Westat

Improving Lives Through Research®

Federal Statistics: Building on the Past, Looking to the Future

Jeri Metzger Mulrow
Vice President, Westat

WESTAT @ FCSM 2024

Agenda

- A bit about me
- Federal Statistics – Important, Invisible, Innovative
- Building on the Past
- Looking to the Future
- Challenges

Federal Committee on Statistical Methodology

2024 Research and Policy Conference

The Relevance, Timeliness, and Integrity of Federal Statistics

October 22-24, 2024
College Park Marriott Hotel and Conference Center
Hyattsville, MD

A bit about me – A Statistician at Heart

Education

BS Mathematics, Montana State University

MS Statistics, Colorado State University

Background

3 Federal Statistical Agencies

Federal Committee on Statistical Methodology

Bureau of Justice Statistics
National Center for Science and Engineering Statistics
Statistics of Income Division

\+ NIST

+ **Westat**, EY, NORC

+ **2026** ASA President

Federal Statistical System – A Large Community

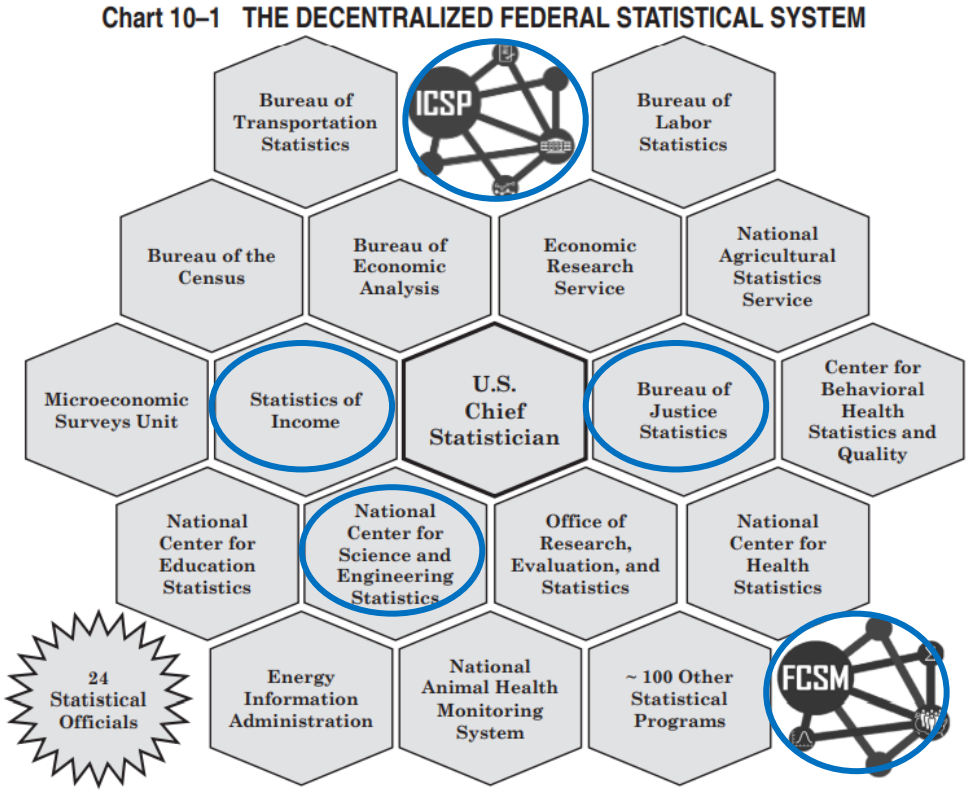
Information Providers

Information Users

Parent Agencies

Statistical Contractors

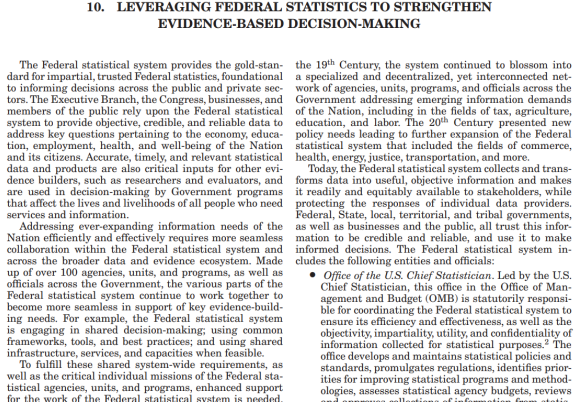
Professional Associations



Analytical Perspectives, Budget of the United States GovernmentFiscal Year 2025: Leveraging Federal Statistics to Strengthen Evidence-Based Decision-Making

Federal Statistics – Important "Gold-Standard"

Analytic Perspectives, 2025



10. LEVERAGING FEDERAL STATISTICS TO STRENGTHEN EVIDENCE-BASED DECISION-MAKING

The Federal statistical system provides the gold-standard for impartial, trusted Federal statistics, foundational to informing decisions across the public and private sectors. The Executive Branch, the Congress, businesses, and members of the public rely upon the Federal statistical system to provide objective, credible, and reliable data to address key questions pertaining to the economy, education, employment, health, and well-being of the Nation and its citizens. Accurate, timely, and relevant statistical data and products are also critical inputs for other evidence builders, such as researchers and evaluators, and are used in decision-making by Government programs that affect the lives and livelihoods of all people who need services and information.

