

R Notebook

```
source(file.path('.', 'library.R'))
```

Read in the Data

Create a function to read in the data.

```
read_data <- function(name){  
  #' Load a data set from the data directory using a string lookup  
  #'  
  #' @param name A text string that will match the partial name of a data file in the data directory  
  #' @examples  
  #' d <- read_data('train')  
  my_path <- file.path('.', 'data')  
  my_dir <- list.files(my_path)  
  my_file <- grep(name, my_dir, value=TRUE)  
  d <- read_csv(file.path(my_path, my_file))  
}
```

An indicator field for the presence of a second floor is created.

```
t <- read_data('train') %>%  
  mutate(SalePriceLog = log(SalePrice)) %>%  
  mutate(secondFloor = `2ndFlrSF` > 0) %>%  
  mutate(X1stFlrSF = `1stFlrSF`) %>%  
  mutate(X2ndFlrSF = `2ndFlrSF`) %>%  
  mutate(X3SsnPorch = `3SsnPorch`) %>%  
  select(-`1stFlrSF`, -`2ndFlrSF`, `3SsnPorch` )
```

```
## Parsed with column specification:  
## cols(  
##   .default = col_character(),  
##   Id = col_integer(),  
##   MSSubClass = col_integer(),  
##   LotFrontage = col_integer(),  
##   LotArea = col_integer(),  
##   OverallQual = col_integer(),  
##   OverallCond = col_integer(),  
##   YearBuilt = col_integer(),  
##   YearRemodAdd = col_integer(),  
##   MasVnrArea = col_integer(),  
##   BsmtFinSF1 = col_integer(),  
##   BsmtFinSF2 = col_integer(),  
##   BsmtUnfSF = col_integer(),  
##   TotalBsmtSF = col_integer(),  
##   `1stFlrSF` = col_integer(),  
##   `2ndFlrSF` = col_integer(),  
##   LowQualFinSF = col_integer(),  
##   GrLivArea = col_integer(),  
##   BsmtFullBath = col_integer(),  
##   BsmtHalfBath = col_integer(),  
##   FullBath = col_integer()
```

```
## # ... with 18 more columns
## )
```

```
## See spec(...) for full column specifications.
```

The columns contain data and each column has a data class. There are qualitative and quantitative fields.

```
sapply(t,class)
```

```
##      Id      MSSubClass      MSZoning      LotFrontage      LotArea
## "integer" "integer" "character" "integer" "integer"
##      Street      Alley      LotShape      LandContour      Utilities
## "character" "character" "character" "character" "character"
##      LotConfig      LandSlope      Neighborhood      Condition1      Condition2
## "character" "character" "character" "character" "character"
##      BldgType      HouseStyle      OverallQual      OverallCond      YearBuilt
## "character" "character" "integer" "integer" "integer"
##      YearRemodAdd      RoofStyle      RoofMatl      Exterior1st      Exterior2nd
## "integer" "character" "character" "character" "character"
##      MasVnrType      MasVnrArea      ExterQual      ExterCond      Foundation
## "character" "integer" "character" "character" "character"
##      BsmtQual      BsmtCond      BsmtExposure      BsmtFinType1      BsmtFinSF1
## "character" "character" "character" "character" "integer"
##      BsmtFinType2      BsmtFinSF2      BsmtUnfSF      TotalBsmtSF      Heating
## "character" "integer" "integer" "integer" "character"
##      HeatingQC      CentralAir      Electrical      LowQualFinSF      GrLivArea
## "character" "character" "character" "integer" "integer"
##      BsmtFullBath      BsmtHalfBath      FullBath      HalfBath      BedroomAbvGr
## "integer" "integer" "integer" "integer" "integer"
##      KitchenAbvGr      KitchenQual      TotRmsAbvGrd      Functional      Fireplaces
## "integer" "character" "integer" "character" "integer"
##      FireplaceQu      GarageType      GarageYrBlt      GarageFinish      GarageCars
## "character" "character" "integer" "character" "integer"
##      GarageArea      GarageQual      GarageCond      PavedDrive      WoodDeckSF
## "integer" "character" "character" "character" "integer"
##      OpenPorchSF      EnclosedPorch      3SsnPorch      ScreenPorch      PoolArea
## "integer" "integer" "integer" "integer" "integer"
##      PoolQC      Fence      MiscFeature      MiscVal      MoSold
## "character" "character" "character" "integer" "integer"
##      YrSold      SaleType      SaleCondition      SalePrice      SalePriceLog
## "integer" "character" "character" "integer" "numeric"
##      secondFloor      X1stFlrSF      X2ndFlrSF      X3SsnPorch
## "logical" "integer" "integer" "integer"
```

