3.5.2.6 The COUNT command 2 – Sexism

[Screenshots: SPSS15, last updated: 24 Aug 2011, previous version 1 October 2009]

Exemplar: Fifth form survey fifthx.sav
Variable to be derived: Negative attitudes to women (Sexism)
Source variables: V248, V252, V253, V256, V261, V251, V255, V259, V260

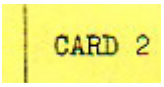
COUNT creates additive indices. It counts, for each case, the number of times a specified value or set of values occurs for a variable or variables on a criterion variable list and enters this as the value for a new variable defined by the **COUNT** command.

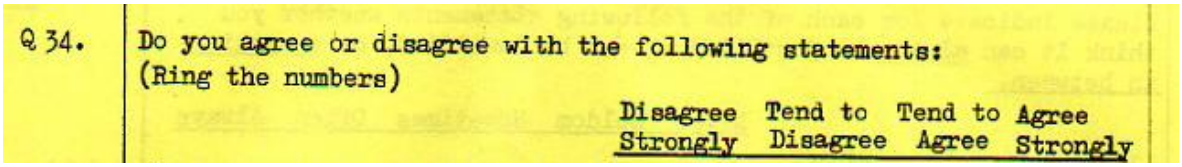
The general format is:

COUNT **<newvar> = <criterion varlist> (<value list>)**
/ <newvar> = <criterion varlist> (<value list>) .

…in which the user has to supply the new variable name(s), the variables to be scanned and the list, **including the round brackets**, of value(s) to be included in the search. You can create several derived variables within a single **COUNT** command, provided you stay in the specification field and provided you separate each new specification with a slash '**/**'.

In the fifth form survey we have already used **COUNT** to create a new variable **STATQUO** by counting the number of **Tend to Agree** or **Agree Strongly** responses to four items in question Q.34 (items o,p,q,r) replicated from a scale developed by Himmelweit to measure "attachment to status quo" among teenagers.


CARD 2


Q 34. Do you agree or disagree with the following statements: (Ring the numbers)

	Disagree Strongly	Tend to Disagree	Tend to Agree	Agree Strongly	
o) It is best to be like the others and not to stand out from the rest.	1	2	3	4	(75)
p) People who are content with what they have will have a better life than those who are always trying to improve their position.	1	2	3	4	(76)
q) We are all born to our various social positions and it won't do to change them.	1	2	3	4	(77)
r) The greatest source of happiness in life is to be satisfied with whatever you have.	1	2	3	4	(78)

Sexism

Question Q.33 is more complex. It consists of 14 statements measuring opinions about women, some negative, some positive, with which pupils can agree or disagree on a 4-point scale. (Codes 1-4 on columns 48 to 61 of card 2 read into SPSS using positional naming convention as **v248 to v261**)

Q33. Here are some statements made about women, We would like to know if you agree or disagree with them. (Please put a ring round the number which indicates your answer).

	Disagree Strongly	Disagree	Agree	Agree Strongly	
a) Careers are fine for women but real fulfilment is a home and kids.	1	2	3	4	(48)
b) Women should not expect men to pay for them when dating etc.	1	2	3	4	(49)
c) Half of all top jobs should be reserved for women.	1	2	3	4	(50)
d) It is a good thing that women can become airline pilots, plumbers etc.	1	2	3	4	(51)
e) Women are too emotional.	1	2	3	4	(52)
f) Women are not as ambitious as men.	1	2	3	4	(53)
g) Women are as intelligent as men.	1	2	3	4	(54)
h) Women do not need to be beautiful to be successful	1	2	3	4	(55)
j) Husbands rather than wives should have the final voice in family matters.	1	2	3	4	(56)
k) There is no difference in brain-power between men and women	1	2	3	4	(57)
l) If women are paid as much as men they should pay for themselves when dating etc.	1	2	3	4	(58)
m) Women should get equal pay for doing the same work as men.	1	2	3	4	(59)
n) Beauty contests are degrading to women and should stop.	1	2	3	4	(60)
o) Romantic love is dead	1	2	3	4	(61)

Again, it is possible to construct a crude index of "Sexism" from nine of these items, five of which are negative (a,e,f,j,o) and four positive (d,h,m,n). To be a sexist a pupil must **agree** with the negative items (a,e,f,j and o) and **disagree** with the positive items (d,h,m and n). Thus we need to construct an index which **counts** the number of **agreements** (codes 3 and 4) with the first set together with the number of **disagreements** (codes 1 and 2) with the second. This will yield a score in the range 0 to 9 where 9 indicates high sexism.

To use **COUNT** to generate a new variable containing the number of **agreements** with negative

items (a,e,f,j,o) and also **disagreements** with positive items (d,h,m,n) we first need to decide on a name. Remember, variable names in SPSS must begin with a letter. In 1981 they were limited to 8 characters, but now you can have up to 64. However, it's best to keep names short as otherwise the **Data Editor** gets very unwieldy. Let us call it **SEXISM**. Next, we have to specify the existing variables to be included in the criterion list for **SEXISM**. The anti-women items are **V248, V252, V253, V256** and **V261**, and the pro-women items are **V251, V255, V259** and **V260**.

A small table helps:

New variable name	Variables in criterion list	Values to be counted
SEXISM	**V248, V252, V253, V256, V261** **V251, V255, V259, V260**	**3** and **4** **1** and **2**

The SPSS command is therefore:

```
COUNT       SEXISM   =   V248 V252 V253 V256 V261 (3 , 4)  V251 V255 V259 V260 (1 , 2) .
```

Note the full stop to terminate the command.

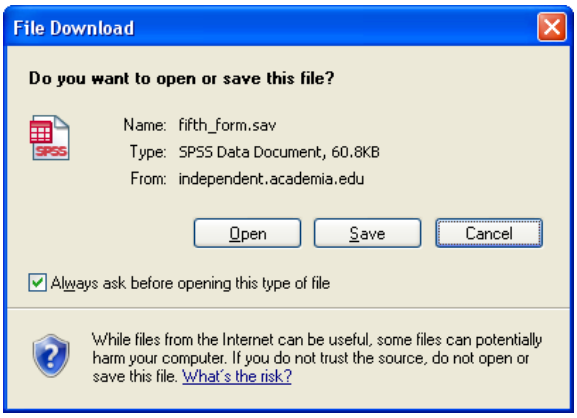
This line is a trifle long and so the author's preference in such cases is to break up the command over more lines and tab the sub-commands and specifications inwards to make it easier to edit or spot mistakes, e.g.

```
COUNT       SEXISM   =   V248 V252 V253 V256 V261 (3,4)
                         V251 V255 V259 V260 (1,2) .
```

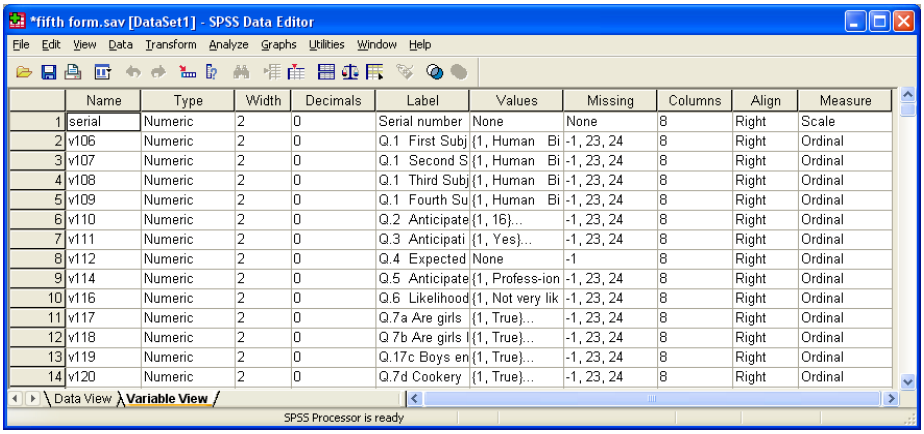
This command creates a new variable **SEXISM** with values in the range 0 to 9 where 9 indicates high sexism (negative attitudes to women).

