

Resultados

Cícero

2022-09-30

R Markdown

Variáveis testadas:

SalePrice x LotArea : Significativa

SalePrice x ExterQual : Siginficativa

SalePrice x ExterCond : significativa

SalePrice x OverallCond : Meio meh

SalePrice x Neighborhood : significativa para algumas vizinhas

SalePrice x TotalBsmtSF : significativa

SalePrice x MiscVal : Poha nenhuma

SalePrice x MiscFeature : Poha nenhuma

```
dados <- read.csv('train.csv')
modeloLotArea <- lm(log(SalePrice) ~ LotArea,dados)
summary(modeloLotArea)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ LotArea, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.53664 -0.23781 -0.02546  0.24052  1.44451
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.192e+01  1.468e-02  811.56  <2e-16 ***
## LotArea      1.030e-05  1.013e-06   10.17  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3861 on 1458 degrees of freedom
## Multiple R-squared:  0.06621,    Adjusted R-squared:  0.06557
## F-statistic: 103.4 on 1 and 1458 DF,  p-value: < 2.2e-16
```

```
modeloExterQual <- lm(log(SalePrice) ~ ExterQual,dados)
summary(modeloExterQual)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ ExterQual, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45228 -0.16806  0.00424  0.17967  1.20986
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.76404    0.04071  313.57  <2e-16 ***
## ExterQualFa -1.45950    0.08838  -16.51  <2e-16 ***
## ExterQualGd -0.45276    0.04282  -10.57  <2e-16 ***
## ExterQualTA -0.92606    0.04186  -22.12  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2935 on 1456 degrees of freedom
## Multiple R-squared:  0.4611, Adjusted R-squared:  0.46
## F-statistic: 415.3 on 3 and 1456 DF,  p-value: < 2.2e-16
```

```
modeloExterCond <- lm(log(SalePrice) ~ ExterCond,dados)
summary(modeloExterCond)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ ExterCond, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.58284 -0.25252 -0.02605  0.24266  1.49140
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.11973    0.22562  53.717  < 2e-16 ***
## ExterCondFa -0.66456    0.23740  -2.799  0.00519 **
## ExterCondGd -0.15028    0.22793  -0.659  0.50980
## ExterCondPo -0.87468    0.45125  -1.938  0.05277 .
## ExterCondTA -0.07665    0.22589  -0.339  0.73442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3908 on 1455 degrees of freedom
## Multiple R-squared:  0.04551, Adjusted R-squared:  0.04289
## F-statistic: 17.34 on 4 and 1455 DF,  p-value: 6.54e-14
```

```
modeloOverallCond <- lm(log(SalePrice) ~ OverallCond,dados)
summary(modeloOverallCond)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ OverallCond, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.58618 -0.24495 -0.01824  0.24207  1.51604
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.097836   0.053410 226.509  <2e-16 ***
## OverallCond -0.013234   0.009394  -1.409   0.159
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3993 on 1458 degrees of freedom
## Multiple R-squared:  0.001359,    Adjusted R-squared:  0.0006743
## F-statistic: 1.984 on 1 and 1458 DF,  p-value: 0.1591
```

```
modeloNeighborhood <- lm(log(SalePrice) ~ Neighborhood,dados)
summary(modeloNeighborhood)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ Neighborhood, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16116 -0.14728 -0.00389  0.14475  1.36721
