LECTURE 01: ANALYSIS OF VARIANCE (ANOVA)

This lecture presents an overview of Analysis of Variance (ANOVA) as a key statistical method for comparing group means across multiple conditions to identify statistically significant differences. ANOVA is widely applicable in fields like educational psychology, applied linguistics, didactics, and sociolinguistics, where it enables researchers to assess complex variables simultaneously and determine whether differences in group means are meaningful or due to random chance. Key assumptions for ANOVA, such as normality, homogeneity of variance, and independence, are discussed to ensure accurate results. The document also covers one-way and two-way ANOVA, including examples and a step-by-step guide for conducting a two-way ANOVA in SPSS, providing practical insights on setting up, analyzing, and interpreting results, especially in studies with multiple variables and interactions.

By the end of this lecture, you will be familiar with setting up and interpreting ANOVA results and will understand its critical role in educational and social science research, especially when handling multifactorial data with multiple group comparisons and potential interactions.

1. DEFINITION

Analysis of Variance (ANOVA) is a statistical test used to evaluate whether there are significant differences between the mean scores of three or more groups. By comparing group means in a single test, ANOVA allows researchers to assess if the differences observed among these groups are likely due to random variation or if they represent real, meaningful distinctions. This makes ANOVA especially useful for studies where multiple groups or conditions are involved, as it provides a structured way to understand variation within and between groups while minimizing the chance of error that would arise from running multiple independent tests.

ANOVA's versatility and ability to handle multiple variables make it a valuable tool for researchers and analysts across various fields. By comparing means and partitioning variance, ANOVA provides a robust way to understand the relationships between variables and identify significant differences among groups.

2. ASSUMPTIONS OF ANOVA

The accuracy of ANOVA depends on several key assumptions about the data. When these assumptions are met, ANOVA can provide reliable and meaningful results for comparing group means. However, if they are violated, the test results may be misleading. Here are the core assumptions for ANOVA, along with an explanation of why each one matters:

A. Normality of the Data

ANOVA assumes that the data within each group follows a normal (bell-shaped) distribution. This is particularly important when sample sizes are small, as non-normal data in small samples can lead to incorrect inferences about group differences. In larger samples, the central limit theorem suggests that even non-normally distributed data can approximate a normal distribution, reducing this issue.

Example: In applied linguistics, a researcher comparing vocabulary retention scores across different teaching methods assumes that the distribution of scores within each group (such as flashcards, contextual learning, and computer-assisted instruction) is approximately normal. If one group has a highly skewed score distribution, the ANOVA results may inaccurately reflect differences between groups.

B. Homogeneity of Variance

ANOVA also assumes that the variances (the spread or dispersion of scores) across the groups being compared are roughly equal. This is known as the assumption of *homogeneity of variance*. If one group has a much larger variance than others, it could unduly influence the ANOVA results, leading to false conclusions about group differences.

Example: In educational psychology, when comparing the effect of different study techniques on exam performance, researchers assume that each group's performance variance is similar. If students in the "self-study" group show a very wide range of scores while the "group study" and "tutor-assisted" groups show much narrower ranges, the results of the ANOVA might be biased, as differences could reflect variance rather than a true effect of the study method itself.

C. Independence of Observations

ANOVA requires that each observation is independent of the others. This means that the outcome for one participant should not influence the outcome for another. This assumption is essential for ensuring that the groups are genuinely comparable and that observed differences are not due to interactions among participants.

Example: In sociolinguistics, suppose a study examines code-switching behavior across different community groups. If participants in the same group influence each other's language use during data collection (for example, by conversing with each other), this could violate the independence assumption. To properly use ANOVA, researchers must collect data in a way that ensures each participant's responses are independently observed.

Why These Assumptions Matter?

If the assumptions of normality, homogeneity of variance, or independence are violated, ANOVA results can become unreliable:

- **Non-normal data** can inflate or deflate the calculated F-ratio, leading to Type I or Type II errors (i.e., finding significant differences that don't exist, or missing those that do).
- **Unequal variances** make it harder to detect real differences and can distort the comparison between group means, especially if the group sizes are also unequal.
- **Non-independent observations** introduce dependencies that bias the results, leading to over- or under-estimations of group differences.

3. TYPES OF ANOVA:

ANOVA is like several two-sample ***t-tests***. However, it results in fewer ***type I errors***. ANOVA groups differences by comparing each group's means and includes spreading the variance into diverse sources. Analysts use a one-way ANOVA with collected data about one independent variable and one dependent variable. A two-way ANOVA uses two independent variables. The independent variable should have at least three different groups or categories. ANOVA determines if the dependent variable changes according to the level of the independent variable.

A. One-Way ANOVA

1. Uses one independent variable or factor
2. Assesses the impact of a single categorical variable on a continuous dependent variable, identifying significant differences among group means
3. Does not account for interactions

B. Two-Way ANOVA

1. Uses two independent variables or factors
2. Used to not only understand the individual effects of two different factors but also how the combination of these two factors influences the outcome
3. Can test for interactions between factors

A one-way ANOVA evaluates the impact of a single factor on a sole response variable. It determines whether all the samples are the same. The one-way ANOVA is used to determine whether there are any statistically significant differences between the means of three or more independent groups.

A two-way ANOVA is an extension of the one-way ANOVA. With a one-way, there is one independent variable affecting a dependent variable. With a two-way ANOVA, there are two independent variables. For example, a two-way ANOVA allows a company to compare worker productivity based on two independent variables, such as salary and skill set. It's utilized to see the interaction between the two factors and test the effect of two factors simultaneously.