The Kaggle data description page give the descriptions of these fields and they can be added to the data set.

```
var.labels <- c(
  SalePrice      = "The property's sale price in dollars." ,
  MSSubClass     = "The building class",
  MSZoning       = "The general zoning classification",
  LotFrontage    = "Linear feet of street connected to property",
  LotArea        = "Lot size in square feet",
  Street         = "Type of road access",
  Alley          = "Type of alley access",
  LotShape       = "General shape of property",
  LandContour    = "Flatness of the property",
  Utilities      = "Type of utilities available",
```

```

LotConfig      = "Lot configuration",
LandSlope      = "Slope of property",
Neighborhood   = "Physical locations within Ames city limits",
Condition1     = "Proximity to main road or railroad",
Condition2     = "Proximity to main road or railroad (if a second is present)",
BldgType       = "Type of dwelling",
HouseStyle     = "Style of dwelling",
OverallQual    = "Overall material and finish quality",
OverallCond    = "Overall condition rating",
YearBuilt      = "Original construction date",
YearRemodAdd   = "Remodel date",
RoofStyle      = "Type of roof",
RoofMatl       = "Roof material",
Exterior1st    = "Exterior covering on house",
Exterior2nd    = "Exterior covering on house (if more than one material)",
MasVnrType     = "Masonry veneer type",
MasVnrArea     = "Masonry veneer area in square feet",
ExterQual      = "Exterior material quality",
ExterCond      = "Present condition of the material on the exterior",
Foundation     = "Type of foundation",
BsmtQual       = "Height of the basement",
BsmtCond       = "General condition of the basement",
BsmtExposure   = "Walkout or garden level basement walls",
BsmtFinType1   = "Quality of basement finished area",
BsmtFinSF1     = "Type 1 finished square feet",
BsmtFinType2   = "Quality of second finished area (if present)",
BsmtFinSF2     = "Type 2 finished square feet",
BsmtUnfSF      = "Unfinished square feet of basement area",
TotalBsmtSF    = "Total square feet of basement area",
Heating        = "Type of heating",
HeatingQC      = "Heating quality and condition",
CentralAir     = "Central air conditioning",
Electrical     = "Electrical system",
X1stFlrSF     = "First Floor square feet",
X2ndFlrSF     = "Second floor square feet",
LowQualFinSF   = "Low quality finished square feet (all floors)",
GrLivArea      = "Above grade (ground) living area square feet",
BsmtFullBath   = "Basement full bathrooms",
BsmtHalfBath   = "Basement half bathrooms",
FullBath       = "Full bathrooms above grade",
HalfBath       = "Half baths above grade",
BedroomAbvGr  = "Number of bedrooms above basement level",
KitchenAbvGr   = "Number of kitchens",
KitchenQual    = "Kitchen quality",
TotRmsAbvGrd  = "Total rooms above grade (does not include bathrooms)",
Functional     = "Home functionality rating",
Fireplaces     = "Number of fireplaces",
FireplaceQu    = "Fireplace quality",
GarageType     = "Garage location",
GarageYrBlt    = "Year garage was built",
GarageFinish   = "Interior finish of the garage",
GarageCars     = "Size of garage in car capacity",
GarageArea     = "Size of garage in square feet",

```

```

GarageQual = "Garage quality",
GarageCond = "Garage condition",
PavedDrive = "Paved driveway",
WoodDeckSF = "Wood deck area in square feet",
OpenPorchSF = "Open porch area in square feet",
EnclosedPorch = "Enclosed porch area in square feet",
X3SsnPorch = "Three season porch area in square feet",
ScreenPorch = "Screen porch area in square feet",
PoolArea = "Pool area in square feet",
PoolQC = "Pool quality",
Fence = "Fence quality",
MiscFeature = "Miscellaneous feature not covered in other categories",
MiscVal = "$Value of miscellaneous feature",
MoSold = "Month Sold",
YrSold = "Year Sold",
SaleType = "Type of sale",
SaleCondition = "Condition of sale",
SalePriceLog = "Log of Sale Price",
secondFloor = "House has Second Floor"
)
t <- Hmisc::upData(t, labels = var.labels)

```

```

## Input object size: 769632 bytes; 84 variables 1460 observations
## New object size: 811560 bytes; 84 variables 1460 observations

```

```
Hmisc::label(t)
```

```

##                               Id
##                               ""
##                               MSSubClass
##                               "The building class"
##                               MSZoning
##                               "The general zoning classification"
##                               LotFrontage
##                               "Linear feet of street connected to property"
##                               LotArea
##                               "Lot size in square feet"
##                               Street
##                               "Type of road access"
##                               Alley
##                               "Type of alley access"
##                               LotShape
##                               "General shape of property"
##                               LandContour
##                               "Flatness of the property"
##                               Utilities
##                               "Type of utilities available"
##                               LotConfig
##                               "Lot configuration"
##                               LandSlope
##                               "Slope of property"
##                               Neighborhood
##                               "Physical locations within Ames city limits"
##                               Condition1
##                               "Proximity to main road or railroad"

