



Growing Up in Scotland: Data Workshops 2014

Handout Pack

Authors: Paul Bradshaw and Joan Corbett Date: January 2014







1	Accessing the GUS data at UKDS	2
2	Useful indicators and outcomes measures	5
2.1	Standard analysis variables	5
	2.1.1 Individual or household level parent and child socio-economic and demographic	_
	Variables	5 6
2.2	Indicators of child outcomes	0 6
2.3	Cognitive development and ability	7
2.4	Social, emotional and behavioural development	7
2.5	Physical health and development	8
3	Data Management in SPSS	.10
3.1	Practical tips on SPSS set-up	10
	3.1.1 Displaying variable names and commands in analysis output	10
3.2	Examples of frequencies & cross-tabs via menu	12
	3.2.1 Simple trequency	12 13
3.3	Recoding variables via syntax	16
3.4	Computing a Derived Variable using syntax	17
4	How to Create and Merge Datasets in SPSS	.19
4.1	Creating bespoke datasets using the 'Keep' and 'drop' Commands	19
	4.1.1 The KEEP command	19
4.0	4.1.2 The DROP command	19
4.2	Merging Datasets	20
5	Data Management - Additional tips	.24
5.1	To change the order of variable names in dialogue boxes	24
5.2 5.3	COMPLITE a variable via menu	25 28
6	Coning with Complex Samples	30
61	Creating a complex samples plan file in SPSS	30
6.2	Basic analysis using the Complex Samples (CS) Module in SPSS	34
7	Workshop Two – Practical exercises	.38
7.1	Exercise 1: Weighted Frequencies	38
7.2	Exercise 2: Weighted Crosstabs	39
8	Workshop Two – Answers to the practical exercises	.41
8.1	Exercise 1: Weighted Frequencies	41
8.2	Exercise 2: Weighted Crosstabs	44

1 Accessing the GUS data at UKDS

a) On the UKDS home page, type GUS into the 'Search the Data Catalogue' box on the left-hand side and click on 'GO'



🕑 🗢 🙋 http:	://discover.ukdataservice.ac.uk/?q:	=GUS&sf=Data+Cata	logue%7CSeries	&searchType=dat	a		🖄 🖆	VKDS		
Edit View F	Favorites Tools Help									
Favorites 🛛 👍 🉋	🕑 Free Hotmail 🙋 Web Slice Galle	ery 🔻								
🔹 🏉 UK Data Se	ervice Disc 🗙 <mark>8</mark> Google						6	• 🗟 • 🖻 🖶 • P	age 🔻 Safety 🕯	🔹 Tools 🔻 🕜
							:	Site Search FAQ Help (Contact	
	UK Data Service	About us	Get data	Use data	Manage data	Deposit data	News	and Events		
		Discover								
		Discover								
		Discover								
	Discover									
	Variable and question	Search and brow	wse our data col	lections, support	guides, case studie	s, and related publ	ications.			
	Туре 😑	Q GUS					GO			
	Data collections (5)				I		1			
	Series (1)			Reset filters	Clear search	Auto-complete	Help			
	Case studies (1)									
	ESRC outputs (11)	Case study	C Data coll	ection 📑 Cer	isus data (S) Se	eries record	ESRC output		guide	
	Support / how to						Sorted b	y: Relevance	-	
	guides (0)	SEARCHR	ESULIS							
	Refine	Displaying 1-6 of	6 results						1	
	Subject +	SN 7432	Growing Up in Sc	otland: Cohort 2. S	weep 1, 2011					
	Date +	Scottish C	Centre for Social F	Research						
		+ Full record	d							
					₩.	Download/Order	Get full DDI XM	L Similar data collect	tions	
		C SN 7145 Scottish C	Growing Up in Sc Centre for Social F	otland: Sweep 5 Po Research	ostcodes, 2009-2010	: Secure Access				
		+ Full record	d							
					₩.	Download/Order	Get full DDI XM	L Similar data collect	tions	
		C SN 5760 Scottish C	Growing Up in Sc Centre for Social F	otland: Cohort 1, S Research	weeps 1-6, 2005-201	1			*****	
		+ Full record	d							
									6	👻 🕀 100%

c) This will take you to a screen listing datasets which have 'GUS' in their description.

d) You can then choose to look at the Full record or Download/Order the datasets. **Anyone** who is browsing the UKDS site **has access to the description and documentation** of datasets in order to help them decide whether a dataset would be of use in their analysis.



e) The various documentation files (including a list of the variables) can be downloaded as pdf files. The study information gives details of copyright and how to acknowledge and cite the data in publications

f) Downloading the datasets requires to be registered with UKDS. Information about how to register can be found on the UKDS site

http://ukdataservice.ac.uk/get-data/how-to-access/

2 Useful indicators and outcomes measures

This section provides some information on some of the variables that have been more commonly used in GUS analysis. They include key measures of demographic, household and socio-economic characteristics, as well as indicators of child outcomes.

2.1 Standard analysis variables

Note that generally speaking, the 'standard' analysis variables are available on ALL datasets. As a result, the sweep prefix (e.g. 'a' for sweep 1, 'b' for sweep 2 etc) has been removed in the table below and replaced with an asterisk.

2.1.1 Individual or household level parent and child socio-economic and demographic variables

Varia	ble name	Deta	I I						
D*HG	inkd2	Num	per of children in the household						
D*HG	inmad	Num	per of adults in the household						
D*HG	Snmk2	Num	Imber of children in household - Banded						
D*HG	Sprim	Whet	nether child was mothers first-born						
D*HG	Grsp05	Whet	her respondent is childs mother (including adoptive, foster and						
		step-	mothers)						
D*HG	Grsp06	Whet	her respondent is childs father (including adoptive, foster and step-						
		fathe	rs)						
D*HG	Grsp07	Who	is the respondent in relation to the child?						
D*HG	Frsp08	Who	is the respondents partner in relation to the child?						
D*HG	mag2	Age o	of mother at time of interview						
D*HG	imag3	Age o	of mother at interview (banded)						
D*HG	imag4	Age o	of mother at birth of sample child (years)						
D*HG	imag5	Age o	of mother at birth of sample child (banded)						
D*Me	th07	Ethni	city of Respondent						
D*Ye	th07	Ethni	city of Partner						
D*Ms	ta01	Resp	ondent's employment status						
D*Ms	ta02	Resp	ondent's employment status: ver2						
D*Ys	ta01	Partn	er's employment status						
D*Ys	ta02	Partn	er's employment status: ver2						
D*Ms	ta10	Hous	ehold employment: ver1						
D*Ms	ta11	Hous	ehold employment and family type						
D*Ws	sta02	Moth	er's employment status						
D*Ws	sta03	Moth	er's employment status						
D*Me	du01	Highest Education level of Respondent							
D*Me	du02	Highest Education level of Respondent - Banded							
D*Me	du04	Highe	est Education level of Respondent – Banded SCQF						
D*Ye	du01	Highe	est Education level of Partner						
D*Ye	du02	Highe	est Education level of Partner - Banded						
	D*Medu04		Highest Education level of Respondent – Banded SCQF						
	MfWinc09		Household income – full version						

Household income - banded
Equivalised household income – banded (quintiles)
Respondent NSSEC - 5 Category
Respondent NSSEC - 5 Category
Household NSSEC - 5 Category (NS-SEC is derived each sweep for the
main respondent and his or her resident partner. Where there are two
such classifications, 'Household NS-SEC' takes the higher classifications
as its reference category)
Derived tenure at sweep 6 - full
-

2.1.2 Area-level demographic variables

AL*URin1	ALe - Urban-rural classification (Scotland)
AL*URin2	ALe - SG urban-rural classification
AL*SNim2	ALe - SIMD 2009 quintiles
AL*Low15	ALe - Flag lowest 15% datazones
AL*Carst	ALe - Deciles of Carstairs scores
AL*HBdBc	ALe - Health Boards aggreg as Sw1 (Birth Cohort)

Variations on area-level variable names at sweep 3 only:

DcURind1	Urban-rural classification (Scotland)
DcURind2	SG urban-rural classification
DcADsco2 *	SIMD 2006 quintiles
DcADsco3	Flag lowest 15% datazones (SIMD 2006)

2.2 Indicators of child outcomes

With the cohort children at the centre of GUS, much of the analysis is often focused on determining, in specific contexts, what factors contribute to or are associated with more positive or negative child outcomes. This type of analysis presents a challenge in that the distinction between input and outcome variables is not often simple or easily defined. Outcomes are often closely interlinked and any single outcome can influence others and may well be seen as an independent variable in a particular analytical context. Furthermore, some factors which may normally be classified as independent variables may also be considered as outcomes in certain contexts. Certain independent variables may also sometimes be treated as mediators or moderators.

Thus, the classification of variables as a specific type often depends on the perspective adopted for a particular analysis rather than on defined objective distinctions. A child outcome as measured at one wave of a longitudinal study may itself become an input variable in a subsequent wave. For example, a number of cohort studies have shown that cognitive ability at a previous wave – a measure widely defined as a key outcome variable – is a significant predictor of the same ability at subsequent waves; childhood obesity can be considered an outcome related to diet and physical activity and also a risk factor for subsequent development of diabetes or cardiovascular disease.

GUS collects measures of child outcomes across five domains: cognitive development and ability, social development and behaviour, affective and emotional development, health and physical development, and education. A brief outline of the focus of each of these domains along with examples of the types of experiences, characteristics and information used to measure outcomes within them is provided below. Variable names have been included in the tables. Note that these have the source and sweep prefix removed (e.g. 'Ma') – this should be added according to the sweep where the data was collected (Ma for sweep 1, Mb for sweep 2 etc).

2.3 Cognitive development and ability

Cognition is the study of thought processes or mental activity by which children and adults acquire and deal with knowledge. It is a wide-ranging concept encompassing a number of facets, such as visual perception, memory, language and higher executive functions such as planning and decision making. These developmental achievements are highly interdependent; many of the cognitive functions children acquire depend on a certain degree of maturation within the growing brain before they can be performed adequately, or in some cases, at all. Similarly, children's ability to form words and start to speak is linked to physical changes in the palate and tongue as well as neurological changes within the brain.

