Flipped Assignment 9

Group 5

2022/2/24

Input Data

```
setwd('G:/OneDrive - Texas Tech University/IE 5344 Statistical Data Analysis/Flipped Assignment 9')
data <- read.csv('data-table-B8.csv', header = TRUE)
head(data)
```

```
##   x1  x2    y
## 1  0  10  7.5
## 2  0  50 15.0
## 3  0  85 22.0
## 4  0 110 28.6
## 5  0 140 31.6
## 6  0 170 34.0
```

Part a.

```
fit1 <- lm(y~x1+x2+x1:x2, data)
summary(fit1)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x1:x2, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0753 -3.6781  0.4395  3.1321  8.8448
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.50128    1.89347   6.602 1.92e-07 ***
## x1          256.73740   73.72914   3.482  0.00146 **
## x2            0.09879    0.01193   8.281 1.84e-09 ***
## x1:x2         0.76127    0.51026   1.492  0.14551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.696 on 32 degrees of freedom
## Multiple R-squared:  0.8518, Adjusted R-squared:  0.8379
## F-statistic: 61.31 on 3 and 32 DF,  p-value: 2.318e-13
```

```r
fit1$coefficients
```

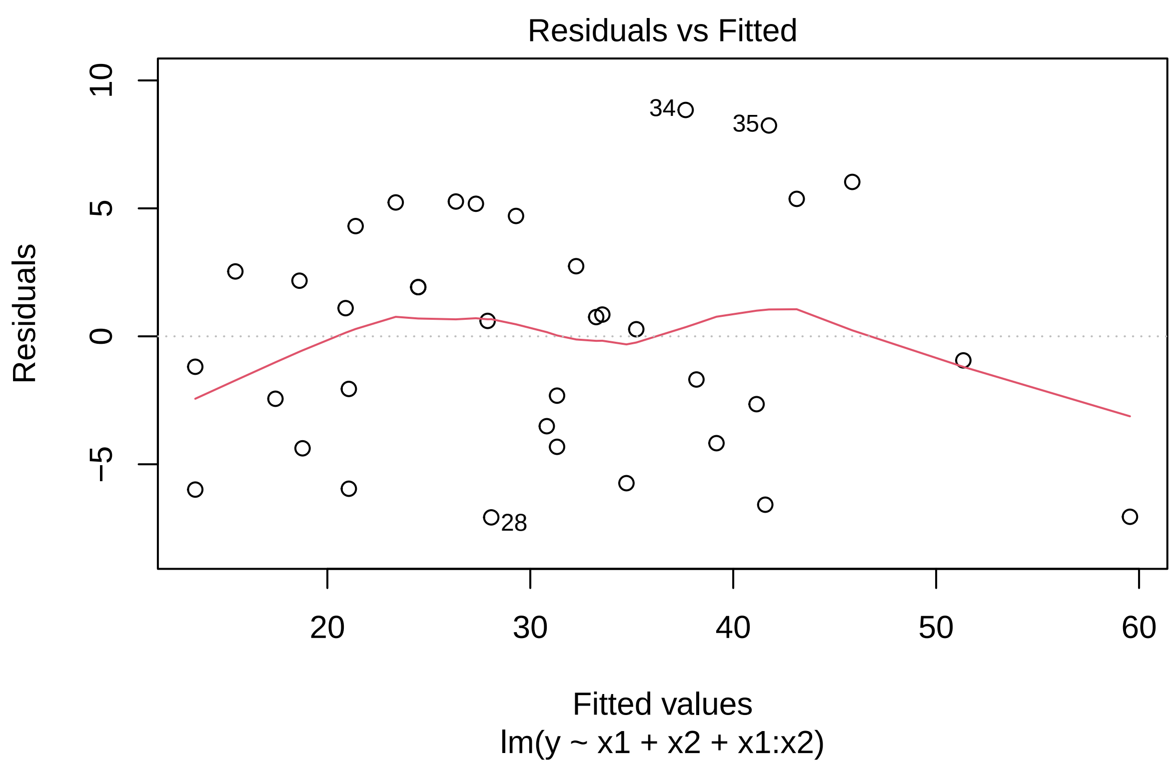
```
##   (Intercept)            x1            x2         x1:x2
##   12.50128449 256.73740096    0.09879204    0.76127041
```

So $\hat{y} = 12.5013 + 256.7374x_1 + 0.0988x_2 + 0.7613x_1x_2$.

