

Προσδιορισμός Στατιστικών Παραμέτρων (Συνόλου Αριθμητικών Δεδομένων)

Μαργαρώνης Κώστας¹ – Μόρμορης Μανόλης² – Μπελεχάκη Αγγελική³

⁽¹⁾ Οικονομολόγος Καθηγήτης Δ/βάθμιας Εκπ/σης

⁽²⁾ Καθηγήτης Πληροφορικής Δ/βάθμιας Εκπ/σης

⁽³⁾ Υπευθ. ΠΛΗ.ΝΕ.Τ. Β' Αθήνας

Περίληψη

Με αυτό το διδακτικό πρόβλημα ζητείται από τους μαθητές να υπολογίσουν τις στατιστικές παραμέτρους ενός συνόλου αριθμητικών δεδομένων. Πρόκειται για τυπικούς υπολογισμούς που υλοποιούνται μέσω λογιστικών φύλλων.

Με χρήση των ηλεκτρονικών λογιστι-

κών φύλλων οι μαθητές μπορούν εύκολα να υπολογίσουν τα μέτρα θέσης, τα μέτρα διασποράς, άλλες στατιστικές παραμέτρους όπως την εξίσωση παλινδρόμησης, τον συντελεστή συσχέτισης, τις χρονολογικές σειρές κ.α.

Περιγραφή του διδακτικού προβλήματος

Το διδακτικό πρόβλημα συνίσταται στη δυσκολία υπολογισμού από τους μαθητές των στατιστικών παραμέτρων ενός συνόλου αριθμητικών δεδομένων. Η δυσκολία αυτή γίνεται ακόμη μεγαλύτερη όταν το σύνολο των αριθμητικών δεδομένων είναι μεγάλο. Οι υπολογισμοί αυτοί αποτελούν τυπικούς υπολογισμούς που υλοποιούνται μέσω πληροφοριακών (μηχανογραφικών) συστημάτων. Η παρουσίαση μεγάλων αλγορίθμων ή ακόμη περισσότερο η ανάπτυξη αντίστοιχων προγραμμάτων από τους μαθητές είναι ιδιαίτερα χρονοβόρα ή και δύσκολη. Με τη χρήση των ηλεκτρονικών λογιστικών φύλλων οι μαθητές μπορούν εύκολα αλλά και παραστατικά να υπολογίσουν:

- Τα μέτρα θέσης (μέσο εύρος, μέση τιμή, διάμεσος, επικρατούσα τιμή και τεταρτημόρια)
- Τα μέτρα διασποράς (εύρος μεταβολής, μέση απόλυτη απόκλιση, διακύμανση, τυπική απόκλιση)
- Άλλες στατιστικές παραμέτρους (εξίσωση παλινδρόμησης, συντελεστής συσχέτισης, χρονολογικές σειρές κ.α)

Πλαίσιο Διδασκαλίας

Το αντίστοιχο γνωστικό πεδίο περιλαμβάνεται στο μάθημα: Οικονομικά Μαθηματικά και Στατιστική. Το μάθημα απευθύνεται στους μαθητές της Α΄ τάξης του τομέα Οικονομίας και Διοίκησης των Τ.Ε.Ε. Με τη χρήση Ηλεκτρονικού λογιστικού φύλλου το μάθημα σε μαθητές που γνωρίζουν τη χρήση αντιστοίχων εφαρμογών μπορεί να διαρκέσει δύο διδακτικές ώρες.

Η χρησιμότητα των ΤΠΕ

Οι τεχνολογίες της Πληροφορικής και των Επικοινωνιών, ιδιαίτερα τα τελευταία χρόνια υποστηρίζουν πλειάδα μαθημάτων και γνωστικών περιοχών σε όλες τις βαθμίδες της εκπαίδευσης (*Η Πληροφορική στο Σχολείο, Παιδαγωγικό Ινστιτούτο, 1998*). Εκπαιδευτικό λογισμικό χρησιμοποιείται όλο και πιο συχνά στην εκπαίδευση. Για τις περισσότερο γνωστές μάλιστα τυπικές μηχανογραφικές εφαρμογές, λογισμικό γενικής χρήσης, όπως ηλεκτρονικά λογιστικά φύλλα μπορεί να χρησιμοποιηθεί για να υποστηρίξει τη διδασκαλία αλλά και να απαλλάξει τους μαθητές από πολλούς επαναλαμβανόμενους υπολογισμούς, σε μαθήματα στα οποία προέχει ο υπολογισμός του αποτελεσμάτος και όχι η εξοικείωση με τις αριθμητικές πράξεις.

Παρουσίαση και σχολιασμός της διδασκαλίας

Εισαγωγικά

Οι στατιστικές παράμετροι είναι αντιπροσωπευτικοί αριθμοί με τις οποίες αντικαθιστούμε το σύνολο των αριθμητικών δεδομένων του πληθυσμού που εξετάζουμε και βοηθούν στην εξαγωγή συμπερασμάτων.

Τα μέτρα θέσης είναι στατιστικές παράμετροι οι οποίες καθορίζουν μια τιμή της μεταβλητής, γύρω από την οποία τείνουν να συγκεντρωθούν οι παρατηρήσεις μας.

