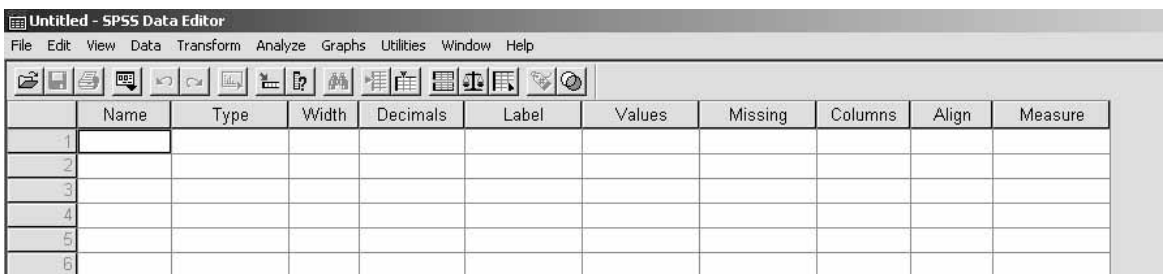
5.2. Heterogeneity in a goodness of fit test

EXAMPLE 5.2. The genetics of tepal colour in *Allium schoenoprasum*

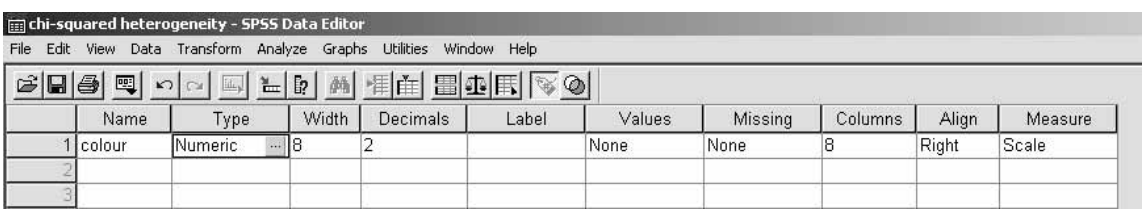
BOX 5.3. How to calculate a chi-squared test for heterogeneity

Step 1. Set up the variables.
(i) When SPSS starts, select the 'Type in Data' option.
(ii) Then choose 'Variable View' from the tabs at the bottom left. You will see a screen something like this:

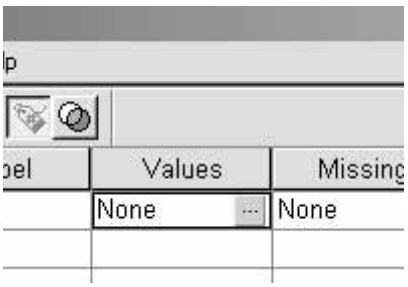


Each row represents a variable for the analysis.

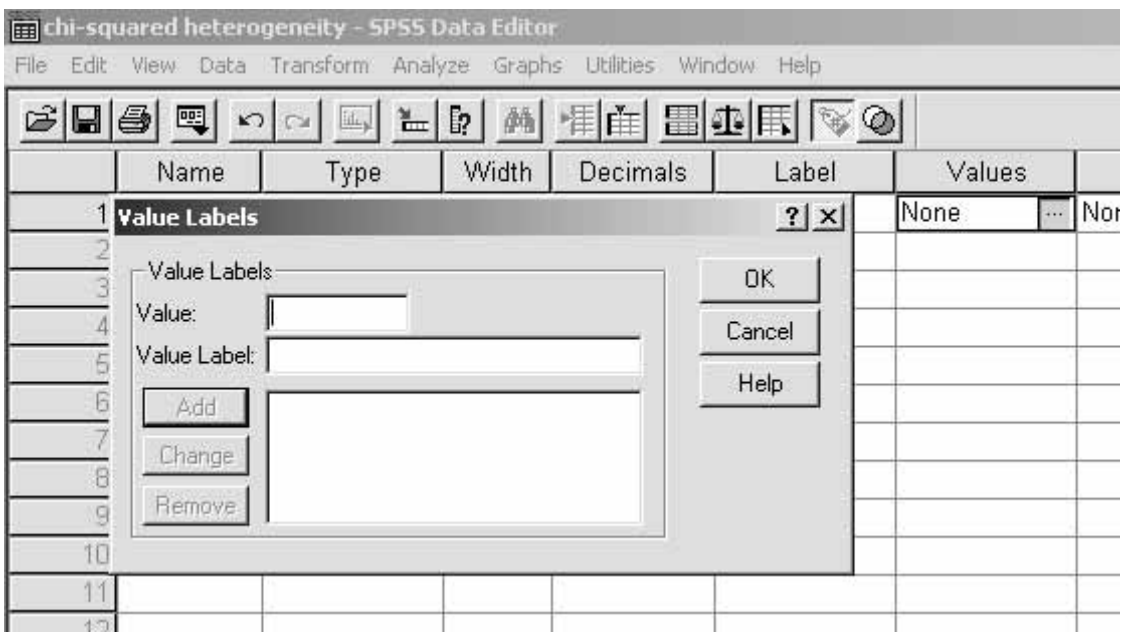
(iii) In the name for variable 1, type 'colour'. Press 'return', and the characteristics of the variable will be given default values.



