QDA Miner 3.2 (with WordStat & Simstat)
Distinguishing features and functions
Christina Silver & Ann Lewins

This document is meant to be read in conjunction with the 'Choosing a CAQDAS Package Working Paper' which provides a more general commentary of common CAQDAS functionality. This document does not provide an exhaustive account of all the features and functions provided by QDA Miner but is designed to highlight some of its distinguishing elements. Here we review version 3.2 of QDA Miner and make some reference to the WordStat and SimStat modules. The Comment section at the end details our opinions on certain aspects of functionality and usability. See also Lewins & Silver (2007) *Using Software in Qualitative Research: A Step-by-Step Guide*, Sage Publications and software developer website.

Background   http://www.provalisresearch.com/

Normand Peladeau developed QDA Miner as a mixed-method qualitative data analysis tool to facilitate the analysis of textual and, to a limited extent, visual (graphic) data. It contains features characteristic of mainstream CAQDAS packages ■ QDA Miner can be integrated with WordStat a content analysis and text-mining module (and also with SimStat a statistical data analysis tool). Utilising the suite of three packages extends the range of exploration and analytic techniques to the more mathematical, *content* and *statistical* analyses of textual data. Potentially handling thousands of cases, the combined software spans the CAQDAS, Text–retriever and Text-based Manager categories of software (see Choosing a CAQDAS package working paper) ■ QDA Miner is therefore part of a new generation of qualitative software, crossing methodological boundaries.
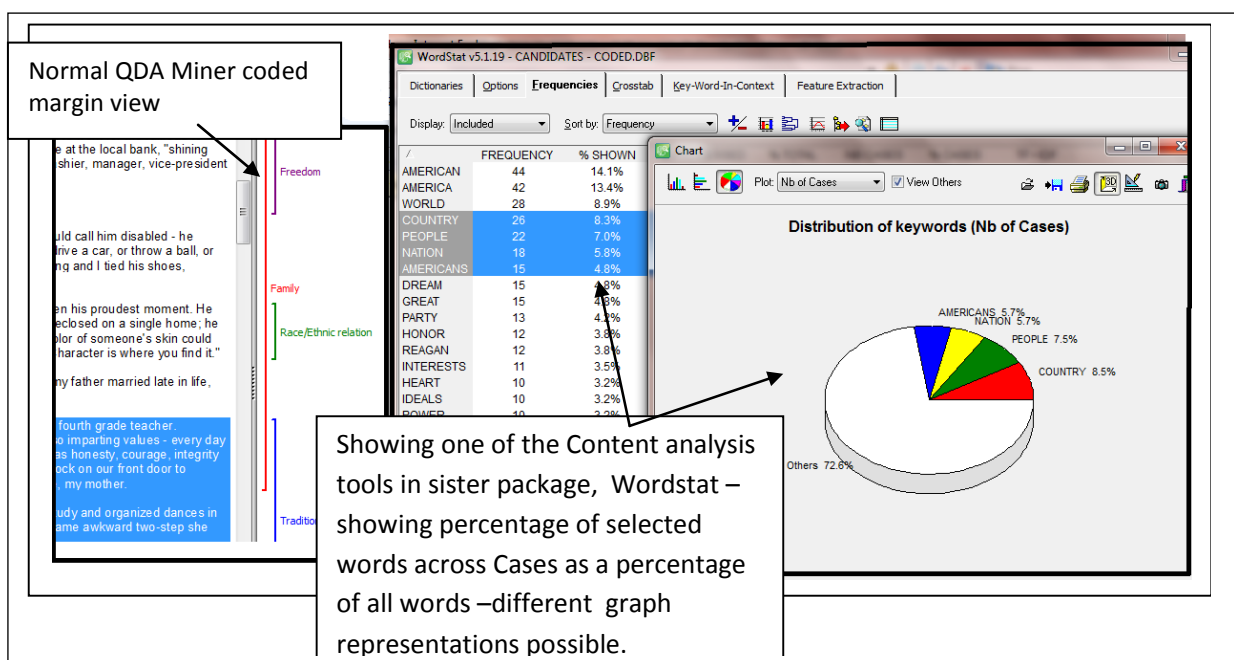
Minimum System Specifications (recommended by developer)

Windows 98, 2000 or XP, Vista (or Mac with Virtual PC, Parrallels, Boot Camp, etc.)
Memory: From 64MB (Win 98) to 512MB (Vista)
20MB of hard disk space

