ΚΕΦΑΛΑΙΟ ΕΒΔΟΜΟ

Έλεγχος για τις παραμέτρους θέσης περισσοτέρων των δύο πληθυσμών με ανεξάρτητα δείγματα

Έστω $Y_{j1},\dots,Y_{jn_j}$, j το πλήθος $j=1,...,k$, $k\geq 2$ τυχαία ανεξάρτητα δείγματα μεγέθους n_j από έναν πληθυσμό με μέση τιμή μ_j και διακύμανση σ_j^2, άγνωστη. Ενδιαφερόμαστε για τον έλεγχο, σε επίπεδο σημαντικότητας α, της μηδενικής υπόθεσης

$$H_0 : \mu_1 = \mu_2 = ... = \mu_k ,$$

ως προς την εναλλακτική $H_a : \mu_i \neq \mu_j$, για τουλάχιστον ένα ζεύγος με $i \neq j, i, j = 1,...,k$, $k > 2$.

Επομένως, το ενδιαφέρον τώρα επικεντρώνεται στον έλεγχο ότι οι μέσες τιμές της εξαρτημένης μεταβλητής δύο ή περισσότερων ομάδων (επιπέδων του παράγοντα) δε διαφέρουν στατιστικά σημαντικά. Επομένως, γίνεται αντιληπτό ότι αποτελεί γενίκευση του προβλήματος της σύγκρισης των μέσων τιμών δύο πληθυσμών σε περισσότερους από δύο ανεξάρτητους πληθυσμούς (με ανεξάρτητα δείγματα).

Το παραπάνω πρόβλημα ελέγχεται υπό κάποιες υποθέσεις με τον παραμετρικό έλεγχο του t-test. Όταν κάποιες από τις υποθέσεις αυτές δεν ικανοποιείται και δεν υπάρχει τρόπος διόρθωσης του προβλήματος ο έλεγχος ανάγεται σε αυτόν ότι οι πληθυσμιακές διάμεσοι είναι ίσες μεταξύ τους. Τα αποτελέσματα του τελευταίου ελέγχου γενικεύονται για τον δοθέν έλεγχο όταν τα δεδομένα είναι συμμετρικά.

7.1 Μεθοδολογία-Υλοποίηση στο S.P.S.S.

Η μεθοδολογία που θα χρησιμοποιηθεί για τη στατιστική ανάλυση ενός τέτοιου προβλήματος εξαρτάται από το αν πληρούνται ή όχι κάποιες προϋποθέσεις, τις οποίες και πρέπει αρχικά να ελέγξει ο ερευνητής. Πιο συγκεκριμένα, ελέγχουμε
α) αν το ποσοστό των ακραίων τιμών στις διαθέσιμες δειγματικές παρατηρήσεις από καθένα από τους k το πλήθος πληθυσμούς ξεπερνά το 10% αυτών, και

β) αν οι πληθυσμοί από τους οποίους λαμβάνονται τα τυχαία δείγματα μπορούμε να ισχυριστούμε ότι περιγράφονται ικανοποιητικά από την κανονική κατανομή.

Ανάλογα με τα αποτελέσματα των παραπάνω ελέγχων προβαίνουμε σε παραμετρικό έλεγχο ή σε μη παραμετρικό έλεγχο. Στη συνέχεια παρουσιάζονται όλα τα πιθανά αποτελέσματα των α) και β), τα διάφορα βήματα της ανάλυσης και οι αποφάσεις στις οποίες οδηγούμαστε.

1. Αρχικά ελέγχουμε αν υπάρχουν ακραίες τιμές στις διαθέσιμες δειγματικές τιμές σε καθένα από τους k σε πλήθος πληθυσμούς. Αν το ποσοστό των ακραίων τιμών σε καθένα από αυτά τα δείγματα δε ξεπερνά το 10%, τότε προχωρούμε στο βήμα 2. Αν το ποσοστό των ακραίων τιμών σε κάποιο από αυτά τα δείγματα ξεπερνά το 10%, τότε δοκιμάζουμε μήπως ο μετασχηματισμός του λογαρίθμου διορθώνει το πρόβλημα. Αν το πρόβλημα αυτό διορθώνεται, τότε μεταβαίνουμε στο βήμα 2, σε διαφορετική περίπτωση συμπεραίνουμε ότι θα χρησιμοποιηθεί ο μη παραμετρικός έλεγχος (βλέπε βήμα 4).

2. Στο βήμα 2, χρησιμοποιώντας το τεστ των Shapiro-Wilk καθώς και γραφικούς τρόπους, ελέγχουμε αν οι διαθέσιμες δειγματικές παρατηρήσεις από καθένα από τους k πληθυσμούς (είτε οι αρχικές είτε οι μετασχηματισμένες που έχουν προκύψει από το βήμα 1) προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή. Αν ο έλεγχος της κανονικότητας μας υποδεικνύει ότι η υπόθεσης της κανονικότητας δεν απορρίπτεται (p-τιμή>α), τότε η ανάλυση θα συνεχιστεί με τον παραμετρικό έλεγχο (βλέπε βήμα 3). Αν η υπόθεση της κανονικότητας απορρίπτεται για έναν από αυτούς ή και για τους k υπό εξέταση πληθυσμούς (τεστ Shapiro-Wilk, p-τιμή<α), τότε ελέγχουμε αν το πρόβλημα της μη κανονικότητας διορθώνεται μετασχηματίζοντας κατάλληλα τα δεδομένα (Box-Cox μετασχηματισμός) και επανελέγχοντας την ύπαρξη ακραίων τιμών, δηλαδή ξεκινώντας την ανάλυση από το βήμα 1. Αν με κάποιο μετασχηματισμό των δεδομένων επιτυγχάνεται η κανονικότητα όλων των πληθυσμών, συνεχίζουμε την ανάλυση παραμετρικά (βήμα 3). Σε αντίθετη περίπτωση, αν το πλήθος των δειγματικών παρατηρήσεων (μη λαμβάνοντας υπόψη αυτές που έχουν αφαιρεθεί στο βήμα 1) εκείνου ή εκείνων των πληθυσμών που δεν περιγράφονται από την κανονική κατανομή είναι μεγάλο (συνήθως μεγαλύτερο του 30), κάνοντας χρήση του Κεντρικού Οριακού Θεωρήματος, προβαίνουμε στον παραμετρικό έλεγχο της υπό έλεγχο υπόθεσης (βλέπε βήμα 3), όπου τα αποτελέσματα θα είναι προσεγγιστικά. Στην περίπτωση τώρα που το πρόβλημα της μη κανονικότητας κάποιου ή και των k πληθυσμών δε διορθώνεται (τεστ Shapiro-Wilk, p-τιμή <α), και ταυτόχρονα το πλήθος των δειγματικών παρατηρήσεων (μη λαμβάνοντας

υπόψη αυτές που έχουν αφαιρεθεί στο βήμα 1) από αυτόν τον πληθυσμό ή από αυτούς τους πληθυσμούς ανάλογα είναι μικρό (συνήθως μικρότερο του 30), συνεχίζεται η περαιτέρω ανάλυση μη παραμετρικά (βήμα 4).

3. Παραμετρικός έλεγχος: Σε αυτήν την περίπτωση η περαιτέρω ανάλυση επηρεάζεται από το αποτέλεσμα του έλεγχου της ισότητας των k το πλήθος πληθυσμιακών διακυμάνσεων.

ι) Ειδικότερα, αν η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων δεν απορρίπτεται (τεστ του Levene, p-τιμή>α), τότε ο έλεγχος της υπόθεσης $H_0 : \mu_1 = \mu_2 = ... = \mu_k$, γίνεται με τη βοήθεια του F στατιστικού του πίνακα ΑΝΑΔΙΑ, όπου

$$F = \frac{MS_{tr}}{MS_{res}} \overset{H_0}{\sim} F_{k-1,n-k},$$

με

$$MS_{tr} = \frac{SS_{tr}}{k-1} = \frac{\sum_{j=1}^{k} n_j \left(\overline{y}_{j.} - \overline{y}_{..}\right)^2}{k-1},$$

και

$$MS_{res} = \frac{SS_{res}}{n-k} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} n_j \left(y_{ji} - \overline{y}_{..}\right)^2 - \sum_{j=1}^{k} n_j \left(\overline{y}_{j.} - \overline{y}_{..}\right)^2}{n-k},$$

όπου $n = \sum_{j=1}^{k} n_j$, $\overline{y}_{j.} = \frac{\sum_{i=1}^{n_j} y_{ji}}{n_j}$, $\overline{y}_{..} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} y_{ij}}{n}$. Η κρίσιμη περιοχή του ελέγχου με επίπεδο σημαντικότητας a δίνεται από τη σχέση: $F \geq F_{k-1,n-k,a}$.

ii) Αν η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων απορρίπτεται (τεστ του Levene, p-τιμή<α), τότε ο έλεγχος της υπόθεσης $H_0 : \mu_1 = \mu_2 = ... = \mu_k$, γίνεται με το στατιστικό τεστ των Brown-Forsythe ή του Welch.

Αν χρησιμοποιώντας είτε το F στατιστικό του πίνακα ΑΝΑΔΙΑ είτε το στατιστικό τεστ των Brown-Forsythe ή του Welch, συμπεράνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές της μέσης τιμής της ποσοτικής μεταβλητής ως προς τις διάφορες κατηγορίες της ποιοτικής μεταβλητής (p-τιμή<α), τότε προκύπτει εύλογα το ερώτημα ποιο μ_i διαφέρει στατιστικά σημαντικά από τα υπόλοιπα. Για το σκοπό αυτό θα πρέπει να γίνουν οι

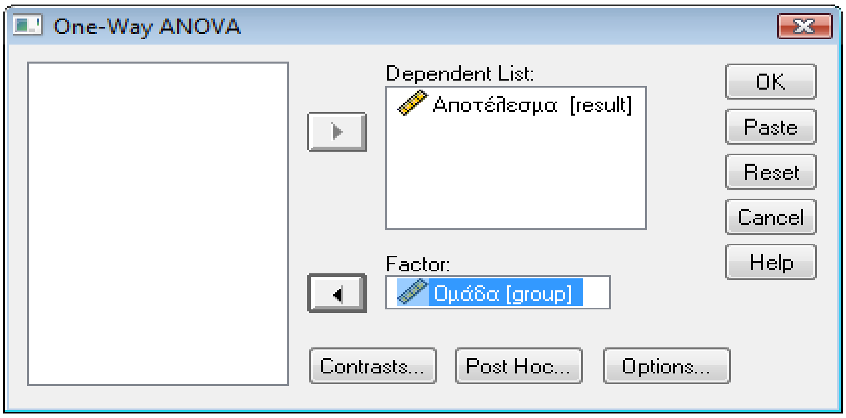
$\binom{k}{2} = \frac{k(k-1)}{2}$ το πλήθος έλεγχοι της μορφής: $H_0 : \mu_i = \mu_j$, $i, j = 1,...,k$, $i \neq j$. Το συνηθέστερο λάθος που γίνεται είναι να πραγματοποιηθούν οι παραπάνω $\binom{k}{2}$ έλεγχοι χρησιμοποιώντας το t τεστ. Η χρήση του t τεστ έχει ως συνέπεια την αύξηση της πιθανότητας εσφαλμένης απόρριψης κάποιας από τις $\binom{k}{2}$ υποθέσεις. Η πιθανότητα εσφαλμένης απόρριψης δεν είναι πλέον ίση με το επίπεδο σημαντικότητας a κάθε ελέγχου, αλλά περίπου ίση με $1-(1-a)^{\binom{k}{2}}$ (βλέπε Καρακώστας (2002)). Για την αντιμετώπιση αυτού του προβλήματος έχουν προταθεί διάφορες μεθοδολογίες που είναι γνωστές ως Πολλαπλές Συγκρίσεις (Multiple Comparisons ή Post hoc tests). Άλλες από αυτές χρησιμοποιούνται όταν δεν έχει απορριφθεί η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων και άλλες στην περίπτωση που έχει απορριφθεί η υπόθεσης της ομοσκεδαστικότητας των πληθυσμών. Στην πρώτη περίπτωση ανήκουν μεταξύ άλλων, η ελάχιστη σημαντική διαφορά (έχουν ασκηθεί κριτικές ότι το επίπεδο σημαντικότητας αυξάνει όσο αυξάνει ο αριθμός των ομάδων k), η μέθοδος του Bonferroni (προτιμάται όταν το k είναι μικρό, επιτυγχάνει να διατηρεί το επίπεδο σημαντικότητας μικρότερο του a), του Sidak, του Scheffe (κατάλληλη μέθοδος όχι μόνο για σύγκριση ζευγών αλλά και περισσότερων). Στη δεύτερη περίπτωση, χρησιμοποιούνται οι μεθοδολογίες των Tamhane's T2 (συντηρητικό τεστ), Dunnett's T3, Games-Howell (τείνει να δίνει στατιστικά σημαντικές διαφορές) και Dunnett's C.

<u>Υλοποίηση στο S.P.S.S.</u>

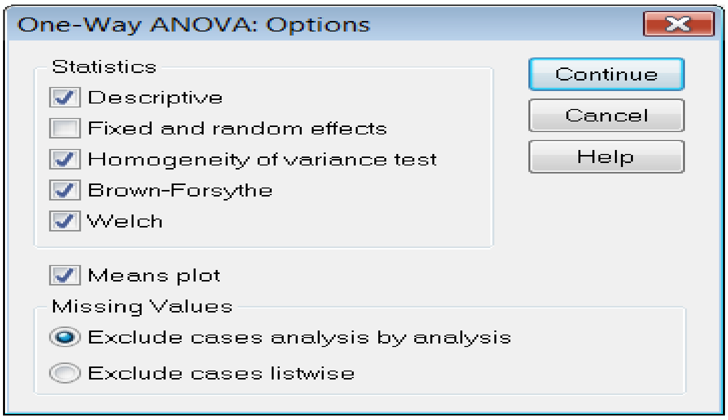
Επιλέγουμε από το αρχικό παράθυρο του S.P.S.S.

i. Analyze $\rightarrow$ Compare Means$\rightarrow$ One-Way Anova.

ii. Στο παράθυρο διαλόγου που προκύπτει στο πλαίσιο Dependent List τοποθετούμε την εξαρτημένη ποσοτική μεταβλητή (έστω Αποτέλεσμα), ενώ στο πλαίσιο Factor τον παράγοντα που πιθανώς επηρεάζει την εξαρτημένη μεταβλητή (έστω Ομάδα).



Πατώντας το πλαίσιο Options από το πλαίσιο Statistics επιλέγουμε τα ακόλουθα:



- Descriptive. Δίνονται στο Output πληροφορίες για κάθε εξαρτημένη μεταβλητή, η οποία έχει δηλωθεί στο πλαίσιο Dependent List, ως προς τα διάφορα επίπεδα του παράγοντα που έχει δηλωθεί στο πλαίσιο Factor. Οι πληροφορίες αυτές αφορούν περιγραφικά μέτρα, όπως το πλήθος των πειραματικών μονάδων, τη μέση τιμή, την τυπική απόκλιση, το τυπικό σφάλμα για τη μέση τιμή, τη μέγιστη και την ελάχιστη τιμή. Επιπλέον υπολογίζεται το 95% διάστημα εμπιστοσύνης για τη μέση τιμή κάθε εξαρτημένης μεταβλητής για κάθε επίπεδο του παράγοντα
- Homogeneity of variance test. Υπολογίζει το στατιστικό του Levene για τον έλεγχο τις ισότητας των διακυμάνσεων. Επισημαίνεται ότι ο έλεγχος αυτός δεν είναι ανεξάρτητος της υπόθεσης της κανονικότητας.
- Brown-Forsythe και Welch. Υπολογίζει το στατιστικό των Brown-Forsythe και του Welch αντίστοιχα για τον έλεγχο της ισότητας των μέσων τιμών. Τα στατιστικά αυτά είναι καταλληλότερα από το F στατιστικό του πίνακα ΑΝΑΔΙΑ, όταν η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων απορρίπτεται.

