- **The data are given in the table below:**

Subject	Day 0	Day 1	Day 2	Day 7	Subject	Day 0	Day 1	Day 2	Day 7
1	108	63	45	42	9	106	65	49	49
2	112	75	56	52	10	110	70	46	47
3	114	75	51	46	11	120	85	60	62
4	129	87	69	69	12	118	78	51	56
5	115	71	52	54	13	110	65	46	47
6	122	80	68	68	14	132	92	73	63
7	105	71	52	54	15	127	90	73	68
8	117	77	54	61					

- **The subjects are not grouped (single group).**
- **There is one repeated measures factor Time – with levels:**
  - Day 0,
  - Day 1,
  - Day 2,
  - Day 7

THE ANOVA TABLE FOR ENZYME EXPERIMENT

Source	SS	df	MS	F	p-value
Subject	4221.100	14	301.507	32.45	0.0000
Day	36282.267	3	12094.089	1301.66	0.0000
ERROR	390.233	42	9.291		

The Subject Source of variability is modelling the variability between subjects.

The ERROR Source of variability is modelling the variability within subjects.

EXAMPLE (2)

- We might compare a drug that is supposed to reduce cholesterol to placebo where cholesterol is measured every two months over a 12-month period.

- This differs from a block design approach in that we are interested in comparing treatment groups not just looking at whether there has been a change over time a SINGLE group of subjects, i.e.

$$H_o : \mu_{Time\,1} = \mu_{Time\,2} = \Lambda\ \mu_{Time\,T}$$

- **HYPOTHETICAL CHOLESTEROL STUDY**

<table>
<tr><th>Group</th><th>Initial</th><th>2 mo.</th><th>4 mo.</th><th>6 mo.</th><th>8 mo.</th><th>10 mo.</th><th>12 mo.</th></tr>
<tr><td>DRUG</td><td colspan="7" rowspan="2">The individual(s) who measure cholesterol levels at each follow-up are blind to which group the subjects are in.</td></tr>
<tr><td>Placebo</td></tr>
</table>

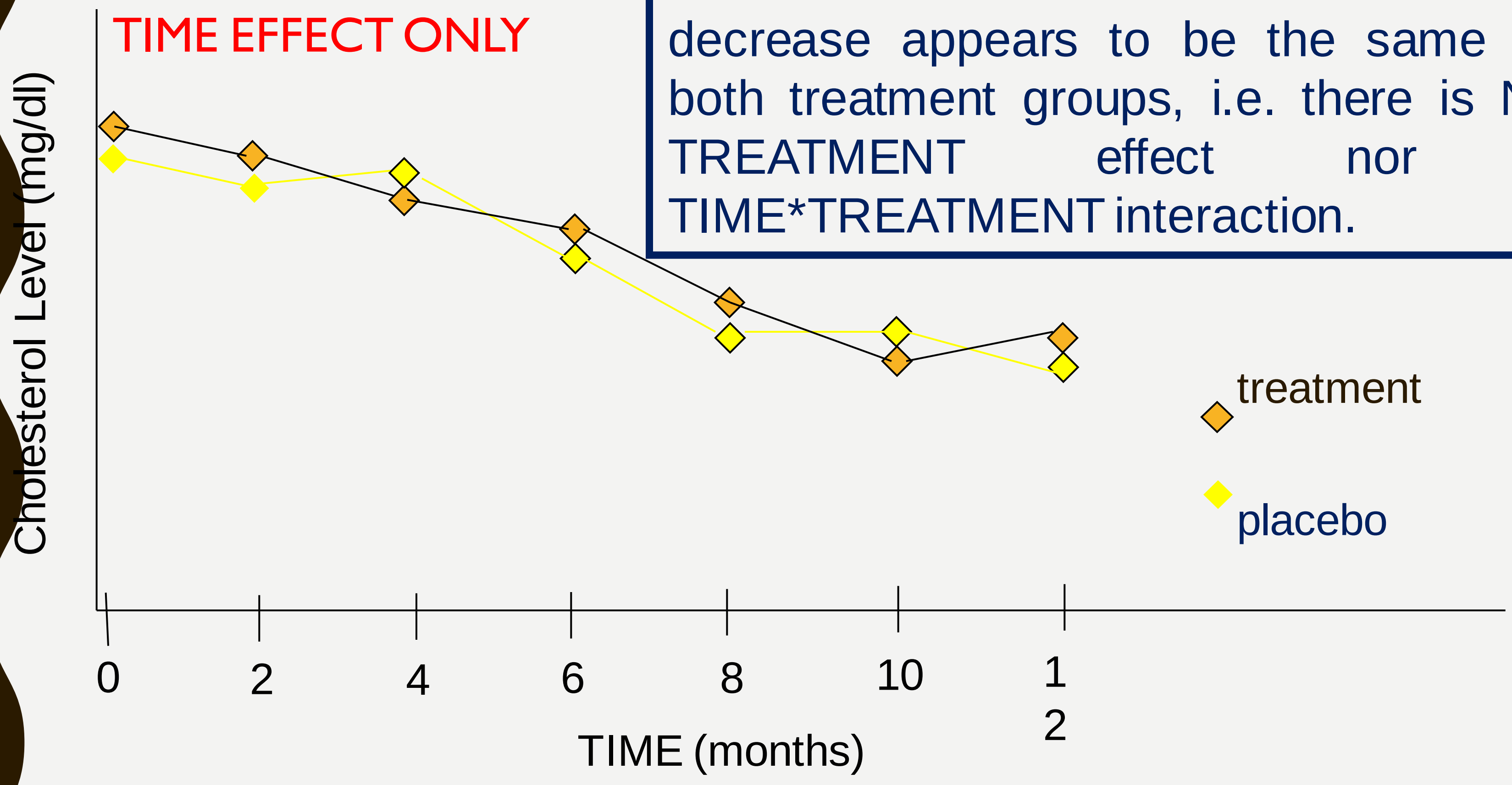
QUESTIONS OF INTEREST:

(1) Is there a change in the cholesterol levels of subjects over time, i.e. is there a TIME EFFECT? (within-subjects effect)

(2) Is there a TREATMENT EFFECT? (between-subjects effect)

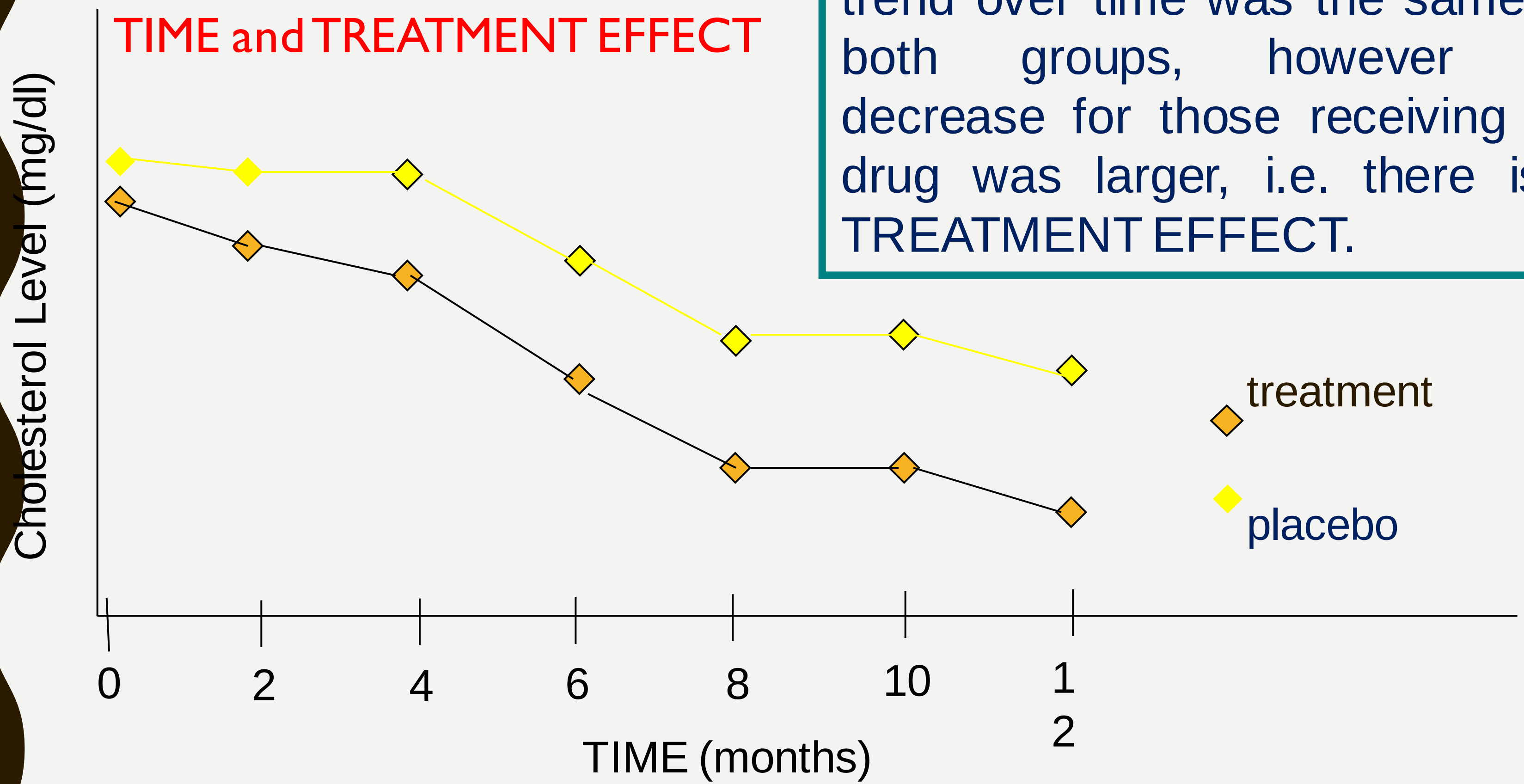
(3) Is the effect of TIME the same for both TREATMENTS? (within-subjects effect)

PROFILE PLOTS ILLUSTRATING THE QUESTIONS OF INTEREST



Cholesterol Levels for both groups decreased over TIME however the decrease appears to be the same for both treatment groups, i.e. there is NO TREATMENT effect nor a TIME*TREATMENT interaction.

