

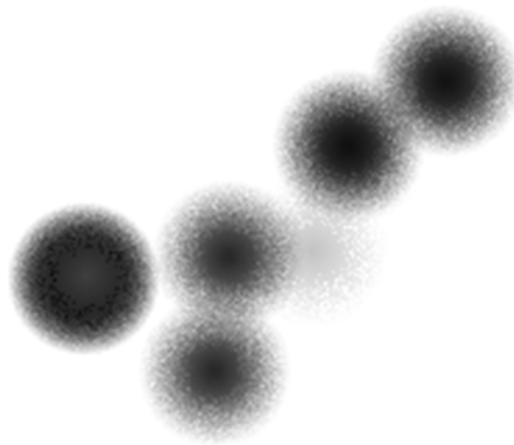
Dynamic Pattern Synthesis

for

Management, Business and Economics

Using IBM SPSS and Microsoft Excel

Philip Haynes



Dynamic Pattern Synthesis

for Management, Business and Economics

Using IBM SPSS and Microsoft Excel

Philip Haynes



White Horse Books



First published in 2018

By White Horse Books

Chichester, UK

www.whb.co.uk

© 2018 Philip Haynes

The right of Philip Haynes to be identified as the author of this work has been asserted by him in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

Published in collaboration with the ESRC CECAN (Centre for the Evaluation of Complexity Across the Nexus) programme

This work is licensed under the Creative Commons Attribution – Non Commercial – No Derivatives 4.0 International License. To view a copy of the license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Printing and photocopying of the entire work, or sections of the work, for the purposes of educational learning, teaching, and training is permitted, provided that the work is properly cited and attributed (for example by inclusion of the title and preliminary pages in any distribution). This permission includes making a digital copy of the complete work available on an educational, or similar, website. The purpose or aim of any distribution must be for the learning and education of others. No attempt should be made to alter the representation of the author or the first publisher (White Horse Books).

Under the terms of the Creative Common Licence, adaptations of the work and/or the use, or reprint, or photocopying, or electronic digital distribution of the work, or part of the work, for the purpose of making a commercial profit, are not permitted.

To cite this work in additional and new - but related - materials, the following text is suggested:

Haynes, P. (2018) *Dynamic Pattern Synthesis for Management, Business and Economics*. Chichester: White Horse Books

Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

British Library Cataloguing-in- Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-9534065-2-4

Contents

List of Figures	6
List of Tables	7
Introduction	8
Chapter one: What is Dynamic Pattern Synthesis (DPS)	9
Exploring complexity	9
Previous influences.....	11
Sampling with DPS.....	11
Replication.....	12
Inferential and probability statistics	12
Effect.....	12
Choosing cases and variables.....	12
The training dataset.....	13
Chapter two: Stage one: Cluster Analysis.....	15
A first HCA cluster model in SPSS	15
Interpreting the cluster agglomeration	17
Interpreting the cluster icicle plot.....	17
Interpreting the cluster dendrogram	19
Chapter three: Stage two: Qualitative Comparative Analysis (QCA)	21
Converting the scale data to crisp set categories and truth tables.....	21
Reconstituting the truth table outcome variable.....	28
Chapter four: Repeating the DPS with longitudinal data	30
2016	30
2017	35
Chapter five: Concluding the DPS.....	39
Variable changes	39
Case and cluster changes.....	39
The qualitative DPS summary	44
Conclusion.....	44
References	45
The datasets	46
Social Media.....	46

List of Figures

<i>Figure 2.1</i> HCA menu.....	15
<i>Figure 2.2</i> HCA sub menu	15
<i>Figure 2.3</i> HCA plots menu	16
<i>Figure 2.4</i> HCA methods menu	16
<i>Figure 2.5</i> HCA icicle plot.....	18
<i>Figure 2.6</i> HCA Dendrogram, 2015.....	20
<i>Figure 3.1</i> Excel custom sort, by cluster	24
<i>Figure 3.2</i> Custom sort, drop down menu.....	24
<i>Figure 3.3</i> Custom sort sub menu	25
<i>Figure 3.4</i> Final set up of the custom sort menu	25
<i>Figure 3.5</i> Rearrangement of custom sort to consider debtors as the outcome variable	28
<i>Figure 4.1</i> Dendrogram for 2016 data.....	30
<i>Figure 4.2</i> Dendrogram for 2017 data.....	35
<i>Figure 5.1</i> Dendrogram comparison: 2015, 2016, 2017.....	40

List of Tables

<i>Table 1.1</i> The complexity domain in research	10
<i>Table 2.1</i> HCA Agglomeration Schedule	17
<i>Table 3.1</i> Scale data converted to binary crisp set scores.....	22
<i>Table 3.2</i> Case by variable ‘truth table’	23
<i>Table 3.3</i> Truth table with three cluster memberships.....	26
<i>Table 3.4</i> Truth table with four cluster memberships.....	27
<i>Table 3.5</i> Truth table with debtors as the outcome variable	29
<i>Table 4.1</i> Scale data converted to binary crisp set scores, 2016	31
<i>Table 4.2</i> Truth table with five cluster memberships, 2016	32
<i>Table 4.3</i> Truth table with debtors as the outcome variable, 2016.....	34
<i>Table 4.4</i> Scale data converted to binary crisp set scores, 2017	36
<i>Table 4.5</i> Truth table with four cluster memberships, 2017	37
<i>Table 4.6</i> Truth table with debtors as the outcome variable, 2017.....	38
<i>Table 5.1</i> Variable mean changes: 2015-17	39
<i>Table 5.2</i> Setup for calculating a longitudinal truth table	42
<i>Table 5.3</i> Longitudinal truth table, sorted by variable patterns	43
<i>Table 5.4</i> Longitudinal truth table with Debtors outcome, 2015-17	43

Introduction

The primary purpose of this short book is to make available a training and instruction manual for those who want to teach themselves, or others, the skills to carry out the method of Dynamic Pattern Synthesis (DPS). The book assumes the basic ability to use Microsoft Excel and IBM SPSS and for the reader to have access to these items of software. It is certainly possible to compute, analyse and teach a DPS exercise with other software also.

This particular book is targeted at those designing, using and teaching research in management, business and economics. It will be of interest to those working in the private, public and non-government sectors. While the example dataset is at the 'meso' (middle), level and concerns business organisations, the DPS method can also be used with macro (country) level and micro (individual) level data.

The development of the book has been influenced by the important work of the ESRC CECAN (Centre for the Evaluation of Complexity Across the Nexus) programme and the method described here is relevant to the ESRC CECAN agenda of promoting methods that can be used to evaluate government policy in complex environments.

For those seeking a more detailed and theoretical account of the methodology that lies behind DPS, you are advised to consult the book: Haynes, P. (2017) *Social Synthesis: Finding Dynamic Patterns in Complex Social Systems* Oxon: Routledge.

In order to work through the example in the book, there is access via a supporting website to download the datasets. This is explained at the end of the book, on page 46.

Dynamic Pattern Synthesis (DPS) seeks to model social science data over time. It does this with small samples of data. This enables it to get a sense of realism while observing both similarity and difference. It borrows from principles of both quantitative and qualitative research methods.

The method was forged from three different areas of research: complexity theory, cluster analysis, and Qualitative Comparative Analysis (QCA). In addition, DPS has been influenced by critical realism and the desire to search for partial causal mechanisms, while always placing these in a dynamic social context. These influences are discussed in chapter one.

Chapter two explains the contribution of cluster analysis to DPS. Chapter three progresses to include the validation and theorisation of the same data model with QCA. This includes consideration of how to promote one variable to an outcome status, for the purposes of evaluation research. Chapter four introduces the longitudinal element, where the consecutive models are linked over time. Chapter five concludes on the qualitative interpretation of what a full DPS model reveals, including the possibility of considering the longitudinal results on a single outcome variable.

Chapter one: What is Dynamic Pattern Synthesis (DPS)

Dynamic Pattern Synthesis (DPS) is a mixed method designed to examine complex patterns in longitudinal datasets. The method combines the strengths of cluster analysis and Qualitative Comparative Analysis (QCA) and looks at dynamic social and economic change over time.

Dynamic Pattern Synthesis is designed to be used with relatively small samples of cases, but can be scaled up to compare several such samples. The method was first presented by the author in 2014 at an international research seminar held at the University of Warwick, UK. The seminar was part of the UK Economic and Social Research (ESRC) Council seminar series on *Complexity and Methods in the Social Sciences*. Before that, the author had been using separate examples of cluster analysis and QCA and beginning to pilot ways of combining them (Haynes & Haynes, 2016; Haynes, 2015, 2014, 2012).

This book is designed to be a practical ‘how to do it guide’. It shows the novice exactly how to compute and calculate a DPS model. For those who want more theoretical background and to consider real world examples of DPS, it is suggested that you read the author’s previous monograph: (2017) *Social Synthesis: Finding Dynamic Patterns in Complex Social Systems*, Oxon: Routledge.

Dynamic Pattern Synthesis has five stages:

Stage 1 The exploration of case patterns

Stage 2 The exploration of variable patterns

Stage 3 The exploration of longitudinal patterns

Stage 4 Theorising about dynamic patterns over time and their qualitative meaning

Stage 5 Reworking the DPS to focus on a specific outcome variable.

Exploring complexity

Dynamic Pattern Synthesis is derived from a view of the world offered by complexity theory. Complexity theory explains that in some areas of science and social science any causal effects are very contingent on the context. For example, causal mechanisms might vary according to the historical time point or the spatial location in which they are situated. In this sense, DPS has some similarities with the approaches of *critical realism* and the idea of *realistic evaluation* (Pawson & Tilley, 1997)

Table 1.1 illustrates the scientific issues when researching the complexity domain. In a simple domain, there is stability in existing cause and effect relationships. An example is the gravitational effect that creates tides. These are predictable to very precise times and can be published in standard tide tables. A complicated domain has strong elements of prediction, but there is the possibility that occasional phenomena and events might disrupt the predictability. Examples are flying an aircraft. Although computerised aircraft are highly predictable and safe in their mechanical

predictability, there is still the very small chance that they can experience major disruption caused by human error and interference, or an exceptional weather event.

In the complex domain, disruptions to cause and effect become much more likely. This includes disruptions when trying to replicate any cause and effect in a different time and place. This is exactly the domain where DPS is designed to operate. Research needs to search for patterns and to examine how consistent and replicable those patterns are over time and space. An example in science research is studying the behaviour of a group of animals, like a herd of elephants or flock of birds. An example in social science is the use of psychological therapy, where a specific therapy may work in some situations, but it is difficult to generalise this therapy to multiple places and over time, as society and its cultures and resulting behaviour changes. Rather than demonstrating and replicating cause and effect, it may be possible to identify patterns about when the therapy is more likely to work or not work, but these patterns will be subject to disruptions and not completely reliable.

Finally, arguably the most difficult task for researchers is to research the domain of chaos. Here instability is the norm and so any pattern replication will be very short term. While pattern analysis may still be relevant, the focus may have to be on single or very small numbers of cases. Weather forecasting is an example in scientific research where there is much instability. Rather than trying to make statements of prediction, such as it will rain, it may be better to say there is a 70% probability of rain.

