

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΔΠΜΣ

«ΕΠΙΣΤΗΜΕΣ ΤΗΣ ΑΓΩΓΗΣ ΜΕΣΩ ΚΑΙΝΟΤΟΜΩΝ ΤΕΧΝΟΛΟΓΙΩΝ
ΚΑΙ ΒΙΟΙΑΤΡΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ»

«Μεθοδολογία Εκπαιδευτικής Έρευνας και
Στατιστική »

Σημειώσεις Εργαστηρίου SPSS

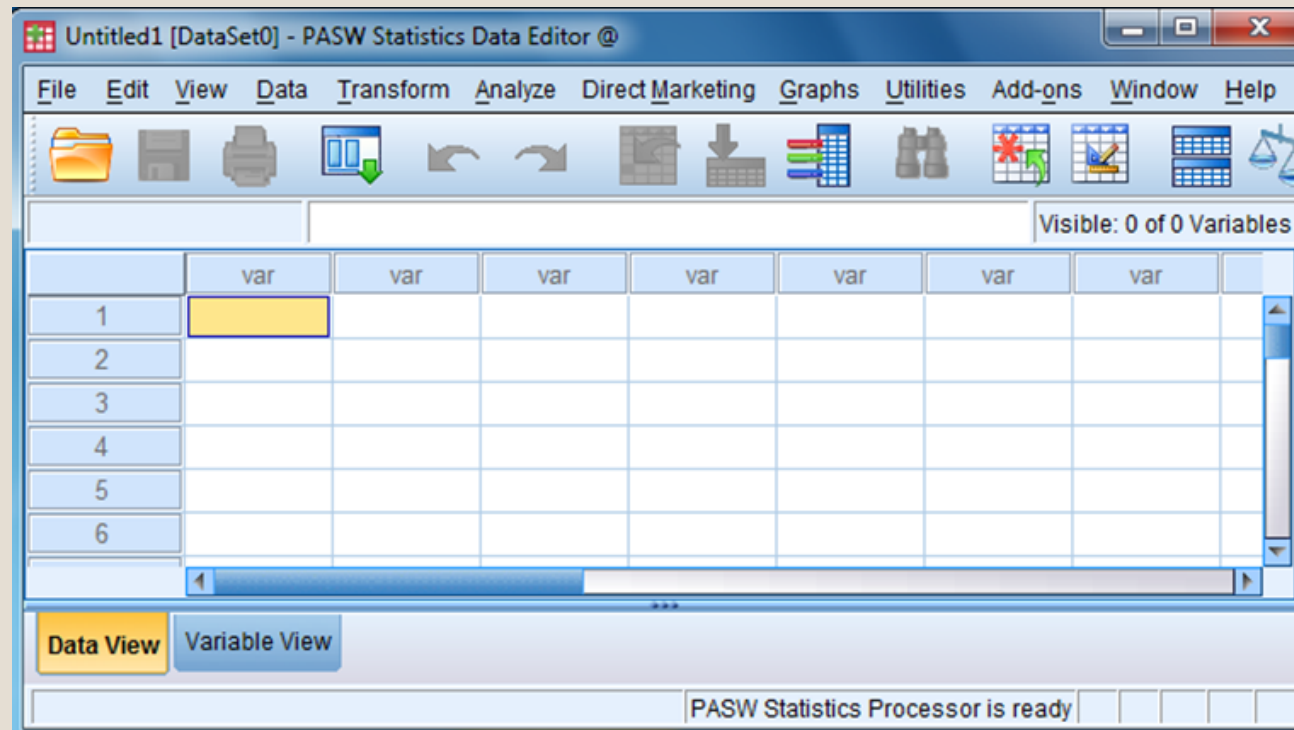
1^ο μέρος: Εισαγωγή στο SPSS, Περιγραφική Στατιστική

ΕΥΣΤΑΘΙΑ ΠΑΠΑΓΕΩΡΓΙΟΥ - ΠΑΝΑΓΙΩΤΑ ΛΑΛΟΥ

1.1 Το Περιβάλλον του SPSS

Ξεκινώντας το πρόγραμμα εμφανίζεται η παρακάτω εικόνα δηλαδή ένας editor παρόμοιος με λογιστικού φύλλου Excel προκειμένου να πραγματοποιηθεί νέα καταχώρηση δεδομένων.

Περιλαμβάνει τη γραμμή (menu) επιλογών, τη γραμμή συντομευμένων ενεργειών (εικονίδια) και δύο “καρτέλες” τη **Data View** και τη **Variable View**



- **Data View :** «Φύλλο δεδομένων»

Είναι ένα λογιστικό φύλλο δεδομένων, στο οποίο ο χρήστης καταχωρεί είτε πληκτρολογώντας είτε εισάγοντας απ' ευθείας από αρχείο άλλου λογισμικού τις τιμές των μεταβλητών τις οποίες θέλει να επεξεργαστεί. Κάθε στήλη αποτελεί ένα διαφορετικό χαρακτηριστικό (δηλαδή μία μεταβλητή). Κάθε γραμμή αντιστοιχεί στο ίδιο άτομο / αντικείμενο του δείγματος.

- **Variable View:** «Φύλλο ρυθμίσεων»

Στο φύλλο αυτό διαμορφώνουμε κατάλληλα τις ρυθμίσεις που αφορούν τις μεταβλητές που υπάρχουν Data View. Κάθε γραμμή του φύλλου, είναι οι ρυθμίσεις μιας μεταβλητής. (Π,χ η 1^η γραμμή είναι τα χαρακτηριστικά της 1^{ης} μεταβλητής δηλαδή της 1^{ης} στήλης του Data View, η 2^η γραμμή τα χαρακτηριστικά της 2^{ης} στήλης κλπ)

*Δεδομένα_στο_excel2.sav [DataSet1] - PASW Statistics Data Editor @

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

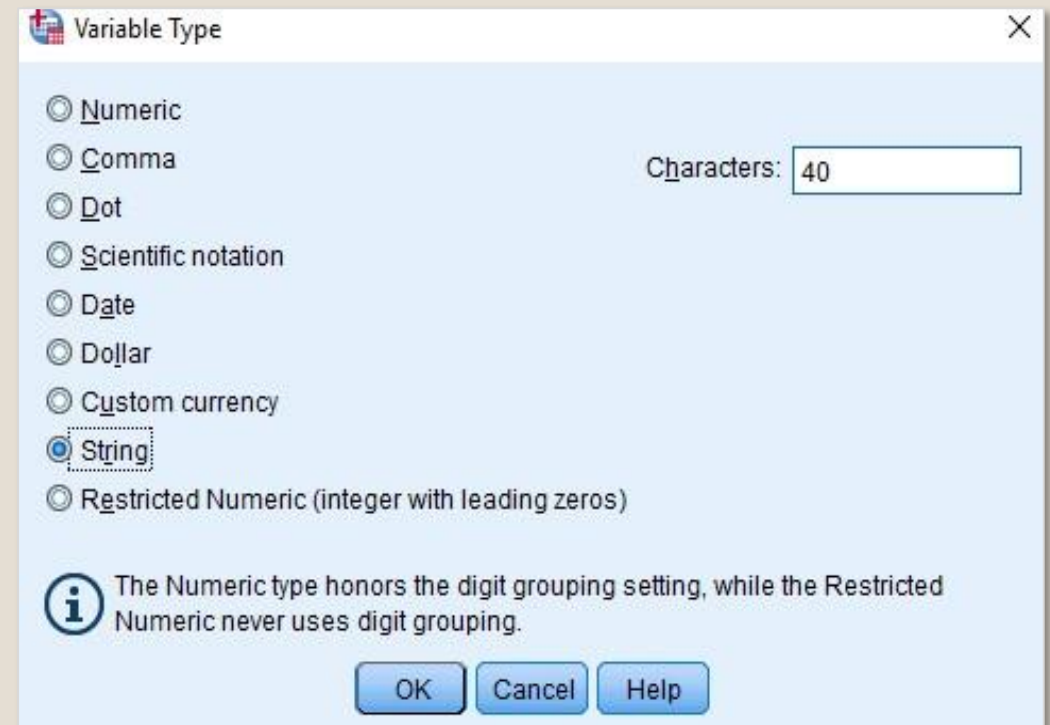
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	ID	String	7	0		None	None	7	Left	Nominal
2	sex	Numeric	11	0	Φύλο	{1, Άνδρας}...	None	11	Right	Nominal
3	age	Numeric	11	0	Ηλικία	None	None	11	Right	Scale
4	education	Numeric	11	0	Μορφωτικό επίπεδο	None	None	11	Right	Ordinal
5	Q1	Numeric	11	0	Ερώτηση 1 ερωτηματολογίου	None	None	11	Right	Ordinal
6	Q2	Numeric	11	0	Ερώτηση 2 ερωτηματολογίου	None	None	11	Right	Ordinal
7										
8										
9										
10										
11										
12										

Data View Variable View

PASW Statistics Processor is ready

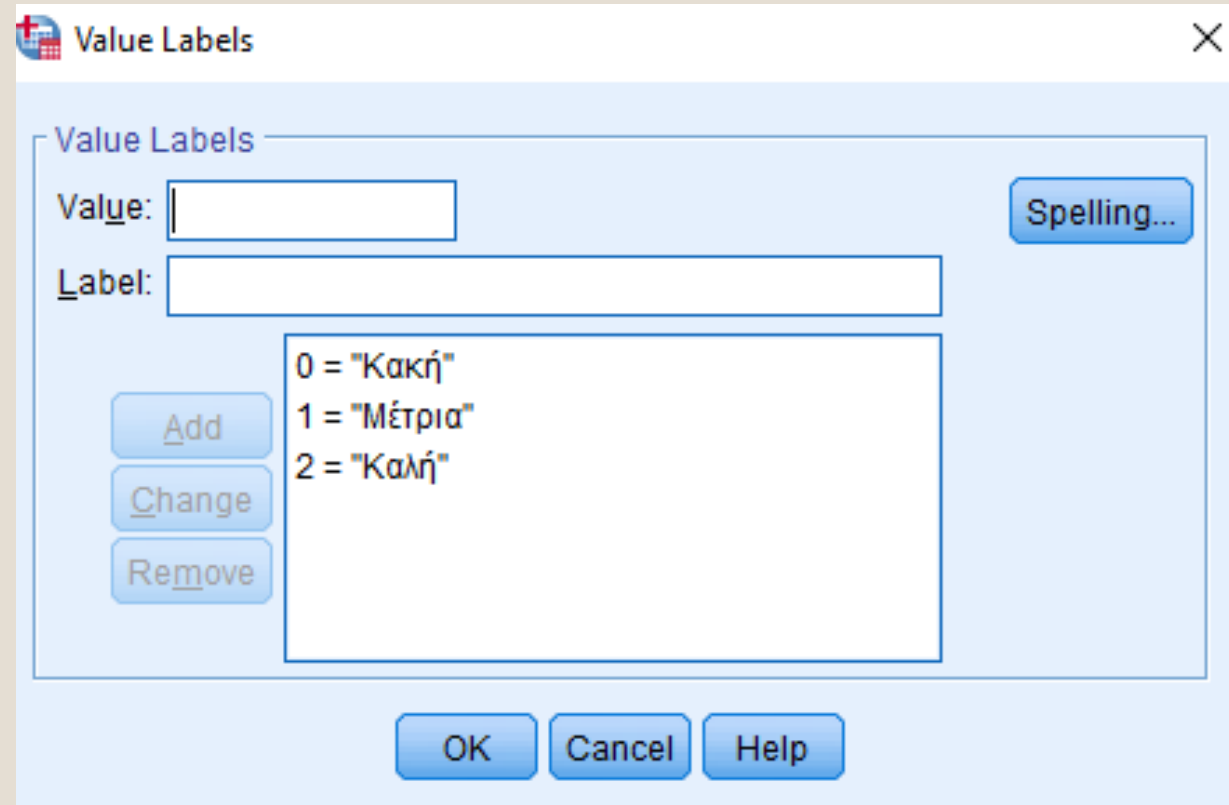
Οι ρυθμίσεις που διαμορφώνονται στο Variable View για κάθε μεταβλητή είναι:

- **Name:** Εδώ δίνεται ένα «κωδικό όνομα» στη μεταβλητή. Κάθε φορά που εισαγουμε τιμές μιας μεταβλητής, το SPSS δίνει ένα κωδικό όνομα στη μεταβλητή (π.χ. VAR00001). Επειδή η ονομασία που δίνει το SPSS δεν είναι αντιπροσωπευτική του χαρακτηριστικού, είναι προτιμότερο να αλλάζουμε το όνομα. Αν και η εφαρμογή υποστηρίζει ελληνικούς και λατινικούς χαρακτήρες, προτείνεται η χρήση λατινικών χαρακτήρων για να υπάρχει μελλοντική συμβατότητα του αρχείου δεδομένων με παλαιότερες εκδόσεις. Το όνομα δεν θα πρέπει να περιέχει κενά και σημεία στίξης.
- **Type:** Σε αυτή τη στήλη καθορίζεται ο τύπος δεδομένων της μεταβλητής. Η αρχική επιλογή είναι «Numeric», ενώ όταν η μεταβλητή δέχεται αλφαριθμητικά δεδομένα τότε καθορίζεται ως «String».



