

# Spark Joy in Your Data

## A KonMari-Inspired Guide to Data Cleaning

This cheat sheet summarizes the key steps and Excel techniques covered in our workshop. It's designed to help you during the hands-on portion and serve as a quick reference afterward.

### KonMari for Data

Step	Action	Why It Matters
<b>0. Backup Your Data</b>	Always backup your data before you do any cleaning.	If something goes wrong, you can go back to the original data.
<b>1. Commit to Tidying</b>	Greet your data. Understand its source & purpose.	Recognize data cleaning as a core part of analysis—not just a chore.
<b>2. Imagine Your Ideal Dataset</b>	Visualize a clean, organized spreadsheet.	Stay motivated to remove the unnecessary and fix what's broken.
<b>3. Finish Discarding First</b>	Remove duplicates, irrelevant rows or columns.	Free up your spreadsheet (and your mind) from what doesn't "spark joy."
<b>4. Tidy by Category</b>	Fix missing data, split/merge columns, standardize categories.	Group similar tasks (all missing data, all name columns) for efficiency.
<b>5. Everything Has a Home</b>	Use logical folder structures, labeled tabs, & data dictionaries.	Prevent confusion & wasted time searching for files/columns.
<b>6. Ask If It Sparks Joy</b>	Spot-check with pivot tables, charts, or filters.	Ensure the final dataset is correct, consistent, and ready for analysis.

### Excel Tools

Name	Location
Find & Replace	Home → Find & Select → Replace
Remove Duplicates	Data → Remove Duplicates
Hide/Delete Column/Row	Right click/Control click for menu
Data Formatting	Right click/Control click column for menu

## Handy Excel Functions

Name	What it Does	Base formula
TRIM	Removes leading, trailing, and extra spaces from text.	=TRIM(text)
CLEAN	Removes non-printable or control characters from text.	=CLEAN(text)
VALUE	Converts a text string that looks like a number into a true numeric value.	=VALUE(text)
TEXT	Converts a numeric value into text, applying the specified format (e.g., currency, dates).	=TEXT(value, format_text)
LOWER	Converts all letters in the text string to lowercase.	=LOWER(text)
UPPER	Converts all letters in the text string to uppercase.	=UPPER(text)
PROPER	Capitalizes the first letter of each word in the text string.	=PROPER(text)
ROUND	Rounds a number to a specified number of decimal places.	=ROUND(number, num_digits)
ROUNDUP	Rounds a number up (away from zero) to the specified number of decimal places.	=ROUNDUP(number, num_digits)
ROUNDDOWN	Rounds a number down (toward zero) to the specified number of decimal places.	=ROUNDDOWN(number, num_digits)
CONCAT/&	Joins multiple text strings into a single string (similar to `&` operator).	=CONCAT(text1, [text2], ...)
LEFT	Extracts a specified number of characters from the start (left) of a text string.	=LEFT(text, [num_chars])
RIGHT	Extracts a specified number of characters from the end (right) of a text string.	=RIGHT(text, [num_chars])
MID	Extracts a specified number of characters from the middle of a text string.	=MID(text, start_num, num_chars)
FIND	Finds the position of a substring in text (case-sensitive).	=FIND(find_text, within_text, [start_num])
SEARCH	Finds the position of a substring in text (not case-sensitive).	=SEARCH(find_text, within_text, [start_num])
LEN	Returns the number of characters in a text string.	=LEN(text)
VLOOKUP	Looks up a value in the leftmost column of a table and returns a value from a specified column.	=VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])
XLOOKUP	Looks up a value in a range or table (in any column) and returns a corresponding result.	=XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])