Table 1.1 The complexity domain in research

	<i>Simple</i>	<i>Complicated</i>	<i>Complex</i>	<i>Chaos</i>
<i>Scientific prediction</i>	Predictable	Bounded prediction	Temporary forecasting	Short term probabilities
<i>Dynamic</i>	Stable	Stable with occasional disruptions	Mix of stability and instability	Unstable
<i>Research evidence</i>	Cause and effect	Linear trends and statistical controls	Pattern analysis	Individual case studies Event probabilities
<i>Science example</i>	Tide tables	Flying an aircraft	Animal behaviour	Weather
<i>Social science example</i>	National controls on price of alcohol influence national levels of consumption	Local level of alcohol consumption and emergency admissions	The use of a psychological therapy to manage depression	Employment attendance for excessive consumers of alcohol

Previous influences

Dynamic Pattern Synthesis is developed from previous research methods that are well tested and widely used. In this sense, it is an incremental development of previous good practice. The two major influences on DPS are: cluster analysis and Qualitative Comparative Analysis (QCA).

Cluster Analysis (CA) was first developed to categorise animal and plant species into similar types (Everitt, 1993). Multiple measurements can be combined in calculations of case-based similarity and difference, to give evidence for possible configurations about which cases are most likely to be similar. A strength of cluster analysis is its ability to compute from multiple variables with a range of scale measures. A weakness is that some small differences between cases and clusters might be mathematical patterns which have little substantive usefulness or meaning in real life.

Qualitative Comparative Analysis (QCA) was developed to theorise about causality in comparative sociology and political science, where the number of countries being compared was relatively small and where different configurations of political and social influence might still lead to the same result (Ragin, 1987). For example, countries could become stable democracies after following different historical paths with only some elements of shared similarity and other patterns of diverse influences.

Sampling with DPS

The fundamental sampling principle in DPS is to choose a group of cases that can reasonably be compared together. Therefore, the first sample needs some degree of expected similarity. These will be cases that share common attributes, but where it is also reasonable to expect them to have differences that are interesting and important to understand. For example, if interested in macro political economy, one would start by comparing a group of countries based on continental geography, or with broadly similar economies. For example, this might be all countries that are members of the Euro currency.

The sampling strategy with DPS is, therefore, purposeful, and not inferential. Inferential sampling is where the researcher takes a random sample from a large population and then uses statistical inference to predict whether sample results can be generalised to the larger population.

DPS sampling is purposeful because there is a deliberate attempt to compare a small group of cases. In this sense, the sampling method is more similar to the approach taken in qualitative social science research, rather than in quantitative research. The sampling method is not inferential and does not involve taking a random sample from a large population.

Replication

Having found a dynamic pattern in a first small sample, it is then reasonable to compare the results with another group of cases. As with any principle of research replication, the next sample group should be chosen so it is a similar size (so that patterns can still be managed and observed) and with the purpose of finding an interesting and logical further comparison. For example, if having studied all the members of the Euro currency, one might then conduct a DPS on all the other countries in the European Community 28 that are not in the single currency.

Inferential and probability statistics

In general, the author does not advocate the use of inferential statistics with DPS, because DPS is not designed to generalise from a sample back to a larger population. Nevertheless, there may be situations where an exception is made. For example, where a small sub sample has been drawn from a previous inferential sample of a national population and the researcher wants to calculate if results found in the small sample could reasonably be inferred back to the original population. In this situation, having decided that clusters were robust and valid, standard statistical operations like ANOVA could be used to see if the differences in mean scores between clusters are statistically significant or not, this when handling the groups as independent samples. Similarly, there are occasions when repeat measure inferential statistics might be used to examine whether cluster changes over time are chance effects, or statistically significant.

Effect

Exploring the effect of variables on clusters is one key element of DPS. It takes place in the second stage of the method. Crisp set QCA is used to diagnose the effect of variables on cluster membership. This is the preferred method in DPS. QCA is used instead of using standard statistical procedures that examine correlation and association for effect. Standard approaches to understanding effect tend to seek an absolute aggregate or average effect measure that best describes all cases. The use of QCA allows for a more complex and appropriate understanding of how effects can be experienced differently and become more relative experiences for individual and small groups of cases.

Choosing cases and variables

The choice of cases is determined by purposeful sampling, as referred to above. The choice of variables is based upon finding as reliable a set of indicators as possible. Ideally, this will be secondary longitudinal dataset from a reputable database and source. At least three, time points are required, in order to examine changing patterns over time.

DPS is flexible enough for additional categorical variables to be added at stage 2. This is because at stage 2, cluster scale variables are reduced to crisp set binary categories. This gives flexibility in understanding the influence of variables on cluster membership. At this stage, additional binary categories can also be added to the model, alongside the scale variables entered at stage 1.

The training dataset

The training dataset used in this book is available at the supporting website (see page 46). It is a fictional dataset that compared 12 research businesses. The dataset is fictional and purely for the purpose for training and teaching the method.

The 12 organisational cases are:

JB Alpha

Cosign Research

Mini Max

System Synthesis

Open Thinking

LKS Data

Strategy Statistics

Visual Research

Ashton Algorithms

Linear Logics

Sun Focus

New Perspectives

The 11 variables (each with a measurement for each of the three years) are:

Business Name

Capital Expenditure 2015, as a percentage of income

Capital Expenditure 2016, as a percentage of income

Capital Expenditure 2017, as a percentage of income

Annual income growth 2015, percentage change from previous year

Annual income growth 2016, percentage change from previous year

Annual income growth 2017, percentage change from previous year

Postgraduate level qualifications 2015, percentage of the workforce

Postgraduate level qualifications 2016, percentage of the workforce
Postgraduate level qualifications 2017, percentage of the workforce
Gender pay gap 2015, percentage of gross income
Gender pay gap 2015, percentage of gross income
Gender pay gap 2015, percentage of gross income
Marketing Expenditure as a percentage of income 2015
Marketing Expenditure as a percentage of income 2016
Marketing Expenditure as a percentage of income 2017
Number of staff per line manager 2015, ratio
Number of staff per line manager 2016, ratio
Number of staff per line manager 2017, ratio
Overseas business 2015, percentage of customers
Overseas business 2016, percentage of customers
Overseas business 2017, percentage of customers
Customers retained 2015, percentage
Customers retained 2016, percentage
Customers retained 2017, percentage
Late payment invoices over one year, 2015, percentage of customers
Late payment invoices over one year, 2016, percentage of customers
Late payment invoices over one year, 2017, percentage of customers
Staff turnover 2015, percentage of staff
Staff turnover 2016, percentage of staff
Staff turnover 2017, percentage of staff
Employee absence with illness 2015, average days absent
Employee absence with illness 2016, average days absent
Employee absence with illness 2017, average days absent

Chapter two: Stage one: Cluster Analysis

The first stage of DPS involves using a type of cluster analysis called Hierarchical Cluster Analysis (HCA).

A first HCA cluster model in SPSS

Using the SPSS file menu, select the following drop-down menus.

File/Analyse/Classify/Hierarchical Cluster.

This reveals the following sub command menu (figure 2.1).

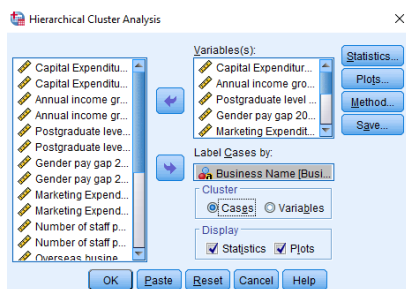


Figure 2.1 HCA menu

Ensure that all 11 variables for the first year of interest, 2015, are entered into the **variable(s)** box.

Ensure that the case labels, Business Name, are entered into the **Label Cases by:** box

Next, select the **Statistics** sub menu

This reveals the sub command menu below (figure 2.2).

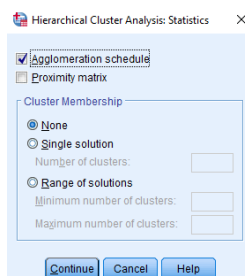


Figure 2.2 HCA sub menu

It is satisfactory for DPS to use the default arrangements. The already selected **agglomeration schedule** will help us to decide about the optimal number of clusters in a DPS model.

Select **Continue** and return to the previous menu.

From here, select Plots. The sub command window below is revealed (figure 2.3).

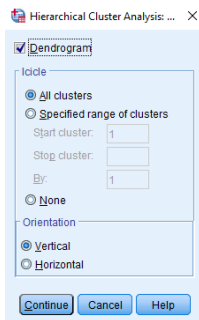


Figure 2.3 HCA plots menu

Tick select the **dendrogram** option.

Leave all other options for the **Icicle** plot in their default values.

Select **Continue**. This returns you to the main hierarchical cluster analysis sub command.

Next, select the **Method** sub menu

This reveals the following sub command menu (figure 2.4).

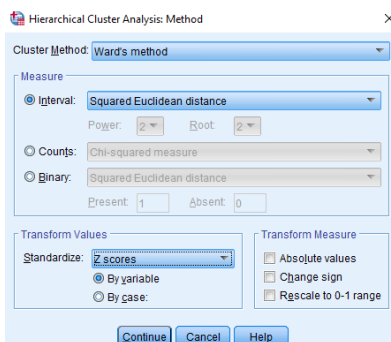


Figure 2.4 HCA methods menu

Select the Cluster Method, Ward's method. Keep the default measure for interval data (**Squared Euclidean Distance**).

Transform the values so they are standardized by selecting **z scores**. This is **by variable**, so keep that default option.

Select **Continue**.

This returns you to the Hierarchical cluster analysis sub command window.

Select **OK**. This will **run** the cluster analysis and the results will be immediately displayed in the SPSS output window.

Interpreting the cluster agglomeration

The first output revealed in the output window is the cluster agglomeration.

This is shown below (table 2.1).

Table 2.1 HCA Agglomeration Schedule

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	5	6	3.236	0	0	7
2	3	9	6.775	0	0	4
3	11	12	11.279	0	0	6
4	2	3	16.898	0	2	6
5	4	8	23.494	0	0	8
6	2	11	31.935	4	3	10
7	5	7	41.172	1	0	9
8	4	10	53.755	5	0	10
9	1	5	68.612	0	7	11
10	2	4	85.274	6	8	11
11	1	2	121.000	9	10	0

The agglomeration illustrates the hierarchical process of combining the 12 cases into a single group. At stage one cases 5 and 6 are combined. These cases are not joined with further cases until stage 7.

The computer calculates a coefficient at each stage. The gradient of the changing coefficient after each stage can be used to guide the optimal number of clusters in a DPS model. For example, when the coefficient gradient changes substantially, this is one argument for taking the stage as the optimum number of substantive clusters. This needs to be checked against an observation of the qualitative meaning of the clusters.

To assist in this gradient judgement, the column containing the coefficients can be cut and pasted in a Microsoft Excel worksheet and a plot constructed to observe the change in gradients.

Interpreting the cluster icicle plot

The cluster icicle plot provides a visual method of seeing the stages of the cluster agglomeration and which cases are most alike according to the variables used in the model. The visual method allows the observer to see the names of the cases, rather than just the case numbers evident in the agglomeration schedule.

In the icicle plot below (figure 2.5), the strongest pairing is cases 5 and 6, *LKS Data* and *Open Thinking*. These are the cases judged to be most alike by the mathematical formula for assessing similarity and they are the first to be combined.

Next, cases 3 and 9, *Mini Max* and *Ashton Algorithms*, are combined.

Next cases 11 and 12, *Sun Focus* and *New Perspectives* are combined.

Next, the case *Cosign Research* joins with the pairing of *Mini Max* and *Ashton Algorithms*, forming the first cluster to have three case members.



Figure 2.5 HCA icicle plot

Interpreting the cluster dendrogram

The second element of the graphical output produced is the cluster dendrogram. This is arguably the single most useful piece of output as it allows the researcher to see clearly the agglomeration of clusters and the overall agglomeration structure.

