Block 3: Analysing two variables (and sometimes three)

3.1.4.1a Income differences work-through

[Draft only: 20 July 2013]

Research question:

Is there a difference between the gross earnings (from paid work) of men and women? What other variables might account for differences in earnings? What other variables might affect earnings regardless of gender? What effect do they have by themselves? What happens to any differences in earnings between men and women when controlling for these other variables?

Exemplar: British Social Attitudes 1989

Previous: 2.3.1.6.1 Specimen answer for conditional frequencies homework [Tasks 1 and 2]
2.3.1.6.2 Specimen answer for conditional frequencies homework [Tasks 3 and 4]

File: 2.3.1.6.1.sav

In the previous exercises we produced a contingency table of **sex** by **v1727** in which epsilons (percentage point differences) were calculated between men and women for each earnings group.

sex Q901a Sex of respondent * v1727 Q.918b Income group of respondent (if working) Crosstabulation

% within sex Q901a Sex of respondent

	Q.918b Income group of respondent (if working)														Total
	Under £2000	£2000 < £2999	£3000 < £3999	£4000 < £4999	£5000 < £5999	£6000 < £6999	£7000 < £7999	£8000 < £9999	£10000 < £11999	£12000 < £1999	£15000 < £17999	£18000 < £19999	£20000 < £24000	£24000 or more	
	%	%	%	%	%	%	%	%	%	%	%	%	%	%	n = 100%
Total	**5.2**	**5.7**	**5.8**	**6.0**	**7.4**	**7.2**	**8.1**	**11.6**	**11.2**	**12.2**	**7.1**	**3.7**	**1.9**	**7.0**	**1560**
Men	0.3	0.8	0.9	2.4	5.4	5.3	8.7	13.4	14.1	17.4	10.9	5.7	3.2	11.6	**874**
Women	11.4	12.0	12.1	10.5	9.9	9.6	7.3	9.3	7.4	5.7	2.3	1.2	0.1	1.2	**686**
Epsilon	**-11.1**	**-11.2**	**-11.2**	**-8.1**	**-4.5**	**-4.3**	**+1.4**	**+4.1**	**+6.7**	**+11.7**	**+8.6**	**+4.5**	**+3.1**	**+10.4**	

Step 1: Choosing cutting points for earnings

For demonstration purposes this is a rather large table to work with so we need to reduce the earnings from fourteen to fewer categories to make the table easier to read, not just when we produce two-way tables for test variables, but especially when producing three-way tables of sex by earnings controlling for test variables. In three- or four-way tables the number of table cells can rapidly get very large, and the cell counts consequently very small. We shall also need to reduce the number of categories in some of the test variables to keep the cell counts large enough to serve as a reasonable base for percentages. A general rule of thumb is that cell counts should be at least 40, at which level moving a single case from one category to another makes a net difference of 5 percentage points (it takes 2.5% from the source category adds 2.5% to the target category).

Before we continue, you should think about where the cutting points should be to reduce earnings groups from fourteen categories to four or even three. Your criteria should make both statistical and sociological sense. Statistically it's better to have categories of approximately equal size, especially if we're going to introduce a third variable: sociologically, it's sometimes better to define very high and very low cut-off points such as top 10% and bottom 10%, but this may mean very small groups at the extremes. Sometimes there will be standard groupings from other sources (age group, social grade, terminal education age) with which you may want to make comparisons. Usually we compromise to find an empirical solution that makes sense, but we'll always be constrained by the size of the sample. In this case we have unequal distributions for men and women, so cutting points also need to take this into account.

This is where cumulative frequencies come in handy. In tutorial 2.3.1.6.2 Specimen answer for conditional frequencies homework [Tasks 3 and 4] we started with a simple frequency count for the whole sample which we then partitioned into separate frequency counts for men and women.

All

v1727 Q.918b Income group of respondent (if working)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Under £2000	81	2.7	5.2	**5.2**
	2 £2000 < £2999	89	2.9	5.7	**10.9**
	3 £3000 < £3999	91	3.0	5.8	**16.7**
	4 £4000 < £4999	93	3.1	6.0	**22.7**
	5 £5000 < £5999	115	3.8	7.4	**30.1**
	6 £6000 < £6999	112	3.7	7.2	**37.2**
	7 £7000 < £7999	126	4.2	8.1	**45.3**
	8 £8000 < £9999	181	6.0	11.6	**56.9**
	9 £10000 < £11999	174	5.8	11.2	**68.1**
	10 £12000 < £1999	191	6.3	12.2	**80.3**
	11 £15000 < £17999	111	3.7	7.1	**87.4**
	12 £18000 < £19999	58	1.9	3.7	**91.2**
	13 £20000 < £24000	29	1.0	1.9	**93.0**
	14 £24000 or more	109	3.6	7.0	**100.0**
	Total	1560	51.6	100.0	
Missing	98 Don't know	17	.6		
	99 Not answered	108	3.6		
	System	1340	44.3		
	Total	1465	48.4		
Total		3025	100.0		

