Repeated Measures Analysis of Variance

Nikos Comoutos – Antonis Hatzigeorgiadis

Introduction to repeated measures designs

Repeated measures is a term used when the same participants participate in all conditions of an experiment.

Compares several means, when those means have come for the same participants: for example, if you measured people's enjoyment when attending STATS module each month over a year-long course.

or

For example the effects of alcohol on enjoyment. We give a questionnaire assessing the enjoyment after 1 beer, 2 beers, and 3 beers. Are the scores in different conditions independent? In Anova the accuracy of F-test depends upon the assumption that scores in different conditions are independent.

In repeated-measures this assumption is violated. So in order to have an F-test that is valid we must have an important assumption that is called **SPHERICITY**.

Tests of repeated measures

- **Paired samples t-test**

- **One-way repeated measures ANOVA** (one repeated factor)

- **Two-way repeated measures ANOVA** (one repeated factor – one independent factor)

- **Two-way repeated measures ANOVA** (two repeated factors)

1 group – 2 measures – 1 variable → paired samples t-test

…Differences in speed before and after an intervention (cycling)

Speed: 100m before: 14sec…..intervention(cycling)……Speed: 100m after:12sec

Time – Intervention (independent, 2 levels – pre / post)

Speed (dependent)

We conducted paired samples t-test to examine the…

1 Group– 3 or more measures– one variable → one-way repeated measures ANOVA

…Differences in speed **before**, **after** an intervention (cycling), **after** an intervention (cycling), and **after** an intervention (cycling)

Speed: 100m before: **14sec**…..intervention(cycling-1week)……Speed:100m after:**13,8sec**….intervention(cycling-1week)……Speed: 100m after:**13,5**sec….intervention(cycling-1week)……Speed: 100m after:**13**sec

Time (independent, 4 levels – prior / post / post / post)

Speed (dependent)

One-way repeated measures ANOVA was conducted to examine differences in speed…

2 (or more) groups – 2 (or more measures) – 1 variable → two-way repeated measures ANOVA

Differences in body fat after an aerobic training and a strength training

Time (independent, 2 levels – before / after)

Training method (independent, 2 levels – aerobic / weigh training)

Body fat percentage (dependent)

Two –way repeated measures ANOVA was conducted to examine…

2 (or more) groups – 2 (or more) meaures –
2 (or more) variables → two-way repeated measures MANOVA

Differences in body fat and body mass index after an aerobic training and a strength training
Time (independent, 2 levels – before / after)
Training method (independent, 2 levels – aerobic / weigh training)
Body fat percentage (dependent)
BMI (dependent)

Two –way repeated measures MANOVA was conducted to examine…

Example 1.

One way repeated measures ANOVA with one repeated factor

- Null hypothesis: No differences in the mistakes across time for cyclists as a function of fatigue.

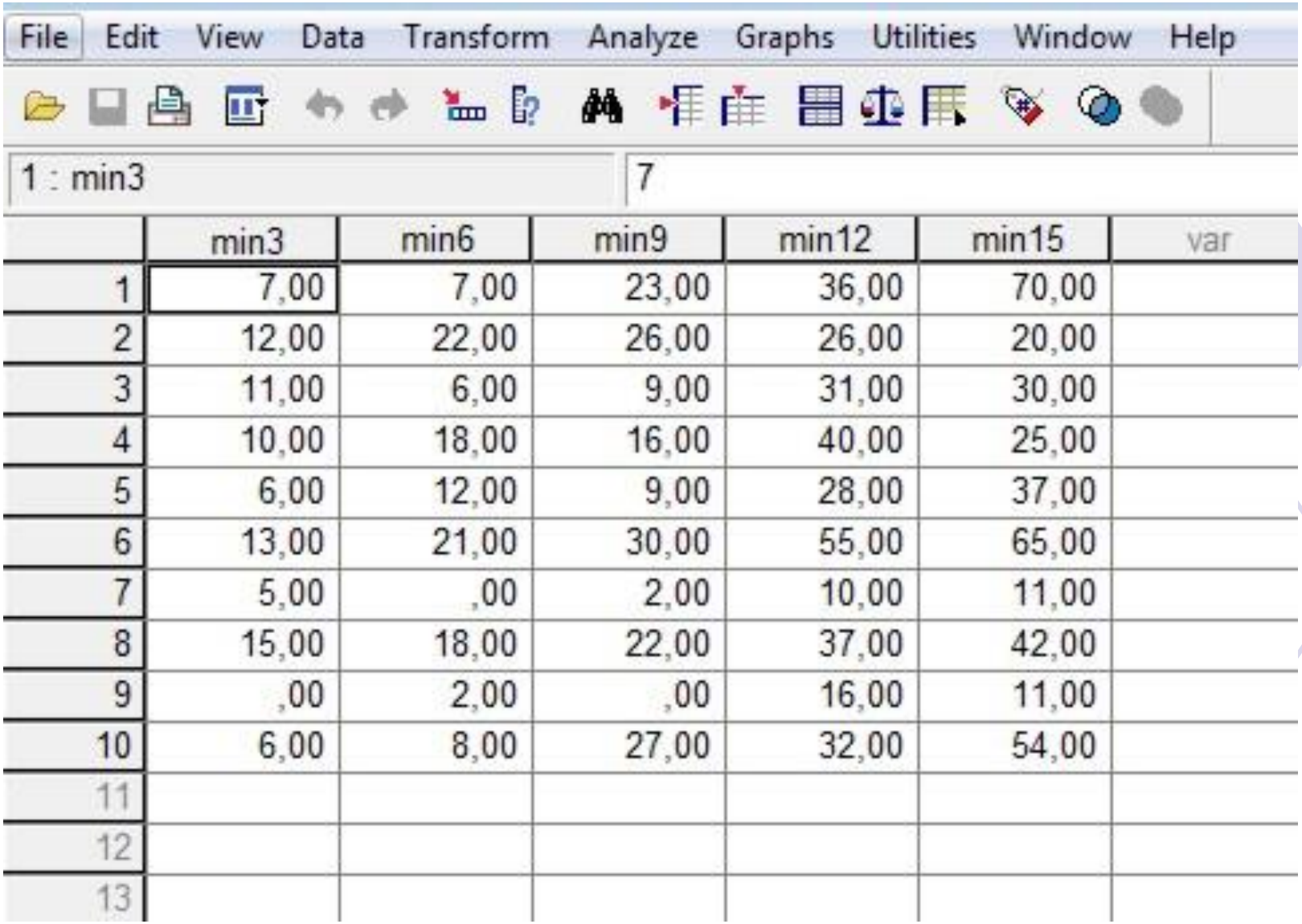
$Ho = \mu_{min3} = \mu_{min6} = \mu_{min9} = \mu_{min12} = \mu_{min15}$

Assumptions

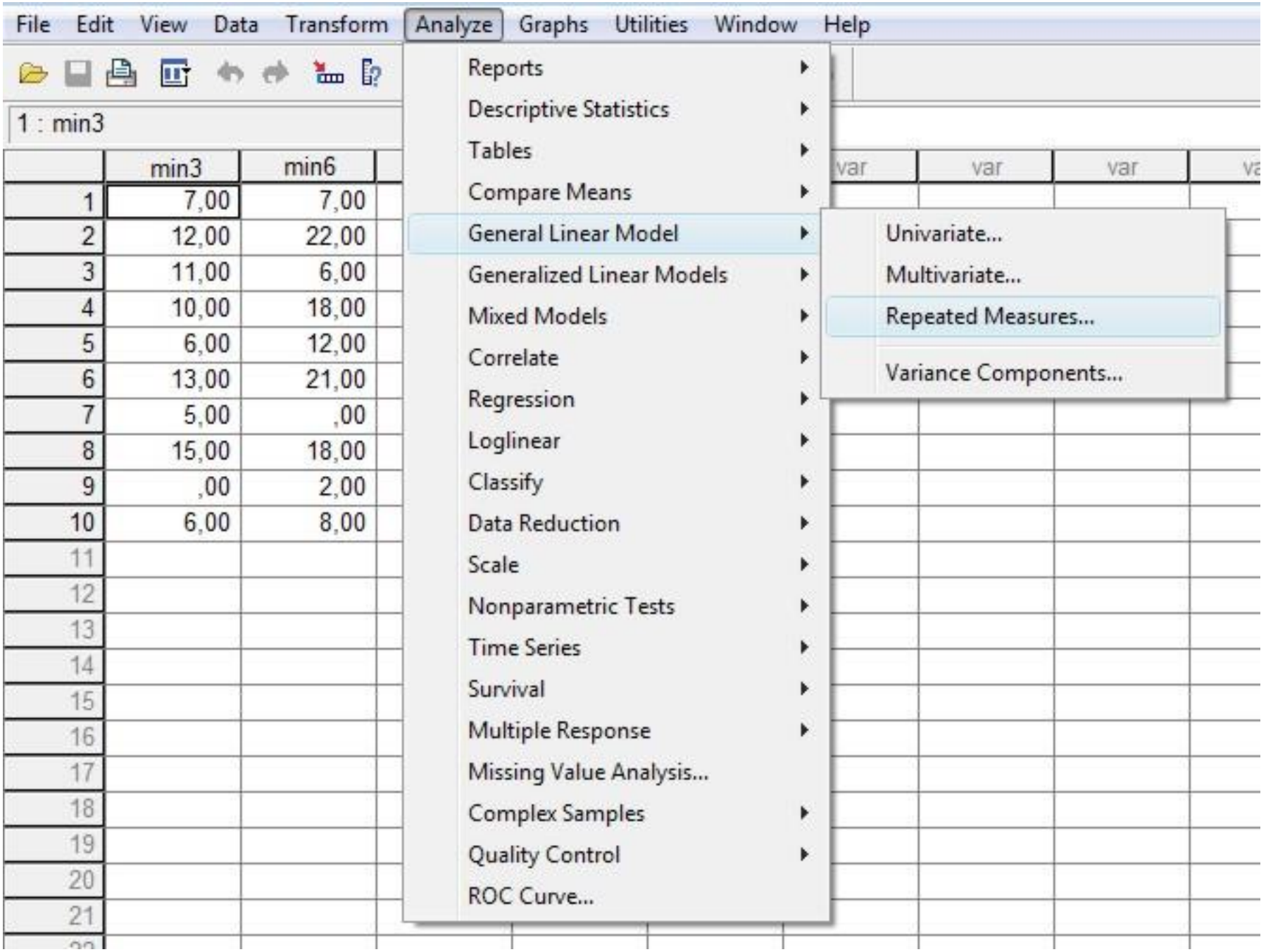
- **Normality**: The dependent variable from which the sample of participants in the population is drawn. Normally distributed.

