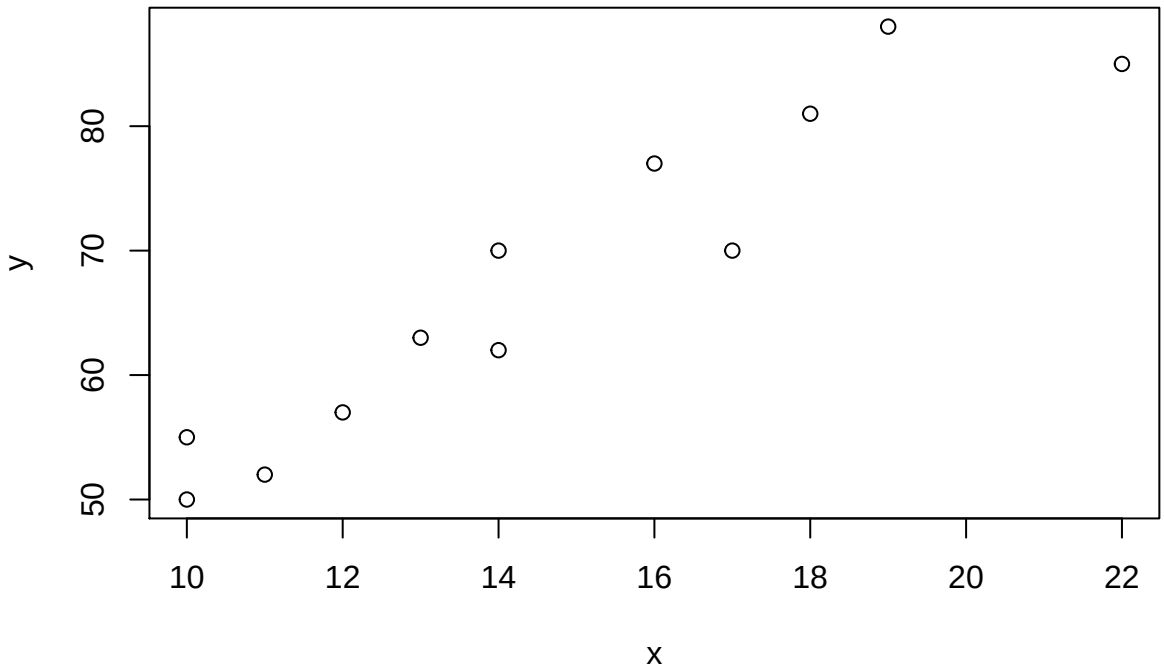
Weighted-Average-01

SP

15/10/2019

```
x = c(16,14,22,10,14,17,10,13,19,12,18,11) # total number of reponses in completing a lesson
y = c(77,70,85,50,62,70,55,63,88,57,81,52) # cost of computer time in cents
```

```
library(MASS)
plot(x,y)
```



