2010s Movies

2024-02-21

```
movies <- read.csv("movies.csv")

head(movies)
```

```
##   Number                          Film Month.of.Release Year.of.Release
## 1      1 Star Wars: The Force Awakens               12            2015
## 2      2            Avengers: Endgame                4            2019
## 3      3                Black Panther                2            2018
## 4      4       Avengers: Infinity War                4            2018
## 5      5                Jurassic World                6            2015
## 6      6        Marvel's The Avengers                5            2012
##   Domestic.Gross.in.Millions Tomatometer Audience.Score MPA.Rating
## 1                   936.6622        0.93           0.85          3
## 2                   858.3730        0.94           0.90          3
## 3                   700.4266        0.96           0.79          3
## 4                   678.8155        0.85           0.92          3
## 5                   653.4066        0.71           0.78          3
## 6                   623.3579        0.91           0.91          3
##   Percent.Inflation.for.Month.and.Year Budget.in.Millions
## 1                                  0.7                245
## 2                                  2.0                356
## 3                                  2.2                200
## 4                                  2.5                321
## 5                                  0.2                150
## 6                                  1.7                220
```

```
summary(movies)
```

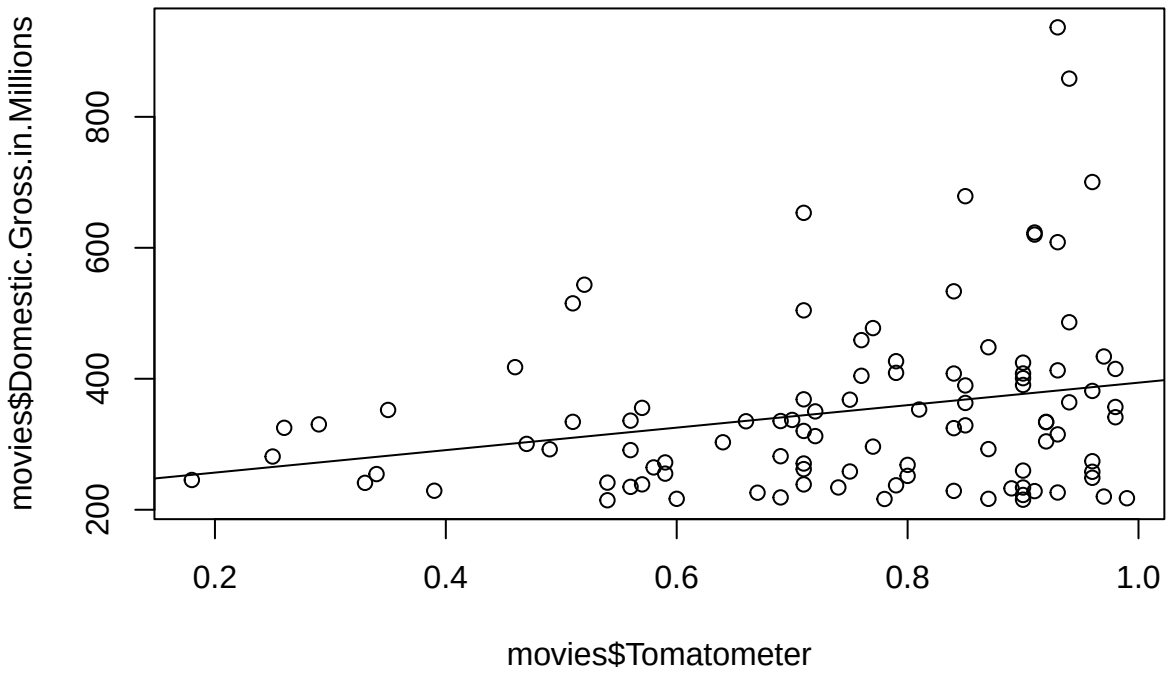
```
##      Number           Film           Month.of.Release Year.of.Release
##  Min.   :  1.00   Length:100         Min.   : 2.00    Min.   :2010
##  1st Qu.: 25.75   Class :character   1st Qu.: 5.00    1st Qu.:2013
##  Median : 50.50   Mode  :character   Median : 7.00    Median :2015
##  Mean   : 50.50                      Mean   : 7.22    Mean   :2015
##  3rd Qu.: 75.25                      3rd Qu.:11.00    3rd Qu.:2017
##  Max.   :100.00                      Max.   :12.00    Max.   :2019
##  Domestic.Gross.in.Millions  Tomatometer     Audience.Score     MPA.Rating
##  Min.   :214.5              Min.   :0.1800   Min.   :0.4200   Min.   :1.00
##  1st Qu.:251.0              1st Qu.:0.6300   1st Qu.:0.7175   1st Qu.:2.00
##  Median :324.8              Median :0.7900   Median :0.8350   Median :3.00
##  Mean   :350.6              Mean   :0.7471   Mean   :0.7805   Mean   :2.76
##  3rd Qu.:405.4              3rd Qu.:0.9025   3rd Qu.:0.8825   3rd Qu.:3.00
```

```
##  Max.   :936.7              Max.   :0.9900   Max.   :0.9500   Max.   :4.00
##  Percent.Inflation.for.Month.and.Year Budget.in.Millions
##  Min.   :-0.20                        Min.   : 35.0
##  1st Qu.: 1.20                        1st Qu.:117.5
##  Median : 1.75                        Median :167.5
##  Mean   : 1.75                        Mean   :163.8
##  3rd Qu.: 2.20                        3rd Qu.:200.0
##  Max.   : 3.60                        Max.   :356.0
```

```
gross_tomatometer_lm <- lm(movies$Domestic.Gross.in.Millions ~ movies$Tomatometer)

gross_budget_lm <- lm(movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)

gross_inflation_lm <- lm(movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.and.Yea

gross_audience_lm <- lm(movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)
```

Domestic Gross predicted by Tomatometer

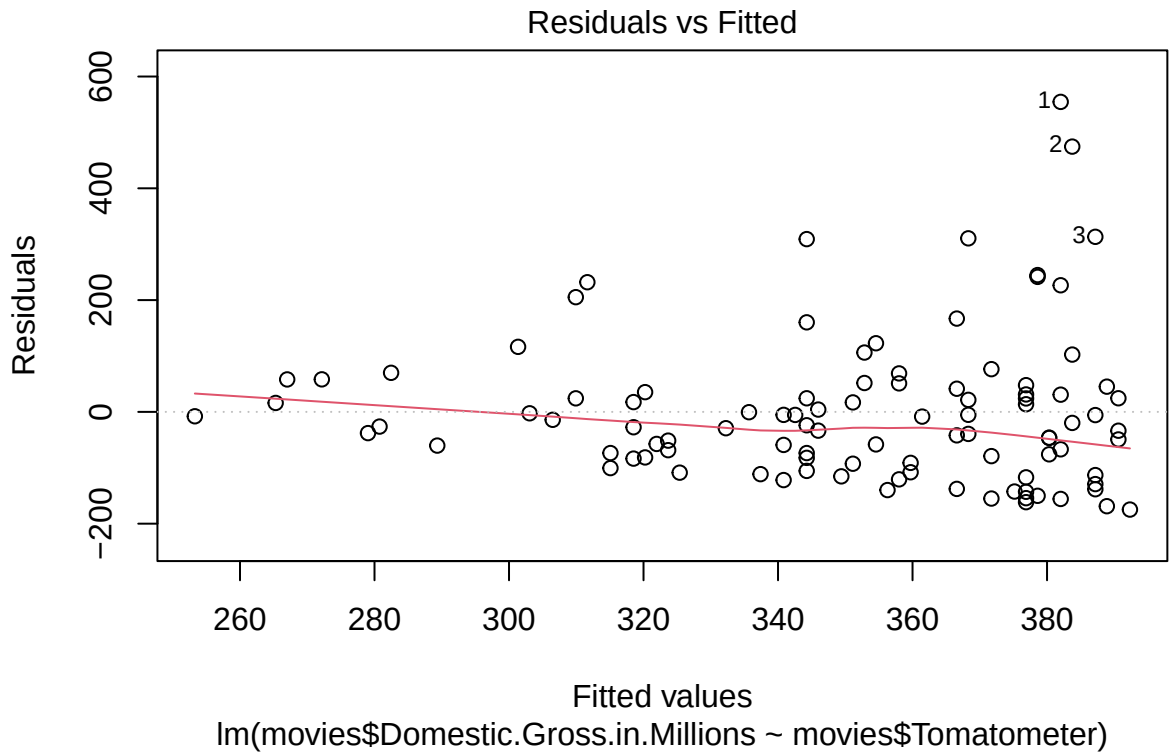
```
plot(movies$Domestic.Gross.in.Millions ~ movies$Tomatometer)

abline(gross_tomatometer_lm)
```