- **Sphericity** (that there is a homogeneity of covariance –that is correlations among all combinations of trials are equal. **Mauchly's test** gives an overall single assessment of sphericity. This assumption need to be considered only when you have three or more repeated- measures conditions. If the value in the column labelled *Sig.* is less than .05 then the assumption is violated. If the significance of Mauchly's test is greater than .05 then the assuption of sphericity has been met and the F-ratios generated by the repeated –measures ANOVA can be accepted **and look at the row labeled Sphericity Assumed.** If it is violated **read the row labeled Greenhouse-Geisser** (will make an adjustment to the degrees of freedom which consequently raises the critical (table) value of F and counters the risk of a type I error.

We can use a multivariate technique Sphericity is not an issue for MANOVA & we don't meet the risk of type I error. However MANOVA REQUIRES LARGER SAMPLE SIZES AND IS LESS POWERFUL THAN UNIVARIATE ANALYSIS.



	min3	min6	min9	min12	min15	var
1	7,00	7,00	23,00	36,00	70,00	
2	12,00	22,00	26,00	26,00	20,00	
3	11,00	6,00	9,00	31,00	30,00	
4	10,00	18,00	16,00	40,00	25,00	
5	6,00	12,00	9,00	28,00	37,00	
6	13,00	21,00	30,00	55,00	65,00	
7	5,00	,00	2,00	10,00	11,00	
8	15,00	18,00	22,00	37,00	42,00	
9	,00	2,00	,00	16,00	11,00	
10	6,00	8,00	27,00	32,00	54,00	
11						
12						
13						

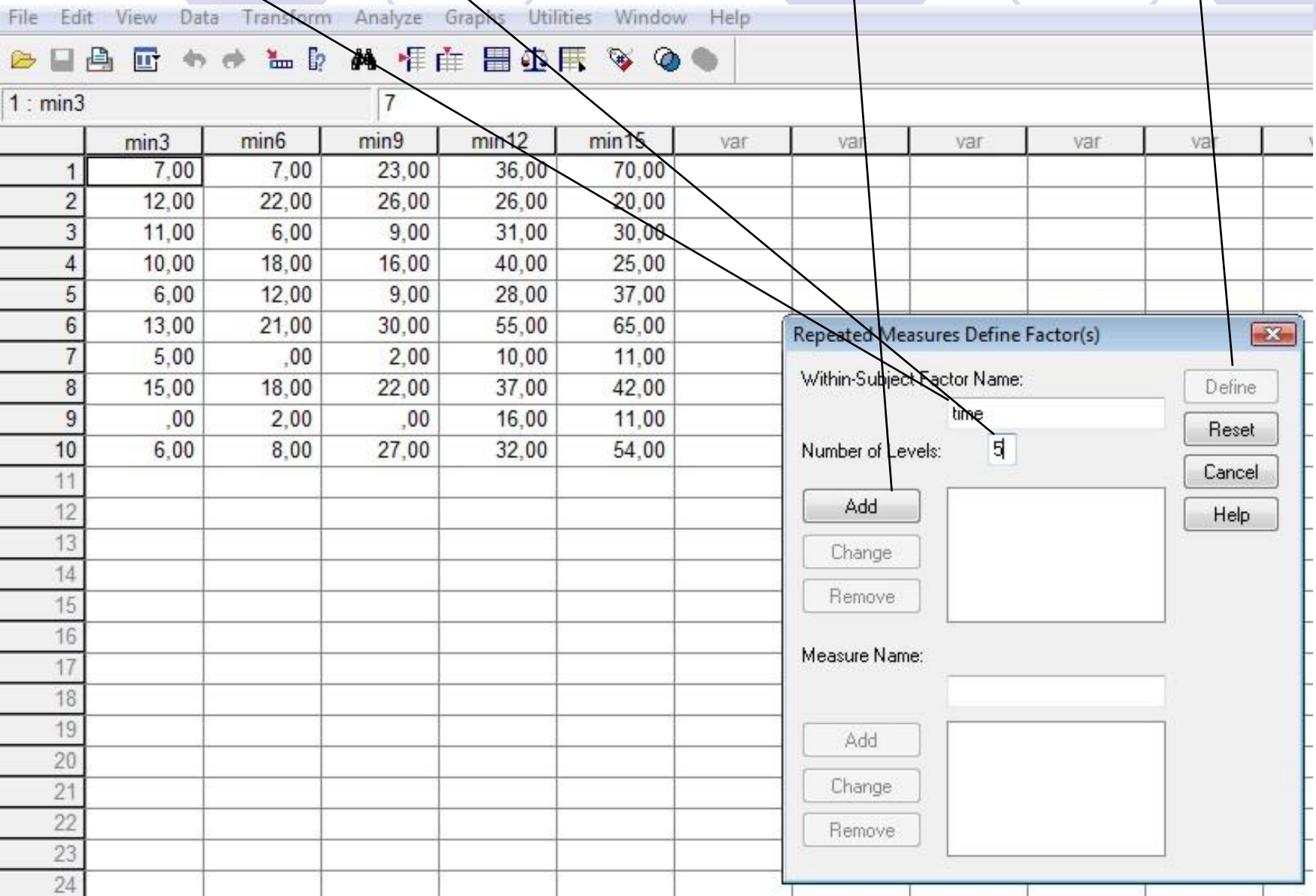

File Edit View Data Transform Analyze Graphs Utilities Window Help
1 : min3
min3 min6
1 7,00 7,00
2 12,00 22,00
3 11,00 6,00
4 10,00 18,00
5 6,00 12,00
6 13,00 21,00
7 5,00 ,00
8 15,00 18,00
9 ,00 2,00
10 6,00 8,00
var var var
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Missing Value Analysis...
Complex Samples
Quality Control
ROC Curve...
Univariate...
Multivariate...
Repeated Measures...
Variance Components...

Write e.g., TIME

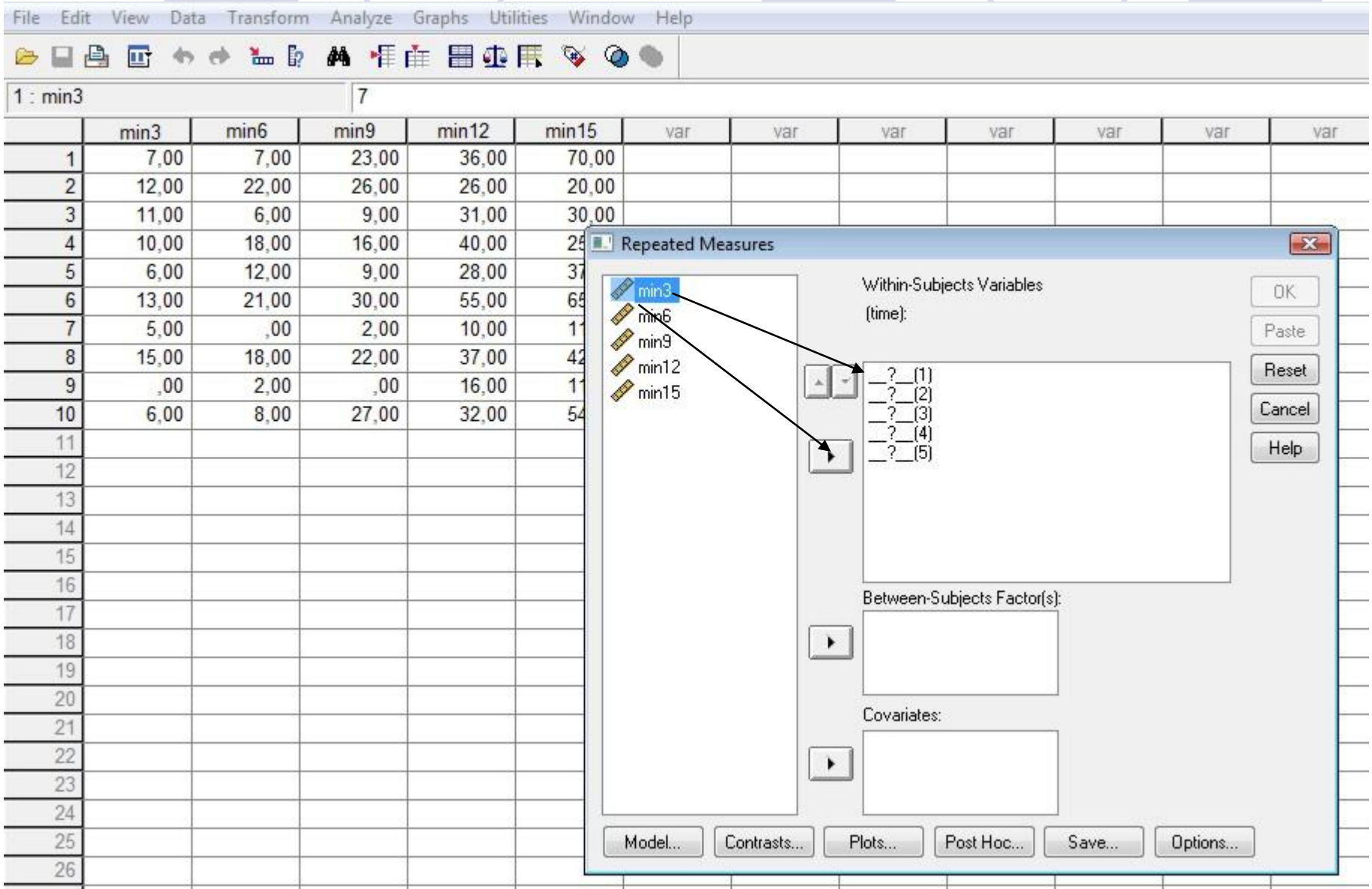
Number of trials (e.g., **5**)

