

Best Practices for Working with Data in Excel

Chantal Ripp, Research Librarian (Data)
University of Ottawa

Best Practice	Additional Guidance
Organize your data <ul style="list-style-type: none">• Implement a well-organized, logical and efficient folder structure• Keep a raw copy of your data in a separate folder• Use file naming practices and version control• Consider using a cloud solution like Microsoft 365, which enables collaboration and build in version history	File naming and organization of data (uOttawa Research Data Management) Microsoft 365 (uOttawa's Information Technology)
Format your data / Prepare your data <ul style="list-style-type: none">• Freeze the header row• Give your worksheet and dataset meaningful names• Avoid merged cells• Tip: If your dataset doesn't come with a sequential uniqueID, consider adding an index column for sorting & referencing	Ways to format a worksheet - Excel (Microsoft Support)
Data quality control <ul style="list-style-type: none">• Each cell should contain just one piece of information• Use build in features to perform basic quality control on your data (ie Data Validation – drop-down menus; restrict cell formats; etc)• Turn formulas into static values after use, if formulas are no longer required• Fill in all cells. Use some common code for missing data (e.g. NA).• Tip: Distinguish columns you add (e.g. by color) to facilitate your workflow and identify what to track in your documentation/change log. This should only be to facilitate documenting your workflow and not highlight values in your data which can be lost if you convert your data to other file formats (e.g. CSV)	Guidelines for organizing and formatting data on a worksheet (Microsoft Support)
Document the steps you take <ul style="list-style-type: none">• Document the steps you took (e.g. Change log - who did the edits, when, what assumptions you made, what calculations you performed, what formulas you used etc)• Create a data dictionary (e.g. Define what each column holds - units, whether it's been normalized, how it's been calculated etc).	OSF Support. How to Make a Data Dictionary . CCO: Public Domain.

Further Readings

Karl W. Broman & Kara H. Woo (2018) Data Organization in Spreadsheets, *The American Statistician*, 72:1, 2-10, DOI: [10.1080/00031305.2017.1375989](https://doi.org/10.1080/00031305.2017.1375989)

Christie Bahlai & Aleksandra Pawlik (2015). Common mistakes by spreadsheet users. *Data Carpentry*. <https://datacarpentry.org/2015-03-09-ISI-CODATA/lessons/excel/ecology-examples/02-common-mistakes.html>