

# Σύντομο εγχειρίδιο του SPSS 13.0

<b>1.0 ΤΟ ΠΕΡΙΒΑΛΛΟΝ ΤΟΥ SPSS</b>	<b>4</b>	
ΣΧΗΜΑ 1.1	Η ΕΙΣΑΓΩΓΙΚΗ ΟΘΟΝΗ ΤΟΥ SPSS	4
ΣΧΗΜΑ 1.2	Η ΑΡΧΙΚΗ ΟΘΟΝΗ ΤΟΥ SPSS	5
ΣΧΗΜΑ 1.3	ΤΟ ΜΕΝΟΥ ΕΠΙΛΟΓΩΝ ΤΟΥ [FILE]	7
<b>2.0 ΕΙΣΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ</b>	<b>8</b>	
ΣΧΗΜΑ 2.1	ΕΙΣΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ	8
ΣΧΗΜΑ 2.2	ΕΝΑΛΛΑΓΗ ΚΑΤΑΣΤΑΣΗΣ ΠΡΟΒΟΛΗΣ	9
ΣΧΗΜΑ 2.3	ΚΑΤΑΣΤΑΣΗ ΠΡΟΒΟΛΗΣ [VARIABLE VIEW]	9
ΠΙΝΑΚΑΣ 2.1	ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ SYSTEM MISSING ΠΕΡΙΠΤΩΣΕΩΝ ΣΤΟ OUTPUT	11
ΠΙΝΑΚΑΣ 2.2	ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ DISCRETE MISSING VALUES ΣΤΟ OUTPUT	11
ΣΧΗΜΑ 2.4	Η ΑΛΛΑΓΗ ΤΟΥ ΟΝΟΜΑΤΟΣ ΤΩΝ ΜΕΤΑΒΛΗΤΩΝ	12
ΣΧΗΜΑ 2.5	ΤΟ ΜΕΝΟΥ [FILE] ΓΙΑ ΤΗΝ ΑΠΟΘΗΚΕΥΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ	13
ΣΧΗΜΑ 2.6	ΤΟ ΜΕΝΟΥ ΕΠΙΛΟΓΩΝ ΤΥΠΟΥ ΜΟΡΦΟΠΟΙΗΣΗΣ ΤΟΥ ΑΡΧΕΙΟΥ ΔΕΔΟΜΕΝΩΝ	13
<b>3.0 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ</b>	<b>15</b>	
<b>3.1 ΚΑΤΑΝΟΜΗ ΣΥΧΝΟΤΗΤΑΣ</b>	<b>15</b>	
ΠΙΝΑΚΑΣ 3.1	ΚΑΤΑΝΟΜΗ ΣΥΧΝΟΤΗΤΑΣ	16
ΠΙΝΑΚΑΣ 3.2	ΣΧΕΤΙΚΗ ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ	18
<b>3.2 ΔΕΙΚΤΕΣ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ</b>	<b>19</b>	
<b>3.2 ΔΕΙΚΤΕΣ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ</b>	<b>19</b>	
ΣΧΗΜΑ 3.4	ΕΝΤΟΛΗ FREQUENCIES - ΠΡΟΕΙΔΟΠΟΙΗΤΙΚΟ ΜΗΝΥΜΑ	19
ΣΧΗΜΑ 3.5	ΕΝΤΟΛΗ FREQUENCIES – ΕΠΙΛΟΓΗ STATISTICS	20
ΠΙΝΑΚΑΣ 3.3	ΔΕΙΚΤΕΣ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ	20
<b>3.3 ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ</b>	<b>21</b>	
ΠΙΝΑΚΑΣ 3.4	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ	21
ΣΧΗΜΑ 3.6	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ – DEFINE SETS	22
ΣΧΗΜΑ 3.7	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ – ΟΡΙΣΜΟΣ SET	23
ΣΧΗΜΑ 3.8	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ – FREQUENCIES	23
ΠΙΝΑΚΑΣ 3.5	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ – ΑΠΑΝΤΗΣΕΩΝ	24
ΣΧΗΜΑ 3.9	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ - DEFINE RANGES	25
ΣΧΗΜΑ 3.10	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ - CROSSTABS ΜΕΤΑ ΤΟ DEFINE RANGES	26
ΣΧΗΜΑ 3.11	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ - CROSSTABS - OPTIONS	26
ΠΙΝΑΚΑΣ 3.11	ΑΝΑΛΥΣΗ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ - CROSSTABS - ΑΠΟΤΕΛΕΣΜΑΤΑ	27
<b>4.0 ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΜΕΤΑΒΛΗΤΩΝ</b>	<b>28</b>	

<b>4.1 COMPUTE</b>		<b>28</b>
ΣΧΗΜΑ 4.1	Η ΕΝΤΟΛΗ [COMPUTE]	28
<b>4.2 RECODE</b>		<b>29</b>
ΣΧΗΜΑ 4.2	Η ΠΡΩΤΗ ΟΘΟΝΗ ΤΗΣ ΕΝΤΟΛΗΣ [RECODE]	29
ΣΧΗΜΑ 4.3	Η ΔΕΥΤΕΡΗ ΟΘΟΝΗ ΤΗΣ ΕΝΤΟΛΗΣ [RECODE]	30
<b>4.3 WEIGHT CASES</b>		<b>31</b>
ΠΙΝΑΚΑΣ 4.1	ΔΕΥΤΕΡΟΓΕΝΗ ΔΕΔΟΜΕΝΑ	31
<b>5.0 ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ</b>		<b>34</b>
<b>5.1 ΤΕΣΤ Χ<sup>2</sup></b>		<b>35</b>
ΣΧΗΜΑ 5.1	Η ΕΝΤΟΛΗ CROSSTABS	35
ΣΧΗΜΑ 5.2	Η ΕΝΤΟΛΗ CROSSTABS – ΕΠΙΛΟΓΗ STATISTICS	36
ΣΧΗΜΑ 5.3	Η ΕΝΤΟΛΗ CROSSTABS – ΕΠΙΛΟΓΗ CELLS	36
ΠΙΝΑΚΑΣ 5.1	ΠΙΝΑΚΑΣ ΔΙΠΛΗΣ ΕΙΣΟΔΟΥ	37
ΠΙΝΑΚΑΣ 5.2	ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΤΕΣΤ Χ <sup>2</sup>	38
ΠΙΝΑΚΑΣ 5.3	ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΤΕΣΤ Χ <sup>2</sup> ΓΙΑ ΠΙΝΑΚΑ 2Χ2	39
<b>5.2 ΤΟ Τ-ΤΕΣΤ</b>		<b>40</b>
ΠΙΝΑΚΑΣ 5.4	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΟ Τ-ΤΕΣΤ ΓΙΑ ΕΞΑΡΤΗΜΕΝΑ ΔΕΙΓΜΑΤΑ	40
ΣΧΗΜΑ 5.4	ΤΟ Τ-ΤΕΣΤ ΓΙΑ ΕΞΑΡΤΗΜΕΝΑ ΔΕΙΓΜΑΤΑ	41
ΠΙΝΑΚΑΣ 5.5	ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ Τ-ΤΕΣΤ ΓΙΑ ΕΞΑΡΤΗΜΕΝΑ ΔΕΙΓΜΑΤΑ	41
ΠΙΝΑΚΑΣ 5.6	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΟ Τ-ΤΕΣΤ ΑΝΕΞΑΡΤΗΤΩΝ ΔΕΙΓΜΑΤΩΝ	42
ΣΧΗΜΑ 5.5	ΤΟ Τ-ΤΕΣΤ ΓΙΑ ΑΝΕΞΑΡΤΗΤΑ ΔΕΙΓΜΑΤΑ	43
ΣΧΗΜΑ 5.6	ΚΑΘΟΡΙΣΜΟΣ ΤΩΝ ΟΜΑΔΩΝ	43
ΣΧΗΜΑ 5.7	ΤΟ Τ-ΤΕΣΤ ΓΙΑ ΑΝΕΞΑΡΤΗΤΑ ΔΕΙΓΜΑΤΑ ΜΕΤΑ ΤΟΝ ΚΑΘΟΡΙΣΜΟ ΤΩΝ ΟΜΑΔΩΝ	44
ΠΙΝΑΚΑΣ 5.7	ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ Τ-ΤΕΣΤ ΑΝΕΞΑΡΤΗΤΩΝ ΔΕΙΓΜΑΤΩΝ	44
<b>5.3 ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ</b>		<b>46</b>
ΠΙΝΑΚΑΣ 5.8	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ	46
ΣΧΗΜΑ 5.8	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ	47
ΣΧΗΜΑ 5.9	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ – ΕΠΙΛΟΓΗ OPTIONS	48
ΠΙΝΑΚΑΣ 5.9	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ	48
ΣΧΗΜΑ 5.10	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΜΟΝΗΣ ΚΑΤΕΥΘΥΝΣΗΣ – ΕΠΙΛΟΓΗ POST HOC	50
ΠΙΝΑΚΑΣ 5.10	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΙΣ ΕΚ ΤΩΝ ΥΣΤΕΡΩΝ ΚΑΤΑ ΖΕΥΓΗ ΣΥΓΚΡΙΣΕΙΣ	51
<b>5.4 ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ</b>		<b>52</b>

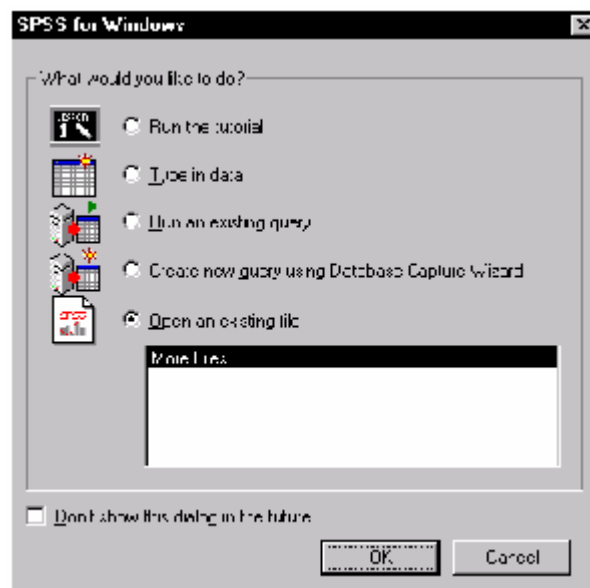
ΠΙΝΑΚΑΣ 5.10	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ	52
ΣΧΗΜΑ 5.10	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ – ΑΡΧΙΚΗ ΟΘΟΝΗ	53
ΣΧΗΜΑ 5.10	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ – ΚΑΘΟΡΙΣΜΟΣ ΕΠΙΠΕΔΩΝ	54
ΣΧΗΜΑ 5.12	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ	55
ΣΧΗΜΑ 5.12	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ – ΕΠΙΛΟΓΗ ΟΡΤΙΩΝ	56
ΠΙΝΑΚΑΣ 5.11	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ	56
<b>5.5 WILCOXON</b>		<b>59</b>
ΣΧΗΜΑ 5.13	ΤΕΣΤ WILCOXON	59
ΠΙΝΑΚΑΣ 5.12	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΟ ΤΕΣΤ WILCOXON	60
<b>5.6 MANN-WHITNEY U</b>		<b>61</b>
ΣΧΗΜΑ 5.14	ΤΕΣΤ MANN-WHITNEY U	61
ΠΙΝΑΚΑΣ 5.13	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΟ ΤΕΣΤ MANN-WHITNEY U	62
ΣΧΗΜΑ 5.15	H ΕΝΤΟΛΗ MEANS	63
<b>5.7 KRUSKAL-WALLIS H</b>		<b>65</b>
ΣΧΗΜΑ 5.16	KRUSKAL-WALLIS H	65
ΣΧΗΜΑ 5.17	KRUSKAL-WALLIS H - ΚΑΘΟΡΙΣΜΟΣ ΤΩΝ ΟΜΑΔΩΝ	66
ΣΧΗΜΑ 5.18	KRUSKAL-WALLIS H ΜΕΤΑ ΤΟΝ ΚΑΘΟΡΙΣΜΟ ΤΩΝ ΟΜΑΔΩΝ	66
<b>5.9 FRIEDMAN</b>		<b>68</b>
ΣΧΗΜΑ 5.19	ΤΕΣΤ FRIEDMAN	68
ΠΙΝΑΚΑΣ 5.16	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΟ ΤΕΣΤ FRIEDMAN	69
<b>5.10 ΣΥΝΑΦΕΙΑ PEARSON R ΚΑΙ SPEARMAN RHO</b>		<b>70</b>
ΠΙΝΑΚΑΣ 5.17	ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΤΟΝ ΥΠΟΛΟΓΙΣΜΟ ΣΥΝΑΦΕΙΩΝ	70
ΣΧΗΜΑ 5.20	ΣΥΝΑΦΕΙΕΣ – PEARSON R	70
ΠΙΝΑΚΑΣ 5.18	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΟ ΔΕΙΚΤΗ ΣΥΝΑΦΕΙΑΣ PEARSON R	71
ΣΧΗΜΑ 5.20	ΣΥΝΑΦΕΙΕΣ – PEARSON R	72
ΠΙΝΑΚΑΣ 5.19	ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΟ ΔΕΙΚΤΗ ΣΥΝΑΦΕΙΑΣ SPEARMAN RHO	73

## 1.0 Το περιβάλλον του SPSS

Το SPSS έχει την ίδια γενική μορφή που έχουν σχεδόν όλες οι εφαρμογές που «τρέχουν» σε περιβάλλον Windows. Ο χρήστης μπορεί να εκτελέσει σχεδόν οποιαδήποτε στατιστική ανάλυση και επεξεργασία των δεδομένων του σε ένα χρηστικό παραθυρικό περιβάλλον. Ωστόσο, είναι δυνατή η στατιστική επεξεργασία των δεδομένων και μέσα από ένα προγραμματιστικό περιβάλλον, στο οποίο όμως δεν θα γίνει αναφορά σε αυτό το σύντομο οδηγό.

Για να ξεκινήσετε την εφαρμογή, πηγαίνετε στο μενού [**Εναρξη => Προγράμματα => SPSS for Windows**] και επιλέξτε το εικονίδιο του **SPSS 13.0 for Windows**. Η εφαρμογή θα ξεκινήσει σύντομα και θα εμφανιστεί το παρακάτω πλαίσιο διαλόγου (βλ. Σχήμα 1.1):

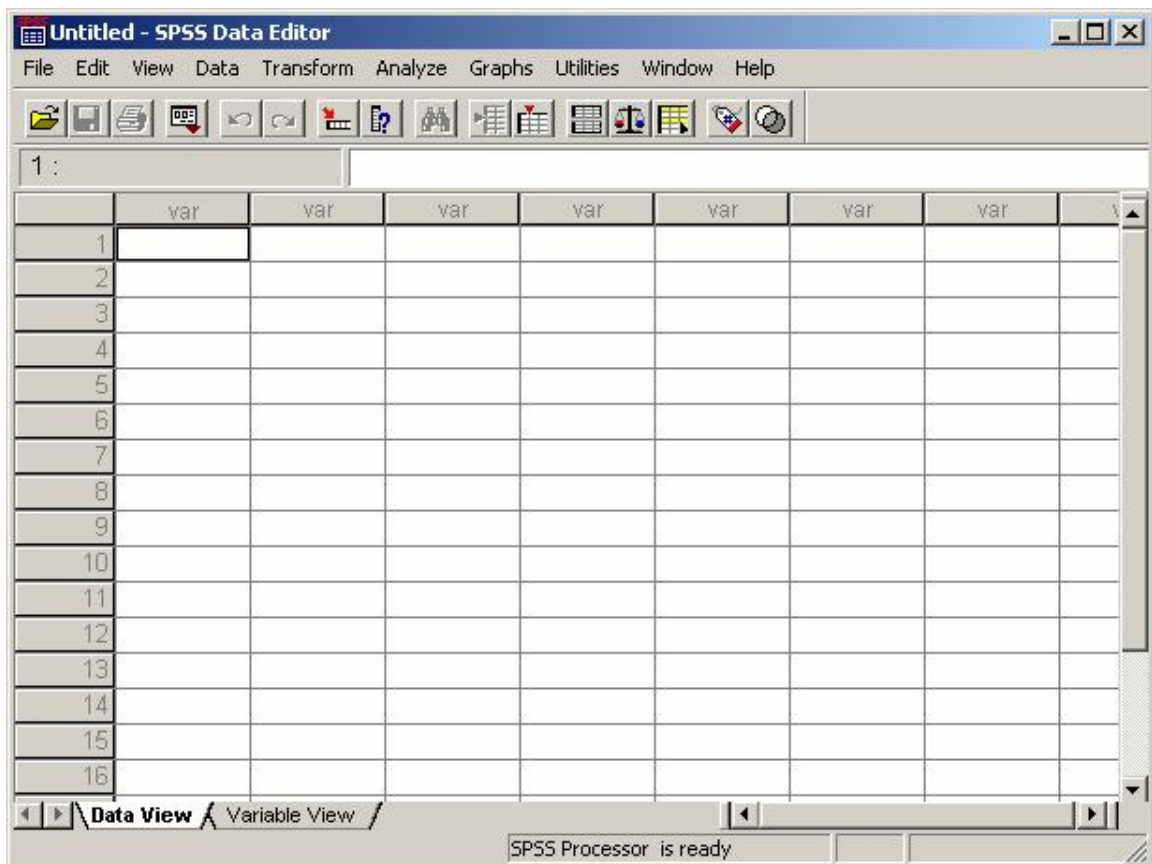
Σχήμα 1.1 Η εισαγωγική οθόνη του SPSS



Όπως βλέπετε, έχετε πέντε διαφορετικές επιλογές: να παρακολουθήσετε το πρόγραμμα εκμάθησης της εφαρμογής [**Run the tutorial**], να εισάγετε δεδομένα πληκτρολογώντας τα [**Type in data**], να εκτελέσετε ένα ερώτημα [**Run an existing query**], να δημιουργήσετε ένα ερώτημα χρησιμοποιώντας τον Database Capture Wizard [**Create new query using Database Capture Wizard**], ή να ανοίξετε ένα αρχείο [**Open an existing file**].

Επιλέξτε το **[Type in data]** και κάντε κλικ στο κουμπί **[OK]**. Η οθόνη που εμφανίζεται (Σχήμα 1.2) είναι παρόμοια με αυτήν πολλών εφαρμογών λογιστικού φύλλου. Αν επιλέξετε την επιλογή **[Don't show this dialog in the future]** στην οθόνη του Σχήματος 1.1 ανοίγοντας το SPSS θα ξεκινάτε από αυτήν την οθόνη.

**Σχήμα 1.2** Η αρχική οθόνη του SPSS



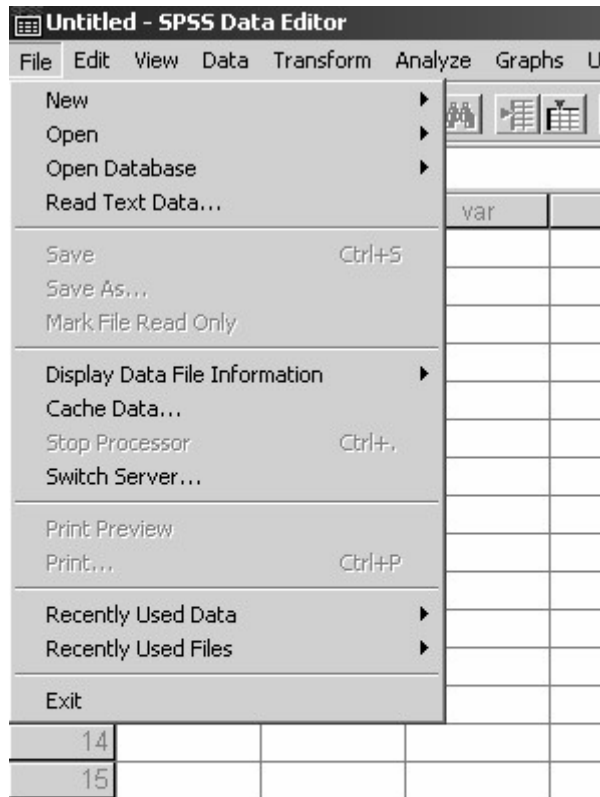
Όπως βλέπετε, στη γραμμή μενού υπάρχουν μια σειρά από επιλογές σχετικές με τις διάφορες λειτουργίες του προγράμματος. Υπάρχουν επίσης εικονίδια συντόμευσης για άμεση πρόσβαση σε λειτουργίες που γίνονται συχνά. Τοποθετώντας το δείκτη του ποντικιού πάνω από αυτά για λίγο εμφανίζεται μια σύντομη επεξήγηση της λειτουργίας τους. Το Σχήμα 1.2 παρουσιάζει ένα άδειο φύλλο δεδομένων (data sheet). Τα δεδομένα μπορούν να εισαχθούν είτε πληκτρολογώντας τα είτε εισάγοντάς τα από κάποιο αρχείο.

Ρίχνοντας μια πρόχειρη ματιά στο μενού **[File]** διαπιστώνουμε ότι πολλές από τις επιλογές του μας είναι ήδη γνωστές από άλλες παραθυρικές εφαρμογές (Σχήμα 1.3). Για

παράδειγμα, η επιλογή **[New]** χρησιμοποιείται για να δηλωθεί ο τύπος του παραθύρου που θα ανοίξει. Οι διάφορες εναλλακτικές επιλογές στο **[New]** είναι:

- **[Data]** Προεπιλεγμένο παράθυρο με ένα άδειο φύλλο δεδομένων έτοιμο για αναλύσεις.
- **[Syntax]** Εκεί κάποιος μπορεί να γράψει κώδικα αντί να χρησιμοποιήσει τα μενού επιλογών. Η χρήση των σχετικών εντολών περιγράφεται στα εγχειρίδια του SPSS.
- **[Output]** Κάθε φορά που «τρέχουμε» μια εντολή, το αποτέλεσμα της κατευθύνεται σε ένα ξεχωριστό παράθυρο εξόδου. Μπορούμε μάλιστα να έχουμε πολλαπλά παράθυρα εξόδου **[Output]** ανοιχτά προκειμένου να οργανώνουμε τα αποτελέσματα των διαφόρων αναλύσεων που εκτελούμε. Αργότερα αυτά τα αποτελέσματα μπορεί να αποθηκευτούν ή/και να εκτυπωθούν.
- **[Draft Output]** Παρόμοιο με το **[Output]** αλλά χρησιμοποιείται μία παλαιότερη μορφή παρουσίασης των αποτελεσμάτων.
- **[Script]** Το παράθυρο αυτό δίνει στο χρήστη τη δυνατότητα να συντάξει πλήρη προγράμματα σε μια γλώσσα που μοιάζει αρκετά με τη *BASIC*. Τα προγράμματα αυτά έχουν πρόσβαση στις λειτουργίες του SPSS. Με αυτή την πρόσβαση είναι δυνατόν κάποιος να δημιουργήσει δικές του διαδικασίες – που δεν αποτελούν μέρος του SPSS – αξιοποιώντας τις λειτουργίες του SPSS. Και πάλι, αυτή η δυνατότητα ξεφεύγει από το αντικείμενο αυτών των σημειώσεων.

