

Centre for Longitudinal Studies, Institute of Education

Teaching students quantitative methods using resources from the British Birth Cohorts

Measuring signs of psychological distress or depression in teenagers and adults

The Malaise Inventory

MalaiseExercises.doc: The SPSS exercises detailed here use data from the 1970 British Cohort Study (BCS70). Students could replicate these exercises using data from the 1958 National Child Development Study (NCDS) for cross-cohort comparisons. After a short summary of the two data files, *b034malaise.sav* and *n042malaise.sav* the document provides a number of step-by-step exercises for students to work through and additional exercises to complete. The exercises use the drop down menus but the SPSS syntax to run the commands also included. The **RutterBehaviourExercises.doc** contains additional exercises. By replacing the Rutter behaviour score variables with the Malaise score variables these exercises could be easily adapted for students using either the *b034malaise.sav* or *n042malaise.sav* data files. The techniques covered in this document range from introductory descriptive statistics to multiple and logistic regression and include:

- a) Frequencies
- b) Means
- c) Cross-tabulations
- d) Constructing summary scores
- e) Recoding continuous into dichotomous variables
- f) Correlations
- g) Linear regression
- h) Logistic regression.

SPSS Data Files: BCS70 data 'b034malaise.sav', NCDS data 'n042malaise.sav'

BCS70 data

This longitudinal dataset includes information from the 1970 British Cohort Study (BCS70) for looking at the relationship between socio-economic circumstances in childhood and mental well-being in adulthood over time. Data from birth are provided by the cohort members' mother; data collected at age 5, age 10 and age 16 are provided by the mother (or father) and the cohort member (child) and data from collected at age 26, 30 and 34 are provided by the cohort member themselves. There are 18,732 cohort members (individuals) in the dataset, 9,740 are men and 8,984 are women. Eight babies did not have a sex recorded at their premature birth.

NCDS data

This longitudinal dataset includes information from the 1958 National Child Development Study (NCDS) for looking at the relationship between socioeconomic circumstances in childhood and mental well-being in adulthood over time. Information included in this file which was collected at birth is provided by the cohort members' mother; information collected at age 7, age 11 and age 16 is provided from the mother (or father) and at age 23, 33 and 42 the cohort member themselves provided the information. There are 18,560 cases in the dataset, 8,960 are women and 9,593 are men. Seven babies did not have a sex recorded at their premature birth.

Step-by-Step Exercises using SPSS

Exploring the data

1. How many cohort members were depressed in their thirties? Were more depressed at age 30 or age 34? (Hint: frequency of b30malg and b34malg variables).

From the Analyse drop-down menu select Descriptive Statistics and then Frequencies. Select b30malg b34malg and click on the \blacktriangleright button to move the two variables into the Variable(s) box. Click on OK.

💼 b034malaise	iii b034malaise - SPSS Data Editor								
File Edit View	Data Transform Analyz	e Graphs Utilities Wind	dow Help						
			<u> </u>						
6 : b16mal03	0								
	key	bsex	b16mal	b26mal	b26malg	b30mal	b		
1	10	female	3	not stated	not stated	2			
2	20	Frequencies	Variable(s):			not stated	nc		
3	30	Image: white the second se	▲ b30ma	lg OK Past	ore of 8+	4			
4	40	 ✤ b16mal ✤ b26mal 		Rese	et .	not stated	nc		
5	50	 ✤ b26malg ♠ b30mal 		Cano	d core of 0-	1			
6	60	 ✤ b34mal ♣ b1634in ♠ b2524in 	~	Hel	· ·	not stated	nc		
7	70	Display frequency tal	bles		ot stated	4			
8	80		Statistics Cha	rts Format	ot stated	3			
9	90	temale	not stated	not stated	not stated	0			
10	100	male	not stated	8.00	score of 8+	9			
4.4	440	<i>c</i> 1							

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

FREQUENCIES VARIABLES=b30malg b34malg ORDER= ANALYSIS.

2. Do men or women have higher mean malaise scores in adulthood? Is the mean score consistently higher for men or women at age 26, 30 and 34? (Hint: Means of b26mal b30mal b34mal by bsex)

From the Analyse drop-down menu select Compare Means and then Means.

🖬 b	034	nalaise [Dat	aSet1] - S	PSS Data	Editor	r						
File	Edit	View Da	ata	Transform	Analyze 🤇	Graphs	Utilities	W	indow Help				
		🖣 📴	•	۰ ک	Reports Descriptive Statistics			5 📎 🥝					
1:	кеу]1	Tables			•					
		keγ		bsex	Compare	e Mean	s	•	Means			lg 🛛	b34r
	1		10	femal	General	Linear	Model		One-Sampl	e T Test		D-7	
	2		20	femal	Mixed M	odels				nt-Samples T T	est	ed	not st
	3		30	mal	Correlat	Correlate			Paired-Sam	Samples T Test D-7			
	4		4N	femal	Regress	Regression			One-Way /	ANOVA		ed	not st
	5		50	mal	Classifu	r		ľ	score of 0-	1		0-7	
	6		60	mal	Data Re	ductior	ì	ĥ		not stated	not stat	ted	not st
	7		70	mal	Scale			۲	not stated	4		0-7	
	8		80	femal	Nonpara	metric	Tests	۲	not stated	3		0-7	
	9		90	femal	Time Ser	ries		۲	not stated	0		0-7	not st
	10	1	00	mal	Survival	n		t	score of 8+	9		8+	
	11	1	10	femal	Missipa V	Multiple Response Missing Value Applysis				not stated	not stat	ted	not st
	12	1	20	mal	Missing value Analysis Complex Samples				score of O-	1		0-7	
	13	1	30	mal	not statedor			Ū	score of O-	0		0-7	
	14	1	40	mal	e	5	2.0	0	score of O-	3		0-7	

Click on b26mal and then b30mal and b34mal and move the variables via the ▶ button into the 'Dependent list' box. Click on bsex and move into the 'Independent list' box. Click on OK.



