



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΙΓΑΙΟΥ
ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ

ΠΟΣΟΤΙΚΕΣ ΜΕΘΟΔΟΙ ΣΤΗ ΓΕΩΓΡΑΦΙΑ: ΕΡΓΑΣΤΗΡΙΟ

Εισαγωγή στη χρήση του SPSS 15.0

Παναγιώτης Στρατάκης
Γεωγράφος, ΜΑ, Υπ. Δρ. Ε.Μ.Π

Μυτιλήνη Δεκέμβριος 2008

ΠΕΡΙΕΧΟΜΕΝΑ

1. Εισαγωγή.....	2
2. Εισαγωγή δεδομένων-Ορισμός μεταβλητών	3
3. Δημιουργία νέων μεταβλητών-Κατανομές συχνοτήτων	5
3.1 Δημιουργία νέων ποιοτικών μεταβλητών με επανακωδικοποίηση.....	5
3.2 Δημιουργία νέων ποσοτικών μεταβλητών με υπολογισμούς	7
3.3 Κατανομές συχνοτήτων.....	8
4. Δημιουργία γραφημάτων	10
5. Περιγραφικά στατιστικά	12
5.1 Μέτρα κεντρικής τάσης-Μέτρα διασποράς	12
6. Έλεγχοι Υποθέσεων	14
6.1 Σύγκριση μέσων όρων ενός δείγματος και ενός γνωστού πληθυσμού	14
6.2 Σύγκριση μέσων όρων δυο ανεξάρτητων δειγμάτων	15
6.3 Σύγκριση μέσων όρων δυο εξαρτημένων δειγμάτων	16
7. Διασταυρώσεις μεταβλητών και έλεγχος χ^2	18
7.1 Δημιουργία πίνακα διπλής εισόδου	18
7.2 Έλεγχος ανεξαρτησίας χ^2	19
8. Συντελεστές συσχέτισης ποιοτικών δεδομένων.....	21
9. Δειγματοληψία.....	22
9.1 Απλή τυχαία Δειγματοληψία.....	22

1. Εισαγωγή

Το παρόν εγχειρίδιο απευθύνεται στους φοιτητές του Τμήματος Γεωγραφίας και αποτελεί μια σύννοψη των διαδικασιών που διδάχθηκαν κατά τη διάρκεια του εργαστηριακού μαθήματος 'Ποσοτικές Μέθοδοι στη Γεωγραφία'.

Στόχος του εγχειριδίου είναι η συμβολή στην εκμάθηση της χρήσης του στατιστικού πακέτου SPSS (Statistical Package for Social Sciences). Το εγχειρίδιο είναι συμπληρωματικό με την παρακολούθηση του εργαστηρίου στην οποία γίνεται αναλυτική παρουσίαση και ερμηνεία των αποτελεσμάτων που προκύπτουν από τις στατιστικές διαδικασίες.

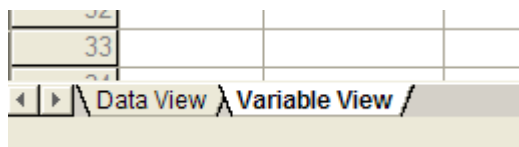
Το εργαστήριο καλύπτει μια εισαγωγή στις τεχνικές στατιστικής επεξεργασίας τόσο ποιοτικών όσο και ποσοτικών δεδομένων που είναι χρήσιμες στους φοιτητές στο στάδιο της ανάλυσης των δεδομένων.

Με την επιτυχή παρακολούθηση του εργαστηρίου οι φοιτητές θα είναι σε θέση να χρησιμοποιήσουν τις βασικές λειτουργίες του λογισμικού SPSS έτσι ώστε να εισάγουν, να επεξεργάζονται και να ερμηνεύουν δεδομένα για μια απλή στατιστική ανάλυση.

2. Εισαγωγή δεδομένων-Ορισμός μεταβλητών

Στο SPSS υπάρχουν τρεις διαφορετικοί τρόποι εισαγωγής δεδομένων. Ο **πρώτος** είναι η εισαγωγή μέσω πληκτρολόγησης των δεδομένων και αφορά αυτά που προέρχονται από πρωτογενή έρευνα π.χ. ερωτηματολόγια από έρευνα πεδίου.

Ο **δεύτερος** τρόπος είναι η εισαγωγή από ένα αρχείο excel με την διαδικασία αντιγραφή-επικόλληση (copy-paste) και αφορά δεδομένα που προέρχονται από δευτερογενείς πηγές π.χ. Εθνική Στατιστική Υπηρεσία Ελλάδος. Με αυτούς τους δυο τρόπους πριν ξεκινήσουμε την διαδικασία πρέπει να προετοιμάσουμε το αρχείο του SPSS και να ορίσουμε της μεταβλητές στο **Variable View**.

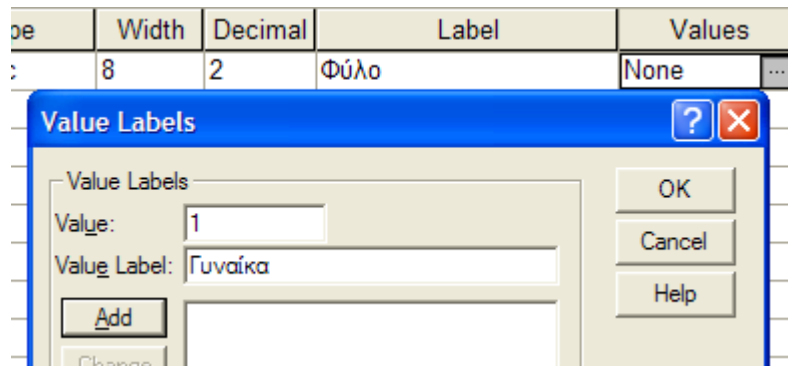


Στον ορισμό των μεταβλητών θα πρέπει να δώσουμε αρχικά την ονομασία (Name) της μεταβλητής (π.χ. Φύλο), το είδος της (type) (π.χ. αριθμητική (Numeric), αλφαριθμητική (String)), την ετικέτα (label) και την κλίμακα μέτρησης (measure) δηλαδή αν είναι ποσοτική (Scale) ή ποιοτική (Ordinal= διατάξιμη, Nominal=ονομαστική).

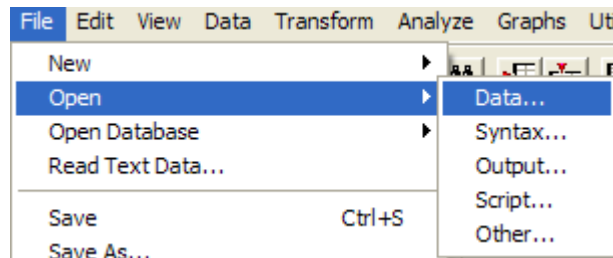
 A screenshot of the SPSS Variable View window. The window title is 'File Edit View Data Transform Analyze Graphs Utilities Window Help'. Below the title bar is a toolbar with various icons. The main area is a table with the following columns: Name, Type, Width, Decimal, Label, Values, Missing, Columns, Align, and Measure. The first row is for variable 1, with Name 'Φύλο', Type 'Numeric', Width '8', Decimal '2', Label 'Φύλο', Values 'None', Missing 'None', Columns '8', Align 'Right', and Measure 'Scale'. The second, third, and fourth rows are empty. The Measure column has a dropdown menu open, showing options: Scale, Ordinal, and Nominal.

