

Interpreting Geographical Data

Social Analysis

Part II

with

Microsoft

Excel



Purpose

This document has been prepared to support the development of skills in the use of Excel for social research. It provides a series of exercises to develop your skills in using Excel for social analysis Part II introduces a larger dataset with information about the 326 local authority districts in England. It provides ways of navigating and manipulating a larger dataset and exploring it to report key findings.

Author(/s) and funding

This document has been produced by Dr Nigel de Noronha. Material included in this guidance comes from work funded by the Nuffield Foundation as part of the Q-Step funding to the University of Warwick and initial work by the author with colleagues Richard Conibere and Rebecca Bromley to develop courses for Q-Step students at the University of Manchester.

Contents

Purpose	2
Author(/s) and funding	2
Using Excel for Social Analysis	4
Working with larger datasets.....	4
Further resources.....	14
Answers to questions in this section	15
Appendix 1 - Dataset.....	17
Local Authority.....	17
Social Class	17
Ethnicity	18
Highest Level of Qualification	19

Tables and Figures

Figure 1 – highest level of qualification achieved by local authority	10
Table 1 – local authority variables	17
Table 2 – social class variables.....	18
Table 3 – ethnic group population.....	18
Table 4 – highest level of qualification categories	19

Using Excel for Social Analysis

The exercises in this part of the toolkit are designed to help you develop a number of skills in Excel. At the end of this section you will be able to ...

- ... load existing data and carry out a range of analysis of individual variables to answer a range of research questions
- ... work with complex datasets, preparing worksheets and variables for a series of analyses
- ... present your findings in professional looking tables and graphs

These exercises begin on the assumption that you are familiar with opening Excel documents, and are able to work with simple Excel formulae. Please refer to the introductory exercises if you are unsure of any of these.

Working with larger datasets

To begin open the workbook:

- Download the file ‘LA.xlsx’ from and open it

Here is a screenshot of the first worksheet in the file:

LA	Region	Urban-rural	Population	Over16	noqual	l1qual	l2qual	appqual	l3qual	l4qual	othqual	prnoqual
1 Adur	South East	Largely urban	61182	50579	12936	8063	8324	2207	5655	11122	2272	26%
2 Allerdale	North West	Hamlet or dispersed	96422	80155	21636	10960	12969	4020	9236	18254	3080	27%
3 Amber Valley	East Midlands	Town and fringe	122309	100841	27220	13986	15627	4683	12086	23346	3893	27%
5 Arun	South East	Largely urban	149518	126164	31364	17871	21159	5019	14610	28732	7409	25%
6 Ashfield	East Midlands	Other urban	119497	96698	30345	15846	15901	4174	11626	14595	4211	31%
7 Ashford	South East	Town and fringe	117956	93411	19777	14578	16639	3602	11229	23007	4579	21%
8 Aylesbury Vale	South East	Village	174137	138567	24059	19099	22857	5014	16894	44291	6353	17%
9 Babergh	East	Hamlet or dispersed	87740	71821	16367	10385	12560	2818	8337	18539	2815	23%
10 Barking and Dagenham	London	Mainly urban	185911	137613	38396	20904	20617	2888	12759	28715	13334	28%
11 Barnet	London	Mainly urban	356386	282152	43833	27240	34028	3778	28742	113815	30716	16%
12 Barnsley	Yorkshire and Humberside	Other urban	231221	188335	60890	26913	30550	7716	21342	32724	8200	32%
13 Barrow-in-Furness	North West	Other urban	69087	56796	14422	7710	8912	5229	7791	10785	1947	25%
14 Basildon	East	Other urban	174497	139198	37084	25629	24568	4986	14961	25939	6031	27%
15 Basingstoke and Deane	South East	Town and fringe	167799	133947	23226	19647	21918	5152	16902	40793	6309	17%
16 Bassetlaw	East Midlands	Village	112863	92702	26393	13513	15245	3847	10453	18761	4490	28%
17 Bath and North East Somerset	South West	Town and fringe	176016	146439	25140	16175	20828	5472	24095	48910	5819	17%
18 Bedford	East	Town and fringe	157479	125871	25645	17313	19129	4939	14882	35417	8546	20%
19 Bexley	London	Mainly urban	231997	184431	42366	31049	32694	6951	21334	40146	9891	23%
20 Birmingham	West Midlands	Mainly urban	1073045	828363	233835	110158	113967	17654	107913	190335	54501	28%
21 Blaby	East Midlands	Largely urban	93915	76334	16964	11097	12257	4360	9378	18988	3290	22%
22 Blackburn with Darwen	North West	Other urban	147489	113122	32485	16211	17221	4428	12923	22416	7438	29%
23 Blackpool	North West	Largely urban	142065	116699	36559	18129	19993	4886	13546	18098	5488	31%
24 Bolsover	East Midlands	Town and fringe	75866	62192	20455	9520	10195	2308	7064	9853	2797	33%
25 Bolton	North West	Mainly urban	276786	219710	57888	31152	34731	9396	26408	48688	11447	26%
26 Boston	East Midlands	Town and fringe	64637	53275	17421	7705	8126	1745	4954	8060	5264	33%
27 Bournemouth	South West	Largely urban	183491	155224	29720	20105	24086	5852	24412	40951	10098	19%
28 Bracknell Forest	South East	Largely urban	113205	89743	14639	13755	15191	3032	11389	27197	4540	16%
29 Bradford	Yorkshire and Humberside	Mainly urban	522452	400121	110892	56413	59350	15131	45101	87397	25837	28%