Question topics		Sw1		Sv	Sw2 Sw3		Sw4		Sw5	Sw6	
	BC1	BC2	СС	BC1	СС	BC1	CC	BC1	СС	BC1	BC1
Child communication indicators/gestures (reaches out to give you something, waves bye-bye, extends arms to be picked up, nods head for yes), <i>Dbab01</i> , <i>Dbab09 to Dbab12</i>	~	~									
Communication and Symbolic Behaviour Scales - Infant/Toddler Checklist CSBS01 to CSBS24 (Derived score and scale variables are on the sw2 dataset, see sw2 user guide for details)		~									
Child's speech and language (whether child can be understood) Dspe01 to Dspe04			~	~	~	\checkmark		\checkmark		\checkmark	
Concerns about child's development, learning or behaviour Dspe04 to Dspe11; at sw6:Dcon01 to Dcon03		~	~	\checkmark	~	~	~	\checkmark	~	\checkmark	\checkmark
BAS Picture Similarities PicRaw to PicSPT (Derived score and scale variables are on the sw3 and sw5 datasets, see user guide for details)											
BAS Naming Vocabulary NamRaw to NamPT (Derived score and scale variables are on the sw3 and sw5 datasets, see user guide for details)											

2.4 Social, emotional and behavioural development

The study of social, emotional and behavioural development is concerned with a range of issues including how a child behaves generally, and in specific contexts or situations, the nature of their relationships and interactions with others - such as parents, siblings, other family members and peers – their neurological development and mental health. It comprises measurement of a considerable range of general concepts such as self-esteem and confidence, attachment, pro- and anti-social behaviour, peer relationships and temperament, as well as more specific child

psychiatric conditions such as conduct disorder, autism spectrum conditions and attention-deficit hyperactivity disorder.

A child's social, emotional and behavioural development has significant implications for current and later social functioning and for educational and employment success. If emotional development is fostered at a young age, children are more likely to settle well into school, work cooperatively, confidently and independently, and behave appropriately; a child with poor social and emotional development is at risk of poor relationships with peers, academic problems, later involvement in crime and developing physical health and adult mental-health problems.

Key to social and emotional development is the child's early relationship with parents or caregivers. Efforts to support parents in understanding and fulfilling their children's emotional needs can help to provide a secure base from which children grow into well-rounded, capable adults with robust mental health.

Question topics		Sw1		Sw	2	Sw	3	Sw	/4	Sw5	Sw6
	BC1	BC2	CC	BC1	CC	BC1	СС	BC1	CC	BC1	BC1
Goodman's Strength and Difficulties Questionnaire SDQ01 to SDQ25 (derived scale variables available on sw2 – 6 datasets – see user guides for details)			~				~	\checkmark	~	\checkmark	\checkmark
Concerns about child's development, learning or behaviour Dcon01 to Dcon03		~	~	\checkmark	~	\checkmark	~	\checkmark	~	\checkmark	\checkmark
Child and respondent's sleep Dsle01, Dsle02, Dsle03	\checkmark	\checkmark	~			\checkmark					~
Parent-child relationship											
Condon Infant-Maternal Attachment Mfee01 to Mfee06	\checkmark	\checkmark									
Six 'Insecure attachment' items from NLSY 'How My Child Usually Acts' temperament scale Hatt01 to Hatt05 Parent-child relationship (Pianta's scale) Ppia01 to Ppia15						✓				~	
Problems managing child's behaviour and relationships <i>Tsib01, Tbeh01</i>			~		~	\checkmark		\checkmark			

2.5 Physical health and development

Changes in body size are the most obvious manifestation of physical growth. During infancy changes in growth are extremely rapid, thus height, weight and other physical measures such as waist circumference and body mass index are examples of outcome measures within this domain. Beyond growth, key measures of physical development in infancy and childhood are concerned with control of motor behaviour which is central to physical development amongst infants. The study of acquisition of motor skills usefully distinguishes between gross motor development – that is, motor skills which help children to explore their environment such as crawling and walking – and fine motor development – skills which cover smaller movement sequences like reaching and grasping. Beyond infancy, motor skills are concerned with locomotor movements, which include walking, running, jumping, hopping, skipping and climbing; manipulative

movements, including throwing, catching and kicking; and stability movements which include bending, stretching and balancing.

Aside from physical development, this domain is also concerned with measures of physical health. This includes experience of short and longer term health problems, illnesses, conditions and disabilities. Physical health can also be measured through accidental injury, hospital admissions and use of other health services.

Question topics		Sw1		Sw	2	Sw	3	Sw	/4	Sw5	Sw6
	BC1	BC2	CC	BC1	СС	BC1	CC	BC1	СС	BC1	BC1
Child's general health <i>Hgen01</i>	\checkmark	\checkmark	\checkmark	\checkmark	~	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Child health problems (not necessarily involving medical intervention) (number and type) <i>Hprb02 – Hprb95</i>	\checkmark	~	~	\checkmark	~	\checkmark	~	\checkmark	~	\checkmark	\checkmark
Child accidents that have involved medical intervention <i>Haca01 – Hace95</i>	\checkmark	~	~	\checkmark	~	\checkmark	~	\checkmark	✓	\checkmark	\checkmark
Child's long-standing illness or disability (existence, nature, whether limiting) <i>Hlsi01/Hlsa01/Hlsb01/Hlsc01</i>	\checkmark	~	~	\checkmark	~	\checkmark	~	\checkmark	~	\checkmark	\checkmark
Height, Weight and BMI A range of BMI measurements are available on the dataset. See user guides at sw4 and sw6 for more details on variables					~			\checkmark	~		V
Child gross & motor development milestones Dbab02 to Dbab08, Dtod01 – Dtod14	~	~	~			\checkmark					

3 Data Management in SPSS

3.1 Practical tips on SPSS set-up

3.1.1 Displaying variable names and commands in analysis output

🖬 çəss	W3B_30.sav	/ [DataSet2] - S	PSS Data	Editor						×
File Edit	View Data	Transform Analy	/ze Graph	s Utilities Win	dow Help					
	🖽 🔶	🔿 📥 🕼	A 情日	1 🖩 🤹 🖡	, 😵 🔕 🌢					
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	^
1	ldnumber	Numeric	8	2	Serial number	None	None	8	Right	-
2	SampType	Numeric	2	0	Sample Type	{1, Birth}	-91	8	Right	1
3	DcXmnth1	Numeric	2	0	Dc Month of in	{1, January}	-91	10	Right	1
4	DcXqurt1	Numeric	2	0	Dc Quarter of i	{1, January to	-91	10	Right	
5	Outcome	Numeric	3	0	Dc Final outco	{110, Full inter	-91	8	Right	
6	McHGrsp03	Numeric	2	0	Mc - Same res	{-9, Refusal}	-91	8	Right	
7	DcHGrsp01	Numeric	2	0	Dc - Whether r	{0, Not mentio	-91	8	Right	
8	DcHGrsp02	Numeric	2	0	Dc - Whether r	{0, Not mentio	-91	8	Right	
9	DcHGnmad	Numeric	2	0	Dc Number of	None	-91	10	Right	
10	DcHGnmad	Numeric	2	0	Dc Number of	{0, None}	-91	11	Right	
11	DcHGnmkd	Numeric	2	0	Dc Number of	None	-91	10	Right	
12	DcHGnmk2	Numeric	2	0	Dc Number of	{1, One}	-91	10	Right	
13	DcHGhsiz	Numeric	2	0	Dc Household	None	-91	10	Right	
14	DcHGnmsb	Numeric	2	0	Dc - Number of	None	-91	8	Right	
15	DcHGnp01	Numeric	2	0	Dc - Number of	{0, Not mentio	-91	8	Right	
16	DcHGnp02	Numeric	2	0	Dc - Natural m	{0, Not mentio	-91	8	Right	
17	DcHGnp03	Numeric	2	0	Dc - Natural fat	{0, Not mentio	-91	8	Right	
18	DcHGnp04	Numeric	2	0	Dc - Responde	(O, Responden	-91	8	Right	
19	DcHGrsp04	Numeric	2	0	Dc - Family Ty	{0, Lone Paren	-91	11	Right	
20	DcHGrsp05	Numeric	2	0	Dc Whether re	{0, No}	-91	11	Right	
a di Ni	ata View) Va	riable View /	-	-		(C) (1)			>	
			S	PSS Processor is	ready					:

In the menu click on 'Edit / Options':

The 'Options' dialog box opens at the 'General' tab. Select the option 'Display names' in the 'Variable Lists':

Data	Currer	юу	Sci	ripts
General Viewer Draft V	/iewer Output La	abels Chart:	s Interactive	Pivot Table
Variable Lists Display labels Dis Alphabetical Eile Session Journal Record syntax in Journal Append Ov F:\SPSSJ0~1.JNL Bro	play names	Jutput No scientific in tables /iewer Type at S (Regular Measurement Sy Language:	notation for small nu itartup: Draft stem: Points English	mbers
Temporary directory: C:\DOCUME~1\MIREIL~1\Lt Recently used file list: Ø Open syntax window at start-	DCALS~1\T	Notification: Raise viewe Scroll to new Sound: Sound: Sound: Sound: Sound:	r window v output one O System t pund Browse	beep

In the 'Viewer' tab tick 'Display commands in the log' at the bottom left. This means that if you run some syntax, it will be displayed in the output screen which opens up automatically when a dataset is opened. This helps spot any errors in the syntax you write.

Options			X
Data	Currency	Scripts	
General Viewer Draft Vie	ewer Output Labels Charl	ts Interactive P	ivot Tables
Initial Output State Item: Log Contents are initially: Shown Hidden Justification: Align left Center Align right Display commands in the log	Title Font Arial 14 Text Output Page Size Width: Standard (80 characters) Standard (80 characters) Wide (132 characters) Custom: 255 Text Output Font Courier New 10	B I U Length: Standard (59 lines) Infinite Custom: 59 B I U Monospaced for	L L
	OK Cance	el Apply	Help

In the 'Output Labels' tab select the following options in the drop down lists:

Options					Þ	
Data		Currency		Scripts		
General Viewer	Draft Viewer	Output Labels	Charts	Interactive	Pivot Tables	
Outline Labeling Variables in item labels Names and Labels	shown as:					
Variable values in item	labels shown as:	_				
Values and Labels	1	1				
Variables in labels show Names and Labels Variable values in label	vn as: s shown as:	•				
Values and Labels	•	/				
		ОК	Cancel	Apply	Help	

This means that both the variable names, labels and values are shown in any output produced by your analysis.

3.2 Examples of frequencies & cross-tabs via menu

3.2.1 Simple frequency

This example demonstrates how to run a simple frequency which will show the proportion of parents who reported attending a parent and toddler group when the cohort child was aged three. It is based on BC1 Sweep 3 dataset.

