



2011 Census Microdata Teaching File – Possibilities for exploring the data

Introduction

The Teaching File is a sample of data from the 2011 Census and represents an ideal resource for training and research.

www.ons.gov.uk/census



The Teaching File is a sample of data from the 2011 Census and represents an ideal resource for training and research. In particular, it offers the following opportunities not available from the standard 2011 Census statistical bulletins, these include:

- **Custom variable combinations**

Users can create their own tables based on whichever variables they choose.

- **More detailed analysis**

Variables can be examined in more detail than is available from the standard release tables.

- **Statistical analysis**

Person-level data provide additional potential for analysis not possible with aggregate data.

Here, we provide some examples of some of the types of data exploration that could be conducted with the Teaching File data. We will show one of the ways in which custom variable combinations can be created and how more detailed information can be obtained from the 2011 Census Teaching File. We will not be addressing statistical analysis in this document.

These examples are intended to illustrate to students how the raw, person-level data can be used differently from the published statistical tables. This document is not intended to act as a 'how to' guide or to provide instruction in any way but to give users an idea of the potential that this resource provides.

The Teaching File is available in csv format and can readily be used with a variety of software packages. For these examples, we have used Microsoft Office – Excel 2007 to manipulate the data but other programmes could be used.

The population base (that is, the group of individuals for whom records have been included) for the Teaching File includes usual residents, short-term residents and students counted at their non term-time address (see [Teaching File User Guide](#) for further details). The population base for the following examples of data exploration includes all residents.



Custom Variable Combinations

The teaching file dataset enables any user to create tables summarising data from any chosen set of variables.

Person ID	Region	Residence	Family	Co	Populatio	Sex	Age	Marital St	Student	Country o	Health	Ethnic Grc	Religion	Economic
7394816	E12000001	H	2	1	2	6	2	1	2	1	2	1	2	5
7394745	E12000001	H	5	1	1	4	1	2	1	1	1	1	2	1
7395066	E12000001	H	3	1	2	4	1	2	1	1	1	1	1	1
7395329	E12000001	H	3	1	2	2	1	2	1	2	1	2	1	1
7394712	E12000001	H	3	1	1	5	4	2	1	1	1	1	2	1
7394750	E12000001	H	2	1	1	6	2	2	1	2	1	1	1	1
7394871	E12000001	H	5	1	2	4	3	2	1	2	1	1	1	1
7394832	E12000001	H	3	1	2	1	1	2	1	2	1	2	1	-9
7394719	E12000001	H	2	1	1	7	2	2	1	1	1	1	2	5
7394840	E12000001	H	1	1	2	6	4	2	1	3	2	2	5	1
7394510	E12000001	H	1	1	1	3	1	2	1	2	1	1	1	1
7394711	E12000001	H	2	1	1	1	1	1	1	1	1	1	1	1
7395070	E12000001	H	1	1	2	2	1	2	1	2	1	2	1	4
7394869	E12000001	H	1	1	2	2	1	1	1	1	2	1	1	4
7394971	E12000001	H	5	1	2	2	1	2	1	2	1	2	1	4
7395019	E12000001	H	1	1	2	2	1	1	1	1	2	1	1	4
7395069	E12000001	H	3	1	1	2	1	2	1	1	1	1	1	2
7394933	E12000001	H	3	1	1	2	1	1	1	1	1	1	1	6
7395059	E12000001	H	1	1	1	4	1	2	1	3	1	1	1	1
7394875	E12000001	H	2	1	1	1	1	1	1	1	1	1	2	-9
7394764	E12000001	H	2	1	1	7	2	2	1	3	1	2	5	2
7394645	E12000001	H	5	1	2	5	4	2	1	3	1	2	2	2
7394857	E12000001	H	2	1	1	4	2	2	1	1	1	1	1	1

Use Pivot Tables to create summaries of custom combinations of variables.

Row Labels	Column Labels	1	2	Grand Total
5	-9	57809	109618	167427
6	1	107997	216056	324053
7	2	28790	23684	52474
8	3	10760	14698	25458
9	4	6587	15060	21647
10	5	40605	113750	154355
11	6	12620	24272	36892
12	7	1715	32460	34175
13	8	9222	17538	26760
14	9	4464	11208	15672
15	Grand Total	280569	578344	858913

For further information on how to create pivot tables, see the [annex](#).
Source : Screenshots provided by RSSCSE, February 2014.