Click **add**

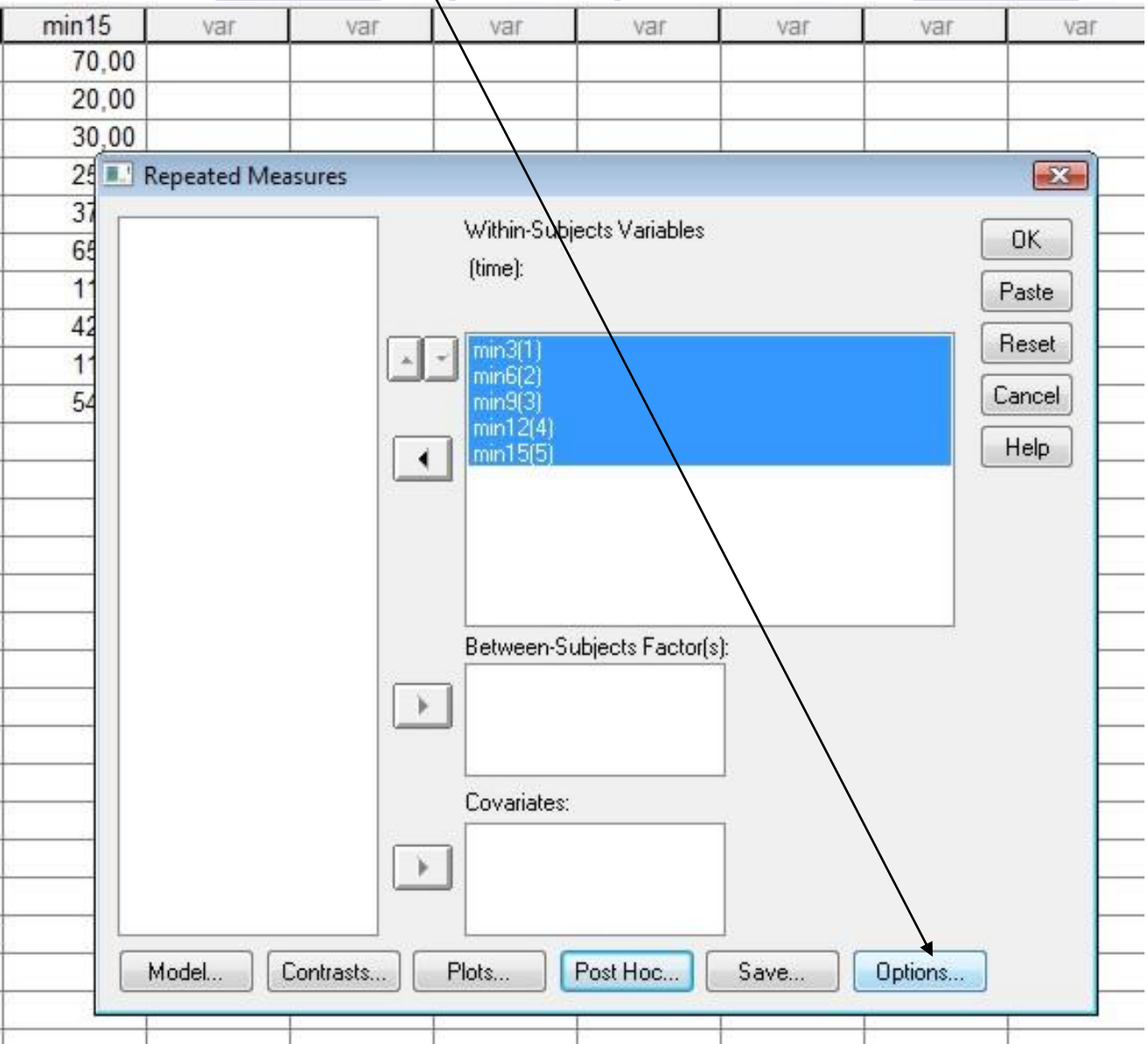
Click **define**



From left to the right…



Click Options



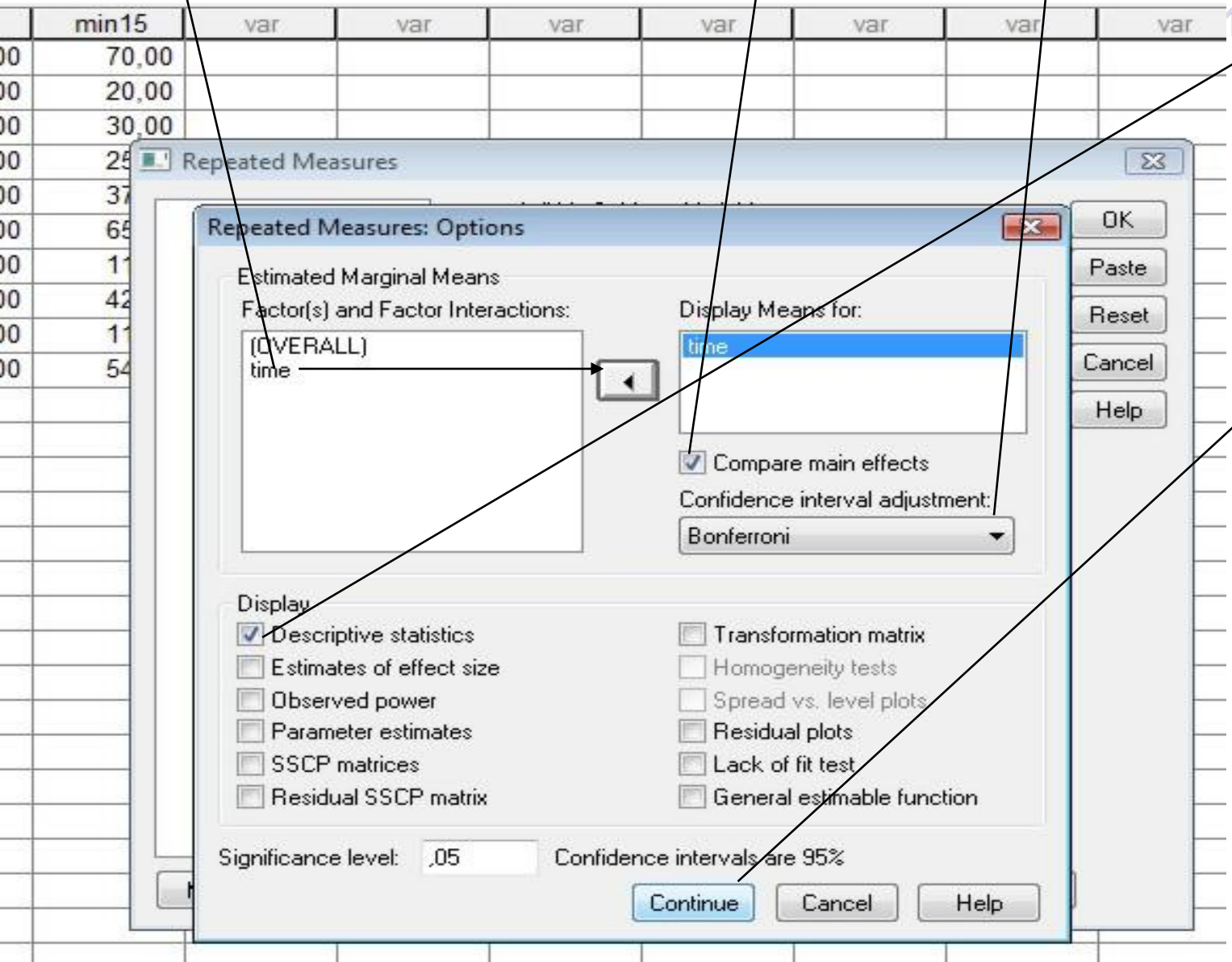
1)click time from left to the right with the arrow (2)