ClussW3B_30.sav [DataSci1] - SPSS Data Editor File Edit View Data Transform Analyze Gruphs Utilities Window Help Image: Type With Decimals Label Values Missing Columns Align 1 Idnumber Numeric 8 2 Serial number None 8 Right 2 SampType Numeric 2 0 Sample Type {1, Bith} 9 - 1 8 Right 3 DcXmnth1 Numeric 2 0 Dc Month of in {1, January} 9 - 1 10 Right 4 DcXqurt1 Numeric 2 0 Dc Month of in {1, January} 9 - 1 8 Right 5 Outcome Numeric 2 0 Dc Austronth of in {1, January} 9 - 1 8 Right 6 McHGrsp03 Numeric 2 0 Mc - Same res {9, Refusa} 9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio 9 - 1 8 Right 9 DcHGrsp02 Nume	I) Sele										
File Edt View Data Transform Analyze Griphs Utilities Window Help Image Type Width Decimals Label Values Missing Columns Align 1 Idnumber Numeric 8 2 Serial number None 8 Right 2 SampType Numeric 2 0 Sample Type 1, Birth) 9 - 1 8 Right 3 DcXmnth1 Numeric 2 0 Dc Quarter of i 1, January) 9 - 1 10 Right 4 DcXqurt1 Numeric 2 0 Dc Guarter of i 1, January to 9 - 1 8 Right 5 Outcome Numeric 2 0 Dc Final outco 110, Full inter 9 - 1 8 Right 6 McHGrspD3 Numeric 2 0 Dc - Whether r (0, Not mentio 9 - 1 8 Right 7 DcHGrspD1 Numeric 2 0 Dc - Whether r (0, Not mentio 9 - 1 8 Right 9 DcHGrspD2 Numeric 2 <th>GUSS</th> <th>W3B_30.sav</th> <th>v [Data\$=:1] - \$</th> <th>iPSS Data</th> <th>Editor</th> <th></th> <th></th> <th></th> <th></th> <th></th> <th>×</th>	GUSS	W3B_30.sav	v [Data\$=:1] - \$	iPSS Data	Editor						×
Image Type Width Decimals Label Values Missing Columns Align 1 [dnumber] Numeric 8 2 Serial number None None 8 Right 2 SampType Numeric 2 0 Sample Type 1, Bith) 9 - 1 8 Right 3 DcXmnth1 Numeric 2 0 Dc Month of in {1, January} 9 - 1 10 Right 4 DcXqurt1 Numeric 2 0 Dc Guarter of i {1, January to 9 - 1 10 Right 5 Outcome Numeric 2 0 Dc Final outco (110, Full inter 9 - 1 8 Right 6 McHGrsp03 Numeric 2 0 Dc - Whether r {0, Not mentio 9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio 9 - 1 8 Right 9 DcHGrsp02 Numeric 2 0 Dc Number of None 9 - 1 10 Right 10 DeHG	File Edit	View Data	Transform Anal	yze Gr <mark>u</mark> ph	is Utilities Wir	idow Help					
NameTypeWidthDecimalsLabelValuesMissingColumnsAlign1IdnumberNumeric82Serial numberNone8Right2SampTypeNumeric20Sample Type{1, Birth}918Right3DcXmnth1Numeric20Dc Month of in{1, January}9110Right4DcXqut1Numeric20Dc Cuarter of i{1, January to-9110Right5OutcomeNumeric30Dc Final outco{110, Full inter -918Right6McHGrsp03Numeric20Dc Venther r{0, Not mentio-918Right7DcHGrsp01Numeric20Dc Venther r{0, Not mentio-918Right9DcHGrsp02Numeric20Dc Venther r{0, Not mentio-918Right9DcHGrsp02Numeric20Dc Number ofNone-9110Right10DcHGrsp02Numeric20Dc Number ofNone-9110Right11DcHGrsp04Numeric20Dc Number ofNone-9110Right11DcHGrsp04Numeric20Dc Number ofNone-9110Right12DcHGrsp04Numeric20Dc		A 🖬 🔶		44 唐	1 🖽 🖬 🖬	. 😪 🙆 🌰					
NameTypeWildinDecimalsLaberValuesWildinsColumnsAlign1IdnumberNumeric82Serial numberNone8Right2SampTypeNumeric20Sample Type{1, Birth}918Right3DcXmnth1Numeric20Dc Month of in{1, January}9110Right4DcXqur1Numeric20Dc Quarter of i{1, January to9110Right5OutcomeNumeric30Dc Final outco{10, Full inter-918Right6McHGrsp03Numeric20Dc - Whether r{0, Not mentio-918Right7DcHGrsp01Numeric20Dc - Whether r{0, Not mentio-918Right9DcHGrsp02Numeric20Dc Number ofNone-918Right9DcHGrmadNumeric20Dc Number ofNone-9110Right10DcHGrmadNumeric20Dc Number ofNone-9110Right11DcHGrmadNumeric20Dc Number ofNone-9110Right11DcHGrmadNumeric20Dc Number ofNone-9110Right12DcHGrmadNumeric20Dc Numbe		Nama	Tune	1000446	Desimals	Lahal		Minning	Columna	Alian	~
Idnumber Numeric 8 2 Senal number None None 8 Right 2 SampType Numeric 2 0 Sample Type {1, Bith}, -91 8 Right 3 DcXmnth1 Numeric 2 0 Dc Month of in {1, January} -91 10 Right 4 DcXqut1 Numeric 2 0 Dc Month of in {1, January to -91 10 Right 5 Outcome Numeric 3 0 Dc Final outco {110, Full inter -91 8 Right 6 McHGrsp03 Numeric 2 0 Mc - Same res {-9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio -91 8 Right 9 DcHGrsp02 Numeric 2 0 Dc Number of None -91 10 Right 10 DcHGrmad		Ivame	Туре	vviutri	Decimais	Laber	Values	wissing	Columns	Align	. =
2 Samplype Numeric 2 0 Sample type [1, Birth] -9 - 1 8 Right 3 DcXmnth1 Numeric 2 0 Dc Month of in [1, January] -9 - 1 10 Right 4 DcXqurt1 Numeric 2 0 Dc Quarter of i [1, January to -9 - 1 10 Right 5 Outcome Numeric 3 0 Dc Final outco [10, Full inter -9 - 1 8 Right 6 McHGrsp03 Numeric 2 0 Mc - Same res {-9, Refusal} -9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r (0, Not mentio -9 - 1 8 Right 9 DcHGrmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 10 DcHGrmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 11 <td< td=""><td>1</td><td>ldnumber</td><td>Numeric</td><td>8</td><td>2</td><td>Serial number</td><td>None</td><td>None</td><td>8</td><td>Right</td><td>-</td></td<>	1	ldnumber	Numeric	8	2	Serial number	None	None	8	Right	-
3 DcXmnth1 Numeric 2 0 Dc Month of in {1, January}	2	SampType	Numeric	2	0	Sample Type	{1, Birth}	-91	8	Right	
4 DcXqurt1 Numeric 2 0 Dc Quarter of i {1, January to -91 10 Right 5 Outcome Numeric 3 0 Dc Final outco {110, Full inter -91 8 Right 6 McHGrsp03 Numeric 2 0 Mc - Same res {-9, Refusal} 91 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio -91 8 Right 8 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -91 8 Right 9 DcHGrsp02 Numeric 2 0 Dc Number of None -91 10 Right 9 DcHGrnmad Numeric 2 0 Dc Number of None -91 10 Right 10 DcHGrnmad Numeric 2 0 Dc Number of None -91 10 Right <t< td=""><td>3</td><td>DcXmnth1</td><td>Numeric</td><td>2</td><td>0</td><td>Dc Month of in</td><td>{1, January}</td><td>-91</td><td>10</td><td>Right</td><td></td></t<>	3	DcXmnth1	Numeric	2	0	Dc Month of in	{1, January}	-91	10	Right	
5 Outcome Numeric 3 0 Dc Final outco {110, Full inter -9 - 1 8 Right 6 McHGrsp03 Numeric 2 0 Mc - Same res {-9, Refusal} -9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 8 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 11 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 12 DcHGnmkd Numeric 2 0 Dc Number of None -9 - 1 10 Right	4	DcXqurt1	Numeric	2	0	Dc Quarter of i	{1, January to	-91	10	Right	
6 McHGrsp03 Numeric 2 0 Mc - Same res {-9, Refusal} -9 - 1 8 Right 7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 8 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 10 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 11 Right 11 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 12 DcHGnmk2 Numeric 2 0 Dc Household None -9 - 1 10 Right 13 DcHGnsb Numeric 2	5	Outcome	Numeric	3	0	Dc Final outco	{110, Full inter	-91	8	Right	
7 DcHGrsp01 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 8 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 10 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 10 Right 11 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 10 Right 12 DcHGnmkd Numeric 2 0 Dc Number of {1, One} -9 - 1 10 Right 13 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsiz Numeric 2 0 Dc - Number of {0, Not mentio -9 - 1 8 Right 1	6	McHGrsp03	Numeric	2	0	Mc - Same res	{-9, Refusal}	-91	8	Right	
8 DcHGrsp02 Numeric 2 0 Dc - Whether r {0, Not mentio -9 - 1 8 Right 9 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 10 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 10 Right 11 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 11 Right 12 DcHGnmkd Numeric 2 0 Dc Number of None -9 - 1 10 Right 13 DcHGnkz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsb Numeric 2 0 Dc - Number of None -9 - 1 8 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of Not mentio -9 - 1 8 Right 16 D	7	DcHGrsp01	Numeric	2	0	Dc - Whether r	{0, Not mentio	-91	8	Right	
9 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 10 Right 10 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 11 Right 11 DcHGnmad Numeric 2 0 Dc Number of None -9 - 1 11 Right 12 DcHGnmkd Numeric 2 0 Dc Number of None -9 - 1 10 Right 13 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsiz Numeric 2 0 Dc - Number of None -9 - 1 10 Right 15 DcHGnsb Numeric 2 0 Dc - Number of None -9 - 1 8 Right 16 DcHGnp01 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 17 DcHGnp03 </td <td>8</td> <td>DcHGrsp02</td> <td>Numeric</td> <td>2</td> <td>0</td> <td>Dc - Whether r</td> <td>{0, Not mentio</td> <td>-91</td> <td>8</td> <td>Right</td> <td></td>	8	DcHGrsp02	Numeric	2	0	Dc - Whether r	{0, Not mentio	-91	8	Right	
10 DcHGnmad Numeric 2 0 Dc Number of {0, None} -9 - 1 11 Right 11 DcHGnmkd Numeric 2 0 Dc Number of None -9 - 1 10 Right 12 DcHGnmkd Numeric 2 0 Dc Number of {In One} -9 - 1 10 Right 13 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsb Numeric 2 0 Dc - Number of None -9 - 1 8 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of (0, Not mentio -9 - 1 8 Right 16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 17	9	DcHGnmad	Numeric	2	0	Dc Number of	None	-91	10	Right	
11 DcHGnmkd Numeric 2 0 Dc Number of None -9 - 1 10 Right 12 DcHGnmk2 Numeric 2 0 Dc Number of {1, One} -9 - 1 10 Right 13 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of {0, Not mentio -9 - 1 8 Right 16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 17 DcHGnp03 Numeric 2 0 Dc - Natural fat (0, Not mentio -9 - 1 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responden -9 - 1 8 Right 19	10	DcHGnmad	Numeric	2	0	Dc Number of	{0, None}	-91	11	Right	
12 DcHGnmk2 Numeric 2 0 Dc Number of {1, One} -9 - 1 10 Right 13 DcHGhsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnsbs Numeric 2 0 Dc - Number of None -9 - 1 8 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of (0, Not mentio) -9 - 1 8 Right 16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio) -9 - 1 8 Right 17 DcHGnp03 Numeric 2 0 Dc - Natural m (0, Not mentio) -9 - 1 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responden) -9 - 1 8 Right 19 DcHGrsp05 Numeric 2 0 Dc - Family Ty (0, Lone Paren) -9 - 1 11 Right	11	DcHGnmkd	Numeric	2	0	Dc Number of	None	-91	10	Right	-
13 DcHGhsiz Numeric 2 0 Dc Household None -9 - 1 10 Right 14 DcHGnmsb Numeric 2 0 Dc - Number of None -9 - 1 8 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of (0, Not mentio -9 - 1 8 Right 16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 17 DcHGnp03 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responden -9 - 1 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty (0, Lone Parei -9 - 1 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re (0, No) -9 - 1 11 Right <td>12</td> <td>DcHGnmk2</td> <td>Numeric</td> <td>2</td> <td>0</td> <td>Dc Number of</td> <td>{1, One}</td> <td>-91</td> <td>10</td> <td>Right</td> <td>1</td>	12	DcHGnmk2	Numeric	2	0	Dc Number of	{1, One}	-91	10	Right	1
14 DcHGnmsb Numeric 2 0 Dc - Number of None -9 - 1 8 Right 15 DcHGnp01 Numeric 2 0 Dc - Number of (0, Not mentio -9 - 1 8 Right 16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 17 DcHGnp03 Numeric 2 0 Dc - Natural m (0, Not mentio -9 - 1 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responden -9 - 1 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty (0, Lone Parei -9 - 1 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re (0, No) -9 - 1 11 Right	13	DcHGhsiz	Numeric	2	0	Dc Household	None	-91	10	Right	1
15 DcHGnpD1 Numeric 2 0 Dc - Number of {0, Not mentio -91 8 Right 16 DcHGnpD2 Numeric 2 0 Dc - Natural m {0, Not mentio -91 8 Right 17 DcHGnpD3 Numeric 2 0 Dc - Natural m {0, Not mentio -91 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde {0, Responden -91 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty {0, Lone Paren -9 - 1 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re {0, No) -9 - 1 11 Right	14	DcHGnmsb	Numeric	2	0	Dc - Number of	None	-91	8	Right	1
16 DcHGnp02 Numeric 2 0 Dc - Natural m (0, Not mentio -91 8 Right 17 DcHGnp03 Numeric 2 0 Dc - Natural fat (0, Not mentio -91 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responde - 91) 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty (0, Lone Paren - 91) 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Wether re (0, No) -91 11 Right	15	DcHGnp01	Numeric	2	0	Dc - Number of	{0, Not mentio	-91	8	Right	
17 DcHGnp03 Numeric 2 0 Dc - Natural fat (0, Not mentio -91 8 Right 18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responde -91 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty (0, Lone Paren -91 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re (0, No) -91 11 Right	16	DcHGnp02	Numeric	2	0	Dc - Natural m	{0, Not mentio	-91	8	Right	1
18 DcHGnp04 Numeric 2 0 Dc - Responde (0, Responden -91 8 Right 19 DcHGrsp04 Numeric 2 0 Dc - Family Ty (0, Lone Paren -91 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re (0, No) -91 11 Right	17	DcHGnp03	Numeric	2	0	Dc - Natural fat	{0, Not mentio	-91	8	Right	1
19 DcHGrsp04 Numeric 2 0 Dc - Family Ty {0, Lone Paren 91 11 Right 20 DcHGrsp05 Numeric 2 0 Dc Whether re {0, No} -91 11 Right	18	DcHGnp04	Numeric	2	0	Dc - Responde	{O, Responden	-91	8	Right	1
20 DcHGrsp05 Numeric 2 0 Dc Whether re {0, No}91 11 Right	19	DcHGrsp04	Numeric	2	0	Dc - Family Ty	{0, Lone Paren	-91	11	Right	1
	20	DcHGrsp05	Numeric	2	0	Dc Whether re	{0, No}	-91	11	Right	1
A Data View A Variable View / Data View /	a đi	ta View λV a	riable View /	-	-		lia u i			>	ř
SPSS Processor is ready	00,0			S	PSS Processor is	ready					9

