136_Liz_Project_Step5_Log Transformation - 0.8

Hyunkyung Kim

November 6, 2018

```r
#install.packages("glmnet")
#install.packages("mlbench")
#install.packages("Boruta")
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

library(tidyverse)

## -- Attaching packages -------------------------------------------------
## --------------------------------------- tidyverse 1.2.1 --

## v tibble  1.4.2      v purrr   0.2.5
## v tidyr   0.8.1      v dplyr   0.7.7
## v readr   1.1.1      v stringr 1.3.1
## v tibble  1.4.2      v forcats 0.3.0

## -- Conflicts ----------------------------------------------------------
## -------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## x purrr::lift()   masks caret::lift()

library(psych)

##
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha

library(glmnet)

## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following object is masked from 'package:tidyr':
##
##     expand
```

```
## Loading required package: foreach

##
## Attaching package: 'foreach'

## The following objects are masked from 'package:purrr':
##
##     accumulate, when

## Loaded glmnet 2.0-16

library(mlbench)
library(Boruta)

## Loading required package: ranger

library(MASS) # stepwise regression

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select

library(leaps) # all subsets regression
library(randomForest)

## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:ranger':
##
##     importance

## The following object is masked from 'package:psych':
##
##     outlier

## The following object is masked from 'package:dplyr':
##
##     combine

## The following object is masked from 'package:ggplot2':
##
##     margin

library(caret)
library(tidyverse)
library(psych)
```

```
library(glmnet)
library(mlbench)
library(Boruta)
library(MASS) # stepwise regression
library(leaps) # all subsets regression
library(randomForest)
library(MLmetrics)

##
## Attaching package: 'MLmetrics'

## The following object is masked from 'package:psych':
##
##     AUC

## The following objects are masked from 'package:caret':
##
##     MAE, RMSE

## The following object is masked from 'package:base':
##
##     Recall

library(e1071)
library(car) # VIF

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## The following object is masked from 'package:dplyr':
##
##     recode

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

Import Clean Data

```
H_Clean<-read.csv( file = "C:\\Users\\Hyunkyung Kim\\Desktop\\CKME999\\136\\d
ataset\\all\\H_clean.csv")
H_Clean$MSSubClass<-as.factor(H_Clean$MSSubClass)
Train<-H_Clean[!is.na(H_Clean$SalePrice),]
Test<-H_Clean[is.na(H_Clean$SalePrice),]
actual<-read.csv( file = "C:\\Users\\Hyunkyung Kim\\Desktop\\CKME999\\136\\da
taset\\all\\AMES_test.csv")[1461:2919,2] # Test Price
```

Utilities has extremely low variance where only 1 out 2919 observation is different. This is not really useful and I will drop this.

```
H_Eng<-H_Clean[,names(H_Clean)!="Utilities"]
```

Add Full bath and half baths into one BathAbvGrd and BsmtFullBath+BsmtHalfBath into BsmtBath

```
H_Eng$BathAbvGrd<-H_Eng$FullBath+0.5*H_Eng$HalfBath
H_Eng$BsmtBath<-H_Eng$BsmtFullBath+0.5*H_Eng$BsmtHalfBath

plot( H_Eng$BathAbvGrd, H_Eng$SalePrice)
abline(lm(H_Eng$SalePrice~H_Eng$BathAbvGrd, data=H_Eng), col='red')
```
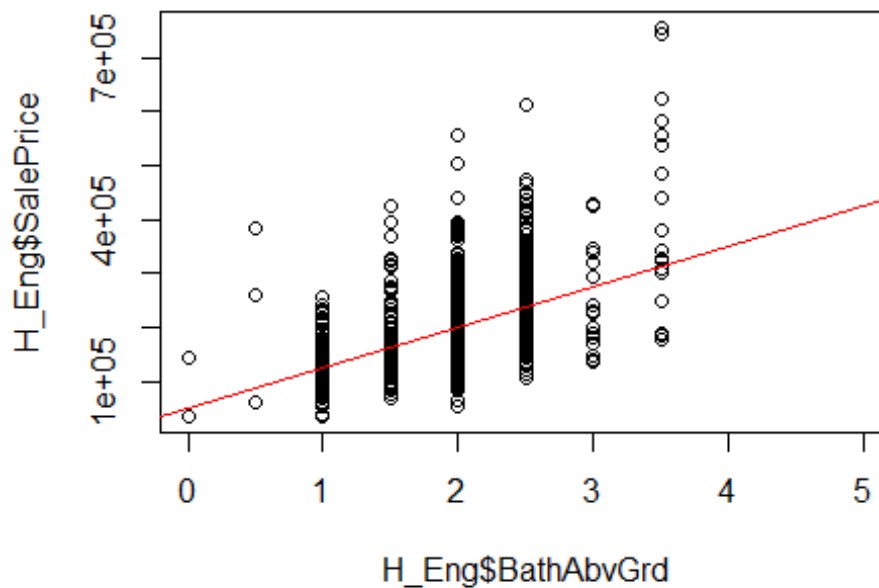


```
plot(H_Eng$BsmtBath, H_Eng$SalePrice)
abline(lm(H_Eng$SalePrice~H_Eng$BsmtBath, data=H_Eng), col='red')
```
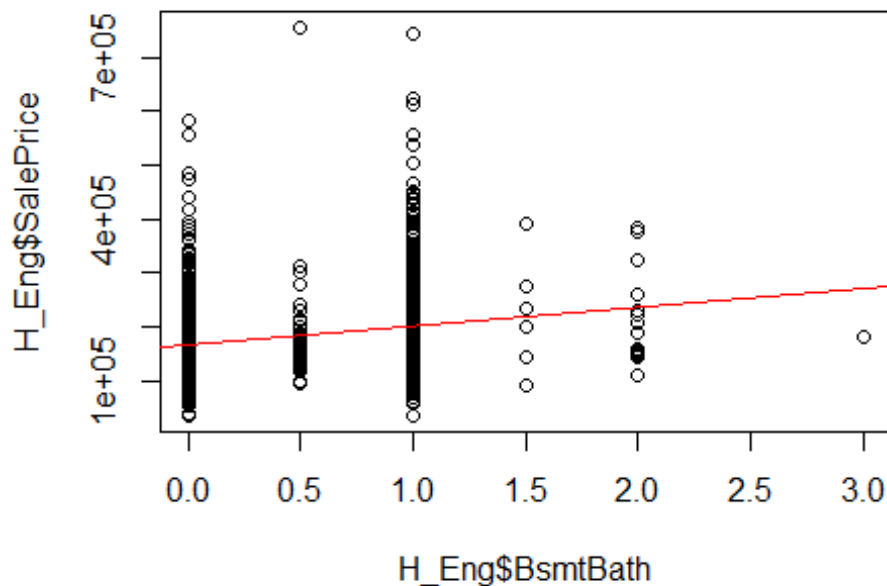
```r
# Remove original columns
#H_Eng<-H_Eng[,!names(H_Clean) %in% c("FullBath",'HalfBath','BathAbvGrd','Bsm
tBath')]
H_Eng<-subset(H_Eng, select=-c(FullBath,HalfBath,BathAbvGrd,BsmtBath))
```

Log Transformation due to Skewness

As Discussed in Step1, we have a few variables including response variable that thas high skewness.

```r
H_num<-H_Eng[,sapply(H_Eng,is.numeric)] #Numerical
H_num<-H_num[-c(1)]
H_cat<-H_Eng[,sapply(H_Eng[,-1],is.factor)] # Categorical
Check<-describe(H_num)[,c(11,12)]
Check[Check$skew>0.75 & Check$skew<1 ,]

##              skew kurtosis
## ExterQual    0.79     0.06
## BsmtUnfSF    0.92     0.40
## X2ndFlrSF    0.86    -0.43
## TotRmsAbvGrd 0.76     1.16

Check[Check$skew<(-1) ,]

##             skew kurtosis
## LandSlope  -4.97    26.51
## BldgType   -2.16     3.16
```

```
## BsmtQual   -1.25    4.04
## BsmtCond   -3.63   17.06
## Functional -4.96   26.95
## GarageQual -3.27   10.14
## GarageCond -3.39   10.62
## PavedDrive -2.98    7.10
```

Below items all high skewness and may need transformation.- Skew over 1 - Removed ordinal items.

-         `skew kurtosis`

LotFrontage 1.33 9.12 LotArea 12.82 264.31 MasVnrArea 2.61 9.31 ExterCond 1.32 6.27 TotalBsmtSF 1.16 9.10 X1stFlrSF 1.47 6.94 LowQualFinSF 12.08 174.51 GrLivArea 1.27 4.11 BsmtHalfBath 3.93 14.82 KitchenAbvGr 4.30 19.73 PavedDrive -2.98 7.10 WoodDeckSF 1.84 6.72 OpenPorchSF 2.53 10.91 EnclosedPorch 4.00 28.31 X3SsnPorch 11.37 149.05 ScreenPorch 3.94 17.73 PoolArea 16.89 297.91 PoolQC 15.86 255.66 MiscVal 21.94 562.72 SalePrice 1.88 6.50

- now add

ExterQual 0.79 0.06 ** not ordinal items. BsmtUnfSF 0.92 0.40 X2ndFlrSF 0.86 -0.43 TotRmsAbvGrd 0.76 1.16

Subset non-normal variables.
```
H_SkewedVar<-H_Eng[,c(

'LotFrontage',
'LotArea',
'MasVnrArea',
'TotalBsmtSF',
'X1stFlrSF',
'LowQualFinSF',
'GrLivArea',
'BsmtHalfBath',
'KitchenAbvGr',
'OpenPorchSF',
'EnclosedPorch',
'X3SsnPorch',
'ScreenPorch',
'PoolArea',
'MiscVal',
'SalePrice')]

H_Lo<-H_Eng
```

Apply Log only to the skewed items

```
H_Lo[,c(
```

```r
  'LotFrontage',
  'LotArea',
  'MasVnrArea',
  'BsmtFinSF1',
  'BsmtFinSF2',
  'TotalBsmtSF',
  'X1stFlrSF',
  'LowQualFinSF',
  'GrLivArea',
  'BsmtHalfBath',
  'KitchenAbvGr',
  'OpenPorchSF',
  'EnclosedPorch',
  'X3SsnPorch',
  'ScreenPorch',
  'PoolArea',
  'MiscVal',
  'SalePrice',

  'BsmtUnfSF',
  'X2ndFlrSF',
  'TotRmsAbvGrd'
)]<-log(1+H_Lo[,c(

  'LotFrontage',
  'LotArea',
  'MasVnrArea',
  'BsmtFinSF1',
  'BsmtFinSF2',
  'TotalBsmtSF',
  'X1stFlrSF',
  'LowQualFinSF',
  'GrLivArea',
  'BsmtHalfBath',
  'KitchenAbvGr',
  'OpenPorchSF',
  'EnclosedPorch',
  'X3SsnPorch',
  'ScreenPorch',
  'PoolArea',
  'MiscVal',
  'SalePrice',
  'BsmtUnfSF',
  'X2ndFlrSF',
  'TotRmsAbvGrd'
)])

head(H_Lo)
```

```
##    Id MSSubClass MSZoning LotFrontage  LotArea Street  Alley LotShape
## 1  1     2SNEW       RL    4.189655 9.042040   Pave NoAlley      Reg
## 2  2     1SNEW       RL    4.394449 9.169623   Pave NoAlley      Reg
## 3  3     2SNEW       RL    4.234107 9.328212   Pave NoAlley      IR1
## 4  4     2SOLD       RL    4.110874 9.164401   Pave NoAlley      IR1
## 5  5     2SNEW       RL    4.442651 9.565284   Pave NoAlley      IR1
## 6  6     1NHFin       RL    4.454347 9.555064   Pave NoAlley      IR1
##    LandContour LotConfig LandSlope Neighborhood Condition1 Condition2
## 1         Lvl    Inside         3      CollgCr       Norm       Norm
## 2         Lvl       FR2         3      Veenker      Feedr       Norm
## 3         Lvl    Inside         3      CollgCr       Norm       Norm
## 4         Lvl    Corner         3      Crawfor       Norm       Norm
## 5         Lvl       FR2         3      NoRidge       Norm       Norm
## 6         Lvl    Inside         3      Mitchel       Norm       Norm
##    BldgType HouseStyle OverallQual OverallCond YearBuilt YearRemodAdd
## 1        5     2Story           7           5      2003         2003
## 2        5     1Story           6           8      1976         1976
## 3        5     2Story           7           5      2001         2002
## 4        5     2Story           7           5      1915         1970
## 5        5     2Story           8           5      2000         2000
## 6        5     1.5Fin           5           5      1993         1995
##    RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType MasVnrArea
## 1     Gable  CompShg     VinylSd     VinylSd    BrkFace   5.283204
## 2     Gable  CompShg     MetalSd     MetalSd       None   0.000000
## 3     Gable  CompShg     VinylSd     VinylSd    BrkFace   5.093750
## 4     Gable  CompShg     Wd Sdng     Wd Shng       None   0.000000
## 5     Gable  CompShg     VinylSd     VinylSd    BrkFace   5.860786
## 6     Gable  CompShg     VinylSd     VinylSd       None   0.000000
##    ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure
## 1         4         3      PConc        4        3            1
## 2         3         3     CBlock        4        3            4
## 3         4         3      PConc        4        3            2
## 4         3         3     BrkTil        3        4            1
## 5         4         3      PConc        4        3            3
## 6         3         3       Wood        4        3            1
##    BsmtFinType1 BsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1            6   6.561031            1          0  5.017280    6.753438
## 2            5   6.886532            1          0  5.652489    7.141245
## 3            6   6.188264            1          0  6.075346    6.825460
## 4            5   5.379897            1          0  6.293419    6.629363
## 5            6   6.486161            1          0  6.196444    7.044033
## 6            6   6.597146            1          0  4.174387    6.680855
##    Heating HeatingQC CentralAir Electrical X1stFlrSF X2ndFlrSF LowQualFinSF
## 1    GasA         5          Y      SBrkr  6.753438  6.751101            0
## 2    GasA         5          Y      SBrkr  7.141245  0.000000            0
## 3    GasA         5          Y      SBrkr  6.825460  6.765039            0
## 4    GasA         4          Y      SBrkr  6.869014  6.629363            0
## 5    GasA         5          Y      SBrkr  7.044033  6.960348            0
## 6    GasA         5          Y      SBrkr  6.680855  6.340359            0
##    GrLivArea BsmtFullBath BsmtHalfBath BedroomAbvGr KitchenAbvGr
```

```
## 1  7.444833                 1     0.0000000            3     0.6931472
## 2  7.141245                 0     0.6931472            3     0.6931472
## 3  7.488294                 1     0.0000000            3     0.6931472
## 4  7.448916                 1     0.0000000            3     0.6931472
## 5  7.695758                 1     0.0000000            4     0.6931472
## 6  7.217443                 1     0.0000000            1     0.6931472
##   KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType
## 1           4     2.197225          7          0           0     Attchd
## 2           3     1.945910          7          1           3     Attchd
## 3           4     1.945910          7          1           3     Attchd
## 4           4     2.079442          7          1           4     Detchd
## 5           4     2.302585          7          1           3     Attchd
## 6           3     1.791759          7          0           0     Attchd
##   GarageYrBlt GarageFinish GarageCars GarageArea GarageQual GarageCond
## 1        2003            2          2        548          3          3
## 2        1976            2          2        460          3          3
## 3        2001            2          2        608          3          3
## 4        1998            1          3        642          3          3
## 5        2000            2          3        836          3          3
## 6        1993            1          2        480          3          3
##   PavedDrive WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch ScreenPorch
## 1          2          0    4.127134      0.000000   0.000000           0
## 2          2        298    0.000000      0.000000   0.000000           0
## 3          2          0    3.761200      0.000000   0.000000           0
## 4          2          0    3.583519      5.609472   0.000000           0
## 5          2        192    4.442651      0.000000   0.000000           0
## 6          2         40    3.433987      0.000000   5.771441           0
##   PoolArea PoolQC   Fence    MiscFeature  MiscVal MoSold YrSold SaleType
## 1        0      0 NoFence NoMiscFeature 0.000000      2   2008       WD
## 2        0      0 NoFence NoMiscFeature 0.000000      5   2007       WD
## 3        0      0 NoFence NoMiscFeature 0.000000      9   2008       WD
## 4        0      0 NoFence NoMiscFeature 0.000000      2   2006       WD
## 5        0      0 NoFence NoMiscFeature 0.000000     12   2008       WD
## 6        0      0   MnPrv          Shed 6.552508     10   2009       WD
##   SaleCondition SalePrice
## 1        Normal  12.24770
## 2        Normal  12.10902
## 3        Normal  12.31717
## 4       Abnorml  11.84940
## 5        Normal  12.42922
## 6        Normal  11.87061
```

head(H_Eng)

```
##   Id MSSubClass MSZoning LotFrontage LotArea Street   Alley LotShape
## 1  1      2SNEW       RL          65    8450   Pave NoAlley      Reg
## 2  2      1SNEW       RL          80    9600   Pave NoAlley      Reg
## 3  3      2SNEW       RL          68   11250   Pave NoAlley      IR1
## 4  4      2SOLD       RL          60    9550   Pave NoAlley      IR1
## 5  5      2SNEW       RL          84   14260   Pave NoAlley      IR1
```

```
## 6    6    1NHFin         RL         85    14115    Pave NoAlley         IR1
##     LandContour LotConfig LandSlope Neighborhood Condition1 Condition2
## 1          Lvl    Inside         3       CollgCr       Norm       Norm
## 2          Lvl       FR2         3       Veenker      Feedr       Norm
## 3          Lvl    Inside         3       CollgCr       Norm       Norm
## 4          Lvl    Corner         3       Crawfor       Norm       Norm
## 5          Lvl       FR2         3       NoRidge       Norm       Norm
## 6          Lvl    Inside         3       Mitchel       Norm       Norm
##     BldgType HouseStyle OverallQual OverallCond YearBuilt YearRemodAdd
## 1          5     2Story           7           5      2003         2003
## 2          5     1Story           6           8      1976         1976
## 3          5     2Story           7           5      2001         2002
## 4          5     2Story           7           5      1915         1970
## 5          5     2Story           8           5      2000         2000
## 6          5     1.5Fin           5           5      1993         1995
##     RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType MasVnrArea
## 1       Gable  CompShg      VinylSd      VinylSd    BrkFace        196
## 2       Gable  CompShg      MetalSd      MetalSd       None          0
## 3       Gable  CompShg      VinylSd      VinylSd    BrkFace        162
## 4       Gable  CompShg      Wd Sdng      Wd Shng       None          0
## 5       Gable  CompShg      VinylSd      VinylSd    BrkFace        350
## 6       Gable  CompShg      VinylSd      VinylSd       None          0
##     ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure
## 1           4         3      PConc        4        3            1
## 2           3         3     CBlock        4        3            4
## 3           4         3      PConc        4        3            2
## 4           3         3     BrkTil        3        4            1
## 5           4         3      PConc        4        3            3
## 6           3         3       Wood        4        3            1
##     BsmtFinType1 BsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1              6        706            1          0       150         856
## 2              5        978            1          0       284        1262
## 3              6        486            1          0       434         920
## 4              5        216            1          0       540         756
## 5              6        655            1          0       490        1145
## 6              6        732            1          0        64         796
##     Heating HeatingQC CentralAir Electrical X1stFlrSF X2ndFlrSF LowQualFinSF
## 1      GasA         5          Y       SBrkr       856       854            0
## 2      GasA         5          Y       SBrkr      1262         0            0
## 3      GasA         5          Y       SBrkr       920       866            0
## 4      GasA         4          Y       SBrkr       961       756            0
## 5      GasA         5          Y       SBrkr      1145      1053            0
## 6      GasA         5          Y       SBrkr       796       566            0
##     GrLivArea BsmtFullBath BsmtHalfBath BedroomAbvGr KitchenAbvGr
## 1        1710            1            0            3            1
## 2        1262            0            1            3            1
## 3        1786            1            0            3            1
## 4        1717            1            0            3            1
## 5        2198            1            0            4            1
## 6        1362            1            0            1            1
```