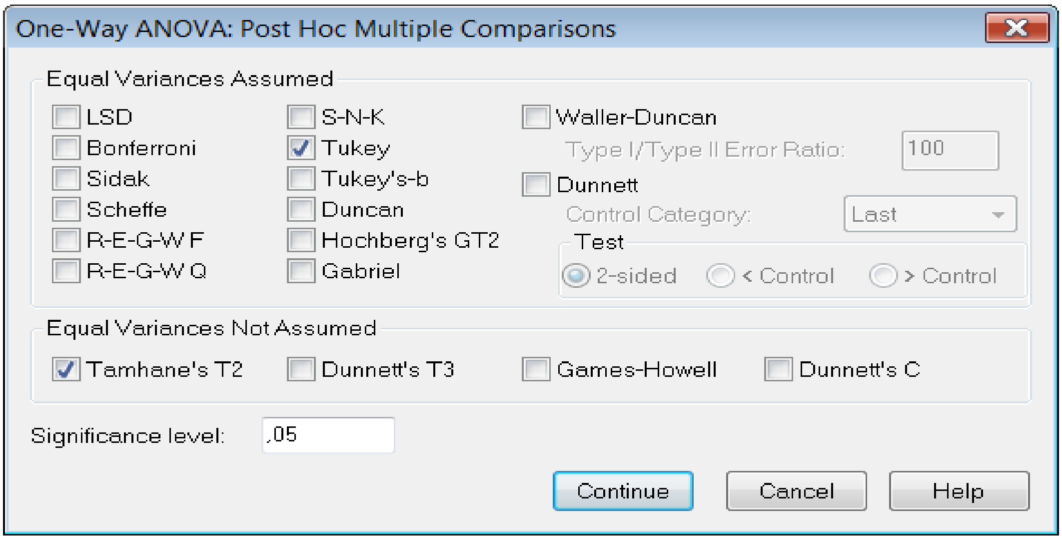
- Επιπλέον, επιλέγουμε το πλαίσιο Means plot προκειμένου να μας δώσει το λογισμικό ένα γράφημα των μέσων τιμών της εξαρτημένης μεταβλητής ως προς κάθε παράγοντα (για το σχηματισμό μίας εικόνας για τις διαφορές μεταξύ των επιπέδων).

Πληκτρολογούμε Continue και επανερχόμαστε στο αρχικό παράθυρο διαλόγου One-Way Anova. Παρατηρούμε ότι είναι διαθέσιμα και άλλα δύο πλαίσια τα Contrasts και Post Hoc.

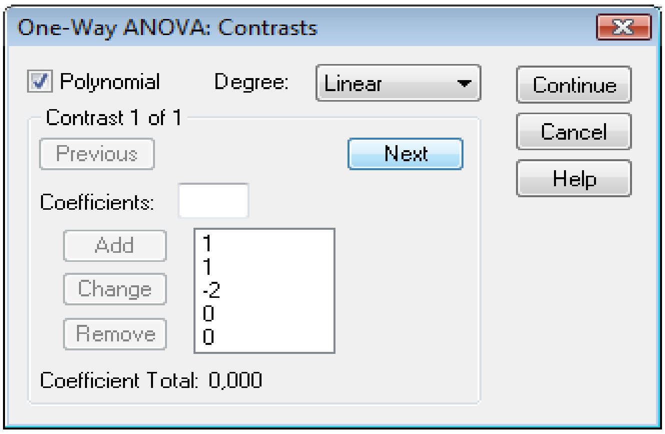
Από την επιλογή Post Hoc Multiple Comparisons έχουμε ένα πλήθος μεθόδων πολλαπλών συγκρίσεων που άλλες εφαρμόζονται στην περίπτωση που ισχύει η ισότητα των πληθυσμιακών διακυμάνσεων (Equal Variances Assumed) και άλλες όταν είναι άνισες (Not Equal Variances Assumed).

Δεν θα δώσουμε λεπτομέρειες για κάθε έλεγχο. Απλά αναφέρουμε ότι οι πιο δημοφιλείς από αυτούς τους ελέγχους είναι οι: LSD (ελάχιστη σημαντική διαφορά), Bonferroni, Tukey, Scheffe, Tamhane, Dunnett με επίπεδο σημαντικότητας που καθορίζεται στο πλαίσιο Significance level (συνήθως είναι 0.05 ή 0.01).

Ας υποθέσουμε ότι επιλέγουμε τα ακόλουθα:



Πολλές φορές ο ερευνητής δεν ενδιαφέρεται μόνο για τη σύγκριση μεταξύ δύο οποιωνδήποτε μέσων τιμών, αλλά είναι πιθανό να ενδιαφέρεται για τη σύγκριση περισσοτέρων των δύο. Στο πλαίσιο αυτό ενδιαφέρεται για τον έλεγχο κάποιας γραμμικής σχέσης της μορφής $\sum_{i=1}^{k} c_i \mu_i = 0$, όπου μ_i, $i=1,\ldots,k$, $k \geq 2$, είναι η πληθυσμιακή μέση τιμή του $i=1,\ldots,k$, $k \geq 2$, πληθυσμού και c_i κατάλληλοι συντελεστές τέτοιοι ώστε να ικανοποιείται η σχέση: $\sum_{i=1}^{k} c_i = 0$, $k \geq 2$. Τότε λέμε ότι πρόκειται για τον έλεγχο μίας γραμμικής αντίθεσης. Το λογισμικό του S.P.S.S. μας δίνει τη δυνατότητα για τη διενέργεια τέτοιων ελέγχων από το παράθυρο διαλόγου One-Way Anova και την επιλογή Contrasts με την εισαγωγή των κατάλληλων συντελεστών. Στο πλαίσιο Degree διατηρούμε την επιλογή Linear. Έπειτα από το πλαίσιο Coefficients καθορίζουμε τους συντελεστές της γραμμικής αντίθεσης. Η σειρά των συντελεστών αυτών είναι σημαντική. Τοποθετούνται κατά αύξουσα σειρά σε αντιστοιχία με τις τιμές της ποιοτικής μεταβλητής. Έστω ότι θέλουμε να ελέγξουμε για παράδειγμα τις δύο πρώτες ομάδες με την τρίτη. Γίνεται αντιληπτό ότι οι συντελεστές θα είναι 1 1 -2 0 0.



Ερμηνεία αποτελεσμάτων

Στον πίνακα Descriptives, μας δίνονται για την Επίδοση ως προς τις πέντε διαφορετικές μεθόδους διδασκαλίας, το πλήθος των πειραματικών μονάδων (N), η μέση τιμή (Mean), η τυπική απόκλιση (Std. Deviation), το τυπικό σφάλμα για τη μέση τιμή (Std. Error), η ελάχιστη (Min) και η μέγιστη τιμή (Max). Επιπλέον υπολογίζεται το 95% διάστημα εμπιστοσύνης για τη μέση επίδοση ως προς τις πέντε μεθόδους διδασκαλίας (95% Confidence Interval for Mean). Παρατηρούμε ότι η μέση επίδοση των μαθητών της τρίτης ομάδας φαίνεται να είναι μεγαλύτερη-καλύτερη, ενώ αυτή των μαθητών της πέμπτης ομάδας δείχνει να είναι η χειρότερη. Μένει να επιβεβαιωθούν, στη συνέχεια, και στατιστικά αυτές οι αρχικές παρατηρήσεις.

Descriptives

Αποτέλεσμα

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
Ομάδα Α	9	19,6667	4,21307	1,40436	16,4282	22,9051	14	24
Ομάδα Β	9	18,3333	3,57071	1,19024	15,5886	21,0780	13	23
Ομάδα Γ	9	27,4444	2,45515	,81838	25,5572	29,3316	23,	30
Ομάδα Δ	9	23,4444	3,08671	1,02890	21,0718	25,8171	19,	28
Ομάδα Ε	9	16,1111	3,62093	1,20698	13,3278	18,8944	10	21
Total	45	21,0000	5,21362	,77720	19,4337	22,5663	10	30

Από τον πίνακα Test of Homogeneity of Variances μας δίνεται η τιμή, οι βαθμοί ελευθερίας και η p-τιμή του στατιστικού τεστ του Levene για τον έλεγχο της υπόθεσης των ίσων διακυμάνσεων. Συμπεραίνουμε ότι η υπόθεση της ισότητας των διακυμάνσεων δε μπορεί να απορριφθεί (p-τιμή=0.201>0.05). Επομένως πρέπει να χρησιμοποιήσουμε το F τεστ από τον πίνακα ΑΝΑΔΙΑ (άρα είναι λανθασμένο να χρησιμοποιήσουμε τα τεστ των Brown-Forsythe και Welch) για τον έλεγχο της υπόθεσης ότι δε διαφέρουν οι μέσες τιμές της επίδοσης ως προς τις 5 διαφορετικές μεθοδολογίες διδασκαλίας.

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1,569	4	40	,201

Στον πίνακα ANOVA μας δίνεται ο πίνακας ΑΝΑΔΙΑ και όλες οι πληροφορίες που περιέχονται σε αυτόν. Από την τιμή και την p-τιμή του F στατιστικού τεστ προκύπτει ότι η υπόθεση της ισότητας των μέσων τιμών απορρίπτεται (p-τιμή<0.001).

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	722,667	4	180,667	15,268	,000
Within Groups	473,333	40	11,833		
Total	1196,000	44			

Άρα υπάρχει στατιστικά σημαντική διαφορά στη μέση επίδοση των μαθητών ως προς τις 5 μεθόδους διδασκαλίας. Για να διαπιστώσουμε ποιες μέθοδοι διδασκαλίας διαφέρουν στατιστικά σημαντικά μεταξύ τους θα αποφανθούμε χρησιμοποιώντας τα αποτελέσματα των πολλαπλών συγκρίσεων με τη μέθοδο του Tukey (είναι μία από τις μεθόδους πολλαπλών συγκρίσεων που ισχύει όταν η υπόθεση της ομοσκεδαστικότητας των πληθυσμών δεν απορρίπτεται).

Με μία πρώτη ματιά κάποιος ίσως απογοητευτεί καθώς υπάρχει πληθώρα αριθμών στον πίνακα των πολλαπλών συγκρίσεων (πίνακας Multiple Comparisons). Όμως, η ερμηνεία αυτών των αποτελεσμάτων είναι εύκολη. Στις δύο πρώτες στήλες κάθε πίνακα μας δηλώνεται ποιος πληθυσμός συγκρίνεται με ποιον και αυτό γίνεται για κάθε συνδυασμό των πληθυσμών. Το αποτέλεσμα της σύγκρισης κάθε συνδυασμού φαίνεται στην αντίστοιχη γραμμή. Έτσι, θέλοντας να συγκρίνουμε την Ομάδα Γ με την Ομάδα Ε με την μέθοδο π.χ. του Tukey θα ερμηνεύσουμε τα αποτελέσματα της γραμμής με **bold** γραφή.

Παρατήρηση (βλέπε Καρακώστας, 2002, σελ. 116)

Πολλές φορές οι πολλαπλές συγκρίσεις οδηγούν σε μη συμβατά αποτελέσματα. Για παράδειγμα μπορούμε να πάρουμε ότι το Α δε διαφέρει στατιστικά σημαντικά από το Β, το Α είναι σημαντικά καλύτερο από το Γ και το Β δε διαφέρει σημαντικά από το Γ.

Μεταξύ άλλων παρατηρούμε ότι η μέση διαφορά στην επίδοση των μαθητών των ομάδων Γ και Ε είναι ίση με 11.3333, με τυπικό σφάλμα 1.6261. Δηλαδή η μέση επίδοση των μαθητών της τρίτης ομάδας είναι υψηλότερη κατά 11.3333 βαθμούς. Επιπλέον δίπλα σε αυτή την τιμή υπάρχει ένα αστεράκι (*). Αυτό μας υποδεικνύει (βλέπε υποσημείωση του πίνακα) ότι υπάρχει στατιστικά σημαντική διαφορά στη μέση επίδοση των μαθητών που ακολουθούν τον τρόπο διδασκαλίας Γ και αυτών της Ε. Η επίδοση αυτών της Γ ομάδας είναι στατιστικά σημαντικά μεγαλύτερη (λαμβάνοντας υπόψη και το αποτέλεσμα από τη στήλη Mean Difference). Τέλος, στη στήλη 95% Confidence Interval μας δίνεται το 95% διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών.

Multiple Comparisons

Dependent Variable: Αποτέλεσμα

Tukey HSD

(I) Ομάδα	(J) Ομάδα	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower	Upper
Ομάδα Α	Ομάδα Β	1,33333	1,62161	,922	Lower	Upper
	Ομάδα Γ	-7,77778(*)	1,62161	,000	-12,4093	-3,1463
	Ομάδα Δ	-3,77778	1,62161	,157	-8,4093	,8537
	Ομάδα Ε	3,55556	1,62161	,203	-1,0759	8,1870
Ομάδα Β	Ομάδα Α	-1,33333	1,62161	,922	-5,9648	3,2981
	Ομάδα Γ	-9,11111(*)	1,62161	,000	-13,7426	-4,4796
	Ομάδα Δ	-5,11111(*)	1,62161	,024	-9,7426	-,4796
	Ομάδα Ε	2,22222	1,62161	,650	-2,4093	6,8537
Ομάδα Γ	Ομάδα Α	7,77778(*)	1,62161	,000	3,1463	12,4093
	Ομάδα Β	9,11111(*)	1,62161	,000	4,4796	13,7426
	Ομάδα Δ	4,00000	1,62161	,119	-,6315	8,6315
	Ομάδα Ε	**11,33333(*)**	**1,62161**	**,000**	**6,7019**	**15,9648**
Ομάδα Δ	Ομάδα Α	3,77778	1,62161	,157	-,8537	8,4093
	Ομάδα Β	5,11111(*)	1,62161	,024	,4796	9,7426
	Ομάδα Γ	-4,00000	1,62161	,119	-8,6315	,6315
	Ομάδα Ε	7,33333(*)	1,62161	,000	2,7019	11,9648
Ομάδα Ε	Ομάδα Α	-3,55556	1,62161	,203	-8,1870	1,0759
	Ομάδα Β	-2,22222	1,62161	,650	-6,8537	2,4093
	Ομάδα Γ	-11,33333(*)	1,62161	,000	-15,9648	-6,7019
	Ομάδα Δ	-7,33333(*)	1,62161	,000	-11,9648	-2,7019

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

Επιπρόσθετα, από τον πίνακα Homogeneous Subsets μας δίνονται οι «πιθανές» ομογενείς ομάδες μέσων μέσω της μεθόδου του Tukey. Έτσι, για το παράδειγμά μας έχουμε τις ομάδες των {Ε, Β, Α}, των {Α, Δ} και {Δ, Γ}. Από τις αντίστοιχες p-τιμές συμπεραίνουμε ότι οι ομάδες αυτές είναι ομογενείς (δηλαδή η υπόθεση της ισότητας των μέσων τιμών εντός αυτών δεν απορρίπτεται).

Homogeneous Subsets

Αποτέλεσμα

	Ομάδα	N	Subset for alpha = .05		
		1	2	3	1
Tukey HSD(a)	Ομάδα Ε	9	16,1111		
	Ομάδα Β	9	18,3333		
	Ομάδα Α	9	19,6667	19,6667	
	Ομάδα Δ	9		23,4444	23,4444
	Ομάδα Γ	9			27,4444
	Sig.		,203	,157	,119

Means for groups in homogeneous subsets are displayed.
a Uses Harmonic Mean Sample Size = 9,000.