Example – Sex and Employment

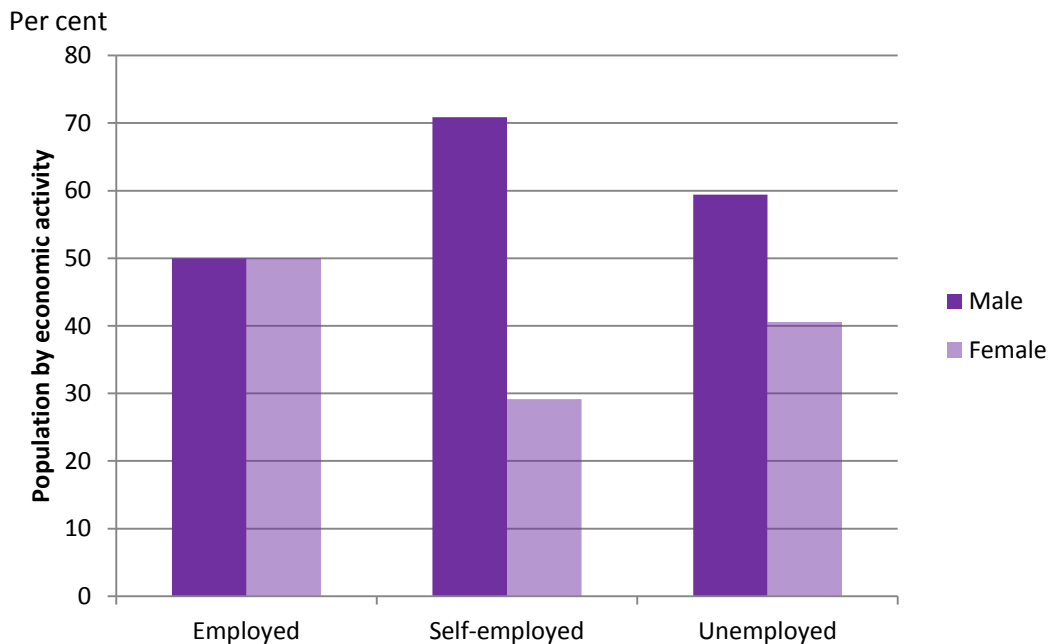
In this example we used the **Sex, Hours Worked per Week, Industry** and **Economic Activity** variables from the Teaching File to explore the effect that a person's sex had on their employment status. We did this by comparing these variables against each other.

Initially we compared the sex breakdowns for people who were unemployed (economically active – so looking for work), working as an employee or working self-employed (Figure 1).

All employment questions refer to a person's main job, which is the job in which they worked the most hours. We found the following results:

- More men are unemployed than women (59.4% male, 40.6% female).
- Of those working who are self-employed, 70.9% are male and 29.1% are female.

Figure 1. Sex differences among unemployed, employees and self-employed individuals.



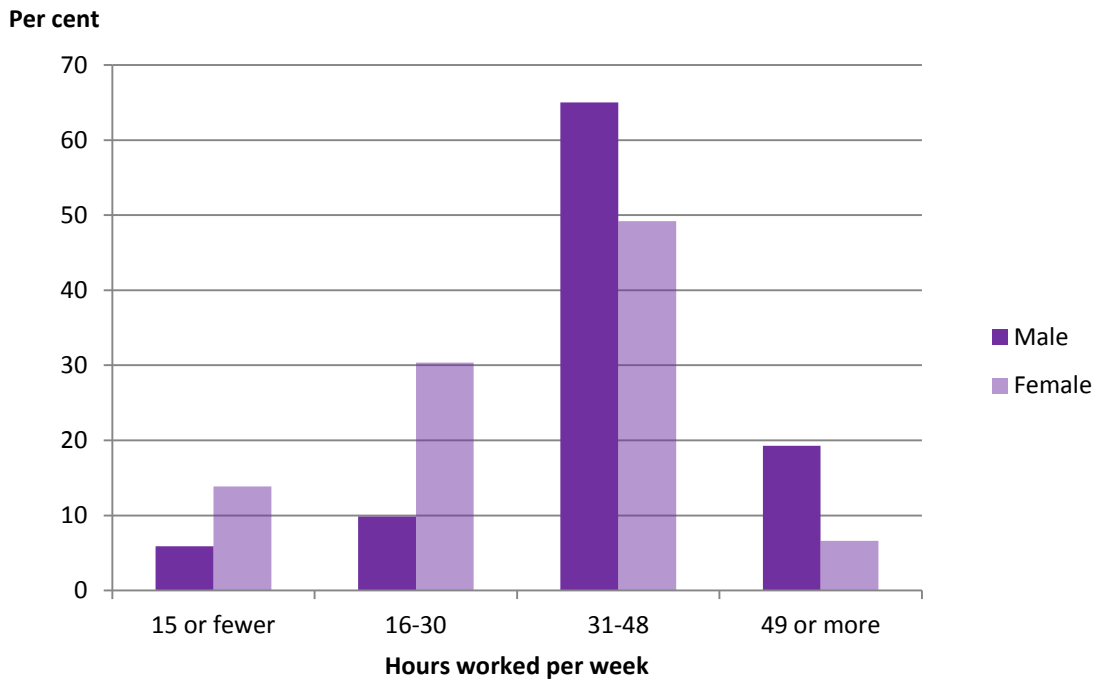
Source: England and Wales, 2011. All residents.

We then compared the sex breakdown against number of hours worked per week (Figure 2):

- A greater proportion of women (44.2%) than men (15.8%) work part-time (30 hours or fewer per week).
- In contrast, 6.6% of women and 19.3% of men work more than 49 hours per week.



