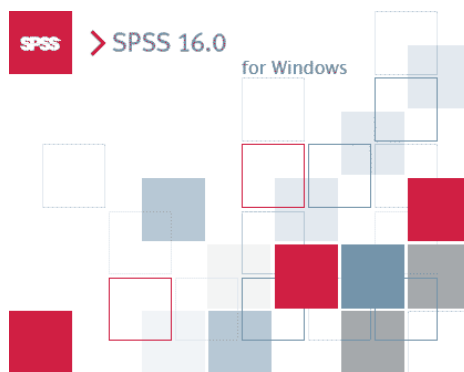


# Σύντομο Εγχειρίδιο SPSS 16.0

**Πέτρος Ρούσσοι & Γιώργος Ευσταθίου**

Πρόγραμμα Ψυχολογίας, Τμήμα ΦΠΨ, ΕΚΠΑ



## Περιεχόμενα

<b>Δυο λόγια εισαγωγικά .....</b>	<b>3</b>
<b>1.0 Το περιβάλλον του SPSS .....</b>	<b>3</b>
<b>2.0 Εισαγωγή και διαχείριση δεδομένων .....</b>	<b>6</b>
<b>3.0 Το παράθυρο Viewer .....</b>	<b>12</b>
<b>4.0 Περιγραφική Στατιστική .....</b>	<b>14</b>
4.1 Κατανομή συχνότητας .....	15
4.2 Δείκτες κεντρικής τάσης και διασποράς .....	17
4.3 Ανάλυση πολλαπλών απαντήσεων .....	19
<b>5.0 Μετασχηματισμοί μεταβλητών .....</b>	<b>24</b>
5.1 Compute .....	24
5.2 Recode .....	25
5.3 Weight Cases .....	27
<b>6.0 Επαγωγική Στατιστική .....</b>	<b>29</b>
6.1 Τεστ $\chi^2$ .....	30
6.2 Το t-τεστ .....	34
6.3 Ανάλυση διακύμανσης μονής κατεύθυνσης (One-way ANOVA) .....	39
6.4 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων (Repeated measures ANOVA) .....	44
6.5 Wilcoxon .....	48
6.6 Mann-Whitney U .....	50
6.7 Kruskal-Wallis H .....	53
6.8 Friedman .....	55
6.9 Δείκτες Συνάφειας – Pearson’s r και Spearman’s Rho .....	56
6.10 Απλή γραμμική παλινδρόμηση .....	60
6.11 Διαδικασία ελέγχου κανονικότητας ερευνητικών δεδομένων .....	63

## Δυο λόγια εισαγωγικά

Το εγχειρίδιο αυτό προσφέρει μια σύντομη περιήγηση στο λογισμικό **SPSS 16.0** για το λειτουργικό σύστημα **Windows Vista**. Αν χρησιμοποιήσετε το SPSS σε διαφορετικό λειτουργικό σύστημα, οι οθόνες που παρουσιάζονται εδώ όπως και οι εντολές μπορεί να διαφέρουν. Είναι σχεδιασμένο για τους φοιτητές μας που παρακολουθούν τα μαθήματά μας στη Μεθοδολογία Έρευνας και τη Στατιστική με χρήση του SPSS. Για πιο διεξοδική περιγραφή του SPSS και των δυνατοτήτων του, θα πρέπει να ανατρέξετε σε πιο αναλυτικά εγχειρίδια (αλλά και στα ηλεκτρονικά εκπαιδευτικά βοηθήματα – tutorials που εγκαταστάθηκαν στον υπολογιστή σας μαζί με το SPSS). Απαντήσεις σε απορίες σας μπορείτε να βρείτε επίσης στο Σύστημα Βοήθειας [**Help**] του SPSS.

Προσέξτε ότι δεν πρόκειται για ένα εγχειρίδιο στατιστικής. Θα πρέπει να το χρησιμοποιήσετε παράλληλα με κάποιο από τα πολλά βιβλία στατιστικής που κυκλοφορούν στο εμπόριο και παρέχουν αναλυτικό εννοιολογικό και θεωρητικό υπόβαθρο των στατιστικών μεθόδων.

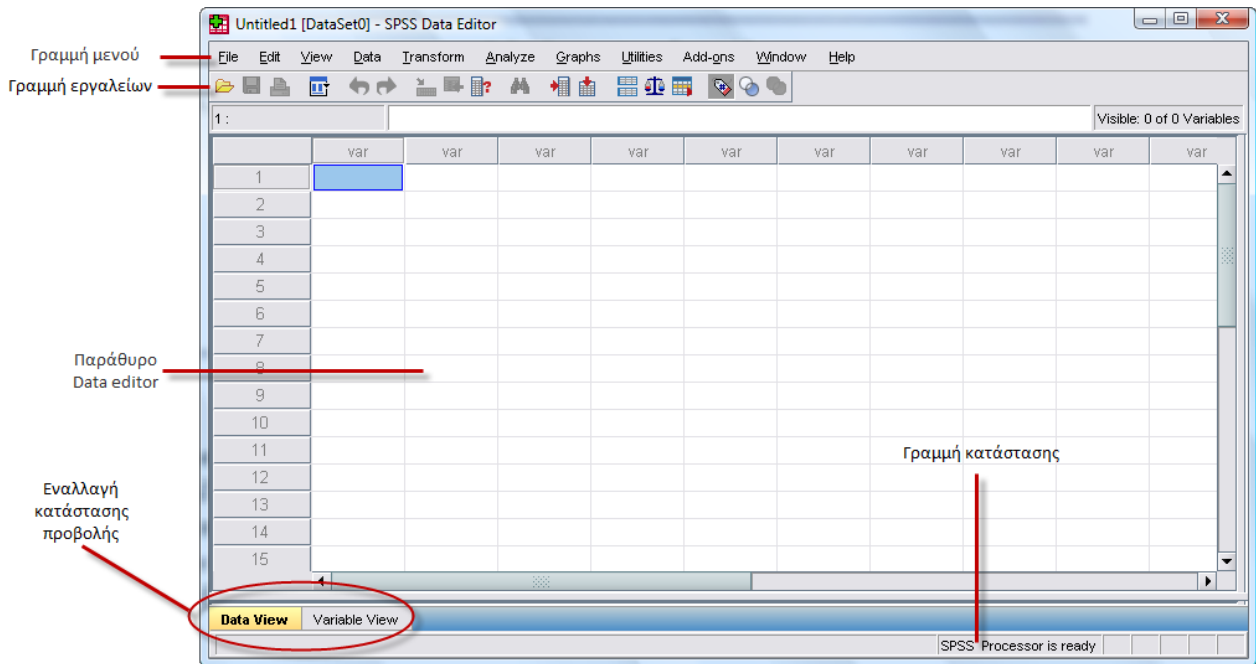
### 1.0 Το περιβάλλον του SPSS

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 γίνεται αναφορά σε αυτόν το σύντομο οδηγό.

Για να ξεκινήσετε την εφαρμογή, πηγαίνετε στο μενού [**Εναρξη --> Όλα τα προγράμματα --> SPSS**] και επιλέξτε το εικονίδιο του **SPSS 16.0**. Η εφαρμογή θα ξεκινήσει σχετικά σύντομα (πρόκειται για μια εφαρμογή αρκετά απαιτητική σε πόρους του συστήματος) και θα εμφανιστεί η παρακάτω οθόνη (βλ. Σχήμα 1.1), η οποία θα σας θυμίσει την εμφάνιση ενός λογιστικού φύλλου (π.χ., MS Excel). Το κελί (cell) της πρώτης γραμμής της πρώτης στήλης έχει πιο έντονο περίγραμμα και γαλανό χρώμα. Πρόκειται για το ενεργό κελί, όπου θα προστεθεί ό,τι πληκτρολογήσετε. Για τη μετακίνηση μεταξύ των κελιών μπορείτε να χρησιμοποιήσετε το ποντίκι ή τα πλήκτρα με τα βέλη (→, ←, ↓, ↑) στο πληκτρολόγιό σας.

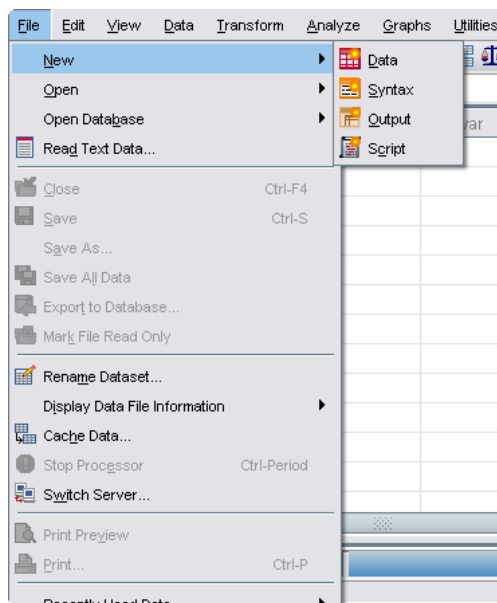
Όπως βλέπετε, στη **γραμμή μενού** υπάρχουν μια σειρά από επιλογές σχετικές με τις διάφορες λειτουργίες του λογισμικού. Υπάρχουν επίσης στη **γραμμή εργαλείων** ορισμένα εικονίδια συντόμευσης για άμεση πρόσβαση σε λειτουργίες που γίνονται συχνά. Τοποθετώντας για λίγο το δείκτη του ποντικιού πάνω από αυτά εμφανίζεται μια σύντομη επεξήγηση της λειτουργίας τους. Το Σχήμα 1.1 παρουσιάζει ένα άδειο φύλλο δεδομένων (data sheet). Τα δεδομένα εισάγονται είτε πληκτρολογώντας τα είτε εισάγοντάς τα απευθείας από κάποιο αρχείο.

**Σχήμα 1.1** Η αρχική οθόνη του SPSS (παράθυρο SPSS Data Editor – Επεξεργαστής Δεδομένων SPSS)



Ρίχνοντας μια πρόχειρη ματιά στο μενού **[File]** διαπιστώνουμε ότι πολλές από τις επιλογές του μας είναι ήδη γνωστές από άλλες παραθυρικές εφαρμογές (Σχήμα 1.2). Για παράδειγμα, η επιλογή **[New]** χρησιμοποιείται για να δηλωθεί ο τύπος του νέου αρχείου που θα δημιουργήσει. Οι διάφορες εναλλακτικές επιλογές στο **[New]** είναι:

**Σχήμα 1.2** Το μενού επιλογών του **[File]**



- **[Data]** Προεπιλεγμένο παράθυρο με ένα άδειο φύλλο δεδομένων έτοιμο για πέρασμα νέων στοιχείων.
- **[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 (Viewer). 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.
- **[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 στόχο του παρόντος εγχειριδίου.

Επίσης, στο μενού **[File]** υπάρχουν τρεις ξεχωριστές διέξοδοι για την ανάγνωση δεδομένων από υπάρχοντα αρχεία (Σχήμα 1.2). Η πρώτη είναι η επιλογή **[Open]**. Όπως και σε πολλές άλλες εφαρμογές (π.χ., MS Word, Excel, κ.λπ.) 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, μπορείτε για παράδειγμα να έχετε 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 είναι η επιλογή **[Open Database]**. Αυτή επιτρέπει την εισαγωγή δεδομένων που είναι αποθηκευμένα σε αρχεία του MS Excel ή της Access. Τέλος, η τρίτη επιλογή **[Read Text Data]**, όπως υποδηλώνει και το όνομά της, προορίζεται για την ανάγνωση αρχείων που έχουν αποθηκευτεί σε μορφοποίηση ASCII.

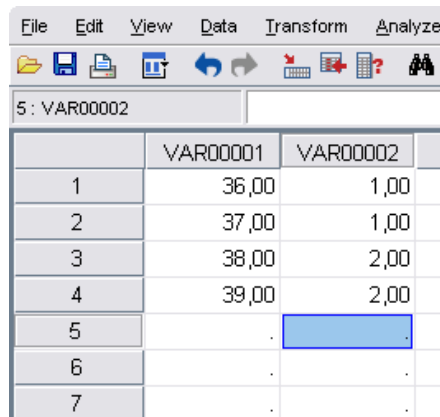
Προσέξτε επίσης ότι στο κάτω αριστερά τμήμα της οθόνης υπάρχουν δύο καρτέλες (**Data View & Variable View**). Στην πρώτη από αυτές **[Data View]** τοποθετείτε τα αρχικά δεδομένα σας, ενώ στη δεύτερη **[Variable View]** καθορίζετε τα χαρακτηριστικά των μεταβλητών της έρευνάς σας. Η μετάβαση από τη μία στην άλλη προβολή γίνεται πατώντας στο όνομα της καρτέλας που θέλετε να μεταβείτε.

## 2.0 Εισαγωγή και διαχείριση 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 πρακτική 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.

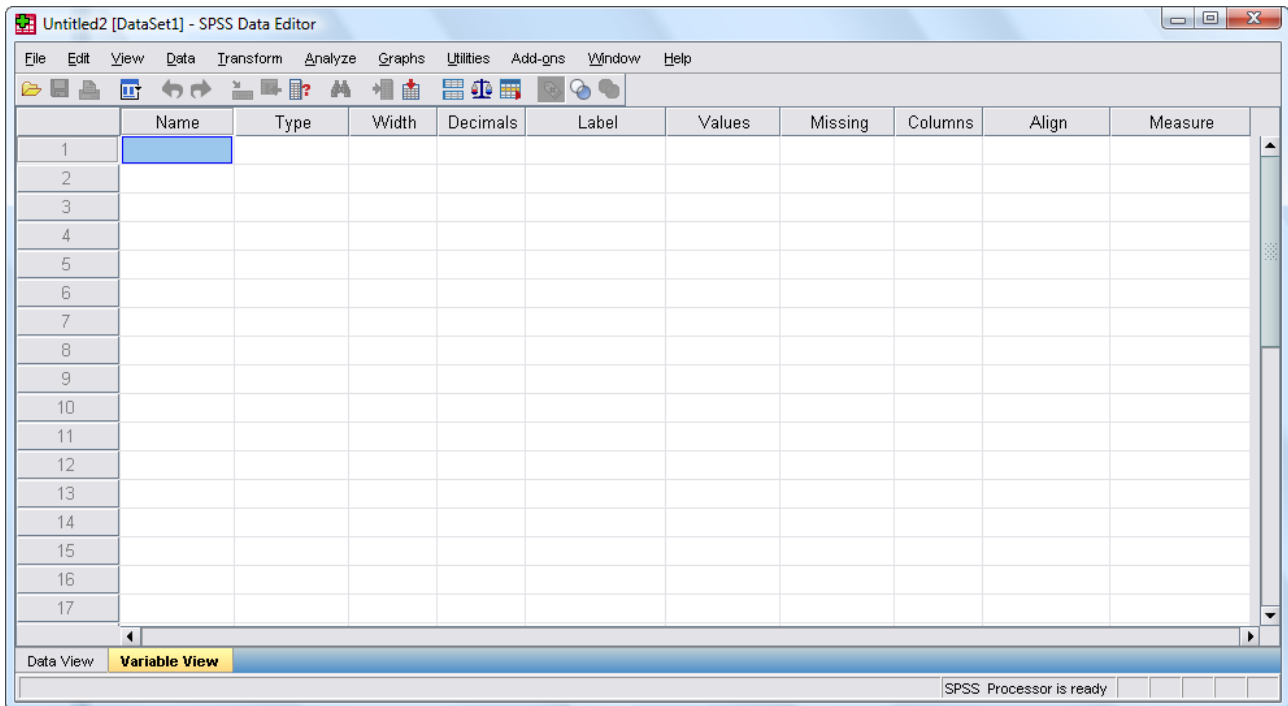
**Σχήμα 2.1** Εισαγωγή δεδομένων



	VAR00001	VAR00002
1	36,00	1,00
2	37,00	1,00
3	38,00	2,00
4	39,00	2,00
5	.	.
6	.	.
7	.	.

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, θα πρέπει να μεταβείτε από την κατάσταση προβολής δεδομένων [Data View] στην κατάσταση προβολής μεταβλητών [Variable View] (βλ. Σχήμα 1.1 και 2.2).

Σχήμα 2.2 Κατάσταση προβολής μεταβλητών [Variable View]



Στο Σχήμα 2.2 παρουσιάζεται η κατάσταση προβολής μεταβλητών **[Variable View]**. Εδώ έχουμε αντιστροφή της παρουσίασης των μεταβλητών, οι οποίες πλέον εμφανίζονται στις σειρές του πίνακα, ενώ στις διάφορες στήλες μπορείτε να καθορίσετε διάφορα χαρακτηριστικά για κάθε μεταβλητή.

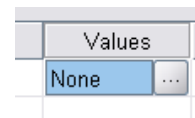
Από αριστερά προς τα δεξιά για κάθε μεταβλητή μπορείτε να καθορίσετε τα ακόλουθα:

- **[Name]** Εδώ καθορίζετε το όνομα της μεταβλητής. Οι νεώτερες εκδόσεις του SPSS δέχονται ελληνικούς χαρακτήρες, αλλά προτείνεται η χρήση ονομάτων με μέγεθος έως 8 λατινικούς χαρακτήρες χωρίς κενά, προκειμένου να διατηρηθεί η συμβατότητα του αρχείου σας με παλαιότερες εκδόσεις του SPSS. Στην περίπτωση που δώσετε το ίδιο όνομα σε δυο μεταβλητές θα λάβετε ένα προειδοποιητικό μήνυμα ότι κάτι τέτοιο δεν επιτρέπεται. Αν θέλετε να δώσετε ένα αναλυτικό όνομα με ελληνικούς χαρακτήρες μπορείτε να χρησιμοποιήσετε την επιλογή **[Label]** (βλ. παρακάτω). **Προσοχή!** Το όνομα της μεταβλητής δεν επιτρέπεται να ξεκινάει από αριθμό και δεν μπορεί να περιέχει κενά.
- **[Type]** Εδώ επιλέγετε 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 οποιονδήποτε συνδυασμό γραμμάτων και αριθμών), τα οποία δεν μπορούν να χρησιμοποιηθούν σε κάποια είδη αναλύσεων. Οι **[dot]**, **[comma]** και **[scientific notation]** αφορούν

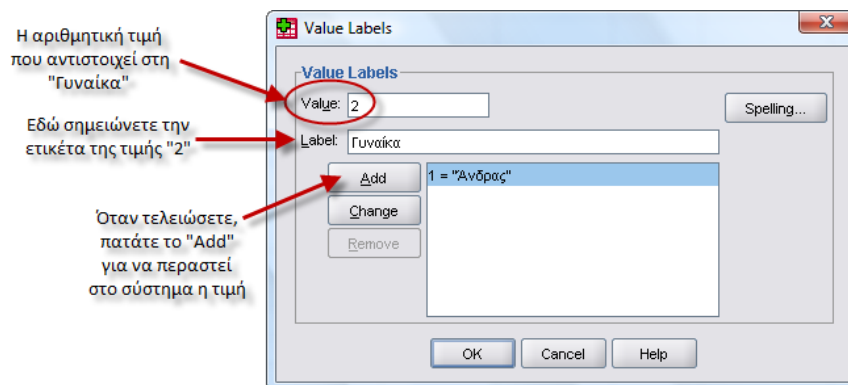
στον τρόπο παρουσίασης των αριθμών στα κελιά, ενώ οι υπόλοιπες επιλογές δεν απαιτούν κάποια ιδιαίτερη εξήγηση.

- **[Width]** Εδώ επιλέγετε το μέγιστο αριθμό χαρακτήρων που μπορείτε να καταχωρίσετε για κάθε περίπτωση στη μεταβλητή σας.
- **[Decimals]** Εδώ επιλέγετε το μέγιστο αριθμό δεκαδικών που μπορείτε να καταχωρίσετε για κάθε περίπτωση στη μεταβλητή σας.
- **[Label]** Εδώ μπορείτε να δώσετε ένα «δεύτερο όνομα» στη μεταβλητή με ελληνικούς χαρακτήρες, κενά ή ό,τι άλλο θέλετε με μήκος έως 256 χαρακτήρες. Είναι σημαντικό να προσέξετε την ορθογραφία των περιγραφών σας, πεζά – κεφαλαία γράμματα κ.λπ., ώστε να βελτιώσετε 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 κωδικοποιήσετε τις κατηγορίες/συνθήκες των μεταβλητών σε αριθμητική μορφή. Για παράδειγμα, το «άνδρας» και «γυναίκα» μπορεί να κωδικοποιηθεί ως "1" και "2" αντίστοιχα. Προκειμένου να αποφύγετε τη σύγχυση, συνιστάται να χρησιμοποιείτε 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.

Για να ονομάσετε τις τιμές μιας μεταβλητής, επιλέξτε το κατάλληλο κελί στη στήλη Values. Θα εμφανιστούν τρεις τελείες (...). Πατήστε πάνω στις τελείες και θα ανοίξει το πλαίσιο διαλόγου που ακολουθεί (Σχήμα 2.3).



**Σχήμα 2.3** Το πλαίσιο διαλόγου Value Labels



- **[Missing]** Εδώ μπορείτε να ορίσετε συγκεκριμένες τιμές που το SPSS θα εκλαμβάνει ως ελλείψεις κατά την ανάλυση των δεδομένων. Αν για μία περίπτωση αφήσετε 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 παράθυρο Viewer ως system missing (βλ. Πίνακα 2.1).

