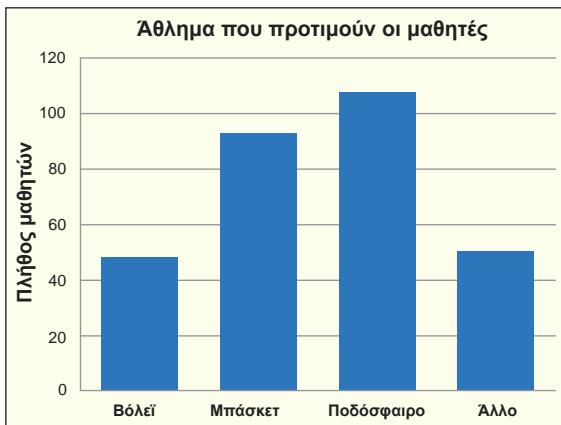


## ΕΝΟΤΗΤΑ 2.2 : ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

### Διερεύνηση

Μια εταιρεία αθλητικών ειδών, πριν επενδύσει σε αθλητικά είδη που προτιμούν οι μαθητές, αποφάσισε να κάνει μια έρευνα. Για τον λόγο αυτό επέλεξε, με τυχαίο τρόπο, δείγμα τριακοσίων μαθητών απ' όλη την Ελλάδα. Ο υπεύθυνος που έκανε την έρευνα, μετά την επεξεργασία των στοιχείων που συγκέντρωσε, παρουσίασε στον διευθυντή της εταιρείας τον παρακάτω πίνακα και τα δυο διαγράμματα.

Άθλημα	Πλήθος μαθητών	Ποσοστό
Βόλεϊ	48	16%
Μπάσκετ	93	31%
Ποδόσφαιρο	108	36%
Άλλο	51	17%
<b>Σύνολο</b>	<b>300</b>	<b>100%</b>



- 1) Ποια είναι η μεταβλητή της έρευνας και ποιο το είδος της;
- 2) Πώς προκύπτουν τα αντίστοιχα ποσοστά για κάθε άθλημα;
- 3) Ποιο είναι το ύψος της κάθε μπάρας στο ραβδόγραμμα;
- 4) Ποια είναι η γωνία του κάθε κυκλικού τομέα στο κυκλικό διάγραμμα;

### Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

- Έστω  $X$  μια μεταβλητή με τιμές  $x_1, x_2, x_3, \dots, x_k$ . **Συχνότητα**  $v_i$  μιας τιμής  $x_i$  λέγεται ο φυσικός αριθμός που εκφράζει πόσες φορές εμφανίζεται η τιμή αυτή στο σύνολο των παρατηρήσεων του δείγματος. Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων μας δίνει το μέγεθος  $n$  του δείγματος. Δηλαδή:

$$v_1 + v_2 + v_3 + \dots + v_k = n$$

**Ορισμοί**

- Η **σχετική συχνότητα**  $f_i$  μιας τιμής  $x_i$  ορίζεται ως ο λόγος της αντίστοιχης συχνότητας  $v_i$  προς το μέγεθος  $v$  του δείγματος. Δηλαδή:

$$f_i = \frac{v_i}{v} \text{ για } i = 1, 2, 3, \dots, \kappa$$

- Σύμφωνα με τους παραπάνω ορισμούς, ο πίνακας του παραδείγματος που αναφέραμε με μεταβλητή «το άθλημα που προτιμούν οι μαθητές» μετατρέπεται, όπως φαίνεται στη συνέχεια και λέγεται πίνακας **κατανομής συχνοτήτων και σχετικών συχνοτήτων**.

Η σχετική συχνότητα μπορεί να εκφραστεί και ως ποσοστό  $f_i\%$ .