[NB: You won't actually be able to do this next bit without immediate direct access to SPSS: don't worry just follow the tutorial anyway.]

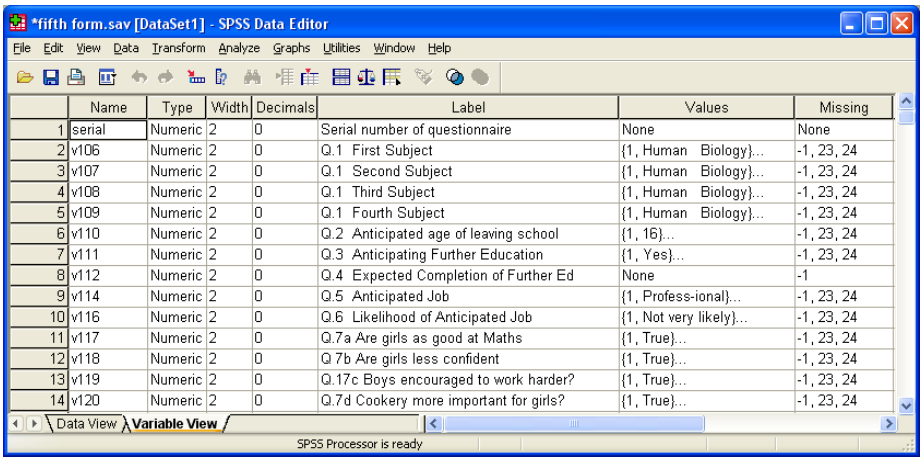
First we need to access the SPSS saved file for the fifth form survey, so if you haven't already got it open, start by downloading Fifth Form Survey



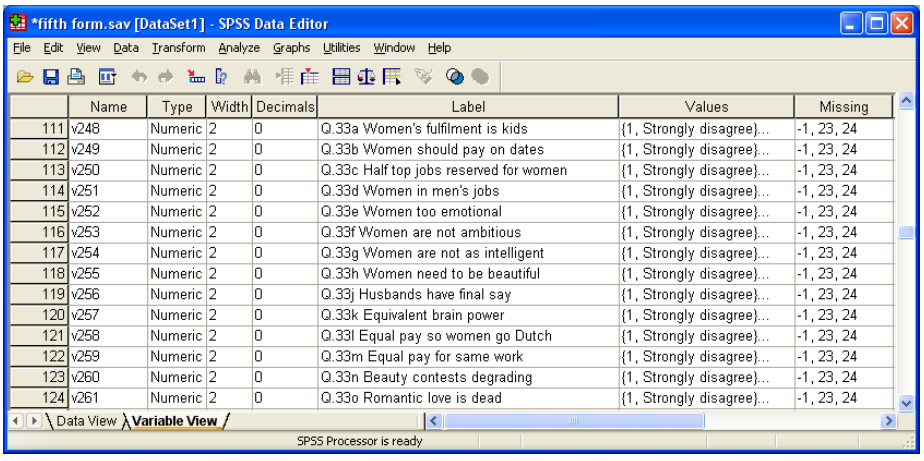
Click on **Open** to get[1]:



Use the mouse to drag the column separators around and make the display easier to read, then reduce the pane to include only essential basic information:



Scroll down to find variables **V248** to **V261**



Now you know why positional variable names and informative variable labels are so important!

The variables and question numbers are so much easier to find inside the Data Editor and you can work straight from the original questionnaire to the data set and *vice versa*.

[1] The size of the Data Editor displayed will depend on how it was left when last used. You can drag the corners of the pane to reduce or enlarge it, and the edges up or down and left or right, using the left mouse button. This pane is adjusted to display only the first few variables in the file.

To create our new variable **STATQUO** (attachment to status quo) refer back to the table:

New variable name	Variables in criterion list	Values to be counted
SEXISM	**V248, V252, V253, V256, V261** **V251, V255, V259, V260**	**3** and **4** **1** and **2**

..and also to the general format of the **COUNT** command:

COUNT **<newvar> = <criterion varlist> (<value list>)**

..where **<value list>** must contain a single value, a range of values, or several individual values separated by commas.

If you want to use the drop-down menus to create **SEXISM** feel free to try. In fact, using the menus, you cannot get **COUNT** to derive a single variable from two criterion lists with different values: you have to create two intermediate variables, one for each list, and then add them together with **COMPUTE** to get the final score. (See fully worked example on pages 11 – 18)

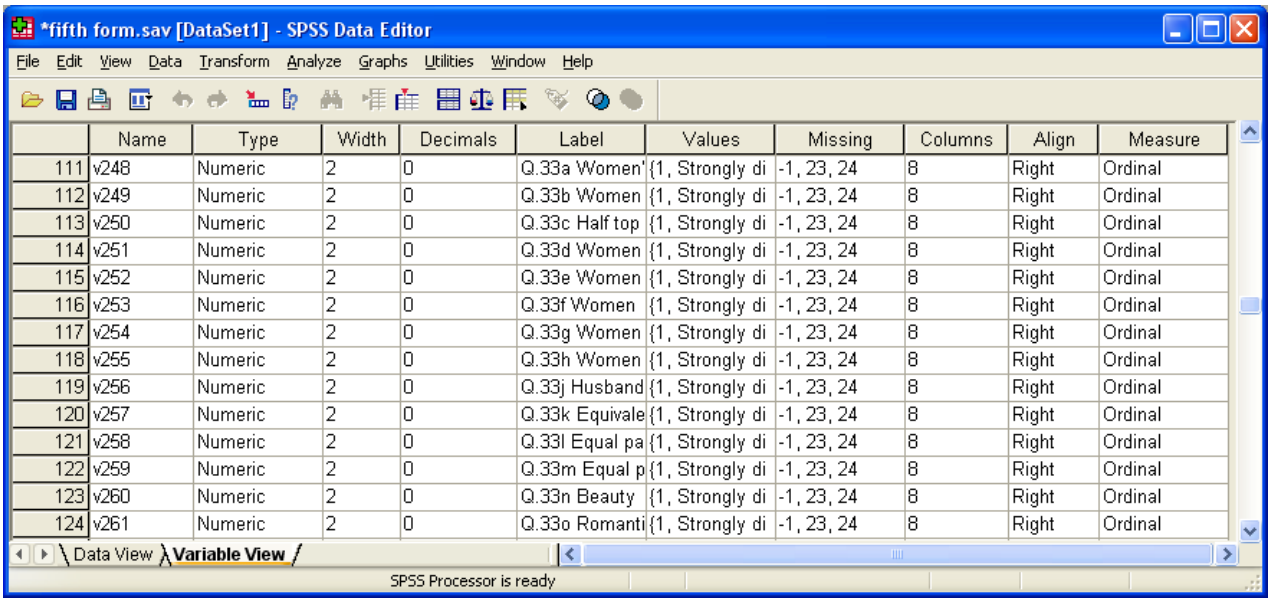
The syntax way

Our SPSS command is:

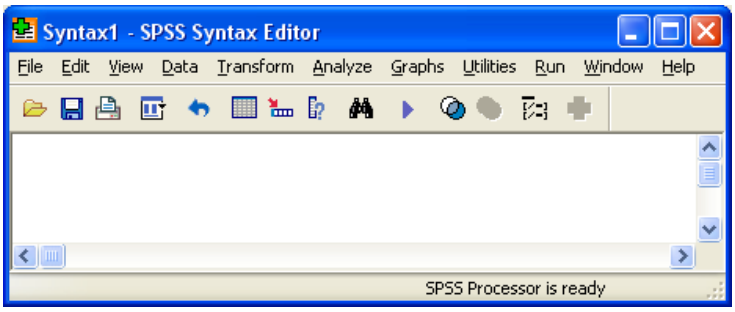
```
COUNT       SEXISM    =    V248 V252 V253 V256 V261 (3,4)
                           V251 V255 V259 V260 (1,2).
```

It is possible to generate syntax using PASTE from the SPSS drop-down menus, but it's extremely confusing and tiresome, albeit error free. My preference is to use commands written directly to a syntax file, which is simpler, quicker and much easier to follow and understand.

