

# Assignment1.R

muj\_m

2022-04-06

```
library(plyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarise
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(Hmisc)
```

```
## Warning: package 'Hmisc' was built under R version 4.1.2
```

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

```
## Loading required package: survival
```

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

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:dplyr':
##
##     src, summarize
```

```
## The following objects are masked from 'package:plyr':  
##  
##     is.discrete, summarize
```

```
## The following objects are masked from 'package:base':  
##  
##     format.pval, units
```

```
library(ggplot2)  
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.1.2
```

```
## corrplot 0.92 loaded
```

```
library(RColorBrewer)  
library(leaps)
```

```
## Warning: package 'leaps' was built under R version 4.1.2
```

```
library(car)
```

```
## Loading required package: carData
```

```
##  
## Attaching package: 'car'
```

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

```
library(performance)
```

```
## Warning: package 'performance' was built under R version 4.1.2
```

```
#1# Loading the dataset.  
setwd("C:\\\\Users\\\\muj_m\\\\Desktop\\\\aly_6015")  
ds <- read.csv('AmesHousing.csv', header=TRUE)  
  
#renaming variables correctly  
colnames(ds)[1] <- 'Order'  
colnames(ds)[25] <- "Exterior.1"  
colnames(ds)[26] <- "Exterior.2"  
colnames(ds)[45] <- "1st.Flr.SF"  
colnames(ds)[46] <- "2nd.Flr.SF"  
colnames(ds)[71] <- "3-Ssn.Porch"  
  
#2# Exploratory data analysis for the data  
str(ds)
```

```

## 'data.frame': 2930 obs. of 82 variables:
## $ Order      : int 1 2 3 4 5 6 7 8 9 10 ...
## $ PID        : int 526301100 526350040 526351010 526353030 527105010 527105030 52712
7150 527145080 527146030 527162130 ...
## $ MS.SubClass : int 20 20 20 20 60 60 120 120 120 60 ...
## $ MS.Zoning   : chr "RL" "RH" "RL" "RL" ...
## $ Lot.Frontage : int 141 80 81 93 74 78 41 43 39 60 ...
## $ Lot.Area     : int 31770 11622 14267 11160 13830 9978 4920 5005 5389 7500 ...
## $ Street       : chr "Pave" "Pave" "Pave" "Pave" ...
## $ Alley        : chr NA NA NA NA ...
## $ Lot.Shape    : chr "IR1" "Reg" "IR1" "Reg" ...
## $ Land.Contour : chr "Lvl" "Lvl" "Lvl" "Lvl" ...
## $ Utilities    : chr "AllPub" "AllPub" "AllPub" "AllPub" ...
## $ Lot.Config   : chr "Corner" "Inside" "Corner" "Corner" ...
## $ Land.Slope    : chr "Gtl" "Gtl" "Gtl" "Gtl" ...
## $ Neighborhood : chr "NAmes" "NAmes" "NAmes" "NAmes" ...
## $ Condition.1  : chr "Norm" "Feedr" "Norm" "Norm" ...
## $ Condition.2  : chr "Norm" "Norm" "Norm" "Norm" ...
## $ Bldg.Type    : chr "1Fam" "1Fam" "1Fam" "1Fam" ...
## $ House.Style   : chr "1Story" "1Story" "1Story" "1Story" ...
## $ Overall.Qual : int 6 5 6 7 5 6 8 8 8 7 ...
## $ Overall.Cond : int 5 6 6 5 5 6 5 5 5 5 ...
## $ Year.Built   : int 1960 1961 1958 1968 1997 1998 2001 1992 1995 1999 ...
## $ Year.Remod.Add: int 1960 1961 1958 1968 1998 1998 2001 1992 1996 1999 ...
## $ Roof.Style   : chr "Hip" "Gable" "Hip" "Hip" ...
## $ Roof.Matl    : chr "CompShg" "CompShg" "CompShg" "CompShg" ...
## $ Exterior.1   : chr "BrkFace" "VinylSd" "Wd Sdng" "BrkFace" ...
## $ Exterior.2   : chr "Plywood" "VinylSd" "Wd Sdng" "BrkFace" ...
## $ Mas.Vnr.Type : chr "Stone" "None" "BrkFace" "None" ...
## $ Mas.Vnr.Area : int 112 0 108 0 0 20 0 0 0 0 ...
## $ Exter.Qual   : chr "TA" "TA" "TA" "Gd" ...
## $ Exter.Cond   : chr "TA" "TA" "TA" "TA" ...
## $ Foundation   : chr "CBlock" "CBlock" "CBlock" "CBlock" ...
## $ Bsmt.Qual    : chr "TA" "TA" "TA" "TA" ...
## $ BsmtCond     : chr "Gd" "TA" "TA" "TA" ...
## $ Bsmt.Exposure: chr "Gd" "No" "No" "No" ...
## $ BsmtFin.Type.1: chr "BLQ" "Rec" "ALQ" "ALQ" ...
## $ BsmtFin.SF.1 : int 639 468 923 1065 791 602 616 263 1180 0 ...
## $ BsmtFin.Type.2: chr "Unf" "LwQ" "Unf" "Unf" ...
## $ BsmtFin.SF.2 : int 0 144 0 0 0 0 0 0 0 0 ...
## $ Bsmt.Unf.SF  : int 441 270 406 1045 137 324 722 1017 415 994 ...
## $ Total.Bsmt.SF: int 1080 882 1329 2110 928 926 1338 1280 1595 994 ...
## $ Heating       : chr "GasA" "GasA" "GasA" "Gasa" ...
## $ Heating.QC   : chr "Fa" "TA" "TA" "Ex" ...
## $ Central.Air   : chr "Y" "Y" "Y" "Y" ...
## $ Electrical    : chr "SBrkr" "SBrkr" "SBrkr" "SBrkr" ...
## $ 1st.Flr.SF    : int 1656 896 1329 2110 928 926 1338 1280 1616 1028 ...
## $ 2nd.Flr.SF    : int 0 0 0 0 701 678 0 0 0 776 ...
## $ Low.Qual.Fin.SF: int 0 0 0 0 0 0 0 0 0 0 ...
## $ Gr.Liv.Area   : int 1656 896 1329 2110 1629 1604 1338 1280 1616 1804 ...
## $ Bsmt.Full.Bath: int 1 0 0 1 0 0 1 0 1 0 ...
## $ Bsmt.Half.Bath: int 0 0 0 0 0 0 0 0 0 0 ...
## $ Full.Bath     : int 1 1 1 2 2 2 2 2 2 2 ...
## $ Half.Bath     : int 0 0 1 1 1 0 0 0 1 ...
## $ Bedroom.AbvGr : int 3 2 3 3 3 3 2 2 2 3 ...

```

```
## $ Kitchen.AbvGr : int 1 1 1 1 1 1 1 1 1 1 1 ...
## $ Kitchen.Qual : chr "TA" "TA" "Gd" "Ex" ...
## $ TotRms.AbvGrd : int 7 5 6 8 6 7 6 5 5 7 ...
## $ Functional : chr "Typ" "Typ" "Typ" "Typ" ...
## $ Fireplaces : int 2 0 0 2 1 1 0 0 1 1 ...
## $ Fireplace.Qu : chr "Gd" NA NA "TA" ...
## $ Garage.Type : chr "Attchd" "Attchd" "Attchd" "Attchd" ...
## $ Garage.Yr.Blt : int 1960 1961 1958 1968 1997 1998 2001 1992 1995 1999 ...
## $ Garage.Finish : chr "Fin" "Unf" "Unf" "Fin" ...
## $ Garage.Cars : int 2 1 1 2 2 2 2 2 2 2 ...
## $ Garage.Area : int 528 730 312 522 482 470 582 506 608 442 ...
## $ Garage.Qual : chr "TA" "TA" "TA" "TA" ...
## $ Garage.Cond : chr "TA" "TA" "TA" "TA" ...
## $ Paved.Drive : chr "P" "Y" "Y" "Y" ...
## $ Wood.Deck.SF : int 210 140 393 0 212 360 0 0 237 140 ...
## $ Open.Porch.SF : int 62 0 36 0 34 36 0 82 152 60 ...
## $ Enclosed.Porch : int 0 0 0 0 0 0 170 0 0 0 ...
## $ 3-Ssn.Porch : int 0 0 0 0 0 0 0 0 0 0 ...
## $ Screen.Porch : int 0 120 0 0 0 0 0 144 0 0 ...
## $ Pool.Area : int 0 0 0 0 0 0 0 0 0 0 ...
## $ Pool.QC : chr NA NA NA NA ...
## $ Fence : chr NA "MnPrv" NA NA ...
## $ Misc.Feature : chr NA NA "Gar2" NA ...
## $ Misc.Val : int 0 0 12500 0 0 0 0 0 0 0 ...
## $ Mo.Sold : int 5 6 6 4 3 6 4 1 3 6 ...
## $ Yr.Sold : int 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
## $ Sale.Type : chr "WD" "WD" "WD" "WD" ...
## $ Sale.Condition : chr "Normal" "Normal" "Normal" "Normal" ...
## $ SalePrice : int 215000 105000 172000 244000 189900 195500 213500 191500 236500 189000 ...
```

```
summary(ds)
```

```

##      Order          PID          MS.SubClass        MS.Zoning
##  Min.   : 1.0  Min.   :5.263e+08  Min.   : 20.00  Length:2930
##  1st Qu.: 733.2 1st Qu.:5.285e+08  1st Qu.: 20.00  Class  :character
##  Median :1465.5 Median :5.355e+08  Median : 50.00  Mode   :character
##  Mean   :1465.5 Mean   :7.145e+08  Mean   : 57.39
##  3rd Qu.:2197.8 3rd Qu.:9.072e+08  3rd Qu.: 70.00
##  Max.   :2930.0  Max.   :1.007e+09  Max.   :190.00
##
##      Lot.Frontage     Lot.Area       Street        Alley
##  Min.   : 21.00  Min.   : 1300  Length:2930    Length:2930
##  1st Qu.: 58.00  1st Qu.: 7440  Class  :character  Class  :character
##  Median : 68.00  Median : 9436  Mode   :character  Mode   :character
##  Mean   : 69.22  Mean   :10148
##  3rd Qu.: 80.00  3rd Qu.:11555
##  Max.   :313.00  Max.   :215245
##  NA's   :490
##      Lot.Shape     Land.Contour     Utilities     Lot.Config
##  Length:2930    Length:2930    Length:2930    Length:2930
##  Class  :character  Class  :character  Class  :character  Class  :character
##  Mode   :character  Mode   :character  Mode   :character  Mode   :character
##
##      Land.Slope     Neighborhood Condition.1 Condition.2
##  Length:2930    Length:2930    Length:2930    Length:2930
##  Class  :character  Class  :character  Class  :character  Class  :character
##  Mode   :character  Mode   :character  Mode   :character  Mode   :character
##
##      Bldg.Type     House.Style Overall.Qual Overall.Cond
##  Length:2930    Length:2930    Min.   : 1.000  Min.   :1.000
##  Class  :character  Class  :character  1st Qu.: 5.000  1st Qu.:5.000
##  Mode   :character  Mode   :character  Median : 6.000  Median :5.000
##                                Mean   : 6.095  Mean   :5.563
##                                3rd Qu.: 7.000  3rd Qu.:6.000
##                                Max.   :10.000  Max.   :9.000
##
##      Year.Built  Year.Remod.Add Roof.Style     Roof.Mat1
##  Min.   :1872   Min.   :1950   Length:2930    Length:2930
##  1st Qu.:1954   1st Qu.:1965   Class  :character  Class  :character
##  Median :1973   Median :1993   Mode   :character  Mode   :character
##  Mean   :1971   Mean   :1984
##  3rd Qu.:2001   3rd Qu.:2004
##  Max.   :2010   Max.   :2010
##
##      Exterior.1    Exterior.2     Mas.Vnr.Type     Mas.Vnr.Area
##  Length:2930    Length:2930    Length:2930    Min.   : 0.0
##  Class  :character  Class  :character  Class  :character  1st Qu.: 0.0
##  Mode   :character  Mode   :character  Mode   :character  Median : 0.0
##                                Mean   : 101.9
##                                3rd Qu.: 164.0
##                                Max.   :1600.0

```

```

##                                     NA's :23
##   Exter.Qual      Exter.Cond      Foundation      Bsmt.Qual
##   Length:2930     Length:2930     Length:2930     Length:2930
##   Class :character  Class :character  Class :character  Class :character
##   Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##                                     NA's :1
##   Bsmt.Cond      Bsmt.Exposure    BsmtFin.Type.1    BsmtFin.SF.1
##   Length:2930     Length:2930     Length:2930     Min.   : 0.0
##   Class :character  Class :character  Class :character  1st Qu.: 0.0
##   Mode  :character  Mode  :character  Mode  :character  Median  : 370.0
##                                     Mean   : 442.6
##                                     3rd Qu.: 734.0
##                                     Max.   :5644.0
##                                     NA's :1
##   BsmtFin.Type.2    BsmtFin.SF.2    Bsmt.Unf.SF    Total.Bsmt.SF
##   Length:2930     Min.   : 0.00    Min.   : 0.0    Min.   : 0
##   Class :character  1st Qu.: 0.00    1st Qu.: 219.0  1st Qu.: 793
##   Mode  :character  Median : 0.00    Median : 466.0  Median : 990
##                                     Mean   : 49.72    Mean   : 559.3  Mean   :1052
##                                     3rd Qu.: 0.00    3rd Qu.: 802.0  3rd Qu.:1302
##                                     Max.   :1526.00   Max.   :2336.0  Max.   :6110
##                                     NA's :1       NA's :1       NA's :1
##   Heating          Heating.QC      Central.Air      Electrical
##   Length:2930     Length:2930     Length:2930     Length:2930
##   Class :character  Class :character  Class :character  Class :character
##   Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##                                     NA's :1
##   1st.Flr.SF      2nd.Flr.SF      Low.Qual.Fin.SF    Gr.Liv.Area
##   Min.   : 334.0   Min.   : 0.0    Min.   : 0.000    Min.   : 334
##   1st Qu.: 876.2   1st Qu.: 0.0    1st Qu.: 0.000    1st Qu.:1126
##   Median :1084.0   Median : 0.0    Median : 0.000    Median :1442
##   Mean   :1159.6   Mean   : 335.5   Mean   : 4.677    Mean   :1500
##   3rd Qu.:1384.0   3rd Qu.: 703.8   3rd Qu.: 0.000    3rd Qu.:1743
##   Max.   :5095.0   Max.   :2065.0   Max.   :1064.000   Max.   :5642
##                                     NA's :2       NA's :2
##   Bsmt.Full.Bath  Bsmt.Half.Bath  Full.Bath        Half.Bath
##   Min.   :0.0000   Min.   :0.00000  Min.   :0.000    Min.   :0.0000
##   1st Qu.:0.0000   1st Qu.:0.00000  1st Qu.:1.000    1st Qu.:0.0000
##   Median :0.0000   Median :0.00000  Median :2.000    Median :0.0000
##   Mean   :0.4314   Mean   :0.06113  Mean   :1.567    Mean   :0.3795
##   3rd Qu.:1.0000   3rd Qu.:0.00000  3rd Qu.:2.000    3rd Qu.:1.0000
##   Max.   :3.0000   Max.   :2.00000  Max.   :4.000    Max.   :2.0000
##   NA's   :2       NA's   :2
##   Bedroom.AbvGr   Kitchen.AbvGr   Kitchen.Qual    TotRms.AbvGrd
##   Min.   :0.000    Min.   :0.000    Length:2930     Min.   : 2.000
##   1st Qu.:2.000    1st Qu.:1.000    Class :character  1st Qu.: 5.000
##   Median :3.000    Median :1.000    Mode  :character  Median : 6.000
##   Mean   :2.854    Mean   :1.044                    Mean   : 6.443
##   3rd Qu.:3.000    3rd Qu.:1.000                    3rd Qu.: 7.000
##   Max.   :8.000    Max.   :3.000                    Max.   :15.000

```