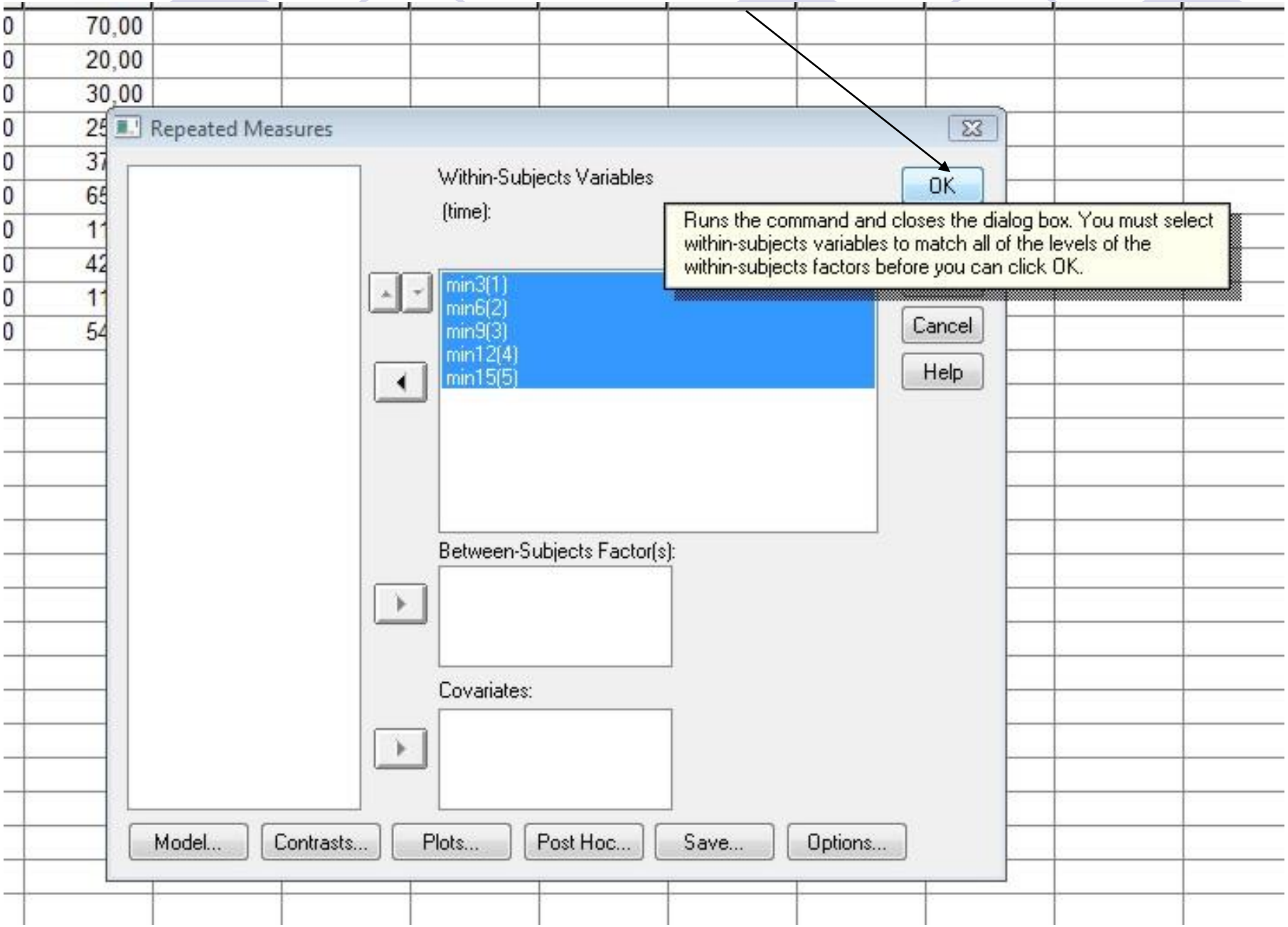
Click Compare main effects(3)

Click Bonferroni (4)

Click Descriptive (5)

Click continue (6)




70,00
20,00
30,00
Repeated Measures
Within-Subjects Variables
(time):
OK
Runs the command and closes the dialog box. You must select within-subjects variables to match all of the levels of the within-subjects factors before you can click OK.
min3(1)
min6(2)
min9(3)
min12(4)
min15(5)
Cancel
Help
Between-Subjects Factor(s):
Covariates:
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...

There is a significant difference between the 3 to 6 to 9 to 12 and to15 minutes but we don't know exactly where this difference is(e.g., between 3 to 6?)

There is a violation

We will use the corrected df

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
time	Pillai's Trace	,866	9,694[a]	4,000	6,000	,009
	Wilks' Lambda	,134	9,694[a]	4,000	6,000	,009
	Hotelling's Trace	6,463	9,694[a]	4,000	6,000	,009
	Roy's Largest Root	6,463	9,694[a]	4,000	6,000	,009

a. Exact statistic

b.
Design: Intercept
Within Subjects Design: time

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhous e-Geisser	Huynh-Feldt	Lower-bound
time	,024	27,594	9	,001	,371	,428	,250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in

The table labelled Tests of within-subjects effects shows the main result of your ANOVA

Corrected df

There are significant differences between the means

5-1 = 4

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	6115,880	4	1528,970	18,359	,000
	Greenhouse-Geisser	6115,880	1,485	4117,754	18,359	,000
	Huynh-Feldt	6115,880	1,710	3575,916	18,359	,000
	Lower-bound	6115,880	1,000	6115,880	18,359	,002
Error(time)	Sphericity Assumed	2998,120	36	83,281		
	Greenhouse-Geisser	2998,120	13,367	224,289		
	Huynh-Feldt	2998,120	15,393	194,776		
	Lower-bound	2998,120	9,000	333,124		

We report there df and the results from Greenhouse-Geisser because sphericity is violated

1528.970 ÷.371 = 4117.75

4 x .371(Greenhouse-Geisser, see the previous table) = 1.484

There are significant differences between 1 and 4 and 5

There are significant differences between 2 and 4 and 5

There are significant differences between 3 and 4 and 5

Pairwise Comparisons

Measure: MEASURE_1

(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-2,900	1,656	1,000	-9,011	3,211
	3	-7,900	2,718	,174	-17,929	2,129
	4	-22,600*	3,194	,001	-34,386	-10,814
	5	-28,000*	6,354	,017	-51,445	-4,555
2	1	2,900	1,656	1,000	-3,211	9,011
	3	-5,000	2,380	,651	-13,783	3,783
	4	-19,700*	2,848	,001	-30,209	-9,191
	5	-25,100*	6,457	,037	-48,926	-1,274
3	1	7,900	2,718	,174	-2,129	17,929
	2	5,000	2,380	,651	-3,783	13,783
	4	-14,700*	2,633	,003	-24,416	-4,984
	5	-20,100*	4,792	,023	-37,782	-2,418
4	1	22,600*	3,194	,001	10,814	34,386
	2	19,700*	2,848	,001	9,191	30,209
	3	14,700*	2,633	,003	4,984	24,416
	5	-5,400	4,525	1,000	-22,094	11,294
5	1	28,000*	6,354	,017	4,555	51,445
	2	25,100*	6,457	,037	1,274	48,926
	3	20,100*	4,792	,023	2,418	37,782
	4	5,400	4,525	1,000	-11,294	22,094

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Bonferroni.

ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΓΓΡΑΦΗΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

One-way repeated ANOVA
significant main-effect

One way repeated measures ANOVA was conducted to examine the differences in the mistakes across time for cyclists as a function of fatigue. The results showed a significant effect of time*, F (1.49, 13.67) = 18.36, p < .001. Pairwise comparison showed significant differences between the 3rd and 6th minute, the 9th to 12th, and from 12th to 15th showing that* the more tired the participants, the more mistakes they made.

One way repeated measures ANOVA with 1-one repeated factor and 1-one independent factor

Main effect for the repeated factor (time): is there a change (difference) in our dependent variable for the whole sample regardless of the independent factor (group)?

Main effect for the independent factor (group):are there significant differences for the independent factor (group) regardless of the dependent factor (time)?

Interaction effect: Is there an interaction between the group and the time (differences between the different levels of the time and the groups)?

Example 2.