Using Excel for Social Analysis (Part II)

In Part 1, the data were small enough to fit on the screen. For the rest of the section the dataset is much larger, covering the 326 Local Authorities (LAs) in England. Also the variables provided for each LA are more comprehensive, with the rows often extending beyond the edge of the screen. To help you navigate these larger data sets here are some useful keyboard short-cuts:

- <Home> - go to start of row
- <End> - go to end of row.
- <Ctrl-Home> - to return to cell A1, the top left cell of the worksheet.
- <Ctrl-End> - to go to the end of the last row, the bottom right cell of the worksheet.
- <Ctrl-Up> - go to top of current column
- <Ctrl-Down> - go to bottom of current column.
- <Ctrl-Page Down> - move to next worksheet.
- <Ctrl-Page Up> - move to previous worksheet

Note that the keyboard short-cuts can also be used to select data by simultaneously holding down the <Shift> key. A common task is to select all the data in the sheet e.g. in preparation for sorting the data or to specify the data needed for a graph. Using the keyboard you can do this by:

1. <Ctrl-Home> to move to cell A1.
2. <Ctrl-Shift-End> to select all data from A1 to the end of the worksheet.

Or with:

- <Ctrl-A> to select All.

Exploring the data

Now that you have opened the <LA.xlsx> spreadsheet, have a look at the different worksheets. Further detailed information on the data is provided in Appendix 1.

<**Qualifications**> contains a record for each local authority in England. They are in alphabetical order. The region and urban-rural classification of the local authority is included as well as the population and the population aged 16 or over. The counts of highest qualification are for people aged 16 or over.

<**Ethnic diversity**> contains a record for each local authority in England in alphabetical order. The region and urban-rural classification of the local authority is included as well as the population. Population counts for each of the ethnic groups are included and the proportion of the total population is calculated in the next set of columns (*they start with the prefix 'pr'*).

<**Social class**> contains two records for each local authority in England. One contains the information for males and the other for females. The region and urban-rural classification of the local authority is included as well as the population, the working age population (16-74) and the population aged 16 or over. The counts of occupation based social class are for males or females of working age.

<**Ethnicity Table**> contains a pivot chart collating populations by ethnic group using data from the <Ethnic diversity> worksheet. The worksheet is only partially complete, ready for your input in the fourth exercise.

<**lookup**> has tables to translate the column headings in the worksheets into more meaningful descriptions to improve the presentation of information.

Exploring the Highest Level of Qualification

The worksheet <Qualifications> has a count of the population with each different level of qualification. The population of different local authorities varies significantly across England. The Isles of Scilly has a population of around two thousand whilst Birmingham is over one million. To enable comparison between districts we will calculate the proportion of the eligible population with each level of qualification.

- Type in new column headings for each of the qualification levels in the columns to the right of the data. You could name them 'pr' and the current name e.g. in the first column type `prnoqual`, in the second type `prl1qual` ...

We now need to calculate the proportion of the eligible population with each level of qualification. To help us do this quickly for all the local authorities we will use the \$ operator in the formula as shown in the box.

Fixing an absolute row and or column in a formula

We often use a formula to enable us to copy the same action e.g. create a new value. By default the column and row changes as the formula is copied. This can be fixed by entering the \$ sign before the column letter and / or the row number.

