Growing Up in Scotland study – Data Workshop – Dec 2011/Jan 2012

Workshop Two – Practical exercises

This document provides worked examples of some very basic commands which can be used to explore and analyse the GUS data.

Exercise 1: Weighted frequencies

Frequencies are a very quick and simple way to obtain a descriptive overview of single or multiple variables allowing an assessment of the distribution of characteristics or responses across the population.

Say you want to find out what proportion of 10 month old children live in single parent families. To obtain the answer to this question simply requires a frequency to be run on the sweep 1 birth cohort family type variable – *dahgnp04*

As only sweep 1 data is being considered, the correct weight to use is the sweep 1 birth cohort weight – *dawtbrth*

The syntax command is:

Weight by dawtbrth.
fre dahgnp04.
exe.

This should produce the following result (on the reduced workshop dataset):

(D) Respondent is living with spouse/partner (SW1)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Respondent is not living with spouse/partner	325	20.6	20.6	20.6
	Respondent is living with spouse/partner	1255	79.4	79.4	100.0
	Total	1580	100.0	100.0	

Note that in syntax "frequency" can be abbreviated to both "fre" and "freq", and "execute" can be abbreviated to "exe".

See if you can find the variables and write out the command to find the answers to the following questions:

1. What proportion of children aged 34 months lived in an area classed as 'large urban'?
2. What proportion of mothers of children aged 4-5 years were employed full-time?
3. Has the general health of children aged 3-4 years changed between 2007 and 2009? (Tip: general health variable is named MaHgen01 at sweep 1)

Exercise 2: Weighted crosstabs

Crosstabs are another quick and simple way to get descriptive results from the data. Crosstabs permit the comparison of responses, circumstances, characteristics or outcomes amongst different children or families. The different groups are defined according to their value or response on an independent variable.

For example, you may want to compare the annual household income according to maternal age across families of children aged just under 5 years old. To get this information requires a crosstab to be run on the sweep 3 child cohort data - or sweep 5 birth cohort data - incorporating household income (dc/dewinc01) and maternal age (dc/dehgmag3). For this example we are using the child cohort data.

As only sweep 3 data is being considered, the correct weight to use is the sweep 3 child cohort cross-sectional weight – *dcwtchld*.

The syntax command is:
weight by dcwtchld.
cross dchgmag3 by dcwinc01
/cells = count row
/count = truncate cell.
exe.

This should produce the following result:

DcHGmag3 Dc Age of childs natural mother at interview (banded) * DcWinc01 Dc Household income - banded

Crosstabulation

| | | | DcWinc01 Dc Household income - banded | | | | |
			1 Up to £14999 per year	2 From £15000 to £25999 per year	3 From £26000 to £43999	4 £44,000 and above	Total
DcHGmag3 Dc Age of childs natural mother at interview (banded)	2 20 to 29	Count	68	54	22	12	156
		% within DcHGmag3 Dc Age of childs natural mother at interview (banded)	43.6%	34.6%	14.1%	7.7%	100.0%
	3 30 to 39	Count	47	82	108	97	334
		% within DcHGmag3 Dc Age of childs natural mother at interview (banded)	14.1%	24.6%	32.3%	29.0%	100.0%
	4 40 or	Count	13	25	30	57	125

	% within DcHGmag3 Dc Age of childs natural mother at interview (banded)	10.4%	20.0%	24.0%	45.6%	100.0%
Total	Count	128	161	160	166	615
	% within DcHGmag3 Dc Age of childs natural mother at interview (banded)	20.8%	26.2%	26.0%	27.0%	100.0%

Try finding the correct variables and writing out the commands to find the answers to the following questions:

1. How does car ownership vary amongst families of 34 month-old children who live in areas with different levels of deprivation?
2. To what extent does the weather in Scotland affect how often 4 year old children play outdoors? (Hint: look at how playing outside varies by month/quarter of interview)
3. What proportion of families who use non-parental childcare when their child is aged 10 months is still doing so when the child is aged 34 months?