One way repeated measures ANOVA with 1-one repeated factor and 1-one independent factor

- Significant differences in performance between the baseline trial and post intervention trial using self-talk

Ho = Control = Experimental

Ho = $\mu_{pre} = \mu_{post}$

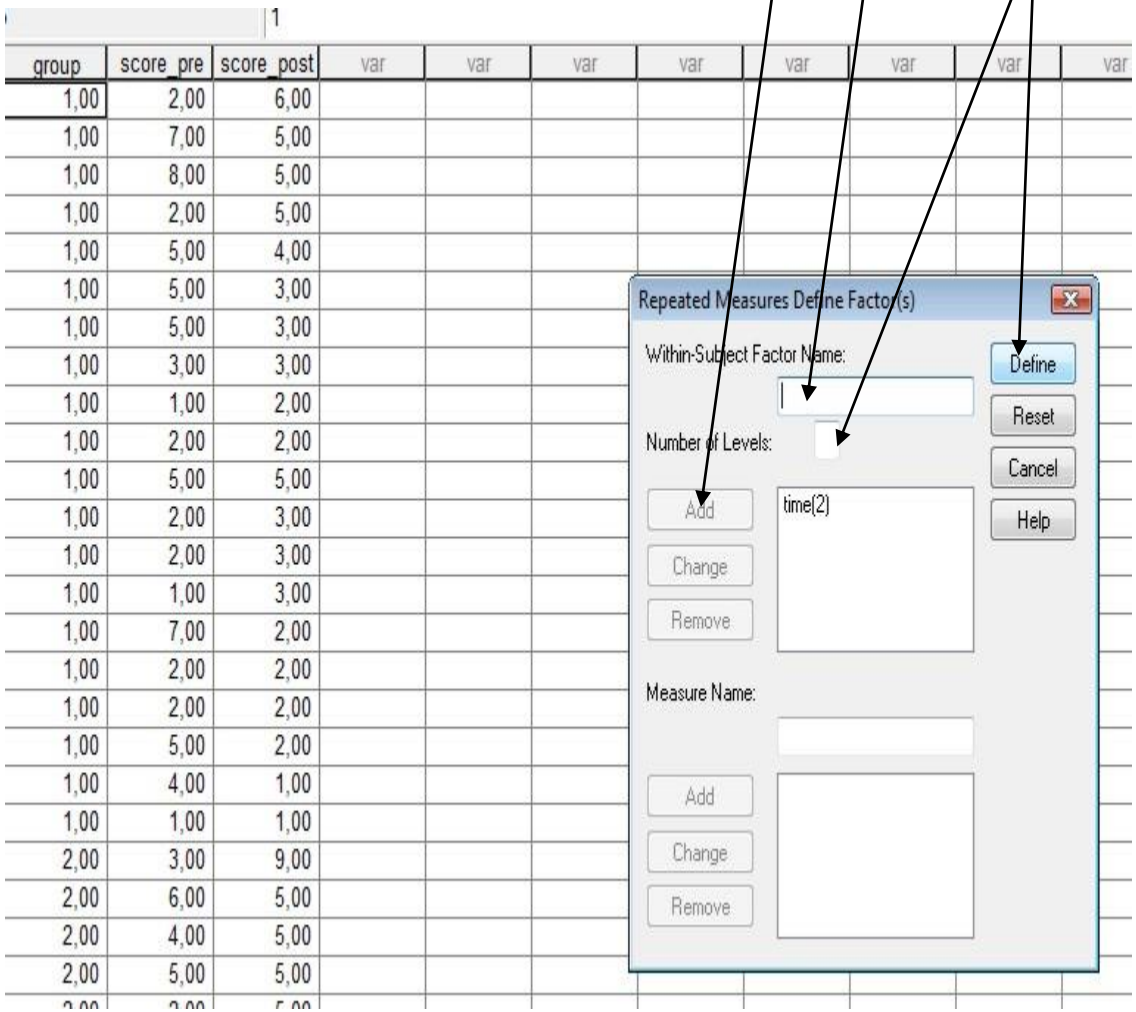
File Edit View Data Transform Analyze Gra

(1) Analyze (2) general linear model (3) repeated

1 : group | 1

	group	score_pre	score_post
1	1,00	2,00	6,00
2	1,00	7,00	5,00
3	1,00	8,00	5,00
4	1,00	2,00	5,00
5	1,00	5,00	4,00
6	1,00	5,00	3,00
7	1,00	5,00	3,00
8	1,00	3,00	3,00
9	1,00	1,00	2,00
10	1,00	2,00	2,00
11	1,00	5,00	5,00
12	1,00	2,00	3,00
13	1,00	2,00	3,00
14	1,00	1,00	3,00
15	1,00	7,00	2,00
16	1,00	2,00	2,00
17	1,00	2,00	2,00
18	1,00	5,00	2,00
19	1,00	4,00	1,00
20	1,00	1,00	1,00
21	2,00	3,00	9,00
22	2,00	6,00	5,00
23	2,00	4,00	5,00
24	2,00	5,00	5,00
25	2,00	2,00	5,00
26	2,00	3,00	5,00
27	2,00	2,00	5,00
28	2,00	6,00	4,00
29	2,00	7,00	4,00
30	2,00	1,00	3,00
31	2,00	6,00	8,00
32	2,00	5,00	7,00

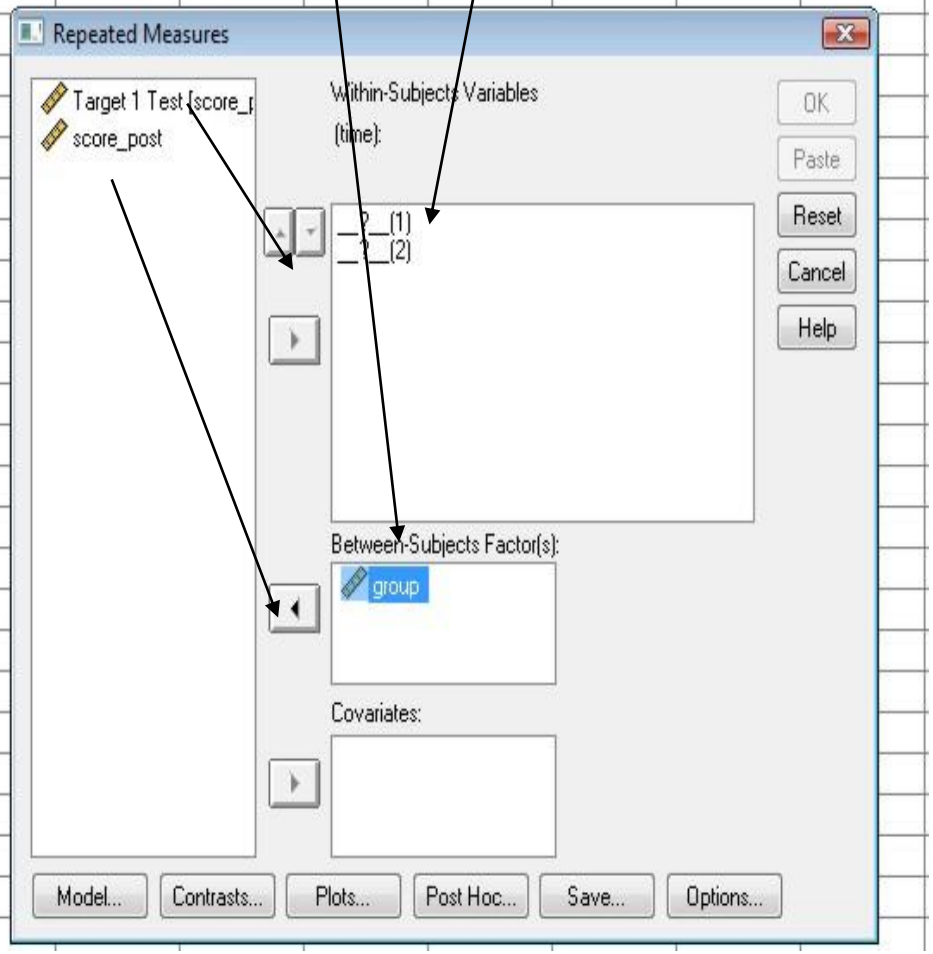
(1) Within subject: time,
(2) number of levels: 2,
(3) add και (4) define

