# Scores Based On Hours Studied

Maria Sanjai

2020-12-19

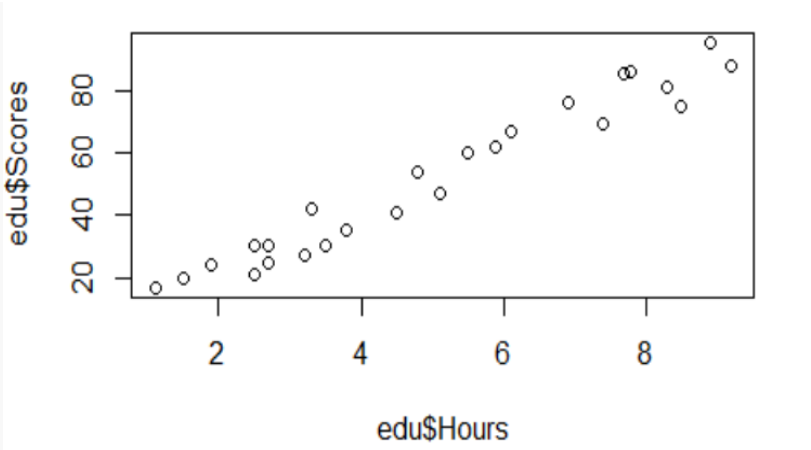
Question: Find the score when studied for 9.25 hours?

We can find the predicted value of scores after studying for 9.25 hours by using the linear regression model.

```
#Importing Dataset:
library(readxl)
edu <- read_excel("C:/Users/Maria sanjai/Desktop/edu.xlsx")
View(edu)
edu
```

```
## # A tibble: 25 x 2
##    Hours Scores
##    <dbl>  <dbl>
##  1   2.5     21
##  2   5.1     47
##  3   3.2     27
##  4   8.5     75
##  5   3.5     30
##  6   1.5     20
##  7   9.2     88
##  8   5.5     60
##  9   8.3     81
## 10   2.7     25
## # ... with 15 more rows
```

```
#plotting scatter plot
plot(edu$Hours,edu$Scores)
```



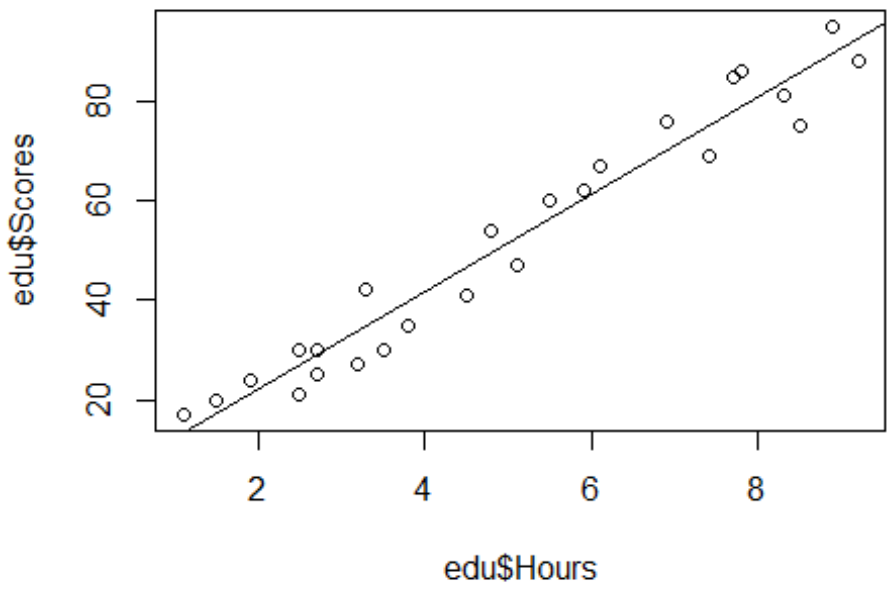

```r
#Finding the correlation:
cor(edu$Scores,edu$Hours)
```

```
## [1] 0.9761907
```

```r
#Simple linear regression model:
r<-lm(Scores~Hours,data=edu)
r
```

```
##
## Call:
## lm(formula = Scores ~ Hours, data = edu)
##
## Coefficients:
## (Intercept)        Hours
##       2.484        9.776
```

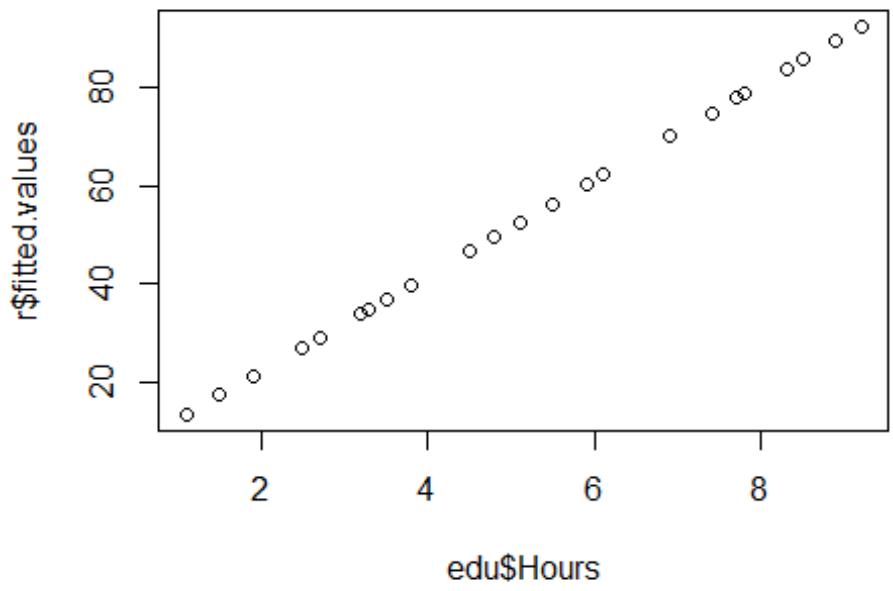
```r
#construction of regression line:
abline(r)
```




```r
#To find the summary of linear regression
summary(r)
```

```
##
## Call:
## lm(formula = Scores ~ Hours, data = edu)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.578  -5.340   1.839   4.593   7.265
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4837     2.5317   0.981    0.337
## Hours         9.7758     0.4529  21.583   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.603 on 23 degrees of freedom
## Multiple R-squared:  0.9529, Adjusted R-squared:  0.9509
## F-statistic: 465.8 on 1 and 23 DF,  p-value: < 2.2e-16
```

```r
#To plot the fitted line:
plot(edu$Hours,r$fitted.values)
```




```
#To predict a value of scores based on the hours studied:
#we can do it two ways 1)Manually using slope intercept formula OR 2)Using pr
edict command.

#1)Manually using slope intercept formula:
intercept=coef(r)[1]
intercept

## (Intercept)
##    2.483673

slope=coef(r)[2]
slope

##    Hours
## 9.775803

hours1=2.5
predicted_value1= intercept+slope*hours1
predicted_value1

## (Intercept)
##    26.92318

hours2=5.1
predicted_value2= intercept+slope*hours2
predicted_value2
```

```
## (Intercept)
##    52.34027

hours3=9.25
predicted_value3= intercept+slope*hours3
predicted_value3

## (Intercept)
##    92.90985

#Using predict command:
hours=data.frame(Hours=c(2.5,5.1,9.25))
predict(r,hours)

##        1        2        3
## 26.92318 52.34027 92.90985

predict(r,list(Hours=9.25))

##        1
## 92.90985
```