ANALYSIS OF THE DATASET:
AGE, WEIGHT, AND HEIGHT RELATIONSHIPS

Albert J. Páez, Yurley K. Torrado

1 Introduction

This report presents an analysis of a dataset containing information on age, weight, height, and gender distribution. The study includes scatter plots to explore relationships between age, weight, and height, a pie chart to show gender distribution, and a bar chart to highlight age distribution among the first ten individuals.

2 Scatter Plots and Correlation Matrix

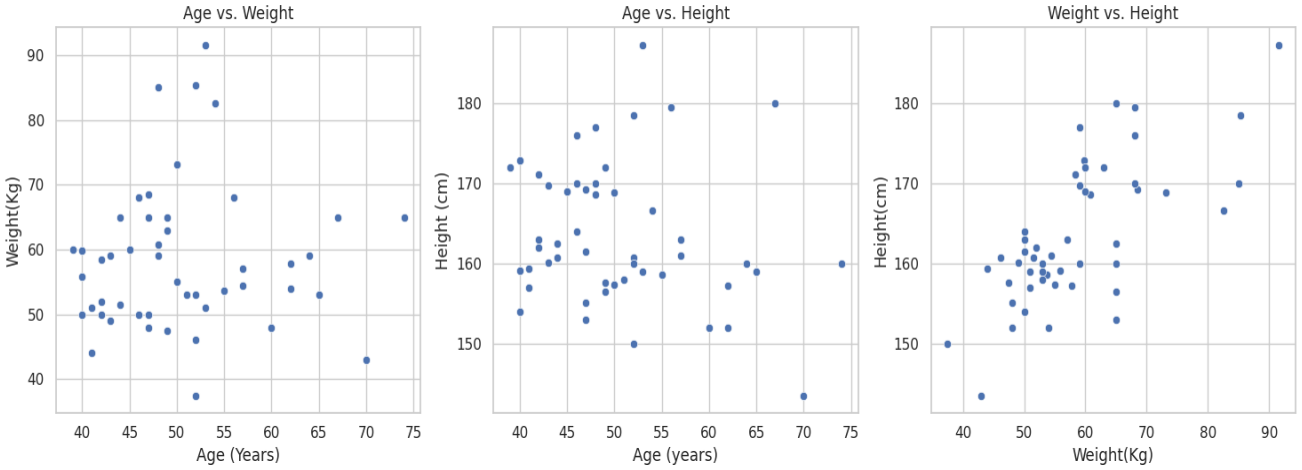

Figure 1. Scatter Plots.

	Age (years)	**Weigth (Kg)**	**Heigth (cm)**
Age (years)	1.000000	0.047952	-0.182663
Weigth (Kg)	0.047952	1.000000	0.706547
Heigth (cm)	-0.182663	0.706547	1.000000

Table 1. Correlation Matrix.

2.1. Age vs. Weight

- The scatter plot illustrates a weak correlation (0.0479) between age and weight.
- The data points appear widely scattered without forming a clear trend, suggesting that weight does not significantly increase or decrease with age.
- This lack of correlation implies that other factors, such as lifestyle or genetics, may have a greater impact on weight variation than age alone

2.2. Age vs. Height

- The correlation coefficient is -0.1827, indicating a weak negative relationship.
- While there is a slight tendency for height to decrease with age, this effect is not strong enough to be considered a defining trend.
- This observation aligns with general biological trends where height may decrease slightly with aging due to factors such as spinal compression and posture changes, though not significantly in this dataset.

2.3. Weight vs. Height

- A strong positive correlation (0.7065) is observed, indicating that as height increases, weight also tends to be higher.
- This relationship is expected as taller individuals generally have larger body frames and more muscle mass, contributing to higher weight measurements.
- The scatter plot visually confirms this trend, showing a clear upward trajectory in the data points.