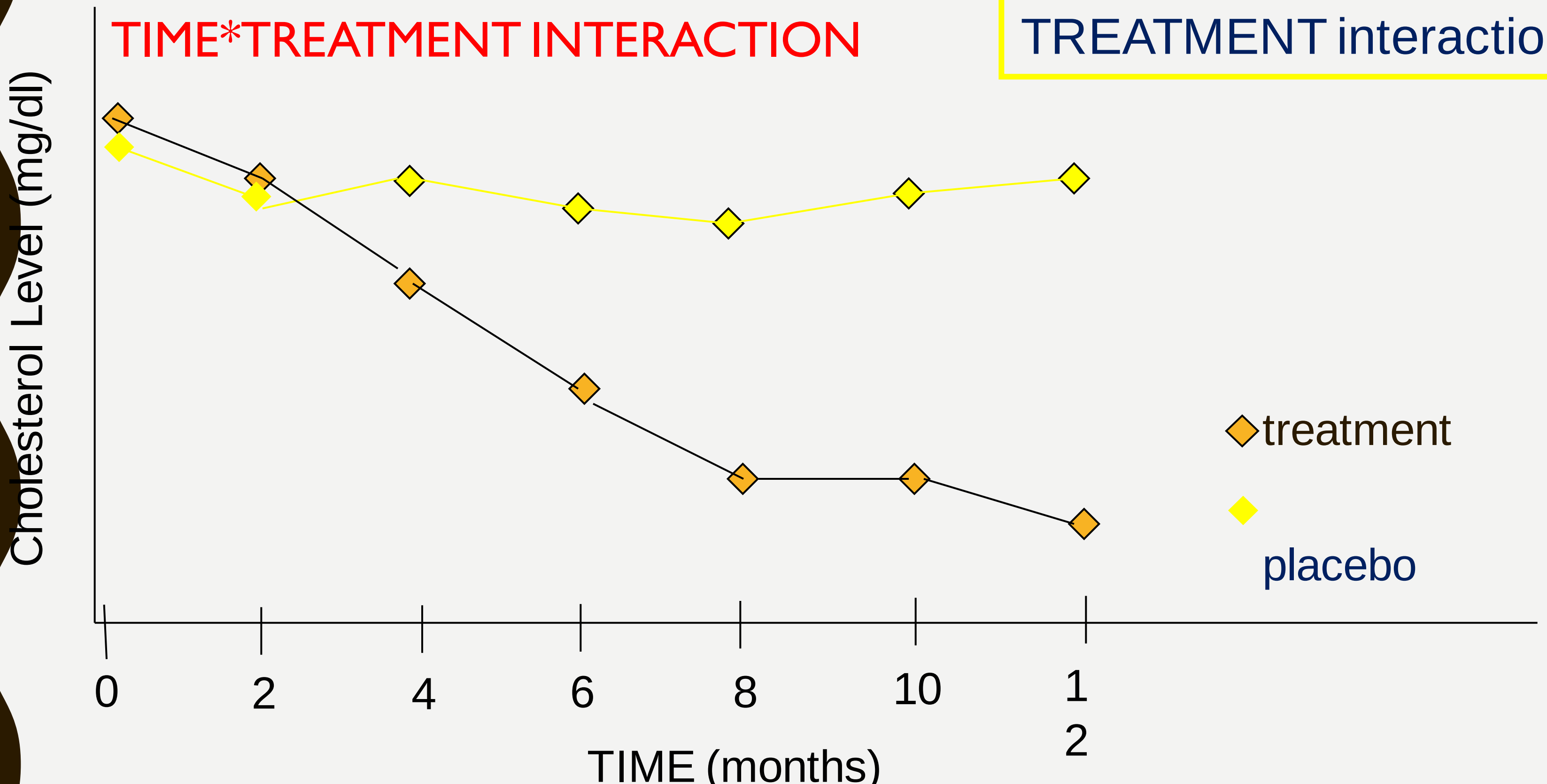
PROFILE PLOTS ILLUSTRATING THE QUESTIONS OF INTEREST



Cholesterol Levels for both groups decreased over TIME and the trend over time was the same for both groups, however the decrease for those receiving the drug was larger, i.e. there is a TREATMENT EFFECT.

PROFILE PLOTS ILLUSTRATING THE QUESTIONS OF INTEREST

Here the effect of time is NOT the same for both groups. Thus we say that there is TIME and TREATMENT interaction.

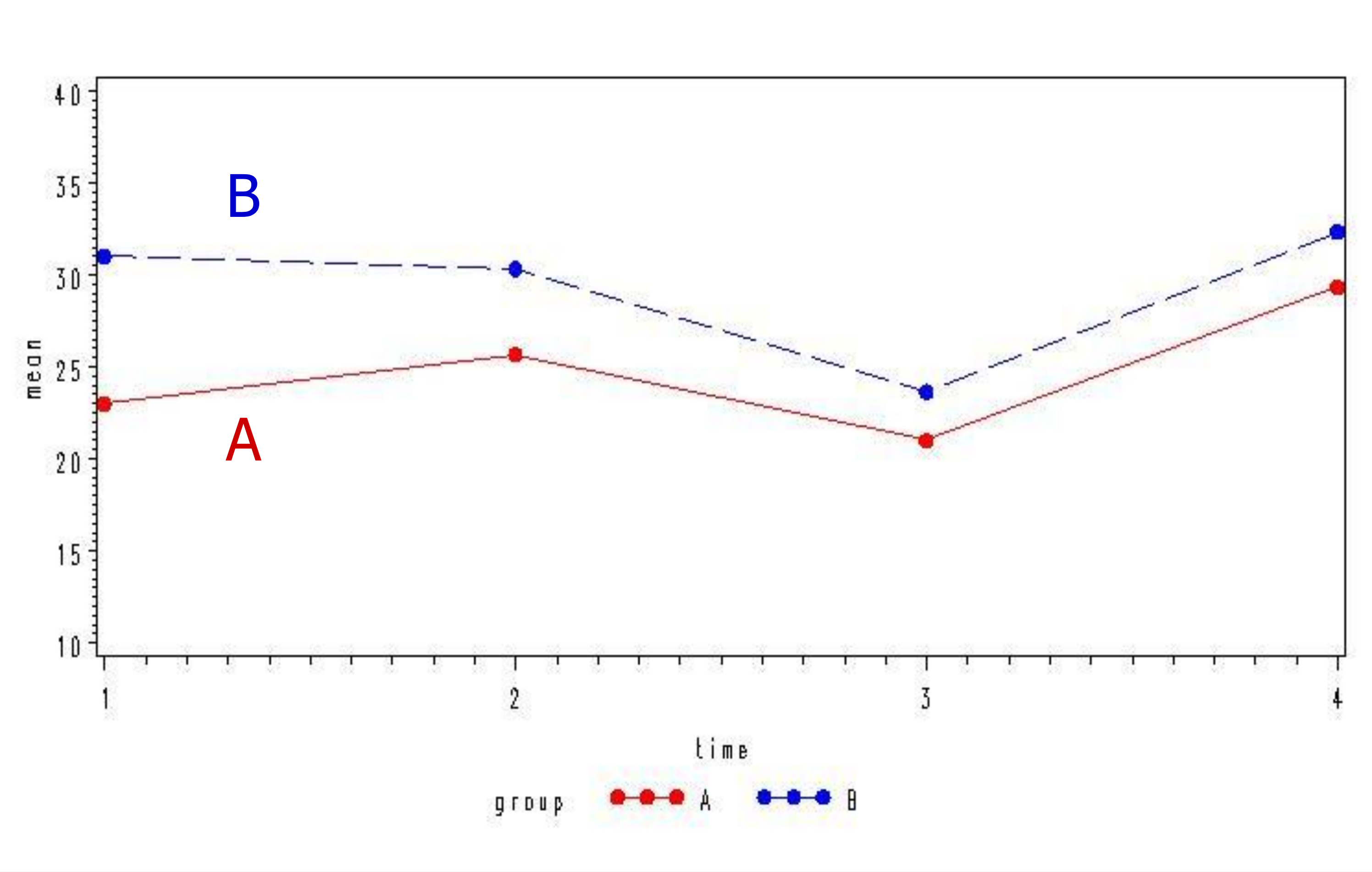


EXAMPLE (3)

Two treatment groups with four measurements taken over equally spaced time intervals (e.g., A = treatment B = placebo)

ID	Group	Time 1	Time 2	Time 3	Time 4
1	A	31	29	15	26
2	A	24	28	20	32
3	A	14	20	28	30
4	B	38	34	30	34
5	B	25	29	25	29
6	B	30	28	16	34

MEAN PROFILE PLOTS BY GROUP



QUESTIONS OF INTEREST

(1) OVERALL, ARE THERE SIGNIFICANT DIFFERENCES BETWEEN TIME POINTS?