4. APPLICABILITY TO OUR FIELD OF INTEREST:

ANOVA is highly valuable across fields such as educational psychology, applied linguistics, didactics, and sociolinguistics, as it allows for the comparison of multiple groups in a single test. Below are examples from each field to illustrate ANOVA’s application.

A. Educational Psychology

Researchers might study the impact of different teaching styles on student motivation. For example, a study could compare traditional lecture-based teaching, interactive group activities, and project-based learning to understand their effects on students' intrinsic motivation. By applying ANOVA, researchers can determine if these teaching methods lead to statistically significant differences in motivation levels among students. Rather than conducting separate tests for each pair of teaching methods (which increases the risk of Type I error), ANOVA enables a simultaneous comparison, revealing if certain teaching styles are more effective in fostering motivation and thus informing pedagogical strategies.

B. Applied Linguistics

ANOVA can be used to examine the effectiveness of various vocabulary teaching methods, such as flashcards, contextual learning, and computer-assisted instruction. Suppose researchers want to assess which method best supports vocabulary retention among language learners. By using ANOVA, they can compare the average retention scores across all three groups at once, helping to identify which method (if any) has a statistically significant advantage in supporting long-term word retention. This type of analysis is valuable for language instructors seeking evidence-based strategies to improve vocabulary acquisition.

C. Didactics

ANOVA can help compare different instructional designs in terms of student engagement or knowledge retention. For instance, a didactics researcher might examine the impact of lecture-based, problem-based, and flipped classroom approaches on student performance in a science course. ANOVA allows for a comparison of the mean test scores across these instructional designs, helping educators determine if certain approaches lead to significantly higher student performance. This type of analysis supports the refinement of instructional methods, allowing teachers to adopt practices that are empirically shown to enhance learning outcomes.

D. Sociolinguistics

In sociolinguistics, ANOVA can be employed to investigate language use patterns across social groups. For example, a researcher may want to explore differences in code-switching frequency among monolingual, bilingual, and multilingual speakers in a community. ANOVA enables the comparison of the mean code-switching instances among these groups, highlighting whether certain linguistic profiles are associated with a greater tendency to switch languages in conversation. This can deepen our understanding of how social and linguistic factors interact, informing theories about language usage in multilingual communities and the social dynamics of communication.

5. RUNNING A TWO-WAY ANOVA ON SPSS

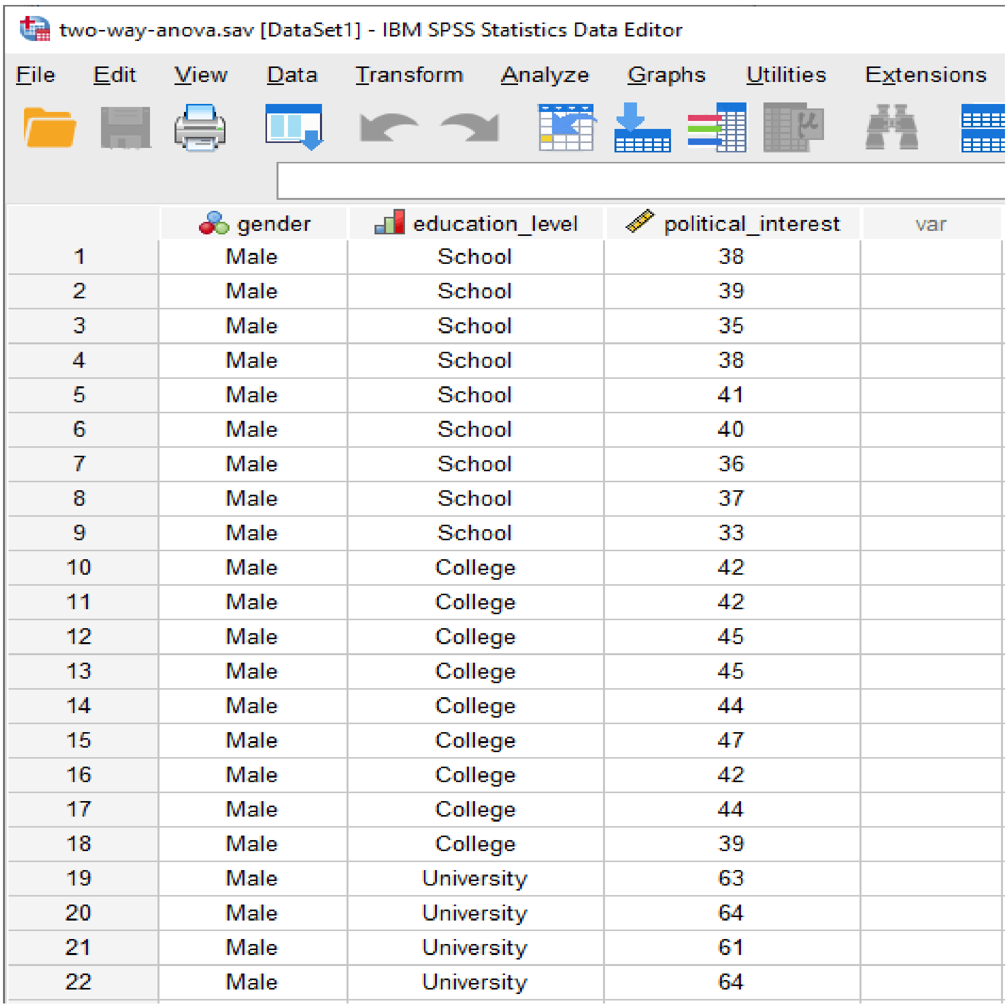
Example

A researcher was interested in whether an individual's interest in politics was influenced by their level of education and gender. They recruited a random sample of participants to their study and asked them about their interest in politics, which they scored from 0 to 100, with higher scores indicating a greater interest in politics. The researcher then divided the participants by gender (Male/Female) and then again by level of education (School/College/University). Therefore, the dependent variable was "interest in politics", and the two independent variables were "gender" and "education".