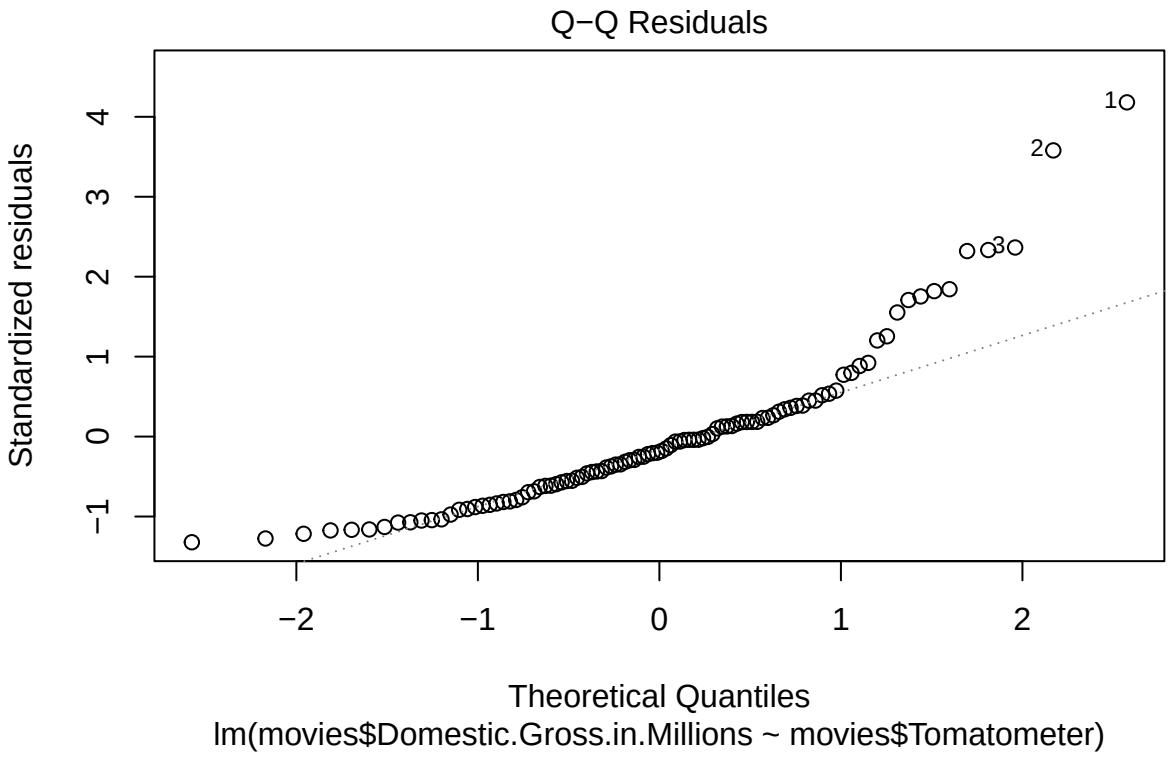


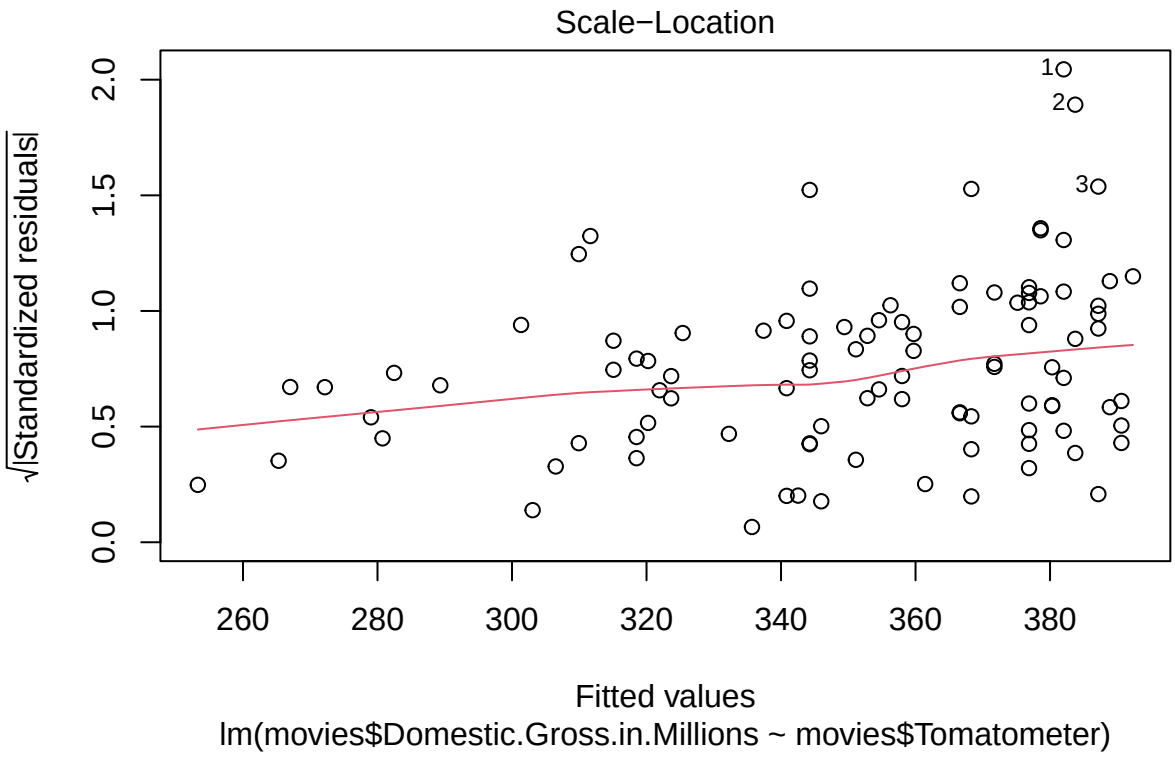
```r
summary(gross_tomatometer_lm)
```

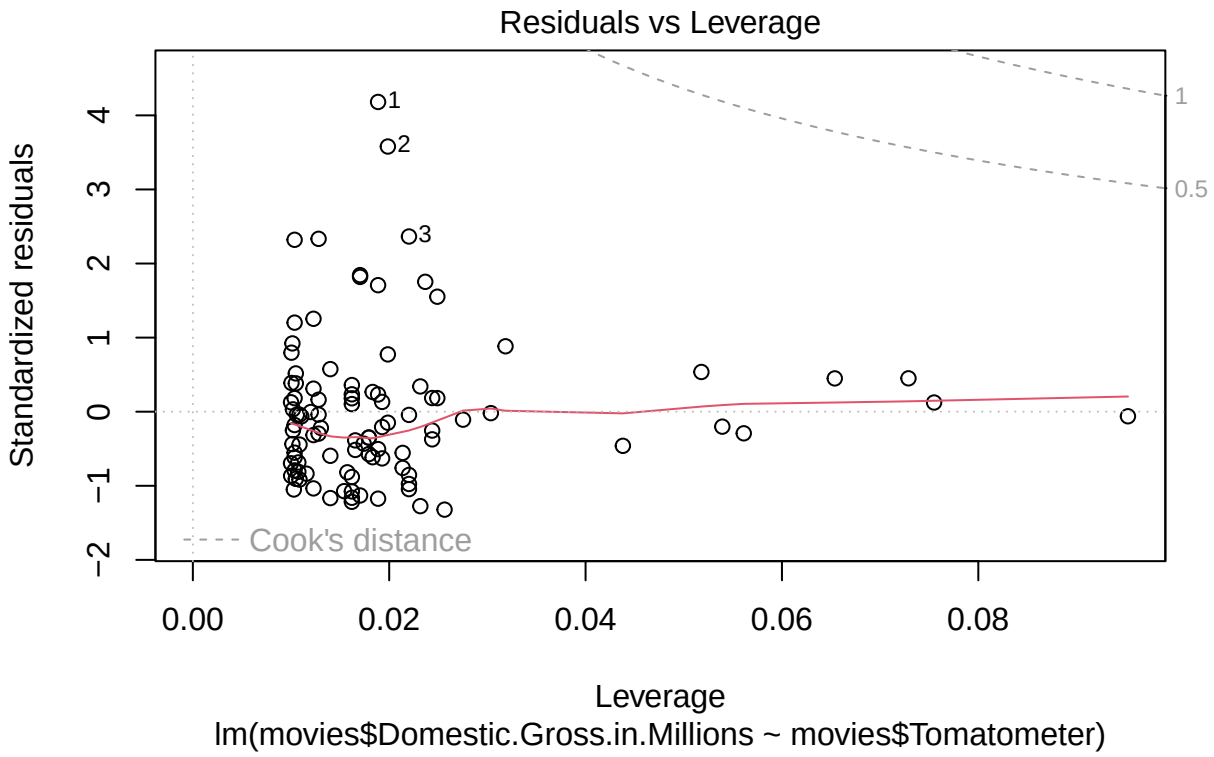
```
##
## Call:
## lm(formula = movies$Domestic.Gross.in.Millions ~ movies$Tomatometer)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -174.75  -85.51  -25.12   42.36  554.63
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          222.39      53.23   4.178 6.38e-05 ***
## movies$Tomatometer   171.66      68.95   2.490   0.0145 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 133.9 on 98 degrees of freedom
## Multiple R-squared:  0.05948,    Adjusted R-squared:  0.04989
## F-statistic: 6.198 on 1 and 98 DF,  p-value: 0.01447
```

