Answers to Sample exam

PART A

Q1. *Students should use the codebook shown in Part A of the appendix of the* SPSS Survival Manual *as a model.*

SPSS Variable name	Full variable name	Coding instruction	Measurement level
id	Identification number	Assign unique	scale
		Identification number to	
		each case	
sex	Sex	1=male 2=female	nominal
age	Age	Age in years	scale
educ	Education level	1=year 10 2=year12	ordinal
		3=University/college	
diet	Diet to lose weight	1= yes 2=no	nominal
mast1 to mast7	Mastery scale items	1= strongly disagree	scale
		4=strongly agree	

An example of a possible answer is shown below.

Q2. Students should follow the instructions described in Chapter 4 of the SPSS Survival Manual (see the section 'Defining the Variables').

1. Click on the Variable View tab in the Data Editor window. Click in the cell under the heading

Values for the educ variable. Click on the blue box to open the Value Labels dialogue box.

2. Click in the box marked **Value**. Type in 1. Click in the box marked **Label**. Type in Year 10. Click on **Add**.

3. Repeat for the values 2 (Year 12) and 3 (University/college. Click on OK.

Q3. The datafile created by the student should be consistent with the codebook that they developed in *Q1*, showing the same variable names, labels, measurement level and value labels.

Q4.

(a)

STEP 1: Checking for errors.

To identify cases with values outside of the possible range run **Frequencies** and request the minimum and maximum values. If either of these values are outside the possible range of values (as specified in the codebook) sort the cases for this variable (either ascending or descending) to identify the ID of the out of range cases.

(b) Students should show that they have run Frequencies on each of the variables requesting the min and max values. Any values that fall out of this range should be noted.

Copyright © Julie Pallant 2015.

Q5. Students' files should show that they have correctly reverse scored items 1 3 4 6 and 7 of the Mastery Scale (shown in the Sample Questionnaire as items 1 to 7). These should have been rescored so that 1=4, 2=3, 3=2, 4=1, consistent with the procedures presented in Chapter 8 of the *SPSS Survival Manual*. These reverse scored items should appear as separate variables in the datafile (preferably with a R or REV in the name to indicate that they have been reversed). The total score should then have been calculated by adding scores for all items (using the reverse scored items 1 3 4 5 7 and the original item 2 and 5), with a possible min value of 7 and a max value of 28.

The mean and SD value for this new total scale score should have been calculated using either the SPSS **Frequency** procedure or **Descriptives** procedure as described in Chapter 6 of the *SPSS Survival Manual*.

Q6. The new education variable should have been created by following the instructions in Chapter 8 in the section 'Collapsing the number of categories of a categorical variable'. The new variable should contain only two values as specified in the question.

Q7. The new age group variable (agegp4) should be included in the list of variables (usually at the bottom of the list of variables). It should have been created by following the instructions shown in Chapter 8 of the *SPSS Survival Manual* in the section 'Collapsing a continuous variable into groups'. It should only contain 4 values, and each should have a value label attached that matches the values as specified in the question. These values should have been specified in the Variable View tab of the Data Editor window.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

PART B

Q1

(a) The assumption being tested is linearity, checking that there is a linear (straight-line) relationship between variable X and Y. The procedure is described in Chapter 11 of the *SPSS Survival Manual*.
(b) The scatterplot suggests that there is a curvilinear relationship between the two variables. It is therefore not appropriate to calculate a Pearson correlation coefficient as it would tend to underestimate the true strength of the association.

Q2

(a) The parametric technique that would be appropriate is an independent-samples t-test. There are two groups (males and females), which are independent of one another, being compared on a continuous measure (weekly coffee consumption).

(b) The main assumption to be tested is equal variances in coffee consumption scores for males and females. This is tested using Levene's test of equality of variance which is reported as part of the t-test output.

(c) The first value to check for is the p value for Levene's test. If this is non-significant (p>.05) then we can assume no violation of the assumption of equal variances.

If this is the case you would then move across the first row of the t-test results table to obtain the p value which compares male and female coffee consumption. If the p value is less than .05 we can conclude that there is a significant difference in the coffee consumption of males and females. The mean consumption values for males and females are presented in the Group Statistics table provided by SPSS.

(d) The equivalent non-parametric test is a Mann-Whitney U test.

Q3

(a) The results shown are from a Two-way between groups ANOVA.

