Digital literacy training

NVivo11

Advanced Coding and Data Analysis

2018
# Table of Contents

**NVivo Advanced Coding and Data Analysis** ................................................................. 1

**Deciding on a Coding Scheme: Word Frequency Queries** ........................................... 1

- Word Frequency Results ............................................................................................... 2
- Excluding Irrelevant Words from Searches ................................................................. 2
- Word Clouds .................................................................................................................. 2
- Tree Map ....................................................................................................................... 3
- Cluster Analysis ........................................................................................................... 3
- Looking at the Results in Context ................................................................................ 4

**Initial Mass Coding: Using Text Search Queries** ....................................................... 4

- Text Query Results Preview ......................................................................................... 5
- Saving the Query Preview Results as a Node .............................................................. 5
- Expanding the Context of References in Results or a Node ....................................... 5
- Word Trees .................................................................................................................... 6

**Analysis** ....................................................................................................................... 7

- Coding Queries ............................................................................................................ 7
- Basic Coding Queries: Wizard ...................................................................................... 7
- Advanced Coding Queries ........................................................................................... 7
- Editing a Query .............................................................................................................. 8
- Compound Queries ....................................................................................................... 8
- Creating a Compound Query ....................................................................................... 8
- Matrix Coding Queries ............................................................................................... 9
- Creating a Basic Matrix Coding Query ....................................................................... 10
- Saving Matrices ......................................................................................................... 10
- Editing Node Matrices ............................................................................................... 10
- Exporting Node Matrices ........................................................................................... 11
- Cluster Analysis ......................................................................................................... 11
- Editing Cluster Analysis Results ............................................................................... 13

**Maps** .............................................................................................................................. 13

- Concept Maps ............................................................................................................ 13
- Add a note to a model ............................................................................................... 14
- Project Maps ............................................................................................................... 14

**Graphs** .......................................................................................................................... 15

- Explore Diagrams ...................................................................................................... 15
- Comparison Diagrams ............................................................................................... 15

**Exporting Results and Project Items** ..................................................................... 16

- Exporting Classification Sheets ............................................................................... 16
- Exporting Nodes (Including the Entire Content) ......................................................... 17
- Exporting the Codebook ............................................................................................ 17
- Exporting the List View of the Nodes Area ............................................................... 17
- Running Reports and Extracts ............................................................................... 17

**Collaboration** .............................................................................................................. 18

**Interrater Reliability** ................................................................................................. 19

**Mac vs PC Versions of NVivo** .................................................................................. 19

**Other resources** ........................................................................................................ 20

---

**Digital literacy training**
NVivo Advanced Coding and Data Analysis

This course will show you some more advanced NVivo features and how to use the most commonly needed NVivo tools to analyse your qualitative data. For further information about NVivo’s full capabilities, please refer to the NVivo 11 help website: http://help-nv11.qsrinternational.com/

This session will show you how to:

- use **Word Frequency Queries** to decide on a coding scheme
- initially mass Code your text data using **Text Search Queries**
- analyse and explore your data further using **Queries**: Coding, Compound and Matrix
- analyse your data using **Cluster Analysis**
- **Export** results and project items, and run **Reports** and **Extracts**
- create and edit special **NVivo Graphs**
- create **Maps** to visualise your project structure
- **Collaborate** in an NVivo Project

Open the sample project to explore how NVivo can be used to analyse your data. When you open NVivo, at the top of the **Welcome Screen** there is an option to open a new **Sample Project**. Select the sample project which is a two year study (2008-2009) documenting community perceptions of development and land-use change on coastal communities in the Down East area of Carteret County, North Carolina, USA.

**Important**: Save the Project under a new name so you can alter the data without changing the sample project for other computer users. Go to **File** choose **Copy Project** and save the file as **Sample Project (2)** in your **My Documents** folder. Saved Projects can be opened on any computer with NVivo and can be altered any collaborator (a record of who changes what is kept according to User ID).

**Deciding on a Coding Scheme:**
**Word Frequency Queries**

In the initial stages of deciding on a coding scheme to use, a **Word Frequency Query** could be useful. This will produce results of the words that occur the most often in the text documents you have (including memos, transcripts etc.). You can search the entire content of your qualitative data, or you can narrow it down to specific groups of people or individual questions, etc..

**Example: Word Frequency Query for Interview Question 4: "Community & Environmental Change".**

On the **Query** tab, click on **Word Frequency**. This will open a new **Detail View** with the word frequency options. Choose where to search in the **Search in** (e.g., **Selected Items** and look in the **Nodes** folder, then within the **Autocoded Responses** folder select **Question 4**). Then decide if you want to **limit the number of results** (by default to the 1000 most frequent) and the largest **letter-length** of the words you want to limit results to (default is 3 letters so you don’t get words like “to”, “a”, “an” or “I”). In the **Grouping Area**, decide on the grouping criteria by sliding the indicator (e.g., **With synonyms** - see table below for descriptions of the 5 levels).

**Note**: The higher the similarity, the longer it will take to run.
The five **Text Match levels:**

<table>
<thead>
<tr>
<th>Grouping Level</th>
<th>Returns</th>
<th>Example (sport)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact matches</td>
<td>Exact matches only</td>
<td>sport</td>
</tr>
<tr>
<td>With stemmed words</td>
<td>+ Words with same stem</td>
<td>sport, sporting</td>
</tr>
<tr>
<td>With synonyms</td>
<td>+ Synonyms</td>
<td>sport, sporting, play, fun</td>
</tr>
<tr>
<td>With specializations</td>
<td>+ Specializations</td>
<td>sport, sporting, play, fun, running, basketball</td>
</tr>
<tr>
<td>With generalizations</td>
<td>+ Generalizations</td>
<td>sport, sporting, play, fun, running, basketball, recreation, business</td>
</tr>
</tbody>
</table>

If you want to save the Query details and run it again later from the **Queries** Area, click **Add to Project**, give the Query a name and select **OK**. Then click **Run Query**.