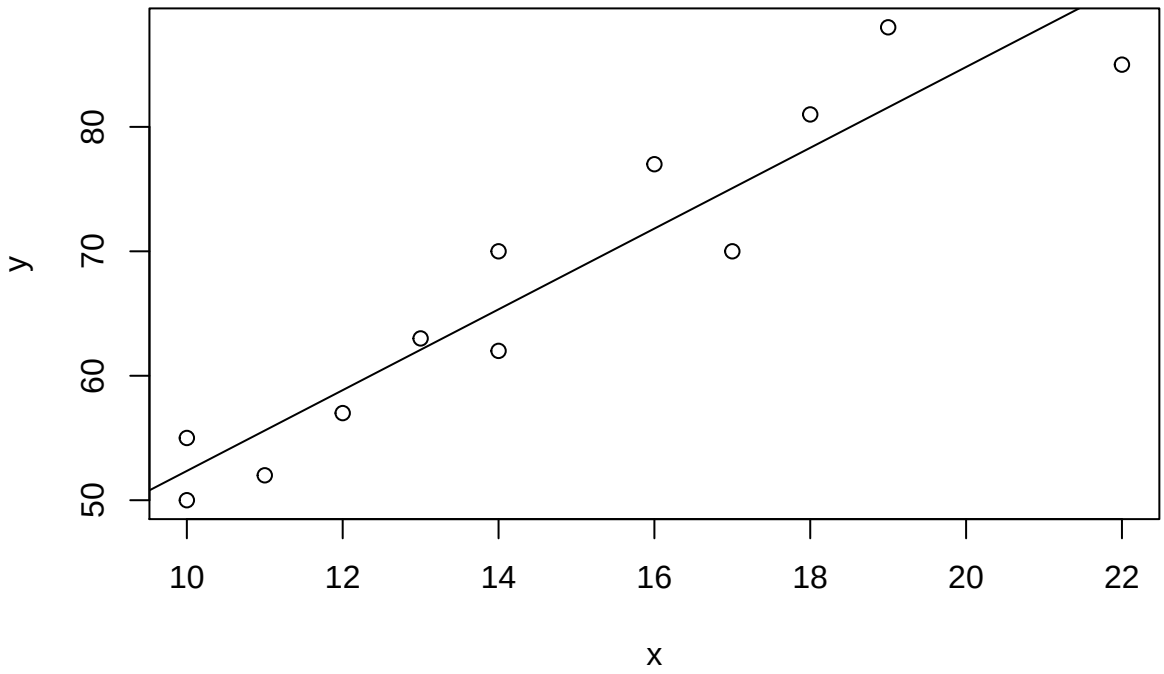

```
model1 = lm(y~x)
model1
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Coefficients:
## (Intercept)            x
##      19.895        3.246
```

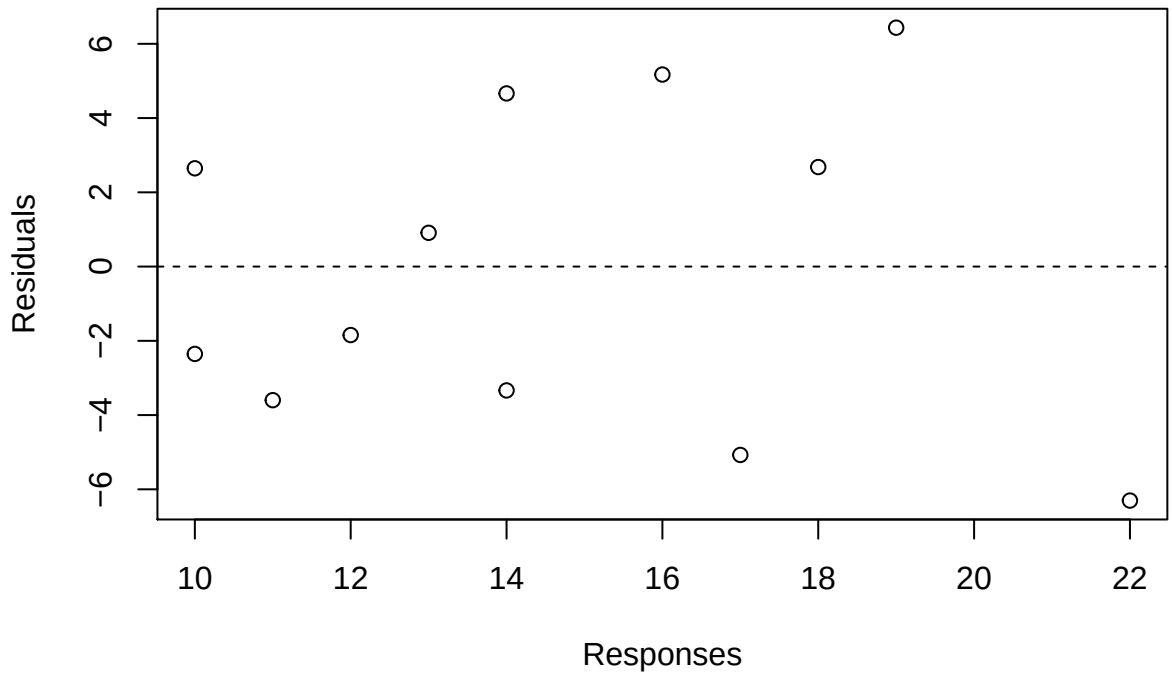
```r
summary(model1)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3025 -3.4018 -0.4674  3.1765  6.4349
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  19.8950     5.4104   3.677  0.00427 **
## x             3.2458     0.3581   9.065 3.88e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.51 on 10 degrees of freedom
## Multiple R-squared:  0.8915, Adjusted R-squared:  0.8807
## F-statistic: 82.18 on 1 and 10 DF,  p-value: 3.878e-06
```

```r
plot(x,y)
abline(model1)
```



```r
#diagnostic plot
plot(x, model1$residuals, xlab = "Responses", ylab = "Residuals")
abline(h=0,lty=2)
```

