



The University of Manchester

Faculty of Humanities

Microsoft Excel Advanced Training



**A step-change in
quantitative social
science skills**

Funded by the
Nuffield Foundation,
ESRC and HEFCE

The University of Manchester
is one of 15 universities taking
part in the Q-Step programme



Contents

Contents.....	2
Aims.....	3
About This Workbook	3
Getting Started.....	4
Practical exercises	5
(1) Exploring the Highest Level of Qualification	6
(2) Exploring Ethnic Diversity	11
(3) Social Class – Pivot Tables	12
(4) Exploring Ethnicity and Place	17
Further resources.....	21
Appendix 1 - Dataset.....	22
Local Authority.....	22
Social Class	22
Ethnicity	23
Highest Level of Qualification	24
Appendix 2 – Feedback / Answers.....	25
Exploring qualifications.....	25
Exploring ethnic diversity.....	25

Aims

At the end of the session / workbook you should have acquired advanced skills in using Excel. You will be able to ...

- ... work with complex datasets, preparing worksheets and variables for a series of analyses
- ... analyse the interactions between two variables and produce pivot tables
- ... develop more complex variables by constructing indices
- ... present your findings in professional looking tables and graphs

About This Workbook

This workbook is designed to enable you to work at your own pace through the exercises. It is designed to develop your learning from the Excel Basic workbook. If you do not have this course material and did not complete the course please ask the tutors for a copy to check if there are any areas you need to work through before starting this one. During the taught session tutors will be available to help you with issues as they arise. Instructions to follow are broken down into bullet points marked with arrows and choices are enclosed in <> markers e.g.

- Select the <File> option from the top menu
- Select <Open> from the drop down menu displayed
- ...

Where you are required to type in information it is shown in Courier font e.g. type
`=sum (...)`.

This workbook has been prepared by Richard Conibere and Nigel de Noronha for the Q-Step programme at the University of Manchester.

Getting Started

We will be using one workbook for all the exercises in the session. So to begin open the workbook:

- Select the <File> option from the top menu and then <Open> from the dropdown menu.
- Choose the file 'LA.xlsx' which is available in the 'C:\Work' folder.

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

In the Basic Training the data were small enough to fit on the screen. For this advanced training the data set 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.

Practical exercises

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.

(1) 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 `pr11qual` ...

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 `b2` 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

Entering `b$2` in a formula will mean that as the cell is copied the column letter will change but the row number will always refer to row 2

- 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
- 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.

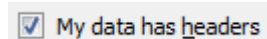
Sorting the Information

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)



- 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.

