

House_Price Project for King County Assessor

2022-11-30

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
# Load training data
train <- readRDS("data/train.rds")
dat <- readRDS("data/train.rds")

ZipGroup <- dat$ZipCode

mod4 <- lm(AdjSalePrice ~ SqFtTotLiving + BldgGrade + ZipGroup, data = train)

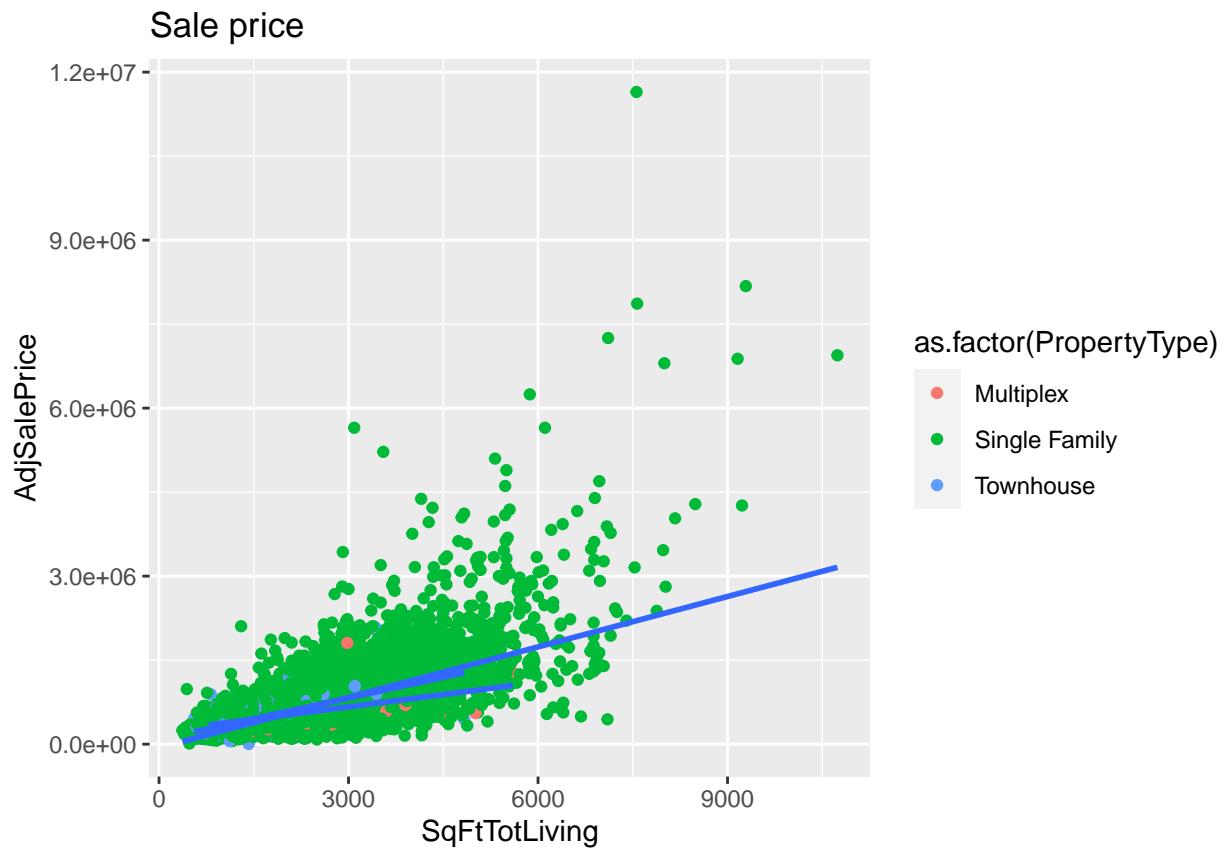
summary(mod4)

##
## Call:
## lm(formula = AdjSalePrice ~ SqFtTotLiving + BldgGrade + ZipGroup,
##      data = train)
##
## Residuals:
##       Min     1Q   Median     3Q    Max 
## -1247420 -120491  -21479   88469  9466886 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -6.419e+07  3.534e+06 -18.16   <2e-16 ***
## SqFtTotLiving 1.867e+02  3.169e+00  58.92   <2e-16 ***
## BldgGrade     1.157e+05  2.453e+03  47.16   <2e-16 ***
## ZipGroup      6.472e+02  3.601e+01  17.97   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 262900 on 20336 degrees of freedom
## Multiple R-squared:  0.5391, Adjusted R-squared:  0.539 
## F-statistic:  7929 on 3 and 20336 DF,  p-value: < 2.2e-16
```

Including Plots

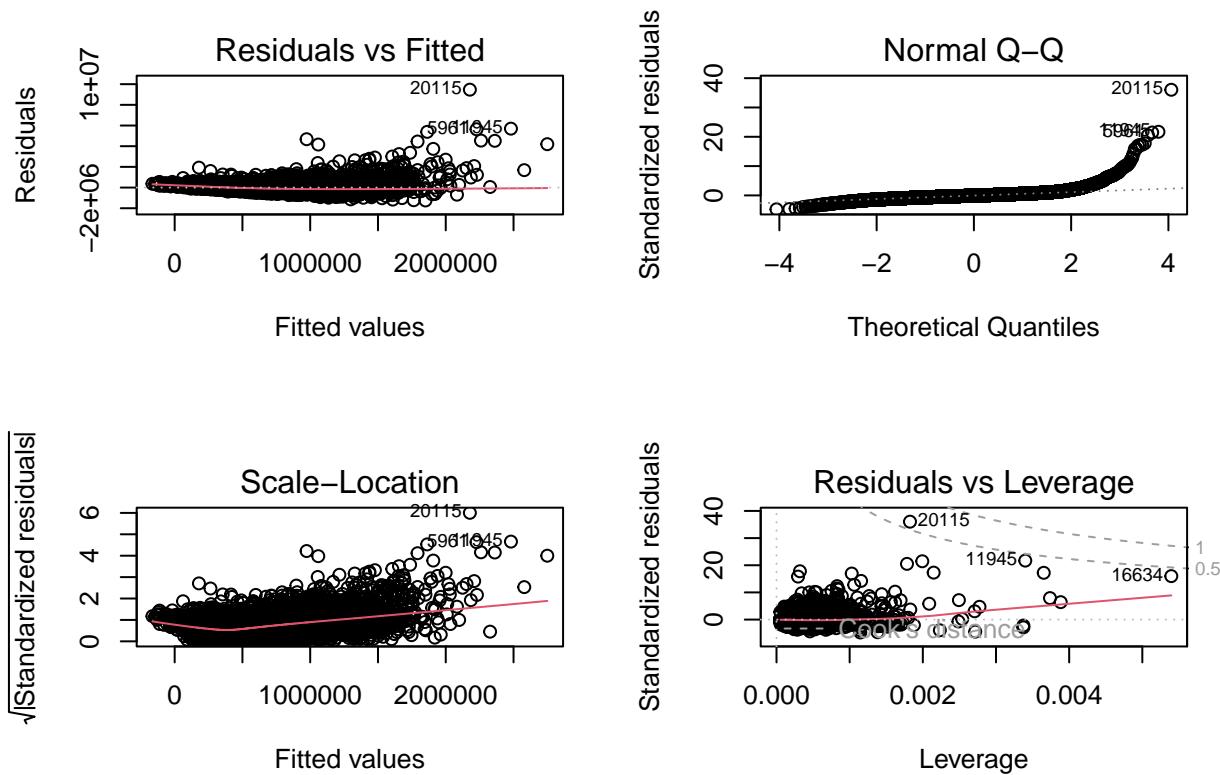
Plots for the first model:

```
## `geom_smooth()` using formula = 'y ~ x'
```



Outlier Identification

```
par(mfrow = c(2,2))
plot(mod4)
```



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