**Note:** A Text/Word Search Query does not search text within PDFs created by scanning paper documents (each page is a single image). Consider using optical character recognition (OCR) to convert the scanned images to text before importing.

**Note:** Only words in the content fields are included in Queries when searching audio and video transcripts and only in codable fields in datasets.

### Word Frequency Results

From the results in the tabs in the **Detail View** you can decide on an appropriate coding scheme and even begin to create Nodes. The **Summary** tab shows you a list of all the most frequent words, their frequency **Count**, other summary data and if you have chosen a **Grouping** level higher than **Exact matches**, you will also see the **Similar Words** it has included as belonging to each Word Count (e.g., the word count for Change also includes the word count for Changed, Changes, Changing, Shift). There are also tabs for graphical representations of your Query: **Word Cloud**, **Tree Map** and **Cluster Analysis** (the latter only appears if there is enough data to run it).

### Excluding Irrelevant Words from Searches

In the results you might see words that you do not want to include in the Query search in future, such as “any”. Right-Click on the word in the results and select **Add to Stop Words List**. If you have included more than exact matches in the text match level section (i.e., >level 1), you must remove any words in the box which you still wish to be included in future searches. When you are ready, click **OK**. Every time you run a Query from now on, these words will not be included in results. You can edit these in **Project Properties** at any time: on the **File** tab under **Info** select **Project Properties** and on the **General** tab click the **Stop Word List** button and edit the words to exclude.

### Word Clouds

The **Word Cloud** displays up to 100 of the most frequent words with each varying in font size: the largest fonts related to the most frequently occurring words, and the smallest the least frequent.
Tree Map

A Hierarchy Map is a diagram that shows hierarchical data as a set of nested rectangles of varying sizes. Like a Bar Chart, you can use them to compare the number of coding references: the largest rectangles represent Nodes with the most number of coding references, and the smallest rectangles the least. As with all graphics, Tree Maps provide links to actual references.

Exercise: Run a Word Frequency Query on all the Interviews.

Cluster Analysis

The Cluster Analysis tab contains a tree-diagram describing the relationships between the words in the Query based on their co-occurrence. This can help by providing more context.
Looking at the Results in Context

If you think that one of the results might be a good Node to have in your Coding Scheme, you can view the actual references found in the Query results in more context. Double-click on the word in the table, or right-click and Choose Runs Text Search Query. This runs a Text Search for all the words in the group and opens a new tab in the Detail View containing all the references containing the word(s). Double-clicking on the words in any of the graphs in a Word Frequency Query will also run this Query.

Example: Create a Node for Change from the results.

If you decide to create a Node for the principal word in your Word Frequency Query result, right-click on the word in the Detail View (e.g., Change) and select Create as Node. This will automatically code all references you see in the results to a new (or existing) Node. Please see Saving the Query Preview Results as a Node below for further details on creating Nodes from results.

Initial Mass Coding: Using Text Search Queries

Text Search Queries can be used to help start the coding process for a large amount of textual data by finding instances of text matching specific search terms based on your pre-defined coding scheme (internal text sources ONLY). Information coded to Nodes this way can always be modified later through uncoding if necessary.

Note: Consider checking the spelling before running Queries. You will find the Spelling button on the Home tab.

Example: Create an initial, automatically coded Node for Water Quality.

On the Query tab, click on Text Search and this will open up a new Detail View window with the Query options at the top. Choose where to Search in (e.g., Selected Folders) and in which folders (e.g., in Interviews and Surveys). In the Search for area, enter the word(s)/phrase you want to look for. You can create more complex Queries using the special features displayed when you click the Special button:

- **Wildcards**: If you want to search for words that have alternative spellings (e.g., colour/color), or are commonly confused with similar words (compliment/complement), Wildcards can help to find every possibility. The Wildcard * is used to stand in for one or more missing characters (e.g., g*t will find get, great, gradient etc.). The Wildcard ? is used to stand in for a single missing character (e.g., g?t will find get, gut, got etc.).

- **Fuzzy Search**: Also useful to pick up misspellings and find words with similar stems. The symbol ~ behind a word will look for similar words (e.g., analyze~ will find analyzes, analyse, etc.).

Note: If you use the above functions you will be limited to Exact match only.

- **Boolean Operators**: When you want to find a specific concept you can use a combination of words and/or excluding a word(s) in your search.
  - AND (or &) will search for BOTH terms (water AND quality).
  - OR (or |) will search for EITHER of the terms (water OR lake OR sea OR river).
  - NOT (- or !) will search for the first term and EXCLUDE any results that also contain the second (water NOT tap, water -tap).
  - REQUIRED (+), will search for items containing the required term and the second term is optionally included in results if and only if it occurs together with the first term (+Water river).
  - Double Quotes “ ” should be used to enclose a group of words to search for a specific phrase. If you search for a phrase, the results will not include words with
the same stem. For example, if you search for "alternative energy", you will not be able find "alternative renewable energies".

- **Round Brackets ( )** can be used to group clauses to form sub Queries.

- **Near**: If you want to find words that co-occur but you want to ensure they are mentioned in the same sentence rather than several sentences away, you can specify the distance either side of the required word (e.g., "water AND quality"~20) will find instances of water and quality within 20 words of each other only. 

**Note**: you need to enclose the search terms in **double quotes** before adding the ~ for a **Near Search**. In addition, the tilde (~) denotes proximity ONLY in this search.