```
##    KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType
## 1            4            8          7          0           0     Attchd
## 2            3            6          7          1           3     Attchd
## 3            4            6          7          1           3     Attchd
## 4            4            7          7          1           4     Detchd
## 5            4            9          7          1           3     Attchd
## 6            3            5          7          0           0     Attchd
##    GarageYrBlt GarageFinish GarageCars GarageArea GarageQual GarageCond
## 1        2003            2          2        548          3          3
## 2        1976            2          2        460          3          3
## 3        2001            2          2        608          3          3
## 4        1998            1          3        642          3          3
## 5        2000            2          3        836          3          3
## 6        1993            1          2        480          3          3
##    PavedDrive WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch ScreenPorch
## 1           2          0          61             0          0           0
## 2           2        298           0             0          0           0
## 3           2          0          42             0          0           0
## 4           2          0          35           272          0           0
## 5           2        192          84             0          0           0
## 6           2         40          30             0        320           0
##    PoolArea PoolQC   Fence   MiscFeature MiscVal MoSold YrSold SaleType
## 1         0      0 NoFence NoMiscFeature       0      2   2008       WD
## 2         0      0 NoFence NoMiscFeature       0      5   2007       WD
## 3         0      0 NoFence NoMiscFeature       0      9   2008       WD
## 4         0      0 NoFence NoMiscFeature       0      2   2006       WD
## 5         0      0 NoFence NoMiscFeature       0     12   2008       WD
## 6         0      0   MnPrv         Shed     700     10   2009       WD
##    SaleCondition SalePrice
## 1        Normal    208500
## 2        Normal    181500
## 3        Normal    223500
## 4        Abnorml   140000
## 5        Normal    250000
## 6        Normal    143000
```

Feature Selection done on the engineered training set.

```
nulllog<-lm(SalePrice~1,data=H_Lo[1:1460,])
fulllog<-lm(SalePrice~.,data=H_Lo[1:1460,])

StepF_Log<-stepAIC(nulllog, scope=list(lower=nulllog, upper=fulllog), directi
on='forward', trace=F)
StepB_Log<-stepAIC(fulllog, direction='backward', trace=F)
StepS_Log<-stepAIC(fulllog,direction='both', trace=F)

summary(StepF_Log)

##
## Call:
## lm(formula = SalePrice ~ OverallQual + GrLivArea + Neighborhood +
```

```
##      MSSubClass + BsmtFinSF1 + OverallCond + GarageCars + YearBuilt +
##      RoofMatl + LotArea + MSZoning + BsmtExposure + Condition2 +
##      Functional + SaleCondition + KitchenQual + Condition1 + Exterior1st +
##      BsmtFullBath + Fireplaces + BsmtQual + HeatingQC + PoolQC +
##      ScreenPorch + Heating + LotConfig + GarageArea + Foundation +
##      WoodDeckSF + CentralAir + SaleType + X1stFlrSF + Street +
##      YearRemodAdd + MoSold + KitchenAbvGr + MasVnrType, data = H_Lo[1:1460,

##      ])
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -0.67331 -0.04917  0.00068  0.05551  0.50000
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -8.097e-01  7.983e-01  -1.014 0.310655
## OverallQual         4.659e-02  4.274e-03  10.899  < 2e-16 ***
## GrLivArea           4.392e-01  2.944e-02  14.917  < 2e-16 ***
## NeighborhoodBlueste 1.907e-02  8.603e-02   0.222 0.824572
## NeighborhoodBrDale  2.763e-02  4.965e-02   0.557 0.577914
## NeighborhoodBrkSide 4.474e-02  4.046e-02   1.106 0.269035
## NeighborhoodClearCr 1.648e-02  4.001e-02   0.412 0.680473
## NeighborhoodCollgCr -3.839e-02  3.212e-02  -1.195 0.232136
## NeighborhoodCrawfor 9.331e-02  3.713e-02   2.513 0.012088 *
## NeighborhoodEdwards -7.222e-02  3.542e-02  -2.039 0.041683 *
## NeighborhoodGilbert -3.391e-02  3.417e-02  -0.992 0.321296
## NeighborhoodIDOTRR  -2.144e-02  4.656e-02  -0.460 0.645295
## NeighborhoodMeadowV -1.054e-01  5.285e-02  -1.994 0.046365 *
## NeighborhoodMitchel -6.177e-02  3.588e-02  -1.722 0.085390 .
## NeighborhoodNAmes   -3.797e-02  3.422e-02  -1.110 0.267317
## NeighborhoodNoRidge 7.158e-02  3.605e-02   1.986 0.047255 *
## NeighborhoodNPkVill 6.739e-02  4.930e-02   1.367 0.171905
## NeighborhoodNridgHt 8.237e-02  3.240e-02   2.542 0.011134 *
## NeighborhoodNWAmes  -5.001e-02  3.531e-02  -1.416 0.156903
## NeighborhoodOldTown -2.880e-02  4.152e-02  -0.694 0.488056
## NeighborhoodSawyer  -2.287e-02  3.586e-02  -0.638 0.523797
## NeighborhoodSawyerW -2.553e-02  3.478e-02  -0.734 0.463138
## NeighborhoodSomerst -4.250e-03  4.020e-02  -0.106 0.915819
## NeighborhoodStoneBr 1.277e-01  3.636e-02   3.511 0.000461 ***
## NeighborhoodSWISU   -4.765e-03  4.143e-02  -0.115 0.908452
## NeighborhoodTimber  -2.777e-02  3.608e-02  -0.770 0.441593
## NeighborhoodVeenker 3.504e-03  4.655e-02   0.075 0.940019
## MSSubClass1NHFin    5.831e-02  2.910e-02   2.004 0.045286 *
## MSSubClass1NHUnf    1.128e-01  4.492e-02   2.511 0.012154 *
## MSSubClass1SFA      1.994e-02  6.314e-02   0.316 0.752153
## MSSubClass1SNEW     9.145e-02  2.938e-02   3.113 0.001894 **
## MSSubClass1SOLD     6.271e-02  3.563e-02   1.760 0.078599 .
## MSSubClass1SPUD     6.010e-02  2.887e-02   2.082 0.037573 *
## MSSubClass2FCONV    2.927e-02  3.801e-02   0.770 0.441454
```

```
## MSSubClass2NHS           4.378e-02  4.227e-02   1.036 0.300471
## MSSubClass2SNEW          5.799e-02  2.454e-02   2.363 0.018271 *
## MSSubClass2SOLD          6.631e-02  3.303e-02   2.007 0.044912 *
## MSSubClassDUPL           4.895e-02  3.879e-02   1.262 0.207186
## MSSubClassMLPUD          7.054e-02  4.559e-02   1.547 0.122017
## MSSubClassSPL            5.637e-02  3.063e-02   1.840 0.065918 .
## MSSubClassSPLF           8.082e-02  3.706e-02   2.181 0.029377 *
## BsmtFinSF1               9.266e-03  1.332e-03   6.959 5.36e-12 ***
## OverallCond              3.774e-02  3.560e-03  10.602  < 2e-16 ***
## GarageCars               2.515e-02  9.481e-03   2.653 0.008085 **
## YearBuilt                2.036e-03  3.299e-04   6.173 8.90e-10 ***
## RoofMatlCompShg          1.701e+00  1.242e-01  13.695  < 2e-16 ***
## RoofMatlMembran          1.818e+00  1.666e-01  10.914  < 2e-16 ***
## RoofMatlMetal            1.777e+00  1.667e-01  10.661  < 2e-16 ***
## RoofMatlRoll             1.674e+00  1.678e-01   9.974  < 2e-16 ***
## RoofMatlTar&Grv          1.666e+00  1.282e-01  12.995  < 2e-16 ***
## RoofMatlWdShake          1.721e+00  1.354e-01  12.708  < 2e-16 ***
## RoofMatlWdShngl          1.877e+00  1.318e-01  14.240  < 2e-16 ***
## LotArea                  8.194e-02  1.066e-02   7.684 2.99e-14 ***
## MSZoningFV               4.389e-01  5.244e-02   8.370  < 2e-16 ***
## MSZoningRH               3.956e-01  5.214e-02   7.587 6.13e-14 ***
## MSZoningRL               3.879e-01  4.401e-02   8.813  < 2e-16 ***
## MSZoningRM               3.598e-01  4.124e-02   8.725  < 2e-16 ***
## BsmtExposure             1.671e-02  3.550e-03   4.707 2.78e-06 ***
## Condition2Feedr          2.491e-02  9.715e-02   0.256 0.797699
## Condition2Norm           2.507e-02  8.226e-02   0.305 0.760631
## Condition2PosA           3.921e-01  1.383e-01   2.836 0.004637 **
## Condition2PosN          -6.227e-01  1.170e-01  -5.324 1.19e-07 ***
## Condition2RRAe          -2.724e-02  1.375e-01  -0.198 0.843007
## Condition2RRAn          -1.320e-02  1.358e-01  -0.097 0.922583
## Condition2RRNn           8.330e-02  1.139e-01   0.731 0.464740
## Functional               3.697e-02  4.750e-03   7.783 1.42e-14 ***
## SaleConditionAdjLand     1.300e-01  5.993e-02   2.170 0.030219 *
## SaleConditionAlloca      5.613e-02  3.653e-02   1.537 0.124581
## SaleConditionFamily      8.511e-03  2.712e-02   0.314 0.753670
## SaleConditionNormal      6.695e-02  1.256e-02   5.330 1.15e-07 ***
## SaleConditionPartial    -5.873e-02  6.579e-02  -0.893 0.372186
## KitchenQual              2.334e-02  6.917e-03   3.374 0.000762 ***
## Condition1Feedr          3.796e-02  2.171e-02   1.749 0.080518 .
## Condition1Norm           8.490e-02  1.797e-02   4.723 2.57e-06 ***
## Condition1PosA           4.559e-02  4.345e-02   1.049 0.294167
## Condition1PosN           8.633e-02  3.237e-02   2.667 0.007744 **
## Condition1RRAe          -4.985e-02  3.849e-02  -1.295 0.195502
## Condition1RRAn           5.332e-02  2.979e-02   1.790 0.073692 .
## Condition1RRNe           1.441e-02  7.893e-02   0.183 0.855129
## Condition1RRNn           1.215e-01  5.511e-02   2.205 0.027601 *
## Exterior1stAsphShn      -4.854e-02  1.124e-01  -0.432 0.666044
## Exterior1stBrkComm      -1.753e-01  8.393e-02  -2.089 0.036904 *
## Exterior1stBrkFace       9.787e-02  3.152e-02   3.105 0.001940 **
## Exterior1stCBlock       -1.103e-02  1.120e-01  -0.099 0.921546
```

```
## Exterior1stCemntBd    5.570e-02  3.256e-02    1.710 0.087414 .
## Exterior1stHdBoard    6.282e-03  2.849e-02    0.220 0.825528
## Exterior1stImStucc   -1.561e-02  1.104e-01   -0.141 0.887640
## Exterior1stMetalSd    4.073e-02  2.764e-02    1.473 0.140888
## Exterior1stPlywood    4.895e-03  2.990e-02    0.164 0.869985
## Exterior1stStone      7.646e-04  8.570e-02    0.009 0.992882
## Exterior1stStucco     4.125e-02  3.513e-02    1.174 0.240559
## Exterior1stVinylSd    2.591e-02  2.790e-02    0.929 0.353310
## Exterior1stWd Sdng    4.984e-03  2.754e-02    0.181 0.856409
## Exterior1stWdShing    1.146e-02  3.453e-02    0.332 0.739923
## BsmtFullBath          2.922e-02  7.338e-03    3.982 7.21e-05 ***
## Fireplaces            2.217e-02  5.849e-03    3.790 0.000158 ***
## BsmtQual              2.217e-02  6.444e-03    3.439 0.000601 ***
## HeatingQC             1.575e-02  4.197e-03    3.753 0.000182 ***
## PoolQC                4.825e-02  1.189e-02    4.059 5.21e-05 ***
## ScreenPorch           8.223e-03  2.101e-03    3.914 9.52e-05 ***
## HeatingGasA           6.869e-02  1.106e-01    0.621 0.534675
## HeatingGasW           1.462e-01  1.134e-01    1.289 0.197547
## HeatingGrav          -1.699e-02  1.175e-01   -0.145 0.885050
## HeatingOthW          -1.253e-02  1.365e-01   -0.092 0.926908
## HeatingWall           1.384e-01  1.262e-01    1.096 0.273064
## LotConfigCulDSac      1.710e-02  1.382e-02    1.238 0.215995
## LotConfigFR2         -3.745e-02  1.766e-02   -2.120 0.034157 *
## LotConfigFR3         -9.599e-02  5.644e-02   -1.701 0.089213 .
## LotConfigInside      -1.167e-02  7.656e-03   -1.524 0.127847
## GarageArea            9.522e-05  3.216e-05    2.961 0.003123 **
## FoundationCBlock      2.099e-02  1.368e-02    1.534 0.125311
## FoundationPConc       3.769e-02  1.498e-02    2.516 0.011971 *
## FoundationSlab        2.639e-02  3.517e-02    0.750 0.453174
## FoundationStone       1.113e-01  4.586e-02    2.426 0.015384 *
## FoundationWood       -1.032e-01  6.399e-02   -1.613 0.107055
## WoodDeckSF            6.905e-05  2.544e-05    2.714 0.006727 **
## CentralAirY           3.686e-02  1.592e-02    2.316 0.020726 *
## SaleTypeCon           6.035e-02  7.946e-02    0.759 0.447694
## SaleTypeConLD         1.290e-01  4.255e-02    3.033 0.002467 **
## SaleTypeConLI         1.109e-03  5.139e-02    0.022 0.982780
## SaleTypeConLw        -1.158e-04  5.269e-02   -0.002 0.998246
## SaleTypeCWD           4.248e-02  5.742e-02    0.740 0.459520
## SaleTypeNew           1.737e-01  6.816e-02    2.548 0.010935 *
## SaleTypeOth           7.019e-02  6.508e-02    1.079 0.281003
## SaleTypeWD           -6.954e-03  1.841e-02   -0.378 0.705701
## X1stFlrSF             5.117e-02  2.742e-02    1.866 0.062244 .
## StreetPave            9.488e-02  4.896e-02    1.938 0.052824 .
## YearRemodAdd          4.240e-04  2.317e-04    1.830 0.067513 .
## MoSold               -1.787e-03  1.065e-03   -1.677 0.093813 .
## KitchenAbvGr         -9.023e-02  5.928e-02   -1.522 0.128227
## MasVnrTypeBrkFace     5.423e-02  2.892e-02    1.875 0.060993 .
## MasVnrTypeNone        4.872e-02  2.850e-02    1.709 0.087639 .
## MasVnrTypeStone       6.871e-02  3.077e-02    2.233 0.025703 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1047 on 1327 degrees of freedom
## Multiple R-squared:  0.9376, Adjusted R-squared:  0.9313
## F-statistic: 150.9 on 132 and 1327 DF,  p-value: < 2.2e-16
```

```r
summary(StepB_Log)
```

```
##
## Call:
## lm(formula = SalePrice ~ MSZoning + LotArea + Street + LotConfig +
##     Neighborhood + Condition1 + Condition2 + BldgType + OverallQual +
##     OverallCond + YearBuilt + YearRemodAdd + RoofMatl + Exterior1st +
##     Foundation + BsmtQual + BsmtExposure + BsmtFinSF1 + Heating +
##     HeatingQC + CentralAir + X1stFlrSF + GrLivArea + BsmtFullBath +
##     KitchenAbvGr + KitchenQual + Functional + Fireplaces + GarageType +
##     GarageFinish + GarageCars + GarageArea + GarageQual + WoodDeckSF +
##     ScreenPorch + PoolQC + MoSold + SaleType + SaleCondition,
##     data = H_Lo[1:1460, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67962 -0.04954  0.00030  0.05688  0.49948
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -8.418e-01  7.374e-01  -1.142 0.253804
## MSZoningFV            4.154e-01  5.146e-02   8.072 1.53e-15 ***
## MSZoningRH            4.039e-01  5.176e-02   7.803 1.21e-14 ***
## MSZoningRL            3.807e-01  4.367e-02   8.717  < 2e-16 ***
## MSZoningRM            3.597e-01  4.108e-02   8.757  < 2e-16 ***
## LotArea              8.103e-02  1.024e-02   7.915 5.16e-15 ***
## StreetPave           7.749e-02  4.950e-02   1.566 0.117695
## LotConfigCulDSac     1.535e-02  1.376e-02   1.115 0.264923
## LotConfigFR2         -3.713e-02  1.759e-02  -2.111 0.034975 *
## LotConfigFR3         -9.979e-02  5.625e-02  -1.774 0.076277 .
## LotConfigInside      -1.194e-02  7.679e-03  -1.555 0.120272
## NeighborhoodBlueste  -9.840e-03  8.292e-02  -0.119 0.905557
## NeighborhoodBrDale   7.673e-03  4.470e-02   0.172 0.863722
## NeighborhoodBrkSide  2.544e-02  3.878e-02   0.656 0.511902
## NeighborhoodClearCr  7.317e-03  3.915e-02   0.187 0.851789
## NeighborhoodCollgCr  -3.904e-02  3.117e-02  -1.252 0.210681
## NeighborhoodCrawfor  9.905e-02  3.590e-02   2.759 0.005882 **
## NeighborhoodEdwards  -7.581e-02  3.387e-02  -2.239 0.025346 *
## NeighborhoodGilbert  -4.219e-02  3.280e-02  -1.286 0.198591
## NeighborhoodIDOTRR   -3.832e-02  4.470e-02  -0.857 0.391425
## NeighborhoodMeadowV  -1.066e-01  4.701e-02  -2.267 0.023568 *
## NeighborhoodMitchel  -6.324e-02  3.463e-02  -1.826 0.068031 .
## NeighborhoodNAmes    -3.963e-02  3.279e-02  -1.209 0.227046
## NeighborhoodNoRidge  6.553e-02  3.505e-02   1.870 0.061751 .
```

