

Collapsing across response categories

Steps in transformation: Collapsing four response categories into three

The frequency distribution for the original variable is illustrated below.

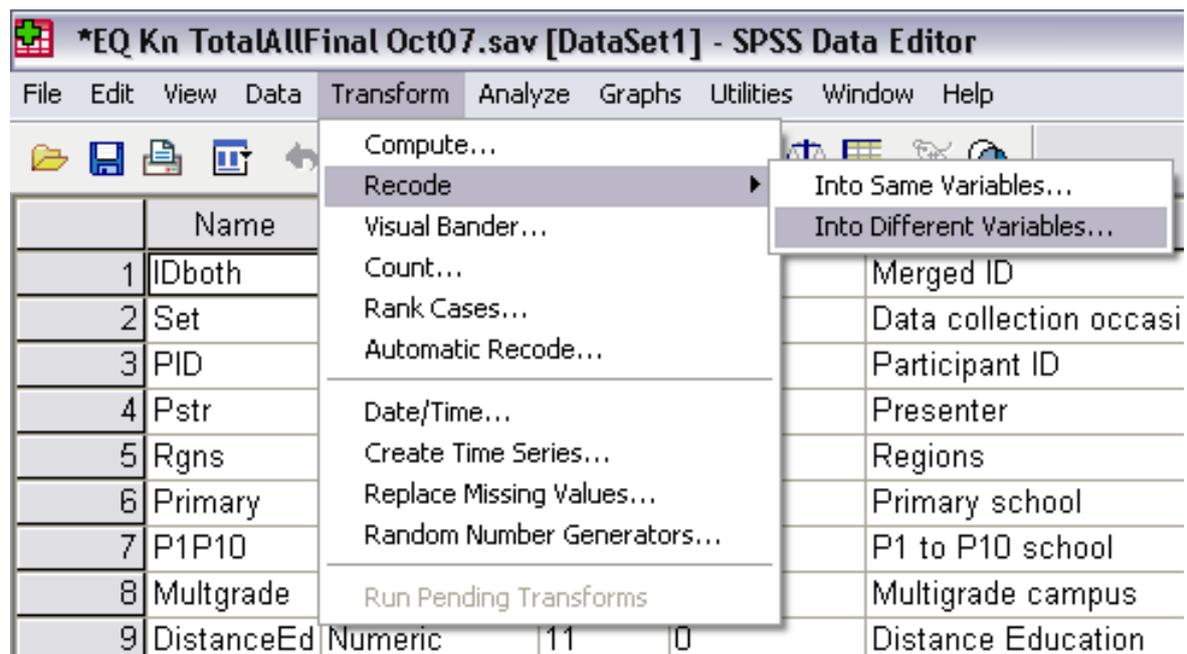
Statistics

Number of teachers (gpd)

N	Valid	Missing
	3914	191

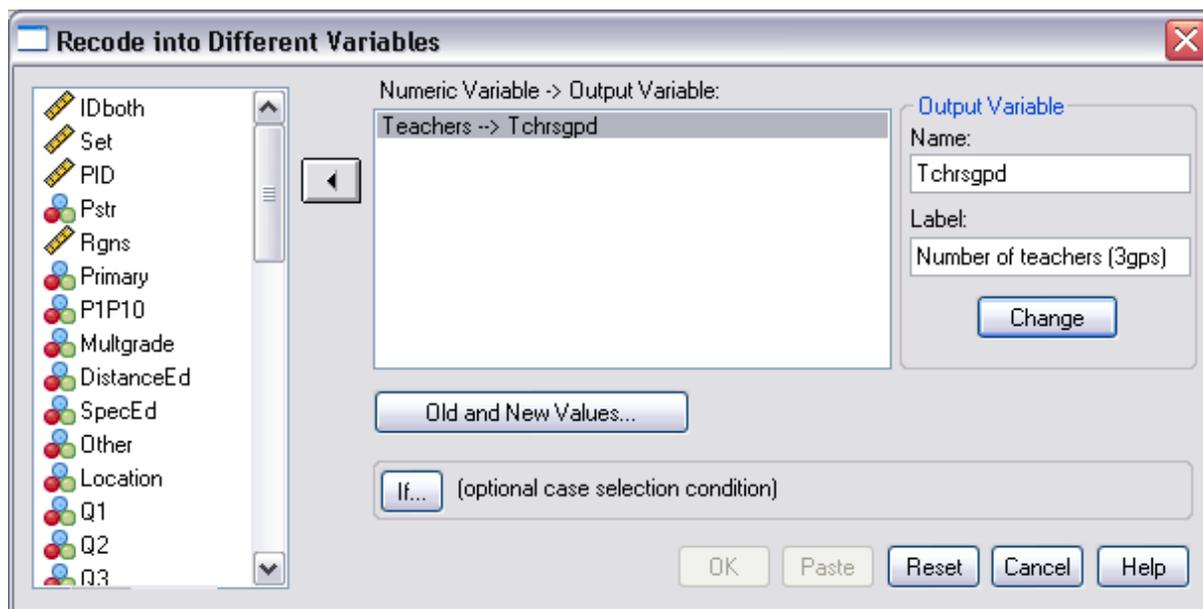
Number of teachers (gpd)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<5 tchrs	553	13.5	14.1	14.1
	5-10 tchrs	423	10.3	10.8	24.9
	11-20 tchrs	932	22.7	23.8	48.7
	>20 tchrs	2006	48.9	51.3	100.0
	Total	3914	95.3	100.0	
Missing	999	191	4.7		
Total		4105	100.0		