- **Width:** Καθορίζεται το μέγιστο πλήθος ψηφίων (χαρακτήρων) που θα εισαχθούν για τη τιμή της μεταβλητής.
- **Decimals:** Καθορίζεται ο αριθμός των δεκαδικών ψηφίων.
- **Label:** Ετικέτα της μεταβλητής, όπου περιγράφουμε (συνοπτικά συνήθως) το περιεχόμενο της. Εδώ, σε αντίθεση με το name μπορούμε να χρησιμοποιούμε κενά και σημεία στίξης

- **Values:** Εισάγουμε ετικέτα σε κάθε τιμή, της μεταβλητής. Έτσι μπορούμε, κυρίως στην περίπτωση ονομαστικών (Nominal) ή διατεταγμένων μεταβλητών (Ordinal) να εισάγουμε επεξηγήσεις για κάθε τιμή τους. Για παράδειγμα, για την μεταβλητή «κλινική εικόνα ασθενή» εάν αυτή έχει κωδικοποιηθεί ως «1-κακή», «2-μέτρια», και «3-καλή» τότε οι τιμές 1, 2 και 3 καταγράφονται στην θέση Value ενώ οι περιγραφές «Κακή», «Μέτρια» και «Καλή» στο πεδίο Value Label



- **Missing:** Σε αυτή τη στήλη καθορίζονται οι τιμές οι οποίες θα ληφθούν από την εφαρμογή ως ελλείπουσες παρατηρήσεις. Συνήθως οι καταχωρητές των δεδομένων στις περιπτώσεις απουσίας απαντήσεων καταγράφουν τιμές «99» ή «999» ή κάποια άλλη τιμή.
- **Columns:** Καθορίζεται το πλάτος της στήλης στην οποία καταχωρείται η μεταβλητή (Column Width)
- **Align:** καθορίζεται ο τρόπος στοίχισης των δεδομένων της μεταβλητής, με τις επιλογές Left, Right και Center.

Missing Values

No missing values

Discrete missing values

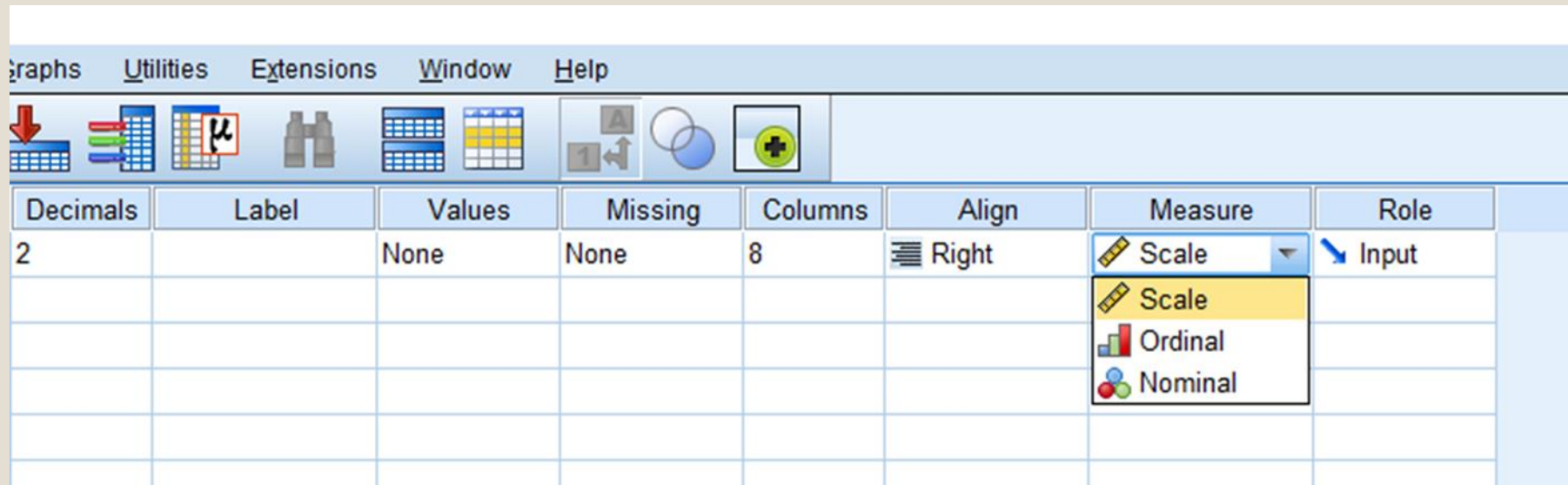
Range plus one optional discrete missing value

Low: High:

Discrete value:

OK Cancel Help

- **Measure** Καθορίζεται το είδος των μεταβλητών (όσον αφορά την στατιστική τους επεξεργασία), με τις επιλογές



Scale: Μεταβλητές κλίμακας ή ποσοτικές μεταβλητές. Πρόκειται για μεταβλητές που αντιστοιχούν σε ποσοτικά μεγέθη, τα οποία δηλαδή μπορούν να μετρηθούν όπως π.χ. ύψος, βάρος, ηλικία κ.τ.λ.

Ordinal: Διατάξιμες μεταβλητές. Πρόκειται για κατηγορικές μεταβλητές με τιμές οι οποίες ορίζουν μία σχέση διάταξης, όπως π.χ. ικανοποιημένος, αδιάφορος, δυσαρεστημένος. Μπορεί να είναι αλφαριθμητικές είτε αριθμητικές (συνήθως ακέραιες), με τιμές οι οποίες αντιστοιχούν σε συγκεκριμένες διατεταγμένες κατηγορίες.

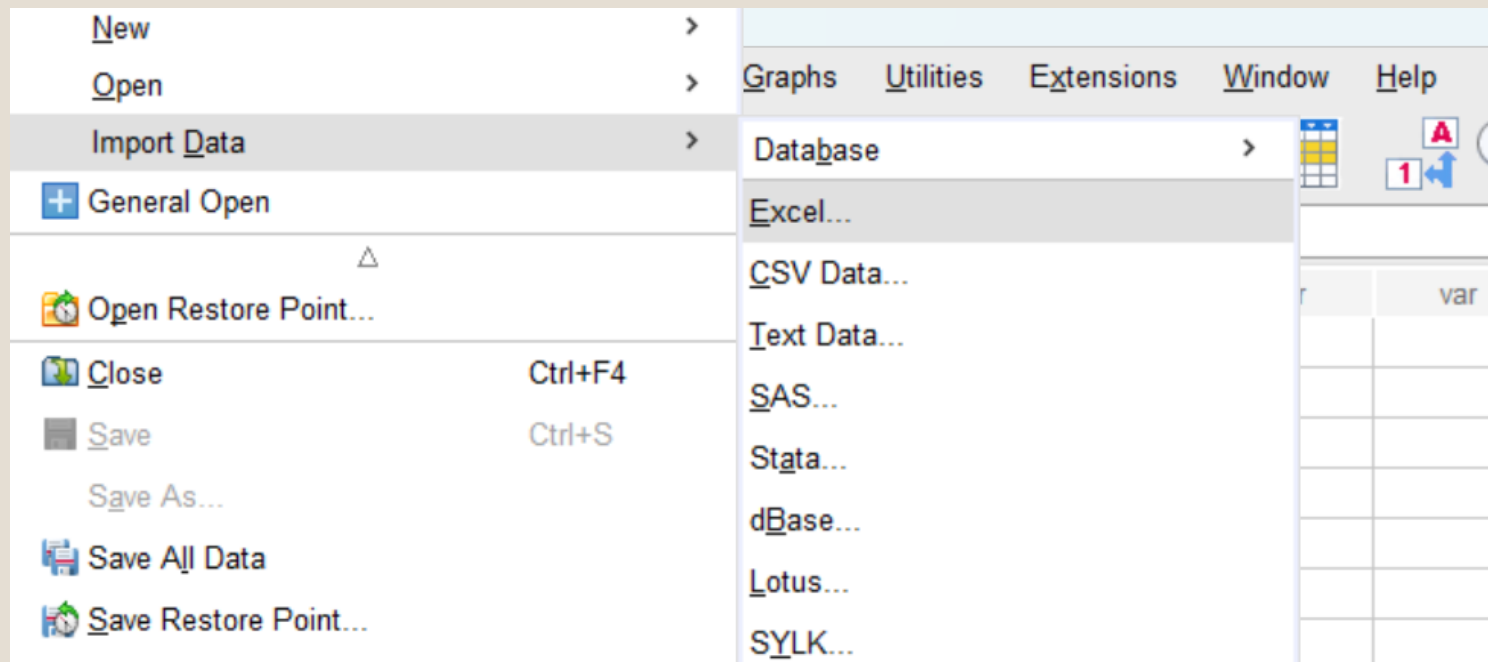
Nominal: Ονομαστικές μεταβλητές. Πρόκειται για κατηγορικές ποιοτικές μεταβλητές, δηλαδή μεταβλητές που αντιστοιχούν σε ποιοτικά μεγέθη, όπου όμως δεν υπάρχει κάποιου είδους διάταξη, όπως π.χ. ο τόπος γέννησης, το φύλο, κ.τ.λ. Μπορεί να είναι αλφαριθμητικές (π.χ. φύλο με τιμές Α ή Θ για άνδρες ή γυναίκες) είτε αριθμητικές, με τιμές οι οποίες αντιστοιχούν σε συγκεκριμένες κατηγορίες, (π.χ. φύλο με τιμές 1 για άνδρες και 0 για γυναίκες)

- **Role:** Ο ρόλος της μεταβλητής σε μία ανάλυση. Η ιδιότητα αυτή χρησιμοποιείται σε ορισμένες μόνο στατιστικές διαδικασίες.

