

QDA Miner v3.2 addendum

This short document presents the new features and describe all changes introduced in the version 3.2 of QDA Miner.

Creating Virtual Grouping of Codes



Codebook categories are used mainly to organize codes in a hierarchical way. By default, categories cannot be assigned to text segments and frequency statistics can only be computed on codes, not on their categories. This restriction has been implemented to maintain the independence of observations required by some routines available in QDA Miner, like clustering or multidimensional scaling. However, one may sometimes wish to obtain descriptive statistics at a category level. While one could transform codes within a category into a single code by merging all codes underneath this category (using the `CODES | MERGE` command), such a solution results in a permanent loss of information since all coding made using those codes are replaced with a single code. The `VIRTUAL GROUPING` function allows one to temporarily group all codes in a category and instruct QDA Miner to compute frequencies and obtain statistics at the category level without reporting statistics on codes underneath it. For example, if a category named `POSITIVE` contains five codes, by default, the `CODING FREQUENCIES` command will report statistics on each of those. If you virtually group codes in this category, the same command will no longer display statistics to each code, but only a single line of statistics for the `POSITIVE` category. Export coding statistics will also result in a single variable containing frequency data for the whole category.

The `VIRTUAL MERGE` function will affect the following features of QDA Miner:

- ANALYSIS | CODING FREQUENCIES
- ANALYSIS | CODING CO-OCCURRENCES
- ANALYSIS | CODING SEQUENCES
- ANALYSIS | CODING BY VARIABLES
- ANALYSIS | CODING AGREEMENT
- VARIABLES | COMPUTE COMMAND
- FILE | EXPORT | STATISTICS

The presence of virtual groupings will not affect the text or coding retrieval features, as well as any coding or other codebook management procedures.

To virtually group codes into their parent category:

- In the codebook, highlight the category on which you would like to obtain statistics.
- Run the `CODES | VIRTUAL GROUPING` command. The category icon will be transformed into a  icon while icons of all codes underneath it will be replaced with a  icon.

To ungroup codes in a virtually grouped category:

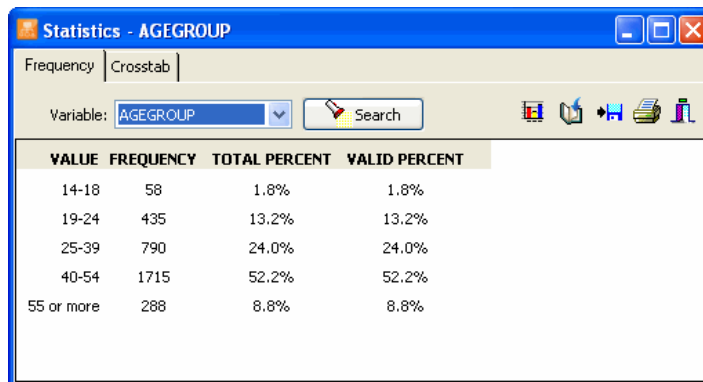
- In the codebook, highlight the category you would like to ungroup.
- Run the CODES | VIRTUAL GROUPING command. The original icons of the highlight category and codes will be restored, indicating that they are no longer grouped.

Variables Statistics

The new VARIABLES | STATISTICS command allows one to quickly obtain the frequency and cross-frequency distribution of numerical, categorical, date and short-string variables. The univariate frequency table includes the frequency count for each value of the selected variable as well as the percentage of the count over all cases and over valid cases only. The contingency table describes the distribution of two variables simultaneously by displaying either the frequency or the row, columns or total percentages. Several types of charts may be created to illustrate the distribution or cross-frequency distribution of variables. You will find below a list of those charts.

To obtain the frequency distribution of a variable:

- From the main QDA Miner screen, in the Variables window, highlight the variable on which you would like to obtain distribution statistics.
- Select the STATISTICS command from the VARIABLES menu, or right-click and select the STATISTICS command.




