ΤΕΙ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΑΝΑΛΥΣΗ ΚΑΙ ΠΡΟΒΛΕΨΗ ΧΡΟΝΟΣΕΙΡΩΝ ΟΙΚΟΝΟΜΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

ΜΠΡΙΝΙΑΣ ΒΑΣΙΛΗΣ Α.Μ. 15060

ΕΠΟΠΤΕΥΩΝ ΚΑΘΗΓΗΤΗΣ
ΜΕΓΑΡΙΤΗΣ ΑΘΑΝΑΣΙΟΣ

ΜΕΣΟΛΛΟΓΓΙ 2015

ΠΡΟΛΟΓΟΣ

Η εισαγωγή των νέων τεχνολογιών αλλάζει τα μέχρι τώρα δεδομένα δημιουργώντας νέο τρόπο σκέψης και διαφορετική προσέγγιση στα δρώμενα της κοινωνίας και της λειτουργίας των δομών της. Εφαρμόζοντας τα νέα τεχνολογικά επιτεύγματα στον τομέα των επικοινωνιών και διαχειριζόμενος σωστά τον σημερινό καταιγισμό πληροφορίας ουσιαστικά άλλαξε τη καθημερινότητά του μέσου ανθρώπου.

Η αξιόπιστη, συνεχής επικοινωνία και παροχή πληροφοριών αποτελεί πλέον αναπόσπαστο τμήμα της σύγχρονης ζωής και απαίτηση κάθε επιχείρησης για τις σύγχρονες υπηρεσίες. Σε ένα τέτοιο κοινωνικό γίγνεσθαι ο συνδυασμός της μετάδοσης της πληροφορίας με τις υπηρεσίες οικονομικών αναλύσεων έχουν ήδη ανοίξει τον δρόμο προς την κοινωνία του μέλλοντος, εκεί όπου η διαχείριση και ανάλυση μεγάλου όγκου οικονομικών δεδομένων γίνεται εφικτή με απλό, εύκολο, γρήγορο, οικονομικό και ποιοτικό τρόπο.

Σκοπός της διπλωματικής εργασίας είναι η πρόβλεψη και η ανάλυση χρονοσειρών σε οικονομικά δεδομένα.

ΠΕΡΙΛΗΨΗ

Το θέμα της παρούσας πτυχιακής εργασίας είναι η Ανάλυση Και Πρόβλεψη Χρονοσειρών οικονομικών Δεδομένων.

Στο $1^{ο}$ κεφάλαιο γίνεται μια εισαγωγή στη στατιστική για να μας βοηθήσει να κατανοήσουμε τι είναι η «Στατιστική», στην συνέχεια παρατίθεται μια ιστορική αναδρομή ως αναφορά το πότε ξεκίνησαν οι άνθρωποι να την χρησιμοποιούν. Αμέσως μετά βλέπουμε με ποιον τρόπο σχετίζονται τα μαθηματικά με την στατιστική, και βλέπουμε μερικές βασικές έννοιες της «Στατιστικής», και τέλος έχουμε μια παρουσίαση στατιστικών δεδομένων.

Στο $2^{ο}$ κεφάλαιο επικεντρωνόμαστε στις χρονοσειρές. Στην αρχή έχουμε κάποια γενικά στοιχεία που μας αναφέρουν τι είναι χρονοσειρά, αμέσως μετά βλέπουμε την γραφική παρουσίαση των χρονοσειρών, και στην συνέχεια παρουσιάζονται οι συνιστώσες των χρονοσειρών δηλαδή ότι μια χρονοσειρά χωρίζεται σε τέσσερις διαφορετικές συνιστώσες.

Το θέμα του $3^{ου}$ κεφαλαίου είναι οι μέθοδοι προσδιορισμού της μακροχρόνιας τάσης. Πρώτη μέθοδος είναι η χάραξη της με το ελεύθερο χέρι, δεύτερη είναι η μέθοδος δυο μέσων και τέλος η μέθοδος δυο κινητών μέσων.

Στο $4^{ο}$ κεφάλαιο αναφερόμαστε στις τεχνικές εξομάλυνσης, η οποία είναι ο απλούστερος τρόπος για να πετύχουμε την μείωση της τυχαίας μεταβλητότητας.

Στο τελευταίο κεφάλαιο γίνεται ανάλυση των χρονοσειρών με τη χρήση του Microsoft Excel.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1 ΕΙΣΑΓΩΓΗ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ

1.1 ΕΙΣΑΓΩΓΗ

Ο όρος «Στατιστική» ενδεχομένως να προέρχεται από τη λατινική λέξη “status” (πολιτεία, κράτος) η οποία, χρησιμοποιήθηκε αρχικά για το χαρακτηρισμό αριθμητικών δεδομένων που αναφέρονται κυρίως στον πληθυσμό μιας χώρας. Μπορεί όμως να προέρχεται από την αρχαία ελληνική λέξη στατίζω (τοποθετώ, ταξινομώ, συμπεραίνω). Πολλές απόψεις έχουν διατυπωθεί μέχρι σήμερα για το αντικείμενο της Στατιστικής.

Στη βιβλιογραφία, ως **Στατιστική** ορίζουμε την επιστήμη που ασχολείται με επιστημονικές μεθόδους συλλογής, οργάνωσης, παρουσίασης, ανάλυσης και ερμηνείας αριθμητικών δεδομένων καθώς επίσης και με την εξαγωγή συμπερασμάτων για κάποιο άγνωστο χαρακτηριστικό ενός γενικότερου συνόλου με βάση τις πληροφορίες που περιέχονται σε ένα μέρος (υποσύνολο) από το σύνολο αυτό.

Παρότι ο ορισμός αυτός αντιπροσωπεύει πράγματι ένα μεγάλο μέρος των δραστηριοτήτων της Στατιστικής δεν αποτελεί το αποκλειστικό αντικείμενο της επιστήμης αυτής. Αυτή η πλευρά των δραστηριοτήτων της Στατιστικής αποτελεί αυτό που ονομάζουμε **Περιγραφική Στατιστική**. Η άλλη διάσταση της Στατιστικής είναι εκείνη, η οποία ασχολείται με τη συμπερασματολογία. Για αυτή τη πλευρά της Στατιστικής θα μπορούσαμε να δίναμε τον ορισμό ότι η Στατιστική είναι η προσπάθεια εξαγωγής συμπερασμάτων κάτω από συνθήκες αβεβαιότητας.

Θα μπορούσε να ισχυρισθεί κανείς ότι η Στατιστική συμπερασματολογία είναι περισσότερο σημαντική. Ο λόγος, ίσως, είναι ότι χρειάζεται περισσότερο πολύπλοκα «εργαλεία» προκειμένου να αναπτυχθεί. Τα τελευταία χρόνια όμως και η περιγραφική Στατιστική βρίσκει όλο και περισσότερες εφαρμογές. Αυτό οφείλεται στο γεγονός ότι σε όλες σχεδόν τις επιστήμες υπάρχει ανάγκη ποσοτικής προσέγγισης των διαφόρων εννοιών και μεθόδων. Αυτό γίνεται με συγκέντρωση, ανάλυση και παρουσίαση των υπαρχόντων στοιχείων.

Και οι δύο βασικές δραστηριότητες της Στατιστικής που αναπτύχθηκαν προηγουμένως χρειάζονται ικανότητα στατιστικής σκέψης. Η ικανότητα στατιστικής σκέψης πρέπει βέβαια να συνδυάζεται και με γνώση του αντικειμένου από το οποίο προέρχονται τα προς ανάλυση στοιχεία.

Είναι γνωστό ότι πολλοί ισχυρίζονται πως η Στατιστική είναι ένας τρόπος για να λέει κανείς ψέματα. Στην πραγματικότητα, δεν είναι η Στατιστική υπεύθυνη για αυτό αλλά οι μη κατάλληλες στατιστικές μέθοδοι που χρησιμοποιούνται για την ανάλυση συγκεκριμένων δεδομένων ή η παραπλανητική ερμηνεία δεδομένων γίνεται από μη στατιστικούς για εξυπηρέτηση συγκεκριμένων σκοπιμοτήτων. Είναι δηλαδή αποτέλεσμα λανθασμένης στατιστικής σκέψης. Γενικά θα μπορούσε να πει κανείς ότι η στατιστική σκέψη μας παρέχει τη δυνατότητα βασικής κατανόησης των στατιστικών μεθόδων και μας βοηθά να ανακαλύψουμε επαναλαμβανόμενες διαδικασίες με την αξιοποίηση διαθέσιμων δεδομένων.

Ένα από τα σημαντικότερα βήματα για να μεγιστοποιήσει κανείς τη χρησιμότητα των στατιστικών εννοιών και μεθόδων βρίσκεται στην επιλογή των δεδομένων. Όσο πιο σχετικά είναι τα δεδομένα με το πρόβλημα που μας ενδιαφέρει να αντιμετωπίσουμε τόσο χειρότερη είναι η ανάλυσή τους. Η ενασχόληση με τον τρόπο συλλογής των υπό ανάλυση δεδομένων και η ανάλυση που θα επακολουθήσει επηρεάζεται από τη γνώση του αντικειμένου από εκείνον, ο οποίος ενδιαφέρεται να πάρει κάποια απόφαση και από τις πληροφορίες που είναι αναγκαίες προκειμένου να ληφθεί απόφαση για ένα συγκεκριμένο πρόβλημα.

1.2 ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

Με την εμφάνιση της Στατιστικής και στα πρώτα στάδια της ανάπτυξής της οι άνθρωποι την ταύτισαν με την παράθεση τεράστιων πινάκων με δεδομένα σχετικά με τους θανάτους, τις γεννήσεις, τους φόρους, τα προϊόντα, τους άνδρες σε στρατεύσιμη ηλικία κτλ., προσπαθώντας έτσι να περιγράψουν διάφορα δημογραφικά, οικονομικά και πολιτικά φαινόμενα. Έχει εξακριβωθεί ότι η πρώτη απογραφή πληθυσμού έγινε στη Κίνα από τον αυτοκράτορα Yao το 2238 πχ. . Επίσης, στοιχειώδεις απογραφές φαίνεται να έχουν πραγματοποιηθεί από τους Σίνες, τους Αιγυπτίους και τους Πέρσες. Ο όρος Στατιστική αναφέρεται επίσης και από το Σωκράτη (Ξενοφώντος «Απομνημονεύματα») και από τον Αριστοτέλη στη «Πολιτεία». Στους Ρωμαίους η πρώτη απογραφή πληθυσμού έγινε επί Ρωμύλου (753-717 πχ) και η τελευταία από τον αυτοκράτορα Βεσπασιανό το 73 πχ. Στην αρχαιότητα, η συγκέντρωση στατιστικών στοιχείων είχε στόχο τον εντοπισμό των πολιτών που είχαν υποχρέωση να υπηρετήσουν ως πολεμιστές ή να πληρώσουν φόρο.

Στην Αγγλία, η πρώτη καθολική απογραφή του πληθυσμού και του πλούτου γενικά έγινε το 1085 από το Γουλιέλμο το Κατακτητή. Πιο συγκεκριμένα, η πραγματικά σπουδαία στατιστική απογραφή στην εποχή του Γουλιέλμου του Κατακτητή, στο τέλος του 11ου αιώνα, αναφέρεται σε διάφορες μονάδες παραγωγής της Αγγλίας όπως μεταλλεία, ιχθυοτροφεία κ.ά. Από το 16ο έως το 19ο αιώνα, η ραγδαία ανάπτυξη του εμπορίου ώθησε τις πολιτειακές αρχές στη μελέτη οικονομικών δεδομένων, όπως είναι το εξαγωγικό εμπόριο, το πλήθος και η δυναμικότητα των βιομηχανιών κτλ.

Συστηματική συλλογή δεδομένων για τον πληθυσμό και την οικονομία άρχισε κατά τη διάρκεια της Αναγέννησης στις πόλεις Βενετία και Φλωρεντία στην Ιταλία, και γρήγορα επεκτάθηκε και σε άλλες χώρες της Δυτικής Ευρώπης. Ο μεγάλος ρυθμός θνησιμότητας στην Ευρώπη οφειλόταν στις επιδημικές ασθένειες, στους πολέμους και στις λιμοκτονίες. Στις αρχικές καταγραφές των θανάτων από την πανώλη, τη φοβερή ασθένεια που εμφανίστηκε το 1348 και κράτησε πάνω από 400 χρόνια, προστέθηκαν στη συνέχεια και οι θάνατοι από άλλες αιτίες. Στα 1620 ο Άγγλος εμπορευόμενος Graunt από δειγματοληπτική έρευνα που έκανε σε οικογένειες του Λονδίνου βρήκε ότι σε κάθε 88 άτομα υπήρχαν 3 θάνατοι. Χρησιμοποιώντας τους καταλόγους του Λονδίνου, που έδιναν 13.200 θανάτους το 1620, εκτίμησε τον πληθυσμό του Λονδίνου το έτος αυτό στα 387.200 άτομα.

Το 1853, γράφεται από το Sansovino το πρώτο βιβλίο στατιστικού περιεχομένου και λίγο αργότερα εισάγεται από το Korning η στατιστική στην ανώτερη παιδεία. Την ίδια εποχή εμφανίζεται ενδιαφέρον για τις ασφάλειες ζωής και ο περίφημος Άγγλος αστρονόμος Halley, χρησιμοποιώντας τα ληξιαρχικά βιβλία γεννήσεων και θανάτων της πόλεως Breslaou, παρουσιάζει το πρώτο πίνακα θνησιμότητας. Το ρεύμα αυτό των δημογραφικών μελετών επεκτείνεται και στη Γερμανία, όπου ο Siissmilch συγκεντρώνει στοιχεία από τα ληξιαρχικά βιβλία των εφημερίων της Πρωσίας και καταλήγει το 1741 στο συμπέρασμα ότι το ποσοστό γέννησης των αγοριών είναι 51% και των κοριτσιών 49%, ενώ τα δύο φύλλα έχουν ίσα ποσοστά κατά την εποχή του γάμου. Το φαινόμενο αυτό για το συγγραφέα δεν είναι τυχαίο γεγονός αλλά νόμος θείας προέλευσης που αποσκοπεί στη διαιώνιση του είδους. Μέχρι την εποχή αυτή η Στατιστική έχει περιγραφικό χαρακτήρα και ασχολείται κυρίως με θέματα Δημογραφίας.

Η Στατιστική θα ξεφύγει από τον περιγραφικό χαρακτήρα της με την ανάπτυξη ενός νέου κλάδου των Πιθανοτήτων, ο οποίος προήλθε από τη μελέτη των τυχερών παιχνιδιών (χαρακτηριστική μάλιστα είναι η αλληλογραφία ανάμεσα στους Γάλλους μαθηματικούς Pascal και Fermat, με αφορμή ερωτήματα που έθεσε στον Pascal ο ιππότης De Mere για τα παιχνίδια του κύβου). Από τους θεμελιωτές των Πιθανοτήτων αναφέρουμε τον Bernoulli, ο οποίος στο βιβλίο του «η τέχνη των προβλέψεων» διατυπώνει το περίφημο νόμο των μεγάλων αριθμών και το Γάλλο μαθηματικό Laplace, στον οποίο οφείλεται η εφαρμογή των Πιθανοτήτων στη σπουδή των φυσικών φαινομένων με πολυσύνθετες αιτίες.

Η Θεωρία των Πιθανοτήτων πρωτοεμφανίστηκες στα μέσα του 17^{ου} αιώνα σε σχέση με προβλήματα που αναφέρονταν σε τυχερά παιχνίδια, τα οποία δεν προσαρμόζονταν στο πλαίσιο των Μαθηματικών της εποχής εκείνης. Αυτό έθεσε τις βάσεις σημαντικών εννοιών όπως η Πιθανότητα και η μαθηματική μέση τιμή. Η μεγάλη όμως άνθιση των πιθανοτήτων εμφανίστηκε στα μέσα του 19^{ου} αιώνα και στις αρχές του 20^{ου} αιώνα κυρίως στη Ρωσία. Επιστήμονες όπως οι Chebyshev, Markov και Lyapunov και στη συνέχεια οι Bernstein, Kolmogorov, Khinchin και Chedenko διαμόρφωσαν τη Θεωρία των Πιθανοτήτων με τη μορφή που γνωρίζουμε σήμερα.

Σημαντικό ρόλο στην ανάπτυξη των πιθανοτήτων έπαιξαν τα παιχνίδια που στηρίζονταν στο ζάρι και στα χαρτιά. Τα παιχνίδια με τα χαρτιά έγιναν δημοφιλή στην Ευρώπη το 14^{ο} αιώνα, ενώ παιχνίδια με ζάρια παίζονταν στην αρχαία Αίγυπτο από την εποχής της πρώτης Δυναστείας (η λέξη «ζάρι» προέρχεται από την αραβική λέξη al zar. Επίσης από εκεί πηγάζει και η αγγλική λέξη hazard που χαρακτηρίζει την αβεβαιότητα των πιθανοτήτων). Σύμφωνα με την ελληνική παράδοση το ζάρι ανακαλύφθηκε από τον Παλαμήδη, όταν ο τελευταίος προσπαθούσε να βρει τρόπο για να διασκεδάσει τους κουρασμένους Έλληνες στρατιώτες που συμμετείχαν στην πολιορκία της Τροίας και περίμεναν τη μεγάλη μάχη. Ο Παυσανίας αναφέρει στα γραπτά του μία εικόνα ζωγραφισμένη που δείχνει τον Παλαμήδη και Θερσίτη να παίζουν ζάρια.

Η σύγχρονη ανάπτυξη των Πιθανοτήτων βέβαια έχει μικρή - αν όχι καμία - σχέση με τα ζάρια, δεδομένου ότι τέτοια τυχερά παιχνίδια αναφέρονται σε ισοπίθανα ενδεχόμενα, κάτι που δύσκολα συναντάται στα προβλήματα που αντιμετωπίζει η σύγχρονη θεωρία Πιθανοτήτων. Αυτός είναι ο λόγος που η Συνδυαστική, η βάση της Θεωρίας Πιθανοτήτων που αναφέρεται σε ισοπίθανα ενδεχόμενα, δεν είναι πια τόσο απαραίτητη στην ανάπτυξη της θεωρίας των Πιθανοτήτων. Όπως παρατήρησε και ο Laplace «είναι αξιοσημείωτο ότι μία επιστήμη που ξεκίνησε με προβλήματα θα γίνονταν το πιο σημαντικό αντικείμενο της ανθρώπινης γνώσης…».

Στη νέα περίοδο της Στατιστικής ο Βέλγος αστρονόμος Quetelet επεκτείνει την εφαρμογή της Στατιστικής στη σπουδή των φυσικών, διανοητικών και ηθικών ιδιοτήτων του ανθρώπου που παίρνει την πρωτοβουλία για τη σύγκληση του πρώτου Διεθνούς Συνεδρίου Στατιστικής που έγινε στις Βρυξέλλες το 1853, ενώ αργότερα ο Galton εφαρμόζει τη Στατιστική στη Βιολογία και ειδικότερα στα προβλήματα της κληρονομικότητας. Η προσπάθεια του Galton συνεχίσθηκε από τον Άγγλο μαθηματικό Pearson, στον οποίο οφείλεται κατά πολύ η σημερινή ανάπτυξη και θέση της Στατιστικής.

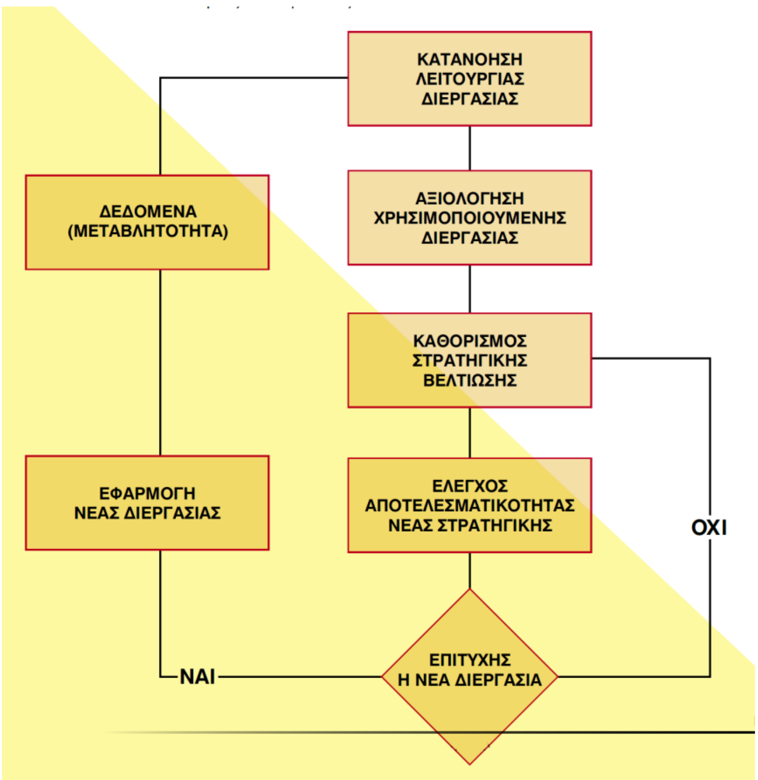
1.3 ΜΑΘΗΜΑΤΙΚΑ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗ

Συνήθως, οι Πιθανότητες θεωρούνται ένας κλάδος των Μαθηματικών, που μελετά τις κανονικότητες τυχαίων φαινομένων. Υπάρχουν πολλοί, οι οποίοι συμμερίζονται την άποψη του κορυφαίου Ρώσου καθηγητή Gnedenko, ότι δηλαδή τα Μαθηματικά θα έπρεπε να θεωρούνται ως μία ειδική περίπτωση της Θεωρίας Πιθανοτήτων.

Η σημασία της επιστήμης των Πιθανοτήτων έχει τονισθεί και από δύο άλλους κορυφαίους επιστήμονες: Ο μεγάλος φιλόσοφος Russel σε μία διάλεξη το 1929 παρατήρησε ότι «οι πιθανότητες είναι οι πιο σημαντική έννοια των μοντέρνων επιστημών διότι κανείς δεν έχει την παραμικρή ιδέα για το τι σημαίνουν». Εξάλλου το 1900 στο διεθνές συνέδριο των μαθηματικών στο Παρίσι ο μαθηματικός Hilbert χαρακτήρισε το πρόβλημα της θεμελίωσης της Θεωρίας των Πιθανοτήτων ως «ένα από τα 23 άλυτα σημαντικά προβλήματα των Μαθηματικών».

Είναι ενδιαφέρον ότι παρά το γεγονός ότι η Θεωρία των Πιθανοτήτων είχε παρουσιάσει στις αρχές του 20^{ου} αιώνα πολλά εξαιρετικά ενδιαφέροντα αποτελέσματα, δεν περιλαμβανόταν ως θεωρία στις περιοχές των Μαθηματικών, κυρίως λόγω της έλλειψης αξιωματικής θεμελίωσης. Από τότε βέβαια έχει περάσει αρκετός χρόνος ώστε να δικαιωθεί η Θεωρία Πιθανοτήτων, όχι μόνο ως κλάδος των Μαθηματικών αλλά και ως ένας ανεξάρτητος επιστημονικός κλάδος.

