

LA County Arts Datathon: Data Rescue Hackathon

Erin Meyer and Homeyra Banaeefar, ICPSR

April 29, 2026

Session Overview

- Importance of preserving federal data
- Introduction to ICPSR and DataLumos
- Data rescue activity

Federal Data

- Information created, collected, and maintained by the federal government. Includes data files, reports, summary statistics, and more.

The Home of the U.S. Government's Open Data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and more.

308,721 DATASETS AVAILABLE

Value of Federal Data

- Historical record
- Informed decision-making, research, and innovation
- Economic planning, disaster response, public safety
- Public trust

How have you used federal data?

Value to Arts and Culture

- Museums & galleries:
 - Audience mapping, school/community partnerships, conservation risk planning
- Performing arts:
 - Touring market selection, pricing & access, transit-informed scheduling
- Libraries & archives:
 - Community needs, digitization & service grants

Value to Arts and Culture, cont.

- Festivals & public art:
 - Site selection, permitting/transport, weather & hazard readiness
- Arts education:
 - Identify underserved schools, tailor programs, measure outcomes
- Creative placemaking/cultural districts:
 - Document need & change, support equitable investment
- Artists & researchers:
 - Source material, data-driven storytelling

LA County Open Data Portal

LA County at a Glance

Total Population

10,014,009

Source:
[USA Census 2020 Redistricting Counties](#)

Total Housing Units

3,591,981

Source:
[USA Census 2020 Redistricting Counties](#)

Median Household Income (\$)

87,760

Source:
[ACS Median Household Income Variables - Boundaries](#)

Poverty Rate (%)

13.6

Source:
[ACS Poverty Status Variables - Boundaries](#)

Median Rent (\$)

1,747

Source:
[ACS Housing Units Occupancy Variables - Boundaries](#)

Unemployment Rate (%)

5.5

Source:
[Bureau of Labor Statistics Monthly Unemployment \(latest 14 months\)](#)

Open Access

- Federal government is not mandated to preserve all of their data
- Open data initiatives promote transparency

GSA Open Data

GSA provides a wide variety of data sets that can be **freely used, re-used and redistributed by anyone**. Opening this data increases public participation in government, creates opportunities for economic development, and informs decision making in both the private and public sectors. So please, explore some of our public data—innovate, create, and make sure to give us your feedback.

[Explore GSA data on Data.gov](#)

Inter-university Consortium for Political and Social Research (ICPSR)

- Data archive of more than 350,000 files of research in the social and behavioral sciences
- Specialized data collections in education, aging, criminal justice, substance use, and more
- Leader in data stewardship
- Summer Program in Quantitative Methods of Social Research