Τέτοια μέτρα είναι:

- Το μέσο εύρος
- Η μέση τιμή
- Η διάμεσος
- Η επικρατούσα τιμή
- Τα τεταρτημόρια

Τα μέτρα διασποράς είναι παράμετροι που δείχνουν το πόσο είναι «απλωμένα» τα δεδομένα γύρω από κάποιο μέτρο θέσης. Ο προσδιορισμός αυτών των μέτρων είναι απαραίτητος προκειμένου να καταλήξουμε σε αξιόπιστα συμπεράσματα για τα δεδομένα μας και για να τα συγκρίνουμε με άλλα ομοειδή δεδομένα. Τα μέτρα διασποράς είναι:

- Το εύρος μεταβολής
- Η μέση απόλυτη απόκλιση
- Η διακύμανση
- Η τυπική απόκλιση

Εφαρμογή

Οι ετήσιες αποδοχές 100 υπαλλήλων ενός οργανισμού δίνονται παρακάτω σε χιλιάδες Δρχ. ($340,75 \text{ Δρχ} = 1\text{€}$):

3000, 3200, 3400, 3600, 3800, 4000, 4200, 4400, 4600, 4800, 5800, 5600, 5400, 5200, 5000, 4800, 4600, 4400, 4200, 4000, 5000, 5200, 5400, 5600, 5800, 6000, 6200, 6400, 6600, 6800, 6000, 6200, 6400, 6600, 6800, 7000, 7200, 7400, 7600, 7800, 8800, 8600, 8400, 8200, 8000, 7800, 7600, 7400, 7200, 7000, 8000, 8200, 8400, 8600, 8800, 9000, 9200, 9400, 9600, 10000, 3550, 3500, 3450, 3400, 3350, 3300, 3250, 3200, 3150, 3100, 3600, 3650, 3700, 3750, 3800, 3850, 3900, 3950, 4000, 4050, 4100, 4150, 4200, 4250, 4300, 4350, 4400, 4450, 4500, 4550, 4600, 4650, 4700, 4750, 4800, 4850, 4900, 4950, 5000, 5050

1. Να βρεθεί η επικρατούσα τιμή
2. Να ομαδοποιηθούν τα παραπάνω δεδομένα σε κλάσεις ίσου πλάτους
3. Να υπολογισθεί η μέση τιμή
4. Να υπολογιστούν η διάμεσος, το πρώτο και το τρίτο τεταρτημόριο
5. Να υπολογισθεί η διακύμανση
6. Να υπολογιστεί η τυπική απόκλιση

Διδακτική Υπόδειξη 1

Για να προχωρήσουμε στη λύση του προβλήματος μας πρέπει πρώτα απ' όλα να καταχωρήσουμε τα δεδομένα μας σε ένα φύλλο του Excel.

Υπολογισμός επικρατούσας τιμής

Η επικρατούσα τιμή υπολογίζεται από τη συνάρτηση `mode` του excel.

Αν υποθέσουμε ότι οι εκατό τιμές που αντιστοιχούν στους ετήσιους μισθούς των υπαλλήλων του οργανισμού έχουν καταχωρηθεί στη στήλη Α από τη δεύτερη έως την εκατοστή πρώτη γραμμή (Σχ. 1), καταχωρούμε στο κελί C14 τη συνάρτηση `MODE` ως εξής: `=MODE(A2:A101)`. Έτσι βρέθηκε η επικρατούσα τιμή για τα δεδομένα της εφαρμογής μας και είναι ίση με 4000 δηλαδή 4.000.000 Δρχ.

Διδακτική υπόδειξη 2

Ομαδοποίηση

Για την ομαδοποίηση των δεδομένων χρησιμοποιούμε επτά κλάσεις ίσου εύρους $c=1000$. Έτσι η πρώτη κλάση περιλαμβάνει τα εισοδήματα που είναι ίσα ή μεγαλύτερα από 3000 και μικρότερα από 4000. Η δεύτερη περιλαμβάνει τα εισοδήματα που είναι μεγαλύτερα ή ίσα από 4000 και μικρότερα από 5000. Όμοια φτιάχνουμε και τις υπόλοιπες κλάσεις μέχρι την τελευταία που περιλαμβάνει εισοδήματα μεγαλύτερα ή ίσα από 9000 και μέχρι το ποσό των 10000.

Οι κλάσεις καταχωρούνται στη στήλη Β του φύλλου μας με επικεφαλίδα «Ετήσιο εισόδημα».

Υπολογισμός συχνοτήτων

Συχνότητα μιας τιμής x_i της μεταβλητής X είναι ο φυσικός αριθμός n_i που φανερώνει πόσες φορές παρουσιάζεται στο δείγμα (ή σε ολόκληρο τον πληθυσμό) η τιμή αυτή. (Δ. Καραγεώργου κ.α., 2000)

Η στήλη C περιέχει τον αριθμό των υπαλλήλων που αντιστοιχεί στη κλάση της στήλης Β (συχνότητα). Αν, δηλαδή, η κλάση που είναι καταχωρη-

μένη στο κελί B2 είναι [3000,4000), τότε στο κελί C2 πρέπει να περιέχεται ο αριθμός των υπαλλήλων με ετήσιο εισόδημα μεταξύ 3000 και 4000. Αυτό γίνεται ως εξής:

Θεωρώντας ότι στη στήλη A από το κελί 2 έως το 101 έχουμε καταχωρήσει τα ετήσια εισοδήματα των υπαλλήλων, χρησιμοποιώντας τη συνάρτηση COUNTIF μπορούμε να μετρήσουμε τα περιεχόμενα των κελιών που βρίσκονται ανάμεσα στις 3000 και 4000.

