

Mixed Methods for Survey Data Analysis

Data Preparation and Importing:

	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

Here is an example Excel worksheet after editing to prepare its data for import into a MAXQDA project.

- 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 by-passes 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.

Which column contains the labels for ...

... the document group?

... the document name?

Mark the columns to be imported and automatically coded as text (open-ended questions).
Mark the columns to be imported as variables (closed questions).

Column	Preview data type	Code	Variable
More Paid work	No	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Unpaid work	Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Studying or further education	No	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More Leisure time or personal care	Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>
More time Other	Time with my family	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Why more time on this particular activity?	I believe life has no meaning, we nee...	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Flexible work arrangements	Yes	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Types flexible work arrangements available	I am able to work from home, increas...	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Satisfied flexible work	Very satisfied	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Motivated at work	Often	<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

- Use the pull-down menus to select the column headings for Columns A & B (see above).

- I have scrolled down the list to show the 2 open questions in this data ticked in the "Code" column.

- I recommend leaving this unticked.

- This is where you separate variables from qualitative texts. The texts to be analysed have ticks in the "Code" column, the short or quantitative data go into the "Variable" column.
- MAXQDA tries to do this for you and most of its ticks will be correct, but you do need to check them all (this screenshot was taken before any editing was done).
- 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.

When you click the "OK" button the program moves on to setting the formats for the Variable data when they are imported into your project.

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 time	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

Import Cancel

- this is a pure number, so Integer is correct.

- a range of numbers like this has to be processed as text.

Again, MAXQDA offers its best guesses for the middle column, based on the data in the second row of the spreadsheet. You can edit these with the pull-down menu options for each row in this table. Only the items ticked as "Variable" in the previous dialog are involved in this stage. You may now regret the way you collected some of your data as "Text" is the only practical option for several numerically based fields.

Click on "Import" to set the function running.

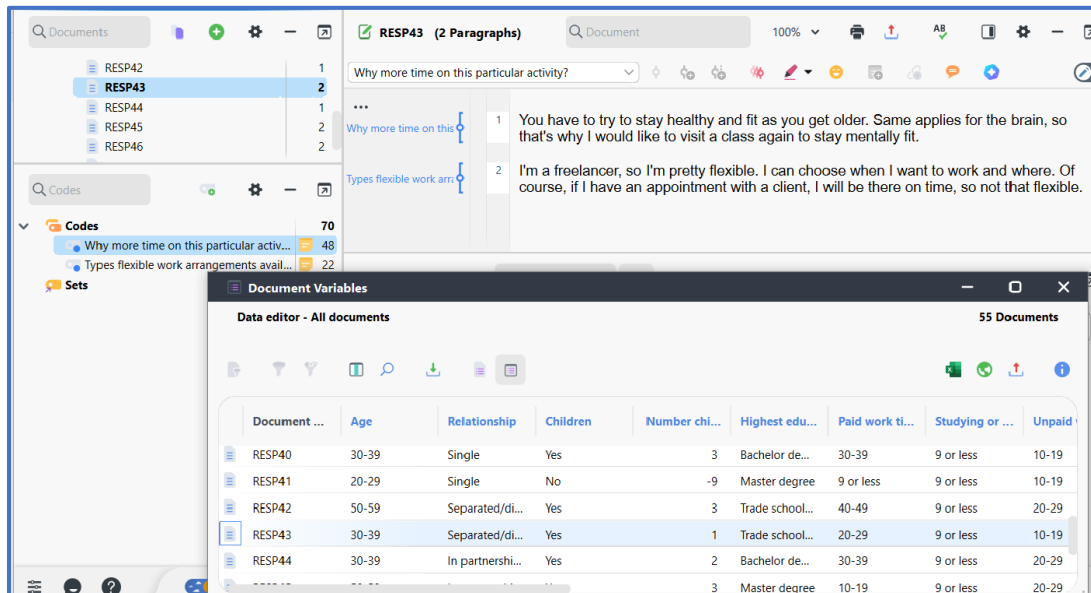
Import Documents from Excel Spreadsheet

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.

The final step in the importing stage is to check that what has arrived in your project is what you expected.



