Mohamed Khider University of Biskra
Faculty of SNVSTU

Practical Protocol 01 – Analysis of Variance (ANOVA)
Biostatistics, M1 Biology

Dr. Ben Gherbal Hanane
hanane.bengherbal@univ-biskra.dz

The objective of this protocol is to perform both one-way and two-way analyses of variance (ANOVA) using the SPSS software. First, a one-way analysis of variance (ANOVA-1) is conducted. Subsequently, a two-way analysis of variance (ANOVA-2) is performed using SPSS. To ensure clarity and understanding, the different steps of the analysis are also demonstrated through a concrete example.

Exercise 1 *We aim to compare three types of feed based on their effect on milk production. We randomly assign 15 cows as follows:*

- A_1 *to the first 5 cows,*
- A_2 *to the next 5 cows,*
- A_3 *to the last 5 cows.*

A_1	A_2	A_3
38	42	30
40	45	32
41	43	41
35	44	34
36	39	33

At a significance level $\alpha = 5\%$, *test the hypothesis that the feeds have no effect on milk production.*

Solution Using SPSS

1. Data Entry: Enter all values into a single variable called "Production", and create a second variable named "Feed" to indicate group membership.

	Production	feed
1	38,00	1,00
2	40,00	1,00
3	41,00	1,00
4	35,00	1,00
5	36,00	1,00
6	42,00	2,00
7	45,00	2,00
8	43,00	2,00
9	44,00	2,00
10	39,00	2,00
11	30,00	3,00
12	32,00	3,00
13	41,00	3,00
14	34,00	3,00
15	33,00	3,00

Figure 1: Data entry in SPSS

2. Running the Test: Navigate in SPSS through:

`Analyze` → `Compare Means` → `One-Way ANOVA`

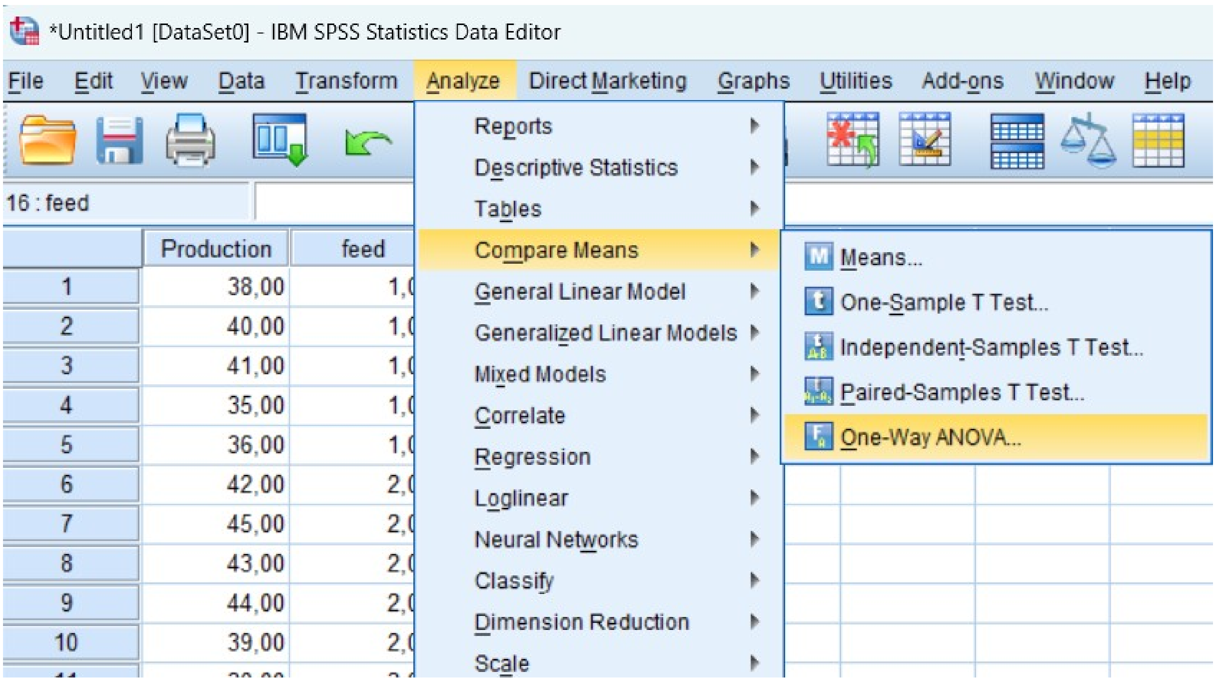


Figure 2: One-Way ANOVA menu in SPSS

3. Assigning Variables:

- Dependent List: This space is reserved for the quantitative variable that we want to analyze. In our example, it corresponds to the variable "Production".
- Factor: This box is reserved for the qualitative variable (the factor) that indicates the grouping of the observations. In our example, this box is reserved for the grouping variable "feed".

