

CDC Data Science Upskilling Program

Developing a future-ready CDC workforce through hands-on data science training and team-based projects



Participants in the Data Science Upskilling (DSU) program apply modern data science methods to advance CDC priorities and improve public health practice. This program supports [CDC's priority for Public Health Data Modernization](#) by expanding data science capabilities across the agency.

Why this Program Matters

DSU transforms the way public health professionals address and solve real-world challenges by training staff on new data science skills so they can grow in their current positions and stay effective as technology changes. The program:

- **Advances CDC priorities** through staff-driven, real-world projects.
- **Builds data science capacity** by upskilling interdisciplinary teams from novice to advanced levels.
- **Accelerates data modernization** through practical, scalable solutions that improve decision-making and efficiency.
- **Strengthens CDC's long-term analytical capacity** by equipping staff with modern data visualization, artificial intelligence (AI), and machine learning (ML) skills.

By the Numbers



552
individual learners
upskilled since 2019



138
data science projects
led by DSU teams
between 2019 and 2025



100%
of 2024 end-of-year
survey respondents said
DSU increased their data
science knowledge



What Participants Do

Over a 10-month period, DSU teams develop practical data science skills while they:

- Design and deliver projects that support state, tribal, local, territorial (STLT), and international health agencies.
- Apply AI and machine learning to improve forecasting accuracy and situational awareness during outbreaks and public health emergencies.
- Generate data that improve public health recommendations.
- Learn data science tools that improve efficiency, reduce costs, and enhance responsiveness.

Public health data science is increasingly combining visualization and statistical methods with AI and ML.

Across 112 CDC data science upskilling projects (2019–2023):



54% used data visualization



51% used statistical methods



42% used AI or ML



52% supported public health decision-making

Source: *Decoding Data Science Upskilling: Insights From 5 Years of Data Science Projects at the Centers for Disease Control and Prevention, 2019-2023*, Journal of Public Health Management & Practice

A Closer Look

DSU participants apply new skills to mission-critical, team-based projects that strengthen public health performance measurement, effectiveness, accountability, and continuous improvement. DSU teams have:



Supported faster, more effective emergency response at STLT levels by optimizing Firefighter Reporting System data.



Accelerated outbreak identification and response by building a tool to detect enteric disease clusters, which happen when a group of people in the same area or connected by a common event get a stomach illness around the same time.



Identified public health priorities within HIV/AIDS programs and provided targeted technical assistance by developing search tools and analyzing data from multiple sources to assess the historical impact of funding.



Used data to inform prevention and programmatic activities at the national, state, and local levels by creating a surveillance dashboard that displays commercial menthol tobacco use.



Guided interventions to improve pedestrian infrastructure and public health outcomes by applying computer vision and satellite imagery to classify street-level scenes, extract sidewalk networks, and identify areas needing intervention.



Monitored progress and accountability for the Orphans and Vulnerable Children (OVC) program through a team at CDC Nigeria by developing a data visualization dashboard.