Using the 2011 Census Microdata Teaching File

Paul Norman
School of Geography
University of Leeds
LS2 9JT
p.d.norman@leeds.ac.uk

Introduction

Various Samples of Anonymised Records (SARs) have been extracted from the 1991 and 2001 Censuses and these have provided an invaluable research resource. Three levels of microdata are planned to be extracted from the 2011 Census: a Teaching file; Safeguarded files and Virtual Microdata Laboratory (VML) files available in an ONS safe setting. The SARs complement census area data: by allowing the user to derive custom variables; to create versatile crosstabulations; and to use statistical models which, for example, help avoid the ecological fallacy (Norman & Boyle 2010).

Whilst large and complex files, the SARs microdata are not so much of a challenge to the user as the Office for National Statistics (ONS) Longitudinal Study for England & Wales (and the Scotland and Northern Ireland equivalents) and allow a study coverage of more than just England & Wales, or Scotland or Northern Ireland. Other datasets which a user might access in census years and during inter- and post-censal years are area data from administrative records and large scale government surveys such as the Labour Force Survey (LFS). The SARs though are a much larger sample than a source such as the LFS and cover both household and institutional residents. The SARs have underpinned a wide range of research which includes: Boyle et al. (2002), Norman & Purdam (2013) Norman (2008) and Norman et al. (2014).

The Office for National Statistics (ONS) has devised a 2011 Census Microdata Teaching File to help users learn how to use microdata and prepare initial analyses for the use of the other SAR releases. This document first steps through the use of a 'customised' version of the 2011 Census Microdata Teaching File and provides SPSS syntax to first read in the original file to SPSS and label the variables and then to revise some of the variables to improve their utility in the context here.

The original 2011 Microdata Teaching File is available from ONS via:
http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/census-microdata/microdata-teaching-file/index.html

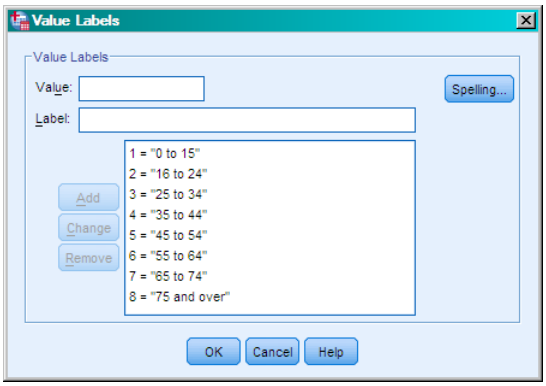
Calculating demographic measures using cross-sectional microdata

Please use this section of the document in conjunction with **Censusmicroteaching11ew-PN.sav** which is a slightly revised version of the Microdata Teaching file for 2011 provided by ONS. This is very similar to the Samples of Anonymised Records (SARs) as available for the 1991 and 2001 Censuses and is a parallel resource to the 2011 SARs. Using this teaching file is good preparation for using the SARs and other large survey data. The SPSS syntax used to read the csv file provided by ONS (source noted above) is included in this document, along with the syntax used to redefine variables used here. An introduction to using SPSS is provided by Norman et al. (2008).

Exploring the dataset: frequency counts

Open Censusmicroteaching11ew-PN.sav and familiarise yourself with the 'Data View' and 'Variable View' windows.

There are Value Labels for categories of all variables except the person ID.



Request frequency counts for a variety of variables to get a feel for the numbers and category distributions in the dataset. Try, age-group, ethnicity and others.