From plots it looks like some differences over time, in particular times 3 and 4 look different.

(2) DO THE TWO GROUPS DIFFER AT ANY TIME POINTS, i.e. IS THERE A TREATMENT EFFECT?

From plots it looks like the groups differ at baseline and there are some difference everywhere else.

(3) DO THE TWO GROUPS DIFFER IN THEIR RESPONSES OVER TIME, I.E IS THERE A TIME*TREATMENT INTERACTION?

Their response profiles looks similar over time, though A and B are closer by the end.

NULL HYPOTHESIS SIGNIFICANCE TESTING

- ***Step 1:***

  H_0 – All of the condition means are equal
- ***Step 2:***

  H_{alt} – At least one condition mean is significantly different from the others
- ***Step 3:***

  Collect your data
- ***Step 4:***

  Run the ANOVA
- ***Step 5:***

  Obtain the F statistic and associated p value
- ***Step 6:***

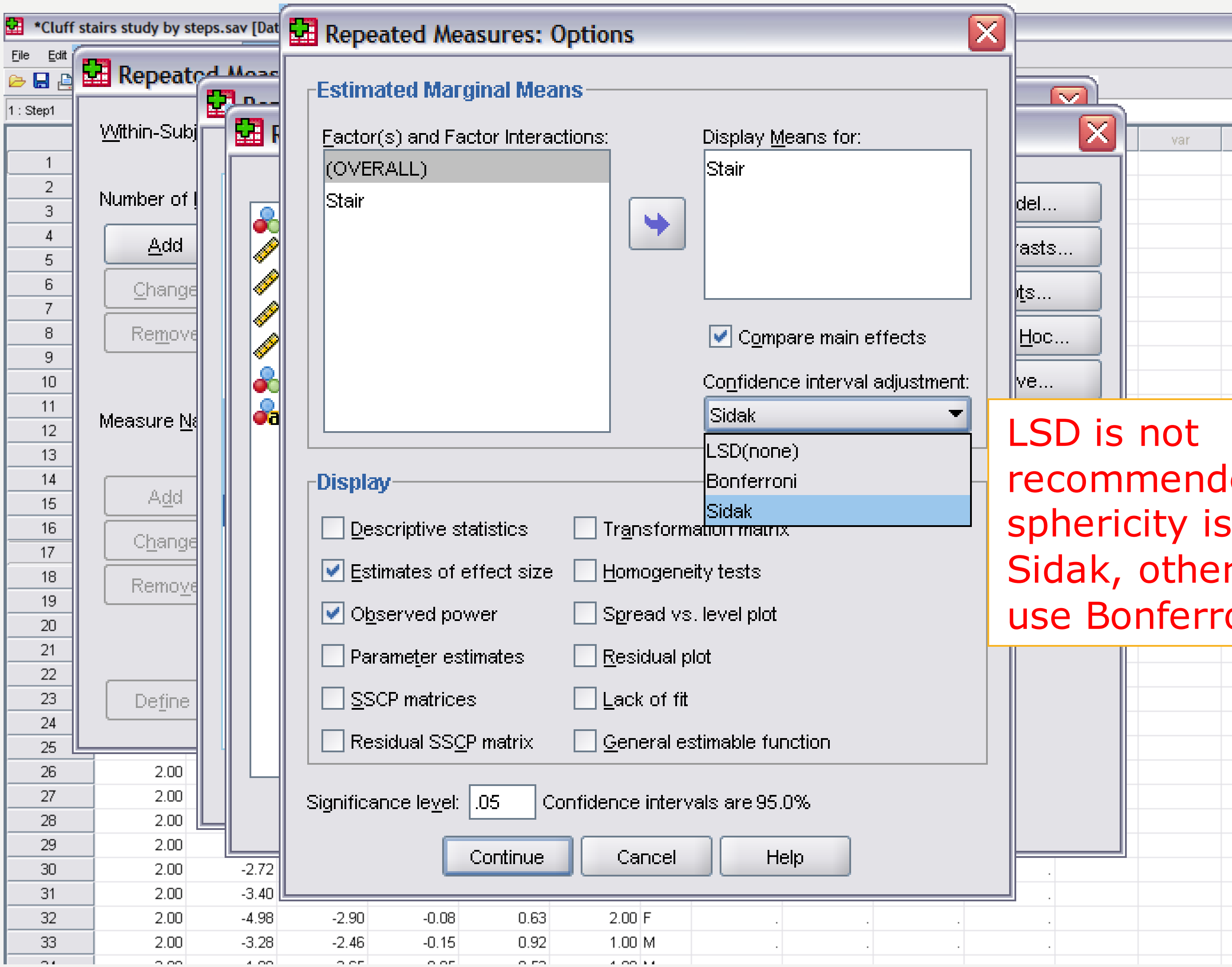
Decide whether to reject or fail to reject H_0 on the basis of the p value

POST HOC TESTING

- **Significant F value**
  - **At least one condition mean is significantly different from the others**
- **But which one?**
- **Posthoc tests:**
  - **Bonferroni**
  - **Tukey**
  - **Sidak**

GENERAL EXAMPLE

REPEATED-MEASURES ANOVA



LSD is not recommended, if sphericity is OK use Sidak, otherwise use Bonferroni.

REPEATED-MEASURES ANOVA: TEST FOR SPHERICITY

SPSS uses Mauchly's Test of Sphericity.

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subject...	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Stair	.585	7.893	5	.163	.723	.841	.333

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 Tests of Within-Subjects Effects table.

b. Design:
Within Subjects Design: Stair

Since the p-value (Sig. = 0.163) is greater than $\alpha = 0.05$, we accept the null hypothesis that covariance are equal and can "assume sphericity".

REPEATED-MEASURES ANOVA: RESULTS

SPSS shows results for four different assumptions. We can choose the first.

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Stair	Sphericity Assumed	17.422	3	5.807	27.706	.000
	Greenhouse-Geisser	17.422	2.170	8.029	27.706	.000
	Huynh-Feldt	17.422	2.523	6.907	27.706	.000
	Lower-bound	17.422	1.000	17.422	27.706	.000
Error(Stair)	Sphericity Assumed	10.061	48	.210		
	Greenhouse-Geisser	10.061	34.717	.290		
	Huynh-Feldt	10.061	40.360	.249		
	Lower-bound	10.061	16.000	.629		

Since the p-value (Sig. = .000) is less than $\alpha = 0.05$, the null hypothesis is rejected and conclude there is a significant difference across stair steps. Note, a p-value of 0.000 is written $p<0.0005$.

REPEATED-MEASURES ANOVA: TEST FOR BEST FIT

SPSS shows results of fitting polynomials from linear to degree *k*-1.