Έτσι λοιπόν καταχωρούμε στο κελί C2: =COUNTIF(A2:A101;"<4000").

Αν ακολουθήσουμε την ίδια μέθοδο τότε καταχωρούμε στα κελιά:

C3 = COUNTIF(A2:A101;"<5000")-SUM(C2:C2)

C4 = COUNTIF(A2:A101;"<6000")- SUM(C2:C3)

C5 = COUNTIF(A2:A101;"<7000")- SUM(C2:C4)

C6 = COUNTIF(A2:A101;"<8000")- SUM(C2:C5)

C7 = COUNTIF(A2:A101;"<9000")- SUM(C2:C6)

C8 = COUNTIF(A2:A101;"<10000")- SUM(C2:C7)

Τα περιεχόμενα των κελιών της στήλης C αντιστοιχούν στις τιμές της μεταβλητής v_i

Διδακτική υπόδειξη 2

Κέντρο κλάσης

Η στήλη D περιέχει τους κεντρικούς όρους (κεντρική τιμή) της κλάσης της γραμμής. Π.χ. στη γραμμή 2, όπου η κλάση είναι [3000,4000) το κέντρο κλάσης είναι 3.500 το οποίο και καταχωρείται στο κελί D2.

Τα περιεχόμενα των κελιών της στήλης D αντιστοιχούν στις τιμές της μεταβλητής x_i

Διδακτική υπόδειξη 3

Υπολογισμός αθροιστικών συχνοτήτων

Αθροιστική συχνότητα μιας τιμής x_i λέγεται το άθροισμα των συχνοτήτων των τιμών της μεταβλητής που είναι μικρότερες ή ίσες με τη τιμή αυτή. (Δ. Καραγεώργου κ.α., 2000)

Η στήλη H περιέχει τις αθροιστικές συχνότητες. Και εδώ θα χρησιμοποιήσουμε τη συνάρτηση COUNTIF ως εξής:

H2= COUNTIF(A2:A101;"<4000")

H3= COUNTIF(A2:A101;"<5000")

H4 = COUNTIF(A2:A101;"<6000")

H5 = COUNTIF(A2:A101;"<7000")

H6 = COUNTIF(A2:A101;"<8000")

H7 = COUNTIF(A2:A101;"<9000")

H8 = COUNTIF(A2:A101;"<10000")

Τα περιεχόμενα των κελιών της στήλης Η αντιστοιχούν στις τιμές της μεταβλητής N_i

Διδακτική υπόδειξη 2

Υπολογισμός άλλων παραμέτρων

Προκειμένου να υπολογίσουμε τις στατιστικές παραμέτρους πρέπει να δημιουργήσουμε στήλες για τον υπολογισμό των $\Sigma n_i x_i$, Σx_i^2 , Σn_i και $\Sigma n_i x_i^2$ (περιέχονται στους τύπους με τους οποίους υπολογίζουμε τις ζητούμενες παραμέτρους).

Υπολογισμός του γινομένου $n_i x_i$

Στη στήλη Ε περιέχονται τα γινόμενα του αριθμού των υπαλλήλων για κάθε κλάση επί τα κέντρα της συγκεκριμένης κλάσης ($n_i x_i$), έτσι στο κελί Ε2 καταχωρούμε τη σχέση = C2 * D2 (το άθροισμα της στήλης μας δίνει $\Sigma n_i x_i$).

Υπολογισμός του x_i^2

Στη στήλη F περιέχονται τα κέντρα κλάσης υψωμένα στο τετράγωνο (x_i^2). Καταχωρούμε στο κελί F2 τη σχέση = D2 * D2. Το άθροισμα της στήλης μας δίνει Σx_i^2

Υπολογισμός του $n_i x_i^2$

Στο κελί G2 καταχωρούμε τη σχέση = C2 * F2, δηλαδή το γινόμενο του αριθμού των υπαλλήλων αυτής της κλάσης επί το κέντρο της κλάσης στο τετράγωνο ($n_i x_i^2$).

Χρησιμοποιώντας τη μέθοδο αντιγραφής των τύπων, αντιγράφουμε τους τύπους σε όλες τις γραμμές μέχρι την τελευταία κλάση.

Διδακτική υπόδειξη 4

Στη γραμμή 10 και για τις στήλες C, E, F και G καταχωρούμε τα αθροίσματα των κελιών των γραμμών 2 έως 8 για τις αντίστοιχες στήλες.

Αναλυτικά καταχωρούμε:

Στο κελί C10 τη σχέση =SUM(C2:C8)

Στο κελί E10 τη σχέση =SUM(E2:E8)

Στο κελί F10 τη σχέση =SUM(F2:F8)

Στο κελί G10 τη σχέση =SUM(G2:G8)

Έτσι βλέπουμε $\Sigma n_i = 100$, $\Sigma n_i x_i = 554.000$, $\Sigma x_i^2 = 323.750.000$ και $\Sigma n_i x_i^2 = 3.411.000.000$.