The screenshot displays the MAXQDA interface. On the left, the 'Documents' pane shows a list of documents (RESP42-RESP46) with 'RESP43' selected. Below it, the 'Codes' pane shows two codes: 'Why more time on this particular activ...' with a count of 48, and 'Types flexible work arrangements avail...' with a count of 22. The main window shows document 'RESP43 (2 Paragraphs)' with two paragraphs of text. The first paragraph is 'Why more time on this particular activity?' and the second is 'Types flexible work arrangements available...'. The 'Document Variables' window is open, showing a 'Data editor - All documents' table with 55 documents. The table has columns for Document ID, Age, Relationship, Children, Number children, Highest education, Paid work time, Studying or..., and Unpaid work time.

Document ...	Age	Relationship	Children	Number chi...	Highest edu...	Paid work ti...	Studying or ...	Unpaid
RESP40	30-39	Single	Yes	3	Bachelor de...	30-39	9 or less	10-19
RESP41	20-29	Single	No	-9	Master degree	9 or less	9 or less	10-19
RESP42	50-59	Separated/di...	Yes	3	Trade school...	40-49	9 or less	20-29
RESP43	30-39	Separated/di...	Yes	1	Trade school...	20-29	9 or less	10-19
RESP44	30-39	In partnershi...	Yes	2	Bachelor de...	30-39	9 or less	20-29
.....	3	Master degree	10-19	9 or less	20-29

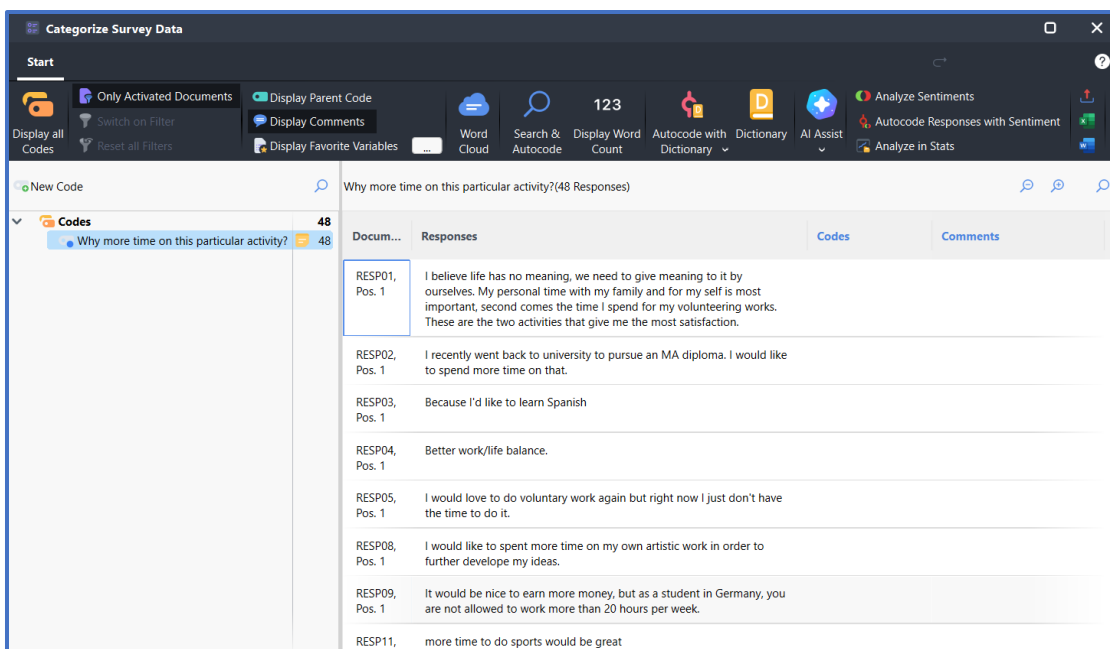
- I have selected one document (RESP43) in the Document System to display its responses to both open questions in the Document Browser.
- I have also opened the **Variables > Data Editor for Document Variables** and scrolled that to show the same case in the lower window above.
- The Code System is showing 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.

Coding the open question responses:

The next step is to apply codes to the text data that is stored in the Document System, ie the responses to your open-ended questions. You have a wide spectrum of ways to work with this data from fully manual (reading one at a time, deciding what codes are relevant, and applying those codes by drag and drop), through a variety of semi-automatic tools to asking AI Assist to help. This session will not look at these in great detail, there are other recorded webinars and sessions you can access for that, but we will just look at some of them briefly.