**Πίνακας 2.1 Παρουσίαση των system missing περιπτώσεων στο παράθυρο Viewer**

**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 ορίσετε συγκεκριμένες τιμές στο πεδίο missing, τότε το SPSS τις εμφανίζει ως απλώς missing και αντί της ένδειξης system αναφέρει την τιμή που ορίσατε στο πεδίο **[discrete missing values]** ή το **[label]** που έχετε ορίσει για τη συγκεκριμένη τιμή. Στο παράδειγμα του Πίνακα 2.2 η τιμή που ορίσατε είναι η 8, δηλαδή είπατε στο SPSS να θεωρήσει κάθε καταχώριση στη συγκεκριμένη μεταβλητή που έχει τιμή 8 ως ελλιπή περίπτωση.

**Πίνακας 2.2 Παρουσίαση των discrete missing values στο παράθυρο Viewer**

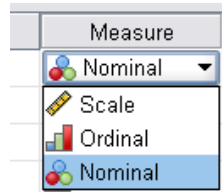
**Φύλο**

		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]** Με αυτή τη ρύθμιση μπορείτε να διαφοροποιήσετε το μέγεθος της κάθε στήλης όπως θα εμφανίζεται στην κατάσταση προβολής **[Data view]**. Η αλλαγή του πλάτους της στήλης από αυτή την επιλογή δεν πρόκειται να επηρεάσει το μέγεθος της μεταβλητής, όπως αυτό ρυθμίστηκε προηγουμένως μέσω της ρύθμισης **[Width]**.
- **[Align]** Πρόκειται για τη ρύθμιση του τρόπου με τον οποίο θα στοιχιστούν τα δεδομένα στη στήλη. Έχετε τρεις επιλογές: **Left** (Αριστερή στοίχιση), **Right** (Δεξιά στοίχιση) και **Center** (Στοίχιση στο κέντρο).

- **[Measure]** Αυτή η ρύθμιση είναι από τις πιο σημαντικές καθώς δηλώνετε στο SPSS με ποιο τύπο κλίμακας έχει γίνει η μέτρηση της μεταβλητής. Υπάρχουν τρεις επιλογές: **Scale**, **Ordinal** και **Nominal**. Αυτές αντιστοιχούν σε **Κλίμακα Ίσων Διαστημάτων** (ή αναλογική), **Ιεραρχική** (ή τακτική) και **Ονομαστική** (ή κατηγορική) αντίστοιχα (βλ. Σχήμα 2.4).

**Σχήμα 2.4** Το μενού επιλογών του Measure



Στο παράδειγμα του Σχήματος 2.1, μπορείτε να αντικαταστήσετε το "VAR00001" με "RT" (Reaction Time – Χρόνος Αντίδρασης) και το "VAR00002" με "freq" (frequency), χρησιμοποιώντας την επιλογή **[Name]** (βλ. Σχήμα 2.2). Η εικόνα που θα έχετε στην κατάσταση προβολής δεδομένων **[Data View]** είναι η ακόλουθη (βλ. Σχήμα 2.5). Αν ορίσετε και κάποιο **[Label]** για τις μεταβλητές σας, μπορείτε να το δείτε στην κατάσταση προβολής δεδομένων **[Data View]** αφήνοντας το δείκτη του ποντικιού σας χωρίς να τον κουνάτε πάνω στο όνομα μίας μεταβλητής.

**Σχήμα 2.5** Η αλλαγή του ονόματος των μεταβλητών

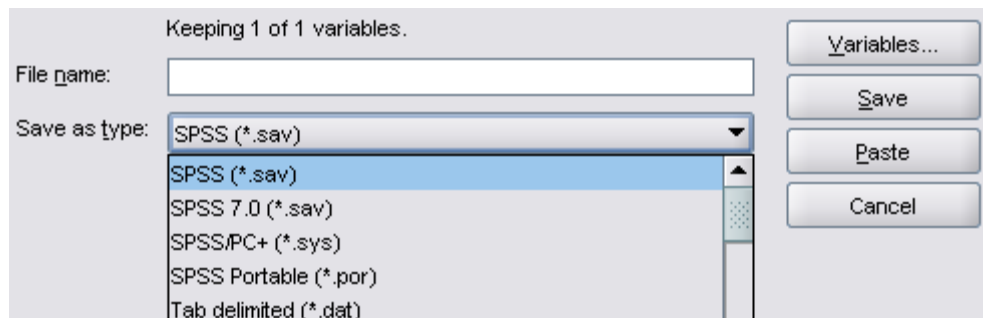
 A screenshot of the SPSS Data Editor window titled '\*Untitled1 [DataSet0] - SPSS Data Editor'. The window shows a data table with the following content:
 

	rt	freq	var
1	36,00	1,00	
2	37,00	1,00	
3	38,00	2,00	
4	39,00	2,00	
5	.	.	
6	.	.	

Αφού έχετε πληκτρολογήσει τα δεδομένα σας πρέπει να αποθηκεύσετε το αρχείο. Στην πραγματικότητα απαιτείται να αποθηκεύετε πολύ συχνά κατά τη διάρκεια της πληκτρολόγησης των δεδομένων, προκειμένου να αποφύγετε την απώλειά τους λόγω κάποιου «ατυχήματος». 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 **[Save As]**.

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



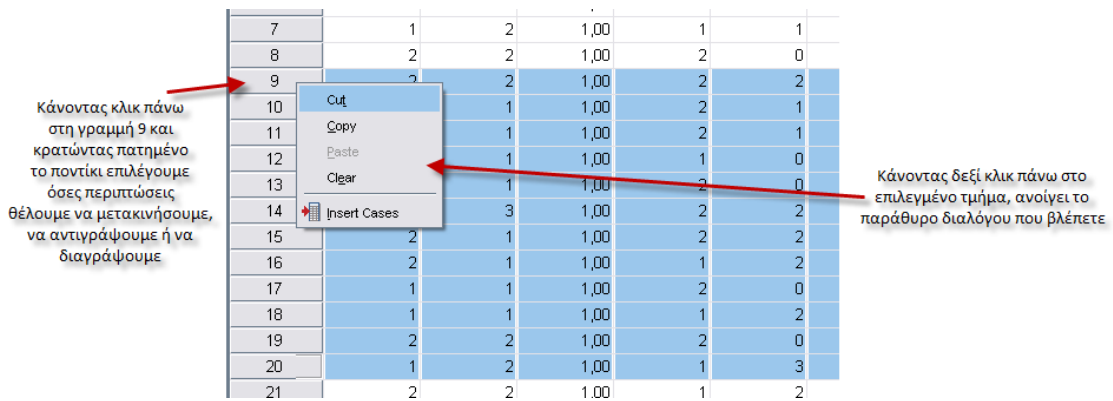
Αν σκοπεύετε να εργαστείτε μόνο στο SPSS, τότε είναι προτιμότερο να αποθηκεύσετε τα δεδομένα σας με τη μορφοποίηση SPSS (\*.sav). Αυτή η μορφοποίηση επιτρέπει ταχύτερη ανάγνωση και επεξεργασία του αρχείου των δεδομένων. Ωστόσο, αν πρόκειται να αναλύσετε ή να χρησιμοποιήσετε τα δεδομένα σας και μέσα από κάποια άλλη εφαρμογή (π.χ., ένα λογιστικό φύλλο), θα ήταν προτιμότερο να τα αποθηκεύσετε και 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 του υπολογιστή ή σε κάποιο φορητό μέσο αποθήκευσης δεδομένων.

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

Συχνά προκύπτει η ανάγκη να μετακινήσετε μέσα στο αρχείο δεδομένων ορισμένες περιπτώσεις ή να τις διαγράψετε ή να τις αντιγράψετε (copy & paste) σε ένα άλλο αρχείο δεδομένων. Για να επιλέξετε μία ή περισσότερες περιπτώσεις (cases) κάντε κλικ πάνω στον αριθμό της γραμμής (περίπτωσης) που θέλετε να μετακινήσετε και παρατηρήστε ότι όλα τα κελιά της συγκεκριμένης περίπτωσης γίνονται γαλάζια (πράγμα που σημαίνει ότι έχουν επιλεγεί). Με δεξί κλικ πάνω στην επιλεγμένη περιοχή ανοίγει το παράθυρο διαλόγου που βλέπετε στο Σχήμα 2.7. Από εκεί μπορείτε να επιλέξετε ή την αποκοπή (Cut) ή την αντιγραφή (Copy) ή τη διαγραφή (Clear) της περίπτωσης αυτής. Αν θέλετε να επιλέξετε περισσότερες από μία περιπτώσεις, κάντε κλικ στην πρώτη από αυτές και κρατώντας πατημένο το ποντίκι σύρετε προς τα κάτω επιλέγοντας όσες γραμμές (περιπτώσεις) θέλετε.

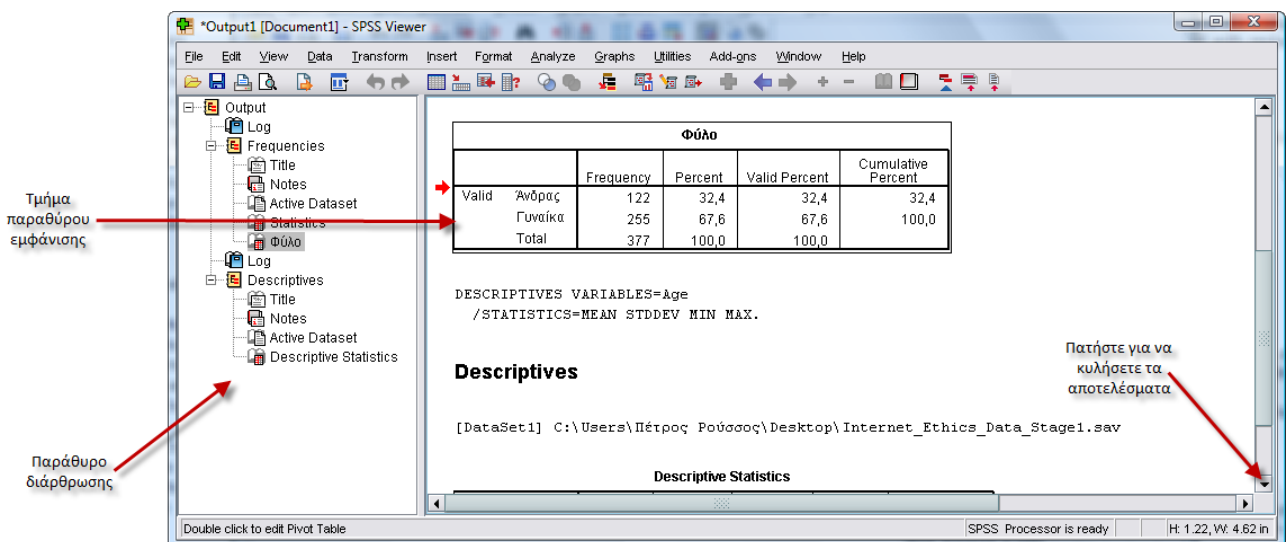
**Σχήμα 2.7** Επιλογή περιπτώσεων για μετακίνηση, αντιγραφή ή διαγραφή



### 3.0 Το παράθυρο Viewer

Τα αποτελέσματα όλων των εργασιών και των αναλύσεων που πραγματοποιούμε στο SPSS παρουσιάζονται σε ένα χωριστό παράθυρο, το οποίο ονομάζεται SPSS Viewer. Όπως φαίνεται στο Σχήμα 3.1, το παράθυρο Viewer χωρίζεται σε δύο μέρη ή **τμήματα παραθύρου** (panes). Το αριστερό μέρος (**τμήμα διάρθρωσης**) περιέχει μια διαρθρωτική άποψη όλων των διαφορετικών μερών των αποτελεσμάτων στο παράθυρο Viewer, είτε φαίνονται τη συγκεκριμένη στιγμή είτε όχι. Το δεξιό μέρος (το **τμήμα εμφάνισης**) περιέχει τα αποτελέσματα (πίνακες, διαγράμματα κ.λπ.).

**Σχήμα 3.1** Το παράθυρο Viewer



Για να αλλάξετε το μέγεθος των δύο τμημάτων παραθύρου, δείξτε με το ποντίκι σας στη διαχωριστική γραμμή μεταξύ τους, πατήστε το αριστερό πλήκτρο του ποντικιού και χωρίς να το αφήσετε σύρετε τη γραμμή προς τα αριστερά ή προς τα δεξιά.

Πολλοί αγνοούν το τμήμα διάρθρωσης και απλώς κυλούν τα αποτελέσματα που εμφανίζονται στη δεξιά πλευρά του παραθύρου Viewer. Ωστόσο, το τμήμα διάρθρωσης προσφέρει αρκετές ευκολίες:

- Καταρχάς, τα διάφορα τμήματα των αποτελεσμάτων συνδέονται με εικονίδια «βιβλίου» στο τμήμα διάρθρωσης. Κάθε εικονίδιο αντιπροσωπεύει ένα ειδικό μέρος των αποτελεσμάτων, όπως είναι ένας πίνακας. Έτσι, αν πατήσετε σε ένα από αυτά τα εικονίδια, θα παρουσιαστεί αμέσως στο τμήμα εμφάνισης το αντίστοιχο τμήμα της ανάλυσης.
- Τα εικονίδια αυτά χρησιμοποιούνται επίσης για να κρύψουν ή για να εμφανίσουν προσωρινά μέρη των αποτελεσμάτων. Προσέξτε ότι τα περισσότερα από τα εικονίδια αυτά είναι εικονίδια «ανοιχτού βιβλίου», ενώ κάποια μοιάζουν περισσότερο με κλειστά βιβλία. Ένα εικονίδιο «κλειστού βιβλίου» αντιπροσωπεύει ένα κρυμμένο μέρος των αποτελεσμάτων. Τα κρυμμένα μέρη δεν εμφανίζονται στο τμήμα εμφάνισης, αλλά μπορείτε να τα αποκαλύψετε όποια στιγμή θέλετε να τα δείτε. Για να κρύψετε ή να αποκαλύψετε ένα μέρος των αποτελεσμάτων διπλοπατήστε στο εικονίδιο του ανοιχτού ή κλειστού βιβλίου αντίστοιχα.
- Αν θέλετε να κρύψετε όλα τα στοιχεία μιας εργασίας (π.χ., **[Frequencies]**), πατήστε στο μικρό πλαίσιο με το σύμβολο «πλην» στα αριστερά του ονόματος της εργασίας. Ολόκληρο το αντίστοιχο μέρος της διάρθρωσης συμπύσσεται και το σύμβολο «πλην» αλλάζει σε «συν» για να δείξει ότι εδώ κρύβεται μια ενότητα των αποτελεσμάτων. Πατήστε το σύμβολο «συν» για να εμφανίσετε και πάλι ολόκληρη την κρυμμένη ενότητα.
- Μπορείτε, τέλος, να αναδιατάξετε τα αποτελέσματα (πατήστε με το αριστερό πλήκτρο του ποντικιού σε ένα εικονίδιο και κρατώντας το πατημένο σύρετέ το σε διαφορετική θέση της διάρθρωσης) ή να διαγράψετε μέρος τους (πατήστε στο εικονίδιο και στη συνέχεια πατήστε το πλήκτρο **[Delete]**).

Το τμήμα εμφάνισης (display pane) παρουσιάζει όσο μέρος των αποτελεσμάτων μπορεί να χωρέσει η οθόνη σας. Για να δείτε περισσότερα, κυλήστε τα περιεχόμενα του τμήματος παραθύρου ή χρησιμοποιήστε το τμήμα διάρθρωσης για να μεταφερθείτε σε διάφορες θέσεις του. Το τμήμα εμφάνισης περιέχει αντικείμενα διαφόρων ειδών: πίνακες αριθμών (στην πραγματικότητα πρόκειται για έναν ειδικό τύπο πίνακα που ονομάζεται **συγκεντρωτικός πίνακας** – pivot table), διαγράμματα και τμήματα κειμένου, όπως είναι οι τίτλοι ή οι σημειώσεις.

Πατώντας σε ένα από τα τμήματα των αποτελεσμάτων αυτό αμέσως περικλείεται από ένα μαύρο πλαίσιο και εμφανίζεται ένα κόκκινο βελάκι αριστερά του για να δείξει ότι το αντίστοιχο τμήμα έχει επιλεγεί. Κάνοντας δεξί κλικ πάνω του ή επιλέγοντας **[Copy]** από το μενού **[Edit]** μπορείτε να αντιγράψετε το τμήμα αυτό και να το επικολλήσετε σε μια άλλη εφαρμογή (π.χ., στο MS Word όπου γράφετε την ερευνητική σας αναφορά). **Προσοχή!** Αν το αντικείμενο που επιλέξατε για αντιγραφή είναι πίνακας μπορείτε να τον επεξεργαστείτε στο πρόγραμμα επεξεργασίας κειμένου που χρησιμοποιείτε. Αν όμως αυτό είναι ένα

διάγραμμα, τότε επεξεργασία δεν μπορεί να γίνει σε άλλη εφαρμογή καθώς εξάγεται ως αντικείμενο (*object*).

Διπλοπατώντας σε ένα από τα τμήματα των αποτελεσμάτων ανοίγει ένα πρόγραμμα επεξεργασίας ειδικά σχεδιασμένο για την τροποποίησή του. Η εμφάνιση του αντικειμένου αλλάζει για να δείξει ότι το επεξεργάζεστε. Αν το αντικείμενο είναι διάγραμμα, ανοίγει ένα ειδικό παράθυρο επεξεργασίας γραφημάτων (**Chart Editor**) που παρέχει ένα ισχυρό σύνολο εργαλείων για να αλλάξετε την εμφάνιση του διαγράμματος.

Μια ιδιαιτερότητα του SPSS είναι ότι τα αρχικά δεδομένα και τα αποτελέσματα των εργασιών και των αναλύσεων που πραγματοποιείτε δεν αποθηκεύονται στο ίδιο αρχείο. Είδαμε παραπάνω ότι τα δεδομένα μας τα αποθηκεύσαμε σε ένα αρχείο με κατάληξη “.sav”. Τα αποτελέσματα αποθηκεύονται σε χωριστά αρχεία με κατάληξη “.spv” (π.χ., “output.spv”). **Προσοχή!** Ως την έκδοση 15 το SPSS δημιουργούσε αρχεία με κατάληξη “.sps”, τα οποία δεν μπορεί να ανοίξει το SPSS 16 (όπως και το αντίθετο). Στην περίπτωση που έχετε αποθηκευμένα αρχεία “.sps” από παλαιότερες εκδόσεις του SPSS, τα οποία θέλετε να ανοίξετε και να επεξεργαστείτε, μπορείτε να εγκαταστήσετε το SPSS 15.0 Smart Viewer (διαθέσιμο στο CD εγκατάστασης του SPSS 16.0).

## 4.0 Περιγραφική Στατιστική

Ας υποθέσουμε ότι έχετε περάσει ήδη τα δεδομένα σας στην εφαρμογή και έχετε δημιουργήσει ένα αρχείο \*.sav. Στο Σχήμα 4.1 βλέπετε ένα μόνο τμήμα ενός τυπικού αρχείου δεδομένων του SPSS.

Σχήμα 4.1 Ο επεξεργαστής δεδομένων του SPSS

Ο επεξεργαστής κελιών εμφανίζει την τιμή του επιλεγμένου κελιού

Επιλεγμένο κελί

Αφήνοντας το δείκτη του ποντικιού πάνω στο όνομα της μεταβλητής εμφανίζεται η περιγραφή

Οι γραμμές είναι οι περιπτώσεις

Οι στήλες είναι οι μεταβλητές

	Gender	Age	Domicile	Education	Int_Access	Int_Job	Int_Know	Int_Conn
1	Γυναίκα	22	Αστικό κέντρο	Μορφωτικό επίπεδο	Σπίτι	Μεγάλη	Επαρκείς	ADSL
2	Άνδρας	28	Αστικό κέντρο	Ανώτατη	Σπίτι	Μεγάλη	Πολύ καλές	ADSL
3	Άνδρας	29	Αστικό κέντρο	Ανώτερη	Σπίτι	Μεγάλη	Επαρκείς	ADSL
4	Γυναίκα	21	Αστικό κέντρο	Ανώτατη	Σπίτι	Μικρή	Επαρκείς	ADSL
5	Άνδρας	29	Αστικό κέντρο	Ανώτατη	Σπίτι	Μικρή	Ελάχιστες	ISDN-PSTN E
6	Γυναίκα	27	Αστικό κέντρο	Ανώτατη	12	Μικρή	Επαρκείς	ISDN-PSTN E
7	Άνδρας	26	Αστικό κέντρο	Ανώτατη	Σπίτι	Μικρή	Επαρκείς	ADSL
8	Άνδρας	31	Υπαίθριος	Ανώτατη	12	Μεγάλη	Πολύ καλές	ISDN-PSTN
9	Άνδρας	32	Αστικό κέντρο	Υποχρεωτική	12	Μεγάλη	Πολύ καλές	ADSL
10	Γυναίκα	26	Αστικό κέντρο	Ανώτατη	Σπίτι	Μεγάλη	Επαρκείς	ADSL
11	Άνδρας	26	Αστικό κέντρο	Ανώτερη	Σπίτι	Όχι	Επαρκείς	ADSL
12	Άνδρας	34	Αστικό κέντρο	Ανώτερη	Σπίτι	Όχι	Επαρκείς	ADSL
13	Άνδρας	28	Αστικό κέντρο	Ανώτερη	Σπίτι	Όχι	Επαρκείς	ADSL
14	Άνδρας	23	Αστικό κέντρο	Ανώτερη	Σπίτι	Μεγάλη	Πολύ καλές	ADSL
15	Άνδρας	29	Αστικό κέντρο	Ανώτατη	Σπίτι	Μεγάλη	Επαρκείς	ADSL

Όταν έχετε ολοκληρώσει το πέρασμα των δεδομένων σας, το επόμενο βήμα είναι η εξαγωγή ορισμένων περιγραφικών στατιστικών δεικτών ή μέτρων. Για να περιγράψετε μία μεταβλητή, πρέπει να γνωρίζετε την κλίμακα μέτρησής της. Συνοπτικά ισχύουν τα ακόλουθα:

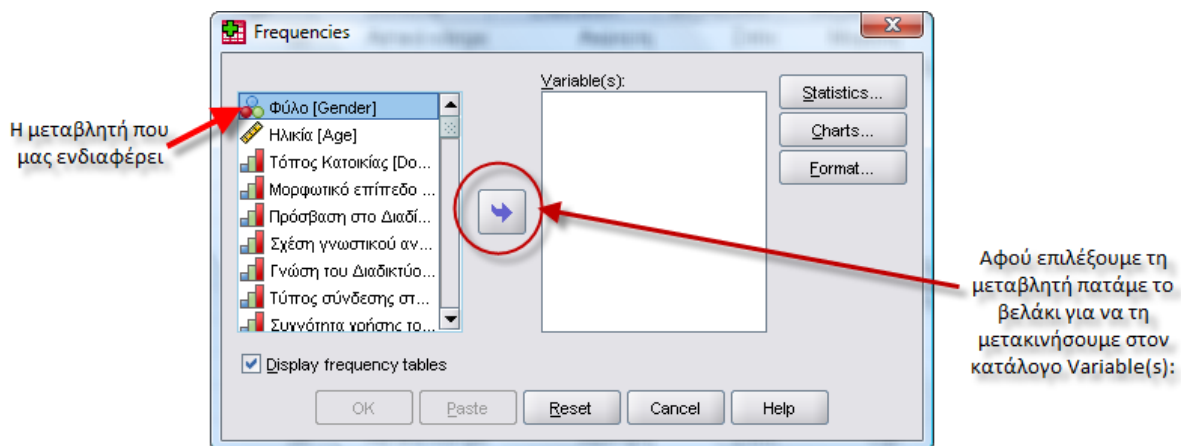
Μεταβλητή (Κλίμακα μέτρησης)	Περιγραφική στατιστική
Κατηγορική ή ονομαστική	Κατανομή συχνότητας
Τακτική	Δείκτες κεντρικής τάσης & διασποράς (Διάμεσος, εύρος)
Αριθμητική	Δείκτες κεντρικής τάσης & διασποράς (Μέσος όρος, τυπική απόκλιση)

#### 4.1 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 4.2 που ακολουθεί, ο χρήστης ενδιαφέρεται για τη μεταβλητή «Φύλο (Gender)», 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 Viewer με τις πληροφορίες για τις συχνότητες καθεμίας από τις τιμές της μεταβλητής που επιλέχτηκε.

