DEMONSTRATION OF TWO-WAY ANOVA

Dr Kyaw Oo

Scenario

- A researcher was interested in whether an employee's **salary** in a job
- was influenced by their level of **education** and **gender**.
- They recruited a random sample of participants to their study and asked them about their salary.
- The researcher then divided the participants by gender (**Male/Female**) and then again
- by level of education (**<=10/11-14/>14**).
- Therefore, the dependent variable was "**salary**", and the two independent variables were "**gender**" and "**education**".

- In SPSS Statistics, we separated the individuals into their appropriate groups by using **two columns** representing the two independent variables, and labelled them **GENDER** and **Edu_Level**.
- For Gender, we coded "males" as **1** and "females" as **2**, and
- for Edu_Level, we coded "<=10" as **1**, "11-14" as **2** and ">14" as **3**.
- The participants' interest in **SALARY** – the dependent variable – was entered under the variable name, salary.

- The setup for this example can be seen: DATA SET

- The 14 steps below show you how to analyse your data using a two-way ANOVA in SPSS Statistics Assumptions, have not been violated.
- At the end of these 14 steps, we show you how to interpret the results from this test.

You will be presented with the Analyze→General Linear Model→**Univariate** dialogue box, as shown:

- Transfer the dependent variable, **salary**, into the Dependent Variable: box, and
- Transfer both independent variables, Gender and Edu_Level, into the Fixed Factor(s): box. (by using the arrow button.)

- Note: For this analysis, you will not need to worry about the Random Factor(s):, Covariate(s): or WLS Weight: boxes.

- Click on the **Plot button**.
- You will be presented with the **Univariate:**
- **Profile Plots** dialogue box, as shown:

- Transfer the independent variable, Edu_Level, from the Factors: box into the Horizontal Axis: box, and
- Transfer the other independent variable, Gender, into the Separate Lines: box.
- You will be presented with the following screen:

- Note: It can help to put the independent variable with the greater number of groups in the Horizontal Axis: box.

- Click the Add button.
- You will see that "**Edu_Level*Gender**" has been added to the Plots: box, as shown:

- Click the Continue button.
- This will return you to the **Univariate** dialogue box.

- Click Post-hoc button.
- You will be presented with the **Univariate: Post Hoc Multiple Comparisons for Observed Means** dialogue box, as shown:

- Transfer Edu_Level from the Factor(s): box to the Post Hoc Tests for: box.
- This will make the –Equal Variances Assumed– area become active (lose the "grey sheen") and present you with some choices for which post hoc test to use.
- For this example, we are going to select Tukey, which is a good, all-round post hoc test.

You will finish up with the following screen:

Note: You only need to transfer independent variables that have more than two groups into the Post Hoc Tests for: box. This is why we do not transfer Gender.

Click Continue button to return to the **Univariate** dialogue box.

- Click Option button.
- This will present you with the **Univariate: Options** dialogue box, as shown:

- Transfer **Gender**, **Edu_Level** and **Gender*Edu_Level** from the Factor(s) and Factor Interactions: box into the Display Means for: box.
- In the –Display– area, tick the Descriptive Statistics option. You will presented with the following screen:

Click Continue button to return to the **Univariate** dialogue box.

Click OK button to return to the **Univariate** dialogue box.

Descriptive statistics

You can find appropriate descriptive statistics for when you report the results of your two-way ANOVA in the aptly named "**Descriptive Statistics**" table, as shown below:

Descriptive Statistics

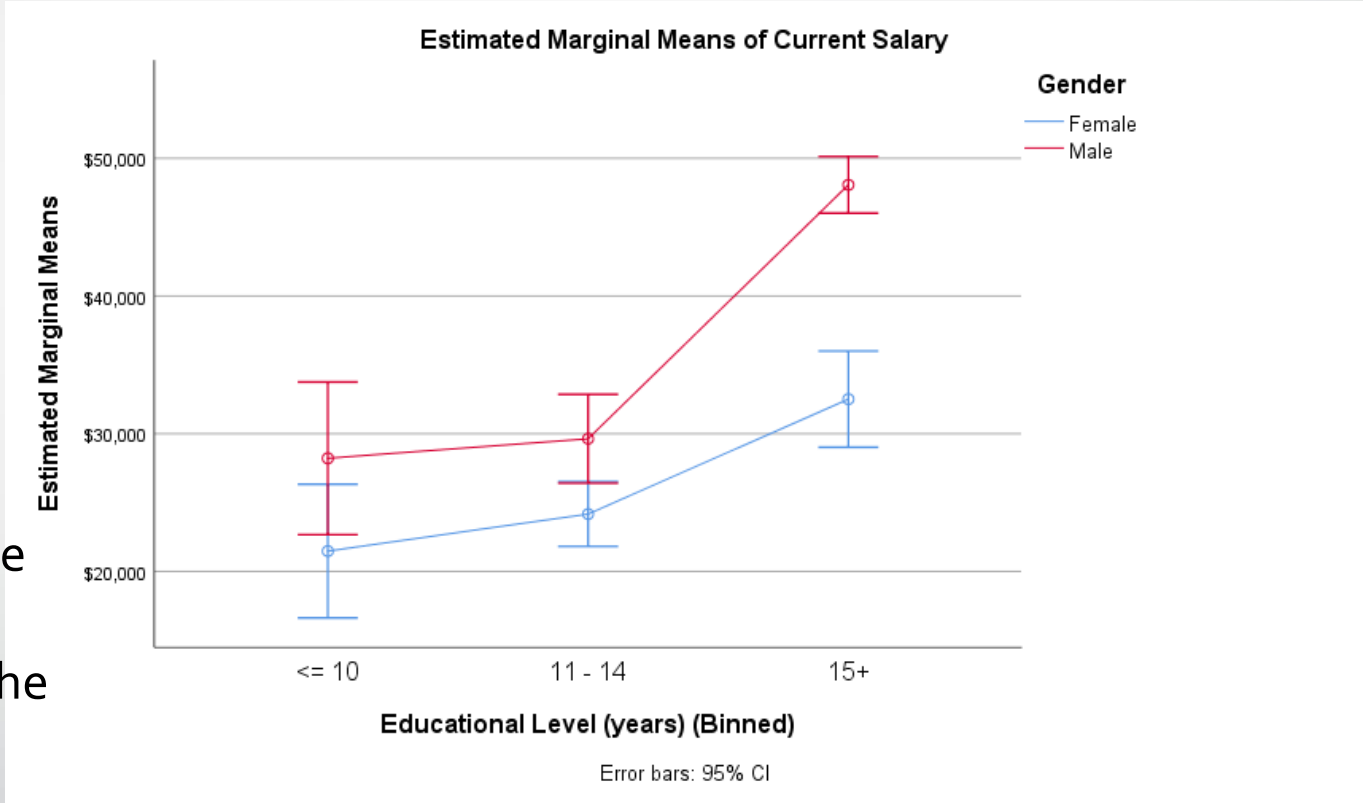
Dependent Variable: Current Salary

Gender	Educational Level (years) (Binned)	Mean	Std. Deviation	N
Female	<= 10	$21,475.00	$4,204.077	30
	11 - 14	$24,165.70	$4,486.400	128
	15+	$32,507.50	$10,010.459	58
	Total	$26,031.92	$7,558.021	216
Male	<= 10	$28,213.04	$3,676.158	23
	11 - 14	$29,633.82	$5,880.984	68
	15+	$48,071.74	$21,149.925	167
	Total	$41,441.78	$19,499.214	258
Total	<= 10	$24,399.06	$5,190.482	53
	11 - 14	$26,062.81	$5,639.289	196
	15+	$44,059.62	$20,088.569	225
	Total	$34,419.57	$17,075.661	474

- This table provides the mean and standard deviation for each combination of the groups of the independent variables (what is sometimes referred to as each "cell" of the design).
- In addition, the table provides "Total" rows, which allows means and standard deviations for groups only split by one independent variable, or none at all, to be known.
- This might be more useful if you do not have a statistically significant interaction.

Plot of the results

The plot of the mean "interest in politics" score for each combination of groups of "Gender" and "Edu_level" are plotted in a line graph, as shown below:



- An interaction effect can usually be seen as a set of non-parallel lines.
- You can see from this graph that the lines do not appear to be parallel.
- You might expect there to be a statistically significant interaction, which we can confirm in the next section.

