

Spotlight Session: Analysing Answers to Open-Ended Questions from Surveys

Excel format for data preparation:

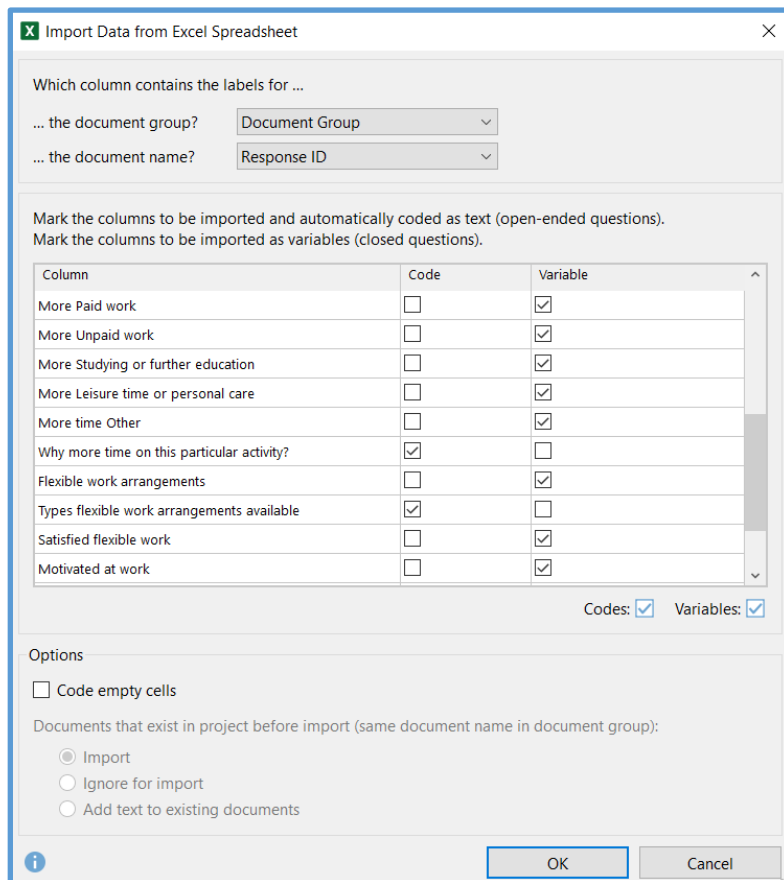
	A	B	M	N	O	P	Q	R	S	T
1	Document Gr	Response ID	More Unpaid	More Studyin	More Leisure	More time Ot	Why more time on this particular activity?	Flexible work	Types flexible work arrangements available	Satisfied flexi
5	Survey	RESP04	No	No	Yes		Better work/life balance.	No		Neither
6	Survey	RESP05	Yes	No	Yes		I would love to do voluntary work again but right now I just don't have the time to do it.	Yes	I can work from home if I want to	Satisfied
7	Survey	RESP06	No	Yes	No			Yes	Flexible working hours and option to work from home from time to time.	Neither
8	Survey	RESP07	No	Yes	No			I do not know		Neither
9	Survey	RESP08	Yes	No	No		I would like to spent more time on my own artistic work in order to further develop my ideas.	Yes	The possibility to rearrange working hours if need be.	Satisfied
10	Survey	RESP09	No	No	No		It would be nice to earn more money, but as a student in Germany, you are not allowed to work more than 20 hours per week.	No		Very dissatisf
11	Survey	RESP10	No	Yes	No			No		Satisfied
12	Survey	RESP11	No	No	Yes		more time to do sports would be great	Yes	Home Office	Satisfied
13	Survey	RESP12	Yes	Yes	No		Because I like it	No		Dissatisfied

Notes:

- Insert Column A which controls the grouping of the texts in MAXQDA's "Document System" window. Enter the same phrase for all rows if you want all texts to be in a single group, here "Survey".
- Column B contains the labels for the Document System in MAXQDA, it should have a unique identifier for each respondent. This is probably an output from your survey program.
- The labels at the top of each column (Row 1) will appear exactly like this in MAXQDA, so use labels that will be meaningful in that different context. Remember that shorter labels are better than long ones.
- You may edit these labels but you should not edit the response cells.
- If you use Survey Monkey for data collection, there is a special routine in MAXQDA to collect the data from your Survey Monkey account which bypasses this step. Full instructions are available in the MAXQDA Online Manual.