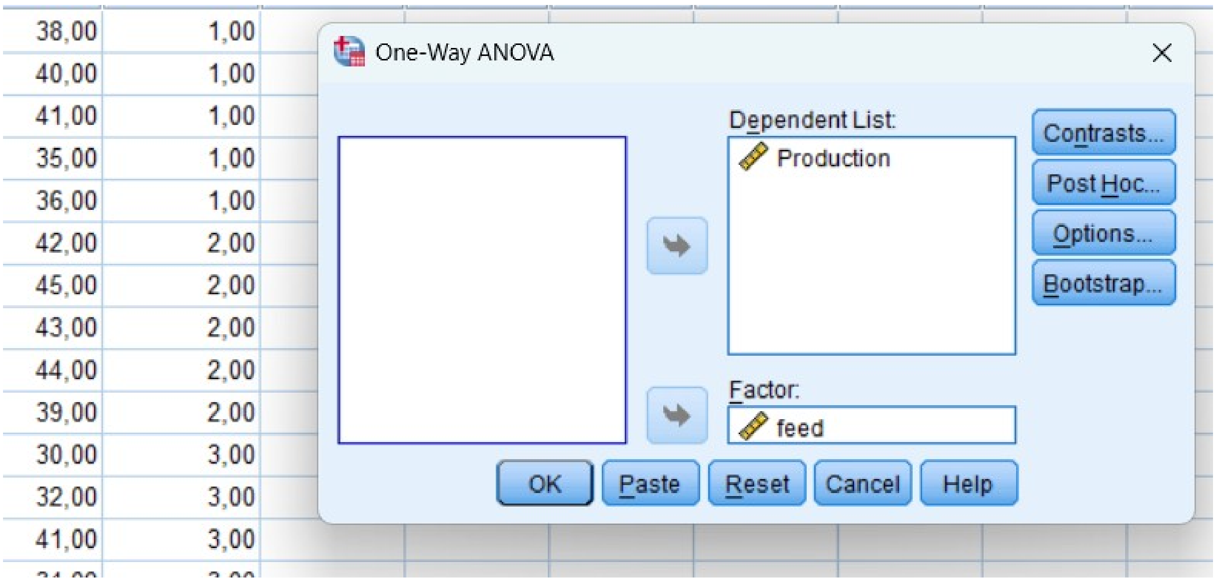
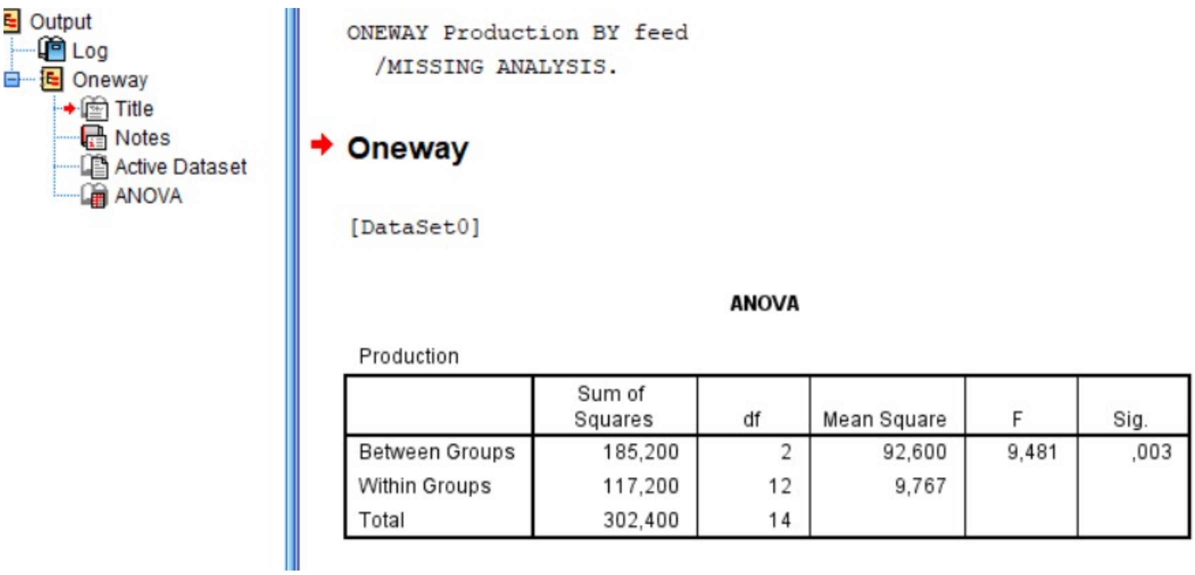


Figure 3: Variable assignment window



	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	185,200	2	92,600	9,481	,003
Within Groups	117,200	12	9,767		
Total	302,400	14			

Figure 4: ANOVA results in SPSS

4. Results Output: Interpretation of the results:

Compare the significance value (p-value) with $\alpha = 0.05$:

$$\begin{cases} \alpha < \text{p-value: } H_0 \text{ accepted (no significant effect)} \\ \alpha \geq \text{p-value: } H_0 \text{ rejected (significant effect)} \end{cases}$$

Thus, in our example, the feed type has a significant effect on milk production.

