

ΣΤΑΤΙΣΤΙΚΗ

ΕΝΟΤΗΤΑ 1^η

ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΒΑΣΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΕΝΝΟΙΕΣ

Η ΣΤΑΤΙΣΤΙΚΗ

Η Στατιστική είναι η επιστήμη, η οποία ασχολείται:

- με τη συλλογή,
- με την παρουσίαση
- με την επεξεργασία

αριθμητικών ή / και μη αριθμητικών στοιχείων με απώτερο σκοπό να εξάγει χρήσιμα συμπεράσματα για τους ευρύτερους πληθυσμούς, στους οποίους ανήκουν τα στοιχεία αυτά.

ΤΟΜΕΙΣ ΣΤΑΤΙΣΤΙΚΗΣ

Η Στατιστική χωρίζεται σε δύο μεγάλους τομείς:

- την **Περιγραφική Στατιστική**
- την **Επαγωγική Στατιστική**

- **Στόχος της Περιγραφικής Στατιστικής:** Η περιγραφή είτε του υπό μελέτη πληθυσμού στόχου στο σύνολό του είτε του διαθέσιμου και κατάλληλα επιλεγμένου δείγματος.

- **Στόχος της Επαγωγικής Στατιστικής** είναι, η κατάλληλη χρήση του δείγματος προκειμένου να εξαγάγουμε συμπεράσματα για τον γεννήτορα πληθυσμό.

οι έννοιες Πληθυσμός και Δείγμα

- Με τον όρο **Πληθυσμό** σε μια έρευνα, εννοούμε ένα σύνολο που μπορεί να αποτελείται από ανθρώπους, ζώα ή αντικείμενα, τα χαρακτηριστικά των οποίων θέλουμε να μελετήσουμε.
- Με τον όρο **Δείγμα** σε μια έρευνα, καλούμε μια συλλογή από στοιχειώδεις ή πρωταρχικές μονάδες δειγματοληψίας, τις οποίες επιλέγουμε έτσι ώστε να αποτελούν μια αντιπροσωπευτική εικόνα του πληθυσμού.

Μεταβλητή - Παράμετροι

- Ένα κατάλληλα επιλεγμένο **δείγμα** μπορεί να μελετηθεί ως προς ένα ή περισσότερα χαρακτηριστικά του.
- Τα χαρακτηριστικά τα οποία μας ενδιαφέρουν τα ονομάζουμε **μεταβλητές**.
- Οι **μεταβλητές** διακρίνονται σε **κατηγορικές, ποιοτικές και ποσοτικές (ασυνεχείς ή συνεχείς)**, ανάλογα με τις τιμές που μπορούν να πάρουν και το είδος της μέτρησης που επιδέχονται.
- Η κατάταξη των **μεταβλητών** έχει τη βάση της στην ισχύουσα ταξινόμηση των **κλιμάκων μέτρησης των μεταβλητών**. Η διάκριση μεταξύ των διαφόρων κλιμάκων μέτρησης υπαγορεύεται από την ύπαρξη ή μη των μαθηματικών ιδιοτήτων της διάταξης, της απόστασης, της ύπαρξης του μηδενός και της αναλογικότητας.

ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Περιγραφική Στατιστική

- **Στόχος της Περιγραφικής Στατιστικής**

Η περιγραφή είτε του υπό μελέτη πληθυσμού στόχου στο σύνολό του είτε του διαθέσιμου και κατάλληλα επιλεγμένου δείγματος.