The dendrogram confirms the strong pairings already discovered, but it also allows one to see the overall possibilities for interpreting the hierarchical structure.

If one is interested in homogenous patterns between cases that require a high degree of similarity between cases, there is a seven cluster structure that includes several outliers.

This is: *Open Thinking* and *LKS Data*, *Strategy Statistics* (outlier), *JB Alpha* (outlier), *Sun Focus* and *New Perspectives*; *Mini Max*, *Ashton Algorithms* and *Cosign Research*, *System Synthesis* and *Visual Research* and *Linear Logics* (Outlier).

Alternatively, there are three clusters that are relatively heterogeneous with less internal similarity.

Open Thinking, LKS Data, Strategy Statistics and *JB Alpha*

Sun Focus, New Perspectives, Mini Max, Ashton Algorithms and *Cosign Research*

System Synthesis, Visual Research and *Linear Logics*

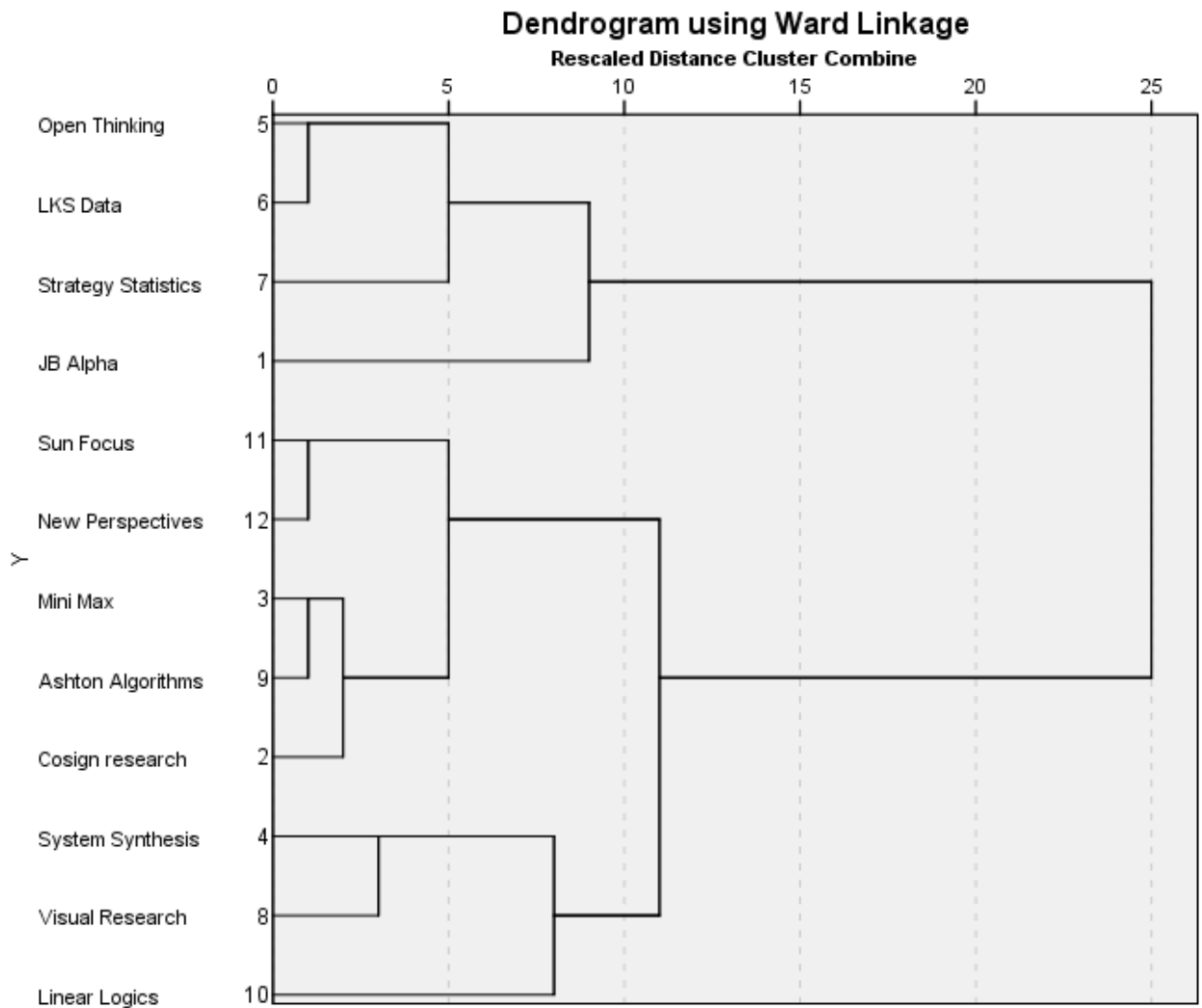


Figure 2.6 HCA Dendrogram, 2015

Next, you can use QCA to test the hypothesis that three useful clusters exist.

Chapter three: Stage two: Qualitative Comparative Analysis (QCA)

The first operation necessary to make the cluster analysis data suitable for QCA is to transform all variable scores to a simple binary score of 0 or 1, where 1 is above the threshold and 0 is below the threshold. This form of QCA is referred to as 'crisp set' (cs) (Rihoux & Ragin, 2009).

In the research example for this chapter, the threshold set to achieve this conversion is the median average. It is also possible for a researcher to use the mean average, or to make a qualitative judgement based on all information that is available about the central tendency and dispersion of the variable scores.

Converting the scale data to crisp set categories and truth tables

If the researcher is using a fixed algorithm for setting the threshold an Excel spreadsheet can be used to automatically convert the scale data into the two binary categories (0,1).

Excel Formula

Place this formula in the cell where you want the crisp set score (0 or 1) to appear.

```
=IF(B2>B$15,1,0)
```

In the example above, the original scale value for a single case score is in B2. The value of this cell is then compared to the threshold set in cell B15. If the value in B2 is greater than B15, a score of 1 is computed. If the value of B2 is less than B15, a score of 0 is computed.

This produced the results in table 3.1, for the first year of data, 2015.

Table 3.1 Scale data converted to binary crisp set scores

<i>Business Name</i>	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	continuecustomers2015	debtors2015	staffturnover2015	sicknessdays2015
JB Alpha	12.3	2.9	72.0	2.0	5.0	0.10	0.0	90.0	2.0	30.0	6.0
Cosign Research	11.1	3.0	54.0	3.0	4.3	0.03	6.0	84.0	2.0	15.0	4.0
Mini Max	4.5	4.0	32.0	3.0	5.2	0.02	0.0	86.0	3.0	16.0	7.0
System Synthesis	9.2	13.7	34.0	7.0	8.1	0.01	12.0	82.0	3.0	13.0	6.0
Open Thinking	8.7	15.6	67.0	1.0	4.2	0.05	6.0	100.0	0.5	16.0	5.0
LKS Data	3.1	8.9	76.0	1.0	4.0	0.05	5.0	98.0	1.0	8.0	4.0
Strategy Statistics	2.1	6.9	90.0	1.0	4.6	0.04	3.0	89.0	1.0	21.0	9.0
Visual Research	9.8	20.3	43.0	3.0	5.7	0.05	8.0	84.0	3.0	2.0	7.0
Ashton Algorithms	7.1	2.8	56.0	1.0	7.2	0.03	4.0	77.0	3.5	14.0	6.0
Linear Logics	7.4	2.3	42.0	8.0	6.1	0.05	23.0	76.0	3.0	9.0	3.0
Sun Focus	5.7	7.1	56.0	2.0	3.7	0.04	4.0	69.0	5.0	7.0	4.0
New Perspectives	4.7	7.3	45.0	4.0	2.3	0.04	11.0	80.0	3.0	11.0	6.0
<i>Mean</i>	<i>7.1</i>	<i>7.9</i>	<i>55.6</i>	<i>3.0</i>	<i>5.0</i>	<i>0.04</i>	<i>6.8</i>	<i>84.6</i>	<i>2.5</i>	<i>13.5</i>	<i>5.6</i>
<i>Median</i>	<i>7.3</i>	<i>7.0</i>	<i>55.0</i>	<i>2.5</i>	<i>4.8</i>	<i>0.04</i>	<i>5.5</i>	<i>84.0</i>	<i>3.0</i>	<i>13.5</i>	<i>6.0</i>
<i>Standard Deviation</i>	<i>3.1</i>	<i>5.6</i>	<i>17.0</i>	<i>2.2</i>	<i>1.5</i>	<i>0.02</i>	<i>6.0</i>	<i>8.5</i>	<i>1.2</i>	<i>6.9</i>	<i>1.6</i>
JB Alpha	1	0	1	0	1	1	0	1	0	1	1
Cosign Research	1	0	0	1	0	0	1	1	0	1	0
Mini Max	0	0	0	1	1	0	0	1	1	1	1
System Synthesis	1	1	0	1	1	0	1	0	1	0	1
Open Thinking	1	1	1	0	0	1	1	1	0	1	0
LKS Data	0	1	1	0	0	1	0	1	0	0	0
Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1
Visual Research	1	1	0	1	1	1	1	1	1	0	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1
Linear Logics	1	0	0	1	1	1	1	0	1	0	0
Sun Focus	0	1	1	0	0	1	0	0	1	0	0
New Perspectives	0	1	0	1	0	1	1	0	1	0	1

Having converted the scale data to binary thresholds, DPS can use these measures to validate the clusters discovered in chapter 2. This is done by constructing a QCA binary ‘truth table’ (Rihoux & Ragin, 2009). A truth table examines the binary score patterns in relation to the outcome variable. The outcome variable below is the cluster groupings proposed in chapter 2.

Adding the clusters in a new column results in the table below (table 3.2).

Table 3.2 Case by variable ‘truth table’

	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	continuecustomers2015	debtors2015	staffturnover2015	sicknesdays2015	cluster
JB Alpha	1	0	1	0	1	1	0	1	0	1	1	1
Cosign Research	1	0	0	1	0	0	1	1	0	1	0	2
Mini Max	0	0	0	1	1	0	0	1	1	1	1	2
System Synthesis	1	1	0	1	1	0	1	0	1	0	1	3
Open Thinking	1	1	1	0	0	1	1	1	0	1	0	1
LKS Data	0	1	1	0	0	1	0	1	0	0	0	1
Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1	1
Visual Research	1	1	0	1	1	1	1	1	1	0	1	3
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1	2
Linear Logics	1	0	0	1	1	1	1	0	1	0	0	3
Sun Focus	0	1	1	0	0	1	0	0	1	0	0	2
New Perspectives	0	1	0	1	0	1	1	0	1	0	1	2

The Excel sort menu can be used to manipulate the table, to make analysis easier.

First select the data, including the case labels in column A, and the variable names in Row 1, as shown in figure 3.1.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	Continuecustomers2015	Debtors2015	Staffturnover2015	Sicknesdays2015	Cluster
2	JB Alpha	1	0	1	0	1	1	0	1	0	1	1	1
3	Cosign research	1	0	0	1	0	0	1	1	0	1	0	2
4	Minute Max	0	0	0	1	1	0	0	1	1	1	1	2
5	System Synthesis	1	1	0	1	1	0	1	0	1	0	1	3
6	Open Thinking	1	1	1	0	0	1	1	1	0	1	0	1
7	LKS Data	0	1	1	0	0	1	0	1	0	0	0	1
8	Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1	1
9	Visual Research	1	1	0	1	1	1	1	1	1	0	1	3
10	Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1	2
11	Linear Logics	1	0	0	1	1	1	1	0	1	0	0	3
12	Sun Focus	0	1	1	0	0	1	0	0	1	0	0	2
13	New Perspectives	0	1	0	1	0	1	1	0	1	0	1	2

Figure 3.1 Excel custom sort, by cluster

Then select **Sort & Filter** from the **Home** menu at the top of the screen. When you see the drop-down menu, select **Custom Sort** (figure 3.2).