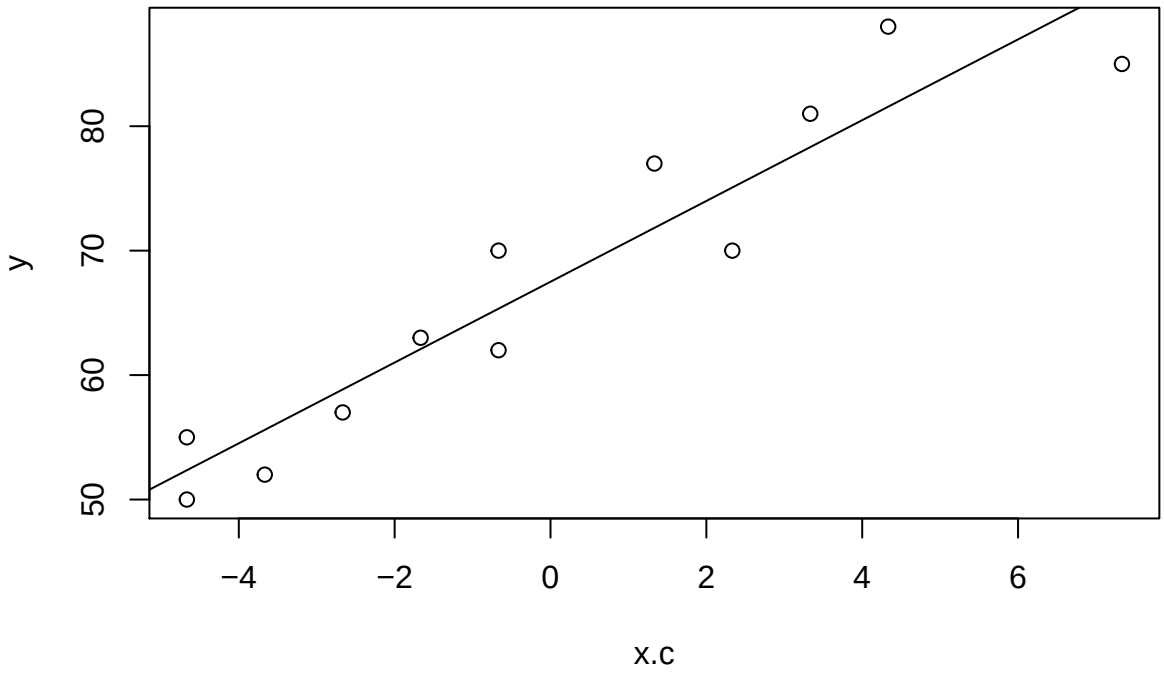



```r
#mean centrered OLS regression
x.c = x-mean(x)
olsfit.c=lm(y~x.c)
summary(olsfit.c)
```