Do this via Analyse > Descriptive Statistics > Frequency Counts

Age

		Frequency	Percent
Valid	0 to 15	106832	18.8
	16 to 24	72785	12.8
	25 to 34	75948	13.3
	35 to 44	78641	13.8
	45 to 54	77388	13.6
	55 to 64	65666	11.5
	65 to 74	48777	8.6
	75 and over	43704	7.7
	Total	569741	100.0

Ethnic group

		Frequency	Percent
Valid	White	483477	84.9
	Mixed	12209	2.1
	Asian and Asian British	42712	7.5
	Black or Black British	18786	3.3
	Chinese or Other ethnic group	5753	1.0
	Total	562937	98.8
Missing	No code required (student/pupil living away /not resident in EW)	6804	1.2
Total		569741	100.0

Exploring the dataset: cross-tabulations

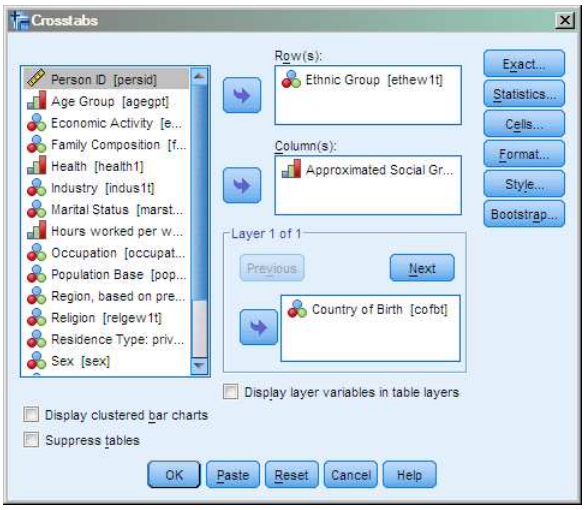
Whilst still at the mercy of the variables selected for any version of the SARs and the categories of data detail, the microdata allow the versatile cross-tabulation of variables with specifications not available in the area tables. Whether or not the following cross-tabulations are actually available in the area tables doesn't matter for this exercise, but variations in the age-structure of ethnic groups would be of interest in demographic analysis, as would age by self-reported health.

Try these cross-tabulations: Analyses > Descriptive Statistics > Crosstabs (select age as the rows and the other variable as the columns)

		Ethnic Group					
		White	Mixed	Asian and Asian British	Black or Black British	Chinese or Other ethnic group	Total
Age Group	0 to 15	83572	5392	10411	5077	1362	105814
	16 to 24	54696	2225	6887	2564	841	67213
	25 to 34	60327	1873	9065	3166	1332	75763
	35 to 44	66457	1162	6710	3280	1018	78627
	45 to 54	68926	885	4393	2576	600	77380
	55 to 64	61116	343	2927	938	338	65662
	65 to 74	46291	188	1449	678	171	48777
	75 and over	42092	141	870	507	91	43701
Total		483477	12209	42712	18786	5753	562937

		Health					
		Very good health	Good health	Fair health	Bad health	Very bad health	Total
Age Group	0 to 15	83580	19167	2360	513	194	105814
	16 to 24	43701	19821	2986	535	170	67213
	25 to 34	42591	27193	4554	1137	288	75763
	35 to 44	37417	30902	7373	2320	615	78627
	45 to 54	28963	32573	10818	3900	1126	77380
	55 to 64	17261	28263	13716	5035	1387	65662
	65 to 74	8185	20585	14225	4485	1297	48777
	75 and over	3273	13240	18448	6633	2107	43701
Total		264971	191744	74480	24558	7184	562937

You can dig a bit deeper using three way cross-tabulations. There may be different levels of occupational attainment (Social Grade) by both ethnic group and by country of birth.



Ethnic group (rows), Social Grade (columns) and Country of Birth (Layer) gives the following output.

Ethnic Group * Approximated Social Grade * Country of Birth Crosstabulation

Count

Country of Birth			Approximated Social Grade				Total
			AB	C1	C2	DE	
UK	Ethnic Group	White	65220	127604	68390	98624	359838
		Mixed	726	2131	567	1102	4526
		Asian and Asian British	1855	4300	687	1764	8606
		Black or Black British	674	2264	566	964	4468
		Chinese or Other ethnic group	165	355	65	128	713
	Total		68640	136654	70275	102582	378151
Non UK	Ethnic Group	White	7238	10026	4906	8920	31090
		Mixed	384	787	270	569	2010
		Asian and Asian British	4261	7457	2772	7604	22094
		Black or Black British	1190	3426	1277	2916	8809
		Chinese or Other ethnic group	607	1292	436	1149	3484
	Total		13680	22988	9661	21158	67487
Total	Ethnic Group	White	72458	137630	73296	107544	390928
		Mixed	1110	2918	837	1671	6536
		Asian and Asian British	6116	11757	3459	9368	30700
		Black or Black British	1864	5690	1843	3880	13277
		Chinese or Other ethnic group	772	1647	501	1277	4197
	Total		82320	159642	79936	123740	445638