- **Εργαλεία της Περιγραφικής Στατιστικής**

- Πίνακες Συχνοτήτων
- Περιγραφικά Μέτρα
- Διαγράμματα

Πίνακες Συχνοτήτων

- Τα δεδομένα μιας ανάλυσης κατά τη συλλογή τους έχουν το χαρακτήρα των απλών δεδομένων
 - Όταν οι τιμές της ποσοτικής μεταβλητής X , οι οποίες αναφέρονται στις N επί μέρους μονάδες του πληθυσμού ή του δείγματος που μελετάται, εξετάζονται μεμονωμένα, τότε τα διαθέσιμα δεδομένα καλούνται απλά.
- Ο πίνακας συχνοτήτων απλής ή διπλής εισόδου δημιουργείται όταν οι τιμές της μεταβλητής X , οι οποίες αναφέρονται στις N επί μέρους μονάδες του πληθυσμού ή του δείγματος που μελετάται ομαδοποιούνται και τα δεδομένα παρουσιάζονται υπό μορφή μιας συνεχούς ή ασυνεχούς κατανομής συχνότητας.

Περιγραφικά Μέτρα

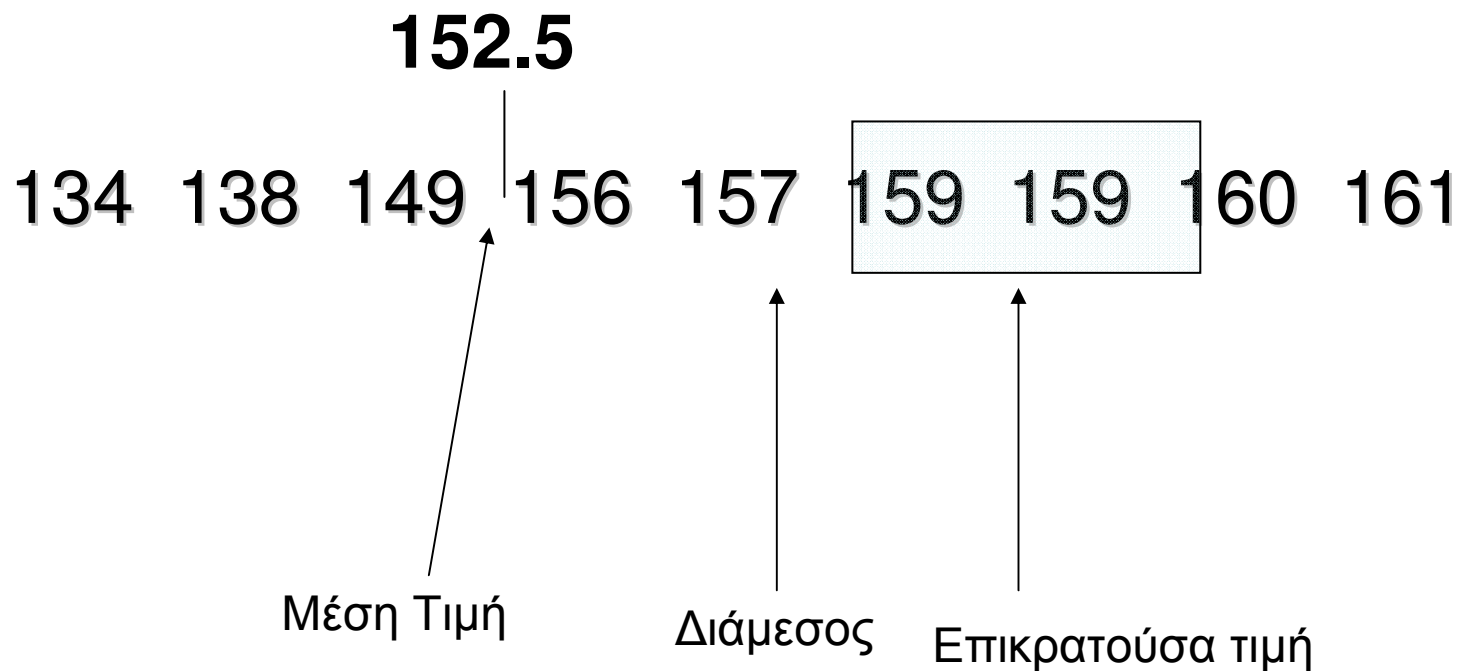
- Τα περιγραφικά στατιστικά μέτρα είναι κατάλληλοι δείκτες οι οποίοι συνοψίζουν τις πληροφορίες που εμπεριέχονται στα δεδομένα μας.
- Έχουμε τρεις κατηγορίες περιγραφικών μέτρων:
 - τα μέτρα θέσης (μέση τιμή, διάμεσος, κορυφή),
 - τα μέτρα διασποράς (εύρος, διακύμανση, τυπ. απόκλιση)
 - τα μέτρα μορφής (συντ. ασυμμετρίας και κύρτωσης)
- Επίσης, ανάλογα με το αν οι τιμές μας προέρχονται από έναν πληθυσμό ή ένα δείγμα, μιλάμε για πληθυσμιακά μέτρα ή δειγματικά μέτρα, αντίστοιχα.

