



2020 Census Data Product Supplemental Demographic and Housing Characteristics File (S-DHC)

PRERELEASE WEBINAR

SEPTEMBER 16, 2024

Closed captioning: click on the "cc" button in the lower left corner of the Webex window.

SPEAKERS

- Stacy Gimbel Vidal, Public Information Office
- Evan Brassell, Assistant Division Chief for Housing Characteristics, Social, Economic and Housing Statistics Division
- Michael Hawes, Senior Statistician for Scientific Communication
- Scott Holan, Senior Research Fellow

AGENDA

- Data release timing and logistics
- Data product overview
- Data quality and guidance
- Accessing and exploring the data
- Q&A

TIMING AND LOGISTICS

Media Embargo – Tuesday, September 17, 2024

- Begins at 10 a.m. EDT

Public Release – Thursday, September 19, 2024

- Release at 12:01 a.m. EDT
- Tables live on data.census.gov and API by 10 a.m. EDT

Click the [Embargo Subscribers](#) link on the press kit to request embargo access.

DATA PRODUCT OVERVIEW

KEY POINTS

- Tables combine characteristics about people and households to provide:
 - The average size of families and households.
 - Counts of people living in certain types of households.
- Most of the tables are also available by race and Hispanic origin.

TABLES

- Average household size by age.*
- Household type for the population in households.
- Population under the age of 18 years by relationship and household type.*
- Population in families by age.*
- Average family size by age.*
- Family type and age for own children under 18 years.
- Total population in occupied housing units by tenure.*
- Average household size of occupied housing units by tenure.*

*Tables marked with an asterisk are also available by race and Hispanic origin.

Note: The final table shells are included on the “SDHC Tables” tab of the [2020 Census Data Product Crosswalk](#).

RACE AND ETHNICITY GROUPS

The tables repeated by major race and ethnicity groups are identified with the letters A through I:

- A. White alone
- B. Black or African American alone
- C. American Indian and Alaska Native (AIAN) alone
- D. Asian alone
- E. Native Hawaiian and Other Pacific Islander (NHPI) alone
- F. Some Other Race alone
- G. Two or More Races
- H. Hispanic or Latino
- I. White alone, not Hispanic or Latino

EXAMPLE DIFFERENCES BETWEEN DHC AND S-DHC TABLES

DHC (separate tables)

- How many married couple households had children living in them.
- The number of children.

Household characteristic

Person characteristic

S-DHC

- How many children live in married couple families.

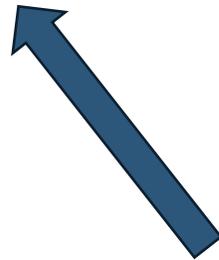
Person characteristic

Household characteristic

ANOTHER EXAMPLE

DHC

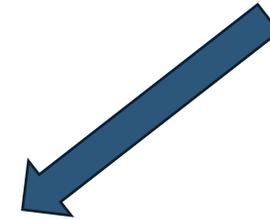
- How many households owned their home with a mortgage or a loan.



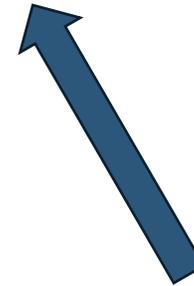
Household
characteristic

S-DHC

- How many people live in a home that's owned with a mortgage or a loan.



Person
characteristic



Household
characteristic

GEOGRAPHY

- Nation
- States, the District of Columbia, and Puerto Rico

The S-DHC is provided for limited geographies because of concerns about meeting data quality and accuracy standards given confidentiality considerations.

RACE AND ETHNICITY: MAKING COMPARISONS

- Consider improvements to how we collected, coded and tabulated race data.
 - Blog: [Improvements to the 2020 Census Race and Hispanic Origin Question Designs, Data Processing, and Coding Procedures](#)
- Use caution when comparing 2020 Census Asian, NHPI and AIAN race data with 2019 or earlier American Community Survey (ACS) data.
 - Improvements were implemented in the ACS in 2020.
- The 2020 Census Hispanic origin data from the ethnicity question **are** comparable with the 2010 Census and with the ACS regardless of data year.