Exploring derived variables

Note that extra variables have been added to the ONS supplied microdata teaching file. These are a simplified general health variable, a simplified religion variable and a derived individual multiple deprivation variable. The syntax for their derivation is available in the course resources. Check the frequency counts of these variables and cross-tabulations with others.

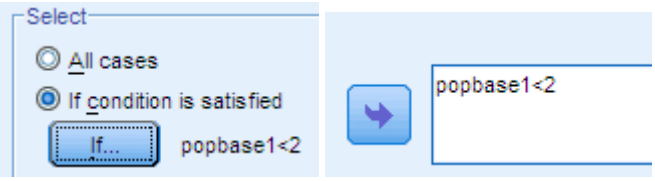
Using a filter

A microdata file allows the selection of a subset of the data to thereby focus the study on a particular subgroup of interest. This could be all females over the age of 55, for example.

Inspection of the frequency counts and crosstabulations above will show that the overall totals on the tables can be different. This is because some data may be missing for a variety of reasons (not applicable, no answer, etc.).

It would appear that the specification of this microdata teaching file has not provided well enough specified data to allow for students and short term visitors to be included in analyses. In this instance, this reminds that any data source may not necessarily be ideal for the research being carried out and that some adjustments may be needed and later some caveats reporting.

Click on:  and then 'If condition is satisfied' and the logical selection of popbase<2 which would then have only persons who are 'Usual residents' in any subsequent analyses (they are not deleted from the file, unless you request this, so don't!).



The output window has the SPSS syntax for this filter which can be a good way of also requesting data subsets. The following is the syntax for omitting students and short-term visitors .

```
* Omits students & short-term visitors .
USE ALL.
COMPUTE filter_$=(popbase1<2).
VARIABLE LABELS filter_$ 'popbase1<2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
```

This is a filter to select just females who are over age 55 which uses two variables in the logic