Μέτρα θέσεως

- Τα κυριότερα μέτρα θέσεως είναι η μέση τιμή, η διάμεσος και η κορυφή.
 - Η μέση τιμή για ποσοτικές μεταβλητές υπολογίζει τον μέσο όρο των δεδομένων.
 - Η διάμεσος είναι η τιμή των διατεταγμένων δεδομένων που ισαπέχει από τη μέγιστη και ελάχιστη τιμή.
 - Η κορυφή είναι η τιμή με τη μέγιστη συχνότητα εμφάνισης.

Σύγκριση των Μέτρων θέσεως

- Έστω οι ακόλουθες τιμές HDL 9 ατόμων:



Μέτρα διασποράς

- Τα κυριότερα μέτρα διασποράς είναι το εύρος, η διακύμανση και η τυπική απόκλιση.
 - Το εύρος ορίζεται ως η διαφορά της μέγιστης από την ελάχιστη παρατήρηση
 - Η διακύμανση είναι το άθροισμα των τετραγώνων των διαφορών κάθε παρατήρησης από τη μέση τιμή
 - Η τυπική απόκλιση είναι η τετραγωνική ρίζα της διακύμανσης

Σύγκριση των Μέτρων Διασποράς

- Έστω οι ακόλουθες τιμές HDL 9 ατόμων:

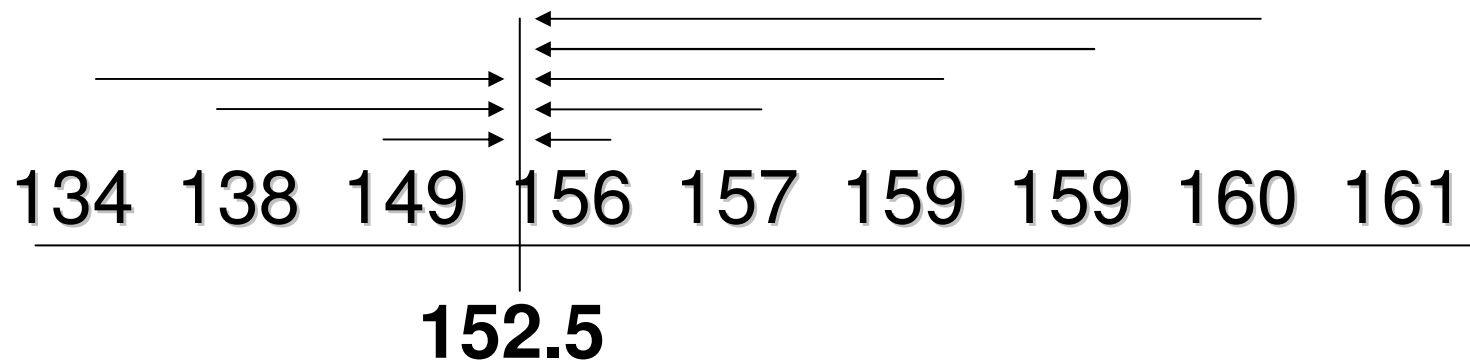


Μέτρα μορφής της κατανομής των δεδομένων

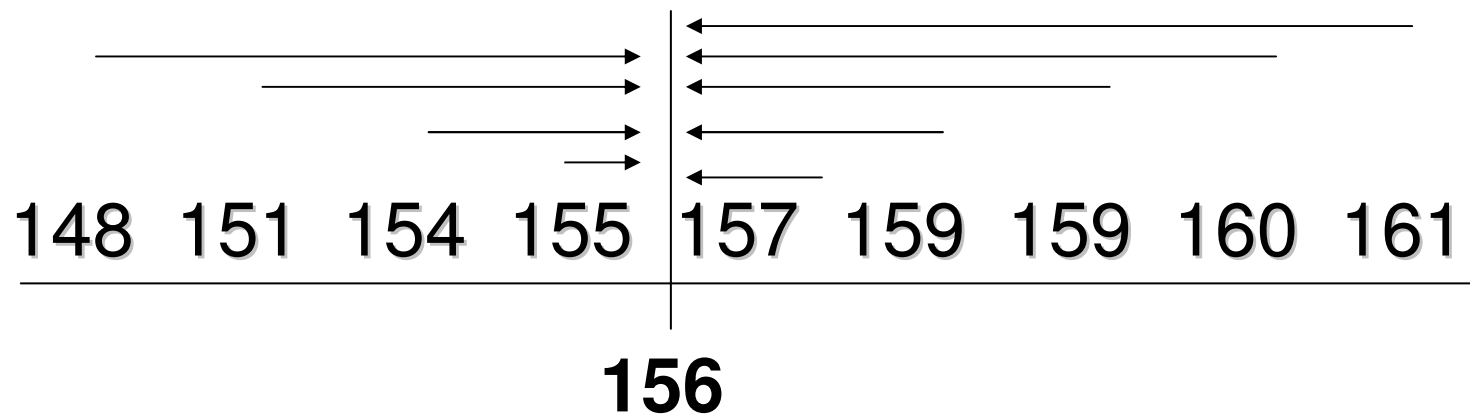
- Τα κυριότερα μέτρα μορφής της κατανομής των δεδομένων είναι ο συντελεστής ασυμμετρίας και ο συντελεστής κύρτωσης.
 - Ο συντελεστής ασυμμετρίας μας δηλώνει πόσο ασύμμετρη είναι η κατανομή των δεδομένων.
 - Ο συντελεστής κύρτωσης δηλώνει το βαθμό συγκέντρωσης των παρατηρήσεων γύρω από ένα κεντρικό σημείο.

Παρουσίαση του Συντελεστή Ασυμμετρίας

- Έστω οι ακόλουθες τιμές HDL 9 ατόμων:



- Έστω οι ακόλουθες τιμές HDL 9 ατόμων:



Διαγράμματα

- Τα διαγράμματα χρησιμοποιούνται ευρύτατα στη στατιστική για την παρουσίαση των δεδομένων με παραστατικό και εύληπτο τρόπο.
- Ανάλογα με τον τύπο της μεταβλητής έχουμε και τα κατάλληλα διαγράμματα.
- Στην περίπτωση μελέτης ενός χαρακτηριστικού χρησιμοποιούμε συνήθως
 - το λεγόμενο ιστόγραμμα για τις ποσοτικές μεταβλητές.
 - το ραβδόγραμμα για τις ποιοτικές μεταβλητές.
 - το διάγραμμα πίτας για τις κατηγορικές μεταβλητές.
- Για δύο χαρακτηριστικά χρησιμοποιούμε
 - τα διαγράμματα διασποράς για τις ποσοτικές μεταβλητές.
 - τα συνδυασμένα ραβδογράμματα για τις ποιοτικές / κατηγορικές μεταβλητές.