Καθοριστικό ρόλο στην ανάλυση δεδομένων παίζει η αντίληψη της έννοιας της μεταβλητότητας. Η μεταβλητότητα είναι αναπόφευκτη σε όλες τις πλευρές της ανθρώπινης δραστηριότητας. Για οποιοδήποτε θέμα που αναφέρεται στον άνθρωπο και στο περιβάλλον του είναι προφανές ότι υπάρχει μεταβλητότητα. Το βάρος των ανθρώπων, η θερμοκρασία του περιβάλλοντος, το περιεχόμενο σε μία συσκευασία είναι μερικά από τα πράγματα που εμφανίζουν μεταβλητότητα. Κατανόηση της μεταβλητότητας και των λόγων που την προκαλούν είναι απαραίτητα για την ερμηνεία των δεδομένων. Θα μπορούσε να ισχυρισθεί κανείς ότι η κατανόηση και ερμηνεία της μεταβλητότητας σε ένα σύνολο δεδομένων είναι ακριβώς αυτό με το οποίο ασχολείται η Στατιστική. Η έννοια της μεταβλητότητας είναι ίσως αυτή ακριβώς που οδήγησε στη σημαντική ανάπτυξη και αξιοποίηση των μεθόδων των Πιθανοτήτων και της Στατιστικής τα τελευταία χρόνια σε κατεύθυνση διαφορετική από εκείνη των Μαθηματικών. Τα Μαθηματικά, όπως είναι γνωστό, ασχολούνται με συγκεκριμένες και σαφώς καθορισμένες διαδικασίες όπου ένα σύνολο συγκεκριμένων υποθέσεων μπορεί να οδηγήσει σε ένα μονοσήμαντο αποτέλεσμα. Αντίθετα, η Στατιστική δημιουργήθηκε από την ανάγκη μελέτης φαινομένων, που υπό συνθήκες είναι δυνατόν να καταλήξουν σε διαφορετικά αποτελέσματα λόγω της ύπαρξης της μεταβλητότητας.



Εικόνα 1.1 Χρήση της Στατιστικής σκέψης σε συνδυασμό με τη γνώση του αντικειμένου και την κρίση αυτού που λαμβάνει τις αποφάσεις

Στην καθομιλουμένη, Στατιστική σημαίνει συστηματική απαρίθμηση και παρουσίαση αριθμητικών δεδομένων ή στοιχείων, τα οποία προέρχονται από πολλές παρατηρήσεις ή μετρήσεις. Στην επιστημονική γλώσσα, η λέξη Στατιστική έχει ευρύτερη σημασία και κάποιος μπορεί να ισχυρισθεί ότι είναι η διαδικασία συλλογισμών που αναγνωρίζει ότι υπάρχει μεταβλητότητα σε όλα τα φαινόμενα και ότι η μελέτη της μεταβλητότητας οδηγεί σε νέες γνώσεις και καλύτερες αποφάσεις.

Μία από τις κυριότερες εφαρμογές της Στατιστικής είναι η χρήση μεθόδων για την υποβοήθηση της λήψης αποφάσεων. Στο πλαίσιο αυτό στατιστική σκέψη είναι ο τρόπος σκέψης που μας επιτρέπει να καταλάβουμε και τελικά να βελτιώσουμε κάποιες διεργασίες μέσω ενδελεχούς μελέτης της μεταβλητότητας των δεδομένων. Στην Εικόνα 1, παρουσιάζεται πως η χρήση της Στατιστικής σκέψης σε συνδυασμό με τη γνώση του αντικειμένου και τη κρίση αυτού που λαμβάνει τις αποφάσεις, όπως αυτά συνδέονται με το σχεδιασμό, τη λήψη αποφάσεων και τη βελτίωση συστημάτων.

1.4 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

Μεταβλητές ονομάζονται τα χαρακτηριστικά εκείνα, ως προς τα οποία εξετάζουμε έναν πληθυσμό. Οι δυνατές τιμές που μπορεί να πάρει μια μεταβλητή λέγονται τιμές της μεταβλητής. Οι μεταβλητές χωρίζονται σε δύο κατηγορίες, στις ποιοτικές και στις ποσοτικές.

- **Ποιοτικές** ή **Ονομαστικές** (nominal) μεταβλητές είναι εκείνες που δεν επιδέχονται μέτρηση και οι τιμές τους δεν είναι αριθμοί.
- **Ποσοτικές** ή **Αριθμητικές** μεταβλητές είναι εκείνες που επιδέχονται μέτρηση και οι τιμές τους είναι αριθμοί.

Οι ποσοτικές μεταβλητές διακρίνονται σε συνεχείς και διακριτές (ασυνεχείς).

- **Συνεχείς** (scale) είναι οι ποσοτικές μεταβλητές που μπορούν να πάρουν οποιαδήποτε τιμή ενός διαστήματος (α, β).
- **Διακριτές** (ordinal) είναι οι ποσοτικές μεταβλητές που παίρνουν μόνο μεμονωμένες τιμές.

Οι συνεχείς μεταβλητές μπορούν να διαχωριστούν περαιτέρω σε διαστημικές (interval) και αναλογικές (ratio), χωρίς ωστόσο να είναι σημαντικός διαχωρισμός υπό την έννοια πως δεν επηρεάζεται η υλοποίηση μίας στατιστικής διαδικασίας από το χαρακτηρισμό μίας συνεχούς μεταβλητής ως διαστημικής ή αναλογικής. Μία μεταβλητή χαρακτηρίζεται αναλογική αν ο λόγος των τιμών μεταξύ δύο παρατηρήσεων είναι δυνατό να ερμηνευθεί σε όρους της έννοιας που καταγράφει η μεταβλητή. Παράδειγμα αναλογικής μεταβλητής είναι το βάρος.

Στατιστικός πληθυσμός ή απλά **πληθυσμός** ονομάζεται κάθε σύνολο, τα στοιχεία του οποίου εξετάζουμε ως προς ένα ή περισσότερα χαρακτηριστικά τους. Τα στοιχεία του πληθυσμού ονομάζονται μονάδες ή άτομα. Ο πληθυσμός μπορεί να είναι θεωρητικός ή πραγματικός (μετρήσιμος).

Παραδείγματα πραγματικού πληθυσμού είναι:

- Οι πολίτες της πόλης του Μεσολογγίου (γνωστός ο αριθμός τους κάθε στιγμή)
- Οι φοιτητές ενός τμήματος Α.Ε.Ι. (ομοίως)
- Οι δημόσιοι υπάλληλοι που απασχολούνται στο Υ.Π.Ε.Π.Θ. (ομοίως)

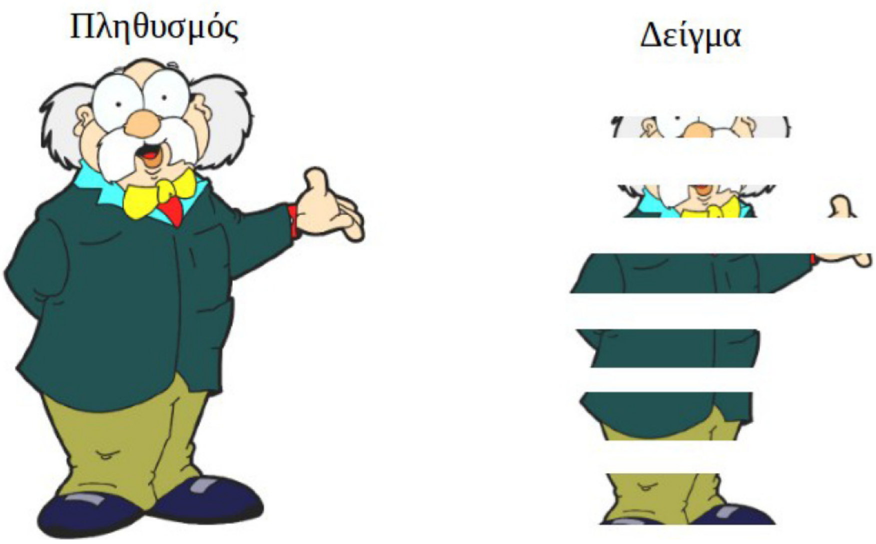
Παραδείγματα θεωρητικού πληθυσμού είναι:

- Οι καπνιστές που θα ασθενήσουν από καρκίνο τον επόμενο μήνα (δεν είναι γνωστός ο αριθμός τους)
- Όσοι θα υποβάλλουν αίτηση στο ΑΕΠ για μεταπτυχιακό το 2015.

Στην περίπτωση του πραγματικού πληθυσμού πρέπει να είμαστε σε θέση να γνωρίζουμε το μέγεθός του (ενημέρωση από τα κατάλληλα αρχεία) ενώ στην περίπτωση του θεωρητικού πληθυσμού πρέπει να είμαστε σε θέση να εκτιμήσουμε το μέγεθος του βάσει των παλαιότερων μετρήσεων, αλλά και το ρυθμό μεταβολής του. **Δείγμα** ονομάζεται το υποσύνολο του πληθυσμού το οποίο μπορούμε να καταγράψουμε υπό τους περιορισμούς (υλικούς και χρονικούς) της έρευνάς μας.

1.4.1 Συλλογή Στατιστικών Δεδομένων

Οι κυριότερες μέθοδοι συλλογής στατιστικών δεδομένων είναι η απογραφή και η δειγματοληψία. **Απογραφή** είναι μια μέθοδος συλλογής στατιστικών δεδομένων, που ακολουθούμε για να πάρουμε όλες τις απαραίτητες πληροφορίες, για έναν πληθυσμό εξετάζοντας όλα τα άτομα του πληθυσμού ως προς τα χαρακτηριστικά που μας ενδιαφέρουν.



Εικόνα 1.2 **Δειγματοληψία**

Δειγματοληψία ονομάζεται η διαδικασία καταγραφής ενός υποσυνόλου του πληθυσμού. Προχωρούμε σε δειγματοληψία γιατί η απογραφή είναι δύσκολη, οικονομικά και χρονικά ασύμφορη και πολλές φορές αδύνατη. Για αυτόν το λόγο επιλέγουμε μια μικρή ομάδα, δηλαδή ένα υποσύνολο του πληθυσμού το οποίο ονομάζεται δείγμα. Συλλέγουμε τις παρατηρήσεις από το δείγμα και στη συνέχεια γενικεύουμε τα συμπεράσματα για ολόκληρο τον πληθυσμό. Τα συμπεράσματα όμως, που θα προκύψουν από τη μελέτη του δείγματος θα είναι αξιόπιστα, δηλαδή θα ισχύουν με ικανοποιητική προσέγγιση για ολόκληρο τον πληθυσμό, μόνο όταν η επιλογή του δείγματος έχει γίνει με τέτοιο τρόπο, ώστε το δείγμα να είναι αντιπροσωπευτικό. Ένα δείγμα θεωρείται αντιπροσωπευτικό, όταν κάθε άτομο του πληθυσμού έχει δυνατότητα να επιλεγεί και αυτό μπορεί να συμβεί με την ίδια πιθανότητα για όλους.

Οι λόγοι για τον οποίο συμβαίνει μία δειγματοληψία είναι οι οικονομικοί και χρονικοί περιορισμοί που υπάρχουν αλλά και η περιορισμένη πρόσβαση στον πληθυσμό. Οι περιορισμοί αυτοί δεν μειώνουν την αξία της δειγματοληψίας καθώς μπορεί να δώσει ακριβή και αξιόπιστα αποτελέσματα ιδιαίτερα όταν ο πληθυσμός που μελετούμε είναι ομοιογενής ως προς το χαρακτηριστικό που μας ενδιαφέρει. Επίσης, η δειγματοληψία μπορεί να είναι περισσότερο αξιόπιστη από μία απογραφή, όταν η γνώση του ερωτώμενου πως πρόκειται για απογραφή αυξάνει τη μεροληψία της απόκρισης π.χ. οι αποκρίσεις των αλλοδαπών στην εθνική απογραφή, οι οποίες ίσως να είναι πιο ειλικρινείς σε δειγματοληψία ή ακόμα στην περίπτωση όπου δεν υπάρχουν αξιόπιστοι κατάλογοι του πληθυσμού όπως στις μη αναπτυγμένες χώρες. Τέλος, η δειγματοληψία μειώνει το κόστος της έρευνας σε πραγματικούς πληθυσμούς, δηλαδή σε πληθυσμούς των οποίων το μέγεθος είναι γνωστό κάθε μία χρονική στιγμή.

Η επιλογή του αντιπροσωπευτικού δείγματος είναι "εκ των ων ουκ άνευ". Αποτελεί πολύ σοβαρή και δύσκολη διαδικασία. Ο κακός σχεδιασμός και η εκτέλεση της στατιστικής έρευνας, η μη αντιπροσωπευτικότητα του δείγματος, ο μη σωστός καθορισμός του μεγέθους του δείγματος αποτελούν μερικά βασικά μειονεκτήματα στη διαδικασία επιλογής ενός δείγματος.

Το **σφάλμα** μίας δειγματοληψίας διαχωρίζεται σε τυχαίο και συστηματικό. **Τυχαίο σφάλμα** δειγματοληψίας ονομάζεται η διαφορά μεταξύ των μετρήσεων του δείγματος και των πραγματικών μετρήσεων το οποίο θα υπάρχει στην έρευνά μας και δεν μπορούμε να το υπολογίζουμε επακριβώς εκτός αν καταφέρουμε να κάνουμε μία τέλεια εκτελεσμένη απογραφή! Το τυχαίο σφάλμα προκύπτει με φυσικό τρόπο καθώς η μέση τιμή (ή άλλα στατιστικά) του υποσυνόλου του πληθυσμού που επιλέγουμε ως δείγμα είναι πρακτικά αδύνατο να είναι ίση με τη μέση τιμή του πληθυσμού, λόγω των τυχαίων σφαλμάτων της δειγματοληψίας. Αν η δειγματοληψία γίνει με κάποια πιθανοθεωρητική μέθοδο τότε το σφάλμα μπορεί να εκτιμηθεί ενώ αν γίνει με κάποια μη πιθανοθεωρητική μέθοδο (όπως συμβαίνει συχνά στην πράξη) τότε ο υπολογισμός του δεν είναι δυνατός. **Συστηματικό σφάλμα δειγματοληψίας** ονομάζεται το σφάλμα που εμφανίζεται λόγω των σφαλμάτων σχεδίασης της δειγματοληψίας, όπως π.χ. αν μετράς την ευχαρίστηση από την απόκτηση ενός προϊόντος και έχεις δύο ομάδες που ρωτάνε, με τη μία να έχει μία πολύ όμορφη γυναίκα ως συνεντευξιαστή και την άλλη έναν άνδρα. Το συστηματικό σφάλμα είναι διαφορετικό από το τυχαίο σφάλμα, καθώς οφείλεται αποκλειστικά στον σχεδιασμό της έρευνας.

Η επιλογή του μεγέθους του δείγματος δεν είναι καθόλου εύκολη εργασία. Ένα δείγμα μεγέθους 30 μπορεί να είναι αρκετό αν ο πληθυσμός είναι μεγέθους 100 και ομοιογενής ενώ ένα δείγμα 100.000 κατοίκων του νομού Θεσσαλονίκης ή Αττικής μπορεί να μην είναι.

1.5 ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Το πλήθος των δεδομένων ενός δείγματος που εξετάζονται για την μελέτη ενός φαινομένου ονομάζεται **μέγεθος του δείγματος** και συμβολίζεται ν. Έστω λοιπόν ένα δείγμα μεγέθους ν και η μεταβλητή X που παίρνει τιμές $x_1, x_2, \dots, x_\kappa$, όπου $\kappa < \nu$. Ο φυσικός αριθμός ν_i, $i = 1, \dots \kappa$ που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της εξεταζόμενης μεταβλητής στο σύνολο των παρατηρήσεων ονομάζεται **συχνότητα**. Αν $\nu_1, \nu_2, \dots, \nu_\kappa$ οι συχνότητες που αντιστοιχούν στις τιμές $x_1, x_2, \dots, x_\kappa$ της μεταβλητής X, τότε:

$$\nu_1 + \nu_2 + \cdots + \nu_\kappa = \nu.$$

Αν διαιρέσουμε τη συχνότητα ν_i, $i = 1, 2, \dots, \kappa$, με το μέγεθος του δείγματος προκύπτει η **σχετική συχνότητα** f_i της τιμής x_i. Για τη σχετική συχνότητα ισχύουν οι σχέσεις:

- $0 \leq f_i \leq 1$
- $f_1 + f_2 + \cdots + f_k = 1$

Όταν η μεταβλητή είναι ποσοτική, ορίζουμε την **αθροιστική συχνότητα** N_i, η οποία εκφράζει το πλήθος των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i. Αν $x_1 < x_2 < \cdots < x_\kappa$ ισχύει:

$$N_i = \nu_1 + \nu_2 + \cdots + \nu_i.$$

Ομοίως, ορίζεται η **αθροιστική σχετική συχνότητα** F_i, η οποία εκφράζει το ποσοστό των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i:

$$F_i = f_1 + f_2 + \cdots + f_i.$$

Τα δεδομένα παρουσιάζονται σύντομα και με σαφήνεια σε **πίνακες συχνοτήτων**, δηλαδή πίνακες όπου αναγραφονται σε στήλες οι τιμές της μεταβλητής, οι συχνότητες εμφάνισης κάθε τιμής της μεταβλητής, κτλ.

Παράδειγμα: Έστω ρωτήθηκαν 30 μαθητές μιας τάξης ενός λυκείου σχετικά με την προτίμηση τους ως προς τις ποδοσφαιρικές ομάδες της Θεσσαλονίκης. Τα δεδομένα παρουσιάζονται στον παρακάτω πίνακα συχνοτήτων.

Ομάδα x_i	**Συχνότητα** ν_i	**Σχετική συχνότητα** f_i	**Ποσοστά** $f_i\%$
ΠΑΟΚ	6	0,2	20
ΑΡΗΣ	12	0,4	40
ΗΡΑΚΛΗΣ	9	0,3	30
ΑΠΟΛΛΩΝ ΚΑΛΑΜΑΡΙΑΣ	3	0,1	10
Σύνολο	$\boldsymbol{\nu = 30}$	**1**	**100**

Πίνακας 1.1 Παράδειγμα πίνακα συχνοτήτων και σχετικών συχνοτήτων

Όταν τα δεδομένα είναι αριθμητικά και το πλήθος των τιμών της εξεταζόμενης μεταβλητής είναι μεγάλο ή η μεταβλητή είναι συνεχής, τότε ταξινομούμε τα δεδομένα σε ένα πλήθος ομάδων που τις ονομάζουμε κλάσεις διαστημάτων και η διαδικασία ταξινόμησης τών δεδομένων σε κλάσεις λέγεται **ομαδοποίηση**. Οι κλάσεις είναι διαστήματα της μορφής $[\alpha, \beta)$ (διάστημα κλειστό αριστερά, ανοιχτό δεξιά). Τα άκρα του διαστήματος τα λέμε **άκρα** της κλάσης. **Κέντρο** μίας κλάσης $[\alpha, \beta)$ ονομάζεται ο αριθμός

$$(\alpha + \beta)/2.$$

Πλάτος μιας κλάσης $[\alpha, \beta)$ ονομάζεται ο αριθμός

$$c = \beta - \alpha.$$

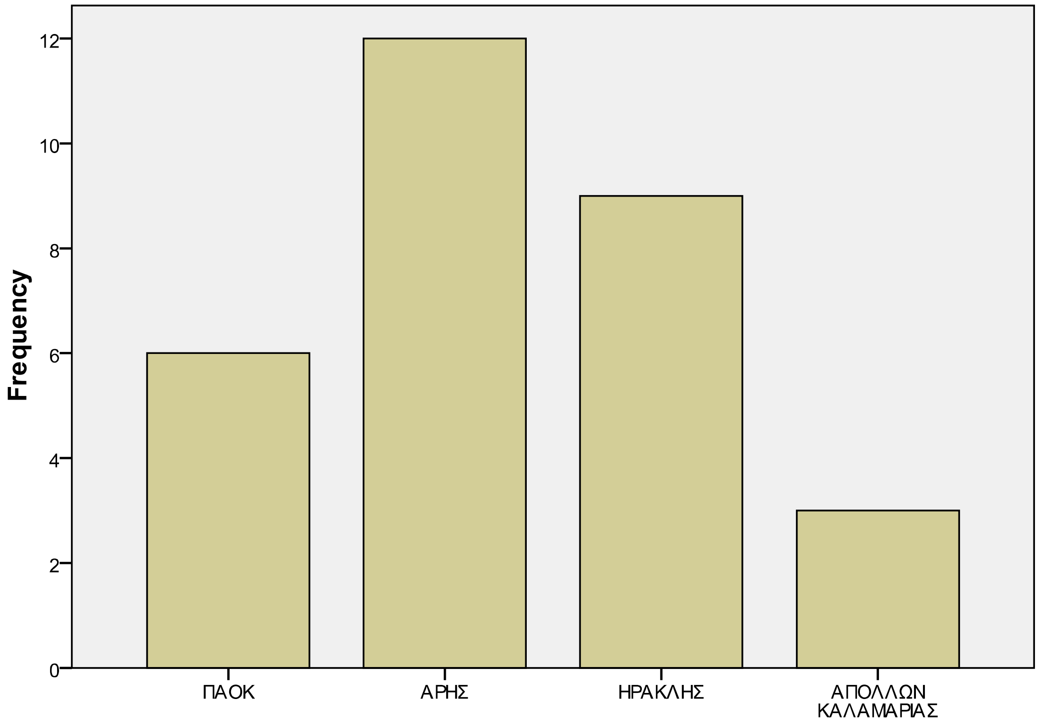
Εύρος του δείγματος είναι η διαφορά της μικρότερης από τη μεγαλύτερη παρατήρηση του δείγματος.

1.5.1 Γραφικοί μέθοδοι για την παρουσίαση των δεδομένων

1.5.1.1 Ραβδόγραμμα

Τα δεδομένα ενός πίνακα συχνοτήτων μπορούν να παρασταθούν γραφικά με ένα ραβδόγραμμα. Το ραβδόγραμμα αποτελείται από ορθογώνια παραλληλόγραμμα (ράβδους) που οι βάσεις τους βρίσκονται πάνω στον οριζόντιο ή τον κατακόρυφο άξονα. Κάθε τιμή της μεταβλητής X αντιστοιχεί σε ένα ορθογώνιο παραλληλόγραμμο το ύψος του οποίου είναι ίσο με την αντίστοιχη συχνότητα ή σχετική συχνότητα. Έτσι έχουμε αντίστοιχα το ραβδόγραμμα συχνοτήτων και το ραβδόγραμμα σχετικών συχνοτήτων. Τόσο η απόσταση μεταξύ των ορθογωνίων παραλληλογράμμων όσο και το μήκος των βάσεών τους καθορίζονται αυθαίρετα.

Παράδειγμα: Τα δεδομένα του Πίνακα 1.1 μπορούν να παρασταθούν γραφικά με το παρακάτω ραβδόγραμμα συχντοτήτων.