Σχήμα 1.3 Το μενού επιλογών του [File]



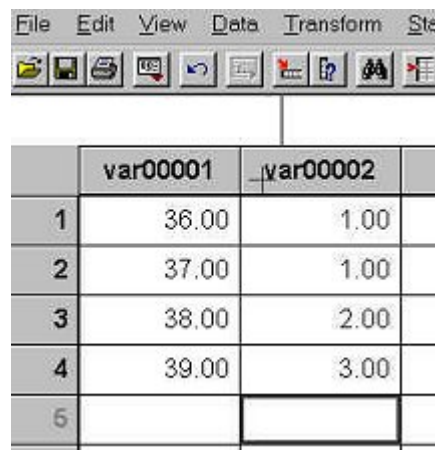
Επίσης, στο μενού **[File]** υπάρχουν δύο ξεχωριστές διέξοδοι για την ανάγνωση δεδομένων από υπάρχοντα αρχεία (Σχήμα 1.3). Η πρώτη είναι η επιλογή **[Open]**. Όπως και σε πολλές άλλες εφαρμογές (π.χ., Word, Excel, ...) το SPSS έχει τη δική του μορφοποίηση για τα δεδομένα που αποθηκεύονται μέσα από αυτό. Στην περίπτωση αυτή, η αποδεκτή προέκταση ονόματος αρχείου είναι το "sav". Έτσι, μπορούμε για παράδειγμα να έχουμε αποθηκευμένο ένα αρχείο με το όνομα "data1.sav". Ωστόσο, αυτή η μορφοποίηση δεν είναι αναγνώσιμη από άλλα προγράμματα, γιατί είναι σε δυαδική μορφή (binary format). Το πλεονέκτημα είναι ότι όλες οι αλλαγές στη μορφοποίηση διατηρούνται και το αρχείο μπορεί να αναγνωστεί ταχύτερα, μέσα από την επιλογή **[Open]**, η οποία προορίζεται για αρχεία αποθηκευμένα σε μορφοποίηση SPSS. Η δεύτερη επιλογή, **[Read Text Data]**, όπως υποδηλώνει και το όνομά της προορίζεται για την ανάγνωση αρχείων που έχουν αποθηκευτεί σε μορφοποίηση ASCII.

## 2.0 Εισαγωγή Δεδομένων

Για να εισάγετε δεδομένα στην εφαρμογή, απλώς κάντε κλικ με το αριστερό πλήκτρο του ποντικιού στο πάνω αριστερά κελί του φύλλου δεδομένων. Αρχίστε να πληκτρολογείτε τα δεδομένα όπως βλέπετε στο Σχήμα 2.1. Συνήθως τα δεδομένα εισάγονται οριζοντίως. Οι γραμμές του πίνακα αντιστοιχούν σε περιπτώσεις (συμμετέχοντες) και οι στήλες σε μεταβλητές. Έτσι, η συνήθης λογική είναι να εισάγουμε τα δεδομένα ενός συμμετέχοντα για όλες τις μεταβλητές.

Στο Σχήμα 2.1 η πρώτη στήλη αναπαριστά «Χρόνο αντίδρασης σε 100στά του δευτερολέπτου» και η δεύτερη δείχνει «Συχνότητα». Όπως είπαμε, οι στήλες αντιπροσωπεύουν τις μεταβλητές και οι σειρές τις περιπτώσεις.

**Σχήμα 2.1** Εισαγωγή δεδομένων

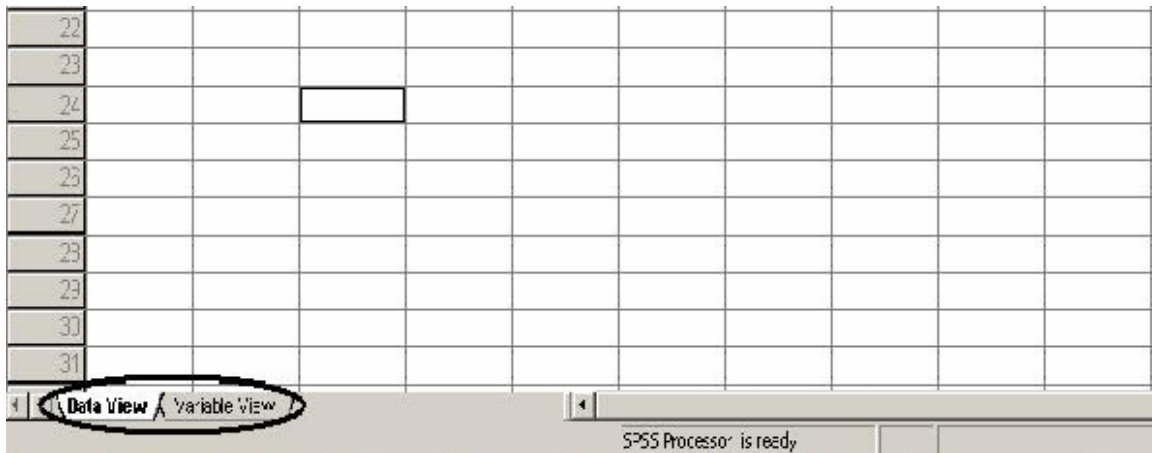


	var00001	_var00002
1	36.00	1.00
2	37.00	1.00
3	38.00	2.00
4	39.00	3.00
5		
6		

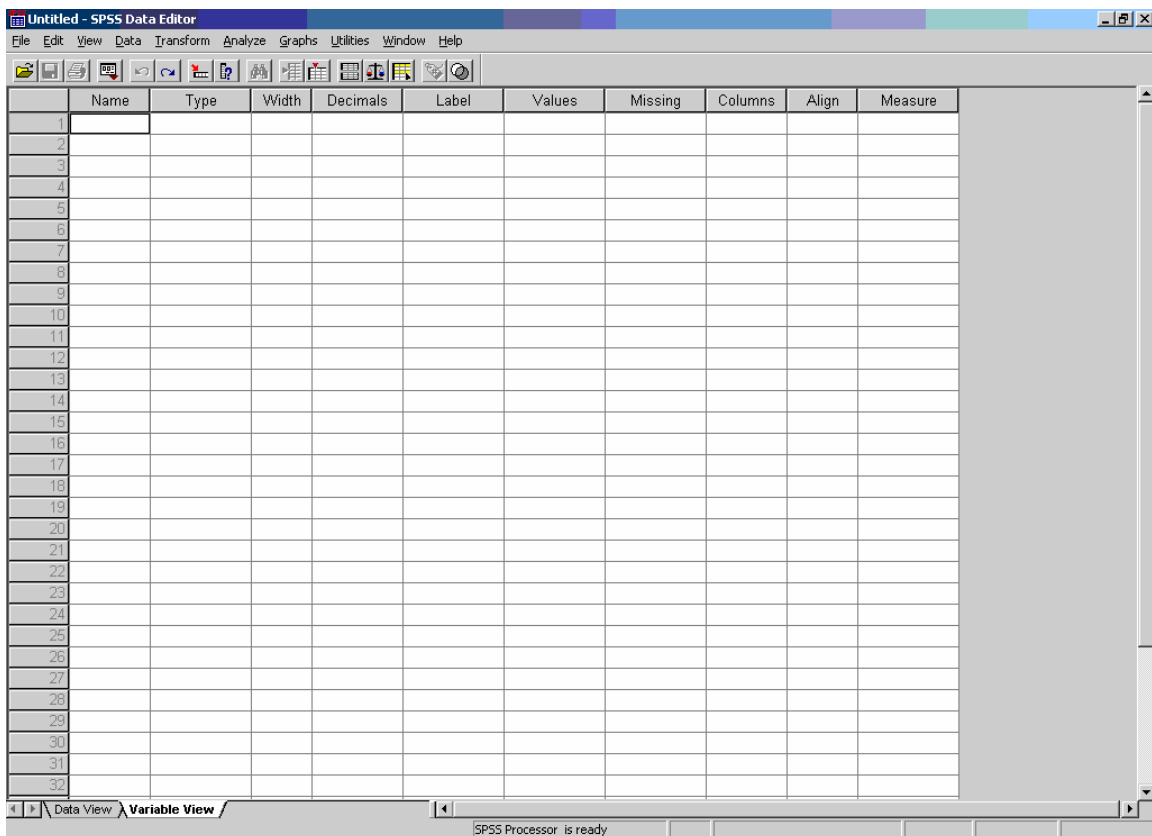
Αν πληκτρολογείτε δεδομένα για πρώτη φορά, όπως στο παράδειγμα αυτό, τότε τα ονόματα των μεταβλητών θα δημιουργηθούν αυτόματα (π.χ., var00001, var00002,...). Προφανώς, δεν είναι πολύ βοηθητικά τα ονόματα αυτά. Προκειμένου να αλλάξετε τα ονόματα αυτά, θα πρέπει να μεταβούμε από την κατάσταση προβολής **[Data View]** στην κατάσταση προβολής **[Variable View]** (βλ. Σχήμα 2.2).



Σχήμα 2.2 Εναλλαγή κατάστασης προβολής



Σχήμα 2.3 Κατάσταση προβολής [Variable View]



Στο Σχήμα 2.3 παρουσιάζεται η κατάσταση προβολής **[Variable View]**. Εδώ έχουμε αντιστροφή της παρουσίασης των μεταβλητών, οι οποίες πλέον εμφανίζονται στις σειρές του πίνακα, ενώ στις διάφορες στήλες μπορούμε να καθορίσουμε διάφορα χαρακτηριστικά για κάθε μεταβλητή.

Από αριστερά προς τα δεξιά για κάθε μεταβλητή μπορούμε να καθορίσουμε τα ακόλουθα:

- **[Name]** Εδώ μπορούμε να καθορίσουμε το όνομα της μεταβλητής. Οι νεώτερες εκδόσεις του SPSS δέχονται ελληνικούς χαρακτήρες, αλλά προτείνεται η χρήση ονομάτων μεταβλητών μέχρι 8 λατινικούς χαρακτήρες χωρίς κενά, προκειμένου να διατηρηθεί η συμβατότητα του αρχείου μας με παλαιότερες εκδόσεις του SPSS. Αν θέλουμε να δώσουμε ένα αναλυτικό όνομα με ελληνικούς χαρακτήρες μπορούμε να χρησιμοποιήσουμε την επιλογή **[Label]** (βλ. παρακάτω).
- **[Type]** Εδώ επιλέγουμε το είδος των δεδομένων για κάθε μεταβλητή. Εξορισμού οι νέες μεταβλητές έχουν ως type το **[numeric]**, δηλαδή θεωρείται ότι το είδος των δεδομένων είναι αριθμοί. Ενδιαφέρον παρουσιάζει και η επιλογή **[string]** που σημαίνει ότι η μεταβλητή θα δεχθεί αλφαριθμητικά δεδομένα (δηλαδή οποιοδήποτε συνδυασμό γραμμάτων και αριθμών), τα οποία δεν μπορούν να χρησιμοποιηθούν σε κάποια είδη αναλύσεων. Οι **[dot]**, **[comma]** και **[scientific notation]** αφορούν στον τρόπο παρουσίασης των αριθμών στα κελιά, ενώ οι υπόλοιπες επιλογές δεν απαιτούν κάποια ιδιαίτερη εξήγηση.
- **[Width]** Εδώ επιλέγουμε τον μέγιστο αριθμό χαρακτήρων που μπορούμε να καταχωρίσουμε για κάθε περίπτωση στην μεταβλητή μας.
- **[Decimals]** Εδώ επιλέγουμε το μέγιστο αριθμό δεκαδικών που μπορούμε να καταχωρήσουμε για κάθε περίπτωση στην μεταβλητή μας.
- **[Label]** Εδώ μπορούμε να δώσουμε ένα «δεύτερο όνομα» στη μεταβλητή με ελληνικούς χαρακτήρες, κενά ή ό,τι άλλο θέλουμε με μήκος έως 256 χαρακτήρες. Είναι σημαντικό να προσέξουμε την ορθογραφία των περιγραφών μας, πεζά – κεφαλαία γράμματα κλπ., ώστε να βελτιώσουμε την αναγνωσιμότητα των αποτελεσμάτων, καθώς το **[Label]** (για όσες μεταβλητές έχει οριστεί) χρησιμοποιείται στους πίνακες που θα κατασκευάσει το SPSS.
- **[Values]** Συχνά, απαιτείται (για λόγους ευκολίας) να κωδικοποιήσουμε τις κατηγορίες των μεταβλητών σε αριθμητική μορφή. Για παράδειγμα, το “άνδρας” και “γυναίκα” μπορεί να κωδικοποιηθεί ως “1” και “2” αντίστοιχα. Προκειμένου να αποφύγουμε τη σύγχυση, συνιστάται να χρησιμοποιούμε ετικέτες για τις διάφορες τιμές. Για παράδειγμα, η τιμή **Value:1** θα αντιστοιχούσε στην

ετικέτα/**Value label**: Άντρας. Παρόμοια, το **Value:2** θα μπορούσε να κωδικοποιηθεί με το **Value Label**: Γυναίκα.

- **[Missing]** Εδώ μπορούμε να ορίσουμε συγκεκριμένες τιμές που το SPSS θα εκλαμβάνει ως ελλείψεις κατά την ανάλυση των δεδομένων. Αν για μία περίπτωση αφήσουμε κενό το κελί που αντιστοιχεί στη στήλη μίας μεταβλητής το SPSS αυτόματα θα θεωρήσει ότι η συγκεκριμένη περίπτωση είναι ελλιπής ως προς τη συγκεκριμένη μεταβλητή. Οι περιπτώσεις αυτές εμφανίζονται στο *output* ως system missing (βλ. Πίνακα 2.1)

**Πίνακας 2.1 Παρουσίαση των system missing περιπτώσεων στο output**

**Φύλο**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Άντρας	636	41,9	41,9	41,9
	Γυναίκα	881	58,0	58,1	100,0
	Total	1517	99,9	100,0	
Missing	System	2	,1		
	Total	1519	100,0		

- Αντίθετα, στην περίπτωση που ορίσουμε συγκεκριμένες τιμές στο πεδίο missing, τότε το SPSS τις εμφανίζει ως απλώς missing και αντί της ένδειξης system αναφέρει την τιμή που ορίσαμε στο πεδίο **[discrete missing values]** ή το **[label]** που έχουμε ορίσει για τη συγκεκριμένη τιμή. Στο παράδειγμα του Πίνακα 2.2 η τιμή που ορίσαμε είναι η 8, δηλαδή είπαμε στο SPSS να θεωρήσει κάθε καταχώρηση στη συγκεκριμένη μεταβλητή που έχει τιμή 8 ως ελλιπή περίπτωση.

**Πίνακας 2.2 Παρουσίαση των discrete missing values στο output**

**Φύλο**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Άντρας	636	41,9	41,9	41,9
	Γυναίκα	881	58,0	58,1	100,0
	Total	1517	99,9	100,0	
Missing	8	2	,1		
Total		1519	100,0		

- **[Columns]** Εδώ επιλέγουμε το πλάτος της στήλης της μεταβλητής, όπως θα εμφανίζεται στην κατάσταση προβολής **[Data view]**. Δεν επηρεάζεται ο πραγματικός αριθμός χαρακτήρων που μπορεί να δεχθεί η μεταβλητή (βλ. επιλογή **[Width]**).
- **[Align]** Εδώ επιλέγουμε τον τρόπο που θα στοιχίζονται τα δεδομένα στα κελιά.
- **[Measure]** Εδώ επιλέγουμε την κλίμακα μέτρησης της μεταβλητής.

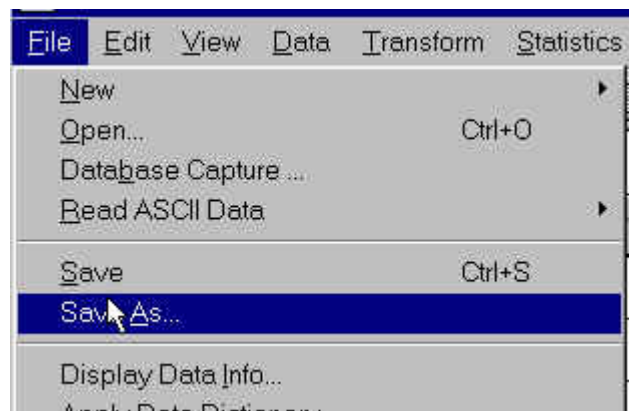
Στο παράδειγμα του Σχήματος 2.1, μπορούμε να αντικαταστήσουμε το "var00001" με "RT" (Reaction Time – Χρόνος Αντίδρασης) και το "var00002" με "freq" (frequency), χρησιμοποιώντας την επιλογή **[Name]** (βλ. Σχήμα 2.3). Η εικόνα που θα έχουμε στην κατάσταση προβολής **[Data View]** είναι η ακόλουθη (βλ. Σχήμα 2.4). Αν ορίσουμε και κάποιο **[Label]** για τις μεταβλητές μας, μπορούμε να το δούμε στην κατάσταση προβολής **[Data View]** αφήνοντας το δείκτη του ποντικιού μας χωρίς να τον κουνάμε πάνω στο όνομα μίας μεταβλητής.

**Σχήμα 2.4** Η αλλαγή του ονόματος των μεταβλητών

freq			
	rt	freq	var
1	36.00	1.00	
2	37.00	1.00	
3	38.00	2.00	
4	40.00	3.00	
5			

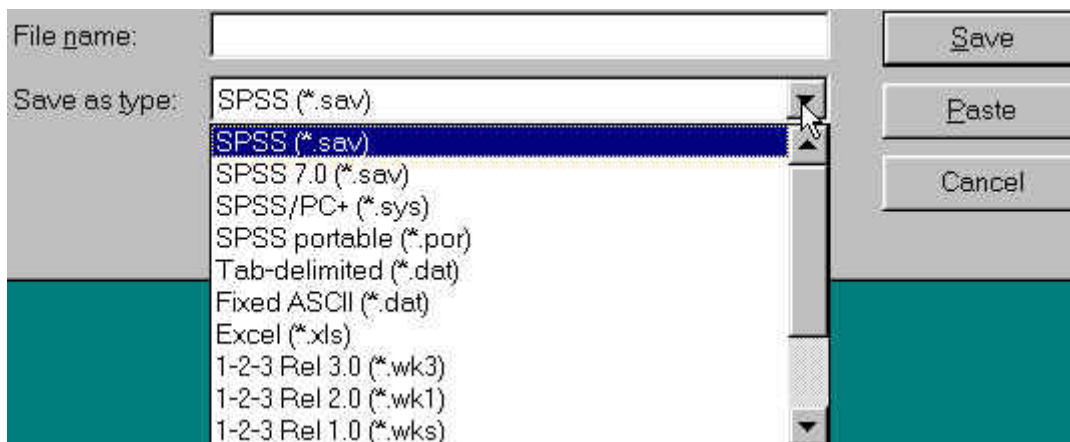
Αφού έχουμε πληκτρολογήσει τα δεδομένα μας πρέπει να αποθηκεύσουμε το αρχείο. Στην πραγματικότητα απαιτείται να αποθηκεύουμε πολύ συχνά κατά τη διάρκεια της πληκτρολόγησης των δεδομένων, προκειμένου να αποφύγουμε την απώλειά τους λόγω κάποιου «ατυχήματος» (βλ. Σχήμα 2.5).

**Σχήμα 2.5** Το μενού [File] για την αποθήκευση των δεδομένων



Το SPSS προσφέρει μια ευρεία γκάμα επιλογών από μορφοποιήσεις για την αποθήκευση των δεδομένων (βλ. Σχήμα 2.6). Ο κατάλογος των διαθέσιμων μορφοποιήσεων είναι προσβάσιμος μέσω του **Save as type:** στο πλαίσιο διαλόγου **SaveAs**.

**Σχήμα 2.6** Το μενού επιλογών τύπου μορφοποίησης του αρχείου δεδομένων



Αν σκοπεύετε να εργαστείτε μόνο στο SPSS, τότε είναι προτιμότερο να αποθηκεύσετε με τη μορφοποίηση *SPSS (\*.sav)*. Αυτή η μορφοποίηση επιτρέπει ταχύτερη ανάγνωση

και επεξεργασία του αρχείου των δεδομένων. Ωστόσο, αν πρόκειται να αναλύσετε ή να χρησιμοποιήσετε τα δεδομένα σας και μέσα από κάποια άλλη εφαρμογή (π.χ., ένα λογιστικό φύλλο), θα σας συμβούλευα να τα αποθηκεύσετε σε ένα πιο διαδεδομένο τύπο μορφοποίησης [π.χ., Excel(\*.xls), 1-2-3 Rel 3.0 (\*.wk3)].

Αφού επιλέξετε τύπο αρχείου, πληκτρολογήστε το όνομα του αρχείου, χωρίς την προέκταση (π.χ., sav, xls). Μπορείτε επίσης να αποθηκεύσετε το αρχείο σας σε ένα σχετικό κατάλογο στο σκληρό δίσκο ή τη δισκέτα σας.

*Προσοχή! Δεν είναι σπάνιο να ψάχνουμε ένα αρχείο μας για ώρες επειδή το αποθηκεύσαμε σε έναν άσχετο κατάλογο.*

### 3.0 Περιγραφική Στατιστική

Ας υποθέσουμε ότι έχετε περάσει ήδη τα δεδομένα σας στην εφαρμογή και έχετε δημιουργήσει ένα αρχείο \*.sav. Το επόμενο βήμα είναι η εξαγωγή ορισμένων περιγραφικών στατιστικών δεικτών.

Για να γνωρίζουμε πώς πρέπει να περιγράψουμε μία μεταβλητή, πρέπει να γνωρίζουμε την κλίμακα μέτρησής της. Συνοπτικά ισχύουν τα ακόλουθα:

Μεταβλητή (Κλίμακα μέτρησης)	Περιγραφική στατιστική
Κατηγορική	Κατανομή συχνότητας
Τακτική	Δείκτες κεντρικής τάσης και διασποράς (Διάμεσος, εύρος)
Αριθμητική	Δείκτες κεντρικής τάσης και διασποράς (Μέσος όρος, τυπική απόκλιση)

Κάθε φορά που «τρέχουμε» μια στατιστική επεξεργασία, τα αποτελέσματα στέλνονται στην οθόνη εξόδου (Output). Αν τρέξουμε αρκετές επεξεργασίες κατά τη διάρκεια μιας εργασίας μας, τα αποτελέσματα θα συσσωρευτούν με τη σειρά στο παράθυρο αυτό. Ωστόσο, μπορούμε να οργανώσουμε καλύτερα τα αποτελέσματά μας ανοίγοντας ένα καινούριο παράθυρο πριν να τρέξουμε κάθε επεξεργασία - **[File=>New=>Output]**. Επιπλέον, τα περιεχόμενα αυτών των παραθύρων μπορούν να αποθηκευτούν για μελλοντική χρήση ή για αντιγραφή σε κάποιο έγγραφο του Word.

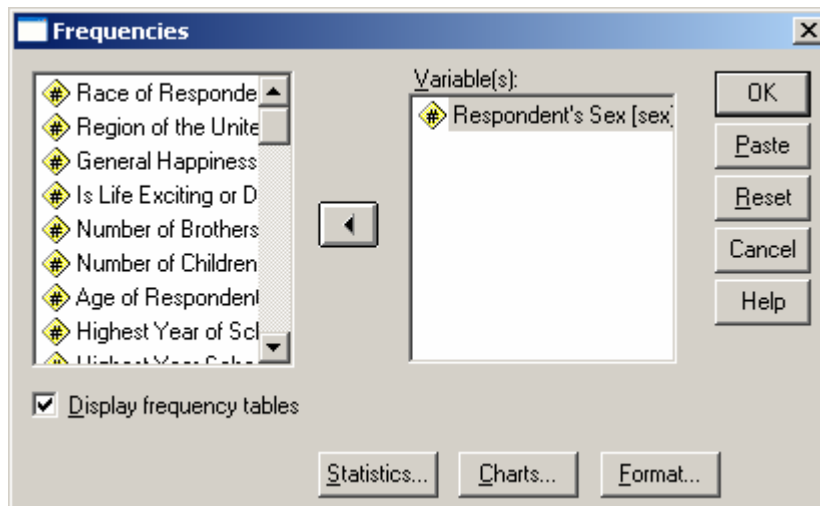
#### 3.1 Κατανομή συχνότητας

Για να ξεκινήσετε, κάντε κλικ στα **[Analyze=>Descriptive Statistics=>Frequencies]**.