Exercise 2 *A biology researcher wants to evaluate the effect of two factors on plant growth: the type of fertilizer* (F_1, F_2, F_3) *and the type of soil* (S_1, S_2)*. Each combination of fertilizer and soil is tested on* 5 *plants (repetitions), and plant height is measured after* 4 *weeks. The results are:*

	F_1	F_2	F_3
S_1	12, 14, 13, 15, 14	17, 18, 18, 17, 18	20, 21, 21, 20, 21
S_2	11, 12, 12, 11, 12	16, 17, 16, 16, 17	20, 20, 21, 20, 20

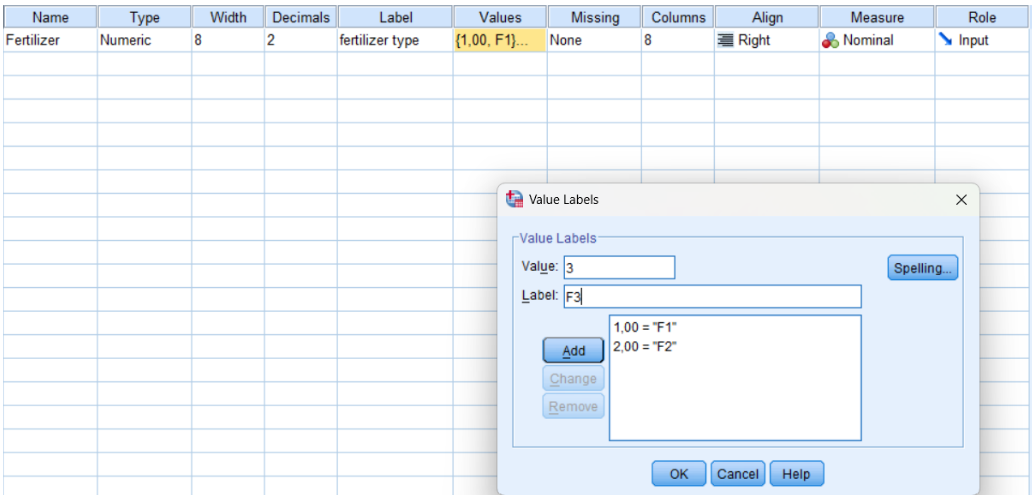
At significance level $\alpha = 4\%$*, test:*

- *the effect of fertilizer on growth;*
- *the effect of soil on growth;*
- *the interaction between fertilizer and soil.*

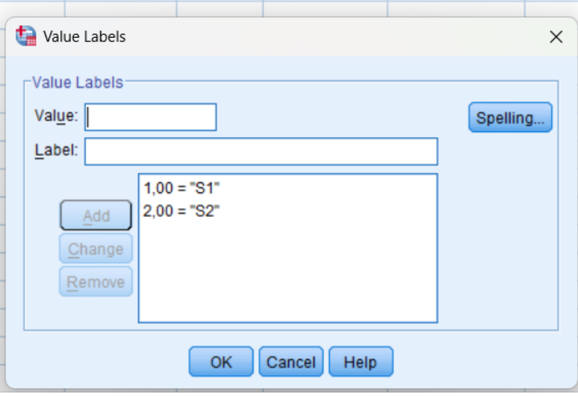
Solution with SPSS

Open SPSS and load your data: In the bottom bar choose **Variable View** and define the following variables: two qualitative variables and one quantitative variable: *fertilizer type*, *soil type*, and *plant growth*.

Choose the categories for each qualitative variable in **Value Labels**. For example: assign value 1 to F_1, value 2 to F_2, and value 3 to F_3; similarly for the soil variable.



Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
Fertilizer	Numeric	8	2	fertilizer type	{1,00, F1}...	None	8	Right	Nominal	Input
Soil	Numeric	8	2	soil type	None	None	8	Right	Nominal	Input



Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
Fertilizer	Numeric	8	2	fertilizer type	{1,00, F1}...	None	8	Right	Nominal	Input
Soil	Numeric	8	2	soil type	{1,00, S1}...	None	8	Right	Nominal	Input
Growth	Numeric	8	2	plant growth	None	None	8	Right	Scale	Input

Enter the data in **Data View** as follows:

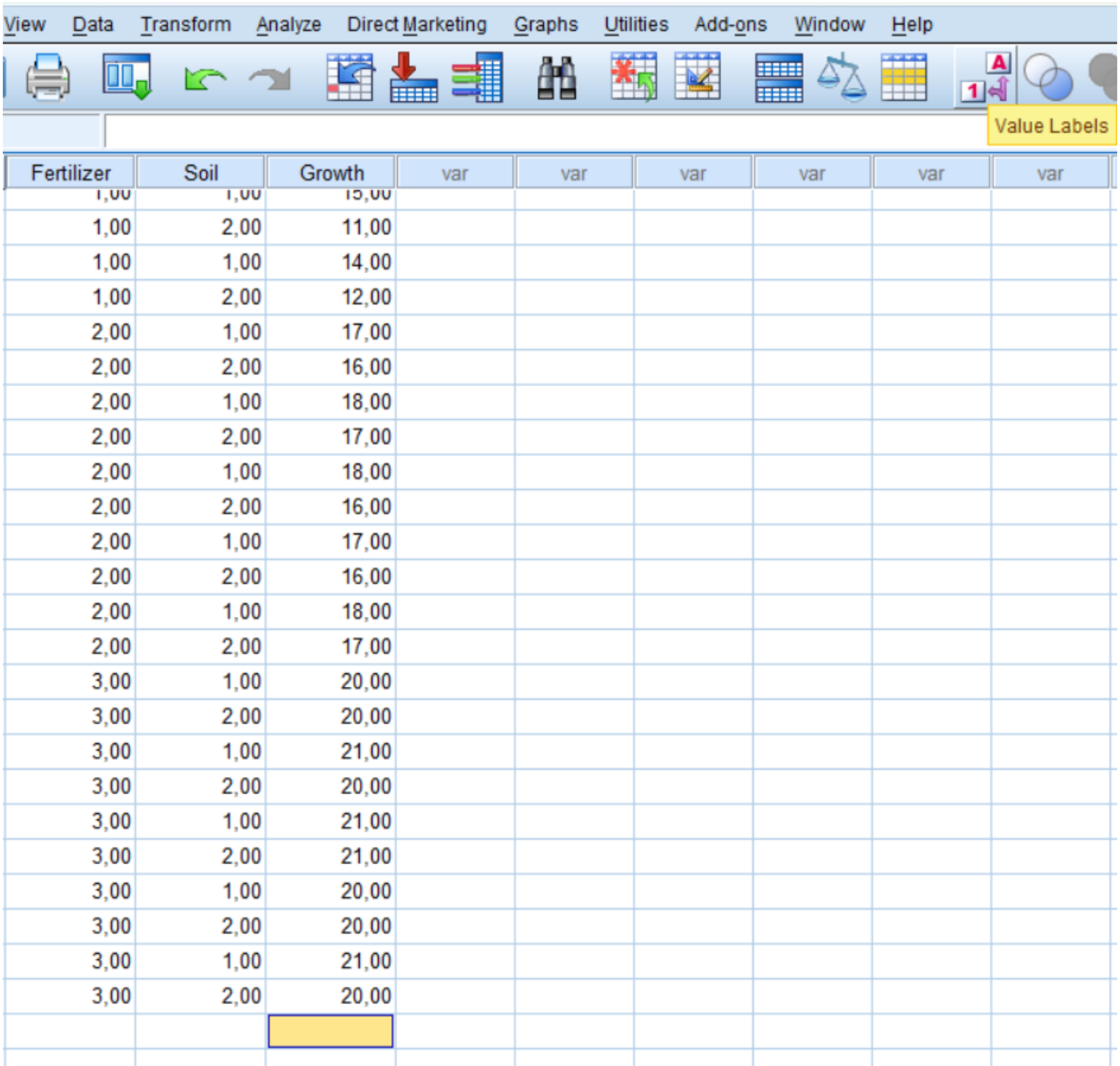
Fertilizer	Soil	Growth	var	var
F1	S1	12,00		
F1	S2	11,00		
F1	S1	14,00		
F1	S2	12,00		
F1	S1	13,00		
F1	S2	12,00		
F1	S1	15,00		
F1	S2	11,00		
F1	S1	14,00		
F1	S2	12,00		

F1
F2
F3

Fertilizer	Soil	Growth	var	var
F1	S1	14,00		
F1	S2	12,00		
F2		.		
F2	S1	.		
F2	S2	.		
F2	.	.		
F2	.	.		
F2	.	.		
F2	.	.		
F2	.	.		
F2	.	.		
F2	.	.		

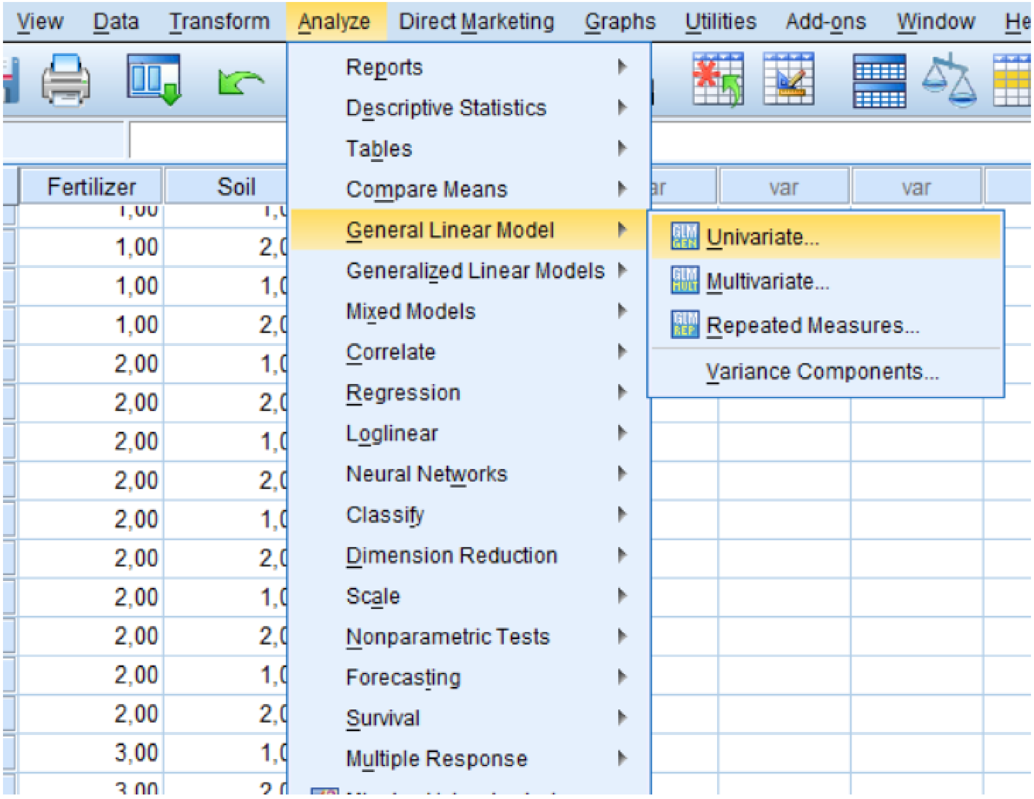
Fertilizer	Soil	Growth	var	var
F1	S1	15,00		
F1	S2	11,00		
F1	S1	14,00		
F1	S2	12,00		
F2	S1	17,00		
F2	S2	16,00		
F2	S1	18,00		
F2	S2	17,00		
F2	S1	18,00		
F2	S2	16,00		
F2	S1	17,00		
F2	S2	16,00		
F2	S1	18,00		
F2	S2	17,00		
F3	S1	20,00		
F3	S2	20,00		
F3	S1	21,00		
F3	S2	20,00		
F3	S1	21,00		
F3	S2	21,00		
F3	S1	20,00		
F3	S2	20,00		
F3	S1	21,00		
F3	S2	20,00		

By clicking on **Value Labels**, you can display the coding of the categories.

