Exercise Instructions

Objective: Analysis of the dataset ObesityDataSet.*csv* using R.

Question 1

- Check the variable types and the dimensions of the dataset.
- Convert categorical variables into factors (e.g., Gender, SMOKE).

Question 2

Create two new categorical variables:

- Age_f:
    - 1 for Minors (<18 years old).
    - 2 for Adults (≥18 years old).
- Weight_f:
    - Categories based on quartiles (Q1, Q2, Q3):
        - "Flyweight" (≤Q1),
        - "Lightweight" (>Q1 & ≤Q2),
        - "Middleweight" (>Q2 & ≤Q3),
        - "Heavyweight" (>Q3).

Question 3

- Isolate the quantitative variables into a new dataframe.
- Calculate descriptive statistics (mean, sd, etc.) for all quantitative variables.

Question 4

For the variable Age:

- Construct:
    - A histogram with a probability density function.
    - A boxplot.
    - A QQ-plot.
- Comment on the distribution.

Question 5

For the categorical variable MTRANS:

- Create:
    - A frequency/relative frequency table.
    - A bar chart and a pie chart with frequency labels.

Question 6

- Scatterplot for the relationship between Height and Weight.
- Add Age_f as a distinguishing element (color/shape of points).

Question 7

- Boxplot of Weight by Gender.

Question 8

- Contingency table for Age_f and Weight_f.
- Grouped & stacked bar chart for the two variables.