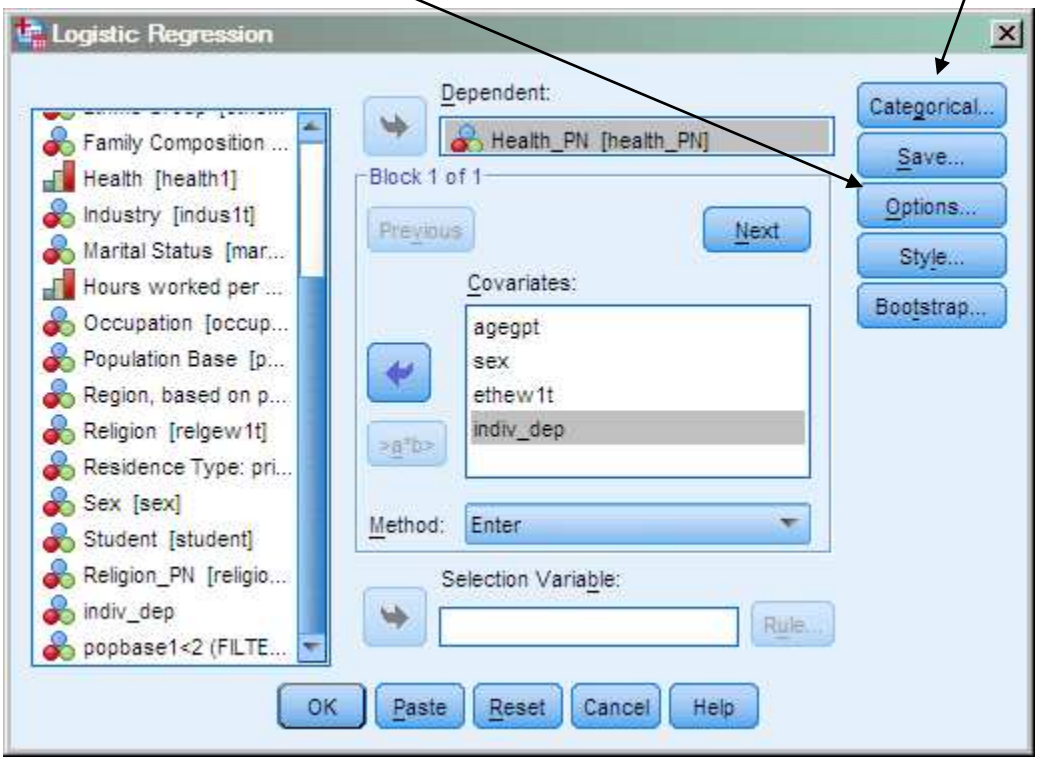
```
* Selects Females aged 55+ .
USE ALL.
COMPUTE filter_$=( sex=2 AND agegpt>6).
VARIABLE LABELS filter_$ 'agegpt>6 AND sex=2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Binary logistic regression using microdata

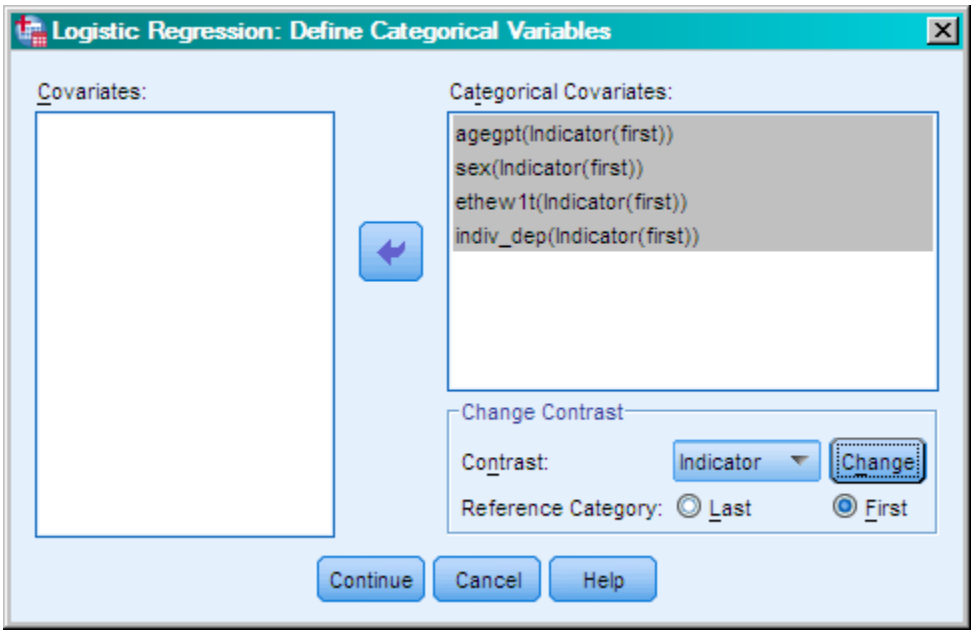
As statistical techniques, linear and logistic regression both model the relationship between an outcome / response / dependent variable and one or more predictor / explanatory variables. The fundamental difference between the models is that linear regression is used with continuous response variables and logistic regression with categorical response variables. Since applying a conventional regression model to predict proportions outside the range 0-1 may be illogical, a binary logistic regression is applicable where the dependent variable has two possible values (e.g. yes/no; presence/absence). Binary logistic regression is also appropriate where the dependent variable is whether or not an individual reported poor health..

Having acquainted ourselves with our dataset using frequencies and cross-tabulations, the logistic regression modelling can begin. Rather than the five level general health categories, this has been dichotomised to 0 Very good to fair health and 1 Bad / Very Bad health.

Use the SPSS menu route: Analyse > Regression > Binary logistic …
Select Health_PN as the Dependent variable age, sex ethnic group and individual deprivation as the 'Covariates'. In the Options, select 'CI for exp (B)' & click Continue. Click on Categorical & a sub-dialogue appears.



When you click on Categorical as suggested above, the following dialogue box appears. The variables selected will be listed in the left-hand panel. Select and move them to the right (you would not do this with continuous variables). With the variables selected on the right, click on First and then Change. The first category of each variable is now the 'reference' or 'base' level with which the other categories are compared.



Click on Continue to take you back to the Logistic Regression dialogue. Now click on OK to run the model. The results will appear in the Output window. Scroll down to the final output table entitled 'Variables in the Equation'.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	agegpt			21767.285	7	.000			