Addressing ever-expanding information needs of the Nation efficiently and effectively requires more seamless collaboration within the Federal statistical system and across the broader data and evidence ecosystem. Made up of over 100 agencies, units, and programs, as well as officials across the Government, the various parts of the Federal statistical system continue to work together to become more seamless in support of key evidence-building needs. For example, the Federal statistical system is engaging in shared decision-making; using common frameworks, tools, and best practices; and using shared infrastructure, services, and capacities when feasible.

To fulfill these shared system-wide requirements, as well as the critical individual missions of the Federal statistical agencies, units, and programs, enhanced support for the work of the Federal statistical system is needed.

the 19th Century, the system continued to blossom into a specialized and decentralized, yet interconnected network of agencies, units, programs, and officials across the Government addressing emerging information demands of the Nation, including in the fields of tax, agriculture, education, and labor. The 20th Century presented new policy needs leading to further expansion of the Federal statistical system that included the fields of commerce, health, energy, justice, transportation, and more.

Today, the Federal statistical system collects and transforms data into useful, objective information and makes it readily and equitably available to stakeholders, while protecting the responses of individual data providers. Federal, State, local, territorial, and tribal governments, as well as businesses and the public, all trust this information to be credible and reliable, and use it to make informed decisions. The Federal statistical system includes the following entities and officials:

- *Office of the U.S. Chief Statistician.* Led by the U.S. Chief Statistician, this office in the Office of Management and Budget (OMB) is statutorily responsible for coordinating the Federal statistical system to ensure its efficiency and effectiveness, as well as the objectivity, impartiality, utility, and confidentiality of information collected for statistical purposes.[2] The office develops and maintains statistical policies and standards, promulgates regulations, identifies priorities for improving statistical programs and methodologies, assesses statistical agency budgets, reviews

"The Federal statistical system provides the gold-standard for impartial, trusted Federal statistics, foundational to informing decisions across the public and private sectors. .... Accurate, timely, and relevant statistical data and products are also critical inputs for other evidence builders, such as researchers and evaluators, and are used in decision-making by Government programs that affect the lives and livelihoods of all people who need services and information."

ap_10_statistics_fy2025.pdf (whitehouse.gov)

Federal Statistics – Important Social Indicators

U.S. Social Indicators

98 ways to measure the state of the Nation

24 Economic Indicators

GDP, CPI, employment, corn for grain production, R&D spending, etc.

9 Demographic Indicators

Total population, Foreign born population, Ave. family size, etc.

19 Socioeconomic Indicators

Math & Reading achievement scores, Poverty rate, Real median income, Food insecure HHs, etc.

17 Health Indicators

Life expectancy, Infant mortality, Disability, Cigarette smokers, Vaccination coverage, etc.

14 Safety & Civic

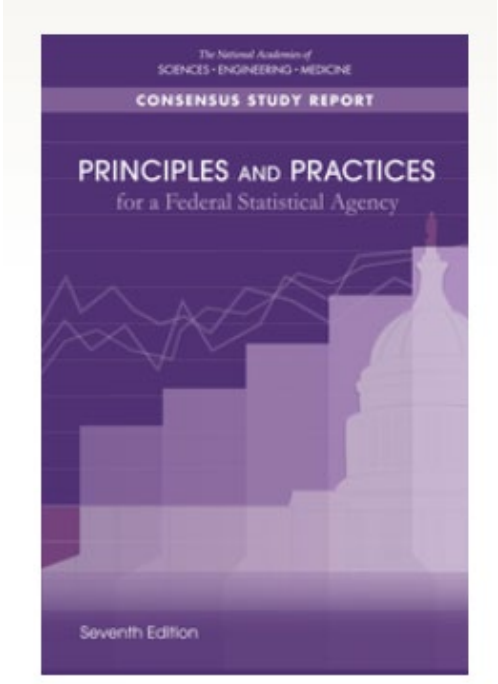
Property & violent crimes rates, Seat belt use, Highway fatalities, Veterans, Military personnel on active duty, etc.

15 Environment & Energy Indicators

Ground level ozone, Net greenhouse gas emissions, energy consumption, coal production, etc.

https://www.performance.gov/explore/social-indicators/

Federal Statistics – Important Indispensable Role



THE VALUE OF NATIONAL STATISTICS IN THE UNITED STATES

The United Nations General Assembly in 2014 formally endorsed *Fundamental Principles of Official Statistics* (see Appendix C).[1] The first of these principles accords worldwide recognition to the indispensable role of official statistics:

> Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information (United Nations Statistical Commission, 2014, March 3, p. 1-2).

https://www.nationalacademies.org/our-work/7th-edition-of-principles-and-practices-for-a-federal-statistical-agency

Federal Statistics – Invisible to the Public

Q: Where did those statistics come from?