- When your data is ready, close the Excel workbook, open the analysis project in MAXQDA and use the option **Import > Survey Data > Import Data from Excel Spreadsheet**. You will need to navigate to the folder with the Excel workbook and select it there to start the process.

Dialog during import routine in MAXQDA:



Column	Code	Variable
More Paid work	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Unpaid work	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Studying or further education	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Leisure time or personal care	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More time Other	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Why more time on this particular activity?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Flexible work arrangements	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Types flexible work arrangements available	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Satisfied flexible work	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Motivated at work	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Codes: ☒ Variables: ☒

Options

☐ Code empty cells

Documents that exist in project before import (same document name in document group):

☒ Import
☐ Ignore for import
☐ Add text to existing documents

OK Cancel

Notes:

- This is where you separate variables from qualitative texts. The texts to be analysed are marked in the “Code” column, the short or quantitative data go into the “Variable” column.
- At the top of this dialog, select the header labels from columns A and B in the Excel workbook from the pull-down menus.
- I recommend leaving the option “Code empty cells” box blank (as above) because then you get a simple count of all collected responses for each question and non-responses are ignored.
- A similar version of this dialog appears as part of the special routine for importing data from Survey Monkey.

- Formatting variables during import routine:

Import variables

Corresponding data field

Source	Target	Preview data type
<input checked="" type="checkbox"/> Age	New variable: Text	30-39
<input checked="" type="checkbox"/> Relationship	New variable: Text	Separated/divorced
<input checked="" type="checkbox"/> Children	New variable: Text	Yes
<input checked="" type="checkbox"/> Number children	New variable: Integer	1
<input checked="" type="checkbox"/> Highest education level	New variable: Text	Master degree
<input checked="" type="checkbox"/> Paid work time	New variable: Text	30-39
<input checked="" type="checkbox"/> Studying or further education ...	New variable: Text	9 or less
<input checked="" type="checkbox"/> Unpaid work time	New variable: Text	9 or less
<input checked="" type="checkbox"/> Leisure time / personal care	New variable: Text	10-19
<input checked="" type="checkbox"/> More Paid work	New variable: Text	No
<input checked="" type="checkbox"/> More Unpaid work	New variable: Text	Yes
<input checked="" type="checkbox"/> More Studying or further ...	New variable: Text	No
<input checked="" type="checkbox"/> More Leisure time or personal ...	New variable: Text	Yes
<input checked="" type="checkbox"/> More time Other	New variable: Text	Time with my family

Import **Cancel**

Notes:

- For each of the closed questions, which are being imported as variables, you should now check that MAXQDA has correctly interpreted the variable type. See the middle column above ("Target") which are mainly shown as "Text" with just one "Integer" variable.
- This is the point at which you may regret the way you phrased some of your questions as the data may be harder to manipulate afterwards.

Import Documents from Excel Spreadsheet

Imported texts: 55
Imported codes: 2
Imported variables: 20
Ignored documents: 0

Copy **Close**

When the import has completed take the opportunity to copy this summary and paste it into a Free Memo or the project Logbook as a record of this process.

- Checking data after completing the import:

The screenshot displays the MAXQDA software interface. At the top, the 'Document Browser' shows a list of documents (RESP41, RESP42, RESP43, RESP44) with RESP43 selected. Below it, the 'Code System' window shows a list of codes with their respective counts: 'Why more time on this particular acti...' (48), 'Types flexible work arrangements av...' (22), and 'Sets' (0). The main window shows the 'Document Variables' table, which lists various demographic and contextual variables for each document. The 'Data Editor' window at the bottom shows a table of data for all documents, with RESP43 highlighted in green.