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    12.169416   0.063998 190.153  < 2e-16 ***
## NeighborhoodBlueste -0.342880   0.197256  -1.738  0.08238 .
## NeighborhoodBrDale  -0.621551   0.091910  -6.763 1.97e-11 ***
## NeighborhoodBrkSide -0.489689   0.072775  -6.729 2.47e-11 ***
## NeighborhoodClearCr  0.070484   0.081132   0.869  0.38513
## NeighborhoodCollgCr -0.005774   0.067527  -0.086  0.93187
## NeighborhoodCrawfor  0.037243   0.073899   0.504  0.61435
## NeighborhoodEdwards -0.457104   0.069225  -6.603 5.65e-11 ***
## NeighborhoodGilbert -0.013612   0.070549  -0.193  0.84703
## NeighborhoodIDOTRR  -0.722527   0.077315  -9.345  < 2e-16 ***
## NeighborhoodMeadowV -0.694893   0.090507  -7.678 2.98e-14 ***
## NeighborhoodMitchel -0.235468   0.074275  -3.170  0.00156 **
## NeighborhoodNAMES    -0.301371   0.066372  -4.541 6.08e-06 ***
## NeighborhoodNoRidge  0.506585   0.076118   6.655 4.01e-11 ***
## NeighborhoodNPkVill -0.302938   0.108776  -2.785  0.00542 **
## NeighborhoodNridgHt  0.449996   0.070711   6.364 2.64e-10 ***
## NeighborhoodNWAmes  -0.038807   0.071060  -0.546  0.58508
## NeighborhoodOldTown -0.465551   0.068644  -6.782 1.73e-11 ***
## NeighborhoodSawyer  -0.357948   0.070970  -5.044 5.15e-07 ***
## NeighborhoodSawyerW -0.078726   0.072635  -1.084  0.27861
## NeighborhoodSomerst  0.127080   0.070039   1.814  0.06982 .
## NeighborhoodStoneBr  0.416070   0.082951   5.016 5.94e-07 ***
```

```
## NeighborhoodSWISU    -0.330981    0.082951   -3.990 6.94e-05 ***
## NeighborhoodTimber    0.194040    0.076994    2.520 0.01184 *
## NeighborhoodVeenker    0.174760    0.102106    1.712 0.08719 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2639 on 1435 degrees of freedom
## Multiple R-squared:  0.5708, Adjusted R-squared:  0.5636
## F-statistic: 79.52 on 24 and 1435 DF,  p-value: < 2.2e-16

modeloTotalBsmtSF <- lm(log(SalePrice) ~ TotalBsmtSF,dados)
summary(modeloTotalBsmtSF)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ TotalBsmtSF, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.85724 -0.19104 -0.02711  0.22317  1.01939
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.143e+01  2.159e-02  529.73  <2e-16 ***
## TotalBsmtSF  5.574e-04  1.886e-05   29.56  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.316 on 1458 degrees of freedom
## Multiple R-squared:  0.3747, Adjusted R-squared:  0.3743
## F-statistic: 873.7 on 1 and 1458 DF,  p-value: < 2.2e-16
```

```
modeloMiscVal <- lm(log(SalePrice) ~ MiscVal,dados)
summary(modeloMiscVal)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ MiscVal, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56451 -0.24965 -0.02053  0.24898  1.50972
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.202e+01  1.050e-02 1145.679  <2e-16 ***
## MiscVal      -1.612e-05  2.108e-05  -0.765    0.445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3995 on 1458 degrees of freedom
## Multiple R-squared:  0.0004008, Adjusted R-squared: -0.0002848
## F-statistic: 0.5846 on 1 and 1458 DF,  p-value: 0.4446
```

```

modeloMiscFeature <- lm(log(SalePrice) ~ MiscFeature,dados)
summary(modeloMiscFeature)

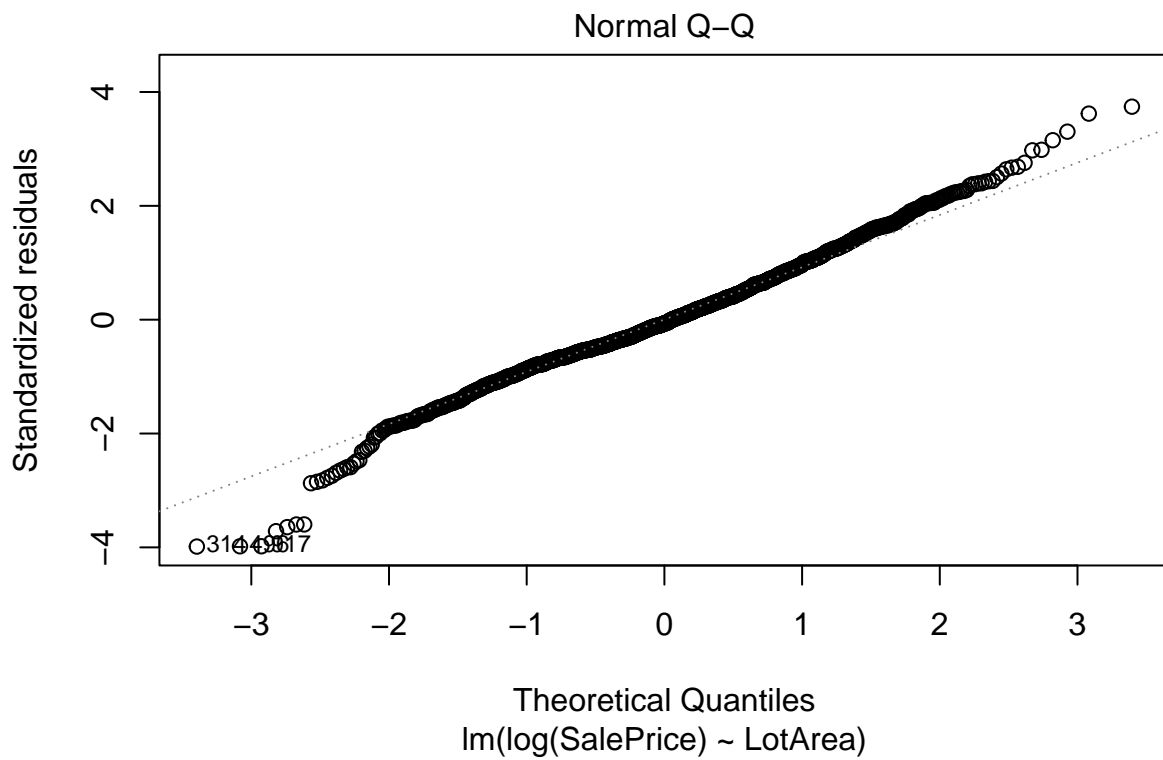
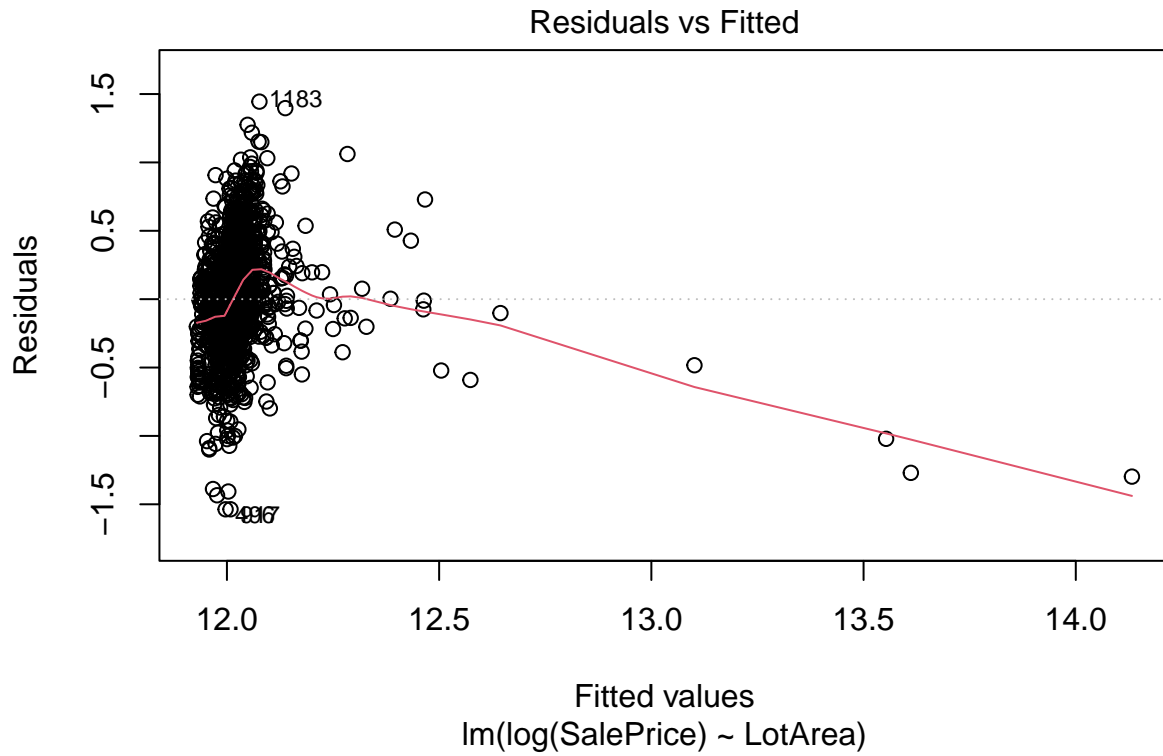
```