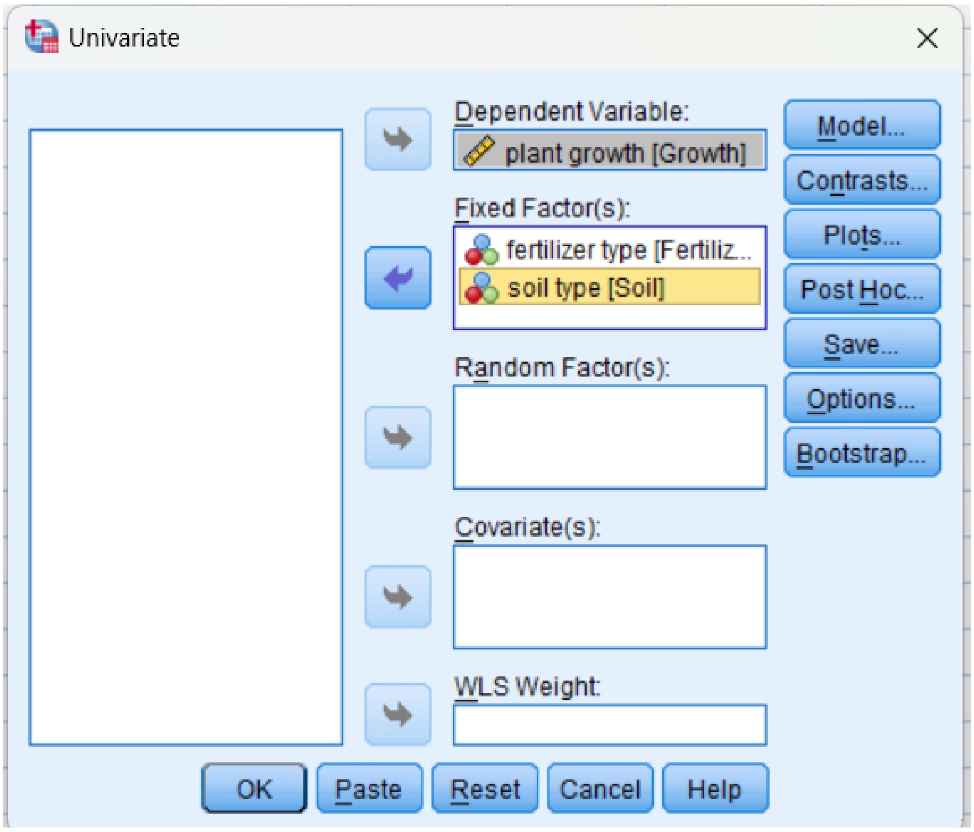


Fertilizer	Soil	Growth	var	var	var	var	var	var
1,00	1,00	15,00						
1,00	2,00	11,00						
1,00	1,00	14,00						
1,00	2,00	12,00						
2,00	1,00	17,00						
2,00	2,00	16,00						
2,00	1,00	18,00						
2,00	2,00	17,00						
2,00	1,00	18,00						
2,00	2,00	16,00						
2,00	1,00	17,00						
2,00	2,00	16,00						
2,00	1,00	18,00						
2,00	2,00	17,00						
3,00	1,00	20,00						
3,00	2,00	20,00						
3,00	1,00	21,00						
3,00	2,00	20,00						
3,00	1,00	21,00						
3,00	2,00	21,00						
3,00	1,00	20,00						
3,00	2,00	20,00						
3,00	1,00	21,00						
3,00	2,00	20,00						