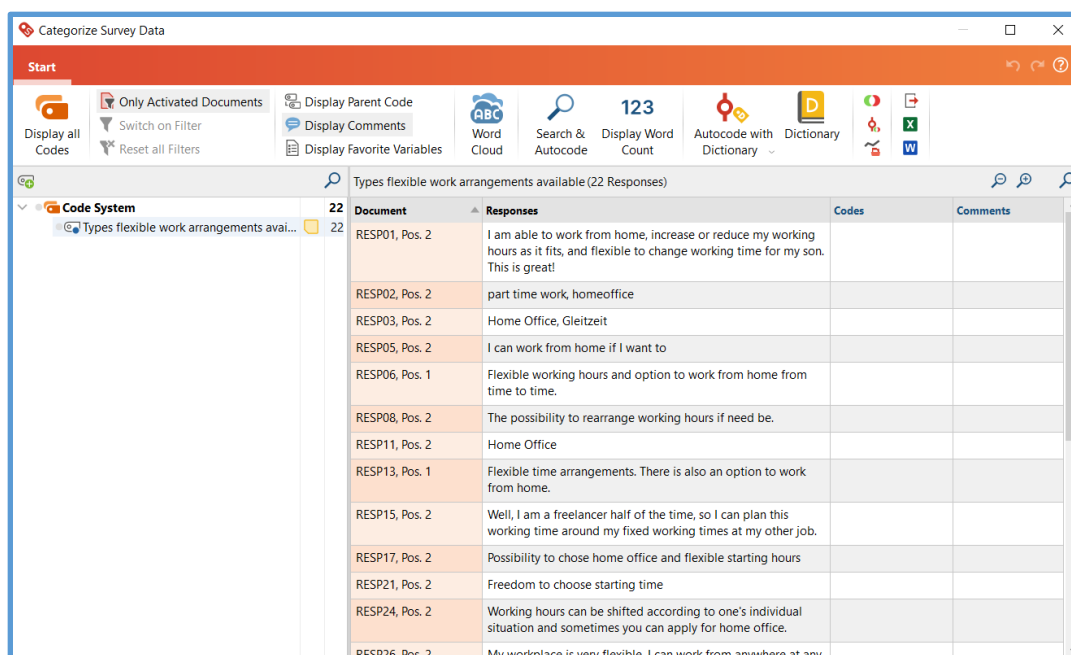
Group	Document name	Age	Relationship	Number children	Children	Highest education	Paid work time	Studying or further education	Unpaid work time
	RESP42	50-59	Separated/divorced	3	Yes	Trade school or less	40-49	9 or less	20-29
	RESP43	30-39	Separated/divorced	1	Yes	Trade school or less	20-29	9 or less	10-19
	RESP44	30-39	In partnership/...	2	Yes	Bachelor degree	30-39	9 or less	20-29
	RESP45	50-59	In partnership/...	3	Yes	Master degree	10-19	9 or less	20-29
	RESP46	40-49	Separated/divorced	3	Yes	Master degree	30-39	9 or less	10-19
	RESP47	60 or older	In partnership/...	1	Yes	PhD	40-49	9 or less	20-29

Notes:

- I have selected one document (RESP43) in the Document System to display its responses to both open questions in the Document Browser.
- I have also opened the **Variables > Data Editor for Document Variables** and scrolled that to show the same case in the lower window above.
- The Code System is showing the code names for the 2 open questions and these can also be seen applied to the responses in the Document System.
- At this stage it is a good idea to paste the full question text into the Code Memo for each of the open questions for easy reference during analysis.
- The numbers beside each code in the Code System show how many responses were found for each question (this would not be correct if you ticked "Code empty cells" during the import).
- Once you have checked that all of the data you expected has been imported successfully, you are ready to begin the analysis.

Analysing the responses to one question:

- Activate all of the Survey documents in the Document System (or use a function like **Mixed Methods > Activate Documents by Variables** if you want to work with a specific subset of them).
- Start the analysis with the function **Analysis > Categorize Survey Data**. When asked, drag in the code for an open question from the Code System window, accept the default tick for "Only activated documents", and click on "OK".



