NVivo beyond the basics: queries with nodes

Workbook

Edition 2
January 2017
Document Reference: 3829-2017
NVivo beyond the basics: queries with nodes

Contents

1. Introduction ......................................................................................................................... 1
2. NVivo interface orientation ............................................................................................... 2
3. Boolean logic ..................................................................................................................... 3
4. Coding queries ................................................................................................................... 3
5. Compound queries ............................................................................................................. 6
6. Matrix queries ................................................................................................................... 8
7. Why use case nodes? And getting started ......................................................................... 8
8. Creating and editing a classification ................................................................................ 9
9. Creating case nodes .......................................................................................................... 10
10. Importing classification sheets ....................................................................................... 12
11. Running matrix queries using cases and attributes ....................................................... 14
1. Introduction

This course is aimed at researchers who already know how to do basic importing and coding of sources, and who would like to be able to run different types of queries. The queries covered in this workbook are queries that allow you to test and build on your coding, and to interrogate your data according to attributes to allow you to make comparisons. Bazeley calls these types of queries ‘theory building’ queries. In this course we will use the sample project data that is shipped with NVivo, but if you would like to use your own data sources, that is welcome.

The screenshots are from NVivo 11 Pro, but are largely easily recognisable in other versions. No particular methodological approach to analysis is adopted in the activities; rather the aim is to introduce key mechanisms of the software which may be applied as required by different analytical approaches.

Why use NVivo?

NVivo is a large and complex piece of software which is most helpful when working with large amounts of data, particularly where the data include different formats. It is useful for managing and organising projects with many separate data sources to support more transparent and systematic approaches to coding and data interpretation. For projects with small datasets, NVivo may be unnecessarily complex.

Course aims

By the end of this course, you will be able to

- Run coding queries
- Run compound queries
- Create case nodes
- Create and import classifications, attributes and values to use with case nodes
- Run matrix queries to extract coded data by nodes and attributes

Further resources

- Follow QSR on twitter for information about free training seminars, updates and tips: @QSRInt
- Mantra practical on research data management: [http://datalib.edina.ac.uk/mantra/softwarepracticals.html](http://datalib.edina.ac.uk/mantra/softwarepracticals.html)
- LinkedIn NVivo Users Group – good source of advice from experienced users. Free, but requires a LinkedIn account and application to join the group [https://www.linkedin.com/groups/NVivo-Users-Group-145388/about](https://www.linkedin.com/groups/NVivo-Users-Group-145388/about)
2. NVivo interface orientation

The NVivo interface is divided into different sections, the contents of which vary depending on the elements or tools being used. The image below has numbers next to the main elements we will be using on this course, and the name of each element is listed under the image.

1. **Tabs** – the tabs along the top of the NVivo window give access to different operations or tools.
2. **Ribbon** – like recent versions of Microsoft Office, the operations or tools available within each tab are presented in a ribbon along the top.
3. **Upper navigation view** – the upper section of the navigation view shows the contents of the element selected in the lower section of the navigation view (in this case, the folders of Sources created for this project).
4. **Lower navigation view** – the lower section of the navigation view lists the core organisational and analytical aspects of the database.
5. **List view** – this pane lists the contents of whatever is selected in the upper navigation pane (3).
6. **Detail view** – this pane shows in detail the item selected in the list pane (5).

The sizes of the different sections can be increased or decreased by clicking and dragging on the dividing lines. The layout can also be changed by using the options in the View menu > Workspace ribbon.
3. Boolean logic

When formulating queries, you will need to think carefully about how the different elements of your queries should be combined in order to get useful or meaningful results. There are three basic combining options:

- **AND** (sometimes **ALL** in NVivo) requires all the different essential elements (e.g. nodes or attributes) to be present in the results. This will narrow your search.

- **OR** (sometimes **ANY** in NVivo) is an ‘either or’ combination that can be used for synonymous or interchangeable elements. This will broaden your search.

- **NOT** may be used to exclude an element from your search results. This will narrow your search.

4. Coding queries

**Why use coding queries?**

Queries using nodes let you interrogate your data in a systematic way that is time-consuming to do manually. What if you wanted to identify all the coded sections that are linked to one, two or more nodes or concepts so that you can analyse more closely the intersection of those ideas? It is possible to run queries on single nodes, combining nodes, combining nodes and text search terms, and more. We will start with a basic query and move to more advanced queries.

**What do you need to do first in order to use coding queries?**

At the most basic level, in order to run coding queries it’s necessary to have coded the relevant sources in your project using the appropriate node(s).

If you are using your own NVivo project, make sure you have done some coding with at least two nodes, ideally with a bit of overlap in selections, before moving to the next step.

