

Block 3: Analysing two variables (and sometimes three)

3.1.4.2 Income differences - Build a working file

[Draft only: 22 July 2013]

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 effect do they have by themselves?

What happens to any differences in earnings between men and women when controlling for these other variables?

Previous session: [3.1.4.1 Income differences work-through](#)

Exemplar: British Social Attitudes 1989

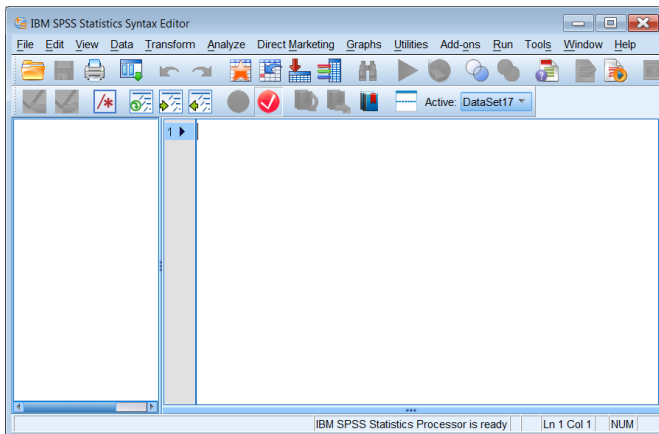
		Question	record	column(s)	Name
Dependent variable:	Personal gross earnings	Q.918b	17	27	v1727
Independent variable:	Sex	Q.901a	14	11	v1411
Test variables:	Work				
	Employee or self-employed	Q.23	2	71	v271
	Hours worked, employee	Q.24	2	75	v275
	Hours worked, self-employed	Q.46a	4	61	v461
	Public or 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

We are now going to build up a working file from scratch with the variables we want.

Raw data file: [bsa89.txt](#) [Download and save to **e:\weebly downloads\bsa89**]

Step 1: Read in raw data

Open SPSS and click on **Cancel**, then **File** > **New** > **Syntax**:



This exercise is **not a typing test**: it will take you through all stages of building up a working file from scratch by defining variables, reading in raw data from a major survey, adding dictionary information, performing data checks and finally saving your *.sav and *.sps files. It will be good practice for you to type out the syntax line-by-line, but if you feel confident working in syntax, just copy/paste the syntax direct into your **Syntax Editor**, then run the commands in stages.

Title '3.1.4.1 Income differences workthrough'.

subtitle 'Get variables from raw data'.

data list file 'e:\weebly downloads\bsa89\bsa89.txt'

records 23

/2 v271 71 v275 75

/4 v463 63

/14 v1411 11 v1412 12-13

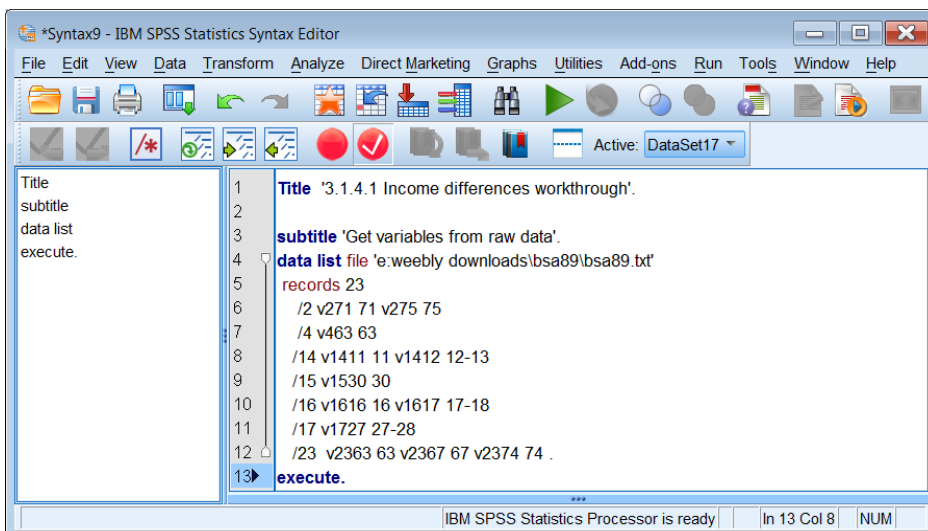
/15 v1530 30

/16 v1616 16 v1617 17-18

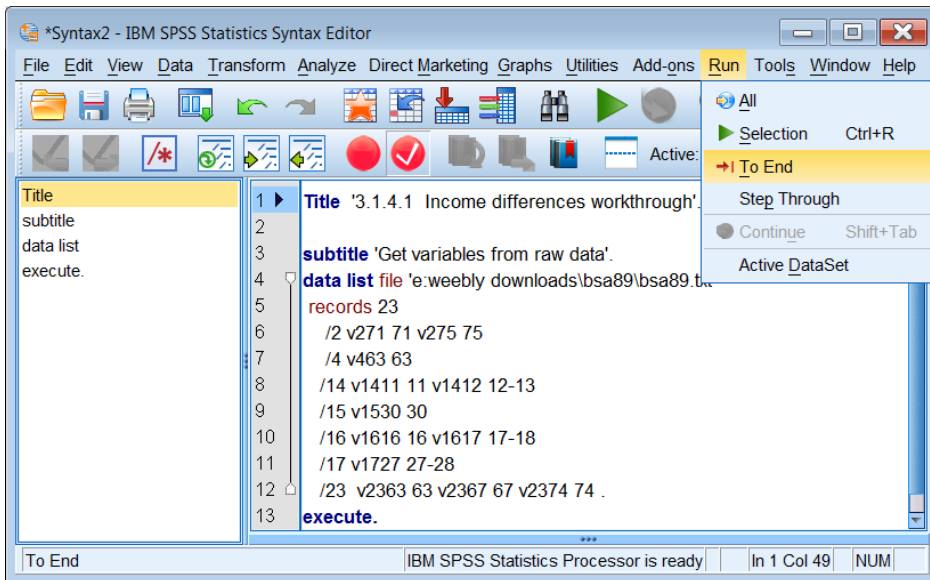
/17 v1727 27-28

/23 v2363 63 v2367 67 v2374 74 .

execute.



Place the cursor in the **title** line, then click on **Run** > **→| To End**



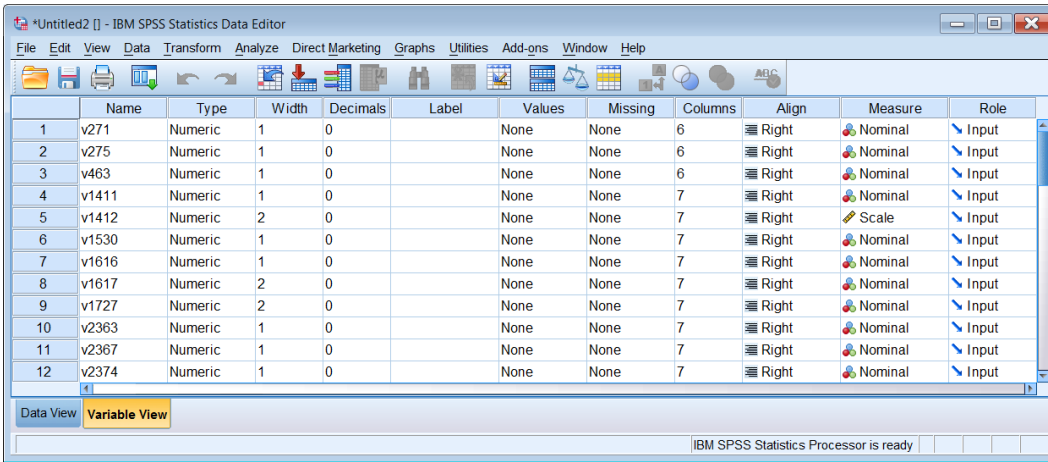
Your syntax is repeated in the viewer:

```
data list file 'e:\weebly downloads\bsa89\bsa89.txt'
  records 23
    /2 v271 71 v275 75
    /4 v463 63
    /14 v1411 11 v1412 12-13
    /16 v1616 16 v1617 17-18
    /17 v1727 27-28
    /23 v2363 63 v2367 67 v2374 74 .

Data List will read 23 records from E:\weebly downloads\bsa89\bsa89.dat

Variable      Rec   Start   End   Format
v271           2     71     71   F1.0
v275           2     75     75   F1.0
v463           4     63     63   F1.0
v1411          14    11     11   F1.0
v1412          14    12     13   F2.0
v1616          16    16     16   F1.0
v1617          16    17     18   F2.0
v1727          17    27     28   F2.0
v2363          23    63     63   F1.0
v2367          23    67     67   F1.0
v2374          23    74     74   F1.0
execute.
```

