```r
#Requirement 1: Preprocessing and Exploratory Data Analysis
housing <- read.table("AmesHousing.txt", header = TRUE, sep = "\t")

# For simplicity, impute missing values with the median (you may choose another method)
housing$Lot.Frontage[is.na(housing$Lot.Frontage)] <- median(housing$Lot.Frontage, na.rm = TRUE)
housing$Mas.Vnr.Area[is.na(housing$Mas.Vnr.Area)] <- median(housing$Mas.Vnr.Area, na.rm = TRUE)
housing$BsmtFin.SF.1[is.na(housing$BsmtFin.SF.1)] <- median(housing$BsmtFin.SF.1, na.rm = TRUE)

library(ggplot2)
library(corrplot)
```

corrplot 0.92 loaded

```r
library(carData)
library(lattice)
library(car)
library(caret)
library(Matrix)
library(glmnet)
```
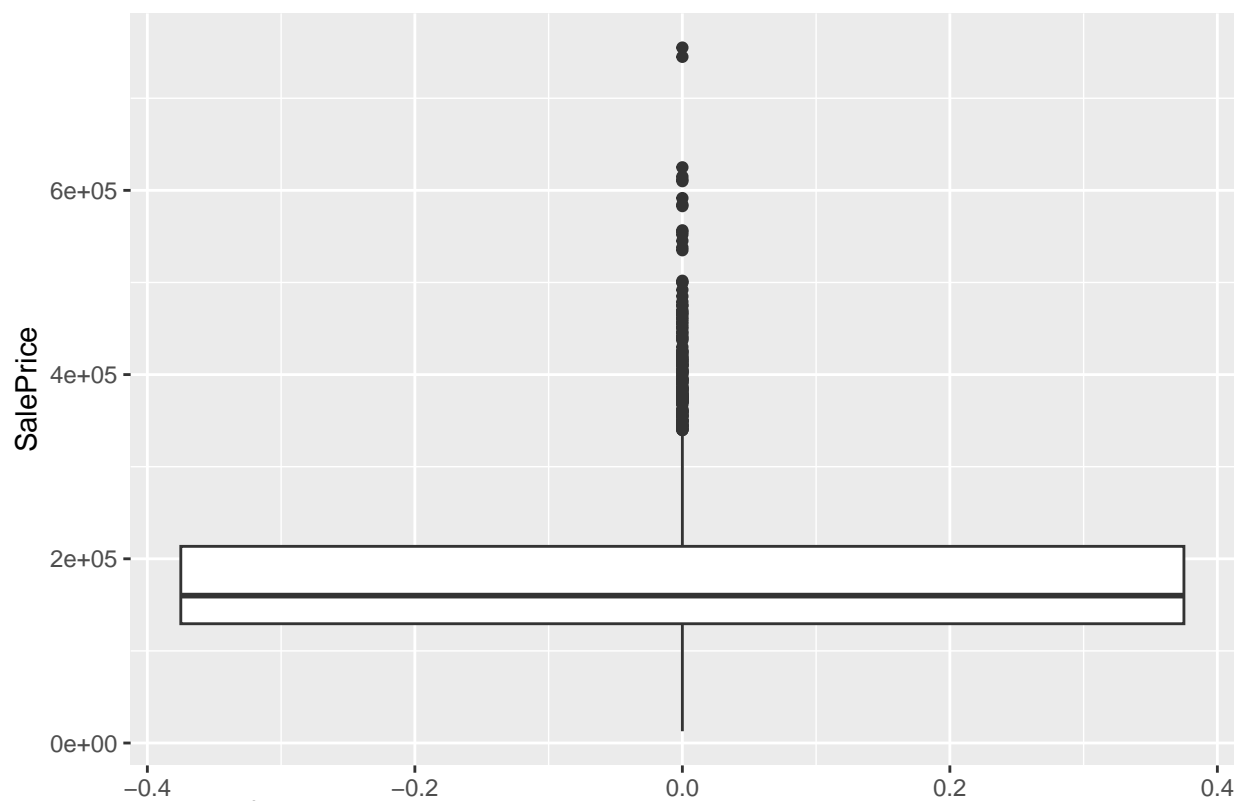
Loaded glmnet 4.1-8

```r
# Boxplot for selected continuous variables
selected_vars <- c("SalePrice","Lot.Frontage", "Lot.Area", "Overall.Qual", "Year.Built", "Overall.Qual"

# Iterate through the selected variables and plot boxplots
for (var in selected_vars) {
  boxplot <- ggplot(housing, aes(y = .data[[var]])) +
    geom_boxplot() +
    ggtitle(paste("Boxplot of", var))
  plot(boxplot)
}
```
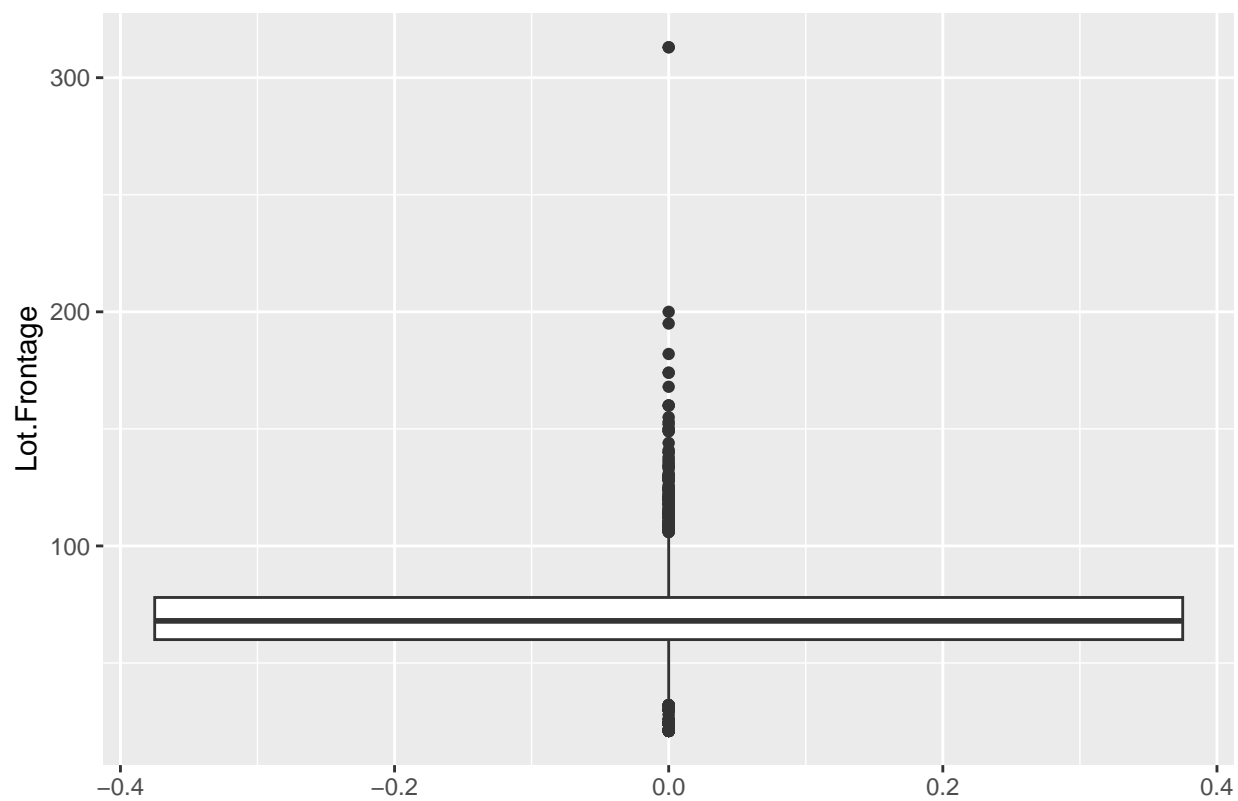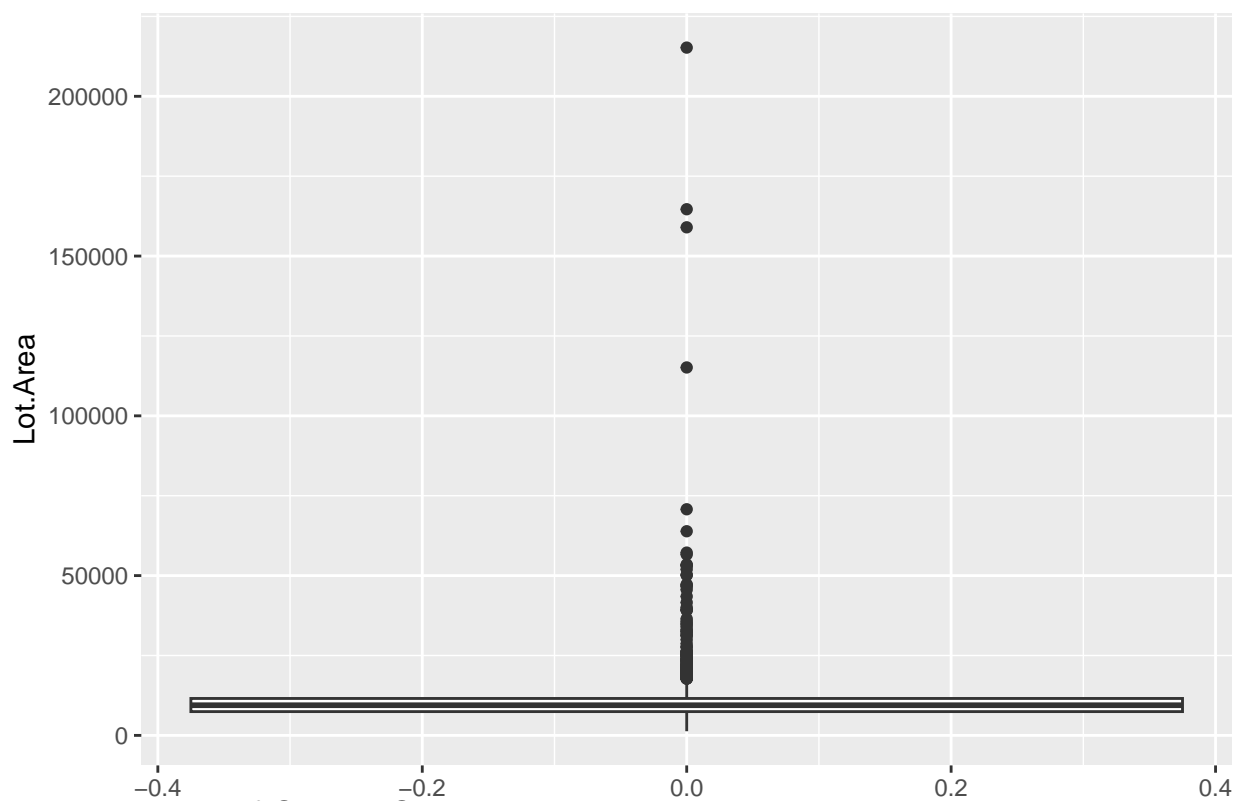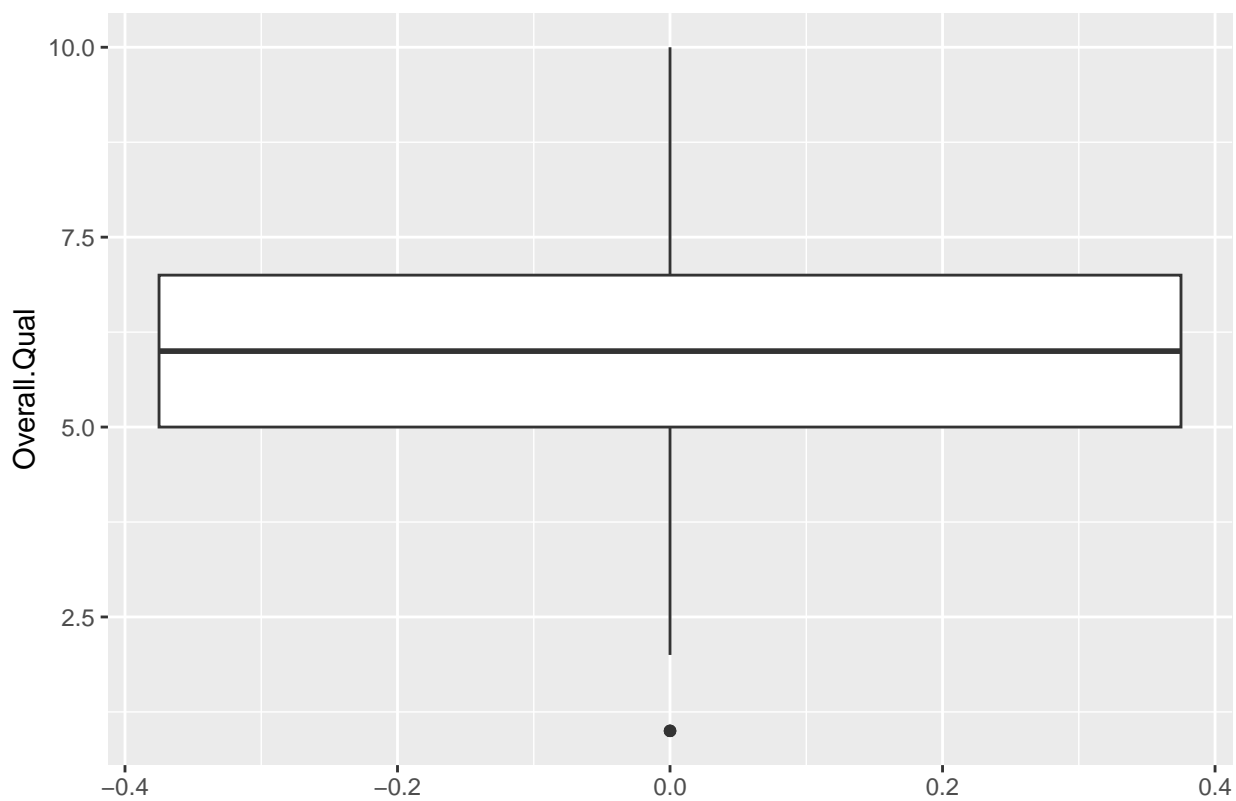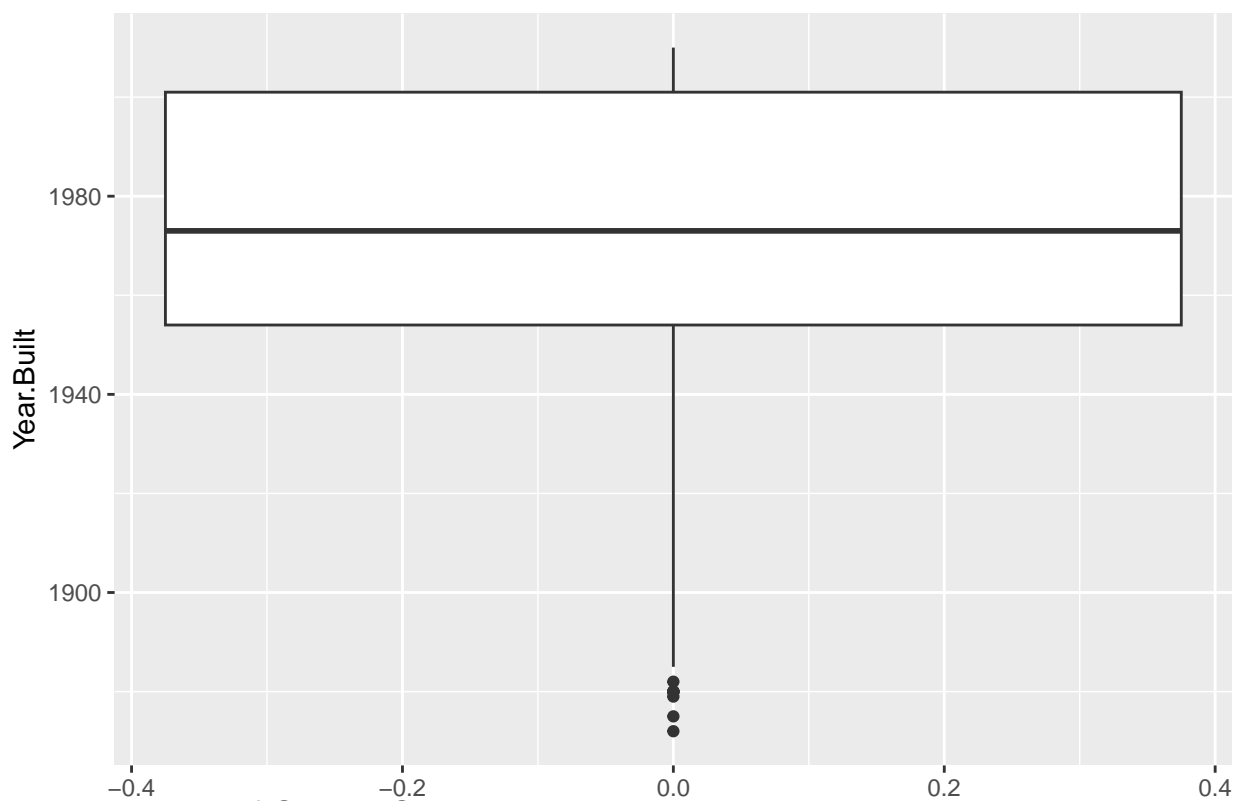
Boxplot of SalePrice

