**Ποσοτική Ανάλυση Γλωσσικών Δεδομένων**

**Άσκηση 3**

Έστω το ακόλουθο πείραμα:

Τα παιδιά μιας τάξης έχουν χωριστεί σε τρεις ομάδες και εξετάζονται στο ίδιο τεστ. Κατά τη βαθμολόγηση του τεστ, εξετάζονται 3 ειδών λάθη (er1, er2, er3) και προκύπτει η τελική βαθμολογία (score). Τα δεδομένα του πειράματος παρουσιάζονται στον ακόλουθο πίνακα:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **groupid** | **er1** | **er2** | **er3** | **score** |
| **1** | 7 | 4 | 27 | 144,6 |
| **2** | 30 | 26 | 24 | 33,2 |
| **3** | 36 | 4 | 27 | 86,6 |
| **1** | 12 | 20 | 26 | 86,8 |
| **3** | 34 | 7 | 29 | 81,2 |
| **1** | 0 | 18 | 26 | 116,8 |
| **3** | 27 | 18 | 1 | 67,8 |
| **1** | 28 | 5 | 27 | 99,6 |
| **3** | 20 | 19 | 18 | 75,4 |
| **2** | 26 | 22 | 16 | 54,8 |
| **3** | 14 | 4 | 7 | 134,6 |
| **2** | 33 | 13 | 18 | 67,4 |
| **1** | 27 | 17 | 11 | 68,8 |
| **3** | 18 | 24 | 10 | 66 |
| **2** | 38 | 7 | 14 | 76,2 |
| **1** | 28 | 18 | 7 | 64,6 |
| **2** | 29 | 16 | 29 | 64,2 |
| **3** | 27 | 25 | 22 | 42,6 |
| **1** | 4 | 4 | 20 | 152 |
| **3** | 21 | 12 | 17 | 94,6 |
| **2** | 27 | 17 | 5 | 70 |
| **2** | 38 | 26 | 8 | 20,4 |
| **1** | 14 | 9 | 17 | 117,6 |
| **2** | 39 | 30 | 2 | 7,6 |
| **3** | 8 | 12 | 10 | 122 |
| **1** | 1 | 19 | 6 | 115,8 |
| **2** | 24 | 24 | 24 | 51,2 |
| 2 | 16 | 27 | 12 | 60,6 |
| 1 | 5 | 26 | 2 | 87,6 |
| 3 | 23 | 29 | 25 | 38 |

1. Προβάλλετε ομαδοποιημένα τα δεδομένα για το 3ο είδος λαθών (er3)
2. Ελέγξτε την ορθότητα καθεμίας από τις ακόλουθες προτάσεις:

Α) Υπάρχει κάποια ομάδα παιδιών που διαφοροποιείται ως προς τη μέση τιμή των λαθών είδους 1 (er1)

Β) Υπάρχει κάποια ομάδα παιδιών που διαφοροποιείται ως προς τη μέση τιμή των λαθών είδους 2 (er2)

Γ) Υπάρχει κάποια ομάδα παιδιών που διαφοροποιείται ως προς τη μέση τιμή των λαθών είδους 3 (er3)

1. Για τις ορθές προτάσεις του προηγούμενου ερωτήματος, βρείτε ποια ομάδα διαφοροποιείται.
2. Κάποιος ισχυρίζεται ότι η τελική βαθμολογία μπορεί να εκτιμηθεί αν εκφραστεί ως γραμμικός συνδυασμός των πληθών των λαθών, δηλαδή

score = a \* er1 + b \* er2 + c \* er3 + d

Υπολογίστε τους συντελεστές a, b, c και τον σταθερό όρο d και ελέγξτε αν όντως υπολογίζεται η τελική βαθμολογία.