	Name	Type	Width	Decimal	Label	Values	Missing	Columns	Align	Measure
1	Φύλο	Numeric	8	2	Φύλο	None	None	8	Right	Scale
2										Scale
3										Ordinal
4										Nominal

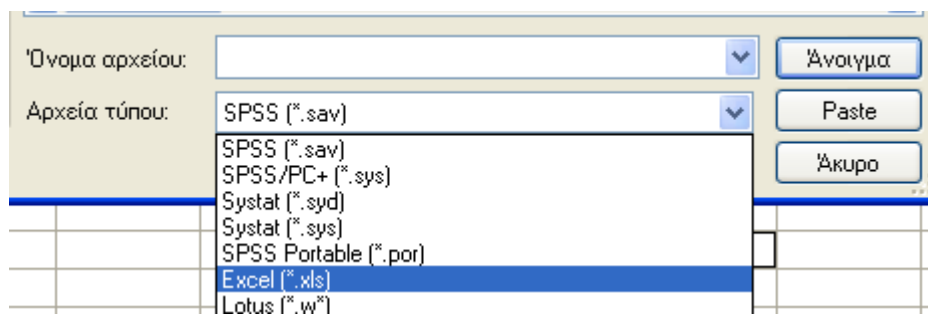
Αν θέλουμε να εισάγουμε μια ονομαστική μεταβλητή χρησιμοποιώντας αριθμητικό είδος θα πρέπει να δώσουμε τιμές (Values) και ετικέτες στις τιμές (Value Label) π.χ. 1=Γυναίκα που θα μας βοηθήσουν κατά τη διάρκεια της εισαγωγής και θα εμφανιστούν αργότερα στα αποτελέσματα στατιστικών επεξεργασιών.



Ο **τρίτος** τρόπος είναι η αυτόματη εισαγωγή ολόκληρου του αρχείου excel στο SPSS με την διαδικασία **File-Open-Data**.

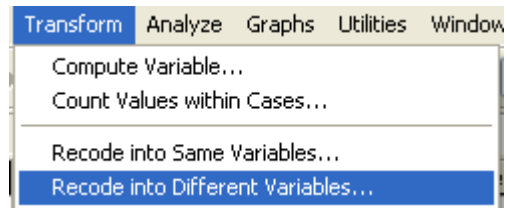


Σε αυτή τη διαδικασία θα πρέπει να προσέξουμε και να αλλάξουμε τον τύπο του αρχείου σε excel γιατί το λογισμικό αναζητά τα αρχεία δεδομένων με προέκταση .sav.



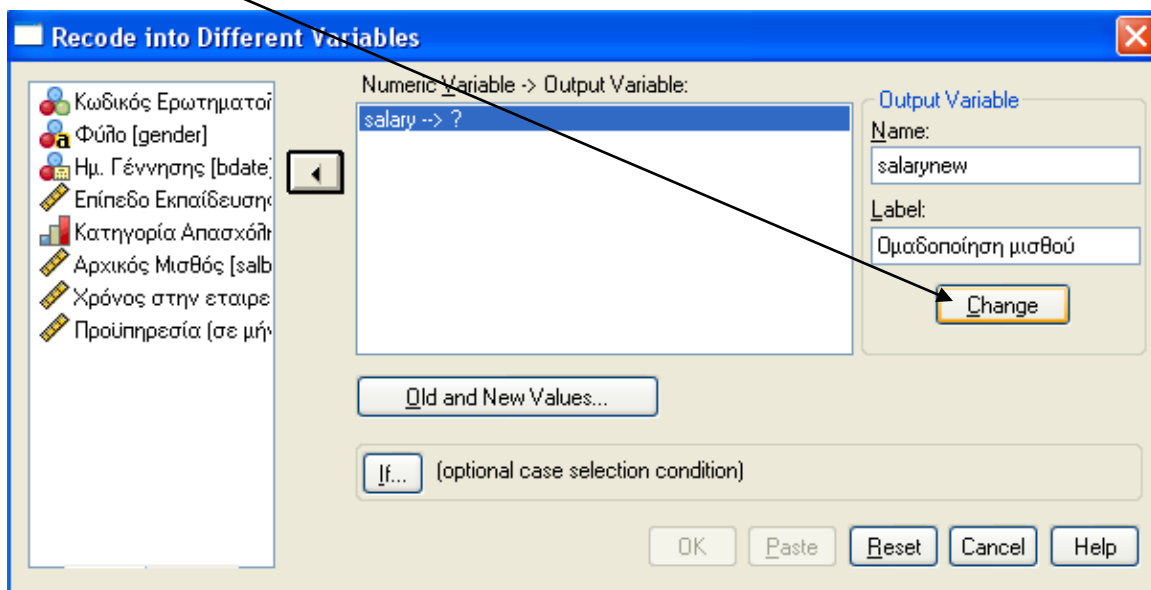
3. Δημιουργία νέων μεταβλητών-Κατανομές συχνοτήτων

3.1 Δημιουργία νέων ποιοτικών μεταβλητών με επανακωδικοποίηση

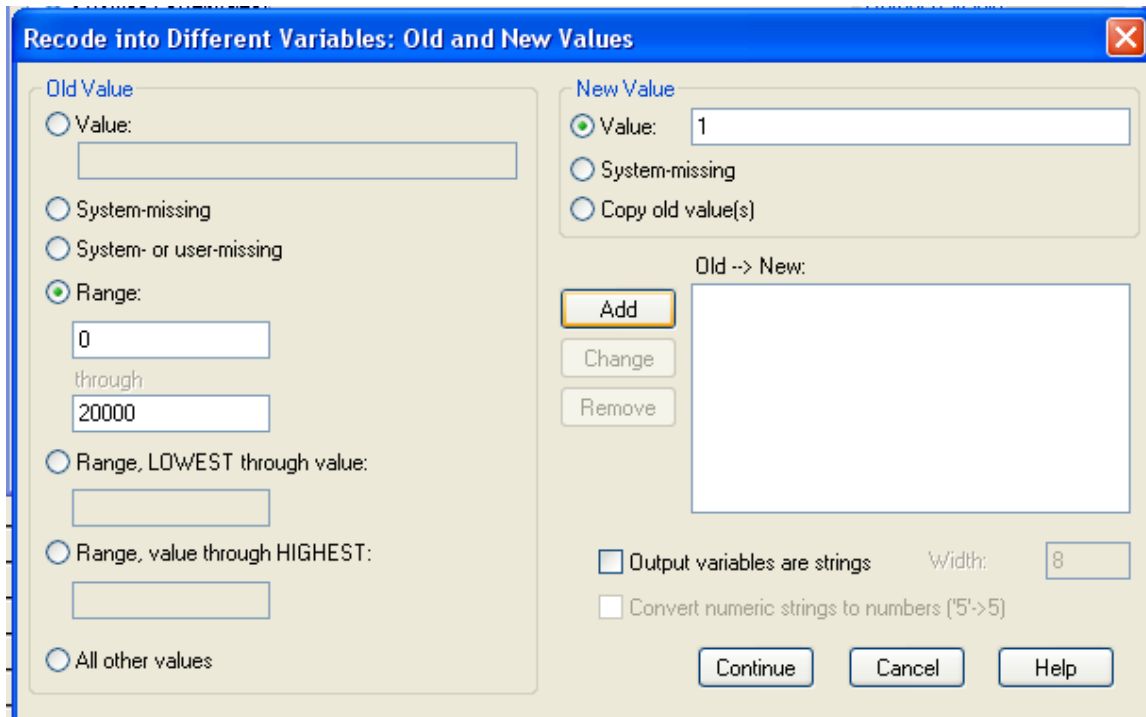


Με τη διαδικασία **Transform**→**Recode Into Different Variables**, θα προκύψει μία νέα μεταβλητή στα δεδομένα σας (Data Editor), η οποία θα περιλαμβάνει τον κωδικό για κάθε τιμή της μεταβλητής, ανάλογα με την ομάδα τιμών στην οποία ανήκει.

Η νέα αυτή μεταβλητή εμφανίζεται στο παρακάτω menu ως **output variable** και θα πρέπει να της δώσετε όνομα (π.χ. salarynew) και να επιλέξετε το κουμπι **Change**.



