

Post hoc tests

Στην ανάλυση διασποράς, η απόρριψη της μηδενικής υπόθεσης $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ οδηγεί στην αποδοχή της εναλλακτικής

Για να εντοπιστούν οι μείζονες τιμές που διαφέρουν \rightarrow multiple comparison test

\rightarrow L.S.D (Least Significant Difference)
(μόνο μετά από απόρριψη της μηδενικής $H_0: \mu_0 = \mu_1 = \dots = \mu_k$)

\rightarrow Duncan

\rightarrow Tukey (ίδιο μέγεθος δειγμάτων)

\rightarrow Scheffe

\rightarrow Bonferroni

Το SPSS διαδίδει 18 Post-hoc tests

Εάν έχουμε δείγματα ανισού μεγέθους

\rightarrow Gabriel's (μικρό δείγμα)

\rightarrow Hochberg's (μεγάλο δείγμα)

Εάν έχουμε ανισες διακυβάνσεις

\rightarrow Games-Howell

Αόκμοη LSD

Δίνεται ο παρακάτω πίνακας ανάλυσης διακύμανσης, για δείγματα μεγέθους 8 από 4 θεραπείες - παράγοντες

Μηχι μεταβλ	DL	SS	MS	F
Παράγοντες	3	6,716	2,239	
Σφάλμα	28	3,362	120	
Σύνολο	31	10,078		

- α) Να υπολογιστεί το F στατιστικό και να εκφραστεί το συμπέρασμα
- β) Δίνονται η μέση τιμή για τις 4 θεραπείες, $\bar{y}_1 = 85$, $\bar{y}_2 = 66,5$, $\bar{y}_3 = 59$, $\bar{y}_4 = 95,5$. Να υπολογιστεί το LSD ανάμεσα στα ζεύγη σε επίπεδο σημαντικότητας 5%
- γ) Ποια είναι τα μη-στατιστικά σημαντικά ζεύγη;

$$\alpha) F_{\text{statistic}} = \frac{2239}{120} = 18,66$$

Οι βαθμοί ελευθερίας είναι 3 και 28

$$\alpha_{p\alpha} F_{3,28, \alpha=0,01} = 4,568$$

$\alpha_{\text{fact}} F_{\text{st.}} > F_{3,28}$, απορρίπτουμε τη μηδενική υπόθεση ότι οι μέσες τιμές όλων των πληθυσμών είναι ίσες.

β) Ο τύπος για τον υπολογισμό της

$$LSD = t_{\alpha/2} \sqrt{MSE \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= t_{28, 0,0025} \cdot \sqrt{MSW \left(\frac{1}{8} + \frac{1}{8}\right)}$$

$$= 2,048 \sqrt{120 \cdot \frac{1}{8} + \frac{1}{8}}$$

$$= 11,2$$

γ) Τοποθετούμε σε αύφουσα σειρά τις μέσες τιμές

$$\bar{y}_3 < \bar{y}_2 < \bar{y}_1 < \bar{y}_4$$

$$\begin{array}{cccc} 59 & 66,5 & 85 & 95,5 \\ \underbrace{\quad\quad} & \underbrace{\quad\quad} & \underbrace{\quad\quad} & \\ 7,5 & 18,5 & 10,5 & \end{array}$$

Όταν $|\bar{y}_i - \bar{y}_j| < LSD$ σημαίνει ότι δεν εντονίζεται στατιστικά σημαντική διαφορά

Μη στατιστικά σημαντικά είναι τα
 ζεύγη 2-3 και 1-4.

Άσκηση - Tukey

Δίνεται ο παρακάτω πίνακας άκρων

Μηγή βιολογία	df	SS	MS	F	p-value
Παραγοντός	2	137,20	68,60	4,314	0,0387
Σφάλμα	12	190,00	15,90		
Σύνολο	14	328,00			

Οι παράγοντες είναι 3 με μέσες τιμές $\bar{x}_A = 81,2$, $\bar{x}_B = 88,2$, $\bar{x}_C = 82,6$ και 5 παρατηρήσεις ανά παράγοντα.