If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

MEANS TABLES=b26mal b30mal b34mal By bsex /CELLS MEAN COUNT STDDEV.

3. Were cohort members with a low birth weight (<2515 grams) more or less likely than those with a normal birth weight (2515 grams +) to be depressed at age 30? (Hint: cross-tab of b0bwghtg and b30malg).

From the Analyse drop-down menu select Descriptive Statistics and then Crosstabs. Select b0bwghtg and click the \blacktriangleright button to place the variable into the Row(s) box. Click on b30malg and click the \blacktriangleright button to place into column(s) box. Click on the 'Cells' button and select 'row' under the percentage heading. Click on the continue button. Click on OK.

🖬 b034m	b034malaise [DataSet1] - SPSS Data Editor							
File Edit '	view Data Transform An	alyze Grapi	ns Utilities V	Vindow Help				
) 📴 🔶 🔶 🗽 🖗	M 📲	nt 🗄 🗗	🖪 🐼				
1 : key	10							
	Crosstabs				X	b30ma	lg	ł
1			Devide				0-7	
2	🛷 key 🔮	N	How(s):		OK	not stat	ed	nc
3	🧳 bsex		🛷 bUbwg	ghtg	Paste		0-7	
4	🛷 b16mal				Recet	not stat	ed	nc
5	✓ b26mal		Column(s):		Tieset		0-7	
6	→ b26maig		🧳 b30ma	alg	Cancel	not stat	ed	nc
7	- Desmail	Crosstab	s: Cell Disp	av			0-7	
8		CT CONTRACT	on com one p)			0-7	
9	🧳 b1634in	Counts			Cont	inue 📃	0-7	nc
10	🥒 🔗 b2634in	🔽 Obse	rved		Can		8+	
11		Expe	cted				ed	nd
12	bUmumedg				He		0-7	
13		Percent	ages	Residuals			0-7	
14	Display clustered bar	Bow		Unstandardi	zed		0-7	
15	Suppress tables	Colur	nn	Standardize	 d		0+	
10		Tota		Adjusted sta	ndardized		81	
17	Exact][at	hot	
19	190 female	Noninte	ger Weights			a	8+	
20	200 female	n 💿 Ro	und cell counts	ORour	nd case weights		8+	
21	210 female	n O Tru	ncate cell cour	its 🔿 Truno	cate case weigh	nts	8+	nc
22	220 female	O No	adjustments			at	ed	nc
23	230 male					at	ed	nc
24	240 female		7.00	score of O-	7		0-7	

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

CROSSTABS /TABLES=b0bwghtg BY b30malg /FORMAT= AVALUE TABLES /CELLS= COUNT ROW /COUNT ROUND CELL.

- 4. Is there a relationship between family social class and mental well-being? Are men and women whose father worked in a professional occupation when they were born more or less likely than those with a father who worked in an unskilled manual job when they were born to be depressed at age 30? (Hint: the same cross-tab procedure as above using b0fsoc and b30malg variables).
- 5. Does academic achievement relate to mental well-being? Are boys or girls who have no academic qualifications by age 34 the most likely to be depressed at age 34? Is the relationship between academic achievement and mental well-being stronger for men or women? Who is the least likely of all to be depressed at age 34? (Hint: cross-tab of b34hq5 by b34malg by bsex).

From the Analyse drop-down menu select Descriptive Statistics and then Crosstabs. Select b34hq5 and click the ▶ button to place the variable into the Row(s) box. Click on b34malg and click the ▶ button to place into the column(s) box. Click on bsex and click on the bottom ▶ button and put into the 'Layer' box. Click on the 'Cells' button and select 'row' under the percentage heading. Click on the continue button. Now click on OK.

🖬 6034	malaise [Dat	aSet1] - SPS	SS Data Edit	or						
File Edit	View Data	Transform A	nalyze Graph	hs Utilities V	/indow Help					
🕞 🔲	A 🖬 🔶	🔶 🗽 🕼	M 📲	nt 🗄 🕮	FF 🗟 Ø					
1:key		10								
	keγ	bsex	b16mal	b26mal	b26malg	b30mal	b30malg	b34mal	b34malg	b1
1	10	female	3	not stated	not stated	2	0-7	.00	0-3	info
2	20	female	6-10-10-10-10-10-10-10-10-10-10-10-10-10-			not stated	not stated	pot stated	not stated	
3	30	male	Crosst	abs					0-3	info
4	40	female			_	Bow(s)			not stated	
5	50	male	🖉 key	^		A h24haE		UK	0-3	info
6	60	male	💞 b16m	al 📃		🞸 034hqo		Paste	not stated	info
7	70	male		al -I-				Beset	0-3	info
8	80	female		alg		Column(s):			0-3	info
9	90	female	✓ 030m	alo		🧳 b34malg		Cancel	not stated	info
10	100	male		al				Help	4+	info
11	110	female	🖋 b1634	4in	-Laver 1 of	1			not stated	
12	120	male	🧳 Ь2634	4in	Previous		Next		i 0-3	info
13	130	male	🚽 b0mu	med					i 0-3	info
14	140	male	💉 b0mu	medg		🛷 bsex			i 0-3	info
15	150	female	DUdad	ied					i 0-3	info
16	160	female		jeag 💌					0-3	info
17	170	female	🗌 Display	clustered bar cl	harts				4+	info
18	180	male		ss tables					J 0-3	info
19	190	female		33 (db)C3					4+	info
20	200	female	E	Exact	Statistics.	Cells	Format		0-3	info
21	210	female							not stated	info
22	220	female				not stated	not stated	not stated	not stated	
23	230	male				not stated	not stated	not stated	not stated	
24	240	famala		1 7.00	coore of 0	7	07	3.00	0.2	linfo

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

CROSSTABS /TABLES=b34hq5 BY b34malg BY bsex /FORMAT= AVALUE TABLES /CELLS= COUNT ROW /COUNT ROUND CELL.