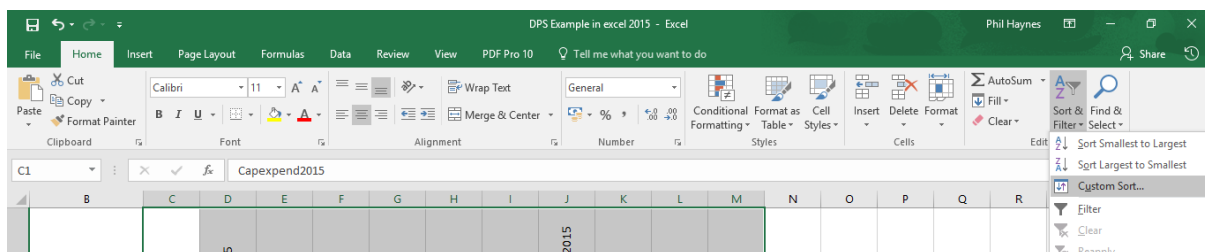


Figure 3.2 Custom sort, drop down menu

The **Custom Sort** sub menu is revealed, as shown in figure 3.3.

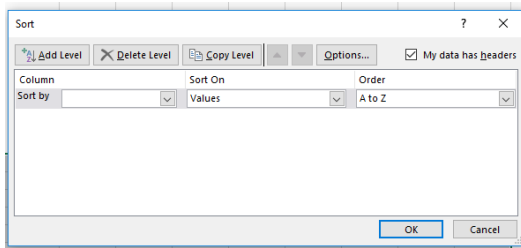


Figure 3.3 Custom sort sub menu

Check that the option **My data has headers** is selected and ticked.

The **Custom Sort** menu allows you to sort by Cluster. The drop-down **Sort by** menu will reveal a list of all the possible variables, including the cluster that the case belongs to.

Use the **Add Level** to Sort by cluster first. Cluster then becomes the first level in your sort.

Keep the default option of **Sort On: Values**.

Ensure that the **Order** is **smallest to largest**, so that cluster 1 will be shown at the top of the resulting sort table.

Next, add each of the other variables to the sort menu in turn, until all variables are listed.

The resulting sort menu is shown below (figure 3.4).

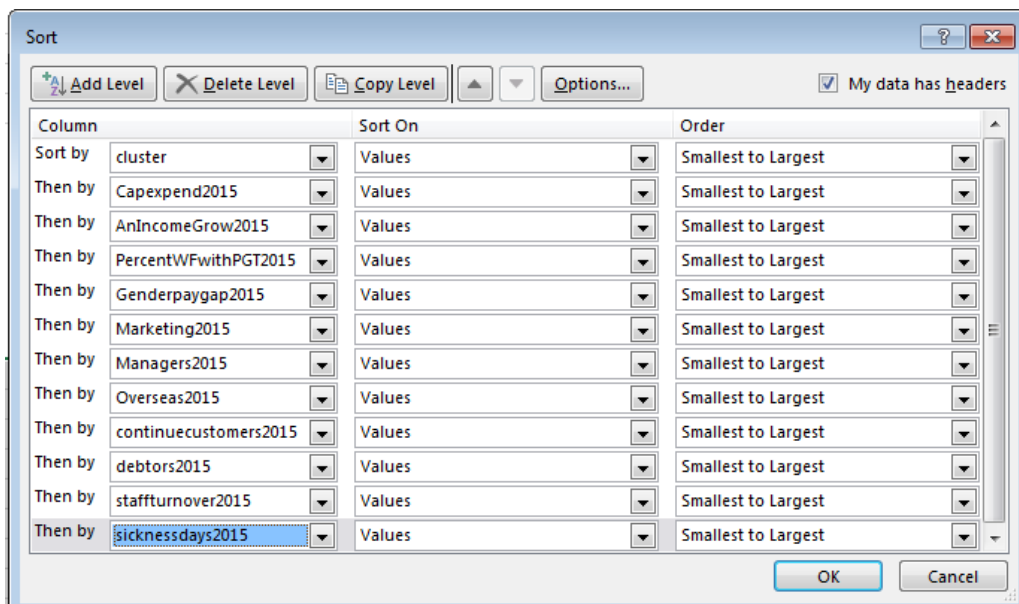


Figure 3.4 Final set up of the custom sort menu

Now click on **OK** and run the sort.

The resulting table is shown below (table 3.3).

Table 3.3 Truth table with three cluster memberships

	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	continuecustomers2015	debtors2015	staffturnover2015	sicknesdays2015	cluster
Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1	1
LKS Data	0	1	1	0	0	1	0	1	0	0	0	1
JB Alpha	1	0	1	0	1	1	0	1	0	1	1	1
Open Thinking	1	1	1	0	0	1	1	1	0	1	0	1
Mini Max	0	0	0	1	1	0	0	1	1	1	1	2
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1	2
New Perspectives	0	1	0	1	0	1	1	0	1	0	1	2
Sun Focus	0	1	1	0	0	1	0	0	1	0	0	2
Cosign Research	1	0	0	1	0	0	1	1	0	1	0	2
Linear Logics	1	0	0	1	1	1	1	0	1	0	0	3
System Synthesis	1	1	0	1	1	0	1	0	1	0	1	3
Visual Research	1	1	0	1	1	1	1	1	1	0	1	3

Where a group of cases in a cluster share the same threshold score for a variable, this is called a PRIME IMPLICANT (Rihoux & Ragin, 2009). It is evidence of the validity or usefulness of the cluster. It shows coherence in the cluster, as evidenced by the shared variable characteristics for that cluster.

In the above example, clusters 1 and 3 have several shared prime implicants and therefore coherence is proven. In table 3.3, prime implicants are shown in bold text. However, cluster 2, has no prime implicants and seems to lack coherence. There are several ‘near misses’, where only one case threshold observation is different to the other cases. (For more information about near misses, see Haynes, 2017, pages 65-67.)

One strategy with such near misses, is to check the scale variable for the single uncharacteristic score in that cluster. If it is close to the threshold set, it can be over ruled, and that variable score accepted as satisfactory to be a prime implicant for the cluster.

In this situation (table 3.3), however, the previous cluster analysis dendrogram (figure 2.6), revealed separate clusters that constructed the present cluster 2, with a sub grouping of the pair, *Sun Focus* and *New Perspectives*, and these separate from the other three businesses, *Mini Max*, *Ashton Algorithms* and *Cosign Research*. Given this evidence from the detailed analysis, the truth table sort is re run to test for four clusters.

The resulting sort, with the prime implicants indicated in bold text, is shown below (table 3.4).

Table 3.4 Truth table with four cluster memberships

	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	continuecustomers2015	debtors2015	staffturnover2015	sicknesdays2015	cluster
Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1	1
LKS Data	0	1	1	0	0	1	0	1	0	0	0	1
JB Alpha	1	0	1	0	1	1	0	1	0	1	1	1
Open Thinking	1	1	1	0	0	1	1	1	0	1	0	1
New Perspectives	0	1	0	1	0	1	1	0	1	0	1	2
Sun Focus	0	1	1	0	0	1	0	0	1	0	0	2
Mini Max	0	0	0	1	1	0	0	1	1	1	1	3
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1	3
Cosign Research	1	0	0	1	0	0	1	1	0	1	0	3
Linear Logics	1	0	0	1	1	1	1	0	1	0	0	4
System Synthesis	1	1	0	1	1	0	1	0	1	0	1	4
Visual Research	1	1	0	1	1	1	1	1	1	0	1	4

This 4 cluster structure appears to be more valid and coherent. The pairing of cases in cluster 2 share 7 prime implicants and cluster 3 shares 3 prime implicants.

The truth table reveals a pattern of case clusters and their relationship with variable scores. The pattern can be summarised using Boolean algebra, where upper case words represent above threshold scores and lower-case words represent below threshold scores.

Cluster 1: PGT * genderpay * MANAGERS * CONTINUING * debtors

Cluster 2: cappexpend * INCOME * marketing * MANAGERS * continuing * DEBTORS * staffturnover

Cluster 3: income * managers* STAFFTURNOVER

Cluster 4: CAPPEXPEND * pgt * GENDERPAY * MARKETING * OVERSEAS * DEBTORS * staffturnover

Having confirmed the cluster structure, it is also possible to explore causality with regard to an outcome variable. For example, if we are interested in why some businesses have more debtors than others, we can make debtors the outcome variable.

Reconstituting the truth table outcome variable

The default approach with DPS is to use the number of clusters as the outcome variable in the QCA stage. This enables the variable patterns to be checked against case clusters, to ensure a logical number of clusters is chosen. If the researcher has decided on the optimum number of clusters that is useful, it is then possible to reconstitute the truth table so that you can consider a different outcome. First, reopen the **Custom Sort** menu and ensure that the chosen outcome variable (debtors) is the first level of the new sort.

You can now order the other variables according to your own preference. The logic applied in the example menu below (figure 3.5) is to put the variables that the researcher thinks are most likely to influence the outcome, higher up the sorting hierarchy.

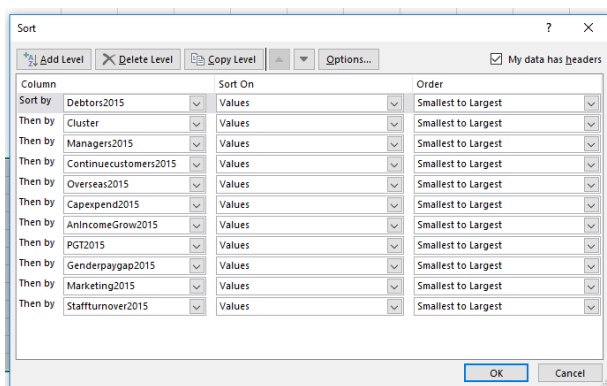


Figure 3.5 Rearrangement of custom sort to consider debtors as the outcome variable

Having clicked on OK, the resulting truth table is produced (table 3.5). The presentation in table 3.5 also required the outcome variable, Debtors, to be cut and pasted into the last column of the table, for ease of presentation.

Table 3.5 Truth table with debtors as the outcome variable

	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	Continuecustomers2015	Staffturnover2015	Sicknessdays2015	Cluster	Debtors2015
Strategy Statistics	0	0	1	0	0	1	0	1	1	1	1	0
LKS Data	0	1	1	0	0	1	0	1	0	0	1	0
JB Alpha	1	0	1	0	1	1	0	1	1	1	1	0
Open Thinking	1	1	1	0	0	1	1	1	1	0	1	0
Cosign Research	1	0	0	1	0	0	1	1	1	0	2	0
Mini Max	0	0	0	1	1	0	0	1	1	1	2	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	2	1
New Perspectives	0	1	0	1	0	1	1	0	0	1	3	1
Sun Focus	0	1	1	0	0	1	0	0	0	0	3	1
Linear Logics	1	0	0	1	1	1	1	0	0	0	4	1
System Synthesis	1	1	0	1	1	0	1	0	0	1	4	1
Visual Research	1	1	0	1	1	1	1	1	0	1	4	1

There is one prime implicant that correlates with a below threshold number of debtors. This is above threshold scores for the percentage of customers who are continuing. There are several other variables where there are near misses. One of these is the cluster number match with outcome. In fact, all members of cluster 1 have a lower number of debtors. The cluster exception is cluster 2 where one member, *Cosign Research*, also has a relatively low number of debtors.

