

ΕΙΣΑΓΩΓΗ ΣΤΗ ΧΡΗΣΗ ΤΟΥ
ΣΤΑΤΙΣΤΙΚΟΥ ΠΑΚΕΤΟΥ
MINITAB 12.0

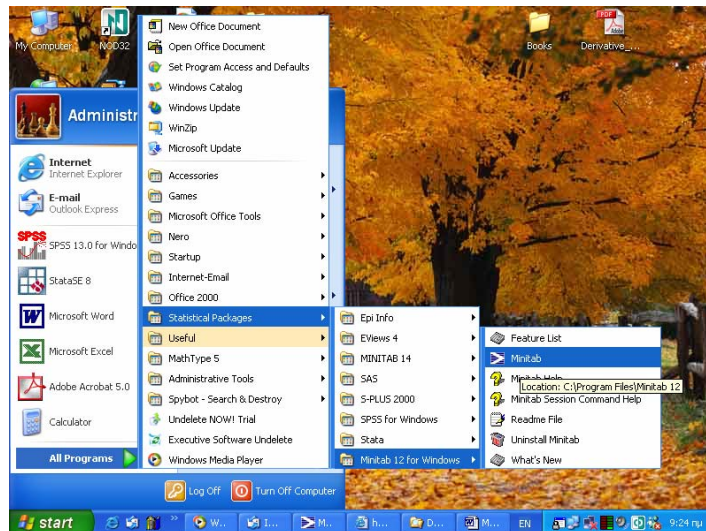
Περιεχόμενα

⊙	Εισαγωγή στο Στατιστικό πακέτο Minitab 12.0	1
❖	Αρχικά	1
❖	Παράθυρο του Minitab.....	1
	Status bar και Toolbar	2
	Menu bar	2
	Session και worksheet window.....	2
	Graph window	3
❖	Ενεργοποίηση παραθύρου.....	4
❖	Μεγιστοποίηση και ελαχιστοποίηση παραθύρου.....	4
❖	Εισαγωγή δεδομένων.....	5
	Πληκτρολόγηση.....	5
	Αντιγραφή-επικόλληση	7
	Άνοιγμα Δεδομένων από αρχείο.....	8
❖	Έλεγχος στοιχείων/Αρχείου.....	18
❖	Δημιουργία μεταβλητών	19
❖	Χειρισμός μεταβλητών.....	19
	Αλλαγή τύπου.....	19
	Κωδικοποίηση	20
❖	Επιλογή δεδομένων	21
❖	Στατιστική Ανάλυση.....	23
	Περιγραφική στατιστική	23
	Κατανομή συχνοτήτων	25
❖	Γραφική παράσταση δεδομένων	27
❖	Επεξεργασία γραφημάτων	33
❖	Χρήση Εντολών.....	35
❖	Ρυθμίσεις.....	36
❖	Βοήθεια	37
❖	Αποθήκευση.....	37
❖	Εκτύπωση.....	39
⊙	Εφαρμογή με χρήση του Στατιστικού πακέτου Minitab 12.0.....	40

© Εισαγωγή στο Στατιστικό πακέτο Minitab 12.0

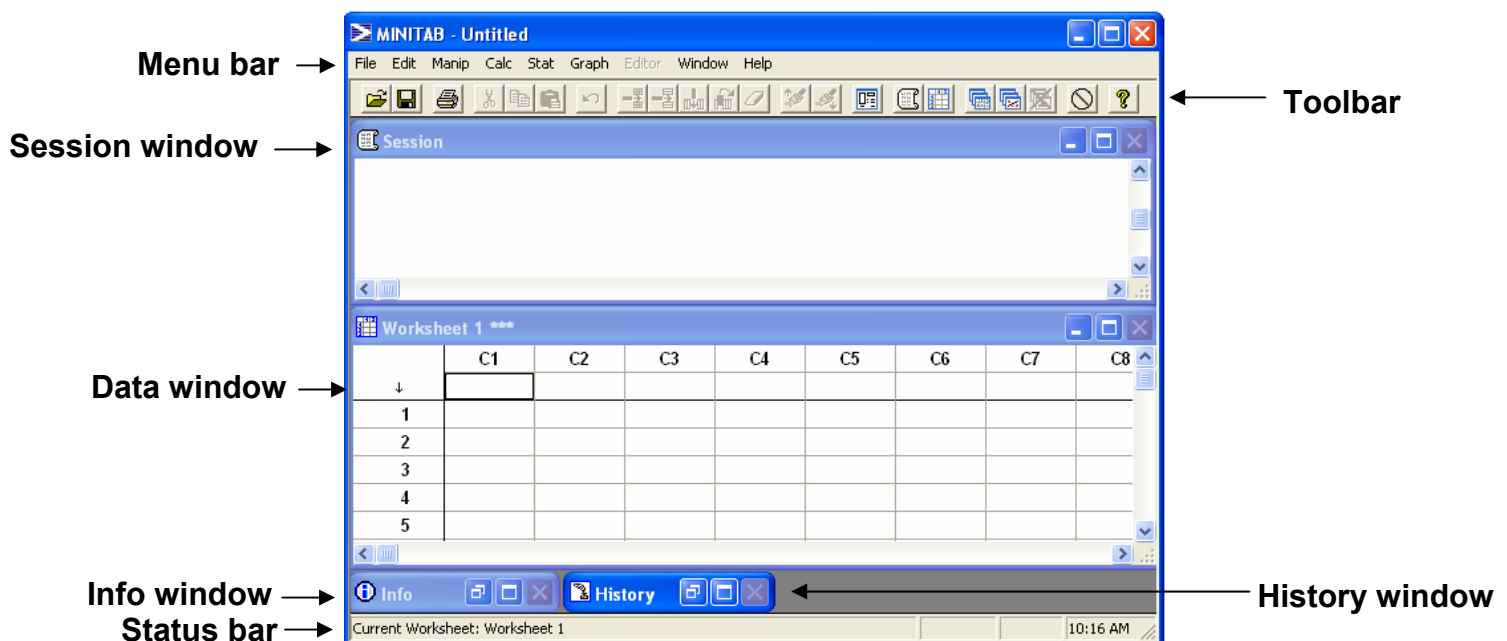
❖ Αρχικά

Ανοίγουμε το *Minitab* κάνοντας διπλό κλικ στο εικονίδιο **Minitab** στην οθόνη ή πηγαίνοντας στο Start/Statistical Packages/Minitab 12.



Όταν το πρόγραμμα “φορτώσει”, εμφανίζεται στη οθόνη το βασικό παράθυρο του Minitab

❖ Παράθυρο του Minitab



Status bar και Toolbar

Το **Status bar** παρουσιάζει χρήσιμα μηνύματα που εξηγούν την τρέχουσα στιγμή ποιό στοιχείο έχει επιλεγεί από το Menu bar ή ποιό εργαλείο έχει επιλεγεί από τη ράβδο εργαλείων.

Παρατηρούμε την αλλαγή κειμένου επιλέγοντας με το δείκτη του ποντικιού για παράδειγμα τη θέση Stat του Menu bar και οποιαδήποτε από τις υποκαρτέλες που εμφανίζονται ή κατά τον ίδιο τρόπο επιλέγοντας οποιοδήποτε εικονίδιο από το Tool bar.

Το **toolbar** βρίσκεται ακριβώς κάτω από το **menu bar** και εμφανίζει εικονίδια χρήσιμα για λειτουργίες που χρησιμοποιούνται συχνά.

Για να μια σύντομη περιγραφή του κάθε εικονιδίου αρκεί να μετακινήσουμε το δείκτη του ποντικιού πάνω σε κάθε εικονίδιο.

Menu bar

Το **Menu bar** περιέχει εννέα κυλιόμενα μενού, μέσω των οποίων μπορούμε να εκτελέσουμε μία εντολή ως εξής:

1. Κάνουμε κλικ στο όνομα του στοιχείου, όπως File.
2. Κάνουμε κλικ στην εντολή που θέλουμε από το συγκεκριμένο μενού, όπως New ή πατάμε το γράμμα που συνδέεται με αυτή την εντολή, N για το New).
3. Αν η εντολή απαιτεί να δηλώσουμε κάποιες επιπλέον πληροφορίες, τότε το Minitab ανοίγει ένα πλαίσιο διαλόγου, διαφορετικά εκτελεί την εντολή άμεσα.

Παρακάτω, στην παράγραφο **Περιγραφική Στατιστική**, θα δούμε αναλυτικά τη χρήση εντολών για την Περιγραφική στατιστική μεταβλητών.

Session και worksheet window

Στο παράθυρο του Minitab εμφανίζονται δύο βασικά παράθυρα:

- ✚ το **Session window**, όπου εμφανίζονται οι εντολές και τα αποτελέσματα. Στο παράθυρο αυτό μπορούμε να παρέμβουμε γράφοντας κάποιες σημειώσεις ή διαγράφοντας κάποια αποτελέσματα και
- ✚ το **Data window ή worksheet**, όπου εισάγουμε, βλέπουμε και επεξεργαζόμαστε τα δεδομένα.

Αρχικά το worksheet είναι άδειο. Κάθε γραμμή των δεδομένων θα περιέχει στοιχεία-δεδομένα για ένα συγκεκριμένο άτομο / αντικείμενο. Κάθε στήλη (C1, C2,...) αντιστοιχεί σε μία μεταβλητή.

Τα κελιά κάτω από τους τίτλους των στηλών (C1, C2, C3, ...)

Worksheet 1 ***			
	C1	C2	C3
↓	id	type	class

είναι για τη λεκτική περιγραφή- ονομασία των μεταβλητών, δεν πρέπει να εισάγουμε δεδομένα.

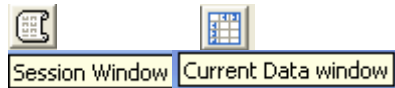
Graph window

Ένα επιπλέον παράθυρο του Minitab είναι το Graph window, το οποίο περιέχει γράφημα των δεδομένων.

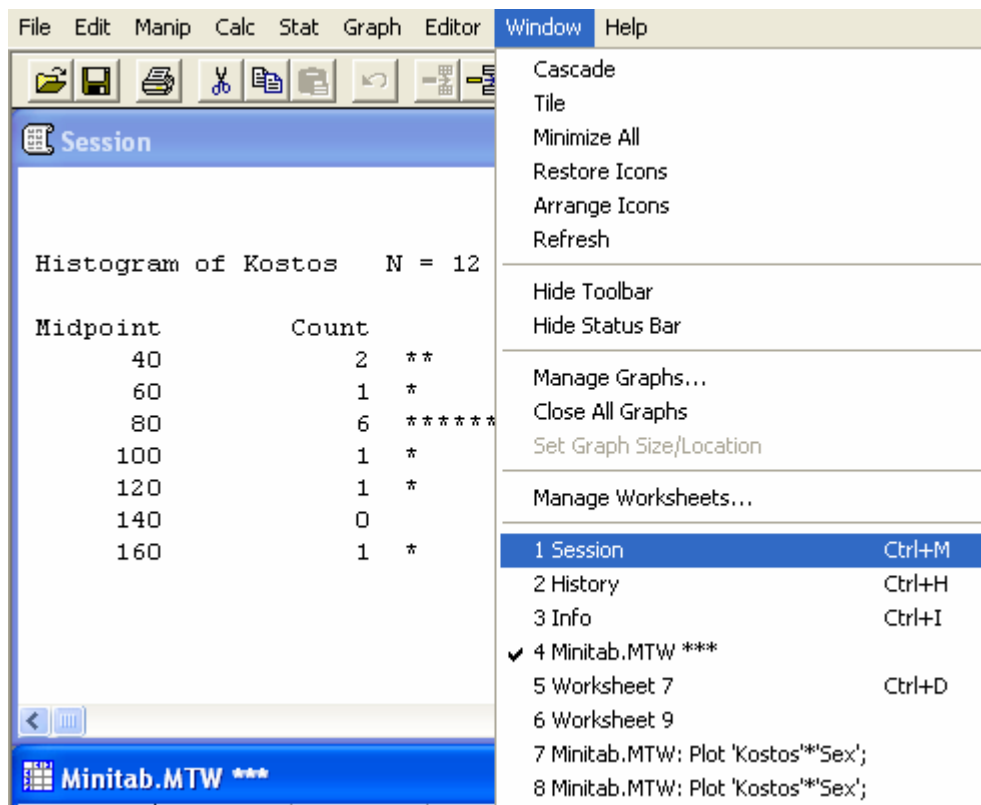
Στο Minitab υπάρχει μόνο ένα Session window ενώ μπορούν να εμφανιστούν πολλά Data και Graph windows. Παρότι πολλά παράθυρα μπορεί να είναι ορατά ταυτόχρονα, μόνο ένα παράθυρο είναι ενεργό. Στο παράθυρο που είναι ενεργό η μπάρα που περιέχει τον τίτλο του είναι φωτισμένη.

❖ Ενεργοποίηση παραθύρου

Για να κάνουμε ένα παράθυρο ενεργό απλά κάνουμε κλικ με το ποντίκι επάνω του. Μπορούμε επίσης να ενεργοποιήσουμε ένα παράθυρο κάνοντας κλικ στο κατάλληλο εικονίδιο του toolbar.




Ένας τρίτος τρόπος είναι χρησιμοποιώντας κυλιόμενο menu κάτω από την εντολή Window.



Το σύμβολο ✓ μπροστά από το όνομα δηλώνει ποιο παράθυρο είναι ενεργό. Κάνοντας κλικ σε οποιοδήποτε από τα παράθυρα μπορούμε να το ενεργοποιήσουμε.

❖ Μεγιστοποίηση και ελαχιστοποίηση παραθύρου

Με τα τρία αυτά κουμπιά  τα οποία εμφανίζονται στην άκρη δεξιά από κάθε παράθυρο ελέγχουμε το μέγεθος του παραθύρου. Το πρώτο κουμπί ελαχιστοποιεί το παράθυρο, ώστε να μην είναι πλέον ορατό. Το παράθυρο μπορεί να γίνει πάλι ορατό μέσω της εντολής Windows και της διαδικασίας που περιγράφηκε στην ενεργοποίηση παραθύρου. Το μεσαίο

μας δίνει τη δυνατότητα εναλλαγής μεταξύ μεγιστοποίησης ενός παραθύρου και επαναφοράς του στο ενδιάμεσο μέγεθος. Το δεξί κουμπί κλείνει ένα παράθυρο. Αν το περιεχόμενο του παραθύρου δεν έχει αποθηκευθεί τότε αυτό χάνεται.



Το ίδιο σετ κουμπιών υπάρχει και στην κορυφή της οθόνης τα οποία αντιστοιχούν στις ίδιες λειτουργίες για το πρόγραμμα Minitab.

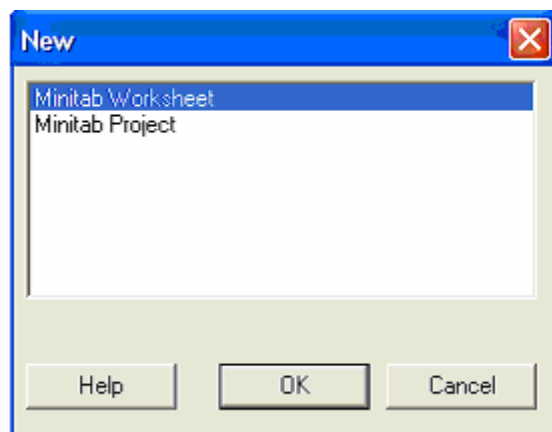
❖ Εισαγωγή δεδομένων

Το Minitab δίνει τη δυνατότητα να εισάγουμε δεδομένα με αρκετούς τρόπους όπως:

1. Πληκτρολογώντας δεδομένα κατευθείαν στο data window
2. Με αντιγραφή και επικόλληση
3. Άνοιγμα δεδομένων από ένα αρχείο.

Πληκτρολόγηση

Για να εισάγουμε δεδομένα σε ένα φύλο εργασίας δεδομένων του Minitab θα πρέπει να ανοίξουμε ένα κενό φύλο εργασίας δεδομένων επιλέγοντας File/New και από το παράθυρο διαλόγου



να επιλέξουμε Minitab Worksheet.

Για να καταχωρήσουμε δεδομένα σε στήλες θα πρέπει το **βέλος κατεύθυνσης δεδομένων** που εμφανίζεται στο αριστερό πεδίο πριν τα ονόματα των στηλών να έχει κατεύθυνση προς τα κάτω ενώ αν θέλουμε να εισάγουμε δεδομένα σε γραμμές θα πρέπει να έχει κατεύθυνση προς τα

δεξιά. Τα δεδομένα εισάγονται με πληκτρολόγηση και πάτημα του πλήκτρου Tab ή Enter.

Worksheet 2 ***		
Βέλος Κατεύθυνσης δεδομένων ↓	C1	C2
1	1	
2	5	
3	7	
4		
5		

Εισαγωγή κατά στήλες

Worksheet 2 ***			
→	C1	C2	C3
1	1	5	7
2			

Εισαγωγή κατά γραμμές

Πατώντας **Ctrl+Enter** μετακινούμε το ενεργό κελί στην κορυφή της επόμενης στήλης ή γραμμής ανάλογα με την κατεύθυνση του βέλους κατεύθυνσης δεδομένων.

Παράδειγμα εισαγωγής δεδομένων

Δεκατέσσερις άνθρωποι ρωτήθηκαν πόσο κοστίζει το ακριβότερο τους ζευγάρι παπούτσια ενώ καταγράφηκε και το φύλο κάθε ατόμου καθώς και το πόσα ζευγάρια παπούτσια αγόρασαν την προηγούμενη χρονιά, όπως φαίνεται παρακάτω.

No	Κόστος	Φύλο	Πλήθος	No	Κόστος	Φύλο	Πλήθος
1	70	Γυναίκα	2	7	70	Γυναίκα	3
2	80	Γυναίκα	4	8	150	Άνδρας	1
3	90	Γυναίκα	2	9	45	Γυναίκα	4
4	85	Άνδρας	5	10	70	Άνδρας	6
5	30	Γυναίκα	2	11	110	Άνδρας	3
6	65	Γυναίκα	4	12	85	Άνδρας	3
7	70	Γυναίκα	3	13	85	Άνδρας	2
8	150	Άνδρας	1	14	80	Άνδρας	2

Εισάγουμε τα στοιχεία στο παράθυρο στοιχείων Minitab. Κάνουμε κλικ στο πρώτο κενό κελί κάτω ακριβώς από τη από τη στήλη C1 και πληκτρολογούμε σε αυτό το διάστημα τη λέξη **kostos** (χρήση greeklish). Πληκτρολογούμε επίσης τη λέξη **sex (φύλο)** ως τίτλο για τη στήλη C2.

↓	C1	C2-T	C3
	Κόστος	Φύλο	No
1	70	F	2
2	80	F	4
3	90	F	2
4	85	M	5
5	30	F	2
6	65	F	4
7	70	F	3
8	150	M	1
9	45	F	4
10	70	M	6
11	110	M	3
12	85	M	3
13	85	M	2
14	80	M	2

Κατόπιν εισάγουμε τις τιμές των στοιχείων όπως μας δόθηκαν. Μπορούμε να κινηθούμε από κελί σε κελί μέσα στο παράθυρο στοιχείων με τη χρησιμοποίηση του ποντικιού ή με τη χρησιμοποίηση των βελών στο πληκτρολόγιό μας. Σημειώστε ότι η "C2" αλλάζει "στο C2-T" όταν αρχίζουμε να εισάγουμε τα στοιχεία για το φύλο. Αυτό σημαίνει ότι το Minitab θεωρεί το φύλο μια μεταβλητή "κειμένου" ("κατηγορική" ή "ποιοτική"), παρά μια ποσοτική μεταβλητή.

Παρατηρούμε τη μορφή των στοιχείων. Κάθε σειρά περιέχει όλες τις πληροφορίες για ένα άτομο και κάθε στήλη αντιπροσωπεύει την πληροφορία (ή τη μεταβλητή) που έχει καταγραφεί. Αυτή είναι η συνηθισμένη μορφή που χρησιμοποιείται για όλους τους τύπους στατιστικών αναλύσεων.