Α	Β	Γ	Δ	Ε	ΣΤ	Ζ	Η	Θ	ΙΑ	ΙΒ
1	Είδος τιμής	Αριθμός αποτελεσμάτων	Ρυθμός κλάσης	Συνολικός αριθμός	Μέση τιμή	Τετάρτη	Μέση τιμή	Μέση τιμή	Μέση τιμή	Μέση τιμή
2	3000	1.000	0.010	25	3.500	0.050	12.250.000	287.750.000	25	25
3	3200	1.400	0.014	30	4.500	0.070	135.000	201.750.000	55	55
4	3400	1.800	0.018	57	5.500	0.090	310.500	332.750.000	92	92
5	3600	2.200	0.022	79	6.500	0.110	517.500	422.250.000	113	113
6	3800	2.600	0.026	105	7.500	0.130	787.500	567.500.000	149	149
7	4000	3.000	0.030	135	8.500	0.150	1.155.000	727.500.000	194	194
8	4200	3.400	0.034	169	9.500	0.170	1.602.500	857.250.000	253	253
9	4400	3.800	0.038	207	10.500	0.190	2.167.500	967.750.000	332	332
10	4600	4.200	0.042	249	11.500	0.210	2.862.500	1.057.250.000	431	431
11	4800	4.600	0.046	295	12.500	0.230	3.687.500	1.127.250.000	550	550
12	5000	5.000	0.050	345	13.500	0.250	4.652.500	1.177.250.000	695	695
13	5200	5.400	0.054	399	14.500	0.270	5.757.500	1.207.250.000	874	874
14	5400	5.800	0.058	457	15.500	0.290	7.002.500	1.217.250.000	1084	1084
15	5600	6.200	0.062	519	16.500	0.310	8.387.500	1.207.250.000	1303	1303
16	5800	6.600	0.066	585	17.500	0.330	9.912.500	1.167.250.000	1538	1538
17	6000	7.000	0.070	655	18.500	0.350	1.158.750	1.097.250.000	1793	1793
18	6200	7.400	0.074	729	19.500	0.370	1.391.250	997.250.000	2122	2122
19	6400	7.800	0.078	807	20.500	0.390	1.608.750	857.250.000	2529	2529
20	6600	8.200	0.082	889	21.500	0.410	1.811.250	676.250.000	3018	3018
21	6800	8.600	0.086	975	22.500	0.430	2.008.750	455.250.000	3593	3593
22	7000	9.000	0.090	1065	23.500	0.450	2.201.250	195.250.000	4258	4258
23	7200	9.400	0.094	1159	24.500	0.470	2.388.750	55.250.000	5017	5017
24	7400	9.800	0.098	1257	25.500	0.490	2.571.250	0.000	5870	5870
25	7600	10.200	0.102	1359	26.500	0.510	2.748.750	0.000	6819	6819
26	7800	10.600	0.106	1465	27.500	0.530	2.921.250	0.000	7864	7864
27	8000	11.000	0.110	1575	28.500	0.550	3.088.750	0.000	9007	9007
28	8200	11.400	0.114	1689	29.500	0.570	3.251.250	0.000	10248	10248
29	8400	11.800	0.118	1807	30.500	0.590	3.408.750	0.000	11587	11587
30	8600	12.200	0.122	1929	31.500	0.610	3.561.250	0.000	13026	13026
31	8800	12.600	0.126	2055	32.500	0.630	3.708.750	0.000	14565	14565
32	9000	13.000	0.130	2185	33.500	0.650	3.851.250	0.000	16204	16204
33	9200	13.400	0.134	2319	34.500	0.670	3.988.750	0.000	17943	17943
34	9400	13.800	0.138	2457	35.500	0.690	4.121.250	0.000	19782	19782
35	9600	14.200	0.142	2599	36.500	0.710	4.248.750	0.000	21721	21721
36	9800	14.600	0.146	2745	37.500	0.730	4.371.250	0.000	23760	23760
37	10000	15.000	0.150	2895	38.500	0.750	4.488.750	0.000	25900	25900
38	10200	15.400	0.154	3049	39.500	0.770	4.601.250	0.000	28141	28141
39	10400	15.800	0.158	3207	40.500	0.790	4.708.750	0.000	30482	30482
40	10600	16.200	0.162	3369	41.500	0.810	4.811.250	0.000	32923	32923
41	10800	16.600	0.166	3535	42.500	0.830	4.908.750	0.000	35464	35464
42	11000	17.000	0.170	3705	43.500	0.850	5.001.250	0.000	38105	38105
43	11200	17.400	0.174	3879	44.500	0.870	5.088.750	0.000	40846	40846
44	11400	17.800	0.178	4057	45.500	0.890	5.171.250	0.000	43687	43687
45	11600	18.200	0.182	4239	46.500	0.910	5.248.750	0.000	46628	46628
46	11800	18.600	0.186	4425	47.500	0.930	5.321.250	0.000	49669	49669
47	12000	19.000	0.190	4615	48.500	0.950	5.388.750	0.000	52810	52810