Το αποτέλεσμα είναι ένα καινούριο παράθυρο διαλόγου που επιτρέπει στο χρήστη να επιλέξει τις μεταβλητές που τον ενδιαφέρουν. Προσέξτε επίσης και τα υπόλοιπα ενεργά κουμπιά στα άκρα του παραθύρου. Τα κουμπιά με τον τίτλο **[Statistics...]** και **[Charts...]** είναι τα πιο σημαντικά. Στο Σχήμα 3.1 που ακολουθεί, ο χρήστης ενδιαφέρεται για τη

μεταβλητή **sex**, την επιλέγει και κάνει κλικ στο βέλος που δείχνει δεξιά. Το αποτέλεσμα αυτής της ενέργειας είναι η μεταφορά της μεταβλητής στη λίστα των μεταβλητών **Variable(s)**:. Στο σημείο αυτό, κάνοντας κλικ στο κουμπί **[OK]**, θα ανοίξει ένα παράθυρο output με τις πληροφορίες για τις συχνότητες καθεμιάς από τις τιμές της μεταβλητής.

### Σχήμα 3.1 Η εντολή Frequencies



Τα αποτελέσματα της εντολής Frequencies παρουσιάζονται στον Πίνακα 3.1:

### Πίνακας 3.1 Κατανομή συχνότητας

#### Φύλο

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Άντρας	636	41,9	41,9	41,9
	Γυναίκα	881	58,0	58,1	100,0
	Total	1517	99,9	100,0	
Missing	System	2	,1		
Total		1519	100,0		

Η στήλη **Frequency** μας δείχνει την απόλυτη συχνότητα των τιμών της μεταβλητής στο δείγμα μας προσέφεραν μία συγκεκριμένη τιμή. Για παράδειγμα, στον Πίνακα 3.1 φαίνεται ότι 636 από τους συμμετέχοντές μας ήταν άντρες.

Η στήλη **Percent** και **Valid Percent** μας δείχνουν τη σχετική συχνότητα των τιμών της μεταβλητής στο δείγμα μας, η πρώτη επί του συνόλου των περιπτώσεων και η δεύτερη



επί του συνόλου των περιπτώσεων που έδωσαν έγκυρες απαντήσεις. Το σύνολο των περιπτώσεων στο παράδειγμά μας είναι 1519 (συμπεριλαμβανομένων των 2 system missing περιπτώσεων), ενώ το σύνολο των έγκυρων περιπτώσεων είναι 1517. Οι δύο στήλες είναι ίσες όταν δεν υπάρχουν ελλιπείς περιπτώσεις.

Η στήλη **cumulative Percent** αφορά στην αθροιστική σχετική συχνότητα των τιμών της μεταβλητής στο δείγμα μας και δεν έχει νόημα για όλες τις μεταβλητές (παρότι το SPSS υπολογίζει την αθροιστική σχετική συχνότητα κάθε φορά που θα του ζητήσουμε να μας κατασκευάσει μία κατανομή συχνότητας χρησιμοποιώντας την εντολή Frequencies). Η αθροιστική σχετική συχνότητα προφανώς δεν έχει εφαρμογή σε μεταβλητές που έχουν μόνο δύο κατηγορίες όπως στο παράδειγμά μας, αλλά και γενικώς σε όσες μεταβλητές δεν έχει νόημα η άθροιση των σχετικών συχνοτήτων. Στις περιπτώσεις αυτές θα πρέπει να μην συμπεριλαμβάνεται στα αποτελέσματα που θα παρουσιάσουμε.

Στον Πίνακα 3.2 παρουσιάζεται μία κατανομή συχνότητας όπου η σχετική αθροιστική συχνότητα έχει νόημα. Πρόκειται για μία κατανομή συχνότητας του αριθμού των παιδιών των συμμετεχόντων. Ο λόγος που δεν περιγράφουμε τη μεταβλητή ως μετρημένη σε μία αριθμητική κλίμακα είναι ότι έχουμε επιτρέψει την απάντηση «περισσότερα από 8». Εδώ η σχετική αθροιστική συχνότητα 44,7% για την απάντηση 1 παιδί σημαίνει ότι το 44,7% των συμμετεχόντων είχαν ένα ή κανένα παιδί και προκύπτει από την άθροιση των σχετικών συχνοτήτων των απαντήσεων κανένα παιδί και ένα παιδί ( $27,8 + 16,9 = 44,7$ ). Αντίστοιχα, η σχετική αθροιστική συχνότητα 99,2% για την απάντηση 4 παιδιά σημαίνει ότι το 99,2% των συμμετεχόντων είχαν από κανένα έως 4 παιδιά, ή από 4 παιδιά έως κανένα.

Πίνακας 3.2 Σχετική αθροιστική συχνότητα

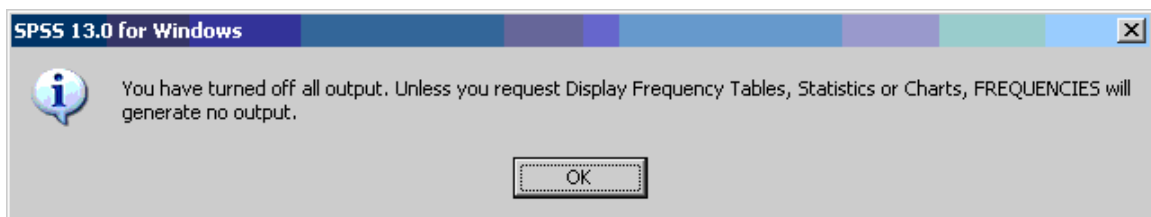
		Number of Children			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	419	27,6	27,8	27,8
	1	255	16,8	16,9	44,7
	2	375	24,7	24,9	69,5
	3	215	14,2	14,2	83,8
	4	127	8,4	8,4	92,2
	5	54	3,6	3,6	95,8
	6	24	1,6	1,6	97,3
	7	23	1,5	1,5	98,9
	Περισσότερα από 8	17	1,1	1,1	100,0
	Total	1509	99,3	100,0	
Missing	Δεν απαντώ	8	,5		
	System	2	,1		
	Total	10	,7		
Total		1519	100,0		

### 3.2 Δείκτες κεντρικής τάσης και διασποράς

Για να ζητήσουμε από το SPSS να υπολογίσει τους δείκτες κεντρικής τάσης και διασποράς που επιθυμούμε, μπορούμε να χρησιμοποιήσουμε και πάλι την εντολή **Frequencies** (υπάρχουν και άλλοι τρόποι, αλλά τα αποτελέσματα που θα πάρουμε δεν αλλάζουν).

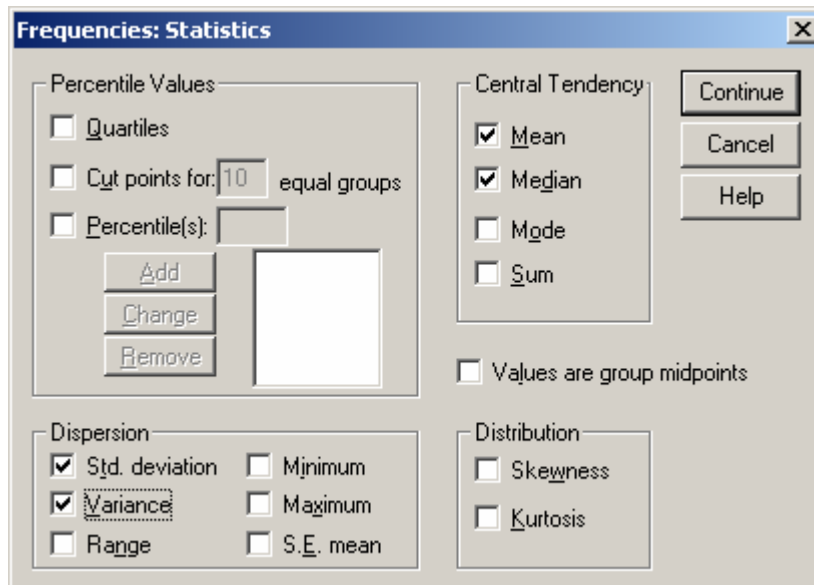
Κάνουμε πάλι κλικ στα **[Analyze=>Descriptive Statistics=>Frequencies]** και βεβαιωνόμαστε ότι στην πρώτη οθόνη που θα εμφανιστεί (βλ. Σχήμα 3.1) απενεργοποιούμε την επιλογή **[display frequency tables]**, ώστε το SPSS να μην κατασκευάσει κατανομές συχνότητες για τις μεταβλητές που θα συμμετάσχουν στις αναλύσεις. Αυτή η κίνηση θα οδηγήσει στην εμφάνιση του ακόλουθου μηνύματος (βλ. Σχήμα 3.4), το οποίο μας ενημερώνει ότι αν συνεχίσουμε χωρίς να επιλέξουμε κάποιους δείκτες, η ανάλυσή μας δε θα έχει κανένα output.

Σχήμα 3.4 Εντολή **Frequencies** - Προειδοποιητικό μήνυμα



Κάνουμε κλικ στην επιλογή **OK** και στη συνέχεια κάνουμε κλικ στο πλήκτρο **[Statistics]** που οδηγεί στην εμφάνιση της οθόνης του Σχήματος 3.5. Στην οθόνη αυτή επιλέγουμε το **Mean** από το πεδίο **Central Tendency** και **Std. deviation** από το πεδίο **Dispersion** για τις αριθμητικές μεταβλητές και **Median** από το πεδίο **Central Tendency** και **Range** από το πεδίο **Dispersion** για τις τακτικές μεταβλητές.

Σχήμα 3.5 Εντολή Frequencies – επιλογή Statistics



Το output που θα παίρναμε, αν διαλέγαμε ταυτόχρονα όλες τις παραπάνω επιλογές είναι το ακόλουθο (βλ. Πίνακα 3.3).

Πίνακας 3.3 Δείκτες κεντρικής τάσης και διασποράς

Statistics		
Ηλικία συμμετεχόντων		
N	Valid	1514
	Missing	5
Mean		45,63
Median		41,00
Std. Deviation		17,808
Range		71

Στον Πίνακα 3.3 βλέπουμε ότι είχαμε 1514 έγκυρες απαντήσεις και 5 ελλιπείς περιπτώσεις.

**Mean** είναι ο μέσος όρος που δείχνει την τιμή που θα είχαν όλες οι περιπτώσεις αν θα είχαν την ίδια τιμή.

**Median** είναι η διάμεσος και αποτελεί τη μεσαία τιμή μίας κατανομής. Αυτό σημαίνει ότι κατά το μέγιστο το 50% των περιπτώσεων λαμβάνουν τιμές μικρότερες της διαμέσου και κατά το μέγιστο το 50% μεγαλύτερες της διαμέσου. Το ακριβές ποσοστό των περιπτώσεων που βρίσκονται κάτω ή πάνω από τη διάμεσο καθορίζεται από το ποσοστό των τιμών που ισούνται με τη διάμεσο.

**Std. Deviation** είναι η τυπική απόκλιση. Η τυπική απόκλιση δείχνει πόσο πολύ “απλώνονται” οι τιμές της κατανομής. Η τυπική απόκλιση εκφράζεται σε μονάδες μέτρησης ίδιες με αυτές του μέσου όρου, και όσο μικρότερη είναι τόσο πιο κοντά στο μέσο όρο είναι η περισσότερες τιμές της κατανομής. Αν πολλές από τις τιμές της κατανομής βρίσκονται μακριά από το μέσο όρο, τότε η τυπική απόκλιση είναι μεγαλύτερη.

**Range** είναι το εύρος. Το εύρος είναι η έκταση του μικρότερου διαστήματος που περιέχει όλα τα δεδομένα και προκύπτει από τη διαφορά της μικρότερης παρατηρούμενης τιμής από τη μεγαλύτερη παρατηρούμενη τιμή.

### 3.3 Ανάλυση πολλαπλών απαντήσεων

Η ανάλυση πολλαπλών απαντήσεων χρησιμοποιείται όταν σε μία ερώτηση έχουμε επιτρέψει στους συμμετέχοντές μας να δώσουν περισσότερες από μία απαντήσεις. Για παράδειγμα μπορεί να ρωτήσουμε τους συμμετέχοντες μας «ποιες από τις παρακάτω πόλεις έχετε επισκεφθεί;». Στην περίπτωση αυτή θέλουμε να μπορούμε να καταχωρήσουμε στο SPSS όλους τους πιθανούς συνδυασμούς πόλεων: Να μην έχει επισκεφθεί καμία, να έχει επισκεφθεί μία, δύο ή και τις τρεις πόλεις.

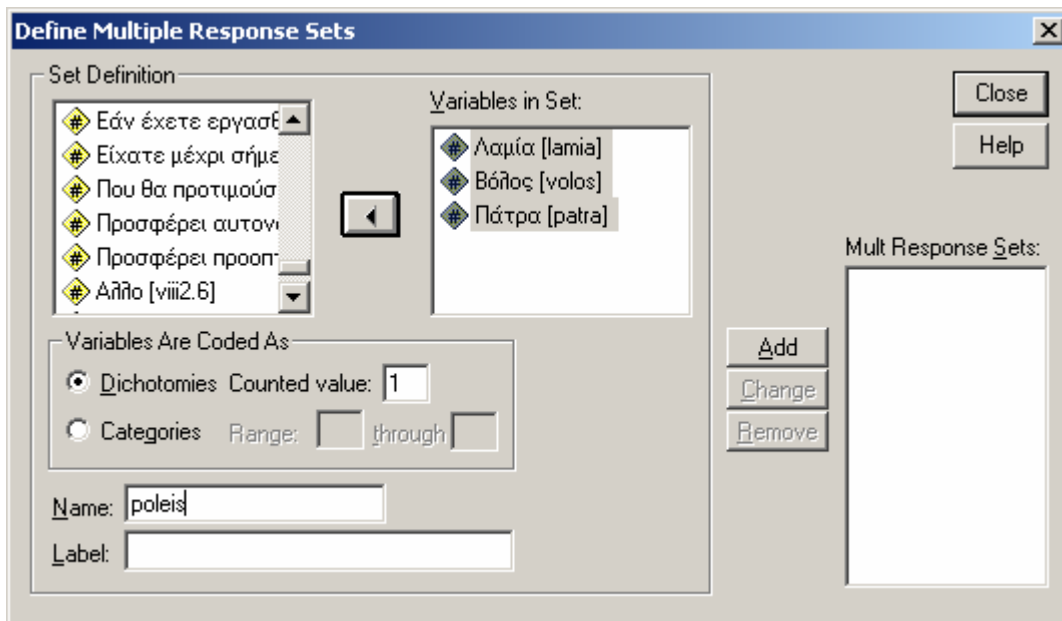
Για να το επιτύχουμε αυτό καταχωρίζουμε κάθε απάντηση ως μία μεταβλητή και κωδικοποιούμε 1 για το «ναι» και 0 για το «όχι». Έτσι ένας συμμετέχοντας που έχει επισκεφθεί και τις τρεις πόλεις θα είχε 1 – 1 – 1 στις τρεις μεταβλητές του Πίνακα 3.4, ένας συμμετέχων που δεν είχε επισκεφθεί καμία πόλη 0 – 0 – 0 και ένας συμμετέχων που είχε επισκεφθεί μόνο τον Βόλο 0 – 1 – 0.

**Πίνακας 3.4 Καταχώρηση δεδομένων για την ανάλυση πολλαπλών απαντήσεων**

lamia	volos	patra
1	1	1
0	0	0
0	1	0
0	1	1
0	0	1
1	0	0
1	1	0
...	...	...
...	...	...
0	1	1

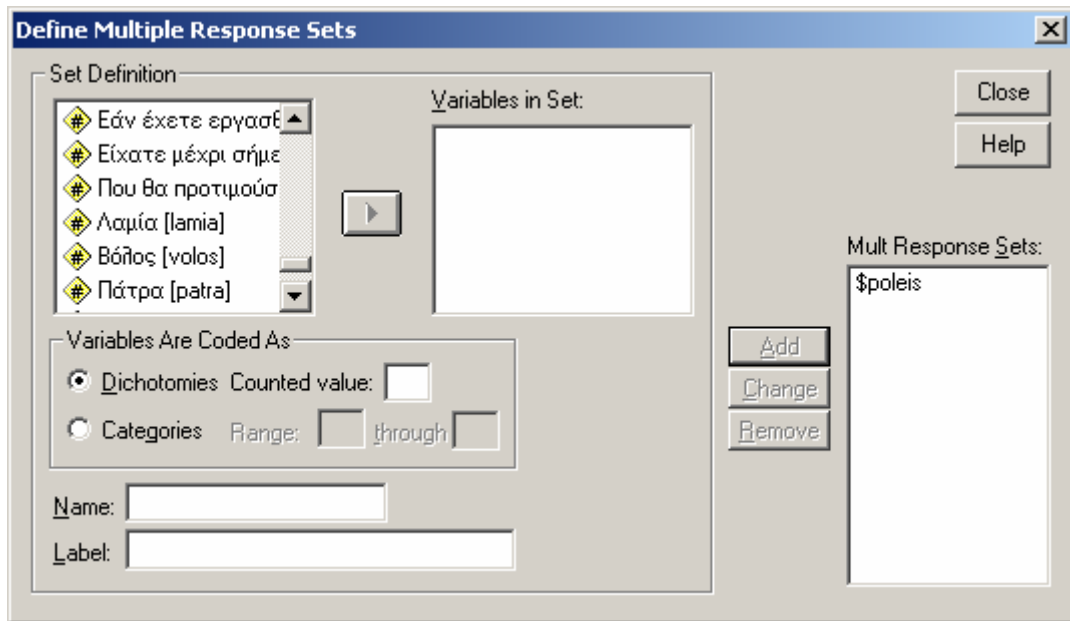
Προκειμένου να εκτελέσουμε μία ανάλυση πολλαπλών απαντήσεων επιλέγουμε **[Analyze=>Multiple Response=>Define Sets...]** και εμφανίζεται η οθόνη του Σχήματος 3.6:

**Σχήμα 3.6 Ανάλυση πολλαπλών απαντήσεων – Define Sets**



Στην οθόνη του Σχήματος 3.6 επιλέγουμε τις μεταβλητές – απαντήσεις που αντιστοιχούν στην ερώτησή μας και δημιουργούμε ένα set απαντήσεων μεταφέροντας τις σχετικές μεταβλητές στη λίστα **Variables in Set:**. Στην συνέχεια στο πεδίο **Variables Are Coded As** επιλέγουμε **Dichotomies** και στην συνέχεια στο πεδίο **Counted Value:** καταχωρούμε τον κωδικό που αντιστοιχεί στο «ναι» (1 στο παράδειγμά μας). Δίνουμε ένα όνομα στο set που δημιουργήσαμε στο πεδίο **Name:** και πατάμε το πλήκτρο **[Add]** στα δεξιά της οθόνης. Η οθόνη που θα πάρουμε είναι η ακόλουθη (βλ. Σχήμα 3.7):

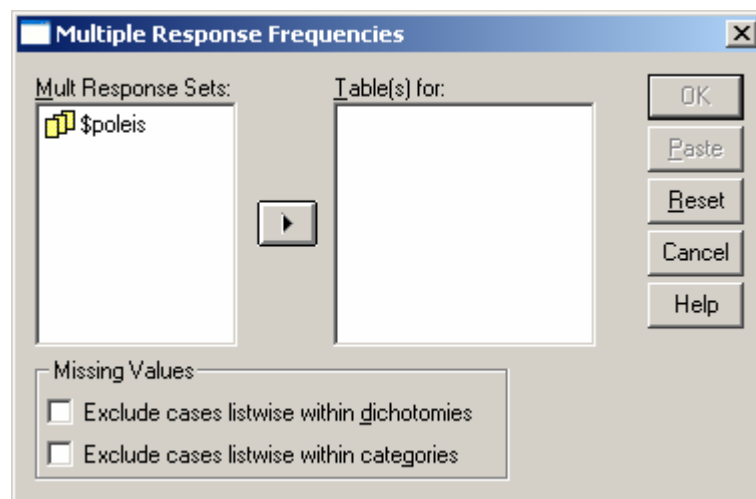
**Σχήμα 3.7** Ανάλυση πολλαπλών απαντήσεων – ορισμός set



Όπως βλέπουμε στο Σχήμα 3.7, στη λίστα **Mult Response Sets:** έχει προστεθεί το set που μόλις δημιουργήσαμε (\$poleis στο παράδειγμά μας). Επαναλαμβάνουμε την ίδια διαδικασία για όλα τα sets απαντήσεων υπάρχουν στο αρχείο δεδομένων μας και πατάμε το πλήκτρο **[Close]**.

Στη συνέχεια επιλέγουμε **[Analyze=>Multiple Response=>Frequencies...]** προκειμένου να δημιουργήσουμε μία κατανομή πολλαπλών απαντήσεων για το set απαντήσεων που δημιουργήσαμε. Η οθόνη που παίρνουμε είναι η ακόλουθη (βλ. Σχήμα 3.8):

**Σχήμα 3.8** Ανάλυση πολλαπλών απαντήσεων – frequencies



Στην οθόνη του Σχήματος 3.8 επιλέγουμε τα sets απαντήσεων που μας ενδιαφέρουν και τα μεταφέρουμε στη λίστα **Table(s) for:** και στη συνέχεια πατάμε το πλήκτρο **[OK]**. Στην έξοδο θα πάρουμε τα αποτελέσματα που βλέπουμε στον Πίνακα 3.5:

**Πίνακας 3.5 Ανάλυση πολλαπλών απαντήσεων – απαντήσεων**

\$poleis Frequencies				
		Responses		Percent of Cases
		N	Percent	
\$poleis <sup>a</sup>	Λαμία	329	47,1%	67,0%
	Βόλος	206	29,5%	42,0%
	Πάτρα	163	23,4%	33,2%
Total		698	100,0%	142,2%

a. Dichotomy group tabulated at value 1.

Οι δύο πρώτες στήλες του Πίνακα 3.5 αφορούν στις απόλυτες και σχετικές συχνότητες επί των απαντήσεων (Responses) και όχι επί των συμμετεχόντων. Όπως βλέπουμε συνολικά στην ερώτησή μας δόθηκαν 698 θετικές απαντήσεις, 329 για τη Λαμία, 206 για το Βόλο και 163 για την Πάτρα. Το 47,1% των απαντήσεων ήταν θετικές απαντήσεις για τη Λαμία, το 29,5% για το Βόλο και το 23,4% για την Πάτρα.

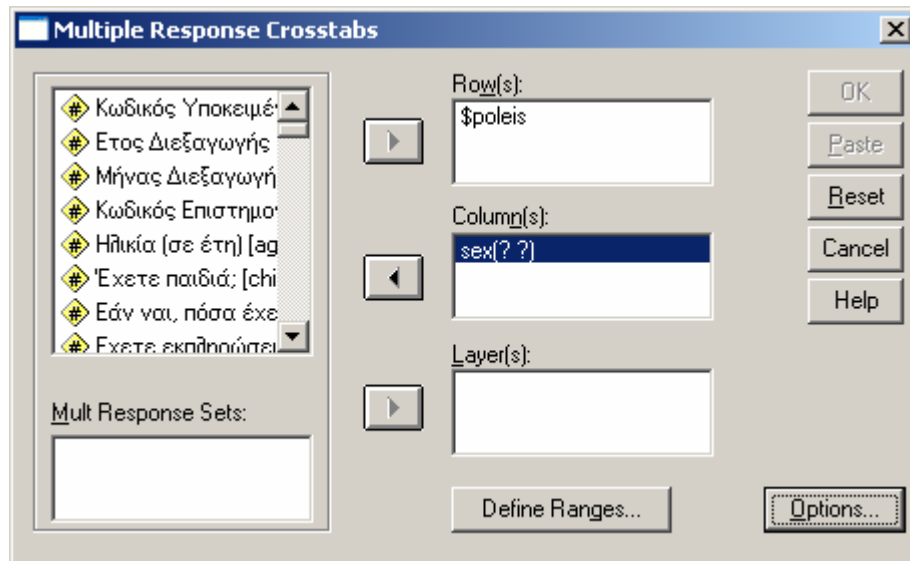
Στην τρίτη στήλη βλέπουμε σχετικές συχνότητες επί των συμμετεχόντων (Cases). Όπως είδαμε κάθε συμμετέχων μπορούσε να δώσει περισσότερες από μία θετικές απαντήσεις (από καμία μέχρι τρεις) και αυτός είναι ο λόγος που το οι σχετικές συχνότητες αθροίζουν 142,2% και όχι 100%. Αυτό σημαίνει ότι κατά μέσο όρο ο κάθε συμμετέχων προσέφερε 1,42 θετικές απαντήσεις.