```r
plot(gross_tomatometer_lm)
```



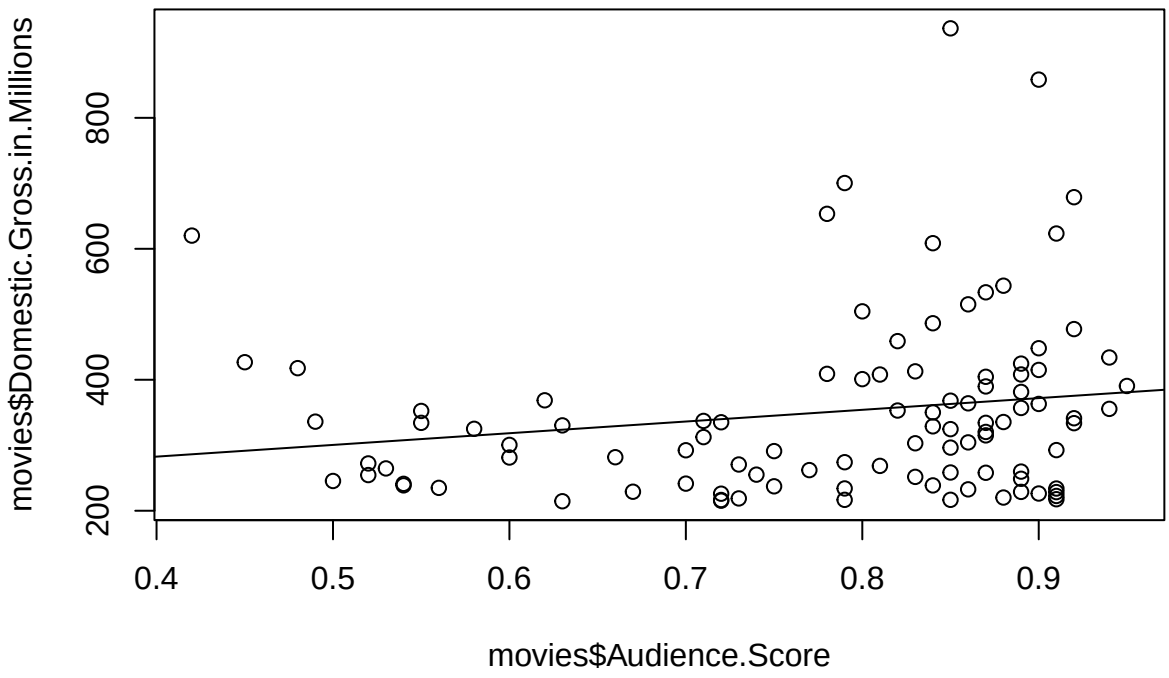

Q−Q Residuals
Standardized residuals
−1
0
1
2
3
4
1
2
3
−2
−1
0
1
2
Theoretical Quantiles
lm(movies$Domestic.Gross.in.Millions ~ movies$Tomatometer)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
1
2
3
260
280
300
320
340
360
380
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Tomatometer)



Domestic Gross predicted by Audience Score

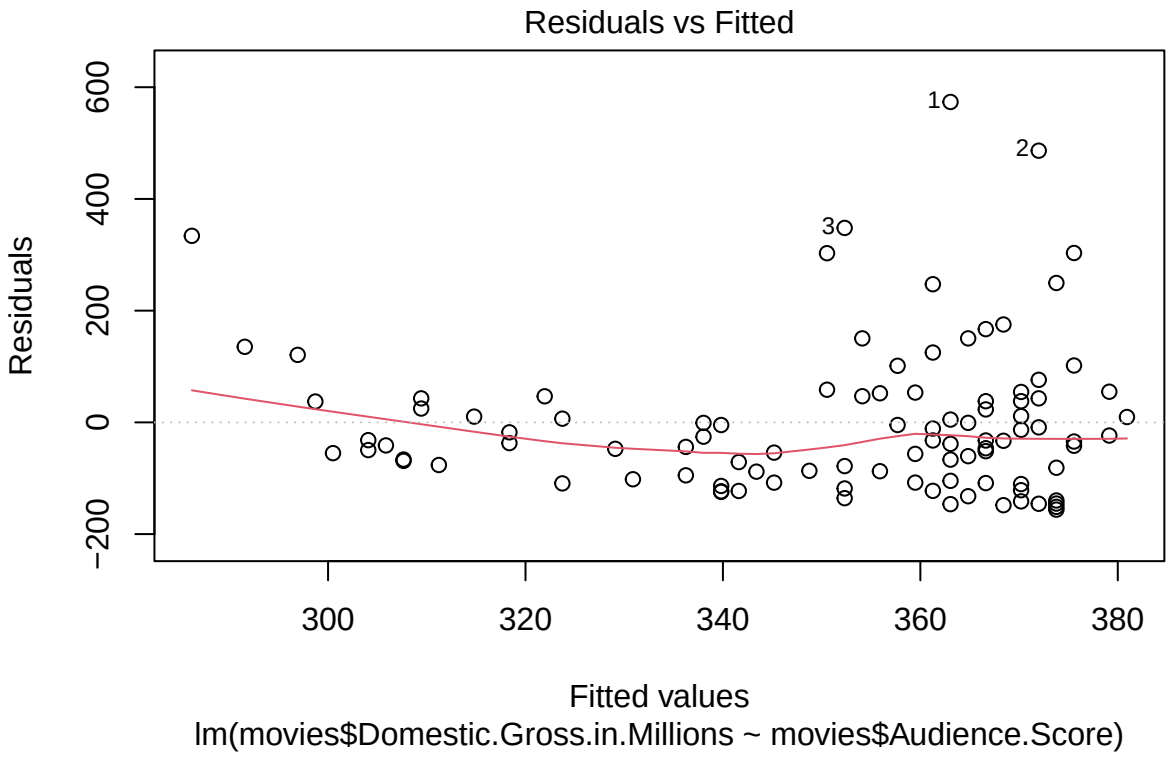
```
plot(movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)

abline(gross_audience_lm)
```