48	12200	19.400	0.194	4809	49.500	0.970	5.451.250	0.000	56051	56051
49	12400	19.800	0.198	5007	50.500	0.990	5.508.750	0.000	59492	59492
50	12600	20.200	0.202	5209	51.500	1.010	5.561.250	0.000	63133	63133
51	12800	20.600	0.206	5415	52.500	1.030	5.608.750	0.000	66974	66974
52	13000	21.000	0.210	5625	53.500	1.050	5.651.250	0.000	71015	71015
53	13200	21.400	0.214	5839	54.500	1.070	5.688.750	0.000	75256	75256
54	13400	21.800	0.218	6057	55.500	1.090	5.721.250	0.000	79697	79697
55	13600	22.200	0.222	6279	56.500	1.110	5.748.750	0.000	84338	84338
56	13800	22.600	0.226	6505	57.500	1.130	5.771.250	0.000	89179	89179
57	14000	23.000	0.230	6735	58.500	1.150	5.788.750	0.000	94220	94220
58	14200	23.400	0.234	6969	59.500	1.170	5.798.750	0.000	99461	99461
59	14400	23.800	0.238	7207	60.500	1.190	5.801.250	0.000	104902	104902
60	14600	24.200	0.242	7449	61.500	1.210	5.798.750	0.000	110543	110543
61	14800	24.600	0.246	7695	62.500	1.230	5.788.750	0.000	116384	116384
62	15000	25.000	0.250	7945	63.500	1.250	5.771.250	0.000	122425	122425
63	15200	25.400	0.254	8199	64.500	1.270	5.748.750	0.000	128666	128666
64	15400	25.800	0.258	8457	65.500	1.290	5.711.250	0.000	135107	135107
65	15600	26.200	0.262	8719	66.500	1.310	5.661.250	0.000	141748	141748
66	15800	26.600	0.266	8985	67.500	1.330	5.598.750	0.000	148589	148589
67	16000	27.000	0.270	9255	68.500	1.350	5.521.250	0.000	155630	155630
68	16200	27.400	0.274	9529	69.500	1.370	5.431.250	0.000	162871	162871
69	16400	27.800	0.278	9807	70.500	1.390	5.328.750	0.000	170312	170312
70	16600	28.200	0.282	10089	71.500	1.410	5.213.750	0.000	177953	177953
71	16800	28.600	0.286	10375	72.500	1.430	5.086.250	0.000	185794	185794
72	17000	29.000	0.290	10665	73.500	1.450	4.946.250	0.000	193835	193835
73	17200	29.400	0.294	10959	74.500	1.470	4.793.750	0.000	202076	202076
74	17400	29.800	0.298	11257	75.500	1.490	4.628.750	0.000	210517	210517
75	17600	30.200	0.302	11559	76.500	1.510	4.451.250	0.000	219158	219158
76	17800	30.600	0.306	11865	77.500	1.530	4.261.250	0.000	227999	227999
77	18000	31.000	0.310	12175	78.500	1.550	4.058.750	0.000	237040	237040
78	18200	31.400	0.314	12489	79.500	1.570	3.843.750	0.000	246281	246281
79	18400	31.800	0.318	12807	80.500	1.590	3.616.250	0.000	255722	255722
80	18600	32.200	0.322	13129	81.500	1.610	3.376.250	0.000	265363	265363
81	18800	32.600	0.326	13455	82.500	1.630	3.123.750	0.000	275204	275204
82	19000	33.000	0.330	13785	83.500	1.650	2.858.750	0.000	285245	285245
83	19200	33.400	0.334	14119	84.500	1.670	2.581.250	0.000	295486	295486
84	19400	33.800	0.338	14457	85.500	1.690	2.291.250	0.000	305927	305927
85	19600	34.200	0.342	14799	86.500	1.710	2.008.750	0.000	316568	316568
86	19800	34.600	0.346	15145	87.500	1.730	1.723.750	0.000	327409	327409
87	20000	35.000	0.350	15495	88.500	1.750	1.436.250	0.000	338450	338450
88	20200	35.400	0.354	15849	89.500	1.770	1.146.250	0.000	349691	349691
89	20400	35.800	0.358	16207	90.500	1.790	853.750	0.000	361132	361132
90	20600	36.200	0.362	16569	91.500	1.810	558.750	0.000	372773	372773
91	20800	36.600	0.366	16935	92.500	1.830	263.750	0.000	384614	384614
92	21000	37.000	0.370	17305	93.500	1.850	0.000	0.000	396655	396655
93	21200	37.400	0.374	17679	94.500	1.870	0.000	0.000		

