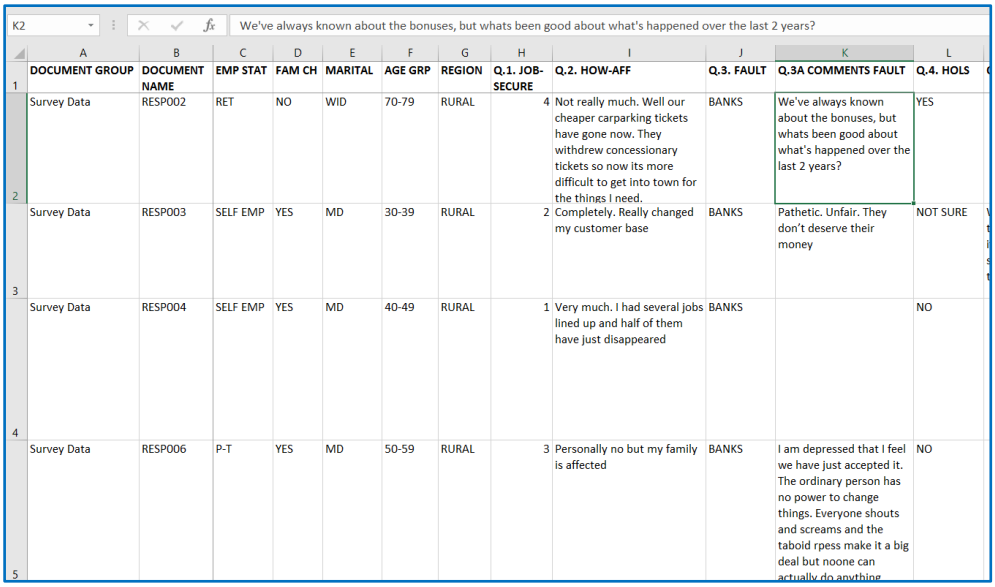
MAXDAYS
MAXQDA VIRTUAL CONFERENCE

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

Excel format for data preparation:



K2 — We've always known about the bonuses, but whats been good about what's happened over the last 2 years?

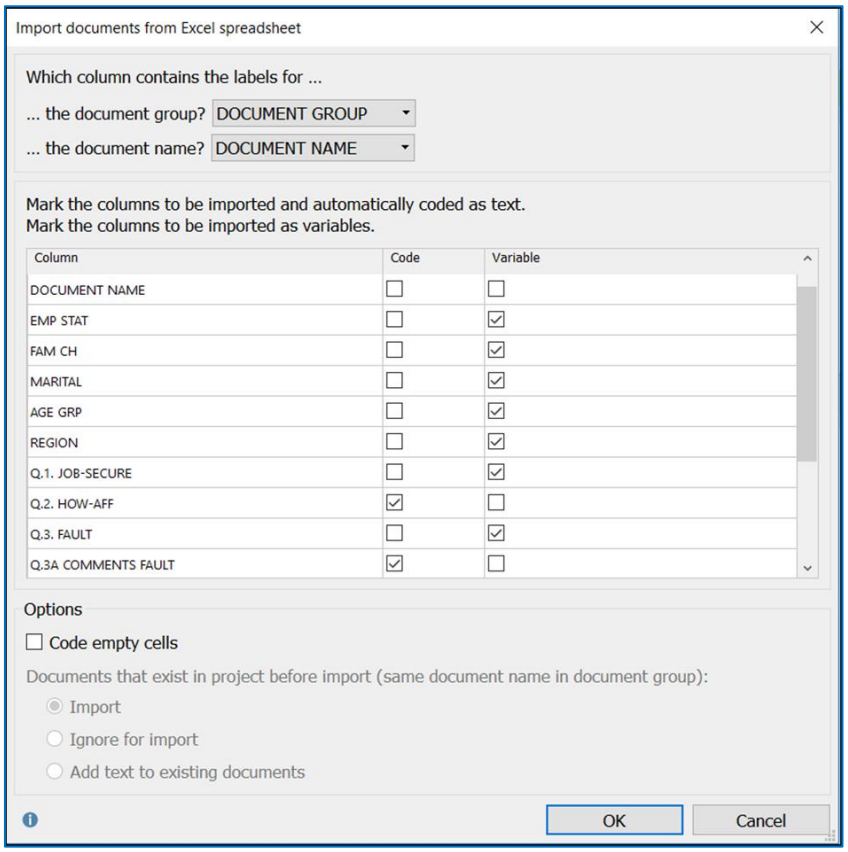
	A	B	C	D	E	F	G	H	I	J	K	L
1	DOCUMENT GROUP	DOCUMENT NAME	EMP STAT	FAM CH	MARITAL	AGE GRP	REGION	Q.1. JOB-SECURE	Q.2. HOW-AFF	Q.3. FAULT	Q.3A COMMENTS FAULT	Q.4. HOLS
2	Survey Data	RESP002	RET	NO	WID	70-79	RURAL	4	Not really much. Well our cheaper carparking tickets have gone now. They withdrew concessionary tickets so now its more difficult to get into town for the things I need.	BANKS	We've always known about the bonuses, but whats been good about what's happened over the last 2 years?	YES
3	Survey Data	RESP003	SELF EMP	YES	MD	30-39	RURAL	2	Completely. Really changed my customer base	BANKS	Pathetic. Unfair. They don't deserve their money	NOT SURE
4	Survey Data	RESP004	SELF EMP	YES	MD	40-49	RURAL	1	Very much. I had several jobs lined up and half of them have just disappeared	BANKS		NO
5	Survey Data	RESP006	P-T	YES	MD	50-59	RURAL	3	Personally no but my family is affected	BANKS	I am depressed that I feel we have just accepted it. The ordinary person has no power to change things. Everyone shouts and screams and the taboid rpess make it a big deal but noone can actually do anything	NO

