

Block 2: Analysing one variable

Interval and ratio variables

2.2.1.5 Specimen answer for frequencies exercise

[4 December 2010]

Exemplar: [British Social Attitudes](#) (1986 survey)

- Question:**
1. What is the distribution of the number of people in a household, including the respondent? What is the average number of persons per household?
 2. What is the age distribution of the sample and what is its average age? What shape does the distribution have? Where are the cutting points for the oldest 10% and 25% and the youngest 10% and 25%?

File: **mybsa86_4.sav**SPSS commands¹ used:**TITLE****FREQUENCIES**

For clarity and ease of understanding SPSS I've used SPSS/PASW 18 colour coding.

Commands are shown in **dark blue**, **sub-commands** in **green** and **keywords** in **dark red**. SPSS displays your own type in black (it's quicker in lower case and with abbreviated commands, but use the full syntax if you prefer). The blank lines and tabs are not necessary, but are used here for clarity. If you write your syntax in a separate Word file and the copy it across, SPSS doesn't mind what font or colour you use.

```
title 'Frequencies for interval scale variables (bsa, 1986)' .
```

```
frequencies v1508
  /barchart
  /statistics mean median mode .
```

```
frequencies v1512
  /format notable
  /histogram
  /statistics all
  /percentiles 10,25,75,90 .
```

Likewise, SPSS sub-commands and specifications need only be one space after the main command, but it is clearer if you use tabs.

Remember that commands always start in column one. If your specification continues over more than one line, you must leave at least one space in the first column. The first column can only be used for commands. **Don't forget the full stops (periods) !**

¹ General formats:

```
TITLE '<Any text>' .
```

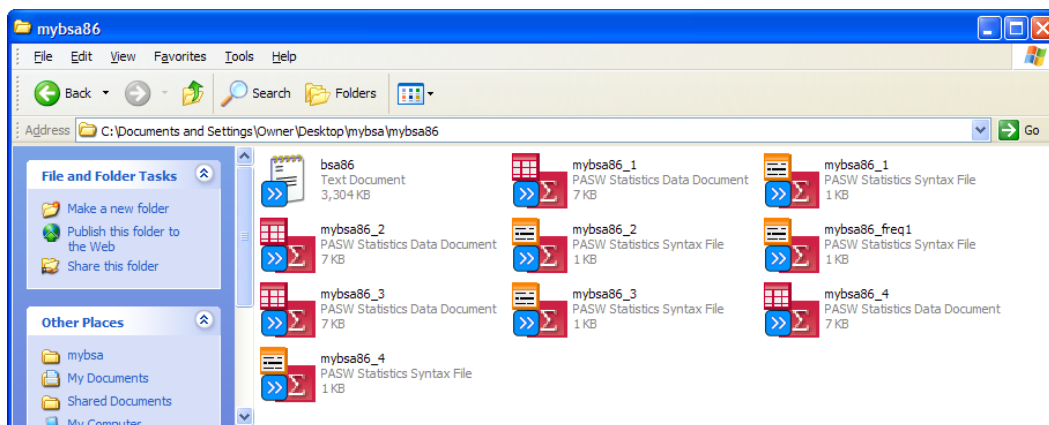
```
FREQUENCIES <varlist>
  /STATISTICS <statistics list>
  / <graphics options>
  / <other options>
```

Note that with variables occupying data fields of more than one column, the first column in the field determines the variable name. Thus the correct **positional** names for number of persons and age of respondent are **V1508** and **V1512**: there are no such variables as V1509 or V1513. Note further that SPSS normally reads only the first 4 characters of commands and the first 3 of keywords in the specification field. The specification field is **free-format**. Provided SPSS can make sense of what you want (it looks for spaces, slashes and beginning of next command) you can continue on the same line, and in either **UPPER** or **lower** case. Variable names will always be printed back exactly as stored (in older versions they may be in **UPPER** case): text in primes will always be printed back exactly as originally keyed in.