```
## NeighborhoodNPkVill    5.743e-02  4.748e-02   1.210 0.226613
## NeighborhoodNridgHt    8.593e-02  3.117e-02   2.757 0.005910 **
## NeighborhoodNWAmes    -5.136e-02  3.421e-02  -1.502 0.133463
## NeighborhoodOldTown   -4.040e-02  4.006e-02  -1.008 0.313425
## NeighborhoodSawyer    -2.729e-02  3.449e-02  -0.791 0.428845
## NeighborhoodSawyerW   -2.980e-02  3.377e-02  -0.882 0.377705
## NeighborhoodSomerst    3.134e-03  3.821e-02   0.082 0.934647
## NeighborhoodStoneBr    1.209e-01  3.549e-02   3.407 0.000675 ***
## NeighborhoodSWISU     -1.386e-02  4.019e-02  -0.345 0.730344
## NeighborhoodTimber    -2.921e-02  3.487e-02  -0.838 0.402276
## NeighborhoodVeenker   -6.048e-04  4.569e-02  -0.013 0.989441
## Condition1Feedr        4.201e-02  2.147e-02   1.957 0.050578 .
## Condition1Norm         8.757e-02  1.762e-02   4.969 7.61e-07 ***
## Condition1PosA         4.867e-02  4.301e-02   1.132 0.257983
## Condition1PosN         8.905e-02  3.183e-02   2.798 0.005217 **
## Condition1RRAe        -4.265e-02  3.844e-02  -1.110 0.267408
## Condition1RRAn         5.609e-02  2.946e-02   1.904 0.057171 .
## Condition1RRNe         1.901e-02  7.871e-02   0.242 0.809184
## Condition1RRNn         1.326e-01  5.540e-02   2.393 0.016867 *
## Condition2Feedr        1.235e-02  9.649e-02   0.128 0.898160
## Condition2Norm         1.392e-02  8.227e-02   0.169 0.865629
## Condition2PosA         3.476e-01  1.361e-01   2.555 0.010740 *
## Condition2PosN        -6.445e-01  1.170e-01  -5.510 4.30e-08 ***
## Condition2RRAe        -2.485e-02  1.356e-01  -0.183 0.854615
## Condition2RRAn        -4.150e-02  1.351e-01  -0.307 0.758813
## Condition2RRNn         7.438e-02  1.138e-01   0.654 0.513530
## BldgType               1.241e-02  4.665e-03   2.660 0.007909 **
## OverallQual            4.741e-02  4.237e-03  11.191  < 2e-16 ***
## OverallCond            3.683e-02  3.550e-03  10.374  < 2e-16 ***
## YearBuilt              2.032e-03  2.917e-04   6.964 5.17e-12 ***
## YearRemodAdd           4.225e-04  2.307e-04   1.832 0.067236 .
## RoofMatlCompShg        1.690e+00  1.228e-01  13.757  < 2e-16 ***
## RoofMatlMembran        1.834e+00  1.651e-01  11.106  < 2e-16 ***
## RoofMatlMetal          1.754e+00  1.647e-01  10.646  < 2e-16 ***
## RoofMatlRoll           1.688e+00  1.647e-01  10.246  < 2e-16 ***
## RoofMatlTar&Grv        1.657e+00  1.267e-01  13.079  < 2e-16 ***
## RoofMatlWdShake        1.691e+00  1.332e-01  12.689  < 2e-16 ***
## RoofMatlWdShngl        1.853e+00  1.302e-01  14.231  < 2e-16 ***
## Exterior1stAsphShn    -6.924e-02  1.119e-01  -0.619 0.536325
## Exterior1stBrkComm    -1.830e-01  8.421e-02  -2.174 0.029902 *
## Exterior1stBrkFace     9.149e-02  3.123e-02   2.930 0.003453 **
## Exterior1stCBlock     -2.764e-02  1.128e-01  -0.245 0.806408
## Exterior1stCemntBd     4.938e-02  3.242e-02   1.523 0.128046
## Exterior1stHdBoard     8.119e-04  2.839e-02   0.029 0.977191
## Exterior1stImStucc    -2.129e-02  1.101e-01  -0.193 0.846737
## Exterior1stMetalSd     3.305e-02  2.761e-02   1.197 0.231491
## Exterior1stPlywood    -1.392e-03  2.979e-02  -0.047 0.962749
## Exterior1stStone      -7.851e-03  8.523e-02  -0.092 0.926623
## Exterior1stStucco      2.964e-02  3.478e-02   0.852 0.394151
## Exterior1stVinylSd     1.973e-02  2.778e-02   0.710 0.477668
```

```
## Exterior1stWd Sdng      -1.649e-03  2.754e-02   -0.060 0.952274
## Exterior1stWdShing      -1.151e-03  3.434e-02   -0.034 0.973275
## FoundationCBlock         2.648e-02  1.350e-02    1.961 0.050077 .
## FoundationPConc          4.013e-02  1.494e-02    2.686 0.007314 **
## FoundationSlab           4.113e-02  3.506e-02    1.173 0.240986
## FoundationStone          1.240e-01  4.557e-02    2.721 0.006587 **
## FoundationWood          -9.518e-02  6.423e-02   -1.482 0.138653
## BsmtQual                 2.283e-02  6.412e-03    3.561 0.000383 ***
## BsmtExposure             1.552e-02  3.330e-03    4.661 3.46e-06 ***
## BsmtFinSF1               9.064e-03  1.327e-03    6.829 1.29e-11 ***
## HeatingGasA              5.030e-02  1.101e-01    0.457 0.648003
## HeatingGasW              1.320e-01  1.130e-01    1.168 0.242828
## HeatingGrav             -3.260e-02  1.170e-01   -0.279 0.780578
## HeatingOthW             -1.976e-02  1.363e-01   -0.145 0.884753
## HeatingWall              9.832e-02  1.252e-01    0.785 0.432585
## HeatingQC                1.603e-02  4.181e-03    3.834 0.000132 ***
## CentralAirY              3.858e-02  1.564e-02    2.467 0.013731 *
## X1stFlrSF                9.998e-02  1.491e-02    6.705 2.97e-11 ***
## GrLivArea                3.930e-01  1.549e-02   25.374  < 2e-16 ***
## BsmtFullBath             2.993e-02  7.300e-03    4.099 4.39e-05 ***
## KitchenAbvGr            -9.581e-02  4.387e-02   -2.184 0.029127 *
## KitchenQual              2.508e-02  6.870e-03    3.651 0.000272 ***
## Functional               3.854e-02  4.633e-03    8.319  < 2e-16 ***
## Fireplaces               2.054e-02  5.798e-03    3.543 0.000409 ***
## GarageTypeAttchd         8.810e-02  4.718e-02    1.867 0.062058 .
## GarageTypeBasment        6.644e-02  5.352e-02    1.241 0.214698
## GarageTypeBuiltIn        9.228e-02  4.895e-02    1.885 0.059602 .
## GarageTypeCarPort        5.106e-02  5.953e-02    0.858 0.391177
## GarageTypeDetchd         1.003e-01  4.684e-02    2.141 0.032456 *
## GarageTypeNoGarage       1.674e-01  6.140e-02    2.726 0.006490 **
## GarageFinish             7.124e-03  5.180e-03    1.375 0.169245
## GarageCars               2.887e-02  9.821e-03    2.940 0.003344 **
## GarageArea               8.932e-05  3.300e-05    2.706 0.006889 **
## GarageQual               1.954e-02  1.283e-02    1.523 0.128096
## WoodDeckSF               6.922e-05  2.533e-05    2.733 0.006368 **
## ScreenPorch              8.744e-03  2.087e-03    4.189 2.99e-05 ***
## PoolQC                   4.495e-02  1.171e-02    3.840 0.000129 ***
## MoSold                  -1.671e-03  1.061e-03   -1.575 0.115476
## SaleTypeCon              6.443e-02  7.932e-02    0.812 0.416777
## SaleTypeConLD            1.076e-01  4.186e-02    2.570 0.010275 *
## SaleTypeConLI           -1.200e-02  5.158e-02   -0.233 0.816077
## SaleTypeConLw           -2.313e-03  5.238e-02   -0.044 0.964782
## SaleTypeCWD              3.390e-02  5.725e-02    0.592 0.553935
## SaleTypeNew              1.561e-01  6.795e-02    2.297 0.021780 *
## SaleTypeOth              6.692e-02  6.507e-02    1.028 0.303923
## SaleTypeWD              -1.314e-02  1.845e-02   -0.712 0.476450
## SaleConditionAdjLand     1.241e-01  5.794e-02    2.142 0.032336 *
## SaleConditionAlloca      5.460e-02  3.637e-02    1.501 0.133527
## SaleConditionFamily      1.594e-02  2.706e-02    0.589 0.555899
## SaleConditionNormal      6.918e-02  1.253e-02    5.522 4.02e-08 ***
```

```
## SaleConditionPartial -4.354e-02  6.558e-02  -0.664 0.506854
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1046 on 1335 degrees of freedom
## Multiple R-squared:  0.9373, Adjusted R-squared:  0.9314
## F-statistic: 160.8 on 124 and 1335 DF,  p-value: < 2.2e-16

summary(StepS_Log)

##
## Call:
## lm(formula = SalePrice ~ MSZoning + LotArea + Street + LotConfig +
##     Neighborhood + Condition1 + Condition2 + BldgType + OverallQual +
##     OverallCond + YearBuilt + YearRemodAdd + RoofMatl + Exterior1st +
##     Foundation + BsmtQual + BsmtExposure + BsmtFinSF1 + Heating +
##     HeatingQC + CentralAir + X1stFlrSF + GrLivArea + BsmtFullBath +
##     KitchenAbvGr + KitchenQual + Functional + Fireplaces + GarageType +
##     GarageFinish + GarageCars + GarageArea + GarageQual + WoodDeckSF +
##     ScreenPorch + PoolQC + MoSold + SaleType + SaleCondition,
##     data = H_Lo[1:1460, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67962 -0.04954  0.00030  0.05688  0.49948
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -8.418e-01  7.374e-01  -1.142 0.253804
## MSZoningFV            4.154e-01  5.146e-02   8.072 1.53e-15 ***
## MSZoningRH            4.039e-01  5.176e-02   7.803 1.21e-14 ***
## MSZoningRL            3.807e-01  4.367e-02   8.717  < 2e-16 ***
## MSZoningRM            3.597e-01  4.108e-02   8.757  < 2e-16 ***
## LotArea              8.103e-02  1.024e-02   7.915 5.16e-15 ***
## StreetPave           7.749e-02  4.950e-02   1.566 0.117695
## LotConfigCulDSac     1.535e-02  1.376e-02   1.115 0.264923
## LotConfigFR2        -3.713e-02  1.759e-02  -2.111 0.034975 *
## LotConfigFR3        -9.979e-02  5.625e-02  -1.774 0.076277 .
## LotConfigInside     -1.194e-02  7.679e-03  -1.555 0.120272
## NeighborhoodBlueste -9.840e-03  8.292e-02  -0.119 0.905557
## NeighborhoodBrDale   7.673e-03  4.470e-02   0.172 0.863722
## NeighborhoodBrkSide  2.544e-02  3.878e-02   0.656 0.511902
## NeighborhoodClearCr  7.317e-03  3.915e-02   0.187 0.851789
## NeighborhoodCollgCr -3.904e-02  3.117e-02  -1.252 0.210681
## NeighborhoodCrawfor  9.905e-02  3.590e-02   2.759 0.005882 **
## NeighborhoodEdwards -7.581e-02  3.387e-02  -2.239 0.025346 *
## NeighborhoodGilbert -4.219e-02  3.280e-02  -1.286 0.198591
## NeighborhoodIDOTRR  -3.832e-02  4.470e-02  -0.857 0.391425
## NeighborhoodMeadowV -1.066e-01  4.701e-02  -2.267 0.023568 *
## NeighborhoodMitchel -6.324e-02  3.463e-02  -1.826 0.068031 .
```

```
## NeighborhoodNAmes       -3.963e-02  3.279e-02  -1.209 0.227046
## NeighborhoodNoRidge       6.553e-02  3.505e-02   1.870 0.061751 .
## NeighborhoodNPkVill       5.743e-02  4.748e-02   1.210 0.226613
## NeighborhoodNridgHt       8.593e-02  3.117e-02   2.757 0.005910 **
## NeighborhoodNWAmes       -5.136e-02  3.421e-02  -1.502 0.133463
## NeighborhoodOldTown      -4.040e-02  4.006e-02  -1.008 0.313425
## NeighborhoodSawyer       -2.729e-02  3.449e-02  -0.791 0.428845
## NeighborhoodSawyerW      -2.980e-02  3.377e-02  -0.882 0.377705
## NeighborhoodSomerst       3.134e-03  3.821e-02   0.082 0.934647
## NeighborhoodStoneBr       1.209e-01  3.549e-02   3.407 0.000675 ***
## NeighborhoodSWISU        -1.386e-02  4.019e-02  -0.345 0.730344
## NeighborhoodTimber       -2.921e-02  3.487e-02  -0.838 0.402276
## NeighborhoodVeenker      -6.048e-04  4.569e-02  -0.013 0.989441
## Condition1Feedr           4.201e-02  2.147e-02   1.957 0.050578 .
## Condition1Norm            8.757e-02  1.762e-02   4.969 7.61e-07 ***
## Condition1PosA            4.867e-02  4.301e-02   1.132 0.257983
## Condition1PosN            8.905e-02  3.183e-02   2.798 0.005217 **
## Condition1RRAe           -4.265e-02  3.844e-02  -1.110 0.267408
## Condition1RRAn            5.609e-02  2.946e-02   1.904 0.057171 .
## Condition1RRNe            1.901e-02  7.871e-02   0.242 0.809184
## Condition1RRNn            1.326e-01  5.540e-02   2.393 0.016867 *
## Condition2Feedr           1.235e-02  9.649e-02   0.128 0.898160
## Condition2Norm            1.392e-02  8.227e-02   0.169 0.865629
## Condition2PosA            3.476e-01  1.361e-01   2.555 0.010740 *
## Condition2PosN           -6.445e-01  1.170e-01  -5.510 4.30e-08 ***
## Condition2RRAe           -2.485e-02  1.356e-01  -0.183 0.854615
## Condition2RRAn           -4.150e-02  1.351e-01  -0.307 0.758813
## Condition2RRNn            7.438e-02  1.138e-01   0.654 0.513530
## BldgType                  1.241e-02  4.665e-03   2.660 0.007909 **
## OverallQual               4.741e-02  4.237e-03  11.191  < 2e-16 ***
## OverallCond               3.683e-02  3.550e-03  10.374  < 2e-16 ***
## YearBuilt                 2.032e-03  2.917e-04   6.964 5.17e-12 ***
## YearRemodAdd              4.225e-04  2.307e-04   1.832 0.067236 .
## RoofMatlCompShg           1.690e+00  1.228e-01  13.757  < 2e-16 ***
## RoofMatlMembran           1.834e+00  1.651e-01  11.106  < 2e-16 ***
## RoofMatlMetal             1.754e+00  1.647e-01  10.646  < 2e-16 ***
## RoofMatlRoll              1.688e+00  1.647e-01  10.246  < 2e-16 ***
## RoofMatlTar&Grv           1.657e+00  1.267e-01  13.079  < 2e-16 ***
## RoofMatlWdShake           1.691e+00  1.332e-01  12.689  < 2e-16 ***
## RoofMatlWdShngl           1.853e+00  1.302e-01  14.231  < 2e-16 ***
## Exterior1stAsphShn       -6.924e-02  1.119e-01  -0.619 0.536325
## Exterior1stBrkComm       -1.830e-01  8.421e-02  -2.174 0.029902 *
## Exterior1stBrkFace        9.149e-02  3.123e-02   2.930 0.003453 **
## Exterior1stCBlock        -2.764e-02  1.128e-01  -0.245 0.806408
## Exterior1stCemntBd        4.938e-02  3.242e-02   1.523 0.128046
## Exterior1stHdBoard        8.119e-04  2.839e-02   0.029 0.977191
## Exterior1stImStucc       -2.129e-02  1.101e-01  -0.193 0.846737
## Exterior1stMetalSd        3.305e-02  2.761e-02   1.197 0.231491
## Exterior1stPlywood       -1.392e-03  2.979e-02  -0.047 0.962749
## Exterior1stStone         -7.851e-03  8.523e-02  -0.092 0.926623
```

```
## Exterior1stStucco      2.964e-02  3.478e-02    0.852 0.394151
## Exterior1stVinylSd     1.973e-02  2.778e-02    0.710 0.477668
## Exterior1stWd Sdng    -1.649e-03  2.754e-02   -0.060 0.952274
## Exterior1stWdShing    -1.151e-03  3.434e-02   -0.034 0.973275
## FoundationCBlock       2.648e-02  1.350e-02    1.961 0.050077 .
## FoundationPConc        4.013e-02  1.494e-02    2.686 0.007314 **
## FoundationSlab         4.113e-02  3.506e-02    1.173 0.240986
## FoundationStone        1.240e-01  4.557e-02    2.721 0.006587 **
## FoundationWood        -9.518e-02  6.423e-02   -1.482 0.138653
## BsmtQual               2.283e-02  6.412e-03    3.561 0.000383 ***
## BsmtExposure           1.552e-02  3.330e-03    4.661 3.46e-06 ***
## BsmtFinSF1             9.064e-03  1.327e-03    6.829 1.29e-11 ***
## HeatingGasA            5.030e-02  1.101e-01    0.457 0.648003
## HeatingGasW            1.320e-01  1.130e-01    1.168 0.242828
## HeatingGrav           -3.260e-02  1.170e-01   -0.279 0.780578
## HeatingOthW           -1.976e-02  1.363e-01   -0.145 0.884753
## HeatingWall            9.832e-02  1.252e-01    0.785 0.432585
## HeatingQC              1.603e-02  4.181e-03    3.834 0.000132 ***
## CentralAirY            3.858e-02  1.564e-02    2.467 0.013731 *
## X1stFlrSF              9.998e-02  1.491e-02    6.705 2.97e-11 ***
## GrLivArea              3.930e-01  1.549e-02   25.374  < 2e-16 ***
## BsmtFullBath           2.993e-02  7.300e-03    4.099 4.39e-05 ***
## KitchenAbvGr          -9.581e-02  4.387e-02   -2.184 0.029127 *
## KitchenQual            2.508e-02  6.870e-03    3.651 0.000272 ***
## Functional             3.854e-02  4.633e-03    8.319  < 2e-16 ***
## Fireplaces             2.054e-02  5.798e-03    3.543 0.000409 ***
## GarageTypeAttchd       8.810e-02  4.718e-02    1.867 0.062058 .
## GarageTypeBasment      6.644e-02  5.352e-02    1.241 0.214698
## GarageTypeBuiltIn      9.228e-02  4.895e-02    1.885 0.059602 .
## GarageTypeCarPort      5.106e-02  5.953e-02    0.858 0.391177
## GarageTypeDetchd       1.003e-01  4.684e-02    2.141 0.032456 *
## GarageTypeNoGarage     1.674e-01  6.140e-02    2.726 0.006490 **
## GarageFinish           7.124e-03  5.180e-03    1.375 0.169245
## GarageCars             2.887e-02  9.821e-03    2.940 0.003344 **
## GarageArea             8.932e-05  3.300e-05    2.706 0.006889 **
## GarageQual             1.954e-02  1.283e-02    1.523 0.128096
## WoodDeckSF             6.922e-05  2.533e-05    2.733 0.006368 **
## ScreenPorch            8.744e-03  2.087e-03    4.189 2.99e-05 ***
## PoolQC                 4.495e-02  1.171e-02    3.840 0.000129 ***
## MoSold                -1.671e-03  1.061e-03   -1.575 0.115476
## SaleTypeCon            6.443e-02  7.932e-02    0.812 0.416777
## SaleTypeConLD          1.076e-01  4.186e-02    2.570 0.010275 *
## SaleTypeConLI         -1.200e-02  5.158e-02   -0.233 0.816077
## SaleTypeConLw         -2.313e-03  5.238e-02   -0.044 0.964782
## SaleTypeCWD            3.390e-02  5.725e-02    0.592 0.553935
## SaleTypeNew            1.561e-01  6.795e-02    2.297 0.021780 *
## SaleTypeOth            6.692e-02  6.507e-02    1.028 0.303923
## SaleTypeWD            -1.314e-02  1.845e-02   -0.712 0.476450
## SaleConditionAdjLand   1.241e-01  5.794e-02    2.142 0.032336 *
## SaleConditionAlloca    5.460e-02  3.637e-02    1.501 0.133527
```