Tests of Within-Subjects Contrasts

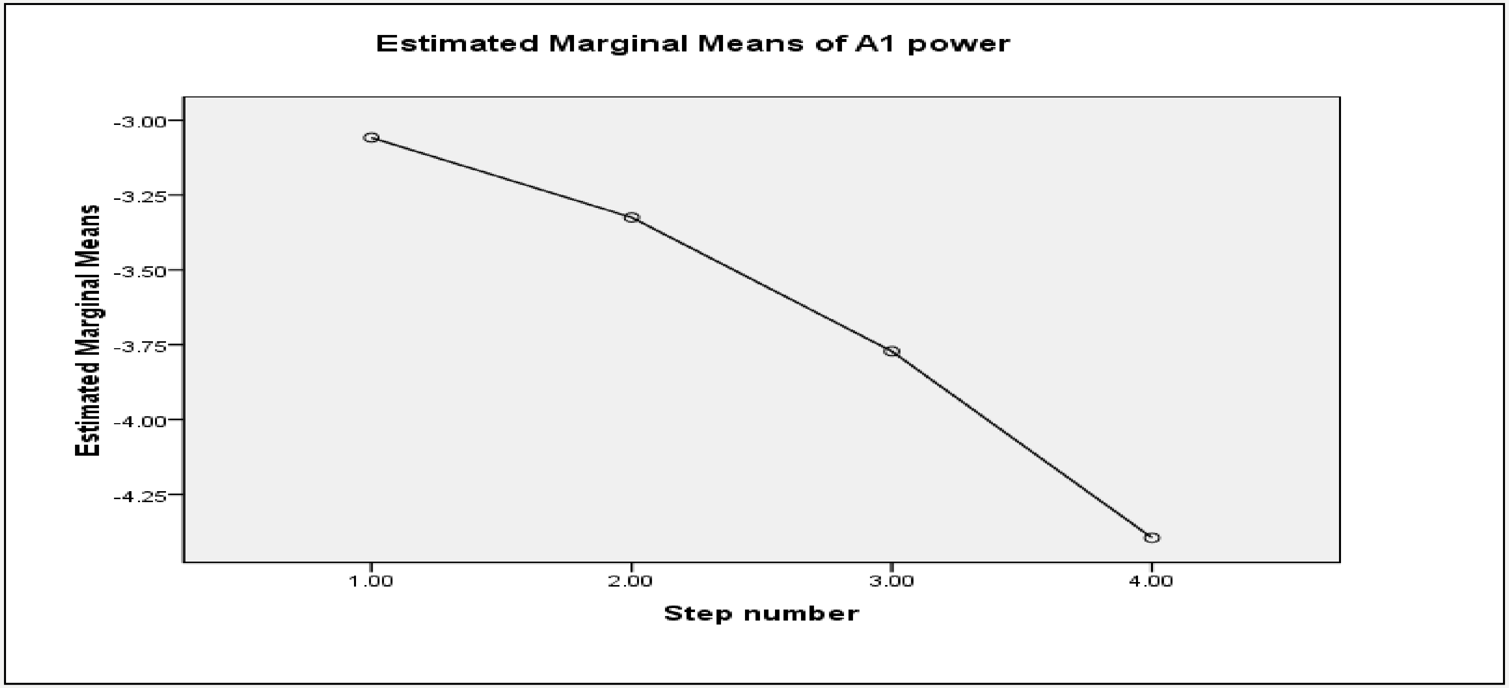
Measure:MEASURE_1

Source	Stair	Type III Sum of Squares	df	Mean Square	F	Sig.
Stair	Linear	16.882	1	16.882	71.498	.000
	Quadratic	.540	1	.540	3.207	.092
	Cubic	2.175E-5	1	2.175E-5	.000	.992
Error(Stair)	Linear	3.778	16	.236		
	Quadratic	2.696	16	.169		
	Cubic	3.587	16	.224		

Since there are only 4 steps, SPSS only tests to a cubic (3rd degree) fit. In this example a linear fit was best. Note, this statistic makes no sense if the DV is not ordered, such as time, age, or date.

REPEATED-MEASURES ANOVA: PLOT OF MARGINAL MEANS

SPSS can plot the group means. This plot shows the means for each step.



Looks like a linear increase in A1 power as people descend the stairs. Note, A1 is a negative power.

REPEATED-MEASURES ANOVA: POST HOC TESTS

- Since there is a significant F we can do *post hoc* testing. If not significant this step IS NOT DONE.
- We will use the **Sidak** *post hoc* test. **Bonferroni** is too conservative. Choose from the **Options…** menu, NOT the **Post Hoc…** menu!

Pairwise Comparisons

Measure:MEASURE_1

(I) Stair	(J) Stair	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	.266	.197	.730	-.326	.858
	3	.714*	.144	.001	.282	1.145
	4	1.336*	.192	.000	.762	1.911
2	1	-.266	.197	.730	-.858	.326
	3	.447*	.132	.023	.051	.843
	4	1.070*	.121	.000	.708	1.432
3	1	-.714*	.144	.001	-1.145	-.282
	2	-.447*	.132	.023	-.843	-.051
	4	.623*	.140	.002	.204	1.042
4	1	-1.336*	.192	.000	-1.911	-.762
	2	-1.070*	.121	.000	-1.432	-.708
	3	-.623*	.140	.002	-1.042	-.204

Based on estimated marginal means

a. Adjustment for multiple comparisons: Sidak.

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

Pairwise Comparisons

Measure:MEASURE_1

(I) Stair	(J) Stair	Mean Difference (I-J)	Std. Error	Sig.a	95% Confidence Interval for Differencea Lower Bound	Upper Bound
1	2	.266	.197	.730	-.326	.858
	3	.714*	.144	.001	.282	1.145
	4	1.336*	.192	.000	.762	1.911
2	1	-.266	.197	.730	-.858	.326
	3	.447*	.132	.023	.051	.843
	4	1.070*	.121	.000	.708	1.432
3	1	-.714*	.144	.001	-1.145	-.282
	2	-.447*	.132	.023	-.843	-.051
	4	.623*	.140	.002	.204	1.042
4	1	-1.336*	.192	.000	-1.911	-.762
	2	-1.070*	.121	.000	-1.432	-.708
	3	-.623*	.140	.002	-1.042	-.204

Based on estimated marginal means

a. Adjustment for multiple comparisons: Sidak.

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

The results now show that steps 1 and 2 are not significantly different for each other but are different from 3 and 4 and steps 3 and 4 are different from all the other steps. This is a better result than the factorial ANOVA.

EXAMPLE:

REPEATED MEASURES

- Subjects undergoing orthodontic treatment, were examined to see the amount of pain perception due to orthodontic treatment. Pain was measured after placing the appliance in different five times (3, 6, 9, 12, 15 minute) periods for the purpose of collecting data. The goal of the experiment was to determine if the pain will progress during treatment or would it stop at any point.

Data file: RM Orthodontic Treatment

Subject	Minute 3	Minute 6	Minute 9	Minute 12	Minute 15
1	7	7	23	36	70
2	12	22	26	26	20
3	11	6	9	31	30
4	10	18	16	40	25
5	6	12	9	28	37
6	13	21	30	55	65
7	5	0	2	10	11
8	15	18	22	37	42
9	0	2	0	16	11
10	6	8	27	32	54
ΣC	85	114	164	311	365
Mean =	8.5	11.4	16.4	31.1	36.5

Repeated Measures ANOVA: Data Entry

RM Orthodontic Treatment.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Visible: 6 of 6 Variables

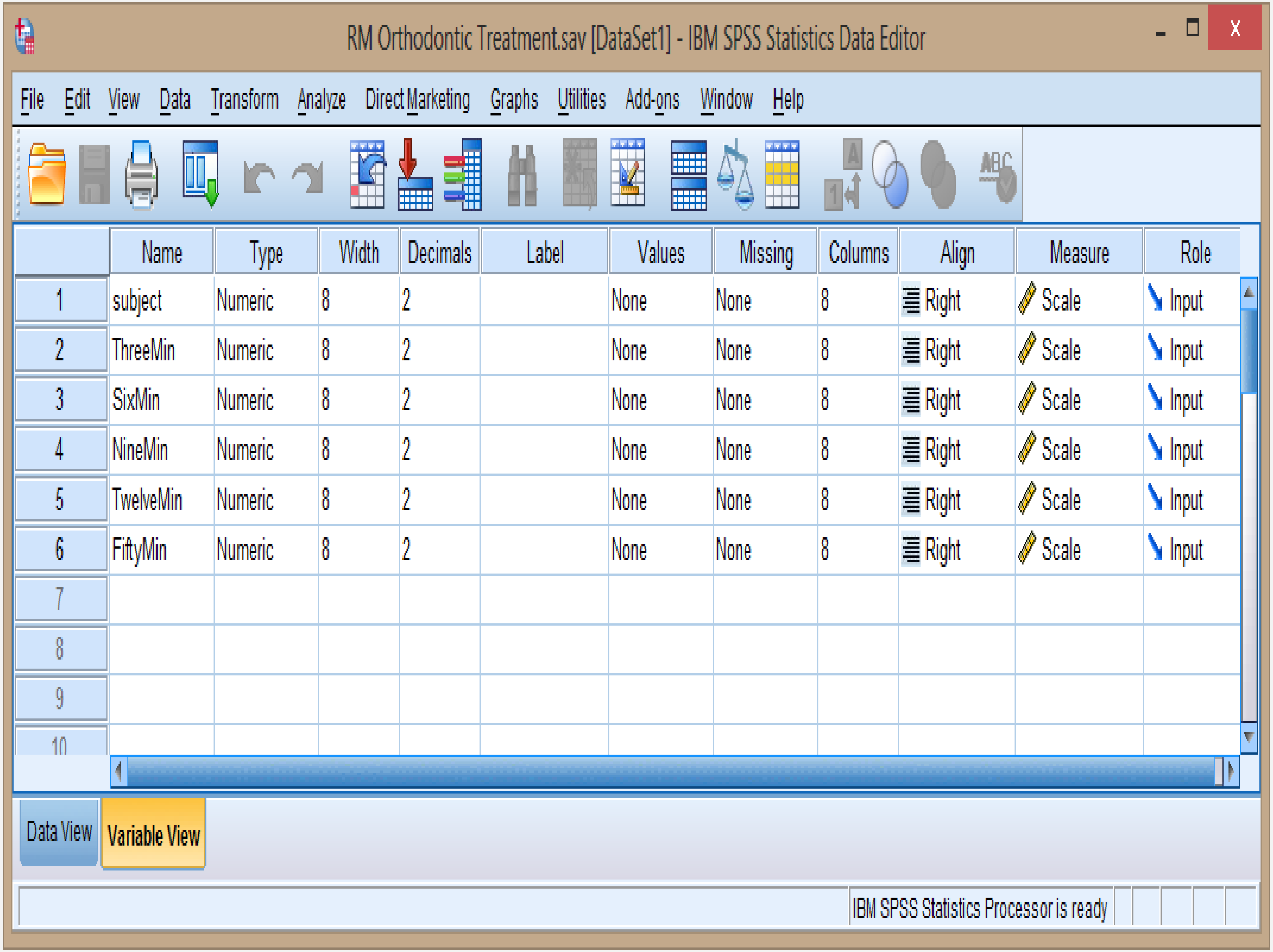
	subject	ThreeMin	SixMin	NineMin	TwelveMin	FiftyMin	var
1	1.00	7.00	7.00	23.00	36.00	70.00	
2	2.00	12.00	22.00	26.00	26.00	20.00	
3	3.00	11.00	6.00	9.00	31.00	30.00	
4	4.00	10.00	18.00	16.00	40.00	25.00	
5	5.00	6.00	12.00	9.00	28.00	37.00	
6	6.00	13.00	21.00	30.00	55.00	65.00	
7	7.00	5.00	.00	2.00	10.00	11.00	
8	8.00	15.00	18.00	22.00	37.00	42.00	
9	9.00	.00	2.00	.00	16.00	11.00	
10	10.00	6.00	8.00	27.00	32.00	54.00	
11							
12							
13							