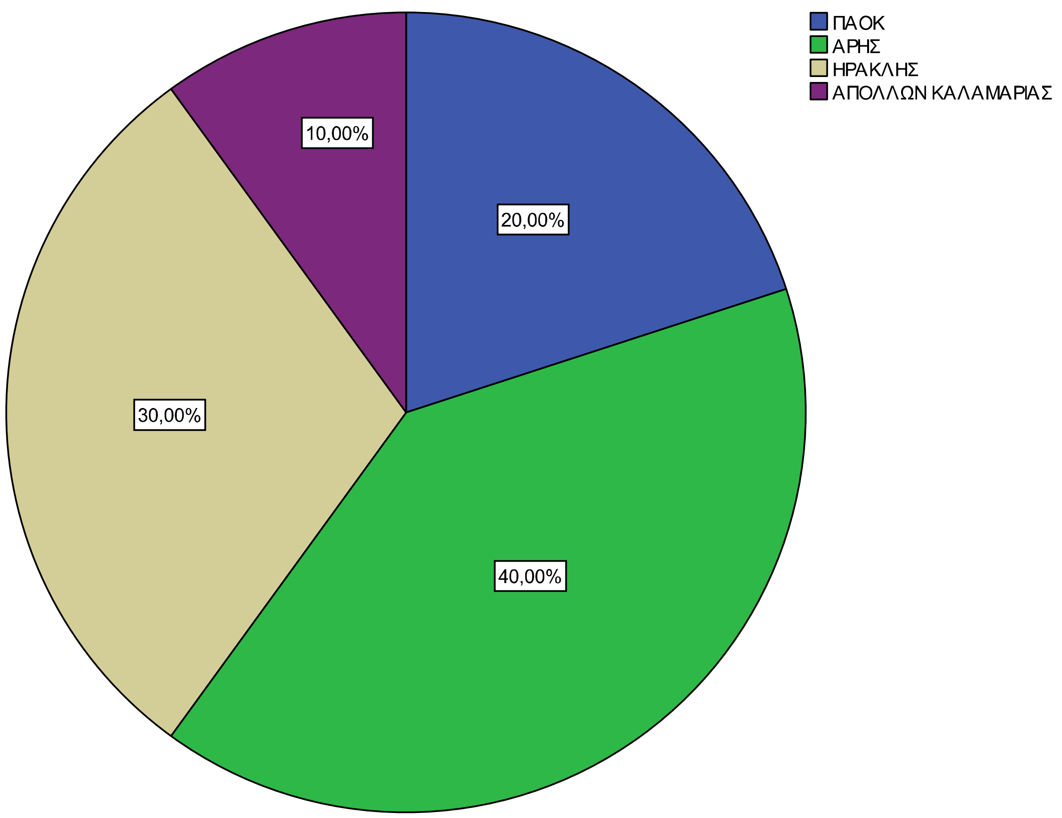
Εικόνα 1.3 **Παράδειγμα ραβδογράμματος**

1.5.1.2 Κυκλικό Διάγραμμα

Χρησιμοποιείται για τη γραφική παράσταση τόσο των ποιοτικών όσο και των ποσοτικών μεταβλητών, όταν οι διάφορες τιμές της μεταβλητής είναι λίγες. Το κυκλικό διάγραμμα περιγράφει το ποσοστό του συνολικού αριθμού παρατηρήσεων που περιέχει κάθε κατηγορία, διαιρώντας ένα κύκλο σε κυκλικούς τομείς έτσι ώστε το εμβαδόν κάθε κυκλικού τομέα να είναι ίσο με τη συχνότητα της αντίστοιχης κατηγορίας. Αν μ_i^o είναι το αντίστοιχο τόξο ενός κυκλικού τομέα στο κυκλικό διάγραμμα συχνοτήτων, τότε:

$$\mu_i^o = 360^o f_i$$

Παράδειγμα: Τα δεδομένα του Πίνακα 1.1 μπορούν να παρασταθούν γραφικά με το παρακάτω κυκλικό διάγραμμα.



Εικόνα 1.4 Παράδειγμα κυκλικού διάγραμματος

1.5.1.3 Φυλλόγραμμα

Το διάγραμμα αυτό δίνει την δυνατότητα ανασύστασης και ανάκλησης των μετρήσεων των αρχικών δεδομένων του δείγματος με ακρίβεια πράγμα το οποίο δεν επιτυγχάνεται με το ιστόγραμμα ή τους πίνακες συχνοτήτων. Χρησιμοποιείται για την επεξεργασία μέτριου αριθμού παρατηρήσεων (περίπου 150). Η παρουσίαση του σχήματος μοιάζει με εκείνου του ιστογράμματος αλλά η τεχνική κατάρτισης δεν είναι η ίδια. Το **φυλλογράφημα** εμφανίζει τα δεδομένα σε όλο το εύρος των παρατηρημένων μετρήσεων, παρουσιάζει την συγκέντρωση των παρατηρήσεων (συχνότητες), δείχνει την μορφή της κατανομής, εμφανίζει τυχόν ακραίες και εκτροπές παρατηρήσεις και επιτρέπει την επισήμανση της απουσίας συγκεκριμένων τιμών ή μετρήσεων.

Παράδειγμα: Δίνονται οι ηλικίες 10 ατόμων:

27 34 34 43 21 38 46 38 22 35.

Να γίνει φυλλόγραμμα των δεδομένων.

1. Διατάσουμε τα δεδομένα σε αύξουσα σειρά

   21 22 27 34 34 35 38 38 43 46.

2. Θεωρούμε ότι κάθε παρατήρηση αποτελείται από δύο τμήματα, το αρχικό ψηφίο και το επόμενο ψηφίο π.χ. 21 : 2 αρχικό, 1 επόμενο.

3. Κατασκευάζουμε τον παρακάτω πίνακα για όλα τα νούμερα

Κορμός	Φύλλο
2	1 2 7
3	4 4 5 8 8
4	3 6

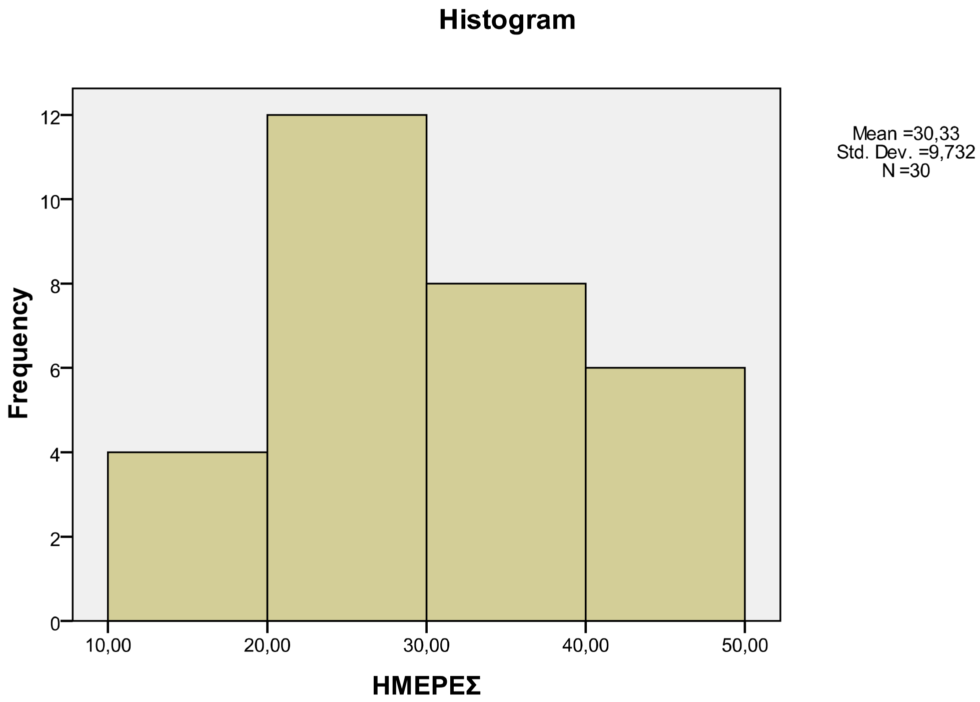
1.5.1.4 Ιστόγραμμα

Η γραφική παράσταση ενός δείγματος με ομαδοποιημένα δεδομένα γίνεται με το **ιστόγραμμα**. Στον οριζόντιο άξονα του συστήματος ορθογωνίων αξόνων σημειώνουμε τα όρια των κλάσεων. Στη συνέχεια κατασκευάζουμε διαδοχικά ορθογώνια, το καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος ισο με τη συχνότητα (ή σχετική συχνότητα) της κλάσης αυτής.

Παράδειγμα: Ο παρακάτω πίνακας δείχνει το πλήθος των λυκείων και τις αντίστοιχες ημέρες κατάληψης κατά τις μαθητικές κινητοποιήσεις 2009-2010.

Ημέρες	**Κέντρο κλάσης** x_i	**Συχνότητα** ν_i
[10,20)	15	4
[20,30)	25	12
[30,40)	35	8
[40,50)	45	6
	Σύνολο	$\boldsymbol{\nu = 30}$

Στον οριζόντιο άξονα ενός συστήματος ορθογωνίων αξόνων σημειώνουμε τα όρια των κλάσεων. Στη συνέχεια, κατασκευάζουμε διαδοχικά ορθογώνια (ιστούς), καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος ίσο με την αντίστοιχη συχνότητα της κλάσης αυτής.



Εικόνα 1.5 Παράδειγμα ιστογράμματος

1.5.2 Περιγραφικά Μέτρα Στατιστικών Δεδομένων

Ο πίνακας συχνοτήτων και οι γραφικές παραστάσεις των δεδομένων μας δίνουν μια συνοπτική εικόνα των δεδομένων. Τα περιγραφικά στατισικά μέτρα είναι αριθμητικά μέτρα που χαρακτηρίζουν ποσοτικά την κατανομή της τυχαίας μεταβλητής, και υπολογίζονται από τις παρατηρήσεις του δείγματος.

1.5.2.1 *Μέτρα Θέσης*

Μέτρα θέσης ονομάζουμε κάποιες χαρακτηριστικές τιμές (αριθμητικά μεγέθη) που δείχνουν τη θέση του «κέντρου» των παρατηρήσεων. Τα κυριότερα μέτρα θέσης είναι η μέση τιμή, η διάμεσος και η επικρατούσα τιμή.

- **Μέση τιμή**: Έστω δείγμα μεγέθους ν και η μεταβλητή X που παίρνει τιμές $x_1, x_2, \dots, x_\kappa$, όπου $\kappa < \nu$. Η δειγματική μέση τιμή δίνεται από τον τύπο
$$\bar{x} = \frac{1}{\nu}\sum_{i=1}^{\kappa} \nu_i \cdot x_i.$$
- **Διάμεσος**: Διάμεσος ενός δείγματος ν παρατηρήσεων οι οποίες έχουν διαταχθεί σε αύξουσα σειρά ορίζεται ως μεσαία παρατήρηση, όταν το ν είναι περιττός αριθμός ή ως μέσος όρος (ημιάθροισμα) των δύο μεσαίων παρατηρήσεων, όταν το ν είναι άρτιος αριθμός.
-
- **Επικρατούσα τιμή:** Επικρατούσα τιμή ονομάζεται η παρατήρηση με την μεγαλύτερη συχνότητα.

1.5.2.2 *Μέτρα Διασποράς*

Μέτρα διασποράς ονομάζουμε τα μέτρα (αριθμητικά μεγέθη) που δείχνουν πως κατανέμονται οι τιμές του δείγματος γύρω από τις «κεντρικές τιμές». Διασπορά ονομάζουμε τη συγκέντρωση ή την απομάκρυνση των στατιστικών δεδομένων γύρω από μια κεντρική τιμή.

- **Εύρος**: Το εύρος είναι μεγαλύτερη παρατήρηση – μικρότερη παρατήρηση.

- **Διασπορά ή διακύμανση**: Έστω δείγμα μεγέθους ν και η μεταβλητή X που παίρνει τιμές $x_1, x_2, \dots, x_\kappa$, όπου $\kappa < \nu$. Η δειγματική διασπορά δίνεται από τον τύπο
$$s^2 = \frac{1}{\nu - 1}\sum_{i=1}^{\kappa} \nu_i \cdot (x_i - \bar{x})^2,$$

όπου $\bar{x}$ είναι η μέση τιμή.

Ο τύπος της διακύμανσης, μπορεί με κατάλληλες πράξεις να πάρει και την ισοδύναμη μορφή:

$$s^2 = \frac{1}{\nu - 1}\left[\sum_{i=1}^{\kappa} \nu_i x_i^2 - \frac{1}{\nu}\left(\sum_{i=1}^{\kappa} \nu_i x_i\right)^2\right] = \frac{1}{\nu - 1}\left[\sum_{i=1}^{\kappa} \nu_i x_i^2 - \nu \bar{x}^2\right].$$

- **Τυπική απόκλιση**: τυπική απόκλιση είναι η θετική τετραγωνική ρίζα της διακύμανσης.

2 ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΧΡΟΝΟΣΕΙΡΕΣ

2.1 ΓΕΝΙΚΑ

Κάθε μεταβλητή της οποίας η τιμή καταγράφεται σε διαδοχικά χρονικά σημεία δημιουργεί μια χρονοσειρά (time series). Στο κεφαλαίο αυτό θα εξετάσουμε την ανάλυση των χρονοσειρών, με στόχο να αναγνωριστούν τάσεις ή πρότυπα που επαναλαμβάνονται ώστε να γίνουν προβλέψεις.

Στη σημερινή εποχή συναντάμε την εφαρμογή των μεθόδων αυτών τόσο στην οικονομία όσο και στη διοίκηση, όπως:

1 Με σκοπό να προγραμματίσουν τις πολιτικές τους οι κυβερνήσεις έχουν ανάγκη τις προβλέψεις μεγεθών όπως τα επιτόκια, η ανεργία και ο πληθωρισμός.
2 Οι επιχειρήσεις που ασχολούνται με τις οικοδομικές εργασίες έχουν ως σκοπό να γνωρίζουν τα επιτόκια στεγαστικών δανείων, τη ζήτηση νέων κατοίκων και τις τιμές των οικοδομικών υλικών.
3 Κάθε επιχείρηση έχει σκοπό να γνωρίζει τις προβλέψεις της ζήτησης των προϊόντων της και το ποσοστό της αγοράς που της αντιστοιχεί.
4 Τόσο τα Πανεπιστήμια όσο και τα κολέγια θέλουν να γνωρίζουν τους αριθμούς των υποψηφίων σπουδαστών κατά ειδικότητα.

Στο χώρο λήψης αποφάσεων η πρόβλεψη (forecasting) είναι μια πρακτική που χρησιμοποιείται ευρέως. Βασιζόμενοι σε γεγονότα παλαιότερων περιόδων σε αυτό το κεφαλαίο θα εστιάσουμε στην πρόβλεψη της πορείας χρονοσειρών, όπως πωλήσεις ή ποσοστά ανεργίας. Μια πρόβλεψη τέτοιου υποδείγματος απαρτίζει ένα χρήσιμο αντικείμενο σε κάθε διεξαγωγή λήψης αποφάσεων.

Για το πόσο σημαντικές είναι αυτές οι προβλέψεις μπορούμε εύκολα να το διαπιστώσουμε από το γεγονός ότι καμία επιχείρηση, ιδιωτική ή δημόσια δεν μπορούν να υπερπηδήσουν τη διακήρυξη προγραμμάτων για το μέλλον.

2.2 ΓΡΑΦΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΧΡΟΝΟΣΕΙΡΩΝ

Σύμφωνα με τα όσα προείπαμε η χρονοσειρά απαρτίζει συλλογή δεδομένων των οποίων οι τιμές (μετρήσεις ή απαριθμήσεις) αναφέρονται σε τακτές χρονικές περιόδους (μήνες, εξάμηνα, έτη) καλύπτουν μια μακροχρόνια περίοδο. Ο συμβολισμός αυτών των χρονικών περιόδων που αναφέρονται οι τιμές των στοιχείων συμβολίζονται με t (t= 1,2,3,…,n) και οι αντίστοιχες τιμές της μεταβλητής που εξετάζουμε με Yt. Τρείς είναι οι βασικοί τρόποι με τους οποίους πετυχαίνουμε τη γραφική παρουσίαση τέτοιων στοιχείων και πιο ειδικά με Πολυγωνικές Γραμμές, Κάθετα Ραβδογράμματα και Διαγράμματα Επιφανειών. Με τον όρο χρονόγραμμα (time series chart) εννοούμε την παρακολούθηση των τιμών μιας μεταβλητής με γραφικές μεθόδους. Αξίζει να σημειώσουμε ότι στην πλειονότητα των κοινωνικό–οικονομικών μεγεθών που μεταβάλλονται ως προς το χρόνο, θα ήταν λάθος να θεωρηθεί ότι ο χρόνος αποτελεί και την αιτία των παρατηρουμένων μεταβολών.

2.2.1 Πολυγωνικές Γραμμές

Η πολυγωνική γραμμή (line) είναι η πιο διαδεδομένη γραφική παρουσίαση χρονικών στοιχείων επειδή έχει τη δυνατότητα να αναπαριστά άμεσα την πορεία που έχει καταγράψει μέσα στο χρόνο το υπό μελέτη μέγεθος. Μέσα από αυτό το διάγραμμα μπορούμε να δούμε αν υπάρχει κάποια νομοτέλεια ως προς την εξέλιξη της στο χρόνο, όπως η μακροχρόνια τάση (ανοδική ή καθοδική) ή περιοδικές διακυμάνσεις (εποχικότητα, κυκλικότητα) που παριστούν πολλά φαινόμενα. Για να πετύχουμε τη σωστή γραφική απεικόνιση θα χρειαστούμε τη βοήθεια των ορθογωνίων συντεταγμένων ακλουθώντας τα παρακάτω τεχνικά βήματα:

<u>Οδηγίες Κατασκευής Πολυγωνικών Γραμμών</u>

- Στον οριζόντιο άξονα, που στην περίπτωση μας ονομάζεται αξονας του χρόνου, λαμβάνονται ισομήκη τμήματα τα οποία παριστούν συμβατικά τις διαδοχικές χρονικές στιγμές ή περιόδους (t=1,2,3,…,n).
- Στον κάθετο άξονα και με την κατάλληλη σε κάθε περίπτωση κλίμακα μετριούνται οι αντίστοιχες τιμές Y_t της υπό μελέτη μεταβλητής. Η κλίμακα μέτρησης του άξονα των Y_t πάντα θα αρχίζει από το μηδέν.
- Απεικονίζονται στο επίπεδο τα διαθέσιμα αριθμητικά ζεύγη τιμών (t,Y_t).
- Στη συνέχεια ενώνουμε αυτά τα σημεία (t,Y_t) με ευθύγραμμα τμήματα και λαμβάνουμε μια τεθλασμένη ή πολυγωνική γραμμή, η οποία παρουσιάζει τις μεταβολές που παρατηρούνται κατά τη διάρκεια του χρόνου στις φυσικές τιμές της Y_t.

<u>Πρακτικές συμβουλές</u>

- Σε ένα σχεδιάγραμμα μπορούμε να απεικονίσουμε περισσότερες από μια χρονοσειρές. Στην περίπτωση αυτή, όλα τα στοιχεία θα πρέπει να παραπέμπουν στις ίδιες χρονικές στιγμές ή περιόδους και να αναφέρονται στις ίδιες μονάδες μέτρησης.
- Αν ένα σχεδιάγραμμα περιλαμβάνει περισσότερες από μια πολυγωνικές γραμμές , θα πρέπει να σχεδιαστούν με διαφορετικές γραμμοσκιάσεις ή χρώματα. Επεξηγηματικό υπόμνημα στην περίπτωση αυτή κρίνεται απαραίτητο.
- Η γραφική απεικόνιση πολλών ταυτοχρόνως χρονοσειρών (άνω των τεσσάρων) θα πρέπει γενικά να αποφεύγεται διότι δυσχεραίνεται η κατανόηση της εξέλιξης των φαινομένων ιδίως στις περιπτώσεις που οι πολυγωνικές γραμμές τέμνονται πολλαπλά μεταξύ τους.

Ο Πίνακας 2.1 παρουσιάζει ετήσια δεδομένα του αριθμού των Ελλήνων μονίμων μεταναστών για την περίοδο 1960 – 1970, δηλαδή για τη δεκαετία της μεγάλης μετανάστευσης. Το Σχήμα 2.1 παρουσιάζει τα στοιχεία της στήλης (1) του πίνακα αυτού. Το πρώτο σημείο προσδιορίζεται από τις συντεταγμένες που προκύπτουν από το πρώτο έτος (1960) και το πλήθος των μεταναστών του έτους αυτού (=47.768), κ.ο.κ.

ΠΙΝΑΚΑΣ 2.1

Μόνιμοι Έλληνες Μετανάστες κατά φύλο: 1960-1970

Έτος	Σύνολο	Άνδρες	Γυναίκες	Αναλογία Φύλων
1960	47768	33278	14490	230
1961	58828	36200	22628	160
1962	84054	51868	32186	161
1963	100072	61966	38106	163
1964	105569	66265	39304	167
1965	117167	65341	51826	126
1966	86896	46369	40527	114
1967	42730	22885	19845	115
1968	50866	27232	23634	115
1969	91552	51633	39919	129
1970	92681	53030	39651	134

Πηγή: ΚΛΕΩΝ ΤΣΙΜΠΟΣ – ΦΩΤΗΣ ΓΕΩΡΓΙΑΚΩΔΗΣ

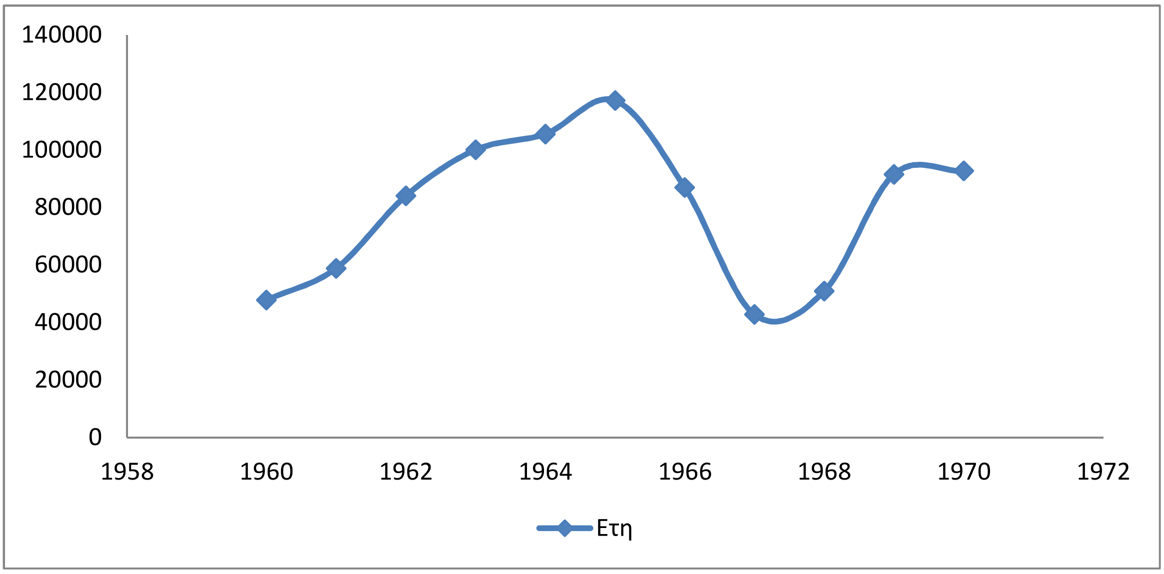
Περιγραφική & Διερευνητική Στατιστική Ανάλυση Δεδομένων, Σελ 167

Το Σχήμα 2.2 παρουσιάζει γραφικά την εξέλιξη των Ελλήνων μεταναστών κατά φύλο – στήλες (3) και (4) του Πίνακα 2.1 – από το οποίο διαπιστώνουμε την πλειονότητα των ανδρών (ιδίως για τα έτη 1960-1965) έναντι του γυναικείου πληθυσμού που πήρε μέρος στην μετανάστευση για την ιδία περίοδο. Πιο συγκεκριμένα στο σχήμα αυτό η πολυγωνική γραμμή έχει αντικατασταθεί από μια ομαλή συνεχή καμπύλη, τεχνική που χρησιμοποιείται, ή επιτυγχάνεται στη συνεχεία, όταν αριθμός των χρησιμοποιούμενων παρατηρήσεων είναι μεγάλος. Ο επιλεκτικός ως προς το φύλο χαρακτήρας της μετανάστευσης παρουσιάζεται επίσης με την πολυγωνική γραμμή του Σχήματος 2.3 το οποίο βασίζεται στα δεδομένα της στήλης (5) του Πίνακα 2.1.

Ο όρος απλή πολυγωνική γραμμή (simple line chart) αναφέρεται στη διαγραμματική παρουσίαση των δεδομένων μιας μόνο χρονοσειράς (Σχήμα 2.1) ενώ ο όρος πολλαπλή πολυγωνική γραμμή (multiple line chart) χρησιμοποιείται στην περίπτωση ταυτόχρονης παρουσίασης περισσοτέρων της μια χρονοσειράς (Σχήμα 2.2).