For this example, enter "water AND quality"~20. **Spread Coding** to a **Broad Context**, so you can see enough information when you review the success of initial coding. Decide on the **match criteria** (e.g., 3). Click **Add to Project** if you want to rerun it more easily later if you obtain new data. Then click on the drop-down menu for **Run**. Select **Run and Save Results** to store the results as a new Node or merge the results into an existing one. Select **Create Results as New Node** and choose where to save the Node (the default is Results so change this to Nodes) by clicking the **Select** button next to the **Location** box. Give the new Node a name (e.g., Water Quality) and (optional) description. Then click **OK**.

**Text Query Results Preview**

In the preview of the Query results, there will be tabs on the right of the **Detail View** for:

- **Summary**: A list of all the sources that contain the word/phrase results.

- **Reference**: The results are opened as a Node preview and the word or phrase is shown with a narrow, broad or custom context (which you can expand: see previous workshop notes).

- **Text, Picture, Audio, Video, or Dataset**: Displays the results found in each type of source (only relevant tabs are available).

**Exercise**: Run a **Text Search Query** for synonyms of the words **Pollution OR Contamination**. Do not **Spread Coding** and do not save the results in order to create a Word tree.

**Saving the Query Preview Results as a Node**

In general, if you run any Query and decide retrospectively that the results in the preview would be valuable to save as initial coding for a Node, click on the Query results in the **Detail View**, then on the **Query** tab click **Store Query Results**. From the **Option** list, choose to either:

- Create the results as a **new Node**

- Merge the results into a **existing Node**

Another method is to right-click somewhere inside the Query results in the Detail View and select **Store Query Results**.

**Note**: by default new Nodes are created in the Results folder unless you choose another location (e.g., Nodes).

**Note**: NVivo will just code the word(s) found in that Query unless you expand the coding.

**Expanding the Context of References in Results or a Node**

Sometimes when you have run a Query, the context for the references in the Nodes can be quite narrow (e.g., only a few words) and you need to see more information to assess whether it can be coded in the way you think.
Within the **Query Results** you can right-click on the reference in the **Detail View**, and select **Coding Context** and choose the level you want:

- **Narrow** to see 5 words, 5 seconds or 5 percent on either side of the coding reference, or (for datasets) see all the text in the cell.
- **Broad** to see the surrounding paragraph, cell or row, or see 20 seconds or 20 percent on either side of the coding reference
- **Custom** to define specific reach settings

You can also do this through the menus by selecting the reference and on the **View** tab click **Node**, then change the **Coding Context**. To expand the context of all references in the Nodes, press **CTRL+A** to select all the text and then change the context.

If you want the coding itself to include a broader/narrower context **permanently** in a specific Nodes, when you are in that Node's **Node View** you can spread/reduce the coding, by clicking on the reference and on the **Analyze** tab click **Spread Coding**. Click the required context reach.

### Word Trees

**Word Tree** tabs are produced when you run **Text Search Queries** (only if you choose **NOT** to spread coding or save the results). The branches of the **Word Tree** represent the various contexts in which the word or phrase occurs. It can be useful for finding recurring themes of phrases that surround the word.

You can click one branch to highlight the other end of the branch to see the actual context. When on the **Word Tree** tab, a main **Word Tree** tab appears at the top of the window. This allows you to change the **Root Term** (e.g. polluted), the **Branch Order** (e.g., **Number of Matches**) and **Context** level (number of words surrounding the root term).

You can also find the project item where a particular occurrence of the word is used, by right-clicking on a branch to see the short-cut menu, then click **Run Text Search Query**.
Analysis

There are three other types of Queries, which are useful for exploring your data in more depth and quantitative analysis:

- **Coding:** explore your data more in depth based on how you coded it (e.g., What did people say about *Water Quality* and *Environmental Impacts*? What did they say about *Impacts* independent of *Water Quality*?).

- **Compound:** combine two Text and/or Coding Queries to explore your data further (e.g., Did people mention terms related to “health” in the Node *Water Quality*?).

- **Matrix:** create cross-tabulations of your Nodes/sources to look for patterns etc. and quantify your results (e.g., how many people said positive things about *Water Quality* and how many said negative things? What themes were more prominent in response to each survey question?).

**Note:** before running any Query check to make sure that Aggregation for the Parent Nodes in your analysis is turned ON if you want to include results for their Child Nodes, and OFF if you do not want to include them. See notes from previous session regarding Aggregation.

Coding Queries

Basic Coding Queries: Wizard

You can run a Coding Query to search for content based on how it’s coded at specific Nodes. You can find content coded at:

- **A particular Node** (e.g., *Water*) [same as opening a Node to view its contents]

- **All of these Nodes** you specify (e.g., *Water Quality* AND *Fishing* OR *Aquaculture*)

- **Any of these Nodes** you specify (e.g., *Habitat* OR *Forest* OR *Landscape*)

- **Any Node where** there are particular Attribute values (e.g., all the responses of the Person classification attribute of Gender equals value Female).

**Example:** How many and which references are coded at both *Water Quality* AND the specific Economic concern of *Fishing* or *Aquaculture*?

On the Query tab, click Query Wizard, then click the Search for content based on how it is coded, and then click Next. Choose All of these Nodes and tick *Water Quality* under the Natural Environment Node, and *Fishing* or *Aquaculture* under the Economy Node.

**Note:** if you tick the Automatically select descendent Nodes (meaning child Nodes) it will look for co-occurrences of all the child Nodes under *Fishing* or *Aquaculture* as well, and as most of these are mutually exclusive, will return no results. This parent Node has Aggregate coding from child Nodes ticked, so it will include information contained within these child Nodes anyway (see in the Node’s Node Properties).

Then click Next, and choose to look for coded content in Items in Selected Folders to restrict the Query, click the Select button and choose Interviews and Surveys folders. Click Next, then choose whether you want to Run this Query Once or Add this Query to Project (and run it). If you choose to add it to your project, you must enter a name and (optional) a description it will be saved in the Queries area. Then click Run.