i) Select 'Analyze' in the menu

ii) Select 'Descriptive statistics -> Frequencies' in the Analyze drop-down menu: the following dialog box will appear:

Frequencies	×
McNMst01 McNMtk01 McNMtk02 McNMtr01 McNMvw01 McNrel03 McObtg01 McObtg01 McObtg02 McObtg0 McObtg0 McObtg0 McObtg0 McObtg0 McObtg0 McObtg0 McObtg0 McObtg0 McOb	Variable(s): OK Paste Reset Cancel Help
	Statistics Charts Format

iii) Browse the list of variables on the left and highlight the one(s) you want and click on the arrow in the middle:

Frequencies			
McNMst01 McNMtk01 McNMtk02 McNMtr01 McNMvw01 McNrel03 McObtg02 McObtg03 Display frequency tables	•	Variable(s):	OK Paste Reset Cancel Help
	Statistic	s Charts Form	at

iv) Click 'OK': the output will show you the frequency table for that variable

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	568	45.6	45.6	45.6
	2 No	678	54.4	54.4	100.0
	Total	1246	100.0	100.0	

McObtg01 Mc - Attend group in last 12 months

3.2.2 Simple crosstab – attendance at parent and toddler groups

If you want to find out whether or not attendance at parent and toddler groups is different amongst parents with different characteristics, the simplest way to do so is using a cross-tabulation.

This next example demonstrates how to run a simple crosstab which compares attendance at parent and toddler groups amongst mothers in different age groups – again, when the cohort child was aged three. It is based on BC1 Sweep 3 dataset.

v) To do cross-tabs, use the same menu as for frequencies above but select 'Crosstabs' instead: Analyse \rightarrow Descriptive statistics \rightarrow Crosstabs

Crosstabs		X
Idnumber SampType DcXmnth1 DcXqurt1 Outcome McHGrsp03 DcHGrsp01 DcHGrsp02 DcHGnmad DcHGnmkd DcHGnmk2 DcHGnmk2	Row(s):	Statistics Cells Format
Display clustered bar charts Suppress tables	Paste Reset Cancel Help	

vi) As before browse and highlight the variables you want to cross-reference and move them to the row and column boxes as applicable:

vii) If you simply want a count of the number of cases in each group then click 'Ok'. If you're interested in percentages (i.e. the % of younger mothers who attend groups compared with the % of older mothers), then click 'Cells' to bring up the 'Cell Display' options:

E	Crosstabs: Cell I	Display 🔀
	Counts	
	Percentages	Residuals Unstandardized Standardized Adjusted standardized
	Noninteger Weights – Round cell counts Truncate cell courts No adjustments	: ⊚ Round case <u>w</u> eights nts ⊚ Truncate case wei <u>gh</u> ts
	Continue	Cancel Help

Under 'Percentages', select 'Column', then click 'Continue'

vii) Click 'OK' in the Crosstabs dialogue box: the following output is displayed. 30% of mothers aged under 20 attended toddler groups compared with 51% of mothers aged 40 or older.

McObtg01 Mc - Attend group in last 12 months * DcHGmag5 Dc Age of natural mother at birth of cohort child (banded) Crosstabulation

			DcHGmag5 Dc Age of natural mother at birth of cohort child (banded)				
			1 Under 20	2 20 to 29	3 30 to 39	4 40 or older	Total
McObtg01	1	Count	66	669	1079	79	1893
Mc - Attend	Yes	% within DcHGmag5 Dc Age	29.9%	42.7%	48.5%	51.0%	45.4%
group in last		of natural mother at birth of					
12 months		cohort child (banded)					
	2 No	Count	155	896	1148	76	2275
		% within DcHGmag5 Dc Age	70.1%	57.3%	51.5%	49.0%	54.6%
		of natural mother at birth of					
		cohort child (banded)					
Total		Count	221	1565	2227	155	4168
		% within DcHGmag5 Dc Age	100.0%	100.0%	100.0%	100.0%	100.0%
		of natural mother at birth of					
		cohort child (banded)					

3.3 Recoding variables via syntax

This example demonstrates a simple recode of the variable indicating the child's general health. The recode collapses the variable from five answer categories into three. It uses data from the BC1 Sweep 3 workshop data file.

i) Check frequencies of the original variable - McHgen01

		Frequency	Percent	Valid Percent	Cumulative Percent			
Valid	1very good,	898	72.1	72.1	72.1			
	2 good,	288	23.1	23.1	95.2			
	3 fair,	56	4.5	4.5	99.7			
	4 bad,	4	.3	.3	100.0			
	Total	1246	100.0	100.0				

McHgen01 Mc - Childs general health

- ii) Open syntax file: go to menu 'File / Open / Syntax ...'
- iii) Type Recode syntax in the syntax file: we want to group the original variables into answer categories Good (1,2) / Fair (3) and Bad (4,5)

RECODE McHgen01 (1 thru 2=1) (3=2) (4 thru 5=3) (else=copy) INTO GenHbdS3. Exe.

iv) and run it. To run the syntax, highlight the syntax text from the 'RECODE' command to the 'Exe' (EXECUTE) command and then click on the large triangle (that looks like a play button) in the toolbar.