```

```

##                                     Condition2
## "Proximity to main road or railroad (if a second is present)"
##                                     BldgType
##                                     "Type of dwelling"
##                                     HouseStyle
##                                     "Style of dwelling"
##                                     OverallQual
## "Overall material and finish quality"
##                                     OverallCond
## "Overall condition rating"
##                                     YearBuilt
## "Original construction date"
##                                     YearRemodAdd
## "Remodel date"
##                                     RoofStyle
## "Type of roof"
##                                     RoofMatl
## "Roof material"
##                                     Exterior1st
## "Exterior covering on house"
##                                     Exterior2nd
## "Exterior covering on house (if more than one material)"
##                                     MasVnrType
## "Masonry veneer type"
##                                     MasVnrArea
## "Masonry veneer area in square feet"
##                                     ExterQual
## "Exterior material quality"
##                                     ExterCond
## "Present condition of the material on the exterior"
##                                     Foundation
## "Type of foundation"
##                                     BsmtQual
## "Height of the basement"
##                                     BsmtCond
## "General condition of the basement"
##                                     BsmtExposure
## "Walkout or garden level basement walls"
##                                     BsmtFinType1
## "Quality of basement finished area"
##                                     BsmtFinSF1
## "Type 1 finished square feet"
##                                     BsmtFinType2
## "Quality of second finished area (if present)"
##                                     BsmtFinSF2
## "Type 2 finished square feet"
##                                     BsmtUnfSF
## "Unfinished square feet of basement area"
##                                     TotalBsmtSF
## "Total square feet of basement area"
##                                     Heating
## "Type of heating"
##                                     HeatingQC
## "Heating quality and condition"

```

```

##                                     CentralAir
##                                "Central air conditioning"
##                                     Electrical
##                                "Electrical system"
##                                     LowQualFinSF
##                                "Low quality finished square feet (all floors)"
##                                     GrLivArea
##                                "Above grade (ground) living area square feet"
##                                     BsmtFullBath
##                                "Basement full bathrooms"
##                                     BsmtHalfBath
##                                "Basement half bathrooms"
##                                     FullBath
##                                "Full bathrooms above grade"
##                                     HalfBath
##                                "Half baths above grade"
##                                     BedroomAbvGr
##                                "Number of bedrooms above basement level"
##                                     KitchenAbvGr
##                                "Number of kitchens"
##                                     KitchenQual
##                                "Kitchen quality"
##                                     TotRmsAbvGrd
##                                "Total rooms above grade (does not include bathrooms)"
##                                     Functional
##                                "Home functionality rating"
##                                     Fireplaces
##                                "Number of fireplaces"
##                                     FireplaceQu
##                                "Fireplace quality"
##                                     GarageType
##                                "Garage location"
##                                     GarageYrBlt
##                                "Year garage was built"
##                                     GarageFinish
##                                "Interior finish of the garage"
##                                     GarageCars
##                                "Size of garage in car capacity"
##                                     GarageArea
##                                "Size of garage in square feet"
##                                     GarageQual
##                                "Garage quality"
##                                     GarageCond
##                                "Garage condition"
##                                     PavedDrive
##                                "Paved driveway"
##                                     WoodDeckSF
##                                "Wood deck area in square feet"
##                                     OpenPorchSF
##                                "Open porch area in square feet"
##                                     EnclosedPorch
##                                "Enclosed porch area in square feet"
##                                     3SsnPorch
##                                     ""

```

```
##                               ScreenPorch
##                               "Screen porch area in square feet"
##                               PoolArea
##                               "Pool area in square feet"
##                               PoolQC
##                               "Pool quality"
##                               Fence
##                               "Fence quality"
##                               MiscFeature
##                               "Miscellaneous feature not covered in other categories"
##                               MiscVal
##                               "$Value of miscellaneous feature"
##                               MoSold
##                               "Month Sold"
##                               YrSold
##                               "Year Sold"
##                               SaleType
##                               "Type of sale"
##                               SaleCondition
##                               "Condition of sale"
##                               SalePrice
##                               "The property's sale price in dollars."
##                               SalePriceLog
##                               "Log of Sale Price"
##                               secondFloor
##                               "House has Second Floor"
##                               X1stFlrSF
##                               "First Floor square feet"
##                               X2ndFlrSF
##                               "Second floor square feet"
##                               X3SsnPorch
##                               "Three season porch area in square feet"
```

Make sure the labels work.

```
label(t$OverallQual)
```

```
## [1] "Overall material and finish quality"
```

```
label(t$SalePriceLog)
```

```
## [1] "Log of Sale Price"
```