Boxplot of Lot.Frontage

Boxplot of Lot.Area

Boxplot of Overall.Qual

Boxplot of Year.Built

Boxplot of Overall.Qual

Boxplot of Overall.Cond

Boxplot of Year.Built
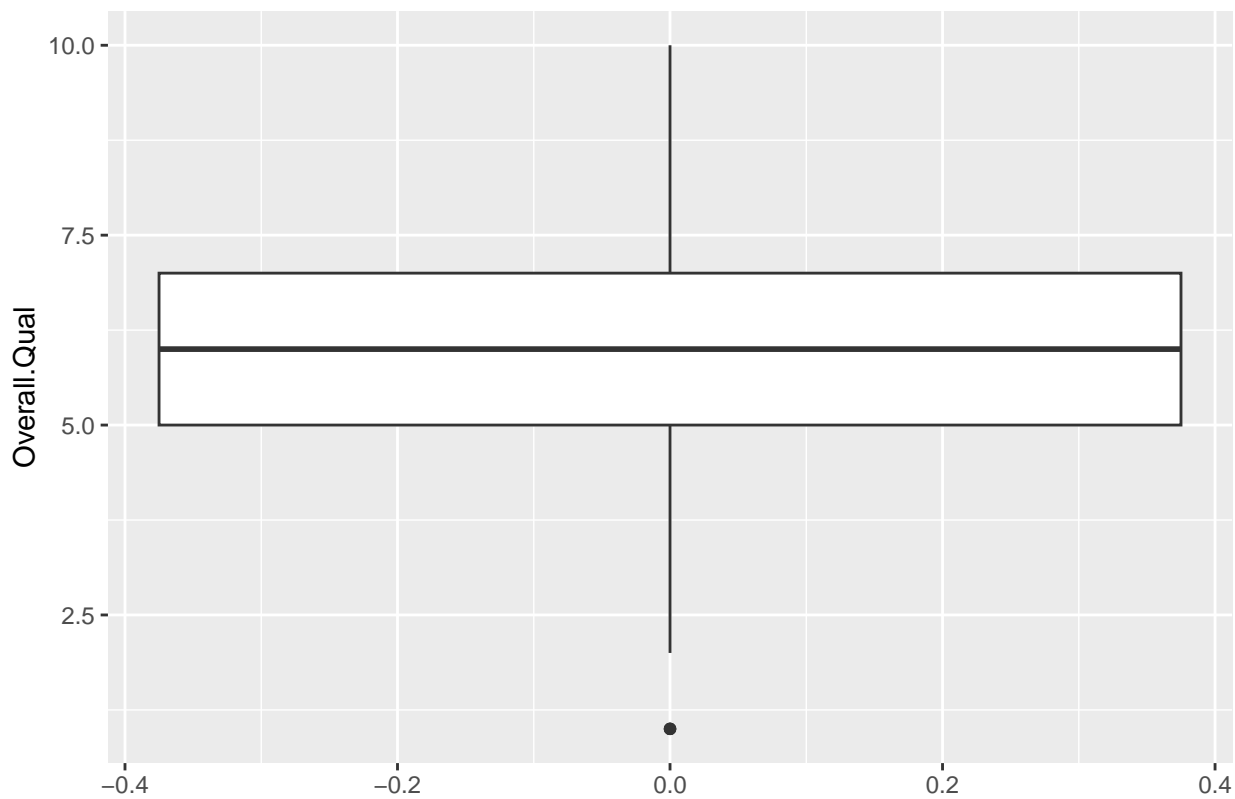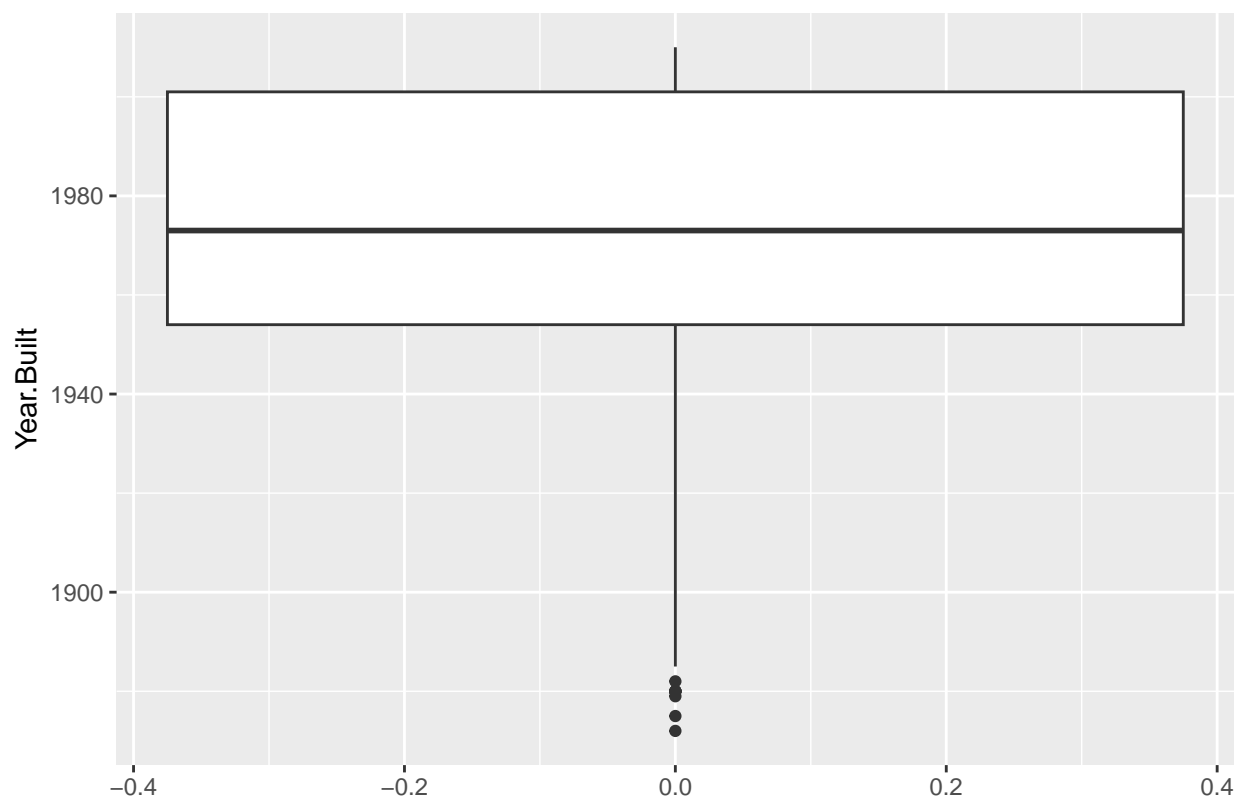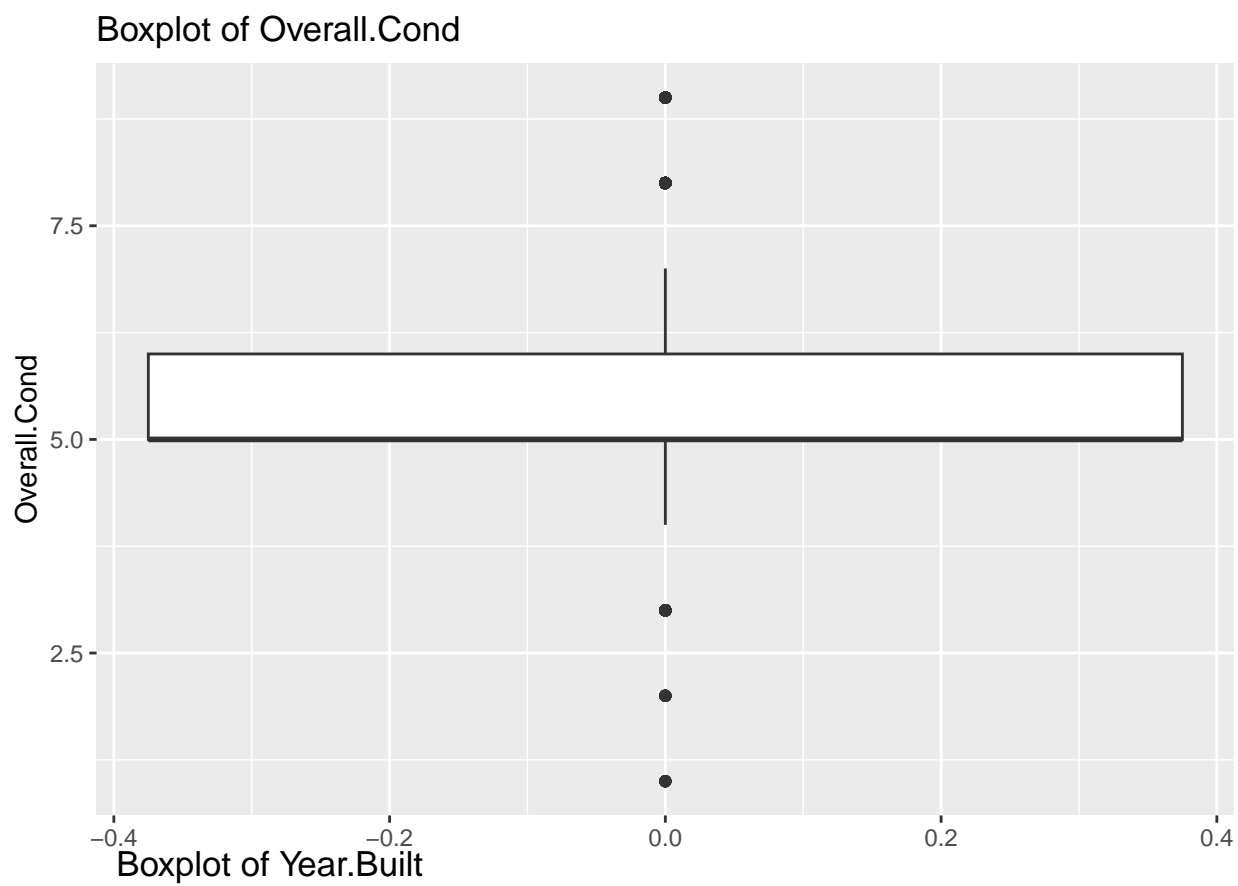
Boxplot of Year.Remod.Add
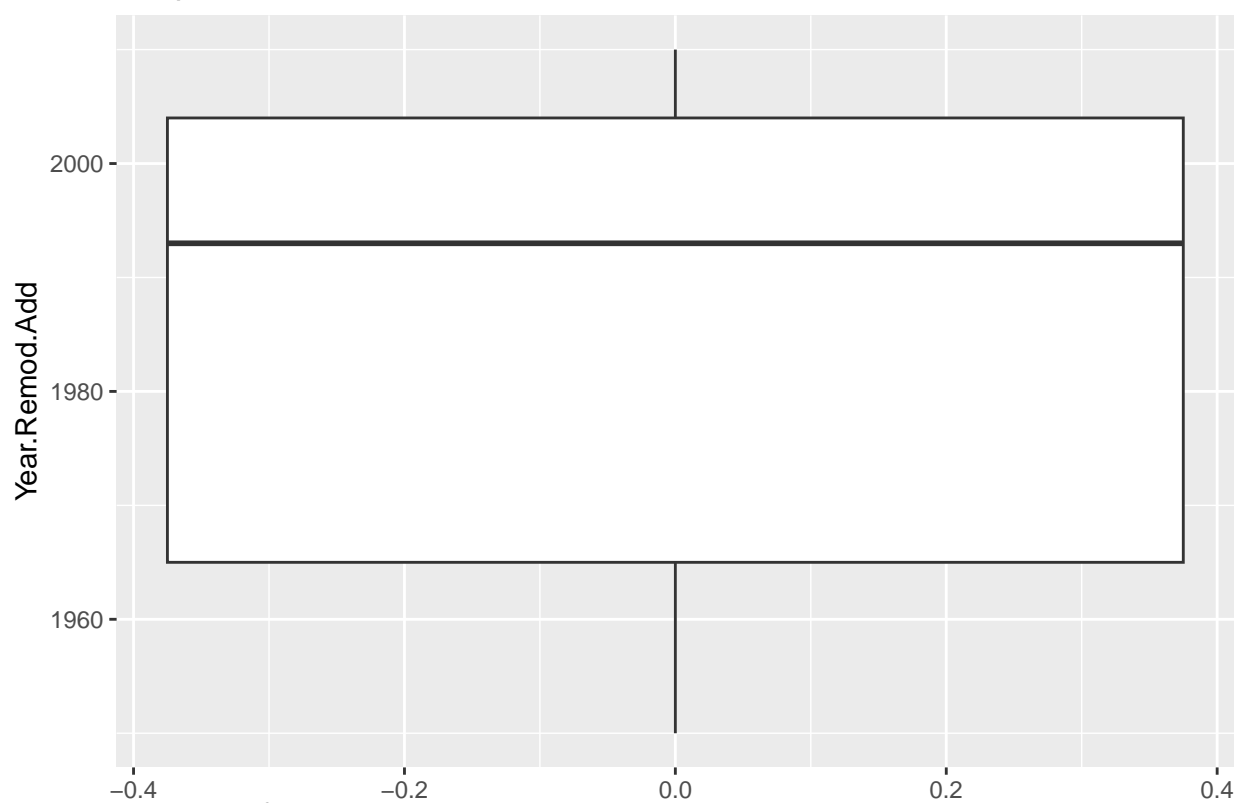