για το τρίτο τεταρτημόριο (κελί G16) = $QUARTILE(A2:A101;3)$

Διδακτική υπόδειξη 3

Στην εφαρμογή αυτή μας έχουν δοθεί τα δεδομένα μη ομαδοποιημένα και γι' αυτό μπορούμε να χρησιμοποιήσουμε την έτοιμη συνάρτηση του EXCEL. Σε περίπτωση όμως που τα δεδομένα είναι στη διάθεσή μας μόνο ομαδοποιημένα, τότε πρέπει να χρησιμοποιήσουμε τον τύπο υπολογισμού της διαμέσου όπως δίνεται στη διδακτική υπόδειξη 1 και να ακολουθήσουμε τα παρακάτω βήματα:

1. Από το φύλλο Excel (Σχ. 1) βλέπουμε ότι ο αριθμός 50 (= $n/2$) βρίσκεται μεταξύ των αθροιστικών συχνοτήτων 23 και 53, άρα η **διάμεσος** βρίσκεται στην κλάση [4000 – 5000)

2. Για τον υπολογισμό της διαμέσου χρειαζόμαστε τα παρακάτω:

i θέση διαμέσου (περιέχεται στο κελί C19)

a_{i-1} το κατώτερο όριο κλάσης που περιέχει τη διάμεσο (περιέχεται στο κελί C20)

$n/2$ το πλήθος των παρατηρήσεων δια δύο (περιέχεται στο κελί C21, αλλά μπορούμε να χρησιμοποιήσουμε τη σχέση = $C10 / 2$)

N_{i-1} η αθροιστική συχνότητα της προηγούμενης κλάσης (περιέχεται στο κελί C22 και D2)

c το πλάτος της κλάσης (περιέχεται στο κελί C23)

v_i συχνότητα της κλάσης (περιέχεται στο κελί C24 και C3)

Έτσι βάσει του τύπου της διαμέσου καταχωρούμε στο κελί F19 τη σχέση:
= $C20 + (C21 - C22) * C23 / C24$

Στη συνέχεια για τον υπολογισμό του πρώτου και του τρίτου τεταρτημορίων, χρησιμοποιούμε τους τύπους και καταχωρούμε στο κελί G15 τη σχέση:
= $C20 + (C10 / 4 - D2) * C23 / C3$ (για το πρώτο τεταρτημόριο)

και στο κελί G16 τη σχέση:

= $C20 + (3 * C10 / 4 - D2) * C23 / C3$ (για το τρίτο τεταρτημόριο)

Υπολογισμός διακύμανσης

Στο κελί C16 καταχωρούμε τη σχέση = $((G10 / C10) - (C15 * C15))$ που είναι ο τύπος της διακύμανσης

Διδακτική υπόδειξη 1

Υπολογισμός τυπικής απόκλισης

Η τυπική απόκλιση είναι ίση με τη θετική τετραγωνική ρίζα της διακύμανσης και άρα στο κελί C17 πληκτρολογούμε: = $SQRT((G10 / C10) - (C15 * C15))$.

Διδακτική υπόδειξη 1

Επικρατούσα τιμή

Ως *επικρατούσα τιμή* (*mode*) ορίζεται η τιμή της μεταβλητής που παρουσιάζεται τις περισσότερες φορές.

Μέση τιμή

Ορίζουμε ως *μέση τιμή* ή *μέσο αριθμητικό* (*mean*) το πηλίκο του αθροίσματος των τιμών αυτών x_1, x_2, \dots, x_v που παίρνει μια μεταβλητή X διά του πλήθους των τιμών της.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_v}{v}$$

Αν τα δεδομένα δίνονται ομαδοποιημένα τότε ο τύπος είναι:

$$\bar{x} = \frac{\sum x_i v_i}{\sum v_i}$$

Αν εξετάσουμε όλο τον πληθυσμό, η μέση τιμή συμβολίζεται διεθνώς με το ελληνικό γράμμα μ , ενώ αν αναφερόμαστε σε δείγμα πληθυσμού συμβολίζεται με \bar{x} .

Διάμεσος

Διάμεσος ή *διάμεση τιμή* (δ) λέγεται η τιμή της μεταβλητής που χωρίζει το σύνολο των τιμών σε δύο ισοπληθείς ομάδες, όταν οι τιμές που παίρνει η μεταβλητή τοποθετηθούν σε αύξουσα τάξη.

Ως *διάμεσος* ορίζεται η τιμή της μεταβλητής που αντιστοιχεί στη μεσαία παρατήρηση, αν το πλήθος των παρατηρήσεων είναι περιττό ή το ημίθροισμα των τιμών της μεταβλητής που αντιστοιχούν στις δύο μεσαίες παρατηρήσεις, αν το πλήθος είναι άρτιο.

Τεταρτημόρια

Τεταρτημόρια λέγονται οι τιμές της μεταβλητής που χωρίζουν το σύνολο των τιμών της σε 4 ισοπληθείς ομάδες, όταν οι τιμές της μεταβλητής τοποθετηθούν σε αύξουσα τάξη.

Σε ένα σύνολο διατεταγμένων παρατηρήσεων σε αύξουσα τάξη ορίζουμε:

Το πρώτο τεταρτημόριο Q_1 ως τιμή της μεταβλητής, κάτω της οποίας βρίσκεται το πολύ το 25% του συνολικού αριθμού των παρατηρήσεων.

Το δεύτερο τεταρτημόριο Q_2 συμπίπτει με τη διάμεσο, η οποία έχει ήδη ορισθεί, δηλαδή $Q_2 = \delta$

Τέλος το τρίτο τεταρτημόριο Q_3 ορίζεται αντίστοιχα ως η τιμή της μετα-

βλητής, κάτω της οποίας βρίσκεται το πολύ το 75% του συνολικού αριθμού των παρατηρήσεων.