Advanced Coding Queries

If you want to use more advanced Coding Queries features that are not available via the Wizard (e.g., find coded content NEAR content coded at other Nodes), you can use the Coding function on the Create tab.
Example: What do respondents 40 years of age and over say about Community Change?

On the Query tab, click Coding and then a new Detail View will open with Query options. Search in the Selected folders Interviews and Surveys. From the drop-down menu select All of the following are true (for the AND Boolean function), then Coded At and Any case where from the next drop-down menus (to specify Attributes). Click on the box to the right labelled “…” to select the attribute conditions for the Query. Expand the list for Person and select Age Group then click OK. Now you have a new drop-down menu to select the logical argument: choose (> Greater than and from the next box select the value 30-39.

Note: NVivo doesn’t recognise age groups in this format as being numerical data so it doesn’t know what order they should be in, it just assumes you have them in the right order.

Add the next condition to the Query by clicking the + button. Choose Any Selected Node then click the Select button to choose the Node Community Change. Spread Coding to the Broad context and click Run Query.

You can also add conditions such as NEAR and Coded by any by clicking on the drop-down menu next to the +/- buttons before adding the next coding criteria.

The results show you everything that people 40 years and over said about Community Change.

Editing a Query

If you make a mistake in a Query, you can always click the Last Query icon on the Query tab to go back to the Query criteria. You can also alter it if the Detail View of the Query is still displayed.

You can edit the search criteria in the Search for content matching these criteria box:

- Change the order of processing, by selecting an item from the criteria list and click the required up or down arrow next to the +/- buttons.
- Remove an item in the criteria list, by selecting the item and then click the “-“ button.
- Edit an item in the list, by editing the criteria in the Detail View.

Exercise: Find out if the interviewees and survey respondents see a connection between Real Estate Development and the Water Quality. Use a NEAR search to allow results that almost overlap in the Nodes.

Compound Queries

Compound Queries can be used to refine a text search by:

- Combining 2 Text Search Queries to find where one term precedes another (e.g., with related to bad before Quality as opposed to good before Quality).
- Combining 2 Coding Queries to search for content coded at one Node, near content coded at another just in case relevant content doesn’t overlap (e.g., where do climate change and sea levels Nodes co-occur within 20 words).
- Combining a Text Search and Coding Query to search for text surrounded by or near a Node (e.g., do people mention words similar to sceptical in/near content coded at the Node climate change).

Creating a Compound Query

Example: Search for mentions of “weather” near the Node Natural Environment.

On the Query tab, click Compound. This opens a new menu window only, not a Detail View like the other searches. On the Compound Query tab select the first Subquery 1 type: Text Search, then click the Criteria button and enter the Search for criteria (e.g., weather). Choose a match level (e.g., the highest: Similar) and select to Search In: Text.
Annotations or Both (e.g., Text). Then click OK. In the next box, choose the appropriate search terms (e.g., Near Content):

- **AND / OR / NOT** (see Advanced Text Search Queries)
- **NEAR content** - coded at Node A near other content coded at Node B within # words (before or after).
- **PRECEDING content** - coded at Node A when it precedes content coded at Node B.
- **SURROUNDING content** - coded at Node A where it is surrounded by content coded at Node B.

If you choose NEAR or PRECEDING Content, click on the Options button to specify the proximity level required:

- **Overlapping** - when coding at both Nodes/context overlaps.
- **In Custom Context** - within a specified number of words, surrounding paragraph/cell/row, minutes or percentage.
- **In Same Scope Item** - within the same item, such as the same document or dataset.
- **In Same Coding Reference** - within the same coding reference (only available when Nodes are included in the Query scope).

For this example, choose In custom context then click Specify button, and select 20 for the **Number of words**. You can also choose how much content you want to gather in the resulting Node/preview (e.g., tick all 3 boxes) in the Options. You can tick:

- **Finds for first search item** - gathers content matching the criteria defined for the first search item only.
- **Finds for second search item** - as above for the second search item only.
- **Content between finds** - displays all content between the first and second search items. Note: This option is only available if both of the above check boxes are selected and Overlapping is not the chosen Proximity option.
- **Compare proximity between text and non-text items** - gathers content from different components of an audio/video or picture source (e.g., Query both the video and the video transcript. If you do not select this check box, the Query will evaluate proximity within the same type of content: either text or media).

In the SubQuery 2 area, select the second Query type from the drop-down menu (e.g., Coding) and click the second Criteria button and enter the Search for content coded at criteria: All of these Nodes (e.g., Natural Environment). You have the Option to set the Scope of your Query from the Search In list (e.g., Items in selected folders), select the items you want to include by clicking the Select button (e.g., Interviews and Surveys). Click OK. Click the Query Options tab to define the spread coding options and preferences for storing the results (e.g. broad). Click Run.

**Note:** You cannot use the Query Wizard to run a Compound Query.

**Matrix Coding Queries**

You can use Matrix Coding Queries to cross tabulate the frequencies of Nodes and/or Attributes in order to:

- compare what different demographic groups have said about an issue (e.g., are different Townships concerned about different issues?).
- explore the amount of overlap between different themes (e.g., is there a potential relationship between Community Change and the Economy, Infrastructure, Real Estate Development, and Policy?).
- compare attitudes (e.g., do Males have more Positive and less Negative attitudes than Females within the Environmental Change Node?).
They are also particularly useful for further descriptive and inferential statistical analysis using chi square for example (when cells are independent). Most often Nodes and/or Attributes are used in the rows and columns of the matrix, but you can also select Sources. In addition, you can create matrices that only relate to content coded within a specific Node (e.g., a specific survey question).