A: The internet!

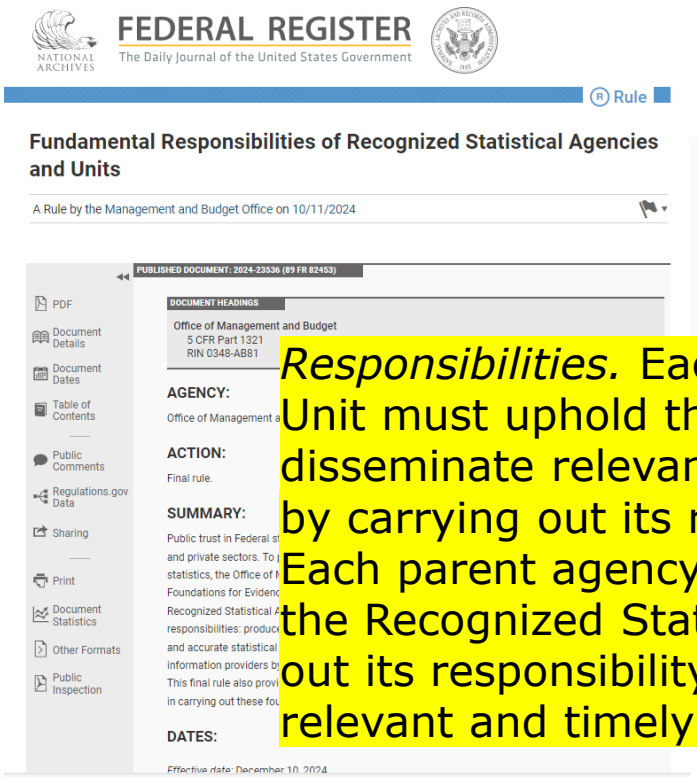
Q: Where did you get that information?

A: Friends on social media

Q: Where did those data come from?

A: Department of Commerce

Federal Statistics – Invisible to Parent Agencies


FEDERAL REGISTER
The Daily Journal of the United States Government

Rule

Fundamental Responsibilities of Recognized Statistical Agencies and Units

A Rule by the Management and Budget Office on 10/11/2024

PUBLISHED DOCUMENT: 2024-23536 (89 FR 82453)

DOCUMENT HEADINGS

Office of Management and Budget
5 CFR Part 1321
RIN 0348-AB81

AGENCY:

ACTION:

Final rule.

SUMMARY:

DATES:

Just Released 10/11/24: Trust Regulation

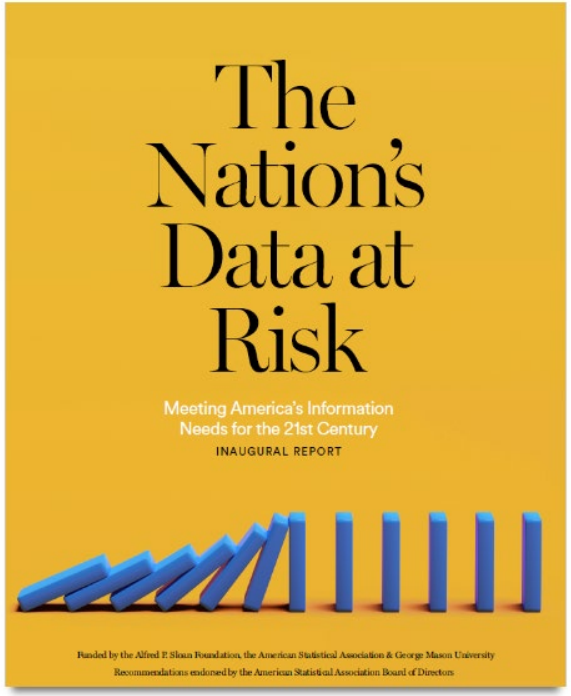
PART 1321—RESPONSIBILITIES OF RECOGNIZED STATISTICAL AGENCIES AND UNITS

1321.1 Purpose.
1321.2 Definitions.
1321.3 General provisions.

Responsibilities. Each Recognized Statistical Agency or Unit must uphold the responsibility to produce and disseminate relevant and timely statistical information by carrying out its requirements under this section. Each parent agency must enable, support, and facilitate the Recognized Statistical Agency or Unit in carrying out its responsibility to produce and disseminate relevant and timely statistical information.