Αντιγραφή-επικόλληση

Το Minitab επιτρέπει:

- ✚ Αντιγραφή δεδομένων από ένα άλλο πρόγραμμα χρήσης φύλλων εργασίας, όπως as Lotus1-2-3 or Microsoft Excel και επικόλληση στο φύλο εργασίας του Minitab .
- ✚ Αντιγραφή δεδομένων από έναν πίνακα του επεξεργαστή κειμένου Word και επικόλληση στο φύλο εργασίας του Minitab .
- ✚ Αντιγραφή δεδομένων από ένα φύλο εργασίας του Minitab και επικόλληση σε ένα άλλο φύλο εργασίας του Minitab.
- ✚ Αντιγραφή δεδομένων από το **Session** window του Minitab επικόλληση στο **Data** window του Minitab

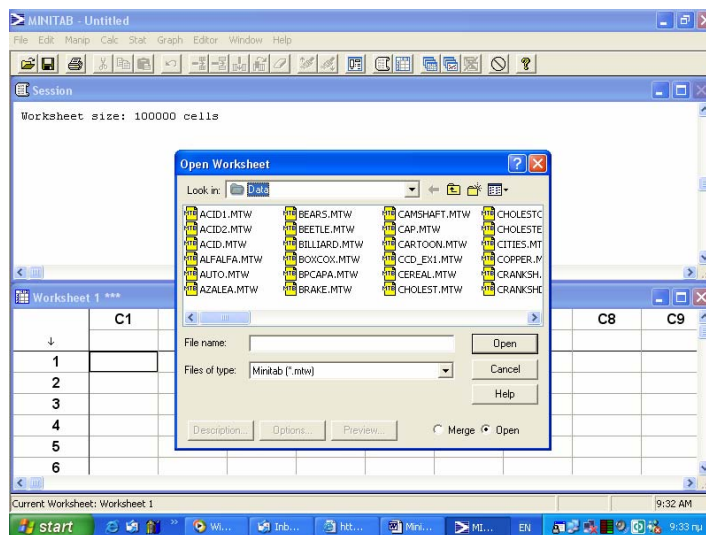
Για να γίνει η αντιγραφή και επικόλληση αρκεί να ακολουθηθούν τα έξης βήματα:

1. Μαρκάρουμε τα δεδομένα.
2. Επιλέγουμε **Edit** από το menu bar και μετά **Copy Cells**.
3. Κάνουμε κλικ στο **Minitab Data** window
4. Επιλέγουμε ένα από τα ακόλουθα:
 - Για να εισάγουμε δεδομένα στο ίδιο worksheet κάνουμε κλικ στο κελί από όπου θέλουμε να ξεκινά η επικόλληση , επιλέγουμε **Edit** και μετά **Paste/Insert Cells**.
 - Για να αντικαταστήσουμε δεδομένα στο τρέχον worksheet μαρκάρουμε τα κελιά με τα δεδομένα που θέλουμε να αντικατασταθούν, επιλέγουμε **Edit** και μετά **Paste Cells**.

Άνοιγμα Δεδομένων από αρχείο

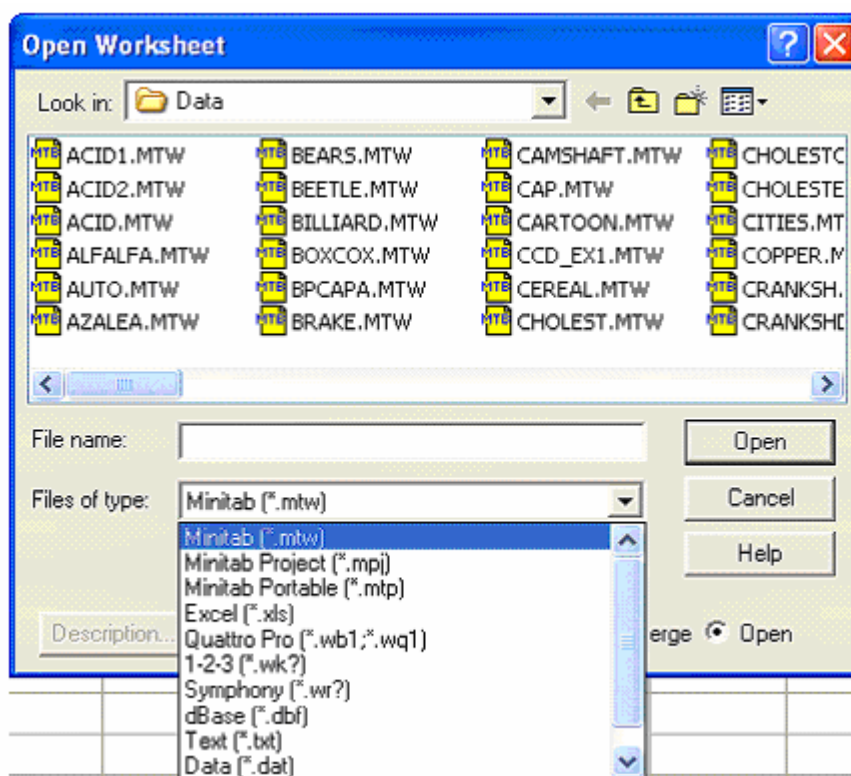
Εισαγωγή αρχείου Minitab

Μπορούμε να ανοίξουμε ένα υπάρχον φύλλο εργασίας δεδομένων του Minitab επιλέγοντας **File>Open Worksheet** και εντοπίζοντας μέσω ενός πλαισίου διαλόγου το φάκελο που είναι αποθηκευμένο το αρχείο.



Τα φύλλα εργασίας δεδομένων του Minitab έχουν κατάληξη “.MTB”.

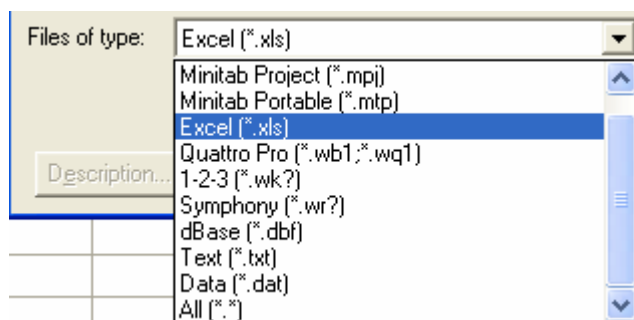
Μπορούμε να ανοίξουμε *διάφορους τύπους αρχείων* με την κατάλληλη επιλογή του τύπου αρχείου στο πλαίσιο των αρχείων καταλόγων (δείτε το πλαίσιο διαλόγου)



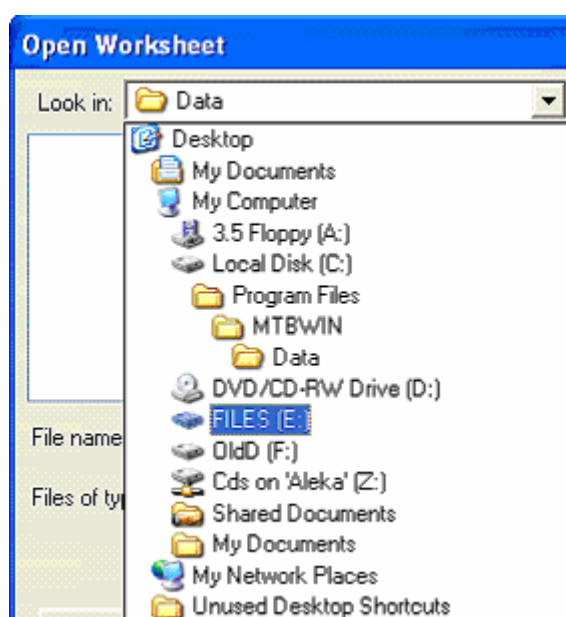
Εισαγωγή φύλλου εργασίας δεδομένων του Excel

Μπορούμε να ανοίξουμε ένα φύλλο εργασίας δεδομένων του Excel ακολουθώντας τα παρακάτω βήματα.

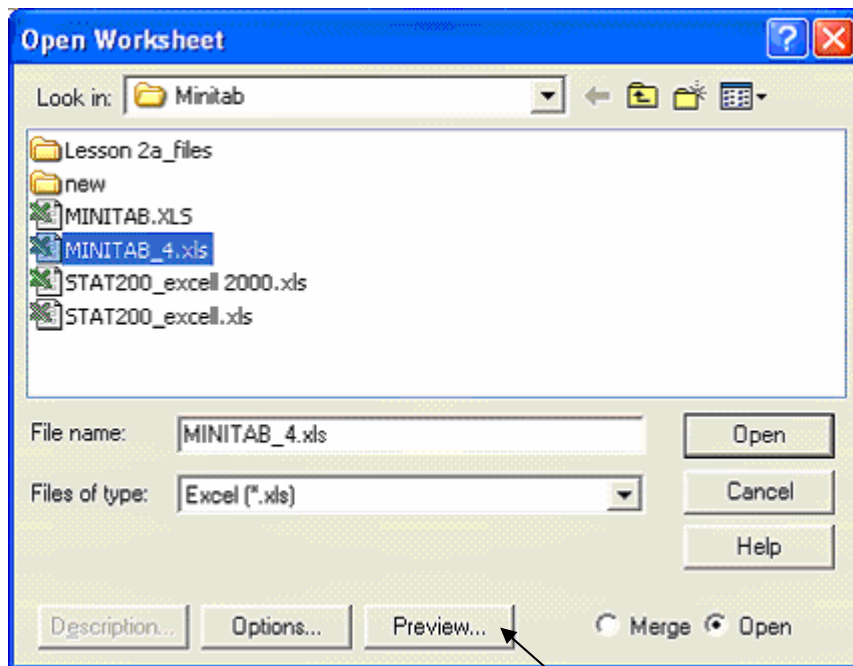
Στο πλαίσιο διαλόγου *Files of Type* επιλέγουμε ως τύπο αρχείου που θέλουμε να ανοίξουμε **Excel**



Στη συνέχεια προσδιορίζουμε τη θέση του αρχείου μέσω του πλαισίου διαλόγου *Look in*,



Αφότου εντοπίσουμε τη θέση όπου βρίσκεται το αρχείο προχωρούμε κάνοντας χρήση της καρτέλας *Preview*



μέσω της οποίας μπορούμε να έχουμε μία εικόνα για τη μορφή των δεδομένων που επιθυμούμε να εισάγουμε¹.

Μέσω της επισκόπηση του αρχείου αντλούμε πληροφορίες τις οποίες θα πρέπει να συμπληρώσουμε στις κατάλληλες θέσεις για να γίνει η ομαλή εισαγωγή των δεδομένων.

Παράδειγμα εισαγωγής αρχείου Excel

Θεωρούμε ότι οι πληροφορίες που μας δόθηκαν στο προηγούμενο παράδειγμα σχετικά με την αγορά παπουτσιών έχουν αποθηκευθεί σε ένα φύλο εργασίας Excel με την ονομασία MIINITAB_4.xls.

¹ Θα πρέπει εδώ να σημειωθεί ότι η έκδοση MINITAB 12 επιτρέπει την εισαγωγή φύλου δεδομένων Excel μόνο της έκδοσης 4.0. Για το λόγο αυτό θα πρέπει να προηγηθεί η μετατροπή του φύλου δεδομένων Excel από οποιαδήποτε μεταγενέστερη έκδοση σε έκδοση Excel 4.0.

	A	B	C
1	Id	Kostos	Filo
2	1	70	F
3	2	80	F
4	3	90	F
5	4	85	M
6	5	30	F
7	6	65	F
8	7	70	F
9	8	150	M
10	9	45	F
11	10	70	M
12	11	110	M
13	12	85	M
14	13	85	M
15	14	80	M
16			

Open Worksheet - Preview - E:\LESSON INTRODUCTION TO STATISTICS

Name: Id, Kostos, Filo

Type: Numeric, Numeric, Text

Column: C1, C2, C3

Row: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14

Preview Data:

Row	Id	Kostos	Filo
2	1	70	F
3	2	80	F
4	3	90	F
5	4	85	M
6	5	30	F
7	6	65	F
8	7	70	F
9	8	150	M
10	9	45	F
11	10	70	M
12	11	110	M
13	12	85	M
14	13	85	M

Help Display data rows only OK

Μέσω της προεπισκόπησης παρατηρούμε ότι κατά την ανάγνωση των δεδομένων το Minitab αναγνωρίζει ότι υπάρχουν 3 στήλες δεδομένων, C1-C2-C3, δηλώνοντας τα ονόματα των μεταβλητών όπως έχουν καθοριστεί, οι τιμές των δεδομένων ξεκινούν στη Γραμμή 2. Επίσης αναγνωρίζει σωστά τον τύπο των μεταβλητών καθώς παρατηρούμε ότι δηλώνει ως text τον τύπο της μεταβλητής Filo που βρίσκεται στη στήλη 3 καθώς περιέχει χαρακτήρες και όχι αριθμητικές τιμές.

Η εισαγωγή του αρχείου αυτού μπορεί να γίνει άμεσα αφότου βγούμε από το παράθυρο της προεπισκόπησης, με χρήση της εντολής Open.

Σε περίπτωση που κατά την προεπισκόπηση δεν έχει γίνει σωστή ανάγνωση του αρχείου, τότε μέσω της καρτέλας Options η οποία βρίσκεται ακριβώς δίπλα στο Preview, μπορούμε να κάνουμε τις κατάλληλες τροποποιήσεις και μετέπειτα να προχωρήσουμε στο άνοιγμα του αρχείου.

Αν για παράδειγμα η πρώτη γραμμή του φύλλου εργασίας του αρχείου Excel είναι κενή, όπως στο Minitab_41.xls, τότε η ανάγνωση του αρχείου δεν γίνεται ορθά.

	A	B	C
1			
2	Id	Kostos	Filo
3	1	70	F
4	2	80	F
5	3	90	F
6	4	85	M
7	5	30	F
8	6	65	F
9	7	70	F
10	8	150	M
11	9	45	F
12	10	70	M
13	11	110	M
14	12	85	M
15	13	85	M
16	14	80	M

Open Worksheet - Preview - E:\LESSON INTRODUCTION TO STATISTICS\SW

Name: [] [] [] []

Type: Text Text Text

Column: C1 C2 C3

Row:

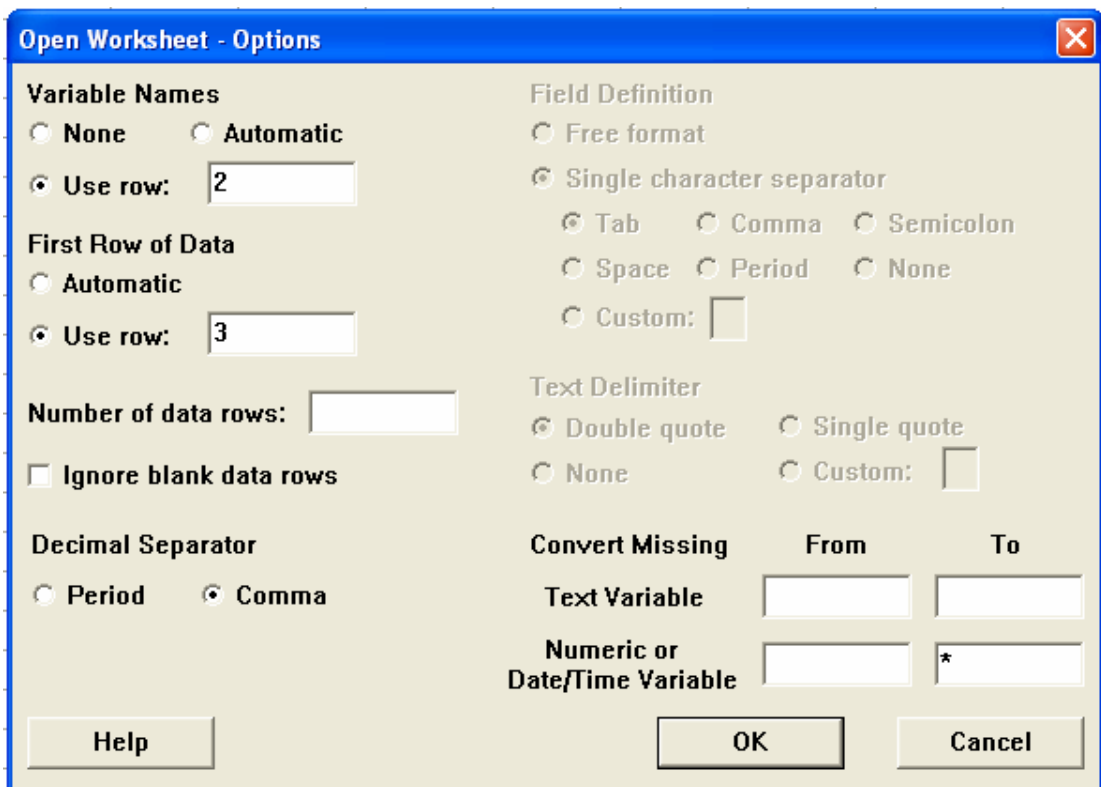
Row	Id	Kostos	Filo
2			
3	1.0000000000	70.0000000000	F
4	2.0000000000	80.0000000000	F
5	3.0000000000	90.0000000000	F
6	4.0000000000	85.0000000000	M
7	5.0000000000	30.0000000000	F
8	6.0000000000	65.0000000000	F
9	7.0000000000	70.0000000000	F
10	8.0000000000	150.0000000000	M
11	9.0000000000	45.0000000000	F
12	10.0000000000	70.0000000000	M
13	11.0000000000	110.0000000000	M
14	12.0000000000	85.0000000000	M

Help Display data rows only OK

Παρατηρώντας τη δομή του αρχείου, στην καρτέλα Options προχωρούμε ως εξής

- Στο πεδίο Variable Names ,κάνουμε χρήση της επιλογής να δηλώσουμε εμείς από ποια γραμμή θα γίνει η ανάγνωση των ονομάτων των μεταβλητών (Row 2)
- Στο πεδίο First Row of Data , κάνουμε χρήση της επιλογής να δηλώσουμε εμείς από ποια γραμμή θα γίνει η ανάγνωση των δεδομένων (Row 3)

Αφότου κλείσουμε την καρτέλα Options μπορούμε είτε να επαναλάβουμε την προεπισκόπηση για να βεβαιωθούμε για την ορθή ανάγνωση του αρχείου είτε να προχωρήσουμε στην άμεση ανάγνωση του αρχείου μέσω της εντολής Open.



Οι υπόλοιπες επιλογές της καρτέλας Options αποτελούν χρήσιμες δυνατότητες για την εισαγωγή δεδομένων με συγκεκριμένη μορφή όπως στην εισαγωγή αρχείου μορφής ASCII (κείμενο) που θα δούμε παρακάτω.

Εισαγωγή αρχείου ASCII (κείμενο)

Μπορούμε να επιλέξουμε να ανοίξουμε ένα αρχείο στοιχείων ASCII (κείμενο) επιλέγοντας μία από τις παρακάτω δύο διαδικασίες:

- File>Other Files>Import Special Text.

Με τον τρόπο αυτό μπορεί να εισαχθεί ένα αρχείο

- a. το οποίο έχει σχηματοποιηθεί με κάποιο ιδιαίτερο τρόπο (special format).
- b. να επισυνάπτουν τα στοιχεία από ένα αρχείο στο κατώτατο σημείο των υπαρχουσών στηλών.
- c.

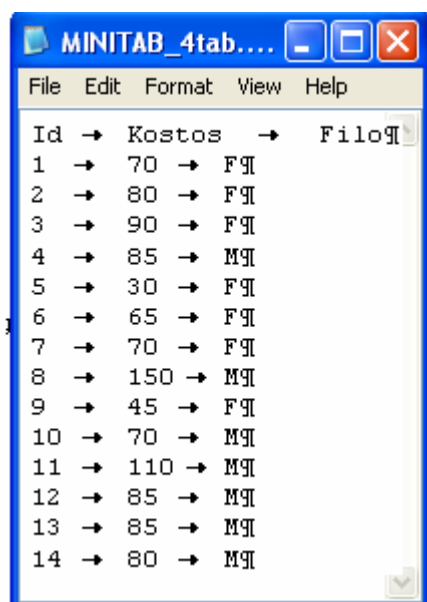
- File > Open Worksheet (ή File) και κατάλληλη χρήση της καρτέλας Options

Με το File > Open Worksheet (ή File) τα στοιχεία τοποθετούνται σε ένα νέο φύλλο εργασίας, ή, όταν χρησιμοποιείται την επιλογή συγχώνευσης (**Merge**), τοποθετούνται σε νέες στήλες στο τρέχον φύλλο εργασίας.

Παράδειγμα εισαγωγής αρχείου ASCII (κείμενο)

- Με χρήση της επιλογής File>Other Files>Import Special Text.