Creating a Basic Matrix Coding Query

Example: Do Commercial Fishers have more Positive, Negative, Mixed or Neutral attitudes than Recreational within the context of Water Quality?

On the Query tab, click Matrix Coding. This will open another menu window instead of a Detail View. On the Rows tab, add rows by selecting either Selected Items or Selected Attributes from the Define More Rows area, then click the Select button and add items (e.g., from the Attitudes Node: tick Mixed, Negative, Neutral and Positive). Then click OK. Then importantly click the Add to List button to move the items into the Name box.

Add columns in the same way by clicking on the Columns tab. To add Attribute values you can choose Attribute Condition from the Define drop-down menu, however it is much simpler to add them as “items”: select the attribute as items from the Case Classifications folder, and choose the appropriate Attribute values from the checklists (e.g., in the Person Classification choose Recreational Fishing and Commercial Fishing value Yes boxes. Note: the reason why there are 2 Fishing attributes with yes/no options is because the Attributes are not mutually exclusive).

Choose where you want to search (e.g., search in Selected Items and choose the Water Quality Node in the Nodes folder under Natural Environment Nodes) then click OK. Choose if you want to run the Query once or add it to the project (and run it). Click Run.

Within the results of a Matrix Query you will see a cross-tabulation known as a Node Matrix. You can double-click on a cell to open the summary of references relating to that cell to look at the context (like running a Coding Query for that cell). In this sense, Matrix Coding Queries can be useful alternatives to Coding Queries.

Matrices can be viewed as charts, by clicking on the Charts tab in the Detail View. You can change the chart type, title, gridlines and rotation (see Creating and Editing Charts below).

Saving Matrices

You can save the matrix retrospectively in the Node Matrices folder in the Nodes area for future reference, or in the Results folder in the Queries area. Select the whole table, right-click on it and select Store Query Results and then select Create Results as New Node Matrix. Select the location Node Matrices, give the matrix a Name and (optional) Description, then click OK.

Exercise: Which Townships are more concerned about Habitat, Landscape and Water Quality Nodes. Include all the townships in rows, and the 3 nodes and columns.

Editing Node Matrices

You can click and drag columns wider or narrower to see more content. Columns can be sorted in ascending or descending order, by selecting the column of interest, on the Layout tab click on Sort By and then click Column. If you need to transpose columns and row, click on the Node matrix in the Detail View then on the Layout tab, click Transpose. Each column and row can be filtered by clicking on the appropriate filter (funnel) symbol. NVivo will show or hide columns/rows based on the filter (e.g., hide rows where number of responses = 0).

You can also apply shading to cells to make it easier to see patterns. When you are in the Matrix Query results, click on the View tab, then the Detail View tab and click on the drop-down menu for Node Matrices and select a shading option (e.g., Blue-White). The darker the shading colour the more coding in that cell. You can also change the information displayed in the cells here (e.g., Number of Nodes Coded as Person). You can also display row or column percentages in the cells of the matrix in the drop-down menu for Node Matrices.
**Note:** By default the Node Matrix displays the number of References in each cell [i.e., the number of text sections coded], and multiple references may be from the same person OR multiple collaborators coded the same content. To see the number of people in each cell making relevant references, click on the View tab, click Node Matrices then under Cell Content choose Cases Coded by → Person.

### Exporting Node Matrices

You can export a Node Matrix to a text or spreadsheet file in order to analyse it statistically or create graphs in Excel (which is probably preferable because you will have more control over the look of charts). In the List View select the Matrix you want to export and on the External Data tab click Items. Select a location to save the file, change the name if necessary, and in the Save as Type box choose the appropriate file format. Click Save.

### Cluster Analysis

Cluster Analysis is helpful in identifying patterns in your data by grouping Sources or Nodes into clusters (maximum of 10 by default). This can help you to visualise the similarities and differences between selected Sources or Nodes (or even between people if your Sources represent individuals). Similarity metrics can be based on the words used, coding conducted or attributes applied. It uses the farthest neighbour (a.k.a. complete linkage) hierarchical clustering method.

You can cluster the following project items:

- **Sources** (e.g., to identify similarities between articles or responses of individual interviewees).
- **Nodes or Case Nodes** (e.g., to identify similarities between the codes you have used or between people/organisations).

You can cluster selected Sources or Nodes based on:

- **Word similarity:** based on the number of words they have in common. **NOTE:** stop words are excluded when using this measure.
- **Coding similarity:** Sources are coded based on the number of Nodes coded they have in common, whereas Nodes are coded based on the number Sources they have in common. **NOTE:** this measure includes ALL of your internal and external sources which contain coding, and you cannot select a subset.
- **Attribute value similarity:** Sources are coded based on the number of Source Classification attribute values they have in common. Only Case Nodes can be coded based on their Case Classification attribute values. **NOTE:** it is best to select a Source or Case Classification to cluster for this measure to ensure they all have the same classification attribute values with which to compare them.
Table of kinds of questions different clustering methods can help you answer:

<table>
<thead>
<tr>
<th>Based On What?</th>
<th>Sources</th>
<th>Nodes/Case Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word Similarity</strong></td>
<td>Which articles are similar in content? Which journals tend to publish similar content?</td>
<td>How similar are selected Nodes in terms of the content coded? Are they related? Are they so similar they should be merged? Which people (Cases) hold similar views according to their survey responses?</td>
</tr>
<tr>
<td><strong>Coding Similarity</strong></td>
<td>Which articles have similar themes? Are there groups of people who are concerned about similar topics?</td>
<td>What group of themes tend to be mentioned together in a lot of articles/interviews? Which authors (if Cases) tend to publish together?</td>
</tr>
<tr>
<td><strong>Attribute Value Similarity</strong></td>
<td>Which articles have similar reference classifications (author, journal, etc.)? Are there distinct groups to analyse separately?</td>
<td>Which Cases are similar in terms of their demographics? Are there distinctive groups of people you might wish to analyse separately?</td>
</tr>
</tbody>
</table>

You can choose from 3 similarity measurements:

- **Pearson’s correlation coefficient**: Ranges from -1 = least to +1 = most similar.
- **Jaccard’s coefficient**: Ranges from 0 = least similar to 1 = most similar.
- **Sørensen’s coefficient**: as above.