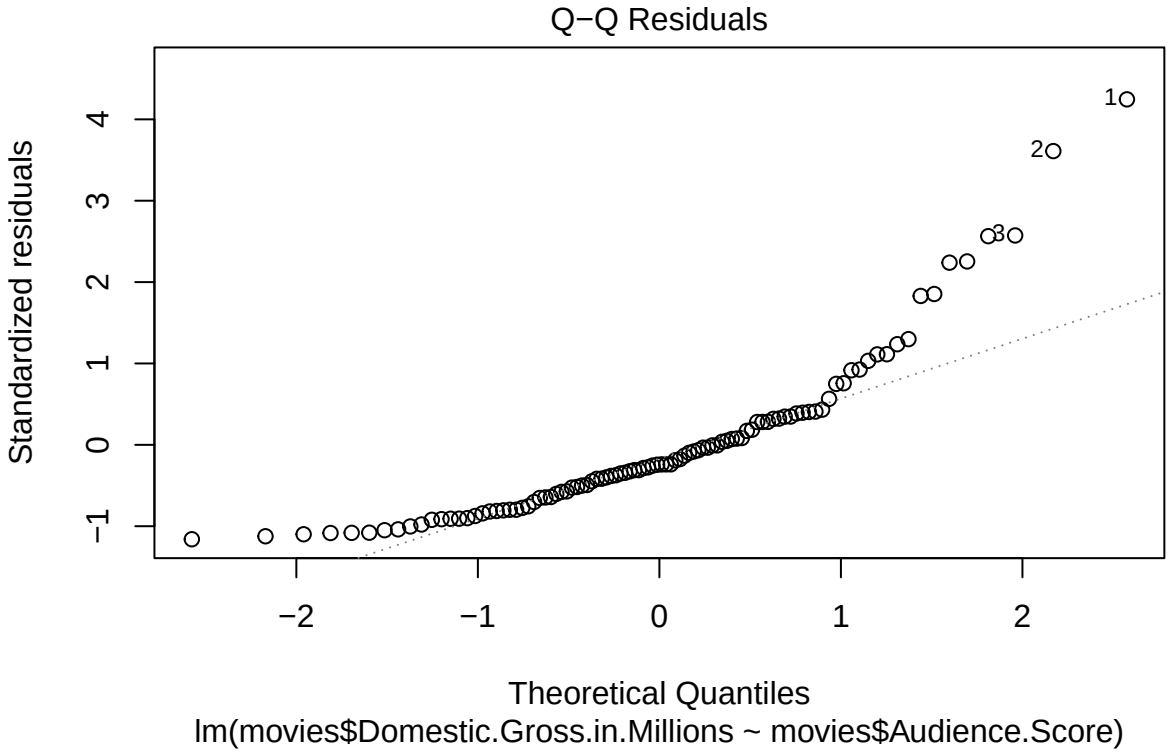


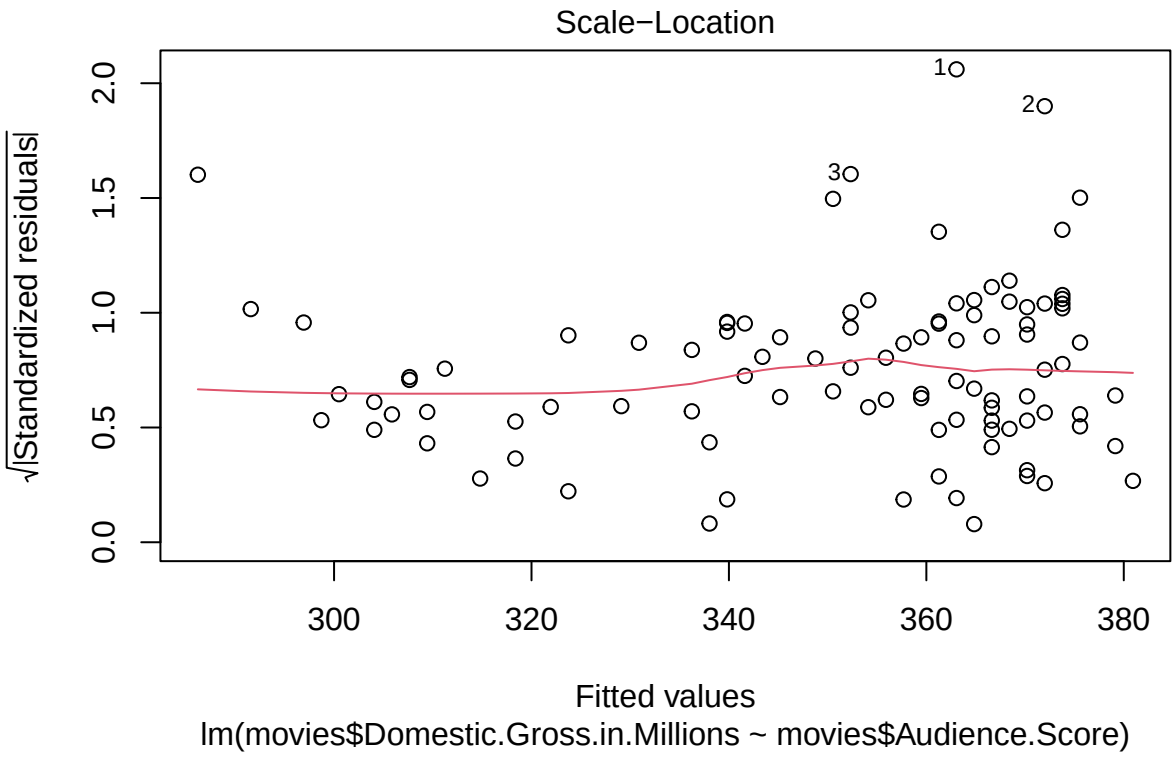

```
summary(gross_audience_lm)
```

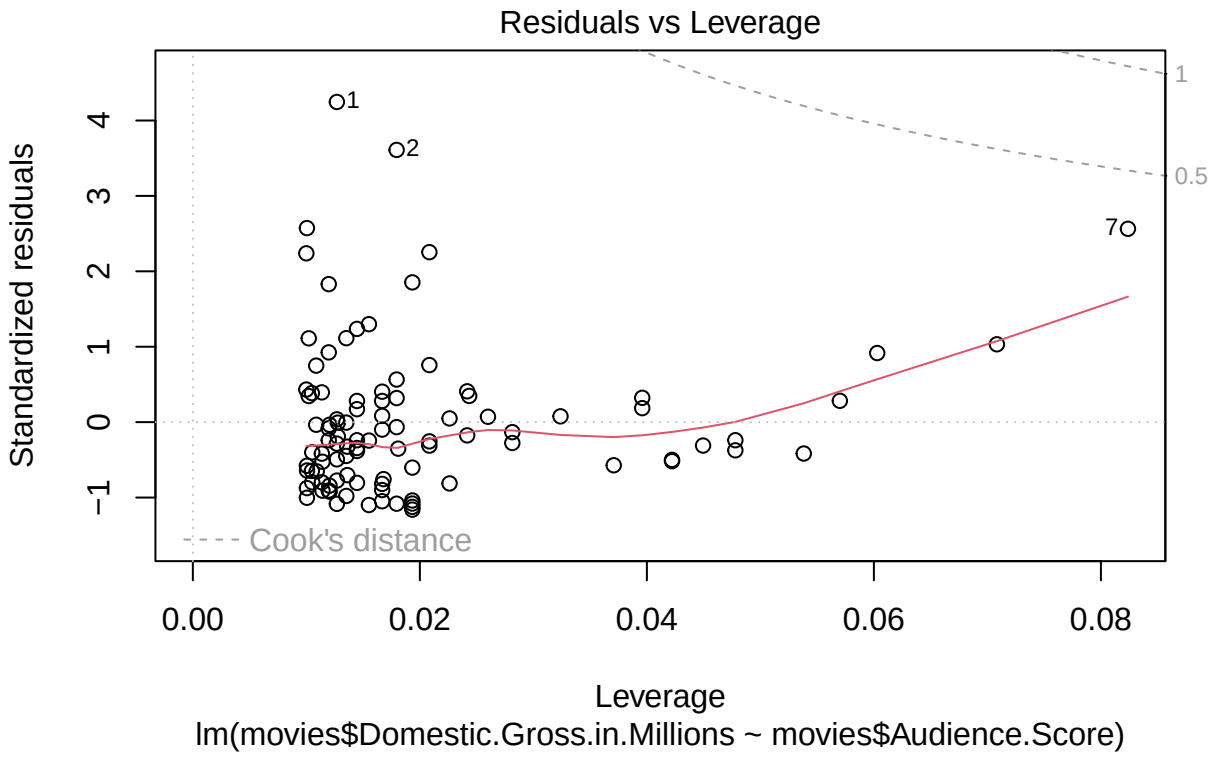
```
##
## Call:
## lm(formula = movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -156.20  -89.90  -32.69   43.93  573.61
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             211.13      80.35   2.627  0.00999 **
## movies$Audience.Score   178.74     101.47   1.762  0.08127 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 136 on 98 degrees of freedom
## Multiple R-squared:  0.03069,    Adjusted R-squared:  0.0208
## F-statistic: 3.103 on 1 and 98 DF,  p-value: 0.08127
```

```
plot(gross_audience_lm)
```