```
## SaleConditionFamily   1.594e-02  2.706e-02    0.589 0.555899
## SaleConditionNormal    6.918e-02  1.253e-02    5.522 4.02e-08 ***
## SaleConditionPartial  -4.354e-02  6.558e-02   -0.664 0.506854
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1046 on 1335 degrees of freedom
## Multiple R-squared:  0.9373, Adjusted R-squared:  0.9314
## F-statistic: 160.8 on 124 and 1335 DF,  p-value: < 2.2e-16
```

Train<-H_Clean[!is.na(H_Clean$SalePrice),-1]

StepF_Log$call

```
## lm(formula = SalePrice ~ OverallQual + GrLivArea + Neighborhood +
##     MSSubClass + BsmtFinSF1 + OverallCond + GarageCars + YearBuilt +
##     RoofMatl + LotArea + MSZoning + BsmtExposure + Condition2 +
##     Functional + SaleCondition + KitchenQual + Condition1 + Exterior1st +
##     BsmtFullBath + Fireplaces + BsmtQual + HeatingQC + PoolQC +
##     ScreenPorch + Heating + LotConfig + GarageArea + Foundation +
##     WoodDeckSF + CentralAir + SaleType + X1stFlrSF + Street +
##     YearRemodAdd + MoSold + KitchenAbvGr + MasVnrType, data = H_Lo[1:1460,

##     ])
```

StepB_Log$call

```
## lm(formula = SalePrice ~ MSZoning + LotArea + Street + LotConfig +
##     Neighborhood + Condition1 + Condition2 + BldgType + OverallQual +
##     OverallCond + YearBuilt + YearRemodAdd + RoofMatl + Exterior1st +
##     Foundation + BsmtQual + BsmtExposure + BsmtFinSF1 + Heating +
##     HeatingQC + CentralAir + X1stFlrSF + GrLivArea + BsmtFullBath +
##     KitchenAbvGr + KitchenQual + Functional + Fireplaces + GarageType +
##     GarageFinish + GarageCars + GarageArea + GarageQual + WoodDeckSF +
##     ScreenPorch + PoolQC + MoSold + SaleType + SaleCondition,
##     data = H_Lo[1:1460, ])
```

StepS_Log$call

```
## lm(formula = SalePrice ~ MSZoning + LotArea + Street + LotConfig +
##     Neighborhood + Condition1 + Condition2 + BldgType + OverallQual +
##     OverallCond + YearBuilt + YearRemodAdd + RoofMatl + Exterior1st +
##     Foundation + BsmtQual + BsmtExposure + BsmtFinSF1 + Heating +
##     HeatingQC + CentralAir + X1stFlrSF + GrLivArea + BsmtFullBath +
##     KitchenAbvGr + KitchenQual + Functional + Fireplaces + GarageType +
##     GarageFinish + GarageCars + GarageArea + GarageQual + WoodDeckSF +
##     ScreenPorch + PoolQC + MoSold + SaleType + SaleCondition,
##     data = H_Lo[1:1460, ])
```

```
#summary(StepB_Log)
#summary(StepS_Log)
identical(StepB_Log$call, StepS_Log$call) # Shows they are same.

## [1] TRUE

identical(StepF_Log$call, StepS_Log$call)

## [1] FALSE

TR<-1:1460
TR_RmvOut<-TR[-c(524,1299)]

a<-(lm(SalePrice~.,data=H_Lo[TR,-1])) # Removing ID
b<-lm(StepF_Log$call,data=H_Lo[TR,-1])
c<-lm(StepS_Log$call,data=H_Lo[TR,-1])
par(mfrow=c(2,2))
plot(a, sub='ALL Var')

## Warning: not plotting observations with leverage one:
##    121, 272, 347, 584, 596, 1004, 1012, 1188, 1231, 1271, 1276, 1299, 1322,
 1371, 1387

## Warning: not plotting observations with leverage one:
##    121, 272, 347, 584, 596, 1004, 1012, 1188, 1231, 1271, 1276, 1299, 1322,
 1371, 1387

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```
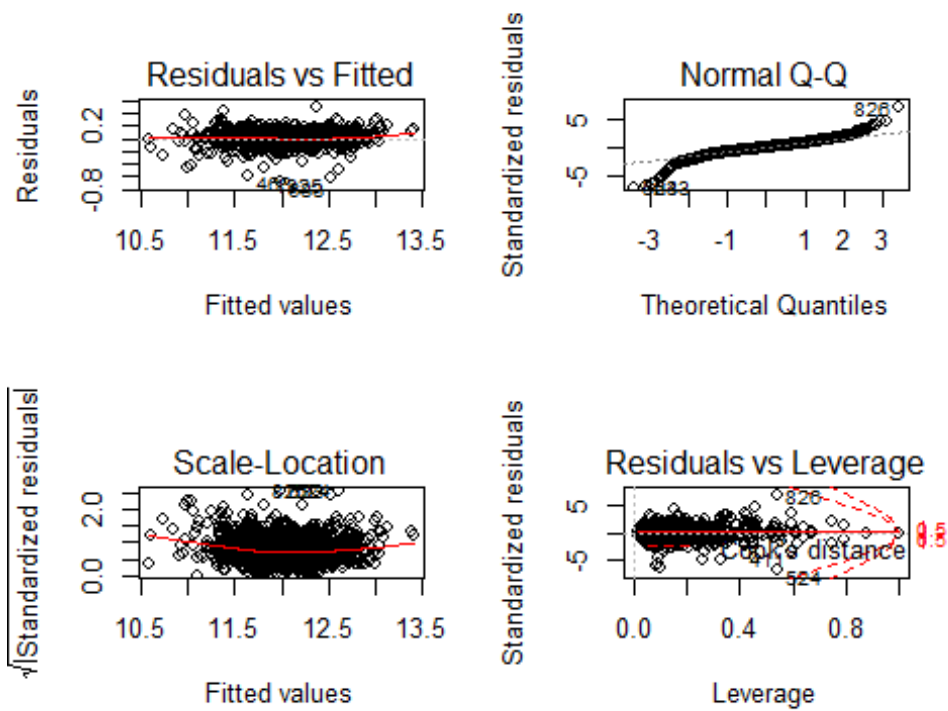
```
plot(b, sub='Foward Sel')
```

```
## Warning: not plotting observations with leverage one:
##    121, 272, 584, 1004, 1012, 1188, 1231, 1276, 1299, 1322, 1371
```

```
## Warning: not plotting observations with leverage one:
##    121, 272, 584, 1004, 1012, 1188, 1231, 1276, 1299, 1322, 1371
```

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage

```
plot(c, sub='Backwards El')

## Warning: not plotting observations with leverage one:
##    121, 272, 584, 1004, 1012, 1188, 1231, 1276, 1299, 1322, 1371

## Warning: not plotting observations with leverage one:
##    121, 272, 584, 1004, 1012, 1188, 1231, 1276, 1299, 1322, 1371
```

Residuals vs Fitted

Residuals

0.4

-0.8

10.5   11.5   12.5   13.5

Fitted values

Normal Q-Q

Standardized residuals

5

-5

826

-3   -1   1 2 3

Theoretical Quantiles

Scale-Location

√|Standardized residuals|

2.0

0.0

10.5   11.5   12.5   13.5

Fitted values

Residuals vs Leverage

Standardized residuals

5

-5

826

0.5

Cook's distance

0.5

52

0.0   0.2   0.4   0.6
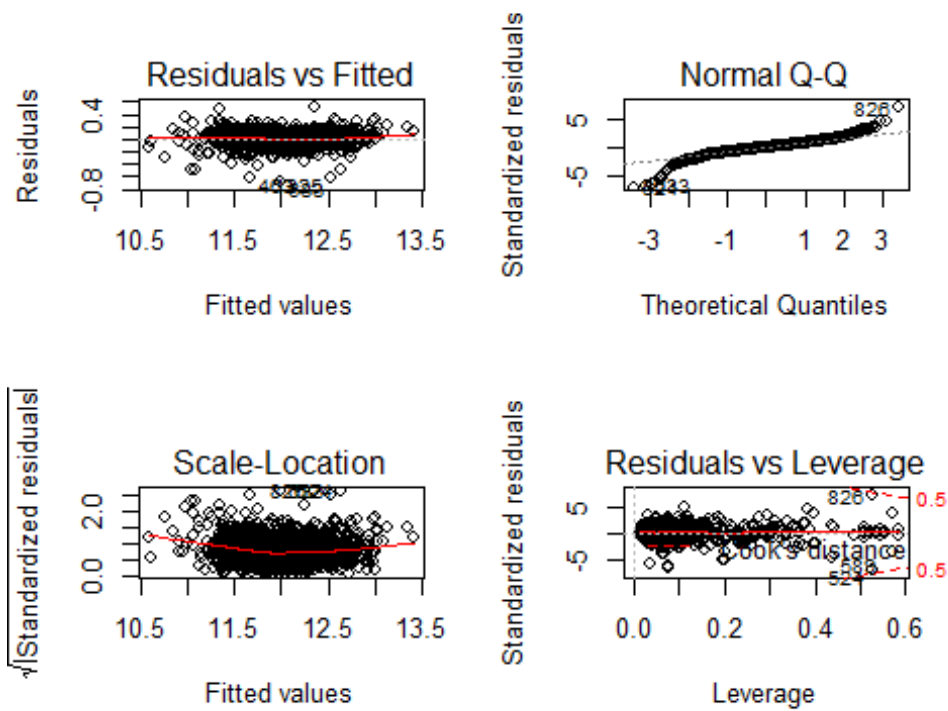
Leverage
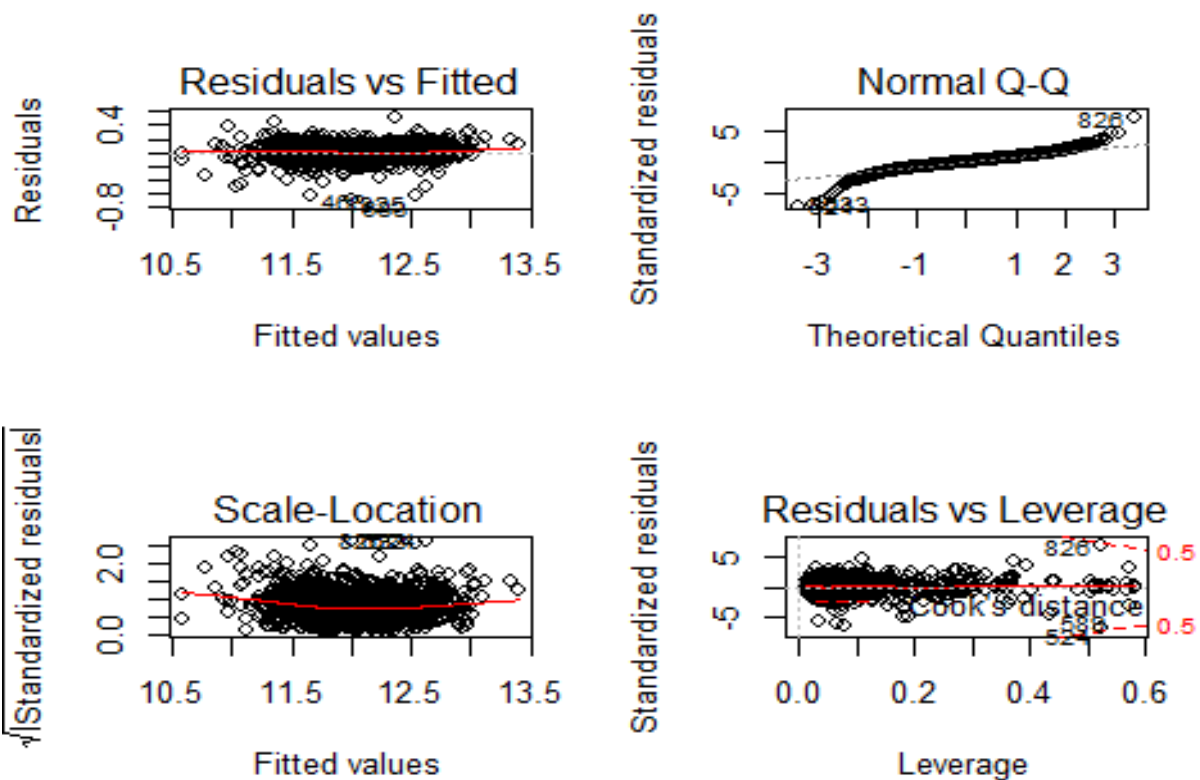
```r
plot(a,which=c(4))
plot(b,which=c(4))


#pdtrain<-predict(a,newdata=H_Lo[!is.na(H_Lo$SalePrice),])
pd_Log_LM_All<-predict(a,newdata=H_Lo[is.na(H_Lo$SalePrice),])

## Warning in predict.lm(a, newdata = H_Lo[is.na(H_Lo$SalePrice), ]):
## prediction from a rank-deficient fit may be misleading

pd_Log_LM_f2<-predict(b,newdata=H_Lo[is.na(H_Lo$SalePrice),])
pd_Log_LM_f3<-predict(c,newdata=H_Lo[is.na(H_Lo$SalePrice),])
#trainSP<-H_Lo$SalePrice[!is.na(H_Lo$SalePrice)]


#plot(exp(trainSP),exp(pdtrain), col='blue', ylab='Prediction')
#points(actual,exp(pdtest), col='red')
#abline(b=1,a=0)
#RMSE(exp(trainSP),exp(pdtrain))
#RMSE(actual,exp(pdtest))

print("RMSLE for All Var - Log")

## [1] "RMSLE for All Var - Log"
```

```
RMSE(pd_Log_LM_All,log(actual))
```

```
## [1] 0.1368791
```

```
RMSE(pd_Log_LM_f2,log(actual))
```

```
## [1] 0.1338353
```

```
RMSE(pd_Log_LM_f3,log(actual))
```

```
## [1] 0.1334131
```
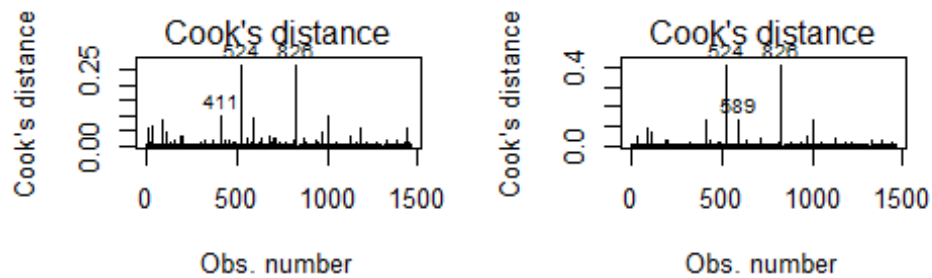