Plot the data

```
t %>% ggplot() +
  aes(X1stFlrSF, SalePriceLog, color=secondFloor) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha=0.35) +
  ggtitle(paste(label(t$SalePrice), label(t$X1stFlrSF))) +
  #geom_smooth(method='loess', formula = y~x) +
  #geom_smooth(method='lm', formula = y~x) +
  xlab(label(t$X1stFlrSF)) +
  ylab(label(t$SalePrice))
```



First Model

First we will experiment with the numeric fields. We will leave the category fields aside for now.

```
t_types <- sapply(t,typeof)
t_types
```

```
##      Id      MSSubClass      MSZoning      LotFrontage      LotArea
## "integer" "integer"    "character" "integer"      "integer"
##      Street      Alley      LotShape      LandContour      Utilities
## "character" "character" "character" "character"    "character"
##      LotConfig      LandSlope      Neighborhood      Condition1      Condition2
## "character" "character" "character" "character"    "character"
##      BldgType      HouseStyle      OverallQual      OverallCond      YearBuilt
## "character" "character" "integer"   "integer"      "integer"
##      YearRemodAdd      RoofStyle      RoofMatl      Exterior1st      Exterior2nd
## "integer"   "character" "character" "character"    "character"
##      MasVnrType      MasVnrArea      ExterQual      ExterCond      Foundation
## "character" "integer"  "character" "character"    "character"
##      BsmtQual      BsmtCond      BsmtExposure      BsmtFinType1      BsmtFinSF1
## "character" "character" "character" "character"    "integer"
##      BsmtFinType2      BsmtFinSF2      BsmtUnfSF      TotalBsmtSF      Heating
## "character" "integer"  "integer"   "integer"      "character"
##      HeatingQC      CentralAir      Electrical      LowQualFinSF      GrLivArea
## "character" "character" "character" "integer"      "integer"
##      BsmtFullBath      BsmtHalfBath      FullBath      HalfBath      BedroomAbvGr
```



```
##      "integer"      "integer"      "integer"      "integer"      "integer"
## KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces
##      "integer"      "character"      "integer"      "character"      "integer"
## FireplaceQu GarageType GarageYrBltd GarageFinish GarageCars
##      "character"      "character"      "integer"      "character"      "integer"
## GarageArea GarageQual GarageCond PavedDrive WoodDeckSF
##      "integer"      "character"      "character"      "character"      "integer"
## OpenPorchSF EnclosedPorch 3SsnPorch ScreenPorch PoolArea
##      "integer"      "integer"      "integer"      "integer"      "integer"
## PoolQC Fence MiscFeature MiscVal MoSold
##      "character"      "character"      "character"      "integer"      "integer"
## YrSold SaleType SaleCondition SalePrice SalePriceLog
##      "integer"      "character"      "character"      "integer"      "double"
## secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch
##      "logical"      "integer"      "integer"      "integer"
```

```
names(t)
```

```
## [1] "Id" "MSSubClass" "MSZoning" "LotFrontage"
## [5] "LotArea" "Street" "Alley" "LotShape"
## [9] "LandContour" "Utilities" "LotConfig" "LandSlope"
## [13] "Neighborhood" "Condition1" "Condition2" "BldgType"
## [17] "HouseStyle" "OverallQual" "OverallCond" "YearBuilt"
## [21] "YearRemodAdd" "RoofStyle" "RoofMatl" "Exterior1st"
## [25] "Exterior2nd" "MasVnrType" "MasVnrArea" "ExterQual"
## [29] "ExterCond" "Foundation" "BsmtQual" "BsmtCond"
## [33] "BsmtExposure" "BsmtFinType1" "BsmtFinSF1" "BsmtFinType2"
## [37] "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF" "Heating"
## [41] "HeatingQC" "CentralAir" "Electrical" "LowQualFinSF"
## [45] "GrLivArea" "BsmtFullBath" "BsmtHalfBath" "FullBath"
## [49] "HalfBath" "BedroomAbvGr" "KitchenAbvGr" "KitchenQual"
## [53] "TotRmsAbvGrd" "Functional" "Fireplaces" "FireplaceQu"
## [57] "GarageType" "GarageYrBltd" "GarageFinish" "GarageCars"
## [61] "GarageArea" "GarageQual" "GarageCond" "PavedDrive"
## [65] "WoodDeckSF" "OpenPorchSF" "EnclosedPorch" "3SsnPorch"
## [69] "ScreenPorch" "PoolArea" "PoolQC" "Fence"
## [73] "MiscFeature" "MiscVal" "MoSold" "YrSold"
## [77] "SaleType" "SaleCondition" "SalePrice" "SalePriceLog"
## [81] "secondFloor" "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch"
```

```
keep_fields <- !(names(t) %in% c('SalePrice','Id'))
keep_fields
```

```
## [1] FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [12] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [23] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [34] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [45] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [56] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [67] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [78] TRUE FALSE TRUE TRUE TRUE TRUE TRUE
```