If the researcher chooses to focus on just cluster 1, the causal relationship is stronger, with three additional prime implicants, an above threshold ratio of managers to staff, a below threshold (smaller) gender pay gap, and an above threshold percentage of staff with postgraduate qualifications.

For cluster 1, we can conclude with an interesting Boolean simplification statement:

CONTINUING * MANAGERS * genderpay * PGT = debtors

This implies evidence that having an adequate ratio of managers who are selected on ability and have good qualifications improves relationships with customers and the contracting of work in such a way that late payments and lost income are reduced.

Chapter four: Repeating the DPS with longitudinal data

Having developed a combined model for 2015, the next step is to repeat the pattern analysis with first 2016, and then 2017, data.

This chapter shows the key results for 2016 and 2017, but does not repeat the detail about how to use SPSS and Excel to carry out the analysis.

2016

Figure 4.1 shows the resulting dendrogram after calculating a HCA cluster analysis for the 2016 data.

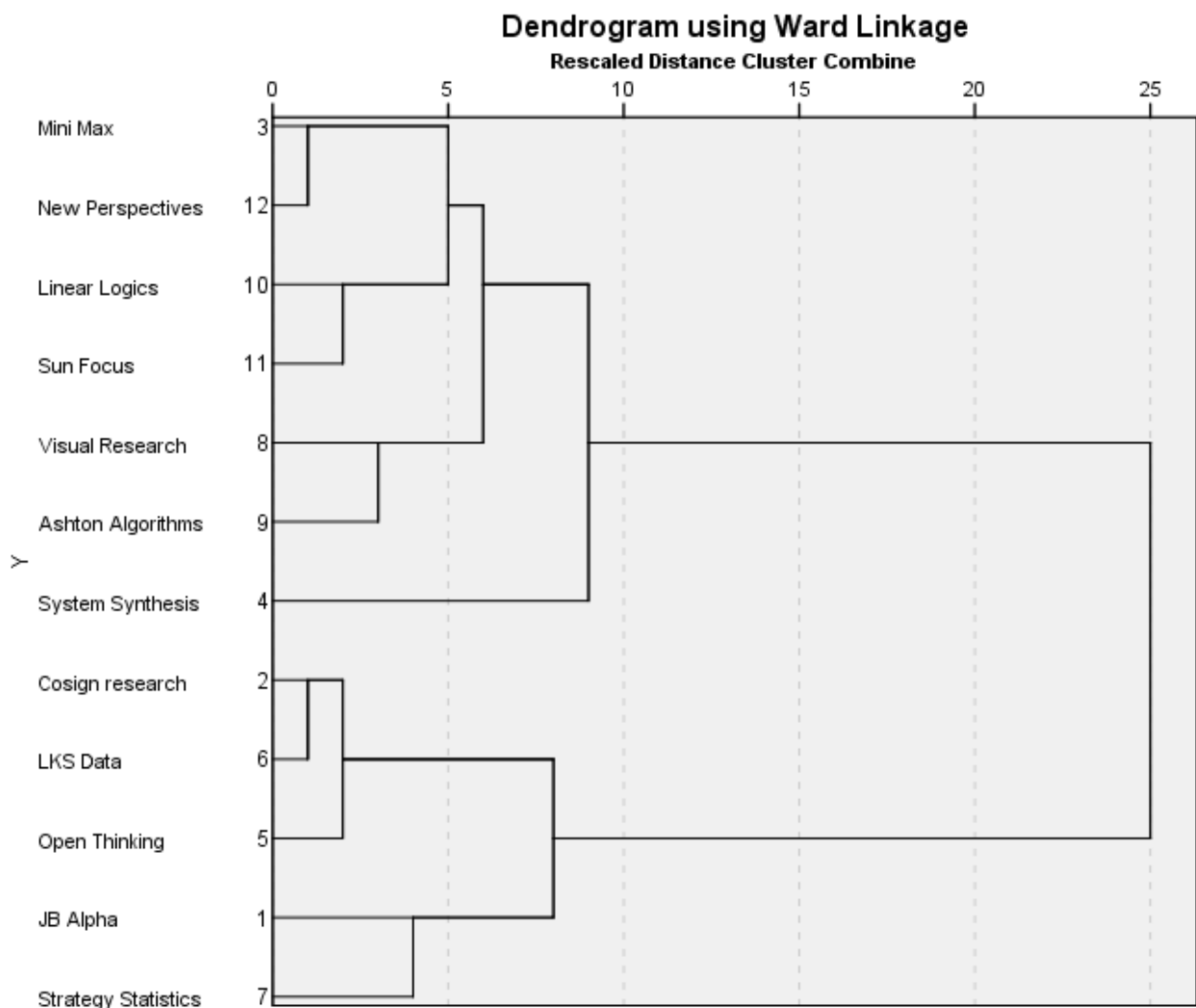


Figure 4.1 Dendrogram for 2016 data

The dendrogram for 2016 data suggests either a three-cluster structure (two clusters and one outlier), or a five-cluster structure (four clusters and one outlier). Table 4.1 shows the conversion of the scale data used for the cluster analysis into binary ‘crisp set’ scores suitable for the QCA.

Table 4.1 Scale data converted to binary crisp set scores, 2016

	Capexpend2016	AnIncomeGrow2016	PGT2016	Genderpaygap2016	Marketing2016	Managers2016	Overseas2016	Continucustomers2016	Debtors2016	Staffturnover2016	Sicknessdays2016
JB Alpha	14.6	3.1	70.0	1.0	4.1	0.09	0.0	92.0	2.0	15.0	5.0
Cosign Research	11.0	4.2	55.0	3.0	5.2	0.03	4.0	90.0	2.0	7.0	3.0
Mini Max	5.5	6.7	49.0	2.0	4.9	0.03	2.0	87.0	3.0	3.0	6.0
System Synthesis	8.4	15.0	40.0	6.0	7.2	0.02	10.0	80.0	4.0	7.0	6.0
Open Thinking	8.3	3.0	65.0	2.0	6.4	0.06	5.0	98.0	1.0	4.0	6.0
LKS Data	4.1	4.5	75.0	2.0	5.2	0.04	2.0	97.0	1.0	7.0	2.0
Strategy Statistics	4.5	5.0	85.0	0.0	5.2	0.05	3.0	92.0	2.0	17.0	5.0
Visual Research	10.2	-5.6	45.0	4.0	5.5	0.04	9.0	87.0	3.0	6.0	7.0
Ashton Algorithms	8.1	-1.0	58.0	2.0	9.0	0.03	4.0	85.0	4.0	9.0	5.0
Linear Logics	8.2	0.6	51.0	3.0	7.1	0.06	20.0	78.0	2.0	5.0	4.0
Sun Focus	6.3	-2.3	59.0	2.0	5.1	0.03	8.0	72.0	2.5	2.0	3.0
New Perspectives	4.6	7.1	52.0	3.0	4.1	0.05	14.0	84.0	2.0	9.0	6.0
<i>Mean</i>	<i>7.8</i>	<i>3.4</i>	<i>58.7</i>	<i>2.5</i>	<i>5.8</i>	<i>0.04</i>	<i>6.8</i>	<i>86.8</i>	<i>2.4</i>	<i>7.6</i>	<i>4.8</i>
<i>Median</i>	<i>8.2</i>	<i>3.7</i>	<i>56.5</i>	<i>2.0</i>	<i>5.2</i>	<i>0.04</i>	<i>4.5</i>	<i>87.0</i>	<i>2.0</i>	<i>7.0</i>	<i>5.0</i>
<i>Standard Deviation</i>	<i>3.0</i>	<i>5.0</i>	<i>12.5</i>	<i>1.4</i>	<i>1.4</i>	<i>0.02</i>	<i>5.5</i>	<i>7.3</i>	<i>0.9</i>	<i>4.3</i>	<i>1.5</i>
JB Alpha	1	0	1	0	0	1	0	1	0	1	0
Cosign research	1	1	0	1	0	0	0	1	0	0	0
Mini Max	0	1	0	0	0	0	0	0	1	0	1
System Synthesis	1	1	0	1	1	0	1	0	1	0	1
Open Thinking	1	0	1	0	1	1	1	1	0	0	1
LKS Data	0	1	1	0	0	0	0	1	0	0	0
Strategy Statistics	0	1	1	0	0	1	0	1	0	1	0
Visual Research	1	0	0	1	1	0	1	0	1	0	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	0
Linear Logics	1	0	0	1	1	1	1	0	0	0	0
Sun Focus	0	0	1	0	0	0	1	0	1	0	0
New Perspectives	0	1	0	1	0	1	1	0	0	1	1

Table 4.2 Truth table with five cluster memberships, 2016

	Capexpend2016	AnIncomeGrow2016	PGT2016	Genderpaygap2016	Marketing2016	Managers2016	Overseas2016	Continuecustomers2016	Debtors2016	Staffturnover2016	Sicknesdays2016	Cluster
Sun Focus	0	0	1	0	0	0	1	0	1	0	0	1
Mini Max	0	1	0	0	0	0	0	0	1	0	1	1
New Perspectives	0	1	0	1	0	1	1	0	0	1	1	1
Linear Logics	1	0	0	1	1	1	1	0	0	0	0	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	0	2
Visual Research	1	0	0	1	1	0	1	0	1	0	1	2
System Synthesis	1	1	0	1	1	0	1	0	1	0	1	3
LKS Data	0	1	1	0	0	0	0	1	0	0	0	4
Open Thinking	1	0	1	0	1	1	1	1	0	0	1	4
Cosign Research	1	1	0	1	0	0	0	1	0	0	0	4
Strategy Statistics	0	1	1	0	0	1	0	1	0	1	0	5
JB Alpha	1	0	1	0	0	1	0	1	0	1	0	5

The first QCA in table 4.2 attempts to validate the five clusters while examining the overlap of clusters. Prime implicants are shown as bold text.

The QCA suggests an overlap in variable definitions of clusters exists between clusters 2 and 3 (indicated by a broken line in table 4.2). The dendrogram in figure 4.1 does confirm some proximity between *Ashton Algorithms* and *Visual Research* with *System Synthesis*.

Boolean algebra can be used to summarise the variable influences on each cluster.

Cluster 1: continuing

Cluster 2: income * MARKETING * managers * continuing * DEBTORS

Cluster 3 (with 2): MARKETING * managers * continuing * DEBTORS

Cluster 4: CONTINUING* debtors * staffturnover

Cluster 5: PGT * gender * marketing * MANAGERS * overseas * CONTINUING* debtors * STAFFTURNOVER * sickness

Cluster 1 is poorly defined, with only one shared prime implicant, that is a below threshold score for the number of continuing customers. Table 4.2 and dendrogram (figure 4.1) show there will be much more case similarity if this larger cluster is separated into two pairs. But this is also a good example where 'near misses' can be assessed. For the capital expenditure variable, only *Linear Logics* scores above threshold, but an examination of table 4.1 shows that this is because it scores at the median point (8.2). Given the established principles of DPS (Haynes, 2017), it is reasonable to overrule the above threshold binary score. Therefore, the score is overruled and below threshold spend for capital expenditure becomes a prime implicant for cluster 1.

The next near miss is postgraduate qualifications, where *Sun Focus* is the only case above threshold. Table 4.1 reveals that *Sun Focus*' score for the percentage of staff with postgraduate qualifications is 59 very close to the mean of 58.7. While it is the first case in the population of cases to score above the threshold set at the median (56.5), it is clearly sitting close to the central tendency. This brings it into line with the other three cases in cluster 1. Again, this adds a further prime implicant to the cluster.