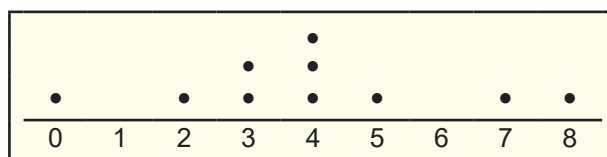
Άθλημα $x_i$	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
Βόλεϊ	48	0,16	16
Μπάσκετ	93	0,31	31
Ποδόσφαιρο	108	0,36	36
Άλλο	51	0,17	17
<b>Σύνολο</b>	<b>300</b>	<b>1,00</b>	<b>100</b>

#### Γραφικές πραστάσεις

- Οι πλέον συνηθισμένοι τρόποι γραφικής παρουσίασης ποιοτικών αλλά και ποσοτικών διακριτών δεδομένων είναι το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.
  - Το **ραβδόγραμμα συχνοτήτων** αποτελείται από ορθογώνιες στήλες, μια για κάθε τιμή της μεταβλητής, όπου το ύψος της κάθε στήλης είναι ίσο με την αντίστοιχη συχνότητα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ραβδόγραμμα σχετικών συχνοτήτων**.
  - Το **κυκλικό διάγραμμα** χρησιμοποιείται επίσης για τη γραφική παράσταση δεδομένων, κυρίως όταν αυτά παίρνουν λίγες τιμές. Η γωνία του κάθε κυκλικού τομέα, είναι ανάλογη της αντίστοιχης σχετικής συχνότητας. Δηλαδή:

$$\alpha_i = 360^\circ \cdot f_i \text{ για } i = 1, 2, 3, \dots, \kappa$$

- Όταν έχουμε λίγες παρατηρήσεις, τότε η κατανομή τους μπορεί να περιγραφεί με το **σημειόγραμμα** στο οποίο οι τιμές παριστάνονται με σημεία υπεράνω ενός άξονα.



- Το **χρονόγραμμα** χρησιμοποιείται για την γραφική απεικόνιση της εξέλιξης σε σχέση με το χρόνο ενός μεγέθους, συνήθως οικονομικού ή δημογραφικού.

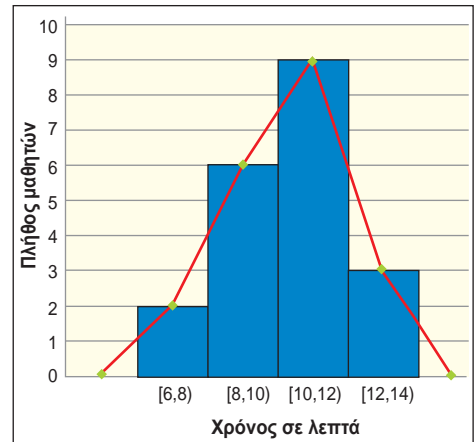
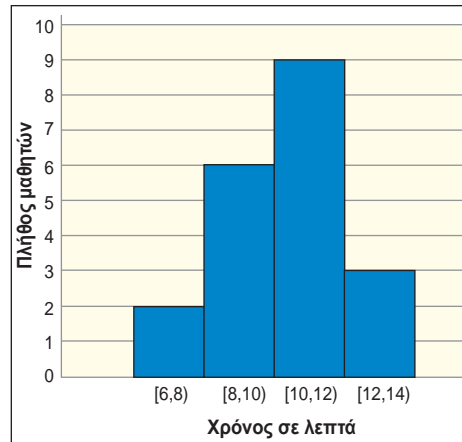


- Στη περίπτωση που έχουμε ποσοτικά συνεχή δεδομένα με πολλές διαφορετικές τιμές, τότε, για την καλύτερη παρουσίασή τους, γίνεται **ομαδοποίηση** αυτών σε κλάσεις, συνήθως ίσου πλάτους. Σχετικός είναι ο παρακάτω πίνακας με τους χρόνους που χρειάστηκαν οι 20 μαθητές ενός τμήματος για να απαντήσουν σε ένα πρόβλημα.