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 Data".
- 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, and remember that shorter labels are better than long ones.
- 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 (click the "(?)" button at the top right-hand corner of the MAXQDA interface).

- 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:

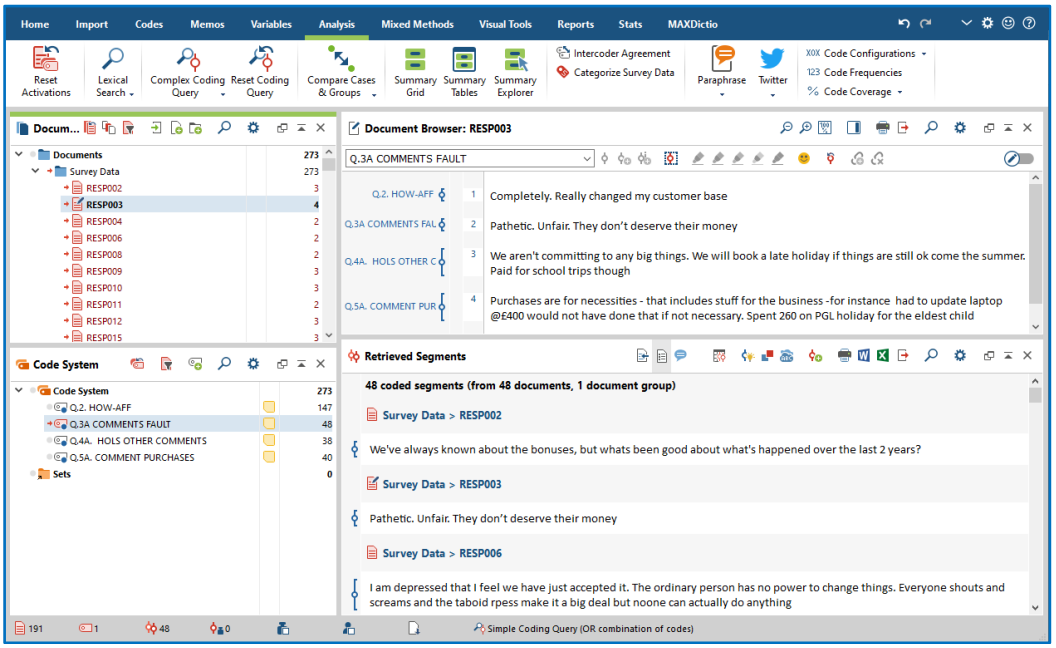


Notes:

- 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 goes into the “Variable” column.
- At the top of this dialog, the program has taken the header labels from columns A & B in the Excel workbook and inserted them in the menu boxes, so you could use different labels for these columns in the workbook but these are the clearest.
- 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.

