**Stage1:Data Collection**

* The provided information is manually entered into a spreadsheet and stored in CSV file format with the name raw\_yield\_data.csv
* The data represents a 10 different persons with their age, weight, height, grip strength and the person has frailty or not(Y,N).

Table

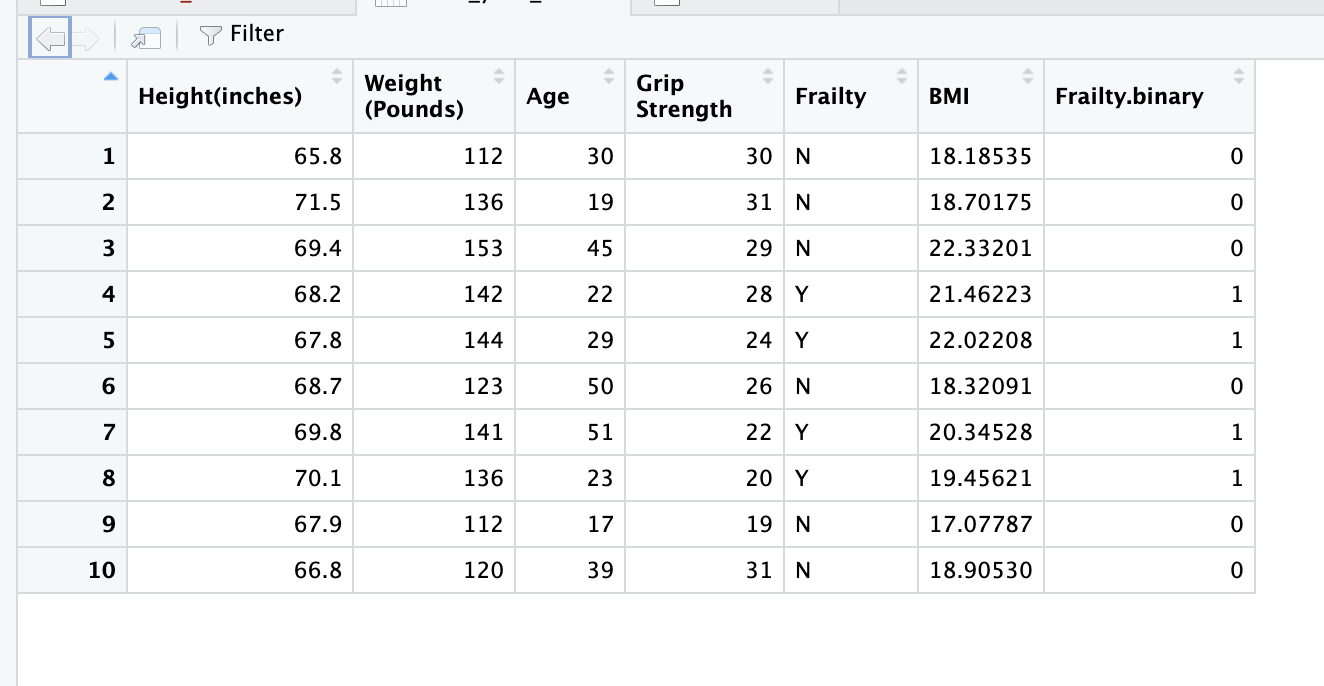
Description automatically generated

**Stage 2 : Data Processing**

* In the given raw data,there is no there is no need to remove or clean to proceed to further analysis.But just in case of large data I have omitted NA values using na.omit() function.
* Here i have added two more columns namely Body Mass Index(BMI) and Frailty. Binary for further analysis.
* BMI =(Weight/Height^2)\*703 and for Frailty.binary Y=1 and N=0.

Graphical user interface, text, application, email

Description automatically generated

****

**Stage3: Data Analysis**

* Here I have used plot, barplot, pairs, ggplot, ggplot2 for data visualization analysis.
* The below plot represents the line plot for BMI for the given 10 persons.

Chart, line chart

Description automatically generated

* Normally to be a healthy person one should have a BMI in between 18.5 to 24.9. If a person has BMI < 18.5 then the person is underweight, similarly if BMI > 24.9 then person is overweight and if BMI > 39.9 then the person is obese.
* lbl <- c(0,18.5,24.9,39.9,Inf)

cols <- c("darkred","darkgreen","red","green")[findInterval(clean\_yield\_data$BMI, vec=lbl)]

barplot(BMI,col = cols)

barplot (Frailty.binary)

Chart, bar chart

Description automatically generated

clean\_yield\_data %>%

ggplot2::ggplot(aes(x = factor(Age), y = BMI,fill=BMI > 18.5)) +

geom\_bar(position = "dodge",

stat = "identity")+

scale\_fill\_manual(values = c("#C00000","#004C99"),

labels=c('TRUE'='>18.5','FALSE'='<18.5'))+

labs(fill='BMI')

Chart, bar chart

Description automatically generated

clean\_yield\_data %>%

ggplot2::ggplot(aes(x = factor(`Grip Strength`), y = BMI,fill=BMI > 18.5)) +

geom\_bar(position = "dodge",

stat = "identity")+

scale\_fill\_manual(values = c("#C00000","#004C99"),

labels=c('TRUE'='Good','FALSE'='Underweight'))+

labs(fill='BMI')

Chart, bar chart

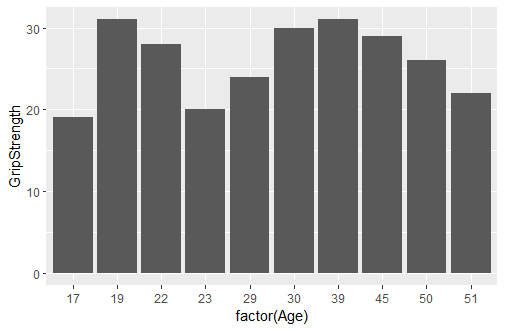
Description automatically generated

* So the blue color represents the person is in good condition. Whereas red color implicates that the person is underweight.

