Δημήτρης Α. Ιωαννίδης
Καθηγητής Πανεπιστημίου

ΣΤΑΤΙΣΤΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ

- Περιγραφική Στατιστική
- Θεωρία Πιθανοτήτων
- Στατιστική Συμπερασματολογία
- Απλή και Πολλαπλή Γραμμική Παλινδρόμηση
- Χρονολογικές Σειρές
- Ανάλυση Διακύμανσης

ΜΕ ΟΔΗΓΟ ΧΡΗΣΗΣ EXCEL

ΠΕΡΙΕΧΕΙ CD-ROM

ΕΚΔΟΣΕΙΣ ΖΗΤΗ

ISBN 978-960-456-278-7

Φωτοστοιχειοθεσία
Εκτύπωση
Βιβλιοδεσία

Π. ΖΗΤΗ & Σια ΟΕ
18ο χλμ Θεσσαλονίκης - Περαίας
Τ.Θ. 4171 • Περαία Θεσσαλονίκης • Τ.Κ. 570 19
Τηλ.: 2392 072.222 - Fax: 2392 072.229 • e-mail: info@ziti.gr

ΕΚΔΟΣΕΙΣ ΖΗΤΗ
www.ziti.gr

ΒΙΒΛΙΟΠΩΛΕΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ - ΚΕΝΤΡΙΚΗ ΔΙΑΘΕΣΗ:
Αρμενοπούλου 27 - 546 35 Θεσσαλονίκη • Τηλ.: 2310 203.720 • Fax 2310 211.305
e-mail: sales@ziti.gr

ΒΙΒΛΙΟΠΩΛΕΙΟ ΑΘΗΝΩΝ - ΕΝΩΣΗ ΕΚΔΟΤΩΝ ΒΙΒΛΙΟΥ ΘΕΣΣΑΛΟΝΙΚΗΣ:
Στοά του Βιβλίου (Πεσμαζόγλου 5) - 105 64 ΑΘΗΝΑ • Τηλ.-Fax: 210 3211.097

ΑΠΟΘΗΚΗ ΑΘΗΝΩΝ - ΠΩΛΗΣΗ ΧΟΝΔΡΙΚΗ:
Ασκληπιού 60 - Εξάρχεια 114 71, Αθήνα • Τηλ.-Fax: 210 3816.650 • e-mail: athina@ziti.gr

ΗΛΕΚΤΡΟΝΙΚΟ ΒΙΒΛΙΟΠΩΛΕΙΟ: www.ziti.gr

Πρόλογος

Η χρησιμοποίηση των στατιστικών μεθόδων μπορεί να συμβάλλει σημαντικά σε μία αποτελεσματικότερη διοίκηση στο Δημόσιο και στον Ιδιωτικό τομέα. Η σημασία της Στατιστικής είναι εμφανής, αφ' ενός από το γεγονός ότι στα προγράμματα των περισσοτέρων πανεπιστημιακών τμημάτων, ιδιαίτερα της αλλοδαπής, περιλαμβάνονται ποικίλα μαθήματα Στατιστικής, και αφ' ετέρου πολλοί δημόσιοι και ιδιωτικοί οργανισμοί διεθνώς διαθέτουν στατιστικά γραφεία που ασχολούνται με την καταγραφή και συλλογή δεδομένων που τους αφορούν, αλλά και συχνά με τη λήψη αποφάσεων ή προβλέψεων.

Σκοπός αυτού του εγχειριδίου είναι να παρουσιασθούν οι πρώτες βασικές αρχές της Στατιστικής μ' ένα μεθοδικό και επεξηγηματικό τρόπο, αποφεύγοντας τις ιδιαίτερες μαθηματικές δυσκολίες, αλλά και να δοθεί η δυνατότητα στον αναγνώστη να εξοικειωθεί με την χρήση ηλεκτρονικών υπολογιστών όταν τις εφαρμόζει. Τα δύο πρώτα κεφάλαια ασχολούνται με εισαγωγικές έννοιες της στατιστικής, και συγκεκριμένα της περιγραφικής. Η κατανόηση αυτών είναι βοηθητική για την ευκολότερη εκμάθηση των εννοιών των επόμενων κεφαλαίων. Στα Κεφάλαια 3, 4 και 5 γίνεται μια σύντομη εισαγωγή στις πιθανότητες, και μελετώνται οι βασικές θεωρητικές κατανομές. Μία εισαγωγή στις πολυδιάστατες κατανομές δίνεται στο Κεφάλαιο 6, με έμφαση τη μελέτη των ανεξάρτητων τυχαίων μεταβλητών. Στο Κεφάλαιο 7 αναλύουμε την έννοια της κανονικής προσέγγισης καθώς και την σχέση της με τον γενικότερο όρο της στατιστικής έρευνας και δειγματοληψίας. Στα υπόλοιπα κεφάλαια διατυπώνονται και αναπτύσσονται οι βασικότερες μέθοδοι για την εξαγωγή στατιστικών συμπερασμάτων σχετικά με τις παραμέτρους που χαρακτηρίζουν έναν πληθυσμό, με βάση πάντα ένα τυχαίο δείγμα που λαμβάνεται απ' αυτόν. Έτσι στο Κεφάλαιο 8 παρουσιάζονται τα απαραίτητα τεχνικά μέσα που θα μας βοηθήσουν στην εξαγωγή αυτών των συμπερασμάτων. Στο Κεφάλαιο 9 αναπτύσσονται διάφοροι Μέθοδοι Σημειοεκτίμησης των παραμέτρων ενός πληθυσμού, ενώ στο Κεφάλαιο 10 η εκτίμηση αυτών γίνεται με τη βοήθεια των Διαστημάτων Εμπιστοσύνης. Άλλες μέθοδοι εξαγωγής στατιστικών συμπερασμάτων είναι αυτές των Ελέγχων Υποθέσεων και δίνονται στο Κεφάλαιο 11. Στα Κεφάλαια 9, 10 και 11 πάντοτε δεχόμαστε ότι ο πληθυσμός που μελετάμε ακολουθεί μία Κανονική κατανομή,

κάτι όμως που μπορούμε να το ελέγξουμε στατιστικά. Το τελευταίο επιτυγχάνεται στο Κεφάλαιο 12 με την βοήθεια των Ελέγχων Καλής Προσαρμογής. Επίσης, στο ίδιο κεφαλαίο αναπτύσσονται και άλλα στατιστικά προβλήματα, όπως αυτό του Ελέγχου της Ανεξαρτησίας μεταξύ δυο ποιοτικών χαρακτηριστικών. Στο Κεφάλαιο 13, με την βοήθεια της Ευθείας Παλινδρόμηση μελετάμε την από κοινού συμπεριφορά δύο μεταβλητών, εκτιμώντας την Ευθεία Παλινδρόμησης, τον βαθμό εξάρτησης των μεταβλητών και δίνοντας την στατιστική αποτελεσματικότητα των συμπερασμάτων μας που την αφορούν. Με τις βασικές αρχές της Ανάλυσης Διακύμανσης, που δίνονται στο Κεφάλαιο 14, συγκρίνουμε τους μέσους πολλών πληθυσμών. Το Κεφάλαιο 15 αποτελεί την γενίκευση του 13 και αφορά την Πολυδιάστατη Παλινδρόμηση. Στο Κεφάλαιο 16 δίνονται οι βασικές αρχές μια Χρονολογικής Σειράς και στην συνέχεια δίνουμε συνοπτικά τα μοντέλα χρονολογικών σειρών κατά Box και Jenkins. Οι βασικοί Μη-Παραμετρικοί έλεγχοι παρουσιάζονται στο Κεφάλαιο 16. Όλα τα παραδείγματα και οι ασκήσεις αποτελούν εφαρμογές της Στατιστικής, χρήσιμες για οικονομολόγους και στελέχη επιχειρήσεων. Τέλος, όλη η Στατιστική Μεθοδολογία που αναπτύσσεται στο βιβλίο μαζί με τα παραδείγματα της γίνεται και με την χρήση του Excel και παρουσιάζεται στο τέλος του βιβλίου σαν ανεξάρτητο κεφάλαιο και με τον τίτλο “Οδηγός Χρήσης του Excel”.

Στη συγγραφή του Excel συνέβαλε σημαντικά ο πτυχιούχος μαθηματικών κ. Πετρίδης Κώστας, τον οποίο και ευχαριστώ θερμά.

Θεσσαλονίκη, Μάρτιος 2011 *Δημήτρης Ιωαννίδης*

Περιεχόμενα

15 Πολυδιάστατη Παλινδρόμηση

16 Χρονολογικές Σειρές

17 Μη Παραμετρικοί Έλεγχοι

Οδηγός χρήσης του Excel

Παράρτημα

1ο Κεφάλαιο

Γενικά περί Στατιστικής

Οι περισσότεροι άνθρωποι όταν αναφέρονται στη λέξη **"Στατιστική"** συνήθως σκέπτονται πίνακες από αριθμούς (αριθμητικά δεδομένα). Ως επί το πλείστον είναι εξοικειωμένοι με αριθμητικά δεδομένα που αναφέρονται σε απογραφές πληθυσμών, σε αναλύσεις αποτελεσμάτων αθλημάτων, σε διάφορα μεγέθη της οικονομίας κ.λπ. Το περιεχόμενο της στατιστικής όμως δεν ανταποκρίνεται μόνο σε κάποιους πίνακες αριθμητικών δεδομένων. Η στατιστική σήμερα αποτελεί έναν ανεξάρτητο επιστημονικό κλάδο, με δικές της μεθόδους ανάλυσης. Το αντικείμενό της είναι πολύ ευρύ, ασχολείται με σχεδιασμούς πειραμάτων, μεθόδους συλλογής δεδομένων, καθώς και με την ανάλυση, παρουσίαση, επεξήγηση των δεδομένων με απώτερο σκοπό τη λήψη αποφάσεων και τη δημιουργία προβλέψεων. Τα αριθμητικά δεδομένα αποτελούν ένα πολύ μικρό μέρος αυτής.

Οι πλέον απλές στατιστικές μέθοδοι είναι αυτές που ασχολούνται με τους μέσους των δεδομένων καθώς και άλλων μέτρων που μπορούν να περιγράψουν τα δεδομένα και να μας δώσουν πληροφορίες σχετικά με τον τρόπο κατανομής των τιμών τους (**Περιγραφική Στατιστική**). Για παράδειγμα, έστω ότι ενδιαφερόμαστε για να αποκτήσουμε γνώση γύρω από το ετήσιο οικογενειακό εισόδημα 1000 νοικοκυριών μιας περιοχής, η επιλογή των οποίων γίνεται κατά έναν αντιπροσωπευτικό τρόπο. Μερικά χαρακτηριστικά αριθμητικά μεγέθη που μπορούν να μας δώσουν πληροφορίες γύρω απ' αυτά είναι ο μέσος όρος αυτών, καθώς και ο τρόπος απόκλισης τους από το μέσο αυτών (Διακύμανση ή Διασπορά).

Η σύγχρονη στατιστική διαθέτει μεθόδους για εξαγωγή συμπερασμάτων και πέρα από ένα συγκεκριμένο αριθμό δεδομένων που γνωρίζουμε, καταλήγοντας σε συμπεράσματα που στηρίζονται στην ανάλυση αυτών.

Έτσι στο προηγούμενο παράδειγμα ένα εύλογο ερώτημα είναι ποιο μπορεί να είναι το μέσο ετήσιο εισόδημα των νοικοκυριών ολόκληρης της περιοχής και όχι

απλά και μόνο των 1000 νοικοκυριών. Εδώ, μπορούμε να διατυπώσουμε την (στατιστική) υπόθεση ότι το μέσο εισόδημα είναι λιγότερο από 14.000 €, οπότε με βάση τη γνώση των 1000 εισοδημάτων να δεχθούμε ή να απορρίψουμε την υπόθεση αυτή (**Έλεγχος Υποθέσεων**). Επίσης μπορούμε να έχουμε μια εκτίμηση για το μέσο εισόδημα όλων των οικογενειών, υπολογίζοντας αριθμητικά τον μέσο των 1000 εισοδημάτων από τα δεδομένα μας δηλ. μια μόνο αριθμητική τιμή για το μέσο εισόδημα όλων των οικογενειών με βάση πάλι τα 1000 γνωστά εισοδήματα (**Σημειοεκτίμηση**). Τέλος αντί να έχουμε μια εκτίμηση του μέσου εισοδήματος, να κατασκευάσουμε ένα διάστημα που η αρχή του θα είναι η κατώτερη μέση εκτίμηση, ενώ το τέλος του θα είναι η ανώτερη μέση εκτίμηση και ανάμεσα στις δύο τιμές να εμπεριέχεται το (άγνωστο) μέσο εισόδημα μ' ένα βαθμό εμπιστοσύνης (**Διαστήματα Εμπιστοσύνης**). Οι τρείς τελευταίες έννοιες που αναφέρθηκαν αποτελούν αντικείμενο της "**Στατιστικής Συμπερασματολογίας**" ή της "**Επαγωγικής Στατιστικής**".