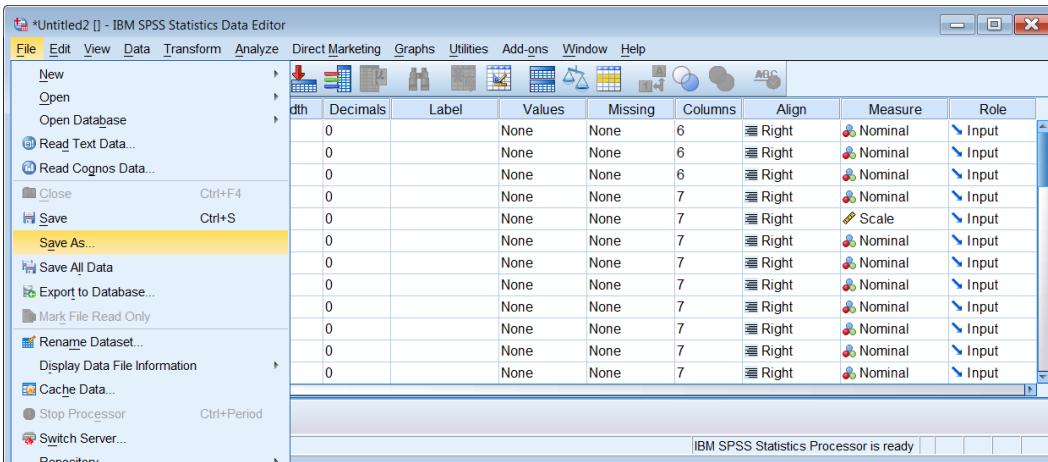
When reading data from 80-column ASCII files, this display is an essential and very useful check to make sure the variables are read in correctly: the **positional** variable names can clearly be seen to tally exactly with the correct records and fields. This run creates a new ***Untitled** data editor with an incremental number, in this case ***Untitled 2**



[!Warning!]

The new working file may disappear if you do anything else in SPSS, so it has to be **saved immediately**. It doesn't display automatically: you have to look for it by hovering over the SPSS icon in the Task Bar. You do **not** need to open a new **Data Editor** before running this job: if you do it will remain empty and there will be a second **Data Editor** containing the data. Even I found this confusing, especially when I was trying to merge data from two SPSS ***.sav** files. However if you work with it from scratch, as now, it seems to be OK, but it's probably safer to save it temporarily, even as **Untitled41!** If you had another ***.sav** file open, make sure you switch to the new ***Untitled** one first.

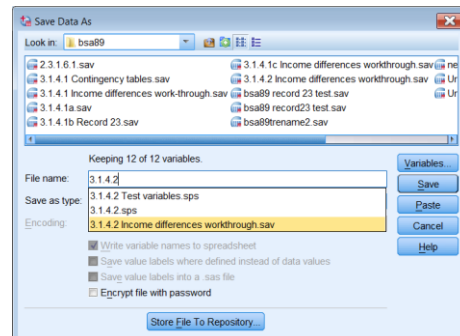
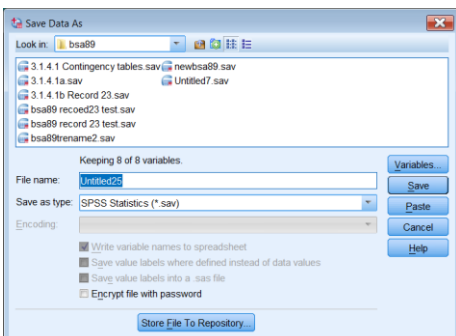
File > Save as



Change ***Untitled** to

3.1.4.2

and click on **Save**



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	v271	Numeric	1	0		None	None	6	Right	Nominal	Input
2	v275	Numeric	1	0		None	None	6	Right	Nominal	Input
3	v463	Numeric	1	0		None	None	6	Right	Nominal	Input
4	v1411	Numeric	1	0		None	None	7	Right	Nominal	Input
5	v1412	Numeric	2	0		None	None	7	Right	Scale	Input
6	v1530	Numeric	1	0		None	None	7	Right	Nominal	Input
7	v1616	Numeric	1	0		None	None	7	Right	Nominal	Input
8	v1617	Numeric	2	0		None	None	7	Right	Nominal	Input
9	v1727	Numeric	2	0		None	None	7	Right	Nominal	Input
10	v2363	Numeric	1	0		None	None	7	Right	Nominal	Input
11	v2367	Numeric	1	0		None	None	7	Right	Nominal	Input
12	v2374	Numeric	1	0		None	None	7	Right	Nominal	Input

Now **save** your syntax file:

```

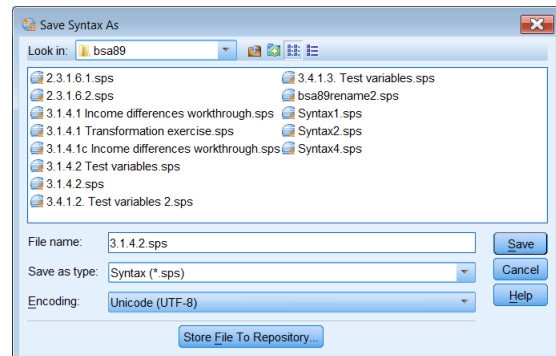
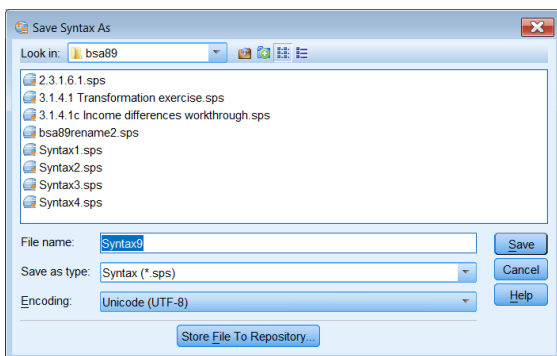
1  Title '3.1.4.1 Income differences workthrough'.
2
3  subtitle 'Get variables from raw data'.
4  data list file 'e:\weebly downloads\bsa89\bsa89.txt'
5  records 23
6  /2 v271 71 v275 75
7  /4 v463 63
8  /14 v1411 11 v1412 12-13
9  /15 v1530 30
10 /16 v1616 16 v1617 17-18
11 /17 v1727 27-28
12 /23 v2363 63 v2367 67 v2374 74 .
13 execute.
  
```

File > Save as

Change ***Syntax** to

3.1.4.2

and click on **Save**



Step 2: Add dictionary information

2.1 Specify variable labels

subtitle 'Add variable labels'.

variable labels

/v271 'Q23: Employee or self-employed'

/v275 'Q24: Hours worked per week [Employee]'

/v463 ' Q.46a Hours worked per week [self-employed]'

/v1411 'Q901a: Sex of respondent'

/v1412 'Q.901b: Age last birthday of respondent'

/v1530 'Age completed full time education'

/v1616 'Q908e: Employee or self-employed'

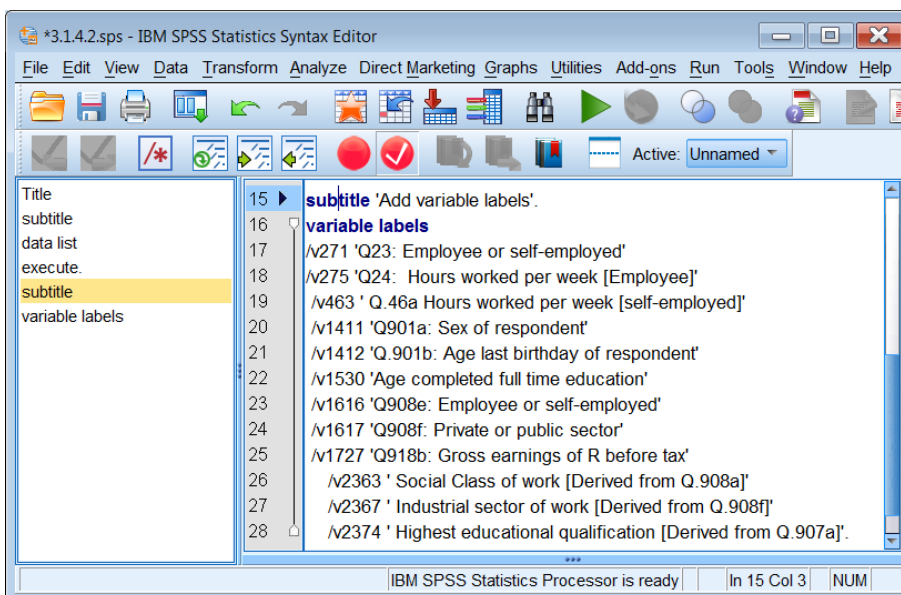
/v1617 'Q908f: Private or public sector'

/v1727 'Q918b: Gross earnings of R before tax'

/v2363 ' Social Class of work [Derived from Q.908a]'

/v2367 ' Industrial sector of work [Derived from Q.908f]'

/v2374 ' Highest educational qualification [Derived from Q.907a]'.