Να γίνουν οι συγκρίσεις ανά ζεύγη με τη μέθοδο Tukey σε επίπεδο σημαντικότητας 5%

Το κριτήριο Tukey ορίζεται ως εξής:

$$T = q_{\alpha, (k, n-k)} \cdot \sqrt{\frac{MSE}{n_i}}$$

k : αριθμός θεραπειών

n : συνολικός αριθμός δείγματος

n_i : το μέγεθος του δείγματος με τις λιγότερες παρατηρήσεις

Από τους πίνακες παίρνουμε

$$q_{0,5}(3, 12) = 3,773$$

$$\alpha \text{ ή } T = 3,773 \cdot \sqrt{\frac{15,9}{5}} = 6,73$$

Υπολογίζουμε και τις απόλυτες διαφορές των μέσων τιμών ανά ζεύγη:

$$|\bar{x}_1 - \bar{x}_2| = |81,2 - 88,2| = 7 > 6,73$$

υπάρχει στατιστικά σημαντική διαφορά

$$|\bar{x}_1 - \bar{x}_3| = |81,2 - 82,6| = 1,4 < 6,73$$

δεν υπάρχει σ.σ διαφορά

$$|\bar{x}_2 - \bar{x}_3| = |88,2 - 82,6| = 5,6 < 6,73$$

δεν υπάρχει σ.σ διαφορά

6.

Μοντέλο Τυχαίων επιδράσεων με
έναν παράγοντα

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

μ : συνολική μέση τιμή

τ_i : είναι ανεξάρτητη τυχαία μεταβλητή
που κατανέμεται κανονικά με μέσο
0 & άγνωστη διακύμανση σ_a^2

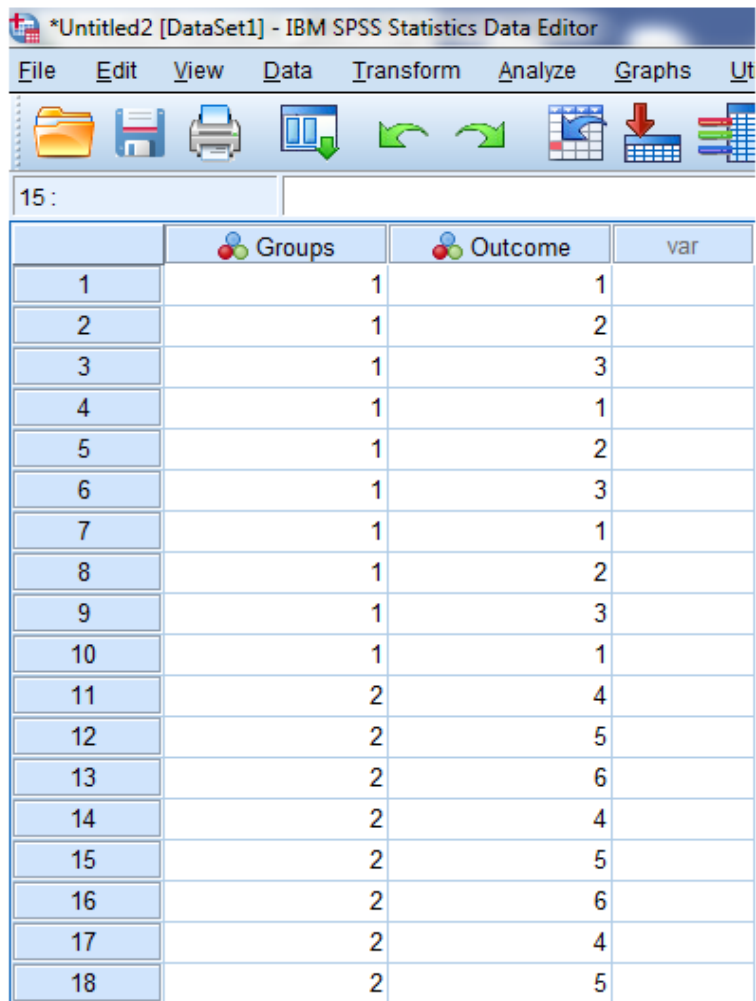
$$\epsilon_{ij} \sim N(0, \sigma^2)$$

Υποθέσεις: $H_0: \sigma_a^2 = 0$ vs $H_1: \sigma_a^2 > 0$

Οι υποθέσεις που εξετάζονται στην προκει-
μένη περίπτωση, όταν δηλαδή έχουμε
παράγοντα τυχαίων επιδράσεων, δεν
αφορούν τις μέσες τιμές ανά επίπεδο
αλλά την διακύμανση του παράγοντα.