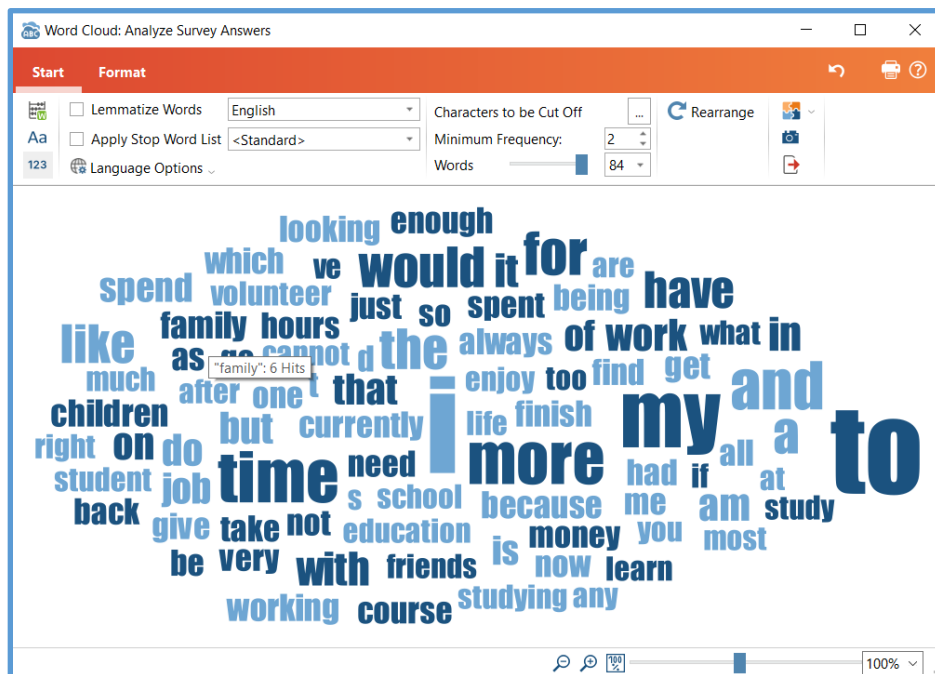
- This is the new version of this function in MAXQDA 2022. We now have access to a range of semi-automation tools like Search & Autocode within the one window so you should be able to do almost all of your analysis here.
- Note that all 22 responses to this question can be read by scrolling down.
- New codes for the themes you identify can be created and should be made sub-codes of the question code in the left panel.
- As you apply further codes to individual (or multiple) responses they will be shown in the Codes column to the right of the responses.
- You can also display selected variable values in columns further to the right to inform your coding decisions (use the command "Display Favorite Variables" from the menu bar inside this window).
- While this window is open the rest of your project is frozen and cannot be changed. All changes you make here will be in the project when you close this window.

You now have decisions to make over your approach to the analysis. Mainly, these concern the extent to which you wish to use automation to speed up your work, although this probably comes with a trade-off in terms of accuracy. The automation tools available within the Categorize Survey Data window are the following:

- Word Cloud and word frequency tables
- Search and autocode with new codes
- Autocode using Dictionaries (essentially multiple search and autocodes carried out simultaneously)
- Sentiment analysis plus autocoding its results

We will briefly review the procedures for each of these tools. It is unlikely that one project would use all of them. It is advisable to check your codings after using automation tools by selecting each code, one at a time, in the left hand panel and reading the full set of responses coded that way in the main panel, to make sure that they all include a common theme or concept.

Word Cloud

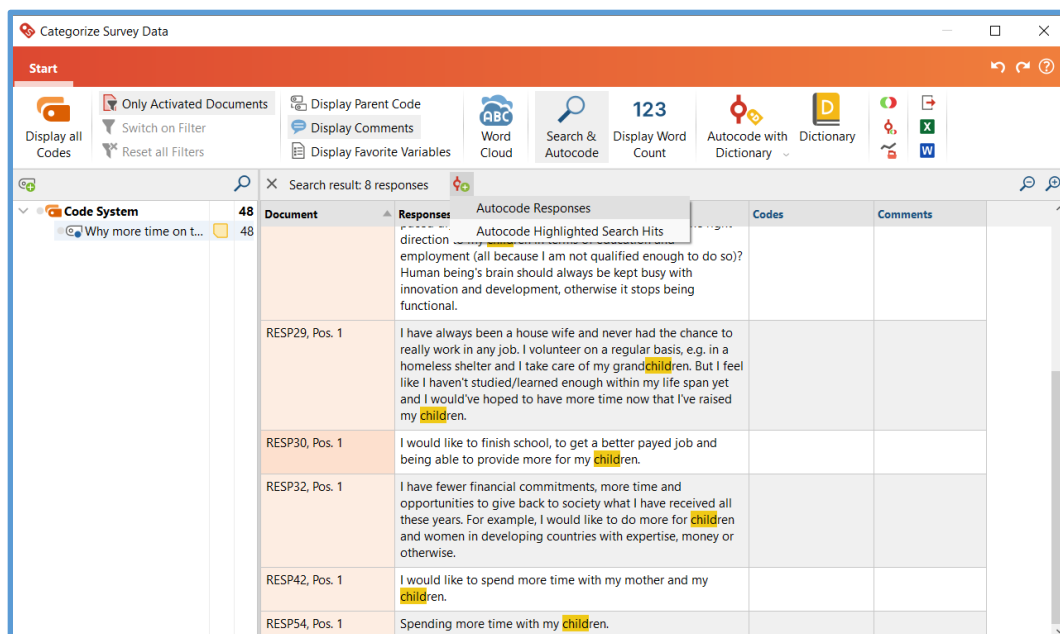


Here is a Word Cloud generated on the full set of responses to the question "Why spend more time on this particular activity" from within the Categorize Survey Data window. By pointing the cursor at the word "family" I can see a pop-up dialog showing it occurs 6 times in this data.