Men only

v1727 Q.918b Income group of respondent (if working)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Under £2000	3	.2	.3	**.3**
	2 £2000 < £2999	7	.5	.8	**1.1**
	3 £3000 < £3999	8	.6	.9	**2.1**
	4 £4000 < £4999	21	1.5	2.4	**4.5**
	5 £5000 < £5999	47	3.4	5.4	**9.8**
	6 £6000 < £6999	46	3.3	5.3	**15.1**
	7 £7000 < £7999	76	5.5	8.7	**23.8**
	8 £8000 < £9999	117	8.4	13.4	**37.2**
	9 £10000 < £11999	123	8.8	14.1	**51.3**
	10 £12000 < £1999	152	10.9	17.4	**68.6**
	11 £15000 < £17999	95	6.8	10.9	**79.5**
	12 £18000 < £19999	50	3.6	5.7	**85.2**
	13 £20000 < £24000	28	2.0	3.2	**88.4**
	14 £24000 or more	101	7.3	11.6	**100.0**
	Total	874	62.7	100.0	
Missing	98 Don't know	7	.5		
	99 Not answered	62	4.5		
	System	450	32.3		
	Total	519	37.3		
Total		1393	100.0		

Women only

v1727 Q.918b Income group of respondent (if working)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Under £2000	78	4.8	11.4	**11.4**
	2 £2000 < £2999	82	5.0	12.0	**23.3**
	3 £3000 < £3999	83	5.1	12.1	**35.4**
	4 £4000 < £4999	72	4.4	10.5	**45.9**
	5 £5000 < £5999	68	4.2	9.9	**55.8**
	6 £6000 < £6999	66	4.0	9.6	**65.5**
	7 £7000 < £7999	50	3.1	7.3	**72.7**
	8 £8000 < £9999	64	3.9	9.3	**82.1**
	9 £10000 < £11999	51	3.1	7.4	**89.5**
	10 £12000 < £1999	39	2.4	5.7	**95.2**
	11 £15000 < £17999	16	1.0	2.3	**97.5**
	12 £18000 < £19999	8	.5	1.2	**98.7**
	13 £20000 < £24000	1	.1	.1	**98.8**
	14 £24000 or more	8	.5	1.2	**100.0**
	Total	686	42.0	100.0	
Missing	98 Don't know	10	.6		
	99 Not answered	46	2.8		
	System	890	54.5		
	Total	946	58.0		
Total		1632	100.0		

For approximately equal-sized groups, percentiles can be useful for deciding on cutting points. For **four** groups, you can use the quartiles and the median:

```
freq v1727 /for not /per 25 50 75.
```

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	1560
	Missing	1465
Percentiles	25	**5.00**
	50	**8.00**
	75	**10.00**

All

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	874
	Missing	519
Percentiles	25	**8.00**
	50	**9.00**
	75	**11.00**

Men only

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	686
	Missing	946
Percentiles	25	**3.00**
	50	**5.00**
	75	**8.00**

Women only

The median for the whole sample is **8** but for men it is **9** and for women **5**. The lower quartile point for the whole sample is **5** but for men it is **8** and for women **3**: the upper quartile point for the whole sample is **8** but for men it is **11** and for women **8**.

If we create **four** new groups:

```
recode v1727 (1 2 3 4 = 1) (5 6 7 = 2) (8 9 10 = 3)
        (11 thru 14 = 4)(else = sysmis) into incr4.
crosstabs sex by incr4.
```

sex Q901a: Sex of respondent * incr4 Q918b Gross income of R (if working) [4 groups] Crosstabulation

			incr4 Q918b Gross income of R (if working) [4 groups]				Total
			1 <£5000	2 <£8000	3 <£15000	4 £15000+	
sex Q901a: Sex of respondent	1 Men	Count	**39**	169	392	274	874
		% within sex Q901a: Sex of respondent	4.5%	19.3%	44.9%	31.4%	100.0%
	2 Women	Count	315	184	154	**33**	686
		% within sex Q901a: Sex of respondent	45.9%	26.8%	22.4%	4.8%	100.0%
Total		Count	354	353	546	307	1560
		% within sex Q901a: Sex of respondent	22.7%	22.6%	35.0%	19.7%	100.0%

. . the cutting points show only 39 men earning under £5000 and only 33 women earning £15,000 or more, and the group totals are uneven.

