model1.R

r2278750

2023-11-26

```
# Multiple Linear Regression for Springbank Drive Data

# Clear Memory
rm(list = ls(all = TRUE))

# Import Data
dat <- read.csv("Springbank Drive Revised.csv", header = TRUE)

# Print column names on the screen
colnames(dat)
```

```
##  [1] "Property.."                      "Address"
##  [3] "Sales.Date"                      "HSETYPE"
##  [5] "One.and.a.Half.Storey"           "Two.Storey"
##  [7] "AGEYR"                           "LFA"
##  [9] "EXTAMEN"                         "Minor.Exterior.Amenities"
## [11] "Two.or.Three.Extra.Amenities"    "More.than.Three.Exterior.Amenities"
## [13] "EXTFINFACTOR"                    "Only.Brick"
## [15] "GAR"                             "Carport"
## [17] "One.Car.Garage"                  "Two.Car.Garage"
## [19] "STSCAPE"                         "Average.View"
## [21] "Good.View"                       "CENAIR"
## [23] "POOL"                            "INTCOND"
## [25] "Average.Interior.Condition"      "Good.Interior.Condition"
## [27] "Excellent.Interior.Condition"    "BSMTFINAREA"
## [29] "BI.AMEN.APPL"                    "LANESRD"
## [31] "TRAFCOUNT"                       "PRICE"
```

```
# Variables for analysis
saleprice <- dat,"PRICE"]  # Sale price of the property
lfa <- dat,"LFA"]  # Area of Frontage
bsmtfinarea <- dat,"BSMTFINAREA"]  # Basement Finished Area
ageyr <- dat,"AGEYR"]  # Age of house in years
lanesrd <- dat,"LANESRD"]  # Indicator for proximity to a four-lane road
onehalfstorey <- dat,"One.and.a.Half.Storey"]  # Indicator for one and a half storey houses
twostorey <- dat,"Two.Storey"]  # Indicator for two-storey houses
avgintcond <- dat,"Average.Interior.Condition"]  # Indicator for houses with average interior conditio
goodintcond <- dat,"Good.Interior.Condition"]  # Indicator for houses with good interior condition
excellentintcond <- dat,"Excellent.Interior.Condition"]  # Indicator for houses with excellent interio
avgview <- dat,"Average.View"]  # Indicator for houses with an average view
goodview <- dat,"Good.View"]  # Indicator for houses with a good view
pool <- dat,"POOL"]  # Indicator for presence of a pool
```

```r
# Generate a correlation matrix with the correct variables
cor(cbind(saleprice, lfa, bsmtfinarea, ageyr, lanesrd, onehalfstorey,
                        twostorey, avgintcond, goodintcond, excellentintcond,
                        avgview, goodview, pool))
```

```
##                   saleprice         lfa   bsmtfinarea        ageyr
## saleprice         1.00000000  0.40025782  0.2172499408 -0.134073554
## lfa               0.40025782  1.00000000 -0.0529680421  0.052779288
## bsmtfinarea       0.21724994 -0.05296804  1.0000000000 -0.505577176
## ageyr            -0.13407355  0.05277929 -0.5055771762  1.000000000
## lanesrd          -0.34800933 -0.17802569 -0.3007915909  0.130351688
## onehalfstorey     0.04097100  0.30732441 -0.3380775920  0.344408666
## twostorey        -0.06795999  0.13879021 -0.2743098231  0.294595253
## avgintcond       -0.21213447 -0.16718949  0.0004322962  0.005684107
## goodintcond       0.24754611  0.19742362  0.0642099698  0.033640944
## excellentintcond  0.17764607 -0.03540225  0.1057009596 -0.272081408
## avgview           0.03464353 -0.11178936  0.1749794978 -0.143459913
## goodview          0.23588527  0.11556279 -0.0619323057 -0.084446706
## pool              0.15390020 -0.03808901  0.1664145984 -0.132126654
##                       lanesrd onehalfstorey   twostorey    avgintcond
## saleprice        -0.348009325   0.040971000 -0.06795999 -0.2121344692
## lfa              -0.178025693   0.307324414  0.13879021 -0.1671894883
## bsmtfinarea      -0.300791591  -0.338077592 -0.27430982  0.0004322962
## ageyr             0.130351688   0.344408666  0.29459525  0.0056841075
## lanesrd           1.000000000   0.004206101  0.32776415  0.2732588620
## onehalfstorey     0.004206101   1.000000000 -0.13124359  0.0640917781
## twostorey         0.327764146  -0.131243592  1.00000000  0.0411909969
## avgintcond        0.273258862   0.064091778  0.04119100  1.0000000000
## goodintcond      -0.236085539  -0.146838801 -0.13449056 -0.6467269782
## excellentintcond -0.187542875   0.102927733 -0.15399810 -0.3279504344
## avgview          -0.025532276  -0.151757382 -0.06639061  0.0737000447
## goodview         -0.272594436   0.083195496 -0.09878983 -0.1580090351
## pool             -0.274873708  -0.121395396 -0.16817499  0.0381001259
##                  goodintcond excellentintcond     avgview    goodview
## saleprice         0.24754611       0.17764607  0.03464353  0.23588527
## lfa               0.19742362      -0.03540225 -0.11178936  0.11556279
## bsmtfinarea       0.06420997       0.10570096  0.17497950 -0.06193231
## ageyr             0.03364094      -0.27208141 -0.14345991 -0.08444671
## lanesrd          -0.23608554      -0.18754287 -0.02553228 -0.27259444
## onehalfstorey    -0.14683880       0.10292773 -0.15175738  0.08319550
## twostorey        -0.13449056      -0.15399810 -0.06639061 -0.09878983
## avgintcond       -0.64672698      -0.32795043  0.07370004 -0.15800904
## goodintcond       1.00000000      -0.25722086  0.03326739  0.06175986
## excellentintcond -0.25722086       1.00000000 -0.05623216  0.24265446
## avgview           0.03326739      -0.05623216  1.00000000 -0.62225830
## goodview          0.06175986       0.24265446 -0.62225830  1.00000000
## pool              0.07681764      -0.05426380  0.10965862 -0.00652692
##                         pool
## saleprice         0.15390020
## lfa              -0.03808901
## bsmtfinarea       0.16641460
## ageyr            -0.13212665
## lanesrd          -0.27487371
## onehalfstorey    -0.12139540
```

```
## twostorey        -0.16817499
## avgintcond        0.03810013
## goodintcond       0.07681764
## excellentintcond -0.05426380
## avgview           0.10965862
## goodview         -0.00652692
## pool              1.00000000
```

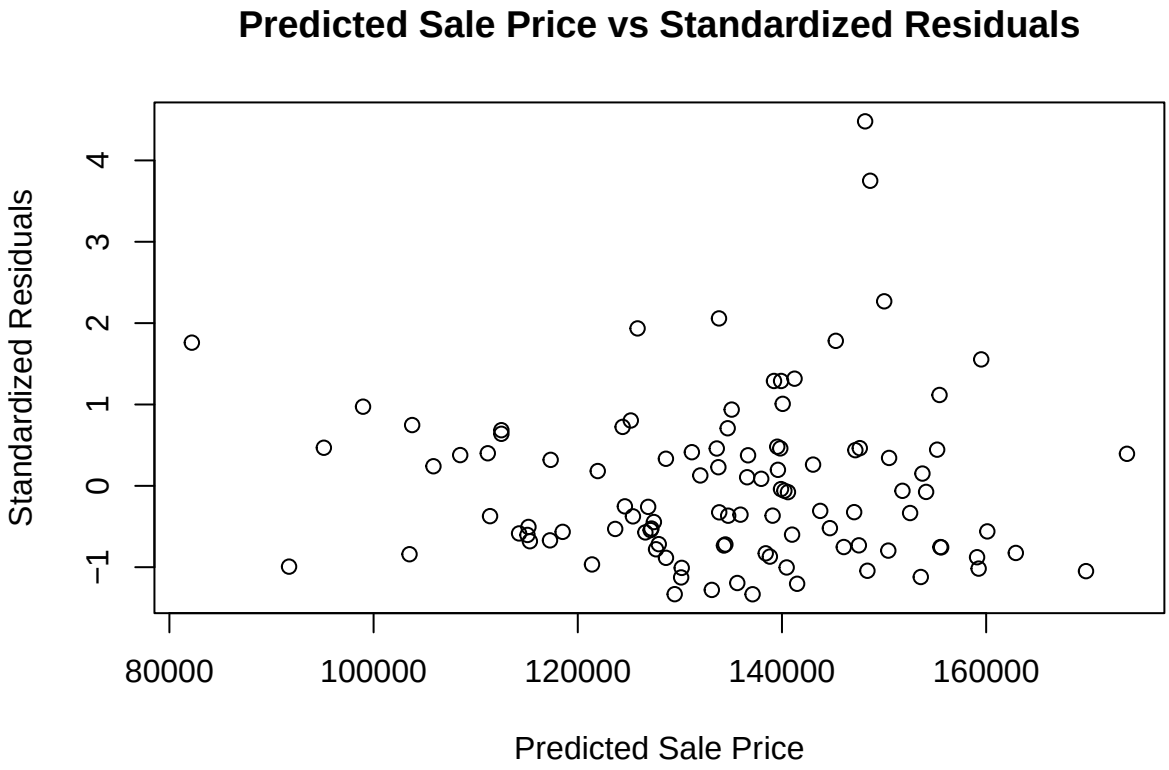
```r
# Multiple Linear Regression: Sales Price on multiple independent variables
mod_mlr <- lm(saleprice ~ lfa + bsmtfinarea + ageyr + lanesrd + onehalfstorey +
                twostorey + avgintcond + goodintcond + excellentintcond + avgview +
                goodview + pool)

# Present Parameter Estimates, Coefficient of Determination, etc.
summary(mod_mlr)
```

```
##
## Call:
## lm(formula = saleprice ~ lfa + bsmtfinarea + ageyr + lanesrd +
##     onehalfstorey + twostorey + avgintcond + goodintcond + excellentintcond +
##     avgview + goodview + pool)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -29591 -16424  -6319   9819 101758
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      65552.19   19333.37   3.391 0.001033 **
## lfa                 37.87      10.68   3.548 0.000617 ***
## bsmtfinarea         16.97      10.33   1.643 0.103770
## ageyr               76.99     197.51   0.390 0.697576
## lanesrd          -5088.84    6273.63  -0.811 0.419397
## onehalfstorey     1605.69   10752.69   0.149 0.881624
## twostorey         6265.54    8035.30   0.780 0.437560
## avgintcond       10079.03    8784.87   1.147 0.254258
## goodintcond      18481.20    9283.56   1.991 0.049509 *
## excellentintcond 23543.33   11510.44   2.045 0.043704 *
## avgview          11001.05    6463.46   1.702 0.092163 .
## goodview         17263.15    8113.22   2.128 0.036059 *
## pool              8038.85    7218.62   1.114 0.268372
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23450 on 91 degrees of freedom
## Multiple R-squared:  0.3686, Adjusted R-squared:  0.2854
## F-statistic: 4.427 on 12 and 91 DF,  p-value: 1.509e-05
```

```r
# Extract standardized residuals and predicted values
standardized_residuals <- rstandard(mod_mlr)
predicted_saleprice <- predict(mod_mlr)

# Plot Predicted Sales Price vs Standardized Residuals
plot(predicted_saleprice, standardized_residuals, xlab = "Predicted Sale Price",
     ylab = "Standardized Residuals", main = "Predicted Sale Price vs Standardized Residuals")
```


Predicted Sale Price vs Standardized Residuals
Standardized Residuals
−1
0
1
2
3
4
80000
100000
120000
140000
160000
Predicted Sale Price