1.2 Καταχώρηση δεδομένων

Όπως ήδη αναφέρθηκε, τα στατιστικά δεδομένα που πρόκειται να επεξεργαστούμε τα πληκτρολογούμε στο φύλλο Data View. Σε κάποιες περιπτώσεις τα δεδομένα είναι διαθέσιμα σε ηλεκτρονικό αρχείο. Αν τα δεδομένα μας είναι αποθηκευμένα σε αρχείο δεδομένων του Excel (*.xlsx) ακολουθούμε την παρακάτω διαδρομή:

File / Import Data / Excel



Άσκηση 1.1: Να καταχωρηθούν τα δεδομένα του παρακάτω πίνακα που αφορούν 20 ασθενείς ενός νοσοκομείου. Για τη καταχώρηση των ποιοτικών μεταβλητών να μπουν ετικέτες (values) στις τιμές:

- Φύλο: 1 → Άντρας , 2 → Γυναίκα
- Οικογενειακή κατάσταση:
1 → Άγαμος – η, 2 → Έγγαμος – η, 3 → Χήρος – α/Διαζευγμένος
- Κάπνισμα: 1 → Καπνιστής, 2 → Μη καπνιστής

Οικ. κατάσταση	Φύλο	Βάρος	Υψος	HDL	CHOL	Κάπνισμα
Άγαμος-η	Άνδρας	92	170	52	201	Καπνιστής
Άγαμος-η	Γυναίκα	78	175	47	230	Μη καπνιστής
Χήρος-α	Γυναίκα	110	168	35	278	Μη καπνιστής
Έγγαμος-η	Άνδρας	80	179	70	197	Μη καπνιστής
Διαζευγμένος-η	Άνδρας	95	180	46	193	Καπνιστής
Άγαμος-η	Γυναίκα	118	159	66	260	Μη καπνιστής
Έγγαμος-η	Γυναίκα	55	171	80	172	Καπνιστής
Έγγαμος-η	Γυναίκα	81	157	62	240	Καπνιστής
Άγαμος-η	Γυναίκα	60	164	54	167	Μη καπνιστής
Άγαμος-η	Άνδρας	99	170	57	234	Μη καπνιστής
Διαζευγμένος-η	Γυναίκα	125	167	49	304	Καπνιστής
Έγγαμος-η	Άνδρας	102	190	69	199	Καπνιστής
Διαζευγμένος-η	Γυναίκα	114	176	41	236	Καπνιστής
Διαζευγμένος-η	Άνδρας	120	170	38	225	Μη καπνιστής
Άγαμος-η	Άνδρας	82	170	89	204	Μη καπνιστής
Διαζευγμένος-η	Γυναίκα	56	166	68	157	Καπνιστής
Έγγαμος-η	Γυναίκα	92	172	51	214	Καπνιστής
Άγαμος-η	Γυναίκα	78	165	45	226	Μη καπνιστής
Έγγαμος-η	Άνδρας	104	175	43	278	Καπνιστής
Άγαμος-η	Άνδρας	80	180	70	254	Μη καπνιστής

Στη καρτέλα **Variable View** διαμορφώνουμε τις ρυθμίσεις για τις τιμές των μεταβλητών που θα καταχωρήσουμε . Θυμίζουμε ότι κάθε γραμμή αντιστοιχεί σε μία μεταβλητή.

Ξεκινάμε στη 1^η γραμμή με τις ρυθμίσεις που αφορούν στη μεταβλητή: Οικ. Κατάσταση.

Στο 1^ο κελί (**Name**) γράφουμε το κωδικό όνομα Status. (προσοχή το name πρέπει να είναι μία λέξη χωρίς κενά),

Πατώντας στο 2^ο κελί (**Type**) ανοίγει το παράθυρο όπου επιλέγουμε ως τύπο των δεδομένων που θα εισαχθούν **Numeric**. Επιλέγουμε Numeric , διότι στα κελιά θα εισάγουμε αριθμητικές τιμές : 1, 2 και 3 αντί για τις λέξεις άγαμος, έγγαμος κλπ.

Στο 3^ο κελί (**Width**) βάζουμε τη τιμή 1, αφού γνωρίζουμε ότι οι τιμές που θα καταχωρήσουμε είναι μονοψήφιοι. Ωστόσο μπορούμε να αφήσουμε την τιμή 8 που έχει ήδη υπάρχει.

Στο 4^ο κελί (**Decimals**) επιλέγουμε ως αριθμό δεκαδικών ψηφίων 0, αφού όλες οι τιμές που θα καταχωρηθούν είναι ακέραιες.

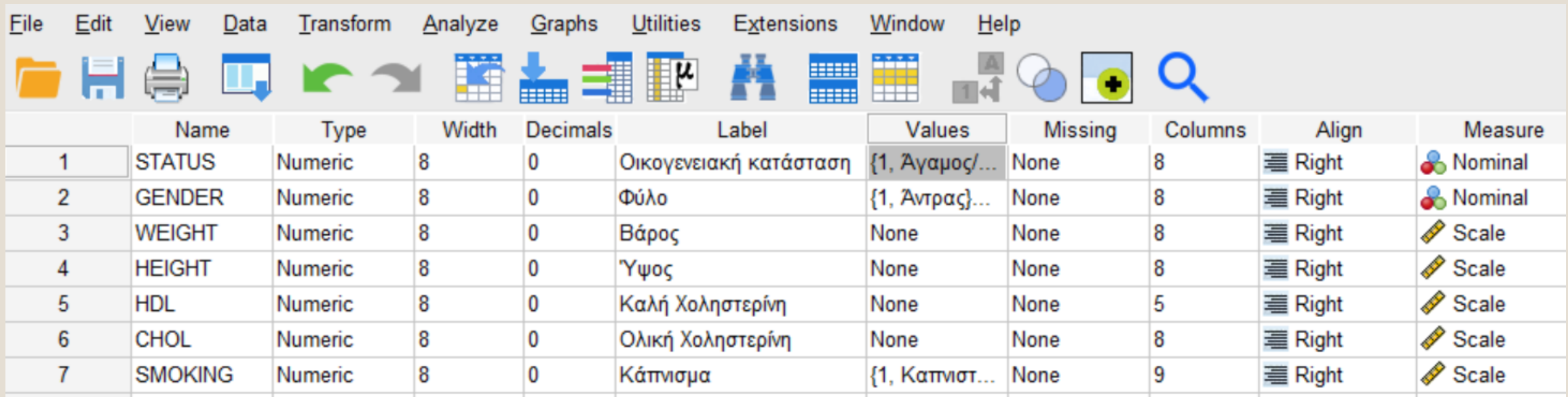
Στο 5^ο κελί (**Label**) γράφουμε την ετικέτα που θέλουμε να δώσουμε στη μεταβλητή:
«Οικογενειακή Κατάσταση»

Πατώντας στο 6^ο κελί (**Values**) ανοίγει παράθυρο στο οποίο δίνουμε ετικέτα για κάθε συγκεκριμένη τιμή της μεταβλητής. Γράφουμε στο Values :1 και στο Label: Άγαμος-η κλπ.

Στη ρύθμιση **Missing** δεν κάνουμε καμία αλλαγή διότι όλες η τιμές της μεταβλητής θέλουμε να συμπεριληφθούν στην ανάλυση.

Στη ρύθμιση **Measure** επιλέγουμε ως είδος μεταβλητής το **Nominal**. (Ονομαστική)

Με όμοιο τρόπο στις επόμενες γραμμές διαμορφώνουμε τις ρυθμίσεις και για τις υπόλοιπες μεταβλητές.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	STATUS	Numeric	8	0	Οικογενειακή κατάσταση	{1, Άγαμος/...	None	8	Right	Nominal
2	GENDER	Numeric	8	0	Φύλο	{1, Άντρας}...	None	8	Right	Nominal
3	WEIGHT	Numeric	8	0	Βάρος	None	None	8	Right	Scale
4	HEIGHT	Numeric	8	0	Ύψος	None	None	8	Right	Scale
5	HDL	Numeric	8	0	Καλή Χοληστερίνη	None	None	5	Right	Scale
6	CHOL	Numeric	8	0	Ολική Χοληστερίνη	None	None	8	Right	Scale
7	SMOKING	Numeric	8	0	Κάπνισμα	{1, Καπνιστ...	None	9	Right	Scale

Αφού τελειώσουμε με τις ρυθμίσεις όλων των μεταβλητών , γυρνάμε στη καρτέλα **Data View** και πληκτρολογούμε τις τιμές. Εκτός από τη πληκτρολόγηση, μπορούμε να χρησιμοποιήσουμε αντιγραφή / επικόλληση για να εισάγουμε τις τιμές.

1.3 Μενού Transform: Επανακωδικοποίηση μεταβλητής

Περιέχει εντολές για την δημιουργία νέων μεταβλητών ή το μετασχηματισμό υπαρχόντων, βασιζόμενες σε ορισμένες συνεχείς κατανομές ή σε μαθηματικές και λογικές σχέσεις υπαρχόντων δεδομένων. Από τις εντολές αυτές θα ασχοληθούμε εκτενέστερα με τις εξής:

1.3.1 Compute Variable

Στόχος είναι η δημιουργία νέων μεταβλητών υπό την εφαρμογή μιας μαθηματικής (αριθμητικής) έκφρασης. Αφου επιλέξουμε

Transform /Compute Variable

- Στο παράθυρο που ανοίγει, στο πεδίο *Target Variable* καταγράφεται η σύντομη ονομασία της νέας μεταβλητής, ενώ στο δεξιό μέρος *Numeric Expression* καταγράφεται η μαθηματική έκφραση (ο τύπος μετασχηματισμού) που θα οδηγήσει στη δημιουργία της νέας μεταβλητής. Αξίζει να σημειωθεί ότι παρέχεται πληθώρα μαθηματικών και στατιστικών συναρτήσεων στο πλαίσιο *Function Group*.

Άσκηση 1.2: Χρησιμοποιώντας τα δεδομένα της άσκησης 1.1, να δημιουργήσετε μια νέα μεταβλητή που θα φαίνεται ο αθηρωματικός δείκτης κάθε ατόμου. Ο αθηρωματικός δείκτης υπολογίζεται από το πηλίκο της ολικής χοληστερίνης προς τη καλή χοληστερίνη:

$$\frac{CHOL}{HDL}$$

Λύση:

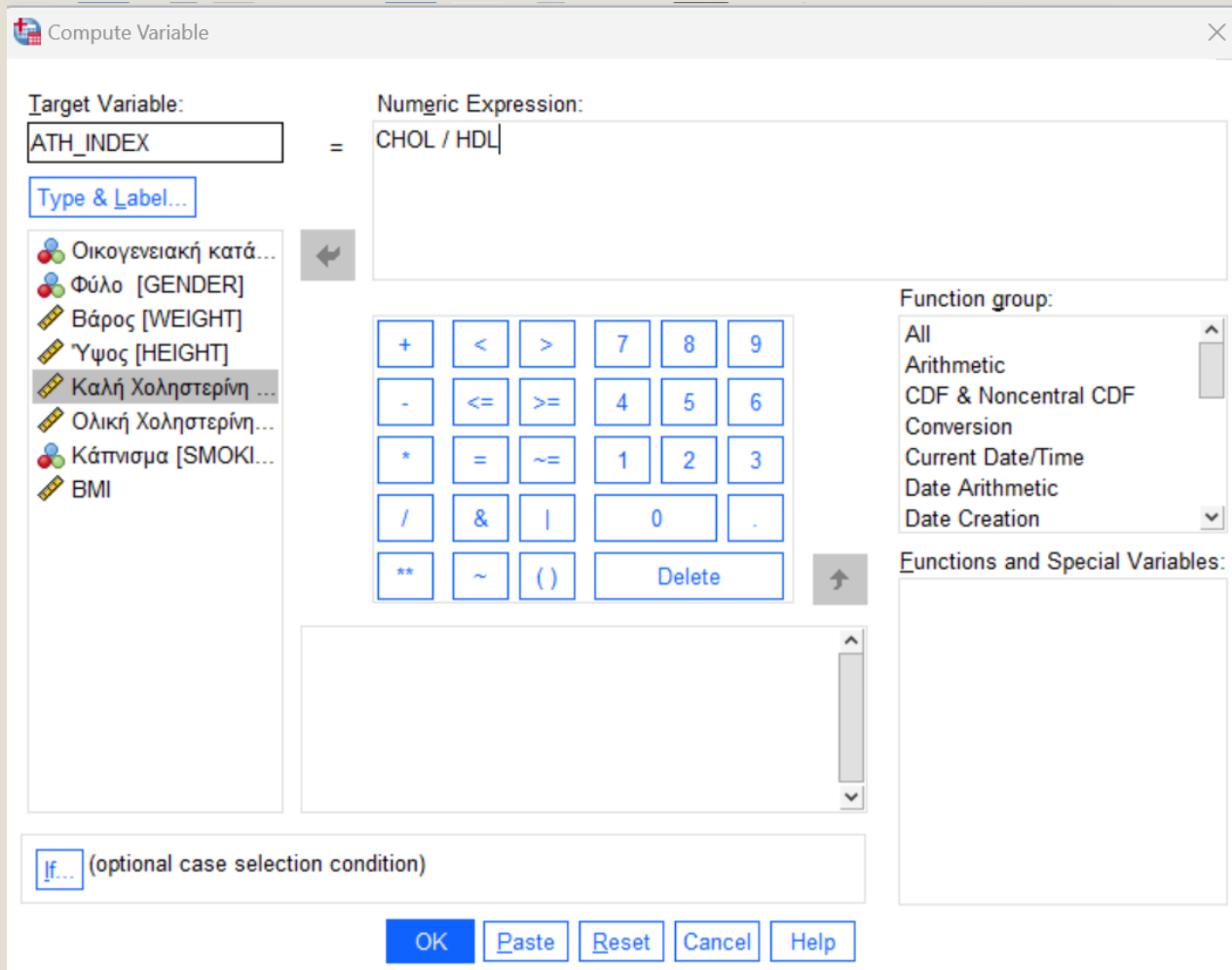
Επιλέγουμε : ***Transform /Compute Variable***

Στο παράθυρο που ανοίγει, στο πεδίο ***Target Variable*** γράφουμε το κωδικό όνομα: ATH_INDEX .

