ANSWERS TO EXERCISES AND REVIEW QUESTIONS

PART FOUR: STATISTICAL TECHNIQUES TO EXPLORE RELATIONSHIPS AMONG VARIABLES

You should review the material in the introduction to Part Four and in Chapters 11, 12, 13, 14 and 15 of the *SPSS Survival Manual* before attempting these exercises.

Correlation

4.1 Using the data file survey.sav follow the instructions in Chapter 11 to explore the relationship between the total mastery scale (measuring control) and life satisfaction (tlifesat). Present the results in a brief report.

Correlations

		tlifesat total life satisfaction	tmast total mastery
tlifesat total life satisfaction	Pearson Correlation	1	.444**
	Sig. (2-tailed)		.000
	N	436	436
tmast total mastery	Pearson Correlation	.444**	1
	Sig. (2-tailed)	.000	
	N	436	436

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

The relationship between mastery and life satisfaction was explored using Pearson's product moment correlation. There was a moderate positive correlation (r=.44, p<.0001) suggesting that people who felt they had control over their lives had higher levels of life satisfaction.

4.2 Use the instructions in Chapter 11 to generate a full correlation matrix to check the intercorrelations among the following variables.

(a) age
(b) perceived stress (tpstress)
(c) positive affect (tposaff)
(d) negative affect (tnegaff)
(e) life satisfaction (tlifesat)

Correlations

		age	tpstress total perceived stress	tposaff total positive affect	tnegaff total negative affect	tlifesat total life satisfaction
age	Pearson Correlation	1	-.127**	.069	-.171**	.059
	Sig. (2-tailed)		.008	.150	.000	.222
	N	439	433	436	435	436
tpstress total perceived stress	Pearson Correlation	-.127**	1	-.442**	.674**	-.494**
	Sig. (2-tailed)	.008		.000	.000	.000
	N	433	433	433	432	433
tposaff total positive affect	Pearson Correlation	.069	-.442**	1	-.294**	.415**
	Sig. (2-tailed)	.150	.000		.000	.000
	N	436	433	436	435	436
tnegaff total negative affect	Pearson Correlation	-.171**	.674**	-.294**	1	-.316**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	435	432	435	435	435
tlifesat total life satisfaction	Pearson Correlation	.059	-.494**	.415**	-.316**	1
	Sig. (2-tailed)	.222	.000	.000	.000	
	N	436	433	436	435	436

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

4.3 Gill, a researcher, is interested in exploring the impact of age on the experience of positive affect (tposaff), negative affect (tnegaff) and perceived stress (tpstress).

(a) Follow the instructions in Chapter 11 of the *SPSS Survival Manual* to generate a condensed correlation matrix which presents the correlations between age with positive affect, negative affect and perceived stress.

Correlations

		tposaff total positive affect	tnegaff total negative affect	tpstress total perceived stress
age	Pearson Correlation	.069	-.171**	-.127**
	Sig. (2-tailed)	.150	.000	.008
	N	436	435	433

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

(b) Repeat the analysis in (a), but first split the sample by sex. Compare the pattern of correlations for males and females. Remember to turn off the **Split File** option after you have finished this analysis.

Correlations
sex sex = MALES

Correlations [a]

		tposaff total positive affect	tnegaff total negative affect	tpstress total perceived stress
age	Pearson Correlation	.061	-.123	-.186*
	Sig. (2-tailed)	.406	.095	.012
	N	185	185	184

*. Correlation is significant at the 0.05 level (2-tailed).

a. sex sex = MALES

sex sex = FEMALES

Correlations [a]

		tposaff total positive affect	tnegaff total negative affect	tpstress total perceived stress
age	Pearson Correlation	.073	-.208*	-.100
	Sig. (2-tailed)	.246	.001	.115
	N	251	250	249

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

a. sex sex = FEMALES

Partial correlation

4.4 Follow the procedures detailed in Chapter 12 of the *SPSS Survival Manual* to calculate the partial correlation between optimism (toptim) and perceived stress (tpstress) while controlling for the effects of age. Compare the zero order correlations with the partial correlation coefficients to see if controlling for age had any effect.