Ο σωστός τρόπος να διαβαστεί η τρίτη στήλη του Πίνακα 3.5 είναι ο ακόλουθος: 67,0% των συμμετεχόντων ανέφεραν μεταξύ των πόλεων που έχουν επισκεφθεί τη Λαμία, 42,0% το Βόλο και 33,2% την Πάτρα.

Επιλέγοντας **[Analyze=>Multiple Response=>Crosstabs...]** μπορούμε να δημιουργήσουμε πίνακες διπλής εισόδου χρησιμοποιώντας συνδυασμούς set απαντήσεων και μεταβλητών. Η οθόνη που εμφανίζεται είναι η ακόλουθη (βλ. Σχήμα 3.9):

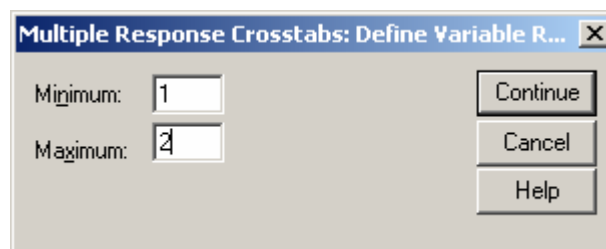


Σχήμα 3.8 Ανάλυση πολλαπλών απαντήσεων – crosstabs



- Στα πεδία **Row(s):** και **Columns:** βάζουμε τις μεταβλητές που θα παρουσιάζονται στον πίνακα διπλής εισόδου. Στο παράδειγμά μας επιλέξαμε το set απαντήσεων που ορίσαμε προηγουμένως (\$roleis) για τις σειρές του πίνακα και τη μεταβλητή φύλο (sex) για τις στήλες του πίνακα.
- Όπως βλέπετε στο Σχήμα 3.8 εμφανίζονται δύο λατινικά ερωτηματικά δίπλα από το όνομα της μεταβλητής sex. Στα ερωτηματικά αυτά πρέπει να αντιστοιχήσουμε τους κωδικούς που αντιστοιχούν στις κατηγορίες της μεταβλητής. Επιλέγουμε τη μεταβλητή ομαδοποίησης και κάνουμε κλικ στο πλήκτρο **[Define Ranges...]**, με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 3.9.

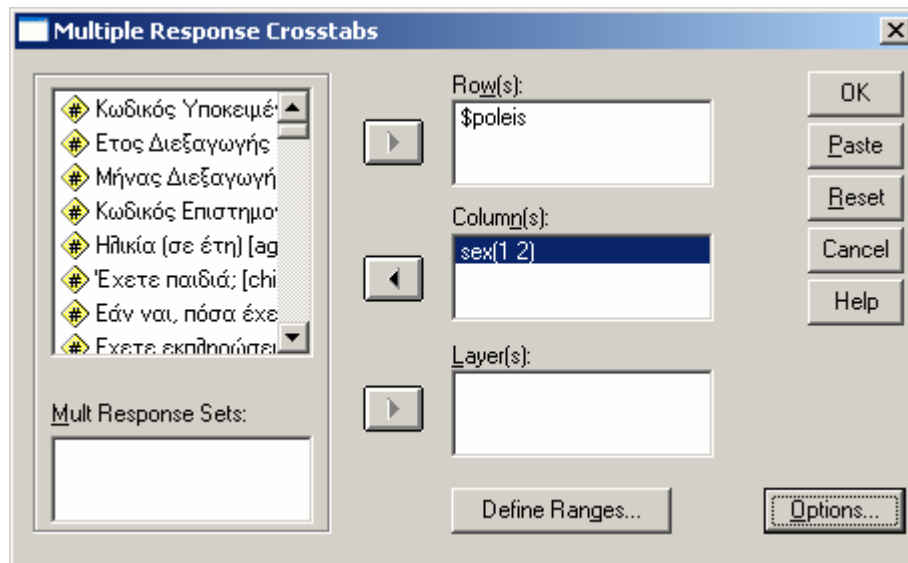
Σχήμα 3.9 Ανάλυση πολλαπλών απαντήσεων - define ranges



- Στην οθόνη του Σχήματος 3.9 ορίζουμε τον ελάχιστον και τον μεγαλύτερο κωδικό που περιλαμβάνει όλες τις κατηγορίες της μεταβλητής. Στο παράδειγμά μας η μεταβλητή φύλο έχει δύο κατηγορίες: 1=Άντρας, 2=Γυναίκα και για αυτό βάζουμε 1 στο πεδίο **Minimum:** και 2 στο πεδίο **Maximum:..** Αν είχαμε μία άλλη μεταβλητή με

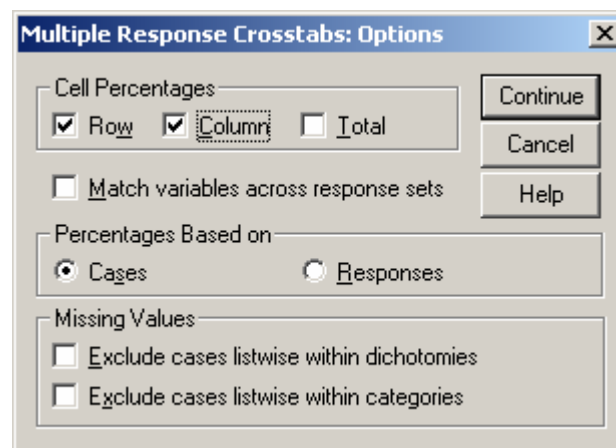
πέντε κατηγορίες, οι οποίες είχαν κωδικοποιηθεί με 1,2,3,4,5 θα βάζαμε 1 στο πεδίο **Minimum:** και 5 στο πεδίο **Maximum:**. Στη συνέχεια πατάμε **[OK]** και η οθόνη μας παίρνει την μορφή που παρουσιάζεται στο Σχήμα 3.10:

**Σχήμα 3.10** Ανάλυση πολλαπλών απαντήσεων - crosstabs μετά το define ranges



- Στην συνέχεια κάνουμε κλικ στο πλήκτρο **[Options...]** και εμφανίζεται η οθόνη του Σχήματος 3.11:

**Σχήμα 3.11** Ανάλυση πολλαπλών απαντήσεων - crosstabs - options



- Στην οθόνη του Σχήματος 3.11 επιλέγουμε **Row** και **Column** στο πεδίο **Cell Percentages**. Στο πεδίο **Percentages Based on** μπορούμε να επιλέξουμε αν τα ποσοστά που θα εμφανιστούν θα βασίζονται στους συμμετέχοντες (Cases) ή στις απαντήσεις (Responses). Η προεπιλογή είναι να εμφανίζονται ποσοστά για τους συμμετέχοντες.
- Πατάμε το πλήκτρο **[Continue]** και επιστρέφουμε στην οθόνη του Σχήματος 3.10, όπου δίνουμε **[OK]** και παίρνουμε στην έξοδο τα ακόλουθα αποτελέσματα (βλ. Πίνακα 3.6):

**Πίνακας 3.11 Ανάλυση πολλαπλών απαντήσεων - crosstabs - αποτελέσματα**

**\$poleis\*sex Crosstabulation**

			Φύλο		Total
			Ανδρας	Γυναίκα	
\$poleis	Λαμία	Count	82	247	329
		% within \$poleis	24,9%	75,1%	
		% within sex	48,0%	77,2%	
	Βόλος	Count	97	109	206
		% within \$poleis	47,1%	52,9%	
		% within sex	56,7%	34,1%	
	Πάτρα	Count	26	137	163
		% within \$poleis	16,0%	84,0%	
		% within sex	15,2%	42,8%	
Total	Count	171	320	491	

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

- Η ερμηνεία του πίνακα γίνεται σύμφωνα με το παράδειγμα που δώσαμε για τις σχετικές συχνότητες επί των συμμετεχόντων (προσέξτε την υποσημείωση: Percentages and totals are based on respondents.). Για την ερμηνεία πινάκων διπλής εισόδου δείτε το κεφάλαιο για το  $\chi^2$ .

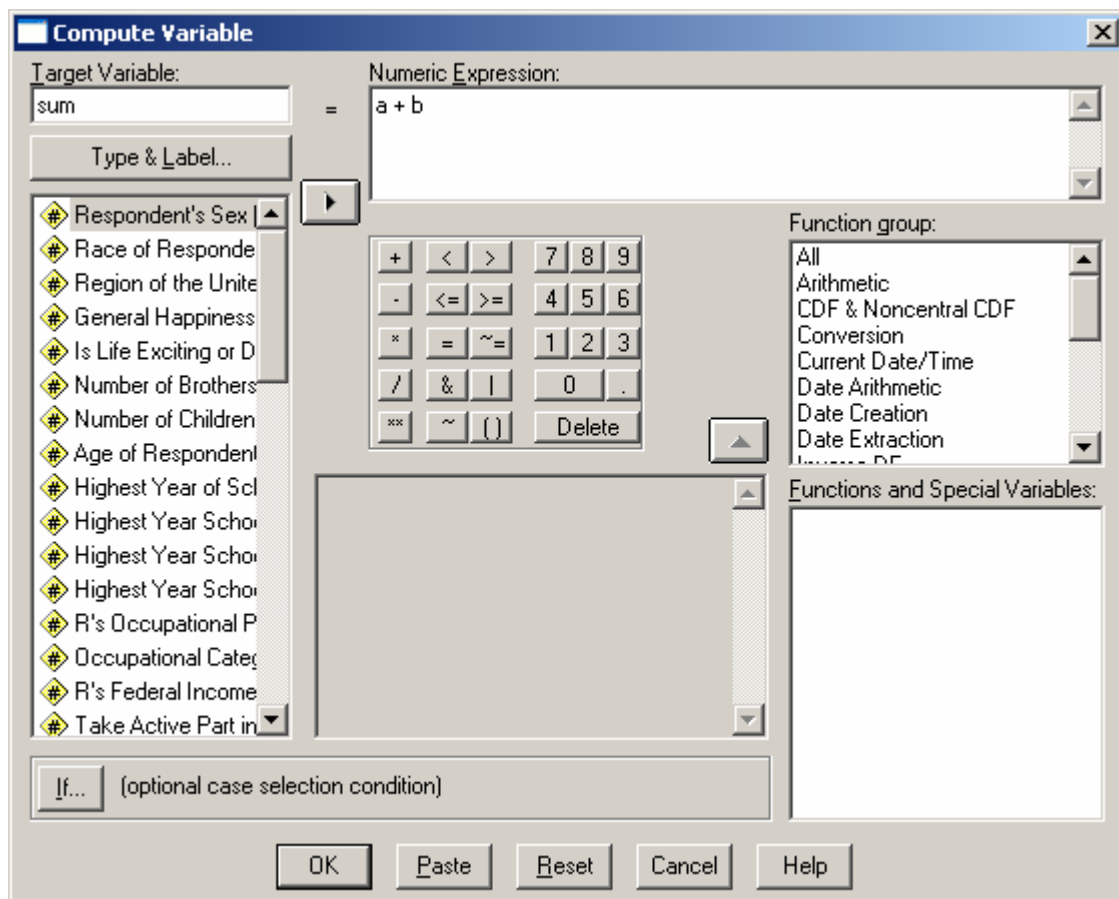
## 4.0 Μετασχηματισμοί μεταβλητών

Αφού είδαμε πώς γίνονται οι περιγραφικές στατιστικές αναλύσεις, μπορούμε να αναφερθούμε σε μία σειρά πολύ σημαντικών δυνατοτήτων του SPSS που αφορά στο μετασχηματισμό των μεταβλητών με σκοπό την αλλαγή υπάρχουσών μεταβλητών ή τη δημιουργία νέων.

### 4.1 Compute

Η απλούστερη εντολή για τη δημιουργία μίας νέας μεταβλητής είναι η εντολή **[Compute]**. Για να την χρησιμοποιήσουμε επιλέγουμε **[Transform=>Compute]** και εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 4.1.

Σχήμα 4.1 Η εντολή **[Compute]**



Στο πεδίο **[Target Variable]** στο πάνω αριστερά τμήμα της οθόνης πληκτρολογούμε το όνομα της νέας μεταβλητής που θέλουμε να κατασκευάσουμε, ενώ στο πεδίο **[Numeric Expression]** τον αριθμητικό τύπο βάση του οποίου θα υπολογίζεται η τιμή της νέας

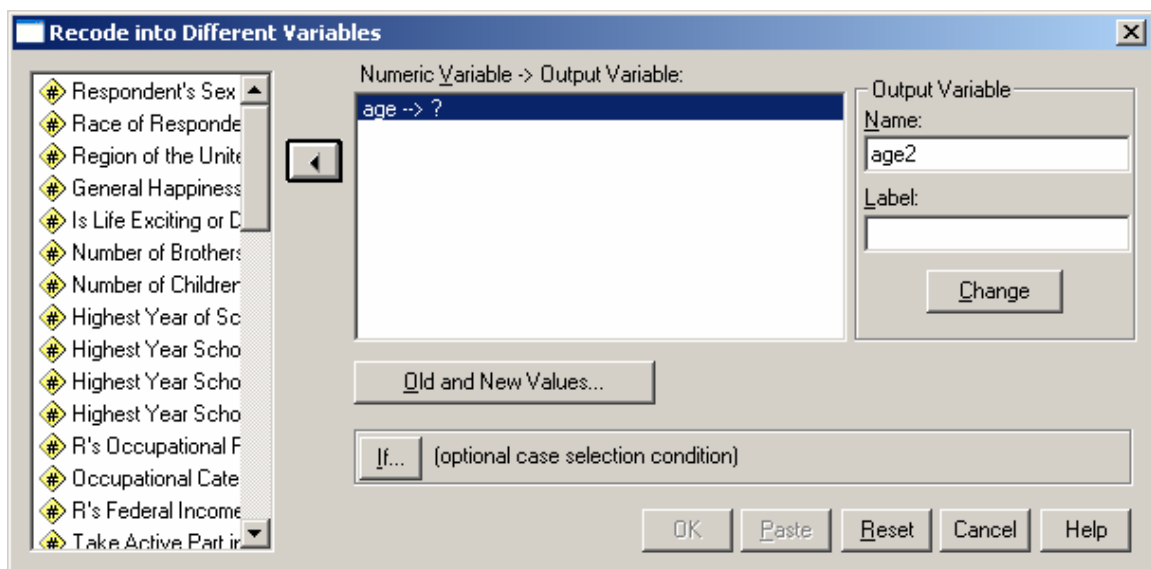
μεταβλητής για κάθε περίπτωση. Στο παράδειγμα του Σχήματος 4.1 έχουμε επιλέξει να δημιουργήσουμε τη νέα μεταβλητή *sum*, η οποία ισούται με το άθροισμα των μεταβλητών *a + b*. Στον αριθμητικό τύπο μπορούμε να χρησιμοποιήσουμε οποιαδήποτε υπάρχουσα μεταβλητή, ένα νούμερο που επιθυμούμε ή συνδυασμό αυτών. Για παράδειγμα, αν υποθέσουμε ότι έχουμε ζητήσει από τους συμμετέχοντές μας το έτος γέννησής τους (μεταβλητή *etos*), μπορούμε να δημιουργήσουμε μία νέα μεταβλητή με την ηλικία τους, η οποία να ισούται με το τρέχον έτος μείον το έτος γέννησης (δηλαδή  $ilikia=2006 - etos$ ).

## 4.2 Recode

Μία δεύτερη πολύ χρήσιμη εντολή του SPSS είναι η εντολή **[Recode]**. Για να τη χρησιμοποιήσουμε επιλέγουμε **[Transform=>Recode=>Into Different Variables]** ή **[Transform=>Recode=>Into Same Variables]**. Η διαφορά μεταξύ των δύο επιλογών έγκειται στο αν οι αλλαγές θα εφαρμοσθούν σε μία υπάρχουσα μεταβλητή ή αν θα δημιουργηθεί μία νέα μεταβλητή, αλλά κατά τα άλλα η λογική είναι απολύτως ίδια. Στο παράδειγμα που ακολουθεί θα χρησιμοποιήσουμε την επιλογή για τη δημιουργία μίας νέας μεταβλητής, που είναι και πιο ασφαλής πρακτική καθώς διατηρούμε αναλλοίωτη και την αρχική μας μεταβλητή.

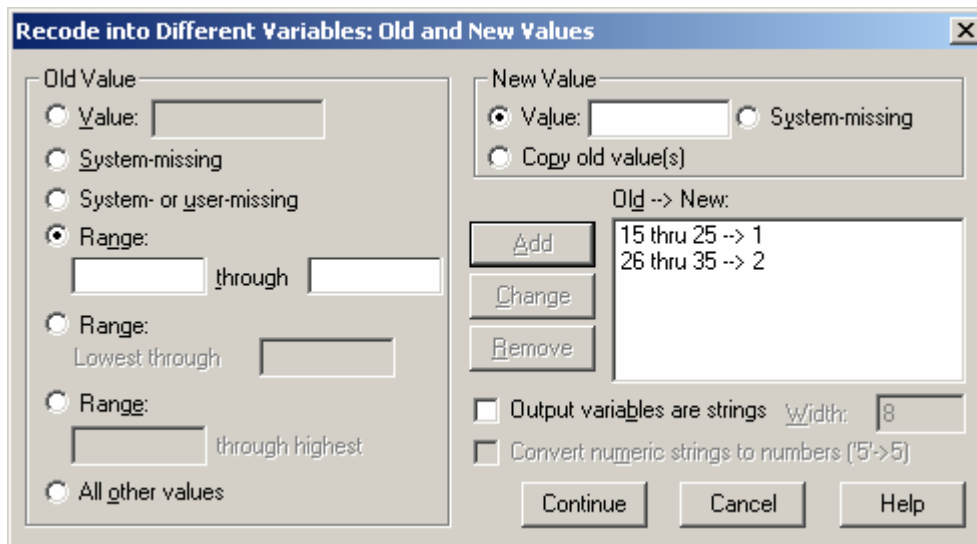
Επιλέγοντας **[Transform=>Recode=>Into Different Variables]** εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 4.2.

**Σχήμα 4.2 Η πρώτη οθόνη της εντολής [Recode]**



Ας υποθέσουμε ότι θέλουμε να μετασχηματίσουμε τις τιμές της αριθμητικής μεταβλητής ηλικία (όπου έχουμε τις ηλικίες εκφρασμένες σε αριθμούς) σε κατηγορίες που αντιπροσωπεύουν ηλικιακές ομάδες. Επιλέγουμε από τη λίστα των μεταβλητών που υπάρχουν στο αρχείο μας τη μεταβλητή *age* και τη μεταφέρουμε στο κεντρικό παράθυρο (βλ. Σχήμα 4.2.). Στη συνέχεια βάζουμε το όνομα της νέας μεταβλητής που θέλουμε να δημιουργήσουμε στο πεδίο **[Output Variable]** στο άνω δεξιά τμήμα της οθόνης (*age2* στο παράδειγμά μας) και κάνουμε κλικ στο πλήκτρο **[Old and New Values]**. Εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 4.3.

**Σχήμα 4.3 Η δεύτερη οθόνη της εντολής [Recode]**



Η οθόνη αυτή χωρίζεται σε δύο πεδία. Στα αριστερά υπάρχει το πεδίο **[Old Value]** όπου δίνουμε τις συγκεκριμένες τιμές ή το εύρος των τιμών της αρχικής μεταβλητής που θέλουμε να αλλάξουμε. Στα δεξιά υπάρχει το πεδίο **[New Value]** όπου καταχωρούμε τις τιμές της νέας μεταβλητής που αντιστοιχούν στις τιμές της παλιάς μεταβλητής που επιλέξαμε στα αριστερά. Όταν ολοκληρωθεί αυτή η διαδικασία, κάνουμε κλικ στο πλήκτρο **[Add]** και η αντιστοίχιση που επιλέξαμε εμφανίζεται στο παράθυρο **[Old→New]**.

Στο παράδειγμα που παρουσιάζεται στον Πίνακα 4.3. Χρησιμοποιήσαμε την επιλογή **[Range]** δύο φορές. Την πρώτη φορά βάλαμε τις τιμές 15 through 25 και στο πεδίο **[New Value]** την τιμή 1 και κάναμε κλικ στο πλήκτρο **[Add]**. Τη δεύτερη φορά βάλαμε τις τιμές 26 through 35 και στο πεδίο **[New Value]** την τιμή 2 και κάναμε κλικ στο πλήκτρο **[Add]**.

Όταν ολοκληρώσουμε τις αντιστοιχίσεις που θέλουμε κάνουμε κλικ στο πλήκτρο **[Continue]** και επιστρέφουμε στην πρώτη οθόνη (βλ. Σχήμα 4.2), όπου κάνουμε πρώτα κλικ στο πλήκτρο **[Change]** και στη συνέχεια στο πλήκτρο **[OK]**. Έχουμε δημιουργήσει μία νέα μεταβλητή με το όνομα *age2*, στην οποία όποιος από τους συμμετέχοντές μας είχε τιμή από 15 έως 25 στην ηλικία θα έχει τιμή 1 και όποιος είχε από 26 έως 35 θα έχει τιμή 2. Η νέα μεταβλητή φαίνεται στην τελευταία στήλη των δεδομένων μας.

### 4.3 Weight Cases

Το SPSS προσφέρει έναν επιπλέον τρόπο μετασχηματισμού δεδομένων μέσω της εντολής **[Weight Cases]**, ο οποίος είναι πολύ χρήσιμος για την καταχώρηση και περαιτέρω στατιστική επεξεργασία δευτερογενών δεδομένων.

Η συγκεκριμένη διαδικασία ονομάζεται **κωδικοποίηση των δεδομένων** στο φύλλο εισαγωγής δεδομένων. Συνήθως τα δεδομένα εισάγονται σε στήλες και συγκεκριμένα μια μεταβλητή σε κάθε στήλη. Ωστόσο, η μέθοδος αυτή δεν είναι επαρκής όταν χρησιμοποιούμε δευτερογενή δεδομένα ή τουλάχιστον δεν είναι η ευκολότερη.

Ας υποθέσουμε ότι γνωρίζουμε ότι σε μία έρευνα με δύο μεταβλητές με δύο κατηγορίες η κάθε μία λάβαμε τα δεδομένα που παρουσιάζονται στον Πίνακα 4.1. Η μεταβλητή 1 έχει τις κατηγορίες A και B και η μεταβλητή 2 τις κατηγορίες X και Ψ. Ο πίνακας περιλαμβάνει τις απόλυτες συχνότητες των τιμών που προέκυψαν για κάθε συνδυασμό κατηγοριών (A-X, A-Ψ, B-X, B-Ψ). Για παράδειγμα, στο πάνω αριστερά κελί υπάρχει η απόλυτη συχνότητα των συμμετεχόντων που έδωσαν την απάντηση A στη Μεταβλητή 1 και την απάντηση X στη Μεταβλητή 2, η οποία ισούται με 153.

