Tool I: Sample Record Layout

The following Sample Record Layout corresponds to the sample spreadsheet in the beginning of Part 3. The record layout is an index or listing of all the variables in a dataset. It lists variables in sequential order and describes the fixed width and data structure of each variable. Record layouts are helpful in understanding the data definitions in a dataset or spreadsheet.

	A	B	C	D	E
1	# VARIABLE	VARIABLE NAME	VARIABLE DEFINITION	VARIABLE FORMAT	VARIABLE TYPE
2	1	SCHOOL NAME	STATE SYSTEM SCHOOL NAME	A60	ALPHANUMERIC

3	2	SCHOOL ID	STATE SYSTEM SCHOOL ID # WHERE FIRST TWO DIGITS ARE:
			01=ELEMENTARY
			02=MIDDLE
			03=SECONDARY
				F5.0	NUMERIC

4	3	GRADE	K THROUGH 12; ALL	A3	ALPHANUMERIC
5	4	TOTAL GRADE ENROLLMENT	TOTAL ENROLLMENT IN GRADE AS OF 20TH DAY	F5.0	NUMERIC
6	5	TOTAL ARTS ENROLLMENT	TOTAL ENROLLMENT ALL ARTS COURSES AS OF 20TH DAY	F5.0	NUMERIC
7	6	VISUAL ARTS ENROLLMENT	ENROLLMENT IN VISUAL ARTS COURSES AS OF 20TH DAY	F5.0	NUMERIC
8	7	MUSIC ARTS ENROLLMENT	ENROLLMENT IN MUSIC ARTS COURSES AS OF 20TH DAY	F5.0	NUMERIC
9	8	DANCE ARTS ENROLLMENT	ENROLLMENT IN DANCE ARTS COURSES AS OF 20TH DAY	F5.0	NUMERIC

10	**9**	**THEATRE ARTS ENROLLMENT**	**ENROLLMENT IN THEATRE ARTS COURSES AS OF 20TH DAY**	**F.5.0**	**NUMERIC**
11	**10**	**PERCENTAGE OF STUDENTS ENROLLED IN ARTS COURSES**	**PERCENTAGE OF STUDENTS TAKING ANY ARTS COURSES**	**F5.2**	**PERCENT**

How to Read This Table

The rows above describe variables in your dataset. These variables often appear as column headings in that dataset.

VARIABLE NAMES AND VARIABLE DESCRIPTIONS.

In this example, the Variable Names are easy to understand, but state agencies commonly abbreviate variable names. For example, "Total Arts Enrollment" might appear as "TArEnroll," or something similar. In that case, you would have to refer to the Variable Description to understand what those abbreviations mean.

VARIABLE FORMAT

In the Variable Format column, "A60" in Row 2, above, indicates two things:

- The "A" tells you that it's an alphanumeric variable, which means a variable that can be either letters, numbers or both. Alphanumeric variables typically identify labels or names.
- The "60" tells you that these variables can't be any longer than 60 characters.

Row 4, whose Variable Format is "A3," follows the same rule: As an indicator of grade level – for example, 9, 10, 11, 12 or all – it's an alphanumeric variable that cannot exceed three characters.

In the Variable format column (Rows 3 and 5 through 11), "F5.0" and "F5.2" tell us different things. The "F" tells us that this variable is quantitative. The numbers following the "F" tell us more about the format of the quantitative variable:

- "F5.0" (in Rows 3 and 5 through 10) tells us that the variable is no more than five characters long and contains no decimals — for example, 66.
- "F5.2" (in row 11) tells us that the variable is no more than five characters long, and that the last two characters are decimals. In this case, the variable describes a percentage — for example, 80.73.

VARIABLE TYPE

The Variable Type column offers more information about the kind of variables in the dataset. Like the Variable Format column, it indicates whether variables are quantitative or alphanumeric. More precisely, it distinguishes between two different kinds of quantitative variables: numbers and percentages. Datasets typically don't use the percentage sign (%) in their quantitative variables, because it can confuse data analysis software.

Many data analysts are familiar with these data formats, so you don't have to commit many of these details to memory.