Part b.

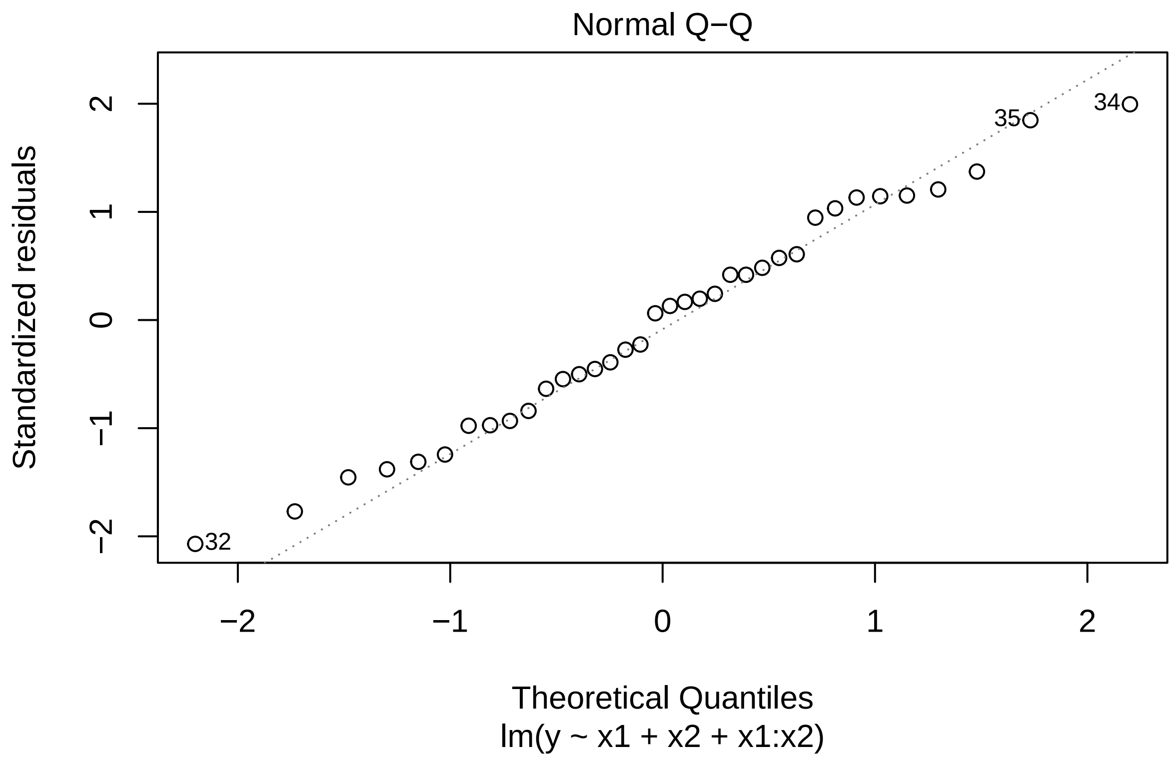
Check for Model Adequacy

```r
plot(fit1,1)
```



