The Frames Program

Lori B. Zehr, PhD

August 7, 2023

Presented at the 2023 Joint Statistical Meetings

Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau.

What You Will Learn Today

- What is the Frames Program?

- What are the benefits of the Frames Program?

- What is the current state of the Enterprise Frames?

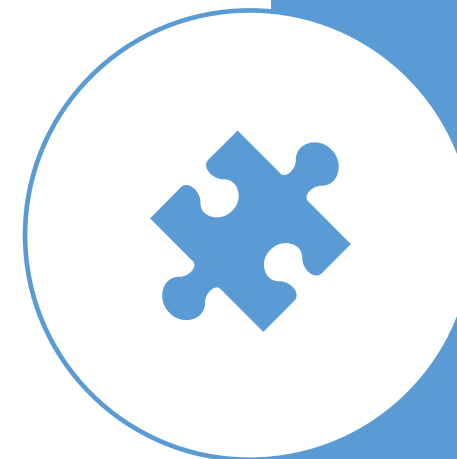- Where are we headed with the Frames Program?

What is a Frame?

A frame refers to a collection of datasets organized around one of four foundational units: Person, Business, Location, Job.

A frame is supported by a collection of standard methods, code, processes, documentation, and subject matter expertise.
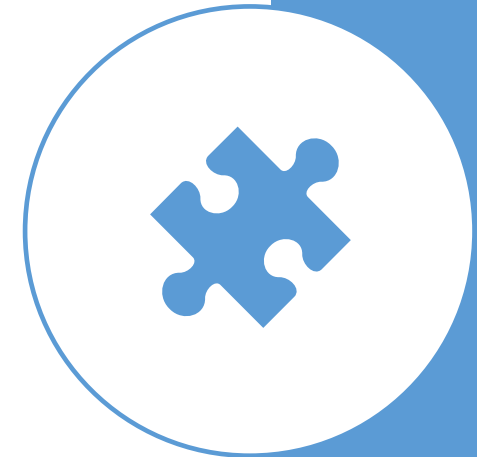
A sampling frame often refers to a subset of the larger foundational dataset.

United States® Census Bureau

Frames Program
The Challenge

- Various frame-like datasets are in use at the Census Bureau.

- Contain information about addresses, businesses, jobs, and people.

- Some information is unique within a particular frame, while other information is duplicative.

  - e.g., group quarters exist in both the Master Address File and Business Register.

- Frames exist in an uncoordinated and unintegrated environment.

- No systematic process exists that allows for the direct linkage of information contained in one frame with information in any other frame.

Frames Program
The Vision

- *To create Enterprise-wide frames linkable in nature, agile in structure, accessible for production or research on a need-to-know basis, and that adhere to best practices in terms of technology usage, data management, and methodology.*

- Linkable in Nature

  - Each frame will include the necessary unique identifiers and keys for linkage to each other.

    - Location information will be available on each person, job, and business record.

    - Person records will be linkable to job records.

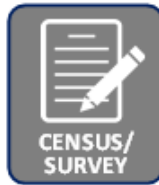   - Job records will be linkable to business records.

United States® Census Bureau

Frames Program

Creating an Infrastructure to Modernize the Census Bureau's Statistical Foundation

