# QDA Miner 6 - Distinguishing features

This document is meant to be read in conjunction with the 'Choosing a CAQDAS Package Working Paper' which provides general commentary of common CAQDAS functionality. This document reviews QDA Miner version 6 which was released in September 2020. This review does not provide an exhaustive account of all QDA Miner's functionality but highlights some of its distinguishing features. Please also see our review of WordStat, content analysis and text mining software also developed by Provalis Research which can be used in combination with QDA Miner. The Comment section at the end of this document details our opinions on certain aspects of functionality and usability. See also Silver & Lewins (2014) *Using Software in Qualitative Research: A Step-by-Step Guide*, Sage Publications and the Provalis Research website. Thanks to Normand Péladeau and Tony Ross for checking the accuracy of this document.

Background   https://provalisresearch.com/products/qualitative-data-analysis-software/   Developed by Normand Péladeau and colleagues at Provalis Research (Montreal, Canada), QDA Miner was released in 2004 as a mixed-methods software providing quantitative tools for analysing qualitative data as well as the common qualitative tools for qualitative data analysis. QDA Miner is part of a 'Prosuite' of three programs developed by Provalis Research that provide integrated text analysis tools - the other two programs are WordStat (content analysis and text mining software – which we review separately) and SimStat (statistical analysis and bootstrapping software). The Prosuite extends the range of exploration and analytic techniques allowing users to cross methodological boundaries. QDA Miner Lite is a free version (with a reduced set of features). For information about the differences between the Full and Lite versions see the feature comparison page on the Provalis Research website.

Minimum System Specifications (recommended by developer)   Operating System: Microsoft Windows XP, 2000, Vista, Windows 7, 8 and 10 ■ Memory: From 256 MB (XP) to 1GB (Vista, Windows 7, 8 and 10) ■ Disk Space: 150 MB of disk space (200 MB for Prosuite). ■ MAC OS & Linux. QDA Miner will run on a Mac OS using virtual machine solution or Boot Camp, and on Linux computers using CrossOver or Wine. For information about running QDA Miner on a Mac computer see https://provalisresearch.com/products/mac-os/

Figure 1. The main QDA Miner interface
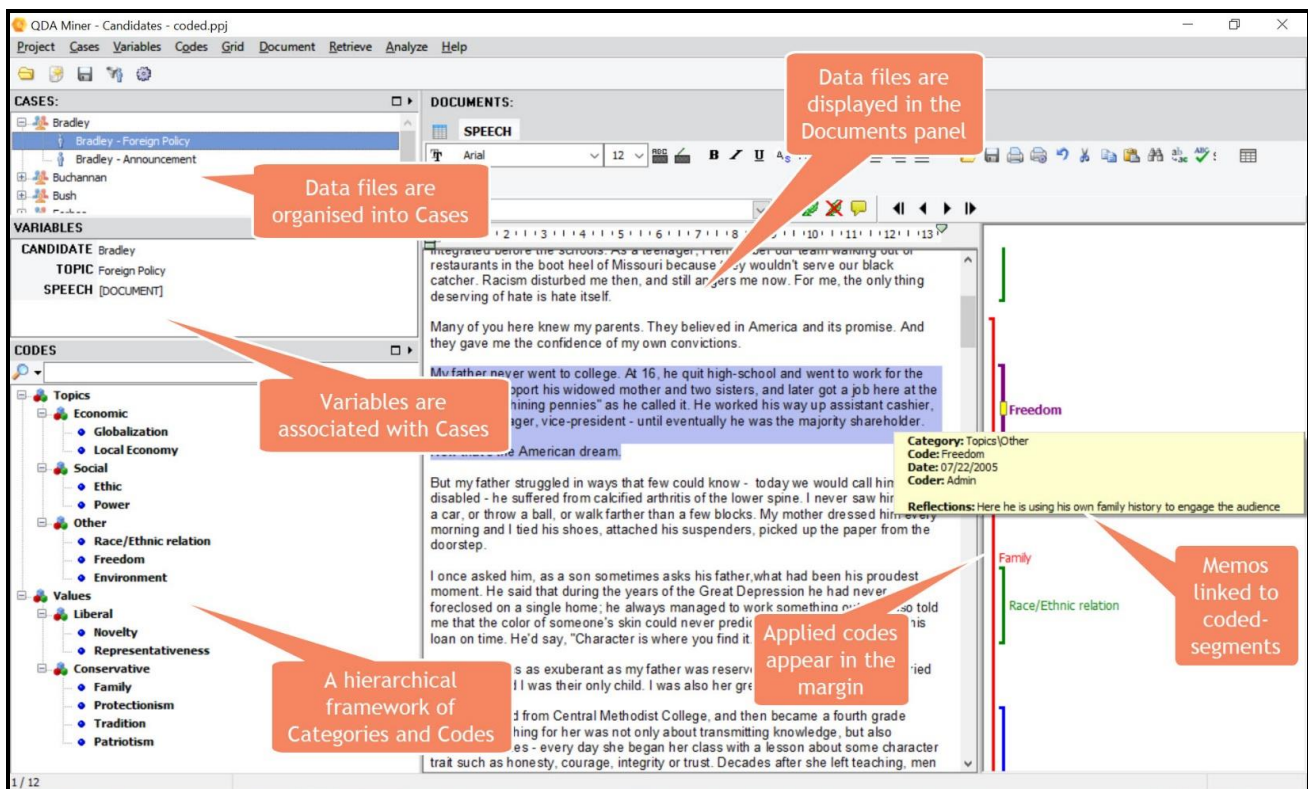
Structure of work in QDA Miner   QDA Miner functions uses an 'internal database structure' so data files are imported into the project file ■ The QDA Miner interface is divided up into resizable areas, the lists of cases, variables and codes and the main document window (with code margin) are viewable simultaneously throughout (see Figure 1). Other windows providing output float on top of the main interface ■ Cases and variables structure: QDA Miner operates using a cases and variables structure whereby Cases comprise one or more Documents (data files) and can have factual characteristics associated with them (e.g. socio-demographic or metadata information) ■ New projects can be created in number of ways, for example,