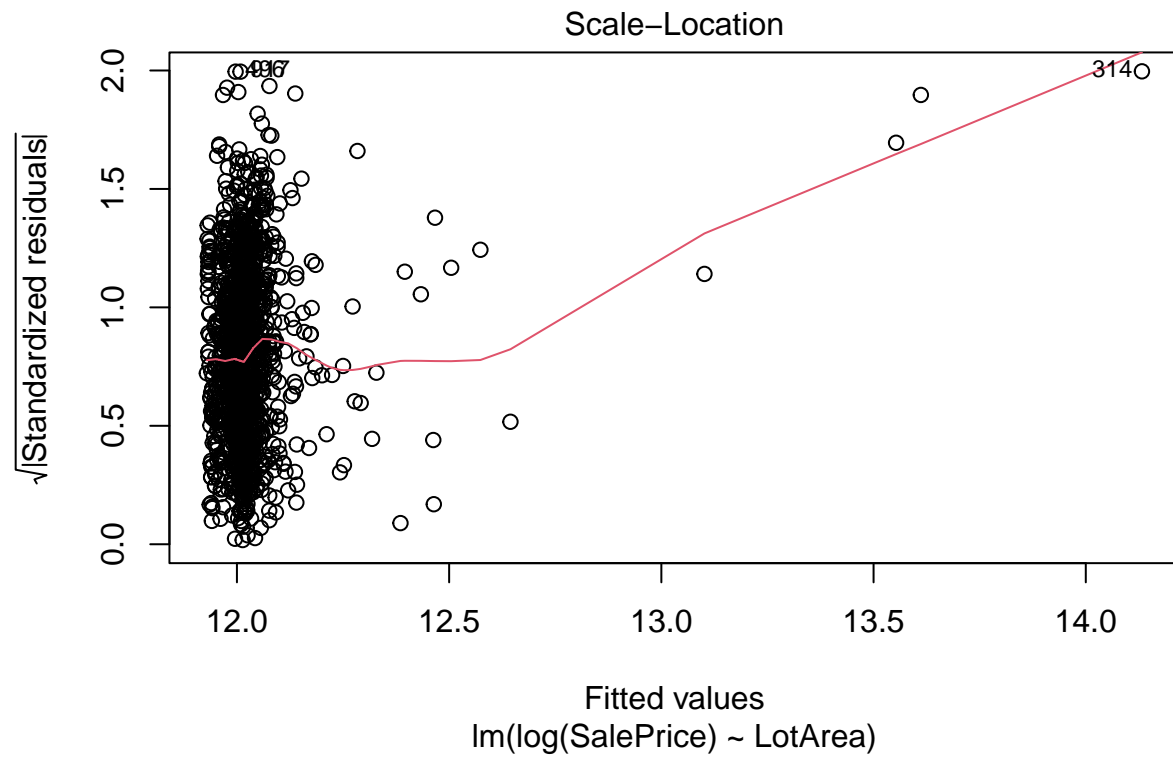
```

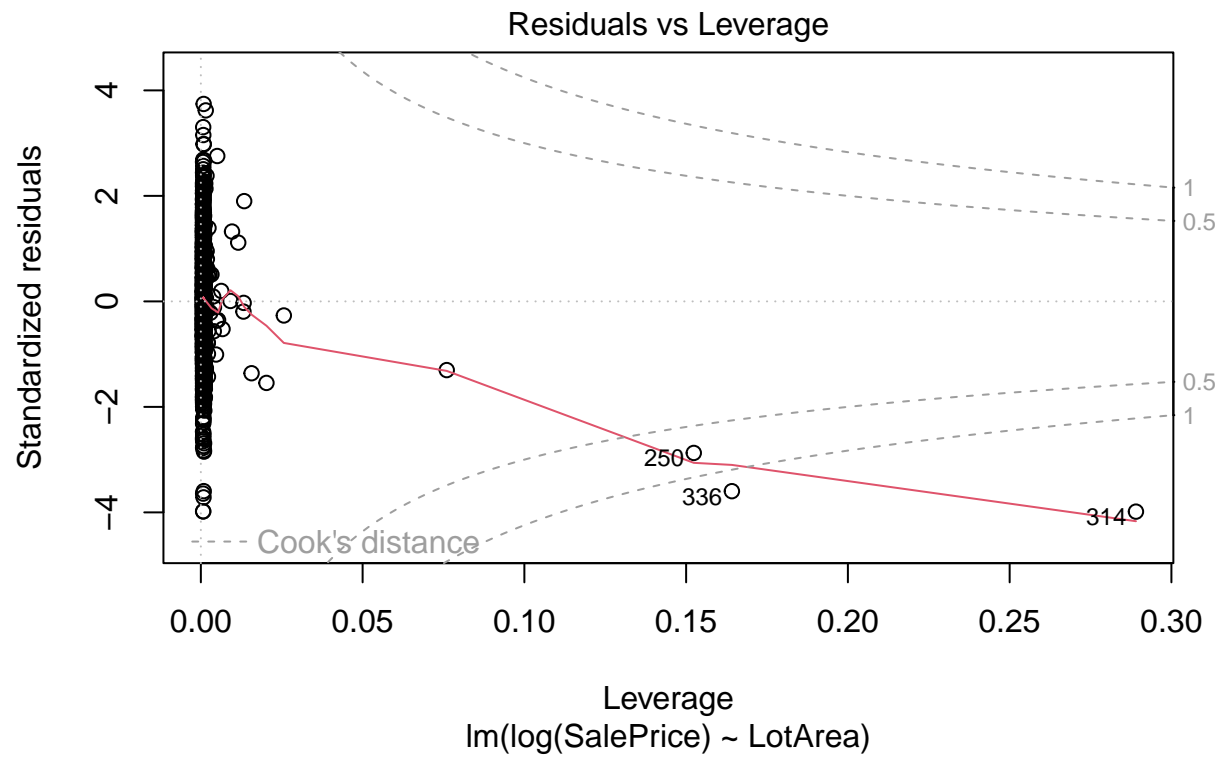
##
## Call:
## lm(formula = log(SalePrice) ~ MiscFeature, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9354 -0.1721  0.0057  0.1839  0.6634
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    12.0416     0.2506  48.055  <2e-16 ***
## MiscFeature0thr -0.6850     0.3544  -1.933   0.0589 .
## MiscFeatureShed -0.1732     0.2556  -0.677   0.5013
## MiscFeatureTenC  0.3877     0.4340   0.893   0.3760
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3544 on 50 degrees of freedom
## (1406 observations deleted due to missingness)
## Multiple R-squared:  0.1251, Adjusted R-squared:  0.07265
## F-statistic: 2.384 on 3 and 50 DF,  p-value: 0.08031

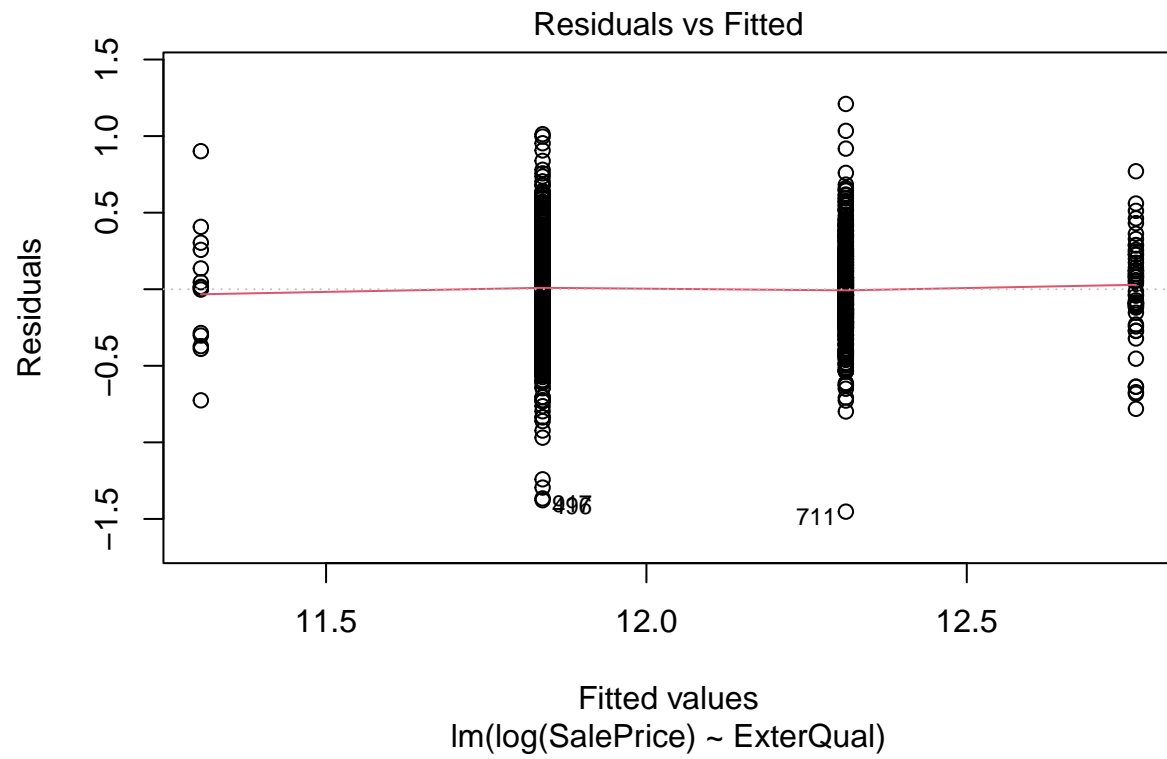
```