The screenshot shows a software window titled "Statistics - AGEGROUP". It has two tabs: "Frequency" (selected) and "Crosstab". Below the tabs, there is a "Variable:" dropdown menu with "AGEGROUP" selected and a "Search" button. To the right of the search button are several icons for chart types. Below this is a table with the following data:

VALUE	FREQUENCY	TOTAL PERCENT	VALID PERCENT
14-18	58	1.8%	1.8%
19-24	435	13.2%	13.2%
25-39	790	24.0%	24.0%
40-54	1715	52.2%	52.2%
55 or more	288	8.8%	8.8%

You may obtain a frequency table on any other numerical, categorical, alphanumeric, and date variable by selecting its name in the **Variable** list box and then clicking the **Search** button.

To chart the distribution of a variable:

Clicking the  button allows one to obtain up to five types of charts to visually display the distribution of specific codes. Some of those charts are available only for numerical and date variables (histograms and box-&-whiskers plots), while others like the bar charts and pie charts will be available in all situations when the total number of values is less than 100.



The vertical bar chart is the default chart used to display the frequencies of distinct values of a nominal or ordinal variable. It is especially useful to compare two or more values.



The horizontal bar chart displays the same information as the vertical bar chart. It is especially useful when the number of values is high and their labels cannot be displayed entirely on the bottom axis.



The pie chart is useful to display the relative frequency of each value and compare individual values to other values and to the whole. Numerical values displayed in pie charts are always expressed in percentages of either the total frequency or case occurrences.



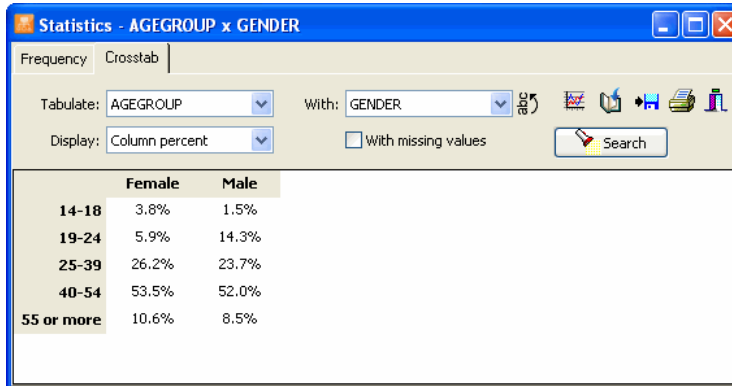
The histogram graphically displays the distribution of a numeric variable. When selected, the program first separates the values into non-overlapping intervals of equal width, then plots bars that represent the frequencies of each interval.



The box-&-whiskers plot can be used to examine the distribution of numerical variables. It is especially useful to detect the presence of outliers and asymmetry in the data distribution. The box includes values that fall between the first and the third quartiles (about 50 percent of the values). The line in the middle of the box represents the median value while the whiskers extend to the farthest observations within 1.5 times the interquartile range measured from the nearest quartiles. Values that are situated farther than 1.5 times the interquartile range but within three times this distance are represented by a dot, while values farther than three times the interquartile range from the nearest quartile are represented by the letter X (for extreme).

To obtain the joint distribution of two variables:


- Select the STATISTICS command, from the VARIABLES menu to display the **Statistics** dialog box (see above).
- Move to the **Crosstab** page. The dialog box will be similar to the one below:




- Select from the **Tabulate** list box the variable you would like to be displayed on the rows of the table.
- Select the variable you would like to be displayed at the top of the table by selecting its name in the **With** list box.
- In the **Display** list box, select the statistics you would like to be displayed in the table.
- Click the **Search** button.

To chart the joint distribution of two variables

Bar charts or line charts are useful for visually comparing the joint distribution of two variables. To produce these types of charts:


- Set the TABULATE, WITH, and DISPLAY options so that the information to be viewed is displayed in the table.
- Click the  button.

To append a copy of a table in the Report Manager:

- Click the  button. A descriptive title will be provided automatically for the table. To edit this title or to enter a new one, hold down the SHIFT keyboard key while clicking this button.

For more information on the Report Manager, see QDA Miner User's Manual page 180..