from a list of documents; from a database/spreadsheet containing both qualitative material and quantitative variables; using a document converter, etc.

Data types and format in QDA Miner

Being designed for mixed-methods projects, QDA Miner and its sister programs can handle large amounts of data in a variety of formats ■ Textual formats: QDA Miner allows the direct import and storage of textual data in a variety of formats, including *.txt; *.rtf; *.doc; *.docx; *.pdf; *.odt *.epub; *.ppt; *.pptx; .xps; and *.html. Textual data is fully editable using standard Windows formatting toolbars. Objects such as tables and graphic elements can be embedded into textual documents ■ Image formats: Range of graphic file formats can be imported, including *.bmp; *.wmf;, *.jpg; *.jpeg, *.gif and *.png. ■ Database formats: several database and spreadsheet file formats (including MS Access, Excel, dBase, Stata, SPSS) and any data file with an ODBC driver (Oracle, MS SQL, etc.) can also be imported. ■ Reference Management formats: Direct import available from Endnote, Mendeley, Zotero and in *.ris format ■ Web Surveys: Direct import from a range of online survey collection tools, including Qualtrics, SurveyMonkey, Survey Gizmo, and Question Pro ■ Online sources: Direct import from a variety of social media platforms, including Twitter, Facebook, Reddit, and YouTube ■ Other data accepted by QDA Miner include from Outlook, Gmail, LexisNexis and Factiva. For a full list of acceptable data types and formats see the Provalis website.

Closeness to data and interactivity in QDA Miner

Documents open with one click from the main listing. By default, one document is viewed at a time but all the documents for the current case are accessible through a tabbed presentation ■ The Grid View allows all documents to be viewed and analysed in a single table ■ Textual data line-wraps in the document window ■ The code 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 the segment within its source context ■ Variables and their values associated with Cases and their corresponding Documents are always on view in the main interface providing relevant contextual information.

Coding scheme in QDA Miner

The QDA Miner coding schema comprises a hierarchical structure of Categories and Codes. It can be completely un-hierarchical (with the proviso that there has to be one top level category, 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 code brackets with colours defined by the user. Viewing coloured codes in the margin allows comparison of codes assigned for example, by different coders, or according to different categories or themes ■ Codes can be edited, annotated, resized, removed etc. from the main listing and from the margin view ■ Codebooks can be imported from other QDA Miner or WordStat projects, and from other CAQDAS packages, using the common CAQDAS exchange standard (REFI-QDA).