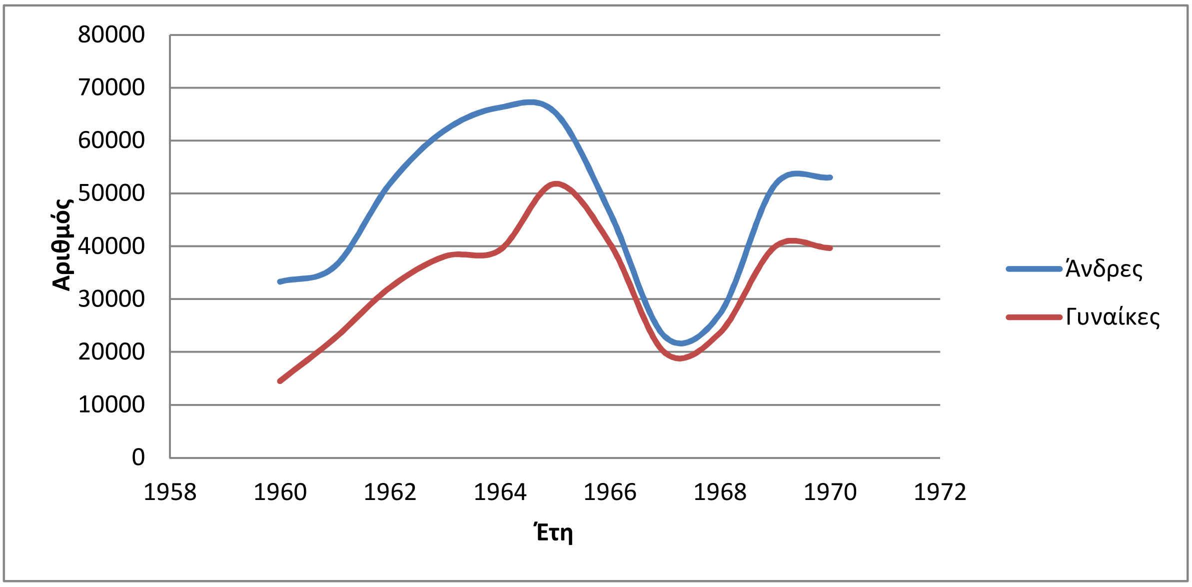
ΣΧΗΜΑ 2.1

Έλληνες μόνιμοι μετανάστες : 1960-70



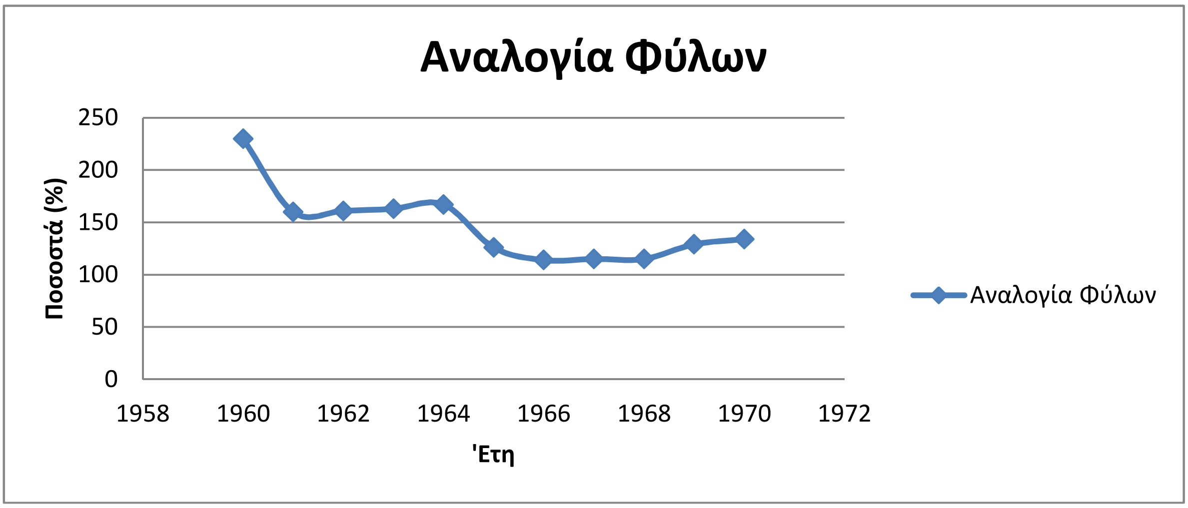
ΣΧΗΜΑ 2.2

Έλληνες μόνιμοι μετανάστες κατά φύλο : 1960-70



ΣΧΗΜΑ 2.3

Άνδρες μετανάστες ανά 100 – Μετανάστριες: 1960-70



2.2.2 Ραβδογράμματα

Η παρουσίαση γραφικά των τιμών **Yt** μιας χρονοσειράς γίνεται και με κάθετα ραβδογραμμάτα (columns). Για τη δημιουργία των απεικονίσεων αυτών η τεχνική που θα πρέπει να χρησιμοποιηθεί στην ουσία είναι ίδια με εκείνη της διαγραμματικής παρουσίασης των ονομαστικών μεταβλητών. Οι βασικές διαφορές εντοπίζονται στα εξής σημεία α) κατά μήκος του οριζόντιου άξονα αντί των συμβολικών τιμών της ονομαστικής μεταβλητής απεικονίζονται οι χρονικές περίοδοι στις οποίες αναφέρονται τα στοιχεία (έτη, μήνες, κ.λ.π.) με ισομήκη ευθύγραμμα τμήματα και β) στον κάθετο άξονα, αντί των απόλυτων ή σχετικών συχνοτήτων της ποσοτικής κατανομής, απεικονίζεται η κλίμακα μέτρησης των τιμών της χρονοσειράς. Το ύψος των ακίδων είναι ανάλογο των τιμών **Yt**.

Το Σχήμα 2.4 απεικονίζει με ραβδογράμματα την εξέλιξη των Ελλήνων μεταναστών της περιόδου 1960-1970-δεδομένα των στηλών (1) και (2) του Πίνακα 1.1. Η σύνδεση των κορυφών των ορθογωνίων (Σχήμα 2.5) δημιουργεί τεθλασμένη γραμμή που διαγράφει πορεία ίδια με εκείνη της πολυγωνικής γραμμής του Σχήματος 1.1.

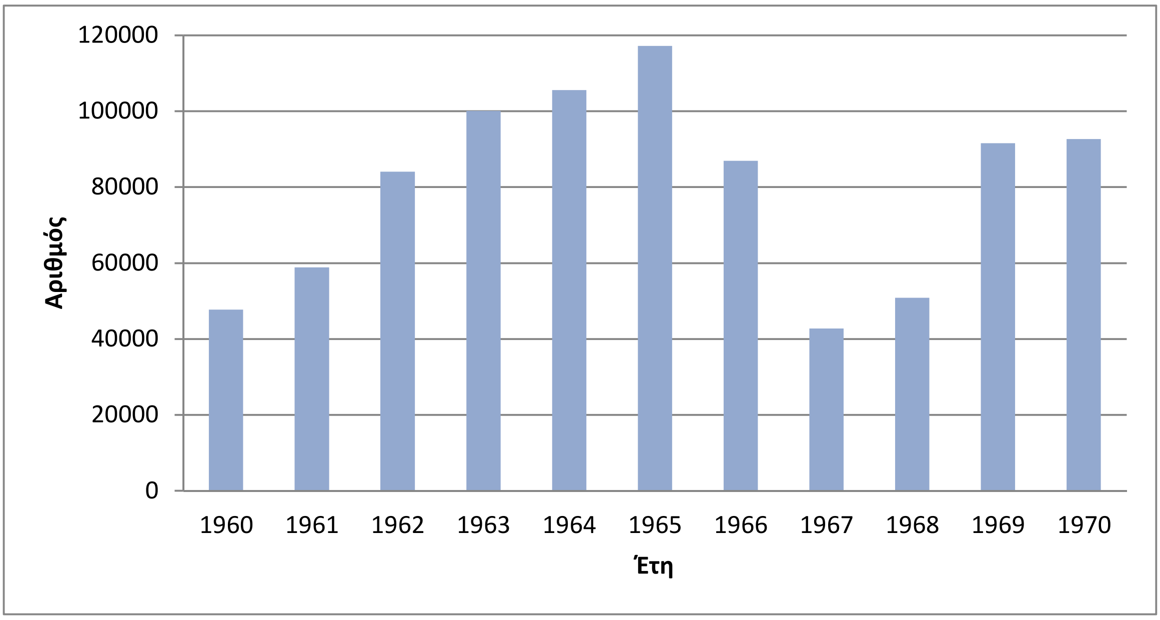
Οι τιμές της **Yt** που απεικονίζονται στα κάθετα ραβδογράμματα, μπορεί να είναι δυο ή τριών διαστάσεων. Σε αυτή την περίπτωση, όταν συνδέσουμε τις κορυφές των τρισδιάστατων ράβδων θα σχηματιστεί ΄΄ πολυγωνική γραμμή τριών διαστάσεων ΄΄ η οποία αποκτά τη μορφή ταινίας (Σχήμα 2.6).

Η απεικόνιση του διαγράμματος της πορείας των μεταναστών κατά φύλο είναι δυνατόν να παρουσιαστεί είτε με σωρευτικά κάθετα ραβδογράμματα (Σχήμα 2.7) είτε με συζυγή κάθετα ραβδογράμματα (μία παραλλαγή τέτοιας απεικόνισης παρουσιάζει το Σχήμα 2.8).

Θα πρέπει να παρατηρήσουμε ότι παρόλο που ο χρόνος (t) είναι συνεχής μεταβλητή, μεταξύ των καθέτων ορθογωνίων παραλληλογράμμων (ράβδων) παρεμβάλλονται ισομήκα κενά τμήματα. Αυτό συμβαίνει διότι στην πράξη θεωρούμε, ότι οι μετρήσεις ή απαριθμήσεις των υπό μελέτη μεγεθών (**Yt**) διενεργούνται σε τακτές ΄΄ διακριτές ΄΄ χρονικές στιγμές.

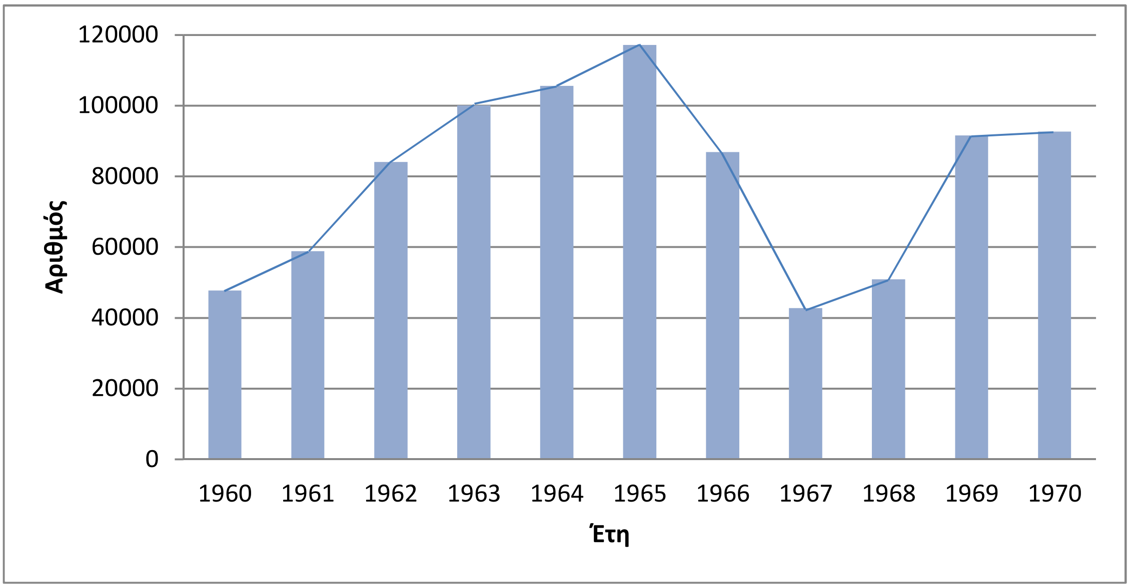
ΣΧΗΜΑ 2.4

Έλληνες μόνιμοι μετανάστες: 1960-70



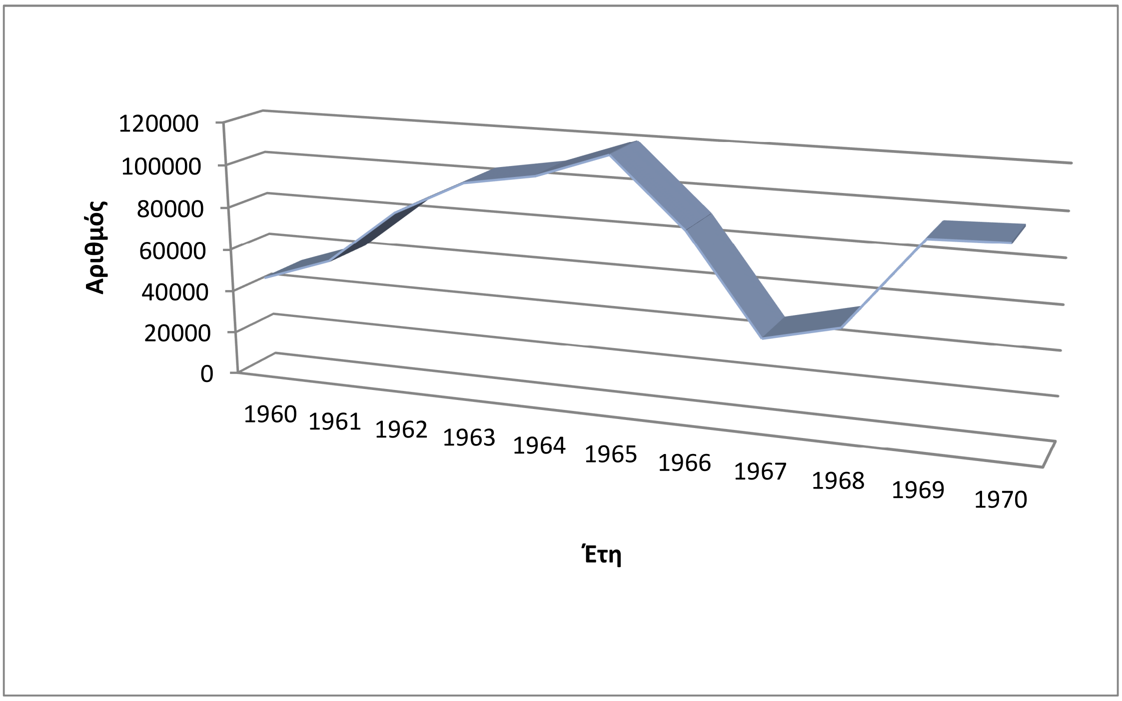
ΣΧΗΜΑ 2.5

Έλληνες μόνιμοι μετανάστες: 1960-70



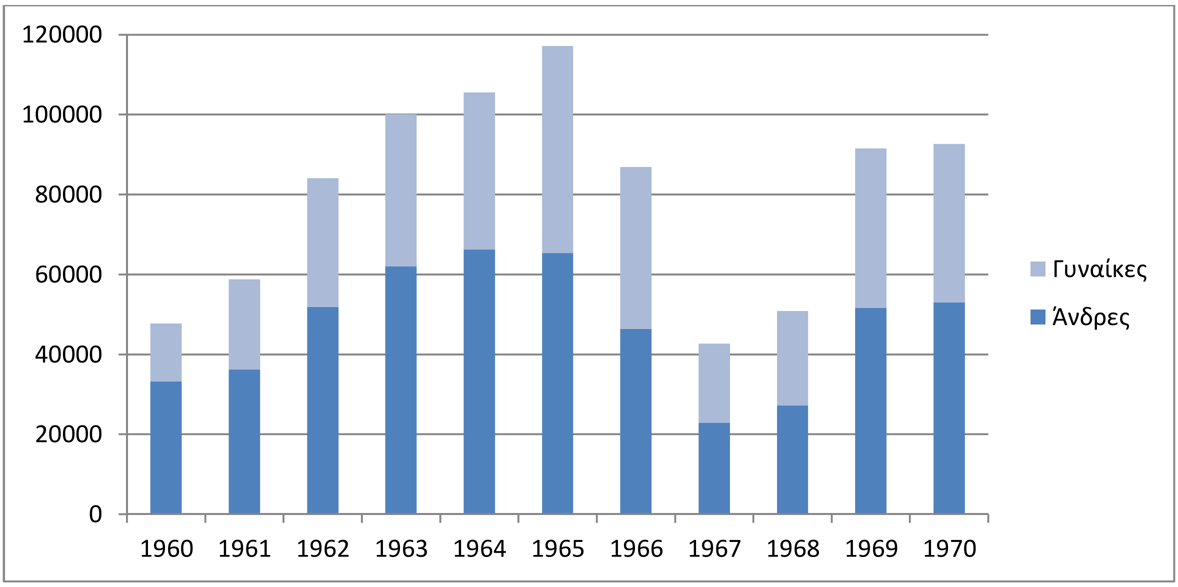
ΣΧΗΜΑ 2.6

Έλληνες μόνιμοι μετανάστες: 1960-70



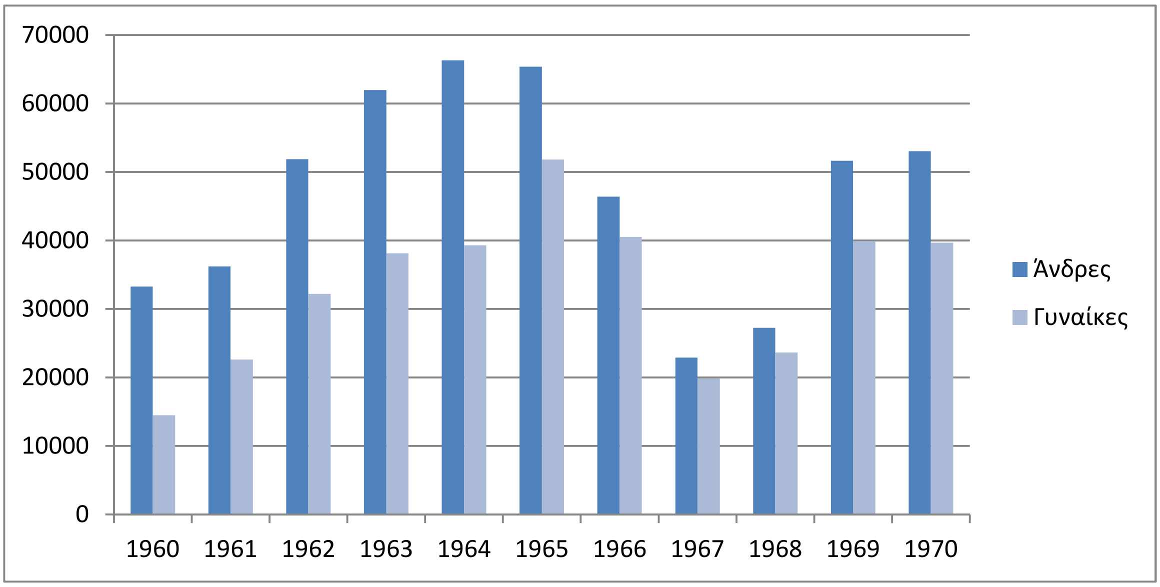
ΣΧΗΜΑ 2.7

Έλληνες μόνιμοι μετανάστες κατά φύλο: 1960-70



ΣΧΗΜΑ 2.8

Έλληνες μόνιμοι μετανάστες κατά φύλο: 1960-70



2.2.3 Διαγράμματα Επιφανειών

Η απεικόνιση χρονολιγικών στοιχείων με **διάγραμμα – επιφάνεια (area)** δίνει σε πολύ αδρές γραμμές την εξέλιξη ενός πληθυσμού ή ενός οικονομικού μεγέθους, και χρησιμοποιείται όταν ο αριθμός των παρατηρήσεων είναι μικρός και ο πληθυσμός (ή το υπό μελέτη μέγεθος) αποτελείται από μικρό επίσης αριθμό υποπληθυσμών (ή διακριτών κατηγοριών). Οι διαγραμματικές παρουσιάσεις που ακολουθούν, βασίζονται στα ελληνικά απογραφικά δεδομένα και αναφέρονται στην κατανομή του πληθυσμού κατά αστικές, ημι-αστικές και αγροτικές περιοχές (Πίνακας 2.2).

ΠΙΝΑΚΑΣ 2.2

Πληθυσμός της Ελλάδος κατά περιοχές: Απογραφές 1951 – 1991

Έτη (1)	Σύνολο (2)	Αστικές (3)	Ημι-αστικές (4)	Αγροτικές (5)
α) Αριθμοί				
1951	7632801	2879994	1130188	3622619
1961	8388553	3628105	1085856	3674592
1971	8768641	4667489	1019421	3081731
1981	9740417	5659528	1125547	2955342
1991	10259900	6036660	1312774	2910466
β) Ποσοστά (%)				
1951	100,00	37,7	14,8	47,5
1961	100,00	43,3	12,9	43,8
1971	100,00	53,2	11,6	35,1
1981	100,00	58,1	11,6	30,3
1991	100,00	58,8	12,8	28,4

Πηγή: ΚΛΕΩΝ ΤΣΙΜΠΟΣ – ΦΩΤΗΣ ΓΕΩΡΓΙΑΚΩΔΗΣ

Περιγραφική & Διερευνητική Στατιστική Ανάλυση Δεδομένων, Σελ 173

Στην πιο απλή της μορφή, η εν λόγω απεικόνιση είναι στην ουσία μια πολυγωνική γραμμή με γραμμοσκιασμένη ή χρωματισμένη την επιφάνεια που περιλαμβάνεται μεταξύ της τεθλασμένης αυτής γραμμής και του οριζόντιου άξονα του χρόνου. Το Σχήμα 2.9 παρουσιάζει τον τρόπο με τον οποίο εξελίχθη το συνολικό μέγεθος του πληθυσμού της Ελλάδος κατά την τεσσαρακονταετία 1951 – 1991. Το Σχήμα 2.10 εμφανίζει και την πορεία των επί μέρους υποπληθυσμών των αστικών, ημιαστικών και αγροτικών περιοχών της χώρας. Εδώ το σχεδιάγραμμα αποκτά αθροιστική μορφή και δημιουργείται με ανάλογο τρόπο με εκείνον που περιγράψαμε κατά την κατασκευή των σωρευτικών ακιδωτών διαγραμμάτων. Το άνω μέρος της επιφάνειας (έχει σχεδιαστεί με συνεχή γραμμή) διαγράφει την πορεία του συνολικού πληθυσμού της χώρας, ενώ οι τρείς διαφορετικού χρώματος επιφάνειες σκιαγραφούν την εξέλιξη των επί μέρους υποπληθυσμών. Στα δύο προαναφερθέντα σχήματα ο κάθετος άξονας μετράει τον πληθυσμό σε απόλυτα μεγέθη. Το Σχήμα 2.11 παρουσιάζει τις μεταβολές στην ποσοστιαία κατανομή του πληθυσμού κατά επίπεδο αστικότητας. Στην προκειμένη περίπτωση, το περίγραμμα της γραφικής απεικόνισης έχει την μορφή ενός παραλληλογράμμου και δείχνει με πιο εύληπτο τρόπο την πορεία της αστικοποίησης του Ελληνικού πληθυσμού και τη μείωση της ποσοστιαίας συμμετοχής του πληθυσμού των αγροτικών περιοχών.