Residuals vs Fitted
Residuals
600
400
200
0
−200
1
2
3
300
320
340
360
380
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)

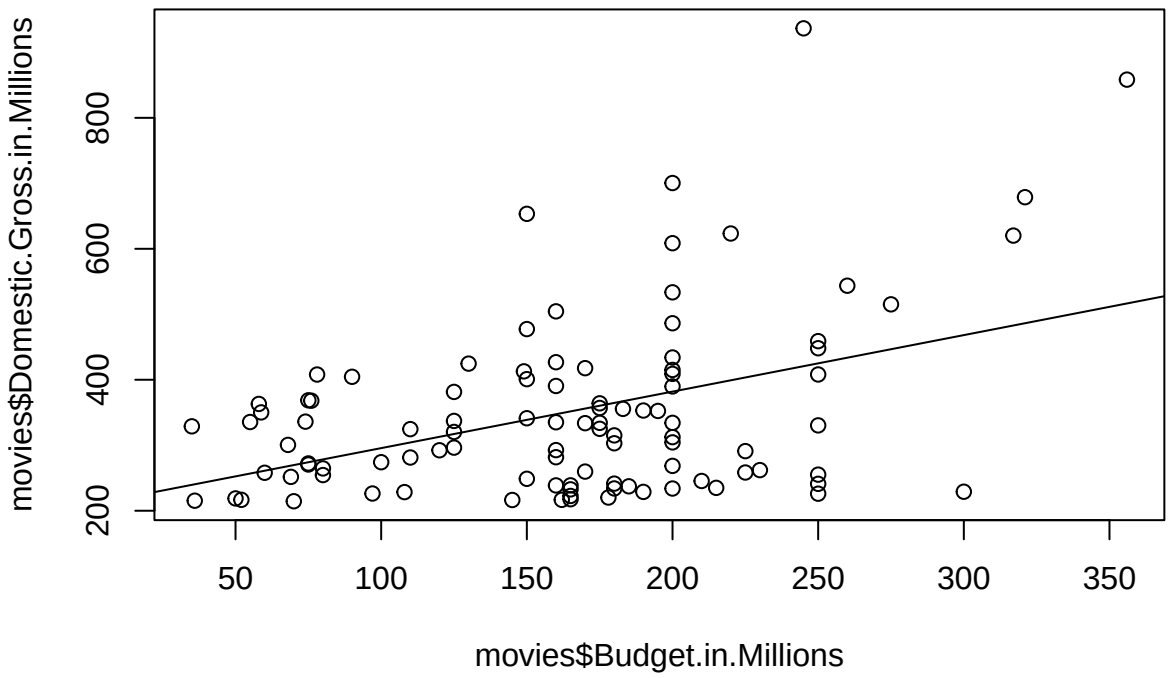

Q−Q Residuals
Standardized residuals
−1
0
1
2
3
4
1
2
3
−2
−1
0
1
2
Theoretical Quantiles
lm(movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
1
2
3
300
320
340
360
380
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Audience.Score)



Domestic Gross predicted by Budget

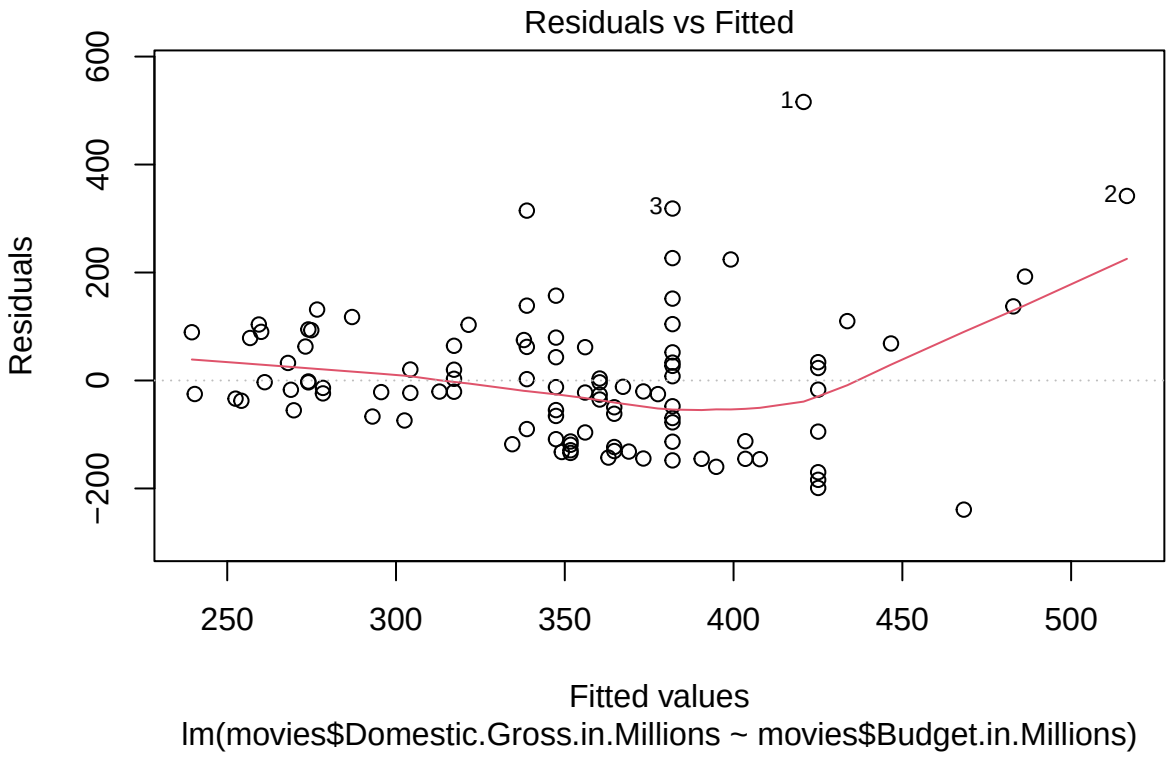
```
plot(movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)

abline(gross_budget_lm)
```