Figure 2. Differences in working hours for men and women

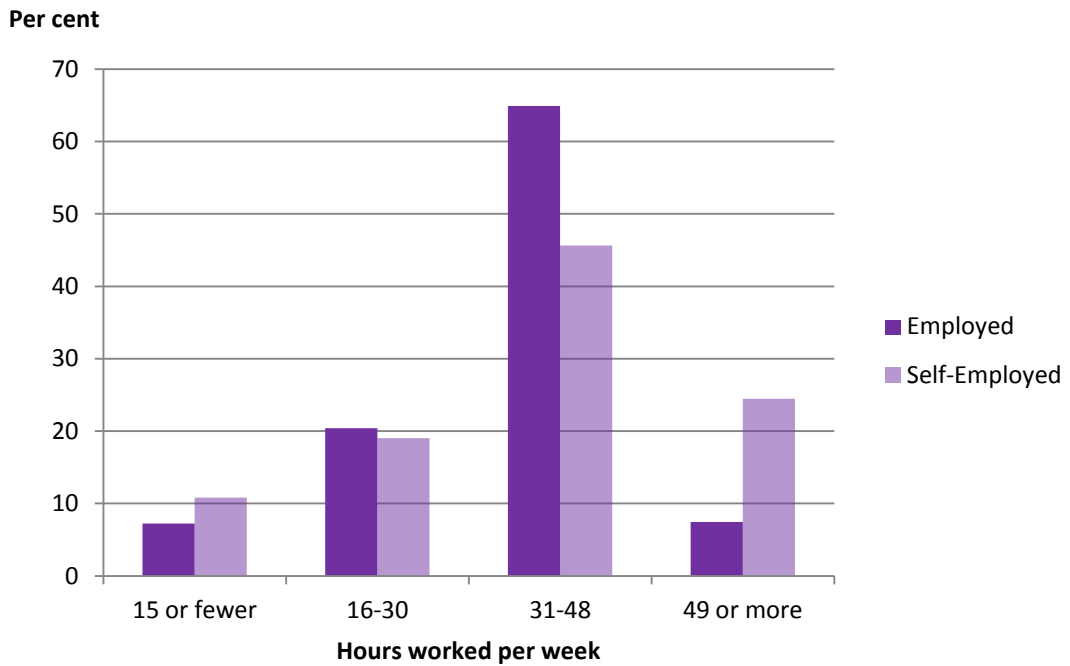


Source: England and Wales, 2011. All residents.

Next, we compared the respondents' economic activity status against the number of hours worked per week (Figure 3):

- A greater percentage of self-employed workers (24.5%) than employees (11.7%) work 49 or more hours per week.

Figure 3. Differences in working hours for employees and self-employed respondents

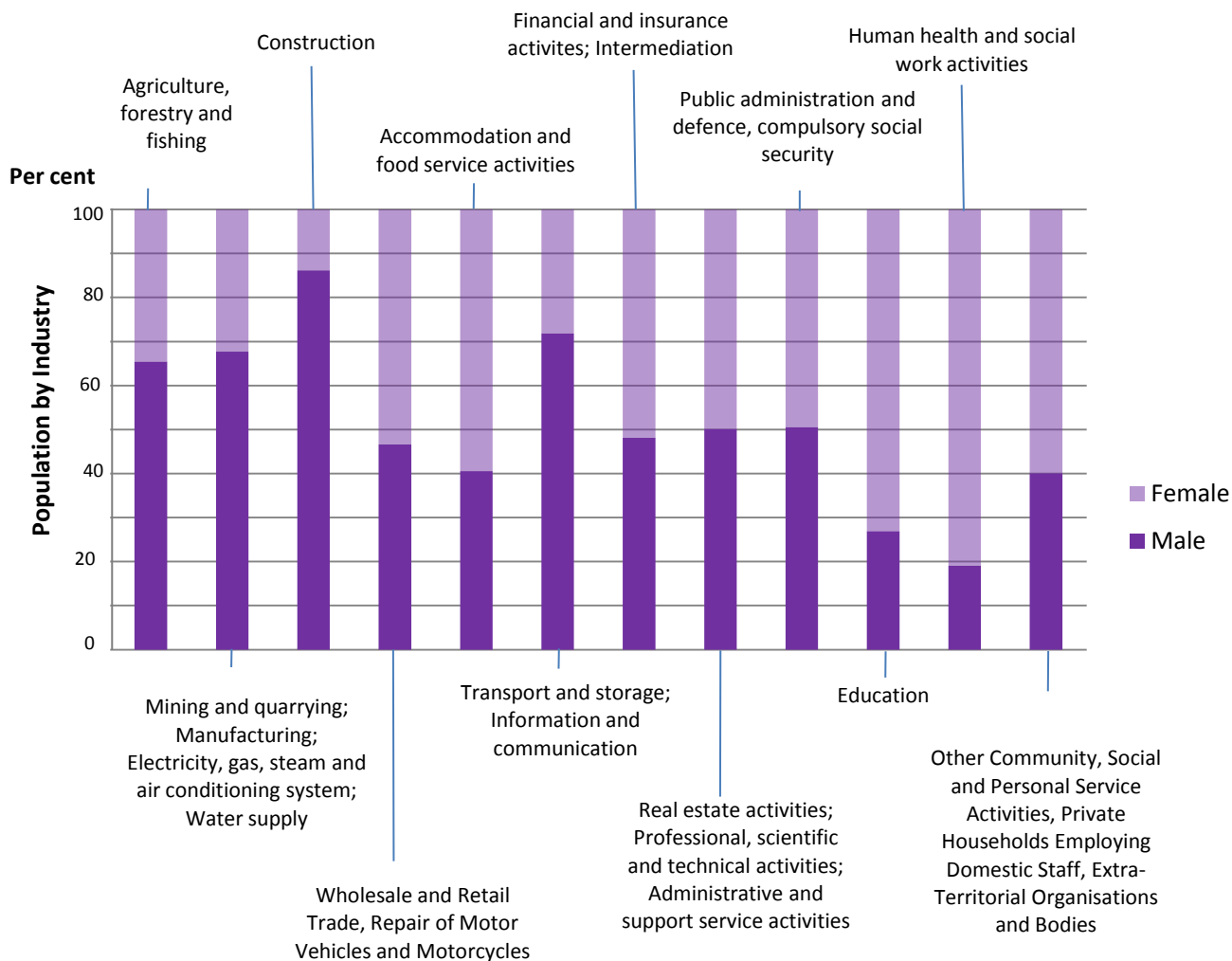


Source: England and Wales, 2011. All residents.

Finally, we looked at the distribution of men and women working in each industry (Figure 4):

- We found very similar proportions of men (49.4%) and women (50.6%) working across all industries as a whole but there was considerable variation within industries.
- ‘Construction’ showed the largest proportion of male workers (79.8%).
- ‘Human health and social work activities’ showed the largest proportion of female workers (81.0%).

Figure 4. Distribution of men and women working by industry



Source: England and Wales, 2011. All residents.

More Detailed Analyses

The Teaching File also enables analyses at a more detailed level than is currently available from tables published on census data by enabling further investigation of a single category within a variable.

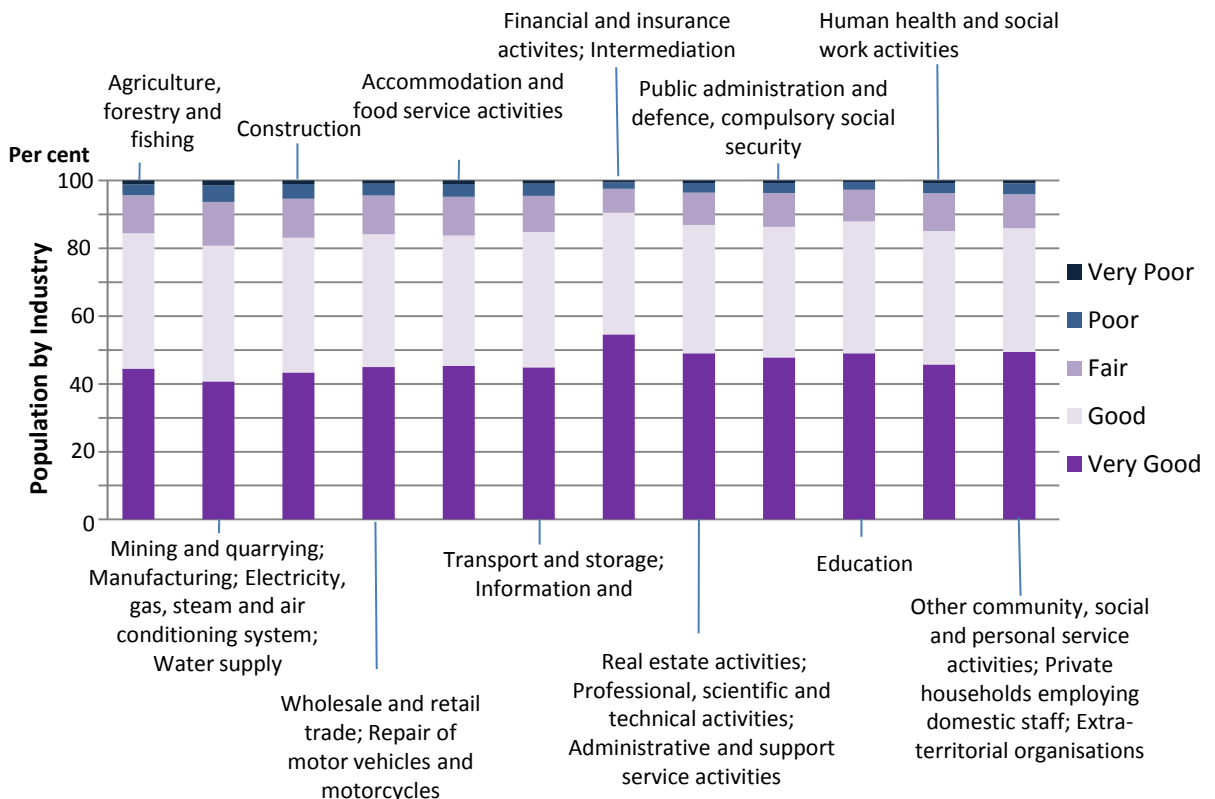
Example – Health

In this example, we used the Teaching File to incorporate general health into our analyses, enabling us to investigate Industry and Employment status in more detail.

In the 2011 Census, respondents were asked to rate their overall health as very good, good, fair, bad or very bad.

Here, we looked at the general health of people across different industries (Figure 5). For those not currently working, industry refers to the last industry in which they worked. The data were filtered to include only individuals aged 16-64 years old.

Figure 5. Self-assessment of general health across industries



There appears to be some variation in reported health across industries, with the Financial and Insurance activities and Intermediation sector reporting 54.6% of people with very good health compared with a mean across all industries of 46.6%.

By choosing one specific industry, in this case we have chosen Education, we can examine the health of respondents in even more detail by looking at the differences in proportions of people within that industry reporting very good health according to the number of hours that they work per week.



Figure 6. Variation in proportion of persons reporting very good health according to working hours within the Education industry