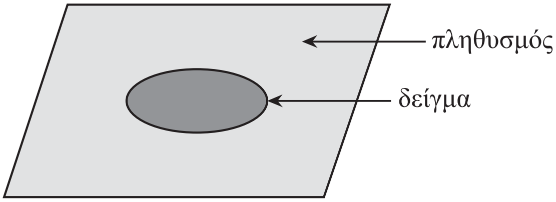
Στηριζόμενοι στα παραπάνω μπορούμε να χαρακτηρίσουμε την περιοχή σαν υψηλών ή μεσαίων ή χαμηλών εισοδημάτων.

Το βοηθητικό επιστημονικό εργαλείο που θα μας δίνει τη δυνατότητα εξαγωγής συμπερασμάτων για όλο το πλήθος των οικογενειών με βάση τα γνωστά σε μας 1000 εισοδήματα είναι η "**Θεωρία Πιθανοτήτων**".

Να συζητήσουμε μερικές ακόμη πτυχές αυτού του παραδείγματος. Κάποιος θα ρωτούσε γιατί δεν μαθαίνουμε όλα τα εισοδήματα των κατοίκων αυτής της περιοχής. Οπωσδήποτε για να γίνει αυτό απαιτεί αρκετό χρόνο και σημαντικά οικονομικά έξοδα. Επειδή όμως λογικό είναι να ελαχιστοποιήσουμε το χρόνο εργασίας μας, και το κόστος που απαιτεί αυτή, περιοριζόμαστε σ' ένα μικρότερο αριθμό εισοδημάτων. Το τελευταίο πρέπει να γίνει μ' ένα τρόπο μεθοδικό. Τα 1000 εισοδήματα που επιλέξαμε κατά ένα τυχαίο τρόπο πρέπει να είναι αντιπροσωπευτικά γι' όλα τα υπόλοιπα εισοδήματα. Εν ολίγοις, κάθε εισόδημα πρέπει να έχει την ίδια δυνατότητα ή τον ίδιο βαθμό βεβαιότητας για να είναι ένα μεταξύ των 1000 εισοδημάτων. Υπάρχει ξεχωριστός κλάδος της στατιστικής που ασχολείται μ' αυτό το πρόβλημα και καλείται "**Θεωρία Δειγματοληψιών**". Θα ασχοληθούμε μ' αυτήν εν συντομία σ' ένα από τα επόμενα κεφάλαια.

Το πλήθος των οικογενειών που δυνητικά αποτελούν αντικείμενα προς εξέταση στην έρευνά μας είναι ο πληθυσμός μας. Γενικά το πλήθος των αντικειμένων ή ατόμων που συνδέονται με τη (στατιστική) έρευνά μας ονομάζεται "**πληθυσμός**". Τα 1000 νοικοκυριά που τα εισοδήματά τους παρατηρήθηκαν αποτελούν το "δείγμα μας". Κάθε υποσύνολο ενός πληθυσμού ονομάζεται **δείγμα**.

Για τη Στατιστική, σε κάθε πληθυσμό θα υπάρχουν χαρακτηριστικά γνωρίσματα, όπως το εισόδημα παρακάτω, που θα αποτελούν αντικείμενο μελέτης και



έρευνας. Τα αποτελέσματα της έρευνάς μας θα είναι συμπεράσματα που θα εξάγονται σχετικά με τα χαρακτηριστικά και θα προκύπτουν μετά την ανάλυση και επεξεργασία του δείγματός μας.

Ας συζητήσουμε μερικά ακόμα παραδείγματα, όπου με τη βοήθεια μεθόδων της Στατιστικής αντλούμε χρήσιμα συμπεράσματα.

➲ Δημοσκοπήσεις:

Έστω ότι ενδιαφερόμαστε να γνωρίσουμε αν το κόμμα A ή το B κερδίσει μία εκλογική αναμέτρηση. Εδώ ο πληθυσμός μας είναι το σύνολο όλων των ψηφοφόρων. Οι διάφορες εταιρίες δημοσκοπήσεων προβλέπουν το ποιο κόμμα θα κερδίσει με βάση ένα πολύ μικρότερο αριθμό ερωτηθέντων ατόμων (Δείγμα).

➲ Ποιοτικός Έλεγχος:

α) Σε ένα εργοστάσιο παραγωγής τσιμέντου, οι σάκοι των τσιμέντων οφείλουν να έχουν βάρος 50 kg, και η πακετοποίηση γίνεται με μια αυτόματη μηχανή. Για να ελέγξει κανείς την αξιοπιστία αυτής της μηχανής, θα έπρεπε να ζυγίζει κάθε σάκο τσιμέντου. Επειδή αυτή η διαδικασία προδικάζει πολύ χρόνο και αρκετά έξοδα, ζυγίζει κανείς κάθε 20^{o} ή 50^{o} σάκο και προσδιορίζει το μέσο βάρος των ζυγισθέντων σάκων. Αν το μέσο βάρος απέχει σημαντικά από τα 50 kg τότε αυτό θα σημαίνει ότι η μηχανή χρειάζεται μια επιδιόρθωση. Οι σάκοι που παρήχθηκαν κατά τη διάρκεια μιας ημέρας αποτελούν τον πληθυσμό μας, ενώ το σύνολο των ζυγισθέντων σάκων το δείγμα μας.

β) Ορισμένα αντικείμενα κατασκευάζονται σύμφωνα με μια βιομηχανική μέθοδο. Θέλουμε να ελέγξομε το πλήθος των ελαττωματικών αντικειμένων που παράγονται κατά τη διάρκεια μιας χρονικής μονάδας παραγωγής. Τα αντικείμενα που παρήχθηκαν κατ' αυτήν τη χρονική μονάδα αποτελούν τον πληθυσμό μας, ενώ τα αντικείμενα που εξετάσαμε το δείγμα μας.

Το βιβλίο αυτό περιέχει τα εξής Κεφάλαια. Το Κεφάλαιο 2 περιέχει τις βασι-

κές γνώσεις της Περιγραφικής Στατιστικής ενώ τα Κεφάλαια 3 έως και 7 ασχολούνται με βασικές αρχές της Θεωρίας Πιθανοτήτων. Τα υπόλοιπα Κεφάλαια περιέχουν τις βασικές μεθόδους της Στατιστικής Συμπερασματολογίας.

2ο Κεφάλαιο

Περιγραφική Στατιστική

2.1 Μεταβλητές και Μετρήσεις

Τα μέλη ενός πληθυσμού θα τα καλούμε **στατιστικές μονάδες**, ενώ το κοινό χαρακτηριστικό γνώρισμα μεταξύ τους θα ονομάζεται **στατιστικό γνώρισμα** ή απλά **γνώρισμα**. Ένα στατιστικό γνώρισμα θα συμβολίζεται στο εξής μ' ένα κεφαλαίο γράμμα X ή Y κ.λπ., που θα καλείται **μεταβλητή**.

Έτσι στο παράδειγμα με τα οικογενειακά εισοδήματα οι στατιστικές μονάδες είναι τα νοικοκυριά και το στατιστικό γνώρισμα το εισόδημα X. Στο παράδειγμα των δημοσκοπήσεων η στατιστική μονάδα είναι ο ψηφοφόρος, ενώ το στατιστικό γνώρισμα είναι η ψήφος X για το A ή το B κόμμα. Στο πρώτο παράδειγμα ποιοτικού ελέγχου οι στατιστικές μονάδες είναι οι σάκοι των τσιμέντων που παράγονται, ενώ το στατιστικό γνώρισμα είναι το βάρος X. Σαν στατιστικό γνώρισμα μπορεί να θεωρηθεί η ηλικία, το βάρος, το φύλο, η θρησκεία κ.λπ.. Γενικά ένα στατιστικό γνώρισμα μπορεί να είναι **ποσοτικό**, δηλ. να επιδέχεται μετρήσεις, ή **ποιοτικό**, δηλ. να επιδέχεται κατηγοριοποιήσεις. Παραδείγματα ποσοτικών στατιστικών γνωρισμάτων είναι το εισόδημα, η ηλικία, το βάρος κ.λπ., ενώ ποιοτικών είναι το φύλλο, η θρησκεία κ.λπ.. Επιπλέον τα ποσοτικά γνωρίσματα διακρίνονται σε **συνεχή** και **διακριτά**. Ένα ποσοτικό γνώρισμα καλείται διακριτό, αν οι μετρήσεις του είναι διακριτοί πραγματικοί αριθμοί ή στην πλέον συνήθη περίπτωση οι θετικοί ακέραιοι, 0, 1, 2, … Έτσι το πλήθος των εργαζομένων σ' ένα εργοστάσιο, το πλήθος των αντικειμένων που παράγονται από μια βιομηχανική μονάδα, είναι διακριτά γνωρίσματα. Συνεχές χαρακτηριστικό είναι εκείνο που όταν το μετρούμε πάνω στις στατιστικές μας μονάδες οι αριθμοί των μετρήσεών του είναι πραγματικοί αριθμοί (το χαρακτηριστικό επιδέχεται μετρήσεις και μεταξύ ακεραίων, δηλ. «έχουμε και δεκαδικούς αριθμούς για με-

τρήσεις»). Τυπικά παραδείγματα γι’ αυτήν την περίπτωση είναι το εισόδημα, ο χρόνος λειτουργίας μιας μηχανής κ.λπ. Στην περίπτωση που έχουμε ποιοτικά χαρακτηριστικά, μπορούμε να έχουμε κατά συμβιβασμό ένα είδος μετρήσεων ανάλογο της διακριτής περίπτωσης. Για παράδειγμα εξετάζοντας το θρήσκευμα των κατοίκων σε μια χώρα, μπορούμε να παραστήσουμε κατά συμβιβασμό

Άθεος = 0, Ορθόδοξος Χριστ. = 1, Καθολικός Χριστ. = 2, Μουσουλμάνος = 3

ή εξετάζοντας τα αντικείμενα μιας βιομηχανικής παραγωγής
έχουμε ελαττωματικό = 0, μη ελαττωματικό = 1.

Βέβαια η εδώ ποσοτικοποίηση που επιτυγχάνεται είναι εντελώς εξωτερική. Μπορούμε να συμβολίσουμε τα θρησκεύματα με άλλους αριθμούς, χωρίς να έχουμε απώλεια σε πληροφορίες. Έτσι μπορούσαμε να έχουμε Άθεος = 3, Μουσουλμάνος = 2 κ.λπ. Αν όμως ρωτούσαμε μια νοικοκυρά να κατατάξει 4 απορρυπαντικά A, B, Γ και Δ σαν πρώτο καλύτερο με 1, δεύτερο καλύτερο με 2, τρίτο καλύτερο με 3, και τελευταίο με 4, τότε αυτοί οι αριθμοί που αντιστοιχούν σ’ αυτές τις κατατάξεις προκύπτουν μέσω από συγκρίσεις και έχουν την δική τους ιεραρχική σημασία. Στην περίπτωση με τις θρησκείες θα λέμε ότι είμαστε στην **ονομαστική κλίμακα**, ενώ στη δεύτερη με τις επιδόσεις, στην **τακτική κλίμακα**.

Οι **μετρήσεις** λαμβάνονται με **ερωτήσεις** (μέσω συνεντεύξεων ή γραπτών εξετάσεων), **παρατηρήσεις** (παρατηρούμε το πλήθος των αυτοκινήτων που περνούν από κάποιο σημείο κατά μία χρονική περίοδο ή παρατηρούμε το χρόνο αναμονής των πελατών στη θυρίδα μιας τράπεζας), και με **πειράματα** (μετρούμε τη διάρκεια ζωής ορισμένων ηλεκτρικών μηχανημάτων). Όλες οι μετρήσεις γίνονται σύμφωνα με κάποια κλίμακα. Εκτός από τις δύο που αναφέραμε, έχουμε την **κλίμακα διαστήματος** και την **κλίμακα πηλίκου**. Έτσι αν ενδιαφερόμαστε για τις επιδόσεις βαθμολογιών μαθητών, και χαρακτηρίζαμε αυτές σε κακή, μέτρια, καλή, πολύ καλή, άριστη, αντιστοιχώντας σ’ αυτές τους αριθμούς 1, 2, 3, 4, 5, θα λέμε ότι βρισκόμαστε σε μια κλίμακα διαστήματος. Επίσης, αντί για τις επιδόσεις μαθητών μπορεί σε θέματα έρευνας αγοράς να ενδιαφερόμασταν για τον βαθμό προτίμησης ή ικανοποίησης σ’ ένα προϊόν. Ο παραπάνω τρόπος βαθμολόγησης στην βιβλιογραφία είναι γνωστός με το όνομα κλίμακα Lickert. Βέβαια μεταξύ της κακής επίδοσης 1 και της μέτριας 2 υπάρχει διαφορά μια μονάδα, που είναι η ίδια και μεταξύ της πολύ καλής 4 και της άριστης 5. Φυσικά δεν μπορούμε να ισχυρισθούμε ότι αυτές με την άριστη βαθμολογία 5 είναι πέντε φορές καλύτερες απ’ αυτές με τη βαθμολογία 1. Για παράδειγμα αν πολλαπλασιάσουμε κάθε επίδοση με 2 και κατόπιν προσθέσουμε τον αριθμό 10 τότε έχουμε την εξής αντιστοιχία

12	κακή
14	μέτρια
16	καλή
18	πολύ καλή
20	άριστη

Οι διαφορές μεταξύ δύο διαδοχικών επιδόσεων παραμένουν ίδιες, αλλά ο λόγος μεταξύ αυτών διαφέρουν. Έτσι ο λόγος 20 προς 12 είναι διαφορετικός απ' αυτόν του 5 προς 1. Η κλίμακα διαστήματος δεν περιέχει κάποια φυσική αρχή και μπορούμε να δεχθούμε σαν τέτοια μια οποιαδήποτε τιμή. Ανάλογο παράδειγμα μπορούμε να δώσουμε με την μέτρηση της θερμοκρασίας. Ως γνωστόν η θερμοκρασία εκφράζεται σε βαθμούς Κελσίου (C) ή σε βαθμούς Φάρεναϊτ (F). Έτσι σε 0°C αντιστοιχούν 32°F που είναι διαφορετικό από 0°F. Σε κάθε περίπτωση όμως το 0° δεν σημαίνει ότι δεν υπάρχει θερμοκρασία, και λέμε ότι το 0 δεν έχει φυσική σημασία για να αποτελέσει μια μοναδική αρχή. Η κλίμακα που περιέχει το φυσικό 0 ονομάζεται κλίμακα πηλίκου. Έτσι μετρώντας την ηλικία των ατόμων μπορούμε να εκφράσουμε με αριθμούς ότι το A άτομο έχει διπλάσια ηλικία από το B. Η τελευταία είναι η ισχυρότερη όλων των κλιμάκων και η πλέον εύχρηστη. Κάθε μέτρηση της κλίμακας πηλίκου, μπορεί να θεωρηθεί σαν μέτρηση της κλίμακας διαστήματος, η της κλίμακας διαστήματος σαν μέτρηση της τακτικής κλίμακας, και της τελευταίας η μέτρηση σαν της ονομαστικής κλίμακας.