Τα Σχήματα 2.12, 2.13 και 2.14 παρουσιάζουν με τρισδιάστατες επιφάνειες την εξέλιξη του Ελληνικού πληθυσμού κατά περιοχές για την εξεταζόμενη περίοδο. Όπως και στην περίπτωση των ραβδογραμμάτων, το αποτέλεσμα είναι ενδιαφέρον όμως δεν προσδίδει πάντα περισσότερες πληροφορίες από πλευράς στατιστικής περιγραφής του φαινομένου.

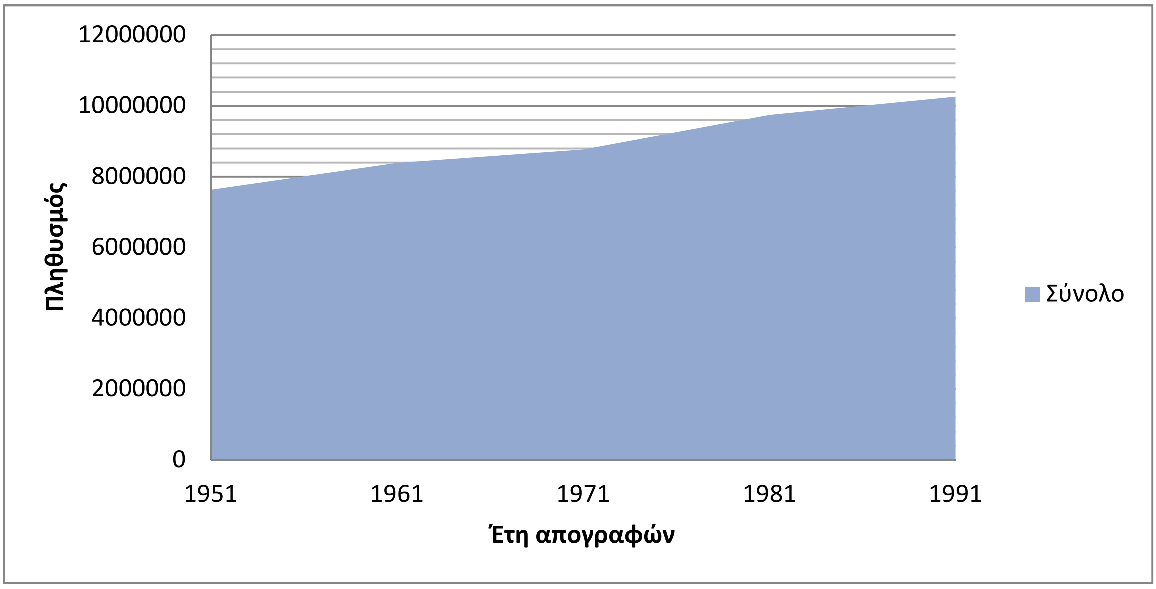
Θα ολοκληρώσουμε την παράγραφο αυτή με τις εξής τρείς επισημάνσεις:

Πρακτικές συμβουλές κατασκευής διαγραμμάτων επιφανειών

- Στην περίπτωση δισδιάστατων επιφανειών (όπως αυτή του Σχήματος 2.10 και 2.11) καλό είναι η επιφάνεια που εφάπτεται του άξονα του χρόνου να αναφέρεται στη μεγαλύτερη, από πλευράς μεγέθους, κατηγορία.
- Στην περίπτωση συζυγών τρισδιάστατων επιφανειών (όπως εκείνη του Σχήματος 2.14) οι επί μέρους υποπληθυσμοί ή υπο-κατηγορίες θα πρέπει να εμφανίζονται κατά σειρά μεγέθους ξεκινώντας από το μικρότερο προς μεγαλύτερο, αν αυτό βέβαια επιτρέπεται από τα στοιχεία.
- Αν κατά διάρκεια της υπό μελέτη χρονικής περιόδου τα στοιχεία παρουσιάζουν σημαντικές διακυμάνσεις (αυξο-μειώσεις) των επί μέρους μεγεθών, καλό είναι να αποφεύγεται η κατασκευή επιφανειών τριών διαστάσεων.

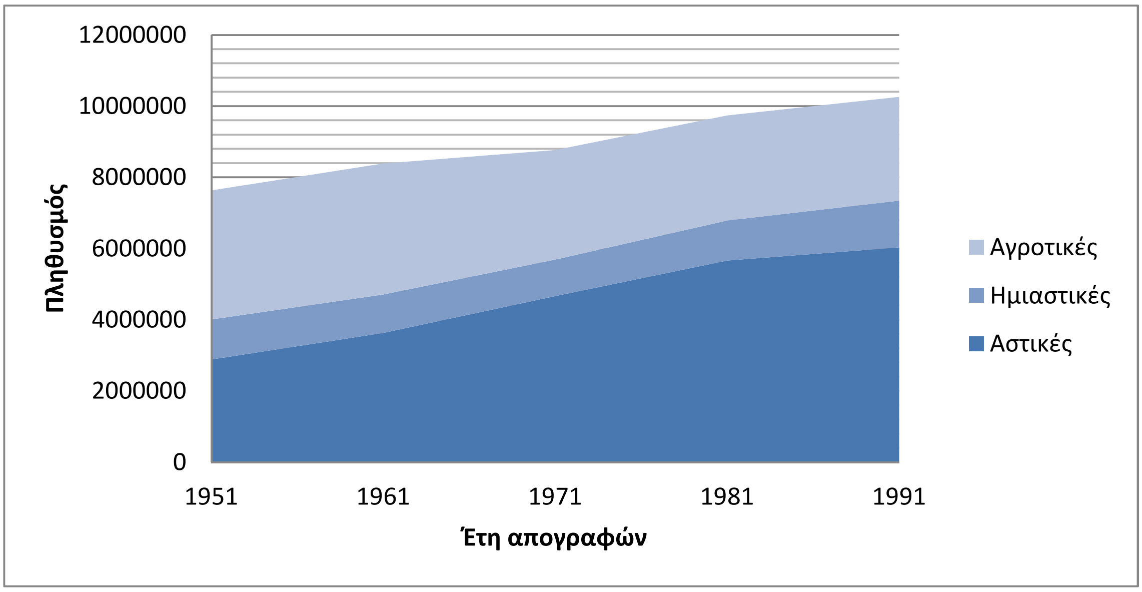
ΣΧΗΜΑ 2.9

Πληθυσμός της Έλλαδος: 1951-91



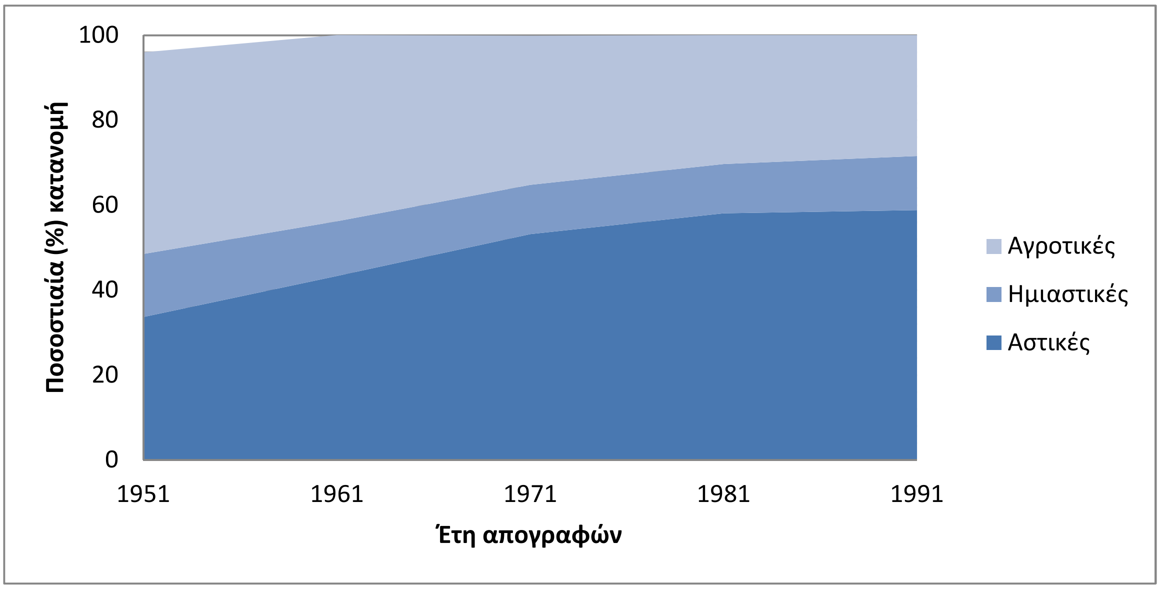
ΣΧΗΜΑ 2.10

Πληθυσμός της Ελλάδος κατά περιοχές: 1951-1991



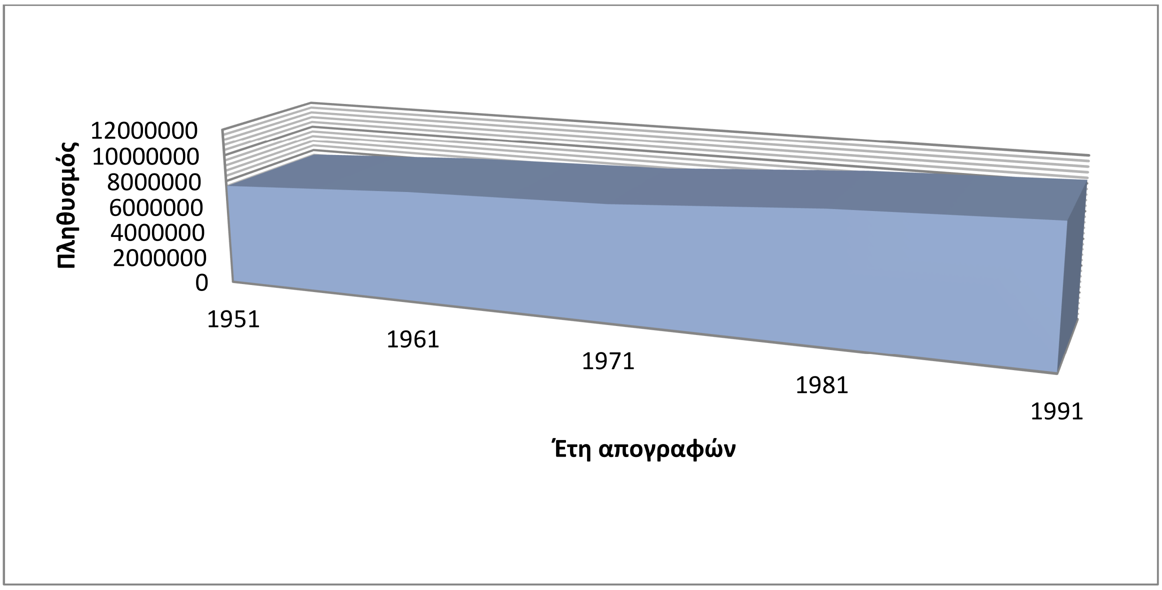
ΣΧΗΜΑ 2.11

Πληθυσμός της Ελλάδος:1951-91



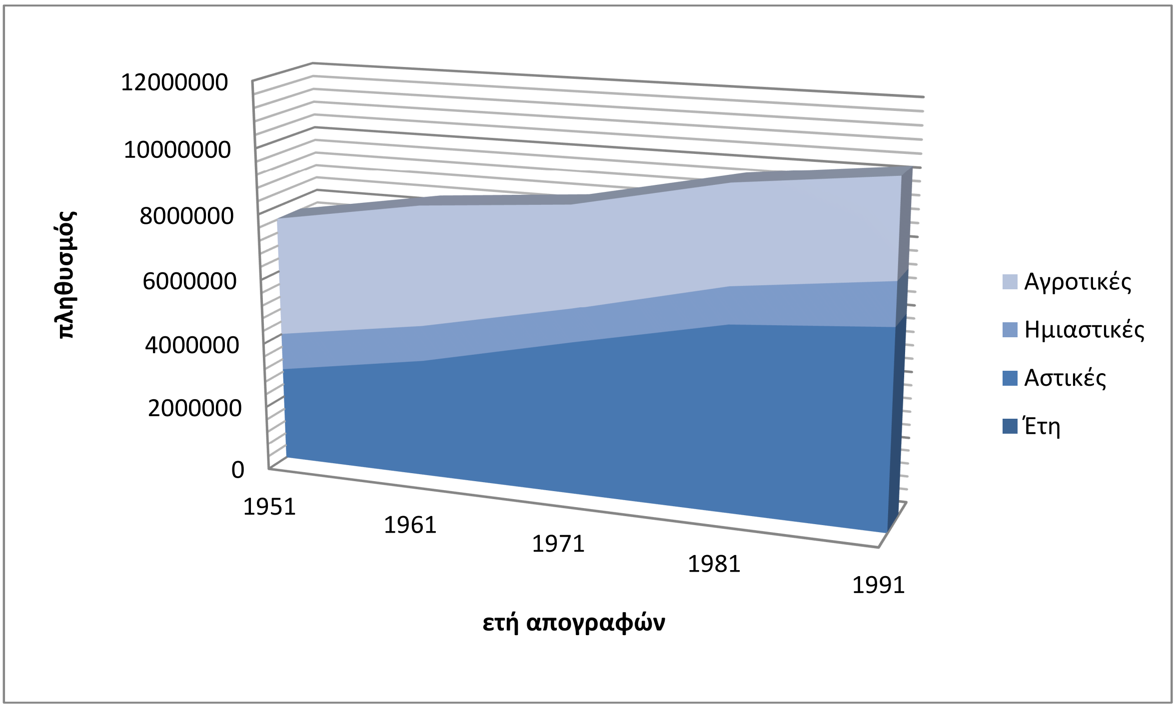
ΣΧΗΜΑ 2.12

Συνολικός Πληθυσμός της Ελλάδος: 1951-91



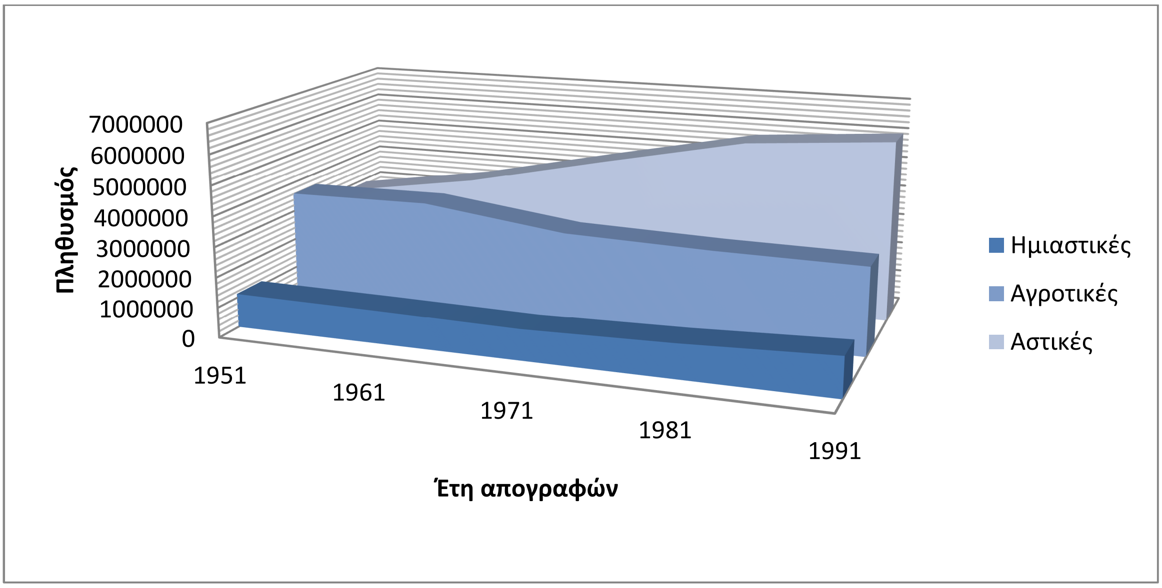
ΣΧΗΜΑ 2.13

Πληθυσμός της Ελλάδος κατά περιοχές: 1951-91



ΣΧΗΜΑ 2.14

Πληθυσμός της Ελλάδος κατά περιοχές: 1951-91



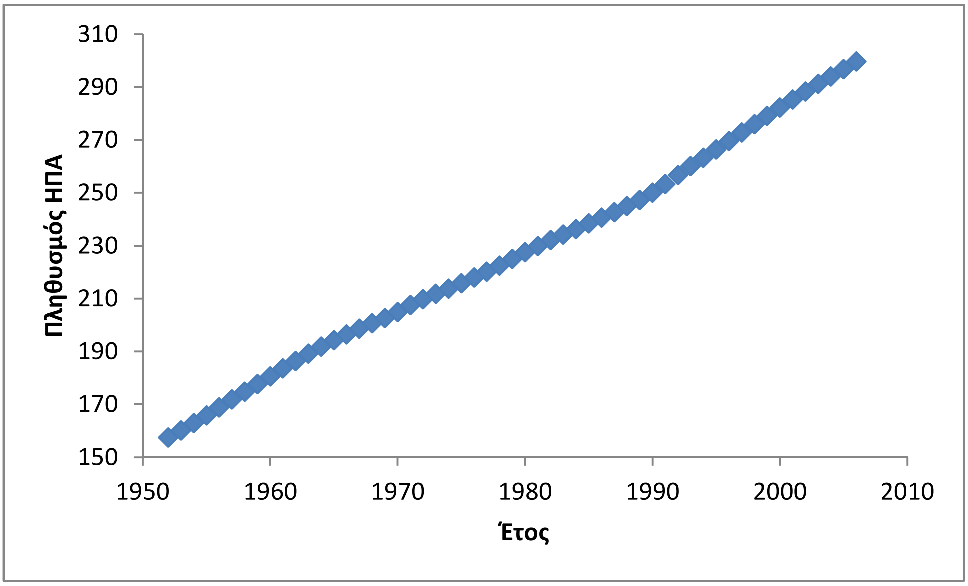
2.3 ΣΥΝΙΣΤΩΣΕΣ ΧΡΟΝΟΣΕΙΡΩΝ

Μια χρονοσειρά μπορεί να έχει τέσσερις διαφορετικές συνιστώσες. Οι συνιστώσες των χρονοσειρών είναι οι εξής:

1. Μακροχρόνια τάση.
2. Κυκλική μεταβλητότητα.
3. Εποχιακή μεταβλητότητα ή Περιοδική μεταβλητότητα.
4. Τυχαία μεταβλητότητα ή Ακανόνιστη μεταβλητότητα.

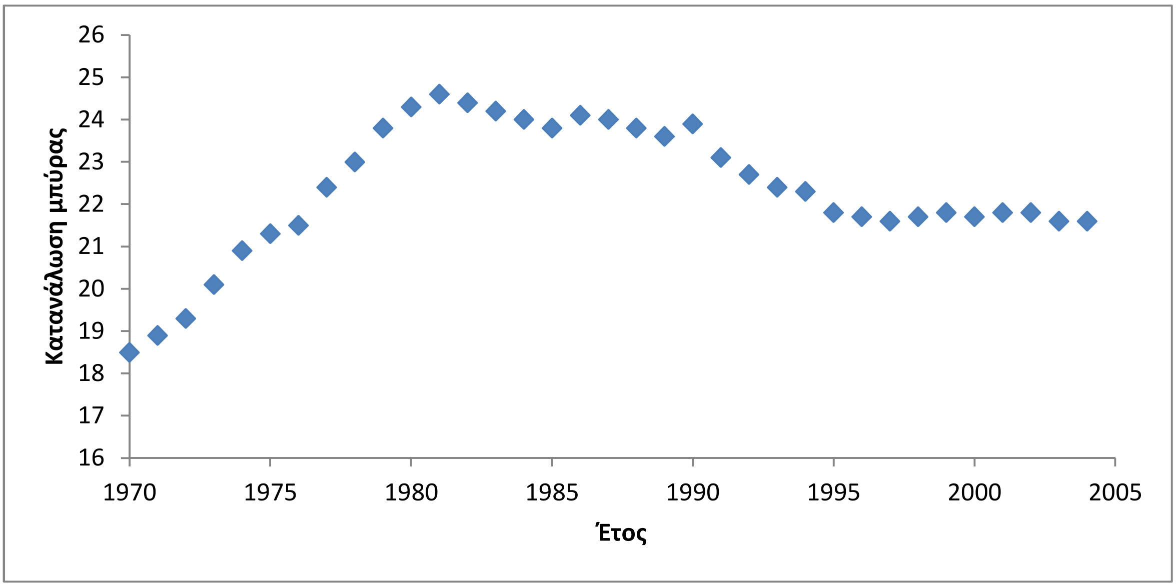
Μια μακροχρόνια τάση (long-term trend) είναι μια σχετικά σταθερή κατεύθυνση που χαρακτηρίζει το σύνολο της χρονοσειράς. Σε αυτό το σημείο θα δανειστούμε τα δεδομένα από το βιβλίο του GERALD KELLER «Στατιστική για οικονομικά και Διοίκηση Επιχειρήσεων σελ.1035». Για παράδειγμα όπως φαίνεται στην Εικόνα 1.1 ο πληθυσμός των ΗΠΑ διατήρησε μια σταθερή ανοδική τάση από 157 εκατομμύρια το 1950 σε 299 εκατομμύρια το 2006.

Εικόνα 2.1 Πληθυσμός των ΗΠΑ, 1950-2006 (σε εκατομμύρια)



Η τάση μιας χρονοσειράς δεν είναι πάντοτε γραμμική. Παραδείγματος χάρη στην Εικόνα 2.2 βλέπουμε την κατά κεφαλή κατανάλωση μπύρας από τον πληθυσμό άνω των 21 ετών στις ΗΠΑ, από το 1970 ως το 2004. Σε αυτή την περίπτωση μπορούμε να παρατηρήσουμε πως η τάση αρχικά είναι μια καμπύλη που αρχικά είναι ανοδική για μια δεκαετία και στη συνέχεια άλλοτε λιγότερο και άλλοτε περισσότερο καθοδική.

Εικόνα 2.2 Κατά κεφαλή κατανάλωση μπύρας στις ΗΠΑ, 1970-2004

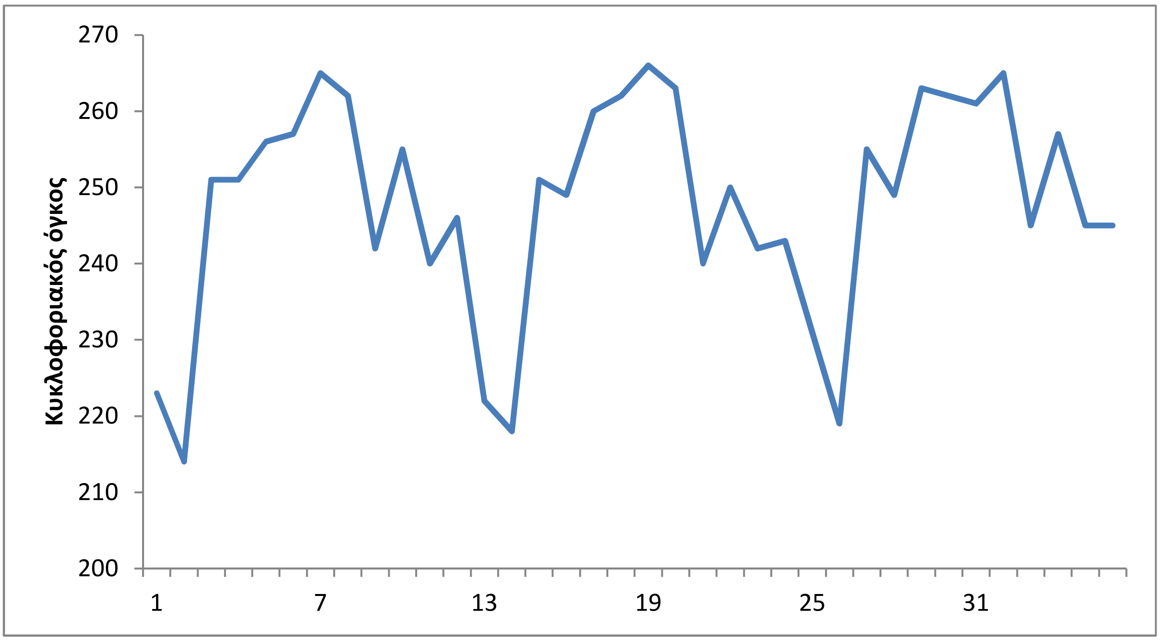


Η κυκλική μεταβλητότητα (cyclical variation) είναι μια επαναλαμβανόμενη κίνηση που διαρκεί ένα μικρότερο ή μεγαλύτερο αριθμό χρόνων. Κλασικό παράδειγμα κυκλικής μεταβλητότητας είναι οι οικονομικοί κύκλοι ανάπτυξης και ύφεσης, ή οι μακροχρόνιες αυξομειώσεις στη ζήτηση ενός προϊόντος.