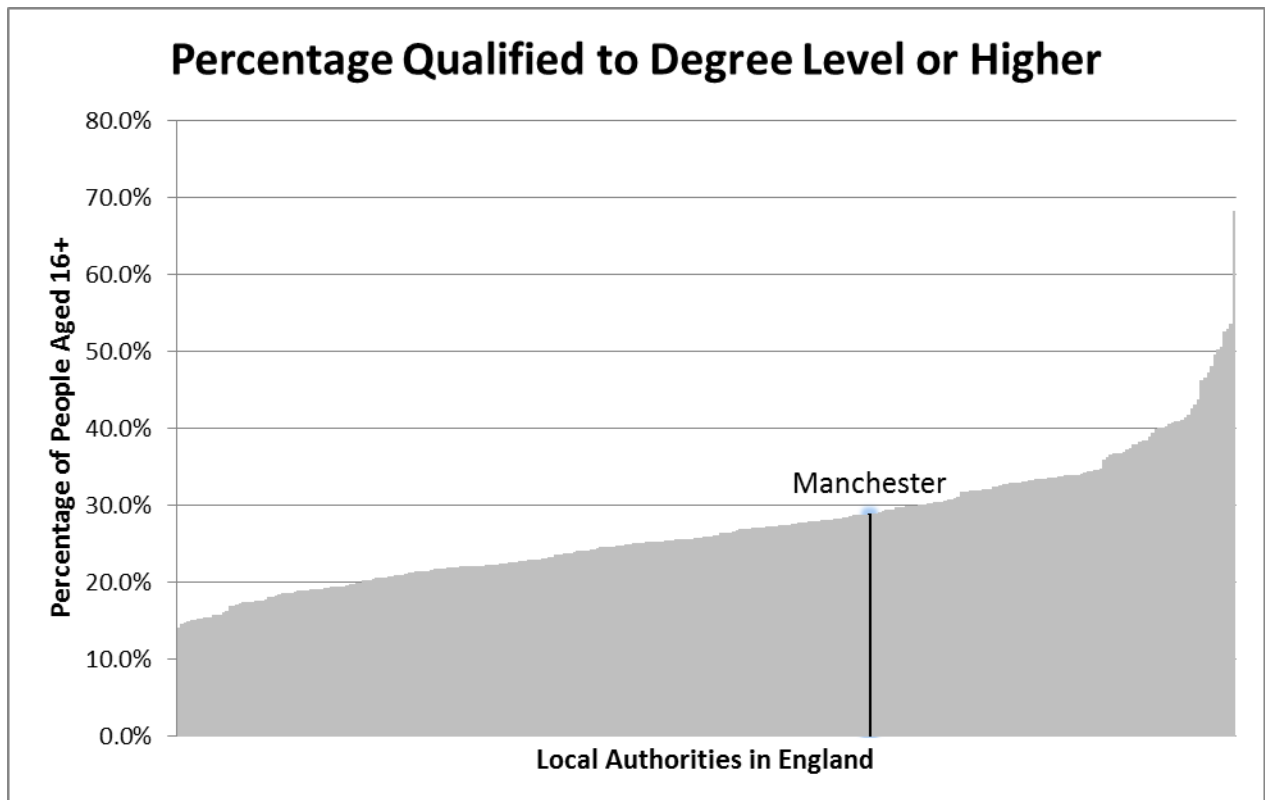
- Click on the titles at the bottom of the graph and press the <Delete> key
- Click on the legend (to the right) and press the <Delete> key
- Click on the title and replace the current label with something like 'Percentage Qualified to Degree Level or Higher'.
- Select the data series, then right click 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:

(2) 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?

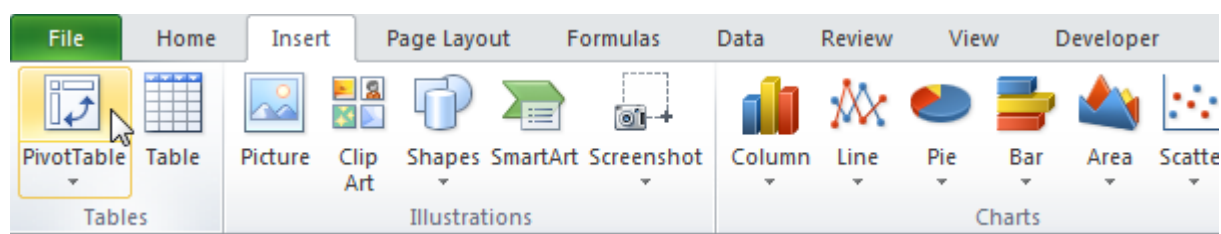
(3) Social Class – Pivot Tables

In this section we will be using Pivot Tables, a powerful feature of Excel that allows data to be grouped and aggregated. A way to view the wood rather than the individual trees. The first exercises aim to introduce a variety of useful features for pivot tables. The final exercise works with the census NS-SEC measure of occupation based social class, aiming to distinguish the occupational roles in which men and women predominate.

Exercise 1 – A Simple Pivot Table

The first task is to count the number of Local Authorities in each Region.

- Don't use the 'Social Class' worksheet yet. Instead click on a cell in the initial Qualifications worksheet.
- On the Insert tab, click PivotTable.



- Click OK in the dialog box. Excel has selected all the data and will, by default, place the pivot table in a new worksheet.
- Excel produces a blank pivot table with a panel on the right hand side of the screen for controlling it. We would like a list of regions with the count of LAs in each so, working on the panel, drag the <Region> field into the Row Labels box and the <LA> field should be dragged into the Σ Values box.

The pivot table and control panel should then look like this:

Row Labels	Count of LA
East	47
East Midlands	40
London	33
North East	12
North West	39
South East	67
South West	37
West Midlands	30
Yorkshire and Humberside	21
Grand Total	326

Drag fields between areas below:

Report Filter	Column Labels
Row Labels	Σ Values
Region	Count of LA

We have quickly produced a summary of the data, counting the number of local authorities in each district. Without pivot tables the process would have been much more painful, sorting the data by region and then counting by hand the number of rows for each.