Σχήμα 4.2 Η εντολή Frequencies



Τα αποτελέσματα της εντολής **Frequencies** παρουσιάζονται στον Πίνακα 4.1:

**Πίνακας 4.1 Κατανομή συχνότητας της μεταβλητής «Φύλο»**

		Φύλο			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Άνδρας	122	32,3	32,4	32,4
	Γυναίκα	255	67,5	67,6	100,0
	Total	377	99,7	100,0	
Missing	System	1	,3		
Total		378	100,0		

Η στήλη **Frequency** δείχνει την απόλυτη συχνότητα των τιμών της μεταβλητής στο δείγμα μας. Για παράδειγμα, στον Πίνακα 4.1 φαίνεται ότι 122 από τους συμμετέχοντες ήταν άντρες. Η στήλη **Percent** και **Valid Percent** δείχνουν τη σχετική συχνότητα των τιμών της μεταβλητής στο δείγμα, 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 378 (συμπεριλαμβανομένης μίας system missing περιπτώσεως), ενώ το σύνολο των έγκυρων περιπτώσεων είναι 377. 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 και δεν έχει νόημα για όλες τις μεταβλητές (παρότι το SPSS υπολογίζει την αθροιστική σχετική συχνότητα κάθε φορά που θα του ζητήσετε να 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 παρουσιάσετε.

Στον Πίνακα 4.2 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 περιγράψουμε 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.



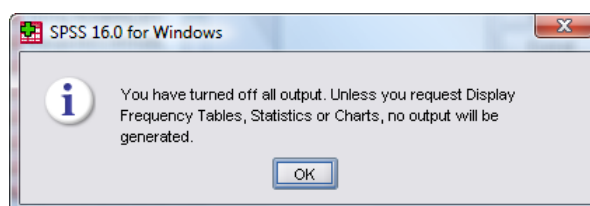
Πίνακας 4.2 Παράδειγμα χρήσης της σχετικής αθροιστικής συχνότητας

		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		

#### 4.2 Δείκτες κεντρικής τάσης και διασποράς

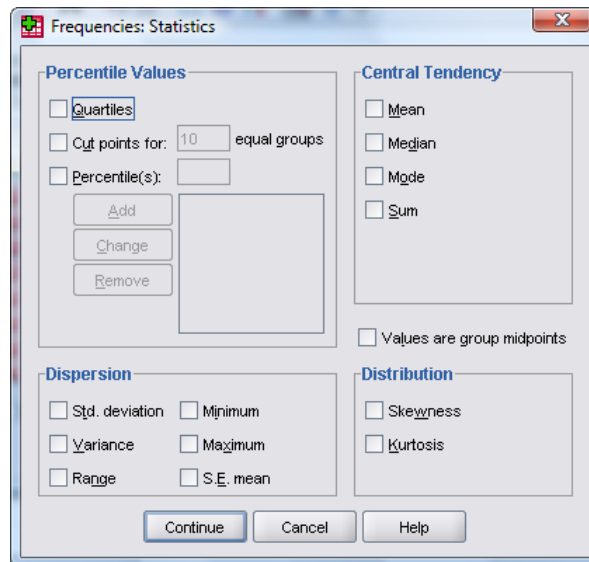
Για να ζητήσετε από το SPSS να υπολογίσει τους δείκτες κεντρικής τάσης και διασποράς που επιθυμείτε, μπορείτε να χρησιμοποιήσετε και πάλι την εντολή *Frequencies* (υπάρχουν και άλλοι τρόποι, αλλά τα αποτελέσματα που θα πάρετε δεν αλλάζουν).

Κάντε πάλι κλικ στα **[Analyze --> Descriptive Statistics --> Frequencies]** και βεβαιωθείτε ότι στην πρώτη οθόνη που θα εμφανιστεί (βλ. Σχήμα 4.2) απενεργοποιείτε 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 4.3), το οποίο σας ενημερώνει ότι αν συνεχίσετε χωρίς να επιλέξετε κάποιους δείκτες, η ανάλυσή σας δε θα έχει κανένα αποτέλεσμα στο παράθυρο Viewer.

Σχήμα 4.3 Εντολή *Frequencies* - Προειδοποιητικό μήνυμα

Κάντε κλικ στο κουμπί **[OK]** και στη συνέχεια κάντε κλικ στο κουμπί **[Statistics]** που οδηγεί στην εμφάνιση της οθόνης του Σχήματος 4.4. Στην οθόνη αυτή επιλέγετε 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 4.4 Εντολή Frequencies – επιλογή Statistics



Το αποτέλεσμα που θα παίρνατε αν διαλέγατε ταυτόχρονα όλες τις παραπάνω επιλογές είναι το ακόλουθο (βλ. Πίνακα 4.3).

Πίνακας 4.3 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 4.3 βλέπετε ότι είχαμε 1514 έγκυρες απαντήσεις και 5 ελλιπείς περιπτώσεις.

**Mean** είναι ο μέσος όρος της κατανομής. **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 γύρω από το μέσο όρο. Η τυπική απόκλιση εκφράζεται σε μονάδες μέτρησης ίδιες με αυτές του μέσου όρου, και όσο μικρότερη είναι τόσο πιο κοντά στο μέσο όρο είναι οι 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.

### 4.3 Ανάλυση πολλαπλών απαντήσεων

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

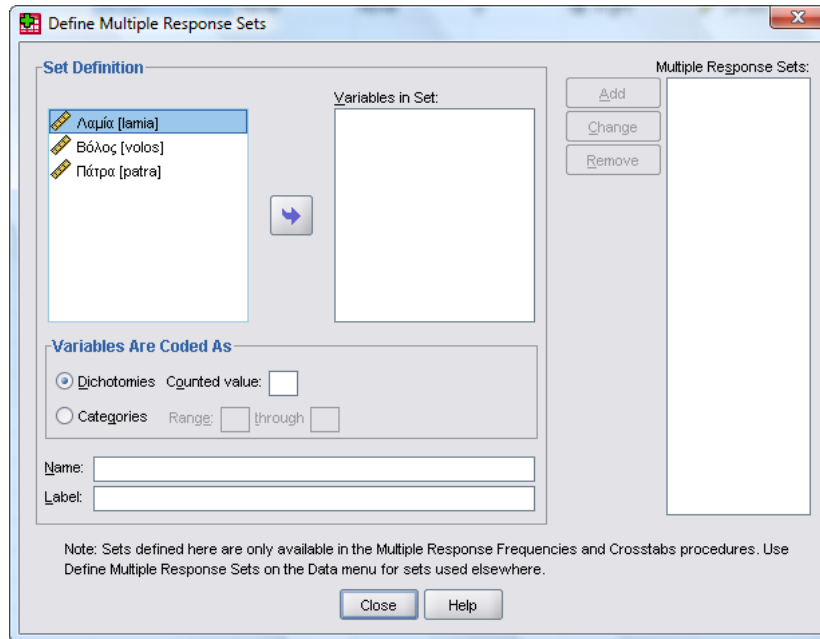
Για να το επιτύχετε αυτό καταχωρίζετε κάθε απάντηση ως μία μεταβλητή και κωδικοποιείτε “1” για το «ναι» και “0” για το «όχι». Έτσι ένα άτομο που έχει επισκεφθεί και τις τρεις πόλεις θα είχε 1 – 1 – 1 στις τρεις μεταβλητές του Πίνακα 4.4, ένα άλλο που δεν έχει επισκεφθεί καμία πόλη 0 – 0 – 0 και ένα που έχει επισκεφθεί μόνο το Βόλο 0 – 1 – 0.

**Πίνακας 4.4 Καταχώριση 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

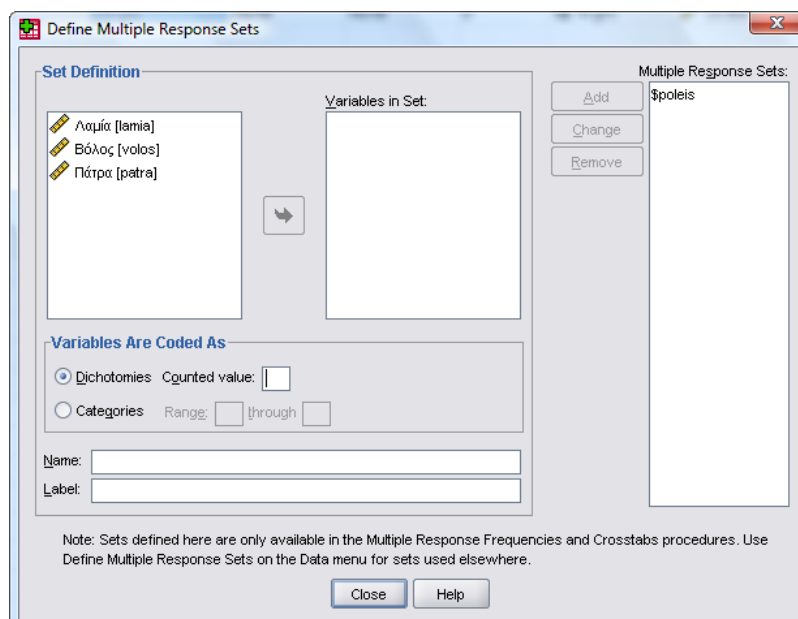
Προκειμένου να εκτελέσετε μία ανάλυση πολλαπλών απαντήσεων επιλέγετε **[Analyze --> Multiple Response --> Define Sets...]** και εμφανίζεται η οθόνη του Σχήματος 4.5:

Σχήμα 4.5 Ανάλυση πολλαπλών απαντήσεων – Define Sets



Στην οθόνη του Σχήματος 4.5 επιλέγετε τις μεταβλητές–απαντήσεις που αντιστοιχούν στην ερώτησή σας και δημιουργείτε ένα set απαντήσεων μεταφέροντας τις σχετικές μεταβλητές στη λίστα **[Variables in Set:]**. Στην συνέχεια στο πεδίο **[Variables Are Coded As]** επιλέγετε **[Dichotomies]** και στη συνέχεια στο πεδίο **[Counted Value:]** καταχωρείτε τον κωδικό που αντιστοιχεί στο «ναι» (“1” στο παράδειγμά μας). Δίνετε ένα όνομα στο set που δημιουργήσατε στο πεδίο **[Name:]** και πατάτε το πλήκτρο **[Add]** στα δεξιά της οθόνης. Η οθόνη που θα πάρετε είναι η ακόλουθη (βλ. Σχήμα 4.6):

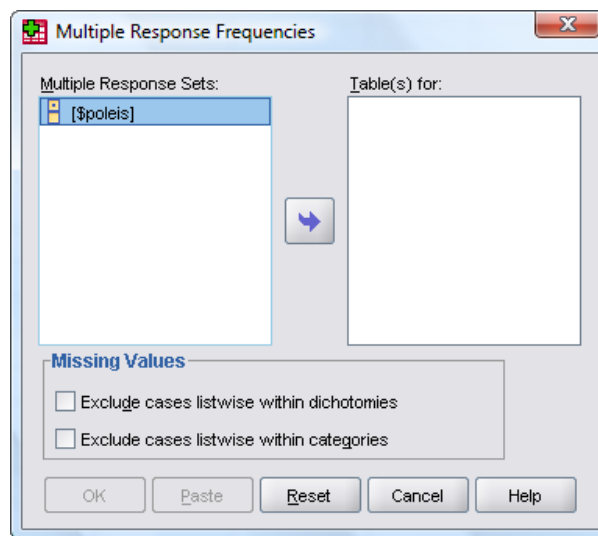
Σχήμα 4.6 Ανάλυση πολλαπλών απαντήσεων – ορισμός set



Όπως βλέπετε στο Σχήμα 4.6, στη λίστα **[Multiple Response Sets:]** έχει προστεθεί το set που μόλις δημιουργήσατε (**\$poleis** στο παράδειγμά μας). Επαναλαμβάνετε την ίδια διαδικασία για όσα sets απαντήσεων υπάρχουν στο αρχείο δεδομένων σας και πατάτε το πλήκτρο **[Close]**.

Στη συνέχεια επιλέγετε **[Analyze --> Multiple Response --> Frequencies...]** προκειμένου να δημιουργήσετε μία κατανομή πολλαπλών απαντήσεων για το set απαντήσεων που δημιουργήσατε. Η οθόνη που παίρνετε είναι η ακόλουθη (βλ. Σχήμα 4.7):

**Σχήμα 4.7** Ανάλυση πολλαπλών απαντήσεων – frequencies



Στο παράθυρο του Σχήματος 4.7 επιλέγετε τα sets απαντήσεων που σας ενδιαφέρουν, τα μεταφέρετε στη λίστα **[Table(s) for:]**, και στη συνέχεια πατάτε το πλήκτρο **[OK]**. Στο παράθυρο Viewer θα πάρετε τα αποτελέσματα που βλέπετε στον Πίνακα 4.5:

**Πίνακας 4.5** 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 4.5 αφορούν στις απόλυτες και σχετικές συχνότητες επί των απαντήσεων (Responses) και όχι επί των συμμετεχόντων. Όπως βλέπετε συνολικά στην ερώτησή μας

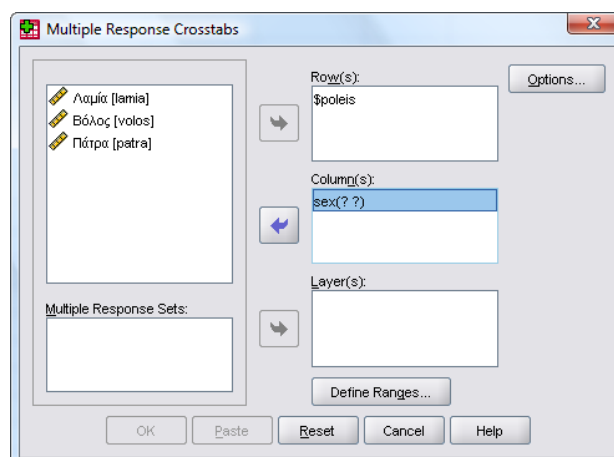
δόθηκαν 698 θετικές απαντήσεις, 329 για τη Λαμία, 206 για το Βόλο και 163 για την Πάτρα. Το 47,1% των απαντήσεων ήταν θετικές απαντήσεις για τη Λαμία, το 29,5% για το Βόλο και το 23,4% για την Πάτρα.

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

Ο σωστός τρόπος να διαβαστεί η τρίτη στήλη του Πίνακα 4.5 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.

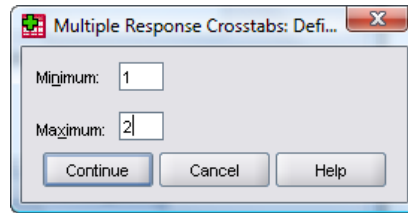
Επιλέγοντας [**Analyze --> Multiple Response --> Crosstabs...**] μπορείτε να δημιουργήσετε πίνακες διπλής εισόδου χρησιμοποιώντας συνδυασμούς set απαντήσεων και μεταβλητών. Η οθόνη που εμφανίζεται παρουσιάζεται στο Σχήμα 4.8.

#### Σχήμα 4.8 Ανάλυση πολλαπλών απαντήσεων – crosstabs



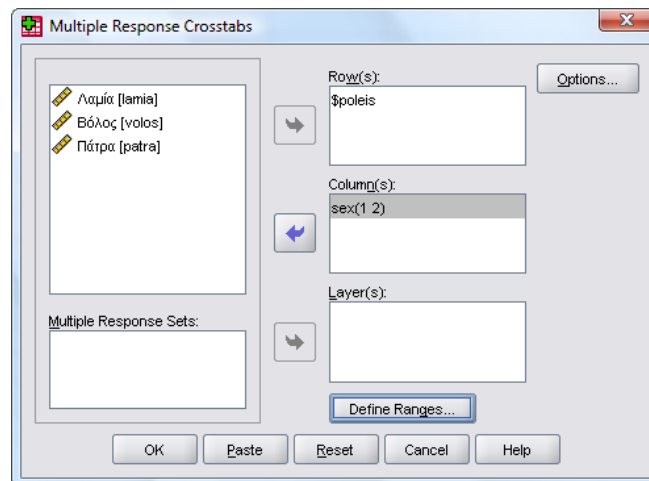
- Στα πεδία **Row(s):** και **Columns:** βάζετε τις μεταβλητές που θα παρουσιάζονται στον πίνακα διπλής εισόδου. Στο παράδειγμά μας επιλέξαμε το set απαντήσεων που ορίσαμε προηγουμένως (\$poleis) για τις σειρές του πίνακα και τη μεταβλητή «φύλο» (sex) για τις στήλες του πίνακα.
- Όπως βλέπετε, στο Σχήμα 4.8 εμφανίζονται δύο λατινικά ερωτηματικά δίπλα από το όνομα της μεταβλητής sex. Στα ερωτηματικά αυτά πρέπει να αντιστοιχίσετε τους κωδικούς που έχετε δώσει στις κατηγορίες της μεταβλητής. Επιλέγετε τη μεταβλητή ομαδοποίησης και κάνετε κλικ στο πλήκτρο [**Define Ranges...**], με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 4.9.

