Growing Up in Scotland: Data Workshops 2014

Handout Pack

Authors: Paul Bradshaw and Joan Corbett
Date: January 2014

ScotCen
Social Research that works for society

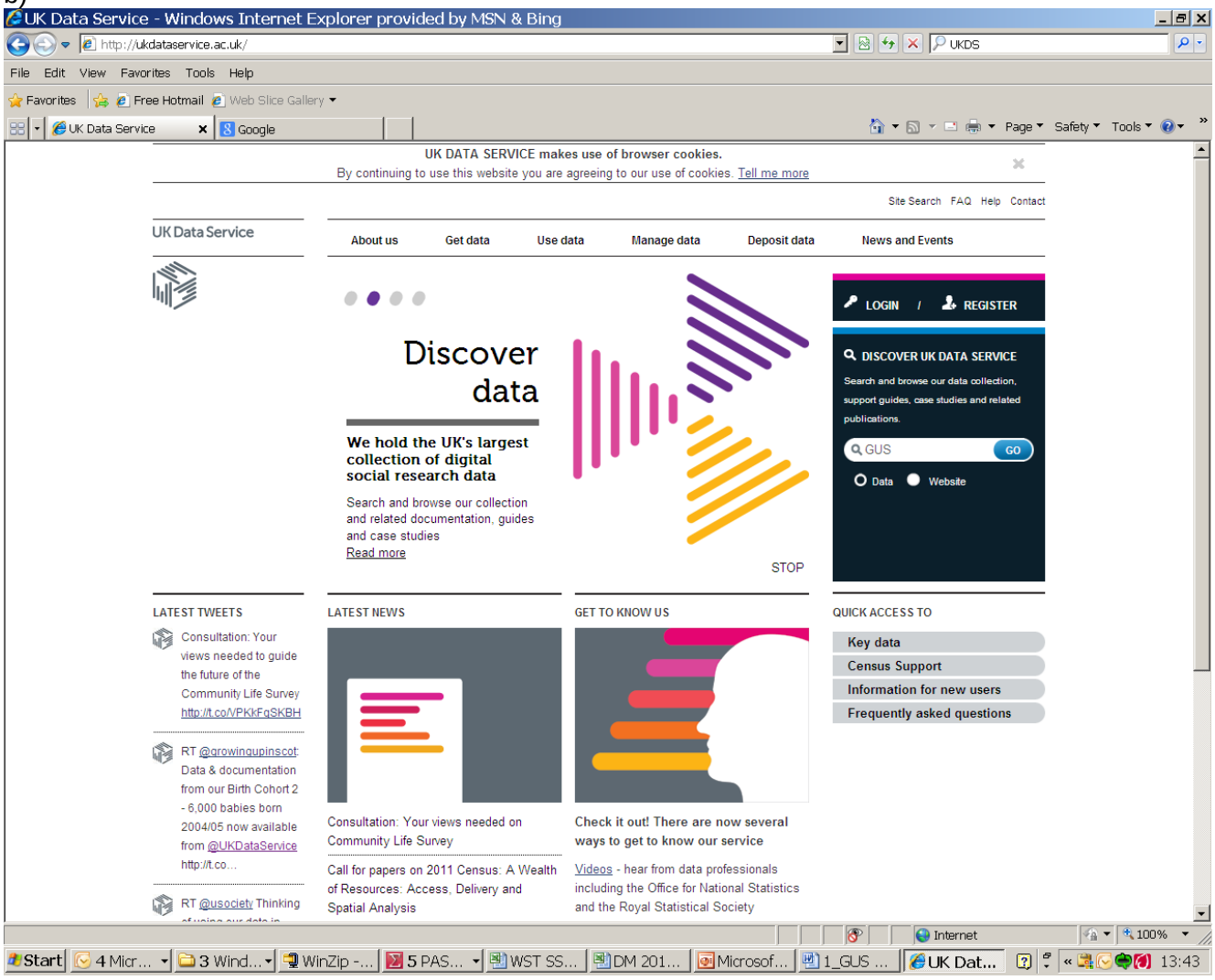
crfr
centre for research on families and relationships

The Scottish Government

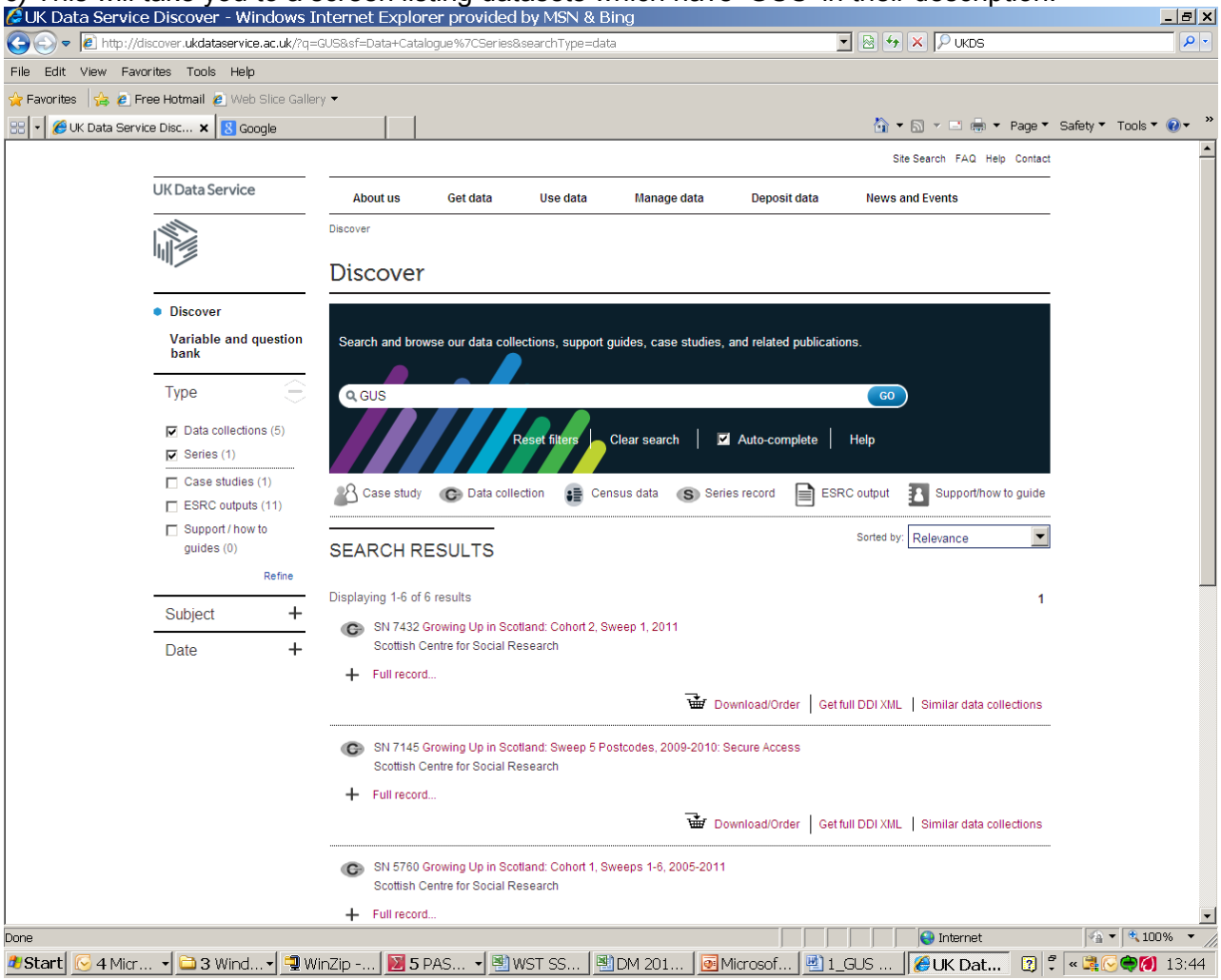
1 Accessing the GUS data at UKDS

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

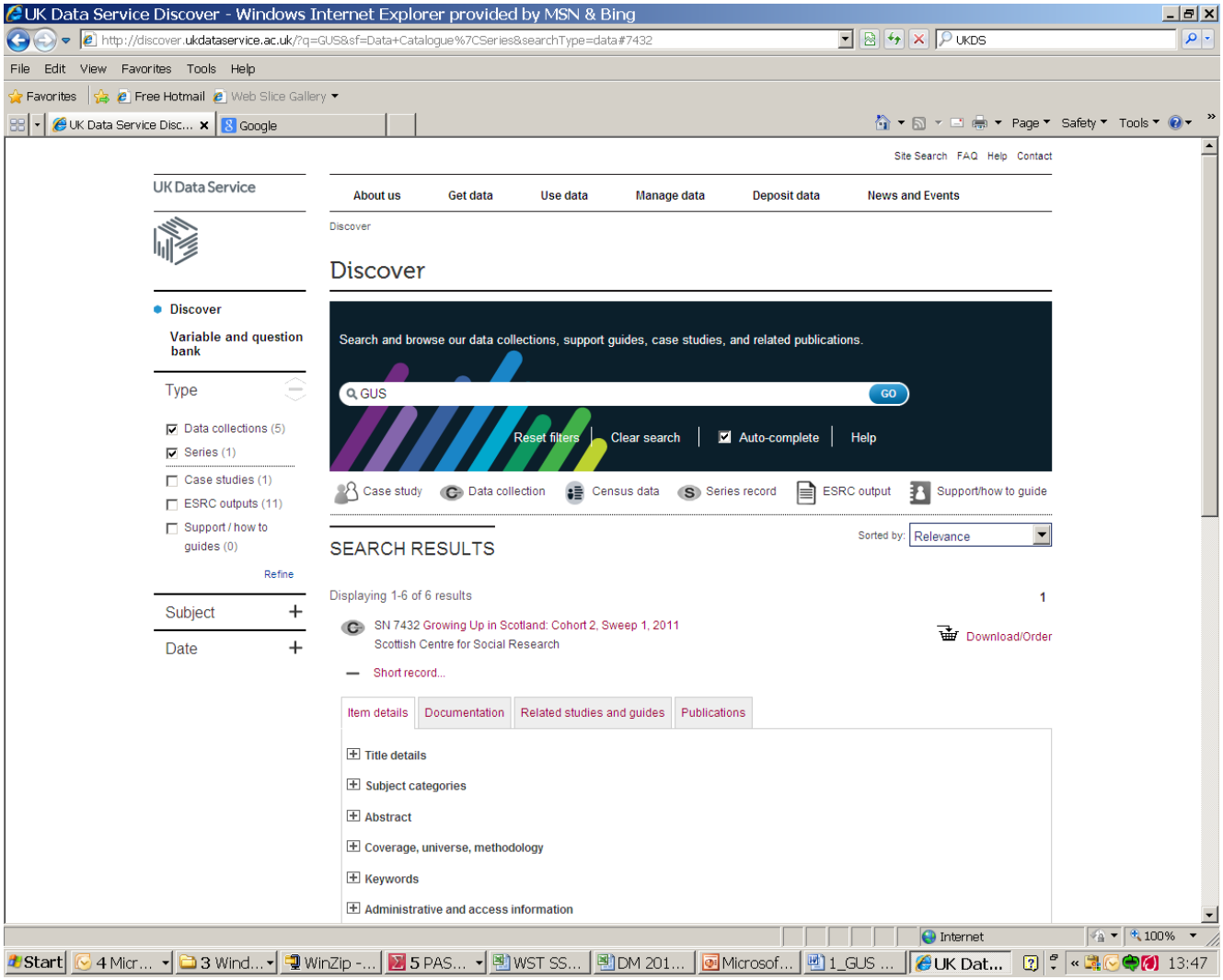
b)



