

Advancing the Frontiers of Social Science: Opportunities and Challenges

Jani Little, Executive Director

Katie Genadek, Expected Administrator

A stylized graphic of a mountain range with three peaks, rendered in dark purple and green, set against a teal background. The mountains are outlined in white and have a jagged, irregular top edge.

ROCKY MOUNTAIN RESEARCH CENTER
A FEDERAL STATISTICAL RESEARCH DATA CENTER

- What is a FSRDC and the FSRDC Program?
- The RMRDC
- The Proposal Process
- How Does the RMRDC Benefit You?
- Available Data and Ideas for RMRDC Research



What is a Federal Statistical Research Data Center (FSRDC)?

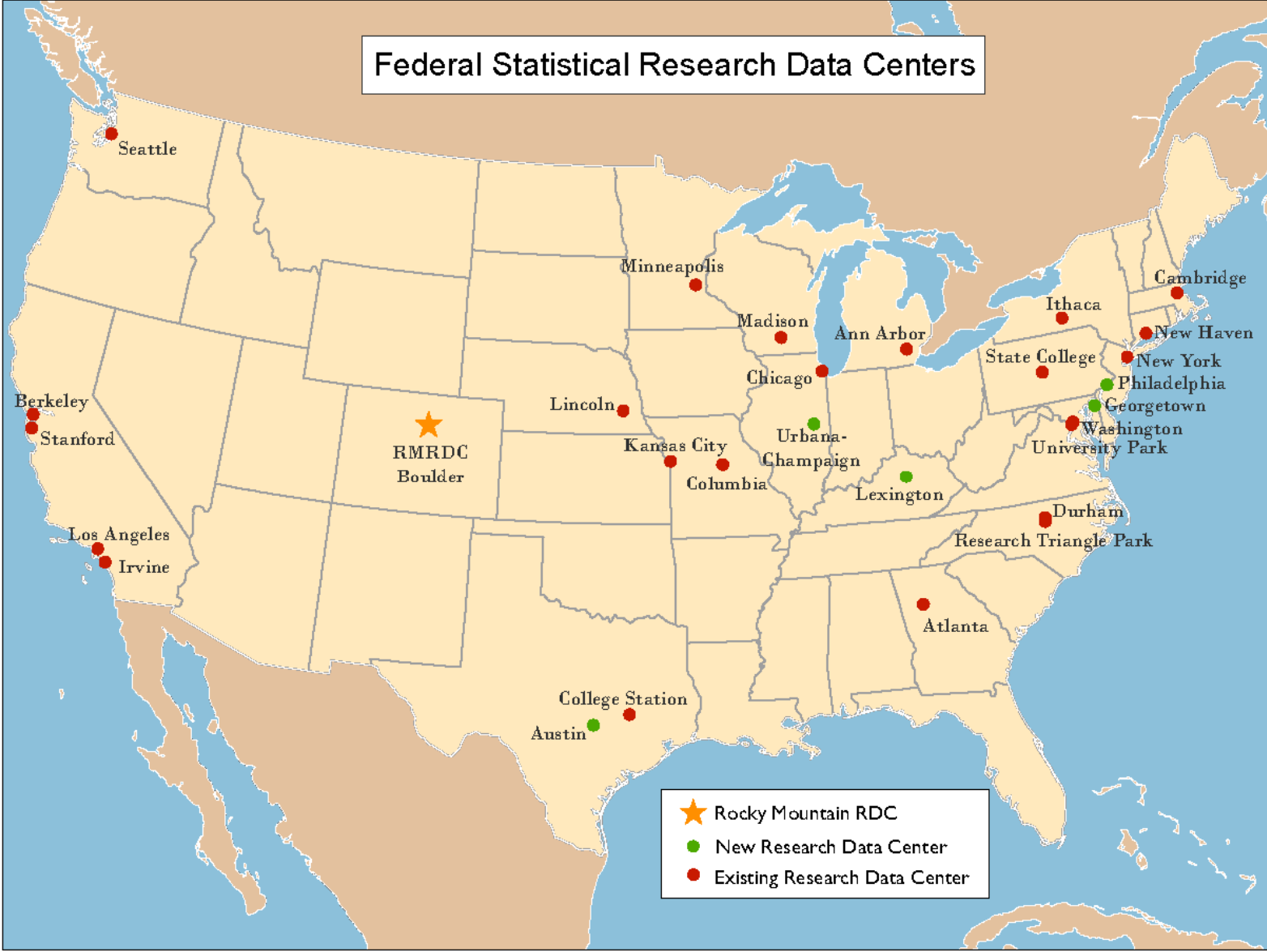
- A secure computing lab that is linked with the Census Bureau's internal computing network, which hosts restricted-access data from multiple federal agencies
- Made possible by a contractual agreement between a leading research institution and the U.S. Census Bureau
- The Census Bureau's Center for Economic Studies (CES) directs all FSRDCs and the FSRDC Program
- FSRDCs are managed by an on-site Census employee—the administrator—who guides researchers on proposal development, enforces security guidelines, and serves as liaison with the research community.

