Question 01

01.

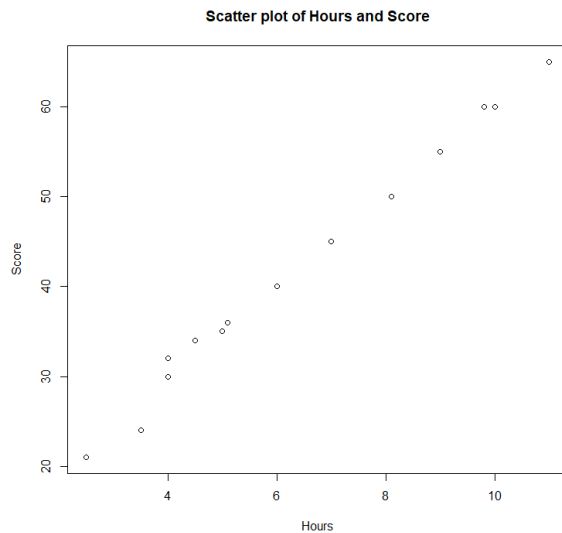Independent variable – hours

Dependent variable – score

02.

```
> exam=read.csv(file.choose())
> attach(exam)
> cor(Hours,Score)
[1] 0.9948127
```

Since the correlation coefficient is positive and close to +1, Hours and Score has a positive strong relationship.

03.

```
> plot(Hours,Score,ylab="Score",xlab="Hours",main="Scatter plot of Hours and Score")
```


Scatter plot of Hours and Score

According to the scatter plot there exists a strong positive relationship between Hours and Score.

04.

> lm(Score~Hours)

Call:

lm(formula = Score ~ Hours)

Coefficients:

(Intercept)   Hours

  8.588   5.185

Therefore,

The fitted simple linear regression model is $\hat{y} = 8.588x + 5.185$

05.

Hypothesis

$$H_0 : \hat{\beta}_1 = 0$$

$$H_1 : \hat{\beta}_1 \neq 0$$

> summary(model1)

Call:

lm(formula = Score ~ Hours)

Residuals:

   Min   1Q Median   3Q   Max
-2.7344 -0.5670 0.1188 0.6373 2.6732

Coefficients:

         Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.588   0.992   8.657 9.33e-07 ***
Hours   5.185   0.147 35.261 2.72e-14 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.514 on 13 degrees of freedom

Multiple R-squared: 0.9897,   Adjusted R-squared: 0.9889

F-statistic: 1243 on 1 and 13 DF, p-value: 2.723e-14

Since p value of hours < 0.05, we can reject the null hypothesis.

This means at 5% level of significance there in enough evidence to conclude that the model is significant.

06.

Intercept-8.588

This is the estimated value for when hours is 0. That means if no hours studied, there will be an 8.588 score approximately.

Slope- 5.185

This is the estimated change when score increases by one unit. That every additional hour you studied there will be an increase of 5.185 of your score.

07.

```
> predict(model1,list(Hours=8.8))
      1
54.21382
```


08.

```
> par(mfrow=c(2,2))
> plot(model1)
```