A. Setup in SPSS Statistics

A researcher had previously discovered that interest in politics is influenced by level of education. When participants were classified into three groups according to their highest level of education; namely "school", "college" or "university", in that order; higher education levels were associated with a greater interest in politics. Having demonstrated this, the researcher was now interested in determining whether this effect of education level on interest in politics was different for males and females (i.e., different depending on your gender). To answer this question, they recruited 60 participants: 30 males and 30 females, equally split by level of education (School/College/University) (i.e., 10 participants in each group). The researcher had participants complete a questionnaire that assessed their interest in politics, which they called the "Political Interest" scale. Participants could score anything between 0 and 100, with higher scores indicating a greater interest in politics.

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 `education_level`. For `gender`, we coded "males" as **1** and "females" as **2**, and for `education_level`, we coded "school" as **1**, "college" as **2** and "university" as **3**. The participants' interest in politics – the dependent variable – was entered under the variable name, `political_interest`. The setup for this example can be seen below:

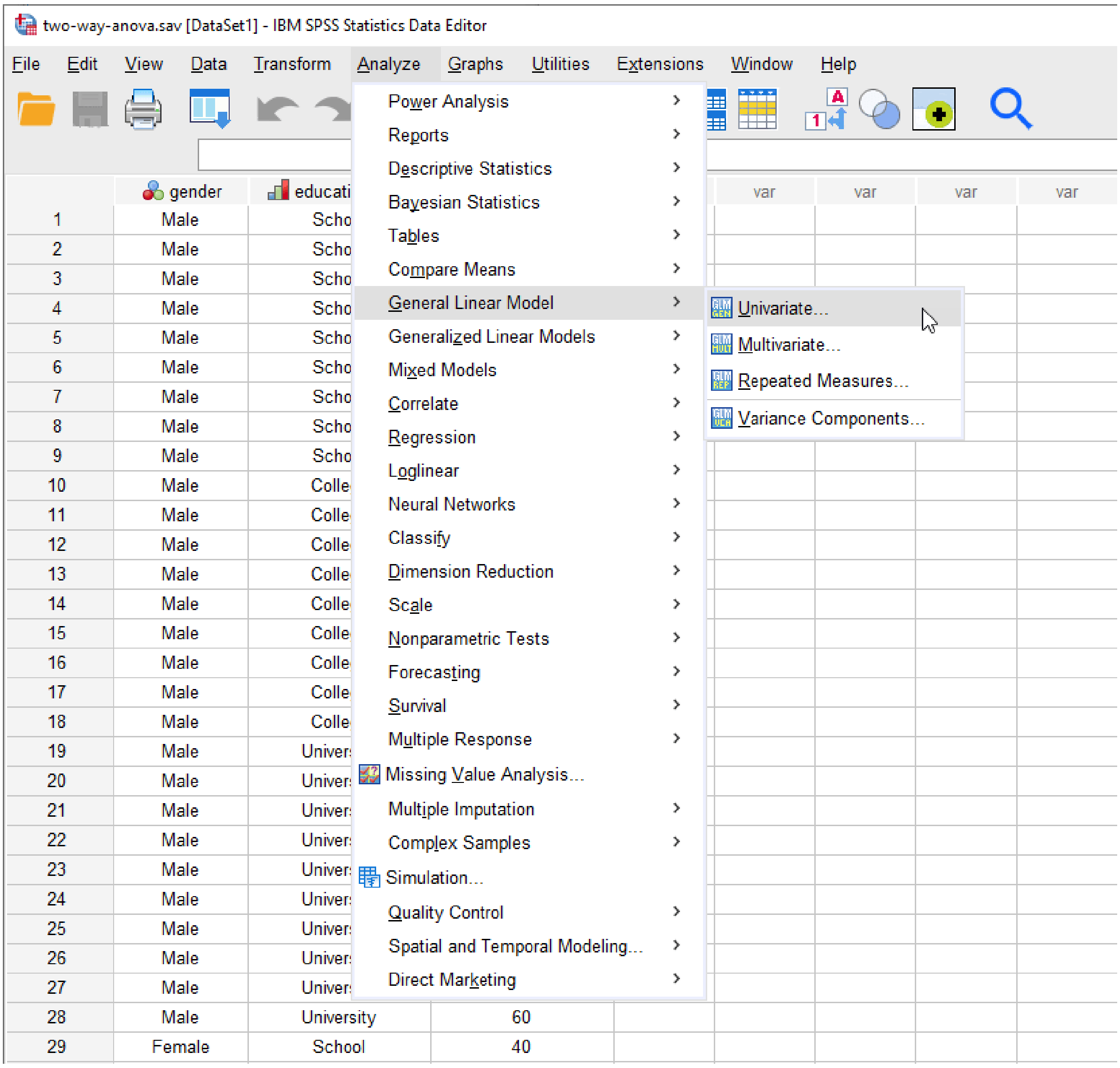


	gender	education_level	political_interest	var
1	Male	School	38	
2	Male	School	39	
3	Male	School	35	
4	Male	School	38	
5	Male	School	41	
6	Male	School	40	
7	Male	School	36	
8	Male	School	37	
9	Male	School	33	
10	Male	College	42	
11	Male	College	42	
12	Male	College	45	
13	Male	College	45	
14	Male	College	44	
15	Male	College	47	
16	Male	College	42	
17	Male	College	44	
18	Male	College	39	
19	Male	University	63	
20	Male	University	64	
21	Male	University	61	
22	Male	University	64	

