Excel Advice: Ignore at your own peril

Prepare your metadata spreadsheet ahead of time, before you collect any data. (You can always add columns for unexpected variables later)

Don't have more than one table in an Excel spreadsheet

Don't use multiple Excel tabs!

Each experimental unit (e.g., a sequencing library) gets a unique identifier and ONE single row in your spreadsheet.

Each variable (e.g., fwd_primer, amplicon_region, nanodrop_ul, latitude, longitude, etc) gets ONE single column. No sharing columns between 2-3 variables (e.g., “SSU_richness” is actually 2 variables: “amplicon_region” and “richness”).

This is not great:

library_id	SSU_richness	ITS_richness
001-1-1	45	21
001-1-2	32	66

Problems:
- 4 sequencing libraries (exp. units) are represented by 2 rows (not unique identifiers)
- Columns 2 and 3 are hiding multiple variables

This is the way:

library_id	amplicon_region	richness
001-1-1_SSU	SSU	45
001-1-1_ITS	ITS	21
001-1-2_SSU	SSU	32
001-1-2_ITS	ITS	66

- unique ids for each library
- each variable gets its own column

Tidy data is “looong”

If you are entering dates in Excel, you are sabotaging yourself
- If you must, use "text format" as YYYY-MM-DD
- Better yet, separate column for YYYY, MM, and DD
- Dates are really 3 different variables anyway ;)

Only ONE value per cell

NEVER merge cells!

Excel Advice: Ignore at your own peril

Don't include summaries or figures in your raw data file

Every row is an observation, every column is a SINGLE variable

Highlighting cells to show information is horrible

Name variables and values CONSISTENTLY (Same spelling and capitalization)
- Better yet, use data validation feature
- Because you can’t trust yourself to never make a mistake
- Avoid spaces and special characters (“_” and “-” are fine)

Missing values? Just leave them blank!

Good data is rectangular. If yours isn't, there's probably a big problem.

Save the final data as a plain text file (csv, or tsv)

Consider making a data dictionary for each project that explains variables, etc.
- This should be rectangular and saved as plain text too!

Create a README.txt file for each project that explains what files are, how study was conducted, and anything else a stranger might need to know 3 years from now in order to analyze your data