Progress 1 - Screen shot of notebook has code and result

A screenshot of a computer

Description automatically generated

Progress 2 - Analysis from info()

1) How many data instances are there? 544

2) How many features are there? 4

3) What are the names? Height, weight, age, male

4) Are there any missing values? No

5) Are there any non-numeric features? No, all features are numeric.

Progress 3 - screenshot of code/result

A screenshot of a computer

Description automatically generated

Progress 4 - analysis

6) Are the data instances sorted on any of the attributes? No, it does not appear that data instances are sorted on any of the attributes.

7) What are the units of height? Centimeters

8) What are the units of weight? Kilograms

9) What are the minimum, median and max age?

Minimum age = 0

Median age = 27

Maximum age = 88

10) What two different features have the highest correlation? Height and weight have the highest correlation.

Progress 5 - Screen shot of graphs

A screenshot of a computer

Description automatically generated

Progress 6 - Analysis

11) Describe each distribution as skew, uniform, bimodal, or gaussian.

Height: Skew left

Weight: Bimodal

Age: Skewed right

12) Look at the age-weight scatter plot and explain the character of the graph.

As weight increases, it becomes much harder to predict the age. The correlation between the two decreases at a certain weight.

13) What does the age histogram tell us about this group of people?

Most people in the group are below the age of 50 and the average age is closer to 0 than 100.

14) How does the age histogram compare with that of people living in modern times?

I believe the age distribution in modern times would have a gaussian distribution.

Progress 7 - Code/Analysis  
1) Plot age v height  
2) Age at change in character of the data set

20 years of age is when the character of the data set changes.

A screenshot of a computer

Description automatically generated

Progress 8 - Screen shot of head

A screenshot of a computer

Description automatically generated

Progress 9 - Screen shot head and value counts

A screenshot of a computer program

Description automatically generated

Progress 10 - Screen shot of age v bmi plot

A screen shot of a computer

Description automatically generated

Progress 11 - Screen shot of data set print

A screenshot of a computer

Description automatically generated

Progress 12 - Screen shot of masked plot

A screen shot of a computer

Description automatically generated

Computation 13 - Ratios computed for

Compute the ratio of male/female for

1) Adults data frame: 164 male/ 182 female = .90

2) Training data frame: 135 male/ 141 female = .95

3) Test data frame: 29 male/ 41 female = .70

Computation 14 - Stratified Ratios computed for

Compute the ratio of male/female for

1) Adults data frame: 164 male/ 182 female = .90

2) Training data frame: 131 male/ 145 female = .90

3) Test data frame: 33 male/ 37 female = .89