```
t_numeric_fields <- t_types[t_types != 'character' & keep_fields ]
t_numeric_fields
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond
```

```
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      YearBuilt    YearRemodAdd    MasVnrArea    BsmtFinSF1    BsmtFinSF2
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      BsmtUnfSF    TotalBsmtSF    LowQualFinSF    GrLivArea    BsmtFullBath
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      BsmtHalfBath    FullBath    HalfBath    BedroomAbvGr    KitchenAbvGr
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      TotRmsAbvGrd    Fireplaces    GarageYrBlt    GarageCars    GarageArea
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      WoodDeckSF    OpenPorchSF    EnclosedPorch    3SsnPorch    ScreenPorch
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      PoolArea      MiscVal      MoSold      YrSold    SalePriceLog
##      "integer"      "integer"      "integer"      "integer"      "double"
##      secondFloor    X1stFlrSF    X2ndFlrSF    X3SsnPorch
##      "logical"      "integer"      "integer"      "integer"
```

```
t_numeric <- t[names(t_numeric_fields)]
sapply(t_numeric,typeof)
```

```
##      MSSubClass    LotFrontage    LotArea    OverallQual    OverallCond
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      YearBuilt    YearRemodAdd    MasVnrArea    BsmtFinSF1    BsmtFinSF2
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      BsmtUnfSF    TotalBsmtSF    LowQualFinSF    GrLivArea    BsmtFullBath
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      BsmtHalfBath    FullBath    HalfBath    BedroomAbvGr    KitchenAbvGr
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      TotRmsAbvGrd    Fireplaces    GarageYrBlt    GarageCars    GarageArea
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      WoodDeckSF    OpenPorchSF    EnclosedPorch    3SsnPorch    ScreenPorch
##      "integer"      "integer"      "integer"      "integer"      "integer"
##      PoolArea      MiscVal      MoSold      YrSold    SalePriceLog
##      "integer"      "integer"      "integer"      "integer"      "double"
##      secondFloor    X1stFlrSF    X2ndFlrSF    X3SsnPorch
##      "logical"      "integer"      "integer"      "integer"
```

```
head(t_numeric)
```

```
## # A tibble: 6 x 39
##      MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
##      <int+lbl> <int+lbl> <int+lbl> <int+lbl> <int+lbl> <int+lbl>
## 1         60         65      8450         7         5      2003
## 2         20         80      9600         6         8      1976
## 3         60         68     11250         7         5      2001
## 4         70         60      9550         7         5      1915
## 5         60         84     14260         8         5      2000
## 6         50         85     14115         5         5      1993
## # ... with 33 more variables: YearRemodAdd <int+lbl>,
## #   MasVnrArea <int+lbl>, BsmtFinSF1 <int+lbl>, BsmtFinSF2 <int+lbl>,
## #   BsmtUnfSF <int+lbl>, TotalBsmtSF <int+lbl>, LowQualFinSF <int+lbl>,
## #   GrLivArea <int+lbl>, BsmtFullBath <int+lbl>, BsmtHalfBath <int+lbl>,
## #   FullBath <int+lbl>, HalfBath <int+lbl>, BedroomAbvGr <int+lbl>,
## #   KitchenAbvGr <int+lbl>, TotRmsAbvGrd <int+lbl>, Fireplaces <int+lbl>,
## #   GarageYrBlt <int+lbl>, GarageCars <int+lbl>, GarageArea <int+lbl>,
## #   WoodDeckSF <int+lbl>, OpenPorchSF <int+lbl>, EnclosedPorch <int+lbl>,
```

```
## # `3SsnPorch` <int>, ScreenPorch <int+lbl>, PoolArea <int+lbl>,
## # MiscVal <int+lbl>, MoSold <int+lbl>, YrSold <int+lbl>,
## # SalePriceLog <dbl+lbl>, secondFloor <lgl+lbl>, X1stFlrSF <int+lbl>,
## # X2ndFlrSF <int+lbl>, X3SsnPorch <int+lbl>
```

```
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "3SsnPorch" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch"
```

```
sapply(t_numeric,typeof)
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond
## "integer" "integer" "integer" "integer" "integer"
## YearBuilt YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2
## "integer" "integer" "integer" "integer" "integer"
## BsmtUnfSF TotalBsmtSF LowQualFinSF GrLivArea BsmtFullBath
## "integer" "integer" "integer" "integer" "integer"
## BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr
## "integer" "integer" "integer" "integer" "integer"
## TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars GarageArea
## "integer" "integer" "integer" "integer" "integer"
## WoodDeckSF OpenPorchSF EnclosedPorch 3SsnPorch ScreenPorch
## "integer" "integer" "integer" "integer" "integer"
## PoolArea MiscVal MoSold YrSold SalePriceLog
## "integer" "integer" "integer" "integer" "double"
## secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch
## "logical" "integer" "integer" "integer" "integer"
```