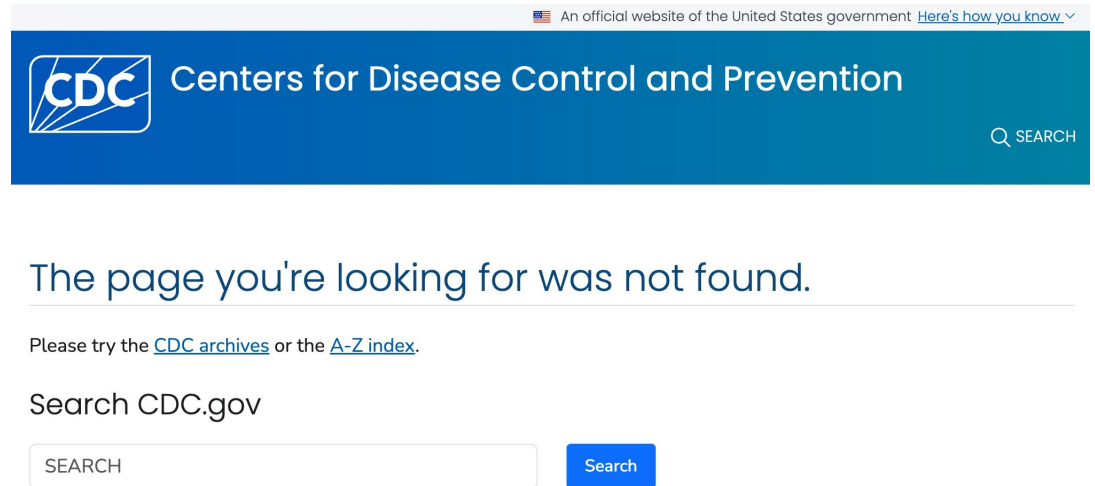
DataLumos Launch

- Launched in 2017 in response to concerns about potential government data loss
- Crowdsourced, self-publishing repository to preserve and disseminate at-risk government data
- 101 deposits made in first four years
- Minimal activity 2019-2024



2025 Administration Change

- January 2025 Executive Orders on sex and gender, DEI, and foreign assistance
- Estimated that more than 8,000 .gov web pages were removed



The screenshot shows the top of the CDC website. At the top right, it says "An official website of the United States government" with a link "Here's how you know". Below that is the CDC logo and the text "Centers for Disease Control and Prevention". On the right side, there is a search icon and the word "SEARCH". The main content area displays the message "The page you're looking for was not found." followed by a horizontal line. Below the line, it says "Please try the [CDC archives](#) or the [A-Z index](#)." At the bottom, there is a search bar with the text "Search CDC.gov" above it. The search bar contains the word "SEARCH" and has a blue "Search" button to its right.

Jumping into Action

- Wrote APIs to scrape content from federal agency websites
- Downloaded data that seemed at risk of being removed
- Hunted down data that had already disappeared
- Connected with other organizations trying to do similar things
- **Revived DataLumos!**

DataLumos Website

Login



Find Data

Upload Data

About DataLumos

A crowd-sourced repository for valuable government data

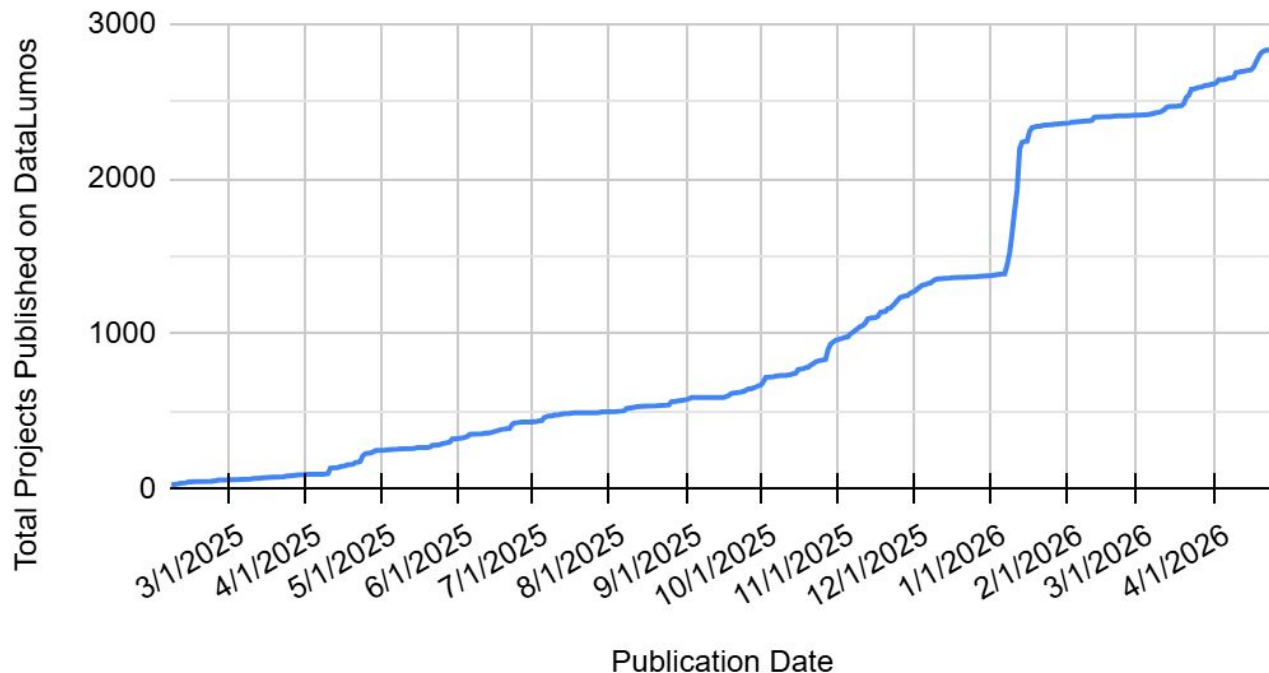
DataLumos is an ICPSR archive for valuable government data resources. ICPSR has a long commitment to safekeeping and disseminating US government and other social science data. DataLumos accepts deposits of public data resources from the community and recommendations of public data resources that ICPSR itself might add to DataLumos. Please consider making a monetary donation to sustain DataLumos.