Στη συνέχεια πατώντας το κουμπί **Old and New Values** θα τοποθετήσουμε τις ομάδες τιμών στο **Range** του Old Value που θέλουμε να δημιουργήσουμε και θα δώσουμε ένα νέο κωδικό για την κάθε ομάδα στο New Value (Value π.χ. 1) και πατώντας το κουμπί **Add** θα συνεχίσουμε.



Μετά την ολοκλήρωση της διαδικασίας θα πρέπει να κάνετε την αντίστροφη διαδικασία για να δώσετε τις ετικέτες (Value Labels) στους κωδικούς των ομάδων που δημιουργήσατε.

Η διαδικασία αυτή πραγματοποιείται στο **Variable View** στο πεδίο **Values**.

Name	Type	Width	Decimals	Label	Values
id	Numeric	4	0	Κωδικός Ερωτηματολογίου	None
gender	String	1	0	Φύλο	{f, Γυναίκα}...
bdate	Date	10	0	Ημ. Γέννησης	None
educ	Numeric	2	0	Επίπεδο Εκπαίδευσης (σε έτη)	{0, 0 (Missing)}
jobcat	Numeric	1	0	Κατηγορία Απασχόλησης	{0, 0 (Missing)}
salary	Dollar	8	0	Σημερινός Μισθός	{\$0, missing}..
salbegin	Dollar	8	0	Αρχικός Μισθός	{\$0, missing}..
jobtime	Numeric	2	0	Χρόνος στην εταιρεία (σε μήνες)	{0, missing}...
prevexp	Numeric	6	0	Προϋπηρεσία (σε μήνες)	{0, missing}...
salarynew	Numeric	8	2	Ομαδοποίηση μισθού	None

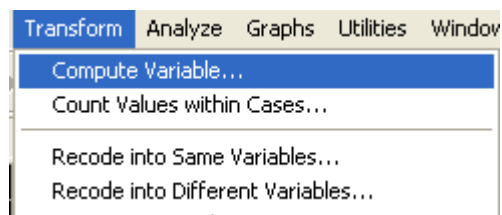
Value Labels

Value Labels

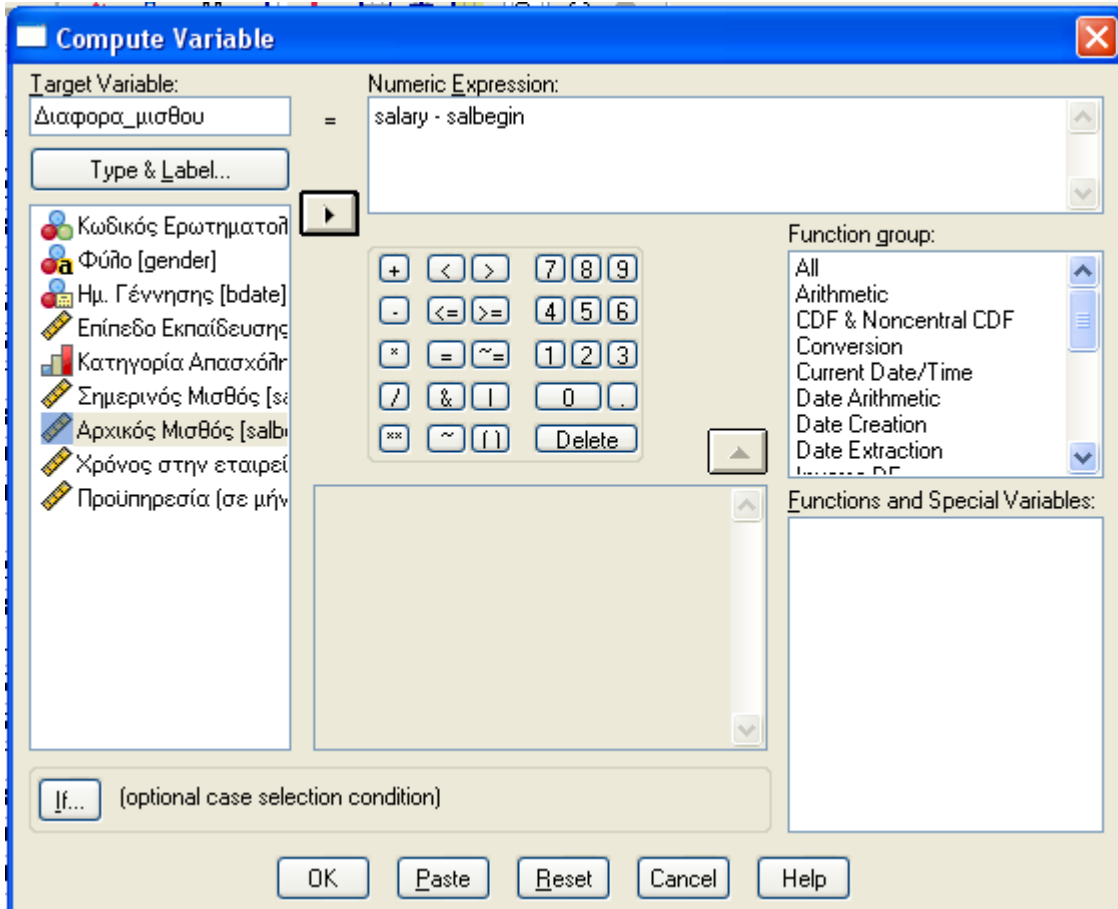
Value:

Label:

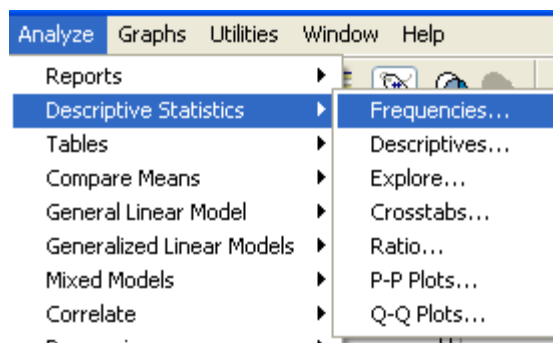
3.2 Δημιουργία νέων ποσοτικών μεταβλητών με υπολογισμούς



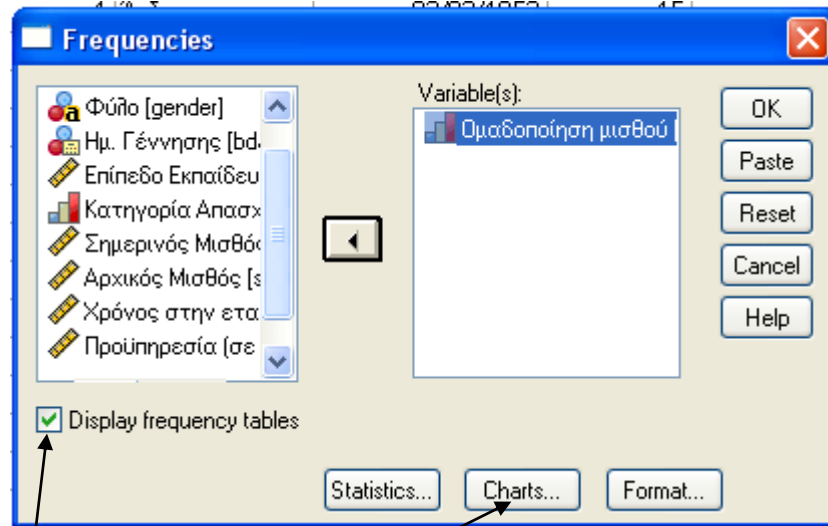
Η διαδικασία αυτή μας βοηθάει να δημιουργήσουμε νέες ποσοτικές μεταβλητές χρησιμοποιώντας τις ήδη υπάρχουσες και κάνοντας τους κατάλληλους υπολογισμούς στο **Numeric Expression**, δίνοντας ένα νέο όνομα στην μεταβλητή που θα δημιουργήσουμε στο **Target Variable** (π.χ. διαφορά μισθού).