HOUSEHOLD TYPE: COMPARISONS WITH 2010 CENSUS

- Inclusion of same-sex couples

2010

Same-sex married couples

If no children or other relatives present – edited and shown as unmarried partners and included in nonfamily households.

2020

Same-sex married couples

Included in estimates of families because they are married couples.

- The household type 2020 Census data are generally comparable with the ACS data.

HOUSEHOLD TYPE: COMPARISONS WITH 2010 CENSUS CONT.

- Updated labels

2010

husband-wife families

male householder, no wife
present

female householder, no husband
present

2020

→ married couple families

→ male householder, no spouse
present

→ female householder, no spouse
present

TENURE: COMPARISONS WITH 2010

- No major changes to housing questions from 2010.
- Housing data are generally comparable to the 2010 Census and the American Community Survey (ACS).

SIMILAR ACS DATA

B11003 | Family Type by Presence and Age of Own Children Under 18 Years

	Estimate	Margin of Error
▼ Total:	84,045,684	±133,737
▼ Married-couple family:	61,421,188	±171,809
▼ With own children of the householder under 18 years:	22,871,453	±86,999
Under 6 years only	5,064,116	±46,631
	4,611,967	±43,016

S1101 | Households and Families

Label	Total		Married-couple family household	
	Estimate	Margin of Error	Estimate	Margin of Error
▼ HOUSEHOLDS				
Total households	131,332,360	±130,190		
Average household size	2.49	±0.01		

B25010B | Average Household Size of Occupied Housing Units by Tenure (Black or African American Alone Householder)

	Estimate	Margin of Error
▼ Average household size --		
▼ Total:	2.41	±0.01
Owner occupied	2.66	±0.01
Renter occupied	2.21	±0.01

SO WHY USE THE S-DHC?

- Based on the 2020 Census data
 - Despite challenges during collection, over a tens of million households directly responded to the census
 - Far more responses than any Census survey, including the ACS

DATA QUALITY AND GUIDANCE

DISCLOSURE AVOIDANCE FOR THE S-DHC

- The Census Bureau is using a disclosure avoidance framework known as *differential privacy* to protect the confidentiality of individuals' census response information.
- Differential privacy protects against disclosure by adding statistical noise—random additions or subtractions—to every published statistic to reduce the likelihood that characteristics about a specific person or household can be accurately inferred using any combination of the published data.
- For the S-DHC, this process is achieved through the PHSafe algorithm.

TRUNCATION

- Linked person and household data have higher disclosure risk than either person or household data alone.
- PHSafe adds an additional disclosure avoidance step that is not applied to other 2020 Census data products to further protect the confidentiality of households' characteristics.
- When the count of people in an enumerated household exceeds a set threshold based on the number of people, individuals are removed or “truncated” at random until the household meets that household size threshold.
- Those thresholds are as follows:
 - 10 people for tables PH2 (population in households), PH4 (population in families) and PH7 (population in occupied housing units);
 - 6 people for tables PH3 (population under 18 years in households) and PH6 (own children under 18 years).

NOISE INFUSION

- Following the truncation step, PHSafe produces preliminary tabulations of characteristics and then adds or subtracts noise to those results to produce a set of preliminary counts referred to as “noisy measurements.”
- Under differential privacy, the amount of noise added or subtracted to protect confidentiality is determined by a “privacy-loss budget.”
- The S-DHC’s privacy-loss budget was set to meet predetermined accuracy targets for the noisy measurements of the truncated (not enumerated) values.
- The noisy measurements then undergo statistical post-processing to produce the published S-DHC statistics.

IMPACTS OF DISCLOSURE AVOIDANCE

User guidance regarding the inconsistencies resulting from noise infusion:

- Don't expect consistency across S-DHC tables;
- Don't expect consistency when aggregating S-DHC data and comparing it to other S-DHC statistics;
- Don't expect consistency between the S-DHC and other 2020 Census data products;
- Use caution when aggregating published counts to produce statistics for custom groups or geographies. Aggregations of the S-DHC data get noisier in proportion to the number of counts that are included in the aggregation.
- Be aware that household size averages are impacted by the truncation of very large households, and that the magnitude of these effects may vary by geography and demographic group.