“This Research Data Center allows us to ***Engage Researchers*** outside of Washington in using this very important data while also ***Protecting the Public’s Rights to Privacy.***”

Former U.S. Census Bureau Director C. Lewis Kincannon



Federal Statistical Research Data Centers



- ★ Rocky Mountain RDC
- New Research Data Center
- Existing Research Data Center

Steve Ruggles, Creator of IPUMS, Historian, and Pro-active
Visionary concludes:

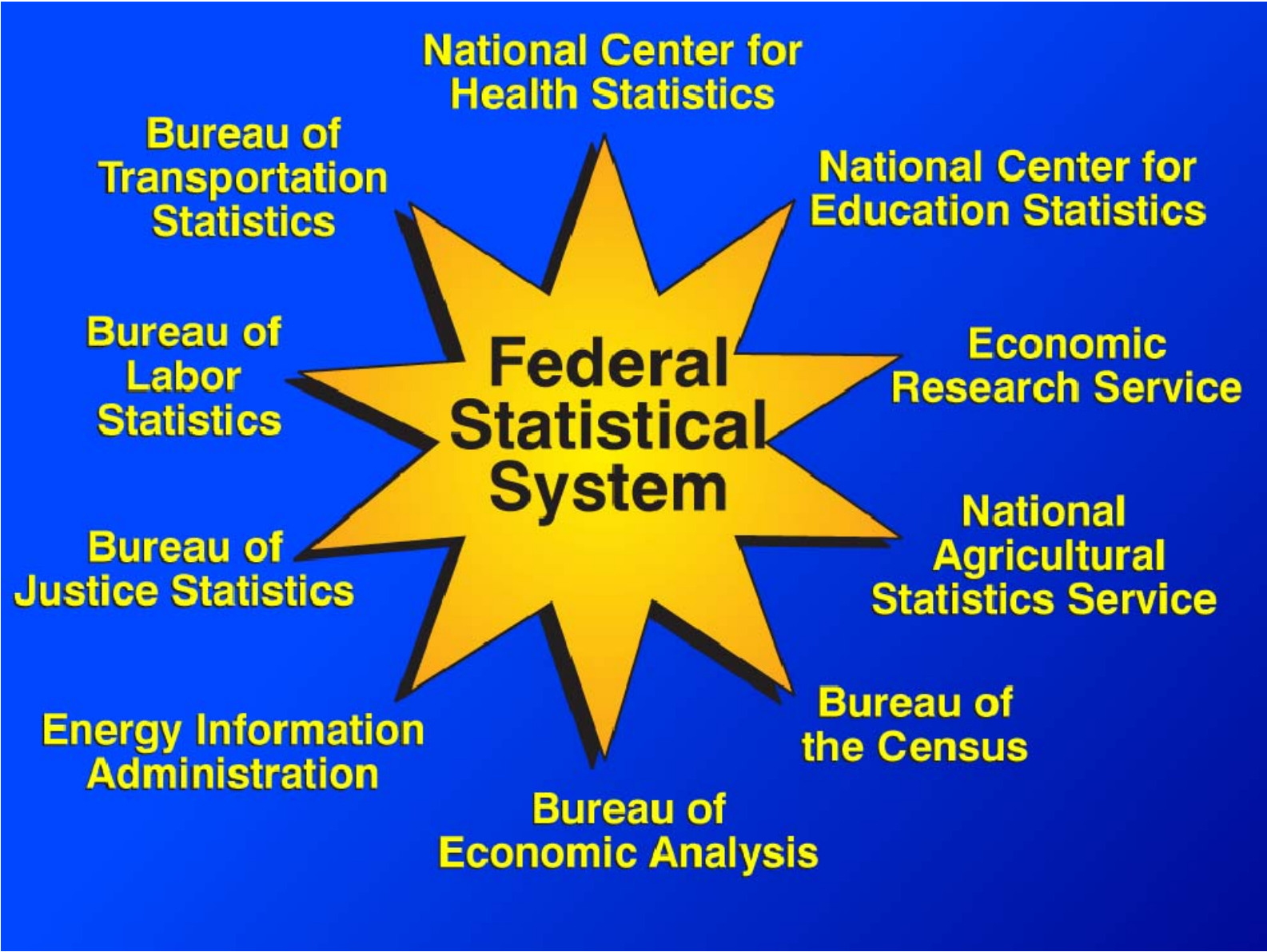
“The most recent milestone in developing a national infrastructure for research is the FSRDC: a trusted institutional mechanism for research access of confidential data.”