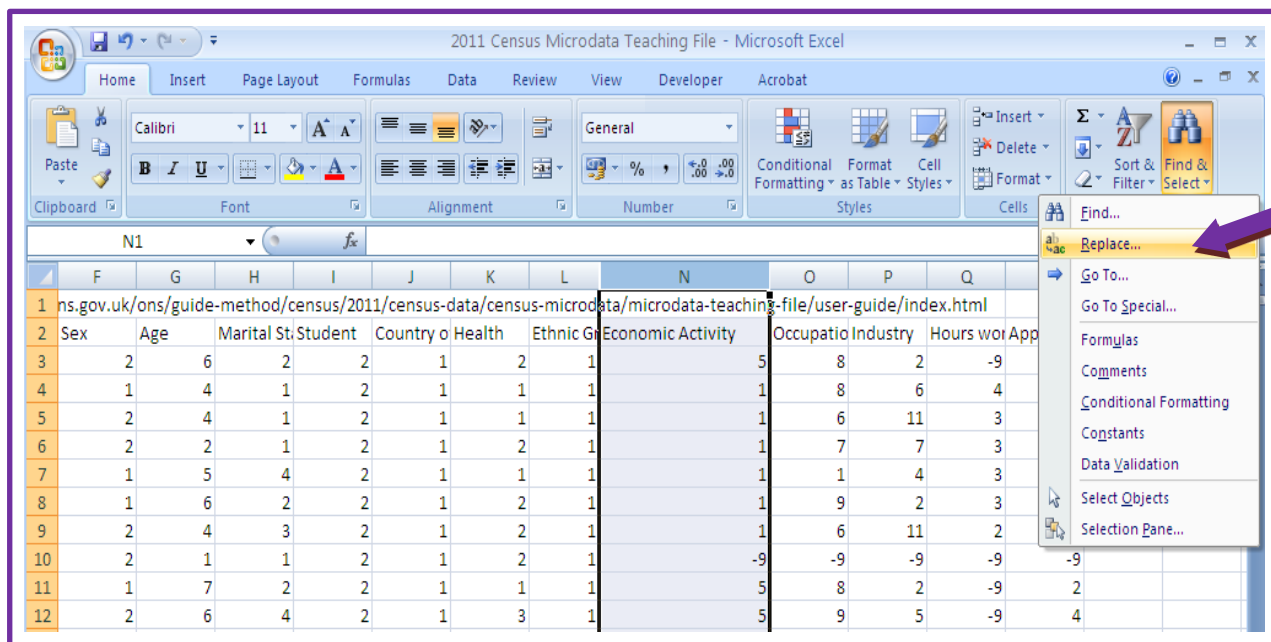
Within the Education industry, 48.1% of people working fewer than 15 hours per week reported very good health, compared with 59.5% of those working more than 49 hours per week.

Annex : How to use pivot tables and draw charts in Excel

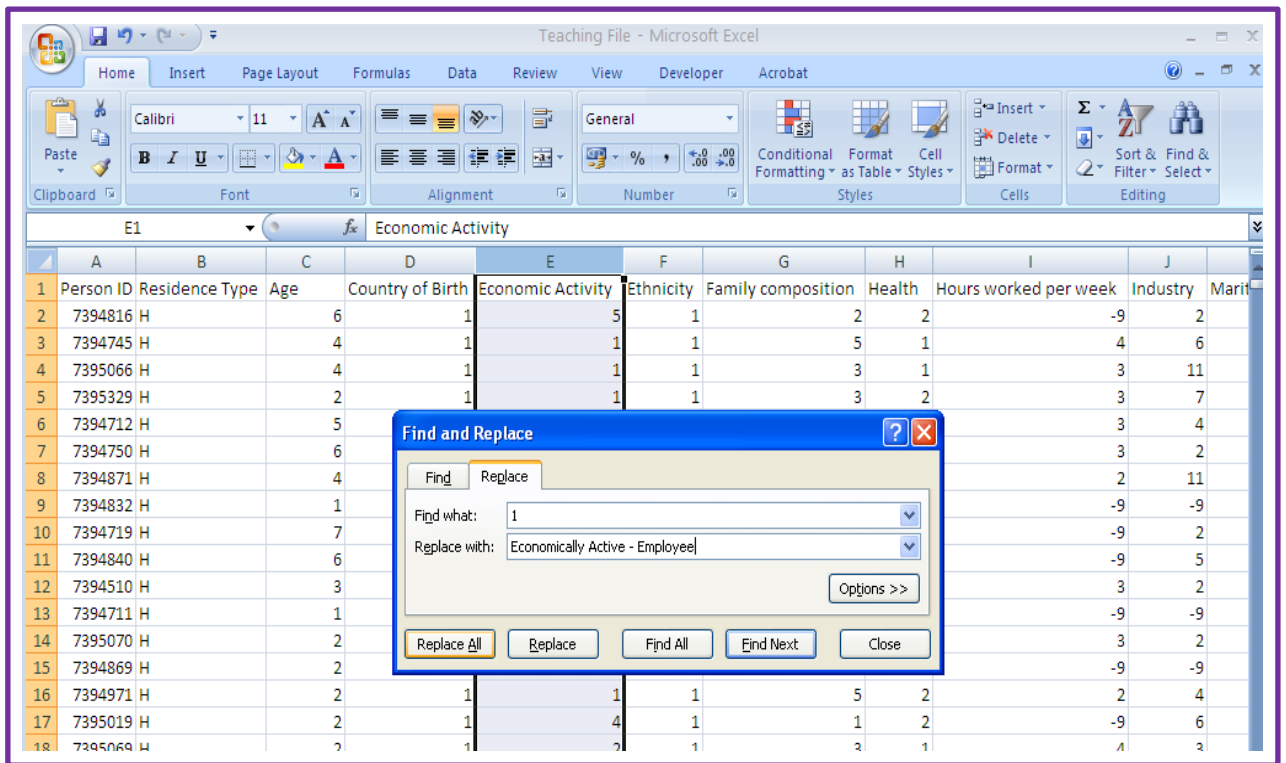
A Pivot Table is a way to present data quickly.

Using a Pivot Table to compare employment status for males and females

Open the [Teaching File dataset](#). To make it easier to examine the data, you may wish to replace the existing category codes with their category names. To do this, select a column (for this example Economic Activity), click **Find & Select > Replace** (or press Ctrl + H).