😫 Recode_Etc - SPSS Syntax Editor 🛛 🔲 🔼					
File Edit View Data Transform Analyze Graphs Utilities Run Window Help					
😕 🖬 📴 🔹 🐘 🕼 🗛 🕟 🍳 🍨 📴 🗰					
******** RECODE SYNTAX **********					
* Sweep 3 general health variable:					
RECODE McHgen01 (1 thru 2=1) (3=2) (4 thru 5=3) (else=copy) INTO GenHbdS3. Exe.					

v) Check frequencies of the new variable – the command is *FREQ GenHbdS3*:

GenHbdS3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1186	95.2	95.2	95.2
	2.00	56	4.5	4.5	99.7
	3.00	4	.3	.3	100.0
	Total	1246	100.0	100.0	

vi) Tidy up the variable label, value labels and output format directly within the variable row in the dataset 'Variable View' or with additional syntax below:

```
VAR LABS GenHbdS3 'Childs general health - banded'.
VAL LABS GenHbdS3
1 'Good/Very Good'
2 'Fair'
3 'Bad/Very Bad'.
FORMATS GenHbdS3 (F2.0).
```

vii) Run the frequencies again:

	Commodee	ennae gene	i ai moultin	Sanaoa	
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Good/Very Good	1186	95.2	95.2	95.2
	2 Fair	56	4.5	4.5	99.7
	3 Bad/Very Bad	4	.3	.3	100.0
	Total	1246	100.0	100.0	

GenHbdS3 Childs general health - banded

3.4 Computing a Derived Variable using syntax

This example demonstrates the computation of a new variable which combines information from two sweeps to provide a summary of how the child's general health status changed – e.g. whether it improved, got worse or stayed the same. It uses data from the BC1 Sweep 2 and Sweep 3 workshop data files.

- i) Create a new variable 'GenHS2S3' coded 0 if no change between Sw2 and Sw3, 1 if child's health not as good at Sw3 and 2 if child's health better at Sw3
- ii) Set up and run Compute syntax:

COMPUTE GenHS2S3=-1.

IF (McHgen01=MbHgen01) GenHS2S3=0.

IF (McHgen01< MbHgen01) GenHS2S3=1.

IF (McHgen01> MbHgen01) GenHS2S3=2.

Exe.

iii) Check the frequencies of the new variable - FREQ GenHS2S3

	GenH5253								
		Frequency	Percent	Valid Percent	Cumulative Percent				
Valid	-1.00	152	11.0	11.0	11.0				
	.00	827	60.1	60.1	71.1				
	1.00	234	17.0	17.0	88.1				
	2.00	164	11.9	11.9	100.0				
	Total	1377	100.0	100.0					

GenHS2S3

iii) Tidy up variable and value labels, output formats and assign missing values:

VAR LABS GenHS2S3 'Childs general health evolution Sw2 to Sw3'.

VAL LABS GenHS2S3

- 0 'No change'
- 1 'better at Sw3'
- 2 ' not as good at Sw3'
- -1 'Not Applicable: no data at Sw2'.

FORMATS GenHS2S3 (F2.0).

MISSING VALUES GenHS2S3 (-1).

					Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	0 No change	827	60.1	67.5	67.5
	1 better at Sw3	234	17.0	19.1	86.6
	2 not as good at Sw3	164	11.9	13.4	100.0
	Total	1225	89.0	100.0	
Missing	-1 Missing data	152	11.0		
Total		1377	100.0		

GenHS2S3 Childs general health evolution Sw2 to Sw3

4 How to Create and Merge Datasets in SPSS

If the original datasets to be merged are large, the process may be slow and unwieldy. Therefore the preferred method for working on multiple sweeps of data is to create bespoke datasets with the necessary variables (using "DROP" or "KEEP" commands) and then merge these datasets together. For this workshop it is the number of cases in the Sweep 1 sample which has been reduced (to approximately 30% of the original sample), and those cases only have been selected (when applicable) in the subsequent datasets for cross-sweep comparison.

4.1 Creating bespoke datasets using the 'Keep' and 'drop' Commands

4.1.1 The KEEP command

- The **KEEP** command allows you to open or save a large data file specifying which of the variables from that file you wish to <u>INCLUDE</u> in that data file.
- The **KEEP** commands can be appended to either the GET FILE or SAVE OUTFILE commands
- Both individual variables and ranges of variables can be specified
- The case unique identifier Idnumber will usually have to be included (to permit later merging with other datasets)

Syntax examples:

GET FILE='C:\temp\GUSSW3B_30.sav'

Keep = idnumber, dcwinc01, dchgmag3 to dcmedu02.

SAVE OUTFILE='C:\temp\Keep Save As Test.sav'

/Keep = idnumber, dcwinc01, dchgmag3 .

4.1.2 The DROP command

- The DROP command allows you to open a large data file specifying which of the variables from that file you wish to REMOVE from your working data file.
- The DROP command can be appended to either the GET FILE or SAVE OUTFILE commands
- Both individual variables and ranges of variables can be specified
- Again, the case unique identifier Idnumber will usually have to be included (to permit later merging with other datasets

Syntax examples:

GET FILE='C:\temp\GUSSW3B_30.sav'

/Drop = samptype to dcwtchd2.

SAVE OUTFILE='C:\temp\Drop Save As Test.sav'

/Drop = dcurind1, dcurind2.

4.2 Merging Datasets

Datasets can be merged using the unique case ID stored in the variable 'IDnumber'. Whenever you are first merging files, it is easier to use the SPSS menus and then paste the syntax (automatically generated and recorded in the output) rather than using the syntax from scratch as it can be quite tricky depending on how large each of your datasets are and how many identical variables are in each already. The datasets to be merged must always be sorted on the same variable before merging otherwise the matching will not proceed.

- i) Open the dataset you want to merge data into: in the example below it is the Sweep 1 birth cohort dataset
- ii) Sort this dataset on the key variable 'IDnumber' in ascending order via the menu: go to Data\Sort Cases:



• Select the variable 'IDnumber' on the left part of the screen

 And move it to the right part of the screen using the arrow – the default option is 'Ascending' order



o Click 'OK'

- iii) Repeat the same process 1) and 2) above with the dataset you want to extract the data from: the Sweep 2 birth cohort in the example below, to be added to the 1st dataset = Sweep 1 birth cohort
- iv) On the menu of the 1st dataset go to: Data\Merge files\Add variables
- v) In the dialogue box, unless the dataset from which you want to merge is already open, select the button for 'An external SPSS data file' and click 'Browse'. If the dataset is open then select it in the 'open dataset' box (as below).

Add Variables to GUSSW1B_v3.sav [DataSet1]					
Select a dataset from the list of open datasets or from a file to merge with the active dataset					
⊙ An open dataset					
GUS_SW2_B.sav [DataSet2]					
🔿 An external SPSS data file					
Browse					
Non-SPSS data files must be opened in SPSS before they can be used as part of a merge.					
Continue Cancel Help					

- vi) If not already open, browse to the dataset of interest and double-click on it
- vii) Click 'Continue'
- viii) The following dialog box will come up; in this example you can see that there is a big list of 'Excluded Variables' on the left, which are the variables shared by both datasets, instead of just the expected variable 'IDnumber'. This is due to the feed forward process: the archived datasets from Sweep 2 include some of the previous sweep variables since original information is only updated when applicable and we want the full information for all cases at each sweep, including those with no changes. To get the full information for this type of variable you need to incorporate the successive sweeps variables.

Add Variables from GUS_SW2_B.sav [Data	Set2]	
Excluded Variables: Idnumber (+) MaCany01 (+) MaCdya01 (+) MaCdyb01 (+) MaCdyc01 (+) MaCdyd01 (+) MaCdye01 (+)	New Active Dataset: Idnumber (*) SampType (*) Interview (*) DaHGnmad (*) DaHGnmkd (*) DaHGnp01 (*) DaHGrsp01 (*) DaHGrsp01 (*)	OK Paste Reset Cancel Help
 Match cases on key variables in sorted files Both files provide cases Non-active dataset is keyed table Active dataset is keyed table Indicate case source as variable: source01 (*) = Active dataset (+) = GUS_SW2_B.sav [DataSet2] 	Variables:	

ix) In this 'Add variables' dialogue box, click the box 'Match cases on key variables in sorted files', and browse to and highlight the variable 'IDnumber':

Add Variables from GUS_SW2_B.sav [Data	Set2]	
Excluded Variables: Idnumber (+) MaCany01 (+) MaCdya01 (+) MaCdyc01 (+) MaCdyc01 (+) MaCdye01 (+) Rename Match cases on key variables in sorted files Both files provide cases Non-active dataset is keyed table Active dataset is keyed table Active dataset is keyed table Carte dataset Match case source as variable: Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset Carte dataset	New Active Dataset: Idnumber (*) SampType (*) Interview (*) DaHGnmad (*) DaHGnmkd (*) DaHGnp01 (*) DaHGrsp02 (*) Key Variables:	OK Paste Reset Cancel Help

x) Click on the arrow next to the 'Key variables' box. 'IDnumber' should now appear in the 'Key variables' box.

Excluded Variables: MaCany01 (+) MaCdya01 (+) MaCdyb01 (+) MaCdyc01 (+) MaCdyc01 (+) MaCdyc01 (+) MaCtma01 (+) Rename		New Active Dataset: SampType (*) Interview (*) DaHGnmad (*) DaHGnmkd (*) DaHGnmkd (*) DaHGnp01 (*) DaHGrsp01 (*) DaHGrsp02 (*) DaHGnp02 (*)	OK Pasl Res Cano Hel
 Match cases on key variables in s Both files provide cases Non-active dataset is keyed table Active dataset is keyed table Indicate case source as variable: (*) = Active dataset (+) = GUS_SW2_B.sav [DataSet2] 	orted files le	Key Variables: Idnumber	

The steps you take next will depend on what dataset you're already working on:

- xi) Under 'Match cases on key..." if you select...
 - 'Both files provide cases' (default option): All cases from the merged dataset will be transferred into the working dataset. If you are working on a later dataset and merging in data from an earlier dataset, choosing this option means that additional cases from the earlier dataset will be merged along with the variables. These cases will have 'missing' data for the variables at the later sweep because they were not achieved at that sweep.
 - 'Non-active dataset is keyed table': Only merged data for those cases already in the working dataset will be transferred. This avoids the above issue if you are working on a later dataset and merging in a variable from an earlier sweep. Only information from those cases in the working (later) dataset will be merged so you won't generate entire cases with 'missing' data which would need to be deleted or filtered out later on.
 - 'Active dataset is keyed table': All cases from the merged dataset will be transferred into the working dataset. This produces the same result as the first scenario.
- xii) In this example you need to select Option 2 'Non-active dataset is keyed table'

Click 'OK' and again 'OK' in the warning message re cases needing to be sorted before merging

5 Data Management - Additional tips

5.1 To change the order of variable names in dialogue boxes

Default option in Edit/Options: General tab = display variables alphabetically in the list: change to 'File' to get the order as per the dataset