Place the cursor on the next **subtitle** command, then click on **Run** >  **To End**

Variable labels will have been entered in the **Data Editor**

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	v271	Numeric	1	0	Q23: Emplo...	None	None	6	Right	Nominal	Input
2	v275	Numeric	1	0	Q24: Hours w...	None	None	6	Right	Nominal	Input
3	v463	Numeric	1	0	Q.46a Hours ...	None	None	6	Right	Nominal	Input
4	v1411	Numeric	1	0	Q901a: Sex of...	None	None	7	Right	Nominal	Input
5	v1412	Numeric	2	0	Q.901b: Age l...	None	None	7	Right	Scale	Input
6	v1530	Numeric	1	0	Age completed...	None	None	7	Right	Nominal	Input
7	v1616	Numeric	1	0	Q908e: Emplo...	None	None	7	Right	Nominal	Input
8	v1617	Numeric	2	0	Q908f: Private...	None	None	7	Right	Nominal	Input
9	v1727	Numeric	2	0	Q918b: Gross ...	None	None	7	Right	Nominal	Input
10	v2363	Numeric	1	0	Social Class ...	None	None	7	Right	Nominal	Input
11	v2367	Numeric	1	0	Industrial sect...	None	None	7	Right	Nominal	Input
12	v2374	Numeric	1	0	Highest educ...	None	None	7	Right	Nominal	Input

Slide the column separators aside to see the labels more clearly:

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1 v271	Numeric	1	0	Q23: Employee or self-employed	None	None	6	Right	Nominal	Input
2 v275	Numeric	1	0	Q24: Hours worked per week [Employee]	None	None	6	Right	Nominal	Input
3 v463	Numeric	1	0	Q.46a Hours worked per week [self-employed]	None	None	6	Right	Nominal	Input
4 v1411	Numeric	1	0	Q901a: Sex of respondent	None	None	7	Right	Nominal	Input
5 v1412	Numeric	2	0	Q.901b: Age last birthday of respondent	None	None	7	Right	Scale	Input
6 v1530	Numeric	1	0	Age completed full time education	None	None	7	Right	Nominal	Input
7 v1616	Numeric	1	0	Q908e: Employee or self-employed	None	None	7	Right	Nominal	Input
8 v1617	Numeric	2	0	Q908f: Private or public sector	None	None	7	Right	Nominal	Input
9 v1727	Numeric	2	0	Q918b: Gross earnings of R before tax	None	None	7	Right	Nominal	Input
10 v2363	Numeric	1	0	Social Class of work [Derived from Q.908a]	None	None	7	Right	Nominal	Input
11 v2367	Numeric	1	0	Industrial sector of work [Derived from Q.908f]	None	None	7	Right	Nominal	Input
12 v2374	Numeric	1	0	Highest educational qualification [Derived from Q.907a]	None	None	7	Right	Nominal	Input

2.2 Specify value labels

subtitle 'Add value labels'.

value labels

v271 v1616

1 'Employee' 2 'Self-employed'

/ v275 v463

1 '10-15' 2 '16-23' 3 '24-29' 4 '30 or more'

/ v1411

1 'Men' 2 'Women'

/ v1530

1 '15 or under' 2 '16' 3 '17' 4 '18'

5 '19 or over' 6 'Still at school' 7 'Still at college'

/ v1617

1 'Private' 2 'Nationalised' 3 'Local Government' 4 'Health Authority'

5 'Civil Service' 6 'Charity or Trust' 7 'Other'

/ v1727

1 'Under £2000' 2 '£2000 - £2999' 3 '£3000 - £3999' 4 '£4000 - £4999'

5 '£5000 - £5999' 6 '£6000 - £6999' 7 '£7000 - £7999' 8 '£8000 - £9999'

9 '£10000 - £11999' 10 '£12000 - £14999' 11 '£15000 - £17999'

12 '£18000 - £19999' 13 '£20000 - £24000' 14 '£24000 or more'

98 "Don't know" 99 'Not answered'

/ v2363

1 'I' 2 'II' 3 'III non-manual' 4 'III manual' 5 'IV' 6 'V'

/ v2367

1 'Private: manufacture' 2 'Private: non-manufacture'

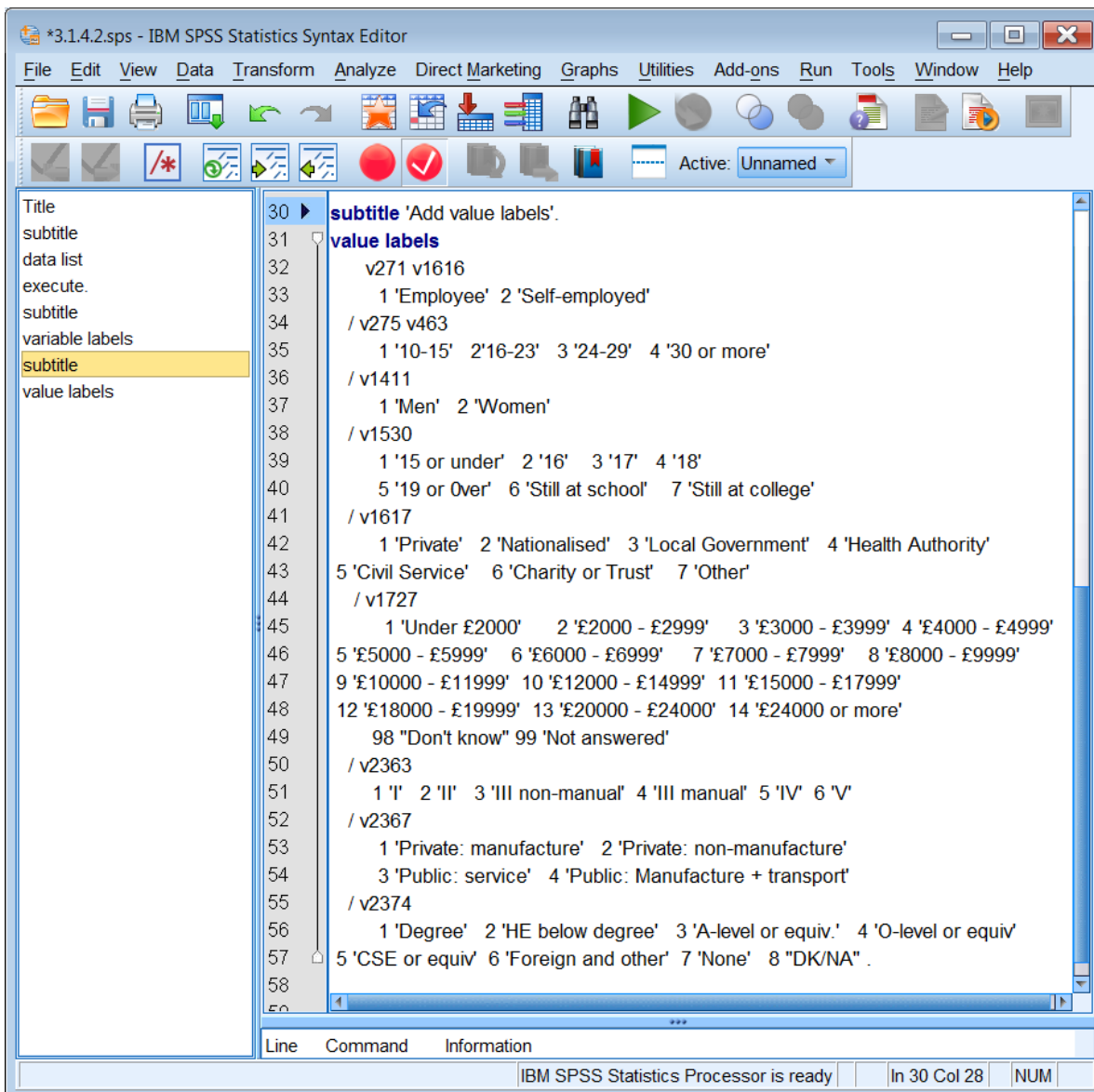
3 'Public: service' 4 'Public: Manufacture + transport'

/ v2374

1 'Degree' 2 'HE below degree' 3 'A-level or equiv.' 4 'O-level or equiv'

5 'CSE or equiv' 6 'Foreign and other' 7 'None' 8 "DK/NA" .

Note the **double primes** for "DK/NA": If you use single primes SPSS will report an error when it finds the forward slash instead of a closing single prime.



