F1000

beFAIR beOpen

Spreadsheets

Spreadsheets are commonly used for data entry, organisation, analysis, and visualisation. By following best practices when using spreadsheets, you help ensure your data is **interoperable** and **reusable** for both humans and machines in the future.

☑ DO

- Keep your raw data raw; calculations and analyses should be done in a copy of the file

- Put variables in columns and observations in rows

- Give each column a descriptive heading that does not include spaces, numbers, or special characters

- Differentiate between zero and null values

- Validate your data

- Keep a separate txt file with a title and a legend describing your dataset, and outlining any steps you take to tidy your data

- Use a version control system and back up your files

- Export each data file in an open non-proprietary format such as CSV or TAB, with a name that appropriately reflects the content of that file

- Check your data thoroughly. Your data should receive the same care as your publications

☒ DO NOT

- Put more than 1 piece of information in a cell

- Use colour coding, embedded charts, comments or tables – your spreadsheet is not a lab book

- Include special (i.e. non alphanumeric) characters within the spreadsheet, including commas

- Use merged or blank cells

- Create multiple worksheets within a spreadsheet

F1000

beFAIR beOpen

Metadata

Each spreadsheet should be accompanied by a data dictionary. A data dictionary is a separate file where each variable is defined, including units and ranges, and often includes other useful information for interpreting the dataset. By helping others (and your future self!) better understand your data, a data dictionary supports **reuse** and reproducibility.

Toolbox

→ **Data Curator**

→ **Open Refine**

→ **Good Tables**

→ **Data Carpentry**

→ **OSF Guides: How to Make a Data Dictionary**

→ **Data Organisation in Spreadsheets**

Caution!

Different versions of spreadsheet software may handle data differently. Be especially cautious where your data contains dates or genes.