**Example: Conduct a Cluster Analysis to see which Nodes are similar in terms of the words coded.**

To run a Cluster Analysis, on the Explore tab, click the Cluster Analysis icon to open the Wizard. Select to cluster by Nodes (to cluster by your thematic Nodes), then click Next. Then in the Nodes box, click the Select button and choose the Nodes you wish to cluster (e.g., all the Nodes in the higher-order folder except for Memorable Quotes and the parent Attitude Node). Then select how you want to cluster them from the drop-down menu (e.g., by Word Similarity), and then choose the Similarity Metric (e.g., Jaccard’s). Click OK.
The Diagram tab shows you a dendrogram (by default) which indicates which Nodes are most similar in terms of the words in their coding. Those that cluster together are more similar than those far apart. In the dendrogram above, it is interesting to note that words coded in the Local Connection and Local Identity are also similar to words used in a Positive context.

On the Summary tab you can see a metric of how similar each pair of Nodes are, from highest to lowest. Items with an index value close to 1 are very similar and as such will appear closer together on the diagram. If you want to see what words Nodes in a cluster have in common, click on the Node names to highlight them, right-click on the selection and choose Run Word Frequency Query.

If you want to see the attribute values of a Source or Case Node in a cluster analysis (e.g., to see if there are demographic similarities between items that cluster together) select the Source or Case Node, right-click on it and select Item Properties.

Editing Cluster Analysis Results

You can change the type of cluster graph when you are in the results area, in the Cluster Analysis tab that appears in the top menu list. Click on the appropriate icon to change the type.

Maps

There are 3 types of Maps in NVivo11:

- **Mind Maps** – For organising your Coding Scheme (and creating Nodes). See *previous course notes*.
- **Concept Maps** – A more free-form way of visually representing the relationships between Project Items.
- **Project Maps** – To get NVivo to visually represent specific Project Items and how they relate to each other within the Project itself.

Dynamic Models can be changed, and Static Models are read-only (all links are removed, but it’s useful for keeping snapshot records of your progress). All newly created models are dynamic.

**Concept Maps**

Concept Maps are free-form rather than strictly hierarchical and allow you to create diagrams with a variety of shapes and connectors, and elements that are connected to Project items (i.e., you can open the project items directly from the map).

On the Explore tab click Concept Map. Enter a name for the model in the Name box, and an (optional) Description, then click OK. In the Detail View click and drag the shapes you want to add from the window pane on the left. Consider using different shapes to denote different elements of your project (e.g., circular for positive and square for negative). Double-click on the shapes to enter labels.

To connect shapes to each other, click the Connector button on the Concept Map tab which changes the cursor to a connector. Click and drag from one shape to another to connect them. If you want to change the directionality of the connector, click on the connector and then the Change Connector button on the tab and select the direction you want. To get your normal cursor back, just click Pointer.

If you wish to add a shape which is a link to a Project Item, click the Project Item button to bring up a menu and tick all the items you want to add, and click OK. You can also click-and-drag project items from the List View to the map.
You can delete a shape by selecting the one you want to delete then on the keyboard click Delete.

**Add a note to a model**

You can add notes to models: on the Model tab, click the Note shape from the Shapes drop-down menu. Double-click the note and in the Text box, enter your notes. Click OK.

**Project Maps**

Project Maps can be used to get NVivo to show you a graphic of the final coding structure of your Project, or how Project Items are connected to each other (e.g., you can look at all the Nodes associated with one Case or Source). You can also click on one of the items and asked to see a list of all associated project items (e.g., Child Nodes, Cases, Sources, Links), to decide if you want to include them in the map.

On the Explore tab click Project Map. Enter a name for the model in the Name box, and an (optional) Description, then click OK. On the Project Map tab, click on the Add Project Items icon to select the items you wish to include in the map (e.g., the Economy Node and all its child Nodes). Then click OK. On the Project Map tab you can deselect any connectors you wish to hide from the map in the Connectors Area of the Project Map tab (e.g., Links).

A quick way to see the Project Items that are associated with an Item already in the Map (and to show this association on the map also), click on the Item and on the Project Map tab, and then click Show Associated Items icon at the top of the window. A new window appears on the left-hand side of the Detail View which lists all valid items you can include. You can click and drag the items you want to add from that pane.
**Explore Diagrams**

You can also create special graphs to show how project items are connected to one specific Source, Node or Case Node. This is very similar to a project map, only the latter can show you how the other Project Items relate to each other as well.

**Note:** I recommend exporting any quantitative data you may have in NVivo to Excel in order to create charts. This is because the charting functions within NVivo do not allow you full control over the design of your graphical representations (e.g., the colour of bars, titles etc.).

---

**Example: Create an Explore Diagram of the Project Items related to Balance.**

Select the Source, Node or Case Node you want to create a graph for in the Node or Source area (e.g., select *Balance*). On the Explore tab, click on the Explore Diagram icon to generate the graph. You will see a new Graph appear in the Detail View, which shows you all the project items in a related to it. As this is a dynamic graph, you can click on any one of the icons to open up that Project Item. You have extra options such as showing or hiding information on the Explore Diagram Tools tab (e.g., Sources coded or Memo Links). You can also change the focus to another Node within the graph itself by right-clicking on a Project Item object, select Change Focus and a graph will be generated with that Node at its centre. If you want to go back to the original, click the Back button on the Explore Diagram tab.

