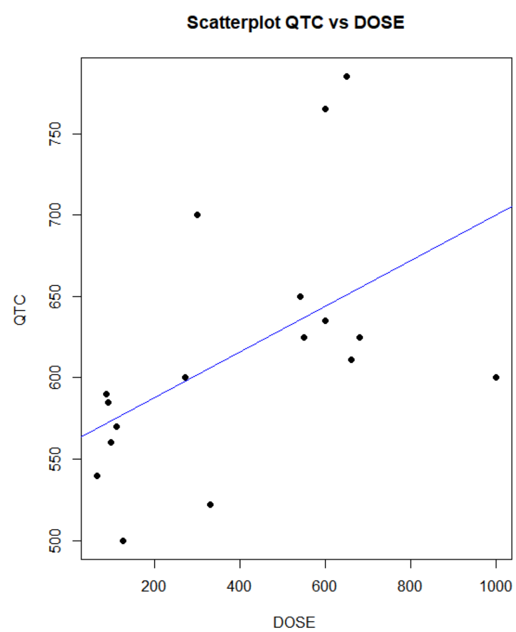
Ex 9.3.3, 9.4.1

```
> hw933 <- read.csv("C:/Users/EXR_C09_S03_03.csv")
> x <- hw933$DOSE
> y <- hw933$QTC
> plot(x, y, main="Scatterplot QTC vs DOSE", xlab="DOSE ", ylab="QTC ", pch=1
9)
> # regression line
> abline(lm(y ~ x, data = hw933), col = "blue")
> relation <- lm(y~x)
```




```
> print(relation)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
   559.9028       0.1399

> print(summary(relation))

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-99.789 -30.026  -8.835  14.559 134.171

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 559.90280   29.12926  19.221 5.61e-12 ***
x             0.13989    0.06033   2.319   0.0349 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 68.28 on 15 degrees of freedom
Multiple R-squared:  0.2639,   Adjusted R-squared:  0.2148
F-statistic: 5.377 on 1 and 15 DF,  p-value: 0.03493
```

- The regression line is: **y=559.9028+0.1399x**
- The coefficient of determination is R-squared = 0.2639
- Hypothesis testing: H0: β_1=0 vs. H1: $\beta_1 \neq 0$

  t-statistic = 2.319

  F-statistics = 5.377

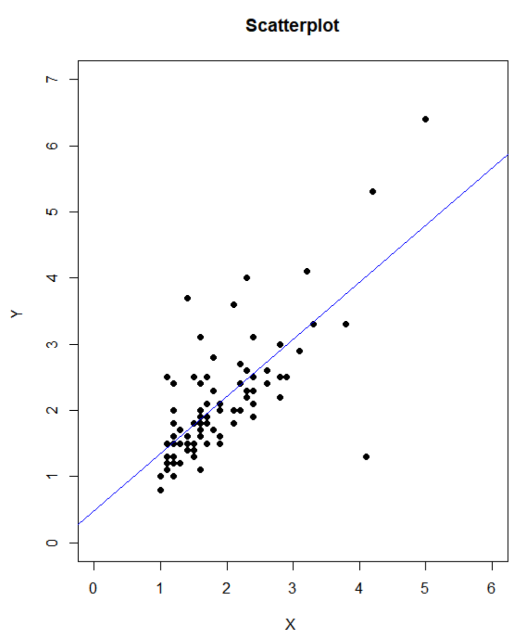
  We reject H0 at significance level 0.05 since the p-value is 0.03493

- 95% confidence interval of β_1 is below (x line)

```
> confint(relation)
                    2.5 %      97.5 %
(Intercept) 497.81525397 621.9903485
x             0.01129888   0.2684738
```

Ex 9.3.4, 9.4.2

```
> hw934 <- read.csv("C:/Users/EXR_C09_S03_04.csv")
> x <- hw934$X
> y <- hw934$Y
> plot(x, y, main="Scatterplot",xlab="X ", ylab="Y ", ylim=c(0,7), xlim=c(0,6
), pch=19)
> # regression line
> abline(lm(y ~ x, data = hw934), col = "blue")
```



```
> relation <- lm(y~x)
> print(summary(relation))

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.7248 -0.3357 -0.1341  0.1306  2.0040 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.48848    0.18167   2.689  0.00858 ** 
x            0.86251    0.08972   9.613 2.24e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.64 on 88 degrees of freedom
Multiple R-squared:  0.5122,   Adjusted R-squared:  0.5067 
F-statistic: 92.41 on 1 and 88 DF,  p-value: 2.244e-15
```

- The regression line is: **y**=0.48848+0.8625x
- The coefficient of determination is R-squared = 0.5122
- Hypothesis testing: H0: $\beta_1=0$ vs. H1: $\beta_1\neq0$

  t-statistic = 9.613

  F-statistics = 92.41

  We reject H0 at significance level 0.05 since the p-value is 2.24e-15<0.05

- 95% confidence interval of β_1 is below (x line)

```
> confint(relation)
                2.5 %    97.5 %
(Intercept) 0.1274485 0.8495109
x           0.6842054 1.0408162
```

QUIZ 5

The following scores represent a nurse's assessment (X) and a physician's assessment (Y) of the condition of 9 patients at time of admission to a trauma center.

X: 18, 13, 18, 15, 10, 12, 8, 4, 7

Y: 22, 19, 20, 16, 12, 15, 10, 7, 6

1) Write the equation of the regression line (round solution to 3 decimal digits).

2) Write the expression of the coefficient of determination end explain its significance (round solution to 3 decimal digits).

3) Test the hypothesis

H_0: \beta_1 =0   v.s. \beta_1 \ne 0

at level of significance 0.01 using the t statistics.

4) Compute the 95% confidence interval of \beta_1 in the form [a,b] (round solution to 3 decimal digits).

Quiz5 solution

```
> y <-c(22, 19, 20, 16, 12, 15, 10, 7, 6)
> x <-c(18, 13, 18, 15, 10, 12, 8, 4, 7)
> plot(x, y, main="Scatterplot")
> abline(lm(y ~ x), col = "blue")
> relation <- lm(y~x)
> print(summary(relation))


Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.9205 -1.1556 -0.0327  0.8444  3.4058 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   1.1345     1.8137   0.626 0.551469    
x             1.1123     0.1446   7.690 0.000117 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.994 on 7 degrees of freedom
Multiple R-squared:  0.8941,	Adjusted R-squared:  0.879 
F-statistic: 59.13 on 1 and 7 DF,  p-value: 0.0001172
```

1) The regression line is: **y=** 1.135 +1.112 x
2) The coefficient of determination is R-squared = 0.894. This indicates that 89.4% of the variability of y is explained by the variability of x
3) Hypothesis testing: H0: $\beta_1=0$ vs. H1: $\beta_1\neq 0$

   t-statistic = 7.690

   We reject H0 at significance level 0.01 since the p-value is 0.0001172 <0.01

4) 99% confidence interval of β_1 is CI = $\beta_1 \pm t(\alpha/2,r)\, s_{\beta 1}$

   ```
   > qt(1-0.01/2,9-2) = 3.499483
   ```

   Hence CI = 1.112 ± (3.499)(0.145) = [0.605,1.619]

   Alternatively:

```
> confint(relation,level=0.99)
                 0.5 %   99.5 %
```

```
(Intercept) -5.2124823 7.481488
x            0.6060942 1.618467
```