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

Variable name	Detail
D*HGnkd2	Number of children in the household
D*HGnmad	Number of adults in the household
D*HGnmk2	Number of children in household - Banded
D*HGprim	Whether child was mothers first-born
D*HGrsp05	Whether respondent is childs mother (including adoptive, foster and step-mothers)
D*HGrsp06	Whether respondent is childs father (including adoptive, foster and step-fathers)
D*HGrsp07	Who is the respondent in relation to the child?
D*HGrsp08	Who is the respondents partner in relation to the child?
D*HGmag2	Age of mother at time of interview
D*HGmag3	Age of mother at interview (banded)
D*HGmag4	Age of mother at birth of sample child (years)
D*HGmag5	Age of mother at birth of sample child (banded)
D*Meth07	Ethnicity of Respondent
D*Yeth07	Ethnicity of Partner
D*Msta01	Respondent's employment status
D*Msta02	Respondent's employment status: ver2
D*Ysta01	Partner's employment status
D*Ysta02	Partner's employment status: ver2
D*Msta10	Household employment: ver1
D*Msta11	Household employment and family type
D*Wsta02	Mother's employment status
D*Wsta03	Mother's employment status
D*Medu01	Highest Education level of Respondent
D*Medu02	Highest Education level of Respondent - Banded
D*Medu04	Highest Education level of Respondent – Banded SCQF
D*Yedu01	Highest Education level of Partner
D*Yedu02	Highest Education level of Partner - Banded
D*Medu04	Highest Education level of Respondent – Banded SCQF
MfWinc09	Household income – full version

D*Winc01	Household income - banded
D*Eqv5	Equivalised household income – banded (quintiles)
D*Msec01	Respondent NSSEC - 5 Category
D*Ysec01	Respondent NSSEC - 5 Category
D*Msec10	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)
D*Zten01	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**			**Sw2**		**Sw3**		**Sw4**		**Sw5**	**Sw6**
	BC1	**BC2**	**CC**	**BC1**	**CC**	**BC1**	**CC**	**BC1**	**CC**	**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) , *Dbab01, Dbab09 to Dbab12*	✓	✓									
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*		✓	✓	✓	✓	✓		✓		✓	
Concerns about child's development, learning or behaviour *Dspe04 to Dspe11; at sw6:Dcon01 to Dcon03*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
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**			**Sw2**		**Sw3**		**Sw4**		**Sw5**	**Sw6**
	BC1	**BC2**	**CC**	**BC1**	**CC**	**BC1**	**CC**	**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)*			✓				✓	✓	✓	✓	✓
Concerns about child's development, learning or behaviour *Dcon01 to Dcon03*		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Child and respondent's sleep *Dsle01, Dsle02, Dsle03*	✓	✓	✓			✓					✓
Parent-child relationship											
Condon Infant-Maternal Attachment *Mfee01 to Mfee06*	✓	✓									
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 *Tsib01, Tbeh01*			✓		✓	✓		✓			

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.

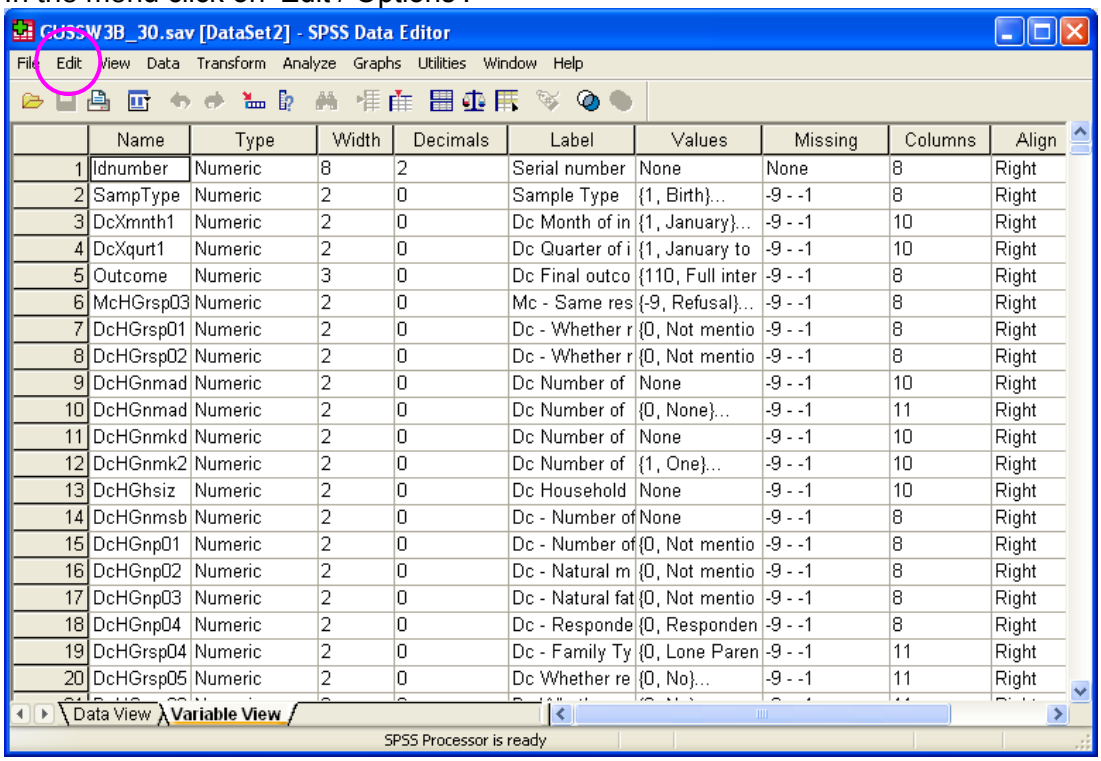
Question topics	**Sw1**			**Sw2**		**Sw3**		**Sw4**		**Sw5**	**Sw6**
	BC1	**BC2**	**CC**	**BC1**	**CC**	**BC1**	**CC**	**BC1**	**CC**	**BC1**	**BC1**
Child's general health *Hgen01*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Child health problems (not necessarily involving medical intervention) (number and type) *Hprb02 – Hprb95*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Child accidents that have involved medical intervention *Haca01 – Hace95*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Child's long-standing illness or disability (existence, nature, whether limiting) *Hlsi01/Hlsa01/Hlsb01/Hlsc01*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
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*					✓			✓	✓		✓
Child gross & motor development milestones *Dbab02 to Dbab08, Dtod01 – Dtod14*	✓	✓	✓			✓					

3 Data Management in SPSS

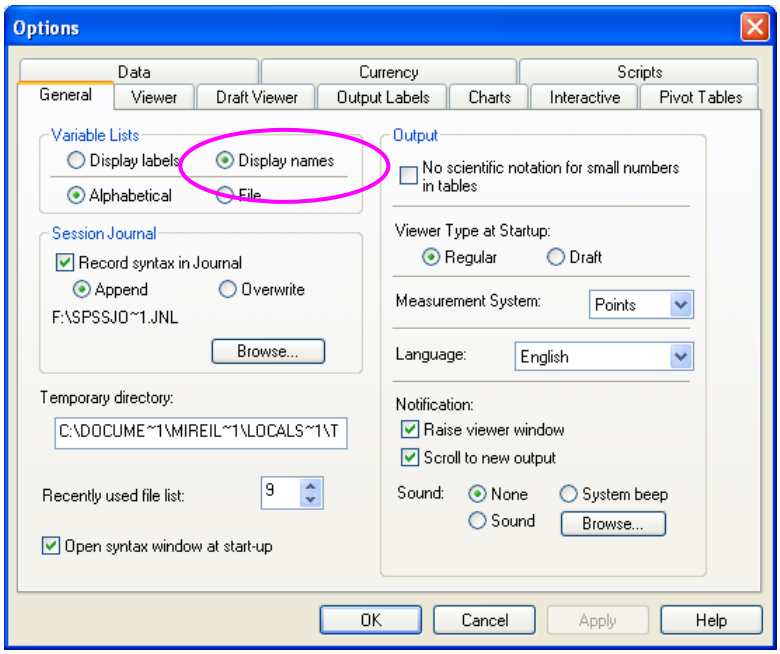
3.1 Practical tips on SPSS set-up

3.1.1 Displaying variable names and commands in analysis output

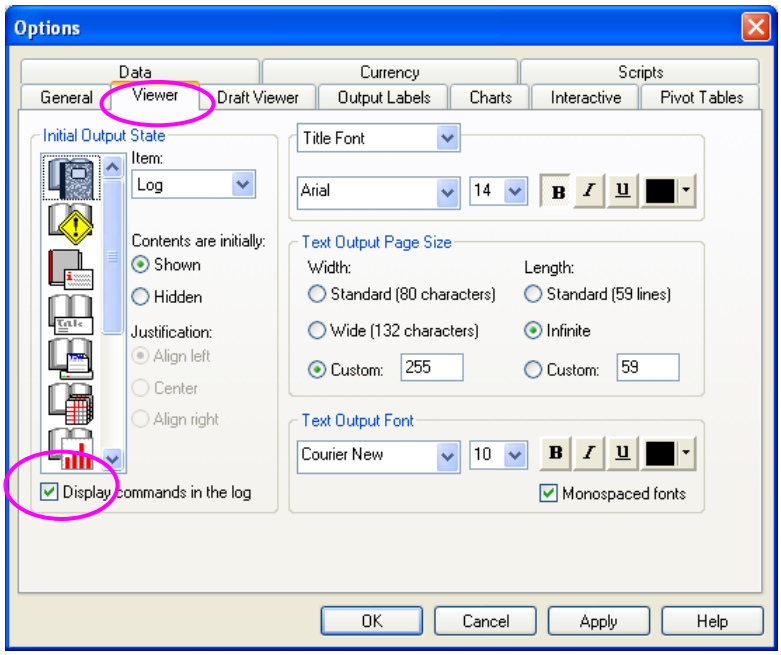
In the menu click on 'Edit / Options':



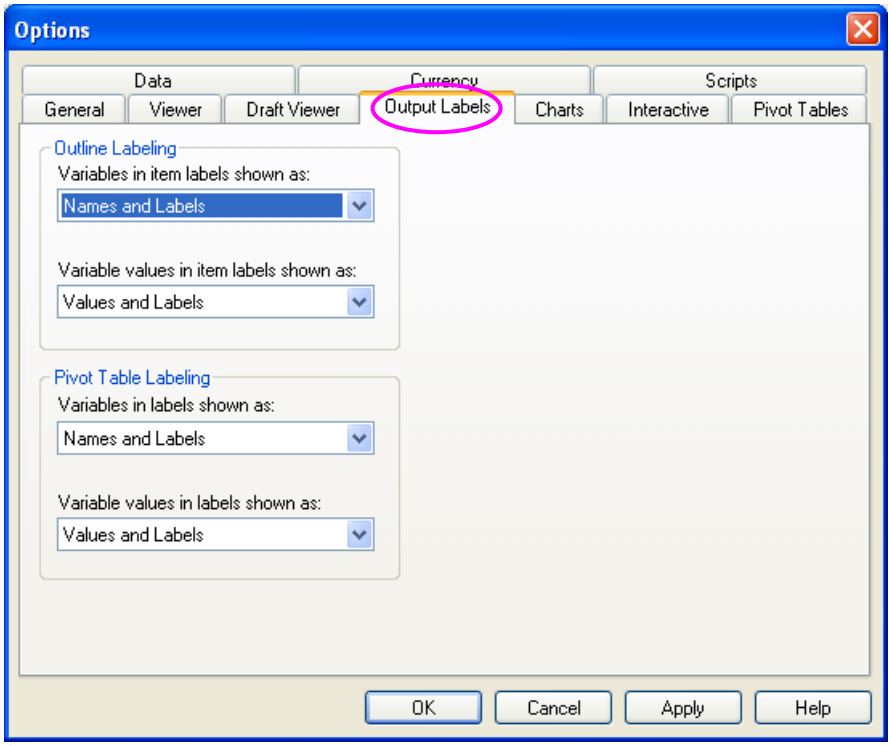
The 'Options' dialog box opens at the 'General' tab. Select the option 'Display names' in the 'Variable Lists':



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.