This is a good way to get a quick overview of the keywords that occur frequently in the responses to this question.

- A double-click on a word in the cloud display opens a search window showing all of its occurrences, and from that window you can autocode those hits to a new or existing code.
- However in this approach new codes are added above the question code (ie not as sub-codes of it, so they should be dragged down to the list below the question code), and your choices of how much text to code are based on words, sentences or paragraphs only.
- A click on the icon at the top left corner of the toolbar converts the cloud to a word frequency table, which some may find easier to interpret, and from which you should be able to search and autocode in the same way.
- With the Format menu you have considerable flexibility to change the layout and colour scheme for the cloud .
- At the right end of the Start menu you will find options to save copies of your cloud for inclusion in your final report.

Search and Autocode



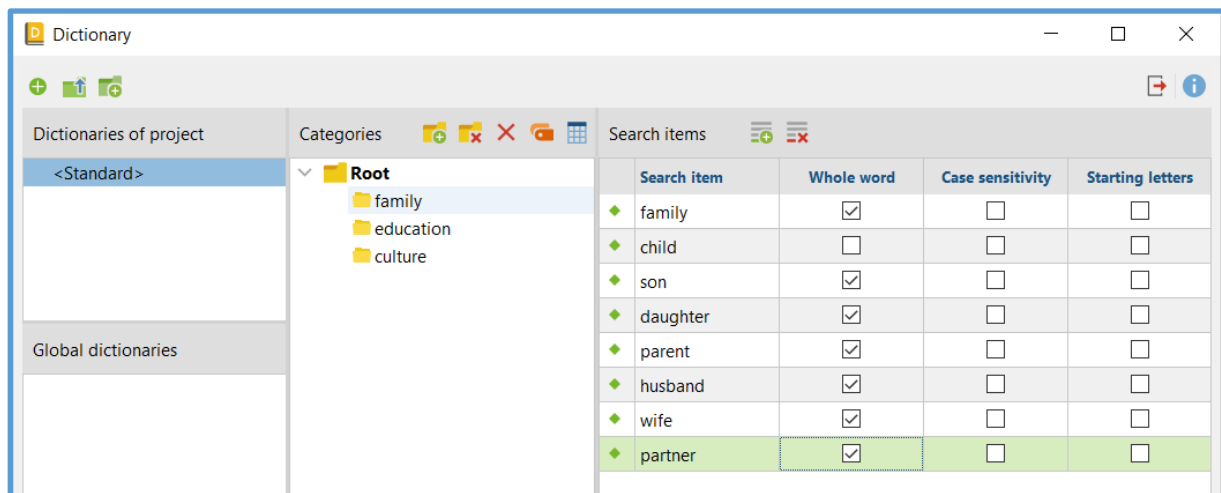
Here I have used the Search & Autocode function to find words beginning with "child". The hits are highlighted in the main panel (not in the normal search results window) where only responses containing the search term are shown. The Autocode command is the "Code with a new code" icon just below the menu bar

and here it is showing the 2 possibilities for the extent of the coding - whole responses or just the highlighted search terms. When you have made this selection you will see the New Code dialog as usual, where clicking on "OK" completes the process and the code is created and applied.

- Note that there is currently no option to apply an already existing code to the search hits in this window, but you could either add an existing code to the segments with manual coding at this point, or create another new code and then merge its contents with the one you wanted to use by dragging across to the code system panel.
- To return to the full set of responses, just click on the "X" beside "Search result: xx responses".
- Note that this is the first time we have been able to autocode entire responses (which may be a couple of words or several paragraphs) with one command.
- New codes created with this routine are placed correctly as subcodes of their question code.

Autocode with Dictionary

To use this facility you need a MAXQDA Plus license.