Corrrelation

So now we search for factors that are significantly correlated and order those by strength of correlation.

```
model_num = 1
num_of_canadate_features = 15
sig <- 0.0010
response <- "SalePriceLog"
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "3SsnPorch" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
```

```
## [37] "X1stFlrSF"      "X2ndFlrSF"      "X3SsnPorch"
```

```
# Calculate the correlation matrix
```

```
res2 <- rcorr(as.matrix(t_numeric))
```

```
res2$P
```

```
##          MSSubClass  LotFrontage      LotArea  OverallQual
## MSSubClass          NA 0.000000e+00 8.194115e-08 2.127765e-01
## LotFrontage 0.000000e+00          NA 0.000000e+00 0.000000e+00
## LotArea      8.194115e-08 0.000000e+00          NA 5.105510e-05
## OverallQual 2.127765e-01 0.000000e+00 5.105510e-05          NA
## OverallCond 2.342026e-02 4.019570e-02 8.296279e-01 4.361720e-04
## YearBuilt   2.875793e-01 1.813803e-05 5.869923e-01 0.000000e+00
## YearRemodAdd 1.211632e-01 2.052218e-03 5.985903e-01 0.000000e+00
## MasVnrArea   3.824712e-01 1.536393e-11 6.995921e-05 0.000000e+00
## BsmtFinSF1   7.598818e-03 2.220446e-16 0.000000e+00 0.000000e+00
## BsmtFinSF2   1.210779e-02 8.388578e-02 2.068605e-05 2.388568e-02
## BsmtUnfSF    6.637585e-08 3.979765e-06 9.203746e-01 0.000000e+00
## TotalBsmtSF 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## LowQualFinSF 7.586520e-02 1.827783e-01 8.552305e-01 2.452458e-01
## GrLivArea    4.213733e-03 0.000000e+00 0.000000e+00 0.000000e+00
## BsmtFullBath 8.939735e-01 4.587932e-04 1.230799e-09 2.094408e-05
## BsmtHalfBath 9.290423e-01 8.022392e-01 6.646032e-02 1.251676e-01
## FullBath     4.502116e-07 3.634426e-12 1.360046e-06 0.000000e+00
## HalfBath     8.788081e-12 6.365788e-02 5.861562e-01 0.000000e+00
## BedroomAbvGr 3.708289e-01 0.000000e+00 4.519868e-06 9.949905e-05
## KitchenAbvGr 0.000000e+00 8.335889e-01 4.971427e-01 1.435740e-12
## TotRmsAbvGrd 1.230183e-01 0.000000e+00 2.460254e-13 0.000000e+00
## Fireplaces   8.175109e-02 0.000000e+00 0.000000e+00 0.000000e+00
## GarageYrBlt  1.566879e-03 1.834203e-02 3.545897e-01 0.000000e+00
## GarageCars   1.255479e-01 0.000000e+00 2.706817e-09 0.000000e+00
## GarageArea   1.592243e-04 0.000000e+00 3.802736e-12 0.000000e+00
## WoodDeckSF   6.310397e-01 2.136523e-03 4.001000e-11 0.000000e+00
## OpenPorchSF  8.158485e-01 1.210208e-07 1.185832e-03 0.000000e+00
## EnclosedPorch 6.458454e-01 7.110477e-01 4.837906e-01 1.276996e-05
## 3SsnPorch    9.414953e-02 1.520968e-02 4.355274e-01 2.461582e-01
## ScreenPorch  3.202578e-01 1.517836e-01 9.924777e-02 1.314582e-02
## PoolArea     7.518384e-01 5.393463e-13 2.979905e-03 1.275642e-02
## MiscVal      7.692691e-01 9.071905e-01 1.459891e-01 2.304123e-01
## MoSold       6.040063e-01 6.982029e-01 9.633078e-01 6.790959e-03
## YrSold       4.137257e-01 7.964813e-01 5.861053e-01 2.963852e-01
## SalePriceLog 4.692451e-03 0.000000e+00 0.000000e+00 0.000000e+00
## secondFloor  0.000000e+00 1.136398e-01 6.543549e-01 1.513700e-11
## X1stFlrSF    0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## X2ndFlrSF    0.000000e+00 5.433170e-03 5.144266e-02 0.000000e+00
## X3SsnPorch   9.414953e-02 1.520968e-02 4.355274e-01 2.461582e-01
##          OverallCond  YearBuilt  YearRemodAdd  MasVnrArea
## MSSubClass 2.342026e-02 2.875793e-01 1.211632e-01 3.824712e-01
## LotFrontage 4.019570e-02 1.813803e-05 2.052218e-03 1.536393e-11
## LotArea     8.296279e-01 5.869923e-01 5.985903e-01 6.995921e-05
## OverallQual 4.361720e-04 0.000000e+00 0.000000e+00 0.000000e+00
## OverallCond          NA 0.000000e+00 4.816091e-03 9.716410e-07
## YearBuilt   0.000000e+00          NA 0.000000e+00 0.000000e+00
## YearRemodAdd 4.816091e-03 0.000000e+00          NA 5.400569e-12
## MasVnrArea   9.716410e-07 0.000000e+00 5.400569e-12          NA
```