Στον πίνακα Contrast Coefficients του Output το λογισμικό μας δίνει τους συντελεστές κάθε γραμμικής αντίθεσης που σχηματίσαμε.

Contrast Coefficients

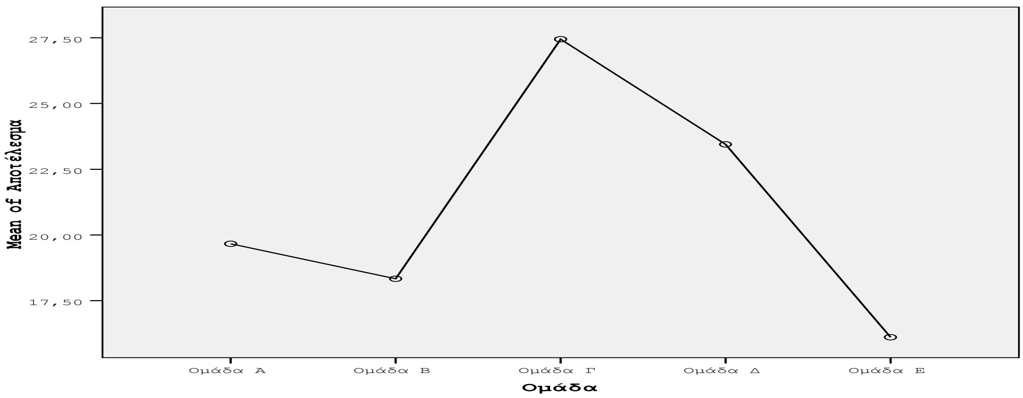
Contrast	Ομάδα				
	Ομάδα Α	Ομάδα Β	Ομάδα Γ	Ομάδα Δ	Ομάδα Ε
1	1	1	-2	0	0

Στον πίνακα Contrast Tests δίνονται τα αποτελέσματα για τους ελέγχους των γραμμικών αντιθέσεων που ζητήσαμε, ανάλογα με το αν ισχύει ή όχι η υπόθεση της ισότητας των διακυμάνσεων. Λόγω του αποτελέσματος του τεστ Levene έχουμε ότι ισχύει επομένως περιοριζόμαστε στα αποτελέσματα του πλαισίου Assume Equal Variances. Για κάθε γραμμική αντίθεση έχουμε στη στήλη Value of Contrast μία εκτίμηση και το αντίστοιχο τυπικό σφάλμα. Τέλος έχουμε την τιμή του t στατιστικού, τους βαθμούς ελευθερίας και την αντίστοιχη p-τιμή. Συμπεραίνουμε ότι η ομάδα Γ διαφέρει στατιστικά σημαντικά από τις ομάδες Α και Β.

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Αποτέλεσμα	Assume equal variances	1	-16,8889	2,80872	-6,013	40	,000
	Does not assume equal variances	1	-16,8889	2,46331	-6,856	22,530	,000

Τέλος, μας δίνεται το γράφημα των μέσων τιμών. Η εικόνα αυτή πολλές φορές μπορεί να μας ξεγελάσει, καθώς θα πρέπει να είμαστε προσεκτικοί στην κλίμακα μέτρησης που χρησιμοποιείται στον κατακόρυφο άξονα. Επιπλέον, λαμβάνοντας υπόψη ότι δε μας δίνεται καμία πρόσθετη και ουσιαστική πληροφορία σε σχέση με αυτές της στήλης Mean του πίνακα Descriptives προτείνεται η αποφυγή δημιουργίας της.



4. <u>Μη παραμετρικός έλεγχος</u>: Σε αυτή την περίπτωση το πρόβλημα ελέγχου αντιμετωπίζεται με το μη παραμετρικό έλεγχο των Kruskal and Wallis (1952), το οποίο αποτελεί επέκταση του τεστ των Mann-Whitney σε περισσότερους από δύο πληθυσμούς. Αποδεικνύεται ότι το στατιστικό

$$KW = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left(R_i - \frac{n_i (n+1)}{2} \right)^2 ,$$

όπου $R_i = \sum_{j=1}^{n_i} R(X_{ij})$, $i = 1,...,k$, $k \geq 3$, με $R(X_{ij})$ $i = 1,...,k$, $j = 1,...,n_i$, $k \geq 3$, οι τάξεις των διαθέσιμων δειγματικών τιμών των k δειγμάτων στο σύνολο των n παρατηρήσεων, στην περίπτωση μη ύπαρξης δεσμών ακολουθεί προσεγγιστικά υπό τη μηδενική υπόθεση X^2 κατανομή με $k-1$ βαθμούς ελευθερίας και η μηδενική υπόθεση απορρίπτεται αν $KW \geq X^2_{k-1,1-a}$.

<u>Σχόλιο</u>: Σε περίπτωση ύπαρξης δεσμών ανάμεσα στις δειγματικές παρατηρήσεις οι Kruskal-Wallis πρότειναν το στατιστικό: $KW^* = \dfrac{KW}{1 - \dfrac{\sum (d_i^3 - d_i)}{n^3 - n}}$, όπου KW είναι το σύνηθες στατιστικό των Kruskal-Wallis υπολογισμένο χρησιμοποιώντας τα midranks και d_i είναι ο αριθμός των δεσμών στο i δείγμα.

Παρατήρηση Το S.P.S.S. διεξάγει δύο ελέγχους για το παραπάνω πρόβλημα. Το τεστ των διαμέσων (median test) που ελέγχει την ισότητα των πληθυσμιακών διαμέσων και το τεστ των Kruskal-Wallis test που ουσιαστικά ελέγχει αν τα δείγματα προέρχονται από τον ίδιο πληθυσμό.

<u>Υλοποίηση στο S.P.S.S.</u> (Kruskal and Wallis (1952))

Σε ένα εργοστάσιο τρεις μηχανές χρησιμοποιούνται για την παραγωγή δοχείων εμφιαλώσεως. Κατά τη διάρκεια μιας εργάσιμης εβδομάδας καταγράφεται ο αριθμός των δοχείων που κατασκευάστηκαν από κάθε μηχανή και τα αποτελέσματα παρατίθενται στον πίνακα που ακολουθεί, με την επιπλέον επισήμανση ότι κάποιες μέρες δεν παρήχθησαν δοχεία από κάποιες μηχανές λόγω ότι είχαν τεθεί εκτός λειτουργίας.

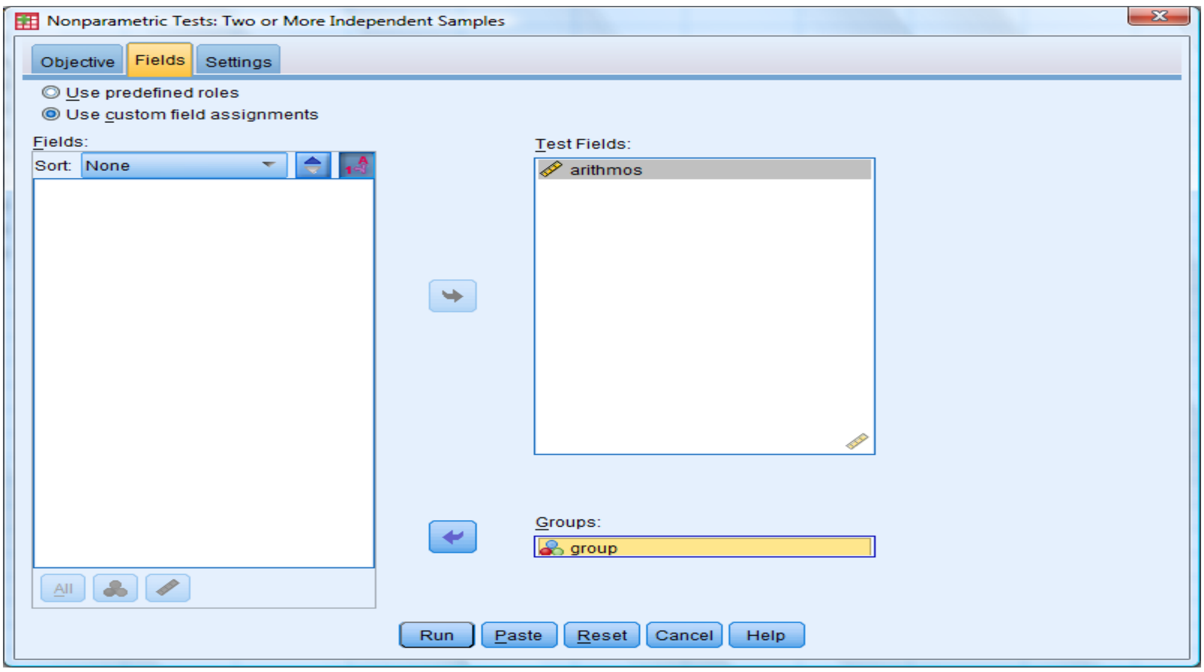
Μηχανή 1	340	345	330	342	338
Μηχανή 2	339	333	344		
Μηχανή 3	347	343	349	355	

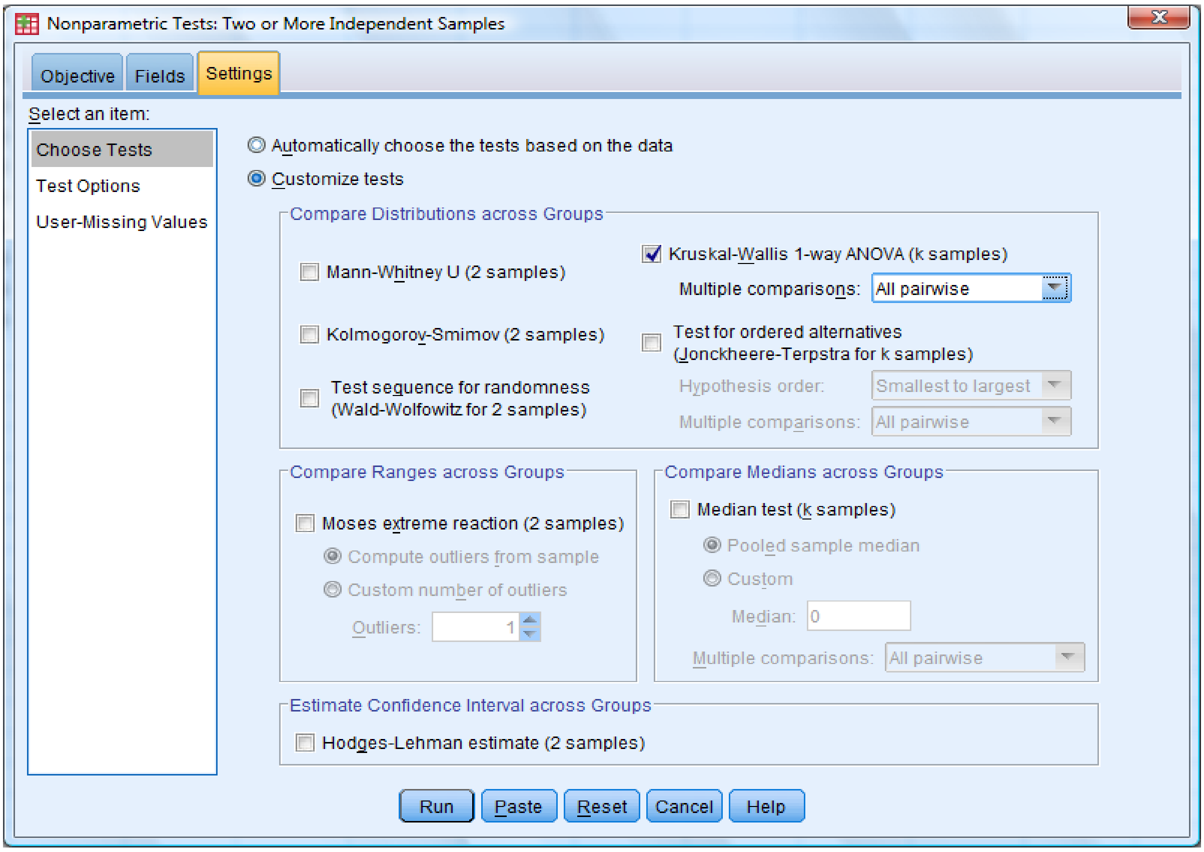
Θέλουμε να ελέγξουμε αν υπάρχει στατιστικά σημαντική διαφορά ως προς την πληθυσμιακή διάμεσο του αριθμού των δοχείων που παράγουν οι τρεις μηχανές. (Σχόλιο: να διαπιστώσετε ότι όντως χρειάζεται μη παραμετρικός έλεγχος)

Από το κύριο μενού επιλέγουμε

i. Analyze→Non Parametric Tests→ Independent Samples.

ii. Στο νέο παράθυρο διαλόγου που προκύπτει επιλέγουμε στο πλαίσιο Objective την επιλογή Customize analysis, έτσι ώστε στη συνέχεια από τα πλαίσια Fields και Settings να καθορίσουμε τον έλεγχο τον οποίο θέλουμε να διενεργηθεί.





Στο πλαίσιο Test Fields τοποθετούμε την υπό μελέτη ποσοτική μεταβλητή, ενώ στο πλαίσιο Groups την ποιοτική μεταβλητή η οποία μας διαχωρίζει τους k πληθυσμούς. Από το πλαίσιο Settings επιλέγουμε Customize Tests και Kruskal-Wallis 1-way Anova.

Ερμηνεία αποτελεσμάτων

Συμπεραίνουμε ότι δεν υπάρχει στατιστικά σημαντική διαφορά στην πληθυσμιακή διάμεσο του αριθμού των δοχείων που παρήχθησαν ως προς τις 3 μηχανές. Τα αποτελέσματα για να γενικευτούν στις μέσες τιμές θα πρέπει να είναι συμμετρικοί οι πληθυσμοί από όπου προέρχονται τα δεδομένα μας. (Σχόλιο: στο συγκεκριμένο παράδειγμα οι πληθυσμοί μπορούν να θεωρηθούν συμμετρικοί. Μπορείτε να το διαπιστώσετε; Επιπλέον ποιο το αποτέλεσμα του Median Test)

Hypothesis Test Summary

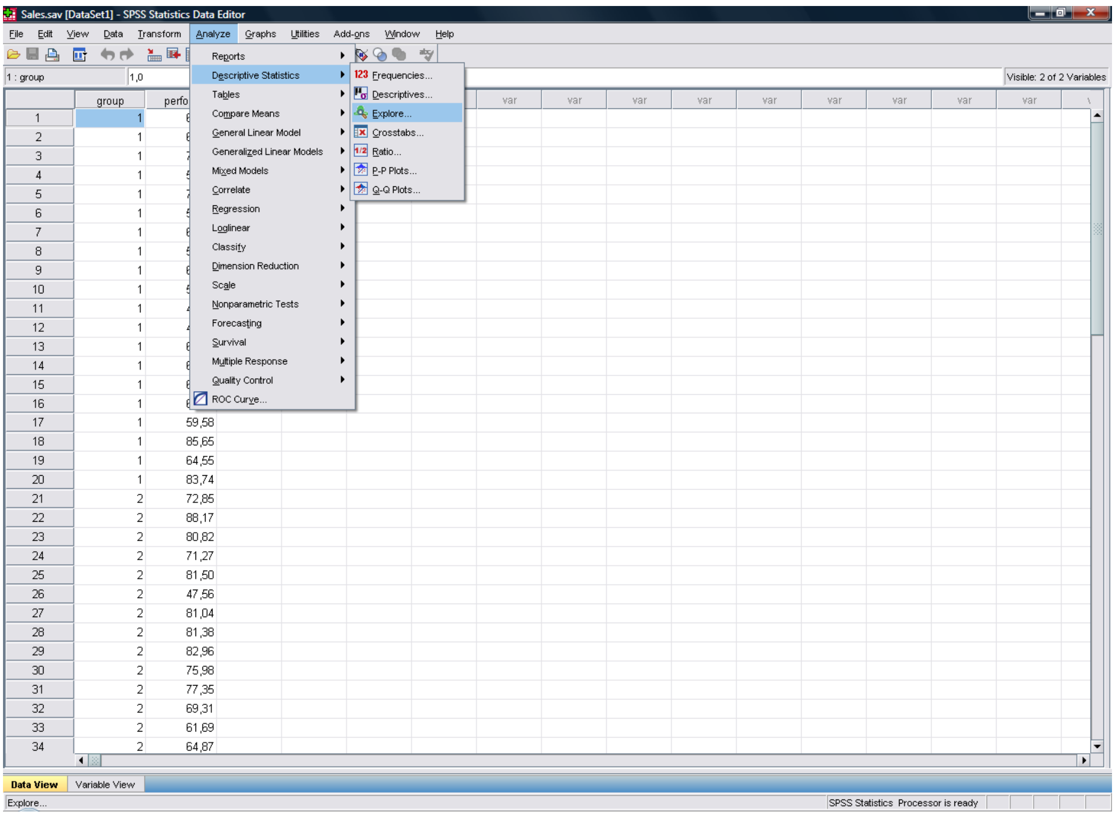
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of arithmos is the same across categories of group.	Independent-Samples Kruskal-Wallis Test	.059	Retain the null hypothesis.