Making comparable summary scores at age 30 and age 34

When all 24 Malaise questions are asked the cut off point to indicate that someone is experiencing depression is a score of 8+, that is 'yes' was answered to at least 8 questions. In the survey that took place in 2004 when cohort members were age 34, just 9 of the 24 Malaise questions were asked. A score of 4+ is the cut off point to indicate that someone is experiencing depression.

In the survey that took place in 2000 when cohort members were age 30 the full 24 questions were asked. A way to test the validity of only using the 9 questions

at age 34 is therefore to see if the same percentage of cohort members are identified as depressed at age 30 from their responses to the 9 questions as they are from responses to all 24 questions (b30malg).

6. Make an overall score variable from 9 questions at age 30 to match the overall score variable at age 34. Are the same percentage of men and women identified as depressed by the reduced number of questions? This can be done through the drop down menus or altering some existing syntax. (Hint: in the syntax below replace the variables starting with b34... with variables starting with b30.... Variable names correspond directly. For example b34mal02 is the same question as b30mal02)

This is the syntax used to construct the overall malaise score at age 34 (b34mal) and the grouped variable (b34malg).

count b34mal = b34mal02 b34mal03 b34mal05 b34mal09 b34mal12 b34mal14 b34mal16 b34mal20 b34mal21 (1).

The syntax below maximises the number of cohort members included in the variables by only excluding those with enough missing values to give them a 'high' (4+) malaise score.

count b34miss = b34mal02 b34mal03 b34mal05 b34mal09 b34mal12 b34mal14 b34mal16 b34mal20 b34mal21 (missing).

compute malmiss =b34mal + b34miss.

if (b34miss > 0 and b34mal <= 3 and malmiss >= 4) b34mal = -1.

If (b34miss = 9) b34mal = -2.

recode b34mal (0 thru 3=1) (4 thru highest = 2) (-1=-1) (-2=-2) into b34malg. missing values b34mal b34malg (-1,-2).

variable labels b34mal 'bcs70 age 34: total Malaise score (9 questions)'.

variable labels b34malg 'bcs70 age 34: total Malaise score - grouped'.

value labels b34mal -1'incomplete info' -2'not stated any questions'.

value labels b34malg 1'0-3' 2'4+' -1'incomplete info' -2'not stated any questions'.

freq b34mal b34malg.

Constructing the score from the drop down menus

STEP 1 Under the *Transform* drop down menu, select *Count*. In the *Target Variable* box type in the name of the new summary variable (b30mal9v). In the *Target Label* box type in an appropriate label to help identify what the variable is. For example 'BCS70 age 30: total Malaise score (9 questions)'. Scroll down the list of variables and select the 9 malaise variables that are the same as those asked at age 34 (b34mal02, b34mal03, b34mal05, b34mal09, b34mal12, b34mal16, b34mal20, b34mal21) and click on the ▶ button to place

them in the *Variables:* box. Once this has been done, click on the *Define Values* button.

🛃 *60 34	Imalaise [Da	taSet1] - SF	PSS Data Edi	tor					
File Edit	View Data	Transform A	Analyze Graph	is Utilities V	Vindow Help				
è 🔒	🔒 📴 🔶	🔶 🖢 🕻	? #4 • ¶	1 🗄 🤹	FF 🐼 Ø				
1 : key		10							
	key	bsex	b16mal	b26mal	b26malg	b30mal	b30malg	b3-	4mal
1	10	female	3	not stated	not stated	2	0-7		.00
2	20	Cour	t Occurrenc	es of Value	s within Cas	P 5			tated r
3	30	coun	it occurrent	es of Fattic	o mitini Guo				.00
4	40	Target V	'ariable:	Ta	arget Label:			ĸ	tated r
5	50	b30mal9	V	ge	: 30: total Malais	e score (9 que:	stions	_	.00
6	60	🚽 ЬЗО	mal10	Nu	umeric Variables	:	Pa	ste	tated r
7	70	🚽 ЬЗО	mal11	d	b30mal14	^	Re	set	1.00
8	80	🚮 ЬЗО	mal13		b30mal16		Car	icel	2.00
9	90	на БЗО	mal15		530mal20				tated r
10	100	5 30	mail 7		b30mal21	~	He	elp	4.00
11	110	= D30	mail8						tated r
12	120	- b30	mal22		Define Value	S			.00
13	130	- b30	mal23	. 6	f (optional o	ase selection c	ondition)		.00
14	140								.00
15	150	temale	14	7.00	score of U-	10	8+		1.00
16	160	female	11	5.00	score of O-	9	8+		2.00
4 - 1	170								

On the left hand side of the new screen, click on the *Value* button and enter 1 into the empty box. Click on the *Add* button on the right hand side of the screen to move the information into the *Values to Count* box. Click on *Continue*, then OK.

🖬 *60 34	4malaise [Da	ıtaSet1] - SI	PSS Data Edi	tor								
File Edit	View Data	Transform /	Analyze Graph	is Utilities V	Vindow Help							
🗁 📙	🔒 🖬 🔶	🔶 🐜 🛙	? #4 • ∏ I	1 🖩 🕩	F 🐼 🧿							
1:key		10										
	key	bsex	b16mal	b26mal	b26malg	b30mal	b30malg	b34mal	b34malg	b16	34in	
1	10	female	3	not stated	not stated	2	0-7	.00	0-3	info at	t least	t inf
2	20	Cour	t Occurrenc	es of Value	s within Cas	29		tated	not stated	n	io info	1
3	30		n occurrenc	cs of Futue	5 WILLING Cus	c.,		.00	0-3	info a	t age	int
4	40	Target V	(ariable:		arget Label: 20. tatal Malais		0	K tated	not stated	n	io info	-
5	50	Doumaid	JV		SU: TOTAL Malais	e score (5 ques	Ra	.00.	0-3	info a	t age	int
6	60	 b30	Imal10		umeric Variables	:		tated	not stated	info at	: least	i
	/0		Imal11		b30mal14		Re	set 1.00	0-3	info at	: least	inf
0	00 00	Count Va	lues within	Cases: Valu	es to Count						least	t inf
10	100										least	t int
10	110	Value					Values to Co	unt			h info	
12	120		3.			Add					least	t int
13	130										least	t inf
14	140	🔘 Syste	em-missing			Lhange					age	int
15	150	🔘 Syste	em- or user-missi	ng		Remove					age	int
16	160	ORang	je:								age	int
17	170			1							least	t inf
18	180	throu	ab.	1							least	t inf
19	190	- Children	ign.	-							age	int
20	200										least	t inf
21	210										least	t inf
22	220	- C Rang	je, LUWEST thr	ough value: T							p info	-
23	230										p into	1 •
24	240									_	least	t inf
25	250	ORang	je, value throug	h HIGHEST:			Continu	le Cance	Help		least	t inf
20	200										least	t inf
28	280										b info	
29	290	male	not stated			not stated	not stated	not stated	not stated	r	io info	
30	300	male	not stated	.00	score of O-	0	0-7	1.00	0-3	info at	t least	t int

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the button on the toolbar to run the command.