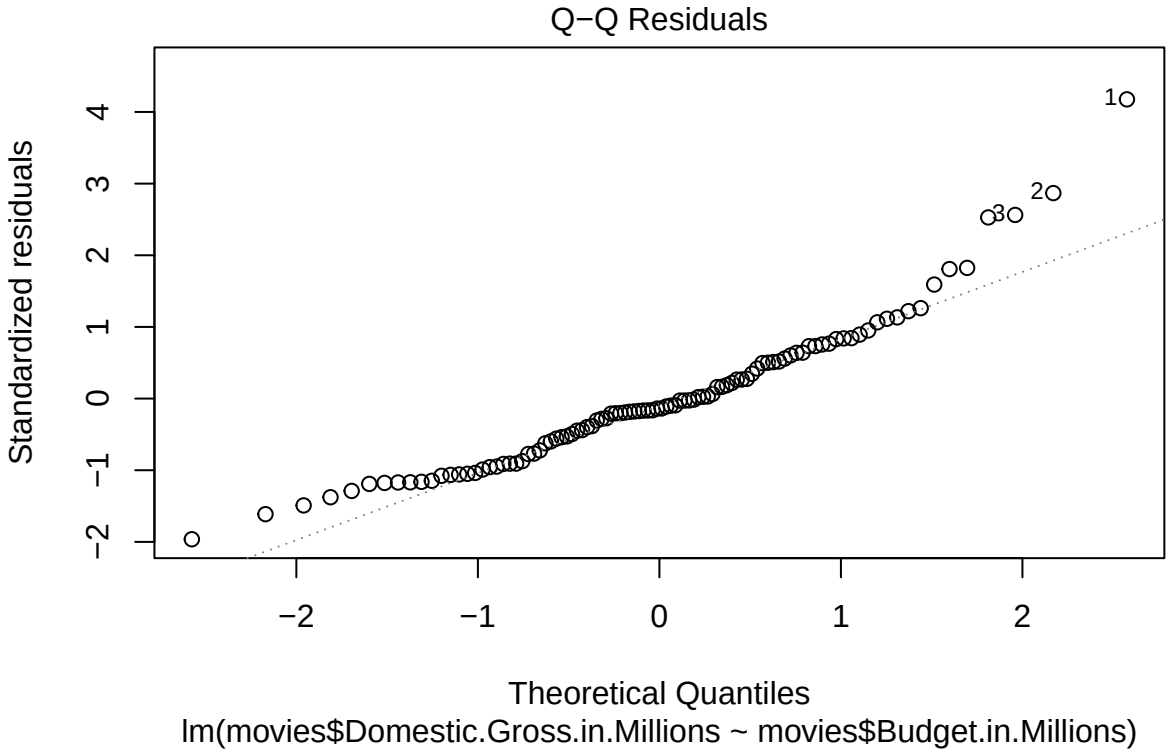


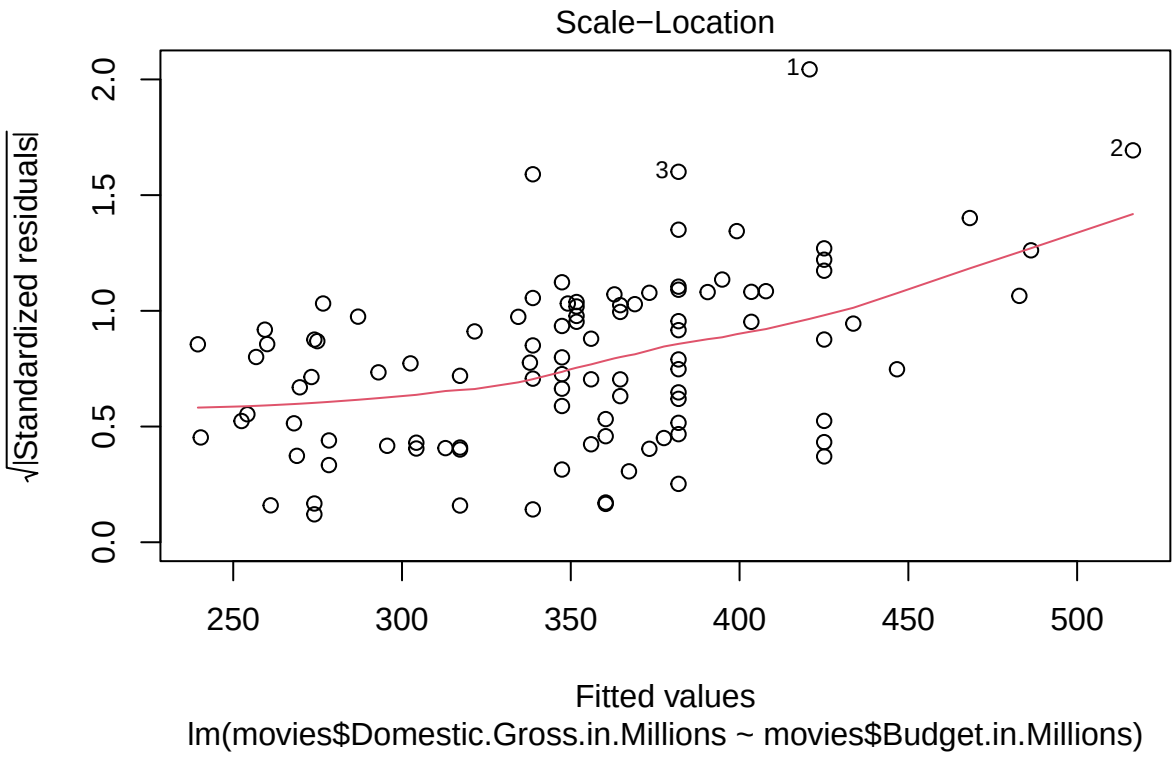

```
summary(gross_budget_lm)
```

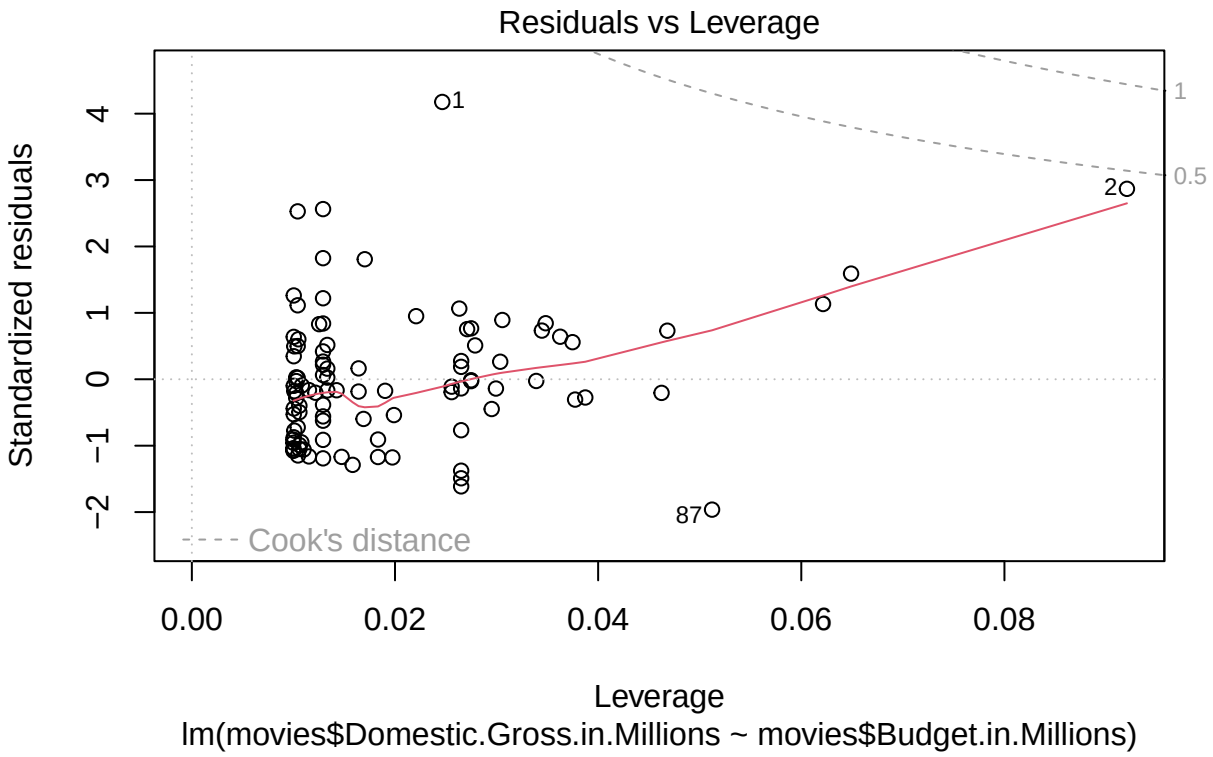
```
##
## Call:
## lm(formula = movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -239.18  -91.18  -17.08   65.34  515.92
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                209.3138    32.9893   6.345 6.91e-09 ***
## movies$Budget.in.Millions   0.8630     0.1864   4.630 1.13e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 125.1 on 98 degrees of freedom
## Multiple R-squared:  0.1795, Adjusted R-squared:  0.1711
## F-statistic: 21.43 on 1 and 98 DF,  p-value: 1.125e-05
```

```
plot(gross_budget_lm)
```


Residuals vs Fitted
Residuals
600
400
200
0
−200
1
2
3
250
300
350
400
450
500
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)