3.3 Κατανομές συχνοτήτων



Η παραπάνω διαδικασία μας βοηθάει να υπολογίσουμε τις κατανομές συχνοτήτων σε ποιοτικά αλλά και ποσοτικά δεδομένα.

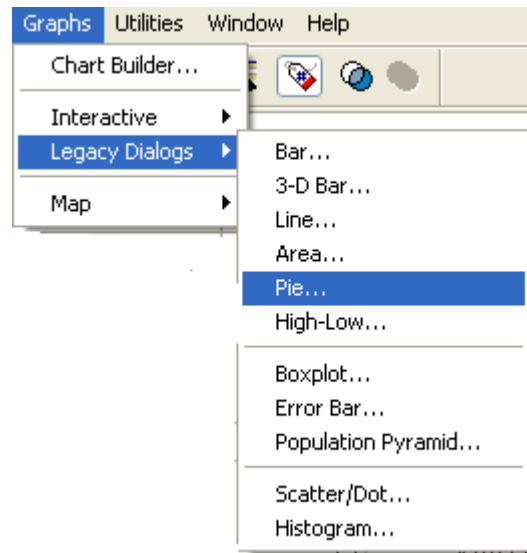


Οι κατανομές συχνοτήτων θα εμφανιστούν αν διατηρήσετε επιλεγμένο το **Display frequency tables**.

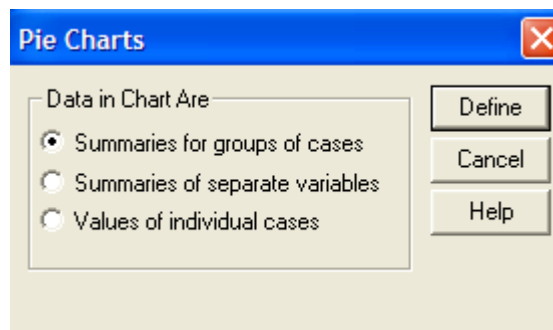
Από την επιλογή **Charts** μπορείτε να επιλέξετε και τη δημιουργία ιστογράμματος συχνοτήτων (**histogram**).

4. Δημιουργία γραφημάτων

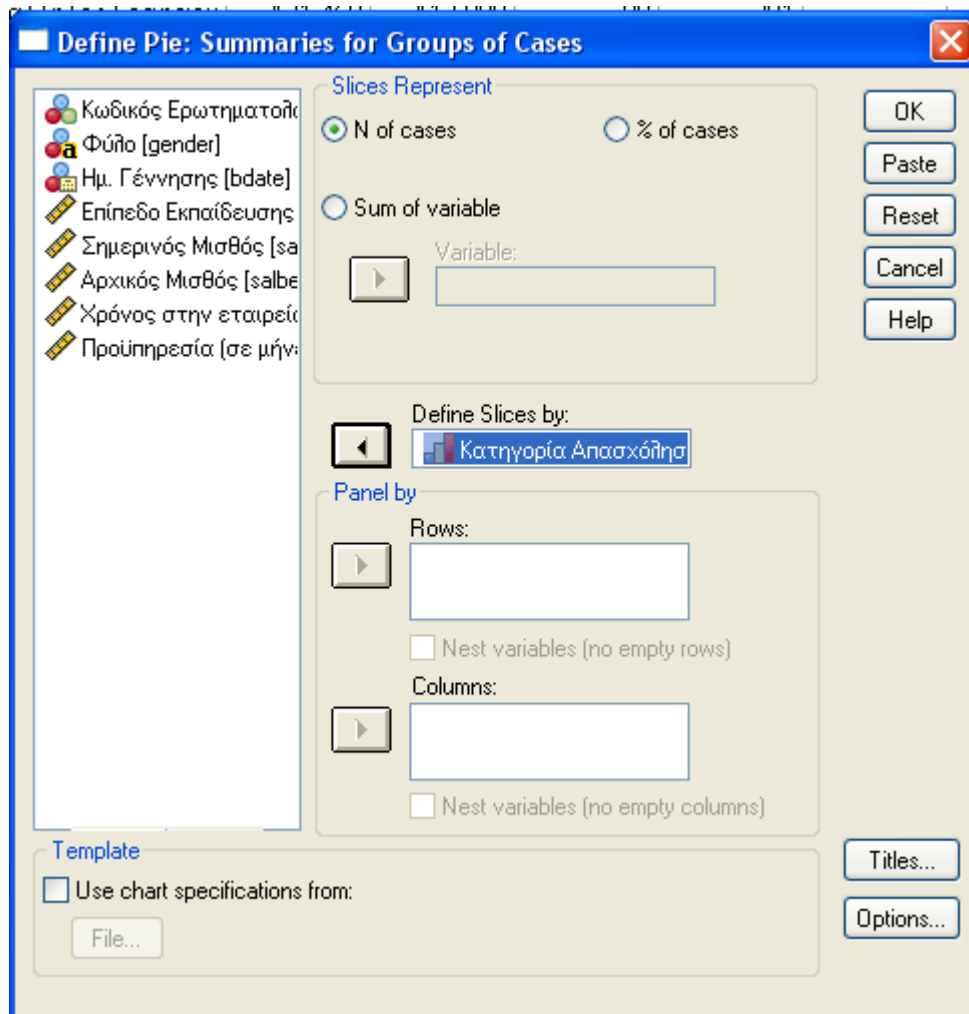
Για τη δημιουργία γραφημάτων χρησιμοποιούμε την εντολή **Graphs→Legacy Dialogs** και ανάλογα με το γράφημα επιλέγουμε **Bar** για ραβδόγραμμα ή **Pie** για κυκλικό διάγραμμα κτλ.



Αν θέλουμε να δημιουργήσουμε ένα γράφημα για μια μεταβλητή επιλέγουμε **Summaries for groups of cases**, ενώ για πάνω από μια μεταβλητές επιλέγουμε **Summaries of separate variables**.



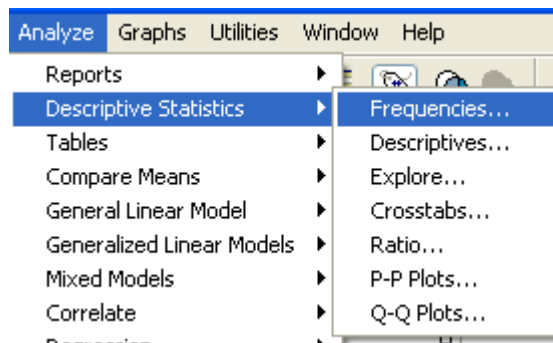
Στη συνέχεια θα τοποθετήσουμε στο **Define slices by** τη μεταβλητή που θέλουμε να χρησιμοποιήσουμε.



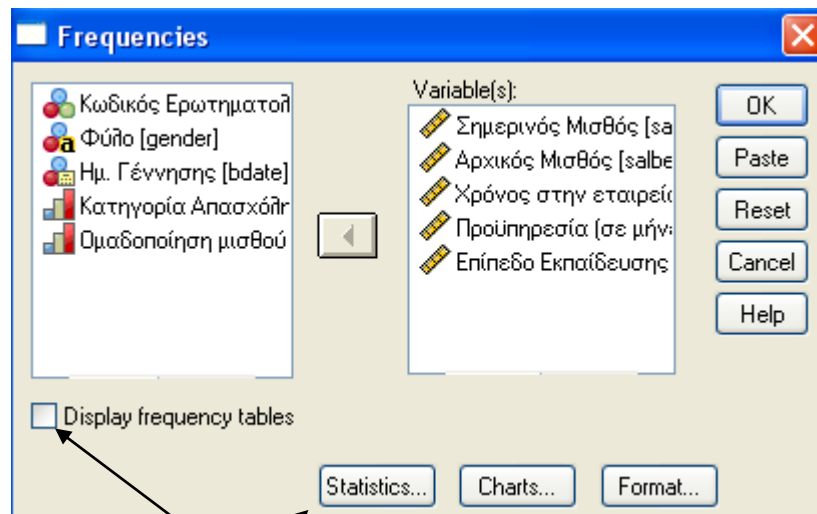
5. Περιγραφικά στατιστικά

Τα περιγραφικά στατιστικά αποτελούν το πρώτο βήμα στην ανάλυση **ποσοτικών δεδομένων** γιατί μας βοηθάνε να περιγράψουμε τη μέση τάση και τη διασπορά των δεδομένων.