Boxplot of Mas.Vnr.Area
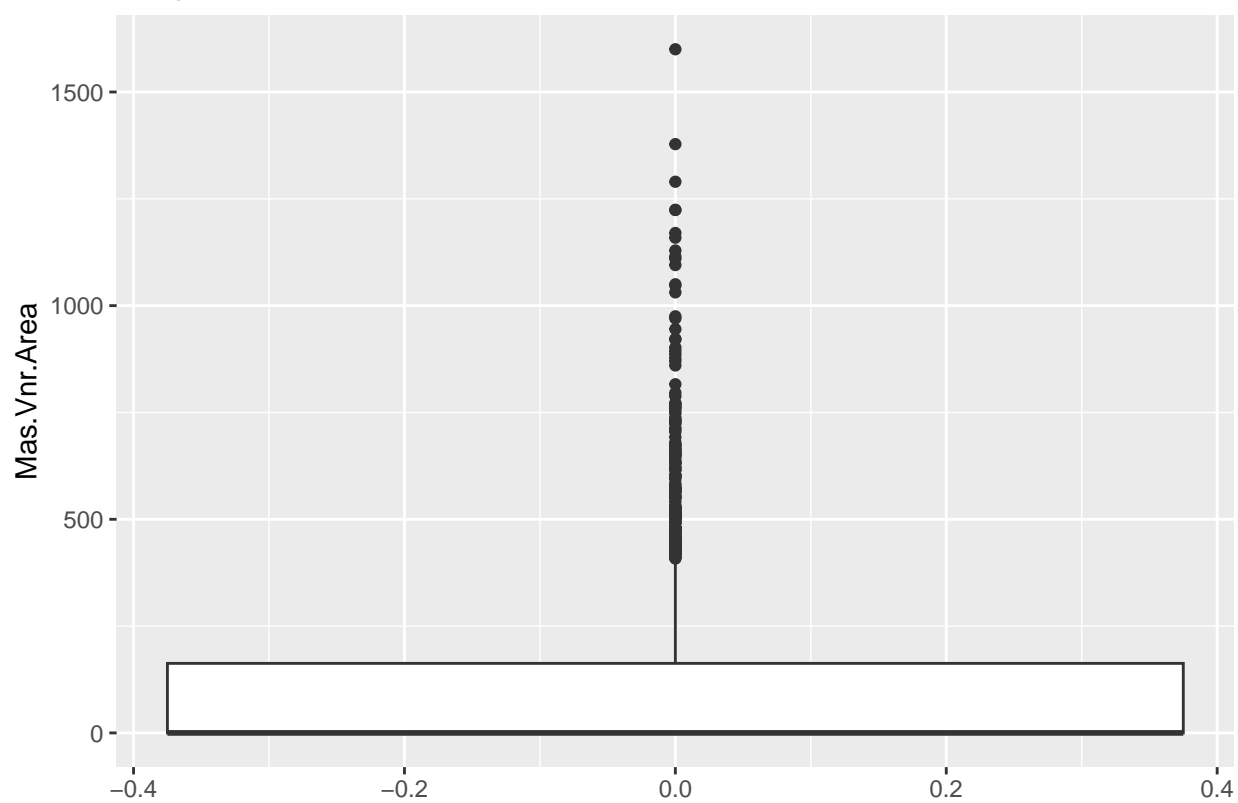
Boxplot of BsmtFin.SF.1
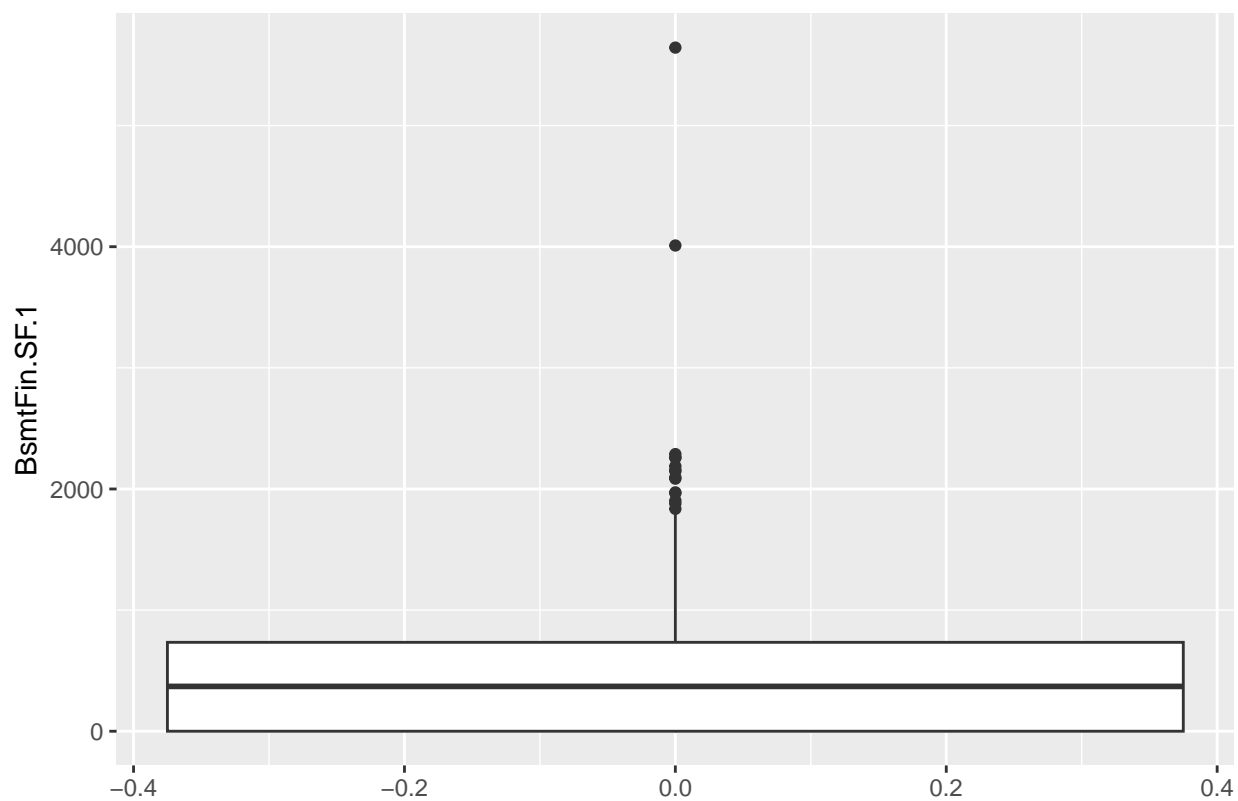

```
## Warning: Removed 1 rows containing non-finite values (`stat_boxplot()`).
```
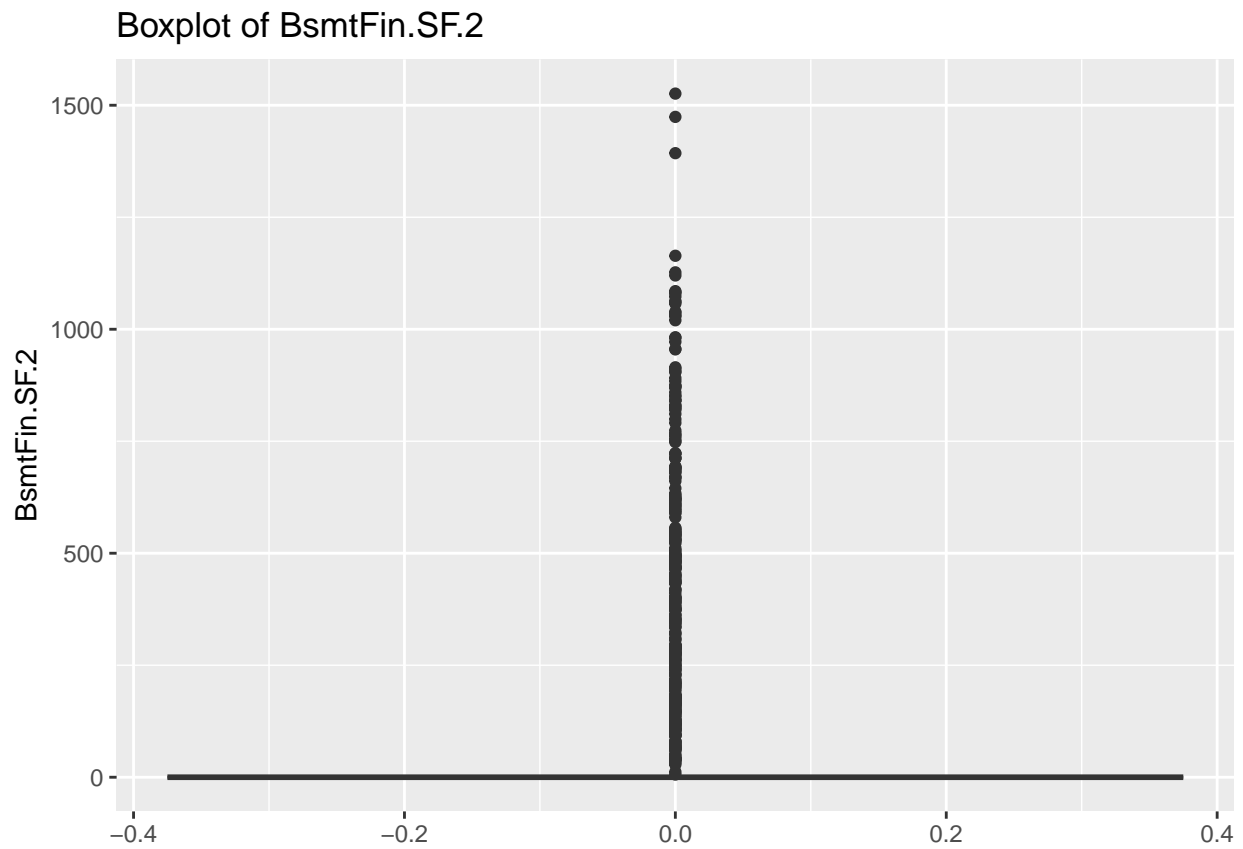
Boxplot of BsmtFin.SF.2

Warning: Removed 1 rows containing non-finite values (`stat_boxplot()`).