For **three** groups, you can use the 33rd and 67th percentiles:

```
freq v1727 /for not /per 33 67.
```

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	1560
	Missing	1465
Percentiles	33	**6.00**
	67	**9.00**

All

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	874
	Missing	519
Percentiles	33	**8.00**
	67	**10.00**

Men only

Statistics
v1727 Q.918b Income group of respondent (if working)

N	Valid	686
	Missing	946
Percentiles	33	**3.00**
	67	**7.00**

Women only

The 33rd percentile point for the whole sample is **6** but for men it is **8** and for women only **3**: the 67th percentile point for the whole sample is **9** but for men it is **10** and for women **7**.

If we create **three** new groups:

```
recode   v1727
        (1 2 3 4 5 = 1) ( 6 7 8 9 = 2) ( 10 thru 14 = 3)
        (else = sysmis) into incr3.
crosstabs sex by incr3 /cel cou row.
```

Q901a Sex of respondent * Q918b Gross income of R (if working) [3 groups] Crosstabulation

			Q918b Gross income of R (if working) [3 groups]			Total
			<£6000	<£12000	£12000+	
Q901a Sex of respondent	Men	Count	86	362	426	874
		% within Q901a Sex of respondent	9.8%	41.4%	48.7%	100.0%
	Women	Count	383	231	72	686
		% within Q901a Sex of respondent	55.8%	33.7%	10.5%	100.0%
Total		Count	469	593	498	1560
		% within Q901a Sex of respondent	30.1%	38.0%	31.9%	100.0%

. . we get 86 men earning under £6,000 and 72 women earning £12,000 or more. It's not much of an improvement in numbers, but it's easier to work with three categories rather than four, so we'll stick with three for the following exercises. These numbers are still quite small, but with these cutting points the group totals are more even-sized, Low (469, 30.1%) Medium (593, 38.0%) and High (498, 31.9%). If we edit the table into a more easily interpretable format, and calculate the epsilons (percentage point differences) between men and women . .

Q901a Sex of respondent * Q918b Gross income of R (if working) [3 groups] Crosstabulation

			Q918b Gross income of R (if working) [3 groups]			Total
			<£6000 %	<£12000 %	£12000+ %	(n = 100%)
Q901a Sex of respondent	Total		**30.1**	**38.0**	**31.9**	1560
	Men		9.8	41.4	48.7	874
	Women		55.8	33.7	10.5	686
		Epsilon	**-46.0**	**+7.7**	**+38.3**	

. . the epsilons have a marked and very large shift from -46.0 through +7.7 to +38.3

Step 2: Choose test variables

Research question:

What other variables might affect earnings regardless of gender? What effect do they have by themselves?

Variables which could affect earnings are the number of hours worked, being self-employed or an employee, skill level required, type of work, educational qualifications held, and whether working in the private or the public sector.

		Question	record	column(s)	Name
Dependent variable:	Personal gross earnings	Q.918b	17	27	**v1727**
Independent variable:	Sex	Q.901a	14	11	**v1411**

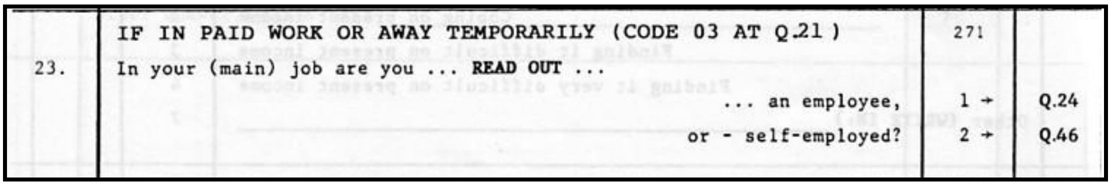
Possible test variables for which data are available in this survey include:

	Work				
Test variables:	Employee or self-employed	Q.23	2	71	**v271**
	Hours worked, employee	Q.24	2	75	**v275**
	Hours worked, self-employed	Q.46a	4	61	**v461**
	Public or private sector	Q.908f	16	17-18	**v1617**
	Level of work	Q.908a	23	61	**v2361**
	Education				
	Terminal Education Age	Q.906a	15	30	**v1530**
	Level of education [derived]	Q.907b	23	74	**v2374**
	Other				
	Age last birthday	Q.901b	14	12-13	**v1412**

Let's have a look at some of these variables:

Work

Employment status (Q.23) is coded direct on record 2 column 71: [**v271**]