Παράδειγμα-Ομαδοποιημένα Δεδομένα

- Τα βάρη (σε gr) ενός τυχαίου δείγματος $n=51$ ποντικών από ένα πληθυσμό ίδιου γένους και ίδιας ηλικίας παρουσιάζονται στον ακόλουθο πίνακα τιμών

(Α. Κυριακούσης, Στατιστικές Μέθοδοι, Εκδ. Σταμούλη, Αθήνα 2000, σελ. 80)

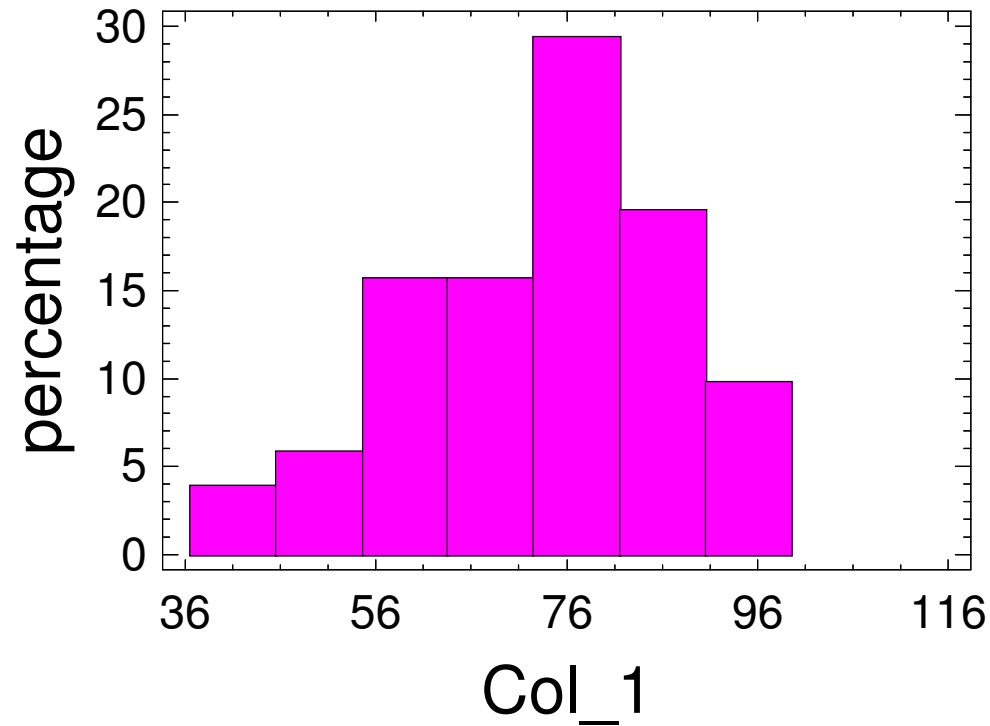
| | | | | | | | | |
|----|----|----|----|----|----|----|----|----|
| 75 | 98 | 42 | 75 | 84 | 87 | 65 | 59 | 63 |
| 86 | 78 | 37 | 99 | 66 | 90 | 79 | 80 | 89 |
| 68 | 57 | 95 | 55 | 79 | 88 | 76 | 60 | 77 |
| 49 | 92 | 83 | 71 | 78 | 53 | 81 | 77 | 58 |
| 93 | 85 | 70 | 62 | 80 | 74 | 69 | 90 | 62 |
| 84 | 64 | 73 | 48 | 72 | 78 | | | |

Πίνακας Συχνοτήτων- Αθροιστικών Συχνοτήτων

| ΚΛΑΣΕΙΣ | ΚΕΝΤΡΙΚΗ ΤΙΜΗ | ΣΥΧΝΟΤΗΤΑ | ΑΘΡΟΙΣΤΙΚΗ ΣΥΧΝΟΤΗΤΑ |
|--------------|------------------|-----------|-------------------------|
| [36.5, 45.5) | 41 | 2 | 2 |
| [45.5, 54.5) | 50 | 3 | 5 |
| [54.5, 63.5) | 59 | 8 | 13 |
| [63.5, 72.5) | 68 | 8 | 21 |
| [72.5, 81.5) | 77 | 15 | 36 |
| [81.5, 90.5) | 86 | 10 | 46 |
| [90.5, 99.5) | 95 | 5 | 51 |