Σύμφωνα με τα παραπάνω εφόσον με μία μεταβλητή X παριστάνουμε ένα γνώρισμα, τότε αυτή μπορεί να είναι του διακριτού ή του συνεχούς τύπου, όταν και το αντίστοιχο γνώρισμα είναι διακριτό ή συνεχές.

Το πλήθος των δυνατών τιμών μιας μεταβλητής του διακριτού τύπου είναι διακριτές αριθμήσιμες τιμές, δηλ. η μεταβλητή παίρνει τιμές $\{0, 1, \ldots, n, \ldots\}$ ή πραγματικές τιμές $\{\alpha_1, \alpha_2, \ldots \alpha_n, \ldots\}$ με α_1, α_2, κ.ο.κ. πραγματικούς αριθμούς σε σειρά διάταξης. Στην περίπτωση που η μεταβλητή μας είναι του συνεχούς τύπου το πλήθος των δυνατών τιμών της δεν μπορεί να μετρηθεί αλλά μπορούμε να ομαδοποιήσουμε τις μετρήσιμες σύμφωνα με το αριθμητικά διαστήματα που τις περικλείει. Έτσι στην περίπτωση που είχαμε τη μεταβλητή X του εισοδήματος, μπορούμε να διακρίνουμε εισοδήματα σύμφωνα με τα διαστήματα (ή κλάσεις) [0, 600), [6.000, 12.000), [12.000, 20) όπου προφανώς μας ικανοποιεί η γνώση του διαστήματος όπου η μεταβλητή X θα παίρνει διάφορες τιμές. Πολλές φορές την ομαδοποίηση των μετρήσεών μας μέσω αριθμητικών διαστημάτων την εφαρμόζουμε και σε περιπτώσεις όπου η μεταβλητή μας είναι διακριτή και με μεγάλο πλήθος δυνατών τιμών. Για παράδειγμα αν εξετάζουμε την ηλικία

των ατόμων μιας χώρας σε έτη, πιθανόν να έχουμε ηλικίες από 1 έτους έως100 ετών. Έχουμε 100 δυνατές τιμές, και ενδιαφερόμαστε να καταγράψουμε τα άτομα του πληθυσμού μας σύμφωνα με τις 100 τιμές. (Πόσα άτομα είναι ηλικίας 1 έτους; κ.ο.κ.). Η έρευνά μας όμως μπορούσε να γίνει με πιο ουσιαστικό και σύντομο τρόπο, αν ενδιαφερόμασταν και ρωτούσαμε πόσων ατόμων οι ηλικίες είναι στο διάστημα $[1, 11)$, $[11, 21)$, … και $[91, 101)$.

Συμπερασματικά λοιπόν αν η μεταβλητή μας είναι του διακριτού τύπου και με μεγάλο πλήθος τιμών ή του συνεχούς τύπου, τότε μας αρκεί να έχουμε την γνώση των αριθμητικών διαστημάτων εντός των οποίων θα περικλείονται οι μετρήσεις μας, και θα μιλάμε για ομαδοποιημένες μετρήσεις ή παρατηρήσεις (σε διαστήματα).

Στην περίπτωση που η μεταβλητή μας είναι διακριτή αλλά με μικρό πλήθος δυνατών τιμών (στην Στατιστική θεωρείται μικρό το πλήθος αν είναι μικρότερο του 20), μπορούμε να ομαδοποιήσουμε τις μετρήσεις σύμφωνα με τις δυνατές τιμές. Στο σημείο αυτό μπορούμε να αναφέρουμε ότι όλες οι μεταβλητές του διακριτού τύπου που θα εξετάσουμε στη συνέχεια θα είναι με μικρό δυνατό πλήθος τιμών, επειδή αν ήταν με μεγάλο πλήθος τιμών θα τη μελετούσαμε με ανάλογο τρόπο όπως τη συνεχή μεταβλητή. Η συζήτηση πότε ένα μέγεθος μπορεί να θεωρηθεί συνεχής ή διακριτό μπορεί να επεκταθεί ακόμη και σε φιλοσοφικό επίπεδο. Εμείς αναφερθήκαμε στις πλέον πρακτικές πλευρές του θέματος.

Είδη Δεδομένων

Έστω ένα ποσοτικό ή ποιοτικό γνώρισμα X που παρατηρείται σε n στατιστικές μονάδες πληθυσμού. Οι αριθμοί $x_1, \dots, x_n$ που προκύπτουν ονομάζονται **παρατηρήσεις ή δεδομένα ή μετρήσεις**. Το x_i θα σημαίνει την παρατήρηση του γνωρίσματος X στην $i^{η}$ στατιστική μονάδα. Τα δεδομένα μας μπορεί να είναι: **Διαστρωματικά**, αυτά που λαμβάνονται από μία μεταβλητή την ίδια χρονική στιγμή και πάνω σε διαφορετικές μονάδες (άτομα, χώρες, εταιρίες, κ.λπ.) τα **χρονολογικά**, αυτά που λαμβάνονται από μία μεταβλητή σε διαδοχικούς χρόνους. Όταν παρατηρούμε τα εισοδήματα 1000 νοικοκυριών σε μια χρονική στιγμή από πέντε διαφορετικές χώρες, τότε αυτά τα δεδομένα είναι διαστρωματικά. Όταν παρατηρούμε το εισόδημα ενός νοικοκυριού σε διαδοχικούς χρόνους, τότε αυτά τα δεδομένα είναι χρονολογικά. Επίσης μπορεί να υπάρξει συνδυασμός των παραπάνω. Παρατηρούμε τα εισοδήματα των 1000 νοικοκυριών από πέντε χώρες σε διαδοχικούς χρόνους.

Αν τα δεδομένα λαμβάνονται με άμεσο τρόπο από τον ερευνητή θα λέμε ότι διαθέτουμε **πρωτογενή δεδομένα**, ενώ αν λαμβάνονται με έμμεσο τρόπο θα

λέμε ότι διαθέτουμε δ**ευτερογενή δεδομένα**. Τα δευτερογενή δεδομένα παρέχονται από διάφορες βάσεις δεδομένων που διατηρούν διάφοροι δημόσιοι και ιδιωτικοί οργανισμοί. Σε εθνικό επίπεδο η Στατιστική Υπηρεσία της χώρας, η Τράπεζα της Ελλάδας, δίνουν ανά τακτά χρονικά διαστήματα στοιχεία γύρω από διάφορα κοινωνοικοοικονομικά μεγέθη, όπως η ανεργία, ο πληθωρισμός, το εισόδημα, το ύψος καταθέσεων, η εκπαίδευση, το εμπόριο κ.ο.κ. Σε διεθνές επίπεδο ο Ο.Η.Ε εκδίδει ανά έτος, ετήσιο Στατιστικό βιβλίο, που περιέχει δεδομένα, σχετικά με τον πληθυσμό της γης, τη μετανάστευση, την εκπαίδευση, τη βιομηχανική παραγωγή, το Εθνικό Ακαθάριστο Εισόδημα των χωρών κ.ο.κ. Σε κάθε ιστοσελίδα ενός πανεπιστημίου, στη θέση βιβλιοθήκη διατίθενται διάφορες ηλεκτρονικές βάσεις δεδομένων. Η ιστοσελίδα για τις βάσεις δεδομένων του Πανεπιστήμιο Μακεδονίας είναι: http//www.lib.uom.gr/dbases/greek/

Άσκηση 2.1.1: Ποια από τα επόμενα στατιστικά γνωρίσματα είναι διακριτά, και ποια συνεχή;

α) Ο αριθμός των παιδιών σ’ ένα νοικοκυριό.
β) Η μηνιαία κατανάλωση σε ηλεκτρική ενέργεια απ’ ένα νοικοκυριό.
γ) Ο αριθμός των πλοίων που καταφθάνουν σ’ ένα λιμάνι.
δ) Η τιμή του χρυσού.

Άσκηση 2.1.2: Ο επόμενος πίνακας μας δίνει πληροφορίες γύρω από πέντε άτομα.

Φύλο	Μισθός	Εκπαίδευση	Έτη προϋπηρεσίας
Α (άνδρας)	1.200 €	Πανεπιστημιακή	8
Γ (γυναίκα)	900 €	Μέση	4
Α	1.100 €	Μέση	5
Γ	1.600 €	Πανεπιστημιακή	10
Γ	950 €	Μέση	6

α) Ποια από τα παραπάνω γνωρίσματα είναι ποιοτικά, και ποια ποσοτικά.
β) Σε ποια στατιστικά γνωρίσματα θα χρησιμοποιήσετε την τακτική κλίμακα και σε ποια την ονομαστική.

Άσκηση 2.1.3: Ποια από τα επόμενα στατιστικά γνωρίσματα είναι διακριτά, και ποια συνεχή;

Άσκηση 2.1.4: Ποια από τα επόμενα στατιστικά γνωρίσματα μας δίνουν χρονολογικά δεδομένα ή διαστρωματικά.

α) Η τιμή μιας μετοχής ανά μέρα τα δύο τελευταία χρόνια.

β) Ο αριθμός των τηλεφωνημάτων που έγινε από κάθε νοικοκυριό χθες.

γ) Η τιμή των σιτηρών τα τελευταία 50 χρόνια.

δ) Οι καταθέσεις σε μια τράπεζα που κάνουν 60 άτομα.

2.2 Διακριτές Κατανομές Συχνοτήτων

Σε πολλές έρευνες μπορούμε να έχουμε δεδομένα από ένα διακριτό μέγεθος με μικρό πλήθος δυνατών τιμών. Αν ο χρόνος X της διεκπεραίωσης σε ημέρες μιας παραγγελίας είναι από 1 ημέρα έως και 5 ημέρες, τότε το πλήθος των δυνατών τιμών της μεταβλητής χρόνος είναι 5, ένας σχετικά μικρός αριθμός. Αν προέκυπτε ένας αριθμός για το πλήθος των δυνατών τιμών μεγαλύτερος του 20 θα μιλούσαμε για έναν σχετικά μεγάλο αριθμό. Στα διακριτά δεδομένα, κυρίως ενδιαφερόμαστε να γνωρίζουμε πόσες φορές επαναλαμβάνεται μια δυνατή τιμή μέσα στα δεδομένα μας, επειδή η γνώση του γεγονότος αυτού μας δίνει τον τρόπο λήψης των μετρήσεων του δείγματός μας.

Παράδειγμα 2.2.1: Παρατηρήθηκε ότι ο κύριος προμηθευτής μιας επιχείρησης για 50 παραγγελίες χρειάσθηκε τους εξής χρόνους (σε ημέρες) διεκπεραίωσης των.

4 5 4 1 5 4 3 4 5 6 6 5 5 4 7 4 6 5 6 4 5 4 7 5 5
6 7 3 7 6 6 7 4 5 4 7 7 5 5 5 5 6 6 4 5 2 5 4 7 5 .

Για το γνώρισμα X [= χρόνος παράδοσης σε ημέρες] παρατηρήθηκαν τα εξής:

$$x_1 = 4\,,\quad x_2 = 5\,,\quad x_3 = 3\,,\quad x_4 = 1\,,\ \ldots,\quad x_{49} = 7\,,\quad x_{50} = 5.$$

Το ενδιαφέρον μας εδώ εστιάζεται στο να βγάλουμε κάποιο συμπέρασμα, στο κατά πόσο γρήγορα διεκπεραιώνονται οι παραγγελίες από τον εν λόγω προμηθευτή. Η διάσπαρτη γνώση των δεδομένων μας δεν μας βοηθά αποτελεσματικά σ' αυτό το σκοπό. Η γνώση όμως το πόσο συχνά εμφανίζεται το 1, το 2, κ.τ.λ. στα δεδομένα μας, μας πληροφορούν καλύτερα για το πόσο γρήγορος είναι ο προμηθευτής μας.