```

## Functional Fireplaces Fireplace.Qu Garage.Type
## Length:2930 Min. :0.0000 Length:2930 Length:2930
## Class :character 1st Qu.:0.0000 Class :character Class :character
## Mode :character Median :1.0000 Mode :character Mode :character
## Mean :0.5993
## 3rd Qu.:1.0000
## Max. :4.0000
##
## Garage.Yr.Blt Garage.Finish Garage.Cars Garage.Area
## Min. :1895 Length:2930 Min. :0.000 Min. : 0.0
## 1st Qu.:1960 Class :character 1st Qu.:1.000 1st Qu.: 320.0
## Median :1979 Mode :character Median :2.000 Median : 480.0
## Mean :1978 Mean :1.767 Mean : 472.8
## 3rd Qu.:2002 3rd Qu.:2.000 3rd Qu.: 576.0
## Max. :2207 Max. :5.000 Max. :1488.0
## NA's :159 NA's :1 NA's :1
## Garage.Qual Garage.Cond Paved.Drive Wood.Deck.SF
## Length:2930 Length:2930 Length:2930 Min. : 0.00
## Class :character Class :character Class :character 1st Qu.: 0.00
## Mode :character Mode :character Mode :character Median : 0.00
## Mean : 93.75
## 3rd Qu.: 168.00
## Max. :1424.00
##
## Open.Porch.SF Enclosed.Porch 3-Ssn.Porch Screen.Porch
## Min. : 0.00 Min. : 0.00 Min. : 0.000 Min. : 0
## 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 0.000 1st Qu.: 0
## Median : 27.00 Median : 0.00 Median : 0.000 Median : 0
## Mean : 47.53 Mean : 23.01 Mean : 2.592 Mean : 16
## 3rd Qu.: 70.00 3rd Qu.: 0.00 3rd Qu.: 0.000 3rd Qu.: 0
## Max. :742.00 Max. :1012.00 Max. :508.000 Max. :576
##
## Pool.Area Pool.QC Fence Misc.Feature
## Min. : 0.000 Length:2930 Length:2930 Length:2930
## 1st Qu.: 0.000 Class :character Class :character Class :character
## Median : 0.000 Mode :character Mode :character Mode :character
## Mean : 2.243
## 3rd Qu.: 0.000
## Max. :800.000
##
## Misc.Val Mo.Sold Yr.Sold Sale.Type
## Min. : 0.00 Min. : 1.000 Min. :2006 Length:2930
## 1st Qu.: 0.00 1st Qu.: 4.000 1st Qu.:2007 Class :character
## Median : 0.00 Median : 6.000 Median :2008 Mode :character
## Mean : 50.63 Mean : 6.216 Mean :2008
## 3rd Qu.: 0.00 3rd Qu.: 8.000 3rd Qu.:2009
## Max. :17000.00 Max. :12.000 Max. :2010
##
## Sale.Condition SalePrice
## Length:2930 Min. : 12789
## Class :character 1st Qu.:129500
## Mode :character Median :160000
## Mean : 180796
## 3rd Qu.:213500

```

```
##          Max.    :755000
##
```

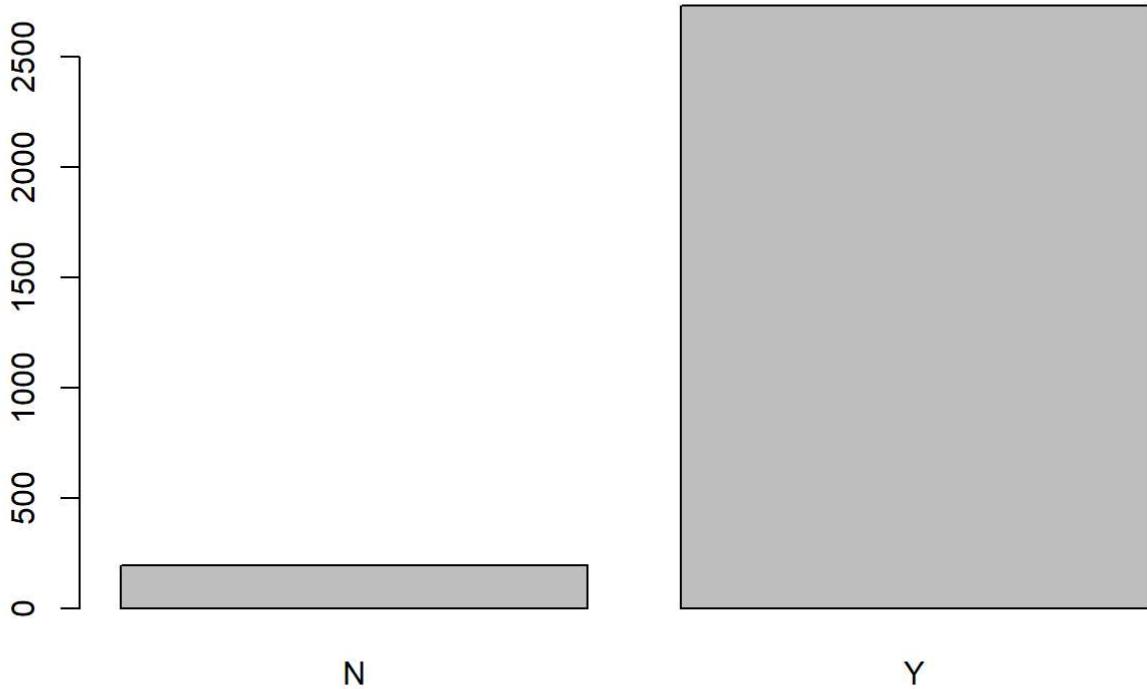
```
#getting all column names
colnames(ds)
```

```
## [1] "Order"           "PID"            "MS.SubClass"      "MS.Zoning"
## [5] "Lot.Frontage"    "Lot.Area"        "Street"          "Alley"
## [9] "Lot.Shape"        "Land.Contour"   "Utilities"        "Lot.Config"
## [13] "Land.Slope"       "Neighborhood"   "Condition.1"     "Condition.2"
## [17] "Bldg.Type"        "House.Style"    "Overall.Qual"   "Overall.Cond"
## [21] "Year.Built"       "Year.Remod.Add" "Roof.Style"      "Roof.Matl"
## [25] "Exterior.1"      "Exterior.2"     "Mas.Vnr.Type"   "Mas.Vnr.Area"
## [29] "Exter.Qual"       "Exter.Cond"     "Foundation"     "Bsmt.Qual"
## [33] "Bsmt.Cond"        "Bsmt.Exposure"  "BsmtFin.Type.1" "BsmtFin.SF.1"
## [37] "BsmtFin.Type.2"   "BsmtFin.SF.2"   "Bsmt.Unf.SF"    "Total.Bsmt.SF"
## [41] "Heating"          "Heating.QC"     "Central.Air"    "Electrical"
## [45] "1st.Flr.SF"       "2nd.Flr.SF"     "Low.Qual.Fin.SF" "Gr.Liv.Area"
## [49] "Bsmt.Full.Bath"   "Bsmt.Half.Bath" "Full.Bath"       "Half.Bath"
## [53] "Bedroom.AbvGr"    "Kitchen.AbvGr"  "Kitchen.Qual"   "TotRms.AbvGrd"
## [57] "Functional"       "Fireplaces"     "Fireplace.Qu"   "Garage.Type"
## [61] "Garage.Yr.Blt"    "Garage.Finish"  "Garage.Cars"    "Garage.Area"
## [65] "Garage.Qual"      "Garage.Cond"    "Paved.Drive"   "Wood.Deck.SF"
## [69] "Open.Porch.SF"    "Enclosed.Porch" "3-Ssn.Porch"    "Screen.Porch"
## [73] "Pool.Area"        "Pool.QC"        "Fence"          "Misc.Feature"
## [77] "Misc.Val"          "Mo.Sold"        "Yr.Sold"        "Sale.Type"
## [81] "Sale.Condition"   "SalePrice"
```

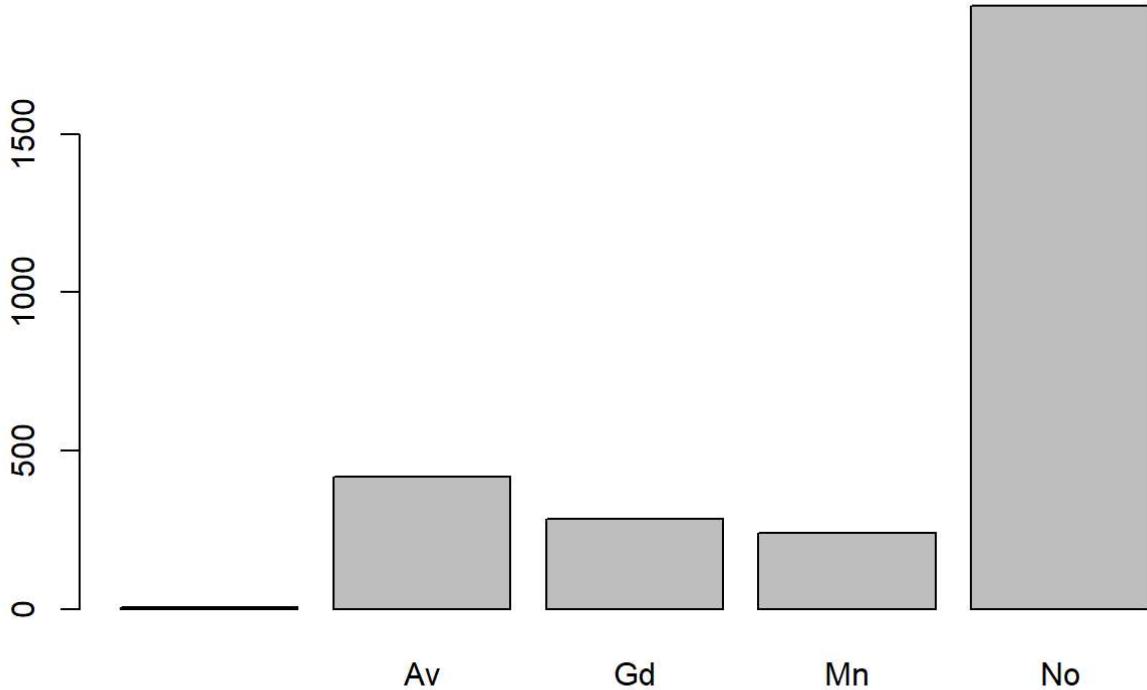
```
dim(ds)
```

```
## [1] 2930    82
```

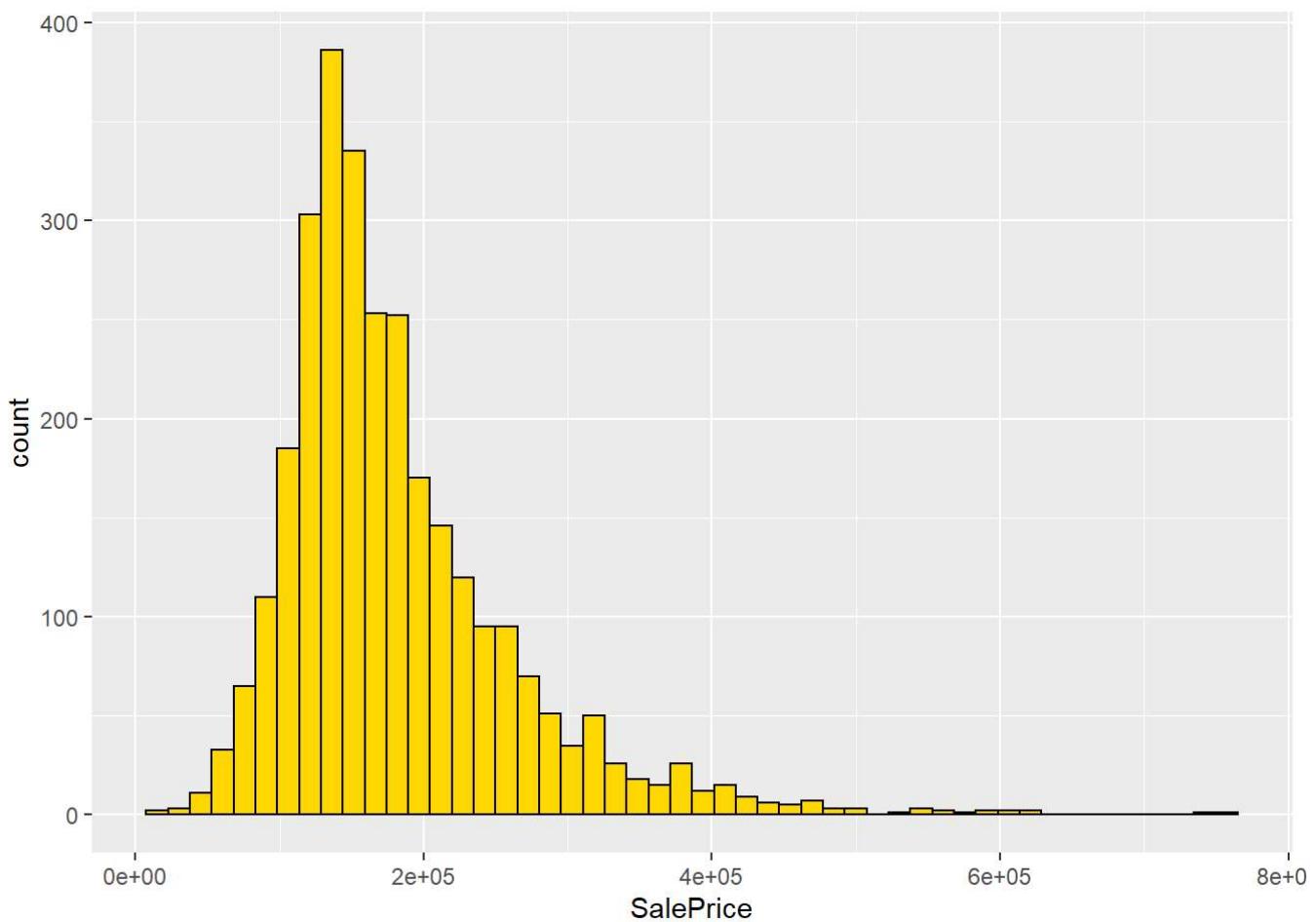
```
#no of house having ac
barplot(table(ds$Central.Air))
```