Boxplot of Bsmt.Unf.SF
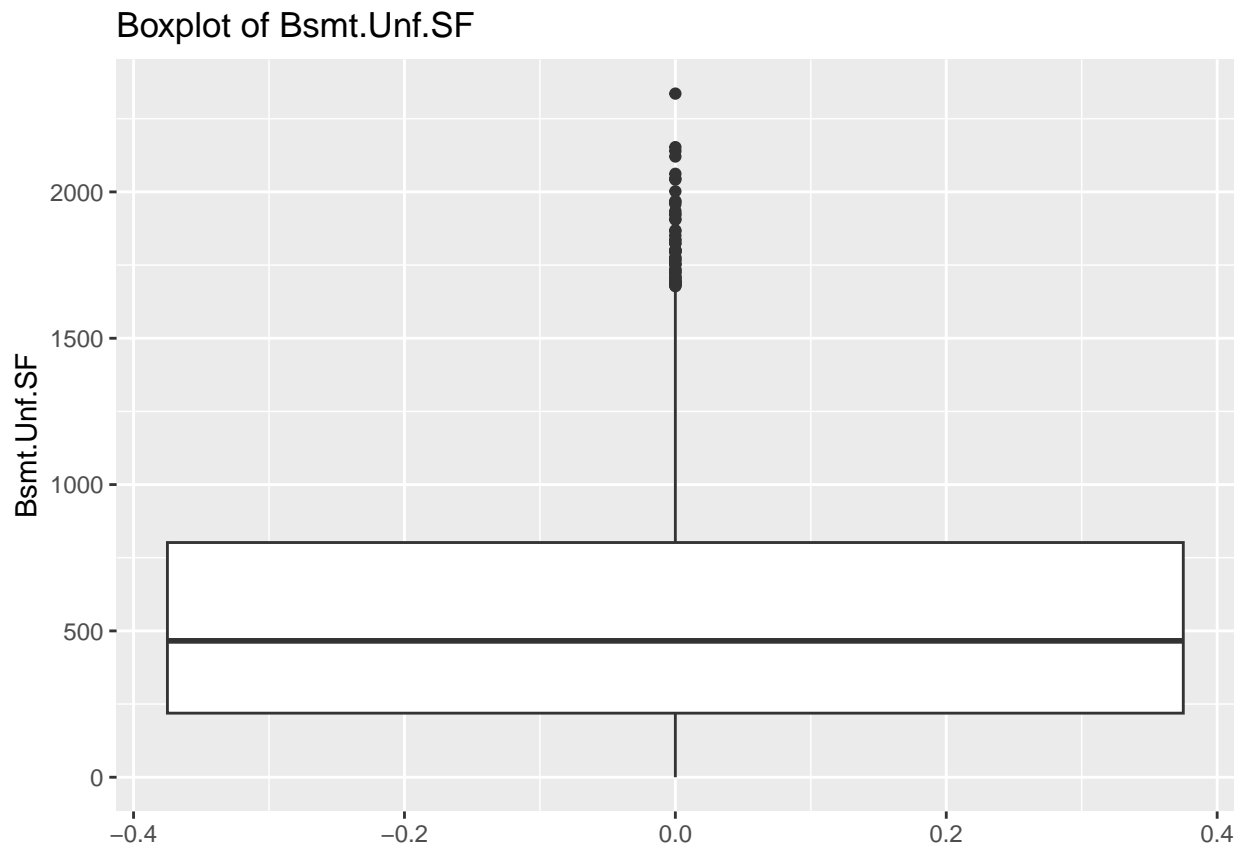
Warning: Removed 1 rows containing non-finite values (`stat_boxplot()`).