Για να έχουμε μια καλύτερη εικόνα των δεδομένων είναι απαραίτητο να γνωρίζουμε με τι τρόπο κατανέμονται αυτές οι τελευταίες τιμές. Έτσι προχωρούμε στην ανάπτυξη των επομένων εννοιών.

2.3 Απόλυτη και Σχετική Συχνότητα

Σε μια λίστα n δεδομένων θεωρούμε τις $\alpha_1, \ldots, \alpha_k$ δυνατές τιμές, που τις θεωρούμε εξ αρχής διατεταγμένες

$$\alpha_1 < \alpha_2 < \ldots < \alpha_k$$

Το πλήθος εμφάνισης του α_j καλείται **απόλυτη συχνότητα** της j-οστής δυνατής τιμής και συμβολίζεται με $h(\alpha_j)$. Ενώ το πηλίκο $f(\alpha_j) = \frac{1}{n} h(\alpha_j)$ θα καλείται **σχετική συχνότητα**. Ισχύουν τα ακόλουθα:

$$\sum_{j=1}^{k} h(\alpha_j) = n \ , \quad \text{και} \quad \sum_{j=1}^{k} f(\alpha_\text{j}) = 1 \, . \tag{2.3.1}$$

Σύμφωνα με τις τελευταίες έννοιες τα δεδομένα του παραδείγματος 2.2.1 μπορούν να παρασταθούν υπό μορφή **πίνακα συχνοτήτων**. Έτσι έχουμε το ακόλουθο:

Παράδειγμα 2.3.1: Να βρεθούν οι απόλυτες και σχετικές συχνότητες των τιμών της μεταβλητής X του Παραδείγματος 2.2.1.

Δυνατές τιμές α_j	α_1=1	α_2=2	α_3=3	α_4=4	α_5=5	α_6=6	α_7=7
Απόλυτη Συχνότητα $h(\alpha_j)$	1	1	2	12	17	9	8
Σχετική Συχνότητα $f(\alpha_j)$	$\frac{1}{50}$	$\frac{1}{50}$	$\frac{2}{50}$	$\frac{12}{50}$	$\frac{17}{50}$	$\frac{9}{50}$	$\frac{8}{50}$

Η σχετική συχνότητα εκφράζεται πολλές φορές σαν ποσοστό, δηλ., 100% $f(\alpha_j)$. Στην περίπτωσή μας η σχετική συχνότητα της τιμής 6 είναι το $\frac{9}{50}$ ή 2%.

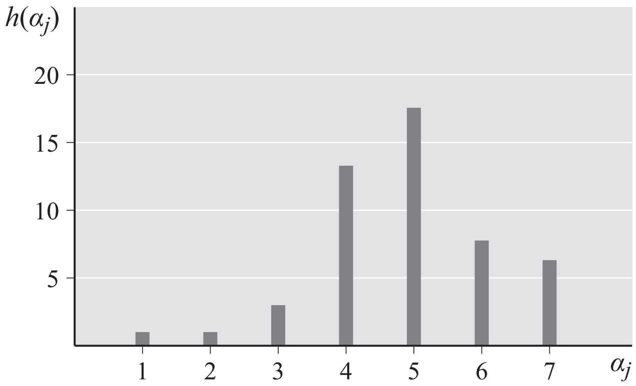
Οι συχνότητες (απόλυτες ή σχετικές) για τις διακριτές τιμές της μεταβλητής X ορίζουν την **διακριτή κατανομή συχνότητας** της μεταβλητής.

Οι συνηθισμένες γραφικές μέθοδοι των κατανομών συχνοτήτων είναι οι εξής:

- **Πίνακας Συχνοτήτων**
- **Ακιδωτά Διαγράμματα**
- **Κυκλικά Διαγράμματα**
- ή στην περίπτωση ομαδοποιημένων δεδομένων τα **ιστογράμματα**.

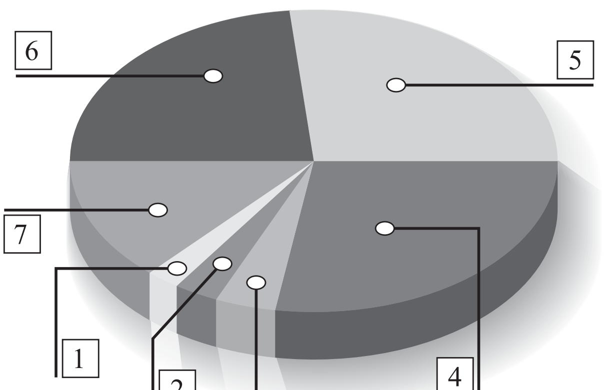
Ο τρόπος παρουσίασης των δεδομένων του Παραδείγματος 2.3.1 έγινε με τον πίνακα συχνοτήτων.

Ένα ακιδωτό διάγραμμα λαμβάνεται σ' ένα ορθογώνιο σύστημα συντεταγμένων, τοποθετώντας τα k σημεία $(\alpha_j, h(\alpha_j)$ ή $f(\alpha_j))$ σ' αυτό. Έτσι το ακιδωτό διάγραμμα του Παραδείγματος 2.3.1, είναι το σχήμα 2.3.1.



Σχήμα 2.3.1

Σ' ένα κυκλικό διάγραμμα θα θεωρούμε τις συχνότητες σαν εμβαδά κυκλικών τομέων, όπου το κάθε εμβαδόν θα είναι ανάλογο της συχνότητας που θα παριστάνει. Το κυκλικό διάγραμμα του Παραδείγματος 2.3.1 δίνεται από το σχήμα 2.3.2.



Σχήμα 2.3.2

Άσκηση 2.3.1: Σ' ένα εργοστάσιο ρωτήθηκαν 40 εργαζόμενοι με τι τρόπο πηγαίνουν στη δουλειά τους. Είχαμε τις ακόλουθες παρατηρήσεις

1 1 2 2 4 3 5 2 2 5 2 4 1 1 2 2 1 2 1 2 2 4 2 5 4 2
2 2 2 2 5 1 1 2 3 1 2 2 1 2

όπου

1 = Δημόσιο μέσο, 2 = Ιδιωτικό, 3 = Μηχανή,
4 = ποδήλατο, 5 = οδοιπορικώς.

Να κατασκευασθούν: το ακιδωτό διάγραμμα, το κυκλικό διάγραμμα, καθώς και ο πίνακας συχνοτήτων.

Άσκηση 2.3.2: Να σχεδιασθεί το κυκλικό διάγραμμα των επόμενων δεδομένων:

Μέσα Μηνιαία Έξοδα σε ευρώ.

Είδος	**1988**	**1990**
Τρόφιμα	640	840
Ρούχα	200	280
Ενοίκιο	420	520
Διάφορα	170	300

2.4 Συνεχείς Κατανομές Συχνοτήτων – Ιστογράμματα

Σε πολλές στατιστικές αναλύσεις η απαρίθμηση των δυνατών τιμών μιας μεταβλητής δεν είναι δυνατή ή και αν είναι δεν έχει νόημα, επειδή

α) Η μεταβλητή είναι συνεχής.

β) Το πλήθος των δυνατών τιμών είναι πολύ μεγάλο, όπως στο παράδειγμα με την ηλικία των παιδιών.

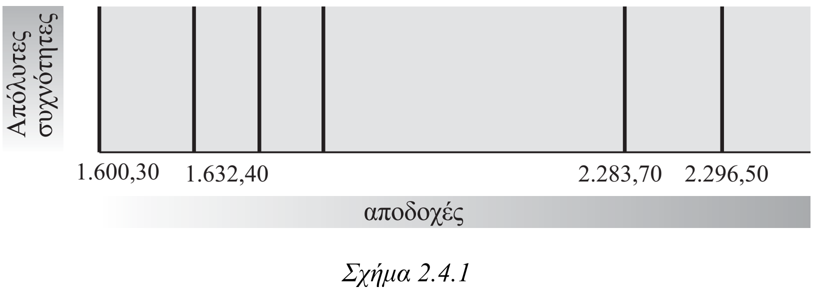
γ) Και αν ακόμη χρησιμοποιούσαμε τον πίνακα συχνοτήτων ή το ακιδωτό διάγραμμα δεν θα είχαμε καμία σαφή εικόνα των δεδομένων μας.

Τα προηγούμενα διαφωτίζονται με το επόμενο παράδειγμα.

Παράδειγμα 2.4.1: Οι μηνιαίες αποδοχές 100 υπαλλήλων μιας εταιρείας το 1987, είναι:

1.600,30	1.819,20	1.920,30	1.970,30	2.091,20
1.632,40	1.819,80	1.926,90	1.979,40	2.091,20
1.673,90	1.821,70	1.928,30	1.982,80	2.096,20
1.685,00	1.825,50	1.929,90	1.996,00	2.097,40
1.711,00	1.826,80	1.939,90	2.000,30	2.100,10
1.715,00	1.839,10	1.939,40	2.006,50	2.109,20
1.722,70	1.840,30	1.940,70	2.017,40	2.111,20
1.738,70	1.848,80	1.942,30	2.025,80	2.119,70
1.742,00	1.852,20	1.942,30	2.039,50	2.140,80
1.751,00	1.857,30	1.943,50	2.043,30	2.149,20
1.751,60	1.859,60	1.948,80	2.049,10	2.149,60
1.768,20	1.872,20	1.950,10	2.049,70	2.158,90
1.771,80	1.883,10	1.952,60	2.050,40	2.159,70
1.782,20	1.886,30	1.952,60	2.050,90	2.162,20
1.784,30	1.892,10	1.952,60	2.058,20	2.186,80
1.789,20	1.893,60	1.958,70	2.069,40	2.198,70
1.790,40	1.901,20	1.959,20	2.078,60	2.223,70
1.791,70	1.913,50	1.960,80	2.082,70	2.248,40
1.800,50	1.913,70	1.969,20	2.083,50	2.269,80
1.817,90	1.914,90	1.969,99	2.083,90	2.296,50

Τότε το ακιδωτό διάγραμμα στο σχήμα 2.4.1 δεν μας δίνει καμιά πληροφορία για τις αποδοχές των υπαλλήλων. Απλά συμπεραίνουμε ότι κάθε τιμή εμφανίζεται μια φορά (βλ. Παράδειγμα 2.4.2).



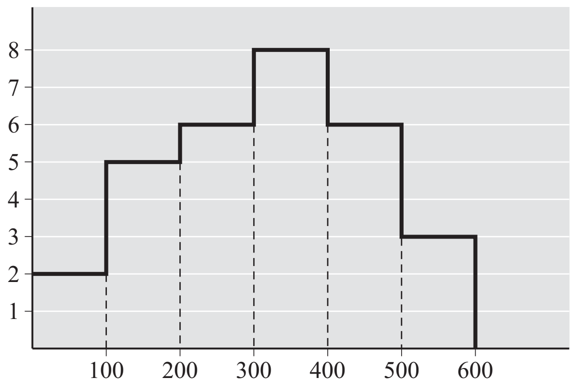
Σχήμα 2.4.1

Σ’ αυτές τις περιπτώσεις ενδείκνυται να χωρίσουμε τον άξονα της μεταβλητής X σε k διαδοχικά διαστήματα, όπου το j-οστό διάστημα είναι το $[\alpha_{j-1}, \alpha_j)$, και η απόλυτη συχνότητα $h(\alpha_j)$ του διαστήματος $[\alpha_{j-1}, \alpha_j)$ ορίζεται σαν το πλήθος των παρατηρήσεων ή των μετρήσεων ή των δεδομένων που περιλαμβάνονται σ’ αυτό. Κατόπιν δημιουργούμε ορθογώνια παραλληλόγραμμα πάνω σε κάθε διάστημα εμβαδού ανάλογου της απόλυτης συχνότητας του διαστήματος. Η γραφική παράσταση που προκύπτει ονομάζεται **Ιστόγραμμα**, ενώ οι συχνότητες (απόλυτες ή σχετικές) των διαστημάτων της μεταβλητής X ορίζουν την **συνεχή κατανομή συχνότητας**.