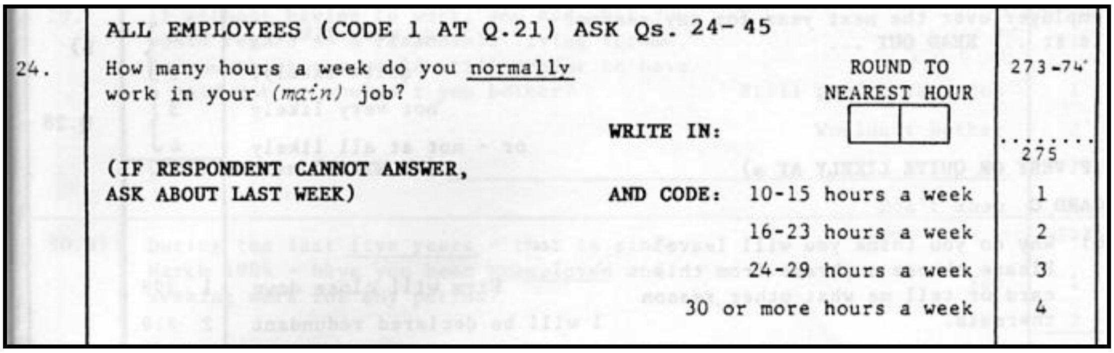
23.	**IF IN PAID WORK OR AWAY TEMPORARILY (CODE 03 AT Q.21)** In your (main) job are you ... **READ OUT** ...	271	
	... an employee,	1 →	Q.24
	or - self-employed?	2 →	Q.46

v271 In your main job are you an employee?Q23NI21

		Frequency
Valid	1 Employee	1458
	2 Self- employed	227
	Total	1685
Missing	100	1339
	System	1
	Total	1340
Total		3025

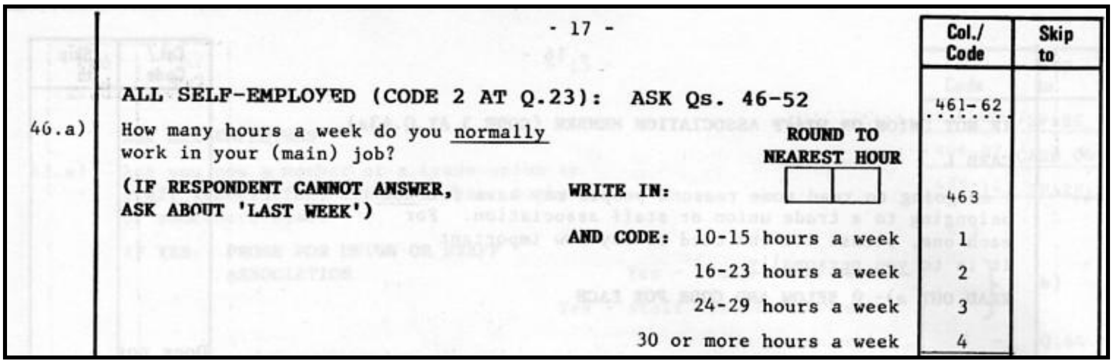
Weekly hours worked are recorded separately for employees and self-employed.

Weekly hours worked by **employees** (Q.24) are grouped and coded direct on record 2 column 75 [**v275**]



	ALL EMPLOYEES (CODE 1 AT Q.21) ASK Qs. 24-45		
24.	How many hours a week do you normally work in your (main) job?	ROUND TO NEAREST HOUR	273-74
		WRITE IN: □□	....
	(IF RESPONDENT CANNOT ANSWER, ASK ABOUT LAST WEEK)		275
		AND CODE: 10-15 hours a week	1
		16-23 hours a week	2
		24-29 hours a week	3
		30 or more hours a week	4

Weekly hours worked by **self-employed** (Q.46) are grouped and coded direct on record 4 column 63 [**v463**]



- 17 -

			Col./ Code	Skip to
	ALL SELF-EMPLOYED (CODE 2 AT Q.23): ASK Qs. 46-52		461-62	
46.a)	How many hours a week do you normally work in your (main) job?	ROUND TO NEAREST HOUR		
	(IF RESPONDENT CANNOT ANSWER, ASK ABOUT 'LAST WEEK')	WRITE IN: □□	463	
		AND CODE: 10-15 hours a week	1	
		16-23 hours a week	2	
		24-29 hours a week	3	
		30 or more hours a week	4	

To tabulate them in the same table you need to use the MULT RESPONSE command:

mult resp groups =

```
ftpt 'Weekly hours worked'
(v275 v463 (1,4))
 /freq ftpt.
```

ftpt Frequencies

		Responses		Percent of Cases
		N	Percent	
ftpt[a]	1 10-15 hours	100	5.9%	5.9%
	2 16-23 hours	138	8.2%	8.2%
	3 24-29 hours	79	4.7%	4.7%
	4 30+ hours	1365	81.2%	81.2%
Total		1682	100.0%	100.0%

a. Group

Creating a single variable for weekly hours worked [**workhours**] is more complex as it requires a conditional transformation to combine information from both **v275** and **v463**.