![Graph Diagram](image)

**Comparison Diagrams**

You can use comparison diagrams to compare two Nodes, Sources or Cases and see which Project Items they do and do not have in common. The following project items can be displayed on a map:

- Nodes
- Sources
- Cases
- Classifications
- Attributes
- Relationships
- Links
**Example: Create a Comparison Diagram of the Nodes common to both of the 2 news articles, and how they differ.**

On the **Explore tab**, click on the **Comparison Diagram** button and choose to **Compare Sources**. Then select the Sources for both **News Articles** and click **OK**. The graph will open up in the **Detail View**. You can simplify the graph by unticking elements such as **Memo Links**, **Relationships** and **Sources Coded**. The graph below shows you which Nodes code both articles, and which do not.

**Exporting Results and Project Items**

The easiest way to export any sources, Nodes, models, report output, as well as Query results and visualizations is when they are displayed in **Detail View**. Click on the **Detail View** you want to export, and on the **External Data** tab click **Items**. If you’re exporting **Text Search**, **Coding** or **Compound Queries**, choose what you want to export (**Entire Content**, **Reference View** or **Summary View**). If you are exporting a **Matrix Coding Query**, a **Query Visualisation** or list of items in the results of a Query, choose the file format for the exported file in the **Save As Type** list. In the **File Name** box, change the name if you like. Click **Browse** to select where you want to save it and click **OK**.

**Exporting Classification Sheets**

**Example: Exporting the Person classification summary sheet.**

If you want to export a spreadsheet containing your demographic or bibliographic information for example, you can export source or Node classification sheets as an Excel Workbook (*xlsx*). Click on the classification sheet you want to export (e.g., the **Person** classification in the **Node Classifications** folder). On the **External Data** tab click **Classification Sheets** in the **Export** group (not the **Import** group). Select the appropriate export options: (e.g., export to **Excel Workbook**). Choose a location to save the file and then click **OK**.
Exporting Nodes (Including the Entire Content)

You can also export the Summary View or the Reference View of individual Nodes by right clicking on the Node, select Export Node and then choosing the appropriate option. In addition, you can export the entire content of a Node as a collection of HTML pages to share your project findings (including audio and video clips) with colleagues who don't have NVivo. The Node is exported as a series of web pages that are stored in your My Documents folder. The Home page has the same name as the exported Node (for example, adapting to climate change.htm). You can open this file to access the 'mini website'.

In the List View, select the Node you want to export. On the External Data tab, click Items and from the Export Options list select Entire Content. Click the Browse button the change the name/location of the exported files under Save As. Select the properties and related content that you want to include and tick the select the Open on Export check box to see the html pages. Click OK.

Note: If you want to move the web pages, make sure you copy the folder containing the .HTM files to the new location. For Nodes that code audio and video sources, the coded sections of media files will be included upon export only if you coded directly on the audio or video timeline.

Exporting the Codebook

The information you provide in the Description Area when creating each Node can also be exported as a table in Word. Right-click in the List View of the Node area, select Export Codebook. You will get a table in a Word document such as the one below.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed</td>
<td>A person's statement that they are indecisive as to whether their attitude is positive or negative. If a person states two different views, such as that they are positive about one aspect and negative about another, this is not mixed, but rather multiple, feelings.</td>
</tr>
<tr>
<td>Negative</td>
<td>A negative opinion or a disagreement where an opinion is not shared with someone or with a group.</td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>specifying positive opinion or agreement with other stakeholders or with an idea. A shared opinion</td>
</tr>
</tbody>
</table>

Exporting the List View of the Nodes Area

You can export all the information shown in the List View of the Nodes Area into an Excel file. Right-click in the List View of the Node area, select Export List.

Running Reports and Extracts

Reports contain easy-to-read information about your project that you can print (e.g., you could run a Coding Summary by Node report to review the content coded at each Node organized by source). When you run a report, the results are displayed in Detail View and can
be printed from there. **Extracts** allow you to export a collection of data to a text, Excel or XML file for analysis in another application. When you run an extract, results are saved as a file.

There are predefined reports and extracts that you can run at any time (see below). You can also build your own reports.

<table>
<thead>
<tr>
<th>Type</th>
<th>Description of Reports and Extracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding Summary by Node</td>
<td>Lists <strong>Nodes</strong> and the sources that are coded at them. No textual content is included in the report. Useful for reviewing coding.</td>
</tr>
<tr>
<td>Coding Summary by Source</td>
<td>Lists <strong>sources</strong> and the Nodes that code them. No textual content coded is included in the report.</td>
</tr>
<tr>
<td>Node Classification Summary</td>
<td>Lists <strong>classifications and attributes</strong> for the <strong>Nodes</strong> in your project. Useful for complementary analysis in statistical applications.</td>
</tr>
<tr>
<td>Node Structure</td>
<td>Lists <strong>Nodes</strong> in your project with information about the folder location, aggregate status, nickname and colour.</td>
</tr>
<tr>
<td>Node Summary</td>
<td>Lists <strong>Nodes</strong> in the project including statistical information about each (e.g., total words coded, number of sources coded).</td>
</tr>
<tr>
<td>Project Summary</td>
<td>Displays <strong>properties</strong> of your project (e.g., project name and description), user profiles, and lists all the items within your project.</td>
</tr>
<tr>
<td>Source Classification</td>
<td>Lists <strong>classifications and attributes</strong> for the <strong>sources</strong> in your project. Useful for exporting bibliographical data to reference management tools like EndNote or Zotero, or other applications that support text, Excel or XML file formats.</td>
</tr>
<tr>
<td>Source Summary</td>
<td>Lists <strong>sources</strong> in the project. For each, it shows source properties and statistics (e.g., number of words, paragraphs), including the number of Nodes that code the source.</td>
</tr>
</tbody>
</table>

**Note:** Node matrices and framework matrices are not included in the Node Summary or Source Summary reports.

To run a predefined report, go to the **Reports** area in the **Navigation View** and select the **Reports** folder. In **List View** select the report you want to run (e.g., **Node Structure**). On the **Explore** tab click **Run Report**. If your report contains filters, the **Filter Options** dialogue box is displayed and you can select any filters you wish to apply to the data. Click **OK**. To run a predefined extract (i.e., save the report as a file), go to the **Extracts** folder instead and run the extract.