Data View | Variable View

IBM SPSS Statistics Processor is ready

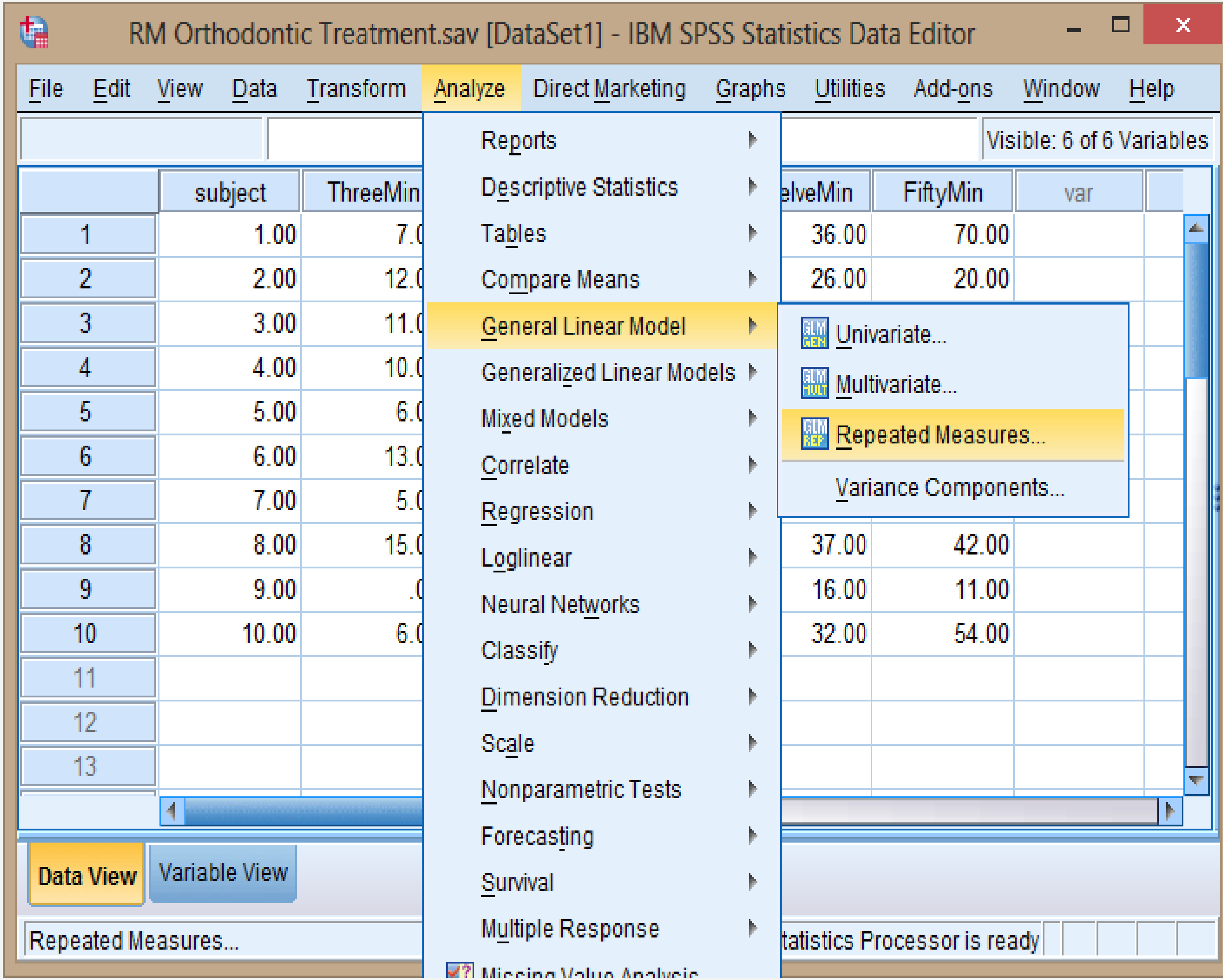
Each level of a within subjects factor is entered as a separate variable.

Fatigue (3, 6, 9, 12, 15 min)

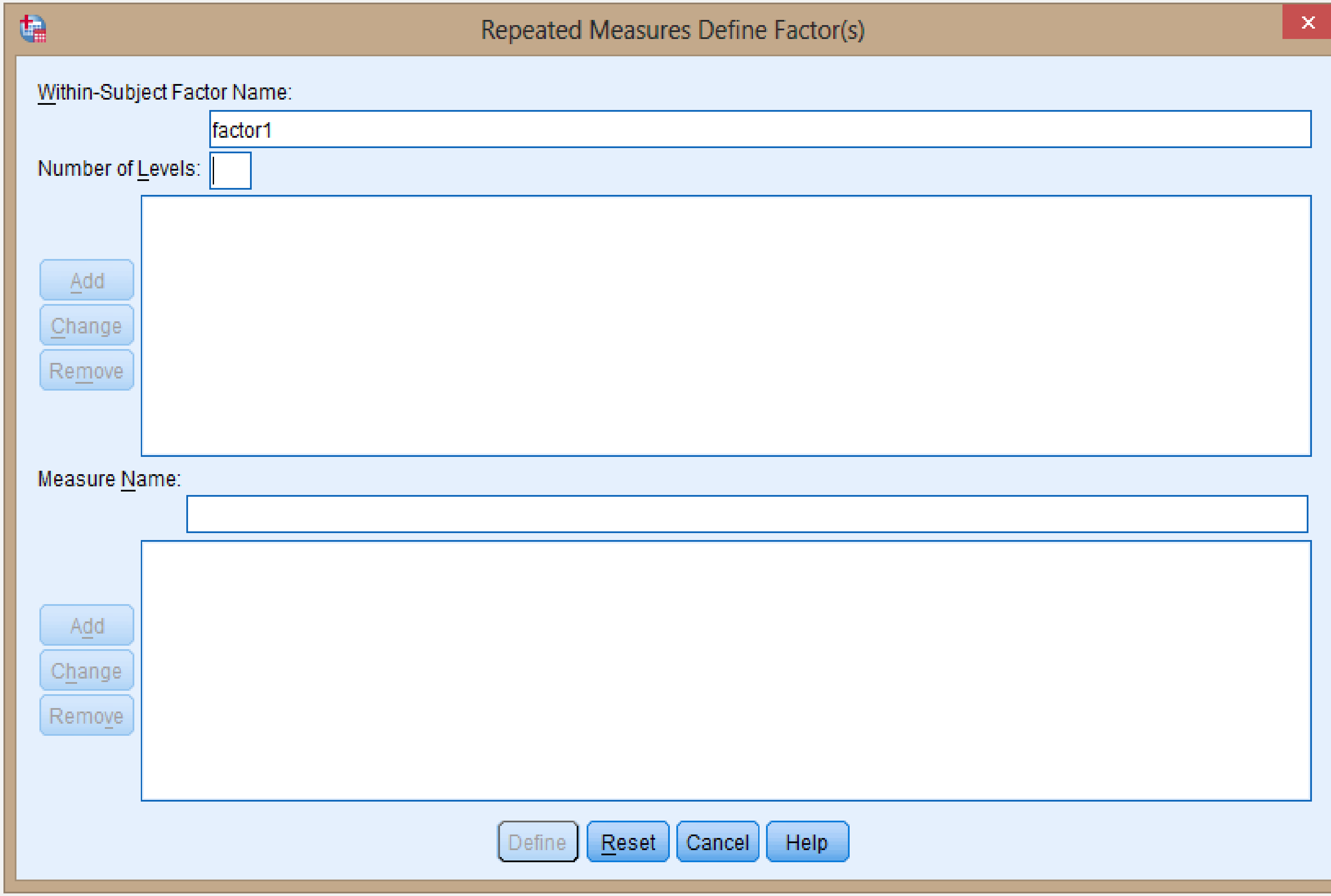

RM Orthodontic Treatment.sav [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Name Type Width Decimals Label Values Missing Columns Align Measure Role
1 subject Numeric 8 2 None None 8 Right Scale Input
2 ThreeMin Numeric 8 2 None None 8 Right Scale Input
3 SixMin Numeric 8 2 None None 8 Right Scale Input
4 NineMin Numeric 8 2 None None 8 Right Scale Input
5 TwelveMin Numeric 8 2 None None 8 Right Scale Input
6 FiftyMin Numeric 8 2 None None 8 Right Scale Input
7
8
9
10
Data View
Variable View
IBM SPSS Statistics Processor is ready

STEPS IN SPSS IS:

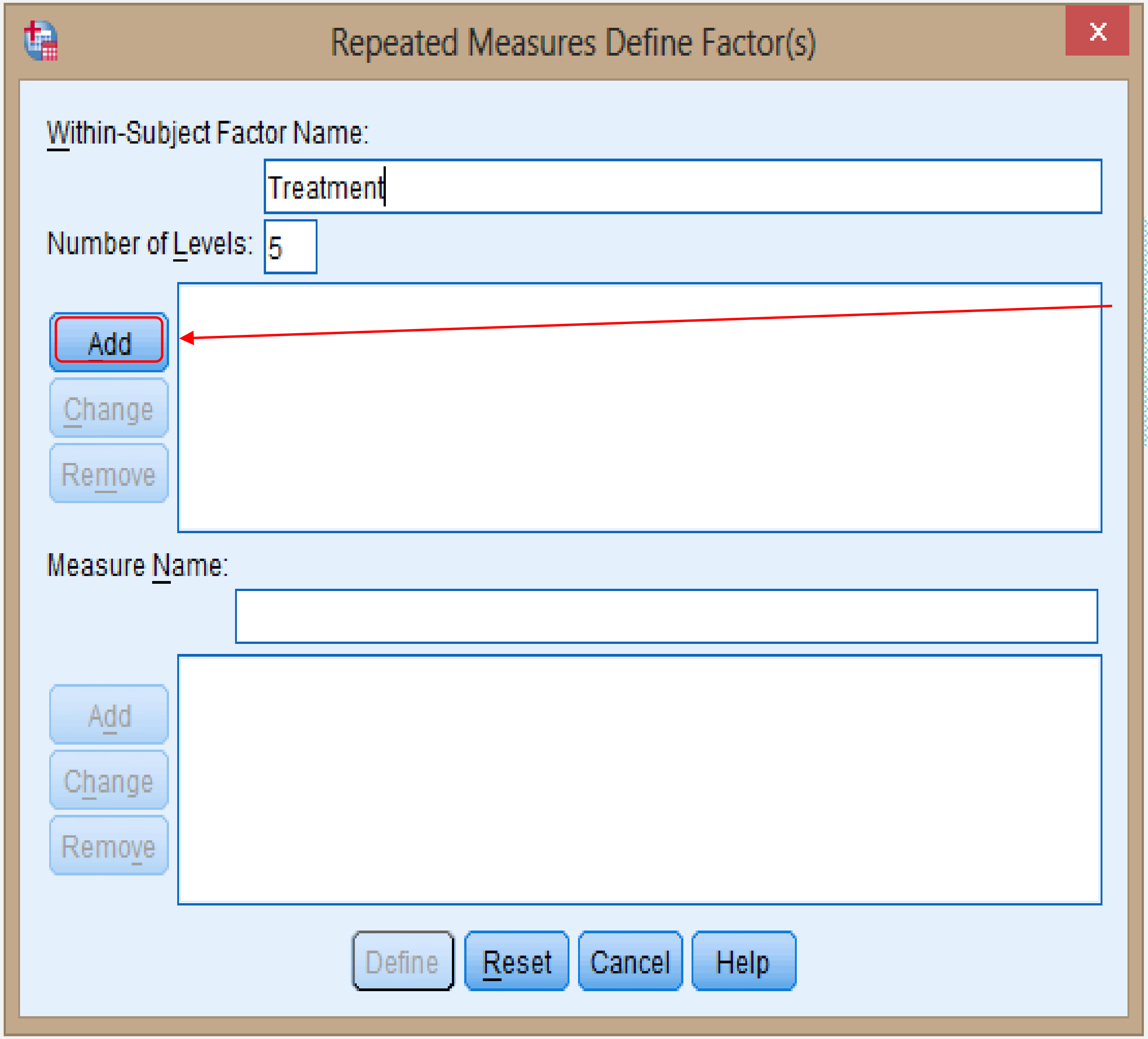
General Linear Model → Repeated Measures



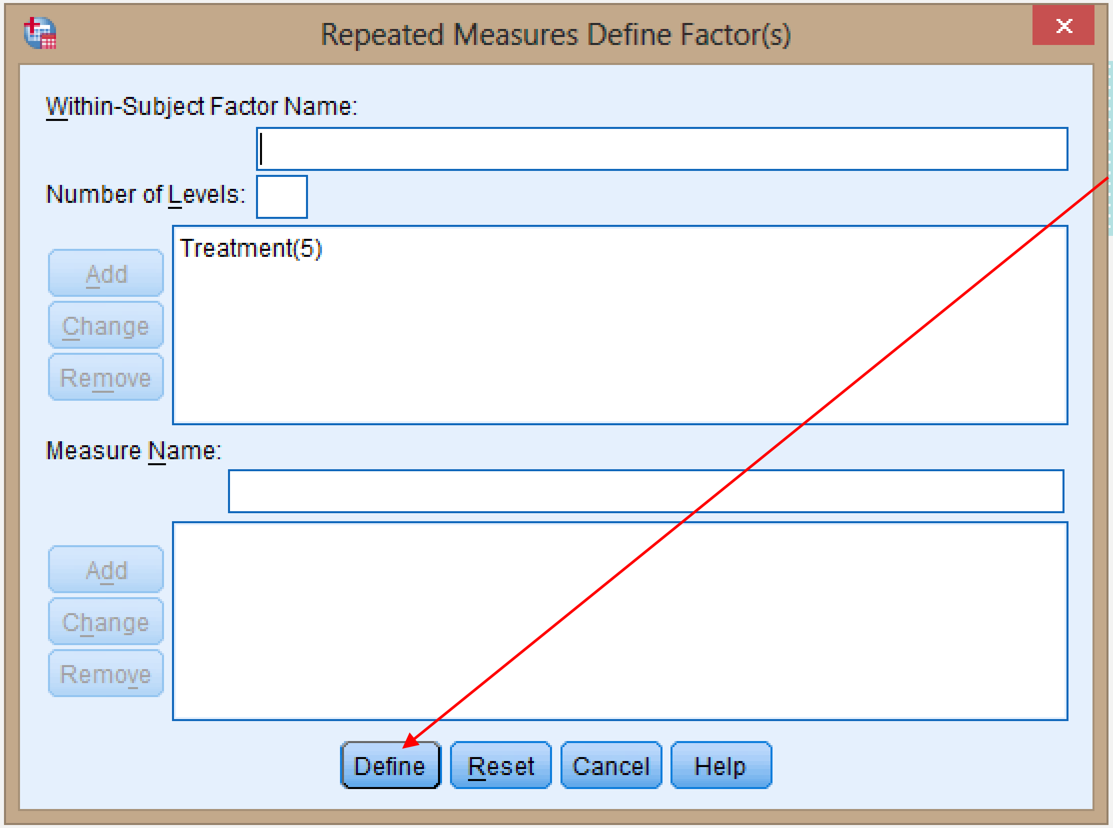
RM Orthodontic Treatment.sav [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Window
Help
Visible: 6 of 6 Variables
subject
ThreeMin
elveMin
FiftyMin
var
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Univariate...
Multivariate...
Repeated Measures...
Variance Components...
Data View
Variable View
Repeated Measures...
tatistics Processor is ready


Repeated Measures Define Factor(s)
Within-Subject Factor Name:
factor1
Number of Levels:
Add
Change
Remove
Measure Name:
Add
Change
Remove
Define
Reset
Cancel
Help

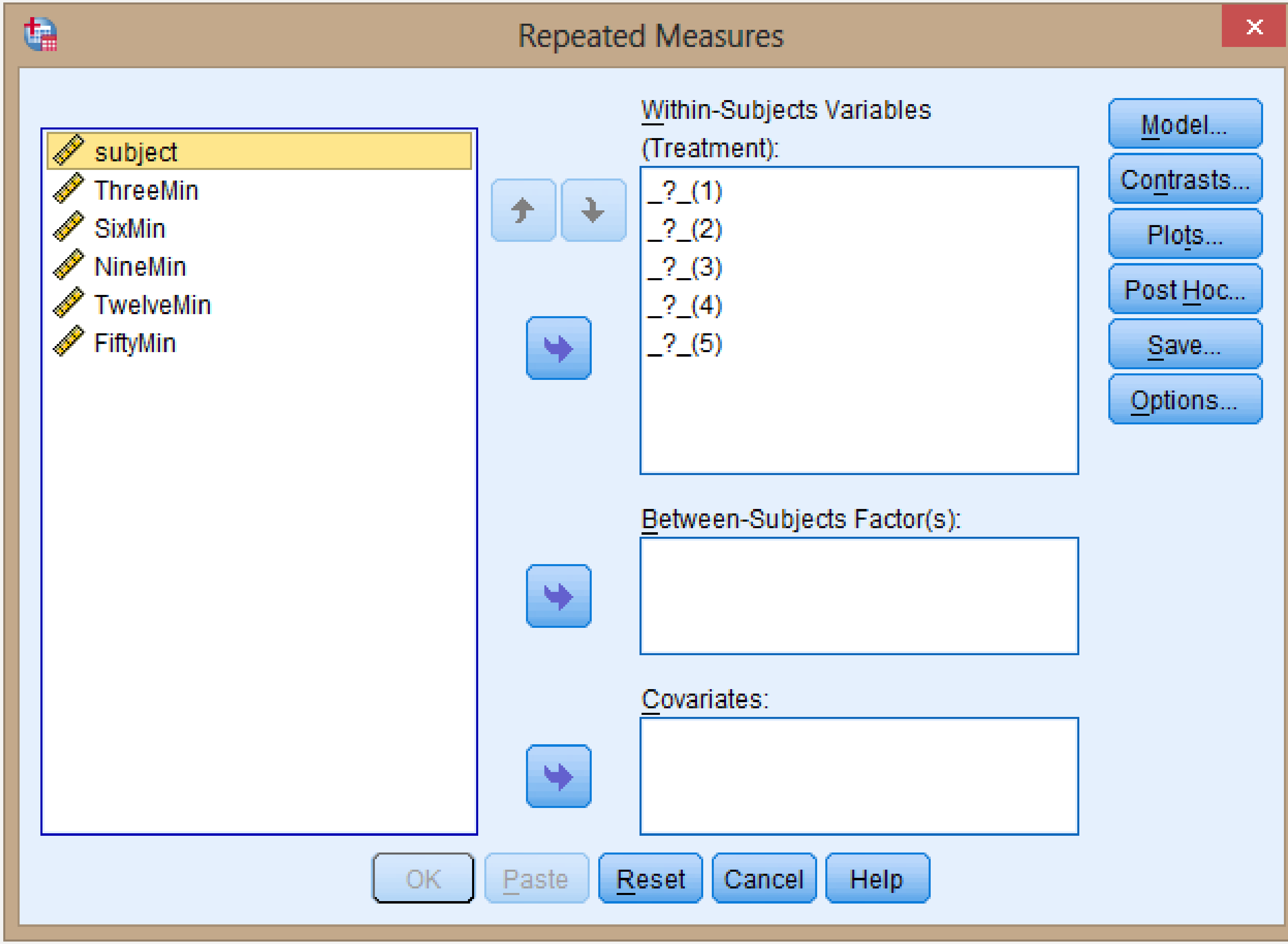
Name and Define the Within Subjects Factors