Try without outliers

```
summary(a)
```

```
##
## Call:
## lm(formula = SalePrice ~ ., data = H_Lo[TR, -1])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67396 -0.04631  0.00222  0.05248  0.48404
##
## Coefficients: (1 not defined because of singularities)
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               5.955e+00  4.708e+00   1.265 0.206185
## MSSubClass1NHFin          7.060e-02  7.237e-02   0.975 0.329515
## MSSubClass1NHUnf         -1.140e-01  1.212e-01  -0.941 0.346844
```

```
## MSSubClass1SFA       -4.537e-02  1.084e-01  -0.418 0.675682
## MSSubClass1SNEW       8.489e-02  6.916e-02   1.227 0.219907
## MSSubClass1SOLD       5.203e-02  7.313e-02   0.711 0.476934
## MSSubClass1SPUD       7.871e-02  4.299e-02   1.831 0.067330 .
## MSSubClass2FCONV      1.934e-02  6.038e-02   0.320 0.748794
## MSSubClass2NHS       -4.257e-02  8.848e-02  -0.481 0.630547
## MSSubClass2SNEW       3.810e-02  6.047e-02   0.630 0.528790
## MSSubClass2SOLD       6.222e-02  6.481e-02   0.960 0.337167
## MSSubClassDUPL        6.346e-02  5.649e-02   1.123 0.261472
## MSSubClassMLPUD       1.219e-01  6.248e-02   1.951 0.051230 .
## MSSubClassSPL         1.022e-01  8.181e-02   1.249 0.211797
## MSSubClassSPLF        8.269e-02  8.303e-02   0.996 0.319520
## MSZoningFV            4.714e-01  5.514e-02   8.550  < 2e-16 ***
## MSZoningRH            4.260e-01  5.486e-02   7.765 1.70e-14 ***
## MSZoningRL            4.245e-01  4.688e-02   9.056  < 2e-16 ***
## MSZoningRM            4.015e-01  4.426e-02   9.073  < 2e-16 ***
## LotFrontage          2.112e-02  1.651e-02   1.279 0.201086
## LotArea              8.364e-02  1.253e-02   6.674 3.72e-11 ***
## StreetPave           6.663e-02  5.275e-02   1.263 0.206823
## AlleyNoAlley         4.428e-03  1.925e-02   0.230 0.818144
## AlleyPave            4.931e-02  2.787e-02   1.769 0.077124 .
## LotShapeIR2          2.871e-02  1.910e-02   1.503 0.133106
## LotShapeIR3          2.952e-02  3.918e-02   0.754 0.451288
## LotShapeReg          1.296e-02  7.352e-03   1.763 0.078229 .
## LandContourHLS       2.931e-02  2.340e-02   1.253 0.210501
## LandContourLow      -1.155e-02  2.863e-02  -0.404 0.686601
## LandContourLvl       2.385e-02  1.683e-02   1.417 0.156687
## LotConfigCulDSac     2.925e-02  1.688e-02   1.733 0.083348 .
## LotConfigFR2        -3.597e-02  1.848e-02  -1.946 0.051862 .
## LotConfigFR3        -1.143e-01  5.762e-02  -1.984 0.047490 *
## LotConfigInside     -9.662e-03  8.228e-03  -1.174 0.240497
## LandSlope           -1.500e-03  1.524e-02  -0.098 0.921600
## NeighborhoodBlueste  6.923e-02  9.074e-02   0.763 0.445615
## NeighborhoodBrDale   2.939e-02  5.373e-02   0.547 0.584419
## NeighborhoodBrkSide  2.225e-02  4.399e-02   0.506 0.613051
## NeighborhoodClearCr  1.264e-02  4.353e-02   0.290 0.771668
## NeighborhoodCollgCr -3.464e-02  3.480e-02  -0.996 0.319660
## NeighborhoodCrawfor  1.004e-01  4.070e-02   2.466 0.013790 *
## NeighborhoodEdwards -8.106e-02  3.807e-02  -2.129 0.033424 *
## NeighborhoodGilbert -3.408e-02  3.674e-02  -0.928 0.353764
## NeighborhoodIDOTRR  -3.391e-02  4.999e-02  -0.678 0.497675
## NeighborhoodMeadowV -9.887e-02  5.692e-02  -1.737 0.082633 .
## NeighborhoodMitchel -5.559e-02  3.881e-02  -1.432 0.152302
## NeighborhoodNAmes   -4.261e-02  3.709e-02  -1.149 0.250884
## NeighborhoodNoRidge  6.784e-02  3.941e-02   1.721 0.085414 .
## NeighborhoodNPkVill  4.150e-02  6.405e-02   0.648 0.517191
## NeighborhoodNridgHt  8.428e-02  3.530e-02   2.388 0.017096 *
## NeighborhoodNWAmes  -4.947e-02  3.821e-02  -1.295 0.195699
## NeighborhoodOldTown -3.949e-02  4.506e-02  -0.876 0.380963
## NeighborhoodSawyer  -2.619e-02  3.835e-02  -0.683 0.494824
```

```
## NeighborhoodSawyerW       -2.121e-02  3.716e-02  -0.571 0.568261
## NeighborhoodSomerst       -5.873e-03  4.241e-02  -0.138 0.889888
## NeighborhoodStoneBr        1.417e-01  3.895e-02   3.639 0.000285 ***
## NeighborhoodSWISU         -3.261e-03  4.527e-02  -0.072 0.942587
## NeighborhoodTimber        -2.480e-02  3.878e-02  -0.639 0.522669
## NeighborhoodVeenker        1.256e-02  4.912e-02   0.256 0.798157
## Condition1Feedr            3.450e-02  2.270e-02   1.520 0.128831
## Condition1Norm             8.692e-02  1.880e-02   4.624 4.15e-06 ***
## Condition1PosA             3.877e-02  4.537e-02   0.855 0.392913
## Condition1PosN             8.596e-02  3.377e-02   2.545 0.011036 *
## Condition1RRAe            -5.332e-02  3.972e-02  -1.342 0.179704
## Condition1RRAn             6.036e-02  3.148e-02   1.917 0.055439 .
## Condition1RRNe             1.501e-02  8.057e-02   0.186 0.852215
## Condition1RRNn             1.030e-01  5.856e-02   1.758 0.078981 .
## Condition2Feedr            1.395e-01  1.151e-01   1.212 0.225759
## Condition2Norm             9.989e-02  1.017e-01   0.982 0.326193
## Condition2PosA             4.314e-01  1.562e-01   2.761 0.005840 **
## Condition2PosN            -5.375e-01  1.318e-01  -4.079 4.82e-05 ***
## Condition2RRAe            -2.115e-01  2.444e-01  -0.865 0.387045
## Condition2RRAn             4.383e-02  1.498e-01   0.293 0.769924
## Condition2RRNn             1.709e-01  1.303e-01   1.311 0.190070
## BldgType                   6.955e-03  1.618e-02   0.430 0.667388
## HouseStyle1.5Unf           2.464e-01  1.049e-01   2.350 0.018933 *
## HouseStyle1Story           2.785e-02  4.523e-02   0.616 0.538078
## HouseStyle2.5Fin           9.312e-02  7.479e-02   1.245 0.213379
## HouseStyle2.5Unf           1.266e-01  7.148e-02   1.771 0.076861 .
## HouseStyle2Story           2.201e-02  3.702e-02   0.595 0.552255
## HouseStyleSFoyer           5.783e-03  5.824e-02   0.099 0.920912
## HouseStyleSLvl            -2.327e-02  6.062e-02  -0.384 0.701174
## OverallQual                4.113e-02  4.645e-03   8.855  < 2e-16 ***
## OverallCond                3.761e-02  3.935e-03   9.558  < 2e-16 ***
## YearBuilt                  2.108e-03  3.831e-04   5.501 4.57e-08 ***
## YearRemodAdd               4.751e-04  2.471e-04   1.923 0.054768 .
## RoofStyleGable             3.689e-03  8.257e-02   0.045 0.964375
## RoofStyleGambrel          -4.374e-02  9.040e-02  -0.484 0.628582
## RoofStyleHip               1.553e-02  8.277e-02   0.188 0.851164
## RoofStyleMansard           2.878e-02  9.601e-02   0.300 0.764430
## RoofStyleShed              2.363e-01  1.702e-01   1.389 0.165220
## RoofMatlCompShg            1.718e+00  1.361e-01  12.628  < 2e-16 ***
## RoofMatlMembran            1.896e+00  1.951e-01   9.720  < 2e-16 ***
## RoofMatlMetal              1.782e+00  1.985e-01   8.975  < 2e-16 ***
## RoofMatlRoll               1.639e+00  1.780e-01   9.207  < 2e-16 ***
## RoofMatlTar&Grv            1.697e+00  1.597e-01  10.628  < 2e-16 ***
## RoofMatlWdShake            1.689e+00  1.520e-01  11.111  < 2e-16 ***
## RoofMatlWdShngl            1.908e+00  1.445e-01  13.209  < 2e-16 ***
## Exterior1stAsphShn        -9.239e-02  1.513e-01  -0.611 0.541588
## Exterior1stBrkComm        -3.157e-01  1.270e-01  -2.486 0.013042 *
## Exterior1stBrkFace         6.916e-02  5.816e-02   1.189 0.234623
## Exterior1stCBlock         -4.370e-02  1.163e-01  -0.376 0.707230
## Exterior1stCemntBd        -5.485e-02  8.587e-02  -0.639 0.523052
```

```
## Exterior1stHdBoard        -2.918e-02  5.863e-02  -0.498 0.618761
## Exterior1stImStucc        -8.538e-02  1.281e-01  -0.666 0.505294
## Exterior1stMetalSd         1.632e-02  6.675e-02   0.245 0.806851
## Exterior1stPlywood        -2.283e-02  5.789e-02  -0.394 0.693406
## Exterior1stStone          -4.263e-02  1.082e-01  -0.394 0.693650
## Exterior1stStucco          2.101e-02  6.371e-02   0.330 0.741642
## Exterior1stVinylSd        -9.268e-03  6.143e-02  -0.151 0.880104
## Exterior1stWd Sdng        -6.144e-02  5.643e-02  -1.089 0.276406
## Exterior1stWdShing        -1.169e-02  6.107e-02  -0.191 0.848175
## Exterior2ndAsphShn         7.348e-02  1.015e-01   0.724 0.469075
## Exterior2ndBrk Cmn         1.191e-01  9.365e-02   1.272 0.203566
## Exterior2ndBrkFace        -6.928e-04  5.936e-02  -0.012 0.990691
## Exterior2ndCBlock                 NA         NA      NA       NA
## Exterior2ndCmentBd         1.088e-01  8.417e-02   1.293 0.196299
## Exterior2ndHdBoard         3.634e-02  5.570e-02   0.652 0.514217
## Exterior2ndImStucc         6.716e-02  6.475e-02   1.037 0.299864
## Exterior2ndMetalSd         2.138e-02  6.473e-02   0.330 0.741251
## Exterior2ndOther          -1.003e-01  1.246e-01  -0.806 0.420606
## Exterior2ndPlywood         2.673e-02  5.410e-02   0.494 0.621380
## Exterior2ndStone           2.620e-02  7.768e-02   0.337 0.735923
## Exterior2ndStucco          2.160e-04  6.173e-02   0.003 0.997209
## Exterior2ndVinylSd         3.331e-02  5.883e-02   0.566 0.571415
## Exterior2ndWd Sdng         7.197e-02  5.378e-02   1.338 0.181012
## Exterior2ndWd Shng         2.253e-02  5.653e-02   0.398 0.690333
## MasVnrTypeBrkFace          4.441e-02  3.024e-02   1.469 0.142194
## MasVnrTypeNone             5.469e-02  4.075e-02   1.342 0.179743
## MasVnrTypeStone            6.418e-02  3.209e-02   2.000 0.045685 *
## MasVnrArea                 3.533e-03  5.384e-03   0.656 0.511827
## ExterQual                  2.236e-03  9.565e-03   0.234 0.815212
## ExterCond                 -1.183e-02  9.597e-03  -1.233 0.217862
## FoundationCBlock           2.309e-02  1.451e-02   1.591 0.111811
## FoundationPConc            4.063e-02  1.560e-02   2.605 0.009310 **
## FoundationSlab             6.109e-02  4.511e-02   1.354 0.175858
## FoundationStone            1.128e-01  4.994e-02   2.259 0.024083 *
## FoundationWood            -1.070e-01  6.747e-02  -1.586 0.113036
## BsmtQual                   1.244e-02  8.206e-03   1.516 0.129666
## BsmtCond                  -6.147e-03  1.126e-02  -0.546 0.585197
## BsmtExposure               1.617e-02  3.858e-03   4.192 2.96e-05 ***
## BsmtFinType1               1.773e-03  3.390e-03   0.523 0.601140
## BsmtFinSF1                 7.132e-03  2.514e-03   2.836 0.004637 **
## BsmtFinType2               2.107e-03  7.842e-03   0.269 0.788264
## BsmtFinSF2                -2.813e-03  3.842e-03  -0.732 0.464181
## BsmtUnfSF                 -4.949e-03  2.700e-03  -1.833 0.067016 .
## TotalBsmtSF                1.701e-02  7.978e-03   2.133 0.033151 *
## HeatingGasA                1.219e-02  1.164e-01   0.105 0.916605
## HeatingGasW                9.816e-02  1.197e-01   0.820 0.412255
## HeatingGrav               -9.541e-02  1.247e-01  -0.765 0.444262
## HeatingOthW               -6.390e-02  1.429e-01  -0.447 0.654914
## HeatingWall                6.492e-02  1.348e-01   0.482 0.630107
## HeatingQC                  1.468e-02  4.361e-03   3.366 0.000786 ***
```

```
## CentralAirY                    4.418e-02  1.745e-02   2.531 0.011485 *
## ElectricalFuseF               -1.448e-02  2.618e-02  -0.553 0.580140
## ElectricalFuseP               -8.806e-03  7.717e-02  -0.114 0.909163
## ElectricalMix                 -2.982e-02  1.187e-01  -0.251 0.801710
## ElectricalSBrkr               -1.736e-02  1.352e-02  -1.284 0.199550
## X1stFlrSF                      3.637e-02  4.532e-02   0.802 0.422420
## X2ndFlrSF                      2.659e-03  5.814e-03   0.457 0.647451
## LowQualFinSF                  -6.090e-03  6.204e-03  -0.982 0.326513
## GrLivArea                      4.224e-01  5.346e-02   7.903 5.97e-15 ***
## BsmtFullBath                   2.946e-02  8.304e-03   3.547 0.000404 ***
## BsmtHalfBath                   5.431e-03  1.973e-02   0.275 0.783140
## BedroomAbvGr                  -6.480e-03  6.196e-03  -1.046 0.295901
## KitchenAbvGr                  -9.608e-02  6.832e-02  -1.406 0.159874
## KitchenQual                    2.518e-02  7.415e-03   3.396 0.000706 ***
## TotRmsAbvGrd                   4.133e-02  3.197e-02   1.293 0.196244
## Functional                     3.419e-02  5.133e-03   6.662 4.04e-11 ***
## Fireplaces                     2.188e-02  9.918e-03   2.206 0.027544 *
## FireplaceQu                   -1.909e-04  3.618e-03  -0.053 0.957917
## GarageTypeAttchd               1.042e-01  5.012e-02   2.080 0.037751 *
## GarageTypeBasment              9.805e-02  5.801e-02   1.690 0.091235 .
## GarageTypeBuiltIn              1.072e-01  5.215e-02   2.055 0.040048 *
## GarageTypeCarPort              1.012e-01  6.536e-02   1.549 0.121702
## GarageTypeDetchd               1.219e-01  4.998e-02   2.439 0.014871 *
## GarageTypeNoGarage             2.345e-01  7.061e-02   3.321 0.000923 ***
## GarageYrBlt                   -8.767e-05  2.709e-04  -0.324 0.746242
## GarageFinish                   6.436e-03  5.461e-03   1.179 0.238796
## GarageCars                     2.747e-02  1.019e-02   2.695 0.007142 **
## GarageArea                     9.212e-05  3.576e-05   2.576 0.010108 *
## GarageQual                     2.422e-02  1.747e-02   1.386 0.165910
## GarageCond                     8.692e-03  1.815e-02   0.479 0.632040
## PavedDrive                     5.508e-03  7.564e-03   0.728 0.466598
## WoodDeckSF                     8.456e-05  2.678e-05   3.158 0.001627 **
## OpenPorchSF                    1.776e-03  1.723e-03   1.031 0.302890
## EnclosedPorch                  2.963e-03  2.052e-03   1.444 0.148940
## X3SsnPorch                     4.407e-03  4.604e-03   0.957 0.338580
## ScreenPorch                    9.393e-03  2.193e-03   4.284 1.98e-05 ***
## PoolArea                      -1.627e-02  2.863e-02  -0.568 0.569815
## PoolQC                         7.997e-02  4.534e-02   1.764 0.078041 .
## FenceGdWo                     -2.448e-02  2.221e-02  -1.102 0.270539
## FenceMnPrv                     1.368e-02  1.815e-02   0.753 0.451404
## FenceMnWw                      5.371e-03  3.750e-02   0.143 0.886134
## FenceNoFence                   1.354e-02  1.652e-02   0.819 0.412747
## MiscFeatureNoMiscFeature -1.975e-01  1.691e-01  -1.168 0.243062
## MiscFeatureOthr               -1.279e-01  1.587e-01  -0.806 0.420367
## MiscFeatureShed               -9.447e-02  1.175e-01  -0.804 0.421646
## MiscFeatureTenC               -1.398e-01  1.847e-01  -0.757 0.449250
## MiscVal                       -1.662e-02  1.355e-02  -1.226 0.220316
## MoSold                        -1.514e-03  1.107e-03  -1.367 0.171914
## YrSold                        -3.445e-03  2.310e-03  -1.491 0.136089
## SaleTypeCon                    8.367e-02  8.066e-02   1.037 0.299810
```

```
## SaleTypeConLD                1.387e-01  4.449e-02   3.117 0.001868 **
## SaleTypeConLI               -6.487e-04  5.276e-02  -0.012 0.990192
## SaleTypeConLw               -2.245e-03  5.436e-02  -0.041 0.967064
## SaleTypeCWD                  6.375e-02  5.870e-02   1.086 0.277688
## SaleTypeNew                  1.621e-01  6.953e-02   2.331 0.019890 *
## SaleTypeOth                  1.761e-02  7.340e-02   0.240 0.810449
## SaleTypeWD                  -5.866e-03  1.903e-02  -0.308 0.757913
## SaleConditionAdjLand         1.331e-01  6.689e-02   1.990 0.046862 *
## SaleConditionAlloca          4.850e-02  3.944e-02   1.230 0.219008
## SaleConditionFamily          1.559e-03  2.790e-02   0.056 0.955453
## SaleConditionNormal          6.190e-02  1.312e-02   4.716 2.67e-06 ***
## SaleConditionPartial        -5.966e-02  6.681e-02  -0.893 0.372070
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1046 on 1246 degrees of freedom
## Multiple R-squared:  0.9415, Adjusted R-squared:  0.9314
## F-statistic: 94.07 on 213 and 1246 DF,  p-value: < 2.2e-16
```

```r
summary(H_Lo$Exterior2nd)
```

```
## AsbShng AsphShn Brk Cmn BrkFace  CBlock CmentBd HdBoard ImStucc MetalSd
##      38       4      22      47       3     126     406      15     447
##   Other Plywood   Stone  Stucco VinylSd Wd Sdng Wd Shng
##       1     270       6      47    1015     391      81
```

Remove Outlier and check

```r
a_o<-(lm(SalePrice~.,data=H_Lo[TR_RmvOut,-1])) # Removing ID
b_o<-lm(StepF_Log$call,data=H_Lo[TR_RmvOut,-1])
c_o<-lm(StepS_Log$call,data=H_Lo[TR_RmvOut,-1])
par(mfrow=c(2,2))
plot(a_o, main='ALL Var')
```

```
## Warning: not plotting observations with leverage one:
##   121, 347, 583, 595, 825, 1003, 1011, 1187, 1230, 1270, 1275, 1320, 1369,
##  1385
```

```
## Warning: not plotting observations with leverage one:
##   121, 347, 583, 595, 825, 1003, 1011, 1187, 1230, 1270, 1275, 1320, 1369,
##  1385
```

```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

ALL Var — Residuals vs Fitted

ALL Var — Normal Q-Q

ALL Var — Scale-Location

ALL Var — Residuals vs Leverage
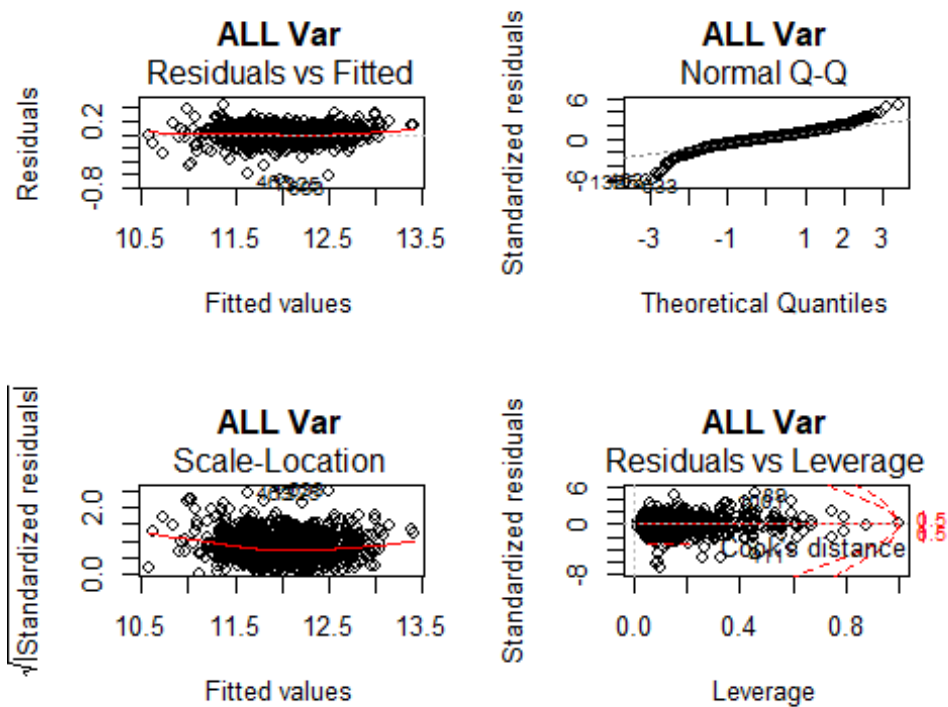
```
plot(b_o, main='Foward Sel')

## Warning: not plotting observations with leverage one:
##    272, 583, 825, 1003, 1011, 1187, 1230, 1275, 1320, 1369

## Warning: not plotting observations with leverage one:
##    272, 583, 825, 1003, 1011, 1187, 1230, 1275, 1320, 1369

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

Foward Sel
Residuals vs Fitted

Residuals

0.2
-0.8

10.5   11.5   12.5   13.5

Fitted values

Foward Sel
Normal Q-Q

Standardized residuals

6
0
-6

-3   -1   1 2 3

Theoretical Quantiles

Foward Sel
Scale-Location

$\sqrt{|\text{Standardized residuals}|}$

2.0
0.0

10.5   11.5   12.5   13.5

Fitted values

Foward Sel
Residuals vs Leverage

Standardized residuals

6
0
-8

Cook's distance

0.0   0.4   0.8

Leverage

```r
plot(c_o, main='Backwards El')

## Warning: not plotting observations with leverage one:
##    121, 272, 825, 1003, 1011, 1187, 1230, 1275, 1320, 1369

## Warning: not plotting observations with leverage one:
##    121, 272, 825, 1003, 1011, 1187, 1230, 1275, 1320, 1369

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

Backwards El — Residuals vs Fitted

Backwards El — Normal Q-Q

Backwards El — Scale-Location

Backwards El — Residuals vs Leverage
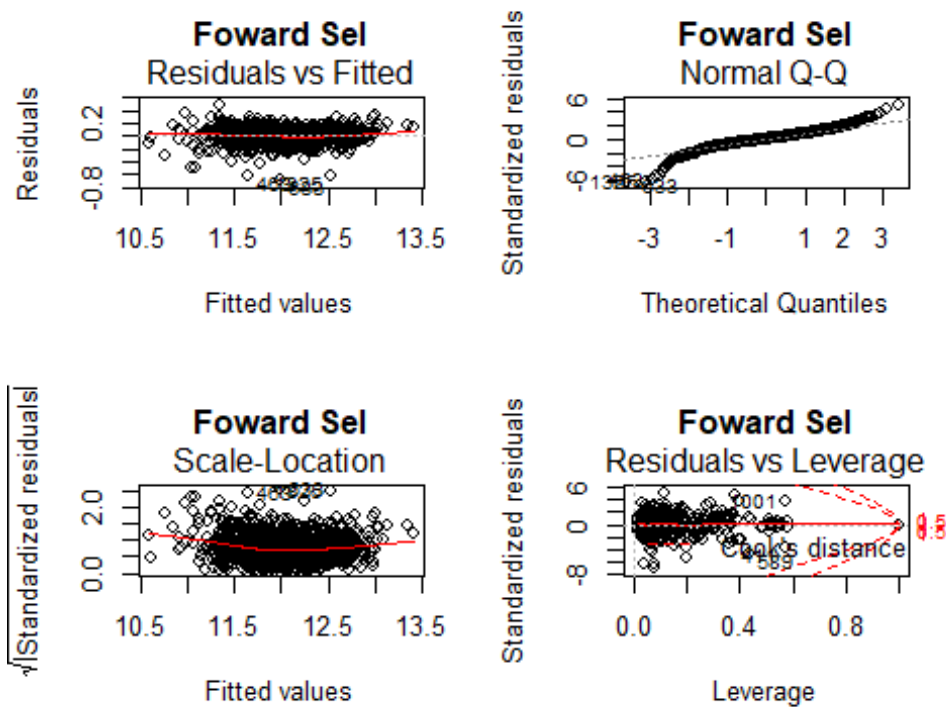
```r
plot(a_o,which=c(4))
plot(b_o,which=c(4))
plot(c_o,which=c(4))

#pdtrain<-predict(a,newdata=H_Lo[!is.na(H_Lo$SalePrice),])
pd_Log_LM_All_o<-predict(a_o,newdata=H_Lo[is.na(H_Lo$SalePrice),])

## Warning in predict.lm(a_o, newdata = H_Lo[is.na(H_Lo$SalePrice), ]):
## prediction from a rank-deficient fit may be misleading

pd_Log_LM_f2_o<-predict(b_o,newdata=H_Lo[is.na(H_Lo$SalePrice),])
pd_Log_LM_f3_o<-predict(c_o,newdata=H_Lo[is.na(H_Lo$SalePrice),])
#trainSP<-H_Lo$SalePrice[!is.na(H_Lo$SalePrice)]


#plot(exp(trainSP),exp(pdtrain), col='blue', ylab='Prediction')
#points(actual,exp(pdtest), col='red')
#abline(b=1,a=0)
#RMSE(exp(trainSP),exp(pdtrain))
#RMSE(actual,exp(pdtest))

print("RMSLE for All Var - Log")

## [1] "RMSLE for All Var - Log"

RMSE(pd_Log_LM_All_o,log(actual))
```

```
## [1] 0.1340808

RMSE(pd_Log_LM_f2_o,log(actual))

## [1] 0.1311971

RMSE(pd_Log_LM_f3_o,log(actual))

## [1] 0.1305666
```
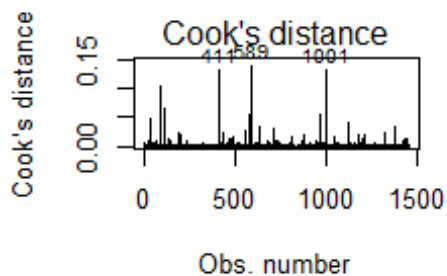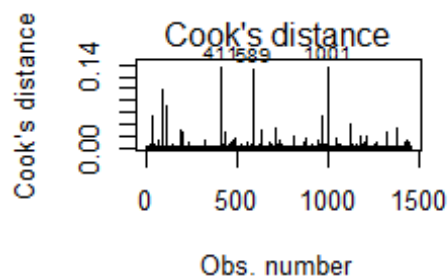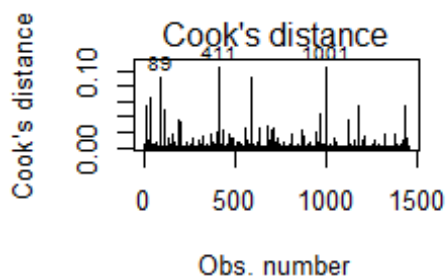

Lasso

Applied - Only done on all variables for Lasso/Ridge/Elasticnet

```
tc<-trainControl(method="cv", number=10)
#rainset -DummyVariables
H_DummyLo_Train<-  as.data.frame(model.matrix(~.-1,data=H_Lo[1:1460,-1])) # R
emove ID

#TestSet   -DummyVariables
H_DummyLo_Test<-   as.data.frame(model.matrix(~.-1,data=H_Lo[1461:2919,-c(1,78
)])) # RemoveID, salePrice
   #as.data.frame

dim(H_DummyLo_Train)

## [1] 1460  216

set.seed(12334)

myLasso1Log <-train(SalePrice~.,
```

```
              data = H_DummyLo_Train
              ,method='glmnet',
              tuneGrid=expand.grid(alpha=1,lambda=seq(0.0001,0.01,length=5
1)), trControl=tc)
myLasso2Log <-train(SalePrice~., data = H_DummyLo_Train[-c(524,1299),]
  ,method='glmnet',tuneGrid=expand.grid(alpha=1,lambda=seq(0.0001,0.01,length
=51)), trControl=tc)
#myLasso3 <-train(f3, data = H_Dummy_Train
#  ,method='glmnet',tuneGrid=expand.grid(alpha=1,lambda=seq(500,1500,length=5
0)), trControl=tc)
#myLasso4 <-train(f4, data = H_Dummy_Train
 #  ,method='glmnet',tuneGrid=expand.grid(alpha=1,lambda=seq(500,1500,length=5
0)), trControl=tc)

set.seed(12334)

plot(myLasso1Log, main='ALL')
```
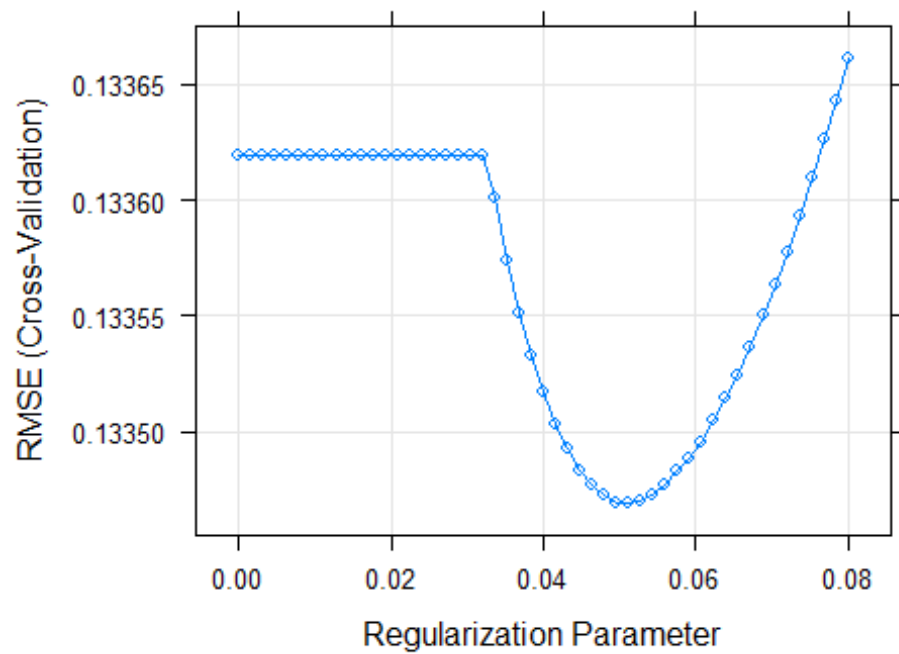

ALL
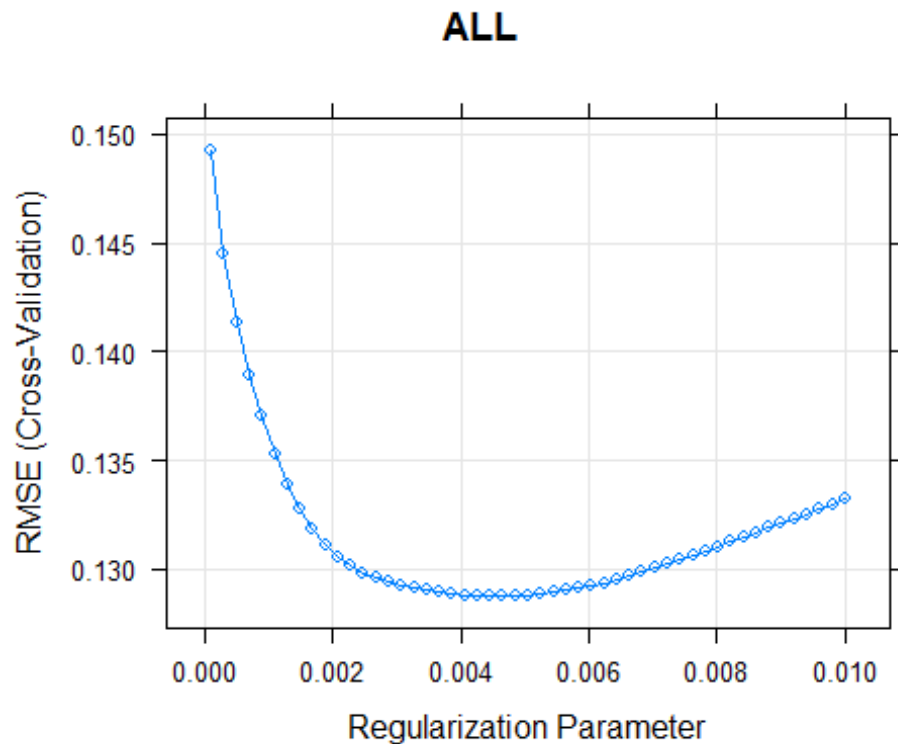
```
plot(myLasso2Log, main='Outlier Removed')
```

Outlier Removed
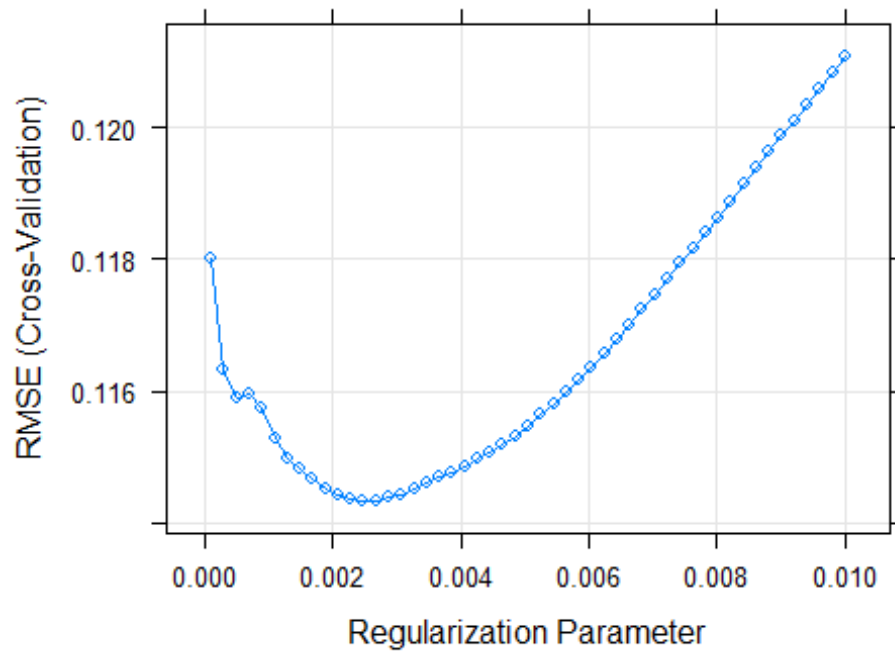

```
plot(varImp(myLasso1Log), Scale=F, top=30)
```
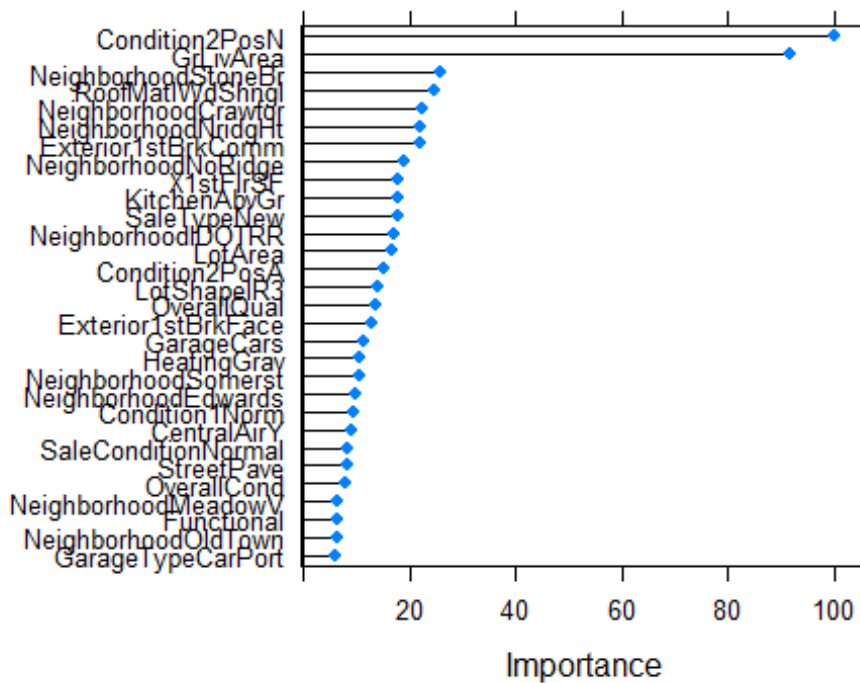


```
plot(varImp(myLasso2Log), Scale=F ,top=30)
```
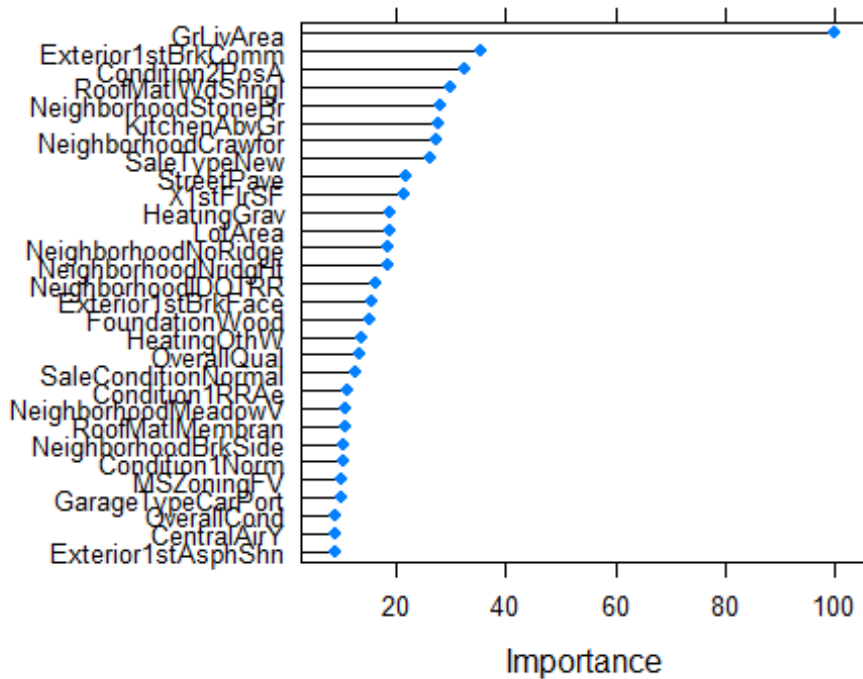
```
varImp(myLasso1Log)