Figure 1. Showing QDA Miner coding window (left) and some possible types of content analysis available if WordStat is integrated (right)

Structure of work in QDA Miner

QDA Miner functions using an 'internal database structure' so documents are imported into the database and all the workings are held in about six project files. This will vary according to the types of data being stored. Being designed for the management of mixed-methods projects, QDA Miner and its sister programs can handle large amounts of data in a variety of formats ■ Data can be imported or a project file can originate and be structured directly from spreadsheet applications, when for example analysing open-ended survey questions ■ The ability to archive and compress the project file facilitates the backing up and movement of projects. The compression reduces file size to between 15% and 20% of the original. ■ The QDA Miner interface is divided up into resizable areas, with lists of variables, documents and codes and the main document window (with coded margin) are viewable simultaneously throughout. Other windows providing output float on top of the main interface - see Figure 1.

Variables structure: QDA Miner operates using a cases and variables structure which is unusual amongst CAQDAS packages. The variable values of the currently selected case are displayed adjacent to the data along with the case from which the data derives and the main codes list (see Figure 1) ■ A new project can be created in number of different ways, for example, from a list of documents; from a database/spreadsheet containing quantitative variables; using a document converter, etc. ■ Variables act as holders for different types of qualitative/categorical and quantitative information. New document type variables can be added at any time to handle additional types of data, e.g. notes for each case. Variables of various types (e.g. numeric, nominal, ordinal, Boolean, string) are used for case filtering and comparison. This information can be imported from a spreadsheet.

Data types and format in QDA Miner

Textual formats: QDA Miner directly stores text saved as 'rich text format' (rtf) and ASCII (txt). However, the Document Conversion Wizard can convert various file formats such as MS Word, WordPerfect, HTML, Adobe Acrobat to work with inside the software. Textual data is fully editable using standard Windows formatting toolbars. Objects such as tables and graphic elements can be embedded into rtf documents ■ QDA Miner can also convert projects created in other CAQDAS packages, including ATLAS.ti, HyperRESEARCH and NVivo.

Database formats: several database and spreadsheet file formats (including MS Access, Excel, dBase, Paradox) and any data file with an ODBC driver (Oracle, MS SQL, etc.) can also be imported.

Closeness to data and interactivity in QDA Miner

Documents open with one click from the main listing (see Figure 1). Only one document can be viewed at a time but all the documents for the current case are accessible through a tabbed presentation ■ Textual data line-wraps in the document window. The codes margin area provides interactive contact with coded text ■ There is good synchronicity between different displays and windows, for example, one click on a text segment in any coded retrieval window or output connects with source data and highlights segment within its source context.

Coding scheme in QDA Miner

The QDA Miner coding schema can be completely un-hierarchical (with the proviso that there has to be one top level code, underneath which all the others hang) or several levels of hierarchy can be used to organize codes ■ The coding schema can be re-organized with drag and drop to move codes within and between categories ■ The margin view displays colour-codeable codes as defined by the user; any Windows colour can be applied to any code. Viewing coloured codes in the margin allows comparison of codes assigned for example, by different coders, or according to broad themes ■ Codes can be edited, annotated, resized, removed etc. from the margin view. Codebooks can be imported from other projects.

Coding processes in QDA Miner

Codes can be assigned to any segment of text, to one or several table cells, or a whole graphic or another embedded object ■ Drag and drop codes on to text or double click on a code to assign it to selection ■ Codes can be applied to whole paragraphs without manual text selection by drag and drop.

Basic retrieval of coded data in QDA Miner

Basic retrieval of coded text is enabled at single and/or multiple codes. This happens in a dialogue box where boxes adjacent to required codes are ticked   ■   Other/new codes can be assigned to any group of retrieved coded segments ■ Tabular output can be generated to list codes providing description and frequency across documents and cases ■ Output of coded data is formatted in interactive tables – click in a cell and the full context is highlighted in main window.

Data organisation in QDA Miner

Documents are organized by Cases for filtering purposes. Cases can be grouped and ordered according to selected variable values, and these orderings follow through to generated outputs. Any numerical, categorical, logical and date data may be used to categorize cases, which can be automatically assigned upon data import. QDA Miner can handle more than 2000 variables and several million cases.