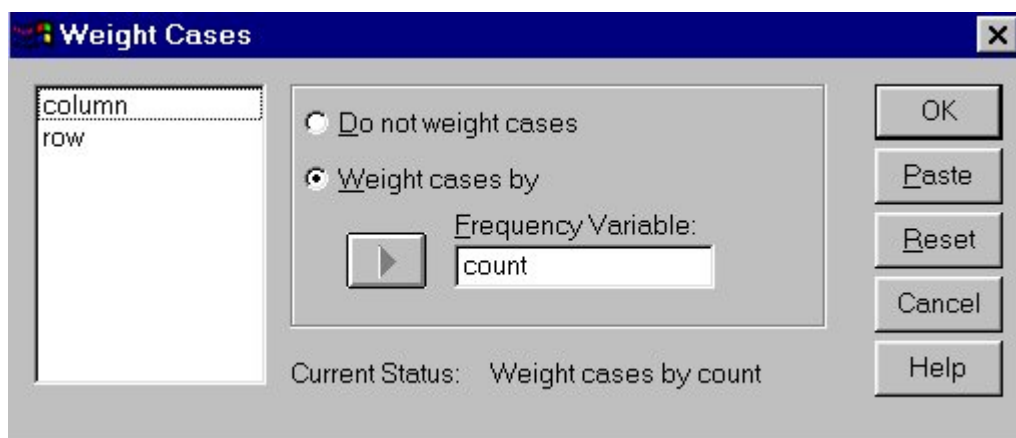
**Πίνακας 4.1 Δευτερογενή δεδομένα**

		Μεταβλητή 1		
		A	B	Σύνολα
Μεταβλ. 2	X	153	24	177
	Ψ	105	76	181
Σύνολα		258	100	358

Ο παραπάνω πίνακας θα μπορούσε να καταχωρηθεί στο SPSS χρησιμοποιώντας 3 μεταβλητές και τέσσερις μόνο γραμμές ( $2 \times 2 = 4$ ), ως εξής:

Row	Column	Count
1	1	153
1	2	24
2	1	105
2	2	76

Στις στήλες *Row* και *Column* δημιουργούμε όλους τους πιθανούς συνδυασμούς των κατηγοριών των δύο μεταβλητών χρησιμοποιώντας τους αριθμούς 1 και 2. Στη στήλη *Count* δίνουμε την απόλυτη συχνότητα που γνωρίζουμε από τον Πίνακα. 4.1. Στη συνέχεια ενημερώνουμε το SPSS ότι η μεταβλητή *Count* αναπαριστά την απόλυτη συχνότητα για κάθε μοναδικό συνδυασμό μεταξύ στήλης και γραμμής. Αυτό γίνεται στην εντολή **WEIGHT** και για να το κάνουμε αυτό επιλέγουμε το **[Data => Weight Cases]**. Στο παράθυρο που ανοίγει, ενεργοποιούμε την επιλογή **Weight cases by**, και μετακινούμε τη μεταβλητή **COUNT** στο κουτί **Frequency Variable**. Αν ξεχάσουμε αυτό το βήμα, η τιμή σε κάθε κελί θα είναι 1 για τον πίνακα αυτόν.



Αν δε χρησιμοποιούσαμε την εντολή **[Weight Cases]** θα έπρεπε να δημιουργήσουμε ένα αρχείο δεδομένων με δύο μεταβλητές (Μεταβλητή 1 και Μεταβλητή 2) και σε κάθε μία από αυτές να καταχωρήσουμε A ή B και X ή Ψ αντίστοιχα για κάθε έναν από τους 358 συμμετέχοντες στην έρευνά μας. Αυτό εκτός από εξαιρετικά χρονοβόρο, θα ήταν εφικτό



μόνο αν είχαμε στη διάθεσή μας τις απαντήσεις του κάθε συμμετέχοντα ξεχωριστά και όχι έναν συγκεντρωτικό πίνακα, όπως π.χ. ο Πίνακας 4.1, όπως συνήθως συμβαίνει με τα δευτερογενή δεδομένα.

## 5.0 Επαγωγική Στατιστική

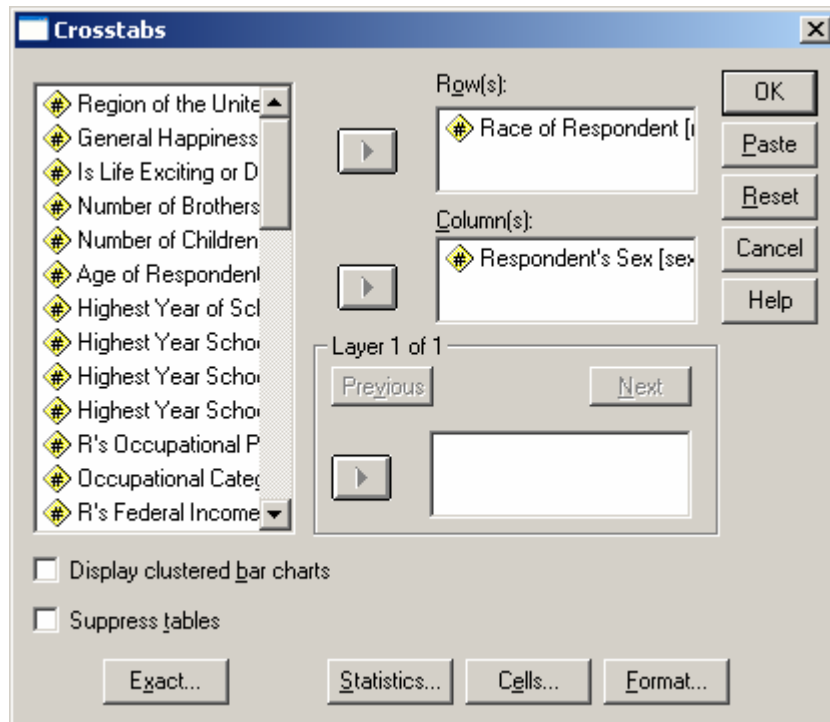
Για να γνωρίζουμε πώς πρέπει να αναλύσουμε δύο ή περισσότερες μεταβλητές, πρέπει να γνωρίζουμε την κλίμακα μέτρησής τους. Συνοπτικά ισχύουν τα ακόλουθα:

Μεταβλητή (Κλίμακα μέτρησης)	×	Μεταβλητή (Κλίμακα μέτρησης)	Επαγωγικό στατιστικό κριτήριο/τεστ
Κατηγορική	×	Κατηγορική	$\chi^2$ (συσχέτιση) [μη παραμετρικό]
Κατηγορική (μέχρι 2 κατηγορίες)	×	Τακτική	Mann-Whitney U (σύγκριση ομάδων) [μη παραμετρικό]
Κατηγορική (πολλές κατηγορίες)	×	Τακτική	Kruskall-Wallis H (σύγκριση ομάδων) [μη παραμετρικό]
Κατηγορική (μέχρι 2 κατηγορίες)	×	Αριθμητική	t-τεστ ανεξάρτητων δειγμάτων (σύγκριση ομάδων) [παραμετρικό]
Κατηγορική (πολλές κατηγορίες)	×	Αριθμητική	One-way ANOVA (σύγκριση ομάδων) [παραμετρικό]
Τακτική(ες)	×	Τακτική(ες)	Spearman Rho (συνάφειες) [μη παραμετρικό]
Τακτική	×	Τακτική	Wilcoxon (σύγκριση μετρήσεων ή μεταβλητών – 2 μεταβλητές ή μετρήσεις) [μη παραμετρικό]
Τακτική(ες)	×	Τακτική(ες)	Friedman (σύγκριση μετρήσεων ή μεταβλητών – πολλές μεταβλητές ή μετρήσεις) [μη παραμετρικό]
Αριθμητική(ες)	×	Αριθμητική(ες)	Pearson r (συνάφειες) [παραμετρικό]
Αριθμητική	×	Αριθμητική	t-τεστ εξαρτημένων δειγμάτων (σύγκριση μετρήσεων ή μεταβλητών – 2 μεταβλητές ή μετρήσεις) [παραμετρικό]
Αριθμητική(ες)	×	Αριθμητική(ες)	Repeated measures ANOVA (σύγκριση μετρήσεων ή μεταβλητών – πολλές μεταβλητές ή μετρήσεις) [παραμετρικό]

## 5.1 Τεστ $\chi^2$

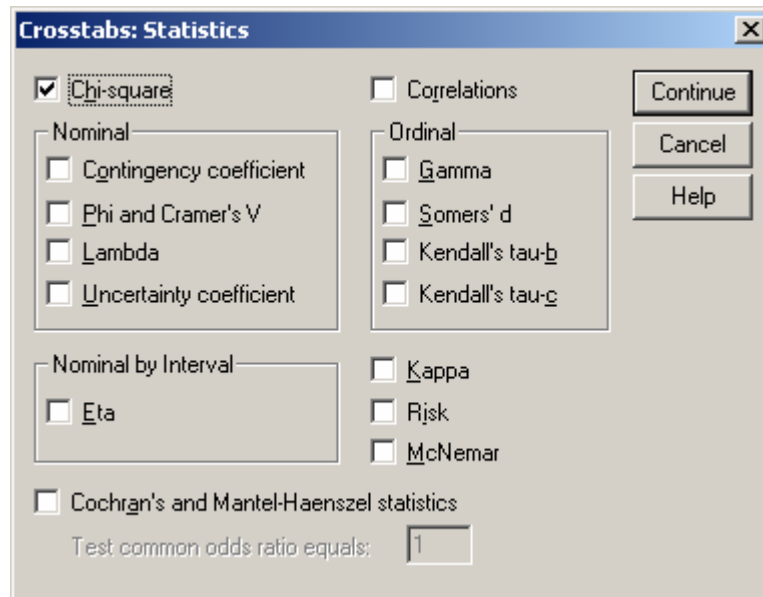
Ο υπολογισμός του τεστ  $\chi^2$  γίνεται μέσω του μενού **[Analyze => Descriptive Statistics => Crosstabs...]**. Η οθόνη που θα εμφανιστεί παρουσιάζεται στο Σχήμα 5.1.

**Σχήμα 5.1 Η εντολή crosstabs**



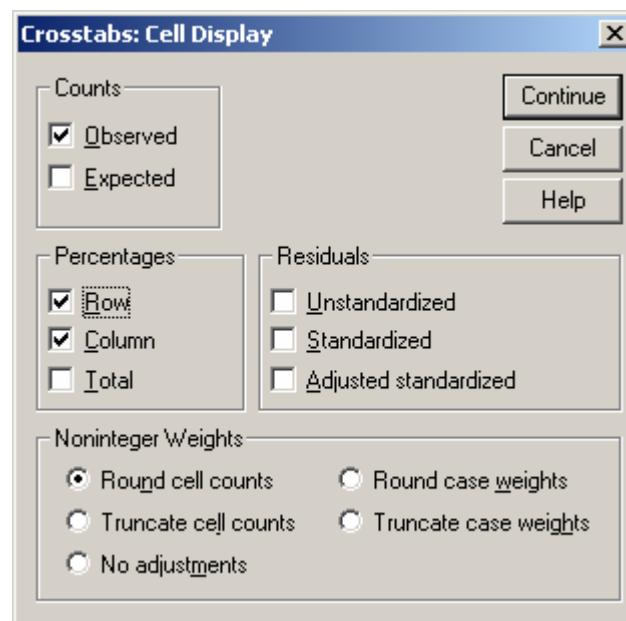
- Στα κουτιά **Row(s):** και **Column(s):** βάζουμε τις μεταβλητές που θα συμμετάσχουν στην ανάλυση. Συνήθως προτιμούμε να βάλουμε στο κουτί **Row(s):** τη μεταβλητή με τις περισσότερες κατηγορίες, ώστε ο πίνακας διπλής εισόδου που θα προκύψει να εκτείνεται καθ' ύψος και όχι κατά πλάτος.
- Στο κάτω τμήμα του κουτιού βρίσκονται τρία κουμπιά, από τα οποία το σημαντικότερο είναι το κουμπί **[Statistics...]**. Πρέπει να κάνουμε κλικ σε αυτό και να επιλέξουμε το **Chi-square**, αλλιώς το στατιστικό κριτήριο δεν θα υπολογιστεί. Από τα περιεχόμενα του κουτιού αυτού προκύπτει ότι μπορούμε να υπολογίσουμε και άλλα κριτήρια παράλληλα προς το  $\chi^2$ . Η οθόνη που εμφανίζεται με το κουμπί **[Statistics...]** παρουσιάζεται στο Σχήμα 5.2.

Σχήμα 5.2 Η εντολή crosstabs – επιλογή statistics



- Αφού επιλέξουμε το **Chi-square**, όπως φαίνεται στο σχήμα, κάνουμε κλικ στο **[Continue]** για να επιστρέψουμε στην αρχική οθόνη (βλ. Σχήμα 5.1). Από εκεί κάνουμε κλικ στο πλήκτρο **[Cells...]** και εμφανίζεται η οθόνη του Σχήματος 5.3.

Σχήμα 5.3 Η εντολή crosstabs – επιλογή cells



- Στην οθόνη του Σχήματος 5.3 είναι σημαντικό να επιλέξουμε το **Row** και το **Column** από το πεδίο **Percentages**, προκειμένου να εμφανίζονται οι σχετικές συχνότητες επί των γραμμών και επί των στηλών στον πίνακα διπλής εισόδου που θα κατασκευάσει το SPSS. Αφού ολοκληρώσουμε τις επιλογές μας, πατάμε πάλι **[Continue]** για να επιστρέψουμε στην αρχική οθόνη (βλ. Σχήμα 5.1).
- Στην αρχική οθόνη (Σχήμα 5.1), δίνουμε **[OK]** για να εκτελεστεί η ανάλυση. Τα αποτελέσματα της ανάλυσης φαίνονται στον Πίνακα 5.1 και 5.2.

**Πίνακας 5.1 Πίνακας διπλής εισόδου**

**Race of Respondent \* Respondent's Sex Crosstabulation**

			Respondent's Sex		Total
			Male	Female	
Race of Respondent	White	Count	545	719	1264
		% within Race of Respondent	43,1%	56,9%	100,0%
		% within Respondent's Sex	85,7%	81,6%	83,3%
	Black	Count	71	133	204
		% within Race of Respondent	34,8%	65,2%	100,0%
		% within Respondent's Sex	11,2%	15,1%	13,4%
	Other	Count	20	29	49
		% within Race of Respondent	40,8%	59,2%	100,0%
		% within Respondent's Sex	3,1%	3,3%	3,2%
Total	Count	636	881	1517	
	% within Race of Respondent	41,9%	58,1%	100,0%	
	% within Respondent's Sex	100,0%	100,0%	100,0%	

- Το σημαντικό στον Πίνακα 5.1 είναι ότι συγκρίνουμε τις σχετικές συχνότητες επί των στηλών και όχι τις σχετικές συχνότητες επί των σειρών. Ο λόγος είναι ότι οι σχετικές συχνότητες επί των σειρών (**% within Race of Respondent** στο παράδειγμά μας) επηρεάζονται από το μέγεθος των επιμέρους ομάδων που προκύπτουν από τις κατηγορίες της μεταβλητής. Για να ξεχωρίζετε τις σχετικές συχνότητες επί των σειρών να θυμάστε ότι πάντα αθροίζουν 100 οριζοντίως. Στον παραπάνω πίνακα συγκρίνουμε καθέτως τις σχετικές συχνότητες επί των στηλών (**% within Respondent's Sex** στο παράδειγμά μας).

- Στον Πίνακα 5.1 παρουσιάζονται τα αποτελέσματα της ανάλυσης. Αυτό που μας ενδιαφέρει η πρώτη σειρά **Pearson Chi-Square**. Στο παράδειγμα μας η συσχέτιση δεν είναι στατιστικώς σημαντική, εφόσον το Sig., δηλαδή το  $p$ , είναι μεγαλύτερο του 0,05. Στο κάτω μέρος του πίνακα υπάρχει μία υποσημείωση που αφορά στις προϋποθέσεις εφαρμογής του τεστ. Το θέμα αυτό αποτελεί αντικείμενο διχογνωμίας, αλλά ένας καλός εμπειρικός κανόνας είναι το 20% των κελιών ή λιγότερο να έχουν θεωρητική συχνότητα (expected count) μικρότερη του 5.

**Πίνακας 5.2 Αποτελέσματα του τεστ  $\chi^2$**

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,011 <sup>a</sup>	2	,082
Likelihood Ratio	5,094	2	,078
Linear-by-Linear Association	2,944	1	,086
N of Valid Cases	1517		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 20,54.

- Ειδικά για τις περιπτώσεις που και οι δύο μεταβλητές έχουν από δύο κατηγορίες (στο αρχικό μας παράδειγμα η μεταβλητή race είχε τρεις κατηγορίες) το SPSS υπολογίζει αυτόματα μία επιπλέον σειρά αποτελεσμάτων (βλ. Πίνακα 5.3). Υπάρχει διχογνωμία ως το κατά πόσο θα πρέπει να χρησιμοποιείται η διόρθωση αυτή (**Continuity Correction**), αν δηλαδή θα πρέπει να διαβάσουμε αυτήν τη γραμμή ή την γραμμή **Pearson Chi-Square**, όπως περιγράψαμε αρχικά.
- Ειδικά για τους πίνακες 2x2 κανένα κελί από τα τέσσερα δεν θα πρέπει να έχει θεωρητική συχνότητα μικρότερη του 5 (είδαμε προηγουμένως ότι ο εμπειρικός κανόνας θέτει ως όριο το 20%, αν έστω και ένα κελί από τα τέσσερα έχει θεωρητική συχνότητα μικρότερη του 5, τότε έχουμε πρόβλημα στο 25% των κελιών).

**Πίνακας 5.3 Αποτελέσματα του τεστ  $\chi^2$  για πίνακα 2x2**

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	35.930 <sup>b</sup>	1	.000		
Continuity Correction <sup>a</sup>	34.532	1	.000		
Likelihood Ratio Fisher's Exact Test	37.351	1	.000	.000	.000
Linear-by-Linear Association	35.830	1	.000		
N of Valid Cases	358				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 49.44.

## 5.2 Το t-τεστ

Υπάρχουν δύο ειδών t-τεστ, ένα για ανεξάρτητα δείγματα και ένα για εξαρτημένα δείγματα. Προκειμένου να ενημερώσουμε το SPSS για αυτή την πληροφορία, χρειάζεται να ακολουθήσουμε δύο διαφορετικούς τρόπους εισαγωγής των δεδομένων. Για τα εξαρτημένα δείγματα, οι δύο μεταβλητές πρέπει να εισαχθούν σε δύο στήλες. Για ανεξάρτητα δείγματα, χρησιμοποιούμε μία μεταβλητή για τις μετρήσεις μας και μία δεύτερη μεταβλητή στην οποία καθορίζουμε την υπο-ομάδα στην οποία ανήκουν οι συμμετέχοντες. Όπως και στον υπολογισμό του στατιστικού κριτηρίου  $\chi^2$  είναι σημαντικό να εισάγουμε τα δεδομένα μας σωστά.

### Το t-τεστ εξαρτημένων δειγμάτων

Το t-τεστ για εξαρτημένα δείγματα εφαρμόζεται στις περιπτώσεις που έχουμε μετρήσει την ίδια μεταβλητή στους ίδιους συμμετέχοντες σε δύο διαφορετικές χρονικές στιγμές (π.χ. πριν την παρέμβαση και μετά την παρέμβαση) ή όταν έχουμε πάρει μετρήσεις από τους ίδιους συμμετέχοντες για δύο μεταβλητές, χρησιμοποιώντας την ίδια κλίμακα μέτρησης. Οι δύο μετρήσεις ή οι δύο μεταβλητές καταχωρίζονται σε ξεχωριστές στήλες όπως φαίνεται στον Πίνακα 5.4.

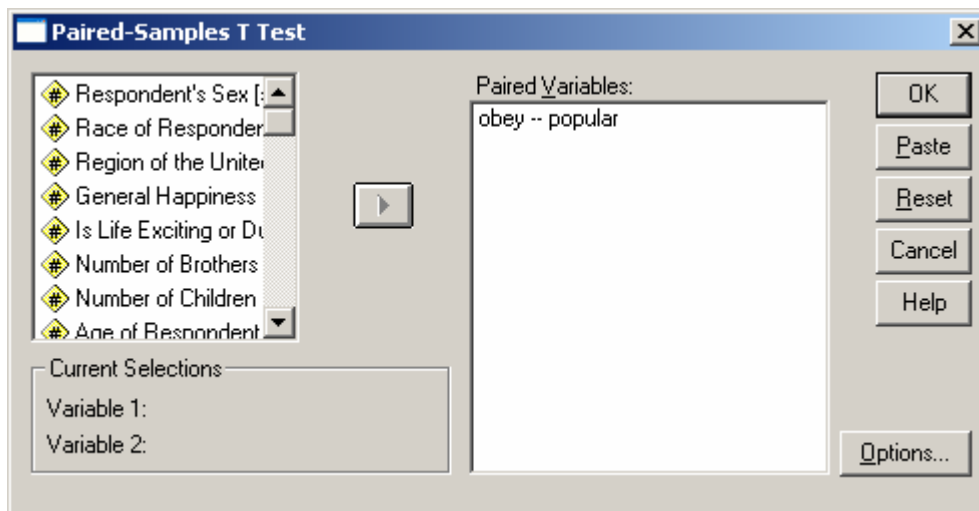
**Πίνακας 5.4 Καταχώρηση δεδομένων για το t-τεστ για εξαρτημένα δείγματα**

Mnths_6	Mnths_24
124	114
94	88
115	102
110	2
116	2
139	2
116	2
110	2
129	2
120	2
105	2
88	2
120	2
120	2
116	2
105	2
...	...
...	...
123	132



Αφού εισαχθούν τα δεδομένα, πρέπει να επιλέξουμε το **[Analyze => Compare Means => Paired-Samples T Test...]**. Εμφανίζεται η οθόνη του Σχήματος 5.4.

**Σχήμα 5.4 Το t-τεστ για εξαρτημένα δείγματα**



Για να επιλέξουμε τις δύο μεταβλητές, κρατάμε το πλήκτρο **[Shift]** πατημένο ενώ χρησιμοποιούμε το ποντίκι για την επιλογή. Η εφαρμογή απαιτεί την επιλογή δύο μεταβλητών. Αφού έχει γίνει αυτό, τις μεταφέρουμε στη λίστα **Paired Variables:**. Με την ολοκλήρωση αυτής της διαδικασίας, πατάμε το **[OK]** και παίρνουμε τα εξής αποτελέσματα στην έξοδο (βλ. Πίνακα 5.5):

**Πίνακας 5.5 Αποτελέσματα του t-τεστ για εξαρτημένα δείγματα**

**Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	To Obey	3,19	982	1,401	,045
	To Be Well Liked or Popular	4,60	982	,756	,024

**Paired Samples Correlations**

		N	Correlation	Sig.
Pair 1	To Obey & To Be Well Liked or Popular	982	-,253	,000

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 To Obey - To Be Well Liked or Popular	-1,405	1,752	,056	-1,515	-1,296	-25,129	981	,000

Ο πρώτος πίνακας μας δίνει περιγραφικά στατιστικά και ο δεύτερος τη συνάφεια για τις δύο μεταβλητές (ή για τις δύο μετρήσεις της ίδιας μεταβλητής). Ο τρίτος πίνακας μας δίνει την τιμή του  $t$ , τους βαθμούς ελευθερίας (στήλη  $df$ ) και το  $p$  (στήλη  $sig.$ ). Η τιμή του  $p$  μας δείχνει ότι στο παράδειγμά μας υπάρχει στατιστικώς σημαντική διαφορά μεταξύ των δύο μεταβλητών.