Example 2 – Pivot Table Grouping by Many Variables

Now we will progress to the 'Social Class' data and try a more complicated example. Suppose you have been asked to identify the region with the largest rural population and to give a breakdown of this population by sex. The rural population are those living in either the 'Village' or 'Hamlet and dispersed' categories of the Urban-rural classification variable.

As before the Regions will be the rows in our pivot table. The results needs to be Filtered to just some of the urban-rural classes and the breakdown by sex requires separate Columns for each region. The Value to be counted is the total population.

- Produce a new pivot table on a new worksheet to the spec given above.

When done the table and panel should look like this:

Urban-rural	(All)		
Sum of Pop	Column Labels		
Row Labels	Female	Male	Grand Total
East	2971158	2875807	5846965
East Midlands	2298729	2234493	4533222
London	4140652	4033289	8173941
North East	1327183	1269703	2596886
North West	3587492	3464685	7052177
South East	4395452	4239298	8634750
South West	2698327	2590608	5288935
West Midlands	2838660	2763187	5601847
Yorkshire and Humberside	2685655	2598078	5283733
Grand Total	26943308	26069148	53012456

Drag fields between areas below:

Report Filter: Urban-rural

Column Labels: Sex

Row Labels: Region

Values: Sum of Pop

The results can be limited to the Rural Areas as follows:

- Click the down arrow on the Urban-rural filter at the top of the pivot table:



- Select the Multiple Items check box at the bottom of the menu.
- Adjust the urban-rural tick boxes so that only 'Village' and 'Hamlet or Dispersed' are selected.

The pivot table will then be limited to the local authorities that are mostly of a rural character.

Finally the results can be sorted so that the region with the greatest population appears at the top of the list:

- Right click on one of the numbers in the Grand Total column and, on the context menu, select <Sort> and then <Largest to Smallest>.

Q1 Returning to the original research question, which region has the greatest population within predominantly rural Local Authorities?

Q2 What is the rural population of men and women in this region?

If you have it right, the South West is the region of interest with a total rural population of 2,445,003 people.

The intention here was mainly to illustrate the process of creating and configuring a pivot table. An analysis like this could be interesting to those involved in planning the provision of council services. Delivery of services like refuse collection will tend to be more expensive in rural areas because of transport costs. However, in reality, a Local Authority may contain a combination of rural and urban areas. Therefore a more sophisticated analysis would use smaller units; electoral wards for example.

Example 3 – Collating Pivot Data for Analysis

The title of the worksheet suggests the pivot table exercise is to focus on Social Class. So, having become familiar with the tools for making and adjusting pivot tables, we now turn to this topic. The table gives data from the National Statistics Socio-economic Classification (NS-SeC). This measure assigns a class to people based on their occupation, whether they are self-employed and the size of the organisation where they work. Investigating sexual inequality in employment, research questions are:

- The highest class position is nssec1 – the higher managerial, administrative and professional workers. To what extent is this class skewed towards male workers?
- Ignoring the long term unemployed/those who have never worked and full-time students, which of the 7 classes are mostly composed of female workers.

The basic procedure for the analysis is to use a pivot table to collate the required data and this is used as a basis for calculating proportions for each class.

- Select the worksheet <Social Class>
- Insert a pivot table, creating a new worksheet.
- Put the Sex variable in Row Names
- Put all of the nssec variables in Σ values

By default the pivot table is showing total counts i.e. the number of males and females in each class group. The counts can be changed to percentages so doing this for the nssec1 variable will be sufficient for answering the first research question:

- Select one of the cells in the pivot table that collates nssec1 e.g. cell B4
- Right click and on the context menu select <Show Values As> and then click <% of Column Total> on the submenu.

So returning to the research question:

Q3 The highest class position is nssec1 – the higher managerial, administrative and professional workers. To what extent is this class skewed towards male workers?

It would be convenient to convert all the variables nssec1...nssec7 to give column percentages. However an inconvenient feature of Excel is that it insists this is done one variable at a time and, additionally, it will not calculate row totals and percentages across multiple variables. So an alternative is to copy the pivot data to make a table describing each of the social class groups.