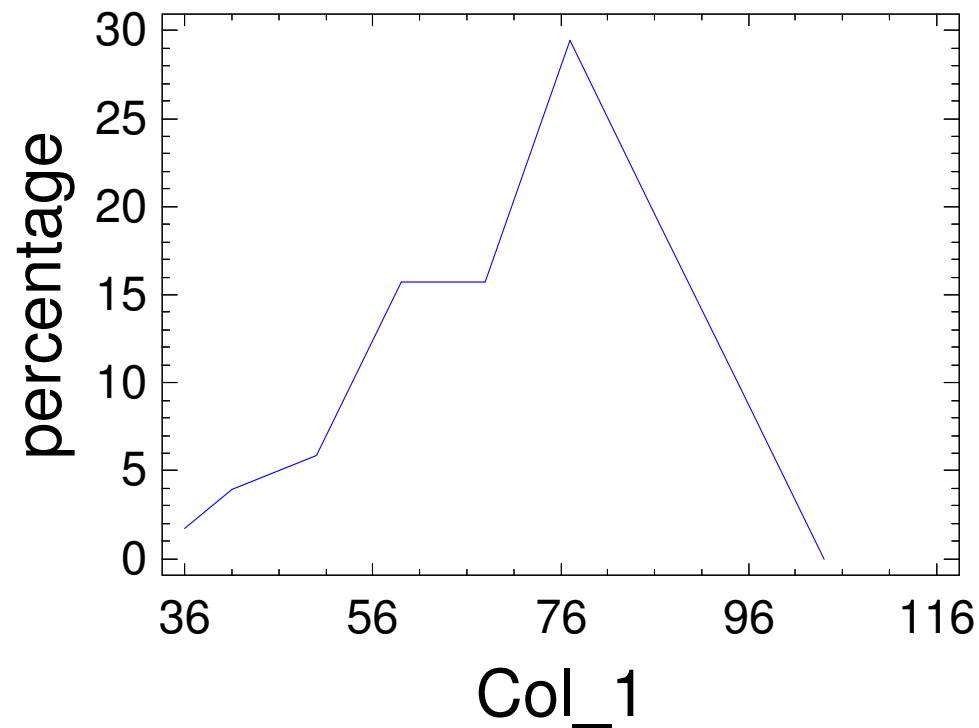
Ιστόγραμμα Σχετικών Συχνοτήτων

Histogram



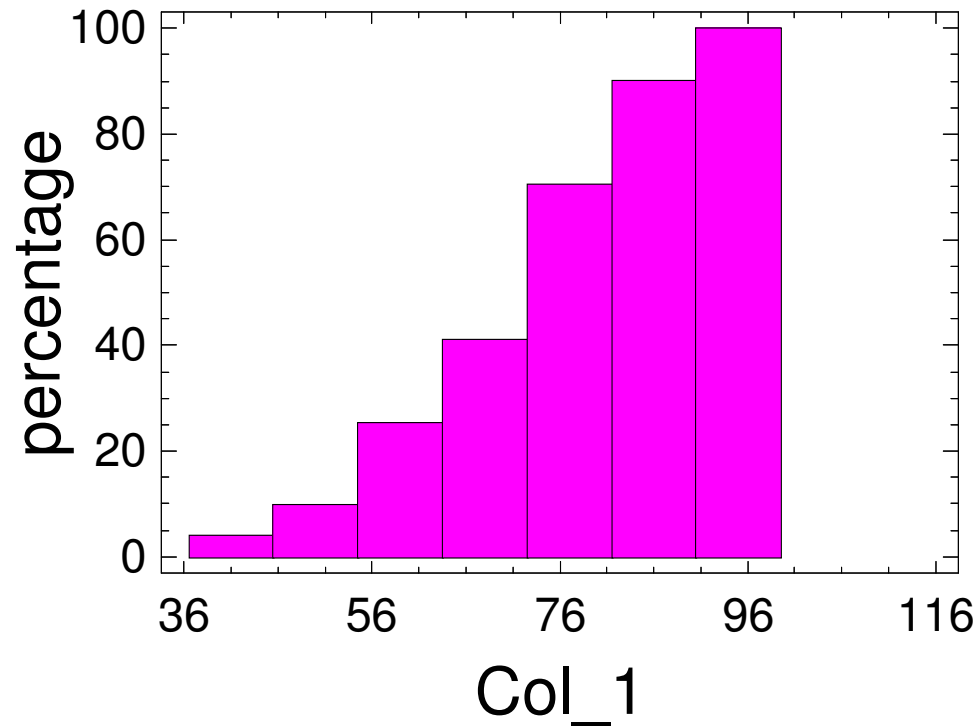
Πολυγωνική Γραμμή Ιστογράμματος Σχετικών Συχνοτήτων