Boxplot of Total.Bsmt.SF
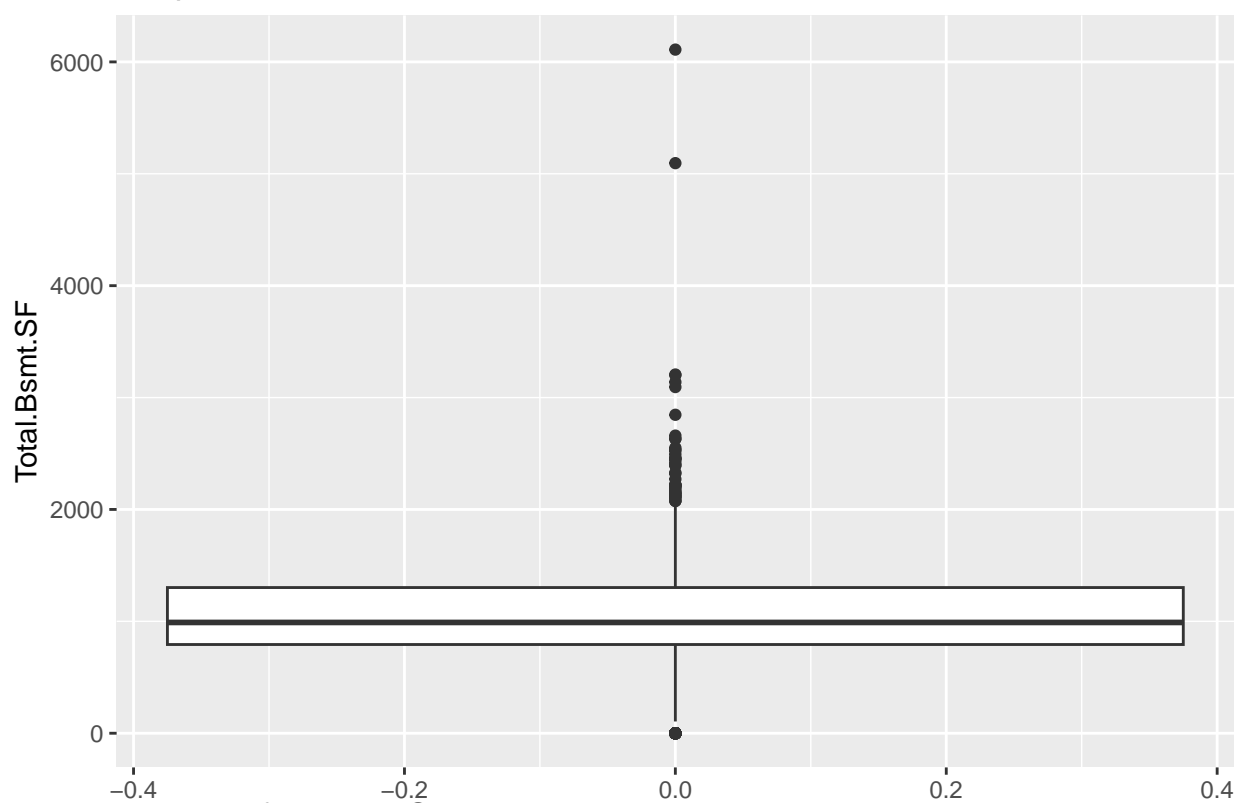
Boxplot of X1st.Flr.SF
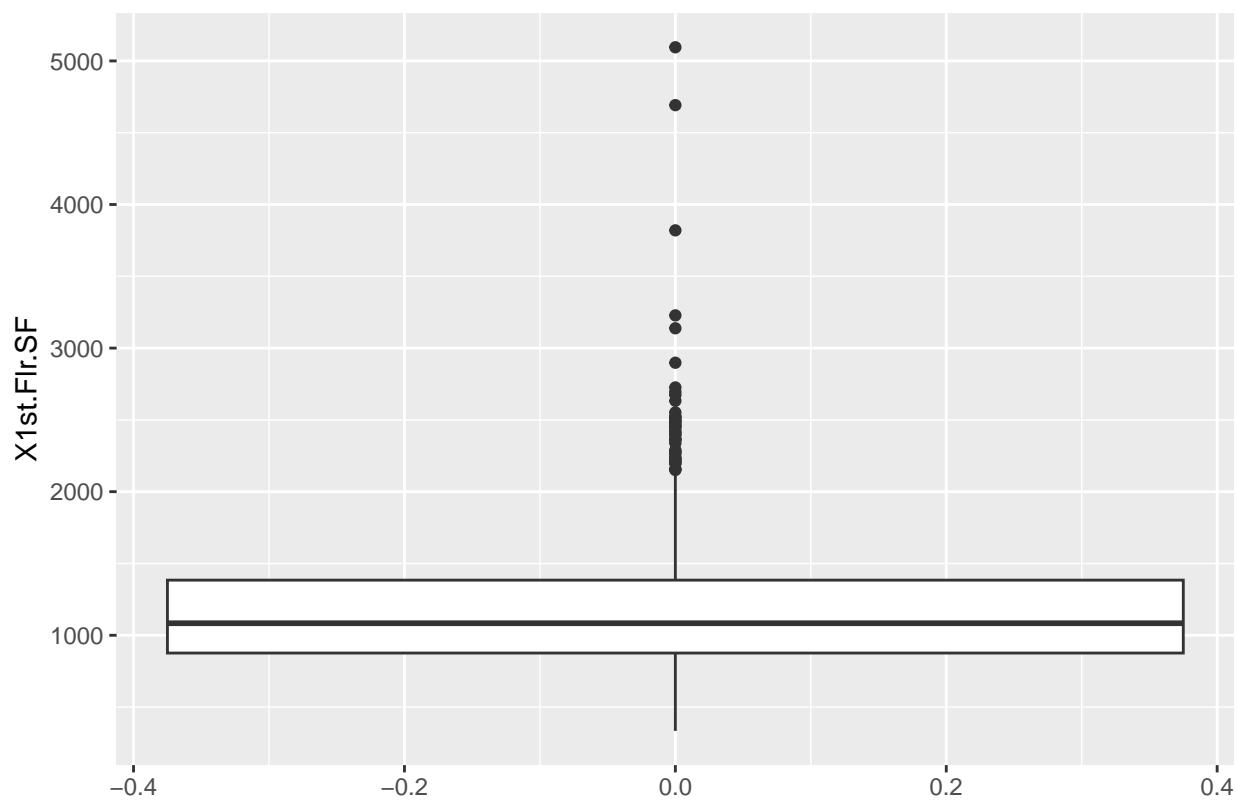
Boxplot of X2nd.Flr.SF
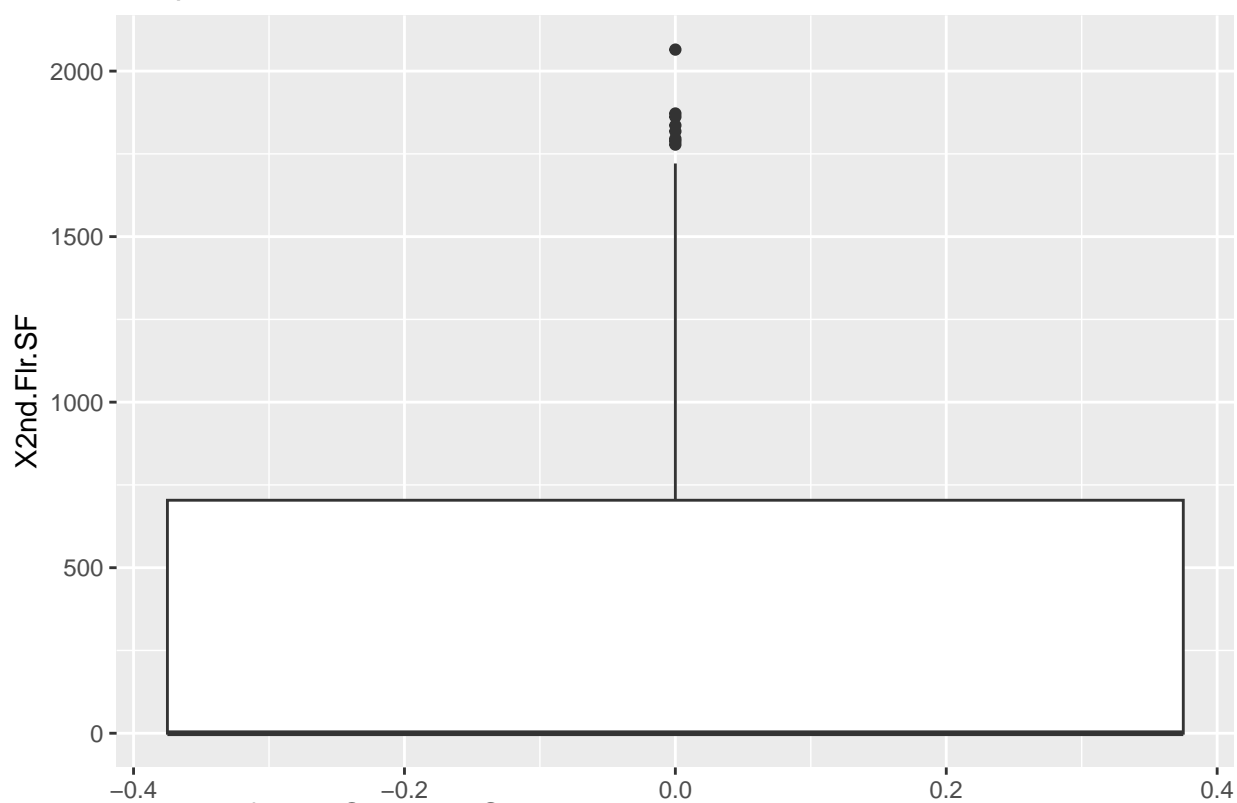
Boxplot of Low.Qual.Fin.SF
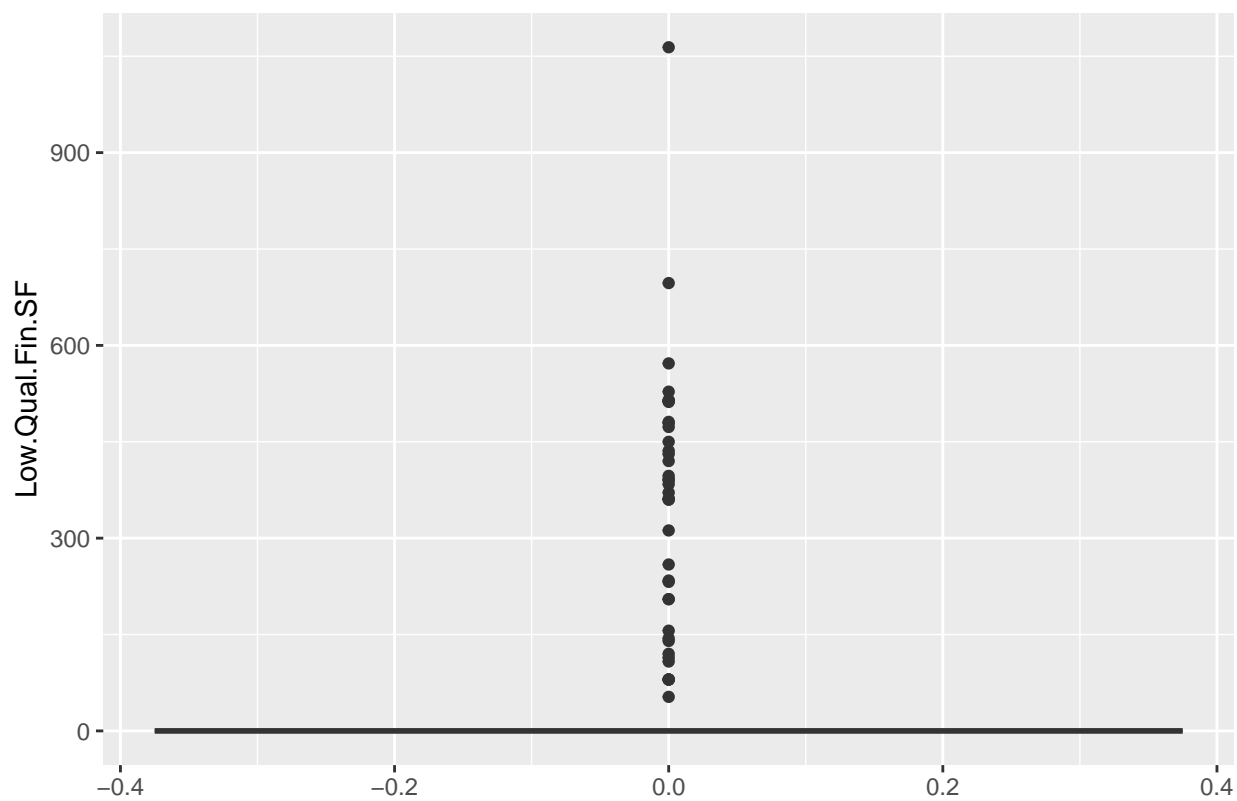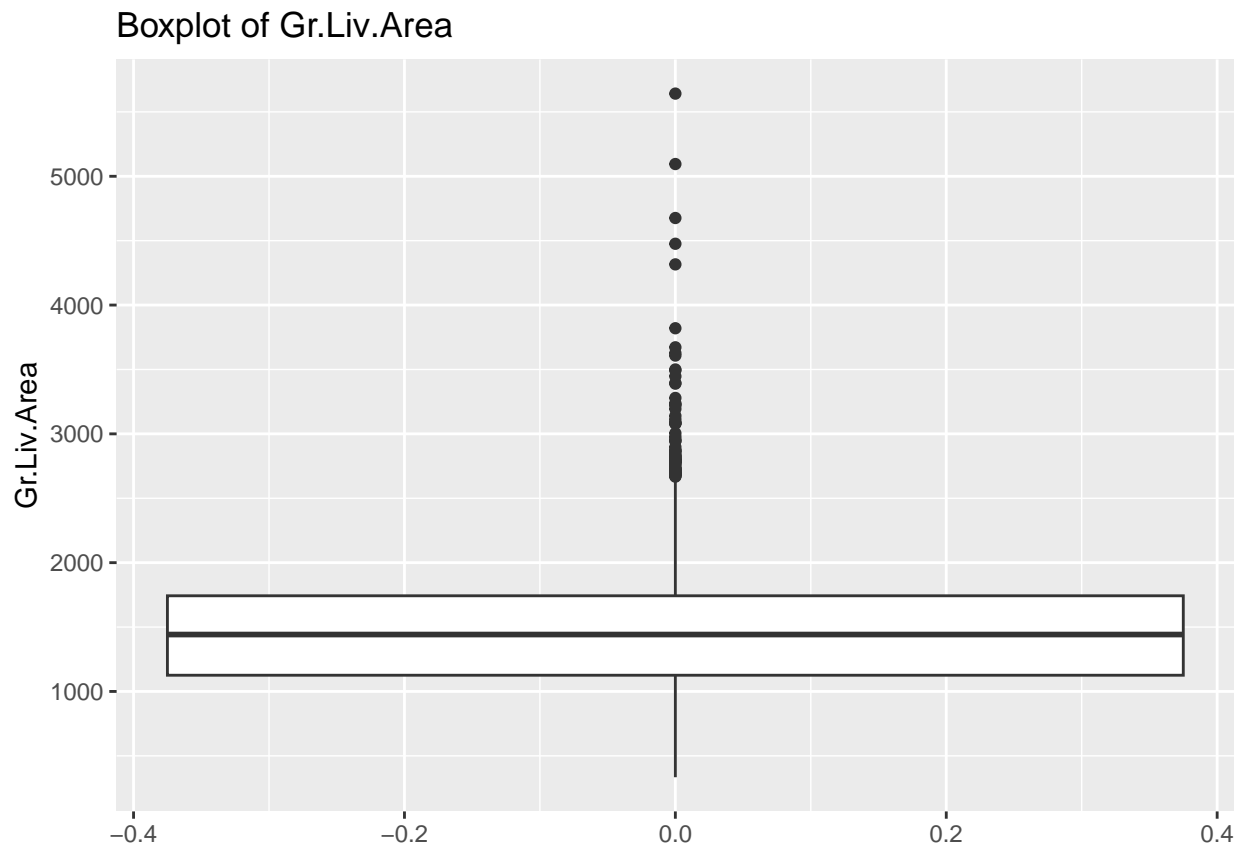
Warning: Removed 2 rows containing non-finite values (`stat_boxplot()`).

Boxplot of Bsmt.Full.Bath


```
## Warning: Removed 2 rows containing non-finite values (`stat_boxplot()`).
```