Και σε αυτή τη περίπτωση, ισχύουν αυτά που αναφέρθηκαν για τη ρύθμιση Name, δηλαδή ότι δεν μπορούμε στο όνομα να χρησιμοποιήσουμε σημεία στίξης και το κενό.

Στο πεδίο ***Numeric Expression*** θα καταγραφεί ο τύπος υπολογισμού του αθηρωματικού δείκτη. Στη λίστα με τις μεταβλητές που υπάρχει στο αριστερό μέρος, επιλέγουμε τη μεταβλητή ολική χοληστερίνη και με το μπλε βελάκι τη στέλνουμε στο ***Numeric Expression*** . Στη συνέχεια στο πληκτρολόγιο που υπάρχει από κάτω επιλέγουμε το πλήκτρο / και τέλος ξανά από τη λίστα μεταβλητών επιλέγουμε τη μεταβλητή Καλή χοληστερίνη και τη στέλνουμε στο ***Numeric Expression*** .

Πατάμε OK, οπότε στο Data View εμφανίζεται η νέα μεταβλητή ATH_INDEX.



CHOL	SMOKING	ATH_INDEX
201	Καπνιστής	3,87
230	Μη καπνιστής	4,89
278	Μη καπνιστής	7,94
197	Μη καπνιστής	2,81
193	Καπνιστής	4,20
260	Μη καπνιστής	3,94
172	Καπνιστής	2,15
240	Καπνιστής	3,87
167	Μη καπνιστής	3,09
234	Μη καπνιστής	4,11
304	Καπνιστής	6,20
199	Καπνιστής	2,88
236	Καπνιστής	5,76
225	Μη καπνιστής	5,92
204	Μη καπνιστής	2,29
157	Καπνιστής	2,31
214	Καπνιστής	4,20
226	Μη καπνιστής	5,02
278	Καπνιστής	6,47
254	Μη καπνιστής	3,63

Άσκηση 1.3: Χρησιμοποιώντας τα δεδομένα της άσκησης 1.1, να δημιουργήσετε μια νέα μεταβλητή που θα φαίνεται ο Δείκτης Μάζας Σώματος κάθε ατόμου. Ο ΔΜΣ υπολογίζεται από τη σχέση:

$$\Delta\text{ΜΣ} = \frac{\text{Βάρος}}{\text{Υψος}^2}$$

(Το Βάρος μετρημένο σε κιλά και το ύψος σε μέτρα)

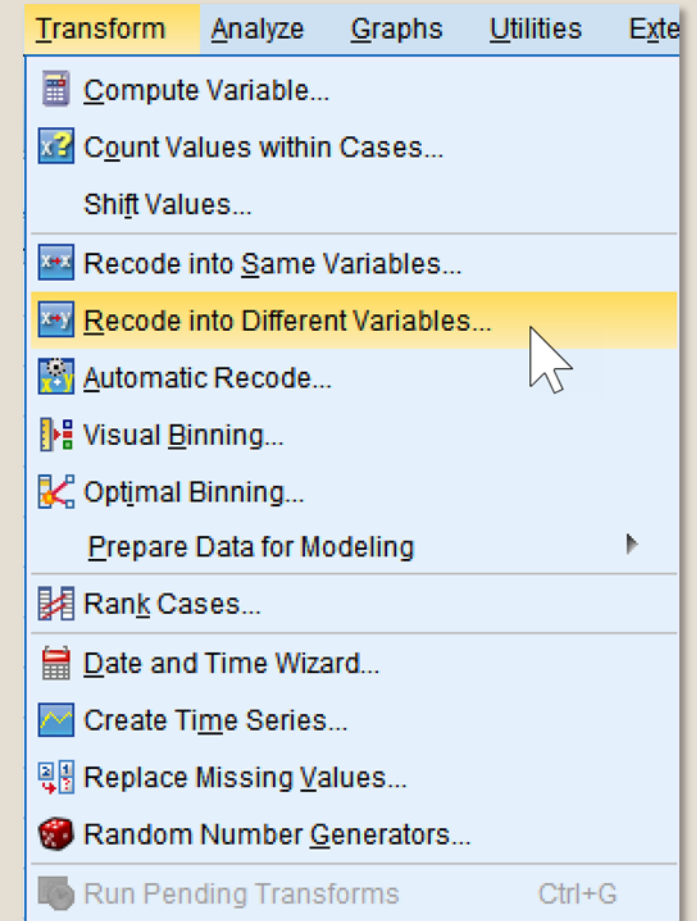
1.3.2 Εντολή Recode

Η εντολή χωρίζεται σε δύο κατηγορίες:

Transform > Recode Into Same Variable - Για την εφαρμογή της επανακωδικοποίησης των τιμών μιας μεταβλητής εντός της ίδιας μεταβλητής.

Transform > Recode Into Different Variable - Για την εφαρμογή της επανακωδικοποίησης των τιμών μιας μεταβλητής σε μια νέα μεταβλητή, διατηρώντας την αρχική μεταβλητή αμετάβλητη.

Από τις δύο κατηγορίες, θα ασχοληθούμε με τη δεύτερη, τη δημιουργία τιμών σε μια νέα μεταβλητή, Συνήθως προτιμούμε τη περίπτωση αυτή, διότι αποφεύγουμε τον κίνδυνο να χάσουμε την αρχική μεταβλητή στη περίπτωση που κάνουμε λάθος.



Άσκηση 1.4: (Συνέχεια της άσκησης 1.2)

Να δημιουργηθεί μια νέα ποιοτική μεταβλητή, στην οποία θα αξιολογείται ο καρδιαγγειακός κίνδυνος βάσει της τιμής του αθηρωματικού δείκτη. Οι τιμές της νέας μεταβλητής ανάλογα με τις τιμές του αθηρωματικού δείκτη φαίνονται στον παρακάτω πίνακα:

Τιμές Αθηρωματικού Δείκτη	Καρδιαγγειακός κίνδυνος
<3,5	Μικρός κίνδυνος
3,5 - 5	Μέτριος κίνδυνος
>5	Μεγάλος κίνδυνος

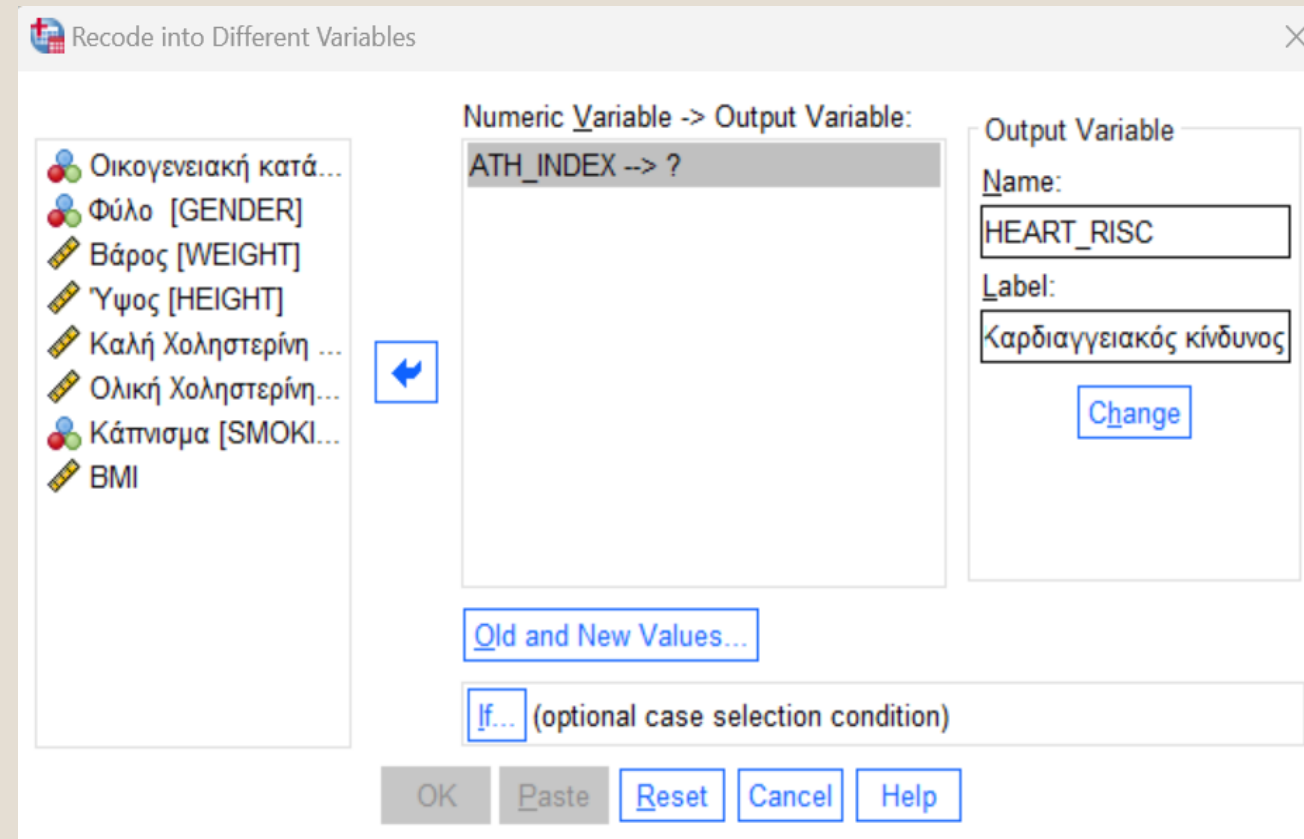
Λύση:

Ακολουθούμε τη διαδρομή: **Transform / Recode into Different Variables**

Στο παράθυρο που ανοίγει, επιλέγουμε από τη λίστα των μεταβλητών αριστερά, τη μεταβλητή που θα χρησιμοποιήσουμε (στη περίπτωση μας τον Αθηρωματικό Δελικτη) για να φτιάξουμε τη νέα μεταβλητή, και τη στέλνουμε με το μπλε βελάκι στο **Numeric Variable**

Στο **Output Variable** γράφουμε το όνομα της νέας μεταβλητής (HEART_RISC) και προαιρετικά την ετικέτα της (Επικινδυνότητα καρδιακού επεισοδίου).

Στη συνέχεια πατάμε το **Old and new Values**



Ο ορισμός των νέων τιμών πραγματοποιείται στο υποπλαίσιο διαλόγου *Old & New Values*.

Old Values

Στο πεδίο αυτό δηλώνονται οι αρχικές τιμές, ή το εύρος τιμών, της υπάρχουσας μεταβλητής που θα επανακωδικοποιηθούν.

New Values

Στο πεδίο αυτό ορίζονται οι τιμές της νέας μεταβλητής. Εάν οι τιμές της νέας μεταβλητής είναι αλφαριθμητικές θα πρέπει πριν την εισαγωγή των τιμών να ενεργοποιήσουμε βάζοντας “τικ” το **Output variables are strings**.

Θα πρέπει επίσης τότε να καθορίσουμε στο **Width** τον μέγιστο αριθμό χαρακτήρων των τιμών.

Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:

3,5

through

5

Range, LOWEST through value:

Range, value through HIGHEST:

All other values

New Value

Value: Μέτριος

System-missing

Copy old value(s)

Old -> New:

Lowest thru 3,49 -> 'Μικρός'

Add

Change

Remove

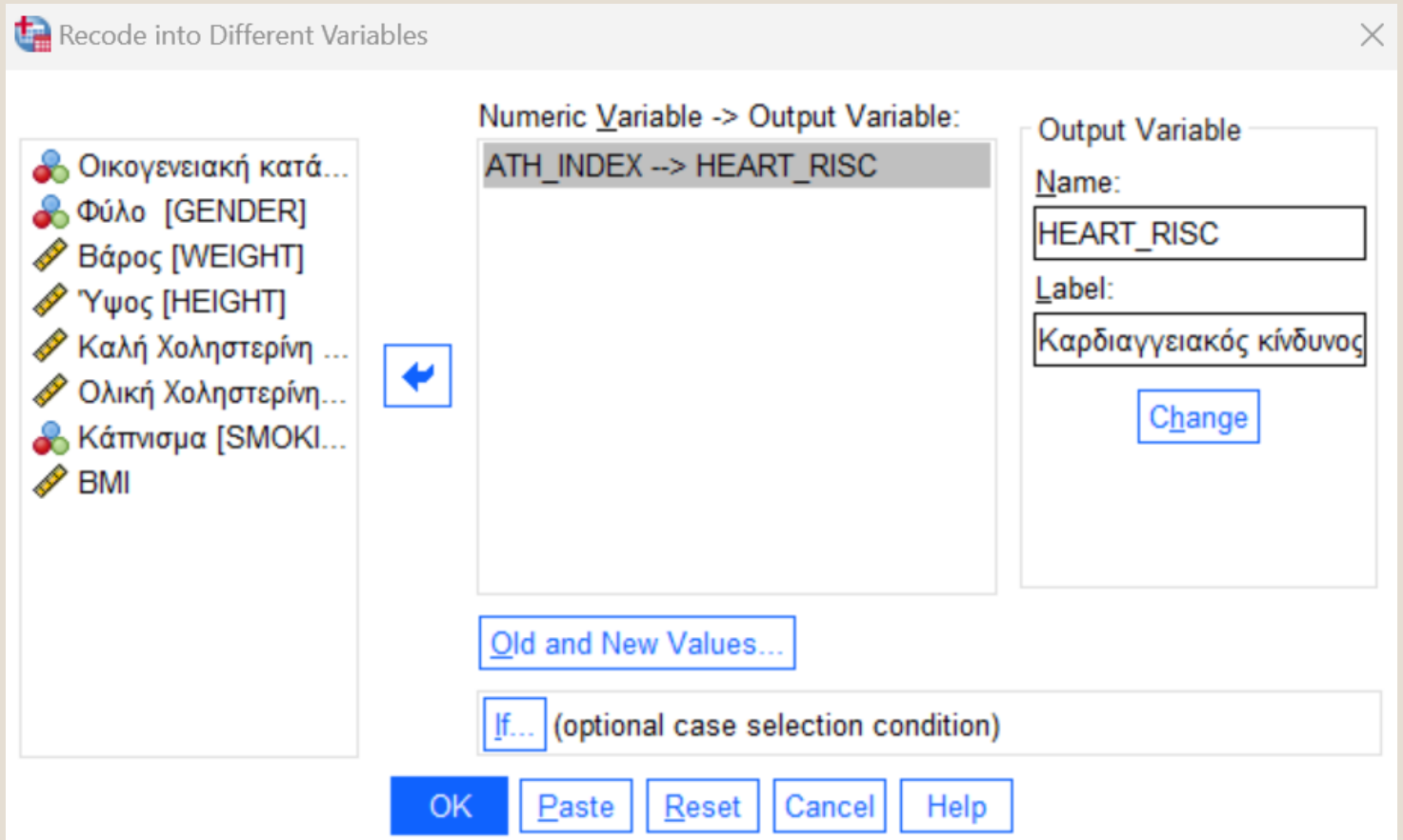
Output variables are strings **Width:** 30

Convert numeric strings to numbers ('5'->5)

Continue Cancel Help

Κάθε φορά που προσθέτουμε μια τιμή πατάμε **Add**, και όταν ολοκληρώσουμε την καταχώρηση τιμών πατάμε **Continue**

Στο παράθυρο που μεταφερόμαστε, πατάμε **Change** και τέλος **OK**.



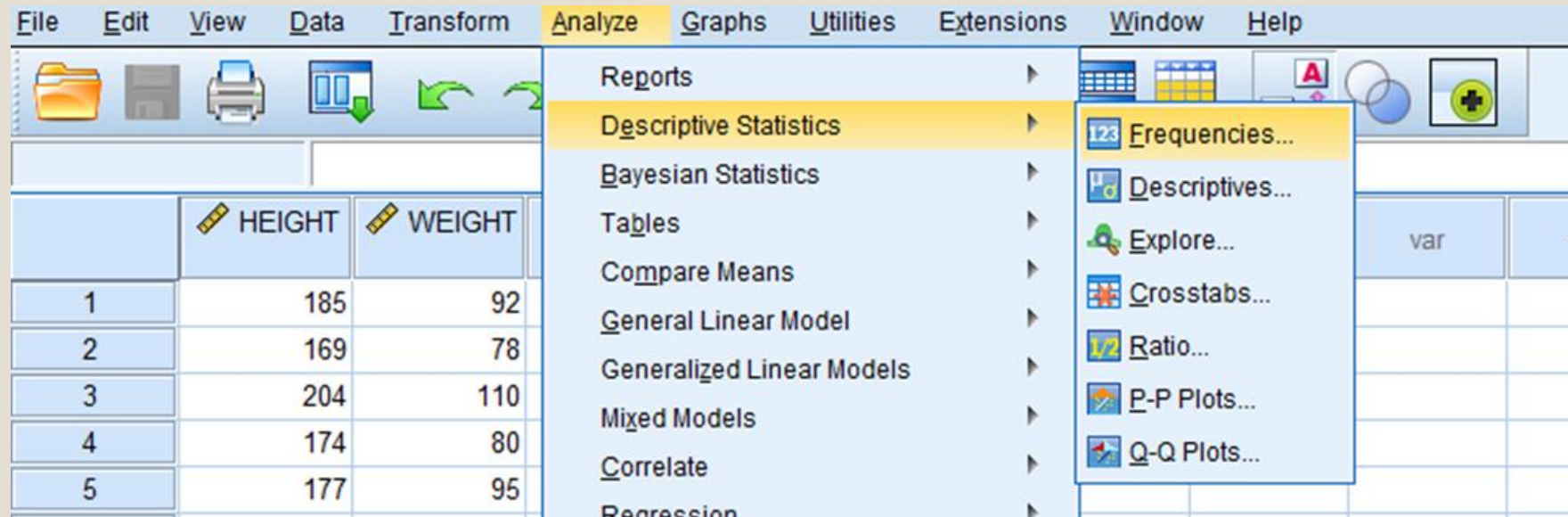
Άσκηση 1.5: Να δημιουργηθεί μια νέα μεταβλητή με τον χαρακτηρισμό του σωματικού βάρους του κάθε ατόμου ανάλογα με την τιμή του Δείκτη Μάζας Σώματος του. Οι κατηγορίες σωματικού βάρους φαίνονται στο παρακάτω πίνακα:

Δ.Μ.Σ	ΚΑΤΗΓΟΡΙΑ
<18,5	Ελλιποβαρής
18,5 - 25	Φυσιολογικός
25 – 30	Υπέρβαρος
>30	Παχύσαρκος

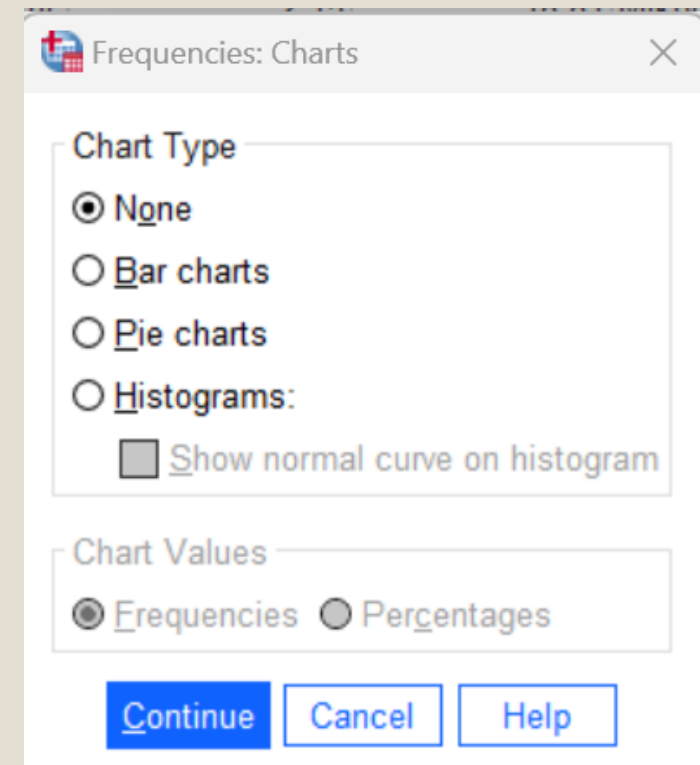
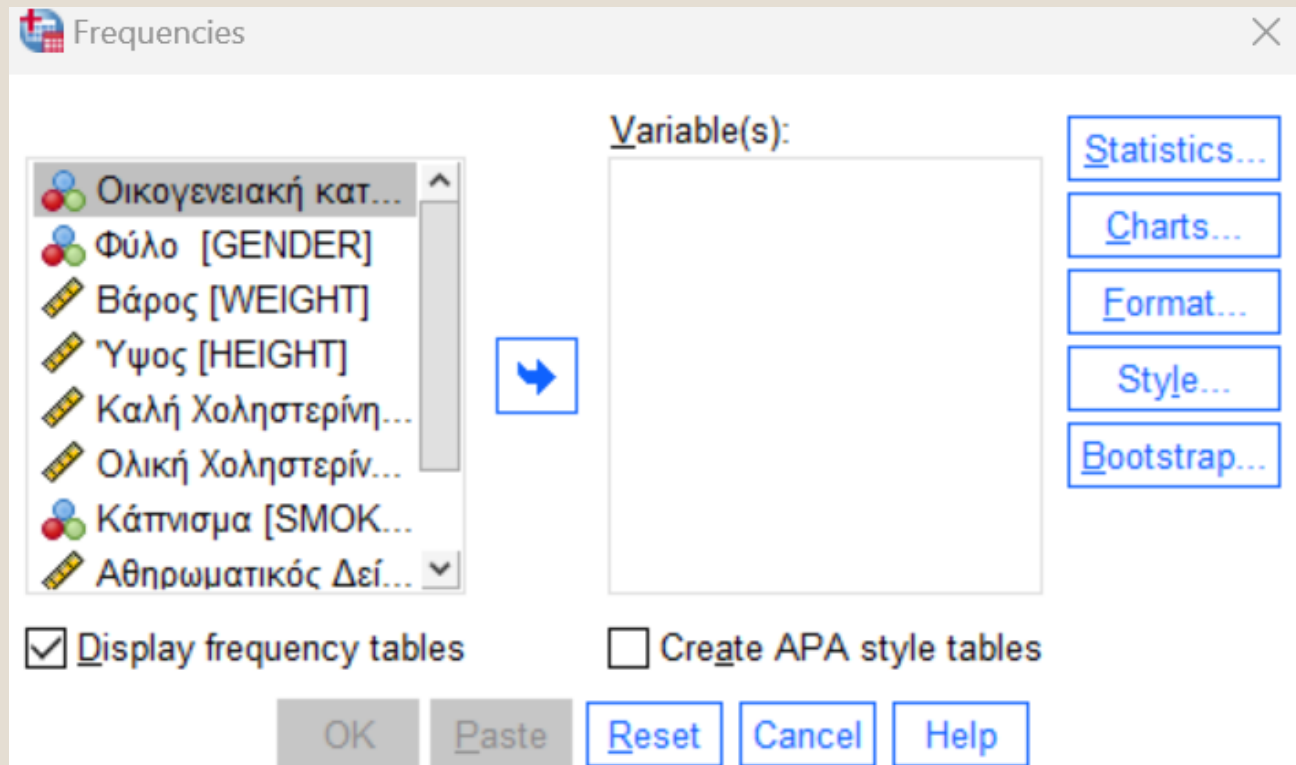
1.4 Πίνακες Συχνοτήτων / Γραφήματα

Για την κατασκευή πίνακα συχνοτήτων ακολουθούμε τη διαδρομή:

Analyze / Descriptive Statistics / Frequencies



- Στο πλαίσιο **Variable(s)** τοποθετούμε τις μεταβλητές για τις οποίες θα εξάγουμε πίνακες (κατανομές) συχνοτήτων. Επίσης από την υποενότητα **Charts** μπορούμε να δημιουργήσουμε κυκλικά διαγράμματα **Pie Chart**, ιστογράμματα **Histograms** ή ραβογράμματα **Bar Chart**.

