

Chapter 7. Exploring Patterns Across Dimensions

In this chapter, you will look at patterns that indicate how one dimension of your data is associated with another dimension. You will build contingency tables showing the relationship across the categories of two dimensions and block charts to examine their patterns. A process for the stepwise comparison of dimensional patterns across your built-in contrast and across multiple streams is introduced.

■ Dimensions

In the last chapter, we looked at the distribution of your data—the ups and downs created by the categories of your coding scheme. In this chapter, we add another dimension by looking at how these distributions are related to the distribution patterns of a second coding scheme. This kind of analysis will help you understand how the distribution of your data in one dimension is associated with its distribution in a second dimension. The nature of the association between the dimensions of your data can yield useful analytic insights. For example, seeing that two codes in different dimensions go up and down together can suggest a relationship worth investigating.

Adding a second dimension to an analysis involves coding your already-coded set of data with a second coding scheme. The goal is often to develop a greater understanding of the differences across your built-in contrasts. If, for example, you have discovered that design and management meetings

are different along the dimension of speaker, you might then begin to wonder how these types of meetings varied along a second dimension—indexicality, for instance. Figure 7.1 shows a sample of data that has been coded in this way both for the dimension of speaker (in column B) and the dimension of indexicality (in column D).

	A	B	C	D
1	T-Unit#	Speaker	Text	Indexicality
2				
3		1 Cheryl	I mean ...	Not Indexed
4		2 Ed	Jesus.	Not Indexed
5		3 Cheryl	See I:	Not Indexed
6		4 Cheryl	see where	Not Indexed
7		5 Cheryl	this little thing is?	Indexed

Figure 7.1: Data coded along two dimensions: speaker and indexicality.

■ Questions of Distribution

By looking at your data across more than one dimension, you can ask two new kinds of questions. To begin with, you can ask how the distribution of data over the categories of the second coding dimension varies across your built-in contrast. This question is a question of distribution and it parallels the question you asked concerning your first dimension. For example, with the two dimensions shown in Figure 7.1, we can ask not only a question about distribution of speaker contribution: *Do design meetings differ from management meetings in the relative contributions made by speaker?* We can also ask a question about the distribution of indexicality: *Are design meetings more or less indexical than management meetings?*

While the distribution of codes from a single dimension can tell us something about the difference between design meetings and management meetings the perspective provided is . . . one dimensional. The qualities that make verbal phenomena significant and analytically interesting are often more complex than can be captured in a single dimension of coding. To get at this complexity, we must ask questions of association.

■ Questions of Association

If we stopped our analysis with answers to questions of distribution, our understanding of the data will be less than complete. We would fail to explore the associations between the two dimensions—how they are interrelated. For this we have to ask questions of association.

Generally, questions of association ask how variations along one dimension are associated with variations along the second dimension. In the case of the data in Figure 7.1, for example, we can ask a question about the association between speaker contribution and indexicality: Do some speakers employ more indexical language than other speakers? Furthermore, we can ask how that association plays out across our built-in contrasts: Does the rate of indexicality of a speaker vary by the kind of meeting they are attending? The procedures outlined in the rest of this chapter are designed to help you answer questions of association.

■ Memo 7.1: Second Dimension

Reflect on possibilities for a second dimension of coding that would help illuminate an aspect of your data. Consider dimensions of coding that help you understand distributions that you have already discovered. What questions of distribution and association could you ask with that second dimension and which seem the most likely to further your analysis?

Exercise 7.1 Test Your Understanding

Decide whether each of the following questions is a question of distribution or a question of association. Then label the dimension being used in each question.

- During what decade did Elvis record his most popular songs?
- How is the productivity of rock stars related to age and gender?
- Are men more likely than women to act aggressively in on-line interactions?
- Are certain topics associated with greater aggression among men than among women?

For Discussion: Discuss with your classmates the kind of data sheets that would allow you to answer each question.

■ Contingency Tables

Seeing associations across dimensions involves the use of contingency tables like the one shown in Figure 7.2. Simply put, a contingency table is a tabular array showing the frequency distributions of two different coding dimensions. The categories of one coding scheme are arrayed across the top of the table (*Indexed* and *Not Indexed*). Down the side are arrayed the categories of the second coding scheme (*Cheryl, Ed, John*). In the cells are listed the frequencies with which the two dimensions intersect. In appearance and effect, a contingency table is a matrix that shows the interrelationship between two coding dimensions. Instead of tracking the distribution of codes within a single dimension (e.g., indexed, not indexed) you are simultaneously tracking the distribution of codes from the first dimension across the second dimension. As seen in Figure 7.2, a portion of the segments that are coded as *Indexed* will be spoken by *Cheryl*, another by *Ed*, and another by *John*. Together, the sum of the indexed segments spoken by *Cheryl, Ed, and John* equals the total number of segments coded as indexed. For example, the upper left-hand cell represents the intersection of *Cheryl* with *Indexed*, and the cell itself tells us that 69 of Cheryl's t-units were coded as *Indexed*.

Contingency tables are useful, functional displays of data that can support a range of analyses. They are relatively easy to construct in Excel (see Excel Procedures 7.1, 7.2, and 7.3 and MAXQDA Procedures 7.1 and 7.2).

	Indexed	Not Indexed	Total
Cheryl	69	73	142
Ed	40	32	72
John	1	2	3
Total	110	107	217

Figure 7.2: A contingency table for data coded in two dimensions: speaker and indexicality.

■ Adding a Second Dimension

To add a second dimension to your analysis, return to the original data sheets

and construct a second coding scheme whereby each segment of data is assigned to one of the categories associated with a second dimension. Refer to the procedures detailed in Chapters 4 and 5 for developing a coding scheme and confirming its reliability. The result will be data that has been coded along two dimensions.

■ Creating Core Contingency Tables

Begin by creating your core contingency tables, one table for each of your data streams. For example, if you have data from four meetings—two design meetings and two management meetings—you must build core contingency tables for all four of these meetings. The results should look similar to Figure 7.3.

Design 1				Management 1			
	Indexed	Not Indexed	Total		Indexed	Not Indexed	Total
Cheryl	49	71	120	Cheryl	69	73	142
Ed	46	55	101	Ed	40	32	72
John	51	72	123	John	1	2	3
Total	146	198	344	Total	110	107	401
Design 2				Management 2			
	Indexed	Not Indexed	Total		Indexed	Not Indexed	Total
Cheryl	59	125	184	Cheryl	36	30	66
Ed	12	39	51	Ed	10	19	29
John	39	127	166	John	71	44	115
Total	110	291	401	Total	117	93	210

Figure 7.3. Core contingency tables for all streams.

Exercise 7.2 Try It Out

Download a copy of the Design and Management meeting data from the book website and build core contingency tables of your own, following the examples described above and shown in the video.



Excel Procedure 7.1: Naming the Data Ranges for Each Dimension

<https://goo.gl/tWqjBL>

1. Go to the tab for the first data stream (e.g., Design 1) and click on the column header to select the first dimension (e.g., Speaker).
2. Click **Insert > Name > Define Name** as shown in Figure 7.4.
3. Label the data range with a name that combines the data stream name and the dimension name (e.g., Design1Speaker) and click **OK** as shown in Figure 7.5.
4. In the same tab, click on the column header to select the second dimension (e.g., Indexicality).
5. Click **Insert > Name > Define Name** and label the data range (e.g., Design1Indexicality) Then click **OK**.
6. Repeat these steps for the data streams on the other tabs until you have all data ranges named and saved.

	A	C	D	E
1	T-Unit	Text	Indexicality	
2				
3		I mean...	Not Indexed	
4		Jesus.	Not Indexed	
5		See I :	Not Indexed	
6		see where	Not Indexed	
7	5 Cheryl	this little thing is?	Indexed	
8	6 Cheryl	It looks to me	Not Indexed	
9	7 Cheryl	like it got wiped out.	Not Indexed	
10	8 Cheryl	That's where	Indexed	
11	9 Cheryl	I wrote it.	Not Indexed	
12	10 Ed	Could be.	Not Indexed	
13	11 Cheryl	And it looks like	Not Indexed	
14		somebody copied it somewhere else or did something.	Not Indexed	

Figure 7.4. Insert a name for the data range.

Define Name

Names in workbook:

Enter a name for the data range:
Design1Speaker

Select the range of cells:
=Design1 data!\$B:\$B

Close OK

Figure 7.5. Choose a name for the data range.



Excel Procedure 7.2: Setting Up a Core Contingency Table

<https://goo.gl/tWqjBL>

You can use the named data ranges to set up and fill the core contingency tables.

1. Insert a worksheet to hold your contingency tables and label it as analysis.
2. Set up a table for your first data stream, with the categories of the first dimension down the side and the categories of the second dimension across the columns.
3. Make sure to include a label for the data stream as show in Figure 7.3



Excel Procedure 7.3: Filling Core Contingency Tables

<https://goo.gl/tWqjbl>

Each cell of the core contingency table is filled with a formula with the following structure:

```
=countifs(  
  Dimension1DataRange,  
  Dimension1Value,  
  Dimension2DataRange,  
  Dimension2Value)
```

To enter this formula into your table:

1. Click in the cell corresponding to the first intersection between your two coding dimensions.

For example: In Figure 7.6 B3 corresponds to the intersection between speaker and indexicality in which Cheryl uses an indexical.