Correlations

Control Variables			toptim total optimism	tpstress total perceived stress	age
-none- [a]	toptim total optimism	Correlation	1.000	-.469	.201
		Significance (2-tailed)	.	.000	.000
		df	0	430	433
	tpstress total perceived stress	Correlation	-.469	1.000	-.127
		Significance (2-tailed)	.000	.	.008
		df	430	0	431
	age	Correlation	.201	-.127	1.000
		Significance (2-tailed)	.000	.008	.
		df	433	431	0
age	toptim total optimism	Correlation	1.000	-.456	
		Significance (2-tailed)	.	.000	
		df	0	429	
	tpstress total perceived stress	Correlation	-.456	1.000	
		Significance (2-tailed)	.000	.	
		df	429	0	
	age	Correlation			
		Significance (2-tailed)			
		df			

a. Cells contain zero-order (Pearson) correlations.

The zero order correlation (not controlling for age) is -.469 indicating a moderate negative correlation between optimism and levels of perceived stress. The partial correlation coefficient (controlling for the effects of age) is -.456, which is only slightly lower. This indicates that the relationship between optimism and perceived stress is not influenced by age.

Multiple regression

4.5 There are three main types of multiple regression analyses. What are they? When would you use each approach?

Standard multiple regression
In standard multiple regression all the independent (or predictor) variables are entered into the equation simultaneously. Each independent variable is evaluated in terms of its predictive power, over and above that offered by all the other independent variables. This approach would be used if you had a set of variables (e.g., various personality scales) and wanted to know how much variance in a dependent variable (e.g., anxiety) they were able to explain as a group or block. This approach would also tell you how much unique variance in the dependent variable that each of the independent variables explained.

Hierarchical multiple regression
In hierarchical regression (also called sequential) the independent variables are entered into the equation in the order specified by the researcher based on theoretical grounds. Variables or sets of variables are entered in steps (or blocks), with each independent variable being assessed in terms of what it adds to the prediction of the dependent variable, after the previous variables are controlled for. For example, if you wanted to know how well optimism predicts life satisfaction, after the effect of age is controlled for, you would enter age in Block 1 and then Optimism in Block 2. Once all sets of variables are entered, the overall model is assessed in terms of its ability to predict the dependent measure. The relative contribution of each block of variables is also assessed.

Stepwise multiple regression
In stepwise regression the researcher provides SPSS with a list of independent variables and then allows the program to select which variables it will enter, and in which order they go into the equation, based on a set of statistical criteria. This would be used when you have a large number of predictor variables, and no underlying theory concerning their possible predictive power.

4.7 As part of the preliminary screening process it is recommended that you inspect the Mahalanobis distances produced by SPSS. What do these tell you?

The Mahalanobis distances produced by SPSS can be used to detect the presence in your datafile of multivariate outliers, people with a strange set of scores on your predictor variables.

4.8 The example used in the *SPSS Survival Manual* to demonstrate the use of standard multiple regression compares two control measures (PCOISS and Mastery) in terms of their ability to predict perceived stress. Repeat this analysis, this time using life satisfaction (tlifesat) as your dependent variable. Use the output to answer the following questions.

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
tlifesat total life satisfaction	22.38	6.770	436
tpcoiss total PCOISS	60.63	11.985	430
tmast total mastery	21.764	3.9696	436

Correlations

		tlifesat total life satisfaction	tpcoiss total PCOISS	tmast total mastery
Pearson Correlation	tlifesat total life satisfaction	1.000	.373	.444
	tpcoiss total PCOISS	.373	1.000	.521
	tmast total mastery	.444	.521	1.000
Sig. (1-tailed)	tlifesat total life satisfaction	.	.000	.000
	tpcoiss total PCOISS	.000	.	.000
	tmast total mastery	.000	.000	.
N	tlifesat total life satisfaction	436	429	436
	tpcoiss total PCOISS	429	430	429
	tmast total mastery	436	429	436

Variables Entered/Removed [b]

Model	Variables Entered	Variables Removed	Method
1	tmast total mastery, tpcoiss total PCOISS [a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: tlifesat total life satisfaction

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.474[a]	.225	.221	5.975

a. Predictors: (Constant), tmast total mastery, tpcoiss total PCOISS

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4407.034	2	2203.517	61.729	.000[a]
	Residual	15206.737	426	35.697		
	Total	19613.771	428			

a. Predictors: (Constant), tmast total mastery, tpcoiss total PCOISS

b. Dependent Variable: tlifesat total life satisfaction

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	2.997	1.774		1.690	.092					
	tpcoiss total PCOISS	.110	.028	.195	3.903	.000	.373	.186	.166	.729	1.372
	tmast total mastery	.584	.085	.342	6.850	.000	.444	.315	.292	.729	1.372

a. Dependent Variable: tlifesat total life satisfaction

(a) Overall, how much of the variance in life satisfaction is explained by these two variables?

The R squared value of .225 indicates that 22.5% of the variance in life satisfaction scores is explained by the two predictor variables (tmast, tpcoiss).

(b) Which of the independent variables (tpcoiss, tmast) is the best predictor of life satisfaction?

Comparison of the standardized coefficient values (beta) indicates that the tmast (beta=.342) is a stronger predictor of life satisfaction than tpcoiss (beta=.195).

(c) Do both variables make a statistically significant contribution to the prediction of life satisfaction?

The probability values (shown in the Sig. Column) are both less than .05, indicating that both predictors make a significant contribution to the equation.

4.9 Follow the instructions in the *SPSS Survival Manual* to perform a hierarchical multiple regression, this time using life satisfaction as the dependent variable.

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
tlifesat total life satisfaction	22.38	6.770	436
tmarlow total social desirability	5.30	2.042	433
age	37.44	13.202	439
tmast total mastery	21.764	3.9696	436
tpcoiss total PCOISS	60.63	11.985	430

Correlations

		tlifesat total life satisfaction	tmarlow total social desirability	age	tmast total mastery	tpcoiss total PCOISS
Pearson Correlation	tlifesat total life satisfaction	1.000	.108	.059	.444	.373
	tmarlow total social desirability	.108	1.000	.268	.154	.295
	age	.059	.268	1.000	-.036	.248
	tmast total mastery	.444	.154	-.036	1.000	.521
	tpcoiss total PCOISS	.373	.295	.248	.521	1.000
Sig. (1-tailed)	tlifesat total life satisfaction	.	.012	.111	.000	.000
	tmarlow total social desirability	.012	.	.000	.001	.000
	age	.111	.000	.	.226	.000
	tmast total mastery	.000	.001	.226	.	.000
	tpcoiss total PCOISS	.000	.000	.000	.000	.
N	tlifesat total life satisfaction	436	431	436	436	429
	tmarlow total social desirability	431	433	433	431	427
	age	436	433	439	436	430
	tmast total mastery	436	431	436	436	429
	tpcoiss total PCOISS	429	427	430	429	430

Variables Entered/Removed [b]

Model	Variables Entered	Variables Removed	Method
1	age, tmarlow total social desirability [a]	.	Enter
2	tmast total mastery, tpcoiss total PCOISS [a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: tlifesat total life satisfaction

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.113[a]	.013	.008	6.742	.013	2.724	2	424	.067
2	.475[b]	.225	.218	5.986	.213	57.911	2	422	.000

a. Predictors: (Constant), age, tmarlow total social desirability

b. Predictors: (Constant), age, tmarlow total social desirability, tmast total mastery, tpcoiss total PCOISS

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	247.684	2	123.842	2.724	.067[a]
	Residual	19274.435	424	45.459		
	Total	19522.118	426			
2	Regression	4398.524	4	1099.631	30.683	.000[b]
	Residual	15123.595	422	35.838		
	Total	19522.118	426			

a. Predictors: (Constant), age, tmarlow total social desirability

b. Predictors: (Constant), age, tmarlow total social desirability, tmast total mastery, tpcoiss total PCOISS

c. Dependent Variable: tlifesat total life satisfaction

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	20.011	1.163		17.204	.000					
	tmarlow total social desirability	.331	.166	.100	1.994	.047	.108	.096	.096	.928	1.077
	age	.016	.026	.032	.636	.525	.059	.031	.031	.928	1.077
	tmast total mastery										
	tpcoiss total PCOISS										
2	(Constant)	2.652	1.917		1.384	.167					
	tmarlow total social desirability	-.026	.152	-.008	-.170	.865	.108	-.008	-.007	.871	1.148
	age	.014	.024	.027	.579	.563	.059	.028	.025	.860	1.163
	tmast total mastery	.594	.087	.348	6.795	.000	.444	.314	.291	.699	1.432
	tpcoiss total PCOISS	.106	.030	.188	3.489	.001	.373	.167	.149	.635	1.574

a. Dependent Variable: tlifesat total life satisfaction

Factor analysis

4.10 There is some controversy in the literature concerning the underlying factor structure of one of the scales included in the questionnaire presented in the appendix of the *SPSS Survival Manual*. The Optimism scale was originally designed as a one-dimension (factor) scale which included some positively worded items and some negatively worded items. Recent studies suggest that it may in fact consist of two factors representing optimism and pessimism.

Conduct a factor analysis using the instructions presented in Chapter 15 to explore the factor structure of the optimism scale (op1 to op6).

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.808
Bartlett's Test of Sphericity	Approx. Chi-Square	720.478
	df	15
	Sig.	.000

Communalities

	Initial	Extraction
op1	1.000	.357
op2	1.000	.538
op3	1.000	.424
op4	1.000	.641
op5	1.000	.537
op6	1.000	.501

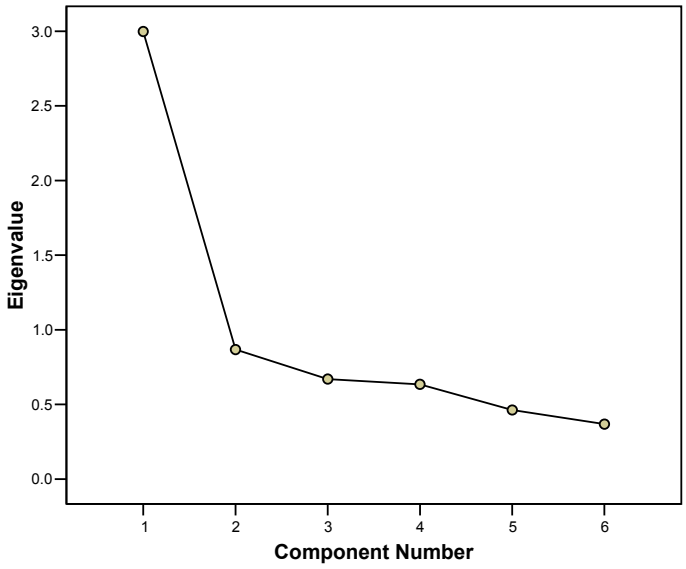
Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.998	49.966	49.966	2.998	49.966	49.966
2	.867	14.458	64.424			
3	.670	11.161	75.584			
4	.634	10.573	86.157			
5	.463	7.709	93.866			
6	.368	6.134	100.000			

Extraction Method: Principal Component Analysis.

Scree Plot



Component Matrix [a]

	Component
	1
op4	.801
op2	.733
op5	.733
op6	.708
op3	.651
op1	.597

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Rotated Component Matrix [a]

Dummy category	

a. Only one component was extracted. The solution cannot be rotated.