## glmnet variable importance
##
##    only 20 most important variables shown (out of 215)
##
##                      Overall
## Condition2PosN       100.00
## GrLivArea             91.43
## NeighborhoodStoneBr   25.92
## RoofMatlWdShngl       24.90
## NeighborhoodCrawfor   22.51
## NeighborhoodNridgHt   21.91
## Exterior1stBrkComm    21.88
## NeighborhoodNoRidge   18.85
## X1stFlrSF             17.91
## KitchenAbvGr          17.90
## SaleTypeNew           17.76
## NeighborhoodIDOTRR    17.18
## LotArea               16.82
## Condition2PosA        15.04
## LotShapeIR3           14.22
## OverallQual           13.63
## Exterior1stBrkFace    12.97
## GarageCars            11.36
## HeatingGrav           10.76
## NeighborhoodSomerst   10.68
```

```
varImp(myLasso2Log)

## glmnet variable importance
##
##    only 20 most important variables shown (out of 215)
##
##                      Overall
## GrLivArea             100.00
## Exterior1stBrkComm     35.56
## Condition2PosA         32.57
## RoofMatlWdShngl        29.77
## NeighborhoodStoneBr    28.16
## KitchenAbvGr           27.61
## NeighborhoodCrawfor    27.29
## SaleTypeNew            26.05
## StreetPave             21.90
## X1stFlrSF              21.38
## HeatingGrav            18.95
## LotArea                18.90
## NeighborhoodNoRidge    18.59
## NeighborhoodNridgHt    18.43
## NeighborhoodIDOTRR     16.19
## Exterior1stBrkFace     15.46
## FoundationWood         15.10
## HeatingOthW            13.54
## OverallQual            13.37
## SaleConditionNormal    12.58

myLasso1Log$bestTune

##    alpha    lambda
## 23     1 0.004456

myLasso2Log$bestTune

##    alpha    lambda
## 13     1 0.002476

#myLasso3$bestTune
#myLasso4$bestTune

plot(myLasso1Log$finalModel, label=T)
```
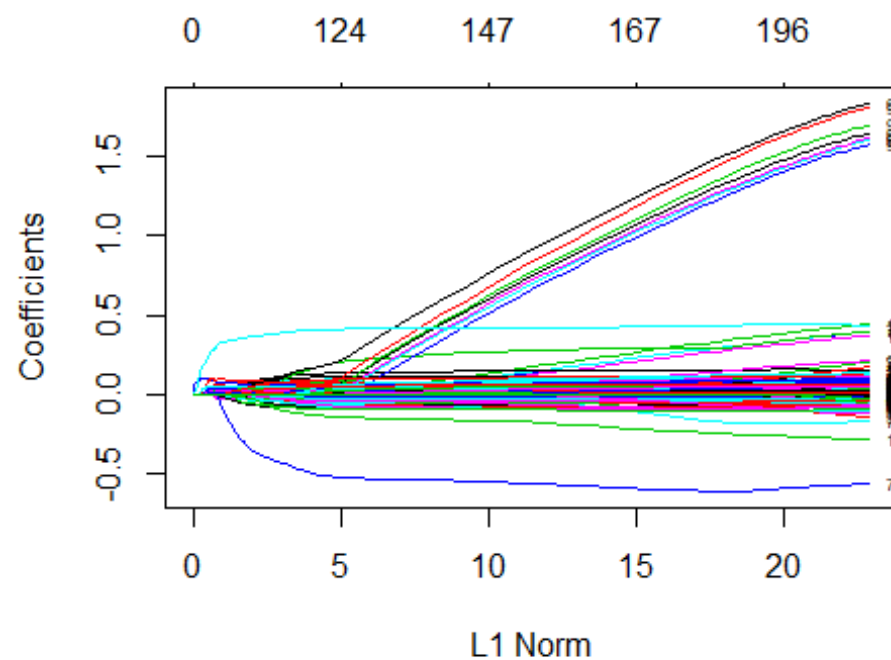
```
plot(myLasso2Log$finalModel ,label=T)
```
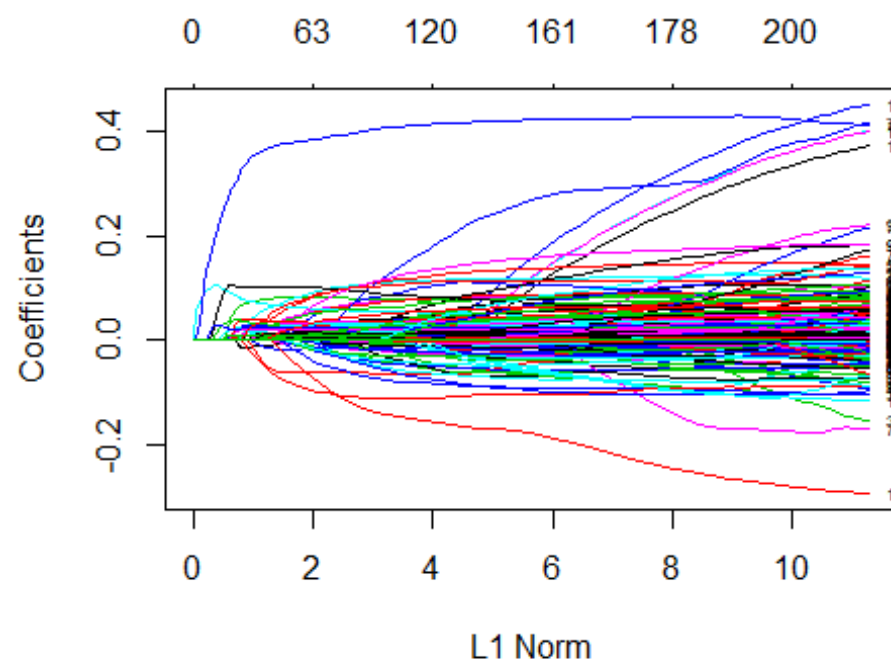
Predict and Measure RMSE of Log

```
PdmyLasso1Log<-predict(myLasso1Log, newdata=H_DummyLo_Test)
PdmyLasso2Log<-predict(myLasso2Log, newdata=H_DummyLo_Test)


RMSE(PdmyLasso1Log,log(actual))

## [1] 0.1269526

RMSE(PdmyLasso2Log,log(actual))

## [1] 0.1242114
```

Ridge of Log
```
set.seed(12334)
myRidge1Log <-train(SalePrice~.,
                 data = H_DummyLo_Train
                 ,method='glmnet',
                 tuneGrid=expand.grid(alpha=0,lambda=seq(0.00001,0.08,length=
51)), trControl=tc)
set.seed(12334)
myRidge2Log <-train(SalePrice~., data = H_DummyLo_Train[-c(524,1299),]
  ,method='glmnet',tuneGrid=expand.grid(alpha=0,lambda=seq(0.0001,0.08,length
=51)), trControl=tc) # without outlier

par(mfrow=c(1,2))
plot(myRidge1Log, main='ALL')
```

ALL


```r
plot(myRidge2Log, main='Outlier Removed')
```
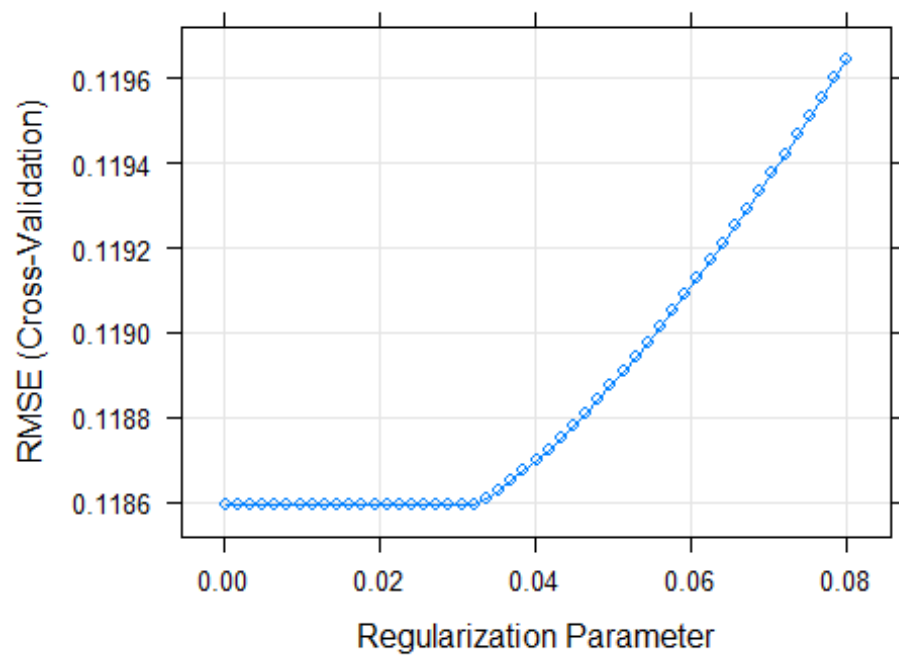
Outlier Removed


```r
plot(varImp(myRidge1Log), Scale=F, top=30,main='ALL')
```

ALL
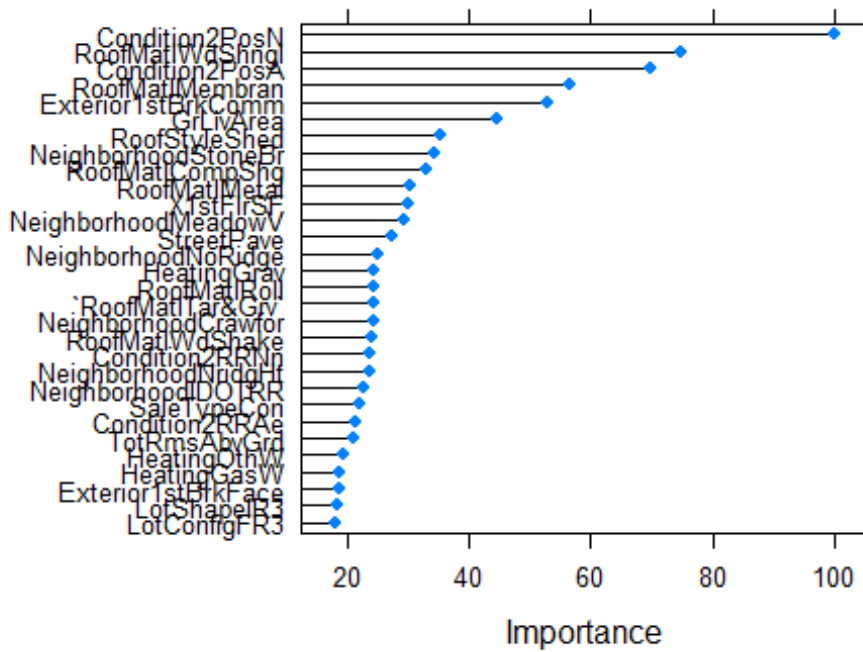

```
plot(varImp(myRidge2Log), Scale=F ,top=30, main='Outlier Removed')
```

Outlier Removed
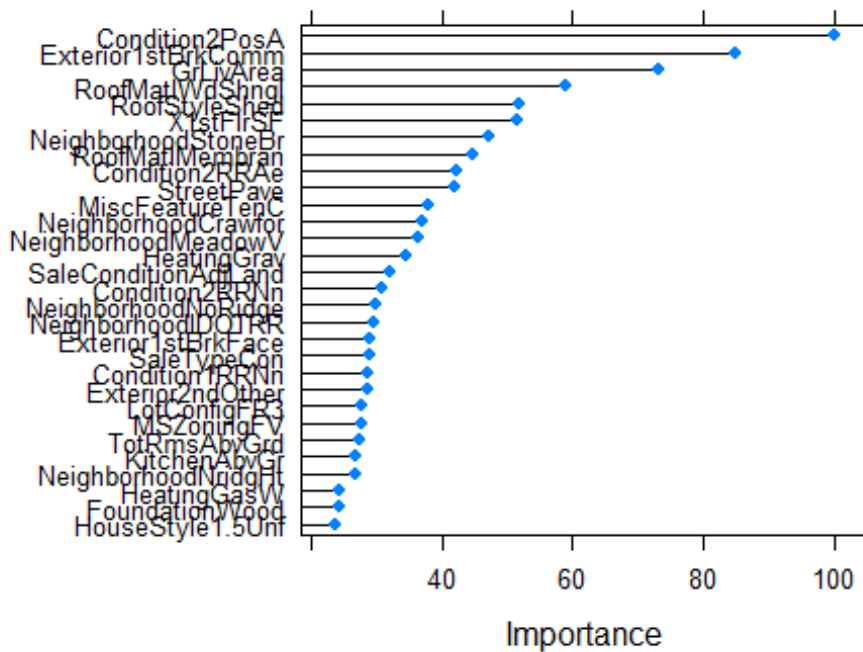

```
varImp(myRidge1Log,main='ALL')
```

```
## glmnet variable importance
##
##    only 20 most important variables shown (out of 215)
##
##                        Overall
## Condition2PosN         100.00
## RoofMatlWdShngl         74.63
## Condition2PosA          69.64
## RoofMatlMembran         56.52
## Exterior1stBrkComm      52.94
## GrLivArea               44.45
## RoofStyleShed           35.25
## NeighborhoodStoneBr     34.26
## RoofMatlCompShg         33.06
## RoofMatlMetal           30.21
## X1stFlrSF               30.10
## NeighborhoodMeadowV     29.23
## StreetPave              27.52
## NeighborhoodNoRidge     25.13
## HeatingGrav             24.53
## RoofMatlRoll            24.53
## `RoofMatlTar&Grv`       24.28
## NeighborhoodCrawfor     24.27
## RoofMatlWdShake         23.93
## Condition2RRNn          23.83
```

`varImp(myRidge2Log,main='Outlier Removed')`

```
## glmnet variable importance
##
##    only 20 most important variables shown (out of 215)
##
##                        Overall
## Condition2PosA         100.00
## Exterior1stBrkComm      84.96
## GrLivArea               73.26
## RoofMatlWdShngl         58.77
## RoofStyleShed           51.75
## X1stFlrSF               51.63
## NeighborhoodStoneBr     47.27
## RoofMatlMembran         44.81
## Condition2RRAe          42.14
## StreetPave              42.07
## MiscFeatureTenC         38.04
## NeighborhoodCrawfor     37.03
## NeighborhoodMeadowV     36.24
## HeatingGrav             34.49
## SaleConditionAdjLand    31.94
## Condition2RRNn          30.73
## NeighborhoodNoRidge     29.89
```

```
## NeighborhoodIDOTRR      29.52
## Exterior1stBrkFace      29.05
## SaleTypeCon             28.96
```

myRidge1Log$bestTune

```
##    alpha     lambda
## 33     0 0.0512036
```

myRidge2Log$bestTune

```
##    alpha  lambda
## 21     0 0.03206
```

```
plot(myRidge1Log$finalModel,main='ALL')
plot(myRidge2Log$finalModel,main='Outlier Removed')
```
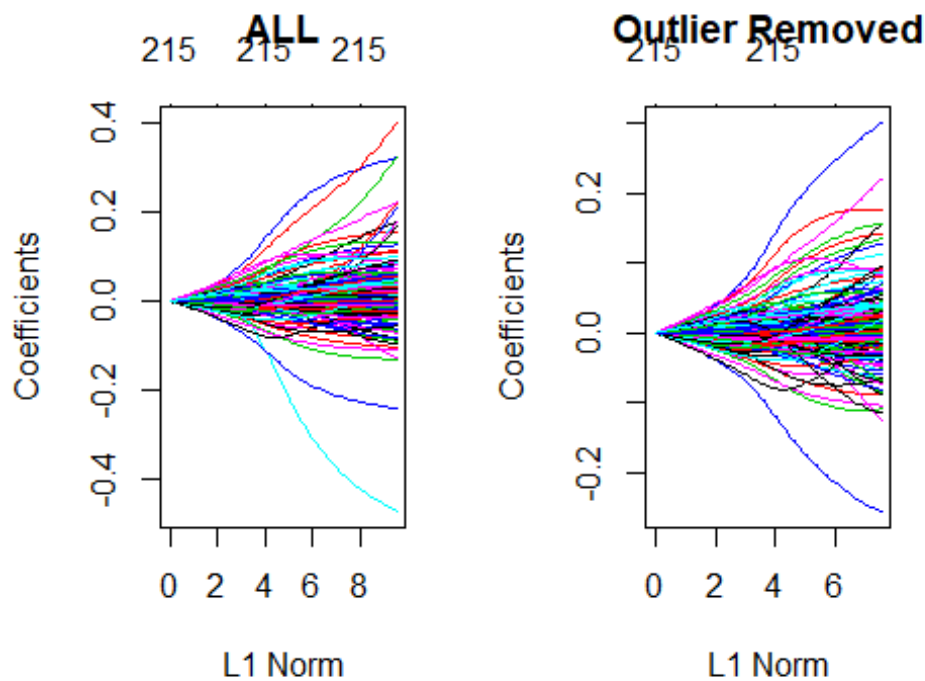


```
PdmyRidge1Log<-predict(myRidge1Log, newdata=H_DummyLo_Test)
PdmyRidge2Log<-predict(myRidge2Log, newdata=H_DummyLo_Test)
```

```
RMSE(PdmyRidge1Log,log(actual))
```

```
## [1] 0.1317548
```

```
RMSE(PdmyRidge2Log,log(actual))
```

```
## [1] 0.1287369
```

```r
set.seed(12334)
myEla1Log<-train(SalePrice~.,data = H_DummyLo_Train,method='glmnet',tuneGrid=
expand.grid(alpha=seq(0, 1,length=11),lambda=seq(0.0001,0.1,length=30)), trCo
ntrol=tc)
set.seed(12334)
myEla2Log<-train(SalePrice~.,data = H_DummyLo_Train[-c(524,1299),],method='gl
mnet',tuneGrid=expand.grid(alpha=seq(0, 1,length=11),lambda=seq(0.0001,0.1,le
ngth=30)), trControl=tc)

par(mfrow=c(1,2))
plot(myEla1Log, main='ALL')
```
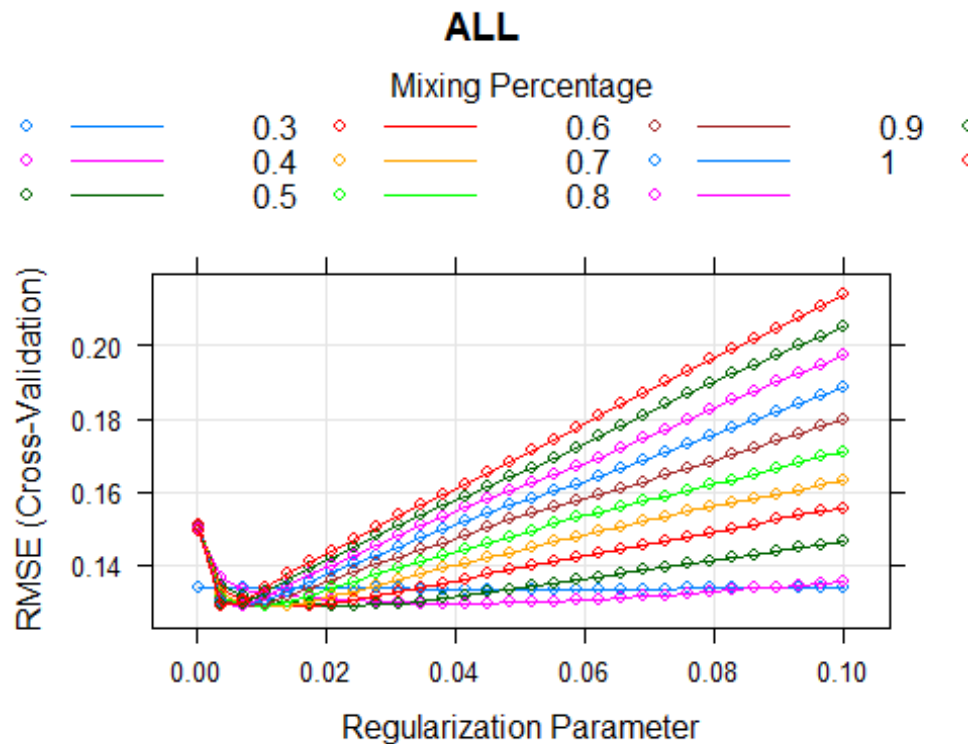