STATISTICAL POST PROCESSING AND S-DHC

- Recall, the privacy of the truncated enumerated counts are protected using a form of differential privacy.
- Applying disclosure avoidance methodology results in a (“noisy”) privacy-protected truncated count.
- Privacy-protected truncated counts can be *statistically post processed* (SPP) – outputs of the SPP inherit the associated confidentiality guarantees.
- Motivation: SPP reduces implausible results, improves accuracy, and provides measures of disclosure avoidance-related uncertainty.

THE SPP APPROACH

- SPP model starts with the (“noisy”) privacy-protected measurements.
- SPP uses a Bayesian approach along with the noisy measurements and logical constraints (such as non-negativity and lower bounds) from the enumerated data.
- SPP then generates a new set of estimates.
- The new estimates are typically more precise (relative to the truncated, enumerated value) than the noisy measurements.
- SPP corrects some of the implausible data and generates a measure of accuracy known as a “credible interval” (CI).

CREDIBLE INTERVALS

- Similar to MOEs, the (90%) CI represents a range of values that contain the truncated, enumerated value with 90 percent probability.
- CIs do not provide estimates of expected error introduced by truncation or non-disclosure-avoidance error (such as coverage error).
- CIs are produced for both the counts and the ratios.
- Importantly, in practice, the SPP approach naturally provides principled measures of accuracy as a byproduct of the procedure.
- Note, “noise” is applied to the numerators and denominators separately.
- Without SPP, measures of accuracy for the ratios would be significantly more difficult.

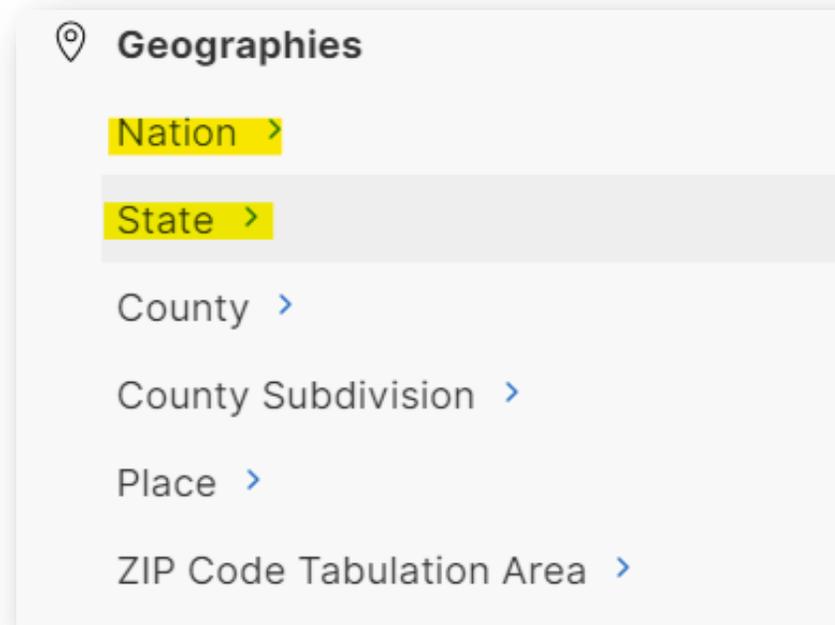
ACCESSING AND EXPLORING THE DATA

DATA RESOURCES

- [Data.census.gov](https://data.census.gov)
- Census Bureau API
- No summary file because of the limited number of tables and geographies.

TIPS FOR ACCESSING ON DATA.CENSUS.GOV

- Select geography.
- Optional: Use the Topics filter to narrow down the tables.



