Questions and answers for Chapter 10

1. You want to find out what factors predict achievement in English. Develop a model that you think can explain this.

   As usual many alternative predictors are possible and present in our data. We will look at the predictors gender and 'the teachers think I'm good at English' in this example.

2. Calculate your model using ANOVA (remember, this limits what variables you can use). Does your model predict grades in English? How strongly does it predict English grades?

Tests of Between-Subjects Effects

Dependent Variable:school grades English

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9262.752[a]	7	1323.250	14.198	.000
Intercept	2155069.676	1	2155069.676	23123.153	.000
gender	133.495	1	133.495	1.432	.232
engsc4	8696.315	3	2898.772	31.103	.000
gender * engsc4	31.995	3	10.665	.114	.952
Error	52564.600	564	93.200		
Total	3576207.854	572			
Corrected Total	61827.352	571			

a. R Squared = .150 (Adjusted R Squared = .139)

When we look at the 'test of between subjects effect box', we can see that our model as a whole (the row labelled 'corrected model') is statistically significant, with an adjusted R squared of .15 (this is given just below the box).

3. Calculate your model using ANOVA. Which individual variables predict English grades? How strong is their effect?

Tests of Between-Subjects Effects

Dependent Variable:school grades English

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9262.752[a]	7	1323.250	14.198	.000	.150
Intercept	2155069.676	1	2155069.676	23123.153	.000	.976
gender	133.495	1	133.495	1.432	.232	.003
engsc4	8696.315	3	2898.772	31.103	.000	.142
gender * engsc4	31.995	3	10.665	.114	.952	.001
Error	52564.600	564	93.200			
Total	3576207.854	572				
Corrected Total	61827.352	571				

a. R Squared = .150 (Adjusted R Squared = .139)

Again looking at the 'Tests of Between-Subject Effects' box, we can see that 'the teachers think I'm good at English'(engsc4) is significant, while gender is not. The effect size measure, Partial Eta Squared, shows a modes effect size for 'the teachers think I'm good at English' at .14, while the effect size for gender is very weak at .003.

4. Calculate your model using ANOVA. Are there any interaction effects? What do they mean?

Tests of Between-Subjects Effects

Dependent Variable:school grades English

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9262.752[a]	7	1323.250	14.198	.000	.150
Intercept	2155069.676	1	2155069.676	23123.153	.000	.976
gender	133.495	1	133.495	1.432	.232	.003
engsc4	8696.315	3	2898.772	31.103	.000	.142
gender * engsc4	31.995	3	10.665	.114	.952	.001
Error	52564.600	564	93.200			
Total	3576207.854	572				
Corrected Total	61827.352	571				

a. R Squared = .150 (Adjusted R Squared = .139)

There are no significant interaction effects in this model, as the gender+engsc4 interaction term is not significant. This means that there is no relationship between English grades and 'the teachers think I'm good at English' that is dependent on gender or differs between boys and girls.

5. Calculate your model using ANOVA. What do the post hoc tests tell you?

Multiple Comparisons

school grades English

Scheffe

(I) teachers think I'm good at English	(J) teachers think I'm good at English	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
disagree strongly	disagree	-3.0054	1.60380	.320	-7.5024	1.4916
	agree	-9.7202*	1.55827	.000	-14.0895	-5.3509
	agree strongly	-11.8352*	1.79892	.000	-16.8793	-6.7912
disagree	disagree strongly	3.0054	1.60380	.320	-1.4916	7.5024
	agree	-6.7148*	.92637	.000	-9.3123	-4.1173
	agree strongly	-8.8299*	1.29076	.000	-12.4491	-5.2106
agree	disagree strongly	9.7202*	1.55827	.000	5.3509	14.0895
	disagree	6.7148*	.92637	.000	4.1173	9.3123
	agree strongly	-2.1151	1.23373	.402	-5.5744	1.3442
agree strongly	disagree strongly	11.8352*	1.79892	.000	6.7912	16.8793
	disagree	8.8299*	1.29076	.000	5.2106	12.4491
	agree	2.1151	1.23373	.402	-1.3442	5.5744

Based on observed means.

The error term is Mean Square(Error) = 93.200.

*. The mean difference is significant at the .05 level.

school grades English

Scheffe[a,b,c]

teachers think I'm good at English	N	Subset	
		1	2
disagree strongly	45	71.3160	
disagree	186	74.3214	
agree	261		81.0362
agree strongly	80		83.1513
Sig.		.221	.535

Means for groups in homogeneous subsets are displayed.

Based on observed means.

The error term is Mean Square(Error) = 93.200.

a. Uses Harmonic Mean Sample Size = 91.054.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Looking at the Scheffe post hoc tests output above, it is clear that the disagree strongly and disagree categories differ significantly from the agree and agree strongly categories. The agree and agree strongly, and the disagree and disagree strongly categories don't differ from each other. Pupils in the disagree and disagree strongly categories have significantly lower grades in English than those in the agree and agree strongly categories.

This is confirmed by the homogeneous subsets, which divide the categories into two groups, agree with agree strongly and disagree with disagree strongly.

6. Can you think of any arguments why you would want to use ANOVA rather than regression?

   There are two main reasons to use ANOVA instead of regression. The first is its neat conceptual fit with experimental and quasi experimental research designs, which makes the method very suitable for analysis of data from these types of research. The second is the use of post-hoc tests which are a useful way of testing exactly which categories of the independent variable are related to different outcomes. The automatic generation of interaction terms is useful, but these can also easily be included in regression models.

7. You want to find out what factors predict achievement in both English and maths. Develop a model that you think can explain this.

   Two variables that may be predictors of both maths and English achievement are 'gender' and 'I'm among the best in my class in all subjects'.

8. Calculate your model using MANOVA. Which individual variables predict combined maths and English grades? How strong is their effect?

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.986	19751.200[a]	2.000	562.000	.000	.986
	Wilks' Lambda	.014	19751.200[a]	2.000	562.000	.000	.986
	Hotelling's Trace	70.289	19751.200[a]	2.000	562.000	.000	.986
	Roy's Largest Root	70.289	19751.200[a]	2.000	562.000	.000	.986
gender	Pillai's Trace	.004	1.240[a]	2.000	562.000	.290	.004
	Wilks' Lambda	.996	1.240[a]	2.000	562.000	.290	.004
	Hotelling's Trace	.004	1.240[a]	2.000	562.000	.290	.004
	Roy's Largest Root	.004	1.240[a]	2.000	562.000	.290	.004
schsc3	Pillai's Trace	.222	23.402	6.000	1126.000	.000	.111
	Wilks' Lambda	.780	24.826[a]	6.000	1124.000	.000	.117
	Hotelling's Trace	.281	26.255	6.000	1122.000	.000	.123
	Roy's Largest Root	.274	51.462[b]	3.000	563.000	.000	.215
gender * schsc3	Pillai's Trace	.015	1.442	6.000	1126.000	.195	.008
	Wilks' Lambda	.985	1.440[a]	6.000	1124.000	.196	.008
	Hotelling's Trace	.015	1.438	6.000	1122.000	.197	.008
	Roy's Largest Root	.011	2.024[b]	3.000	563.000	.109	.011

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept + gender + schsc3 + gender * schsc3

Looking at the 'multivariate tests' box gives us the statistics for the new combined dependent variable (the linear combination of school grades maths and school grades English). Looking at Wilks' Lambda as the independent variables form more than two groups, we can see that 'I am among the best in my class for all subjects (schsc3)' is significantly related to the combined outcome variable, while gender isn't. The Partial Eta Squared effect size measure is .117 for 'I am among the best in my class for all subjects (schsc3)', and only just above 0 for gender. 'I am among the best in my class for all subjects (schsc3)' therefore has a weak to modest relationship with the combined outcome measure.

9. Calculate your model using MANOVA. Are there any interaction effects? What do they mean?

The interaction between gender and 'I am among the best in my class for all subjects (schsc3)' is not significant, suggesting that the relationship between the outcome variable and 'I am among the best in my class for all subjects (schsc3)' doesn't differ depending on pupil gender.