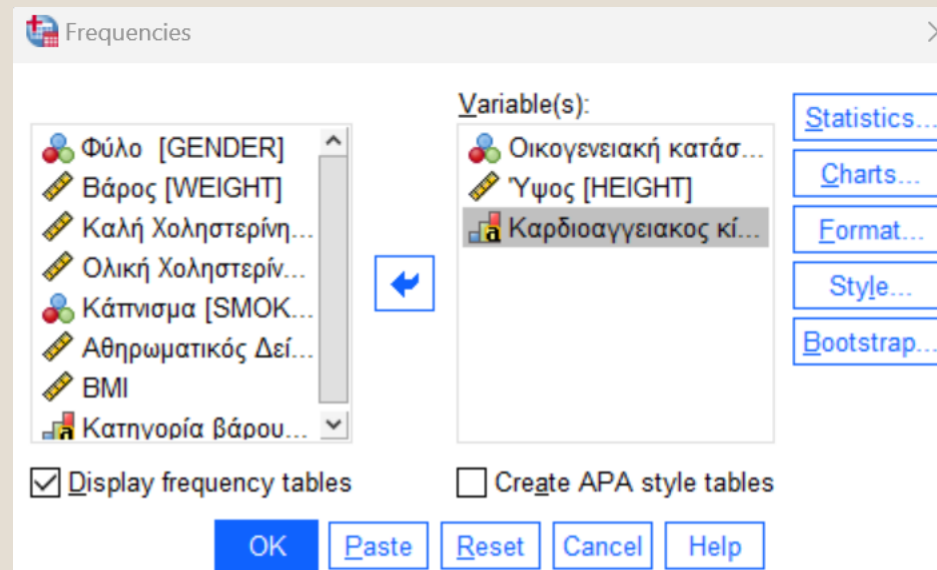


Άσκηση 1.6: Να κατασκευαστεί πίνακας κατανομής συχνοτήτων για τις μεταβλητές οικογενειακή κατάσταση, ύψος και καρδιαγγειακός κίνδυνος της άσκησης 1.4. Στη συνέχεια να δημιουργηθούν τα διαγράμματα : ραβδόγραμμα για την οικογενειακή κατάσταση, ιστόγραμμα για το ύψος και κυκλικό διάγραμμα για τον καρδιαγγειακό κίνδυνο.

Λύση:

Ακολουθούμε τη διαδρομή: **Analyze / Descriptive Statistics / Frequencies.**

Στο παράθυρο που ανοίγει, μεταφέρουμε με το μπλε βελάκι τις 3 μεταβλητές στο πλαίσιο Variable(s) και πατάμε OK.



Στο Output εμφανίζονται οι πίνακες κατανομής συχνοτήτων των τριών μεταβλητών. Στη διπλανή εικόνα βλέπουμε τον πίνακα της μεταβλητής 'ύψος'. Σε αυτόν η στήλη *Frequency* δείχνει τις συχνότητες των τιμών (n_i). Οι στήλες *Percent* και *Valid Percent* δείχνουν τη σχετική συχνότητα (f_i) δηλαδή το ποσοστό της κάθε τιμής. Η διαφορά τους είναι ότι στη στήλη *Percent* φαίνεται το ποσοστό κάθε τιμής επί το σύνολο όλων των τιμών, ενώ στο *Valid Percent* το ποσοστό κάθε τιμής επί των τιμών που βρίσκονται σε ισχύ, εξαιρούνται δηλαδή αυτές που είναι ορισμένες ως *Missing*. Η στήλη *Cumulative Percent* δείχνει τις σχετικές αθροιστικές συχνότητες (F_i).

Στη συνέχεια, για το σχεδιασμό των ζητούμενων γραφικών, ακολουθούμε πάλι την ίδια διαδρομή:

Analyze / Descriptive Statistics / Frequencies

Στο πλαίσιο *Variables* στέλνουμε μία μόνο μεταβλητή κάθε φορά. (Γιατί θέλουμε διαφορετικά διαγράμματα για τη κάθε μια)

		'Υψος			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	157	1	5,0	5,0	5,0
	159	1	5,0	5,0	10,0
	164	1	5,0	5,0	15,0
	165	1	5,0	5,0	20,0
	166	1	5,0	5,0	25,0
	167	1	5,0	5,0	30,0
	168	1	5,0	5,0	35,0
	170	4	20,0	20,0	55,0
	171	1	5,0	5,0	60,0
	172	1	5,0	5,0	65,0
	175	2	10,0	10,0	75,0
	176	1	5,0	5,0	80,0
	179	1	5,0	5,0	85,0
	180	2	10,0	10,0	95,0
190	1	5,0	5,0	100,0	
Total		20	100,0	100,0	

Αφού μεταφέρουμε τη μεταβλητή «Οικογενειακή κατάσταση στο πλαίσιο *Variables(s)* κλικάρουμε το κουμπί **Charts** και στο επόμενο παράθυρο κλικάρουμε το **Bar charts**. Στη συνέχεια **Continue** και τέλος **OK**. Στο Output, εκτός από το Ραβδόγραμμα κατασκευάζεται ξανά και ο πίνακας συχνοτήτων. Εάν επιθυμούμε να σχεδιαστεί το ραβδόγραμμα μόνο, χωρίς πίνακα, θα πρέπει να αφαιρέσουμε το «✓» από το *Display frequency tables*.

Variable(s):
Οικογενειακή κατάσ...

Display frequency tables Create APA style tables

OK Paste Reset Cancel Help

Statistics...
Charts...
Format...

Frequencies: Charts

Chart Type

None

Bar charts

Pie charts

Histograms:

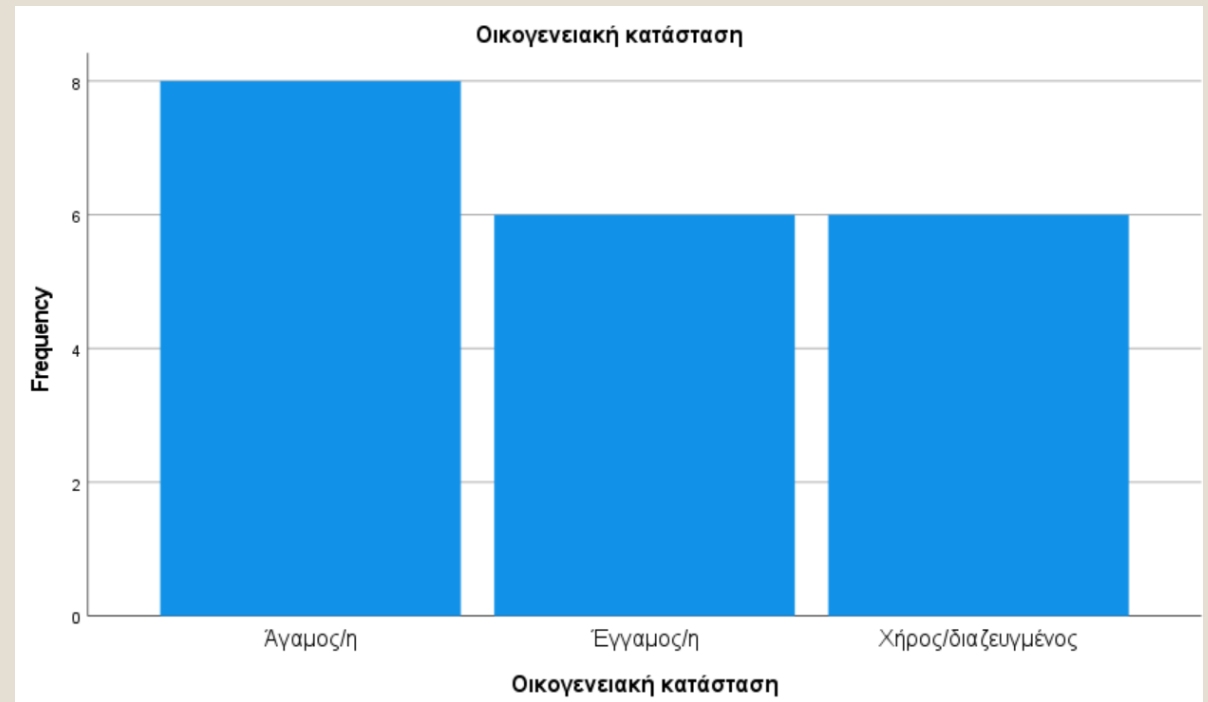
Show normal curve on histogram

Chart Values

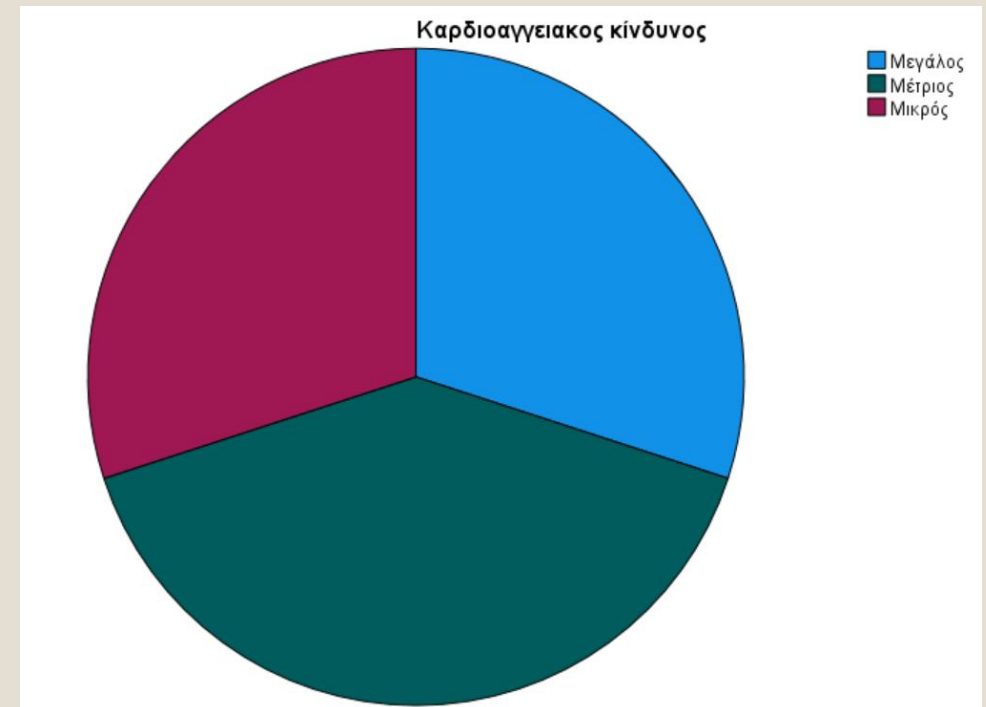
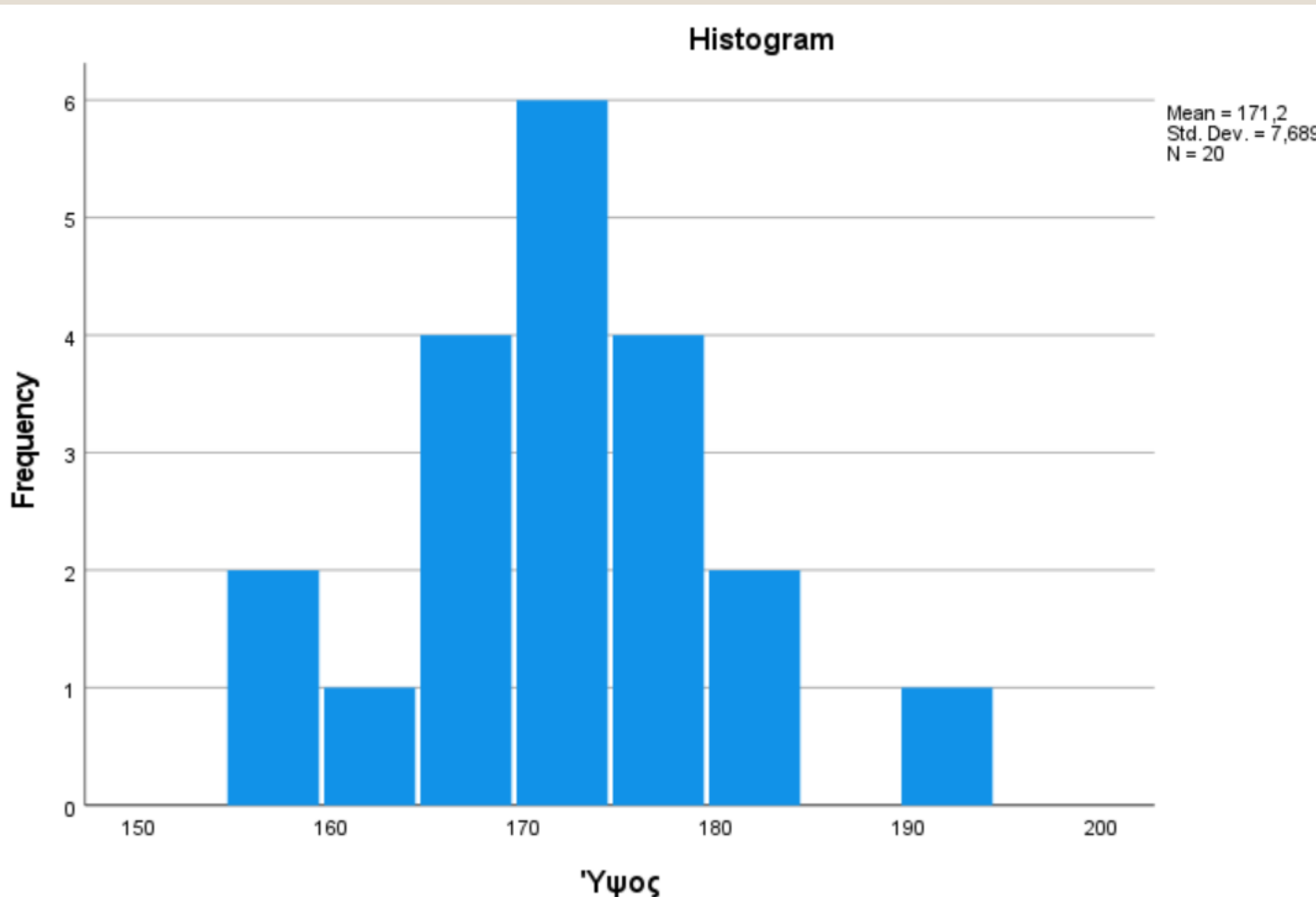
Frequencies Percentages

Continue Cancel Help

41	230	Καπνιστής	5,10
38	225	Μη καπνιστής	5,92
89	204	Μη καπνιστής	2,29
68	157	Καπνιστής	2,31
51	214	Καπνιστής	4,20
45	226	Μη καπνιστής	5,02



Ομοίως, για το ιστόγραμμα της μεταβλητής «ύψος» επιλέγουμε στο παράθυρο του *charts* το **Histograms** ενώ για το κυκλικό διάγραμμα της μεταβλητής «καρδιαγγειακός κίνδυνος» επιλέγουμε **Pie charts**.



1.5 Πίνακες συνάφειας

Ο πίνακας συνάφειας ή αλλιώς πίνακας διπλής εισόδου δίνει μια πρώτη εικόνα για την ύπαρξη σχέσης μεταξύ δύο ποιοτικών μεταβλητών. Ο πίνακας συνάφειας στις γραμμές του περιέχει τις τιμές της μιας ποιοτικής μεταβλητής και στις στήλες του τις τιμές της άλλης, ενώ στο εσωτερικό του, κάθε κελί περιέχει τις συχνότητες εμφάνισης ταυτόχρονα των τιμών των δύο μεταβλητών που αντιστοιχούν σε αυτό το κελί

Άσκηση 1.7: Να γίνει ο πίνακας συνάφειας των μεταβλητών Οικογενειακή κατάσταση και Κάπνισμα της άσκησης 1.4

Λύση:

Ακολουθούμε τη διαδρομή:

Analyze / Descriptive Statistics / Crosstabs

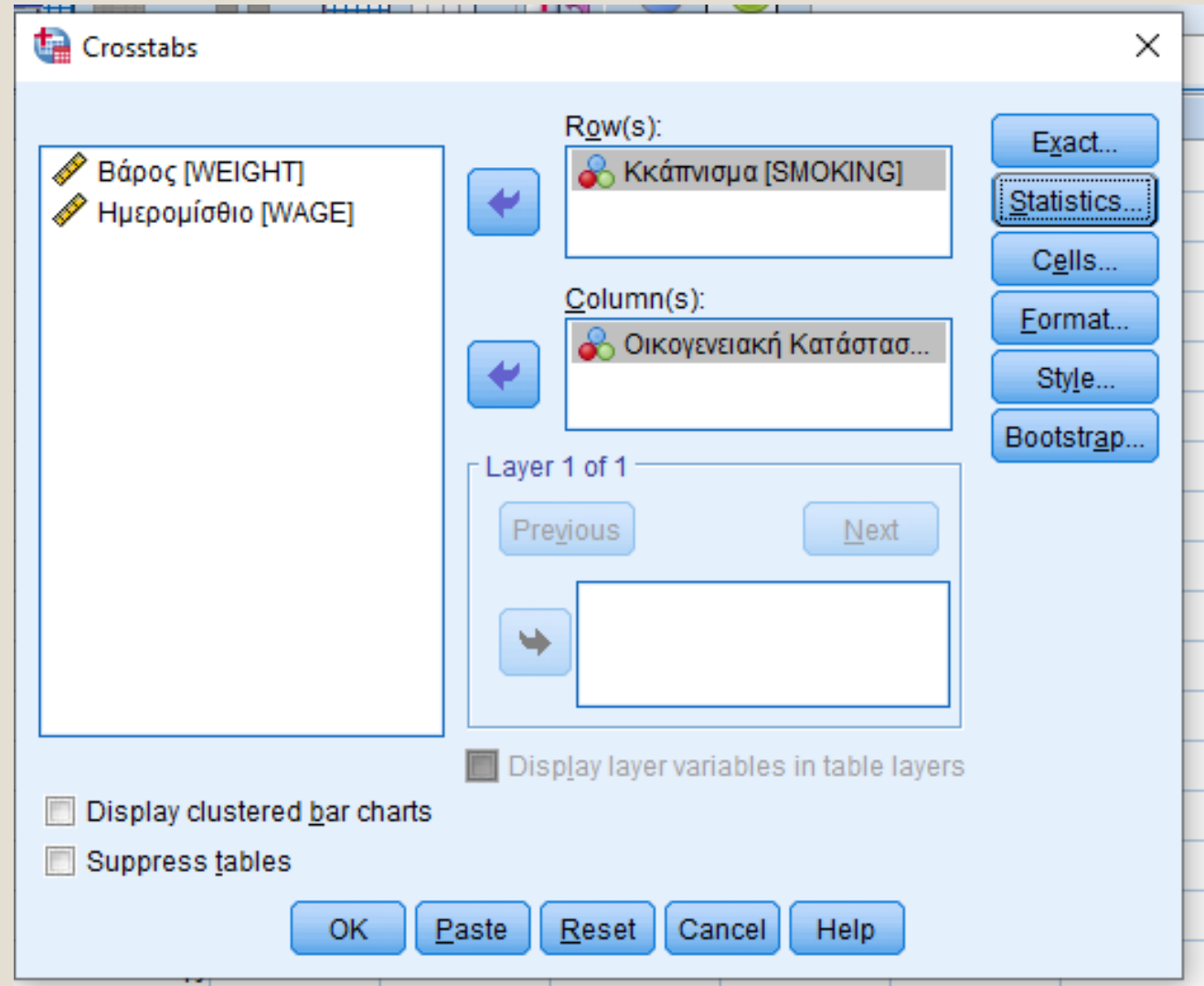
*οικογ_κατασταση.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

123 Frequencies...
 Descriptives...
 Explore...
 Crosstabs...
 Ratio...
 P-P Plots...
 Q-Q Plots...

	STATUS	WEIGHT
1	Αγαμος-η	92.00
2	Αγαμος-η	78.00
3	Χήρος-α	110.00
4	Έγγαμος-η	80.00
5	Διαζευγμέν...	95.00
6	Αγαμος-η	118.00
7	Έγγαμος-η	55.00
8	Έγγαμος-η	81.00
9	Αγαμος-η	60.00
10	Αγαμος-η	99.00
11	Διαζευγμέν...	125.00
12	Έγγαμος-η	102.00
13	Διαζευγμέν...	124.00
14	Διαζευγμέν...	120.00

Στο παράθυρο που ανοίγει μεταφέρουμε τη μία από τις δύο μεταβλητές στο Row(s) και την άλλη στο Column(s). (Ανάλογα με το πως θέλουμε να εμφανίζονται στον πίνακα) και πατάμε OK



- Εάν θέσουμε στο Rows το Κάπνισμα και στο Columns την Οικογενειακή κατάσταση προκύπτει ο πίνακας:

Κάπνισμα * Οικογενειακή κατάσταση Crosstabulation

Count

		Οικογενειακή κατάσταση			Total
		Άγαμος/η	Έγγαμος/η	Χήρος/διαζευγμένος	
Κάπνισμα	Καπνιστής	1	5	4	10
	Μη καπνιστής	7	1	2	10
Total		8	6	6	20

- Εάν θέλουμε στα κελιά εκτός από τις απόλυτες συχνότητες να εμφανίζονται και τα ποσοστά – σχετικές συχνότητες, στο παράθυρο διαλόγου πατάμε το κουμπί **Cells** και στο **Percentages** κλικάρουμε το **Total**.

Crosstabs

Row(s): Κάπνισμα [SMOKING]

Column(s): Οικογενειακή κατάσταση [...]

Layer 1 of 1

Display layer variables in table layers

Display clustered bar charts

Suppress tables

OK Paste Reset Cancel Help

Crosstabs: Cell Display

Counts

Observed

Expected

Hide small counts

Less than 5

z-test

Compare column proportions

Adjust p-values (Bonferroni method)

Percentages

Row

Column

Total

Create APA style table

Residuals

Unstandardized

Standardized

Adjusted standardized

Noninteger Weights

Round cell counts

Round case weights

Truncate cell counts

Truncate case weights

No adjustments

Continue Cancel Help

Γυναίκα	78	165	45	226	Mr. καπνιστής
---------	----	-----	----	-----	---------------

➔ **Crosstabs**

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Κάπνισμα * Οικογενειακή κατάσταση	20	100,0%	0	0,0%	20	100,0%

Κάπνισμα * Οικογενειακή κατάσταση Crosstabulation

		Οικογενειακή κατάσταση			Total	
		Άγαμος/η	Έγγαμος/η	Χήρος/διαζευγμένος		
Κάπνισμα	Καπνιστής	Count	1	5	4	10
	% of Total	5,0%	25,0%	20,0%	50,0%	
	Μη καπνιστής	Count	7	1	2	10
	% of Total	35,0%	5,0%	10,0%	50,0%	
Total	Count	8	6	6	20	
	% of Total	40,0%	30,0%	30,0%	100,0%	

1.6 Μέτρα Θέσης και Διασποράς

Άσκηση 1.8: Στη μεταβλητές «Βάρος» και «Αθηρωματικός δείκτης» της άσκησης 1.4, να βρεθεί η διάμεσος, η μέση τιμή, η επικρατούσα τιμή, το 1ο και 3ο τεταρτημόριο, το 60% και 85% ποσοστημόριο, το εύρος, η διακύμανση και η τυπική απόκλιση.