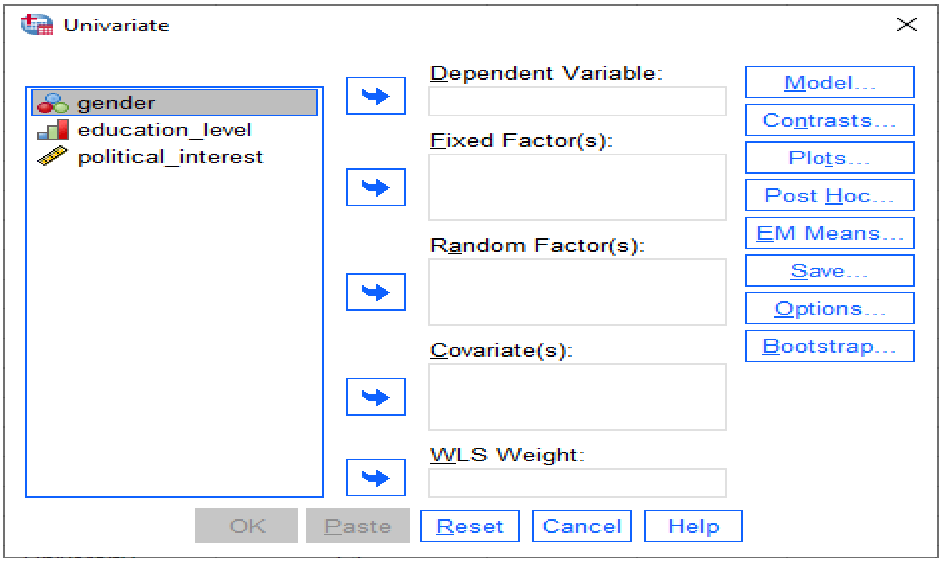
SPSS Statistics procedure for the two-way ANOVA

The **General Linear Model > Univariate...** procedure below shows you how to analyse your data using a two-way ANOVA in SPSS Statistics when the assumptions have not been violated.

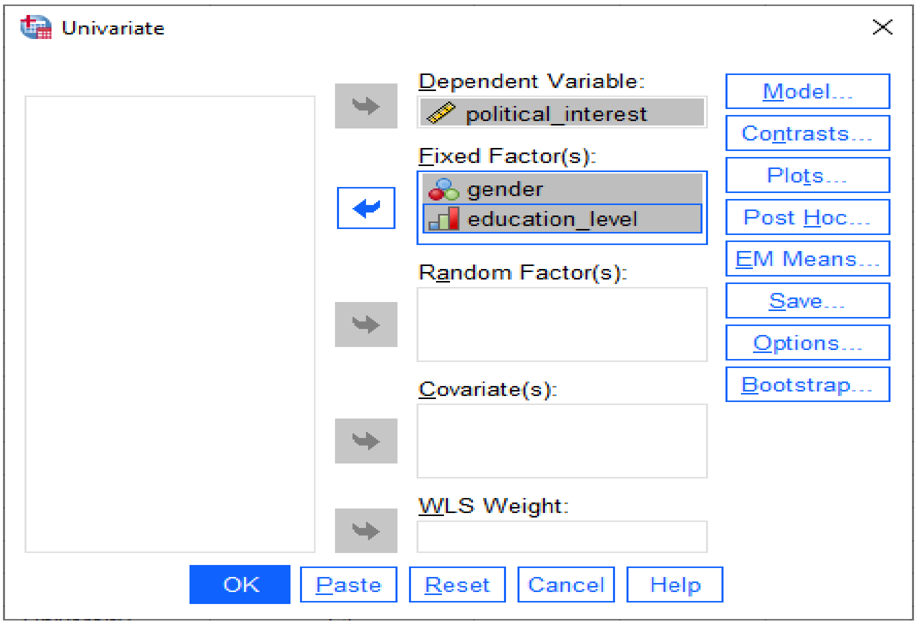
1. Click **Analyze > General Linear Model > Univariate...** on the top menu, as shown below:



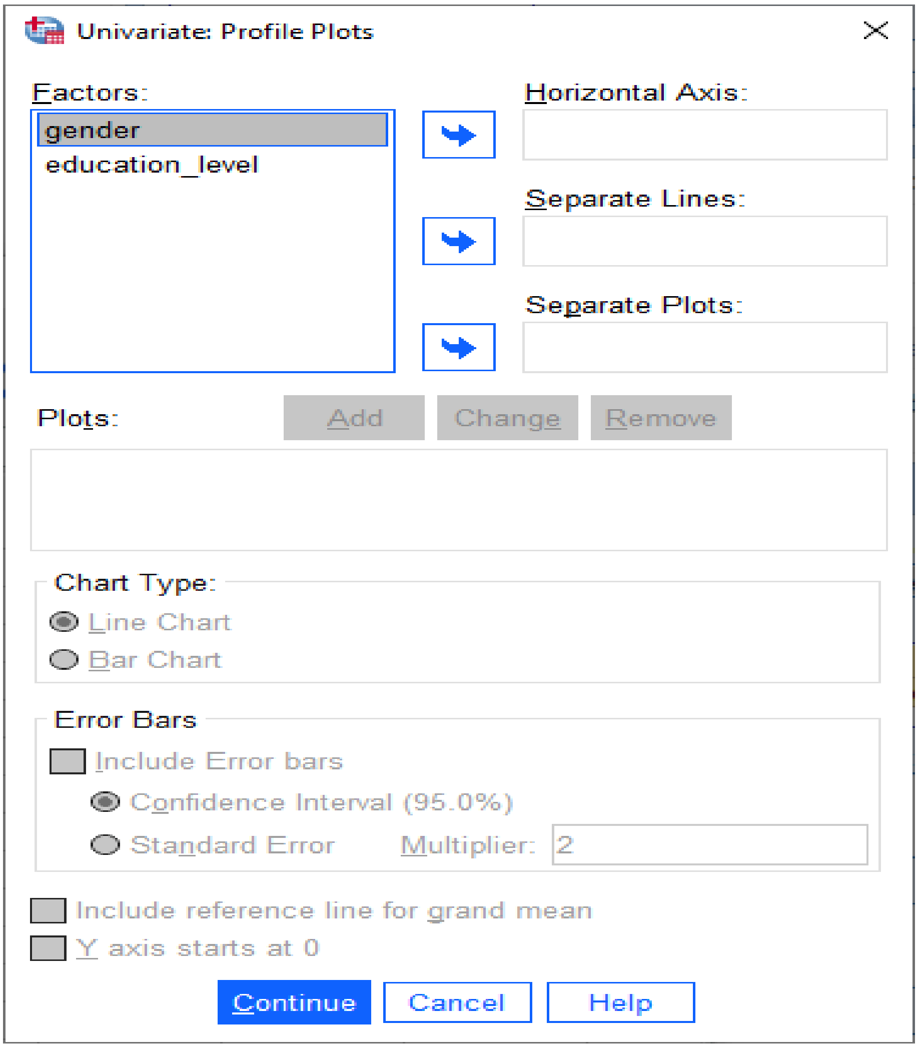
You will be presented with the **Univariate** dialogue box, as shown below:



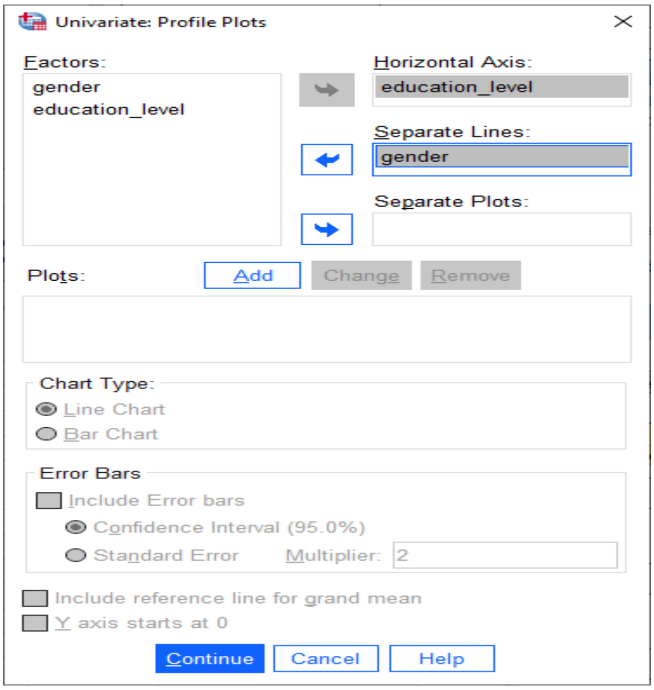
2. Transfer the dependent variable, political_interest, into the Dependent Variable: box, and the independent variables, gender and education_level, into the Fixed Factor(s): box, using the relevant  buttons, as shown below:



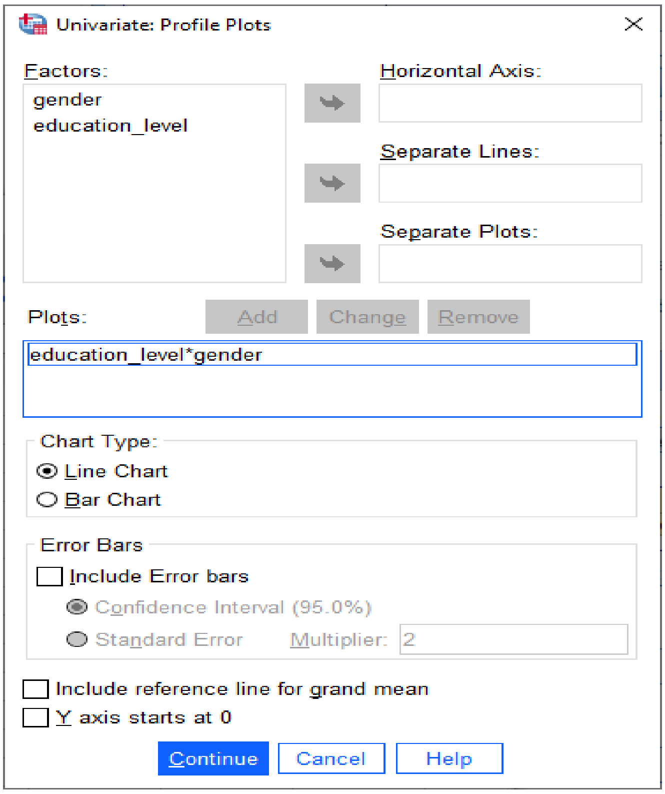
3. Click on the Plots... button. You will be presented with the **Univariate: Profile Plots** dialogue box, as shown below:



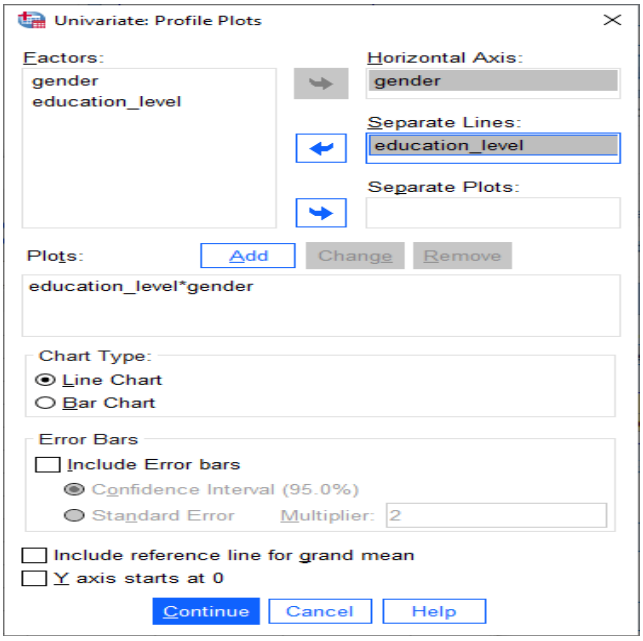
4. Transfer gender from the Factors: box to the Separate Lines: box and education_level into the Horizontal Axis: box, as shown below:



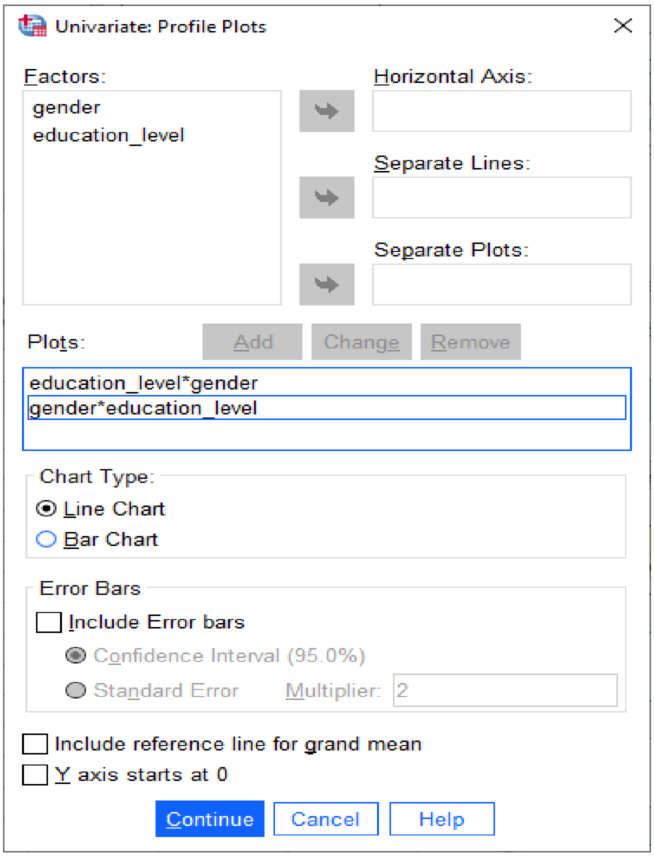
5. Click on the Add button. This will add this profile plot, which it labels "**education_level*gender**", into the Plots: box, as shown below:



6. Transfer education_level from the Factors: box to the Separate Lines: box and gender into the Horizontal Axis: box, as shown below:

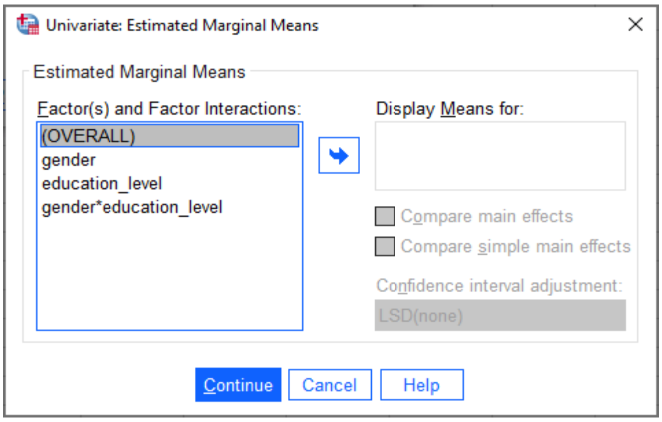


7. Click on the Add button. This will add this profile plot, which it labels "**gender*education_level**", into the Plots: box, as shown below:

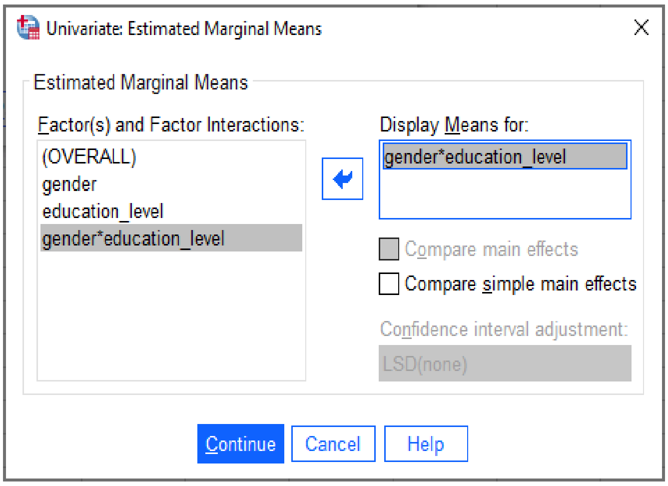


8. Click on the Continue button. You will be returned to the **Univariate** dialogue box.

9. Click on the EM Means... button. You will be presented with the **Univariate: Estimated Marginal Means** dialogue box, as shown below:



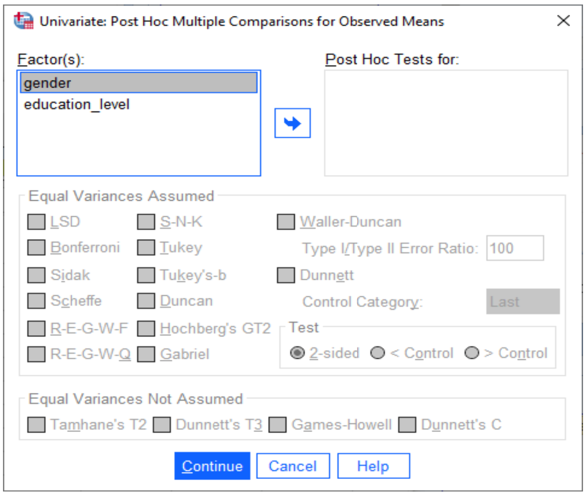
10. Transfer the interaction effect, "**gender*education_level**", from the Factor(s) and Factor Interactions: box to the Display Means for: box by highlighting it and clicking on the → button, as shown below:



11. Click on the Continue button. You will be returned to the **Univariate** dialogue box.

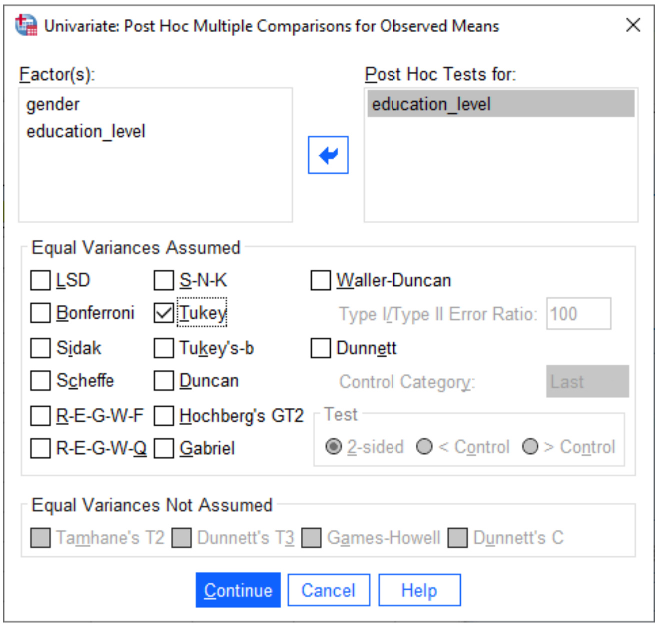
12. Click on the Post Hoc... button. You will be presented with the **Univariate: Post Hoc**

13. **Multiple Comparisons for Observed Means** dialogue box, as shown below:



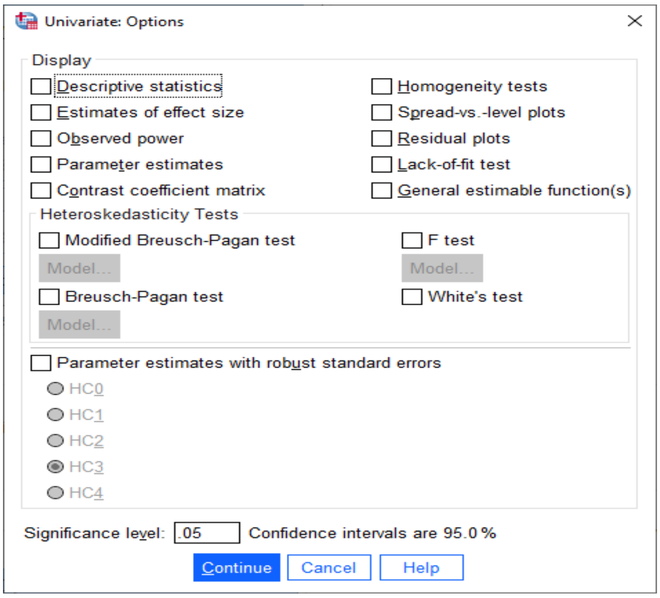
14. Transfer education_level from the Factor(s): box to the Post Hoc Tests for: box using the  button. This will activate the –Equal Variances Assumed– area (i.e., it will no longer be greyed out) and present you with some choices for which post hoc test to use. For this example, we are going to select Tukey, which is a good, all-round post hoc test.

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 .

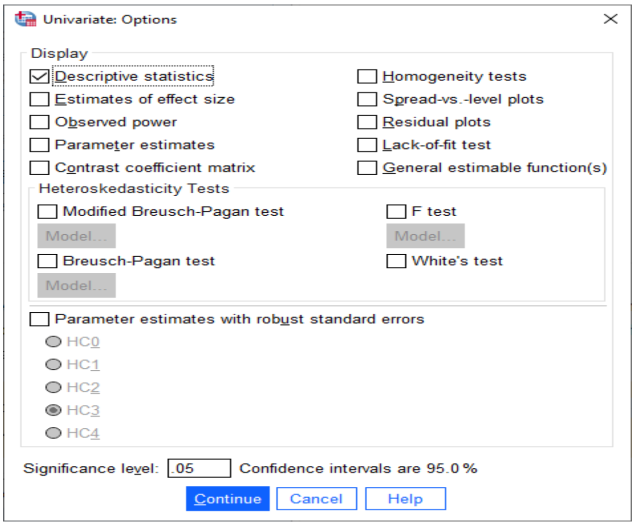


15. Click on the Continue button. You will be returned to the **Univariate** dialogue box.

16. Click on the Options... button. You will be presented with the **Univariate: Options** dialogue box, as shown below:



17. Select the Descriptive statistics option in the –Display– area. You will end up with the following screen:



18. Click on the Continue button. You will be returned to the **Univariate** dialogue box.

19. Click on the OK button.

SPSS Statistics Output

SPSS Statistics generates quite a few tables in its output from a two-way ANOVA. This section shows the main tables required to understand your results from the two-way ANOVA, including descriptives, between-subjects effects, Tukey post hoc tests (multiple comparisons), a plot of the results, and how to write up these results.

