

## Εκφώνηση Άσκησης

Αντικείμενο: Ανάλυση του dataset ObesityDataSet.csv με χρήση R.

### 1. Ερώτημα

- Ελέγξτε τον τύπο των μεταβλητών και τη διάσταση του dataset.
- Μετατρέψτε τις κατηγορικές μεταβλητές σε factors (π.χ., Gender, SMOKE).

### 2. Ερώτημα

Δημιουργήστε δύο νέες κατηγορικές μεταβλητές:

- Age\_f:
  - 1 για Ανήλικους (<18 ετών).
  - 2 για Ενήλικους (≥18 ετών).
- Weight\_f:
  - Κατηγορίες βάσει τεταρτημορίων (Q1, Q2, Q3):
    - "Flyweight" (≤Q1), "Lightweight" (>Q1 & ≤Q2), "Middleweight" (>Q2 & ≤Q3), "Heavyweight" (>Q3).

### 3. Ερώτημα

- Απομονώστε τις ποσοτικές μεταβλητές σε νέο dataframe.
- Υπολογίστε περιγραφικά μέτρα (mean, sd, κλπ.) για όλες τις ποσοτικές μεταβλητές.

### 4. Ερώτημα

Για την μεταβλητή Age:

- Κατασκευάστε:
  - Ιστόγραμμα με συνάρτηση πυκνότητας πιθανότητας.
  - Θηκόγραμμα (boxplot).
  - QQ-plot.
- Σχολιάστε την κατανομή.

### 5. Ερώτημα

Για την κατηγορική μεταβλητή MTRANS:

- Δημιουργήστε:
  - Πίνακα συχνοτήτων/σχετικών συχνοτήτων.
  - Ραβδόγραμμα και πίτα με ετικέτες συχνοτήτων.

#### 6. Ερώτημα

- Scatterplot για τη σχέση μεταξύ Height και Weight.
- Προσθέστε την Age\_f ως διαχωριστικό στοιχείο (χρώμα/σχήμα σημείων).

#### 7. Ερώτημα

- Θηκόγραμμα του Weight ανά Gender.

#### 8. Ερώτημα

- Πίνακας συνάφειας για Age\_f και Weight\_f.
- Ομαδοποιημένο & στοιβαγμένο ραβδόγραμμα για τις δύο μεταβλητές.