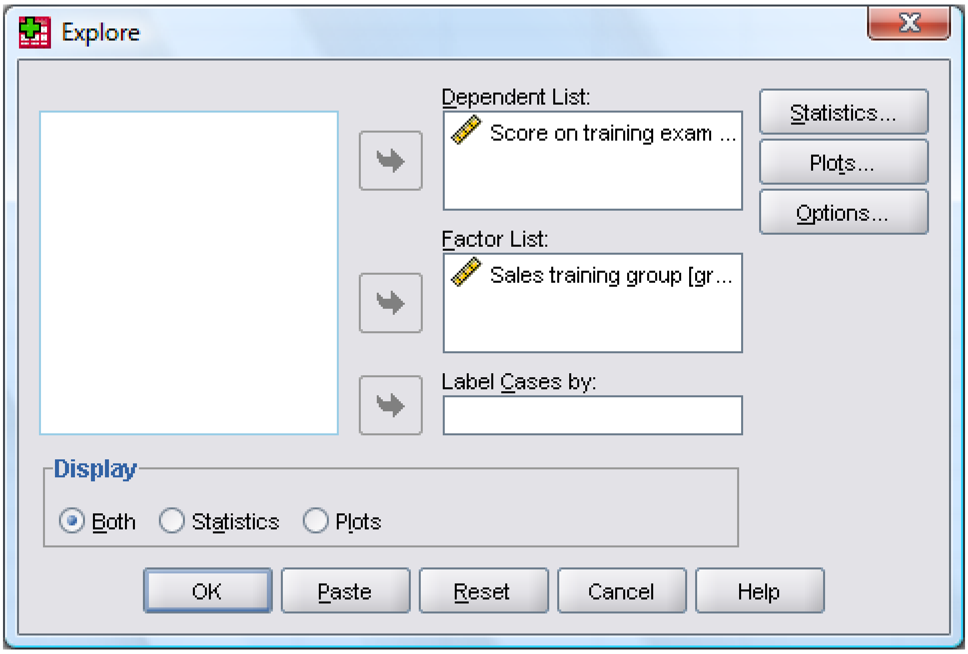
Asymptotic significances are displayed. The significance level is .05.

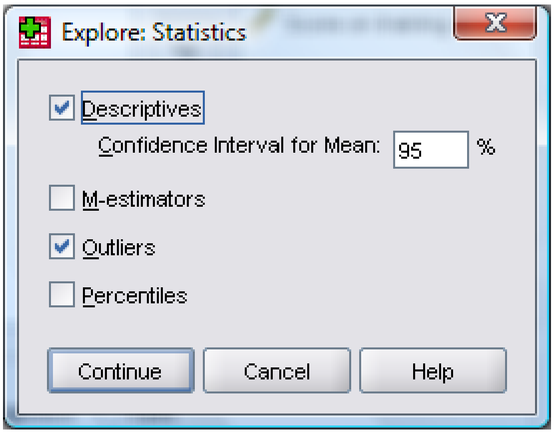
7.2 Παραδείγματα

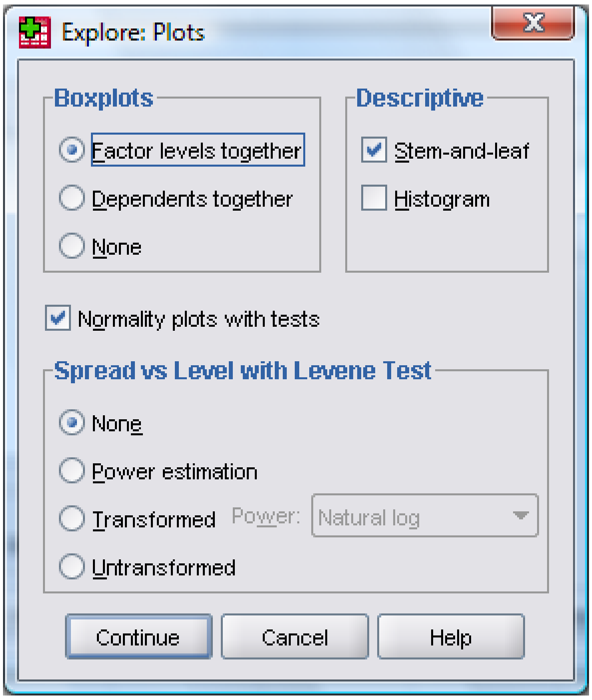
<u>Παράδειγμα 1ο</u> Στο αρχείο Sales.sav* καταγράφονται οι αποδόσεις (στήλη Perform) 60 τυχαία επιλεγμένων πωλητών μιας εταιρείας οι οποίοι έχουν χωριστεί σε τρεις ομάδες (Group). Θέλουμε να ελέγξουμε αν είναι εφικτό υπάρχει στατιστικά σημαντική διαφορά στη μέση απόδοσή των πωλητών ανάλογα με την ομάδα που ανήκουν.

Αρχικά θα ελέγξουμε την ύπαρξη ακραίων τιμών στις δειγματικές τιμές που καταγράφεται η επίδοση των 3 ομάδων πωλητών. Για το σκοπό αυτό ακολουθούμε τα βήματα:

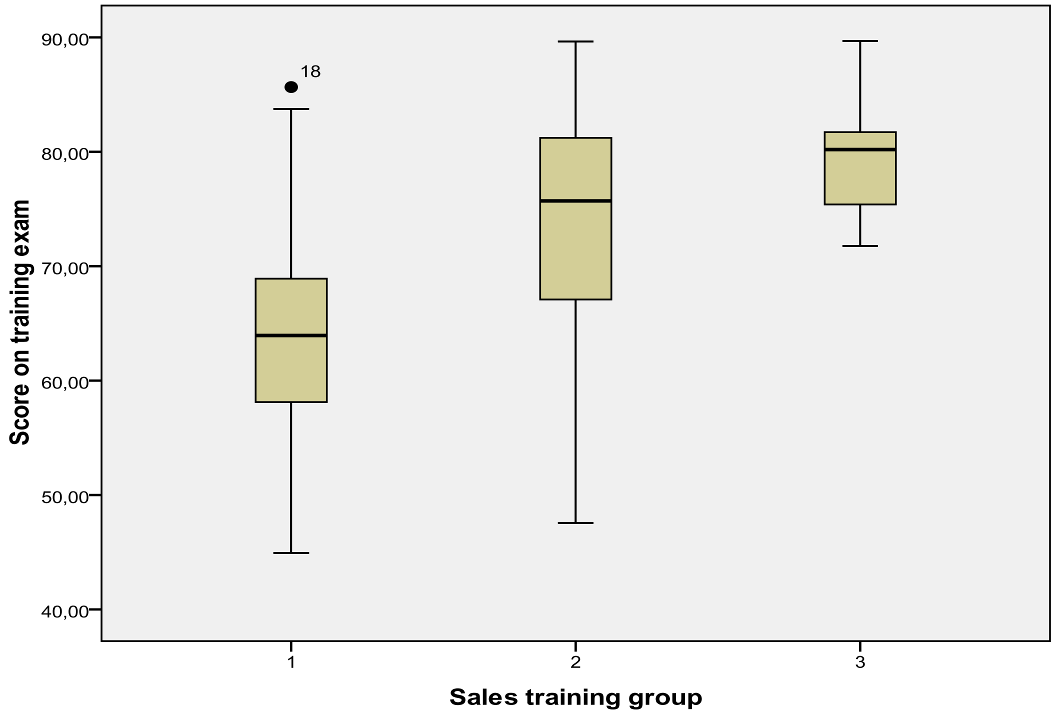



Explore
Dependent List:
Score on training exam ...
Factor List:
Sales training group [gr...
Label Cases by:
Statistics...
Plots...
Options...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help

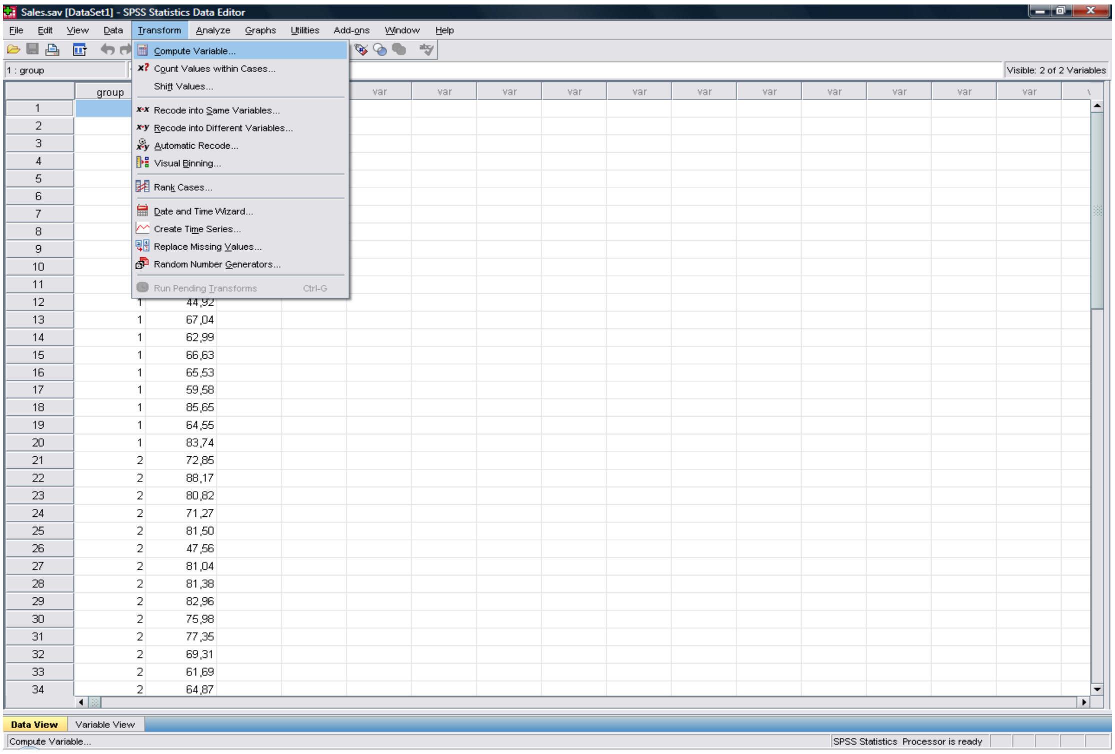

Explore: Statistics
Descriptives
Confidence Interval for Mean: 95 %
M-estimators
Outliers
Percentiles
Continue
Cancel
Help

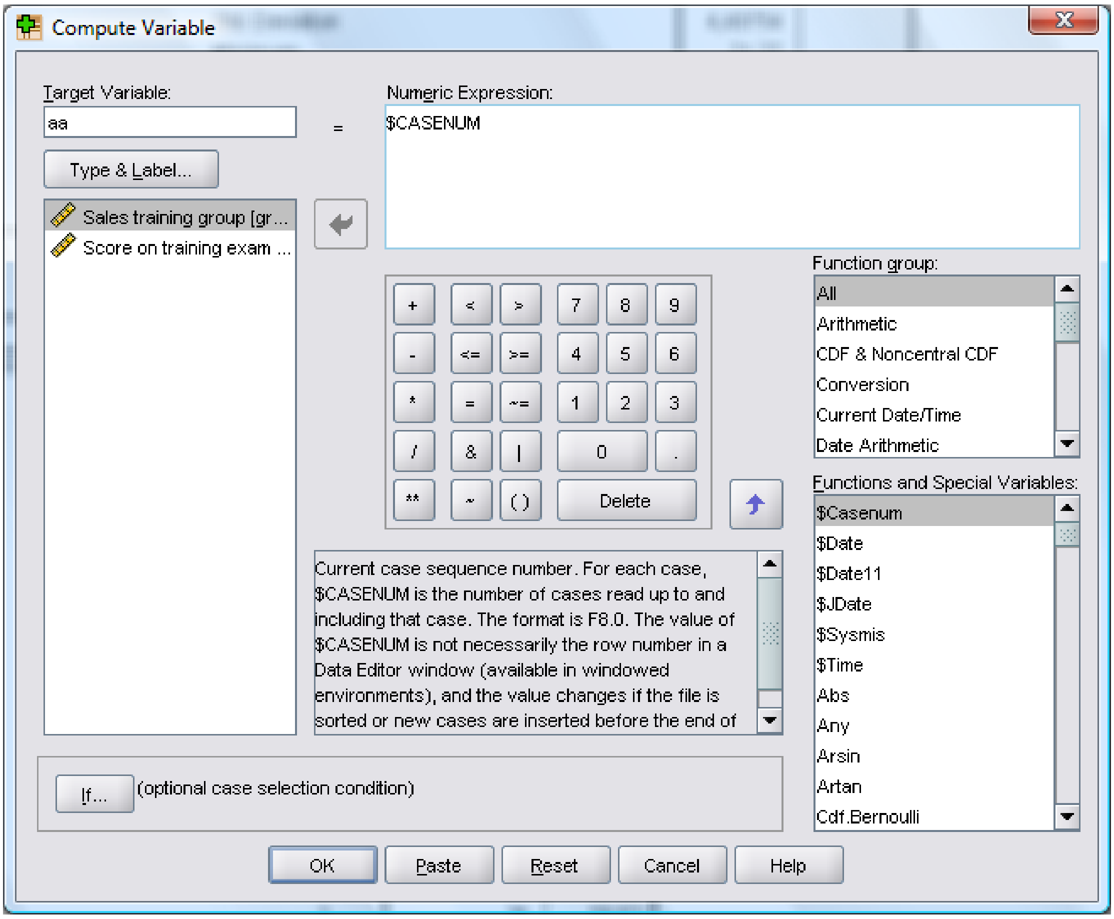

Explore: Plots
Boxplots
Factor levels together
Dependents together
None
Descriptive
Stem-and-leaf
Histogram
Normality plots with tests
Spread vs Level with Levene Test
None
Power estimation
Transformed
Power:
Natural log
Untransformed
Continue
Cancel
Help