Options	X					
Data Cur General Viewer Draft Viewer Output	rency Scripts Labels Charts Interactive Pivot Tables					
Variable Lists Display labels Oisplay names Alphabetical File Session Journal Record syntax in Journal Append Overwrite F:\SPSSJO~1.JNL Browse	Output No scientific notation for small numbers in tables Viewer Type at Startup: Image: Regular O Draft Measurement System: Points Language:					
Temporary directory: C:\DOCUME~1\MIREIL~1\LOCALS~1\T Recently used file list: 9 ♥ Open syntax window at start-up						
ОК	Cancel Apply Help					

5.2 RECODE a variable via the menu

Select 'Transform' and 'Recode into Different Variables' for the example chosen:

🚰 *GUSSW3B_30.sav [DataSet1] - SPSS Data Editor								
File Edit View Data Transfo	rm Analyze Graphs Utilitie	es Window	Help					
🗁 🔚 🏝 📴 🔸 Com	oute Variable It Values within Cases	9	V 🖉 🌢 🗸					
Name	do into Como Variablac		Label	Values	Missing	Columns	Align	Measure
1 Idnumber Reco	ide into Different Variables	, Г	Serial number	None	None	8	Right	Scale
2 SampType Aoto	matic Recode	′ [Sample Type	{1, Birth}	-91	8	Right	Nominal
3 DcXmnth1 Visua	al Binning	[Dc Month of in	{1, January}	-91	10	Right	Scale
4 DcXqurt1 Back	Casas	[Dc Quarter of i	{1, January to	-91	10	Right	Scale
5 Outcome	. Cases	[Dc Final outco	{110, Full inter	-91	8	Right	Nominal
6 McHGrsp03 Date	and Time Wizard		Mc - Same res	{-9, Refusal}	-91	8	Right	Nominal
7 DcHGrsp01 Crea	te Time Series	ſ	Dc - Whether r	{O, Not mentio	-91	8	Right	Nominal
8 DcHGrsp02 Repl	ace Missing Values	1	Dc - Whether r	{O, Not mentio	-91	8	Right	Nominal
9 DcHGnmad	Iom Number Generators		Dc Number of	None	-91	10	Right	Scale
10 DcHGnmad Run	Pending Transforms	Ctrl+G	Dc Number of	{0, None}	-91	11	Right	Scale
11 DcHGnmkd N	umeric 2 0		Dc Number of	None	-91	10	Right	Scale

Select the variable to be recoded from the list on the left-hand side, and type the name of the new variable under 'Output Variable/Name:' (you can also add a Label for the new variable)

🔲 Recode into Differe	ent Variables	×
Mclhea21 Mclhea22 Mclhea23 Mclhea23 Mclhea24 Mclhea95 Mclhea95 Mclhea25 McHnin01 McHnin02 McHnin03 McHnin04 McHnin04 McHnin04 McHnin04 McHlsi05 McHlsi06	Numeric Variable Output Variable McHgen01> ? GenHbdS3 Label: Change Old and New Values If (optional case selection condition) OK Paste Reset Cancel Help	

To change the values click on 'Old and New Values' button, the following dialog box comes up:

Recode into Different Variables: Old and New Values						
Old Value Value: System-missing	New Value Value: System-missing Copy old value(s)					
 System- or user-missing Range: through Range, LOWEST through value: 	Old> New: Add Change Remove					
 Range, value through HIGHEST: All other values 	Output variables are strings Width: 8 Convert numeric strings to numbers ('5'->5) Continue Cancel Help					

Use 'Range' or 'Value' depending on transformation required

Recode into Different Variables: Old and New	v Values 🛛 🗙
 Old Value Value: System-missing Sustem- or user-missing Range: through Range, LOWEST through value: Range, value through HIGHEST: 	New Value Value: System-missing Copy old value(s) Old> New: Add Change Remove Output variables are strings Width: 8
All other values	Continue Cancel Help

Click on 'Add':

Recode into Different Variables: Old and New Values							
Old Value Value: 3 System-missing System- or user-missing Range: through	New Value Value: 2 System-missing Copy old value(s) Old> New: Add Lhange Remove						
 Range, LOWEST through value: Range, value through HIGHEST: All other values 	□ Output variables are strings Width: 8 □ Convert numeric strings to numbers ('5'->5) Continue Cancel Help						

Repeat for other values to be grouped in this example

For missing values, just copy over original codes (and click 'Add'):

Recode into Different Variables: Old and New Values 🛛 🛛 🔀							
Old Value Value: Sustem-missing System- or user-missing Range: through Change, LOWEST through value:	New Value Value: System-missing Copy old value(s) Old> New: Add Change Remove						
 Range, value through HIGHEST: All other values 	Output variables are strings Width: 8 Convert numeric strings to numbers ('5'->5) Continue Cancel Help						

Click on 'Continue'

Click on 'Change' in the next screen

McIhea21 McIhea22 McIhea23 McIhea24 McIhea94 McIhea95 McIhea25 McHnin01 McHnin03 McHnin04 McHnin04 McHnin04 McHisi05 McHisi05



5.3 COMPUTE a variable via menu

Select 'Transform' and 'Compute Variable' for the example chosen – see drop-down list shown for Recode page 1, the following dialog box opens:

Enter name of new variable top left and select the operations or function you will be using in the middle or on the right:

	Compute Variable		×
\langle	Target Variable: Devlpt1 Type & Label	= Numeric Expression:	<
	Control Contr	+ > 789 + < 789 - < 4 Arithmetic CDF & Noncentral CDF Conversion Current Date/Time Z 0 /* /*	
	Temp1 Temp2 Temp3 Temp4 Temp5	SUM[numexpr,numexpr],]. Numeric. Returns the sum of its arguments that have valid values. This function requires two or more arguments, which must be numeric. You can specify a minimum number of valid arguments for this function to be evaluated. Substr[1] Substr[2] Substr[2] Sum Sysmis Time.Hms[1] Time.Hms[3] Tunc	
	If (optional case sel	ection condition)	~

Select the dataset variables required in the operation from the list on the left and click 'OK':



6 Coping with Complex Samples

6.1 Creating a complex samples plan file in SPSS

NOTE: This process requires the Complex Samples Module in SPSS. This module is not supplied in the standard SPSS package. The module is not available on versions of SPSS prior to version 15.

It is common practise to weight your data to account for the non-response bias in your sample, which usually results in disadvantaged groups being underrepresented. If the sample in your data was generated using clustering and stratification and is not a simple random sample, then you should account for the sampling design as you weight your data. If you do not account for the sample design, SPSS assumes it is dealing with a simple random sample. In doing so it will underestimate the real amount of error (the standard error), affecting the confidence intervals and ultimately provide a potentially inaccurate result in significance tests. By not accounting for the complex sample design, SPSS will return differences which are not significant as significant.

To resolve this, and to account for the complex sample design, SPSS requires you to create a complex samples plan file (these files end with '.csaplan').

- 1. You need three components to create a plan file
 - a) A weight variable
 - b) A stratification variable
 - c) A cluster variable

The precise variables to select will depend on the data incorporated in your analysis (i.e. whether it is cross-sectional or longitudinal analysis and which sweeps of data are involved)

The example below illustrates the creation of a plan file for cross-sectional analysis of birth cohort data from sweep 3.

- 2. To weight for cross-sectional analysis for the birth cohort at sweep 3 you need the following:
 - a) Weight = DcWTbrth
 - b) Stratification variable (Strata) = Dcstrat
 - c) Cluster variable = Dcpsu
- 3. On the menu tool bar follow the command: Analyze//Complex samples/prepare for analysis
 - \rightarrow Select: Create a plan file

 \rightarrow Select a destination to save your plan file: your temporary folder, and give your file a name (you'll find using something descriptive incorporating the sweep and weight used is useful)

- \rightarrow Click NEXT
- \rightarrow Insert the relevant variables as shown below in the screenshot, click NEXT

nat define strata or clusters. A sample	weight variable m	ust be selected in the first stage.
ige that will be used in the output.		
⊻ariables:		Strata:
🕹 CAssPrb1 🔺		🔗 DoStrat
💫 CAssPrb2 🦳	•	
💦 CAssPrb3	r	
💑 CAssPrb4		
CAssPrb5		<u>C</u> lusters:
CAssPrb6		n DcPSU
DePicRaw	•	
✓ DePicSPt		
✓ DoNamBaw		Sample Weight:
DcNamVAS		DcWTbrth
DcNamVTS		
🛷 DcNamVPt		
🖉 DcWTchld 📃	Charac 1	-h-h
	ar derine strata of clusters. A sample ge that will be used in the output.	ard derine strata or clusters. A sample weight Variable in ge that will be used in the output.

→ Now select the WR option (sampling with replacement) and UNSELECT the option "Use finite population correction" (FPC). Click Next.

 \rightarrow The summary page shows you the selections you have made. Click Next

 \rightarrow Select 'Paste the syntax generated' and click 'Finish' to see it in your syntax window.

 \rightarrow You should now be able to see the following syntax in your Syntax Window. Select it with the runner and click on the blue play button on the toolbar or click CTRL-R on your keyboard.



Your csplan file is ready. You can use this every time you intend to do cross-sectional analysis of birth cohort data from sweep 3. SPSS will prompt you to select a plan file you have already constructed when you use any of the commands available through the Analyze/Complex Samples commands.

If you wish to look at the child cohort you will have to use the appropriate weight and create another csplan file, and you will also have to create a different csplan file for analysis between sweeps 1, 2 and 3, and so on. However, as suggested above, you can do this by directly editing the above syntax, as well as via the menus.

It is good practise to keep the syntax of your plan files, and select intuitive names for these files so you can identify them instantly.