### Το $t$ -τεστ για ανεξάρτητα δείγματα

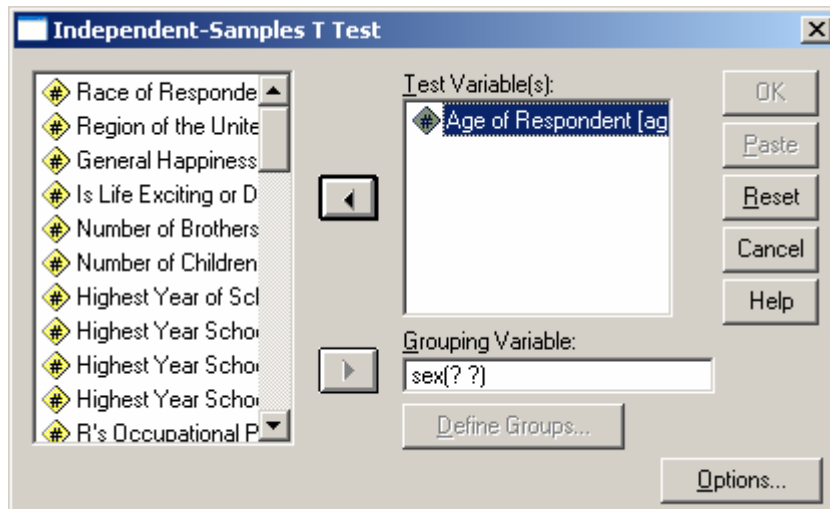
Το  $t$ -τεστ για ανεξάρτητα δείγματα εφαρμόζεται στις περιπτώσεις που θέλουμε να συγκρίνουμε δύο υπο-ομάδες (δείγματα) ως προς μία μεταβλητή (π.χ. την ομάδα των αντρών και την ομάδα των γυναικών ως προς μία μεταβλητή). Χρησιμοποιούμε μία μεταβλητή ομαδοποίησης για να δηλώσουμε την ομάδα στην οποία ανήκουν οι συμμετέχοντες και μία μεταβλητή όπου καταχωρίζουμε τη μέτρηση που μας ενδιαφέρει για κάθε συμμετέχοντα, όπως φαίνεται στον Πίνακα 5.6.

**Πίνακας 5.6 Καταχώρηση δεδομένων για το  $t$ -τεστ ανεξάρτητων δειγμάτων**

Group	Exp_Con
1	96
1	127
1	127
1	119
2	109
2	143
2	104
2	96
...	...
...	...
2	114

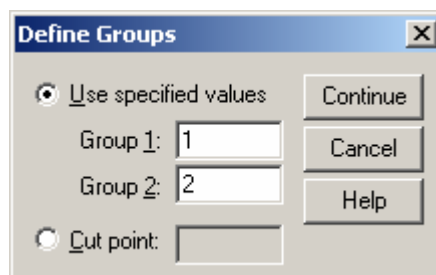
Για να υπολογίσουμε το στατιστικό κριτήριο  $t$ , κάνουμε κλικ στα **[Analyze => Compare Means => Independent-Samples T Test]** και εμφανίζεται η οθόνη του Σχήματος 5.5.

**Σχήμα 5.5 Το t-τεστ για ανεξάρτητα δείγματα**



- Στο κουτί **Test Variable(s):** βάζουμε τη μεταβλητή που περιέχει τις μετρήσεις μας (στο παράδειγμά μας την ηλικία των συμμετεχόντων).
- Στο κουτί **Grouping Variable:** βάζουμε τη μεταβλητή ομαδοποίησης (στο παράδειγμά μας το φύλο των συμμετεχόντων). Όπως βλέπετε στο Σχήμα 5.5 εμφανίζονται δύο λατινικά ερωτηματικά δίπλα από το όνομα της μεταβλητής ομαδοποίησης. Στα ερωτηματικά αυτά πρέπει να αντιστοιχήσουμε τους κωδικούς που αντιστοιχούν στις ομάδες (αν δεν τους θυμόμαστε ανατρέχουμε στην προβολή **[Variable View]**). Επιλέγουμε τη μεταβλητή ομαδοποίησης και κάνουμε κλικ στο πλήκτρο **[Define Groups...]**, με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 5.6.

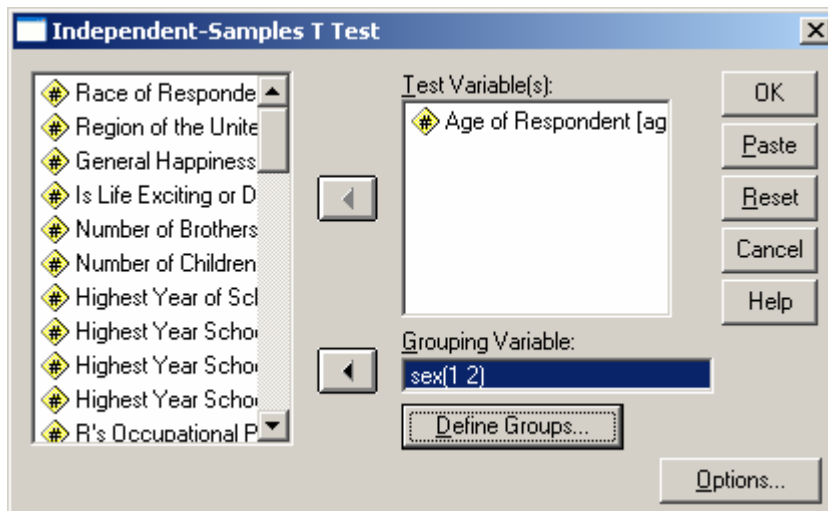
**Σχήμα 5.6 Καθορισμός των ομάδων**



- Στην οθόνη του Σχήματος 5.6 πρέπει να δηλώσουμε πως κωδικοποιήθηκαν οι δύο ομάδες στην μεταβλητή ομαδοποίησης που χρησιμοποιήσαμε. Στο παράδειγμά μας

η μεταβλητή ομαδοποίησης είναι το φύλο και έχουμε κωδικοποιήσει την ομάδα των αντρών με 1 και την ομάδα των γυναικών με 2. Αφού δηλώσουμε τις τιμές, επιλέγουμε το **[Continue]** για να επιστρέψουμε στην αρχική οθόνη (βλ. Σχήμα 5.7, όπου τα λατινικά ερωτηματικά έχουν αντικατασταθεί με τους κωδικούς 1 και 2.

**Σχήμα 5.7 Το t-τεστ για ανεξάρτητα δείγματα μετά τον καθορισμό των ομάδων**



- Πατώντας **[OK]** γίνεται η επεξεργασία των δεδομένων και παίρνουμε στην έξοδο τα ακόλουθα (βλ. Πίνακα 5.7):

**Πίνακας 5.7 Αποτελέσματα του t-τεστ ανεξάρτητων δειγμάτων**

#### Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
Age of Respondent	Male	636	44,18	17,033	,675
	Female	878	46,67	18,288	,617

#### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Age of Respondent	Equal variances assumed	9,697	,002	-2,693	1512	,007	-2,492	,925	-4,307	-.677
	Equal variances not assumed			-2,724	1420,931	,007	-2,492	,915	-4,287	-.698

- Ο πρώτος πίνακας μας δίνει περιγραφικά στοιχεία για τις δύο ομάδες που συγκρίνουμε, δηλαδή μας δίνει ξεχωριστά το μέσο όρο και την τυπική απόκλιση για τους άντρες και τις γυναίκες.
- Ο δεύτερος πίνακας μας δίνει τα αποτελέσματα της ανάλυσης. Οι δύο πρώτες στήλες παρουσιάζουν τα αποτελέσματα του τεστ Levene για την ισότητα των διακυμάνσεων, που είναι και η προϋπόθεση για την εφαρμογή του κριτηρίου t-τεστ. Για να επιτρέπεται η εφαρμογή του τεστ θα πρέπει το Sig. του Levene να υποδηλώνει οι δύο διακυμάνσεις είναι ίσες (δηλαδή να είναι μεγαλύτερο του 0,05). Αν το Levene είναι στατιστικώς σημαντικό (όπως συμβαίνει στο παράδειγμά μας), τότε θα πρέπει να διαβάσουμε τη δεύτερη γραμμή του πίνακα (equal variances not assumed).
- Ορισμένοι προτιμούν να μην κάνουν χρήση της δεύτερης γραμμής του πίνακα και θεωρούν ότι από τη στιγμή που δεν υπάρχει ίσες διακυμάνσεις θα πρέπει να χρησιμοποιηθεί το μη παραμετρικό ισοδύναμο του t-τεστ (βλ. Mann-Whitney U). Άλλοι πάλι, αγνοούν τελείως το τεστ Levene και θεωρούν ότι σύμφωνα με το θεώρημα του κεντρικού ορίου, αν για κάθε ομάδα έχουμε N μεγαλύτερο του 30 (δηλαδή τουλάχιστον 30 άντρες και τουλάχιστον 30 γυναίκες στο παράδειγμά μας), μπορούμε να εφαρμόσουμε το t-τεστ, διαφορετικά θα πρέπει να εφαρμόσουμε το μη παραμετρικό ισοδύναμο.
- Κάνοντας χρήση της δεύτερης γραμμής του πίνακα βλέπουμε ότι η διαφορά στο μέσο όρο της ηλικίας των αντρών και των γυναικών είναι στατιστικώς σημαντική ( $p=0,007$  άρα  $p<0,01$ ).

### 5.3 Ανάλυση Διακύμανσης Μονής Κατεύθυνσης

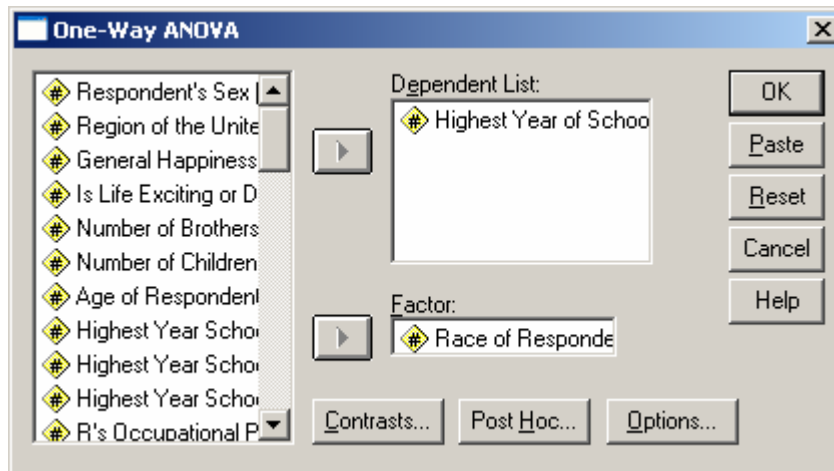
Η ανάλυση διακύμανσης μονής κατεύθυνσης εφαρμόζεται στις περιπτώσεις που θέλουμε να συγκρίνουμε περισσότερες από δύο υπο-ομάδες (δείγματα) ως προς μία μεταβλητή (π.χ. τις τρεις ομάδες που προκύπτουν από την φυλή των συμμετεχόντων ως προς μία μεταβλητή). Χρησιμοποιούμε μία μεταβλητή ομαδοποίησης για να δηλώσουμε την ομάδα στην οποία ανήκουν οι συμμετέχοντες και μία μεταβλητή όπου καταχωρίζουμε τη μέτρηση στη μεταβλητή που μας ενδιαφέρει για κάθε συμμετέχοντα, όπως φαίνεται στον Πίνακα 5.8.

**Πίνακας 5.8** Καταχώρηση δεδομένων για την ανάλυση διακύμανσης μονής κατεύθυνσης

Groups	Scores
1	9
3	8
3	6
...	...
1	7
2	7
3	9
3	6
...	...
...	...
...	...
2	10
2	19
...	...
1	11

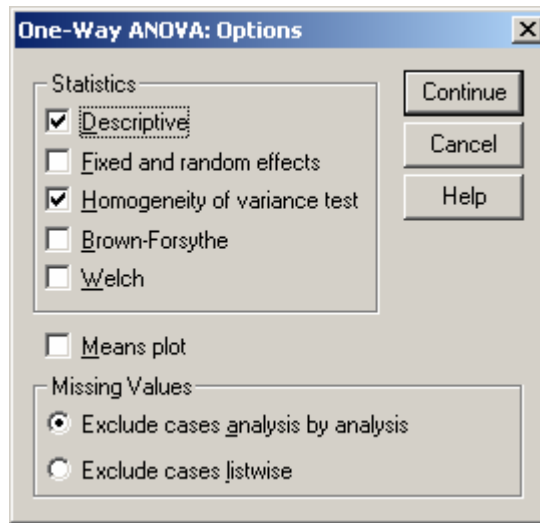
Για να κάνουμε ανάλυση διακύμανσης μονής κατεύθυνσης κάνουμε κλικ στα **[Analyze => Compare Means => One-Way ANOVA]** και εμφανίζεται η οθόνη του Σχήματος 5.8.

Σχήμα 5.8 Ανάλυση διακύμανσης μόνης κατεύθυνσης



- Βάζουμε στο **Dependent list**: τη μεταβλητή που περιλαμβάνει τις μετρήσεις μας (στο παράδειγμα, τα έτη συστηματικής εκπαίδευσης) και στο **Factor**: τη μεταβλητή ομαδοποίησης. Προσέξτε ότι σε αυτήν την περίπτωση δε χρειάζεται να ορίσουμε το πώς έχουμε κωδικοποιήσει τις επιμέρους ομάδες.
- Κάνουμε κλικ στο πλήκτρο **Options...** και εμφανίζεται η οθόνη του Σχήματος 5.9. επιλέγουμε **Descriptive** προκειμένου να πάρουμε περιγραφικά στατιστικά για τις επιμέρους ομάδες και **Homogeneity of variance test** για να πάρουμε το τεστ Levene. Στη συνέχεια επιλέγουμε **Continue** για να επιστρέψουμε στην αρχική οθόνη (Σχήμα 5.8).

Σχήμα 5.9 Ανάλυση διακύμανσης μόνης κατεύθυνσης – επιλογή Options



- Πατήστε το [OK] για να γίνει η ανάλυση. Τα αποτελέσματα θα είναι ως εξής (βλ. Πίνακα 5.9):

Πίνακας 5.9 Αποτελέσματα για την ανάλυση διακύμανσης μόνης κατεύθυνσης

## Descriptives

Highest Year of School Completed

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
White	1262	13,06	2,955	,083	12,89	13,22	0	20
Black	199	11,89	2,677	,190	11,52	12,27	3	20
Other	49	12,47	4,001	,572	11,32	13,62	3	20
Total	1510	12,88	2,984	,077	12,73	13,03	0	20

## Test of Homogeneity of Variances

Highest Year of School Completed

Levene Statistic	df1	df2	Sig.
9,539	2	1507	,000

## ANOVA

Highest Year of School Completed

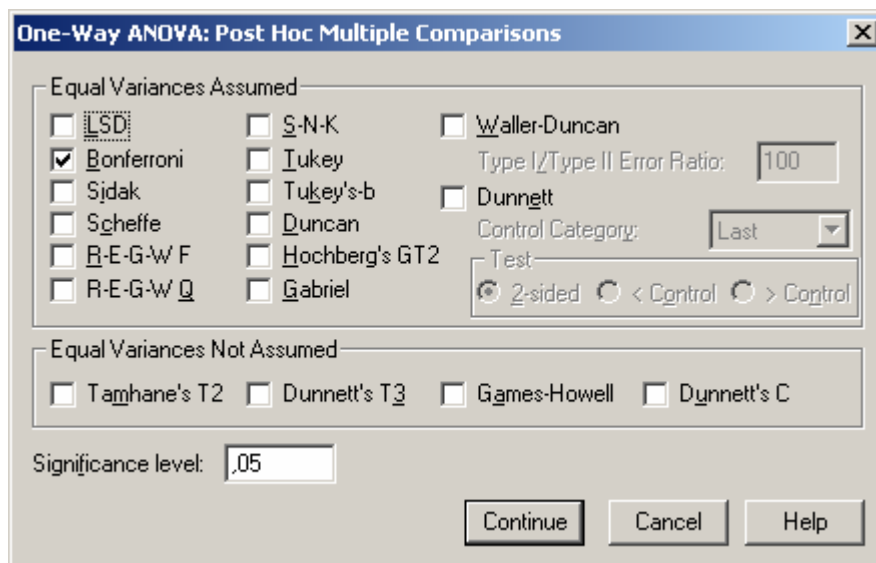
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	240,725	2	120,362	13,746	,000
Within Groups	13195,994	1507	8,756		
Total	13436,719	1509			



- Στον πρώτο πίνακα παρουσιάζονται περιγραφικά στατιστικά για κάθε ομάδα χωριστά (στο παράδειγμά μας για τις τρεις κατηγορίες της φυλής, αλλά και για το σύνολο του δείγματος). Στο δεύτερο πίνακα παρουσιάζονται τα αποτελέσματα για το τεστ Levene. Και σε αυτήν την περίπτωση προϋπόθεση για την εφαρμογή της ανάλυσης διακύμανσης μονής κατεύθυνσης είναι οι διακυμάνσεις να είναι ίσες, κάτι που ισχύει όταν το τεστ Levene είναι στατιστικώς ασήμαντο. Στο παράδειγμά μας η προϋπόθεση της ισότητας των διακυμάνσεων δεν ικανοποιείται και κατά συνέπεια θα πρέπει να χρησιμοποιήσουμε το μη παραμετρικό ισοδύναμο της ανάλυσης διακύμανσης μονής κατεύθυνσης (βλ. παρακάτω). Στον τρίτο πίνακα παρουσιάζονται τα αποτελέσματα της ανάλυσης διακύμανσης μονής κατεύθυνσης.
- Στο παράδειγμά μας είδαμε ότι οι προϋποθέσεις για την εφαρμογή της ανάλυσης διακύμανσης μονής κατεύθυνσης δεν ικανοποιούνται, αλλά υπάρχουν ορισμένοι που και σε αυτήν την περίπτωση αγνοούν το τεστ Levene και βασιζόμενοι στο θεώρημα του κεντρικού ορίου, προχωρούν σε ανάλυση διακύμανσης μονής κατεύθυνσης αν το N είναι μεγαλύτερο από τριάντα για κάθε μία από τις υπό σύγκριση ομάδες.
- Ας υποθέσουμε ότι ακολουθούμε αυτή τη λογική και θεωρούμε ότι η ανάλυσή του παραδείγματός μας είναι έγκυρη. Όπως φαίνεται στην τελευταία στήλη του τρίτου πίνακα, υπάρχει στατιστικώς σημαντική διάφορα σε επίπεδο  $p < 0,001$ . Αυτό σημαίνει ότι όντως οι τρεις ομάδες που προκύπτουν από τη μεταβλητή φυλή διαφέρουν κατά στατιστικώς σημαντικό τρόπο ως προς τα έτη συστηματικής εκπαίδευσης που έχουν λάβει. Αυτή η διαπίστωση δεν είναι αρκετή, καθώς σε αντίθεση με ότι συνέβαινε στο t-τεστ, εδώ συγκρίνουμε περισσότερες από δύο ομάδες και κατά συνέπεια δεν γνωρίζουμε που ακριβώς βρίσκονται οι στατιστικώς σημαντικές διαφορές. Διαφέρουν όλες οι ομάδες μεταξύ τους ή μόνο κάποιες από αυτές;
- Η απάντηση στο ερώτημα αυτό γίνεται με εκ των υστέρων κατά ζεύγη συγκρίσεις (post hoc), δηλαδή με μία διαδικασία που διερευνούμε την ύπαρξη πιθανών διαφορών σε όλους τους δυνατούς συνδυασμούς ζευγών ομάδων. Όπως είδαμε η μεταβλητή φυλή έχει τρεις κατηγορίες: White, Black και Other. Άρα οι πιθανοί συνδυασμοί είναι οι ακόλουθοι:
  - White × Black
  - White × Other

- Black × Other
- Επειδή οι ίδιες ομάδες ή κατηγορίες συμμετέχουν σε περισσότερες από μία αναλύσεις, είναι πιθανό να εκληφθούν ως στατιστικώς σημαντικές διαφορές που δεν είναι, για τυχαίους λόγους. Για το λόγο αυτό κατά τις εκ των υστέρων κατά ζεύγη συγκρίσεις εφαρμόζουμε κάποια «διόρθωση» που λαμβάνει υπόψη τον αριθμό των συγκρίσεων κατά τον υπολογισμό της στατιστικής σημαντικότητας. Το SPSS προσφέρει μεγάλο αριθμό διορθώσεων, κάθε μία από τις οποίες έχει πλεονεκτήματα ανάλογα με τις ανάγκες της ανάλυσής μας. Η απλούστερη διόρθωση που μπορούμε να εφαρμόσουμε είναι η διόρθωση Bonferroni.
- Για να ζητήσουμε από το SPSS κατά ζεύγη συγκρίσεις, κάνουμε κλικ στο πλήκτρο **[Post Hoc]** από την αρχική οθόνη, το οποίο ανοίγει το ακόλουθο παράθυρο διαλόγου (βλ. Σχήμα 5.10):

**Σχήμα 5.10** Ανάλυση διακύμανσης μονής κατεύθυνσης – επιλογή Post Hoc



- Επιλέγουμε τη διόρθωση Bonferroni και πατάμε **[Continue]** για να επιστρέψουμε στην αρχική οθόνη. Δίνοντας **[OK]** εκτελείται η ανάλυση δίνοντάς μας τα ακόλουθα επιπλέον αποτελέσματα (βλ. Πίνακα 5.10).

**Πίνακας 5.10 Αποτελέσματα για τις εκ των υστέρων κατά ζεύγη συγκρίσεις****Multiple Comparisons**

Dependent Variable: Highest Year of School Completed

Bonferroni

(I) Race of Respondent	(J) Race of Respondent	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
White	Black	1,162*	,226	,000	,62	1,70
	Other	,587	,431	,520	-,45	1,62
Black	White	-1,162*	,226	,000	-1,70	-,62
	Other	-,575	,472	,670	-1,71	,56
Other	White	-,587	,431	,520	-1,62	,45
	Black	,575	,472	,670	-,56	1,71

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

- Στην πρώτη στήλη του πίνακα φαίνεται η σύγκριση στην οποία αντιστοιχεί η κάθε σειρά, ενώ στην τέταρτη στήλη μας δίνεται η στατιστική σημαντικότητα. Από την επισκόπηση του Πίνακα 5.10 προκύπτει ότι στατιστικώς σημαντική διαφορά υπάρχει μόνο μεταξύ των κατηγοριών White και Other και όχι για τους υπόλοιπους συνδυασμούς κατηγοριών.

#### 5.4 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων

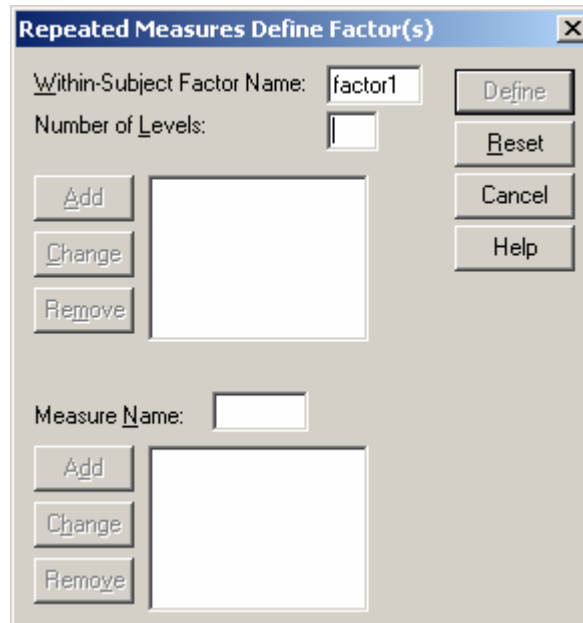
Η ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων εφαρμόζεται στις περιπτώσεις που έχουμε μετρήσει την ίδια μεταβλητή στους ίδιους συμμετέχοντες σε περισσότερες από δύο διαφορετικές χρονικές στιγμές (π.χ. πριν την παρέμβαση, στο μέσο της παρέμβασης και μετά την παρέμβαση) ή όταν έχουμε πάρει μετρήσεις από τους ίδιους συμμετέχοντες για περισσότερες από δύο μεταβλητές, χρησιμοποιώντας την ίδια κλίμακα μέτρησης. Οι μετρήσεις ή οι μεταβλητές καταχωρίζονται σε ξεχωριστές στήλες όπως φαίνεται στον Πίνακα 5.10.