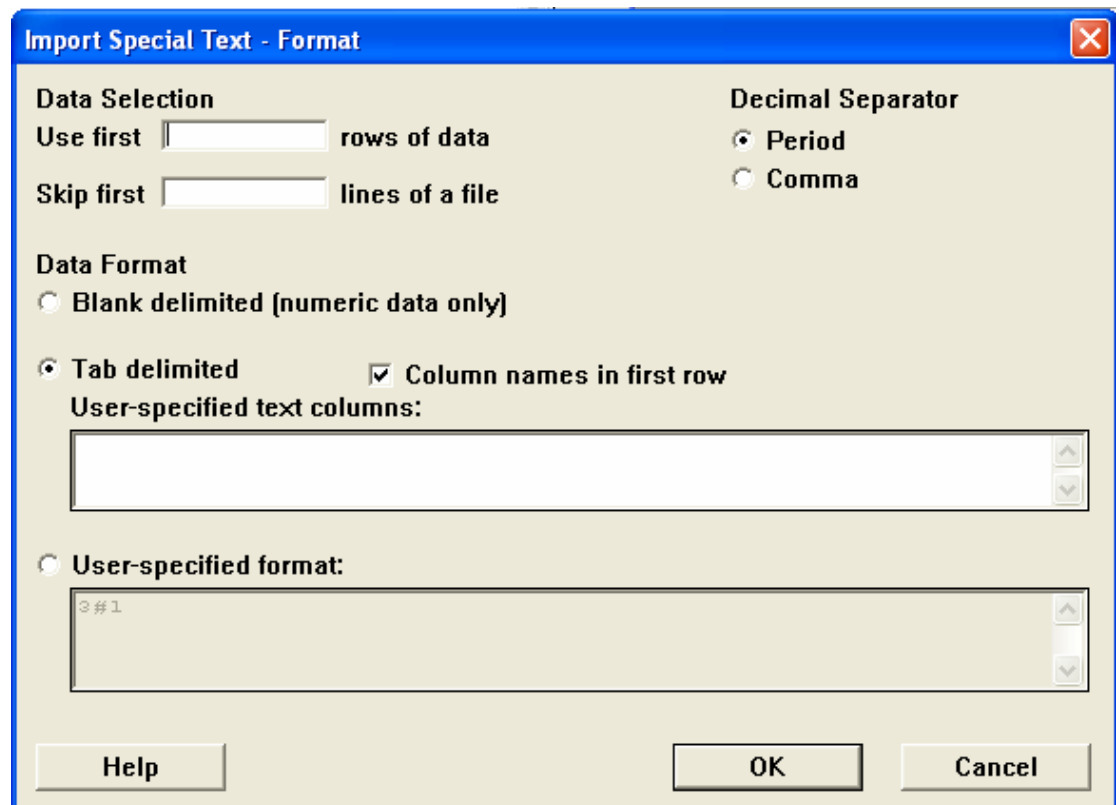
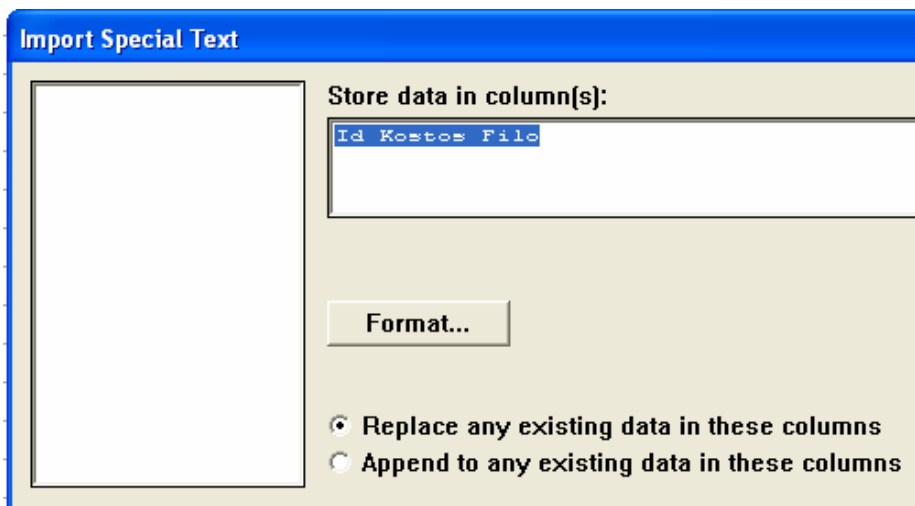
Θεωρούμε ότι οι πληροφορίες που μας δόθηκαν σε προηγούμενο παράδειγμα σχετικά με την αγορά παπουτσιών έχουν αποθηκευθεί σε ένα αρχείο κειμένου με την ονομασία MIINITAB_4tab.txt, όπου οι μεταβλητές χωρίζονται από το σύμβολο της παραγράφου → .



Id	Kostos	Filoq
1	70	F
2	80	F
3	90	F
4	85	M
5	30	F
6	65	F
7	70	F
8	150	M
9	45	F
10	70	M
11	110	M
12	85	M
13	85	M
14	80	M

Για να γίνει ορθά η ανάγνωση του αρχείου δίνουμε τις στήλες που θα εισαχθούν τα δεδομένα, C1 C2 C3 ή τα ονόματά τους και στην καρτέλα Format δηλώνουμε ότι η μορφή των δεδομένων καθορίζεται ως :

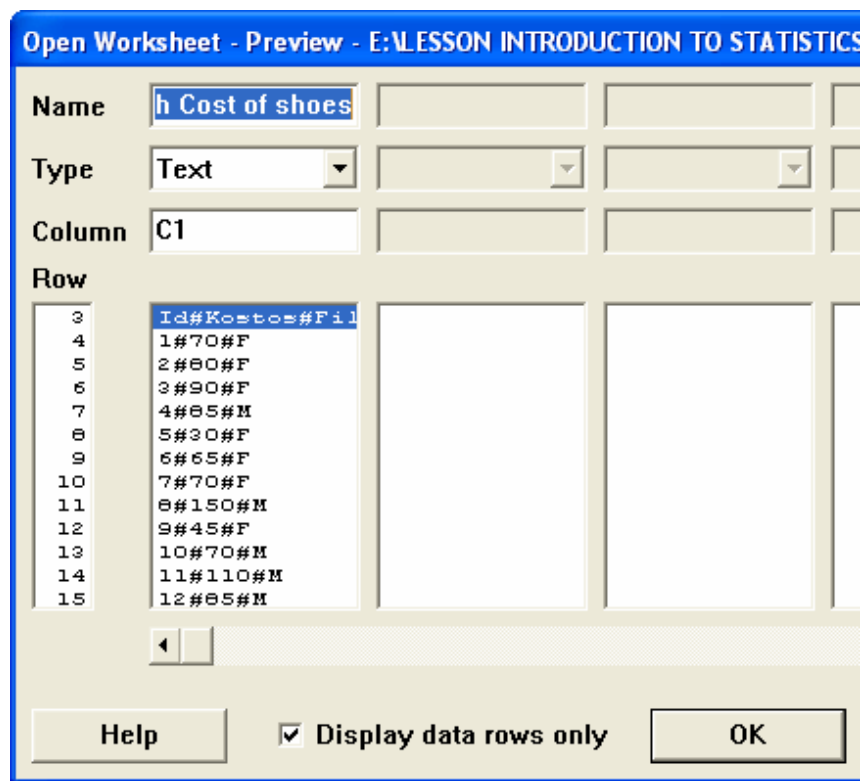
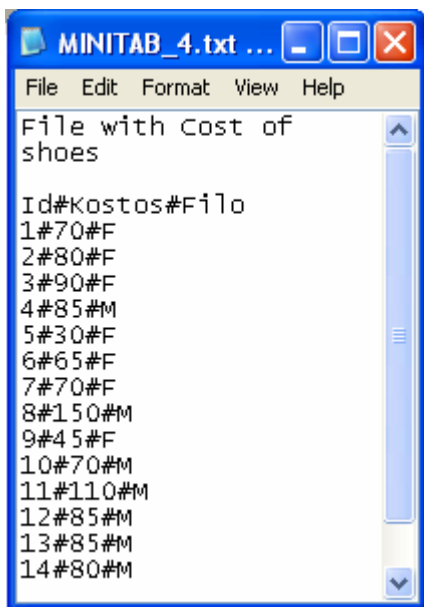
- Tab delimited (οι στήλες χωρίζονται από την παράγραφο)
- Τα ονόματα των στηλών βρίσκονται στην 1^η γραμμή



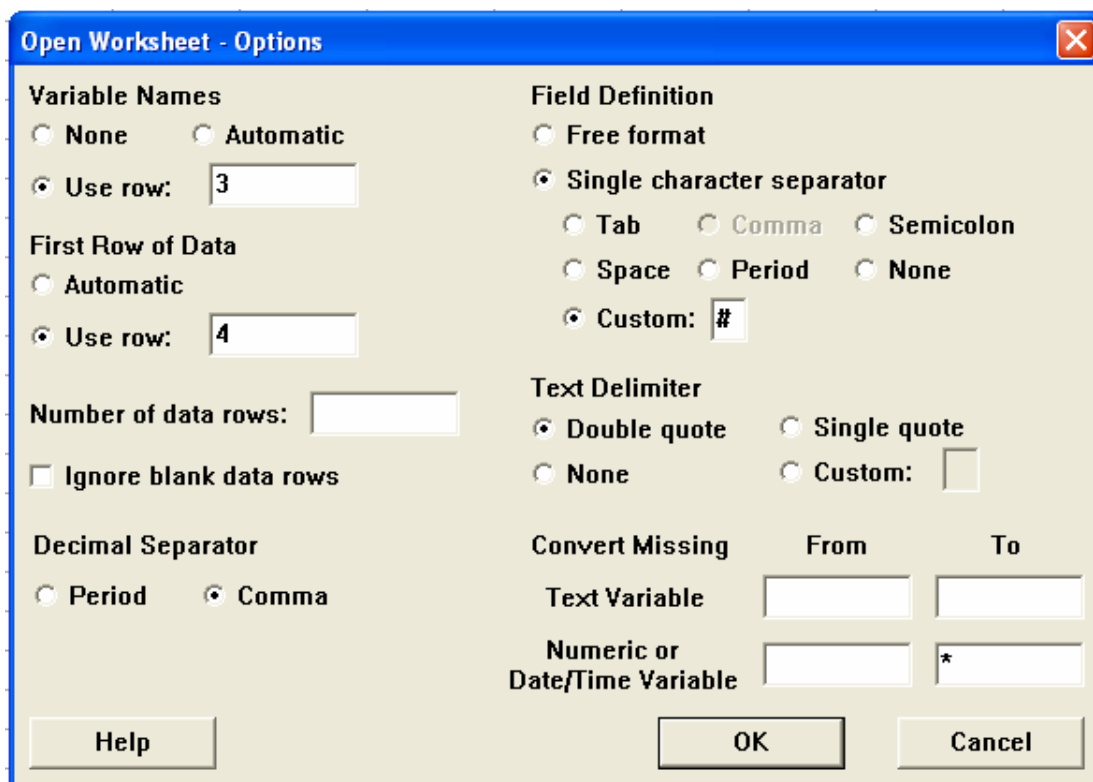
- Με χρήση της επιλογής File > Open Worksheet (ή File)

Θεωρούμε ότι οι πληροφορίες που μας δόθηκαν σε προηγούμενο παράδειγμα σχετικά με την αγορά παπουτσιών έχουν αποθηκευθεί σε ένα αρχείο κειμένου με την ονομασία MIINITAB_4.txt.

Παρατηρούμε ότι στο πάνω μέρος του κειμένου αναφέρονται κάποιες πληροφορίες για το κείμενο, ενώ οι μεταβλητές χωρίζονται από το σύμβολο #.




Για να γίνει ορθά η ανάγνωση του αρχείου στην καρτέλα Options κάνουμε τις εξής ρυθμίσεις:



- Στο πεδίο Variable Names ,κάνουμε χρήση της επιλογής να δηλώσουμε εμείς από ποια γραμμή θα γίνει η ανάγνωση των ονομάτων των μεταβλητών (Row 3)

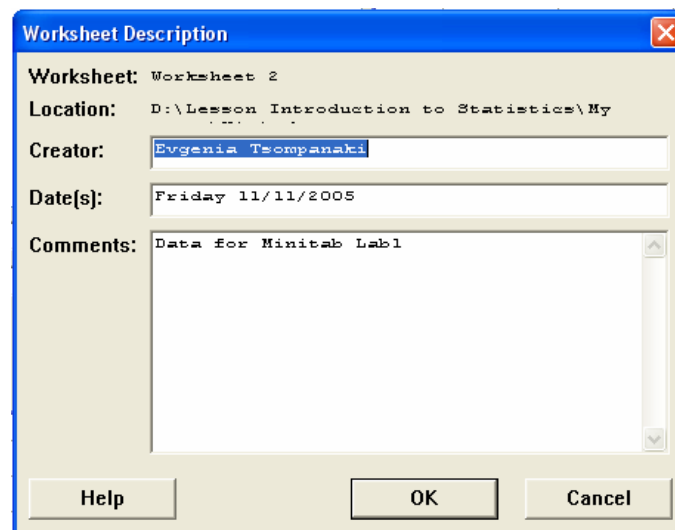
- Στο πεδίο First Row of Data , κάνουμε χρήση της επιλογής να δηλώσουμε εμείς από ποια γραμμή θα γίνει η ανάγνωση των δεδομένων (Row 4)
- Στο πεδίο Single character separator , κάνουμε χρήση της επιλογής να δηλώσουμε εμείς το σύμβολο που χωρίζει τις στήλες των δεδομένων(#)

❖ Έλεγχος στοιχείων/Αρχείου

Ένας τρόπος να ελέγξουμε τα στοιχεία μας είναι μεγιστοποιώντας το παράθυρο πληροφοριών .

Θα πρέπει να σημειώσουμε ότι μπορούμε να αποθηκεύσουμε το παράθυρο πληροφοριών Info επιλέγοντας **File/Save Info Window As** .

Επίσης χρήσιμη είναι η δυνατότητα συνοπτικής περιγραφής του αρχείου δεδομένων, μέσω της επιλογής Editor/Worksheet Description.

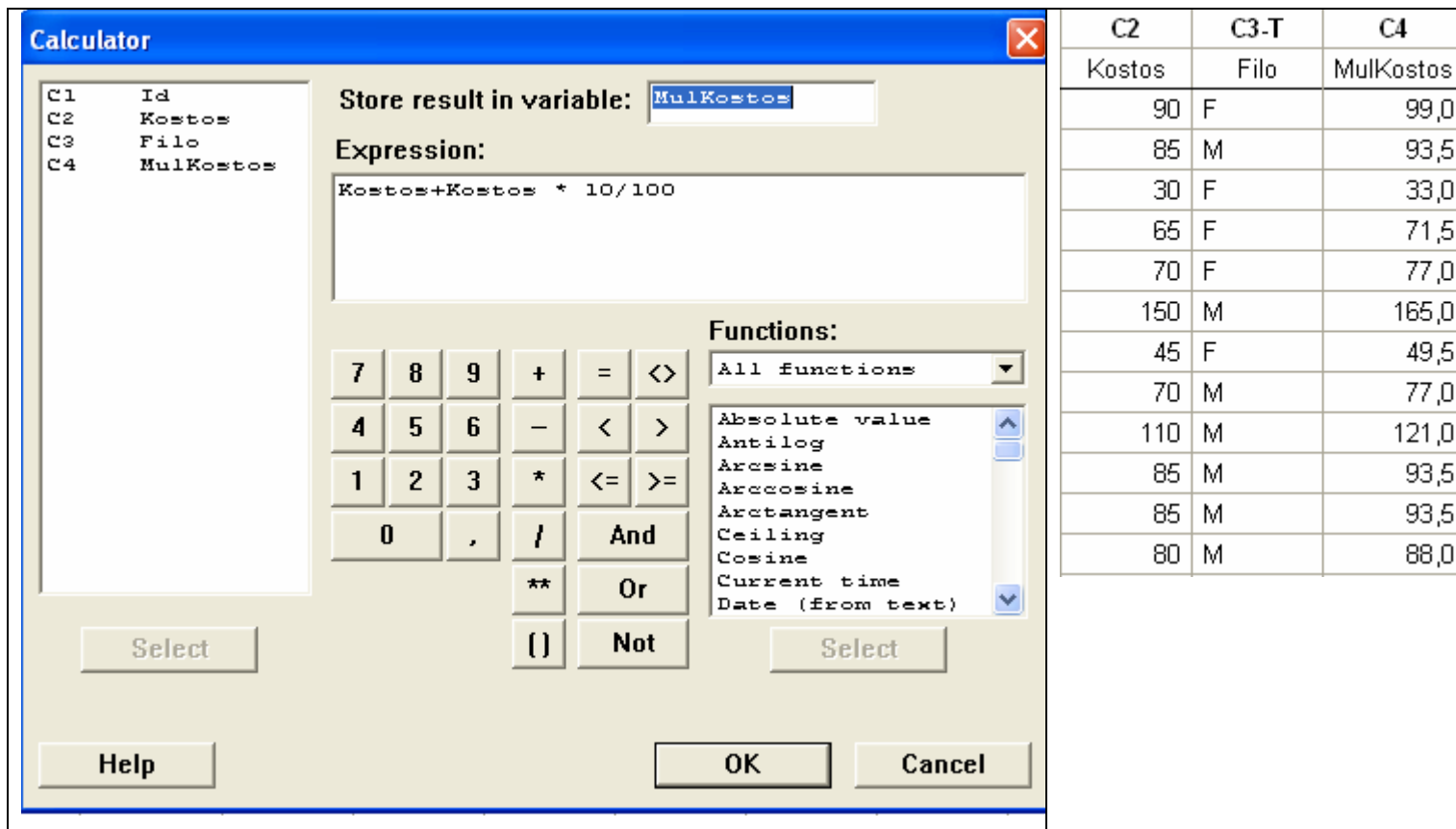


Η περιγραφή του αρχείου δεδομένων μπορεί να γίνει και κατά την αποθήκευση του αρχείου, όπως θα δούμε στην ενότητα Αποθήκευση.

❖ Δημιουργία μεταβλητών

Μπορούμε να δημιουργήσουμε ή να μετασχηματίσουμε μεταβλητές στο Minitab μέσω της επιλογής **Calk**.

Αν για τα δεδομένα της αγοράς παπουτσιών υποθέσουμε ότι την επόμενη χρονιά υπάρχει μία αύξηση της τιμής των παπουτσιών κατά 10%, τότε στη μεταβλητή MulKostos εμφανίζονται οι υπολογισμένες μέσω της σχέσης $MultKostos=Kostos+Kostos*10/100$, οι νέες τιμές



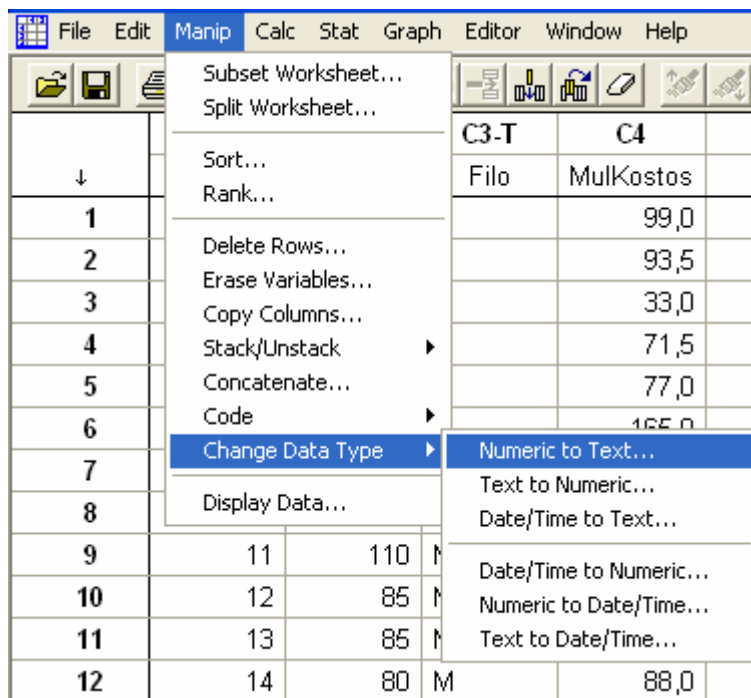
The image shows the Minitab Calculator dialog box and a data table. The calculator window is titled "Calculator" and has a list of variables on the left: C1 Id, C2 Kostos, C3 Filo, and C4 MulKostos. The "Store result in variable:" field is set to "MulKostos". The "Expression:" field contains the formula "Kostos+Kostos * 10/100". The "Functions:" list is open, showing various mathematical functions. The "Select" button is highlighted. The data table to the right has three columns: C2, C3-T, and C4. The rows represent individual data points with values for Kostos, Filo, and MulKostos.