Παράδειγμα ANOVA με Post Hoc ελέγχους στο SPSS

Τα δεδομένα:



The screenshot shows the IBM SPSS Statistics Data Editor interface. The title bar reads '*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, and Ut. The toolbar contains icons for file operations (Folder, Save, Print, Copy, Paste) and data manipulation (Undo, Redo, Sort, Filter, Download, Refresh). The data grid shows 18 rows and 4 columns. The first column is labeled '15 :'. The second column is 'Groups', the third is 'Outcome', and the fourth is 'var'.

	Groups	Outcome	var
1	1	1	
2	1	2	
3	1	3	
4	1	1	
5	1	2	
6	1	3	
7	1	1	
8	1	2	
9	1	3	
10	1	1	
11	2	4	
12	2	5	
13	2	6	
14	2	4	
15	2	5	
16	2	6	
17	2	4	
18	2	5	

Έλεγχος κανονικότητας:

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

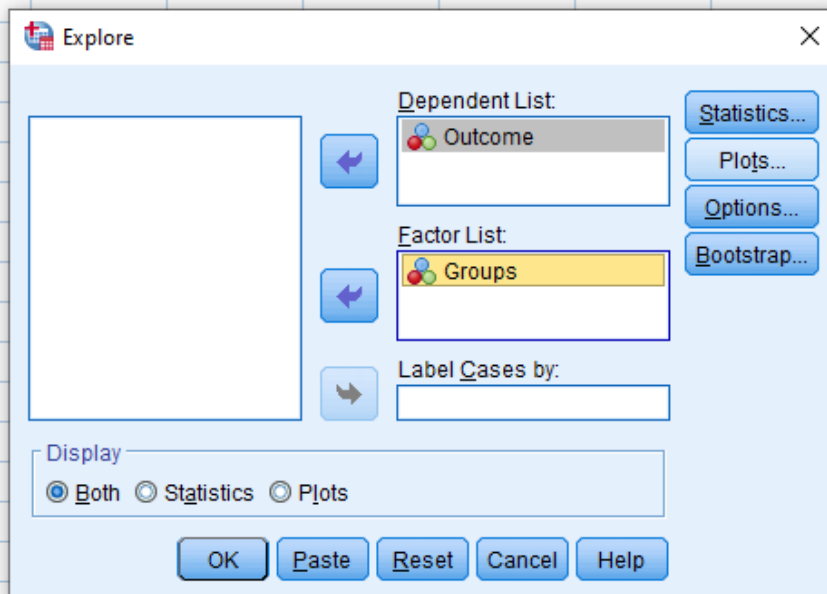
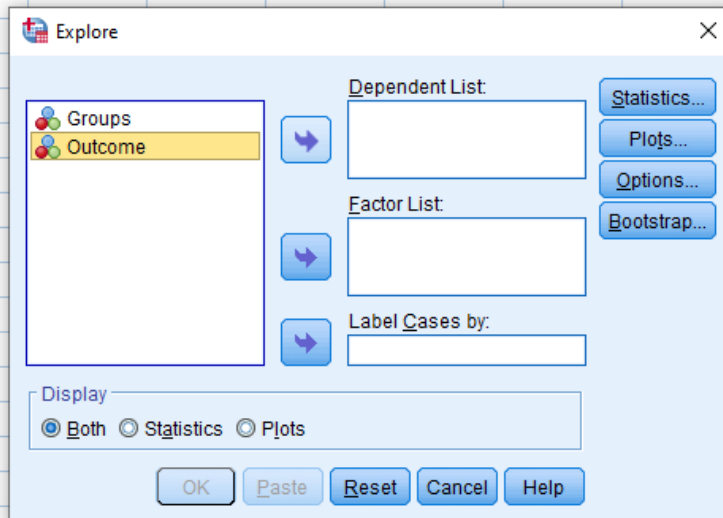
File Edit View Data Transform **Analyze** Graphs Utilities Extensions Window Help