5.1 Μέτρα κεντρικής τάσης-Μέτρα διασποράς



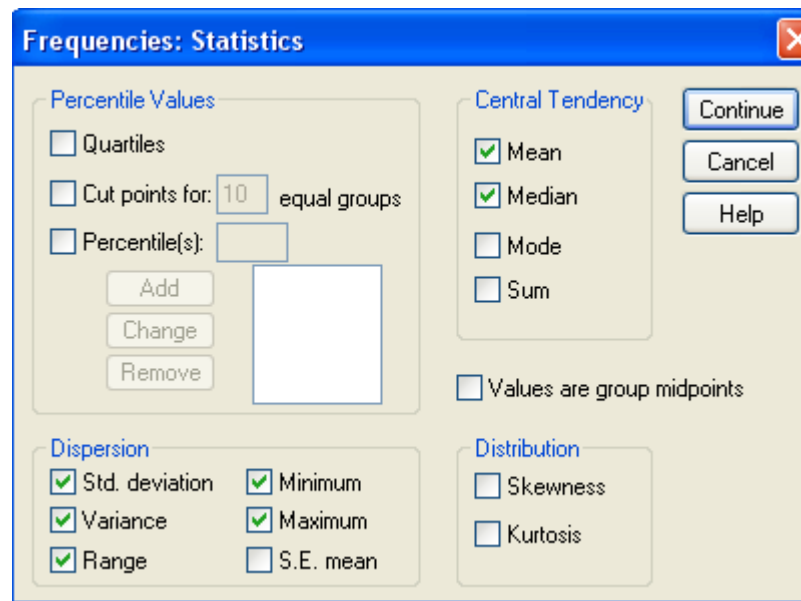
Η παραπάνω εντολή είναι μια από τις επιλογές που υπάρχουν για τον υπολογισμό των περιγραφικών στατιστικών.



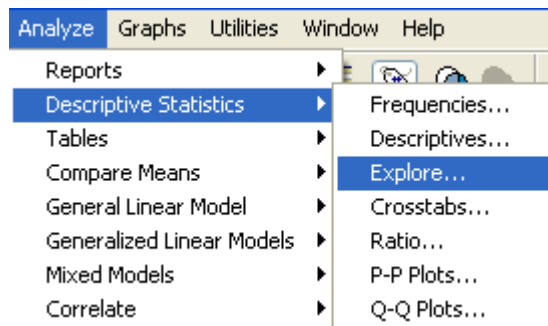
Θα πρέπει να αποεπιλέξετε το **Display frequency tables**.

Με την επιλογή **statistics**, προκύπτει το παρακάτω menu, όπου επιλέγουμε **mean** για τον αριθμητικό μέσο, **median** για τη διάμεσο, **range** για το εύρος, **minimum**,

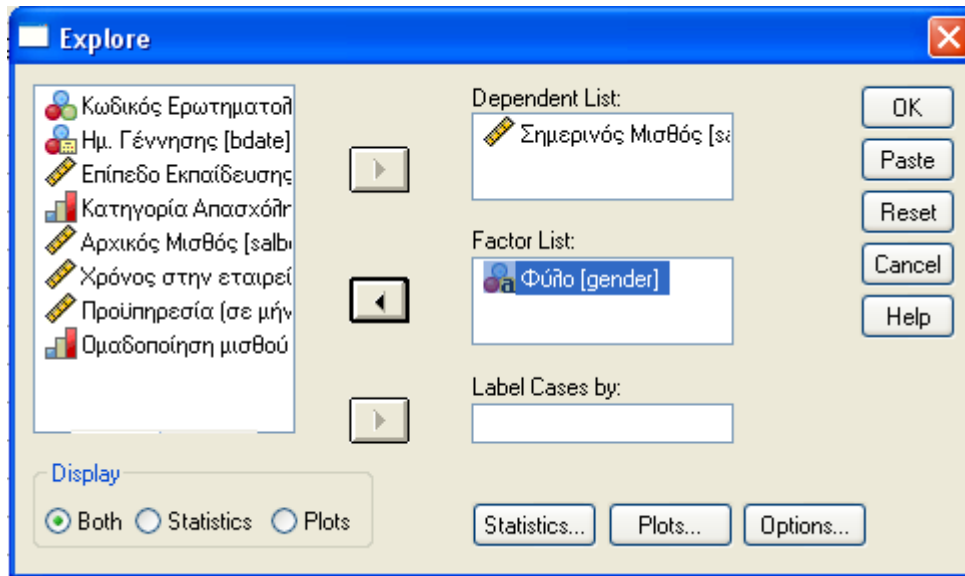
maximum, **Std. Deviation** για την τυπική απόκλιση και **Variance** για τη διασπορά.



Εναλλακτικά μπορείτε να υπολογίσετε τα ίδια μέτρα και επιπλέον το **διάστημα εμπιστοσύνης του μέσου όρου** με την εντολή **Analyze -Descriptive Statistics – Explore**.



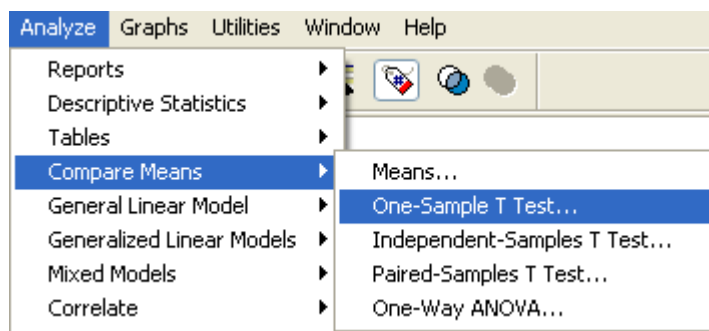
Στην συγκεκριμένη διαδικασία μπορούμε να χρησιμοποιήσουμε και μια ποιοτική μεταβλητή στο **Factor List** για να υπολογίσουμε τα περιγραφικά στατιστικά χωρίζοντας το δείγμα μας σε δυο τουλάχιστον υποομάδες (π.χ. άνδρες, γυναίκες).



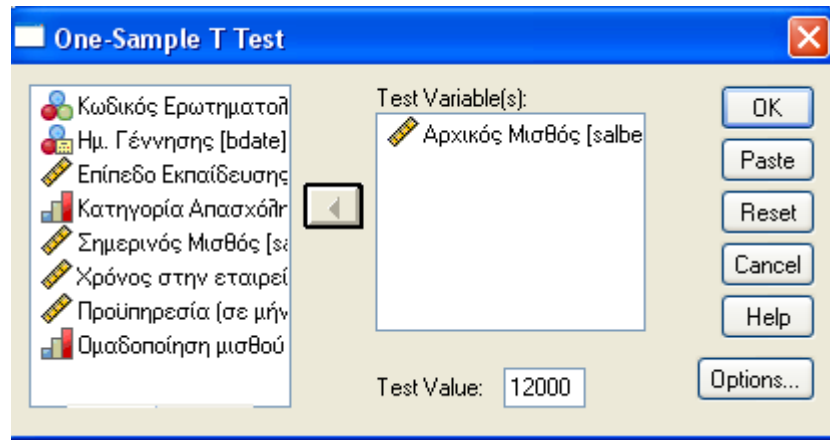
6. Έλεγχοι Υποθέσεων

Οι έλεγχοι υποθέσεων εντάσσονται στην επαγωγική στατιστική την οποία χρησιμοποιούμε όταν επιλέγουμε ένα δείγμα και προσπαθούμε να γενικεύσουμε τα αποτελέσματά μας σε όλο τον υπό μελέτη πληθυσμό.