Σύμφωνα με τα παραπάνω, αν n είναι το πλήθος των διατεταγμένων παρατηρήσεων, τότε η θέση που κατέχει το πρώτο τεταρτημόριο Q_1 είναι $(n+1)/4$ και η θέση που κατέχει το τρίτο τεταρτημόριο Q_3 είναι $3(n+1)/4$.

Ο υπολογισμός των τεταρτημορίων σε ομαδοποιημένα δεδομένα γίνεται όπως ο αντίστοιχος της διαμέσου ή γραφικά με τη βοήθεια του διαγράμματος των αθροιστικών συχνοτήτων. Η θέση των τεταρτημορίων Q_1 και Q_3 σε ομαδοποιημένα δεδομένα δεχόμαστε ότι βρίσκεται στη $n/4$ και $3n/4$ θέση αντίστοιχα.

Παρακάτω δίνονται οι τύποι υπολογισμού της διαμέσου, του πρώτου και του τρίτου τεταρτημορίου.

$$\delta = a_{i-1} + (n/2 - N_{i-1}) \cdot c / n_i$$

$$Q_1 = a_{i-1} + (n/4 - N_{i-1}) \cdot c / n_i$$

$$Q_3 = a_{i-1} + (3n/4 - N_{i-1}) \cdot c / n_i$$

Οπου:

i	θέση διαμέσου (ή τεταρτημορίου)
a_{i-1}	το κατώτερο όριο κλάσης που περιέχει τη διάμεσο (ή το τεταρτημόριο)
$n/2$	το πλήθος των παρατηρήσεων διά δύο (για τα τεταρτημόρια $n/4$, $3n/4$)
N_{i-1}	η αθροιστική συχνότητα της προηγούμενης κλάσης
c	το πλάτος της κλάσης
n_i	συχνότητα της κλάσης

Εύρος μεταβολής

Είναι η διαφορά της μικρότερης τιμής από τη μεγαλύτερη τιμή της μεταβλητής X .

Μέση Απόλυτη Απόκλιση

Η *Μέση Απόλυτη Απόκλιση* μας δείχνει πόσο, κατά μέσο όρο, απέχουν οι παρατηρήσεις μας από τη μέση τιμή δηλαδή είναι ο μέσος αριθμητικός των αποκλίσεων των τιμών μιας μεταβλητής, σε απόλυτες τιμές, από τον μέσο αριθμητικό.

Διακύμανση

Ως *διακύμανση* των τιμών της μεταβλητής X ορίζουμε τη μέση τιμή των

τετραγώνων των αποκλίσεων των τιμών της μεταβλητής από το μέσο αριθμητικό της.

$$S^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 \quad (\text{για αταξινόμητα δεδομένα})$$

$$S^2 = \frac{\sum v_i x_i^2}{\sum v_i} - \bar{x}^2 \quad (\text{για ταξινομημένα δεδομένα})$$

Διδακτική υπόδειξη 2

MODE

Η συνάρτηση αυτή ορίζει την επικρατούσα τιμή από ένα πεδίο τιμών.

Σύνταξη: Mode(περιοχή κελιών)

Παράδειγμα

Mode(a1:a8)

Θα πάρουμε σαν αποτέλεσμα την επικρατούσα τιμή των κελιών a1, a2, a3, a4, a5, a6, a7 και a8.

COUNTIF

Η συνάρτηση αυτή ορίζει το πλήθος των κελιών σε ένα σύνολο τιμών που εκπληρώνει κάποιο κριτήριο.

Σύνταξη: Countif (περιοχή κελιών;κριτήριο)

Παράδειγμα

Countif(a1:a8;">55")

Ορίζει το πλήθος των κελιών ανάμεσα στα κελιά A1 και A8 που η τιμή τους είναι μεγαλύτερη από 55.

MEDIAN

Ορίζει τη διάμεσο για τους αριθμούς που έχει δώσει ο χρήστης. Είναι ο αριθμός που βρίσκεται στη μέση από ένα σύνολο αριθμών δηλαδή πρέπει οι μισοί αριθμοί να έχουν τιμές μικρότερες από τη διάμεσο και οι άλλοι μισοί μεγαλύτεροι.

Σύνταξη: **MEDIAN**(αριθμός1, αριθμός2,...)

Αριθμός1, αριθμός2,... είναι 1 έως 30 αριθμοί για τους οποίους ζητάμε

την διάμεσο. Τα ορίσματα πρέπει να είναι είτε αριθμοί ή ονόματα ή πίνακες ή αναφορές που περιέχουν αριθμούς.

Παρατηρήσεις

Εάν το πλήθος των αριθμών που αναζητούμε τη διάμεσο είναι μονός αριθμός, τότε η συνάρτηση MEDIAN υπολογίζει τον μεσαίο αριθμό ενώ αν είναι ζυγός υπολογίζει τον μέσο όρο των δύο αριθμών που βρίσκονται στη μέση.

Παραδείγματα

MEDIAN(1, 2, 3, 4, 5) ορίζει τη τιμή 3

MEDIAN(1, 2, 3, 4, 5, 6) ορίζει τη τιμή 3.5, την μέση τιμή των 3 και 4.