```
#no of house having good exposure  
barplot(table(ds$Bsmt.Exposure))
```

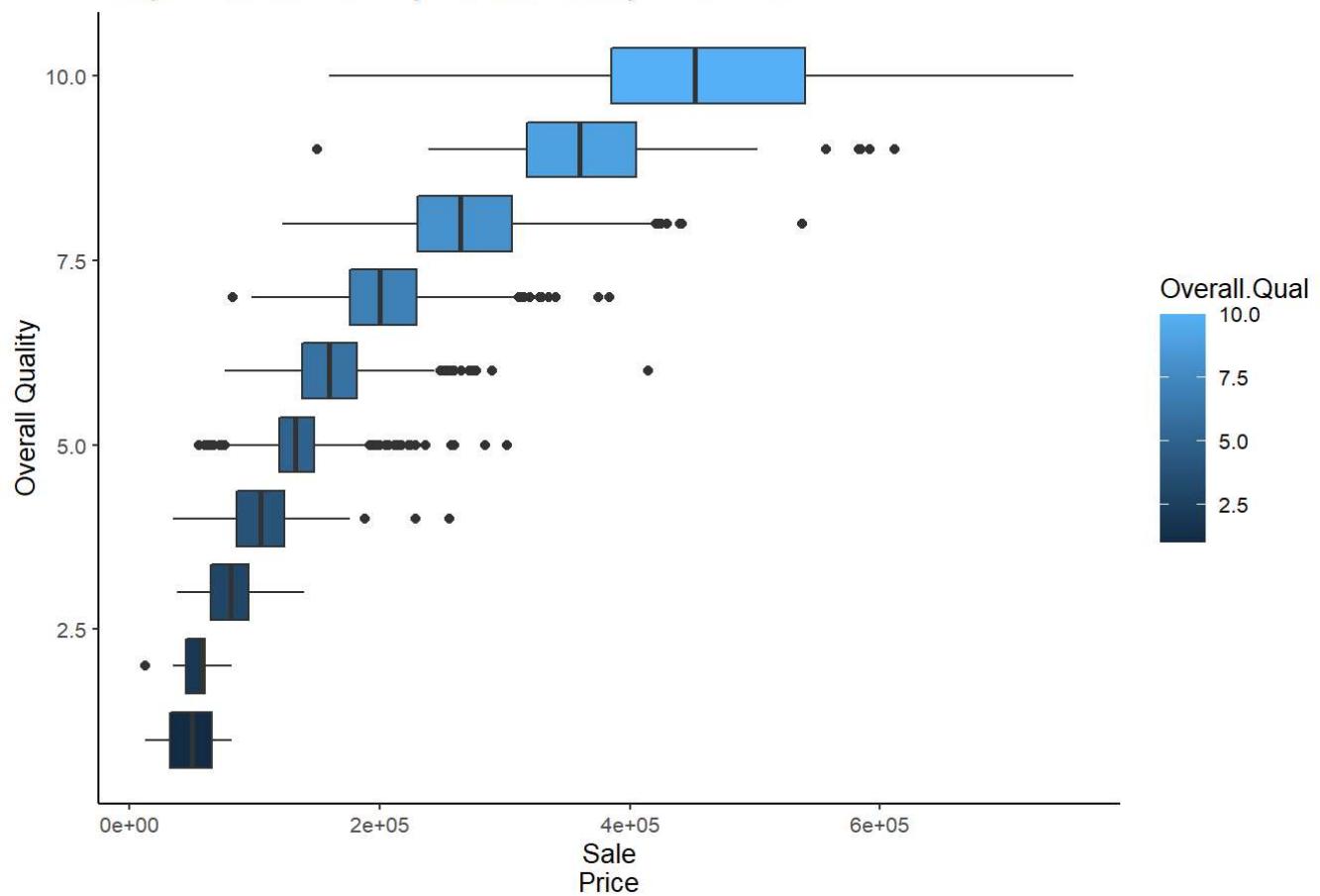


```
#plotting sales price
ggplot(ds, aes(x = SalePrice)) +
  geom_histogram(bins = 50, fill="gold",color="black")
```



```
ggplot(ds,aes(SalePrice, Overall.Qual , group=Overall.Qual, fill=Overall.Qual))+geom_boxplot()
() +labs(title = "Boxplot of Sale Price by Overall Quality of the houses", x="Sale
Price", y="Overall Quality") +theme_classic(base_size = 10)
```

Boxplot of Sale Price by Overall Quality of the houses



```

#3# missing data being filled with mean
ds$Lot.Frontage[is.na(ds$Lot.Frontage)] <-
  mean(ds$Lot.Frontage, na.rm = T)

ds$Mas.Vnr.Area[is.na(ds$Mas.Vnr.Area)] <-
  mean(ds$Mas.Vnr.Area, na.rm = T)

ds$Garage.Yr.Blt[is.na(ds$Garage.Yr.Blt)] <-
  mean(ds$Garage.Yr.Blt, na.rm = T)

ds$Garage.Area[is.na(ds$Garage.Area) == T] <- mean(ds$Garage.Area,
  na.rm=TRUE)

ds$Bsmt.Half.Bath[is.na(ds$Bsmt.Half.Bath) == T] <- mean(ds$Bsmt.Half.Bath,
  na.rm=TRUE)

ds$Bsmt.Full.Bath[is.na(ds$Bsmt.Full.Bath) == T] <- mean(ds$Bsmt.Full.Bath,
  na.rm=T)

ds$Garage.Cars[is.na(ds$Garage.Cars)] <-
  mean(ds$Garage.Cars, na.rm = T)

ds$Total.Bsmt.SF[is.na(ds$Total.Bsmt.SF)] <-
  mean(ds$Total.Bsmt.SF, na.rm = T)

#4#Correlation based on numerical variables.
num_cols <- sapply(ds, is.numeric)
data_num <- ds[,num_cols]