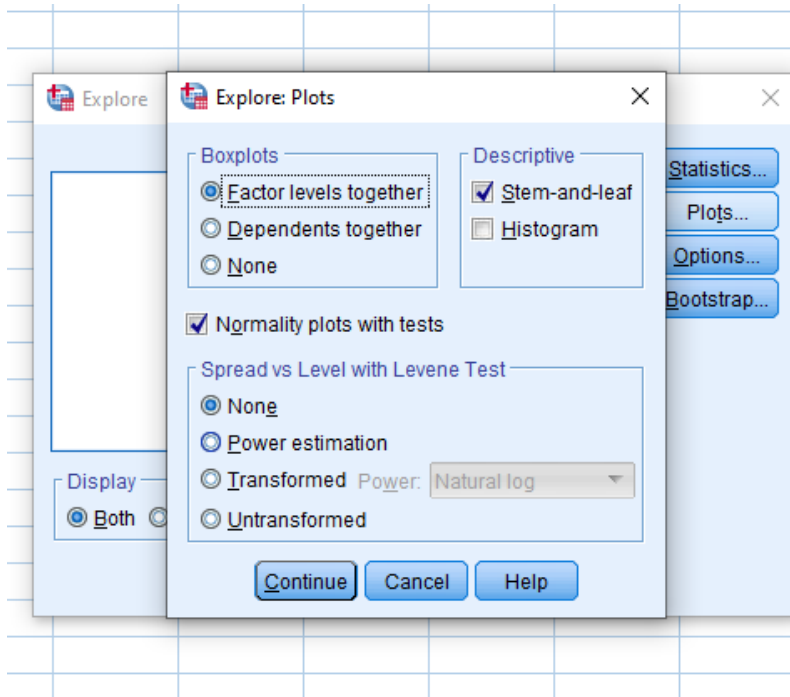
Reports
Descriptive Statistics
 Bayesian Statistics
 Tables
 Compare Means
 General Linear Model
 Generalized Linear Models
 Mixed Models
 Correlate
 Regression
 Loglinear
 Neural Networks
 Classify
 Dimension Reduction
 Scale
 Nonparametric Tests
 Forecasting
 Survival
 Multiple Response

123 Frequencies...
 Descriptives...
Explore...
 Crosstabs...
 TURF Analysis
 Ratio...
 P-P Plots...
 Q-Q Plots...

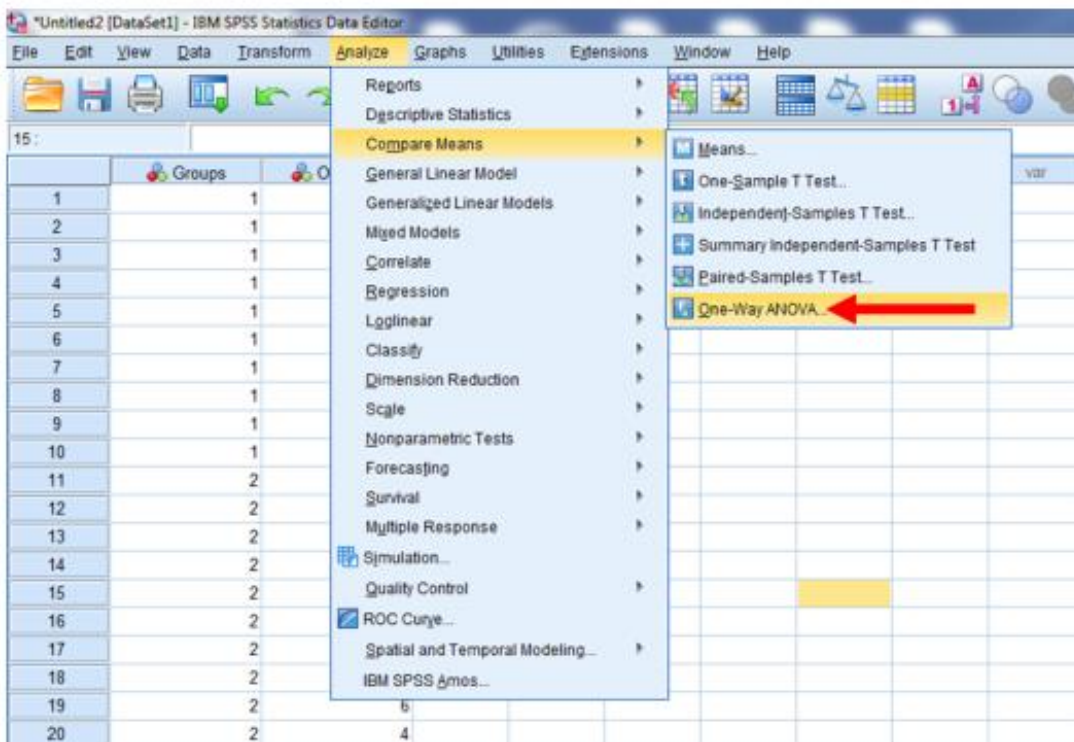
	Groups	
1	1	
2	1	
3	1	
4	1	
5	1	
6	1	
7	1	
8	1	
9	1	
10	1	
11	2	
12	2	
13	2	
14	2	
15	2	
16	2	