- <Copy> the descriptions of the NS-SeC classes 1 to 7 from the Lookup worksheet and <Paste> them in column A beneath the Pivot Table.
- Make Column A wider, as appropriate, to show the full class descriptions.
- Select the pivot table data including the Row Labels Grand Total row.
- Do a <Paste Special> - <Transpose> to place the data next to the class descriptions.
- Calculate two further columns, working out the proportion of males and females in each class.
- Format the proportions as percentages to 1 d.p.

To check your progress, the sheet below shows the work mostly done, with only the percentage columns remaining:

	A	B	C	D	E	F	G	H
1								
2								
3		Sum of nssec1	Sum of nssec2	Sum of nssec3	Sum of nssec4	Sum of nssec5	Sum of nssec6	Sum of nssec7
4	Female	1407660	4404270	3676483	1089293	771637	3390047	1815652
5	Male	2638163	3727837	1295561	2573318	1904481	2040816	2461831
6	Grand Total	4045823	8132107	4972044	3662611	2676118	5430863	4277483
7								
8								
9								
10		Female	Male	Total	% Female	% Male		
11	Higher managerial, administrative or professional	1407660	2638163	4045823				
12	Lower managerial, administrative or professional	4404270	3727837	8132107				
13	Intermediate	3676483	1295561	4972044				
14	Small employer and self-employed	1089293	2573318	3662611				
15	Lower supervisory and technical	771637	1904481	2676118				
16	Semi-routine	3390047	2040816	5430863				
17	Routine	1815652	2461831	4277483				

Finally, with the percentages calculated, it will now be possible to answer the second research question:

Q4 Ignoring the long term unemployed/those who have never worked and full-time students, which of the 7 classes are mostly composed of female workers?

(4) Exploring Ethnicity and Place

In this section we are going to develop your skills so that you can produce good quality tables to support your analysis and reporting. This is a common task – having, produced some data in Excel or SPSS it needs to be *pretty printed*, making it presentable as a table for a report.

In this exercise the aim is to produce a table, showing the counts and percentages of each ethnic group in Local Authorities characterised by different urban and rural settings.

So you know what you're aiming for, here is the first row of the final table:

Ethnicity	Urban/Rural Setting of Local Authority		
	Mainly/Largely Urban (%)	Other Urban, Town & Fringe (%)	Village, Hamlet or dispersed (%)
White British	17,721,127 (41.9%)	14,133,402 (33.4%)	10,424,707 (24.7%)
...

To begin:

- Select the worksheet <Ethnicity Table> and review what has already been done to get you started.

Currently the ethnicity variables have been collated in a pivot table, drawing the data from the <Ethnic Diversity> worksheet. The contents of the pivot table have been transposed and, moving to the right, some of the groups have been aggregated.

The next task is to finish preparing the data:

- Starting at row 14, complete column M, combining the populations of the 'Village' and 'Hamlet or dispersed' categories.
- Complete columns N to P, calculating the row percentages.
 - For example row 14 is the wb (White British) group so cell N14 will contain the proportion of the whole White British Population living in a Largely Urban setting.
 - Once the proportions have been calculated they can be formatted as percentages to 1 decimal place.

Presenting Text pt I – the VLOOKUP function

The area where the final table should be built starts at cell R14. Up to now the ethnic groups have been described with short codes e.g. bcar for 'Black Caribbean' or ban for 'Bangladeshi'. As with the other exercises, the full description of the codes is given on the <lookup> worksheet.

Whilst we use short codes as a convenient shorthand during the analysis, they won't mean a lot to anyone else! Presenting a good finished table requires proper descriptions. The VLOOKUP function helps, using a lookup table to translate the codes to something more legible. It is much less painful than having to re-type all the descriptions every time you produce a table.

VLOOKUP - using a table to derive values

A lookup table is created with the value expected in a cell and additional columns to provide the other value(/s) related to it. The key or first field needs to be ordered. The format of the function is:

`=VLOOKUP (cell, values, column)`

cell holds the value you are looking up

values contains the table that holds the key and the related values

column the position of the related value relative to the first column i.e. the second column in the values table is referenced by 2.

Noting that:

- Our first code to lookup is 'wb' in cell J14
- The lookup table is on the <lookup> worksheet in cell range A13:B26

➤ In cell R14 enter the formula `=VLOOKUP (J14, lookup!A13:B26, 2)`

Note in the formula above 'lookup!' directs Excel to retrieve the values from the <lookup> worksheet.