	agegpt(1)	.328	.054	37.446	1	.000	1.388	1.250	1.542
	agegpt(2)	.977	.046	443.439	1	.000	2.656	2.425	2.909
	agegpt(3)	1.660	.042	1531.964	1	.000	5.261	4.841	5.717
	agegpt(4)	2.235	.041	2999.461	1	.000	9.345	8.627	10.123
	agegpt(5)	2.703	.040	4495.797	1	.000	14.929	13.795	16.157
	agegpt(6)	2.954	.041	5298.681	1	.000	19.192	17.724	20.781
	agegpt(7)	3.618	.040	8206.801	1	.000	37.265	34.459	40.299
	sex(1)	.011	.012	.898	1	.343	1.011	.988	1.036
	ethew1t			196.939	4	.000			
	ethew1t(1)	.187	.058	10.501	1	.001	1.206	1.077	1.351
	ethew1t(2)	.307	.026	143.766	1	.000	1.359	1.293	1.429
	ethew1t(3)	.084	.039	4.727	1	.030	1.087	1.008	1.173
	ethew1t(4)	.453	.064	50.521	1	.000	1.573	1.388	1.782
	indiv_dep			1929.561	3	.000			
	indiv_dep(1)	.474	.013	1336.702	1	.000	1.606	1.566	1.647
	indiv_dep(2)	.653	.021	999.061	1	.000	1.921	1.845	2.001
	indiv_dep(3)	.711	.048	221.364	1	.000	2.036	1.854	2.236
	Constant	-5.204	.039	17922.653	1	.000	.005		

Log odds
Positive: outcome more likely
Negative: outcome less likely

Odds ratio
>1: outcome more likely
<1: outcome less likely

Significant?
p<0.05: variable significant
95% CIs not inc 1: variable significant

The output above is less than helpful for what each category represents so look back up the output window to the listings in other outputs or use the Variable window.

It is straightforward to see the health gradient with increasing age and with increasing individual deprivation. You have to imagine that males are reported since where there are only two categories, that output only has the contrasting category. As it happens, there is no difference in odds for females compared with males (the reference). Ethnic groups are compared with the reference category, White ethnicity.

Try adding in Country of Birth to see whether there is any difference for UK and not-UK born. Similarly, try Region.

Summary

ONS have devised a 2011 Census Microdata Teaching File. This document has noted the source of the file and below provides SPSS syntax through which to read the original csv file format and to revise the variable definitions to improve their utility (in the context here). It is hope this document provides a useful resource for people to know about the SARs and to learn to use census microdata.

SPSS syntax 1: To read in 2011 Census Microdata Teaching File, define and label variables