Coding processes in QDA Miner

Manual coding: Codes can be assigned to any segment of text, to one or several table cells, to part of or a whole graphic image or another embedded object ■ Drag-and-drop or double click to assign codes to selections ■ Codes can be applied to whole paragraphs without manual text selection by drag-and-drop ■ Grid View offers alternative way of displaying and coding responses stored as Documents ■ **Computer Assistance with coding:** QDA Miner also includes several powerful tools that provide coding suggestions to the user which can be accepted or rejected as relevant. These include Cluster Extraction which automatically groups similar text segments to facilitate identification of data that likely relates to analogous concepts or topics and Query by example which finds similar text passages from a user-specified example, and suggests other text segments as candidates for coding in descending order of similarity. The tool learns from the user marking finds as relevant or irrelevant and results are then further refined ■ **Geotagging:** QDA Miner also allows Documents or coded segments (textual or image) to be associated with geographic locations and if relevant for a specific time-point or -period to be added to situate data in both time and space dimensions. These can be visualised in maps and time-line views.

## Basic retrieval of coded data in QDA Miner

Basic retrieval of coded data 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 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 ■ Filtering also provides a straightforward and flexible way to retrieve coded data for sub-sets of Cases or in relation to how documents have been coded.

Data organisation in QDA Miner

Data files (Documents) are organized by Cases and Variables ■ Cases can be grouped, ordered and filtered 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 ■ Variables store different types of qualitative/categorical and quantitative information and can be 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 or added within a project at any point. 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 ■ New variables can be added at any time to handle additional information Variables types can be changed, and their values transformed or recoded (including numeric transformation, binning, date transformation) ■ The variable values of the currently selected Case are always displayed adjacent to the data along with the Case from which the data derives and the main codes list (see Figure 1)

Figure 2. Cluster extraction, Quote Matrix and Report Manager in QDA Miner
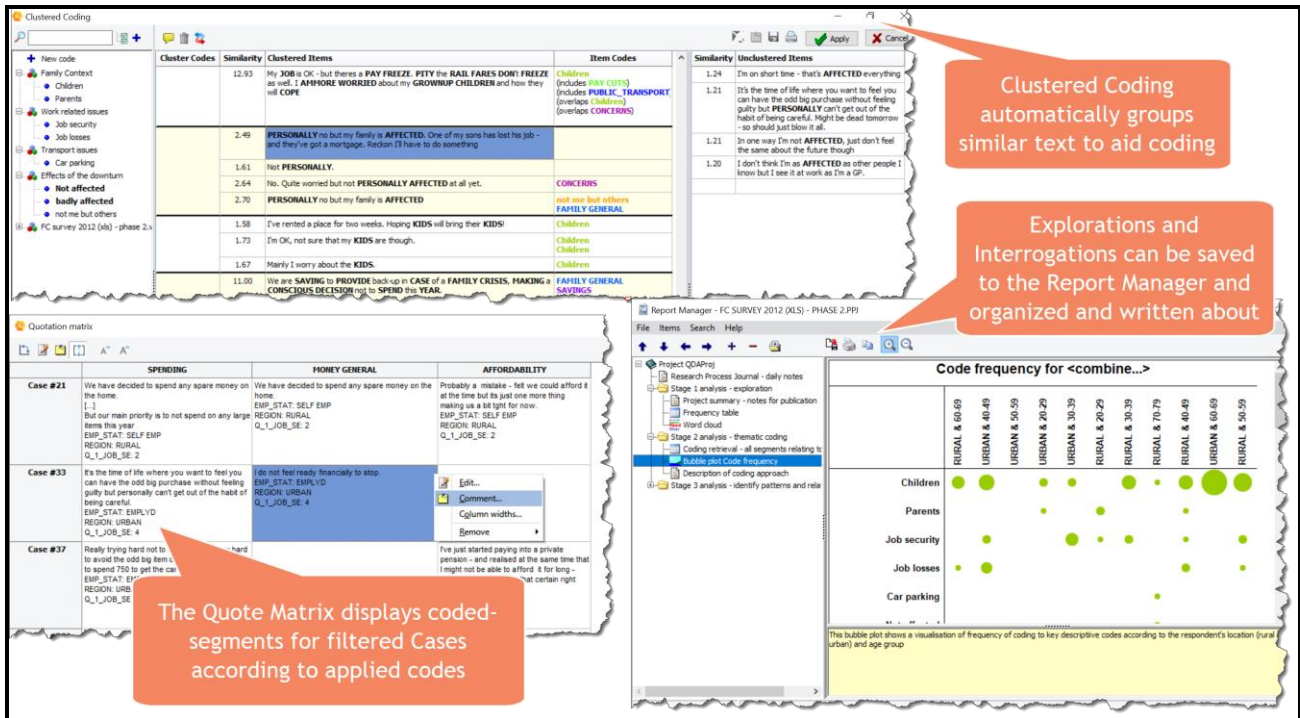
Writing tools in QDA Miner