Έτσι αν είχαμε τον ακόλουθο πίνακα συχνοτήτων:

Διαστήματα	Απόλυτες Συχνότητες
[0, 100)	2
[100, 200)	5
[200, 300)	6
[300, 400)	8
[400, 500)	6
[500, 600)	3

λαμβάνουμε το παρακάτω Ιστόγραμμα (σχ. 2.4.2)



Σχήμα 2.4.2

Τα ερωτήματα που δημιουργούνται κατά τη κατασκευή ενός Ιστογράμματος είναι τα εξής: Πρώτον, ποιο είναι το κατάλληλο μήκος των διαστημάτων, και αν αυτά πρέπει να είναι ίσα. Πόσα διαστήματα πρέπει να λάβουμε; Από που πρέπει

να ξεκινήσουμε τη σκιαγράφησή του;

Στην περίπτωση που δεχθούμε ότι έχουμε διαστήματα ίσου μήκους, τότε το πλήθος k αυτών πρέπει να είναι ένας αριθμός μεταξύ 5 και 20 ή να ικανοποιεί τη σχέση $k = 1 + 3{,}3 \ln n$, όπου n το πλήθος των δεδομένων μας. Το μήκος ℓ του διαστήματος δίνεται από τη σχέση

$$\ell = \frac{\text{μέγιστο}\{x_i\} - \text{ελάχιστο}\{x_i\}}{k}.$$

Σε αρκετές περιπτώσεις λαμβάνουμε διαστήματα άνισου μήκους, επειδή υπάρχει ο κίνδυνος στην περίπτωση ίσων διαστημάτων να υπάρξουν τέτοια που να μην συμπεριλαμβάνουν καμία μέτρηση. Για παράδειγμα αν μελετάμε τα ετήσια εισοδήματα φορολογουμένων με εισόδημα κάτω των 250.000 € και θεωρήσουμε διαστήματα ανά 10.000 €, υπάρχει κίνδυνος επιλέγοντας ένα διάστημα π.χ. το διάστημα [120.000, 130.000) να μην υπάρχει φορολογούμενος που το εισόδημά του να εμπίπτει σε αυτό το διάστημα.

Παράδειγμα 2.4.2: Να κατασκευασθεί το Ιστόγραμμα του Παραδείγματος 2.4.1.

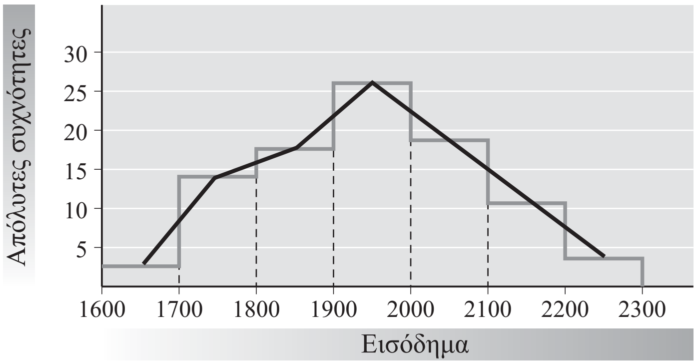
Λύση: Από τη σχέση $k = 1{+}3{,}3 \ln n$, έχουμε για $n{=}100$, ότι $k \approx 7$, ενώ

$$\ell = \frac{2.296{,}50 - 1.600{,}30}{7} = 99{,}5 \approx 100 .$$

Αν θεωρήσουμε σαν αρχή το 1.600, τότε λαμβάνουμε τον εξής πίνακα συχνοτήτων

Διαστήματα	**Απόλυτες Συχνότητες**
[1.600, 1.700)	4
[1.700, 1.800)	14
[1.800, 1.900)	18
[1.900, 2.000)	28
[2.000, 2.100)	20
[2.100, 2.200)	12
[2.200, 2.300)	4

Με βάση τον πίνακα συχνοτήτων έχουμε την εξής γραφική παράσταση του Ιστογράμματος (σχ. 2.4.3).



Σχήμα 2.4.3

Αν ενώσουμε τα μέσα των πλευρών των ορθογωνίων που είναι παράλληλες προς τον άξονα των εισοδημάτων σχηματίζουμε το λεγόμενο **Πολύγωνο Συχνοτήτων**.

Αν αντί για τις απόλυτες συχνότητες θεωρούσαμε τις σχετικές συχνότητες, τότε το Ιστόγραμμα που θα προκύψει θα έχει εμβαδόν ίσον με μονάδα, όταν το μήκος κάθε διαστήματος θεωρείται μοναδιαίο.

Άσκηση 2.4.1: Είκοσι φοιτητές ρωτήθηκαν για τα μηνιαία έξοδά τους και είχαμε

1000, 580, 520, 350, 620, 800, 120, 600, 550, 420, 470, 200, 560, 480, 1000, 600, 1150, 800, 250, 650.

Να σχηματίσετε το Ιστόγραμμα, καθώς και το Πολύγωνο Συχνοτήτων.

2.5 Αθροιστικές Κατανομές Συχνοτήτων

Πολλές φορές μας ενδιαφέρει να γνωρίζουμε πόσες παρατηρήσεις είναι ίσες, ή μικρότερες ή μεγαλύτερες από κάποιον αριθμό. Επίσης, ποιο ποσοστό παρατηρήσεων είναι ίσο ή μικρότερο ή μεγαλύτερο από κάποιον αριθμό.

Οι απαντήσεις σ' αυτά τα ερωτήματα δίδονται με τη βοήθεια της αθροιστικής κατανομής συχνοτήτων. Η συνάρτηση που αθροίζει εκείνες τις απόλυτες συχνότητες των δυνατών τιμών α_j που είναι μικρότερες του x συμβολίζεται με $H(x)=\sum_{j:\alpha_j\le x} h(\alpha_j)$ και ονομάζεται **απόλυτη αθροιστική κατανομή συχνοτήτων**, ενώ η $F(x)=\dfrac{H(x)}{n}$ **σχετική αθροιστική κατανομή συχνοτήτων**. Αν η

μεταβλητή μας είναι του διακριτού τύπου δημιουργούμε τον ακόλουθο πίνακα για τον εύκολο υπολογισμό των συχνοτήτων και των αθροιστικών κατανομών.

Τιμή Μεταβλητής	**Απόλυτη Συχνότητα**	**Σχετική Συχνότητα**	**Απόλυτη Αθροιστική Κατανομή**	**Σχετική Αθροιστική Κατανομή**
α_1	$h_1 = h(\alpha_1)$	$f_1 = f(\alpha_1)$	$H_1=H(\alpha_1)=h_1$	$F_1 = f_1$
α_2	$h_2 = h(\alpha_2)$	$f_2 = f(\alpha_2)$	$H_2=H(\alpha_2)=h_1+h_2$	$F_2 = f_1+f_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
α_k	$h_k = h(\alpha_k)$	$f_k = f(\alpha_k)$	$H_k=H(\alpha_k)= h_1+ h_2+\ldots+ h_k$	$F_k = f_1+\ldots+f_k$

Με εντελώς ανάλογο τρόπο εργαζόμαστε αν η μεταβλητή μας είναι του συνεχούς τύπου. Αντικαθιστούμε τις δυνατές τιμές $\alpha_1, \alpha_2, \ldots, \alpha_k$ με τα διαστήματα $[\alpha_1, \alpha_2)$, $[\alpha_2, \alpha_3), \ldots$ και $[\alpha_{k-1}, \alpha_k)$, και υπολογίζουμε τις απόλυτες συχνότητες $h_1, \ldots h_k$ μετρώντας το πλήθος των μετρήσεων που περιλαμβάνονται στα παραπάνω διαστήματα, αντίστοιχα. Όμοια για τις σχετικές συχνότητες, και τις αθροιστικές κατανομές κάνουμε τις ανάλογες τροποποιήσεις. Τέλος η αθροιστική σχετική συχνότητα εκφράζεται και ως ποσοστό, δηλ. 100% $F(x)$, που δείχνει το ποσοστό των μετρήσεων που δεν ξεπερνούν το x.

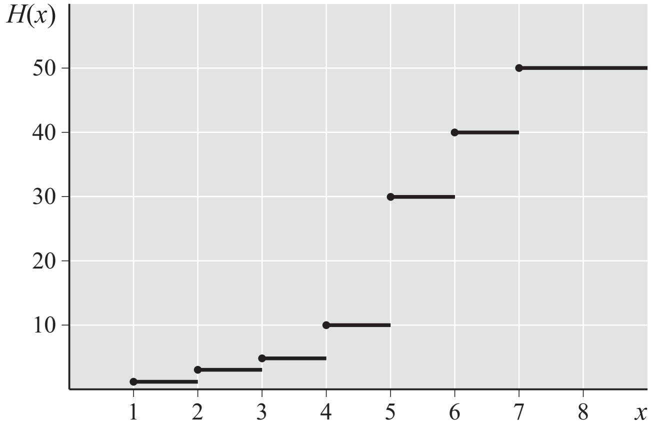
Παράδειγμα 2.5.1: Να βρεθεί η αθροιστική κατανομή συχνοτήτων (απόλυτη και σχετική) του Παραδείγματος 2.2.1.

Λύση: Η απόλυτη κατανομή συχνότητας $H(x)$ συμβολίζει το πλήθος των παραγγελιών που χρειάσθηκαν χρόνο λιγότερο ή ίσο με x. Έτσι οι τιμές της θα εξαρτώνται ανάλογα σε ποιο διάστημα θα κυμαίνεται το x.

$$H(x)=\begin{cases} 0 & x<1 \\ 1 & 1\le x<2 \\ 2 & 2\le x<3 \\ 4 & 3\le x<4 \\ 16 & 4\le x<5 \\ 33 & 5\le x<6 \\ 42 & 6\le x<7 \\ 50 & 7\le x \end{cases}$$

(καμία παραγγελία δεν έγινε σε χρόνο μικρότερο του 1)
(μία παραγγελία έγινε σε χρόνο μεγαλύτερο ή ίσο του 1 και μικρότερο του 2)

Η γραφική παράσταση της $H(x)$ δίνεται στο σχήμα 2.5.1.



Σχήμα 2.5.1

Άσκηση 2.5.1: Σε μια ασφαλιστική εταιρεία 100 άτομα κάνουν ασφάλεια ζωής με τον εξής τρόπο.

Ηλικία	**Ποσά**	**Απόλυτες συχνότητες**	**Σχετικές συχνότητες**
[20, 30)	70.000	26	$\frac{26}{100}$
[30, 40)	60.000	33	$\frac{33}{100}$
[40, 50)	40.000	21	$\frac{21}{100}$
[50, 60)	25.000	14	$\frac{14}{100}$
[60, 70)	10.000	6	$\frac{6}{100}$

1) Να γίνουν τα ιστογράμματα και τα ακιδωτά διαγράμματα.
2) Ποιο ποσοστό ασφαλισμένων είναι 50 ετών και γηραιότερο;
3) Ποιο ποσοστό είναι ασφαλισμένο για 40.000 και λιγότερο;
4) Να βρεθούν οι αθροιστικές κατανομές.

Άσκηση 2.5.2: Ποια η αθροιστική κατανομή απόλυτης ή (σχετικής) συχνότητας του Παραδείγματος 2.4.1.

Άσκηση 2.5.3: Οι 150 επιβάτες των λεωφορείων μιας γραμμής ρωτήθηκαν για το χρόνο αναμονής τους στη στάση. Λάβαμε τα εξής δεδομένα:

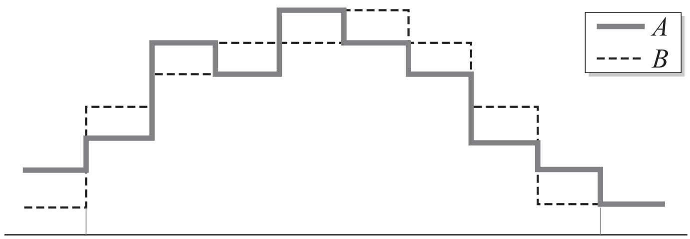
Χρόνος αναμονής: (σε λεπτά)	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12
Απόλυτη συχνότητα:	3	9	18	30	24	21	12	9	3	6	9	6

1) Να υπολογισθεί η αθροιστική κατανομή συχνοτήτων.

2) Η συνάρτηση $A(x)=\sum_{j:\alpha_j>x} h(\alpha_j)=n-H(x)$ ονομάζεται **αφαιρετική κατανομή συχνοτήτων**. Στην άσκηση αυτή δείχνει το πλήθος των επιβατών που ανέμειναν περισσότερο από x λεπτά.

2.6 Μέτρα Θέσης (ή Κεντρικής Τάσης)

Έστω ότι ενδιαφερόμαστε να συγκρίνουμε τους μισθούς των εργαζομένων μεταξύ δύο επιχειρήσεων A και B. Μπορούμε να δημιουργήσουμε τα ιστογράμματα ή τα πολύγωνα συχνοτήτων των μισθών και να τα συγκρίνουμε. Αυτού του είδους η σύγκριση είναι αρκετά δύσκολη, π.χ. ενδέχεται να λαμβάναμε τα ιστογράμματα του σχήματος 2.6.1. Από τη σύγκριση αυτών δεν έχουμε σαφή ένδειξη κατά πόσο οι μισθοί των εργαζομένων στην A επιχείρηση είναι μικρότεροι ή μεγαλύτεροι των αντίστοιχων εργαζομένων στην B επιχείρηση. Στο συγκεκριμένο παράδειγμα το ευκολότερο που μπορούσαμε να κάνουμε είναι να συγκρίνουμε τους μέσους μισθούς.



Σχήμα 2.6.1 Σύγκριση δύο Ιστογραμμάτων

Το Excel είναι το πλέον διαδεδομένο πρόγραμμα που χρησιμοποιείται στη Λογιστική και Χρηματοοικονομική καθώς και στη διαχείριση κάθε μορφής παραγωγής και έρευνα αγοράς. Το λογιστικό φύλλο του Excel δίνει την δυνατότητα στο να διαχειρίζονται δεδομένα που εκφράζονται με αριθμούς (ποσοτικά) αλλά και με λέξεις ή γράμματα (ποιοτικά). Επίσης στο ίδιο φύλλο μπορούμε να έχουμε τα γραφήματα που προκύπτουν από την ανάλυση των δεδομένων μας. Το Excel διαθέτει ένα ικανό σύνολο από εντολές ώστε να δημιουργεί φυσικό υπολογιστικό περιβάλλον για κάθε στατιστική μεθοδολογία.