```
* Jo Wathan (jo.wathan@manchester.ac.uk) read in for 2011 microdata teaching file.
*******************************************************************************.
**** 2011 Census Microdata Teaching File for England and Wales *************.
**** SPSS read in file ******************************************************.
*******************************************************************************.
* last updated 13/08/14.
* The data are a copy of the census microdata teaching data file for england and wales
produced by the Office for National Statistics
* The data source is http://www.ons.gov.uk/ons/guide-method/census/2011/census-
data/2011-census-prospectus/release-plans-for-2011-census-statistics/subsequent-
releases-of-specialist-products/microdata/index.html
* Crown copyright.

*read in data from folder.
GET DATA
  /TYPE=TXT
  /FILE="2011 Census Microdata Teaching File.csv"
  /DELCASE=LINE
  /DELIMITERS=","
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=3
  /IMPORTCASE=ALL
  /VARIABLES=
  v1 F7.0
  V2 A9
  V3 A1
  V4 F2.0
  V5 F1.0
  V6 F1.0
  V7 F1.0
  V8 F1.0
  V9 F1.0
  V10 F2.0
  V11 F2.0
  V12 F2.0
  V13 F2.0
  V14 F2.0
  V15 F2.0
  V16 F2.0
  V17 F2.0
  V18 F2.0.
CACHE.
EXECUTE.
DATASET NAME DataSet2 WINDOW=FRONT.

*start by turning string variables into numerics.
if v2 = 'E12000001'region=1.
```

```
if v2 = 'E12000002'region=2.
if v2 = 'E12000003'region=3.
if v2 = 'E12000004'region=4.
if v2 = 'E12000005'region=5.
if v2 = 'E12000006'region=6.
if v2 = 'E12000007'region=7.
if v2 = 'E12000008'region=8.
if v2 = 'E12000009'region=9.
if v2 = 'W92000004'region=10.
execute.
if v3= 'C'restype=1.
if v3= 'H'restype=2.
execute.

*check that worked OK.
crosstabs v2 by region /missing=INCLUDE.
crosstabs v3 by restype /missing = include.

*drop the original two string variables.
DELETE VARIABLES v2 v3.

*tidy up the format of the new variables.
formats region (f2.0).
formats restype (f1.0).
execute.


*rename remaining variables.
RENAME VARIABLE
   (v1 =persid)
   (v4=famcomp)
   (v5=popbase1)
   (v6=sex)
   (v7=agegpt)
   (v8=marstat1t)
   (v9=student)
   (v10=cofbt)
   (v11=health1)
   (v12 =ethew1t)
   (v13=relgew1t)
   (v14=econac1t)
   (v15=occupat1t)
   (v16=indus1t)
   (v17=numhrs)
   (v18=aprxsocgr).

variable level region restype persid famcomp popbase1 sex marstat1t student health1
cofbt occupat1t ethew1t relgew1t econac1t indus1t (nominal)
   / agegpt health1 numhrs aprxsocgr (ordinal).

*label all the variables*.
variable labels persid 'Person ID '.
variable labels  region 'Region, based on previous GOR '.
variable labels restype 'Residence Type: private or communal '.
variable labels famcomp 'Family Composition '.
variable labels popbase1 'Population Base '.
variable labels sex 'Sex '.
variable labels agegpt 'Age Group '.
variable labels marstat1t 'Marital Status '.
variable labels student 'Student '.
variable labels cofbt 'Country of Birth '.
variable labels health1 'Health '.
variable labels ethew1t 'Ethnic Group '.
```

```
variable labels relgew1t 'Religion '.
variable labels econac1t 'Economic Activity '.
variable labels occupat1t 'Occupation '.
variable labels indus1t  'Industry '.
variable labels numhrs 'Hours worked per week '.
variable labels aprxsocgr 'Approximated Social Grade '.

*label all the values*.

value labels region
1 'North East '
2 'North West '
3 'Yorkshire and the Humber '
4 'East Midlands '
5 'West Midlands '
6 'East of England '
7 'London '
8 'South East '
9 'South West '
10 'Wales '.

value labels restype
1 'C Resident in a communal establishment '
2 'H Not resident in a communal establishment '.

value labels famcomp
1 ' Not in a family '
2 ' Married/same-sex civil partnership couple family '
3 ' Cohabiting couple family '
4 ' Lone parent family (male head '
5 ' Lone parent family (female head '
6 ' Other related family '
-9 ' No code required (Not resident in private hhd) '.

value labels popbase1
1 ' Usual resident '
2 ' Student living away from home during term-time '
3 ' Short-term resident '.

value labels sex
1 ' Male '
2 ' Female '.

value labels agegpt
1 ' 0 to 15 '
2 ' 16 to 24 '
3 ' 25 to 34 '
4 ' 35 to 44 '
5 ' 45 to 54 '
6 ' 55 to 64 '
7 ' 65 to 74 '
8 ' 75 and over '.