6.1 Σύγκριση μέσω όρων ενός δείγματος και ενός γνωστού πληθυσμού



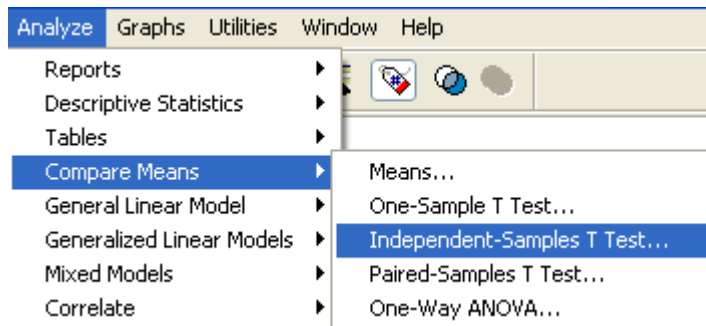
Ο πρώτος έλεγχος υποθέσεων αφορά την σύγκριση του μέσου όρου ενός δείγματος με τον μέσο όρο ενός γνωστού πληθυσμού.

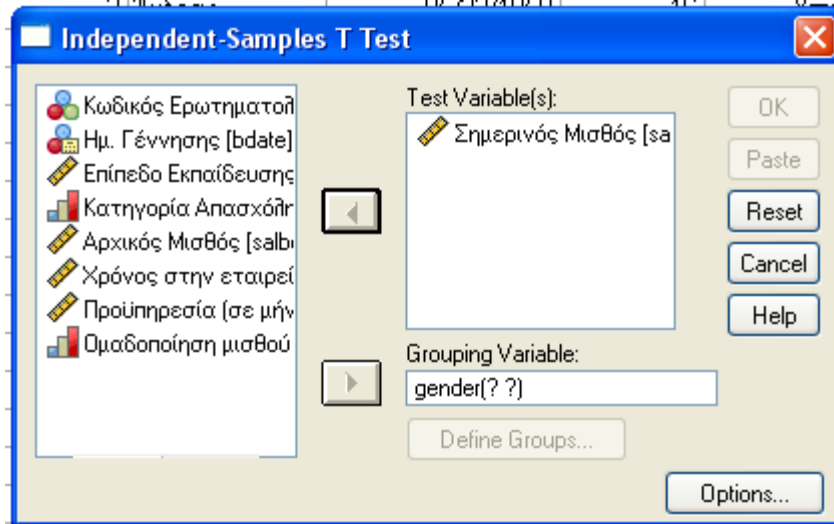


Θα πρέπει να τοποθετήσουμε την ποσοτική μεταβλητή που μελετάμε στο **Test Variable** και την τιμή του γνωστού μέσου όρου του πληθυσμού στο **Test Value**.

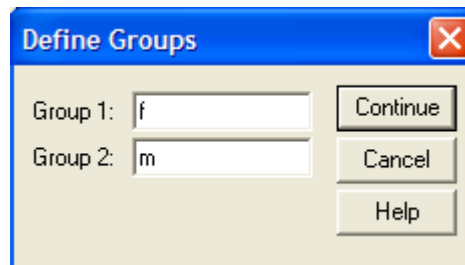
6.2 Σύγκριση μέσων όρων δυο ανεξάρτητων δειγμάτων

Ο δεύτερος έλεγχος υποθέσεων αφορά την σύγκριση των μέσων όρων δυο ανεξάρτητων δειγμάτων.



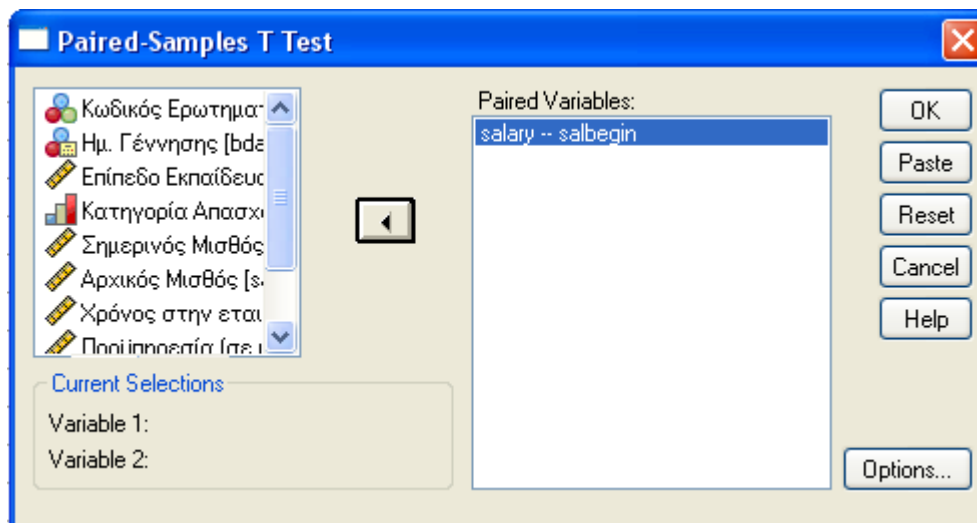
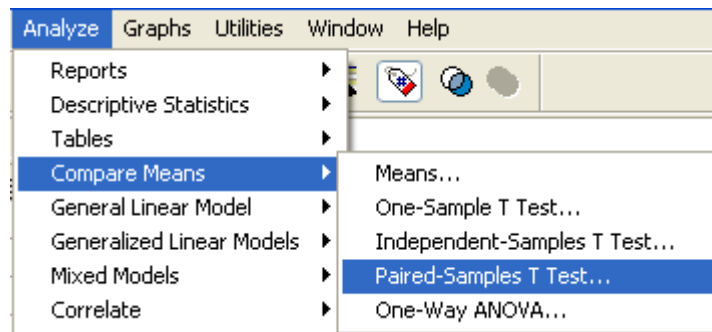


Στο **Grouping Variable** τοποθετούμε την ποιοτική μεταβλητή που χωρίζει το δείγμα μας σε δυο υποομάδες (π.χ. φύλο) και στο Define Groups ορίζουμε τον τρόπο καταχώρησης της ποιοτικής μεταβλητής (π.χ. f,m ή 1,2).



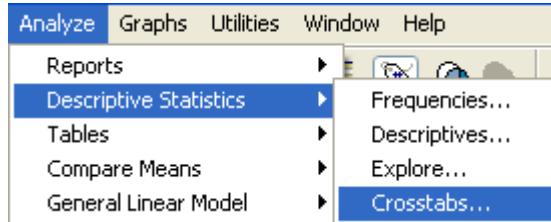
6.3 Σύγκριση μέσων όρων δυο εξαρτημένων δειγμάτων

Ο τρίτος έλεγχος υποθέσεων αφορά την σύγκριση των μέσων όρων δυο εξαρτημένων δειγμάτων.