Ο σκοπός του Οδηγού του *Excel είναι να βοηθηθεί ο σπουδαστής για την καλύτερη εκμάθηση της στατιστικής επιλύοντας παραδείγματα και ασκήσεις από το βιβλίο «Στατιστική Μέθοδολογίαι» με την βοήθεια του. Ο σπουδαστής δεν χρειάζεται να έχει ιδιαίτερες γνώσεις σε υπολογιστές. Από την στιγμή που ανοίγει ένα φύλλο* Excel μπορεί ψηλαφιστά να φθάνει στην επίλυση των προβλημάτων που τον ενδιαφέρουν.

Στο πρώτο κεφάλαιο αναπτύσσονται κάποιες βασικές γνώσεις, όπως εισαγωγή δεδομένων, πράξεων κτλ. Στην συνέχεια τα κεφάλαια που ακολουθούν έχουν αντίστοιχη ανάπτυξη όπως η διάρθρωση του βιβλίου «Στατιστική Μέθοδολογία». Από την στιγμή που λύνεται ένα παράδειγμα από το βιβλίο καλό είναι παράλληλα να επιλύεται και με το Excel.

Κάθε κεφάλαιο περιέχει τον τρόπο ανάπτυξης των ενεργειών που πρέπει να προχωρήσουμε μέσω του Excel για να εφαρμόσουμε μια συγκεκριμένη στατιστική μέθοδο, λυμένα παραδείγματα, αλλά και επεξηγήσεις. Πολλά από τα δεδομένα υπάρχουν στο συνοδευτικό CD.

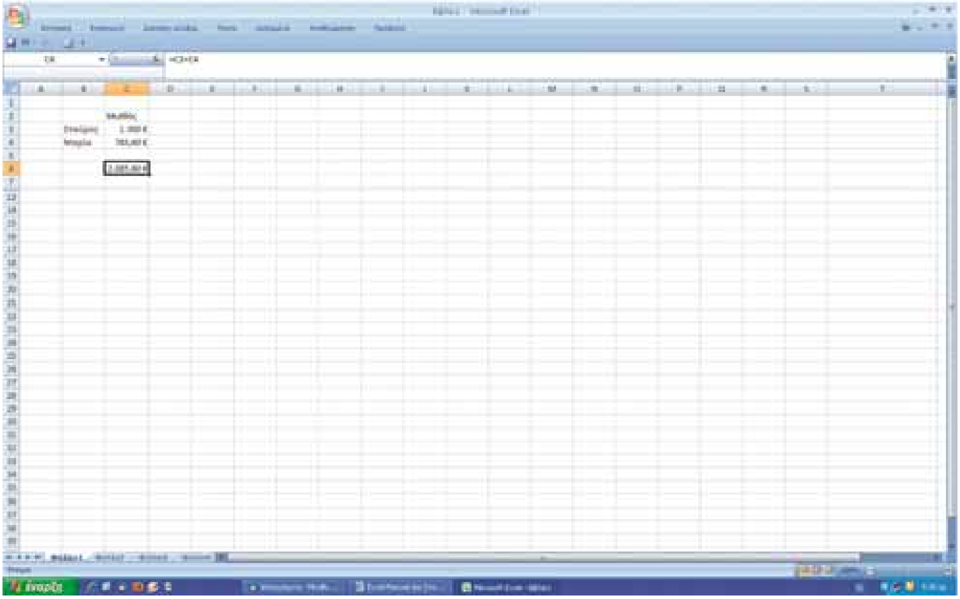
1 Βασικές Αρχές του Excel

Το Excel διαθέτει ένα ισχυρό υπολογιστικό περιβάλλον με μεγάλες δυνατότητες εφαρμογών του στην στατιστική. Μπορεί κανείς σε ένα λογιστικό φύλλο να εισάγει τα δεδομένα του, να τα επεξεργάζεται και συγχρόνως να γράφει κείμενο.

1.1 Περιήγηση και Εισαγωγή Δεδομένων στο Excel

Άνοιγμα Αρχείων

Με την ενεργοποίηση του Excel που γίνεται με το κλικ του κέρσορα του ποντικιού στο εικονίδιο με την ένδειξη **Microsoft Office Excel 2007, ξετυλίγεται στην οθόνη του υπολογιστή το λογιστικό φύλλο του Excel.** Στο άνω μέρος εμφανίζεται ο τίτλος «Βιβλίο 1». Ένα λογιστικό φύλλο περιέχει στήλες με τίτλους σε σειρά τα κεφαλαία γράμματα της Λατινικής Αλφαβήτου. Μετά από την στήλη με το τελευταίο γράμμα Z, οι στήλες φέρνουν σαν τίτλο συνδυασμούς δυο Λατινικών γραμμάτων (π.χ., AA, AB,….) έως την τελευταία στήλη, που είναι η 256η. Επιπλέον ένα λογιστικό φύλλο περιέχει γραμμές, αριθμημένες από το 1 έως το 16.384. Οι 256 στήλες διασταυρώνονται με τις 16.384 γραμμές και δημιουργούν 256*16.834 κελιά. Το κελί A1 εντοπίζεται στην διασταύρωση της πρώτη στήλης και της πρώτης γραμμής, ενώ το κελί Z26 εντοπίζεται στην διασταύρωση της 26η στήλης και 26η γραμμής. Με το άνοιγμα του Excel ενεργοποιείται το κελί A1. Μπορούμε να περιηγηθούμε ελεύθερα ένα λογιστικό φύλλο του Excel, μεταπηδώντας από κελί σε κελί και να εισάγουμε κείμενο ή αριθμούς ή συναρτήσεις απλά πιέζοντας το πλήκτρο Enter . Η περιήγηση μπορεί να γίνει με την μετακίνηση του κέρσορα του ποντικιού ή απλά πιέζοντας τα πλήκτρα ←, ↑, →, and ↓. Παρακάτω δίνεται ένα λογιστικό φύλλο που περιέχει κείμενο, κάποιους αριθμούς και την εισαγωγή συνάρτησης. Στην περίπτωση μας βλέπουμε την συνάρτηση που εισάγουμε κοιτάζοντας στην «γραμμή τύπων» που βρίσκεται στο άνω δεξιό μέρος του λογιστικού φύλλου (εδώ, “=C3+C4”).



Για την εισαγωγή μιας συνάρτησης ξεκινάμε πάντα με την εισαγωγή του συμβόλου του “=”. Ο υπολογισμός του αθροίσματος των μισθών του Σταύρου και της Μαρίας, μπορεί να γίνει άμεσα μέσω της εισαγωγής της συνάρτησης «Sum» από την «γραμμή τύπων» (ή από το πλήκτρο fx) πιέζοντας το πλήκτρο Enter ή ακολουθώντας τα επόμενα βήματα:

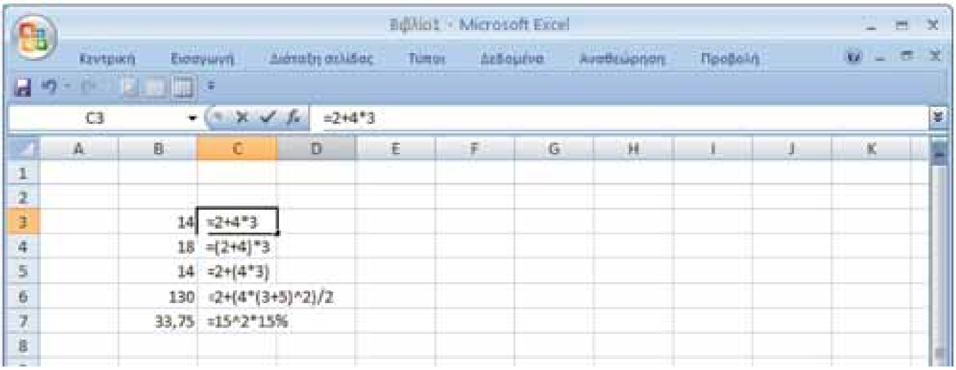
1. Επιλέγουμε το κελί C6.
2. Εισάγουμε το =
3. Εισάγουμε το μισθό του Σταύρου με την λειτουργία της αντιγραφής από το κελί C3 (ή το πληκτρολογούμε εκ νέου)
4. Εισάγουμε το +
5. Εισάγουμε το μισθό της Μαρίας με την λειτουργία της αντιγραφής από το κελί C4 (ή το πληκτρολογούμε το εκ νέου)
6. Πιέζουμε το πλήκτρο Enter.

1.2 Βασικές Πράξεις

Οποιοδήποτε κελί μπορεί να περιέχει έναν μαθηματικό τύπο στον οποίο να σημειώνονται οι βασικές πράξεις με αριθμούς που αναφέρονται σε άλλα κελιά. Αναφέρουμε τις βασικότερες αριθμητικές πράξεις με το Excel.

1. Επιλέγουμε ένα κελί και εισάγουμε το =

2. Χρησιμοποιούμε τα σύμβολα “+” πρόσθεση, “–” για αφαίρεση, “*” για πολλαπλασιασμό, “/” για διαίρεση, τον εκθέτη “^” για να υψώσουμε σε δύναμη και τέλος το “%” για ποσοστά. Μερικά παραδείγματα με τις βασικές πράξεις να σημειώνονται στην στήλη D ενώ τα αποτελέσματα σημειώνονται στο αντίστοιχο κελί της στήλης B.

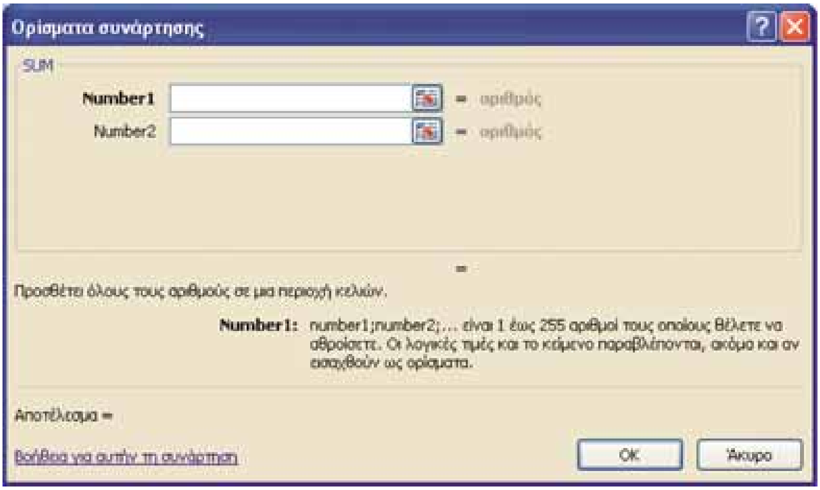


3. Σειρά εκτέλεσης πράξεων:
   α. Η ύψωση σε δύναμη “^” εκτελείται πρώτη
   β. Πολλαπλασιασμός “*”και διαίρεση “/” ακολουθούν. Οι πράξεις στην παρένθεση προηγούνται και μάλιστα από την εσωτερικότερη προς την εξωτερικότερη. Στο επόμενο κλάσμα, 2 / (3 * 4), υπολογίζεται πρώτα ο παρανομαστής και ακολουθεί μετά η διαίρεση
   γ. Η πρόσθεση “+” και η αφαίρεση “–“εκτελούνται τελευταίες. Έτσι η παράσταση 6 + 4 * 2 ^ 3 σε πρώτη φάση γίνεται 6 + 4 * 8 μετά 6 + 32 και τελικά 38
   δ. Προσοχή στο σύμβολο του «μείον» προηγείται όλων των άλλων. Έτσι το –2 ^ 4 είναι σαν να υπολογίζει (–2) ^ 4 που δίνει 16, ενώ αν γράψετε –(2 ^ 4) λαμβάνετε –16
4. Τα ποσοστά λαμβάνονται σαν να διαιρείς με 100. Αν εισάγεις “20%” σ’ ένα κελί λαμβάνεις 0.20 και απλοποιούνται οι πολλαπλασιασμοί.