	A	B	C	D
1	Design 1			
2		Indexed	Not Indexed	Total
3	Cheryl	49		49
4	Ed			0
5	John			0
6	Total	49	0	49

Figure 7.6. Add formula to the intersection of Dimension 1 and 2.

2. In the formula bar type
=countifs(
3. Then click **Insert > Name > Paste** and choose the data range for the first dimension (e.g., “DesignSpeaker”).
4. Type a comma.
5. Then click on the cell holding the name of the first coding category for the first dimension. This is the criteria for counting in the first data range.

Continued . . .



Excel Procedure 7.3: Filling Core Contingency Tables (continued)

<https://goo.gl/tWgjbL>

For example (see Figure 7.10), we click on A3 “Cheryl” meaning that we have supplied the first argument to look in the data range we named “Design1Speaker” and find every instance of “Cheryl.”

6. Type another comma.
7. Then click **Insert > Name > Past** and choose the data range for the second dimension (e.g., “Design1Indexicality”).
8. Type another comma.
9. Then click on the cell holding the name of the first coding category in the second dimension (e.g., Indexed).
10. Type) and hit enter.
11. Next, edit the formula to put a \$ in front of the column for the first dimension and in front of the row for the second dimension. This “fixes” or keeps their values constant.

In our example, the formula will now say:

```
=countifs(Design1Speaker,$A3,  
Design1Indexicality,B$2)
```

This formula looks in the data range Design1Speaker for the word Cheryl and then it looks in the data range Design1Indexicality to find the word indexed. It counts the number of times both criteria are met in the data ranges specified.

12. Drag this formula across the columns and then down the rows to fill your table.
13. Continue the same process for each of your data streams.



MAXQDA Procedure 7.1: Creating a Contingency Table with Two Dimensions

<https://goo.gl/tWgjbL>

The following procedure can be used to create contingency tables with two dimensions, like those shown in Figures 7.2 and 7.27. This can be used to produce charts such as those in Figures 7.11, 7.28, and 7.29.

1. Activate the document holding your first data stream.
2. Activate the codes for your first dimension.
3. Select **Visual Tools > Code Relations Browser**.
4. As shown in Figure 7.7, select **Activated codes** for rows.
5. Select **Choose top level codes** for columns.
6. Choose **Co-occurrence of codes** for type of analysis.
7. Make sure that **Only for activated documents** is checked.
8. Click **OK**.
9. In the next input window, select the categories for your second dimension as shown in Figure 7.8.

The pop-up window will now show a core contingency table with colored squares in the cells.

10. Click on the **Display nodes as values** icon to change squares to values.
11. Click on the **Sum** icon to add marginal sums.
12. Click on the **Open as Excel table** icon to open the table in Excel.
13. Repeat these steps for the rest of your documents/data streams.

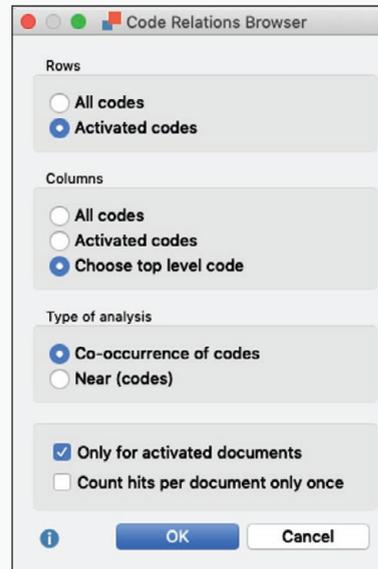


Figure 7.7: Input for a core contingency table in the Code Relations Browser in MAXQDA.

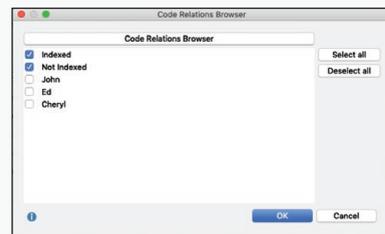


Figure 7.8: Selecting the codes for the second dimension for a core contingency table in the Code Relations Browser in MAXQDA.



MAXQDA Procedure 7.2: Creating a Contingency Table with One Dimension

<https://goo.gl/tWgjbL>

The following procedure can be used to create contingency tables with one dimension, like those shown in Figure 7.20 and on the top of Figure 7.23. These tables can be used to produce charts such as the two top graphs in Figures 7.25.

1. Activate the required documents/data streams.
2. Activate the codes of one dimension.
3. Select **Codes > Overview of Codes** to create a table like that shown in Figure 7.9.

Parent code	Code	Coded segments of all d...	Coded segment...	Author	Creation date	Code alias	% Coded segme...	% Coded segme...	Documents
	Ed	257	101	Cheryl Geisler	.../18 9:56 AM		21.80	23.65	4
	John	409	118	Cheryl Geisler	.../18 10:07 AM		34.69	27.63	4
	Cheryl	513	208	Cheryl Geisler	.../18 9:55 AM		43.51	48.71	4

Figure 7.9: Unedited version of a contingency table with one dimension.

4. Click on the **Only activate codes** icon to see just one dimension.
5. Click on the **Open as Excel table** icon to open the table in Excel.
6. In Excel, label the table appropriately.
7. Delete from the table all columns except for the column with the codes and the column labeled **Coded segments of activated documents** as shown in Figure 7.10.

Management x Speaker	Coded segments of activated documents
Ed	101
John	118
Cheryl	208

Figure 7.10: Contingency table with one dimension made using Overview of Codes.

■ Graphing Dimensions

The rest of this chapter outlines a stepwise process for analyzing graphs constructed from the core contingency tables described in the last section. Before you go on to this analysis, however, we introduce the general procedure for constructing these graphs (see Excel Procedure 7.4 and MAXQDA Procedure 7.3).

Just as your analysis adds a dimension with the addition of a second coding scheme, so too do your graphing techniques. The basic graph used to explore data coded along two dimensions is called a block chart. As you can see from the example in Figure 7.11, a block chart is a three-dimensional graph. Along the x-axis are arrayed the coding categories of the first coding scheme. Along the z-axis are arrayed the coding categories of the second coding scheme. The third or y-axis shows the number of segments. The graph itself shows the distribution of the segments across the two coding dimensions.

In Figure 7.11, for example, we see that in the Management₁ meeting, Cheryl was the most frequent speaker, Ed the second most frequent, and John hardly spoke at all. We also see that this pattern holds true both for contributions that were indexed (shown in blue) and those that were not indexed (shown in orange).

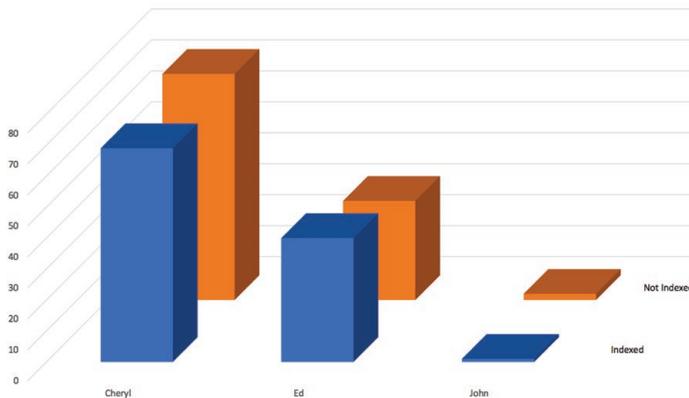


Figure 7.11: A block chart showing two-dimensional data for the Management₁ Stream: speaker and indexicality.

Excel Procedure 7.4: Creating a Block Chart

<https://goo.gl/tWgjbL>

Creating a block chart for data coded along two dimensions involves the same basic process as creating a frequency chart.

1. Select the data in the contingency table, including both column and row headers, but excluding the marginals, as shown to the left of Figure 7.12.

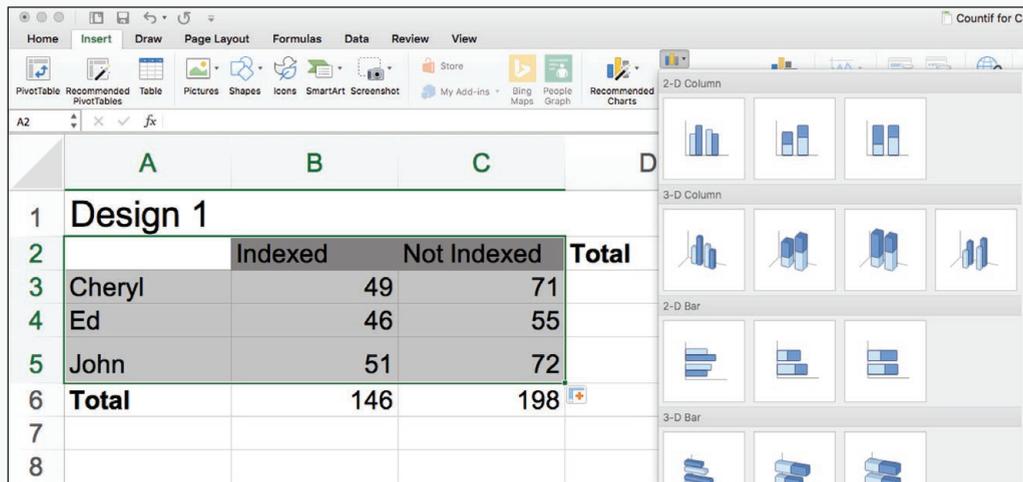


Figure 7.12. Selecting the data and the 3-D column graph to create a block chart.