MAXDAYS
MAXQDA VIRTUAL CONFERENCE

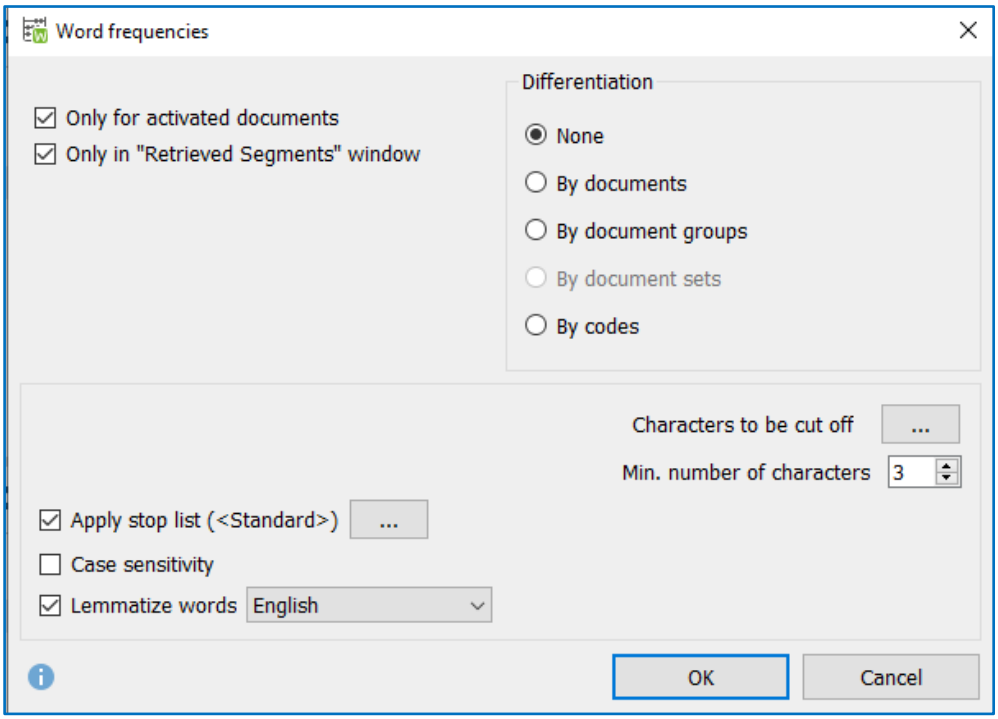
Setup for analysing responses to one question:



Notes:

- I have activated the entire Document Group "Survey Data" in the Document System and activated the Code "Q3A Comments Fault" in the Code System to get this display. All 48 responses to Q3A are now readable in MAXQDA's "Retrieved Segments" window.
- Note that 143 cases in this survey (of 191 people) did not answer this question.
- The Document Browser shows all the responses to the 4 open questions from a single respondent ("RESP003").
- From this display you have 3 possible approaches to start the analysis:
    - you could work manually by coding the responses directly from the "Retrieved Segments" window
    - you can use a more efficient approach with the option **Analysis > Categorize Survey Data**
    - or, you can use a more automated approach as detailed in the next section.
- Before starting your coding work, it is good practice to enter the full text of each question into the Code Memo for that question. This means that you can remind yourself of that wording at any stage during the analysis. The wording for the Q3A question in this data was *"Please expand on your understanding of who/what is at fault"*.

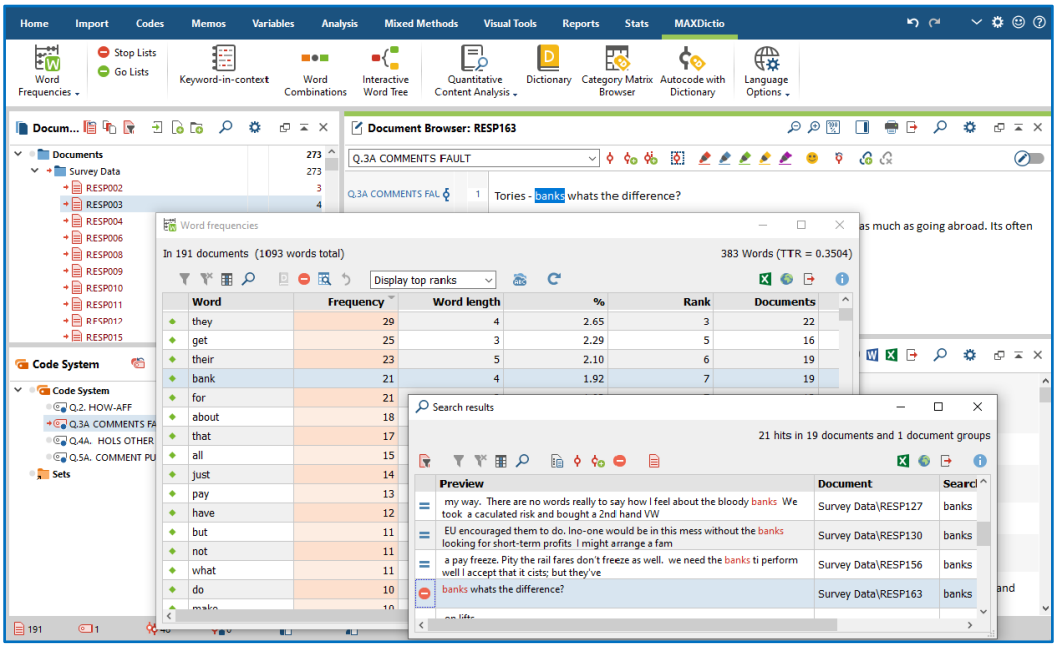
Using **MAXDictio > Word Frequencies** to analyse the contents of the “Retrieved Segments” window:



Notes:

- You need to have MAXQDA Plus to use this routine.
- The activated documents and retrieved segments settings limit the routine to the single question that is currently appearing in MAXQDA's "Retrieved Segments" window.
- Differentiation is not helpful in this particular application, so leave the default on "None".
- "Stop lists" can be helpful at avoiding counting trivial words like "of", "and", etc. You can create and edit your own stop lists for particular purposes, but it may not be necessary for a one-off survey project.
- The "Lemmatize" function will combine words with a common stem, in this example the count for "bank" includes "Banks" and "banking" which is useful in this particular application.
- The "Min. number of characters" setting will exclude words shorter than this length from the count, so "by" will not be counted here but "pay" and "job" will be included. Setting this higher might exclude some important words.
- Click on "OK" to run the frequency count and see the display shown below.

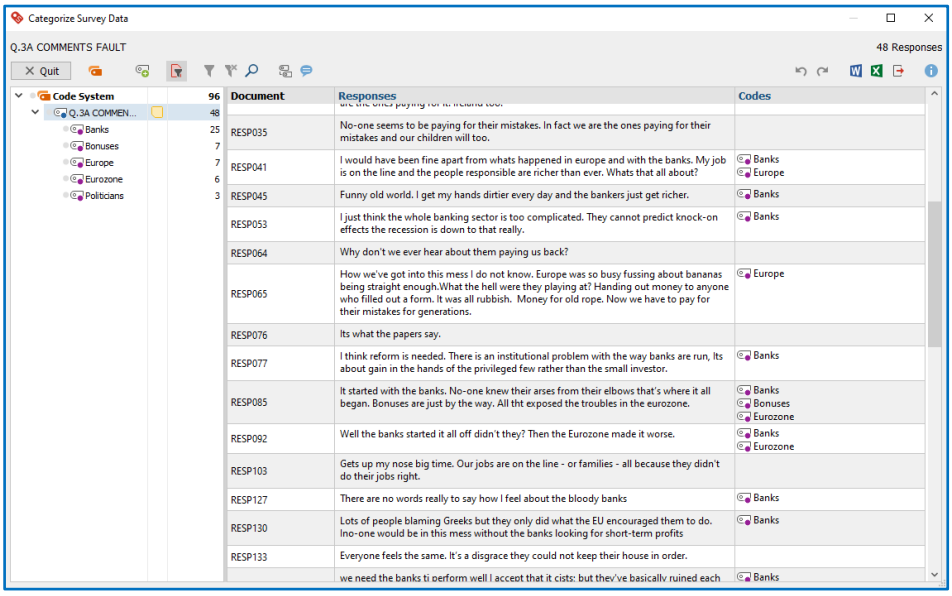
Frequency table and search results:



Notes:

- In the "Word frequencies" table, select a word of interest with a click to highlight its row (e.g. “bank”). Use a right-click (Windows) and select "Search results" from the context menu to display the "Search results" table. Click on a row in the results table to see its full text highlighted in the Document Browser behind these 2 windows.
- Often this display (effectively keyword in context) is enough to identify whether the word has been used in a consistent way in each instance. If an item is inconsistent, exclude it from the next step by highlighting it and clicking on the "no entry" button (as here for case 163).
- Use the 7th or 8th buttons in the toolbar of the results table to autocode all of the remaining hits in that table with an existing code or a new code.
- In the autocode dialog, think about how much text you want to apply the code to. The option "Only search string" will just code the selected word each time. More usually with this kind of data, coding the whole paragraph containing the hit will give you just one coded segment per response - allowing you to identify how many respondents mentioned the concept.
- Repeat the search and autocode approach for the most common themes arising for this question (I found "bonuses", "Europe", "Euro", and "Politicians" for this data), then close all of these windows and switch to the **Categorize Survey Data** tool in the **Analysis** menu.