```r
plot(myEla2Log, main='Outlier Removed')
```

Outlier Removed

Mixing Percentage

	0.3		0.6		0.9
	0.4		0.7		1
	0.5		0.8		
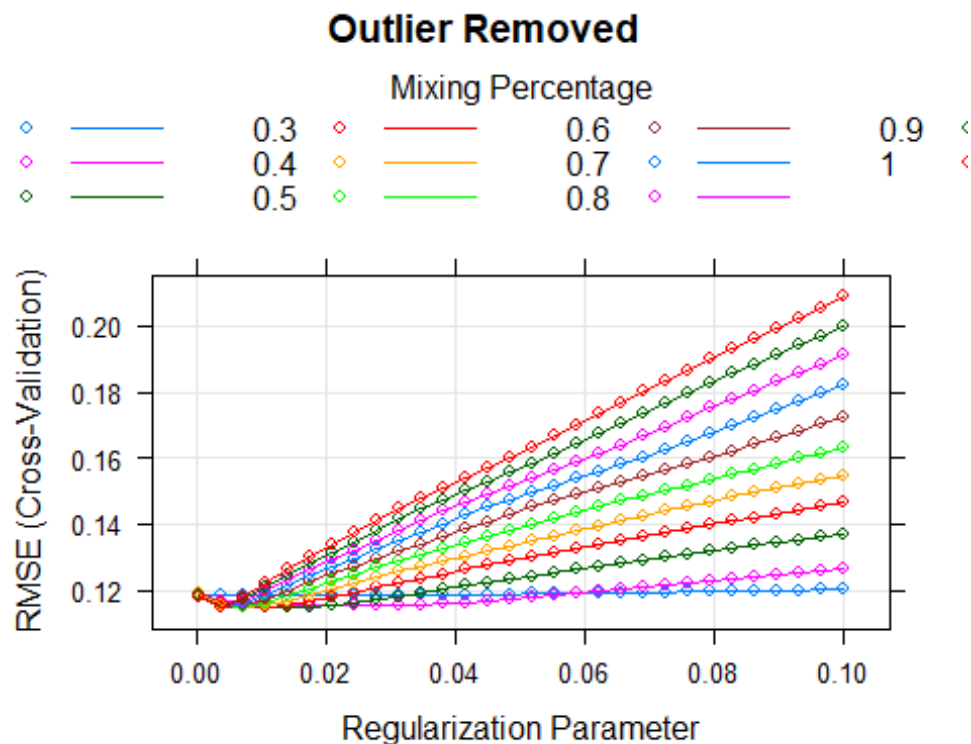

```
#plot(varImp(myEla1Log),top=25)
#varImp(myEla1Log)
myEla1Log$bestTune
```

```
##      alpha        lambda
## 213    0.7 0.006989655
```

```
myEla2Log$bestTune
```

```
##      alpha        lambda
## 242    0.8 0.003544828
```

```
PdEla1Log<-predict(myEla1Log, newdata=H_DummyLo_Test)
PdEla2Log<-predict(myEla2Log, newdata=H_DummyLo_Test)
RMSE(PdEla1Log, log(actual))
```

```
## [1] 0.1273111
```

```
RMSE(PdEla2Log, log(actual))
```

```
## [1] 0.1243352
```

Origial Lasso/Ridge/Elastic - removing outlier didn't help and actually gave worse results (barely) but once log transformed, it improved the results although little.

Random Forest for Log

Do a quick tuning.

```
set.seed(100)
tuneRF10<-tuneRF(H_Lo[1:1460,-c(1,78)],H_Lo[1:1460,78], mtryStart=16,StepFact
or=1 ,improve=0.0001, Trace=T,Plot=T,ntreeTry=500, doBest = T)

## mtry = 16  OOB error = 0.01840168
## Searching left ...
## mtry = 8      OOB error = 0.01917974
## -0.04228188 1e-04
## Searching right ...
## mtry = 32     OOB error = 0.01822353
## 0.009681035 1e-04
## mtry = 64     OOB error = 0.01908944
## -0.04751592 1e-04
```
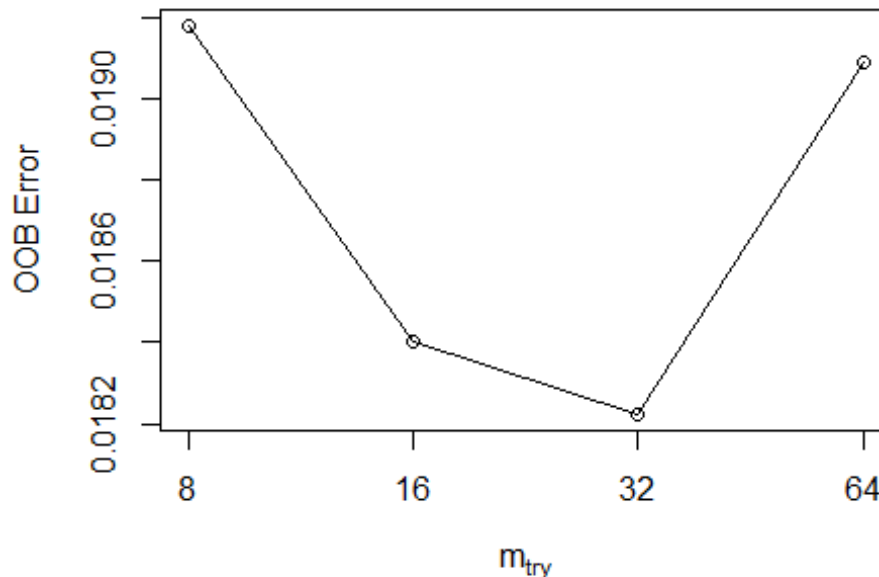

Best ntree=32. Apply rf1=randomForest(SalePrice~.,data=Train, ntree=500,mtry=32)

```
RFLog=randomForest(SalePrice~.,data=H_Lo[TR,-1],  mtry = 32,ntree=500)
RFLog1<-randomForest(SalePrice~.,data=H_Lo[TR_RmvOut,-1],  mtry = 32,ntree=50
0)
PdRFLog<-predict(RFLog,newdata= H_Lo[1461:2919,-78])
PdRFLog1<-predict(RFLog1,newdata= H_Lo[1461:2919,-78])
RMSE(PdRFLog,log(actual))

## [1] 0.1397059

RMSE(PdRFLog1,log(actual))

## [1] 0.1407842
```

SVM Tuning - some are texted out due to running time

```
#SV1logtune1<-tune.svm(SalePrice~.,data=H_Lo[TR,-1],gamma=0.1*c(0.0001, 0.000
5,0.001,0.005, 0.01, 0.15,0.2, 0.3,0.5) ,cost=1:10)

#print(SV1logtune1)
#plot(SV1logtune1)
#sqrt(SV1logtune1$best.performance)
```

Parameter tuning of 'svm':

- sampling method: 10-fold cross validation

- best parameters: gamma cost 0.001 4

- best performance: 0.01571751

Need more fining tuning up to 0.02 - 0.001 4 '

```
#SV1logtune2<-tune.svm(SalePrice~.,data=H_Lo[TR,-1],gamma=0.002*(1:10) ,cost=
1:10)

#print(SV1logtune2)
#plot(SV1logtune2)
#sqrt(SV1logtune2$best.performance)
```

Finally try for 0.0005 to 0.005 and cost between 1 and 7 - not done due to running time.

```
#SV1logtune3<-tune.svm(SalePrice~.,data=H_Lo[TR,-1],gamma=0.0005*(1:10) ,cost
=1:7)

#print(SV1logtune3)
#plot(SV1logtune3)
```

Parameter tuning of 'svm':

- sampling method: 10-fold cross validation

- best parameters: gamma cost 0.0035 2

- best performance: 0.01617483

[1] 0.1271803

```
SVLog1<-svm(SalePrice~.,data=H_Lo[TR,-1], gamma= 0.0035, cost=2)
SVLog2<-svm(SalePrice~.,data=H_Lo[TR_RmvOut,-1], gamma= 0.0035, cost=2 )
```

Plot all to check.

```
pdSVLog1<-predict(SVLog1, newdata=H_Lo[1461:2919,-78])
pdSVLog2<-predict(SVLog2, newdata=H_Lo[1461:2919,-78])
RMSE(pdSVLog1,log(actual))
```
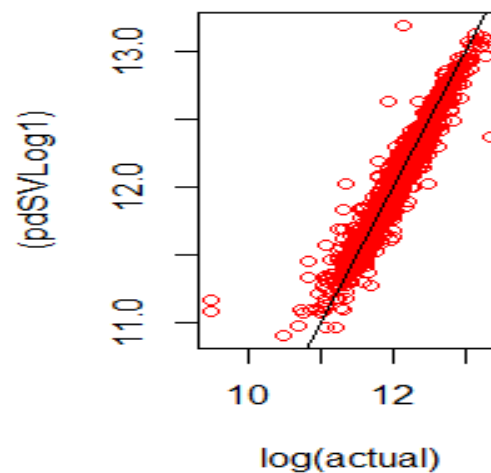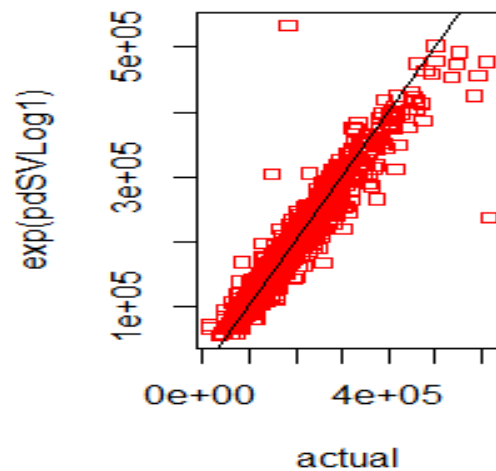
```
## [1] 0.1263945

RMSE(pdSVLog2,log(actual))

## [1] 0.1254436

par(mfrow=c(1,2))
plot(actual, exp(pdSVLog1), pch=22,col='red')
abline(a=0,b=1)

plot(log(actual),(pdSVLog1),col='red')
abline(a=0,b=1)
```
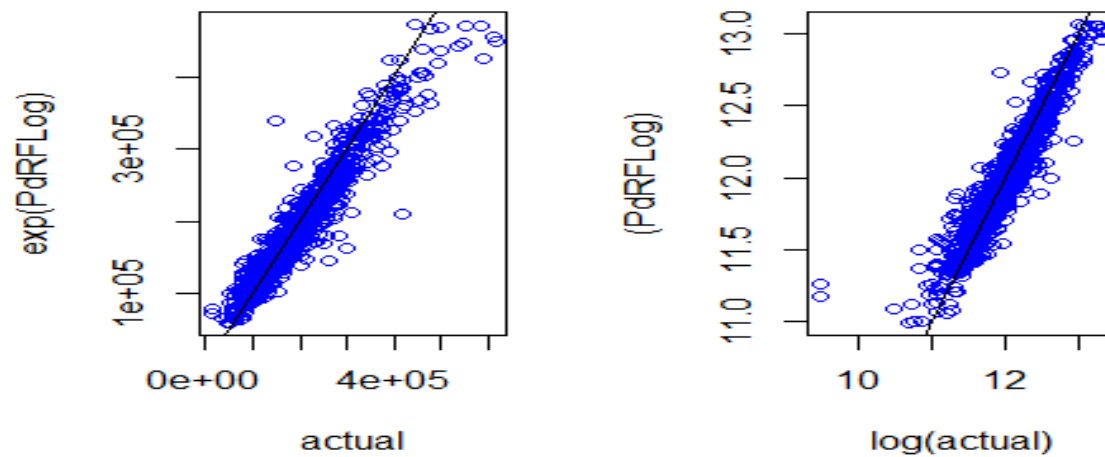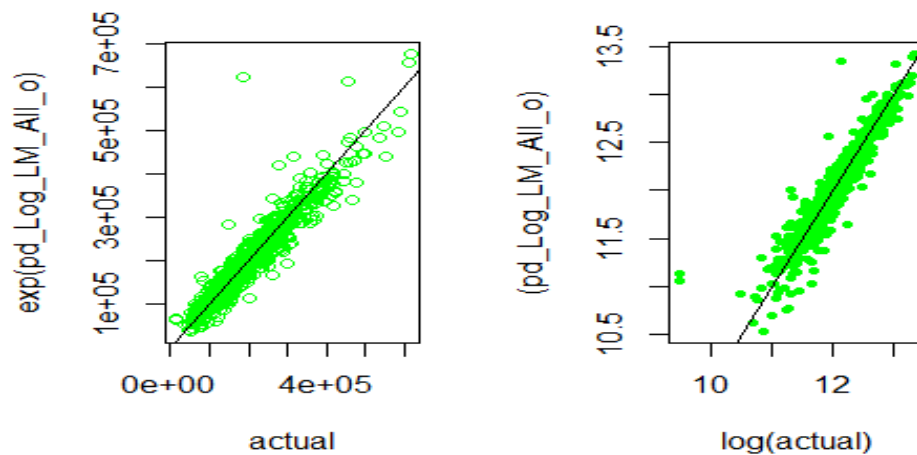

SVM - Logged Plot

```
par(mfrow=c(1,2))
plot(actual, exp(PdRFLog), col='blue')
abline(a=0,b=1)
plot(log(actual),(PdRFLog),col='blue')
abline(a=0,b=1)
```
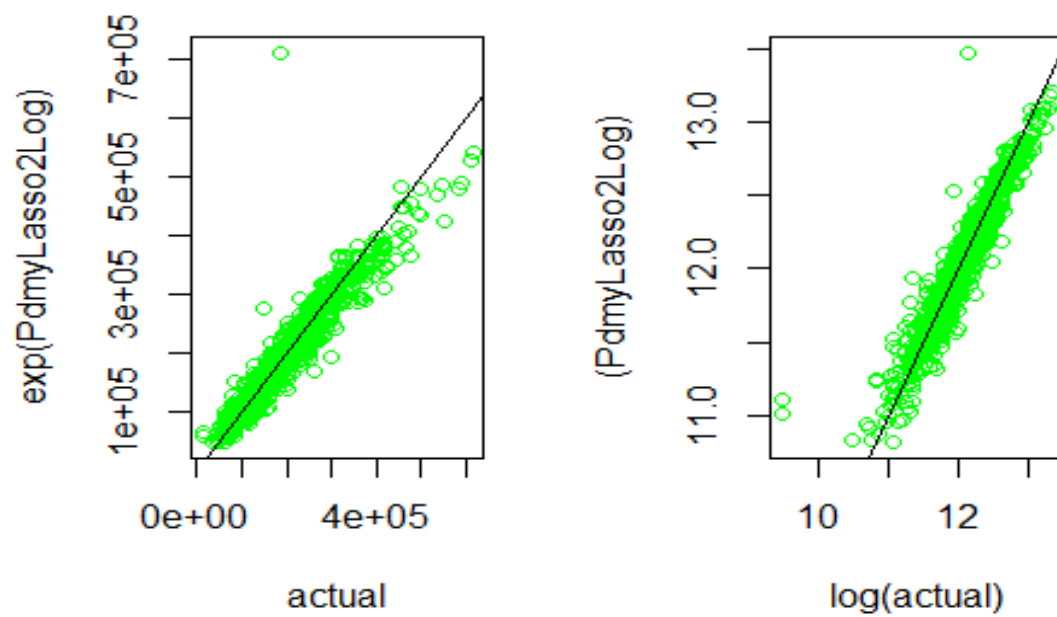
Random Forest – actual/Pred

```r
par(mfrow=c(1,2))
plot(actual, exp(pd_Log_LM_All_o), col='green')
abline(a=0,b=1)
plot(log(actual),(pd_Log_LM_All_o), pch=20,col='green')
abline(a=0,b=1)
```
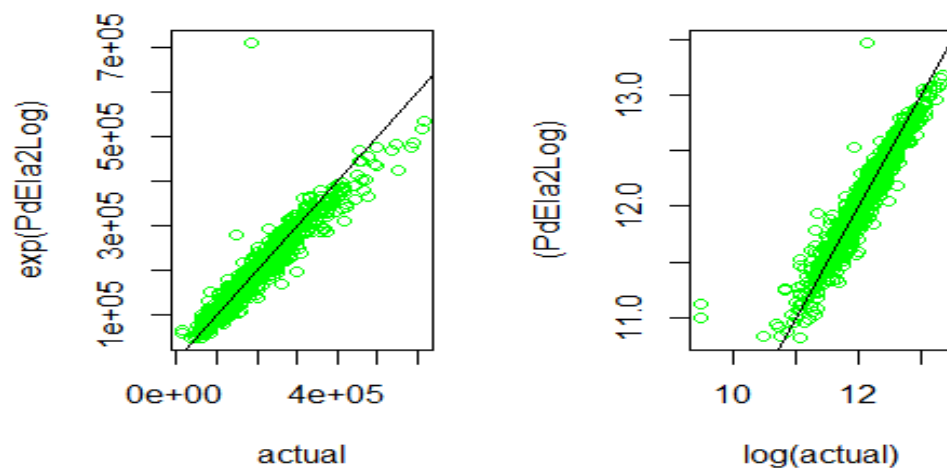

Linear Regression without outliers.

```r
par(mfrow=c(1,2))
plot(actual, exp(PdmyLasso2Log), col='green')
abline(a=0,b=1)
plot(log(actual),(PdmyLasso2Log),col='green')
abline(a=0,b=1)
```

Lasso without outliers.

```r
par(mfrow=c(1,2))
plot(actual, exp(PdEla2Log), col='green')
abline(a=0,b=1)
plot(log(actual),(PdEla2Log),col='green')
abline(a=0,b=1)
```


Elasticnet w/o outliers.