cors <- cor(data_num,use =
  "pairwise")
round(cors,2)

```

##	Order	PID	MS.SubClass	Lot.Frontage	Lot.Area	Overall.Qual
## Order	1.00	0.17	0.01	-0.01	0.03	-0.05
## PID	0.17	1.00	0.00	-0.09	0.03	-0.26
## MS.SubClass	0.01	0.00	1.00	-0.39	-0.20	0.04
## Lot.Frontage	-0.01	-0.09	-0.39	1.00	0.37	0.20
## Lot.Area	0.03	0.03	-0.20	0.37	1.00	0.10
## Overall.Qual	-0.05	-0.26	0.04	0.20	0.10	1.00
## Overall.Cond	-0.01	0.10	-0.07	-0.07	-0.03	-0.09
## Year.Built	-0.05	-0.34	0.04	0.12	0.02	0.60
## Year.Remod.Add	-0.08	-0.16	0.04	0.09	0.02	0.57
## Mas.Vnr.Area	-0.03	-0.23	0.00	0.20	0.13	0.43
## BsmtFin.SF.1	-0.03	-0.10	-0.06	0.20	0.19	0.28
## BsmtFin.SF.2	0.00	0.00	-0.07	0.04	0.08	-0.04
## Bsmt.Unf.SF	0.01	-0.09	-0.13	0.11	0.02	0.27
## Total.Bsmt.SF	-0.03	-0.19	-0.22	0.33	0.25	0.55
## 1st.Flr.SF	-0.01	-0.14	-0.25	0.42	0.33	0.48
## 2nd.Flr.SF	0.00	0.00	0.30	0.03	0.03	0.24
## Low.Qual.Fin.SF	0.01	0.06	0.03	0.01	0.00	-0.05
## Gr.Liv.Area	-0.01	-0.11	0.07	0.35	0.29	0.57
## Bsmt.Full.Bath	-0.04	-0.04	0.01	0.10	0.13	0.17
## Bsmt.Half.Bath	0.02	0.00	0.00	-0.02	0.03	-0.04
## Full.Bath	-0.04	-0.17	0.13	0.17	0.13	0.52
## Half.Bath	-0.04	-0.17	0.18	0.04	0.04	0.27
## Bedroom.AbvGr	0.02	0.01	-0.02	0.22	0.14	0.06
## Kitchen.AbvGr	-0.02	0.08	0.26	0.01	-0.02	-0.16
## TotRms.AbvGrd	0.00	-0.07	0.03	0.32	0.22	0.38
## Fireplaces	-0.02	-0.11	-0.05	0.23	0.26	0.39
## Garage.Yr.Blt	-0.05	-0.25	0.08	0.07	-0.01	0.54
## Garage.Cars	-0.04	-0.24	-0.05	0.29	0.18	0.60
## Garage.Area	-0.04	-0.21	-0.10	0.34	0.21	0.56
## Wood.Deck.SF	-0.01	-0.05	-0.02	0.10	0.16	0.26
## Open.Porch.SF	0.02	-0.07	-0.01	0.15	0.10	0.30
## Enclosed.Porch	0.03	0.16	-0.02	0.01	0.02	-0.14
## 3-Ssn.Porch	-0.02	-0.02	-0.04	0.03	0.02	0.02
## Screen.Porch	0.00	-0.03	-0.05	0.07	0.06	0.04
## Pool.Area	0.05	0.00	0.00	0.16	0.09	0.03
## Misc.Val	-0.01	-0.01	-0.03	0.04	0.07	0.01
## Mo.Sold	0.13	-0.05	0.00	0.01	0.00	0.03
## Yr.Sold	-0.98	0.01	-0.02	-0.01	-0.02	-0.02
## SalePrice	-0.03	-0.25	-0.09	0.34	0.27	0.80
##		Overall.Cond	Year.Built	Year.Remod.Add	Mas.Vnr.Area	
## Order		-0.01	-0.05	-0.08	-0.03	
## PID		0.10	-0.34	-0.16	-0.23	
## MS.SubClass		-0.07	0.04	0.04	0.00	
## Lot.Frontage		-0.07	0.12	0.09	0.20	
## Lot.Area		-0.03	0.02	0.02	0.13	
## Overall.Qual		-0.09	0.60	0.57	0.43	
## Overall.Cond		1.00	-0.37	0.05	-0.14	
## Year.Built		-0.37	1.00	0.61	0.31	
## Year.Remod.Add		0.05	0.61	1.00	0.20	
## Mas.Vnr.Area		-0.14	0.31	0.20	1.00	
## BsmtFin.SF.1		-0.05	0.28	0.15	0.30	
## BsmtFin.SF.2		0.04	-0.03	-0.06	-0.02	
## Bsmt.Unf.SF		-0.14	0.13	0.16	0.09	
## Total.Bsmt.SF		-0.17	0.41	0.30	0.40	

## 1st.Flr.SF	-0.16	0.31	0.24	0.39	
## 2nd.Flr.SF	0.01	0.02	0.16	0.12	
## Low.Qual.Fin.SF	0.01	-0.14	-0.06	-0.06	
## Gr.Liv.Area	-0.12	0.24	0.32	0.40	
## Bsmt.Full.Bath	-0.04	0.21	0.13	0.14	
## Bsmt.Half.Bath	0.08	-0.03	-0.05	0.02	
## Full.Bath	-0.21	0.47	0.46	0.26	
## Half.Bath	-0.09	0.27	0.21	0.19	
## Bedroom.AbvGr	-0.01	-0.06	-0.02	0.08	
## Kitchen.AbvGr	-0.09	-0.14	-0.14	-0.05	
## TotRms.AbvGrd	-0.09	0.11	0.20	0.28	
## Fireplaces	-0.03	0.17	0.13	0.27	
## Garage.Yr.Blt	-0.30	0.79	0.63	0.25	
## Garage.Cars	-0.18	0.54	0.43	0.36	
## Garage.Area	-0.15	0.48	0.38	0.37	
## Wood.Deck.SF	0.02	0.23	0.22	0.17	
## Open.Porch.SF	-0.07	0.20	0.24	0.14	
## Enclosed.Porch	0.07	-0.37	-0.22	-0.11	
## 3-Ssn.Porch	0.04	0.02	0.04	0.01	
## Screen.Porch	0.04	-0.04	-0.05	0.07	
## Pool.Area	-0.02	0.00	-0.01	0.00	
## Misc.Val	0.03	-0.01	0.00	0.04	
## Mo.Sold	-0.01	0.01	0.02	0.00	
## Yr.Sold	0.03	-0.01	0.03	-0.02	
## SalePrice	-0.10	0.56	0.53	0.51	
##	BsmtFin_SF.1	BsmtFin_SF.2	Bsmt.Unf.SF	Total.Bsmt.SF	1st.Flr.SF
## Order	-0.03	0.00	0.01	-0.03	-0.01
## PID	-0.10	0.00	-0.09	-0.19	-0.14
## MS.SubClass	-0.06	-0.07	-0.13	-0.22	-0.25
## Lot.Frontage	0.20	0.04	0.11	0.33	0.42
## Lot.Area	0.19	0.08	0.02	0.25	0.33
## Overall.Qual	0.28	-0.04	0.27	0.55	0.48
## Overall.Cond	-0.05	0.04	-0.14	-0.17	-0.16
## Year.Built	0.28	-0.03	0.13	0.41	0.31
## Year.Remod.Add	0.15	-0.06	0.16	0.30	0.24
## Mas.Vnr.Area	0.30	-0.02	0.09	0.40	0.39
## BsmtFin_SF.1	1.00	-0.05	-0.48	0.54	0.46
## BsmtFin_SF.2	-0.05	1.00	-0.24	0.09	0.08
## Bsmt.Unf.SF	-0.48	-0.24	1.00	0.41	0.30
## Total.Bsmt.SF	0.54	0.09	0.41	1.00	0.80
## 1st.Flr.SF	0.46	0.08	0.30	0.80	1.00
## 2nd.Flr.SF	-0.16	-0.10	0.00	-0.21	-0.25
## Low.Qual.Fin.SF	-0.07	0.00	0.05	-0.02	-0.01
## Gr.Liv.Area	0.21	-0.02	0.24	0.44	0.56
## Bsmt.Full.Bath	0.64	0.16	-0.40	0.33	0.26
## Bsmt.Half.Bath	0.08	0.10	-0.11	0.01	0.01
## Full.Bath	0.08	-0.08	0.27	0.32	0.37
## Half.Bath	-0.01	-0.03	-0.03	-0.06	-0.10
## Bedroom.AbvGr	-0.12	-0.03	0.19	0.05	0.11
## Kitchen.AbvGr	-0.09	-0.04	0.07	-0.04	0.08
## TotRms.AbvGrd	0.05	-0.05	0.25	0.28	0.39
## Fireplaces	0.30	0.07	0.00	0.33	0.41
## Garage.Yr.Blt	0.19	-0.07	0.17	0.34	0.25
## Garage.Cars	0.26	-0.01	0.18	0.44	0.44
## Garage.Area	0.31	0.00	0.16	0.49	0.49
## Wood.Deck.SF	0.22	0.10	-0.04	0.23	0.23

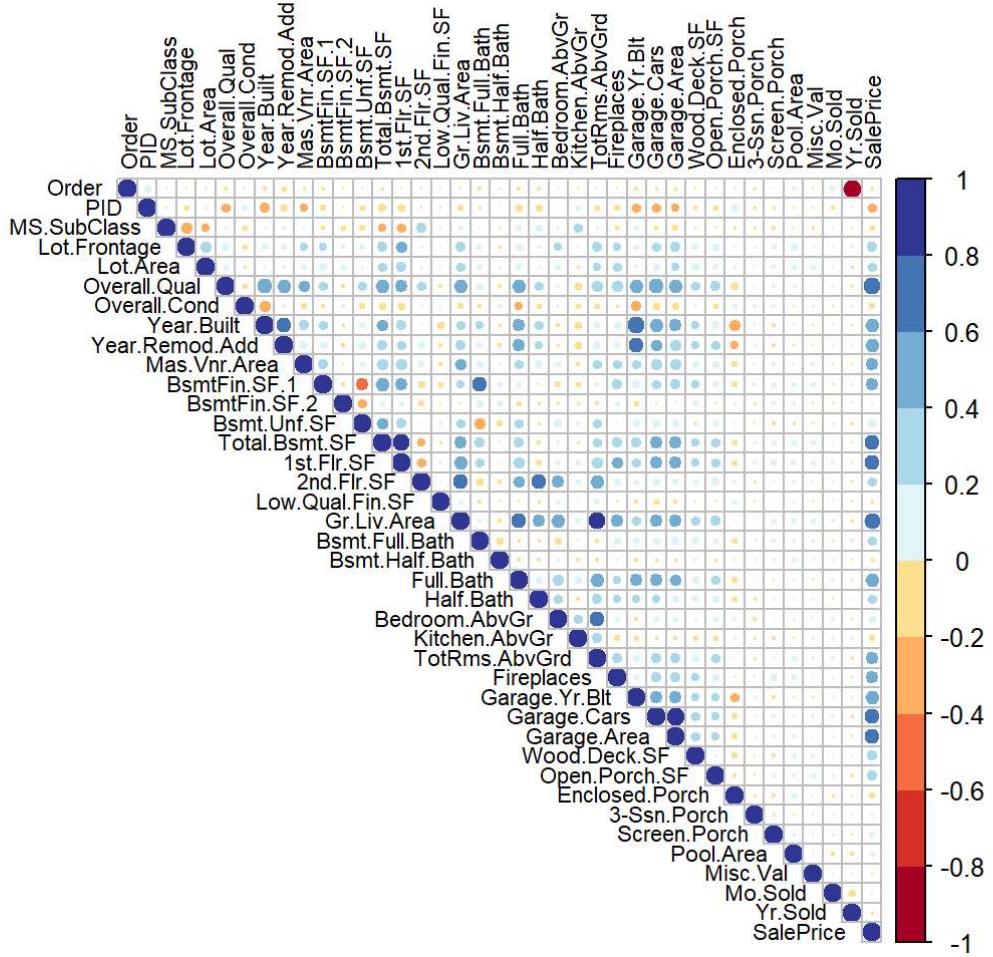
## Open.Porch.SF	0.12	-0.01	0.12	0.25	0.24
## Enclosed.Porch	-0.10	0.03	0.01	-0.09	-0.07
## 3-Ssn.Porch	0.05	-0.02	-0.01	0.04	0.04
## Screen.Porch	0.10	0.06	-0.05	0.08	0.10
## Pool.Area	0.08	0.04	-0.03	0.07	0.12
## Misc.Val	0.09	-0.01	-0.01	0.08	0.09
## Mo.Sold	0.00	-0.01	0.02	0.02	0.04
## Yr.Sold	0.02	0.01	-0.04	-0.01	-0.01
## SalePrice	0.43	0.01	0.18	0.63	0.62
##	2nd.Flr.SF	Low.Qual.Fin.SF	Gr.Liv.Area	Bsmt.Full.Bath	
## Order	0.00	0.01	-0.01	-0.04	
## PID	0.00	0.06	-0.11	-0.04	
## MS.SubClass	0.30	0.03	0.07	0.01	
## Lot.Frontage	0.03	0.01	0.35	0.10	
## Lot.Area	0.03	0.00	0.29	0.13	
## Overall.Qual	0.24	-0.05	0.57	0.17	
## Overall.Cond	0.01	0.01	-0.12	-0.04	
## Year.Built	0.02	-0.14	0.24	0.21	
## Year.Remod.Add	0.16	-0.06	0.32	0.13	
## Mas.Vnr.Area	0.12	-0.06	0.40	0.14	
## BsmtFin.SF.1	-0.16	-0.07	0.21	0.64	
## BsmtFin.SF.2	-0.10	0.00	-0.02	0.16	
## Bsmt.Unf.SF	0.00	0.05	0.24	-0.40	
## Total.Bsmt.SF	-0.21	-0.02	0.44	0.33	
## 1st.Flr.SF	-0.25	-0.01	0.56	0.26	
## 2nd.Flr.SF	1.00	0.02	0.66	-0.16	
## Low.Qual.Fin.SF	0.02	1.00	0.10	-0.05	
## Gr.Liv.Area	0.66	0.10	1.00	0.06	
## Bsmt.Full.Bath	-0.16	-0.05	0.06	1.00	
## Bsmt.Half.Bath	-0.06	-0.01	-0.04	-0.15	
## Full.Bath	0.40	0.00	0.63	-0.02	
## Half.Bath	0.61	-0.04	0.43	-0.04	
## Bedroom.AbvGr	0.50	0.07	0.52	-0.16	
## Kitchen.AbvGr	0.07	0.00	0.12	-0.02	
## TotRms.AbvGrd	0.59	0.10	0.81	-0.04	
## Fireplaces	0.17	-0.01	0.45	0.17	
## Garage.Yr.Blt	0.08	-0.05	0.26	0.14	
## Garage.Cars	0.18	-0.07	0.49	0.16	
## Garage.Area	0.13	-0.05	0.48	0.18	
## Wood.Deck.SF	0.09	-0.02	0.25	0.19	
## Open.Porch.SF	0.18	0.00	0.34	0.08	
## Enclosed.Porch	0.06	0.09	0.00	-0.07	
## 3-Ssn.Porch	-0.03	0.00	0.01	0.03	
## Screen.Porch	0.01	0.01	0.09	0.05	
## Pool.Area	0.04	0.04	0.14	0.04	
## Misc.Val	-0.01	-0.01	0.07	0.00	
## Mo.Sold	0.01	0.01	0.04	0.00	
## Yr.Sold	-0.02	0.00	-0.03	0.04	
## SalePrice	0.27	-0.04	0.71	0.28	
##	Bsmt.Half.Bath	Full.Bath	Half.Bath	Bedroom.AbvGr	Kitchen.AbvGr
## Order	0.02	-0.04	-0.04	0.02	-0.02
## PID	0.00	-0.17	-0.17	0.01	0.08
## MS.SubClass	0.00	0.13	0.18	-0.02	0.26
## Lot.Frontage	-0.02	0.17	0.04	0.22	0.01
## Lot.Area	0.03	0.13	0.04	0.14	-0.02
## Overall.Qual	-0.04	0.52	0.27	0.06	-0.16

## Overall.Cond	0.08	-0.21	-0.09	-0.01	-0.09
## Year.Built	-0.03	0.47	0.27	-0.06	-0.14
## Year.Remod.Add	-0.05	0.46	0.21	-0.02	-0.14
## Mas.Vnr.Area	0.02	0.26	0.19	0.08	-0.05
## BsmtFin.SF.1	0.08	0.08	-0.01	-0.12	-0.09
## BsmtFin.SF.2	0.10	-0.08	-0.03	-0.03	-0.04
## Bsmt.Unf.SF	-0.11	0.27	-0.03	0.19	0.07
## Total.Bsmt.SF	0.01	0.32	-0.06	0.05	-0.04
## 1st.Flr.SF	0.01	0.37	-0.10	0.11	0.08
## 2nd.Flr.SF	-0.06	0.40	0.61	0.50	0.07
## Low.Qual.Fin.SF	-0.01	0.00	-0.04	0.07	0.00
## Gr.Liv.Area	-0.04	0.63	0.43	0.52	0.12
## Bsmt.Full.Bath	-0.15	-0.02	-0.04	-0.16	-0.02
## Bsmt.Half.Bath	1.00	-0.05	-0.06	0.02	-0.06
## Full.Bath	-0.05	1.00	0.16	0.36	0.17
## Half.Bath	-0.06	0.16	1.00	0.25	-0.04
## Bedroom.AbvGr	0.02	0.36	0.25	1.00	0.24
## Kitchen.AbvGr	-0.06	0.17	-0.04	0.24	1.00
## TotRms.AbvGrd	-0.05	0.53	0.35	0.67	0.29
## Fireplaces	0.04	0.23	0.20	0.08	-0.11
## Garage.Yr.Blt	-0.06	0.48	0.22	-0.04	-0.08
## Garage.Cars	-0.03	0.48	0.23	0.09	-0.04
## Garage.Area	-0.02	0.41	0.18	0.07	-0.06
## Wood.Deck.SF	0.05	0.18	0.12	0.03	-0.09
## Open.Porch.SF	-0.03	0.26	0.18	0.08	-0.07
## Enclosed.Porch	-0.01	-0.12	-0.08	0.05	0.03
## 3-Ssn.Porch	0.03	0.02	-0.02	-0.05	-0.02
## Screen.Porch	0.04	-0.02	0.04	0.01	-0.06
## Pool.Area	0.07	0.03	0.00	0.04	-0.01
## Misc.Val	0.04	-0.01	0.03	0.00	0.03
## Mo.Sold	0.02	0.05	0.00	0.05	0.04
## Yr.Sold	-0.02	0.00	0.00	-0.02	0.04
## SalePrice	-0.04	0.55	0.29	0.14	-0.12
##	TotRms.AbvGrd	Fireplaces	Garage.Yr.Blt	Garage.Cars	Garage.Area
## Order	0.00	-0.02	-0.05	-0.04	-0.04
## PID	-0.07	-0.11	-0.25	-0.24	-0.21
## MS.SubClass	0.03	-0.05	0.08	-0.05	-0.10
## Lot.Frontage	0.32	0.23	0.07	0.29	0.34
## Lot.Area	0.22	0.26	-0.01	0.18	0.21
## Overall.Qual	0.38	0.39	0.54	0.60	0.56
## Overall.Cond	-0.09	-0.03	-0.30	-0.18	-0.15
## Year.Built	0.11	0.17	0.79	0.54	0.48
## Year.Remod.Add	0.20	0.13	0.63	0.43	0.38
## Mas.Vnr.Area	0.28	0.27	0.25	0.36	0.37
## BsmtFin.SF.1	0.05	0.30	0.19	0.26	0.31
## BsmtFin.SF.2	-0.05	0.07	-0.07	-0.01	0.00
## Bsmt.Unf.SF	0.25	0.00	0.17	0.18	0.16
## Total.Bsmt.SF	0.28	0.33	0.34	0.44	0.49
## 1st.Flr.SF	0.39	0.41	0.25	0.44	0.49
## 2nd.Flr.SF	0.59	0.17	0.08	0.18	0.13
## Low.Qual.Fin.SF	0.10	-0.01	-0.05	-0.07	-0.05
## Gr.Liv.Area	0.81	0.45	0.26	0.49	0.48
## Bsmt.Full.Bath	-0.04	0.17	0.14	0.16	0.18
## Bsmt.Half.Bath	-0.05	0.04	-0.06	-0.03	-0.02
## Full.Bath	0.53	0.23	0.48	0.48	0.41
## Half.Bath	0.35	0.20	0.22	0.23	0.18

## Bedroom.AbvGr	0.67	0.08	-0.04	0.09	0.07
## Kitchen.AbvGr	0.29	-0.11	-0.08	-0.04	-0.06
## TotRms.AbvGrd	1.00	0.30	0.15	0.36	0.33
## Fireplaces	0.30	1.00	0.09	0.32	0.29
## Garage.Yr.Blt	0.15	0.09	1.00	0.49	0.47
## Garage.Cars	0.36	0.32	0.49	1.00	0.89
## Garage.Area	0.33	0.29	0.47	0.89	1.00
## Wood.Deck.SF	0.15	0.23	0.22	0.24	0.24
## Open.Porch.SF	0.24	0.16	0.22	0.20	0.23
## Enclosed.Porch	0.02	0.00	-0.29	-0.13	-0.11
## 3-Ssn.Porch	-0.03	0.02	0.02	0.02	0.03
## Screen.Porch	0.03	0.17	-0.06	0.04	0.06
## Pool.Area	0.07	0.10	-0.01	0.03	0.05
## Misc.Val	0.06	0.01	-0.01	-0.02	0.01
## Mo.Sold	0.04	0.03	0.02	0.05	0.04
## Yr.Sold	-0.03	-0.01	-0.01	-0.02	-0.01
## SalePrice	0.50	0.47	0.51	0.65	0.64
##	Wood.Deck.SF	Open.Porch.SF	Enclosed.Porch	3-Ssn.Porch	
## Order	-0.01	0.02	0.03	-0.02	
## PID	-0.05	-0.07	0.16	-0.02	
## MS.SubClass	-0.02	-0.01	-0.02	-0.04	
## Lot.Frontage	0.10	0.15	0.01	0.03	
## Lot.Area	0.16	0.10	0.02	0.02	
## Overall.Qual	0.26	0.30	-0.14	0.02	
## Overall.Cond	0.02	-0.07	0.07	0.04	
## Year.Built	0.23	0.20	-0.37	0.02	
## Year.Remod.Add	0.22	0.24	-0.22	0.04	
## Mas.Vnr.Area	0.17	0.14	-0.11	0.01	
## BsmtFin.SF.1	0.22	0.12	-0.10	0.05	
## BsmtFin.SF.2	0.10	-0.01	0.03	-0.02	
## Bsmt.Unf.SF	-0.04	0.12	0.01	-0.01	
## Total.Bsmt.SF	0.23	0.25	-0.09	0.04	
## 1st.Flr.SF	0.23	0.24	-0.07	0.04	
## 2nd.Flr.SF	0.09	0.18	0.06	-0.03	
## Low.Qual.Fin.SF	-0.02	0.00	0.09	0.00	
## Gr.Liv.Area	0.25	0.34	0.00	0.01	
## Bsmt.Full.Bath	0.19	0.08	-0.07	0.03	
## Bsmt.Half.Bath	0.05	-0.03	-0.01	0.03	
## Full.Bath	0.18	0.26	-0.12	0.02	
## Half.Bath	0.12	0.18	-0.08	-0.02	
## Bedroom.AbvGr	0.03	0.08	0.05	-0.05	
## Kitchen.AbvGr	-0.09	-0.07	0.03	-0.02	
## TotRms.AbvGrd	0.15	0.24	0.02	-0.03	
## Fireplaces	0.23	0.16	0.00	0.02	
## Garage.Yr.Blt	0.22	0.22	-0.29	0.02	
## Garage.Cars	0.24	0.20	-0.13	0.02	
## Garage.Area	0.24	0.23	-0.11	0.03	
## Wood.Deck.SF	1.00	0.04	-0.12	0.00	
## Open.Porch.SF	0.04	1.00	-0.06	-0.01	
## Enclosed.Porch	-0.12	-0.06	1.00	-0.03	
## 3-Ssn.Porch	0.00	-0.01	-0.03	1.00	
## Screen.Porch	-0.05	0.05	-0.06	-0.03	
## Pool.Area	0.09	0.06	0.09	-0.01	
## Misc.Val	0.06	0.08	0.01	0.00	
## Mo.Sold	0.02	0.03	-0.02	0.03	
## Yr.Sold	0.00	-0.04	0.00	0.02	

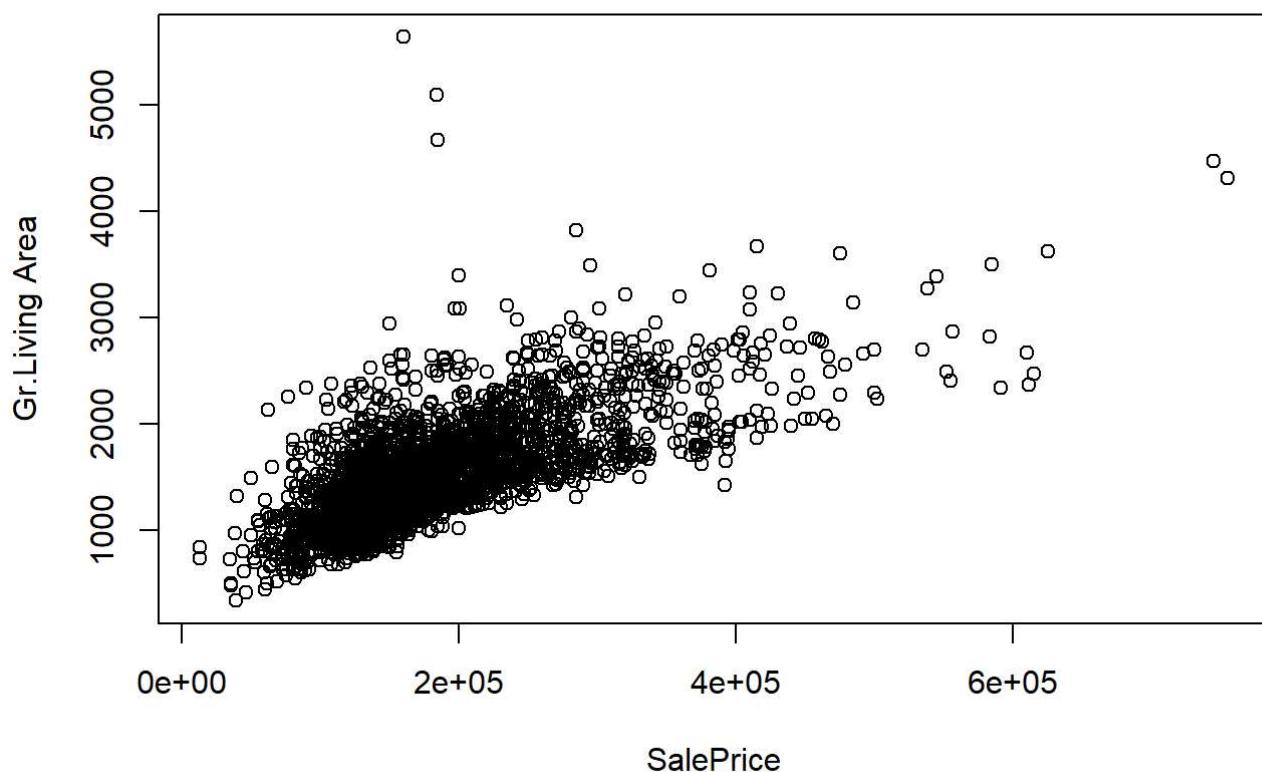
## SalePrice	0.33	0.31	-0.13	0.03		
## Screen.Porch						SalePrice
## Order	0.00	0.05	-0.01	0.13	-0.98	-0.03
## PID	-0.03	0.00	-0.01	-0.05	0.01	-0.25
## MS.SubClass	-0.05	0.00	-0.03	0.00	-0.02	-0.09
## Lot.Frontage	0.07	0.16	0.04	0.01	-0.01	0.34
## Lot.Area	0.06	0.09	0.07	0.00	-0.02	0.27
## Overall.Qual	0.04	0.03	0.01	0.03	-0.02	0.80
## Overall.Cond	0.04	-0.02	0.03	-0.01	0.03	-0.10
## Year.Built	-0.04	0.00	-0.01	0.01	-0.01	0.56
## Year.Remod.Add	-0.05	-0.01	0.00	0.02	0.03	0.53
## Mas.Vnr.Area	0.07	0.00	0.04	0.00	-0.02	0.51
## BsmtFin.SF.1	0.10	0.08	0.09	0.00	0.02	0.43
## BsmtFin.SF.2	0.06	0.04	-0.01	-0.01	0.01	0.01
## Bsmt.Unf.SF	-0.05	-0.03	-0.01	0.02	-0.04	0.18
## Total.Bsmt.SF	0.08	0.07	0.08	0.02	-0.01	0.63
## 1st.Flr.SF	0.10	0.12	0.09	0.04	-0.01	0.62
## 2nd.Flr.SF	0.01	0.04	-0.01	0.01	-0.02	0.27
## Low.Qual.Fin.SF	0.01	0.04	-0.01	0.01	0.00	-0.04
## Gr.Liv.Area	0.09	0.14	0.07	0.04	-0.03	0.71
## Bsmt.Full.Bath	0.05	0.04	0.00	0.00	0.04	0.28
## Bsmt.Half.Bath	0.04	0.07	0.04	0.02	-0.02	-0.04
## Full.Bath	-0.02	0.03	-0.01	0.05	0.00	0.55
## Half.Bath	0.04	0.00	0.03	0.00	0.00	0.29
## Bedroom.AbvGr	0.01	0.04	0.00	0.05	-0.02	0.14
## Kitchen.AbvGr	-0.06	-0.01	0.03	0.04	0.04	-0.12
## TotRms.AbvGrd	0.03	0.07	0.06	0.04	-0.03	0.50
## Fireplaces	0.17	0.10	0.01	0.03	-0.01	0.47
## Garage.Yr.Blt	-0.06	-0.01	-0.01	0.02	-0.01	0.51
## Garage.Cars	0.04	0.03	-0.02	0.05	-0.02	0.65
## Garage.Area	0.06	0.05	0.01	0.04	-0.01	0.64
## Wood.Deck.SF	-0.05	0.09	0.06	0.02	0.00	0.33
## Open.Porch.SF	0.05	0.06	0.08	0.03	-0.04	0.31
## Enclosed.Porch	-0.06	0.09	0.01	-0.02	0.00	-0.13
## 3-Ssn.Porch	-0.03	-0.01	0.00	0.03	0.02	0.03
## Screen.Porch	1.00	0.03	0.01	0.03	-0.01	0.11
## Pool.Area	0.03	1.00	0.01	-0.04	-0.05	0.07
## Misc.Val	0.01	0.01	1.00	0.01	0.01	-0.02
## Mo.Sold	0.03	-0.04	0.01	1.00	-0.16	0.04
## Yr.Sold	-0.01	-0.05	0.01	-0.16	1.00	-0.03
## SalePrice	0.11	0.07	-0.02	0.04	-0.03	1.00