Σχήμα 4.9 Ανάλυση πολλαπλών απαντήσεων - define ranges



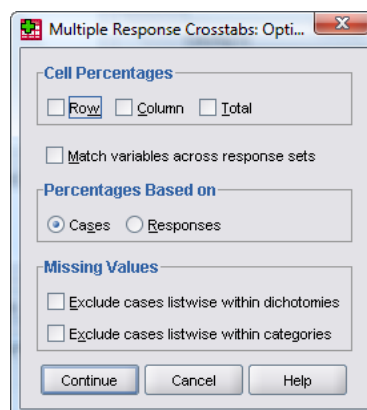
- Στην οθόνη του Σχήματος 4.9 ορίζετε τον ελάχιστο και το μεγαλύτερο κωδικό που περιλαμβάνει όλες τις κατηγορίες της μεταβλητής. Στο παράδειγμά μας, η μεταβλητή φύλο έχει δύο κατηγορίες: 1=Άντρας και 2=Γυναίκα. Για αυτό βάζουμε 1 στο πεδίο **[Minimum:]** και 2 στο πεδίο **[Maximum:]**. Αν είχατε μία άλλη μεταβλητή με πέντε κατηγορίες, οι οποίες είχαν κωδικοποιηθεί με 1, 2, 3, 4, 5 θα βάζατε 1 στο πεδίο **[Minimum:]** και 5 στο πεδίο **[Maximum:]**. Στη συνέχεια πατάτε **[OK]** και η οθόνη παίρνει τη μορφή που παρουσιάζεται στο Σχήμα 4.10:

Σχήμα 4.10 Ανάλυση πολλαπλών απαντήσεων - crosstabs μετά το define ranges



- Στην συνέχεια κάνετε κλικ στο πλήκτρο **[Options...]** και εμφανίζεται η οθόνη του Σχήματος 4.11:

Σχήμα 4.11 Ανάλυση πολλαπλών απαντήσεων - crosstabs - options



- Στην οθόνη του Σχήματος 4.11 επιλέγετε **[Row]** και **[Column]** στο πεδίο **[Cell Percentages]**. Στο πεδίο **[Percentages Based on]** μπορείτε να επιλέξετε αν τα ποσοστά που θα εμφανιστούν θα βασίζονται στους συμμετέχοντες (Cases) ή στις απαντήσεις (Responses). Η προεπιλογή είναι να εμφανίζονται ποσοστά για τους συμμετέχοντες.
- Πατάτε το πλήκτρο **[Continue]** και επιστρέφετε στην οθόνη του Σχήματος 4.10, όπου δίνετε **[OK]** και παίρνετε στο παράθυρο Viewer τα ακόλουθα αποτελέσματα (βλ. Πίνακα 4.6):

**Πίνακας 4.6 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  $\chi^2$ .

## 5.0 Μετασχηματισμοί μεταβλητών

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

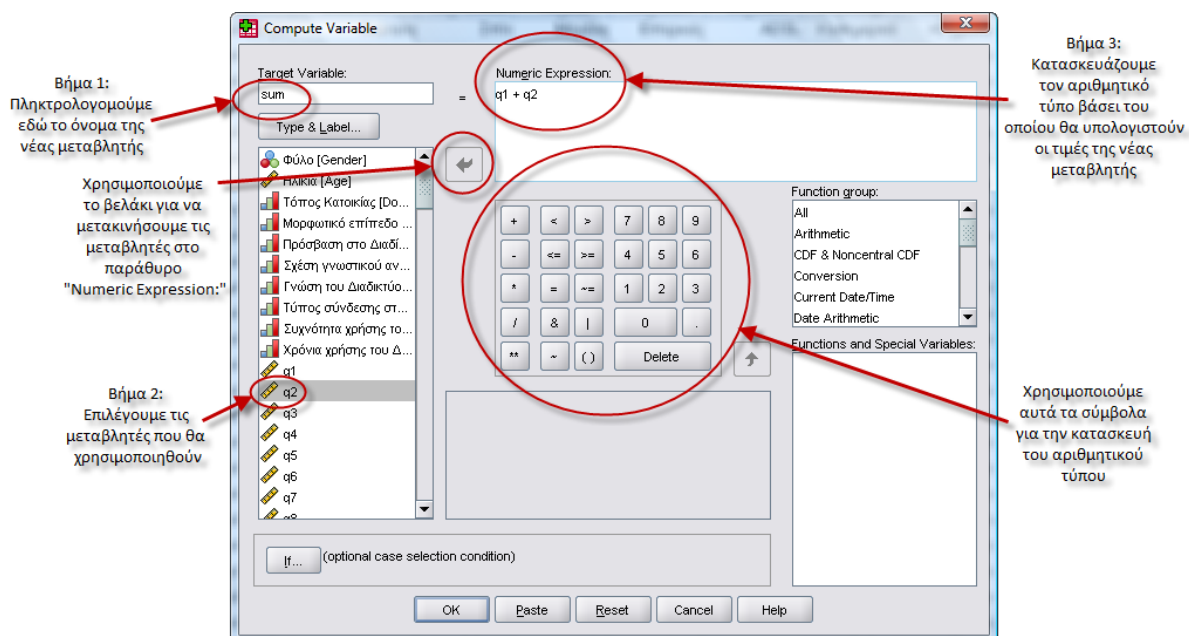
### 5.1 Compute

Η απλούστερη εντολή για τη δημιουργία μίας νέας μεταβλητής είναι η εντολή **[Compute]**. Για να τη χρησιμοποιήσετε επιλέγετε **[Transform --> Compute]** και εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 5.1.



Στο πεδίο **[Target Variable]** στο πάνω αριστερά τμήμα της οθόνης πληκτρολογείτε το όνομα της νέας μεταβλητής που θέλετε να κατασκευάσετε, ενώ στο πεδίο **[Numeric Expression]** τον αριθμητικό τύπο βάσει του οποίου θα υπολογίζεται η τιμή της νέας μεταβλητής για κάθε περίπτωση. Στο παράδειγμα του Σχήματος 5.1 έχουμε επιλέξει να δημιουργήσουμε τη νέα μεταβλητή "sum", η οποία ισούται με το άθροισμα των μεταβλητών  $q1 + q2$ . Στον αριθμητικό τύπο μπορείτε να χρησιμοποιήσετε οποιαδήποτε υπάρχουσα μεταβλητή, ένα νούμερο που επιθυμείτε ή συνδυασμό αυτών. Για παράδειγμα, αν υποθέσουμε ότι έχετε ζητήσει από τους συμμετέχοντές σας το έτος γέννησής τους (μεταβλητή "etos"), μπορείτε να δημιουργήσετε μία νέα μεταβλητή με την ηλικία τους, η οποία να ισούται με το τρέχον έτος μείον το έτος γέννησης (δηλαδή  $ilikia = 2008 - etos$ ).

**Σχήμα 5.1 Η εντολή [Compute]**

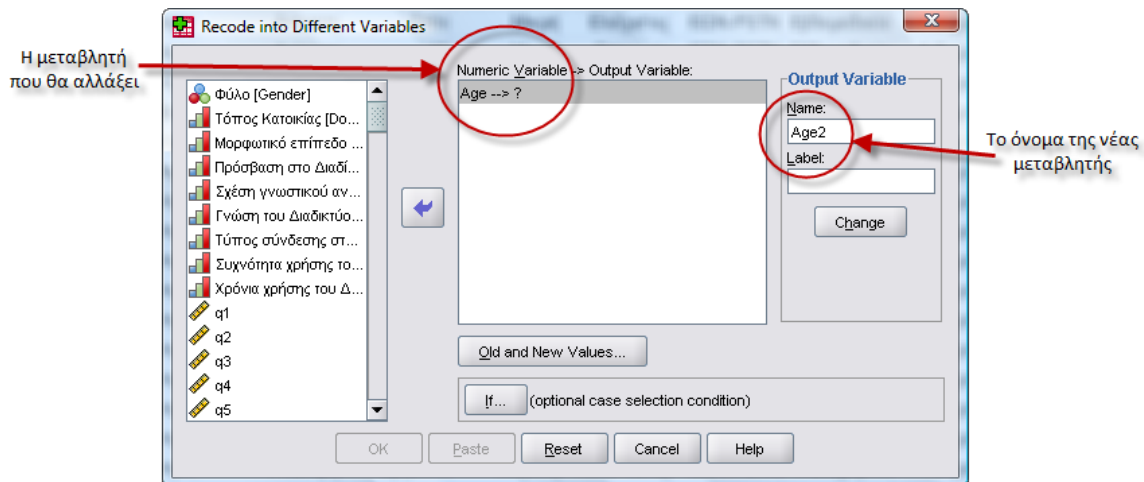


## 5.2 Recode

Μία δεύτερη πολύ χρήσιμη εντολή του SPSS είναι η εντολή **[Recode]**. Για να τη χρησιμοποιήσετε επιλέγετε **[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 διατηρείται αναλλοίωτη και η αρχική μεταβλητή.

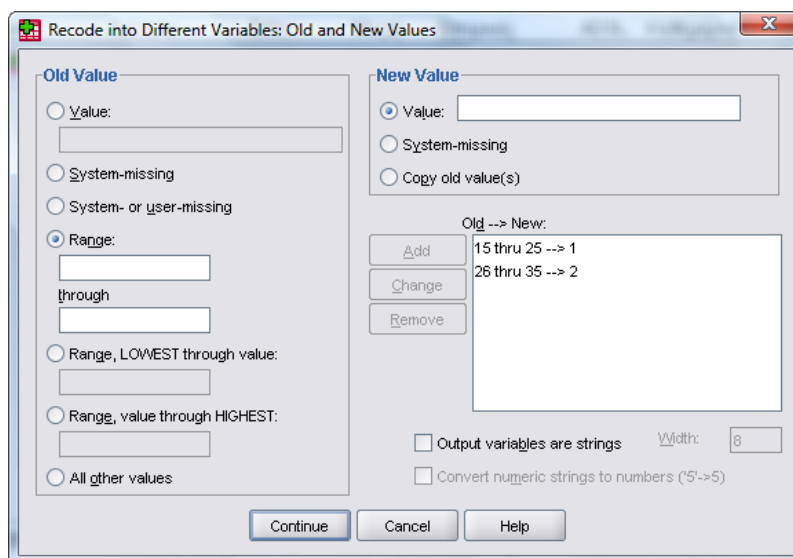
Επιλέγοντας **[Transform --> Recode --> Into Different Variables]** εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 5.2.

Σχήμα 5.2 Η πρώτη οθόνη της εντολής [Recode]



Ας υποθέσουμε ότι θέλετε να μετασηματίσετε τις τιμές της αριθμητικής μεταβλητής ηλικία (όπου οι ηλικίες είναι εκφρασμένες σε αριθμούς) σε κατηγορίες που αντιπροσωπεύουν ηλικιακές ομάδες. Επιλέγετε από τη λίστα των μεταβλητών που υπάρχουν στο αρχείο τη μεταβλητή “Age” και τη μεταφέρετε στο κεντρικό παράθυρο (βλ. Σχήμα 5.2.). Στη συνέχεια βάζετε το όνομα της νέας μεταβλητής που θέλετε να δημιουργήσετε στο πεδίο **[Output Variable]** στο άνω δεξιά τμήμα της οθόνης (Age2 στο παράδειγμά μας) και κάνετε κλικ στο πλήκτρο **[Old and New Values]**. Εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 5.3.

Σχήμα 5.3 Η δεύτερη οθόνη της εντολής [Recode]



Η οθόνη αυτή χωρίζεται σε δύο πεδία. Στα αριστερά υπάρχει το πεδίο **[Old Value]** όπου δίνετε τις συγκεκριμένες τιμές ή το εύρος των τιμών της αρχικής μεταβλητής που θέλετε να αλλάξετε. Στα δεξιά

υπάρχει το πεδίο **[New Value]** όπου καταχωρείτε τις τιμές της νέας μεταβλητής που αντιστοιχούν στις τιμές της παλιάς μεταβλητής που επιλέξατε στα αριστερά. Όταν ολοκληρωθεί αυτή η διαδικασία, κάνετε κλικ στο πλήκτρο **[Add]** και η αντιστοίχιση που επιλέξατε εμφανίζεται στο παράθυρο **[Old→New]**.

Στο παράδειγμα που παρουσιάζεται στο Σχήμα 5.3 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 ολοκληρώσετε τις αντιστοιχίσεις που θέλετε κάντε κλικ στο πλήκτρο **[Continue]** και επιστρέψετε στην πρώτη οθόνη (βλ. Σχήμα 5.2), όπου κάνετε πρώτα κλικ στο πλήκτρο **[Change]** και στη συνέχεια στο πλήκτρο **[OK]**. Έχουμε έτσι δημιουργήσει μία νέα μεταβλητή με το όνομα “Age2”, στην οποία όποιος από τους συμμετέχοντες 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.

### 5.3 Weight Cases

Το SPSS προσφέρει έναν επιπλέον τρόπο μετασχηματισμού δεδομένων μέσω της εντολής **[Weight Cases]**, ο οποίος είναι πολύ χρήσιμος για την καταχώριση 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 χρησιμοποιείτε δευτερογενή δεδομένα ή τουλάχιστον δεν είναι η ευκολότερη.

Ας υποθέσουμε ότι γνωρίζετε πως σε μία έρευνα με δύο μεταβλητές που είχαν δύο κατηγορίες η καθεμία προέκυψαν τα δεδομένα που παρουσιάζονται στον Πίνακα 5.1.

**Πίνακας 5.1 Δευτερογενή δεδομένα**

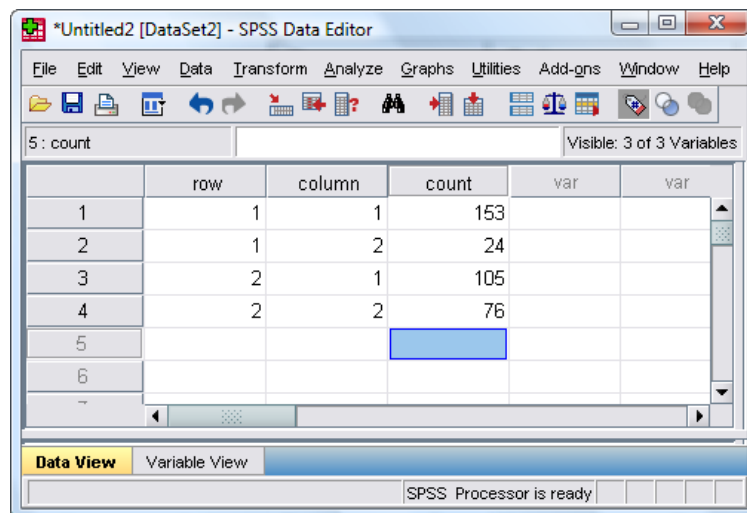
		Μεταβλητή 1		
		A	B	Σύνολα
Μεταβλητή 2	X	153	24	177
	Ψ	105	76	181
Σύνολα		258	100	358

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. Ο παραπάνω πίνακας θα μπορούσε να καταχωριστεί στο SPSS (σε ένα νέο αρχείο δεδομένων) 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

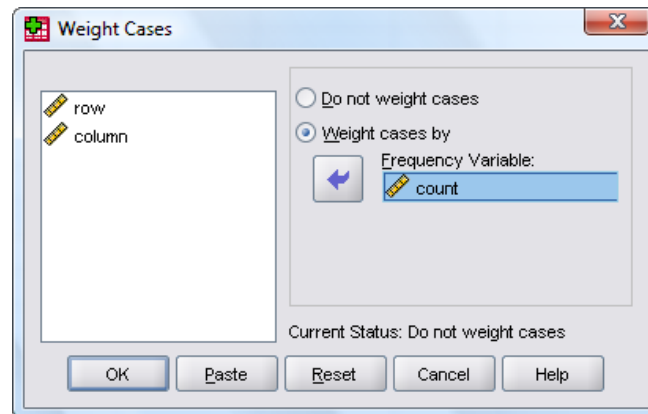
Στις στήλες *Row* και *Column* δημιουργείτε όλους τους πιθανούς συνδυασμούς των κατηγοριών των δύο μεταβλητών χρησιμοποιώντας τους αριθμούς 1 και 2. Στη στήλη *Count* δίνετε την απόλυτη συχνότητα που γνωρίζετε από τον Πίνακα 5.1 (βλ. Σχήμα 5.4).

#### Σχήμα 5.4 Ο Επεξεργαστής Δεδομένων SPSS με τα δεδομένα του Πίνακα 5.1



Στη συνέχεια ενημερώνετε 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 κάνετε αυτό επιλέγετε το **[Data --> Weight Cases]**. Στο παράθυρο που ανοίγει, ενεργοποιείτε την επιλογή **[Weight cases by]**, και μετακινείτε τη μεταβλητή *COUNT* στο κουτί **[Frequency Variable]** (βλ. Σχήμα 5.5). Αν ξεχάσετε αυτό το βήμα, η τιμή σε κάθε κελί θα είναι 1 για τον πίνακα αυτόν.

Σχήμα 5.5 Το παράθυρο της εντολής [Weight Cases]



Αν δε χρησιμοποιούσατε την εντολή **[Weight Cases]** θα έπρεπε να δημιουργήσετε ένα αρχείο δεδομένων με δύο μεταβλητές (Μεταβλητή 1 και Μεταβλητή 2) και σε κάθε μία από αυτές να καταχωρίσετε Α ή Β και Χ ή Ψ αντίστοιχα για κάθε έναν από τους 358 συμμετέχοντες στην έρευνα. Αυτό εκτός από εξαιρετικά χρονοβόρο, θα ήταν εφικτό μόνο αν είχατε στη διάθεσή σας τις απαντήσεις του κάθε συμμετέχοντα ξεχωριστά και όχι έναν συγκεντρωτικό πίνακα, όπως π.χ. ο Πίνακας 5.1, όπως συνήθως συμβαίνει με τα δευτερογενή δεδομένα.

## 6.0 Επαγωγική Στατιστική

Για να αποφασίσετε πώς πρέπει να αναλύσετε δύο ή περισσότερες μεταβλητές, πρέπει να γνωρίζετε την κλίμακα μέτρησής τους. Συνοπτικά ισχύουν τα ακόλουθα:

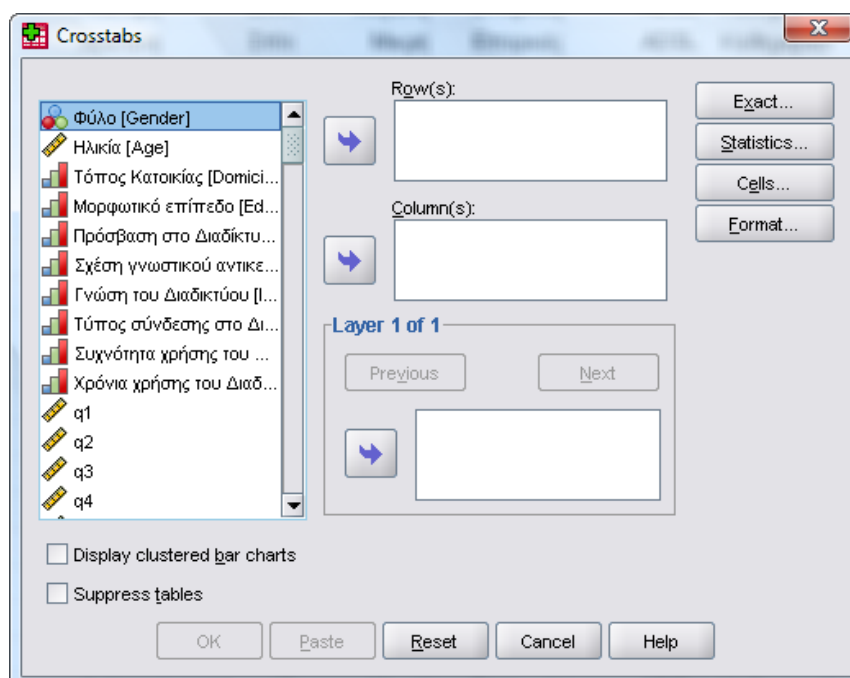
Μεταβλητή (Κλίμακα μέτρησης)	×	Μεταβλητή (Κλίμακα μέτρησης)	Επαγωγικό στατιστικό κριτήριο/τεστ
Κατηγορική (ονομαστική)	×	Κατηγορική (ονομαστική)	$\chi^2$ (συσχέτιση) [μη παραμετρικό]
Κατηγορική (ονομαστική) (μέχρι 2 κατηγορίες)	×	Τακτική (ιεραρχική)	Mann-Whitney U (σύγκριση ομάδων) [μη παραμετρικό]
Κατηγορική (ονομαστική) (πολλές κατηγορίες)	×	Τακτική (ιεραρχική)	Kruskal-Wallis H (σύγκριση ομάδων) [μη παραμετρικό]
Κατηγορική (ονομαστική) (μέχρι 2 κατηγορίες)	×	Αριθμητική (ίσων διαστημάτων)	t-τεστ ανεξάρτητων δειγμάτων (σύγκριση ομάδων) [παραμετρικό]
Κατηγορική (ονομαστική) (πολλές κατηγορίες)	×	Αριθμητική (ίσων διαστημάτων)	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]

## 6.1 Τεστ $\chi^2$

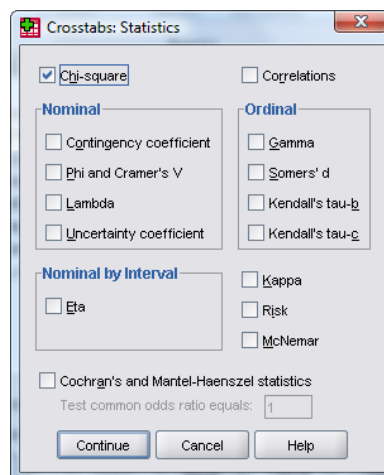
Ο υπολογισμός του τεστ  $\chi^2$  γίνεται μέσω του μενού [Analyze --> Descriptive Statistics --> Crosstabs...]. Η οθόνη που θα εμφανιστεί παρουσιάζεται στο Σχήμα 6.1.

Σχήμα 6.1 Η εντολή crosstabs



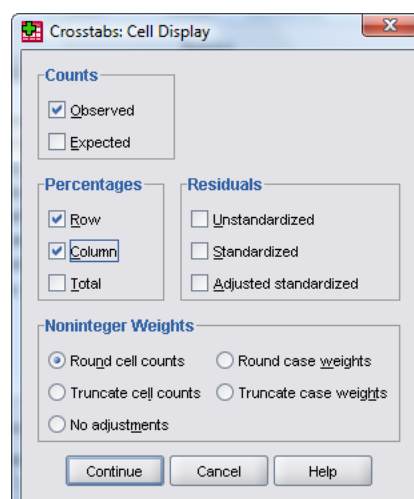
- Στα πλαίσια **[Row(s):]** και **[Column(s):]** βάζετε τις μεταβλητές που θα συμμετάσχουν στην ανάλυση. Συνήθως προτιμούμε να βάλουμε στο κουτί **[Row(s):]** τη μεταβλητή με τις περισσότερες κατηγορίες, ώστε ο πίνακας διπλής εισόδου που θα προκύψει να εκτείνεται καθ' ύψος και όχι κατά πλάτος.
- Στο δεξιό τμήμα του κουτιού βρίσκονται τέσσερα κουμπιά, από τα οποία το σημαντικότερο είναι το κουμπί **[Statistics...]**. Πρέπει να κάνετε κλικ σε αυτό και να επιλέξετε το **[Chi-square]**, αλλιώς το στατιστικό κριτήριο δε θα υπολογιστεί. Από τα περιεχόμενα του κουτιού αυτού προκύπτει ότι μπορείτε να υπολογίσετε και άλλα κριτήρια παράλληλα προς το  $\chi^2$ . Η οθόνη που εμφανίζεται με το κουμπί **[Statistics...]** παρουσιάζεται στο Σχήμα 6.2.