Examples

Entering $\$B\2 in a formula will mean that when it is copied elsewhere it will refer to B2

Entering $\$B2$ in a formula will mean that as the cell is copied the row number will change but it will always refer to column B

- In the cell below *prnoqual* type in $=F2/\$E2$. This means that as we copy this formula it will always refer to the column *population aged 16 or over*. This is the population group that highest level of qualification is based on. As the formula is copied down it will change to reflect the *population aged 16 or over* for the relevant local authority.
- <Copy> the cell
- Highlight the area you want to copy to (*no qualifications on the first row to other qualifications on the last row*)
- <Paste> the cell
- Click on <Home> on the top menu and click the <%> button near the middle of the ribbon to format results as percentages.

You should now have added seven new fields to each local authority showing the percentage of the eligible population with each level of qualification. However making sense of all this information is challenging. Excel offers some useful tools to help you do this.

Freezing Columns and Rows

You can freeze the heading and row titles and then scroll around in the data.

- Move the cursor to cell B2
- Select <View> on the top menu

Using Excel for Social Analysis (Part II)

- Select <Split> on the ribbon

A grey frame now appears around the data with the heading row and local authority names outside it. When you move the cursor to the right the column heading and cells within the grey box will move. When you move the cursor down the column heading remains and the local authority name and cells move down. This enables you to focus on the information you are interested in. Scroll across and down so that you can see the pattern of qualifications for Sunderland and Surrey Heath. Michael Gove, former Minister of State for Education in the last coalition government, is MP for Surrey Heath.

What does the pattern of qualifications suggest is different about Sunderland and Surrey Heath?

You can turn the split screen off by selecting <Split> again. It operates like an on/off switch. Keep it turned on for the rest of this exercise.

Sorting the Information

We can order the information by any of the columns to help us make more sense of it.

- Select <Data> from the top menu
- Select <Sort> from the ribbon.

This opens a dialogue box. Let's take a look at the areas where more of the population have no qualifications.

- First check that the data has headers option is ticked (top right of the dialog box)

My data has headers

- Select *prnoqual* to sort by, values to Sort on and Largest to Smallest to Order by

Have a look at the ordered list of local authorities. What do you know about the places near the top and bottom of the list?

It is worth noting that over time we have increasingly seen higher level qualifications. This means that many older people who have had equivalent jobs / incomes and wealth are likely to be less well qualified than their younger equivalents.

Repeat the exercise for those with the highest qualifications (pr14qual) and ordered by the smallest to largest. What do you know about the places near the top and bottom of the list?

Graphing the Information

Information ordered in this way can be easily presented to help tell the story of how a particular local authority compares to the rest of the country.

- Click on the column letter above the local authority, hold the <Ctrl> key and click on the column letter for the percentage with the highest level of qualification
- Select <Insert> on the top menu
- Select <Columns> from the ribbon and a set of options are displayed, select the top left image under the 2D-Column heading.

An ascending column chart is now displayed. You may wish to change the colour of this. Next we want to highlight the local authority that we are going to focus on. The easiest way to do this is to enlarge the graph until the name of the authority you want can be seen at the bottom. Find Manchester (the proportion is 29%).

- Select the individual column for Manchester (the first time you click will select all the columns, the second the individual one). *This can be tricky and you will need to dramatically increase the size the graph to find and select it.*
- On the top menu select <Format>
- Select <Shape Fill> and choose a colour that stands out against the other columns
- Right click on the Manchester column again to show a menu
- Select <Add Data Label> and 29% is displayed above the column
- Right click on the Manchester column to show the menu
- Select <Format Data Label> and change the Label Options to Category Name rather than Value

You can now reduce the size of the graph and do some tidying up. Get rid of the labels on the bottom axis, legend and type in a better title.

- Click on the titles at the bottom of the graph and press the <Delete> key
- Click on the legend (to the right) and press the <Delete> key
- Click on the title and replace the current label with something like 'Percentage Qualified to Degree Level or Higher'.
- Select the data series, then right click and choose <Format Data Series>. In the resulting dialog box reduce the gap between the bars to zero.