Identifying the correct GUS weight

Sweep 1 Weight

- Use for any cross-sectional analysis of SWEEP 1 DATA ONLY
 - Birth cohort (either BC1 or BC2): dawtbrth
 - Child cohort: dawtchld

Sweep 2 Main interview weight

- Use for any cross-sectional analysis of SWEEP 2 DATA ONLY
- Use for any longitudinal analysis involving Sweep 1 AND Sweep 2 data
 - Birth cohort: dbwtbrth
 - Child cohort: dbwtchld

Sweep 2 Partner interview weight

- Use for any analysis of Partner interview data
 - Birth cohort: dbwtbrtp
 - Child cohort: dbwtchlp

Sweep 3 Longitudinal weight

- Use for any analysis of more than one sweep of data, when data from sweep 3 is included (e.g. data from sweep 1 and sweep 3, or from sweep 2 and sweep 3, or from sweeps 1, 2 and 3)
 - Birth cohort: dcwtbth2
 - Child cohort: dcwtchd2

Sweep 3 Cross-sectional weight

- Use for any cross-sectional analysis of SWEEP 3 DATA ONLY
 - Birth cohort: dcwtbrth
 - Child cohort: dcwtchld

Sweep 4 Longitudinal weight

- Use for any analysis of more than one sweep of data, when data from sweep 4 is included (e.g. data from sweep 1 and sweep 4, or from sweep 2 and sweep 4, or from sweeps 1, 2, 3 and 4)
 - Birth cohort: **ddwtbth2**

• Child cohort: ddwtchd2

Sweep 4 Cross-sectional weight

- Use for any cross-sectional analysis of SWEEP 4 DATA ONLY
 - Birth cohort: ddwtbrth
 - o Child cohort: ddwtchld

Sweep 5 Longitudinal weight

- Use for any analysis of more than one sweep of data, when data from sweep 5 is included (e.g. data from sweep 1 and sweep 5, or from sweep 3 and sweep 5, or from sweeps 3, 4 and 5)
 - o Birth cohort: dewtbth2

Sweep 5 Cross-sectional weight

Use for any cross-sectional analysis of SWEEP 5 DATA ONLY

 Birth cohort: dewtbrth

Sweep 6 Longitudinal weight

- Use for any analysis of more than one sweep of data, when data from sweep 6 is included (e.g. data from sweep 1 and sweep 6, or from sweep 3 and sweep 6, or from sweeps 4, 5 and 6)
 - Birth cohort: dfwtbth2

Sweep 6 Cross-sectional weight

- Use for any cross-sectional analysis of SWEEP 6 DATA ONLY
 - Birth cohort: **dfwtbrth**

6.2 Basic analysis using the Complex Samples (CS) Module in SPSS

Having created your CS plan file, you can now use it as a component in your analysis allowing you to check statistical significance whilst correcting for the complex sample design in GUS.

This worksheet illustrates how to run, and check significance in, a simple crosstabulation using the CS module. The example is cross-sectional using the birth cohort data from sweep 3 (and therefore corresponds with the CS plan file created in the previous worksheet). Note that unless you are testing significance, there is no need to use the CS analysis approach, a simple weighted table or crosstab will do.

1. Open the sweep 3 birth cohort dataset.

S	On the	monu	tool har	coloct.	Analyz	n/Com		Some		Crocot	aha
∠.	On the	menu	luui bai	Select.	Analyze	=/ C0111	piex c	Sam	162/1	010221	aus

🚉 GU	SS₩3B_30.sa	iv [DataSet	t1] - SPSS I	Data E	ditor						
File Ed	t View Data	Transform	Analyze Gr	raphs	Utilities	۷	Vindow	Help			
<u> </u>	🖹 🖬 🖕	•	Reports Descriptiv	/e Statis	⊧ tics		E	<u>s</u>]		
	Name	Туре	Tables		•	s		Label	− Va	alues	
	1 Idnumber	Numeric	Compare	Means	•	Γ	Se	rial number	None		No
	2 SampType	Numeric	General L	inear Mo	odel 🕨	Γ	Sa	mple Type	{1, Bir	th}	-9 -
	3 DcXmnth1	Numeric	Correlate	!		Γ	Dc	Month of in	{1, Jar	nuary}	-9 -
	4 DcXqurt1	Numeric	Classify	n		Γ	Dc	Quarter of i	{1, Jar	nuary to	-9 -
	5 Outcome	Numeric	Data Red	luction		Γ	Dc	Final outco	{110,	Full inter	-9 -
	McHGrsp03	Numeric	Scale	action	•		Mc	- Same res	{-9, Re	efusal}	-9 -
	7 DcHGrsp01	Numeric	Nonparan	netric Te	ests 🕨	Γ	Dc	- Whether r	{0, No	t mentio	-9 -
	B DcHGrsp02	Numeric	Time Serie	es	•		Dc	- Whether r	{0, No	t mentio	-9 -
	DcHGnmad	Numeric	Multiple R	esponse	e 🕨	1	Dc	Number of	None	1	-9 -
1	DcHGnmad	Numeric	Complex !	Samples	; 🕨		Select	: a Sample		ne}	-9 -
1	1 DcHGnmkd	Numeric	Quality C	ontrol	•	_	Prepa	re for Analysis			-9 -
1	2 DcHGnmk2	Numeric	ROC Curv	ve		-				e}	-9 -
1	3 DcHGhsiz	Numeric	2	0			Descri	intives			-9 -
1	4 DcHGnmsb	Numeric	2	0			Cross	tabs			-9 -
1	5 DcHGnp01	Numeric	2	0			Ratios	5		t mentio	-9 -
1	DcHGnp02	Numeric	2	0]—				t mentio	-9 -
1	7 DcHGnp03	Numeric	2	0			Logist	ai Linear Mode is Regression	H	t mentio	-9 -
1	B DcHGnp04	Numeric	2	0			Ordin	al Regression		sponden	-9 -
1	DcHGrsp04	Numeric	2	0		_		- i anniy i y	י. עט, בטו	ne Paren	-9 -
2	DcHGrsp05	Numeric	2	0			Dc	Whether re	{0, No	}	-9 -
2	1 DoHGren06	Numeric	2	Π			De	Whether re	(O No	1	-9 -

3. In the 'Complex Samples Plan for Crosstabs Analysis' dialog box, the plan file you have just created may already appear in the 'File' box, if not, click 'Browse' and navigate to the correct plan file. [REMEMBER – the plan file you require is dependent on the data you are using and the analysis you are running]. Once you have selected your plan file, click 'Continue' (Ignore 'Joint Probabilities')

4. The 'Complex Samples Crosstabs' dialog box which appears is almost identical to the standard Crosstabs dialog box. Select your Row variables and Column variables (and nested variables under 'Subpopulations') as necessary. For this example we want to compare household income by family type:

- Under the variable list find and select the household income variable (DcWinc01) once selected click the black arrow next to the 'Column' box.
- Back in the variable list find and select the family type variable (DcHGrsp04)once selected click the black arrow next to the 'Row' box

		•	
Complex Samples Cr	osstabs		×
Variables: McWinc07 McWinc08 McWben01 McWben2b McWben02	•	Ro <u>w</u> s: CHGrsp04	OK <u>P</u> aste <u>R</u> eset Cancel
McWben03 McWben04 McWben05 McWben06 McWben07 McWben07 McWben08 McWben08 McWben09 McWben09 McWben10	•	Column: DcWinc01 Subpopulations:	Help
AcWben11	- -	Each combination of categorie defines a subpopulation.	25
<u>S</u> tatistics	<u>M</u> issing Valu	Jes Options	

 Click on the 'Statistics' tab. In the 'Statistics' dialogue box, under 'Cells' unselect 'Population Size' and select 'Column' percent. Under 'Statistics', unselect 'Standard Error' and select 'Confidence Interval' (keep the level at 95%). At the bottom, select 'Test of independence of rows and columns'. Click 'Continue'.

Complex Samples Crosstabs: Sta	atistics	×						
Cells Population size	☑ Column percent	Continue						
☐ <u>R</u> ow percent	I able percent	Cancel Help						
Statistics								
Standard error	Unweighted count							
Confidence interval	🔲 <u>D</u> esign effect							
Le <u>v</u> el (%): 95	Sguare root of design effect							
Coefficient of variation	🗖 Residua <u>l</u> s							
Expected values	Adjusted residuals							
Summaries for 2-by-2 Tables								
□ <u>O</u> dds ratio	🔲 Risk differe <u>n</u> ce							
☐ Relative ris <u>k</u>								
✓ Test of independence of rows and columns								

• Back in the 'Crosstabs' dialog box, click 'Paste' to paste your syntax, which should look like this:

CSTABULATE /PLAN FILE = 'C:\temp\sw3bcxs.csaplan' /TABLES VARIABLES = DcHGrsp04 BY DcWinc01 /CELLS COLPCT /STATISTICS SE /TEST INDEPENDENCE /MISSING SCOPE = TABLE CLASSMISSING = EXCLUDE.

5. Select the syntax with your mouse and click on the blue play button on the toolbar or click CTRL-R on your keyboard. SPSS will return the following output:

Dc - Family	Туре									
5				Dc Household income - banded						
				Up to	From £15000	From				
				£14999	to £25999 per	£26000 to	£44,000			
				per year	year	£43999	and above	Total		
Lone Parent	% within Dc Household	Estimate		62.9%	15.5%	2.7%	1.2%	20.2%		
	income - banded	95% Confidence Interval	Lower	57.6%	11.4%	1.4%	.4%	17.8%		
			Upper	67.9%	20.6%	5.1%	3.6%	22.9%		
Couple Family	% within Dc Household	Estimate		37.1%	84.5%	97.3%	98.8%	79.8%		
	income - banded	95% Confidence Interval	Lower	32.1%	79.4%	94.9%	96.4%	77.1%		
			Upper	42.4%	88.6%	98.6%	99.6%	82.2%		
Total	% within Dc Household	Estimate		100.0%	100.0%	100.0%	100.0%	100.0%		
	income - banded	95% Confidence Interval	Lower	100.0%	100.0%	100.0%	100.0%	100.0%		
			Upper	100.0%	100.0%	100.0%	100.0%	100.0%		

Dc - Family Type * Dc Household income - banded

Tests of Independence

	-	Chi-Square	Adjusted F	df1	df2	Sig.
Dc Household income - banded * Dc - Family Type	Pearson	463.872	159.212	2.838	184.482	.000
	Likelihood Ratio	452.276	155.232	2.838	184.482	.000

The adjusted F is a variant of the second-order Rao-Scott adjusted chi-square statistic. Significance is based on the adjusted F and its degrees of freedom.

6. The significance value is returned in the final column of the 'Tests of Independence' variable. At a significance of p < 0.001, we can conclude that family type is

significantly associated with household income. And from the data, we can see that lone parent families are considerably more likely to be on lower incomes than are couple families.

To run further crosstabs with significance testing on the sweep 3 birth cohort data, all you now need to do is copy and paste the syntax you've created and adjust the row and column variables as necessary.

/TABLES VARIABLES = <u>DcHGrsp04</u> BY <u>DcWinc01</u>

If you are running analysis on a different dataset, or combination of data, and have already prepared an appropriate plan file, all you need to do is change the plan file reference in the syntax

/PLAN FILE = <u>'C:\temp\sw3bcxs.csaplan'</u>

7 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.

7.1 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. This is the same for both BC1 and BC2.

The syntax command is:

Weight by dawtbrth.

fre dahgnp04.

exe.