2. Highlight the cells that include the data you want to graph.
3. Choose the **Insert** tab and choose 3-D column from under the column chart dropdown menu.

Block charts offer the same set of options that you have available for frequency graphs. The chart shown in Figure 7.11, for example, has the following:

- For the chart **Title**, we used the names of the two dimensions as well as the name of the stream (i.e., Speaker x Indexicality x Management₁); this helps us to keep track of what we have graphed.
- For **Gridlines**, we have removed the default gridlines.
- For the **Legend**, we have deselected the **Show legend** box since it provides redundant information.

Create a chart style for your preferred options in a block graph to save considerable time.



MAXQDA Procedure 7.3: Making a Block Chart from MAXQDA Data

<https://goo.gl/tWgjbL>

Once MAXQDA data has been moved into Excel using MAXQDA procedure 7.1 or 7.2, you can use Excel to create a block chart as follows:

1. Select the data in the contingency table, including both column and row headers, but excluding the marginals, as shown to the left of Figure 7.12.
2. Highlight the cells that include the data you want to graph.
3. Choose the **Insert** tab and choose **3-D column** from under the column chart dropdown menu.

Block charts offer the same set of options that you have available for other graphs. The chart shown in Figure 7.11, for example, has the following:

- For the chart **Title**, we used the names of the dimensions as well as the name of the stream (i.e., Speaker x Indexicality x Management).
- For **Gridlines**, we have removed the default gridlines.
- For the **Legend**, we have deselected the **Show legend** box since it provides redundant information.

■ Rotating the Chart

Although block charts have the value of showing distribution data in two dimensions, they have the flaw of being difficult to read and interpret. The forward plane of the graph often obscures the data on the planes behind it. Furthermore, the dimensionality of the graph often makes it difficult to compare magnitudes across planes.

To deal with difficulties in viewing patterns in a block chart, you will find it useful to rotate the chart (see Excel Procedure 7.5 and MAXQDA Procedure 7.4).

Rotating charts can help to refine your understanding of the patterns in three-dimensional block charts. In Figure 7.13, for example, while it was easy to see the contours across the dimension of speaker contribution, it was harder to compare across the dimension of indexicality. Are there more, less, or about the same numbers of indexed units as non-indexed? The view in Figure 7.13 made it hard to tell. Once the graph is rotated to the view shown in Figure 7.14, however, the answer is more easy to come by: The levels are more or less the same.

When you want to compare block charts, it is important that they have the same degree of rotation.

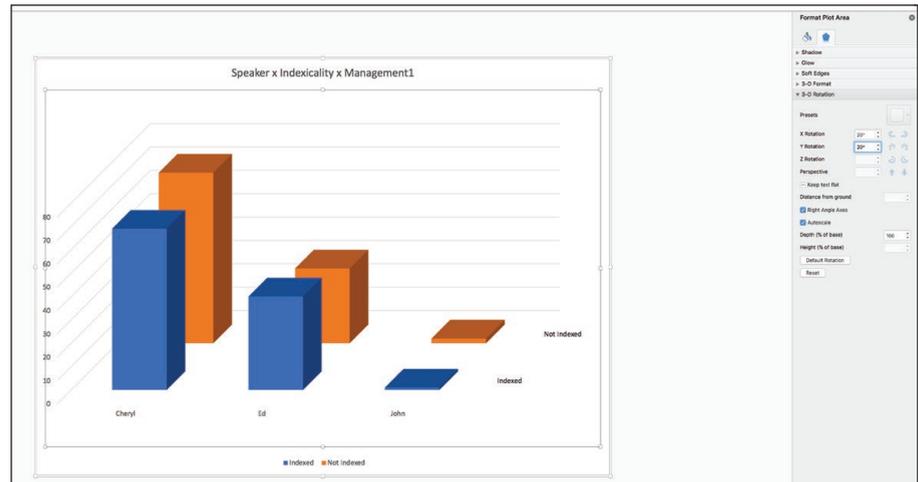


Figure 7.13. Rotating a block chart.

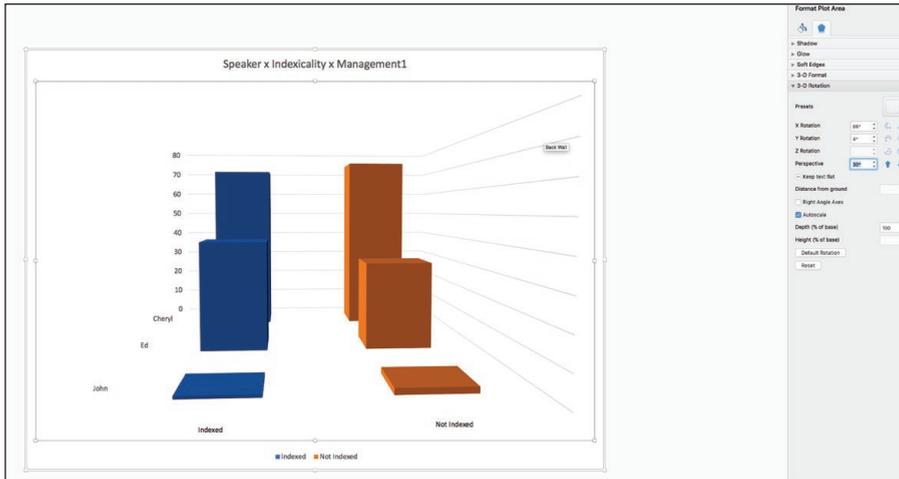


Figure 7.14: A rotated view of the same data shown in Figure 7.13.



Excel Procedure 7.5: Rotating a Block Chart in Excel

<https://goo.gl/tWgjbL>

1. Click on the block chart to select it, then right-click (control+click on Mac) to access the context menu.
2. Choose **3-D Rotation** from the **context** menu.
3. Adjust the values for X rotation and Y rotation until you can see the data that had been hidden (See Figure 7.13).



MAXQDA Procedure 7.4: Rotating a Block Chart for MAXQDA Data

<https://goo.gl/tWgjbL>

1. Click on the block chart to select it, then right-click to access the context menu.
2. Choose **3-D Rotation** from the **context** menu.

Adjust the values for X rotation and Y rotation until you can see the data that had been hidden (See Figure 7.13).

■ Refining Patterns

As you create block charts, you may find that the distribution patterns of certain categories are similar to one another and quite different from patterns for other categories. In Figure 7.15, for example, there appear to be three different patterns. First, we can see business, technology, and special interest publications as sharing the same pattern of dominance through the years. Another group of publications seems to have started from relatively modest numbers in 1996 and to have been increasing. And finally, a third group of publications wasn't on the horizon in 1996, but seems to have become a regular, if still small, publication venue by 2000.

When you see possible clusters in the distribution patterns of your data, you can reorder the categories in your data sheet to create block charts that better represent these clusters (see Excel Procedure 7.6 and MAXQDA Procedure 7.5). An example of a clustered block chart is shown in Figure 7.16. Here, we have placed the categories with the strongest incidence toward the left of the chart, the “medium” categories in the middle, and the relatively late-occurring categories toward the right. This kind of clustered display has greater ability to convey that the publication patterns for the articles.

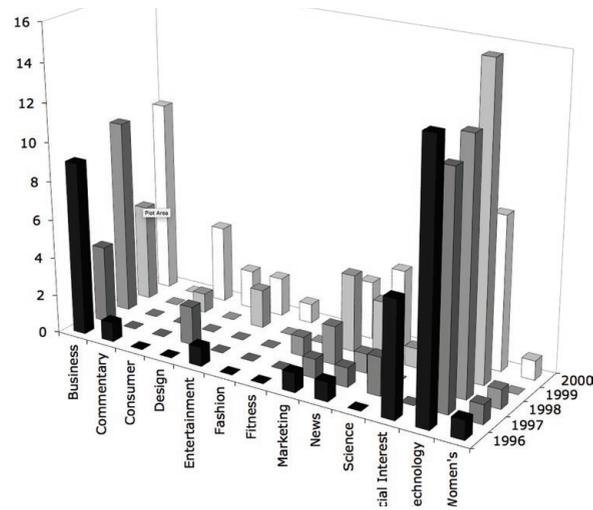


Figure 7.15. A block chart showing how the distribution of articles across publication type varies by year.