TUTORIAL RESOURCES

- How-to guides and short video tutorials on accessing the data (*available with embargo*)
- Webinar post-release

Navigating data.census.gov for 2020 S-DHC Data

Tuesday, September 24, 2024

2-3 p.m. EDT

<https://www.census.gov/data/academy/webinars/2024/data-census-gov-for-2020-s-dhc-data.html>

Technical documentation also provides tips

TECHNICAL RESOURCES

- 2020 Census Data Product Planning Crosswalk - Table shells are available on the “SDHC Tables” tab.
- Technical documentation (*embargo/release day*)
- Brief: Disclosure Avoidance and the Supplemental Demographic and Housing Characteristics File (S-DHC): How PHSafe Works (*release day*)

CONTENT RESOURCES

- Blog: [Understanding the Supplemental DHC File \(census.gov\)](#) (*available today*)
- Tip sheet (*embargo/release day*)
- America Counts story on coupled households (*embargo/release day*)

RECAP: TIMING AND LOGISTICS

Media Embargo – Tuesday, September 17, 2024

- Begins at 10 a.m. EDT

Public Release – Thursday, September 19, 2024

- Release at 12:01 a.m. EDT
- Tables live on data.census.gov and API by 10 a.m. EDT

Click the [Embargo Subscribers](#) link on the press kit to request embargo access.

Q&A

- Subject matter experts:
 - S-DHC product: Alexandra Krause
 - Tenure: Caroline Short
 - Families and households: Rose Kreider

For Further Information

Media

Public Information Office

pio@census.gov

301-763-3030 or

877-861-2010 (U.S. and Canada only)

Non-Media

Customer Service Center

1-800-923-8282 or 301-763-INFO (4636)

DATA DISSEMINATION AND TRAINING

Bring our data experts to you.

Request a free data training for your organization.

Receive our Data Gems.

These short “how-to” videos are an easy and quick way to increase your knowledge of census data. Get them in your inbox!

Get access to our Courses.

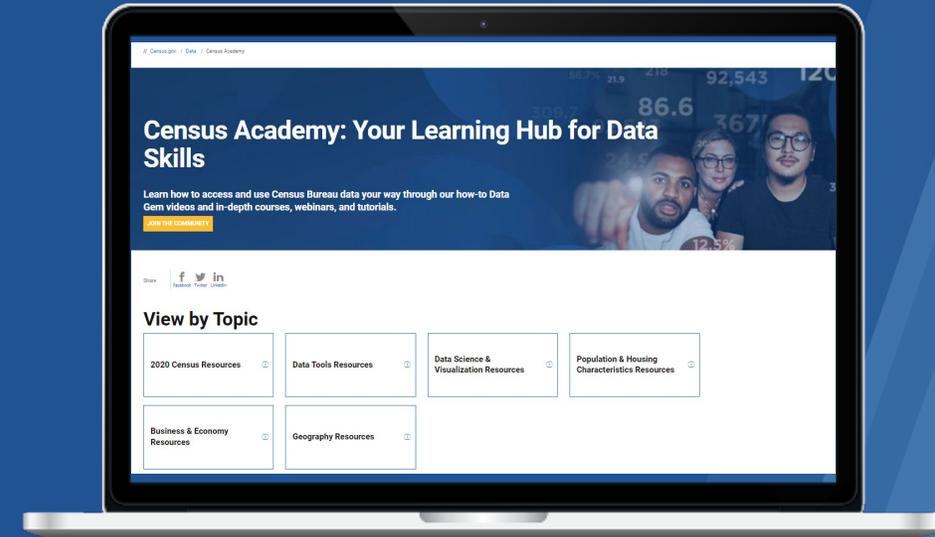
You will learn at your own pace with these video-tutorials designed for different skill levels.

Interact with our instructors via Webinars.

Learn about our data releases and tools while attending these live virtual classes.



[Census.gov/Academy](https://www.census.gov/academy)



Contact us:

Census.askdata@census.gov

1-844-ASK-DATA

Share your ideas and feedback at
census.academy@census.gov

LOOKING AHEAD: 2030 CENSUS DATA PRODUCTS

In the coming months, we'll share plans for:

- Disclosure avoidance research
- Data product planning
- Public engagement

FOR FURTHER INFORMATION

Media

Public Information Office

pio@census.gov

301-763-3030 or

877-861-2010 (U.S. and Canada only)

Non-Media

Customer Service Center

1-800-923-8282 or 301-763-INFO (4636)

Press Kit

<https://www.census.gov/newsroom/press-kits/2024/2020-census-supplemental-dhc.html>

Source Us

U.S. Census Bureau, 2020 Census Supplemental Demographic and Housing Characteristics File



THANK YOU



PIO@CENSUS.GOV



CENSUS.GOV

