# **Resources and Learning Opportunities**

## **Definitions:**

[**Scientific Data**](https://www.nnlm.gov/guides/data-glossary/data)**:** definition from the NIH’s Data Management and Sharing Policy: “The recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications.”

**Data Input:** A way to capture or record raw data

**Data processing or** [**Data Wrangling**](https://www.nnlm.gov/guides/data-glossary/data-wrangling)**:** Converting the recorded (or “raw”) data into a more meaningful form; or preparing data for analysis

[**Data Collection**](https://www.nnlm.gov/guides/data-glossary/data-collection)**:** The process of gathering, capturing, or recording raw data

[**Data Analysis:**](https://www.nnlm.gov/guides/data-glossary/data-analysis)To extract meaningful information from data

[**Metadata:**](https://www.nnlm.gov/guides/data-glossary/metadata)Information that describes, explains, and otherwise contextualizes data

[**Variable:**](https://www.nnlm.gov/guides/data-glossary/variable)A symbol (be it text-based, numerical, date formatted, etc.) that represents information. It is not fixed in value, but can vary based on circumstances

[**Database:**](https://www.nnlm.gov/guides/data-glossary/database) An organized collection of data stored electronically.

## **Spreadsheet How Tos**

**Cell:** An individual rectangle/field in a spreadsheet. Each cell contains a **Value**.

**Value:** The contents of a cell; values can be text-based, numerical, dates, etc. (See Data Type)

**Columns:** Spreadsheets are generally formatted to contain **variables** as their columns/column names

[**Variable**](https://www.nnlm.gov/guides/data-glossary/variable)**:** In research, each variable is a feature or property that is being observed or measured

**Rows**: Spreadsheets are generally formatted for the rows to contain instances of observation

**Data Type:** Go into the formatting settings for each column and format it appropriately for the type of data you have. Most common are:

**Text:** For product names, personal names, etc.

**Numeric**: For simple numeric counts, prices, etc.

**Dates:** Set up according to your preferred format, e.g.: mm/dd/yy; mm/dd/yyyy; yyyy/mm/dd, etc.

### **Primary vs. Secondary Data**

Primary data: Data collected directly by a researcher

Secondary data: Data collected previously from a different source, which can be used to answer other questions than the original

Secondary data analysis: Uses secondary data with the aim of reproducing the original study and attempt to answer its questions, or to answer questions completely different from those of the original study.

### **Quantitative vs. Qualitative Data**

[**Quantitative data**](https://www.nnlm.gov/guides/data-glossary/quantitative-data) is measurable data and is generally expressed numerically.

[**Qualitative data**](https://www.nnlm.gov/guides/data-glossary/qualitative-data), especially in its rawest form, is not measurable, and is generally any kind of non-numeric data. Qualitative data often must be coded to be made quantifiable for analysis.

## **Introductory Classes:**

### On-Demand Classes

* [Data Services On-Demand (4 modules)](https://www.nnlm.gov/training/class-catalog/data-services-demand)
* [Health Data Standards and Terminologies A Tutorial](https://www.nlm.nih.gov/oet/ed/healthdatastandards/index.html)
* [Common Data Elements: Standardizing Data Collection](https://www.nnlm.gov/training/class-catalog/common-data-elements-standardizing-data-collection)

### Webinars/Recordings

* [Making Sense of Numbers: Communicating Numerical Health Information](https://www.nnlm.gov/training/class-catalog/making-sense-numbers-communicating-numerical-health-information)
* [Health Statistics on the Web](https://www.nnlm.gov/training/class-catalog/health-statistics-web)
* [The Research Data Services Landscape: How Do You Start And Where Does Your Library Fit In?](https://www.nnlm.gov/training/class-catalog/research-data-services-landscape-how-do-you-start-and-where-does-your)
* [NNLM YouTube Channel](https://www.youtube.com/%40NnlmGov)

## **Skill Building:**

### Data management

* [Research Data Management On-Demand](https://www.nnlm.gov/training/class/research-data-management)
* [Creating Data Management Plans with the DMPTool](https://www.nnlm.gov/training/class-catalog/creating-data-management-plans-dmptool)
* [Fundamentals of Health Sciences Research Data Management](https://www.nnlm.gov/training/class-catalog/fundamentals-health-sciences-research-data-management)

### Data science

* [Open Tools for Data De-identification](https://www.nnlm.gov/training/class-catalog/open-tools-data-de-identification)
* [Fundamentals of Health Data Science](https://www.nnlm.gov/training/class-catalog/fundamentals-health-data-science)

### NIH Data Management and Sharing Policy

* [NIH Data Management and Sharing Policy Overview](https://www.nnlm.gov/training/class-catalog/nih-data-management-and-sharing-policy-overview)
* [NIH Data Management and Sharing Policy Workshop On-Demand](https://www.nnlm.gov/training/class-catalog/nih-data-management-and-sharing-policy-workshop-demand-0)
* [NIH Data Management and Sharing Requirements Series – practitioners perspectives](https://www.youtube.com/watch?v=M_LDO359BHs)
* [The NIH Data Management and Sharing Policy for non-data librarians – recording available](https://www.youtube.com/watch?v=6JAj5rHpFd0)

## **Deep Dive (non-NNLM)**

* [The Carpentries](https://carpentries.org/)
	+ Library Carpentry
	+ Data Carpentry
	+ Software Carpentry
* [Data Services Continuing Professional Education (DSCPE)](https://dscpe.github.io/)
	+ Cohorts run for 10 weeks in the Fall

## **Resources:**

[Data Literacy Project](https://thedataliteracyproject.org/data-literacy-courses/) (free courses)

[Georgia State University Data Ready! Badges](https://research.library.gsu.edu/dataservices/rds-workshops-recordings) (recordings)

[Top data literacy skills for becoming data literate](https://www.tableau.com/data-insights/data-literacy/skills) (article)

## **Resources for Common Health Data Terms:**

[**NCDS – NNLM Data Glossary**](https://www.nnlm.gov/guides/data-glossary)

[**NIH Common Data Element Repository**](https://cde.nlm.nih.gov/home)

[**Data management in clinical research – an overview**](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3326906/)

[**Glossary of Commonly Used Healthcare Terminology**](https://portal.ct.gov/-/media/OHS/Health-IT-Advisory-Council/Miscellaneous/20161117_Glossary_of_Terms_FINAL.pdf)

[**Health and Human Services. Assistant Secretary for Planning and Evaluation, Office of Science & Data Policy: Glossary of Common Data-Related Terms**](https://aspe.hhs.gov/sites/default/files/documents/248275672affa2035cccfbcdaa3670d2/aspe-working-paper-data-terms.pdf)