To export retrieved codes to disk:

- Click the  button. A Save File dialog box will appear.
- In the **Save as type** list box select the file format under which you would like to save the table. The following formats are supported: ASCII file (*.TXT); Tab delimited file (*.TAB); Comma delimited file (*.CSV); MS Word (*.DOC); HTML file (*.HTM; *.HTML); XML file (*.XML); Excel spreadsheet file (*.XLS).
- Type a valid filename with the proper file extension.
- Click the **Save** button.

The obtained table can also be stored in a new project file, where each row becomes a new case and each column is transformed into a variable. This type of project file may be useful for performing a more detailed analysis of coded segments.

To print the table:

- Click the  button.


Coding by Variables

Several important new features have been added to the **Coding By Variable** dialog box:

Combining values of two variables

You can combine values of two variables by setting the **Tabulate With** list box to **<combine...>**. When this option is selected, you will be presented with a dialog box that will let you choose the two variables you wish to combine.

Displaying column labels vertically

To display column labels vertically, click the  button located on the right of the **Tabulate With** list box.

Searching for selected coded segments

- Select the cells that contain the data for which you would like to retrieve associated coded segments.
- Right-click and choose the RETRIEVE SEGMENTS command.

You may also select entire rows by right-clicking anywhere in the table and enabling the ROW SELECT option. When this option is enabled, multiple but separate rows can also be selected by clicking while holding the CTRL key down.

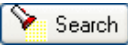

All retrieved segments will be displayed in a separate dialog box.

Creating bubble charts

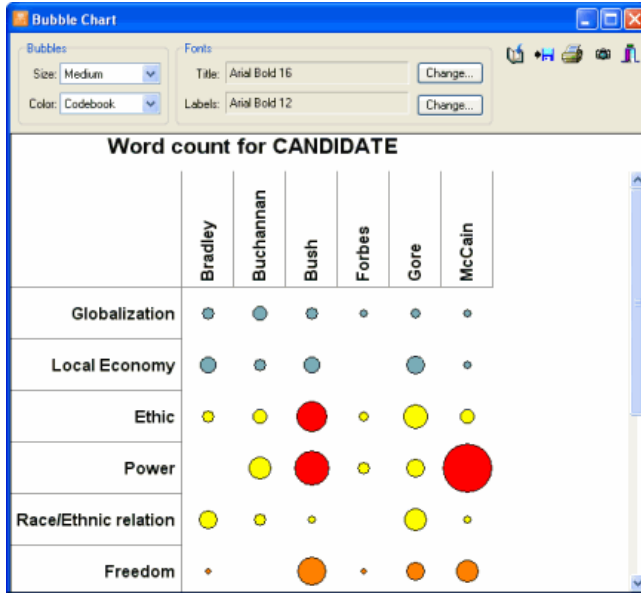
Bubble charts are graphic representations of contingency tables where relative frequencies are represented by circles of different diameters. This type of graphic chart allows one to quickly identify high- and low-frequency cells and is thus especially useful for presentation purposes. Many features of the chart can be customized to highlight specific findings. Rows and columns can be moved freely or deleted, and one can adjust the color of each cell as well as the fonts used in the chart.

The bubble chart represents graphically the underlying table and displays the corresponding measure used (code occurrences or frequencies, word counts or percentages) and will also reflect the currently selected display settings (such as count, percentage of rows or columns, etc.).

To create a bubble chart:

- Select the CODING BY VARIABLES command from the ANALYSIS menu.
- Select the codes that you want to plot, along with the documents to which you wish to restrict the analysis.
- Set the **Tabulate With** option to the desired independent variable.
- Set the **Count** option to either **Code Occurrence**, **Code Frequency**, **Word Count**, or **% of Words**.
- Set the **Display** option to specify how this information will be displayed.
- Click the  Search button to construct the frequency matrix.
- Click the  button.

A dialog box similar to this one will appear:



To adjust the size of the bubbles:

- In the **Bubbles** group box at the top of the window, adjust the **Size** option to Small, Medium or Large.

To adjust the color of the bubbles:

Global adjustment

- To adjust the color of all bubbles, set the **Color** option to either **Codebook** or **Single Color**. Choosing the **Codebook** option will reset the color of bubbles to the color associated with each code in the codebook. Selecting **Single Color** will bring a dialog box that will allow you to select which color should be used for displaying all bubbles in the chart.

