

Answers to Exercise 6.1 (p. 189)

In relation to the **timeseries.sav** dataset you need to begin by creating two additional variables representing the percentages of boys and girls who did not achieve five or more GCSE passes for each of the years listed. How to do this was explained in Chapter 3 (pp. 108-9) in relation to the creation of the two variables 'boys_no' and 'girls_no' respectively. Once you have created these two variables you have the four figures needed to calculate the effect size, Phi, for each year.

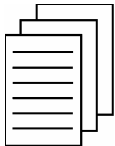
To calculate the effect sizes, you need to begin by creating the special dataset exactly as described in the book on pages 185 – 189. There are two things to note about the figures you are going to use from the **timeseries.sav** dataset:

- They cannot be used exactly as they stand because they are stated to one decimal place. Unfortunately, this procedure only works with whole numbers. As such, you need to multiply all four figures for each year by 10 before entering them into the specially created dataset (you'll see an example of this shortly).
- As we are dealing with percentage figures it needs to be remembered that while the effect size, Phi, will be reliable the Chi-Square statistic and associated 'p value' calculated will not. As explained in the book, to calculate an accurate Chi-Square statistic using this method you need to use actual frequency counts (i.e. the actual numbers of pupils in each of the four categories) rather than just percentages.

If we begin with the first year in the dataset (1974/75) then we need to multiple the four percentage figures by 10 and add them into the specially created dataset so that it looks like this:

	gender	attain	count	var
1	Male	5 or More Passes	222	
2	Male	Less than 5 Passes	778	
3	Female	5 or More Passes	230	
4	Female	Less than 5 Passes	770	
5				
6				

Once you have this, use the **Data → Weight Cases...** procedure to weight cases by the variable 'count'. Once done, use the **Analyze → Descriptive Statistics →**



Photocopiable Resource

Crosstabs... procedure to put 'gender' in rows and 'attain' in columns. Click on the 'Statistics' button to select 'Phi and Cramer's V'. Click 'Continue' and then 'OK' to complete the procedure. The output you should get is as below:

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
gender * attain	2000	100.0%	0	.0%	2000	100.0%

gender * attain Crosstabulation

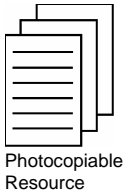
Count		attain		Total
		5 or More Passes	Less than 5 Passes	
gender	Male	222	778	1000
	Female	230	770	1000
Total		452	1548	2000

Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	-.010	.669
Nominal	Cramer's V	.010	.669
N of Valid Cases		2000	

- a. Not assuming the null hypothesis.
- b. Using the asymptotic standard error assuming the null hypothesis.

As can be seen, the effect size (Phi) for this first year is 0.010 (we can ignore the minus sign). If you now do this for all of the other years you should get the figures for Phi that can then be entered into the **timeseries.sav** dataset as an additional variable as shown overleaf.



Photocopiable Resource

*timeseries[1].sav [DataSet5] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

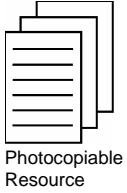
36 : boys_no Visible: 8 of 8 Variables

	year	girls	boys	total	girls no	boys no	cp ratio	phi
1	1974/75	23.0	22.2	22.6	77.0	77.8	1.047	.010
2	1975/76	23.1	22.7	22.9	76.9	77.3	1.023	.005
3	1976/77	23.5	23.4	23.5	76.5	76.6	1.006	.001
4	1977/78	23.6	23.7	23.7	76.4	76.3	.994	.001
5	1978/79	23.9	23.5	23.7	76.1	76.5	1.022	.005
6	1979/80	24.4	23.7	24.1	75.6	76.3	1.039	.008
7	1980/81	25.6	24.5	25.1	74.4	75.5	1.060	.013
8	1981/82	26.8	25.4	26.1	73.2	74.6	1.075	.016
9	1982/83	27.1	25.4	26.3	72.9	74.6	1.092	.019
10	1983/84	27.2	26.3	26.8	72.8	73.7	1.047	.010
11	1984/85	27.4	26.3	26.9	72.6	73.7	1.058	.012
12	1985/86	27.2	26.2	26.7	72.8	73.8	1.052	.011
13	1986/87	27.2	25.6	26.4	72.8	74.4	1.086	.018
14	1987/88	31.7	28.2	30.0	68.3	71.8	1.182	.038
15	1988/89	35.8	29.8	32.8	64.2	70.2	1.314	.064
16	1989/90	38.4	30.8	34.6	61.6	69.2	1.401	.080
17	1990/91	40.3	33.3	36.8	59.7	66.7	1.352	.073
18	1991/92	42.7	34.1	38.4	57.3	65.9	1.440	.088
19	1992/93	45.8	36.8	41.3	54.2	63.2	1.451	.091
20	1993/94	47.8	39.1	43.5	52.2	60.9	1.426	.088
21	1994/95	48.1	39.0	43.6	51.9	61.0	1.450	.092
22	1995/96	49.4	39.9	44.7	50.6	60.1	1.471	.096
23	1996/97	50.0	40.5	45.3	50.0	59.5	1.469	.095
24	1997/98	51.5	41.3	46.4	48.5	58.7	1.509	.102
25	1998/99	53.4	42.8	48.1	46.6	57.2	1.531	.106
26	1999/00	54.6	44.0	49.3	45.4	56.0	1.531	.106
27	2000/01	55.4	44.8	50.1	44.6	55.2	1.531	.106
28	2001/02	57.0	46.4	51.7	43.0	53.6	1.531	.106
29	2002/03	58.2	47.9	53.1	41.8	52.1	1.514	.103
30	2003/04	58.4	48.4	53.4	41.6	51.6	1.497	.100
31	2004/05	60.8	50.8	55.7	39.2	49.2	1.502	.101
32								

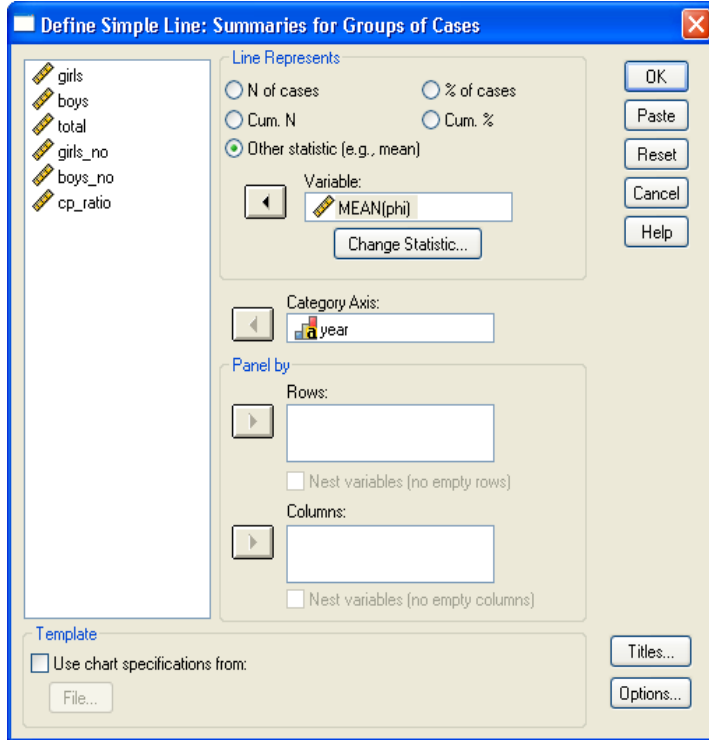
Data View Variable View

SPSS Processor is ready

Finally, you can use the **Graphs** → **Legacy Dialogs** → **Line...** procedure to create a simple line graph for the variable Phi as shown overleaf:



Photocopiable Resource



The resultant graph you should get is as below. In comparing it with Figure 3.21 it can be seen that the use of Phi as a measure of sex differences in attainment gives substantially the same picture as that found with the cross product ratio:

