

The Joy of Data: *Data Dictionary*
wayne.smith@csun.edu
[updated: Friday, June 20, 2025]

"The beginning of wisdom is to call things by their right names."
--Confucious (Kongzi) (551 B.C.E. – 479 B.C.E.)

Introduction:

Documentation is a core part of data literacy. For example, documenting the cells in a spreadsheet is important. Part of documentation is also what we call in this course a “Data Dictionary”. A Data Dictionary describes important characteristics of each column in the spreadsheet, such as Description, Type, Length, Number of Decimals, and other aspects of the data. Data Dictionaries not only help the *creator* of the data but also *others* that might use the data at a different time or in a different way. Data Dictionaries can also go by other names because 1), data dictionary concepts originate in different disciplines at different times, and 2), the purpose and use of Data Dictionaries can vary somewhat by context. The most important thing to remember is that “you can never have too much documentation.” Business students learn to incorporate documentation into their regular data analysis work.

The following are examples of different terms for a “Data Dictionary”.

Domain	Name	Reason/Notes
<i>Business Spreadsheet</i>	Data Dictionary	It’s simple and intuitive. It is similar to a regular dictionary that has a list of words and their meanings.
<i>Database</i>	Schema	A schema is a “plan” for the data. In a database, the schema is not just recommended it’s required for the data to be created, read, updated, and deleted properly.
<i>Research Data</i>	Codebook	Often, especially in surveys and related data collection activities, numbers are used as a placeholder (or “code”) for a character. For example, a “1” might mean female and a “0” might mean male. The codebook explains how each code should be interpreted.
<i>Computer Science</i>	Metadata	The term Metadata means “data about data”. That is, metadata is, itself, data. It is human- and machine-readable.
<i>Library</i>	Catalog	Prior to the Internet, indexes to books were kept in big “card catalogs”. These were file cabinets with a card that listed where a particular book could be located.