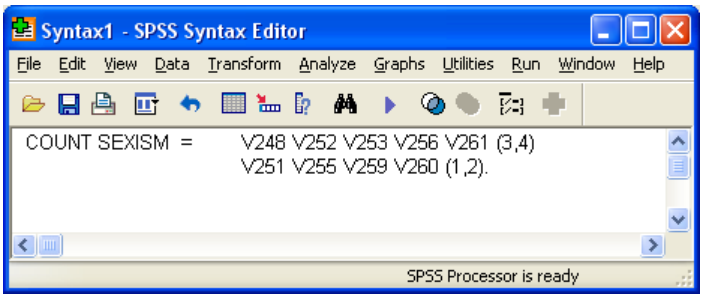
Go back to the Data editor:



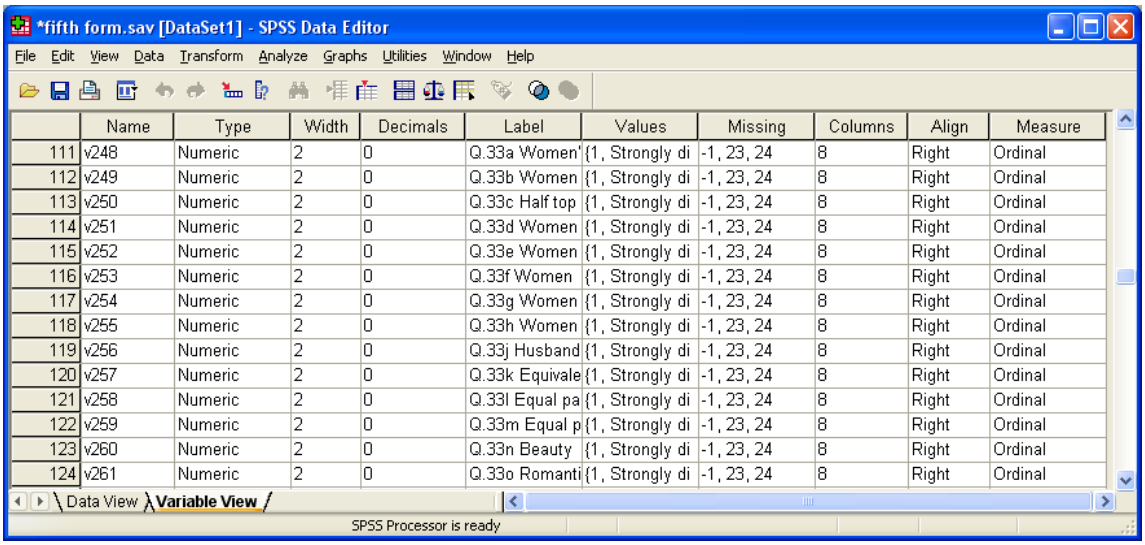
File
…New
…Syntax:



and type in your command:

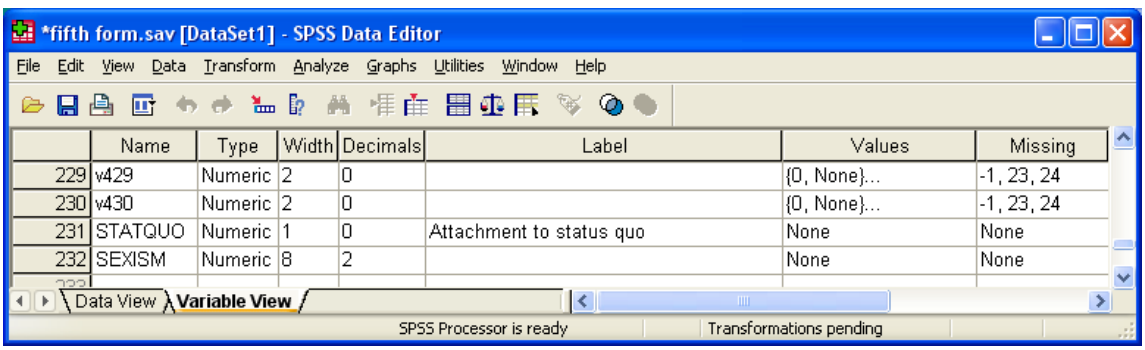


Don't forget the full stop! Leave the cursor in or on the line and click on **Run..Current** or press **[CTRL]+R** to run it. Nothing appears to happen, but if you go back to the data editor:



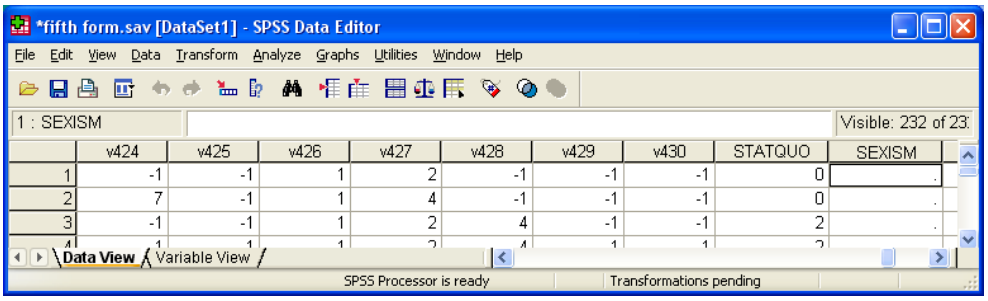
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
111	v248	Numeric	2	0	Q.33a Women'	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
112	v249	Numeric	2	0	Q.33b Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
113	v250	Numeric	2	0	Q.33c Half top	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
114	v251	Numeric	2	0	Q.33d Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
115	v252	Numeric	2	0	Q.33e Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
116	v253	Numeric	2	0	Q.33f Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
117	v254	Numeric	2	0	Q.33g Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
118	v255	Numeric	2	0	Q.33h Women	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
119	v256	Numeric	2	0	Q.33j Husband	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
120	v257	Numeric	2	0	Q.33k Equivale	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
121	v258	Numeric	2	0	Q.33l Equal pa	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
122	v259	Numeric	2	0	Q.33m Equal p	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
123	v260	Numeric	2	0	Q.33n Beauty	{1, Strongly di	-1, 23, 24	8	Right	Ordinal
124	v261	Numeric	2	0	Q.33o Romanti	{1, Strongly di	-1, 23, 24	8	Right	Ordinal

..and scroll down to the last row:

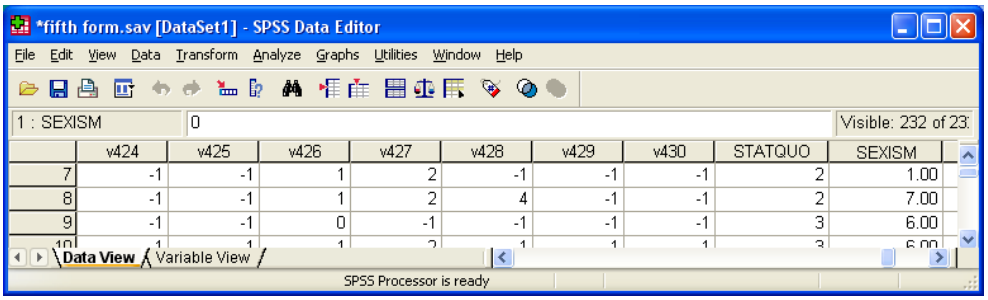


	Name	Type	Width	Decimals	Label	Values	Missing
229	v429	Numeric	2	0		{0, None}...	-1, 23, 24
230	v430	Numeric	2	0		{0, None}...	-1, 23, 24
231	STATQUO	Numeric	1	0	Attachment to status quo	None	None
232	SEXISM	Numeric	8	2		None	None