If you are using the NVivo Sample Project, move to the next step.
Running a coding query

1. Open the Sample Project, or your own project in which you have done coding with at least two nodes.

2. In NVivo, click on the **Query** tab, and click on **Coding**.

3. In the detail view, a **Coding Query** tab appears. You can see a **Coding Criteria** section at the top, and a space for results at the bottom. You may need to click and drag the edges of the coding criteria area to increase the space in order to see the options clearly.

4. The coding query provides several options:
   
   a. Search for a single node in all/selected sources and as applied by all/selected people who have worked on the project. So this is a bit like opening a node and looking at the details of where it has been used, except that the query allows some refinement of results.

   b. You can also search for combinations of nodes in various ways.

   c. Search for data coded using a Case Node with attributes – for example age of respondents or organisation type. This search will work in the Sample project, but if you are using your own data and have not used attributes, the second
option will not work properly. Later in this coursebook, we will look at attributes in detail while constructing matrix queries.

5. To try a node search, use the default to search in All Sources, leave the option All of the following are true, Coded at, All Selected Nodes, and click the ellipse […] button.

6. Browse to find a node to search for – e.g. Natural environment.

7. Click the Run button to start the search. The results will be displayed below the query builder.

**Tip:** If you think you’ll want to run a query again, be sure to select the Add to project option in the query window. The query will be saved in the Queries area, and can be organised into folders and subfolders. When a query is run later, the new results will reflect any changes and updates in the project.

Next, let’s try building a more advanced coding query, in which we do a search on a combination of nodes.

1. Click on one of the + symbols at the right end of the query builder – this adds a new row to the builder.

2. The new row also shows Coded at and All Selected Nodes. Click the ellipse […] button.

3. Browse to find a node to search for – e.g. Real estate development. Click the Run button to start the search.

The results will be displayed below the query builder. Note the tabs at the right hand side of the detail view, which let you explore the results in detail:
Note that because we left the query criteria set to *All of the following are true*, our search results are restricted to those coded extracts that are coded at both Natural environment (or one of its child nodes) AND Real estate development. You can switch on the coding stripes in your results to check.

If you got no results, it is likely that you combined two nodes that do not overlap in your coding – if that happens to you, you can test this query builder by changing the setting to *Any of the following are true*, which will bring up results showing where either of your nodes have been used (an OR combination).

If you want to find areas with combinations of nodes that are not actually overlapping, you can click on the drop-down arrow at the right of one of the rows and select *Near*. In this new row, you can customise how far apart the nodes may be – for example, within a certain number of words, or in the same source.

Bear in mind that you can also save the results of queries like this as nodes. Those results nodes are effectively incremental combinations that can then be used in a more complex matrix query.

5. Compound queries

Why use compound queries?

Compound queries combine node searches and/or text searches. These queries could be used to find places where two nodes are used in close proximity where you may wish, for example, to re-code or merge nodes.

These queries can be used to check the thoroughness of your coding – for example, have you missed any relevant excerpts where an interviewee uses a significant word or phrase? This may be particularly useful if your coding concepts (represented by the nodes in NVivo) have evolved over the course of your research, and you would like to go back and check whether there are important sections of your data that potentially have new relevance in light of your more mature codes.

What do you need to do first in order to use compound queries?

It is necessary to have coded with at least one node.

Running a compound query

We will run a query that combines a node with text, to check on thoroughness of coding.

1. In the *Query* tab of the ribbon, click on *Compound*. Alternatively, if you have selected Queries in the navigation area, you can right-click in the list view and select *New query*, then *Compound*.

2. Select *Add to project* and give the query a name.

3. Click on the *Compound Criteria* tab in the query window.
4. In **Subquery 1**, choose **Text Search Query** and click **Criteria** to specify a word or phrase to search – for example, pollution, in the Sample project. You can experiment with the Finding matches options, depending on how flexible you want your search to be.

5. In the dropdown menu in between the two subqueries, select **AND NOT**.

6. In **Subquery 2**, choose **Coding Query** and click **Criteria** to specify a particular node – for example, in the Sample project, Environmental impacts.

7. Click on the **Query Options** tab, and click on the **Option** dropdown menu. For this query, leave the setting to **Preview Only**, but note that there are other options for dealing with the query results.

8. Still in the **Query Options** tab, under **Spread Coding**, select **Broad Context**. This will allow you to see more text round the search terms, and makes it easier to analyse and judge whether additional coding is needed.

9. Click **Run** and you should be able to examine the results in the detail view.

10. It is possible to do additional coding in the **Reference** tab of the results – it may help to show coding stripes for all nodes before doing new coding.

Try experimenting with other options in compound queries, and compare results and outputs.
6. Matrix queries

Matrix queries are fairly complex, and the rest of the workbook deals with how to build them from scratch, including creating case nodes to work with in the queries.