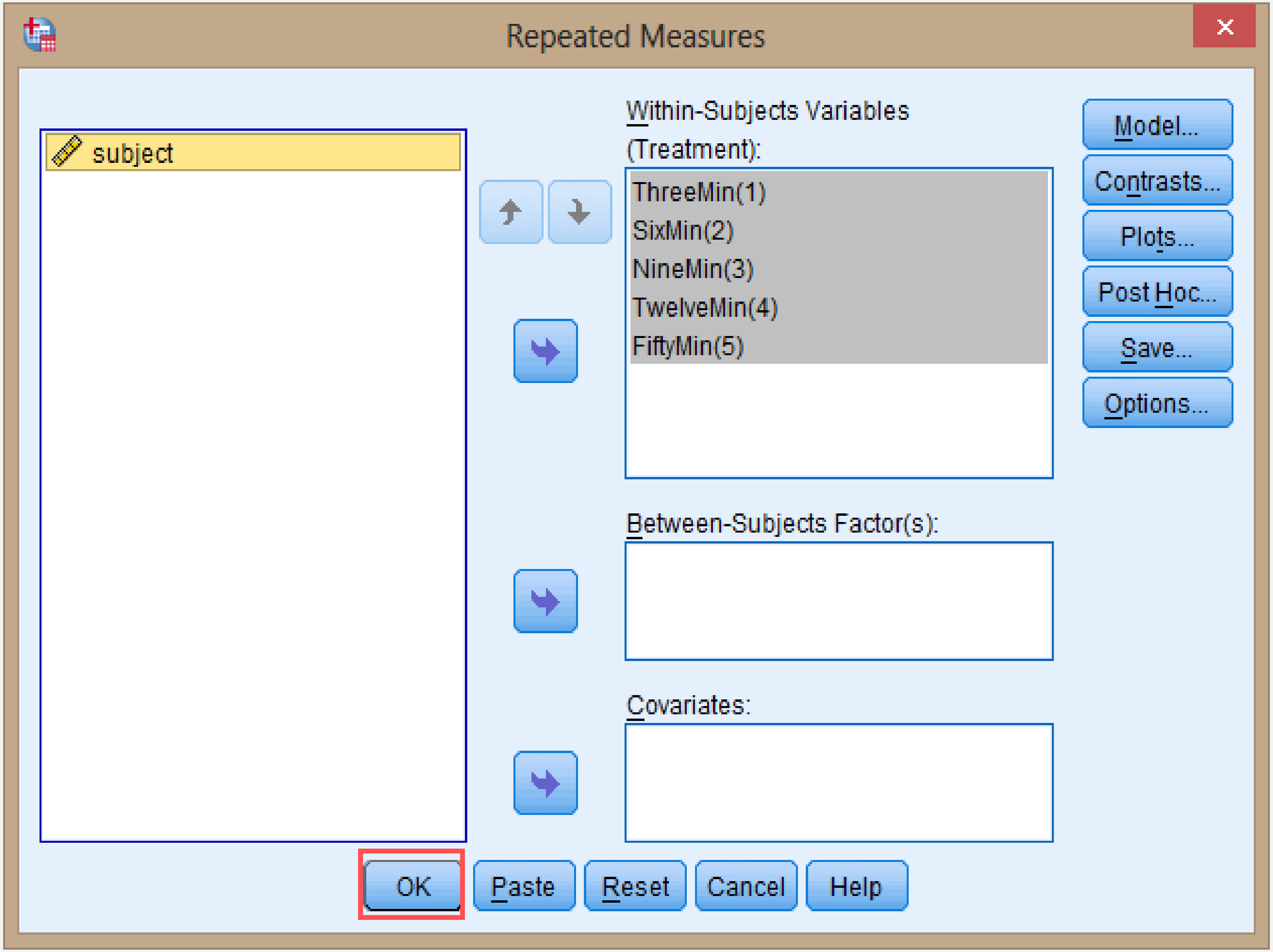


Click Add to enter each within subjects factor.



Click Define to define both Within and Between Subjects Factors.


Repeated Measures
subject
ThreeMin
SixMin
NineMin
TwelveMin
FiftyMin
Within-Subjects Variables
(Treatment):
_?_(1)
_?_(2)
_?_(3)
_?_(4)
_?_(5)
Between-Subjects Factor(s):
Covariates:
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...
OK
Paste
Reset
Cancel
Help



To check ANOVA Assumptions: Sphericity click OK

Repeated Measure ANOVA Assumptions: Sphericity?

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Treatment	.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. Design: Intercept
 Within Subjects Design: Treatment

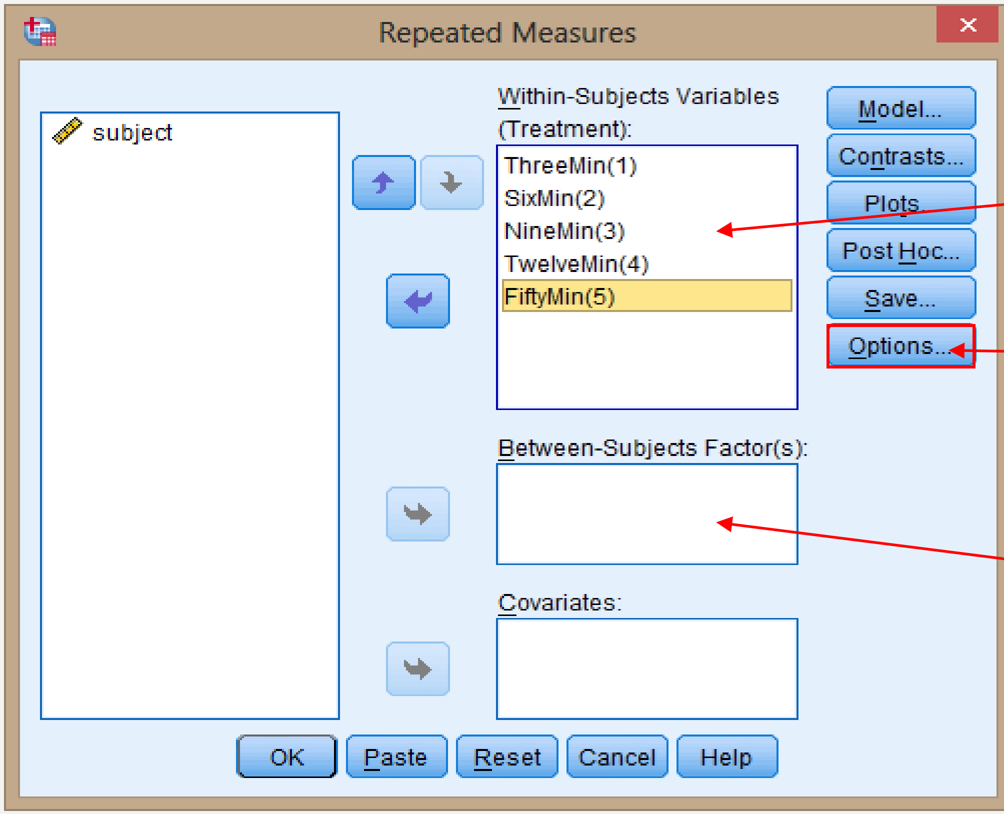
b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

You don’t want this to be significant.