B. *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

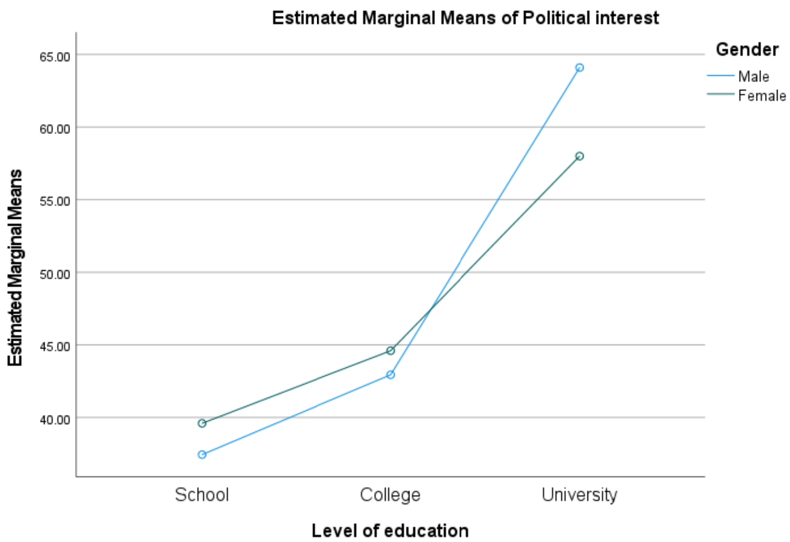
Dependent Variable: Political interest

Gender	Level of education	Mean	Std. Deviation	N
Male	School	37.4444	2.50555	9
	College	42.9444	2.33779	9
	University	64.1000	3.07137	10
	Total	48.7321	12.15447	28
Female	School	39.6000	3.27278	10
	College	44.6000	3.27278	10
	University	58.0000	6.46357	10
	Total	47.4000	9.05767	30
Total	School	38.5789	3.06079	19
	College	43.8158	2.91648	19
	University	61.0500	5.83524	20
	Total	48.0431	10.59099	58

This table is very useful because it 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 "education_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 (with the lines actually crossing). You might expect there to be a statistically significant interaction, which we can confirm in the next section.

C. *Statistical Significance*

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: Political interest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5645.998[a]	5	1129.200	78.538	<.001
Intercept	132091.906	1	132091.906	9187.227	<.001
gender	8.420	1	8.420	.586	.448
education_level	5446.697	2	2723.348	189.414	<.001
gender * education_level	210.338	2	105.169	7.315	.002
Error	747.644	52	14.378		
Total	140265.750	58			
Corrected Total	6393.642	57			

a. R Squared = .883 (Adjusted R Squared = .872)

The particular rows we are interested in are the "gender", "education_level" and "gender*education_level" rows, and these are highlighted above. These rows inform us whether our independent variables (the "gender" and "education_level" rows) and their interaction (the "gender*education_level" row) have a statistically significant effect on the dependent variable, "interest in politics". It is important to first look at the "gender*education_level" interaction as this will determine how you can interpret your results (see our enhanced guide for more information). 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 "education_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 no statistically significant difference in mean interest in politics between males and females (p = .448), but there were statistically significant differences between educational levels (p < .001).

D. *Post hoc Tests – Simple Main Effects in SPSS Statistics*

When you have a **statistically significant interaction**, reporting the **main effects** can be misleading. Therefore, you will need to report the **simple main effects**. In our example, this would involve determining the mean difference in interest in politics between genders at each educational level, as well as between educational level for each gender. You can carry out a simple main effects analysis, and you will need to use SPSS Statistics **syntax**.

When you **do not** have a statistically significant interaction, we explain two options you have, as well as a procedure you can use in SPSS Statistics to deal with this issue.

E. *Multiple Comparisons Table*

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

Multiple Comparisons

Dependent Variable: Political interest
Tukey HSD

(I) Level of education	(J) Level of education	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
School	College	-5.2368*	1.23022	<.001	-8.2049	-2.2688
	University	-22.4711*	1.21475	<.001	-25.4017	-19.5404
College	School	5.2368*	1.23022	<.001	2.2688	8.2049
	University	-17.2342*	1.21475	<.001	-20.1649	-14.3035
University	School	22.4711*	1.21475	<.001	19.5404	25.4017
	College	17.2342*	1.21475	<.001	14.3035	20.1649

Based on observed means.
The error term is Mean Square(Error) = 14.378.
*. The mean difference is significant at the .05 level.

You can see from the table above that 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) School and College, (2) School and University, and (3) College and University. From the results, we can see that there is a statistically significant difference between all three different educational levels ($p < .001$).

F. *Reporting the Results of a Two-way ANOVA*

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

- General

A two-way ANOVA was conducted that examined the effect of gender and education level on interest in politics. There was a statistically significant interaction between the effects of gender and education level on interest in politics, F (2, 52) = 7.315, 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:

- General

Simple main effects analysis showed that males were significantly more interested in politics than females when educated to university level (p = .002), but there were no differences between gender when educated to school (p = .465) or college level (p = .793).

REFERENCES

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.

Lund Research Ltd. (2018). *Two-way ANOVA in SPSS Statistics*. Laerd Statistics. https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics.php