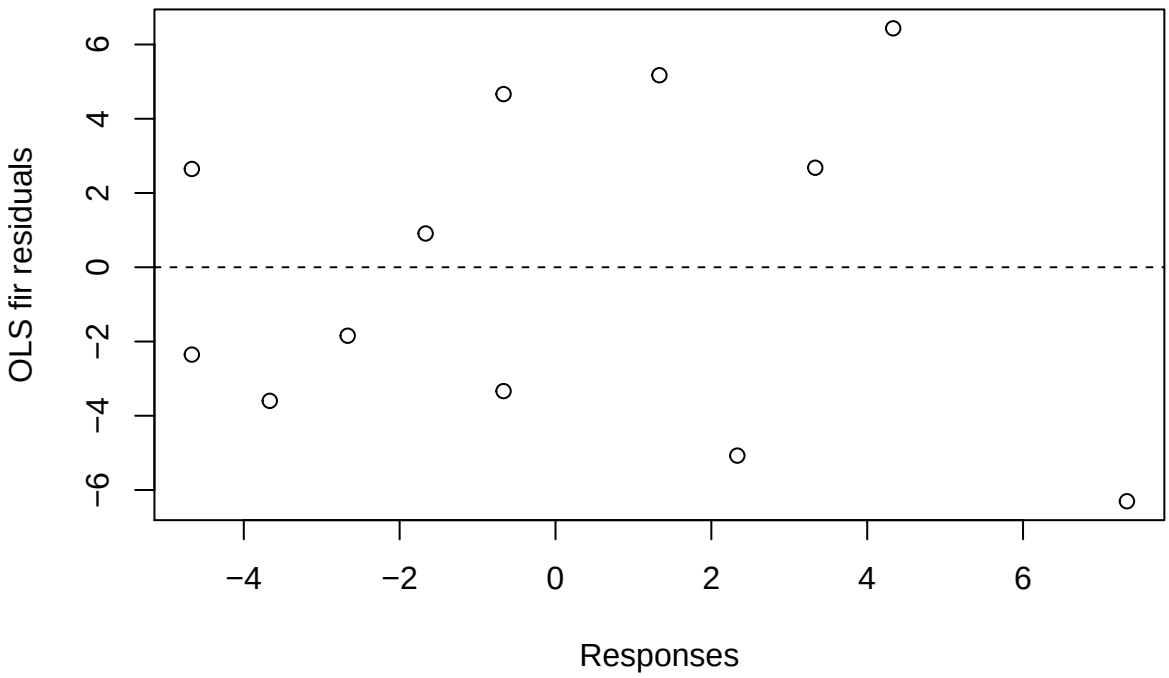
```
##
## Call:
## lm(formula = y ~ x.c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3025 -3.4018 -0.4674  3.1765  6.4349
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  67.5000     1.3020  51.845 1.72e-13 ***
## x.c           3.2458     0.3581   9.065 3.88e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.51 on 10 degrees of freedom
```

```
## Multiple R-squared:  0.8915, Adjusted R-squared:  0.8807
## F-statistic: 82.18 on 1 and 10 DF,  p-value: 3.878e-06
```