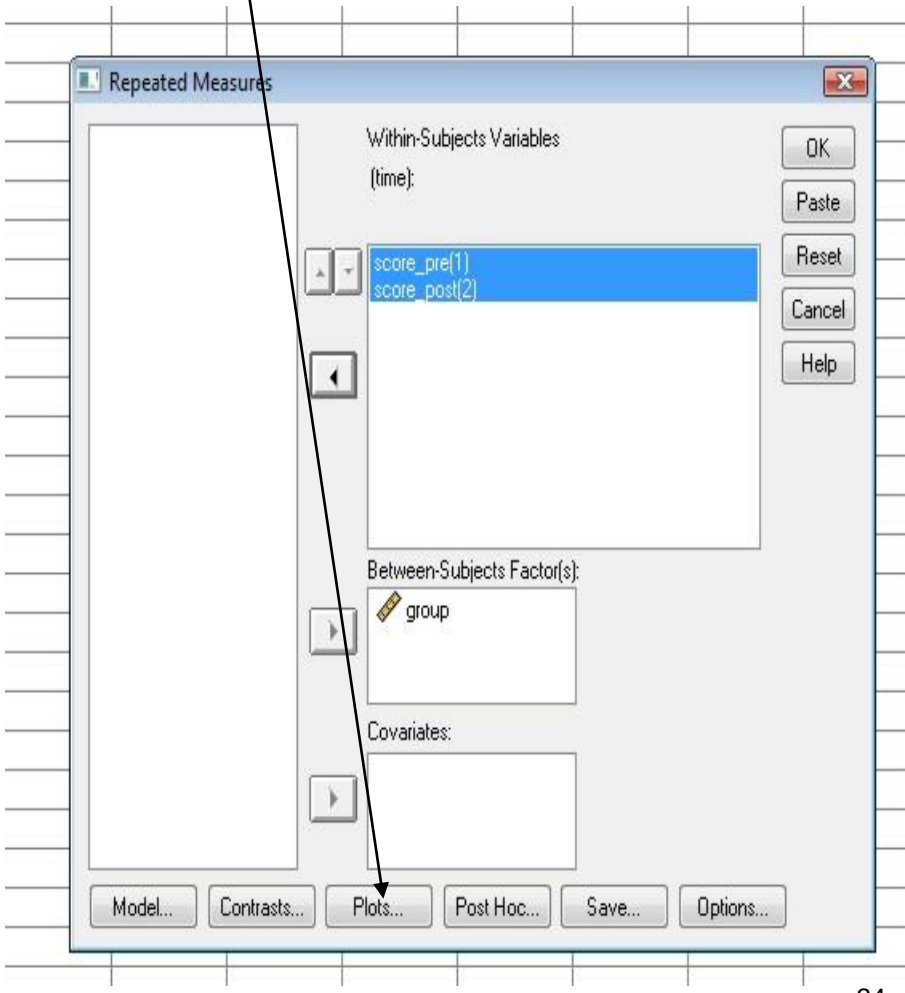


(1) group click to between-subject factor

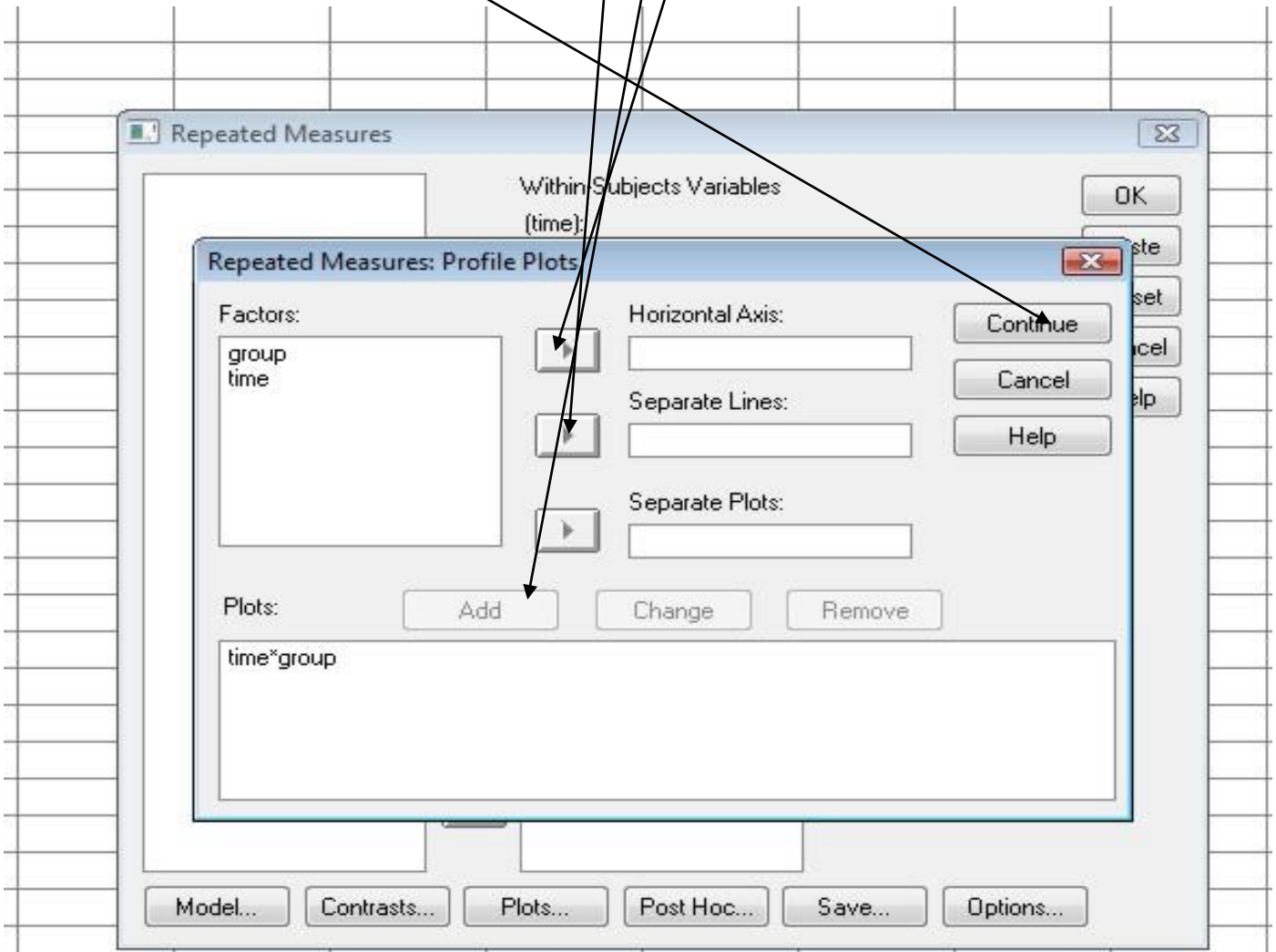
(2) Score pre and score click to the within subject factor

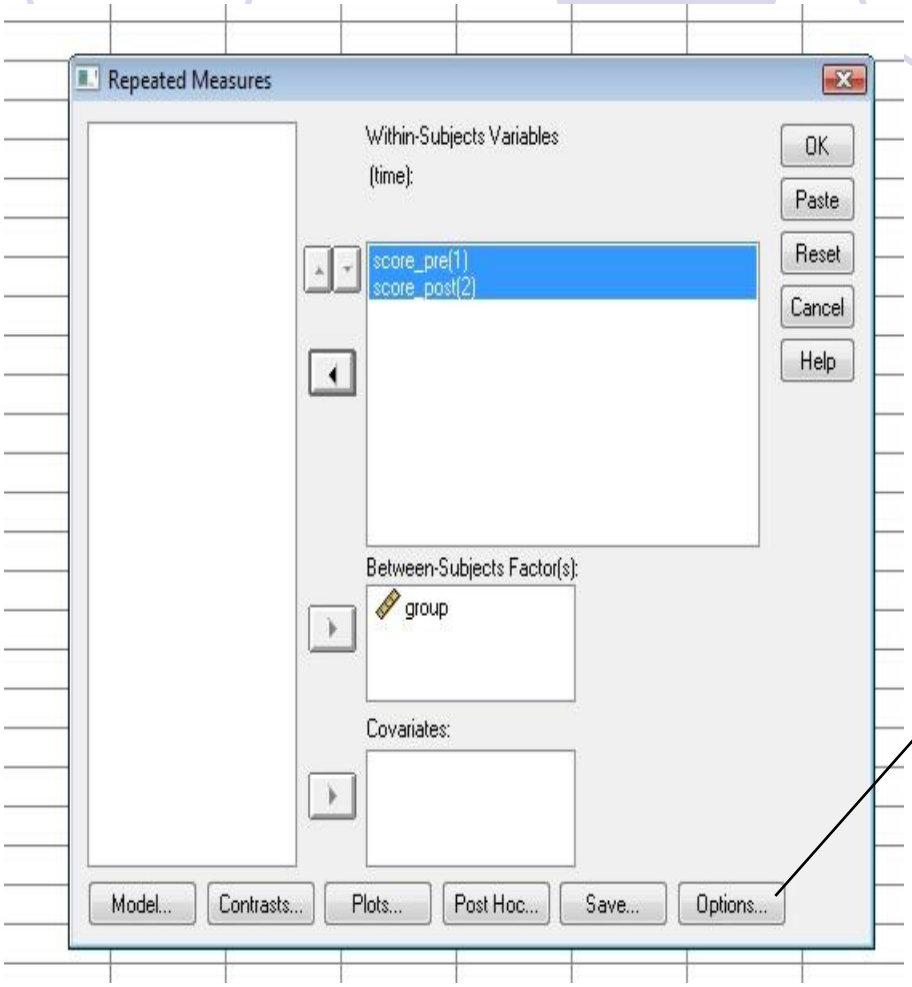
(3)Click Plots





group click to separate lines (1) time click to horizontal axis (2) click add (3) and then click continue (4)



