

Σημειώσεις για το μάθημα

# Εργαστήριο στατιστικής Στατιστικό πακέτο S.P.S.S.

Παπάνα Αγγελική

E-mail: [papanagel@yahoo.gr](mailto:papanagel@yahoo.gr), [agpapanana@gen.auth.gr](mailto:agpapanana@gen.auth.gr)

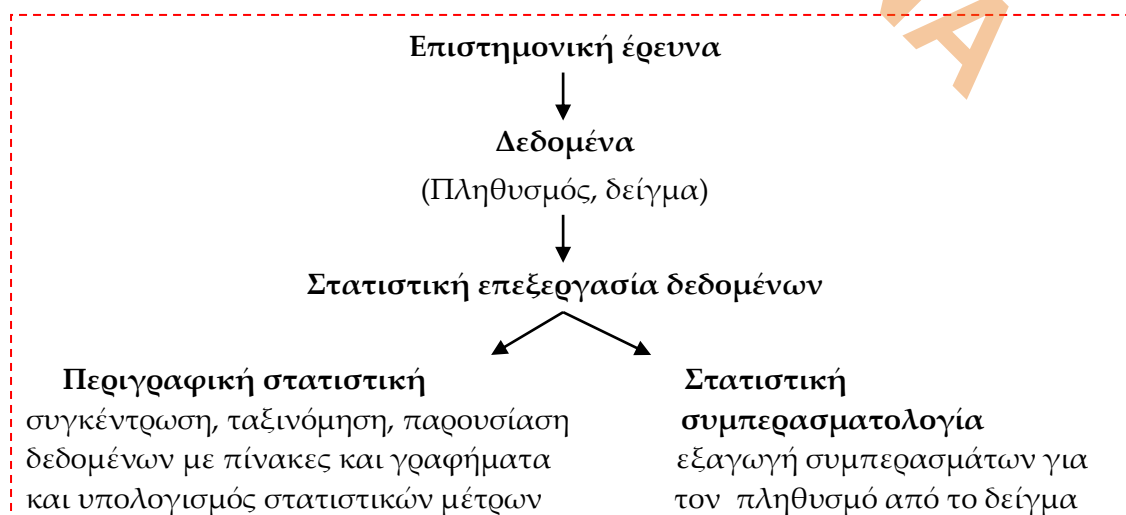
Α.Τ.Ε.Ι. Θεσσαλονίκης  
ΠΑΡΑΡΤΗΜΑ ΚΑΤΕΡΙΝΗΣ  
Τμήμα Τυποποίησης και Διακίνησης Προϊόντων (Logistics)

## Κεφάλαιο 1

### Εισαγωγή στην στατιστική ανάλυση

Η **στατιστική** είναι η επιστήμη που ασχολείται με τη συγκέντρωση, παρουσίαση, αξιολόγηση και επεξεργασία της πληροφορίας, με στόχο την εξαγωγή συμπερασμάτων και χρησιμεύει σε όλες της επιστήμες, όπως στην οικονομία, κοινωνιολογία, ιατρική, βιολογία, ψυχολογία, κτλ. Η στατιστική απαιτεί πολύ καλή γνώση του φαινομένου που μελετά, σωστή οργάνωση των παρατηρήσεων και σωστή καταγραφή των δεδομένων. Οι πληροφορίες που χρησιμοποιεί πρέπει να είναι πλήρεις, ποιοτικά ελεγμένες και χωρίς λάθη. Για την επεξεργασία των δεδομένων απαιτείται θεωρητική γνώση ώστε να επιλεγούν οι σωστές τεχνικές και μέθοδοι για την στατιστική ανάλυση που επιδιώκουμε αλλά και να πληρούνται οι κατάλληλες προϋποθέσεις για την εφαρμογή των στατιστικών τεχνικών.