Histogram



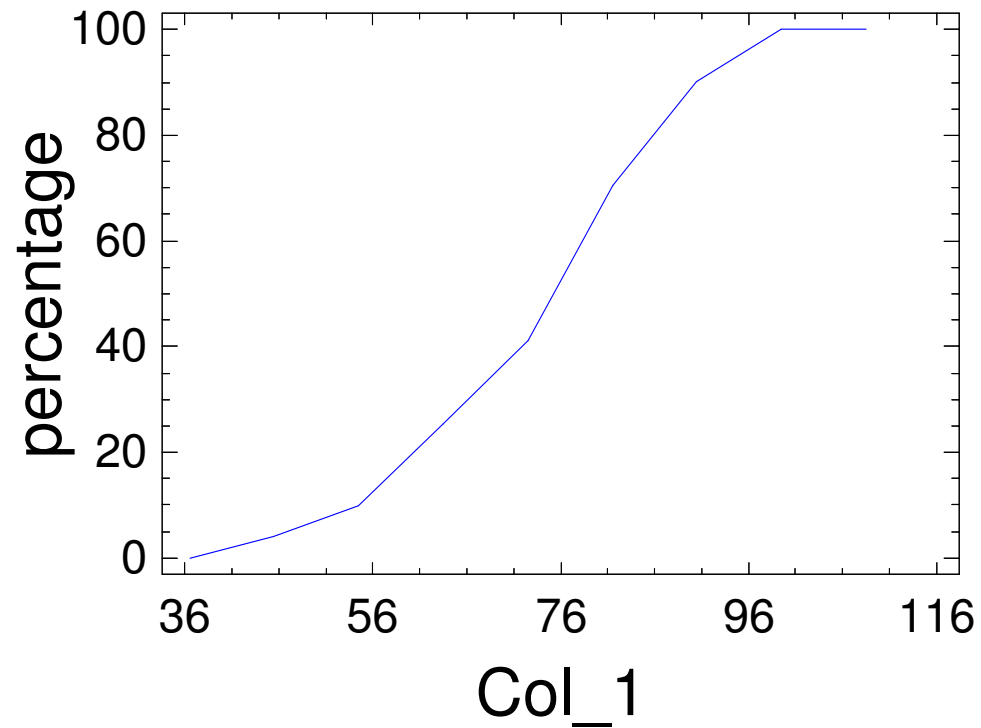
Ιστόγραμμα Αθροιστικών Συχνοτήτων

Histogram



Πολυγωνική Γραμμή Ιστογράμματος Αθροιστικών Συχνοτήτων

Histogram



Θηκόγραμμα