**Δεδομένα σε κλάσεις**

Χρόνος σε λεπτά των μαθητών	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
[6,8)	2	0,10	10
[8,10)	6	0,30	30
[10,12)	9	0,45	45
[12,14)	3	0,15	15
<b>Σύνολο</b>	<b>20</b>	<b>1,00</b>	<b>100</b>

- Η γραφική παρουσίαση ομαδοποιημένων στατιστικών δεδομένων γίνεται με το **ιστόγραμμα συχνοτήτων**. Στο παρακάτω σχήμα φαίνεται η γραφική παρουσίαση των δεδομένων του προηγούμενου πίνακα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ιστόγραμμα σχετικών συχνοτήτων**.
- Αν θεωρήσουμε δυο επιπλέον κλάσεις ίσου πλάτους, μια στην αρχή και μια στο τέλος, με συχνότητα 0 και ενώσουμε τα μέσα των άνω βάσεων, τότε προκύπτει το λεγόμενο **πολύγωνο συχνοτήτων**. Σχετικό είναι το παρακάτω σχήμα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **πολύγωνο σχετικών συχνοτήτων**.



## Εφαρμογές

### Εφαρμογή 1

Σε μια δημοσκόπηση που έγινε για τις δημοτικές εκλογές, 280 άτομα απάντησαν ότι προτιμούν τον υποψήφιο «Α», 320 άτομα τον υποψήφιο «Β» και 200 άτομα τον υποψήφιο «Γ».

- 1) Ποιο ήταν το μέγεθος του δείγματος;
- 2) Να κάνετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- 3) Να παραστήσετε τα δεδομένα με ραβδόγραμμα αλλά και με κυκλικό διάγραμμα.

### Λύση

- 1) Το πλήθος αυτών που απάντησαν είναι:

$$v=280+320+200=800$$

Επομένως το μέγεθος του δείγματος είναι 800 άτομα.

- 2) Οι συχνότητες είναι:  $v_1 = 280$ ,  $v_2 = 320$  και  $v_3 = 200$ . Οι σχετικές συχνότητες είναι:

$$f_1 = \frac{v_1}{v} = \frac{280}{800} = 0,35, \quad f_2 = \frac{v_2}{v} = \frac{320}{800} = 0,40 \quad \text{και} \quad f_3 = \frac{v_3}{v} = \frac{200}{800} = 0,25$$

Ο πίνακας κατανομής συχνοτήτων και σχετικών συχνοτήτων είναι:

Υποψήφιος δήμαρχος	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
Υποψήφιος «Α»	280	0,35	35
Υποψήφιος «Β»	320	0,40	40
Υποψήφιος «Γ»	200	0,25	25
<b>Σύνολο</b>	<b>800</b>	<b>1,00</b>	<b>100</b>

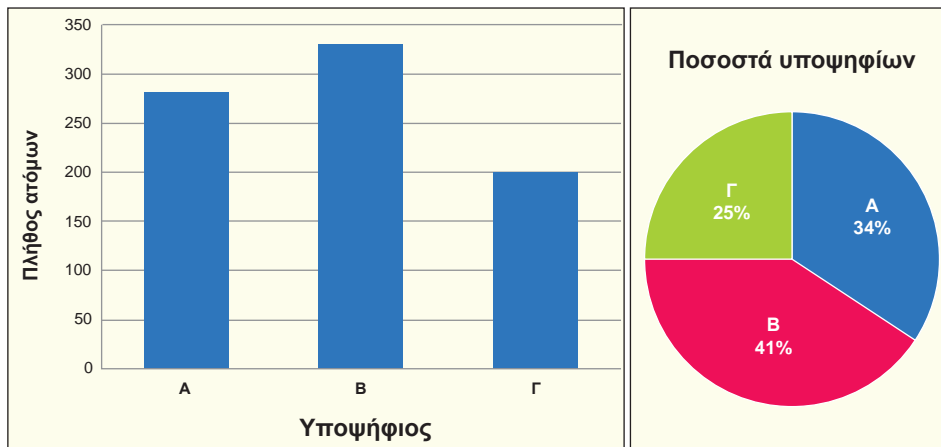
- 3) Για τις γωνίες των κυκλικών τομέων στο κυκλικό διάγραμμα έχουμε:

$$\alpha_1 = 360^\circ \cdot f_1 = 360^\circ \cdot 0,35 = 126^\circ$$

$$\alpha_2 = 360^\circ \cdot f_2 = 360^\circ \cdot 0,40 = 144^\circ$$

$$\alpha_3 = 360^\circ \cdot f_3 = 360^\circ \cdot 0,25 = 90^\circ$$

Παρακάτω φαίνονται το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.