value labels marstat1t
1 ' Single (never married/registered a same-sex civil partnership '
2 ' Married/registered same-sex civil partnership '
3 ' Separated but still legally married/same-sex civil partnership '
4 ' Divorced or same-sex civil partnership legally dissolved '
5 ' Widowed or surviving partner from a same-sex civil partnership '.

value labels student
1 ' Yes '
2 ' No '.
```

```
value labels cofbt
1 ' UK '
2 ' Non UK '
-9 ' No Code required (student/pupil living away) '.

value labels health1
1 ' Very good health '
2 ' Good health '
3 ' Fair health '
4 ' Bad health '
5 ' Very bad health '
-9 ' No code required (student/pupil living away) '.

value labels ethew1t
1 ' White '
2 ' Mixed '
3 ' Asian and Asian British '
4 ' Black or Black British '
5 ' Chinese or Other ethnic group '
-9 'No code required (student/pupil living away /not resident in EW)'.

value labels relgew1t
1 ' No religion '
2 ' Christian '
3 ' Buddhist '
4 ' Hindu '
5 ' Jewish '
6 ' Muslim '
7 ' Sikh '
8 ' Other religion '
9 ' Not stated '
-9 ' No code required (student/pupil living away /not resident in EW)'.

value labels econac1t
1 ' Economically active: Employee '
2 ' Economically active: Self-employed '
3 ' Economically active: Unemployed '
4 ' Economically active: Full-time student '
5 ' Economically inactive: Retired '
6 ' Economically inactive: Student '
7 ' Economically inactive: Looking after home or family '
8 ' Economically inactive: Long-term sick or disabled '
9 ' Economically inactive: Other '
-9 ' No code required (student/pupil living away or under 16yo)'.

value labels occupat1t
1 ' Managers, Directors and Senior Officials '
2 ' Professional Occupations '
3 ' Associate Professional and Technical Occupations '
4 ' Administrative and Secretarial Occupations '
5 ' Skilled Trades Occupations '
6 ' Caring, Leisure and Other Service Occupations '
7 ' Sales and Customer Service Occupations '
8 ' Process, Plant and Machine Operatives '
9 ' Elementary Occupations '
-9 ' No code required  (student/pupil living away /not resident in EW)'.

value labels indus1t
1 ' Agriculture, forestry and fishing '
2 ' Mining/quarrying; Manufac; Elec/gas, steam/conditioning; Water supply '
3 ' Construction '
4 ' Wholesale & retail trade; Repair of motor vehicles and motorcycles '
```

```
5 ' Accommodation and food service activities '
6 ' Transport and storage; Information and communication '
7 ' Financial and insurance activities; Intermediation '
8 ' Real estate; Professional, sci/technical ; Admin and support service activities '
9 ' Public administration and defence; compulsory social security '
10 ' Education '
11 ' Human health and social work activities '
12 ' Other community/social/personal service; priv hhd domestic staff; Extra-
territorial orgns/bodies '
-9 ' No code required  (student/pupil living away /never worked/ not resident in EW)'.

value labels numhrs
1 ' Part-time: 15 or less hours worked '
2 ' Part-time: 16 to 30 hours worked '
3 ' Full-time: 31 to 48 hours worked '
4 ' Full-time: 49 or more hours worked '
-9 ' No code required (under 16yo, not working, student/pupil living away) '.

value labels aprxsocgr
1 ' AB '
2 ' C1 '
3 ' C2 '
4 ' DE '
-9 'No code required (under 16yo, resident in C.Est, student/pupil living away) '.

sort variables by name (a).
* commands for tabulations to check the data were OK have been removed.
*set the -9 values to missing.

missing values famcomp  cofbt health1 ethew1t relgew1t econac1t occupat1t Indus1t
numhrs aprxsocgr (-9).

frequencies famcomp  cofbt health1 ethew1t relgew1t econac1t occupat1t indus1t numhrs
aprxsocgr (-9).