**Πίνακας 5.10 Καταχώρηση δεδομένων για την ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων**

Mnths_6	Mnths_12	Mnths_24
124	45	114
94	12	88
115	32	102
110	56	2
116	87	2
139	45	2
116	34	2
110	65	2
129	65	2
120	76	2
105	89	2
88	08	2
120	66	2
120	88	2
116	32	2
105	23	2
...	...	...
...	...	...
123	343	132

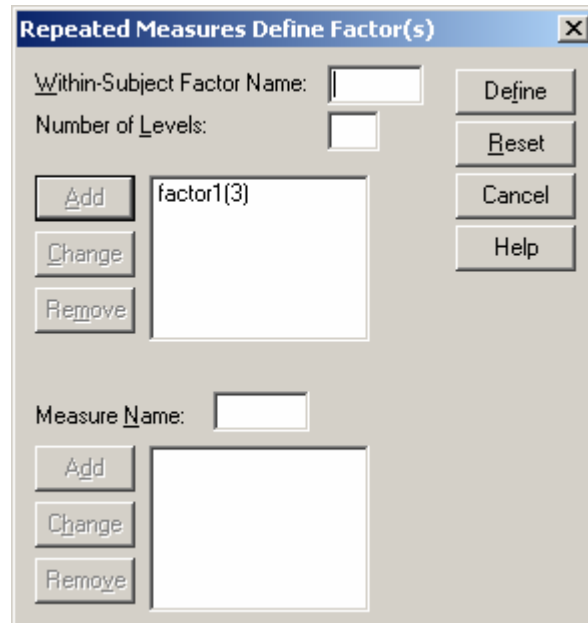
Για να κάνουμε ανάλυση διακύμανσης μονής κατεύθυνσης κάνουμε κλικ στα **[Analyze => General Linear Model => Repeated Measures...]** και εμφανίζεται η οθόνη του Σχήματος 5.10.

**Σχήμα 5.10** Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – αρχική οθόνη



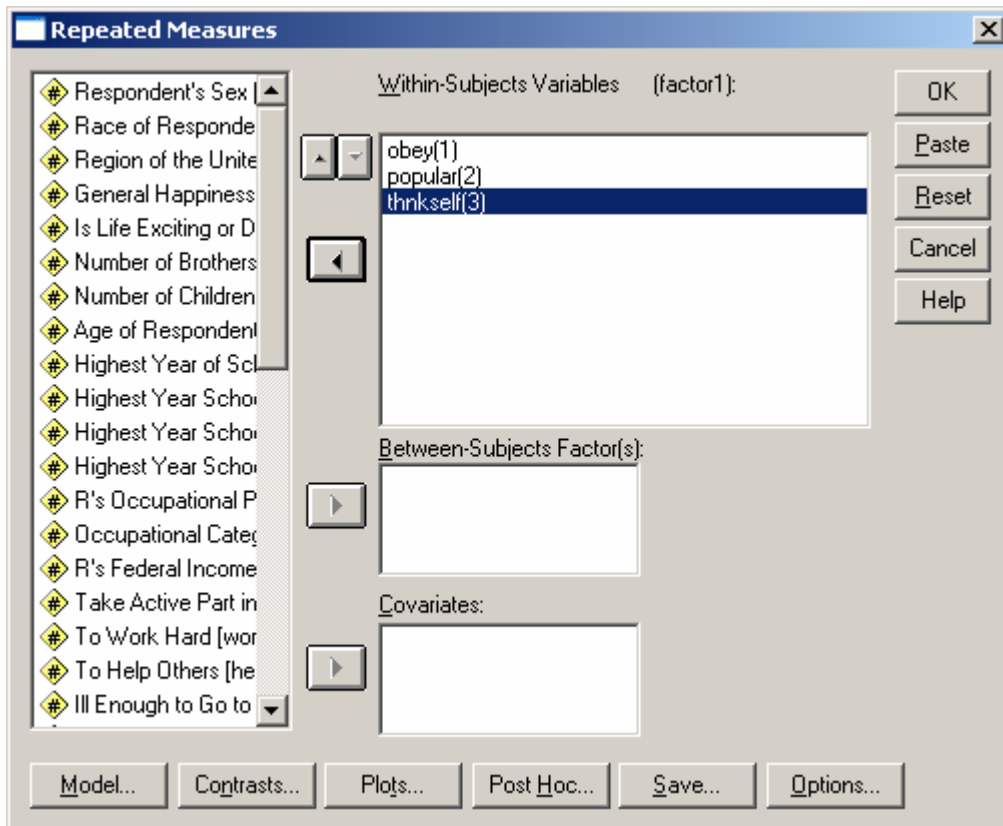
- Στην οθόνη του Πίνακα 5.10 προσθέτουμε τον αριθμό των μεταβλητών ή των μετρήσεων της ίδιας μεταβλητής που θέλουμε να συγκρίνουμε στο πεδίο **Number of Levels** και κάνουμε κλικ στο πλήκτρο **Add**. Στο παράδειγμά μας θα επιλέξουμε τρία επίπεδα και η οθόνη μας θα έχει την ακόλουθη μορφή (βλ. Σχήμα 5.11).

**Σχήμα 5.10** Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – καθορισμός επιπέδων



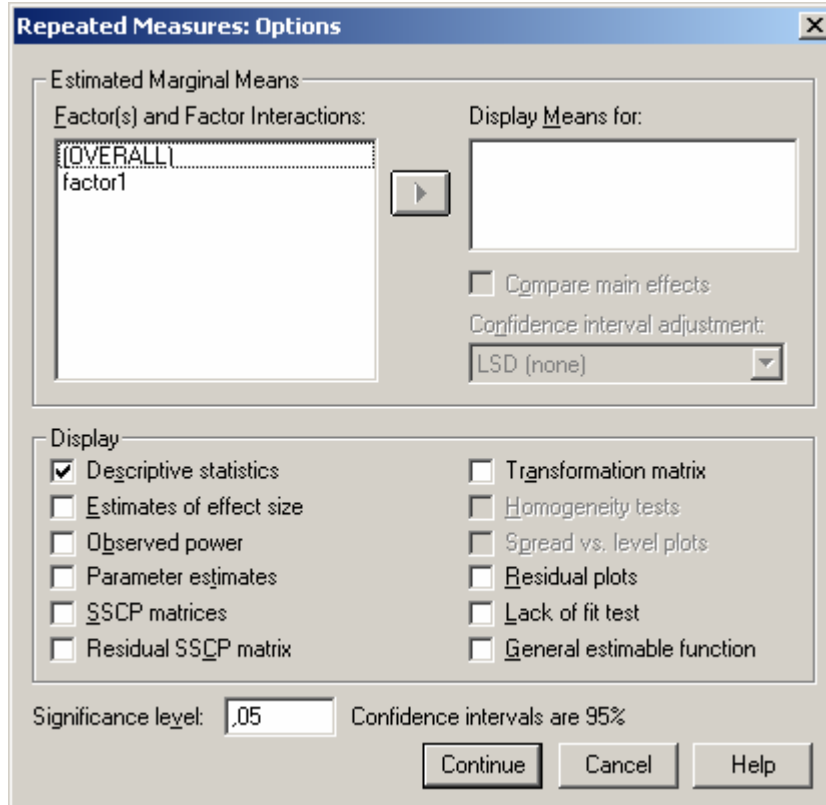
- Στη συνέχεια κάνουμε κλικ στο πλήκτρο **[Define]** και εμφανίζεται η ακόλουθη οθόνη (βλ. Σχήμα 5.12).

Σχήμα 5.12 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων



- Στην οθόνη του σχήματος 5.12 βάζουμε στη λίστα **Within-Subjects Variables** τις μεταβλητές ή τις μετρήσεις της ίδιας μεταβλητής που επιθυμούμε να συγκρίνουμε. Στη συνέχεια κάνουμε κλικ στο πλήκτρο **Options** και εμφανίζεται η ακόλουθη οθόνη (βλ. Σχήμα 5.13).

**Σχήμα 5.12** Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – επιλογή Options



- Στην οθόνη του σχήματος 5.13 επιλέγουμε **Descriptive Statistics** από την επιλογή **Display** και πατάμε [**Continue**] για να επιστρέψουμε στην προηγούμενη οθόνη. Εκεί δίνουμε [**OK**] για να τρέξουμε την ανάλυση. Τα αποτελέσματα που λαμβάνουμε είναι αρκετά εκτεταμένα, για αυτό στον Πίνακα 5.11 παρουσιάζουμε επιλεκτικά ένα μέρος τους.

**Πίνακας 5.11** Αποτελέσματα για την Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων

Descriptive Statistics			
	Mean	Std. Deviation	N
To Obey	3,19	1,401	982
To Be Well Liked or Popular	4,60	,756	982
To Think for Oneself	2,03	1,274	982



**Mauchly's Test of Sphericity<sup>b</sup>**

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse e-Geisser	Huynh-Feldt	Lower-bound
factor1	,741	293,346	2	,000	,794	,796	,500

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: Intercept

Within Subjects Design: factor1

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	3248,149	2	1624,075	887,129	,000
	Greenhouse-Geisser	3248,149	1,589	2044,202	887,129	,000
	Huynh-Feldt	3248,149	1,591	2041,430	887,129	,000
	Lower-bound	3248,149	1,000	3248,149	887,129	,000
Error(factor1)	Sphericity Assumed	3591,851	1962	1,831		
	Greenhouse-Geisser	3591,851	1558,767	2,304		
	Huynh-Feldt	3591,851	1560,884	2,301		
	Lower-bound	3591,851	981,000	3,661		

- Στον πρώτο πίνακα παρουσιάζονται περιγραφικά στατιστικά για τις μεταβλητές ή τις μετρήσεις που συμμετέχουν στην ανάλυση. Στο δεύτερο πίνακα παρουσιάζονται τα στοιχεία για τον έλεγχο της σφαιρικότητας. Όταν η προϋπόθεση της σφαιρικότητας παραβιάζεται (όταν δηλαδή το  $p$  είναι στατιστικώς σημαντικό), τότε στον επόμενο πίνακα κοιτάμε τη σειρά Greenhouse-Geisser και όχι τη σειρά Sphericity Assumed.
- Στον τρίτο πίνακα παρουσιάζονται τα αποτελέσματα της ανάλυσης διακύμανσης. Επειδή δεν πληρείται το κριτήριο της σφαιρικότητας κοιτάμε τη δεύτερη γραμμή για τον factor 1 και διαπιστώνουμε ότι οι διαφορές που παρατηρούμε μεταξύ των τριών μεταβλητών είναι στατιστικώς σημαντικές σε επίπεδο  $p < 0,001$ .
- Επειδή συγκρίνουμε περισσότερες από δύο μεταβλητές ή μετρήσεις ανακύπτει πάλι το πρόβλημα του εντοπισμού του που ακριβώς υπάρχουν στατιστικώς σημαντικές διαφορές. Η λύση είναι και πάλι να κάνουμε εκ των υστέρων κατά ζεύγη συγκρίσεις,

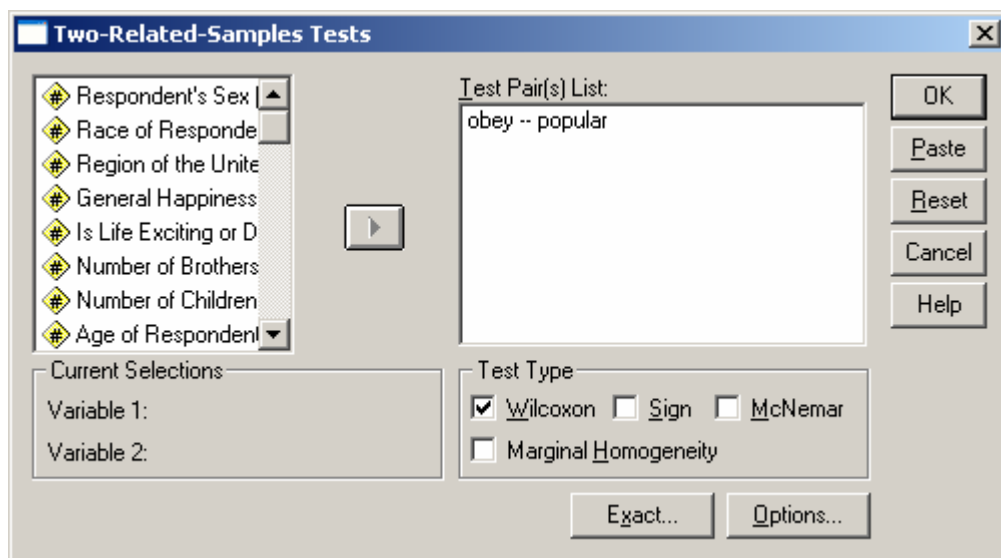
αλλά σε αυτήν την περίπτωση το SPSS δεν προσφέρει κάποια έτοιμη εντολή. Μπορούμε όμως να κάνουμε εύκολα τις εκ των υστέρων συγκρίσεις χρησιμοποιώντας όσα περιγράφονται για το t-τεστ για εξαρτημένα δείγματα για όλους τους πιθανούς συνδυασμούς των μεταβλητών ή των μετρήσεών μας. Η διόρθωση Bonferroni γίνεται εύκολα με το χέρι αν διαιρέσουμε το επίπεδο στατιστικής σημαντικότητας με τον αριθμό των συγκρίσεων που θα κάνουμε. Στο παράδειγμά μας χρησιμοποιούμε τρεις μεταβλητές, άρα έχουμε τρεις συγκρίσεις (η πρώτη με τη δεύτερη, η πρώτη με την τρίτη, η δεύτερη με την τρίτη). Έτσι το επίπεδο στατιστικής σημαντικότητας που πρέπει να χρησιμοποιήσουμε είναι  $0,05 / 3 = 0,17$  περίπου. Αυτό σημαίνει ότι θεωρούμε στατιστικώς σημαντικές τις εκ των υστέρων συγκρίσεις όπου  $p < 0,017$ .

## 5.5 Wilcoxon

Το τεστ Wilcoxon είναι το μη παραμετρικό ισοδύναμο του t-τεστ για εξαρτημένα δείγματα. Εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχουμε καταχωρίσει τα δεδομένα μας με το ίδιο τρόπο.

- Για να εκτελέσουμε το τεστ επιλέγουμε **[Analyze=>Nonparametric Tests=>2 Related Samples]** και εμφανίζεται η οθόνη του Σχήματος 5.13:

Σχήμα 5.13 Τεστ Wilcoxon



- Για να επιλέξουμε τις δύο μεταβλητές, κρατάμε το πλήκτρο **[Shift]** πατημένο ενώ χρησιμοποιούμε το ποντίκι για την επιλογή. Η εφαρμογή απαιτεί την επιλογή δύο μεταβλητών. Αφού έχει γίνει αυτό, τις μεταφέρουμε στη λίστα **Test Pair(s) List**. Με την ολοκλήρωση αυτής της διαδικασίας, πατάμε το **[OK]** και παίρνουμε τα εξής αποτελέσματα στην έξοδο (βλ. Πίνακα 5.12):

Πίνακας 5.12 Αποτελέσματα για το Τεστ Wilcoxon

		Ranks		
		N	Mean Rank	Sum of Ranks
To Be Well Liked or Popular - To Obey	Negative Ranks	205 <sup>a</sup>	337,99	69288,00
	Positive Ranks	777 <sup>b</sup>	532,00	413365,00
	Ties	0 <sup>c</sup>		
	Total	982		

a. To Be Well Liked or Popular < To Obey

b. To Be Well Liked or Popular > To Obey

c. To Be Well Liked or Popular = To Obey

Test Statistics<sup>b</sup>

	To Be Well Liked or Popular - To Obey
Z	-19,742 <sup>a</sup>
Asymp. Sig. (2-tailed)	,000

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

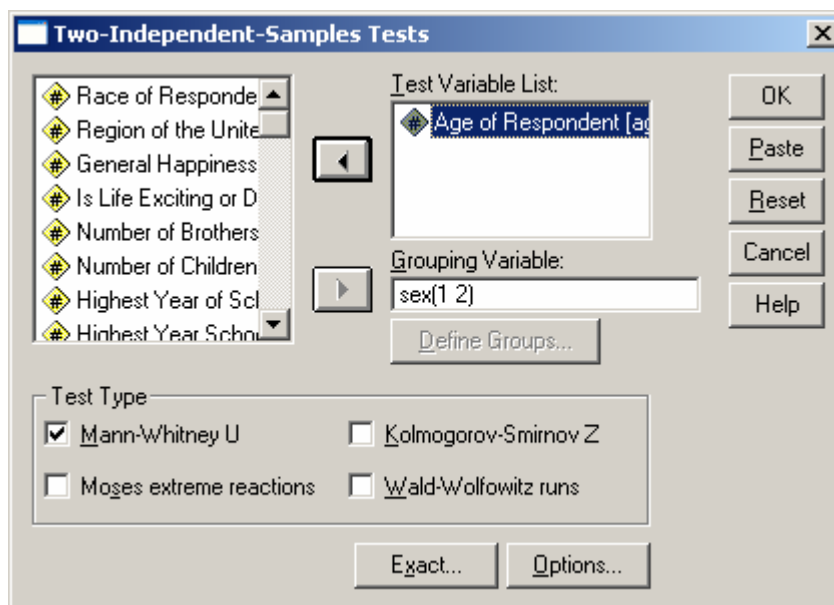
- Στον πρώτο πίνακα παίρνουμε περιγραφικά στατιστικά για τις δύο μεταβλητές ή για τις δύο μετρήσεις της ίδιας μεταβλητής. Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα μας δίνεται η τιμή του κριτηρίου και το p. Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική σε επίπεδο  $p < 0,001$ .
- Ορισμένοι προτιμούν να δίνουν ως περιγραφικά στατιστικά για το τεστ Wilcoxon τις διαμέσους και το εύρος για κάθε μεταβλητή ή για κάθε μέτρηση της ίδιας μεταβλητής. Το SPSS δεν υπολογίζει αυτόματα τα στατιστικά αυτά, αλλά μπορούμε να τα πάρουμε ακολουθώντας τα όσα περιγράφονται στο κεφάλαιο για τους δείκτες κεντρικής τάσης και διασποράς.

## 5.6 Mann-Whitney U

Το τεστ Mann-Whitney U είναι το μη παραμετρικό ισοδύναμο του t-τεστ ανεξάρτητων ομάδων. Εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχουμε καταχωρίσει τα δεδομένα μας με το ίδιο τρόπο.

- Για να εκτελέσουμε το τεστ επιλέγουμε **[Analyze=>Nonparametric Tests=>2 Independent Samples]** και εμφανίζεται η οθόνη του Σχήματος 5.14:

**Σχήμα 5.14 Τεστ Mann-Whitney U**



- Στο κουτί **Test Variable List**: βάζουμε τη μεταβλητή που περιέχει τις μετρήσεις μας (στο παράδειγμά μας την ηλικία των συμμετεχόντων).
- Στο κουτί **Grouping Variable**: βάζουμε τη μεταβλητή ομαδοποίησης (στο παράδειγμά μας το φύλο των συμμετεχόντων) και ακολουθούμε τη διαδικασία που περιγράφηκε στο t-τεστ ανεξάρτητων ομάδων προκειμένου να ορίσουμε τις ομάδες με τη χρήση του πλήκτρου **[Define Groups]**.
- Στη συνέχεια δίνουμε **[OK]** και παίρνουμε στην έξοδο τα αποτελέσματα που παρουσιάζονται στον Πίνακα 5.13.

Πίνακας 5.13 Αποτελέσματα για το Τεστ Mann-Whitney U

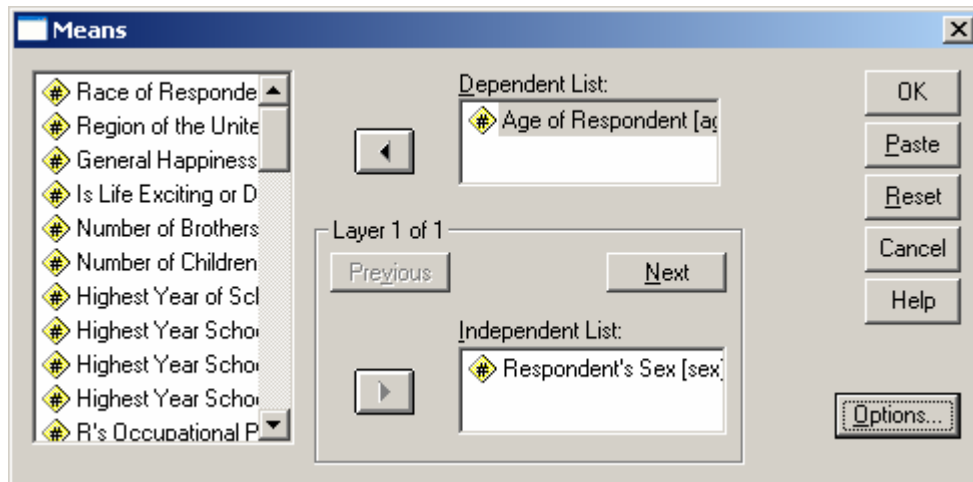
	Respondent's Sex	N	Mean Rank	Sum of Ranks
Age of Respondent	Male	636	726,92	462319,50
	Female	878	779,65	684535,50
	Total	1514		

	Age of Respondent
Mann-Whitney U	259753,500
Wilcoxon W	462319,500
Z	-2,317
Asymp. Sig. (2-tailed)	,021

a. Grouping Variable: Respondent's Sex

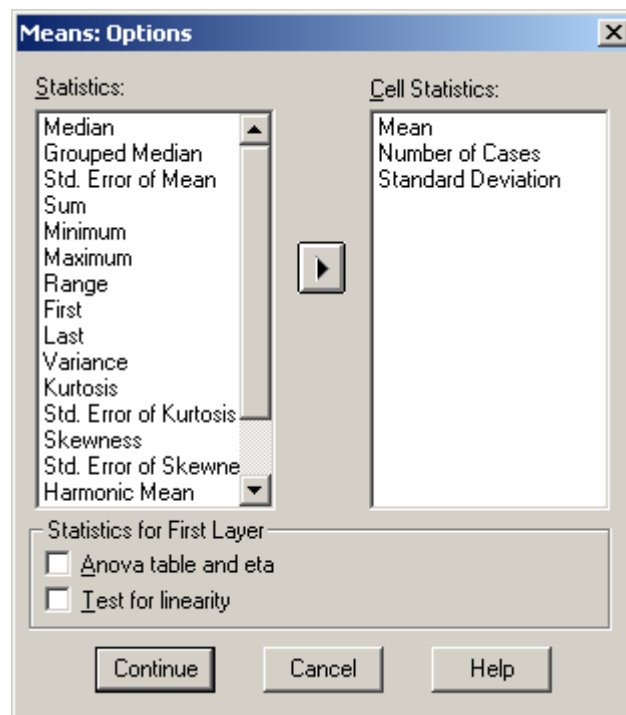
- Στον πρώτο πίνακα παίρνουμε περιγραφικά στατιστικά για τις δύο ομάδες. Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα μας δίνεται η τιμή του κριτηρίου (πρώτη γραμμή) και το p. Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική σε επίπεδο  $p < 0,05$ .
- Ορισμένοι προτιμούν να δίνουν ως περιγραφικά στατιστικά για το τεστ Mann-Whitney U τις διαμέσους και το εύρος για κάθε ομάδα. Το SPSS δεν υπολογίζει αυτόματα τα στατιστικά αυτά, αλλά μπορούμε να τα πάρουμε χρησιμοποιώντας την εντολή [Means].
- Για να χρησιμοποιήσουμε την εντολή αυτή επιλέγουμε **[Analyze=>Compare Means=>Means]** και εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 5.15:

Σχήμα 5.15 Η εντολή Means



- Βάζουμε τη μεταβλητή της μέτρησής μας λίστα **Dependent List** και τη μεταβλητή ομαδοποίησης στη λίστα **Independent List** και στη συνέχεια κάνουμε κλικ στο πλήκτρο **[Options]** με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 5.16.