Οι **μετρήσεις** ή αλλιώς τα **δεδομένα** μιας επιστημονικής έρευνας είναι συγκεκριμένες πληροφορίες, στοιχεία ή γεγονότα, οι οποίες μπορούν να χρησιμοποιηθούν για περαιτέρω αποφάσεις και συμπεράσματα. Γεγονός είναι ότι όλες οι ερευνητικές προσπάθειες οι οποίες παράγουν δεδομένα επιδέχονται στατιστική επεξεργασία. Η επεξεργασία αυτή διακρίνεται σε δύο φάσεις. Οι μέθοδοι της **περιγραφικής στατιστικής** αποτελούν το επιστημονικό εργαλείο για την συγκέντρωση, ταξινόμηση και παρουσίαση των πρωτογενών δεδομένων, ενώ τα εργαλεία της **στατιστικής συμπερασματολογίας** επιτρέπουν την εκτίμηση χαρακτηριστικών του συνόλου των δεδομένων από την μελέτη ενός υποσυνόλου τους .



Η ανάγκη για άμεση εφαρμογή της στατιστικής σε ποικίλες επιστήμες οδήγησε στην δημιουργία στατιστικών πακέτων τα οποία έχουν την δυνατότητα εισαγωγής, επεξεργασίας, ανάλυσης και παρουσίασης δεδομένων σε σύντομο χρονικό διάστημα. Φυσικά χάρη στους υπολογιστές υπάρχουν απεριόριστες πλέον δυνατότητες όσον αφορά τον όγκο των δεδομένων.

Το **S.P.S.S.** (Statistical Package for Social Sciences) είναι ένα στατιστικό πακέτο που έχει πολλές δυνατότητες όσον αφορά την επεξεργασία και παρουσίαση των δεδομένων μιας επιστημονικής έρευνας αλλά και μεγάλη αξιοπιστία. Οι τελευταίες εκδόσεις του **S.P.S.S.** έχουν γραφικό περιβάλλον, πράγμα που το καθιστά πολύ εύκολο για την χειριστή του.

**Σκοπός αυτού του μαθήματος είναι η περιγραφή μεθόδων καταχώρησης, παρουσίασης, ανάλυσης και επεξεργασίας των δεδομένων μιας έρευνας καθώς και η εκμάθηση κάποιων ελέγχων υποθέσεων για διάφορα σύνολα δεδομένων. Περιγράφονται στατιστικές διαδικασίες καθώς και η χρήση του S.P.S.S. σε αυτές.**

Η συνηθισμένη προσέγγιση σε ένα πρόβλημα που απαιτεί στατιστική ανάλυση είναι στην κατεύθυνση της οργάνωσης της παρατήρησης, της καταγραφής των αποτελεσμάτων και της διατύπωσης της θεωρίας. Τα συστατικά αυτής της προσέγγισης είναι ο **πληθυσμός**, το **δείγμα** και η **μεταβλητή**.

❑ **Πληθυσμός** είναι το σύνολο των ατόμων ή αντικειμένων για τα οποία ενδιαφερόμαστε να βγάλουμε συμπεράσματα σε σχέση με κάποιες ιδιότητες που αφορούν τα στοιχεία του.

Ο πληθυσμός πρέπει να είναι καλά ορισμένος, έτσι ώστε κάθε στοιχείο να μπορεί να αποφασιστεί μονοσήμαντα αν είναι μέλος του.

Τα στοιχεία του πληθυσμού αναφέρονται ως **υποκείμενα**.

❑ Όταν ένας πληθυσμός είναι πρακτικά άπειρος ή μεταβάλλεται κατά την διάρκεια πραγματοποίησης μιας έρευνας, τότε μελετάτε κάποιο υποσύνολο του πληθυσμού, το οποίο ονομάζεται **δείγμα**.

Από την μελέτη του δείγματος θέλουμε να εξάγουμε συμπεράσματα για τον πληθυσμό.

Το δείγμα που επιλέγεται πρέπει να είναι **αντιπροσωπευτικό**, δηλαδή πρέπει να είναι τυχαίο, δηλαδή να μην υπάρχει μεροληψία στην επιλογή των υποκειμένων του πληθυσμού.