When the formula is copied down the location of the lookup table must remain constant. The formula can be fixed by:

- Select cell R14 and edit it by pressing <F2> (or double click in cell R14).
- Move the cursor to the reference to cell A13 and press <F4>
- And do the same for the reference to cell B26
- Finish the edit by pressing <return>

The <F4> shortcut changes cell references so they are static, referring to a fixed location. The formula should now read:

=VLOOKUP(J14, lookup!\$A\$13:\$B\$26, 2)

- Copy the formula down as far as row 28 to lookup the descriptions for all the ethnic groups.
- The final row is for the population as a whole so enter 'Total Population' in cell R29.
- Highlight R29 in Bold

The completed column is illustrated on the right

Ethnicity
White British
White Irish
White Gypsy or Irish Traveller
White Other
Mixed
Indian
Pakistani
Bangladeshi
Chinese
Asian other
Black African
Black Caribbean
Black other
Other Ethnic Group
Total Population

Presenting Text pt II – the TEXT and CONCATENATE functions.

For the columns describing the populations for the different urban/rural settings we are aiming to present the population followed by the percentage in brackets. So the first cell will contain both:

- The total White British population residing in a Mainly/Largely Urban setting
- The percentage of the group in this setting (surrounded by brackets)

This first cell, S14, will contain a nicely formatted text string like this:

17,721,127 (41.9%)

Assembling content like this in a cell requires the use of some of Excel's functions for manipulating text:

Text String Functions

We can put together values from different places to provide a formatted string of text. In the next step we are going to create a format for the first cell and then copy it to the others. This will combine the use of the two functions

TEXT(*cell*, *format*) which converts a number to the format specified in the second parameter.

CONCATENATE(*text1*, *text2*, ...) which adds together two or more strings

Below are some examples of the functions usage. For the purpose of these examples, cell A1 contains the value 12345 and cell A2 has the value 0.1

Function usage	Resulting Text
=TEXT(1.5, "£0.00")	£1.50
=TEXT(A1, "#,###")	12,345
=TEXT(A2, "0.0%")	10.0%
=CONCATENATE("Hello", " World")	Hello World
=CONCATENATE("(", TEXT(A1, "#,###"), ")")	(12,345)


=CONCATENATE("s5", "t5")

- Write a formula for cell S14 so the data is presented in the required format. i.e. the cell should say:

17,721,127 (41.9%)

- The formula is more complicated than anything covered so far. It might help to build it up in small steps rather than trying to get the whole thing right first time.
- The formula can then be copied to all the table cells (S14:U28)

The table contents are now complete. All that remains to be done is some tidying up to make the table more presentable i.e. making column headings bold, aligning cell contents and adding lines around the cells.

- Double clicking on the line between the column letters will automatically make the column adjust to fit the largest contents
- Select the table headers and click **B** on the home ribbon to make the headings bold
- Selecting the numerical data (S14:U28), these will look better if the cell contents are aligned to the right. On the home ribbon you can change the alignment of this text.
- Finally add borders to the table using the 'Borders' tool  on the home ribbon.



As a reminder, the first row of the final table should look like this:

Ethnicity	Urban/Rural Setting of Local Authority		
	Mainly/Largely Urban (%)	Other Urban, Town & Fringe (%)	Village, Hamlet or dispersed (%)
White British	17,721,127 (41.9%)	14,133,402 (33.4%)	10,424,707 (24.7%)
...

You may like to experiment with copying and pasting the table into Word. Excel and Word don't always agree about the width required for some text so this might require some final further adjustments to the column widths.

Well done, that's the final exercise completed. If you have time you might like to refer back to the Excel Basic training as well. For example the final exercise illustrates the process of downloading census data, using the same website that was used to acquire the data for this workbook.

Further resources

- Whatever you are trying to do remember that pressing F1 to get help is always an option.
- There are good online training materials like those at Excel Easy:
<http://www.excel-easy.com/>

This training session introduced various Excel features in the context of work to look at research questions across a range of social data. The approach at Excel Easy is much more skill-focussed, providing specific sections on basic skills like 'Printing' and 'Formatting a cell' through to more advanced topics like how to deal with 'Formula errors' or the creation of 'Pivot tables'.

- 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>

Appendix 1 - Dataset

The dataset being used is similar to the one prepared for the Excel basic course. It 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 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

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)

Appendix 2 – Feedback / Answers

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 (2.4%), 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