KIND Community Standards: Excel

AUTHOR

Brendan Clarke, NHS Education for Scotland, brendan.clarke2@nhs.scot

PUBLISHED

06/06/2024

Excel workbooks are frequently troublesome. This document contains a simple check-list to avoid the most serious problems found in Excel files in health and care. The aim is to be:

- universal, to provide effective guidance across health, social care, and housing
- simple and achievable, to keep to the most important guidance only
- realistic and grounded in health, social care, and housing practices

This is version 1, and this document will be fully reviewed in January 2025 by the KIND network community.

Basic standard for Excel workbooks

- Make your data tidy
    - One value per cell
    - One observation per row
    - One variable per column
- Avoid repetition
    - consider restructuring into several tables to avoid repeating data
- Use tables to contain data wherever possible
- Use meaningful sheet, table, and column names (not `Sheet 1`, `Sheet 2`...)
- Use number formatting as appropriate
    - Dates and percentages are especially important to format correctly
    - Text format for CHI numbers, rather than the initial '
- Avoid using colour-based formatting which can easily breach accessibility guidance
    - If conditional format is required, prefer using more accessible formats, such as shape icons sets

Enhanced standard for Excel workbooks

Follow the basic standards above, plus:

- Consider using data validation to make data consistent, especially if you don't really want free text
- Add a documentation / readme worksheet. You might include:
    - a description of the data
    - information about ownership/origin/dates/data sources
    - a key explaining abbreviations
    - explanation of any complex formulas

- advice about correct and consistent method of naming different versions of the workbook
- Consider using lambda formulas to avoid repetition of complex formulas
- Prefer structured referencing and naming to the use of complex A1 referencing
- Consider loading data using Power Query, especially if the workbook is designed to be updated