var





ONE WAY ANOVA



*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

15:

	Groups	Outcome	var	var	var	var	var	var	var
1	1	1							
2	1	2							
3	1	3							
4	1	1							
5	1	2							
6	1	3							
7	1	1							
8	1	2							
9	1	3							
10	1	1							
11	2	4							
12	2	5							
13	2	6							
14	2	4							
15	2	5							

One-Way ANOVA

Dependent List:

Groups Outcome

Factor:

OK Paste Reset Cancel Help

Contrasts Post Hoc Options...

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

15:

	Groups	Outcome	var	var	var	var	var	var	var
1	1	1							
2	1	2							
3	1	3							
4	1	1							
5	1	2							
6	1	3							
7	1	1							
8	1	2							
9	1	3							
10	1	1							
11	2	4							
12	2	5							
13	2	6							
14	2	4							

One-Way ANOVA

Dependent List:

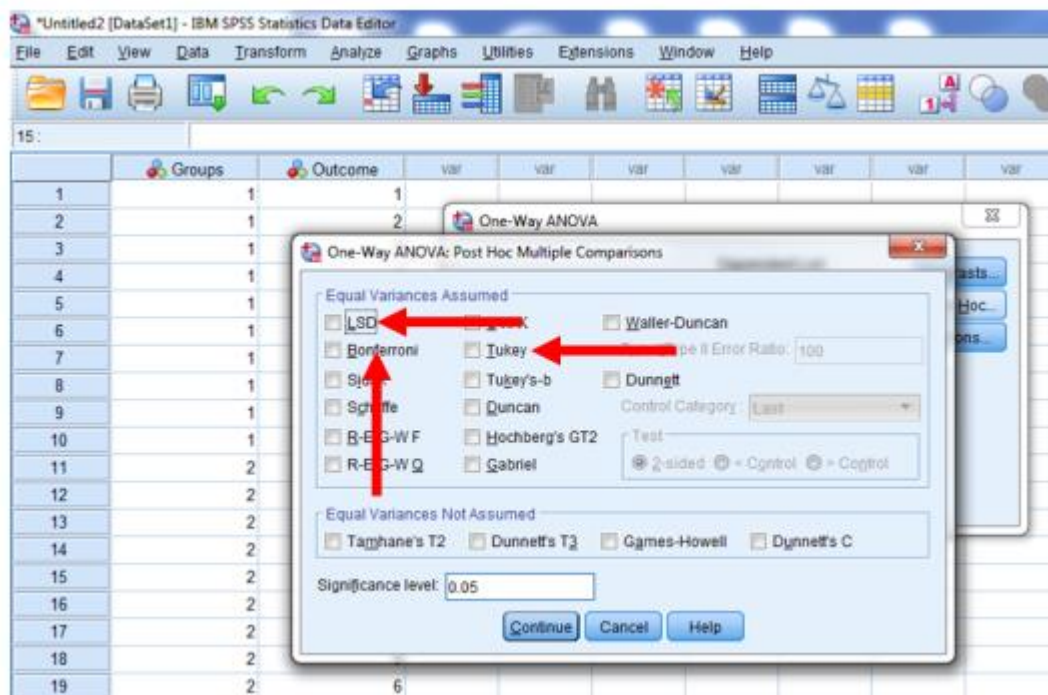
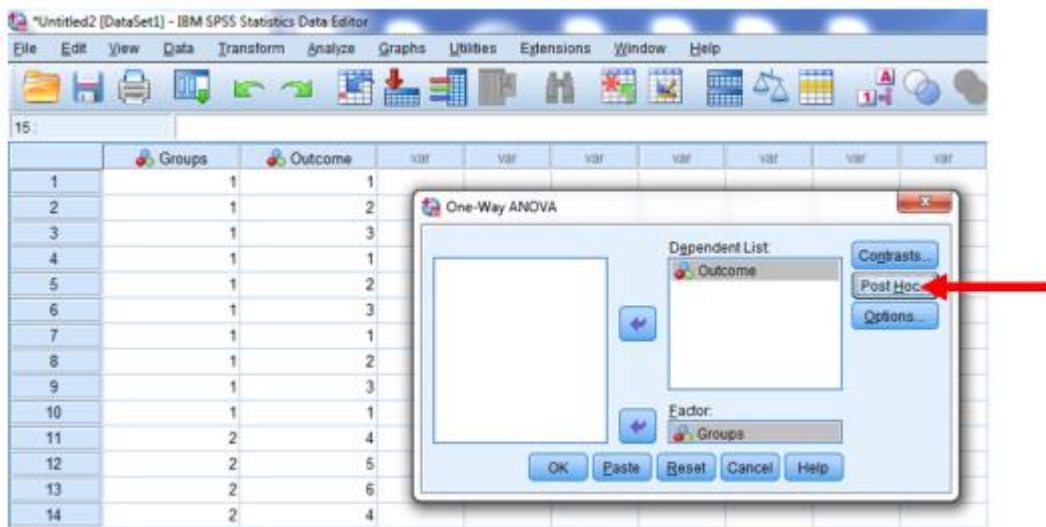
Groups Outcome

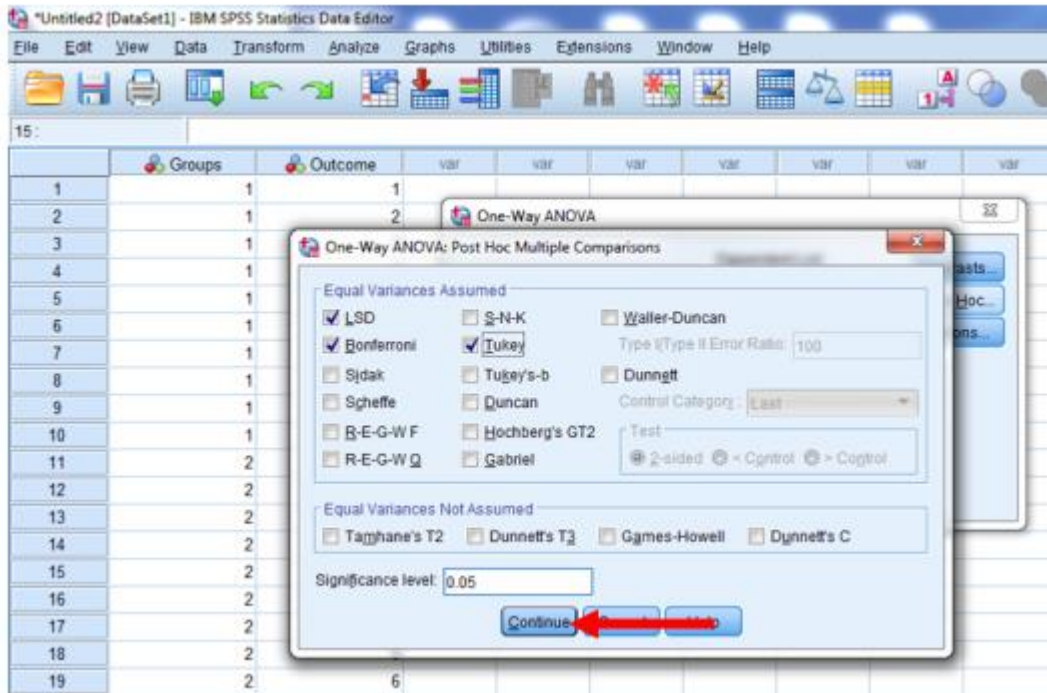
Factor:

OK Paste Reset Cancel Help

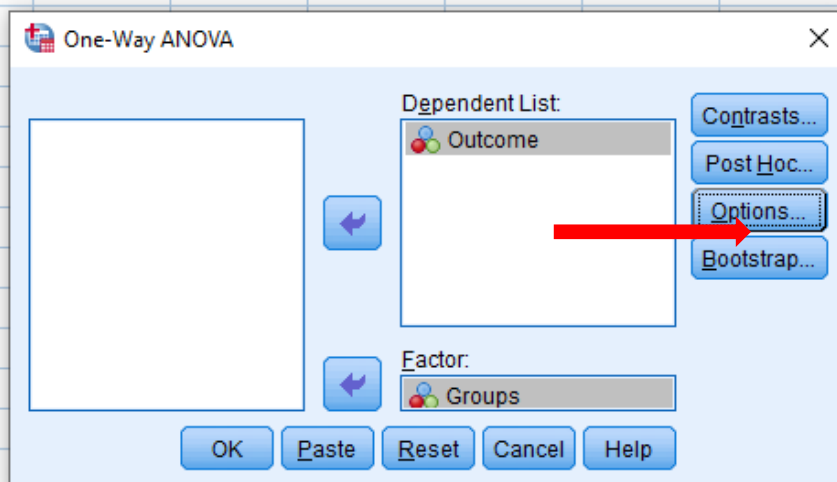
Contrasts Post Hoc Options...

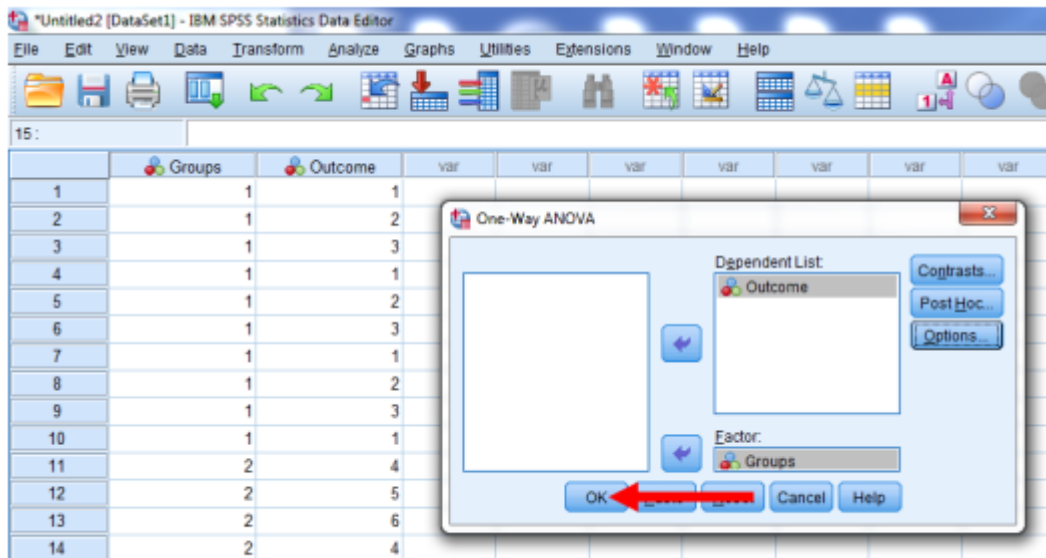
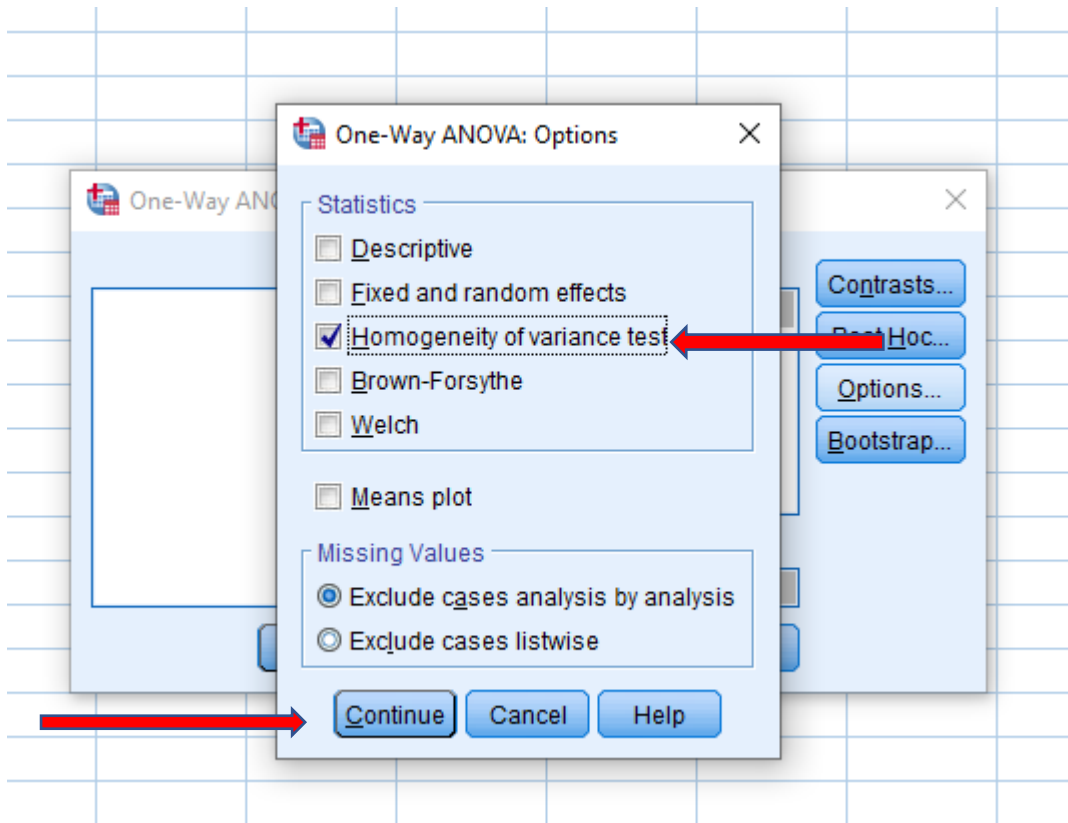
Post Hoc έλεγχου:





Έλεγχος διασπορών:





Τα output από το SPSS:

Tests of Normality

Groups	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Outcome 1	,248	10	,082	,805	10	,017
Outcome 2	,248	10	,082	,805	10	,017
Outcome 3	,248	10	,082	,805	10	,017

a. Lilliefors Significance Correction

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Outcome	Based on Mean	,000	2	27	1,000
	Based on Median	,000	2	27	1,000
	Based on Median and with adjusted df	,000	2	27,000	1,000
	Based on trimmed mean	,000	2	27	1,000

ANOVA					
Outcome					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	180.000	2	90.000	117.391	,000
Within Groups	20.700	27	,767		
Total	200.700	29			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Outcome

	(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	-3.000*	.392	.000	-3.97	-2.03
		3	-6.000*	.392	.000	-6.97	-5.03
	2	1	3.000*	.392	.000	2.03	3.97
		3	-3.000*	.392	.000	-3.97	-2.03
	3	1	6.000*	.392	.000	5.03	6.97
		2	3.000*	.392	.000	2.03	3.97
LSD	1	2	-3.000*	.392	.000	-3.80	-2.20
		3	-6.000*	.392	.000	-6.80	-5.20
	2	1	3.000*	.392	.000	2.20	3.80
		3	-3.000*	.392	.000	-3.80	-2.20
	3	1	6.000*	.392	.000	5.20	6.80
		2	3.000*	.392	.000	2.20	3.80
Bonferroni	1	2	-3.000*	.392	.000	-4.00	-2.00
		3	-6.000*	.392	.000	-7.00	-5.00
	2	1	3.000*	.392	.000	2.00	4.00
		3	-3.000*	.392	.000	-4.00	-2.00
	3	1	6.000*	.392	.000	5.00	7.00
		2	3.000*	.392	.000	2.00	4.00

*. The mean difference is significant at the 0.05 level.