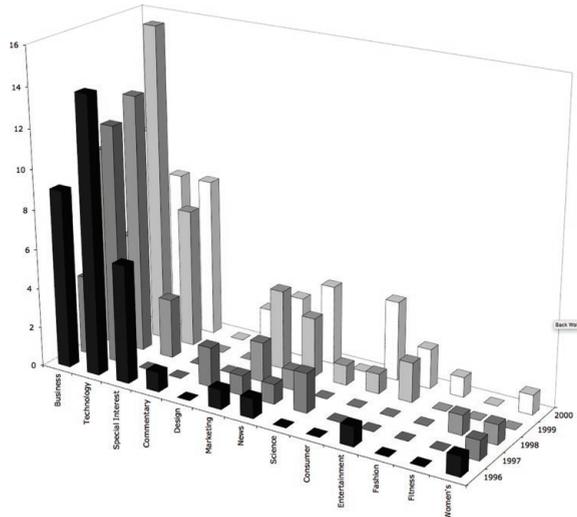


Figure 7.16. Clustering categories to reveal common patterns of distribution.

■ Memo 7.2 Core Contingency Tables

Build a set of core contingency tables for the dimensions of your data that you want to compare. From those tables, build a set of block charts and reflect on the patterns that are revealed. What meaningful relationships are emerging from the data? How can the data be clustered to clarify the visual representation of those relationships?

Exercise 7.3 Try It Out

For a study of PDA users (Personal Digital Assistants), we classified participants into three groups based upon the balance between work and life items in their PDAs: Strong Life, Strong Work, and Integrated. We were interested to understand the relationship between this Work-Life Balance classification and their home situation. The data are shown in the table in Figure 7.18 and available at <https://wac.colostate.edu/books/practice/codingstreams/>.

Use the techniques in this chapter to cluster together those participants with children, with partner but no children, and without family responsibilities.

	Single, living alone	Single, living with partner	Single, living with friends	Single, with children	Married, no children
STRONG LIFE	4	0	0	0	3
INTEGRATED	3	1	0	0	3
STRONG WORK	1	0	0	0	2
Total	8	1	0	0	8

	Married with children at home	Married, children grown	Divorced, with children	Other	Total
STRONG LIFE	2	0	1	0	10
INTEGRATED	9	3	2	0	21
STRONG WORK	6	2	0	0	11
Total	17	5	3	0	42

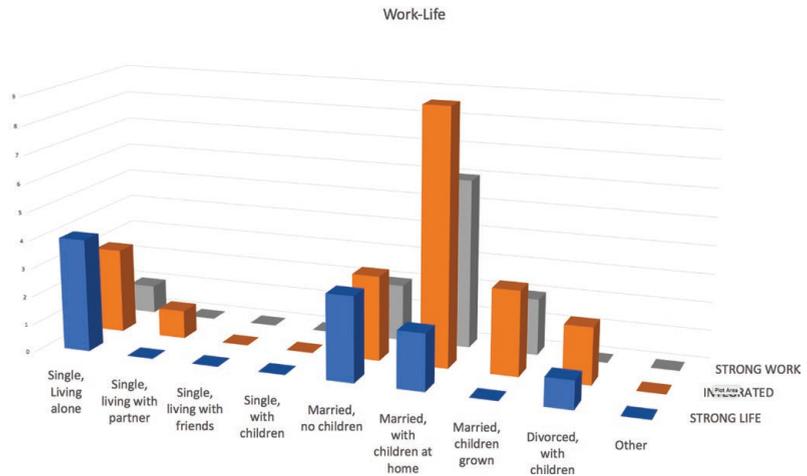


Figure 7.17. Relationship between work-life balance and home situation.

For Discussion: What relationship, if any, do you see between work-life balance and home situation?



Excel Procedure 7.6: Creating Clustered Categories in a Block Chart

<https://goo.gl/tWgjbL>

- Return to your contingency table and insert a cluster value to group the data categories of interest. For example, insert a column to the left to group data categories across the rows or a row across the top to group data categories across the columns (see Figure 7.18).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Cluster Values	1	1	3	2	3	3	3	2	2	2	1	1	3		
	Business	Commentary	Consumer	Design	Entertainment	Fashion	Fitness	Marketing							
1996	9	1	0	0	1	0	0	1							
1997	4	0	0	2	0	0	0	1							
1998	10	0	0	0	0	0	0	1	2						
1999	5	0	1	0	2	0	0	4	3	1	7	16	0		
2000	10	0	4	2	2	1	0	3	4	0	8	8	1		
	1	4	9	5	10	11	12	6	7	8	3	2	13		

Figure 7.18. Assigning a cluster value to categories of data in the totals worksheet.

- Assign a common cluster value to data points that you want to appear close together in the chart. Assign low cluster values to data points that you want to appear on the left of the chart and high cluster values to data points you want to appear on the right (see Figure 7.18).
- Select the entire table (including the cluster values) then click **Data > Sort**.
- In the dialogue box that appears, choose the column holding your cluster values, and then click the **Option button** to choose the direction of the sort.
 - Sort top to bottom to sort by cluster values in a column.
 - Sort left to right to sort by cluster values in a row.

The result will be a clustered set of categories in your contingency table that will automatically update your associated block graphs to look like the one in Figure 7.17.



MAXQDA Procedure 7.5: Creating Clustered Categories in a Block Chart

<https://goo.gl/tWgjbL>

1. Return to your contingency table and insert a cluster value to group the data categories of interest.

For example, insert a column to the left to group data categories across the rows or a row across the top to group data categories across the columns (see Figure 7.18).

2. Assign a common cluster value to data points that you want to appear close together in the chart.

Assign low cluster values to data points that you want to appear on the left of the chart and high cluster values to data points you want to appear on the right (see Figure 7.18).

3. Select the entire table (including the cluster values) then click **Data > Sort**.
4. In the dialogue box that appears click the **Option button** to choose the direction of the sort:
 - Sort top to bottom to sort by cluster values in a column.
 - Sort left to right to sort by cluster values in a row.

The result will be a clustered set of categories in your contingency table that will automatically update your associated block graphs to look like the one in Figure 7.17.

■ Characterizing Dimensions

The analysis of verbal data across dimensions is a complex comparative process that builds up analyses from a basic understanding of how your codes are distributed across all of your data streams. One way to understand it is through the schematic given in Figure 7.19. The base of your analysis starts with characterizing the overall distribution patterns of each dimension (Dimension). From there, we split your data across both sides of the built-in contrast to see if the distribution holds (Dimension x Contrast). If the distribution pattern differs across the contrast, then the contrast may be analytically meaningful. Next, we examine the associations between dimensions and check them across the contrasts (Dimension x Dimension x Contrast). Finally, you take whatever patterns you find across dimensions and contrasts and see whether they hold true across cases (Dimension x Contrast x Stream).

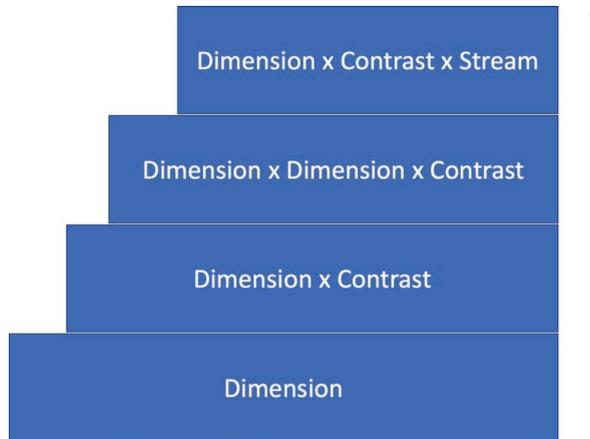


Figure 7.19. Schematic of the complex comparative process of analyzing data across dimensions.

As this schematic indicates, then, the distribution of codes within each dimension is the baseline of our analysis. While many factors arise that can complicate, if not compromise, the adequacy of these baseline characterizations, you will only be able to understand these complications if you begin with

overall characterizations. Each successive layer of complexity in the analysis then becomes meaningful only by comparison to the baseline understanding upon which it is built.

A characterization of the overall pattern for a dimension is a description of the general contours of the data as it is distributed in the categories of your coding scheme—without regard to specific streams and without regard to the built-in contrast. If you have a coding scheme with three categories, for example, you ask yourself how, overall, the data have been placed in those categories: You can think about these distributions in terms of frequencies (e.g., 33 of 100 segments) or proportions (e.g., 33%).

Figure 7.20, for instance, shows the frequency and relative frequency of the coding categories for the categories of speaker (Cheryl, Ed, and John) and for the categories of indexicality (Indexed and Not Indexed). Notice that the totals for each table are identical because exactly the same set of segments has been classified according to both schemes.

■ Calculating Overall Frequencies

To calculate overall frequencies for each dimension and fill tables like those shown in Figure 7.20, sum the appropriate frequencies from the core contingency tables for each case (see Excel Procedure 7.7 and MAXQDA Procedure 7.6). If, for example, we have the same four meetings (two management, two design) described earlier, the overall frequency for Cheryl would be equal to the totals for Cheryl in Management₁, in Management₂, in Design₁, and in Design₂. The values we need to sum are, therefore, in all four core contingency tables, and we need to bring them together in one formula.

