5. Problem Statement Use the below data find the relationship between age (in months) and height (Average height in cm)

• age=16:27

• height=c(61.1,61.2,61.8,62.8,63.5,76.1,77,78.1,78.2,78.8,79.7,79.9)

1) check that age and height have the same number of elements

2) Create a scatterplot to determine the relationship between age and height

3) Create a "linear model" to fits the data above

4) Find the equation of the line of best fit

Answer 1)

Age <- 16:27

Height <- c(61.1,61.2,61.8,62.8,63.5,76.1,77,78.1,78.2,78.8,79.7,79.9)

A <- length(Age)

B <- length(Height)

if (A-B == 0){

print("Length of the vectors Age and Height is equal")

}else{

print("Length of the vectors Age and Height is not equal")

}

OUTPUT:-

[1] "Length of the vectors Age and Height is equal"

Answer 2) Age <- 16:27

Height <- c(61.1,61.2,61.8,62.8,63.5,76.1,77,78.1,78.2,78.8,79.7,79.9)

summary(Height)

summary(Age)

# help(plot)

# ?plot

cor(Age, Height)

plot(Age, Height)

plot(Age, Height, main="Scatter plot", xlab = "Age", ylab = "Hegiht")

plot(Age, Height, main="Scatter plot", xlab = "Age", ylab = "Hegiht", las = 1, xlim = c(0, 25), cex=0.5)

# Create a "linear model" to fits the data above # Answer no 3) starts from this line

abline(lm(Height~Age), col = 4)

lines(smooth.spline(Age, Height))

lines(smooth.spline(Age, Height), lty=2, lwd=5)

Answer 4) Finding the equation of the line of best fit

Step:1 Finding the gradient of the line

Using sxy and sxx

Find the eaquation of the regression line: y=bx+a (b is gradient and a is y intercept)

So gradient b= sxy/sxx=143/307=0.5

Sxy = ∑xy-∑x∑y/n=18758.3-258\*858.2/12=18758.3-18451.3=307

Sxx=∑x\*x-∑x\*∑x/n=5690-258\*258/12=5690-5547=143

Syy=∑y\*y-∑y\*∑y/n=62154.58-736507.24/12=62154.58-61375.60=779

Step 2: finding the value of the intercept

Find mean of x and y

Mean of x=71.5 and Y= 21.5

Step3: Finding the intercept

Y=bx+a

Ῡ=bx(Same as ybar)+a

21.5=71.5\*.5+a

21.5=35.75+a

A= -14.25

Y=0.5\*x+(-14.25) – is the best equation of line to fit.