Using the Categorize Survey Data tool:

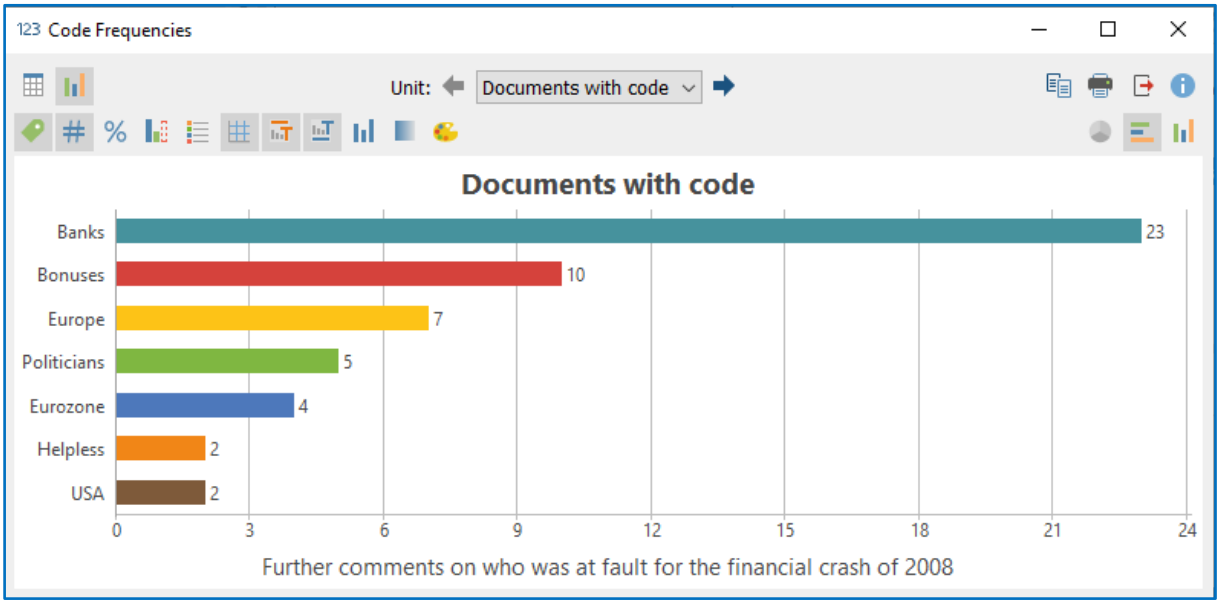


Notes:

- It is necessary to drag the new codes that you have just created into subcodes of the relevant question in the "Code System" window, because that is the basis on which this tool works. Then select the option **Analysis > Categorize Survey Data**, drag in the code of the question that you are analysing and click "OK" to see the screen shown above.
- You can now see all of the initial coding that you have done with the MAXDictio and autocoding tools, so you can re-check that (if you want), but most likely you will want to focus on the segments that have no code attached as yet (blanks in the Codes column).
- Apply a code by dragging from right to left (data segment onto the code).
- As you identify new themes by reading the uncoded segments, you can create new codes for them with the 2nd icon to the right of the "Quit" button. If you want to search all of these segments for a particular word, you have a search tool at the 6th icon. If you want to add comments about any segments, you can open the column for those with the 8th icon.
- This display gives you a strong visual check for uncoded segments. 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.
- By clicking on each thematic code in turn, you can quickly check each subset for consistency, and finally, re-examine the uncategorizable segments.
- For a final attempt with the most difficult comments, quit this tool and activate just the “uncategorizable” code under this question to pull those segments into the Retrieved Segments window. Click on the code brackets

there to see each in the context of the responses by that person to the other open questions in the Document Browser above. To see their variable data (such as their age and answer to the closed “Fault” question) go to the Variables > List of Document Variables table and add some ticks in the column headed “Favorite variable”, this data will then be shown between the Document name and the data segment in the Retrieved Segments window.