❑ Τα χαρακτηριστικά ενός πληθυσμού τα οποία μελετάμε αναφέρονται στην βιβλιογραφία με τον όρο **μεταβλητή**. Οι μεταβλητές χωρίζονται σε δύο βασικές κατηγορίες με βάση τις «τιμές» που παίρνουν.

(1) Οι **ποιοτικές μεταβλητές** παίρνουν «τιμές» που δεν έχουν αριθμητικές ιδιότητες. Διακρίνονται σε **μεταβλητές κατηγορίας (nominal)** και σε **μεταβλητές διάταξης (ordinal)**.

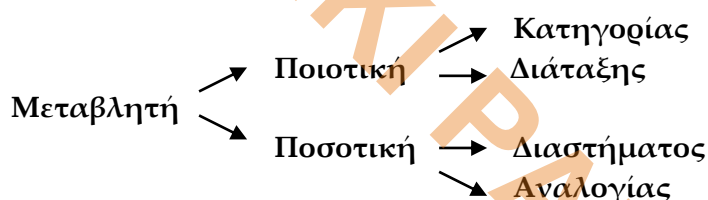
α) **Κατηγορίας** είναι οι μεταβλητές που το σύνολο τιμών τους δεν έχει καμία ιδιότητα, π.χ. χρώμα ματιών, φύλο, τόπος γέννησης.

β) **Διάταξης** είναι οι μεταβλητές που για το σύνολο τιμών τους μπορούμε να ορίσουμε μια διάταξη, π.χ. σειρά κατάταξης σε ένα αγώνισμα, επίπεδο εκπαίδευσης, κλίμακα σεισμών RICHTER, κτλ.

(2) Οι **ποσοτικές μεταβλητές** παίρνουν αριθμητικές «τιμές» και εκφράζονται με μια μονάδα μέτρησης. Διακρίνονται σε **μεταβλητές διαστήματος (interval)** και σε **μεταβλητές αναλογίας (ratio)**.

α) **Διαστήματος** είναι οι μεταβλητές των οποίων οι ίσες διαφορές μεταξύ των τιμών τους συνεπάγονται και ίσες διαφορές για το χαρακτηριστικό που μετράει η μεταβλητή. Η διαστημική κλίμακα επιτρέπει την ιεράρχηση των υποκειμένων π.χ. θερμοκρασία, ηλικία, κλίμακα βαθμών Κελσίου κτλ.

β) **Αναλογίας** είναι οι μεταβλητές των οποίων οι τιμές αντιστοιχούν αναλογικά στην ποσότητα του χαρακτηριστικού που μετρούν. Το μηδέν ανήκει πάντα στο σύνολο τιμών της μεταβλητής και δηλώνει την 'απουσία' του χαρακτηριστικού, π.χ. ταχύτητα, ύψος, ημερήσιος τζίρος μιας εταιρείας κτλ.



Οι ποσοτικές μεταβλητές διακρίνονται σε **συνεχείς (continuous)** και **διακριτές (discrete)**.

**Συνεχής** είναι μια μεταβλητή όταν παίρνει τιμές από ένα διάστημα, δηλαδή για κάθε δυο τιμές της υπάρχει μια τιμή ανάμεσά τους, π.χ. ύψος, βάρος.

**Διακριτή** είναι μια μεταβλητή που για κάθε δυο τιμές της δεν υπάρχει πάντα μια τιμή της μεταβλητής ανάμεσά τους, π.χ. πλήθος παιδιών μιας οικογένειας.

Η επιλογή των στατιστικών τεχνικών εξαρτάται κατά κύριο λόγο από τον τύπο των μεταβλητών που εξετάζονται.

## Κεφάλαιο 2

### Data Editor

Για την καταχώρηση και επεξεργασία των δεδομένων μιας πειραματικής έρευνας, το πρώτο στάδιο είναι η καταχώρηση των δεδομένων στον **Data Editor του S.P.S.S.**, δηλαδή στο περιβάλλον του S.P.S.S.