```
#5#Corr plot for all the numerical variables.
corrplot(cors, type = "upper", col =brewer.pal(n=10, name="RdYlBu"), tl.cex = 0.70, tl.col =
'black')
```

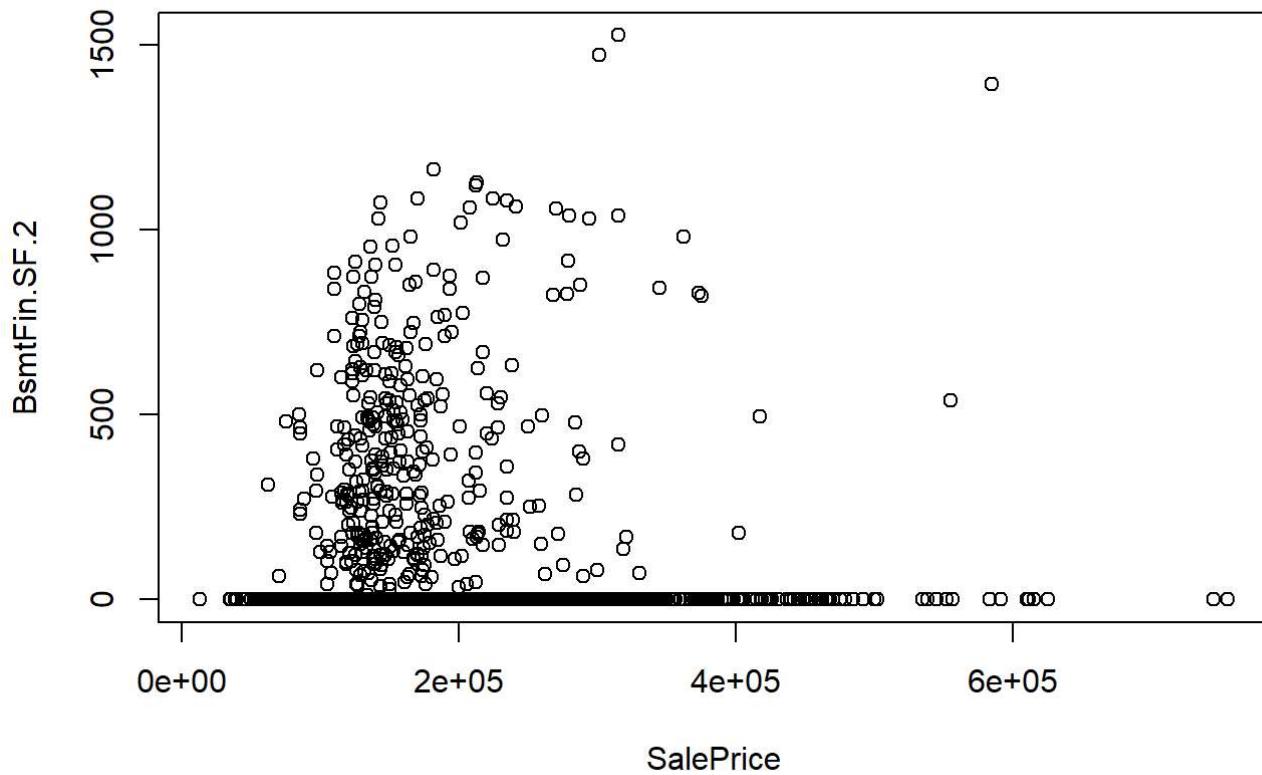


```
#6#pLottting with high, Low & 0.5 correlations
```

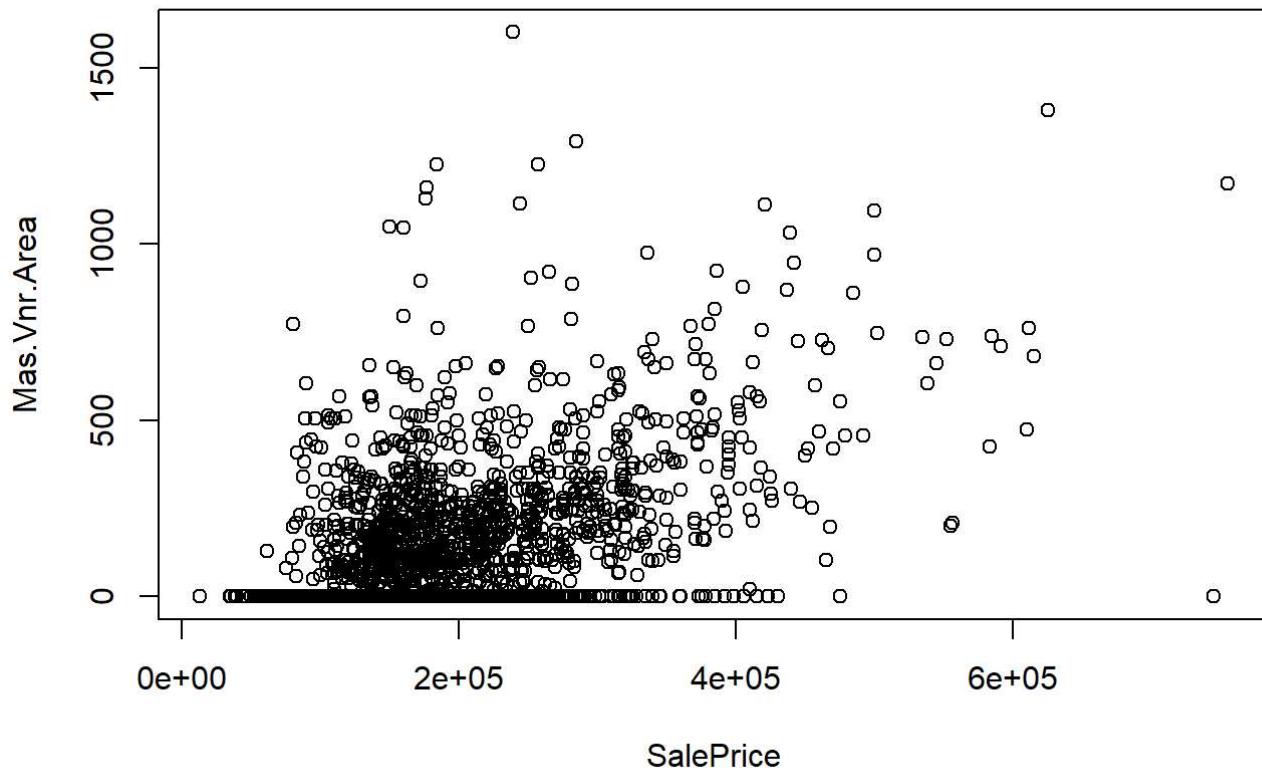
```
#high
plot(data_num$SalePrice, data_num$Gr.Liv.Area, xlab= "SalePrice", ylab= "Gr.Living Area")
```



```
#Low  
plot(data_num$SalePrice, data_num$BsmtFin.SF.2, xlab= "SalePrice", ylab= "BsmtFin.SF.2")
```



