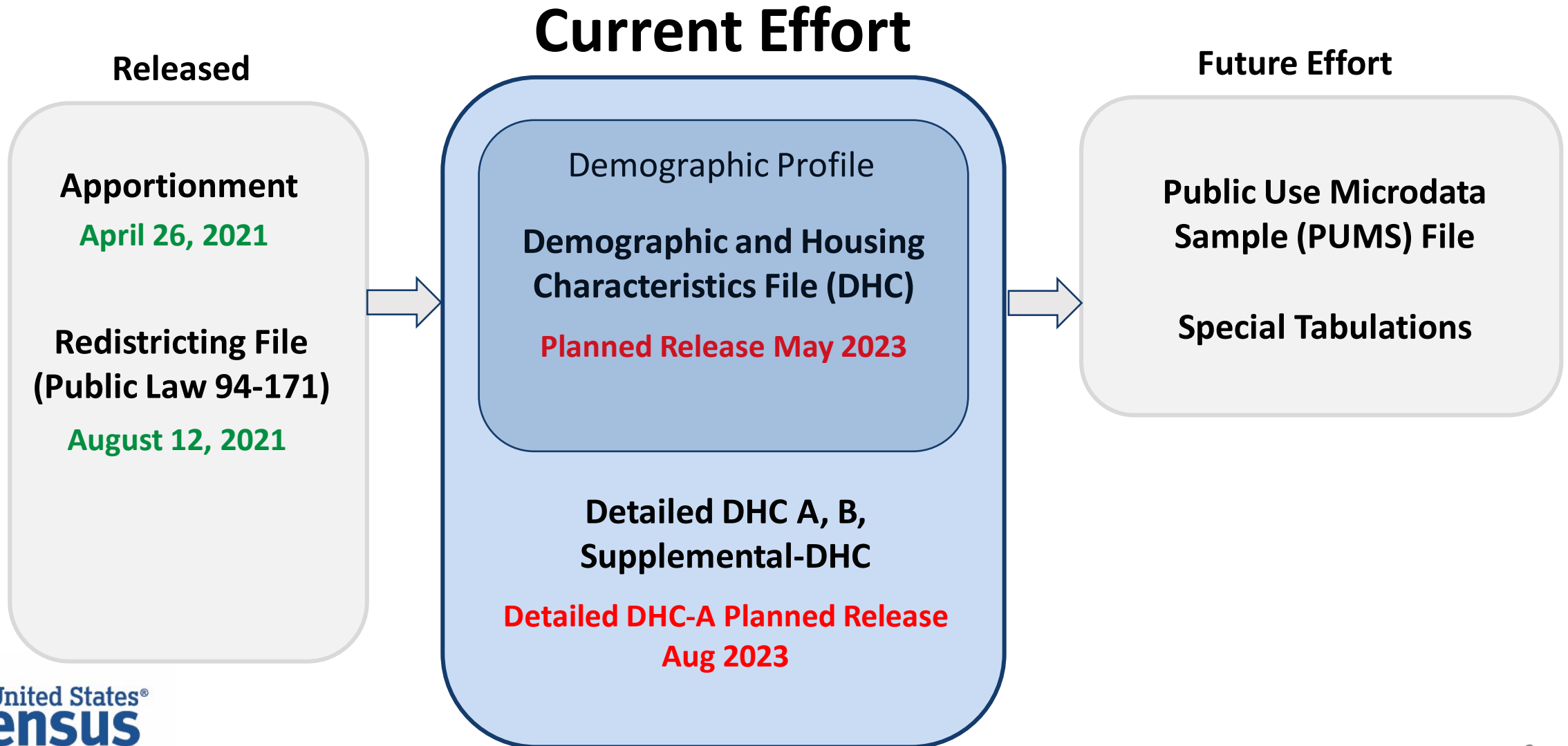


Turning Feedback Into Action: Preparing for the Second Round Demographic and Housing Characteristics File (DHC) Demonstration Data

Jason Devine
Population Division
Census Bureau

Committee on National Statistics
Public Workshop on the Demographic and Housing Characteristics Files
June 21-22, 2022

2020 Census Data Products



What is the Demographic and Housing Characteristics File (DHC)?