Federal Register :: Fundamental Responsibilities of Recognized Statistical Agencies and Units

Federal Statistics – Invisible to Congress



2. THREATS TO FEDERAL STATISTICS AND AGENCIES

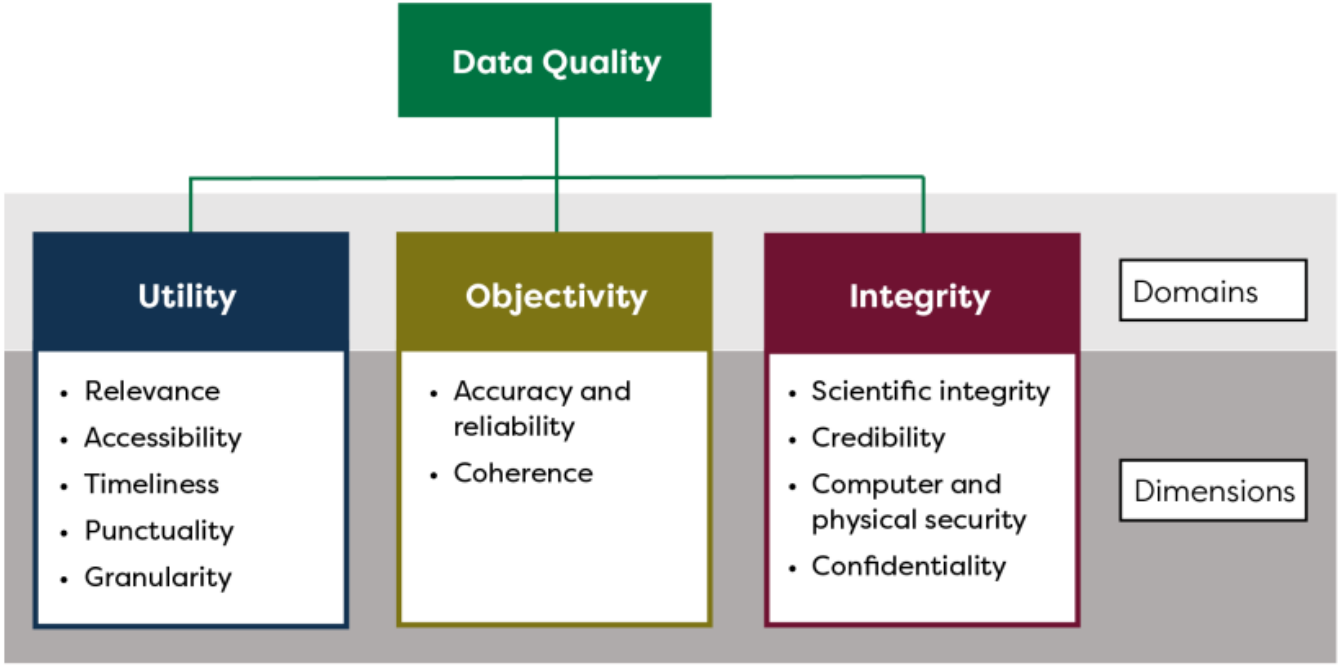
We describe four broad categories of threats that affect the federal statistical agencies' ability to continue to produce high-quality, relevant, timely, and objective data for policy and public use into the 21st century.

Neglect. The first threat is neglect by Congress and the executive branch. Such neglect may be unintentional but, regardless, the consequences can be dire. For example, funding for most of the principal statistical agencies has declined 14% in purchasing power over the past 15 years (see Figure 1 for an illustration for three agencies). In contrast, federal discretionary, nondefense spending, accounting for inflation, has increased 16% (Pierson et al., 2024—tab 10). Moreover, the amount of funding that individual agencies receive owes more to historical happenstance than careful coordination and prioritization across the statistical system. For example, the budget for BTS, despite the importance of the transportation sector in the economy, has been limited since its founding in 1992 to the originally dedicated small amount of Highway Trust Funds supplemented by transfers from other Department of Transportation agencies.

" …funding for most of the principal statistical agencies has declined 14% in purchasing power over the last 15 years… In contrast, federal discretionary, nondefense spending, accounting for inflation, has increased 16% (Pierson et al, 2024)

https://www.amstat.org/policy-and-advocacy/the-nation's-data-at-risk-meeting-american's-information-needs-for-the-21st-century

Building on the Past – FCSM Data Quality Framework



Building on the Past – FCSM DQF Definitions

Domain	Dimension	Definition
Utility	Relevance	Relevance refers to whether the data product is targeted to meet current and prospective user needs.
	Accessibility	Accessibility relates to the ease with which data users can obtain an agency's products and documentation in forms and formats that are understandable to data users.
	Timeliness	Timeliness is the length of time between the event or phenomenon the data describe and their availability.
	Punctuality	Punctuality is measured as the time lag between the actual release of the data and the planned target date for data release.
	Granularity	Granularity refers to the amount of disaggregation available for key data elements. Granularity can be expressed in units of time, level of geographic detail available, or the amount of detail available on any of a number of characteristics (*e.g.* (demographic, socio-economic).
Objectivity	Accuracy and reliability	Accuracy measures the closeness of an estimate from a data product to its true value. Reliability, a related concept, characterizes the consistency of results when the same phenomenon is measured or estimated more than once under similar conditions.
	Coherence	Coherence is defined as the ability of the data product to maintain common definitions, classification, and methodological processes, to align with external statistical standards, and to maintain consistency and comparability with other relevant data.
Integrity	Scientific integrity	Scientific integrity refers to an environment that ensures adherence to scientific standards and use of established scientific methods to produce and disseminate objective data products and one that shields these products from inappropriate political influence.
	Credibility	Credibility characterizes the confidence that users place in data products based simply on the qualifications and past performance of the data producer.
	Computer and physical security	Computer and physical security of data refers to the protection of information throughout the collection, production, analysis, and development process from unauthorized access or revision to ensure that the information is not compromised through corruption or falsification.
	Confidentiality	Confidentiality refers to a quality or condition of information as an obligation not to disclose that information to an unauthorized party.