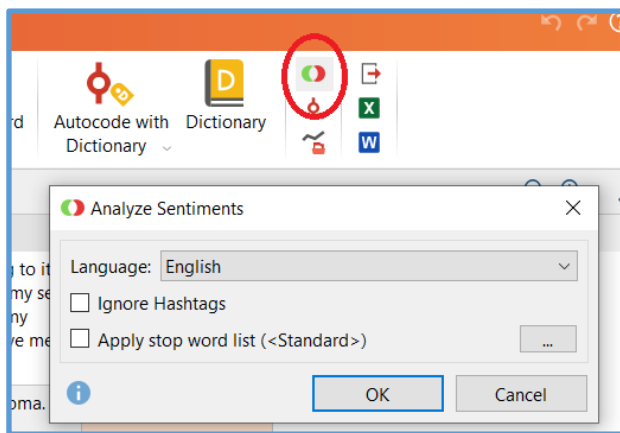
Here is a simple dictionary that can be applied to this data.

- The 3 Categories ("family", "education" & "culture") represent codes that will be added to the code system when this is used.
- For each category a set of keywords can be created as the "Search items" to the right, so here we are showing a set of words for "family".
- Each search term can have its own settings for "Whole word" etc. as shown in the columns to the right. ("children" and "grandchild" will also be hits).

- You can create multiple dictionaries and copy them across different projects. It would probably be inefficient to create a new dictionary for a single project and never use it again, so this approach assumes that there will be more surveys covering similar topics to make this worthwhile.
- Dictionaries can be created, saved and moved between projects with the functions on the left side of this window.
- You should check your dictionary categories and search items carefully before you click on the command "Autocode with Dictionary". With a context menu option (or a double-click) on any category you can temporarily "deactivate" it and then, when the autocode is run, that code need not be applied.
- The autocode will create a new sub-group of codes beneath the question code, headed "Autocode with Dictionary". This is useful so that you are reminded that all of this coding was automatically applied and should be checked before it is reported.

Sentiment Analysis & Autocode

The data in this example is not really appropriate for this approach, but I will use it nevertheless to show the mechanics of this function.



The icon is found just to the right of the Dictionary symbol.

Only German and English are available at present.

Stop word lists could have a big effect on the results.

Here is the setup dialog for Sentiment Analysis.

- Below you can see a screenshot of the display when these settings have been applied to the first open question in this small survey.
- In addition I have applied the "Display Word Count" function on the menu bar and sorted the rows into descending order on the word count, so that we can concentrate on the longer contributions (where sentiments may be detected more easily).

Docum...	Responses	Codes	Sentiment	+ Words	- Words	Difference	Words
RESP28, Pos. 1	One of my hobbies is sailing which is very time-consuming. I almost always have to take time off work to go sailing for a few days because it's not doable during the weekends. I have other hobbies but that is the one I enjoy most and I'm very unhappy with how often I currently get to do it		Neutral	8	7	1	50
RESP27, Pos. 1	I find it hard to work and study simultaneously. I do enjoy studying but my income solely depends on my student job as I receive no support from my parents which is stressful at times. If I had more time to study I could focus more on the things I'm learning	education	Neutral	11	9	2	51
RESP01, Pos. 1	I believe life has no meaning, we need to give meaning to it by ourselves. My personal time with my family and for my self is most important, second comes the time I spend for my volunteering works. These are the two activities that give me the most satisfaction.	family	Slightly Positive	15	4	11	49
RESP32, Pos. 1	I have fewer financial commitments, more time and opportunities to give back to society what I have received all these years. For example, I would like to do more for children and women in developing countries with expertise, money or otherwise.	family	Slightly Positive	9	3	6	41
RESP36, Pos. 1	Participating on an evening course would help me further my career. In order to qualify for my next promotion I will need additional training but I do not have any time to do so right now.	education	Neutral	6	7	1	36

- The sentiment result is expressed as a 5 point Likert type scale between Negative and Positive, with a central value of "Neutral". It is based on counts of positive and negative words from preset lists (in columns "+ Words" and "- Words" above).
- We can see here that respondents 01 and 32 appear as "Slightly Positive" while the others are "Neutral". A reading of the response texts shows that this is pretty subjective, although the other 3 cases in this screenshot do each have something more negative to say.
- With the right sort of data it might be useful to store the result of this analysis by clicking on the icon below the one we just used which is called "Autocode Responses with Sentiment". And a further step might be to subsequently (after closing this Categorize Survey Data function) convert that code to a Document Variable in order to investigate how it interacts with other codings.
- Hide the sentiment analysis columns by clicking again on the menu icon with which you started this function.