Manual adjustment

- To adjust the color of specific bubbles, select the cell or group of cells you would like to alter.
- Right-click and select **Set Color**. A dialog box will appear letting you choose a specific color value.

To move a row or a column:

- Click anywhere on the cell header containing the row or column label and hold the mouse down.
- Drag the mouse cursor over the desired new location and release the mouse button.

To delete selected rows or columns:

- Select a cell range covering the rows or columns you would like to delete.
- Right-click to display a pop-up menu.
- Select the REMOVE command and then choose SELECTED ROWS or SELECTED COLUMNS depending on your desired action.


To adjust the font size and style of the title and labels:

- Click the **Change** button beside the **Title** or **Labels** option. A **Font** dialog box will appear, letting you change the font size, style, and color.

To edit the title:


- Click anywhere in the title region and type in the new title. The height of the title region is automatically adjusted to the number of lines in the title.

To append a copy of the table in the Report Manager:


- Click the  button. A descriptive title will be provided automatically for the table. To edit this title or to enter a new one, hold down the SHIFT keyboard key while clicking this button.

For more information on the Report Manager, see QDA Miner User's Manual page 180.

To create a text report of retrieved segments:

- Click the  button. The sort order of the current table is used to determine the display order in the report. This report is displayed in a text editor dialog box and may be modified, stored on disk in RTF, HTML or plain text format, printed, or cut and pasted into another application. Graphics and tables may also be inserted anywhere in this report.


To export the table to disk:

- Click the  button. A Save File dialog box will appear.
- In the **Save As Type** list box, select the file format under which to save the table. The following formats are supported: ASCII file (*.TXT), Tab delimited file (*.TAB), Comma delimited file (*.CSV), MS Word (*.DOC), HTML file (*.HTM; *.HTML), XML files (*.XML) and Excel spreadsheet file (*.XLS).
- Type a valid file name with the proper file extension.
- Click the **SAVE** button.

To print the table:

- Click the  button.

Enhanced Integration with WordStat

WordStat can now be called from various text or coding retrieval dialog boxes to analyze the retrieved text segments. This feature is available by clicking the  button located on the toolbar above the results table.

Filtering Cases

Two new operators have been added for filtering document variables. The **Is Coded** operator will return true if one or more codes have been assigned to the document. The **Is Uncoded** operator will return true only if the document contains no coded segments.

Printing the Codebook

The following two options have been added to the Print Codebook dialog box:


PRINT DESCRIPTION - Select this option to print not only the code names but also their descriptions. The description is printed in italics just below the code name.

PRINT KEYWORDS - This option prints the keywords and key phrases associated with each code.


Coding Co-occurrences

New search tools have been added to various pages of the CODING CO-OCCURRENCE procedure:


- From the **Dendrogram** page:

To retrieve text segments associated with a specific cluster, click anywhere on a cluster to select it (the selected cluster is displayed with thicker black lines), and then click on the  button to retrieve the associated text segments. This operation retrieves all pairs of coded segments associated with the selected cluster.


- From the **Proximity Plot** page:

When looking at coding co-occurrences, selecting an item enables the  button. Clicking this button retrieves every pair of coded segments co-occurring, allowing one to further explore the factors that may explain this co-occurrence.


- From the **Co-occurrence** table of the **Tables & Statistics** page:

Select the cells containing the data for which you would like to retrieve associated coded segments and click the  button.