Relevance

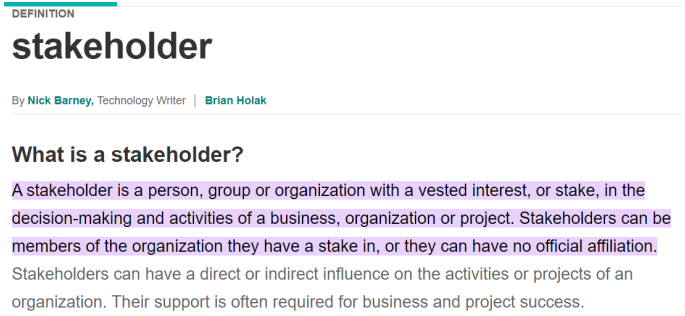
Timeliness

Integrity

Building on the Past – Relevance

Relevance

Relevance refers to whether the data product is targeted to meet current and prospective user needs.



DEFINITION

stakeholder

By **Nick Barney**, Technology Writer | **Brian Holak**

What is a stakeholder?

A stakeholder is a person, group or organization with a vested interest, or stake, in the decision-making and activities of a business, organization or project. Stakeholders can be members of the organization they have a stake in, or they can have no official affiliation. Stakeholders can have a direct or indirect influence on the activities or projects of an organization. Their support is often required for business and project success.

Federal Statistics Users:

- Data providers
- Policy/decision makers
  - Congress
  - White House
  - **Parent Agencies**
  - **State or Local governments**
- **General public**
  - Business owners
  - Individuals

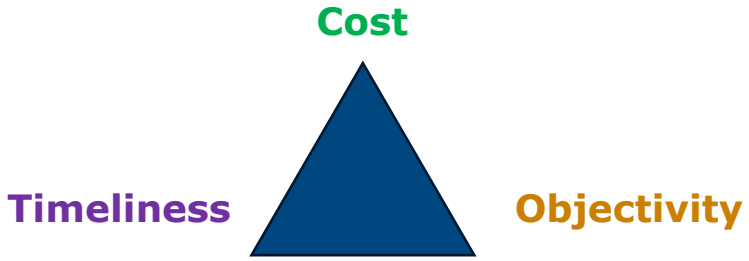
https://www.techtarget.com/searchcio/definition/stakeholder#:~:text=A%20stakeholder%20is%20a%20person,can%20have%20no%20official%20affiliation.

Building on the Past – Timeliness

Utility	Timeliness	Timeliness is the length of time between the event or phenomenon the data describe and their availability.

"It may be better, in the gross affairs of life, to be less precise and more prompt. Quick decisions, though they may contain a grain of error, are often better than precise decisions at the expense of time."
– T.C. Chamberlin, President of the University of Wisconsin, 1890

Different products need different timeliness & different balance depending upon the uses



Building on the Past – Scientific Integrity

Scientific integrity

Scientific integrity refers to an environment that ensures adherence to scientific standards and use of established scientific methods to produce and disseminate objective data products and one that shields these products from inappropriate political influence.

Society benefits from informed judgments supported by ethical statistical practice.

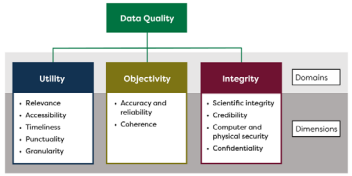
ASA Ethical Guidelines for Statistical Practice

Producers of Federal Statistics must:

- Be ethical and responsible
- Use appropriate measures & methodologies
- Ensure representativeness of population(s) of interest

https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice

Building on the Past – Threats to Data Quality



FCSM Identified Threats

- **Utility**: competing data sources, costs of access and documentation, use of disclosure protections, and delays in data acquisition and processing
- **Objectivity**: accuracy and reliability in the Total Survey Error paradigm and for integrated data products, such as, linkage error, harmonization error and modeling error,
- **Integrity**: lack of scientific integrity, political interference, and data security failures

https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice

Looking to the Future – Select Innovative Activities

- Planned innovations
- COVID-related
  - Adaptations
  - New creations

- Relevant
- Timely
- Scientific Integrity

SOI – Building a tiered access program

- **Open data products** – e.g. annual ZIP code level tax return data, county income data, US population migration data, public use files
- **Synthetic PUFs**
  - TYs 2012-2015 – supplemental for individuals likely below the tax filing threshold
  - TY 2015 – fully synthetic beta Form 1040 file
- **Joint Statistical Research Program** – estab. 2012, permits selected researchers access to federal tax microdata
- **Custom tabulations** for special studies - e.g. College Scorecard
- ***NEW* Secure Query System –** Linking standardized client data to Individual tax data to produce output products after disclosure review