1.3 Συναρτήσεις του Excel

Το Excel διαθέτει μια πλούσια συλλογή από Μαθηματικές, Οικονομικές Στατιστικές, αλλά και άλλου είδους συναρτήσεις. Από την στιγμή που επιλέγηκε ανά κελί στο λογιστικό φύλλο του Excel μπορούμε να πάμε στην καρτέλα ΤΥΠΟΙ (FORMULAS) και να κάνουμε κλικ στο Εισαγωγή της Συνάρτησης, οπότε επι-

λέγουμε την συνάρτηση που επιθυμούμε. Παρακάτω επιλέγουμε την συνάρτηση Χτυπώντας το SUM, εμφανίζεται:

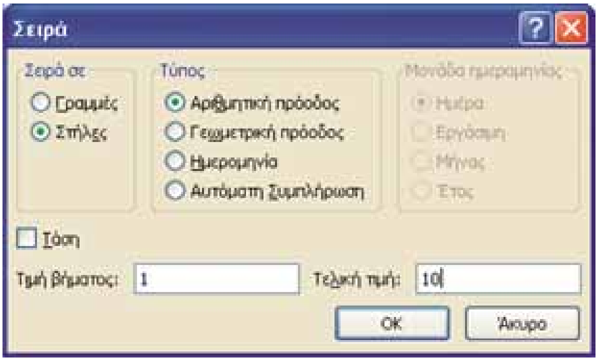


Με βάσει την παραπάνω συνάρτηση αθροίζω τους 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Πριν αθροίσω πρέπει πρώτα να τους εισάγω στο φύλλο Excel. Το τελευταίο μπορεί να γίνει και μέσω εντολών του Excel. Έτσι αναπτύσσουμε το παρακάτω.

1.4 Συμπλήρωση Κελιών με Διαδοχικούς Αριθμούς

Για να τους εισάγω, τους 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, επειδή οι αριθμοί μου είναι διαδοχικοί χρησιμοποιώ Ιδιότητα Γεμίσματος των Κελιών του Excel. Εισάγω τον αριθμό 1 στο D5, από το εικονίδιο **Κεντρική** του μενού, επιλέγω το εικονίδιο κάτω από το **Σ**,

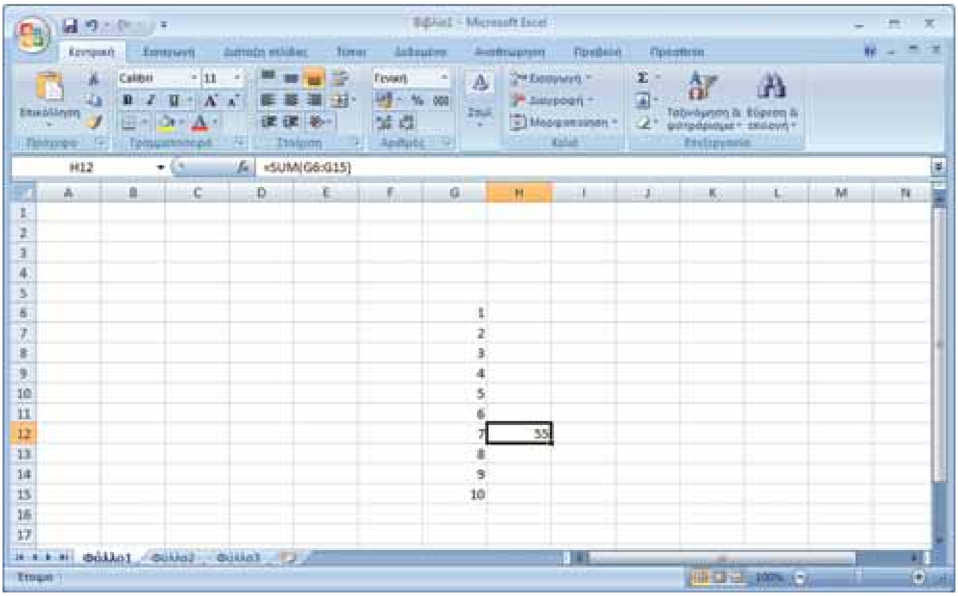
και μέσω του «Σειρά» εισάγω τους 10 αριθμούς μου, όπως υποδεικνύεται παρακάτω.



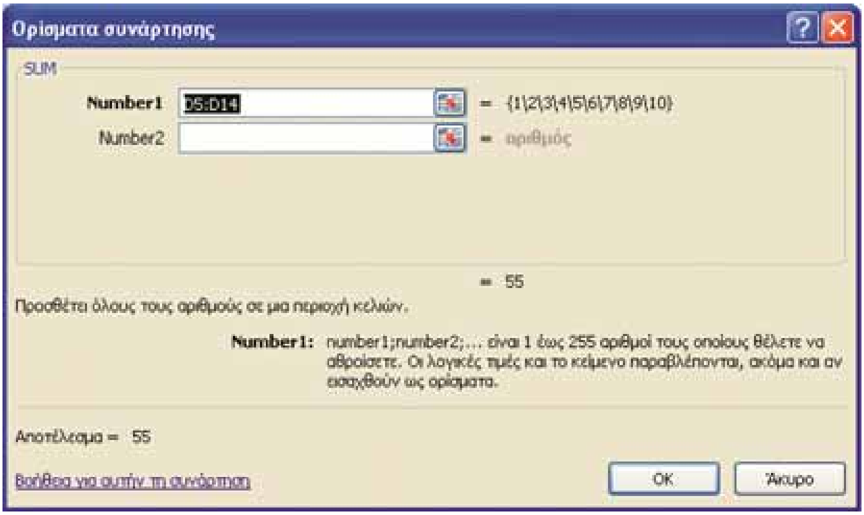
1.5 Υπολογισμός Αθροίσματος

Στη στήλη του D από “D5:D15” εισήχθηκαν οι αριθμοί 1, 2, …, 10.

Στο κελί E11 δίνεται το άθροισμα των 1, 2, 2, 4, 5, 6, 7, 8, 9, 10.



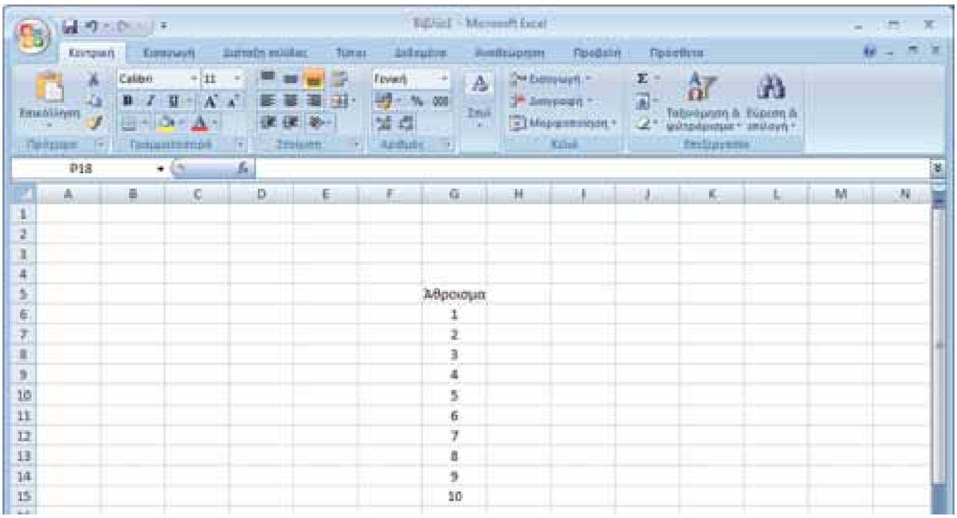
Το άθροισμα υπολογίστηκε μέσω της συνάρτησης Sum, όπως υποδεικνύεται παρακάτω.



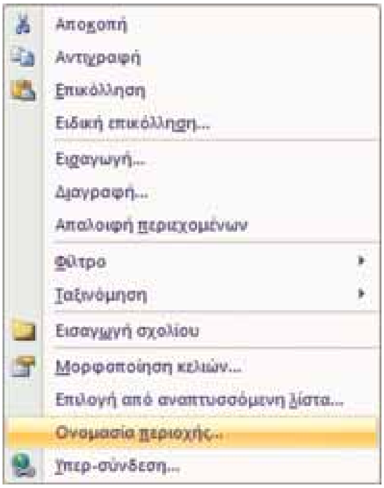
1.6 Ονομασία Δεδομένων "Άθροισμα"

Παραπάνω τα δεδομένα από το 1 έως το 10, αναφέρονται απλά με την αναφορά "D5:D14", είναι όμως πιο πρακτικό στο να δίνουμε ένα όνομα στα δεδομένα μας.

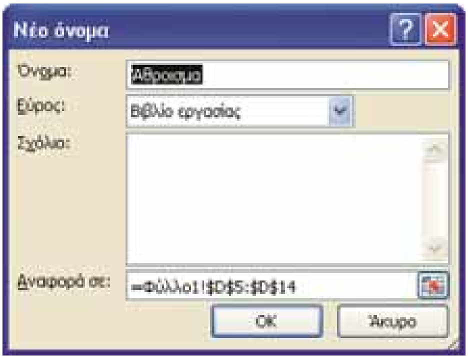
Μπορούμε στα δεδομένα μας από "D5:D14" να δώσουμε ένα όνομα π.χ. "Άθροισμα". Γράφουμε την λέξη "Άθροισμα" στο κελί "D4", μετά μαρκάρουμε τα κελιά από "D5:D14".



Κατόπιν δίνουμε δεξί κλικ με το ποντίκι και επιλέγουμε ονομασία περιοχής



Λαμβάνουμε το επόμενο

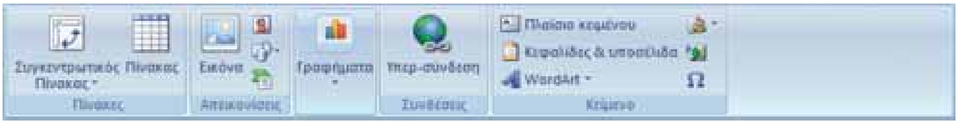


και πατώντας “OK” κατοχυρώνουμε το όνομα “Άθροισμα” για τα δεδομένα μας από “D5:D14”.

1.7 Οδηγός Γραφημάτων με Excel

Μέσω του Excel μπορούμε να λάβουμε πολλών ειδών γραφήματα.

1. Επιλέγουμε τα δεδομένα μας, σε μια στήλη ή σε πολλές. Πολλές φορές είναι χρήσιμο να επιλέξουμε και τα ονόματα των δεδομένων μας. Σε κάθε στήλη αντιστοιχεί ένα όνομα (ή μια ετικέτα).
2. Από το από το εικονίδιο **Εισαγωγή** του μενού, επιλέγω το εικονίδιο **Γραφήματα**

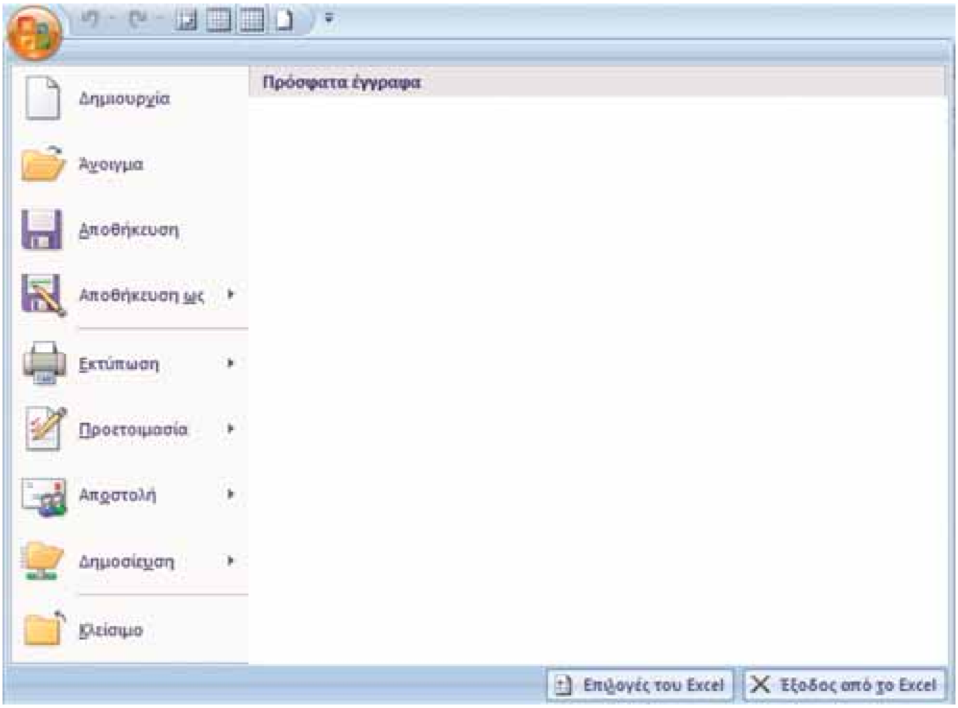


και στην συνέχεια με κλικ στα γραφήματα έχουμε.