Statistical significance of the two-way ANOVA

The actual result of the two-way ANOVA – namely, whether either of the two independent variables or their interaction are statistically significant – is shown in the **Tests of Between-Subjects Effects** table, as shown below:

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.227E+10[a]	5	1.045E+10	57.117	.000
Intercept	2.764E+11	1	2.764E+11	1510.502	.000
gender	6292420377	1	6292420377	34.382	.000
educlevel	1.907E+10	2	9532563870	52.087	.000
gender * educlevel	2383227926	2	1191613963	6.511	.002
Error	8.565E+10	468	183013449.2		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .379 (Adjusted R Squared = .372)

These rows inform us whether our independent variables (the "Gender" and "Edu_Level" rows) and their interaction (the "Gender*Edu_Level" row) have a statistically significant effect on the dependent variable, "salary".

It is important to first look at the "Gender*Edu_Level" interaction as this will determine how you can interpret your results.

You can see from the "**Sig.**" column that we have a statistically significant interaction at the $p = .002$ level.

You may also wish to report the results of "Gender" and "Edu_Level", but again, these need to be interpreted in the context of the interaction result.

We can see from the table above that there was statistically significant difference in mean salary males and females ($p = .000$),

and there were statistically significant differences between educational levels ($p < .001$).

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.227E+10[a]	5	1.045E+10	57.117	.000
Intercept	2.764E+11	1	2.764E+11	1510.502	.000
gender	6292420377	1	6292420377	34.382	.000
educlevel	1.907E+10	2	9532563870	52.087	.000
gender * educlevel	2383227926	2	1191613963	6.511	.002
Error	8.565E+10	468	183013449.2		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .379 (Adjusted R Squared = .372)

- These rows inform us whether our independent variables (the "Gender" and "Edu_Level" rows) and their interaction (the "Gender*Edu_Level" row) have a statistically significant effect on the dependent variable, "salary".
- Important to first look at the "Gender*Edu_Level" interaction as this will determine how you can interpret your results.
- You can see from the "**Sig.**" column that we have a statistically significant interaction at the $p = .002$ level.
- Report the results of "Gender" and "Edu_Level", but again, these need to be interpreted in the context of the interaction result.
- We can see from the table above that there was statistically significant difference in mean salary males and females ($p = .000$),
- and there were statistically significant differences between educational levels ($p < .001$).

Multiple Comparisons

Dependent Variable: Current Salary

Bonferroni

(I) Educational Level (years) (Binned)	(J) Educational Level (years) (Binned)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
<= 10	11 - 14	-$1,663.75	$2,094.475	1.000	-$6,695.97	$3,368.47
	15+	-$19,660.57*	$2,065.545	.000	-$24,623.28	-$14,697.85
11 - 14	<= 10	$1,663.75	$2,094.475	1.000	-$3,368.47	$6,695.97
	15+	-$17,996.82*	$1,321.792	.000	-$21,172.58	-$14,821.05
15+	<= 10	$19,660.57*	$2,065.545	.000	$14,697.85	$24,623.28
	11 - 14	$17,996.82*	$1,321.792	.000	$14,821.05	$21,172.58

Based on observed means.
The error term is Mean Square(Error) = 183013449.227.

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

Multiple Comparisons Table

If you do not have a statistically significant interaction, you might interpret the Bonferroni post hoc test results for the different levels of education, which can be found in the **Multiple Comparisons** table, as shown:

There is some repetition of the results, but regardless of which row we choose to read from, we are interested in the differences between (1) <=10 and 11-14, (2) 11-14 and >14, and (3) <=10 and >14.

From the results, we can see that there is a statistically significant difference between all three different educational levels ($p < .0005$) except (<=10 and 11-14).

Reporting the results of a two-way ANOVA

Emphasize the results from the interaction first before you mention the main effects. For example, you might report the result as:

A two-way ANOVA was conducted that examined the effect of gender and education level on salary.
There was a statistically significant interaction between the effects of gender and education level on salary, F (2, 468) = 6.511, p = .002.

If you had a statistically significant interaction term and carried out the procedure for simple main effects in SPSS Statistics, you would also report these results. Briefly, you might report these as:

Simple main effects analysis showed that males were significantly more salary than females when educated to >14 level (p = .002), but there were no differences between gender when educated to <=10(p = 1) or 11-14 level (p = 1).

Reference: https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics-2.php

How could you decide if salary increment is associating with gender and job category of the employees in the dataset?

Thank.