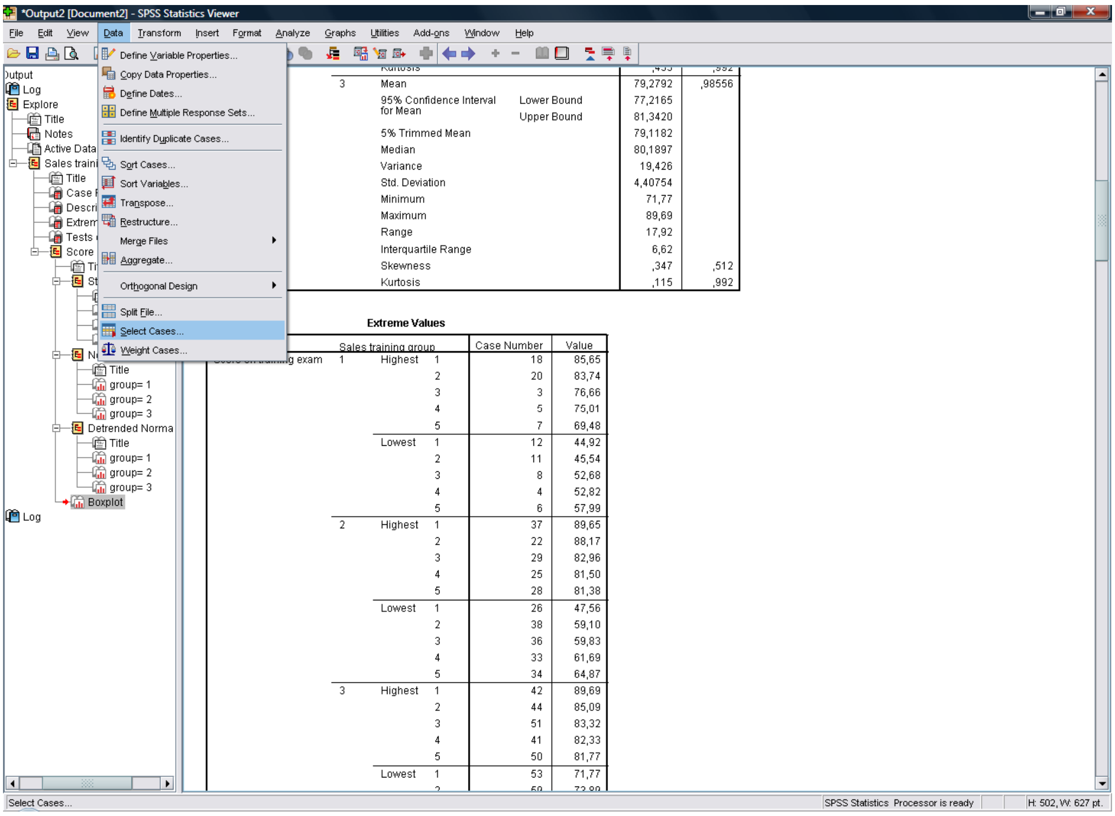
Παρατηρούμε ότι από το θηκόγραμμα που ακολουθεί πρέπει να αποκλειστεί από την περαιτέρω ανάλυση η παρατήρηση με α/α 18 που προέρχεται από την πρώτη ομάδα πωλητών και έχει τιμή απόδοσης 85,65, λόγω ότι είναι ακραία

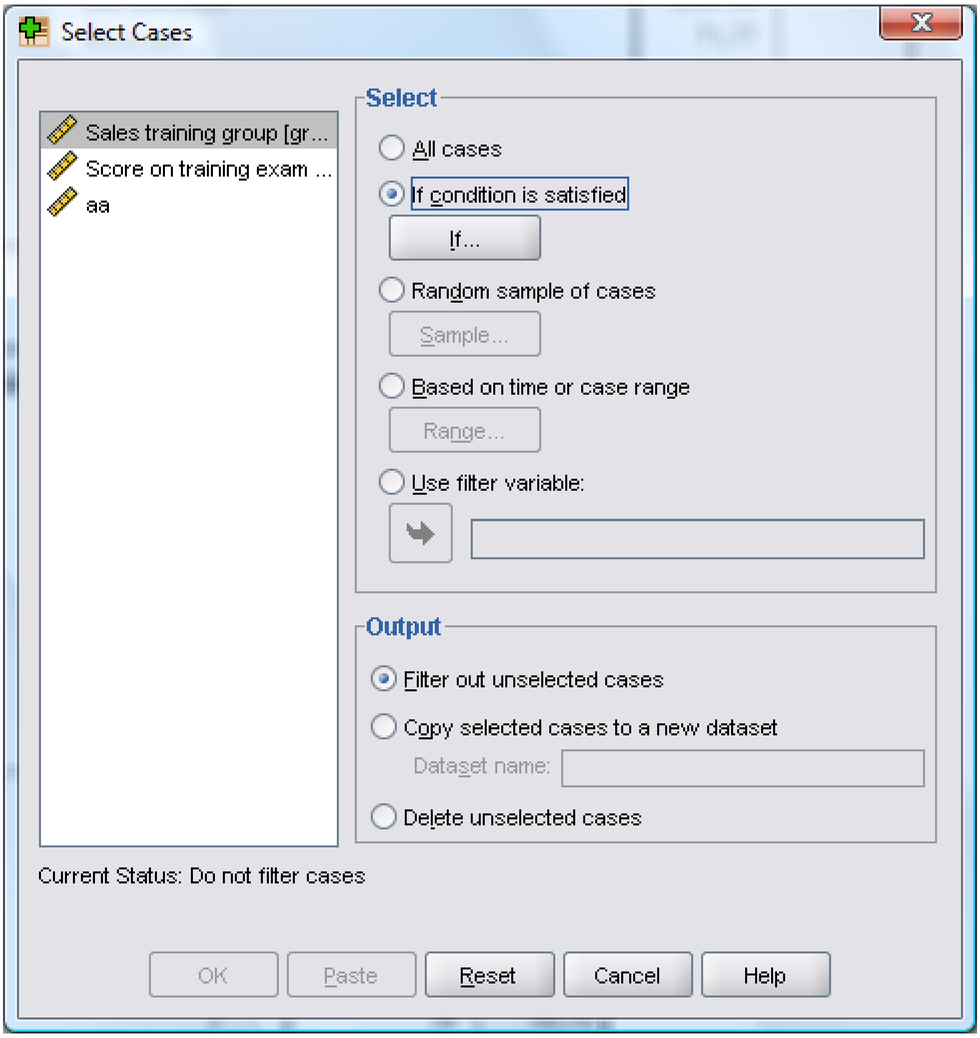


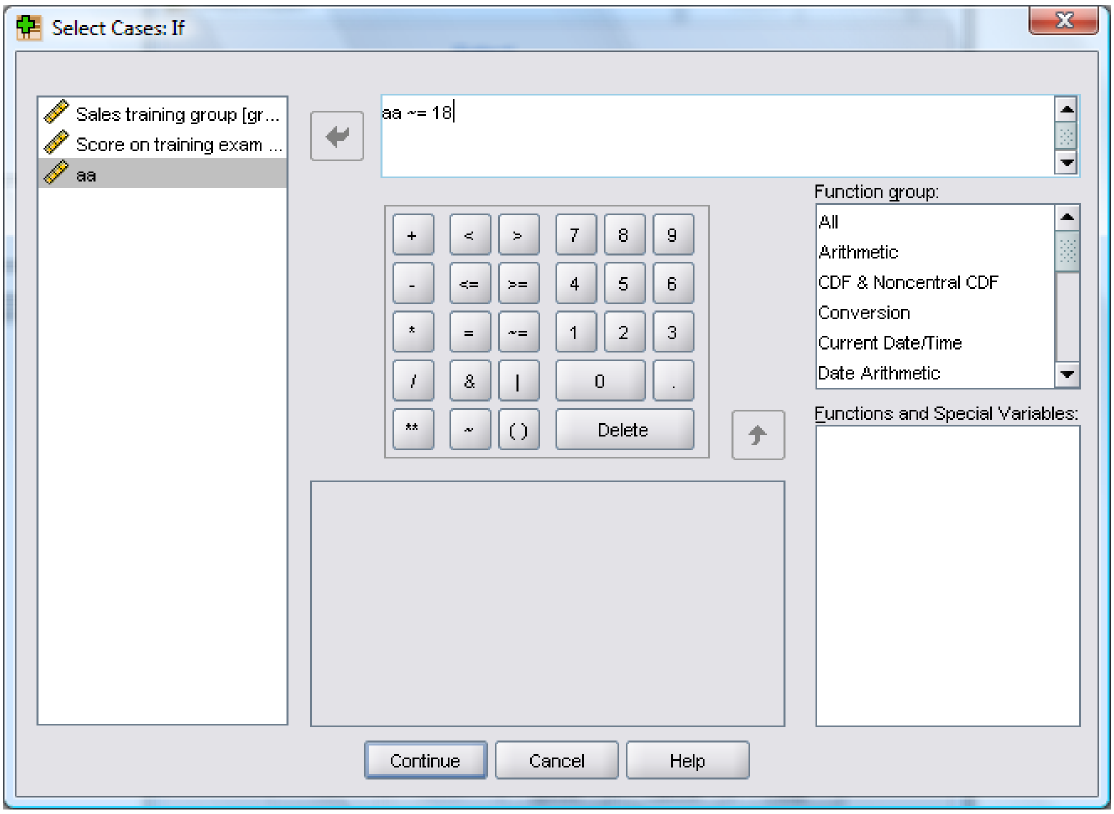
Δημιουργούμε κατά τα γνωστά μία νέα μεταβλητή με τον αύξοντα αριθμό παρατήρησης ως βοηθητική για τον αποκλεισμό της παρατήρησης 18 από την περαιτέρω ανάλυση.



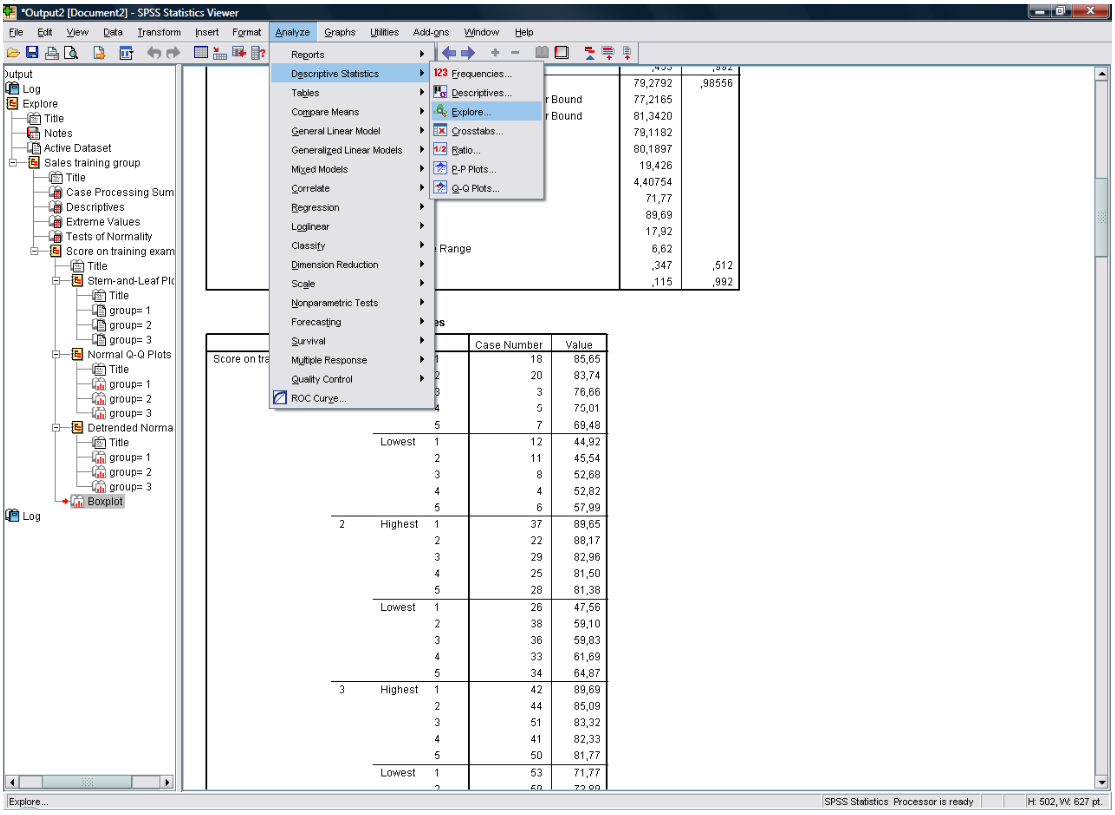

Compute Variable
Target Variable:
aa
=
Numeric Expression:
$CASENUM
Type & Label...
Sales training group [gr...
Score on training exam ...
+ < > 7 8 9
- <= >= 4 5 6
* = ~= 1 2 3
/ & | 0 .
** ~ ( ) Delete
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Functions and Special Variables:
$Casenum
$Date
$Date11
$JDate
$Sysmis
$Time
Abs
Any
Arsin
Artan
Cdf.Bernoulli
Current case sequence number. For each case, $CASENUM is the number of cases read up to and including that case. The format is F8.0. The value of $CASENUM is not necessarily the row number in a Data Editor window (available in windowed environments), and the value changes if the file is sorted or new cases are inserted before the end of
If... (optional case selection condition)
OK
Paste
Reset
Cancel
Help

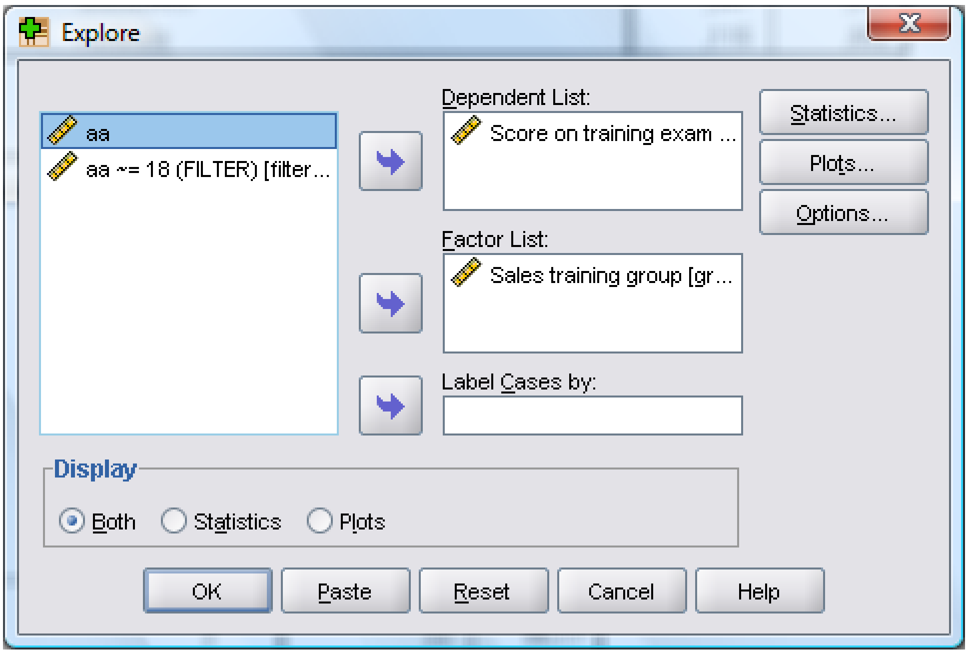

*Output2 [Document2] - SPSS Statistics Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help
Define Variable Properties...
Copy Data Properties...
Define Dates...
Define Multiple Response Sets...
Identify Duplicate Cases...
Sort Cases...
Sort Variables...
Transpose...
Restructure...
Merge Files
Aggregate...
Orthogonal Design
Split File...
Select Cases...
Weight Cases...
Extreme Values
Select Cases...
SPSS Statistics Processor is ready

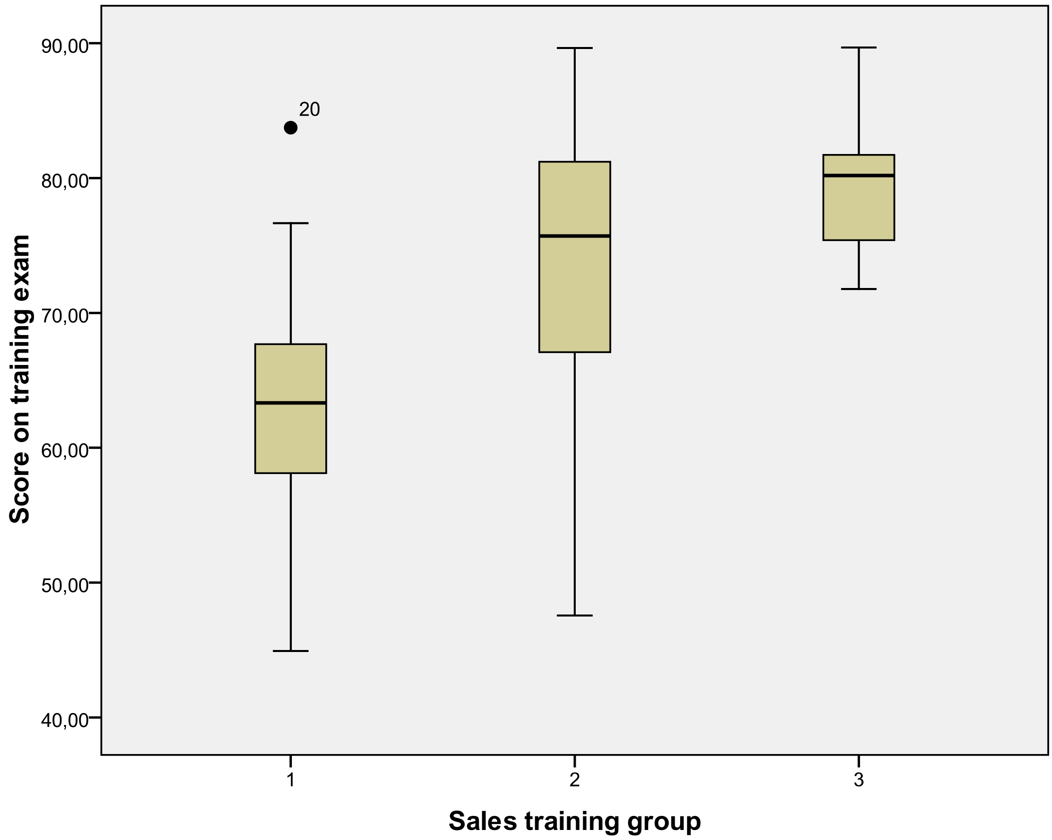

Select Cases
Sales training group [gr...
Score on training exam ...
aa
Select
All cases
If condition is satisfied
If...
Random sample of cases
Sample...
Based on time or case range
Range...
Use filter variable:
Output
Filter out unselected cases
Copy selected cases to a new dataset
Dataset name:
Delete unselected cases
Current Status: Do not filter cases
OK
Paste
Reset
Cancel
Help


Select Cases: If
Sales training group [gr...
Score on training exam ...
aa
aa ~= 18
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Functions and Special Variables:
Delete
Continue
Cancel
Help

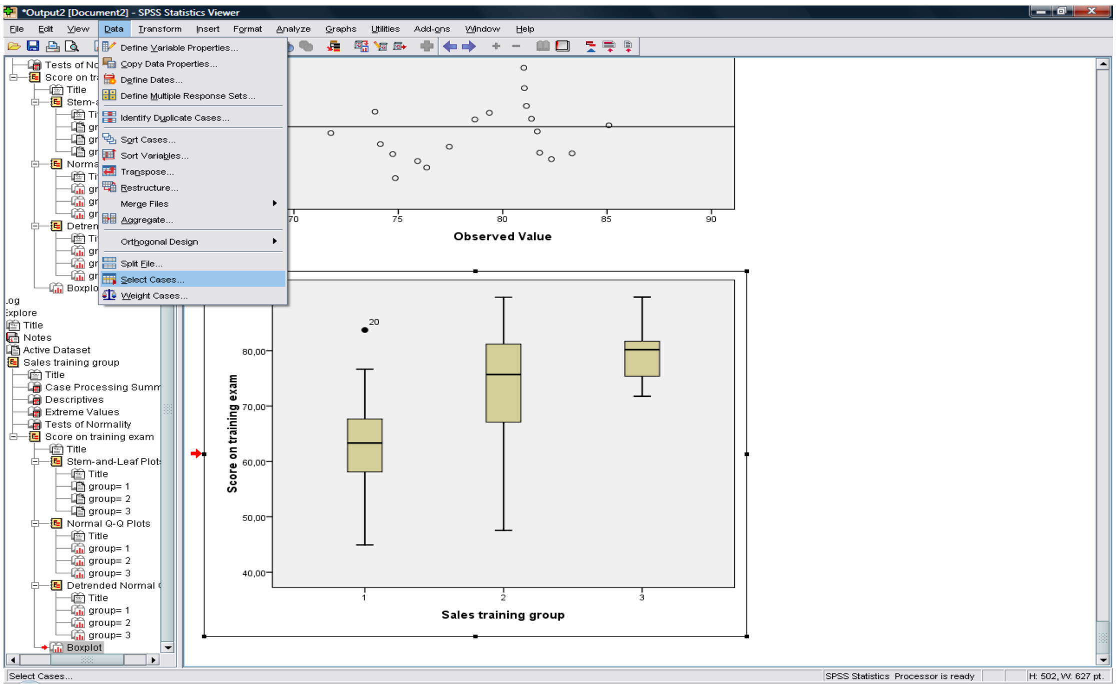
Στη συνέχεια προβαίνουμε πάλι σε έλεγχο ύπαρξης ακραίων τιμών.

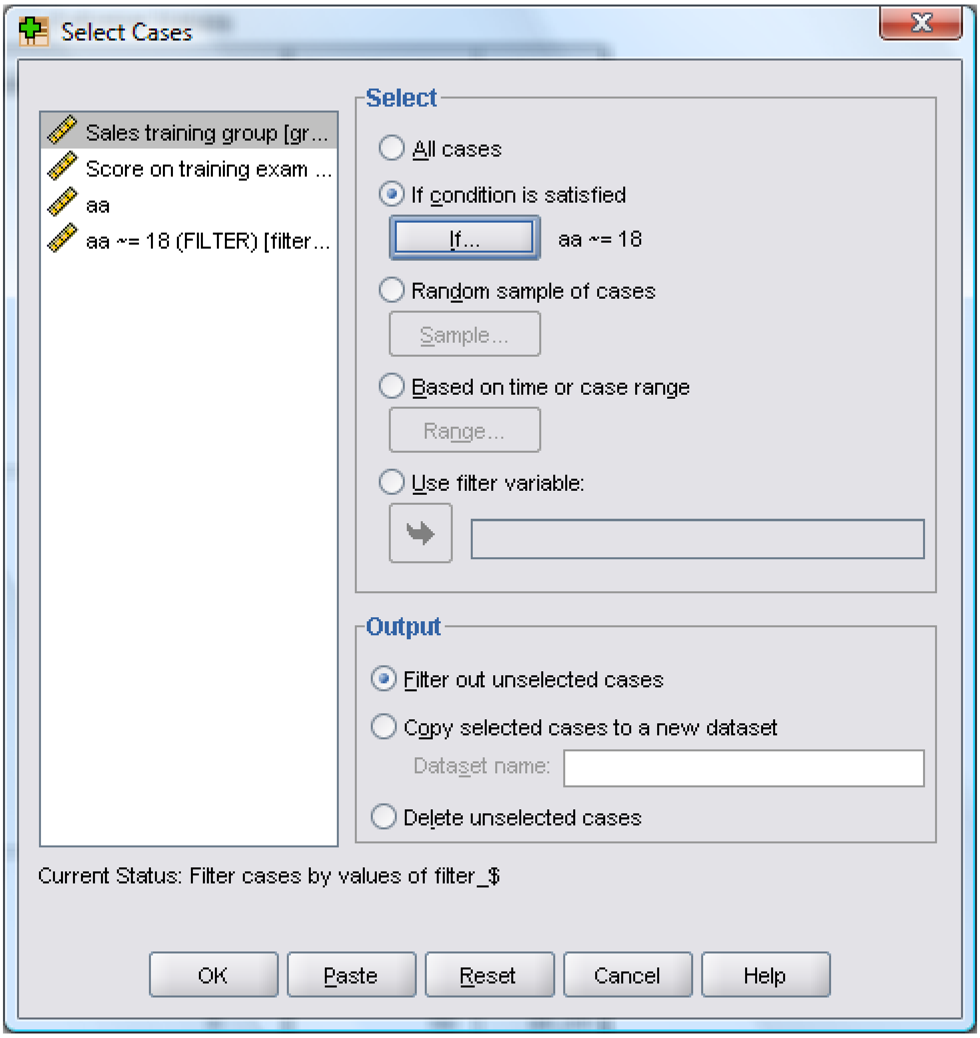


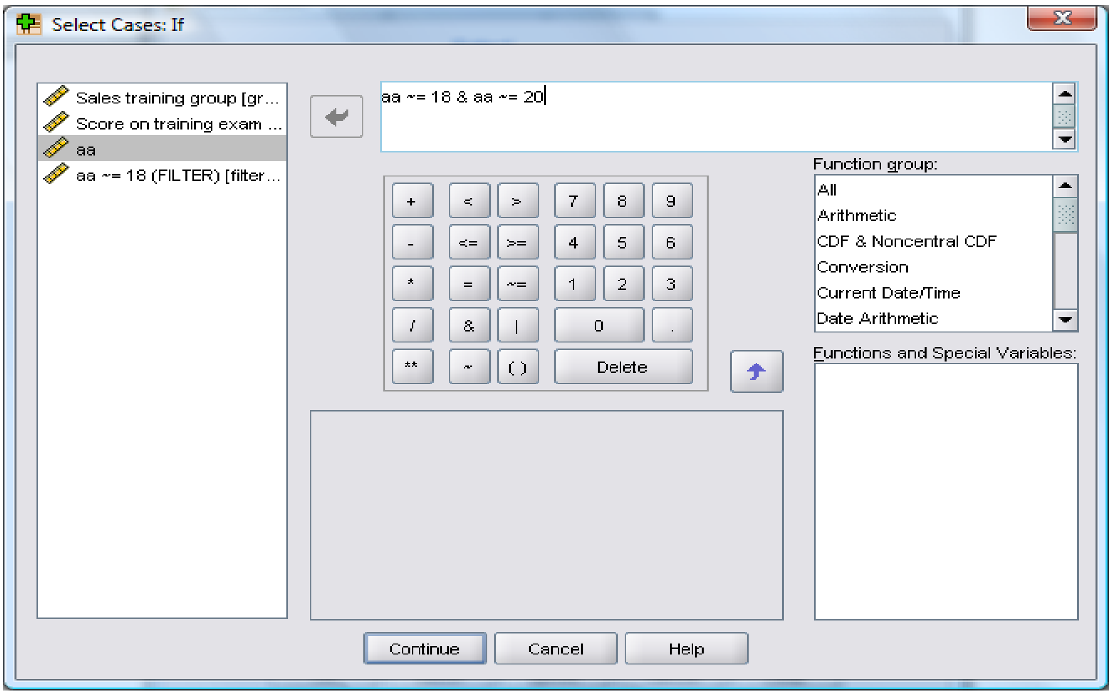




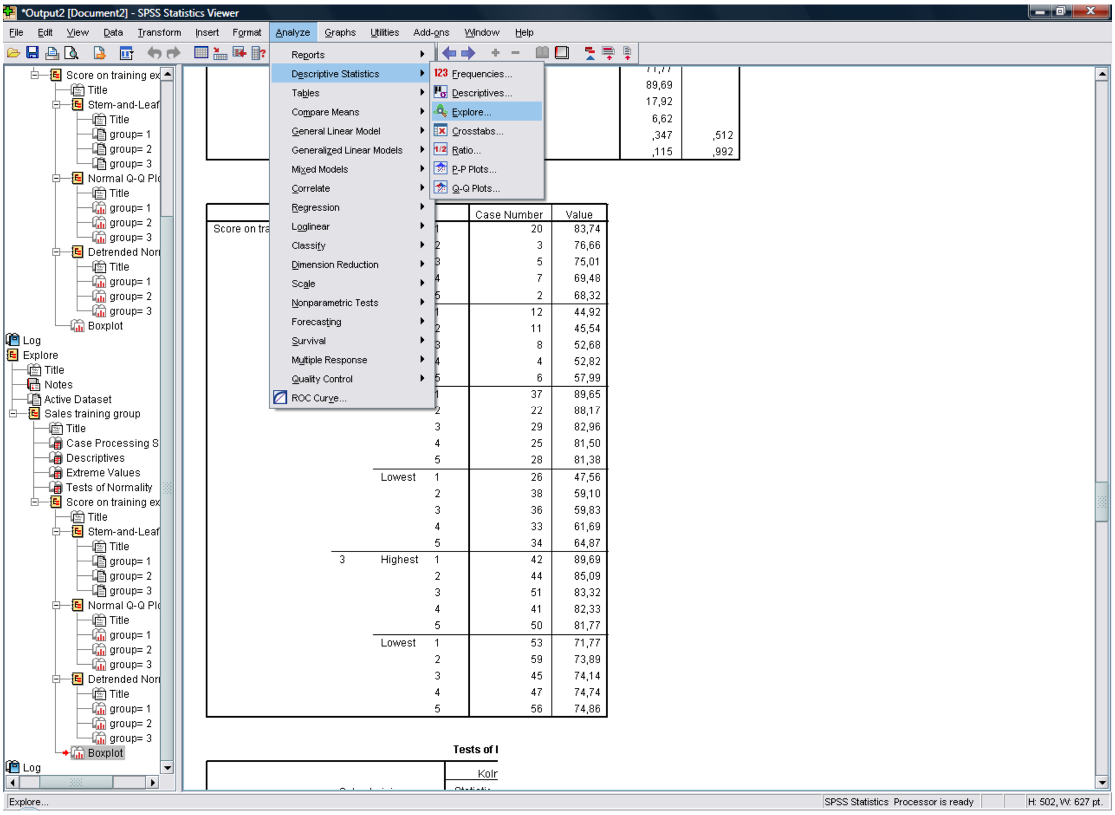
Θα αποκλειστεί επομένως και η παρατήρηση με αύξοντα αριθμό 20 (το ποσοστό των ακραίων είναι τώρα 2/20* 100%=10%) με τιμή 83,74. Για το σκοπό αυτό επιλέγουμε τα ακόλουθα μέσω της διαδικασίας Data Select Cases. Παρατήρηση: Η 20 δεν εμφανιζόταν πριν ως πιθανή ακραία.