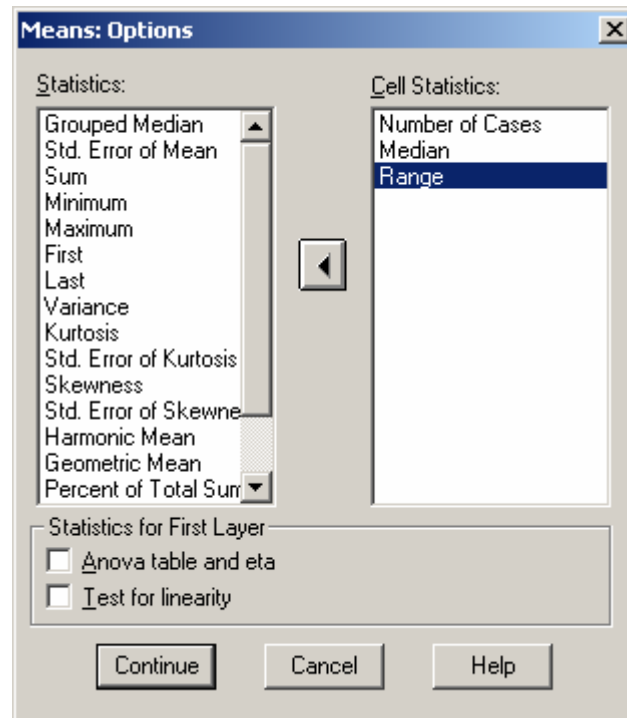
Σχήμα 5.15 Η εντολή Means – επιλογή Options



- Από τη λίστα **Cell Statistics** αφαιρούμε τις επιλογές **Mean** και **Standard Deviation** και προσθέτουμε τις επιλογές **Median** και **Range**. Για να το επιτύχουμε αυτό κάνουμε κλικ στην κάθε επιλογή και στη συνέχεια κάνουμε κλικ στο βελάκι μεταξύ

των δύο λιστών. Η οθόνη μας θα πρέπει να έχει τη μορφή που φαίνεται στο Σχήμα 5.16.

**Σχήμα 5.15 Η εντολή Means – επιλογή Options (II)**



- Στη συνέχεια κάνουμε κλικ στο πλήκτρο **[Continue]** και επιστρέφουμε στην αρχική οθόνη, όπου δίνουμε **[OK]** και λαμβάνουμε στην έξοδο τα όσα παρουσιάζονται στον Πίνακα 5.14:

**Πίνακας 5.14 Η εντολή Means – αποτελέσματα**

**Report**

Age of Respondent

Respondent's Sex	N	Median	Range
Male	636	41,00	71
Female	878	42,00	71
Total	1514	41,00	71

- Βλέπουμε ότι πήραμε τη διάμεσο και το εύρος για κάθε μία από τις κατηγορίες της μεταβλητής ομαδοποίησης, αλλά και για το σύνολο του δείγματος.

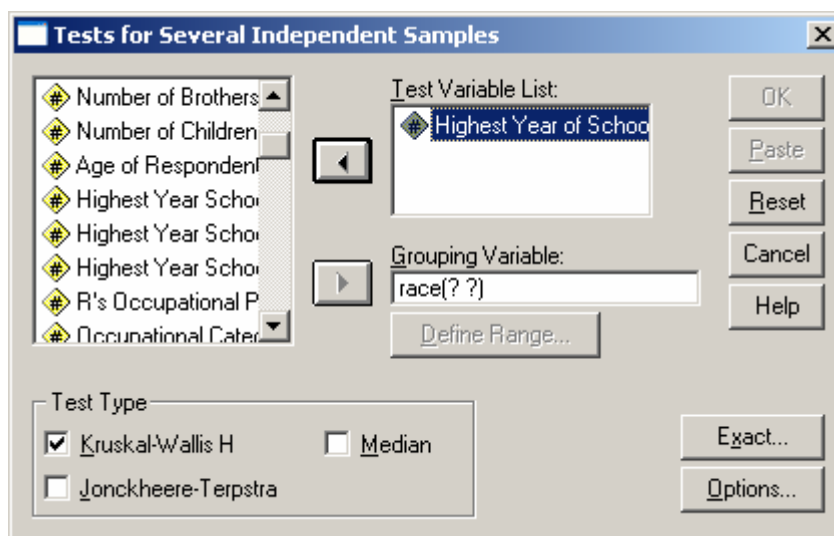


## 5.7 Kruskal-Wallis H

Το τεστ Kruskal-Wallis H είναι το μη παραμετρικό ισοδύναμο της ανάλυσης διακύμανσης μονής κατεύθυνσης. Εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχουμε καταχωρίσει τα δεδομένα μας με το ίδιο τρόπο.

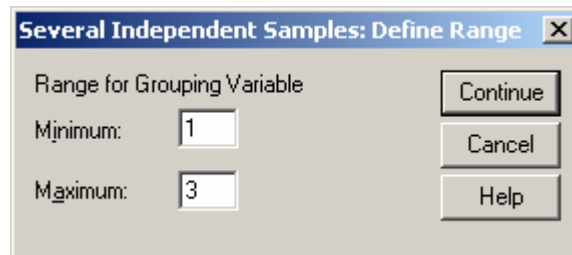
Για να εκτελέσουμε το τεστ επιλέγουμε **[Analyze=>Nonparametric Tests=>Kruskal-Wallis H]** και εμφανίζεται η οθόνη του Σχήματος 5.16:

**Σχήμα 5.16 Kruskal-Wallis H**



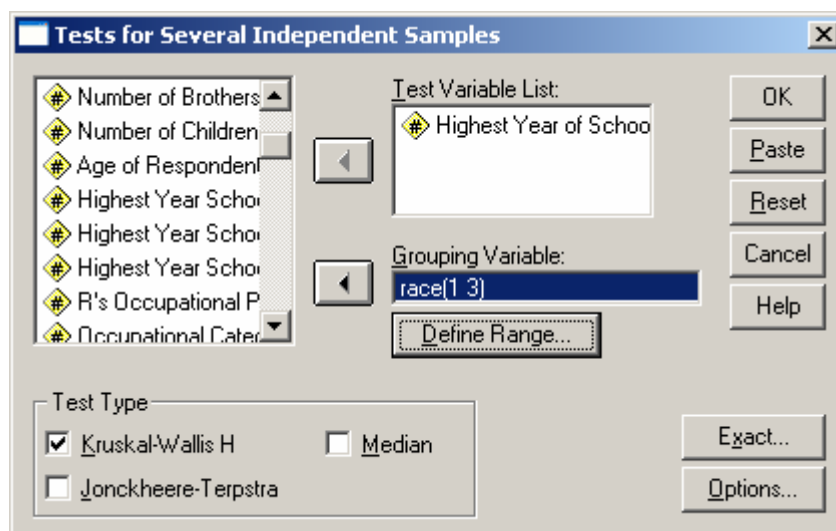
- Στο κουτί **Test Variable List**: βάζουμε τη μεταβλητή που περιέχει τις μετρήσεις μας (στο παράδειγμά μας τα έτη συστηματικής εκπαίδευσης).
- Στο κουτί **Grouping Variable**: βάζουμε τη μεταβλητή ομαδοποίησης (στο παράδειγμά μας τη φυλή των συμμετεχόντων). Όπως βλέπετε στο Σχήμα 5.16 εμφανίζονται δύο λατινικά ερωτηματικά δίπλα από το όνομα της μεταβλητής ομαδοποίησης. Στα ερωτηματικά αυτά πρέπει να αντιστοιχήσουμε τους κωδικούς που αντιστοιχούν στις ομάδες, όπως κάναμε προηγουμένως για το t-τεστ ανεξάρτητων δειγμάτων. Επιλέγουμε τη μεταβλητή ομαδοποίησης και κάνουμε κλικ στο πλήκτρο **[Define Groups...]**, με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 5.17.

Σχήμα 5.17 Kruskal-Wallis H - Καθορισμός των ομάδων



- Στην οθόνη του Σχήματος 5.17 πρέπει να δηλώσουμε πως κωδικοποιήθηκαν οι ομάδες στην μεταβλητή ομαδοποίησης που χρησιμοποιήσαμε. Προσοχή στο γεγονός ότι η οθόνη αυτή είναι διαφορετική από αυτή που χρησιμοποιήσαμε στο t-τεστ ανεξάρτητων δειγμάτων και στο Mann-Whitney U καθώς εδώ καλούμαστε να βάλουμε το εύρος της κωδικοποίησης, δηλαδή να ορίσουμε το μικρότερο και το μεγαλύτερο κωδικό που αντιστοιχούν στις ομάδες μας (στο παράδειγμά μας έχουμε τρεις φυλές κωδικοποιημένες με 1, 2 και 3 και για αυτό τοποθετούμε 1 στο **Minimum** και 3 στο **Maximum**).
- Επιλέγουμε το **[Continue]** για να επιστρέψουμε στην αρχική οθόνη (βλ. Σχήμα 5.18), όπου τα λατινικά ερωτηματικά έχουν αντικατασταθεί με το εύρος κωδικών 1 και 3.

Σχήμα 5.18 Kruskal-Wallis H μετά τον καθορισμό των ομάδων



- Πατώντας **[OK]** γίνεται η επεξεργασία των δεδομένων και παίρνουμε στην έξοδο τα ακόλουθα (βλ. Πίνακα 5.15):

**Πίνακας 5.15 Αποτελέσματα του Kruskal-Wallis H**

Ranks			
	Race of Respondent	N	Mean Rank
Highest Year of School Completed	White	1262	779,33
	Black	199	616,52
	Other	49	706,22
	Total	1510	

**Test Statistics<sup>a,b</sup>**

	Highest Year of School Completed
Chi-Square	25,460
df	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: Race of Respondent

- Στον πρώτο πίνακα παίρνουμε περιγραφικά στατιστικά για τις επιμέρους ομάδες. Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα μας δίνεται η τιμή του κριτηρίου και το p. Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική σε επίπεδο  $p < 0,001$ .
- Ορισμένοι προτιμούν να δίνουν ως περιγραφικά στατιστικά για το τεστ Mann-Whitney U τις διάμεσους και το εύρος για κάθε ομάδα. Το SPSS δεν υπολογίζει αυτόματα τα στατιστικά αυτά, αλλά μπορούμε να τα πάρουμε χρησιμοποιώντας την εντολή **Means**, ακολουθώντας τη διαδικασία που περιγράφηκε για το τεστ Mann-Whitney U.
- Επειδή συγκρίνουμε περισσότερες από δύο επιμέρους ομάδες ανακύπτει πάλι το πρόβλημα του εντοπισμού του πού ακριβώς υπάρχουν στατιστικώς σημαντικές διαφορές. Η λύση είναι και πάλι να κάνουμε εκ των υστέρων κατά ζεύγη συγκρίσεις, αλλά σε αυτήν την περίπτωση το SPSS δεν προσφέρει κάποια έτοιμη εντολή. Μπορούμε όμως να κάνουμε εύκολα τις εκ των υστέρων συγκρίσεις χρησιμοποιώντας όσα περιγράφονται για το Mann-Whitney U για όλους τους πιθανούς συνδυασμούς των επιμέρους ομάδων. Η διόρθωση Bonferroni γίνεται εύκολα με το χέρι αν διαιρέσουμε το επίπεδο στατιστικής σημαντικότητας με τον αριθμό των συγκρίσεων που θα κάνουμε. Στο παράδειγμά μας χρησιμοποιούμε

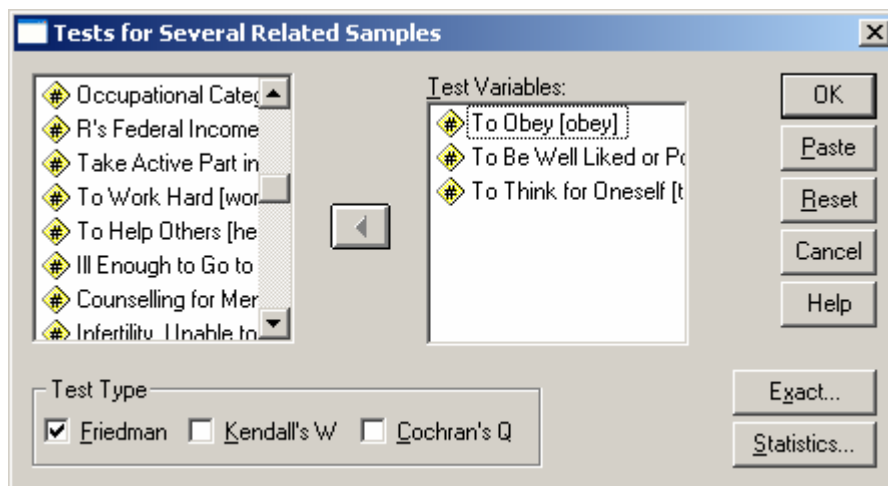
τρεις επιμέρους ομάδες, άρα έχουμε τρεις συγκρίσεις (η πρώτη με τη δεύτερη, η πρώτη με την τρίτη, η δεύτερη με την τρίτη). Έτσι το επίπεδο στατιστικής σημαντικότητας που πρέπει να χρησιμοποιήσουμε είναι  $0,05 / 3 = 0,17$  περίπου. Αυτό σημαίνει ότι θεωρούμε στατιστικώς σημαντικές τις εκ των υστέρων συγκρίσεις όπου  $p < 0,017$ .

## 5.9 Friedman

Το τεστ Friedman είναι το μη παραμετρικό ισοδύναμο για την ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων. Εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχουμε καταχωρίσει τα δεδομένα μας με τον ίδιο τρόπο.

- Για να εκτελέσουμε το τεστ επιλέγουμε **[Analyze=>Nonparametric Tests=>K Related Samples]** και εμφανίζεται η οθόνη του Σχήματος 5.19:

**Σχήμα 5.19 Τεστ Friedman**



- Επιλέγουμε τις μεταβλητές ή τις μετρήσεις που θέλουμε να αναλύσουμε και τις μεταφέρουμε στη λίστα **Test Variables:**. Με την ολοκλήρωση αυτής της διαδικασίας, πατάμε το **[OK]** και παίρνουμε τα εξής αποτελέσματα στην έξοδο (βλ. Πίνακα 5.16):

**Πίνακας 5.16 Αποτελέσματα για το Τεστ Friedman**

Ranks	
	Mean Rank
To Obey	1,89
To Be Well Liked or Popular	2,72
To Think for Oneself	1,39

Test Statistics <sup>a</sup>	
N	982
Chi-Square	884,275
df	2
Asymp. Sig.	,000

a. Friedman Test

- Στον πρώτο πίνακα παίρνουμε περιγραφικά στατιστικά για τις μεταβλητές ή για τις μετρήσεις της ίδιας μεταβλητής. Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα μας δίνεται ο αριθμός των συμμετεχόντων (N), η τιμή του κριτηρίου (Chi-Square), οι βαθμοί ελευθερίας (df) και το p (Asymp. Sig.). Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική σε επίπεδο  $p < 0,001$ .
- Ορισμένοι προτιμούν να δίνουν ως περιγραφικά στατιστικά για το τεστ Wilcoxon τις διαμέσους και το εύρος για κάθε μεταβλητή ή για κάθε μέτρηση της ίδιας μεταβλητής. Το SPSS δεν υπολογίζει αυτόματα τα στατιστικά αυτά, αλλά μπορούμε να τα πάρουμε ακολουθώντας τα όσα περιγράφονται στο κεφάλαιο για τους δείκτες κεντρικής τάσης και διασποράς.

### 5.10 Συνάφεια Pearson r και Spearman Rho

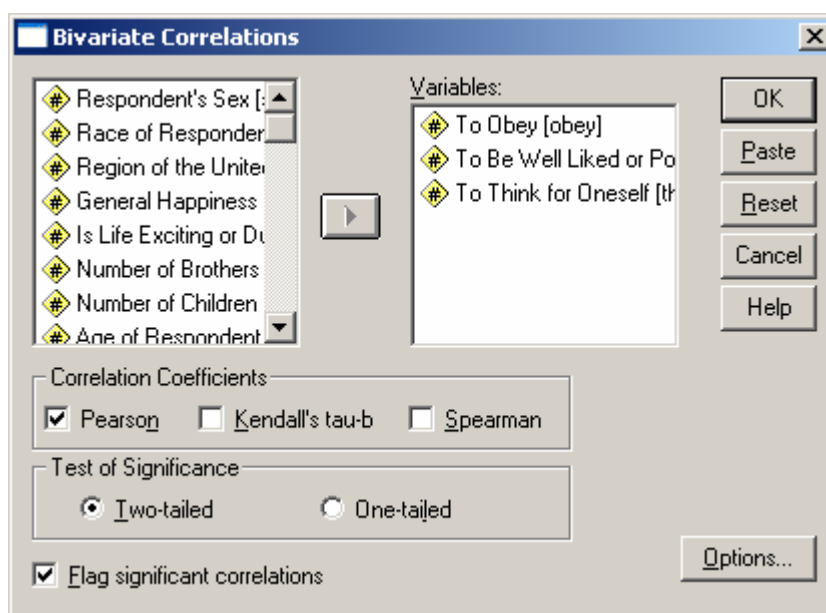
Προκειμένου να υπολογίσουμε συνάφειες με τον Pearson r ή με το μη παραμετρικό ισοδύναμό του Spearman Rho. Οι μεταβλητές καταχωρίζονται σε ξεχωριστές στήλες όπως φαίνεται στον Πίνακα 5.17.

**Πίνακας 5.17 Καταχώρηση δεδομένων για τον υπολογισμό συναφειών**

IQ	GPA
102	2.75
108	4.00
109	2.25
118	3.00
79	1.67
88	2.25
...	...
...	...
85	2.50

Επιλέγουμε [**Analyze => Correlate => Bivariate...**] και εμφανίζεται η οθόνη του Σχήματος 5.20. Στη συνέχεια επιλέγουμε και μετακινούμε τις μεταβλητές που θέλουμε να αναλύσουμε στη λίστα **Variables:** και κάνοντας κλικ στο [**OK**] παίρνουμε τα αποτελέσματα της επεξεργασίας.

**Σχήμα 5.20 Συνάφειες – Pearson r**



Τα αποτελέσματα για το δείκτη Pearson  $r$  (προεπιλογή στην οθόνη του Σχήματος 5.20) είναι τα ακόλουθα (βλ. Πίνακα 5.18):

**Πίνακας 5.18 Αποτελέσματα για το δείκτη συνάφειας Pearson  $r$**

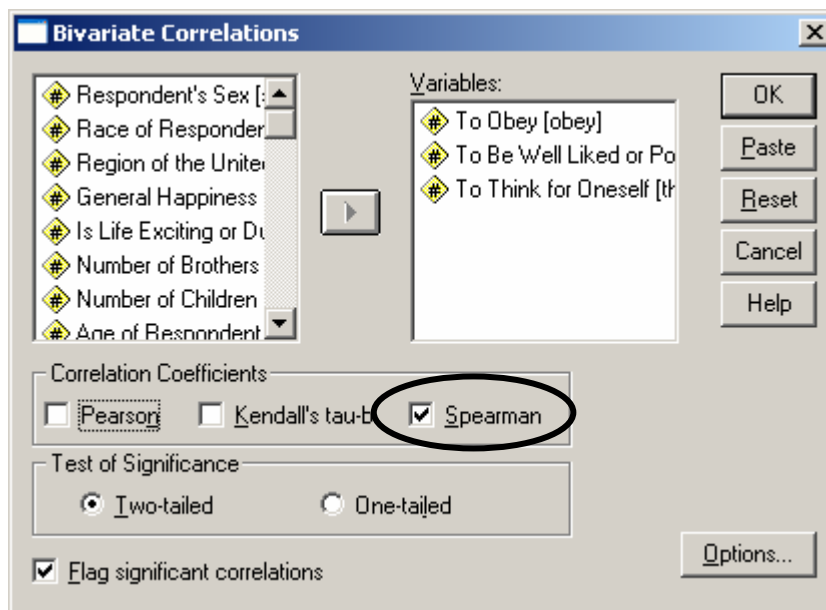
		Correlations		
		To Obey	To Be Well Liked or Popular	To Think for Oneself
To Obey	Pearson Correlation	1	-,253**	-,532**
	Sig. (2-tailed)		,000	,000
	N	982	982	982
To Be Well Liked or Popular	Pearson Correlation	-,253**	1	-,120**
	Sig. (2-tailed)	,000		,000
	N	982	982	982
To Think for Oneself	Pearson Correlation	-,532**	-,120**	1
	Sig. (2-tailed)	,000	,000	
	N	982	982	982

\*\* . Correlation is significant at the 0.01 level (2-tailed).

- Στον πρώτο πίνακα βλέπουμε το δείκτη Pearson  $r$  για όλους τους πιθανούς συνδυασμούς των μεταβλητών που συμμετείχαν στην ανάλυση. Η συνάφεια κάθε μεταβλητής με τον εαυτό της είναι 1. Έτσι σχηματίζεται μία διαγώνιος από 1 στις δύο πλευρές της οποίας επαναλαμβάνονται τα ίδια αποτελέσματα.
- Στον στην πρώτη γραμμή κάθε κελιού μας δίνεται ο δείκτης Pearson  $r$ , στη δεύτερη γραμμή το  $p$  και στην τρίτη ο αριθμός των συμμετεχόντων για τους οποίους υπήρχαν στοιχεία και για τις δύο μεταβλητές.
- Με αστεράκια σημειώνονται οι στατιστικώς σημαντικές συνάφειες (βλ. την υποσημείωση του πίνακα).
- Στην περίπτωση των συναφειών δεν υπολογίζουμε κάποιου άλλου είδους περιγραφικά στατιστικά, καθώς ο δείκτης συνάφειας μας δείχνει το μέγεθος και την κατεύθυνση της συνάφειας.

Στην περίπτωση που θέλουμε να υπολογίσουμε τον μη παραμετρικό δείκτη συνάφειας Spearman Rho, ακολουθούμε τα ίδια ακριβώς βήματα, αλλά στο πεδίο **Correlation Coefficients** αποεπιλέγουμε τον δείκτη Pearson  $r$  και επιλέγουμε τον δείκτη Spearman Rho, ώστε η οθόνη μας να έχει την μορφή του Σχήματος 5.21.

Σχήμα 5.20 Συνάφειες – Pearson  $r$



Τα αποτελέσματα για το δείκτη Spearman Rho είναι τα ακόλουθα (βλ. Πίνακα 5.19):



**Πίνακας 5.19 Αποτελέσματα για το δείκτη συνάφειας Spearman Rho**

**Correlations**

			To Obey	To Be Well Liked or Popular	To Think for Oneself
Spearman's rho	To Obey	Correlation Coefficient	1,000	-,418**	-,529**
		Sig. (2-tailed)	.	,000	,000
		N	982	982	982
	To Be Well Liked or Popular	Correlation Coefficient	-,418**	1,000	-,038
		Sig. (2-tailed)	,000	.	,229
		N	982	982	982
	To Think for Oneself	Correlation Coefficient	-,529**	-,038	1,000
		Sig. (2-tailed)	,000	,229	.
		N	982	982	982

\*\* . Correlation is significant at the 0.01 level (2-tailed).

- Για την ερμηνεία του πίνακα ισχύουν ακριβώς τα ίδια με την ερμηνεία των αποτελεσμάτων του δείκτη Pearson r.