You can also design your own reports using the **Report Wizard**. You can also use the **Report Designer** to modify predefined reports. See the NVivo 11 Help website under **Using the Report Designer** for more information on editing reports.

You can **cancel** a report or extract, if it is taking too long to run by clicking the red cancel button beside the progress bar that is displayed at the bottom of the window when the report or extract is running.

**Collaboration**

It’s always a good idea to have more than one person coding the data to improve the reliability and validity of the research. You can keep track of who coded what and how, by giving each team member accesses the project using their own ‘User Profile’. [You MUST use the same version of NVivo]. The easiest option is for everyone to work on the same project at different times on a **Standalone Project** (a single file) logging-in with different user profiles.
When you open the project, check the **Current User Profile** to make sure you are using your own. This is displayed on the **Status Bar** at the bottom of the NVivo window (with the person icon). If not, you can switch to the correct user profiles without closing the project: click the **General** tab and then under **User** enter the name and initials of your User Profile that you want to switch to.

**Note:** It MUST match your User Profile initials or it will create a new profile for you. Then click **OK**.

If more than one user wants to work on a project at the same time, you can establish a 'master' project, and then give a copy of the master project to each team member. The team members work in the copied projects while you work in the master project. At regular intervals, you can import your team members' work into the master project—refer to **Import selected content into an open project** on the **NVivo 11 Help Site** for more information.

## Interrater Reliability

An important measure of the reliability of your coding scheme is to see how well multiple Coders (or Raters) agree with each other on the coding performed. You can conduct a **Coding Comparison Query** to get two measurements of inter-rater reliability: the **percentage agreement** (the number of units of agreement divided by the total units of measure within the data item) and 'Kappa coefficient' (a statistic which compares your level of agreement to that which could be expected to occur by chance alone). For more information on conducting an interrater reliability analysis see **Run a Coding Comparison Query** on the NVivo 11 help site.

**Example: conduct a Coding Comparison Query between Effie and Henry.**

On the **Query** tab click the **Coding Comparison** and select the users you wish to compare from the **Compare coding between** box (e.g., Effie Reeves and Henry Patterson). In the **At** box click **Select** and choose the Nodes you wish to include in the analysis (e.g., choose **Nodes in selected sets**, then click the **Select** button and choose the **Nodes for coding comparison** folder). Choose the specific sources to use in the **Scope** box (e.g. choose **Sources in selected sets**, then click the **Select** button and choose the **Interviews for coding comparison** folder). Select the measure(s) you want to use: **Kappa and/or Percentage Agreement**. Click **Run**.

You can double-click on a row in the results where there is some disagreement (e.g., less than 80%) to compare the coding conducted and see where they disagreed for that source and Node. There will be a coloured coding stripe for Effie’s coding and another colour for Henry’s with their initials attached. The pink coding stripe represents the aggregated coding for that section (Effie and/or Henry’s coding).

## Mac vs PC Versions of NVivo

If you are intending to use the Apple Mac version of NVivo, please note that it does not have all the capabilities of the PC version and the user interface can be quite different. In newer versions of NVivo, the interface is becoming more similar, however the older versions are vastly different. These notes apply to the PC version only, so if you are working on a Mac and have trouble finding what you need please refer to the online NVivo 11 Help for Mac: [gl.anu.edu.au/n1bj](http://gl.anu.edu.au/n1bj)

If you are working on both PC and Mac versions, you will need to convert your Project files to the appropriate format to the platform you are working on. Information on how to do this can be found at the link above also.
Other resources

Training notes
To access the training notes, visit the Research & learn webpage anulib.anu.edu.au/research-learn and select the skill area followed by the relevant course. You can register for a workshop and find other information.

Research & learn how-to guides
Explore and learn with the ANU Library's how to guides (ql.anu.edu.au/howto). Topics covered are:

- Citations & abstracts
- Data Management
- E-books
- EndNote
- Finding books and more
- Finding journal articles and more
- Finding theses
- Increasing your research impact
- NVivo
- ORCiD (Open Researcher and Contributor ID)
- Topic analysis
- Using Google scholar from off-campus

Subject guides
Find subject-specific guides (ql.anu.edu.au/subjectguides) and resources on broad range of disciplines. Such as:

- History, indigenous studies, linguistics and philosophy
- Criminal, human rights and taxation law
- Biochemistry and molecular biology, neurosciences and psychology
- Asia Pacific, Southeast Asia and East Asian studies
- Engineering
- Astronomy and astrophysics, earth sciences, mathematical sciences and natural hazards

Online learning
Online learning is available through ANU Pulse (ql.anu.edu.au/pulse), which can be accessed from both on and off campus by all ANU staff and students.

IT skills development modules available in ANU Pulse
- Microsoft Office (Access, Excel, OneNote, Outlook, PowerPoint, Project, Visio, Word)
- Microsoft Office (Mac)
- Adobe suite (Illustrator, Photoshop)
- Type IT

Training calendar
Select Events » near the bottom of the Library homepage to access our events calendar with upcoming training opportunities displayed day by day (ql.anu.edu.au/cal).

Feedback
Please provide feedback about today’s workshop via an online feedback form (ql.anu.edu.au/survey)