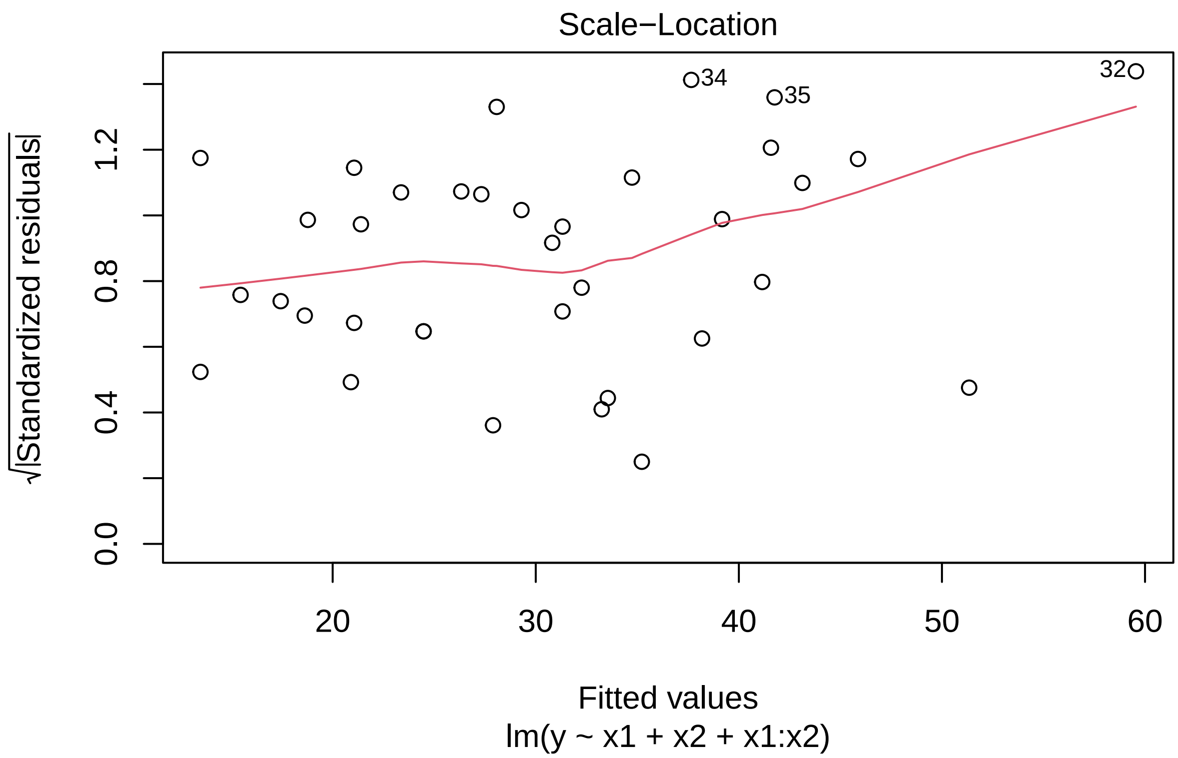
This figure ensures the constant variance.

```r
plot(fit1,2)
```