Recent Advances in the FSRDC Program:

- Expanding Participation by Federal Statistical Agencies
- More Efficient Administration of Program
- Bipartisan Commission on Evidence-Based Policymaking





What is the Commission on Evidence-Based Policy Making?

- The result of discussions between Congress and the Executive Branch on **improving how government uses research findings from survey and administrative data**
- **Appropriated \$10 million to Census Bureau's CARRA unit to build linkages between data sources**

	Researchers and Administrators		Privacy Experts
President	Kathrine Abraham <i>University of Maryland</i> (CHAIR)	Allison Orris OMB	Paul Ohm Georgetown University
Speaker of the House	Ron Haskins <i>Brookings Institution</i> (CO-CHAIR)	Bruce Meyers University of Chicago	Robert Hahn University of Oxford
House Minority Leader	Hilary Hoynes University of California, Berkeley	Sherry Glied New York University	Latanya Sweeney Harvard University
Senate Majority Leader	Kenneth Troske University of Kentucky	Robert Shea Grant Thornton LLP	Kathleen Rice Faerge Baker Daniels LLP
Senate Minority Leader	Jeffrey Liebman Harvard University	Kim Wallin Wallin Ltd.	Robert Groves Georgetown University

Advantages to IBS Researchers:

- Greatly expands the policy and basic questions that can be addressed
- Improves competitive edge for grants and publications
- Improves graduate education (big data/statistical techniques) and placement
- Attracts and retains top researchers and scholars
- Serves as a “hub” that attracts leading researchers to visit and collaborate
- Fosters a vibrant regional research community in the Rocky Mountains
- Fosters collaboration with researchers near other FSRDCs



RMRDC: The Physical Facility

Projected Opening Date: March 1, 2017

Location: IBS Rms 132+136

- 10 thin client workstations to access FSRDC servers
- Secure communications that tunnel over campus internet
- Contains the Administrator's office
- Badge Reader at Entrance
- 24/7 Security System with camera
- no electronic devices allowed
- NOTHING leaves the secure lab without approval



How is the RMRDC Funded?

Costs:

\$300,000 annual operations

\$250,000 initial construction

Contributions:

NSF 3-year award

CU Boulder Administration

IBS in kind

Consortium Partners

Funded Research and External Researchers



CU Faculty, Grad Students, and Affiliated
Researchers

Get free access to the RMRDC services and
secure laboratory

Costs for external researchers: \$15,000 per
year



The RMRDC Consortium

Partner Members:



University of Colorado
Boulder



Denver



Anschutz



UNIVERSITY
OF WYOMING

Other Members:

UC Colorado Springs

Colorado State Government

Colorado School of Mines

Denver University

National Center for Atmospheric Research

National Renewable Energy Laboratory

Requirements for Using Any FSRDC:

- Research projects must undergo a formal approval process with the agency that owns the data, e.g., Census, NCHS, BLS
- Researchers must go through a background investigation that qualifies them for “Special Sworn Status” which makes them an unpaid Census Bureau employee.
- Projects using Census data must provide benefits to the Census Bureau (PPS). All projects must show scientific merit and feasibility.
- Results must be formally reviewed for disclosure violation before they leave the secure facility.