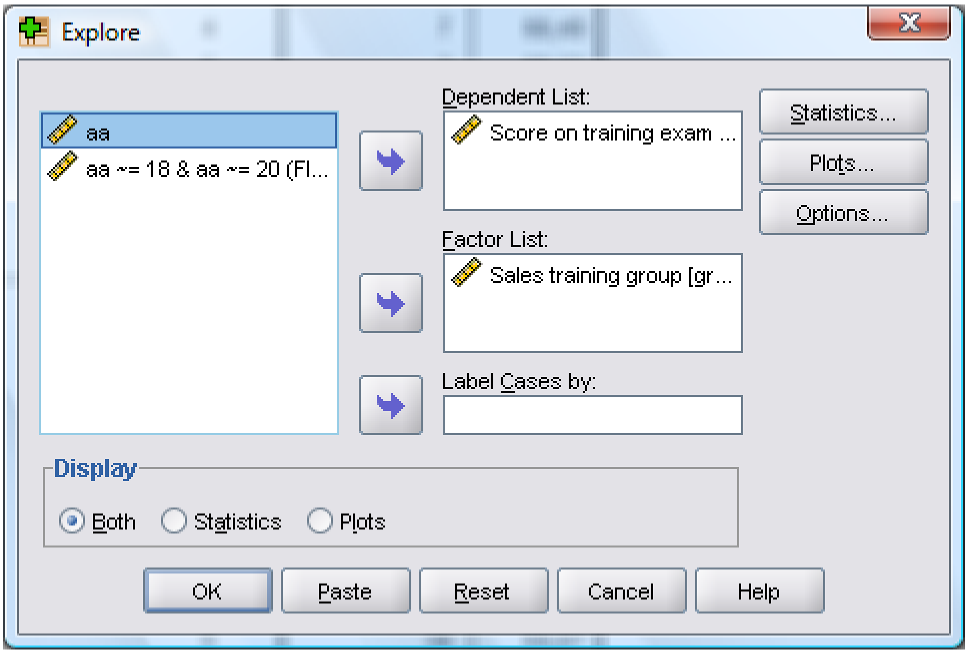


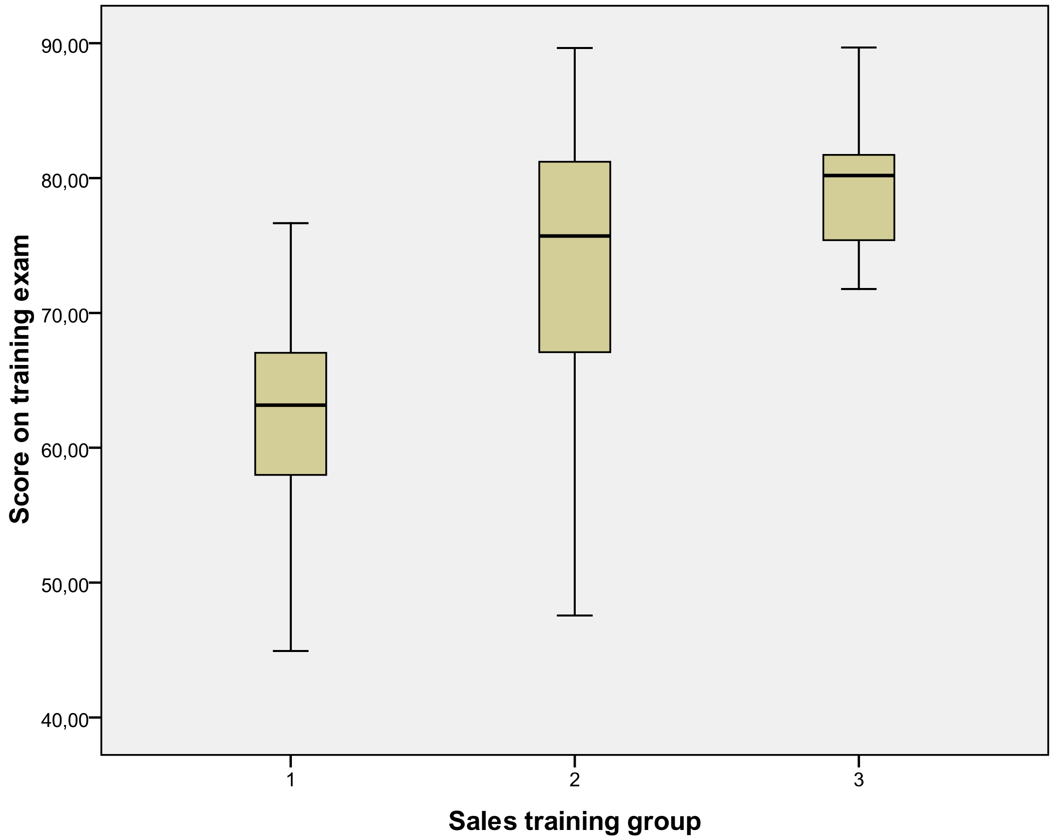

Select Cases
Sales training group [gr...
Score on training exam ...
aa
aa ~= 18 (FILTER) [filter...
Select
All cases
If condition is satisfied
If...
aa ~= 18
Random sample of cases
Sample...
Based on time or case range
Range...
Use filter variable:
Output
Filter out unselected cases
Copy selected cases to a new dataset
Dataset name:
Delete unselected cases
Current Status: Filter cases by values of filter_$
OK
Paste
Reset
Cancel
Help


Select Cases: If
Sales training group [gr...
Score on training exam ...
aa
aa ~= 18 (FILTER) [filter...
aa ~= 18 & aa ~= 20
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Functions and Special Variables:
Delete
Continue
Cancel
Help

Επαναλαμβάνουμε τον έλεγχο ύπαρξης ακραίων τιμών







Συνοψίζοντας, από τα παραπάνω θηκογράμματα προκύπτει ότι υπάρχουν δύο ακραίες τιμές στις δειγματικές τιμές της επίδοσης της πρώτης ομάδας, ενώ δεν υπάρχουν ακραίες τιμές στις δειγματικές τιμές της επίδοσης των υπολοίπων ομάδων. Καθώς το ποσοστό των ακραίων τιμών εντός της πρώτης ομάδας δεν ξεπερνά το 10% (2/10*100%=10%) συνεχίζουμε την περαιτέρω ανάλυση, έχοντας αποκλείσει από την περαιτέρω ανάλυση της προαναφερθείσες δειγματικές τιμές.

Προσοχή: Τις ακραίες τιμές τις αποκλείουμε μία μία για κάθε «ομάδα», ξεκινώντας από την πιο απομακρυσμένη της ομάδας. Το ποσοστό 10% δεν το υπολογίζουμε στο σύνολο των παρατηρήσεων αλλά στον αριθμό των παρατηρήσεων εντός κάθε ομάδας.

Έπειτα ελέγχουμε αν οι δειγματικές τιμές της επίδοσης των τριών ομάδων προέρχονται από κανονικούς πληθυσμούς (τεστ Shapiro Wilk έχει ήδη ζητηθεί η υλοποίηση του στο προηγούμενο βήμα).

Tests of Normality

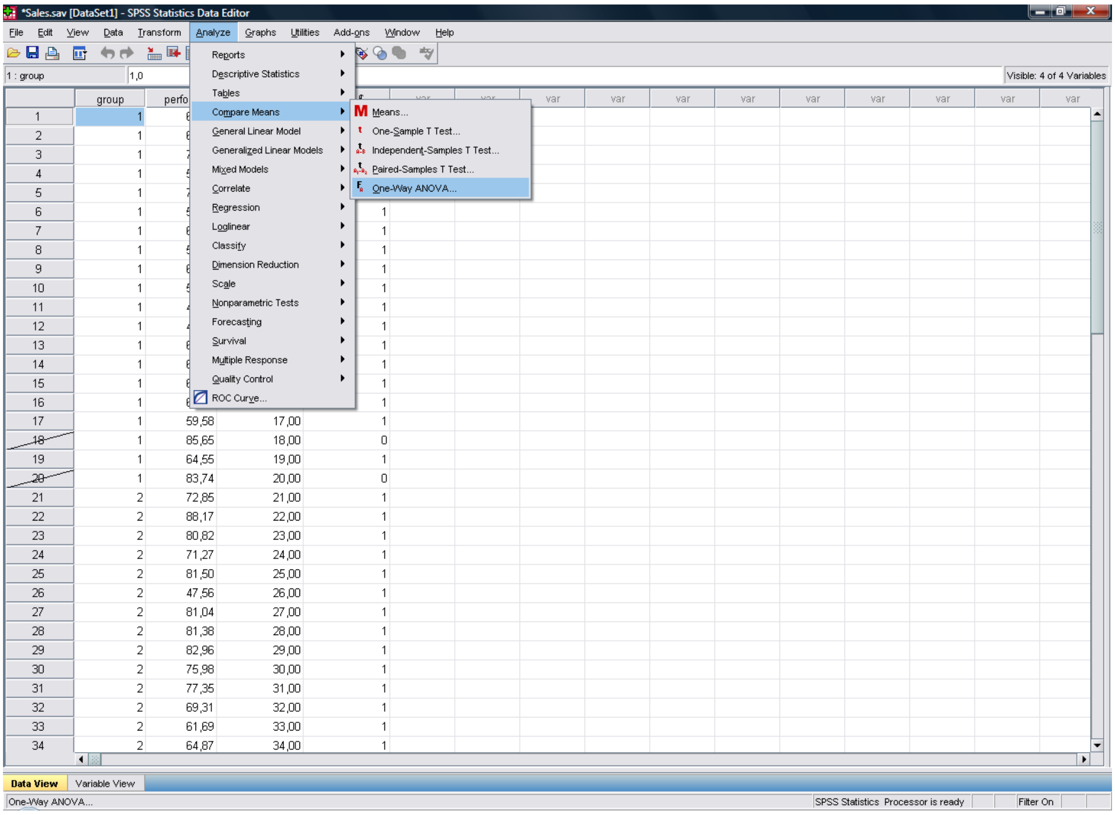
	Sales training group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Score on training exam	1	,111	18	,200*	,964	18	,676
	2	,134	20	,200*	,948	20	,344
	3	,153	20	,200*	,962	20	,582

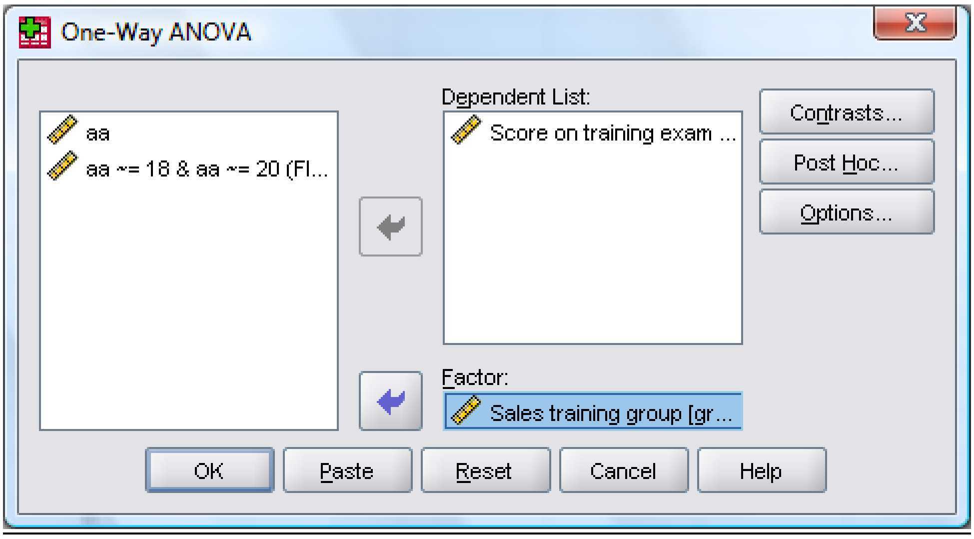
a. Lilliefors Significance Correction

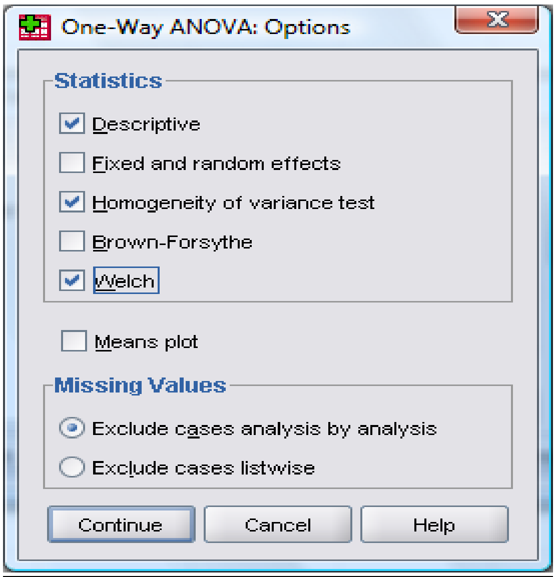
*. This is a lower bound of the true significance.

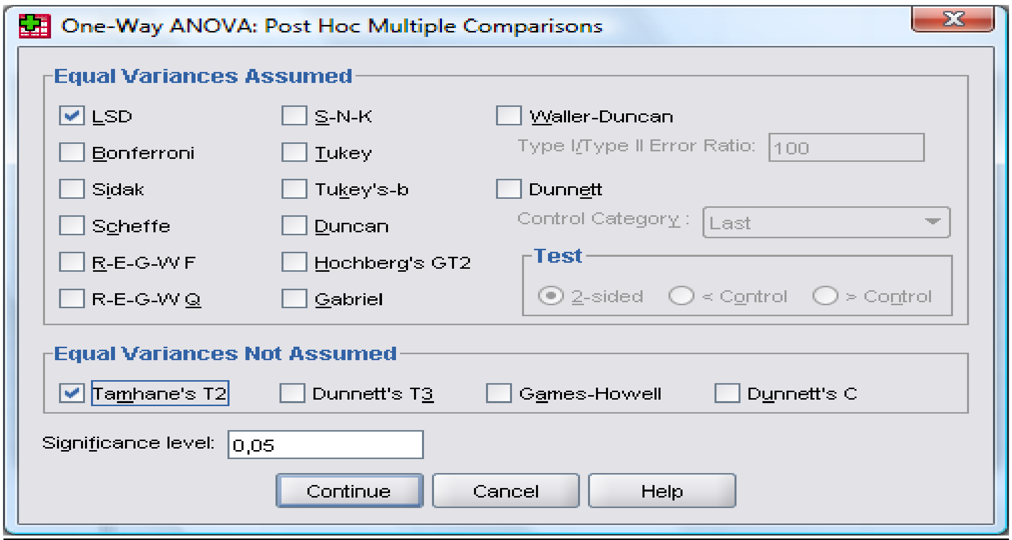
Η υπόθεση της κανονικότητας δεν απορρίπτεται για κάποιον από τους τρεις πληθυσμούς και θα εξετάσουμε την προς έλεγχο υπόθεση παραμετρικά.

Προσοχή; Παρότι δηλώνουμε τις παρακάτω επιλογές δεν χρειάζονται όλες αυτές οι επιλογές, όπως αναλυτικά θα εξηγηθεί.




One-Way ANOVA
aa
aa ~= 18 & aa ~= 20 (FI...
Dependent List:
Score on training exam ...
Factor:
Sales training group [gr...
Contrasts...
Post Hoc...
Options...
OK
Paste
Reset
Cancel
Help


One-Way ANOVA: Options
Statistics
Descriptive
Fixed and random effects
Homogeneity of variance test
Brown-Forsythe
Welch
Means plot
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue
Cancel
Help


One-Way ANOVA: Post Hoc Multiple Comparisons
Equal Variances Assumed
LSD
S-N-K
Waller-Duncan
Bonferroni
Tukey
Type I/Type II Error Ratio:
100
Sidak
Tukey's-b
Dunnett
Scheffe
Duncan
Control Category :
Last
R-E-G-W F
Hochberg's GT2
Test
R-E-G-W Q
Gabriel
2-sided
< Control
> Control
Equal Variances Not Assumed
Tamhane's T2
Dunnett's T3
Games-Howell
Dunnett's C
Significance level:
0,05
Continue
Cancel
Help

Από τη θεωρία γνωρίζουμε ότι αν χρησιμοποιηθεί το F-test του πίνακα Ανάδια ή το τεστ του Welch καθορίζεται από την ικανοποίηση ή όχι της ισότητας των πληθυσμιακών διακυμάνσεων. Η υπόθεση αυτή ελέγχεται από το στατιστικό τεστ του Levene. Όταν η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων απορρίπτεται χρησιμοποιούμε το στατιστικό του Welch. Στο συγκεκριμένο παράδειγμα αν πούμε ότι θέλουμε να εργαστούμε με επίπεδο σημαντικότητας 5% τότε απορρίπτεται η ισότητα των πληθυσμιακών διακυμάνσεων. Και επομένως θα χρησιμοποιηθεί το τεστ του Welch.