Boxplot of Bsmt.Half.Bath
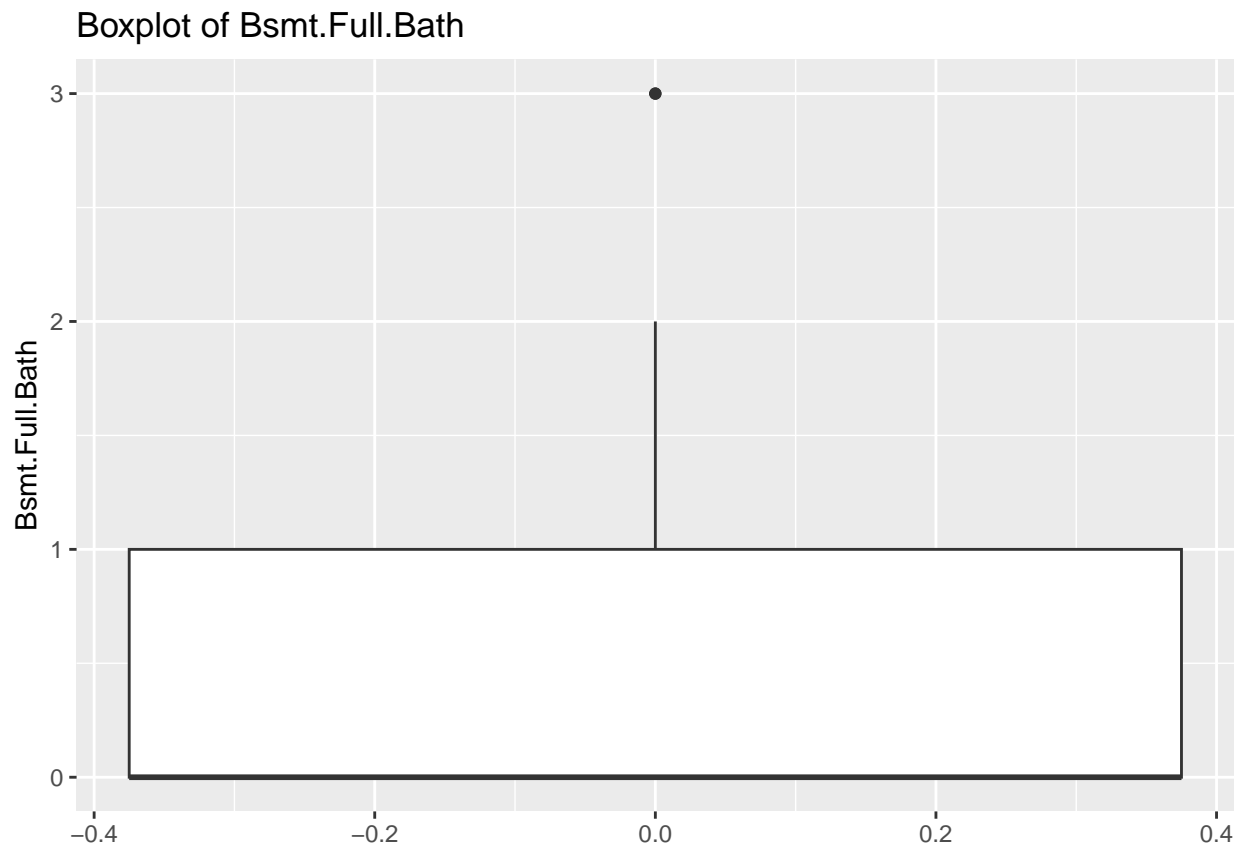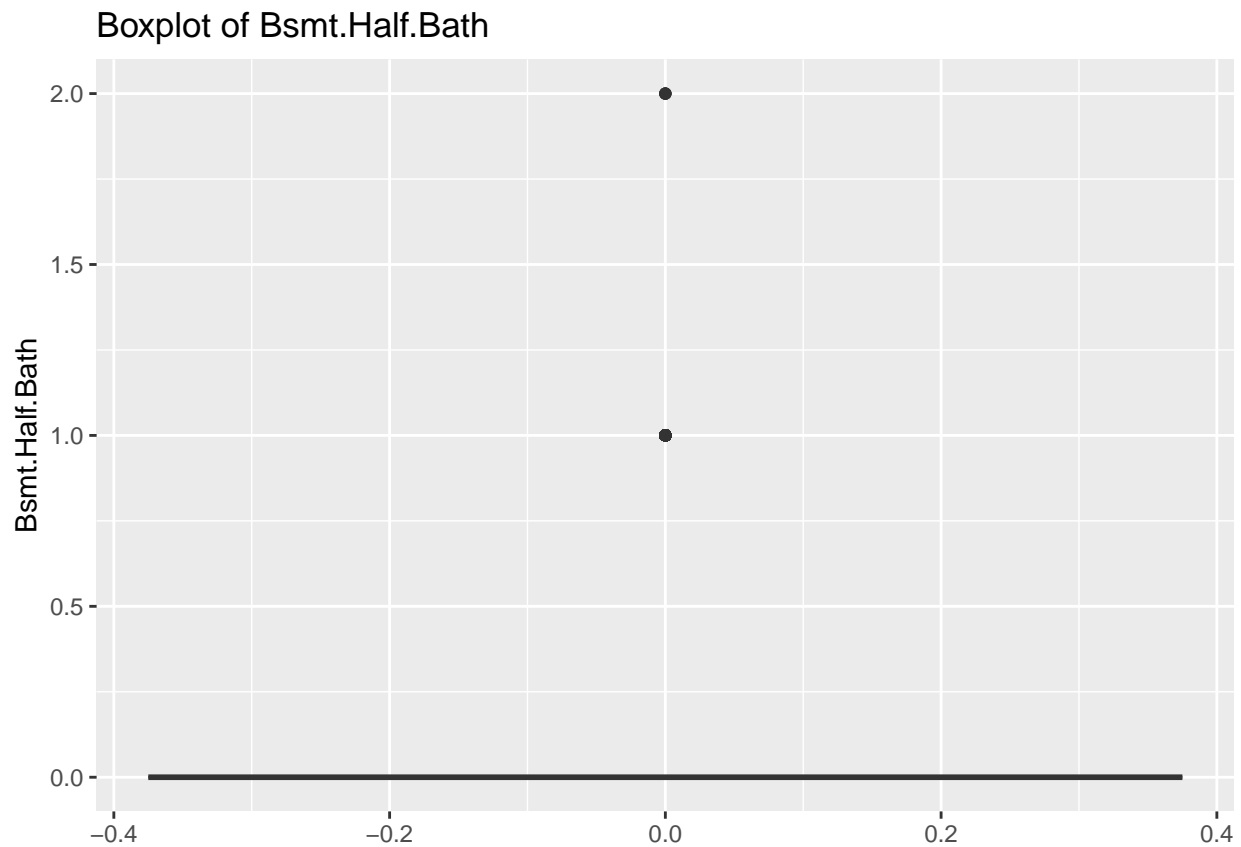

```
# Correlation plot for selected continuous variables
continuous_data <- housing[, selected_vars]
corrplot(cor(continuous_data, use = "complete.obs"), method = "circle")
```
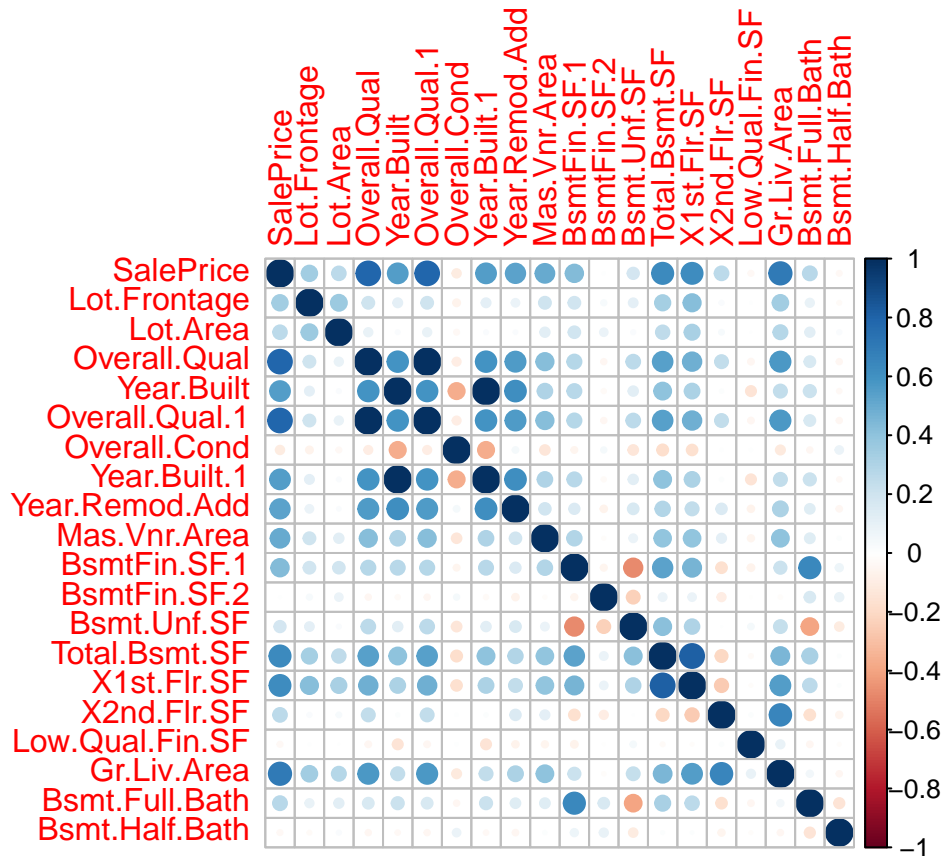
```
#Requirement 2: Regression Model and Diagnostics
# Split Data into Training and Testing Sets
set.seed(2023)
test_indices <- sample(nrow(housing), round(nrow(housing)/4))
train_set <- housing[-test_indices, ]
test_set <- housing[test_indices, ]


# Check for Highly Correlated Predictors
corr_matrix <- cor(train_set[, selected_vars], use = "complete.obs")
high_corr <- findCorrelation(corr_matrix, cutoff = 0.75)
high_corr_vars <- names(train_set[, selected_vars])[high_corr]
high_corr_vars
```

```
## [1] "SalePrice"     "Overall.Qual" "Total.Bsmt.SF" "Year.Built"
```

```
selected_high_corr_vars <- setdiff(selected_vars, c("SalePrice", "Overall.Qual", "Year.Built", "Total.B
```

```
# Fit Regression Model and Perform Diagnostics
lm_model <- lm(SalePrice ~ ., data = train_set[, selected_vars])
summary(lm_model)
```

```
##
## Call:
## lm(formula = SalePrice ~ ., data = train_set[, selected_vars])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -580068  -16912   -2103   13834  258855
##
## Coefficients: (4 not defined because of singularities)
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.282e+06  9.001e+04 -14.238  < 2e-16 ***
## Lot.Frontage    1.064e+02  3.959e+01   2.688 0.007232 **
## Lot.Area        6.196e-01  9.524e-02   6.506 9.55e-11 ***
## Overall.Qual    2.101e+04  8.669e+02  24.233  < 2e-16 ***
## Year.Built      3.851e+02  4.039e+01   9.534  < 2e-16 ***
## Overall.Qual.1        NA         NA      NA       NA
## Overall.Cond    4.685e+03  7.951e+02   5.893 4.38e-09 ***
## Year.Built.1          NA         NA      NA       NA
## Year.Remod.Add  2.136e+02  5.168e+01   4.132 3.72e-05 ***
## Mas.Vnr.Area    2.862e+01  4.747e+00   6.029 1.93e-09 ***
## BsmtFin.SF.1    2.403e+01  3.618e+00   6.643 3.87e-11 ***
## BsmtFin.SF.2    1.613e+01  5.513e+00   2.925 0.003478 **
## Bsmt.Unf.SF     1.058e+01  3.216e+00   3.289 0.001022 **
## Total.Bsmt.SF         NA         NA      NA       NA
## X1st.Flr.SF     6.077e+01  3.619e+00  16.791  < 2e-16 ***
## X2nd.Flr.SF     5.246e+01  2.137e+00  24.553  < 2e-16 ***
## Low.Qual.Fin.SF 2.733e+01  1.640e+01   1.667 0.095749 .
## Gr.Liv.Area           NA         NA      NA       NA
## Bsmt.Full.Bath  7.368e+03  2.056e+03   3.583 0.000347 ***
## Bsmt.Half.Bath -4.544e+03  3.355e+03  -1.354 0.175735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34810 on 2180 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.8062, Adjusted R-squared:  0.8048
## F-statistic: 604.5 on 15 and 2180 DF,  p-value: < 2.2e-16
```

```
# Diagnostics: Check for assumptions, outliers, and multicollinearity
plot(lm_model)
```

Residuals vs Fitted

Residuals

Fitted values
lm(SalePrice ~ .)

Q–Q Residuals

Standardized residuals

Theoretical Quantiles
lm(SalePrice ~ .)

Scale–Location

lm(SalePrice ~ .)