Place the cursor on the next **subtitle** command, then click on **Run** > **→ To End**

Value labels will have been entered in the **Data Editor**

Slide the column separators aside again to see the labels more clearly:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	v271	Numeric	1	0	Q23: Employee or self-employed	{1, Employee}...	None	6	Right	Nominal	Input
2	v275	Numeric	1	0	Q24: Hours worked per week [Employee]	{1, 10-15}...	None	6	Right	Nominal	Input
3	v463	Numeric	1	0	Q.46a Hours worked per week [self-employed]	{1, 10-15}...	None	6	Right	Nominal	Input
4	v1411	Numeric	1	0	Q901a: Sex of respondent	{1, Men}...	None	7	Right	Nominal	Input
5	v1412	Numeric	2	0	Q.901b: Age last birthday of respondent	None	None	7	Right	Scale	Input
6	v1530	Numeric	1	0	Age completed full time education	{1, 15 or under}...	None	7	Right	Nominal	Input
7	v1616	Numeric	1	0	Q908e: Employee or self-employed	{1, Employee}...	None	7	Right	Nominal	Input
8	v1617	Numeric	2	0	Q908f: Private or public sector	{1, Private}...	None	7	Right	Nominal	Input
9	v1727	Numeric	2	0	Q918b: Gross earnings of R before tax	{1, Under £2000}...	None	7	Right	Nominal	Input
10	v2363	Numeric	1	0	Social Class of work [Derived from Q.908a]	{1, I}...	None	7	Right	Nominal	Input
11	v2367	Numeric	1	0	Industrial sector of work [Derived from Q.908f]	{1, Private: manufacture}...	None	7	Right	Nominal	Input
12	v2374	Numeric	1	0	Highest educational qualification [Derived from Q.907a]	{1, Degree}...	None	7	Right	Nominal	Input

The status bar at the bottom indicates: IBM SPSS Statistics Processor is ready

2.3 Specify missing values

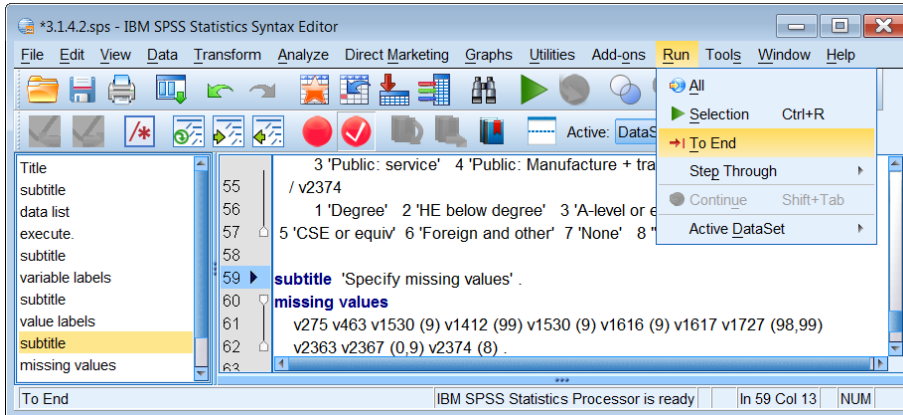
subtitle 'Specify missing values' .

missing values

v275 v463 v1530 (9) v1412 (99) v1530 (9) v1616 (9) v1617 v1727 (98,99)

v2363 v2367 (0,9) v2374 (8) .

Place the cursor on the **subtitle** command and press **Run > →| To End**

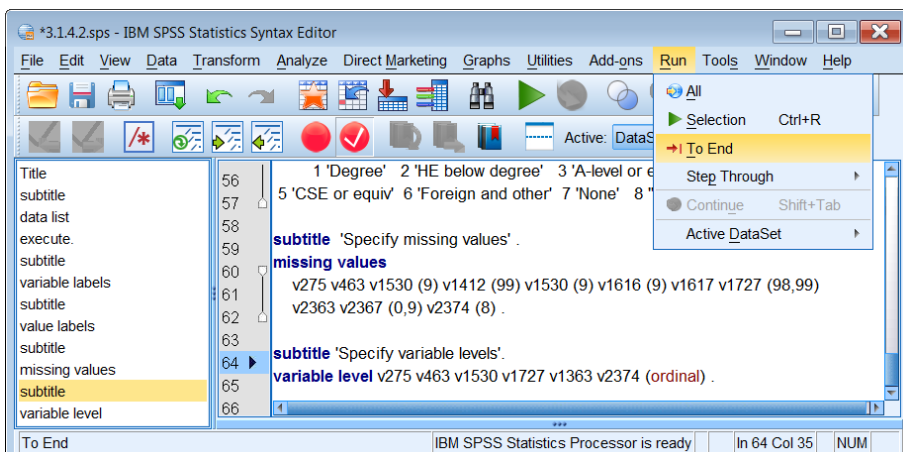


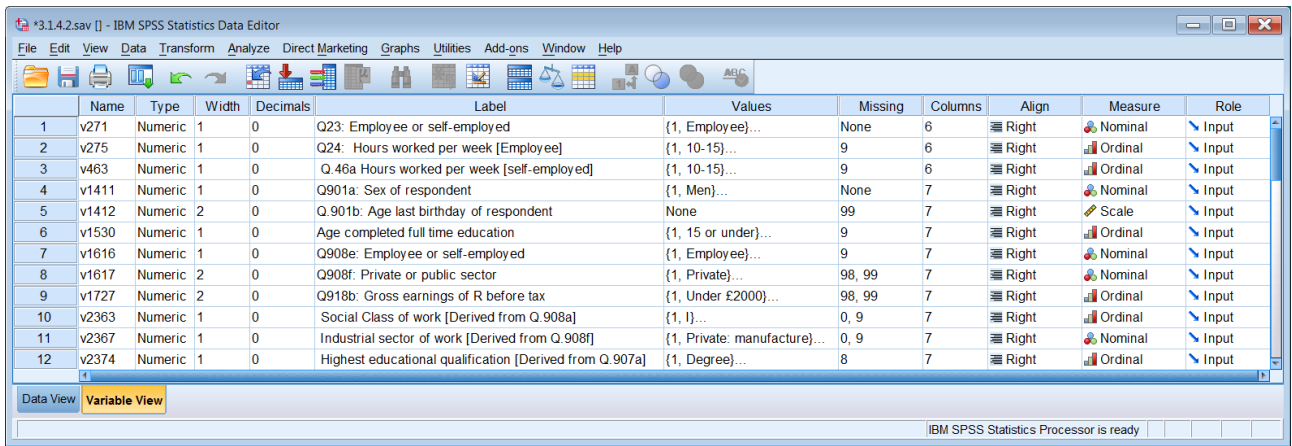
Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1 v271	Numeric	1	0	Q23 Employee or self-employed	{1, Employee}...	None	6	Right	Nominal	Input
2 v275	Numeric	1	0	Q24 Hours worked per week [Employee]	{1, 10-15}...	9	6	Right	Nominal	Input
3 v463	Numeric	1	0	Q 46a Hours worked per week [self-employed]	{1, 10-15}...	9	6	Right	Nominal	Input
4 v1411	Numeric	1	0	Q901a: Sex of respondent	{1, Men}...	None	7	Right	Nominal	Input
5 v1412	Numeric	2	0	Q 901b: Age last birthday of respondent	None	99	7	Right	Scale	Input
6 v1530	Numeric	1	0	Age completed full time education	{1, 15 or under}...	9	7	Right	Nominal	Input
7 v1616	Numeric	1	0	Q908e: Employee or self-employed	{1, Employee}...	9	7	Right	Nominal	Input
8 v1617	Numeric	2	0	Q908f: Private or public sector	{1, Private}...	98, 99	7	Right	Nominal	Input
9 v1727	Numeric	2	0	Q918b: Gross earnings of R before tax	{1, Under £2000}...	98, 99	7	Right	Nominal	Input
10 v2363	Numeric	1	0	Social Class of work [Derived from Q 908a]	{1, I}...	0, 9	7	Right	Nominal	Input
11 v2367	Numeric	1	0	Industrial sector of work [Derived from Q 908f]	{1, Private: manufacture}...	0, 9	7	Right	Nominal	Input
12 v2374	Numeric	1	0	Highest educational qualification [Derived from Q 907a]	{1, Degree}...	8	7	Right	Nominal	Input

Note that, in the **Measure** column, SPSS has set the measurement levels at **Nominal** for all variables except age [v1412] which has been set at **Scale**. This is because the SPSS default is **Nominal** unless the number of values found exceeds a minimum (which can be changed by altering your SPSS settings). For now we can use the **VARIABLE LEVEL** command to change the measurement levels of the **ordinal** variables:

variable level v275 v463 v1530 v1727 v1363 v2374 (ordinal) .