Test of Homogeneity of Variances

Score on training exam

Levene Statistic	df1	df2	Sig.
4,511	2	55	,015

Προκύπτει ότι υπάρχουν στατιστικά σημαντικές διαφορές στη μέση επίδοση των 3 πληθυσμών (τεστ Welch p-τιμή <0.001).

Robust Tests of Equality of Means

Score on training exam

	Statistic[a]	df1	df2	Sig.
Welch	28,831	2	30,976	,000

a. Asymptotically F distributed.

Προβαίνουμε σε πολλαπλές συγκρίσεις και καθώς έχει απορριφθεί η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων θα στηριχθούμε στα αποτελέσματα της μεθόδου Tamhanne.

Multiple Comparisons

Dependent Variable:Score on training exam

	(I) Sales training group	(J) Sales training group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	-11,77844*	2,70979	,000	-17,2090	-6,3479
		3	-17,49002*	2,70979	,000	-22,9206	-12,0595
	2	1	11,77844*	2,70979	,000	6,3479	17,2090
		3	-5,71158*	2,63752	,035	-10,9973	-,4259
	3	1	17,49002*	2,70979	,000	12,0595	22,9206
		2	5,71158*	2,63752	,035	,4259	10,9973
Tamhane	1	2	-11,77844*	3,15221	,002	-19,6737	-3,8832
		3	-17,49002*	2,29786	,000	-23,3790	-11,6010
	2	1	11,77844*	3,15221	,002	3,8832	19,6737
		3	-5,71158	2,56883	,102	-12,2771	,8539
	3	1	17,49002*	2,29786	,000	11,6010	23,3790
		2	5,71158	2,56883	,102	-,8539	12,2771

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

Από τον πίνακα Multiple Comparisons προκύπτει ότι η μέση επίδοση της ομάδας 1 διαφέρει στατιστικά σημαντικά από αυτή των 2 και 3 όντας χειρότερη.

Αναφορά: Θέλουμε να εξετάσουμε αν υπάρχει στατιστικά σημαντική διαφορά στη μέση επίδοση ως προς τις 3 διαφορετικές ομάδες πωλητών.

Για να μπορούμε να αποφανθούμε χρησιμοποιώντας μεθόδους της παραμετρικής στατιστικής θα πρέπει να πληρούνται οι ακόλουθες υποθέσεις:

1. Τα δείγματά μας να είναι τυχαία επιλεγμένα

2. Να μην υπάρχουν ακραίες τιμές στα δειγματικά δεδομένα κάθε πληθυσμού που να ξεπερνούν σε ποσοστό το 10%.
3. Κάθε πληθυσμός να περιγράφεται ικανοποιητικά από την κανονική κατανομή.

Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε τα δείγματά μας και ικανοποιείται.

Έλεγχος ακραίων τιμών

Ο έλεγχος των ακραίων τιμών στο δείγμα των τιμών που περιγράφουν την επίδοση της πρώτης ομάδας πωλητών έδειξε ότι οι παρατηρήσεις 18 και 20 με τιμές στην επίδοση 85.65 και 83.74 αντίστοιχα είναι ακραίες (20 διαθέσιμες, άρα ποσοστό ίσο του 10%) και αποκλείονται από την περαιτέρω ανάλυση. Ο έλεγχος των ακραίων τιμών στα δείγματα των τιμών που καταγράφεται η επίδοση της πρώτης καθώς και της δεύτερης ομάδας πωλητών έδειξε ότι δεν υπάρχουν ακραίες τιμές.

Έλεγχος κανονικότητας

Ο έλεγχος της υπόθεσης ότι τα δεδομένα που καταγράφεται η επίδοση της πρώτης ομάδας πωλητών ακολουθούν κανονική κατανομή δεν απορρίπτεται (τεστ Shapiro-Wilk p-τιμή=0.676)

Ο έλεγχος της υπόθεσης ότι τα δεδομένα που καταγράφεται η επίδοση της δεύτερης ομάδας πωλητών ακολουθούν κανονική κατανομή δεν απορρίπτεται (τεστ Shapiro-Wilk p-τιμή=0.344)

Ο έλεγχος της υπόθεσης ότι τα δεδομένα που καταγράφεται η επίδοση της τρίτης ομάδας πωλητών ακολουθούν κανονική κατανομή δεν απορρίπτεται (τεστ Shapiro-Wilk p-τιμή=0.582

Έλεγχος ίσων διακυμάνσεων

Η υπόθεση της ισότητας των πληθυσμιακών διακυμάνσεων της επίδοσης των 3 διαθέσιμων ομάδων πωλητών απορρίπτεται σε επίπεδο σημαντικότητας 5% (test Levene p-τιμή=0.015>0.05)

Έλεγχος ισότητας πληθυσμιακών μέσων τιμών-Πολλαπλές συγκρίσεις

Επομένως για να ελέγξουμε αν υπάρχει στατιστικά σημαντική διαφορά στη μέση πληθυσμιακή επίδοση των 3 ομάδων πωλητών θα χρησιμοποιήσουμε το τεστ του Welch με με επίπεδο σημαντικότητας 5%.

Υπάρχει στατιστικά σημαντική διαφορά στη μέση πληθυσμιακή επίδοση των 3 ομάδων πωλητών (τεστ Welch p-τιμή<0,001). Θέλοντας να εντοπίσουμε που υπάρχουν οι στατιστικά σημαντικές διαφορές ως προς τη μέση πληθυσμιακή επίδοση των 3 ομάδων πωλητών θα χρησιμοποιήσουμε τη μέθοδο των πολλαπλών συγκρίσεων του Tamhane με επίπεδο σημαντικότητας 5%.

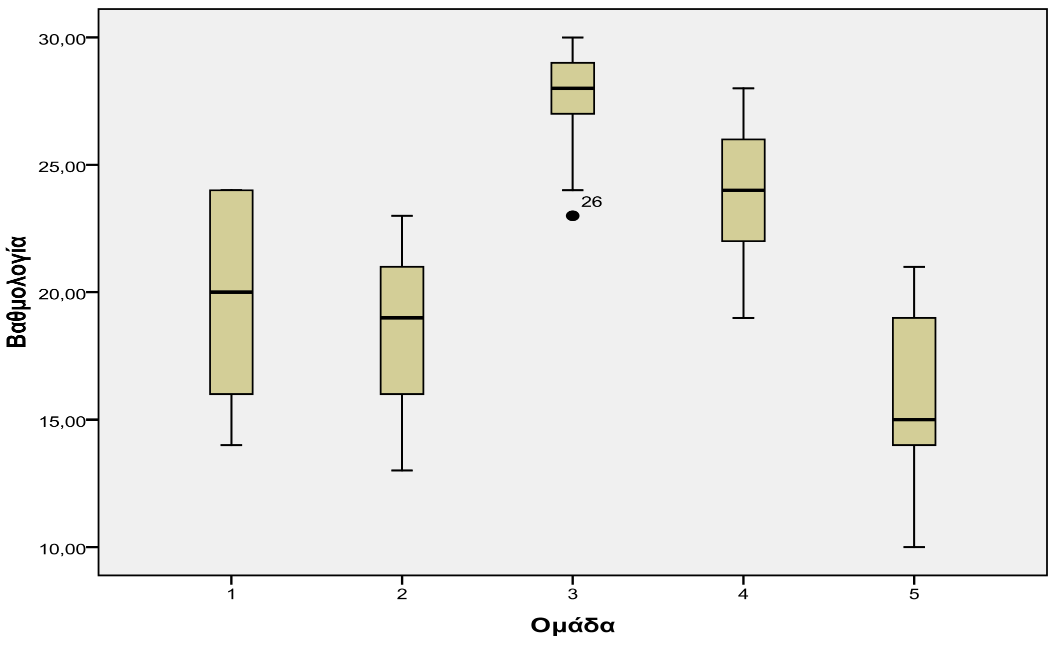
Η μέση επίδοση της πρώτης ομάδας πωλητών διαφέρει στατιστικά σημαντικά από τις αντίστοιχες της δεύτερης και τρίτης ομάδας (βλέπε πίνακα Multiple Comparisons., Tamhane, p=0.002 και p<0.001 αντίστοιχα) και μάλιστα είναι χειρότερη όπως προκύπτει από τον πίνακα που ακολουθεί ή από τον πίνακα των πολλαπλών συγκρίσεων.

	N	Mean
1	18	61,7892
2	20	73,5677
3	20	79,2792
Total	58	71,8818

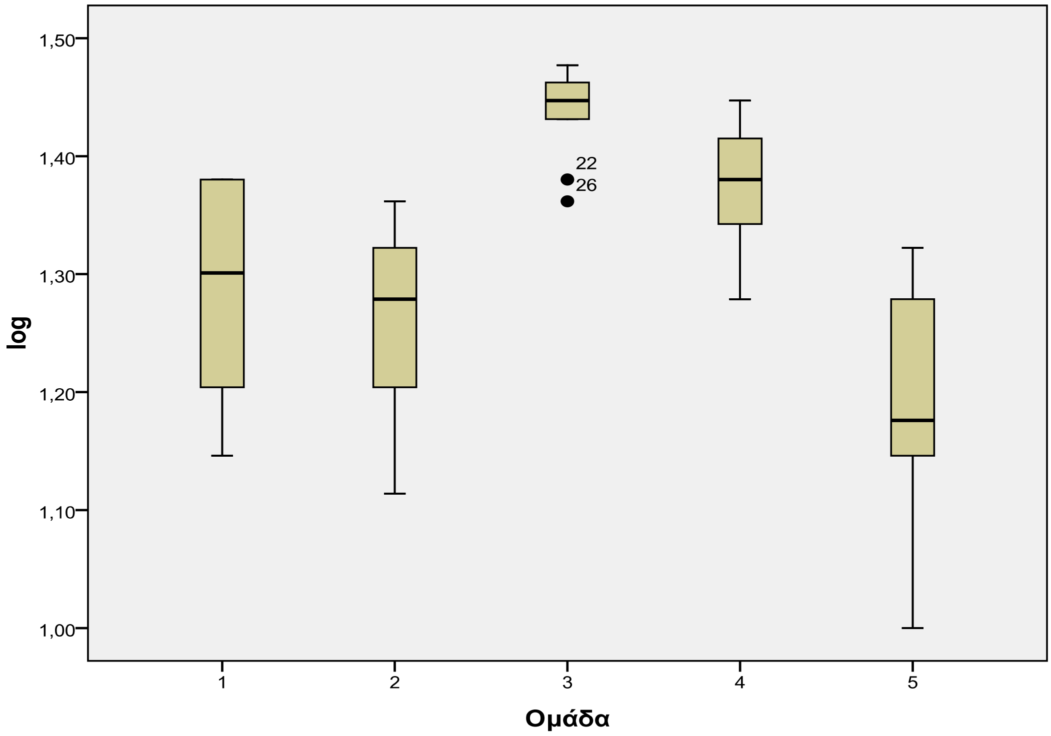
<u>Παράδειγμα 2^ο</u> Αρχείο School.sav[*]

Έχουμε 5 ομάδες μαθητών: οι δύο πρώτες διδάχθηκαν με την ίδια μέθοδο διδασκαλίας ενώ οι τρεις επόμενες με διαφορετική μέθοδο. **Θέλουμε να ελέγξουμε αν είναι εφικτό αν υπάρχει στατιστικά σημαντική διαφορά στη μέση βαθμολογία στις 5 ομάδες.**

Αποτελέσματα:

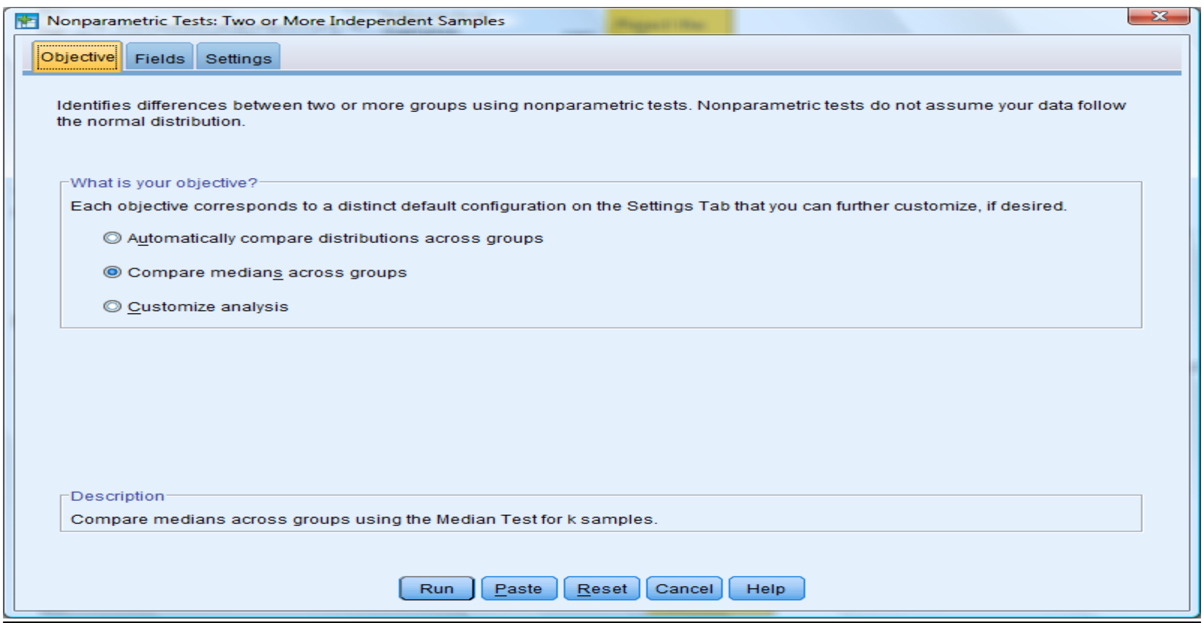


Τουλάχιστον 1 ακραία στις 9 διαθέσιμες δειγματικές της 3^{ης} ομάδας. Επομένως υπάρχει πρόβλημα ακραίων τιμών. Ο μετασχηματισμός του λογαρίθμου δε διορθώνει το πρόβλημα.



Άρα πρέπει να πάμε μη παραμετρικά και να ελέγξουμε αν υπάρχει διαφορά στη διάμεσο της επίδοσης των 5 πληθυσμών.

Θα χρησιμοποιηθεί και ένας εναλλακτικός τρόπος όπως φαίνεται παρακάτω μέσω της διαδικασίας Analyze Nonparametrics Independent Samples



Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Βαθμολογία is the same across categories of Ομάδα.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The medians of Βαθμολογία are the same across categories of Ομάδα.	Independent-Samples Median Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Ποια η καλύτερη ομάδα? Τα αποτελέσματα γενικεύονται?

Η αναφορά των αποτελεσμάτων αφήνεται ως άσκηση.

Άσκηση

Αρχείο Cartoon [*]

Θέλουμε να ελέγξουμε αν η μέση βαθμολογία στα cartoon διαφοροποιείται ανάλογα με το επίπεδο μόρφωσης.