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



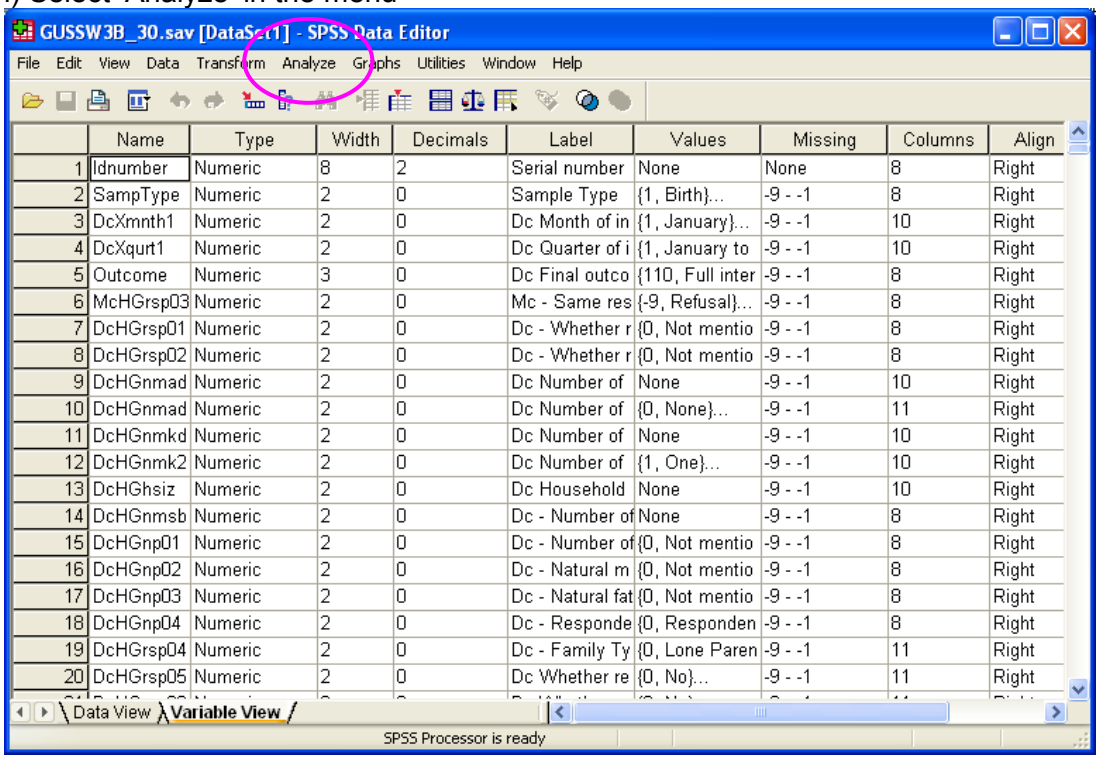
This means that both the variable names, labels and values are shown in any output produced by your analysis.

3.2 Examples of frequencies & cross-tabs via menu

3.2.1 Simple frequency

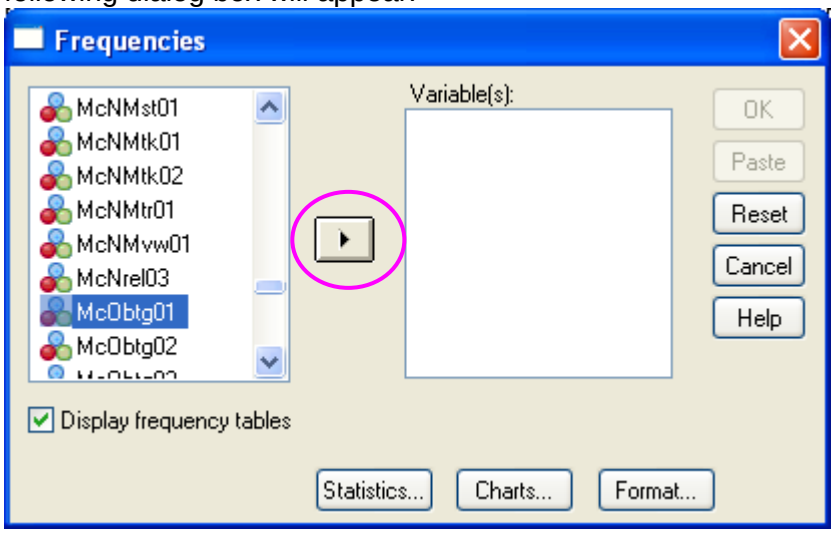
This example demonstrates how to run a simple frequency which will show the proportion of parents who reported attending a parent and toddler group when the cohort child was aged three. **It is based on BC1 Sweep 3 dataset**.

i) Select 'Analyze' in the menu

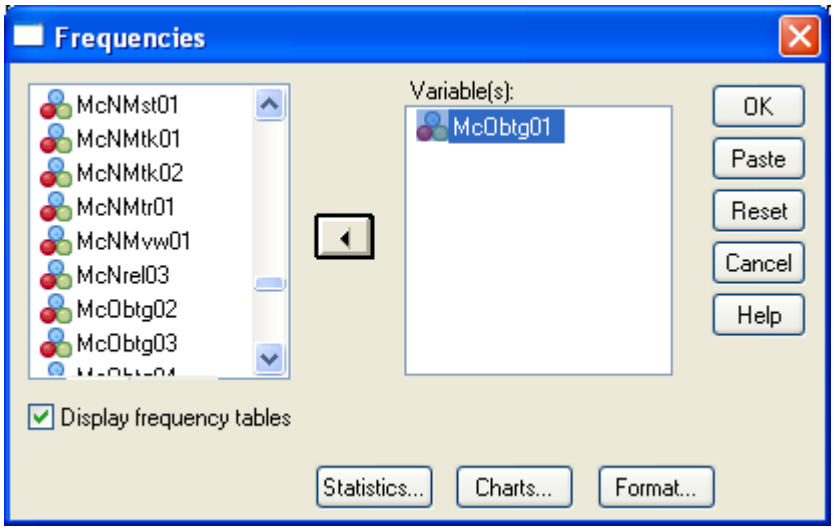


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	Idnumber	Numeric	8	2	Serial number	None	None	8	Right
2	SampType	Numeric	2	0	Sample Type	{1, Birth}...	-9 - -1	8	Right
3	DcXmnth1	Numeric	2	0	Dc Month of in	{1, January}...	-9 - -1	10	Right
4	DcXqurt1	Numeric	2	0	Dc Quarter of i	{1, January to	-9 - -1	10	Right
5	Outcome	Numeric	3	0	Dc Final outco	{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	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	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	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 fat	{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 Paren	-9 - -1	11	Right
20	DcHGrsp05	Numeric	2	0	Dc Whether re	{0, No}...	-9 - -1	11	Right

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



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



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

McObtg01 Mc - Attend group in last 12 months

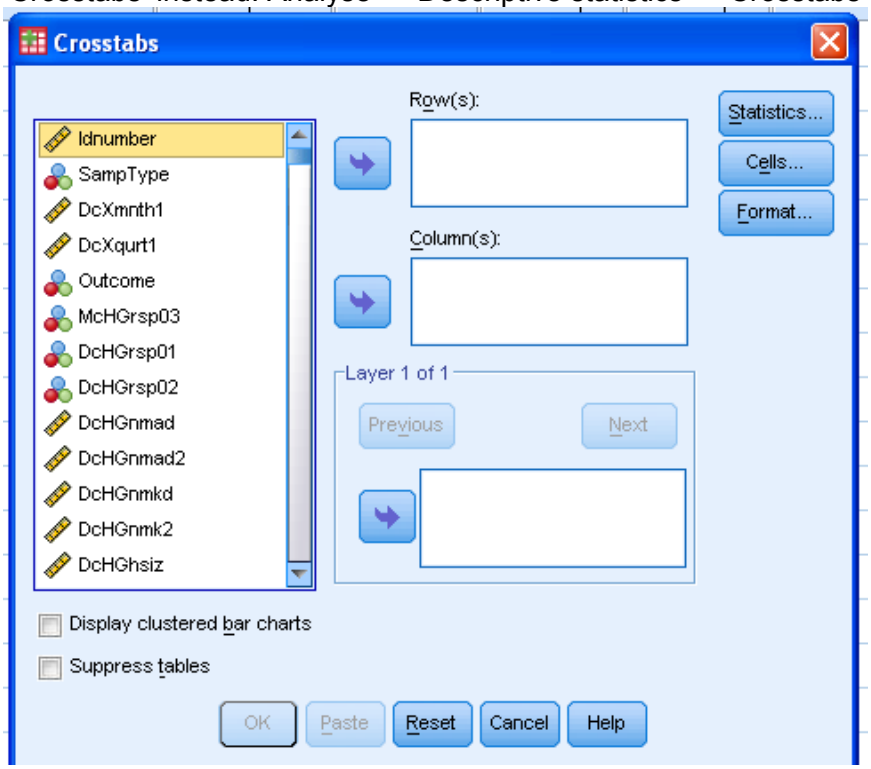
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	568	45.6	45.6	45.6
	2 No	678	54.4	54.4	100.0
	Total	1246	100.0	100.0	

3.2.2 Simple crosstab – attendance at parent and toddler groups

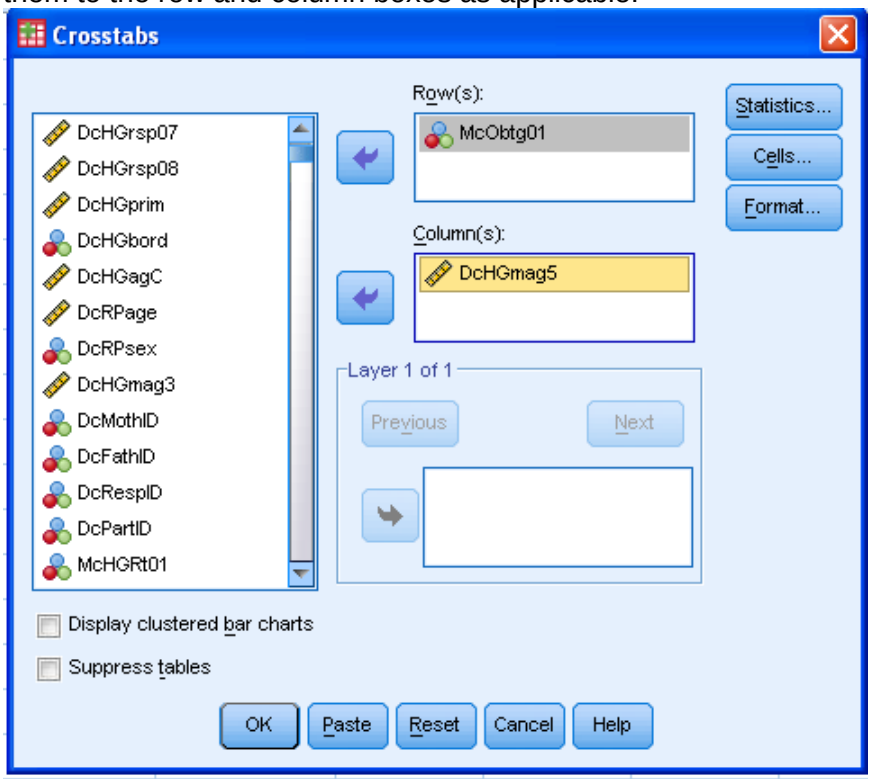
If you want to find out whether or not attendance at parent and toddler groups is different amongst parents with different characteristics, the simplest way to do so is using a cross-tabulation.

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**.

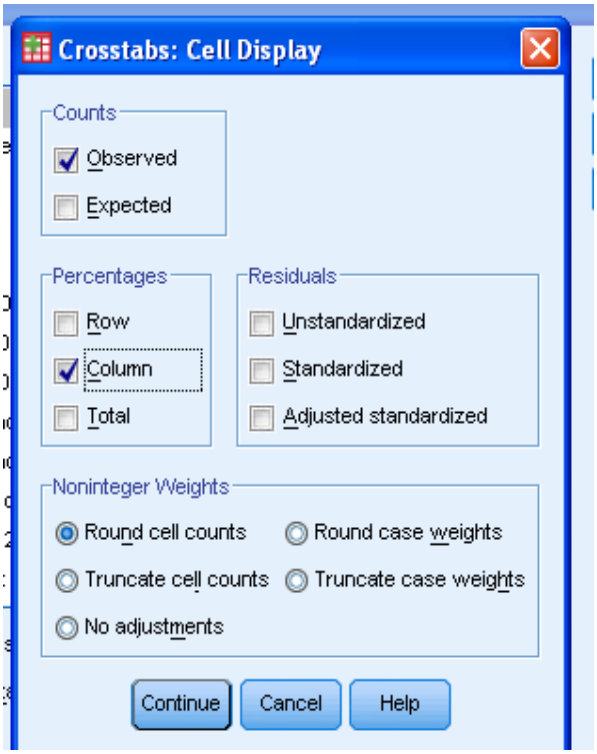
v) To do cross-tabs, use the same menu as for frequencies above but select 'Crosstabs' instead: Analyse → Descriptive statistics → Crosstabs



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:



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

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

McHgen01 Mc - Childs general health

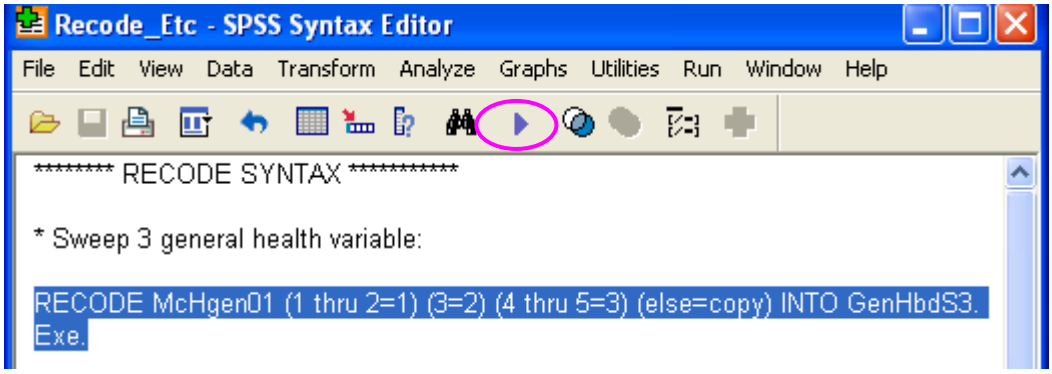
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ...very good,	898	72.1	72.1	72.1
	2 good,	288	23.1	23.1	95.2
	3 fair,	56	4.5	4.5	99.7
	4 bad,	4	.3	.3	100.0
	Total	1246	100.0	100.0	

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.



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:

GenHbdS3 Childs general health - banded

		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	

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*

GenHS2S3

		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	

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).

GenHS2S3 Childs general health evolution Sw2 to Sw3

		Frequency	Percent	Valid Percent	Cumulative 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		