## Εφαρμογή 2

Παρακάτω δίνονται οι χρόνοι, στρογγυλοποιημένοι στο δέκατο του δευτερολέπτου, που απαιτήθηκαν για να τρέξουν 50 αθλητές έναν αγώνα δρόμου 400 m.

52,1	55,3	50	56,4	59,1	54,2	56,7	54,4	57,1	53,7
55,2	55,1	58	59,2	56	55,5	52,5	56,5	58,5	55
55,2	57,3	54,3	53,5	57,9	53	55,4	55,6	52,4	54,5
56,4	59,1	54,2	56,7	55,3	52,4	56,4	54,1	54,3	56,7
54,3	51,5	57	53,2	54,9	55,6	52	55,3	55,1	54,7

- 1) Ποιο είναι το είδος της μεταβλητής, ποιος ο μικρότερος και ποιος ο μεγαλύτερος χρόνος;
- 2) Τι έχετε να παρατηρήσετε για το πλήθος των τιμών των παρατηρήσεων σε σχέση με το πλήθος των παρατηρήσεων;
- 3) Ξεκινώντας από το μικρότερο χρόνο και με βήμα 2 sec, ποιες κλάσεις της μορφής [α,β) δημιουργούνται στις οποίες περιέχονται όλες οι παρατηρήσεις;
- 4) Να παρουσιαστούν οι παρατηρήσεις, ομαδοποιημένες στις παραπάνω κλάσεις ίσου πλάτους, σε έναν πίνακα συχνοτήτων και σχετικών συχνοτήτων.
- 5) Να παραστήσετε τα δεδομένα με ιστόγραμμα συχνοτήτων και με πολύγωνο συχνοτήτων.

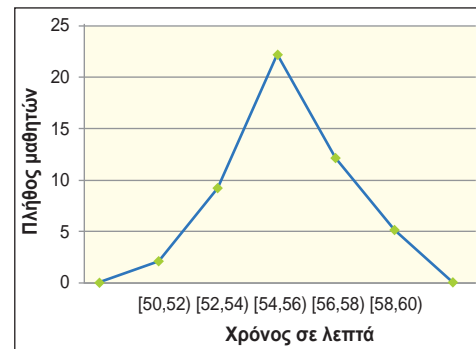
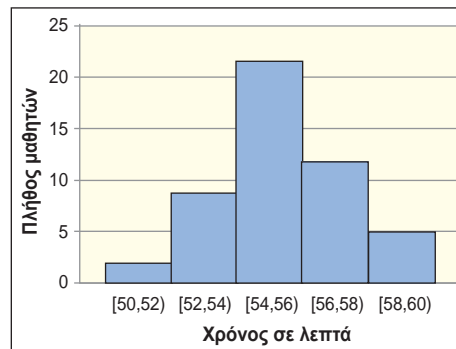
### Λύση

- 1) Η μεταβλητή είναι ποσοτική συνεχής. Ο μικρότερος χρόνος είναι 50 sec και ο μεγαλύτερος 59,2 sec.

- 2) Το πλήθος των τιμών των παρατηρήσεων είναι 37 σε σύνολο 50 παρατηρήσεων. Για να παρουσιαστούν τα δεδομένα, ως έχουν, σε ένα πίνακα συχνοτήτων θα χρειαστούμε 37 γραμμές. Είναι φανερό ότι ένας τέτοιος πίνακας είναι δύσχρηστος και οι πληροφορίες δεν παρουσιάζονται συνοπτικά.
- 3) Οι κλάσεις που δημιουργούνται είναι: [50, 52), [52, 54), [54, 56), [56, 58) και [58, 60)
- 4) Μετά τη διαλογή ο πίνακας συχνοτήτων και σχετικών συχνοτήτων είναι ο παρακάτω.