A Secure Query System to Improve Access to Individual Income Tax Data Amy O'Hara, Stephanie Straus, Ron Borzekowski, Paul Arnsberger, and Barry Johnson NBER Working Paper No. 32969 September 2024 JEL No. H20, I18, I20, I23, I25, I28, I30, I31, J10, J15, J18, J19, J30, J38, J40, Y10

NCSES – Demonstrating a National Secure Data Service

- **NSDS Philosophy –** showcase the critical importance of data and produce value by facilitating evidence building.
- **NSDS as a Service -** coordination and capacity-building services for data users, data providers, and related communities of practice.
- **NSDS as a Place -** legally recognized entity with hardware, software, and administrative infrastructure and capacity to meet its mission

Data Access Alternatives: Artificial Intelligence Supported Interfaces

AI Chat Bot

Westat

Synthetic Data Generation with Large, Real-World Data

Super Computing Environment

https://ncses.nsf.gov/initiatives/national-secure-data-service-demo
ADC | America's Datahub Consortium – Accelerating data and statistical innovation (americasdatahub.org)

BJS – Redesigning a flagship survey

National Crime Victimization Survey

- 50 years (1973 – 2023)
- Annual data collection
- 240,000 individuals in 150,000 HH
- Topics: non-fatal personal crimes and property crimes, whether reported to police or not, reasons reported or not, experience with criminal just system and more

- **Redesign efforts – split sample**
  - Updated crime screening questions
  - Revised rape and sexual assault questions
  - Expanded info on victimization incidents and seeking of help
  - Non-crime questions on police performance & community safety

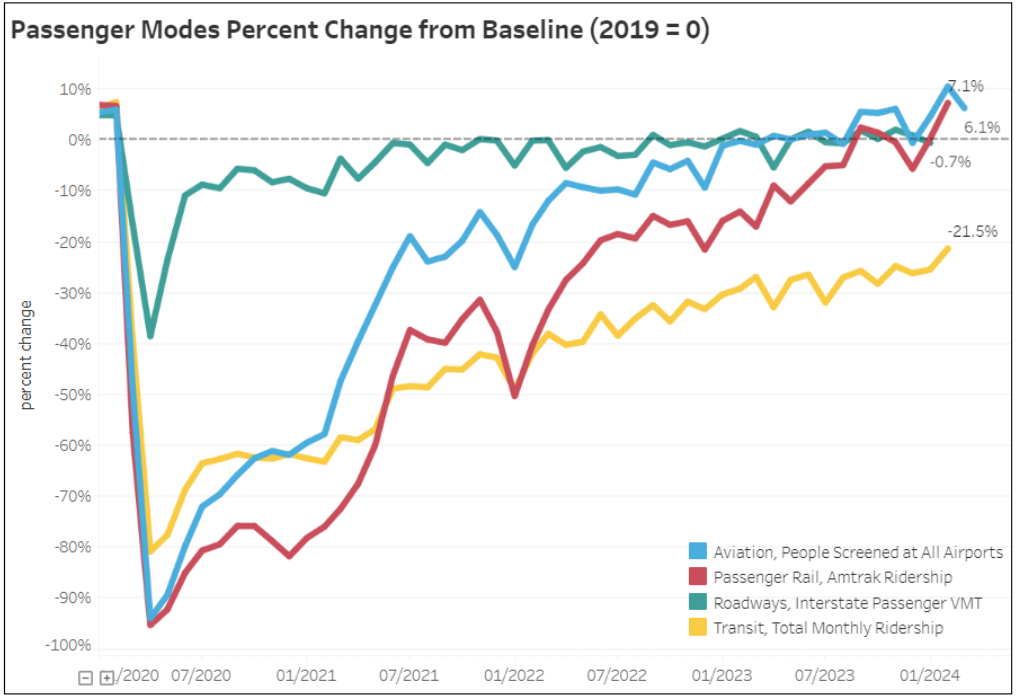
https://bjs.ojp.gov/data-collection/ncvs#8-0

BTS – Adapting to change

Bureau of Transportation Statistics

The Week in Transportation

Shifted from annual and monthly to weekly and daily statistics



Check out BTS' Commercial Airline data

The Week in Transportation | Bureau of Tr

Quick links to popular air carrier statistics | Bureau of Transportation Statistics (bts.gov)