Reporting the results of your analysis:

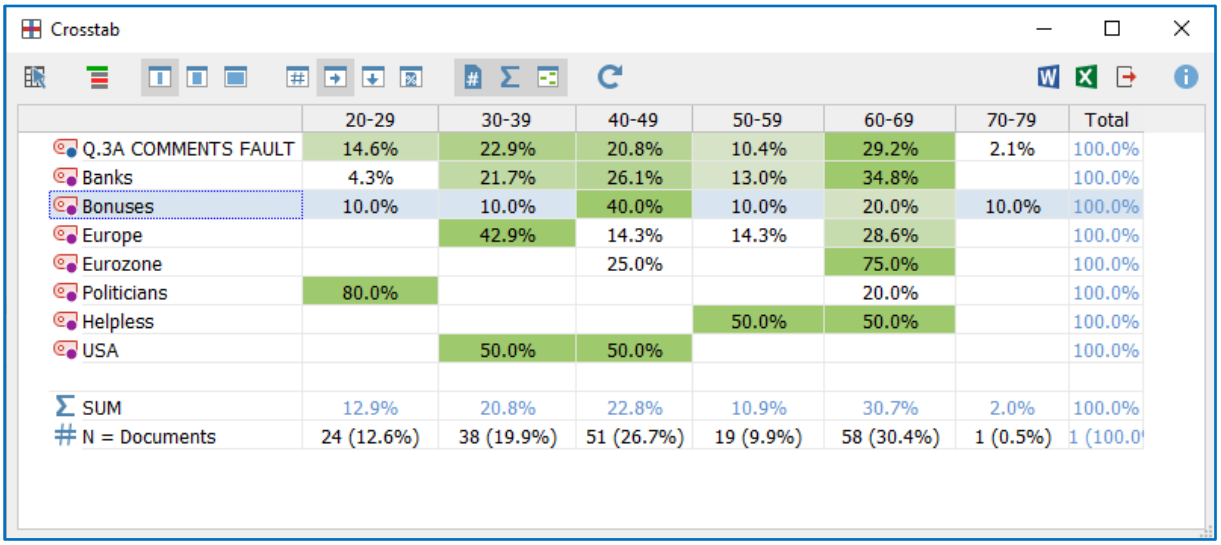


Notes:

- To obtain a display similar to the one above, you need to use the tool **Analysis > Code Frequencies**. It will be best to activate the relevant group of codes in the Code System first so that you can use the option to "Paste activated codes" in the set-up dialog (and you will then need to "Remove" the question code itself, which is the group header, to see the frequencies of just the analysis codes).
- You can explore the table and chart settings, which are all controlled with the toolbars in the display window.
- You can report on the quantitative elements of your survey by using the Variables functions. For example, try **Variables > Document Variable Statistics** and select several of the closed question identifiers in the set-up dialog. You can then scroll forwards and backwards through these in table or chart views. Note that you can re-order the tables by frequency or alphabetically and that this controls the equivalent charts.

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 who was at fault for the economic crash analysed according the age-group of the respondents.



	20-29	30-39	40-49	50-59	60-69	70-79	Total
Q.3A COMMENTS FAULT	14.6%	22.9%	20.8%	10.4%	29.2%	2.1%	100.0%
Banks	4.3%	21.7%	26.1%	13.0%	34.8%		100.0%
Bonuses	10.0%	10.0%	40.0%	10.0%	20.0%	10.0%	100.0%
Europe		42.9%	14.3%	14.3%	28.6%		100.0%
Eurozone			25.0%		75.0%		100.0%
Politicians	80.0%				20.0%		100.0%
Helpless				50.0%	50.0%		100.0%
USA		50.0%	50.0%				100.0%
SUM	12.9%	20.8%	22.8%	10.9%	30.7%	2.0%	100.0%
N = Documents	24 (12.6%)	38 (19.9%)	51 (26.7%)	19 (9.9%)	58 (30.4%)	1 (0.5%)	1 (100.0

Notes:

- There is not really sufficient data in this example to justify any conclusions but in the above table it appears that the comments about bonuses were made in disproportionate numbers by people in their 40's (40.0% in that row compared with 26.7% of the respondents being in that age group as shown in the bottom row).
- This is where the full pay-off for using MAXQDA to analyse your survey data becomes apparent!

Reference:

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

Graham Hughes (2021)