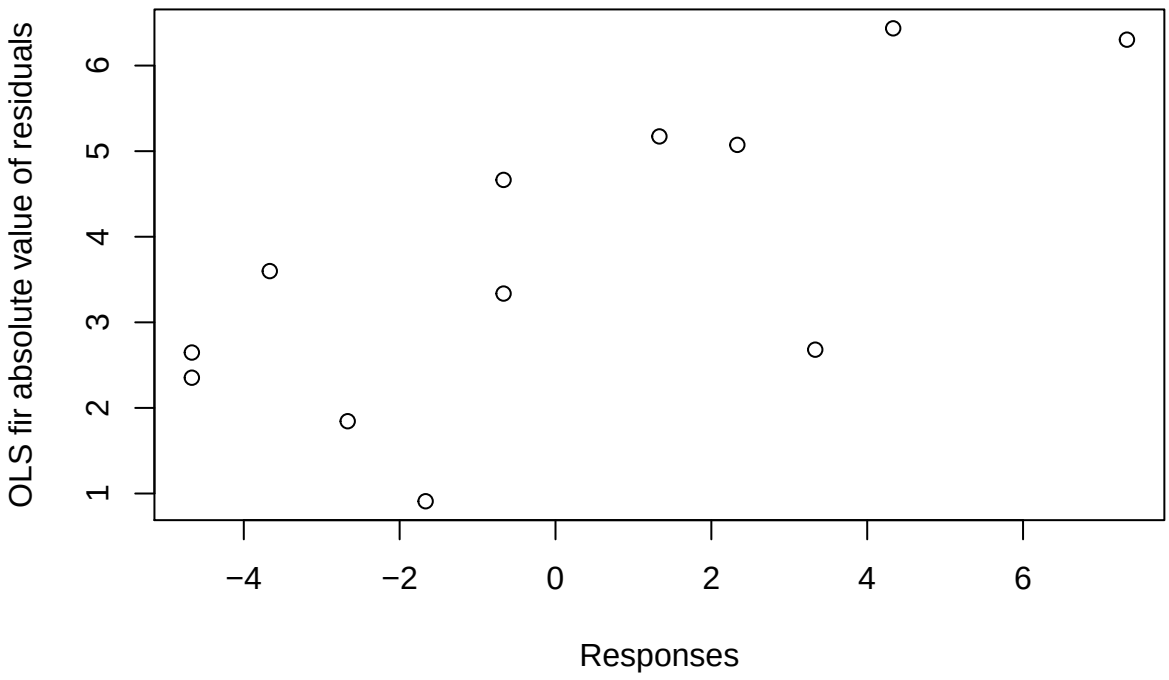
```r
plot(x.c,y)
abline(olsfit.c)
```




```r
#diagnostic plot
plot(x.c, olsfit.c$residuals, xlab = "Responses", ylab = "OLS fir residuals")
abline(h=0,lty=2)
```




```
plot(x.c,abs(olsfit.c$residuals), xlab = "Responses", ylab = "OLS fir absolute value of residuals")
```

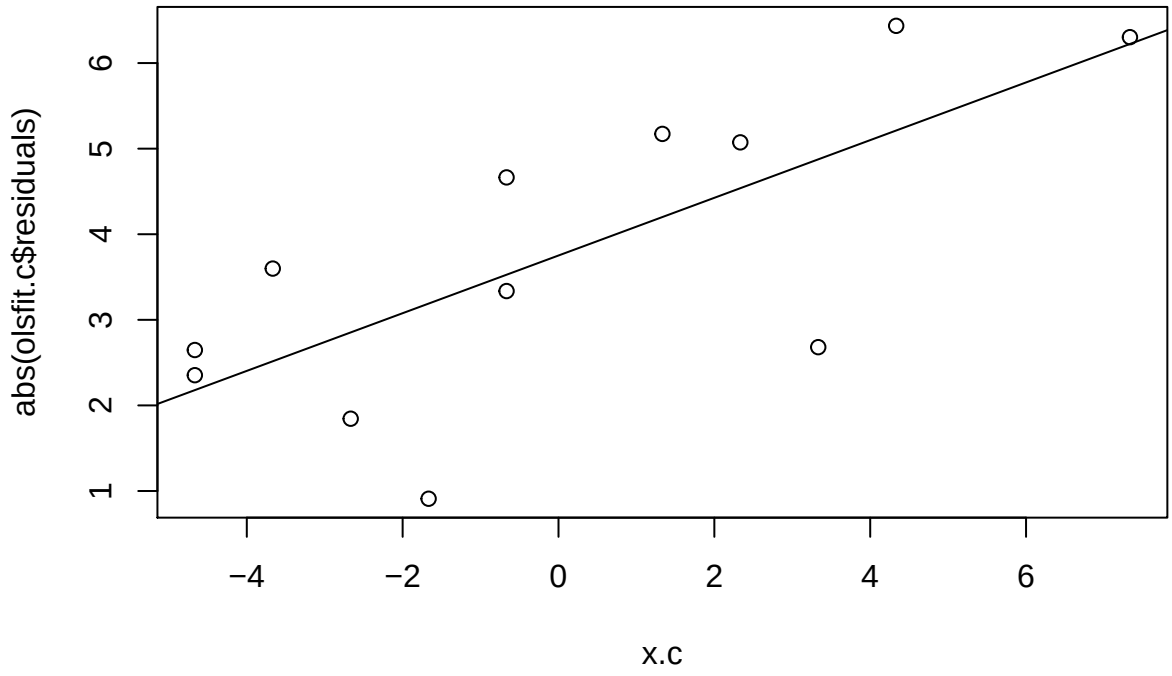



```
#we see a positive relationship between the residuals meaning there could ve non-constasnt variance
```

```
# Absolute value of residuals regression due to possible due possible non-constant variance of error te
sd.func=lm(abs(olsfit.c$residuals)~x.c)
summary(sd.func)
```

```
##
## Call:
## lm(formula = abs(olsfit.c$residuals) ~ x.c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2800 -0.3950  0.3214  0.9994  1.2230
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.751      0.371  10.112 1.43e-06 ***
## x.c            0.337      0.102   3.304  0.00797 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.285 on 10 degrees of freedom
## Multiple R-squared:  0.5218, Adjusted R-squared:  0.474
## F-statistic: 10.91 on 1 and 10 DF,  p-value: 0.007965
```

```r
plot(x.c,abs(olsfit.c$residuals))
abline(sd.func)
```




```r
wights = 1/((sd.func$fitted.values)^2)
wights
```

```
##          1          2          3          4          5          6
## 0.05666828 0.08040079 0.02582272 0.21069605 0.08040079 0.04856310
##          7          8          9         10         11         12
## 0.21069605 0.09828945 0.03681378 0.12288694 0.04208017 0.15802095
```

```r
# mean centered regression with wieghts inserted
wlsfit = lm(y~x.c, weights = wights)
summary(wlsfit)
```