Differences between Census and NCHS Proposals:

	Census Proposals	NCHS Proposals
Approval %	70	nearly 100
Time to Approval	3-4 months on average	less on average
Predominant Purpose Statement (PPS)	Required	Not Required
Data Extract Fee	None	\$1200
Scope	Broad as Possible	Precise



FSRDC Server Software

Gauss

Stata

Matlab & toolboxes

PBS Pro

Intel Composer XE

NX Enterprise

R

SAS

SAS (Dataflux)

SUDAAN

GEODa

Tomlab

Knitro

Madd

QGIS

StatTransfer

Python - Anaconda

Fortran

Perl

Tex/LaTex



RMRDC

Rocky Mountain Research Data Center
www.colorado.edu/rocky-mountain-research-data-center

What Do I Do? When?

Contact the RMRDC Director and Administrator

Review materials at our website, the CES website, and other relevant sites listed on our website.

When? **Start ASAP. Approval is needed from Census and/or other agencies. This process takes about 90 days once your proposal is submitted. IRS takes additional time. SSS takes additional time.**



How Will the RMRDC Staff Help?

We will research information relevant to the researcher's interest

- data availability**
- project budget and timeline**
- contact information**

For Census projects, the Administrator will give invaluable guidance on the Proposal Development Process

- the benefits to Census (PPS)**
- help navigate the project approval process**

Give assistance and training in using the Secure Lab



Upcoming Opportunities

1) The Cornell University Course by Videoconferencing

Taught by John Abowd and Lars Vilhuber in the Spring semester

Understanding Social and Economic Data

The 2016 version is particularly aimed at American Ph.D. students who are interested in using **public-use** and **confidential** U.S. Census Bureau data, and the confidential data of other American statistical agencies that cooperate with the Census Bureau.

2) Proposal Development Workshops

One day events in IBS between semesters

3) Invited speakers with FSRDC expertise



FSRDCs Used to Address Many Research Topics

- Business, Trade, Finance, and Management
- Crime and Crime Victimization
- Demography, Population Distributions and Trends, Migration, and Immigration
- Economics, Labor Markets, Entrepreneurship, Employment and Industry
- Education and Education Policy
- Hazard Mitigation, Environmental Impact Assessment, Pollution Abatement
- Health and Well-Being, Health Insurance, Health Policy
- Housing, Housing Markets, and Residential Patterns
- Poverty and Social Welfare Policy
- Transportation Analysis and Planning
- Urban and Regional Economics and Planning

Advantages Provided to Research?

- Microdata not available publicly on individuals and households, firms and establishments (including Personal Identification Key (PIK) & Employer Identification number (EIN))
- Variables not available in public versions of data sets (e.g., low level geography—block and block group)
- Full population counts or larger samples (Decennial Census, ACS, CPS)
- Full range of response items before they are “edited” and processed (e.g., detailed race answers, income is not top-coded, etc.)
- Finer codings of variables (e.g., 5 digit industry codes)
- Ability to link with external data (e.g., via geocodes, establishment ID, etc.)
- Ability to merge multiple internal data sets via non-public link keys

Major Partners in the FSRDC System

- U.S. Census Bureau
 - [Economic Data](#)
 - [Demographic Data](#)
 - [Longitudinal Employer-Household Dynamics \(LEHD\) Data](#)
- [Bureau of Labor Statistics \(BLS\)](#)
- [National Center for Health Statistics \(NCHS\)](#)
- [Agency for Healthcare Research and Quality \(AHRQ\)](#)
 - [Other Federal Partners](#)

Restricted Census Data*

Economic Data

- Economic Censuses
- Longitudinal Business Database (LBD)
- Annual Survey of Manufactures (ASM)

Longitudinal Employer-Household Dynamics Data (LEHD)

- Employment History File (EHF)
- Employer Characteristics File (ECF)

*<https://www.census.gov/ces/dataproducts/economicdata.html>

Restricted Census Data*

Demographic Data

- Decennial Censuses
- American Community Surveys (ACS)
- Current Population Survey (CPS) & Supplements
- American Housing Surveys (AHS)
- Survey of Income and Program Participation Panels (SIPP)