**Μεταβλητές**

**Περιπτώσεις / Παρατηρήσεις**

	VAR00001	VAR00002	VAR00003	VAR00004	VAR00005
1	1,00	95,00	9,00	30,00	9,00
2	,00	80,00	7,00	20,00	7,00
3	,00	30,00	7,00	5,00	3,00
4	1,00	20,00	4,00	.	2,00
5	1,00	75,00	3,00	15,00	5,00
6	1,00	75,00	6,00	20,00	6,00
7	,00	25,00	3,00	5,00	2,00
8	1,00	25,00	5,00	8,00	2,00
9	,00	45,00	.	10,00	5,00
10	1,00	30,00	6,00	5,00	3,00

SPSS Processor is ready

Δηλαδή σε κάθε στήλη καταχωρούμε τις απαντήσεις της ίδιας ερώτησης, ενώ κάθε γραμμή αντιστοιχεί στις απαντήσεις κάθε υποκειμένου.

## Τα βασικά μενού του Data Editor

### Κύρια γραμμή επιλογής

**File** με βασικές εντολές:

**New** : δημιουργία νέου αρχείου δεδομένων

**Open** : άνοιγμα αρχείων δεδομένων

**Save & Save as** : σώζουμε το αρχείο δεδομένων που δημιουργήσαμε ή ανοίξαμε

**Print** : εκτυπώνουμε

**Exit** : βγαίνουμε από το πρόγραμμα

**Edit** με βασικές εντολές:

**Undo** : αναίρεση της τελευταίας ενέργειας

**Cut** : αποκοπή

**Copy** : αντιγραφή

**Paste** : επικόλληση

**Find** : αναζήτηση εγγραφής

**Options** : Γενικές ρυθμίσεις

**View** με βασικές εντολές:

**Status bar** : γραμμή κατάστασης

**Toolbars** : διαμόρφωση γραμμών εντολών

**Fonts** : διαμόρφωση γραμματοσειρών

**Grid lines** : εμφάνιση - απόκρυψη περιγραμμάτων κελιών

**Value labels** : εμφάνιση – απόκρυψη *labels*

**Data** με βασικές εντολές:

**Insert case** : εισαγωγή νέας γραμμής

**Insert variable** : εισαγωγή νέας στήλης

**Got to case** : γρήγορη μετάβαση στον αριθμό της γραμμής που θέλουμε

**Sort case** : ταξινόμηση δεδομένων κατά αύξουσα ή φθίνουσα σειρά

**Transform** αλλαγές σε μεταβλητές π.χ. επανακωδικοποίηση τιμών, πράξεις μεταξύ μεταβλητών

**Analyze** Στατιστική ανάλυση π.χ. υπολογισμός στατιστικών μέτρων

**Graphs** Γραφικές παραστάσεις

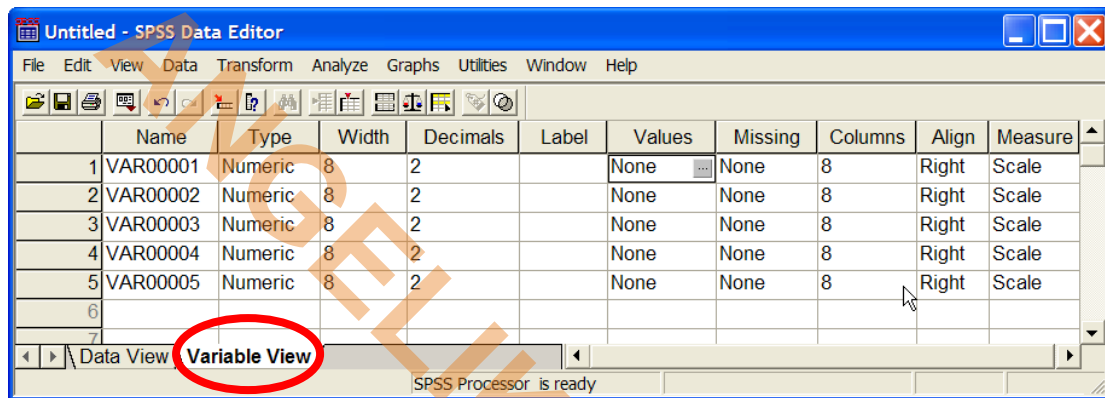
- Utilities** δημιουργία ενός πίνακα περιεχομένων για τις μεταβλητές
- Windows** πληροφορίες για αρχεία σχετικά με το πρόγραμμα π.χ. αρχεία δεδομένων, αποτελεσμάτων, γραφημάτων και εντολών που ενεργοποιήθηκαν κατά τη διάρκεια της στατιστικής ανάλυσης
- Help** βοήθεια