```
#0.5  
plot(data_num$SalePrice, data_num$Mas.Vnr.Area, xlab= "SalePrice", ylab= "Mas.Vnr.Area")
```



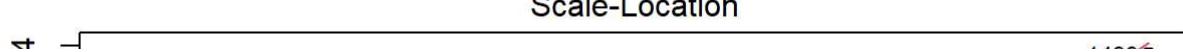
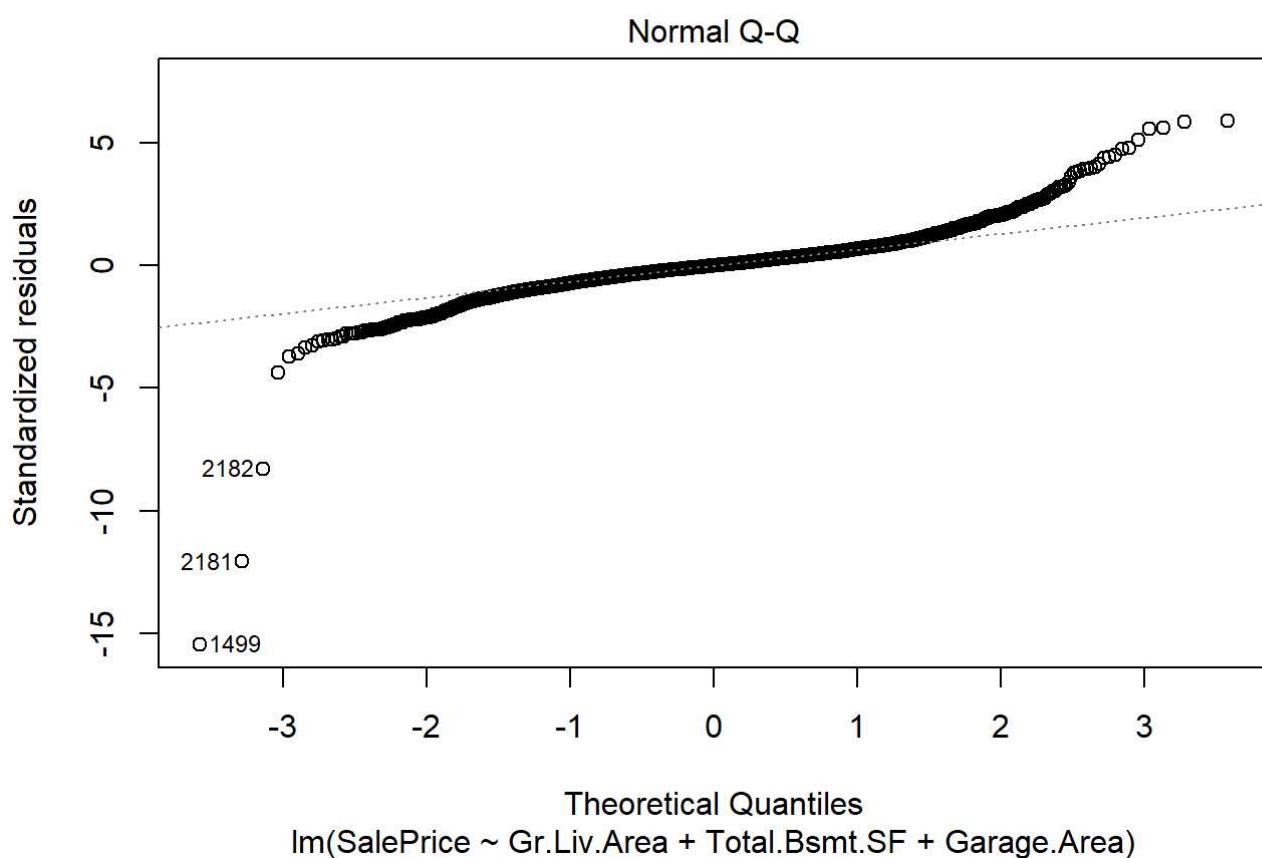
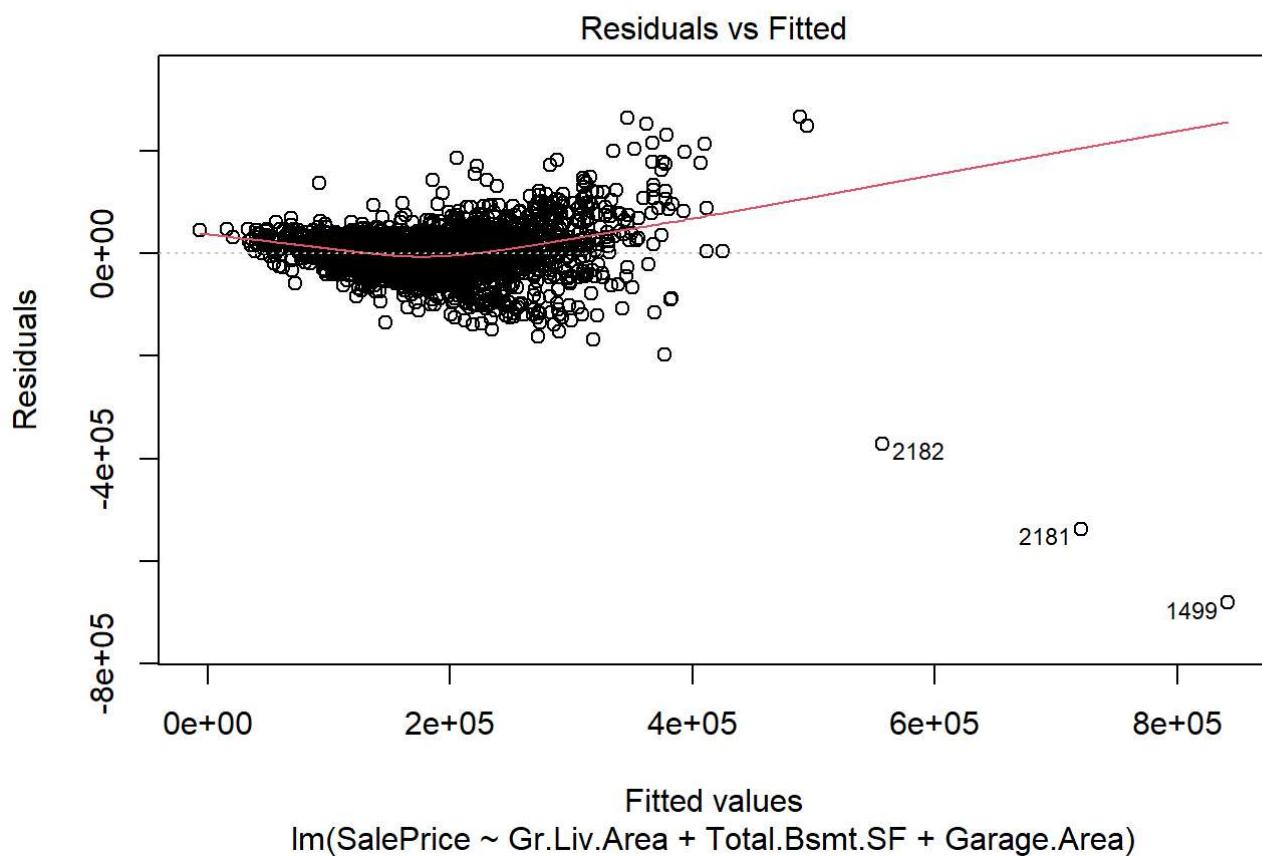
```
#7# fitting our regression model  
fit <- lm(formula = SalePrice ~ Gr.Liv.Area + Total.Bsmt.SF + Garage.Area, data = data_num)  
fit
```

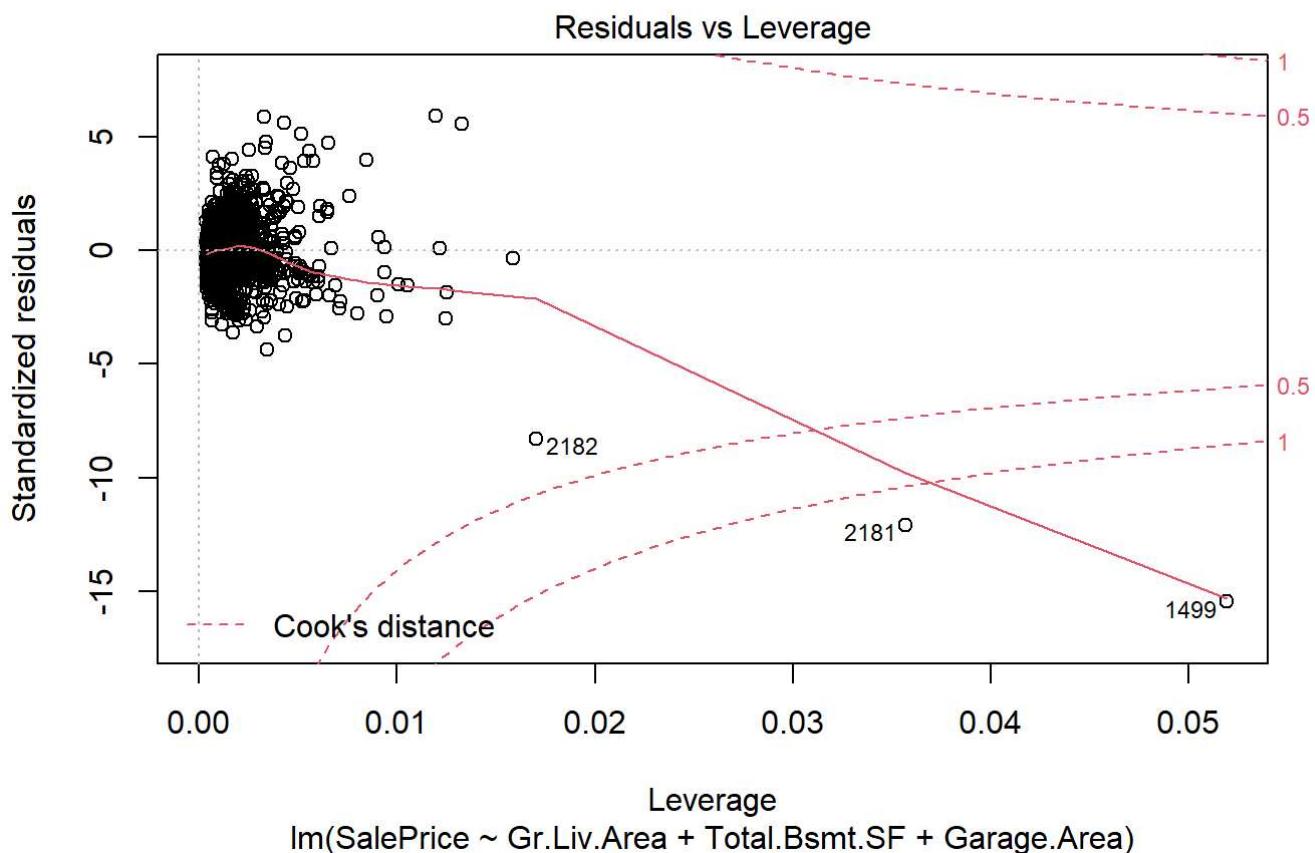
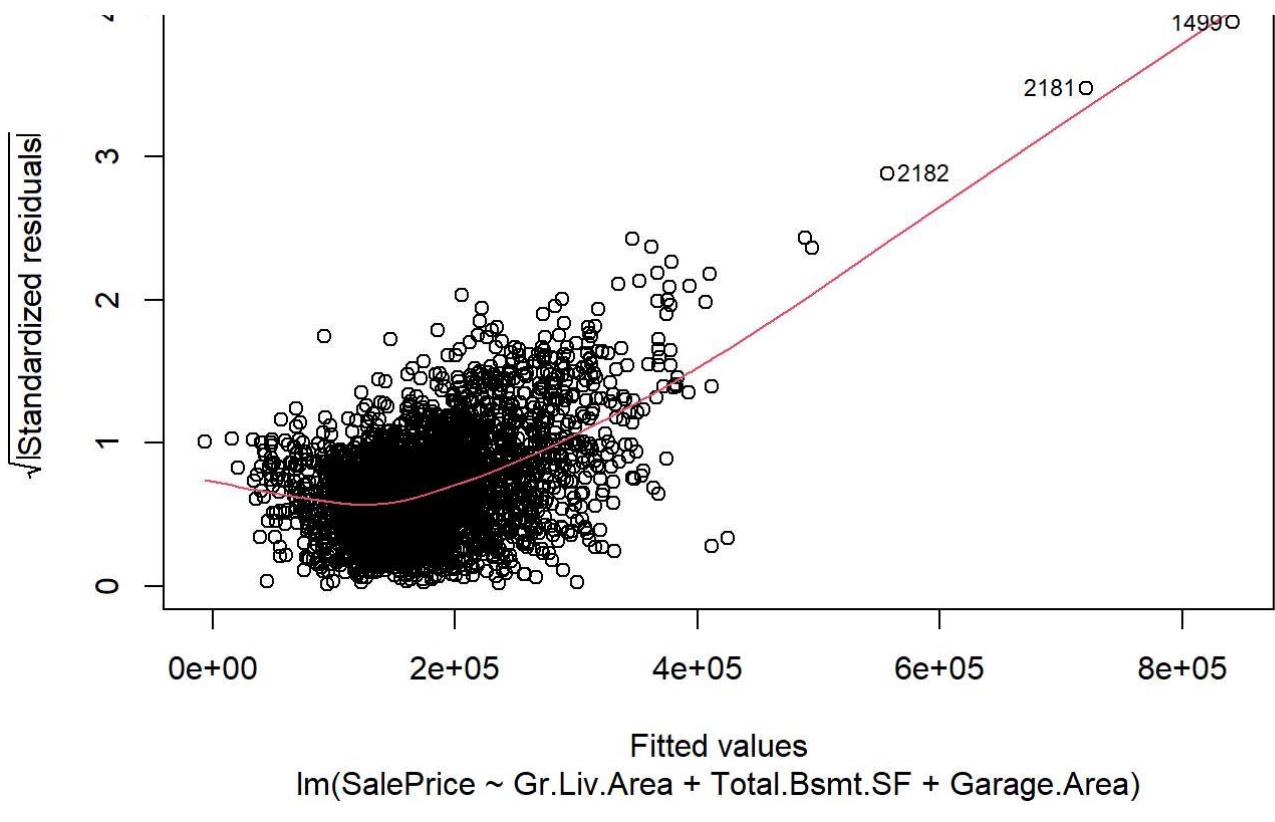
```
##  
## Call:  
## lm(formula = SalePrice ~ Gr.Liv.Area + Total.Bsmt.SF + Garage.Area,  
##      data = data_num)  
##  
## Coefficients:  
##   (Intercept)    Gr.Liv.Area  Total.Bsmt.SF    Garage.Area  
## -29593.64        68.86          54.59        105.14
```

```
summary(fit)
```

```
##  
## Call:  
## lm(formula = SalePrice ~ Gr.Liv.Area + Total.Bsmt.SF + Garage.Area,  
##      data = data_num)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max  
## -681541  -19927     204    19841   266496  
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -29593.644   2830.734  -10.45 <2e-16 ***  
## Gr.Liv.Area     68.862      1.966    35.02 <2e-16 ***  
## Total.Bsmt.SF    54.586      2.257    24.18 <2e-16 ***  
## Garage.Area     105.145      4.736    22.20 <2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 45250 on 2926 degrees of freedom  
## Multiple R-squared:  0.6795, Adjusted R-squared:  0.6791  
## F-statistic:  2068 on 3 and 2926 DF,  p-value: < 2.2e-16
```

```
plot(fit)
```

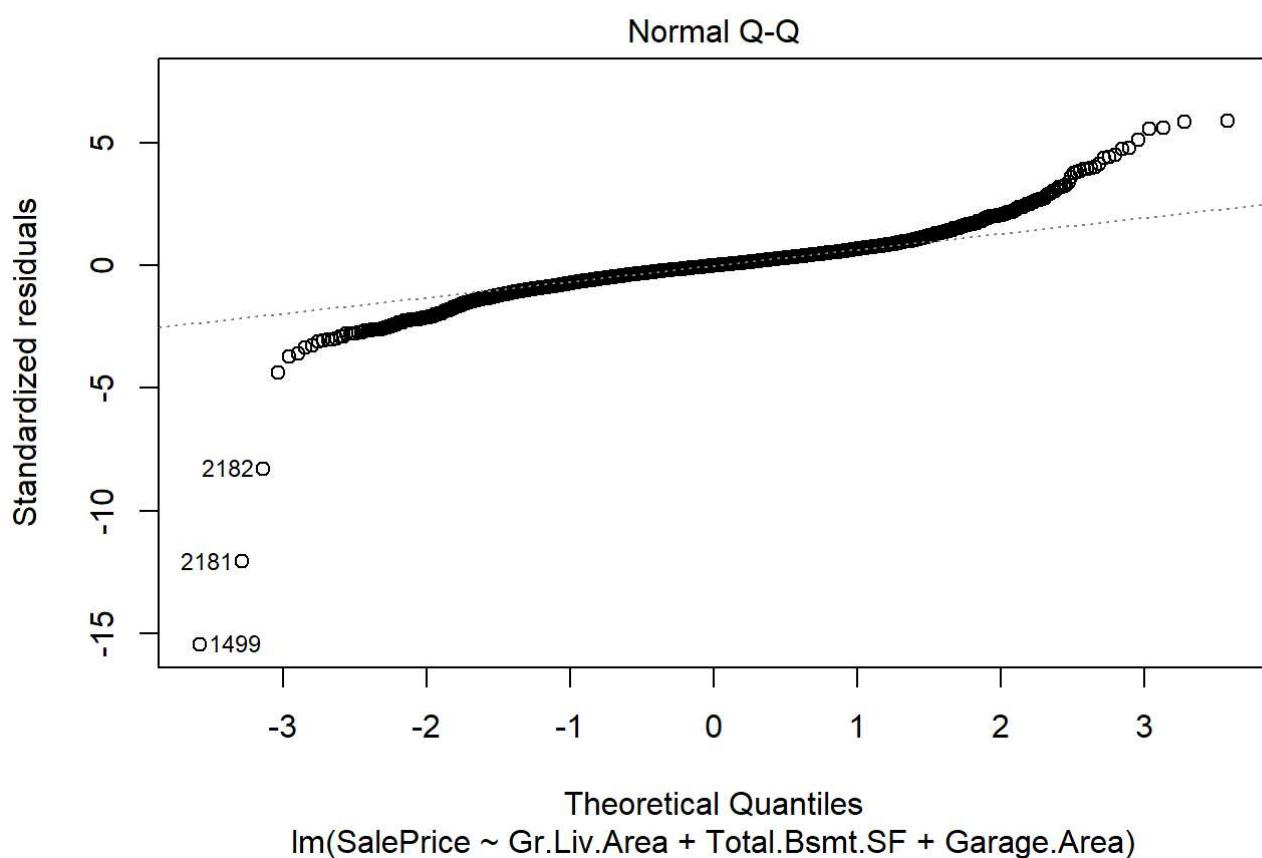
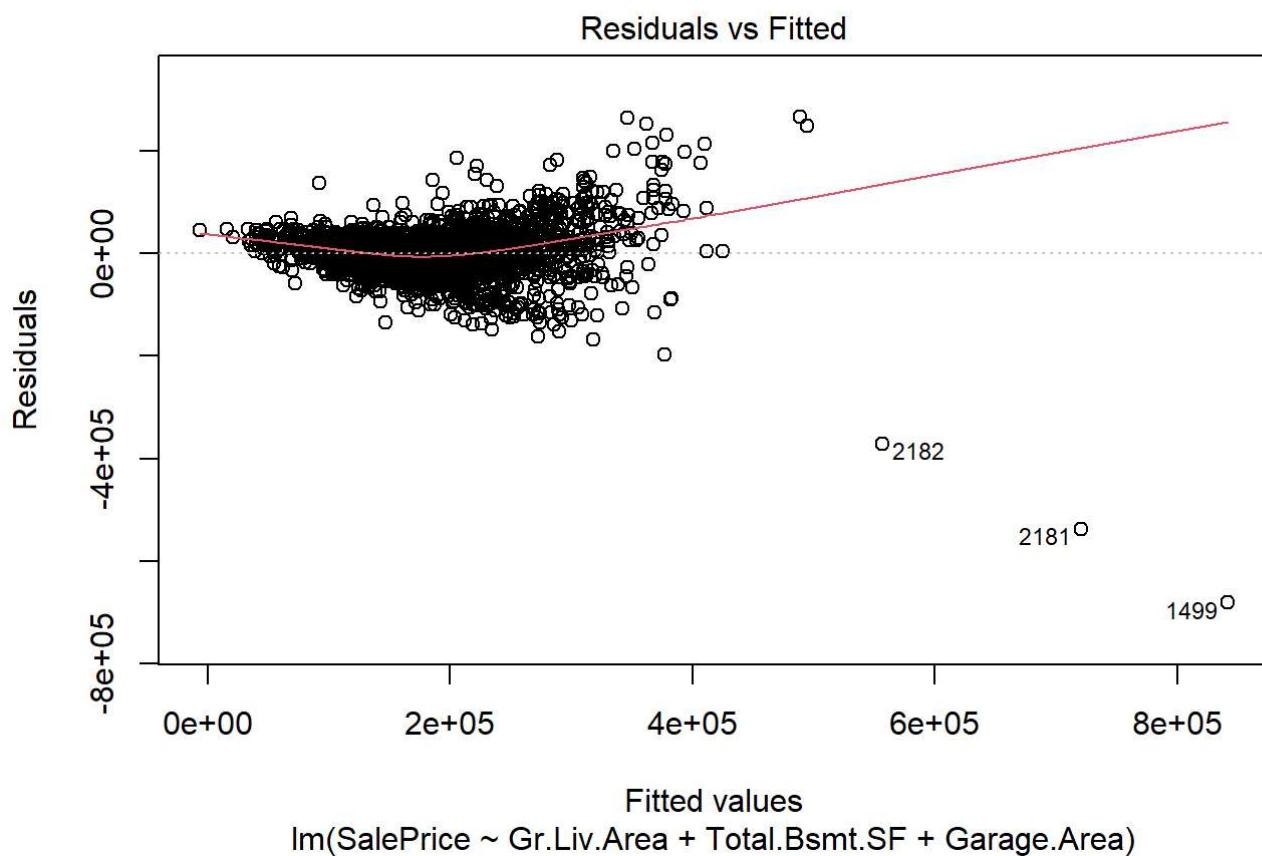


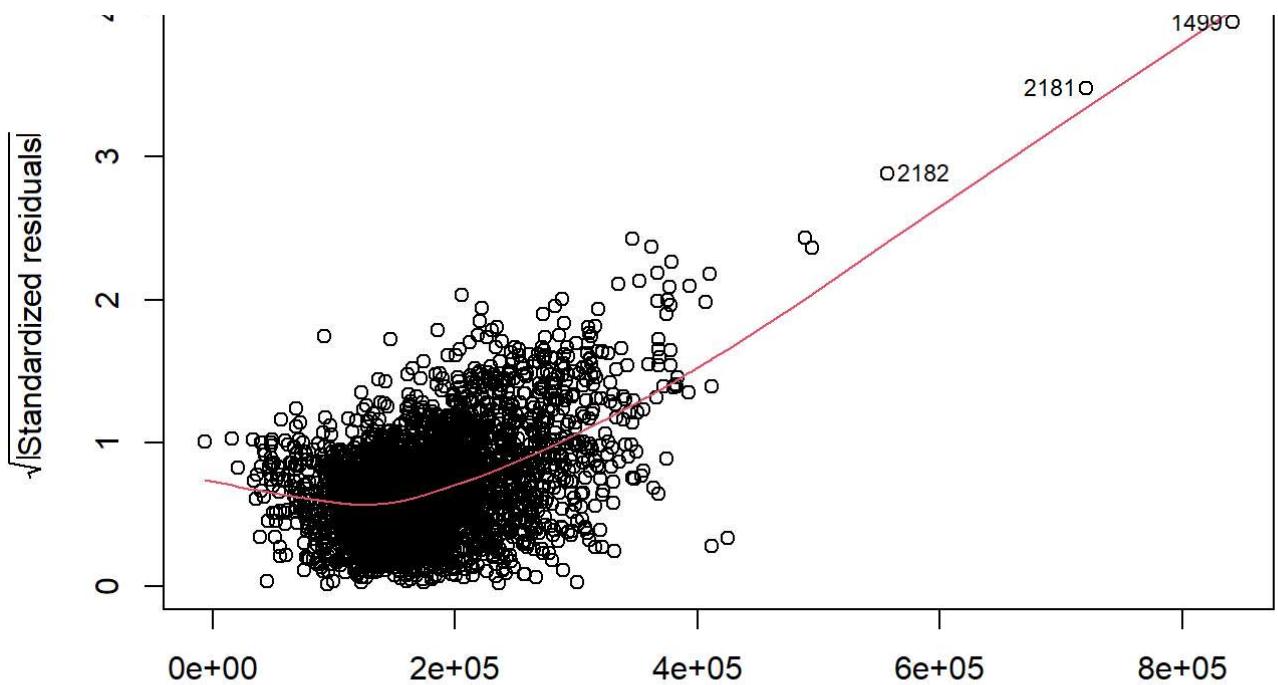