Includes many of the demographic and housing tables previously included in the 2010 Census Summary File 1 with many tables provided at the Census block level

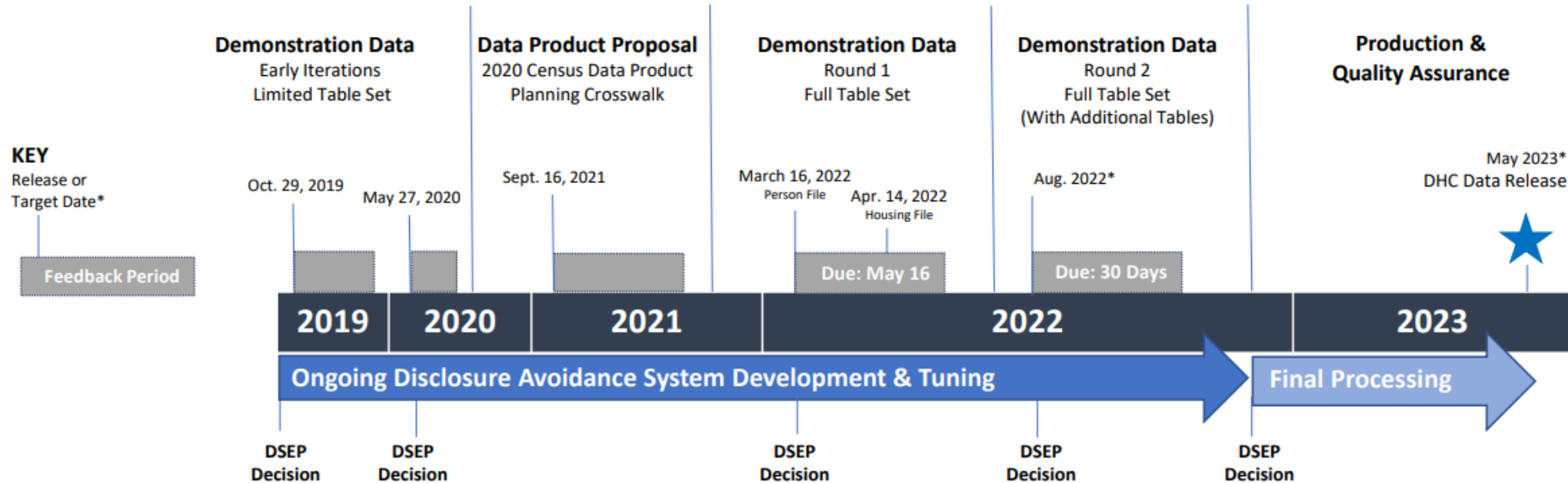
Includes tabulations and cross tabulations of:

- People by age, sex, race, Hispanic origin, and relationship to the householder
- Counts of households and families by type
- Counts of households by the presence of people of various ages
- Counts of households by characteristics of the householder
- Counts of housing by tenure (owner, renter occupied) and vacancy status
- Counts of people living in Group Quarters by characteristics

Not included are tables that provide:

- Counts of people by detailed race, ethnicity, and American Indian and Alaska Native (AIAN) tribes and villages – these will be included in the Detailed DHC A and Detailed DHC B
- Counts of people by household or family types (tables that require complex person-household joins) – these will be included in the Supplemental DHC

Demographic and Housing Characteristics File (DHC) Development and Production Timeline



Demonstration Data

The demonstration data apply the latest version of the disclosure avoidance system (DAS) to 2010 Census data. This allows a side-by-side comparison of DAS impact, quantified by detailed summary metrics. All demonstration data include the person and housing files, but Round 1 split these files into two releases.

DSEP Decision Points

The Data Stewardship Executive Policy Committee (DSEP), composed exclusively of senior career officials, determines the allocation of privacy-loss budget across geography types and queries based on an analysis of data quality and disclosure risk. DSEP makes these determinations prior to every demonstration data release and prior to final release.

Feedback Periods and Due Dates

Data users have at least 30 days to evaluate and submit data feedback to help inform further development. Upcoming feedback deadlines are marked.

Read our blog to learn more:
["2020 Census Data Products: Next Steps for Data Releases"](#)

Where We Are

- The redistricting product has been released and was protected using a differentially private disclosure avoidance system (DAS)
- Concerns from data users regarding the released redistricting data have been mostly about the accuracy of the Census counts and the perception that this accuracy was affected by disclosure avoidance protections
- We are now focused on the DAS to support the Demographic and Housing Characteristics file (DHC) and are following a similar strategy as with the redistricting product by releasing demonstration data
- One additional round of DHC demonstration data is planned, along with a production settings product
- The comment period for the first round DHC demonstration data closed on May 16
- We plan to release a second round of demonstration data for comment in August that will reflect close-to-final DAS settings
- DAS development efforts are now focused on making improvements based on internal analysis and external feedback

Feedback Periods

Six Distinct Comment Periods

1. **2018/2019 Federal Register Notice** – Focused on initial design of the 2020 Census data products and alerted the public to the use of a new disclosure avoidance methodology
2. **2019 October Demonstration Data Product** – Provided demonstration data for an initial assessment of the new disclosure avoidance methodology – results were the focus of the first CNSTAT workshop and were generally considered unacceptable
3. **Redistricting Demonstration Data Products** – May 2020 through June 2021 – Focused on obtaining feedback on the accuracy of the Redistricting Product
4. **2021 September - December Feedback Period on Proposed Tables** – Focused on collecting use cases to inform decisions about the contents of the DHC and Detailed DHC products
5. **2022 March – May DHC First Round Demonstration Product** – Provided first full set of DHC tables based on protected 2010 Census data
6. **2022 DHC Second Round Demonstration Product (Planned August Release)** – Will provide results from the close-to-final version of the Disclosure Avoidance System to support the DHC