ANGELIKI PAPANA

## Κεφάλαιο 3

### ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ

Η καταχώρηση των δεδομένων ενός ερωτηματολογίου στον Data Editor του SPSS γίνεται επιλέγοντας το **Variable View**.



Τα χαρακτηριστικά της μεταβλητής που πρέπει να οριστούν είναι τα παρακάτω:

<b>Name</b>	Όνομα μεταβλητής
<b>Type</b>	Τύπος τιμών που παίρνει η μεταβλητή
<b>Width</b>	Μέγιστος αριθμός ψηφίων ή χαρακτήρων που παίρνουν οι τιμές της μεταβλητής
<b>Decimal</b>	Μέγιστο πλήθος δεκαδικών ψηφίων τιμών της μεταβλητής
<b>Label</b>	Τίτλος ερώτησης
<b>Values</b>	Αν θέλω να κωδικοποιήσω τις τιμές της μεταβλητής
<b>Missing</b>	Αν λείπουν τιμές ή δεν θέλω να συμπεριλάβω κάποιες τιμές
<b>Align</b>	Στοίχιση μέσα στα κελιά
<b>Measure</b>	Scale



## (1) ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ

Χωρίς κωδικοποίηση των τιμών της μεταβλητής

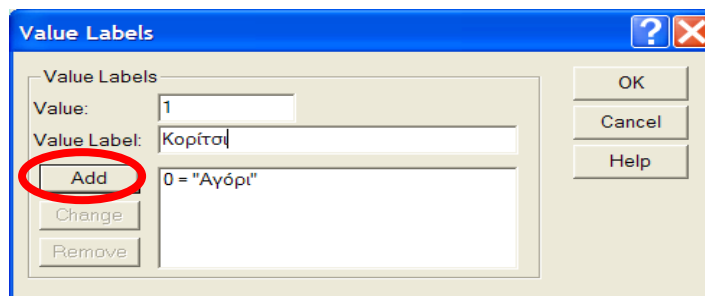
Type	String (characters)
Width	μέγιστο πλήθος χαρακτήρων της μεταβλητής
Decimal	0
Label	Τίτλος ερώτησης
Values	-
Missing	αν δεν θέλω να συμπεριλάβω κάποιες τιμές
Align	στοίχιση μέσα στα κελιά
Measure	nominal ή ordinal

## (2) ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ

Με κωδικοποίηση των τιμών της μεταβλητής

Type	Numeric
Width	1: μονοψήφιο πλήθος κωδικών, 2: διψήφιο πλήθος κωδικών
Decimal	0
Label	Τίτλος ερώτησης
Values	Κωδικοποίηση
Missing	αν δεν θέλω να συμπεριλάβω κάποιες τιμές
Align	στοίχιση μέσα στα κελιά
Measure	nominal ή ordinal

Το παρακάτω πλαίσιο διαλόγου είναι για την κωδικοποίηση των τιμών της μεταβλητής "φύλο". Στο πάνω μέρος, δίπλα στο "Value" γράφουμε τον 'κωδικό' και στο "Value Label" την κάθε 'τιμή' της μεταβλητής, δηλ. 'αγόρι' και 'κορίτσι'. Στη συνέχεια τα 'προσθέτουμε' πατώντας το κουμπί "Add" και συνεχίζουμε με τον επόμενο 'κωδικό'. Εδώ έχουμε ήδη ορίσει την τιμή 0 για το "αγόρι" και μένει να πατήσουμε "Add" για να ορισθεί η τιμή 1 για το "κορίτσι". Όταν τελειώσουμε με όλες τις τιμές, πατάμε "OK".