variable level agegpt to aprxsocgr (ordinal)
   /cofbt to famcomp (nominal)
   /health1 (ordinal)
   /indus1t to marstat1t (nominal)
   /numhrs (ordinal)
   /occupat1t (nominal)
   /persid (scale)
   /popbase1 to student (nominal).
execute.

save outfile "Censusmicroteaching11ew.sav"
```

Syntax 2: Customise 2011 Census Microdata Teaching File variables

```
* Paul Norman (p.d.norman@leeds.ac.uk) revise variables in the original teaching file.
*******************************************************************************.
**** 2011 Census Microdata Teaching File for England and Wales *************.
*******************************************************************************.
* last updated 23/11/14.
* The data are a copy of the census microdata teaching data file for england and wales
produced by the Office for National Statistics
* The data source is http://www.ons.gov.uk/ons/guide-method/census/2011/census-
data/2011-census-prospectus/release-plans-for-2011-census-statistics/subsequent-
releases-of-specialist-products/microdata/index.html
* Crown copyright.

* With Censusmicroteaching11ew.sav open in SPSS .
```

```
* Simplified health variable .
compute health_PN = 0.
if (health1 > 3) health_PN = 1.
variable labels health_PN 'Health_PN '.
value labels health_PN
0 ' Very good to fair health '
1 ' Bad / Very Bad health ' .
execute.

* Recode Religion .
compute religion_PN = 7 .
if (relgew1t = 2) religion_PN = 1.
if (relgew1t = 3) religion_PN = 2.
if (relgew1t = 4) religion_PN = 3.
if (relgew1t = 5) religion_PN = 4.
if (relgew1t = 6) religion_PN = 5.
if (relgew1t = 7) religion_PN = 6.
variable labels religion_PN 'Religion_PN '.
value labels religion_PN
1 'Christian'
2 'Buddhist'
3 'Hindu'
4 'Jewish'
5 'Muslim'
6 'Sikh'
7 'Other, No or Not Stated religion' .
execute.

* Calculate individual deprivation .
compute dep1 = 0.
if (econac1t = 3) dep1 = 1.
compute dep2 = 0.
if (famcomp = 4 or famcomp = 5) dep2 = 1.
compute dep3 = 0.
if (marstat1t = 3 or marstat1t = 4) dep3 = 1.
compute dep4 = 0.
if (numhrs = 1) dep4 = 1.
compute dep5 = 0.
if (occupat1t = 8 or occupat1t = 9) dep5 = 1.
execute.
compute indiv_dep = dep1 + dep2 + dep3 + dep4 + dep5.
execute .
* To have 4 levels .
recode indiv_dep (4=3) (5=3).
execute .
* The above re: .
* unemployment, single parents, separated / divorced .
* few working hours, unrewarding occupation .

save outfile "Censusmicroteaching11ew-PN.sav"
```

References

Boyle P, Norman P & Rees P (2002) Does migration exaggerate the relationship between deprivation and limiting long-term illness? A Scottish analysis. *Social Science & Medicine* 55: 21-31

Norman P & Boyle P (2010) Using Migration Microdata from the Samples of Anonymised Records and the Longitudinal Studies. In *Technologies for Migration and Population Analysis: Spatial Interaction Data Applications* (eds.) John Stillwell, Adam Dennett. IGI Global: Hershey, New York: 133-151.

Norman P & Purdam K (2013) Unpaid caring within and outside the household in England and Wales. *Population, Space & Place* 19: 15-31 DOI: 10.1002/psp.1702

Norman P (2008) Country of birth and country of residence influences on self-reported health: a British analysis using individual-level data. Working Paper 08/05, School of Geography, University of Leeds, Leeds http://eprints.whiterose.ac.uk/76681/

Norman P, Rees P & Wohland P (2014) The use of a new indirect method to estimate ethnic-group fertility rates for subnational projections for England. *Population Studies: A Journal of Demography* 68(1): 43-64 DOI: 10.1080/00324728.2013.810300

Norman P, Wathan J & Higgins V (2008) Introduction to SPSS: a guide for Social Scientists. ESDS government