The next near miss is expenditure on marketing. *Linear Logics* is the outlier, with an above threshold score. An examination of table 4.1 shows that its actual score is 7.1, noticeably above the median and mean. For this reason, an overrule is not appropriate.

The next near miss is overseas business. *Mini Max* is the exception for the cluster, with a below threshold score. It has a score of 2, and this is well below the median (4.5) and mean (6.8) and an overrule is not appropriate.

The final near miss is the staff turnover score for *New Perspectives*. It is the cluster exception by scoring above threshold. Table 4.1 reveals that its scale score is 9, marginally above the median of 7 and the mean of 7.6. Given that 9 is the first score recorded above the threshold and the standard deviation is 4.3 and the maximum 17, it is decided to overrule this above threshold score and to make a below threshold score a prime implicant for the cluster.

Given the overruling process applied, cluster 1 now has four prime implicants:

continuing * capital expenditure * pgt * staffturnover

This is an example of how a qualitative logic can be applied to the quantitative data used in DPS.

For the other clusters:

Cluster 2 and 3 share above threshold debtors and above threshold expenditure on marketing, this is also associated with a below threshold continuing customers and managers.

Cluster 4 has above threshold continuing customers and below median debtors. They have low staff turnover.

Cluster 5 has much similarity across a range of variables and shares above threshold continuing customers and below median debtors.

Table 4.3 Truth table with debtors as the outcome variable, 2016

	Capexpend2016	AnIncomeGrow2016	PGT2016	Genderpaygap2016	Marketing2016	Managers2016	Overseas2016	Continuecustomers2016	Staffturnover2016	Sicknessdays2016	Cluster	Debtors2016
New Perspectives	0	1	0	1	0	1	1	0	1	1	1	0
Linear Logics	1	0	0	1	1	1	1	0	0	0	1	0
LKS Data	0	1	1	0	0	0	0	1	0	0	4	0
Open Thinking	1	0	1	0	1	1	1	1	0	1	4	0
Cosign Research	1	1	0	1	0	0	0	1	0	0	4	0
Strategy Statistics	0	1	1	0	0	1	0	1	1	0	5	0
JB Alpha	1	0	1	0	0	1	0	1	1	0	5	0
Sun Focus	0	0	1	0	0	0	1	0	0	0	1	1
Mini Max	0	1	0	0	0	0	0	0	0	1	1	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	0	2	1
Visual Research	1	0	0	1	1	0	1	0	0	1	2	1
System Synthesis	1	1	0	1	1	0	1	0	0	1	3	1

Table 4.3 shows a rearrangement of the order of the truth table in table 4.2 to consider debtors as the outcome variable.

The businesses with a below threshold number of debtors do not have any prime implicants and there are no near misses.

It is possible to do further analysis of this outcome, related to discrete clusters.

For cluster 4 the relationship is:

CONTINUING * staffturnover = debtors

For the pair in cluster 5 the relationship is:

PGT * genderpay * marketing * MANAGERS * overseas * CONTINUING * STAFFTURNOVER * sickness = debtors

The organisations with above threshold debtors have two prime implicants. These are below threshold on the ratio of managers and continuing customers. A near miss can be considered for staff turnover. The exception for staff turnover is *Ashton Algorithms*. Table 4.1 shows that it scores 9 where the median is 7 and the mean 7.6 (standard deviation = 4.3). It is decided not to overrule the binary score.

The summary Boolean explanation for businesses that have above threshold debtors is:

managers * continuing = DEBTORS

It is possible to do further analysis of the outcome, related to the pairing in cluster 2.

anincomegrowth * MARKETING * managers * continuing = DEBTORS

2017

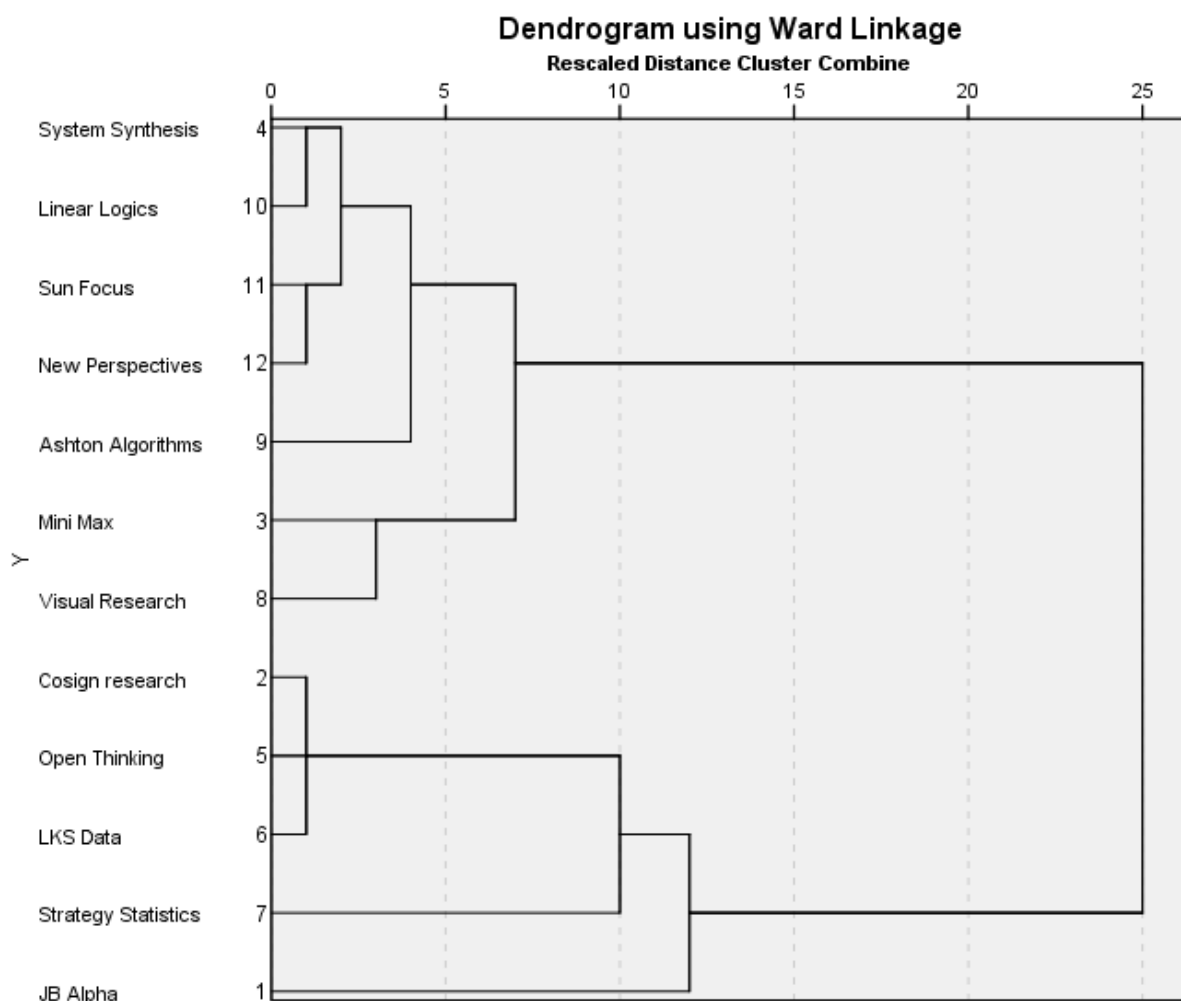


Figure 4.2 Dendrogram for 2017 data

The dendrogram for 2017 (figure 4.2) shows a structure that suggests three clusters with one outlier (*JB Alpha*). *Ashton Algorithms* is situated between clusters 1 and 2, indicating that its group location needs careful consideration at the next stage of analysis. *Strategy Statistics* looks like it is a possible outlier within cluster 3.

Table 4.4 Scale data converted to binary crisp set scores, 2017

	Capexpend2017	AnIncomeGrow2017	PGT2017	Genderpaygap2017	Marketing2017	Managers2017	Overseas2017	Continuecustomers2017	Debtors2017	Staffturnover2017	Sicknessdays2017
JB Alpha	15.1	4.3	65.0	1.0	5.3	0.08	0.0	94.0	0.5	16.0	6.0
Cosign Research	9.9	6.1	59.0	1.0	5.1	0.03	5.0	92.0	1.5	5.0	4.0
Mini Max	7.1	5.3	48.0	2.0	4.7	0.04	0.0	87.0	3.5	4.0	6.0
System Synthesis	7.1	5.0	43.0	7.0	6.2	0.03	12.0	85.0	4.0	8.0	6.0
Open Thinking	8.5	7.0	65.0	3.0	6.3	0.05	8.0	96.0	1.0	9.0	4.0
LKS Data	2.3	5.3	73.0	1.0	5.3	0.04	6.0	97.0	1.0	8.0	3.0
Strategy Statistics	6.3	14.0	84.0	1.0	5.2	0.05	4.0	90.0	5.0	13.0	6.0
Visual Research	12.5	4.0	55.0	6.0	4.3	0.04	6.0	83.0	4.0	8.0	9.0
Ashton Algorithms	8.2	2.3	61.0	4.0	8.2	0.03	6.0	89.0	4.5	7.0	6.0
Linear Logics	6.9	4.7	58.0	5.0	6.3	0.05	14.0	90.0	3.0	8.0	5.0
Sun Focus	6.4	5.6	61.0	4.0	5.9	0.04	10.0	79.0	5.0	7.0	7.0
New Perspectives	5.7	5.4	55.0	5.0	5.0	0.04	12.0	86.0	4.0	13.0	8.0
Mean	8.0	5.8	60.6	3.3	5.7	0.04	6.9	89.0	3.1	8.8	5.8
Median	7.1	5.3	60.0	3.5	5.3	0.04	6.0	89.5	3.8	8.0	6.0
Standard Deviation	3.2	2.7	10.3	2.1	1.0	0.01	4.3	5.1	1.6	3.3	1.6
JB Alpha	1	0	1	0	0	1	0	1	0	1	0
Cosign Research	1	1	0	0	0	0	0	1	0	0	0
Mini Max	0	0	0	0	0	0	0	0	0	0	0
System Synthesis	0	0	0	1	1	0	1	0	1	0	0
Open Thinking	1	1	1	0	1	1	1	1	0	1	0
LKS Data	0	0	1	0	0	0	0	1	0	0	0
Strategy Statistics	0	1	1	0	0	1	0	1	1	1	0
Visual Research	1	0	0	1	0	0	0	0	1	0	1
Ashton Algorithms	1	0	1	1	1	0	0	0	1	0	0
Linear Logics	0	0	0	1	1	1	1	1	0	0	0
Sun Focus	0	1	1	1	1	0	1	0	1	0	1
New Perspectives	0	1	0	1	0	0	1	0	1	1	1

Table 4.5 Truth table with four cluster memberships, 2017

	Capexpend2017	AnIncomeGrow2017	PGT2017	Genderpaygap2017	Marketing2017	Managers2017	Overseas2017	Continucustomers2017	Debtors2017	Staffturnover2017	Sicknessdays2017	Cluster
System Synthesis	0	0	0	1	1	0	1	0	1	0	0	1
Linear Logics	0	0	0	1	1	1	1	1	0	0	0	1
New Perspectives	0	1	0	1	0	0	1	0	1	1	1	1
Sun Focus	0	1	1	1	1	0	1	0	1	0	1	1
Ashton Algorithms	1	0	1	1	1	0	0	0	1	0	0	2
Mini Max	0	0	0	0	0	0	0	0	0	0	0	2
Visual Research	1	0	0	1	0	0	0	0	1	0	1	2
LKS Data	0	0	1	0	0	0	0	1	0	0	0	3
Cosign Research	1	1	0	0	0	0	0	1	0	0	0	3
Open Thinking	1	1	1	0	1	1	1	1	0	1	0	3
Strategy Statistics	0	1	1	0	0	1	0	1	1	1	0	3
JB Alpha	1	0	1	0	0	1	0	1	0	1	0	3

The truth table 4.5 shows the variable relationships with cluster memberships and definitions. *Ashton Algorithms* appears to fit better with cluster 2 in this context. If *Ashton Algorithms* is located to cluster 2, cluster 1 has three prime implicants and cluster two has five (the cluster two pairing of *Mini Max* and *Visual Research* share seven prime implicants). Cluster 3 shares three prime implicants, which it also shared with *JB Alpha*. For this reason, *JB Alpha* is allocated to cluster 3.

