R\_P3\_2048046

**R Akhilandeshwari**

**1/29/2021**

setwd("C:/Users/akhil/OneDrive/Desktop/R/P3")  
getwd()

## [1] "C:/Users/akhil/OneDrive/Desktop/R/P3"

1. **Import the LungCapData into R and attach it**

library("readxl")  
data<-read\_excel("Lungcap.xlsx")  
data

## # A tibble: 725 x 6  
## LungCap Age Height Smoke Gender Caesarean  
## <dbl> <dbl> <dbl> <chr> <chr> <chr>   
## 1 6.48 6 62.1 no male no   
## 2 10.1 18 74.7 yes female no   
## 3 9.55 16 69.7 no female yes   
## 4 11.1 14 71 no male no   
## 5 4.8 5 56.9 no male no   
## 6 6.22 11 58.7 no female no   
## 7 4.95 8 63.3 no male yes   
## 8 7.32 11 70.4 no male no   
## 9 8.88 15 70.5 no male no   
## 10 6.8 11 59.2 no male no   
## # ... with 715 more rows

1. **Find the class and typeof Age and Height**

typeof(data$Age)

## [1] "double"

class(data$Height)

## [1] "numeric"

1. **Find the summary of the dataset**

summary(data)

## LungCap Age Height Smoke   
## Min. : 0.507 Min. : 3.00 Min. :45.30 Length:725   
## 1st Qu.: 6.150 1st Qu.: 9.00 1st Qu.:59.90 Class :character   
## Median : 8.000 Median :13.00 Median :65.40 Mode :character   
## Mean : 7.863 Mean :12.33 Mean :64.84   
## 3rd Qu.: 9.800 3rd Qu.:15.00 3rd Qu.:70.30   
## Max. :14.675 Max. :19.00 Max. :81.80   
## Gender Caesarean   
## Length:725 Length:725   
## Class :character Class :character   
## Mode :character Mode :character   
##

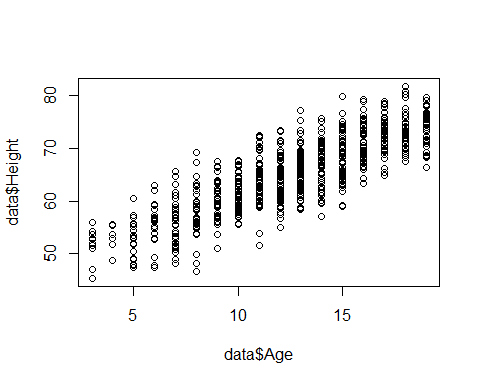
**4.Find the strength of the relationship between Age and Height (Hint:Using Correlation)**

cor(data$Age,data$Height)

## [1] 0.8357368

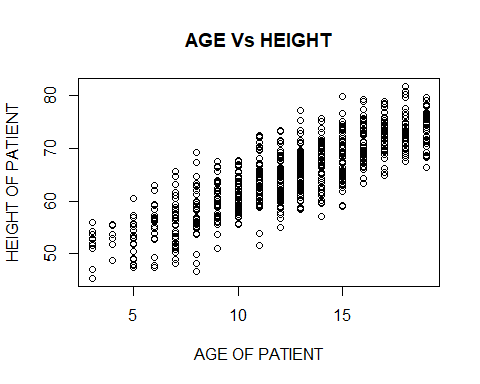
**5.Perform a scatterplot for the above (Qn.4)**

plot(data$Age,data$Height)



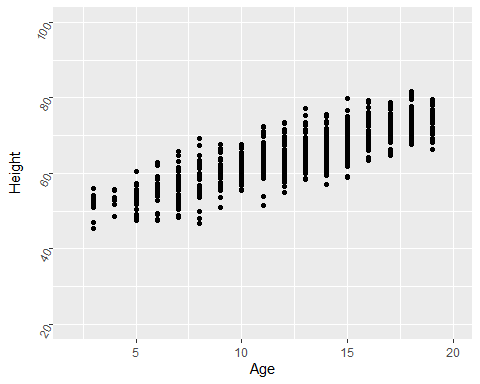
**6. Add a title to the plot and Label x-axis and y-axis**

plot(data$Age,data$Height,xlab="AGE OF PATIENT",ylab="HEIGHT OF PATIENT",main="AGE Vs HEIGHT")



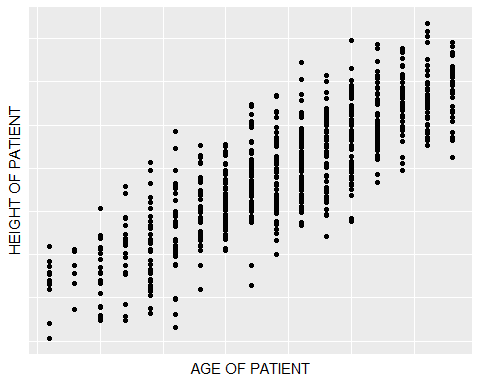
**7.Rotate the values on the y-axis and also change the limits of x-axis and y-axis**

library(ggplot2)  
pl<-ggplot(data,aes(Age, y=Height)) + geom\_point()  
pl+theme(axis.text.y = element\_text(angle=60))+scale\_x\_continuous(limits = c(2, 20))+scale\_y\_continuous(limits = c(20, 100))