QUARTILE

Ορίζει το τεταρτημόριο ενός συνόλου τιμών. Το τεταρτημόριο βοηθά στο πως να μοιραστούν πληθυσμοί σε ομάδες, Για παράδειγμα μπορούμε να χρησιμοποιήσουμε τη συνάρτηση QUARTILE για να βρούμε το πρώτο 25% των εισοδημάτων σε ένα πληθυσμό.

Σύνταξη: **QUARTILE**(πίνακας,τεταρτημόριο)

Το όρισμα *πίνακας* είναι μια περιοχή κελιών με αριθμητικές τιμές για τις οποίες θέλουμε να υπολογίσουμε το τεταρτημόριο.

Το όρισμα *Τεταρτημόριο* ορίζει τις τιμές:

0: η ελάχιστη τιμή

1: το πρώτο τεταρτημόριο

2: η διάμεσος

3: το τρίτο τεταρτημόριο

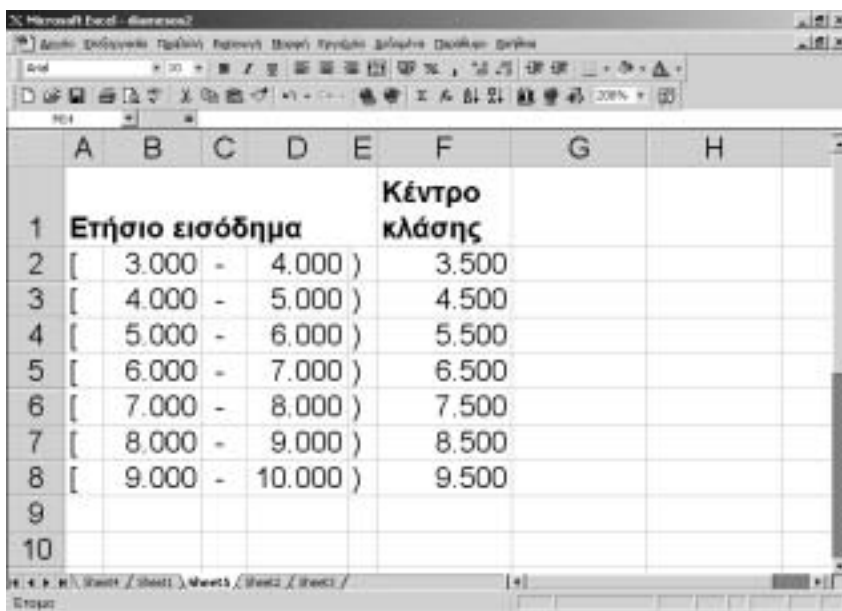
4: η μέγιστη τιμή

Παράδειγμα

QUARTILE({1,2,4,7,8,9,10,12},1) ορίζει 3.5

Διδακτική υπόδειξη 3

Τα κέντρα κλάσης, με τον τρόπο που έχουμε καταχωρήσει τα δεδομένα στο Excel, δεν μπορούν να υπολογιστούν και γι αυτό τα καταχωρούμε. Υπάρχει όμως ένας άλλος τρόπος καταχώρησης των ετησίων εισοδημάτων που βοηθάει στον υπολογισμό τους. Πληκτρολογούμε, λοιπόν τα δεδομένα όπως φαίνεται στο Σχ. 2 και στο κελί F2 πληκτρολογούμε τη σχέση: = (B2+D2)/2. Μετά με τη μέθοδο αντιγραφής τύπων, αντιγράφουμε τη σχέση και στα κελιά F3, F4, F5, F6, F7, F8.



The screenshot shows a Microsoft Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H
1	Ετήσιο εισόδημα					Κέντρο κλάσης		
2	[3.000	-	4.000)	3.500		
3	[4.000	-	5.000)	4.500		
4	[5.000	-	6.000)	5.500		
5	[6.000	-	7.000)	6.500		
6	[7.000	-	8.000)	7.500		
7	[8.000	-	9.000)	8.500		
8	[9.000	-	10.000)	9.500		
9								
10								

Σχήμα 2

Διδακτική υπόδειξη 4

Αντιγραφή των τύπων

Μπορούμε να αντιγράψουμε ή να γεμίσουμε με τύπους μια περιοχή κελιών με τον ίδιο τρόπο που συμπληρώνουμε δεδομένα στις περιοχές. Για να γεμίσει με έναν τύπο μια περιοχή, επιλέγουμε το κελί που περιέχει τον τύπο και σύρουμε τη λαβή συμπλήρωσης προς τα κάτω, προς τα πάνω, προς τα δεξιά ή προς τα αριστερά μέχρι εκεί που θέλουμε. Ο τύπος θα αντιγραφεί αυτόματα στα νέα κελιά με προσαρμοσμένα τα ορίσματα των κελιών.

Ενδεικτική βιβλιογραφία

- Δ. Καραγεώργου – Α.Μ. Κόκλα – Ε. Παπακωνσταντίνου, (2000), *Στατιστική Επιχειρήσεων*, ΟΕΔΒ.
- Ν. Κανελλόπουλου – Σ. Μπουρέλου, (2000), *Εφαρμογές H/Y*, ΟΕΔΒ.
- Β. Παπασπύρου – Α. Μπελεχάκη, (2000), *Χρήση H/Y – EXCEL*, ΟΕΔΒ
- Η Πληροφορική στο Σχολείο, (1998), Παιδαγωγικό Ινστιτούτο