(b) This technique allows us to determine

- i. if there is difference in perceived stress levels between males and females (main effect)
- ii. if there is a difference in perceived stress levels across 5 different age groups (main effect)
- iii. if there is an interaction between age and sex in their effect on perceived stress (interaction effect).

(c) There is a significant main effect for both sex (p=.004) and age group (p=.005) however there is no interaction effect between sex and age group (p=.74).

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

Q4

(a) The test that may be appropriate to use here is a one-way repeated measures analysis of variance. This test is appropriate when you assess a group of participants on the same measure on three or more different occasions and wish to know if there are differences among the three sets of scores.(b) The non-parametric equivalent is a Friedman test.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

PART C

Q1

(a) N/A.

(**b**) The sample consisted of 439 cases ranging in age from 18 to 82 (M=37.44, SD=13.2). Forty-two percent (n=185) were male and 57.9% (n=254) were female. The majority of the sample were in a relationship (married for the first time n=189, 43%; living with a partner n=37 8.4% or in a steady relationship n=37 8.4%). Almost half of the sample had completed tertiary education (undergraduate n=123 28%, postgraduate n=56 12.8%).

Q2. The procedure for calculating Cronbach alpha to assess the internal consistency reliability of a scale is presented in Chapter 9 of the *SPSS Survival Manual*.

- (a) Cronbach alpha for Mastery Scale = .76
- (**b**) Cronbach alpha for Life Satisfaction Scale = .89
- (c) Cronbach alpha for Perceived Control of Internal States Scale = .90
- (d) Cronbach alpha for Perceived Stress Scale = .85

Q3. The procedure for assessing the distribution of scores on a scale is presented in Chapter 6 of the *SPSS Survival Manual*. The main focus is the inspection of the histogram and Normal Q-Q plot as the the Kolmogorov-Smirnov statistic, a statistical measure of normality, is overly sensitive and often indicates non-normality in large samples.

(a) Mastery scale

Inspection of the histogram and Normal Q-Q plot suggests that the distribution of scores on the Mastery scale are slightly skewed, with scores bunching at the high end, and an extended tail at the low end.



Copyright © Julie Pallant 2015.



(b) Life Satisfaction Scale

The histogram suggests that the scores on the Life Satisfaction Scale are reasonably normally distributed. This is also supported by inspection of the Normal Q-Q plot



Copyright © Julie Pallant 2015.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

Observed Value

(c) Perceived Control of Internal States Scale

Overall scores on the Perceived Control of Internal States Scale approximated a normal distribution however there was evidence of a number of outlying cases at the lower end of the distribution (as shown in the Histogram and Normal Q-Q plot).



(d) Perceived Stress Scale

Both the histogram and Normal Q-Q Plot suggest that scores on the Perceived Stress Scale approximates a normal distribution.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.



Q4

(a) The scores on the Perceived Stress Scale were normally distributed therefore it would be appropriate to use an **Independent Samples t-test** to explore gender differences.

The procedure for this test is provided in Chapter 17 of the SPSS Survival Manual.

The Levene's Test for Equality of Variances was not significant (p=.165) indicating no violation of the assumption of equal variances. This means that the results from the first line in the output for the Independent Samples Test can be reported.

The effect size statistic (eta squared) is calculated using the formula provided in Chapter 17 in the section 'Calculating the effect size for for independent samples t-test'.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

Output

Group Statistics

	sex sex	N	Mean	Std. Deviation	Std. Error Mean
tpstress Total perceived	1 MALES	184	25.79	5.414	.399
stress	2 FEMALES	249	27.42	6.066	.384

Independent Samples Test

		Levene for Equ Varia	e's Test ality of ances			t-test f	for Equal	ity of Me	ans	
						Sig. (2-	Mean Differe	Std. Error Differe	95 Confi Interva Diffe	% dence l of the rence
		F	Sig.	t	df	tailed)	nce	nce	Lower	Upper
tpstress Total perceived stress	Equal variances assumed	1.936	.165	- 2.89 8	431	.004	-1.634	.564	-2.742	526
	Equal variances not assumed			- 2.94 8	415. 886	.003	-1.634	.554	-2.723	544

(**b**) Reporting the results in a report

An independent-samples t-test was conducted to compare perceived stress scores for males and females. There was a statistically significant difference in scores for males and females t(431)=-2.89, p=.004. Females recorded higher mean scores (M=27.42, df=6.07) than males (M=25.79, df=5.41). The mean difference in scores was 1.63 (95% CI=-2.74 to -.53). The eta squared value was .019 indicating a small effect size.

Q5

(a) To explore the relationship between scores on the perceived control (PCOISS) and life satisfaction scales we need to use one of the correlational techniques described in Chapter 11 of the *SPSS Survival Manual* (Pearson correlation or Spearman Rank Order Correlation).

There are a number of assumptions to test before deciding on the most appropriate correlation technique to use. These are discussed in the introduction to 'Part Four: Statistical techniques to explore relationships among variables'.

Copyright © Julie Pallant 2015.

Assumption testing

Level of measurement: Both scales have a good range of scores and are often treated as having interval scale measurement, particularly by researchers in psychology and business, who often use Pearson correlation coefficients. Technically however the level of measurement is ordinal (which does not assume equal distances between response options). In the health and medical literature you will often see Spearman Rank Order correlation coefficients used.

Normality: It is necessary to check the distribution of scores for the two scales (PCOISS: perceived control, life satisfaction). The scores for the PCOISS scale are normally distributed (see the answer to Q3 above). An additional histogram should be generated to check the life satisfaction scale. This can be done in a number of ways however the most thorough is by using the Explore procedure (described in the section 'Assessing Normality' in Chapter 6 of the *SPSS Survival Manual*).

Output from Explore



The distribution of scores on the life satisfaction scale approximates a normal curve. Therefore this supports the use of parametric techniques (Pearson correlation coefficient).

Linear relationship. The third assumption to test is the linearity of the relationship between the two variables. This is tested using a scatterplot as described in Chapter 11 of the *SPSS Survival Manual*.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.



Inspection of the scatterplot does not show any signs of curvilinearity. It is therefore appropriate to proceed with the calculation of Pearson correlation coefficient.

Calculation of correlation coefficient

The instructions for generating a Pearson correlation coefficient is described in Chapter 11 of the *SPSS Survival Manual*.

The output is shown below:

		tpcoiss Total PCOISS	tlifesat Total life satisfaction
tpcoiss Total PCOISS	Pearson Correlation	1	.373**
	Sig. (2-tailed)		.000
	Ν	430	429
tlifesat Total life	Pearson Correlation	.373**	1
satisfaction	Sig. (2-tailed)	.000	
	Ν	429	436

Correlations

**. Correlation is significant at the 0.01 level (2-tailed).

(b) Presenting the results in a report

The relationship between perceived control (measured by the Perceived Control of Internal States Scale) and life satisfaction (measured by the Life Satisfaction Scale) was investigated using a Pearson Product Moment correlation coefficient. Preliminary analyses were performed to check the assumptions of normality, linearity and homoscedasticity. There was a moderate positive correlation

Copyright © Julie Pallant 2015.

between the two variables (r=.37, n=429, p<.001) with high levels of perceived control associated with higher life satisfaction.

Q6

(a) The appropriate statistical test to use to address this research question is a two-way between groups analysis of variance which is described in Chapter 19 of the *SPSS Survival Manual*.

Output from 2-way between groups ANOVA

Descriptive Statistics

Dependent Variable: tlifesat Total life satisfaction

agegp3 age 3 groups	sex sex	Mean	Std. Deviation	Ν
1 18 - 29	1 MALES	21.67	6.456	60
	2 FEMALES	23.29	6.624	87
	Total	22.63	6.583	147
2 30 - 44	1 MALES	21.40	6.611	68
	2 FEMALES	21.82	7.083	85
	Total	21.63	6.858	153
3 45+	1 MALES	22.00	6.593	57
	2 FEMALES	23.63	6.978	79
	Total	22.95	6.842	136
Total	1 MALES	21.67	6.525	185
	2 FEMALES	22.90	6.911	251
	Total	22.38	6.770	436

Levene's Test of Equality of Error

Variances^a

Dependent Variable: tlifesat Total life

satisfaction

F	df1	df2	Sig.
.503	5	430	.774

Tests the null hypothesis that the error

variance of the dependent variable is equal

across groups.

a. Design: Intercept + agegp3 + sex + agegp3

* sex

Copyright © Julie Pallant 2015.

	Type III Sum					Partial Eta
Source	of Squares	df	Mean Square	F	Sig.	Squared
Corrected Model	326.421 ^a	5	65.284	1.432	.212	.016
Intercept	211053.805	1	211053.805	4628.341	.000	.915
agegp3	111.391	2	55.695	1.221	.296	.006
sex	159.641	1	159.641	3.501	.062	.008
agegp3 * sex	35.094	2	17.547	.385	.681	.002
Error	19608.136	430	45.600			
Total	238281.000	436				
Corrected Total	19934.557	435				

Tests of Between-Subjects Effects Dependent Variable: tlifesat Total life satisfaction

a. R Squared = .016 (Adjusted R Squared = .005)

(b) Reporting the results

A two-way between-groups analysis of variance was conducted to explore the impact of sex and age on levels of life satisfaction. Participants were divided into three groups according to their age (Group1: 18-29 years; Group 2: 30-44 years; Group 3: 45 years and above). The interaction effect between sex and age group was not statistically significant, F(2,430)=.39, p=.68, with a very small effect size (partial eta squared=.002). There was no significant main effect for either age F(2,430)=1.22, p=.30, or sex F(1,430)=3.5, p=.06, and both effect sizes were very small (age: partial eta squared=.006, sex: partial eta squared=.008).

Q7. The appropriate statistical test to use is a standard multiple regression described in Chapter 13 of the *SPSS Survival Manual*.

	Correl	lations		
		tlifesat Total life satisfaction	toptim Total Optimism	tpcoiss Total PCOISS
Pearson Correlation	tlifesat Total life satisfaction	1.000	.483	.373
	toptim Total Optimism	.483	1.000	.513
	tpcoiss Total PCOISS	.373	.513	1.000
Sig. (1-tailed)	tlifesat Total life satisfaction		.000	.000
	toptim Total Optimism	.000	•	.000
	tpcoiss Total PCOISS	.000	.000	•
Ν	tlifesat Total life satisfaction	436	435	429
	toptim Total Optimism	435	435	428
	tpcoiss Total PCOISS	429	428	430

Output from Regression

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

Model Summary

			Adjusted R	Std. Error of
Model	R	R Square	Square	the Estimate
1	.505ª	.255	.251	5.858

a. Predictors: (Constant), tpcoiss Total PCOISS, toptim Total Optimism

	Coefficients ^a							
_				Standardized				
		Unstandardize	d Coefficients	Coefficients				
Model		B Std. Error		Beta	t	Sig.		
1	(Constant)	3.175	1.662		1.910	.057		
	toptim Total Optimism	.605	.075	.396	8.112	.000		
	tpcoiss Total PCOISS	.096	.028	.170	3.484	.001		

a. Dependent Variable: tlifesat Total life satisfaction

Interpretation of output

The optimism scale was a better predictor of life satisfaction than perceived control (PCOISS). The standardized regression coefficient (beta) for optimism was .396, compared with .17 for the PCOISS. Both were significant unique predictors of life satisfaction (p<.01).

The model explained a total of 25.5% of the variance in life satisfaction scores (R square=.255).

Q8

The appropriate statistical analysis to use to address this question is chi-square test for independence which is presented in Chapter 16 of the *SPSS Survival Manual*.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

			smoke	smoker	
			1 YES	2 NO	Total
agegp5 age 5 groups	1 18 - 24	Count	16	77	93
		% within agegp5 age 5 groups	17.2%	82.8%	100.0%
	2 25 - 32	Count	25	62	87
		% within agegp5 age 5 groups	28.7%	71.3%	100.0%
	3 33 - 40	Count	19	64	83
		% within agegp5 age 5 groups	22.9%	77.1%	100.0%
	4 41 - 49	Count	12	81	93
		% within agegp5 age 5 groups	12.9%	87.1%	100.0%
	5 50+	Count	13	67	80
		% within agegp5 age 5 groups	16.3%	83.8%	100.0%
Total		Count	85	351	436
		% within agegp5 age 5 groups	19.5%	80.5%	100.0%

agegp5 ag	e 5	groups *	smoke smoker	Crosstabulation
		8-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0		

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	8.766 ^a	4	.067
Likelihood Ratio	8.594	4	.072
Linear-by-Linear Association	1.664	1	.197
N of Valid Cases	436		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 15.60.

Symmetric Measures

			Approximate
		Value	Significance
Nominal by Nominal	Phi	.142	.067
	Cramer's V	.142	.067
N of Valid Cases		436	

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.

Presenting results

A chi-square test of independence indicated no significant association between age and smoking status $\Box 2$ (4, *n*=436)=8.77, *p*=.067, *phi*=.14.

The answers to the statistics revision test are intended for us by instructors only. Instructors are free to use, copy and manipulate the material in this manual on the condition that it is not distributed to students under any circumstances except for formal testing.