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 INCLUDE 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.
```

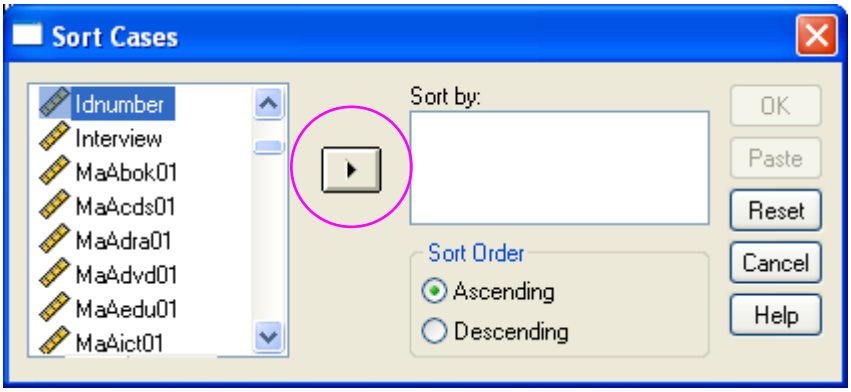
4.2 Merging Datasets

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.

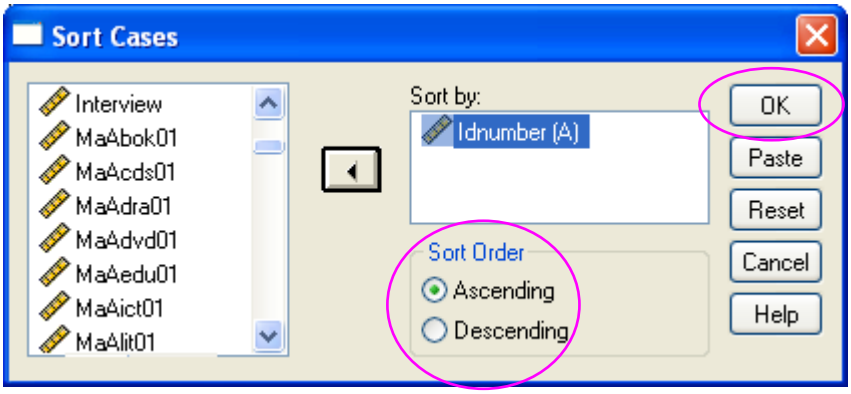
i) Open the dataset you want to merge data into: in the example below it is the Sweep 1 birth cohort dataset

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

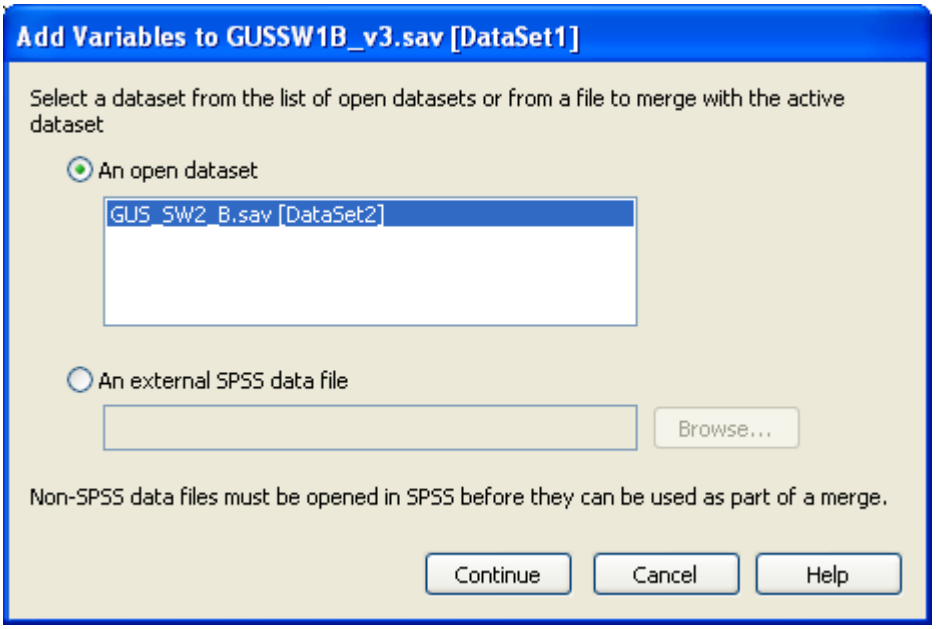


- Click 'OK'

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

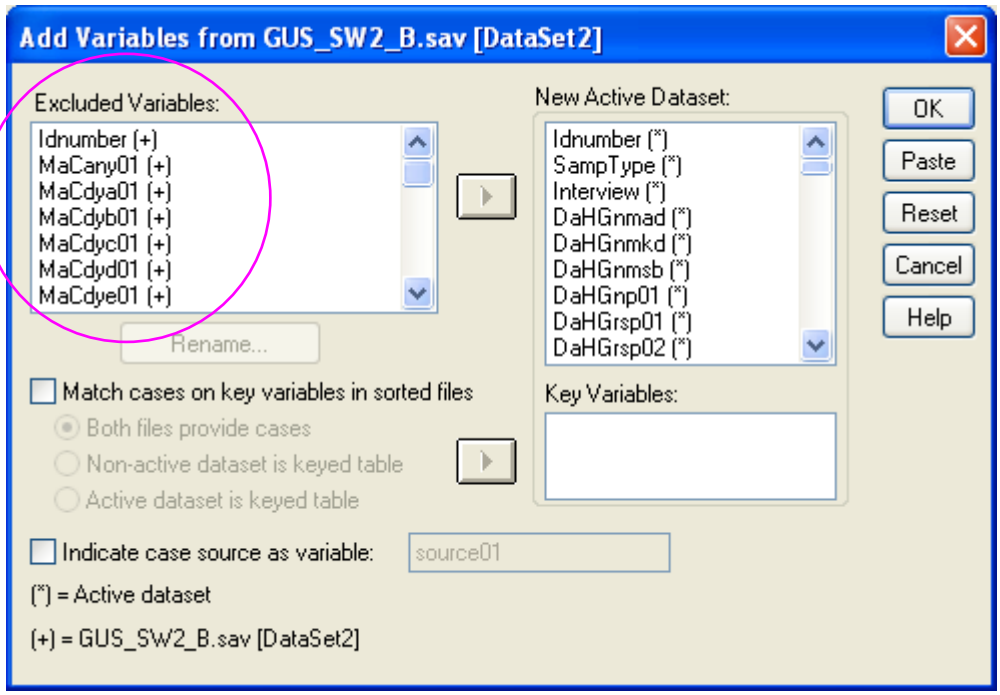
v) In the dialogue box, unless the dataset from which you want to merge is already open, select the button for 'An external SPSS data file' and click 'Browse'. If the dataset is open then select it in the 'open dataset' box (as below).



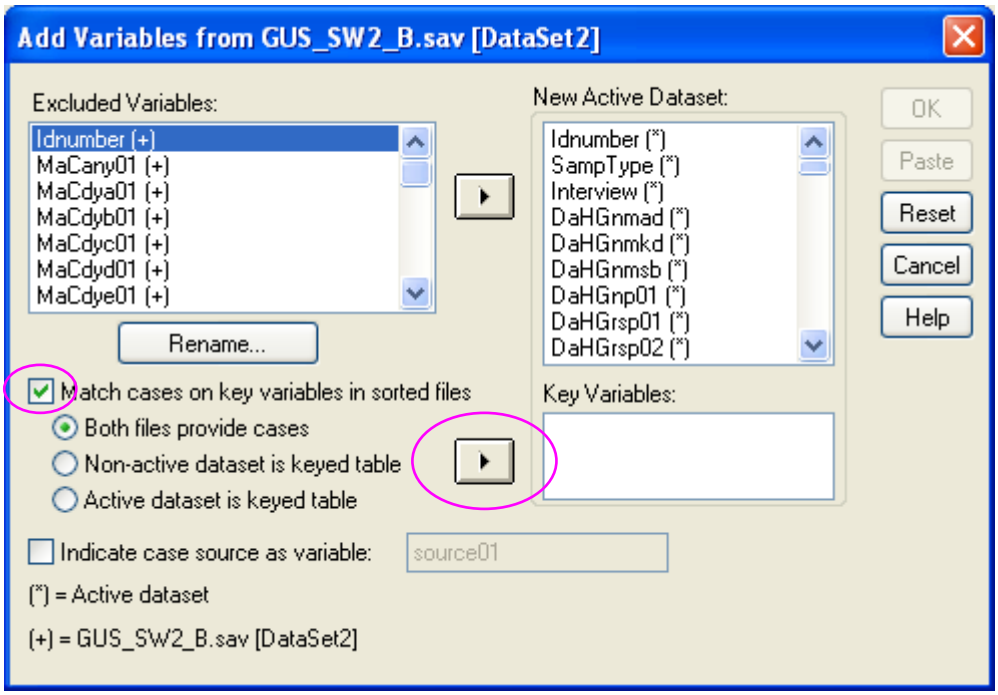
vi) If not already open, browse to the dataset of interest and double-click on it

vii) Click 'Continue'

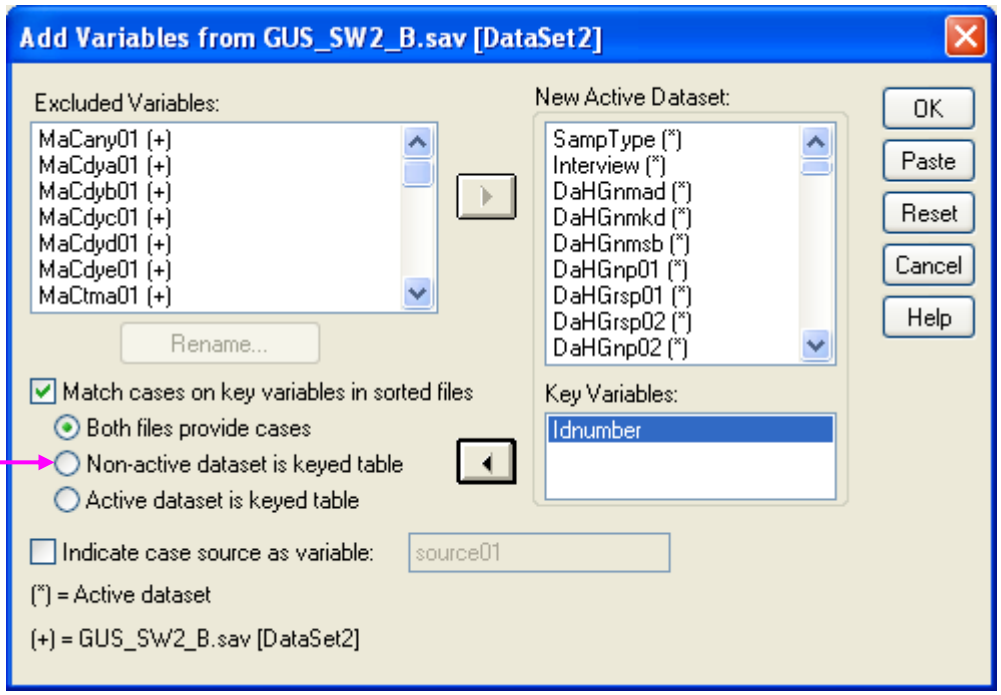
viii) The following dialog box will come up; in this example you can see that there is a big list of 'Excluded Variables' on the left, which are the variables shared by both datasets, instead of just the expected variable 'IDnumber'. This is due to the feed forward process: the archived datasets from 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.