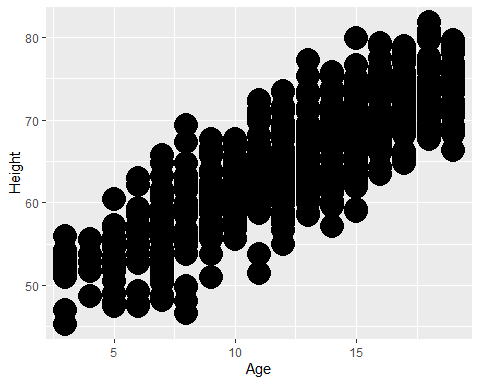
**8. Remove the axes of the plot and relabel these axes**

pl+theme(axis.text.x = element\_blank(),  
 axis.ticks.x = element\_blank(),  
 axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank())+labs(x="AGE OF PATIENT",y="HEIGHT OF PATIENT")



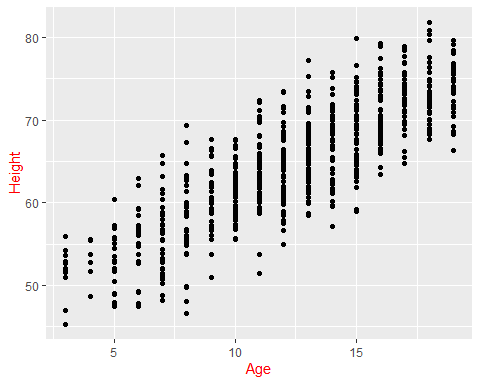
**9.Change the size of the plotting characters**

pl+geom\_point(size=8)



**10.Change the color of the characters to red**

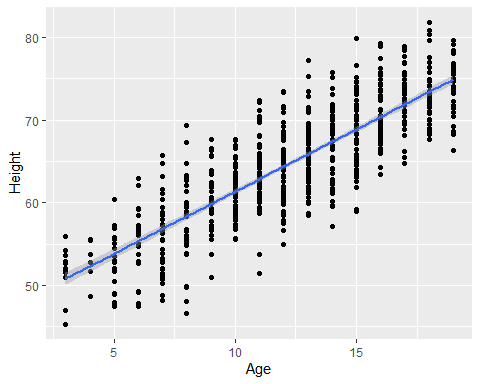
pl+theme(axis.title.x = element\_text(colour = "red"),  
 axis.title.y = element\_text(colour = "red"))



**11.Draw the regression line to the plot predicting height using age**

pl+geom\_smooth(method=lm)

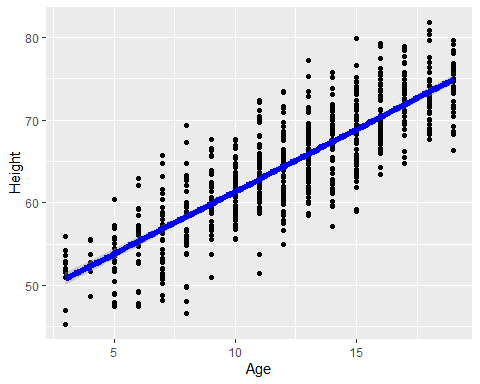
## `geom\_smooth()` using formula 'y ~ x'



**12.Change the color of the line to blue and the width of the line**

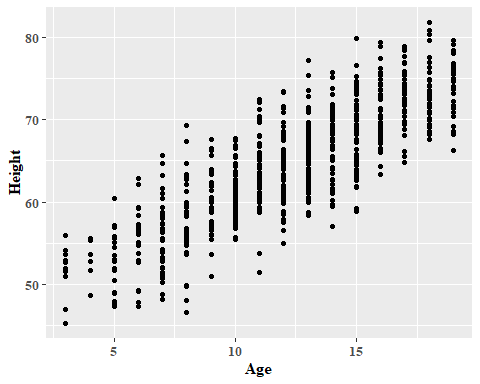
pl+geom\_smooth(method=lm,size=2.3,color="blue")

## `geom\_smooth()` using formula 'y ~ x'



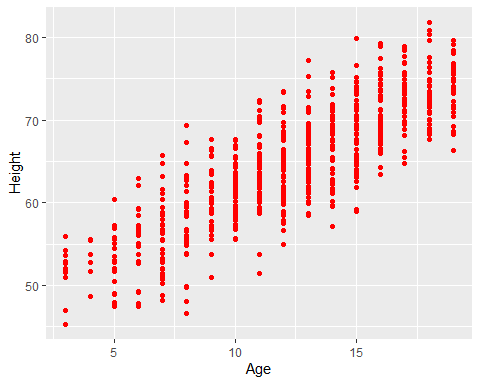
**13.Change the font of the plotting characters**

windowsFonts(A = windowsFont("Times New Roman"))   
pl+theme(text=element\_text(family="A", face="bold", size=12))



**14.Change the color of the plotting characters**

pl+geom\_point(color="red")



**15.Identify gender on the plot for the age male and female**

pl + geom\_point(aes(shape=factor(Gender),colour=factor(Gender)))