This should produce the following result (on the reduced BC1 workshop dataset):

		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	

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

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)

7.2 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	s natural mother at interview	(banded) * DcWinc01 I	Dc Household income -	banded Crosstabulation
--	---------------------------	-------------------------------	-----------------------	-----------------------	------------------------

			DcWin	c01 Dc Househo	old income - ba	anded	
			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 % within DcHGmag3 Dc Age of childs natural mother at interview (banded)	68 43.6%	54 34.6%	22 14.1%	12 7.7%	156 100.0%
	3 30 to 39	Count % within DcHGmag3 Dc Age of childs natural mother at interview (banded)	47 14.1%	82 24.6%	108 32.3%	97 29.0%	334 100.0%
	4 40 or older	Count % within DcHGmag3 Dc Age of childs natural mother at interview (banded)	13 10.4%	25 20.0%	30 24.0%	57 45.6%	125 100.0%
Total		Count % within DcHGmag3 Dc Age of childs natural mother at interview (banded)	128 20.8%	161 26.2%	160 26.0%	166 27.0%	615 100.0%

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?
- 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?

8 Workshop Two – Answers to the practical exercises

8.1 Exercise 1: Weighted Frequencies

1. % of children aged 34 months lived in an area classed as 'large urban'? 38%

Could be using birth cohort data at sweep 3 or child cohort data at sweep 1 and applying the appropriate cross-sectional cohort weight.

Syntax would be: Weight by dcwtbrth. fre dcurind2.

exe.

DcURind2 ALc - SG urban-rural classification

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Large urban	<mark>469</mark>	<mark>37.6</mark>	<mark>37.6</mark>	37.6
	2 Other urban	433	34.7	34.7	72.2
	3 Small, accessible towns	108	8.6	8.6	80.9
	4 Small remote towns	28	2.3	2.3	83.2
	5 Accessible rural	150	12.0	12.0	95.2
	6 Remote rural	60	4.8	4.8	100.0
	Total	1248	100.0	100.0	

For child cohort, syntax would be:

Weight by dawtchld. fre ALaURin2. exe.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Large urban	<mark>321</mark>	<mark>37.5</mark>	<mark>37.5</mark>	37.5
	2 Other urban	266	31.1	31.1	68.6
	3 Small, accessible towns	90	10.5	10.5	79.1
	4 Small remote towns	16	1.9	1.9	81.0
	5 Accessible rural	112	13.1	13.1	94.0
	6 Remote rural	51	6.0	6.0	100.0
	Total	856	100.0	100.0	

ALaURin2 SE urban-rural classification

2. % of mothers of children aged 4-5 years were employed full-time? 15% on BC, 19% on CC – decrease possibly as a result of recession?

Could be using birth cohort sweep 5 or child cohort sweep 3

For birth cohort, syntax would be:

Weight by dewtchld.

fre dewsta02.

exe.

DeWsta02 De Mothers employment status (incl. adopt./foster/step-mothers)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Childs mother working - full-time	<mark>168</mark>	<mark>14.8</mark>	<mark>15.0</mark>	15.0
	2 Childs mother working - part-time	562	49.6	50.2	65.2
	3 Childs mother not working	389	34.3	34.8	100.0
	Total	1119	98.7	100.0	
Missing	-3 Information not available	14	1.3		
Total		1134	100.0		

For child cohort, syntax would be: *Weight by dcwtchld.*

fre dcwsta02.

exe.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Childs mother working - full-time	<mark>125</mark>	<mark>18.3</mark>	<mark>18.8</mark>	18.8
	2 Childs mother working - part-time	300	43.9	45.1	63.9
	3 Childs mother not working	240	35.1	36.1	100.0
	Total	666	97.4	100.0	
Missing	-3 No information	18	2.6		
Total		684	100.0		

DcWsta02 Dc Mothers employment status

3. How has general health of children aged 3-4 years changed between 2007 and 2009? A bit of a change: 2007 - 65% very good and 28% good, changed to 73% and 22% respectively by 2009, overall 'good' increased 93% to 95%.

In the context of this exercise should use child cohort data at sweep 2 and birth cohort data from sweep 4.

Syntax for child cohort would be:

Weight by dbwtchld

fre mbhgen01.

Exe.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Very good	479	64.8	64.8	64.8
	2 Good	208	28.2	28.2	93.1
	3 Fair	48	6.5	6.5	99.6
	4 Bad	3	.4	.4	100.0
	Total	738	100.0	100.0	

MbHgen01 Mb - Child's general health

Syntax for birth cohort would be: Weight by ddwtbrth. fre mdhgen01. Exe.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1very good,	855	71.8	71.8	71.8
	2 good,	267	22.4	22.4	94.2
	3 fair,	61	5.1	5.1	99.3
	4 bad,	5	.4	.4	99.7
	5 or, very bad?	3	.3	.3	100.0
	Total	1191	100.0	100.0	1

MdHgen01 Md - Childs general health

8.2 Exercise 2: Weighted Crosstabs

1. How does car ownership vary amongst families of 34 month-old children who live in areas of different deprivation?

Can use CC sweep 1 or BC sweep 3.

For BC sweep 3, syntax is:

weight by dcwtbrth.
cross dcadsco2 by mczveh01
/cells = count row
/count = truncate cell.
exe.

DcADsco2 ALc - SIMD 2006 quintiles * McZveh01 Mc - Household has cont use of vehicle Crosstabulation

-			McZveh01 Mo has cont us	c - Household e of vehicle	
			1 Yes	2 No	Total
DcADsco2	1 0.9449 - 7.7446	Count	225	7	232
2006 quintiles	- least deprived	% within DcADsco2 ALc - SIMD 2006 quintiles	97.0%	3.0%	100.0%
	2 7.7472 -	Count	217	20	237
	13.5627	% within DcADsco2 ALc - SIMD 2006 quintiles	91.6%	8.4%	100.0%
	3 13.5640 - 21.0436	Count	211	30	241
		% within DcADsco2 ALc - SIMD 2006 quintiles	87.6%	12.4%	100.0%
	4 21.0521 - 33.6982	Count	198	50	248
		% within DcADsco2 ALc - SIMD 2006 quintiles	79.8%	20.2%	100.0%
	5 33.7252 -	Count	156	129	285
	deprived	% within DcADsco2 ALc - SIMD 2006 quintiles	54.7%	45.3%	100.0%
Total		Count	1007	236	1243
		% within DcADsco2 ALc - SIMD 2006 quintiles	81.0%	19.0%	100.0%

For CC sweep 1, syntax is:

weight by dawtchld.
cross ALaSNimd by mazveh01
/cells = count row
/count = truncate cell.
exe.

			MaZveh01 D member household, a or have conti any motor ve	o you, or any s of your t present own nuous use of hicles (SW1)	
			1 Yes	2 No	Total
ALaSNimd	1 0.9449 -	Count	152	3	155
Quintiles	deprived	% within ALaSNimd SIMD 2006 Quintiles	98.1%	1.9%	100.0%
	2 7.7472 -	Count	153	12	165
	13.5627	% within ALaSNimd SIMD 2006 Quintiles	92.7%	7.3%	100.0%
	3 13.5640 - 21.0436	Count	149	24	173
		% within ALaSNimd SIMD 2006 Quintiles	86.1%	13.9%	100.0%
	4 21.0521 -	Count	97	42	139
	33.6982	% within ALaSNimd SIMD 2006 Quintiles	69.8%	30.2%	100.0%
	5 33.7252 -	Count	117	100	217
	89.0941 - most deprived	% within ALaSNimd SIMD 2006 Quintiles	53.9%	46.1%	100.0%
Total		Count	668	181	849
		% within ALaSNimd SIMD 2006 Quintiles	78.7%	21.3%	100.0%

ALaSNimd SIMD 2006 Quintiles * MaZveh01 Do you, or any members of your household, at present own or have continuous use of any motor vehicles (SW1) Crosstabulation

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)

Need to use child cohort data at sweep 3. Equivalent variable is not available for BC at sweep 5.

weight by dcwtchld. cross dcxqurt1 by mcaply02 /cells = count row /count = truncate cell. exe.

		-			Mc - F	Play outdo	ors in last	week		
			0	1	2	3	4	5	6	7
Dc Quarter of interview	January to March	Count	23.24	18.312	28.167	18.199	12.089	13.012	3.415	44.801
		% within Dc Quarter of interview	14.4%	11.4%	17.5%	11.3%	7.5%	8.1%	2.1%	27.8%
	April to June	Count	.000	.843	7.198	8.669	14.818	17.620	3.750	117.149
		% within Dc Quarter of interview	.0%	.5%	4.2%	5.1%	8.7%	10.4%	2.2%	68.9%
	July to September	Count	2.907	2.486	4.829	10.622	14.994	16.582	7.464	143.751
		% within Dc Quarter of interview	1.4%	1.2%	2.4%	5.2%	7.4%	8.1%	3.7%	70.6%
	October to December	Count	10.341	4.913	10.653	5.774	12.370	19.661	10.006	73.950
		% within Dc Quarter of interview	7.0%	3.3%	7.2%	3.9%	8.4%	13.3%	6.8%	50.1%
Total		Count	36.489	26.553	50.847	43.264	54.271	66.874	24.634	379.651
		% within Dc Quarter of interview	5.3%	3.9%	7.4%	6.3%	8.0%	9.8%	3.6%	55.6%

Dc Quarter of interview * Mc - Play outdoors in last week Crosstabulation

3. What proportion of families who used non-parental childcare when their child is aged 10 months is still doing so at age 34 months?

Crosstab on use of childcare at sweep 1 by use of childcare at sweep 3 in birth cohort using longitudinal weight.

weight by dcwtbrth2. cross macany01 by dccany01 /cells = count row /count = truncate cell.

exe.

		Dc Whet uses regula Sv	Dc Whether resp uses regular CCare at Sw3		
			Yes	No	Total
Sw1 Whether using any	Yes	Count	10.306	.898	11.205
childcare for cohort child		% within Sw1 Whether using any childcare for cohort child	92.0%	8.0%	100.0%
	No	Count	8.123	1.978	10.101
		% within Sw1 Whether using any childcare for cohort child	80.4%	19.6%	100.0%
Total		Count	18.429	2.876	21.305
		% within Sw1 Whether using any childcare for cohort child	86.5%	13.5%	100.0%

ScotCen Social Research 73 Lothian Road Edinburgh EH3 9AW T 0131 228 2167 www.scotcen.org.uk

A company Limited by Guarantee Registered in England No.4392418 A Charity registered in England and Wales (1091768) and Scotland (SC038454)