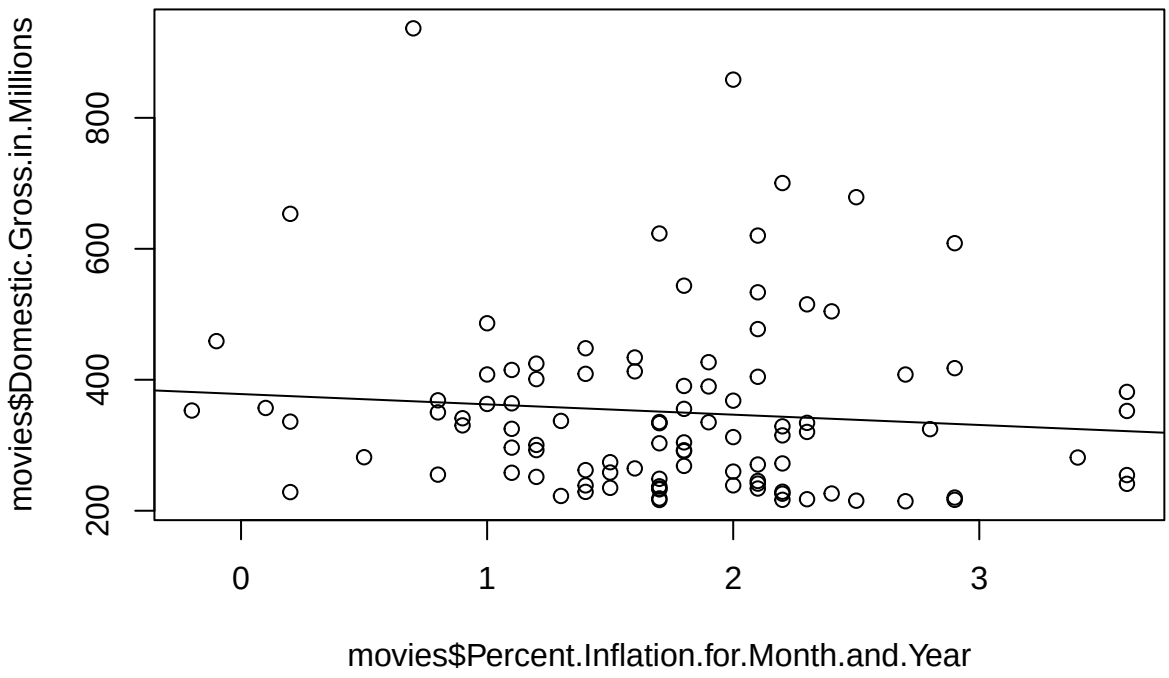

Q−Q Residuals
Standardized residuals
−2
−1
0
1
2
3
4
1
2
3
−2
−1
0
1
2
Theoretical Quantiles
lm(movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
1
2
3
250
300
350
400
450
500
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Budget.in.Millions)



Domestic Gross predicted by Inflation for Month and Year of Release

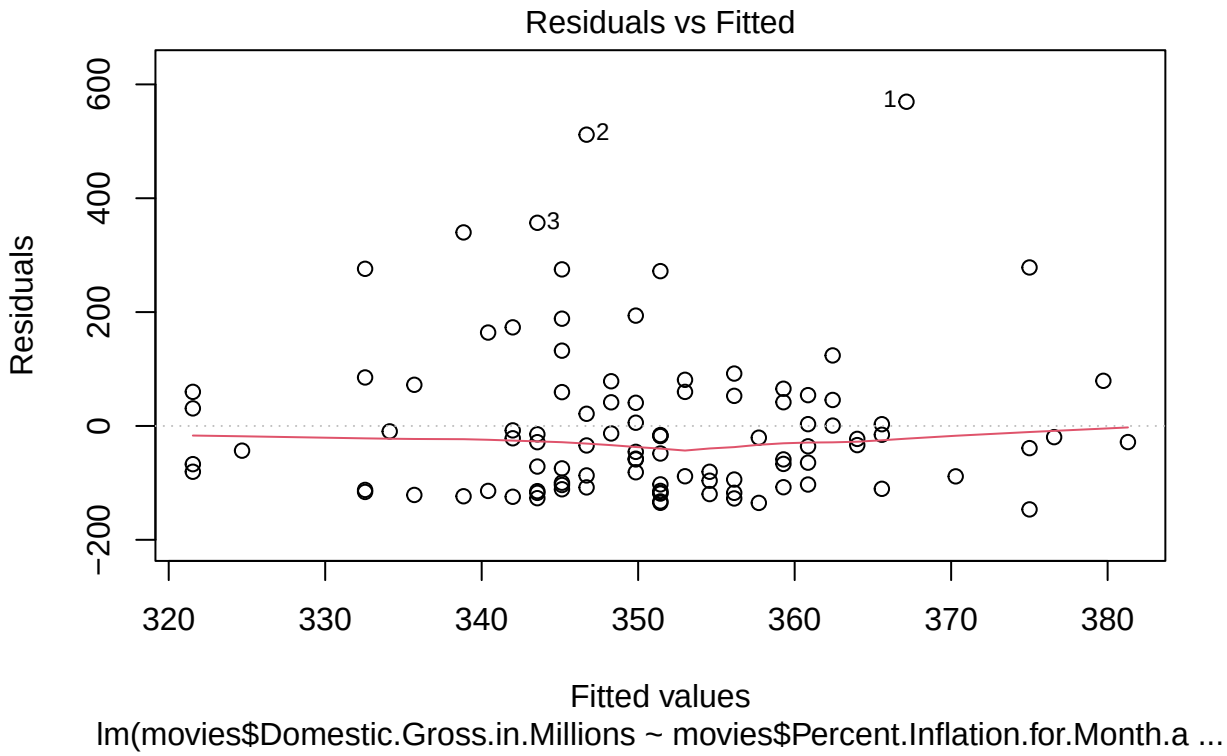
```
plot(movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.and.Year)

abline(gross_inflation_lm)
```

