

Answers to Exercise 3.1 (p. 72)

To create the new variable 'HINCOME2' use the **Recode into Different Variables...** procedure as illustrated in Figure 3.1 (p. 70). However, to work out how to recode the new variable it is useful to see how the original variable has been coded first. To do this use the **Frequencies...** procedure for 'HINCOME' as illustrated in Figure 1.11 (p. 23). This should give you the output below:

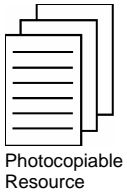
PW19-TOTAL HH INCOME RANGE 2

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 \$5,000 OR LESS	334	2.9	2.9	2.9
2 \$5,001 - \$10K	433	3.7	3.7	6.6
3 \$10,001 - \$15K	589	5.0	5.0	11.6
4 \$15,001 - \$20K	577	4.9	4.9	16.5
5 \$20,001 - \$25K	779	6.7	6.7	23.2
6 \$25,001 - \$30K	612	5.2	5.2	28.4
7 \$30,001 - \$35K	642	5.5	5.5	33.9
8 \$35,001 - \$40K	594	5.1	5.1	39.0
9 \$40,001 - \$45K	426	3.6	3.6	42.7
10 \$45,001 - \$50K	587	5.0	5.0	47.7
11 \$50,001 - \$60K	1033	8.8	8.8	56.5
12 \$60,001 - \$75K	1442	12.3	12.3	68.9
13 \$75,001 - \$100K	1538	13.2	13.2	82.0
14 OVER \$100K	2098	18.0	18.0	100.0
Total	11684	100.0	100.0	

To create the new variable 'HINCOME2' therefore, we need to create three new categories as follows:

- All those currently coded 1 to 6 should become the first category (\$0 - \$30,000)
- All those coded 7 to 11 should become the second category (\$30,001 - \$60,000)
- All those coded 12 to 14 should become the third category (\$60,001 and over)

The actual old and new values should therefore be entered as shown overleaf:

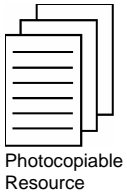


Photocopiable Resource

Once you have run this procedure, do not forget to give the new variable a full title and to label the three new categories (see Figure 1.10 on p. 22 to remind yourself how to do this). As a check to see that you have re-coded the variable properly, run the simple **Crosstabs...** procedure (as illustrated by Figure 3.5 on p. 78), putting 'HINCOME' in rows and 'HINCOME2' in columns. This should give you the table below:

PW19-TOTAL HH INCOME RANGE 2 * TOTAL HH INCOME RANGE Crosstabulation

Count		TOTAL HH INCOME RANGE			Total
		\$0 - \$30,000	\$30,001 - \$60,000	\$60,001 and over	
PW19-TOTAL	1 \$5,000 OR LESS	334	0	0	334
HH INCOME RANGE 2	2 \$5,001 - \$10K	433	0	0	433
	3 \$10,001 - \$15K	589	0	0	589
	4 \$15,001 - \$20K	577	0	0	577
	5 \$20,001 - \$25K	779	0	0	779
	6 \$25,001 - \$30K	612	0	0	612
	7 \$30,001 - \$35K	0	642	0	642
	8 \$35,001 - \$40K	0	594	0	594
	9 \$40,001 - \$45K	0	426	0	426
	10 \$45,001 - \$50K	0	587	0	587
	11 \$50,001 - \$60K	0	1033	0	1033
	12 \$60,001 - \$75K	0	0	1442	1442
	13 \$75,001 - \$100K	0	0	1538	1538
	14 OVER \$100K	0	0	2098	2098
	Total		3324	3282	5078



As can be seen, all of the values seem to have been recoded properly. The final thing to do, therefore, is to run the simple **Frequencies...** procedure again but this time for the new variable 'HINCOME2' to calculate the proportions of the sample falling into each of the three categories.

Remember to apply the weighting variable 'WTCORRCT' before you run the analysis (see Figure 2.5 on p. 52 for how to do this). The results you should get are shown below:

TOTAL HH INCOME RANGE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	\$0 - \$30,000	1819	30.5	30.5	30.5
	\$30,001 - \$60,000	1696	28.5	28.5	59.0
	\$60,001 and over	2446	41.0	41.0	100.0
	Total	5961	100.0	100.0	

Finally, don't forget to save this amended dataset with the new variable 'HINCOME2' in it.