■ Establishing Overall Patterns

Because you will be comparing the overall patterns for each dimension to the patterns across built-in contrasts, the relative frequencies give you the best understanding of the overall patterns for each dimension. We can see from Figure 7.20, for example, that overall, 41% of the segments were

indexed and 59% not indexed, a pattern in which, in general, the language is slightly less indexed than not indexed. Once we know this, we can then go on to see whether this overall pattern of indexicality holds across our built-in contrasts.

Speaker		
	Frequency	Relative Frequency
Cheryl	512	0.43
Ed	269	0.23
John	407	0.34
Total	1188	1
Indexicality		
	Frequency	Relative Frequency
Indexed	483	0.41
Not Indexed	705	0.59
Total	1188	1

Figure 7.20. Overall patterns for the dimension of speaker contribution and indexicality.

■ Memo 7.3: Emerging Patterns

Create contingency tables that sum across the data streams in your study. What are the patterns that begin to emerge? Are these patterns what you expected? What might these patterns be telling you about the phenomenon?

■ Graphing Overall Patterns

Sometimes you will want to create a graphic representation of the overall patterns for a dimension to compare with patterns across contrast. With a two-category coding scheme like indexicality, this is not particularly necessary because the overall pattern of 41% versus 59% is not hard to understand. But for dimensions with more numerous categories, graphing can be helpful.

Although you are working with one-dimension data, you will find it easier to use three-dimensional block charts to facilitate comparison with later block charts. Figure 7.21 shows block charts for the overall relative frequencies for both dimensions of the data seen earlier in Figure 7.20. You



Excel Procedure 7.7: Summing Overall Frequencies

<https://goo.gl/tWqjbL>

1. Build a table to hold the frequency values you wish to track.
2. Begin a sum formula (i.e.,=) and then **control+click** (**command click** on Mac) the values to be summed from each contingency table.

For example, Figure 7.20 shows the result of summing Cheryl's contributions from four contingency tables (Design 1, Design 2, Management 1, and Management 2) = D28+J28+D36+J36.



MAXQDA Procedure 7.6: Summing Overall Frequencies

<https://goo.gl/tWqjbL>

1. Activate the documents you wish to sum across.
2. Activate the codes you want to include.
3. Use the **Codes > Overview of Codes** command to create the table.

The sums will be found in the column labeled Coded segments of activated documents.

4. For a table with two dimensions, use the **Visual Tools > Code Relations Browser** to create the table as described in MAXQDA Procedure 7.1.

may notice that they are slightly different from one another. The dimension of speaker contribution is distributed along the x-axis; that for indexicality along the z-axis. This change facilitates comparison with the block charts we will create in general.

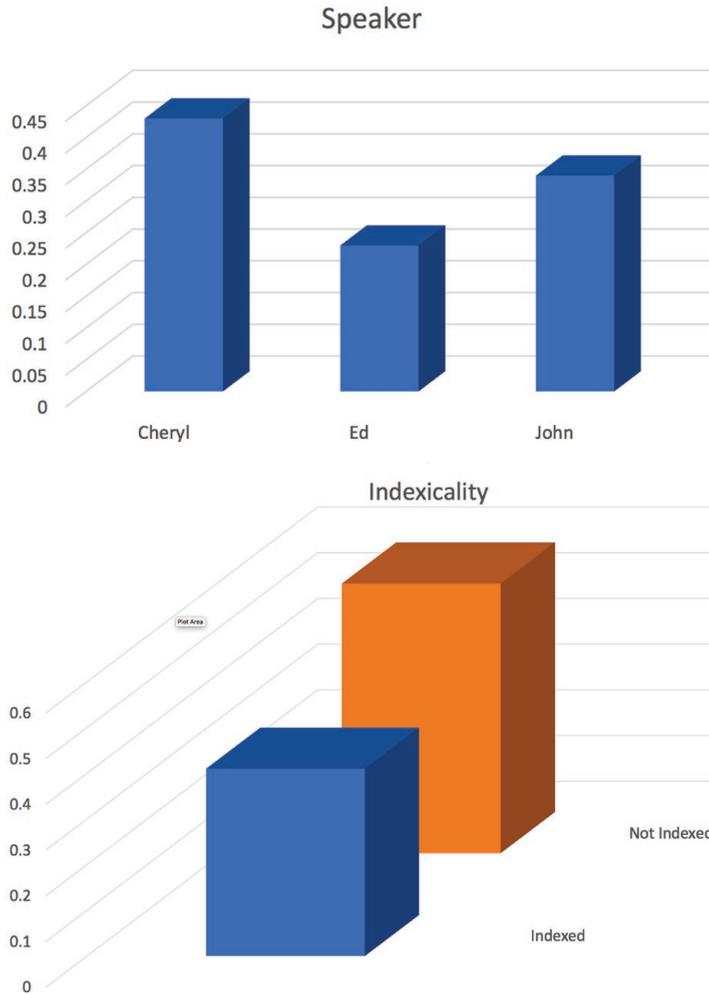


Figure 7.21. Block charts of the overall patterns for speaker contribution and indexicality.

When you create a three-dimensional graph using two-dimensional data, Excel will array that data along the x-axis by default. This will work well for your first dimension, but is not as easy to work with for your second dimension. To move a second dimension from the x-axis to the z-axis, select your graph and choose “switch row/column” from the Insert tab as seen in Figure 7.22.

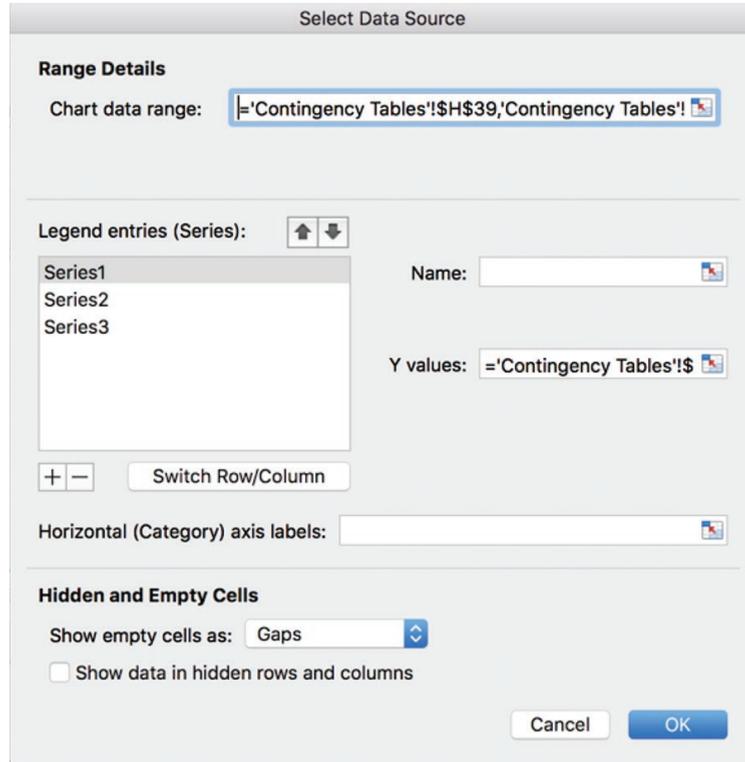


Figure 7.22. Button to switch graph perspective between data on rows and columns.

One final note about producing these block charts. With data laid out as shown in Figure 7.20, you cannot simply drag to select the data to produce block charts like those shown in Figure 7.21. As you can see, the data for relative frequency are not adjacent to the category names on the spreadsheet. While you could copy and rearrange the data in a more appropriate manner, a way does exist to select non-adjacent data (see Excel Procedure 7.8).

■ Checking Patterns across Contrast

Once you understand the overall patterns for each dimension of the data, you can explore how these patterns change across your built-in contrast. Examine each dimension separately.

■ Establishing the Contrast

Begin by constructing a set of contingency tables for each dimension that show the data on either side of your contrast. In Figure 7.23, for example, you can see two sets of contingency tables, one for the dimension of speaker contribution and one for the dimension of indexicality. These have been created by summing the relevant data from our core contingency tables.



Excel Procedure 7.8: Selecting Non-Adjacent Data for Graphing

<https://goo.gl/tWgjbL>

1. From the **Insert** tab, choose a blank 3-d column chart.
2. Right click in the blank chart and select **Series Data**.
3. Click on the graph icon in the field **Chart Data Range** and then select the two non-adjacent data series that you would like to graph.

For example, first choose the data series for the category names (e.g., H37:H39) and then choose the data series for the relative frequencies (e.g., H45:H46).

4. Separate those two data series by a comma and press enter to build your chart with the appropriate labels.
 =ContingencyTables!\$H\$37:\$H\$39,
 Contingency Tables!\$H\$45:\$H\$46

Assign a common cluster value to data points that you want to appear close together in the chart.