Boolean algebra can be used to summarise the variable influences on each cluster.

Cluster 1: capexpend * GENDERPAY * OVERSEAS

Cluster 2: anincomegrowth * managers * overseas * continuing * staffturnover

Cluster 3: genderpay * CONTINUING * sickness

Table 4.6 Truth table with debtors as the outcome variable, 2017

	Capexpend2017	AnIncomeGrow2017	PGT2017	Genderpaygap2017	Marketing2017	Managers2017	Overseas2017	Continuecustomers2017	Staffturnover2017	Sicknessdays2017	Cluster	Debtors2017
Linear Logics	0	0	0	1	1	1	1	1	0	0	1	0
Mini Max	0	0	0	0	0	0	0	0	0	0	2	0
LKS Data	0	0	1	0	0	0	0	1	0	0	3	0
Cosign Research	1	1	0	0	0	0	0	1	0	0	3	0
Open Thinking	1	1	1	0	1	1	1	1	1	0	3	0
JB Alpha	1	0	1	0	0	1	0	1	1	0	3	0
System Synthesis	0	0	0	1	1	0	1	0	0	0	1	1
New Perspectives	0	1	0	1	0	0	1	0	1	1	1	1
Sun Focus	0	1	1	1	1	0	1	0	0	1	1	1
Ashton Algorithms	1	0	1	1	1	0	0	0	0	0	2	1
Visual Research	1	0	0	1	0	0	0	0	0	1	2	1
Strategy Statistics	0	1	1	0	0	1	0	1	1	0	3	1

Table 4.6 shows a rearrangement of the truth table to make the measure of debtors the outcome variable. The prime implicant for below threshold level of debtors is a low level of staff sickness. Having an above threshold number of continuing customers is a near miss with only one business, *Mini Max*, being below threshold (87 – compared to the median of 89.5). This is overruled to make the variable a prime implicant. This can be summarised as:

CONTINUING * sickness = debtors

For those organisations scoring above threshold with the debtors score, a below threshold score for continuing customers can be considered a prime implicant. This is because the near miss for that variable, *Strategy Statistics*, is close to the median with a score of 90 (compared to the median of 89.5, mean = 90, standard deviation = 5.1).

The other near miss is the ratio of managers. The outlier is *Strategy Statistics* and it scores 0.05. The median and mean are 0.04 and the maximum score 0.08. The variable is not upgraded to a prime implicant.

There is no uniform association between cluster membership and the debtor outcome.

Chapter five: Concluding the DPS

Variable changes

When concluding the results of DPS it is important to conclude for the pattern changes of variable scores, case similarities and differences, and the interactions between variable changes with case pattern changes.

A good starting point is to plot the mean variable averages for each year and to conclude on the trend characteristic for each variable. Table 5.1 shows the mean variable changes and the final row indicates a qualitative conclusion about the trend for that variable.

Where overall change over the three years is small, it is appropriate to conclude that the variable trend change is stable, although this may also depend on the range of scale measurement used by that variable.

Noticeable variable trends from the example data set are a decline in annual income growth (although some evidence of a recovery in 2017), an increasing percentage of employees with postgraduate qualifications, an increase in continuing customers, and a decline in staff turnover.

Table 5.1 Variable mean changes: 2015-17

	Capexpend2015	AnIncomeGrow2015	PercentWFwithPGT2015	Genderpaygap2015	Marketing2015	Managers2015	Overseas2015	continuecustomers2015	debtors2015	staffturnover2015	sicknesdays2015
2015	7.1	7.9	55.6	3.0	5.0	0.04	6.8	84.6	2.5	13.5	5.6
2016	7.8	3.4	58.7	2.5	5.8	0.04	6.8	86.8	2.4	7.6	4.8
2017	8.0	5.8	60.6	3.3	5.7	0.04	6.9	89.0	3.1	8.8	5.8
	stable	∨	^	stable	stable	stable	stable	^	stable	∨	stable

Case and cluster changes

The best method for considering cluster stability and case changes between clusters is to compare the dendrograms for each year.

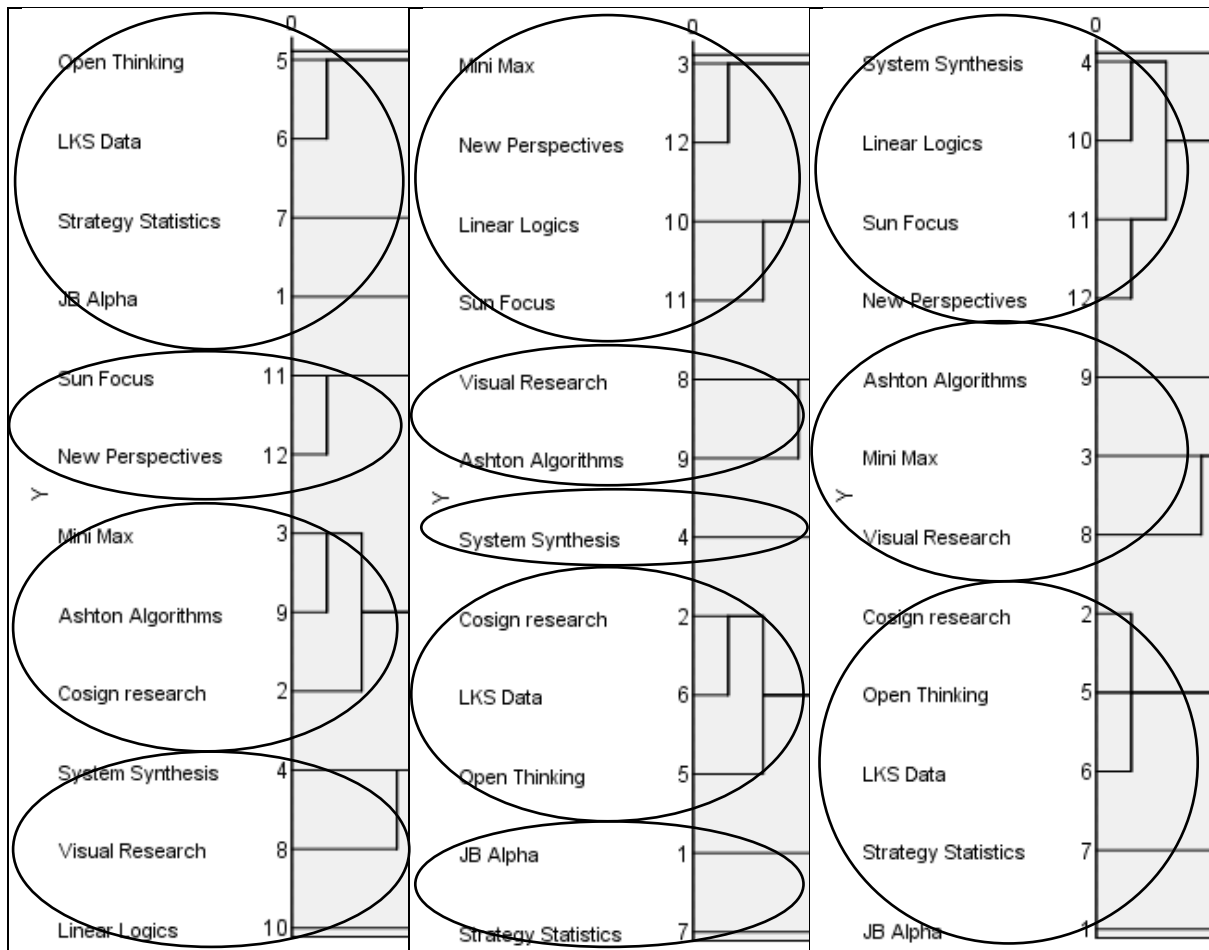


Figure 5.1 Dendrogram comparison: 2015, 2016, 2017

Figure 5.1 shows that:

New Perspectives and *Sun Focus* remain similar and linked in the same cluster across the three-year period.

JB Alpha and *Strategy Statistics* are also both linked as similar and are situated in the same cluster across the three-year period.

There is more consistency in the relationships between cases over time, after 2016.

Further relationships of similarity between 2016 and 2017 include:

Cosign Research, LKS data and Open Thinking

Visual Research and Ashton Algorithms

Finally, another method of analysis for considering the overall characteristics over time of the DPS is the longitudinal truth table. This device uses an excel formula to identify case and variable patterns that are consistently above or below threshold for the entire three-year period. It is therefore a method for identifying stable patterns over the time period.

Table 5.2 shows the excel process and method for computing the longitudinal truth table where the formula used was:

```
=IF(OR(AND(B2=1,B15=1,B28=1)), "ABOVE", IF(AND(B2=0,B15=0,B28=0), "BELOW", " "))
```

In the example above, the formula is for the first comparison, and the cells being compared are B2, B15 and B26. That is B2 is the data for 2015, B15 is the data for 2016, and B28 is the data for 2017.

The result of the Excel formula computations is in the area at the bottom of table 5.2 sub headed: 'OVERALL threshold stability'. Cells scoring ABOVE, are where there is a consistent above threshold score between the case and variable for all three years, and those scoring BELOW, are where there is a consistent below threshold score between the case and variable for all three years. If the cell is empty, with no text present, this indicates there is no consistent score over time for that case and variable relationship.

Table 5.3 shows a re-sorting of the overall threshold stability data, in order to find the consistent case pattern similarities, as defined by variables, over time. The sort is set up to prioritise the order of cases according to the amount of consistent longitudinal scores in each variable. For example, the first variable in the sort, is set as the proportion of employees who have a postgraduate qualification, as this variable is consistently either above or below threshold for all cases across all three years (there are no empty cells in the column).

The results of this re-sort are a structure of similarity and difference that is not directly related to similarities over time indicated in the earlier comparison of years indicated in figure 5.1. Together this evidence suggests a degree of instability in case-based similarity and difference, with it being relatively likely that cases will change over time in relationship to each other, despite more relative stability in the underlying variables that define the characteristics of these cases. This is not particularly surprising when examining meso business cases in a dynamic market environment. Macro studies with DPS tend to illustrate much more case-based stability over time, with countries being relatively less likely to change their similarity and difference to each other over time, despite any variable instability (see Haynes, 2017).

Finally, we can draw conclusions based on any particular variable outcome of interest to the researcher. In table 5.4, the table is re-sorted to show the patterns more clearly for the outcome variable of debtors. This confirms the benefits that continuing customers make to reduce the ratio of debtors.