C2	C3-T	C4
Kostos	Filo	MulKostos
90	F	99,0
85	M	93,5
30	F	33,0
65	F	71,5
70	F	77,0
150	M	165,0
45	F	49,5
70	M	77,0
110	M	121,0
85	M	93,5
85	M	93,5
80	M	88,0

❖ Χειρισμός μεταβλητών

Αλλαγή τύπου

Μπορούμε να αλλάξουμε τους τύπους μεταβλητών μέσω των ακόλουθων μετασχηματισμών, εφόσον αυτό έχει έννοια.



Δεν έχει νόημα να προσπαθήσουμε να αλλάξουμε μία μεταβλητή η οποία περιέχει ονόματα ανθρώπων σε αριθμητικές τιμές. Θα μπορούσαμε όμως να είχαμε μία μεταβλητή με αριθμούς η οποία κατά λάθος εισήχθη ως κείμενο και με τον κατάλληλο χειρισμό να την ορίσουμε ως αριθμητική.

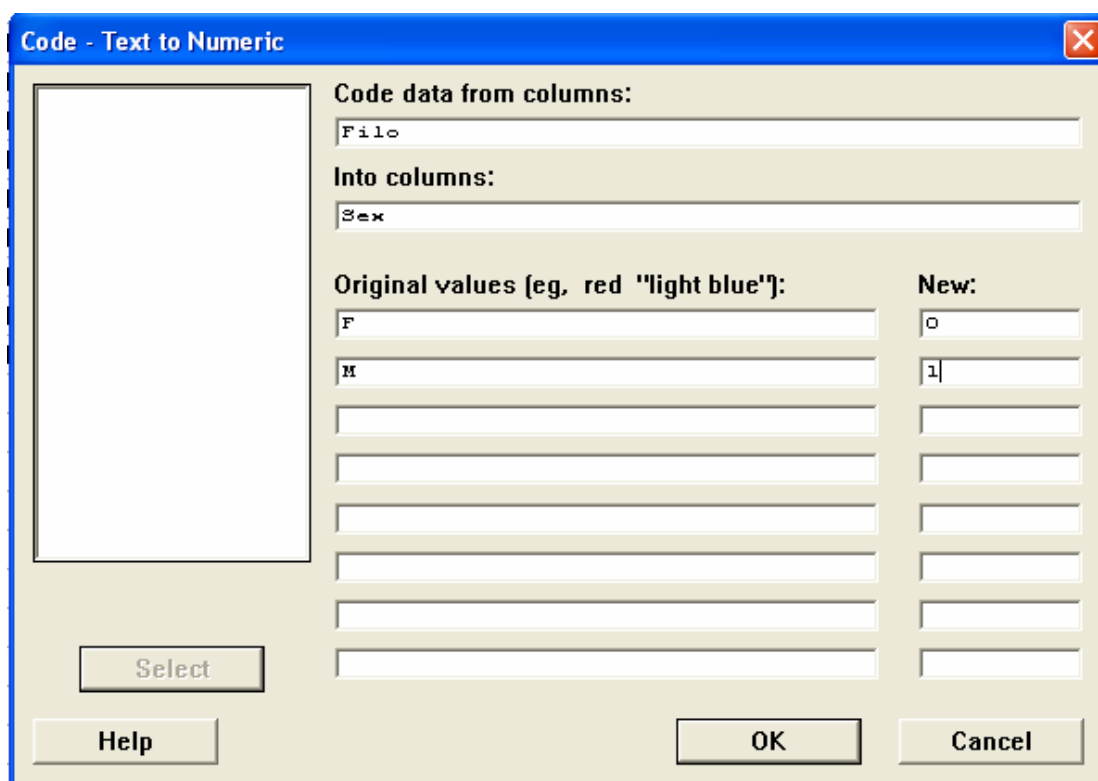
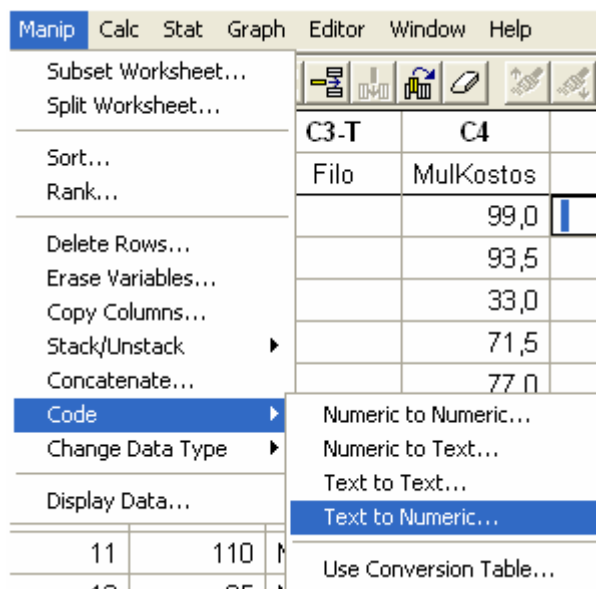
Κωδικοποίηση

Ανάλυση σε οποιοδήποτε αρχείο δεδομένων μπορεί να γίνει μόνο αν τα στοιχεία έχουν δηλωθεί στο αρχείο με τη μορφή αριθμών (τιμές ή κωδικοί).

Στο παράδειγμα μας δεν είναι δυνατόν να πάρουμε κάποια πληροφορία για τη μεταβλητή φύλο των αγοραστών αν αυτή απλά παραμείνει ως Filo με τιμές F για τις γυναίκες και M για τους άνδρες.

Για το λόγο αυτό θα πρέπει να δημιουργήσουμε μία μεταβλητή, ας πούμε τη μεταβλητή Sex κάνοντας την ακόλουθη αντιστοίχιση

Τιμή	Κωδικός
F	0
M	1



❖ Επιλογή δεδομένων

Είναι δυνατό κάποιες φορές να θέλουμε να προχωρήσουμε στην ανάλυση μόνο κάποιων από τα στοιχεία ενός συνόλου δεδομένων τα οποία ικανοποιούν μία συγκεκριμένη συνθήκη. Για παράδειγμα θα θέλαμε ανάλυση μόνο για τις γυναίκες φοιτήτριες ή για τους φοιτητές οι οποίοι ανήκουν σε μία συγκεκριμένη ηλικιακή ή οικονομική τάξη.

Για να γίνει αυτό αφού ενεργοποιήσουμε το Data window στο οποίο θέλουμε να κάνουμε την επιλογή επιλέγουμε Manip <Subset Worksheet

Στο παράθυρο διαλόγου που ανοίγει προσδιορίζουμε στην επιλογή Condition

τις τιμές στις οποίες θέλουμε να εστιάσουμε την προσοχή μας.

Στο δικό μας παράδειγμα έστω ότι θέλουμε να συλλέξουμε πληροφορίες μόνο για τα παπούτσια τα οποία κοστίζουν λιγότερο από 100euro. Για το λόγο αυτό ορίζουμε ως συνθήκη $Kostos \leq 100$.

Subset Worksheet

Name of the New Worksheet

Name:

Include or Exclude

Specify which rows to include

Specify which rows to exclude

Specify Which Rows to Include

Rows that match **Kostos <= 100**

Brushed rows

Row numbers:

Με αυτή τη διαδικασία δημιουργούμε το παρακάτω υποσύνολο των αρχικών δεδομένων σύμφωνα με τη συνθήκη που θέσαμε.

Subset of Minitab.MTW ***					
	C1	C2	C3-T	C4	C5
↓	Id	Kostos	Filo	MulKostos	Sex
1	3	90	F	99,0	0
2	4	85	M	93,5	1
3	5	30	F	33,0	0
4	6	65	F	71,5	0
5	7	70	F	77,0	0
6	9	45	F	49,5	0
7	10	70	M	77,0	1
8	12	85	M	93,5	1
9	13	85	M	93,5	1
10	14	80	M	88,0	1
11					

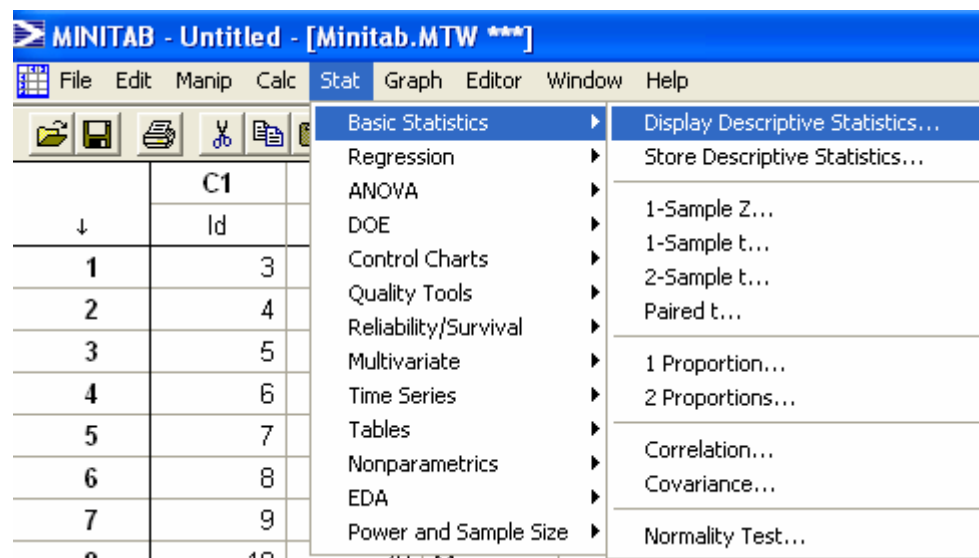
❖ Στατιστική Ανάλυση

Αρχικά θα ασχοληθούμε με βασική στατιστική ανάλυση των στοιχείων. Η έμφαση δίνεται στη χρήση του στατιστικού πακέτου για την απόκτηση των αποτελεσμάτων και όχι στην ερμηνεία των εννοιών.

Για να προβούμε σε βασική στατιστική ανάλυση, κάνουμε κλικ στη λέξη **STAT** στο μενού εντολών.

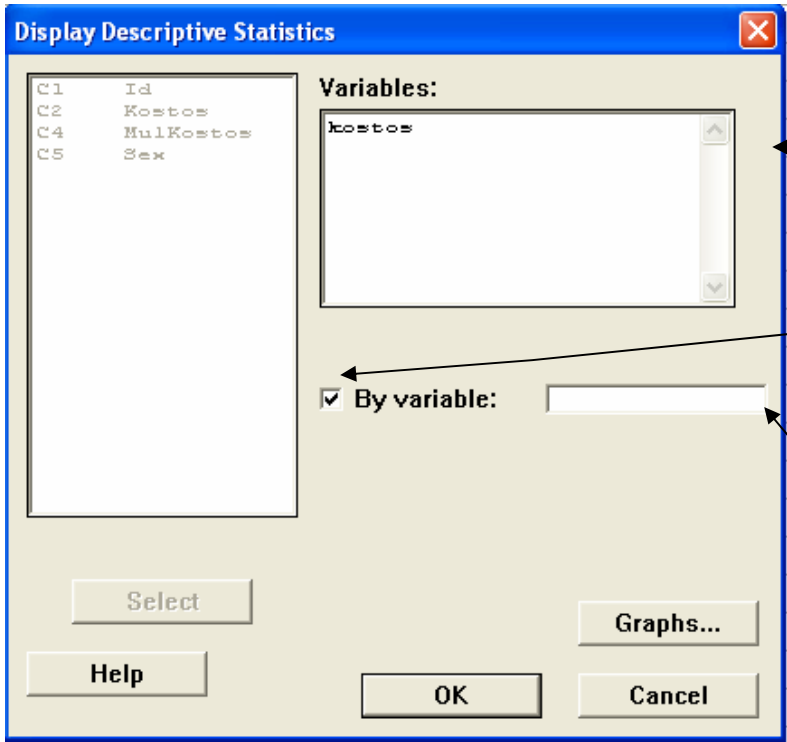
Περιγραφική στατιστική

Στο κυλιόμενο μενού επιλέγουμε **Basic Statistics**, και έπειτα **Display Descriptive Statistics** (Στο μέλλον, μια τέτοια ακολουθία επιλογών θα δηλώνουμε ως **Stat > Basic Statistics > Display Descriptive Statistics**).



Οι επιλογές μας αυτές θα μας δώσουν ένα παράθυρο περιγραφικών στατιστικών με το δρομέα να αναβοσβήνει στο ορθογώνιο πλαίσιο των μεταβλητών.

Παράθυρο εμφάνισης όλων των μεταβλητών του αρχείου



Παράθυρο επιλεγμένων για ανάλυση μεταβλητών. Επιλογή για να μας δοθεί η δυνατότητα εμφάνισης αποτελεσμάτων χωριστά για κάθε τιμή της μεταβλητής που θα δηλωθεί στο παράθυρο

Παρατηρούμε ότι μόνο οι μεταβλητές με αριθμητικές τιμές εμφανίζονται (η μεταβλητή χαρακτήρων Filo δεν εμφανίζεται). Μεταφέρουμε τη μεταβλητή/ες που επιθυμούμε να αναλύσουμε στο δεξί ορθογώνιο πλαίσιο των μεταβλητών. Ο ευκολότερος τρόπος να γίνει αυτό είναι να κάνουμε δύο φορές κλικ στη μεταβλητή kostos στο αριστερό πλαίσιο. Εάν αυτό γίνει, η μεταβλητή kostos θα εμφανιστεί τώρα και στο ορθογώνιο πλαίσιο των μεταβλητών. (Μπορούμε να πληκτρολογήσουμε απλά τη λέξη kostos εκεί όπου αναβοσβήνει ο δρομέας αλλά ο τρόπος αυτός δεν συνιστάται.) Διαφορετικά μπορούμε απλά να επιλέξουμε τη μεταβλητή kostos κάνοντας κλικ επάνω της και κατόπιν να πατήσουμε το κουμπί Select στο παράθυρο επιλογών.

Στο παράθυρο αυτό επιλογών δίνεται η δυνατότητα να ζητήσουμε και γραφήματα για τις μεταβλητές μέσω της επιλογής Graph. Τη διαδικασία δημιουργίας γραφημάτων θα τη δούμε παρακάτω. Πατώντας OK, στο Session Window, είναι τα στατιστικά αποτελέσματα.

Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Kostos	12	80,42	82,50	78,50	30,34	8,76

Variable	Minimum	Maximum	Q1	Q3
Kostos	30,00	150,00	66,25	88,75

Αν είχε ζητηθεί η επιλογή **By variable:** , τότε θα μας είχε δοθεί όλη η παραπάνω πληροφορία αναλυτικά για τις δύο τιμές της μεταβλητής sex:

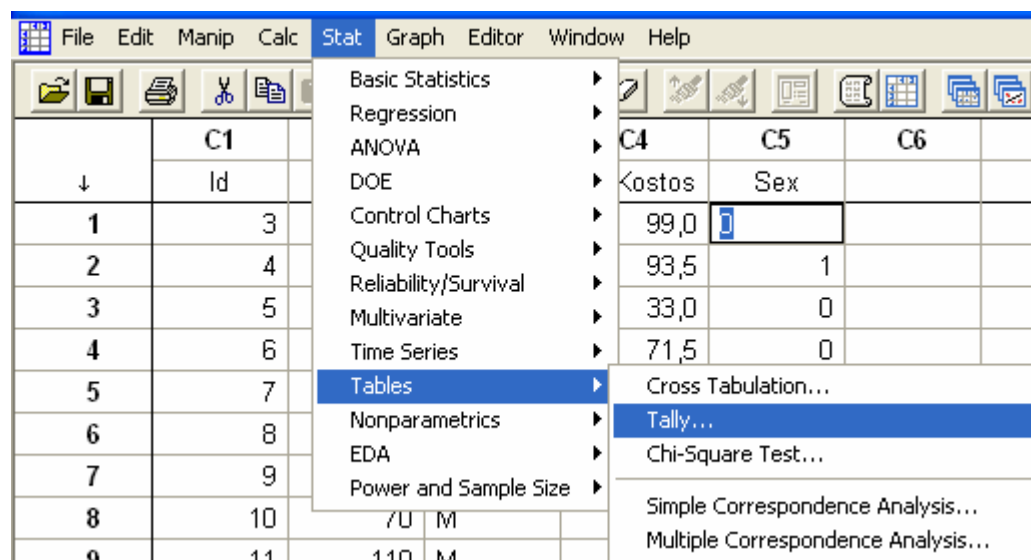
Descriptive Statistics

Variable	Sex	N	Mean	Median	TrMean	StDev
Kostos	0	5	60,0	65,0	60,0	23,2
	1	7	95,0	85,0	95,0	27,1

Variable	Sex	SE Mean	Minimum	Maximum	Q1	Q3
Kostos	0	10,4	30,0	90,0	37,5	80,0
	1	10,2	70,0	150,0	80,0	110,0

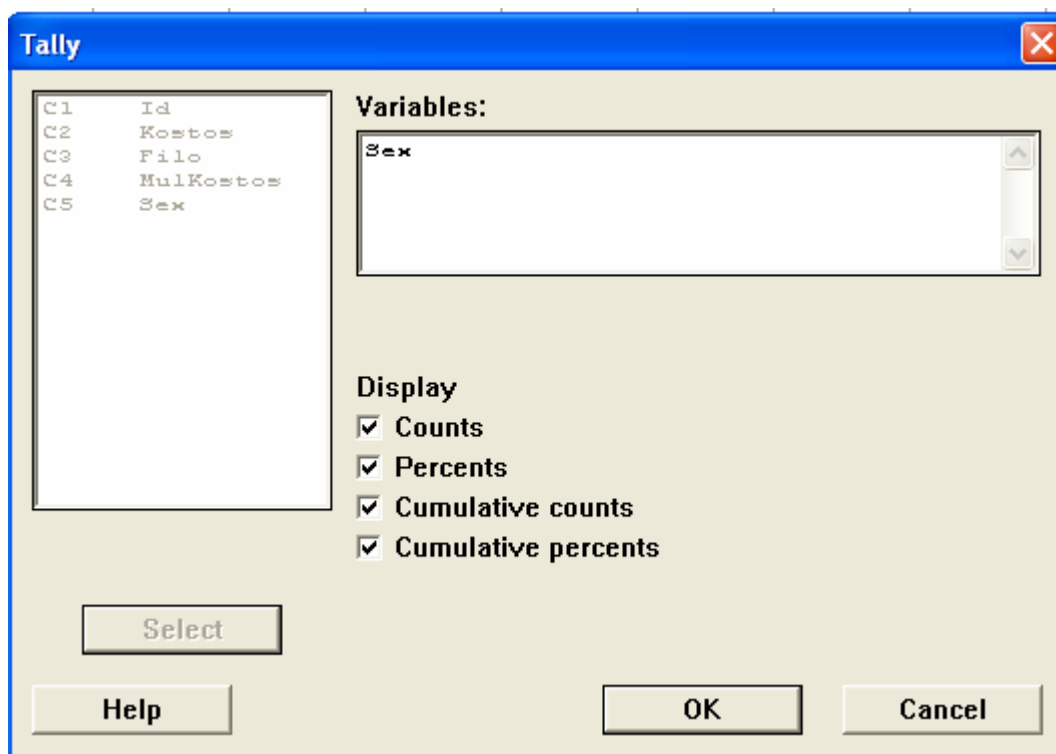
Κατανομή συχνοτήτων

Στο κυλιόμενο μενού επιλέγουμε



Μέσω αυτής της εντολής μπορούμε να ζητήσουμε από το Minitab να μας υπολογίσει συχνότητα, αθροιστική συχνότητα, σχετική συχνότητα και αθροιστική σχετική συχνότητα **Διακριτών** μεταβλητών.