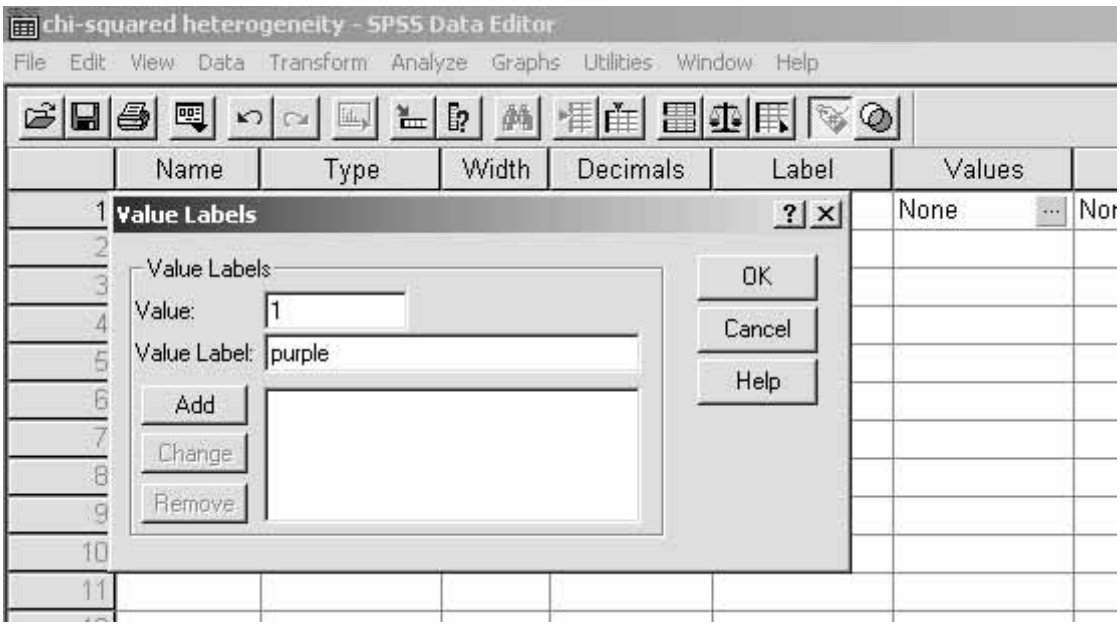
The two possible colours are purple and white, so we need to enter these in the 'values'. Click in the 'Values' cell for row 1: a grey box should appear at the right-hand edge.



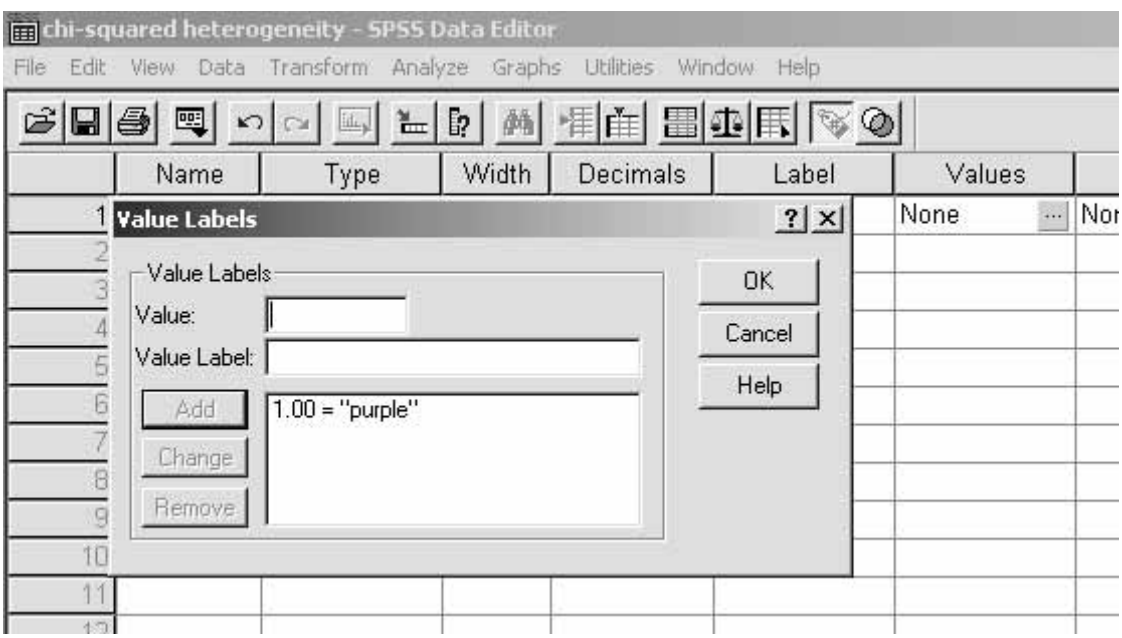
Click on the grey area, and a dialogue box should appear.



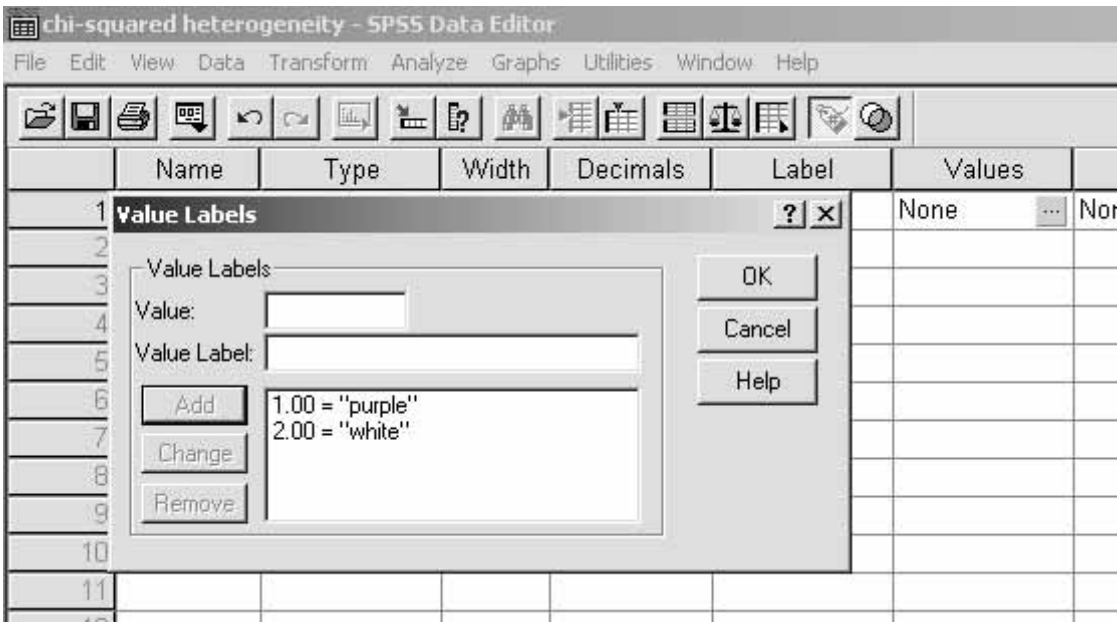
In the 'Value' box, enter '1', and in the 'Value Label' box, enter 'purple'.



Click on 'Add'.

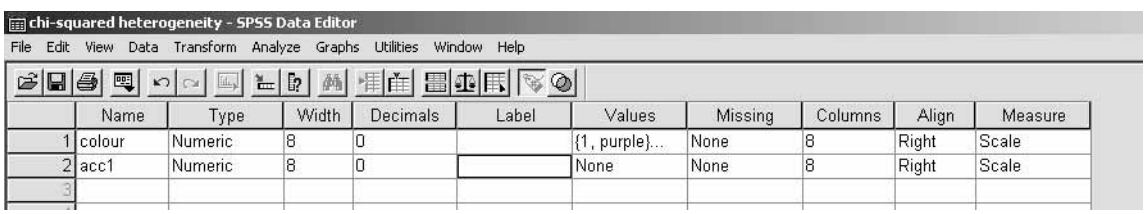


Repeat for '2' and 'white'.



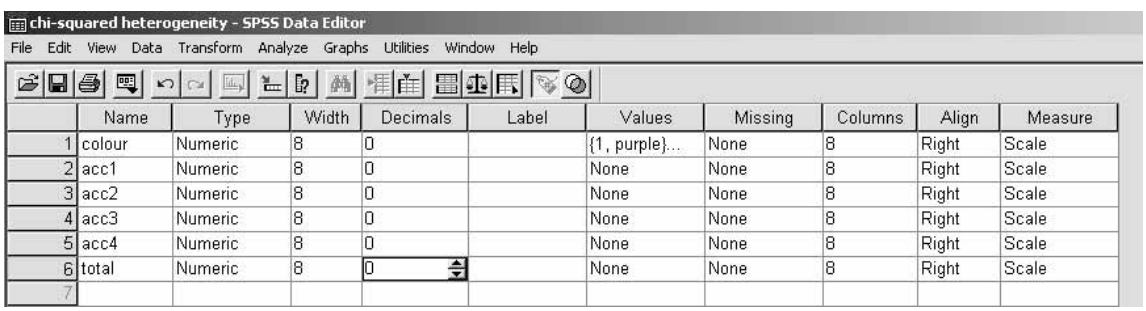
When all the possible colours have been entered, click on 'OK' to return to the variable view. Clicking in the 'Decimals' cell in row 1, and changing the value to zero will change the values to integers.

SPSS limits variable names to eight characters (including spaces), so 'accession 1' is far too long. In the name for row 2, type something shorter, such as 'acc 1'. This will contain the number of flowers of that colour, so set the 'Decimals' to zero again.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	colour	Numeric	8	0		{1, purple}...	None	8	Right	Scale
2	acc1	Numeric	8	0		None	None	8	Right	Scale
3										

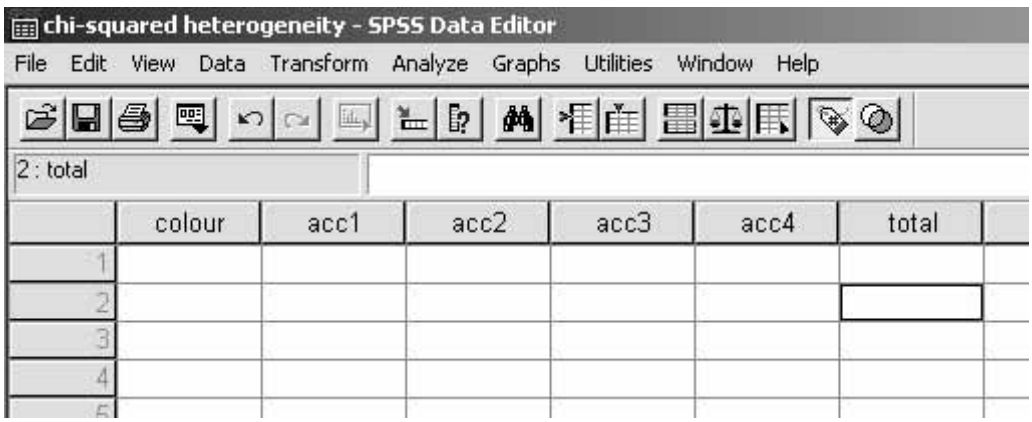
Repeat to set up names for the other three accessions and for the total.



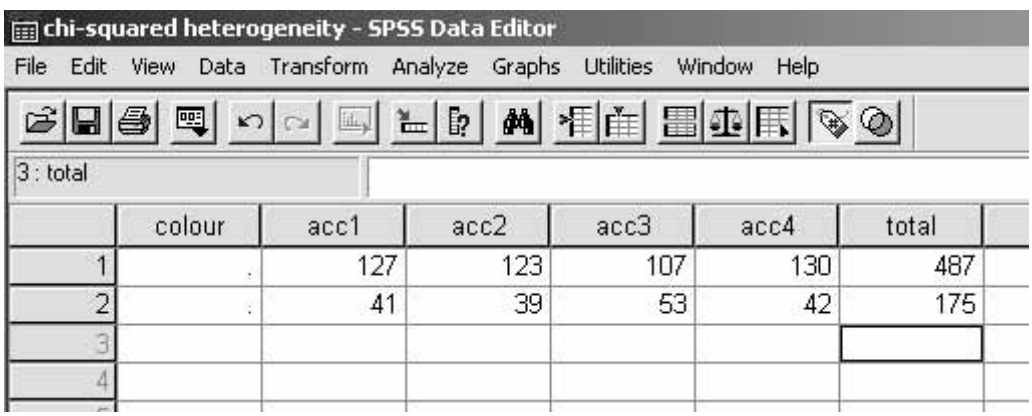
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	colour	Numeric	8	0		{1, purple}...	None	8	Right	Scale
2	acc1	Numeric	8	0		None	None	8	Right	Scale
3	acc2	Numeric	8	0		None	None	8	Right	Scale
4	acc3	Numeric	8	0		None	None	8	Right	Scale
5	acc4	Numeric	8	0		None	None	8	Right	Scale
6	total	Numeric	8	0		None	None	8	Right	Scale
7										

Step 2. Enter the data.

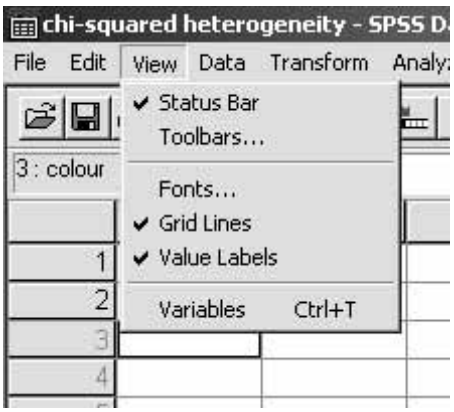
Go to 'Data View' (click on the appropriate tab in the bottom left-hand corner of the screen. You should see something like this:



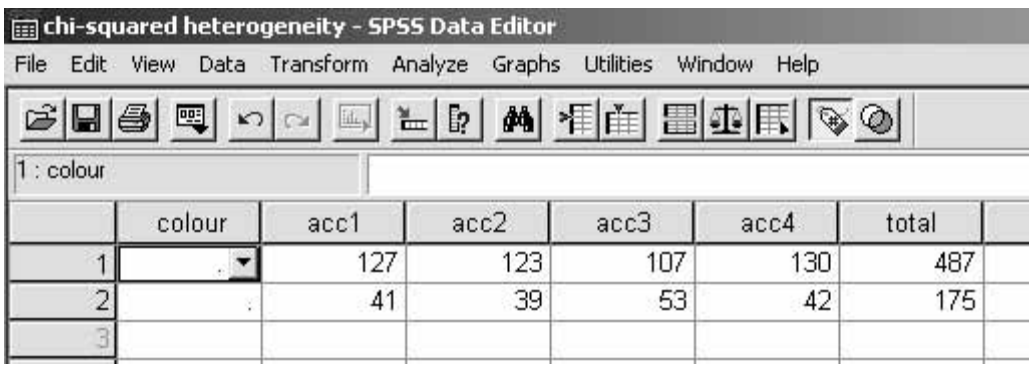
We need to enter some data into the accession columns before we can display the colours. Enter the appropriate numbers, remembering that row 1 is purple, and row 2 is white.



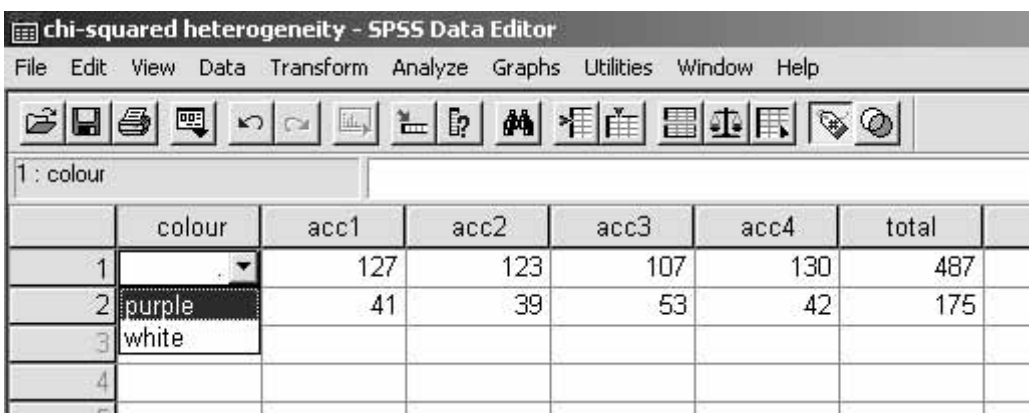
Now we can enter the colours. Check that value labels are enabled by going to 'View' and ensuring that 'Value Labels' is selected.'



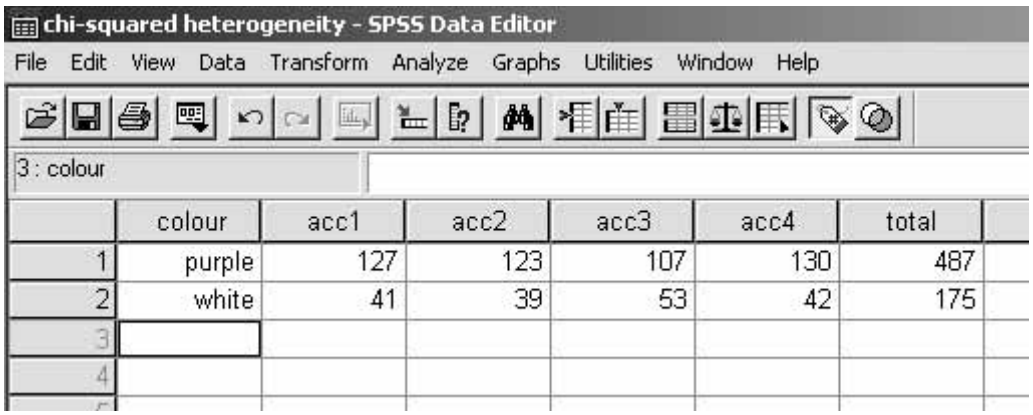
Click in the 'colour' cell of row 1. A drop-down menu should be accessible via an arrow at the right-hand side.



Click on the down arrow, and select 'purple'.