Residuals vs Leverage
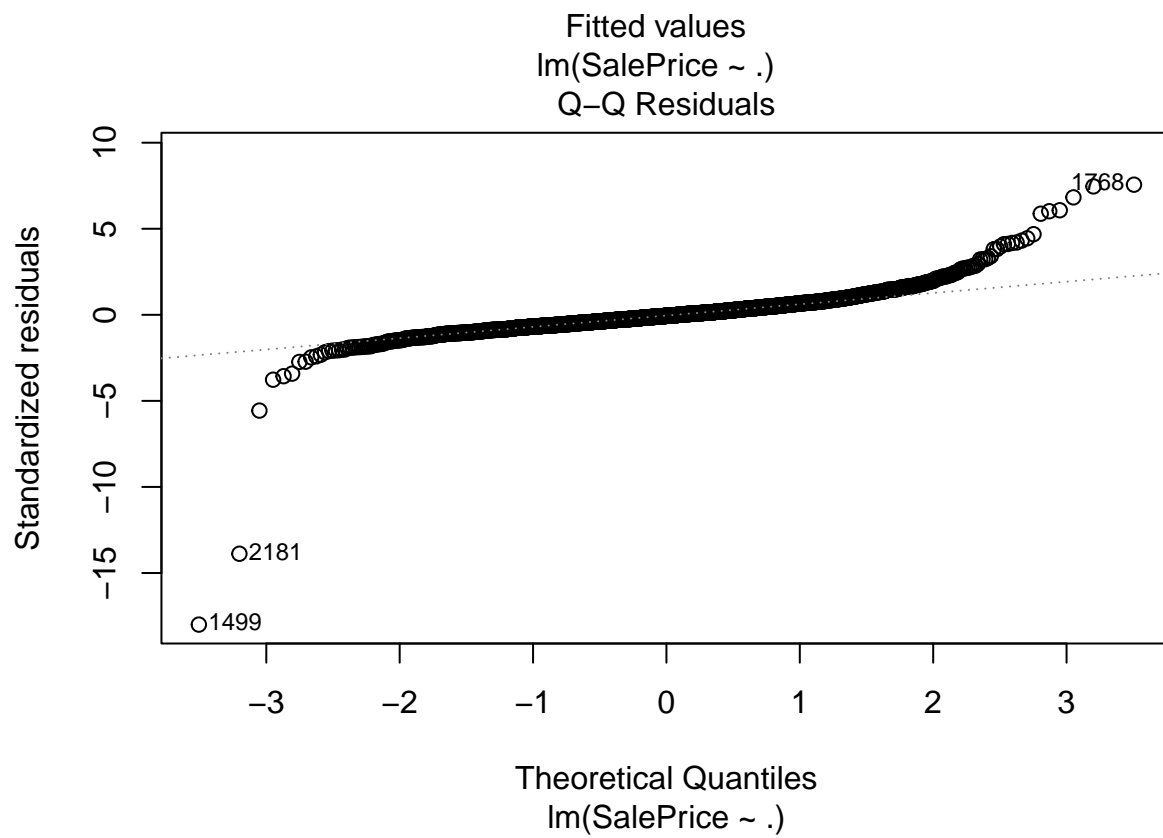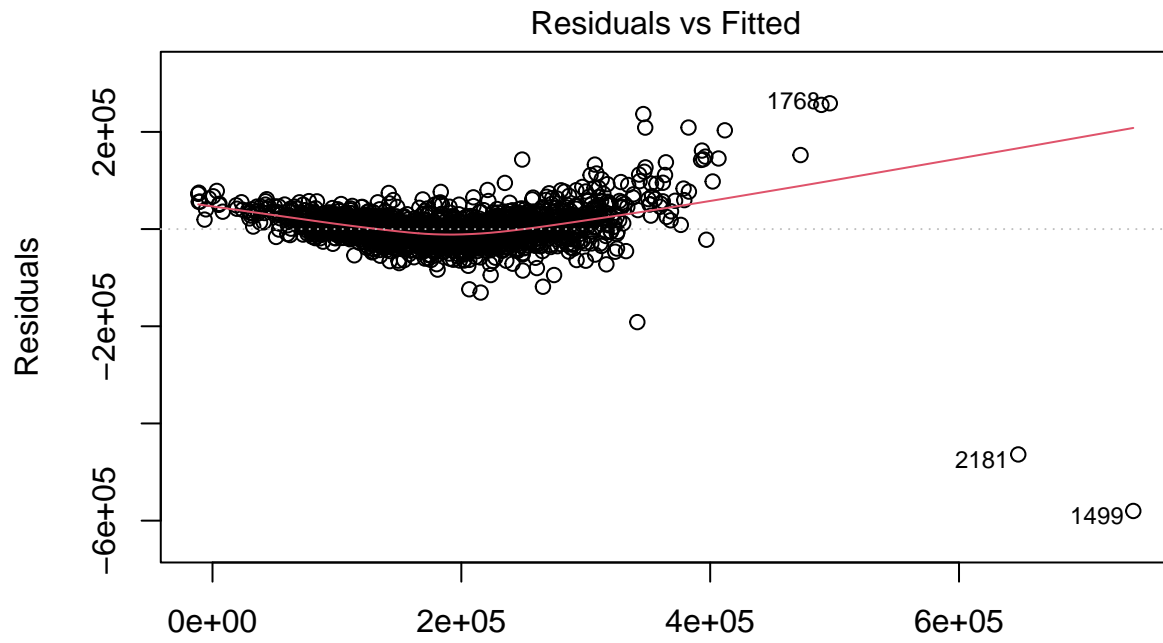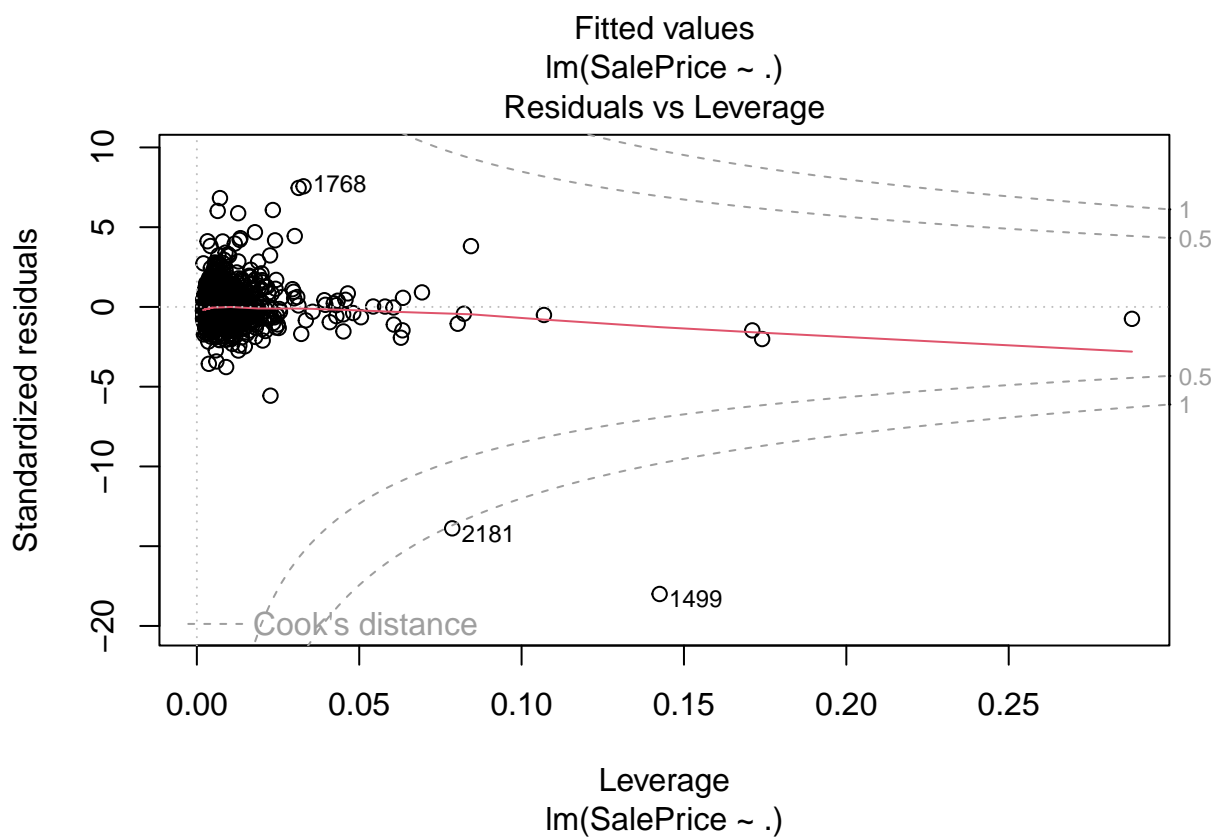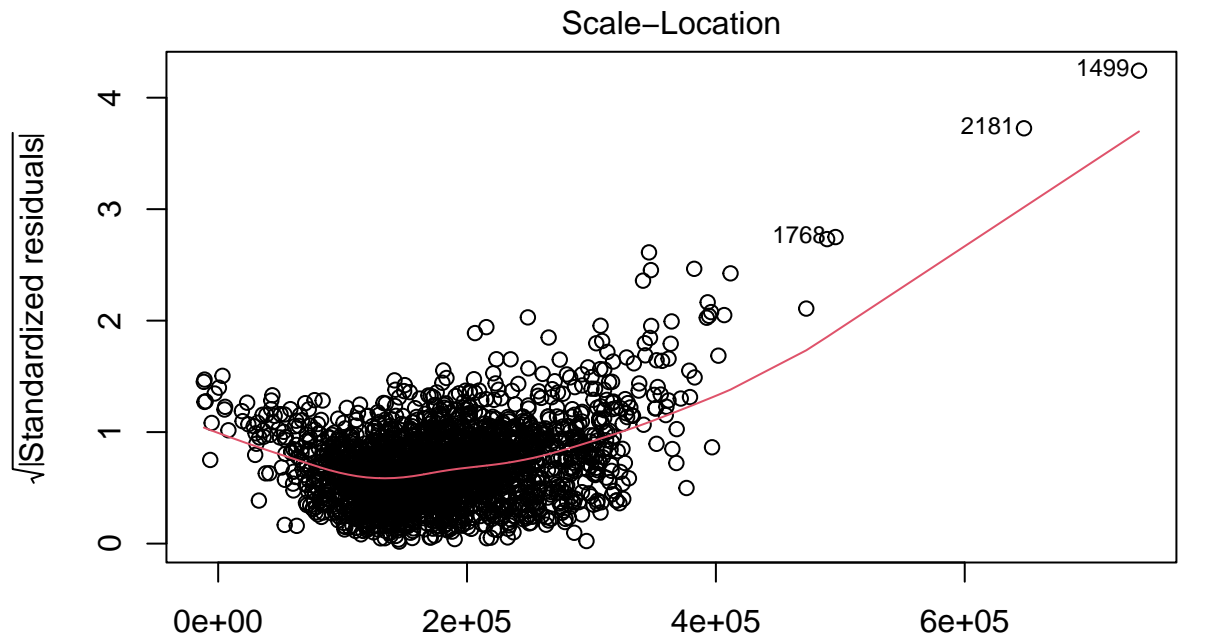
lm(SalePrice ~ .)

```r
#Requirement 3: Model Fitting and Comparison
# Assuming 'selected_vars' contains the names of your predictors and excludes 'SalePrice'
# Ensure the training set does not have missing values in the selected variables
train_set_clean <- na.omit(train_set[, c(selected_vars, "SalePrice")])

# Create the model matrix for predictors
```

```r
x <- model.matrix(~ ., data = train_set_clean[, selected_vars])

# Create the response vector
y <- train_set_clean$SalePrice

# Check if the number of rows in x matches the length of y
nrow(x) == length(y)
```

```
## [1] TRUE
```

```r
# Fit AIC model
aic_model <- step(lm(SalePrice ~ ., data = train_set[, selected_vars]), direction = "both", criterion =
```

```
## Start:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     Total.Bsmt.SF + X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF +
##     Gr.Liv.Area + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     Total.Bsmt.SF + X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF +
##     Bsmt.Full.Bath + Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath +
##     Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Remod.Add + Mas.Vnr.Area +
##     BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF +
##     X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Cond + Year.Remod.Add + Mas.Vnr.Area + BsmtFin.SF.1 +
##     BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF + X2nd.Flr.SF +
##     Low.Qual.Fin.SF + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##                     Df  Sum of Sq        RSS    AIC
## - Bsmt.Half.Bath     1  2.2226e+09  2.6434e+12  45945
## <none>                             2.6412e+12  45946
## - Low.Qual.Fin.SF    1  3.3649e+09  2.6445e+12  45946
```

```
## - Lot.Frontage       1 8.7570e+09 2.6499e+12 45951
## - BsmtFin.SF.2        1 1.0367e+10 2.6515e+12 45952
## - Bsmt.Unf.SF         1 1.3105e+10 2.6543e+12 45955
## - Bsmt.Full.Bath      1 1.5556e+10 2.6567e+12 45957
## - Year.Remod.Add      1 2.0690e+10 2.6619e+12 45961
## - Overall.Cond        1 4.2075e+10 2.6832e+12 45978
## - Mas.Vnr.Area        1 4.4035e+10 2.6852e+12 45980
## - Lot.Area            1 5.1280e+10 2.6924e+12 45986
## - BsmtFin.SF.1        1 5.3463e+10 2.6946e+12 45988
## - Year.Built          1 1.1014e+11 2.7513e+12 46033
## - X1st.Flr.SF         1 3.4160e+11 2.9828e+12 46211
## - Overall.Qual        1 7.1146e+11 3.3526e+12 46467
## - X2nd.Flr.SF         1 7.3035e+11 3.3715e+12 46480
##
## Step:  AIC=45945.48
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Cond + Year.Remod.Add + Mas.Vnr.Area + BsmtFin.SF.1 +
##     BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF + X2nd.Flr.SF +
##     Low.Qual.Fin.SF + Bsmt.Full.Bath
##
##                    Df  Sum of Sq        RSS    AIC
## <none>                            2.6434e+12 45945
## + Bsmt.Half.Bath    1 2.2226e+09 2.6412e+12 45946
## - Low.Qual.Fin.SF   1 3.3661e+09 2.6468e+12 45946
## - BsmtFin.SF.2      1 9.0867e+09 2.6525e+12 45951
## - Lot.Frontage      1 9.1020e+09 2.6525e+12 45951
## - Bsmt.Unf.SF       1 1.3132e+10 2.6565e+12 45954
## - Year.Remod.Add    1 2.0769e+10 2.6642e+12 45961
## - Bsmt.Full.Bath    1 2.1068e+10 2.6645e+12 45961
## - Overall.Cond      1 4.0931e+10 2.6843e+12 45977
## - Mas.Vnr.Area      1 4.3803e+10 2.6872e+12 45980
## - Lot.Area          1 5.0522e+10 2.6939e+12 45985
## - BsmtFin.SF.1      1 5.1292e+10 2.6947e+12 45986
## - Year.Built        1 1.0959e+11 2.7530e+12 46033
## - X1st.Flr.SF       1 3.4457e+11 2.9880e+12 46213
## - Overall.Qual      1 7.1269e+11 3.3561e+12 46468
## - X2nd.Flr.SF       1 7.3397e+11 3.3773e+12 46482
```
```r
# Fit BIC model
bic_model <- step(lm(SalePrice ~ ., data = train_set[, selected_vars]), direction = "both", criterion =
```
```
## Start:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     Total.Bsmt.SF + X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF +
##     Gr.Liv.Area + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     Total.Bsmt.SF + X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF +
##     Bsmt.Full.Bath + Bsmt.Half.Bath
```

```
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Built.1 + Year.Remod.Add +
##     Mas.Vnr.Area + BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF +
##     X1st.Flr.SF + X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath +
##     Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Qual.1 + Overall.Cond + Year.Remod.Add + Mas.Vnr.Area +
##     BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF +
##     X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##
## Step:  AIC=45945.63
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Cond + Year.Remod.Add + Mas.Vnr.Area + BsmtFin.SF.1 +
##     BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF + X2nd.Flr.SF +
##     Low.Qual.Fin.SF + Bsmt.Full.Bath + Bsmt.Half.Bath
##
##                    Df  Sum of Sq         RSS    AIC
## - Bsmt.Half.Bath    1 2.2226e+09  2.6434e+12  45945
## <none>                            2.6412e+12  45946
## - Low.Qual.Fin.SF   1 3.3649e+09  2.6445e+12  45946
## - Lot.Frontage      1 8.7570e+09  2.6499e+12  45951
## - BsmtFin.SF.2      1 1.0367e+10  2.6515e+12  45952
## - Bsmt.Unf.SF       1 1.3105e+10  2.6543e+12  45955
## - Bsmt.Full.Bath    1 1.5556e+10  2.6567e+12  45957
## - Year.Remod.Add    1 2.0690e+10  2.6619e+12  45961
## - Overall.Cond      1 4.2075e+10  2.6832e+12  45978
## - Mas.Vnr.Area      1 4.4035e+10  2.6852e+12  45980
## - Lot.Area          1 5.1280e+10  2.6924e+12  45986
## - BsmtFin.SF.1      1 5.3463e+10  2.6946e+12  45988
## - Year.Built        1 1.1014e+11  2.7513e+12  46033
## - X1st.Flr.SF       1 3.4160e+11  2.9828e+12  46211
## - Overall.Qual      1 7.1146e+11  3.3526e+12  46467
## - X2nd.Flr.SF       1 7.3035e+11  3.3715e+12  46480
##
## Step:  AIC=45945.48
## SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual + Year.Built +
##     Overall.Cond + Year.Remod.Add + Mas.Vnr.Area + BsmtFin.SF.1 +
##     BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF + X2nd.Flr.SF +
##     Low.Qual.Fin.SF + Bsmt.Full.Bath
##
##                    Df  Sum of Sq         RSS    AIC
## <none>                            2.6434e+12  45945
## + Bsmt.Half.Bath    1 2.2226e+09  2.6412e+12  45946
## - Low.Qual.Fin.SF   1 3.3661e+09  2.6468e+12  45946
## - BsmtFin.SF.2      1 9.0867e+09  2.6525e+12  45951
## - Lot.Frontage      1 9.1020e+09  2.6525e+12  45951
## - Bsmt.Unf.SF       1 1.3132e+10  2.6565e+12  45954
```