There are several spaces to write within QDA Miner ■ Code Descriptions: Codes can be defined in associated Description field ■ Comments: Up to 6 separate comments can be attached to each coded segment from the margin area ■ Central Note: One area for writing about the whole analysis project, for example acting as a research process journal ■ Documents are fully editable, so associating a new Document Variable to a Case acts as a memo which can be coded, annotated and searched etc., like any other Document ■ The Report Manager is fully integrated, allowing the user to collate useful tables, charts, results and files in a folder structure that can be individually customised to help organise the write-up (see Figure 2).

Text Searching in QDA Miner

There are many ways to search textual data in QDA Miner ■ Text searching tools include Text Retrieval, Keyword Retrieval, Section Retrieval and Cluster Extraction (see Figure 3.). A Thesaurus facility associated with user-created collections of synonyms (called Categories) allows the program to return a hit every time any of the words or phrases associated with a category is found ■ Search different units of text including the whole document, a paragraph, a sentence, or previously coded segments ■ Auto-coding. Selected hits in retrieved searches are displayed in interactive tabular lists with adjustable amounts of surrounding context and including selected Variables. Selected hits or the entire list of hits can be auto-coded as required ■ The Section Retrieval tool searches for and codes sections repeated within structured documents.

Interrogating the dataset in QDA Miner

There are many ways to interrogate patterns and relationships between codes and variables using QDA Miner's query tools. These include Coding Frequency, Coding Co-occurrences, Case Similarity, Coding Sequences, Coding by Variables and Code Combination ■ All queries can be filtered by Case, Documents and Variables and some can additionally be scoped to coded segments associated with selected Codes. User-specified query criteria can be saved to be re-run at a later stage (either by saving an external query file or by accessing from the Command Log). Results are displayed in interactive floating windows and each interrogation includes alternative visualisations (for some examples see Figure 3.). New coding can be applied from the results windows ■ Coding Sequences extracts and displays common code sequences for any or selected codes to illustrate which codes follow one another and where within the source Documents. Results are presented in a Frequency Matrix and a Frequency List (in both orders – Code A followed by Code B and vice versa) and results can be accessed within their original context from the results ■ Case Similarity and Coding Co-occurrences interrogate relationships between codes by computing similarity and performing cluster analysis and multidimensional scaling on selected codes. When interrogating code co-occurrences the user specifies the basis of the co-occurrence (e.g. within case, within a specified number of paragraphs, or in overlapping segments) and the statistical similarly index to be used (e.g. Jaccard's,

Sorensen's, or Ochiai's coefficient). Case clustering computer the relative frequency of selected codes amongst Cases using cosine coefficients. The user specifies the indicator to be used for clustering (e.g. Code co-occurrence, code frequencies, number of words or percentage of words. Results are listed in tabular format and can be visualized in several ways (see for example Co-occurrence Heatmap and Dendogram in Figure 3.)

Visual tools in QDA Miner   QDA Miner has a large range of 2D and 3D visual tools for exploring, interrogating and displaying data and analytic findings (see Figure 3.). With the exception of multidimensional scaling plots, all visualisations are interactively linked to the underlying qualitative data.

