing
- Military quarters
- Other noninstitutional facilities

We use GQ population data from two sources to estimate subcounty populations:

- 2020 Census counts of the GQ population by facility type for each subcounty area,[9] including any post-2020 additions to the GQ population when available; and

---

[6] The 2020 Census household population totals feature confidentiality protections applied using the 2020 Census Disclosure Avoidance System prior to the processing of the subcounty estimates.
[7] We apply a national-level Household Population Per Housing Unit ratio to new housing estimates in areas with zero values for population or housing as of the 2020 Census.
[8] Greatest mantissa rounding uses the decimal of each value to rank them, and then consecutively adds a single integer to the highest-ranked values (without decimals) until the population sums to the original total.
[9] The 2020 Census counts of the GQ population by facility type feature confidentiality protections applied using the 2020 Census Disclosure Avoidance System prior to the processing of the subcounty estimates.

- A time series of individual GQ facility records from the Group Quarters Report (GQR) that the PEP prepares based upon annual updates from state representatives of the Federal-State Cooperative for Population Estimates (FSCPE).

These two sets of GQ population data are used to derive a time series of GQ population through the following process:

Step 3A – Aggregate the GQ population from the 2020 Census to the subcounty level by the seven facility types.

Step 3B – Sum the facility-level GQ populations from the GQR to the subcounty level by the seven facility types for each estimate date in the time series.

Step 3C – Calculate the year-to-year change indicated by the aggregated GQR time series of population and add this time series of change to the aggregated 2020 Census data. This creates a census-based time series of GQ population at the subcounty level for each of the seven facility types.

Step 4 – Producing the Final Subcounty Population Estimate

To produce the final subcounty resident population estimate, we add the controlled household population estimate to the total GQ population estimate.