Κλάσεις με χρόνους σε sec	Συχνότητα $v_i$	Σχετική συχνότητα $f_i$	Σχετική συχνότητα % $f_i\%$
[50, 52)	2	0,04	4
[52, 54)	9	0,18	18
[54, 56)	22	0,44	44
[56, 58)	12	0,24	24
[58, 60)	5	0,10	10
<b>Σύνολο</b>	<b>50</b>	<b>1,00</b>	<b>100</b>

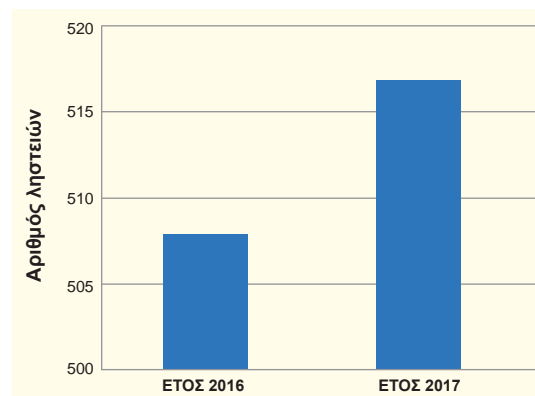
- 5) Παρακάτω φαίνονται το ιστόγραμμα και το πολύγωνο συχνοτήτων.



### Εφαρμογή 3

Σε ένα τηλεοπτικό κανάλι, ένας δημοσιογράφος σχολίασε τη διπλανή γραφική παράσταση ως εξής: «*Η γραφική παράσταση δείχνει ότι σημειώθηκε τεράστια αύξηση του αριθμού των ληστειών το έτος 2017 σε σχέση με το έτος 2016*».

Νομίζετε ότι ο δημοσιογράφος του καναλιού αυτού ερμήνευσε



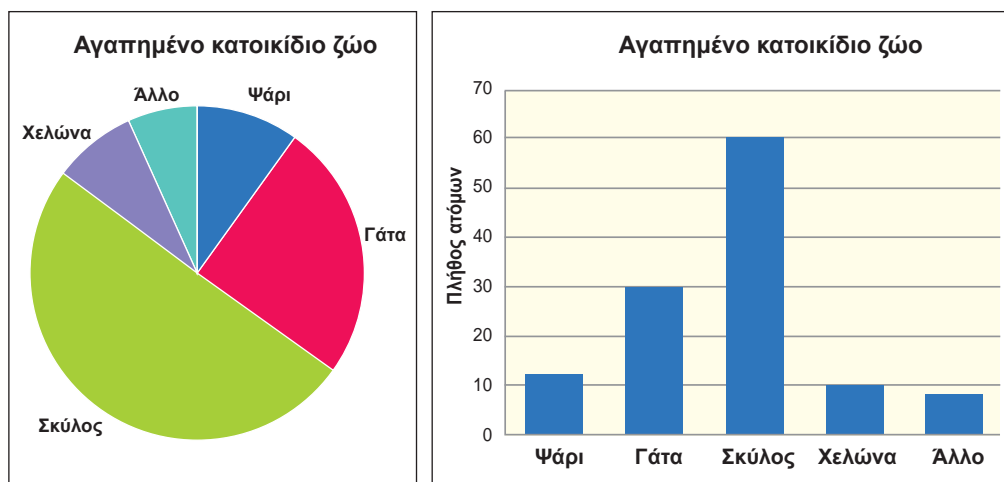
σωστά την γραφική παράσταση; Να γράψετε ένα επιχειρήμα που να τεκμηριώνει την απάντησή σας.

### Λύση

Για τον λόγο ότι η δεύτερη στήλη φαίνεται περίπου διπλάσια στο ύψος από την πρώτη στήλη, δόθηκε η λανθασμένη ερμηνεία. Όμως αυτό που πραγματικά φαίνεται είναι το άνω τμήμα των στηλών, αφού ο άξονας των τεταγμένων αρχίζει από το 500. Έτσι, το πραγματικό ύψος της πρώτης στήλης είναι 508, ενώ της δεύτερης είναι 517. Επομένως ο αριθμός των ληστειών αυξήθηκε μόνο κατά 9 σε σύνολο 500 και πλέον ληστειών. Το γεγονός αυτό δε δικαιολογεί την παραπάνω ερμηνεία.