Work Manually

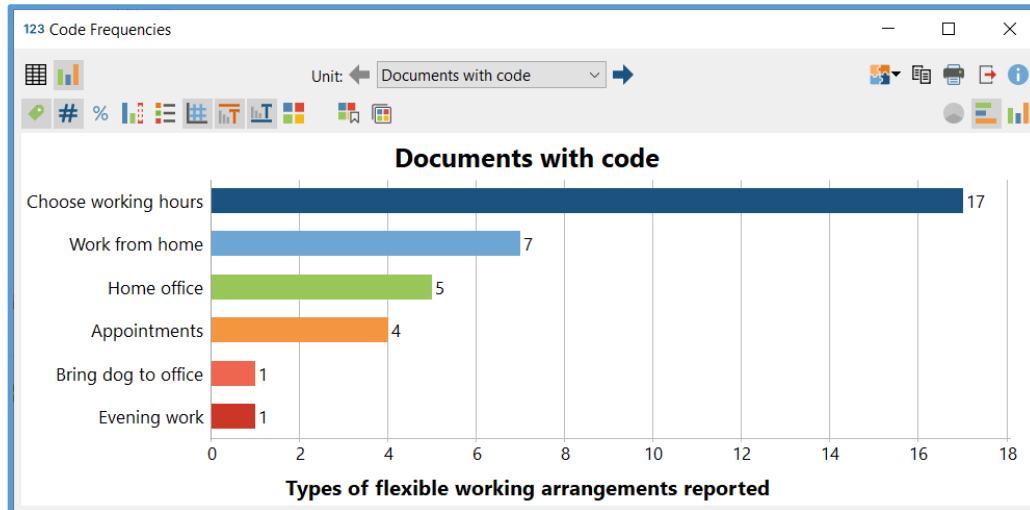
It is unlikely that these automation tools will complete the coding phase of a project on their own. It is always advisable to carry out checks on their accuracy and manually add, edit or remove codes that on a closer reading have not been applied correctly. Some respondents will use different words for a key concept and these may not be picked up in the word searches, others may use a keyword with a

different meaning, and in self-completion surveys there may be lots of mis-spelled words.

- An important facility (shared with the Smart Coding tool in MAXQDA) is that when you click on any code showing in the code system list in the left panel, the display will quickly change to show only the responses attached to that code. This facilitates quick checking for incorrect code applications as such segments should look out of place in the list.
- While checking for incorrect codings you should also look out for omissions by thinking about unusual words in the responses and checking the other codes listed in the "Codes" column to the right.
- Finally look for any responses which have not been coded to any analysis codes yet. An easy way to do this is to select the code for the question itself and then click in the header bar for the "Codes" column to the right of the responses to sort the rows by the contents of that column - any rows with no other codes attached will appear at the top of the table where you can think about further manual coding work.
- Codes can be applied manually by clicking on the response row to highlight it and then dragging that data onto the relevant code in the Code System list on the left. They can be removed from the code currently displayed with a context menu option in the middle panel. And incorrect alternative codes can be removed with a click on the red "X" beside the code name in the Codes column to the right.
- It may be useful to create a code for the ones that cannot be coded to any meaningful code (say "No clear answer" or "uncategorizable") and apply it to these.
- You may find it helpful to display the answers to any relevant closed questions (such as the one immediately preceding the question you are analyzing) by adding their variables to the "favorite variables" list (in **Variables > List of Document Variables**) and clicking the menu option for "Display Favorite Variables".

Reporting the results of your analysis:

When you are satisfied that you can go no further with creating and applying codes to the survey question responses it is time to consider how to report the results of your analysis. Close the Categorize Survey Data window to return to the main MAXQDA interface.



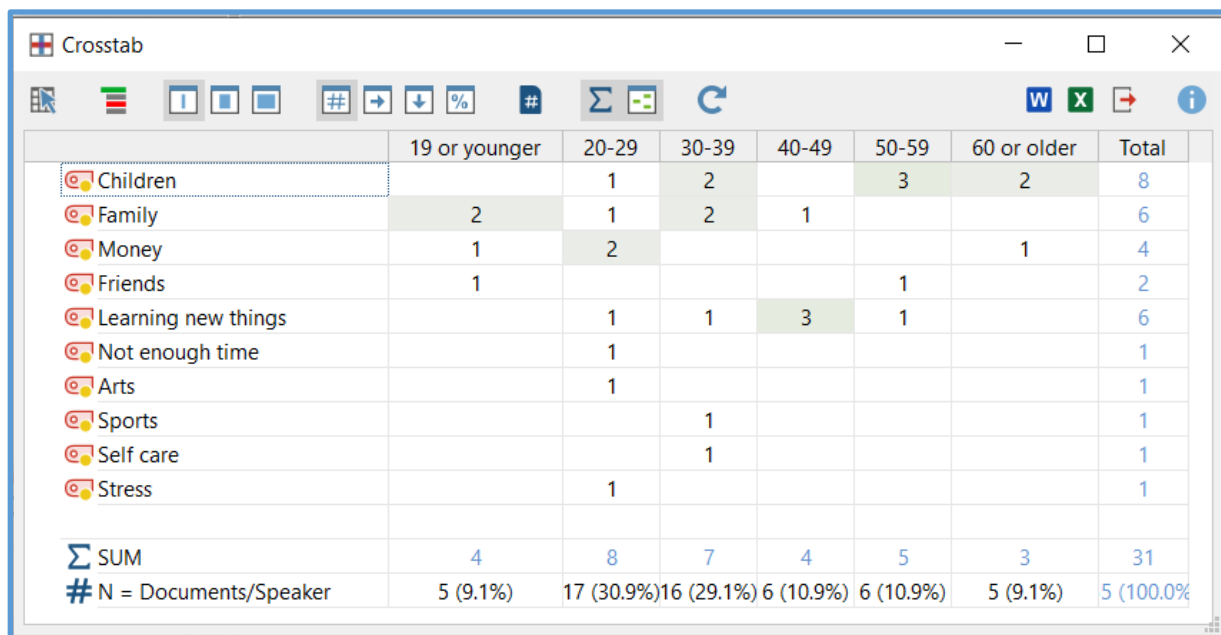
To obtain a chart like the one shown above, use the function **Analysis > Code Frequencies**.

- With all of your Survey responses and the set of analysis codes for one question activated, start the function.
- In the set up dialog use the option for "Insert activated codes" and the tick box for "Only activated documents".
- In the table display which opens first you can choose to order the rows by frequency or alphabetically.
- Toggle between Table and Chart view with the icons at the top left, and in Chart view you have a choice between a pie chart and horizontal or vertical bar charts.
- The choice of units (documents or segments with the code) probably makes no difference if you have generally coded whole responses in the analysis as these will give the same results.
- You can choose whether to show counts or percentages on the scale, and add a subtitle beneath the chart with the "Display text" icon (8th from the left) where I have added "Types of flexible working arrangements reported".

To report the frequencies for your closed questions, use the function **Variables > Document Variable Statistics**, from where you can create more tables and charts looking similar to the one shown above.

Combining data from closed and open questions:

You can combine the quantitative and qualitative elements of your survey by using MAXQDA's **Mixed Methods** tools. For example, using this data, we can create a **Crosstab** of the responses to the question about why they wanted to spend more time on certain activities analysed by the age-group of the respondents.



	19 or younger	20-29	30-39	40-49	50-59	60 or older	Total
Children		1	2		3	2	8
Family	2	1	2	1			6
Money	1	2				1	4
Friends	1				1		2
Learning new things		1	1	3	1		6
Not enough time		1					1
Arts		1					1
Sports			1				1
Self care			1				1
Stress		1					1
SUM	4	8	7	4	5	3	31
# N = Documents/Speaker	5 (9.1%)	17 (30.9%)	16 (29.1%)	6 (10.9%)	6 (10.9%)	5 (9.1%)	5 (100.0%)

- There is not really sufficient data in this example to justify any conclusions but in the above table it appears that the comments about learning new things are out of line with the numbers of respondents in each age group shown in the bottom row.
- Note that in the set up dialog for a Crosstab you can change the order of the columns in the table by dragging the formula rows around, so if the first attempt does not give you a useful sequence then use the "refresh" button (13th from the left) to return to that dialog and change it.
- This is where the full pay-off for using MAXQDA to analyse your survey data becomes apparent!

Graham Hughes (March 2022)

Reference:

Fielding J, Fielding N, and Hughes G: Opening up open-ended survey data using qualitative software: Quality and Quantity (2012)