Why use matrix queries?

In Coding queries, we saw how to run a query to find data at a particular node.

We did not look at it earlier, but it is also possible to retrieve data at a node but restricted to a particular attribute – for example, data coded at the node ‘attitude’ in respondents aged 60 years or older. The latter option in the coding query implies the need to make comparisons between subgroups in your data, and this is where matrix queries are particularly helpful.

How do men compare with women in their views about pollution? How do teenagers’ experiences compare with adults’ experiences? How do not-for-profit organisations conceptualise urbanisation in comparison with commercial organisations?

What do you need to do first in order to use matrix queries?

In order to run detailed matrix queries, you need to have completed at least some coding, created classifications and have assigned attributes (such as demographic data) to your respondents or cases.

In the next sections, we will create case nodes, attributes and coding.

7. Why use case nodes? And getting started

Case nodes are associated with sources, but are not quite the same as the nodes that you would use for coding your data according to the concepts and themes you are analysing.

A case node allows you to manage various types of information about each of your cases (a case may be a person, an organisation, a group, a publication, etc., depending on your research project).

With a case node, you can associate important information about your cases, such as demographic attributes, as well as link multiple sources with a single case, such as multiple interviews or images. With all the associated sources and attributes as part of a case node, it is then possible to run reports and queries that can help with getting a good overview of your study sample and your coding.

We will start with a new project in this course book, but do apply the steps to your own project and data if you prefer.

1. In NVivo, go to the File menu and select New

2. In the New Project dialog, give your project a title and use the Browse button to save it into the Documents folder on your computer.

3. Make sure you have Sources selected in the navigation pane (list at lower left).

4. In the Sources list (list at upper left above the navigation pane), right click on the Internals folder and select New Folder. Name the new folder Interviews.

5. Into your Interviews folder, import the document-type interview of Barbara from the sample data folders.
8. Creating and editing a classification

Before you can create case nodes, you need to create classifications matching the type of case you are dealing with. By default, NVivo has basic classifications for people and organisations, but you can create your own.

1. Select **Classifications** in the navigation pane (list at lower left).
2. In the Classifications list (list at upper left above the navigation pane), select **Case Classifications**.
3. Right click in the (empty) list view pane, and select **New Classification**.
4. Create a new Person classification by selecting **Add one or more predefined classifications to the project** and click the tick-box for **Person**.

By clicking on the [+ ] at the left of your new Person classification, you can open the list of attributes available by default. You can edit those attributes, or add new attributes, along with their possible variables, or values.

1. To add a new attribute, right click over the Person classification row and selecting **New Attribute**.
2. Name the new attribute **Marital Status**.
3. Click on the **Values** tab, and use the **Add** button to add two new values: **Single** and **Married**. Click OK.
9. Creating case nodes

Organising your case nodes
Click on **Nodes** in the lower navigation panel, then you will see **Cases** listed above, in Nodes navigation panel. If it helps to organise cases into clusters, you can create folders and sub-folders for the clusters relevant to your project.

Creating a case node manually
Sometimes you need to create case nodes based on your sources after your sources have already been imported.

1. Go back to the **Sources** button and select the **Interviews** folder.
2. Right click on the row of the source **Barbara** that you imported earlier.
3. In the menu, select **Create As**, then **Create As Cases**.
4. In the **Select Location** window, select on a folder under **Cases** if you made one, then for **Assign to Classification**, use the drop-down menu to select **Person**. Click **OK** to confirm.
5. Now, use the lower navigation panel to view **Nodes**, select **Cases** in the upper navigation panel, and you should then see a new case node named **Barbara** listed within **Cases** (or within a sub-folder, if you made one).
6. Right-click on the case node **Barbara** and select **Open Classification Sheet**

7. You can see each attribute from the Person Classification in each column. These attributes do not yet have values added, but if you scroll to the right, you will find **Marital status**, and you can select from the values you added earlier.

   To add values for the rest of the attributes, you would need to edit the Case Classification as described in section 8 above.

   If you have a lot of cases, it can be tedious to add these attributes and values manually to each case. To save time, the attributes and values for multiple cases can be imported in a previously created classification sheet - in section 10, we look at importing a classification sheet created in Excel.
Creating case nodes as part of importing sources

It can save time to create case nodes at the same time as you import your sources. Be sure that your sources are named exactly as you would like to have your case nodes named.

1. Click on your **Interviews** folder in the upper navigation pane so that it is highlighted. (You need to have Sources selected in the lower navigation pane.)
2. Right-click in the main window, and in the menu that pops up, select Import, then **Import Documents**.
3. In the Import Internals dialog box, use Browse to navigate to the desktop, where you should find the Sample Project Data folder, then open the Interviews folder. (On your own computer, you would browse to wherever you have saved your project files.)
4. Leaving the Barbara interview alone, select the rest of the interview documents, then click Open.
5. Back in the Import Internals dialog box, click the More button if necessary to see additional options.