*<https://www.census.gov/ces/dataproducts/demographicdata.html>

NCHS Restricted Data I: National Health Status Surveys

- National Health and Nutrition Examination Survey (NHANES) I, II, and III
- National Health Interview Survey (NHIS)
- Longitudinal Study on Aging I and II (LSOA)
- National Survey of Family Growth
- National Survey of Children's Health
- National Survey of Early Childhood Health
- National Survey of Children with Special Health Care Needs
- National Asthma Survey

NCHS Restricted Data II: National Health Care Surveys

- National Ambulatory Medical Care Survey
- National Hospital Ambulatory Medical Care Survey
- National Survey of Ambulatory Surgery
- National Hospital Discharge Survey
- National Nursing Home Survey (NNHS)
- National Home and Hospice Care Survey
- National Employer Health Insurance Survey
- National Health Provider Inventory
- National Immunization Survey

NCHS Restricted Data III:

Vital Statistics

- Mortality and Multiple Mortality
- Birth
- Fetal Death
- National Death Index

- [Linked Mortality Data Products](#)
- [Linked Medicare Enrollment and Claims Files Data](#)
- [Linked Medicaid Enrollment and Claims Data](#)
- [Linked Social Security Benefit History Data](#)

Potential IBS Research Projects

Business Impact of Natural Disasters:

Longitudinal Business Database (LBD) annual

Survey of Business Owners (SBO) 2007 and 2012

Economic Census 2007 and 2012

Longitudinal Employer-Household Dynamics (LEHD)



Potential IBS Research Projects (con't)

Neighborhood and employment effects on recidivism:

Lonestar Project Data (external, David Pyrooz, PI)

ACS block group data (neighborhood instability and SES)

2010 Decennial Census block data (race, ethnic, and immigrant enclaves, female headed households)

LEHD (EHF and ECF) (local labor market and individual employment)



Potential IBS Research Projects (con't)

Migration county level: health, amenity, and labor drivers and consequences

older population: Health and Retirement Survey (HRS) (external)

cohorts: National Longitudinal Survey of Youth 1979 (NLSY79) and
National Longitudinal Survey of Youth 1997 (NLSY97)

Migration county level: in and out flows

Decennial Census Longform: 2000 and prior

ACS: 2005 and beyond

Migration neighborhood level: gentrification LEHD (EHF)



Potential IBS Research Projects (con't)

County-level Factors in Declining Teen Births:

Vital Statistics

Decennial Census and ACS

Neighborhood Disadvantage and Suicide:

Geocoded National Health Survey data linked to mortality data



Potential IBS Research Projects (con't)

Changes in State Laws and Changes in Marijuana and Other Drug Use:

National Longitudinal Survey of Youth 1979 (NLSY79)

National Longitudinal Survey of Youth 1997 (NLSY97)



Katie Genadek, Expected RMRDC Administrator

Ph.D. Applied Economics, University of Minnesota, 2012

IPUMS USA Project Manager

Manager of User Support for Integrated:

NHIS

CPS

ACS

IPUMS International

DHS



Current RDC projects of interest – Census Data

- Linking 2010, 2000 full count census to assess race response changes and how is racial identity influenced by neighborhood characteristics
- Interracial couples and children's race – how are they reported, how does that change over time

Current RDC projects of interest – Census Data

- Migration and duration of international student graduates (NSCG merged to census data)
- The effects of 1960 individual housing conditions on urban renewal



Current RDC projects of interest – Health Data

- Effects of Medicaid expansion on insurance coverage duration and access to health care (NHIS)
- Effects of job and family demands on health and health behaviors (MEPS with merged O*NET)
- How does race on death certificates compare to self reported race? How does this impact race-specific mortality rates? (NHIS linked to NDI)



Current RDC Infrastructure projects

Harmonization:

- NHIS data integrated over time, merged with internal data
- Currently harmonizing codes in 1960-2010 Census and ACS data for ease of use over time
- Harmonized geographies in the census data



My (hopeful) future RDC projects

- Effect of Great Recession/housing crisis on Marriage and Divorce: bring in mortgage foreclosure data, merge on to ACS data at finer geographic level
- Child care costs and time spent alone with spouse: ATUS data in RDC, proprietary data on caregiving prices at finer geographic levels
- Life cycle fertility and employment for women in the US over time