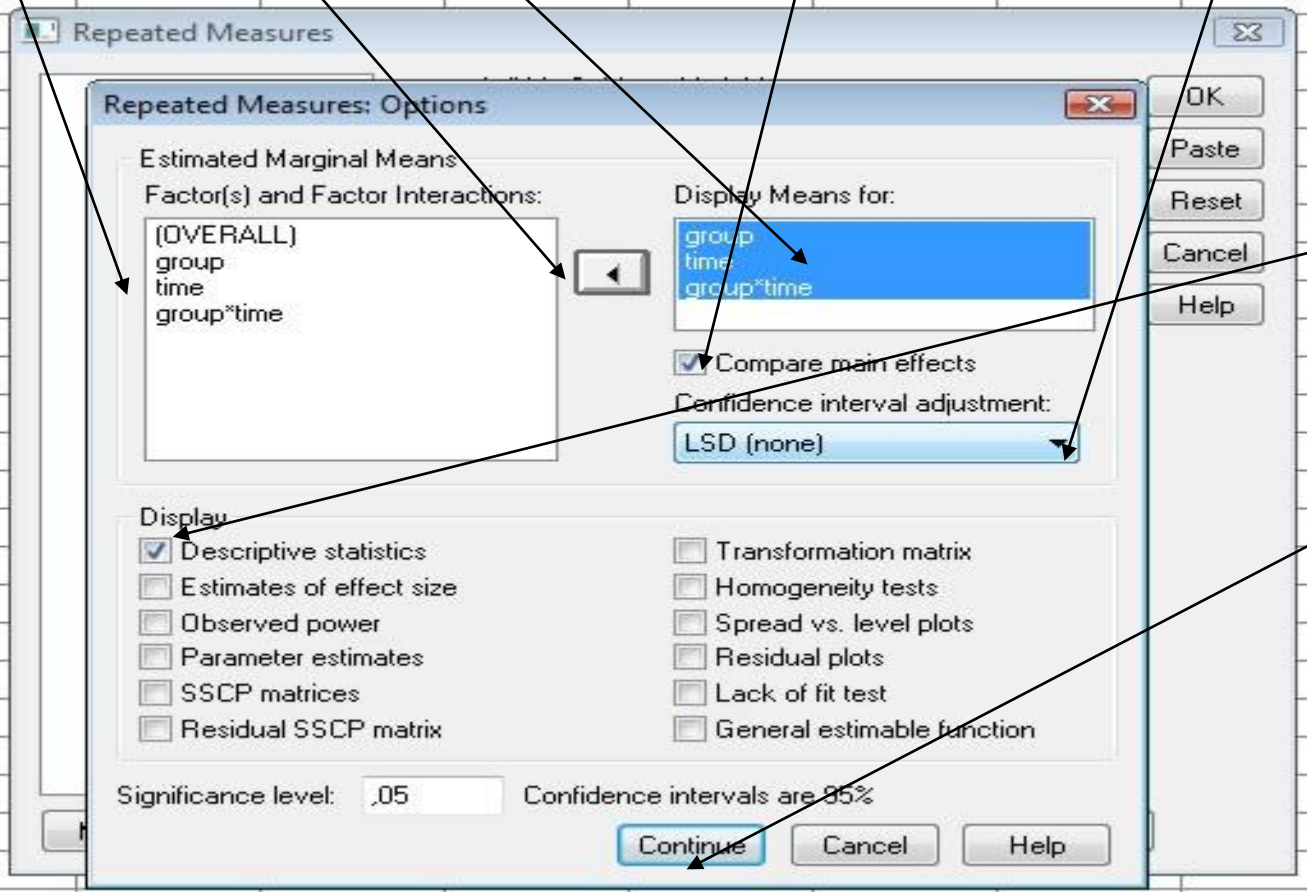


Click options

(1)Click with the (arrow) 3 group, time, group x time in the display mean for

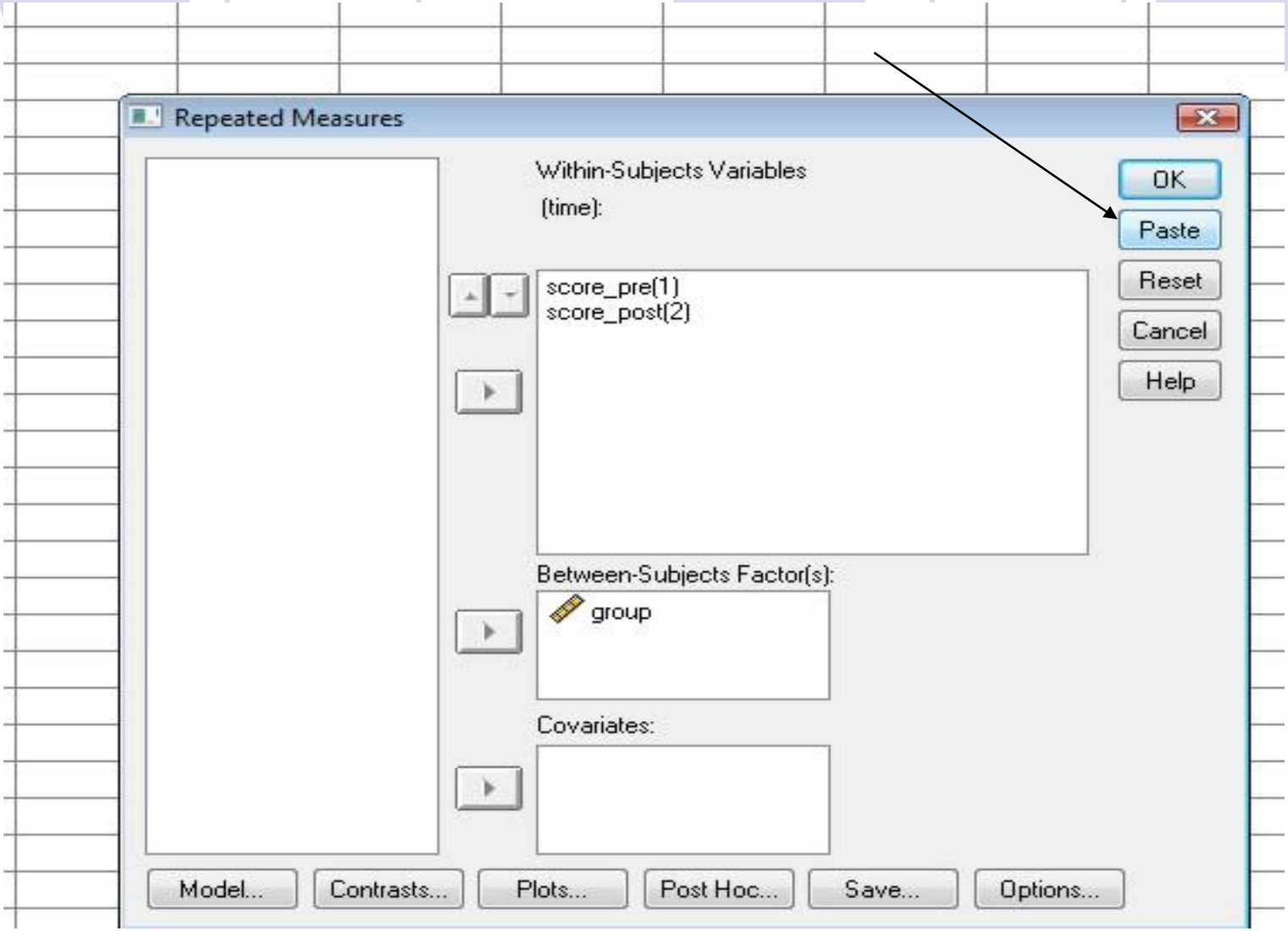
(2) Click compare…

(3) Click LSD…

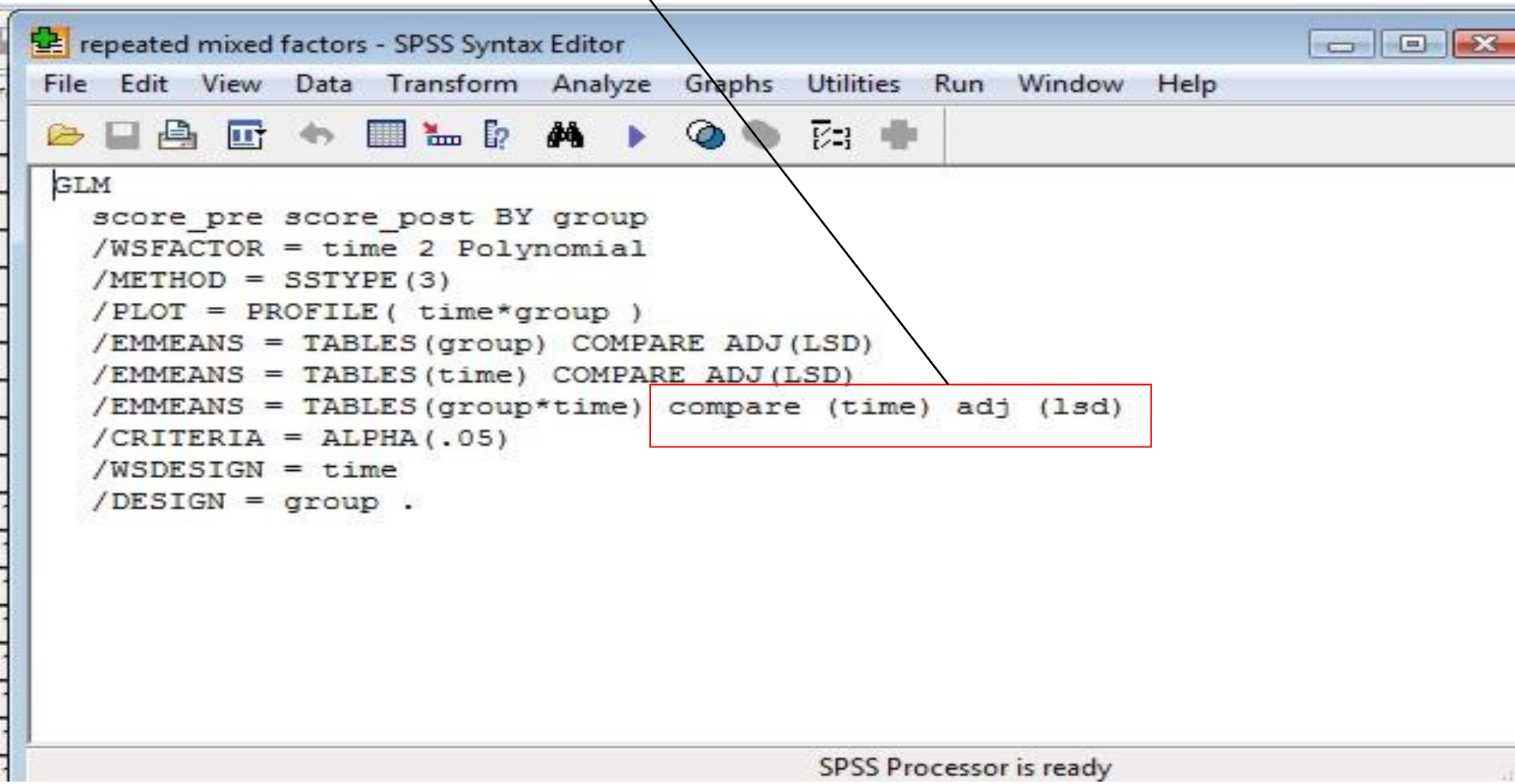


(4) Click Descriptive…

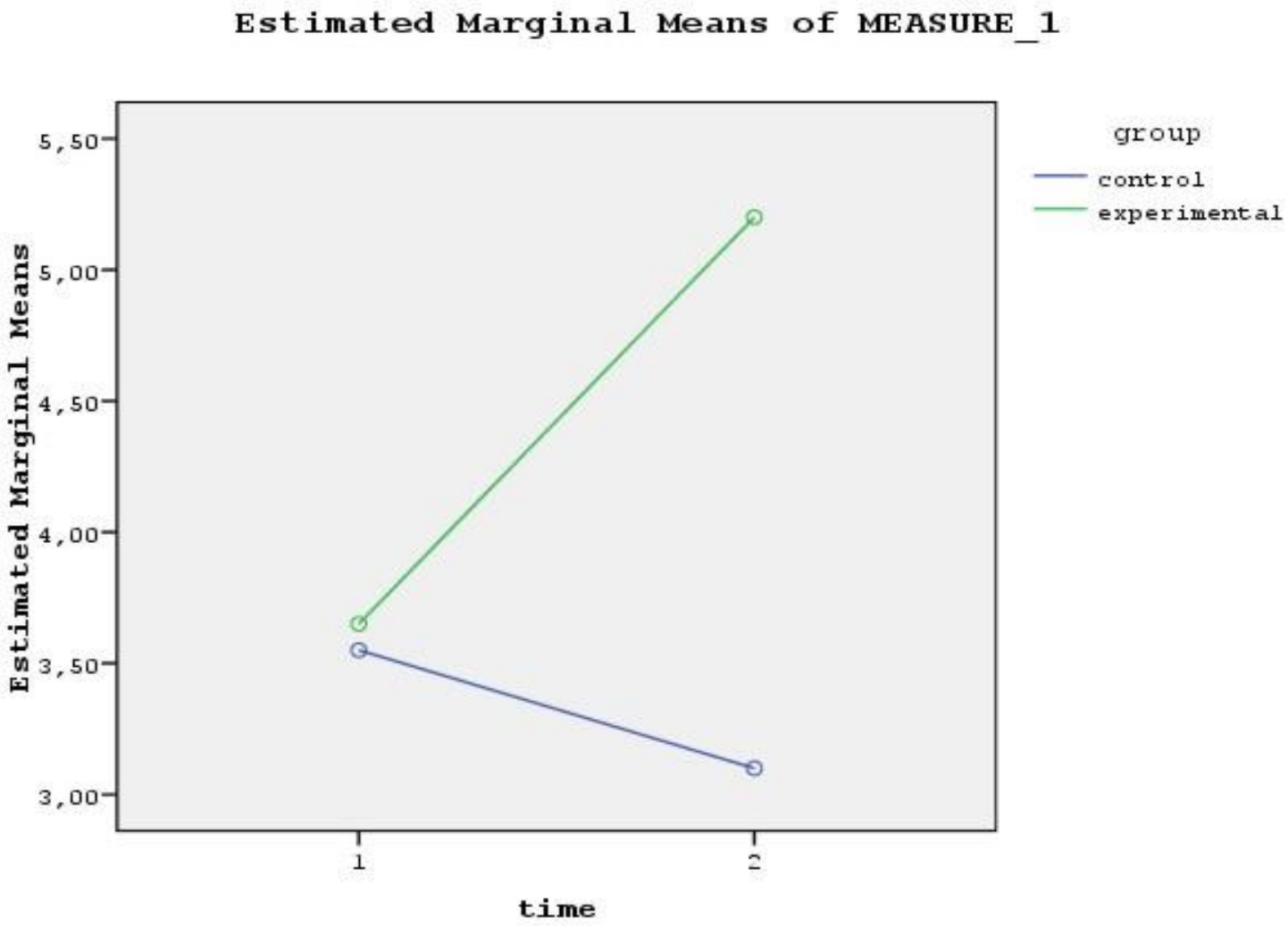
(5) Click continue…


Repeated Measures
Within-Subjects Variables
(time):
score_pre(1)
score_post(2)
Between-Subjects Factor(s):
group
Covariates:
OK
Paste
Reset
Cancel
Help
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...

We write compare (time) adj (lsd)



Profile Plots



There are not significant differences between the 1 σε 2 measure for both groups (total mean for all participants)

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
time	Pillai's Trace	,064	2,585[a]	1,000	38,000	,116
	Wilks' Lambda	,936	2,585[a]	1,000	38,000	,116
	Hotelling's Trace	,068	2,585[a]	1,000	38,000	,116
	Roy's Largest Root	,068	2,585[a]	1,000	38,000	,116
time * group	Pillai's Trace	,184	8,544[a]	1,000	38,000	,006
	Wilks' Lambda	,816	8,544[a]	1,000	38,000	,006
	Hotelling's Trace	,225	8,544[a]	1,000	38,000	,006
	Roy's Largest Root	,225	8,544[a]	1,000	38,000	,006

a. Exact statistic

b. Design: Intercept+group
Within Subjects Design: time

There is significant interaction effect but we don't know in which group there is significant difference between the first and the second measure….so we look in pairwise comparisons (LSD)

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1,000	,000	0	.	1,000	1,000	1,000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

F(1,38) = 8.54, p < .05

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	6,050	1	6,050	2,585	,116
	Greenhouse-Geisser	6,050	1,000	6,050	2,585	,116
	Huynh-Feldt	6,050	1,000	6,050	2,585	,116
	Lower-bound	6,050	1,000	6,050	2,585	,116
time * group	Sphericity Assumed	20,000	1	20,000	8,544	,006
	Greenhouse-Geisser	20,000	1,000	20,000	8,544	,006
	Huynh-Feldt	20,000	1,000	20,000	8,544	,006
	Lower-bound	20,000	1,000	20,000	8,544	,006
Error(time)	Sphericity Assumed	88,950	38	2,341		
	Greenhouse-Geisser	88,950	38,000	2,341		
	Huynh-Feldt	88,950	38,000	2,341		
	Lower-bound	88,950	38,000	2,341		

Estimated Marginal Means

1. group

Mean for the control = 3.33, and experimental = 4.43

There are significant differences between the control and experimental group p < .05

Estimates

Measure: MEASURE_1

group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