Bar Chart or Line Chart

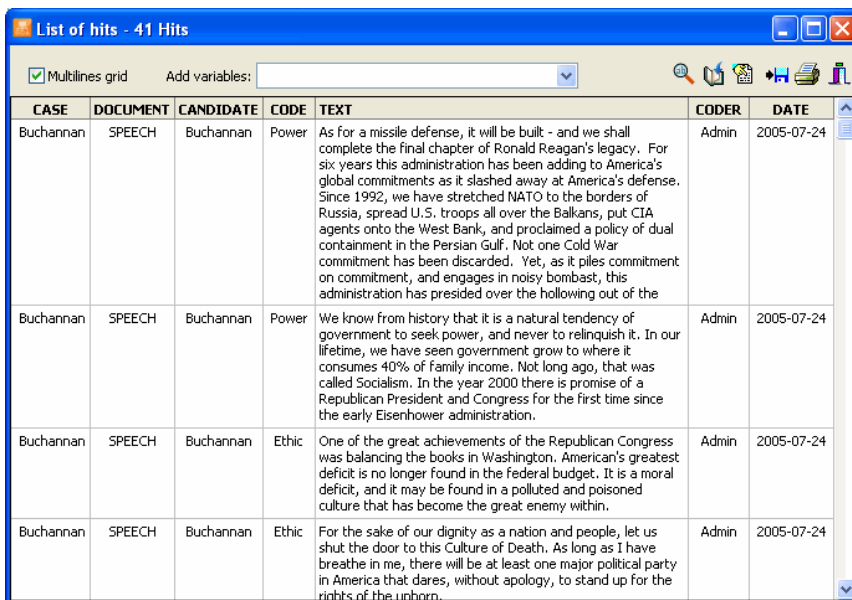
Clicking down the new  button vertically displays the labels on the bottom axis.

Correspondence Analysis

Items distributed in a very similar way among subgroups of the categorical variables may be plotted on top of each other, making them difficult to differentiate. Clicking down the new  button adds some random noises to the location of individual words or keywords, allowing one to clearly identify those words that overlap. To remove the random noises, click the button again.

Using the Retrieved Segments Dialog Box

The Retrieved Segments dialog box is displayed when one searches coded segments from joint frequency or co-occurrence tables as well as from various charting tools (such as heatmaps, bubble charts, bar charts, dendrograms, proximity plots). It lists all coded segments associated with the last retrieval operation in a table that may be customized, saved or printed. The table provides basic information on each hit, including the case and the document variable in which it was found, the code name, the coder's name and coding date as well as the full text of the coded segment. When searching for co-occurrence, the texts of both coded segments are displayed in separate columns (see below).



CASE	DOCUMENT	CANDIDATE	CODE	TEXT	CODER	DATE
Buchanan	SPEECH	Buchanan	Power	As for a missile defense, it will be built - and we shall complete the final chapter of Ronald Reagan's legacy. For six years this administration has been adding to America's global commitments as it slashed away at America's defense. Since 1992, we have stretched NATO to the borders of Russia, spread U.S. troops all over the Balkans, put CIA agents onto the West Bank, and proclaimed a policy of dual containment in the Persian Gulf. Not one Cold War commitment has been discarded. Yet, as it piles commitment on commitment, and engages in noisy bombast, this administration has presided over the hollowing out of the	Admin	2005-07-24
Buchanan	SPEECH	Buchanan	Power	We know from history that it is a natural tendency of government to seek power, and never to relinquish it. In our lifetime, we have seen government grow to where it consumes 40% of family income. Not long ago, that was called Socialism. In the year 2000 there is promise of a Republican President and Congress for the first time since the early Eisenhower administration.	Admin	2005-07-24
Buchanan	SPEECH	Buchanan	Ethic	One of the great achievements of the Republican Congress was balancing the books in Washington. American's greatest deficit is no longer found in the federal budget. It is a moral deficit, and it may be found in a polluted and poisoned culture that has become the great enemy within.	Admin	2005-07-24
Buchanan	SPEECH	Buchanan	Ethic	For the sake of our dignity as a nation and people, let us shut the door to this Culture of Death. As long as I have breathe in me, there will be at least one major political party in America that dares, without apology, to stand up for the rights of the unborn.	Admin	2005-07-24


To sort the list of hits:

- To sort in ascending order any column values, simply click this column header. Clicking a second time on the same column header sorts the rows in descending order.

To add variables to the table:

- Use the **Add Variables** drop-down checklist box to add the values stored in one or more variable to the table of retrieved segments for the specific case from which a text segment originates.

To analyze the text of all retrieved segments using WordStat:

- Click the  button. QDA Miner will minimize itself and run WordStat, allowing you to perform a quantitative content analysis or apply text-mining techniques on the retrieved segments. WordStat will also allow you to perform keyword comparisons across any other column in the table. To include additional data for comparison analysis, use the **Add Variables** option.