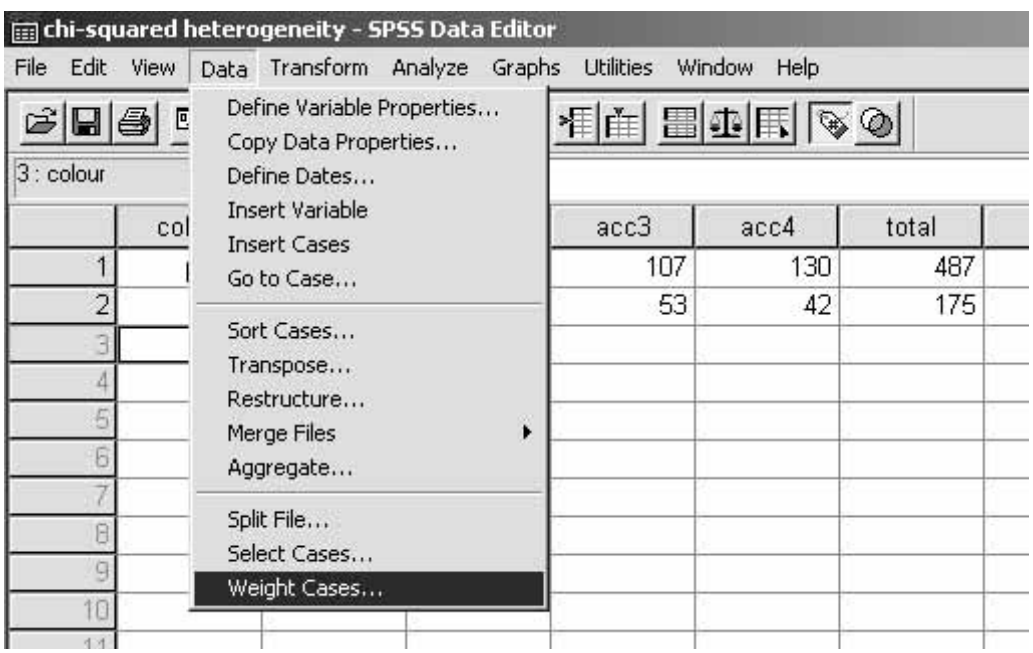


Repeat the procedure for the colour cell in row 2, selecting 'white' in this case.

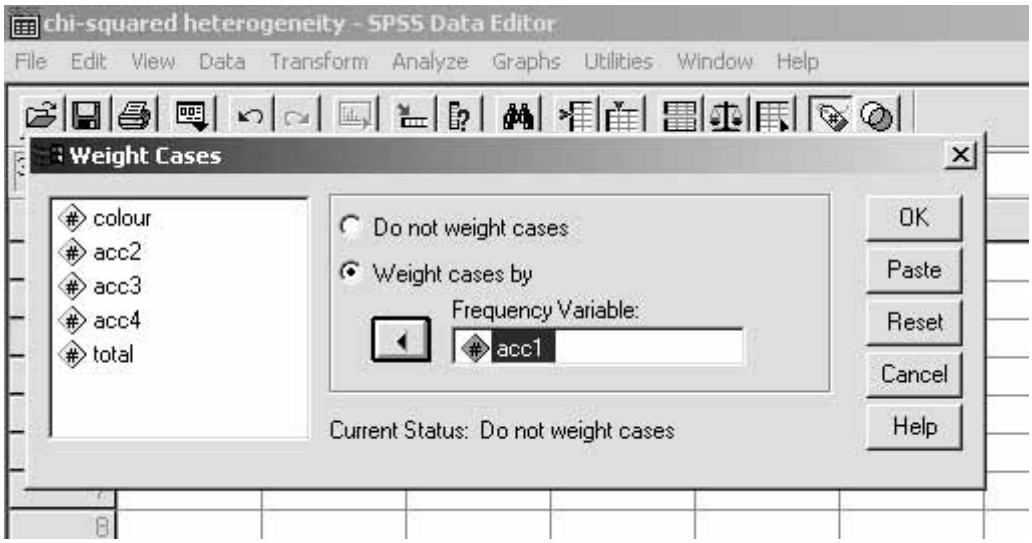


Step 3. Perform the test.
We will have to perform the chi-squared tests separately on each accession, and on the 'total'.

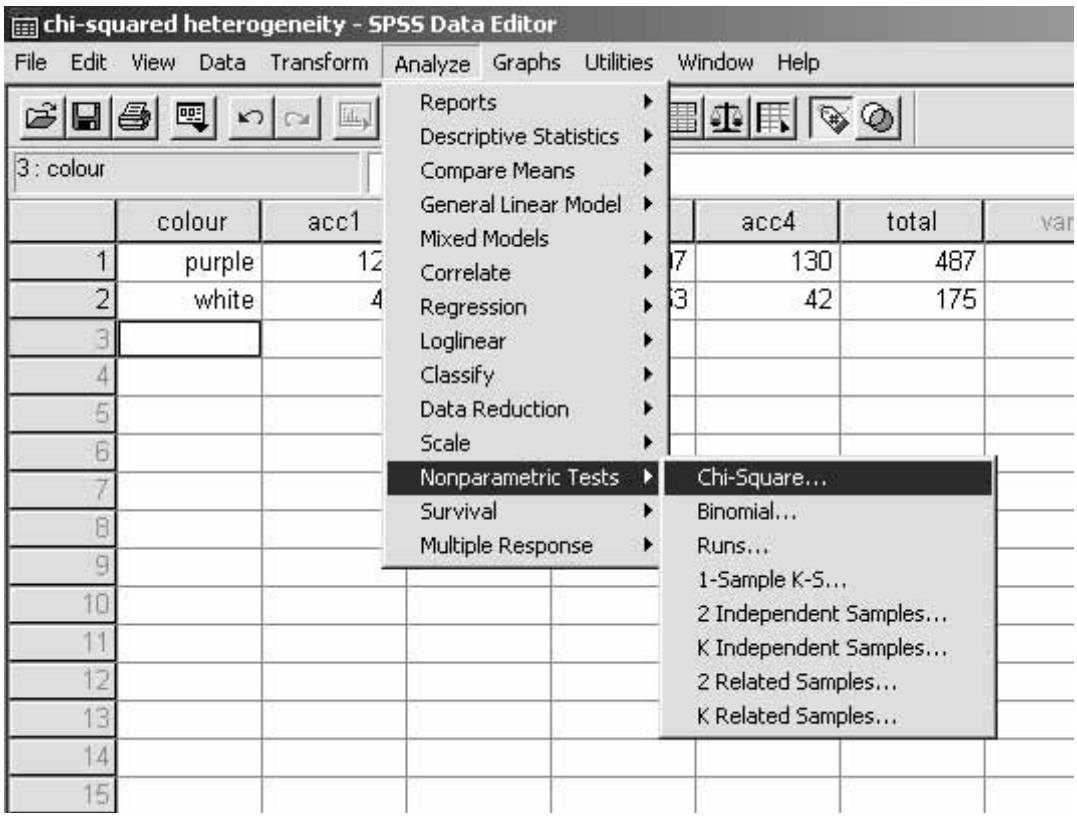
The first step is to weight the data for the set we are using. Go to 'Data', 'Weight Cases'.



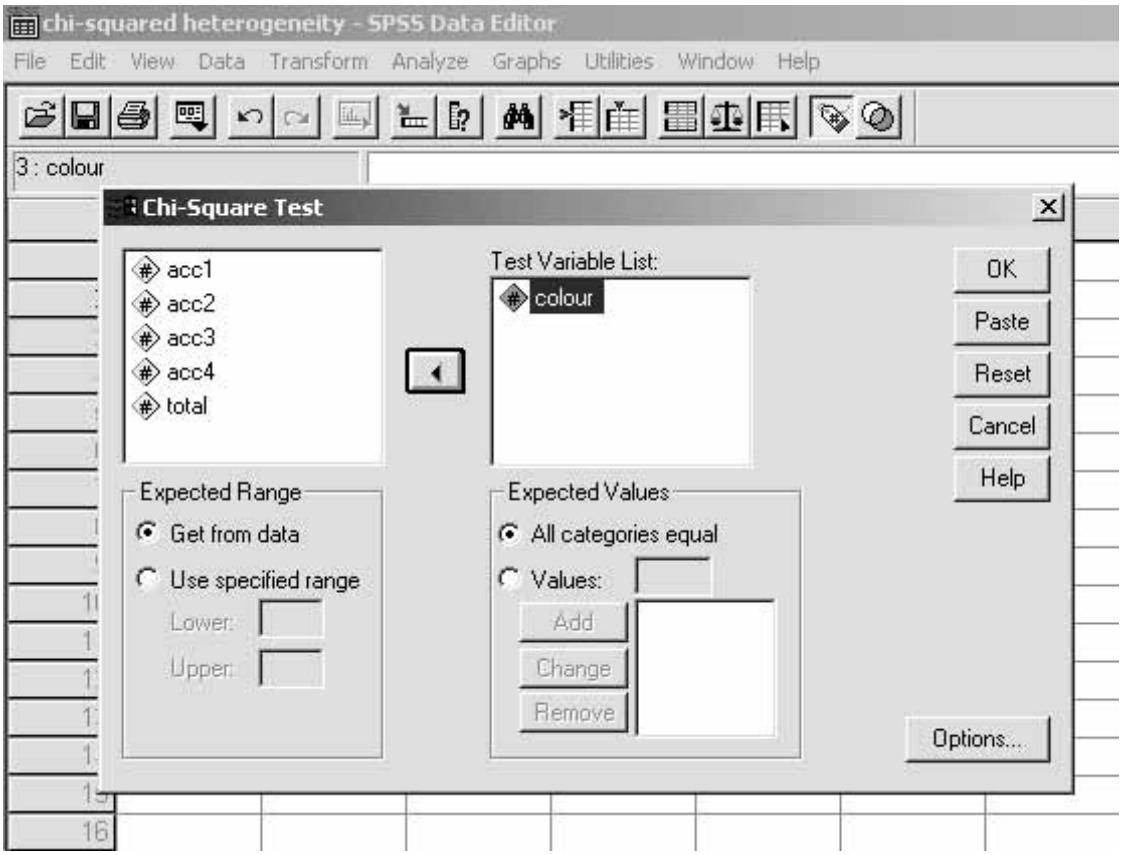
Click on the 'Weight cases by' radio button, click on 'acc1' to highlight it, and then transfer it into the 'frequency variable' window by clicking on the arrow.



Click on 'OK'.
Go to 'Analyze', 'Nonparametric tests', 'Chi-square'.

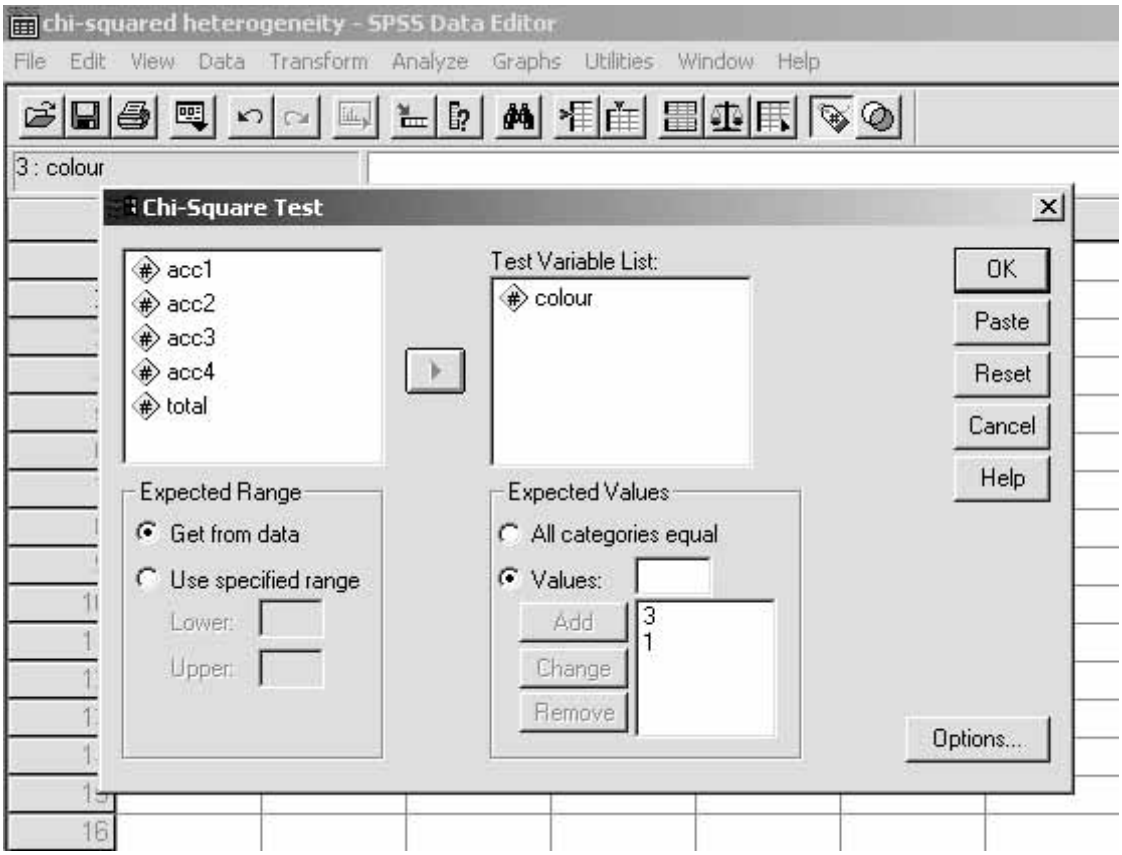


Our test variable is 'colour', so click on this to highlight it, and then click on the arrow to transfer it to the 'Test Variable List' box.



Unfortunately, we have to enter the expected values by hand. Fortunately, SPSS doesn't need exact values – it can use ratios. In this case, we expect three purple flowers for every one white flower.

In the 'Expected Values' section, click on the 'Values' radio button. Type '3' in the little box, and click on 'Add'. Repeat for '1', the expected relative number of white flowers. It is important that these are entered in the correct order – the same as the order of the colours in the 'colours' column.



Click on 'OK'. The results will appear in a separate window.

Chi-Square Test

Frequencies

COLOUR

	Observed N	Expected N	Residual
purple	127	126.0	1.0
white	41	42.0	–1.0
Total	168		

Test Statistics

	COLOUR
Chi-Square [a]	.032
di	1
Asymp.Sig.	.859

a. 0 cells (0%) have epected frequencies less than 5. The minimum expected cell frequency is 42.0.

Step 4. What do the results mean?
This tells us that the value of chi-squared for accession 1 is 0.032, and that the observed numbers of flowers are not significantly different from the expected values at $p = 0.05$.

This process can be repeated for the other three accessions and for the 'total' data, by changing the weighting. From here, the process is arithmetic (see BOX 5.3. in the book) to calculate the chi-squared summed, chi-squared deviation, and chi-squared homogeneity, and finally the use of chi-squared tables (to find the critical value of chi-squared).