Mauchly’s Test of Sphericity indicated that sphericity was violated

Since Sphericity is violated and epsilon (0.371) <0.75, we use the Greenhous-Geisser

Defining Within & Between Subjects Factors

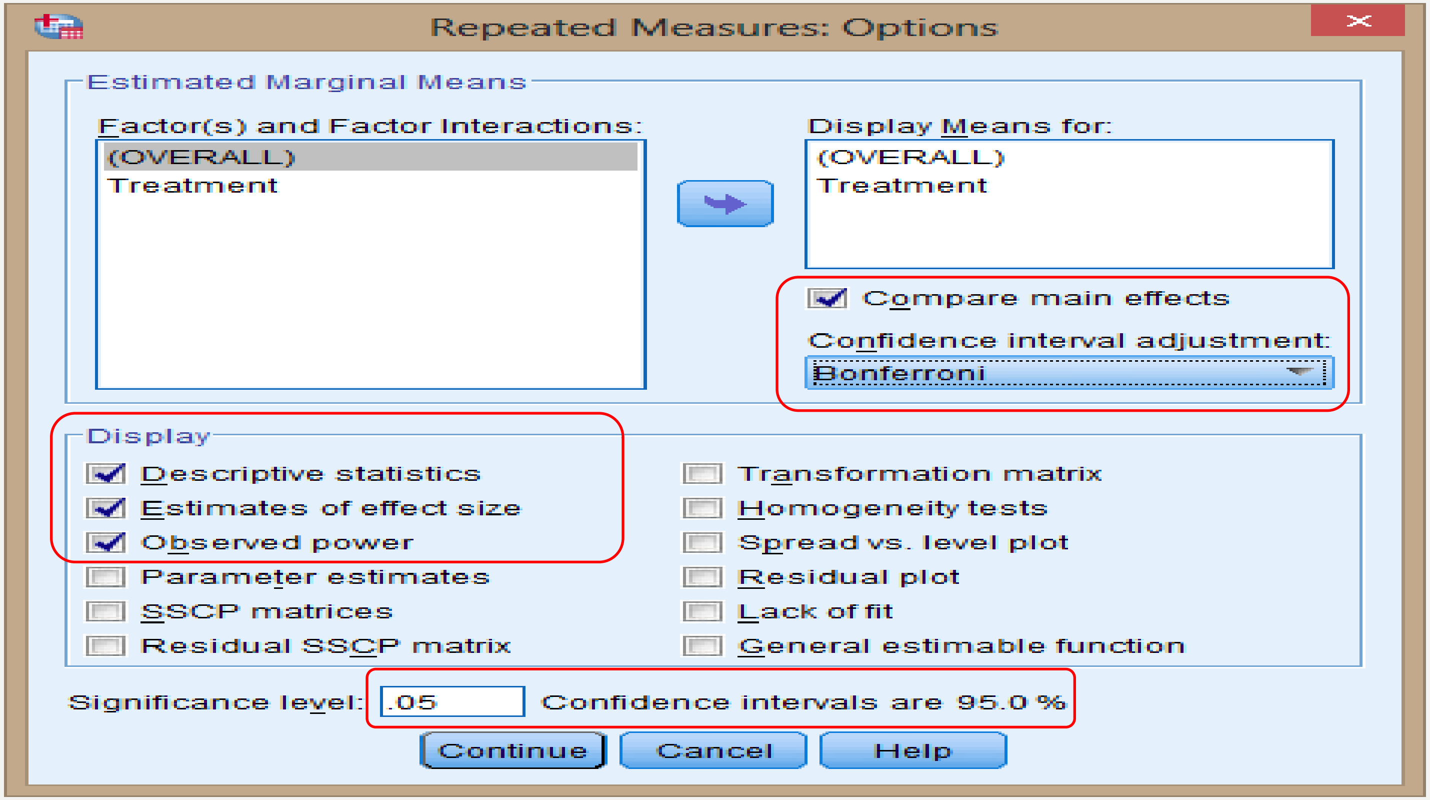


Within Subjects Factors

Click Options

Between Subjects Factors

Repeated Measures Options



SPSS Output

General Linear Model

Within-Subjects Factors

Measure: MEASURE_1

treatmnt	Dependent Variable
1	Minute_3
2	Minute_6
3	Minute_9
4	Minute_12
5	Minute_15

Descriptive Statistics

	Mean	Std. Deviation	N
Minute_3	8.5000	4.50309	10
Minute_6	11.4000	7.96102	10
Minute_9	16.4000	10.80329	10
Minute_12	31.1000	12.55610	10
Minute_15	36.5000	21.13055	10

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
treatmnt	Pillai's Trace	.866	9.694[b]	4.000	6.000	.009	.866	38.777	.934
	Wilks' Lambda	.134	9.694[b]	4.000	6.000	.009	.866	38.777	.934
	Hotelling's Trace	6.463	9.694[b]	4.000	6.000	.009	.866	38.777	.934
	Roy's Largest Root	6.463	9.694[b]	4.000	6.000	.009	.866	38.777	.934

a. Computed using alpha = .05

b. Exact statistic

c.
Design: Intercept
Within Subjects Design: treatmnt

SPSS Output: Within Subjects Factors

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Treatment	Sphericity Assumed	6115.880	4	1528.970	18.359	.000	.671	73.437	1.000
	Greenhouse-Geisser	6115.880	1.485	4117.754	18.359	.000	.671	27.268	.995
	Huynh-Feldt	6115.880	1.710	3575.916	18.359	.000	.671	31.400	.998
	Lower-bound	6115.880	1.000	6115.880	18.359	.002	.671	18.359	.967
Error(Treatment)	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					

a. Computed using alpha = .05

Since Sphericity was violated we use the adjusted values:

$F(1.485, 13.367) = 18.36$,

SPSS Output: Between Subjects Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	21590.420	1	21590.420	45.801	.000	.836	45.801	1.000
Error	4242.580	9	471.398					

a. Computed using alpha = .05

Since p value<0.05, we reject H_0 and accept H_A that at least one time period different in pain sensation.

SPSS Output: Effect Size & Confidence Intervals

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
20.780	3.070	13.834	27.726

Estimates

Measure: MEASURE_1

treatmnt	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	8.500	1.424	5.279	11.721
2	11.400	2.517	5.705	17.095
3	16.400	3.416	8.672	24.128
4	31.100	3.971	22.118	40.082
5	36.500	6.682	21.384	51.616

Post hoc Tests for Main Effects (Treatment means)

Pairwise Comparisons

Measure: MEASURE_1

(I) treatmnt	(J) treatmnt	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	-2.900	1.656	.114	-6.647	.847
	3	-7.900*	2.718	.017	-14.049	-1.751
	4	-22.600*	3.194	.000	-29.826	-15.374
	5	-28.000*	6.354	.002	-42.375	-13.625
2	1	2.900	1.656	.114	-.847	6.647
	3	-5.000	2.380	.065	-10.385	.385
	4	-19.700*	2.848	.000	-26.143	-13.257
	5	-25.100*	6.457	.004	-39.708	-10.492
3	1	7.900*	2.718	.017	1.751	14.049
	2	5.000	2.380	.065	-.385	10.385
	4	-14.700*	2.633	.000	-20.657	-8.743
	5	-20.100*	4.792	.002	-30.941	-9.259
4	1	22.600*	3.194	.000	15.374	29.826
	2	19.700*	2.848	.000	13.257	26.143
	3	14.700*	2.633	.000	8.743	20.657
	5	-5.400	4.525	.263	-15.635	4.835
5	1	28.000*	6.354	.002	13.625	42.375
	2	25.100*	6.457	.004	10.492	39.708
	3	20.100*	4.792	.002	9.259	30.941
	4	5.400	4.525	.263	-4.835	15.635

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).

For Example:

- 4 (12 min) is diff from: 1, 2, 3 (3, 6, 9 min)

============

- 5 (15 min) is diff from: 1, 2, 3 (3,6,9 min)

REPORT

1. Data.

2. Assumptions. We assume that the assumptions for the one way ANOVA repeated measures design are met.

3. Aim : to determine if the pain will progress during treatment or would it stop at any point.

4. Hypotheses.

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

H_A: not all μ's are equal

5. Using one-way ANOVA repeated measures, from Manchly's test of sphericity the Sig less than 0.05 sphericity violated and by Greenhouse the sphericity assumed.

6. Using F test and P-value, we reject H_0 and accept H_A that at least one time period different in pain sensation. So we must do paired test for every pairwise and the results reported as follows from pairwise comparisons table (you may do a similar report for μ_2 with the other means and μ_3 , μ_4 , μ_5) : For example:

- H_0: $\mu_1 = \mu_2$ Versus H_1: $\mu_1 \neq \mu_2$ (since $P=0.114>0.05$, we accept H_0)
- H_0: $\mu_1 = \mu_3$ Versus H_1: $\mu_1 \neq \mu_3$ (since $P=0.017<0.05$, we reject H_0)
- H_0: $\mu_1 = \mu_4$ Versus H_1: $\mu_1 \neq \mu_4$ (since $P=0.000<0.05$, we reject H_0)
- H_0: $\mu_1 = \mu_5$ Versus H_1: $\mu_1 \neq \mu_5$ (since $P=0.002<0.05$, we reject H_0)

….. etc.

SUMMARY

In this lecture, we review some Parametric Testes such as T-test (1-sample, 2-samples and paired samples), one and two way ANOVA and repeated measures.

In the next lectures, we will focus on some Non- Parametric Tests which are the content of STAT 333 course.

End of the Lecture