GUI Data Workshop

Information Sheet 4: How to match 9 year and 13 year files

This information sheet provides a guide on how to match the Child Cohort Wave 1 (at 9 years) and Wave 2 (at 13 years) files. Please note this information sheet is based on SPSS Version 22.

To match the Wave 1 and Wave 2 datafile you must follow these steps:

1. Open the Wave 1 datafile - XGUI Data_9YearCohort.sav

2. Click Data \rightarrow Merge \rightarrow Merge Files \rightarrow Add variables`

ta xo	🙀 XGUI Data_9YearCohort.sav [DataSet3] - IBM SPSS Statistics Data Editor											
<u>F</u> ile	<u>E</u> dit	View	Data 1	ransform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add- <u>o</u> ns	Window	<u>H</u> elp		
🚝			🧊 Defin	e <u>V</u> ariable P	roperties			H 👯		-		
			🏄 Set M	leasuremen	t <u>L</u> evel for	Unknown						
		zio	<u>ि</u> ору	Data Prope	rties			azid08	azid	00	071	d10
1		210		Custom Attr	i <u>b</u> ute				1	1	azı	(
			🗟 D <u>e</u> fin	e Dates				1	. 1		(
3	}	E Define Multiple Response Sets							1	1		(
4	ł	ldentify Duplicate Cases							1	1		(
5	5	Recompare Datasets							1	1		(
6	;		Sort (- Cases					1	1		(
7	'		_	/ariables					1	1		(
8			Trans	_					1	1		(
9				e Files			•	🔝 Add Cas	1	1		(
1			Restr							1		(
1	-		Aggre					Add Vari		1		(
1				Weights					1	1		(
1				ensity Score	Matching				1	י 1		
				Control Mat	_				1	1		
1					ching				1	. 1		(
1	7		Ropy	_					1	1		(
1	8			nto Files					1	1		(
1	9	Split File							1	1		(
2	0		_	t Cases					1	1		(
2	1		III Weigl	ht Cases					1	1		(
2			22000		0		2		1	1		(
2	3		23000		0		2		1	1		(

3. This will bring up the following dialog box. Click browse and select the location of the Wave 2 datafile – XGUI Data_ ChildCohortWave2.sav. Click Continue.



4. In the following dialog box tick 'match cases on key variables in sorted files', 'cases are sorted in order of key variables in both datasets', and using the arrow move the variable 'zid01' into the Key Variables box. Also tick 'Indicate case source as variable' and type a name for the variable into the box (the default variable name is 'source01'. This has been changed to 'wave2' in the example below). This will create a new variable which is coded 1 if the case is in Wave 2 and 0 if the case is not in Wave 2. Click 'ok'. There are 8,568 cases in Wave 1 and 7,525 cases in Wave 2. Longitudinal analysis can be conducted on cases for which we have both Wave 1 and Wave 2 data (7,525 cases).

🔄 Add Variables from P:\NLSCI\chdata\All AMFs\Child Cohort\XGUI Data_ChildCohortWave2.sav				
Excluded Variables:		New Active Dataset:		
Rename Image: Match cases on key variables Image: Cases are sorted in order of key variables in both datasets Image: Cases are sorted in order of key variables in both datasets	*	azid06(*) azid07(*) azid08(*) azid09(*) azid10(*) azid11(*) azid12(*) azid13(*) azid14(*)		
 Non-active dataset is keyed table Active dataset is keyed table 	+	zid01		
 <u>Active dataset is keyed table</u> <u>B</u>oth files provide cases 				
Indicate case source as variable: wave2				
(*)=Active dataset				
(+)=P:\NLSCI\chdata\All AMFs\Child Cohort\XGUI Data_ChildCohortWave2.sav				
OK Paste Reset Cancel Help				

5. The following warning will appear. In order to match the files they both must be presorted by the key variable used for matching – in this case 'zid01'. Click 'ok'.

ta IBM	SPSS Statistics 22
	Warning: Keyed match will fail if data are not sorted in ascending order of Key Variables.
	OK Cancel

6. The matched file will contain 8,568 cases with all the variables from Wave 1 (prefixed with 'a') and all the variables from Wave 2 (prefixed with 'b'). The Wave 2 variables will be system missing (sysmis) for any cases which are not included in Wave 2 (there are 7,525 valid cases in Wave 2). There will also be an identifier variable 'wave2' which will flag which cases responded in Wave 2.

7. This matched file should be saved under a new name (for e.g. XGUI Data_ChildCohortWave1andWave2.sav). <u>Simply saving the file will overwrite the original Wave 1 file.</u>