### Εφαρμογή 4

Ρωτήθηκαν 120 άτομα για το αγαπημένο τους κατοικίδιο ζώο και τα αποτελέσματα παρουσιάζονται με ένα κυκλικό διάγραμμα και ένα ραβδόγραμμα.



- 1) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το ραβδόγραμμα.
- 2) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το κυκλικό διάγραμμα.

### Λύση

- 1) Για το ραβδόγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Πόσες φορές περισσότερα είναι τα άτομα που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος σε σχέση με αυτά που είναι η χελώνα»;
- 2) Για το κυκλικό διάγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Ποιο είναι το ποσοστό των ατόμων που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος»;

## Ασκήσεις - Προβλήματα - Δραστηριότητες

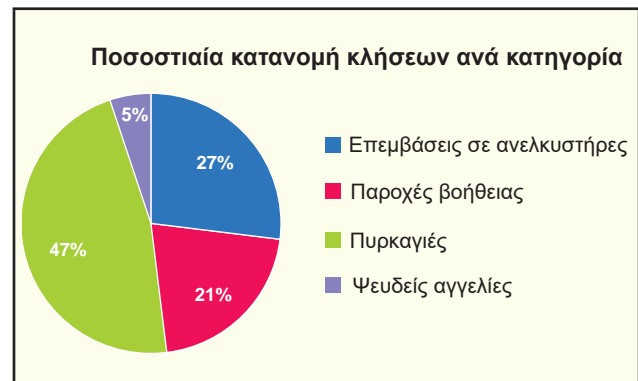
1) Το επάγγελμα του πατέρα 20 μαθητών καταγράφηκε στον διπλανό πίνακα. Να κάνετε πίνακα σχετικών συχνοτήτων και να κατασκευάσετε το αντίστοιχο ραβδόγραμμα συχνοτήτων καθώς και το κυκλικό διάγραμμα.

Επάγγελμα πατέρα	Αριθμός ατόμων
Ιδιωτικός υπάλληλος	6
Δημόσιος υπάλληλος	7
Αυτοαπασχολούμενος	5
Άλλο	2

2) Στον διπλανό πίνακα δίνονται τα καθαρά κέρδη μιας εταιρείας, ανά έτος, από το 2014 έως και το 2017. Να κάνετε χρονόγραμμα όπου να φαίνεται η εξέλιξη των κερδών σε σχέση με τον χρόνο.

Έτος	Κέρδη σε ευρώ
2014	180.000
2015	270.000
2016	230.000
2017	210.000

3) Η στατιστική υπηρεσία της Πυροσβεστικής μας έδωσε το διπλανό κυκλικό διάγραμμα, που παρουσιάζει τα ποσοστά των κλήσεων ανά κατηγορία. Αν το σύνολο των κλήσεων είναι 60.400, να γίνει πίνακας συχνοτήτων.



4) Δίνεται η ποσοστιαία σύνθεση (%) του προσωπικού μιας επιχείρησης, με αριθμό υπαλλήλων 80 άτομα ως προς το μορφωτικό τους επίπεδο.

1) Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.

2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με κυκλικό διάγραμμα.

Τίτλος Σπουδών	Ποσοστό (%)
Μεταπτυχιακό Δίπλωμα	20
Πτυχίο Τριτοβάθμιας Εκπαίδευσης	50
Απολυτήριο Λυκείου	30



5) Οι βαθμοί στην Ιστορία 25 μαθητών, ενός τμήματος της Β΄ τάξης ΓΕΛ, είναι:

1) Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.

16	15	17	16	17
18	17	16	17	18
16	19	17	15	16
17	16	15	17	18
17	14	17	16	19

2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με σημειόγραμμα.

6) Οι πιο κάτω αριθμοί παρουσιάζουν τις ενδείξεις της άνω έδρας ενός ζαριού το οποίο ρίξαμε 30 φορές.