3 Gender Distribution

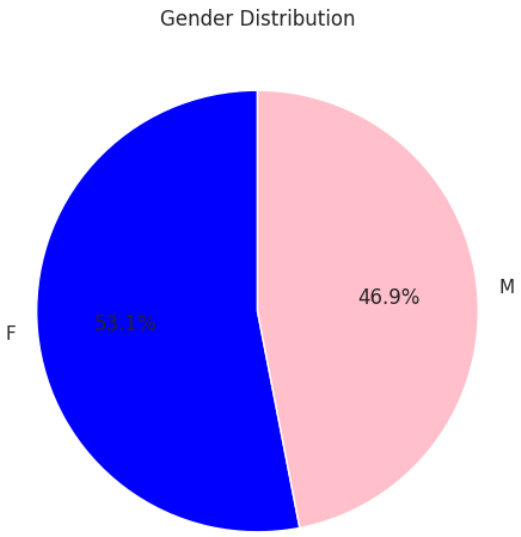

Figure 2. Gender Distribution.

A pie chart was generated to analyze the gender composition in the dataset. The distribution between males and females is nearly balanced, ensuring that the dataset is representative of both genders.

- The visual representation allows for a quick assessment of gender proportions, avoiding potential biases that could arise if one gender were significantly overrepresented.
- A balanced gender ratio ensures that any conclusions drawn from the dataset apply to both males and females equally.

3.1. Age vs. Weight Trend Analysis

- The scatter plot for age vs. weight does not show a clear upward or downward trend. The weak correlation indicates that weight variation is not significantly influenced by age.
- Unlike variables such as height and weight, which show a clearer relationship, age and weight appear to be independent in this dataset.
- The lack of a trend suggests that factors such as lifestyle, diet, and activity level may play a more significant role in weight changes over time rather than age itself.

4 Age Distribution of the First Ten Individuals

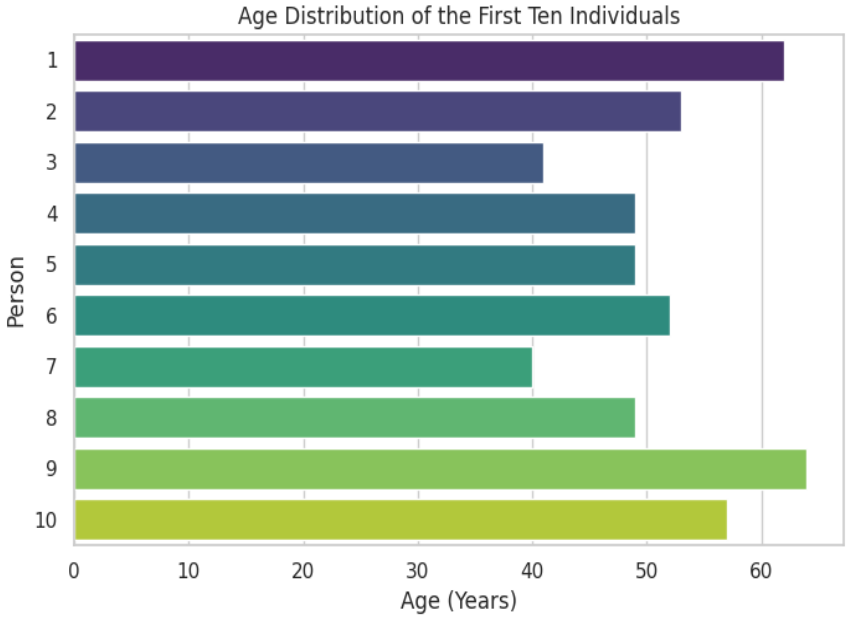

Figure 3. Age Distribution of the First Ten Individuals.

A horizontal bar chart was created to visualize the ages of the first ten individuals in the dataset. The ages vary significantly, without a clear concentration in a specific age range.

- The data distribution suggests diversity in the sampled population, making it more representative of a broader group.
- The variation in age prevents skewed interpretations that could arise from analyzing only a specific age group.

5 Conclusions

- There is a strong correlation between weight and height, confirming the expected relationship that taller individuals tend to weigh more.
- Age does not significantly influence weight or height in this dataset, indicating that these attributes are likely determined by other factors.

- The gender distribution is balanced, ensuring that the dataset provides a representative analysis for both males and females.
- The first ten individuals exhibit a diverse age range, adding variety to the sample and avoiding potential biases.

These findings help in understanding the dataset's structure and relationships between variables, offering insights into basic demographic patterns and possible influencing factors beyond what is directly observable in the dataset.