**Σχήμα 6.2** Η εντολή crosstabs – επιλογή statistics



- Αφού επιλέξετε το **[Chi-square]**, κάνετε κλικ στο **[Continue]** για να επιστρέψετε στην αρχική οθόνη (βλ. Σχήμα 6.1). Από εκεί κάνετε κλικ στο πλήκτρο **[Cells...]** και εμφανίζεται η οθόνη του Σχήματος 6.3.

**Σχήμα 6.3** Η εντολή crosstabs – επιλογή cells



- Στην οθόνη του Σχήματος 6.3 είναι σημαντικό να επιλέξετε 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 ολοκληρώσετε τις επιλογές σας, πατάτε πάλι **[Continue]** για να επιστρέψετε στην αρχική οθόνη (βλ. Σχήμα 6.1).
- Στην αρχική οθόνη (Σχήμα 6.1) δίνετε **[OK]** για να εκτελεστεί η ανάλυση. Τα αποτελέσματα της ανάλυσης φαίνονται στους Πίνακες 6.1 και 6.2.

**Πίνακας 6.1 Πίνακας διπλής εισόδου (ή πίνακας σύμπτωσης)**

**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 6.1 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 6.2 παρουσιάζονται τα αποτελέσματα της ανάλυσης. Αυτό που μας ενδιαφέρει βρίσκεται στην πρώτη σειρά: **Pearson Chi-Square**. Στο παράδειγμά 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 6.2 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 παράδειγμα η μεταβλητή race είχε τρεις κατηγορίες) το SPSS υπολογίζει αυτόματα μία επιπλέον σειρά αποτελεσμάτων (βλ. Πίνακα 6.3). Υπάρχει διχογνωμία ως προς το κατά πόσο θα πρέπει να χρησιμοποιείται η διόρθωση αυτή (**Continuity Correction**), αν δηλαδή θα πρέπει να διαβάσουμε αυτήν τη γραμμή ή τη γραμμή **Pearson Chi-Square**, όπως περιγράψαμε αρχικά.
- Ειδικά για τους πίνακες 2x2 κανένα κελί από τα τέσσερα δε 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 6.3 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	37.351	1	.000		
Fisher's Exact Test				.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.

## 6.2 Το t-τεστ

Υπάρχουν δύο ειδών t-τεστ, ένα για ανεξάρτητα δείγματα και ένα για εξαρτημένα δείγματα. Προκειμένου να ενημερώσετε το SPSS για αυτή την πληροφορία, χρειάζεται να ακολουθήσετε δύο διαφορετικούς τρόπους εισαγωγής των δεδομένων. Για τα εξαρτημένα δείγματα, οι μετρήσεις πρέπει να εισαχθούν σε δύο στήλες του Επεξεργαστή Δεδομένων του SPSS (αυτό συμβαίνει γιατί πρόκειται για σύγκριση δύο ομάδων μετρήσεων που προέρχονται από τα ίδια άτομα). Για ανεξάρτητα δείγματα, χρησιμοποιείτε μία στήλη για τις μετρήσεις (την εξαρτημένη μεταβλητή της έρευνας) και μία δεύτερη στήλη στην οποία καθορίζετε την υπο-ομάδα στην οποία ανήκουν οι συμμετέχοντες (οι συνθήκες της ανεξάρτητης μεταβλητής μας). Όπως και στον υπολογισμό του στατιστικού κριτηρίου  $\chi^2$ , είναι σημαντικό να εισάγετε τα δεδομένα σας σωστά.

### Το t-τεστ εξαρτημένων δειγμάτων

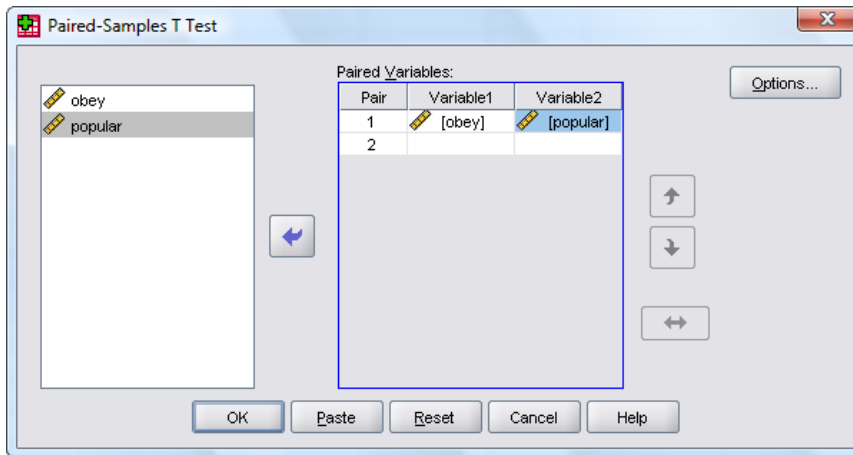
Το t-τεστ για εξαρτημένα δείγματα εφαρμόζεται στις περιπτώσεις που έχετε μετρήσει την ίδια μεταβλητή στους ίδιους συμμετέχοντες σε δύο διαφορετικές χρονικές στιγμές (π.χ., πριν την παρέμβαση και μετά την παρέμβαση) ή όταν έχετε πάρει μετρήσεις από τους ίδιους συμμετέχοντες για δύο μεταβλητές, χρησιμοποιώντας την ίδια κλίμακα μέτρησης. Οι δύο μετρήσεις ή οι δύο μεταβλητές καταχωρίζονται σε ξεχωριστές στήλες όπως φαίνεται στον Πίνακα 6.4.

**Πίνακας 6.4 Καταχώριση 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 εισάγετε τα δεδομένα, πρέπει να επιλέξετε το **[Analyze --> Compare Means --> Paired-Samples T Test...]**. Εμφανίζεται η οθόνη του Σχήματος 6.4.

**Σχήμα 6.4 Το t-τεστ για εξαρτημένα δείγματα**



Η εφαρμογή απαιτεί την επιλογή δύο μεταβλητών: αφού επιλέξετε την πρώτη πατάτε το βελάκι και αυτή πηγαίνει στη θέση **[Variable1]**. Η δεύτερη πηγαίνει στη θέση **[Variable2]**. Βλέπετε ότι μόλις ολοκληρωθεί η συμπλήρωση του ζεύγους των μεταβλητών που θα συγκριθούν, το SPSS δημιουργεί και δεύτερη σειρά για τη σύγκριση και δεύτερου ζεύγους μεταβλητών (στην περίπτωση που θέλετε να πραγματοποιήσετε πολλαπλά t-test εξαρτημένων δειγμάτων). Αφού τελειώσετε με τη συμπλήρωση των μεταβλητών, πατάτε το **[OK]** και παίρνετε τα εξής αποτελέσματα στο παράθυρο Viewer (βλ. Πίνακα 6.5):

**Πίνακας 6.5 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 δίνει περιγραφικούς δείκτες (οι μέσοι όροι και οι τυπικές αποκλίσεις είναι οι σημαντικότεροι) και ο δεύτερος τη συνάφεια για τις δύο μεταβλητές (ή για τις δύο μετρήσεις της ίδιας μεταβλητής). Ο τρίτος πίνακας δίνει την τιμή του t (στήλη t), τους βαθμούς ελευθερίας (στήλη df) και το p [στήλη Sig.(2-tailed)]. Σύμφωνα με τους κανόνες μορφοποίησης της APA, το αποτέλεσμα αυτό θα παρουσιαστεί ως εξής: **t(981)=-25,13, p<.001**. Η τιμή του p δείχνει ότι στο παράδειγμά μας υπάρχει στατιστικώς σημαντική διαφορά μεταξύ των μέσων όρων των δύο μεταβλητών.

**Σημείωση:** το SPSS σε πολλές περιπτώσεις δεν επιτρέπει να καθορίσετε αν η υπόθεση που θα ελέγξει είναι μονής ή διπλής κατεύθυνσης, αλλά προχωράει στον έλεγχο υπόθεσης διπλής κατεύθυνσης. Στις περιπτώσεις εκείνες όπου έχετε διατυπώσει υπόθεση μονής κατεύθυνσης, θα πρέπει να προσαρμόσετε κατάλληλα το επίπεδο στατιστικής σημαντικότητας. Έτσι, αν για παράδειγμα το **Sig. (2-tailed)** που υπολόγισε το SPSS ήταν 0,043 (στατιστικά σημαντικό), θα πρέπει να το προσαρμόσετε διαιρώντας δια δύο ( $0,043/2 = 0,0215$ ).

### Το t-τεστ για ανεξάρτητα δείγματα

Το t-τεστ για ανεξάρτητα δείγματα εφαρμόζεται στις περιπτώσεις που θέλετε να συγκρίνετε δύο ομάδες μετρήσεων από διαφορετικά άτομα (δείγματα) ως προς μία εξαρτημένη μεταβλητή (π.χ., την ομάδα των αντρών και την ομάδα των γυναικών ως προς μία εξαρτημένη μεταβλητή). Περνάτε τα δεδομένα σας στον Επεξεργαστή Δεδομένων του SPSS ως εξής: σε μία στήλη περνάτε τη μεταβλητή ομαδοποίησης (ανεξάρτητη) για να δηλώσετε την ομάδα στην οποία ανήκουν οι συμμετέχοντες και σε μία δεύτερη στήλη (εξαρτημένη μεταβλητή) καταχωρίζετε τη μέτρηση που έχετε για κάθε συμμετέχοντα, όπως φαίνεται στον Πίνακα 6.6.

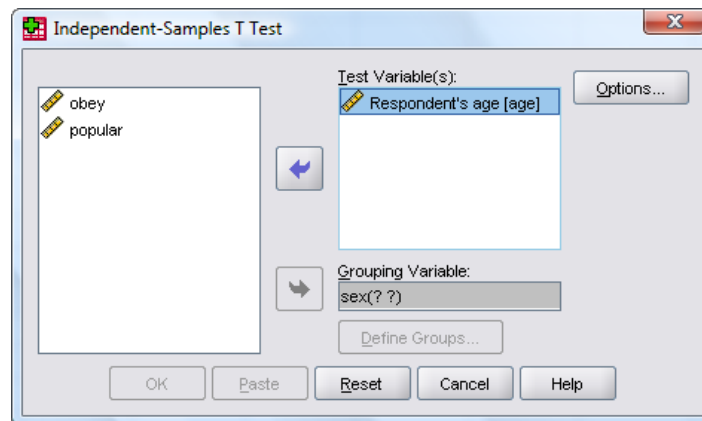
**Πίνακας 6.6 Καταχώριση δεδομένων για το t-τεστ ανεξάρτητων δειγμάτων**

Group	Exp_Con
1	96
1	127
1	127
1	119

2	109
2	143
2	104
2	96
...	...
...	...
2	114

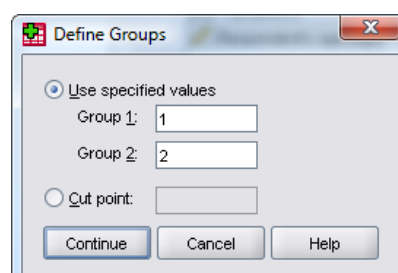
Για να υπολογίσετε το στατιστικό κριτήριο  $t$ , κάντε κλικ στα **[Analyze --> Compare Means --> Independent-Samples T Test]** και εμφανίζεται η οθόνη του Σχήματος 6.5.

**Σχήμα 6.5 Το  $t$ -τεστ για ανεξάρτητα δείγματα**



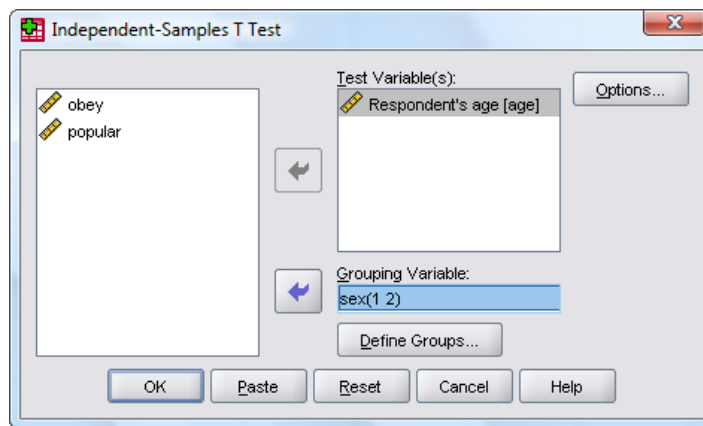
- Στο κουτί **[Test Variable(s):]** βάζετε τη μεταβλητή που περιέχει τις μετρήσεις σας (στο παράδειγμά μας την ηλικία των συμμετεχόντων).
- Στο κουτί **[Grouping Variable:]** βάζετε τη μεταβλητή ομαδοποίησης (στο παράδειγμά μας το φύλο των συμμετεχόντων). Όπως βλέπετε, στο Σχήμα 6.5 εμφανίζονται δύο λατινικά ερωτηματικά δίπλα από το όνομα της μεταβλητής ομαδοποίησης. Στα ερωτηματικά αυτά πρέπει να αντιστοιχήσετε τους κωδικούς των ομάδων (αν δεν τους θυμόσαστε ανατρέξτε στην προβολή **[Variable View]**). Επιλέγετε τη μεταβλητή ομαδοποίησης και κάνετε κλικ στο πλήκτρο **[Define Groups...]**, με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 5.6.

**Σχήμα 6.6 Καθορισμός των ομάδων**



- Στην οθόνη του Σχήματος 6.6 πρέπει να δηλώσετε πώς κωδικοποιήθηκαν οι δύο ομάδες στην μεταβλητή ομαδοποίησης που χρησιμοποιήσατε. 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 δηλώσετε τις τιμές, επιλέγετε το **[Continue]** για να επιστρέψετε στην αρχική οθόνη (βλ. Σχήμα 6.7), όπου τα λατινικά ερωτηματικά έχουν αντικατασταθεί με τους κωδικούς 1 και 2.

**Σχήμα 6.7 Το t-τεστ για ανεξάρτητα δείγματα μετά τον καθορισμό των ομάδων**



- Πατώντας **[OK]** γίνεται η επεξεργασία των δεδομένων και παίρνετε στο παράθυρο Viewer τα ακόλουθα (βλ. Πίνακα 6.7):

**Πίνακας 6.7 Αποτελέσματα του t-τεστ ανεξάρτητων δειγμάτων**

**Group Statistics**

Respondent's Sex		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 δίνει περιγραφικά στοιχεία για τις δύο ομάδες που συγκρίνετε, δηλαδή δίνει ξεχωριστά το μέσο όρο και την τυπική απόκλιση για τους άντρες και τις γυναίκες.

- Ο δεύτερος πίνακας παρουσιάζει τα αποτελέσματα της ανάλυσης. Οι δύο πρώτες στήλες παρουσιάζουν τα αποτελέσματα του κριτηρίου Levene για την ισότητα των διακυμάνσεων των δύο ομάδων, που είναι και μία από τις προϋποθέσεις για την εφαρμογή των παραμετρικών στατιστικών κριτηρίων (και του κριτηρίου t). Για να μπορέσετε να εφαρμόσετε το t-τεστ θα πρέπει το αποτέλεσμα από το κριτήριο Levene να είναι στατιστικά μη σημαντικό (υποδηλώνοντας με αυτό τον τρόπο ότι οι δύο διακυμάνσεις είναι ίσες - δηλαδή το Sig. θα πρέπει να είναι μεγαλύτερο από το 0,05). Αν το αποτέλεσμα του κριτηρίου Levene είναι στατιστικώς σημαντικό (όπως συμβαίνει στο παράδειγμά μας), τότε θα πρέπει να διαβάσετε τη δεύτερη γραμμή του πίνακα (equal variances not assumed). Το SPSS κάνει μια διόρθωση των τιμών του t-τεστ και σε αυτή την περίπτωση θα αναφέρατε: **t(1420,93)=-2,72, p=,007**. Επομένως, η διαφορά των μέσων όρων της ηλικίας των αντρών και των γυναικών είναι στατιστικώς σημαντική (p=0,007).
- Ορισμένοι προτιμούν να μην κάνουν χρήση της δεύτερης γραμμής του πίνακα και θεωρούν ότι από τη στιγμή που δεν υπάρχουν ίσες διακυμάνσεις θα πρέπει να χρησιμοποιηθεί το μη παραμετρικό ισοδύναμο του t-τεστ (βλ. Mann-Whitney U). Άλλοι πάλι, αγνοούν τελείως το κριτήριο Levene και υποστηρίζουν 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 ισοδύναμό του.

### 6.3 Ανάλυση διακύμανσης μονής κατεύθυνσης (One-way ANOVA)

Η ανάλυση διακύμανσης μονής κατεύθυνσης εφαρμόζεται στις περιπτώσεις που συγκρίνονται μετρήσεις που προέρχονται από περισσότερες από δύο ομάδες (δείγματα) ως προς μία εξαρτημένη μεταβλητή (π.χ., τις τρεις ομάδες που προκύπτουν από τη φυλή των συμμετεχόντων ως προς μία εξαρτημένη μεταβλητή). Όπως αναφέρθηκε παραπάνω και στο t-τεστ, σε μια στήλη περνάτε τα στοιχεία της μεταβλητής ομαδοποίησης για να δηλώσετε την ομάδα στην οποία ανήκουν οι συμμετέχοντες και σε μία δεύτερη στήλη καταχωρίζετε τις μετρήσεις στην εξαρτημένη μεταβλητή που σας ενδιαφέρει για κάθε συμμετέχοντα, όπως φαίνεται στον Πίνακα 6.8.

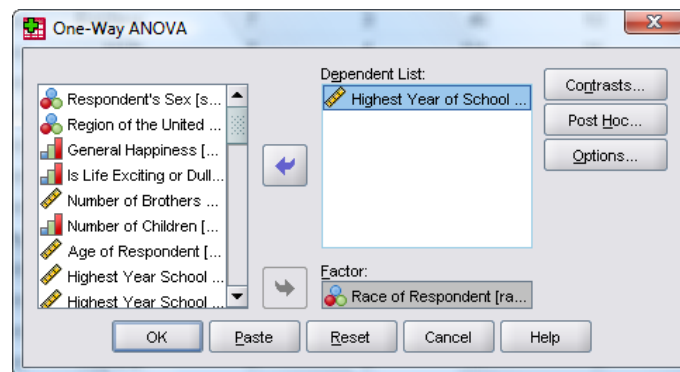
**Πίνακας 6.8 Καταχώριση δεδομένων για την ανάλυση διακύμανσης μονής κατεύθυνσης**

Groups	Scores
1	9
3	8
3	6
...	...
1	7

2	7
3	9
3	6
...	...
...	...
...	...
2	10
2	19
...	...
1	11

Για να κάνετε ανάλυση διακύμανσης μονής κατεύθυνσης, κάντε κλικ στα **[Analyze --> Compare Means --> One-Way ANOVA]** και εμφανίζεται η οθόνη του Σχήματος 6.8.

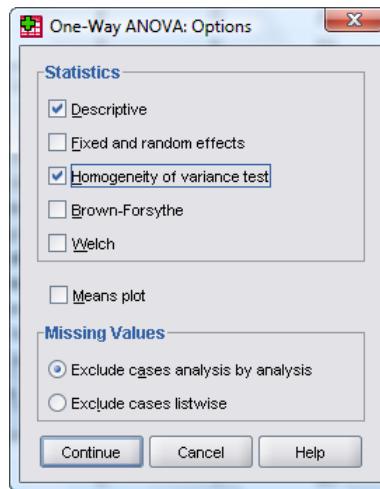
**Σχήμα 6.8** Ανάλυση διακύμανσης μονής κατεύθυνσης



- Βάζετε στο **[Dependent list:]** τη μεταβλητή που περιλαμβάνει τις μετρήσεις σας (στο παράδειγμα, τα έτη συστηματικής εκπαίδευσης) και στο **[Factor:]** τη μεταβλητή ομαδοποίησης (την ανεξάρτητη μεταβλητή σας). Προσέξτε ότι σε αυτήν την περίπτωση δε χρειάζεται να ορίσετε το πώς έχετε κωδικοποιήσει τις επιμέρους ομάδες.
- Κάντε κλικ στο πλήκτρο **[Options...]** και εμφανίζεται η οθόνη του Σχήματος 6.9. Επιλέγετε **[Descriptive]** προκειμένου να πάρετε περιγραφικά στατιστικά για τις επιμέρους ομάδες και **[Homogeneity of variance test]** για να πάρετε το κριτήριο Levene. Στη συνέχεια επιλέγετε **[Continue]** για να επιστρέψετε στην αρχική οθόνη (Σχήμα 6.8).