Table 5.2 Setup for calculating a longitudinal truth table

	Capexpend	AnIncomeGrow	PGT	Genderpaygap	Marketing	Managers	Overseas	Continuecustomers	Debtors	Staffturnover	Sicknessdays
2015											
JB Alpha	1	0	1	0	1	1	0	1	0	1	1
Cosign Research	1	0	0	1	0	0	1	1	0	1	0
Mini Max	0	0	0	1	1	0	0	1	1	1	1
System Synthesis	1	1	0	1	1	0	1	0	1	0	1
Open Thinking	1	1	1	0	0	1	1	1	0	1	0
LKS Data	0	1	1	0	0	1	0	1	0	0	0
Strategy Statistics	0	0	1	0	0	1	0	1	0	1	1
Visual Research	1	1	0	1	1	1	1	1	1	0	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	1
Linear Logics	1	0	0	1	1	1	1	0	1	0	0
Sun Focus	0	1	1	0	0	1	0	0	1	0	0
New Perspectives	0	1	0	1	0	1	1	0	1	0	1
2016											
JB Alpha	1	0	1	0	0	1	0	1	0	1	0
Cosign Research	1	1	0	1	0	0	0	1	0	0	0
Mini Max	0	1	0	0	0	0	0	0	1	0	1
System Synthesis	1	1	0	1	1	0	1	0	1	0	1
Open Thinking	1	0	1	0	1	1	1	1	0	0	1
LKS Data	0	1	1	0	0	0	0	1	0	0	0
Strategy Statistics	0	1	1	0	0	1	0	1	0	1	0
Visual Research	1	0	0	1	1	0	1	0	1	0	1
Ashton Algorithms	0	0	1	0	1	0	0	0	1	1	0
Linear Logics	1	0	0	1	1	1	1	0	0	0	0
Sun Focus	0	0	1	0	0	0	1	0	1	0	0
New Perspectives	0	1	0	1	0	1	1	0	0	1	1
2017											
JB Alpha	1	0	1	0	0	1	0	1	0	1	0
Cosign Research	1	1	0	0	0	0	0	1	0	0	0
Mini Max	0	0	0	0	0	0	0	0	0	0	0
System Synthesis	0	0	0	1	1	0	1	0	1	0	0
Open Thinking	1	1	1	0	1	1	1	1	0	1	0
LKS Data	0	0	1	0	0	0	0	1	0	0	0
Strategy Statistics	0	1	1	0	0	1	0	1	1	1	0
Visual Research	1	0	0	1	0	0	0	0	1	0	1
Ashton Algorithms	1	0	1	1	1	0	0	0	1	0	0
Linear Logics	0	0	0	1	1	1	1	1	0	0	0
Sun Focus	0	1	1	1	1	0	1	0	1	0	1
New Perspectives	0	1	0	1	0	0	1	0	1	1	1
OVERALL threshold stability											
JB Alpha	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW	ABOVE	BELOW	ABOVE	
Cosign Research	ABOVE		BELOW		BELOW	BELOW		ABOVE	BELOW		BELOW
Mini Max	BELOW		BELOW			BELOW	BELOW				
System Synthesis			BELOW	ABOVE	ABOVE	BELOW	ABOVE	BELOW	ABOVE	BELOW	
Open Thinking	ABOVE		ABOVE	BELOW		ABOVE	ABOVE	ABOVE	BELOW		
LKS Data	BELOW		ABOVE	BELOW	BELOW		BELOW	ABOVE	BELOW	BELOW	BELOW
Strategy Statistics	BELOW		ABOVE	BELOW	BELOW	ABOVE	BELOW	ABOVE		ABOVE	
Visual Research	ABOVE		BELOW	ABOVE					ABOVE	BELOW	ABOVE
Ashton Algorithms		BELOW	ABOVE		ABOVE	BELOW	BELOW	BELOW	ABOVE		
Linear Logics		BELOW	BELOW	ABOVE	ABOVE	ABOVE	ABOVE			BELOW	BELOW
Sun Focus	BELOW		ABOVE					BELOW	ABOVE	BELOW	
New Perspectives	BELOW	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW			ABOVE

Table 5.3 Longitudinal truth table, sorted by variable patterns

	Capexpend	AnIncomeGrow	PGT	Genderpaygap	Marketing	Managers	Overseas	Continuecustomers	Debtors	Staffturnover	Sicknessdays
Ashton Algorithms		BELOW	ABOVE		ABOVE	BELOW	BELOW	BELOW	ABOVE		
Open Thinking	ABOVE		ABOVE	BELOW		ABOVE	ABOVE	ABOVE	BELOW		
JB Alpha	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW	ABOVE	BELOW	ABOVE	
Sun Focus	BELOW		ABOVE					BELOW	ABOVE	BELOW	
LKS Data	BELOW		ABOVE	BELOW	BELOW		BELOW	ABOVE	BELOW	BELOW	BELOW
Strategy Statistics	BELOW		ABOVE	BELOW	BELOW	ABOVE	BELOW	ABOVE		ABOVE	
Linear Logics		BELOW	BELOW	ABOVE	ABOVE	ABOVE	ABOVE			BELOW	BELOW
System Synthesis			BELOW	ABOVE	ABOVE	BELOW	ABOVE	BELOW	ABOVE	BELOW	
Visual Research	ABOVE		BELOW	ABOVE					ABOVE	BELOW	ABOVE
Cosign Research	ABOVE		BELOW		BELOW	BELOW		ABOVE	BELOW		BELOW
New Perspectives	BELOW	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW			ABOVE
Mini Max	BELOW		BELOW			BELOW	BELOW				

Table 5.4 Longitudinal truth table with Debtors outcome, 2015-17

	Capexpend	AnIncomeGrow	PGT	Genderpaygap	Marketing	Managers	Overseas	Continuecustomers	Staffturnover	Sicknessdays	Debtors
Linear Logics		BELOW	BELOW	ABOVE	ABOVE	ABOVE	ABOVE		BELOW	BELOW	
Strategy Statistics	BELOW		ABOVE	BELOW	BELOW	ABOVE	BELOW	ABOVE	ABOVE		
Mini Max	BELOW		BELOW			BELOW	BELOW				
New Perspectives	BELOW	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW		ABOVE	
System Synthesis			BELOW	ABOVE	ABOVE	BELOW	ABOVE	BELOW	BELOW		ABOVE
Ashton Algorithms		BELOW	ABOVE		ABOVE	BELOW	BELOW	BELOW			ABOVE
Visual Research	ABOVE		BELOW	ABOVE					BELOW	ABOVE	ABOVE
Sun Focus	BELOW		ABOVE					BELOW	BELOW		ABOVE
Open Thinking	ABOVE		ABOVE	BELOW		ABOVE	ABOVE	ABOVE			BELOW
Cosign Research	ABOVE		BELOW		BELOW	BELOW		ABOVE		BELOW	BELOW
JB Alpha	ABOVE	BELOW	ABOVE	BELOW		ABOVE	BELOW	ABOVE	ABOVE		BELOW
LKS Data	BELOW		ABOVE	BELOW	BELOW		BELOW	ABOVE	BELOW	BELOW	BELOW

The qualitative DPS summary

The outcome focus in this book has been on the percentage of debtors as defined by percentage of customers with late payment invoices over one year. Nevertheless, DPS allows for the researcher to re-sort the pattern matrix in both truth tables and longitudinal truth tables to consider any available variable as an outcome that might be affected by other variables.

The longitudinal truth table in table 5.4 shows four businesses that consistently have above average debtors and four others that have consistently below average debtors.

Three of the four businesses that have consistent challenges over time with the level of debtors also have consistently scored below average for retaining continuing customers. Two of these cases, *System Synthesis* and *Ashton Algorithms*, also have an ongoing pattern over time with above average expenditure on marketing (perhaps indicating a pressure to secure more new and reliable business) and a below average ratio of managers (suggesting this at least needs exploring in terms of manager's ability to prioritise staff to chase and secure income owed).

The four businesses that have consistently avoided having higher numbers of debtors are also above average in consistently retaining customers. Three of these four are consistent in higher than average capital expenditure. Three cases are also consistent over time in employing a higher proportion of staff with postgraduate qualifications and having a lower gender pay gap. These features might be part of a qualitative explanation about why these businesses can avoid delay in securing payments.

Conclusion

Dynamic Pattern Synthesis is influenced much by the ontology of qualitative approaches but uses quantitative data to explore case heterogeneity. Rather than building a model based on aggregate or average scores, where real cases are considered according to their closeness to typical or 'ideal' models of cases, DPS demonstrates dynamic differences between cases and the limits to their degrees of similarity. Similarity is better balanced with considerations of difference also. This leads to a better sense of judgement about associations and causality, where such relationships are likely to be contextual, especially with regard to their persistence over time and place.

References

- Boulton, J., Allen, P., & Bowman, C. (2015). *Embracing Complexity: Strategic Perspectives for an Age of Turbulence*. Oxford: Oxford University Press.
- Everitt, B. (1993). *Cluster Analysis*. London: Arnold.
- Haynes, P. (2012). *Public policy beyond the financial crisis*. Oxon: Routledge.
- Haynes, P. (2014). Combining the Strengths of qualitative comparative analysis and cluster analysis. *International Journal of Public Policy*, 37(9), 581-590.
- Haynes, P. (2015). *Managing Complexity in the Public Services* (2nd ed.). Oxford: Routledge.
- Haynes, P. (2015). The European Single Currency Project and the Concept of Convergence for European Welfare States - The Ideal and the Reality. *Social Policy and Administration*, 44(4), 466-489.
- Haynes, P. (2017). *Social Synthesis: understanding patterns in complex social systems*. Oxford: Routledge.
- Haynes, P., & Haynes, J. (2016). Convergence and Heterogeneity in Euro Based Economies: Stability and Dynamics. *Economies*, 1-16.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- Ragin, C. (1987). *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.
- Rihoux, B., & Ragin, C. (2009). *Configurational Comparative Methods: Qualitative Comparative Methods (QCA) and Related Techniques*. London: Sage.

The datasets

Access to the data files is via the supporting website at: <http://blogs.brighton.ac.uk/dpsmethod/>

It is possible to download the following files, for use in training and education:

DPS how to do it book database (dataset of fictional organisations)	IBM SPSS
DPS 2015 Data	Microsoft Excel
DPS 2016 Data	Microsoft Excel
DPS 2017 Data	Microsoft Excel

Social Media

@profpdh	Philip Haynes
@cecan	ESRC CECAN (Centre for the Evaluation of Complexity Across the Nexus)

Related hashtags

#ResearchResources

#researchmethods

#DPS

Acknowledgements

My thanks to the following organisations and individuals who have helped to support the development and dissemination of Dynamic Pattern Synthesis.

ESRC CECAN (Centre for the Evaluation of Complexity Across the Nexus)

ESRC Complexity and Methods in the Social Sciences Seminar Programme

Brian Castellani

David Byrne

Emma Uprichard

Jonathan Haynes

University of Brighton Research Office and Impact Awards

School of Applied Social Science, University of Brighton

Proof reading provided by Jane Haynes

About the Author

Philip Haynes is Professor of Public Policy at the University of Brighton where he teaches applied statistics, research methods, social and public policy, and public management. He is a Research Associate with the ESRC CECAN programme. He has published extensively in related areas including *Social Synthesis; Finding Dynamic Patterns in Complex Social Systems* (2017), *Managing Complexity in Public Policy*, 2nd Edition, (2015), and *Public Policy Beyond the Financial Crisis* (2012). He undertakes research and consultancy using applied statistics to explore policy and management related problems.



White Horse Books

whb.co.uk

2018

ISBN: 978-0-9534065-2-4



978-0-9534065-2-4