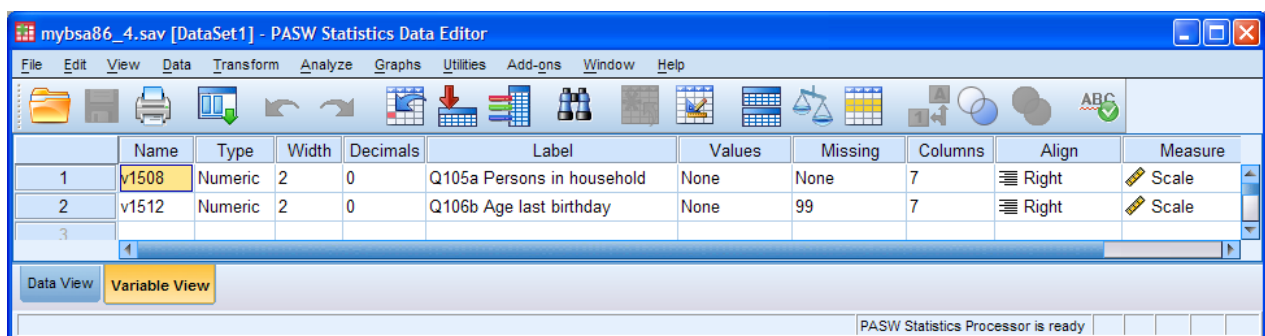
Thus the following will also work (but SPSS will not colour code the syntax):

```
freq      v1508 /hba /sta mea med mod .
freq      v1512 /for not /his /sta all /per 10 25 75 90 .
```

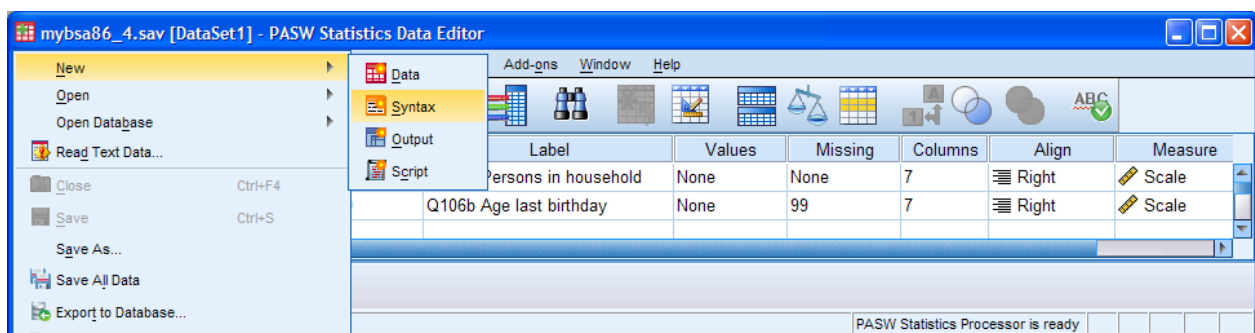
Go to your **mybsa86** folder



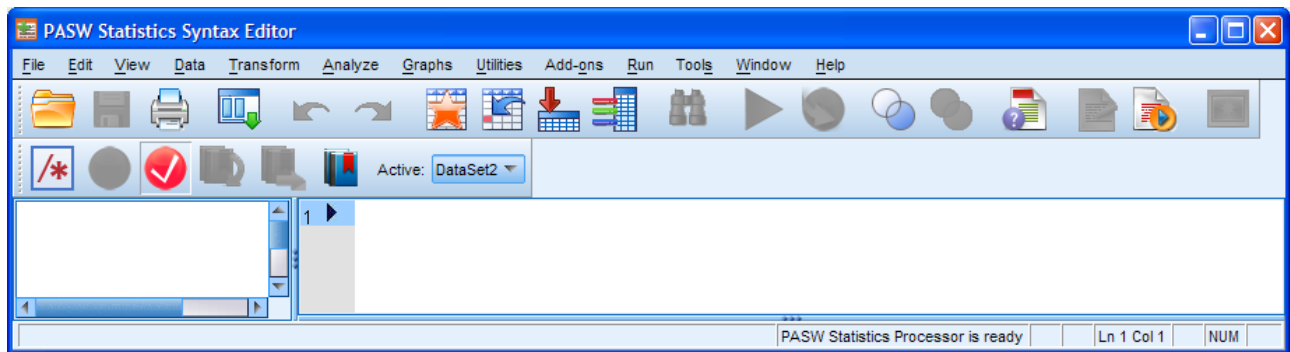
Double-click on **mybsa86_4.sav**.



Click on **File** > **New** > **Syntax**

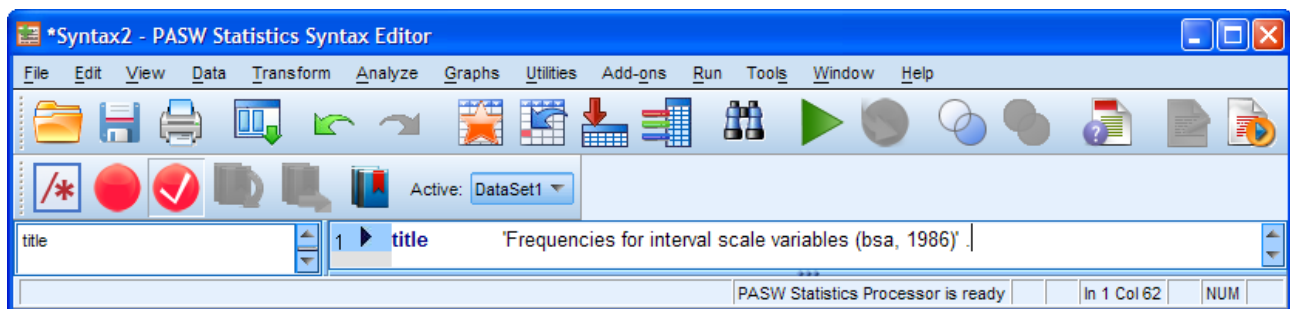


. . . to open a new syntax editor:



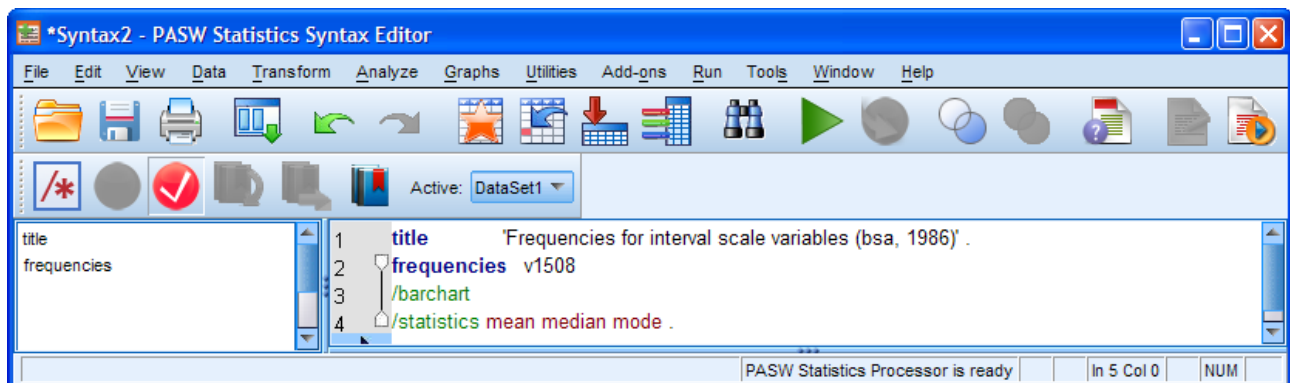
Give your job a title:

title 'Frequencies for interval scale variables (bsa, 1986)' .



Write the first **FREQUENCIES** command:

frequencies v1508
/barchart
/statistics mean median mode .



Run the command to produce:

Statistics

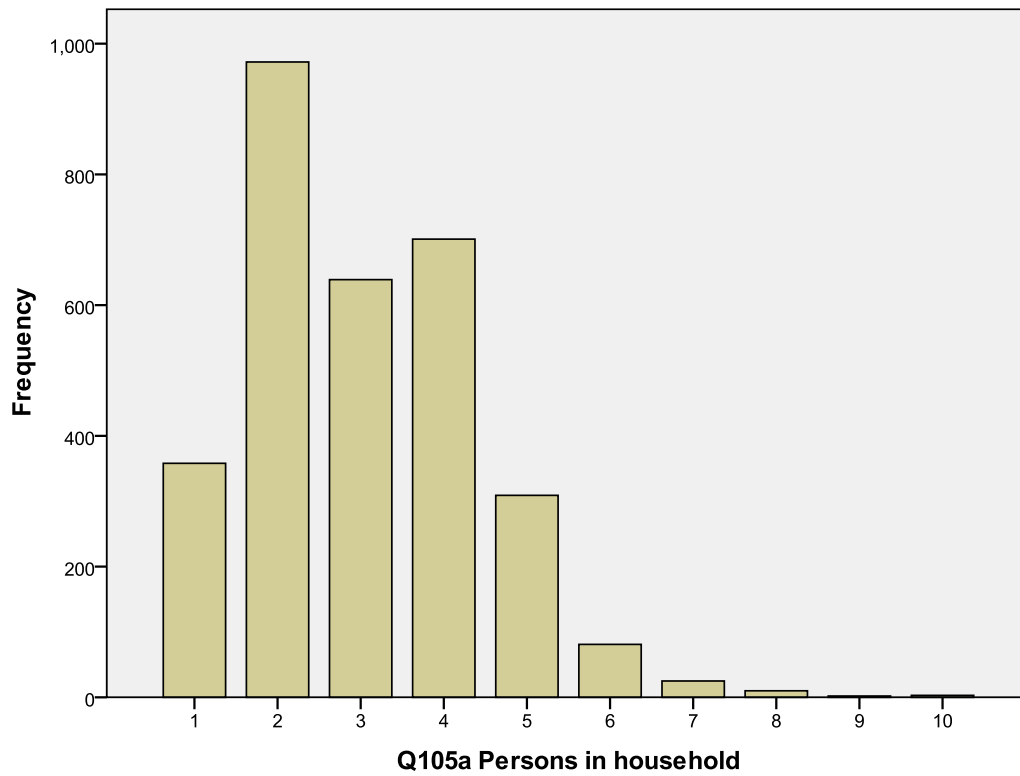
Q105a Persons in household

N	Valid	3100
	Missing	0
Mean		3.02
Median		3.00
Mode		2

Q105a Persons in household

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	358	11.5	11.5	11.5
	2	972	31.4	31.4	42.9
	3	639	20.6	20.6	63.5
	4	701	22.6	22.6	86.1
	5	309	10.0	10.0	96.1
	6	81	2.6	2.6	98.7
	7	25	.8	.8	99.5
	8	10	.3	.3	99.8
	9	2	.1	.1	99.9
	10	3	.1	.1	100.0
Total		3100	100.0	100.0	

Q105a Persons in household

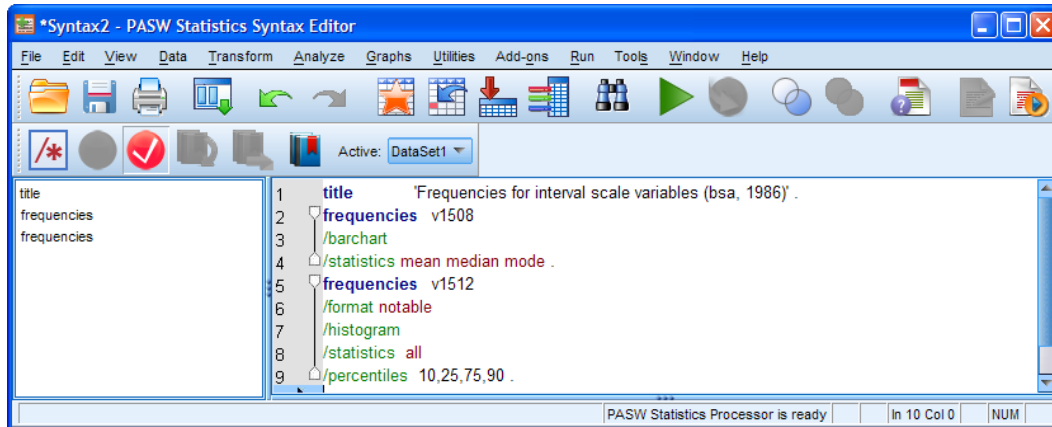


Now write the second **FREQUENCIES** command:

```

frequencies v1512
/format notable
/histogram
/statistics all
/percentiles 10,25,75,90 .

```



Run it, and when you get your results, look for something SPSS has missed (or rather gives a warning about). What is it?

Running this job produces:

Statistics

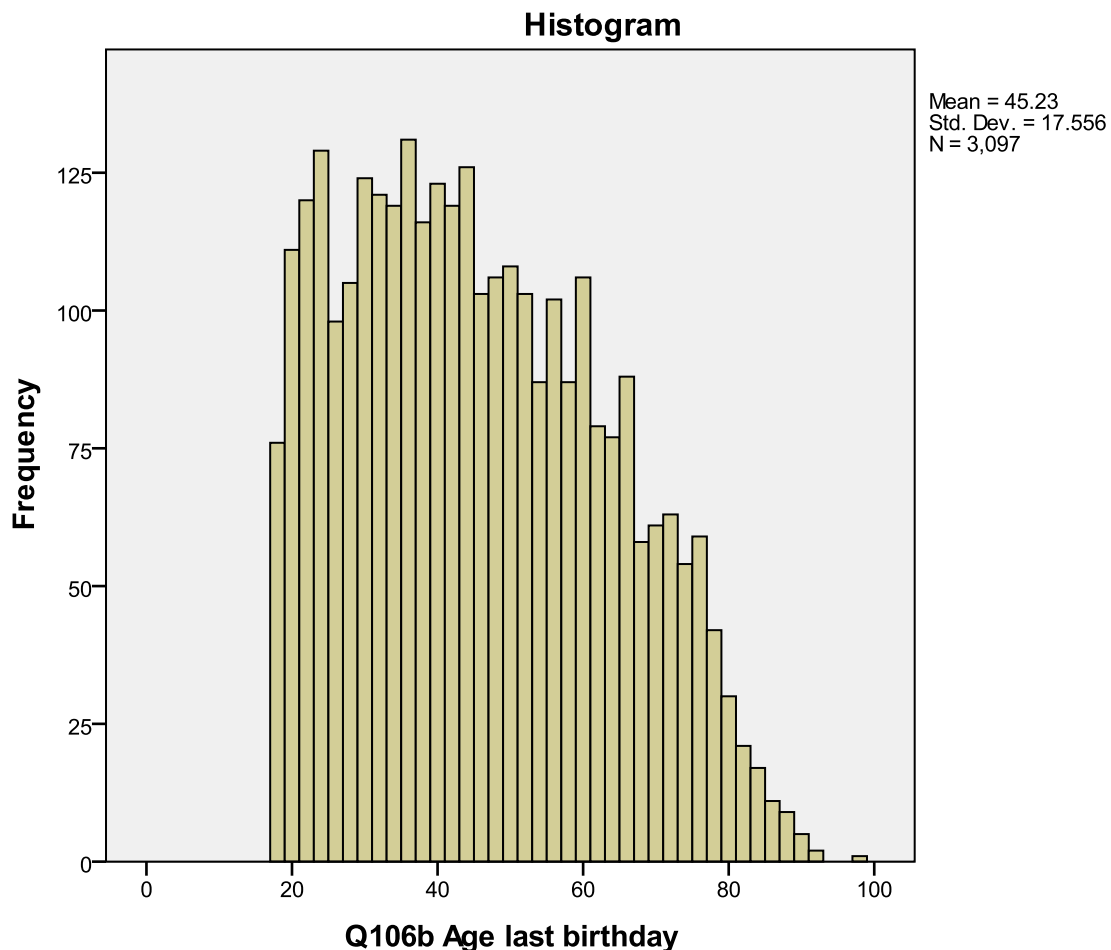
Q106b Age last birthday

N	Valid	3097
	Missing	3
Mean		45.23
Std. Error of Mean		.315
Median		43.00
Mode		18 ^a
Std. Deviation		17.556
Variance		308.209
Skewness		.334
Std. Error of Skewness		.044
Kurtosis		-.827
Std. Error of Kurtosis		.088
Range		79
Minimum		18
Maximum		97
Sum		140080
Percentiles	10	23.00
	25	31.00
	75	59.00
	90	71.00

a. Multiple modes exist. The smallest value is shown

[NB: This is the first and also the very last time you'll be using **STATISTICS ALL**, especially if you don't know what they are or what they are for! At this level you won't be using most of them anyway. I can explain in non-mathematical language what they are and how they work, but that's another tutorial to write. For now you are referred to the [Statistics notes to accompany course](#), the recommended [textbooks](#) or to the explanations available from the SPSS menus via **Help > Statistics Coach**]

The histogram looks like this:



It differs from the barchart in that the bars are touching. This is because there is a known underlying metric (age in years last birthday) and the width of the bars therefore has meaning.

Running frequencies on actual age produces this huge table! It may be useful to inspect on-screen, but it's far too big to print. This is why the old specification **/FORMAT = CONDENSE** was so useful (and saved trees!). Now you can see why **/FORMAT NOTABLE** is sometimes needed.

Q106b Age of respondent last birthday

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18	76	2.5	2.5	2.5
	19	53	1.7	1.7	4.2
	20	58	1.9	1.9	6.0
	21	56	1.8	1.8	7.8

22	64	2.1	2.1	9.9
23	64	2.1	2.1	12.0
24	65	2.1	2.1	14.1
25	51	1.6	1.6	15.7
26	47	1.5	1.5	17.2
27	48	1.5	1.5	18.8
28	57	1.8	1.8	20.6
29	62	2.0	2.0	22.6
30	62	2.0	2.0	24.6
31	74	2.4	2.4	27.0
32	47	1.5	1.5	28.5
33	60	1.9	1.9	30.5
34	59	1.9	1.9	32.4
35	66	2.1	2.1	34.5
36	65	2.1	2.1	36.6
37	56	1.8	1.8	38.4
38	60	1.9	1.9	40.4
39	68	2.2	2.2	42.6
40	55	1.8	1.8	44.3
41	54	1.7	1.7	46.1
42	65	2.1	2.1	48.2
43	76	2.5	2.5	50.6
44	50	1.6	1.6	52.2
45	52	1.7	1.7	53.9
46	51	1.6	1.6	55.6
47	48	1.5	1.5	57.1
48	58	1.9	1.9	59.0
49	64	2.1	2.1	61.1
50	44	1.4	1.4	62.5
51	44	1.4	1.4	63.9
52	59	1.9	1.9	65.8
53	39	1.3	1.3	67.1
54	48	1.5	1.5	68.6
55	51	1.6	1.6	70.3
56	51	1.6	1.6	71.9
57	50	1.6	1.6	73.5
58	37	1.2	1.2	74.7
59	52	1.7	1.7	76.4
60	54	1.7	1.7	78.1
61	41	1.3	1.3	79.5
62	38	1.2	1.2	80.7
63	30	1.0	1.0	81.7

64	47	1.5	1.5	83.2
65	41	1.3	1.3	84.5
66	47	1.5	1.5	86.0
67	26	.8	.8	86.9
68	32	1.0	1.0	87.9
69	29	.9	.9	88.8
70	32	1.0	1.0	89.9
71	33	1.1	1.1	90.9
72	30	1.0	1.0	91.9
73	23	.7	.7	92.6
74	31	1.0	1.0	93.6
75	27	.9	.9	94.5
76	32	1.0	1.0	95.5
77	19	.6	.6	96.2
78	23	.7	.7	96.9
79	21	.7	.7	97.6
80	9	.3	.3	97.9
81	11	.4	.4	98.2
82	10	.3	.3	98.5
83	9	.3	.3	98.8
84	8	.3	.3	99.1
85	8	.3	.3	99.4
86	3	.1	.1	99.5
87	4	.1	.1	99.6
88	5	.2	.2	99.7
89	2	.1	.1	99.8
90	3	.1	.1	99.9
91	1	.0	.0	99.9
92	1	.0	.0	100.0
97	1	.0	.0	100.0
Total	3097	99.9	100.0	
Missing 99	3	.1		
Total	3100	100.0		

End of session

Now do the homework exercises on the same variables from the 1989 survey.

Homework: 2.2.1.6 Homework exercises

[\[Back to Block 2 menu\]](#)