Support DataLumos

ICPSR

DataLumos Projects (through 4/26/2026)

DataLumos Growth Since January 2025



Data Rescue Hackathons

- Events where attendees work together to tackle a data rescue task
 - Downloading datasets
 - Creating or reviewing metadata
 - Adding rescued files to a repository like DataLumos



Targeted Data Rescue Initiatives

- NIOSH data
- USAID education data
- HIFLD Open database
- Department of Education DataLab tables
- Tracking changes in SOGI availability

How to Stay Involved

- Participate in today's activity, attend future events, or host one of your own!
- Preserve data in DataLumos
- [Help us enhance the metadata for DataLumos projects](#)
- [Sign up to volunteer with the Data Rescue Project](#)
- Add to the collection of real world examples of federal data benefits: [essentialdata.us](#)

Thanks

- Trent Alexander
- Homeyra Banaeefar
- Aalap Doshi
- Libby Hemphill
- Lynette Hoelter
- Jeannette Jackson
- Bella Kan
- Jennifer Koski
- John Kubale
- Maggie Levenstein
- Kelly Ogden-Schuette
- David Thomas
- Alexandra Toma

Questions?

icpsr-data-rescue@umich.edu

Data Rescue Activity

Today's Data

1. Current Population Survey: Annual Social and Economic Supplement
 - Economic status of the US population. Occupation and industry variables can flag arts and culture workers
2. American Time Use Survey
 - How people spend time, including arts/entertainment, socializing, volunteering, and media
3. Consumer Expenditure Survey
 - Household spending, including entertainment, admissions, AV equipment, reading materials

Activity Instructions

myumi.ch/qw8ZZ

1-2 Volunteers to Report Out

- What did you learn in your breakout session that can help you build and maintain knowledge?
- What is one practice or tool you learned about that you'd like to put to use?
- What is one question about building and maintaining knowledge in the arts that you're leaving with today?