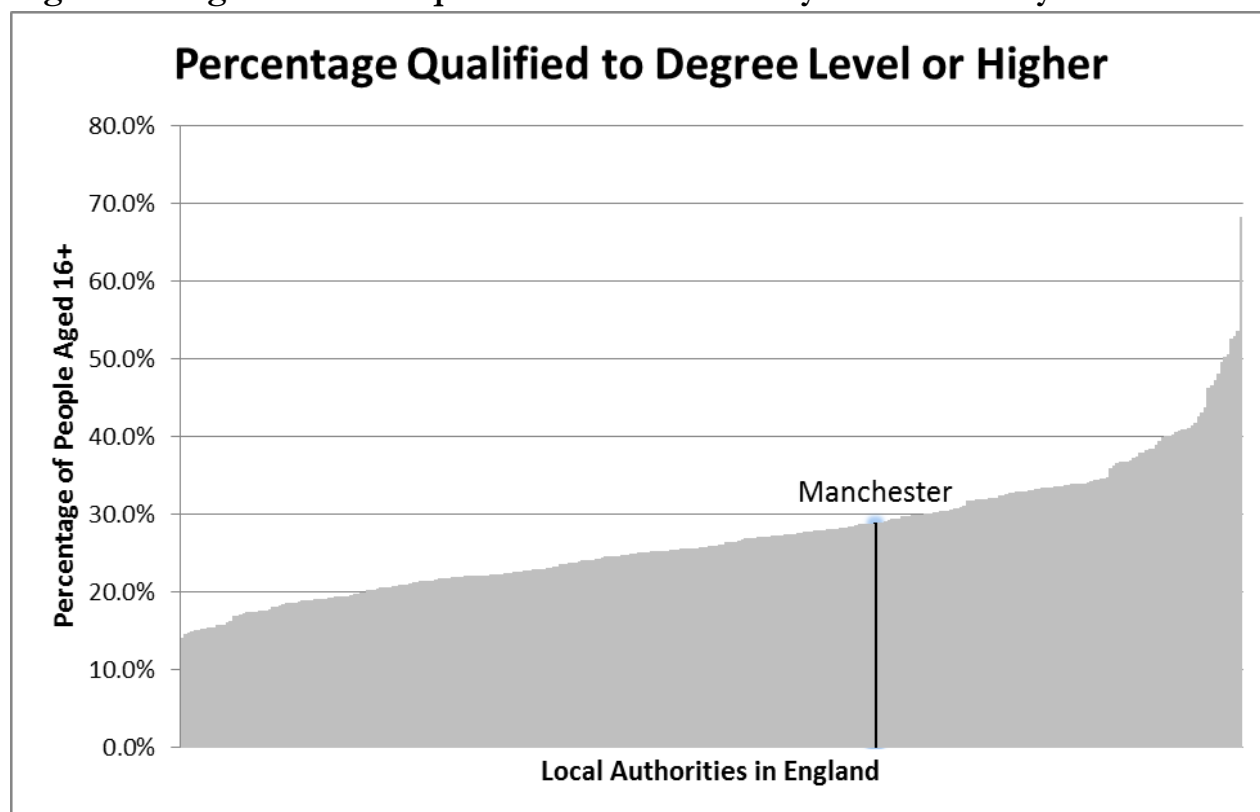
You can add axis titles in the layout menu.

- Select <Layout> on the top menu

- Add a vertical axis title with <Rotated Title> which says something like *'Percentage of People Aged 16+'*
- Add a horizontal axis title which says something like *'Local Authorities in England'*

The graph is now complete and can be copied and pasted into a Word document as part of a report. When doing so you should provide a title and cite the source e.g.

Figure 1 – highest level of qualification achieved by local authority



Source: Census 2011 standard table KS501EW

Using Filter to explore the data

Data filter allow you to adopt a more exploratory approach to the information.

- Select <Data> on the top menu
- Click on <Filter>. Arrows are displayed at the head of each column. This is an on/off switch

The arrow next to each field provides access to a list of the different values that this variable holds. You can select individual values, number calculations such as greater than 10% or the top 10. When using filters the effect is cumulative, the second and subsequent ones are only applied to the remaining rows displayed. Unless this is what you want you will need to turn a particular filter off before using another one. This can be done by initially choosing <Select all> in the drop down menu before applying the filter.

Handy tip: When the filter tool is applied <copy> will only act on the displayed information. It can be used to make a copy of the information that you have selected through the use of the filter tool.

Use the filter tool to answer the following questions

To work out which variable you need to use to answer each question, see the <layout> worksheet. The full description for all the education level variables is given at the bottom of the sheet.

In which local authorities do more than 50% of the population (aged 16+) have a degree level qualification or higher?

Which local authority has the highest percentage of apprentices?