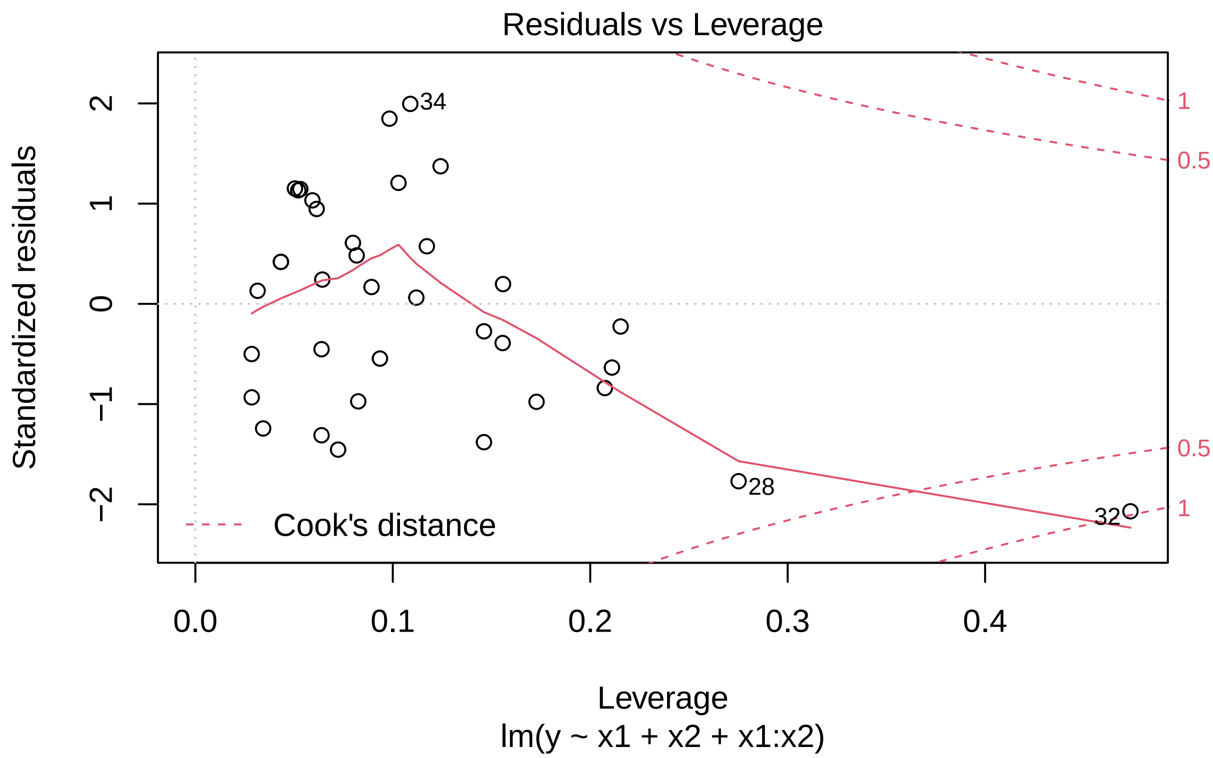
This figures ensures the normality.

```
plot(fit1,3)
```



We don't think there is any outliers.

```
plot(fit1,5)
```



We don't think there is any outliers. So this model has model adequacy and no transformation is needed.

Part c.

```
summary(fit1)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x1:x2, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0753 -3.6781  0.4395  3.1321  8.8448
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.50128    1.89347   6.602 1.92e-07 ***
## x1          256.73740   73.72914   3.482  0.00146 **
## x2            0.09879    0.01193   8.281 1.84e-09 ***
## x1:x2         0.76127    0.51026   1.492  0.14551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.696 on 32 degrees of freedom
## Multiple R-squared:  0.8518, Adjusted R-squared:  0.8379
```

```
## F-statistic: 61.31 on 3 and 32 DF,  p-value: 2.318e-13
```

Reject H_0 because $p - value < 0.05$. So we can conclude that least one of these regressors contributes significantly to the model.

Part d.

```r
fit2 <- lm(y~x1+x2, data)
anova(fit2,fit1)
```

```
## Analysis of Variance Table
##
## Model 1: y ~ x1 + x2
## Model 2: y ~ x1 + x2 + x1:x2
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     33 754.74
## 2     32 705.66  1    49.084 2.2259 0.1455
```

So we drop $x_1 x_2$ because we don't reject H_0 for $p - value > 0.05$.

```r
summary(fit2)
```

```
##
## Call:
## lm(formula = y ~ x1 + x2, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.7716 -4.1656  0.0802  3.8323  8.3349
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.109e+01  1.669e+00   6.642 1.48e-07 ***
## x1          3.501e+02  3.968e+01   8.823 3.38e-10 ***
## x2          1.089e-01  9.983e-03  10.912 1.74e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.782 on 33 degrees of freedom
## Multiple R-squared:  0.8415, Adjusted R-squared:  0.8319
## F-statistic:  87.6 on 2 and 33 DF,  p-value: 6.316e-14
```

So from the summary table, we keep x_1 and x_2 as these regressors are significant under level of 0.0001. So fit2 is the best model.