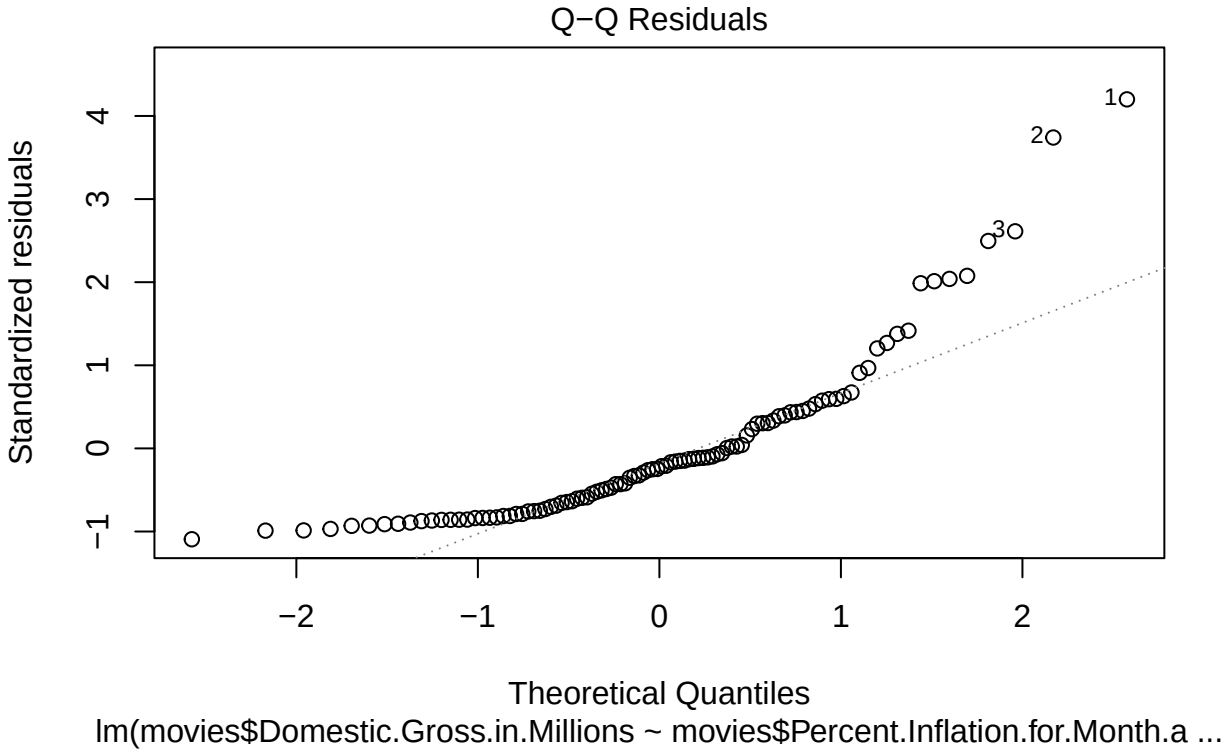


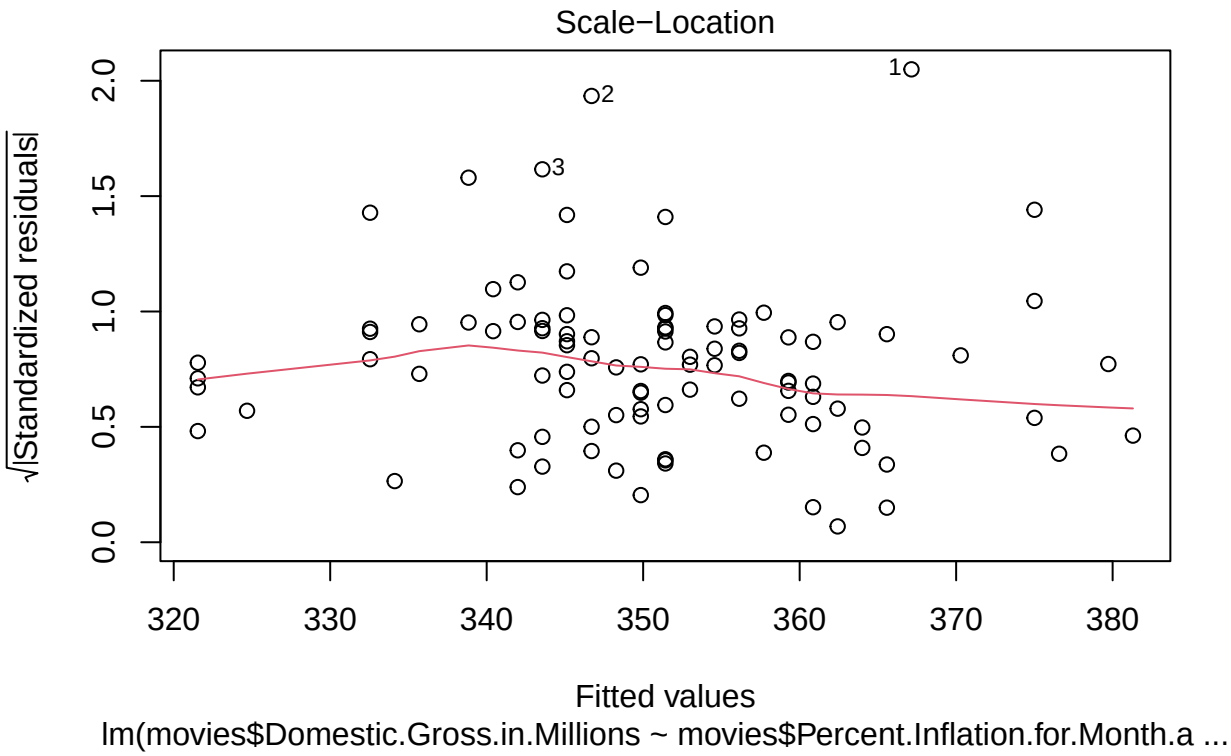

```
summary(gross_inflation_lm)
```

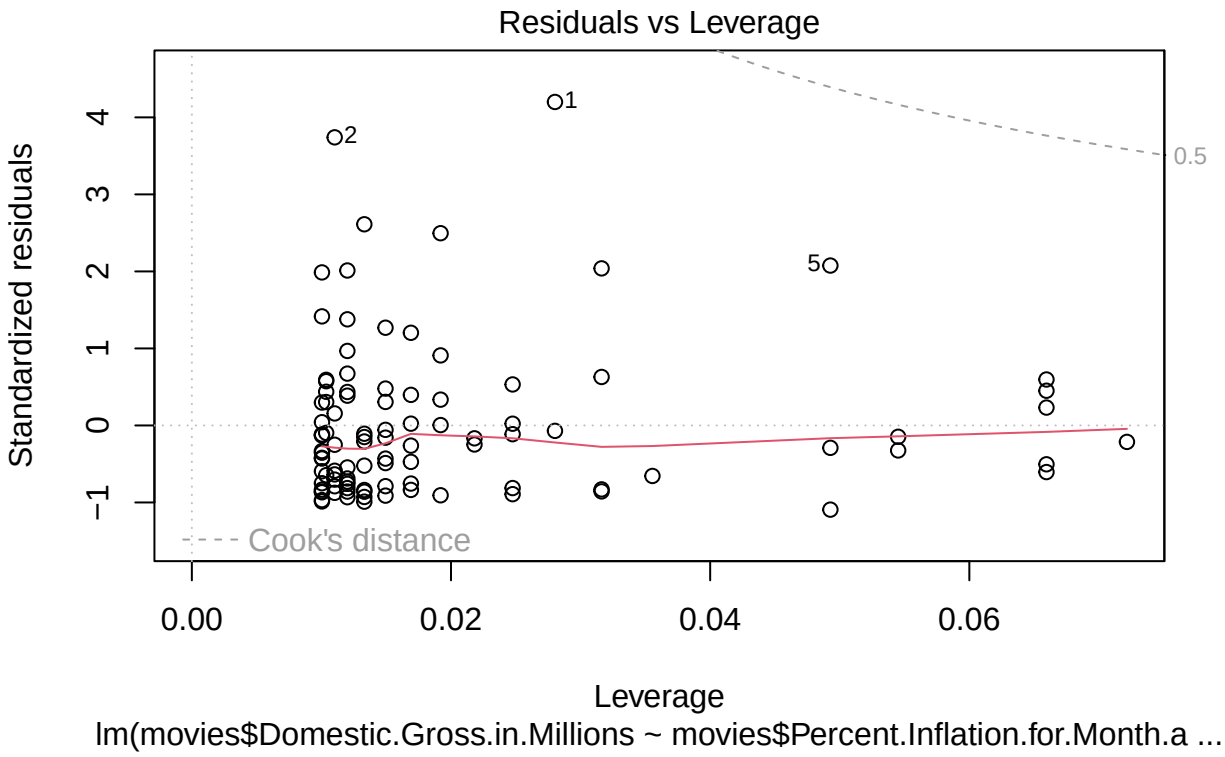
```
##
## Call:
## lm(formula = movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.and.Year)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -146.58 -102.72  -31.07   53.19  569.52
##
## Coefficients:
##                                           Estimate Std. Error t value
## (Intercept)                                 378.15      33.71  11.218
## movies$Percent.Inflation.for.Month.and.Year  -15.73      17.59  -0.894
##                                           Pr(>|t|)
## (Intercept)                                 <2e-16 ***
## movies$Percent.Inflation.for.Month.and.Year   0.373
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 137.5 on 98 degrees of freedom
## Multiple R-squared:  0.008094,   Adjusted R-squared:  -0.002028
## F-statistic: 0.7997 on 1 and 98 DF,  p-value: 0.3734
```

```r
plot(gross_inflation_lm)
```




Q−Q Residuals
Standardized residuals
−1
0
1
2
3
4
−2
−1
0
1
2
Theoretical Quantiles
lm(movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.a ...
1
2
3


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
320
330
340
350
360
370
380
1
2
3
Fitted values
lm(movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.a ...


Residuals vs Leverage
Standardized residuals
Leverage
lm(movies$Domestic.Gross.in.Millions ~ movies$Percent.Inflation.for.Month.a ...
Cook's distance
0.5
1
2
5
−1
0
1
2
3
4
0.00
0.02
0.04
0.06