Speaker x Management			Speaker x Design			Speaker		
Cheryl	208		Cheryl	304		Cheryl	512	
Ed	101		Ed	168		Ed	269	
John	118		John	289		John	407	
Total	427		Total	761		Total	1188	
Indexicality x Management			Indexicality x Design			Indexicality		
Indexed	Not Indexed	Total	Indexed	Not Indexed	Total	Indexed	Not Indexed	Total
227	200	427	256	505	761	483	705	1188

Figure 7.23. Contingency tables across a built-in contrast.

■ Making the Comparison

Using these contingency tables, we then construct a set of block charts for each dimension and make comparisons with the baseline pattern for that dimension. Two outcomes are possible for these comparisons. On the one hand, the patterns on either side of the contrast may mirror the overall pattern for that dimension. In this case, we have evidence that the contrast may be irrelevant to the dimension. If, on the other hand, the patterns on either side of the built-in contrast shift as we move from one side of the contrast to another, we have evidence that the dimension may be relevant to the contrast. In other words, the dimension may highlight something significant about the contrast that bears further analysis.

For example, Figure 7.24 shows a comparison between the overall pattern for the dimension of speaker contribution and the patterns across the built-in contrast of management versus design. In Graph (A) we see a distribution of the relative frequencies for the dimension of speaker. In Graph (B) we see a distribution of the relative frequencies for the dimension of speaker in design meetings. In Graph C is the distribution for the dimension of speaker in management meetings. We can compare the distributions in B and C to the distribution shown in A so long as the scale of the axes are the same (see Excel Procedure 7.9).

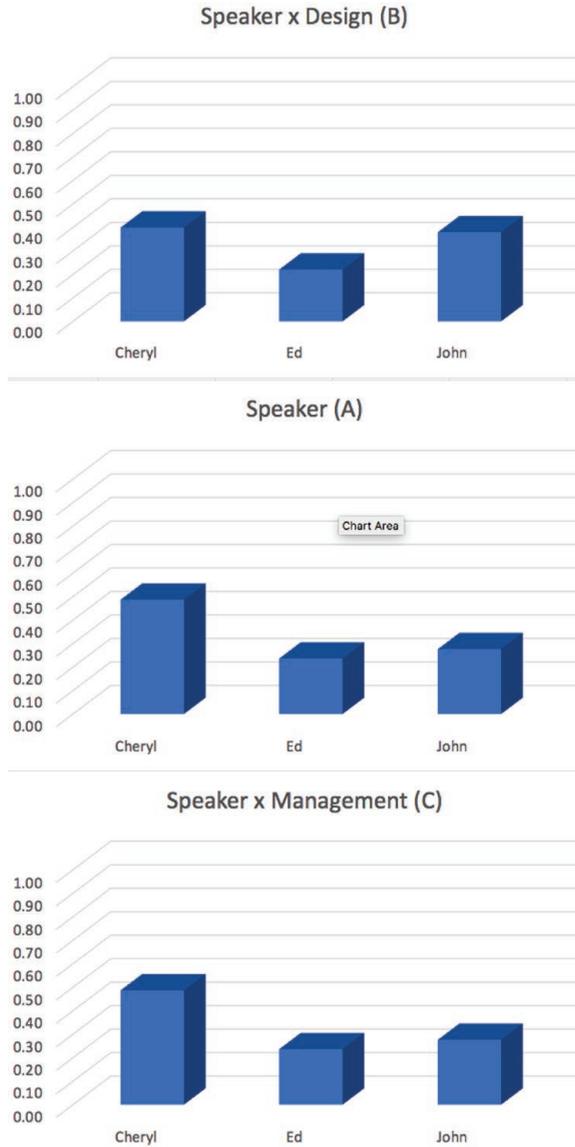


Figure 7.24. Checking the dimension of speaker across contrast of management versus design meetings.

After adjusting the axes, we can see that the contour of the distribution for talk by Cheryl, Ed, and John in the management meetings (B) pretty much mirrors the overall contour (A). The same is true for the design meetings (C) shown at the top.

The general order of speakers—Cheryl, John, Ed—and the general magnitude of the difference between them—Cheryl in the 40%’s, John in the 20%’s and Ed in the 30%’s—seem to suggest that the dimension of speaker contribution is not providing any meaningful insight into the differences between design and management meetings in the data set. As we shall see later on, this evidence is complicated by patterns revealed by later stages of our analysis, but at this stage, it is important to understand the preliminary evidence.

The sample comparisons across the contrast suggests a different story when we look at the second dimension of indexicality. Figure 7.25 shows this comparison. Here, unlike for speaker contribution, we do see substantial differences between the patterns on either side of the contrast and the baseline patterns. As you can see, indexicality appears to be lower, across the board, in design meetings than in management meetings. A look at the relative frequencies confirms this association:

	Overall	Management	Design
Indexed	0.41	0.53	0.34
Not Indexed	0.59	0.47	0.66

In the management meetings, the majority of the units are indexed; the opposite is the case in the design meetings, where the majority of units are not indexed.

In a situation where the baseline patterns are not borne out by the patterns across the built-in contrast, you find evidence that the contrast is associated with the dimension. What this means is that the overall pattern for a dimension (A)—41% indexed, for example—is not an adequate characterization of the data on either side of your built-on contrast—in this case, of neither the management data (C) (where the ratio was 53% vs. 47%) nor of the design data (B) (where the ratio was 34% vs. 66%). Such a pattern suggests, then, that your built-in contrast makes a difference—is associated—with this dimension.

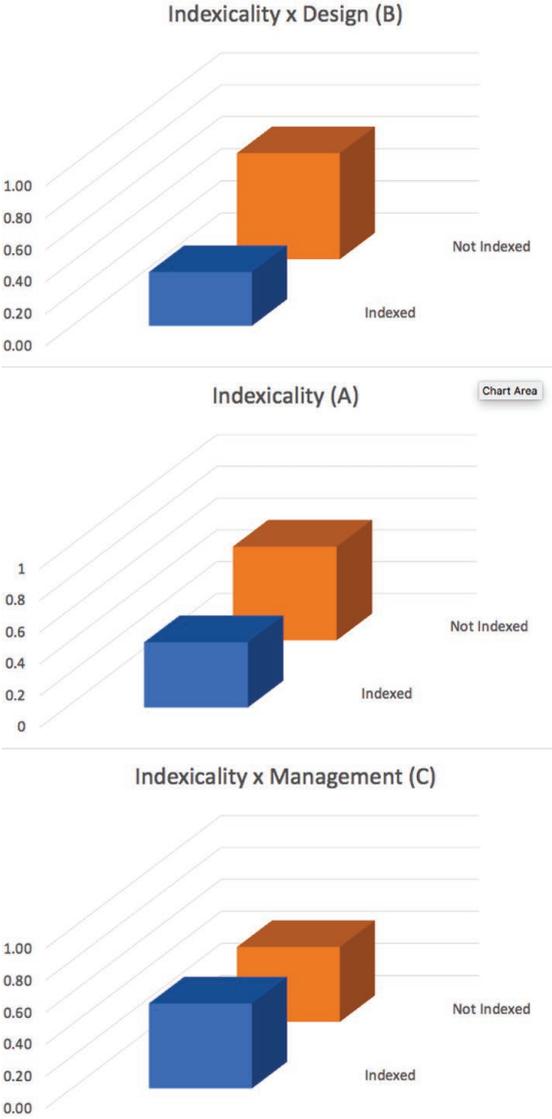


Figure 7.25. Checking the dimension of indexicality across contrast of management versus design meetings.

■ Checking Patterns across Dimensions

Once you have compared the overall pattern within each dimension with the patterns by contrast, the next step is to compare the patterns across dimensions. Looked at in isolation, dimensions of data may appear to be unassociated when a deeper look shows an association. Association might mean that a coding category from one dimension might *always* apply to the same data segments as a coding category from another dimension. Or it could mean that a coding category from one dimension is *never* applied to the same data segments as a coding category from another dimension. Or other patterns might become visible, which would complicate the analysis further. It might mean, in other words, that the categories within the dimensions are associated with each other.

■ Establishing the Associations

The foundation for checking patterns across dimensions is an understanding of the baseline associations across dimensions. The reason for making this kind of check is similar to the reason for checking the baseline distribution for each separate dimension. A baseline distribution showing how multiple dimensions intersect will allow you to determine whether the intersection of those dimensions is meaningful for understanding the built-in contrast. For example, if we know how often Cheryl, Ed, and John use indexed language across both the Management and Design meetings, we can compare that pattern to the ones for Management and Design meetings separately. Any variation from the baseline would indicate that the contrast is meaningfully influential on patterns of indexed speech. A contingency table that shows variations in the two dimensions of your analysis, such as that shown at the top of Figure 7.27, will help you to answer this question.

Developing a contingency table of the baseline associations requires summing data from the core contingency tables both across data streams and across contrasts (see Excel Procedure 7.7 and MAXQDA Procedure 7.6). For example, if we began with four core contingency tables, one for each of the two design meetings and one for each of the two management meetings, the contingency table of baseline associations will sum across those four contingency tables to produce one contingency table, the one shown to the right in Figure 7.27. A block chart of these baseline associations is shown in Figure 7.28.