```
#8#
#creating formula for our model

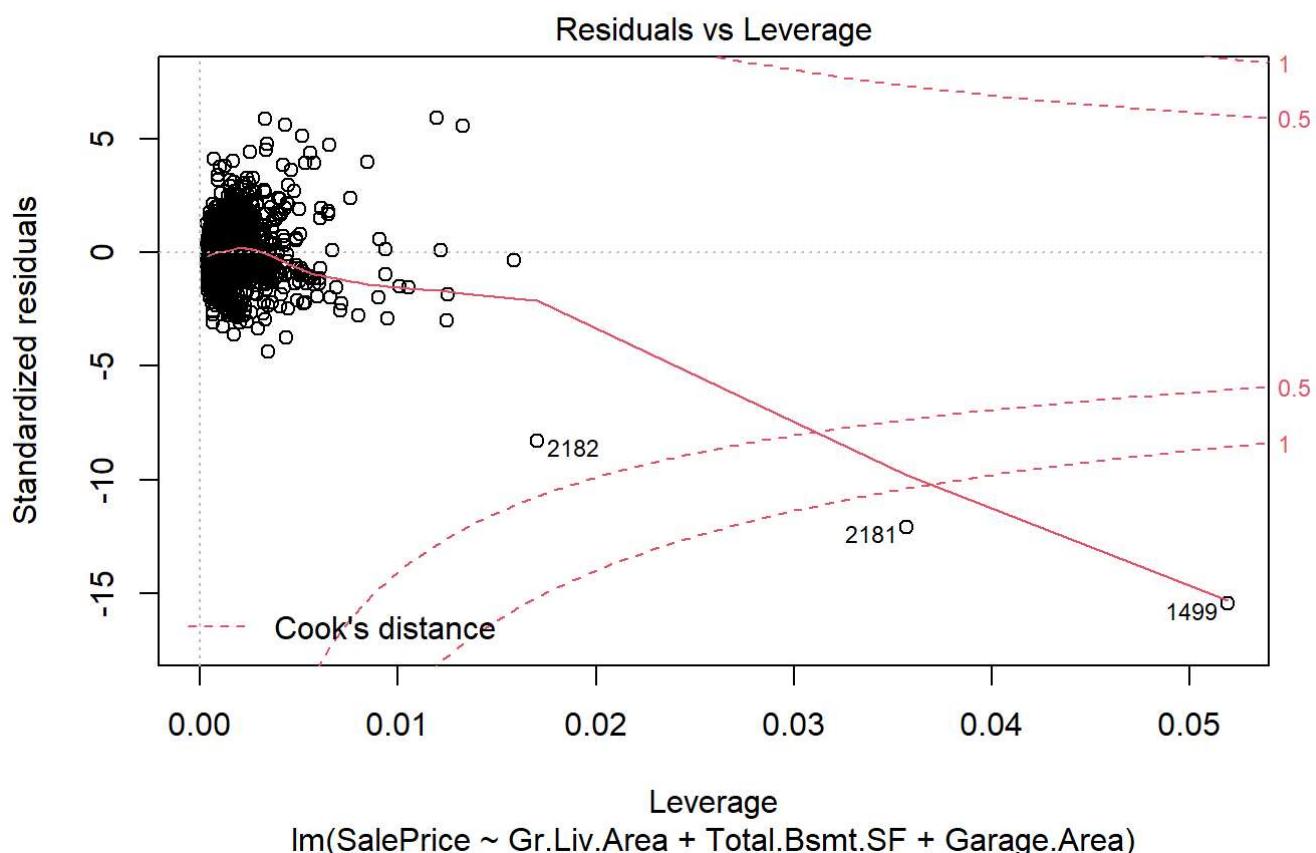
#9# Different plots generated for our model
plot(fit)
```







Fitted values  
 $\text{lm}(\text{SalePrice} \sim \text{Gr.Liv.Area} + \text{Total.Bsmt.SF} + \text{Garage.Area})$



Residuals vs Leverage  
 $\text{lm}(\text{SalePrice} \sim \text{Gr.Liv.Area} + \text{Total.Bsmt.SF} + \text{Garage.Area})$

```
#10# Checking for multicollinearity
#we use check_collinearity function to check for multicollinearity
mc <- check_collinearity(fit)
mc
```

```

## # Check for Multicollinearity
##
## Low Correlation
##
##           Term  VIF Increased SE Tolerance
##     Gr.Liv.Area 1.41      1.19      0.71
## Total.Bsmt.SF 1.41      1.19      0.71
## Garage.Area   1.48      1.22      0.67

```

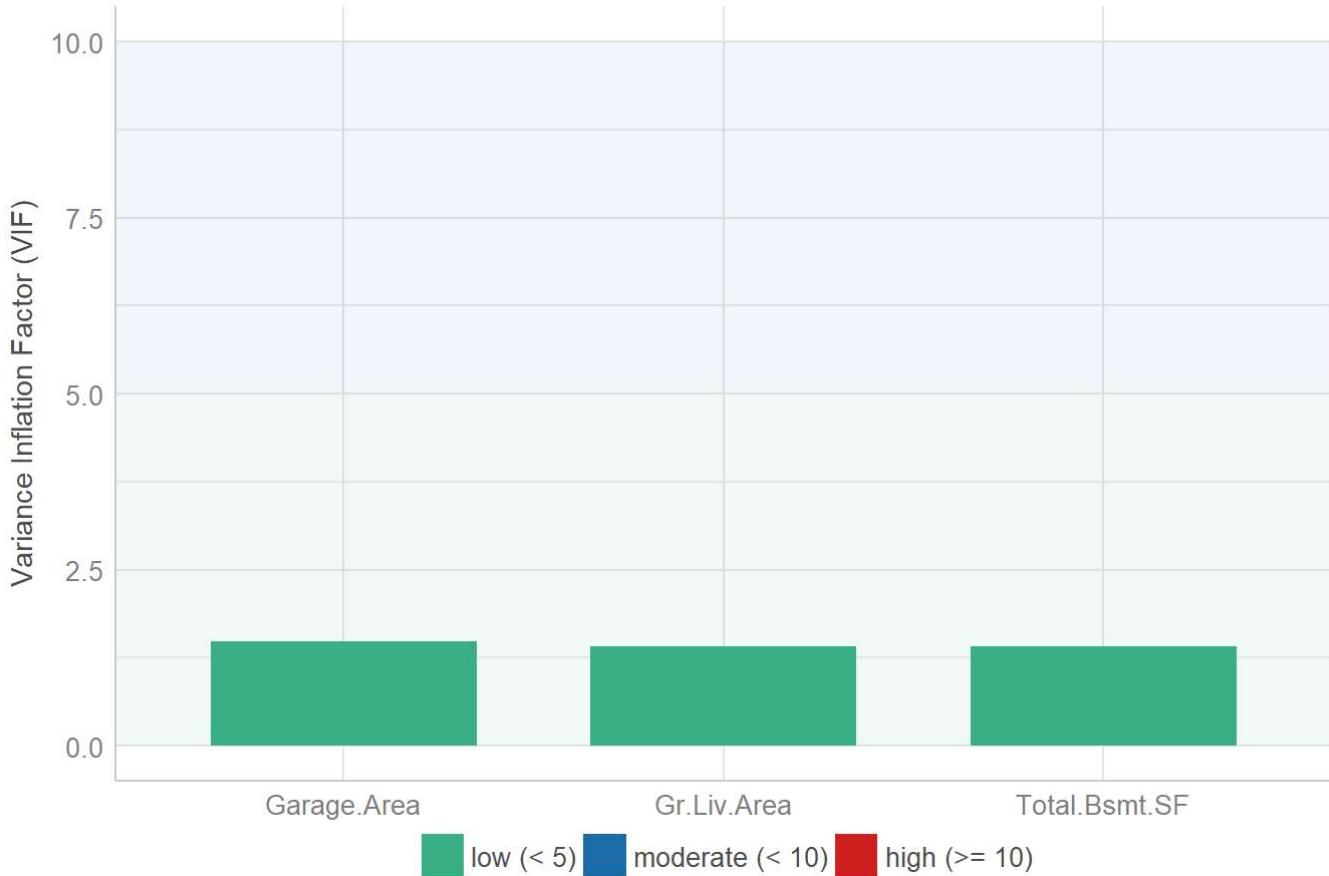
```

#we plot it to see the vif value.
plot(mc)

```

## Collinearity

Higher bars (>5) indicate potential collinearity issues



```
#11# We check for outliers in our model.
```

```

#ar1
array1 <- boxplot.stats(data_num$Gr.Liv.Area)$out
paste0("The minimum value of the outliers is ", min(array1), " The maximum
value of the outliers is ", max(array1))

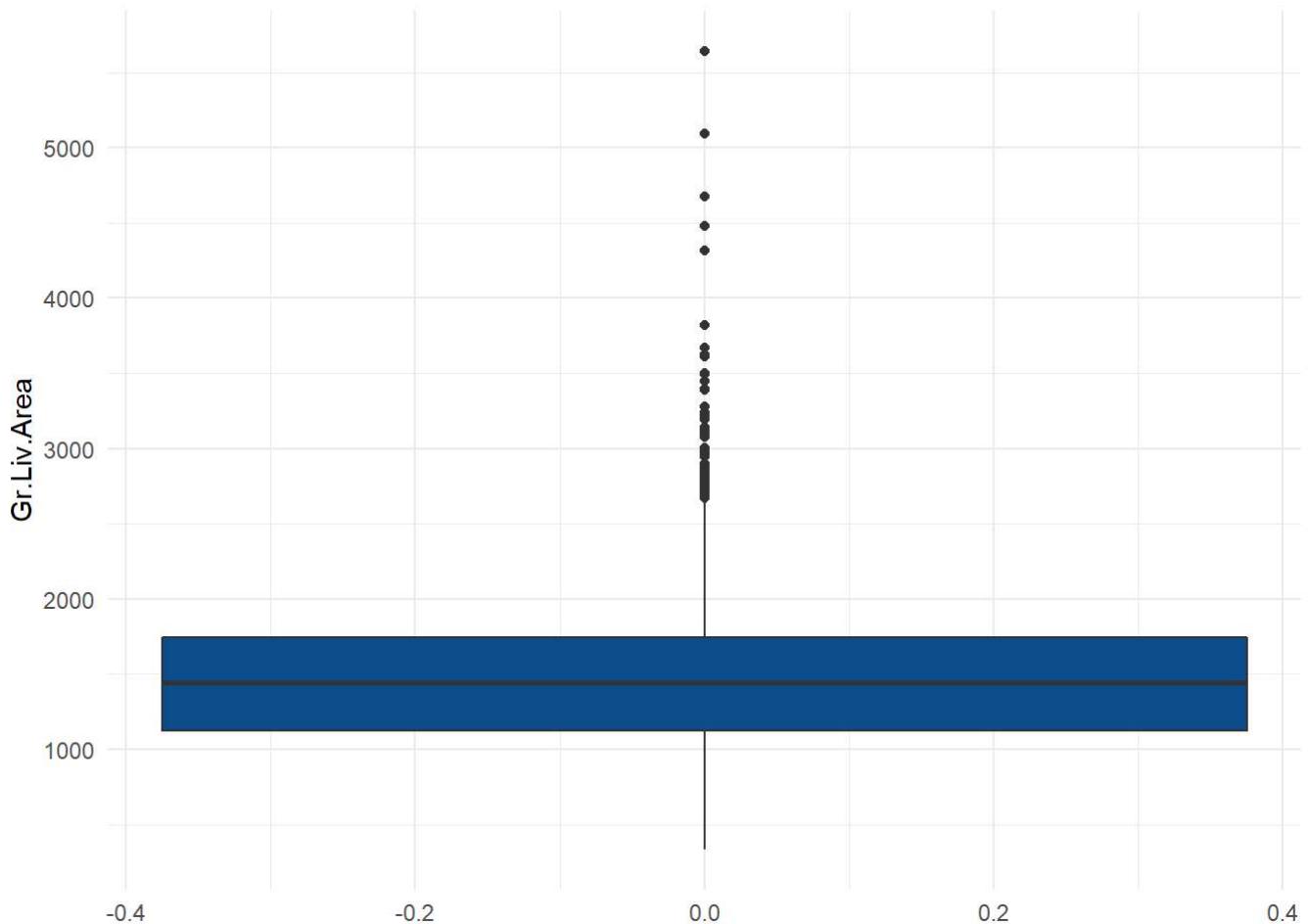
```

```
## [1] "The minimum value of the outliers is 2673 The maximum\nvalue of the outliers is 5642"
```