Η επιλογή μας αυτή μας μεταφέρει στο παρακάτω πλαίσιο διαλόγου:



Κάνοντας διπλό κλικ μεταφέρουμε τη μεταβλητή Sex στο δεξί παράθυρο για τις προς ανάλυση μεταβλητές. Επιλέγοντας όλα τα αποτελέσματα που μπορούν μέσω αυτού του παραθύρου να μας δοθούν, ένα \checkmark στο κουτί κάθε επιλογής δηλώνει ότι έχει γίνει η επιλογή, έχουμε το παρακάτω αποτέλεσμα στο Session window:

Summary Statistics for Discrete Variables

Sex	Count	CumCnt	Percent	CumPct
0	5	5	41,67	41,67
1	7	12	58,33	100,00
N=	12			

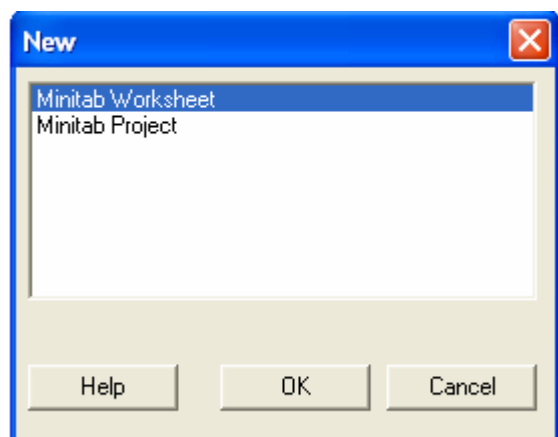
Η 1^η στήλη απλά παρουσιάζει όλες τις τιμές της μεταβλητής Sex, η 2^η στήλη δίνει τη **συχνότητα** αντίστοιχα για κάθε τιμή της μεταβλητής, η 3^η στήλη την **αθροιστική συχνότητα**, η 4^η στήλη τη **σχετική συχνότητα** ως ποσοστό και η τελευταία στήλη τη **σχετική αθροιστική συχνότητα** ως ποσοστό για τις τιμές της μεταβλητής.

Μπορούμε μέσω της διαδικασίας αντιγραφής- επικόλλησης να μεταφέρουμε αυτή την πληροφορία σε ένα νέο φύλλο εργασίας του Minitab, ώστε να δημιουργήσουμε τα απαραίτητα γραφήματα.

Για να γίνει αυτό κάνουμε ενεργό το Session κάνοντας κλικ επάνω του. Με χρήση του ποντικιού μαρκάρουμε τις δύο γραμμές αποτελεσμάτων

0	5	5	41,67	41,67
1	7	12	58,33	100,00

επιλέγουμε Edit> Copy. Για να μεταφέρουμε τα δεδομένα αυτά σε ένα νέο φύλλο εργασίας, επιλέγουμε File>New και εμφανίζεται το πλαίσιο διαλόγου



Επιλέγουμε **Minitab Worksheet** και πατάμε OK. Έτσι δημιουργείτε ένα νέο φύλλο εργασίας με τις αντίστοιχες στήλες (τα ονόματα των στηλών έχουν προστεθεί μετά την επικόλληση).

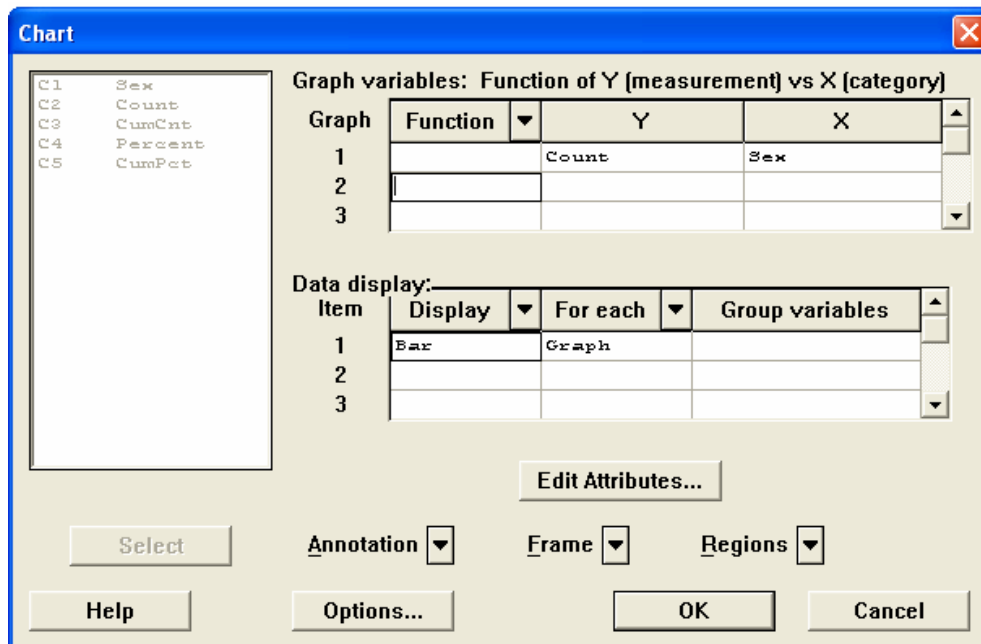
	C1	C2	C3	C4	C5
↓	Sex	Count	CumCnt	Percent	CumPct
1	0	1	1	50	50
2	1	1	2	50	100

❖ Γραφική παράσταση δεδομένων

Το Minitab μας δίνει τη δυνατότητα δημιουργίας διάφορων τύπων γραφημάτων είτε άμεσα μέσω της εντολής Graph και την κατάλληλη επιλογή από το κυλιόμενο παράθυρο ή ως επιλογή μέσω των εντολών ανάλυσης και τα παράθυρα διαλόγου.

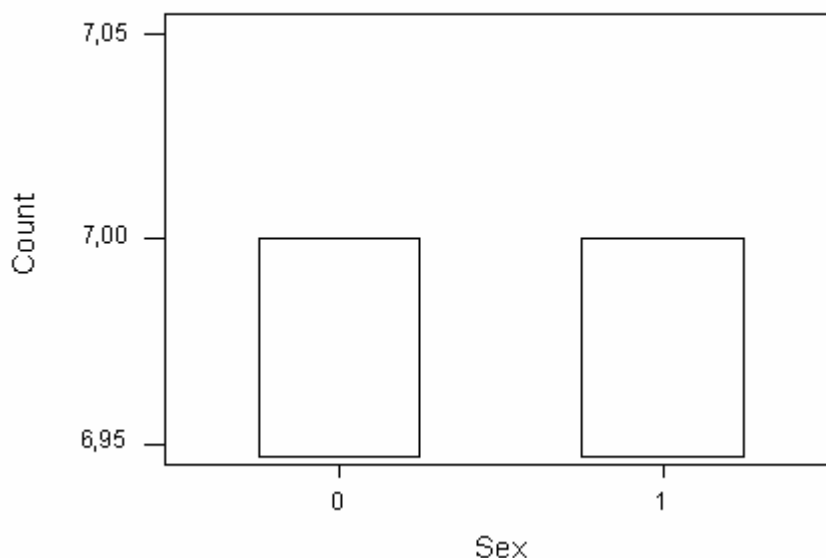
Bar graph

Για να δημιουργήσουμε ένα ραβδόγραμμα επιλέγουμε Graph < Chart και εμφανίζεται το ακόλουθο παράθυρο



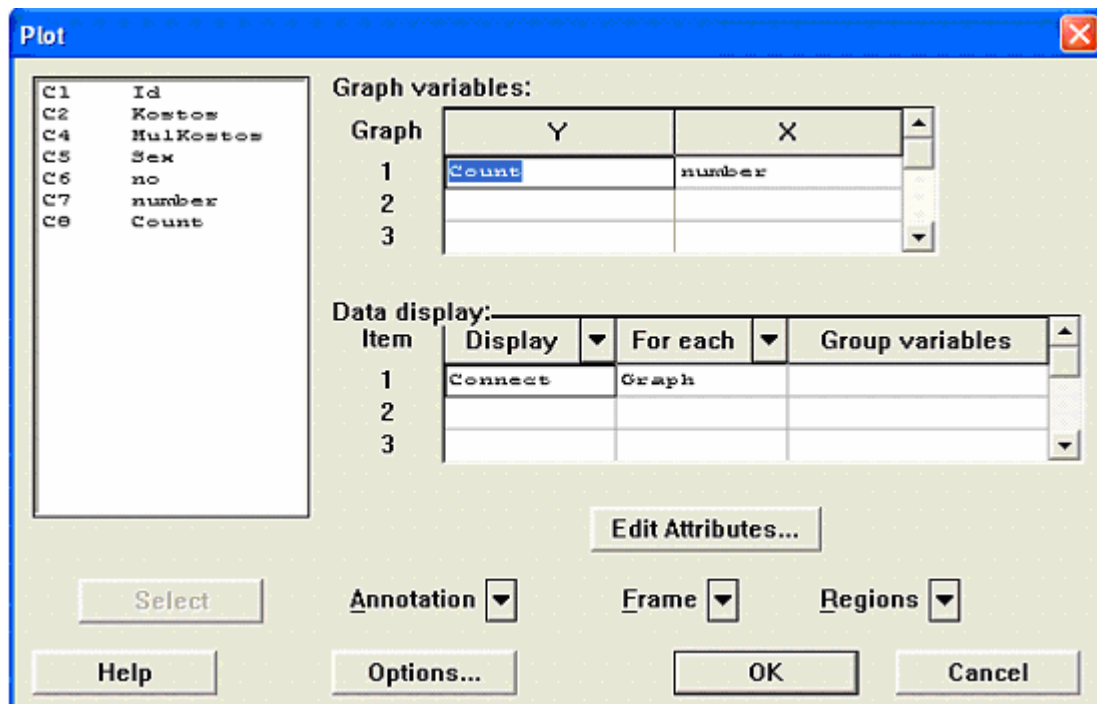
Για το παράδειγμά μας θα θέλαμε να δημιουργήσουμε ένα ραβδόγραμμα της μεταβλητής sex, για την οποία σε προηγούμενη ενότητα παρουσιάσαμε πως μπορούμε να αποθηκεύσουμε την κατανομή συχνοτήτων της (βλ. Κατανομή συχνοτήτων).

Συμπληρώνουμε τα πεδία όπως φαίνεται στο παράθυρο και δημιουργείται το παρακάτω γράφημα:

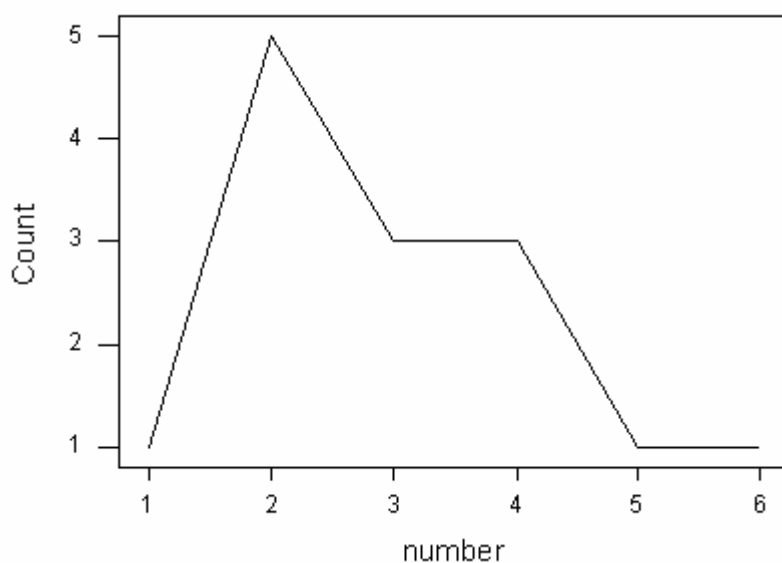


Frequency polygon

Για να δημιουργήσουμε ένα πολύγωνο συχνοτήτων επιλέγουμε Graph < Plot και εμφανίζεται το ακόλουθο παράθυρο

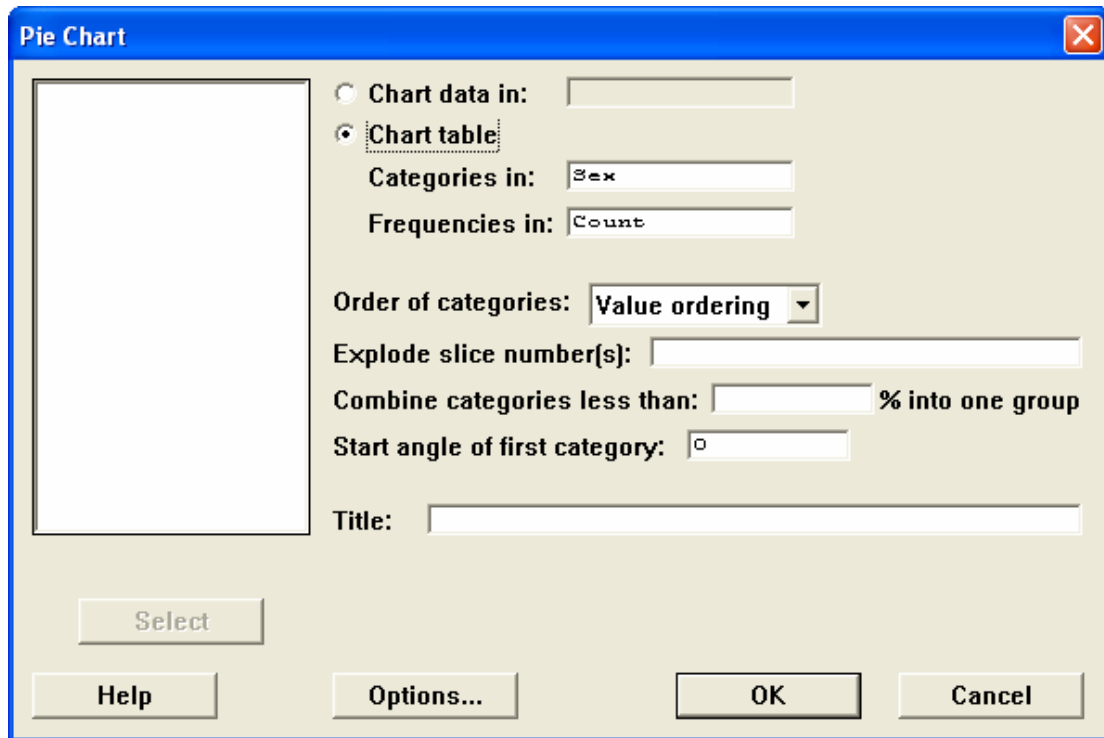


όπου συμπληρώνουμε στο πεδίο της Y μεταβλητής τη συχνότητα που έχουμε υπολογίσει και αποθηκεύσει για τη μεταβλητή no, την οποία στο πεδίο X αναφέρουμε ως number. Το γράφημα που προκύπτει είναι:

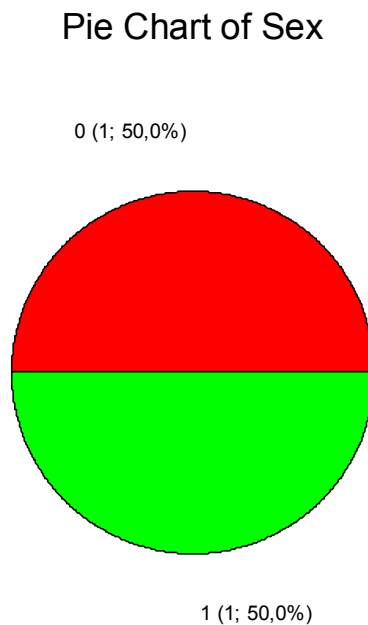


Pie chart

Για να δημιουργήσουμε ένα κυκλικό γράφημα, επιλέγουμε Graph < Pie Chart και εμφανίζεται το ακόλουθο παράθυρο

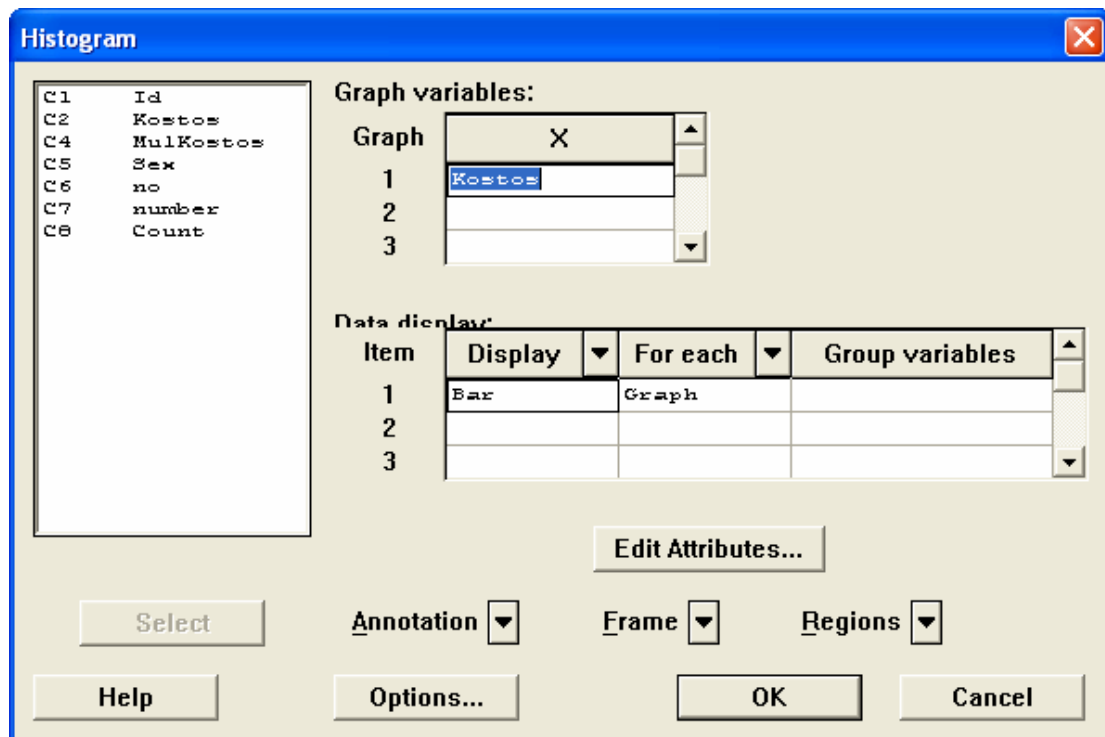


Συμπληρώνουμε τα πεδία όπως φαίνεται στο παράθυρο και δημιουργείται το παρακάτω γράφημα:

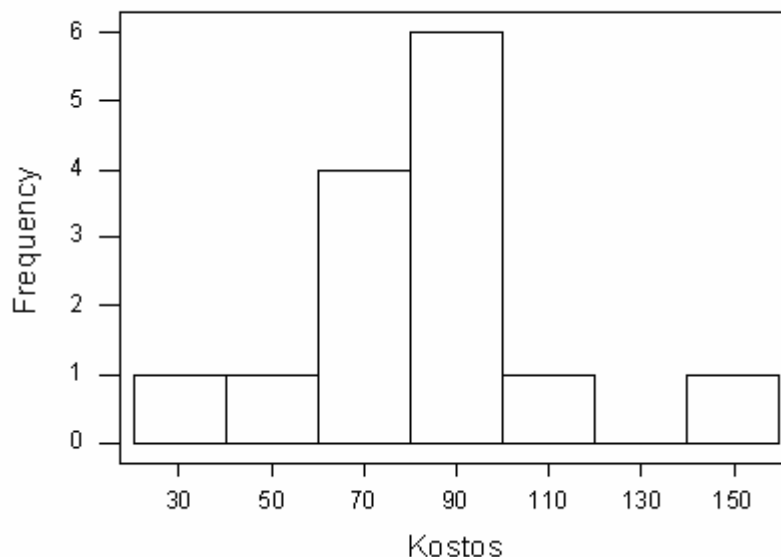


Histogram

Για να δημιουργήσουμε ένα ιστόγραμμα, επιλέγουμε Graph < Histogram και εμφανίζεται το ακόλουθο παράθυρο

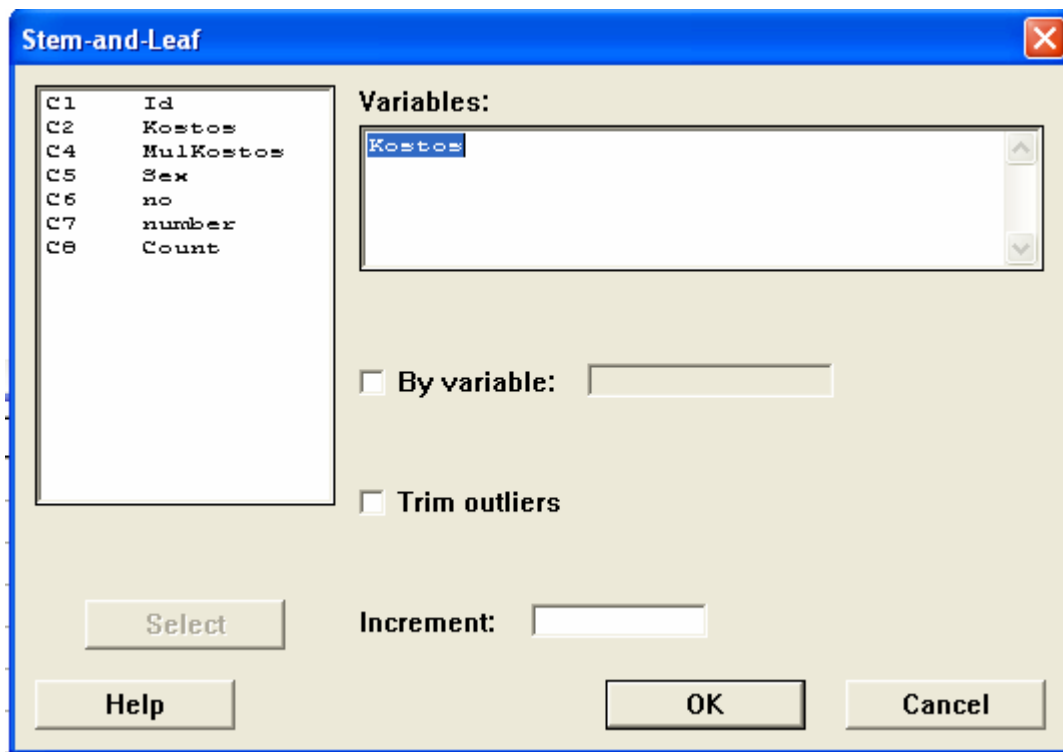


Συμπληρώνουμε τα πεδία όπως φαίνεται στο παράθυρο και δημιουργείται το παρακάτω γράφημα:

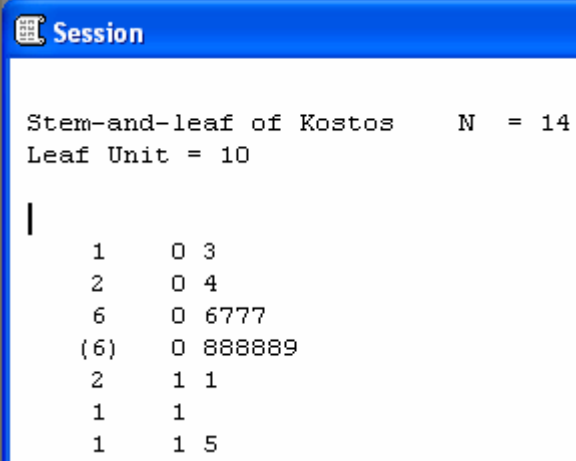


Stem and Leaf

Για να δημιουργήσουμε ένα διάγραμμα μίσχου -φύλλου, επιλέγουμε Graph < Stem and Leaf και εμφανίζεται το ακόλουθο παράθυρο



Συμπληρώνουμε τα πεδία όπως φαίνεται στο παράθυρο και δημιουργείται το παρακάτω γράφημα στο παράθυρο του Session

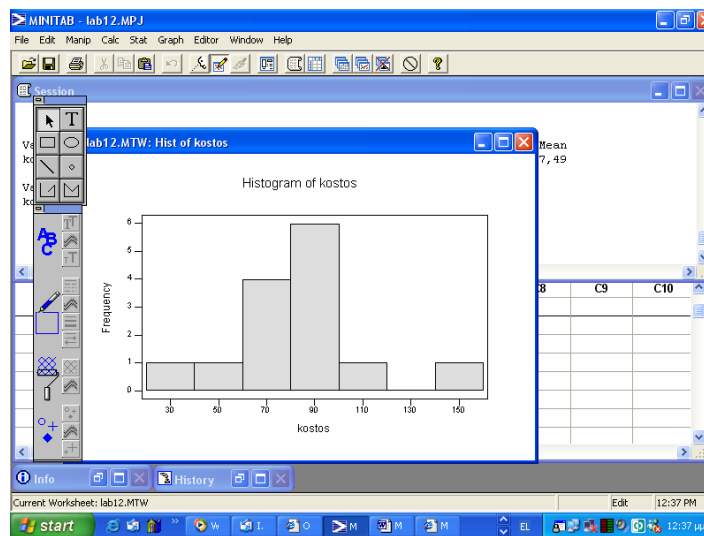


```
Session
Stem-and-leaf of Kostos    N = 14
Leaf Unit = 10

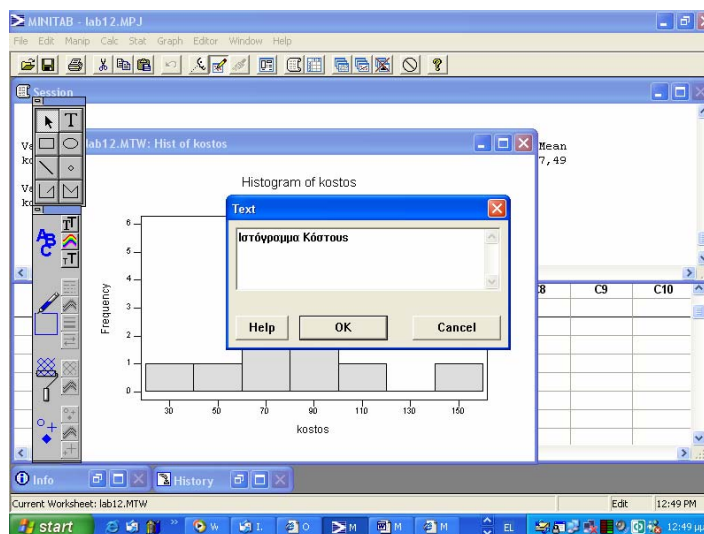
|
 1   0 3
 2   0 4
 6   0 6777
(6)  0 888889
 2   1 1
 1   1
 1   1 5
```

❖ Επεξεργασία γραφημάτων

Μετά την παραγωγή γραφήματος και την εμφάνισή του στην οθόνη μας, πιθανόν να θέλουμε να παρέμβουμε κάνοντας κάποιες τροποποιήσεις. Για να γίνει αυτό κάνουμε διπλό κλικ στο γράφημα και ενεργοποιούνται δύο παλέτες επιλογών, η παλέτα εργαλείων **Tool Palette** η οποία είναι στην κορυφή και η παλέτα χαρακτηριστικών **Attribute palette**.

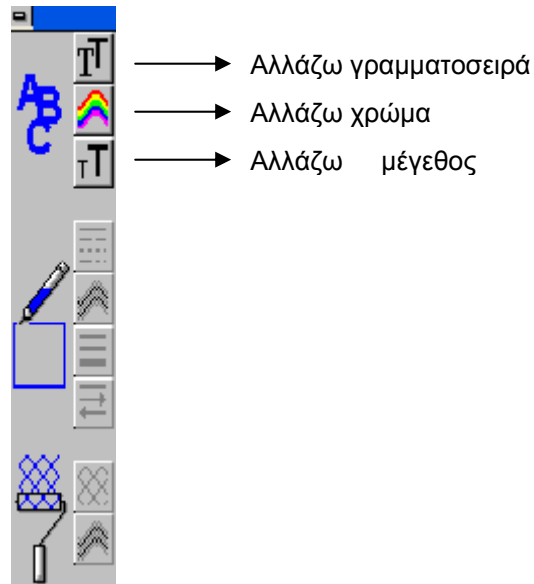


1. Για να τροποποιήσουμε ως πούμε τον τίτλο του γραφήματος
 - κάνουμε διπλό κλικ στον τίτλο και εμφανίζεται ένα μικρό παράθυρο όπου δακτυλογραφούμε το κείμενο

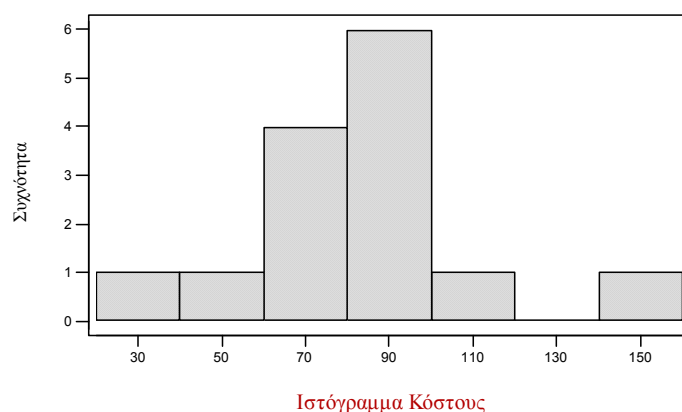


Για να τροποποιήσουμε τη γραμματοσειρά του τίτλου του γραφήματος κάνουμε κλικ στον τίτλο και εμφανίζεται ένα μικρό κουτί τετραγώνων. Στην παλέτα χαρακτηριστικών επιλέγουμε

- ο μία γραμματοσειρά ελληνικών Times new Roman Greek
- ο χρώμα για τον τίτλο
- ο μέγεθος γραμματοσειράς



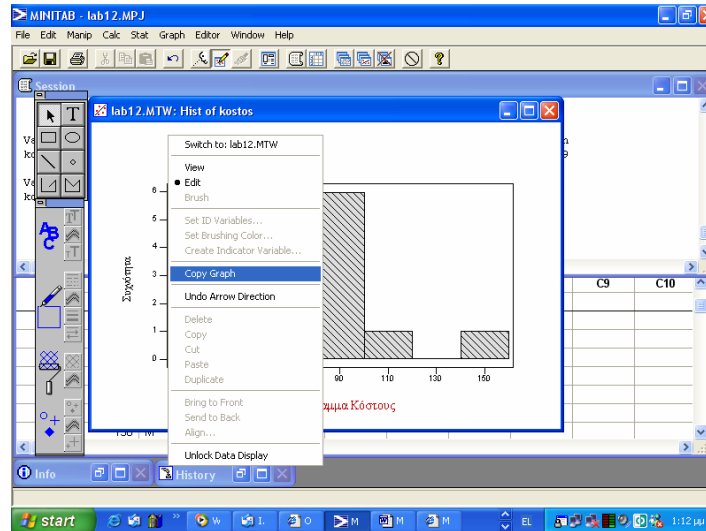
Μπορούμε έπειτα να σύρουμε τον τίτλο για να βάλουμε το κείμενο στην επιθυμητή θέση. Μετά τις αλλαγές έχουμε το Γράφημα 1 .



Γράφημα 1 Ιστόγραμμα του Κόστους των Υποδημάτων

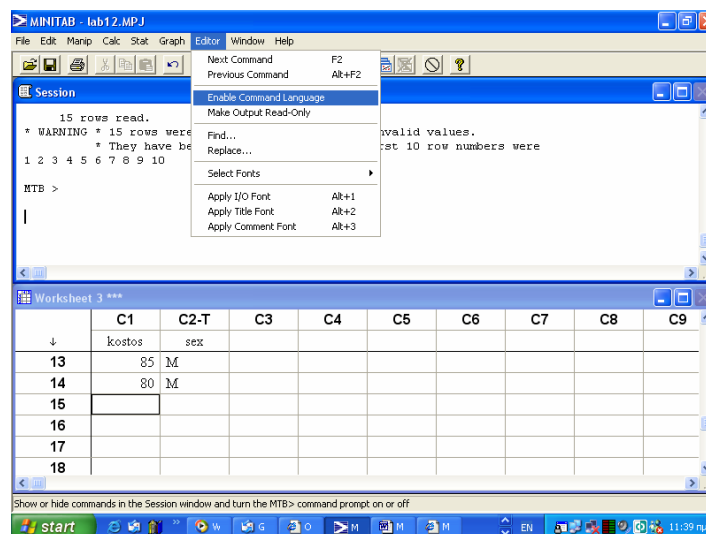
Προκειμένου να τεθεί η γραφική παράστασή μας σε ένα έγγραφο του Microsoft Word θα πρέπει αρχικά να κάνουμε ένα αντίγραφο από αυτό.

Μπορούμε να το κάνουμε αυτό κάνοντας δεξί κλικ στο γράφημα και επιλογή **Copy Graph** (αντιγραφή γραφικής παράστασης). Με τον τρόπο αυτό ένα αντίγραφο της γραφικής παράστασης αποθηκεύεται στο " clipboard " του υπολογιστή μας. Από εκεί μπορεί να επικολληθεί στο έγγραφο εκθέσεων μας-report.



❖ Χρήση Εντολών

Οι ίδιες πληροφορίες μπορούν να αποκτηθούν και με τη χρήση εντολών. Για το σκοπό αυτό θα πρέπει στο **Session** να εμφανίζεται το **Minitab prompt (MTB>)**. Η ενεργοποίησή του μπορεί να γίνει επιλέγοντας **Editor/Enable Command Language**.



Με τον τρόπο αυτό για κάθε ενέργεια που πραγματοποιούμε στο Minitab μέσω των παραθύρων εμφανίζεται η σύνταξη της αντίστοιχης **εντολής** στο Minitab. Έτσι είναι δυνατή η εκμάθηση βασικών εντολών.

Για παράδειγμα για να μας δοθούν περιγραφικά μέτρα και το ιστόγραμμα για τη μεταβλητή “Κόστος” η σύνταξη των εντολών/υποεντολών είναι:

```
MTB > Describe 'Kostos';  
SUBC>GHist.
```

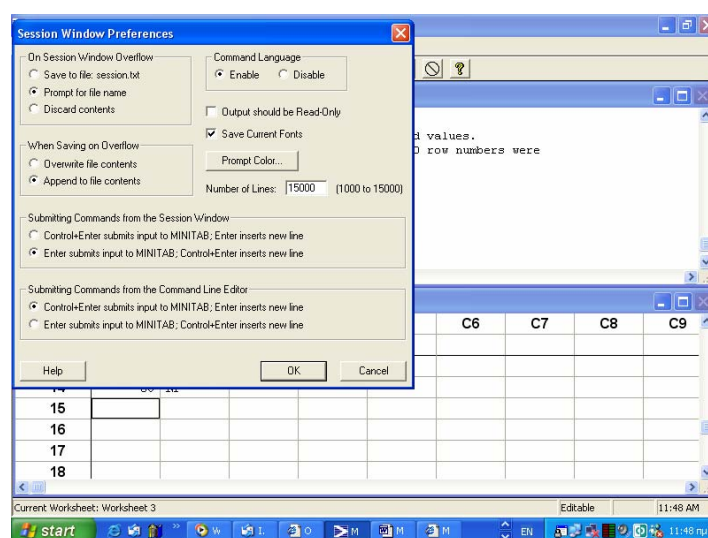
Όπως μπορούμε να παρατηρήσουμε η πρώτη γραμμή εντολών χωρίζεται από τη δεύτερη με το σύμβολο ; Με το σύμβολο αυτό λέμε στο Minitab ότι υπάρχουν και άλλα στοιχεία που θα ζητήσουμε με την εντολή αυτή.

Με το SUBC> prompt προσδιορίζουμε την υποεντολή για το συγκεκριμένο γράφημα, το ιστόγραμμα της μεταβλητής. Η τελεία . δηλώνει στο Minitab ότι έχουμε ολοκληρώσει με τις εντολές/υποεντολές.

❖ Ρυθμίσεις

Καλό είναι κάποιες ρυθμίσεις σας να τις “σώζουμε” έτσι ώστε θα είναι σε ισχύ κάθε φορά που χρησιμοποιούμε το Minitab.

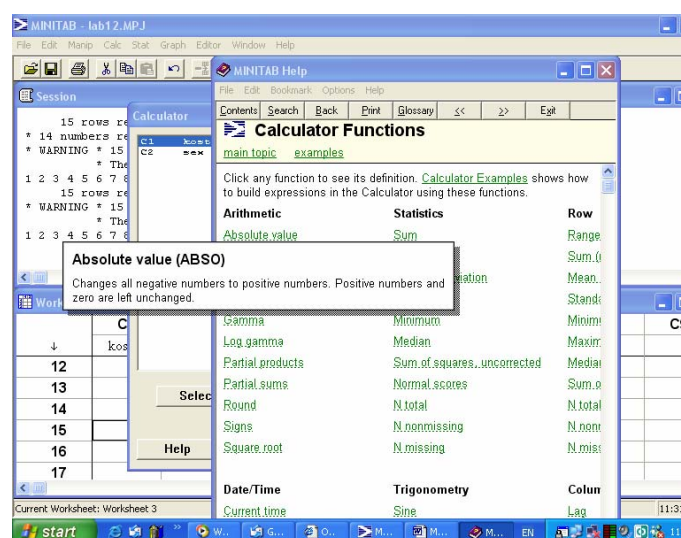
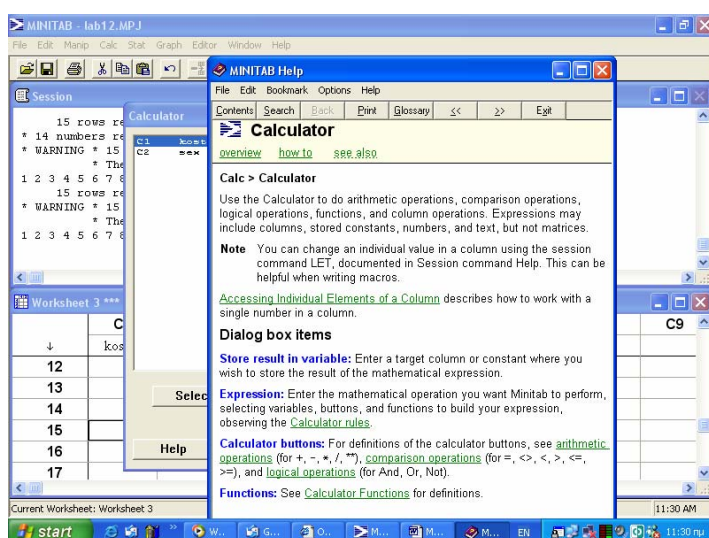
Για παράδειγμα για να είναι πάντα ενεργοποιημένο το Minitab prompt (MTB>) επιλέγουμε **Edit>Preferences>Session Window** και επιλέγουμε **Enable Command Language**, όπως φαίνεται στο παράθυρο.



Μπορούμε να αλλάξουμε την προτίμησή μας οποιαδήποτε στιγμή ακολουθώντας τα ίδια βήματα.

❖ Βοήθεια

Θα πρέπει να τονιστεί ιδιαίτερα ότι είναι πολύ σημαντική η χρήση της βοήθειας **Help** κάθε προγράμματος που χρησιμοποιούμε. Το Minitab παρέχει σημαντική βοήθεια σε κάθε επιλεγόμενο παράθυρο. Ας δούμε στο συγκεκριμένο μενού επιλογών τη βοήθεια που θα μπορούσαμε να έχουμε για παράδειγμα στη χρήση της εντολής Calc:

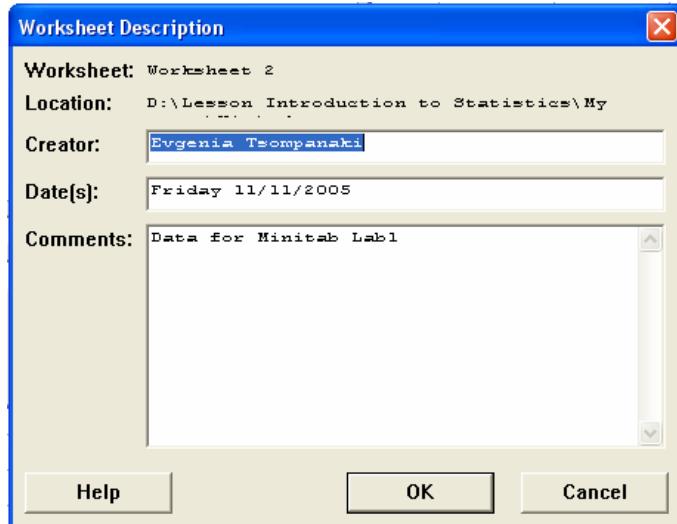
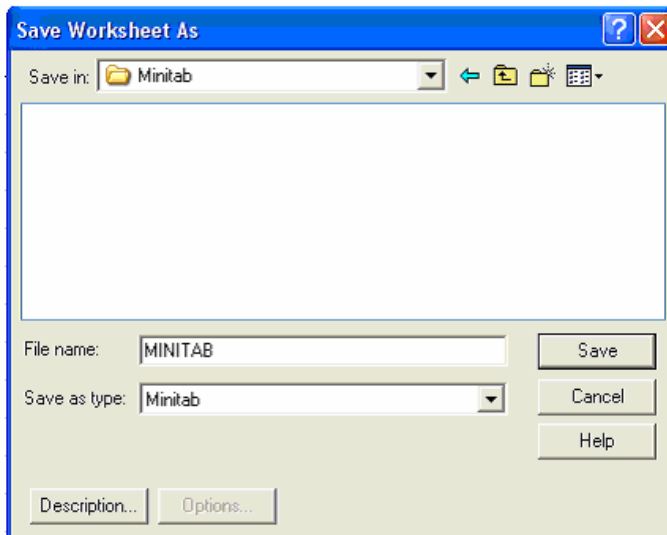


❖ Αποθήκευση

Για να αποθηκεύσουμε τη δουλειά μας στο Minitab υπάρχουν δύο επιλογές

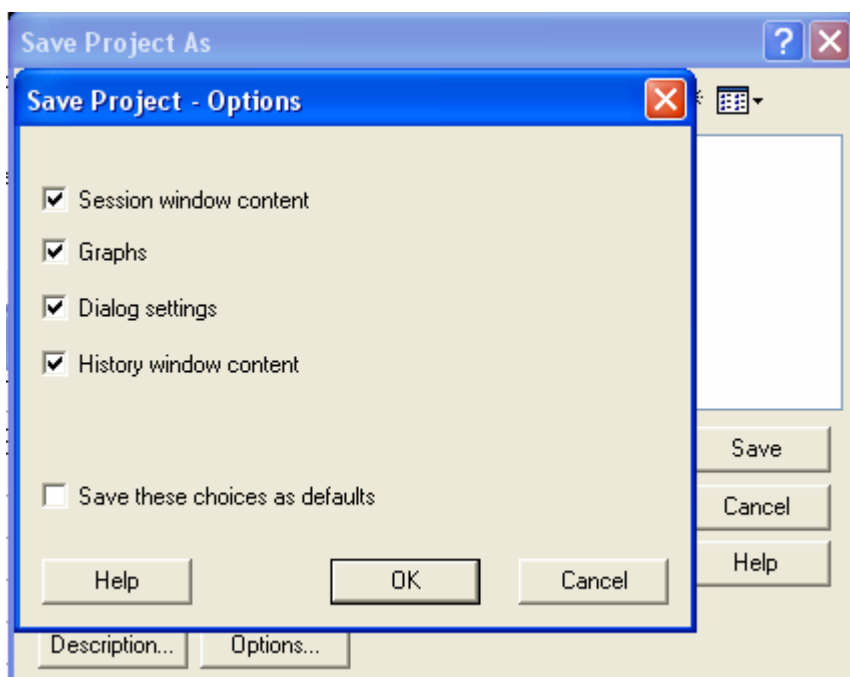
- **File > Save Current Worksheet As** με την οποία απλά αποθηκεύουμε τα δεδομένα μας σε ένα αρχείο του Minitab.

Η αποθήκευση του αρχείου μπορεί να συνοδεύεται με δήλωση και στοιχείων για την περιγραφή του αρχείου δεδομένων, μέσω της επιλογής **Description**

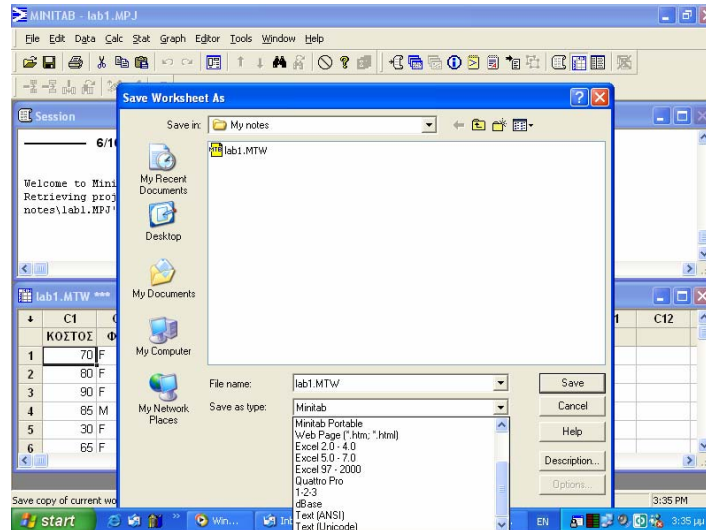


- **File > Save Project As**, με την οποία αποθηκεύονται επίσης τα γραφήματα και τα αποτελέσματα του Session.

Μέσω της επιλογής **Options** μπορούμε να επιλέξουμε τα γραφήματα, τα αποτελέσματα του Session, οι ρυθμίσεων που έχουμε επιλέξει και το παραθύρου History να αποτελούν αποθηκεύονται κάθε φορά που αποθηκεύουμε ένα project.



Υπάρχει η δυνατότητα να αποθηκεύσουμε το αρχείο των δεδομένων μας και ως άλλου τύπου αρχείο με την κατάλληλη επιλογή του τύπου αρχείου στο πλαίσιο των αρχείων καταλόγων (δείτε το πλαίσιο διαλόγου), ώστε να είναι δυνατή η άμεση επεξεργασία του με άλλα προγράμματα.



❖ Εκτύπωση

Μπορούμε να τυπώσουμε οποιοδήποτε παράθυρο με τη μετάβαση στις επιλογές αρχείων **File** και την επιλογή της **Print**. Αυτό ισχύει για Microsoft Word, Minitab, και για οποιοδήποτε πρόγραμμα λογισμικού εάν χρησιμοποιούμε το λειτουργικό σύστημα παραθύρων της Microsoft.

**© Εφαρμογή με χρήση του Στατιστικού πακέτου
Minitab 12.0**

Περιγραφή Δεδομένων

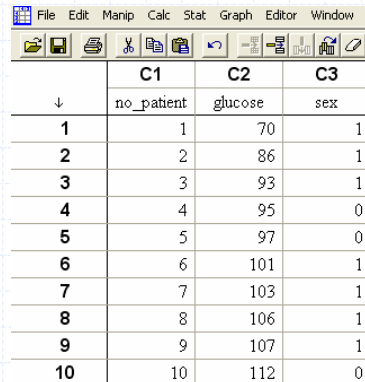
Σε μία έρευνα μετρήθηκε η ποσότητα γλυκόζης (mg%) στο αίμα εκατό ατόμων για τη διερεύνηση της ύπαρξης διαβήτη ενώ καταγράφηκε και το φύλο του ασθενούς

Δεδομένα

A/A	Γλυκόζη (mg%)	Φύλο	A/A	Γλυκόζη (mg%)	Φύλο	A/A	Γλυκόζη (mg%)	Φύλο	A/A	Γλυκόζη (mg%)	Φύλο
1	70	Γ	26	101	A	51	85	Γ	76	102	A
2	86	Γ	27	103	Γ	52	90	A	77	106	A
3	93	Γ	28	106	A	53	94	Γ	78	107	A
4	95	A	29	110	Γ	54	96	A	79	111	Γ
5	97	A	30	116	A	55	98	A	80	121	Γ
6	101	Γ	31	83	A	56	101	Γ	81	86	Γ
7	103	Γ	32	88	Γ	57	104	A	82	92	A
8	106	Γ	33	93	A	58	106	Γ	83	94	A
9	107	Γ	34	96	Γ	59	111	A	84	97	A
10	112	A	35	98	A	60	116	Γ	85	100	A
11	79	Γ	36	101	Γ	61	85	A	86	102	Γ
12	87	Γ	37	104	Γ	62	91	Γ	87	106	Γ
13	93	A	38	106	A	63	94	A	88	107	Γ
14	95	A	39	111	A	64	96	Γ	89	112	A
15	97	A	40	116	A	65	98	Γ	90	121	A
16	101	A	41	85	A	66	102	A	91	86	Γ
17	103	Γ	42	89	Γ	67	105	A	92	92	A
18	106	Γ	43	94	Γ	68	106	Γ	93	94	A
19	108	Γ	44	96	Γ	69	111	Γ	94	97	Γ
20	115	A	45	98	A	70	119	A	95	100	Γ
21	80	Γ	46	101	A	71	85	Γ	96	103	A
22	87	A	47	104	A	72	91	A	97	106	A
23	93	Γ	48	106	A	73	94	Γ	98	107	A
24	96	Γ	49	111	Γ	74	97	Γ	99	112	Γ
25	98	Γ	50	116	Γ	75	98	Γ	100	126	Γ

1. Οργάνωση Δεδομένων Ονόματα Μεταβλητών

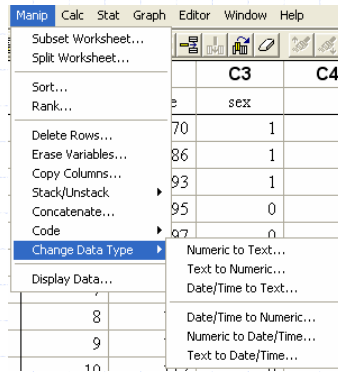
- Ονόματα μεταβλητών
 - Πληκτρολογούμε στα πεδία ανάμεσα στο πρώτο κελί και τα ονόματα των στηλών



	C1	C2	C3
↓	no_patient	glucose	sex
1	1	70	1
2	2	86	1
3	3	93	1
4	4	95	0
5	5	97	0
6	6	101	1
7	7	103	1
8	8	106	1
9	9	107	1
10	10	112	0

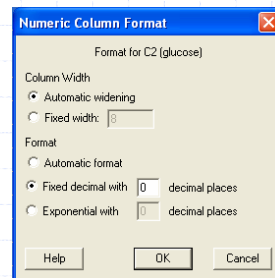
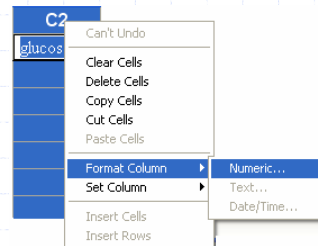
2. Οργάνωση Δεδομένων Τύπος Μεταβλητών

- Διαμόρφωση μεταβλητών
 - Διαμορφώνουμε τις μεταβλητές καθορίζοντας αν είναι απαραίτητο τον **Τύπο** των μεταβλητών



3. Οργάνωση Δεδομένων Δεκαδικά Ψηφία

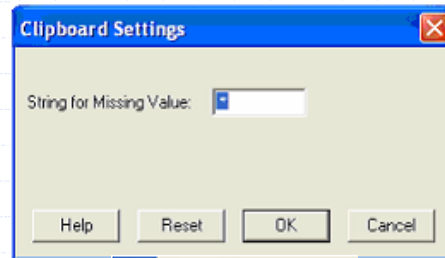
- Διαμόρφωση μεταβλητών
 - Διαμορφώνουμε τις μεταβλητές καθορίζοντας τον **Αριθμό** των **Δεκαδικών** ψηφίων



Κάνουμε δεξί κλικ πάνω στη μεταβλητή που θέλουμε να διαμορφώσουμε

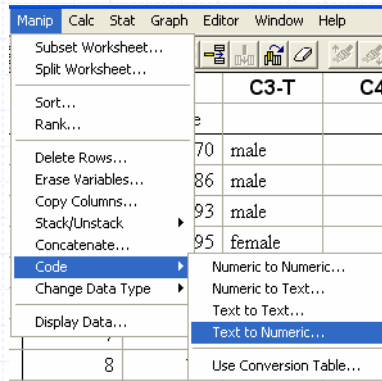
4. Οργάνωση Δεδομένων Missing values

- Διαμόρφωση μεταβλητών
 - Καθορίζουμε το **Σύμβολο** που θα παριστάνει πιθανές ελλείψεις παρατηρήσεις (**missing values**) στο αρχείο των δεδομένων μας



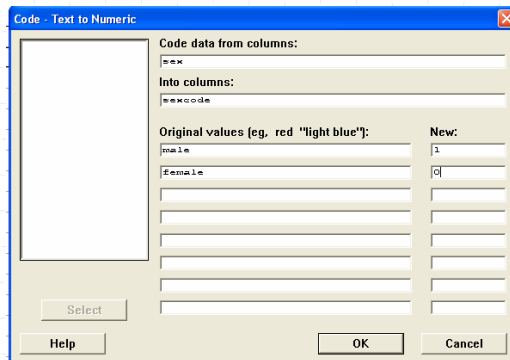
5. Οργάνωση Δεδομένων Κωδικοποίηση μεταβλητών

- Κωδικοποίηση μεταβλητών



Κωδικοποίηση Μεταβλητής Sex

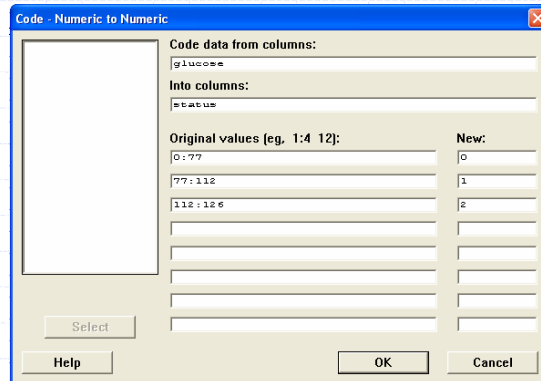
- Female:0
- Male:1



Προσοχή στη χρήση κεφαλαίων & πεζών γραμμάτων

Ομαδοποίηση Μεταβλητής Glucose

- ≤ 77 :0 "No diabetes"
- 77-112 :1 "Borderline"
- > 112 :2 "Diabetes"



Τα διαστήματα είναι κλειστά άνω

Αρχείο Δεδομένων

- Το αρχείο δεδομένων για τους πρώτους 10 ασθενείς

	C1	C2	C3	C4-T	C5	C6-T
↓	no_patient	glucose	status	stcde	sex	sexcode
1	1	70	0	no diabetes	1	male
2	2	86	1	borderline	1	male
3	3	93	1	borderline	1	male
4	4	95	1	borderline	0	female
5	5	97	1	borderline	0	female
6	6	101	1	borderline	1	male
7	7	103	1	borderline	1	male
8	8	106	1	borderline	1	male
9	9	107	1	borderline	1	male
10	10	112	2	diabetes	0	female

Μεταβλητές

- Ποσοτικές -Συνεχείς
 - Ποσότητα γλυκόζης ~ **Glucose**
- Ποιοτικά-Κατηγορικά
 - Φύλο ~ **Sex**
 - Χαρακτηρισμός ατόμου ~ **Status**

1.1 Ανάλυση Δεδομένων "Glucose"

- Μετρήσιμη Μεταβλητή

glucose	Count	CumCnt	Percent	CumPct
70	1	1	1,00	1,00
79	1	2	1,00	2,00
80	1	3	1,00	3,00
83	1	4	1,00	4,00
85	4	8	4,00	8,00
86	3	11	3,00	11,00
87	2	13	2,00	13,00
88	1	14	1,00	14,00
89	1	15	1,00	15,00
90	1	16	1,00	16,00

Κατανομή Συχνοτήτων ΔΕΝ έχει νόημα γι' αυτό το λόγο ομαδοποιούμε τη μεταβλητή αυτή στη μεταβλητή status

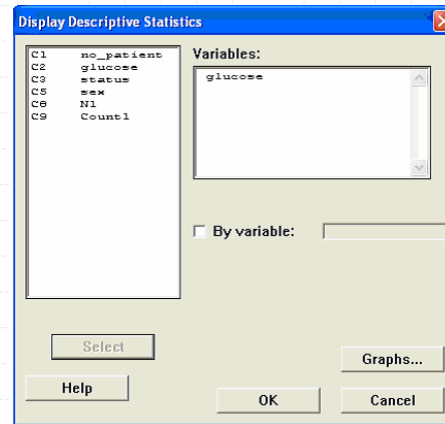
1.2 Ανάλυση Δεδομένων "Glucose"

- Μετρήσιμη Μεταβλητή

Περιγραφικά Μέτρα

- Stat->Basic Statistics-> Display Descriptive Statistics

```
MTB > Describe 'glucose'.
```



Οι εργασίες που κάνω κάνοντας χρήση των μενού μπορούν να γίνουν και με χρήση εντολών στο session του Minitab.

Για να είναι δυνατή η εμφάνιση του Minitab prompt (MTB>) θα πρέπει να έχει προηγηθεί η ενεργοποίησή του επιλέγοντας Editor/Enable Command Language.

1.3 Ανάλυση Δεδομένων "Glucose"

- Περιγραφικά Μέτρα
 - Αποτελέσματα στο Session του Minitab

Variable	N	Mean	Median	TrMean
glucose	100	100,00	100,50	100,00
	StDev	SE Mean	Minimum	
	10,14	1,01	70,00	
	Maximum	Q1	Q3	
	126,00	94,00	106,00	

Mean=100 ~~Αθροίζουμε όλες τις τιμές της γλυκόζης και διαιρούμε με το πλήθος των ασθενών

Median=100,50~~ Πενήντα παρατηρήσεις βρίσκονται πάνω από την τιμή 100,50 και πενήντα κάτω

TrMean=100~~ 5% trimmed mean Αφαιρείται το μικρότερο 5% και το μεγαλύτερο 5% των τιμών και υπολογίζεται ο μέση τιμή για τις υπόλοιπες τιμές.

StDev=10.14 mg% Πώς διασπείρονται /σκορπίζονται οι τιμές της γλυκόζης γύρω από τη μέση τιμή

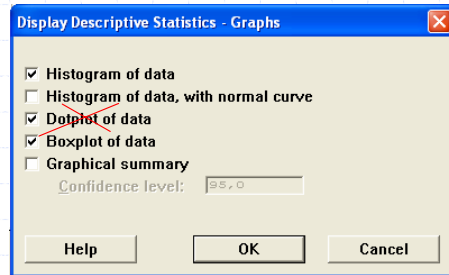
SE Mean=StDev/Sqrt(n) Μετράει τη μεταβλητότητα του εκτιμητή της μέσης τιμής από δείγμα σε δείγμα

Q1=94 25% των τιμών είναι μικρότερες από 94 mg% & 75 % των τιμών είναι μεγαλύτερες από 94 mg%

Q3=106 75 % των τιμών είναι μικρότερες από 106 mg% & 25 % των τιμών είναι μεγαλύτερες από 106 mg%

1.4 Ανάλυση Δεδομένων Γραφικά "Glucose"

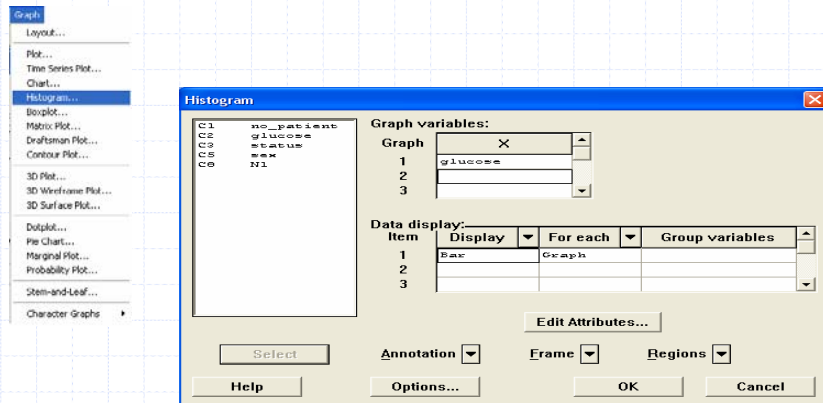
- Stat->Basic Statistics->
Display Descriptive Statistics ->Graphs



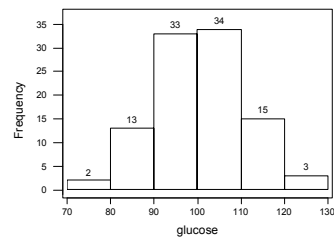
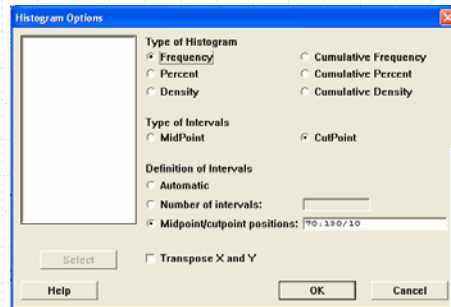
Το διάγραμμα dot plot που εμφανίζεται μέσω αυτού του μενού δεν είναι ακριβές. Για το λόγο αυτό θα πρέπει να δημιουργηθεί το διάγραμμα αυτό μέσω των επιλογών Γραφημάτων.

1.5 Ανάλυση Δεδομένων Γραφικά "Glucose"

Histogram-Ιστογράμμο



1.6 Ανάλυση Δεδομένων Γραφικά "Glucose"



1.7 Ανάλυση Δεδομένων Γραφικά "Glucose"

Histogram-Ιστόγραμμα

```
MTB > Describe 'glucose';  
SUBC> Ghist.
```

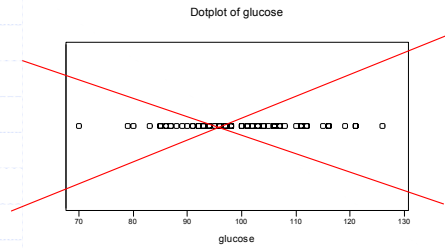


Αν θέλουμε να διατηρήσουμε τις ίδιες κλάσεις με το διάγραμμα μίσχου φύλλου, θα πρέπει να δημιουργήσουμε το ιστόγραμμα από το μενού των γραφημάτων.

1.8 Ανάλυση Δεδομένων Γραφικά "Glucose"

Dot Plot-Διάγραμμα Σημείων

```
MTB > Describe 'glucose';  
SUBC> Gdotplot.
```

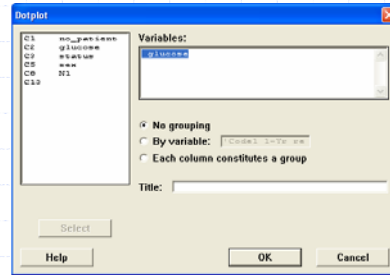


1.9 Ανάλυση Δεδομένων Γραφικά "Glucose"

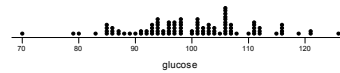
Dot-Plot

Διάγραμμα Σημείων

MTB > %Dotplot 'glucose'.



Dotplot for glucose



Στο γράφημα αυτό η θέση κάθε δεδομένου σε έναν αριθμημένο άξονα καθορίζεται με μία κουκκίδα. Όπου οι τιμές συμπίπτουν οι κουκκίδες τοποθετούνται η μία πάνω από την άλλη.

1.10 Ανάλυση Δεδομένων Γραφικά "Glucose"

Stem-and-Leaf

Διάγραμμα Μίσχου-Φύλλου

MTB > Stem-and-Leaf 'glucose';

SUBC> **Increment 1.**

```
1 7 0
1 7
1 7
1 7
2 7 9
3 8 0
4 8 3
8 8 5555
13 8 66677
15 8 89
18 9 011
24 9 223333
32 9 44444455
42 9 6666677777
48 9 888888
(8) 10 00111111
44 10 2223333
37 10 4445
33 10 6666666667777
19 10 8
18 11 011111
12 11 222
9 11 5
8 11 6666
4 11 9
3 12 11
1 12
1 12
1 12 6
```

1.11 Ανάλυση Δεδομένων Γραφικά "Glucose"

Stem-and-Leaf

Διάγραμμα Μίσχου-Φύλλου

MTB > Stem-and-Leaf 'glucose';
SUBC> **Increment 10.**

```
2  7 09                      Η μικρότερη τιμή γλυκόζης είναι 70
15 8 03555566667789
48 9 011223333444444556666677777888888      Μεγαλύτερη συγκέντρωση τιμών
(34) 10 001111111222333344445666666666677778      μεταξύ 90 και 108
18 11 011111222566669
3  12 116                      Η μεγαλύτερη τιμή γλυκόζης είναι 126
```

Αυξάνοντας το "increment" σε 10 δηλώνοντας δηλαδή ότι η διαφορά μεταξύ των μικρότερων πιθανών τιμών σε δύο διαδοχικές γραμμές είναι μία δεκάδα το διάγραμμα αλλάζει ως προς τον τρόπο παρουσίασης

Η αριστερή στήλη δηλώνει το σύνολο των τιμών (αθροιστικά) από την κορυφή και προς τα κάτω & από το τέλος και προς τα πάνω μέχρι τη διάμεσο

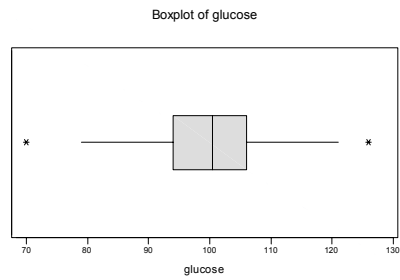
Η τιμή σε παρένθεση,(34), μετράει το πλήθος των τιμών στη γραμμή που εμπεριέχει τη διάμεσο

Στη δεξιά στήλη κάθε ψηφίο (φύλλο) που τοποθετείται δεξιά από το μίσχο αντιπροσωπεύει μία τιμή των δεδομένων

1.12 Ανάλυση Δεδομένων Γραφικά "Glucose"

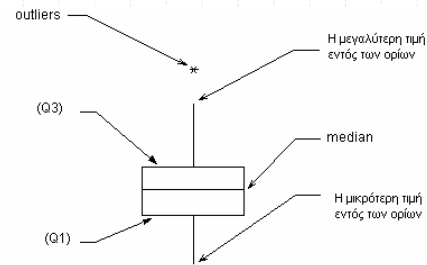
Box Plot- Διάγραμμα Πλαισίου Απολήξεων

```
MTB > Describe 'glucose';  
SUBC> GBoxplot.
```



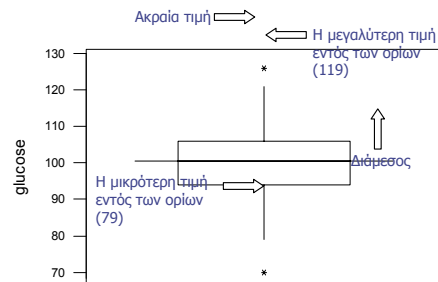
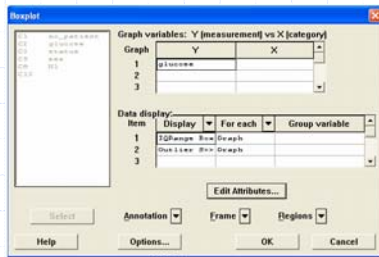
1.13 Ανάλυση Δεδομένων Γραφικά "Glucose"

Box Plot- Διάγραμμα Πλαισίου Απολήξεων



1.14 Ανάλυση Δεδομένων Γραφικά "Glucose"

Box-Plot Διάγραμμα Σημείων Απολήξεων



Το κεντρικό πλαίσιο περιέχει το κεντρικό 50% των δεδομένων.

Το άνω άκρο του πλαισίου δηλώνει το Q3, ενώ το κάτω άκρο το Q1

Οι απολήξεις δείχνουν τη μέγιστη και ελάχιστη τιμή εκτός αν υπάρχουν ακραίες τιμές

Αν η διάμεσος ισαπέχει από τα άκρα του πλαισίου τότε η κατανομή των δεδομένων είναι συμμετρική

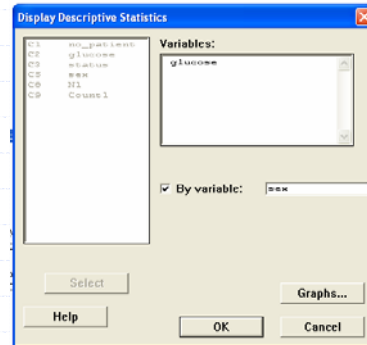
Στο ίδιο συμπέρασμα θα καταλήγαμε συγκρίνοντας τις απολήξεις του πλαισίου.

Γενικά αν η απόληξη του πάνω πλαισίου είναι μεγαλύτερη από αυτή του κάτω, τότε υπάρχουν περισσότερες μεγάλες τιμές προς το θετικό άκρο της κατανομής και έτσι η κατανομή είναι θετικά ασύμμετρη /λοξή. Αν ισχύει το ακριβώς αντίθετο τότε η κατανομή είναι αρνητικά ασύμμετρη /λοξή

1.15 Ανάλυση Δεδομένων "Glucose" Κατά φύλο

- Περιγραφικά Μέτρα Glucose κατά Φύλο
 - Stat->Basic Statistics->Display Descriptive Statistics

```
MTB > Describe 'glucose';  
SUBC> By 'sex';
```



1.16 Ανάλυση Δεδομένων "Glucose" Κατά φύλο

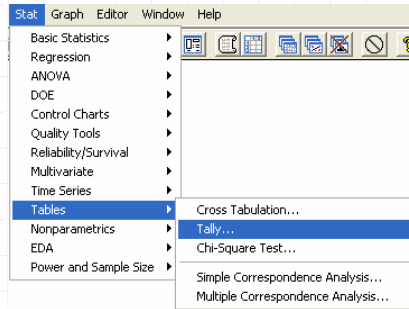
- Περιγραφικά Μέτρα Glucose κατά Φύλο

Variable	sex	N	Mean	Median	TrMean	StDev
glucose	0	48	100,92	101,00	100,82	9,12
	1	52	99,15	99,00	99,22	11,01

Variable	sex	SE Mean	Minimum	Maximum	Q1	Q3
glucose	0	1,32	83,00	121,00	94,00	106,00
	1	1,53	70,00	126,00	93,00	106,00

Ανάλυση Δεδομένων Κατηγορικών Μεταβλητών

- Κατανομή Συχνοτήτων



2.1 Ανάλυση Δεδομένων "Sex"

- Ονομαστική Μεταβλητή
 - Κατανομή Συχνοτήτων

```
MTB > Tally 'sex';  
SUBC> Counts;  
SUBC> CumCounts;  
SUBC> Percents;  
SUBC> CumPercents.
```

Summary Statistics for Discrete Variables

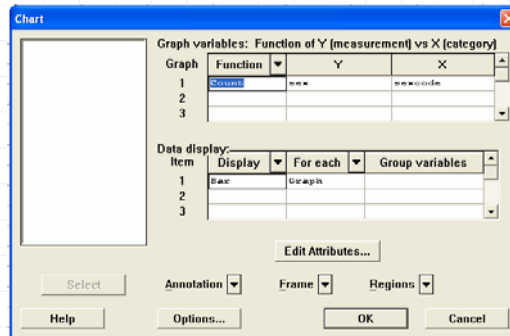
sex	Count	CumCnt	Percent	CumPct
0	48	48	48,00	48,00
1	52	100	52,00	100,00
N=	100			

↑
Επικρατούσα
τιμή

↑
Οι Αθροιστικές Συχνότητες ΔΕΝ έχουν
νόημα για κατηγορικές μεταβλητές

2.2 Ανάλυση Δεδομένων Γραφικά "Sex"

- Bar Chart
 - Graph->Chart

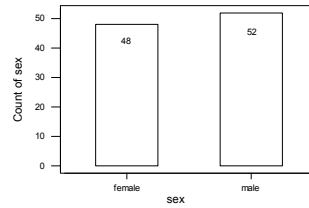


2.3 Ανάλυση Δεδομένων Γραφικά "Sex"

- Bar Chart
 - Graph->Chart

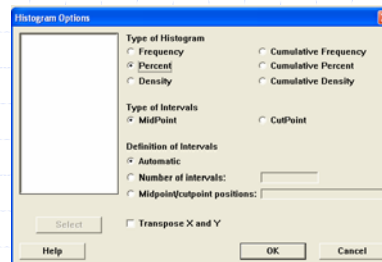
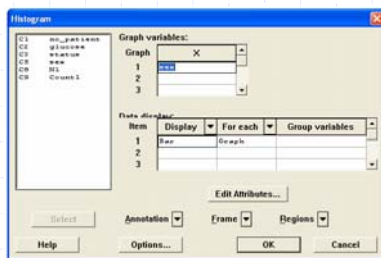
Διαθέσιμες Συναρτήσεις για το Γράφημα

Count	number of values (frequency)
N	number of non-missing values
Nmiss	number of missing values
Mean	mean of values
Median	median of values
Minimum	the smallest value
Maximum	the largest value
Stdev	standard deviation of values
Ssq	sum of squares of values



2.4 Ανάλυση Δεδομένων Γραφικά "Sex"

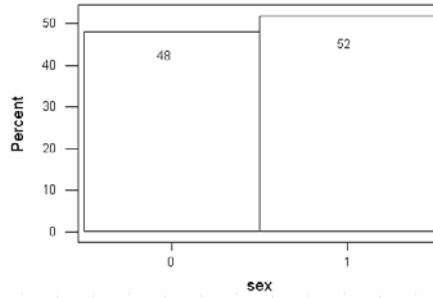
- Bar Chart
 - Graph->Histogram



Για να εμφανίσουμε ποσοστά στο γράφημα

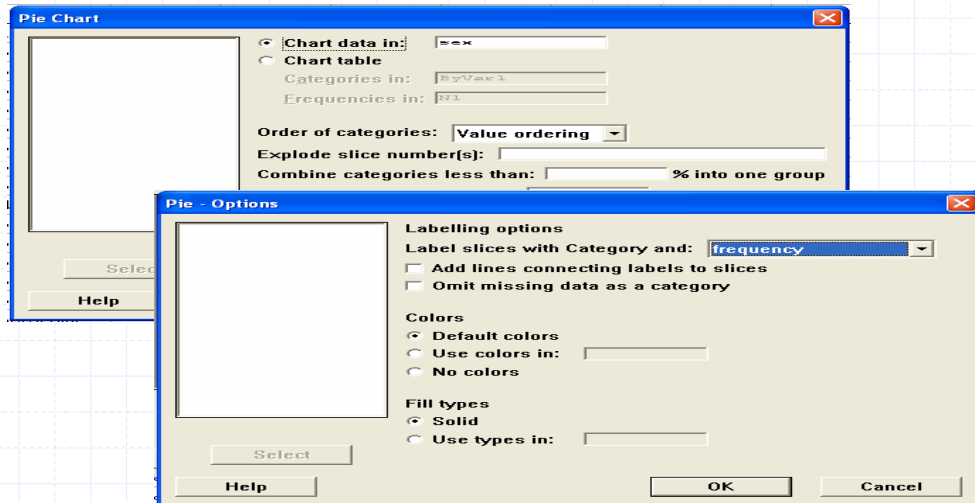
2.5 Ανάλυση Δεδομένων Γραφικά "Sex"

- Bar Chart
 - Graph->Histogram



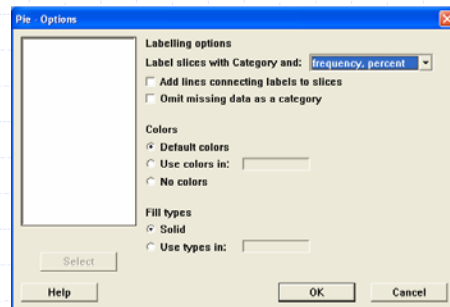
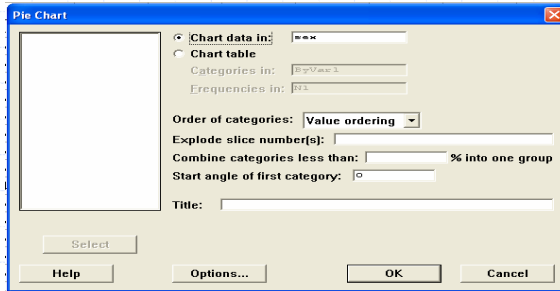
2.6 Ανάλυση Δεδομένων Γραφικά "Sex"

- Pie Chart
 - Graph->Chart



2.7 Ανάλυση Δεδομένων Γραφικά "Sex"

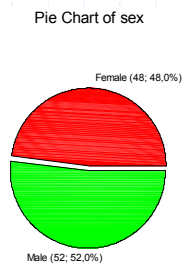
- Pie Chart
 - Graph->Chart



2.8 Ανάλυση Δεδομένων Γραφικά "Sex"

- Pie Chart
 - Graph->Chart

```
MTB > %Pie 'sex' .
```



3.1 Ανάλυση Δεδομένων "Status"

- Διατάξιμη Μεταβλητή
 - Κατανομή Συχνοτήτων

```
MTB > Tally 'status';  
SUBC> Counts;  
SUBC> CumCounts;  
SUBC> Percents;  
SUBC> CumPercents.
```

Summary Statistics for Discrete Variables

status	Count	CumCnt	Percent	CumPct
0	1	1	1,00	1,00
1	87	88	87,00	88,00
2	12	100	12,00	100,00
N=	100			

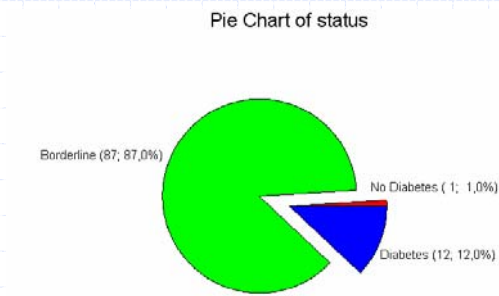
88% των ατόμων έχουν τιμές γλυκόζης στο αίμα τους <112 mg%

Η Επικρατούσα Τιμή στη μεταβλητή status είναι 1

3.2 Ανάλυση Δεδομένων Γραφικά "Status"

- Pie Chart
 - Graph->Chart

```
MTB > %Pie 'status'.
```



3.3 Ανάλυση Δεδομένων Γραφικά "Status"

- Bar Chart

```
MTB > Chart Count( status ) * 'stcode';
```

```
SUBC> Bar;
```

```
SUBC> ScFrame;
```

```
SUBC> ScAnnotation;
```

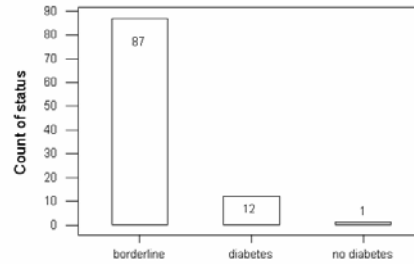
```
SUBC> Symbol;
```

```
SUBC> Type 0;
```

```
SUBC> Label;
```

```
SUBC> Offset 0,0 0,0125;
```

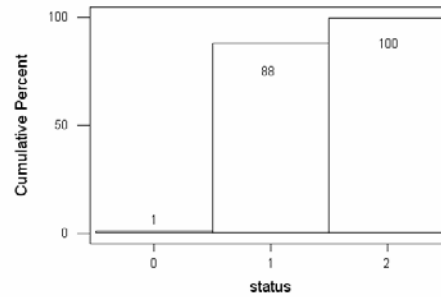
```
SUBC> Placement 0 1.
```



3.4 Ανάλυση Δεδομένων Γραφικά "Status"

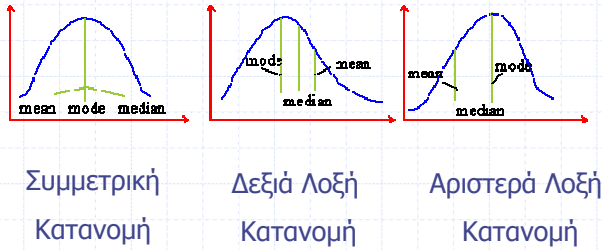
• Bar Chart

```
MTB > Histogram 'status';  
SUBC> Cumulative;  
SUBC> Percent;  
SUBC> MidPoint;  
SUBC> Bar;  
SUBC> ScFrame;  
SUBC> ScAnnotation;  
SUBC> Symbol;  
SUBC> Type 0;  
SUBC> Label;  
SUBC> Offset 0,0 0,0125;  
SUBC> Placement 0 1;  
SUBC> Axis 1;  
SUBC> Axis 2.
```



Ανάλυση Δεδομένων

•Κατανομές



Οι δύο κατανομές διαφέρουν στη συμμετρία. Η κατανομή στα αριστερά είναι συμμετρική ενώ η κατανομή στα δεξιά ασύμμετρη /λοξή.

Η συμμετρία προσδιορίζεται από τη ουρά με το μεγαλύτερο μήκος.

Αν η ουρά με το μεγαλύτερο μήκος βρίσκεται προς το μέρος των μεγάλων αριθμών, τότε η κατανομή καλείται δεξιά ή θετική λοξή, ενώ στην αντίθετη περίπτωση καλείται αριστερά ή αρνητικά λοξή.