You should consider the analysis methodology you want to use, the consequences of the way you collected the responses, the objectives of your research questions, and the quality of the data you are using to inform your decisions about how to handle this task. Some experimental work with a small sub-sample of your data may be useful before you commit to your chosen approach, to make sure it will work.

A good place to start is the function **Analysis > Categorize Survey Data** .

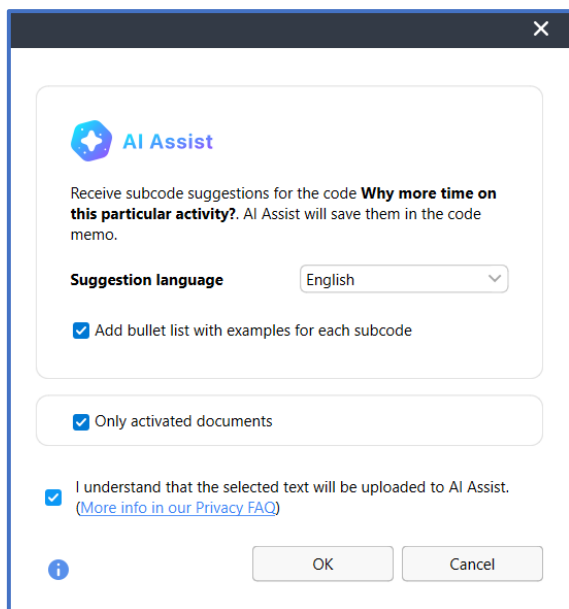


Docum...	Responses	Codes	Comments
RESP01, Pos. 1	I believe life has no meaning, we need to give meaning to it by ourselves. My personal time with my family and for my self is most important, second comes the time I spend for my volunteering works. These are the two activities that give me the most satisfaction.		
RESP02, Pos. 1	I recently went back to university to pursue an MA diploma. I would like to spend more time on that.		
RESP03, Pos. 1	Because I'd like to learn Spanish		
RESP04, Pos. 1	Better work/life balance.		
RESP05, Pos. 1	I would love to do voluntary work again but right now I just don't have the time to do it.		
RESP08, Pos. 1	I would like to spent more time on my own artistic work in order to further developpe my ideas.		
RESP09, Pos. 1	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.		
RESP11, Pos. 1	more time to do sports would be great		

I activated all the Survey documents and dragged the code for the first open question into the setup dialog to get the above screen. Here you can now work in many different ways with the convenience that you can easily view all the segments already attached to any particular code with a single click in the left section.

You can create codes manually in the left section - it is useful to make all the thematic codes derived from one question's responses as sub-codes of that question's code as this makes reporting the basic statistics easy. You can drag segments onto those thematic codes within this window, and also change or edit any of these decisions.

With the Word Cloud function in the toolbar, you have access to word frequency tools to start identifying common themes on that basis. With the Search & Autocode function you can partially automate the process and allocate codes to multiple responses which share the use of key words. With the Dictionary function you can take this a step further and store multiple search and autocode commands into a single step (this is useful if you plan to repeat a survey in successive waves or rounds of data collection). Or you can use AI Assist from this toolbar to summarise the responses or to suggest multiple subcodes for you. In all probability, you will use several of these options, including manual corrections of the segments which use unusual language or mislead the automated tools. We will look at the AI Assist functions which are the newest innovations in this area.



- The bullet list with examples will help you to apply those suggestions which you choose to adopt.


- The upload tick is required for this function to work. Please check the FAQs to understand the implications of this.

Below is an extract from the Code Memo, which is where the suggestions are put. You may find it useful to run the same request a couple more times, as AI Assist may come up with some slightly different suggestions, giving you more choice.

Why more time on this particular activity?