BLS – Adapting to change

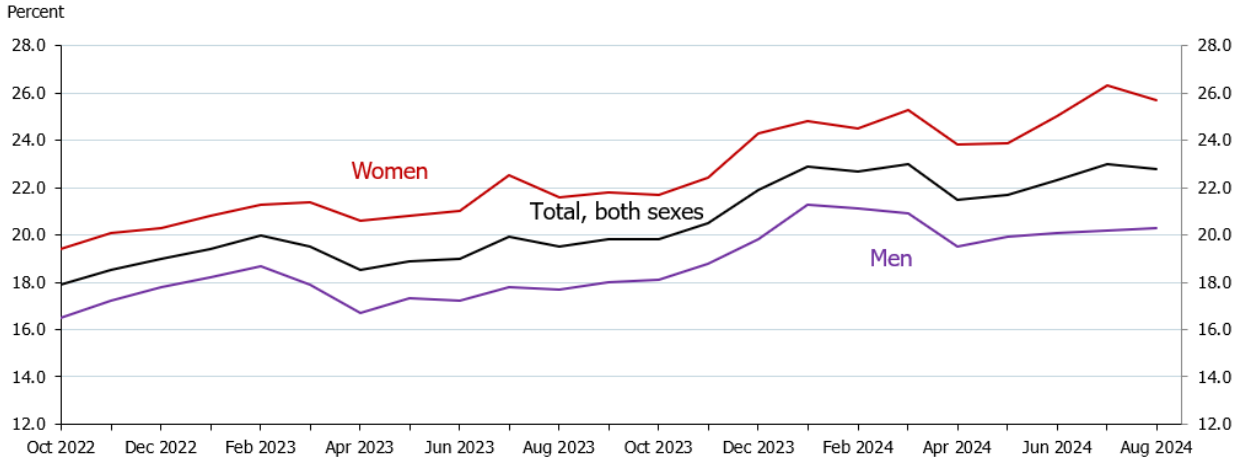
Measuring Telework - Current Population Survey

May 2020: At any time in the LAST 4 WEEKS, did you telework or work at home for pay BECAUSE OF THE CORONAVIRUS PANDEMIC?
Yes/No

October 2022 Q1: At any time LAST WEEK did you telework or work at home for pay?
Yes/No

A-13

Telework or work at home rate for pay by sex
Not seasonally adjusted, October 2022–August 2024