6. Tick the box to select **Code sources at case nodes located under**.
7. Click the Select button to choose the People folder within Cases.
8. For **Assign new nodes to classification**, click the Select button to choose the Person classification.
9. Click OK and you should see the new sources in the Interviews folder, as well as the new case nodes within Nodes > Cases.

10. Right-click over the cases list and select **Open Classification Sheet**. The attribute values can then be manually assigned to each case.
10. Importing classification sheets

It can save time to populate classifications using an excel spreadsheet that can be imported into NVivo – or you may have spreadsheets of attributes as a result of survey responses.

The sample data includes a classification excel spreadsheet for interview participants (look in the folder named Other).

When you use a spreadsheet to manage your classifications, it is essential that the left-most column matches exactly the names of case nodes you would like the attributes assigned to. By importing all of the interview documents earlier, we included two files for which there will not be an exact match – so you can see what happens.

The first row in the classification sheet gives the attribute names in each column, then for each following case row, the case’s value for each attribute is given in the corresponding cells.

1. Make sure you have selected Classifications in the lower left navigation panel, and have Case Classifications selected in the upper left navigation panel.

2. Right click on the Person classification row (or the empty space) in the classifications list view and select Import Classification Sheets.

3. Browse to open the excel classification sheet in the sample data folder, Other, then click Next.

4. In Step 2, make sure Classification type is set to Case Classification. Select all three tick-boxes for

   a. Create new attributes if they do not exist

   b. Update the classification of existing sources or nodes

   c. Replace attribute values of existing sources or nodes that appear in this file.

Click Next.
5. In Step 3, specify that the nodes are represented As names. If you are placing the new cases into a folder within Nodes > Cases, you’ll need to click the Select button to select the relevant folder. Make sure there is a tick in the box for Create new nodes if they do not exist (you may not always want to use this option).

6. In the Wizard’s Step 4, we will not make any changes in this instance, but with your own data, depending your preferences for defaults for date formats and labels for missing data, you may wish to make changes. Click Finish.

The Classification sheet now includes the imported attributes with values, as well as the empty default attributes from our earlier step in manually creating the Person classification. It is possible to delete unwanted attributes by right-clicking on them in the list view and selecting Delete.

If you look in the list of Cases, you should see two cases where the names of sources we had imported earlier did not match the names in the classification sheet (Ken and Helen). You can merge cases by right clicking on one, selecting Cut, then right clicking on the case node you would like to merge with, and selecting Merge into selected case.
11. Running matrix queries using cases and attributes

Next we will look at how to extract coded data according to particular attributes - this is valuable for letting you compare, for example, views by gender or age, based on classifications, attributes and values in combination with coding using nodes.

In order to get a better understanding of how your cases and classifications work, we need to have a bit of coding done - we'll do that next.

1. Click on the Nodes button in the navigation pane (lower left area of NVivo). Right click on the Nodes folder at the top of the navigation pane and create a new folder named Themes.

2. Within the Themes folder, create at least three nodes (Fishing, Tourism, and Ecosystem, or any concepts you would like) and code one or two sections of at least three interviews - Barbara, Charles and Dorothy (or others of your choice and remember which ones you worked on for future steps). For this exercise, just select some of the text for coding without spending too much time choosing the text - we just need to have some coding done for running a query.

Once you have a bit of thematic coding, we are ready to run a query.

1. Click on the Queries button in the navigation pane (lower left area of NVivo). Select the Queries folder at the top of the navigation pane.

2. Right click in the list window and select New Query, then select Matrix Coding.

3. Tick the box to select Add to Project (this will let you save and re-run or edit the query).

4. Give the query a name, such as Themes by Attribute.

5. Click on the tab for Matrix Coding Criteria.

6. In the Rows tab area, we need to define the row – click the Select button and navigate to select the thematic nodes created earlier, click OK and back at the Matrix Coding Query window, click Add to List.
7. Next, click on the **Columns** tab, and under **Define More Columns**, click the **Select** button and navigate to select a few of the Case Classification attribute values you know have been assigned from our earlier steps (such as under **Gender** and **Years in Down East**), click **OK** and back at the Matrix Coding Query window, click **Add to List**.

8. The **Node Matrix** tab options can be left as they are. Click the **Run** button to generate your query report.

The report is displayed as a table showing the number of times each thematic node has been used in a source with particular attributes.

At the right, a Chart display option will show a graphic of the frequencies.

If you would like to see the coded data indicated in each cell of the matric for comparative analysis, double click to open a report extracting the coded data elements.