Writing tools in QDA Miner

Annotations can be attached to codes in the code list or at the code mark in the margin area. Coded segments can also be annotated ■ A code-able notes file is automatically created and linked to each case. Editable memo files can be created as documents, and then coded, annotated and searched etc., like any other file by creating new document type variable ■ A Report Manager inside the software allows user to collate useful tables, charts, results and files in a folder structure that can be individually customised to help structure the write-up (see Figure 2).

Figure 2: Showing the Report Manager and its folders - displaying a selected Text retrieval search stored within the folder structure.
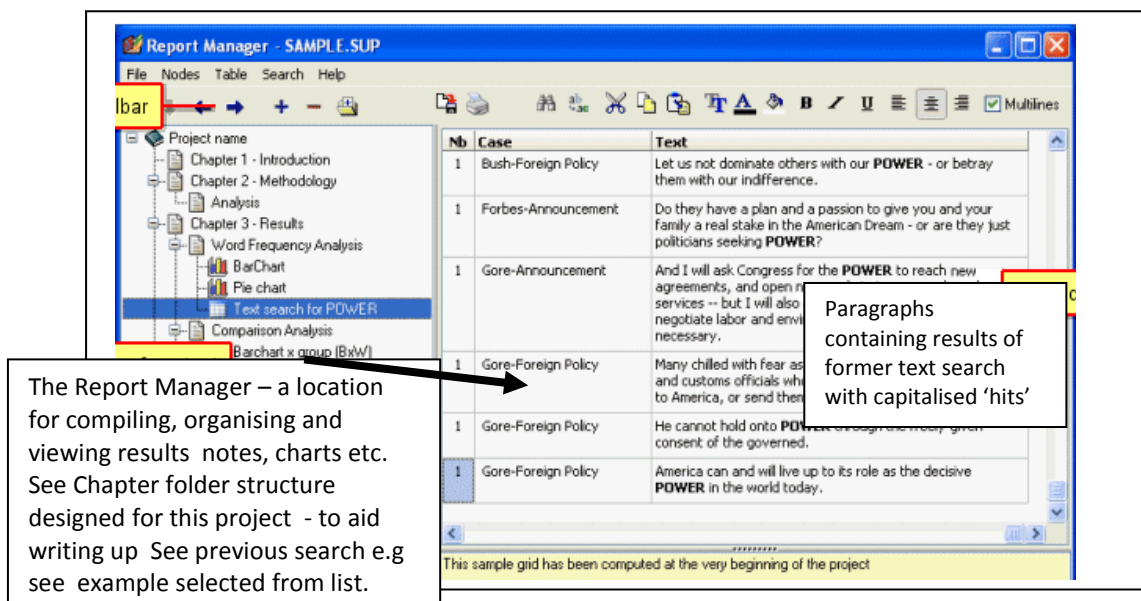
Searching and interrogating the dataset in QDA Miner

Text searching tools: ■ a Thesaurus facility associated with user created collections of synonyms (into word categories) allows the program to return a hit every time any one of the words or phrase associated with or previously coded category item is found ■ Search different units of text; the whole document, a paragraph, a sentence text segments associated with a specific or several codes

The Code Sequence retrieval tool: exploration of common code sequences to gauge some idea of what codes are following each other in the text. All or selected codes can be requested and overlapping codes can be included or excluded as required. ■ A tabular list of sequences is produced between pairs of codes (in both orders – Code A followed by Code B and vice versa) together with frequency or cross tabulations of the recurring sequences, say across cases. This then allows the user to select a pair of codes in a particular sequence producing a result window, listing full text of each coded segment for both codes -interactively

connected to full context in source file ■ New coding can be applied to one result or all results, in the results pane.

Auto-coding/text searching: interactive tabular list of hits highlights the original text segment ■   Coding may be done on selected hits or on the entire list of hits (see Figure 2). ■ The Section Retrieval tool searches for and codes sections repeated within structured documents.

Output in QDA Miner

Tabular outputs can be printed or exported to Excel, HTML, comma or tab-delimited files. Coding retrieval results may also be exported as a new QDA Miner project   ■   Text reports may be saved to disk in Rich Text, MS Word™, ASCII or HTML format   ■   Various graphical displays can be generated and exported, for example, cluster plots, heatmaps etc. ■   Export to BMP, WMF, PNG and JPG format ■ Whole project export to spreadsheet and database file formats e.g. Quattro Pro, Lotus, Excel, Paradox, dBase.