Για να εμφανίζονται οι απαντήσεις του ερωτηματολογίου και όχι οι κωδικοί, από το κύριο μενού επιλογών: **View** → **Value Labels** (τσεκάρισμα)

### (3) ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΔΙΑΚΡΙΤΩΝ ΠΟΣΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ

Type	Numeric
Width	Μέγιστος αριθμός ψηφίων αριθμού
Decimal	Μέγιστος αριθμός δεκαδικών ψηφίων αριθμού
Label	Τίτλος ερώτησης
Values	-
Missing	αν δεν θέλω να συμπεριλάβω κάποιες τιμές
Align	στοίχιση μέσα στα κελιά
Measure	Scale

### (4) ΚΑΤΑΧΩΡΗΣΗ ΔΕΔΟΜΕΝΩΝ ΣΥΝΕΧΩΝ ΠΟΣΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΟΜΑΔΟΠΟΙΗΣΗ ΤΩΝ ΤΙΜΩΝ ΜΕΤΑΒΛΗΤΩΝ

#### 1<sup>Η</sup> ΜΕΘΟΔΟΣ

- Κωδικοποίηση ομάδων.

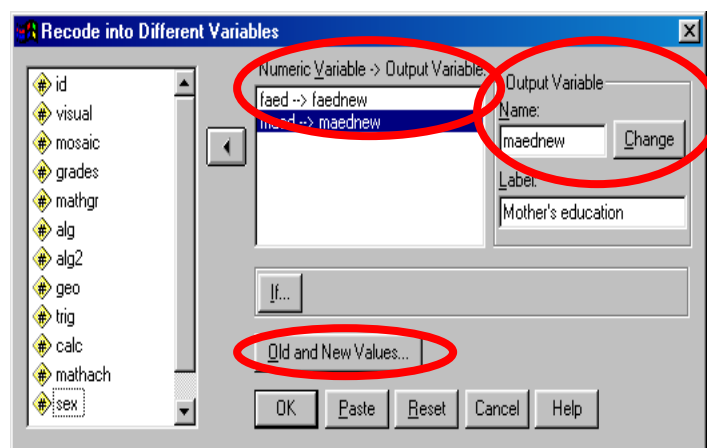
- Variable View
  - Name
  - Type (numeric)
  - Label
  - Values (κωδικοποίηση απαντήσεων)
  - Measure (ordinal)

#### 2<sup>Η</sup> ΜΕΘΟΔΟΣ

Εισαγωγή δεδομένων κανονικά σε μια στήλη.

Από το μενού του Data Editor :

- Transform
- Recode
- Into Different Variables



## (5) ΚΑΤΑΧΩΡΗΣΗ ΠΟΙΟΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΕΡΩΤΗΣΕΙΣ ΠΟΛΛΑΠΛΩΝ ΑΠΑΝΤΗΣΕΩΝ

### 1<sup>Η</sup> ΜΕΘΟΔΟΣ

Δημιουργούμε τόσες μεταβλητές όσο είναι το μέγιστο πλήθος πολλαπλών απαντήσεων. Για κάθε μια από τις μεταβλητές αυτές καθορίζουμε το είδος τους.

**Variable View** → **Name**  
→ **Type** (numeric)  
→ **Label**  
→ **Values** (κωδικοποίηση απαντήσεων)  
→ **Measure** (nominal or ordinal)