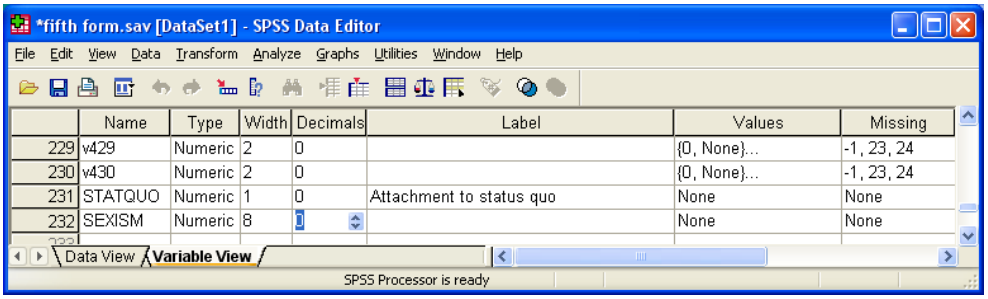
Your new variable **SEXISM** has been defined in the last row of the file. If you saved your own file last time, **STATQUO** will also be there. If you click on **Data View** then press **[CTRL] + ►** to see the last column, you will see that no values have been calculated yet.



This is because SPSS is still waiting for a statistical command such as **FREQUENCIES** or for an interim **EXECUTE**. Go back to the syntax file, type **EXECUTE** on the next line then run it. Now check the Data Editor again. The **SEXISM** column has now filled up with the scores generated by **COUNT**.

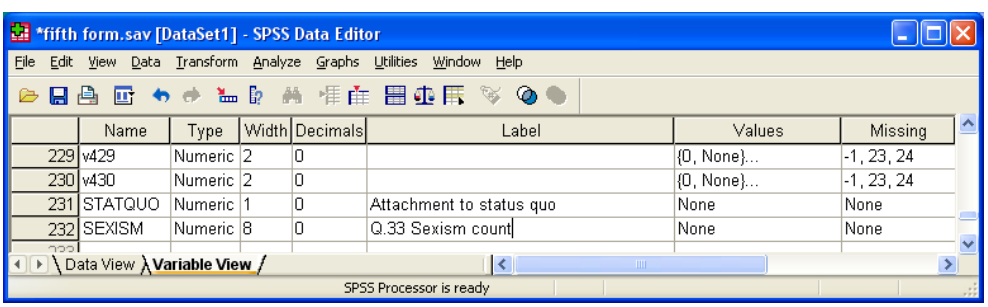


We don't actually need two decimal places for the integer variable **SEXISM** so the file can be tidied up by clicking on **Variable View** and manually changing 2 to 0 in the Data Editor **Decimals** column for **SEXISM**.

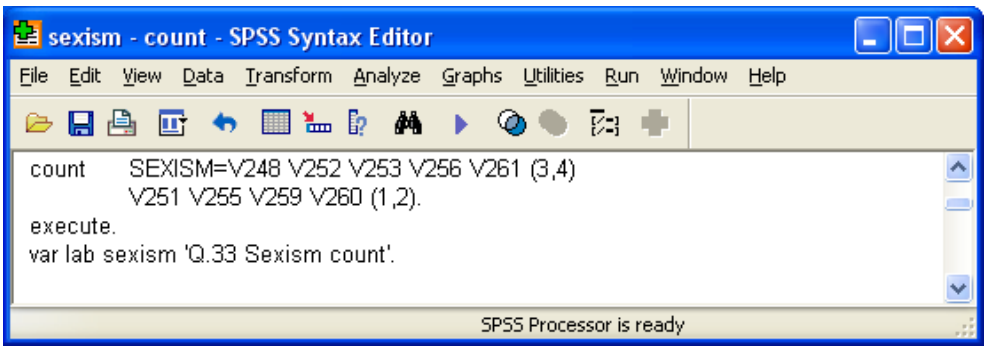


Although the **SEXISM** column has now filled up with scores generated by **COUNT**, if you go back to **Variable View** there is nothing to tell you (or other users) what **SEXISM** is, so it needs a label.

You can write labels directly into the **Label** column in the Data Editor (easy for just one or two variables)

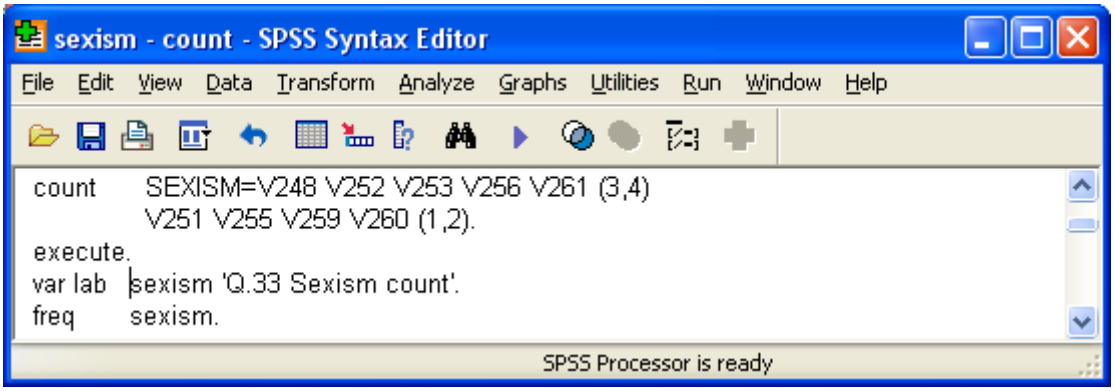


…but if you have to label a large number of variables it's probably better to use syntax.



Syntax is easy to edit, but for very long setup files it's preferable to compose all the SPSS syntax in Word (or other wordprocessor) edit it and then, when you're happy with it, copy it across to a syntax file. Remember, you can use lower case and abbreviations for commands and subcommands. SPSS normally only reads the first 3 or 4 characters anyway. So for **VARIABLE LABELS** you can simply write **var lab**. You soon get used to it and you're less likely to get RSI!

So far, so good. The new variable **SEXISM** has been saved at the end of the file, but we also need an analysis to check what the distribution looks like. A simple frequency count will do. Again, it is sufficient to write **freq** instead of **FREQUENCIES** Add a line **freq sexism.**



Run the job to get:

Q.33 Sexism count

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	14	9.9	9.9	9.9
	1	33	23.2	23.2	33.1
	2	32	22.5	22.5	55.6
	3	21	14.8	14.8	70.4
	4	11	7.7	7.7	78.2
	5	14	9.9	9.9	88.0
	6	11	7.7	7.7	95.8
	7	6	4.2	4.2	100.0
	Total	142	100.0	100.0	

This measure is quite crude because it only takes into account information from half the response on each item in the scale. It is also potentially inaccurate because it gives a score to every case even if one or more items is missing. Thus a score of 5 on the 0 to 9 sexism scale could be 5 out

of 9 or 5 out of 5 or anywhere in between.

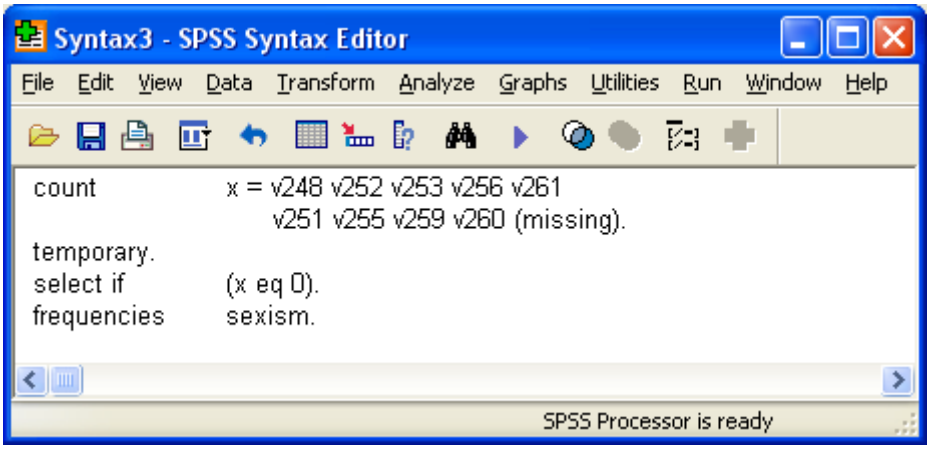
We still need to think about how to deal with the problem of missing values which are coded as ' **-1**' for all items on Q.34 and which have already been declared as missing in the file.