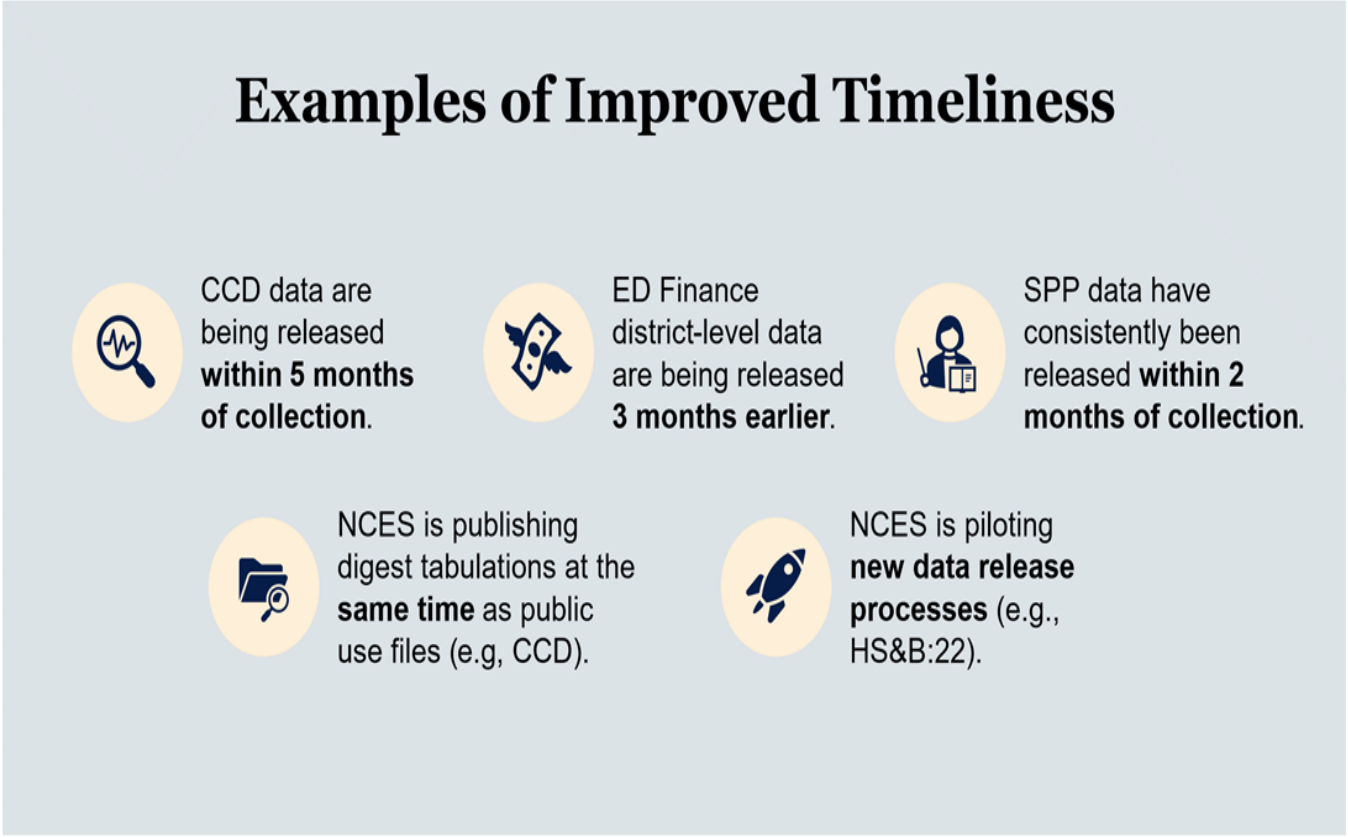


Note: Seasonally adjusted data are not available. Data online at https://data.bls.gov/timeseries/LNU0201B46B8&series_id=LNU0201B491&series_id=LNU0201B487.
Source: Bureau of Labor Statistics, Current Population Survey, September 6, 2024.

CPS Home : U.S. Bureau of Labor Statistics (bls.gov)

NCES – Adapting to change

NCES
National Center for Education Statistics



Quick turnaround surveys

United States® Census Bureau

Household Pulse Survey

USDA ERS

Business Trends and Outlooks Survey

NCES National Center for Education Statistics

School Pulse Panel

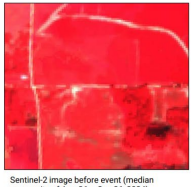
Centers for Disease Control and Prevention National Center for Health Statistics

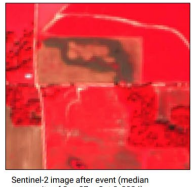
NCHS Rapid Surveys System

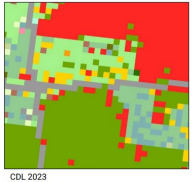
https://www.census.gov/programs-surveys/household-pulse-survey/data.html; https://www.census.gov/programs-surveys/btos.html; https://nces.ed.gov/surveys/spp/; RDC - RSS (cdc.gov)
;

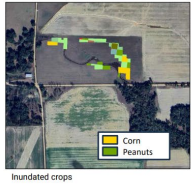
NASS – Near real-time estimates

Georgia

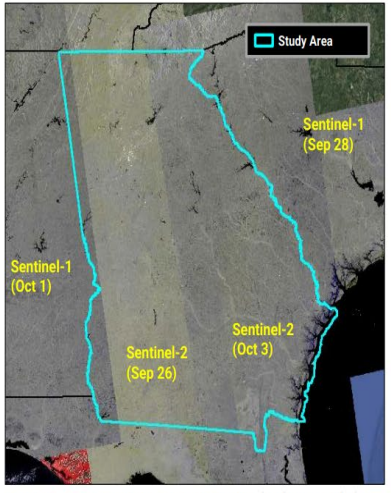

Sentinel-2 image before event (median composite of Aug 26 – Sep 26, 2024)


Sentinel-2 image after event (median composite of Sep 27 – Oct 3, 2024)


CDL 2023



Inundated crops



Assessment Dates (after inundation): 9/26/24 to 10/5/24
Reference Dates (before inundation): 8/26/24 to 9/26/24

Percent of Crop Acres Inundated by Hurricane Helene
September 2024
Georgia

Crop Type	Total Statewide Acres	Minimal Percent Inundated†
Corn*	485,000	0.31%
Cotton*	1,110,000	0.24%
Peanuts*	775,000	0.23%
Pecans***	148,000	0.19%
Soybeans*	160,000	0.21%
Total (selected commodities)	**2,678,000**	**0.25%**

*Acres Planted, NASS 2023

***Acres Bearing, NASS 2023

†Percent of acres impacted based on 1) all available post-event image acquisitions as of October 5, 2024, and 2) raw pixel counts from the 2023 CDL which are not official NASS estimates. Therefore, the amount of cropland affected by storm inundation may be different than these estimates indicate.

Conference Themes

Relevancy

Timeliness

Integrity

Federal Statistics – Looking Ahead

Opportunities

- Expanded use of data science & AI methodologies to create efficiencies
  - GenAI
  - NLP
  - ML
- Growing of data linkage work
  - Broader context & coverage
  - Answer complex questions
- Better, more comprehensive and expanded data access methods

Challenges

- Be ethical and responsible
- Apply appropriate measures & methodologies
- Ensure representativeness of population(s) of interest
- Safeguard equity and inclusion principles
- Be vigilant & safeguard TRUST

Practice & Profession of Statistics – Looking Ahead

https://www.amstat.org/

Thank you!

Jeri Mulrow
Vice President & Data Solutions Sector Lead
jerimulrow@westat.com
240-429-9181

westat.com