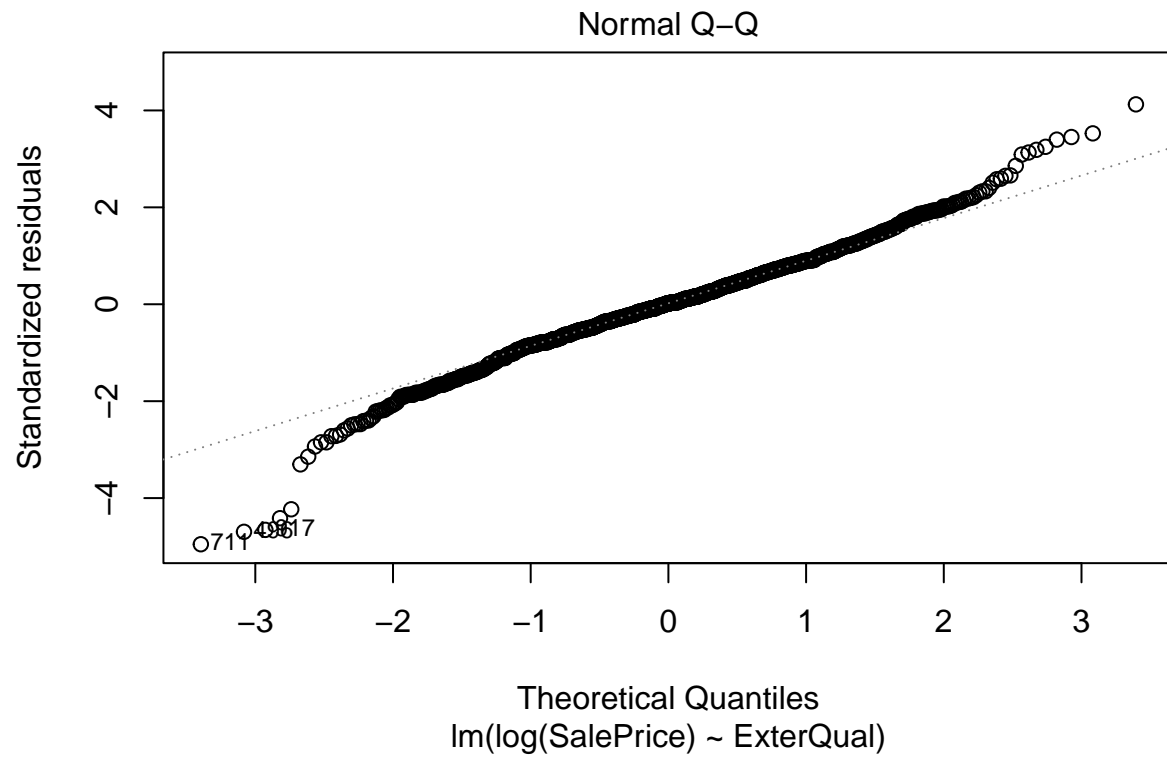
Lm Plots

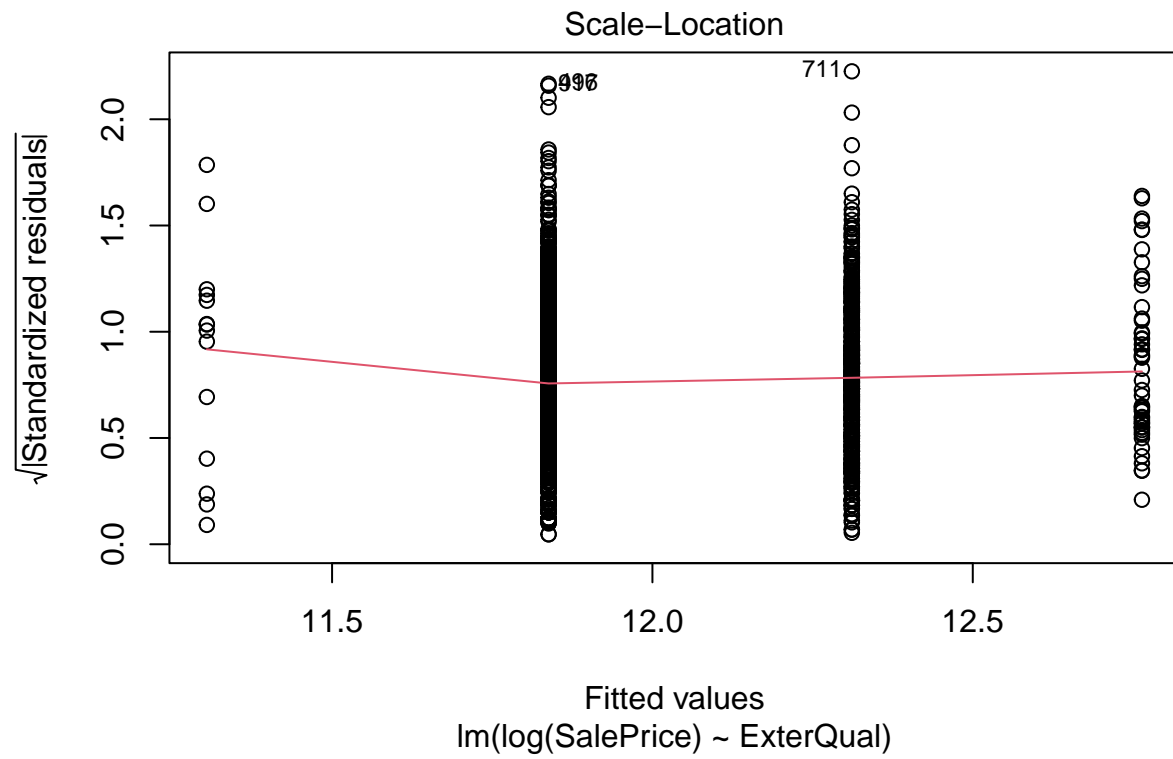


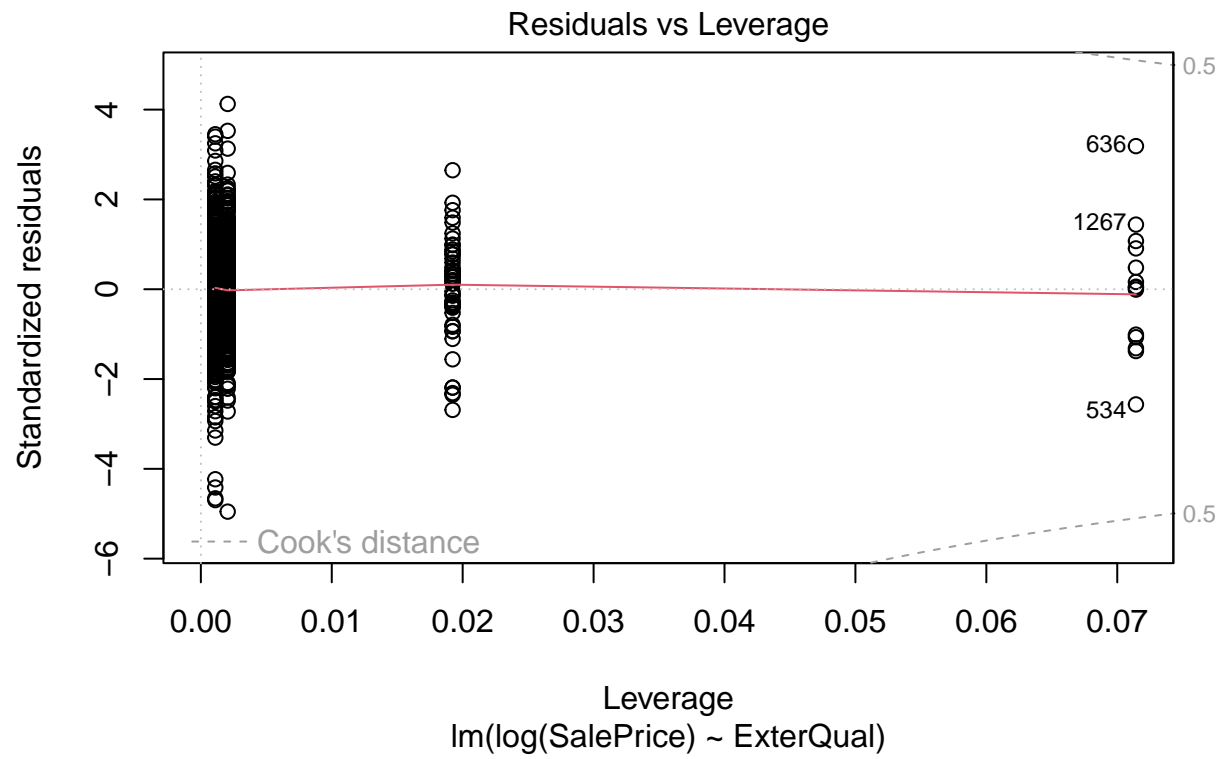


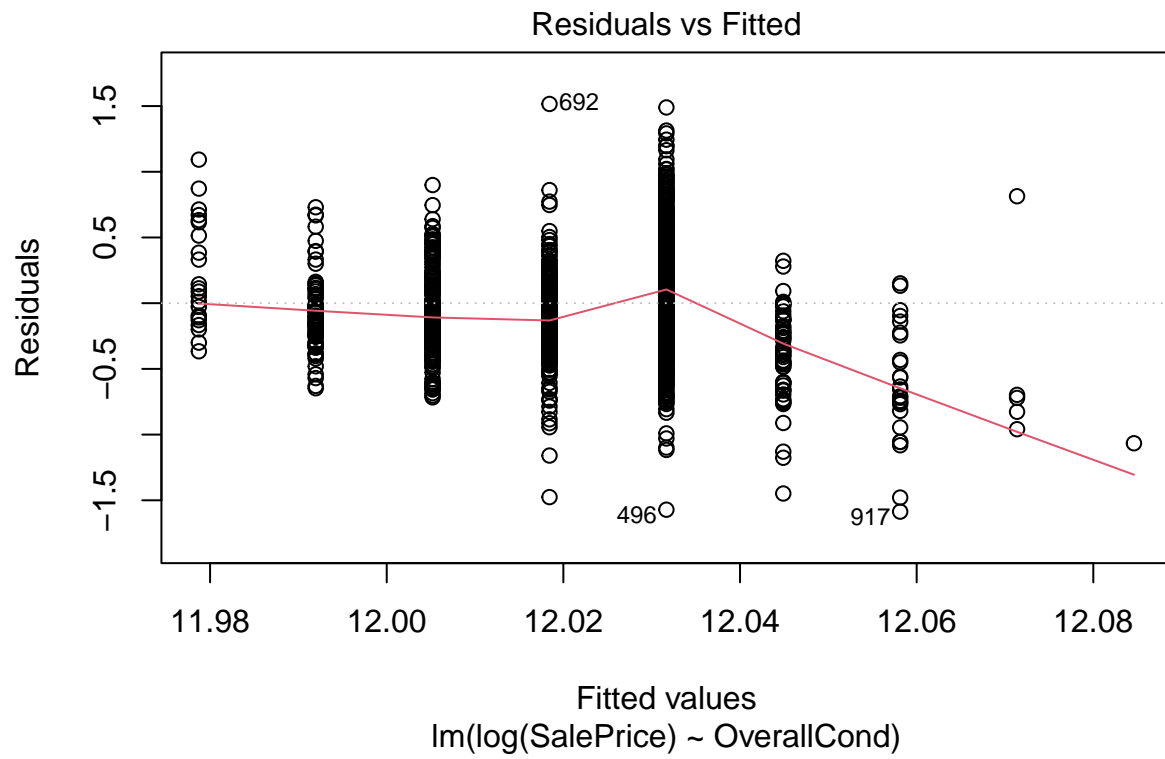


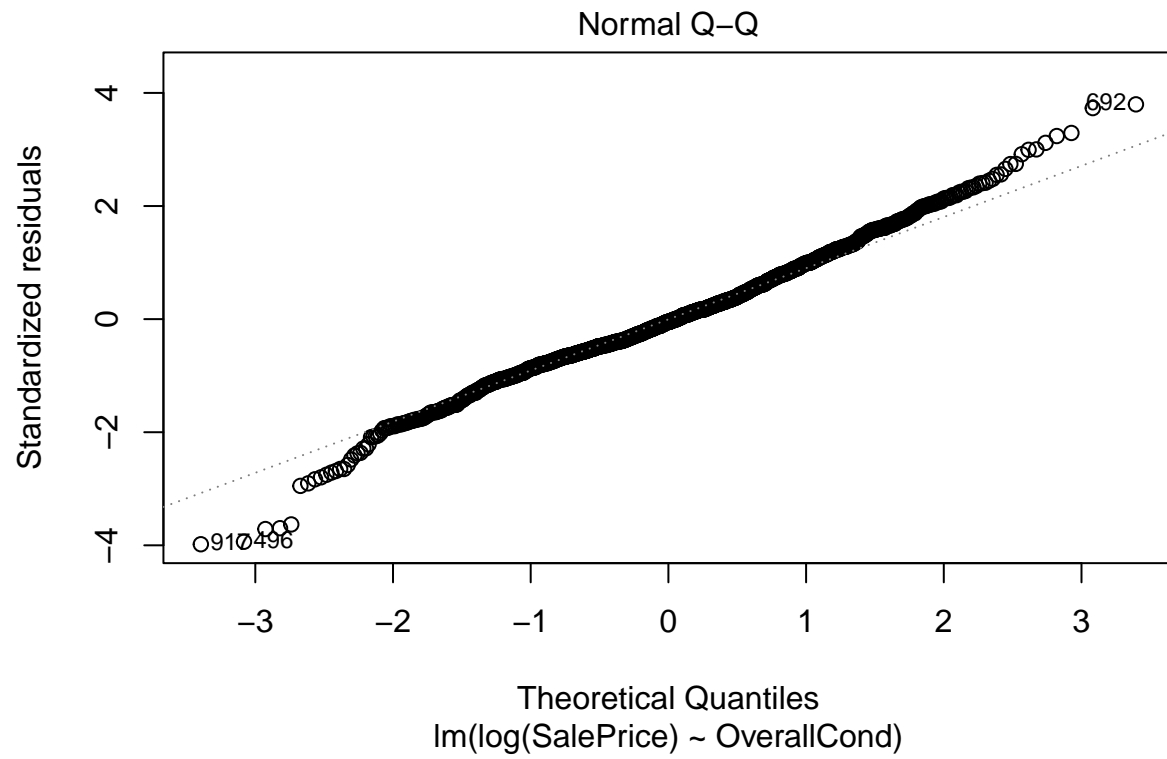


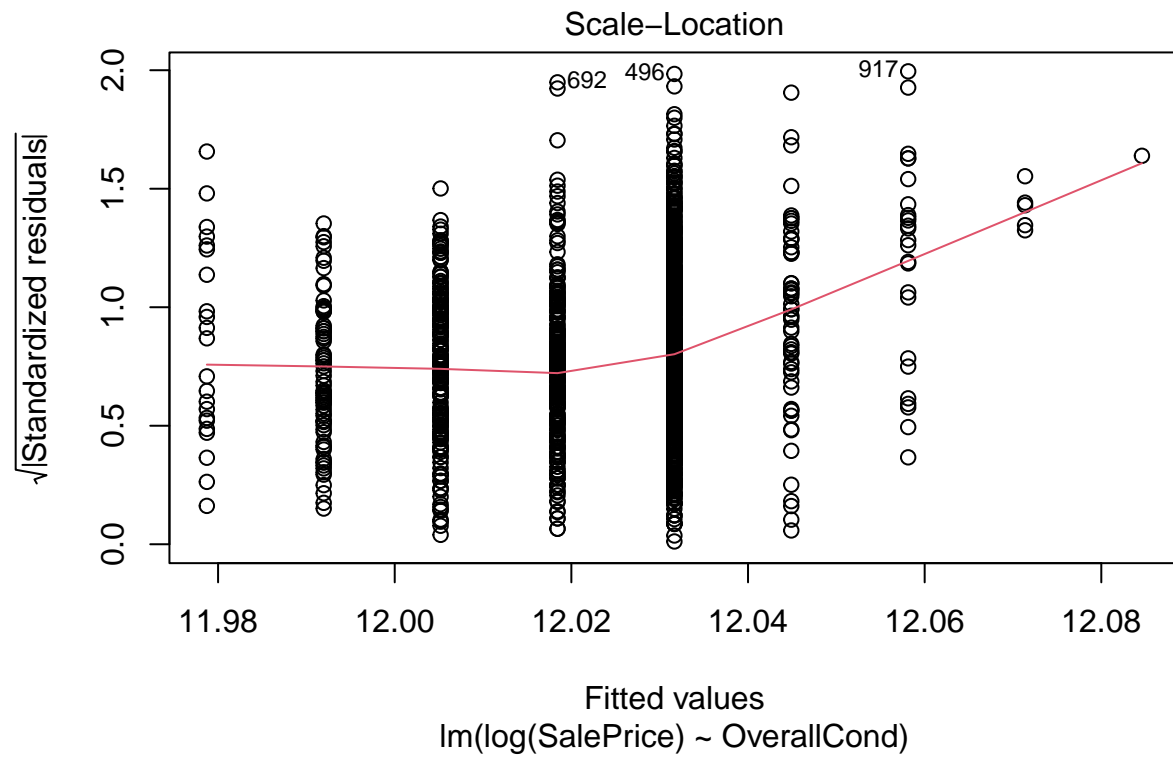