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':



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



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.

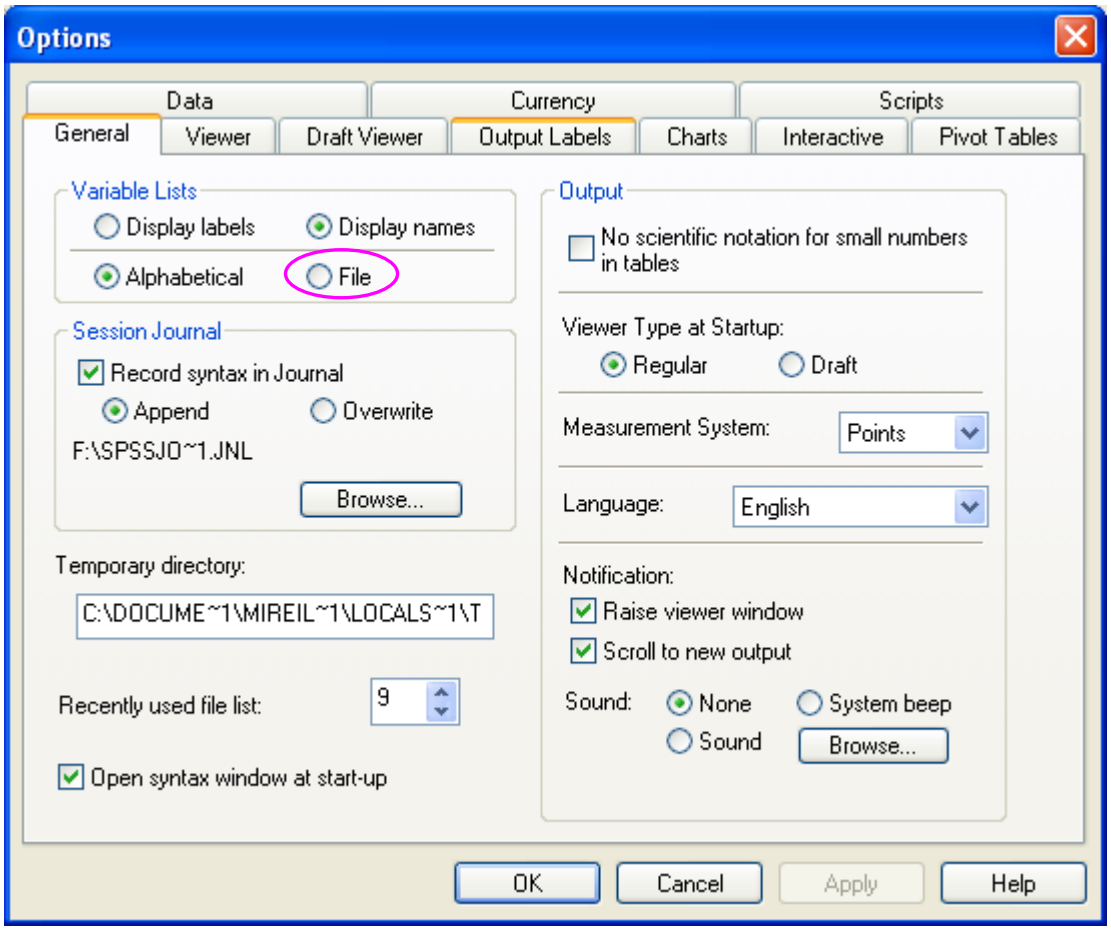
xii) In this example you need to select Option 2 'Non-active dataset is keyed table'

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

5 Data Management - Additional tips

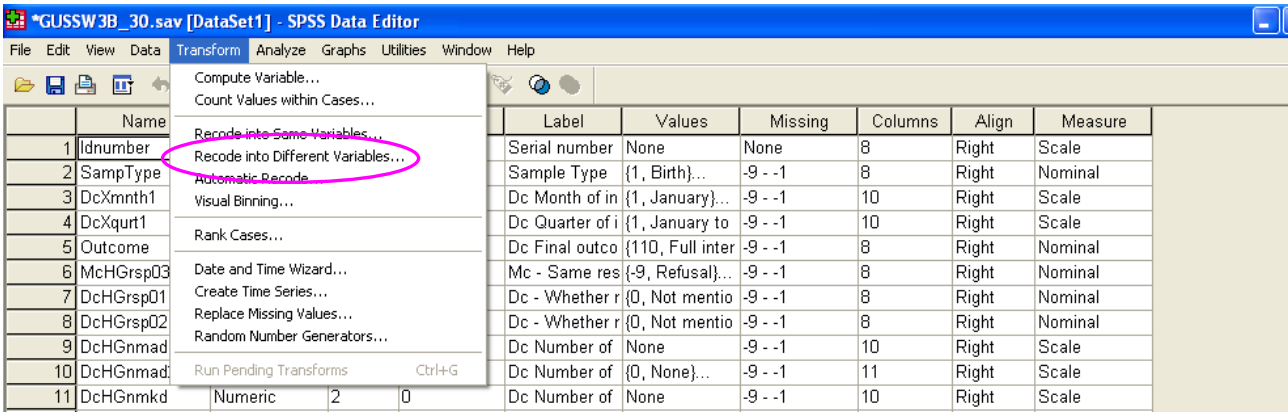
5.1 To change the order of variable names in dialogue boxes

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

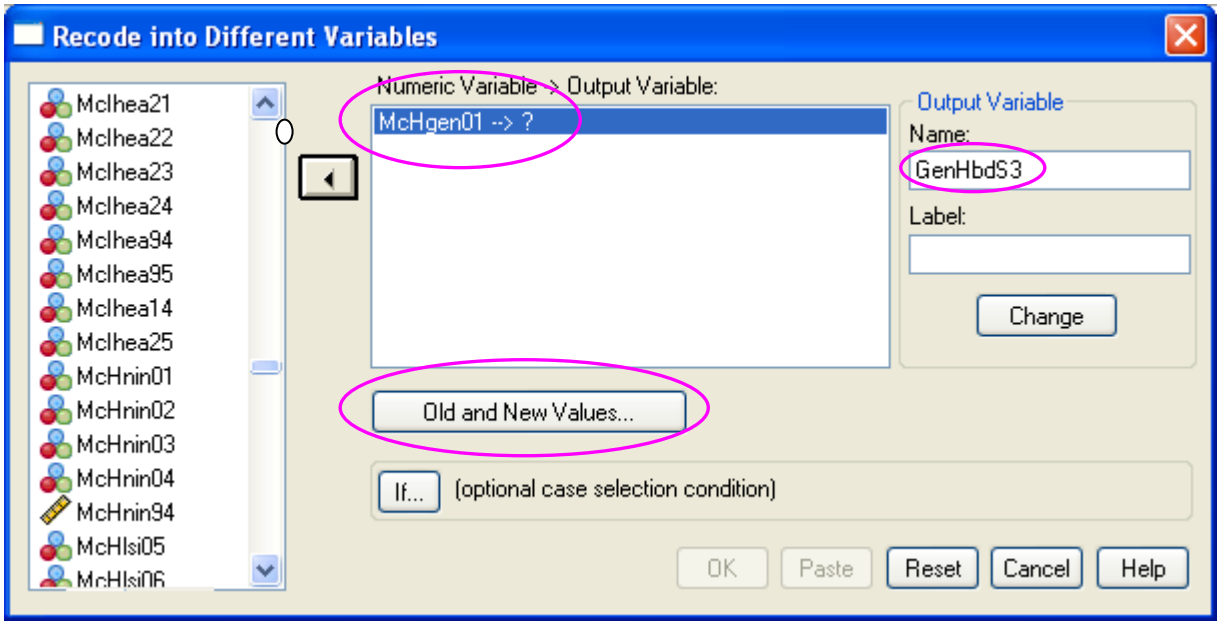


5.2 RECODE a variable via the menu

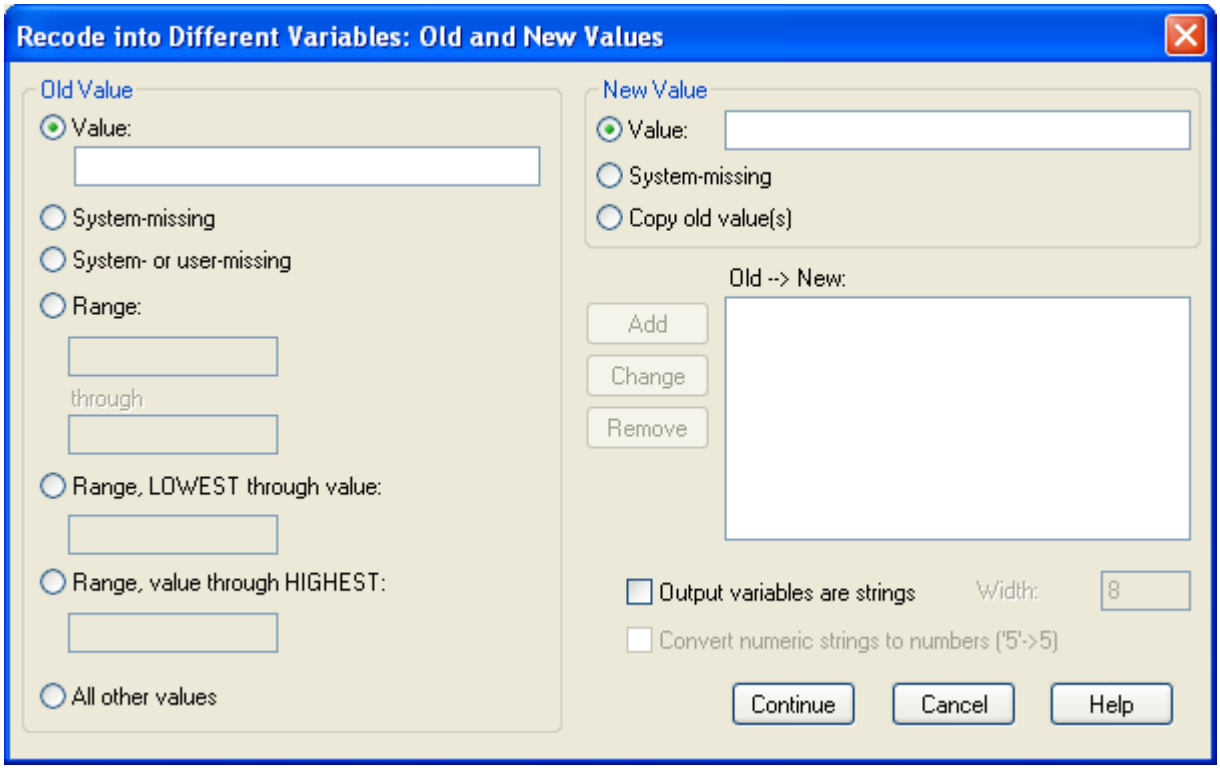
Select ‘Transform’ and ‘Recode into Different Variables’ for the example chosen:



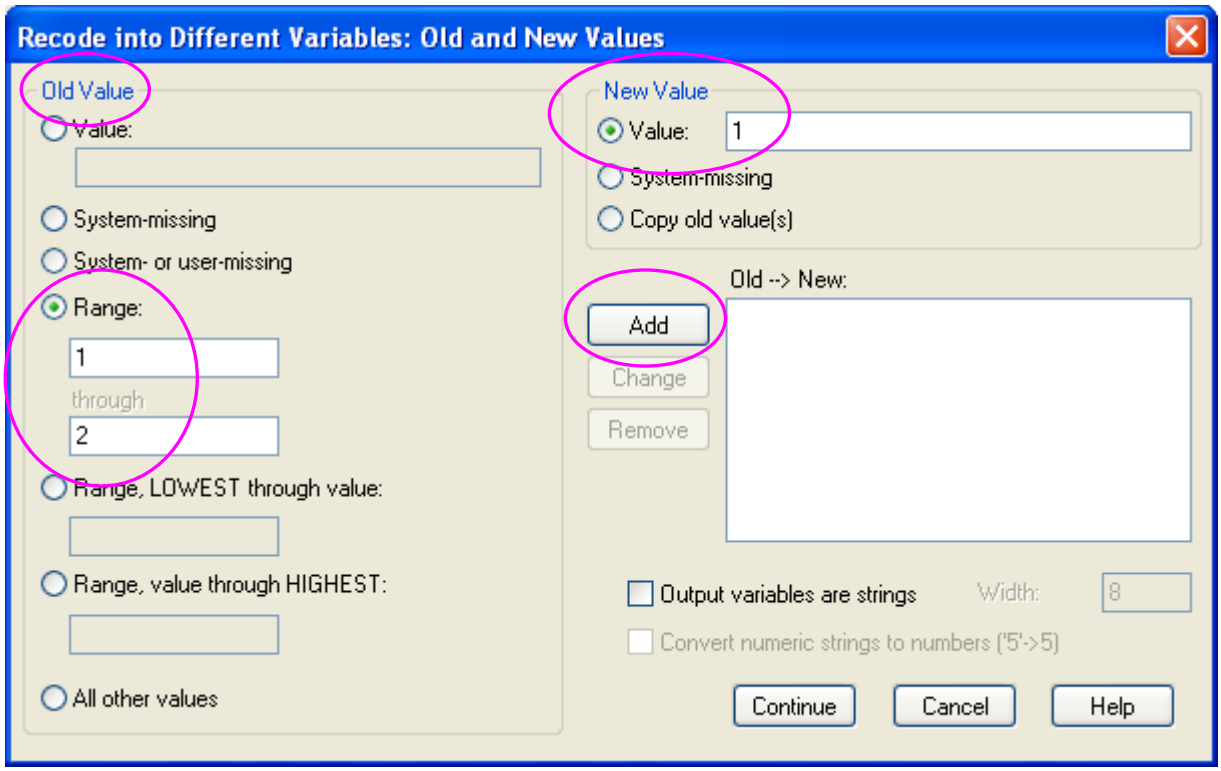
Select the variable to be recoded from the list on the left-hand side, and type the name of the new variable under ‘Output Variable/Name:’ (you can also add a Label for the new variable)



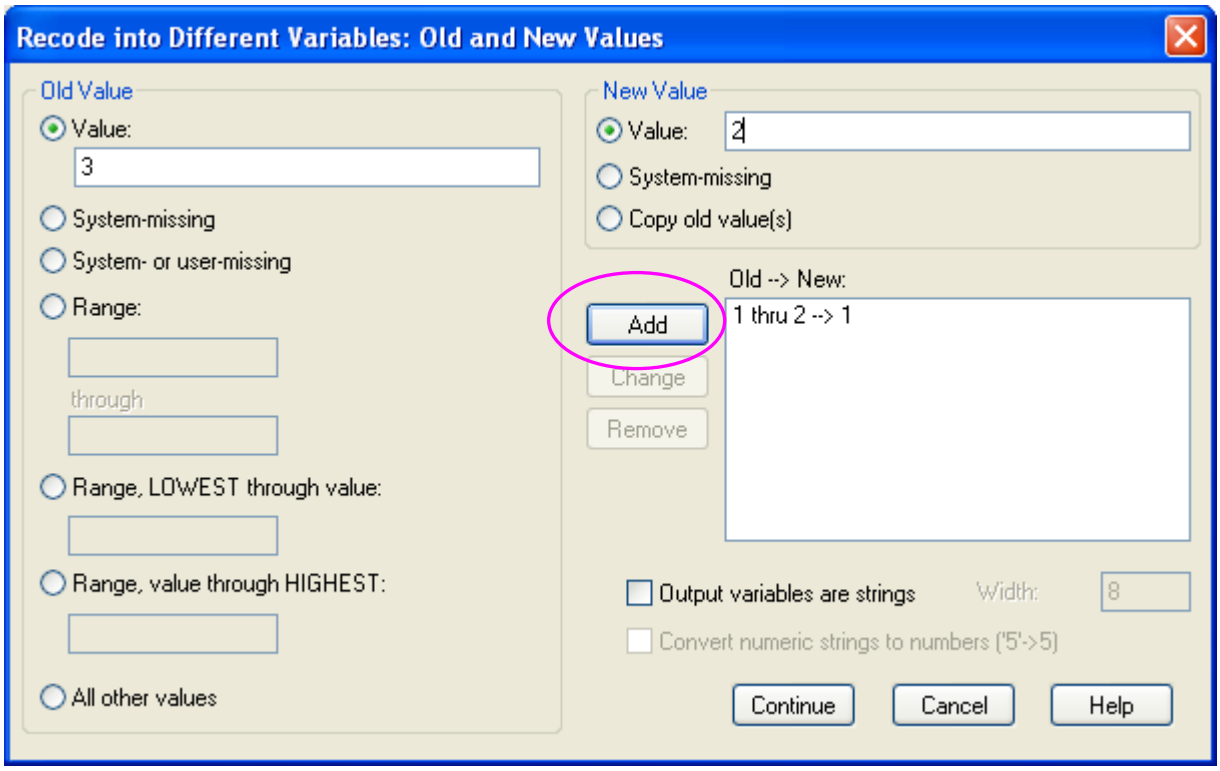
To change the values click on ‘Old and New Values’ button, the following dialog box comes up:



Use ‘Range’ or ‘Value’ depending on transformation required

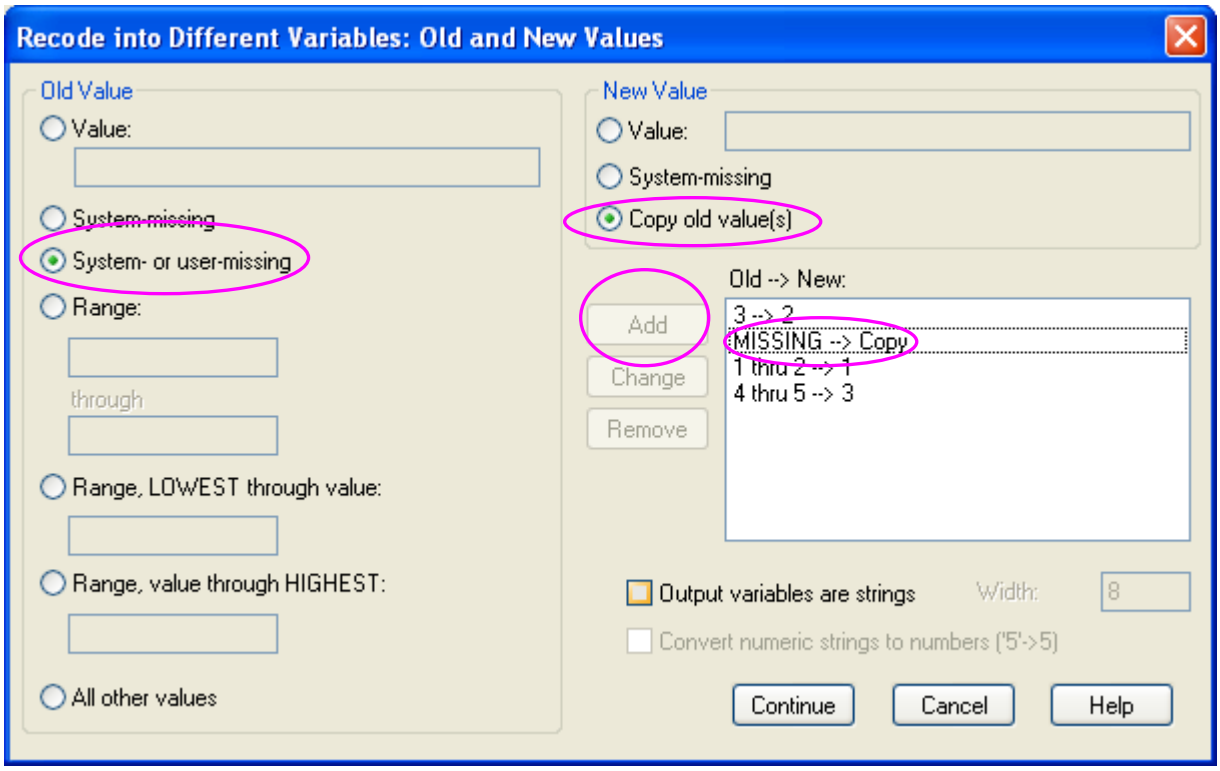


Click on ‘Add’:



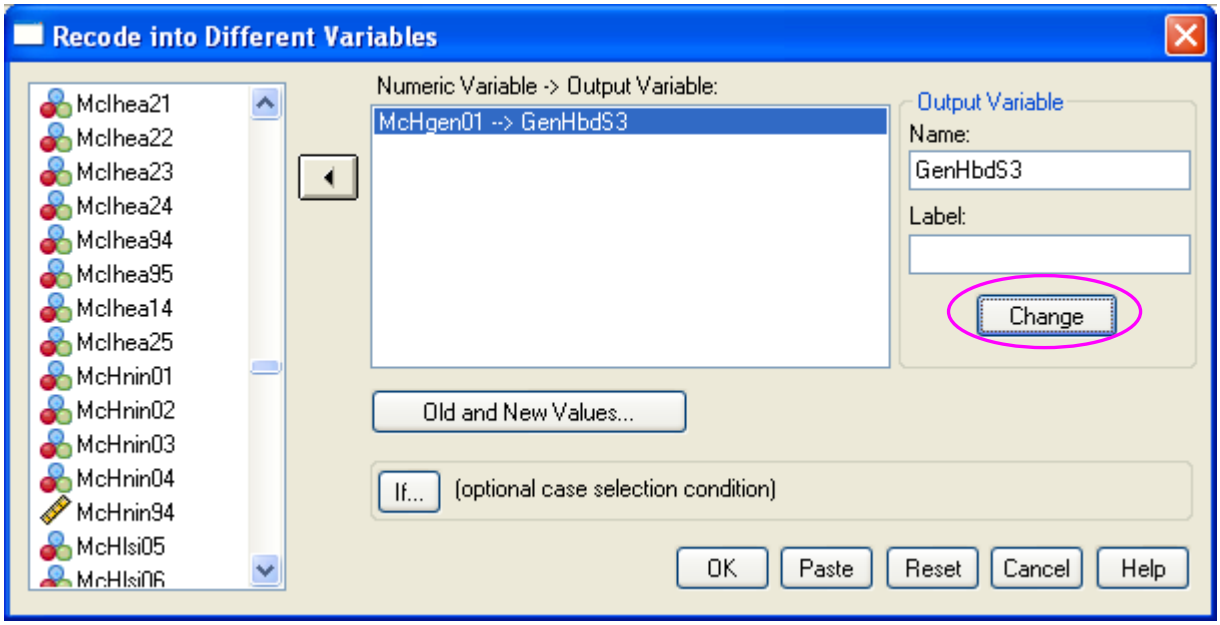
Repeat for other values to be grouped in this example

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



Click on ‘Continue’

Click on ‘Change’ in the next screen

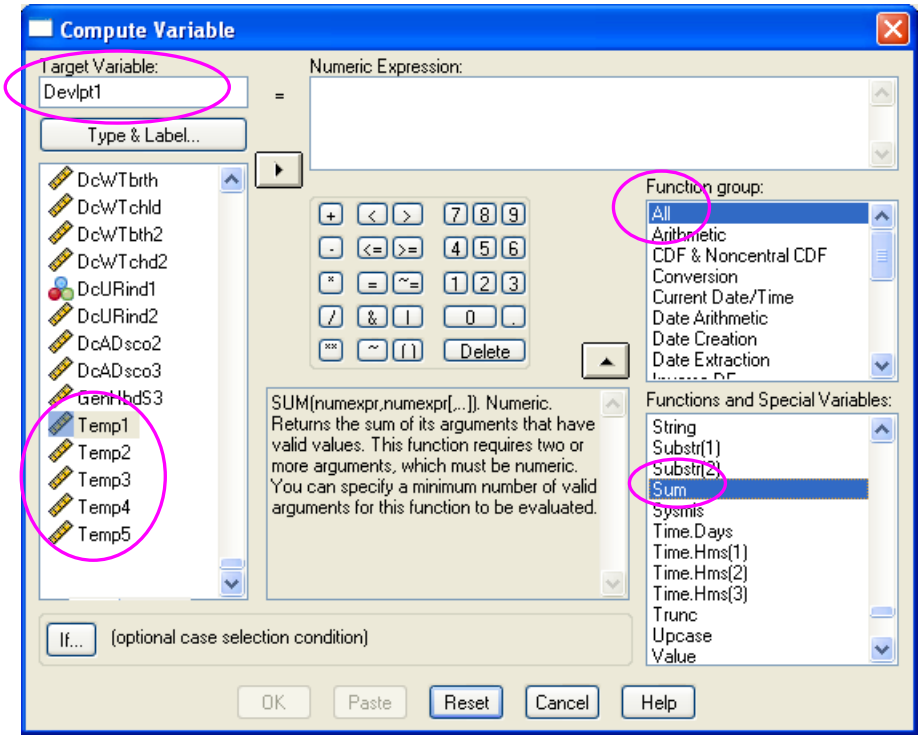


Then ‘OK’

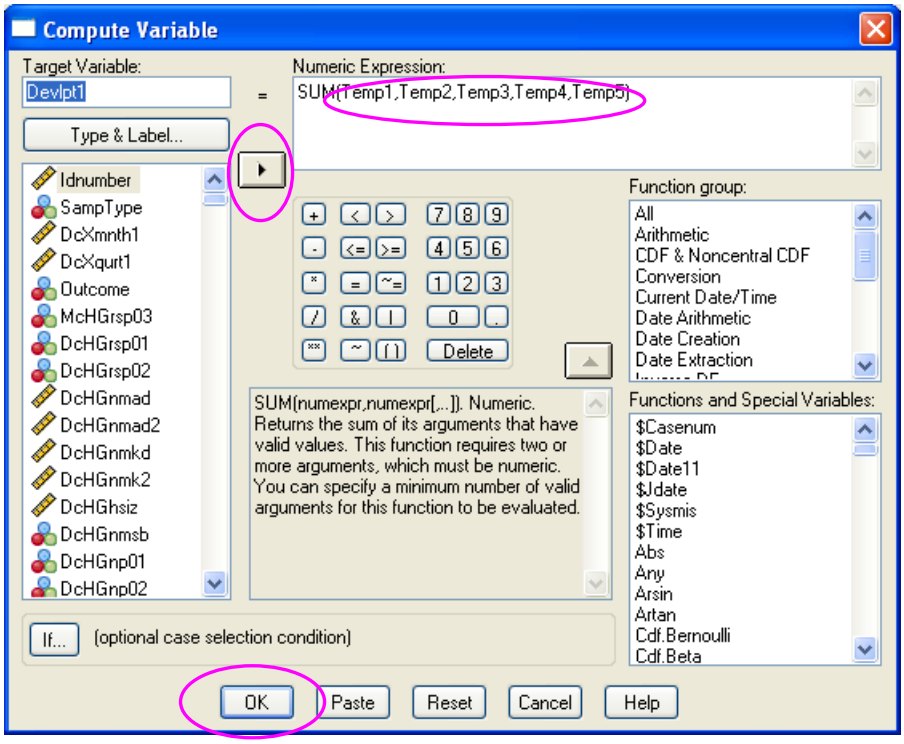
5.3 COMPUTE a variable via menu

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:



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').

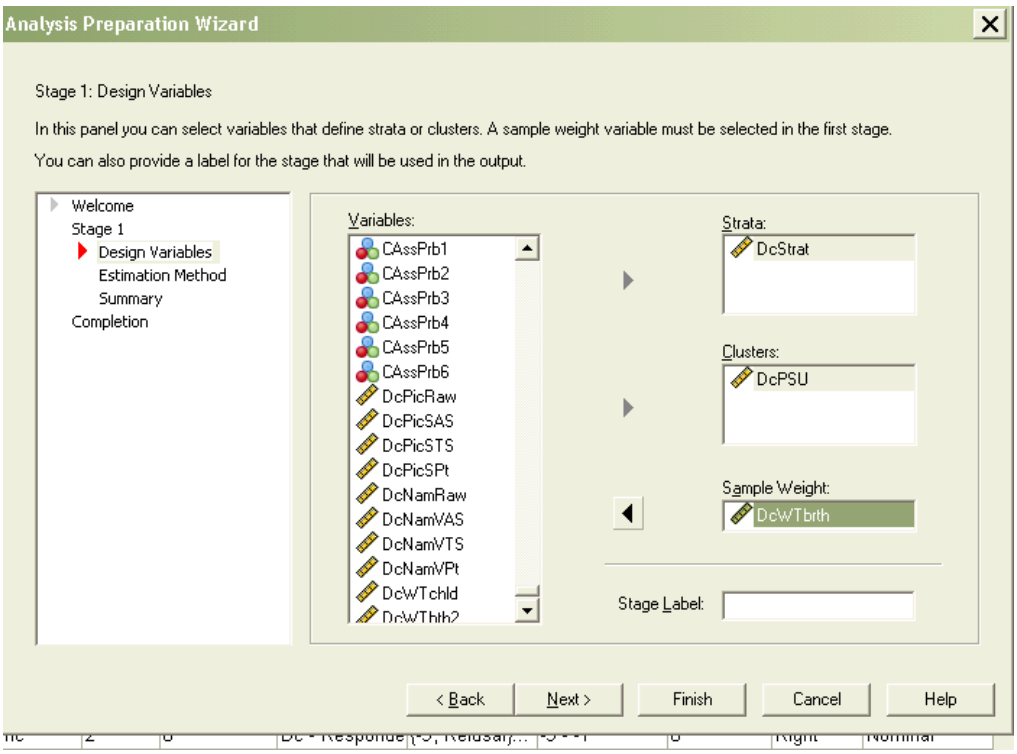
1. You need three components to create a plan file
   a) A weight variable
   b) A stratification variable
   c) A cluster variable

   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
   → Select: Create a plan file
   → 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)
   → Click NEXT
   → Insert the relevant variables as shown below in the screenshot, click NEXT



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

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

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

→ 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.