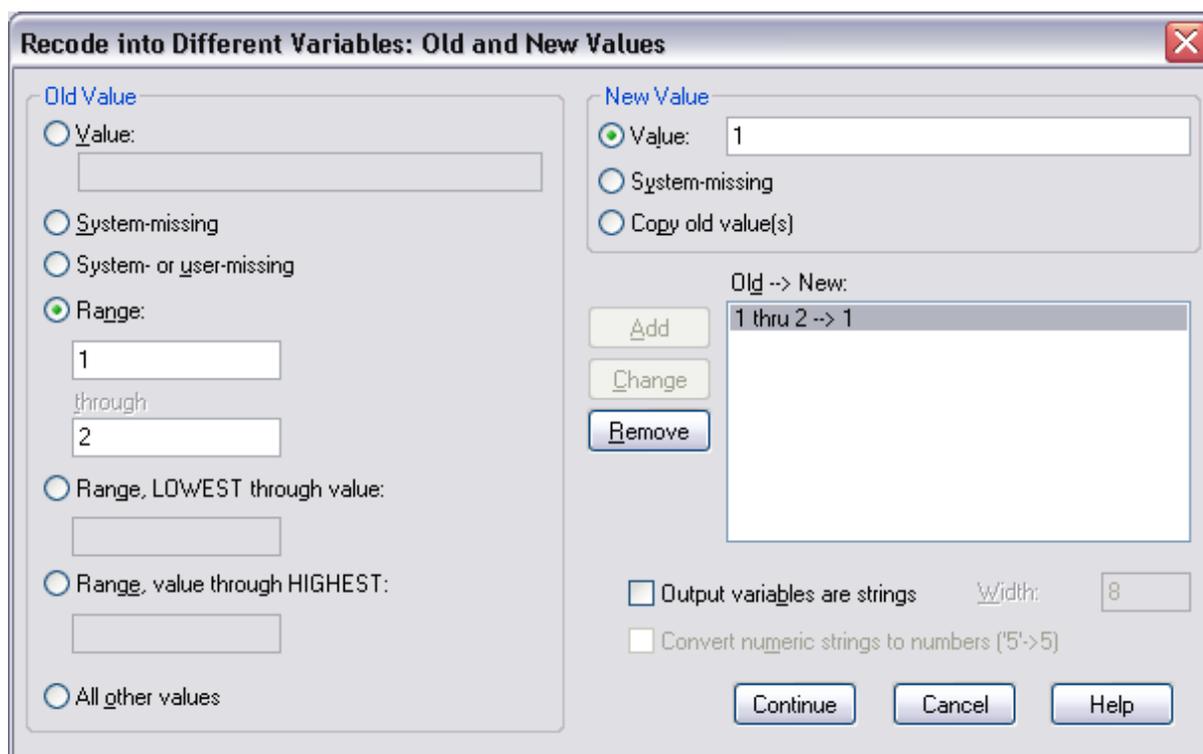


This variable includes four steps each representing a range of teachers working at the school, the first two of which could be collapsed to produce a three-step variable (<10, 11-20,>20 teachers). It is safer to select the choice of recoding into different variables. Changes to same variable become irreversible.

Collapsing across response categories

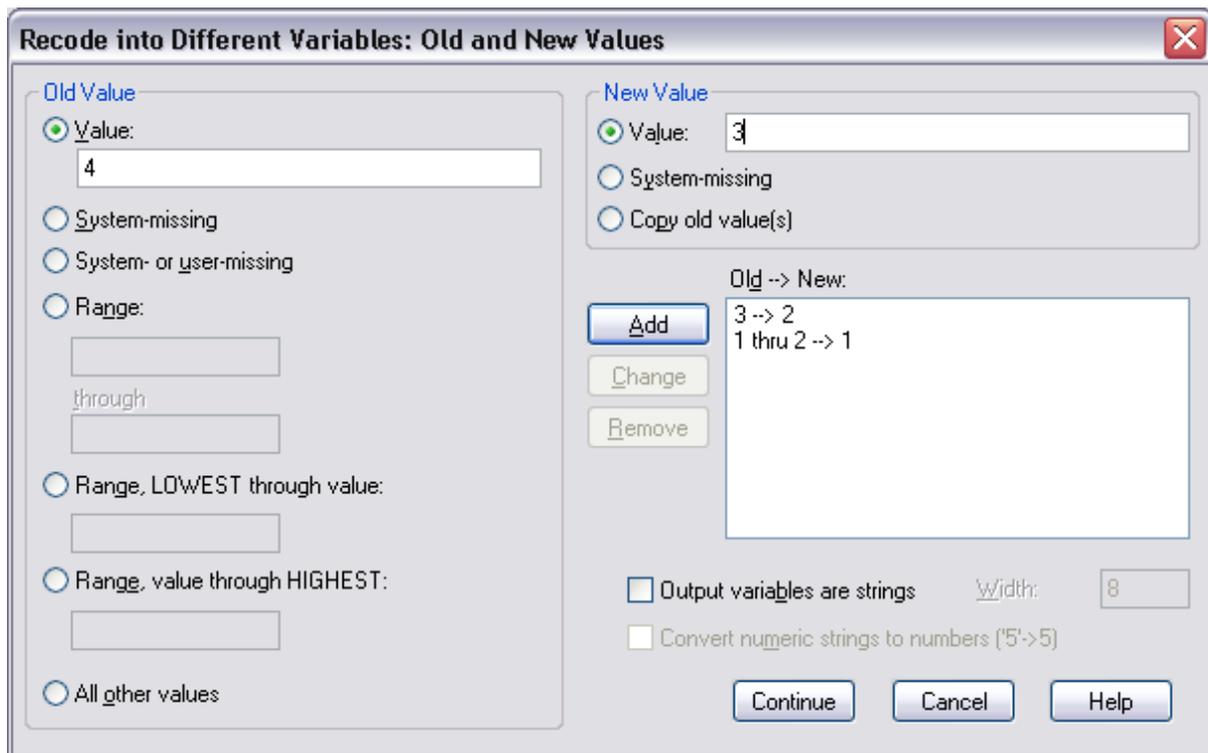


Select variable to be changed from list to left, type new name and label (not essential) into Output variable boxes to right. Click change to obtain result above.



Clicking on Old and New Values opens dialogue box above. Here, the first two steps (1 (<5) to 2 (5-10) are being consolidated on left and given New Value on right (at top). Clicking Add produces the transformation listed in the Old->New window in lower right of dialogue box.

Collapsing across response categories



The dialog box is titled "Recode into Different Variables: Old and New Values". It is divided into two main sections: "Old Value" and "New Value".

Old Value:

- Value: 4
- System-missing
- System- or user-missing
- Range: [] through []
- Range, LOWEST through value: []
- Range, value through HIGHEST: []
- All other values

New Value:

- Value: 3
- System-missing
- Copy old value(s)

Old -> New:

- 3 -> 2
- 1 thru 2 -> 1

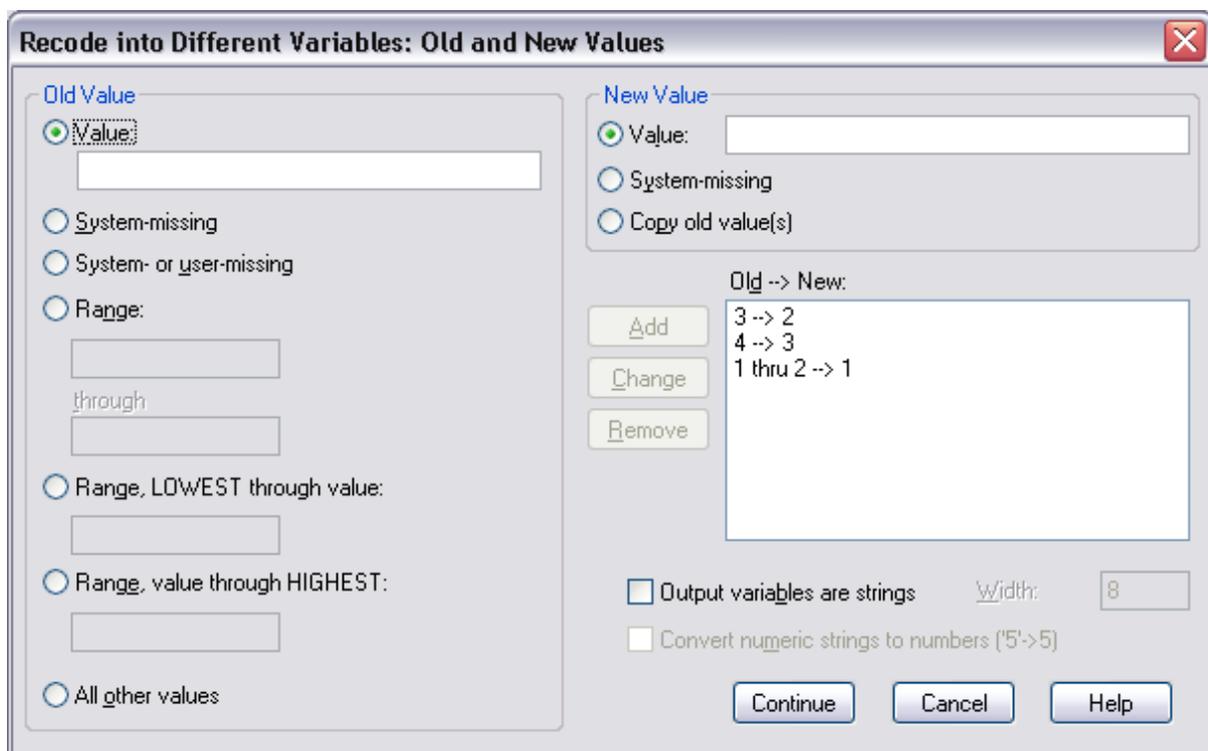
Buttons: Add, Change, Remove

Options:

- Output variables are strings Width: 8
- Convert numeric strings to numbers ('5'>5)

Buttons: Continue, Cancel, Help

Here, 4 (>20) is being recoded to 3 (>20). Note that the new values range from 1-3. The Add button has not yet been clicked to confirm this transformation.



The dialog box is titled "Recode into Different Variables: Old and New Values". It is divided into two main sections: "Old Value" and "New Value".

Old Value:

- Value: []
- System-missing
- System- or user-missing
- Range: [] through []
- Range, LOWEST through value: []
- Range, value through HIGHEST: []
- All other values

New Value:

- Value: []
- System-missing
- Copy old value(s)

Old -> New:

- 3 -> 2
- 4 -> 3
- 1 thru 2 -> 1

Buttons: Add, Change, Remove

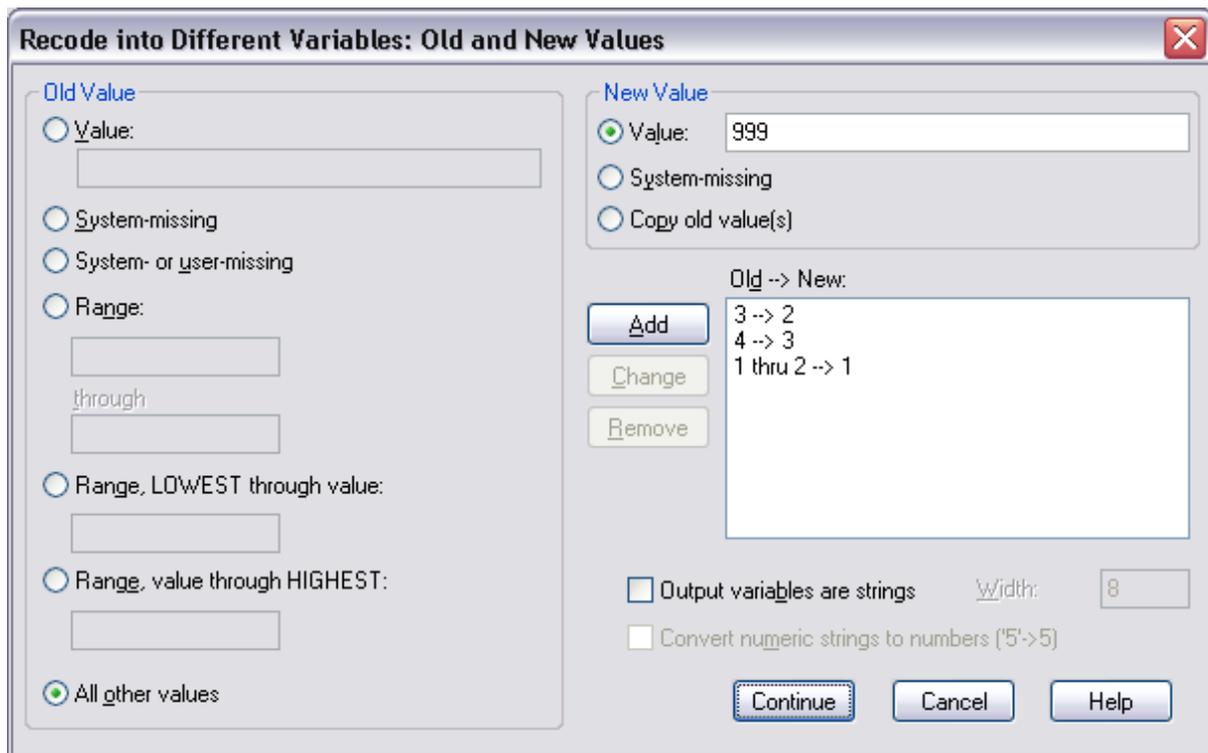
Options:

- Output variables are strings Width: 8
- Convert numeric strings to numbers ('5'>5)

Buttons: Continue, Cancel, Help

This illustrates the same transformation after clicking on Add button.

Collapsing across response categories



The dialog box is titled "Recode into Different Variables: Old and New Values". It is divided into two main sections: "Old Value" and "New Value".

Old Value section:

- Value: []
- System-missing
- System- or user-missing
- Range: [] through []
- Range, LOWEST through value: []
- Range, value through HIGHEST: []
- All other values

New Value section:

- Value: 999
- System-missing
- Copy old value(s)

Old -> New list:

- 3 -> 2
- 4 -> 3
- 1 thru 2 -> 1

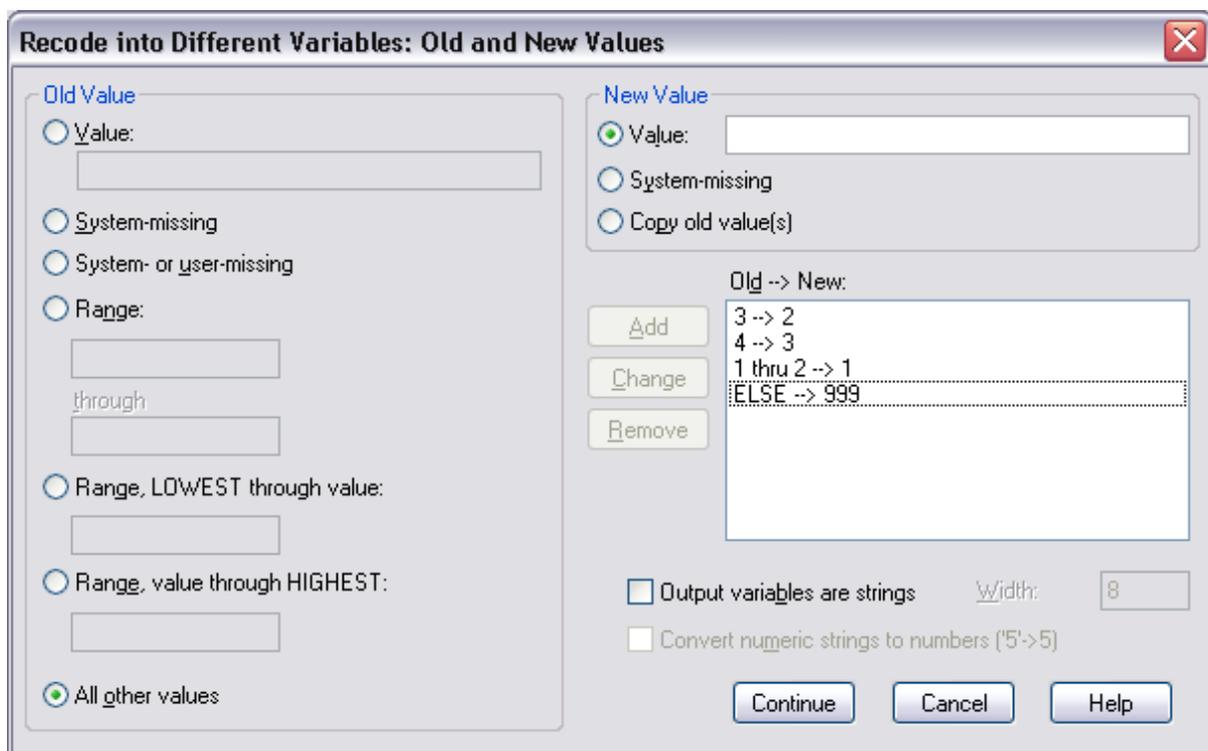
Buttons: Add, Change, Remove

Options:

- Output variables are strings Width: 8
- Convert numeric strings to numbers ('5'>5)

Buttons: Continue, Cancel, Help

Usually a good idea to code missing values with a specific Missing Values code (e.g., 999) Add button not yet clicked.



The dialog box is titled "Recode into Different Variables: Old and New Values". It is divided into two main sections: "Old Value" and "New Value".

Old Value section:

- Value: []
- System-missing
- System- or user-missing
- Range: [] through []
- Range, LOWEST through value: []
- Range, value through HIGHEST: []
- All other values

New Value section:

- Value: []
- System-missing
- Copy old value(s)

Old -> New list:

- 3 -> 2
- 4 -> 3
- 1 thru 2 -> 1
- ELSE -> 999

Buttons: Add, Change, Remove

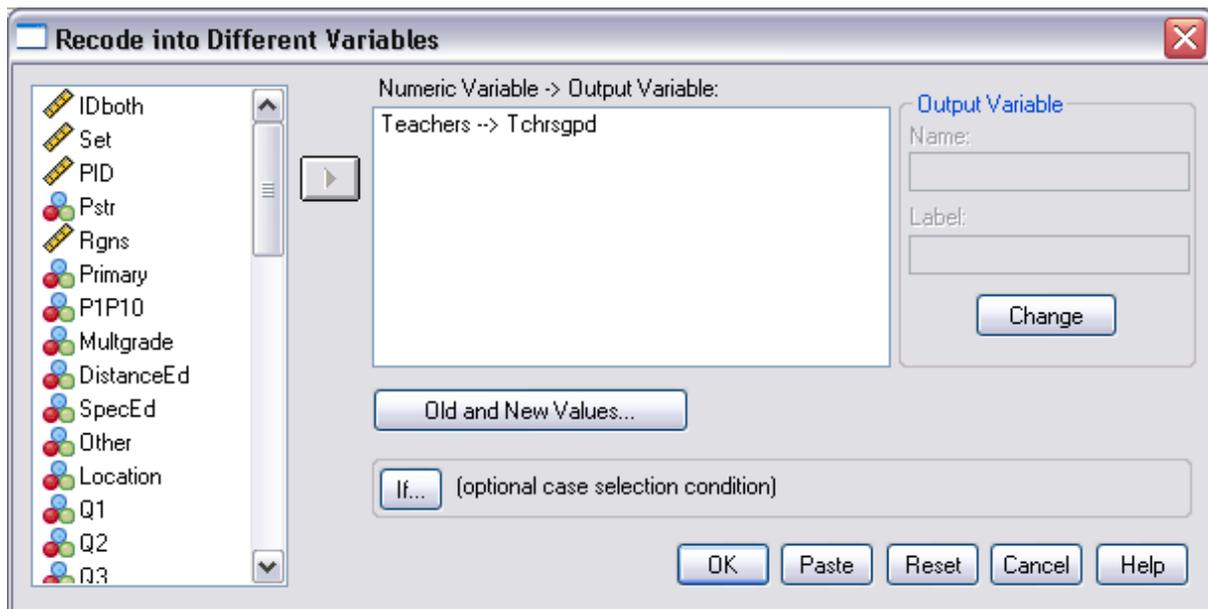
Options:

- Output variables are strings Width: 8
- Convert numeric strings to numbers ('5'>5)

Buttons: Continue, Cancel, Help

Add button clicked for the above.

Collapsing across response categories



Clicking Continue in Old and New Values dialogue box returns one to the initial Recode window. Now it's time to click OK.

38	QP11	Numeric	11	0	Can support spelling	None	999	8	Right	Nominal
39	QP12	Numeric	11	0	Respond in timely manner	None	999	8	Right	Nominal
40	QP13	Numeric	11	0	Support learning needs	None	999	8	Right	Nominal
41	QP14	Numeric	11	0	Bld on home practice	None	999	8	Right	Nominal
42	Tchrsgrp	Numeric	8	2	Number of teachers (3gps)	None	None	10	Right	Scale
43										
44										
45										

Clicking OK produces default version of new variable (final row).

37	QP10	Numeric	11	0	Can support writing	None	999	0	Right	Nominal
38	QP11	Numeric	11	0	Can support spelling	None	999	8	Right	Nominal
39	QP12	Numeric	11	0	Respond in timely manner	None	999	8	Right	Nominal
40	QP13	Numeric	11	0	Support learning needs	None	999	8	Right	Nominal
41	QP14	Numeric	11	0	Bld on home practice	None	999	8	Right	Nominal
42	Tchrsgrp	Numeric	8	0	Number of teachers (3gps)	{1, <=10 tchrs}	999	10	Right	Scale
43										
44										
45										

This can be tidied up by adjusting the number of decimals to 0 and by providing labels, values, and missing values coding.

Collapsing across response categories

The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and 'Frequencies...' is selected under 'Descriptive Statistics'. The data view shows a list of variables with their names and types. The 'Teachers' variable is highlighted in the list.

ID	Name	Type
1	IDboth	Numeric
2	Set	Numeric
3	PID	Numeric
4	Pstr	Numeric
5	Rgns	Numeric
6	Primary	Numeric
7	P1P10	Numeric
8	Multgrade	Numeric
9	DistanceEd	Numeric
10	SpecEd	Numeric
11	Other	Numeric
12	Location	Numeric
13	Teachers	Numeric
14	Q1	Numeric
15	Q2	Numeric
16	Q3	Numeric

Now the work should be checked by using frequencies to compare the original versus the recoded variable.

The screenshot shows the 'Frequencies' dialog box. The 'Variable(s):' list contains 'Teachers' and 'Tchrsgpd'. The 'Display frequency tables' checkbox is checked. The 'Statistics...', 'Charts...', and 'Format...' buttons are visible at the bottom.

From the list to the left, add the original variable (Teachers) and the collapsed variable (Tchrsgpd) into the Variable/s window. Default settings are fine. Click OK.

Collapsing across response categories

Frequencies

Statistics			
		Number of teachers (gpd)	Number of teachers (3gps)
N	Valid	3914	3914
	Missing	191	191

Frequency Table

Number of teachers (gpd)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<5 tchrs	553	13.5	14.1	14.1
	5-10 tchrs	423	10.3	10.8	24.9
	11-20 tchrs	932	22.7	23.8	48.7
	>20 tchrs	2006	48.9	51.3	100.0
	Total	3914	95.3	100.0	
Missing	999	191	4.7		
Total		4105	100.0		

Number of teachers (3gps)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<=10 tchrs	976	23.8	24.9	24.9
	11-20 tchrs	932	22.7	23.8	48.7
	>20 tchrs	2006	48.9	51.3	100.0
	Total	3914	95.3	100.0	
Missing	999	191	4.7		
Total		4105	100.0		

Fortunately the outcomes seem satisfactory.