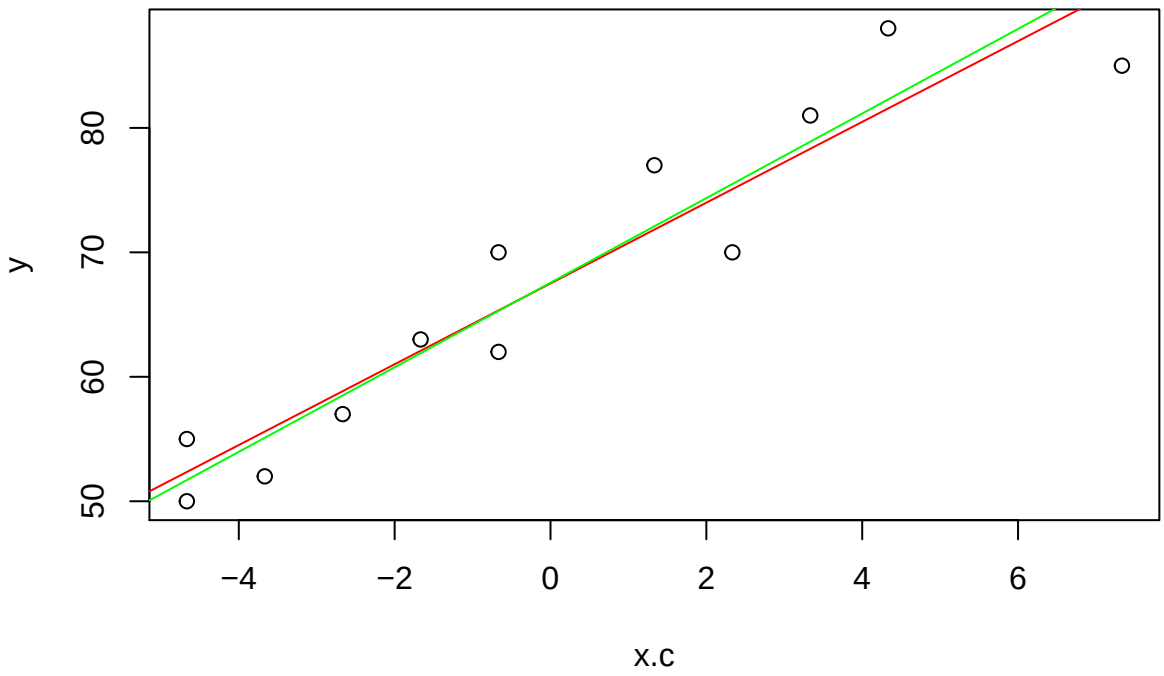
```
##
## Call:
## lm(formula = y ~ x.c, weights = wights)
##
## Weighted Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23265 -1.00195 -0.09015  1.11349  1.51379
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  67.5629     1.3000   51.97 1.68e-13 ***
## x.c           3.3987     0.3581    9.49 2.56e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.146 on 10 degrees of freedom
## Multiple R-squared:  0.9001, Adjusted R-squared:  0.8901
## F-statistic: 90.06 on 1 and 10 DF,  p-value: 2.563e-06
```