Η εποχιακή μεταβλητότητα ή περιοδική μεταβλητότητα (seasonal variation) είναι μια επαναλαμβανόμενη κίνηση που είναι παρούσα σε κάθε ημερολογιακή περίοδο. Ο όρος εποχιακή προέρχεται από τον ετήσιο κύκλο των εποχών (π.χ. ζήτηση τουριστικών υπηρεσιών) αλλά η περίοδος μπορεί να έχει οποιαδήποτε διάρκεια π.χ. μηνιαία, εβδομαδιαία κ.τ.λ. Για παράδειγμα η πληρότητα ενός ξενοδοχείου έχει ένα μηνιαίο κύκλο.

Στην Εικόνα 2.3 βλέπουμε το διάγραμμα του μηνιαίου κυκλοφοριακού όγκου των ΗΠΑ (σε δισεκατομμύρια μίλια) από το 2000 ως το 2002. Το διάγραμμα συνδυάζει μια μακροχρόνια ανοδική τάση και μια εποχιακή ετήσια μεταβλητότητα. Με απλά λόγια, τους θερινούς μήνες οι Αμερικανοί οδηγούν περισσότερο, ενώ συνολικά κάθε χρόνο οδηγούν λίγο περισσότερο από τον προηγούμενο.

Εικόνα 2.3 Κυκλοφοριακός όγκος στις ΗΠΑ, 2000-2002



Η τυχαία μεταβλητότητα ή ακανόνιστη μεταβλητότητα (random variation) είναι η μεταβλητότητα που δεν εξηγείται από τις προηγούμενες συνιστώσες και οφείλεται σε απρόβλεπτα γεγονότα. Η τυχαία μεταβλητότητα έχει την ικανότητα να καλύπτει την ομοιομορφία των υπόλοιπων συνιστωσών. Ένας από τους στόχους της ανάλυσης είναι η απομόνωση της τυχαίας μεταβλητότητας, ώστε να μπορούν να αναγνωριστούν οι προβλέψιμες συνιστώσες μιας χρονοσειράς.

2.3.1 Μακροχρόνια τάση

Αν για μια μεγάλη περίοδο οι τιμές μιας χρονολογικής σειράς τείνουν να αυξηθούν ή να μειωθούν ή να παρουσιάσουν σύνθετη μορφή, τότε λέμε ότι η σειρά των παρατηρήσεων παρουσιάζει μακροχρόνια τάση. Δηλαδή, τάση είναι η μακροχρόνια αύξηση ή μείωση που παρατηρείται στα δεδομένα μιας χρονοσειράς.

Το χαρακτηριστικό της μακροχρόνιας τάσης είναι ότι έχουμε μακροχρόνια και σταθερή κίνηση των οικονομικών μεγεθών που επηρεάζεται από γενικότερους παράγοντες και ωθεί τα δεδομένα μας προς ορισμένη κατεύθυνση. Στις βιομηχανικές χώρες η τάση αυτή ανταποκρίνεται στο φαινόμενο της οικονομικής μεγέθυνσης, το οποίο μεταφράζεται από μια γρήγορη αύξηση της παραγωγής και των συναλλαγών, ιδιαίτερα από το τέλος του δευτέρου παγκοσμίου πολέμου μέχρι σήμερα. Ιδιαίτερα παρατηρήθηκε ότι σε μερικές χώρες υπήρχαν μακροχρόνιες μεταβολές, οι οποίες περιελάμβαναν ανοδικές και καθοδικές φάσεις. Σ’ αυτές τις μεταβολές δόθηκε το όνομα «κύκλοι Kondratieff» από το όνομα του Ρώσου οικονομολόγου και στατιστικού, ο οποίος μελέτησε διεξοδικά τα φαινόμενα αυτά. Οι κύκλοι αυτοί αναφέρονται σε μια μεγάλη χρονική περίοδο της τάξης 50 έως και 60 χρόνων. Από την παραπάνω ανάλυση προκύπτει ότι, αν για μια μεγάλη χρονική περίοδο οι τιμές της μεταβλητής Ψ μιας χρονολογικής σειράς αυξάνονται ή μειώνονται, τότε η σειρά των παρατηρήσεων της μεταβλητής παρουσιάζει μακροχρόνια τάση. Από το Διάγραμμα 2.1 φαίνεται ότι η καμπύλη η οποία έρχεται από τα κάτω αριστερά προς τα πάνω δεξιά έχει αυξητική μακροχρόνια τάση, η οποία θα μπορούσε να εκφράσει τις πωλήσεις αυτοκινήτων τα τελευταία πενήντα χρόνια σε παγκόσμιο επίπεδο. Διάφοροι παράγοντες επέδρασαν προς την κατεύθυνση αυτή, όπως η συνεχής καλυτέρευση του βιοτικού επιπέδου, η τεχνολογική πρόοδος, η αύξηση του πληθυσμού, η αύξηση των μετακινήσεων του πληθυσμού κ.τ.λ. Αντίθετα η καμπύλη η οποία κατέρχεται από τα πάνω αριστερά προς τα κάτω δεξιά έχει καθοδική μακροχρόνια τάση, η οποία θα μπορούσε να εκφράσει π.χ. την κατανάλωση φωταερίου στην Ελλάδα. Διάφοροι παράγοντες επέδρασαν προς την κατεύθυνση αυτή. Ο σημαντικότερος απ’ αυτούς θεωρείται η εγκατάσταση και η λειτουργία δικτύου ηλεκτρικού ρεύματος σ’ όλη τη χώρα. Γενικά θα μπορούσαμε να χαρακτηρίσουμε ως παράγοντες που επενεργούν μειωτικά πάνω στη μακροχρόνια τάση:

a) Την εμφάνιση υποκαταστημάτων αγαθών (λάδι ελιάς – σπορέλαιο, βούτυρο – μαργαρίνη).
b) Τη μεταβολή της ζήτησης ενός άλλου αγαθού, που ικανοποιεί τις ίδιες ανάγκες (σιδηρόδρομος – αεροπλάνο, βενζίνη super – βενζίνη αμόλυβδη).

Διαγραμα 2.1

2.3.2 Περιοδικές μεταβολές

Οι περιοδικές μεταβολές είναι εκείνες που επαναλαμβάνονται κατά ορισμένα χρονικά διαστήματα μέσα σε ορισμένη χρονική περίοδο. Πιο συγκεκριμένα μια εποχιακή μεταβλητότητα είναι μια χρονολογική σειρά διακυμάνσεων, που συμβαίνει από έτος σε έτος κατά το ίδιο χρονικό σημείο, του έτους ή, περισσότερο γενικά, μια διακύμανση που συμβαίνει από χρονικό διάστημα σε χρονικό διάστημα (μήνας, εβδομάδα, έτος) περίπου στο ίδιο σημείο μέσα στο χρονικό διάστημα. Τέτοιες διακυμάνσεις πρέπει να λαμβάνονται υπόψη όταν καταρτίζουμε προγράμματα παραγωγής, πωλήσεων, αποθεμάτων. Αυτές επίσης, πρέπει να λαμβάνονται υπόψη όταν προσπαθούμε να εκτιμήσουμε αν μια μεταβολή στις πωλήσεις είναι μια αναμενόμενη εποχιακή μεταβολή ή αν αυτό είναι συνέπεια μιας μεταβολής στην τάση της συγκεκριμένης επιχειρηματικής δραστηριότητος.

2.3.3 Μέτρηση της Εποχιακής μεταβλητότητας

Στο τέλος κάθε έτους, η Επιχείρηση Β διαμορφώνει κατ΄ εκτίμηση ένα επίπεδο πωλήσεων για το ερχόμενο έτος ως μια βάση για το σχεδιασμό των οικονομικών απαιτήσεων. Για να υπολογιστεί από χρόνο σε χρόνο εάν οι πραγματικές πωλήσεις είναι σύμφωνες με τις προγραμματισμένες θεωρήθηκε απαραίτητο να κατανεμηθεί η ετήσια ποσότητα μεταξύ των διαφόρων περιόδων – πωλήσεων του έτους. Με δεδομένο ότι είναι γνωστό ότι οι τελευταίοι μήνες της ανοίξεως και του καλοκαιριού είναι περίοδοι χαμηλών πωλήσεων σε σύγκριση με την αρχή και το τέλος του χειμώνα, και ότι η περίοδος του τέλους του χειμώνα και της αρχής της άνοιξης καταλαμβάνει μια ενδιάμεση θέση στην κλίμακα του ύψους των πωλήσεων του παρελθόντος επιβεβαίωσε την ύπαρξη αυτού του εποχιακού υποδείγματος. Έτσι αποφασίστηκε να διαιρεθεί κάθε έτος σε τρείς περιόδους, θεωρούμενη αυτή η υποδιαίρεση ως η πλέον αξιόπιστη βάση για εποχική ανάλυση. Οι πωλήσεις για την περίοδο 1983-1989 παρουσιάζονται στον Πίνακα 1.3 . Η περίοδος 1 περιλαμβάνει τον Ιανουάριο, Φεβρουάριο, Μάρτιο και Απρίλιο. Περίοδος 2: Μάϊος, Ιούνιος, Ιούλιος και Αύγουστος. Περίοδος 3: Σεπτέμβριος, Οκτώβριος, Νοέμβριος και Δεκέμβριος.

ΠΙΝΑΚΑΣ 2.3

Πωλήσεις της Επιχειρήσεως Β, 1983-1989, κατά περίοδο

Έτος	Πωλήσεις κατά περίοδο (χιλιάδες μονάδων)			Σύνολο
	Περίοδος 1	Περίοδος 2	Περίοδος 3	
1983	19,1	15,6	27,3	62
1984	22,8	18	29,9	70,7
1985	28,9	20,7	33,3	82,9
1986	26,5	17,5	28,1	72,1
1987	24,8	16,1	28,3	69,2
1988	23,8	19	30,4	73,2
1989	28,6	22,8	35,4	86,8

2.3.4 Κυκλικές μεταβολές

Οι κυκλικές μεταβολές αναφέρονται σε διαδοχικές περιόδους «ευημερίας» και «δυσπραγίας», οι οποίες έχουν διάρκεια πολύ μεγαλύτερη του έτους. Οι κυμάνσεις αυτές ονομάσθηκαν κυκλικές, παρά περιοδικές. Κι αυτό γιατί η χρονική τους διάρκεια ήταν μερικά χρόνια, η οποία δεν προέκυψε ότι παρουσίαζε κανονική περιοδικότητα. Στην οικονομική επιστήμη οι κυκλικές μεταβολές είναι γνωστές με τα ονόματα των οικονομικών ή επιχειρηματικών κύκλων. Ο πιο παραδοσιακός κύκλος είναι αυτός που φέρνει το όνομα «μέγας κύκλος» ή κύκλος Juglar η διάρκεια του οποίου είναι από 7 μέχρι 11 χρόνια, δηλαδή μιας μέσης διάρκειας 9 χρόνων. Ο κύκλος αυτός περιλαμβάνει τέσσερις φάσεις την εξάπλωση (expansion), την κρίση (crise), τη χαλάρωση (depression) και την ανάκαμψη (reprise). Τέτοιες κυκλικές κυμάνσεις καταγράφηκαν το 1882, το1890, το 1900, το 1907, το 1921 αλλά και αργότερα. Αντίθετα με τον παραπάνω κύκλο, στις Η.Π.Α. εμφανίστηκαν κύκλοι μικρότερης διάρκειας οι οποίοι ονομάστηκαν «μικροί κύκλοι» ή κύκλοι Kitchin. Η διάρκεια των κύκλων αυτών είναι από μερικούς μήνες μέχρι σαράντα μήνες. Οι παράγοντες που προκαλούν τις κυκλικές κυμάνσεις δεν έχουν εξακριβωθεί. Έχουν διατυπωθεί πολλές θεωρίες, οι οποίες όμως φαίνεται ότι δεν έχουν αποδώσει. Μια απ' αυτές τις θεωρίες αναφέρεται στην επίδραση των ηλιακών κηλίδων.

«Ο οικονομολόγος W.S. Jevons διάσημος ως ένας από τους πρωτεργάτες της ΄΄νεοκλασσσικής΄΄ οικονομικής, ανέπτυξε μια θεωρία που έλεγε ότι οι γενικές οικονομικές διακυμάνσεις προκαλούνται από τις αυξομειώσεις των συγκομιδών στη γεωργία, οι οποίες με τη σειρά τους προκαλούνται από κλιματολογικούς κύκλους που είχαν σχέση με τη δράση των ηλιακών κηλίδων».

2.3.5 Τυχαία μεταβλητότητα ή Ακανόνιστη μεταβλητότητα

Πολλές φορές συμβαίνει να εμφανίζονται στα δεδομένα της χρονολογικής σειράς μικρές ή μεγάλες μεταβολές και χωρίς καμία κανονικότητα ή ρυθμό μεταβολής, οι οποίες ονομάζονται ακανόνιστες ή τυχαίες μεταβολές. Οι ακανόνιστες μεταβολές διακρίνονται σε δύο κατηγορίες: στις συμπτωματικές και στις τυχαίες. Οι συμπτωματικές μεταβολές προέρχονται από εξαιρετικά και απρόβλεπτα γεγονότα, όπως είναι οι σεισμοί, οι θύελλες, οι απεργίες, οι επιδημίες, οι πόλεμοι κ.τ.λ., ενώ οι τυχαίες μεταβολές οφείλονται σε πολυάριθμους άγνωστους παράγοντες ή, όπως συνήθως λέγεται, στην τύχη. Οι παραπάνω ακανόνιστες ή τυχαίες μεταβολές εμφανίζονται κυρίως στην αγροτική παραγωγή, η οποοία επηρεάζεται σημαντικά από τις καιρικές συνθήκες. Αλλά ανάλογα «δυστυχήματα» μπορούν να συμβούν και στη βιομηχανία για τεχνικούς λόγους, για καλυτέρευση των μεθόδων παραγωγής, για μεταβολές στο τελωνειακό καθεστώς, για μακροχρόνιες απεργιακές κινητοποιήσεις, για πολεμικές συρράξεις κ.τ.λ.

Οι κυμάνσεις αυτές που είναι ακανόνιστες και δύσκολα μπορούν να μελετηθούν, τις ορίζουμε ως υπολειμματικές μεταβολές (variations residuelles), δηλαδή είναι εκείνες οι μεταβολές οι οποίες εμφανίζονται μετά την εξάλειψη όλων των άλλων μεταβολών.

3 ΜΕΘΟΔΟΙ ΠΡΟΣΔΙΟΡΙΣΜΟΥ ΤΗΣ ΜΑΚΡΟΧΡΟΝΙΑΣ ΤΑΣΗΣ

Οι κυριότερες μέθοδοι που χρησιμοποιούμε για τον προσδιορισμό της τάσης (trend) είναι οι παρακάτω:

Α. Πρακτικοί τρόποι

- Χάραξη της τάσης με το χέρι.
- Μέθοδος των δύο μέσων σημείων.
- Μέθοδος των κινητών μέσων.

Β. Μαθηματική μέθοδος προσδιορισμού της τάσης.

3.1 ΧΑΡΑΞΗ ΤΗΣ ΤΑΣΗΣ ΜΕ ΤΟ ΕΛΕΥΘΕΡΟ ΧΕΡΙ

Για να χαραχθεί η τάση με το ελεύθερο χέρι θα πρέπει να κατασκευάσουμε πρώτα την τεθλασμένη γραμμή των εμπειρικών στοιχείων και μετά χαράζουμε με το χέρι μια γραμμή που να περνάει ανάμεσα από την τεθλασμένη γραμμή των εμπειρικών στοιχείων, κατά τέτοιο τρόπο που το άθροισμα των εμβαδών που βρίσκονται κάτω από την τάση να είναι περίπου ίσο με το άθροισμα των εμβαδών που βρίσκονται άνω από την τάση.

3.2 ΜΕΘΟΔΟΣ ΤΩΝ ΔΥΟ ΜΕΣΩΝ ΣΗΜΕΙΩΝ

Για την εφαρμογή της μεθόδου αυτής χωρίζουμε τις παρατηρήσεις, οι οποίες αποτελούν τη χρονολογική σειρά, σε δύο ίσες ομάδες και υπολογίζουμε τους αριθμητικούς μέσους των δύο ομάδων. Μετά χαράζουμε την ευθεία η οποία περνάει από τα σημεία των αριθμητικών μέσων.

Παράδειγμα 3.1

Στον Πίνακα 3.1 δίνεται η παραγωγή ενός εργοστασίου, σε χιλιάδες τόνους, κατά τη διάρκεια των ετών 2004 μέχρι το 2009.

Πίνακας 3.1

Έτη	Παραγωγή (Yi)
2004	5
2005	8
2006	12
2007	15
2008	20
2009	23

Ζητείται να προσδιοριστεί με τη μέθοδο των δύο μέσων σημείων η ευθεία της τάσης.

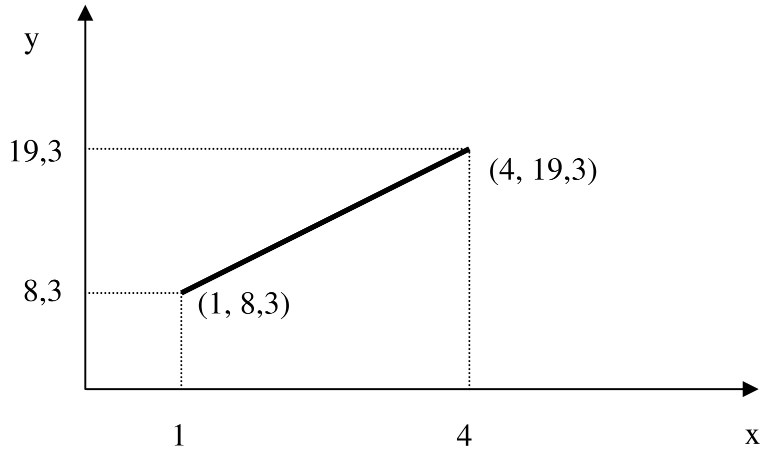
Λύση:

Διαιρούμε τη χρονολογική σειρά σε δύο ομάδες και προσδιορίζουμε τους μέσους αριθμητικούς όπως φαίνεται στον Πίνακα 3.2

Πίνακας 3.2

Έτη	Xi	Yi
2004	0	5
2005	1} 3/3=1	8} 25/3=8,3
2006	2	12
2007	3	15
2008	4} 12/3=4	20} 58/3=19,3
2009	5	23

Η τάση που ζητούμε είναι αυτή που περνάει από τα σημεία των συντεταγμένων m_1 (1, 8,3) και m_2 (4, 19,3).



Διάγραμμα 3.1

3.3 ΜΕΘΟΔΟΣ ΤΩΝ ΚΙΝΗΤΩΝ ΜΕΣΩΝ

Όταν οι παρατηρήσεις μιας χρονολογικής σειράς παρουσιάζουν κυκλικές μεταβολές ρυθμικά επαναλαμβανόμενες, η εκτίμηση της τάσης είναι δυνατή με τη μέθοδο των κινητών μέσων. Με τη μέθοδο αυτή προκύπτει μια νέα σειρά παρατηρήσεων, απαλλαγμένη από την κυκλική συνιστώσα, και ταυτόχρονα γίνεται και εξομάλυνση των ακανόνιστων μεταβολών. Ο αριθμός των όρων που χρησιμοποιούνται στον κινητό μέσο καθορίζει το βαθμό της ομαλότητας. Συνήθως, όσο περισσότεροι όροι χρησιμοποιούνται για τον υπολογισμό του κινητού μέσου, τόσο καλύτερη ομαλότητα προκύπτει.

Διακρίνουμε δύο περιπτώσεις:

α) Όταν ο αριθμός των όρων (**n**) είναι μονός αριθμός π.χ. n = 3, n = 5, θα έχουμε τον Πίνακα 3.3.

Πίνακας 3.3

Χρόνος	**Τιμές**	**Κινητοί μέσοι 3 όρων**	**Κινητοί μέσοι 5 όρων**
1	α_1		
2	α_2	$\alpha'_2 = \dfrac{\alpha_1 + \alpha_2 + \alpha_3}{3}$	
3	α_3	$\alpha'_3 = \dfrac{\alpha_2 + \alpha_3 + \alpha_4}{3}$	$\alpha'_3 = \dfrac{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5}{5}$
4	α_4	$\alpha'_3 = \dfrac{\alpha_2 + \alpha_3 + \alpha_4}{3}$	$\alpha'_4 = \dfrac{\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6}{5}$
5	α_5	…	…
6	α_6	…	…

Οι κινητοί αυτοί μέσοι ονομάζονται εξομαλυνθείσες τιμές τάξης 1. Μετά τον υπολογισμό των εξομαλυντικών τιμών της τάξης 1, μπορούμε να υπολογίσουμε τις εξομαλυντικές τιμές της 2ης τάξης. Οι τιμές αυτές προκύπτουν με τον ίδιο τρόπο όπως οι τιμές της τάξης 1, με τη διαφορά ότι οι κινητοί μέσοι υπολογίζονται με τη βοήθεια των εξομαλυντικών τιμών τάξης 1 και όχι με τη βοήθεια των ακατέργαστων αρχικών τιμών.

Έτσι υπολογίζοντας τους κινητούς μέσους τριών όρων (τους οποίους παριστάνουμε με α''_1, α''_2, ...) επί των προηγούμενων κινητών μέσων τάξης 1 τριών όρων (α'_1, α'_2, ...), θα έχουμε:

$$\alpha''_3 = \frac{\alpha'_2+\alpha'_3+\alpha'_4}{3} = \frac{\frac{\alpha_1+\alpha_2+\alpha_3}{3}+\frac{\alpha_2+\alpha_3+\alpha_4}{3}+\frac{\alpha_3+\alpha_4+\alpha_5}{3}}{3} =$$

$$= \frac{\alpha_1+2\alpha_2+3\alpha_3+2\alpha_4+\alpha_5}{9}$$

Με τον ίδιο τρόπο υπολογίζουμε το α''_4:

$$\alpha''_4 = \frac{\alpha'_3+\alpha'_4+\alpha'_5}{3} = \frac{\alpha_1+2\alpha_3+3\alpha_4+2\alpha_5+\alpha_6}{9}$$

Επίσης, οι κινητοί μέσοι της 2ης τάξης των 5 όρων θα είναι:

$$\alpha''_5 = \frac{\alpha'_3+\alpha'_4+\alpha'_5+\alpha'_6+\alpha'_7}{5} = \frac{\alpha_1+2\alpha_2+3\alpha_3+4\alpha_4+5\alpha_5+4\alpha_6+3\alpha_7+2\alpha_8+\alpha_9}{25}$$

Κατά τον ίδιο τρόπο εργαζόμαστε και για τους υπόλοιπους όρους.