Figure 3: Selected visual tools in QDA Miner

Word Cloud

Link Analysis

Code Frequency Chart

Co-occurrence HeatMap

Dendogram

Pie Chart in Report Manager
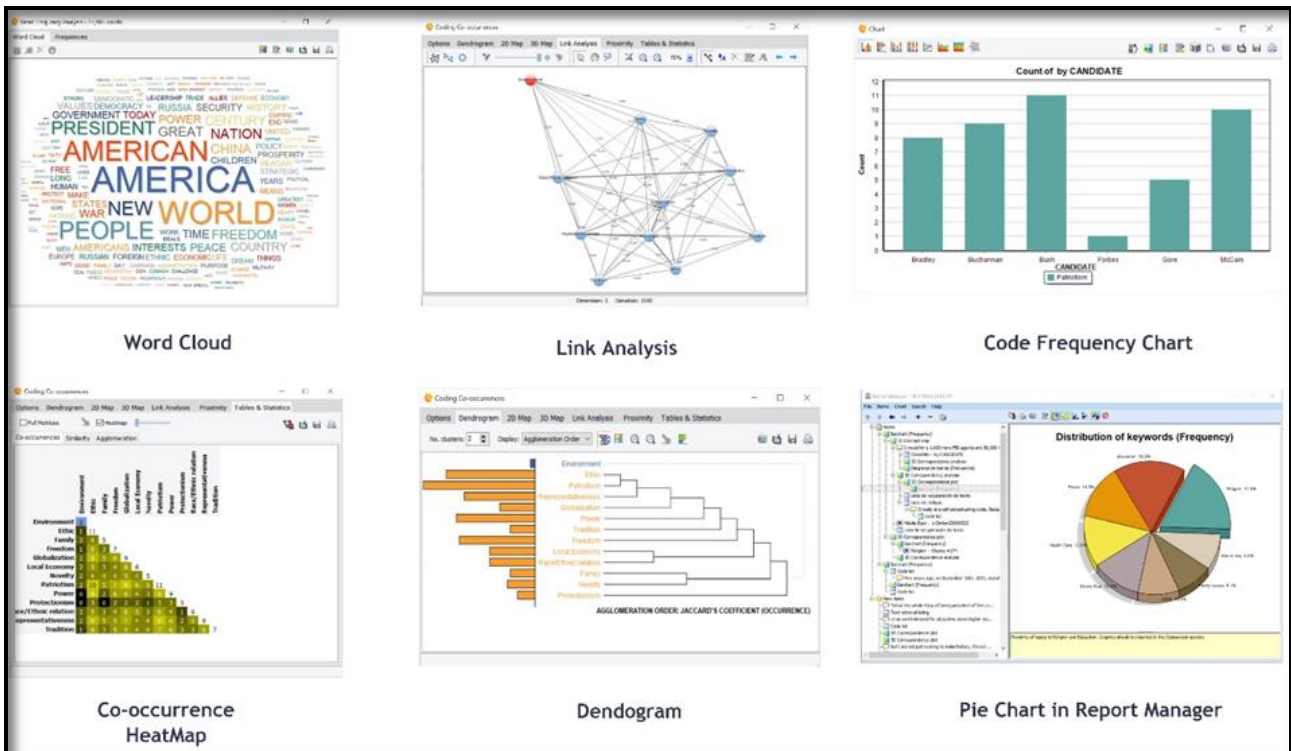
Output in QDA Miner   There are many ways to generate output from QDA Miner ■ Tabular outputs can be printed or exported to Excel, HTML, comma or tab-delimited files and exported directly to the data visualization platform, Tableau if other visualisations are required. Coding retrieval results may also be exported as a new QDA Miner project ■ Text reports may be saved to disk in Rich Text, MS Word, ASCII or HTML format ■ Visualisations such as those illustrated in Figure 3 can all be exported in several formats, e.g. BMP, WMF, PNG and JPG ■ Whole project export to spreadsheet and database file formats e.g. Excel, dBase, SPSS, Stata.

WordStat and SimStat used in combination with QDA Miner   WordStat is a content analysis and text mining program which as well as being a standalone software, can be opened from within a QDA Miner project to explore and analyse text found in specific documents or in selected coded segments. It performs descriptive analyses of word occurrences or can investigate relationships between words or categories of words and other numeric or categorical variables - including lemmatization, stemming, stop lists, hierarchical categorization of words, word patterns and phrases ■ SimStat is a statistical analysis program which can be used alongside QDA Miner and WordStat 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. (note many numeric transformation and re-coding tasks can also be performed in QDA Miner and WordStat).

Team-working in QDA Miner   Although QDA Miner is a single user program, it has a Multi-user logon that allows several analysts to work on the same QDA Miner project (not concurrently) with either full or restricted access (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 one project ■ The Command Log feature allows changes made by different team-members to be tracked which also allows earlier interrogations to be performed again if updated results are

required in, for example, a longitudinal 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 manual coding processes, annotating coded-segments and straightforward retrievals of coded data. In particular, the ability to apply multiple comments to an individual coded-segment and to organize them both by labelling and applying colour is particularly useful and also unique.*

- *Both straightforward and complex retrieval of coded data is easy to accomplish without recourse to learning how to build complicated queries, because of the simple tick-box logic in dialogue boxes.*

- *Synchronicity between retrieval windows and the main display windows is excellent, enabling full context to be seen for any retrieved segment along with the values of key variables. Retrieval formats are also fully interactive, tidy, clearly labelled and compact. The way query and retrieval dialogs are kept in tabbed pages in the same window as their outputs greatly facilitates refining and improving queries.*

- *The Code Sequence Analyzer is a very useful and unique tool that is 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 – or filtered subsets of the it.*

- *Powerful and flexible tools to assist with coding. QDA Miner has a number of very powerful and sophisticated tools that assist the user with coding which are both transparent and adaptable. As a result an excellent balance between computer-assistance (including some machine learning) with the value of human interpretation is achieved.*

- *Excellent methodological crossover between quantitative, text mining and conventional qualitative functionality makes QDA Miner ideal for mixed-methods analysis and projects taking more quantitative approaches to qualitative data. Nevertheless, QDA Miner is also ideal for projects at the purely qualitative and interpretive end of the methodological spectrum and is therefore a good choice for researchers who engage in a range of project types and use different analytic methods.*

- *The central 'variable-based' structure of the software is unique and QDA Miner is a particularly useful tool for mixed-methods analysis and when working with large sets of textual and image data. However, for some purposes (e.g. longitudinal projects) it can be restrictive that variables can only be associated with cases and not the individual documents that relate to them. This architectural structure may also affect the way data is viewed particularly with open ended survey 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.*

- *The thesaurus facility and the search category tool are powerful and easy to use and together extend the potential of text searching by catching interpretively meaningful similar words.*

- *The export features are well organized and easy to use. When viewing the results table of a coded retrieval, output in a variety of formats can be 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 – this is particularly useful for secondary analysis or for projects that need to analyse datasets from multiple perspectives.*

- *The Report Manager is a unique and incredibly useful 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.*

- *Geocoding is implemented powerfully and it is unique to be able to fully integrate space and time through the linking of geographical locations and corresponding temporal dimensions. Visualisations for working with the results are also engaging and dynamic.*

Further Reading

- Lewins & Silver (2014) *Using Software in Qualitative Research: A Step-by-Step Guide*, Sage Publications
- Silver C. (2014) '**Software Review: QDA Miner (WordStat and SimStat)**'. Journal for Mixed Methods Research 1(2)
- Lewis, R. B., & Maas, S. M. (2007). **QDA Miner 2.0: Mixed-model qualitative data analysis software**. Field methods, 19(1), 87-108.
- O'Kane, P., Smith, A., & Lerman, M. P. (2019). **Building transparency and trustworthiness in inductive research through computer-aided qualitative data analysis software**. Organizational Research Methods, 1094428119865016.
- Péladeau, N. (in press, forthcoming 2021). **Mixing Beyond Mixed Methods: QDA Miner, SimStat and WordStat**. In A. Onwuegbuzie & B. Johnson (Eds). The Routledge Reviewer's Guide to Mixed Methods Analysis. Routledge.
- Rudd, M. A. (2017). **What a Decade (2006–15) of Journal Abstracts Can Tell Us about Trends in Ocean and Coastal Sustainability Challenges and Solutions**. Frontiers in Marine Science, 4, 170.
- Du Plessis, C. (2017). **The role of content marketing in social media content communities**. South African Journal of Information Management, 19(1), 1-7.