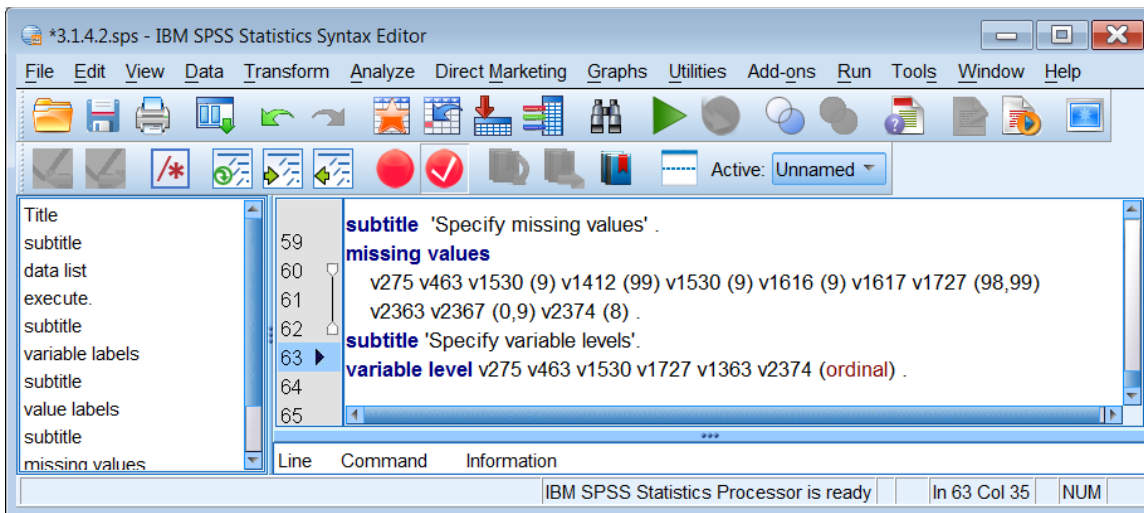
Place the cursor on the **subtitle** command and press **Run > →| To End**





At this point **save the working file** with [CTRL]S

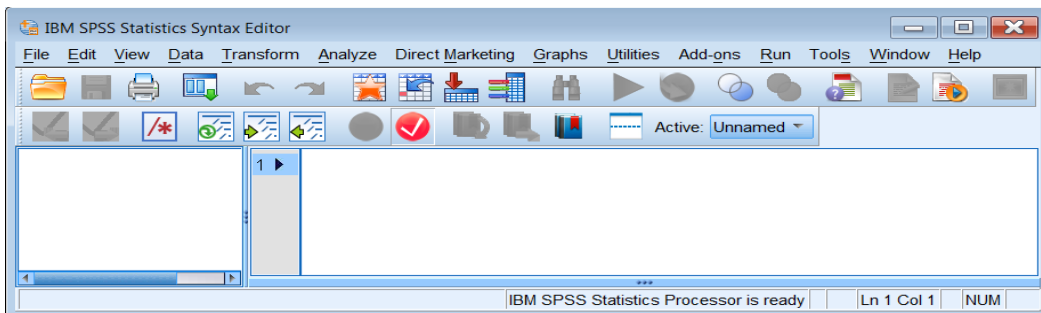
Go back to the **Syntax Editor**



.. and **save the syntax file** with [CTRL]S

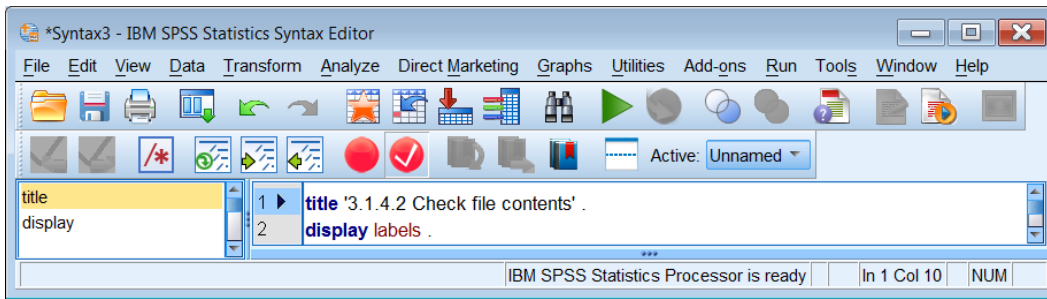
Step 3: Check the file contents

File > New > Syntax:



**title '3.1.4.2 Check file contents' .
display labels .**

[NB: The **DISPLAY** command is **not available** from the drop-down menus]



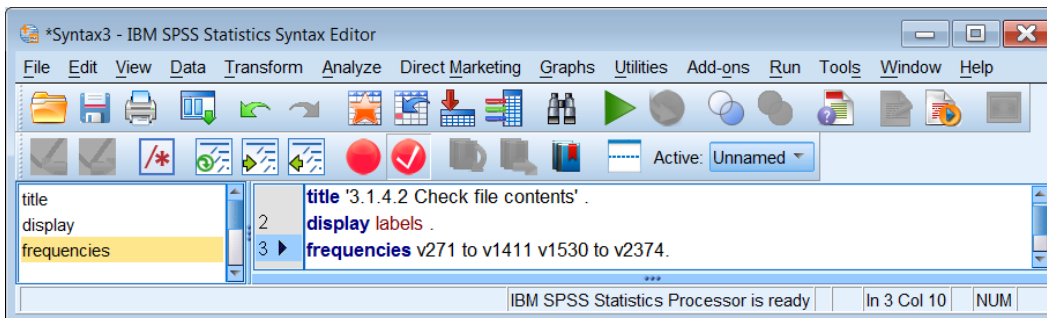
Place the cursor on the **title** command and press **Run** > **→ To end**

Variable	Position	Label
v271	1	Q23: Employee or self-employed
v275	2	Q24: Hours worked per week [Employee]
v463	3	Q.46a Hours worked per week [self-employed]
v1411	4	Q901a: Sex of respondent
v1412	5	Q.901b: Age last birthday of respondent
v1530	6	Age completed full time education
v1616	7	Q908e: Employee or self-employed
v1617	8	Q908f: Private or public sector
v1727	9	Q918b: Gross earnings of R before tax
v2363	10	Social Class of work [Derived from Q.908a]
v2367	11	Industrial sector of work [Derived from Q.908f]
v2374	12	Highest educational qualification [Derived from Q.907a]

Variables in the working file

Run **frequencies** for all variables except age [v1412]

frequencies v271 to v1411 v1530 to v2374.



Place the cursor on the **frequencies** command and press the green triangle:

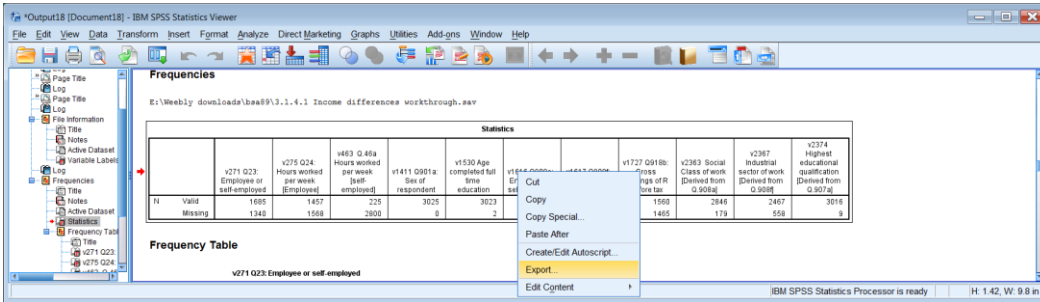
The **Statistics** table in the viewer is far too wide to display when copied direct to a Word page:

The screenshot shows the IBM SPSS Statistics Viewer window displaying the output of the 'frequencies' command. The 'Statistics' table is as follows:

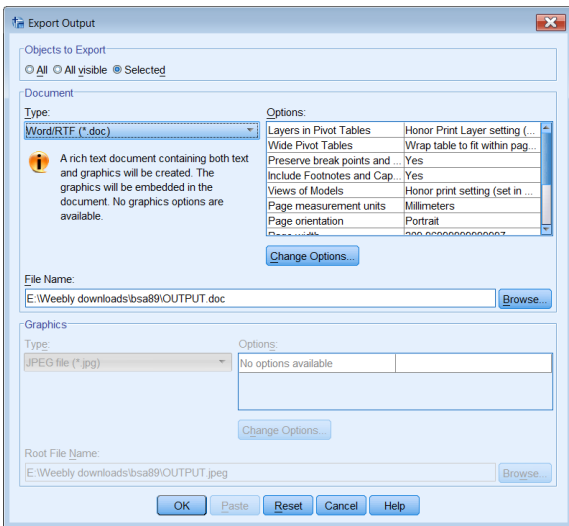
		v271 Q23: Employee or self-employed	v275 Q24: Hours worked per week [Employee]	v463 Q.46a Hours worked per week [self- employed]	v1411 Q901a: Sex of respondent	v1530 Age completed full time education	v1616 Q908e: Employee or self-employed	v1617 Q908f: Private or public sector	v1727 Q918b: Gross earnings of R before tax	v2363 Social Class of work [Derived from Q.908a]	v2367 Industrial sector of week [Derived from Q.908f]	v2374 Highest educational qualification [Derived from Q.907a]
N	Valid	1685	1457	225	3025	3023	2877	2562	1560	2846	2467	3016
	Missing	1340	1588	2800	0	2	148	463	1465	178	558	9

.. but it can be **exported** to a temporary Word RTF *.doc file in the folder from which it was created.