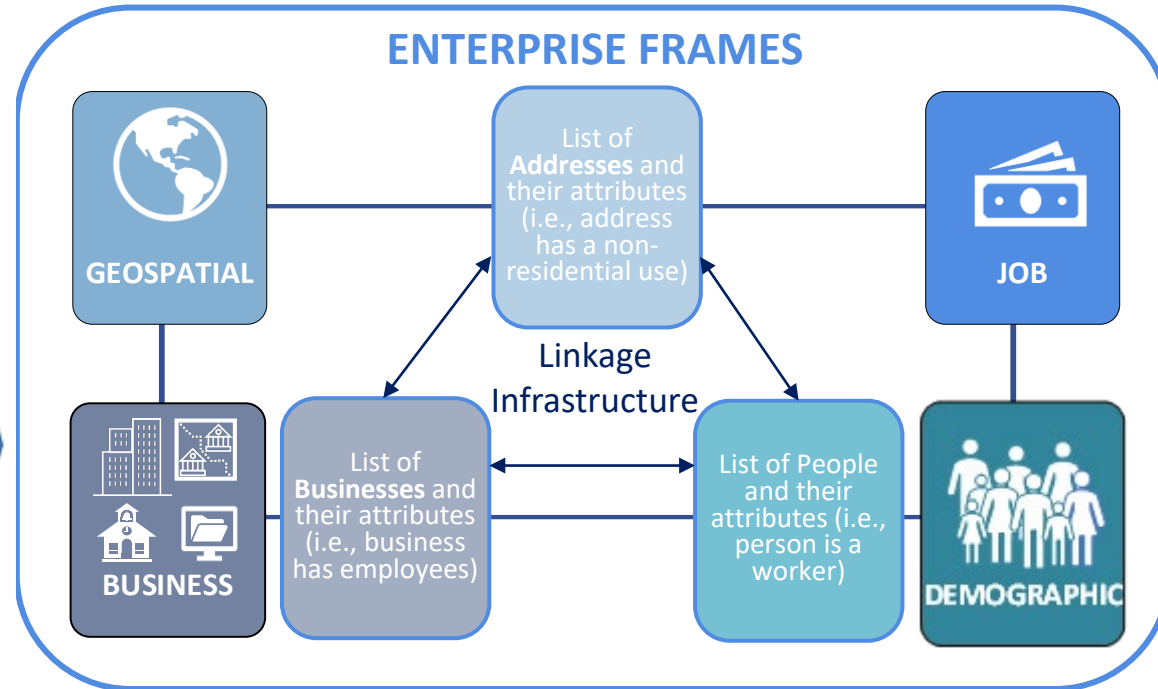**Data Input Categories**

- ADMIN RECORDS
- CENSUS/SURVEY
- PUBLIC RECORDS
- THIRD PARTY

Data Ingest and Collection

ENTERPRISE FRAMES

- GEOSPATIAL
- List of **Addresses** and their attributes (i.e., address has a non-residential use)
- JOB
- BUSINESS
- List of **Businesses** and their attributes (i.e., business has employees)
- Linkage Infrastructure
- List of People and their attributes (i.e., person is a worker)
- DEMOGRAPHIC

PRODUCTS/SERVICES

Programmatic and Research Activities

- ENTERPRISE INITIATIVE SUPPORT
- SAMPLING FRAMES
- QUALITY IMPROVEMENT RESEARCH
- DATA PRODUCTS
- DATA EXTRACTS

United States® Census Bureau

6

Geospatial Frame

Current State

- National address list of 200M+ addresses prioritizing residential addresses.

- Linked to the Topologically Integrated Geographic Encoding and Referencing (TIGER) System.

- Serves as the source of address and location information for the Decennial Census, American Community Survey, and other demographic surveys.

- Updated using information from the U.S. Postal Service; tribal, state, and local governments; and Census data collection operations.

Future State

- Stores a comprehensive set of residential, **business**, and **mixed-use** addresses.

- Provides "Master Address File IDs" (MAFIDs) and geocodes (e.g., latitude and longitude) for all addresses in the frame.

- Can be accessed with to-be-developed Address Matching Tools to link new sources of address information to the Geospatial Frame.

- Linkable to other Enterprise Frames.

Business Frame

Current State

- National list of 30M+ businesses.

- Serves as the enumeration list for the Economic Census and the source of sampling for Economic Surveys.

- Updated using data from the Internal Revenue Service (IRS) tax records; the Bureau of Labor Statistics; the Social Security Administration; the Economic Census; and some Census surveys.

- Updated using geographic information from the Census Bureau's Master Address File.

Future State

- Collection of rich, harmonized business data stored in a central location.

- Incorporates the Longitudinal Business Database, resulting in consistent methodology.

- Connects the Governments Master Address File, ensuring improved quality for the Government sector.

- Linkable to other Enterprise Frames.

Job Frame

Current State

- An inventory of 14+ billion "job records" as part of the Longitudinal Employer-Household Dynamics Program (LEHD).

- Each "job record" is a relationship between a person and a firm in which at least $1 of wages were paid in a given quarter.

- Covers >97% of private employment and most federal, state, and local government jobs.

- Supports public-use data products.

- Additionally, the LEHD Program maintains a business frame that is independent of the BR, known as the Employer Characteristics File (ECF).

Future State

- Expands on the job frame that was built as part of the LEHD.

- Contains two linked databases of jobs:
    - Quarterly Job Frame, derived from quarterly unemployment insurance wage records supplied by states
    - Annual Job Frame, derived from W-2 tax records

- Adds self-employment jobs.

- Focuses explicitly on job records.

- Linkable to other Enterprise Frames.

Demographic Frame

Current State

- Several disparate person frames exist within the Census Bureau.
- Each was developed to serve a specific purpose:
  - Evaluate administrative data coverage
  - Support 2020 Decennial Census

Future State

- Consists of demographic, social, and economic characteristics of individuals derived from census, survey, administrative, and third-party data sources.
- Includes unique person-level identifiers used to link individuals across datasets.
- Links information about individuals across time and space.
- Can be aggregated to family and household units.
- Linkable to other Enterprise Frames.

Frames Program
Benefits

- Fosters an environment to do more work virtually.

- Reduce burden on respondents (individuals and businesses) by re-using data that already exist in one or more frames.

- Reduce duplication, manual efforts, and one-off data cleaning projects.

- Allows more time for production and research.

- Allows for ability to more easily trace data and code through the linked frames, facilitating reproducibility.

- Expands the Census Bureau's ability to report on changes and trends affecting the Nation's population, economy, and communities.

For More Information

lori.b.zehr@census.gov