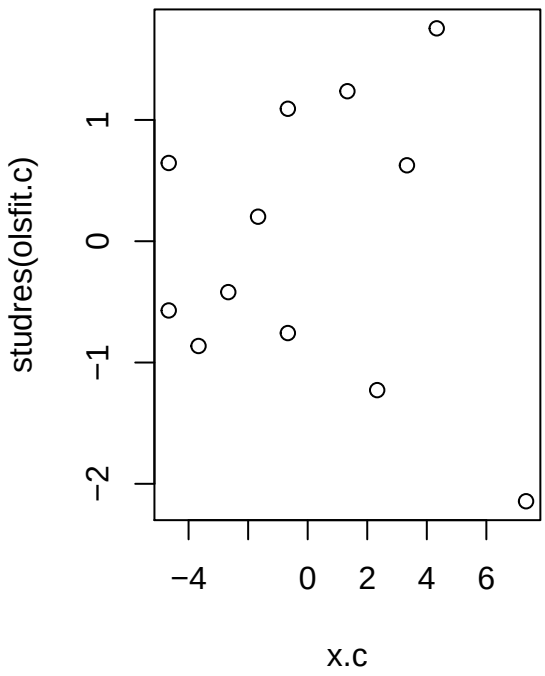
```
summary(olsfit.c)
```

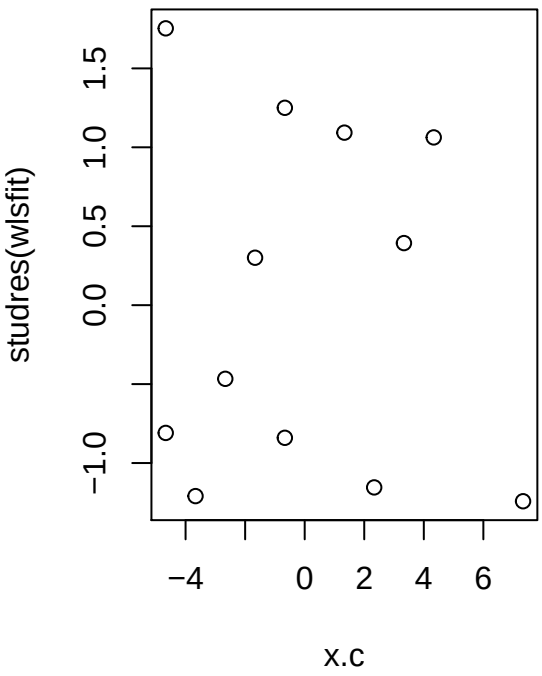
```
##
## Call:
## lm(formula = y ~ x.c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3025 -3.4018 -0.4674  3.1765  6.4349
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  67.5000     1.3020  51.845 1.72e-13 ***
## x.c           3.2458     0.3581   9.065 3.88e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.51 on 10 degrees of freedom
## Multiple R-squared:  0.8915, Adjusted R-squared:  0.8807
## F-statistic: 82.18 on 1 and 10 DF,  p-value: 3.878e-06
```

```
#model comparision
plot(x.c,y)
abline(olsfit.c, col = "red")
abline(wlsfit, col="green")
```




```
# comparing the residuals
par(mfrow=c(1,2))
plot(x.c,studres(olsfit.c))
plot(x.c,studres(wlsfit))
```


studres(olsfit.c)
−2
−1
0
1
−4
0
2
4
6
x.c


studres(wlsfit)
−1.0
0.0
0.5
1.0
1.5
−4
0
2
4
6
x.c