SPSS ignores these when using **COUNT** and can therefore give very distorted scores. It is possible to leave out pupils with missing items by counting the number of items with missing values and then selecting out only those pupils with no items missing.

With **COUNT** you'd have to do something like:

```
COUNT       X=V248 v252 v253 v256 V261
            V251 v255 v259 v260  (MISSING).
SELECT IF   (X EQ 0).
```

before tabulating the **SEXISM** score. The full sequence for this example might then be:



Note the use of **temporary** which limits transformations to the next procedure, otherwise you run the risk of permanent changes to the data which you may not be able to retrieve.

The run produces the following tables:

Statistics

Q.34 Sexism count

N	Valid	114
	Missing	0

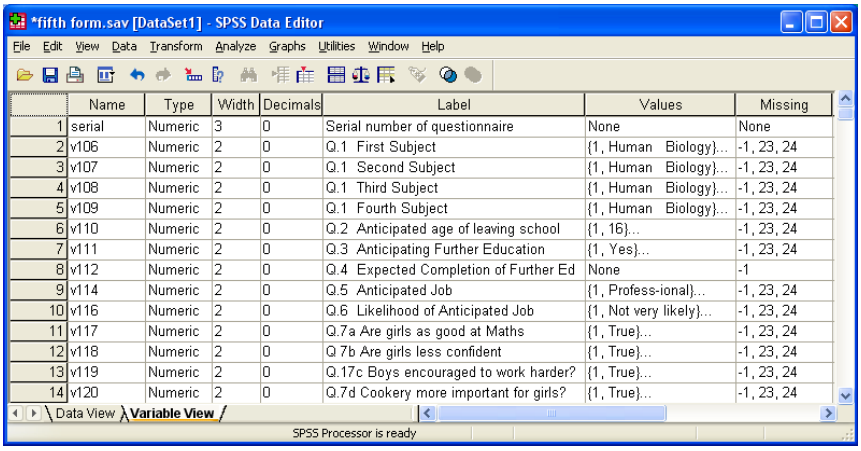
Q.34 Sexism count

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6	5.3	5.3	5.3
	1	31	27.2	27.2	32.5
	2	26	22.8	22.8	55.3
	3	15	13.2	13.2	68.4
	4	10	8.8	8.8	77.2
	5	11	9.6	9.6	86.8
	6	10	8.8	8.8	95.6
	7	5	4.4	4.4	100.0
	Total	114	100.0	100.0	

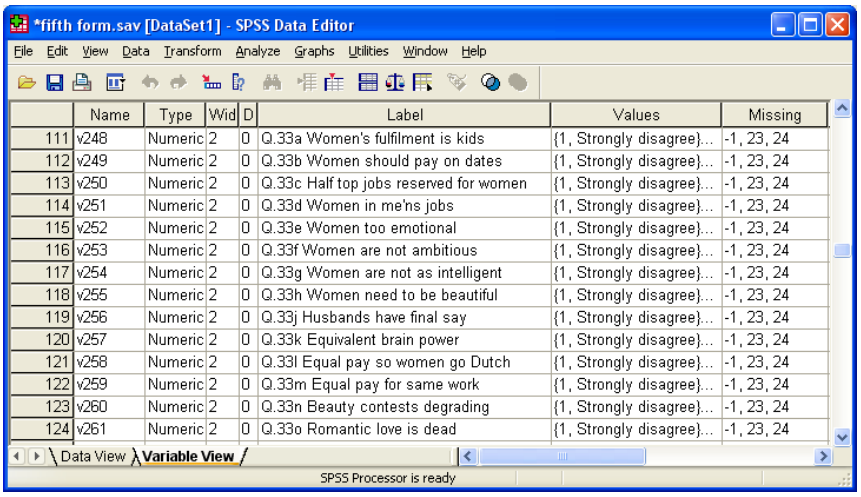
Note that the sample size has been drastically reduced from 142 to 114. This is what can happen when you combine many variables with missing values into one variable: it's why statisticians sometimes recommend substituting missing values with some central value such as mean, median or mode. At other times it's best to leave such cases out.

The point-and-click way

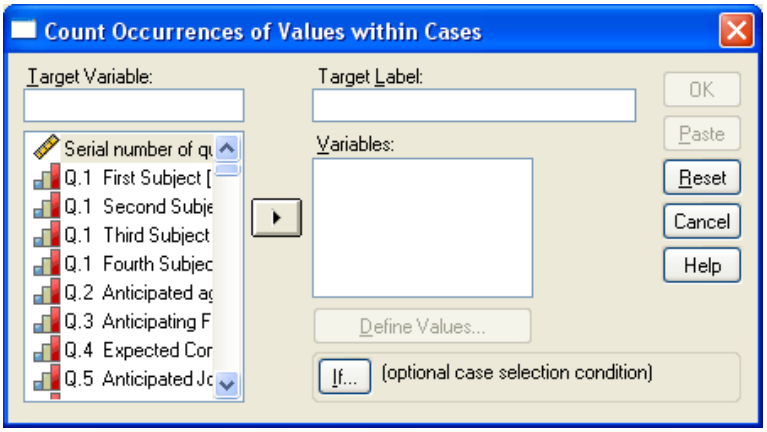
Go back to the **Data Editor**



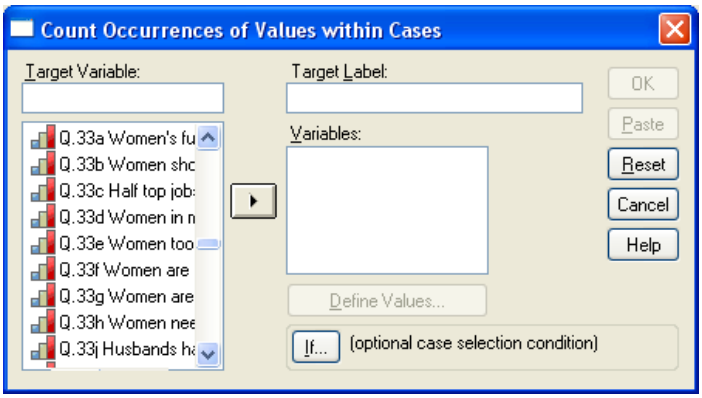
…and scroll to find **V248…V261**:



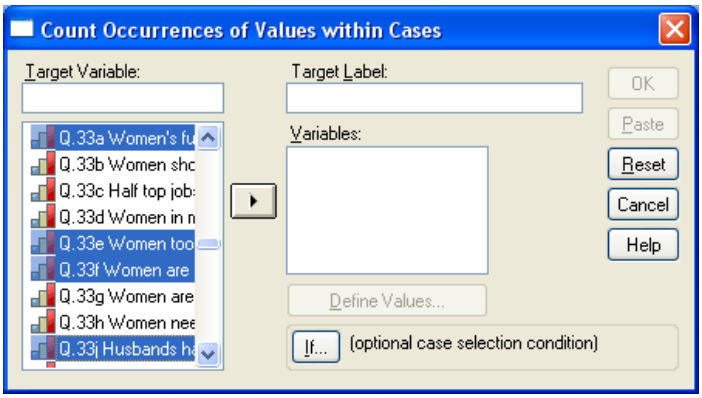
Transform
…Count values within cases



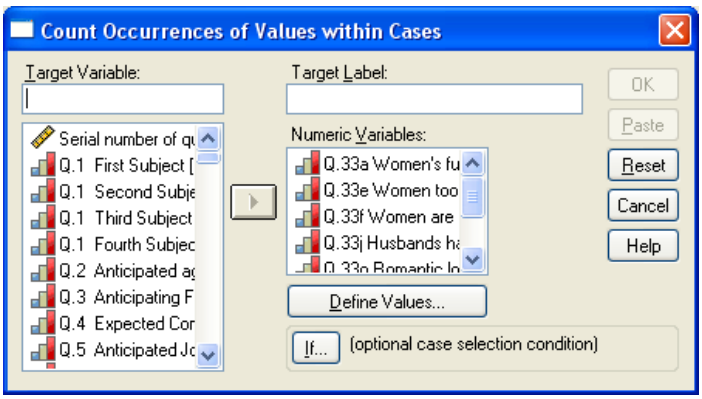
Scroll to find the items in Q.33:



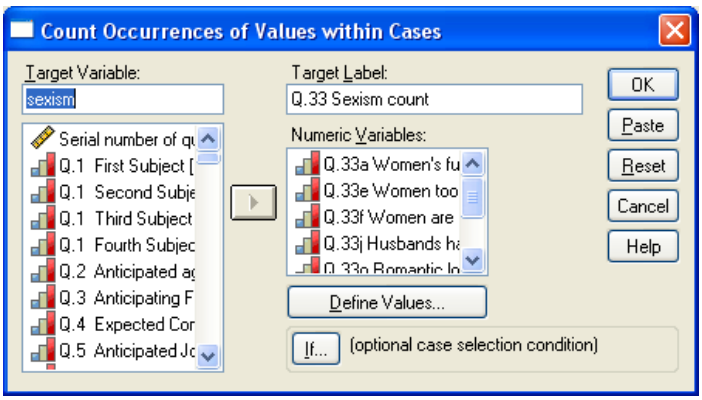
... highlight them one at a time and click on ▶ or highlight all five with **[CTRL] + left click**



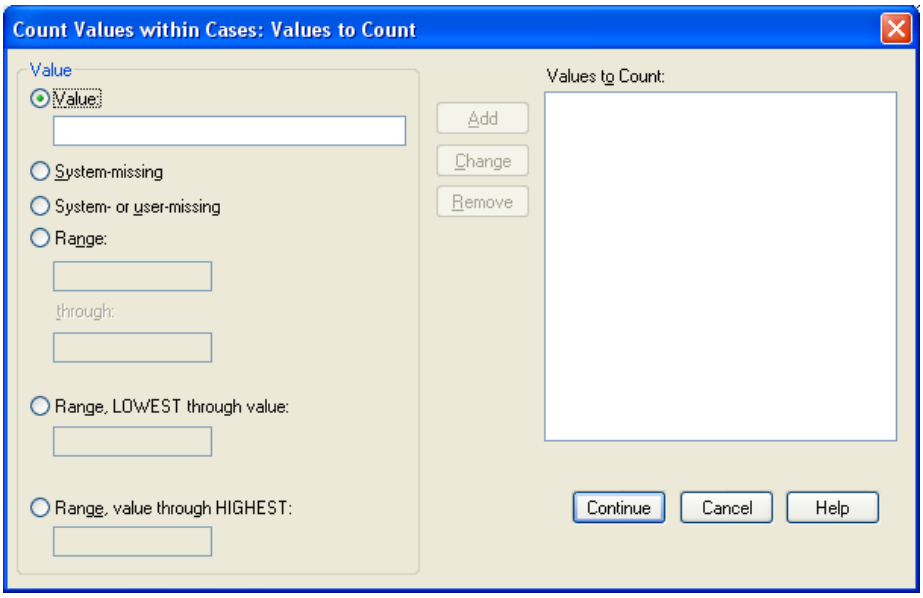
…and click on ▶ to drag them across to the **Variables** box.



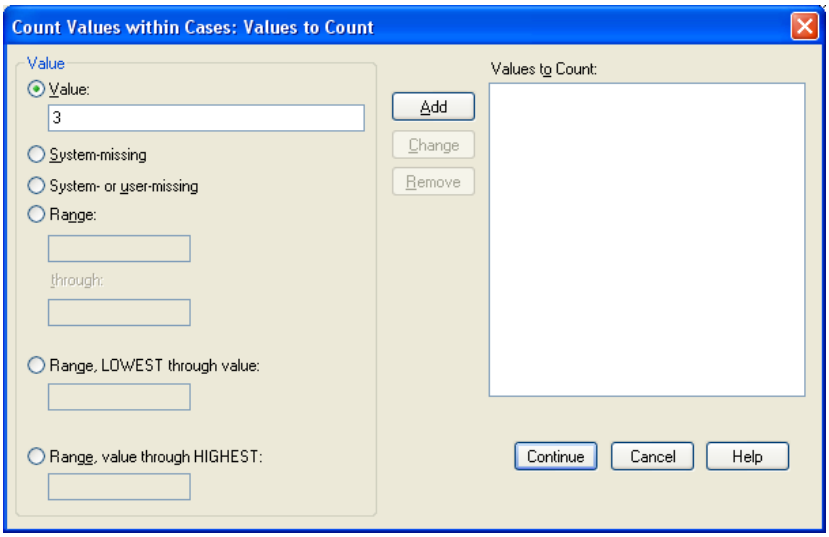
Now write **SEXISM** in the **Target Variable** box and your own label in the **Target Label** box.



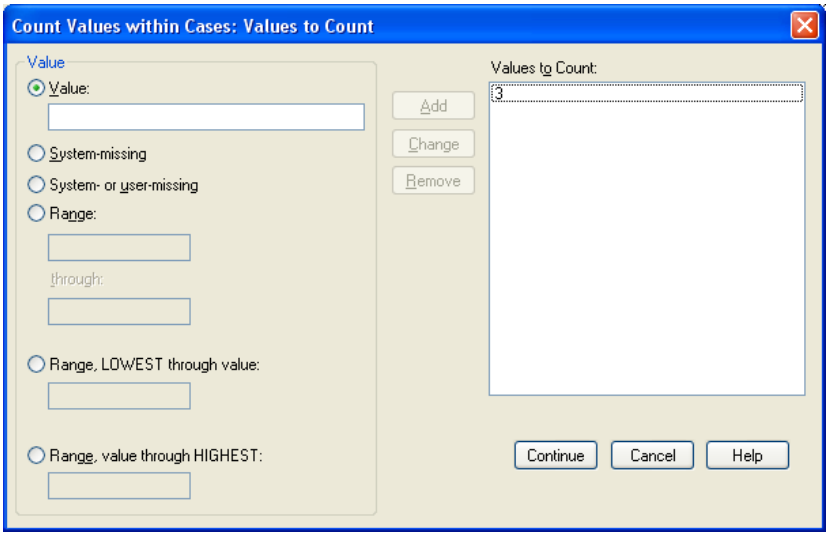
Click on **Define Values**:



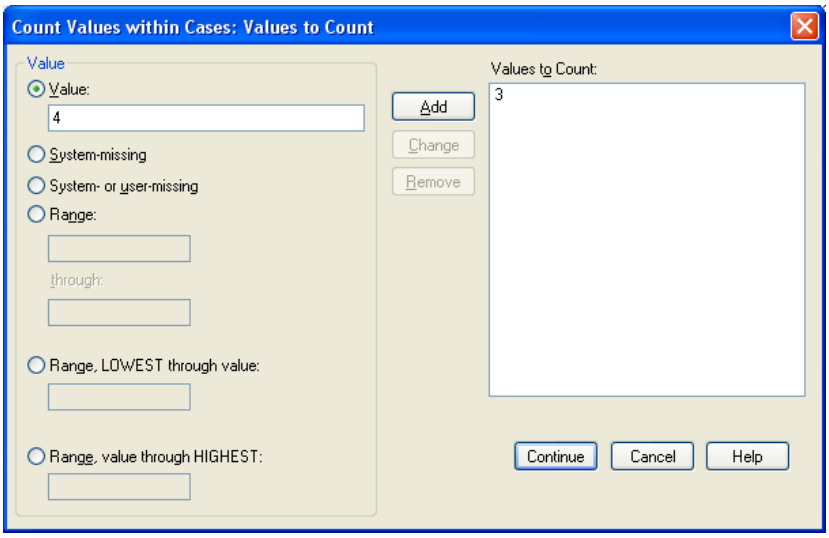
The dialog is already set for entry of values so type **3** in the **Value** box