19/08/2024 15:39 - Graham Hughes
AI subcode suggestions

1. Childcare and Family Responsibilities
 - Spending more time with my children
 - Providing for my family, siblings, and parents
 - Caring for my grandchildren
2. Career Transition and Development
 - Pursuing higher education and training
 - Seeking a new job or career change
 - Applying for a promotion or taking on more responsibilities
3. Work-Life Balance
 - Reducing work hours or commute time
 - Seeking early retirement
 - Prioritizing personal hobbies and interests
4. Volunteering and Community Engagement
 - Dedicating time to volunteer work
 - Giving back to society and helping others
 - Participating in social projects
5. Personal Growth and Lifelong Learning
 - Continuing education and skill development
 - Exploring new hobbies and interests

Code summary 

- The numbered line is a suggested code name, followed by 3 content examples which may be useful for Search & Autocode processes.

- There were 7 suggestions from this query.

You can easily see several areas of overlap within these suggestions, so it would be difficult to apply all of them in this exact structure. However, you may find that several fit quite well with your main research questions and so it would be worth using those as the next step. For illustration, I will use the first two.

My approach is as follows:

- Create the selected new codes as sub-codes of the question code.
- Use copy & paste to transfer the content examples into their respective code memos.
- With the main question code selected (so that you have all of its responses in the window) explore Search & Autocode to attach the most common ideas to codes.
- With this kind of data it is generally best to apply the thematic codes to the whole response each time (rather than just keywords or sentences).
- Manually check and edit the results, including adding codes to relevant segments which used different words. Remember that some responses may be linked to several subcodes, but every response should be linked to at least one subcode.

Note that, when you use this autocode function, you cannot directly autocode into your existing codes as you have to autocode into a code with the name of the words you searched on. But then, within this function window, it is quick to use the context menus to move all of the newly coded segments from the temporary autocode to the desired code and delete the unwanted autocode.

You may find it useful to practise using these functions on some unimportant data (such as the example projects included with MAXQDA) before using them on your real data. Ideally, you want to be able to add, edit and remove codings while thinking more about the meanings in the data than the processes in the software.

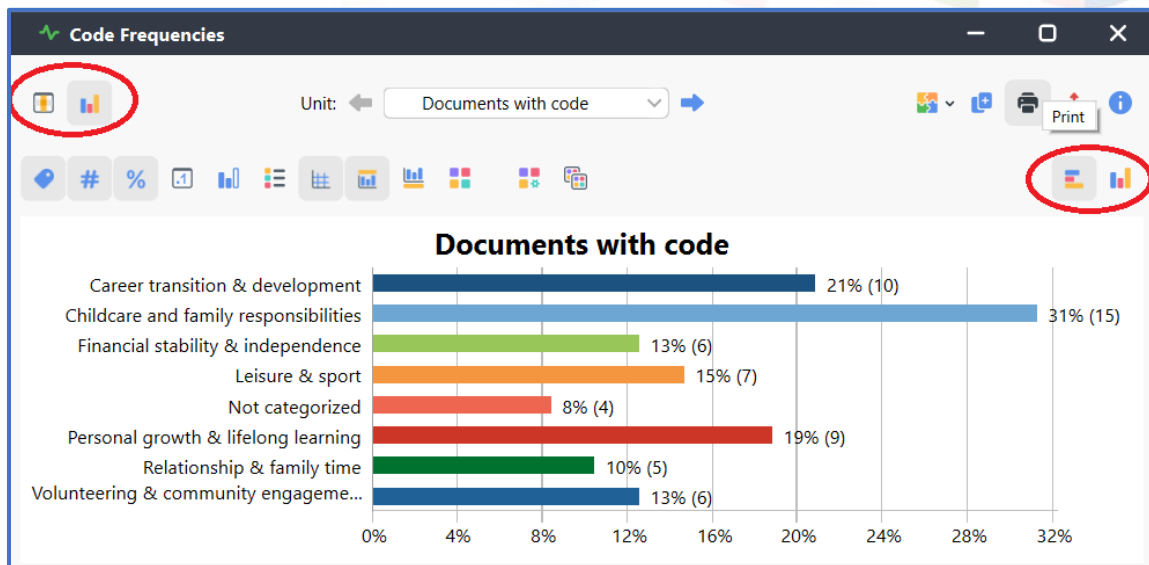
Using data from both open and closed question responses:

This is the point at which the benefit of importing the variable data as well as the open questions becomes apparent. If your survey sample is large enough and sufficiently randomly selected, then you may reasonably make inferences from how various subgroups of your cases raise some of the themes you identify in the coded responses. This can be done quite easily, either as quantitative frequency tables or, more excitingly, with informative text quotations to show differing meanings.