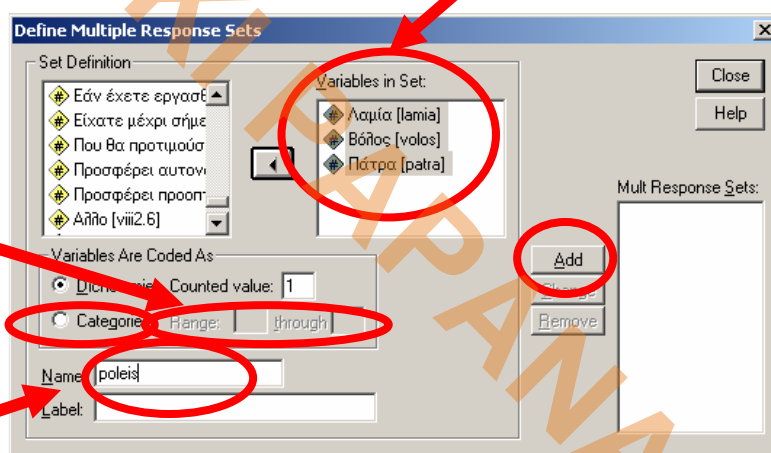
Από το μενού → **Analyze**  
**Multiple Response**  
**Define Variable Sets**

Το παράθυρο που ανοίγει είναι το παρακάτω:

Μεταφέρουμε τις μεταβλητές που αναφέρονται στην ίδια ερώτηση

Εύρος των κωδικών της μεταβλητής

Καθορίζουμε το όνομα του 'set'



### 2<sup>Η</sup> ΜΕΘΟΔΟΣ

Κωδικοποίηση απαντήσεων με 0 ή 1. Δημιουργούμε τόσες μεταβλητές όσο είναι το πλήθος των πολλαπλών απαντήσεων. Για κάθε μια από τις μεταβλητές αυτές καθορίζουμε το είδος τους. Από το μενού επιλέγουμε:

**Analyze** → **Multiple Response** → **Define Variable Sets**

Στο παράθυρο που ανοίγει (το ίδιο με πριν), επιλέγουμε: **Variables are Coded as** και **τσεκάρουμε**:

☒ **Dichotomies**    **Counted Values: 1**

## Κεφάλαιο 4

### ΠΑΡΟΥΣΙΑΣΗ ΔΕΔΟΜΕΝΩΝ

#### ΠΑΡΟΥΣΙΑΣΗ ΔΕΔΟΜΕΝΩΝ

Στατιστικοί πίνακες /  
Πίνακες συχνотήτων

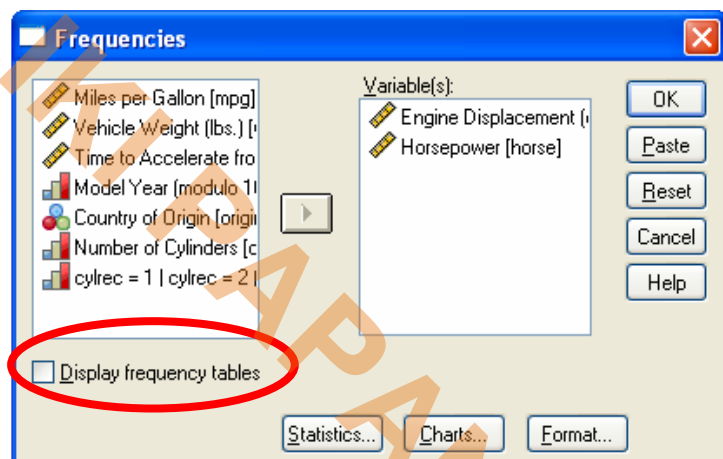
Γραφικές  
παραστάσεις

#### (1) ΠΙΝΑΚΕΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

##### 1.1 Πίνακες κατανομής συχνотήτων: απλές μεταβλητές

Από το μενού

Analyze  
Descriptive Statistics  
Frequencies



#### Output

gender ?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	8	53,3	53,3	53,3
	K	7	46,7	46,7	100,0
	Total	15	100,0	100,0	

❑ **Frequency:** συχνότητα (εκφράζει πλήθος)

❑ **Percent:** ποσοστό (σχετική συχνότητα)

❑ **Valid percent:** πραγματική σχετική συχνότητα

Αν δεν υπάρχουν ελλειπούσες τιμές, το Valid Percent ταυτίζεται με το Percent

❑ **Cumulative percent:** αθροιστικό ποσοστό

Αν υπάρχουν ελλειπούσες τιμές:

❑ **Valid percent = Frequency / Valid cases**

❑ **Cumulative percent:** υπολογίζεται με βάση το Valid percent

## 1.2 Πίνακες κατανομής συχνοτήτων: μεταβλητές με πολλαπλές απαντήσεις

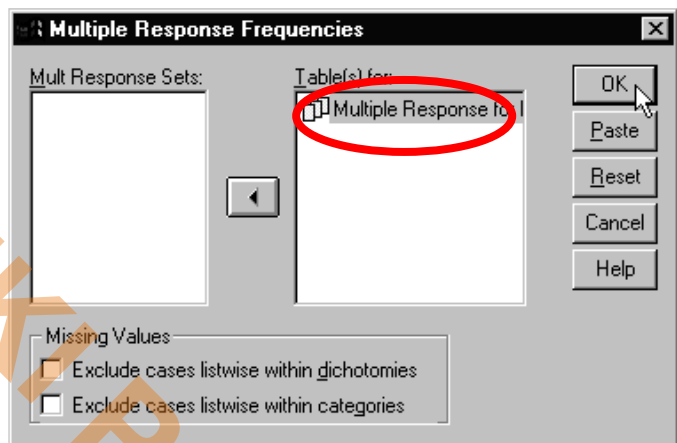
Καταρχήν δηλώνουμε ποιες μεταβλητές αναφέρονται στην ίδια ερώτηση.

Από το μενού του Data Editor επιλέγουμε:

**Analyze → Multiple Response → Define Sets**

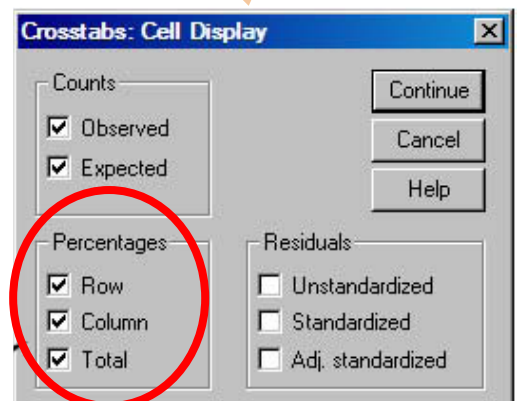
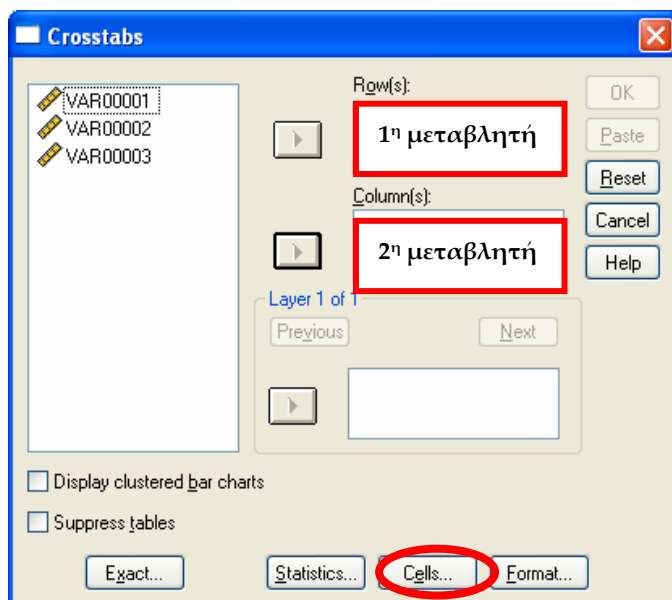
Έπειτα επιλέγουμε από το μενού

**Analyze  
Multiple Response  
Frequencies**



## 1.3 Πίνακες συνάφειας: για 2 μεταβλητές

Από το μενού → **Analyze  
Descriptive Statistics  
Crosstabs**



#### 1.4 Πίνακες συνάφειας: για 1 απλή μεταβλητή και 1 μεταβλητή με πολλαπλές απαντήσεις

Από το μενού → **Analyze**  
**Multiple Response**  
**Crosstabs**