Σχήμα 6.9 Ανάλυση διακύμανσης μονής κατεύθυνσης – επιλογή Options



- Πατήστε το **[OK]** για να γίνει η ανάλυση. Τα αποτελέσματα θα είναι ως εξής (βλ. Πίνακα 6.9):

Πίνακας 6.9 Αποτελέσματα για την ανάλυση διακύμανσης μονής 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

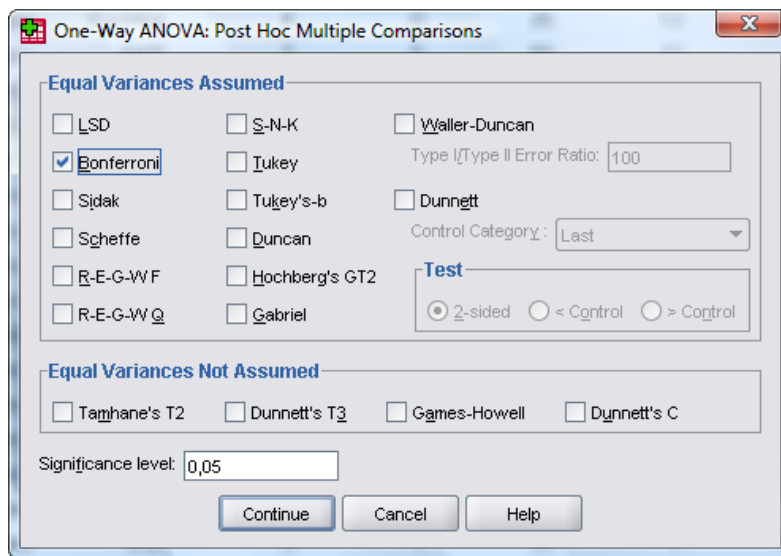
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	240,725	2	120,362	13,746	,000
Within Groups	13195,994	1507	8,756		
Total	13436,719	1509			

- 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 κριτήριο 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 κριτήριο Levene είναι στατιστικώς μη σημαντικό. Στο παράδειγμά μας η προϋπόθεση της ισότητας των διακυμάνσεων δεν ικανοποιείται και κατά συνέπεια θα πρέπει να χρησιμοποιήσετε 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 κριτήριο 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 ανάλυση του παραδείγματός μας είναι έγκυρη. Το αποτέλεσμα θα παρουσιαστεί ως εξής: **F(2, 1507)=13,75, p<.001**. Επομένως, υπάρχει στατιστικώς σημαντική επίδραση της ανεξάρτητης μεταβλητής στην εξαρτημένη. Αυτό σημαίνει ότι 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 ό,τι συνέβαινε στο t-τεστ, εδώ συγκρίνονται περισσότερες από δύο ομάδες και κατά συνέπεια δεν γνωρίζουμε πού 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 διερευνάται η ύπαρξη πιθανών διαφορών σε όλους τους δυνατούς συνδυασμούς ζευγών ομάδων. Όπως αναφέρθηκε προηγουμένως, 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 ορισμένες μη σημαντικές διαφορές να εκληφθούν ως στατιστικώς σημαντικές, για τυχαίους λόγους. Για το λόγο αυτό κατά τις εκ των υστέρων κατά ζεύγη συγκρίσεις εφαρμόζεται κάποια «διόρθωση»

που λαμβάνει υπόψη τον αριθμό των συγκρίσεων κατά τον υπολογισμό της στατιστικής σημαντικότητας. Το SPSS προσφέρει μεγάλο αριθμό διορθώσεων, καθεμία από τις οποίες έχει πλεονεκτήματα ανάλογα με τις ανάγκες της ανάλυσής μας. Η απλούστερη διόρθωση που μπορείτε να εφαρμόσετε είναι η διόρθωση Bonferroni.

- Για να ζητήσετε από το SPSS κατά ζεύγη συγκρίσεις, κάντε κλικ στο πλήκτρο **[Post Hoc]** από την αρχική οθόνη, το οποίο ανοίγει το ακόλουθο παράθυρο διαλόγου (βλ. Σχήμα 6.10):

**Σχήμα 6.10** Ανάλυση διακύμανσης μονής κατεύθυνσης – επιλογή Post Hoc



- Επιλέγετε τη διόρθωση **Bonferroni** και πατάτε **[Continue]** για να επιστρέψετε στην αρχική οθόνη. Δίνοντας **[OK]** εκτελείται η ανάλυση δίνοντας τα ακόλουθα επιπλέον αποτελέσματα (βλ. Πίνακα 6.10).

**Πίνακας 6.10** 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 δίνεται η στατιστική σημαντικότητα. Από την επισκόπηση του Πίνακα 6.10 προκύπτει ότι στατιστικώς σημαντική διαφορά υπάρχει μόνο μεταξύ των κατηγοριών White και Other και όχι για τους υπόλοιπους συνδυασμούς κατηγοριών.

#### 6.4 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων (Repeated measures ANOVA)

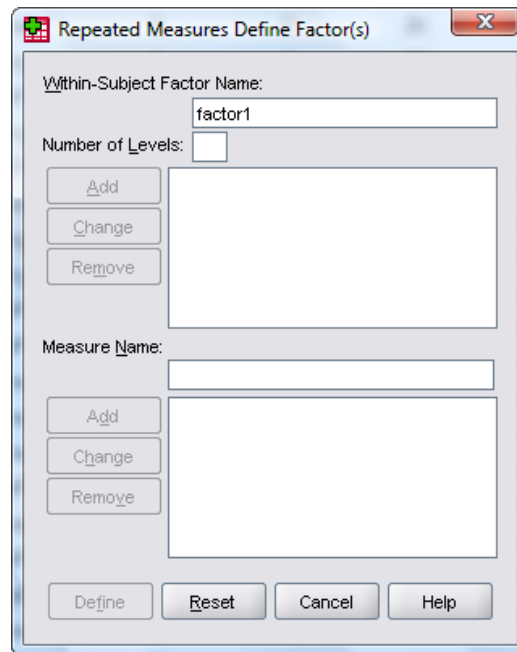
Η ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων εφαρμόζεται στις περιπτώσεις που έχετε μετρήσει την ίδια μεταβλητή στους ίδιους συμμετέχοντες σε περισσότερες από δύο διαφορετικές χρονικές στιγμές (π.χ. πριν την παρέμβαση, στο μέσο της παρέμβασης και μετά την παρέμβαση) ή όταν έχετε πάρει μετρήσεις 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 6.11.

**Πίνακας 6.11 Καταχώριση 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

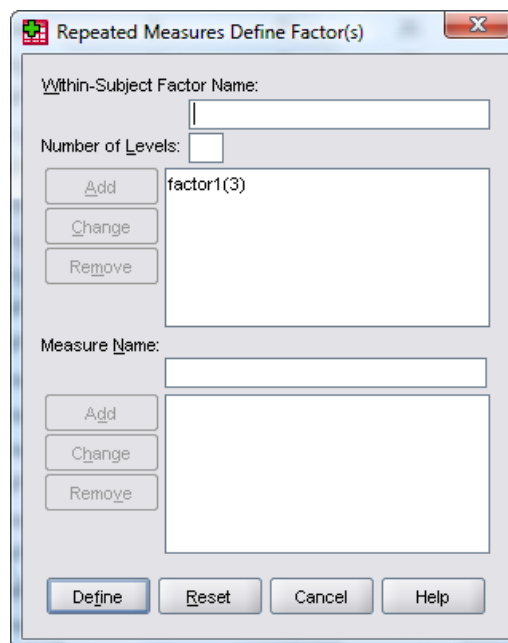
Για να κάνετε ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων κάνετε κλικ στα **[Analyze --> General Linear Model --> Repeated Measures...]** και εμφανίζεται η οθόνη του Σχήματος 6.11.

Σχήμα 6.11 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – αρχική οθόνη



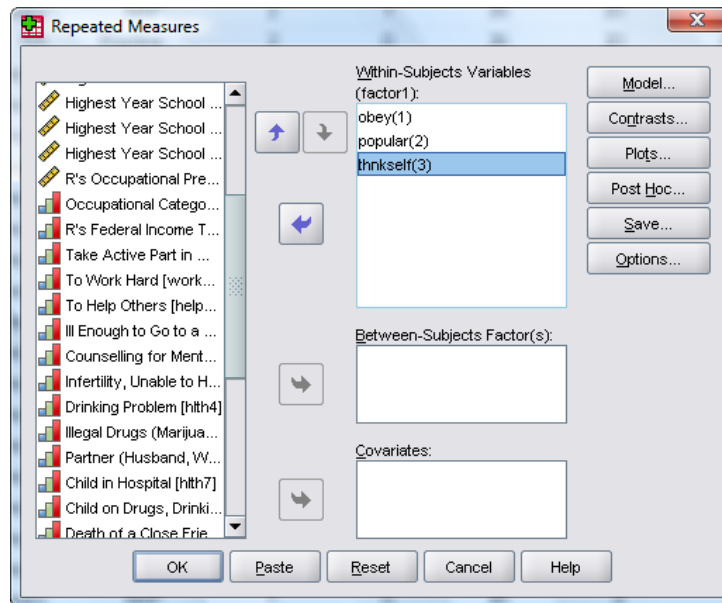
- Στην αρχική οθόνη προσθέτετε τον αριθμό των συνθηκών της ανεξάρτητης μεταβλητής που θέλετε να συγκρίνετε στο πεδίο **[Number of Levels]** και κάνουμε κλικ στο πλήκτρο **[Add]**. Στο παράδειγμα αυτό θα επιλέξετε τρία επίπεδα και η οθόνη θα έχει την ακόλουθη μορφή (βλ. Σχήμα 6.12).

Σχήμα 6.12 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – καθορισμός επιπέδων



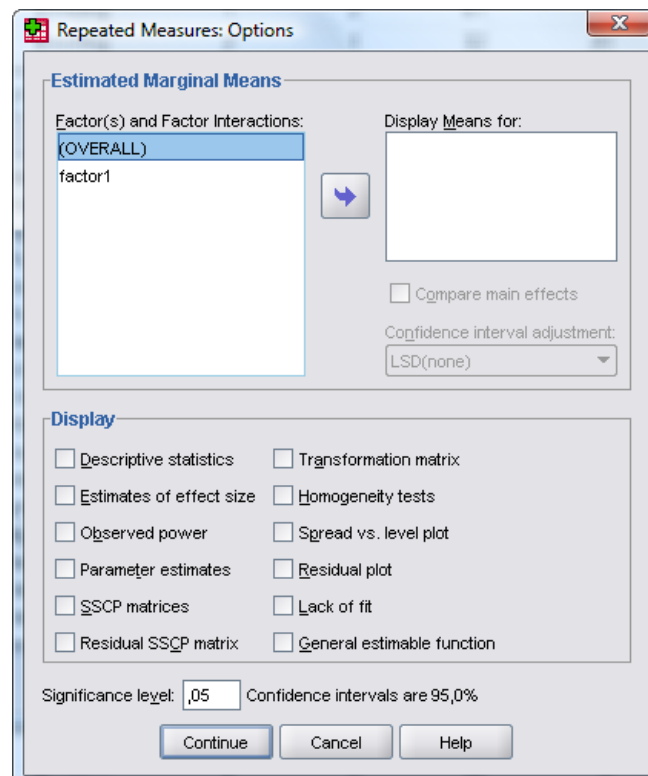
- Στη συνέχεια κάνετε κλικ στο πλήκτρο **[Define]** και εμφανίζεται η ακόλουθη οθόνη (βλ. Σχήμα 6.13).

Σχήμα 6.13 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων



- Στην οθόνη του σχήματος 5.12 βάζετε στη λίστα **[Within-Subjects Variables (factor1):]** τις συνθήκες της ανεξάρτητης μεταβλητής που επιθυμείτε να συγκρίνετε. Στη συνέχεια κάνετε κλικ στο πλήκτρο **[Options]** και εμφανίζεται η ακόλουθη οθόνη (βλ. Σχήμα 6.14).

Σχήμα 6.14 Ανάλυση διακύμανσης επαναλαμβανόμενων μετρήσεων – επιλογή Options



- Στην οθόνη του σχήματος 6.14 επιλέγετε **[Descriptive Statistics]** από την επιλογή **[Display]** και πατάτε **[Continue]** για να επιστρέψετε στην προηγούμενη οθόνη. Εκεί δίνετε **[OK]** για να τρέξετε την ανάλυση. Τα αποτελέσματα που προκύπτουν είναι αρκετά εκτεταμένα, για αυτό στον Πίνακα 6.12 παρουσιάζουμε επιλεκτικά ένα μέρος τους.

**Πίνακας 6.12 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 οι περιγραφικοί στατιστικοί δείκτες για τις συνθήκες της ανεξάρτητης μεταβλητής που συμμετέχει 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 κοιτάτε τη σειρά

Greenhouse-Geisser και όχι τη σειρά Sphericity Assumed. Στο παράδειγμά μας η συγκεκριμένη προϋπόθεση δεν πληρείται.

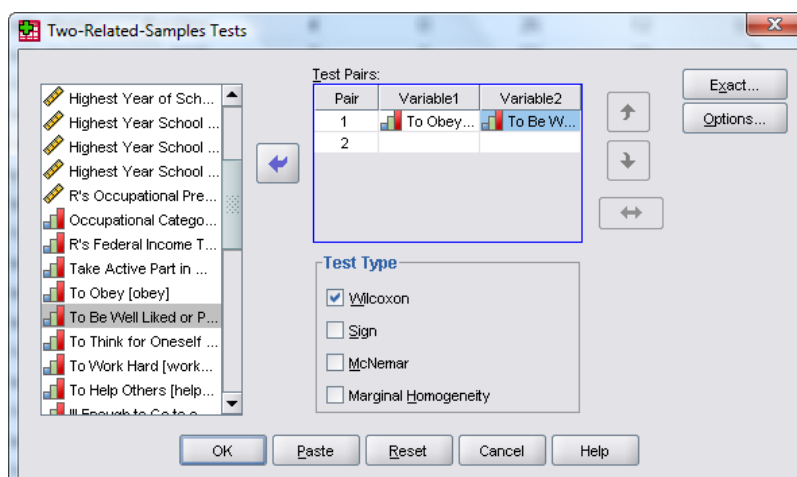
- Στον τρίτο πίνακα παρουσιάζονται τα αποτελέσματα της ανάλυσης διακύμανσης. Επειδή δεν πληρείται το κριτήριο της σφαιρικότητας, κοιτάτε τη δεύτερη γραμμή για τον factor 1 και διαπιστώνετε ότι οι διαφορές των τριών μέσων όρων είναι στατιστικώς σημαντικές ( $p < 0,001$ ).
- Επειδή συγκρίνετε περισσότερους από δύο μέσους όρους ανακύπτει και πάλι το ερώτημα σχετικά με το πού ακριβώς υπάρχουν στατιστικώς σημαντικές διαφορές. Η λύση είναι και πάλι να κάνετε εκ των υστέρων κατά ζεύγη συγκρίσεις, αλλά σε αυτήν την περίπτωση το SPSS δεν προσφέρει κάποια έτοιμη εντολή. Μπορείτε όμως να κάνετε 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 σας. Η διόρθωση Bonferroni γίνεται εύκολα με το χέρι αν διαιρέσετε το επίπεδο στατιστικής σημαντικότητας με τον αριθμό των συγκρίσεων που θα κάνετε. Στο παράδειγμα αυτό χρησιμοποιείτε τρεις μεταβλητές, άρα έχετε 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 χρησιμοποιήσετε είναι  $0,05 / 3 = 0,17$  περίπου. Αυτό σημαίνει ότι θεωρείτε στατιστικώς σημαντικές τις εκ των υστέρων συγκρίσεις όπου  $p < 0,017$ .

## 6.5 Wilcoxon

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

- Για να εκτελέσετε το τεστ επιλέγετε **[Analyze --> Nonparametric Tests --> 2 Related Samples]** και εμφανίζεται η οθόνη του Σχήματος 6.15:

Σχήμα 6.15 Τεστ Wilcoxon





- Η εφαρμογή απαιτεί την επιλογή δύο μεταβλητών: αφού επιλέξετε την πρώτη πατάτε το βελάκι και αυτή πηγαίνει στη θέση **[Variable1]**. Η δεύτερη πηγαίνει στη θέση **[Variable2]**. Βλέπετε ότι μόλις ολοκληρωθεί η συμπλήρωση του ζεύγους των μεταβλητών που θα συγκριθούν, το SPSS δημιουργεί και δεύτερη σειρά για τη σύγκριση και δεύτερου ζεύγους μεταβλητών (στην περίπτωση που θέλετε να πραγματοποιήσετε πολλαπλά κριτήρια Wilcoxon). Αφού τελειώσετε με τη συμπλήρωση των μεταβλητών, πατάτε το **[OK]** και παίρνετε τα εξής αποτελέσματα στο παράθυρο Viewer (βλ. Πίνακα 6.13):

**Πίνακας 6.13 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

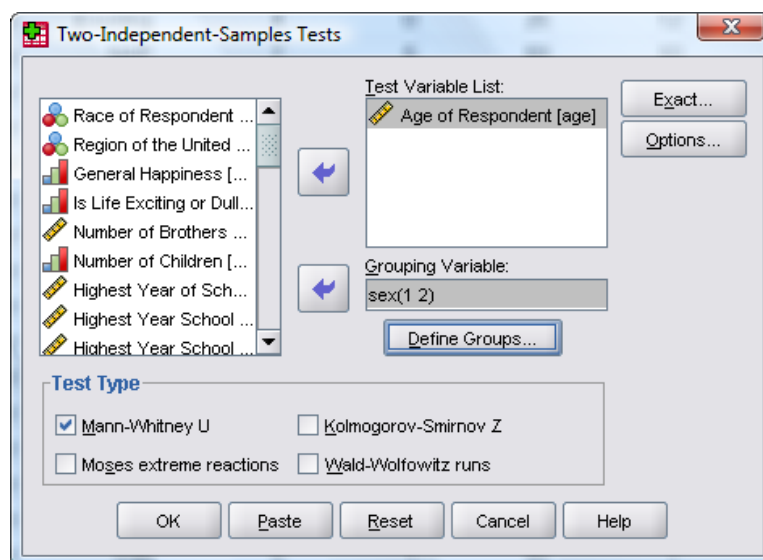
- Στον πρώτο πίνακα παίρνετε περιγραφικούς στατιστικούς δείκτες για τις δύο μεταβλητές ή για τις δύο μετρήσεις της ίδιας μεταβλητής. Το SPSS δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα δίνεται η τιμή του κριτηρίου και το p. Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική ( $p < 0,001$ ).
- 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 μπορείτε να τα πάρετε ακολουθώντας τα όσα περιγράφονται στο κεφάλαιο για τους δείκτες κεντρικής τάσης και διασποράς.

## 6.6 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 έχετε καταχωρίσει τα δεδομένα σας με το ίδιο τρόπο.

- Για να εκτελέσετε το τεστ επιλέγετε **[Analyze --> Nonparametric Tests --> 2 Independent Samples]** και εμφανίζεται η οθόνη του Σχήματος 6.16:

**Σχήμα 6.16 Τεστ Mann-Whitney U**



- Στο πλαίσιο **[Test Variable List:]** βάζετε την εξαρτημένη μεταβλητή (στο παράδειγμά μας την ηλικία των συμμετεχόντων).
- Στο πλαίσιο **[Grouping Variable:]** βάζετε τη μεταβλητή ομαδοποίησης (την ανεξάρτητη μεταβλητή, στο παράδειγμά μας το φύλο των συμμετεχόντων) και ακολουθείτε τη διαδικασία που περιγράφηκε στο t-τεστ ανεξάρτητων ομάδων προκειμένου να ορίσετε τις ομάδες με τη χρήση του πλήκτρου **[Define Groups]**.
- Στη συνέχεια δίνετε **[OK]** και παίρνετε στο παράθυρο Viewer τα αποτελέσματα που παρουσιάζονται στον Πίνακα 6.14.