control	3,325	,316	2,686	3,964
experimental	4,425	,316	3,786	5,064

Pairwise Comparisons

Measure: MEASURE_1

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
control	experimental	-1,100*	,447	,018	-2,004	-,196
experimental	control	1,100*	,447	,018	,196	2,004

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

2. time

There are not significant differences between the first and second measure for both groups p = .12

Estimates

Measure: MEASURE_1

time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	3,600	,324	2,944	4,256
2	4,150	,231	3,683	4,617

Pairwise Comparisons

Measure: MEASURE_1

(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-,550	,342	,116	-1,243	,143
2	1	,550	,342	,116	-,143	1,243

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

3. group * time

Means for control and experimental group

There are no significant differences in the control group between pre and post measures (p = .36)

There are significant differences for the experimental group between pre and post measure p < .05

Estimates

Measure: MEASURE_1

group	time	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
control	1	3,550	,458	2,623	4,477
	2	3,100	,326	2,439	3,761
experimental	1	3,650	,458	2,723	4,577
	2	5,200	,326	4,539	5,861

For the experimental group we notice that the mean of second measurement is bigger than the mean of the first measurement

Pairwise Comparisons

Measure: MEASURE_1

group	(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
control	1	2	,450	,484	,358	-,529	1,429
	2	1	-,450	,484	,358	-1,429	,529
experimental	1	2	-1,550*	,484	,003	-2,529	-,571
	2	1	1,550*	,484	,003	,571	2,529

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Reporting the results

Two-way repeated ANOVA
with one repeated factor and one independent factor

Non significant main effect (time)*
Significant interaction effect (time x group)

Two-way repeated measures ANOVA was conducted to examine differences in performance before and after the self-talk intervention between the control and experimental group The results showed significant interaction effect, F (1, 38) = 8.54, p < .01. Pairwise comparisons showed that performance for the experimental group increased significantly (p < .01) whereas for control group performance did not show any significant change (p = .36).

* When there is not main effect but there is significant interaction effect we do not present the main effect but we present and discuss the interaction effect.