SPSS

statstutor

community project

encouraging academics to share statistics support resources

stcp-marshall-ANOVA2

The following resources are associated:
Checking normality in SPSS, ANOVA in SPSS, Interactions and the SPSS dataset 'Diet.sav'

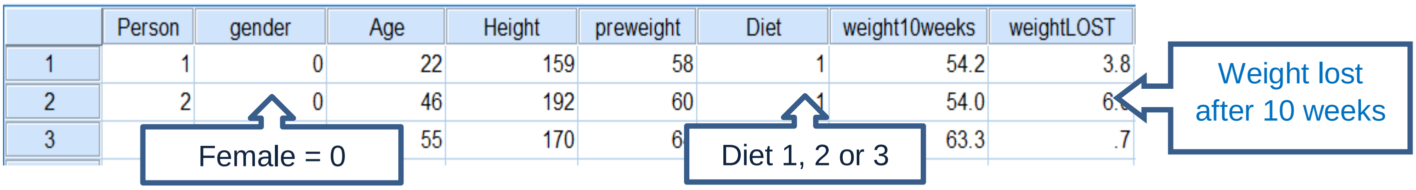
Two-way (between-groups) ANOVA

Dependent variable: Continuous (scale),

Independent variables: Two categorical (grouping factors)

Common Applications: Comparing means for combinations of two independent categorical variables (factors).

Data: The data set 'Diet.sav' contains information on 78 people who undertook one of three diets. There is background information such as age, gender and height as well as weight lost on the diet (a positive value means they lost weight). The aim of the study was to see which diet was best for losing weight but it was also thought that best diets for males and females may be different so the independent variables are diet and gender.



	Person	gender	Age	Height	preweight	Diet	weight10weeks	weightLOST
1	1	0	22	159	58	1	54.2	3.8
2	2	0	46	192	60	1	54.0	6.
3			55	170	6		63.3	.7

There are three hypotheses with a two-way ANOVA. There are the tests for the main effects (diet and gender) as well as a test for the interaction of the two between diet and gender.

Checking the assumptions for two-way ANOVA

Assumptions	How to check	What to do if the assumption is not met
Residuals should be normally distributed	Use the **Save** menu to request the standardised residuals to be added to the dataset, then use *Analyze* → *Descriptive Statistics* → *Explore* to produce a histogram of residuals	If the residuals are very skewed, the results of the ANOVA are less reliable. There is no equivalent non-parametric test in SPSS but transforming the dependent variable or a separate ANOVA by gender are options.
Homogeneity of variance (Levene's test)	Use the **Options** menu to select *Homogeneity tests* for equality of variances. If $p > 0.05$, equal variances can be assumed	If $p < 0.05$, the results of the ANOVA are less reliable. There is no equivalent test but comparing the p-values from the ANOVA with 0.01 instead of 0.05 is acceptable.

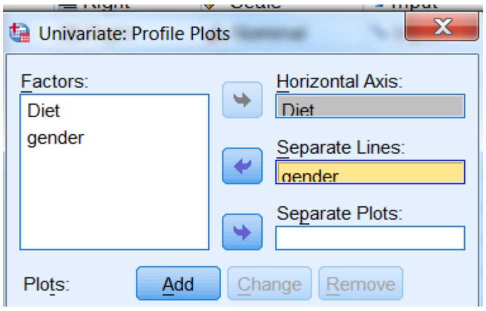
© Ellen Marshall
Sheffield Hallam University

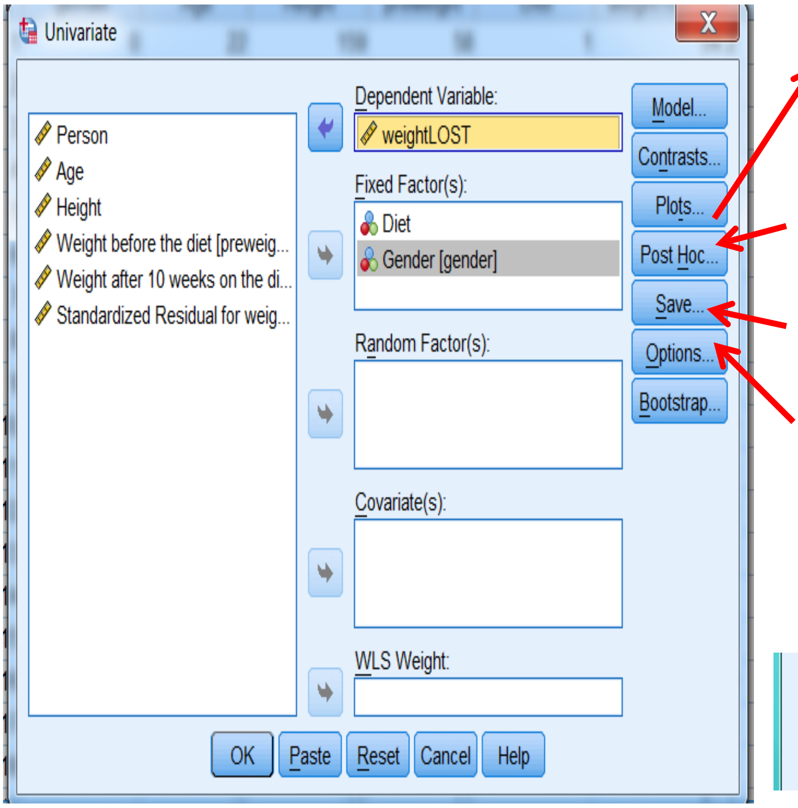
Reviewer: Alun Owen
Coventry University

Steps in SPSS

To carry out an ANOVA, select *Analyze → General Linear Model → Univariate*
Put the dependent variable (weight lost) in the *Dependent Variable* box and the independent variables (Diet and Gender) in the *Fixed Factors* box.

In the **Plots** menu, move Diet to the *Horizontal Axis* box, Gender to the *Separate Lines* box and click *Add*.

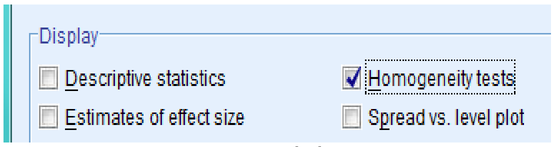




Request *Tukey's* from the **Post hoc** menu for both factors.

Ask for *standardised residuals* via the **Save** menu.

Select *Homogeneity tests* from the **Options** menu to check the assumption of equal variances for each combination of diet/ gender.



For more detailed notes on carrying out an ANOVA and the interpretation, see the **ANOVA in SPSS** resource.

The output

Checking the assumptions for this data

Check equality of variances

Levene's Test of Equality of Error Variancesa

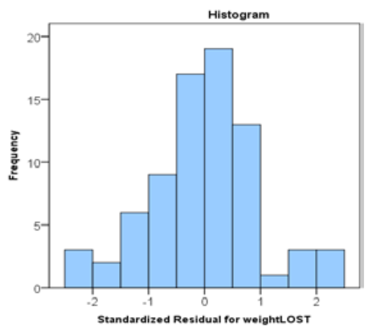
Dependent Variable: weightLOST

F	df1	df2	Sig.
.382	5	70	.860

As p > 0.05, equal variances can be assumed.

Check normality of residuals

(Using *Analyze → Descriptive Statistics → Explore* to produce the histogram). The residuals are normally distributed.



Tests of Between-Subjects Effects

Dependent Variable: weightLOST

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	94.600[a]	5	18.920	3.519	.007
Intercept	1144.438	1	1144.438	212.874	.000
Diet	49.679	2	24.840	4.620	.013
gender	.428	1	.428	.080	.779
Diet * gender	33.904	2	16.952	3.153	.049
Error	376.329	70	5.376		
Total	1654.350	76			
Corrected Total	470.929	75			

a. R Squared = .201 (Adjusted R Squared = .144)

Main effect of diet p = 0.013

Main effect of gender p = 0.779

Interaction effect p = 0.049

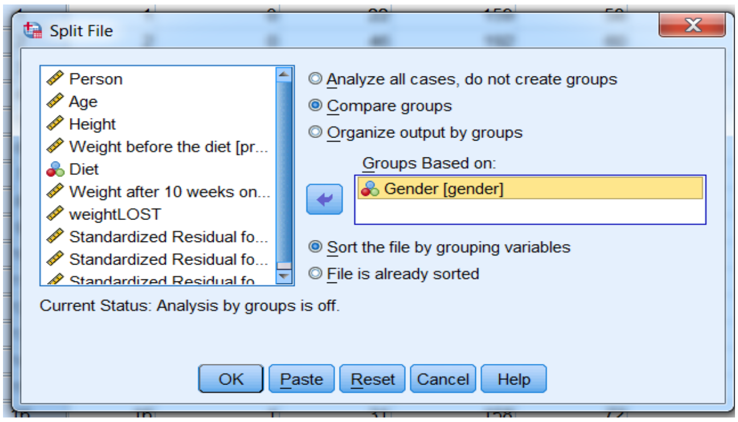
The results of the two-way ANOVA and post hoc tests are reported in the same way as one way ANOVA for the main effects and the interaction e.g. there was a statistically significant interaction between the effects of Diet and Gender on weight loss [$F(2, 70)=3.153$, $p = 0.049$]. Since the interaction effect is significant ($p = 0.049$), interpreting the main effects can be misleading.

To easiest way to interpret the interaction is to use the plot from the output known as a means or interaction plot which shows the means for each combination of diet and gender (see the **Interactions** resource for more details on interactions).

The means (or interaction) plot clearly shows a difference between males and females in the way that diet affects weight lost, since the lines are not parallel. The differences between the mean weight lost on the diets is much bigger for females.

As the main effect of diet combines the male and female data, the overall means for each diet understate the diet differences for females. Some people just use the interaction plot to describe the combined effect of diet and gender but others prefer to carry out one way ANOVA's for each group of one factor.
For this example, it makes sense to look at the differences between the diets by gender.

To carry out separate ANOVA's by gender, use *Data → Split File*



Select '*Compare groups*' and move Gender to the '*Groups Based on*' box. After clicking '*OK*', all analyses and charts will appear separately for males and females until the split is cancelled by going back to this box and selecting '*Analyze all cases, do not create groups*'.

Run a one-way ANOVA for Diet: *Analyze → General Linear Model → Univariate*

The results appear separately for males and females.

Tests of Between-Subjects Effects

Dependent Variable: WeightLOST

Gender	Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Female	Corrected Model	92.320[b]	2	46.160	10.640	.000
	Intercept	635.277	1	635.277	146.438	.000
	Diet	**92.320**	**2**	**46.160**	**10.640**	**.000**
	Error	173.528	40	4.338		
	Total	917.540	43			
	Corrected Total	265.848	42			
Male	Corrected Model	2.002[c]	2	1.001	.148	.863
	Intercept	524.420	1	524.420	77.577	.000
	Diet	**2.002**	**2**	**1.001**	**.148**	**.863**
	Error	202.801	30	6.760		
	Total	736.810	33			
	Corrected Total	204.802	32			

There was a difference between the mean weight lost on the 3 diets for females ($F(2,40)=10.64$, $p < 0.001$) but not for males ($F(2,30)=0.148$, $p = 0.863$). Only the post hoc tests for females should be interpreted.

You should also report the mean weight lost for each diet for females.

Tukey's post hoc tests were carried out for females. Diet 3 was significantly different to diet 1 ($p = 0.002$) and diet 2 ($p < 0.001$) but there is no evidence to suggest that diets 1 and 2 differ ($p = 0.841$). For females, the mean diet lost on diet 3 was 5.88kg compared to only 3.05kg and 2.61kg on diets 1 and 2 respectively.

Gender	(I) Diet	(J) Diet	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Female	1	2	.443	.7872	.841	-1.473	2.359
		3	-2.830*	.7740	.002	-4.714	-.946
	2	1	-.443	.7872	.841	-2.359	1.473
		3	-3.273*	.7740	.000	-5.157	-1.389

Normality checks and Levene's test were carried out and the assumptions were met.