.

```
do if   (v271 = 1).
compute workhours =  v275 .
else if   (7271 = 2).
compute workhours =  v463 .
end if.

var lab workhours 'Weekly hours worked' .
val lab workhours
      1 '10-15'
      2'16-23'
      3 '24-29'
      4 '30 or more' .
freq workhours.
```

workhours Weekly hours worked

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 10-15	100	3.3	5.9	5.9
	2 16-23	138	4.6	8.2	14.1
	3 24-29	79	2.6	4.7	18.8
	4 30 or more	1365	45.1	81.2	100.0
	Total	1682	55.6	100.0	
Missing	System	1343	44.4		
Total		3025	100.0		

Likewise the question on **self-employment** appears twice, once at Q.23 [**v271**] and as a check at Q.908e [**v1616**] but if you cross-tabulate the two, they don't actually tally: three cases are not coded the same in each variable.

```
crosstabs      v271 by v1616.
```

v271 In your main job are you an employee?Q23NI21 * v1616 Respondent:employee or selfemployedQ908eNI907e Crosstabulation

Count

		v1616 Respondent:employee or selfemployedQ908eNI907e		Total
		1 Employee	2 Self- employed	
v271 In your main job are you an employee?Q23NI21	1 Employee	1456	1	1457
	2 Self- employed	2	225	227
Total		1458	226	1684

This should have been picked up by the research team at the data entry and data cleaning stages: we can identify the three cases, but we can't do the check because we haven't got the original questionnaires.

Step 3: Check test variables

We now need to run a few checks on the test variables to get a feel for the data and see how the cell counts work out. First let's look at weekly hours worked by full- or part-time work:

crosstabs   workhours by v271 v1616.

workhours Weekly hours worked * v271 Q23: Employee or self-employed Crosstabulation

Count

		v271 Q23: Employee or self-employed		Total
		1 Employee	2 Self-employed	
workhours Weekly hours worked	1 10-15	88	12	100
	2 16-23	119	19	138
	3 24-29	74	5	79
	4 30 or more	1176	189	1365
Total		1457	225	1682

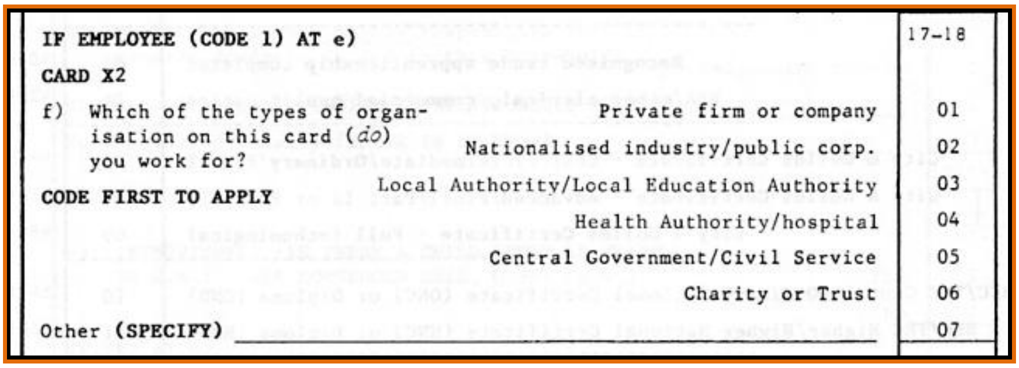
workhours Weekly hours worked * v1616 Q908e: Employee or self-employed Crosstabulation

Count

		v1616 Q908e: Employee or self-employed		Total
		1 Employee	2 Self-employed	
workhours Weekly hours worked	1 10-15	88	12	100
	2 16-23	120	18	138
	3 24-29	74	5	79
	4 30 or more	1175	189	1364
Total		1457	224	1681

It doesn't matter which one you use as they are both the same, so let's stick with **v271**. Regardless of cell counts, the obvious split will be part-time (values 1, 2 and 3) and full-time (4)

Employment sector is asked at Q.908f and is coded directly on record 16 columns 17-18 [**v1617**]



IF EMPLOYEE (CODE 1) AT e)		17-18
CARD X2		
f) Which of the types of organisation on this card (do) you work for?	Private firm or company	01
	Nationalised industry/public corp.	02
CODE FIRST TO APPLY	Local Authority/Local Education Authority	03
	Health Authority/hospital	04
	Central Government/Civil Service	05
	Charity or Trust	06
Other (SPECIFY) ________________		07

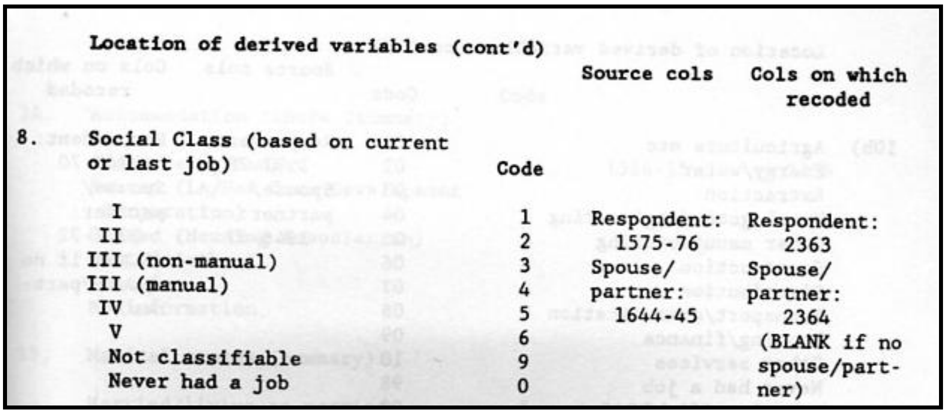
[NB: The following tables have been edited to keep only the frequency counts and cumulative percentages.**]**

v1617 Q908f: Private or public sector

		Frequency	Cumulative Percent
Valid	1 Private	1633	63.7
	2 Nationalised	179	70.7
	3 Local Government	340	84.0
	4 Health Authority	161	90.3
	5 Civil Service	146	96.0
	6 Charity or Trust	33	97.3
	7 Other	70	100.0
	Total	2562	

Sector could be dichotomised in to Private (value 1) and Other (2 thru 7)

Classification of work into (Registrar General's) **Social Class** is based on post-coding of the detailed questions at Q.908: the data are on record 23 column 63 [v2363]



Location of derived variables (cont'd)

	Code	Source cols	Cols on which recoded
8. Social Class (based on current or last job)			
I	1	Respondent: 1575-76 Spouse/ partner: 1644-45	Respondent: 2363 Spouse/ partner: 2364 (BLANK if no spouse/partner)
II	2		
III (non-manual)	3		
III (manual)	4		
IV	5		
V	6		
Not classifiable	9		
Never had a job	0		

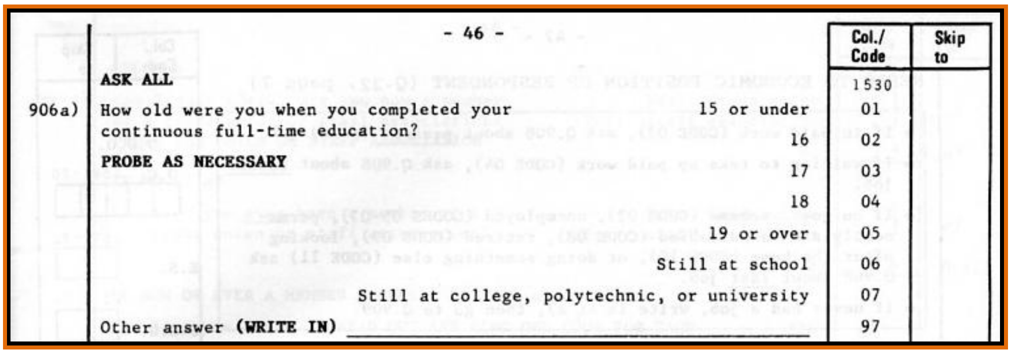
v2363 Social Class of work

		Frequency	Cumulative Percent
Valid	1 I	121	4.3
	2 II	642	26.8
	3 III non-manual	724	52.2
	4 III manual	658	75.4
	5 IV	545	94.5
	6 V	156	100.0
	Total	2846	

Social class can be dichotomised into Non-manual (White collar, values 1, 2 and 3) and Manual (Blue-collar, 4, 5 and 6)

Education

Terminal education age (TEA) is a standard question in many social surveys. This is asked at Q.906a and the data are directly entered on record 15 column 30 [v1530]. It could be used as well as, or instead of, highest educational qualification.