Πίνακας 6.14 Αποτελέσματα για το Τεστ Mann-Whitney U

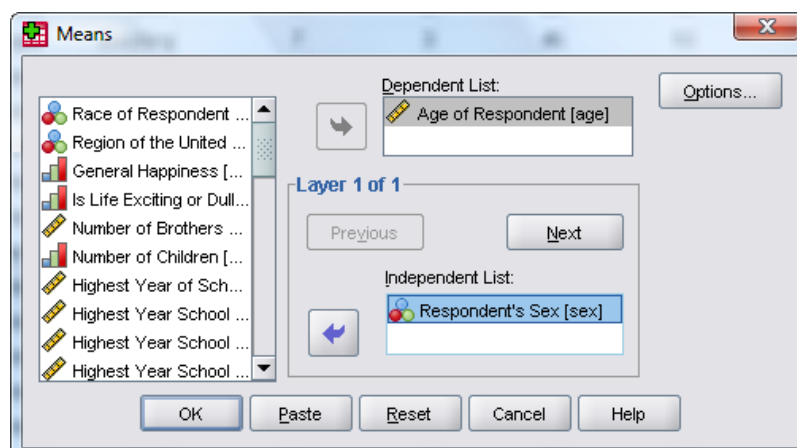
Ranks				
	Respondent's Sex	N	Mean Rank	Sum of Ranks
Age of Respondent	Male	636	726,92	462319,50
	Female	878	779,65	684535,50
	Total	1514		

Test Statistics <sup>a</sup>	
	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

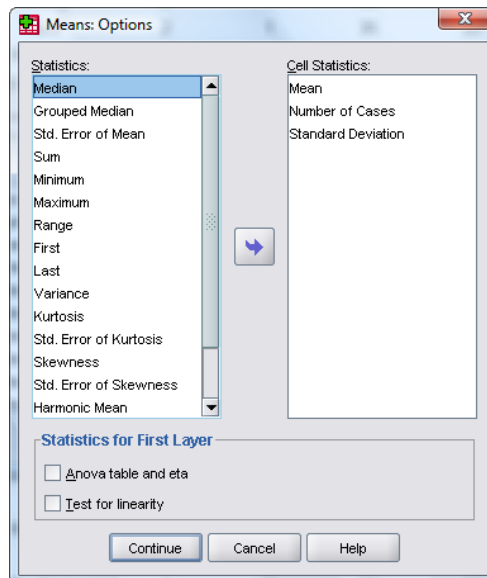
- Στον πρώτο πίνακα παρουσιάζονται περιγραφικά στατιστικά για τις δύο ομάδες. Το SPSS δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα δίνεται η τιμή του κριτηρίου (πρώτη γραμμή) και το  $p$ . Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική ( $p=0,021$ ).
- 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 μπορείτε να τα πάρετε χρησιμοποιώντας την εντολή **[Means]**.
- Για να χρησιμοποιήσουμε την εντολή αυτή επιλέγετε **[Analyze --> Compare Means --> Means]** και εμφανίζεται η οθόνη που παρουσιάζεται στο Σχήμα 6.17:

Σχήμα 6.17 Η εντολή Means



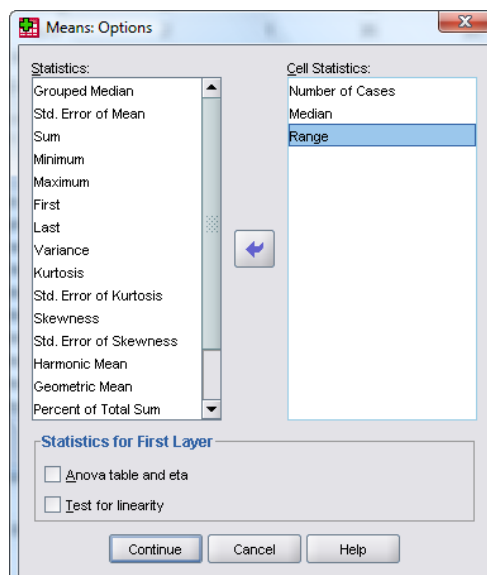
- Βάζετε την εξαρτημένη μεταβλητή στο πλαίσιο [**Dependent List:**] και την ανεξάρτητη μεταβλητή στο παράθυρο [**Independent List:**] και στη συνέχεια κάνετε κλικ στο πλήκτρο [**Options**] με αποτέλεσμα να εμφανιστεί η οθόνη του Σχήματος 6.18:

**Σχήμα 6.18** Η εντολή Means – επιλογή Options



- Από τη λίστα [**Cell Statistics:**] αφαιρείτε τις επιλογές [**Mean**] και [**Standard Deviation**] και προσθέτετε τις επιλογές [**Median**] και [**Range**]. Για να το επιτύχετε αυτό κάνετε κλικ στην κάθε επιλογή και στη συνέχεια κάνετε κλικ στο βελάκι μεταξύ των δύο λιστών. Η οθόνη σας θα πρέπει να έχει τη μορφή που φαίνεται στο Σχήμα 6.19:

**Σχήμα 6.19** Η εντολή Means – επιλογή Options (II)



- Στη συνέχεια κάνετε κλικ στο πλήκτρο **[Continue]** και επιστρέφετε στην αρχική οθόνη, όπου δίνετε **[OK]** και λαμβάνετε στο παράθυρο Viewer όσα παρουσιάζονται στον Πίνακα 6.15:

**Πίνακας 6.15 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

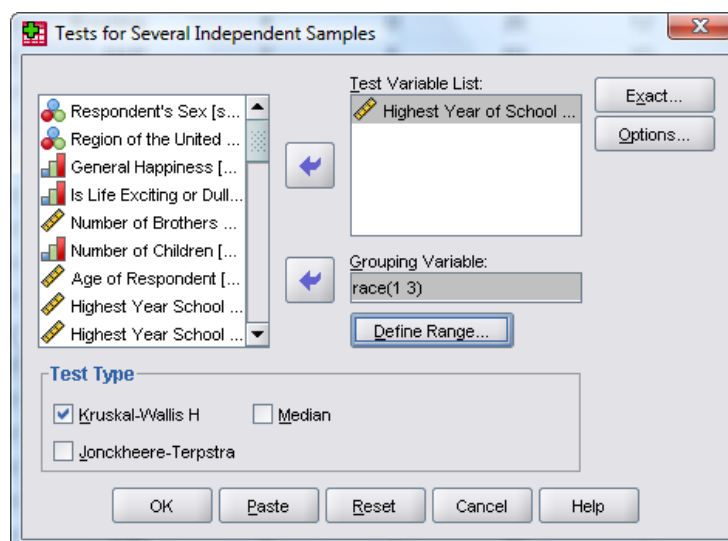
- Βλέπετε ότι πήρατε τη διάμεσο και το εύρος για κάθε μία από τις κατηγορίες της μεταβλητής ομαδοποίησης, αλλά και για το σύνολο του δείγματος.

## 6.7 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 έχετε καταχωρίσει τα δεδομένα σας με τον ίδιο τρόπο.

Για να εκτελέσετε το τεστ επιλέγετε **[Analyze --> Nonparametric Tests --> K Independent Samples]** και εμφανίζεται η οθόνη του Σχήματος 6.20:

**Σχήμα 6.20 Kruskal-Wallis H**



- Στο παράθυρο **[Test Variable List:]** βάζετε την εξαρτημένη μεταβλητή (στο παράδειγμά μας τα έτη συστηματικής εκπαίδευσης).

- Στο παράθυρο [**Grouping Variable:**] βάζετε την ανεξάρτητη μεταβλητή (στο παράδειγμά μας τη φυλή των συμμετεχόντων). Όπως κάνατε και στο t-τεστ για εξαρτημένα δείγματα, πρέπει να ορίσετε τους κωδικούς που αντιστοιχούν στις ομάδες (μέσω του κουμπιού [**Define Range...**]).
- Πατώντας [**OK**] γίνεται η επεξεργασία των δεδομένων και παίρνετε στο παράθυρο Viewer τα ακόλουθα (βλ. Πίνακα 6.16):

**Πίνακας 6.16 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 παίρνετε περιγραφικούς στατιστικούς δείκτες για τις επιμέρους ομάδες. Το SPSS δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα δίνεται η τιμή του κριτηρίου και το p. Στο παράδειγμά μας η διαφορά είναι στατιστικώς σημαντική ( $p < 0,001$ ).
- 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 μπορείτε να τα πάρετε χρησιμοποιώντας την εντολή **Means**, ακολουθώντας τη διαδικασία που περιγράφηκε στην ενότητα του κριτηρίου Mann-Whitney U.
- Επειδή συγκρίνετε μετρήσεις από περισσότερες από δύο ομάδες αντιμετωπίζετε και πάλι το ερώτημα πού ακριβώς υπάρχουν στατιστικώς σημαντικές διαφορές. Η λύση είναι και πάλι να κάνετε εκ των υστέρων κατά ζεύγη συγκρίσεις, αλλά σε αυτήν την περίπτωση το SPSS δεν προσφέρει κάποια έτοιμη εντολή. Μπορείτε όμως να κάνετε εύκολα τις εκ των υστέρων συγκρίσεις χρησιμοποιώντας όσα περιγράφονται για το Mann-Whitney U για όλους τους πιθανούς συνδυασμούς των επιμέρους

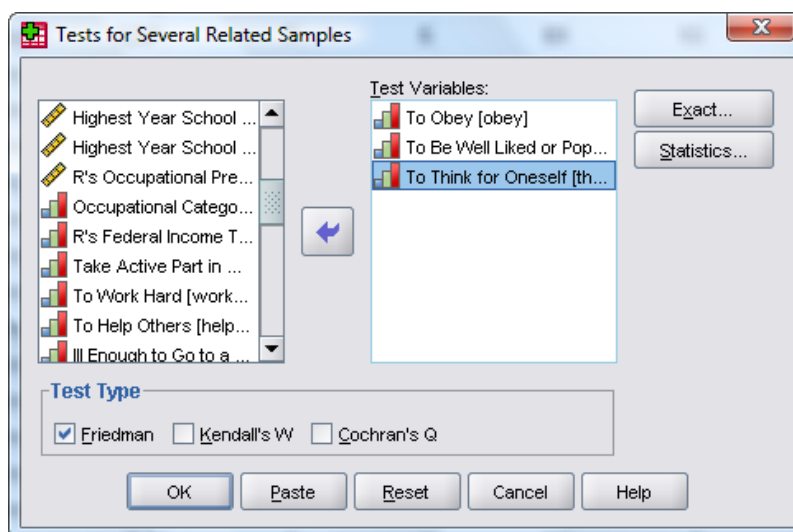
ομάδων. Η διόρθωση Bonferroni γίνεται εύκολα με το χέρι αν διαιρέσετε το επίπεδο στατιστικής σημαντικότητας με τον αριθμό των συγκρίσεων που θα κάνετε. Στο παράδειγμά μας χρησιμοποιείτε τρεις επιμέρους ομάδες, άρα έχετε τρεις συγκρίσεις (η πρώτη με τη δεύτερη, η πρώτη με την τρίτη, και η δεύτερη με την τρίτη). Έτσι το επίπεδο στατιστικής σημαντικότητας που πρέπει να χρησιμοποιήσετε είναι  $0,05 / 3 = 0,17$  περίπου. Αυτό σημαίνει ότι θεωρείτε στατιστικώς σημαντικές τις εκ των υστέρων συγκρίσεις όπου  $p < 0,17$ .

## 6.8 Friedman

Το κριτήριο Friedman είναι το μη παραμετρικό ισοδύναμο για την ανάλυση διακύμανσης επαναληπτικών μετρήσεων. Εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχετε καταχωρίσει τα δεδομένα σας με τον ίδιο τρόπο.

- Για να εκτελέσετε το τεστ επιλέγετε **[Analyze --> Nonparametric Tests --> K Related Samples]** και εμφανίζεται η οθόνη του Σχήματος 6.21:

Σχήμα 6.21 Τεστ Friedman



- Επιλέγετε τις μεταβλητές ή τις μετρήσεις που θέλετε να αναλύσετε και τις μεταφέρετε στη λίστα **[Test Variables:]**. Με την ολοκλήρωση αυτής της διαδικασίας, πατάτε το **[OK]** και παίρνετε τα εξής αποτελέσματα (βλ. Πίνακα 6.17):

**Πίνακας 6.17** 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 παρουσιάζονται περιγραφικοί στατιστικοί δείκτες για τις μεταβλητές ή για τις μετρήσεις της ίδιας μεταβλητής. Το SPSS δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank).
- Στο δεύτερο πίνακα 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 ( $p < 0,001$ ).
- 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 μπορείτε να τα πάρετε ακολουθώντας τα όσα περιγράφονται στο κεφάλαιο για τους δείκτες κεντρικής τάσης και διασποράς.

## 6.9 Δείκτες Συνάφειας – Pearson's r και Spearman's Rho

Προκειμένου να υπολογίσετε συνάφειες με το δείκτη συσχέτισης Pearson's r ή με το μη παραμετρικό ισοδύναμό του Spearman's Rho, οι μεταβλητές καταχωρίζονται σε ξεχωριστές στήλες, όπως φαίνεται στον Πίνακα 6.18.

**Πίνακας 6.18** Καταχώριση δεδομένων για τον υπολογισμό συναφειών

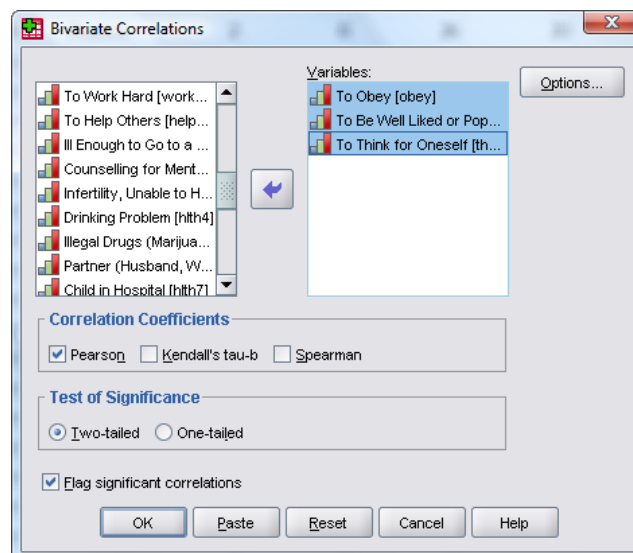
IQ	GPA
102	2.75
108	4.00



109	2.25
118	3.00
79	1.67
88	2.25
...	...
...	...
85	2.50

Επιλέγετε [**Analyze --> Correlate --> Bivariate...**] και εμφανίζεται η οθόνη του Σχήματος 6.22. Στη συνέχεια επιλέγετε και μετακινείτε τις μεταβλητές που θέλετε να αναλύσετε στη λίστα [**Variables:**] και κάνοντας κλικ στο [**OK**] παίρνετε τα αποτελέσματα της επεξεργασίας (βλ. Πίνακα 6.19).

**Σχήμα 6.22** Συνάψεις – Δείκτης συσχέτισης Pearson's r



**Πίνακας 6.19** Αποτελέσματα για το δείκτη συσχέτισης Pearson's r

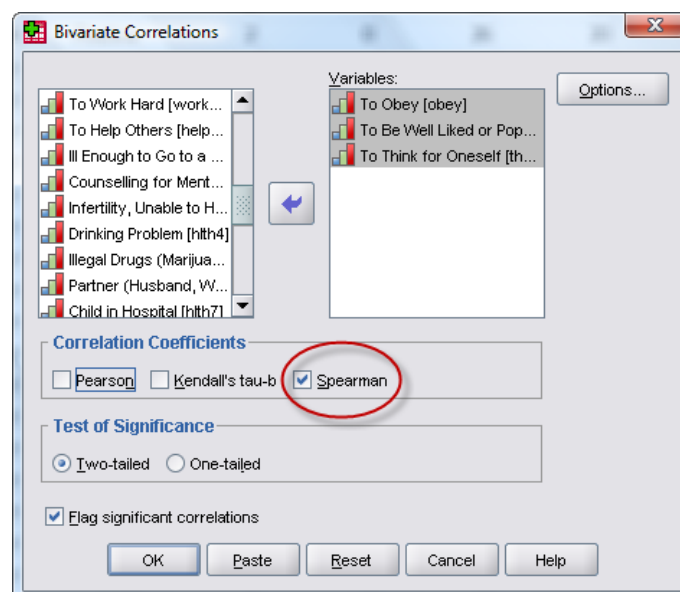
		<b>Correlations</b>		
		To Obey	To Be Well Liked or Popular	To Think for Oneself
To Obey	Pearson Correlation	1	-,253**	-,532**
	Sig. (2-tailed)		,000	,000
	N	982	982	982
To Be Well Liked or Popular	Pearson Correlation	-,253**	1	-,120**
	Sig. (2-tailed)	,000		,000
	N	982	982	982
To Think for Oneself	Pearson 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 βλέπετε το δείκτη Pearson's  $r$  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.
- Στην πρώτη γραμμή κάθε κελιού δίνεται ο δείκτης Pearson's  $r$ , 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 υπολογίζονται κάποιου άλλου είδους περιγραφικά στατιστικά, καθώς ο δείκτης συνάφειας δείχνει το μέγεθος και την κατεύθυνση της συνάφειας.

Στην περίπτωση που θέλετε να υπολογίσετε το μη παραμετρικό δείκτη συνάφειας Spearman Rho, ακολουθείτε τα ίδια ακριβώς βήματα, αλλά στο πεδίο **[Correlation Coefficients]** αποεπιλέγετε το δείκτη Pearson's  $r$  και επιλέγετε το δείκτη Spearman's Rho, ώστε η οθόνη σας να έχει την μορφή του Σχήματος 6.23.

**Σχήμα 6.23** Συνάφειες – Δείκτης συσχέτισης Spearman's Rho



Τα αποτελέσματα για το δείκτη Spearman's Rho είναι τα ακόλουθα (βλ. Πίνακα 6.20):

Πίνακας 6.20 Αποτελέσματα για το δείκτη συσχέτισης Spearman's Rho

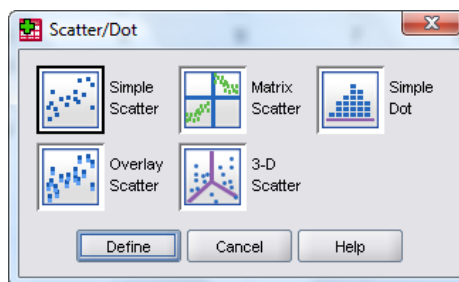
			Correlations		
			To Obey	To Be Well Liked or Popular	To Think for Oneself
Spearman's rho	To Obey	Correlation Coefficient	1,000	-,418**	-,529**
		Sig. (2-tailed)	.	,000	,000
		N	982	982	982
	To Be Well Liked or Popular	Correlation Coefficient	-,418**	1,000	-,038
		Sig. (2-tailed)	,000	.	,229
		N	982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

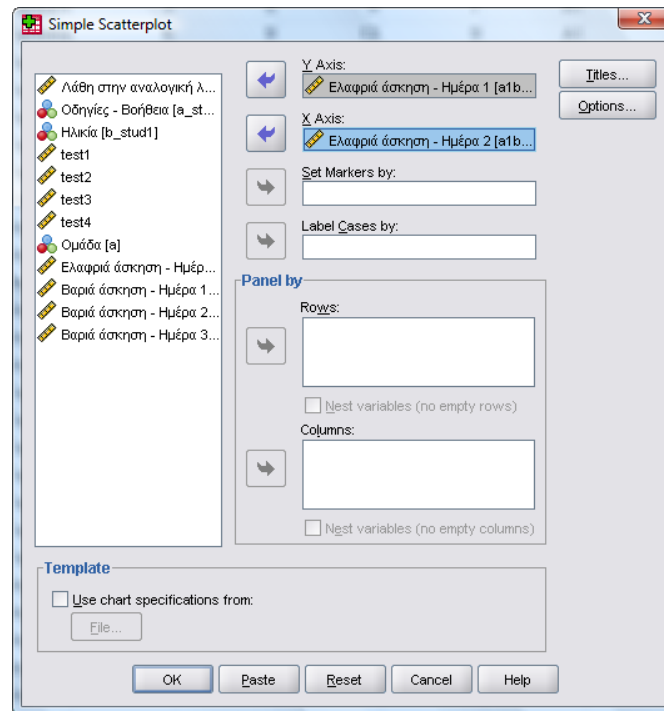
Για τη δημιουργία ενός διαγράμματος σκεδασμού (scatterplot), το οποίο προσφέρει οπτική απεικόνιση του είδους και του μεγέθους της σχέσης μεταξύ των δύο μεταβλητών, επιλέγετε **[Graphs --> Legacy Dialogs --> Scatter/Dot...]**. Στη πλαίσιο του Σχήματος 6.24 κάντε κλικ στο **[Define]** διατηρώντας επιλεγμένο το **[Simple Scatter]**.

Σχήμα 6.24 Το πλαίσιο διαλόγου Scatter/Dot



Επιλέγετε τις μεταβλητές που θα τοποθετηθούν στον οριζόντιο και στον κάθετο άξονα και τις μετακινείτε στα πλαίσια **[X Axis:]** και **[Y Axis:]** αντίστοιχα (βλ. Σχήμα 6.25). Τέλος, κάνετε κλικ στο **[OK]** για να πάρετε στο παράθυρο Viewer το διάγραμμα σκεδασμού.

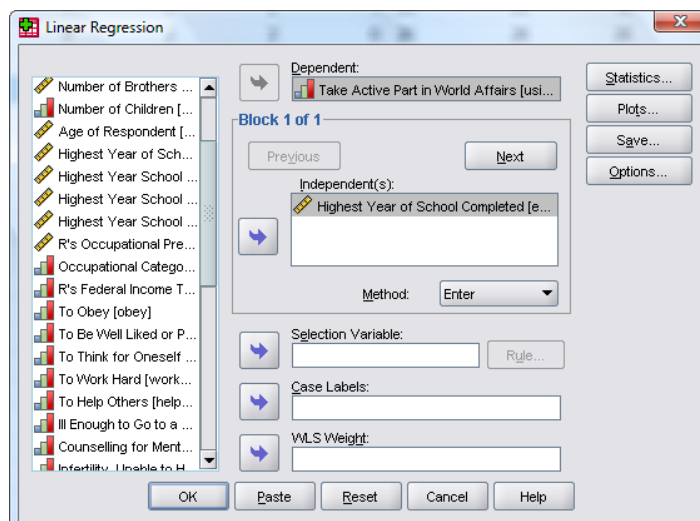
Σχήμα 6.25 Το πλαίσιο διαλόγου Simple Scatterplot



## 6.10 Απλή γραμμική παλινδρόμηση

Για την πραγματοποίηση της απλής γραμμικής παλινδρόμησης επιλέγετε **[Analyze --> Regression --> Linear...]**. Ανοίγει το παράθυρο διαλόγου που βλέπετε στο Σχήμα 6.26. Στο πλαίσιο κειμένου **[Dependent:]** μετακινείτε την εξαρτημένη μεταβλητή σας (τη μεταβλητή κριτήριο) και στο πλαίσιο **[Independent(s):]** την ανεξάρτητη μεταβλητή (ή προβλεπτική μεταβλητή). Στην περίπτωση που πραγματοποιείτε πολλαπλή παλινδρόμηση, στο πλαίσιο αυτό μετακινείτε όλες τις προβλεπτικές σας μεταβλητές.

Σχήμα 6.26 Το πλαίσιο διαλόγου Linear Regression



Κάνοντας κλικ στο **[OK]** παίρνετε τα αποτελέσματα της επεξεργασίας (βλ. Πίνακα 6.21).