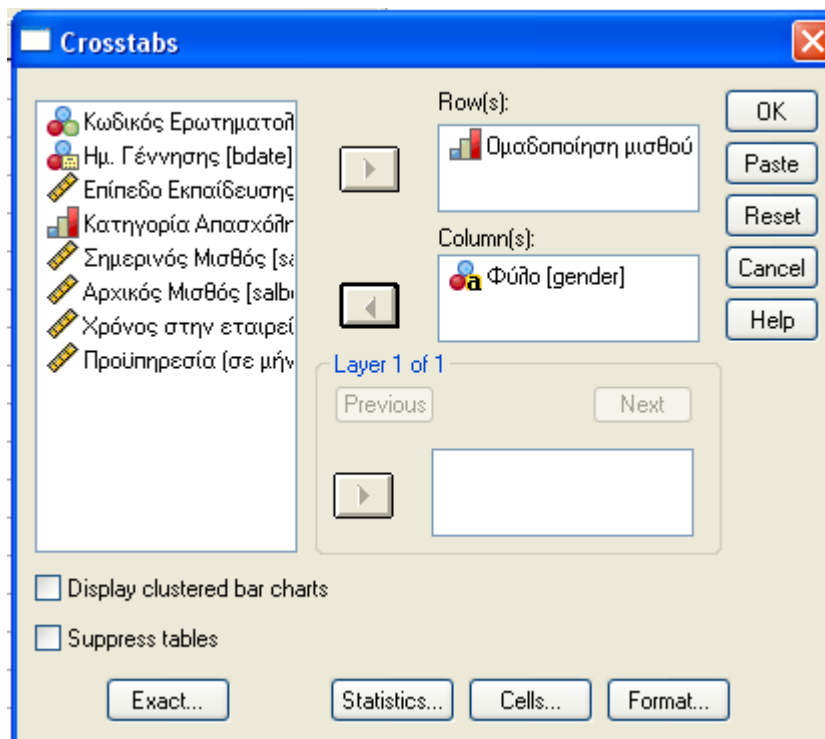


7. Διασταυρώσεις μεταβλητών και έλεγχος χ^2

7.1 Δημιουργία πίνακα διπλής εισόδου

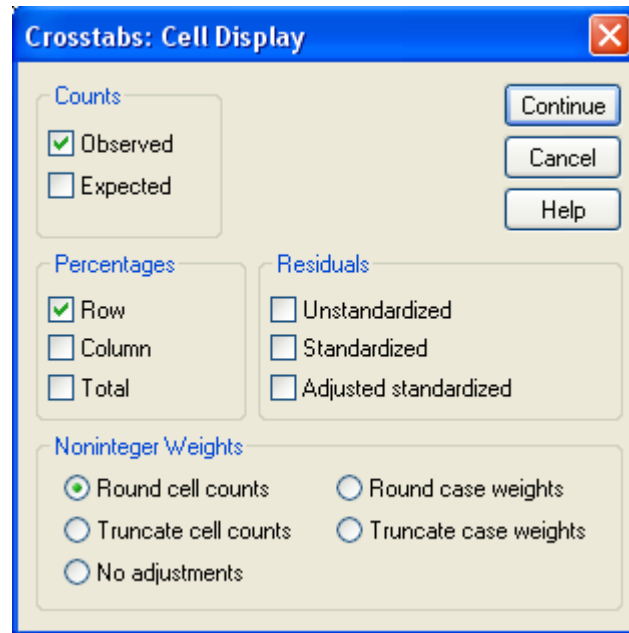


Με την διασταύρωση μεταβλητών δημιουργούμε ένα πίνακα διπλής εισόδου που μας βοηθάει να περιγράψουμε την σχέση που υπάρχει ανάμεσα σε δυο **ποιοτικές μεταβλητές**.

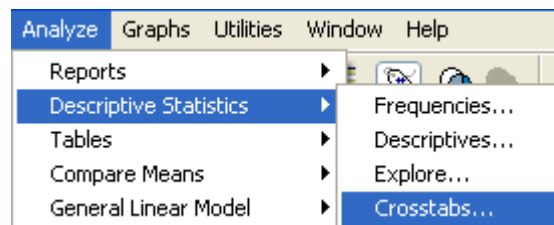


Στις γραμμές (Rows) τοποθετούμε την εξαρτημένη μεταβλητή και στις στήλες (Columns) την ανεξάρτητη. Αν δεν υπάρχει σχέση αιτίου-αιτιατού ανάμεσα στις δυο μεταβλητές τότε στις γραμμές τοποθετούμε την μεταβλητή που θέλουμε να δώσουμε έμφαση.

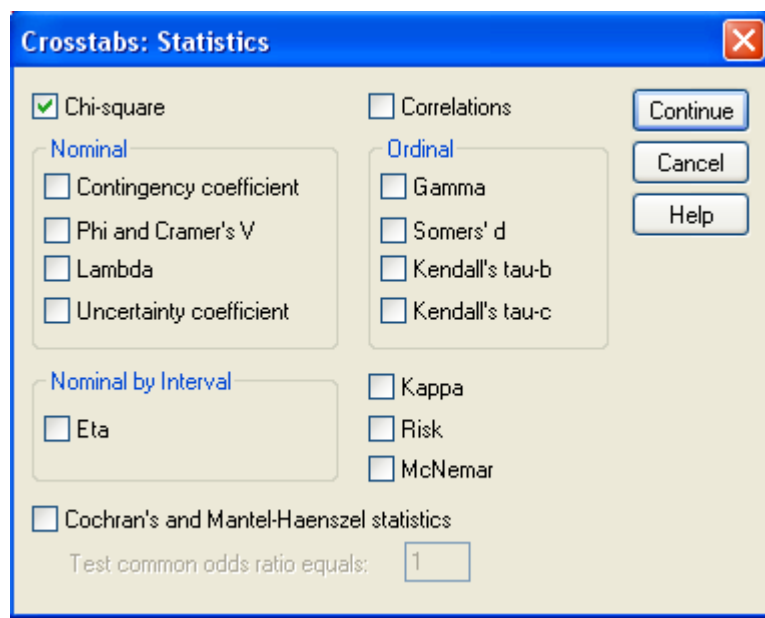
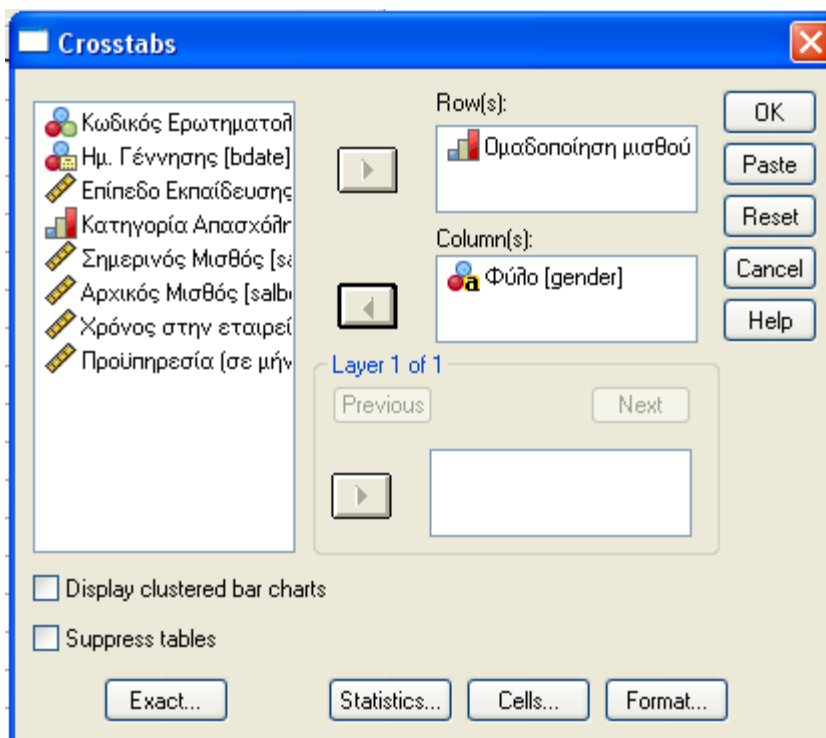
Για τον υπολογισμό και των ποσοτών στον πίνακα διπλής εισόδου θα επιλέξετε το **Cells** και εκτός από το Observed στο Counts θα επιλέξετε και το Row από το Percentages για να εμφανιστούν και τα ποσοστά σε σχέση με την εξαρτημένη μεταβλητή.