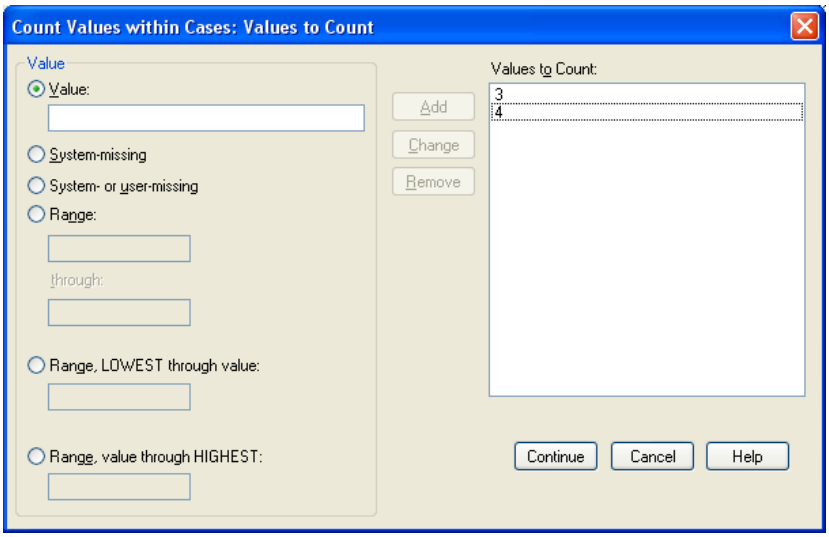


…and click on **Add** to transfer it to the **Values to Count** box:

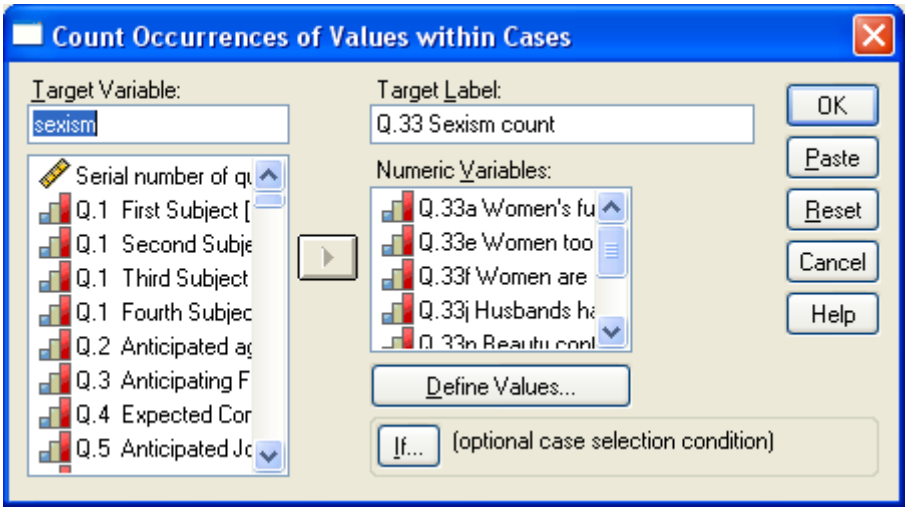


Now do the same with the value **4**

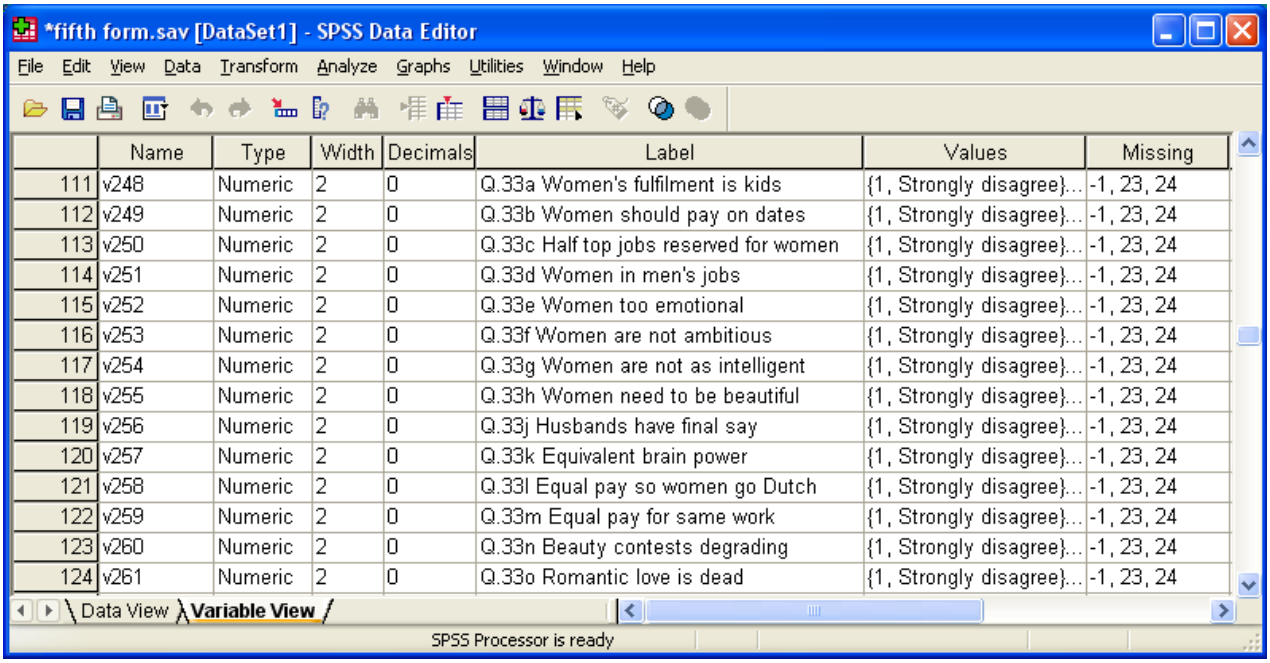




..and click **Continue**

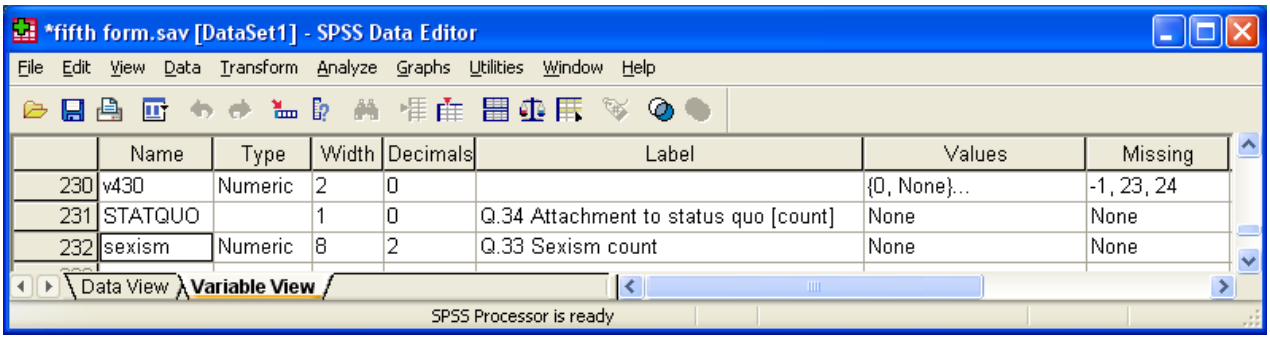


Bit of a surprise, eh? If you click **Define Values** again, you'll see your selection is still there, but it still gave me a bit of a fright. Now click OK :



Another surprise! What's going on? Don't worry, your new variable has been created and appended in the last row of the **Data Editor**.