β) Όταν ο αριθμός των όρων είναι ζυγός, π.χ. n = 2, θα έχουμε τον Πίνακα 3.4.

Πίνακας 3.4

Χρόνος	**Τιμές**	**Κινητοί μέσοι 2 όρων**	**Κινητοί μέσοι 2 όρων «κεντρικοί»**
1	α_1	$m'_1 = \frac{\alpha_1+\alpha_2}{2}$	
2	α_2	$m'_2 = \frac{\alpha_2+\alpha_3}{2}$	$\alpha'_2 = \frac{m'_1+m'_2}{2} = \frac{\alpha_1+2\alpha_2+\alpha_3}{4}$
3	α_3	$m'_3 = \frac{\alpha_3+\alpha_4}{2}$	$\alpha'_3 = \frac{m'_2+m'_3}{2} = \frac{\alpha_2+2\alpha_3+\alpha_4}{4}$
4	α_4	$m'_4 = \frac{\alpha_4+\alpha_5}{2}$	$\alpha'_2 = \frac{m'_1+m'_2}{2} = \frac{\alpha_1+2\alpha_2+\alpha_3}{4}$
5	α_5	…	…
…	…	…	…

4. ΤΕΧΝΙΚΕΣ ΕΞΟΜΑΛΥΝΣΗΣ

Αν μπορέσουμε να αναγνωρίσουμε τις συνιστώσες που υπάρχουν σε μια χρονοσειρά, θα είμαστε σε θέση να κάνουμε καλύτερες προβλέψεις. Δυστυχώς, η παρουσία της τυχαίας μεταβλητότητας συγκαλύπτει τις άλλες συνιστώσες και κάνει δύσκολη την αναγνώρισή τους.

Ένας από τους απλούστερους τρόπους για την μείωση της τυχαίας μεταβλητότητας είναι η **εξομάλυνση** (smoothing). Στην ενότητα αυτή θα περιγράψουμε δυο μεθόδους εξομάλυνσης:

- ✓ τον **κυλιόμενο μέσο** (moving average) και
- ✓ την **εκθετική εξομάλυνση** (exponential smoothing).

4.1 ΚΥΛΙΟΜΕΝΟΣ ΜΕΣΟΣ

Ο κυλιόμενος μέσος σε ένα χρονικό σημείο είναι ο αριθμητικός μέσος ενός χρονικού διαστήματος που έχει ως κέντρο το συγκεκριμένο σημείο. Για παράδειγμα, ο κυλιόμενος μέσος τριών περιόδων είναι ο μέσος της τρέχουσας περιόδου, της προηγούμενης και της επόμενης. Σε μια χρονοσειρά μπορούμε να υπολογίσουμε κυλιόμενο μέσο τριών περιόδων για κάθε περίοδο εκτός από την πρώτη και την τελευταία. Όμοια, ο κυλιόμενος μέσος πέντε περιόδων ο αριθμητικός μέσος της τρέχουσας περιόδου, των δυο προηγούμενων και των δυο επόμενων.

Παράδειγμα 4.1

Προσπαθώντας να προβλέψει τις μελλοντικές πωλήσεις, ο διευθυντής πέντε πρατηρίων καυσίμων έχει καταγράψει τις τριμηνιαίες πωλήσεις (σε χιλιάδες γαλόνια) για τα τελευταία 4 χρόνια, όπως φαίνεται στον παρακάτω πίνακα.

Πίνακας 4.1

Περίοδος	Έτος	Τρίμηνο	Πωλήσεις
1	1	1	39
2		2	37
3		3	61
4		4	58
5	2	1	18
6		2	56
7		3	82
8		4	27
9	3	1	41
10		2	69
11		3	49
12		4	66
13	4	1	54
14		2	42
15		3	90
16		4	66

Θα υπολογίσουμε τον κυλιόμενο μέσο τριών και πέντε περιόδων και θα σχεδιάσουμε το γράφημα της χρονολογικής σειράς και των κυλιόμενων μέσων.

Λύση:

Για τον υπολογισμό του κυλιόμενου μέσου τριών περιόδων προσθέτουμε κάθε τριάδα διαδοχικών τιμών και διαιρούμε δια 3. Έτσι, ο κυλιόμενος μέσος για την περίοδο 2 θα είναι:

$$\frac{39+37+61}{3}=45{,}7$$

Για τον υπολογισμό του κυλιόμενου μέσου πέντε περιόδων προσθέτουμε κάθε πεντάδα διαδοχικών τιμών και διαιρούμε δια 5, αρχίζοντας από την περίοδο 3:

$$\frac{39+37+61+58+18}{5}=42,6$$

Με τον τρόπο αυτό συμπληρώνουμε τον πίνακα των κυλιόμενων μέσων ως εξής:

Πίνακας 4.2

Περίοδος	Έτος	Τρίμηνο	Πωλήσεις	Κυλιόμενος μέσος 3	Κυλιόμενος μέσος 5
1	1	1	39	-	-
2		2	37	45,7	-
3		3	61	52,0	42,6
4		4	58	45,7	46,0
5	2	1	18	44,0	55,0
6		2	56	52,0	48,2
7		3	82	55,0	44,8
8		4	27	50,0	55,0
9	3	1	41	45,7	53,6
10		2	69	53,0	50,4
11		3	49	61,3	55,8
12		4	66	56,3	56,0
13	4	1	54	54,0	60,2
14		2	42	62,0	63,6
15		3	90	66,0	-
16		4	66	-	-

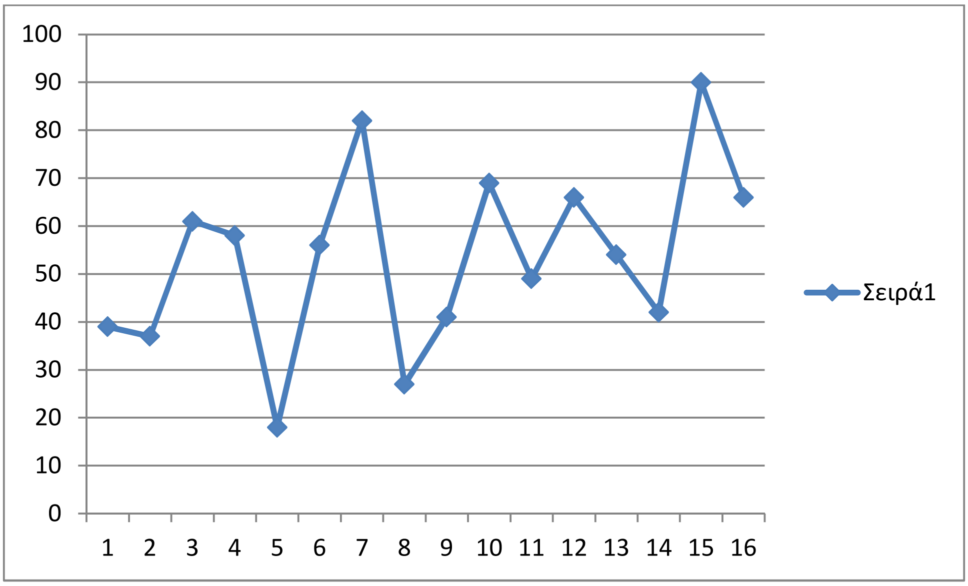
Παρατηρούμε ότι ο κυλιόμενος μέσος τοποθετείται στο κέντρο του διαστήματος των τριών ή πέντε περιόδων. Για το λόγο αυτό επιλέγουμε περιττό αριθμό περιόδων για τον κυλιόμενο μέσο.

Ερμηνεία:

Για να δούμε πως ο κυλιόμενος μέσος απομονώνει ένα μέρος της τυχαίας μεταβλητότητας, αρκεί να συγκρίνουμε τα διαγράμματα των Εικόνων 4.1 και 4.2.

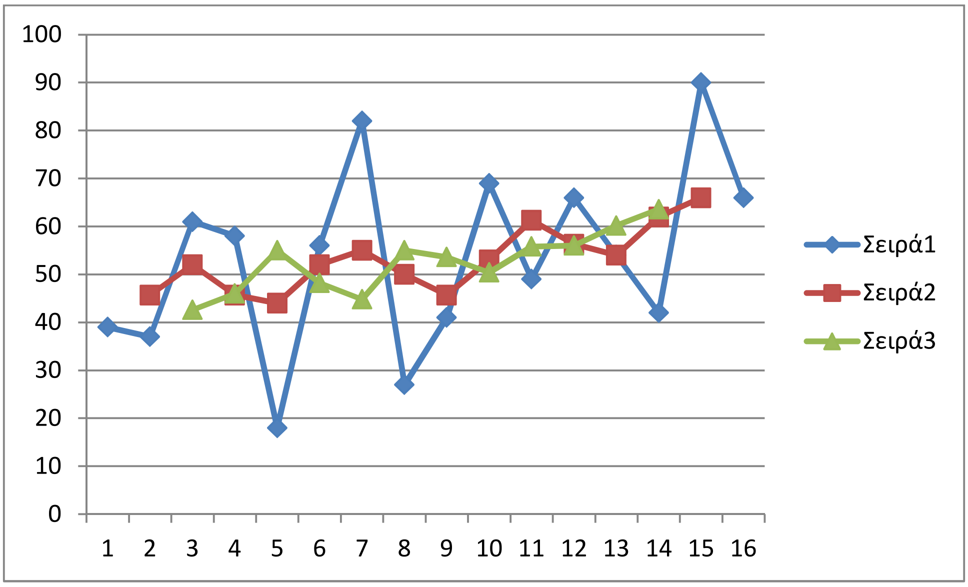
Η Εικόνα 4.1 είναι ένα διάγραμμα της χρονολογικής σειράς των τριμηνιαίων πωλήσεων βενζίνης. Στο σχήμα αυτό υπάρχει αρκετή τυχαία μεταβλητότητα, και δεν είναι εύκολο να αναγνωρίσουμε κάποια σταθερή συνιστώσα.

Εικόνα 4.1 Τριμηνιαίες πωλήσεις βενζίνης



Στην Εικόνα 4.2 μαζί με το διάγραμμα των πωλήσεων υπάρχουν τα διαγράμματα του κυλιόμενου μέσου τριών και πέντε περιόδων, όπου είναι φανερό ένα μέγιστο στο τρίτο τρίμηνο κάθε έτους. Επίσης είναι ορατή μια ελαφρά ανοδική μακροχρόνια τάση.

Εικόνα 4.2 Τριμηνιαίες πωλήσεις και κυλιόμενος μέσος 3 και 5 περιόδων



Παρατηρούμε ότι ο κυλιόμενος μέσος πέντε περιόδων δίνει μια ομαλότερη γραμμή. Δυστυχώς όμως στον κυλιόμενο μέσο των πέντε περιόδων έχει αφαιρεθεί και ένα μέρος της εποχιακής μεταβλητότητας, και έτσι ο ετήσιος κύκλος του ελάχιστου στο πρώτο τρίμηνο και του μέγιστου στο τρίτο τρίμηνο είναι λιγότερο ορατός. Γενικά, όσο μεγαλύτερο είναι το διάστημα ενός κυλιόμενου μέσου τόσο περισσότερη μεταβλητότητα εξουδετερώνεται, αφήνοντας να φανεί μόνο η μακροχρόνια τάση.

4.2 ΕΠΙΚΕΝΤΡΩΜΕΝΟΣ ΚΥΛΙΟΜΕΝΟΣ ΜΕΣΟΣ

Αν ο αριθμός περιόδων είναι άρτιος, δεν υπάρχει στο παράθυρο του κυλιόμενου μέσου μια κεντρική περίοδος στην οποία θα αντιστοιχεί ο κυλιόμενος μέσος.

Για παράδειγμα, ας υποθέσουμε ότι πρέπει να υπολογίσουμε τον κυλιόμενο μέσο 4 περιόδων για τον εξής πίνακα:

Πίνακας 4.3

Περίοδος	Τιμή
1	15
2	27
3	20
4	14
5	25
6	11

Ο πρώτος κυλιόμενος μέσος υπολογίζεται:

$$\frac{15+27+20+14}{4}=19,0$$

Σε μια χρονοσειρά έξι περιόδων υπάρχουν τρεις διαδοχικές τετράδες τιμών, έτσι μπορούν να υπολογιστούν δυο ακόμη κυλιόμενοι μέσοι τεσσάρων περιόδων, ως εξής:

$$\frac{27+20+14+25}{4}=21,5$$

$$\frac{20+14+25+11}{4}=17,5$$

Ο πρώτος κυλιόμενος μέσος αντιπροσωπεύει τις περιόδους 1-4, άρα θα πρέπει να τοποθετηθεί μεταξύ των περιόδων 2 και 3. Αυτό δημιουργεί προβλήματα, τόσο για τη γραφική παράσταση όσο και για την αντιστοίχιση των κυλιόμενων μέσων στις περιόδους.Το πρόβλημα λύνεται υπολογίζοντας τον κυλιόμενο μέσο δεύτερης τάξης δυο περιόδων, που ονομάζεται επικεντρωμένος (centered) κυλιόμενος μέσος:

$$\frac{19.0+21.5}{2}=20,25$$

Ο κυλιόμενος μέσος 19,0 βρίσκεται μεταξύ των περιόδων 2 και 3 ενώ ο κυλιόμενος μέσος 21,5 βρίσκεται μεταξύ των περιόδων 3 και 4, συνεπώς ο επικεντρωμένος κυλιόμενος μέσος αντιστοιχεί στην περίοδο 3. Η διαδικασία αυτή συνοψίζεται στον παρακάτω πίνακα:

Πίνακας 4.4

Περίοδος	Τιμή	Κυλιόμενος μέσος 4	Επικεντρωμένος
1	15	-	-
2	27	19	-
3	20	21,5	20,25
4	14	17,5	19,5
5	25	-	-
6	11		-

4.3 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ

Η μέθοδος του κυλιόμενου μέσου έχει δυο μειονεκτήματα. Πρώτον, δεν μπορούν να υπάρχουν κυλιόμενοι μέσοι για τις πρώτες και τελευταίες περιόδους, ανάλογα με το μέγεθος του παραθύρου. Έτσι, αν η χρονοσειρά έχει σχετικά μικρό αριθμό παρατηρήσεων, θα υπάρχουν στον πίνακα πολλά κενά. Δεύτερον, ο κυλιόμενος μέσος έχει τοπικό χαρακτήρα και δεν λαμβάνει υπόψη τιμές που βρίσκονται έξω από το παράθυρό του. Για παράδειγμα, ο κυλιόμενος μέσος 5 περιόδων για την περίοδο 4 αγνοεί την τιμή της περιόδου 1.

Τα προβλήματα αυτά λύνονται με τη χρήση της εκθετικής εξομάλυνσης.

Εκθετική εξομάλυνση χρονοσειράς

$$S_t = wy_t + (1-w)\mathrm{S}_{\mathrm{t-1}}$$

για t≥2, όπου:

t Αριθμός της περιόδου

y_t Τιμή χρονοσειράς για την περίοδο t

S_t Τιμή εξομάλυνσης για την περίοδο t

S_{t-1} Τιμή εξομάλυνσης για την περίοδο $t-1$

w Συντελεστής απόσβεσης, $0 \le w \le 1$

Η τιμή εξομάλυνσης για την πρώτη περίοδο ορίζεται ως:

$$S_1 = y_1$$

Στη συνέχεια χρησιμοποιείται επανειλημμένα ο παραπάνω τύπος για τον υπολογισμό των τιμών εξομάλυνσης:

$S_2 = wy_2 + (1 - w)S_1 = wy_2 + (1 - w)y_1$

$S_3 = wy_3 + (1 - w)S_2 = wy_3 + w(1 - w)y_2 + (1 - w)^2 y_1$

Γενικά για κάθε τιμή της χρονοσειράς έχουμε:

$$S_t = wy_t + w(1 - w)y_{t-1} + \cdots + (1 - w)^{t-1} y_1$$

Όπως φαίνεται στον τύπο, κάθε τιμή εξομάλυνσης λαμβάνει υπόψη όλες τις προηγούμενες τιμές της χρονοσειράς.

Ο συντελεστής απόσβεσης (damping factor) w καθορίζει το βαθμό της εξομάλυνσης. Αν η σταθερά w είναι κοντά στη μονάδα η εξομάλυνση είναι πολύ μικρή, ενώ αν είναι κοντά στο μηδέν η εξομάλυνση είναι πολύ μεγάλη.

5 ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ ΜΕ ΤΗ ΧΡΗΣΗ EXCEL

5.1 ΠΡΟΣΑΡΜΟΓΗ ΓΡΑΜΜΩΝ ΤΑΣΕΩΣ ΣΕ ΧΡΟΝΟΣΕΙΡΕΣ

Οι γραμμές τάσεως χρησιμοποιούνται για την απόδοση γραφικής αναπαράστασης των τάσεων. Το Excel παρέχει αυτόματα τη δυνατότητα εισαγωγής γραμμής τάσης στα διαγράμματα, μέσω της επιλογής TrendLine.

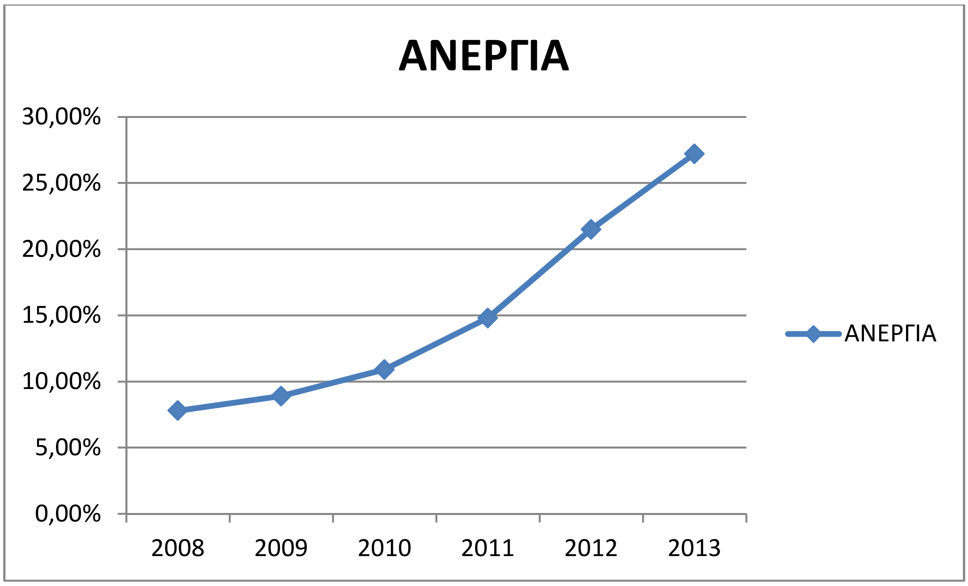
Μπορούμε να επιλέξουμε έξι διαφορετικούς τύπους γραμμών τάσεως:

- ✓ γραμμικές γραμμές τάσης
- ✓ πολυωνυμικές γραμμές τάσης
- ✓ γραμμές τάσης αύξησης
- ✓ εκθετικές γραμμές τάσης
- ✓ γραμμές τάσης κυλιόμενου μέσου.

Παράδειγμα 5.1

Στον παρακάτω διάγραμμα παρουσιάζεται η Πορεία της ανεργίας στην Ελλάδα κατά το μήνα Ιανουάριο 2008-2013.

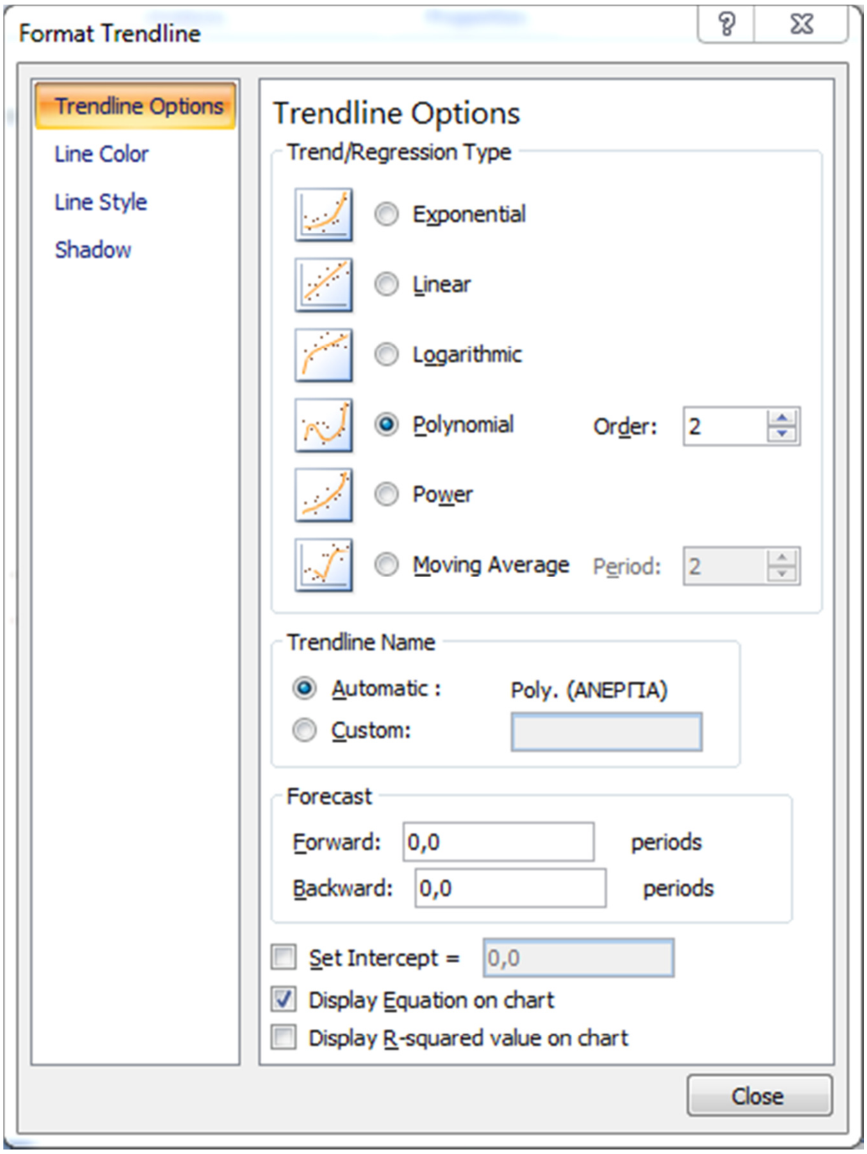
Διάγραμμα 5.1



Πηγή: ΕΛ.ΣΤΑΤ.

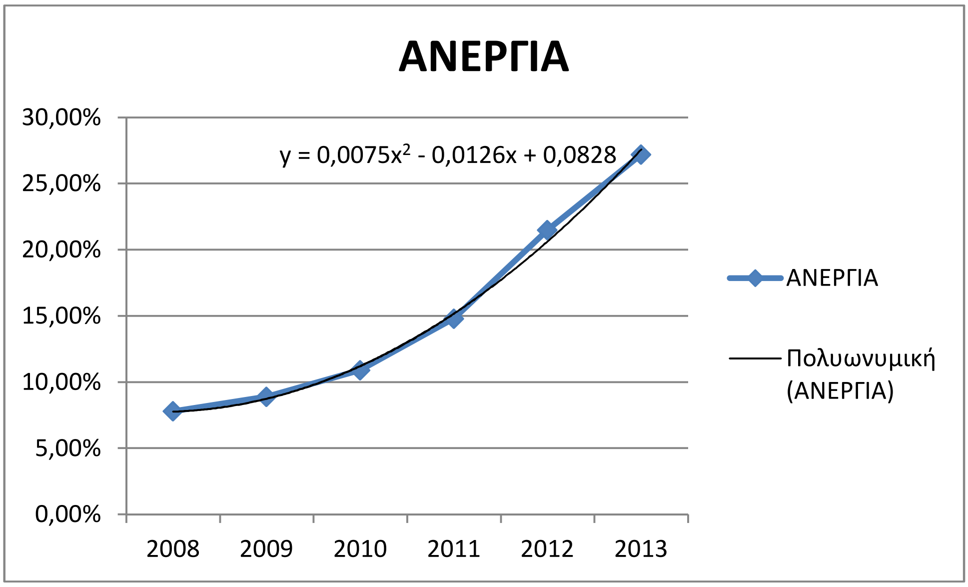
Από το μενού Chart Tools επιλέγεται η εντολή Trendline.

