

Using Excel to Organize Qualitative Data for Analysis

A Step by Step Guide

Formatting

1. Open a new Workbook in Excel.
2. Use the “Page Setup” function under “File” to set left and right margins to “0” and top and bottom margins to “0.25”.
3. Use the “Page Setup” function to reformat the page to “Landscape” orientation.

Template Development

4. Enter a title at the top of the page.
5. If desired, use the “Header/Footer” function under “View” to add other parameters relevant to your study (date, organization, etc.).
6. Leave a line blank to enter the unique question for each worksheet.
7. Starting at the left, label a column for participant identification code (PID).
8. Label columns for each face sheet code.
9. Label one or more columns for analytical coding.
10. Adjust the width of each of the coding columns to minimize the space they occupy on the worksheet. Make sure column width is adequate to accommodate the widest anticipated entry.
11. Create a wide column with the remaining width of the worksheet. Don’t let it overflow onto the next page. Label it “Responses.”
12. Use the “Copy” function by right-clicking the tab at the bottom of the worksheet to make a copy of the template for each question included in your study. Do this by clicking “Copy or Move” and then clicking on the check box in front of “Copy.” A new worksheet will appear that looks exactly like the one you just created. Do this again and again until you have one worksheet for every question in your study.
13. For each sheet you just created, insert one of your questions in the blank space you left at the top of each worksheet. (If you forgot to leave a blank line, don’t worry . . . just use the “Insert” button to add another row where you want it.) Now go to the tab at the bottom of each question worksheet, right click on it and select “Rename.” Give the tab

a name to match the Question at the top. I usually just use the number of the question. When you're done you should see all of your tabs at the bottom with the labels you gave them and be able to click on any one of them to take you to the associated page.

Data Entry

14. The goal is to transfer every response to each question to a single cell on the spreadsheet. The form of data you will be transferring over and your proficiency in importing data will determine your approach. (I understand there is a way to configure Word data using tab-delineators to make it importable into Excel in one fell swoop. I admit I tried it once and gave up. If you're proficient in this way or know someone who is that would be ideal.) For the rest of us here are some other options (The good news is that the manual approach gives us one more pass through the data.):
 - Transcript-based:
 - Copy and paste each response into its own cell on the spreadsheet.
 - Use voice-recognition software to speak each response into a separate cell.
 - Recording (digital or tape) based:
 - Transcribe responses directly into spreadsheet.
 - Note-based:
 - Use voice-recognition software to speak each response into a separate cell.
 - Type or copy notes into cells on spreadsheet.
15. Simultaneously add identification codes and face sheet codes to each entry. For identification codes, I assign each study participant a unique number based on their position in the study. If they are focus group participant number two in group number five, for example, their unique identifier would be "5.2" (Note: Identification and face sheet codes must be replicated for each response entry made by individual study participants.)
16. Separate long responses into "meaning units," i.e., bracket multivalent response entries into separate codable units consistent with study intent. (Make sure that bracketing occurs at normal and logical breaks in the response rather than at places of convenience.) Enter each meaning unit on a separate line and attach corresponding identification and face sheet codes to each meaning unit.
17. Repeat steps 14 through 16 above until all of the ID, face sheet, and response data is entered.
18. For ease in working with the data (seeing it all on one screen) and printing out in usable format, use the "Wrap Text" function to wrap the text in each of the response cells. Do this by highlighting all of the response cells and clicking on "Format," "Cells," "Alignment," and "Wrap Text."
19. If you had not cleaned data prior to entering it, now would be a good time to go back and do it. Otherwise, you're ready for coding.

Using Excel to Code, Categorize and Compare Qualitative Data

A Step by Step Guide

Data Familiarity

1. Print a copy of each spreadsheet in your project Workbook and read through the data.
2. Read through again, this time with pencil in hand. Make notes to yourself in the margins (maybe the beginnings of some patterns or even large coding categories).
3. Go back through the spreadsheets, but this time with a large (4 X 6) sticky pad and start jotting down what you believe might be some of the major coding categories. If you have a co-researcher ask them to go through the same process in parallel.

Refinement

4. Since the categories emerge from the data in an inductive analysis process, the only way to test coding categories is to try applying them to the data. When you think you might have a fairly comprehensive list of mutually exclusive categories on your sticky pad, assign each one a letter code. I like to use the first letter of the word or phrase I used to name the category until I run into alphabetical duplicates. Then I go to the second letter of the word or phrase. (Please develop a system that works for you.)

Armed with your set of letter codes, you're ready to take pencil to spreadsheet and assign one of your codes to each entry. You'll soon start to find responses that:

- Don't fit in any category
 - Fit in more than one category
 - Are a major subcategory of one of your categories
 - Generate a whole new category
 - Are the only one of their kind
5. Make adjustments in coding categories to accommodate your new insights. Each time you move through the data, category criteria should become less vague and more precise. Try to be as open, flexible, and unattached as you can. If you're working with a co-researcher now might be a good time to compare notes and reconcile categories (if possible) and develop clear inclusion criteria for each category.
 6. Now it's time to categorize responses on your spreadsheet. Once you have coded each response, highlight every column on the spreadsheet including the column headings. Then click on "Data" at the top of the sheet and then "Sort." You'll want to sort the data by the codes you just entered in the coding column so choose that column heading from the drop-down menu. Choose "Ascending" to sort from "A" to "Z." Then click "OK." Walla! All of your data will be automatically grouped by the codes

you just assigned. Now you will have a chance to look at each category of data and decide if there's convergence or not. Again, you may need to add, delete, change, subdivide, or collapse categories. Note: resist trying to fit data into established categories for convenience sake. After you adjust codes, "Sort" again and see what you think.

7. Proceed with this constant continuous refinement process until you find a reasonable fit between data and categories. You'll know when you're there. (And you'll appreciate all of your hard work when you sit down to write the report.)

Making Comparisons

8. If you faithfully entered face sheet and identification codes with each response, then Excel's "Sort" function will enable you to make comparisons among subsets of your data. For example, if you entered the gender of each of the participant associated with each response then you will be able to sort responses by gender to compare responses of men to those of women for any particular question. You could do a sort for any of the face sheet variables you entered and make comparisons. Like a kaleidoscope, the picture changes each time you re-sort.
9. The "Sort" function also enables you to cross-tabulate responses in multi-part questions. For instance, you may ask a group of teachers with which you are exploring overregulation in the classroom a question like this:

"On a scale of 1 to 10, to what extent do you feel burdened, overwhelmed, or your hands are tied behind back by outside rules and policies? Give a specific example."

You might code responses to the first part of the question (on a scale from 1 to 10) in column D on your Excel spreadsheet and responses to the second half in column E. Once you sort by column D you can do cross comparisons with the data in column E. You may even generate a version of a two-by-two table to display relationships you detect in the data.

Inter-question cross-tabulations can also be facilitated with Excel using the "Find" function. If a segment of study participants answer a question in one direction and you want to know how that might relate to their responses on a different question you can use their unique identifiers to "find" their responses to the alternate question and compare them.

10. Final note: Excel has much more capability than I have tapped into for organizing qualitative data. I understand there are functions like "If/Then" "Lookup," and "Concatenate" that can be used with text. There is also an "Auto Filter" and "Advanced Filter" which sound like they could be useful for sifting through data and performing constant comparative analyses. I encourage you to explore Excel's capabilities further and let me know what you learn.