Design 1				Design Meetings			
	Indexed	Not Indexed	Total		Indexed	Not Indexed	Total
Cheryl	49	71	120	Cheryl	108	196	304
Ed	46	55	101	Ed	58	94	152
John	51	72	123	John	90	199	289
Total	146	198	344	Total	256	489	745
Design 2				Management Meetings			
	Indexed	Not Indexed	Total		Indexed	Not Indexed	Total
Cheryl	59	125	184	Cheryl	105	103	208
Ed	12	39	51	Ed	50	51	101
John	39	127	166	John	72	46	118
Total	110	291	401	Total	227	200	427
Management 1				Management 2			
	Indexed	Not Indexed	Total		Indexed	Not Indexed	Total
Cheryl	69	73	142	Cheryl	36	30	66
Ed	40	32	72	Ed	10	19	29
John	1	2	3	John	71	44	115
Total	110	107	217	Total	117	93	210

Figure 7.27. Basic associations across dimensions.

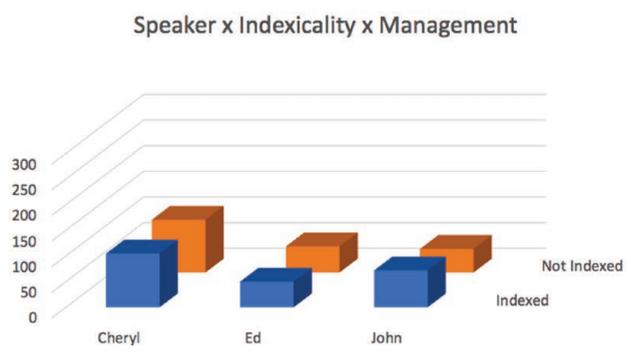
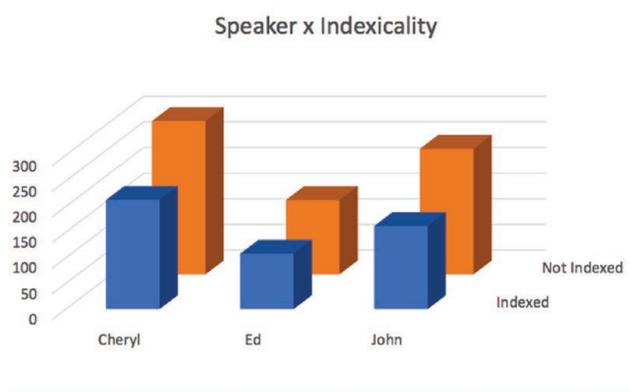
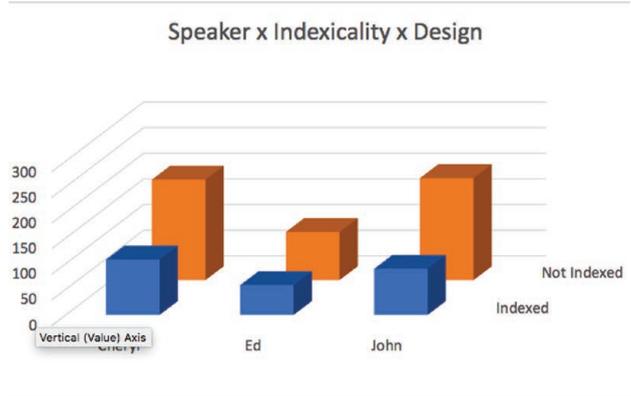


Figure 7.28. Comparing the basic associations across dimensions with the associations across contrasts.



Excel Procedure 7.9: Formatting a Graph Axis

<https://goo.gl/tWgjbL>

1. On your graph, click once on the Y axis (vertical) to select it.
2. Right click on the selected axis and choose **Format Axis** (Figure 7.26).
3. Adjust the bounds to go from a minimum of 0.0 to a maximum of 1.0 (for relative frequencies).

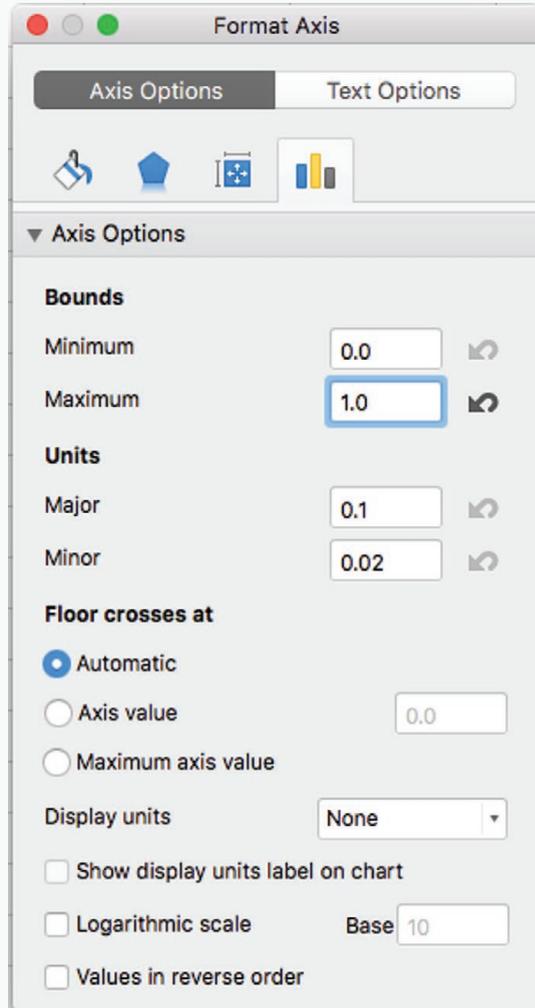


Figure 7.26. Formatting the horizontal (Y) axis.

■ Checking the Patterns

After establishing this baseline association across dimensions, your next step is to look at the two-dimensional patterns formed on either side of your built-in contrast. In Figure 7.27, we see an example of contingency tables that represent the intersections of the first and second dimensions (Speaker and Indexicality) with the built-in contrast (Management versus Design).

To construct these contingency tables, return to your core contingency tables and sum across data streams, but not across the contrast. In the case of our sample data, then, the original four contingency tables (two design meetings and two management meetings) are collapsed to produce two: one for design and one for management.

Block charts are then constructed from these contingencies like those shown in Figure 7.28. Make sure to adjust their scale to be equivalent if necessary.

■ Interpreting Comparisons

To make comparisons between the baseline two-dimensional contours and the contours found in the data across the built-in contrast, trace the baseline contour and then see whether that same contour is found in the data by contrast.

For example, as shown in Figure 7.28, we see that the baseline contour for speaker contribution in the indexed data (in the front plane of the chart) goes from a high with Cheryl, moves to a low for Ed, and then rises to an intermediate value for John. Next, we look to see whether this same contour is repeated in the management data. Here we find the same contour, albeit at a lower rate: starting from a high with Cheryl, moving to a low for Ed, rising to an intermediate value for John. In terms of the indexed data, then, the management data looks a lot like the data overall.

Next, we compare the contours for the non-indexed data. In the baseline associations, we see, as we have noted earlier, that across all speakers, contributions are less likely to be indexed than not indexed. In the management data, however, this is not the case; there, the data appear to be about equally

indexed and not indexed. That is, the orange columns in the back plane are not that much higher than the blue columns in the front plane. Thus this contour looks quite unlike the unlike the baseline patterns of falling and rising that held true for the overall pattern.

Furthermore, John's contributions do not appear to mirror the general pattern. Indeed, in the non-indexed data, they appear to be no higher than Ed's. That is, the contour here is relatively flat as we move from Ed's column to John's column. Again, this is quite different from either the overall pattern or the pattern for indexed data where John's contributions were higher than Ed's.

Having seen two ways in which the management data do not parallel the two-dimensional contours of the baseline data, we next look to see what is going on in the design data. Here we see what looks like a more expected picture. That is, we see a contour that goes from a high with Cheryl, moves to a low for Ed, and then rises to an intermediate value for John in both data that is indexed and data that is not indexed.

Clearly then, something appears to be going on in the management meetings along both dimensions. In the management meetings, talk is relatively more indexed than usual and John's talk stands out as particularly highly indexed. You can see how this comparison focuses attention on John and the management meetings as potentially interesting from an analytic standpoint. These unusual patterns suggest that the two dimensions, speaker contribution and indexicality, have some association in this data set.

■ Checking Patterns across Data Streams

The last stage in an analysis across dimensions is to look at the patterns by comparing the data streams. That is, for any pattern that has emerged in the earlier stages of analysis, we need to know whether that pattern holds true for all streams in our data.

Two outcomes are possible. In the first situation, we may find that the patterns suggested in earlier stages of analysis hold true in specific data streams. In this situation, we can report the more general patterns as good characteri-

zations of our data set. In the second situation, we may find that the generalizations suggested in earlier stages do not hold true of specific data streams. In this situation, we must acknowledge that the more general patterns do not provide an adequate characterization of our data set, that there are differences across the streams.