- 46 -

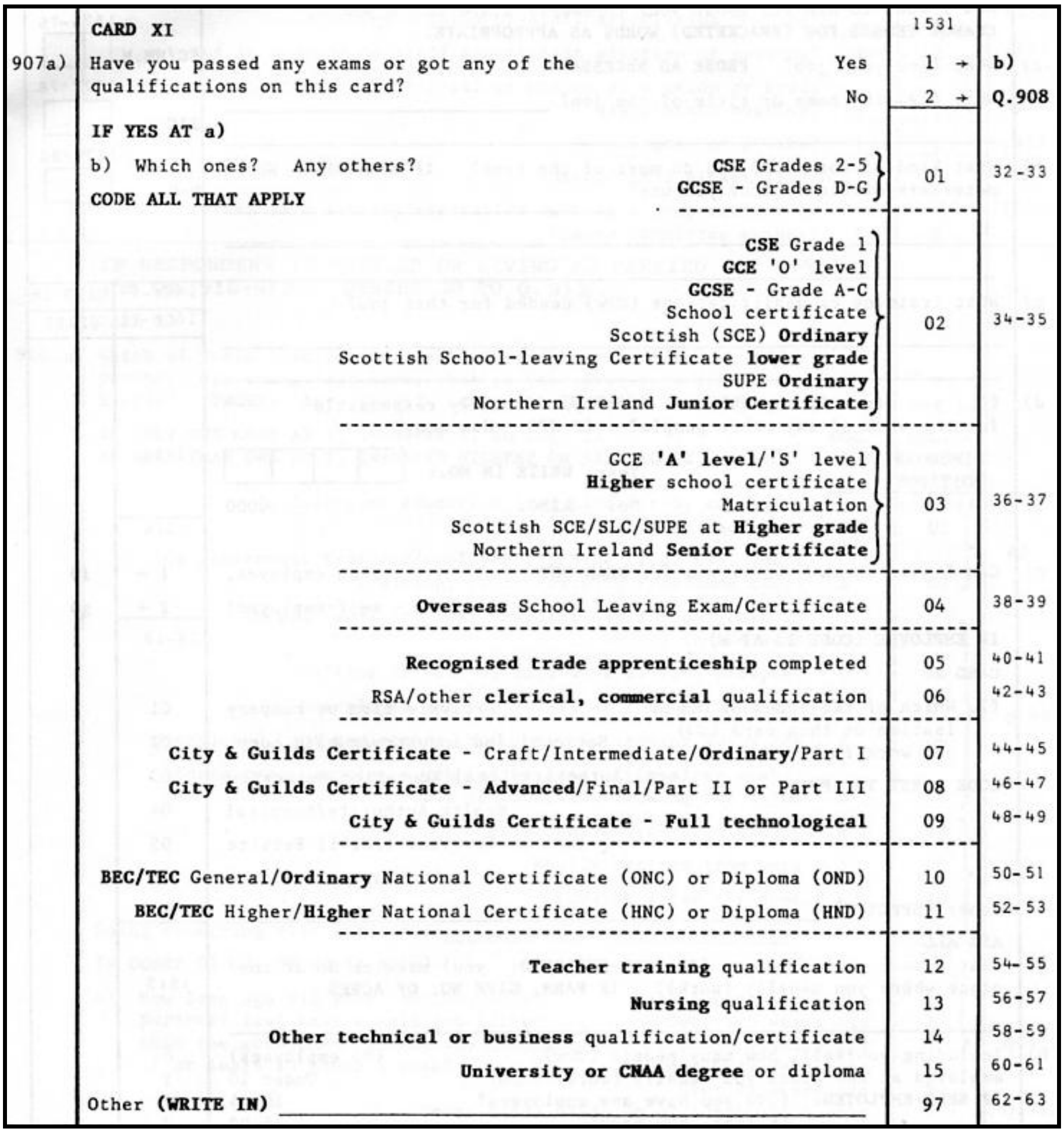
			Col./ Code	Skip to
	ASK ALL		1530	
906a)	How old were you when you completed your continuous full-time education? PROBE AS NECESSARY	15 or under	01	
		16	02	
		17	03	
		18	04	
		19 or over	05	
		Still at school	06	
		Still at college, polytechnic, or university	07	
	Other answer (WRITE IN) ______		97	

v1530 Age completed full time education

		Frequency	Cumulative Percent
Valid	1 15 or under	1421	47.0
	2 16	753	71.9
	3 17	219	79.1
	4 18	198	85.7
	5 19 or 0ver	370	97.9
	6 Still at school	7	98.1
	7 Still at college	55	99.9
	9	2	100.0
	Total	3025	

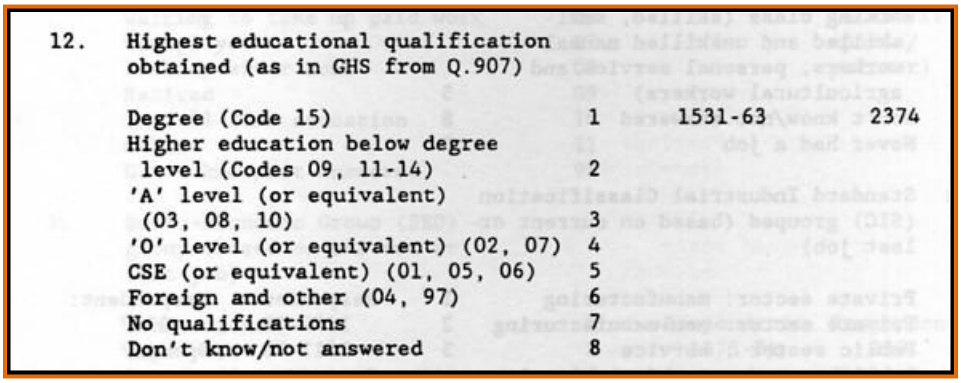
TEA could be dichotomised into 15 or under and 16 or over, but this would lose the distinction for later leaving ages, Perhaps a better choice would be three groups, 15 or under (value 1) 16 and 17 (2 and 3) and 18 or over (4 thru 7). Note that value 9 has not been declared as missing.