First, let us look at the basic reporting of the frequencies of the codes we have just been applying to the open questions. This can be done simply with the function **Analysis > Code Frequencies**. It is best to exclude the overall frequency of the responses to the question itself, so activate all the Survey documents and the whole group of analysis codes for the question being examined before starting the function. The dialog is then easy to complete by clicking on the box for "Insert activated codes" to bring them all in and then use the "X" button on the row for the question code itself to exclude that and leave all the thematic codes. The dialog should default to tick the box for "Only for activated documents" which is what you want.

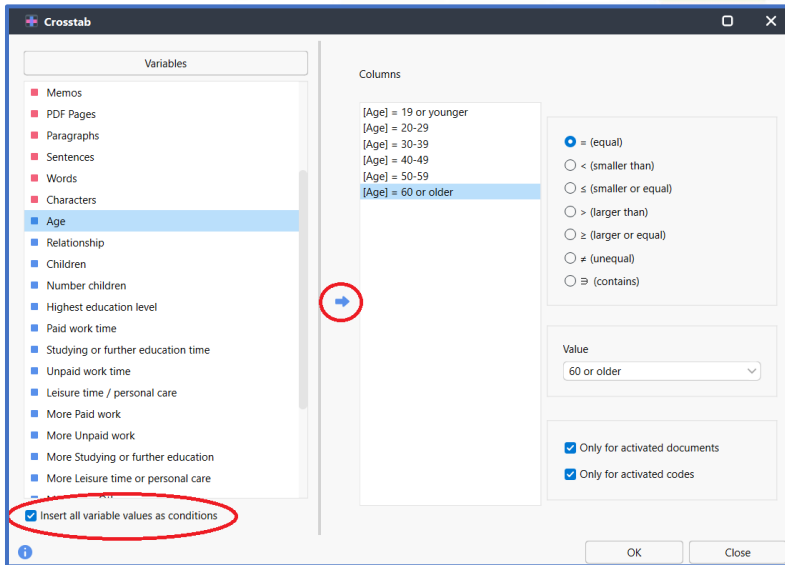
In the table which opens, you can choose the sequence in which to display the results, alphabetical order of the code names or descending frequency of the responses. Then click on the "Chart view" button (2nd on the toolbar). I find that the

horizontal bar chart is useful for this kind of display because the code names are clearer.



In the screenshot above, I have circled the pair of icons in the top left corner which switch back and forth between table view and chart view. On the right side I have circled the icons which switch between horizontal and vertical bar charts. Other toolbar icons allow you to add or remove percentages and frequencies, change colour schemes and edit titles to allow you to create the visual appearance you want.

To start looking at interactions between data from closed questions and these themes we can use the function **Mixed Methods > Crosstab**. We can use the same activations as before. The dialog to set this up is shown below. I first ticked the box (circled) at the bottom left for "Insert all variable values as conditions", then selected the variable I want to use - this time "Age" - and finally clicked on the arrow in the middle of the dialog to move "Age" and all of its values into the middle part of the dialog to become the "columns" in the table. The program has defaulted to add ticks on the right for only activated documents and codes, because it detected that some of each had been activated.



Below is the table generated from this dialog:

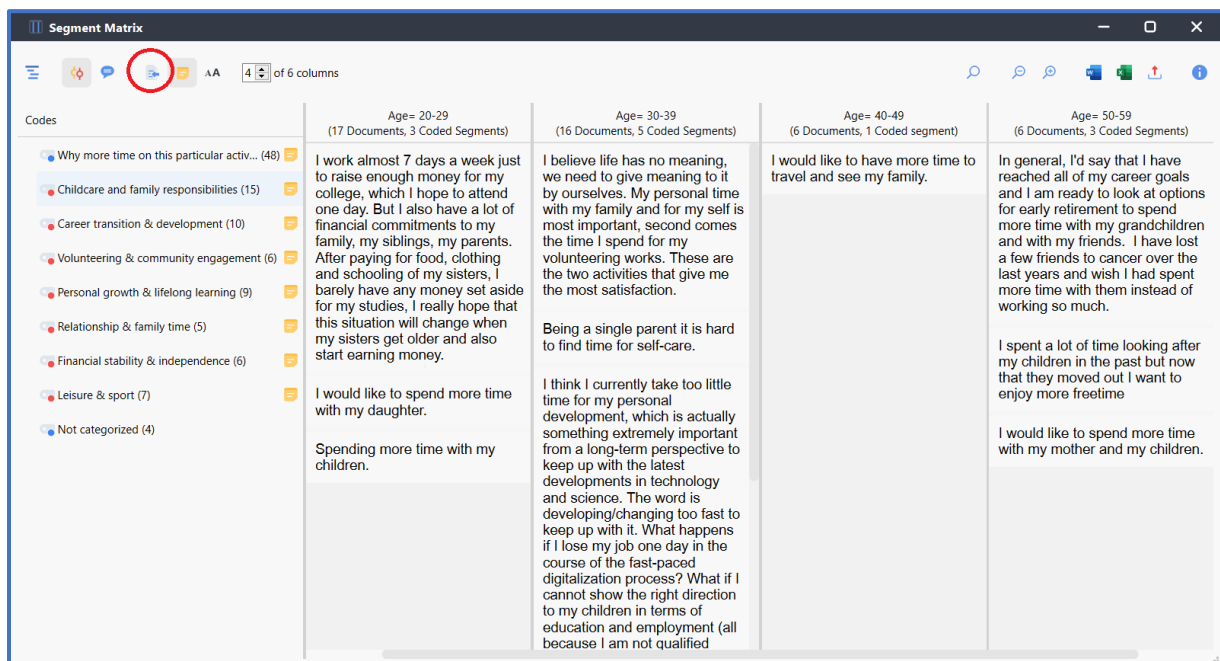
	19 or younger	20-29	30-39	40-49	50-59	60 or older	Total
Why more time on this particular activity	8.3%	27.1%	31.3%	12.5%	12.5%	8.3%	100.0%
Childcare and family responsibilities	13.3%	20.0%	33.3%	6.7%	20.0%	6.7%	100.0%
Career transition & development	10.0%	20.0%	30.0%	20.0%	20.0%		100.0%
Volunteering & community engagement		16.7%	33.3%	16.7%		33.3%	100.0%
Personal growth & lifelong learning		33.3%	33.3%	11.1%		22.2%	100.0%
Relationship & family time	20.0%	20.0%	20.0%		40.0%		100.0%
Financial stability & independence	33.3%	33.3%			33.3%		100.0%
Leisure & sport		14.3%	28.6%	28.6%	14.3%	14.3%	100.0%
Not categorized		25.0%	75.0%				100.0%
SUM	9.1%	24.5%	30.9%	11.8%	14.5%	9.1%	100.0%
# N = Documents	5 (9.1%)	17 (30.9%)	16 (29.1%)	6 (10.9%)	6 (10.9%)	5 (9.1%)	55 (100.0%)

I have clicked on the 2 circled icons. The first of these reduced the width of the table by reducing the information shown in each column header, I would need to make sure that any subsequent presentation of this table clearly states that the columns refer to the ages of the respondents. The second has set the display to "Row percentages".

We need to be careful in this type of table to account for the fact that different numbers of respondents fall into each age-range used, so raw frequency numbers can be misleading. The top row of data shows the percentages for each age group

that responded to this question, the bottom row of data ("# N = Documents") shows the percentages for each group of participants in the survey some of whom did not respond here. Where a percentage in the rest of the table is very different from both of these it might be an indication of something interesting. The sample size here is not sufficient for these differences to be reliable but, hopefully, you will collect more data and find useful results in this way.

The final step is to convert the frequency table into a quotations table. This is done by clicking on the first icon in the toolbar of the Crosstab table ("Segment Matrix").



This table displays all of the source texts behind the Crosstab table. By clicking on a code in the first column, you display all of its segments in the table. Scroll to the right with the slider at the bottom of the screen to see the columns for older age-groups. Each column of texts is independently scrollable (see the vertical slider on the "30-39" column). The circled toolbar icon toggles the display of Document Names on/off. In this table you can consider how the language used by respondents of different ages differs for any of the themes identified in your coding, which can be more interesting than simple frequency differences.