

DAS Updates

Michael Hawes

Senior Advisor for Data Access and Privacy
Research and Methodology Directorate
U.S. Census Bureau

CNSTAT Expert Meeting on Disclosure Avoidance
June 1, 2020

Any opinions and viewpoints expressed in this presentation are the author's own, and do not represent the opinions or viewpoints of the U.S. Census Bureau.

**Shape
your future
START HERE >**

United States[®]
**Census
2020**

Agile Development Process

Recurring six week cycle of:

Four week development “sprint”

- 2-3 Executive Priorities
- Defined development and research tasks for each team member
- Clearly established completion criteria

Two week “inter-sprint” period

- Evaluation of completed sprint
- Establishing executive priorities for next sprint
- Identifying tasks, LOE, and completion criteria for next sprint

Spring 2020 Sprint Schedule

Sprint II (March 2 – March 31)

Sprint III (April 15 – May 13)

Sprint IV (May 21 – June 17)

...development sprints will continue throughout the remainder of 2020

Sprint II (March 2020)

Executive Priorities

1. Implement multi-stage post-processing to improve population totals for legal and political entities.

Modify the TopDown Algorithm to process tabulations for each geographic level in the following sequential order:

- Total population
- Tabulations necessary for the PL94-171 file
- Tabulations necessary to support the Population Estimates program and most demographic use cases
- The remaining tabulations necessary to produce the Demographic and Housing Characteristics file and Demographic Profiles

Each pass of the algorithm will be constrained to the values established in the prior passes.

2. Branch this version of the algorithm as “Decennial Production Version 1”

3. Identify and test additional approaches to mitigating post-processing error

Sprint III (April-May 2020)

Executive Priorities

1. Produce a full national DAS run using the new multi-stage version of TDA
2. Continue testing and implementing accuracy improvements for DHC-P and DHC-H, prioritizing:
 - Total population for legal and political entities
 - Core Population Estimates queries (3M)
 - DHC-H occupancy/tenure and household size
 - PL94-171 redistricting data
 - Balance of DHC-P and DHC-H

Sprint IV (May-June 2020)

Executive Priorities

- **Preparation for Test Readiness Review (TRR)**
 - Produce MDF_Per and MDF_Unit, with ability to have variants of multipass
 - Ability to manipulate geographic spine
 - Focus on readers/writers.
 - Documentation and QA necessary for TRR
- **Evaluate control of privacy-loss budget over wide range to assess accuracy**

Ongoing Data User Engagement

CNSTAT Expert Meetings and Small Groups

- Priorities and Tradeoffs
- Accuracy Metrics
- Geographic Hierarchy
- Post-processing
- Supporting data use cases with the “noisy measurements”

Advisory Committees

- National Advisory Committee Working Group
- Census Scientific Advisory Committee Working Group

FSCPE Working Group

Department of Justice

AIAN Tribal Listening Sessions

Feedback Requested

Evaluate the Census Bureau's collection of use cases:

- What are we missing?
- *Especially important for uses of Group II data products (detailed race, ethnicity, tribal data and person-household joins)*

Prioritize Use Cases

- For Group I data products: Input on how PLB should be allocated across data products and tabulations
- For Group II data products: Input on necessary geographic and variable detail for Group II data products

Evaluate and Review the Census Bureau's Detailed Summary Metrics

Suggestions for Communications Strategy and Materials about DP

Questions?