British Social Attitudes 1983-2014: Beta version of cumulative file

Copyright © 2016 John F Hall

Incompatible type and format attributes have been resolved and data from all 32 waves are now in a cumulated file. The latest version of the combined file is **bsa1983-2014jfh2.sav** sorted in reverse year order: size 1.04 gb; number of cases 95630; number of variables 10765.

Some variables are unique to a single wave (but there could be duplicates with different names in other years). Other variables have been used in two or more waves, but their metadata will always be taken from the most recent wave in which they figured. The variable **[year]** was in position 848, but I have moved it to position 1. In the cumulative file, blocks of variables are in order within the year in which they were last used. The file was created in reverse year order with an intermediate variable **[yearorder]** using SPSS command: **COMPUTE** yearorder = 2015 – year.

Examples

1: Variable used in four waves: most recently 1995

year Year of Interview * brtirsde R on British/Irish govt side in arg B107NI95b Crosstabulation Count

		brtirsde R on British/Irish govt side in arg B107NI95b								
				1 British	2 Irish	3 It	4	8 Don't know	9 Not	
		-2	-1	Govt	Govt	depends	Neither	/Can't say	Answered	Total
year Year	1989	0	0	908	74	171	189	163	11	1516
of	1991	0	0	834	94	151	241	142	11	1473
Interview	1993	0	1461	824	72	159	295	123	11	2945
	1995	2406	0	634	99	208	224	57	5	3633
Total		2406	1461	3200	339	689	949	485	38	9567

[NB: In this and the following tables codes -2, -1, 8 and 9 not yet declared as missing.]

2: Variable used in two waves: most recently 1995

year Year of interview * crowded1 Limit crowds:cut down/close carparksB485 Crosstabulation Count

	crowded1 Limit crowds:cut down/close carparksB485								
		1 Strongly	2 In	3 Neither	4	5 Strongly	8 Don't	9 Not	
		favour	favour	favour/against	Against	against	know	answered	Total
year Year of	1994	132	363	252	303	100	14	1	1165
interview	1995	116	404	271	304	55	21	1	1172
Total		248	767	523	607	155	35	2	2337

3: Variable used in three waves: most recently 1995

year Year of interview * jobchk Have you *ever* had a job? Q923 Crosstabulation

Count	

		jobchk	jobchk Have you *ever* had a job? Q923					
		-1 Skp,work +wait work	1 Yes	2 No, never	9 Not answered	Total		
year Year of	1993	1381	1462	94	8	2945		
interview	1994	1702	1679	86	2	3469		
	1995	1747	1773	110	3	3633		
Total		4830	4914	290	13	10047		

(Draft only: 18 April 2016)

Some variables seem to be used only once:

4: Variable used in one wave only: 1994

year Year of interview * censrds2 Censorship necessary/unnecessary[2]B235h Crosstabulation Count

		censrd	censrds2 Censorship necessary/unnecessary[2]B235h						
					9 Not				
		1 Necessary	2 Unnecessary	8 Don't Know	answered	Total			
year Year of interview	1994	694	173	99	9	975			
Total		694	173	99	9	975			

5: Variable used in one wave only: 1995

year Year of interview * closeurp Europe - R feel close to? A201e Crosstabulation
Count

		C	closeurp Europe - R feel close to? A201e						
						8 Can't			
				3 Not	4 Not	choose/			
		1 Very	2	very	close at	not	9 Not		
		close	Close	close	all	apply	answered	Total	
year Year of interview	1995	38	155	392	322	52	99	1058	
Total		38	155	392	322	52	99	1058	

Building the file

SPSS command sequence was something like the following, starting with 2014 and repeated for each wave:

COMPUTE yearorder = 2015 – year.

ADD FILES file *

/file ' <dataset for previous wave> ' /by yearorder.

This automatically detects and type incompatibilities which can be corrected before:

```
SAVE outfile '~~~ bsa1983-2014jfh2.sav'
/keep year yearorder <first var> to <lastvar>
/by yearorder.
```

Inconsistent coding

The tables below show that the variable **[pressrun]** is coded in two different ways. In 1986 there are only two response categories (1 = Well run; 2 = Not well run) but in 1987 there are four for the same question (1 = Very well run; 2 = Well run; 3 = Not very well run; 4 = Not at all well run).

These codes are incompatible: the only way to compare them is to combine the 1987 codes 1 with 2 and 3 with 4 and change the labels accordingly. In both cases, codes 8 and 9 are not declared as missing, but this is true for all the rating scales in all waves.

year Year of interview * pressrun BQ217 WELL RUN: THE PRESS Crosstabulation

		pressrun BQ217 WELL RUN: THE PRESS						
		1 WELL RUN	2 NOT WELL RUN	3	4	8 DK	9 NA	Total
year Year of	1986	632	647	0	0	13	29	1321
interview Total	1987	25 657	449 1096	<mark>507</mark> 507	185 185	5 18	41 70	1212 2533

Checking on values 3 and 4 for 1987:

temp. select if year = 1987. freq pressrun.

		Frequency
Valid	1 VERY WELL RUN	22
	2 WELL RUN	441
	3 NOT VRY WELL RUN	495
	4 NOT AT ALL WELL RUN	179
	8 DK	5
	9 NA	40
	Total	1181
Missing	0 NO SELF-COMPLETN	194
	System	1391
	Total	1585
Total		2766

I hope there aren't too many of these!

Measurement levels and missing values

A number of serious problems with measurement levels and missing values remain to be resolved.

Incorrect **measurement levels** may be due to SPSS automatically applying "heuristic" levels to variables depending on the number of values encountered, but they could equally be due to automatic archiving software at UKDS. See: <u>British Social Attitudes 2009 - 2014 Measurement levels</u>

Missing values are declared for all variables as negative, mostly in the range **Lo thru -1**, but the original positive values have not been changed and consequently are not treated as missing for analysis purposes. See: <u>British Social Attitudes 2009 to 2014</u> <u>Missing values</u>