Λύση:

Ακολουθούμε τη διαδρομή:

Analyze / Descriptive Statistics / Frequencies

Επιλέγουμε τις μεταβλητές «Βάρος» και «Αθηρωματικός δείκτης» και πατάμε το κουμπί ***Statistics***

Frequency dialog box showing variable selection:

- Available variables: Οικογενειακή κατ..., Φύλο [GENDER], Ύψος [HEIGHT], Καλή Χοληστερίνη..., Ολική Χοληστερίν..., Κάπνισμα [SMOK...], BMI, Καρδιοαγγειακος ...
- Selected variables: Βάρος [WEIGHT], Αθηρωματικός Δείκτ...
- Buttons: Statistics..., Charts..., Format..., Style..., Bootstrap...
- Options: Display frequency tables, Create APA style tables
- Buttons: OK, Paste, Reset, Cancel, Help

Γυναίκα	114	176	41	236	Καπνιστής
Αντρας	120	170	38	225	Μη καπνιστής
Αντρας	82	170	89	204	Μη καπνιστής
Γυναίκα	56	166	68	157	Καπνιστής
Γυναίκα	92	172	51	214	Καπνιστής
...

Frequencies: Statistics dialog box showing statistical options:

- Percentile Values:
 - Quartiles
 - Cut points for: 10 equal groups
 - Percentile(s): []
 - Buttons: Add, Change, Remove
- Central Tendency:
 - Mean
 - Median
 - Mode
 - Sum
 - Values are group midpoints
- Dispersion:
 - Std. deviation
 - Variance
 - Range
 - Minimum
 - Maximum
 - S.E. mean
- Distribution:
 - Skewness
 - Kurtosis
- Buttons: Continue, Cancel, Help

Στο παράθυρο που ανοίγει επιλέγουμε:

- *Mean* (Μέση τιμή)
- *Median* (Διάμεσος)
- *Mode* (Επικρατούσα τιμή)
- *Quartiles* (Τεταρτημόρια)
- *Percentiles* (Ποσοστημόρια) όπου πρέπει να γράψουμε τα ποσοστά 60 και 85 και να πατήσουμε το κουμπί **Add**
- *Std deviation* (Τυπική απόκλιση)
- *Variance* (Διακύμανση)
- *Range* (Εύρος)

Στη συνέχεια πατάμε *Continue* και μετά **OK**

Οπότε στο Output εμφανίζεται ο πίνακας **Statistics** με τα περιγραφικά στατιστικά που ζητήσαμε:

→ **Frequencies**

		Statistics	
		Βάρος	Αθροισματικό Δείκτης
N	Valid	20	20
	Missing	0	0
Mean		91,05	4,2774
Median		92,00	4,0223
Mode		78 ^a	2,15 ^a
Std. Deviation		20,814	1,56437
Variance		433,208	2,447
Range		70	5,79
Percentiles	25	78,50	2,9362
	50	92,00	4,0223
	60	97,40	4,1959
	75	108,50	5,5726
	85	117,40	6,1616

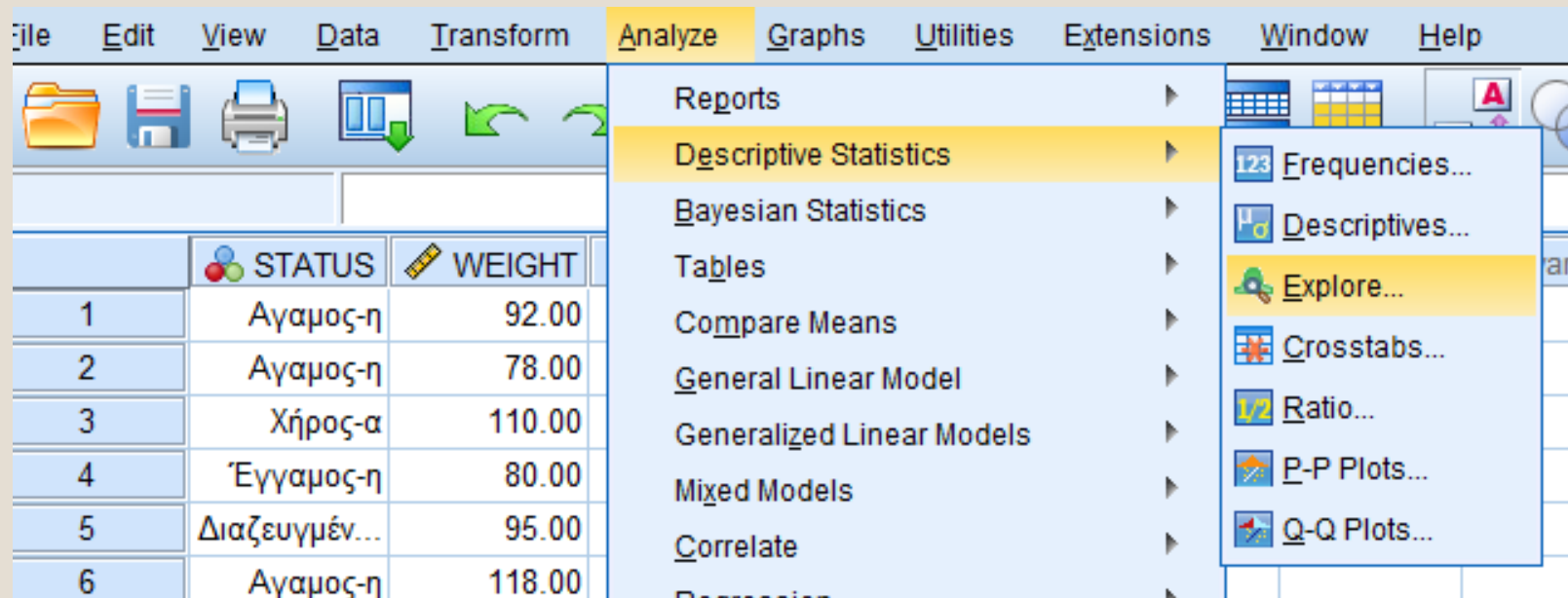
a. Multiple modes exist. The smallest value is shown

Άσκηση 1.9: Στο αρχείο της άσκησης 1.4, να βρεθούν τα κυριότερα μέτρα θέσης της μεταβλητης βάρος, αλλά χωριστά για καπνιστές και μη καπνιστές.

Λύση

Η διαδρομή στο SPSS είναι:

Analyze / Descriptive Statistics / Explore

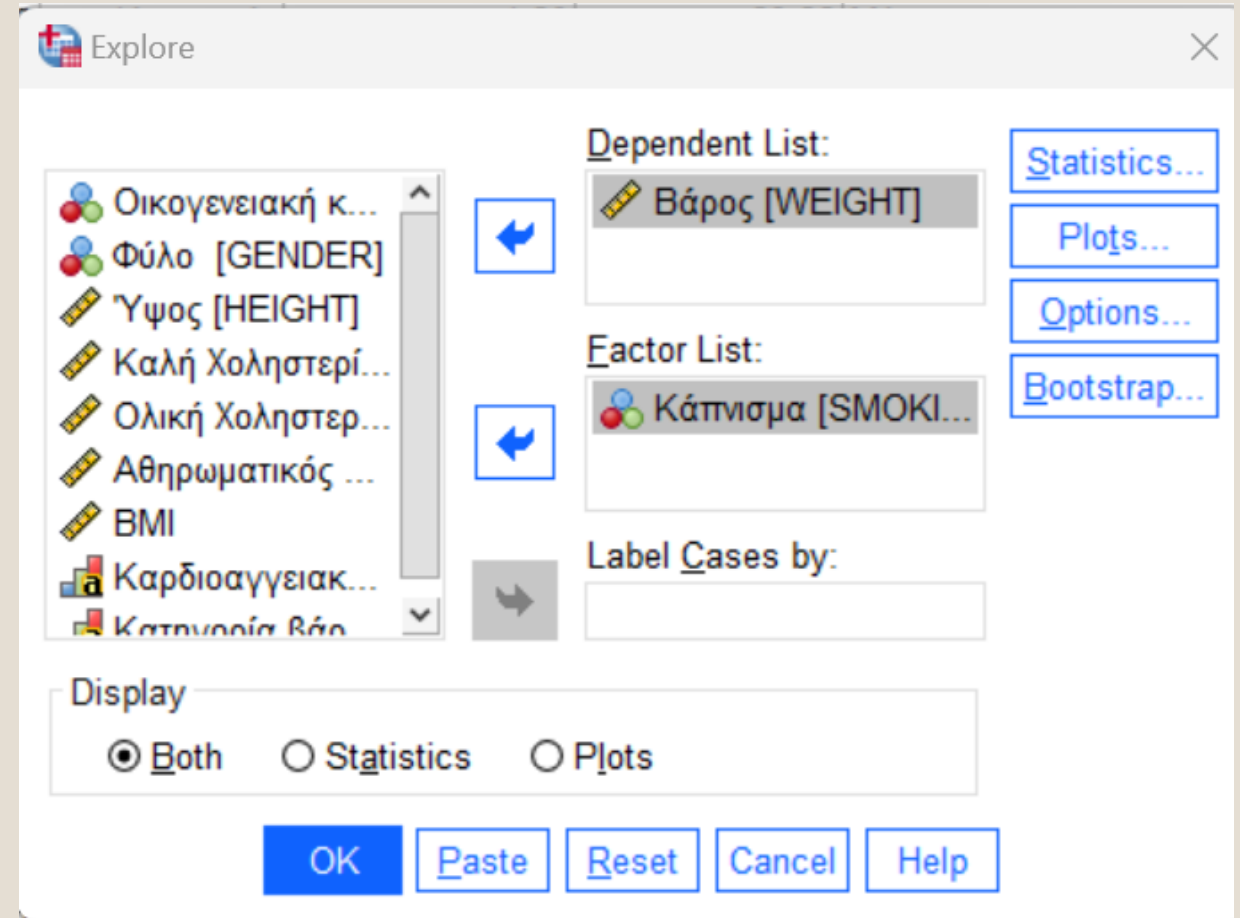


- Στο παράθυρο που ανοίγει, τοποθετούμε τη ποσοτική μεταβλητή (WEIGHT) στο πλαίσιο **Dependent List** και τη ποιοτική μεταβλητή (SMOKING) στο πλαίσιο **Factor List**

Στο **Display** κλικάρουμε

- **Both** εάν θέλουμε μαζί με τα μέτρα να σχεδιαστούν τα αντίστοιχα θηκογράμματα (Boxplots)
- **Statistics** εάν θέλουμε να εμφανιστούν μόνο τα περιγραφικά μέτρα
- **Plots** εάν θέλουμε να εμφανιστούν μόνο γραφήματα

Τέλος πατάμε **OK**.



Descriptives

Κάπνισμα		Statistic	Std. Error		
Βάρος	Καπνιστής	Mean	91,60	7,157	
		95% Confidence Interval for Mean	Lower Bound	75,41	
			Upper Bound	107,79	
		5% Trimmed Mean	91,78		
		Median	93,50		
		Variance	512,267		
		Std. Deviation	22,633		
		Minimum	55		
		Maximum	125		
		Range	70		
		Interquartile Range	32		
		Skewness	-,524	,687	
		Kurtosis	-,227	1,334	
		Μη καπνιστής	Μη καπνιστής	Mean	90,50
95% Confidence Interval for Mean	Lower Bound			76,16	
	Upper Bound			104,84	
5% Trimmed Mean	90,56				
Median	81,00				
Variance	401,611				
Std. Deviation	20,040				
Minimum	60				
Maximum	120				
Range	60				
Interquartile Range	34				
Skewness	,333			,687	
Kurtosis	-1,061			1,334	

Επειδή είχαμε επιλέξει **Both** το πρόγραμμα εμφάνισε και τα αντίστοιχα γραφήματα:

