

# Predicting House Prices

## Kaggle Competition

This Kaggle competition is about predicting house prices based on a set of around 80 predictor variables. Please read the brief description of the project and get familiar with the various predictors. We will have to do some initial cleaning to successfully work with these data. Overall, we (in teams) will use the provided training dataset to build a multiple linear regression model for predicting house prices. Once we have settled on a final model, we will use it with the predictors available in the testing dataset to predict house prices. The goal of the competition mentions that our predictions  $\hat{y}_i$  for the houses in the testing data are compared to the (withheld) true selling prices  $y_i^{\text{test}}$  via  $\sum_i (\log \hat{y}_i - \log y_i^{\text{test}})^2$ . Because selling prices are typically right-skewed, I think as a first step we will log-transform the selling prices of the houses in the training data to obtain a more bell-shaped distribution. However, although we will build a model for the log-prices, we will still have to submit the price of a house (and not the log-price) to Kaggle, together with the ID of the house.

## Loading and inspecting the train and test datasets

```
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.1.3

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2

## Warning: package 'ggplot2' was built under R version 4.1.3
## Warning: package 'tibble' was built under R version 4.1.3
## Warning: package 'tidyr' was built under R version 4.1.3
## Warning: package 'readr' was built under R version 4.1.3
## Warning: package 'dplyr' was built under R version 4.1.3
## Warning: package 'stringr' was built under R version 4.1.3
## Warning: package 'forcats' was built under R version 4.1.3

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
## Load Training Data
path_traindata <- 'https://raw.githubusercontent.com/bklingen/Price-Prediction/main/train.csv'
train <- read_csv(path_traindata)
```

```
## Rows: 1460 Columns: 81
## -- Column specification -----
## Delimiter: ","
## chr (43): MSZoning, Street, Alley, LotShape, LandContour, Utilities, LotConf...
## dbl (38): Id, MSSubClass, LotFrontage, LotArea, OverallQual, OverallCond, Ye...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
dim(train)
```

```
## [1] 1460 81
```

```
## Load Testing Data
path_testdata <- 'https://raw.githubusercontent.com/bklingen/Price-Prediction/main/test.csv'
test <- read_csv(path_testdata)
```

```
## Rows: 1459 Columns: 80
## -- Column specification -----
## Delimiter: ","
## chr (43): MSZoning, Street, Alley, LotShape, LandContour, Utilities, LotConf...
## dbl (37): Id, MSSubClass, LotFrontage, LotArea, OverallQual, OverallCond, Ye...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
dim(test)
```

```
## [1] 1459 80
```

```
SalePrice <- train$SalePrice
```

This makes sense: We have one less column in test data because of the missing house prices.

But, are the column names the same? Let's find the "difference" between two sets: All the column names that are in the test data but not in the train data:

```
setdiff(colnames(test), colnames(train))
```

```
## character(0)
```

OK, good, and now the other way around:

```
setdiff(colnames(train), colnames(test))
```

```
## [1] "SalePrice"
```

OK, great. So no surprises there. All predictors that exist in the train data set also appear in the test dataset.

Let's see how many quantitative and how many categorical predictors we have in the training dataset, at least at face value:

```
train_quantPredictors = train %>% select(where(is.numeric)) %>% select(-SalePrice)
train_catPredictors = train %>% select(where(is.character))
dim(train_quantPredictors)
```

```
## [1] 1460 37
```

```
dim(train_catPredictors)
```

```
## [1] 1460 43
```

Let's quickly do the same split for the test data:

```
test_quantPredictors = test %>% select(where(is.numeric))
test_catPredictors = test %>% select(where(is.character))
```

Let's transform the categorical predictors into factors, which should make it easier to combine categories, create a category like "other", etc.

```
train_catPredictors = train_catPredictors %>% transmute_all(as.factor)
test_catPredictors = test_catPredictors %>% transmute_all(as.factor)
```

First, let's see the category names and frequency for each variable:

```
for(i in 1:ncol(train_catPredictors)) {
  print(colnames(train_catPredictors)[i])
  print("----")
  print(as.data.frame(fct_count(unlist(train_catPredictors[,i]))))
  print("-----")
}
```

```
## [1] "MSZoning"
## [1] "----"
##      f      n
## 1 C (all)  10
## 2      FV   65
## 3      RH   16
## 4      RL 1151
## 5      RM  218
## [1] "-----"
## [1] "Street"
## [1] "----"
##      f      n
## 1 Grvl    6
## 2 Pave 1454
## [1] "-----"
```

```

## [1] "Alley"
## [1] "-----"
##      f      n
## 1 Grvl   50
## 2 Pave   41
## 3 <NA> 1369
## [1] "-----"
## [1] "LotShape"
## [1] "-----"
##      f      n
## 1 IR1 484
## 2 IR2  41
## 3 IR3  10
## 4 Reg 925
## [1] "-----"
## [1] "LandContour"
## [1] "-----"
##      f      n
## 1 Bnk   63
## 2 HLS   50
## 3 Low   36
## 4 Lvl 1311
## [1] "-----"
## [1] "Utilities"
## [1] "-----"
##      f      n
## 1 AllPub 1459
## 2 NoSeWa  1
## [1] "-----"
## [1] "LotConfig"
## [1] "-----"
##      f      n
## 1 Corner 263
## 2 CulDSac 94
## 3      FR2 47
## 4      FR3  4
## 5 Inside 1052
## [1] "-----"
## [1] "LandSlope"
## [1] "-----"
##      f      n
## 1 Gtl 1382
## 2 Mod  65
## 3 Sev  13
## [1] "-----"
## [1] "Neighborhood"
## [1] "-----"
##      f      n
## 1 Blmngtn 17
## 2 Blueste  2
## 3 BrDale  16
## 4 BrkSide 58
## 5 ClearCr 28
## 6 CollgCr 150

```

```

## 7  Crawfor  51
## 8  Edwards 100
## 9  Gilbert  79
## 10 IDOTRR  37
## 11 MeadowV 17
## 12 Mitchel 49
## 13  NAmes  225
## 14 NoRidge 41
## 15 NPKvill  9
## 16 NridgHt 77
## 17  NWAmes 73
## 18 OldTown 113
## 19  Sawyer 74
## 20 SawyerW 59
## 21 Somerst 86
## 22 StoneBr 25
## 23  SWISU  25
## 24  Timber 38
## 25 Veenker 11
## [1] "-----"
## [1] "Condition1"
## [1] "-----"
##      f      n
## 1 Artery   48
## 2  Feedr   81
## 3   Norm 1260
## 4   PosA    8
## 5   PosN   19
## 6   RRAe   11
## 7   RRAn   26
## 8   RRNe    2
## 9   RRNn    5
## [1] "-----"
## [1] "Condition2"
## [1] "-----"
##      f      n
## 1 Artery    2
## 2  Feedr    6
## 3   Norm 1445
## 4   PosA    1
## 5   PosN    2
## 6   RRAe    1
## 7   RRAn    1
## 8   RRNn    2
## [1] "-----"
## [1] "BldgType"
## [1] "-----"
##      f      n
## 1   1Fam 1220
## 2  2fmCon  31
## 3 Duplex   52
## 4  Twnhs   43
## 5  TwnhsE 114
## [1] "-----"

```

```

## [1] "HouseStyle"
## [1] "-----"
##      f      n
## 1 1.5Fin 154
## 2 1.5Unf  14
## 3 1Story 726
## 4 2.5Fin   8
## 5 2.5Unf  11
## 6 2Story 445
## 7 SFoyer  37
## 8  SLvl  65
## [1] "-----"
## [1] "RoofStyle"
## [1] "-----"
##      f      n
## 1   Flat  13
## 2   Gable 1141
## 3 Gambrel  11
## 4    Hip  286
## 5 Mansard   7
## 6   Shed   2
## [1] "-----"
## [1] "RoofMat1"
## [1] "-----"
##      f      n
## 1 ClyTile   1
## 2 CompShg 1434
## 3 Membran   1
## 4   Metal   1
## 5    Roll   1
## 6 Tar&Grv  11
## 7 WdShake   5
## 8 WdShngl   6
## [1] "-----"
## [1] "Exterior1st"
## [1] "-----"
##      f      n
## 1 AsbShng  20
## 2 AsphShn   1
## 3 BrkComm   2
## 4 BrkFace  50
## 5  CBlock   1
## 6 CemntBd  61
## 7 HdBoard 222
## 8 ImStucc   1
## 9 MetalSd 220
## 10 Plywood 108
## 11  Stone   2
## 12 Stucco  25
## 13 VinylSd 515
## 14 Wd Sdng 206
## 15 WdShing  26
## [1] "-----"
## [1] "Exterior2nd"

```

```

## [1] "----"
##      f      n
## 1 AsbShng 20
## 2 AsphShn  3
## 3 Brk Cmn  7
## 4 BrkFace 25
## 5  CBlock  1
## 6 CmentBd 60
## 7 HdBoard 207
## 8 ImStucc 10
## 9 MetalSd 214
## 10 Other  1
## 11 Plywood 142
## 12 Stone   5
## 13 Stucco  26
## 14 VinylSd 504
## 15 Wd Sdng 197
## 16 Wd Shng  38
## [1] "-----"
## [1] "MasVnrType"
## [1] "----"
##      f      n
## 1 BrkCmn 15
## 2 BrkFace 445
## 3   None 864
## 4   Stone 128
## 5   <NA>  8
## [1] "-----"
## [1] "ExterQual"
## [1] "----"
##      f      n
## 1 Ex   52
## 2 Fa   14
## 3 Gd 488
## 4 TA 906
## [1] "-----"
## [1] "ExterCond"
## [1] "----"
##      f      n
## 1 Ex    3
## 2 Fa   28
## 3 Gd  146
## 4 Po    1
## 5 TA 1282
## [1] "-----"
## [1] "Foundation"
## [1] "----"
##      f      n
## 1 BrkTil 146
## 2 CBlock 634
## 3 PConc 647
## 4  Slab  24
## 5 Stone   6
## 6  Wood   3

```

```

## [1] "-----"
## [1] "BsmtQual"
## [1] "-----"
##      f      n
## 1   Ex 121
## 2   Fa  35
## 3   Gd 618
## 4   TA 649
## 5 <NA>  37
## [1] "-----"
## [1] "BsmtCond"
## [1] "-----"
##      f      n
## 1   Fa  45
## 2   Gd  65
## 3   Po   2
## 4   TA 1311
## 5 <NA>  37
## [1] "-----"
## [1] "BsmtExposure"
## [1] "-----"
##      f      n
## 1   Av 221
## 2   Gd 134
## 3   Mn 114
## 4   No 953
## 5 <NA>  38
## [1] "-----"
## [1] "BsmtFinType1"
## [1] "-----"
##      f      n
## 1  ALQ 220
## 2  BLQ 148
## 3  GLQ 418
## 4  LwQ  74
## 5  Rec 133
## 6  Unf 430
## 7 <NA>  37
## [1] "-----"
## [1] "BsmtFinType2"
## [1] "-----"
##      f      n
## 1  ALQ  19
## 2  BLQ  33
## 3  GLQ  14
## 4  LwQ  46
## 5  Rec  54
## 6  Unf 1256
## 7 <NA>  38
## [1] "-----"
## [1] "Heating"
## [1] "-----"
##      f      n
## 1 Floor  1

```



```

## 2 GasA 1428
## 3 GasW 18
## 4 Grav 7
## 5 OthW 2
## 6 Wall 4
## [1] "-----"
## [1] "HeatingQC"
## [1] "----"
## f n
## 1 Ex 741
## 2 Fa 49
## 3 Gd 241
## 4 Po 1
## 5 TA 428
## [1] "-----"
## [1] "CentralAir"
## [1] "----"
## f n
## 1 N 95
## 2 Y 1365
## [1] "-----"
## [1] "Electrical"
## [1] "----"
## f n
## 1 FuseA 94
## 2 FuseF 27
## 3 FuseP 3
## 4 Mix 1
## 5 SBrkr 1334
## 6 <NA> 1
## [1] "-----"
## [1] "KitchenQual"
## [1] "----"
## f n
## 1 Ex 100
## 2 Fa 39
## 3 Gd 586
## 4 TA 735
## [1] "-----"
## [1] "Functional"
## [1] "----"
## f n
## 1 Maj1 14
## 2 Maj2 5
## 3 Min1 31
## 4 Min2 34
## 5 Mod 15
## 6 Sev 1
## 7 Typ 1360
## [1] "-----"
## [1] "FireplaceQu"
## [1] "----"
## f n
## 1 Ex 24

```

```

## 2 Fa 33
## 3 Gd 380
## 4 Po 20
## 5 TA 313
## 6 <NA> 690
## [1] "-----"
## [1] "GarageType"
## [1] "-----"
## f n
## 1 2Types 6
## 2 Attchd 870
## 3 Basment 19
## 4 BuiltIn 88
## 5 CarPort 9
## 6 Detchd 387
## 7 <NA> 81
## [1] "-----"
## [1] "GarageFinish"
## [1] "-----"
## f n
## 1 Fin 352
## 2 RFn 422
## 3 Unf 605
## 4 <NA> 81
## [1] "-----"
## [1] "GarageQual"
## [1] "-----"
## f n
## 1 Ex 3
## 2 Fa 48
## 3 Gd 14
## 4 Po 3
## 5 TA 1311
## 6 <NA> 81
## [1] "-----"
## [1] "GarageCond"
## [1] "-----"
## f n
## 1 Ex 2
## 2 Fa 35
## 3 Gd 9
## 4 Po 7
## 5 TA 1326
## 6 <NA> 81
## [1] "-----"
## [1] "PavedDrive"
## [1] "-----"
## f n
## 1 N 90
## 2 P 30
## 3 Y 1340
## [1] "-----"
## [1] "PoolQC"
## [1] "-----"

```

```

##      f      n
## 1   Ex      2
## 2   Fa      2
## 3   Gd      3
## 4 <NA> 1453
## [1] "-----"
## [1] "Fence"
## [1] "-----"
##      f      n
## 1 GdPrv    59
## 2  GdWo    54
## 3 MnPrv   157
## 4  MnWw    11
## 5 <NA> 1179
## [1] "-----"
## [1] "MiscFeature"
## [1] "-----"
##      f      n
## 1 Gar2      2
## 2 Othr      2
## 3 Shed     49
## 4 TenC      1
## 5 <NA> 1406
## [1] "-----"
## [1] "SaleType"
## [1] "-----"
##      f      n
## 1   COD     43
## 2   Con      2
## 3 ConLD      9
## 4 ConLI      5
## 5 ConLw      5
## 6   CWD      4
## 7   New    122
## 8   Oth      3
## 9   WD   1267
## [1] "-----"
## [1] "SaleCondition"
## [1] "-----"
##      f      n
## 1 Abnorml  101
## 2 AdjLand    4
## 3 Alloca    12
## 4 Family    20
## 5 Normal  1198
## 6 Partial   125
## [1] "-----"

```

## Handle Categorical Features

### MSZoning (Mei)

There are no null/missing values in the training set, but there are a few in the test set

```
sum(is.na(train$MSZoning))
```

```
## [1] 0
```

```
sum(is.na(test$MSZoning))
```

```
## [1] 4
```

Although there are 8 potential categories for this variable, there only exist 5 unique ones in the training and test set.

```
fct_count(train$MSZoning)
```

```
## # A tibble: 5 x 2
##   f           n
##   <fct>   <int>
## 1 C (all)    10
## 2 FV         65
## 3 RH         16
## 4 RL       1151
## 5 RM        218
```

```
fct_count(test$MSZoning)
```

```
## # A tibble: 6 x 2
##   f           n
##   <fct>   <int>
## 1 C (all)    15
## 2 FV         74
## 3 RH         10
## 4 RL       1114
## 5 RM        242
## 6 <NA>         4
```

```
mszoning.collapse <- function(x) fct_collapse(x,
  "FV" = c("FV"),
  "RL" = c("RL", "RP"),
  "RO" = c("RM", "RH"),
  other_level = "other")
```

```
train <- train %>% mutate(MSZoning = as.factor(MSZoning), MSZoning = mszoning.collapse(MSZoning))
test <- test %>% mutate(MSZoning = as.factor(MSZoning), MSZoning = mszoning.collapse(MSZoning))
```

```
fct_count(train$MSZoning)
```

```
## # A tibble: 4 x 2
##   f           n
##   <fct> <int>
## 1 FV         65
## 2 RO        234
## 3 RL       1151
## 4 other      10
```

## MSSubClass (Mei)

There are no null/missing values

```
sum(is.na(train$MSSubClass))
```

```
## [1] 0
```

```
sum(is.na(test$MSSubClass))
```

```
## [1] 0
```

Assuming the 1/2 story refers to a basement level as “(un)finished” terminology typically refers to, the categories will be split as follows (counts in parenthesis): - 1-STORY 1946 & NEWER single-family (536) - 1-STORY single-family other - 30 1-STORY 1945 & OLDER (69) - 40 1-STORY W/FINISHED ATTIC ALL AGES (4) - 45 1-1/2 STORY - UNFINISHED ALL AGES (12) - 50 1-1/2 STORY FINISHED ALL AGES (144) - multi-level single-family non PUD - 60 2-STORY 1946 & NEWER (299) - 70 2-STORY 1945 & OLDER (60) - 75 2-1/2 STORY ALL AGES (16) - 80 SPLIT OR MULTI-LEVEL (58) - 85 SPLIT FOYER (20) - other - 90 DUPLEX - ALL STYLES AND AGES (52) - 120 1-STORY PUD (Planned Unit Development) - 1946 & NEWER (87) - 150 1-1/2 STORY PUD - ALL AGES - 160 2-STORY PUD - 1946 & NEWER (63) - 180 PUD - MULTILEVEL - INCL SPLIT LEV/FOYER (10) - 190 2 FAMILY CONVERSION - ALL STYLES AND AGES (30)

```
mssubclass.collapse <- function(x) fct_collapse(x,  
  "1-story single-family 1946 & newer" = c("20"),  
  "1-story single-family other" = c("30", "40", "45", "50"),  
  "multi-level single-family non PUD" = c("60", "70", "75", "80", "85"),  
  other_level = "other")
```

```
train <- train %>% mutate(MSSubClass = as.factor(MSSubClass), MSSubClass = mssubclass.collapse(MSSubClass))  
test <- test %>% mutate(MSSubClass = as.factor(MSSubClass), MSSubClass = mssubclass.collapse(MSSubClass))
```

```
fct_count(train$MSSubClass)
```

```
## # A tibble: 4 x 2  
##   f                n  
##   <fct>          <int>  
## 1 1-story single-family 1946 & newer    536  
## 2 1-story single-family other          229  
## 3 multi-level single-family non PUD    453  
## 4 other                          242
```

## Condition1/Condition2 (Mei)

There are no null/missing values

```
sum(is.na(train$Condition1))
```

```
## [1] 0
```

```
sum(is.na(test$Condition1))
```

```
## [1] 0
```

```
sum(is.na(train$Condition2))
```

```
## [1] 0
```

```
sum(is.na(test$Condition2))
```

```
## [1] 0
```

Collapse similar locations together: - All the railroad related locations - All the park related locations - All the street related locations This results in only 4 categories: - Normal - Near railroad - Near park - Near arterial or feeder street

```
condition.collapse <- function(x) fct_collapse(x,
  RR = c("RRNn", "RRAn", "RRNe", "RR Ae"),
  Pos = c("PosN", "PosA"),
  St = c("Artery", "Feedr"))

train <- train %>% mutate_at(vars(Condition1, Condition2), condition.collapse)
test <- test %>% mutate_at(vars(Condition1, Condition2), condition.collapse)
```

```
fct_count(train$Condition1)
```

```
## # A tibble: 4 x 2
##   f         n
##   <fct> <int>
## 1 St      129
## 2 Norm    1260
## 3 Pos      27
## 4 RR       44
```

## Richard's Features

### RoofStyle

combine flat, shed as other; gambrel, mansard, gable as gable; leave others as is

```
roof_price <- train %>% group_by(RoofStyle) %>% summarize(count=n(),
  mean(SalePrice), sd(SalePrice))
```

```
roof_price
```

```
## # A tibble: 6 x 4
##   RoofStyle count 'mean(SalePrice)' 'sd(SalePrice)'
##   <chr>      <int>          <dbl>          <dbl>
```

```
## 1 Flat      13      194690      62523.
## 2 Gable    1141      171484.      66331.
## 3 Gambrel   11      148909.      67014.
## 4 Hip      286      218877.      111550.
## 5 Mansard    7      180568.      58058.
## 6 Shed       2      225000      49497.
```

```
train$RoofStyle <- fct_collapse(train$RoofStyle, Other = c("Flat", "Shed"))
train$RoofStyle <- fct_collapse(train$RoofStyle, Gable = c("Gable", "Gambrel", "Mansard"))
```

Let's do the same on the testing dataset:

```
test$RoofStyle <- fct_collapse(test$RoofStyle, Other = c("Flat", "Shed"))
test$RoofStyle <- fct_collapse(test$RoofStyle, Gable = c("Gable", "Gambrel", "Mansard"))
```

## BldgType

Combine 2FmCon, Duplex as multifamily; leave others as is

```
bldg_price <- train %>% group_by(BldgType) %>% summarize(count=n(),
  mean(SalePrice), sd(SalePrice))
```

```
bldg_price
```

```
## # A tibble: 5 x 4
##   BldgType count 'mean(SalePrice)' 'sd(SalePrice)'
##   <chr>    <int>          <dbl>          <dbl>
## 1 1Fam      1220      185764.      82649.
## 2 2fmCon     31      128432.      35459.
## 3 Duplex     52      133541.      27833.
## 4 Twnhs     43      135912.      41013.
## 5 TwnhsE    114      181959.      60626.
```

```
train$BldgType <- fct_collapse(train$BldgType, MultiFam = c("2fmCon", "Duplex"))
```

Let's do the same on the testing dataset:

```
test$BldgType <- fct_collapse(test$BldgType, MultiFam = c("2fmCon", "Duplex"))
```

## HouseStyle

Combine 1.5Fin, 1Story, split foyer, split level as less than 2 story; 2.5fin, 2Story as two story or greater; leave 1.5Unf and 2.5Unf as is since they drag down property values

```
style_price <- train %>% group_by(HouseStyle) %>% summarize(count=n(),
  mean(SalePrice), sd(SalePrice))
```

```
style_price
```

```
## # A tibble: 8 x 4
##   HouseStyle count 'mean(SalePrice)' 'sd(SalePrice)'
##   <chr>      <int>      <dbl>      <dbl>
## 1 1.5Fin      154      143117.    54278.
## 2 1.5Unf       14      110150     19036.
## 3 1Story      726      175985.    77056.
## 4 2.5Fin        8      220000     118212.
## 5 2.5Unf       11      157355.    63934.
## 6 2Story      445      210052.    87339.
## 7 SFoyer       37      135074.    30481.
## 8 SLvl        65      166703.    38305.
```

```
train$HouseStyle <- fct_collapse(train$HouseStyle, Less2story = c("1Story", "1.5Fin", "SFoyer", "SLvl"))
train$HouseStyle <- fct_collapse(train$HouseStyle, EqMore2story = c("2Story", "2.5Fin"))
```

And on the test data:

```
test$HouseStyle <- fct_collapse(test$HouseStyle, Less2story = c("1Story", "1.5Fin", "SFoyer", "SLvl"))
test$HouseStyle <- fct_collapse(test$HouseStyle, EqMore2story = c("2Story", "2.5Fin"))
```

```
## Warning: Unknown levels in 'f': 2.5Fin
```

Kyle:

```
cleanpool <- as.character(train_catPredictors$PoolQC)
cleanpool[is.na(cleanpool)] <- "none"
cleanpool <- as.factor(cleanpool)
```

```
cleanfence <- as.character(train_catPredictors$Fence)
cleanfence[is.na(cleanfence)] <- "none"
cleanfence <- as.factor(cleanfence)
```

```
cleanfunc <- as.character(train_catPredictors$Functional)
cleanfunc[cleanfunc == 'Min1' | cleanfunc == 'Min2'] <- "Minor"
cleanfunc[cleanfunc == 'Maj1' | cleanfunc == 'Maj2'] <- "Major"
cleanfunc[cleanfunc == 'Sev' | cleanfunc == 'Sal'] <- "Severe"
cleanfunc <- as.factor(cleanfunc)
```

```
train$PoolQC <- cleanpool
train$Fence <- cleanfence
train$Functional <- cleanfunc
```

We need to do the same for the test dataset, so I just copied the code block and replaced “train” by “test”:

```
cleanpool <- as.character(test_catPredictors$PoolQC)
cleanpool[is.na(cleanpool)] <- "none"
cleanpool <- as.factor(cleanpool)
```



```
cleanfence <- as.character(test_catPredictors$Fence)
cleanfence[is.na(cleanfence)] <- "none"
cleanfence <- as.factor(cleanfence)
```

```
cleanfunc <- as.character(test_catPredictors$Functional)
cleanfunc[cleanfunc == 'Min1' | cleanfunc == 'Min2'] <- "Minor"
cleanfunc[cleanfunc == 'Maj1' | cleanfunc == 'Maj2'] <- "Major"
cleanfunc[cleanfunc == 'Sev' | cleanfunc == 'Sal'] <- "Severe"
cleanfunc <- as.factor(cleanfunc)
```

```
test$PoolQC <- cleanpool
test$Fence <- cleanfence
test$Functional <- cleanfunc
```

### Mileva: Heating, Electrical, FireplaceQu, HeatingQC, CentralAir

The processing for the Heating, Electrical, and FireplaceQu predictors is below. The HeatingQC and CentralAir predictors did not require any additional processing.

```
# Heating: Collapsed categories with low frequencies into "other"
heating <- as.factor(train_catPredictors$Heating)
heating <- fct_other(heating, keep=c("GasA", "GasW"))
train$Heating <- heating
```

```
# Electrical: Collapsed similar categories together and handled missing values
electrical <- as.character(train_catPredictors$Electrical)

electrical <- fct_collapse(electrical, Fuse=c("FuseA", "FuseF", "FuseP"))
electrical <- fct_collapse(electrical, Other=c("Mix"))
electrical[is.na(electrical)] <- "Other"

train$Electrical <- electrical
```

```
# Fireplace: Handled missing values
fireplace <- as.character(train_catPredictors$FireplaceQu)
fireplace[is.na(fireplace)] <- "none"
train$FireplaceQu <- as.factor(fireplace)
```

Need to do the same for test dataset:

```
# Heating: Collapsed categories with low frequencies into "other"
heating <- as.factor(test_catPredictors$Heating)
heating <- fct_other(heating, keep=c("GasA", "GasW"))
test$Heating <- heating
```

```
# Electrical: Collapsed similar categories together and handled missing values
electrical <- as.character(test_catPredictors$Electrical)

electrical <- fct_collapse(electrical, Fuse=c("FuseA", "FuseF", "FuseP"))
electrical <- fct_collapse(electrical, Other=c("Mix"))
```

```
## Warning: Unknown levels in 'f': Mix
```

```
electrical[is.na(electrical)] <- "Other"
```

```
## Warning in '[<-factor'('*tmp*', is.na(electrical), value = "Other"): invalid  
## factor level, NA generated
```

```
test$Electrical <- electrical
```

```
# Fireplace: Handled missing values  
fireplace <- as.character(test_catPredictors$FireplaceQu)  
fireplace[is.na(fireplace)] <- "none"  
test$FireplaceQu <- as.factor(fireplace)
```

Thomas: RoofMatl, Exterior1st/Exterior2nd, SaleType

## RoofMatl - Dropped

1434/1460 entries in the training set are CompShg.

The off-materials aren't meaningfully different price-wise as an 'other' group. Wood Shingles ('wdshngl') does contain 2 houses in the 99th percentile sale price, but with only 6 entries I don't think it's safe to include.

I think we're better off dropping this one.

```
train <- select(train, -c(RoofMatl))  
test <- select(test, -c(RoofMatl))
```

## Exterior1st/2nd

Fixed the following label mis-matches between columns: Exterior1st - WdShng, CemntBd, BrkComm, Exterior2nd - Wd Shng, CmentBd, Brk Cmn

~90% of these two variables matched. In the ~10% that didn't match, Exterior1st is generally a better predictor of sale price than Exterior2nd. I converted Exterior2nd into a boolean, TRUE if Exterior1st != Exterior2nd.

I combined the bottom half of Exterior1st's categories into an 'Other' category. (This leaves 7, but Brick Face/Cement Board seem to be decent categories for predicting sale price, so I didn't want to drop them.)

```
train$Exterior2nd[train$Exterior2nd=='Wd Shng'] <- 'WdShing'  
train$Exterior2nd[train$Exterior2nd=='CmentBd'] <- 'CemntBd'  
train$Exterior2nd[train$Exterior2nd=='Brk Cmn'] <- 'BrkComm'  
train$Exterior2nd <- train$Exterior1st!=train$Exterior2nd  
train$Exterior1st <- fct_collapse(train$Exterior1st, Other = c("AsbShng", "AsphShn", "CBlock", "ImStucc", "Brk Face", "CmentBd", "BrkComm"))  
  
test$Exterior2nd[test$Exterior2nd=='Wd Shng'] <- 'WdShing'  
test$Exterior2nd[test$Exterior2nd=='CmentBd'] <- 'CemntBd'  
test$Exterior2nd[test$Exterior2nd=='Brk Cmn'] <- 'BrkComm'  
test$Exterior2nd <- test$Exterior1st!=test$Exterior2nd  
test$Exterior1st <- fct_collapse(test$Exterior1st, Other = c("AsbShng", "AsphShn", "CBlock", "ImStucc", "Brk Face", "CmentBd", "BrkComm"))
```

```
## Warning: Unknown levels in 'f': ImStucc, Stone
```

Bernhard: I also changed ExterCond:

```
table(train$ExterCond)
```

```
##
##   Ex   Fa   Gd   Po   TA
##    3   28  146    1 1282
```

```
table(test$ExterCond)
```

```
##
##   Ex   Fa   Gd   Po   TA
##    9   39  153    2 1256
```

Po and Ex are rather uncommon, so we collapse them all into “other”:

```
train$ExterCond = fct_collapse(train$ExterCond, other=c("Ex", "Po"))
test$ExterCond = fct_collapse(test$ExterCond, other=c("Ex", "Po"))

summary(train$ExterCond)
```

```
## other    Fa    Gd    TA
##     4     28   146  1282
```

```
summary(test$ExterCond)
```

```
## other    Fa    Gd    TA
##    11     39   153  1256
```

## SaleType

WD, New, and Court deed/estate were the three most common categories, and all 3 were significant when using SaleType as sole predictor. Combined the other categories into ‘Other’.

```
train$SaleType <- fct_collapse(train$SaleType, Other = c("ConLD", "ConLw", "ConLI", "CWD", "Oth", "Con"))
test$SaleType <- fct_collapse(test$SaleType, Other = c("ConLD", "ConLw", "ConLI", "CWD", "Oth", "Con"))
```

## Marina: Neighborhood, GarageType, GarageFinish, GarageQual, GarageCond

```
### Neighborhood ###
# Collapse categories with low frequencies into "other"

#Explore counts
train_catPredictors %>% count(Neighborhood, sort = TRUE)
```

```
## # A tibble: 25 x 2
##   Neighborhood      n
##   <fct>          <int>
## 1 NAmes          225
## 2 CollgCr        150
## 3 OldTown        113
## 4 Edwards        100
## 5 Somerst         86
## 6 Gilbert         79
## 7 NridgHt         77
## 8 Sawyer          74
## 9 NWAmes          73
## 10 SawyerW        59
## # ... with 15 more rows
```

```
#Factorize
neighborhood <- as.factor(train_catPredictors$Neighborhood)

#Convert to "Other" any category that represents less than 2% of the data
neighborhood <- fct_collapse(neighborhood, Other = c("MeadowV", "BrDale", "Veenker", "NPkVill", "Blueste
levels(neighborhood) #New levels of the factor
```

```
## [1] "Other" "BrkSide" "CollgCr" "Crawfor" "Edwards" "Gilbert" "Mitchel"
## [8] "NAmes" "NoRidge" "NridgHt" "NWAmes" "OldTown" "Sawyer" "SawyerW"
## [15] "Somerst" "Timber"
```

```
#Update column with new values
train$Neighborhood <- neighborhood
```

Need to do the same on test data:

```
#Factorize
neighborhood <- as.factor(test_catPredictors$Neighborhood)

#Convert to "Other" any category that represents less than 2% of the data
neighborhood <- fct_collapse(neighborhood, Other = c("MeadowV", "BrDale", "Veenker", "NPkVill", "Blueste
levels(neighborhood) #New levels of the factor
```

```
## [1] "Other" "BrkSide" "CollgCr" "Crawfor" "Edwards" "Gilbert" "Mitchel"
## [8] "NAmes" "NoRidge" "NridgHt" "NWAmes" "OldTown" "Sawyer" "SawyerW"
## [15] "Somerst" "Timber"
```

```
#Update column with new values
test$Neighborhood <- neighborhood
```

Anyone sees the issue??

```
table(train$Neighborhood)
```

```
##
##   Other BrkSide CollgCr Crawfor Edwards Gilbert Mitchel  NAmes NoRidge NridgHt
##   187     58     150     51     100     79     49     225     41     77
##  NWAmes OldTown  Sawyer SawyerW Somerst  Timber
##     73     113     74     59     86     38
```

```
table(test$Neighborhood)
```

```
##
##   Other BrkSide CollgCr Crawfor Edwards Gilbert Mitchel  NAmes NoRidge NridgHt
##   201     50     117     52     94     86     65     218     30     89
##  NWAmes OldTown  Sawyer SawyerW Somerst  Timber
##     58     126     77     66     96     34
```

```
### GarageType ###
```

```
#Explore counts
```

```
train_catPredictors %>% count(GarageType, sort = TRUE)
```

```
## # A tibble: 7 x 2
##   GarageType      n
##   <fct>        <int>
## 1 Attchd         870
## 2 Detchd         387
## 3 BuiltIn         88
## 4 <NA>           81
## 5 Basement        19
## 6 CarPort         9
## 7 2Types         6
```

```
#Handle NAs
```

```
#According to the data description, NA means no garage.
```

```
#Change NA category to "none" to avoid issues.
```

```
garageType <- as.character(train_catPredictors$GarageType)
```

```
garageType[is.na(garageType)] <- "none"
```

```
garageType <- as.factor(garageType)
```

```
#Collapse into "Other" categories that represent less than 5% of the data
```

```
garageType <- garageType %>%
  fct_lump(prop=0.05, other_level='Other')
```

```
#levels(garageType) #New levels of the factor
```

```
#Update column with new values
```

```
train$GarageType <- garageType
```

Attention!! Need to do the same on the test data:

```
garageType <- as.character(test$GarageType)
garageType[is.na(garageType)] <- "none"
garageType <- as.factor(garageType)
```

```

garageType <- garageType %>%
  fct_lump(prop=0.05, other_level='Other')
levels(garageType)

## [1] "Attchd" "BuiltIn" "Detchd" "none" "Other"

levels(train$GarageType)

## [1] "Attchd" "BuiltIn" "Detchd" "none" "Other"

test$GarageType <- garageType

### GarageFinish ###

#Explore counts
train_catPredictors %>% count(GarageFinish, sort = TRUE)

## # A tibble: 4 x 2
##   GarageFinish      n
##   <fct>         <int>
## 1 Unf           605
## 2 RFn           422
## 3 Fin           352
## 4 <NA>          81

#Handle NAs
#According to the data description, NA means no garage.
#Change NA category to "none" to avoid issues.
garageFinish <- as.character(train_catPredictors$GarageFinish)
garageFinish[is.na(garageFinish)] <- "none"
garageFinish <- as.factor(garageFinish)

#No need to collapse categories

#Update column with new values
train$GarageFinish <- garageFinish

```

Need to do the same for the test data:

```

#Handle NAs
#According to the data description, NA means no garage.
#Change NA category to "none" to avoid issues.
garageFinish <- as.character(test_catPredictors$GarageFinish)
garageFinish[is.na(garageFinish)] <- "none"
garageFinish <- as.factor(garageFinish)
#No need to collapse categories

#Update column with new values
test$GarageFinish <- garageFinish

```

```
### GarageQual ###
```

```
#Explore counts
```

```
train_catPredictors %>% count(GarageQual, sort = TRUE)
```

```
## # A tibble: 6 x 2
##   GarageQual      n
##   <fct>         <int>
## 1 TA           1311
## 2 <NA>           81
## 3 Fa           48
## 4 Gd           14
## 5 Ex            3
## 6 Po            3
```

```
#Handle NAs
```

```
#According to the data description, NA means no garage.
```

```
#Change NA category to "none" to avoid issues.
```

```
garageQual <- as.character(train_catPredictors$GarageQual)
```

```
garageQual[is.na(garageQual)] <- "none"
```

```
garageQual <- as.factor(garageQual)
```

```
#Collapse categories:
```

```
# - Let's collapse Ex (Excellent) and Gd (Good) into 1 category: Gd
```

```
# - Let's collapse Fa (Fair) and Po (Poor) into 1 category: Po
```

```
# - None and TA remains the same
```

```
garageQual <- fct_collapse(garageQual, Gd = c("Ex", "Gd"))
```

```
garageQual <- fct_collapse(garageQual, Po = c("Fa", "Po"))
```

```
#Update column with new values
```

```
train$GarageQual <- garageQual
```

Need to do the same for test data:

```
#Handle NAs
```

```
#According to the data description, NA means no garage.
```

```
#Change NA category to "none" to avoid issues.
```

```
garageQual <- as.character(test_catPredictors$GarageQual)
```

```
garageQual[is.na(garageQual)] <- "none"
```

```
garageQual <- as.factor(garageQual)
```

```
#Collapse categories:
```

```
# - Let's collapse Ex (Excellent) and Gd (Good) into 1 category: Gd
```

```
# - Let's collapse Fa (Fair) and Po (Poor) into 1 category: Po
```

```
# - None and TA remains the same
```

```
garageQual <- fct_collapse(garageQual, Gd = c("Ex", "Gd"))
```

```
## Warning: Unknown levels in 'f': Ex
```

```
garageQual <- fct_collapse(garageQual, Po = c("Fa", "Po"))
```

```
#Update column with new values  
test$GarageQual <- garageQual
```

```
### GarageCond ###
```

```
#Explore counts  
train_catPredictors %>% count(GarageCond, sort = TRUE)
```

```
## # A tibble: 6 x 2  
##   GarageCond      n  
##   <fct>         <int>  
## 1 TA           1326  
## 2 <NA>           81  
## 3 Fa            35  
## 4 Gd             9  
## 5 Po             7  
## 6 Ex             2
```

```
#Handle NAs  
#According to the data description, NA means no garage.  
#Change NA category to "none" to avoid issues.  
garageCond <- as.character(train_catPredictors$GarageCond)  
garageCond[is.na(garageCond)] <- "none"  
garageCond <- as.factor(garageCond)
```

```
#Collapse categories:  
# - Let's collapse Ex (Excellent) and Gd (Good) into 1 category: Gd  
# - Let's collapse Fa (Fair) and Po (Poor) into 1 category: Po  
# - None and TA remains the same
```

```
garageCond <- fct_collapse(garageCond, Gd = c("Ex", "Gd"))  
garageCond <- fct_collapse(garageCond, Po = c("Fa", "Po"))
```

```
#Update column with new values  
train$GarageCond <- garageCond
```

Need to do the same with test data:

```
#Handle NAs  
#According to the data description, NA means no garage.  
#Change NA category to "none" to avoid issues.  
garageCond <- as.character(test_catPredictors$GarageCond)  
garageCond[is.na(garageCond)] <- "none"  
garageCond <- as.factor(garageCond)
```

```
#Collapse categories:  
# - Let's collapse Ex (Excellent) and Gd (Good) into 1 category: Gd  
# - Let's collapse Fa (Fair) and Po (Poor) into 1 category: Po  
# - None and TA remains the same
```



```
garageCond <- fct_collapse(garageCond, Gd = c("Ex", "Gd"))
garageCond <- fct_collapse(garageCond, Po = c("Fa", "Po"))
```

```
#Update column with new values
test$GarageCond <- garageCond
```

Note: We also need to discuss the NA's in the numerical variable GarageYrBlt, see later.

### Paul: LotShape, LotConfig, LandContour

Fortunately there are no NA values in either the test or train sets.

```
sum(is.na(train$LotShape))
```

```
## [1] 0
```

```
sum(is.na(test$LotShape))
```

```
## [1] 0
```

```
sum(is.na(train$LotConfig))
```

```
## [1] 0
```

```
sum(is.na(test$LotConfig))
```

```
## [1] 0
```

```
sum(is.na(train$LandContour))
```

```
## [1] 0
```

```
sum(is.na(test$LandContour))
```

```
## [1] 0
```

```
fct_count(train$LotShape)
```

```
## # A tibble: 4 x 2
##   f         n
##   <fct> <int>
## 1 IR1     484
## 2 IR2      41
## 3 IR3      10
## 4 Reg    925
```

```
fct_count(test$LotShape)
```

```
## # A tibble: 4 x 2
##   f         n
##   <fct> <int>
## 1 IR1     484
## 2 IR2      35
## 3 IR3       6
## 4 Reg     934
```

```
fct_count(train$LotConfig)
```

```
## # A tibble: 5 x 2
##   f         n
##   <fct> <int>
## 1 Corner    263
## 2 CulDSac    94
## 3 FR2        47
## 4 FR3         4
## 5 Inside   1052
```

```
fct_count(test$LotConfig)
```

```
## # A tibble: 5 x 2
##   f         n
##   <fct> <int>
## 1 Corner    248
## 2 CulDSac    82
## 3 FR2        38
## 4 FR3        10
## 5 Inside   1081
```

```
fct_count(train$LandContour)
```

```
## # A tibble: 4 x 2
##   f         n
##   <fct> <int>
## 1 Bnk      63
## 2 HLS      50
## 3 Low      36
## 4 Lvl     1311
```

```
fct_count(test$LandContour)
```

```
## # A tibble: 4 x 2
##   f         n
##   <fct> <int>
## 1 Bnk      54
## 2 HLS      70
## 3 Low      24
## 4 Lvl     1311
```

All of these variables are highly imbalanced. In each there is one category that represents a “regular” shape, configuration, or land contour, which amount for ~2/3 or more of the total instances. Thus, I collapsed all of the less represented “irregular” categories into one.

```
train$LotShape <- fct_collapse(train$LotShape, Irregular = c("IR1", "IR2", "IR3"))
train$LotConfig <- fct_collapse(train$LotConfig, Other = c("Corner", "CulDSac", "FR2", "FR3"))
train$LandContour <- fct_collapse(train$LandContour, NonLvl = c("Bnk", "HLS", "Low"))
```

```
fct_count(train$LotShape)
```

```
## # A tibble: 2 x 2
##   f         n
##   <fct>   <int>
## 1 Irregular 535
## 2 Reg      925
```

```
fct_count(train$LotConfig)
```

```
## # A tibble: 2 x 2
##   f         n
##   <fct>   <int>
## 1 Other   408
## 2 Inside 1052
```

```
fct_count(train$LandContour)
```

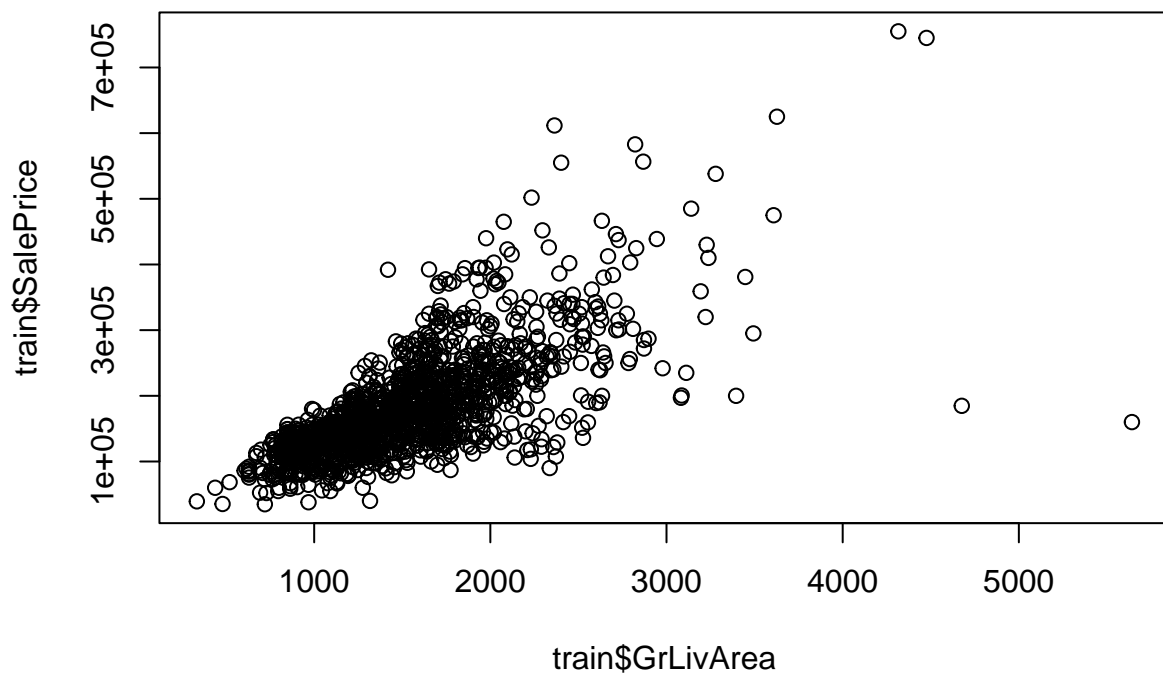
```
## # A tibble: 2 x 2
##   f         n
##   <fct>   <int>
## 1 NonLvl  149
## 2 Lvl    1311
```

Need to do the same for the test data:

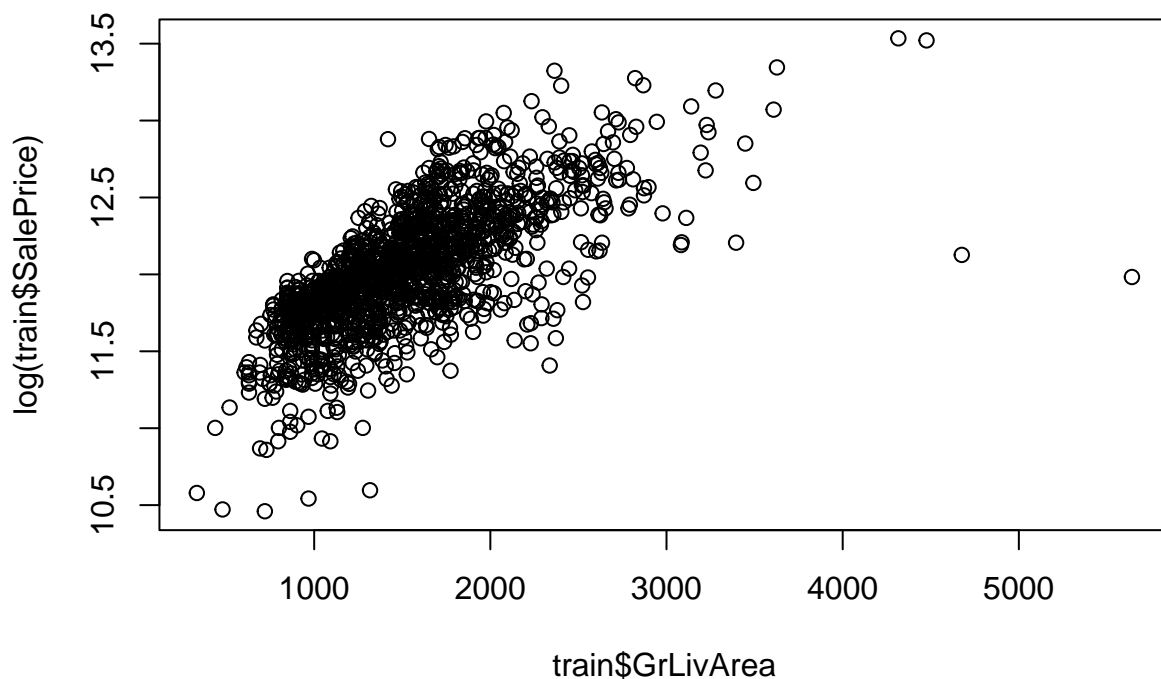
```
test$LotShape <- fct_collapse(test$LotShape, Irregular = c("IR1", "IR2", "IR3"))
test$LotConfig <- fct_collapse(test$LotConfig, Other = c("Corner", "CulDSac", "FR2", "FR3"))
test$LandContour <- fct_collapse(test$LandContour, NonLvl = c("Bnk", "HLS", "Low"))
```

**First Try for building a predictive model, using just one variable, but as a smooth function:**

```
library(splines)
plot(train$SalePrice ~ train$GrLivArea)
```

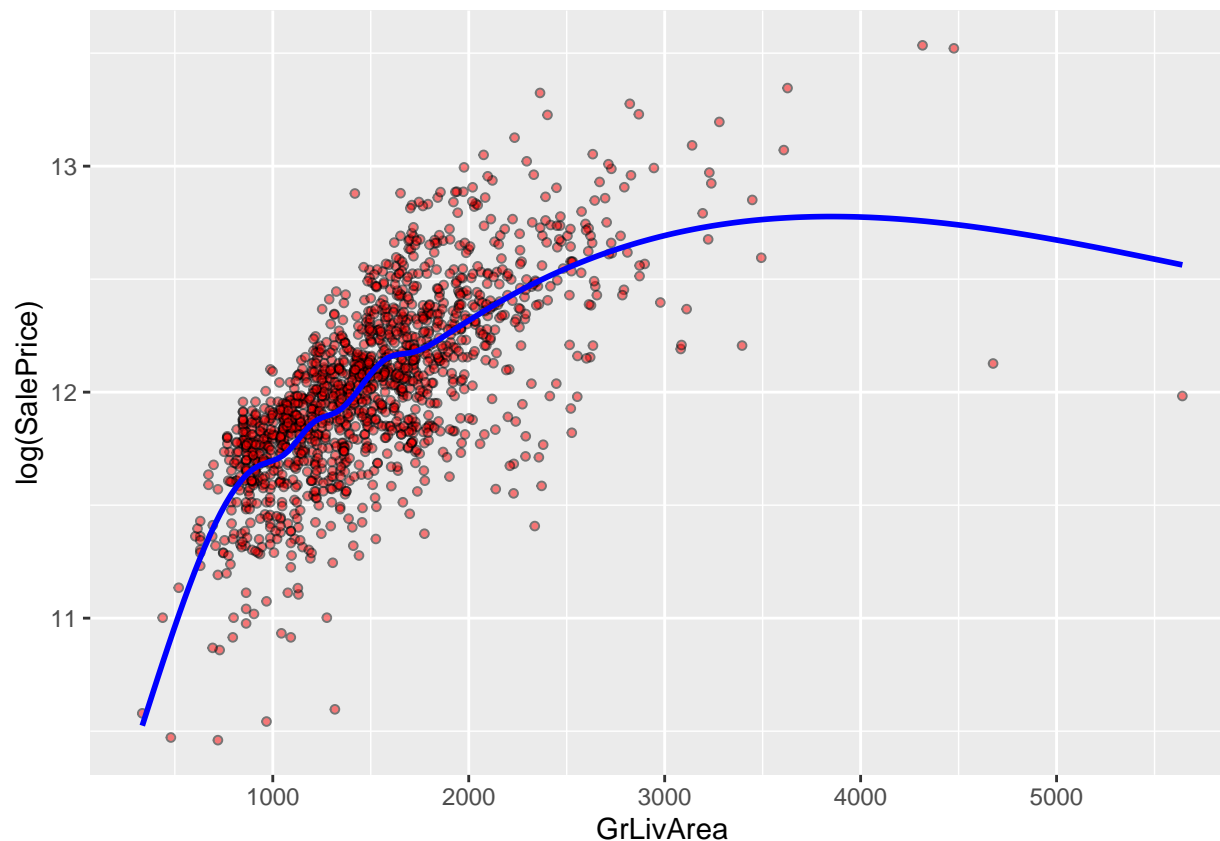


```
plot(log(train$SalePrice) ~ train$GrLivArea)
```



Better to log-transform response

```
fit1 = lm(log(SalePrice) ~ ns(GrLivArea, df=10), data=train)
seqGrLivArea = seq(min(train$GrLivArea), max(train$GrLivArea), length.out=200)
predictedSpline = predict(fit1, newdata = data.frame(GrLivArea=seqGrLivArea))
ggplot(data=train,
       aes(x=GrLivArea,
           y=log(SalePrice))
) +
  geom_point(pch=21, fill="red", size=1.2, alpha=0.5) +
  geom_line(
    data = data.frame(
      x = seqGrLivArea,
      y = predictedSpline
    ),
    aes(
      x=x,
      y=y
    ),
    color = "blue", size = 1
  )
```



Preparing the dataset with the predicted sale prices for the test data:

```
predicted.SalePrice = exp(predict(fit1, newdata=data.frame(GrLivArea = test$GrLivArea)))
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice)
# write.csv(file='C:\\Teaching\\NewCollege\\StatsTopics\\Submission1.csv', SubmitDF, row.names = FALSE)
```

Submitting this file to the Kaggle competition, I obtained a “Prediction error”, measures as

$$\sum (\log(\hat{y}_i) - \log(y_i))^2$$

of 0.28857, where  $\hat{y}_i$  is my prediction of the sale price of the  $i$ th house in the test data, and  $y_i$  is the actual sale price only known to Kaggle.

## Second Try, including all predictors!

If we include all predictors, one issue is that a few predictors might have a lot of NA values, and then the corresponding observation is not used in the fit. (You find this out when you try to fit the full model.) Let’s see which variables have the most NA’s.

```
train %>%
  summarize(across(everything(), ~sum(is.na(.x)))) %>%
  sort(decreasing=TRUE)
```

```
## Warning in xtfm.data.frame(x): cannot xtfm data frames
```

```
## # A tibble: 1 x 80
##   MiscFe~1 Alley LotFr~2 Garag~3 BsmtE~4 BsmtF~5 BsmtQ~6 BsmtC~7 BsmtF~8 MasVn~9
##   <int> <int>   <int>   <int>   <int>   <int>   <int>   <int>   <int>   <int>
## 1    1406  1369     259     81     38     38     37     37     37     8
## # ... with 70 more variables: MasVnrArea <int>, Id <int>, MSSubClass <int>,
## #   MSZoning <int>, LotArea <int>, Street <int>, LotShape <int>,
## #   LandContour <int>, Utilities <int>, LotConfig <int>, LandSlope <int>,
## #   Neighborhood <int>, Condition1 <int>, Condition2 <int>, BldgType <int>,
## #   HouseStyle <int>, OverallQual <int>, OverallCond <int>, YearBuilt <int>,
## #   YearRemodAdd <int>, RoofStyle <int>, Exterior1st <int>, Exterior2nd <int>,
## #   ExterQual <int>, ExterCond <int>, Foundation <int>, BsmtFinSF1 <int>, ...
```

```
dim(train)
```

```
## [1] 1460   80
```

For the variable `MiscFeature`, almost all values are missing. However, looking in the data description file, this actually means that the house simply doesn't have any other features. So, we set the NA's to "none", in both the train and test datasets. The same applies to `Alley`, where an NA means "none":

```
train$MiscFeature = fct_explicit_na(train$MiscFeature, na_level="none")
test$MiscFeature = fct_explicit_na(test$MiscFeature, na_level="none")

train$Alley = fct_explicit_na(train$Alley, na_level="none")
test$Alley = fct_explicit_na(test$Alley, na_level="none")
```

For `LotFrontage`, the missing values are genuine. (But lets hope that the value being missing has no connection to the sales price of a house.)

Another issue with fitting a full model is the number of unique values a predictor has. If it only has **one unique value (or one unique factor level)**, then it doesn't vary, i.e., it is a constant. This causes issues because then the design matrix  $X$  is not full rank. The column for the intercept is a column of all 1's, and then each column for a predictor which is constant is also a column of a fixed number. This causes a linear dependency between these columns, and the design matrix is not full rank.

First, lets turn the character variables into factors, both in the training and testing data. This will pay off later:

```
train = train %>% mutate(across(where(is.character), as.factor))
test = test %>% mutate(across(where(is.character), as.factor))
```

Let's find the predictors which have constant values throughout:

```
train %>%
  summarize(across(everything(), ~length(unique(.x)))) %>%
  sort()
```

```
## Warning in xtfrm.data.frame(x): cannot xtfrm data frames
```

```
## # A tibble: 1 x 80
##   Street LotShape LandCo~1 Utili~2 LotCo~3 Exter~4 Centr~5 Alley LandS~6 RoofS~7
##   <int>   <int>   <int>   <int>   <int>   <int>   <int> <int>   <int>   <int>
```

```
## 1      2      2      2      2      2      2      2      3      3      3
## # ... with 70 more variables: Heating <int>, Electrical <int>,
## #   BsmtHalfBath <int>, HalfBath <int>, PavedDrive <int>, MSSubClass <int>,
## #   MSZoning <int>, Condition1 <int>, Condition2 <int>, BldgType <int>,
## #   HouseStyle <int>, ExterQual <int>, ExterCond <int>, BsmtFullBath <int>,
## #   FullBath <int>, KitchenAbvGr <int>, KitchenQual <int>, Fireplaces <int>,
## #   GarageFinish <int>, GarageQual <int>, GarageCond <int>, PoolQC <int>,
## #   SaleType <int>, MasVnrType <int>, BsmtQual <int>, BsmtCond <int>, ...
```

There doesn't seem to be a variable that has only one unique value or one unique factor level. So we should be good to go.

Having done/checked all that, we are ready to fit the full model with all variables. However, using `> fit2 = lm(log(SalePrice) ~ . , data=train %>% select(-Id, -SalePrice))`, I ran into a problem, where R shows the error message `contrasts can be applied only to factors with 2 or more levels`.

With trial and error, I saw that we can fit a model with the first 8 predictors, but when we include 'Utilities', there is an issue

```
fit2 = lm(log(train$SalePrice) ~ . , data=train[,2:9])
summary(fit2)
```

```
##
## Call:
## lm(formula = log(train$SalePrice) ~ . , data = train[, 2:9])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36843 -0.20396 -0.03778  0.18938  1.10208
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)      1.178e+01  1.766e-01  66.695
## MSSubClass1-story single-family other -1.999e-01  3.060e-02  -6.531
## MSSubClassmulti-level single-family non PUD 1.505e-01  2.361e-02   6.373
## MSSubClassother      1.592e-02  3.126e-02   0.509
## MSZoningR0      -3.892e-01  5.575e-02  -6.982
## MSZoningRL      -2.201e-01  5.187e-02  -4.243
## MSZoningother     -9.451e-01  1.177e-01  -8.028
## LotFrontage       2.851e-03  4.789e-04   5.954
## LotArea          7.202e-06  1.346e-06   5.352
## StreetPave       1.959e-01  1.528e-01   1.282
## AlleyPave       1.152e-01  7.898e-02   1.458
## Alleynone       9.683e-02  5.118e-02   1.892
## LotShapeReg     -1.762e-01  2.200e-02  -8.009
## LandContourLvl    4.205e-02  3.304e-02   1.273
##
##              Pr(>|t|)
## (Intercept)    < 2e-16 ***
## MSSubClass1-story single-family other  9.67e-11 ***
## MSSubClassmulti-level single-family non PUD 2.65e-10 ***
## MSSubClassother    0.6106
## MSZoningR0       4.82e-12 ***
## MSZoningRL       2.37e-05 ***
## MSZoningother    2.36e-15 ***
```



```
## LotFrontage                3.44e-09 ***
## LotArea                    1.05e-07 ***
## StreetPave                 0.2002
## AlleyPave                  0.1451
## Alleynone                  0.0587 .
## LotShapeReg                2.74e-15 ***
## LandContourLvl            0.2033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3245 on 1187 degrees of freedom
## (259 observations deleted due to missingness)
## Multiple R-squared:  0.3984, Adjusted R-squared:  0.3918
## F-statistic: 60.47 on 13 and 1187 DF, p-value: < 2.2e-16
```

What is going on with utilities:

```
summary(train$Utilities)
```

```
## AllPub NoSeWa
##    1459      1
```

We see that it is almost constant! There is only one observation with a different utility type. Probably, that observation has some missing values on some other variables, and hence is removed from the design matrix, making it an all constant predictor. Let's check:

```
train[train$Utilities == 'NoSeWa',]
```

```
## # A tibble: 1 x 80
##   Id MSSubClass MSZon~1 LotFr~2 LotArea Street Alley LotSh~3 LandC~4 Utili~5
##   <dbl> <fct>      <fct>      <dbl>  <dbl> <fct> <fct> <fct>  <fct>  <fct>
## 1   945 1-story si~ RL        NA   14375 Pave   none Irregu~ Lvl    NoSeWa
## # ... with 70 more variables: LotConfig <fct>, LandSlope <fct>,
## # Neighborhood <fct>, Condition1 <fct>, Condition2 <fct>, BldgType <fct>,
## # HouseStyle <fct>, OverallQual <dbl>, OverallCond <dbl>, YearBuilt <dbl>,
## # YearRemodAdd <dbl>, RoofStyle <fct>, Exterior1st <fct>, Exterior2nd <lgl>,
## # MasVnrType <fct>, MasVnrArea <dbl>, ExterQual <fct>, ExterCond <fct>,
## # Foundation <fct>, BsmtQual <fct>, BsmtCond <fct>, BsmtExposure <fct>,
## # BsmtFinType1 <fct>, BsmtFinSF1 <dbl>, BsmtFinType2 <fct>, ...
```

There we go, `LotFrontage` is NA for this particular house, so it is removed, and the remaining houses all have the same utility type.

## Removing/Replacing Missing Values

It is best to get the training dataset that has no missing values (since they will be discarded in the fitting process of the full model anyway), and then check if any other predictors are constant. Which variables still have a lot of missing values:

```
train %>%
  summarize(across(everything(), ~sum(is.na(.x)))) %>%
  sort(decreasing = TRUE)
```

```
## Warning in xtfrm.data.frame(x): cannot xtfrm data frames
```

```
## # A tibble: 1 x 80
##   LotFro~1 Garag~2 BsmtE~3 BsmtF~4 BsmtQ~5 BsmtC~6 BsmtF~7 MasVn~8 MasVn~9 Id
##   <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>
## 1     259      81      38      38      37      37      37      8      8      0
## # ... with 70 more variables: MSSubClass <int>, MSZoning <int>, LotArea <int>,
## #   Street <int>, Alley <int>, LotShape <int>, LandContour <int>,
## #   Utilities <int>, LotConfig <int>, LandSlope <int>, Neighborhood <int>,
## #   Condition1 <int>, Condition2 <int>, BldgType <int>, HouseStyle <int>,
## #   OverallQual <int>, OverallCond <int>, YearBuilt <int>, YearRemodAdd <int>,
## #   RoofStyle <int>, Exterior1st <int>, Exterior2nd <int>, ExterQual <int>,
## #   ExterCond <int>, Foundation <int>, BsmtFinSF1 <int>, BsmtFinSF2 <int>, ...
```

For now, I'm going to drop `LotFrontage` from consideration, although we could impute values. I'm also going to drop `GarageYrBlt` from consideration, because it has around 80 missing values for those garages where there is no information. Since we have info on the garage from other variables, I rather keep 81 observations in the dataset, but not include `GarageYrBlt`. So, I'm going to drop `GarageYrBlt` from the list of predictors:

```
train = train %>% select(-LotFrontage, -GarageYrBlt)
test = test %>% select(-LotFrontage, -GarageYrBlt)
```

We now need to handle the Basement values. We need to replace the NA's with "none":

```
train$BsmtQual = fct_explicit_na(train$BsmtQual, na_level="none")
train$BsmtCond = fct_explicit_na(train$BsmtCond, na_level="none")
train$BsmtExposure = fct_explicit_na(train$BsmtExposure, na_level="none")
train$BsmtFinType1 = fct_explicit_na(train$BsmtFinType1, na_level="none")
train$BsmtFinType2 = fct_explicit_na(train$BsmtFinType2, na_level="none")

test$BsmtQual = fct_explicit_na(test$BsmtQual, na_level="none")
test$BsmtCond = fct_explicit_na(test$BsmtCond, na_level="none")
test$BsmtExposure = fct_explicit_na(test$BsmtExposure, na_level="none")
test$BsmtFinType1 = fct_explicit_na(test$BsmtFinType1, na_level="none")
test$BsmtFinType2 = fct_explicit_na(test$BsmtFinType2, na_level="none")

train %>%
  summarize(across(everything(), ~sum(is.na(.x)))) %>%
  sort(decreasing = TRUE)
```

```
## Warning in xtfrm.data.frame(x): cannot xtfrm data frames
```

```
## # A tibble: 1 x 78
##   MasVnrType MasVnr~1 Id MSSub~2 MSZon~3 LotArea Street Alley LotSh~4 LandC~5
##   <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>
## 1      8      8      0      0      0      0      0      0      0      0
```

```
## # ... with 68 more variables: Utilities <int>, LotConfig <int>,
## #   LandSlope <int>, Neighborhood <int>, Condition1 <int>, Condition2 <int>,
## #   BldgType <int>, HouseStyle <int>, OverallQual <int>, OverallCond <int>,
## #   YearBuilt <int>, YearRemodAdd <int>, RoofStyle <int>, Exterior1st <int>,
## #   Exterior2nd <int>, ExterQual <int>, ExterCond <int>, Foundation <int>,
## #   BsmtQual <int>, BsmtCond <int>, BsmtExposure <int>, BsmtFinType1 <int>,
## #   BsmtFinSF1 <int>, BsmtFinType2 <int>, BsmtFinSF2 <int>, ...
```

For MasVnrType, I will introduce a new category “missing”, but for MasVnrArea I will just input 0 for those 8 missing areas:

```
summary(train$MasVnrType)
```

```
## BrkCmn BrkFace      None      Stone      NA's
##      15      445      864      128       8
```

```
train$MasVnrType = fct_explicit_na(train$MasVnrType, na_level="missing")
train$MasVnrArea[is.na(train$MasVnrArea)] = 0

test$MasVnrType = fct_explicit_na(test$MasVnrType, na_level="missing")
test$MasVnrArea[is.na(test$MasVnrArea)] = 0
```

We now have no missing predictor values in the training data:

```
dim(train)
```

```
## [1] 1460   78
```

```
dim(train %>% drop_na())
```

```
## [1] 1460   78
```

Let’s now revisit check if any predictors are constant:

```
train %>%
  summarize(across(everything(), ~length(unique(.x)))) %>%
  sort()
```

```
## Warning in xtfrm.data.frame(x): cannot xtfrm data frames
```

```
## # A tibble: 1 x 78
##   Street LotShape LandCo~1 Utili~2 LotCo~3 Exter~4 Centr~5 Alley LandS~6 RoofS~7
##   <int>    <int>    <int>    <int>    <int>    <int>    <int> <int>    <int>    <int>
## 1      2      2      2      2      2      2      2      3      3      3
## # ... with 68 more variables: Heating <int>, Electrical <int>,
## #   BsmtHalfBath <int>, HalfBath <int>, PavedDrive <int>, MSSubClass <int>,
## #   MSZoning <int>, Condition1 <int>, Condition2 <int>, BldgType <int>,
## #   HouseStyle <int>, ExterQual <int>, ExterCond <int>, BsmtFullBath <int>,
## #   FullBath <int>, KitchenAbvGr <int>, KitchenQual <int>, Fireplaces <int>,
## #   GarageFinish <int>, GarageQual <int>, GarageCond <int>, PoolQC <int>,
## #   SaleType <int>, MasVnrType <int>, BsmtQual <int>, BsmtCond <int>, ...
```

Seems fine, although for Utilities:

```
summary(train$Utilities)
```

```
## AllPub NoSeWa
##    1459      1
```

This means we also need to drop Utilities from the test data.

```
train = train %>% select(-Utilities)
test = test %>% select(-Utilities)
```

## NA's in Test Data

Just like in the training dataset, we might have some NA's in the test data:

```
isNAtest = apply(test,1,function(x) any(is.na(x)))
sum(isNAtest)
```

```
## [1] 11
```

We still have 11 observations with at least one missing predictor. This is a problem since when we use all predictors, we will not be able to obtain a predicted sales price for these 11 houses. Which predictors have the most missing values:

```
test %>%
  summarize(across(everything(), ~sum(is.na(.x)))) %>%
  sort(decreasing = TRUE)
```

```
## Warning in xtfrm.data.frame(x): cannot xtfrm data frames
```

```
## # A tibble: 1 x 76
##   MSZoning BsmtFullBath BsmtHa~1 Funct~2 Exter~3 Exter~4 BsmtF~5 BsmtF~6 BsmtU~7
##   <int>      <int>      <int>  <int>  <int>  <int>  <int>  <int>  <int>
## 1      4          2          2      2      1      1      1      1      1
## # ... with 67 more variables: TotalBsmtSF <int>, KitchenQual <int>,
## #   GarageCars <int>, GarageArea <int>, SaleType <int>, Id <int>,
## #   MSSubClass <int>, LotArea <int>, Street <int>, Alley <int>, LotShape <int>,
## #   LandContour <int>, LotConfig <int>, LandSlope <int>, Neighborhood <int>,
## #   Condition1 <int>, Condition2 <int>, BldgType <int>, HouseStyle <int>,
## #   OverallQual <int>, OverallCond <int>, YearBuilt <int>, YearRemodAdd <int>,
## #   RoofStyle <int>, MasVnrType <int>, MasVnrArea <int>, ExterQual <int>, ...
```

MSZoning:

```
summary(train$MSZoning)
```

```
##   FV   R0   RL other
##   65  234 1151   10
```

```
summary(test$MSZoning)
```

```
##      FV      R0      RL other  NA's  
##      74     252   1114     15     4
```

```
test$MSZoning = fct_explicit_na(test$MSZoning, na_level="other")  
summary(test$MSZoning)
```

```
##      FV      R0      RL other  
##      74     252   1114     19
```

### BsmtFullBath:

```
summary(train$BsmtFullBath)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.  
## 0.0000 0.0000 0.0000 0.4253 1.0000 3.0000
```

```
summary(test$BsmtFullBath)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.      NA's  
## 0.0000 0.0000 0.0000 0.4345 1.0000 3.0000         2
```

```
test$BsmtFullBath[is.na(test$BsmtFullBath)] = 0
```

### BsmtHalfBath:

```
summary(train$BsmtHalfBath)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.  
## 0.00000 0.00000 0.00000 0.05753 0.00000 2.00000
```

```
summary(test$BsmtHalfBath)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.      NA's  
## 0.0000 0.0000 0.0000 0.0652 0.0000 2.0000         2
```

```
test$BsmtHalfBath[is.na(test$BsmtHalfBath)] = 0
```

### Functional:

```
summary(train$Functional)
```

```
## Major Minor  Mod Severe  Typ  
##    19    65    15      1 1360
```

```
summary(test$Functional)
```

```
## Major Minor Mod Severe Typ NA's  
##      9    70    20      1 1357    2
```

```
train$Functional = train$Functional == "Typ"  
test$Functional = test$Functional == "Typ"  
test$Functional[is.na(test$Functional)] = TRUE  
summary(train$Functional)
```

```
##      Mode FALSE TRUE  
## logical    100 1360
```

```
summary(test$Functional)
```

```
##      Mode FALSE TRUE  
## logical    100 1359
```

### Exterior1st:

```
summary(train$Exterior1st)
```

```
## Other BrkFace CemntBd HdBoard MetalSd Plywood VinylSd Wd Sdng  
##      78      50      61      222      220      108      515      206
```

```
summary(test$Exterior1st)
```

```
## Other BrkFace CemntBd HdBoard MetalSd Plywood VinylSd Wd Sdng NA's  
##      78      37      65      220      230      113      510      205      1
```

```
test$Exterior1st <- fct_explicit_na(test$Exterior1st, na_level="Other")  
summary(test$Exterior1st)
```

```
## Other BrkFace CemntBd HdBoard MetalSd Plywood VinylSd Wd Sdng  
##      79      37      65      220      230      113      510      205
```

### Exterior2nd:

```
summary(train$Exterior2nd)
```

```
##      Mode FALSE TRUE  
## logical  1323  137
```

```
summary(test$Exterior2nd)
```

```
##      Mode  FALSE    TRUE   NA's  
## logical   1327    131     1
```

```
test$Exterior2nd[is.na(test$Exterior2nd)] = FALSE
```

## BsmtFinSF1

```
summary(train$BsmtFinSF1)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.0     0.0   383.5   443.6   712.2   5644.0
```

```
summary(test$BsmtFinSF1)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##      0.0     0.0   350.5   439.2   753.5   4010.0     1
```

```
test$BsmtFinSF1[is.na(test$BsmtFinSF1)] = 0
```

## BsmtFinSF2

```
summary(train$BsmtFinSF2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.00     0.00     0.00   46.55     0.00  1474.00
```

```
summary(test$BsmtFinSF2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##      0.00     0.00     0.00   52.62     0.00  1526.00     1
```

```
test$BsmtFinSF2[is.na(test$BsmtFinSF2)] = 0
```

## BsmtUnfSF

```
summary(train$BsmtUnfSF)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.0    223.0   477.5   567.2   808.0   2336.0
```

```
summary(test$BsmtUnfSF)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##      0.0   219.2   460.0   554.3   797.8  2140.0     1
```

```
test$BsmtUnfSF[is.na(test$BsmtUnfSF)] = 460
```

## TotalBsmtSF

```
summary(train$TotalBsmtSF)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.0   795.8   991.5  1057.4  1298.2  6110.0
```

```
summary(test$TotalBsmtSF)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##       0     784     988    1046    1305    5095     1
```

```
test$TotalBsmtSF[is.na(test$TotalBsmtSF)] = 988
```

## KitchenQual

```
summary(train$KitchenQual)
```

```
## Ex Fa Gd TA  
## 100 39 586 735
```

```
summary(test$KitchenQual)
```

```
## Ex Fa Gd TA NA's  
## 105 31 565 757 1
```

```
test$KitchenQual[is.na(test$KitchenQual)] = "TA"
```

## GarageCars

```
summary(train$GarageCars)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.000   1.000   2.000   1.767   2.000   4.000
```



```
summary(test$GarageCars)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##    0.000   1.000   2.000   1.766   2.000   5.000     1
```

```
test$GarageCars[is.na(test$GarageCars)] = 1.766
```

## GarageArea

```
summary(train$GarageArea)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##      0.0   334.5   480.0   473.0   576.0  1418.0
```

```
summary(test$GarageArea)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##      0.0   318.0   480.0   472.8   576.0  1488.0     1
```

```
test$GarageArea[is.na(test$GarageArea)] = 480
```

## SaleType

```
summary(train$SaleType)
```

```
##   COD Other   New    WD  
##   43    28   122  1267
```

```
summary(test$SaleType)
```

```
##   COD Other   New    WD  NA's  
##   44    39   117  1258     1
```

```
test$SaleType[is.na(test$SaleType)] = "Other"
```

## Fitting the model with almost all variables

We can now fit the full model:

```
SalePrice = train$SalePrice  
HouseId = train$Id #just in case we need it  
train = train %>% select(-Id, -SalePrice)  
fit2 = lm(log(SalePrice) ~ . , data=train)
```

We can now try to predict the sales price based on the variables in the test data, since we have addressed all missing values in the test data:

```
predicted.SalePrice2 = exp(predict(fit2, newdata=test))
```

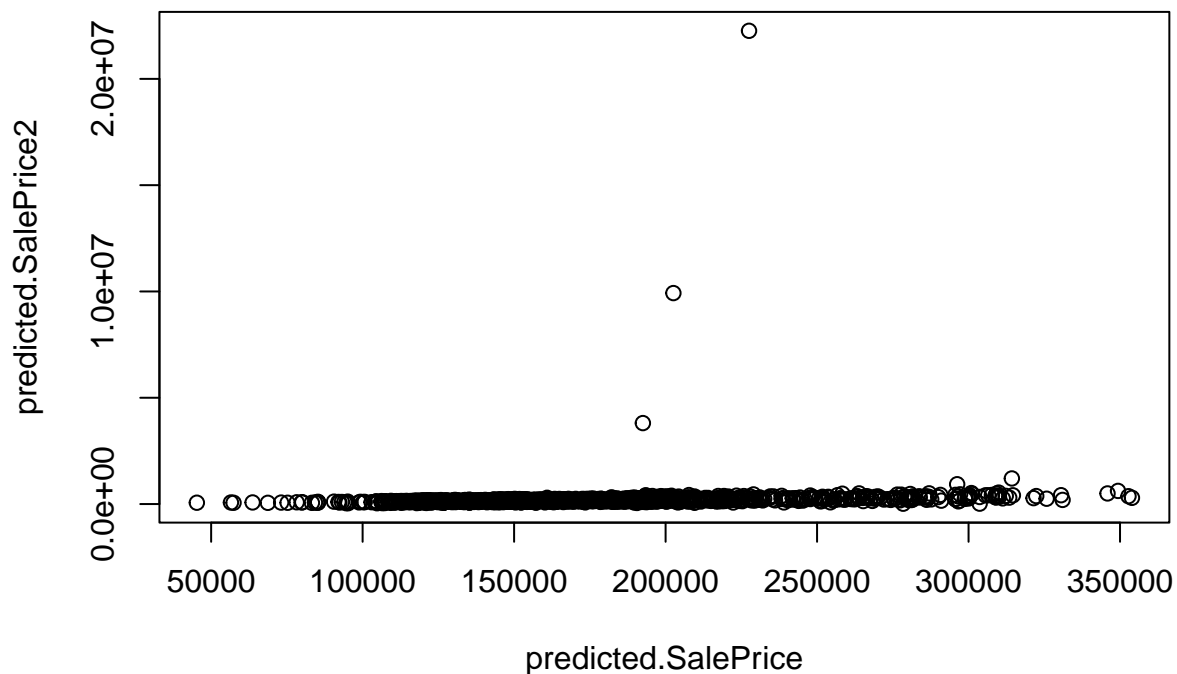
```
## Warning in predict.lm(fit2, newdata = test): prediction from a rank-deficient
## fit may be misleading
```

Preparing the dataset with the predicted sale prices for the test data:

```
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice2)
# write.csv(file='C:\\Teaching\\NewCollege\\StatsTopics\\Submission2.csv', SubmitDF, row.names = FALSE)
```

Interestingly, using all these variables, the prediction score did not go down by much. It is now 0.26450. What is the relationship between our predictions based on the two models:

```
plot(predicted.SalePrice2 ~ predicted.SalePrice)
```



This is pretty telling. Just for a few houses (three), we predicted a much higher price with the second model compared to the first. Which houses are these:

```
SubmitDF %>% slice_max(SalePrice,n=8)
```

```
##      Id  SalePrice
## 1140 2600 22261168.2
## 1044 2504  9924031.0
##  961 2421  3808022.8
```

```
## 1090 2550 1204635.6
## 1251 2711 932354.8
## 1223 2683 618442.0
## 1168 2628 538556.1
## 20 1480 516352.8
```

For the house with ID 2600 in the test data, we predicted a sales price of over 22 million! The error alone in this prediction could be huge! To find out, I'm replacing just the prediction for the 5 most expensive predicted prices with the maximum sales price found in the training data.

```
SubmitDF$SalePrice[SubmitDF$Id %in% c(2600, 2504, 2421, 2550, 2711)] = max(SalePrice)
# write.csv(file='C:\\Teaching\\NewCollege\\StatsTopics\\Submission3.csv', SubmitDF, row.names = FALSE)
```

Yes, the prediction error went down to 0.19609!

## Where our work begins

## using AIC to determine best model

```
library(stats)
```

```
step(fit2)
```

```
## Start: AIC=-6169.59
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
## LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
## Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
## OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
## Exterior2nd + MasVnrType + MasVnrArea + ExterQual + ExterCond +
## Foundation + BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 +
## BsmtFinSF1 + BsmtFinType2 + BsmtFinSF2 + BsmtUnfSF + TotalBsmtSF +
## Heating + HeatingQC + CentralAir + Electrical + '1stFlrSF' +
## '2ndFlrSF' + LowQualFinSF + GrLivArea + BsmtFullBath + BsmtHalfBath +
## FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
## TotRmsAbvGrd + Functional + Fireplaces + FireplaceQu + GarageType +
## GarageFinish + GarageCars + GarageArea + GarageQual + GarageCond +
## PavedDrive + WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
## ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
## MoSold + YrSold + SaleType + SaleCondition
##
##
## Step: AIC=-6169.59
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
## LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
## Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
## OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
## Exterior2nd + MasVnrType + MasVnrArea + ExterQual + ExterCond +
## Foundation + BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 +
## BsmtFinSF1 + BsmtFinType2 + BsmtFinSF2 + BsmtUnfSF + TotalBsmtSF +
## Heating + HeatingQC + CentralAir + Electrical + '1stFlrSF' +
```

```

##      '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##      FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##      TotRmsAbvGrd + Functional + Fireplaces + FireplaceQu + GarageType +
##      GarageFinish + GarageCars + GarageArea + GarageQual + GarageCond +
##      PavedDrive + WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
##      ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
##      MoSold + YrSold + SaleType + SaleCondition
##
##
## Step:   AIC=-6169.59
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##      LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##      Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##      OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
##      Exterior2nd + MasVnrType + MasVnrArea + ExterQual + ExterCond +
##      Foundation + BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 +
##      BsmtFinSF1 + BsmtFinType2 + BsmtFinSF2 + BsmtUnfSF + Heating +
##      HeatingQC + CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' +
##      LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##      BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##      Functional + Fireplaces + FireplaceQu + GarageType + GarageFinish +
##      GarageCars + GarageArea + GarageQual + GarageCond + PavedDrive +
##      WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
##      ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
##      MoSold + YrSold + SaleType + SaleCondition
##
##
##           Df Sum of Sq   RSS   AIC
## - BsmtFinType2    6   0.07592 17.003 -6175.1
## - ExterQual        3   0.01714 16.945 -6174.1
## - FireplaceQu      5   0.06920 16.997 -6173.6
## - RoofStyle        2   0.00468 16.932 -6173.2
## - MasVnrType       4   0.05216 16.980 -6173.1
## - BsmtCond         3   0.03242 16.960 -6172.8
## - Fence            4   0.06691 16.994 -6171.8
## - PavedDrive       2   0.02086 16.948 -6171.8
## - GarageType       3   0.04548 16.973 -6171.7
## - Fireplaces       1   0.00018 16.928 -6171.6
## - MiscVal          1   0.00028 16.928 -6171.6
## - MasVnrArea       1   0.00043 16.928 -6171.5
## - OpenPorchSF      1   0.00047 16.928 -6171.5
## - Exterior2nd      1   0.00053 16.928 -6171.5
## - LotShape         1   0.00099 16.928 -6171.5
## - MoSold           1   0.00143 16.929 -6171.5
## - Electrical       2   0.02594 16.953 -6171.4
## - BsmtQual         3   0.04929 16.977 -6171.3
## - Alley            2   0.02657 16.954 -6171.3
## - BedroomAbvGr     1   0.00360 16.931 -6171.3
## - GarageFinish     2   0.02912 16.957 -6171.1
## - BsmtHalfBath     1   0.00758 16.935 -6170.9
## - YrSold           1   0.01434 16.942 -6170.4
## - LandContour      1   0.01735 16.945 -6170.1
## - Street           1   0.02276 16.950 -6169.6
## <none>                                16.927 -6169.6
## - GarageQual       2   0.05324 16.981 -6169.0

```

```

## - LandSlope      2    0.05499 16.983 -6168.9
## - '3SsnPorch'    1    0.03537 16.963 -6168.5
## - GarageCond     2    0.06136 16.989 -6168.3
## - KitchenAbvGr   1    0.04921 16.977 -6167.3
## - GarageArea     1    0.05675 16.984 -6166.7
## - LotConfig      1    0.05780 16.985 -6166.6
## - EnclosedPorch  1    0.06167 16.989 -6166.3
## - BsmtFinSF2     1    0.06234 16.990 -6166.2
## - MSSubClass     3    0.11028 17.038 -6166.1
## - BsmtFinType1   5    0.15741 17.085 -6166.1
## - HeatingQC      4    0.14271 17.070 -6165.3
## - TotRmsAbvGrd   1    0.07331 17.001 -6165.3
## - Heating        2    0.10324 17.031 -6164.7
## - GarageCars     1    0.09236 17.020 -6163.6
## - Foundation     5    0.18771 17.115 -6163.5
## - HalfBath       1    0.10218 17.030 -6162.8
## - BsmtUnfSF      1    0.10512 17.033 -6162.5
## - FullBath       1    0.11336 17.041 -6161.8
## - LowQualFinSF   1    0.11363 17.041 -6161.8
## - CentralAir     1    0.11787 17.045 -6161.5
## - YearRemodAdd    1    0.12225 17.050 -6161.1
## - SaleType       3    0.17149 17.099 -6160.9
## - KitchenQual    3    0.17552 17.103 -6160.5
## - HouseStyle     3    0.18528 17.113 -6159.7
## - BsmtExposure   4    0.21772 17.145 -6158.9
## - WoodDeckSF     1    0.14815 17.076 -6158.9
## - BsmtFinSF1     1    0.15522 17.083 -6158.3
## - BldgType       3    0.20334 17.131 -6158.2
## - ExterCond      3    0.21661 17.144 -6157.0
## - YearBuilt      1    0.17095 17.099 -6156.9
## - Exterior1st    7    0.34555 17.273 -6154.1
## - Functional     1    0.21279 17.140 -6153.3
## - MiscFeature    4    0.29623 17.224 -6152.3
## - Condition1     3    0.27700 17.204 -6151.9
## - BsmtFullBath   1    0.24641 17.174 -6150.5
## - LotArea        1    0.27049 17.198 -6148.4
## - ScreenPorch    1    0.33983 17.267 -6142.6
## - SaleCondition  5    0.47470 17.402 -6139.2
## - Condition2     3    0.57202 17.500 -6127.1
## - '1stFlrSF'     1    0.91551 17.843 -6094.7
## - OverallCond    1    0.99855 17.926 -6087.9
## - Neighborhood   15   1.52367 18.451 -6073.8
## - '2ndFlrSF'     1    1.30606 18.234 -6063.1
## - MSZoning       3    1.36053 18.288 -6062.7
## - OverallQual    1    1.60281 18.530 -6039.5
## - PoolArea       1    1.79485 18.722 -6024.5
## - PoolQC         3    2.34172 19.269 -5986.4
##
## Step:  AIC=-6175.05
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
##   Exterior2nd + MasVnrType + MasVnrArea + ExterQual + ExterCond +

```

```
## Foundation + BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 +
## BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC +
## CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +
## BsmtFullBath + BsmtHalfBath + FullBath + HalfBath + BedroomAbvGr +
## KitchenAbvGr + KitchenQual + TotRmsAbvGrd + Functional +
## Fireplaces + FireplaceQu + GarageType + GarageFinish + GarageCars +
## GarageArea + GarageQual + GarageCond + PavedDrive + WoodDeckSF +
## OpenPorchSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
## PoolArea + PoolQC + Fence + MiscFeature + MiscVal + MoSold +
## YrSold + SaleType + SaleCondition
```

```
##
##          Df Sum of Sq    RSS      AIC
## - ExterQual      3    0.01651 17.020 -6179.6
## - FireplaceQu    5    0.06505 17.069 -6179.5
## - RoofStyle      2    0.00278 17.006 -6178.8
## - BsmtCond       3    0.03221 17.036 -6178.3
## - MasVnrType     4    0.05597 17.059 -6178.3
## - PavedDrive     2    0.01899 17.022 -6177.4
## - Electrical     2    0.02312 17.027 -6177.1
## - Fireplaces     1    0.00005 17.003 -6177.0
## - MasVnrArea     1    0.00031 17.004 -6177.0
## - LotShape       1    0.00065 17.004 -6177.0
## - MiscVal        1    0.00068 17.004 -6177.0
## - OpenPorchSF    1    0.00080 17.004 -6177.0
## - Exterior2nd    1    0.00108 17.005 -6177.0
## - MoSold         1    0.00120 17.005 -6176.9
## - Alley          2    0.02607 17.029 -6176.8
## - Fence          4    0.07292 17.076 -6176.8
## - GarageType     3    0.05011 17.053 -6176.8
## - GarageFinish   2    0.02747 17.031 -6176.7
## - BsmtQual       3    0.05186 17.055 -6176.6
## - BedroomAbvGr   1    0.00583 17.009 -6176.6
## - BsmtHalfBath   1    0.00704 17.011 -6176.4
## - YrSold         1    0.01562 17.019 -6175.7
## - LandContour    1    0.01738 17.021 -6175.6
## - Street         1    0.02275 17.026 -6175.1
## <none>                17.003 -6175.1
## - GarageQual     2    0.04947 17.053 -6174.8
## - LandSlope      2    0.05528 17.059 -6174.3
## - '3SsnPorch'    1    0.03238 17.036 -6174.3
## - GarageCond     2    0.05990 17.063 -6173.9
## - BsmtFinType1   5    0.13953 17.143 -6173.1
## - KitchenAbvGr   1    0.04910 17.052 -6172.8
## - GarageArea     1    0.05274 17.056 -6172.5
## - EnclosedPorch  1    0.06073 17.064 -6171.8
## - LotConfig      1    0.06302 17.067 -6171.7
## - MSSubClass     3    0.11309 17.116 -6171.4
## - BsmtFinSF2     1    0.06748 17.071 -6171.3
## - HeatingQC      4    0.14302 17.146 -6170.8
## - TotRmsAbvGrd   1    0.07865 17.082 -6170.3
## - Heating        2    0.10691 17.110 -6169.9
## - Foundation     5    0.18950 17.193 -6168.9
## - GarageCars     1    0.09758 17.101 -6168.7
## - BsmtUnfSF      1    0.10566 17.109 -6168.0
```

```

## - HalfBath      1  0.10579 17.109 -6168.0
## - LowQualFinSF  1  0.11331 17.117 -6167.4
## - CentralAir    1  0.11385 17.117 -6167.3
## - FullBath      1  0.11612 17.120 -6167.1
## - YearRemodAdd  1  0.12308 17.127 -6166.5
## - KitchenQual   3  0.17741 17.181 -6165.9
## - HouseStyle    3  0.18190 17.185 -6165.5
## - WoodDeckSF    1  0.14314 17.147 -6164.8
## - SaleType      3  0.19201 17.195 -6164.7
## - BldgType      3  0.19719 17.201 -6164.2
## - BsmtFinSF1    1  0.15826 17.162 -6163.5
## - ExterCond     3  0.21829 17.222 -6162.4
## - BsmtExposure  4  0.24613 17.250 -6162.1
## - YearBuilt     1  0.18147 17.185 -6161.6
## - Exterior1st   7  0.36012 17.364 -6158.5
## - MiscFeature   4  0.29539 17.299 -6157.9
## - Functional    1  0.22589 17.229 -6157.8
## - Condition1    3  0.27530 17.279 -6157.6
## - BsmtFullBath  1  0.25334 17.257 -6155.5
## - LotArea       1  0.25702 17.261 -6155.1
## - ScreenPorch   1  0.33740 17.341 -6148.4
## - SaleCondition  5  0.47714 17.481 -6144.6
## - Condition2    3  0.56941 17.573 -6133.0
## - '1stFlrSF'    1  0.96760 17.971 -6096.2
## - OverallCond   1  1.01946 18.023 -6092.0
## - Neighborhood  15 1.56492 18.568 -6076.5
## - MSZoning       3  1.35142 18.355 -6069.4
## - '2ndFlrSF'    1  1.30148 18.305 -6069.4
## - OverallQual    1  1.60728 18.611 -6045.2
## - PoolArea      1  1.78152 18.785 -6031.6
## - PoolQC        3  2.32886 19.332 -5993.6
##
## Step:  AIC=-6179.64
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
##   Exterior2nd + MasVnrType + MasVnrArea + ExterCond + Foundation +
##   BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 + BsmtFinSF1 +
##   BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC + CentralAir +
##   Electrical + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath +
##   BsmtHalfBath + FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr +
##   KitchenQual + TotRmsAbvGrd + Functional + Fireplaces + FireplaceQu +
##   GarageType + GarageFinish + GarageCars + GarageArea + GarageQual +
##   GarageCond + PavedDrive + WoodDeckSF + OpenPorchSF + EnclosedPorch +
##   '3SsnPorch' + ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature +
##   MiscVal + MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq  RSS    AIC
## - FireplaceQu    5  0.06748 17.087 -6183.9
## - RoofStyle      2  0.00384 17.024 -6183.3
## - BsmtCond       3  0.03127 17.051 -6183.0
## - MasVnrType     4  0.05521 17.075 -6182.9
## - PavedDrive     2  0.01704 17.037 -6182.2

```

## - Electrical	2	0.02266	17.043	-6181.7
## - Fireplaces	1	0.00002	17.020	-6181.6
## - MasVnrArea	1	0.00025	17.020	-6181.6
## - LotShape	1	0.00033	17.020	-6181.6
## - OpenPorchSF	1	0.00078	17.021	-6181.6
## - Exterior2nd	1	0.00088	17.021	-6181.6
## - MiscVal	1	0.00100	17.021	-6181.6
## - MoSold	1	0.00131	17.021	-6181.5
## - Alley	2	0.02553	17.046	-6181.4
## - GarageType	3	0.04913	17.069	-6181.4
## - Fence	4	0.07384	17.094	-6181.3
## - BsmtHalfBath	1	0.00674	17.027	-6181.1
## - BedroomAbvGr	1	0.00726	17.027	-6181.0
## - GarageFinish	2	0.03113	17.051	-6181.0
## - BsmtQual	3	0.05598	17.076	-6180.8
## - YrSold	1	0.01482	17.035	-6180.4
## - LandContour	1	0.01623	17.036	-6180.2
## - Street	1	0.01906	17.039	-6180.0
## <none>			17.020	-6179.6
## - GarageQual	2	0.04938	17.069	-6179.4
## - LandSlope	2	0.05585	17.076	-6178.9
## - '3SsnPorch'	1	0.03360	17.053	-6178.8
## - GarageCond	2	0.06204	17.082	-6178.3
## - BsmtFinType1	5	0.14112	17.161	-6177.6
## - KitchenAbvGr	1	0.04866	17.069	-6177.5
## - GarageArea	1	0.05339	17.073	-6177.1
## - LotConfig	1	0.06297	17.083	-6176.2
## - EnclosedPorch	1	0.06449	17.084	-6176.1
## - MSSubClass	3	0.11164	17.132	-6176.1
## - BsmtFinSF2	1	0.07452	17.095	-6175.3
## - HeatingQC	4	0.14955	17.169	-6174.9
## - Heating	2	0.10618	17.126	-6174.6
## - TotRmsAbvGrd	1	0.08282	17.103	-6174.5
## - GarageCars	1	0.09838	17.118	-6173.2
## - Foundation	5	0.19630	17.216	-6172.9
## - HalfBath	1	0.10667	17.127	-6172.5
## - CentralAir	1	0.10780	17.128	-6172.4
## - FullBath	1	0.11447	17.134	-6171.8
## - BsmtUnfSF	1	0.11697	17.137	-6171.6
## - LowQualFinSF	1	0.12196	17.142	-6171.2
## - YearRemodAdd	1	0.12342	17.143	-6171.1
## - HouseStyle	3	0.18491	17.205	-6169.9
## - WoodDeckSF	1	0.14247	17.162	-6169.5
## - BldgType	3	0.19575	17.216	-6168.9
## - KitchenQual	3	0.20086	17.221	-6168.5
## - SaleType	3	0.20103	17.221	-6168.5
## - ExterCond	3	0.21281	17.233	-6167.5
## - BsmtFinSF1	1	0.16734	17.187	-6167.4
## - BsmtExposure	4	0.24487	17.265	-6166.8
## - YearBuilt	1	0.19346	17.213	-6165.1
## - Exterior1st	7	0.35989	17.380	-6163.1
## - Functional	1	0.22055	17.241	-6162.8
## - MiscFeature	4	0.29416	17.314	-6162.6
## - Condition1	3	0.27193	17.292	-6162.5



```

## - LotArea      1  0.25083 17.271 -6160.3
## - BsmtFullBath 1  0.25193 17.272 -6160.2
## - ScreenPorch  1  0.33969 17.360 -6152.8
## - SaleCondition 5  0.48188 17.502 -6148.9
## - Condition2    3  0.56815 17.588 -6137.7
## - '1stFlrSF'    1  0.95854 17.979 -6101.6
## - OverallCond   1  1.02791 18.048 -6096.0
## - Neighborhood 15  1.59959 18.619 -6078.5
## - MSZoning       3  1.35342 18.373 -6073.9
## - '2ndFlrSF'    1  1.31763 18.338 -6072.8
## - OverallQual    1  1.71760 18.738 -6041.3
## - PoolArea       1  1.80760 18.828 -6034.3
## - PoolQC         3  2.34827 19.368 -5996.9
##
## Step:  AIC=-6183.86
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + RoofStyle + Exterior1st +
##   Exterior2nd + MasVnrType + MasVnrArea + ExterCond + Foundation +
##   BsmtQual + BsmtCond + BsmtExposure + BsmtFinType1 + BsmtFinSF1 +
##   BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC + CentralAir +
##   Electrical + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath +
##   BsmtHalfBath + FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr +
##   KitchenQual + TotRmsAbvGrd + Functional + Fireplaces + GarageType +
##   GarageFinish + GarageCars + GarageArea + GarageQual + GarageCond +
##   PavedDrive + WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
##   ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
##   MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq  RSS    AIC
## - RoofStyle      2   0.00218 17.090 -6187.7
## - MasVnrType      4   0.05104 17.139 -6187.5
## - BsmtCond        3   0.02964 17.117 -6187.3
## - PavedDrive      2   0.01692 17.104 -6186.4
## - Electrical      2   0.02063 17.108 -6186.1
## - Fence           4   0.07037 17.158 -6185.9
## - LotShape        1   0.00017 17.088 -6185.8
## - MasVnrArea      1   0.00047 17.088 -6185.8
## - OpenPorchSF     1   0.00065 17.088 -6185.8
## - MiscVal         1   0.00133 17.089 -6185.7
## - Alley           2   0.02505 17.113 -6185.7
## - MoSold          1   0.00166 17.089 -6185.7
## - Exterior2nd     1   0.00176 17.089 -6185.7
## - GarageType      3   0.05459 17.142 -6185.2
## - BedroomAbvGr    1   0.00778 17.095 -6185.2
## - BsmtHalfBath    1   0.00846 17.096 -6185.1
## - GarageFinish     2   0.03706 17.125 -6184.7
## - LandContour     1   0.01520 17.103 -6184.6
## - YrSold           1   0.01600 17.103 -6184.5
## - BsmtQual        3   0.06533 17.153 -6184.3
## - Street          1   0.02034 17.108 -6184.1
## <none>                                17.087 -6183.9
## - GarageQual      2   0.05556 17.143 -6183.1

```

```

## - '3SsnPorch'      1  0.03691 17.124 -6182.7
## - GarageCond       2  0.06299 17.150 -6182.5
## - LandSlope        2  0.06346 17.151 -6182.4
## - BsmtFinType1     5  0.14384 17.231 -6181.6
## - KitchenAbvGr     1  0.05110 17.139 -6181.5
## - LotConfig        1  0.05885 17.146 -6180.8
## - GarageArea       1  0.06168 17.149 -6180.6
## - MSSubClass       3  0.10915 17.197 -6180.6
## - EnclosedPorch    1  0.06737 17.155 -6180.1
## - BsmtFinSF2       1  0.07190 17.159 -6179.7
## - HeatingQC        4  0.14829 17.236 -6179.2
## - Heating          2  0.10270 17.190 -6179.1
## - TotRmsAbvGrd     1  0.09001 17.177 -6178.2
## - GarageCars       1  0.09745 17.185 -6177.6
## - Foundation       5  0.19439 17.282 -6177.3
## - HalfBath         1  0.10561 17.193 -6176.9
## - BsmtUnfSF        1  0.11052 17.198 -6176.4
## - FullBath         1  0.11106 17.198 -6176.4
## - CentralAir       1  0.11183 17.199 -6176.3
## - LowQualFinSF     1  0.11887 17.206 -6175.7
## - YearRemodAdd     1  0.12215 17.210 -6175.5
## - Fireplaces       1  0.12344 17.211 -6175.3
## - HouseStyle       3  0.18607 17.273 -6174.0
## - BldgType         3  0.19693 17.284 -6173.1
## - WoodDeckSF       1  0.15179 17.239 -6172.9
## - BsmtFinSF1       1  0.15321 17.241 -6172.8
## - KitchenQual      3  0.20379 17.291 -6172.5
## - BsmtExposure     4  0.23203 17.320 -6172.2
## - SaleType         3  0.21477 17.302 -6171.6
## - ExterCond        3  0.21553 17.303 -6171.6
## - YearBuilt        1  0.18693 17.274 -6170.0
## - Exterior1st      7  0.34662 17.434 -6168.5
## - Functional       1  0.22131 17.309 -6167.1
## - MiscFeature      4  0.29267 17.380 -6167.1
## - Condition1       3  0.27026 17.358 -6166.9
## - LotArea          1  0.24174 17.329 -6165.3
## - BsmtFullBath     1  0.24438 17.332 -6165.1
## - ScreenPorch      1  0.33283 17.420 -6157.7
## - SaleCondition    5  0.49152 17.579 -6152.5
## - Condition2       3  0.57240 17.660 -6141.8
## - '1stFlrSF'       1  0.96046 18.048 -6106.0
## - OverallCond      1  1.00172 18.089 -6102.7
## - Neighborhood    15  1.61004 18.698 -6082.4
## - MSZoning         3  1.35405 18.442 -6078.5
## - '2ndFlrSF'       1  1.33777 18.425 -6075.8
## - OverallQual      1  1.76188 18.849 -6042.6
## - PoolArea         1  1.82919 18.917 -6037.4
## - PoolQC          3  2.39662 19.484 -5998.2
##
## Step:  AIC=-6187.67
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd +

```

```
## MasVnrType + MasVnrArea + ExterCond + Foundation + BsmtQual +
## BsmtCond + BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 +
## BsmtUnfSF + Heating + HeatingQC + CentralAir + Electrical +
## '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
## FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
## TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
## GarageCars + GarageArea + GarageQual + GarageCond + PavedDrive +
## WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
## ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
## MoSold + YrSold + SaleType + SaleCondition
##
```

	Df	Sum of Sq	RSS	AIC
## - MasVnrType	4	0.05199	17.142	-6191.2
## - BsmtCond	3	0.02938	17.119	-6191.2
## - PavedDrive	2	0.01648	17.106	-6190.3
## - Electrical	2	0.02081	17.110	-6189.9
## - LotShape	1	0.00024	17.090	-6189.7
## - Fence	4	0.07069	17.160	-6189.6
## - OpenPorchSF	1	0.00056	17.090	-6189.6
## - MasVnrArea	1	0.00069	17.090	-6189.6
## - Alley	2	0.02464	17.114	-6189.6
## - Exterior2nd	1	0.00168	17.091	-6189.5
## - MoSold	1	0.00170	17.091	-6189.5
## - MiscVal	1	0.00194	17.091	-6189.5
## - GarageType	3	0.05392	17.143	-6189.1
## - BedroomAbvGr	1	0.00775	17.097	-6189.0
## - BsmtHalfBath	1	0.00842	17.098	-6189.0
## - GarageFinish	2	0.03760	17.127	-6188.5
## - LandContour	1	0.01530	17.105	-6188.4
## - YrSold	1	0.01614	17.106	-6188.3
## - Street	1	0.01977	17.109	-6188.0
## - BsmtQual	3	0.06777	17.157	-6187.9
## <none>			17.090	-6187.7
## - GarageQual	2	0.05564	17.145	-6186.9
## - '3SsnPorch'	1	0.03697	17.127	-6186.5
## - GarageCond	2	0.06355	17.153	-6186.3
## - LandSlope	2	0.06764	17.157	-6185.9
## - BsmtFinType1	5	0.14346	17.233	-6185.5
## - KitchenAbvGr	1	0.05111	17.141	-6185.3
## - LotConfig	1	0.05876	17.148	-6184.7
## - GarageArea	1	0.06094	17.151	-6184.5
## - MSSubClass	3	0.11108	17.201	-6184.2
## - EnclosedPorch	1	0.06775	17.157	-6183.9
## - BsmtFinSF2	1	0.07143	17.161	-6183.6
## - HeatingQC	4	0.14784	17.237	-6183.1
## - Heating	2	0.10221	17.192	-6183.0
## - TotRmsAbvGrd	1	0.09247	17.182	-6181.8
## - GarageCars	1	0.09774	17.187	-6181.3
## - Foundation	5	0.19490	17.285	-6181.1
## - HalfBath	1	0.10582	17.195	-6180.7
## - CentralAir	1	0.11049	17.200	-6180.3
## - BsmtUnfSF	1	0.11084	17.200	-6180.2
## - FullBath	1	0.11093	17.201	-6180.2
## - LowQualFinSF	1	0.11845	17.208	-6179.6

```

## - YearRemodAdd      1    0.12211 17.212 -6179.3
## - Fireplaces        1    0.12283 17.212 -6179.2
## - HouseStyle        3    0.18652 17.276 -6177.8
## - BldgType          3    0.19595 17.286 -6177.0
## - WoodDeckSF        1    0.15340 17.243 -6176.6
## - BsmtFinSF1        1    0.15460 17.244 -6176.5
## - KitchenQual       3    0.20524 17.295 -6176.2
## - BsmtExposure      4    0.23173 17.321 -6176.0
## - SaleType          3    0.21737 17.307 -6175.2
## - ExterCond         3    0.21766 17.307 -6175.2
## - YearBuilt         1    0.18888 17.279 -6173.6
## - Exterior1st       7    0.34889 17.439 -6172.2
## - MiscFeature       4    0.29099 17.381 -6171.0
## - Condition1        3    0.27129 17.361 -6170.7
## - Functional        1    0.22515 17.315 -6170.6
## - BsmtFullBath      1    0.24560 17.335 -6168.8
## - LotArea           1    0.25600 17.346 -6168.0
## - ScreenPorch       1    0.33415 17.424 -6161.4
## - SaleCondition     5    0.49510 17.585 -6156.0
## - Condition2        3    0.57250 17.662 -6145.6
## - '1stFlrSF'        1    0.96546 18.055 -6109.4
## - OverallCond       1    1.00519 18.095 -6106.2
## - Neighborhood     15    1.60831 18.698 -6086.4
## - MSZoning          3    1.35291 18.442 -6082.4
## - '2ndFlrSF'        1    1.33593 18.425 -6079.8
## - OverallQual       1    1.77977 18.869 -6045.0
## - PoolArea          1    1.89818 18.988 -6035.9
## - PoolQC           3    2.43267 19.522 -5999.4
##
## Step:  AIC=-6191.24
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd +
##   MasVnrArea + ExterCond + Foundation + BsmtQual + BsmtCond +
##   BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF +
##   Heating + HeatingQC + CentralAir + Electrical + '1stFlrSF' +
##   '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##   FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + PavedDrive +
##   WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
##   ScreenPorch + PoolArea + PoolQC + Fence + MiscFeature + MiscVal +
##   MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS   AIC
## - BsmtCond      3    0.02729 17.169 -6194.9
## - PavedDrive    2    0.01742 17.159 -6193.8
## - Fence         4    0.06884 17.210 -6193.4
## - LotShape      1    0.00007 17.142 -6193.2
## - Electrical    2    0.02360 17.165 -6193.2
## - OpenPorchSF   1    0.00048 17.142 -6193.2
## - MiscVal       1    0.00115 17.143 -6193.1
## - Exterior2nd   1    0.00130 17.143 -6193.1

```

## - Alley	2	0.02487	17.166	-6193.1
## - MoSold	1	0.00196	17.144	-6193.1
## - MasVnrArea	1	0.00318	17.145	-6193.0
## - GarageType	3	0.05258	17.194	-6192.8
## - BsmtHalfBath	1	0.00693	17.148	-6192.6
## - BedroomAbvGr	1	0.00720	17.149	-6192.6
## - YrSold	1	0.01185	17.153	-6192.2
## - LandContour	1	0.01512	17.157	-6192.0
## - GarageFinish	2	0.03948	17.181	-6191.9
## - Street	1	0.01939	17.161	-6191.6
## <none>			17.142	-6191.2
## - BsmtQual	3	0.07124	17.213	-6191.2
## - GarageQual	2	0.05273	17.194	-6190.8
## - GarageCond	2	0.05947	17.201	-6190.2
## - '3SsnPorch'	1	0.03856	17.180	-6190.0
## - BsmtFinType1	5	0.13763	17.279	-6189.6
## - LandSlope	2	0.06851	17.210	-6189.4
## - KitchenAbvGr	1	0.05489	17.197	-6188.6
## - LotConfig	1	0.05809	17.200	-6188.3
## - MSSubClass	3	0.10550	17.247	-6188.3
## - GarageArea	1	0.06003	17.202	-6188.1
## - BsmtFinSF2	1	0.06863	17.210	-6187.4
## - EnclosedPorch	1	0.07014	17.212	-6187.3
## - Heating	2	0.10459	17.246	-6186.4
## - HeatingQC	4	0.15666	17.298	-6186.0
## - HalfBath	1	0.09620	17.238	-6185.1
## - Foundation	5	0.19417	17.336	-6184.8
## - TotRmsAbvGrd	1	0.10264	17.244	-6184.5
## - GarageCars	1	0.10470	17.246	-6184.3
## - FullBath	1	0.10792	17.250	-6184.1
## - CentralAir	1	0.10833	17.250	-6184.0
## - BsmtUnfSF	1	0.10941	17.251	-6183.9
## - LowQualFinSF	1	0.11859	17.260	-6183.2
## - YearRemodAdd	1	0.12007	17.262	-6183.0
## - Fireplaces	1	0.12709	17.269	-6182.5
## - HouseStyle	3	0.18125	17.323	-6181.9
## - BldgType	3	0.18864	17.330	-6181.3
## - KitchenQual	3	0.20276	17.344	-6180.1
## - WoodDeckSF	1	0.15681	17.298	-6179.9
## - BsmtFinSF1	1	0.16094	17.302	-6179.6
## - BsmtExposure	4	0.23863	17.380	-6179.1
## - SaleType	3	0.21729	17.359	-6178.8
## - ExterCond	3	0.22342	17.365	-6178.3
## - YearBuilt	1	0.19264	17.334	-6176.9
## - Exterior1st	7	0.35676	17.498	-6175.2
## - MiscFeature	4	0.28600	17.428	-6175.1
## - Condition1	3	0.27232	17.414	-6174.2
## - Functional	1	0.22850	17.370	-6173.9
## - LotArea	1	0.23984	17.381	-6173.0
## - BsmtFullBath	1	0.24436	17.386	-6172.6
## - ScreenPorch	1	0.34342	17.485	-6164.3
## - SaleCondition	5	0.51370	17.655	-6158.1
## - Condition2	3	0.57349	17.715	-6149.2
## - '1stFlrSF'	1	0.95321	18.095	-6114.2

```

## - OverallCond      1      1.01311 18.155 -6109.4
## - Neighborhood    15      1.61158 18.753 -6090.0
## - MSZoning         3      1.33086 18.473 -6088.1
## - '2ndFlrSF'       1      1.32197 18.464 -6084.8
## - OverallQual      1      1.83889 18.980 -6044.5
## - PoolArea         1      1.89012 19.032 -6040.5
## - PoolQC           3      2.42048 19.562 -6004.4
##
## Step:  AIC=-6194.92
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd +
##   MasVnrArea + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##   BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageType + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + PavedDrive + WoodDeckSF +
##   OpenPorchSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
##   PoolArea + PoolQC + Fence + MiscFeature + MiscVal + MoSold +
##   YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS     AIC
## - Fence      4    0.06299 17.232 -6197.6
## - PavedDrive  2    0.01900 17.188 -6197.3
## - Electrical  2    0.01911 17.188 -6197.3
## - Alley       2    0.02330 17.192 -6196.9
## - LotShape    1    0.00003 17.169 -6196.9
## - OpenPorchSF 1    0.00049 17.169 -6196.9
## - Exterior2nd 1    0.00112 17.170 -6196.8
## - MiscVal     1    0.00152 17.170 -6196.8
## - MoSold      1    0.00225 17.171 -6196.7
## - MasVnrArea  1    0.00340 17.172 -6196.6
## - BsmtHalfBath 1    0.00724 17.176 -6196.3
## - BedroomAbvGr 1    0.00749 17.176 -6196.3
## - GarageType  3    0.05529 17.224 -6196.2
## - YrSold      1    0.01206 17.181 -6195.9
## - LandContour 1    0.01544 17.184 -6195.6
## - GarageFinish 2    0.03941 17.208 -6195.6
## - Street      1    0.01810 17.187 -6195.4
## <none>                17.169 -6194.9
## - BsmtQual    3    0.07303 17.242 -6194.7
## - GarageQual  2    0.05559 17.224 -6194.2
## - '3SsnPorch' 1    0.03683 17.206 -6193.8
## - LandSlope   2    0.07115 17.240 -6192.9
## - GarageCond  2    0.07158 17.241 -6192.8
## - MSSubClass  3    0.09912 17.268 -6192.5
## - KitchenAbvGr 1    0.05343 17.222 -6192.4
## - BsmtFinType1 5    0.14929 17.318 -6192.3
## - LotConfig   1    0.05873 17.228 -6191.9
## - GarageArea  1    0.06296 17.232 -6191.6
## - EnclosedPorch 1    0.06973 17.239 -6191.0

```

```

## - BsmtFinSF2      1    0.07262 17.241 -6190.8
## - Heating         2    0.10263 17.271 -6190.2
## - HeatingQC       4    0.16515 17.334 -6188.9
## - HalfBath        1    0.09959 17.268 -6188.5
## - Foundation      5    0.19806 17.367 -6188.2
## - GarageCars      1    0.10439 17.273 -6188.1
## - TotRmsAbvGrd    1    0.10619 17.275 -6187.9
## - FullBath        1    0.10726 17.276 -6187.8
## - YearRemodAdd     1    0.10749 17.276 -6187.8
## - BsmtUnfSF       1    0.11017 17.279 -6187.6
## - LowQualFinSF    1    0.11263 17.282 -6187.4
## - CentralAir      1    0.11495 17.284 -6187.2
## - Fireplaces      1    0.12498 17.294 -6186.3
## - HouseStyle      3    0.17416 17.343 -6186.2
## - BldgType        3    0.19119 17.360 -6184.7
## - KitchenQual     3    0.20155 17.370 -6183.9
## - WoodDeckSF      1    0.15442 17.323 -6183.8
## - BsmtFinSF1      1    0.15806 17.327 -6183.5
## - SaleType        3    0.20723 17.376 -6183.4
## - BsmtExposure    4    0.24191 17.411 -6182.5
## - ExterCond       3    0.23421 17.403 -6181.1
## - YearBuilt       1    0.20450 17.373 -6179.6
## - MiscFeature     4    0.28356 17.453 -6179.0
## - Exterior1st     7    0.35617 17.525 -6178.9
## - Condition1      3    0.27919 17.448 -6177.4
## - Functional      1    0.23305 17.402 -6177.2
## - LotArea         1    0.23641 17.405 -6176.9
## - BsmtFullBath    1    0.24986 17.419 -6175.8
## - ScreenPorch     1    0.34095 17.510 -6168.2
## - SaleCondition   5    0.50507 17.674 -6162.6
## - Condition2      3    0.57541 17.744 -6152.8
## - '1stFlrSF'      1    0.95022 18.119 -6118.3
## - OverallCond     1    1.08552 18.254 -6107.4
## - Neighborhood   15    1.60086 18.770 -6094.8
## - MSZoning        3    1.33622 18.505 -6091.5
## - '2ndFlrSF'      1    1.30545 18.474 -6089.9
## - OverallQual     1    1.85362 19.023 -6047.2
## - PoolArea        1    1.89054 19.059 -6044.4
## - PoolQC          3    2.42139 19.590 -6008.3
##
## Step:  AIC=-6197.57
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd +
##   MasVnrArea + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##   BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageType + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + PavedDrive + WoodDeckSF +
##   OpenPorchSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
##   PoolArea + PoolQC + MiscFeature + MiscVal + MoSold + YrSold +

```

```

##      SaleType + SaleCondition
##
##      Df Sum of Sq      RSS       AIC
## - PavedDrive      2    0.01900 17.251 -6200.0
## - Electrical      2    0.02114 17.253 -6199.8
## - Alley           2    0.02283 17.255 -6199.6
## - LotShape        1    0.00000 17.232 -6199.6
## - OpenPorchSF     1    0.00030 17.232 -6199.5
## - MoSold          1    0.00136 17.233 -6199.5
## - Exterior2nd     1    0.00138 17.233 -6199.5
## - MiscVal         1    0.00197 17.234 -6199.4
## - MasVnrArea      1    0.00281 17.235 -6199.3
## - BedroomAbvGr    1    0.00588 17.238 -6199.1
## - GarageType      3    0.05390 17.286 -6199.0
## - BsmtHalfBath    1    0.00873 17.241 -6198.8
## - YrSold          1    0.01139 17.243 -6198.6
## - LandContour     1    0.01222 17.244 -6198.5
## - GarageFinish    2    0.03997 17.272 -6198.2
## - Street          1    0.01650 17.248 -6198.2
## <none>                                17.232 -6197.6
## - BsmtQual        3    0.07499 17.307 -6197.2
## - GarageQual      2    0.05768 17.290 -6196.7
## - '3SsnPorch'     1    0.03607 17.268 -6196.5
## - LandSlope       2    0.06939 17.301 -6195.7
## - GarageCond      2    0.06999 17.302 -6195.7
## - KitchenAbvGr    1    0.04917 17.281 -6195.4
## - MSSubClass      3    0.09975 17.332 -6195.1
## - BsmtFinType1    5    0.14891 17.381 -6195.0
## - LotConfig       1    0.05816 17.290 -6194.6
## - GarageArea      1    0.06631 17.298 -6194.0
## - EnclosedPorch   1    0.06809 17.300 -6193.8
## - BsmtFinSF2      1    0.07687 17.309 -6193.1
## - Heating         2    0.10307 17.335 -6192.9
## - Foundation      5    0.19288 17.425 -6191.3
## - TotRmsAbvGrd    1    0.10241 17.334 -6190.9
## - HeatingQC       4    0.17448 17.406 -6190.9
## - GarageCars      1    0.10351 17.335 -6190.8
## - YearRemodAdd    1    0.10425 17.336 -6190.8
## - HalfBath        1    0.11213 17.344 -6190.1
## - FullBath        1    0.11267 17.345 -6190.1
## - LowQualFinSF    1    0.11442 17.346 -6189.9
## - CentralAir      1    0.11774 17.350 -6189.6
## - HouseStyle      3    0.17086 17.403 -6189.2
## - BsmtUnfSF       1    0.12588 17.358 -6188.9
## - Fireplaces      1    0.12791 17.360 -6188.8
## - BldgType        3    0.19110 17.423 -6187.5
## - WoodDeckSF      1    0.14525 17.377 -6187.3
## - KitchenQual     3    0.19707 17.429 -6187.0
## - SaleType        3    0.21166 17.444 -6185.7
## - BsmtExposure    4    0.23690 17.469 -6185.6
## - BsmtFinSF1      1    0.17107 17.403 -6185.1
## - ExterCond       3    0.22316 17.455 -6184.8
## - YearBuilt       1    0.19499 17.427 -6183.1
## - MiscFeature     4    0.28512 17.517 -6181.6

```



```

## - Exterior1st      7    0.35935 17.591 -6181.4
## - Condition1       3    0.28264 17.515 -6179.8
## - LotArea          1    0.23715 17.469 -6179.6
## - Functional       1    0.24581 17.478 -6178.9
## - BsmtFullBath     1    0.25419 17.486 -6178.2
## - ScreenPorch      1    0.34435 17.576 -6170.7
## - SaleCondition     5    0.52038 17.752 -6164.1
## - Condition2       3    0.58066 17.812 -6155.2
## - '1stFlrSF'       1    0.93427 18.166 -6122.5
## - OverallCond      1    1.07870 18.311 -6110.9
## - Neighborhood    15    1.59110 18.823 -6098.6
## - MSZoning         3    1.33481 18.567 -6094.6
## - '2ndFlrSF'       1    1.28748 18.519 -6094.4
## - PoolArea         1    1.89943 19.131 -6046.9
## - OverallQual      1    1.90705 19.139 -6046.3
## - PoolQC           3    2.43455 19.666 -6010.6
##
## Step:  AIC=-6199.96
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + Alley +
##   LotShape + LandContour + LotConfig + LandSlope + Neighborhood +
##   Condition1 + Condition2 + BldgType + HouseStyle + OverallQual +
##   OverallCond + YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd +
##   MasVnrArea + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##   BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageType + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + WoodDeckSF + OpenPorchSF +
##   EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##   MiscFeature + MiscVal + MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS     AIC
## - Alley      2    0.02033 17.271 -6202.2
## - Electrical  2    0.02041 17.271 -6202.2
## - LotShape    1    0.00001 17.251 -6202.0
## - OpenPorchSF 1    0.00018 17.251 -6201.9
## - Exterior2nd 1    0.00142 17.252 -6201.8
## - MoSold      1    0.00155 17.252 -6201.8
## - MasVnrArea  1    0.00265 17.253 -6201.7
## - MiscVal     1    0.00328 17.254 -6201.7
## - BedroomAbvGr 1    0.00597 17.257 -6201.5
## - GarageType  3    0.05481 17.306 -6201.3
## - BsmtHalfBath 1    0.00904 17.260 -6201.2
## - YrSold      1    0.01077 17.262 -6201.0
## - GarageFinish 2    0.03691 17.288 -6200.8
## - LandContour 1    0.01373 17.265 -6200.8
## - Street      1    0.01644 17.267 -6200.6
## <none>                17.251 -6200.0
## - BsmtQual    3    0.07579 17.327 -6199.6
## - '3SsnPorch' 1    0.03510 17.286 -6199.0
## - GarageQual  2    0.06020 17.311 -6198.9
## - LandSlope   2    0.06893 17.320 -6198.1
## - GarageCond  2    0.07296 17.324 -6197.8

```

```

## - MSSubClass      3    0.09919 17.350 -6197.6
## - KitchenAbvGr    1    0.05240 17.303 -6197.5
## - BsmtFinType1     5    0.15674 17.408 -6196.8
## - LotConfig        1    0.06258 17.314 -6196.7
## - EnclosedPorch    1    0.06688 17.318 -6196.3
## - GarageArea       1    0.06975 17.321 -6196.1
## - BsmtFinSF2       1    0.07458 17.326 -6195.7
## - Heating          2    0.09845 17.349 -6195.7
## - Foundation       5    0.19276 17.444 -6193.7
## - YearRemodAdd     1    0.09855 17.349 -6193.6
## - TotRmsAbvGrd     1    0.09914 17.350 -6193.6
## - GarageCars       1    0.10050 17.351 -6193.5
## - HeatingQC        4    0.17483 17.426 -6193.2
## - FullBath         1    0.11089 17.362 -6192.6
## - HalfBath         1    0.11148 17.362 -6192.6
## - CentralAir       1    0.12147 17.372 -6191.7
## - BsmtUnfSF        1    0.12305 17.374 -6191.6
## - LowQualFinSF     1    0.12503 17.376 -6191.4
## - HouseStyle       3    0.17354 17.424 -6191.3
## - Fireplaces       1    0.12986 17.381 -6191.0
## - BldgType         3    0.19274 17.444 -6189.7
## - WoodDeckSF       1    0.14571 17.397 -6189.7
## - KitchenQual      3    0.19867 17.450 -6189.2
## - SaleType         3    0.20891 17.460 -6188.4
## - BsmtExposure     4    0.23710 17.488 -6188.0
## - BsmtFinSF1       1    0.16637 17.417 -6187.9
## - ExterCond        3    0.22941 17.480 -6186.7
## - MiscFeature      4    0.28577 17.537 -6184.0
## - Exterior1st      7    0.36729 17.618 -6183.2
## - YearBuilt        1    0.22773 17.479 -6182.8
## - LotArea          1    0.23439 17.485 -6182.3
## - Condition1       3    0.29238 17.543 -6181.4
## - Functional       1    0.24940 17.500 -6181.0
## - BsmtFullBath     1    0.25041 17.501 -6180.9
## - ScreenPorch      1    0.34763 17.599 -6172.8
## - SaleCondition    5    0.51321 17.764 -6167.2
## - Condition2       3    0.57776 17.829 -6157.9
## - '1stFlrSF'       1    0.95717 18.208 -6123.1
## - OverallCond      1    1.08251 18.333 -6113.1
## - Neighborhood    15    1.60344 18.854 -6100.2
## - '2ndFlrSF'       1    1.30745 18.558 -6095.3
## - MSZoning         3    1.36564 18.616 -6094.7
## - PoolArea         1    1.89333 19.144 -6049.9
## - OverallQual      1    1.89998 19.151 -6049.4
## - PoolQC          3    2.43128 19.682 -6013.5
##
## Step:  AIC=-6202.24
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotShape +
##   LandContour + LotConfig + LandSlope + Neighborhood + Condition1 +
##   Condition2 + BldgType + HouseStyle + OverallQual + OverallCond +
##   YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd + MasVnrArea +
##   ExterCond + Foundation + BsmtQual + BsmtExposure + BsmtFinType1 +
##   BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC +
##   CentralAir + Electrical + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +

```

```

##      BsmtFullBath + BsmtHalfBath + FullBath + HalfBath + BedroomAbvGr +
##      KitchenAbvGr + KitchenQual + TotRmsAbvGrd + Functional +
##      Fireplaces + GarageType + GarageFinish + GarageCars + GarageArea +
##      GarageQual + GarageCond + WoodDeckSF + OpenPorchSF + EnclosedPorch +
##      '3SsnPorch' + ScreenPorch + PoolArea + PoolQC + MiscFeature +
##      MiscVal + MoSold + YrSold + SaleType + SaleCondition
##
##      Df Sum of Sq    RSS      AIC
## - Electrical      2   0.02057 17.292 -6204.5
## - LotShape        1   0.00000 17.271 -6204.2
## - OpenPorchSF     1   0.00040 17.272 -6204.2
## - MoSold          1   0.00126 17.273 -6204.1
## - Exterior2nd     1   0.00129 17.273 -6204.1
## - MiscVal         1   0.00294 17.274 -6204.0
## - MasVnrArea      1   0.00387 17.275 -6203.9
## - BedroomAbvGr   1   0.00474 17.276 -6203.8
## - GarageType      3   0.05360 17.325 -6203.7
## - BsmtHalfBath    1   0.00930 17.280 -6203.5
## - YrSold          1   0.01093 17.282 -6203.3
## - LandContour     1   0.01123 17.282 -6203.3
## - GarageFinish    2   0.03632 17.308 -6203.2
## - Street          1   0.01719 17.288 -6202.8
## <none>                                17.271 -6202.2
## - BsmtQual        3   0.07321 17.344 -6202.1
## - '3SsnPorch'     1   0.03432 17.305 -6201.3
## - GarageQual      2   0.06390 17.335 -6200.8
## - LandSlope       2   0.06786 17.339 -6200.5
## - GarageCond      2   0.06959 17.341 -6200.4
## - MSSubClass      3   0.10354 17.375 -6199.5
## - KitchenAbvGr    1   0.05723 17.328 -6199.4
## - BsmtFinType1    5   0.15262 17.424 -6199.4
## - LotConfig       1   0.05942 17.331 -6199.2
## - GarageArea      1   0.06540 17.337 -6198.7
## - EnclosedPorch   1   0.07225 17.343 -6198.1
## - BsmtFinSF2      1   0.07248 17.344 -6198.1
## - Heating         2   0.09705 17.368 -6198.1
## - TotRmsAbvGrd    1   0.09395 17.365 -6196.3
## - GarageCars      1   0.10581 17.377 -6195.3
## - YearRemodAdd     1   0.10599 17.377 -6195.3
## - Foundation      5   0.20197 17.473 -6195.3
## - HeatingQC       4   0.17815 17.449 -6195.3
## - FullBath        1   0.11009 17.381 -6195.0
## - HalfBath        1   0.11060 17.382 -6194.9
## - CentralAir      1   0.11763 17.389 -6194.3
## - BsmtUnfSF       1   0.12187 17.393 -6194.0
## - HouseStyle      3   0.17327 17.445 -6193.7
## - LowQualFinSF    1   0.12699 17.398 -6193.5
## - Fireplaces      1   0.12822 17.399 -6193.4
## - WoodDeckSF      1   0.14270 17.414 -6192.2
## - BldgType        3   0.19282 17.464 -6192.0
## - KitchenQual     3   0.19998 17.471 -6191.4
## - SaleType        3   0.20448 17.476 -6191.1
## - BsmtFinSF1      1   0.16582 17.437 -6190.3
## - BsmtExposure    4   0.23801 17.509 -6190.3

```

```

## - ExterCond      3  0.22704 17.498 -6189.2
## - MiscFeature    4  0.28629 17.558 -6186.2
## - YearBuilt      1  0.21850 17.490 -6185.9
## - Exterior1st    7  0.37554 17.647 -6184.8
## - LotArea        1  0.23373 17.505 -6184.6
## - Condition1     3  0.29417 17.565 -6183.6
## - Functional     1  0.24980 17.521 -6183.3
## - BsmtFullBath   1  0.25639 17.528 -6182.7
## - ScreenPorch    1  0.34741 17.619 -6175.2
## - SaleCondition   5  0.50410 17.775 -6170.2
## - Condition2     3  0.58210 17.853 -6159.8
## - '1stFlrSF'     1  0.97013 18.241 -6124.5
## - OverallCond    1  1.08078 18.352 -6115.6
## - Neighborhood   15 1.60177 18.873 -6102.8
## - MSZoning        3  1.36679 18.638 -6097.0
## - '2ndFlrSF'     1  1.34542 18.617 -6094.7
## - PoolArea       1  1.89907 19.170 -6051.9
## - OverallQual     1  1.90894 19.180 -6051.2
## - PoolQC         3  2.43196 19.703 -6015.9
##
## Step: AIC=-6204.5
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotShape +
##   LandContour + LotConfig + LandSlope + Neighborhood + Condition1 +
##   Condition2 + BldgType + HouseStyle + OverallQual + OverallCond +
##   YearBuilt + YearRemodAdd + Exterior1st + Exterior2nd + MasVnrArea +
##   ExterCond + Foundation + BsmtQual + BsmtExposure + BsmtFinType1 +
##   BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC +
##   CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath +
##   BsmtHalfBath + FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr +
##   KitchenQual + TotRmsAbvGrd + Functional + Fireplaces + GarageType +
##   GarageFinish + GarageCars + GarageArea + GarageQual + GarageCond +
##   WoodDeckSF + OpenPorchSF + EnclosedPorch + '3SsnPorch' +
##   ScreenPorch + PoolArea + PoolQC + MiscFeature + MiscVal +
##   MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq  RSS    AIC
## - LotShape      1  0.00007 17.292 -6206.5
## - OpenPorchSF   1  0.00042 17.292 -6206.5
## - MoSold        1  0.00128 17.293 -6206.4
## - Exterior2nd   1  0.00161 17.293 -6206.4
## - MiscVal       1  0.00274 17.294 -6206.3
## - MasVnrArea    1  0.00419 17.296 -6206.1
## - GarageType    3  0.05293 17.345 -6206.0
## - BedroomAbvGr  1  0.00562 17.297 -6206.0
## - BsmtHalfBath  1  0.00930 17.301 -6205.7
## - LandContour   1  0.01066 17.302 -6205.6
## - GarageFinish  2  0.03511 17.327 -6205.5
## - YrSold        1  0.01217 17.304 -6205.5
## - Street        1  0.01683 17.309 -6205.1
## <none>                17.292 -6204.5
## - BsmtQual      3  0.07580 17.368 -6204.1
## - '3SsnPorch'   1  0.03506 17.327 -6203.5
## - GarageQual    2  0.06263 17.354 -6203.2
## - LandSlope     2  0.06828 17.360 -6202.7

```

```

## - GarageCond      2    0.07007 17.362 -6202.6
## - KitchenAbvGr    1    0.05152 17.343 -6202.2
## - MSSubClass      3    0.10161 17.393 -6201.9
## - LotConfig       1    0.05701 17.349 -6201.7
## - BsmtFinType1    5    0.15348 17.445 -6201.6
## - GarageArea      1    0.06388 17.356 -6201.1
## - Heating         2    0.09137 17.383 -6200.8
## - BsmtFinSF2      1    0.07249 17.364 -6200.4
## - EnclosedPorch   1    0.07384 17.366 -6200.3
## - TotRmsAbvGrd    1    0.09461 17.386 -6198.5
## - YearRemodAdd     1    0.09948 17.391 -6198.1
## - GarageCars      1    0.10357 17.395 -6197.8
## - Foundation      5    0.19931 17.491 -6197.8
## - FullBath        1    0.10597 17.398 -6197.6
## - CentralAir      1    0.10685 17.399 -6197.5
## - HeatingQC       4    0.17874 17.471 -6197.5
## - HalfBath        1    0.10776 17.399 -6197.4
## - BsmtUnfSF       1    0.12524 17.417 -6196.0
## - HouseStyle      3    0.17339 17.465 -6195.9
## - Fireplaces      1    0.12615 17.418 -6195.9
## - LowQualFinSF    1    0.12747 17.419 -6195.8
## - WoodDeckSF      1    0.14253 17.434 -6194.5
## - BldgType        3    0.19181 17.484 -6194.4
## - KitchenQual     3    0.20056 17.492 -6193.7
## - SaleType        3    0.20255 17.494 -6193.5
## - BsmtExposure    4    0.23445 17.526 -6192.8
## - BsmtFinSF1      1    0.16676 17.459 -6192.5
## - ExterCond       3    0.23229 17.524 -6191.0
## - YearBuilt       1    0.21358 17.505 -6188.6
## - MiscFeature     4    0.28886 17.581 -6188.3
## - LotArea         1    0.23422 17.526 -6186.9
## - Exterior1st     7    0.38426 17.676 -6186.4
## - Condition1      3    0.28822 17.580 -6186.4
## - Functional      1    0.25539 17.547 -6185.1
## - BsmtFullBath    1    0.25596 17.548 -6185.0
## - ScreenPorch     1    0.34767 17.640 -6177.4
## - SaleCondition   5    0.50891 17.801 -6172.2
## - Condition2      3    0.58181 17.874 -6162.2
## - '1stFlrSF'      1    0.96713 18.259 -6127.0
## - OverallCond     1    1.07683 18.369 -6118.3
## - Neighborhood   15    1.59159 18.883 -6105.9
## - MSZoning        3    1.35879 18.651 -6100.1
## - '2ndFlrSF'      1    1.35947 18.651 -6096.0
## - PoolArea        1    1.90915 19.201 -6053.6
## - OverallQual     1    1.91092 19.203 -6053.5
## - PoolQC         3    2.44126 19.733 -6017.7
##
## Step:  AIC=-6206.5
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + Exterior2nd + MasVnrArea + ExterCond +
##   Foundation + BsmtQual + BsmtExposure + BsmtFinType1 + BsmtFinSF1 +
##   BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC + CentralAir +

```

```

##      '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##      FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##      TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##      GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +
##      OpenPorchSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
##      PoolArea + PoolQC + MiscFeature + MiscVal + MoSold + YrSold +
##      SaleType + SaleCondition
##
##      Df Sum of Sq    RSS      AIC
## - OpenPorchSF      1    0.00042 17.292 -6208.5
## - MoSold            1    0.00130 17.293 -6208.4
## - Exterior2nd       1    0.00161 17.294 -6208.4
## - MiscVal           1    0.00275 17.295 -6208.3
## - MasVnrArea        1    0.00420 17.296 -6208.1
## - GarageType        3    0.05297 17.345 -6208.0
## - BedroomAbvGr      1    0.00563 17.297 -6208.0
## - BsmtHalfBath      1    0.00932 17.301 -6207.7
## - LandContour       1    0.01060 17.302 -6207.6
## - GarageFinish      2    0.03522 17.327 -6207.5
## - YrSold            1    0.01224 17.304 -6207.5
## - Street            1    0.01685 17.309 -6207.1
## <none>                                17.292 -6206.5
## - BsmtQual          3    0.07584 17.368 -6206.1
## - '3SsnPorch'       1    0.03502 17.327 -6205.5
## - GarageQual        2    0.06274 17.355 -6205.2
## - LandSlope         2    0.06826 17.360 -6204.7
## - GarageCond        2    0.07006 17.362 -6204.6
## - KitchenAbvGr      1    0.05146 17.343 -6204.2
## - MSSubClass        3    0.10167 17.393 -6203.9
## - BsmtFinType1      5    0.15383 17.446 -6203.6
## - LotConfig         1    0.06076 17.353 -6203.4
## - GarageArea        1    0.06394 17.356 -6203.1
## - Heating           2    0.09133 17.383 -6202.8
## - BsmtFinSF2        1    0.07256 17.364 -6202.4
## - EnclosedPorch     1    0.07386 17.366 -6202.3
## - TotRmsAbvGrd      1    0.09461 17.387 -6200.5
## - YearRemodAdd      1    0.09941 17.391 -6200.1
## - GarageCars        1    0.10351 17.395 -6199.8
## - Foundation        5    0.19928 17.491 -6199.8
## - FullBath          1    0.10600 17.398 -6199.6
## - CentralAir        1    0.10678 17.399 -6199.5
## - HeatingQC         4    0.17873 17.471 -6199.5
## - HalfBath          1    0.10783 17.400 -6199.4
## - BsmtUnfSF         1    0.12526 17.417 -6198.0
## - HouseStyle        3    0.17354 17.465 -6197.9
## - Fireplaces        1    0.12620 17.418 -6197.9
## - LowQualFinSF      1    0.12761 17.419 -6197.8
## - WoodDeckSF        1    0.14293 17.435 -6196.5
## - BldgType          3    0.19181 17.484 -6196.4
## - KitchenQual       3    0.20069 17.492 -6195.6
## - SaleType          3    0.20252 17.494 -6195.5
## - BsmtExposure      4    0.23526 17.527 -6194.8
## - BsmtFinSF1        1    0.16684 17.459 -6194.5
## - ExterCond         3    0.23227 17.524 -6193.0

```

```

## - YearBuilt      1  0.21534 17.507 -6190.4
## - MiscFeature    4  0.28961 17.581 -6190.2
## - LotArea        1  0.23579 17.528 -6188.7
## - Exterior1st    7  0.38421 17.676 -6188.4
## - Condition1     3  0.28836 17.580 -6188.4
## - Functional     1  0.25532 17.547 -6187.1
## - BsmtFullBath   1  0.25593 17.548 -6187.0
## - ScreenPorch    1  0.34902 17.641 -6179.3
## - SaleCondition   5  0.50918 17.801 -6174.1
## - Condition2     3  0.58213 17.874 -6164.2
## - '1stFlrSF'     1  0.96782 18.260 -6129.0
## - OverallCond    1  1.07867 18.370 -6120.1
## - Neighborhood   15  1.59750 18.889 -6107.5
## - MSZoning        3  1.36205 18.654 -6101.8
## - '2ndFlrSF'     1  1.35952 18.651 -6098.0
## - PoolArea       1  1.91144 19.203 -6055.4
## - OverallQual     1  1.91179 19.204 -6055.4
## - PoolQC         3  2.44344 19.735 -6019.5
##
## Step:  AIC=-6208.46
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + Exterior2nd + MasVnrArea + ExterCond +
##   Foundation + BsmtQual + BsmtExposure + BsmtFinType1 + BsmtFinSF1 +
##   BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC + CentralAir +
##   '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##   FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +
##   EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##   MiscFeature + MiscVal + MoSold + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS     AIC
## - MoSold      1  0.00140 17.294 -6210.3
## - Exterior2nd  1  0.00165 17.294 -6210.3
## - MiscVal      1  0.00272 17.295 -6210.2
## - MasVnrArea   1  0.00401 17.296 -6210.1
## - GarageType   3  0.05269 17.345 -6210.0
## - BedroomAbvGr 1  0.00567 17.298 -6210.0
## - BsmtHalfBath 1  0.00934 17.302 -6209.7
## - LandContour  1  0.01067 17.303 -6209.6
## - GarageFinish  2  0.03550 17.328 -6209.5
## - YrSold       1  0.01263 17.305 -6209.4
## - Street       1  0.01678 17.309 -6209.0
## <none>                17.292 -6208.5
## - BsmtQual     3  0.07698 17.369 -6208.0
## - '3SsnPorch'  1  0.03484 17.327 -6207.5
## - GarageQual   2  0.06296 17.355 -6207.2
## - LandSlope    2  0.06862 17.361 -6206.7
## - GarageCond   2  0.07038 17.363 -6206.5
## - KitchenAbvGr 1  0.05174 17.344 -6206.1
## - MSSubClass   3  0.10167 17.394 -6205.9
## - BsmtFinType1 5  0.15500 17.447 -6205.4

```

```

## - LotConfig      1    0.06156 17.354 -6205.3
## - GarageArea     1    0.06426 17.357 -6205.0
## - Heating        2    0.09197 17.384 -6204.7
## - EnclosedPorch  1    0.07345 17.366 -6204.3
## - BsmtFinSF2     1    0.07385 17.366 -6204.2
## - TotRmsAbvGrd   1    0.09452 17.387 -6202.5
## - YearRemodAdd    1    0.09941 17.392 -6202.1
## - GarageCars     1    0.10320 17.395 -6201.8
## - Foundation     5    0.20119 17.494 -6201.6
## - CentralAir     1    0.10636 17.399 -6201.5
## - FullBath       1    0.10663 17.399 -6201.5
## - HeatingQC      4    0.17877 17.471 -6201.4
## - HalfBath       1    0.10983 17.402 -6201.2
## - HouseStyle     3    0.17369 17.466 -6199.9
## - Fireplaces     1    0.12706 17.419 -6199.8
## - LowQualFinSF   1    0.12774 17.420 -6199.7
## - BsmtUnfSF      1    0.12827 17.421 -6199.7
## - WoodDeckSF     1    0.14259 17.435 -6198.5
## - BldgType       3    0.19218 17.485 -6198.3
## - KitchenQual    3    0.20053 17.493 -6197.6
## - SaleType       3    0.20210 17.494 -6197.5
## - BsmtExposure   4    0.23542 17.528 -6196.7
## - BsmtFinSF1     1    0.16877 17.461 -6196.3
## - ExterCond      3    0.23225 17.524 -6195.0
## - YearBuilt      1    0.21499 17.507 -6192.4
## - MiscFeature    4    0.29003 17.582 -6192.2
## - LotArea        1    0.23658 17.529 -6190.6
## - Exterior1st    7    0.38380 17.676 -6190.4
## - Condition1     3    0.28799 17.580 -6190.3
## - Functional     1    0.25550 17.548 -6189.0
## - BsmtFullBath   1    0.25765 17.550 -6188.9
## - ScreenPorch    1    0.35020 17.642 -6181.2
## - SaleCondition  5    0.50880 17.801 -6176.1
## - Condition2     3    0.58412 17.876 -6166.0
## - '1stFlrSF'     1    0.96961 18.262 -6130.8
## - OverallCond    1    1.08217 18.374 -6121.8
## - Neighborhood   15    1.59863 18.891 -6109.4
## - MSZoning        3    1.36267 18.655 -6103.7
## - '2ndFlrSF'     1    1.36136 18.654 -6099.8
## - OverallQual    1    1.91163 19.204 -6057.4
## - PoolArea       1    1.91966 19.212 -6056.8
## - PoolQC         3    2.44563 19.738 -6021.3
##
## Step:  AIC=-6210.34
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + Exterior2nd + MasVnrArea + ExterCond +
##   Foundation + BsmtQual + BsmtExposure + BsmtFinType1 + BsmtFinSF1 +
##   BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC + CentralAir +
##   '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##   FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +

```



```

##      EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##      MiscFeature + MiscVal + YrSold + SaleType + SaleCondition
##
##      Df Sum of Sq      RSS       AIC
## - Exterior2nd      1    0.00168 17.295 -6212.2
## - MiscVal          1    0.00259 17.296 -6212.1
## - MasVnrArea       1    0.00386 17.297 -6212.0
## - GarageType       3    0.05238 17.346 -6211.9
## - BedroomAbvGr     1    0.00542 17.299 -6211.9
## - BsmtHalfBath      1    0.00957 17.303 -6211.5
## - LandContour       1    0.01047 17.304 -6211.5
## - GarageFinish      2    0.03496 17.329 -6211.4
## - YrSold            1    0.01403 17.308 -6211.2
## - Street           1    0.01675 17.310 -6210.9
## <none>                                17.294 -6210.3
## - BsmtQual         3    0.07617 17.370 -6209.9
## - '3SsnPorch'      1    0.03524 17.329 -6209.4
## - GarageQual       2    0.06242 17.356 -6209.1
## - LandSlope        2    0.06803 17.362 -6208.6
## - GarageCond       2    0.06985 17.363 -6208.5
## - KitchenAbvGr     1    0.05090 17.345 -6208.1
## - MSSubClass       3    0.10206 17.396 -6207.8
## - BsmtFinType1     5    0.15425 17.448 -6207.4
## - LotConfig        1    0.06121 17.355 -6207.2
## - GarageArea       1    0.06397 17.358 -6207.0
## - Heating          2    0.09259 17.386 -6206.5
## - EnclosedPorch    1    0.07248 17.366 -6206.2
## - BsmtFinSF2       1    0.07346 17.367 -6206.2
## - TotRmsAbvGrd     1    0.09332 17.387 -6204.5
## - YearRemodAdd     1    0.09979 17.393 -6203.9
## - GarageCars       1    0.10341 17.397 -6203.6
## - Foundation       5    0.20059 17.494 -6203.5
## - HeatingQC        4    0.17775 17.471 -6203.4
## - CentralAir       1    0.10669 17.400 -6203.4
## - FullBath         1    0.10731 17.401 -6203.3
## - HalfBath         1    0.10947 17.403 -6203.1
## - HouseStyle       3    0.17415 17.468 -6201.7
## - LowQualFinSF     1    0.12713 17.421 -6201.6
## - BsmtUnfSF        1    0.12762 17.421 -6201.6
## - Fireplaces       1    0.12783 17.422 -6201.6
## - WoodDeckSF       1    0.14339 17.437 -6200.3
## - BldgType         3    0.19190 17.486 -6200.2
## - KitchenQual      3    0.20044 17.494 -6199.5
## - SaleType         3    0.20299 17.497 -6199.3
## - BsmtExposure     4    0.23486 17.529 -6198.6
## - BsmtFinSF1       1    0.16837 17.462 -6198.2
## - ExterCond        3    0.23348 17.527 -6196.8
## - YearBuilt        1    0.21517 17.509 -6194.3
## - MiscFeature      4    0.29276 17.586 -6193.8
## - LotArea          1    0.23610 17.530 -6192.5
## - Exterior1st      7    0.38329 17.677 -6192.3
## - Condition1       3    0.28984 17.584 -6192.1
## - Functional       1    0.25570 17.549 -6190.9
## - BsmtFullBath     1    0.25853 17.552 -6190.7

```

```

## - ScreenPorch      1    0.35083 17.645 -6183.0
## - SaleCondition    5    0.51027 17.804 -6177.9
## - Condition2       3    0.58447 17.878 -6167.8
## - '1stFlrSF'       1    0.97241 18.266 -6132.5
## - OverallCond      1    1.08079 18.375 -6123.8
## - Neighborhood    15    1.60594 18.900 -6110.7
## - MSZoning         3    1.36251 18.656 -6105.6
## - '2ndFlrSF'       1    1.36387 18.657 -6101.5
## - PoolArea         1    1.92829 19.222 -6058.0
## - OverallQual      1    1.93483 19.229 -6057.5
## - PoolQC           3    2.45192 19.746 -6022.8
##
## Step:  AIC=-6212.2
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + MasVnrArea + ExterCond + Foundation +
##   BsmtQual + BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 +
##   BsmtUnfSF + Heating + HeatingQC + CentralAir + '1stFlrSF' +
##   '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##   FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +
##   EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##   MiscFeature + MiscVal + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS     AIC
## - MiscVal      1    0.00244 17.298 -6214.0
## - MasVnrArea   1    0.00402 17.299 -6213.9
## - GarageType   3    0.05234 17.348 -6213.8
## - BedroomAbvGr 1    0.00541 17.301 -6213.7
## - BsmtHalfBath 1    0.00986 17.305 -6213.4
## - LandContour  1    0.01070 17.306 -6213.3
## - GarageFinish 2    0.03530 17.331 -6213.2
## - YrSold        1    0.01435 17.310 -6213.0
## - Street        1    0.01671 17.312 -6212.8
## <none>                17.295 -6212.2
## - BsmtQual      3    0.07694 17.372 -6211.7
## - '3SsnPorch'   1    0.03466 17.330 -6211.3
## - GarageQual    2    0.06234 17.358 -6210.9
## - LandSlope     2    0.06824 17.364 -6210.5
## - GarageCond    2    0.06970 17.365 -6210.3
## - KitchenAbvGr  1    0.05086 17.346 -6209.9
## - MSSubClass    3    0.10108 17.396 -6209.7
## - BsmtFinType1  5    0.15335 17.449 -6209.3
## - LotConfig     1    0.06069 17.356 -6209.1
## - GarageArea    1    0.06391 17.359 -6208.8
## - Heating       2    0.09173 17.387 -6208.5
## - EnclosedPorch 1    0.07187 17.367 -6208.1
## - BsmtFinSF2    1    0.07333 17.369 -6208.0
## - TotRmsAbvGrd  1    0.09396 17.389 -6206.3
## - YearRemodAdd  1    0.09868 17.394 -6205.9
## - GarageCars    1    0.10251 17.398 -6205.6
## - CentralAir    1    0.10541 17.401 -6205.3

```

```

## - Foundation      5    0.20171 17.497 -6205.3
## - FullBath        1    0.10653 17.402 -6205.2
## - HeatingQC       4    0.17966 17.475 -6205.1
## - HalfBath        1    0.10977 17.405 -6205.0
## - HouseStyle      3    0.17357 17.469 -6203.6
## - BsmtUnfSF       1    0.12811 17.424 -6203.4
## - LowQualFinSF    1    0.12855 17.424 -6203.4
## - Fireplaces      1    0.12938 17.425 -6203.3
## - WoodDeckSF      1    0.14293 17.438 -6202.2
## - BldgType        3    0.19553 17.491 -6201.8
## - KitchenQual     3    0.20023 17.496 -6201.4
## - SaleType        3    0.20550 17.501 -6201.0
## - BsmtExposure    4    0.23319 17.529 -6200.6
## - BsmtFinSF1      1    0.16850 17.464 -6200.0
## - ExterCond       3    0.23232 17.528 -6198.7
## - YearBuilt       1    0.21991 17.515 -6195.8
## - MiscFeature     4    0.29280 17.588 -6195.7
## - LotArea         1    0.23648 17.532 -6194.4
## - Exterior1st     7    0.38356 17.679 -6194.2
## - Condition1      3    0.29139 17.587 -6193.8
## - Functional      1    0.25862 17.554 -6192.5
## - BsmtFullBath    1    0.26143 17.557 -6192.3
## - ScreenPorch     1    0.35164 17.647 -6184.8
## - SaleCondition   5    0.51049 17.806 -6179.7
## - Condition2      3    0.58322 17.879 -6169.8
## - '1stFlrSF'      1    0.97131 18.267 -6134.4
## - OverallCond     1    1.08913 18.384 -6125.0
## - Neighborhood    15    1.61352 18.909 -6112.0
## - MSZoning         3    1.36209 18.657 -6107.5
## - '2ndFlrSF'      1    1.36219 18.657 -6103.5
## - PoolArea        1    1.92715 19.223 -6060.0
## - OverallQual     1    1.94204 19.237 -6058.8
## - PoolQC          3    2.45025 19.746 -6024.8
##
## Step:  AIC=-6213.99
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + MasVnrArea + ExterCond + Foundation +
##   BsmtQual + BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 +
##   BsmtUnfSF + Heating + HeatingQC + CentralAir + '1stFlrSF' +
##   '2ndFlrSF' + LowQualFinSF + BsmtFullBath + BsmtHalfBath +
##   FullBath + HalfBath + BedroomAbvGr + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageType + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +
##   EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##   MiscFeature + YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS   AIC
## - MasVnrArea    1    0.00418 17.302 -6215.6
## - GarageType    3    0.05190 17.350 -6215.6
## - BedroomAbvGr  1    0.00530 17.303 -6215.5
## - BsmtHalfBath  1    0.00969 17.308 -6215.2
## - LandContour   1    0.01077 17.309 -6215.1

```

## - GarageFinish	2	0.03582	17.334	-6215.0
## - YrSold	1	0.01436	17.312	-6214.8
## - Street	1	0.01728	17.315	-6214.5
## <none>			17.298	-6214.0
## - BsmtQual	3	0.07625	17.374	-6213.6
## - '3SsnPorch'	1	0.03471	17.332	-6213.1
## - GarageQual	2	0.06187	17.360	-6212.8
## - LandSlope	2	0.06758	17.365	-6212.3
## - GarageCond	2	0.07049	17.368	-6212.1
## - KitchenAbvGr	1	0.05124	17.349	-6211.7
## - MSSubClass	3	0.10078	17.399	-6211.5
## - BsmtFinType1	5	0.15201	17.450	-6211.2
## - LotConfig	1	0.06105	17.359	-6210.9
## - GarageArea	1	0.06359	17.361	-6210.6
## - Heating	2	0.09243	17.390	-6210.2
## - EnclosedPorch	1	0.07066	17.368	-6210.0
## - BsmtFinSF2	1	0.07418	17.372	-6209.7
## - TotRmsAbvGrd	1	0.09317	17.391	-6208.2
## - YearRemodAdd	1	0.10086	17.399	-6207.5
## - GarageCars	1	0.10245	17.400	-6207.4
## - CentralAir	1	0.10514	17.403	-6207.1
## - HeatingQC	4	0.17788	17.476	-6207.1
## - FullBath	1	0.10683	17.405	-6207.0
## - HalfBath	1	0.10800	17.406	-6206.9
## - Foundation	5	0.20525	17.503	-6206.8
## - HouseStyle	3	0.17276	17.471	-6205.5
## - LowQualFinSF	1	0.12679	17.425	-6205.3
## - Fireplaces	1	0.12990	17.428	-6205.1
## - BsmtUnfSF	1	0.13023	17.428	-6205.0
## - WoodDeckSF	1	0.14234	17.440	-6204.0
## - BldgType	3	0.19638	17.494	-6203.5
## - KitchenQual	3	0.19995	17.498	-6203.2
## - SaleType	3	0.20486	17.503	-6202.8
## - BsmtExposure	4	0.23374	17.532	-6202.4
## - BsmtFinSF1	1	0.17152	17.469	-6201.6
## - ExterCond	3	0.22989	17.528	-6200.7
## - MiscFeature	4	0.29047	17.588	-6197.7
## - YearBuilt	1	0.22006	17.518	-6197.5
## - LotArea	1	0.23653	17.534	-6196.2
## - Exterior1st	7	0.38625	17.684	-6195.8
## - Condition1	3	0.29058	17.588	-6195.7
## - Functional	1	0.25626	17.554	-6194.5
## - BsmtFullBath	1	0.26326	17.561	-6193.9
## - ScreenPorch	1	0.35541	17.653	-6186.3
## - SaleCondition	5	0.51212	17.810	-6181.4
## - Condition2	3	0.58560	17.883	-6171.4
## - '1stFlrSF'	1	0.96889	18.267	-6136.4
## - OverallCond	1	1.09369	18.392	-6126.5
## - Neighborhood	15	1.61670	18.915	-6113.5
## - MSZoning	3	1.36805	18.666	-6108.9
## - '2ndFlrSF'	1	1.36072	18.659	-6105.4
## - PoolArea	1	1.93128	19.229	-6061.5
## - OverallQual	1	1.95580	19.254	-6059.6
## - PoolQC	3	2.45428	19.752	-6026.3

```

##
## Step:  AIC=-6215.64
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + ExterCond + Foundation + BsmtQual +
##   BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF +
##   Heating + HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##   BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageType + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + WoodDeckSF + EnclosedPorch +
##   '3SsnPorch' + ScreenPorch + PoolArea + PoolQC + MiscFeature +
##   YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS   AIC
## - GarageType      3   0.05136 17.353 -6217.3
## - BedroomAbvGr    1   0.00552 17.308 -6217.2
## - BsmtHalfBath     1   0.01010 17.312 -6216.8
## - LandContour      1   0.01119 17.313 -6216.7
## - GarageFinish     2   0.03655 17.338 -6216.6
## - YrSold           1   0.01412 17.316 -6216.5
## - Street           1   0.01771 17.320 -6216.1
## <none>                        17.302 -6215.6
## - BsmtQual         3   0.07900 17.381 -6215.0
## - '3SsnPorch'      1   0.03540 17.337 -6214.7
## - GarageQual       2   0.06027 17.362 -6214.6
## - LandSlope        2   0.06790 17.370 -6213.9
## - GarageCond       2   0.06943 17.371 -6213.8
## - KitchenAbvGr     1   0.05166 17.354 -6213.3
## - MSSubClass       3   0.10027 17.402 -6213.2
## - BsmtFinType1     5   0.15497 17.457 -6212.6
## - LotConfig        1   0.06126 17.363 -6212.5
## - GarageArea       1   0.06384 17.366 -6212.3
## - Heating          2   0.09142 17.393 -6211.9
## - EnclosedPorch    1   0.07109 17.373 -6211.7
## - BsmtFinSF2       1   0.07372 17.376 -6211.4
## - TotRmsAbvGrd     1   0.09300 17.395 -6209.8
## - YearRemodAdd     1   0.09936 17.401 -6209.3
## - CentralAir       1   0.10446 17.406 -6208.9
## - GarageCars       1   0.10477 17.407 -6208.8
## - FullBath         1   0.10685 17.409 -6208.7
## - HeatingQC        4   0.17907 17.481 -6208.6
## - Foundation       5   0.20426 17.506 -6208.5
## - HalfBath         1   0.10904 17.411 -6208.5
## - HouseStyle       3   0.17108 17.473 -6207.3
## - LowQualFinSF     1   0.12538 17.427 -6207.1
## - Fireplaces       1   0.13207 17.434 -6206.5
## - BsmtUnfSF        1   0.13272 17.435 -6206.5
## - WoodDeckSF       1   0.14367 17.446 -6205.6
## - BldgType         3   0.19418 17.496 -6205.3
## - KitchenQual      3   0.20048 17.503 -6204.8
## - SaleType         3   0.20455 17.506 -6204.5
## - BsmtExposure     4   0.23754 17.540 -6203.7

```

```

## - BsmtFinSF1      1   0.17474 17.477 -6203.0
## - ExterCond       3   0.22774 17.530 -6202.5
## - MiscFeature     4   0.29002 17.592 -6199.4
## - YearBuilt       1   0.22454 17.526 -6198.8
## - Exterior1st     7   0.38207 17.684 -6197.8
## - LotArea         1   0.23825 17.540 -6197.7
## - Condition1      3   0.29054 17.593 -6197.3
## - BsmtFullBath    1   0.26052 17.562 -6195.8
## - Functional      1   0.26110 17.563 -6195.8
## - ScreenPorch     1   0.36006 17.662 -6187.6
## - SaleCondition   5   0.51332 17.815 -6183.0
## - Condition2      3   0.58151 17.884 -6173.4
## - '1stFlrSF'      1   0.97966 18.282 -6137.2
## - OverallCond     1   1.09132 18.393 -6128.3
## - Neighborhood   15   1.64576 18.948 -6113.0
## - MSZoning         3   1.36388 18.666 -6110.9
## - '2ndFlrSF'      1   1.37958 18.681 -6105.6
## - PoolArea        1   1.92874 19.231 -6063.3
## - OverallQual     1   1.97226 19.274 -6060.0
## - PoolQC          3   2.45020 19.752 -6028.3
##
## Step: AIC=-6217.31
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + ExterCond + Foundation + BsmtQual +
##   BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF +
##   Heating + HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
##   BedroomAbvGr + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageFinish + GarageCars + GarageArea +
##   GarageQual + GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' +
##   ScreenPorch + PoolArea + PoolQC + MiscFeature + YrSold +
##   SaleType + SaleCondition
##
##           Df Sum of Sq   RSS   AIC
## - BedroomAbvGr    1    0.00595 17.359 -6218.8
## - BsmtHalfBath    1    0.00937 17.363 -6218.5
## - LandContour     1    0.01096 17.364 -6218.4
## - YrSold           1    0.01636 17.370 -6217.9
## - GarageFinish     2    0.04487 17.398 -6217.5
## <none>                17.353 -6217.3
## - Street           1    0.02441 17.378 -6217.3
## - LandSlope        2    0.05720 17.410 -6216.5
## - BsmtQual         3    0.08178 17.435 -6216.5
## - '3SsnPorch'      1    0.03597 17.389 -6216.3
## - GarageQual       2    0.06346 17.417 -6216.0
## - GarageCond       2    0.06883 17.422 -6215.5
## - MSSubClass       3    0.09436 17.448 -6215.4
## - KitchenAbvGr     1    0.05563 17.409 -6214.6
## - GarageArea       1    0.05644 17.410 -6214.6
## - LotConfig        1    0.05779 17.411 -6214.5
## - BsmtFinType1     5    0.15603 17.509 -6214.2
## - Heating          2    0.08971 17.443 -6213.8

```

```

## - EnclosedPorch 1 0.07184 17.425 -6213.3
## - BsmtFinSF2 1 0.08210 17.435 -6212.4
## - YearRemodAdd 1 0.09637 17.450 -6211.2
## - TotRmsAbvGrd 1 0.09816 17.451 -6211.1
## - HeatingQC 4 0.17365 17.527 -6210.8
## - CentralAir 1 0.10612 17.459 -6210.4
## - GarageCars 1 0.10697 17.460 -6210.3
## - FullBath 1 0.11140 17.465 -6210.0
## - HouseStyle 3 0.15966 17.513 -6209.9
## - HalfBath 1 0.11408 17.467 -6209.7
## - LowQualFinSF 1 0.11862 17.472 -6209.4
## - Foundation 5 0.21571 17.569 -6209.3
## - Fireplaces 1 0.13282 17.486 -6208.2
## - KitchenQual 3 0.18433 17.538 -6207.9
## - WoodDeckSF 1 0.14853 17.502 -6206.9
## - BldgType 3 0.20033 17.554 -6206.6
## - BsmtUnfSF 1 0.15267 17.506 -6206.5
## - BsmtExposure 4 0.22923 17.583 -6206.2
## - SaleType 3 0.20602 17.559 -6206.1
## - ExterCond 3 0.22610 17.579 -6204.4
## - BsmtFinSF1 1 0.20230 17.556 -6202.4
## - MiscFeature 4 0.30167 17.655 -6200.2
## - YearBuilt 1 0.23045 17.584 -6200.1
## - Exterior1st 7 0.37611 17.729 -6200.0
## - Condition1 3 0.28589 17.639 -6199.5
## - LotArea 1 0.24372 17.597 -6199.0
## - Functional 1 0.26220 17.616 -6197.4
## - BsmtFullBath 1 0.26429 17.618 -6197.2
## - ScreenPorch 1 0.36357 17.717 -6189.0
## - SaleCondition 5 0.54247 17.896 -6182.4
## - Condition2 3 0.59495 17.948 -6174.1
## - '1stFlrSF' 1 0.97822 18.332 -6139.2
## - OverallCond 1 1.11787 18.471 -6128.2
## - Neighborhood 15 1.64841 19.002 -6114.8
## - MSZoning 3 1.39125 18.745 -6110.7
## - '2ndFlrSF' 1 1.42710 18.780 -6103.9
## - PoolArea 1 1.97149 19.325 -6062.2
## - OverallQual 1 2.02622 19.380 -6058.1
## - PoolQC 3 2.48948 19.843 -6027.6
##
## Step: AIC=-6218.81
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
## LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
## BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
## YearRemodAdd + Exterior1st + ExterCond + Foundation + BsmtQual +
## BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF +
## Heating + HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' +
## LowQualFinSF + BsmtFullBath + BsmtHalfBath + FullBath + HalfBath +
## KitchenAbvGr + KitchenQual + TotRmsAbvGrd + Functional +
## Fireplaces + GarageFinish + GarageCars + GarageArea + GarageQual +
## GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
## PoolArea + PoolQC + MiscFeature + YrSold + SaleType + SaleCondition
##
## Df Sum of Sq RSS AIC

```

## - BsmtHalfBath	1	0.00815	17.367	-6220.1
## - LandContour	1	0.01171	17.371	-6219.8
## - YrSold	1	0.01640	17.376	-6219.4
## - GarageFinish	2	0.04720	17.407	-6218.8
## <none>			17.359	-6218.8
## - Street	1	0.02451	17.384	-6218.8
## - LandSlope	2	0.05694	17.416	-6218.0
## - '3SsnPorch'	1	0.03671	17.396	-6217.7
## - BsmtQual	3	0.08684	17.446	-6217.5
## - GarageQual	2	0.06484	17.424	-6217.4
## - MSSubClass	3	0.09001	17.449	-6217.3
## - GarageCond	2	0.06941	17.429	-6217.0
## - KitchenAbvGr	1	0.05210	17.411	-6216.4
## - LotConfig	1	0.05747	17.417	-6216.0
## - BsmtFinType1	5	0.15459	17.514	-6215.9
## - GarageArea	1	0.05901	17.418	-6215.9
## - Heating	2	0.08650	17.446	-6215.6
## - EnclosedPorch	1	0.07058	17.430	-6214.9
## - BsmtFinSF2	1	0.08340	17.443	-6213.8
## - TotRmsAbvGrd	1	0.09487	17.454	-6212.9
## - YearRemodAdd	1	0.09999	17.459	-6212.4
## - CentralAir	1	0.10395	17.463	-6212.1
## - HeatingQC	4	0.17660	17.536	-6212.0
## - FullBath	1	0.10611	17.465	-6211.9
## - GarageCars	1	0.10642	17.466	-6211.9
## - HouseStyle	3	0.16078	17.520	-6211.4
## - HalfBath	1	0.11507	17.474	-6211.2
## - LowQualFinSF	1	0.11782	17.477	-6210.9
## - Foundation	5	0.21521	17.575	-6210.8
## - Fireplaces	1	0.13703	17.496	-6209.3
## - KitchenQual	3	0.19298	17.552	-6208.7
## - BldgType	3	0.19461	17.554	-6208.5
## - WoodDeckSF	1	0.14943	17.509	-6208.3
## - BsmtUnfSF	1	0.15009	17.509	-6208.2
## - SaleType	3	0.20539	17.565	-6207.6
## - BsmtExposure	4	0.23163	17.591	-6207.5
## - ExterCond	3	0.22862	17.588	-6205.7
## - BsmtFinSF1	1	0.20424	17.564	-6203.7
## - YearBuilt	1	0.22941	17.589	-6201.6
## - MiscFeature	4	0.30280	17.662	-6201.6
## - Exterior1st	7	0.38327	17.743	-6200.9
## - Condition1	3	0.29132	17.651	-6200.5
## - LotArea	1	0.24394	17.603	-6200.4
## - Functional	1	0.26051	17.620	-6199.1
## - BsmtFullBath	1	0.26231	17.622	-6198.9
## - ScreenPorch	1	0.36168	17.721	-6190.7
## - SaleCondition	5	0.54441	17.904	-6183.7
## - Condition2	3	0.59032	17.950	-6176.0
## - '1stFlrSF'	1	0.97675	18.336	-6140.9
## - OverallCond	1	1.11686	18.476	-6129.8
## - Neighborhood	15	1.66117	19.020	-6115.4
## - MSZoning	3	1.38661	18.746	-6112.6
## - '2ndFlrSF'	1	1.43115	18.790	-6105.2
## - PoolArea	1	1.96648	19.326	-6064.1



```

## - OverallQual      1    2.03946 19.399 -6058.6
## - PoolQC           3    2.48538 19.845 -6029.5
##
## Step:  AIC=-6220.13
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LandContour +
##   LotConfig + LandSlope + Neighborhood + Condition1 + Condition2 +
##   BldgType + HouseStyle + OverallQual + OverallCond + YearBuilt +
##   YearRemodAdd + Exterior1st + ExterCond + Foundation + BsmtQual +
##   BsmtExposure + BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF +
##   Heating + HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' +
##   LowQualFinSF + BsmtFullBath + FullBath + HalfBath + KitchenAbvGr +
##   KitchenQual + TotRmsAbvGrd + Functional + Fireplaces + GarageFinish +
##   GarageCars + GarageArea + GarageQual + GarageCond + WoodDeckSF +
##   EnclosedPorch + '3SsnPorch' + ScreenPorch + PoolArea + PoolQC +
##   MiscFeature + YrSold + SaleType + SaleCondition
##
##              Df Sum of Sq   RSS     AIC
## - LandContour    1    0.01235 17.380 -6221.1
## - YrSold          1    0.01718 17.385 -6220.7
## - GarageFinish    2    0.04588 17.413 -6220.3
## <none>                                17.367 -6220.1
## - Street          1    0.02545 17.393 -6220.0
## - LandSlope        2    0.05665 17.424 -6219.4
## - '3SsnPorch'      1    0.03851 17.406 -6218.9
## - GarageQual       2    0.06283 17.430 -6218.9
## - BsmtQual         3    0.08917 17.457 -6218.7
## - MSSubClass       3    0.09172 17.459 -6218.4
## - GarageCond       2    0.06972 17.437 -6218.3
## - KitchenAbvGr     1    0.05310 17.421 -6217.7
## - LotConfig        1    0.05746 17.425 -6217.3
## - GarageArea       1    0.05895 17.426 -6217.2
## - BsmtFinType1     5    0.15994 17.527 -6216.7
## - Heating          2    0.08857 17.456 -6216.7
## - EnclosedPorch    1    0.07146 17.439 -6216.1
## - BsmtFinSF2       1    0.08866 17.456 -6214.7
## - TotRmsAbvGrd     1    0.09616 17.464 -6214.1
## - YearRemodAdd     1    0.10163 17.469 -6213.6
## - FullBath         1    0.10300 17.470 -6213.5
## - CentralAir       1    0.10456 17.472 -6213.4
## - HeatingQC        4    0.17744 17.545 -6213.3
## - GarageCars       1    0.10689 17.474 -6213.2
## - HalfBath         1    0.11410 17.482 -6212.6
## - HouseStyle       3    0.16448 17.532 -6212.4
## - Foundation       5    0.21418 17.582 -6212.2
## - LowQualFinSF     1    0.11837 17.486 -6212.2
## - Fireplaces       1    0.13752 17.505 -6210.6
## - KitchenQual      3    0.19242 17.560 -6210.0
## - BldgType         3    0.19678 17.564 -6209.7
## - BsmtUnfSF        1    0.14993 17.517 -6209.6
## - WoodDeckSF       1    0.15271 17.520 -6209.3
## - SaleType         3    0.20654 17.574 -6208.9
## - BsmtExposure     4    0.23704 17.605 -6208.3
## - ExterCond        3    0.22986 17.597 -6206.9
## - BsmtFinSF1       1    0.21321 17.581 -6204.3

```

```

## - YearBuilt      1  0.23019 17.598 -6202.9
## - MiscFeature    4  0.30494 17.672 -6202.7
## - Exterior1st    7  0.38015 17.748 -6202.5
## - Condition1     3  0.29151 17.659 -6201.8
## - LotArea        1  0.24473 17.612 -6201.7
## - Functional     1  0.25763 17.625 -6200.6
## - BsmtFullBath   1  0.26071 17.628 -6200.4
## - ScreenPorch    1  0.36266 17.730 -6192.0
## - SaleCondition   5  0.54280 17.910 -6185.2
## - Condition2     3  0.59308 17.960 -6177.1
## - '1stFlrSF'     1  0.97239 18.340 -6142.6
## - OverallCond    1  1.13081 18.498 -6130.0
## - Neighborhood   15 1.65455 19.022 -6117.3
## - MSZoning        3  1.38334 18.751 -6114.2
## - '2ndFlrSF'     1  1.43243 18.800 -6106.4
## - PoolArea       1  1.96832 19.336 -6065.4
## - OverallQual    1  2.03437 19.402 -6060.4
## - PoolQC         3  2.49450 19.862 -6030.2
##
## Step:  AIC=-6221.09
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotConfig +
##   LandSlope + Neighborhood + Condition1 + Condition2 + BldgType +
##   HouseStyle + OverallQual + OverallCond + YearBuilt + YearRemodAdd +
##   Exterior1st + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +
##   BsmtFullBath + FullBath + HalfBath + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + WoodDeckSF + EnclosedPorch +
##   '3SsnPorch' + ScreenPorch + PoolArea + PoolQC + MiscFeature +
##   YrSold + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS     AIC
## - YrSold      1  0.01725 17.397 -6221.6
## - LandSlope    2  0.04553 17.425 -6221.3
## - GarageFinish  2  0.04713 17.427 -6221.1
## <none>                17.380 -6221.1
## - Street       1  0.02433 17.404 -6221.0
## - '3SsnPorch'  1  0.03653 17.416 -6220.0
## - BsmtQual     3  0.08788 17.468 -6219.7
## - GarageQual   2  0.06542 17.445 -6219.6
## - GarageCond   2  0.06940 17.449 -6219.3
## - MSSubClass   3  0.09465 17.474 -6219.2
## - KitchenAbvGr 1  0.05233 17.432 -6218.7
## - LotConfig    1  0.05811 17.438 -6218.2
## - GarageArea   1  0.05829 17.438 -6218.2
## - BsmtFinType1 5  0.16018 17.540 -6217.7
## - Heating      2  0.08949 17.469 -6217.6
## - EnclosedPorch 1  0.07252 17.452 -6217.0
## - BsmtFinSF2   1  0.09102 17.471 -6215.5
## - TotRmsAbvGrd 1  0.09523 17.475 -6215.1
## - YearRemodAdd 1  0.10169 17.482 -6214.6
## - GarageCars   1  0.10487 17.485 -6214.3
## - HeatingQC    4  0.17828 17.558 -6214.2

```

```

## - CentralAir      1  0.10863 17.488 -6214.0
## - FullBath        1  0.10924 17.489 -6213.9
## - HouseStyle      3  0.16152 17.541 -6213.6
## - Foundation      5  0.21225 17.592 -6213.4
## - HalfBath        1  0.11736 17.497 -6213.3
## - LowQualFinSF    1  0.11767 17.497 -6213.2
## - Fireplaces      1  0.13931 17.519 -6211.4
## - KitchenQual     3  0.19202 17.572 -6211.0
## - BldgType        3  0.19220 17.572 -6211.0
## - BsmtUnfSF       1  0.15112 17.531 -6210.5
## - WoodDeckSF      1  0.15226 17.532 -6210.4
## - SaleType        3  0.20508 17.585 -6210.0
## - BsmtExposure    4  0.23170 17.611 -6209.8
## - ExterCond       3  0.22981 17.610 -6207.9
## - BsmtFinSF1      1  0.21549 17.595 -6205.1
## - YearBuilt       1  0.23024 17.610 -6203.9
## - Exterior1st     7  0.37563 17.755 -6203.9
## - MiscFeature     4  0.30599 17.686 -6203.6
## - LotArea         1  0.23645 17.616 -6203.4
## - Condition1      3  0.28911 17.669 -6203.0
## - Functional      1  0.25756 17.637 -6201.6
## - BsmtFullBath    1  0.26218 17.642 -6201.2
## - ScreenPorch     1  0.36629 17.746 -6192.6
## - SaleCondition   5  0.54813 17.928 -6185.8
## - Condition2      3  0.60207 17.982 -6177.4
## - '1stFlrSF'      1  0.97396 18.354 -6143.5
## - OverallCond     1  1.13080 18.511 -6131.1
## - Neighborhood    15  1.64247 19.022 -6119.3
## - MSZoning        3  1.37614 18.756 -6115.8
## - '2ndFlrSF'      1  1.42928 18.809 -6107.7
## - PoolArea        1  2.00300 19.383 -6063.8
## - OverallQual     1  2.02302 19.403 -6062.3
## - PoolQC          3  2.53995 19.920 -6027.9
##
## Step: AIC=-6221.64
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotConfig +
##   LandSlope + Neighborhood + Condition1 + Condition2 + BldgType +
##   HouseStyle + OverallQual + OverallCond + YearBuilt + YearRemodAdd +
##   Exterior1st + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +
##   BsmtFullBath + FullBath + HalfBath + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageFinish + GarageCars +
##   GarageArea + GarageQual + GarageCond + WoodDeckSF + EnclosedPorch +
##   '3SsnPorch' + ScreenPorch + PoolArea + PoolQC + MiscFeature +
##   SaleType + SaleCondition
##
##           Df Sum of Sq  RSS    AIC
## - GarageFinish  2  0.04509 17.442 -6221.9
## - LandSlope     2  0.04615 17.443 -6221.8
## <none>                17.397 -6221.6
## - Street        1  0.02556 17.423 -6221.5
## - '3SsnPorch'   1  0.03584 17.433 -6220.6
## - BsmtQual      3  0.08566 17.483 -6220.5

```

```

## - GarageQual      2    0.06663 17.464 -6220.1
## - GarageCond      2    0.06789 17.465 -6220.0
## - MSSubClass      3    0.09399 17.491 -6219.8
## - KitchenAbvGr    1    0.05680 17.454 -6218.9
## - LotConfig       1    0.05766 17.455 -6218.8
## - GarageArea      1    0.05887 17.456 -6218.7
## - BsmtFinType1    5    0.15761 17.555 -6218.5
## - Heating         2    0.09393 17.491 -6217.8
## - EnclosedPorch   1    0.07274 17.470 -6217.6
## - BsmtFinSF2      1    0.08846 17.485 -6216.2
## - TotRmsAbvGrd    1    0.09578 17.493 -6215.6
## - YearRemodAdd    1    0.09669 17.494 -6215.6
## - GarageCars      1    0.10575 17.503 -6214.8
## - HeatingQC       4    0.17939 17.576 -6214.7
## - FullBath        1    0.10798 17.505 -6214.6
## - CentralAir      1    0.11021 17.507 -6214.4
## - HouseStyle      3    0.16208 17.559 -6214.1
## - Foundation      5    0.21142 17.608 -6214.0
## - HalfBath        1    0.11576 17.513 -6214.0
## - LowQualFinSF    1    0.12200 17.519 -6213.4
## - Fireplaces      1    0.13971 17.537 -6212.0
## - KitchenQual     3    0.18946 17.587 -6211.8
## - BldgType        3    0.19574 17.593 -6211.3
## - BsmtUnfSF       1    0.14813 17.545 -6211.3
## - WoodDeckSF      1    0.15071 17.548 -6211.0
## - SaleType        3    0.20512 17.602 -6210.5
## - BsmtExposure     4    0.23956 17.637 -6209.7
## - ExterCond       3    0.22898 17.626 -6208.6
## - BsmtFinSF1      1    0.21434 17.611 -6205.8
## - Exterior1st     7    0.37163 17.769 -6204.8
## - YearBuilt       1    0.23637 17.633 -6203.9
## - MiscFeature     4    0.31096 17.708 -6203.8
## - LotArea         1    0.23899 17.636 -6203.7
## - Condition1      3    0.29323 17.690 -6203.2
## - BsmtFullBath    1    0.25673 17.654 -6202.3
## - Functional      1    0.26441 17.661 -6201.6
## - ScreenPorch     1    0.36221 17.759 -6193.6
## - SaleCondition   5    0.53932 17.936 -6187.1
## - Condition2      3    0.59997 17.997 -6178.1
## - '1stFlrSF'      1    0.98778 18.385 -6143.0
## - OverallCond     1    1.12551 18.523 -6132.1
## - Neighborhood    15    1.65696 19.054 -6118.8
## - MSZoning        3    1.37396 18.771 -6116.7
## - '2ndFlrSF'      1    1.43458 18.832 -6108.0
## - PoolArea        1    2.02080 19.418 -6063.2
## - OverallQual     1    2.02191 19.419 -6063.1
## - PoolQC         3    2.56348 19.960 -6027.0
##
## Step:  AIC=-6221.86
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotConfig +
##   LandSlope + Neighborhood + Condition1 + Condition2 + BldgType +
##   HouseStyle + OverallQual + OverallCond + YearBuilt + YearRemodAdd +
##   Exterior1st + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +

```

```

## HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +
## BsmtFullBath + FullBath + HalfBath + KitchenAbvGr + KitchenQual +
## TotRmsAbvGrd + Functional + Fireplaces + GarageCars + GarageArea +
## GarageQual + GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' +
## ScreenPorch + PoolArea + PoolQC + MiscFeature + SaleType +
## SaleCondition
##
##           Df Sum of Sq    RSS      AIC
## - LandSlope      2    0.04572 17.488 -6222.0
## <none>                17.442 -6221.9
## - Street          1    0.02449 17.467 -6221.8
## - '3SsnPorch'     1    0.03750 17.480 -6220.7
## - BsmtQual        3    0.08912 17.531 -6220.4
## - GarageCond      2    0.06601 17.508 -6220.3
## - GarageQual      2    0.06983 17.512 -6220.0
## - MSSubClass      3    0.09982 17.542 -6219.5
## - LotConfig       1    0.05592 17.498 -6219.2
## - KitchenAbvGr    1    0.05598 17.498 -6219.2
## - BsmtFinType1     5    0.15421 17.596 -6219.0
## - GarageArea       1    0.06007 17.502 -6218.8
## - Heating          2    0.09242 17.535 -6218.1
## - EnclosedPorch   1    0.07421 17.516 -6217.7
## - BsmtFinSF2       1    0.08596 17.528 -6216.7
## - TotRmsAbvGrd    1    0.09447 17.537 -6216.0
## - YearRemodAdd     1    0.10380 17.546 -6215.2
## - GarageCars       1    0.10683 17.549 -6214.9
## - CentralAir       1    0.10986 17.552 -6214.7
## - HouseStyle       3    0.15864 17.601 -6214.6
## - LowQualFinSF     1    0.12047 17.563 -6213.8
## - Foundation       5    0.22209 17.664 -6213.4
## - FullBath         1    0.12563 17.568 -6213.4
## - HeatingQC        4    0.19802 17.640 -6213.4
## - HalfBath         1    0.12665 17.569 -6213.3
## - BsmtUnfSF        1    0.14156 17.584 -6212.1
## - KitchenQual      3    0.19668 17.639 -6211.5
## - Fireplaces       1    0.14879 17.591 -6211.5
## - SaleType         3    0.20775 17.650 -6210.6
## - WoodDeckSF       1    0.16171 17.604 -6210.4
## - BldgType         3    0.21382 17.656 -6210.1
## - BsmtExposure     4    0.24492 17.687 -6209.5
## - ExterCond        3    0.22563 17.668 -6209.1
## - BsmtFinSF1       1    0.21283 17.655 -6206.2
## - MiscFeature      4    0.30597 17.748 -6204.5
## - Exterior1st      7    0.38742 17.829 -6203.8
## - LotArea          1    0.24396 17.686 -6203.6
## - Condition1       3    0.29720 17.739 -6203.2
## - BsmtFullBath     1    0.26442 17.706 -6201.9
## - YearBuilt        1    0.26570 17.708 -6201.8
## - Functional       1    0.27994 17.722 -6200.6
## - ScreenPorch      1    0.36605 17.808 -6193.5
## - SaleCondition    5    0.54263 17.985 -6187.1
## - Condition2       3    0.60726 18.049 -6177.9
## - '1stFlrSF'       1    1.00977 18.452 -6141.7
## - OverallCond      1    1.11082 18.553 -6133.7

```

```

## - Neighborhood 15 1.63933 19.081 -6120.7
## - MSZoning 3 1.35364 18.796 -6118.7
## - '2ndFlrSF' 1 1.41776 18.860 -6109.8
## - PoolArea 1 2.01390 19.456 -6064.3
## - OverallQual 1 2.07107 19.513 -6060.0
## - PoolQC 3 2.55733 19.999 -6028.1
##
## Step: AIC=-6222.04
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + Street + LotConfig +
## Neighborhood + Condition1 + Condition2 + BldgType + HouseStyle +
## OverallQual + OverallCond + YearBuilt + YearRemodAdd + Exterior1st +
## ExterCond + Foundation + BsmtQual + BsmtExposure + BsmtFinType1 +
## BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC +
## CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath +
## FullBath + HalfBath + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
## Functional + Fireplaces + GarageCars + GarageArea + GarageQual +
## GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
## PoolArea + PoolQC + MiscFeature + SaleType + SaleCondition
##
## Df Sum of Sq RSS AIC
## - Street 1 0.01877 17.507 -6222.5
## <none> 17.488 -6222.0
## - '3SsnPorch' 1 0.03451 17.522 -6221.2
## - GarageCond 2 0.06358 17.551 -6220.7
## - BsmtQual 3 0.09024 17.578 -6220.5
## - GarageQual 2 0.07220 17.560 -6220.0
## - LotConfig 1 0.05070 17.538 -6219.8
## - MSSubClass 3 0.10448 17.592 -6219.3
## - BsmtFinType1 5 0.15493 17.643 -6219.2
## - GarageArea 1 0.06100 17.549 -6219.0
## - KitchenAbvGr 1 0.06104 17.549 -6219.0
## - Heating 2 0.09255 17.580 -6218.3
## - EnclosedPorch 1 0.07507 17.563 -6217.8
## - BsmtFinSF2 1 0.08016 17.568 -6217.4
## - TotRmsAbvGrd 1 0.09418 17.582 -6216.2
## - YearRemodAdd 1 0.10060 17.588 -6215.7
## - GarageCars 1 0.10794 17.596 -6215.1
## - HouseStyle 3 0.15674 17.645 -6215.0
## - CentralAir 1 0.11259 17.600 -6214.7
## - Foundation 5 0.22144 17.709 -6213.7
## - LowQualFinSF 1 0.12601 17.614 -6213.6
## - HeatingQC 4 0.19962 17.688 -6213.5
## - HalfBath 1 0.13136 17.619 -6213.1
## - BsmtUnfSF 1 0.13572 17.624 -6212.8
## - FullBath 1 0.14126 17.629 -6212.3
## - Fireplaces 1 0.14227 17.630 -6212.2
## - KitchenQual 3 0.19476 17.683 -6211.9
## - SaleType 3 0.21877 17.707 -6209.9
## - WoodDeckSF 1 0.17152 17.659 -6209.8
## - BldgType 3 0.22503 17.713 -6209.4
## - ExterCond 3 0.23164 17.720 -6208.8
## - BsmtFinSF1 1 0.21118 17.699 -6206.5
## - BsmtExposure 4 0.28939 17.777 -6206.1
## - MiscFeature 4 0.30817 17.796 -6204.5

```

```

## - Condition1      3    0.28614 17.774 -6204.3
## - Exterior1st     7    0.38781 17.876 -6204.0
## - BsmtFullBath    1    0.25864 17.747 -6202.6
## - YearBuilt       1    0.25901 17.747 -6202.6
## - LotArea         1    0.27009 17.758 -6201.7
## - Functional      1    0.27987 17.768 -6200.9
## - ScreenPorch     1    0.37863 17.866 -6192.8
## - SaleCondition    5    0.56079 18.049 -6186.0
## - Condition2      3    0.61261 18.100 -6177.8
## - '1stFlrSF'      1    1.03251 18.520 -6140.3
## - OverallCond     1    1.10554 18.593 -6134.5
## - MSZoning         3    1.32539 18.813 -6121.4
## - Neighborhood    15    1.71115 19.199 -6115.7
## - '2ndFlrSF'      1    1.43041 18.918 -6109.3
## - PoolArea        1    2.00401 19.492 -6065.6
## - OverallQual      1    2.06236 19.550 -6061.3
## - PoolQC          3    2.54701 20.035 -6029.5
##
## Step:  AIC=-6222.48
## log(SalePrice) ~ MSSubClass + MSZoning + LotArea + LotConfig +
##   Neighborhood + Condition1 + Condition2 + BldgType + HouseStyle +
##   OverallQual + OverallCond + YearBuilt + YearRemodAdd + Exterior1st +
##   ExterCond + Foundation + BsmtQual + BsmtExposure + BsmtFinType1 +
##   BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating + HeatingQC +
##   CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF + BsmtFullBath +
##   FullBath + HalfBath + KitchenAbvGr + KitchenQual + TotRmsAbvGrd +
##   Functional + Fireplaces + GarageCars + GarageArea + GarageQual +
##   GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' + ScreenPorch +
##   PoolArea + PoolQC + MiscFeature + SaleType + SaleCondition
##
##           Df Sum of Sq   RSS   AIC
## <none>                17.507 -6222.5
## - '3SsnPorch'      1    0.03427 17.541 -6221.6
## - GarageCond       2    0.06310 17.570 -6221.2
## - BsmtQual         3    0.09331 17.600 -6220.7
## - GarageQual       2    0.07329 17.580 -6220.4
## - LotConfig        1    0.04985 17.556 -6220.3
## - KitchenAbvGr     1    0.05206 17.559 -6220.1
## - MSSubClass       3    0.10267 17.609 -6219.9
## - GarageArea       1    0.05667 17.563 -6219.8
## - BsmtFinType1     5    0.15931 17.666 -6219.2
## - Heating          2    0.09623 17.603 -6218.5
## - EnclosedPorch    1    0.07763 17.584 -6218.0
## - BsmtFinSF2       1    0.07959 17.586 -6217.9
## - TotRmsAbvGrd     1    0.09521 17.602 -6216.6
## - YearRemodAdd     1    0.10324 17.610 -6215.9
## - HouseStyle       3    0.15535 17.662 -6215.6
## - GarageCars       1    0.10759 17.614 -6215.5
## - CentralAir       1    0.11650 17.623 -6214.8
## - HeatingQC        4    0.19785 17.704 -6214.1
## - Foundation       5    0.22284 17.729 -6214.0
## - LowQualFinSF     1    0.12971 17.636 -6213.7
## - HalfBath         1    0.13022 17.637 -6213.7
## - BsmtUnfSF        1    0.13772 17.644 -6213.0

```

```
## - FullBath      1  0.14009 17.647 -6212.8
## - Fireplaces   1  0.14075 17.647 -6212.8
## - KitchenQual  3  0.19423 17.701 -6212.4
## - SaleType     3  0.21422 17.721 -6210.7
## - BldgType     3  0.21540 17.722 -6210.6
## - WoodDeckSF   1  0.17060 17.677 -6210.3
## - ExterCond    3  0.23104 17.738 -6209.3
## - BsmtFinSF1   1  0.21516 17.722 -6206.6
## - BsmtExposure 4  0.28929 17.796 -6206.5
## - Condition1   3  0.28884 17.796 -6204.6
## - MiscFeature   4  0.31580 17.822 -6204.4
## - Exterior1st  7  0.39142 17.898 -6204.2
## - LotArea      1  0.25418 17.761 -6203.4
## - BsmtFullBath 1  0.25671 17.763 -6203.2
## - YearBuilt    1  0.25868 17.765 -6203.1
## - Functional   1  0.27783 17.784 -6201.5
## - ScreenPorch  1  0.37214 17.879 -6193.8
## - SaleCondition 5  0.55932 18.066 -6186.6
## - Condition2   3  0.61245 18.119 -6178.3
## - '1stFlrSF'   1  1.03669 18.543 -6140.5
## - OverallCond  1  1.11553 18.622 -6134.3
## - Neighborhood 15  1.71222 19.219 -6116.2
## - MSZoning      3  1.47117 18.978 -6110.7
## - '2ndFlrSF'   1  1.43336 18.940 -6109.6
## - PoolArea     1  2.01467 19.521 -6065.4
## - OverallQual   1  2.06631 19.573 -6061.6
## - PoolQC       3  2.55320 20.060 -6029.7
```

```
##
```

```
## Call:
```

```
## lm(formula = log(SalePrice) ~ MSSubClass + MSZoning + LotArea +
##   LotConfig + Neighborhood + Condition1 + Condition2 + BldgType +
##   HouseStyle + OverallQual + OverallCond + YearBuilt + YearRemodAdd +
##   Exterior1st + ExterCond + Foundation + BsmtQual + BsmtExposure +
##   BsmtFinType1 + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF + Heating +
##   HeatingQC + CentralAir + '1stFlrSF' + '2ndFlrSF' + LowQualFinSF +
##   BsmtFullBath + FullBath + HalfBath + KitchenAbvGr + KitchenQual +
##   TotRmsAbvGrd + Functional + Fireplaces + GarageCars + GarageArea +
##   GarageQual + GarageCond + WoodDeckSF + EnclosedPorch + '3SsnPorch' +
##   ScreenPorch + PoolArea + PoolQC + MiscFeature + SaleType +
##   SaleCondition, data = train)
```

```
##
```

```
## Coefficients:
```

```
## (Intercept)
## 1.847e+00
## MSSubClass1-story single-family other
## -4.130e-02
## MSSubClassmulti-level single-family non PUD
## -2.555e-03
## MSSubClassother
## -3.762e-02
## MSZoningR0
## -7.012e-02
## MSZoningRL
```



##	-9.780e-03
##	MSZoningother
##	-4.276e-01
##	LotArea
##	1.613e-06
##	LotConfigInside
##	-1.384e-02
##	NeighborhoodBrkSide
##	1.854e-02
##	NeighborhoodCollgCr
##	-4.025e-02
##	NeighborhoodCrawfor
##	1.025e-01
##	NeighborhoodEdwards
##	-1.017e-01
##	NeighborhoodGilbert
##	-4.125e-02
##	NeighborhoodMitchel
##	-4.772e-02
##	NeighborhoodNames
##	-5.237e-02
##	NeighborhoodNoRidge
##	2.342e-02
##	NeighborhoodNridgHt
##	4.775e-02
##	NeighborhoodNWames
##	-4.811e-02
##	NeighborhoodOldTown
##	-4.161e-02
##	NeighborhoodSawyer
##	-4.788e-02
##	NeighborhoodSawyerW
##	-3.313e-02
##	NeighborhoodSomerst
##	2.247e-02
##	NeighborhoodTimber
##	-3.722e-02
##	Condition1Norm
##	4.809e-02
##	Condition1Pos
##	1.194e-02
##	Condition1RR
##	-9.549e-05
##	Condition2Norm
##	-1.637e-02
##	Condition2Pos
##	-5.349e-01
##	Condition2RR
##	-7.769e-02
##	BldgTypeMultiFam
##	1.154e-02
##	BldgTypeTwnhs
##	-1.056e-01
##	BldgTypeTwnhsE

##	-3.145e-02
##	HouseStyle1.5Unf
##	3.214e-02
##	HouseStyleEqMore2story
##	-4.893e-02
##	HouseStyle2.5Unf
##	3.450e-02
##	OverallQual
##	5.616e-02
##	OverallCond
##	3.710e-02
##	YearBuilt
##	1.448e-03
##	YearRemodAdd
##	7.267e-04
##	Exterior1stBrkFace
##	8.290e-02
##	Exterior1stCemntBd
##	-1.167e-02
##	Exterior1stHdBoard
##	-1.221e-03
##	Exterior1stMetalSd
##	2.603e-02
##	Exterior1stPlywood
##	1.516e-02
##	Exterior1stVinylSd
##	1.910e-02
##	Exterior1stWd Sdng
##	-4.362e-03
##	ExterCondFa
##	-2.045e-01
##	ExterCondGd
##	-1.841e-01
##	ExterCondTA
##	-1.551e-01
##	FoundationCBlock
##	2.262e-02
##	FoundationPConc
##	4.611e-02
##	FoundationSlab
##	2.562e-02
##	FoundationStone
##	1.332e-01
##	FoundationWood
##	-8.077e-02
##	BsmtQualFa
##	-3.080e-02
##	BsmtQualGd
##	-3.814e-02
##	BsmtQualTA
##	-4.593e-02
##	BsmtQualnone
##	-5.443e-02
##	BsmtExposureGd

##	4.968e-02
##	BsmtExposureMn
##	2.434e-03
##	BsmtExposureNo
##	-1.162e-02
##	BsmtExposurenone
##	-3.674e-02
##	BsmtFinType1BLQ
##	2.893e-03
##	BsmtFinType1GLQ
##	3.645e-03
##	BsmtFinType1LwQ
##	-2.072e-02
##	BsmtFinType1Rec
##	-5.323e-03
##	BsmtFinType1Unf
##	-3.582e-02
##	BsmtFinType1none
##	NA
##	BsmtFinSF1
##	9.017e-05
##	BsmtFinSF2
##	7.106e-05
##	BsmtUnfSF
##	6.603e-05
##	HeatingGasW
##	5.207e-02
##	HeatingOther
##	-6.978e-02
##	HeatingQCFa
##	-2.675e-02
##	HeatingQCGd
##	-1.748e-02
##	HeatingQCPo
##	-2.259e-01
##	HeatingQCTA
##	-3.425e-02
##	CentralAirY
##	5.184e-02
##	'1stFlrSF'
##	2.155e-04
##	'2ndFlrSF'
##	2.245e-04
##	LowQualFinSF
##	2.181e-04
##	BsmtFullBath
##	3.918e-02
##	FullBath
##	3.301e-02
##	HalfBath
##	3.030e-02
##	KitchenAbvGr
##	-4.961e-02
##	KitchenQualFa

##	-6.981e-02
##	KitchenQualGd
##	-5.989e-02
##	KitchenQualTA
##	-6.279e-02
##	TotRmsAbvGrd
##	1.097e-02
##	FunctionalTRUE
##	6.328e-02
##	Fireplaces
##	2.080e-02
##	GarageCars
##	3.076e-02
##	GarageArea
##	7.402e-05
##	GarageQualPo
##	-9.262e-02
##	GarageQualnone
##	-5.077e-02
##	GarageQualTA
##	-6.999e-02
##	GarageCondPo
##	3.383e-02
##	GarageCondnone
##	NA
##	GarageCondTA
##	6.793e-02
##	WoodDeckSF
##	1.003e-04
##	EnclosedPorch
##	1.424e-04
##	'3SsnPorch'
##	1.736e-04
##	ScreenPorch
##	3.149e-04
##	PoolArea
##	9.115e-03
##	PoolQCFa
##	-1.020e+00
##	PoolQCGd
##	-1.166e+00
##	PoolQCnone
##	4.820e+00
##	MiscFeatureOthr
##	-6.757e-02
##	MiscFeatureShed
##	-4.940e-03
##	MiscFeatureTenC
##	9.328e-01
##	MiscFeaturenone
##	1.268e-02
##	SaleTypeOthr
##	5.989e-02
##	SaleTypeNew

```
##              1.596e-01
##              SaleTypeWD
##              -1.772e-02
##              SaleConditionAdjLand
##              1.331e-01
##              SaleConditionAlloca
##              -6.110e-03
##              SaleConditionFamily
##              1.939e-02
##              SaleConditionNormal
##              7.522e-02
##              SaleConditionPartial
##              -5.614e-02
```

```
fit3 <- lm(formula = log(SalePrice) ~ MSSubClass + MSZoning + LotArea + LotConfig + Neighborhood + Cond
```

## outputing model

```
predicted.SalePrice3 = exp(predict(fit3, newdata=test))
```

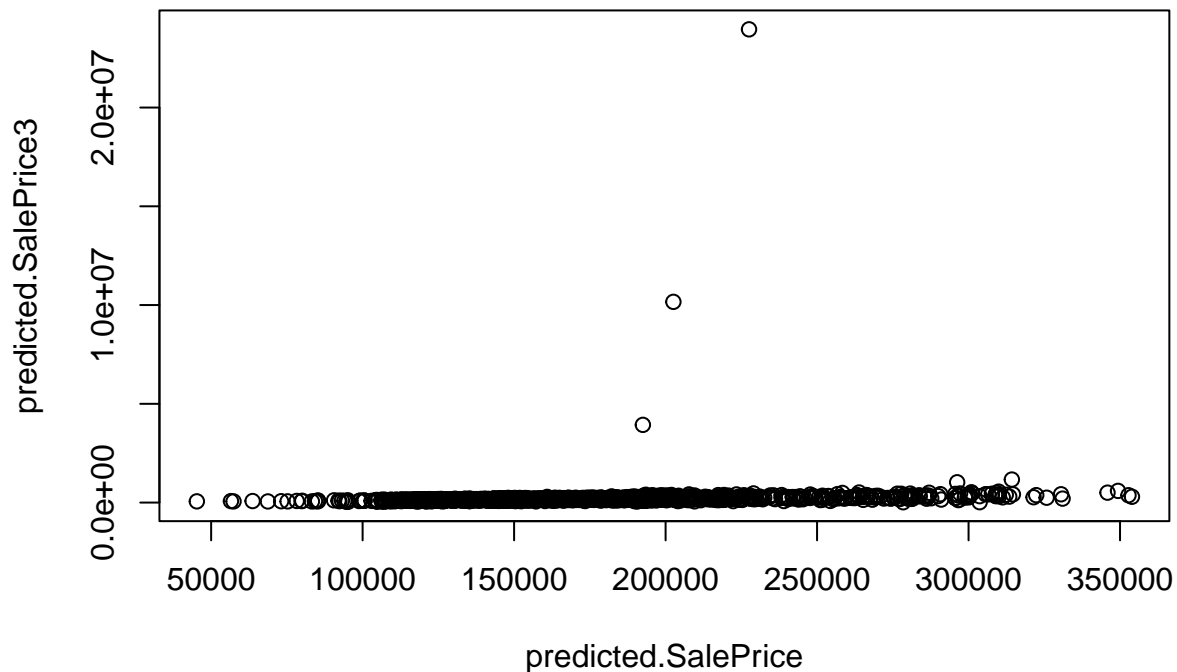
```
## Warning in predict.lm(fit3, newdata = test): prediction from a rank-deficient
## fit may be misleading
```

```
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice3)
write.csv(file='C:/Users/rodri/Desktop/Submission3.csv', SubmitDF, row.names = FALSE)
```

Score: 0.26621 before adjusting of outliers

## adjust for ouliers

```
plot(predicted.SalePrice3 ~ predicted.SalePrice)
```



```
SubmitDF$SalePrice[SubmitDF$Id %in% c(2600, 2504, 2421, 2550, 2711)] = max(SalePrice)
```

```
write.csv(file='C:/Users/rodri/Desktop/Submission3.csv', SubmitDF, row.names = FALSE)
```

Score: 0.19617 after adjusting of outliers.

**View F-test results for full model to get an idea of the most influential predictors**

```
summary(fit2)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ ., data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.88948 -0.05598  0.00292  0.05677  0.80277
##
## Coefficients: (7 not defined because of singularities)
##                                     Estimate Std. Error t value
```

## (Intercept)	7.239e+00	5.129e+00	1.411
## MSSubClass1-story single-family other	-4.503e-02	1.608e-02	-2.801
## MSSubClassmulti-level single-family non PUD	-1.347e-04	1.420e-02	-0.009
## MSSubClassother	-2.349e-02	6.219e-02	-0.378
## MSZoningR0	-8.193e-02	3.497e-02	-2.343
## MSZoningRL	-2.288e-02	3.294e-02	-0.695
## MSZoningother	-4.486e-01	5.443e-02	-8.243
## LotArea	2.108e-06	4.642e-07	4.542
## StreetPave	7.610e-02	5.776e-02	1.317
## AlleyPave	1.348e-02	2.927e-02	0.461
## Alleynone	-1.537e-02	2.031e-02	-0.757
## LotShapeReg	2.077e-03	7.550e-03	0.275
## LandContourLvl	1.564e-02	1.360e-02	1.150
## LotConfigInside	-1.556e-02	7.411e-03	-2.100
## LandSlopeMod	3.274e-02	1.871e-02	1.750
## LandSlopeSev	-3.549e-02	4.840e-02	-0.733
## NeighborhoodBrkSide	4.943e-03	2.189e-02	0.226
## NeighborhoodCollgCr	-4.281e-02	1.769e-02	-2.420
## NeighborhoodCrawfor	9.693e-02	2.159e-02	4.490
## NeighborhoodEdwards	-1.056e-01	1.750e-02	-6.033
## NeighborhoodGilbert	-4.705e-02	2.142e-02	-2.196
## NeighborhoodMitchel	-5.706e-02	2.205e-02	-2.588
## NeighborhoodNames	-5.658e-02	1.627e-02	-3.477
## NeighborhoodNoRidge	1.559e-02	2.581e-02	0.604
## NeighborhoodNridgHt	4.221e-02	2.171e-02	1.944
## NeighborhoodNWAmes	-5.794e-02	2.057e-02	-2.817
## NeighborhoodOldTown	-5.491e-02	2.035e-02	-2.698
## NeighborhoodSawyer	-5.393e-02	1.991e-02	-2.708
## NeighborhoodSawyerW	-3.402e-02	2.061e-02	-1.651
## NeighborhoodSomerst	2.341e-03	3.185e-02	0.074
## NeighborhoodTimber	-3.195e-02	2.406e-02	-1.328
## Condition1Norm	4.861e-02	1.227e-02	3.963
## Condition1Pos	9.915e-03	2.748e-02	0.361
## Condition1RR	4.183e-03	2.339e-02	0.179
## Condition2Norm	-2.198e-02	4.766e-02	-0.461
## Condition2Pos	-5.408e-01	9.171e-02	-5.897
## Condition2RR	-5.480e-02	8.399e-02	-0.652
## BldgTypeMultiFam	3.011e-03	6.464e-02	0.047
## BldgTypeTwnhs	-1.205e-01	6.583e-02	-1.831
## BldgTypeTwnhsE	-5.470e-02	6.323e-02	-0.865
## HouseStyle1.5Unf	3.820e-02	3.546e-02	1.077
## HouseStyleEqMore2story	-5.700e-02	1.922e-02	-2.966
## HouseStyle2.5Unf	3.337e-02	4.419e-02	0.755
## OverallQual	5.259e-02	4.756e-03	11.056
## OverallCond	3.665e-02	4.200e-03	8.727
## YearBuilt	1.267e-03	3.510e-04	3.611
## YearRemodAdd	8.258e-04	2.704e-04	3.053
## RoofStyleGable	2.241e-02	3.966e-02	0.565
## RoofStyleHip	2.406e-02	4.048e-02	0.594
## Exterior1stBrkFace	7.841e-02	2.342e-02	3.348
## Exterior1stCemntBd	-1.604e-02	2.350e-02	-0.683
## Exterior1stHdBoard	-3.172e-03	1.856e-02	-0.171
## Exterior1stMetalSd	2.445e-02	1.730e-02	1.413
## Exterior1stPlywood	1.964e-02	2.091e-02	0.939

## Exterior1stVinylSd	1.749e-02	1.765e-02	0.991
## Exterior1stWd Sdng	-4.572e-03	1.686e-02	-0.271
## Exterior2ndTRUE	-2.393e-03	1.190e-02	-0.201
## MasVnrTypeBrkFace	3.771e-02	3.301e-02	1.143
## MasVnrTypeNone	3.500e-02	3.330e-02	1.051
## MasVnrTypeStone	5.517e-02	3.517e-02	1.568
## MasVnrTypemissing	2.185e-03	5.363e-02	0.041
## MasVnrArea	5.050e-06	2.790e-05	0.181
## ExterQualFa	7.865e-03	4.969e-02	0.158
## ExterQualGd	-7.873e-03	2.360e-02	-0.334
## ExterQualTA	-1.848e-02	2.610e-02	-0.708
## ExterCondFa	-2.159e-01	6.893e-02	-3.132
## ExterCondGd	-1.894e-01	6.447e-02	-2.938
## ExterCondTA	-1.621e-01	6.383e-02	-2.540
## FoundationCBlock	1.951e-02	1.544e-02	1.263
## FoundationPConc	4.400e-02	1.674e-02	2.629
## FoundationSlab	2.064e-02	4.428e-02	0.466
## FoundationStone	1.170e-01	5.254e-02	2.227
## FoundationWood	-7.759e-02	7.298e-02	-1.063
## BsmtQualFa	-1.469e-02	3.066e-02	-0.479
## BsmtQualGd	-2.970e-02	1.636e-02	-1.815
## BsmtQualTA	-3.004e-02	2.001e-02	-1.501
## BsmtQualnone	6.166e-02	1.816e-01	0.340
## BsmtCondGd	3.638e-02	2.576e-02	1.412
## BsmtCondPo	4.950e-02	1.106e-01	0.448
## BsmtCondTA	3.022e-02	2.050e-02	1.474
## BsmtCondnone	NA	NA	NA
## BsmtExposureGd	4.720e-02	1.465e-02	3.221
## BsmtExposureMn	5.301e-03	1.462e-02	0.363
## BsmtExposureNo	-9.468e-03	1.043e-02	-0.907
## BsmtExposurenone	-4.851e-02	1.163e-01	-0.417
## BsmtFinType1BLQ	7.575e-04	1.340e-02	0.057
## BsmtFinType1GLQ	4.957e-03	1.217e-02	0.407
## BsmtFinType1LwQ	-2.449e-02	1.798e-02	-1.362
## BsmtFinType1Rec	-1.621e-03	1.455e-02	-0.111
## BsmtFinType1Unf	-3.555e-02	1.394e-02	-2.551
## BsmtFinType1none	NA	NA	NA
## BsmtFinSF1	8.304e-05	2.414e-05	3.441
## BsmtFinType2BLQ	-5.443e-02	3.698e-02	-1.472
## BsmtFinType2GLQ	-3.601e-04	4.349e-02	-0.008
## BsmtFinType2LwQ	-1.877e-02	3.618e-02	-0.519
## BsmtFinType2Rec	-3.600e-02	3.442e-02	-1.046
## BsmtFinType2Unf	-8.706e-03	3.688e-02	-0.236
## BsmtFinType2none	-7.660e-02	1.252e-01	-0.612
## BsmtFinSF2	9.445e-05	4.332e-05	2.180
## BsmtUnfSF	6.298e-05	2.224e-05	2.831
## TotalBsmtSF	NA	NA	NA
## HeatingGasW	5.391e-02	3.211e-02	1.679
## HeatingOther	-7.529e-02	3.910e-02	-1.926
## HeatingQCFA	-2.735e-02	2.260e-02	-1.210
## HeatingQCGd	-1.603e-02	1.012e-02	-1.584
## HeatingQCPo	-2.469e-01	1.265e-01	-1.951
## HeatingQCTA	-2.723e-02	1.001e-02	-2.719
## CentralAirY	5.375e-02	1.793e-02	2.998



## ElectricalOther	-3.973e-02	1.005e-01	-0.395
## ElectricalSBrkr	-1.828e-02	1.325e-02	-1.380
## '1stFlrSF'	2.149e-04	2.572e-05	8.356
## '2ndFlrSF'	2.368e-04	2.373e-05	9.980
## LowQualFinSF	2.110e-04	7.169e-05	2.944
## GrLivArea	NA	NA	NA
## BsmtFullBath	4.166e-02	9.611e-03	4.335
## BsmtHalfBath	1.100e-02	1.446e-02	0.761
## FullBath	3.115e-02	1.059e-02	2.940
## HalfBath	2.776e-02	9.945e-03	2.792
## BedroomAbvGr	-3.397e-03	6.486e-03	-0.524
## KitchenAbvGr	-5.074e-02	2.619e-02	-1.937
## KitchenQualFa	-6.684e-02	2.950e-02	-2.266
## KitchenQualGd	-6.204e-02	1.708e-02	-3.633
## KitchenQualTA	-5.785e-02	1.920e-02	-3.012
## TotRmsAbvGrd	1.090e-02	4.611e-03	2.365
## FunctionalTRUE	5.743e-02	1.426e-02	4.028
## Fireplaces	1.481e-03	1.252e-02	0.118
## FireplaceQuFa	-3.239e-02	3.398e-02	-0.953
## FireplaceQuGd	-2.097e-02	2.634e-02	-0.796
## FireplaceQunone	-4.845e-02	3.070e-02	-1.578
## FireplaceQuPo	6.684e-03	3.877e-02	0.172
## FireplaceQuTA	-2.187e-02	2.735e-02	-0.800
## GarageTypeBuiltIn	-2.171e-02	1.582e-02	-1.373
## GarageTypeDetchd	-6.658e-03	1.085e-02	-0.614
## GarageTypenone	-5.489e-02	4.732e-02	-1.160
## GarageTypeOther	-2.852e-02	2.340e-02	-1.219
## GarageFinishnone	NA	NA	NA
## GarageFinishRFRn	-5.491e-03	9.518e-03	-0.577
## GarageFinishUnf	-1.722e-02	1.171e-02	-1.470
## GarageCars	2.911e-02	1.097e-02	2.654
## GarageArea	7.626e-05	3.666e-05	2.080
## GarageQualPo	-8.147e-02	4.048e-02	-2.013
## GarageQualnone	NA	NA	NA
## GarageQualTA	-5.889e-02	3.409e-02	-1.727
## GarageCondPo	3.523e-02	4.838e-02	0.728
## GarageCondnone	NA	NA	NA
## GarageCondTA	6.991e-02	4.205e-02	1.662
## PavedDriveP	8.802e-03	2.684e-02	0.328
## PavedDriveY	1.987e-02	1.625e-02	1.223
## WoodDeckSF	9.634e-05	2.866e-05	3.361
## OpenPorchSF	1.049e-05	5.555e-05	0.189
## EnclosedPorch	1.310e-04	6.039e-05	2.169
## '3SsnPorch'	1.798e-04	1.095e-04	1.642
## ScreenPorch	3.058e-04	6.007e-05	5.091
## PoolArea	9.017e-03	7.707e-04	11.700
## PoolQCFa	-9.962e-01	1.727e-01	-5.767
## PoolQCGd	-1.162e+00	1.257e-01	-9.242
## PoolQCnone	4.775e+00	4.277e-01	11.164
## FenceGdWo	-3.600e-02	2.364e-02	-1.523
## FenceMnPrv	-5.225e-03	1.899e-02	-0.275
## FenceMnWw	-3.510e-02	4.025e-02	-0.872
## Fencenone	3.616e-04	1.718e-02	0.021
## MiscFeatureOthr	-5.298e-02	2.557e-01	-0.207

## MiscFeatureShed	2.912e-02	2.565e-01	0.114
## MiscFeatureTenC	9.537e-01	2.988e-01	3.192
## MiscFeaturenene	4.240e-02	2.685e-01	0.158
## MiscVal	2.903e-06	1.985e-05	0.146
## MoSold	3.950e-04	1.196e-03	0.330
## YrSold	-2.619e-03	2.504e-03	-1.046
## SaleTypeOther	6.262e-02	3.101e-02	2.020
## SaleTypeNew	1.117e-01	7.449e-02	1.499
## SaleTypeWD	-1.667e-02	2.042e-02	-0.816
## SaleConditionAdjLand	1.291e-01	6.507e-02	1.984
## SaleConditionAlloca	1.362e-02	4.164e-02	0.327
## SaleConditionFamily	1.999e-02	2.996e-02	0.667
## SaleConditionNormal	7.465e-02	1.378e-02	5.416
## SaleConditionPartial	-1.189e-02	7.165e-02	-0.166
##	Pr(> t )		
## (Intercept)	0.158381		
## MSSubClass1-story single-family other	0.005174	**	
## MSSubClassmulti-level single-family non PUD	0.992436		
## MSSubClassother	0.705674		
## MSZoningR0	0.019297	*	
## MSZoningRL	0.487352		
## MSZoningother	4.09e-16	***	
## LotArea	6.09e-06	***	
## StreetPave	0.187915		
## AlleyPave	0.645141		
## Alleynone	0.449134		
## LotShapeReg	0.783249		
## LandContourLvl	0.250272		
## LotConfigInside	0.035961	*	
## LandSlopeMod	0.080440	.	
## LandSlopeSev	0.463497		
## NeighborhoodBrkSide	0.821396		
## NeighborhoodCollgCr	0.015654	*	
## NeighborhoodCrawfor	7.76e-06	***	
## NeighborhoodEdwards	2.09e-09	***	
## NeighborhoodGilbert	0.028248	*	
## NeighborhoodMitchel	0.009749	**	
## NeighborhoodNames	0.000524	***	
## NeighborhoodNoRidge	0.545875		
## NeighborhoodNridgHt	0.052056	.	
## NeighborhoodNWames	0.004921	**	
## NeighborhoodOldTown	0.007066	**	
## NeighborhoodSawyer	0.006858	**	
## NeighborhoodSawyerW	0.099054	.	
## NeighborhoodSomerst	0.941401		
## NeighborhoodTimber	0.184398		
## Condition1Norm	7.81e-05	***	
## Condition1Pos	0.718332		
## Condition1RR	0.858078		
## Condition2Norm	0.644770		
## Condition2Pos	4.71e-09	***	
## Condition2RR	0.514246		
## BldgTypeMultiFam	0.962852		
## BldgTypeTwnhs	0.067328	.	

## BldgTypeTwnhsE	0.387165
## HouseStyle1.5Unf	0.281534
## HouseStyleEqMore2story	0.003074 **
## HouseStyle2.5Unf	0.450381
## OverallQual	< 2e-16 ***
## OverallCond	< 2e-16 ***
## YearBuilt	0.000317 ***
## YearRemodAdd	0.002309 **
## RoofStyleGable	0.572099
## RoofStyleHip	0.552406
## Exterior1stBrkFace	0.000836 ***
## Exterior1stCemntBd	0.495023
## Exterior1stHdBoard	0.864327
## Exterior1stMetalSd	0.157771
## Exterior1stPlywood	0.347916
## Exterior1stVinylSd	0.321827
## Exterior1stWd Sdng	0.786333
## Exterior2ndTRUE	0.840696
## MasVnrTypeBrkFace	0.253443
## MasVnrTypeNone	0.293501
## MasVnrTypeStone	0.117016
## MasVnrTypemissing	0.967506
## MasVnrArea	0.856377
## ExterQualFa	0.874262
## ExterQualGd	0.738734
## ExterQualTA	0.479006
## ExterCondFa	0.001776 **
## ExterCondGd	0.003359 **
## ExterCondTA	0.011206 *
## FoundationCBBlock	0.206728
## FoundationPConc	0.008672 **
## FoundationSlab	0.641164
## FoundationStone	0.026124 *
## FoundationWood	0.287873
## BsmtQualFa	0.631897
## BsmtQualGd	0.069691 .
## BsmtQualTA	0.133535
## BsmtQualnone	0.734228
## BsmtCondGd	0.158061
## BsmtCondPo	0.654421
## BsmtCondTA	0.140703
## BsmtCondnone	NA
## BsmtExposureGd	0.001310 **
## BsmtExposureMn	0.717001
## BsmtExposureNo	0.364333
## BsmtExposurenone	0.676737
## BsmtFinType1BLQ	0.954930
## BsmtFinType1GLQ	0.683967
## BsmtFinType1LwQ	0.173370
## BsmtFinType1Rec	0.911299
## BsmtFinType1Unf	0.010864 *
## BsmtFinType1none	NA
## BsmtFinSF1	0.000599 ***
## BsmtFinType2BLQ	0.141317

## BsmtFinType2GLQ	0.993395
## BsmtFinType2LwQ	0.604014
## BsmtFinType2Rec	0.295726
## BsmtFinType2Unf	0.813436
## BsmtFinType2none	0.540743
## BsmtFinSF2	0.029405 *
## BsmtUnfSF	0.004706 **
## TotalBsmtSF	NA
## HeatingGasW	0.093408 .
## HeatingOther	0.054381 .
## HeatingQCFa	0.226479
## HeatingQCGd	0.113425
## HeatingQCPo	0.051222 .
## HeatingQCTA	0.006637 **
## CentralAirY	0.002767 **
## ElectricalOther	0.692751
## ElectricalSBrkr	0.167804
## '1stFlrSF'	< 2e-16 ***
## '2ndFlrSF'	< 2e-16 ***
## LowQualFinSF	0.003300 **
## GrLivArea	NA
## BsmtFullBath	1.57e-05 ***
## BsmtHalfBath	0.447094
## FullBath	0.003336 **
## HalfBath	0.005323 **
## BedroomAbvGr	0.600482
## KitchenAbvGr	0.052917 .
## KitchenQualFa	0.023611 *
## KitchenQualGd	0.000291 ***
## KitchenQualTA	0.002643 **
## TotRmsAbvGrd	0.018197 *
## FunctionalTRUE	5.94e-05 ***
## Fireplaces	0.905876
## FireplaceQuFa	0.340721
## FireplaceQuGd	0.426148
## FireplaceQunone	0.114790
## FireplaceQuPo	0.863168
## FireplaceQuTA	0.424026
## GarageTypeBuiltIn	0.170108
## GarageTypeDetchd	0.539421
## GarageTypenone	0.246262
## GarageTypeOther	0.223252
## GarageFinishnone	NA
## GarageFinishRFn	0.564101
## GarageFinishUnf	0.141705
## GarageCars	0.008052 **
## GarageArea	0.037682 *
## GarageQualPo	0.044354 *
## GarageQualnone	NA
## GarageQualTA	0.084323 .
## GarageCondPo	0.466650
## GarageCondnone	NA
## GarageCondTA	0.096707 .
## PavedDriveP	0.743052

```
## PavedDriveY 0.221668
## WoodDeckSF 0.000798 ***
## OpenPorchSF 0.850303
## EnclosedPorch 0.030287 *
## '3SsnPorch' 0.100745
## ScreenPorch 4.09e-07 ***
## PoolArea < 2e-16 ***
## PoolQCFa 1.01e-08 ***
## PoolQCGd < 2e-16 ***
## PoolQCnone < 2e-16 ***
## FenceGdWo 0.128111
## FenceMnPrv 0.783271
## FenceMnWw 0.383336
## Fencenone 0.983214
## MiscFeatureOthr 0.835860
## MiscFeatureShed 0.909626
## MiscFeatureTenC 0.001448 **
## MiscFeaturenone 0.874569
## MiscVal 0.883731
## MoSold 0.741217
## YrSold 0.295807
## SaleTypeOther 0.043633 *
## SaleTypeNew 0.134081
## SaleTypeWD 0.414628
## SaleConditionAdjLand 0.047489 *
## SaleConditionAlloca 0.743555
## SaleConditionFamily 0.504833
## SaleConditionNormal 7.27e-08 ***
## SaleConditionPartial 0.868207
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1145 on 1291 degrees of freedom
## Multiple R-squared:  0.9273, Adjusted R-squared:  0.9178
## F-statistic: 98 on 168 and 1291 DF, p-value: < 2.2e-16
```

## create reduced model using influential predictors

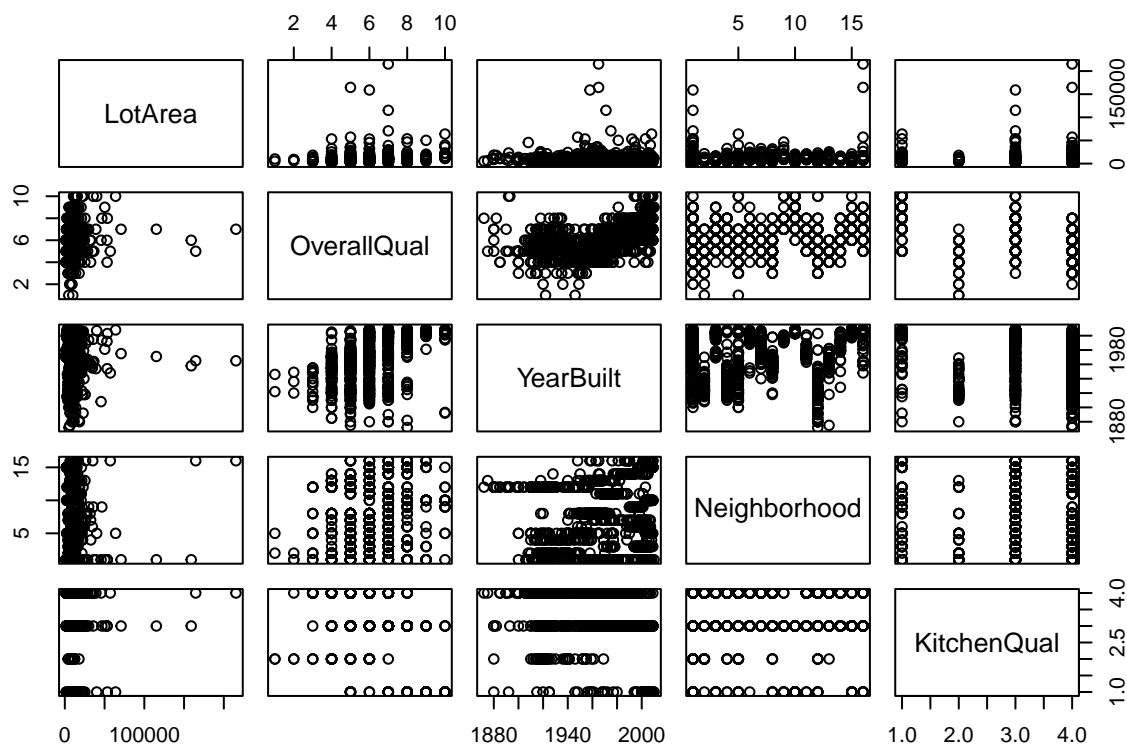
```
train.highF <- lm(log(SalePrice) ~ LotArea + OverallQual + YearBuilt + Neighborhood + KitchenQual + Lot.
summary(train.highF)
```

```
##
## Call:
## lm(formula = log(SalePrice) ~ LotArea + OverallQual + YearBuilt +
##     Neighborhood + KitchenQual + LotArea * Neighborhood, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.00301 -0.09348  0.00486  0.10165  0.63383
##
## Coefficients:
```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.066e+00  6.109e-01   9.929 < 2e-16 ***
## LotArea           8.574e-06  8.084e-07  10.606 < 2e-16 ***
## OverallQual       1.448e-01  5.721e-03  25.304 < 2e-16 ***
## YearBuilt         2.588e-03  3.135e-04   8.258 3.35e-16 ***
## NeighborhoodBrkSide -2.588e-01  7.346e-02  -3.523 0.000440 ***
## NeighborhoodCollgCr -1.107e-01  6.707e-02  -1.650 0.099201 .
## NeighborhoodCrawfor  1.810e-01  6.647e-02   2.723 0.006555 **
## NeighborhoodEdwards  1.075e-01  3.733e-02   2.881 0.004029 **
## NeighborhoodGilbert -2.579e-03  5.871e-02  -0.044 0.964969
## NeighborhoodMitchel  3.986e-02  6.088e-02   0.655 0.512757
## NeighborhoodNames   -3.723e-04  4.419e-02  -0.008 0.993280
## NeighborhoodNoRidge  2.813e-01  6.656e-02   4.226 2.53e-05 ***
## NeighborhoodNridgHt -1.412e-01  7.007e-02  -2.015 0.044055 *
## NeighborhoodNWAmes   1.282e-03  1.000e-01   0.013 0.989777
## NeighborhoodOldTown -2.609e-01  5.088e-02  -5.128 3.33e-07 ***
## NeighborhoodSawyer   -3.728e-02  7.105e-02  -0.525 0.599865
## NeighborhoodSawyerW  -2.319e-01  1.107e-01  -2.095 0.036374 *
## NeighborhoodSomerst  -1.118e-01  5.085e-02  -2.199 0.028032 *
## NeighborhoodTimber    1.527e-01  3.967e-02   3.849 0.000124 ***
## KitchenQualFa       -3.002e-01  4.103e-02  -7.317 4.22e-13 ***
## KitchenQualGd       -1.253e-01  2.328e-02  -5.383 8.54e-08 ***
## KitchenQualTA       -2.192e-01  2.602e-02  -8.423 < 2e-16 ***
## LotArea:NeighborhoodBrkSide  4.032e-05  9.137e-06   4.413 1.10e-05 ***
## LotArea:NeighborhoodCollgCr  1.353e-05  6.547e-06   2.067 0.038939 *
## LotArea:NeighborhoodCrawfor  8.259e-06  4.996e-06   1.653 0.098534 .
## LotArea:NeighborhoodEdwards -1.442e-05  2.803e-06  -5.144 3.07e-07 ***
## LotArea:NeighborhoodGilbert  3.283e-06  4.559e-06   0.720 0.471498
## LotArea:NeighborhoodMitchel -9.223e-07  4.519e-06  -0.204 0.838329
## LotArea:NeighborhoodNames    6.456e-06  3.917e-06   1.648 0.099547 .
## LotArea:NeighborhoodNoRidge  8.359e-07  4.073e-06   0.205 0.837427
## LotArea:NeighborhoodNridgHt  2.533e-05  6.030e-06   4.201 2.82e-05 ***
## LotArea:NeighborhoodNWAmes   9.650e-06  8.173e-06   1.181 0.237942
## LotArea:NeighborhoodOldTown  3.219e-05  5.569e-06   5.780 9.16e-09 ***
## LotArea:NeighborhoodSawyer   8.489e-06  6.323e-06   1.343 0.179601
## LotArea:NeighborhoodSawyerW  2.650e-05  1.081e-05   2.451 0.014373 *
## LotArea:NeighborhoodSomerst  1.763e-05  5.475e-06   3.220 0.001311 **
## LotArea:NeighborhoodTimber  -5.035e-06  1.101e-06  -4.572 5.24e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.185 on 1423 degrees of freedom
## Multiple R-squared:  0.7908, Adjusted R-squared:  0.7855
## F-statistic: 149.4 on 36 and 1423 DF, p-value: < 2.2e-16
```

View crosswise correlations to make sure no predictors are highly correlated

```
train.sub <- train %>% select(LotArea, OverallQual, YearBuilt, Neighborhood, KitchenQual)
pairs(train.sub)
```



# Look at “best” subsets of reduced model. Adjusted  $r^2$  is highest for full reduced model. AIC is lowest for full reduced model but should it be negative? # No further reductions need to be done

```
library(olsrr)
```

```
## Warning: package 'olsrr' was built under R version 4.1.3
```

```
##
```

```
## Attaching package: 'olsrr'
```

```
## The following object is masked from 'package:datasets':
```

```
##
```

```
## rivers
```

```
fit.olsrr <- ols_step_best_subset(train.highF)
```

```
fit.olsrr
```

```
## Best Subsets Regression
```

```
## -----
```

```
## Model Index Predictors
```

```
## -----
```

```
## 1 OverallQual
```

```
## 2 OverallQual LotArea:Neighborhood
```

```
## 3 OverallQual KitchenQual LotArea:Neighborhood
```

```
## 4 OverallQual Neighborhood KitchenQual LotArea:Neighborhood
```

```
##      5      OverallQual YearBuilt Neighborhood KitchenQual LotArea:Neighborhood
##      6      LotArea OverallQual YearBuilt Neighborhood KitchenQual LotArea:Neighborhood
## -----
##
##                                     Subsets Regression Summary
## -----
## Model      R-Square      Adj.      Pred      C(p)      AIC      SBIC      SBC      MS
## -----
## 1          0.6678      0.6676      0.6667      803.6582      -140.1770      -4331.8207      -124.3184      77.4
## 2          0.7408      0.7378      0.7264      308.8195      -470.6935      -4723.8605      -370.2559      60.4
## 3          0.7621      0.7588      0.7472      166.4386      -589.4545      -4851.1613      -473.1582      55.5
## 4          0.7808      0.7754      0.761      41.1939      -679.0062      -4973.5748      -483.4171      51.2
## 5          0.7908      0.7855      0.7693      -25.0000      -745.3485      -5034.6904      -544.4732      48.9
## 6          0.7908      0.7855      0.7693      -23.0000      -745.3485      -5012.4046      -544.4732      48.9
## -----
## AIC: Akaike Information Criteria
## SBIC: Sawa's Bayesian Information Criteria
## SBC: Schwarz Bayesian Criteria
## MSE: Estimated error of prediction, assuming multivariate normality
## FPE: Final Prediction Error
## HSP: Hocking's Sp
## APC: Amemiya Prediction Criteria
```

## predict test set results using trained linear model

```
predicted.SalePrice3 = exp(predict(train.highF, newdata=test))
```

## create data frame of IDs and predicted Sale Prices, write to csv

```
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice3)
write.csv(file='~/Submission_LRB.csv', SubmitDF, row.names = FALSE)
```

.19457 which is better!

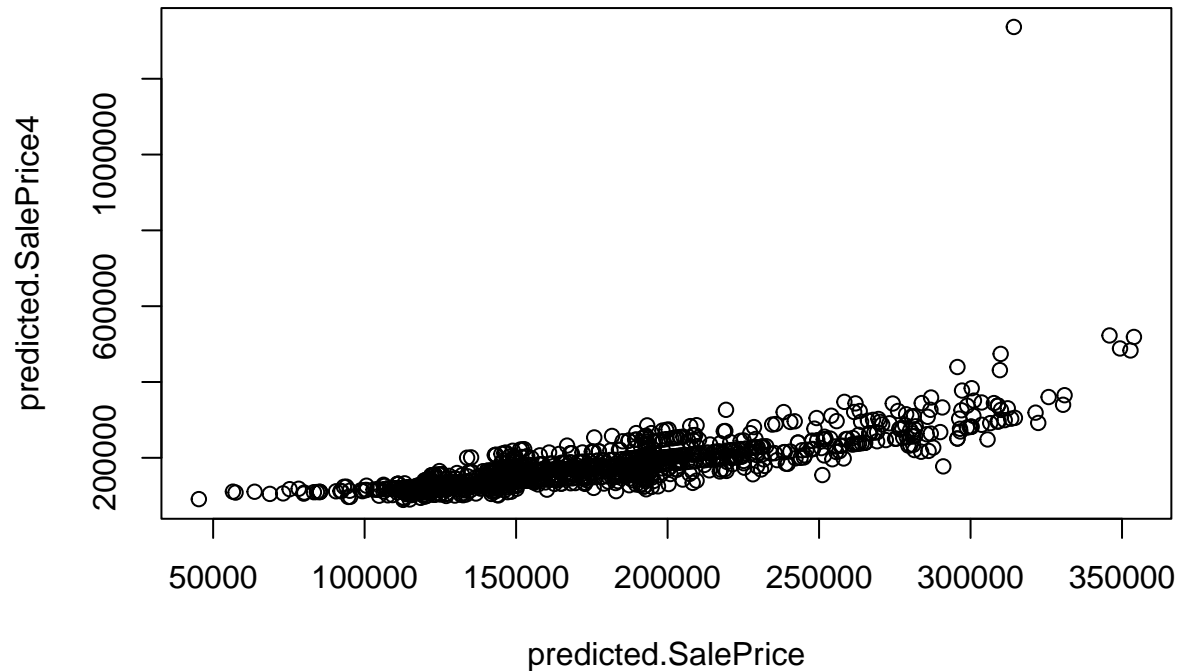
## Next going to try and choose my own predictors.

```
fit4 = lm(formula = log(SalePrice) ~ SaleType+SaleCondition+FullBath+BedroomAbvGr+GrLivArea, data = tr
```

```
predicted.SalePrice4 = exp(predict(fit4, newdata=test))
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice4)
write.csv(file='C:/Users/rodri/Desktop/Submission4.csv', SubmitDF, row.names = FALSE)
```



```
plot(predicted.SalePrice4 ~ predicted.SalePrice)
```



This should be worse than our previous model.  
0.26267 which it is.

Lets Try Splines

using the earth library to do Multivariate Adaptive Regression  
Splines(MARS)

```
library(earth)
```

```
## Warning: package 'earth' was built under R version 4.1.3
```

```
## Loading required package: Formula
```

```
## Loading required package: plotmo
```

```
## Warning: package 'plotmo' was built under R version 4.1.3
```

```
## Loading required package: plotrix
```

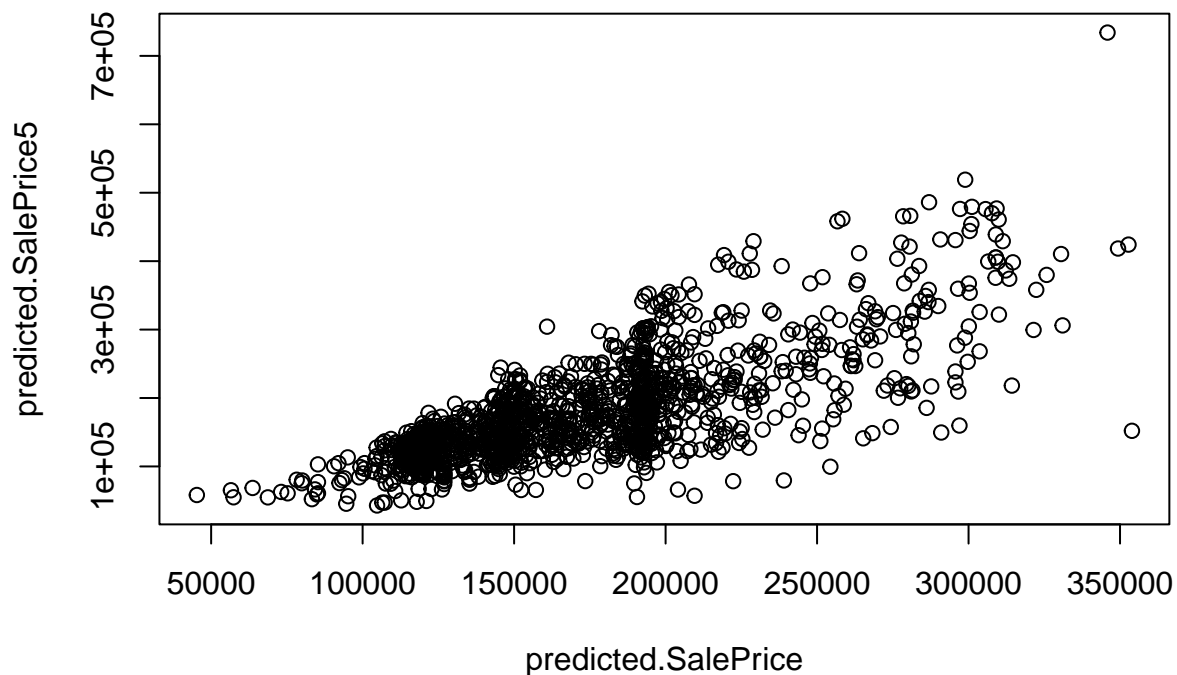
```
## Loading required package: TeachingDemos
```

```
## Warning: package 'TeachingDemos' was built under R version 4.1.3
```

```
fit5 <- earth(log(SalePrice)~.,data=train)
```

```
predicted.SalePrice5 = exp(predict(fit5, newdata=test))  
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice5)  
names(SubmitDF)[2]<-"SalePrice"  
write.csv(file='C:/Users/rodri/Desktop/Submission5.csv', SubmitDF, row.names = FALSE)
```

```
plot(predicted.SalePrice5 ~ predicted.SalePrice)
```



0.13238!!! This is even without model selection.

## Lets try with model tuning

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.1.3
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
## lift
```

making a grid search to find the best fit

```
hyper_grid <- expand.grid(
  degree = 1:3,
  nprune = seq(2, 100, length.out = 10) %>% floor()
)

fit6 <- train(
  x = train,
  y = SalePrice,
  method = "earth",
  metric = "RMSE",
  trControl = trainControl(method = "cv", number = 10),
  tuneGrid = hyper_grid
)
```

[illegible]

```
## Setting row names on a tibble is deprecated.  
## Setting row names on a tibble is deprecated.
```

## looking at the best tune combination

```
fit6$bestTune
```

```
##   nprune degree  
## 5      45      1
```

```
library(vip)
```

```
## Warning: package 'vip' was built under R version 4.1.3
```

```
##  
## Attaching package: 'vip'
```

```
## The following object is masked from 'package:utils':  
##  
##   vi
```

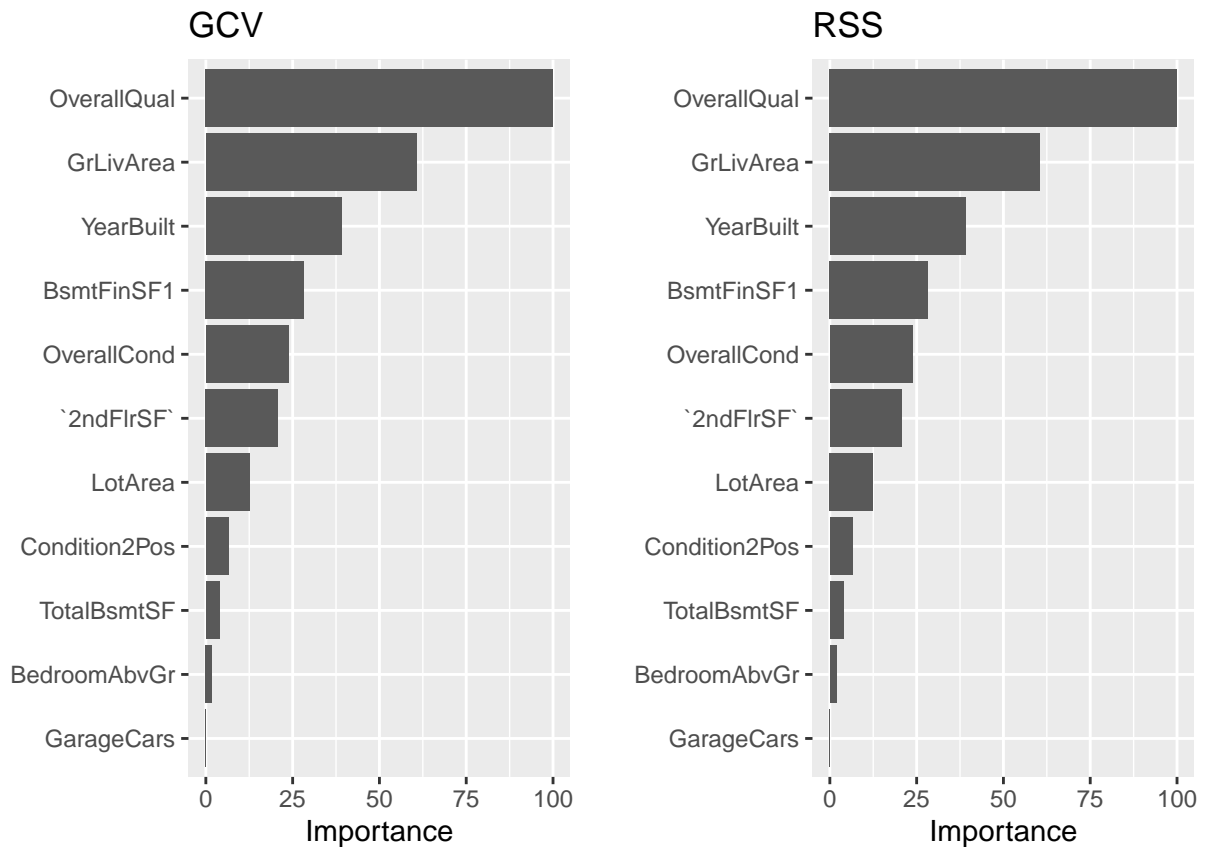
```
library(pdp)
```

```
## Warning: package 'pdp' was built under R version 4.1.3
```

```
##  
## Attaching package: 'pdp'
```

```
## The following object is masked from 'package:purrr':  
##  
##   partial
```

```
p1 <- vip(fit6, num_features = 40, bar = FALSE, value = "gcv") + ggtitle("GCV")  
p2 <- vip(fit6, num_features = 40, bar = FALSE, value = "rss") + ggtitle("RSS")  
gridExtra::grid.arrange(p1, p2, ncol = 2)
```

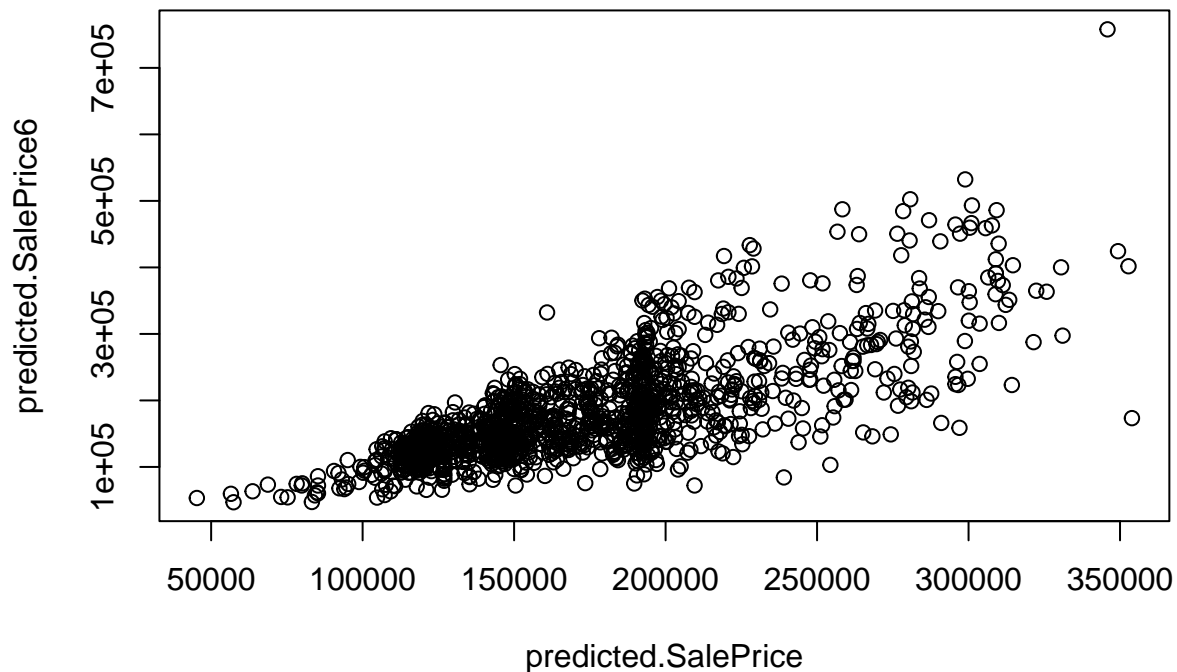


# Trying MARS with our AIC model

```
fit6 <- earth(log(SalePrice)~OverallQual+GrLivArea+YearBuilt+BsmFinSF1+OverallCond+`2ndFlrSF`+LotArea+
```

```
predicted.SalePrice6 = exp(predict(fit6, newdata=test))
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice6)
names(SubmitDF)[2]<-"SalePrice"
write.csv(file='C:/Users/rodri/Desktop/Submission6.csv', SubmitDF, row.names = FALSE)
```

```
plot(predicted.SalePrice6 ~ predicted.SalePrice)
```



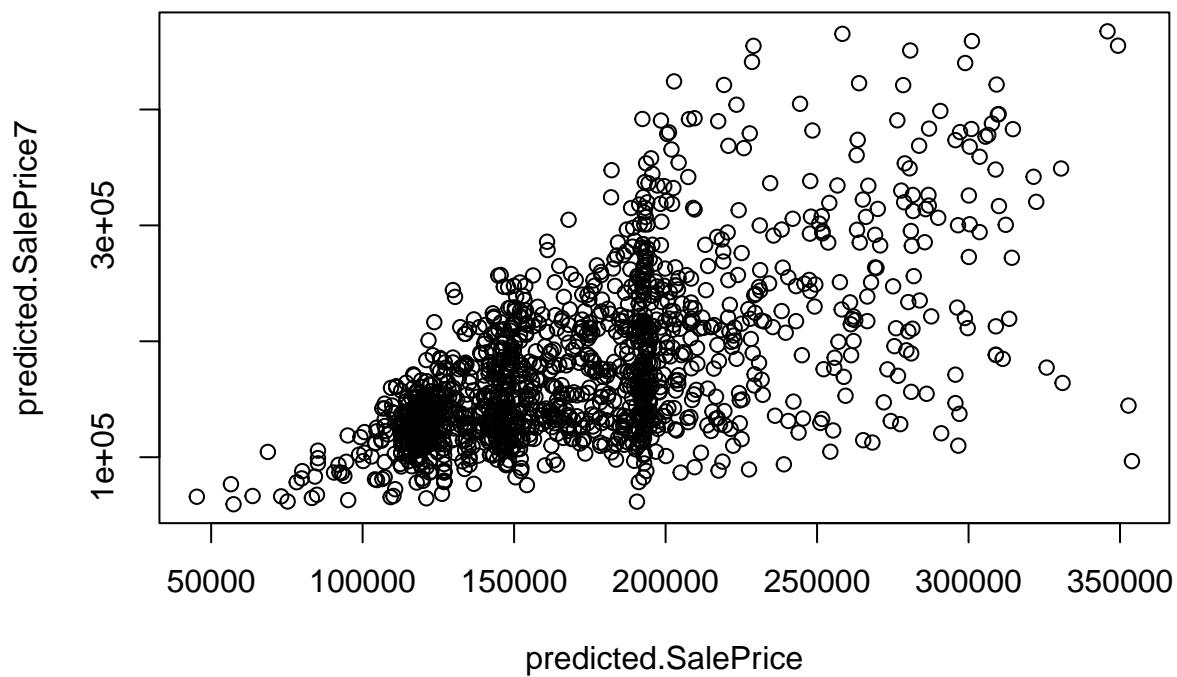
0.13586 worse than before. I think this is pretty good.

**Trying MARS with our trimmed down model which was the best performance model before MARS**

```
fit7 <- earth(log(SalePrice)~ LotArea + OverallQual + YearBuilt + Neighborhood + KitchenQual + LotArea*)
```

```
predicted.SalePrice7 = exp(predict(fit7, newdata=test))
SubmitDF = data.frame(Id=test$Id, SalePrice=predicted.SalePrice7)
names(SubmitDF)[2]<-"SalePrice"
write.csv(file='C:/Users/rodri/Desktop/Submission7.csv', SubmitDF, row.names = FALSE)
```

```
plot(predicted.SalePrice7 ~ predicted.SalePrice)
```



0.18318 interesting how splines plus best performing model turns out to be worse than using all the predictors.