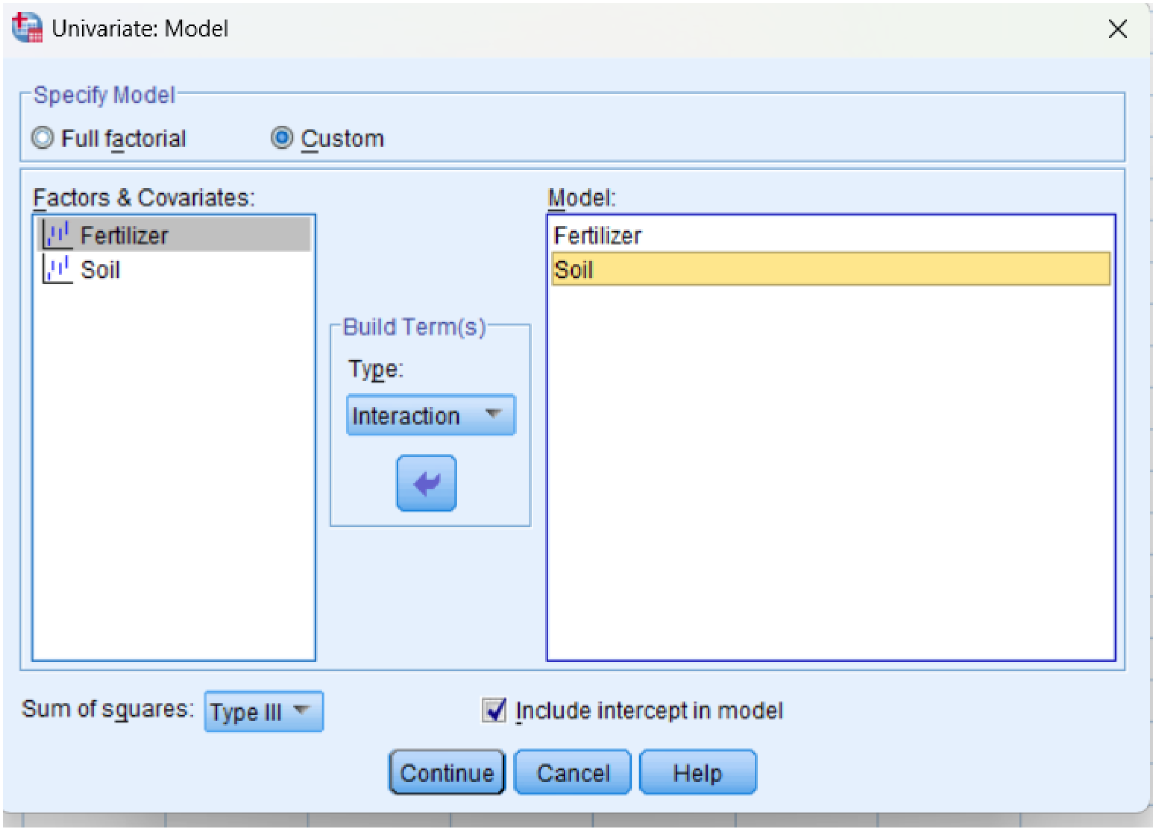
Performing the two-way ANOVA test: **Analyze** $\longrightarrow$ **General Linear Model** $\longrightarrow$ **Univariate** (one dependent variable).

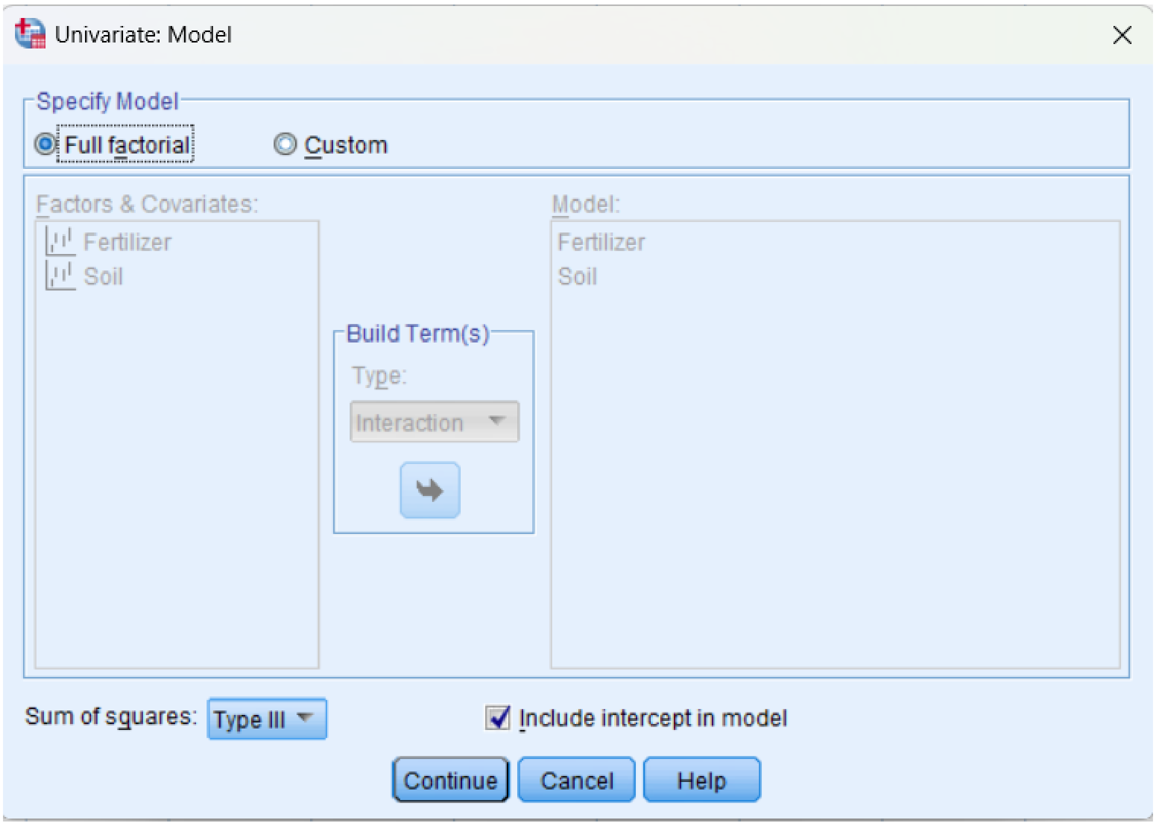


Place **plant growth** in **Dependent Variable**, and place the two factors **fertilizer** and **soil** in **Fixed Factors**.