21

```
## - Year.Remod.Add   1 2.0769e+10 2.6642e+12 45961
## - Bsmt.Full.Bath   1 2.1068e+10 2.6645e+12 45961
## - Overall.Cond     1 4.0931e+10 2.6843e+12 45977
## - Mas.Vnr.Area     1 4.3803e+10 2.6872e+12 45980
## - Lot.Area         1 5.0522e+10 2.6939e+12 45985
## - BsmtFin.SF.1     1 5.1292e+10 2.6947e+12 45986
## - Year.Built       1 1.0959e+11 2.7530e+12 46033
## - X1st.Flr.SF      1 3.4457e+11 2.9880e+12 46213
## - Overall.Qual     1 7.1269e+11 3.3561e+12 46468
## - X2nd.Flr.SF      1 7.3397e+11 3.3773e+12 46482
```

```r
# Fit Lasso model
lasso_model <- glmnet(x, y, alpha = 1)

# Fit Elastic Net model
elastic_net_model <- glmnet(x, y, alpha = 0.5)

# Summarize AIC and BIC models
summary(aic_model)
```

```
##
## Call:
## lm(formula = SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual +
##     Year.Built + Overall.Cond + Year.Remod.Add + Mas.Vnr.Area +
##     BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF +
##     X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath, data = train_set[,
##     selected_vars])
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -577788  -17002   -2041   13882  256052
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -1.281e+06  9.002e+04 -14.224  < 2e-16 ***
## Lot.Frontage     1.084e+02  3.957e+01   2.740  0.00619 **
## Lot.Area         6.146e-01  9.519e-02   6.456 1.32e-10 ***
## Overall.Qual     2.102e+04  8.670e+02  24.249  < 2e-16 ***
## Year.Built       3.841e+02  4.039e+01   9.509  < 2e-16 ***
## Overall.Cond     4.610e+03  7.932e+02   5.811 7.11e-09 ***
## Year.Remod.Add   2.140e+02  5.169e+01   4.140 3.61e-05 ***
## Mas.Vnr.Area     2.854e+01  4.748e+00   6.012 2.15e-09 ***
## BsmtFin.SF.1     2.318e+01  3.564e+00   6.505 9.58e-11 ***
## BsmtFin.SF.2     1.489e+01  5.438e+00   2.738  0.00623 **
## Bsmt.Unf.SF      1.059e+01  3.217e+00   3.292  0.00101 **
## X1st.Flr.SF      6.098e+01  3.616e+00  16.861  < 2e-16 ***
## X2nd.Flr.SF      5.256e+01  2.136e+00  24.609  < 2e-16 ***
## Low.Qual.Fin.SF  2.734e+01  1.640e+01   1.667  0.09575 .
## Bsmt.Full.Bath   8.192e+03  1.965e+03   4.169 3.18e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34810 on 2181 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.806,  Adjusted R-squared:  0.8048
```

```
## F-statistic: 647.3 on 14 and 2181 DF,  p-value: < 2.2e-16
```
```r
summary(bic_model)
```
```
##
## Call:
## lm(formula = SalePrice ~ Lot.Frontage + Lot.Area + Overall.Qual +
##     Year.Built + Overall.Cond + Year.Remod.Add + Mas.Vnr.Area +
##     BsmtFin.SF.1 + BsmtFin.SF.2 + Bsmt.Unf.SF + X1st.Flr.SF +
##     X2nd.Flr.SF + Low.Qual.Fin.SF + Bsmt.Full.Bath, data = train_set[,
##     selected_vars])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -577788  -17002   -2041   13882  256052
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -1.281e+06  9.002e+04 -14.224  < 2e-16 ***
## Lot.Frontage     1.084e+02  3.957e+01   2.740  0.00619 **
## Lot.Area         6.146e-01  9.519e-02   6.456 1.32e-10 ***
## Overall.Qual     2.102e+04  8.670e+02  24.249  < 2e-16 ***
## Year.Built       3.841e+02  4.039e+01   9.509  < 2e-16 ***
## Overall.Cond     4.610e+03  7.932e+02   5.811 7.11e-09 ***
## Year.Remod.Add   2.140e+02  5.169e+01   4.140 3.61e-05 ***
## Mas.Vnr.Area     2.854e+01  4.748e+00   6.012 2.15e-09 ***
## BsmtFin.SF.1     2.318e+01  3.564e+00   6.505 9.58e-11 ***
## BsmtFin.SF.2     1.489e+01  5.438e+00   2.738  0.00623 **
## Bsmt.Unf.SF      1.059e+01  3.217e+00   3.292  0.00101 **
## X1st.Flr.SF      6.098e+01  3.616e+00  16.861  < 2e-16 ***
## X2nd.Flr.SF      5.256e+01  2.136e+00  24.609  < 2e-16 ***
## Low.Qual.Fin.SF  2.734e+01  1.640e+01   1.667  0.09575 .
## Bsmt.Full.Bath   8.192e+03  1.965e+03   4.169 3.18e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 34810 on 2181 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.806,  Adjusted R-squared:  0.8048
## F-statistic: 647.3 on 14 and 2181 DF,  p-value: < 2.2e-16
```
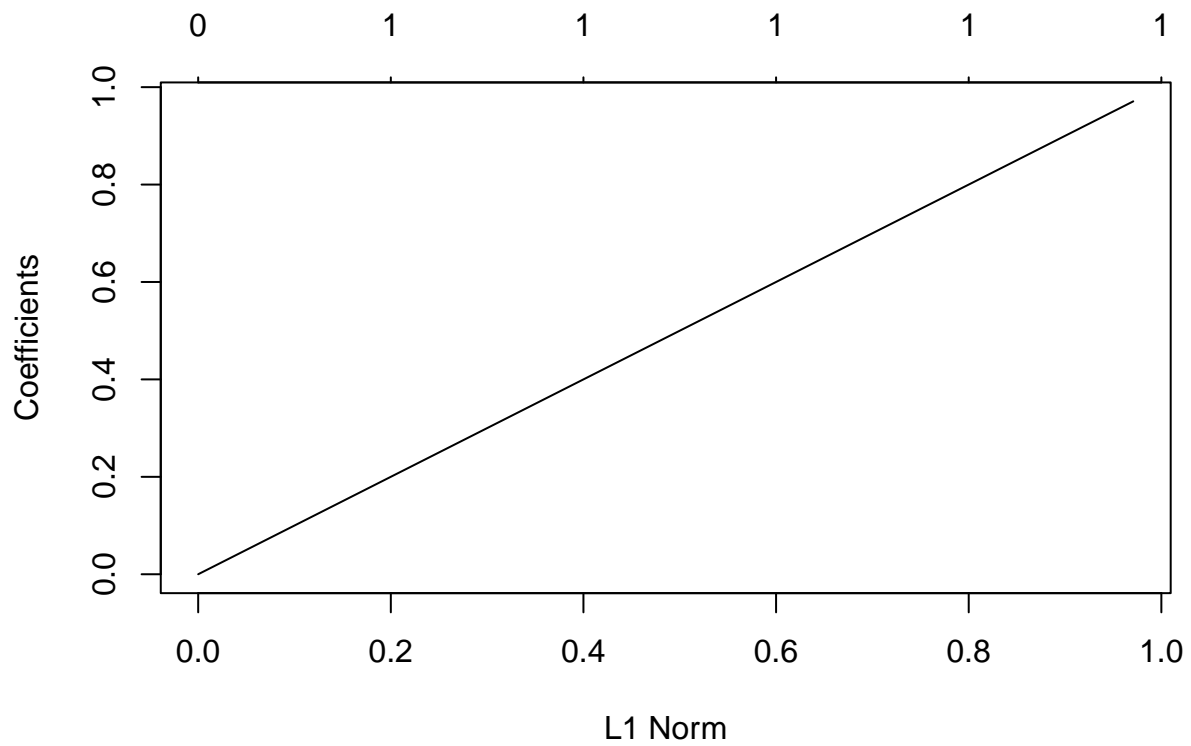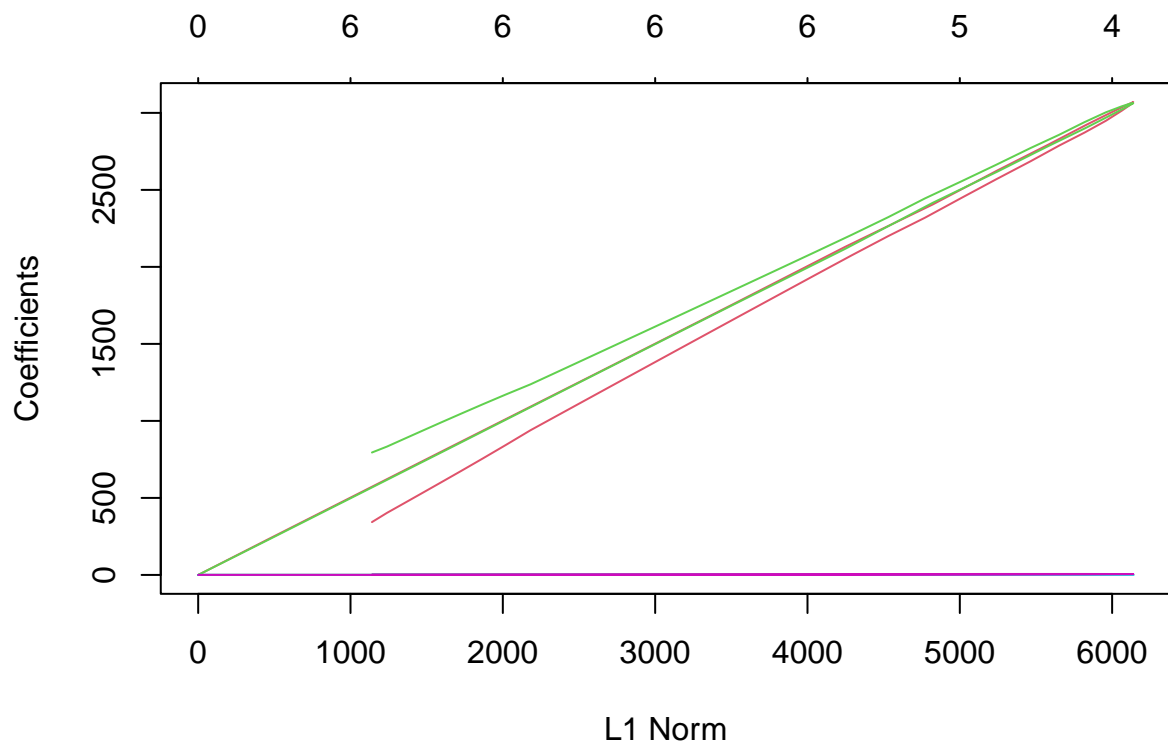```r
# Plot Lasso and Elastic Net models
plot(lasso_model)
```
```
Warning in plotCoef(x$beta, lambda = x$lambda, df = x$df, dev = x$dev.ratio, :
1 or less nonzero coefficients; glmnet plot is not meaningful
```

```r
plot(elastic_net_model)
```



```r
#Requirement 4: Calculate Mean Prediction Error
# Prepare Testing Data
x_test <- model.matrix(~ ., data = test_set[, selected_vars])
y_test <- test_set$SalePrice
```

```r
# Calculate Mean Prediction Error
predict_error <- function(model, x, y) {
  predictions <- predict(model, newx = x)
  mean((predictions - y)^2)
}

# Calculate mean prediction error for each model
mean_error_aic <- predict_error(aic_model, x_test, y_test)
mean_error_bic <- predict_error(bic_model, x_test, y_test)
mean_error_lasso <- predict_error(lasso_model, x_test, y_test)
mean_error_elastic_net <- predict_error(elastic_net_model, x_test, y_test)

# Output the errors
c(AIC = mean_error_aic, BIC = mean_error_bic, Lasso = mean_error_lasso, ElasticNet = mean_error_elastic
```

```
##          AIC          BIC        Lasso   ElasticNet
## 11536668354  11536668354   1040106431   1272404579
```