WordStat and SimStat modules used in combination with QDA Miner

WordStat is an optional quantitative content analysis and text mining module (See Figure 1). It can be opened from within a QDA Miner project to analyze words and phrases found in specific documents or in selected coded segments. It performs simple descriptive analyses of word occurrences or can investigate relationships between words or categories of words and other numeric or categorical variables - with lemmatization, stemming, stop lists, hierarchical categorization of words, word patterns and phrase. ■ Generate Key Word In Context (KWIC) lists, links to lexical databases, vocabulary and phrase extractors ■ Analysis of case or document similarities - clusters of cases based on their similarity of coding.

SimStat is a statistical analysis application which can be used alongside QDA Miner to perform quantitative analysis on any numerical data in a QDA Miner project ■   It can also perform numeric and alphanumeric computations, transformation and recoding of variables, as well as advanced file management procedures such as data file merging, file aggregation, etc.

Team-working in QDA Miner

Multi-user logon allows several coders to work on the same project (not concurrently) with either full access or to limited features and no access to other coders work.   Several users may also work on different copies of a project or different subsets of cases. ■   The merge feature consolidates several project files into master project. ■ The inter-coder reliability tool computes agreement measures and the identification of common disagreements.

Comment on QDA Miner

Simple functions in QDA Miner are handled really well. For example colours for codes as they appear in the margin, can be based on meaning or on differential colour coding for a team members (this function still missing in some other software programs). In single or multiple code retrieval just tick the codes wanted, for clearly labeled retrieval. The synchronization between retrieval and query reports and the main display windows is good, enabling full context to be seen for any retrieved segment along with the values of key variables. Retrieval format is interactive, tidy and compact tabular format. The way query and retrieval dialogs are kept in tabbed pages beside their outputs really facilitates refining and improving queries.

The Code Sequence Analyzer – a very useful and unique tool – easy to use and replaces what would be a complicated query in other software packages. This offers the user the ability to not only list co-occurring or overlapping codes in the data, but separated sequential paired occurrences of all codes with each other across the whole dataset. Results tables are interactively connected to source data.

The central 'variable-based' structure of the software is unique in comparison to other packages and QDA Miner is potentially useful tool for mixed methods approaches. The text mining terminology in the program

can be a bit demanding however. A researcher working interpretatively could make good use of the software, but the package is perhaps more difficult if your focus and experience is very qualitative. The cross over between quantitative text mining, statistical and conventional qualitative functionality, make this potentially ideal for those who also need to have one foot in the more quantitative approaches to qualitative data. Its potential to integrate other analytic elements with more interpretive methods is underlined by its capacity to convert projects from CAQDAS packages including ATLAS.ti, HyperRESEARCH and NVivo.

The case and variable structure may affect the way data is viewed *particularly with open ended questions (i.e. data imported straight from a database or spread sheet) in that while navigating through the data associated with a case you can only see one case's response to a single question at a time rather than the whole context of that case.*

Thesaurus facility and the search category tool *extends the potential of the text search tool by catching meaningfully similar words.*

The export features are well organized *and easy to use. When viewing the results table of a coded retrieval, the report is created at one click of an icon. An option to export coded segments and in the process create a new project with them, is unusual and makes good data reduction sense when dealing with hundreds or thousands of documents or records. The exported code labels become case names, and the segments are contained within a document type variable for each case.*

The Report Manager is a unique tool – *a location internal to the software to house reports and outputs, charts, retrieved segments, memos etc., from current work in QDA Miner. Its folder structure can be designed to suit the user and report files moved around within them. A one click button appends a current output or result to the Report Manager. This is an exceptional compiler and organizer of results from which to write-up.*

Further Reading

- Gibbs, G. (2007) **Analysing Qualitative Data**, part of the Qualitative Research Kit, ed. U. Flick, Sage, London
- di Gregorio, S & Davidson J (2008) **Qualitative Research for Software Users**, McGraw Hill, Open University Press, UK
- Lewins & Silver (2007) *Using Software in Qualitative Research: A Step-by-Step Guide*, Sage Publications
- Lewins, A. (2008) 'CAQDAS: Computer Assisted Qualitative Data Analysis' in (ed) N. Gilbert, **Researching Social Life** (3rd ed), Sage, London
- Lewis and Maas (2007) **QDA Miner 2.0: Mixed-Model Qualitative Data Analysis Software** *Field Methods* 19: 87-108
- Silver, C., & Fielding, N. (2008) 'Using Computer Packages in Qualitative Research', in Willig C & Stainton-Rogers W (eds.) **The Sage Handbook of Qualitative Research in Psychology**, London, Sage Publication