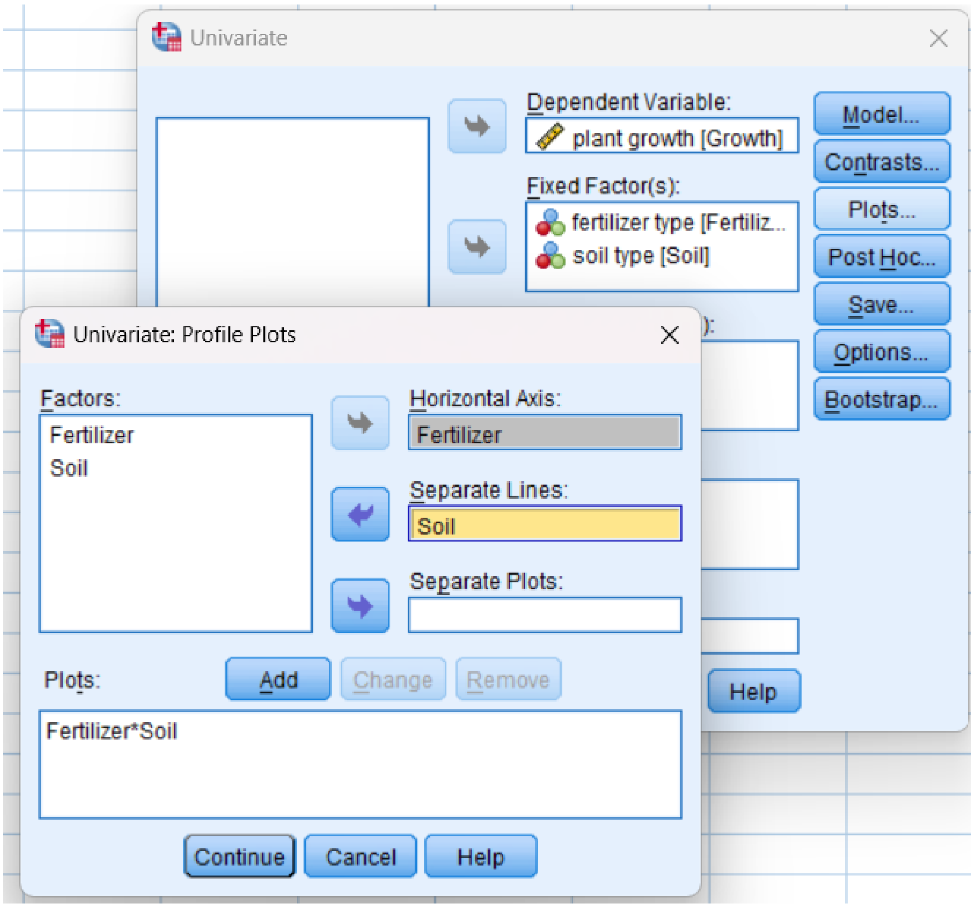


For **Model**: drag the factors into the interaction model.

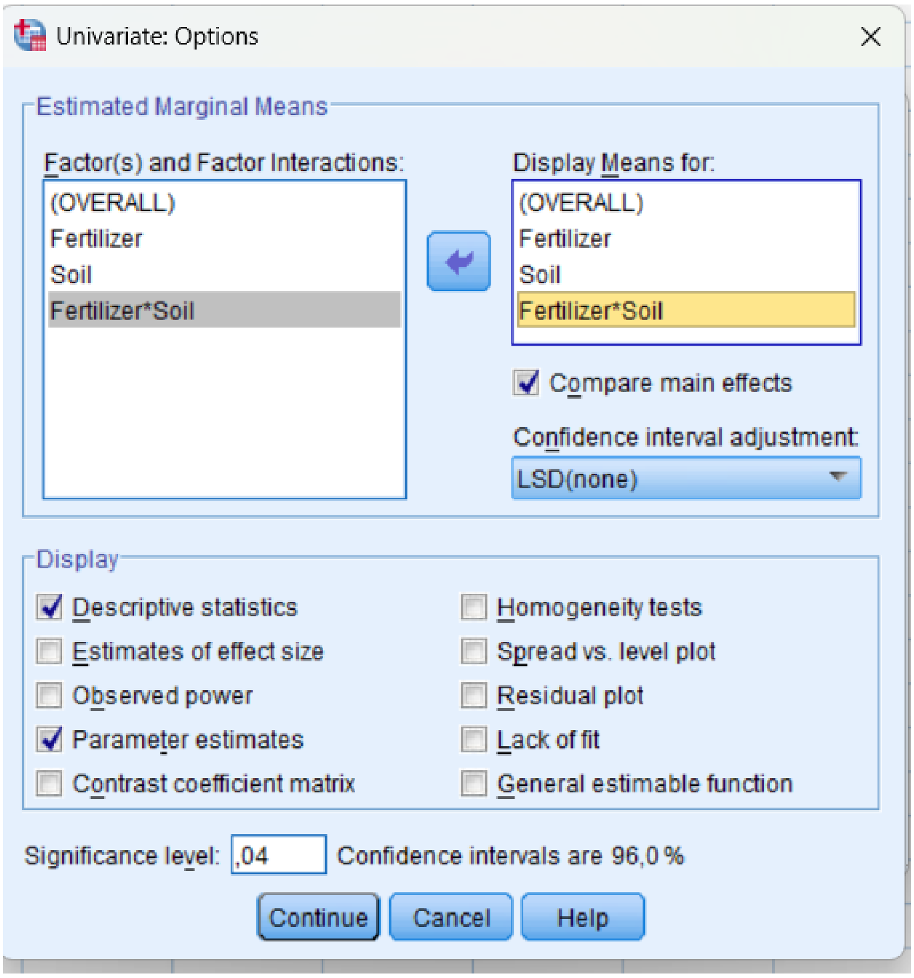




For **Plots**: place **fertilizer** on the horizontal axis and **soil** on separate lines, then click **Add** ⟶ **Continue** (this option allows checking interaction graphically).