Ways We Responded to Feedback on the Redistricting Data

- Feedback on the Redistricting Demonstration Data led to the allocation of a substantial amount of additional privacy loss budget
- Concerns about lower accuracy of data for tribal geographies and other off-spine geographies such as places and Minor Civil Divisions (MCDs) led to changes to the 'optimized spine'
- Several measures in the summary metrics revealed biases in the results that were improved through changes to the post-processing system parameters
- A need for more accuracy in race and ethnicity statistics at many levels of geography led to the allocation of additional privacy loss budget to those queries
- A need for more accurate statistics on occupancy rates led to the allocation of additional privacy loss budget to the housing unit data

Ways We Responded to General Feedback

- Increased our outreach through workshops, webinars, blogs, newsletters, and presentations
- Provided Privacy Protected Microdata files in addition to the Detailed Summary Metrics
- Added tables back into the DHC
- Included two feedback periods for the development of the DAS that will support the DHC
- Produced a Differential Privacy Handbook
- Split the Detailed DHC into Detailed DHC A, B, and the Supplemental DHC

Feedback on the Contents and Design of the DHC

- Feedback received in response to the 2018/2019 Federal Register Notice was considered in the initial design of the 2020 Census data products
- Feedback received on the 2021 September through December feedback period on proposed tables was used to finalize the proposed contents of the DHC
- Over 400 comments on the contents of the DHC and Detailed DHC, many of which included use cases, were received throughout the September through December 2021 comment period
- Most uses that require basic demographic and housing information were already being met by the tables initially proposed for the DHC
- Based on data user needs provided in the comments, many tables are being added back to the DHC by making tables proposed for the county level or above available at the tract level
- Based on data user feedback, no tables were removed from the DHC or moved to a higher geographic level

First Round DHC Demonstration Data Feedback

- Data were provided as person and unit summary files along with detailed summary metrics – the review period extended 30 days for the re-release of the unit files and ended May 16
- 14 comments were received in response to the call for feedback on the accuracy of the first round DHC demonstration data
- 11 of those comments provided some sort of assessment of accuracy with a subset of the 11 providing more detailed analysis
- Most of the comments raised concerns about the accuracy of the data by age and sex with a specific concern for the accuracy of counts of children
- The accuracy of age data for children for school districts was a specific concern of one commenter
- There is an interest in greater accuracy for tables by household type, household size, and tenure, both alone and crossed by race and ethnicity
- Commenters expressed concern about the level of inconsistencies between the unit and person files
- Two commenters requested tables that are already planned to be added back to the DHC

In What Ways Have We Been Successful?

- The allocation of additional privacy loss budget, infusion of greater block-level noise, and improvements to the DAS resulted in substantial increases in accuracy from the first demonstration data release in October of 2019
- We have released multiple demonstration products
- We have had successive comment periods that have benefited from an improved understanding of the new disclosure avoidance methodology by Census staff and data users
- We have held multiple workshops and webinars and have participated at conferences to help inform stakeholders
- We have responded to data user requests to add tables back into the DHC
- We have been able to make improvements to the DAS used for the first round demonstration data
- We provided a thorough summary of comments from the first round demonstration data comment period

In What Ways Have We Been Less Successful?

- Despite outreach efforts, more is needed
- We have released multiple demonstration products, but they have been difficult to use
- We had to split the unit- and person- sourced summary files in the first round DHC demonstration data and then identified an issue in the unit-sourced summary files
- We need to make more information from the late 2021 comment period available
- We have had difficulty using use cases to establish acceptable levels of accuracy for the DHC data
- We have expected data users to accept inconsistencies within the data without enough explanation of why they are there and why they cannot be eliminated
- We have been able to make improvements in accuracy for some of the DHC tables but some areas of lower accuracy likely will not be improved
- We have not been able to clearly establish how much privacy loss is too much

What Is Next

- We will listen to the presenters and discussion at this workshop and use what we hear to inform decisions about improvements for the second round DHC demonstration data
- In August, the second round DHC demonstration data will be released with a 30-day comment period
- Comments from the second round comment period will be reviewed and summarized and used to inform decisions about final improvements to the DAS before final settings are determined
- The Data Stewardship Executive Policy Committee (DSEP) will approve the final settings for the DAS and production of the DHC files will begin
- This effort will conclude with the release of the DHC scheduled for May 2023

Thank You!!