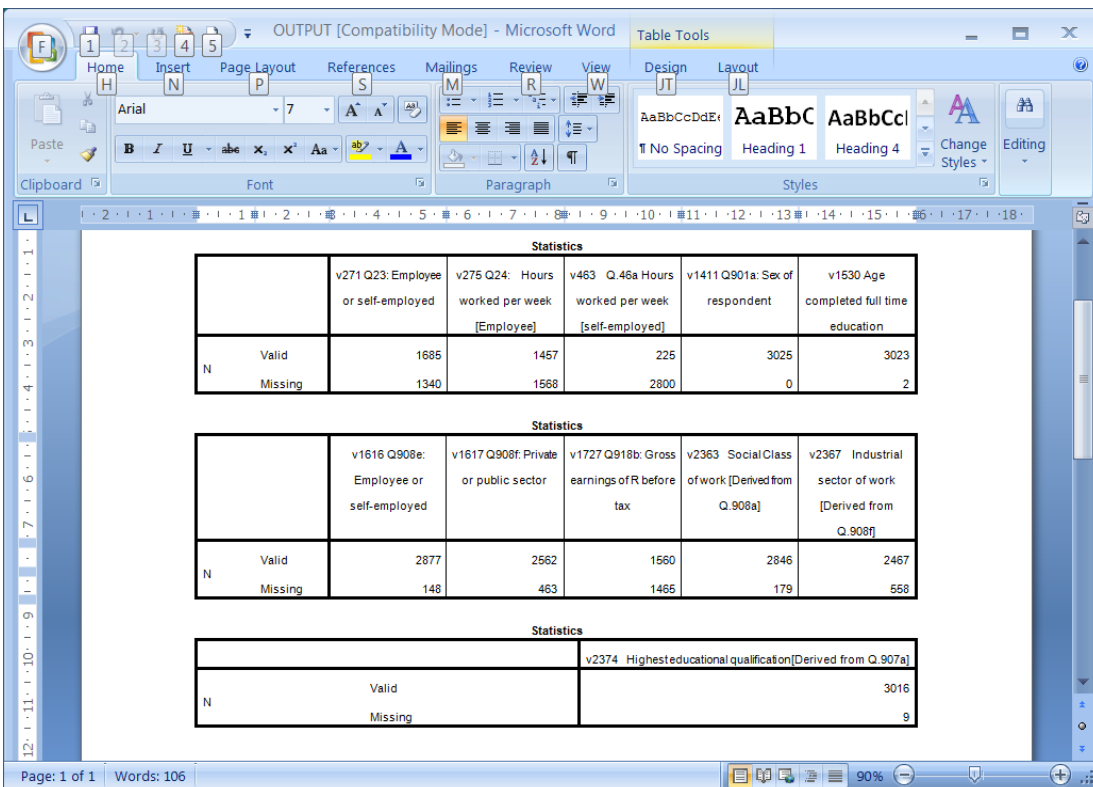
Right click on the table and choose **Export**:



The default target file type is **Word RTF *.doc**:



Click **OK** to save **E:\Weebly downloads\bsa89\OUTPUT.doc** and the table will be split into parts which fit on a Word page:



from which you can copy the tables, as here:

Statistics

		v271 Q23: Employee or self-employed	v275 Q24: Hours worked per week [Employee]	v463 Q.46a Hours worked per week [self-employed]	v1411 Q901a: Sex of respondent	v1530 Age completed full time education
N	Valid	1685	1457	225	3025	3023
	Missing	1340	1568	2800	0	2

Statistics

		v1616 Q908e: Employee or self-employed	v1617 Q908f: Private or public sector	v1727 Q918b: Gross earnings of R before tax	v2363 Social Class of work [Derived from Q.908a]	v2367 Industrial sector of work [Derived from Q.908f]
N	Valid	2877	2562	1560	2846	2467
	Missing	148	463	1465	179	558

Statistics

		v2374 Highest educational qualification [Derived from Q.907a]
N	Valid	3016
	Missing	9

The last one above can be edited on this [Word] page by sliding the column edges, thus:

Statistics

		v2374 Highest educational qualification [Derived from Q.907a]
N	Valid	3016
	Missing	9

I've reduced the following tables to single spacing which makes it look as if the missing values are not properly aligned, but all values seem to be in range. **[NB: Be careful of missing values: they're not always the same in the raw data and the published versions of the main SPSS saved files distributed by the [UK Data Service](#) (UKDS) for [British Social Attitudes 1983 onwards](#).]**

v271 Q23: Employee or self-employed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Employee	1458	48.2	86.5	86.5
	2 Self-employed	227	7.5	13.5	100.0
	Total	1685	55.7	100.0	
Missing	System	1340	44.3		
Total		3025	100.0		

v275 Q24: Hours worked per week [Employee]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 10-15	88	2.9	6.0	6.0
	2 16-23	119	3.9	8.2	14.2
	3 24-29	74	2.4	5.1	19.3
	4 30 or more	1176	38.9	80.7	100.0
	Total	1457	48.2	100.0	
Missing	9	1	.0		
	System	1567	51.8		
	Total	1568	51.8		
Total		3025	100.0		

v463 Q.46a Hours worked per week [self-employed]

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 10-15	12	.4	5.3	5.3
	2 16-23	19	.6	8.4	13.8
	3 24-29	5	.2	2.2	16.0
	4 30 or more	189	6.2	84.0	100.0
	Total	225	7.4	100.0	
Missing	9	2	.1		
	System	2798	92.5		
Total	3025	100.0			

v1411 Q901a: Sex of respondent

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 Men	1393	46.0	46.0	46.0
	2 Women	1632	54.0	54.0	100.0
	Total	3025	100.0	100.0	

v1530 Age completed full time education

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 15 or under	1421	47.0	47.0	47.0
	2 16	753	24.9	24.9	71.9
	3 17	219	7.2	7.2	79.2
	4 18	198	6.5	6.5	85.7
	5 19 or Over	370	12.2	12.2	97.9
	6 Still at school	7	.2	.2	98.2
	7 Still at college	55	1.8	1.8	100.0
Total	3023	99.9	100.0		
Missing	9	2	.1		
Total	3025	100.0			

v1616 Q908e: Employee or self-employed

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 Employee	2567	84.9	89.2	89.2
	2 Self-employed	310	10.2	10.8	100.0
	Total	2877	95.1	100.0	
Missing	9	24	.8		
	System	124	4.1		
	Total	148	4.9		
Total	3025	100.0			

v1617 Q908f: Private or public sector

	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 Private	1633	54.0	63.7	63.7
	2 Nationalised	179	5.9	7.0	70.7
	3 Local Government	340	11.2	13.3	84.0
	4 Health Authority	161	5.3	6.3	90.3
	5 Civil Service	146	4.8	5.7	96.0
	6 Charity or Trust	33	1.1	1.3	97.3
	7 Other	70	2.3	2.7	100.0
Total	2562	84.7	100.0		
Missing	98	1	.0		
	99	4	.1		
	System	458	15.1		
Total	463	15.3			
Total	3025	100.0			

v1727 Q918b: Gross earnings of R before tax

	Frequency	Percent	Valid Percent	Cumulative Percent
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
Valid 8 £8000 - £9999	181	6.0	11.6	56.9
9 £10000 - £11999	174	5.8	11.2	68.1
10 £12000 - £14999	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	
98 Don't know	17	.6		
Missing 99 Not answered	108	3.6		
System	1340	44.3		
Total	1465	48.4		
Total	3025	100.0		

v2363 Social Class of work [Derived from Q.908a]

	Frequency	Percent	Valid Percent	Cumulative Percent
1 I	121	4.0	4.3	4.3
2 II	642	21.2	22.6	26.8
3 III non-manual	724	23.9	25.4	52.2
Valid 4 III manual	658	21.8	23.1	75.4
5 IV	545	18.0	19.1	94.5
6 V	156	5.2	5.5	100.0
Total	2846	94.1	100.0	
0	124	4.1		
Missing 9	55	1.8		
Total	179	5.9		
Total	3025	100.0		

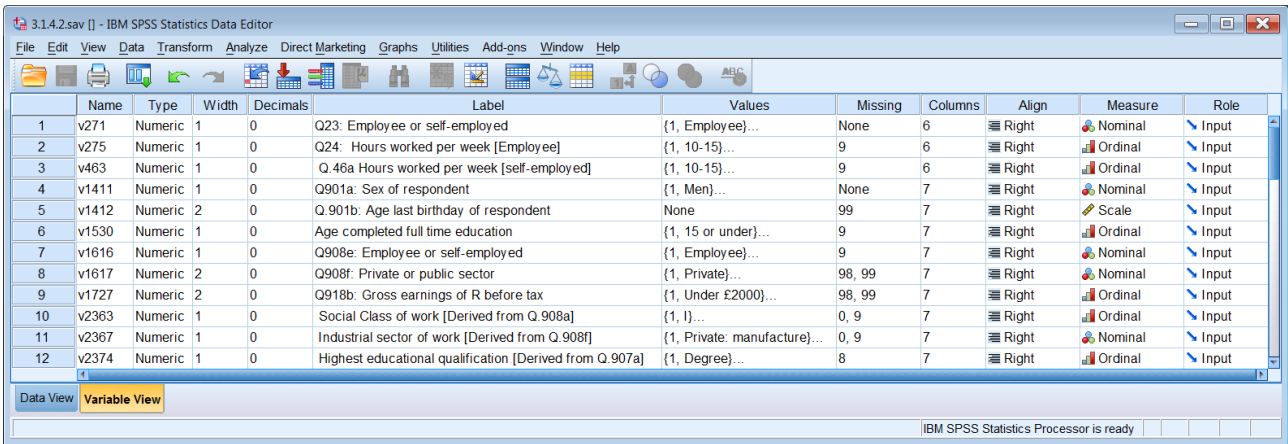
v2367 Industrial sector of work [Derived from Q.908f]