For **Options**: place the factors and the interaction in the box (**Display Means for**), then check: **Descriptive statistics**, **Effect size estimates**, and set the risk to 4%.



Interpretation of results

The first two tables provide an overall view of the descriptive statistics for the dependent variable as well as for the qualitative variables. The second table presents the means for each cell (intersection between level i and level j).

Between-Subjects Factors

		Value Label	N
fertilizer type	1,00	F1	10
	2,00	F2	10
	3,00	F3	10
soil type	1,00	S1	15
	2,00	S2	15

Descriptive Statistics

Dependent Variable: plant growth

fertilizer type	soil type	Mean	Std. Deviation	N
F1	S1	13,6000	1,14018	5
	S2	11,6000	,54772	5
	Total	12,6000	1,34990	10
F2	S1	17,6000	,54772	5
	S2	16,4000	,54772	5
	Total	17,0000	,81650	10
F3	S1	20,6000	,54772	5
	S2	20,2000	,44721	5
	Total	20,4000	,51640	10
Total	S1	17,2667	3,05817	15
	S2	16,0667	3,67359	15
	Total	16,6667	3,37673	30

The two-way ANOVA table (**Tests of Between-Subjects Effects**) allows us to determine whether there is: an effect of **fertilizer** on **plant growth**, an effect of **soil** on **plant growth**, and an **interaction** effect between the two factors on the dependent variable.

Tests of Between-Subjects Effects

Dependent Variable: plant growth

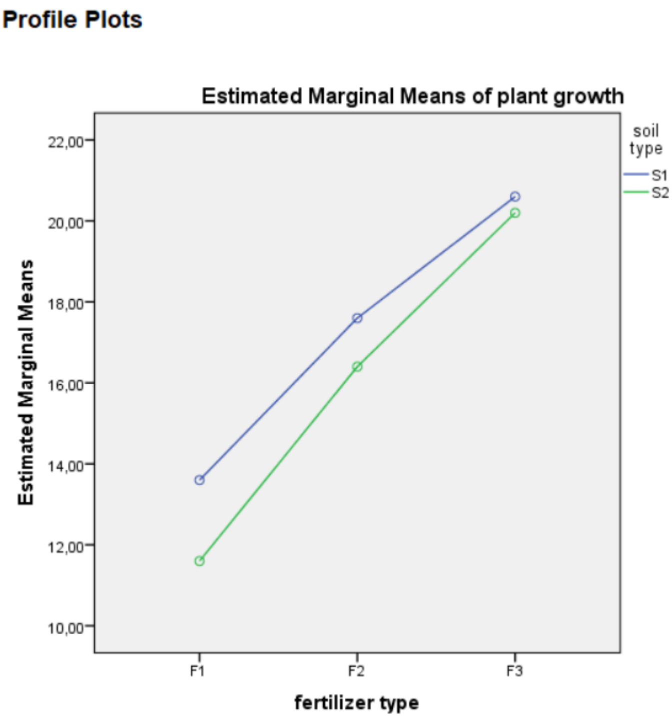
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	319,867[a]	5	63,973	142,163	,000
Intercept	8333,333	1	8333,333	18518,519	,000
Fertilizer	305,867	2	152,933	339,852	,000
Soil	10,800	1	10,800	24,000	,000
Fertilizer * Soil	3,200	2	1,600	3,556	,044
Error	10,800	24	,450		
Total	8664,000	30			
Corrected Total	330,667	29			

a. R Squared = ,967 (Adjusted R Squared = ,961)

For the **interaction** effect: we have ($sig = 0.044 > 0.04$), so the null hypothesis is accepted for the interaction; therefore, there is no interaction effect.

For the factor **fertilizer type**: we note that $sig = 0.000 < 0.04$. Thus, we reject the null hypothesis; there is an effect of fertilizer type on plant growth.

For the factor **soil type**: we note that $sig = 0.000 < 0.04$. Thus, we reject the null hypothesis; there is an effect of soil type on plant growth.



The plot indicates that there is no interaction between the two factors: it is enough to observe that there is no intersection between the line segments.

Exercise 3 *Three laboratories, L_1, L_2 and L_3, are required to measure the phosphorus content in 4 common products P_1, P_2, P_3 and P_4. Each product is measured twice by each laboratory. The results are:*

	P_1	P_2	P_3	P_4
L_1	$19, 17$	$29.8, 29.2$	$34.6, 33.4$	$59.1, 57.1$
L_2	$20.5, 19.5$	$28.3, 27.7$	$35.4, 33.4$	$58.8, 57.8$
L_3	$19.3, 18.7$	$26.9, 26.1$	$34.6, 33.8$	$58.7, 57.7$

Same work for a significance level $\alpha = 5\%$.