Principles of Data Science – 5530

Assignment 1

Nikhil Jagadeesh Sriram

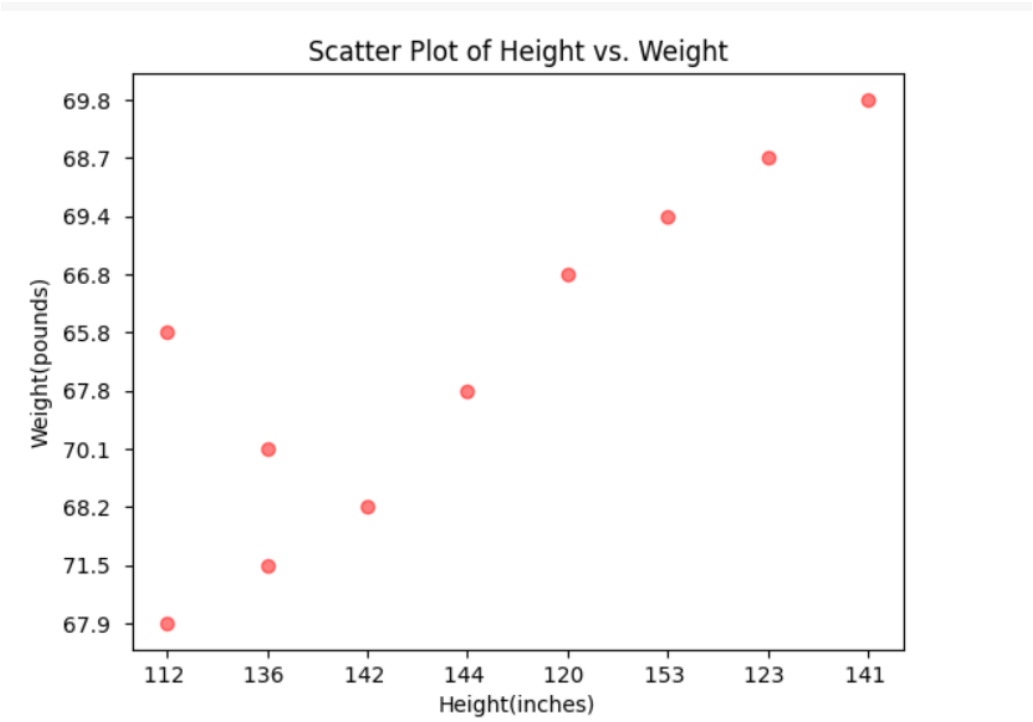
Student ID: 16352573

Question1: Frailty Dataset

Visualizations:

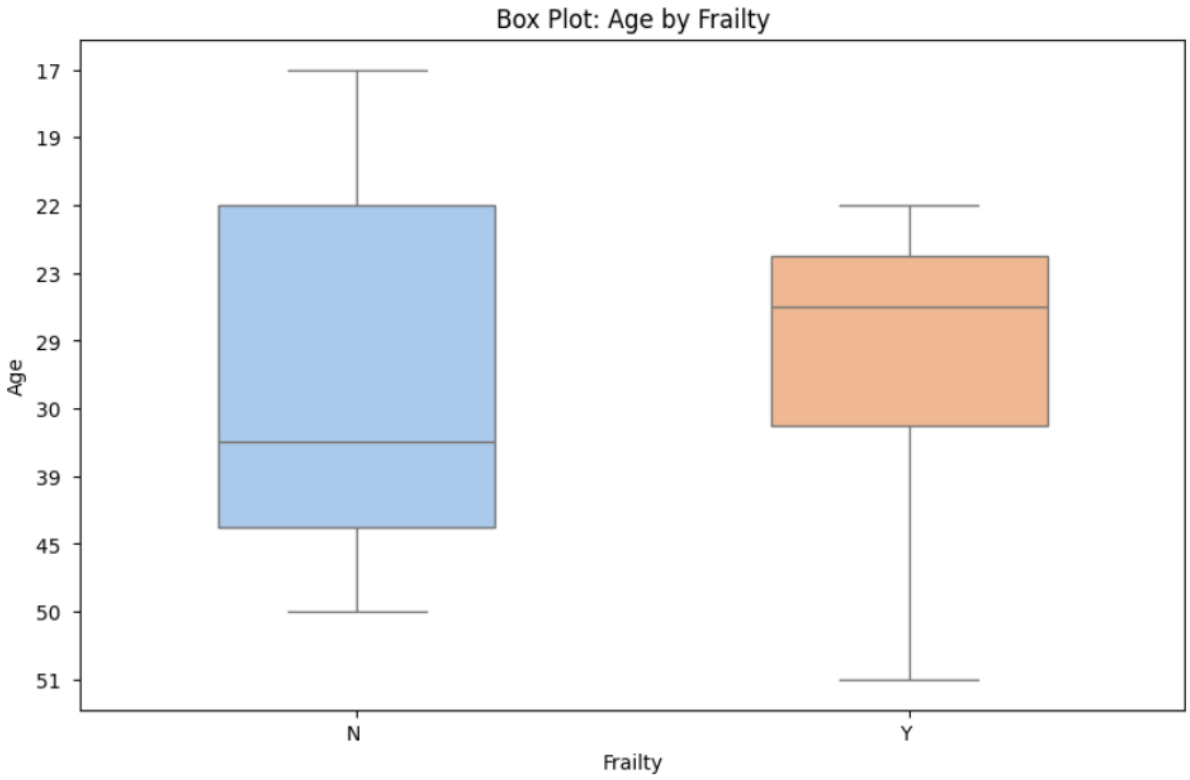
Plot1:

A scatter plot visually represents the relationship between total scores and the parental level of education. This graphical representation allows for an exploration of how total scores vary based on different levels of parental education.



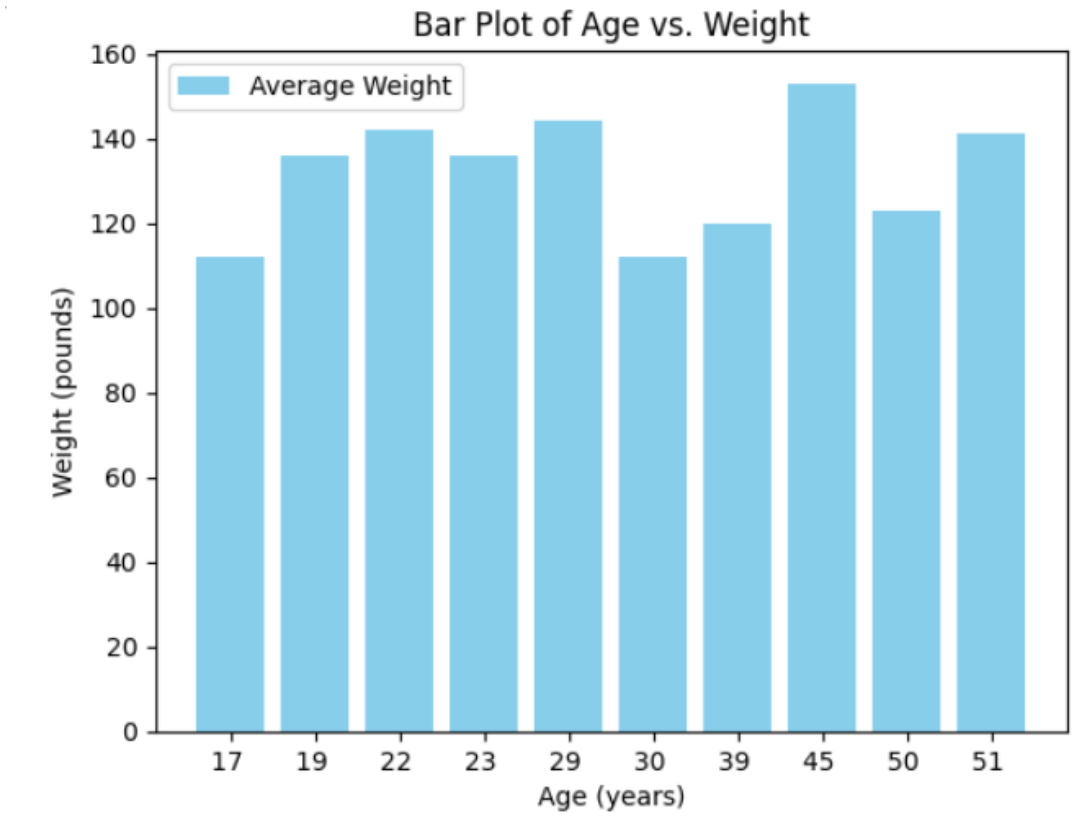
Plot 2:

The box plot visually represents the relationship between frailty and age. On the x-axis, different levels of frailty are displayed, while the y-axis represents the corresponding ages. The box plot provides a concise summary of the distribution of age within each frailty category, showcasing the median, quartiles, and potential outliers.



Plot3:

The bar graph visually represents the relationship between weight and age, where the y-axis corresponds to weight values and the x-axis represents age categories. Each bar on the graph represents the weight within a specific age group, providing a clear depiction of how weight varies across different age ranges. This visualization offers insights into potential trends or patterns in weight distribution across various age categories.



Plot 4:

The line graph depicts the relationship between age and frequency, with age represented on the y-axis and data points on the x-axis. Each point on the line corresponds to a specific data point, showcasing how the frequency of age values changes across the dataset. The upward or downward trend in the line indicates the varying distribution of ages, providing a visual insight into the prevalence of different age groups within the dataset.