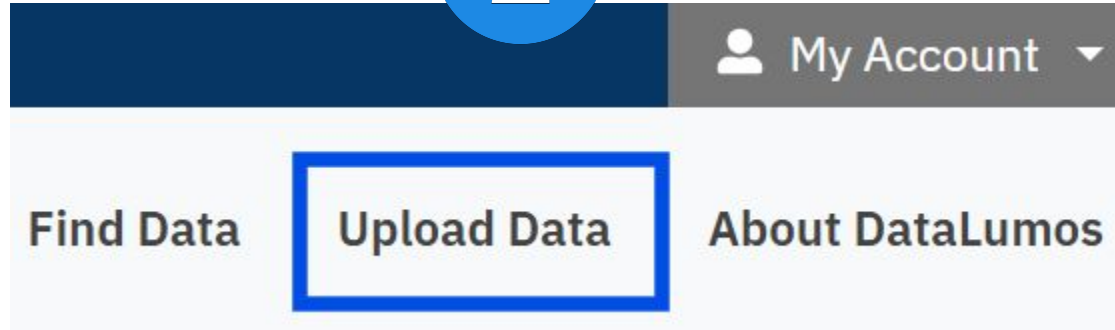
Screenshots

Start a new DataLumos project

1



2



3



Hide inactive

Download data folder

... > public_data > 2juy-3evq ▾ 👤

Type ▾

People ▾

Mo

📁 New folder Alt+C then F

🔗 Open with ▸

📄 Download

✎ Rename Ctrl+Alt+E

👤 Share ▸

📁 Organize ▸

ℹ Folder information ▸

🗑 Move to trash Delete

Name ↑

Last modified ▾

File size

☰ 2juy-3evq_data.zip 👤

May 28, 2025 me

458 KB

📄 about_data_2juy-3evq.htm

May 28, 2025 me

7.4 MB

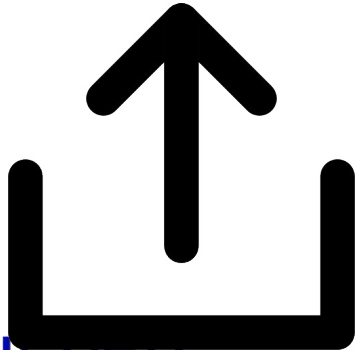
Upload data

HUD FHA Single Family Portfolio Snapshot

Create Folder Upload Files **Import From Zip** more ▾

<input type="checkbox"/>	Name	Type	Size
<input type="checkbox"/>	FHA Single Family Po...		

Project



Project Description

Government Agency/Principal Investigator(s) (required) + add value
United States Department of Housing and Urban Development edit remove

Summary (required)
The Single-Family Portfolio Snapshot consists of a monthly data table and a report generator (Excel pivot table) that can be used to quickly create new reports of interest to the user from the data records. The data records themselves are loan level records using all of the categorical variables highlighted on the report generator table. Users may download and save the Excel file that contains the data records and the pivot table.

The report generator sheet consists of an Excel pivot table that gives individual users some ability to analyze monthly trends on dimensions of interest to them. There are six choice dimensions: property state, property county, loan purpose, loan type, property product type, and downpayment source.

Each report generator selection variable has an associated drop-down menu that is accessed by clicking once on the associated arrows. Only single selections can be made from each menu. For example, users must choose one state or all states, one county or all counties. If a county is chosen that does not correspond with the selected state, the result will be null values.

The data records include each report generator choice variable plus the property zip code, originating mortgagee (lender) number, sponsor-lender name, sponsor number, nonprofit gift provider tax identification number, interest rate, and FHA insurance endorsement year and month. The report generator only provides output for the dollar amount of loans. Users who desire to analyze other data that are available on the data table, for example, interest rates or sponsor number, must first download the Excel file. See the data definitions (PDF in top folder) for details on each data element.

Files switch from .zip to excel in August 2017. edit remove

Funding Sources + add value

Original Distribution URL
https://www.hud.gov/program_offices/housing/rmra/oe/rpts/sfsnap/sfsnap edit remove

+ Scope of Project

+ Methodology

+ Related Publications + add citation

Project Communication Log You are NOT watching comments + add entry watch

Storage Status:

Space:
3.13 GB of 30.00 GB used (10%)

File/Folder:
194 of 1000 used (19%)

Published Versions

[V1 - 2/20/2025 4:19:37 PM]
Edit Permissions

Publish!

Create Folder Upload Files Import From Zip more

Delete Selection

<input type="checkbox"/>	Name	Type	Size	Last Modified	Actions
<input type="checkbox"/>	Family Unification Program Youth Data			2/11/2025 5:34:43 PM	

Project

- Collapse All

datalumos-219103

Publish Project

Share Project

Change Owner

View Log

Report a Problem

Storage Status:

