One-way & two-way ANOVA

Dataset

Our dataset consists of the following variables:

Respondent ID	Preference for blues music	Educational level	Autonomy	Unemployment
Unique number	Value in range 1-25	1 = low 2 = medium 3 = high	Value in range 1-15	1 = yes 2 = no

respnr	blues	educlevel_rec	autonomy	unemp
1012	16,00	1	8	2
1013	.	1	9	1
1016	10,00	1	6	2
1023	.	1	10	2
1024	.	1	9	2
1028	18,00	1	8	2

One-way ANOVA

Question

We're interested in this question:
"What is the effect of educational level on the dimension autonomy?"

To find it out, we can use one-way ANOVA. It is one-way because we have

- one independent categorical variable (Educational level)
- and one dependent continuous variable (Autonomy)

Analysis

In ANOVA,

H0: Group means are equal
H1: At least one group mean is different from others

In the SPSS' ANOVA output we see that $p < 0.05$, which means we need to reject the null hypothesis => at least one group mean is different.

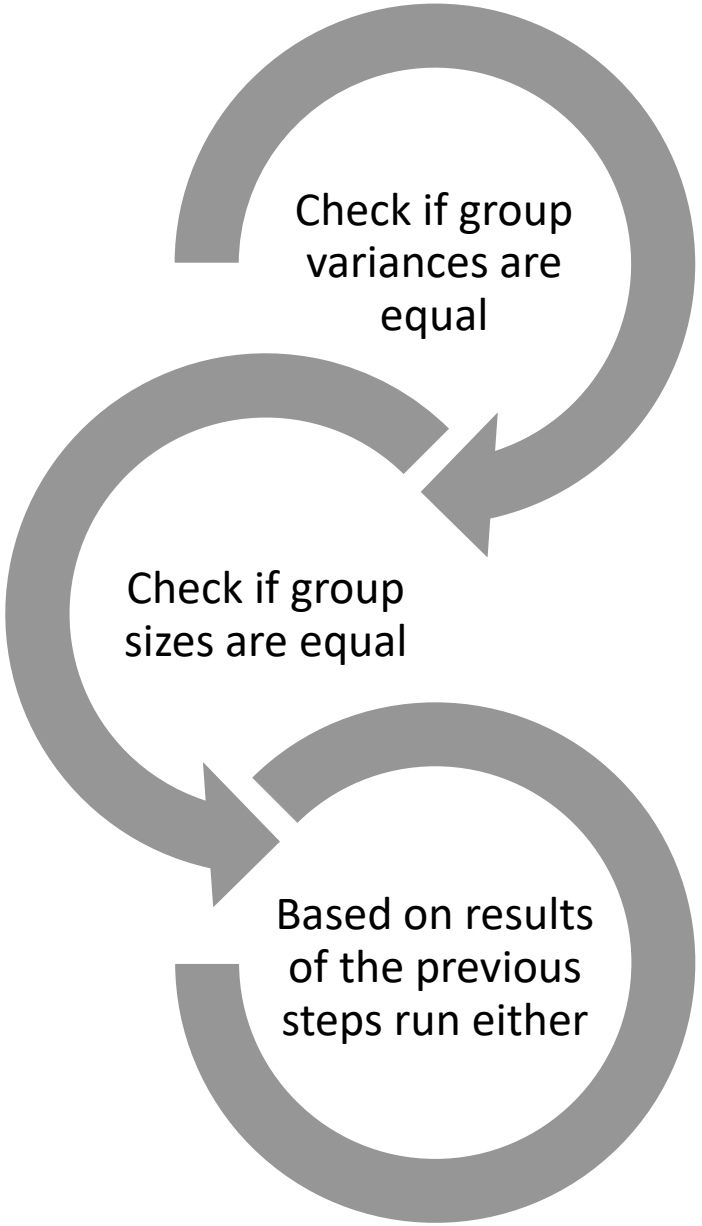
ANOVA

Score on autonomy dimension

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,907	2	35,954	6,397	,002
Within Groups	5232,513	931	5,620		
Total	5304,421	933			

Analysis

To learn which mean(s) exactly are different, we need to conduct Post Hoc analysis. Steps are as follows:



- **Tukey** (both variances & size are equal)
- **Hochberg** (variances equal, size unequal)
- **Games-Howell** (variances unequal, size equal)

Analysis

Equality of group variances can be checked by means of Homogeneity of variance test. For this test,

H0: Group variances are equal
H1: Group variances are different

In the test's output we see that p = 0.544, which means H0 must be accepted => Group variances are equal.

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Score on autonomy dimension	Based on Mean	,610	2	931	,544
	Based on Median	,264	2	931	,768
	Based on Median and with adjusted df	,264	2	923,106	,768
	Based on trimmed mean	,579	2	931	,561

Analysis

To check equality of group sizes we just compare the group sizes.

Descriptives

Score on autonomy dimension

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
low	370	8,50	2,424	,126	8,25	8,75	3	15
medium	353	8,92	2,341	,125	8,68	9,17	3	15
high	211	9,20	2,324	,160	8,89	9,52	3	15
Total	934	8,82	2,384	,078	8,67	8,97	3	15

Group sizes are unequal as the ratio between the smallest and the largest group sizes is > 1.5 (370/211 = 1.75).

Analysis

Summing it up, variances are equal but sizes are unequal - therefore, we should look at the Hochberg's test.

Result

- There's a significant difference in effect on Autonomy between low and high educational levels.
- High educational level has a higher mean score on Autonomy than low.
- The other group differences do not differ.

Results of Hochberg's test

(I) Educational level (recoded)	(J) Educational level (recoded)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
low	medium	-,418	,176	,053	-,84	,00
	high	-,701*	,205	,002	-1,19	-,21
medium	low	,418	,176	,053	,00	,84
	high	-,283	,206	,429	-,78	,21
high	low	,701*	,205	,002	,21	1,19
	medium	,283	,206	,429	-,21	,78

*. The mean difference is significant at the 0.05 level.

Two-way ANOVA

Question

We're interested in this question:

"What are the effects of educational level and unemployment on the preference for blues music?"

This time we should use two-way ANOVA because we have

- **two** independent categorical variables (Educational level & Unemployment)
- and one dependent interval variable (Blues)

Analysis

Educational level has more than two categories, so we don't know which means will be different (if any).

To learn that, we again need to conduct Post Hoc analysis.

Test on equality of variances

$p > 0.05$
=> variances are equal

		Levene Statistic	df1	df2	Sig.
Preference for blues music	Based on Mean	,385	5	764	,859
	Based on Median	,398	5	764	,850
	Based on Median and with adjusted df	,398	5	758,02	,850
	Based on trimmed mean	,394	5	764	,853

Test on equality of group sizes

303/188 = 1.61 > 1.5
=> sizes are unequal

Unemployed	Educational level (recoded)	Mean	Std. Deviation	N
yes	low	16,467	4,74893	15
	medium	12,733	3,84460	15
	high	15,286	3,49830	7
	Total	14,730	4,41996	37
no	low	15,133	4,04366	264
	medium	15,965	4,20703	288
	high	15,160	4,26116	181
	Total	15,467	4,17618	733
Total	low	15,204	4,08602	279
	medium	15,805	4,24245	303
	high	15,165	4,22742	188
	Total	15,431	4,18817	770

Analysis

Summing it up, variances are equal but sizes are unequal - therefore, we should look at the Hochberg's test.

Analysis

Hochberg's test says the effect of Educational level on Preference of blues music is not significant.

(I) Educational level (recoded)	(J) Educational level (recoded)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
low	medium	-,6010	,34547	,227	-1,4276	,2257
	high	,0394	,39287	,999	-,9007	,9795
medium	low	,6010	,34547	,227	-,2257	1,4276
	high	,6404	,38655	,266	-,2846	1,5653
high	low	-,0394	,39287	,999	-,9795	,9007
	medium	-,6404	,38655	,266	-1,5653	,2846

Based on observed means.
The error term is Mean Square(Error) = 17,336.

Analysis

From the resultant ANOVA table we can see that:

- the interaction effect between Educational level & Unemployment is significant ($p < 0.05$)
- the main effects are not significant ($p = 0.176$ and $p = 0.429$), which also corresponds to the insignificance result from the Hochberg's test.

Tests of Between-Subjects Effects

Dependent Variable: Preference for blues music

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	244,390[a]	5	48,878	2,820	,016	,018
Intercept	28495,277	1	28495,3	1644	,000	,683
unemp	10,870	1	10,870	,627	,429	,001
educlevel_rec	60,335	2	30,168	1,740	,176	,005
unemp * educlevel_rec	154,641	2	77,320	4,460	,012	,012
Error	13244,461	764	17,336			
Total	196842,000	770				
Corrected Total	13488,852	769				

a. R Squared = ,018 (Adjusted R Squared = ,012)

Result

Since the interaction effect is significant, we can look at a plot that shows dependence between all three variables.

In the plot **we observe**:

- Employed people have a high preference for blues regardless of their educational level
- Unemployed with medium education have the lowest preference for blues among all groups
- Unemployed with low education prefer blues more than anyone
- All people with high education have an average preference for blues