2	5	6	1	2	5	4	3	2	5
1	3	5	4	1	3	2	6	5	4
1	2	6	2	4	3	1	6	4	5

1) Να κατασκευάσετε πίνακα συχνοτήτων.

2) Να κατασκευάσετε ραβδόγραμμα συχνοτήτων.

7) Στον πιο κάτω πίνακα δίνεται η συγκέντρωση ( $\text{mg}/\text{cm}^3$ ) ενός ρύπου στον αέρα 40 πόλεων της χώρας.

16	24	36	47	23	22	43	27	49	48
12	32	17	38	42	27	31	50	38	21
36	19	28	31	28	25	45	12	57	51
22	23	24	25	24	37	43	25	39	51

1) Να ομαδοποιήσετε τις παρατηρήσεις στις κλάσεις:  $[10,20)$ ,  $[20,30)$ ,  $[30,40)$ ,  $[40,50)$  και  $[50,60)$ .

2) Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.

3) Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων.

8) Οι 50 εργάτες ενός εργοστασίου έχουν τις παρακάτω ηλικίες:

21	43	50	25	55	30	28	40	31	51
18	47	52	34	47	32	27	41	35	54
30	48	36	43	38	33	27	39	41	43
32	22	46	52	29	32	34	34	42	36
35	28	57	56	20	38	27	27	40	35

1) Να ομαδοποιήσετε τις ηλικίες στις κλάσεις:  $[18,28)$ ,  $[28,38)$ ,  $[38,48)$  και  $[48,58)$ .

2) Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.

3) Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων.

## Πρόσθετο υλικό

Η εκμάθηση των βασικών εννοιών της στατιστικής γίνεται πιο ενδιαφέρουσα, όταν μπορούμε να εργαστούμε με πληροφορίες που συλλέγονται από εμάς και αποθηκεύονται σε έναν υπολογιστή. Η απλούστερη συλλογή δεδομένων μπο-

ρεί να αφορά πληροφορίες για τους συμμαθητές σας. Η έρευνα που ακολουθεί μπορεί να σας δώσει αρκετά δεδομένα, με τα οποία μπορείτε να απαντήσετε σε μερικές ενδιαφέρουσες ερωτήσεις. Αποθηκεύστε τα δεδομένα που θα συλλέξετε.

- 1) Φύλο
- 2) Ηλικία
- 3) Ύψος σε cm
- 4) Ύψος του πατέρα σας σε cm
- 5) Ύψος της μητέρας σας σε cm
- 6) Το πέμπτο ψηφίο του αριθμού της ταυτότητάς σας
- 7) Το τελευταίο ψηφίο του αριθμού ενός αυτοκινήτου της οικογένειάς σας
- 8) Χρώμα μαλλιών
- 9) Χρώμα ματιών

Μπορείτε να προσθέσετε μια ή περισσότερες ερωτήσεις που σας ενδιαφέρουν. Μοιράστε ένα έντυπο ερωτηματολόγιου, όπου κάθε μαθητής θα γράφει τις απαντήσεις του. Επεξεργαστείτε τα δεδομένα που συγκεντρώσατε με ένα λογιστικό φύλλο.

- 1) Προσδιορίστε το είδος κάθε μεταβλητής
- 2) Κατασκευάστε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων για κάθε μεταβλητή ξεχωριστά.
- 3) Κατασκευάστε τα αντίστοιχα διαγράμματα ανάλογα με το είδος της μεταβλητής.

### ΥΠΟΔΕΙΓΜΑ

α/α	Φύλο	Ηλικία	Ύψος	Ύψος πατέρα	Ύψος μητέρας	Τρίτο ψηφίο του αριθ. της ταυτότητας	Τελευταίο ψηφίο του αριθ. του αυτοκινήτου	Χρώμα μαλλιών	Χρώμα ματιών
1	Θ	17	167	170	160	8	1	M	M
2	Θ	18	165	172	163	4	7	K	M
3	A	18	184	185	171	5	5	K	K
4	Θ	17	175	179	168	3	8	Ξ	M
5	A	19	171	175	165	9	4	K	K
6	A	18	169	170	167	6	7	M	K