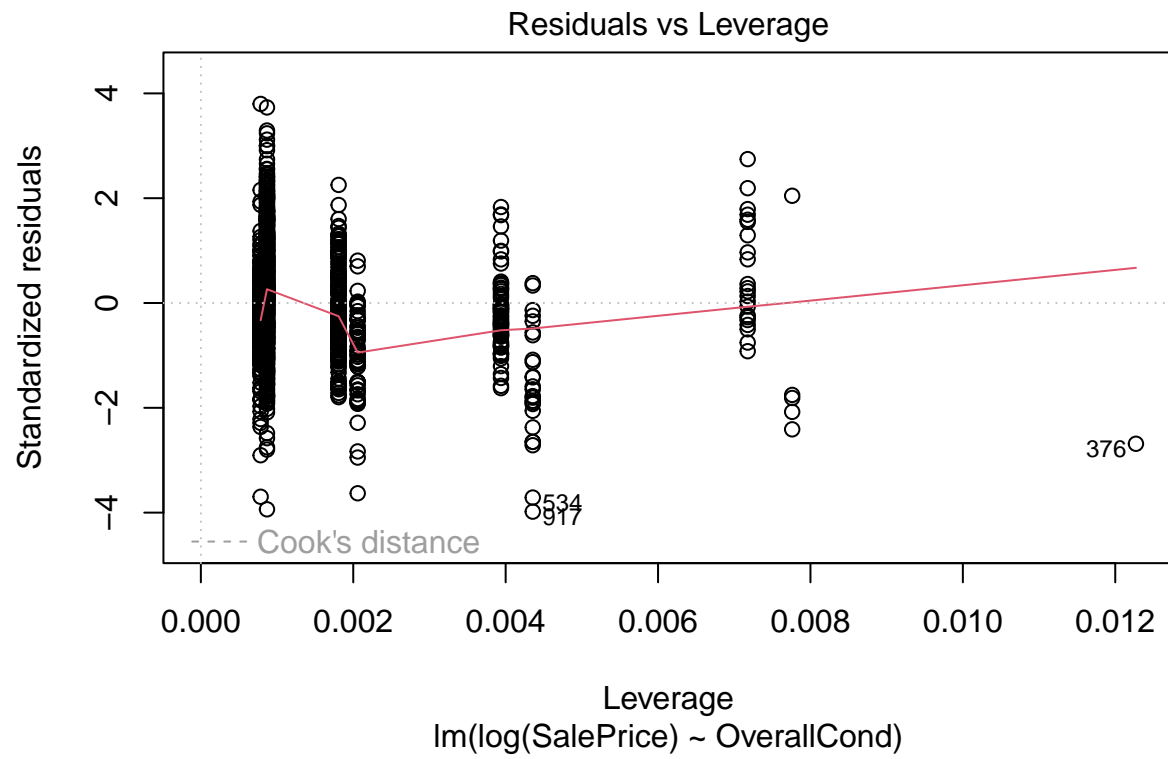


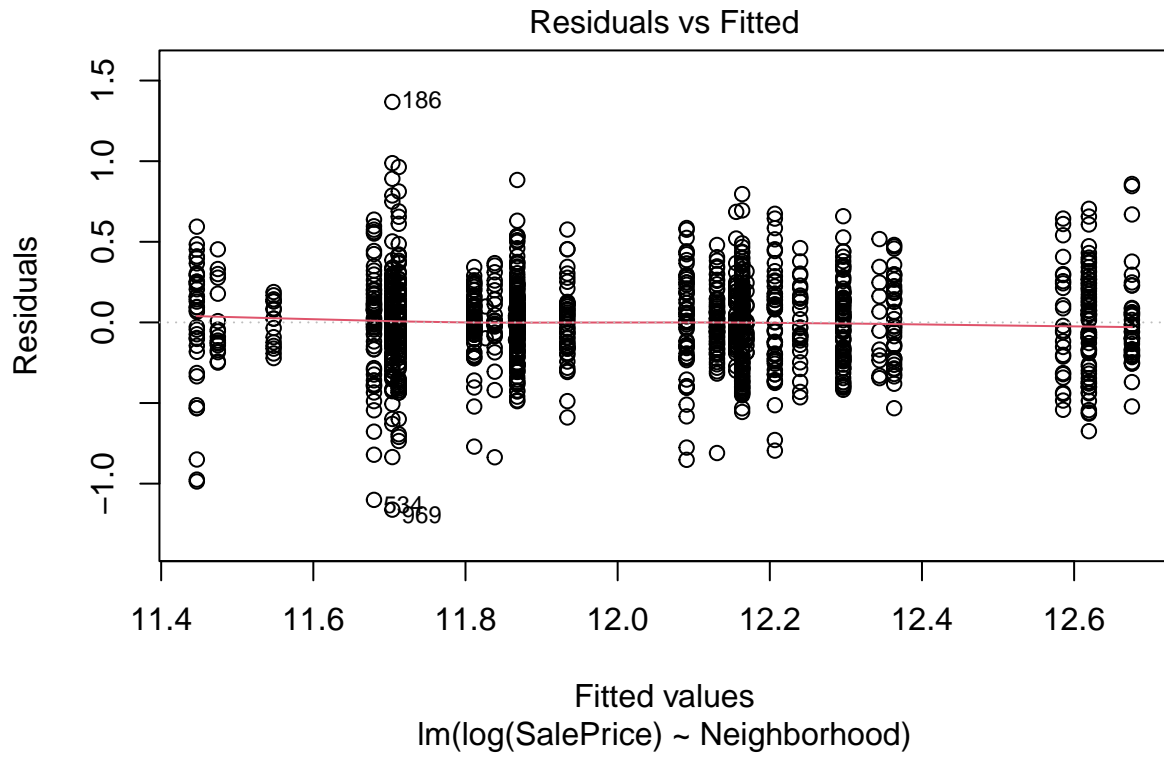


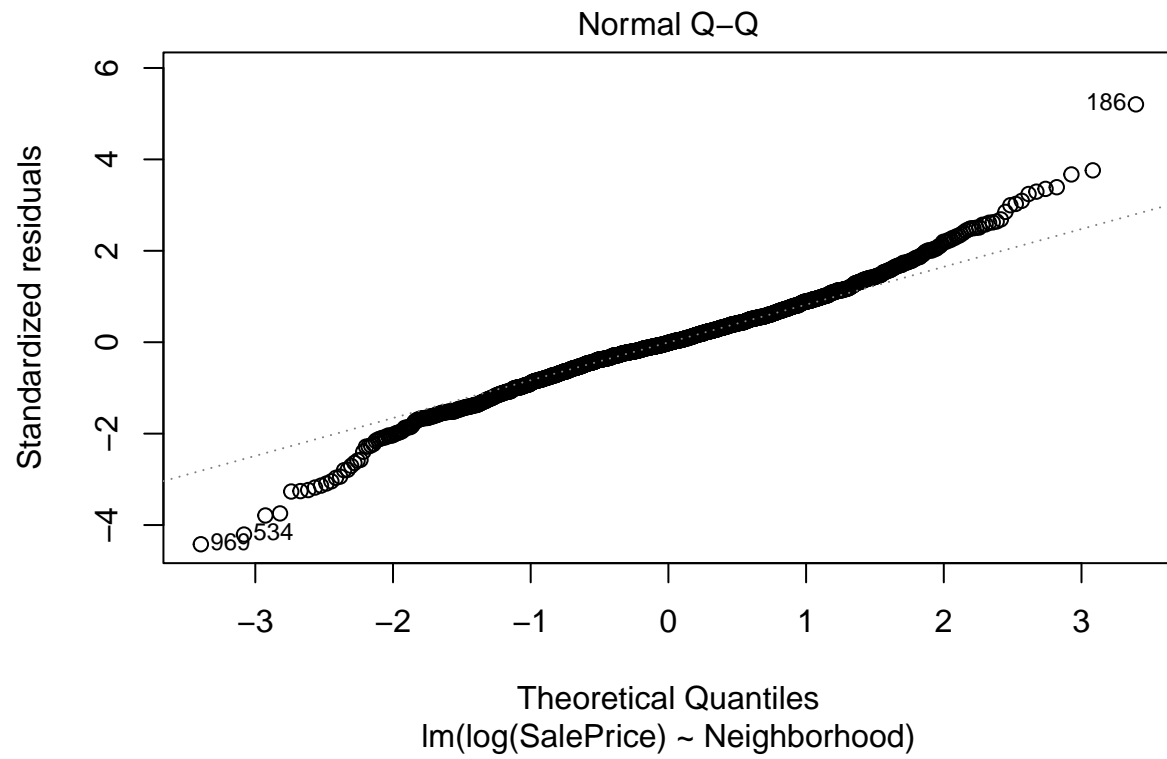


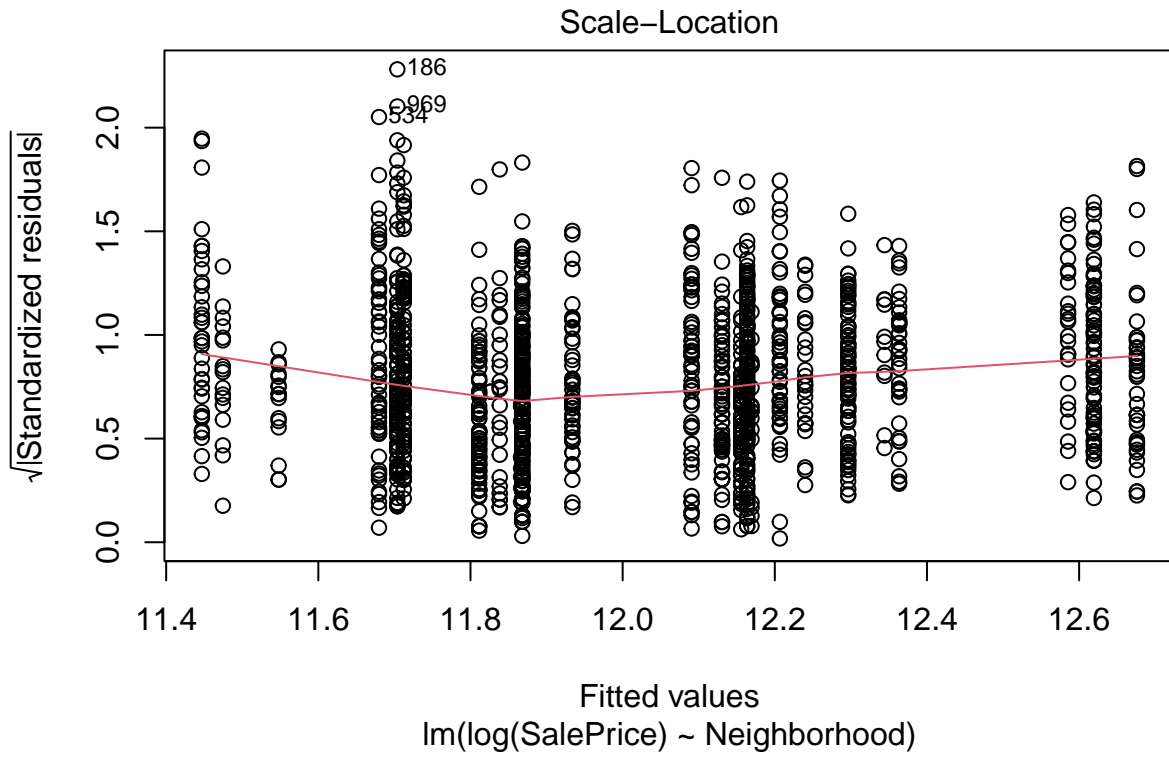


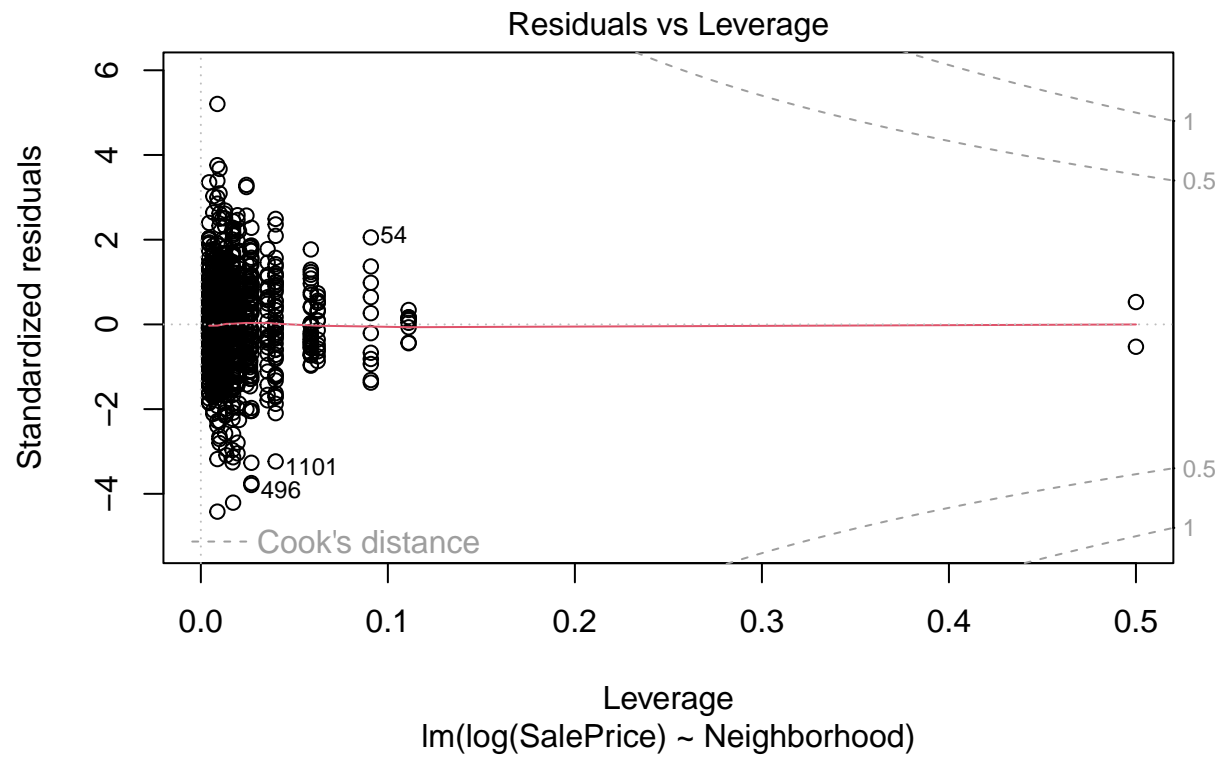


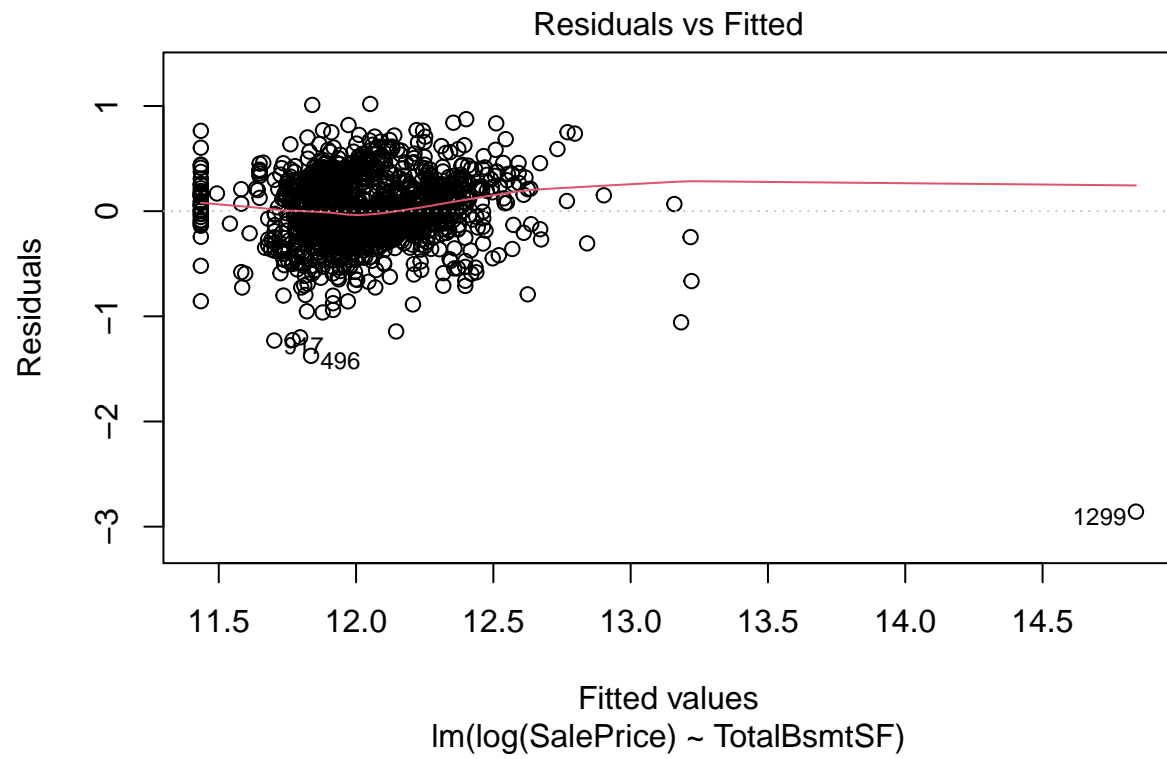


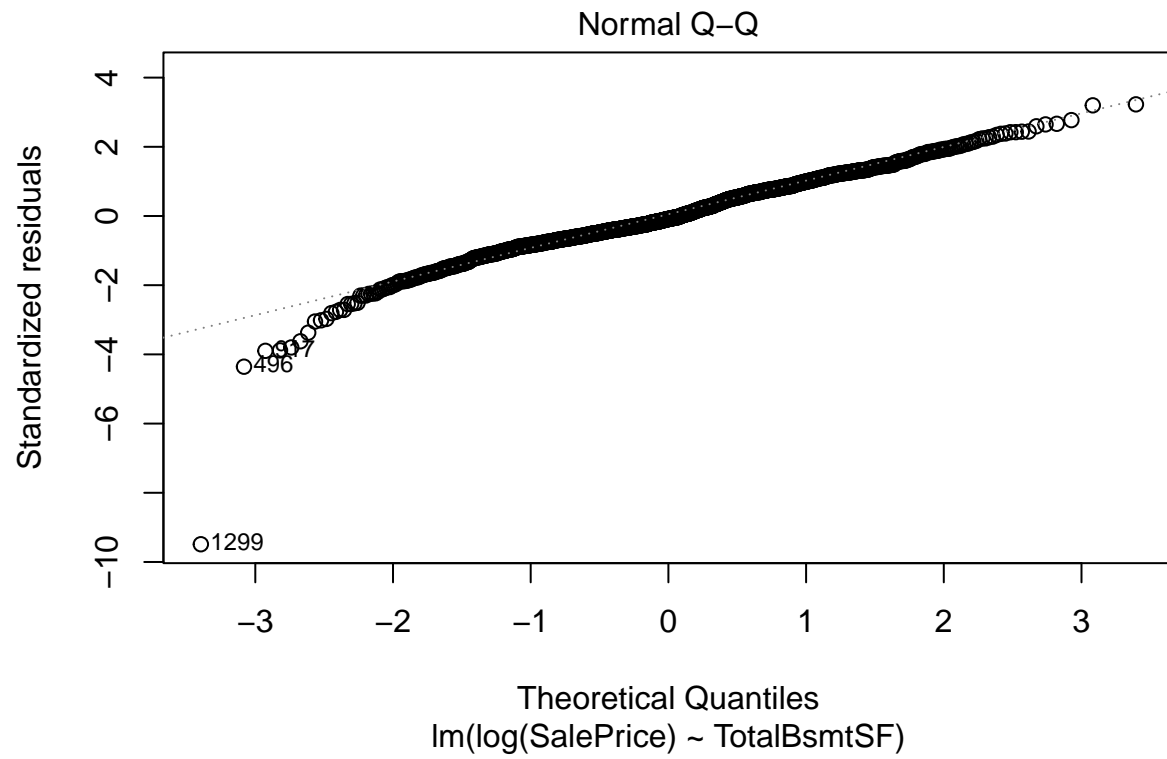


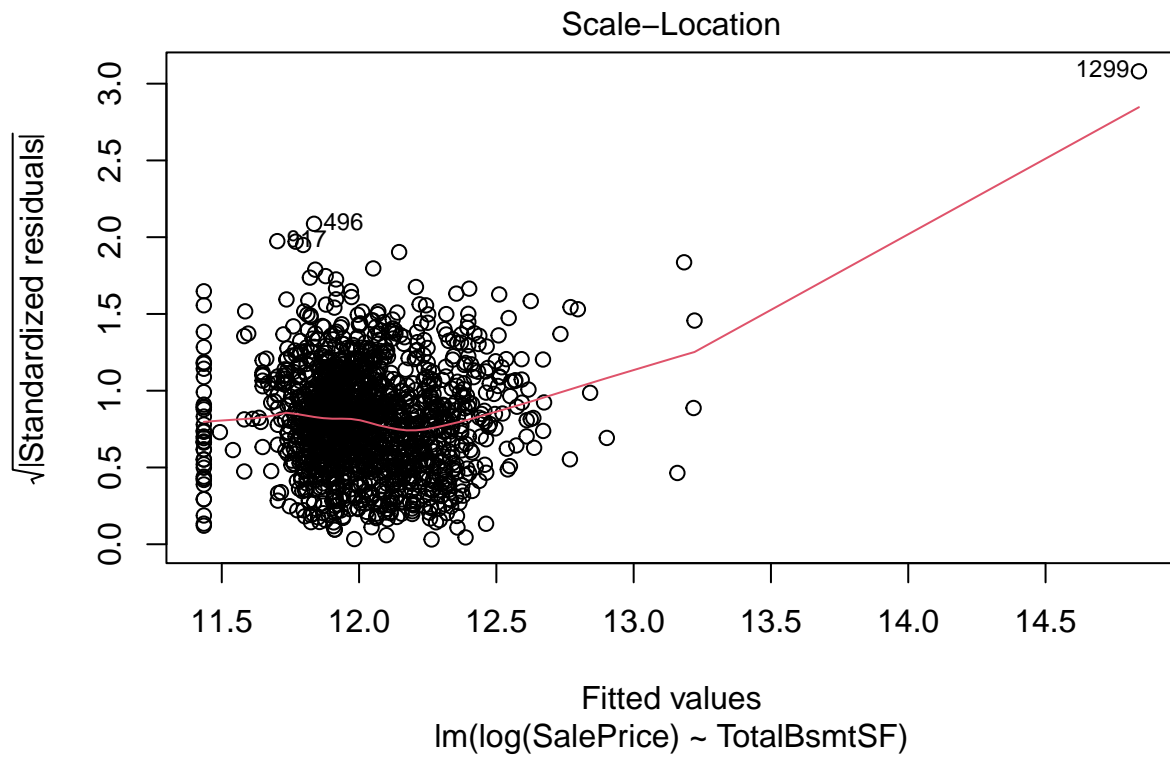


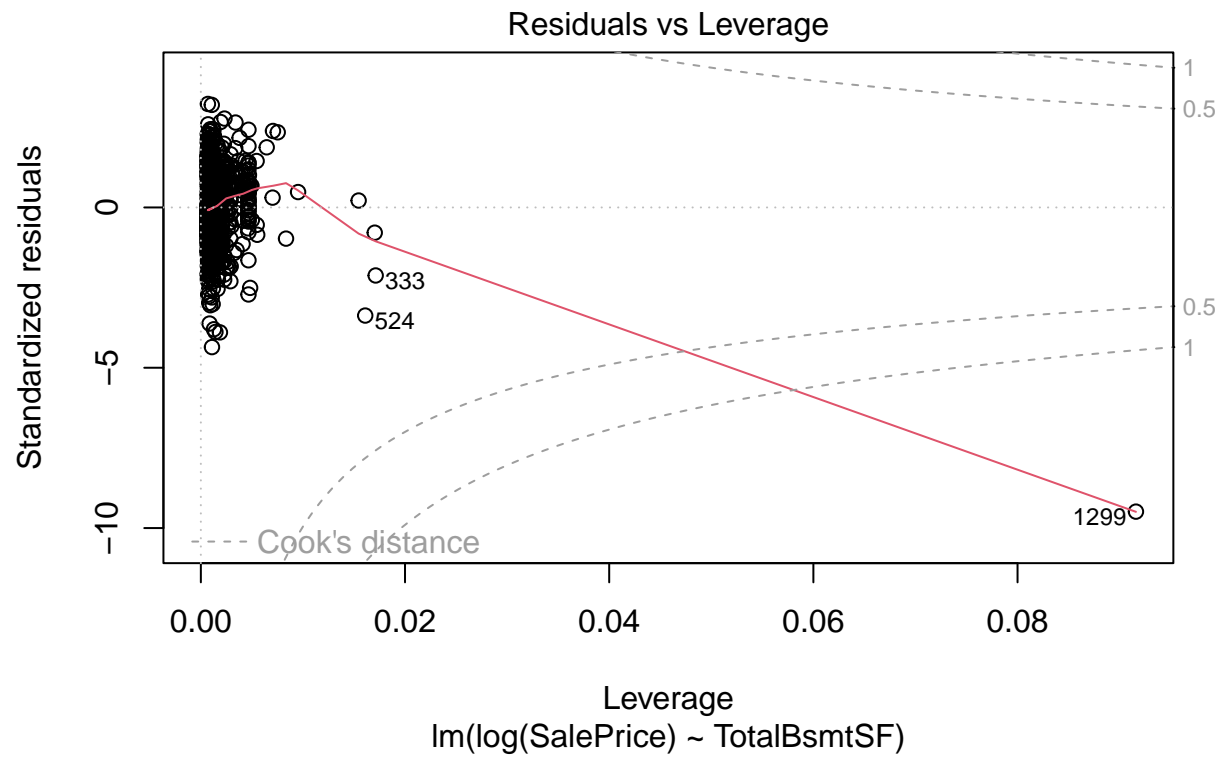












Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.