```

ggplot(data_num, aes(y=Gr.Liv.Area)) + geom_boxplot(fill="#0c4c8a") +
  theme_minimal()

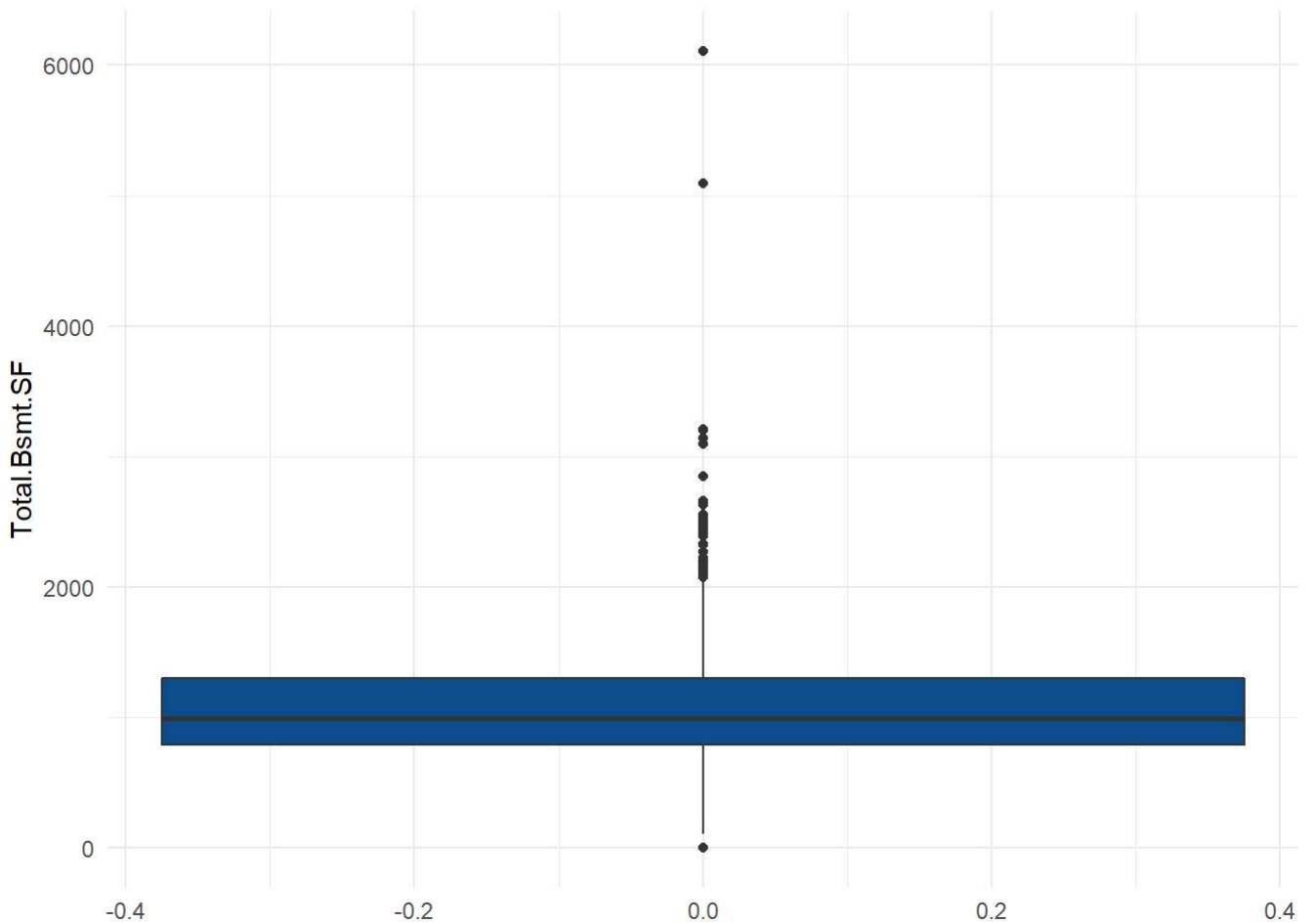
```



```
#arr2
array2 <- boxplot.stats(data_num$Total.Bsmt.SF)$out
paste0("The minimum value of the outliers is at ", min(array2), " The maximum
value of the outliers is at ", max(array2))
```

```
## [1] "The minimum value of the outliers is at 0 The maximum\nvalue of the outliers is at 611
0"
```

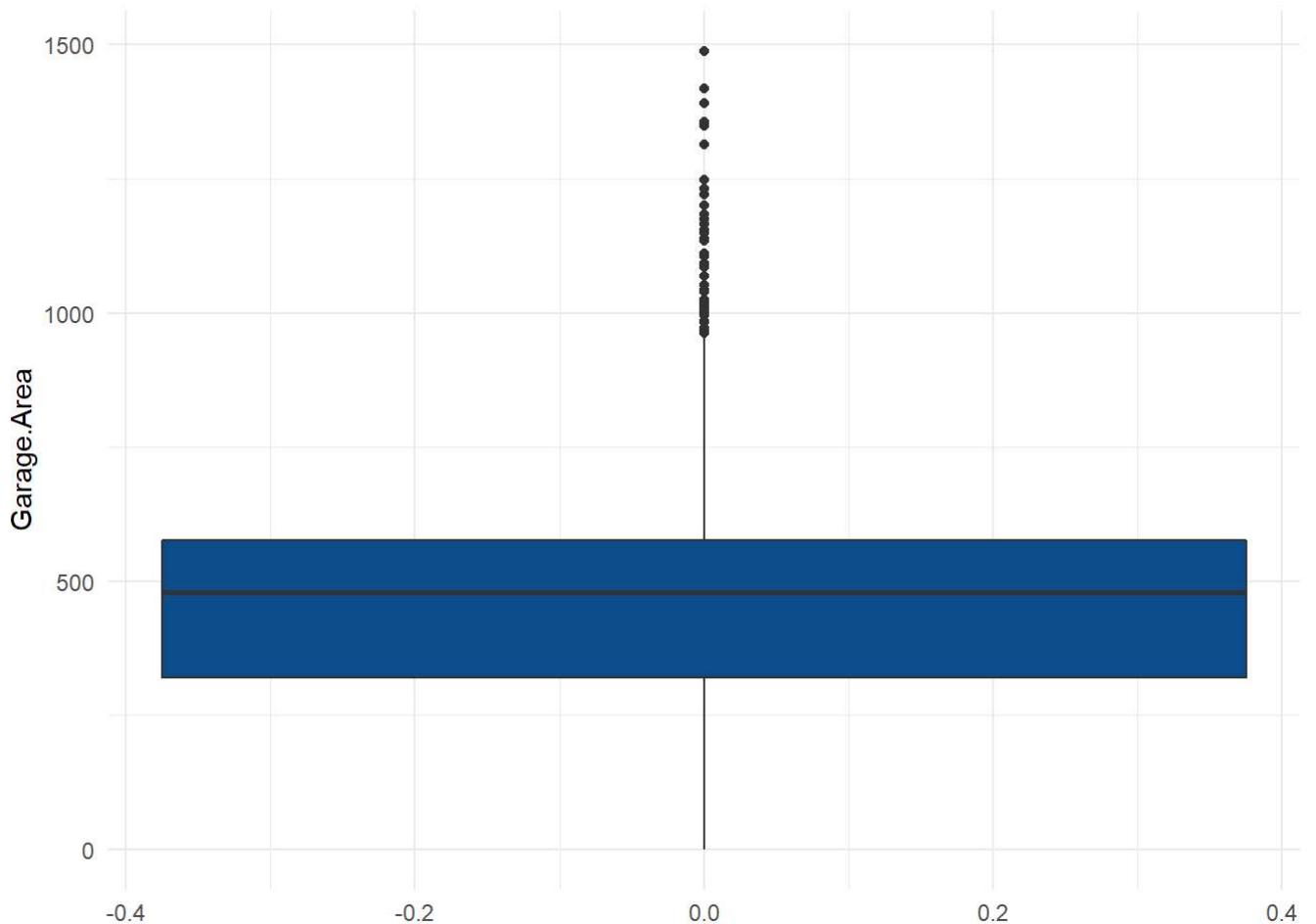
```
ggplot(data_num, aes(y=Total.Bsmt.SF)) + geom_boxplot(fill="#0c4c8a") +
  theme_minimal()
```



```
#arr3  
array3 <- boxplot.stats(data_num$Garage.Area)$out  
paste0("The minimum value of the outliers is at ", min(array3), " The maximum  
value of the outliers is at ", max(array3))
```

```
## [1] "The minimum value of the outliers is at 962 The maximum\nvalue of the outliers is at 1  
488"
```

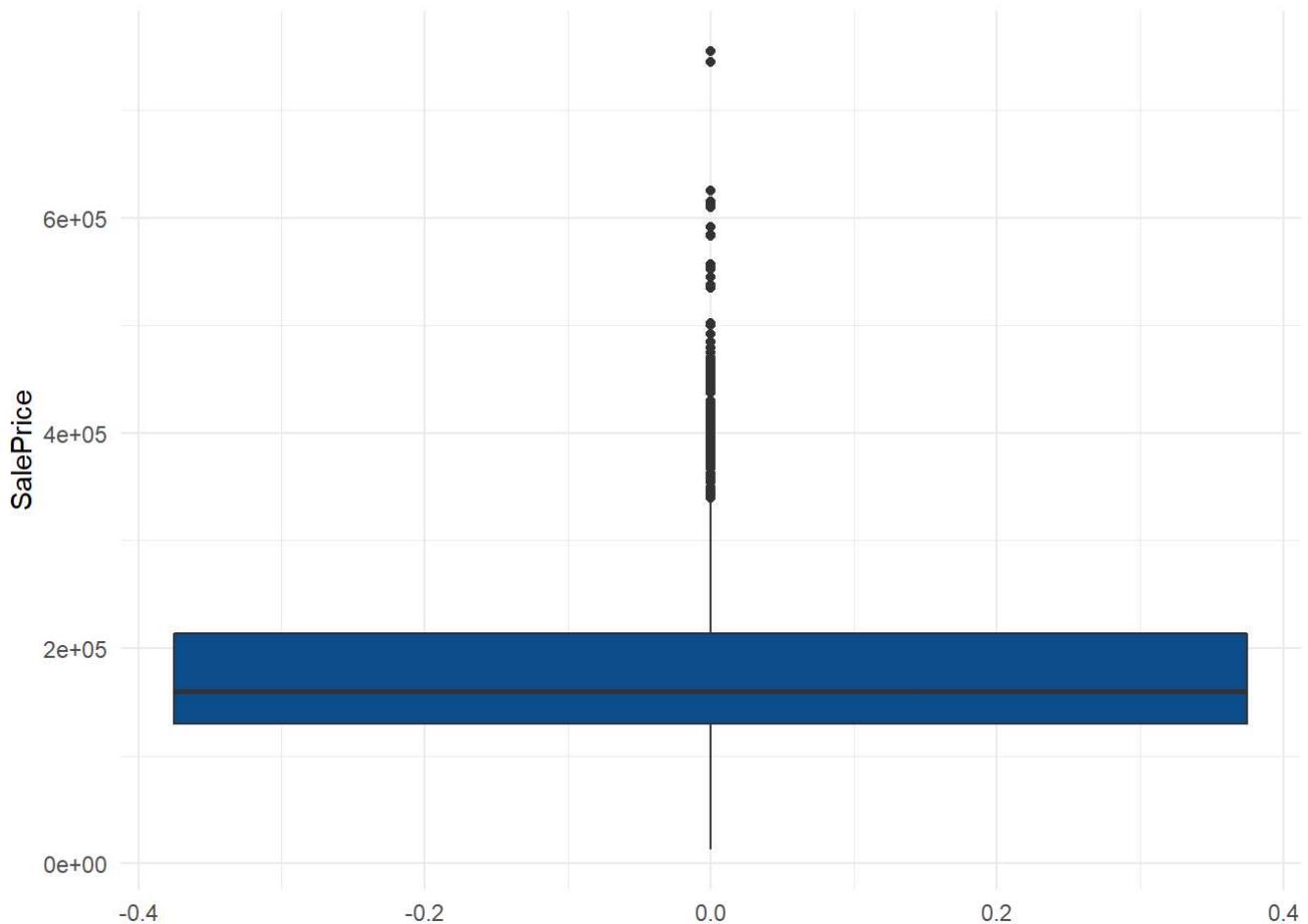
```
ggplot(data_num, aes(y=Garage.Area)) + geom_boxplot(fill="#0c4c8a") +  
  theme_minimal()
```



```
#arr4
array4 <- boxplot.stats(data_num$SalePrice)$out
paste0("The minimum value of the outliers is at ", min(array4), " The maximum
value of the outliers is at ", max(array4))
```

```
## [1] "The minimum value of the outliers is at 339750 The maximum\nvalue of the outliers is a
t 755000"
```

```
ggplot(data_num, aes(y=SalePrice)) + geom_boxplot(fill="#0c4c8a") +
  theme_minimal()
```



```
#12#
#attempt to correct any changes

#13#We try to find the best regression model

#helps us find best subset regression model
#selecting required numerical columns
dn <- data_num[c(4,5,6,7,10,11,12,13,14,18,29,39)]
#we apply regsubsets which gives us best regression all-subset model
leapy <- regsubsets(SalePrice~., data = dn, nbest = 1, method = "exhaustive")
```

```
## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.in =
## force.in, : 1 linear dependencies found
```

```
## Reordering variables and trying again:
```

```
#we use subsets to show which combination is the best
#subsets(leapy, statistic = "adjr2", ylim = c(0.60, 1), xlim = c(0, 4))
```

```
#14#
#creating Linear model for our best subset.
modelx<- lm(SalePrice~Overall.Qual, data = data_num)
summary(modelx)
```

```

## 
## Call:
## lm(formula = SalePrice ~ Overall.Qual, data = data_num)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -197507 -29254 -2283  21658 397493 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -95003.6   3933.8  -24.15 <2e-16 ***
## Overall.Qual 45251.0    628.8   71.96 <2e-16 ***  
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 48020 on 2928 degrees of freedom
## Multiple R-squared:  0.6388, Adjusted R-squared:  0.6387 
## F-statistic:  5179 on 1 and 2928 DF,  p-value: < 2.2e-16

```

```

plot(SalePrice~Overall.Qual, data= data_num)
abline(lm(data_num$SalePrice~data_num$Overall.Qual))

```