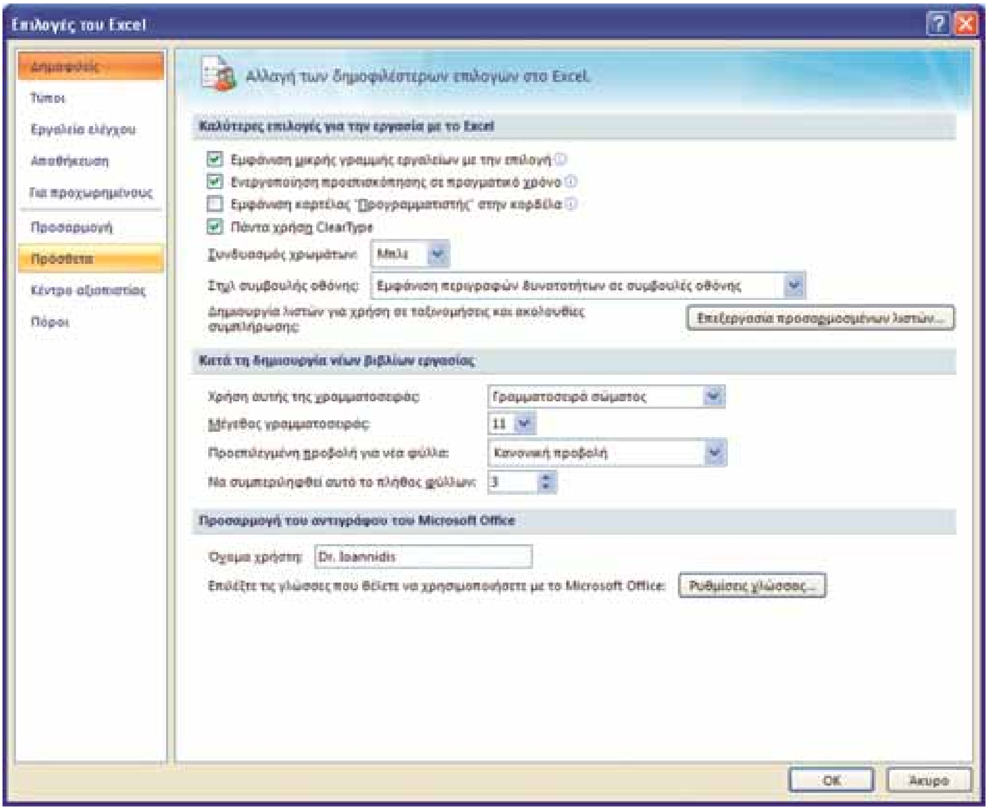
Ιδιαίτερο ενδιαφέρον παρουσιάζει το γράφημα της Διασποράς για δεδομένα από δυο ή περισσότερα μεγέθη (Πολυδιάστατη Ανάλυση).

1.8 Ανάλυση Δεδομένων με Excel

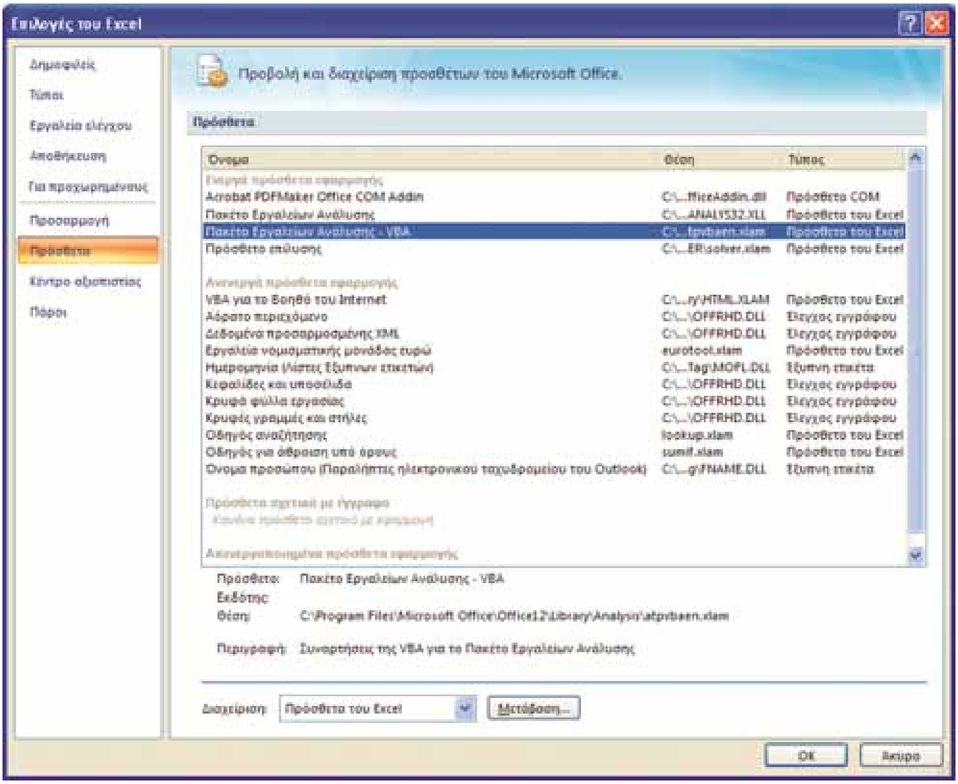
Πέραν των στατιστικών συναρτήσεων, το Excel περιέχει ολοκληρωμένες εντολές Στατιστικής Ανάλυσης. Αν δεν είναι φορτωμένες στο πρόγραμμα του Excel σας μπορείτε να τις εισαγάγετε ως εξής: Πηγαίνετε στο κουμπί του Office και κάντε αριστερό κλικ, εμφανίζεται η επόμενη εικόνα. Στο κάτω αριστερό μέρος επιλέγουμε το: Επιλογές του Excel



Με αριστερό κλικ στο επιλογές έχουμε το:



Με αριστερό κλικ στα πρόσθετα φθάνουμε στο επόμενο και κάνουμε αριστερό κλικ στο Πακέτο Εργαλείο Ανάλυσης



Το “Ανάλυση Δεδομένων” εμφανίζεται με αριστερό κλικ στο εικονίδιο “Δεδομένα” του κύριου μενού.

2 Περιγραφική Στατιστική

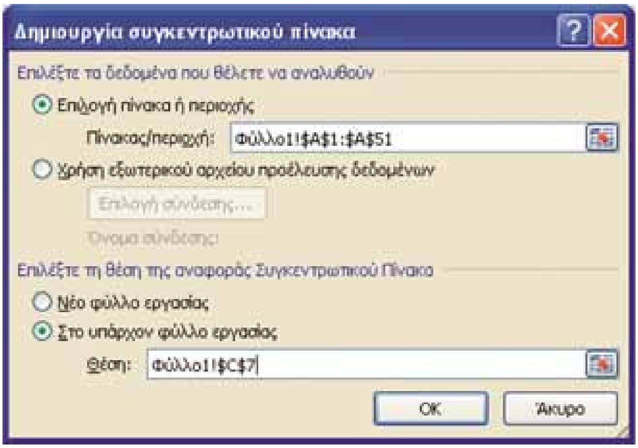
2.1 Ακιδωτό Διάγραμμα

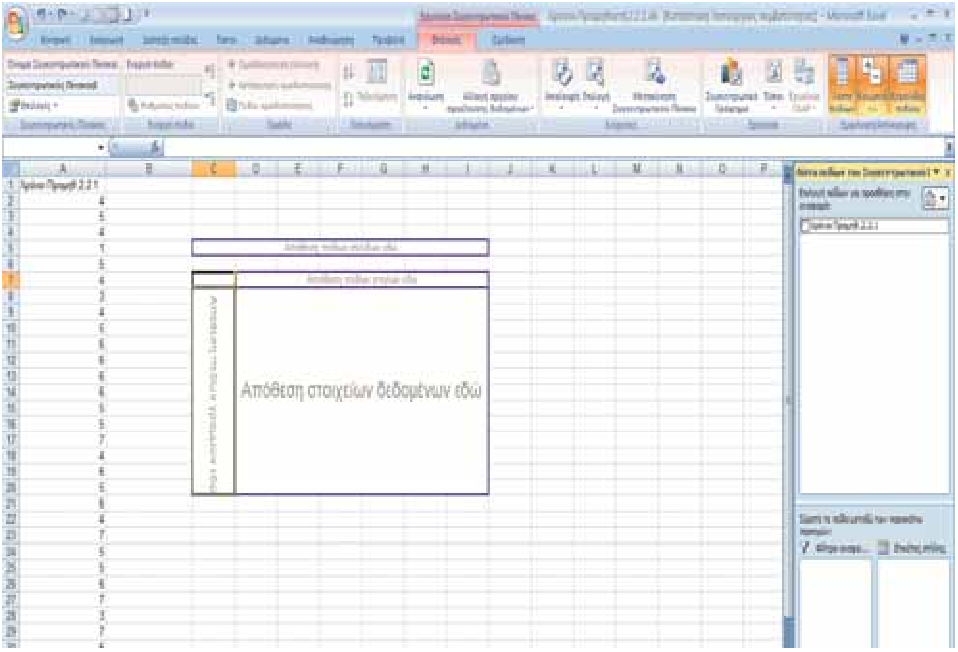
Παράδειγμα 2.3.1 (Στατιστική Μεθοδολογία)

Οδηγίες Excel για την κατασκευή Ακιδωτού Διαγράμματος Παραδείγματος 2.3.1

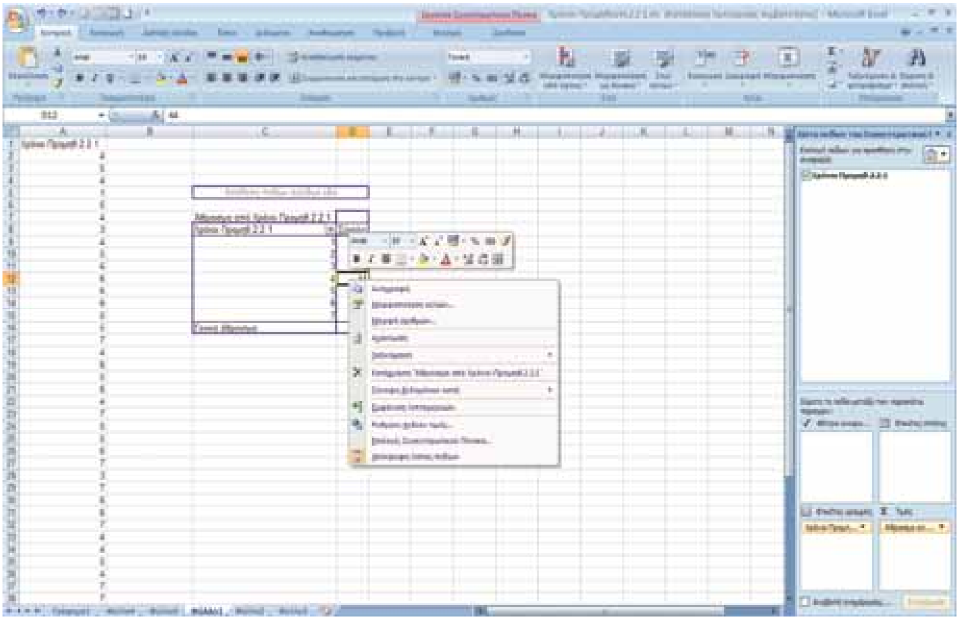
1. Ανοίγουμε το Αρχείο Παράδειγμα 2.2.1
2. Μαρκάρουμε τα δεδομένα μας
3. Επιλέγουμε Συγκεντρωτικού Πίνακα

Εισάγουμε τα δεδομένα μας στο πεδίο: "Πίνακας /περιοχή" και επιλέγουμε για την θέση αναφοράς του συγκεντρωτικού πίνακα το κελί "C7" και "OK".

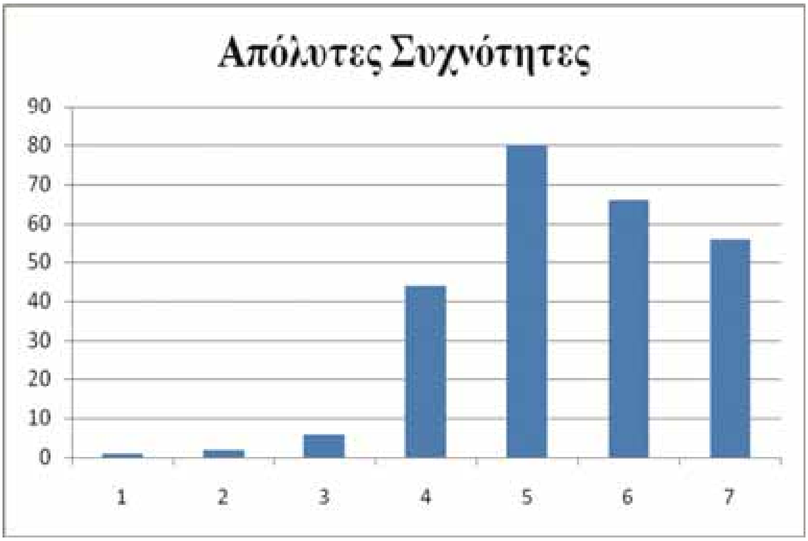




4. Σύρουμε το αρχείο Χρόνοι Προμηθευτή . πρώτα στην Γραμμή και μετά στα δεδομένα. Στη συνέχεια κάνουμε δεξί κλικ πάνω από ένα κελί της στήλης σύνολό και επιλέγουμε στο Ρυθμίσεις Πεδίου Τιμής το “πλήθος”.



5. Από τα Εργαλεία Συγκεντρωτικού Πίνακα επιλέγουμε “Συγκεντρωτικό Γράφημα” και επιλέγουμε από τα στήλη γραφημάτων το στυλ γραφήματος “Στήλη Τμημάτων”.



6. Μπορούμε να αλλάξουμε τον τύπο του γραφήματος κάνοντας δεξί κλικ στο υπάρχον γράφημα, κάνοντας επιλογή Αλλαγή Τύπου Γραφήματος και επιλογή Πίτα.