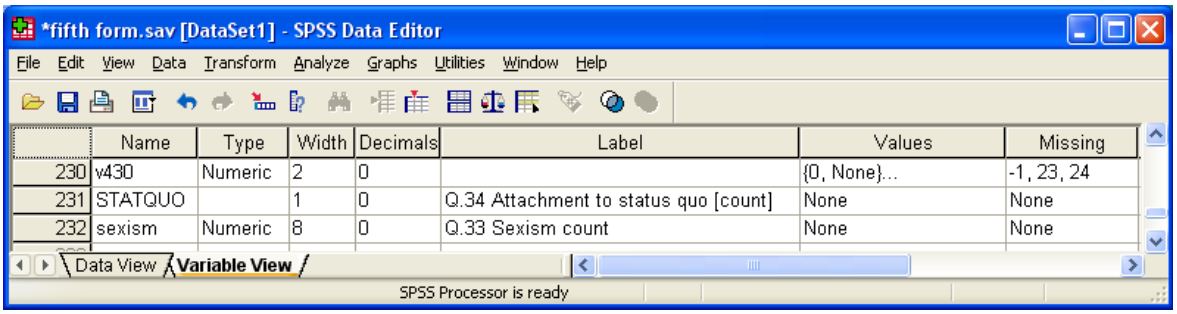
Scroll down to see it:



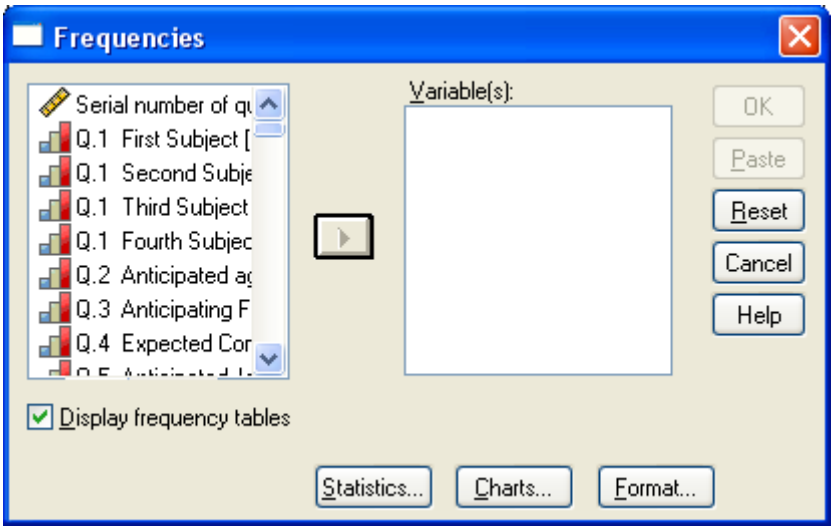
The syntax generated (automatically pasted to the Viewer window) is:

```
COUNT
  sexism = v248 v252 v253 v256 v260  (3)  v248 v252 v253 v256 v260  (4)  .
VARIABLE LABELS sexism 'Q.33 Sexism count' .
EXECUTE .
```

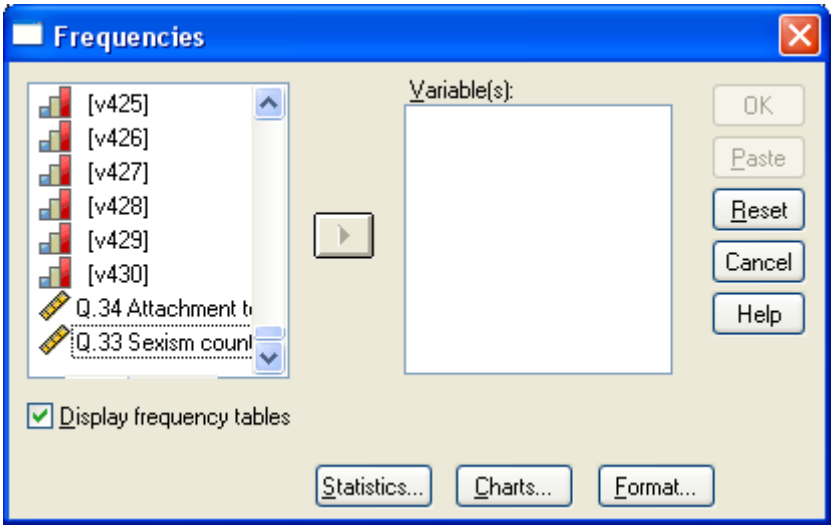
To get the frequency count go back to the **Data Editor** and scroll to the bottom of the file:



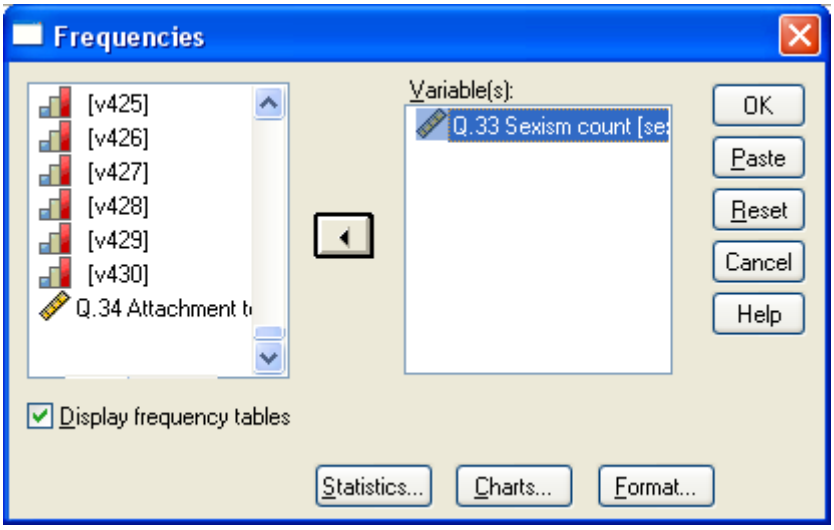
Analyze
…Descriptive statistics
…Frequencies



[CTRL] + [END] to skip to the end of the list:



Highlight and drag **SEXISM** to the **Variables** box:



Click OK for:

Statistics

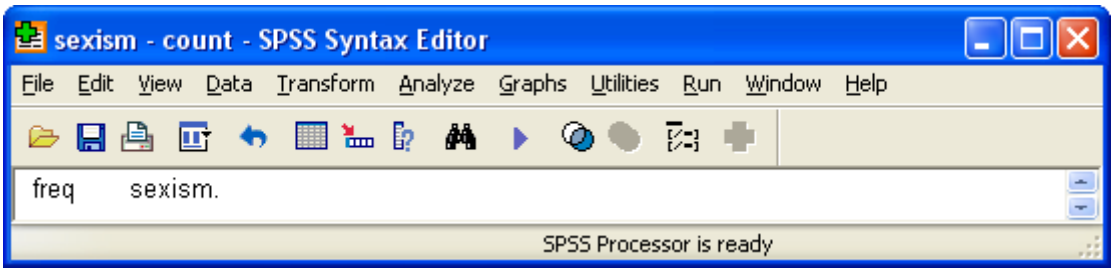
Q.33 Sexism count

N	Valid	142
	Missing	0

Q.33 Sexism count

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	50	35.2	35.2	35.2
	1	32	22.5	22.5	57.7
	2	28	19.7	19.7	77.5
	3	16	11.3	11.3	88.7
	4	12	8.5	8.5	97.2
	5	4	2.8	2.8	100.0
	Total	142	100.0	100.0	

..but it's **so much easier** to add a line to the syntax file and run that instead:



Next tutorial: 3.5.2.7 The COMPUTE command 2 - Sexism

Feedback on ease of understanding and use of tutorial, please, to: johnfhall@orange.fr