Which five authorities have the highest proportion of people with no qualifications?

There are eighteen local authorities where more than 10% of people have other qualifications. Fifteen are in London which has a much higher proportion of people who were educated outside the UK because of migration. The others are Luton, Slough and Forest Heath. Look for information about one of them to help explain why they have such high rates of other qualifications.

Answers:

Exploring Ethnic Diversity

Freezing columns and rows helps us to make more sense of the information about particular local authorities. Sorting can help to identify, highlight and report on the way an individual piece of information varies. Filtering provides a more exploratory approach to the information. Here are a few questions for you to use any combination of these techniques to explore ethnic diversity across England. For this worksheet, the percentages of the different groups have already been calculated for you, using the same method as in the last exercise.

In what types of local authorities are white British less than 50% of the population? What are the relative proportions of other ethnic groups? What other information might be useful to help you explore this further?

Can you identify different stories about the places where the white other group is at least 10% of the population? *(there are at least four stories I can think of)*

Look at the top ten areas for other ethnic groups. How do they differ for ...

... Indians?

... Pakistanis?

... Bangladeshis?

What might these figures suggest?

Further resources

Graphs

Many (many) online resources exist showing different techniques to use when presenting data. For some general creative ideas, see

<https://sites.google.com/site/e90e50charts/> (not all are useful but interesting nonetheless!).

In exercise 1, the process of repeatedly enlarging a chart to find the one datapoint you need could become onerous if, for example, several of the bars needed to be found and highlighted. So, for an alternative approach see this tutorial:

<http://www.k2e.com/tech-update/tips/341-creating-an-excel-chart-with-conditional-formatting>

Recommended book for graph design and data presentation: Show Me the Numbers, Stephen Few, 2nd edition 2012, Atlantic Press.

Summary

This section has involved working with data and using a range of functions within Excel to present tables and charts. The remainder of the toolkit covers getting data from external sources, more advanced graphical options, creating an index and producing area profiles.

Answers to questions in this section

Exploring qualifications

Local authorities where more than 50% of people have a degree level qualification or above

Camden, City of London, Kensington and Chelsea, Richmond upon Thames, Wandsworth, Westminster

Local authority with the highest percentage of apprentices

Barrow-in-Furness

Five local authorities with highest percentage of people with no qualifications

Knowsley, Sandwell, Stoke-on-Trent, Tendring, Walsall

About Luton, Slough and Forest Heath (with thanks to google)

Luton: <http://www.loveluton.org.uk/diversity/>

Slough: <http://www.theguardian.com/uk/2013/jan/13/slough-multi-cultural-integration>

Forest Heath: <http://www.lakenheath.af.mil/>

Exploring ethnic diversity

White other 10% or more of the population

London – a tale of two cities?

The levels of the white other ethnic group in London might suggest two or more stories. First it might reflect a global elite who work in and around financial services and the major corporations hosted in the city. Some of the districts that people live in suggest that alongside the elite there are others working in lower paid possibly more casual service roles such as care services, catering, domestic service, hospitality, retail and transport.

The boroughs identified are:

Barnet, Brent, Camden, City of London, Ealing, Enfield, Hackney, Hammersmith and Fulham, Haringey, Hounslow, Islington, Kensington and Chelsea, Lambeth, Lewisham, Merton, Newham, Richmond upon Thames, Southwark, Tower Hamlets, Waltham Forest, Wandsworth and Westminster

Forest Heath is home to a large United States air force base. The presence of a large number of Americans is likely to explain the higher proportion of the white other ethnic group.

Peterborough and Boston have both been associated with the migration of people from the A8 accession states to support the large scale agricultural production in this part of the country.

Oxford and Cambridge host major global universities. Their students and staff are recruited globally. This provides the most likely explanation for the large proportion of the white other ethnic minority group.