Qualifications from education or training are recorded at Q.907a: they are coded direct in two-column fields on record 15 columns 32 – 63 and are effectively multiple response items [**v1532, v1534 v1536,** ~ ~ ~ **v1562**]



	CARD XI	1531	
907a)	Have you passed any exams or got any of the qualifications on this card? Yes	1 →	b)
	No	2 →	Q.908
	IF YES AT a) b) Which ones? Any others? **CODE ALL THAT APPLY**		
	CSE Grades 2-5 **GCSE** - Grades D-G	01	32-33
	CSE Grade 1 **GCE** 'O' level **GCSE** - Grade A-C School certificate Scottish (SCE) **Ordinary** Scottish School-leaving Certificate **lower grade** SUPE **Ordinary** Northern Ireland **Junior Certificate**	02	34-35
	GCE **'A'** level/'S' level **Higher** school certificate Matriculation Scottish SCE/SLC/SUPE at **Higher grade** Northern Ireland **Senior Certificate**	03	36-37
	Overseas School Leaving Exam/Certificate	04	38-39
	Recognised trade apprenticeship completed	05	40-41
	RSA/other **clerical, commercial** qualification	06	42-43
	City & Guilds Certificate - Craft/Intermediate/**Ordinary**/Part I	07	44-45
	City & Guilds Certificate - **Advanced/Final**/Part II or Part III	08	46-47
	City & Guilds Certificate - Full technological	09	48-49
	BEC/TEC General/**Ordinary** National Certificate (ONC) or Diploma (OND)	10	50-51
	BEC/TEC Higher/**Higher** National Certificate (HNC) or Diploma (HND)	11	52-53
	Teacher training qualification	12	54-55
	Nursing qualification	13	56-57
	Other technical or business qualification/certificate	14	58-59
	University or CNAA degree or diploma	15	60-61
	Other **(WRITE IN)** ________	97	62-63

The data for the **highest qualification** have already been recoded by the research team into a derived variable on record 23 column 74 [**v2374**]



12. Highest educational qualification obtained (as in GHS from Q.907)

Degree (Code 15)	1	1531-63	2374
Higher education below degree level (Codes 09, 11-14)	2		
'A' level (or equivalent) (03, 08, 10)	3		
'O' level (or equivalent) (02, 07)	4		
CSE (or equivalent) (01, 05, 06)	5		
Foreign and other (04, 97)	6		
No qualifications	7		
Don't know/not answered	8		

v2374 Highest educational qualification

		Frequency	Cumulative Percent
Valid	1 Degree	216	7.2
	2 HE below degree	424	21.2
	3 A-level or equiv.	304	31.3
	4 O-level or equiv	536	49.1
	5 CSE or equiv	242	57.1
	6 Foreign and other	11	57.5
	7 None	1283	100.0
	Total	3016	

This is a difficult one: cutting points determined by cell counts don't necessarily yield meaningful educational levels. Let's save the decision on that one until we've seen it tabulated agagainst earnings.

That's probably enough for one session, but you can be thinking of appropriate cutting points to reduce the number of categories in the test variables.

End of session: 3.1.4.1: Income differences work-through

In the following sessions we will go through a set of exercises to answer our original research question:

Is there a difference between the earnings (from paid work) of men and women? What other variables might account for differences in earnings? What other variables might affect earnings regardless of gender? What effect do they have by themselves? What happens to any differences in earnings between men and women when controlling for these other variables?

3.1.4.2: Income differences – Build working file
Builds up a working file by reading in raw data for dependent, independent and test variables, adding dictionary information and checking file contents. The file is then saved.

3.1.4.3: Income differences for test variables
Reduce gross earnings [v1727] to three categories: produce two-way contingency tables to display distributions of earnings within categories of the test variables.

3.1.4.4: Income differences: choose test variables and cutting points
Decide which test variables to use and choose cutting points; recode test variables into derived variables with fewer categories; produce two-way contingency tables to display distributions of earnings within categories of the selected test variables.

3.1.4.5: Income differences - Elaboration
Three-way contingency tables to see what happens to differences in earnings between men and women when controlling for the selected test variables.

Back to: Block 3 Analysing two variables (and sometimes three)
3.1 Two variables (**CROSSTABS**)

Forward to: 3.1.4.2 Income differences - Build a working file