```
* Analysis Preparation Wizard.
CSPLAN ANALYSIS
 /PLAN
FILE='C:\temp\sw3bcxs.csaplan'
 /PLANVARS
ANALYSISWEIGHT=DcWTbrth
 /SRSESTIMATOR TYPE=WR
 /PRINT PLAN
 /DESIGN  STRATA= DcStrat
CLUSTER= DcPSU
   /ESTIMATOR TYPE=WR.
```

You can now use this syntax as the basis for any plan file you need to create.

All you need to change is:

- the filename
- the weight variable
- the strata variable (dastrat, dbstrat, dcstrat etc)
- the cluster variable (dapsu, dbpsu, dcpsu etc)

To suit the datasets you are using and analysis you are running.

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**
  - 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)
  - 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**

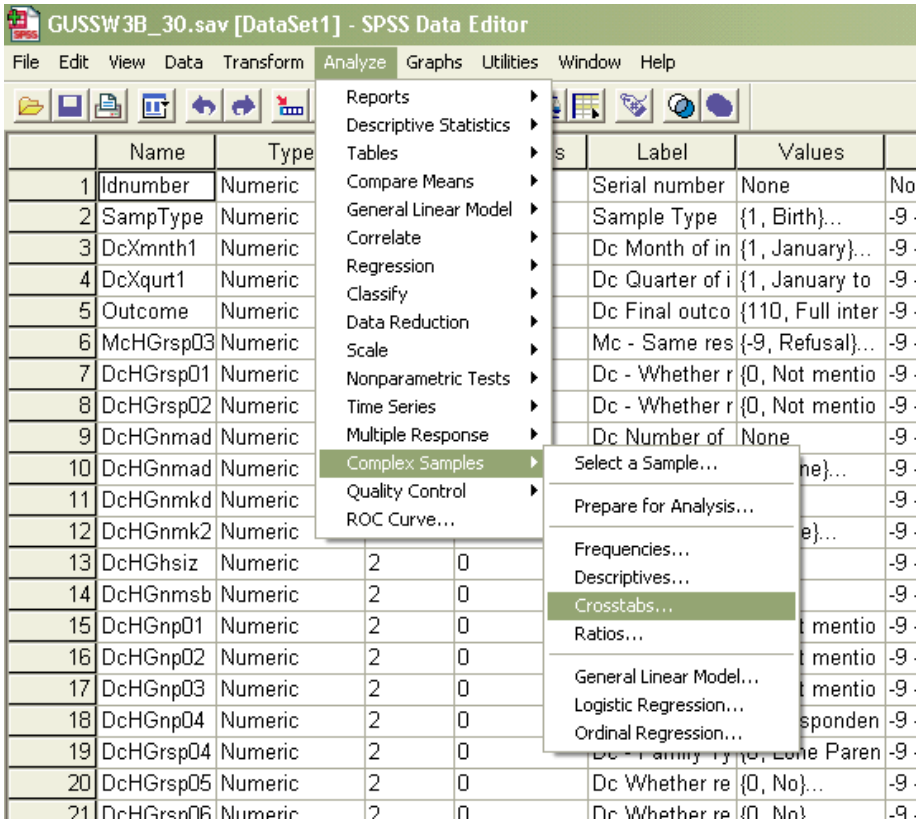
6.2 Basic analysis using the Complex Samples (CS) Module in SPSS

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 cross-tabulation 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.

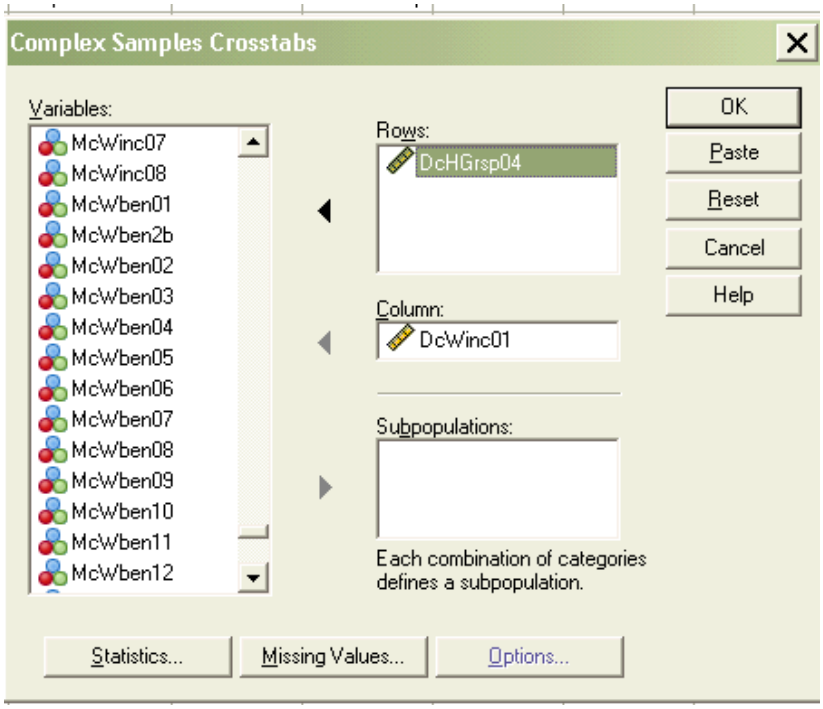
2. On the menu tool bar select: Analyze/Complex Samples/Crosstabs



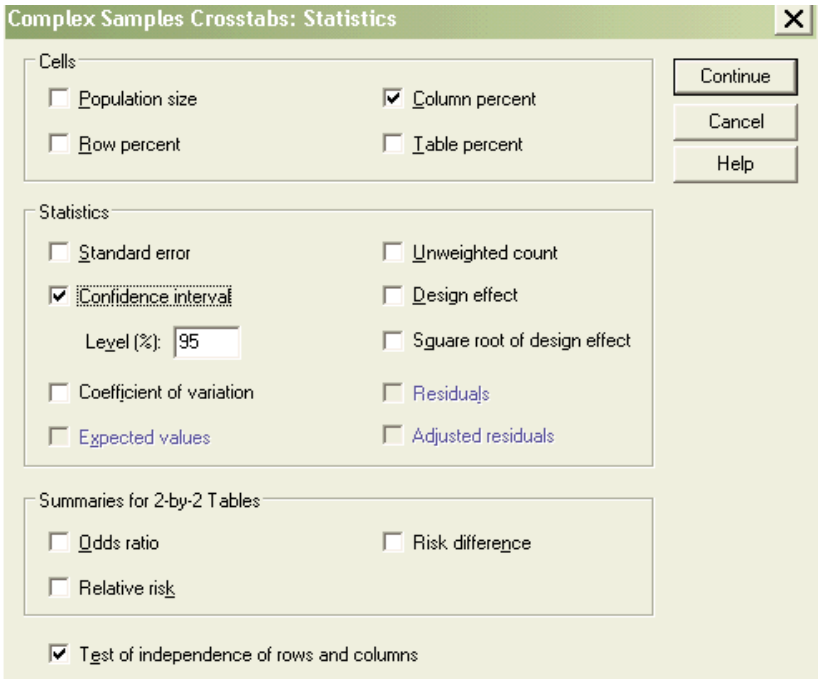
3. In the 'Complex Samples Plan for Crosstabs Analysis' dialog box, the plan file you have just created may already appear in the 'File' box, if not, click 'Browse' and navigate to the correct plan file. [REMEMBER – the plan file you require is dependent on the data you are using and the analysis you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



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



- 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 Type * Dc Household income - banded

Dc - Family Type				Dc Household income - banded				
				Up to £14999 per year	From £15000 to £25999 per year	From £26000 to £43999	£44,000 and above	Total
Lone Parent	% within Dc Household income - banded	Estimate		62.9%	15.5%	2.7%	1.2%	20.2%
		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 income - banded	Estimate		37.1%	84.5%	97.3%	98.8%	79.8%
		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 income - banded	Estimate		100.0%	100.0%	100.0%	100.0%	100.0%
		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%

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 = *DcHGrsp04* BY *DcWinc01*

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 = 'C:\temp\sw3bcxs.csaplan'

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):

(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)

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 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 older	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?

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	469	37.6	37.6	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.

ALaURin2 SE urban-rural classification

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Large urban	321	37.5	37.5	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	

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	168	14.8	15.0	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.

DcWsta02 Dc Mothers employment status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Childs mother working - full-time	125	18.3	18.8	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		

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.

MbHgen01 Mb - Child's general health

		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	

Syntax for birth cohort would be:

Weight by ddwtbrth.

fre mdhgen01.

Exe.

MdHgen01 Md - Childs general health

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ...very 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	

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 Mc - Household has cont use of vehicle		
			1 Yes	2 No	Total
DcADsco2 ALc - SIMD 2006 quintiles	1 0.9449 - 7.7446 - least deprived	Count	225	7	232
		% within DcADsco2 ALc - SIMD 2006 quintiles	97.0%	3.0%	100.0%
	2 7.7472 - 13.5627	Count	217	20	237
		% 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 - 89.0941 - most deprived	Count	156	129	285
		% 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.

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

			MaZveh01 Do you, or any members of your household, at present own or have continuous use of any motor vehicles (SW1)		
			1 Yes	2 No	Total
ALaSNimd SIMD 2006 Quintiles	1 0.9449 - 7.7446 - least deprived	Count	152	3	155
		% within ALaSNimd SIMD 2006 Quintiles	98.1%	1.9%	100.0%
	2 7.7472 - 13.5627	Count	153	12	165
		% 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 - 33.6982	Count	97	42	139
		% within ALaSNimd SIMD 2006 Quintiles	69.8%	30.2%	100.0%
	5 33.7252 - 89.0941 - most deprived	Count	117	100	217
		% 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%

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.
   ```

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

			Mc - Play outdoors 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%

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 Whether resp uses regular CCare at Sw3		
			Yes	No	Total
Sw1 Whether using any childcare for cohort child	Yes	Count	10.306	.898	11.205
		% 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)