Block graphs of individual data streams can be constructed from the core contingency tables with which we started this chapter. For our sample data, this yields four block graphs, two for design meetings and two for management meetings. We can then compare these to the contours for the graphs established in the last section. That is, we can ask, if the patterns of association that seem to hold overall—between the dimensions of speaker contribution and indexicality and across the contrast of management and design—hold for the individual data streams.

In the design data, the pattern we want to confirm is that the general contours shown in the middle (A) of Figure 7.29 parallel the contours in the two stream-specific graphs above (B) and below it (C). This means that

- the contributions should, in general, be less indexed than indexed and
- that Cheryl should make the greatest number of contributions, followed by John, followed by Ed.

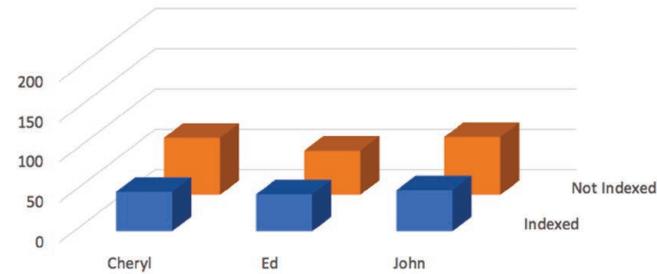
A preliminary comparison suggests that these two patterns do hold in the specific design meetings. That is, the contours of the block chart for Design₁ and Design₂ suggest both that the language is both less indexed than indexed and that the speaker contributions are ordered and of the same magnitude as in the general pattern.

As we noted earlier, the patterns in the management meetings were more complex than those in the design meetings. In particular, we found preliminary evidence that

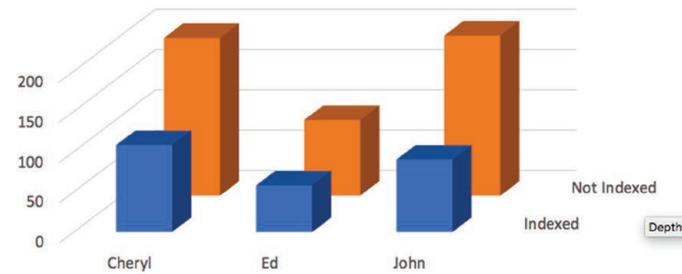
- talk is relatively more indexed than usual, and
- that John's talk was particularly highly indexed.

At this final stage in our analysis, we need to understand how these complex patterns play out in the two management streams. Block graphs for the management data are shown in Figure 7.30.

Speaker x Indexicality x Design 1 (B)



Speaker x Indexicality x Design (A)



Speaker x Indexicality x Design 2 (C)

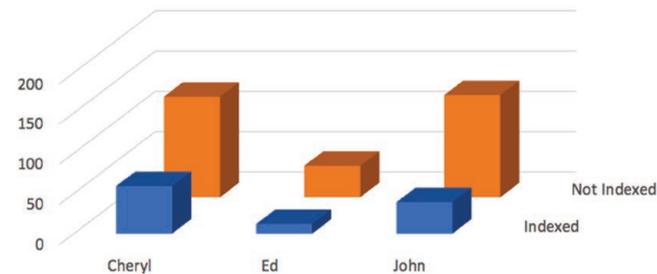
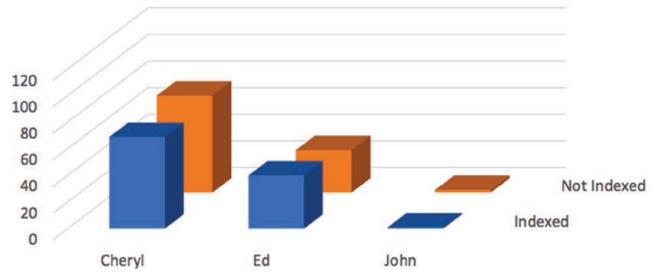
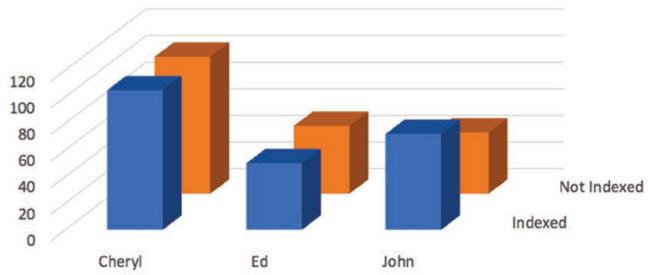


Figure 7.29. Checking patterns across streams in the design data.

Speaker x Indexicality x Management 1 (B)



Speaker x Indexicality x Management (A)



Speaker x Indexicality x Management 2 (C)

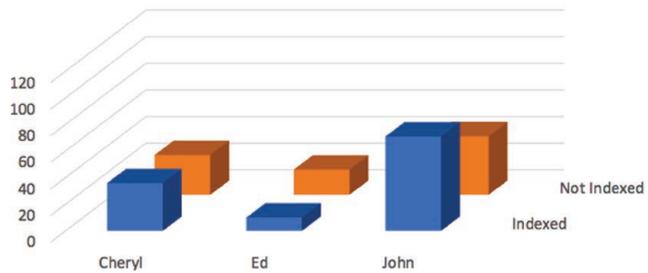


Figure 7.30. Checking patterns across streams in the management data.

Here, unlike in the design data, we find that the general patterns for the management data do not hold for the individual streams. That is, as we see in Figure 7.30, neither Management₁ (B) nor Management₂ (C) look like each other nor like the general pattern (A) already discussed. The patterns for Cheryl and Ed look relatively as expected, but John's talk is very different across the two management meetings. He talks more than anyone else in Management₂ and almost not at all in Management₁. In Management₂, furthermore, his talk is relatively more indexed than not indexed.

What this finding might suggest is that any analysis that looks at the difference between design and management meetings will need to keep in mind that the management meetings differ from each other, particularly on the contributions that John makes. Since our ultimate aim is a descriptive analysis of the data, we are not troubled by the lack of consistency between management meetings, it is just another complexity to account for in the analysis.

Memo 7.4: Comparisons of Data Streams with Baseline Data

Compare your baseline data to your data streams and reflect on whether your streams are consistent with the patterns you have established. If the streams are similar in contour to the baseline, describe those similarities. If the streams are different, describe how they are different.

Putting It All Together

The process presented in this chapter, the process of analyzing across dimensions, involves so many comparisons and distinctions that it is not unusual, when you are done, to lose sight of the big picture. Keeping track of what you find at each stage of analysis can be complex and figuring out the relationships among the stages can be a real challenge.

You will find that the best technique for putting it all together is to write it out. That is, for each level of the analysis, write out any characterizations true

at that level. Then, move to the next level of analysis and see whether those characterizations remain true or must be qualified or withdrawn entirely.

With our sample data, then, we begin with the dimensional analysis. Is there anything we can say about the dimension of speaker contribution that seems to hold overall? Originally it looked like speakers were ordered in terms of relative contribution: Cheryl, John, Ed. This pattern held true, more or less, for the design meetings, but not for management meetings. In Management₂, John dominated; whereas in Management₁, he hardly talked.

What, then, might we say about the dimension of speaker contribution in this data set? The answer would be that while speaker contribution was relatively stable in design meetings, and relatively stable for Cheryl and Ed in management meetings, John's contribution in management meetings was highly variable.

Next we ask about the second dimension. Is there anything we can say about the dimension of indexicality that seems to hold overall? Originally, it looked like the language was generally less indexed than indexed. This pattern held true for the data in both design meetings. It was, however, reversed in the management data, except in Ed's talk in Management₂. In terms of indexicality, then, we see some consistency across our built-in contrast, with talk in the management meetings being more indexed than in the design meetings, with the exception of Ed in Management₂, whose language was not as highly indexed.

Characterizations like these, characterizations built on systematic analysis across dimensions, reflect both possible general statements about a given data set—

Talk in the management meetings was more indexed than in the design meetings.

—and qualifications specific to particular data streams:

The language of Ed in Management₂ meeting was not as highly indexed.

Such characterizations help you to know what is going on in a given data set. They form a rock-solid foundation for the analyses described in the remaining chapters of this book:

- They become the source of questions that you pursue through the temporal analysis described in Chapter 8.
- They form the patterns whose significance you can test in Chapter 9.
- And they become the substratum of the detailed analysis you carry out in Chapter 10.

Take a moment to appreciate what you now know!

■ Memo 7.5: Generalizations

As you complete each level of analysis described in this chapter, write out some generalizations and ponder what kinds of conclusions you might reach? Which of these generalizations seem to be most strongly supported by the data and which are the most analytically interesting to you.

Selected Studies Exploring Patterns across Dimensions

- Karatsolis, A. (2016). Rhetorical patterns in citations across disciplines and levels of participation. *Journal of Writing Research*, 7(3), 425-452. <https://doi.org/10.17239/jowr-2016.07.03.06>
- Swarts, J. (2018). Open-source software in the sciences: The challenge of user support. *Journal of Business and Technical Communication*, 33(1), 60-90. <https://doi.org/10.1177/1050651918780202>