	Frequency	Percent	Valid Percent	Cumulative Percent
1 Private: manufacture	710	23.5	28.8	28.8
2 Private: non-manufacture	935	30.9	37.9	66.7
Valid 3 Public: service	682	22.5	27.6	94.3
4 Public: Manufacture + transport	140	4.6	5.7	100.0
Total	2467	81.6	100.0	
0	124	4.1		
Missing 9	434	14.3		
Total	558	18.4		
Total	3025	100.0		

v2374 Highest educational qualification [Derived from Q.907a]

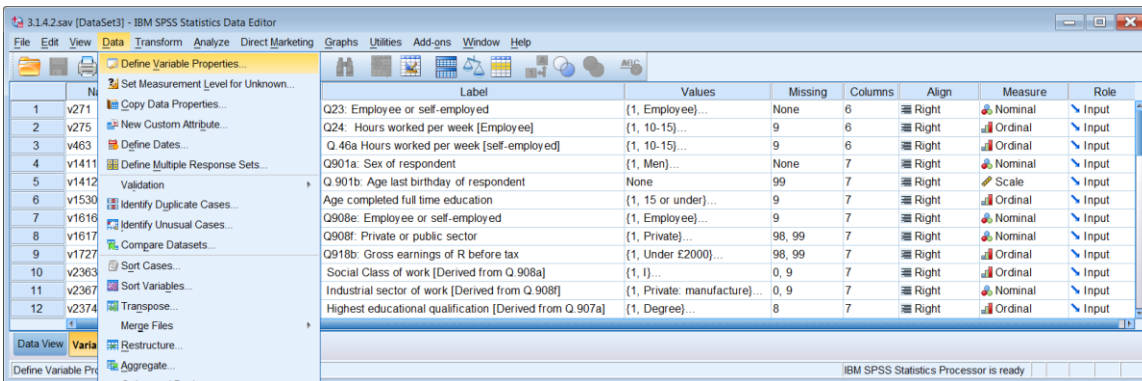
	Frequency	Percent	Valid Percent	Cumulative Percent
1 Degree	216	7.1	7.2	7.2
2 HE below degree	424	14.0	14.1	21.2
3 A-level or equiv.	304	10.0	10.1	31.3
Valid 4 O-level or equiv	536	17.7	17.8	49.1
5 CSE or equiv	242	8.0	8.0	57.1
6 Foreign and other	11	.4	.4	57.5
7 None	1283	42.4	42.5	100.0
Total	3016	99.7	100.0	
Missing 8 DK/NA	9	.3		
Total	3025	100.0		

These all seem to be OK, so go back to your **Data Editor** and press **[CTRL]S** to save the file



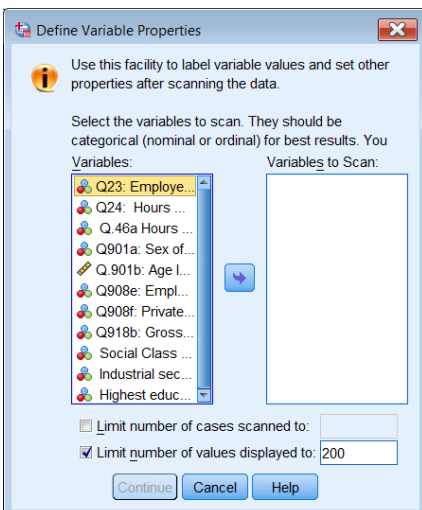
An alternative way of checking the file contents is:

Data > Define variable properties:

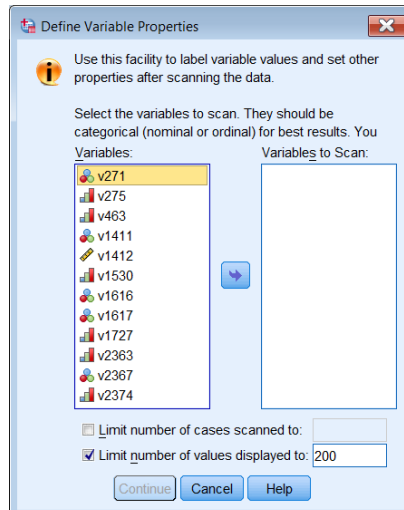


If labels are displayed these can be confusing to navigate: it's much easier to work with variable names.

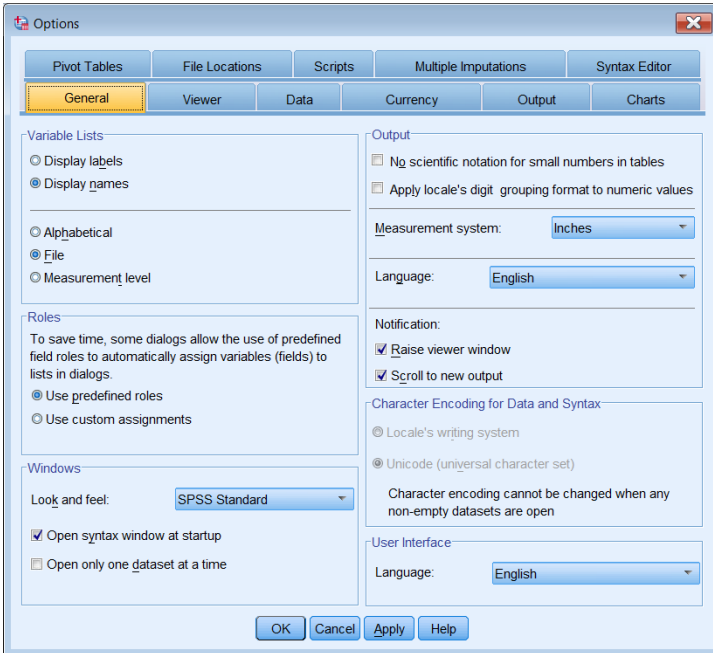
Labels



Names



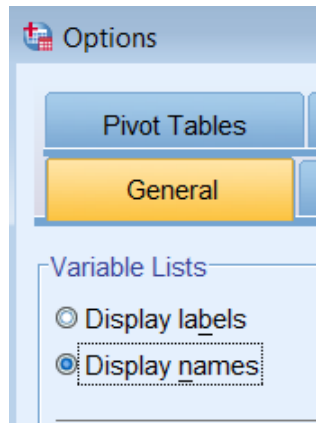
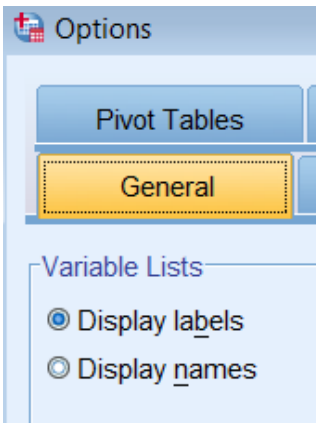
To do this we have to modify the SPSS settings with **Edit > Options**



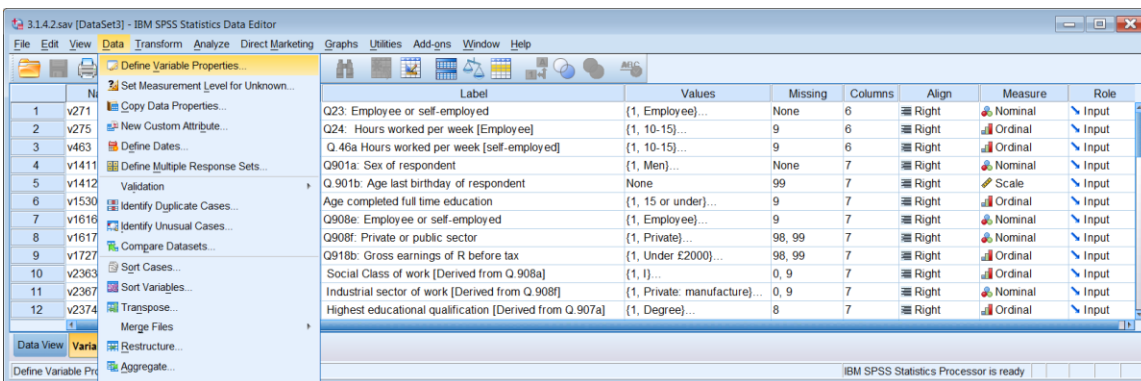
.. then click on the **General** tab and change:

Display labels

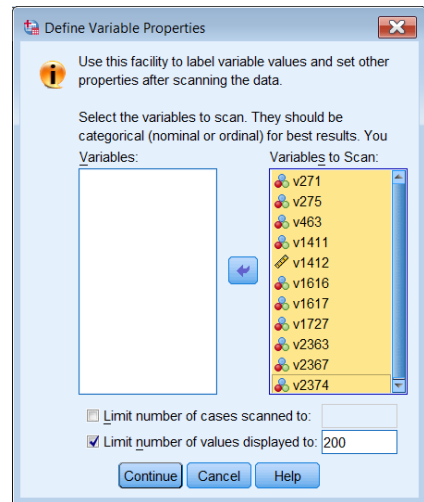
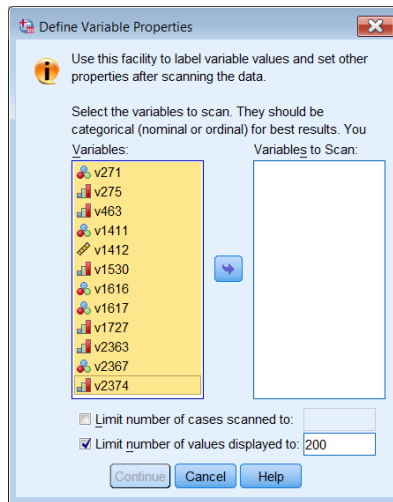
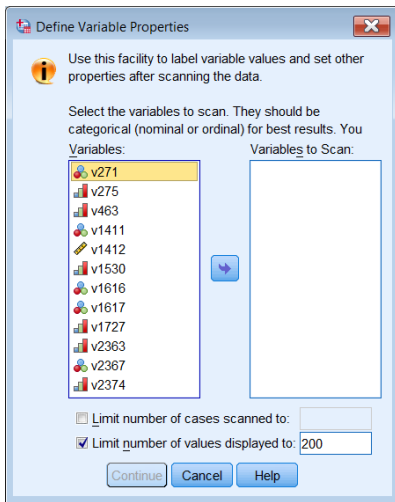
to Display names



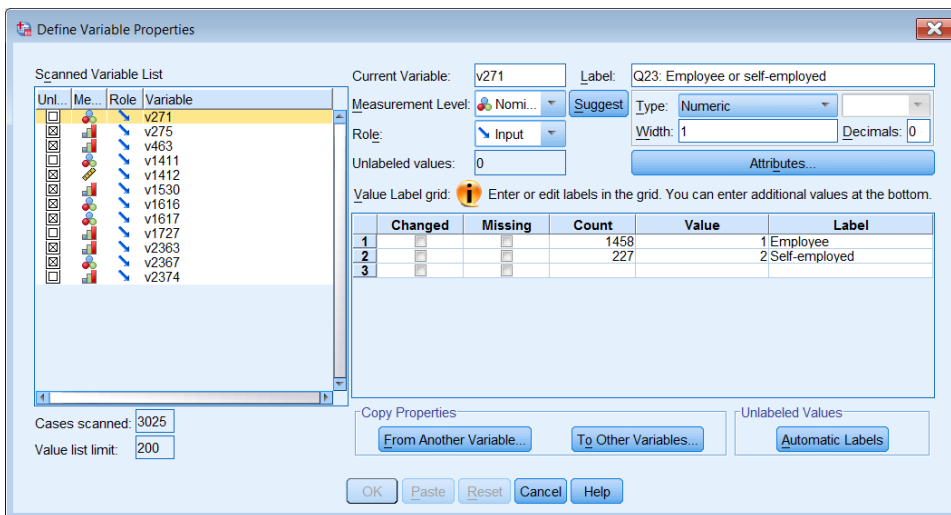
Data > Define variable properties:



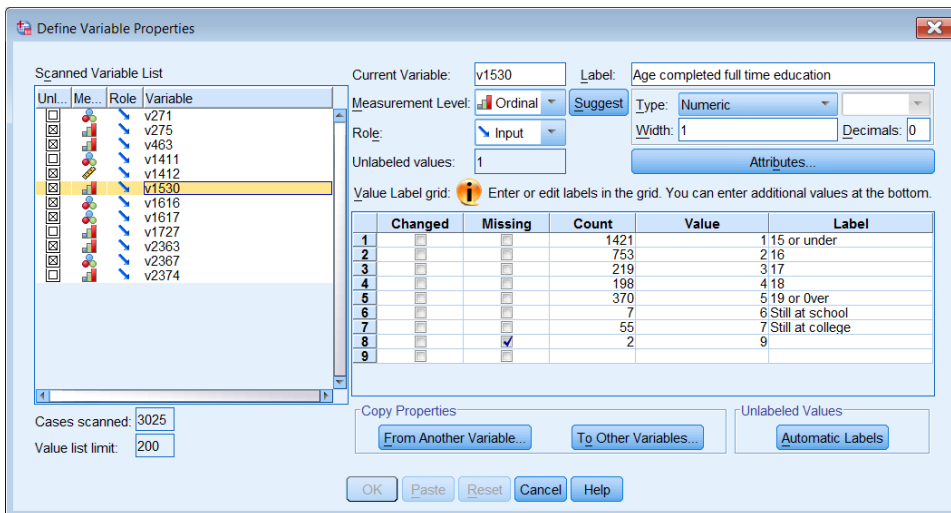
Highlight the variables you want and press  to transfer them to the **Variables to scan** box:



.. then click on **Continue:**



Just click on any variable name to display its properties: if you want, you can change any or all of the properties from this dialog box.



Define Variable Properties

Scanned Variable List

Uni...	Me...	Role	Variable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v271
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v275
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v463
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1411
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1412
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1530
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1616
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1617
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1727
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2363
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2367
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2374

Current Variable: v2363 Label: Social Class of work [Derived from Q.908a]

Measurement Level: Ordinal

Role: Input

Type: Numeric

Width: 1 Decimals: 0

Unlabeled values: 2

Value Label grid: Enter or edit labels in the grid. You can enter additional values at the bottom.

Changed	Missing	Count	Value	Label
<input type="checkbox"/>	<input type="checkbox"/>	124	0	
<input type="checkbox"/>	<input type="checkbox"/>	121	1	I
<input type="checkbox"/>	<input type="checkbox"/>	642	2	II
<input type="checkbox"/>	<input type="checkbox"/>	724	3	III non-manual
<input type="checkbox"/>	<input type="checkbox"/>	658	4	III manual
<input type="checkbox"/>	<input type="checkbox"/>	545	5	IV
<input type="checkbox"/>	<input type="checkbox"/>	156	6	V
<input type="checkbox"/>	<input checked="" type="checkbox"/>	55	9	

Cases scanned: 3025 Value list limit: 200

From Another Variable... To Other Variables... Automatic Labels

OK Paste Reset Cancel Help

Define Variable Properties

Scanned Variable List

Uni...	Me...	Role	Variable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v271
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v275
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v463
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1411
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1412
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1530
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1616
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1617
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	v1727
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2363
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2367
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	v2374

Current Variable: v2374 Label: Highest educational qualification [Derived from Q.907a]

Measurement Level: Ordinal

Role: Input

Type: Numeric

Width: 1 Decimals: 0

Unlabeled values: 0

Value Label grid: Enter or edit labels in the grid. You can enter additional values at the bottom.

Changed	Missing	Count	Value	Label
<input type="checkbox"/>	<input type="checkbox"/>	216	1	Degree
<input type="checkbox"/>	<input type="checkbox"/>	424	2	HE below degree
<input type="checkbox"/>	<input type="checkbox"/>	304	3	A-level or equiv.
<input type="checkbox"/>	<input type="checkbox"/>	536	4	O-level or equiv
<input type="checkbox"/>	<input type="checkbox"/>	242	5	CSE or equiv
<input type="checkbox"/>	<input type="checkbox"/>	11	6	Foreign and other
<input type="checkbox"/>	<input type="checkbox"/>	1283	7	None
<input type="checkbox"/>	<input checked="" type="checkbox"/>	9	8	DK/NA

Cases scanned: 3025 Value list limit: 200

From Another Variable... To Other Variables... Automatic Labels

OK Paste Reset Cancel Help

That's probably enough for one session, but use the tables above to think about appropriate cutting points to reduce the number of categories in the test variables. If your brain is not yet addled, you can continue straight to the next session [3.1.4.3 Income differences for test variables](#).

End of session 3.1.4.2: Income differences working file

Next sessions:

3.1.4.3: Income differences for test variables

Reduce gross earnings [v1727] to three categories: produce two-way contingency tables to investigate income differences for each 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 investigate income differences for the derived test variables.

3.2.4: Income differences - Elaboration

Three-way contingency tables to see what happens to income differences between men and women when controlling for test variables.

Back to: [Block 3 Analysing two variables \(and sometimes three\)](#)
[3.1 Two variables \(CROSSTABS\)](#)
[3.1.4.1 Income differences work-through](#)

Forward to: [3.1.4.3 Income differences for test variables](#) [d]