White British population less than 50% of total

Barking and Dagenham, Barnet, Brent, Camden, Croydon, Ealing, Enfield, Hackney, Hammersmith and Fulham, Haringey, Harrow, Hounslow, Islington, Kensington and Chelsea, Lambeth, Leicester, Lewisham, Luton, Merton, Newham, Redbridge, Slough, Southwark, Tower Hamlets, Waltham Forest, Westminster

Concentration of ethnic groups in local authorities

White other 39 *local authorities in order of proportion*

Haringey (4%), Barnet, Enfield, Westminster, Ealing, Wandsworth, Lambeth, Kensington and Chelsea, Brent, Camden, Hackney, Waltham Forest, Hammersmith and Fulham, Southwark, Newham, Islington, Tower Hamlets, Hounslow, Birmingham, Merton, Lewisham, Manchester, Croydon, Richmond upon Thames, Leeds, Bristol, Greenwich, Harrow, Brighton and Hove, Peterborough, Oxford, Cambridge, Redbridge, Southampton, Hillingdon, Bromley, Bradford, Nottingham, Kingston upon Thames

Indian 16 *local authorities in order of proportion*

Leicester (7.9%), Birmingham, Harrow, Brent, Ealing, Hounslow, Redbridge, Newham, Hillingdon, Wolverhampton, Sandwell, Barnet, Coventry, Croydon, Slough, Bolton

Pakistani 13 *local authorities in order of proportion*

Birmingham (13.2%), Bradford, Manchester, Kirklees, Redbridge, Newham, Luton, Waltham Forest, Slough, Oldham, Leeds, Rochdale, Sheffield

Bangladeshi 8 *local authorities in order of proportion*

Tower Hamlets (18.8%), Newham, Birmingham, Oldham, Redbridge, Luton, Camden, Bradford

Appendix 1 - Dataset

The dataset being used provides information on population, social class, ethnicity and highest level of qualification achieved for each local authority in England. The tables below describe how this data is organised, the main categories used and the population which different variables are drawn from.

Local Authority

Table 1 shows the variable names and, for the regions and urban rural classification, the possible values that are given to each local authority.

Table 1 – local authority variables

Variable name	Description
LA	Local authority
Region	English region
	<i>Possible values</i>
	East
	East Midlands
	London
	North East
	North West
	South East
	South West
	West Midlands
	Yorkshire and Humberside
Urban-rural	Urban-rural classification
	<i>Possible values</i>
	Major urban
	Largely urban
	Other urban
	Town and fringe
	Village
	Hamlet or dispersed

Social Class

The measure of social class is derived from the occupation of the respondent. It is therefore only applied to those identified as of working age which is 16-74 in the census data. Table 2 shows the nine categories used in the analysis together with a way of combining these to higher level categories if required.

Table 2 – social class variables

Variable	Description
nssec1	Higher managerial, administrative or professional
nssec2	Lower managerial, administrative or professional
nssec3	Intermediate
nssec4	Small employer and self-employed
nssec5	Lower supervisory and technical
nssec6	Semi-routine
nssec7	Routine
nssec8	Never worked and long-term unemployed
nssec9	Full time students

Ethnicity

The number of members of each ethnic group is provided as a separate population count for each local authority. The count reflects the number by ethnicity for the whole population. Table 7.3 provides these counts as well as the proportion of the population in the local authority who come from each ethnic group.

Table 3 – ethnic group population

Variable	Description
wb	White British
wi	White Irish
wgt	White Gypsy or Irish Traveller
wo	White Other
mix	Mixed
ind	Indian
pak	Pakistani
ban	Bangladeshi
chi	Chinese
ao	Asian other
bafr	Black African
bcar	Black Caribbean
bo	Black other
oth	other ethnic group

Additionally further columns use these variable names but prefixed by **pr**, giving the proportion/percentage of the ethnic group relative to the total population in that Local Authority.

Highest Level of Qualification

Details of the translation of the highest level of qualification to the categories are shown in table 4.

Table 4 – highest level of qualification categories

Variable	Description
Noqual	No qualifications
L1qual	1-4 O Levels/CSE/GCSEs (any grades), Entry Level, Foundation Diploma, NVQ level 1, Foundation GNVQ, Basic/Essential Skills
L2qual	5+ O Level (Passes)/CSEs (Grade 1)/GCSEs (Grades A*-C), School Certificate, 1 A Level / 2-3 AS Levels/VCEs, Intermediate/Higher Diploma, Intermediate Diploma, NVQ level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/General Diploma, RSA Diploma
Appqual	Apprenticeship
L3qual	2+ A Levels/VCEs, 4+ AS Levels, Higher School Certificate, Progression/Advanced Diploma, NVQ Level 3; Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma
L4qual	Degree (for example BA, BSc), Higher Degree (for example MA, PhD, PGCE), NVQ Level 4-5, HNC, HND, RSA Higher Diploma, BTEC Higher level, Foundation degree
Othqual	Vocational/Work-related Qualifications, Qualifications gained outside the UK (Not stated/level unknown)