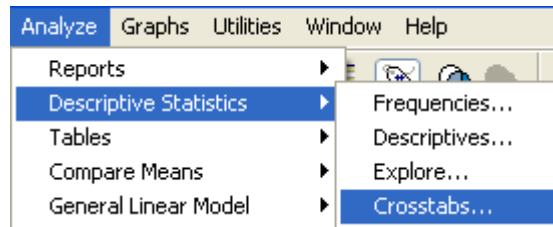
7.2 Έλεγχος ανεξαρτησίας χ^2



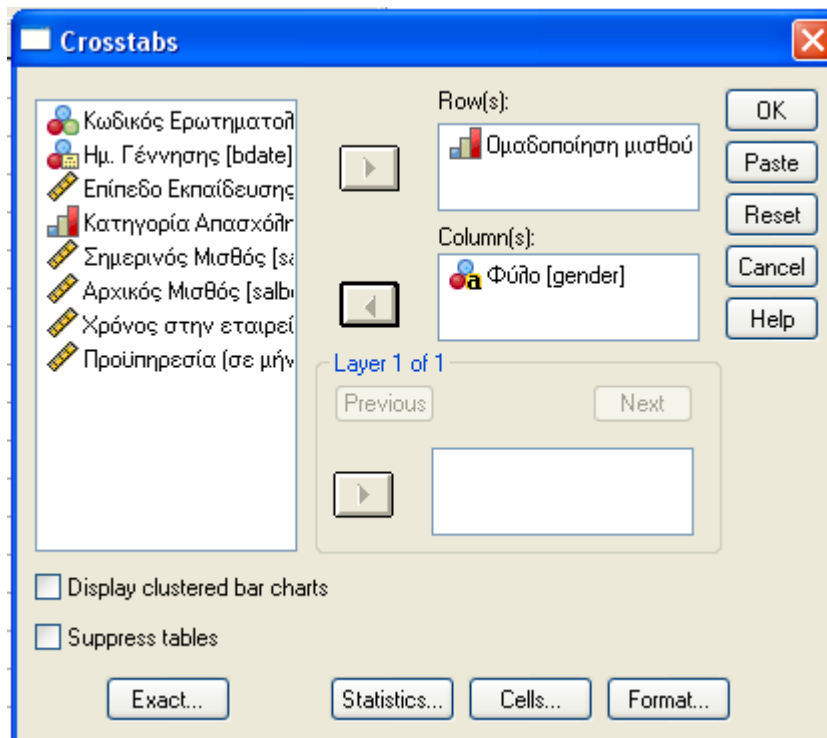
Για να κάνουμε τον έλεγχο ανεξαρτησίας μετά από την παραπάνω εντολή θα πρέπει να πατήσουμε το κουμπί **Statistics** και να επιλέξουμε το **Chi-square**.



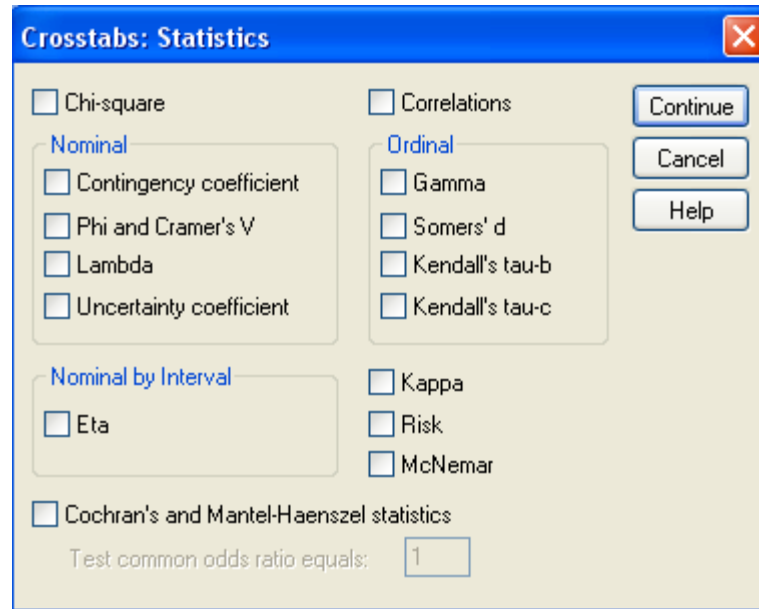
8. Συντελεστές συσχέτισης ποιοτικών δεδομένων



Οι συντελεστές συσχέτισης θα βοηθήνε να αξιολογήσουμε την ένταση της συσχέτισης που υπάρχει ανάμεσα σε δυο **ποιοτικές μεταβλητές**.



Μετά από την παραπάνω εντολή πατάμε το κουμπι **Statistics** και εμφανίζονται όλοι οι συντελεστές συσχέτισης ανάλογα με την κλίμακα μέτρησης των μεταβλητών.

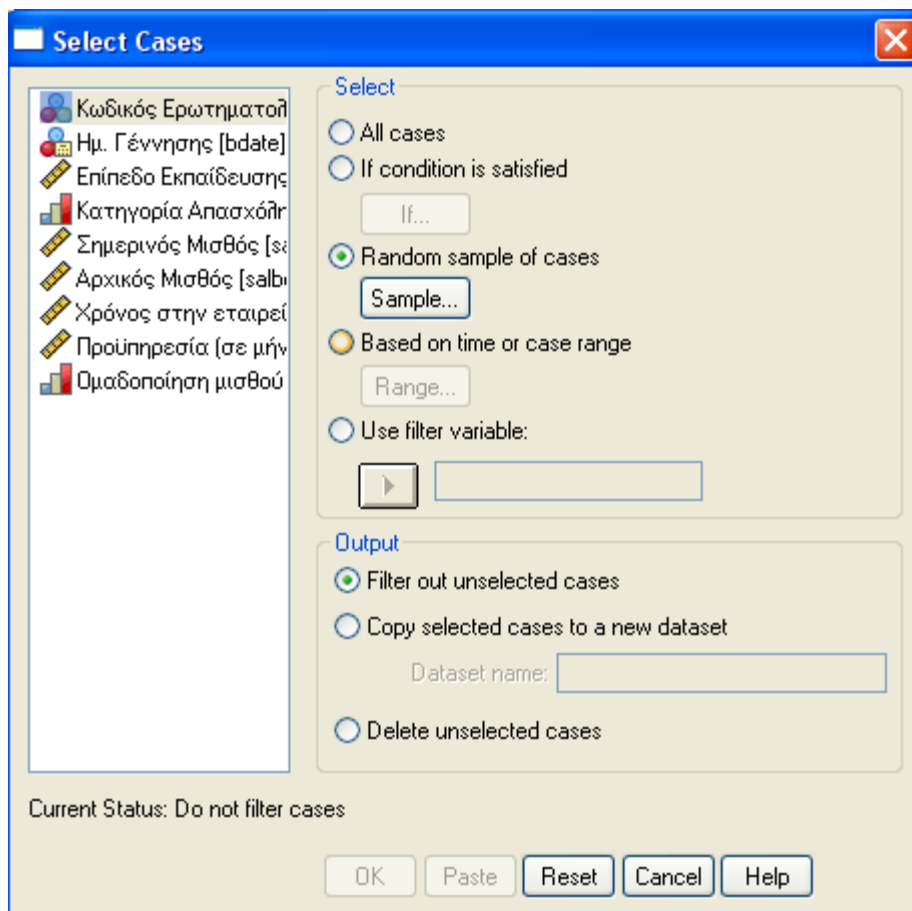


9. Δειγματοληψία

9.1 Απλή τυχαία Δειγματοληψία

Η απλή τυχαία δειγματοληψία χρησιμοποιεί του πίνακες με τους τυχαίους αριθμούς για την επιλογή ενός τυχαίου δείγματος.





Μπορούμε να επιλέξουμε είτε κάποιο συγκεκριμένο ποσοστό δείγματος είτε έναν ακριβή αριθμό.