**Πίνακας 6.21 Αποτελέσματα για την απλή γραμμική παλινδρόμηση**

**Variables Entered/Removed<sup>b</sup>**

Model	Variables Entered	Variables Removed	Method
1	Highest Year of School Completed <sup>a</sup>		. Enter

a. All requested variables entered.

b. Dependent Variable: Take Active Part in World Affairs

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,198 <sup>a</sup>	,039	,038	,423

a. Predictors: (Constant), Highest Year of School Completed

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7,260	1	7,260	40,517	,000 <sup>a</sup>
	Residual	177,920	993	,179		
	Total	185,180	994			

a. Predictors: (Constant), Highest Year of School Completed

b. Dependent Variable: Take Active Part in World Affairs

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1,624	,061		26,736	,000
	Highest Year of School Completed	-,029	,005	-,198	-6,365	,000

a. Dependent Variable: Take Active Part in World Affairs

Στην απλή παλινδρόμηση συνηθίζεται να εμφανίζουμε την εξίσωση της παλινδρόμησης με ένα συντελεστή κλίσης (b) και μια τεταγμένη (α). Το SPSS δεν ακολουθεί ακριβώς αυτή την ορολογία, αλλά στον τελευταίο από τους παραπάνω πίνακες εμφανίζονται όλες οι σχετικές πληροφορίες.

Ο Πίνακας 6.21 περιλαμβάνει τέσσερις πίνακες με αποτελέσματα της παλινδρόμησης όπως εμφανίζονται στο παράθυρο Viewer. Ο πρώτος από αυτούς παρουσιάζει τις μεταβλητές της ανάλυσης (Προβλεπτικός παράγοντας: Highest Year of School Completed & Μεταβλητή κριτήριο: Take Active Part in World Affairs) καθώς και τη μέθοδο με την οποία προστέθηκε η προβλεπτική μεταβλητή στο μοντέλο (Enter).

Ο δεύτερος από τους πίνακες (Model Summary) παρουσιάζει το R (0,198 στο παράδειγμά μας), το οποίο είναι ο δείκτης συσχέτισης μεταξύ των δύο μεταβλητών, το συντελεστή προσδιορισμού  $R^2$  (0,039 στο παράδειγμά μας) και το συντελεστή Adjusted  $R^2$  (0,038 στο παράδειγμά μας), ο οποίος δηλώνει ότι η προβλεπτική μεταβλητή (Highest Year of School Completed) είναι υπεύθυνη για το 4% περίπου της μεταβολής των τιμών της μεταβλητής κριτήριο (Take Active Part in World Affairs). Με ακόμη πιο απλά λόγια, αν γνωρίζουμε τα συνολικά έτη εκπαίδευσης ενός ατόμου μπορούμε να κατανοήσουμε κατά 4% περίπου το φαινόμενο της ενεργούς συμμετοχής του στα διεθνή ζητήματα.

Όπως δείχνει ο τρίτος πίνακας των αποτελεσμάτων, με τον τίτλο ANOVA, η στατιστική σημαντικότητα του λόγου F είναι 0,000 ( $p < ,001$ ). Αυτό το τεστ αξιολογεί την υπόθεση ότι η πρόβλεψη είναι σημαντική ή τυχαία. Στο παράδειγμα αυτό, συμπεραίνετε ότι υπάρχει στατιστικά σημαντική συσχέτιση/πρόβλεψη της εξαρτημένης μεταβλητής από την ανεξάρτητη.

Στον τελευταίο πίνακα (Coefficients) παρουσιάζονται τα εξής: το σύμβολο B είναι η κλίση. Η κλίση της γραμμής παλινδρόμησης ονομάζεται στο SPSS μη κανονικοποιημένος συντελεστής παλινδρόμησης (unstandardized regression coefficient). Ο συντελεστής αυτός στο παράδειγμά μας είναι -0,029. Αυτό σημαίνει ότι για κάθε αύξηση κατά 1 μονάδα στον οριζόντιο άξονα, η τιμή στον κατακόρυφο άξονα αυξάνεται κατά -,029. Η τεταγμένη (α) αναφέρεται ως σταθερά (constant) στο SPSS. Η σταθερά αυτή είναι 1,624 στο παράδειγμά μας και πρόκειται για το σημείο στο οποίο η γραμμή παλινδρόμησης τέμνει τον κατακόρυφο άξονα (Y). Η στήλη με την επικεφαλίδα "Beta" δίνει μια τιμή -0,198. Προσέξτε ότι η τιμή αυτή ταυτίζεται με το δείκτη συσχέτισης μεταξύ των δύο μεταβλητών. Αν μετατρέψετε τις τιμές των δεδομένων σας σε τυπικές τιμές (standard scores), η κλίση της γραμμής παλινδρόμησης και ο συντελεστής συσχέτισης ταυτίζονται. Στον πίνακα υπάρχουν επίσης τα αποτελέσματα από δύο t-test: το πρώτο από αυτά εξετάζει την υπόθεση ότι η τιμή της σταθεράς είναι ίση με το μηδέν και ουσιαστικά έχει ελάχιστη πρακτική αξία καθώς συνεισφέρει ελάχιστες πληροφορίες για το μέγεθος και την κατεύθυνση της πρόβλεψης. Αντίθετα, το δεύτερο t-test ελέγχει την υπόθεση ότι η κλίση b είναι ίση με το μηδέν. Προσέξτε ότι στατιστικά σημαντική κλίση σημαίνει και σημαντική πρόβλεψη. Στο παράδειγμά μας, το αποτέλεσμα του

συγκεκριμένου t-test είναι σημαντικό ( $p < .001$ ), επομένως υπάρχει μια στατιστικά σημαντική πρόβλεψη της εξαρτημένης μεταβλητής από την ανεξάρτητη.

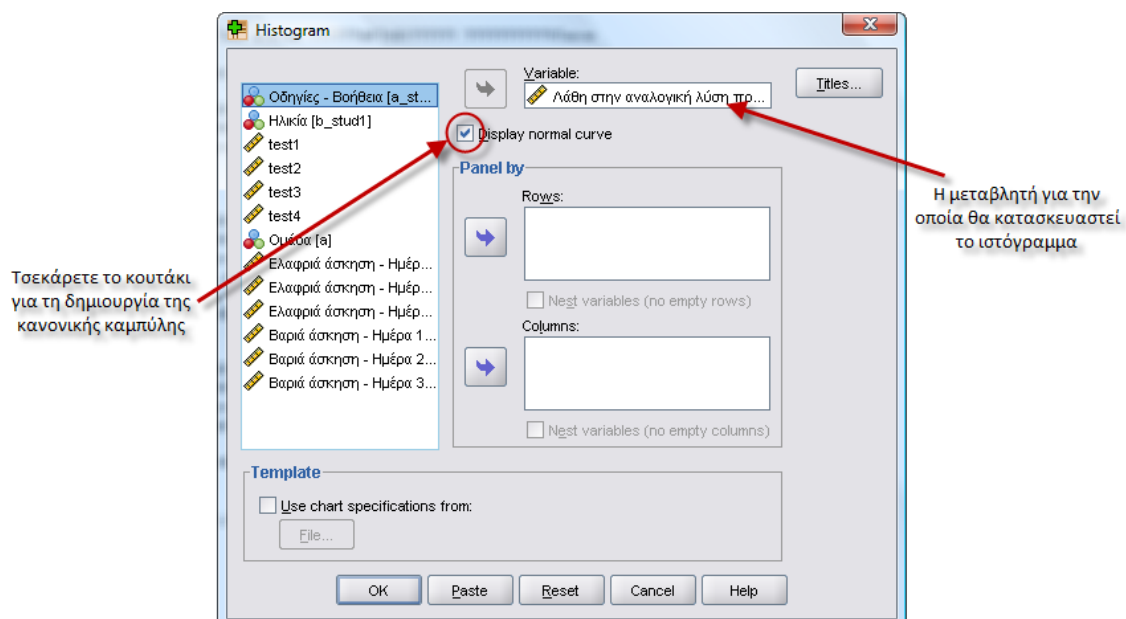
### 6.11 Διαδικασία ελέγχου κανονικότητας ερευνητικών δεδομένων

Μια από τις προϋποθέσεις που πρέπει να πληρούν τα δεδομένα μιας έρευνας προκειμένου να εφαρμόσετε παραμετρικά στατιστικά κριτήρια είναι να σχηματίζουν κανονική κατανομή. Αν και ορισμένοι (π.χ., Miceri, 1989) υποστηρίζουν ότι στην ψυχολογία είναι πολύ σπάνια τα δεδομένα που σχηματίζουν κανονική κατανομή (πολύ συχνά αυτό οφείλεται στη χρήση λανθανουσών –latent– μεταβλητών), δεν είναι πολύ σωστό να αγνοείτε τη συγκεκριμένη προϋπόθεση... Πώς εξετάζετε όμως τη συγκεκριμένη προϋπόθεση;

Δεν υπάρχει μια σαφής απάντηση... Υπάρχουν αρκετές διαφορετικές διαδικασίες, αλλά όχι τόσο σαφή κριτήρια.

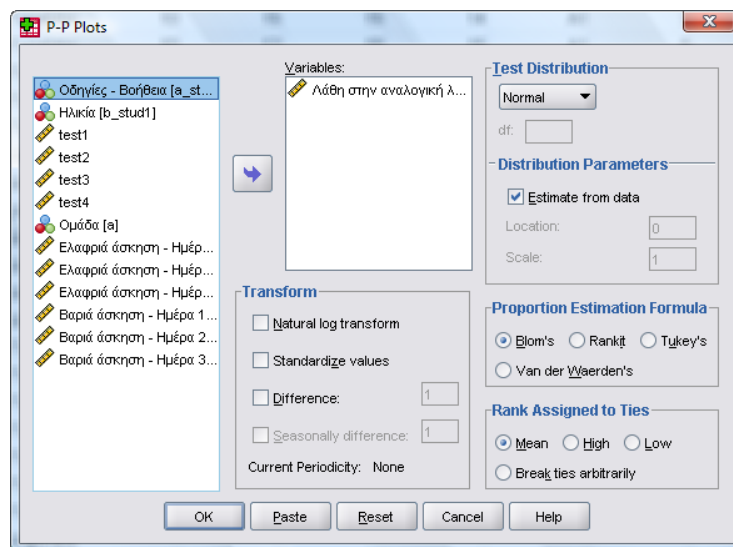
1. **Κατασκευή του ιστογράμματος της μεταβλητής.** Επιλέγετε [**Graphs --> Legacy Dialogs --> Histogram...**]. Στο παράθυρο που ανοίγει (Σχήμα 6.27) θα πρέπει (α) να μετακινήσετε στο πλαίσιο [**Variable:**] τη μεταβλητή για την οποία θα κατασκευαστεί το ιστόγραμμα, και (β) να τσεκάρετε το κουτάκι στην επιλογή για τη δημιουργία της κανονικής καμπύλης πάνω από το ιστόγραμμα [**Display normal curve**]. Βεβαίως, δεν υπάρχουν ασφαλή κριτήρια για το βαθμό απόκλισης του ιστογράμματος από την κανονική καμπύλη...

Σχήμα 6.27 Το πλαίσιο διαλόγου Histogram

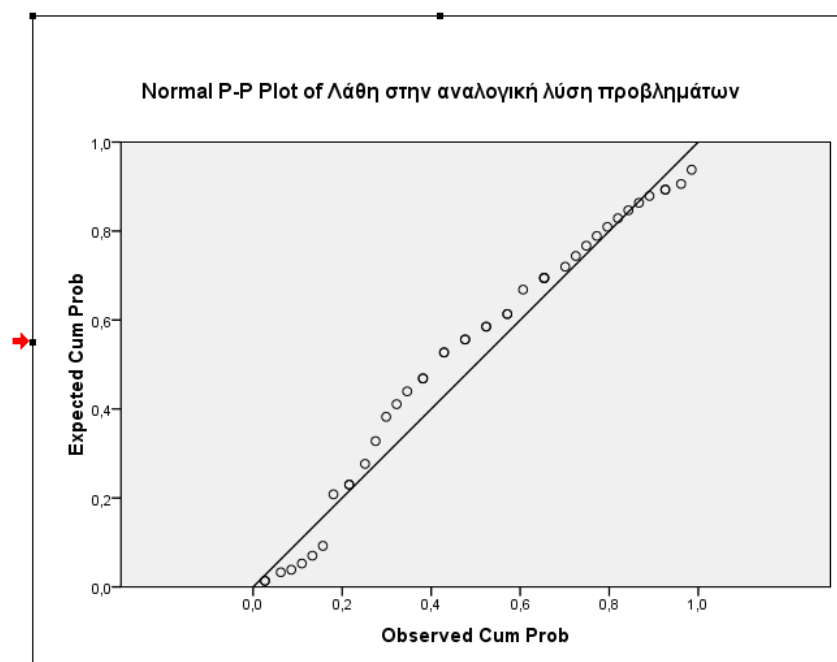


2. **Κατασκευή P-P plots** (Proportion-Proportion). Πρόκειται για ένα διάγραμμα όπου η παρατηρηθείσα αθροιστική σχετική συχνότητα σχεδιάζεται απέναντι στην αναμενόμενη αθροιστική σχετική συχνότητα έτσι όπως θα έδειχναν αν η κατανομή ήταν κανονική. Επιλέγετε **[Analyze --> Descriptive Statistics --> P-P Plots...]** και στο παράθυρο που ανοίγει (Σχήμα 6.28) μεταφέρετε στο πλαίσιο **[Variables:]** τη μεταβλητή για την οποία θα κατασκευαστεί το διάγραμμα. Κάνοντας κλικ στο κουμπί **[OK]** θα δημιουργηθεί ένα διάγραμμα όπως αυτό του Σχήματος 6.29.

Σχήμα 6.28 Το πλαίσιο διαλόγου P-P Plots



Σχήμα 6.29 Ένα παράδειγμα διαγράμματος P-P plot

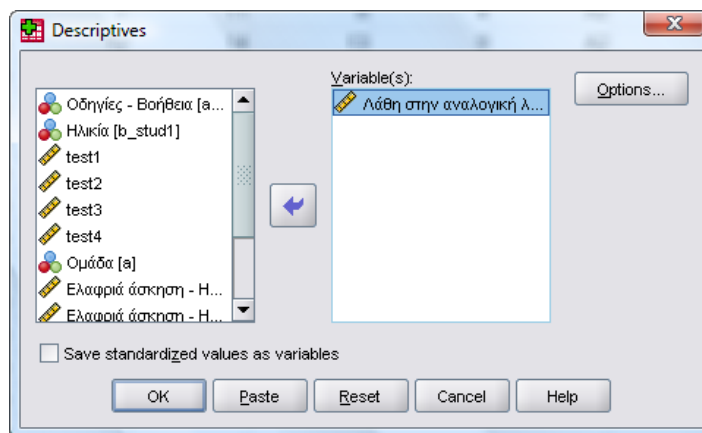




Όλα τα σημεία θα πρέπει να βρίσκονται πάνω στη διαγώνιο εφόσον η μεταβλητή είναι κανονικά κατανομημένη. Κατασκευάζεται εύκολα αλλά δεν υπάρχουν κοινά αποδεκτά κριτήρια για τον καθορισμό της απόστασης που μπορεί να έχουν τα σημεία από τη διαγώνιο ώστε να θεωρηθεί κανονική η κατανομή...

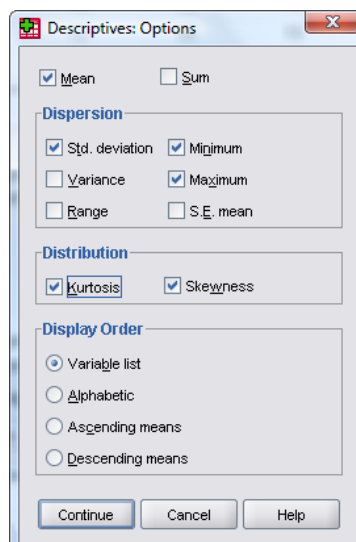
3. **Συμμετρία και κύρτωση (skew & kurtosis).** Επιλέγετε [**Analyze --> Descriptive Statistics --> Descriptives...**] και στο παράθυρο που ανοίγει (Σχήμα 6.30) μεταφέρετε στο πλαίσιο [**Variables:**] τη μεταβλητή για την οποία θέλετε να υπολογιστούν οι συγκεκριμένοι δείκτες.

**Σχήμα 6.30** Το παράθυρο διαλόγου Descriptives



Κάντε κλικ στο κουμπί [**Options...**] και στο παράθυρο διαλόγου που ανοίγει (Σχήμα 6.31) επιλέγετε το [**Kurtosis**] και το [**Skewness**]. Πατάτε [**Continue**], μετά [**OK**] και στο παράθυρο Viewer θα πάρετε ένα πίνακα με όλους τους δείκτες που έχετε ζητήσει.

**Σχήμα 6.31** Το παράθυρο διαλόγου Descriptives: Options



Και οι δύο τιμές πρέπει να είναι μηδενικές για κανονική κατανομή. Ένας απλός κανόνας για να δεχτείτε ότι μια μεταβλητή είναι κανονική είναι να μην ξεφεύγουν οι τιμές της συμμετρίας και της κύρτωσης από το +2 ως το -2. Άλλοι είναι πιο «αυστηροί» και συνιστούν το εύρος από το +1 ως το -1... Αυτός ο κανόνας εφαρμόζεται στις περιπτώσεις μεγάλων δειγμάτων ( $N > 300$ ). Αν το δείγμα είναι μικρό ( $N < 100$ ), υπολογίστε τις τυπικές τιμές για τη συμμετρία και την κύρτωση βάσει του τύπου που ακολουθεί και απορρίψτε ως μη κανονικές τις μεταβλητές εκείνες που έχουν τυπική τιμή μεγαλύτερη από το 1,96:

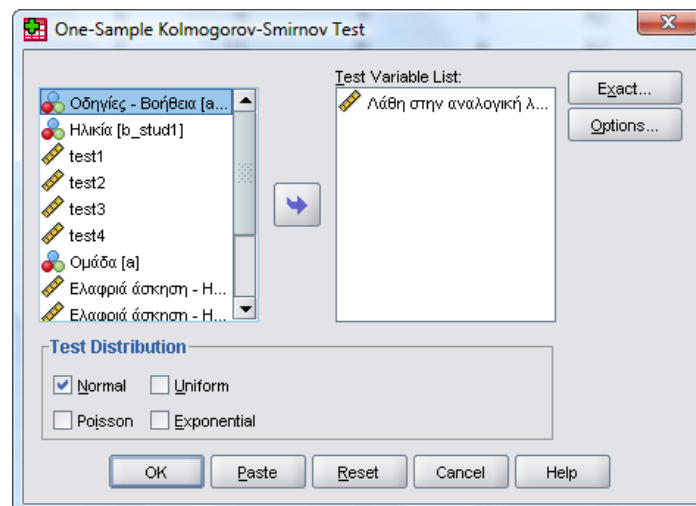
$$Z_{\text{skew}} = \text{Skew} / \text{SE}_{\text{skew}}$$

$$Z_{\text{kurtosis}} = \text{Kurtosis} / \text{SE}_{\text{kurtosis}}$$

Αν το δείγμα σας ήταν μέτριου μεγέθους ( $100 < N < 300$ ), τότε υπολογίστε και πάλι τις τυπικές τιμές για τη συμμετρία και την κύρτωση και απορρίψτε ως μη κανονικές εκείνες τις μεταβλητές που έχουν έστω μία από αυτές τις τυπικές τιμές μεγαλύτερη από 3,29.

4. **Τυπικά τεστ κανονικότητας.** Το γνωστότερο είναι το Kolmogorov-Smirnov (ένα δεύτερο που υπολογίζει το SPSS είναι το Shapiro-Wilk). Επιλέγετε **[Analyze --> Nonparametric Tests --> 1-Sample K-S...]** και μεταφέρετε στο πλαίσιο διαλόγου **[Test Variable List:]** του Σχήματος 6.32 τη μεταβλητή για την οποία θα πραγματοποιηθεί το τεστ.

**Σχήμα 6.32 Το πλαίσιο διαλόγου One-Sample Kolmogorov-Smirnov Test**



Κάνοντας κλικ στο **[OK]** παίρνετε τα αποτελέσματα της επεξεργασίας (βλ. Πίνακα 6.22).

Πίνακας 6.22 Αποτέλεσμα του τεστ Kolmogorov-Smirnov

One-Sample Kolmogorov-Smirnov Test		Λάθη στην αναλογική λύση προβλημάτων
N		42
Normal Parameters <sup>a</sup>	Mean	30,07
	Std. Deviation	13,617
Most Extreme Differences	Absolute	,122
	Positive	,074
	Negative	-,122
Kolmogorov-Smirnov Z		,793
Asymp. Sig. (2-tailed)		,555

a. Test distribution is Normal.

Στο παράδειγμά μας το αποτέλεσμα του τεστ δεν είναι στατιστικά σημαντικό ( $p=.555$ ), επομένως θα πρέπει να δεχτείτε ότι η μεταβλητή που ελέγχθηκε δεν αποκλίνει από την κανονικότητα. Το πρόβλημα με τα τεστ αυτά είναι ότι όσο μεγαλώνει το δείγμα ( $N > 300$ ) τόσο αυξάνει η πιθανότητα να απορριφθεί μια μεταβλητή, η οποία αποκλίνει ελάχιστα από την κανονικότητα. Τα τεστ αυτά είναι πολύ «συντηρητικά» και μπορεί να απορρίψουν ως μη κανονική μια μεταβλητή που απέχει ελάχιστα από την κανονικότητα.