COUNT

b30mal9v = b30mal02 b30mal03 b30mal05 b30mal09 b30mal12 b30mal14 b30mal16 b30mal20 b30mal21 (1) .

VARIABLE LABELS b30mal9v 'BCS70 age 30: total Malaise score (9 questions)'. EXECUTE .

STEP 2 Repeat this process to count the number of missing responses for each cohort member (variable name b30miss). Instead of counting the number of yes

(value 1) responses, select *System- or user-missing* in the Values to count window.

STEP 3 The next stage is to maximise the number of cohort members we can include in the new dichotomous variable we will construct (b30mal9vg).

Under the *Transform* drop down menu, select *Compute*. In the *Target Variable* box type in the name of the new variable that will add together the number of 'yes' responses and the number of 'missing' responses (e.g. malmiss). Scroll down the list of variables to the two new variables you have just created b30mal9v and b30miss. Select b30mal9v and click on the \blacktriangleright button to move it into the *Numeric Expression:* box. Click on the + symbol. Then select b30miss and click on the \blacktriangleright button to now move this variable into the *Numeric Expression:* box. Once this has been done, click on OK.

1	5034m	alaise [Da	taSet1] - 1	SPSS Data Editor			
File	Edit V	'iew Data	Transform	Analyze Graphs Utilities Window Help			
(CH)		. 🕞 🛧	a 🖌	6 🗛 🖷 📺 🖽 🖽 🖼 🐼 🙆			
4.1							
ТСК	.ey				1		
		key	bsex	<u>b16mal b26mal b26malg b30mal b30malg b34mal b34malg</u>	<u>b16</u>	34in	
	1	10	fema	o Singlet stated instictated VIIII III III		t least ir	11
_	2	20	fema	Compute variable		o info	_
	3	30	ma	Target Variable: Numeric Expression:	1000	age II	n
	4	40	fema	maimss = DJUmai9v + DJ4miss	_	o info	_
-	5	50	ma	Type & Label	(270)	age II	n
	5	50	ma	► 530mal22		least	_
		70	ma	Function group:		least in	11
	0	00	ferma	b30mal24	<u></u>	least ir	11
	9	90 100	Terna	<i>i</i> → b34mal02 <i>i</i> → i → i → i → i → i → i → i → i → i →		least ir	
	10	100	sm sm			least li	rı
	10	110	Tema			U Inniu Lagast in	
	12	120	sm	Ø b34mal09 Image: Creation Date Creation Date Extraction		least in	rı
-	1.4	130	sini soci	D34mail2 Laure DE Employee and Considiate	in bila su	least li	rı n
	14	140	forme	Boundary Punctions and Special Var	ables:	age i	n
	10	100	ferna	✓ b34mal20		aye ii	n
	17	170	forma			looct in	ni ni
-	18	120	ma	🖋 b30mal9v		loget in	nt
-	19	100	feres	🖋 b34miss			n
~	20	200	forms			loget in	nt
	21	200	fema			least in	nt
	22	210	fema	(optional case selection condition)		n info	-
	23	220	ma			h info	-
	24	230	fema	OK Paste Reset Cancel Help		least in	n
	25	250	fema			least in	nt
	26	260	mal	e not stated not stated not stated 2 0-7 not stated not stated	info a	t least ir	nt

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

COMPUTE malmiss = b30mal9v + b30miss . EXECUTE .

STEP 4 Recoding the cohort members who have a valid score in b30mal9v to 'missing' if they did not answer all 9 questions and the number of missing answers could effect whether they were assigned a depressed label or not in the new dichotomous variable b30mal9vg.

Under the *Transform* drop down menu, again select *Compute*. In the *Target Variable* box type in the name of the existing variable b30mal9v. In the *Numeric Expression* box type in -1. (We do not want to make a new variable; we are just altering the existing one.) Click on the *If* button at the bottom of the screen. A new screen will open. Click on the button *Include if case satisfies condition:* Either type in the syntax rules (shown below) or select the variables from the variable list and click on the ▶ button to move them into the box. Likewise click on the mathematical symbols and numbers. Click on Continue and then OK. SPSS will ask if you want to change the existing variable. Click Yes.



If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the button on the toolbar to run the command.

IF (b34miss > 0 and b30mal9v <= 3 and malmiss >= 4) b30mal9v = -1. EXECUTE .

Repeat this process to assign a value of -2 to those who did not answer any of the 9 Malaise questions. (HINT: if b30miss = 9, b30mal9v = -2).

STEP 5: Recode the continuous variable into a dichotomous variable.

Under the Transform drop down menu select Recode and then Into Different Variable.

🗑 b034malaise - SPSS Data Editor							
File Edit View Data	Transform Analyze G	iraphs U	tilities Wind	low Help			
<mark>🖻 🗐 🗐 👱</mark> 1 : key	Compute Recode Visual Bander		Into Same	Variables			
	Count Rank Cases Automatic Recode		ex	b16ma	I	b26r	
1	Create Time Series Replace Missing Value	:s	emale		3	not st	
2	Random Number Seed	 ms	emale				
3	30		male		6	1	
4	40	fe	emale				
5	50		male		6		
6	60		male		3		

Scroll down the list of variables and select b30mal9v, click on the \blacktriangleright button to move the variable into the Numeric variable \rightarrow Output variable box. Type the name of the new variable in the Name box, and label in the Label box. Click on the Change button to move the information into the Numeric variable \rightarrow Output variable box. Now click on the Old and New Values button.

🛅 b034malaise	- SPSS Data Editor							
File Edit View [Data Transform Analyz	e Graphs Utilities Wind	dow Help					
🖻 🖬 🎒 🖷								
1 : key	10							
	key	bsex	b16mal	b26mal	b26malg	b30mal		
	10	famala		in at atata d				
1	10	Temale	3	not stated	not stated	2		
2	Recode	into Different Variab	les			lot stated		
3	() ⊕ b30ma ⊕ b30ma	21 🛆 Nu 22	umeric Variable -> Output Va 80mal9v> b30mal9vg	ariable:	Output Variable	4		
4	() the second s	23			b30mal9vg	ot stated		
5	() b34ma () b34ma	02		i	bcs70 age 30: total Malaise	1		
6	() b34ma () b34ma	05			Change	ot stated		
7	 	12	Old and New Values	1		4		
8	 	16	F 1	_		3		
9	() € b34ma	21	<u></u>			0		
10					Teset Lancel Help	9		
11	110	female				not stated		
40	400		wet stated	2.00		4		

Under Old value select Range and type 0 in the first box and 3 in the second box. Under New value select Value and type in 1. Click on the Add button to move the information into the Old \rightarrow New box. Repeat this for the values in the range 4 – 9 in b30mal9v and assign 2 as the new value in b30mal9vg. Under Old Value select Value and type in -1. Under New Value select Copy old value(s), then the Add button. Repeat this to copy the value -2 to the new variable. Click on Continue, then OK.

🛅 b034malaise	- SPSS Data Editor					
File Edit View (Data Transform Ana	lyze Graphs Utilities Win	dow Help			
280	ю си 🏪 📭	M 1 📩 🗏 🕮 🖡	<u>s</u> 🔊 🖉			
1:key	10					
	key	bsex	b16mal	b26mal	b26malg	b30mal
1	1	0 female	3	not stated	not stated	
2	🗖 Reco	de into Different Varial	oles			anot stated
3	🛞 ЬЗО	mal21	umeric Variable -> Output Va 80mal9v> h30mal9vo	ariable:	Output Variable	
4	Recode	into Different Variable	s: Old and New Value	S	Nal9vg	ot state
5		lue:	New Value	C System-missing	Dage 30: total Malaise	
6	C Sy C Sy	stem-missing stem- or user-missing	C Copy old valu	e(s) > New:	Change	not state
7	(• Ra 4	nge: through 9	Add Uth Change	nru 3> 1		
8	C Ra	nge:	Remove			
9	C Ra	nge:	Output variab	les are strings Width: 8		
10		through highest	Convert nume	ric strings to numbers ('5'->5	Cancel Help	
11		outer values	Continue C	Cancel Help		not state
12	12	0 male	not stated	2.00	score of 0-	
13	13	0 male	not stated	.00	score of 0-	

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the button on the toolbar to run the command.

RECODE b30mal9v (-1=Copy) (-2=Copy) (0 thru 3=1) (4 thru 9=2) INTO b30mal9vg. VARIABLE LABELS b30mal9vg 'bcs70 age 30: total Malaise score (9 questions grouped'. EXECUTE.

STEP 6: Assigning missing values and value labels to b30mal and b34mal9vg. In the Variable View window, scroll down the variables to b30mal9v and b30mal9vg. Move across to the Values column and click on the box for b30mal9vg. A new screen will open. Type in 1 in the Value box and 0-3 in the Value Label box. Click on Add. Repeat this for the following 3 values: 2= 4+, -1= incomplete info, -2=not stated any questions. Click on OK.

💼 b034malaise	b034malaise - SPSS Data Editor						
File Edit View	Data Transform Analyz	e Graphs Utilities Window	Help				
288			0				
	Name	Туре	Width	Decimals	Label	Values	Missing
104	b30mal23	Numeric	1	0	bcs70 age 30:	{*, dont know/	-1
105	b30mal24	Numeric	1	0	bcs70 age 30:	{*, dont know/	-1
106	b34mal02	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91
107	b34mal03	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91
108	b34mal05	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91
109	b34mal09	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91
110	b34mal12	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91
111	b34mal14	Numeric	2	0 🔽	lue Labels	?(🗵1
112	b34mal16	Numeric	2	0	Value Labels	<u></u> 0K	<u> </u>
113	b34mal20	Numeric	2	0 v	alue Label: not stated any questio	ns Cancel	1
114	b34mal21	Numeric	2	0	Add -1.00 = "incomplete in 1.00 = "0-3"	fo"	⁻ 1
115	b30mal9∨	Numeric	8	2	Remove		one
116	b30miss	Numeric	8	2			one
117	malmiss	Numeric	8	2		None	None
118	b30mal9∨g	Numeric	8	2	bcs70 age 30:	None 🗖	None
110							

Assign the same value labels to -1 and -2 in b30mal9v.

Click on the cell in the Missing column for variable b30mal9vg, and then select Discrete missing values. Type in -2 and -1 into the separate cells. Repeat this for variables b30mal9v.

Malaica	Evoraicae daa
ivialaise	

b034malaise - SPSS Data Editor										
File Edit View	File Edit View Data Transform Analyze Graphs Utilities Window Help									
			0		0	· · · · · · · · · · · · · · · · · · ·				
	Name	Туре	Width	Decimals	Label	Values	Missing			
104	b30mal23	Numeric	1	0	bcs70 age 30:	{*, dont know/	-1			
105	b30mal24	Numeric	1	0	bcs70 age 30:	{*, dont know/	-1			
106	b34mal02	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
107	b34mal03	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
108	b34mal05	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
109	b34mal09	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
110	b34mal12	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
111	b34mal14	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
112	b34mal16	Numeric	2	0	bcs70 age 30:	{-1, dont know	-91			
113	b34mal20	Numeric	2	О	ssing Values	[?⊠]ow	-91			
114	b34mal21	Numeric	2	0	No missing values	οκ ΙΟΨ	-91			
115	b30mal9∨	Numeric	8	2	-2 -1	Cancel	None			
116	b30miss	Numeric	8	2	Range plus one optional discrete	missing value	None			
117	malmiss	Numeric	8	2	Low: -2.00 High: -1.00)	None			
118	b30mal9∨g	Numeric	8	2	Discrete value:		-2.001.00			
119										

As you can see, once you know the syntax to write, it is a much faster process than using the drop down menus. If you save the syntax file you will also retain a record of how you made the variables.

Now (finally!) compare the frequencies of the original Malaise dichotomous variable and the new dichotomous Malaise variable from the 9 questions. Are the same percentage of cohort members identified as depressed by the reduced number of questions? (Hint: Frequencies of b30malg b30mal9vg). Is this the same for men and women? (Hint: run a cross tab of b30malg b30mal9vg by bsex)

Correlations between scores

7. Depression (malaise) is measured at age 16, 26, 30 and 34. At which two time points, ages, are the scores most strongly associated? Is this the same for men and women? (Hint: correlate the four continuous variables b16mal b26mal b30mal b34mal).

From the Analyse drop-down menu select Correlate and then Bivariate. Select b16mal and click the ► button to place it into the Variables box. Repeat for variables b26mal, b30mal, b34mal. Click OK.

b034malaise - SPSS Data Editor										
≤∎⊜ =										
378 : key	3740									
	key	bsex	b16mal	b26mal	b26malg	b30mal				
372	3680	male	not stated	11.00	score of 8+	not stated				
373	3690	female	7	6.00	score of 0-	5				
374	3700	female	🔲 Bivariate Corre	lations		0				
375	3710	female	() wey () we beex	Variables:	nal	ж 1				
376	3720	male	 	 	nal Pa	t stated				
377	3730	female			Ca	incel 1				
378	3740	female	 	✓	Н	lelp 2				
379	3750	female	Correlation Coefficier	nts		t stated				
380	3760	male	Test of Significance	.endali's tau-b Spear	man	t stated				
381	3770	female	Two-tailed	C One-tailed		t stated				
382	3780	male	✓ Flag significant cor	relations	Option	t stated				
383	3790	female	17	11.00	score of 8+	4				
384	3800	male	not stated	not stated	not stated	not stated				

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the button on the toolbar to run the command.

CORRELATIONS /VARIABLES=b16mal b26mal b30mal b34mal /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE .

Re-run the analyses for men and women separately. (Hint: use Select cases).

From the Data drop down menu, select Select Cases. Click on If condition is satisfied and then the If button. Select the bsex variable and click the \blacktriangleright button to place it into the white box. To select men, add = 1 after bsex. Click on Continue and then OK. Now re-run the Correlate command, first for men then for women. To select women only, add = 2 after bsex.

_							
i	b034malais	e - SPSS Dat	a Editor				
File	Edit View	Data Transfo	orm Analyze	e Graphs Utilities Wind	low Help		
Ē	Select C		<mark>⊷ I⊞ E</mark> a				
37			- Select-				
	 Image: white whi	^	C All cas	ses divers is settisfied		b26mal	b26malg
	(₩) b16mal	_	Se li con				
	b26malg		C Rando	om sample of cases		11.00	score of 8+
	⊕ b30m S	elect Cases:	lf			0 🗵	score of 0-
		🚸 key	^	bsex = 1		<u> </u>	score of 0-
	♣ b163	✤ b16mal				<u> </u>	score of 0-
	b0mu	✤ b26malg ♠ b26malg		+ < > 7	89 Functions:	<u>م</u> ا	not stated
	Doda b0da	✤ b30malg ♠ b30malg		· <= >= 4 × = ~= 1	ABS(numexpr)	value,) 🚔 🕽	not stated
	Current St	 b34malg b1634in 		<u>/ & _</u>	O ARTAN(numexp CDFNORM(zval	ír) lue)	score of 0-
		 b2634in b0mumed 					score of 0-
	38	×			ancer Help	<u>ل</u> ا	not stated
	381		3770	female	not stated	not stated	not stated
	382		3780	male	not stated	5.00	score of 0-
1	383		3790	female	17	11.00	score of 8+

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

USE ALL. COMPUTE filter_\$=(bsex = 1). VARIABLE LABEL filter_\$ 'bsex = 1 (FILTER)'. VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'. FORMAT filter_\$ (f1.0). FILTER BY filter_\$. EXECUTE .

NB: if you are using syntax, a simpler way of selecting men only would be to add the following instruction before an analyses.

Temporary.

Select if (bsex = 1). CORRELATIONS /VARIABLES=b16mal b26mal b30mal b34mal /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE .

Linear Regression

Linear Regression estimates the coefficients by a linear equation, involving one or more independent variables, which best predict the value of the dependent variable, in this example malaise score at age 34.

8. Although the correlation between malaise scores at age 16 and 34 was the weakest among the four scores, the positive correlation was highly significant. Using simple linear and then multiple regression we first predict a high malaise score at age 34 from malaise score at age 16, and then control for birthweight, gender, mother's age of leaving full-time education and own highest qualification by age 34.

🗰 b034malaise - SPSS Data Editor										
File Edit View I	File Edit View Data Transform Analyze Graphs Utilities Window Help									
378 : key 3740										
	bsex	b1	6mal	b26mal	b26malg	b30mal				
	· · · ·									
1	fer	Linear Regre	ssion			2				
2	fer	<pre></pre>	<u> </u>	b34mal		t stated				
		✤ bi6mal	Blo	ck 2 of 2	Pa	iste				
		 		Previous	Next Re	eset 4				
4	fer	♦ b30mal		Independent(s):		t stated				
5	r	✤ b34malg		bsex		^{eip} 1				
6	r	 ✤ b1634in ✤ b2634in 		Method:	<u> </u>	t stated				
7	r	b0mumed b0mumeda	-	Method. JEnter		4				
^		✤ b0daded	F	Selection Variable:	Bule					
8	Ter	 b0dadedg b0fsoc 		Case Labels:	1.1411-000	3				
9	fer	() b0msoc				0				
10	r	b0psoc b0bwght	_	WLS Weight:		9				
11	fer	🚸 b0bwghtg				t stated				
12			Sta	atistics Plots	Save Options	1				
12	<u>ا</u> ر									
13	m	ale not	stated	.00	score of 0-	0				
11	- m		5	2.00	and of a	2				
Click here t	o select D	escriptive]	Click here to	request a plot	of				
Statistics an	Statistics and Casewise				Standardised Predicted Residuals (Y)					
Diagnostics – this will identify against Standardised Predicted Va										
any outliers in the data										

The following syntax is produced when the paste command is selected If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the button on the toolbar to run the command. REGRESSION /DESCRIPTIVES MEAN STDDEV CORR SIG N /MISSING LISTWISE /STATISTICS COEFF OUTS R ANOVA /CRITERIA=PIN(.05) POUT(.10) /NOORIGIN /DEPENDENT b34mal /METHOD=ENTER b16mal /METHOD=ENTER b16mal bsex b0bwght b0mumed b34lefted /SCATTERPLOT=(*ZPRED ,*ZRESID) /CASEWISE PLOT(ZRESID) OUTLIERS(3) .

Fitting real data into the Linear Regression equation

The Linear Regression equation is

(*Predicted*)
$$\hat{Y} = b_0 = b_1(x_1) + b_2(x_2) + \dots + b_p(x_p)$$

Using the regression equation (above), use the statistics produced in the B column in the *Coefficients* table in the SPSS output file and the information below to calculate the estimated Malaise score for a cohort member at age 34.

This particular cohort member

- is female
- had a low birth weight of 2410 grams
- had a very high malaise score at age 16 of 44
- had a mother who left education at age 15,
- had an A/S level qualification (highest academic qualification achieved)

SPSS syntax to identify this cohort member is

temporary. select if (bsex = 2 and b16mal = 44 and b0mumed = 15 and b0bwght = 2410 and b34hq13 = 4). SUMMARIZE /TABLES=key bsex b16mal b0mumed b0bwght b34hq13. /FORMAT=VALIDLIST NOCASENUM TOTAL LIMIT=100 /TITLE='Case Summaries' /MISSING=VARIABLE /CELLS=COUNT .

Compare this with their actual Malaise score at age 34 (Y). How accurate was the predicted \hat{Y} ? (Hint: to find the individual cohort member in the data file, go to the *Data View* window and place the curser on the *Key* column, select *Find* from the *Edit* drop down menu. The Key number to find is 170. Look under b34mal to see the actual score obtained at age 34.

🖩 b034malaise - SPSS Data Editor 📃 🖻 🔀											
File Edit View	File Edit View Data Transform Analyze Graphs Utilibies Window Help										
	I I I I I I I I I I I I I I I I I I I										
17 : key	170										
	key	bsex	b16mal	b2	26mal	b26malg	b30mal	b30malg	b34mal	b34malg	
2	20	female	not stated	not	stated	not stated	not stated	not stated	not stated	not stated	
3	30	male	6		10.00	score of 8+	4	0-7	.00	0-3	
4	40	female	not stated	not	stated	not stated	not stated	not stated	not stated	not stated	
5	50	male	6		1.00	score of 0-	1	0-7	.00	0-3	
6	Find Data in	variable key			stated	not stated	not stated	not stated	not stated	not stated	
7		8			stated	not stated	4	0-7	1.00	0-3	
8	Find what	170		•	stated	not stated	3	0-7	2.00	0-3	
9			Q		stated	not stated	0	0-7	not stated	not stated	
10	I Match	case	`		8.00	score of 8+	9	8+	4.00	4+	
11	Find Next		Stop Can	cel	stated	not stated	not stated	not stated	not stated	not stated	
12	120	mare	not stated		2.00	score of 0-	1	0-7	.00	0-3	
13	130	male	not stated		.00	score of 0-	0	0-7	.00	0-3	
14	140	male	5		2.00	score of 0-	3	0-7	.00	0-3	
15	150	female	14		7.00	score of 0-	10	8+	1.00	0-3	
16	160	female	11		5.00	score of 0-	9	8+	2.00	0-3	
17	170	female	44	not	stated	not stated	11	8+	5.00	4+	
18	180	male	3		.00	score of 0-	not stated	not stated	.00	0-3	
	w V valiable view \					SPSS F	rocessor is ready				
🛃 start	Aicrosoft Outlook .	🛛 🗢 KINGSTON (E:)	W 3 Microsoft		- 🛗 b034ma	ilaise - SPS 🛗 Out	put1 - SPSS Vie 👔	b034malaise - SPS	🕄 🗘 🕢 Norton" 🤇	SE ^N 🔊 🛢 🖸 🖉 15:22	

You could repeat the above analysis using different but similar variables in the data file. For example, replace highest qualification at age 34 with the age the cohort member left full-time education (b34lefted), mothers age of leaving full-time education (b0daded) or family social class at some point in childhood (b0psoc, b5psoc, etc).

Logistic Regression

We can run a similar analysis using logistic regression to predict depression at age 34 from signs of depression at age 16, birth weight, gender, mother's age of leaving full-time education and own highest qualification by age 34.

From the *Analyse* drop down menu select *Regression* and then *Binary Logistic* to open the Logistic Regression dialog box. Transfer the categorical Malaise score variable b34malg into the *Dependent* box, and the (predictor) variable into the *Covariates* box. In the first analysis just move the grouped malaise score at age 16 (b16malg). In the second stage of the analysis include the other measures from birth using the categorical and not continuous measures for age mother left full-time education (b0mumedg), birth weight (b0bwghtg), gender (bsex) and highest qualification at age 34 (b34hq5).

Reference Category

Categorical variables require a reference category to be set. There are different methods of contrasting category membership in a variable. The default method in SPPS is *Indicator*, with the last (highest value) category being set as the reference category. (In the SPSS output, the reference category is represented in the contrast matrix as a row of zeros.)

In our model, this would mean that the reference category in the variable b34hq5 would be set as category 5'higher degree / PGCE'. The relative importance of membership to each of the other qualification groups (values 0'none' up to 4'degree....nursing qual') will be compared against membership to category 5'higher degree / PGCE'.

Click on the *Categorical* button to open the *Define Categorical Variables* box. Move the covariates into the *Categorical Covariates* box. The default makes the last (highest value) category of a variable the reference category. To change the reference category to the first (lowest value) category, click on *First* and then *Change*. Do this for the each of the following variables b16malg, bsex, b0bwghtg, b0meumedg. Click on *Continue* and then OK.

🛅 b034malaise	- SPSS Data Editor										
File Edit View	ile Edit View Data Transform Analyze Graphs Utilities Window Help										
28	N 🗠 🏪 💽	M <u>* i </u>									
15 : key	5 : key 150										
	key	bsex	b16mal	b26mal	b26malg	b30mal	b30				
2	20) female	not stated	not stated	not stated	not stated	not s				
3	🗖 Logisti	c Regression			ore of 8+	4					
4	 				bt stated	not stated	not s				
5	() €10 ± () €10 ±	ogistic Regression: Def	ine Categorical Varia	bles	of 0-	1					
6	 	Lovanates.	b16malg(Indicator(fir base(Indicator(fir	st))	Continue ated	not stated	not s				
7	 	Г	b0mumedg(Indicator) b0bwghtg(Indicator)	(first))	Help ated	4					
8	● b30 ● b34	L. L.	b34hq5(Indicator)		ated	3					
9	() € 534 () € 534		Change Contrast	Change	ated	0					
10	() € 534 () € 516		Reference Categor	y: C Last € First	of 8+	9					
11					ated	not stated	not s				
12					core of 0-	1					
13	130) male	not stated	.00	score of 0-	0					
14	140) male	5	2.00	score of 0_	3					

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

Initial analyses

LOGISTIC REGRESSION b34malg /METHOD = ENTER b16malg /CONTRAST (b16malg)=Indicator(1) /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

Second analyses

LOGISTIC REGRESSION b34malg /METHOD = ENTER b16malg bsex b0mumedg b0bwghtg b34hq5 /CONTRAST (b16malg)=Indicator(1) /CONTRAST (bsex)=Indicator(1) /CONTRAST (b0mumedg)=Indicator(1) /CONTRAST (b0bwghtg)=Indicator(1) /CONTRAST (b34hq5)=Indicator /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

Looking at the Variables in the Equation table

- How much more likely were depressed teenagers to grow up to be a depressed 34 year old?
- After taking into account gender, birth weight, age mother left education and own highest qualification, how much more likely were depressed teenagers to grow up to be depressed at age 34?

• After malaise score at age 16, which variable makes the biggest contribution to the model?

Further stages of analysis

Repeat the above analysis adding in new variables or replacing variables with similar measures. For example, fathers age of leaving full-time education (b0dadedg). Remember to think about whether it makes more sense to set the first or last category as the reference category. If a variable has many categories (e.g. family social class), recode to have fewer categories.

- Is mothers or fathers age of leaving full-time education the stronger measure (bigger coefficient) for predicting depression at age 34?
- Which variable makes a significant contribution to the model when mothers age of leaving full-time education is included, but not when fathers age of leaving full-time education is included?

If you have the inclination, you could also try repeating the above analyses for men and women separately.

• Are the results the same for men and women separately?

Interaction Terms

For each variable in the logistic model, it is assumed that the effect is the same for all values of other variables in the model. For example, the effect of level of a mother's education is the same for men as it is for women. If this is not the case, there is an interaction. You can include a term for the *level of mother's education* (b0mumedg) by gender (bsex) interaction in your model.

To include an interaction term between depression at age 16 and gender in your model click on bsex and then b16malg whilst simultaneously pressing the *Ctrl* key on your PC keyboard. This will mean both variables are highlighted. Click on the >a*b> button to place an interaction term in the Covariates dialog box. Continue as before.

🛅 b034malaise	e - SPSS Data Editor				
File Edit View	Data Transform Analyze	e Graphs Utilities Wind	low Help		
	4 <u>9</u>				
1 : key	10				
	key	bsex	b16mal	b16malg	b2
	2				
1	10	fomalo	3	1.00	not
	Logistic Regression			00	not
	🛞 key 🔼	Dependent:		OK .00	not
		Block 1 of 1		Paste .00	
	b16malg	Previous	Next	Reset .00	not
	 ✤ b26malg 	Covariates:		Cancel .00	
	 	b16malg(Cat) bsex(Cat)			not
		>a*b> b0mumedg(Ca b0bwghtg(Cat	at))	.00	not
	 	Method: Louis		.00	not
	b0mumed	Calcoling Maria	<u> </u>	.00	not
	boldaded	Selection Varia	Bule	.00	not
1		Categorical	Save Options	. 00	
1,		теппане	not stated	-2.00	not
12	120	male	not stated	-2.00	

If PASTE is selected instead of OK, a syntax window will open and the following syntax command will appear in it. Highlight the syntax and then click on the ▶ button on the toolbar to run the command.

LOGISTIC REGRESSION b34malg /METHOD = ENTER b16malg bsex b0mumedg b0bwghtg b34hq5 b16malg*bsex /CONTRAST (b16malg)=Indicator(1) /CONTRAST (bsex)=Indicator(1) /CONTRAST (b0mumedg)=Indicator(1) /CONTRAST (b0bwghtg)=Indicator(1) /CONTRAST (b34hq5)=Indicator /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

• Is the interaction significant? Try repeating this using the continuous malaise score at age 16 variables (b16mal). Are the same results obtained?

Remember all of these exercises can be repeated for cross-cohort comparisons using similar variables in the NCDS data file *n042malaise.sav*.