Εικόνα 5.1 Πλαίσιο διαλόγου Format Tredline



Το ακόλουθο διάγραμμα δείχνει μια πολυωνυμική γραμμή τάσης τάξης 2.

Διάγραμμα 5.2



5.2 ΥΠΟΛΟΓΙΣΜΟΣ ΚΥΛΙΟΜΕΝΟΥ ΜΕΣΟΥ

Το Excel παρέχει αυτόματα τη δυνατότητα εισαγωγής της χρονοσειράς κυλιόμενου μέσου στα διαγράμματα, μέσω της επιλογής TrendLine.

Παράδειγμα 5.2

Στον παρακάτω πίνακα παρουσιάζονται οι εβδομαδιαίες πωλήσεις (σε τεμάχια) ενός προϊόντος τις τελευταίες 60 εβδομάδες.

Πίνακας 5.1

Εβδομάδα	Πωλήσεις	Εβδομάδα	Πωλήσεις
1	98	31	56
2	90	32	82
3	93	33	69
4	87	34	74
5	80	35	36
6	76	36	48
7	79	37	52
8	92	38	43
9	66	39	49
10	70	40	22
11	50	41	26
12	20	42	19
13	40	43	27
14	54	44	29
15	68	45	31
16	60	46	42
17	65	47	48
18	60	48	43
19	92	49	57
20	78	50	62

21	65	51	71
22	43	52	65
23	52	53	73
24	47	54	35
25	30	55	46
26	67	56	68
27	20	57	61
28	23	58	29
29	28	59	33
30	34	60	39

Διάγραμμα 5.3

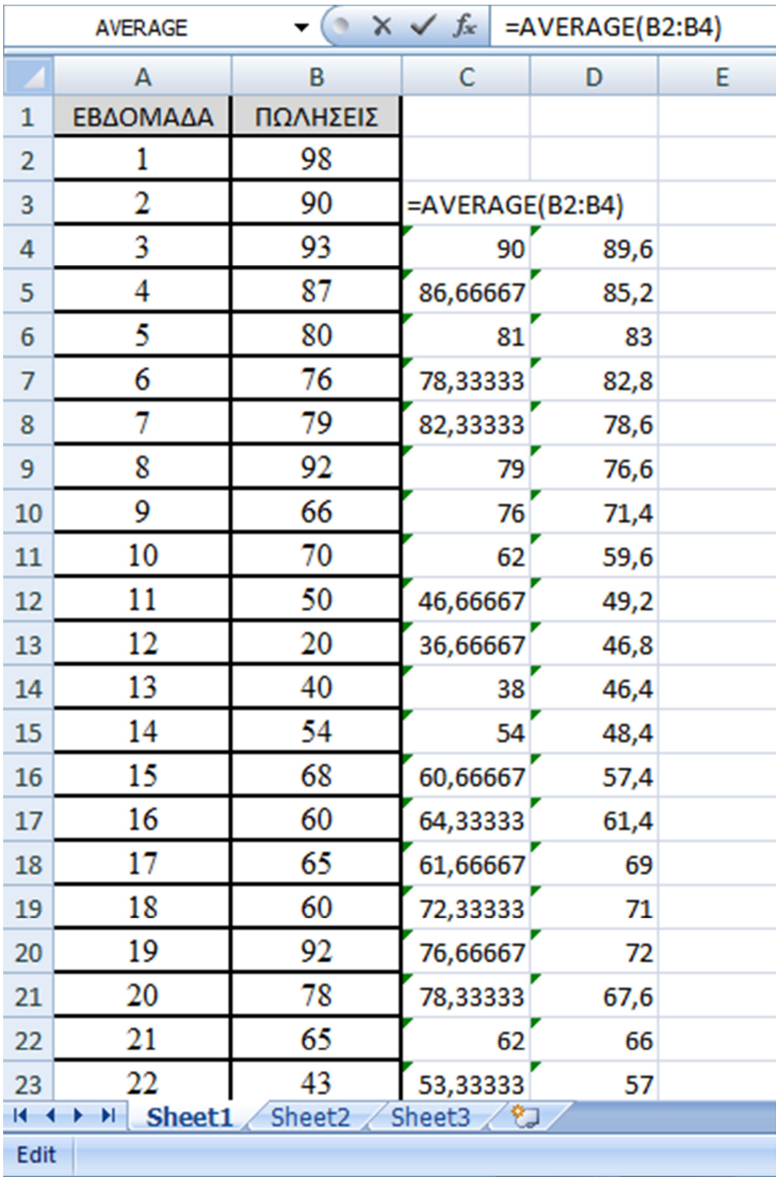
Διάγραμμα 5.4

Για τον υπολογισμό του κυλιόμενου μέσου τριών περιόδων, στο κελί C3 εισάγεται ο τύπος:

=AVERAGE(B2:B4)

και αντιγράφεται προς τα κάτω μέχρι το κελί C60.

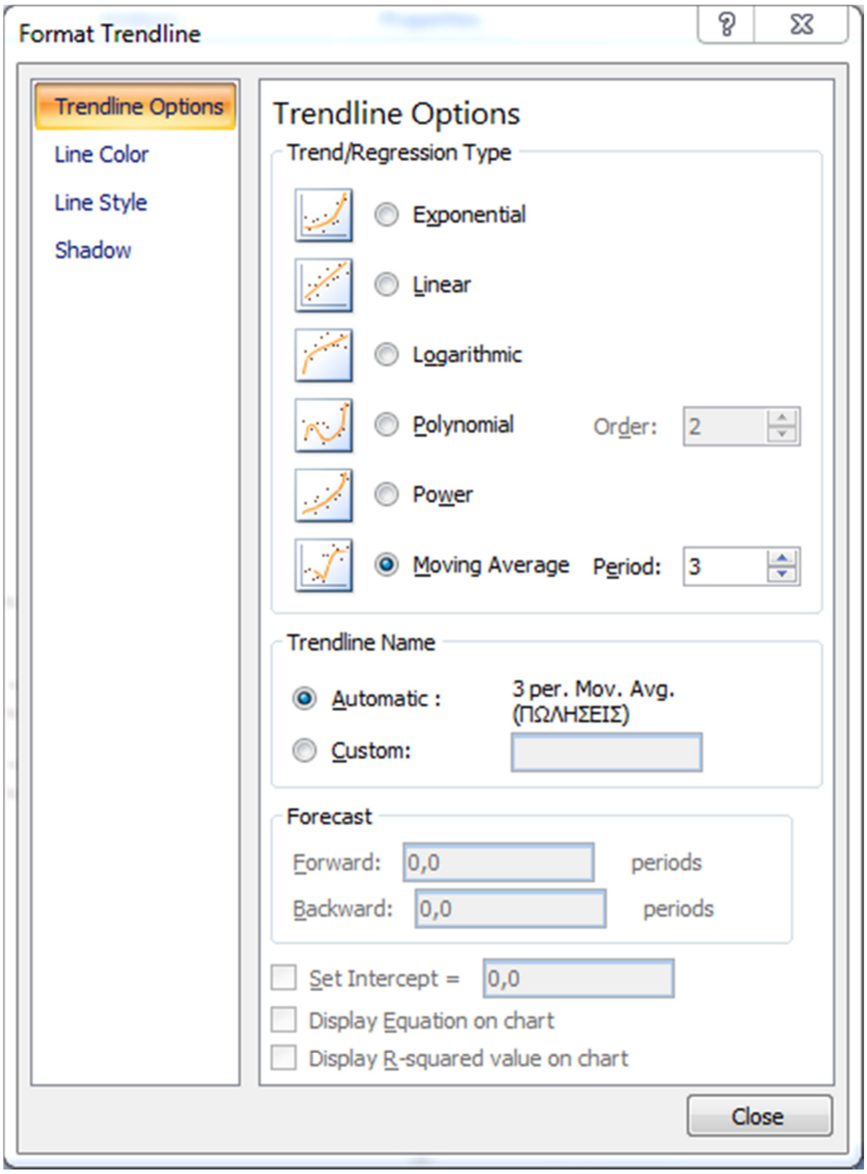
Εικόνα 5.2



	A	B	C	D	E
1	ΕΒΔΟΜΑΔΑ	ΠΩΛΗΣΕΙΣ			
2	1	98			
3	2	90	=AVERAGE(B2:B4)		
4	3	93	90	89,6	
5	4	87	86,66667	85,2	
6	5	80	81	83	
7	6	76	78,33333	82,8	
8	7	79	82,33333	78,6	
9	8	92	79	76,6	
10	9	66	76	71,4	
11	10	70	62	59,6	
12	11	50	46,66667	49,2	
13	12	20	36,66667	46,8	
14	13	40	38	46,4	
15	14	54	54	48,4	
16	15	68	60,66667	57,4	
17	16	60	64,33333	61,4	
18	17	65	61,66667	69	
19	18	60	72,33333	71	
20	19	92	76,66667	72	
21	20	78	78,33333	67,6	
22	21	65	62	66	
23	22	43	53,33333	57	

Από το μενού Chart Tools επιλέγεται η εντολή Trendline.

Εικόνα 5.3 Πλαίσιο διαλόγου Format Tredline



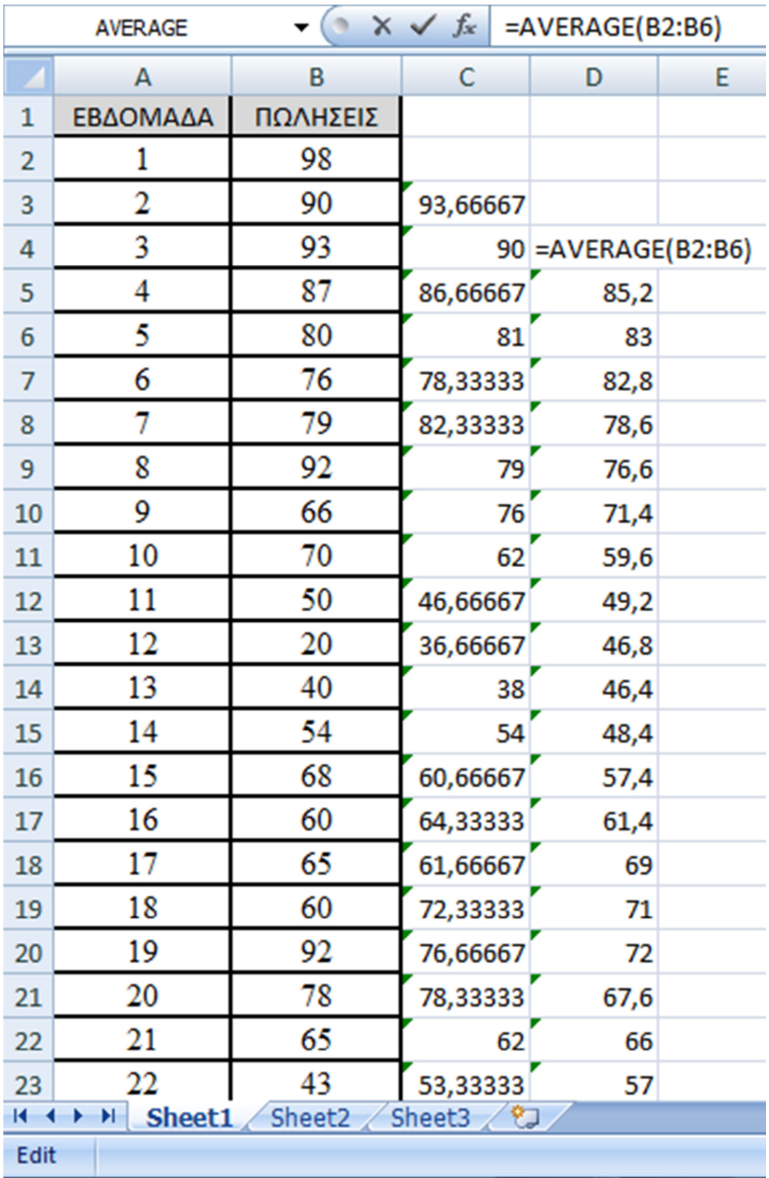
Διάγραμμα 5.5 Κυλιόμενος μέσος τριών περιόδων

Για τον υπολογισμό του κυλιόμενου μέσου πέντε περιόδων, στο κελί D4 εισάγεται ο τύπος:

=AVERAGE(B2:B6)

και αντιγράφεται προς τα κάτω μέχρι το κελί D59.

Εικόνα 5.4



AVERAGE | =AVERAGE(B2:B6)

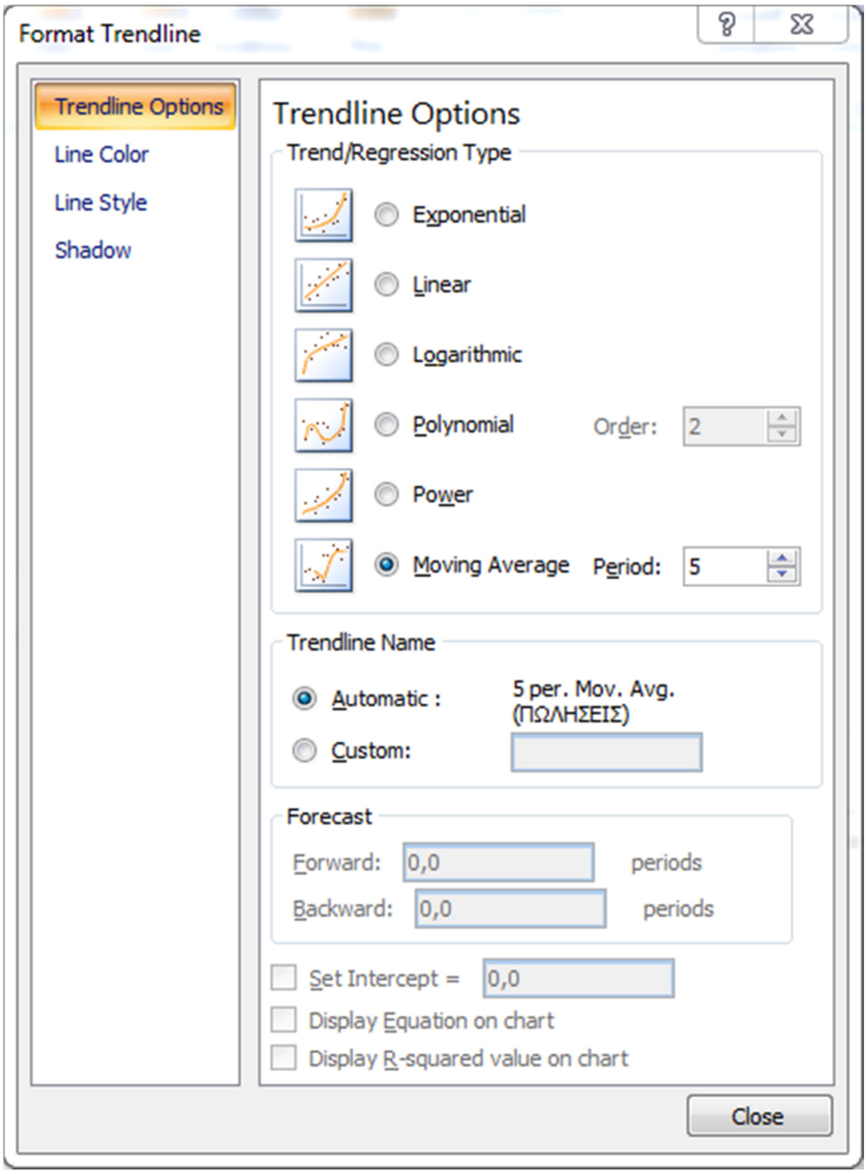
	A	B	C	D	E
1	ΕΒΔΟΜΑΔΑ	ΠΩΛΗΣΕΙΣ			
2	1	98			
3	2	90	93,66667		
4	3	93	90	=AVERAGE(B2:B6)	
5	4	87	86,66667	85,2	
6	5	80	81	83	
7	6	76	78,33333	82,8	
8	7	79	82,33333	78,6	
9	8	92	79	76,6	
10	9	66	76	71,4	
11	10	70	62	59,6	
12	11	50	46,66667	49,2	
13	12	20	36,66667	46,8	
14	13	40	38	46,4	
15	14	54	54	48,4	
16	15	68	60,66667	57,4	
17	16	60	64,33333	61,4	
18	17	65	61,66667	69	
19	18	60	72,33333	71	
20	19	92	76,66667	72	
21	20	78	78,33333	67,6	
22	21	65	62	66	
23	22	43	53,33333	57	

Sheet1 | Sheet2 | Sheet3

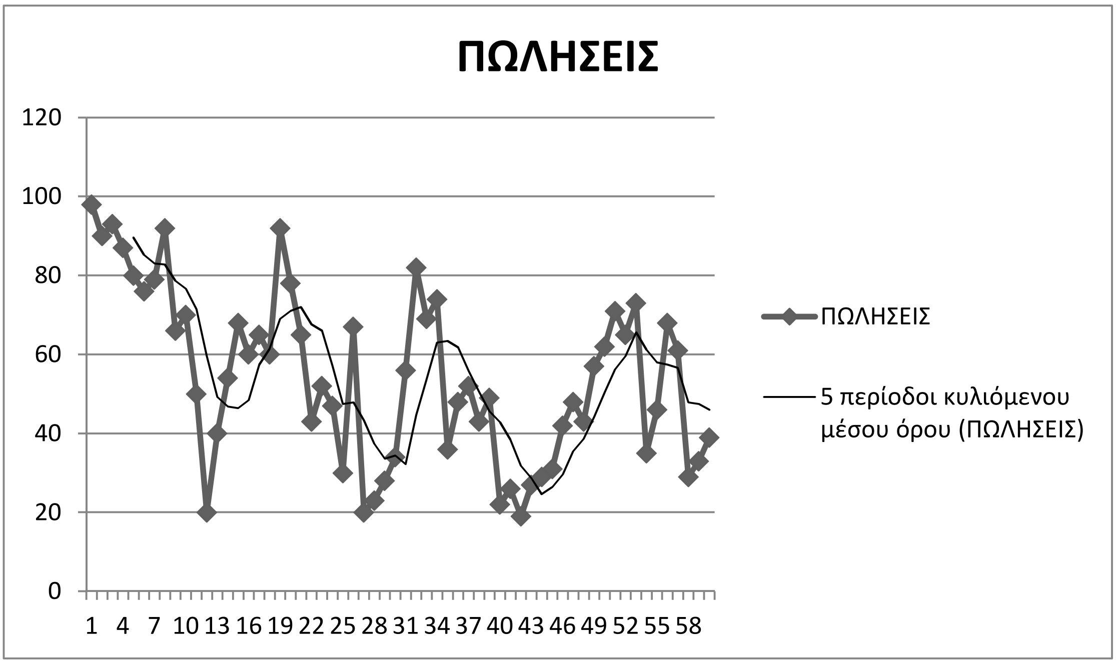
Edit

Από το μενού Chart Tools επιλέγεται η εντολή Trendline.

Εικόνα 5.5 Πλαίσιο διαλόγου Format Tredline



Διάγραμμα 5.6 Κυλιόμενος μέσος πέντε περιόδων



ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα διπλωματική εργασία επικεντρώνεται στην «ανάλυση και πρόβλεψη χρονοσειρών οικονομικών δεδομένων».

Η αβεβαιότητα που πολλές φορές χαρακτηρίζει τη ζήτηση προϊόντων ή υπηρεσιών και, συνεπώς, τις απαιτήσεις σε μηχανές, υλικά, κεφάλαια, ανθρώπινο δυναμικό και, γενικά, δυναμικότητα που θα χρησιμοποιηθεί ώστε να ικανοποιηθεί η ζήτηση κατέστησε αναγκαία την ανάπτυξη μεθόδων πρόβλεψης. Ο προγραμματισμός και ο έλεγχος της παραγωγής, ειδικότερα, απαιτούν εκτιμήσεις όσον αφορά την ποσότητα και το χρόνο που απαιτούν εκτιμήσεις όσον αφορά την ποσότητα και το χρόνο που αναμένεται να ζητηθεί το προϊόν ενός παραγωγικού συστήματος. Οι εκτιμήσεις αυτές θα χρησιμοποιηθούν για την κατάρτιση των προγραμμάτων παραγωγής, προμήθειας πρώτων υλών, απασχόλησης ανθρώπινου δυναμικού κ.λπ.

Τα προγράμματα αυτά θα είναι τόσο περισσότερο αποτελεσματικά, σε σχέση με το σκοπό του παραγωγικού συστήματος, όσο περισσότερο αξιόπιστες είναι οι σχετικές προβλέψεις. Οι κριτικές προβλέψεις είναι πράγματι απαραίτητες καθώς αποτελούν την μοναδική εναλλακτική για την πρόβλεψη συστηματικών αλλαγών από καθιερωμένα πρότυπα και υπάρχουσες σχέσεις. Την ίδια στιγμή, θα πρέπει να είμαστε προσεκτικοί ώστε να αποφεύγουμε τις μεροληψίες και άλλους περιορισμούς που χαρακτηρίζουν την κρίση μας για να μειώνουμε τις αρνητικές τους συνέπειες στις προβλέψεις. Η πρόκληση για τις εταιρείες είναι να εκμεταλλεύονται και τις στατιστικές προβλέψεις αλλά και την μοναδική ικανότητα της ανθρώπινης κρίσης να αντιμετωπίζει συστηματικές αλλαγές σε πρότυπα / σχέσεις.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Λ. Αδαμόπουλος, Χ. Δαμιανού, Α. Σβέρκος, *Μαθηματικά και Στοιχεία Στατιστικής*, ΟΕΔΒ, Αθήνα (1999).

Διαθέσιμο στο δικτυακό τόπο:

http://digitalschool.minedu.gov.gr/modules/ebook/show.php/DSGL-C100/493/3201,13005/

[2] Χ. Γναρδέλλης, *Εφαρμοσμένη Στατιστική*, Εκδόσεις Παπαζήση, Αθήνα (2003).

[3] Π. Κιόχος & Α. Κιόχος, *Στατιστική για τις Επιχειρήσεις και την Οικονομία*, Αυτοέκδοση, Αθήνα, (2010).

[4] Β. Μπένος, *Στατιστική (Περιγραφική Σταστική)*, Εκδόσεις Σταμούλης, Αθήνα (1997).

[5] Κ. Τραχάνας & Α. Τσέβας, *Περιγραφική Στατιστική*, Εκδόσεις Σταμούλης, Αθήνα (1998).

[6] Κ. Τσίμπος & Φ. Γεωργιακώδης, *Περιγραφική και Διερευνητική Στατιστική. Ανάλυση Δεδομένων, τόμος Α΄*, Εκδόσεις Σταμούλης, Αθήνα (2010).

[7] Γ. Χρήστου, *Εισαγωγή στην Οικονομετρία*, Εκδόσεις Gutenberg, Αθήνα (2011).

[8] G. Keller, *Στατιστική για οικονομικά και διοίκηση επιχειρήσεων,* Επίκεντρο, Θεσσαλονίκη (2010).

Πνευματικά δικαιώματα

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1988 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον.

ΜΠΡΙΝΙΑΣ ΒΑΣΙΛΗΣ, [2015]