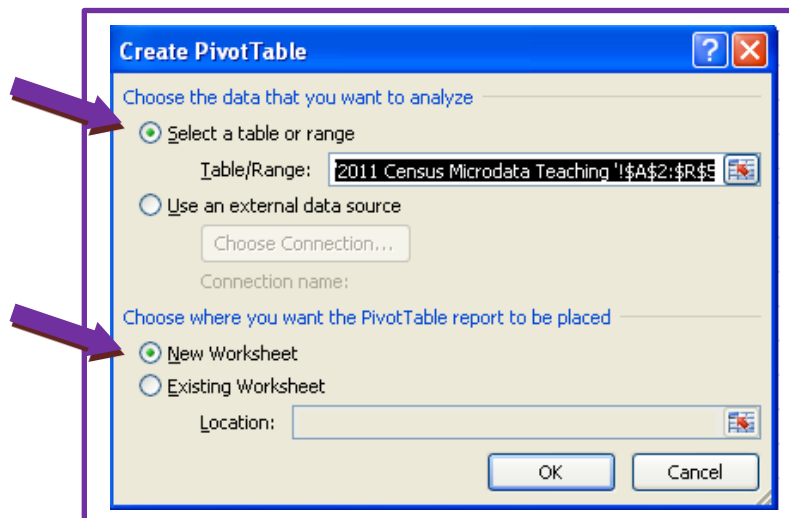


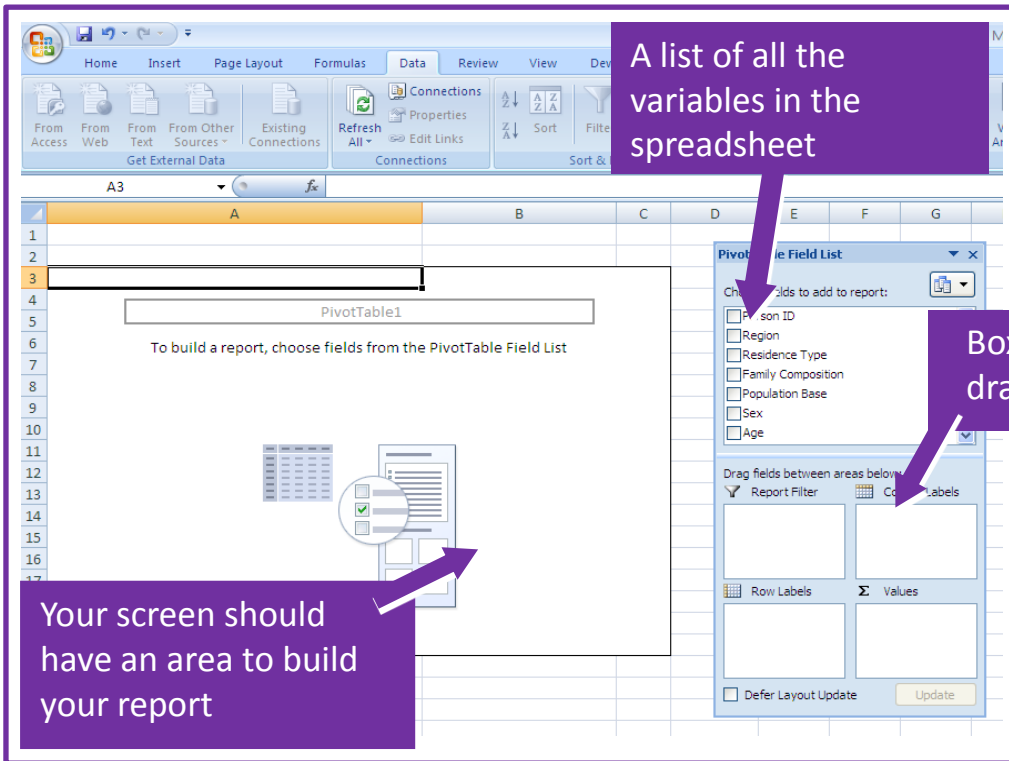
Type a category code (e.g. 1) into the **Find what** box and enter the corresponding category name (e.g. Economically Active – Employee) into the **Replace with** box. Click **Replace All**. Repeat this process for each category in the variable and for each variable that you are interested in.



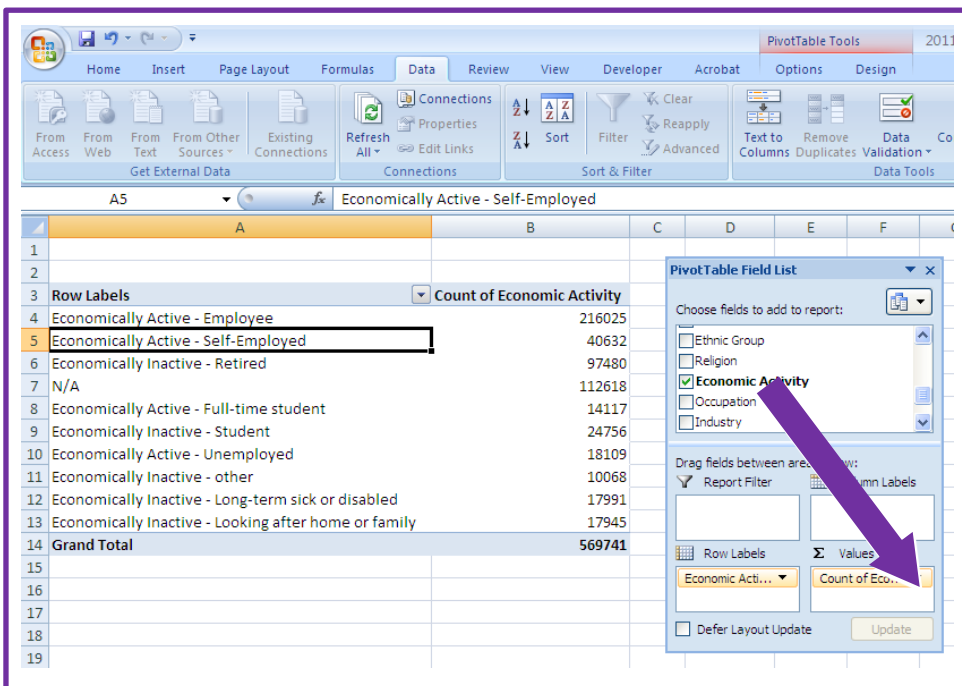
Creating a Pivot Table

1. Select **Insert > PivotTable > PivotTable**
2. Click on **Select a table or range** (the whole data set should be included. There is no need to select any cells before inserting a pivot table).
3. Click on **New Worksheet/OK**





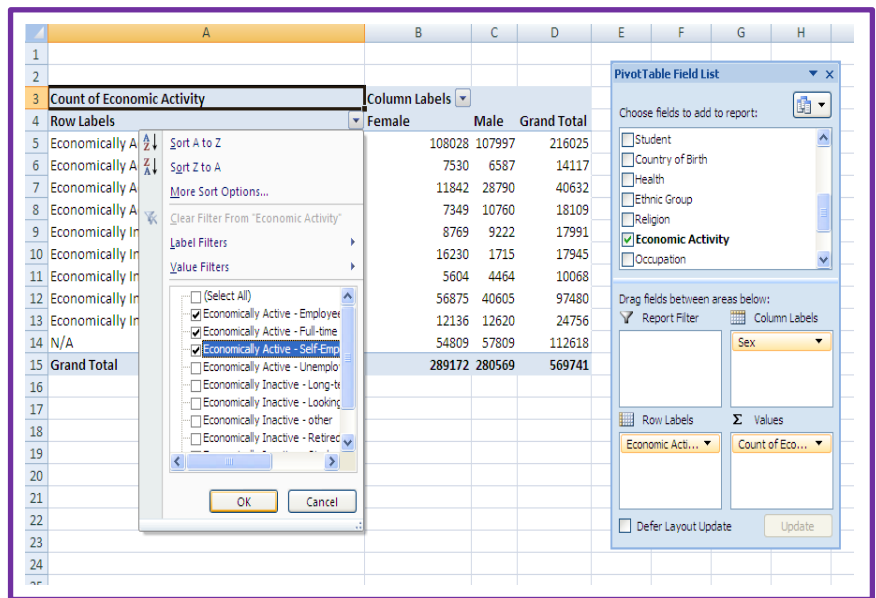
1. Select **Economic Activity** in the Pivot Table Field List (you may need to scroll down). Make sure your cursor is in the table in Column A. If the cursor is outside the table the variables on the right hand side will disappear.
2. Click on **Economic Activity** (so it is ticked). This should appear in the Row Labels.



Drag **Economic Activity** from the Pivot Table Field List to the Σ Values. The frequencies now appear in the table in **column B**.

If you now want to compare sex and economic activity, or more specifically, to see if there is a difference in proportions of male and female workers who are employees, self-employed or unemployed, drag **Sex** into the **Column labels** box.

Click on the arrow to expand **Row Labels** in column and select **Economically Active – Employee, Economically Active – self-employed and Economically Active – unemployed**. Click OK.



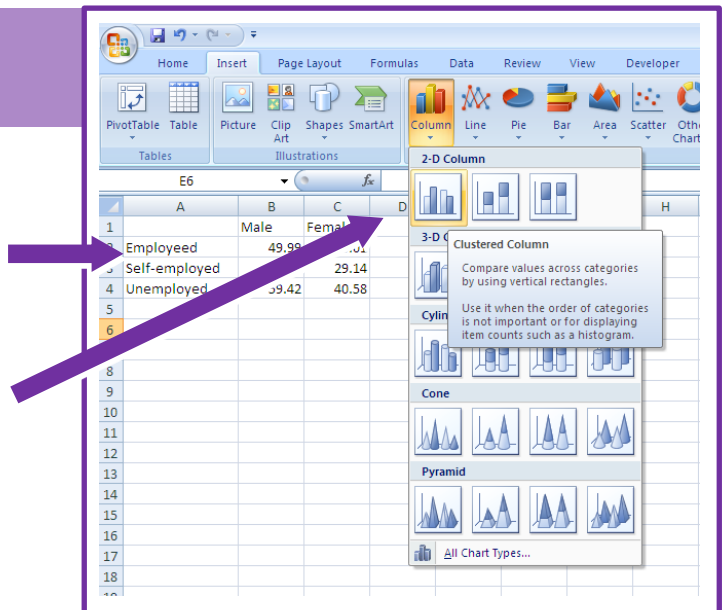
You now have counts for males and females for each selected economic activity category. Counts can be converted into proportions or percentages to aid interpretation of results (e.g. $11842/40632 = 29.1\%$ of self-employed people are female).

Count of Economic Activity	Column Labels		
Row Labels	Female	Male	Grand Total
Economically Active - Employee	108028	107997	216025
Economically Active - Self-Employed	11842	28790	40632
Economically Active - Unemployed	7349	10760	18109
Grand Total	127219	147547	274766

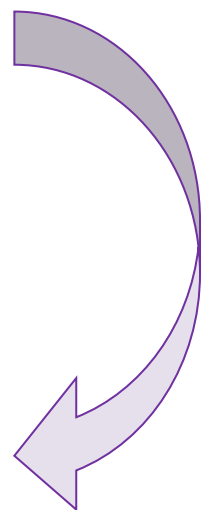
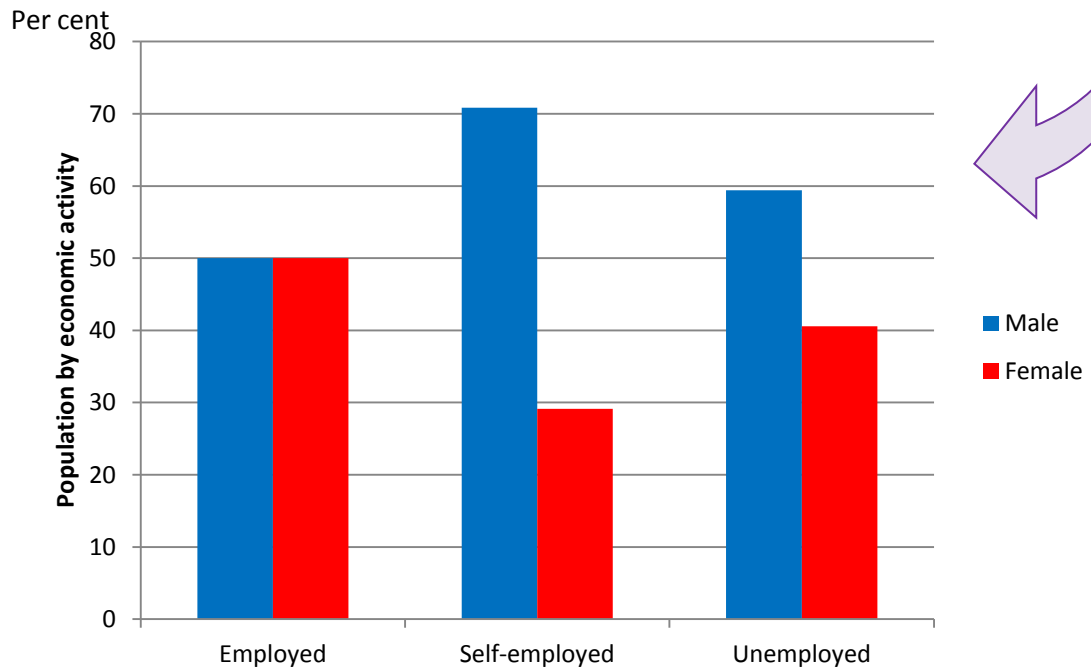
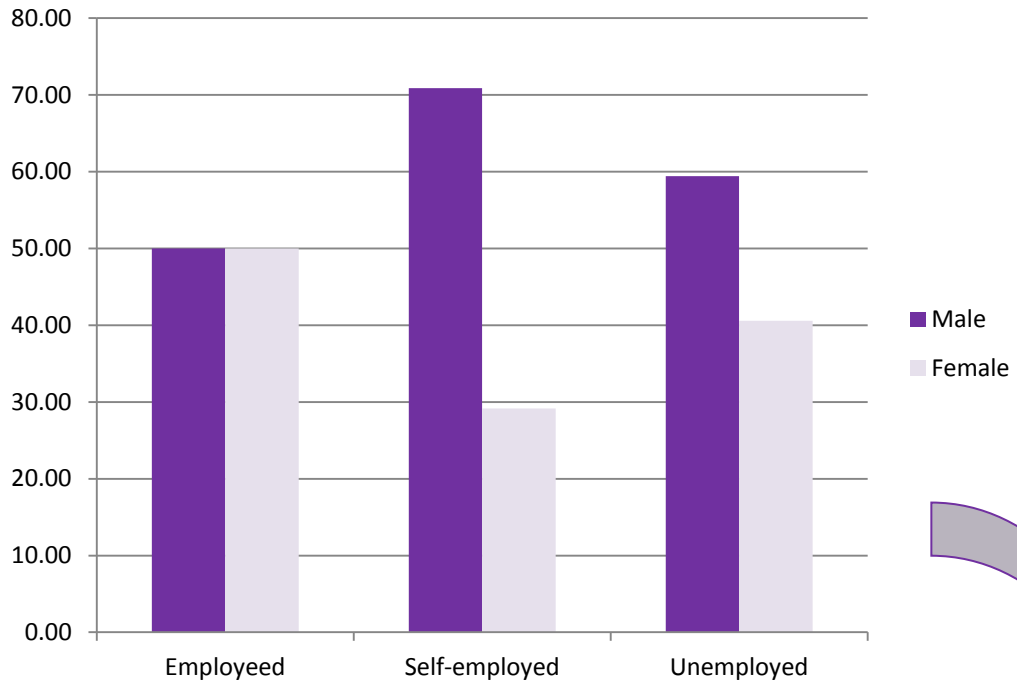
Plotting a bar chart from the pivot table data

To plot a bar chart from the pivot table data, create a new table using per cent values derived from the pivot table data.

Click on **Insert > Column > 2-D Column** and click on the first graph type.



Click on **Chart tools > Design / Layout / Format** to adjust the colour scheme, axis labels and formats as appropriate.



Need further help

Census Customer Services
Office for National Statistics
Segensworth Road
Titchfield
Fareham
Hampshire
PO15 5RR

census.customerservices@ons.gsi.gov.uk

Tel : 01329 444972
Fax : 01633 652981