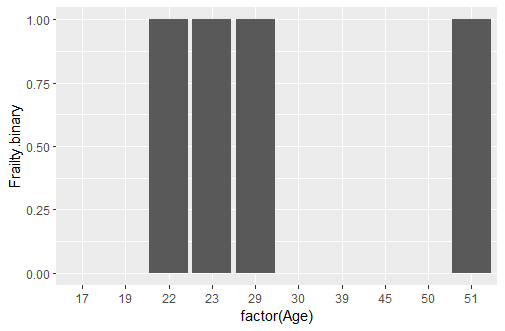
**barplot(Frailty.binary,** **col = "lightblue")** Chart, bar chart

Description automatically generated

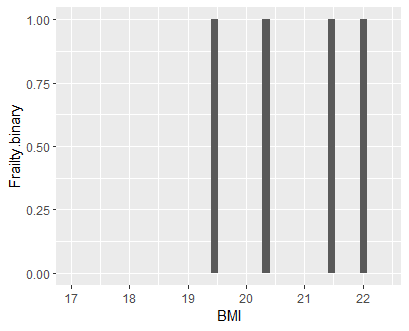
**ggplot(clean\_yield\_data,aes(x = factor(Age), y = `Grip Strength`)) + geom\_bar(stat = "identity")**



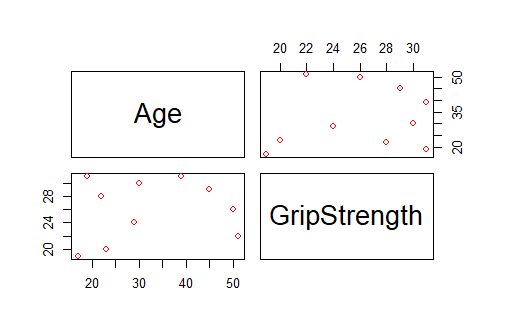
**ggplot(clean\_yield\_data, aes(x = factor(Age), y = Frailty.binary)) + geom\_bar(stat = "identity")**



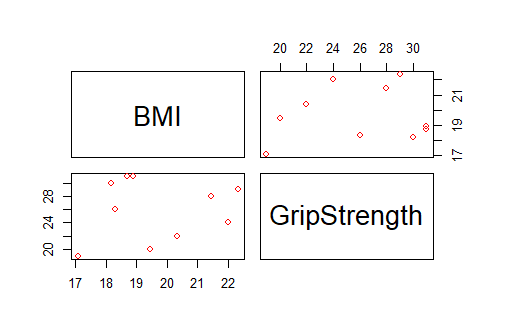
**ggplot(clean\_yield\_data, aes(x = BMI, y = Frailty.binary)) + geom\_bar(stat = "identity")**



**pairs(~Age+`Grip Strength`,data = clean\_yield\_data, col = "darkred")**



**pairs(~BMI+`Grip Strength`,data = clean\_yield\_data, col = "darkred")**



t\_test\_BMI\_Frailty <- with(clean\_yield\_data, t.test(clean\_yield\_data$BMI ~ clean\_yield\_data$Frailty))

capture.output(t\_test\_BMI\_Frailty, file = "Desktop/Spring'23/PDS/Assignment1/Question1\_project/results/test\_results\_BMI\_Frailty.txt")

Graphical user interface, text, application

Description automatically generated

t\_test\_Age\_Frailty <- with(clean\_yield\_data, t.test(clean\_yield\_data$Age ~ clean\_yield\_data$Frailty))

capture.output(t\_test\_Age\_Frailty, file = "Desktop/Spring'23/PDS/Assignment1/Question1\_project/results/test\_results\_Age\_Frailty.txt")

Text

Description automatically generated

t\_test\_GripStrength\_Age <- with(clean\_yield\_data, t.test(clean\_yield\_data$`Grip Strength`, clean\_yield\_data$Age))

capture.output(t\_test\_GripStrength\_Age, file = "Desktop/Spring'23/PDS/Assignment1/Question1\_project/results/test\_results\_GripStrength\_Age.txt")

Graphical user interface, text, application, email

Description automatically generated

t\_test\_BMI\_Age <- with(clean\_yield\_data, t.test(clean\_yield\_data$BMI, clean\_yield\_data$Age))

capture.output(t\_test\_BMI\_Age, file = "Desktop/Spring'23/PDS/Assignment1/Question1\_project/results/test\_results\_BMI\_Age.txt")

Graphical user interface, text, application

Description automatically generated