## BsmtFinSF1	7.741099e-02	0.000000e+00	8.466438e-07	0.000000e+00
## BsmtFinSF2	1.244256e-01	6.066982e-02	9.602934e-03	5.834234e-03
## BsmtUnfSF	1.530091e-07	1.053188e-08	3.104628e-12	1.234288e-05
## TotalBsmtSF	4.685274e-11	0.000000e+00	0.000000e+00	0.000000e+00
## LowQualFinSF	3.303245e-01	1.475930e-12	1.706455e-02	8.467646e-03
## GrLivArea	2.311138e-03	1.665335e-14	0.000000e+00	0.000000e+00
## BsmtFullBath	3.580810e-02	4.964917e-13	4.707214e-06	1.138425e-03
## BsmtHalfBath	6.367484e-06	1.449943e-01	6.376325e-01	3.097834e-01
## FullBath	7.238654e-14	0.000000e+00	0.000000e+00	0.000000e+00
## HalfBath	2.022485e-02	0.000000e+00	1.677547e-12	9.325873e-15
## BedroomAbvGr	6.202024e-01	6.920572e-03	1.211642e-01	8.671369e-05
## KitchenAbvGr	8.754659e-04	1.753175e-11	9.270313e-09	1.520249e-01
## TotRmsAbvGrd	2.779336e-02	2.544832e-04	1.483258e-13	0.000000e+00
## Fireplaces	3.630812e-01	1.423519e-08	1.619633e-05	0.000000e+00
## GarageYrBlt	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## GarageCars	8.424372e-13	0.000000e+00	0.000000e+00	0.000000e+00
## GarageArea	5.945338e-09	0.000000e+00	0.000000e+00	0.000000e+00
## WoodDeckSF	8.987253e-01	0.000000e+00	1.998401e-15	9.347851e-10
## OpenPorchSF	2.133227e-01	3.623768e-13	0.000000e+00	1.548562e-06
## EnclosedPorch	7.159421e-03	0.000000e+00	7.771561e-14	2.570285e-05
## 3SsnPorch	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
## ScreenPorch	3.625216e-02	5.435519e-02	1.389957e-01	1.916276e-02
## PoolArea	9.395944e-01	8.501179e-01	8.238849e-01	6.553522e-01
## MiscVal	8.568156e-03	1.891692e-01	6.945331e-01	2.562107e-01
## MoSold	8.933753e-01	6.359583e-01	4.119179e-01	8.203518e-01
## YrSold	9.321243e-02	6.031268e-01	1.722487e-01	7.548601e-01
## SalePriceLog	1.591324e-01	0.000000e+00	0.000000e+00	0.000000e+00
## secondFloor	1.764981e-02	2.073907e-04	2.739984e-02	1.313690e-02
## X1stFlrSF	3.126349e-08	0.000000e+00	0.000000e+00	0.000000e+00
## X2ndFlrSF	2.690898e-01	6.939294e-01	7.777955e-08	2.121014e-11
## X3SsnPorch	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
##	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF
## MSSubClass	7.598818e-03	1.210779e-02	6.637585e-08	0.000000e+00
## LotFrontage	2.220446e-16	8.388578e-02	3.979765e-06	0.000000e+00
## LotArea	0.000000e+00	2.068605e-05	9.203746e-01	0.000000e+00
## OverallQual	0.000000e+00	2.388568e-02	0.000000e+00	0.000000e+00
## OverallCond	7.741099e-02	1.244256e-01	1.530091e-07	4.685274e-11
## YearBuilt	0.000000e+00	6.066982e-02	1.053188e-08	0.000000e+00
## YearRemodAdd	8.466438e-07	9.602934e-03	3.104628e-12	0.000000e+00
## MasVnrArea	0.000000e+00	5.834234e-03	1.234288e-05	0.000000e+00
## BsmtFinSF1	NA	5.555041e-02	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.555041e-02	NA	6.661338e-16	6.010660e-05
## BsmtUnfSF	0.000000e+00	6.661338e-16	NA	0.000000e+00
## TotalBsmtSF	0.000000e+00	6.010660e-05	0.000000e+00	NA
## LowQualFinSF	1.369761e-02	5.718554e-01	2.821334e-01	2.042388e-01
## GrLivArea	8.881784e-16	7.128488e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	0.000000e+00	1.083816e-09	0.000000e+00	0.000000e+00
## BsmtHalfBath	9.972478e-03	6.687447e-03	2.463966e-04	9.904177e-01
## FullBath	2.529072e-02	3.470204e-03	0.000000e+00	0.000000e+00
## HalfBath	8.707324e-01	2.195850e-01	1.163181e-01	6.227931e-02
## BedroomAbvGr	3.949967e-05	5.481779e-01	1.485798e-10	5.394644e-02
## KitchenAbvGr	1.950234e-03	1.196093e-01	2.506158e-01	8.449363e-03
## TotRmsAbvGrd	9.051660e-02	1.785383e-01	0.000000e+00	0.000000e+00
## Fireplaces	0.000000e+00	7.308663e-02	4.880435e-02	0.000000e+00

## GarageYrBlt	1.013706e-08	1.069187e-03	9.254819e-13	0.000000e+00
## GarageCars	0.000000e+00	1.439250e-01	0.000000e+00	0.000000e+00
## GarageArea	0.000000e+00	4.864925e-01	1.690648e-12	0.000000e+00
## WoodDeckSF	3.108624e-15	9.454565e-03	8.391603e-01	0.000000e+00
## OpenPorchSF	1.867888e-05	9.060156e-01	7.585228e-07	0.000000e+00
## EnclosedPorch	9.005499e-05	1.628382e-01	9.228150e-01	2.587670e-04
## 3SsnPorch	3.125032e-01	2.520757e-01	4.278958e-01	1.533765e-01
## ScreenPorch	1.778536e-02	6.748805e-04	6.310420e-01	1.232140e-03
## PoolArea	7.033176e-08	1.111537e-01	1.802008e-01	1.354119e-06
## MiscVal	8.915451e-01	8.504155e-01	3.627455e-01	4.804779e-01
## MoSold	5.482075e-01	5.614175e-01	1.827454e-01	6.143911e-01
## YrSold	5.835463e-01	2.259961e-01	1.150733e-01	5.676653e-01
## SalePriceLog	0.000000e+00	8.536300e-01	0.000000e+00	0.000000e+00
## secondFloor	0.000000e+00	7.216583e-06	1.797090e-01	0.000000e+00
## X1stFlrSF	0.000000e+00	2.020320e-04	0.000000e+00	0.000000e+00
## X2ndFlrSF	1.455248e-07	1.453587e-04	8.645239e-01	1.894085e-11
## X3SsnPorch	3.125032e-01	2.520757e-01	4.278958e-01	1.533765e-01
## LowQualFinSF		GrLivArea	BsmtFullBath	BsmtHalfBath
## MSSubClass	7.586520e-02	4.213733e-03	8.939735e-01	9.290423e-01
## LotFrontage	1.827783e-01	0.000000e+00	4.587932e-04	8.022392e-01
## LotArea	8.552305e-01	0.000000e+00	1.230799e-09	6.646032e-02
## OverallQual	2.452458e-01	0.000000e+00	2.094408e-05	1.251676e-01
## OverallCond	3.303245e-01	2.311138e-03	3.580810e-02	6.367484e-06
## YearBuilt	1.475930e-12	1.665335e-14	4.964917e-13	1.449943e-01
## YearRemodAdd	1.706455e-02	0.000000e+00	4.707214e-06	6.376325e-01
## MasVnrArea	8.467646e-03	0.000000e+00	1.138425e-03	3.097834e-01
## BsmtFinSF1	1.369761e-02	8.881784e-16	0.000000e+00	9.972478e-03
## BsmtFinSF2	5.718554e-01	7.128488e-01	1.083816e-09	6.687447e-03
## BsmtUnfSF	2.821334e-01	0.000000e+00	0.000000e+00	2.463966e-04
## TotalBsmtSF	2.042388e-01	0.000000e+00	0.000000e+00	9.904177e-01
## LowQualFinSF	NA	2.399685e-07	7.173345e-02	8.235244e-01
## GrLivArea	2.399685e-07	NA	1.834039e-01	4.700979e-01
## BsmtFullBath	7.173345e-02	1.834039e-01	NA	1.374508e-08
## BsmtHalfBath	8.235244e-01	4.700979e-01	1.374508e-08	NA
## FullBath	9.783903e-01	0.000000e+00	1.368378e-02	3.719872e-02
## HalfBath	3.011218e-01	0.000000e+00	2.379416e-01	6.375543e-01
## BedroomAbvGr	5.275460e-05	0.000000e+00	7.237188e-09	7.558104e-02
## KitchenAbvGr	7.739872e-01	1.282742e-04	1.129357e-01	1.473008e-01
## TotRmsAbvGrd	4.904244e-07	0.000000e+00	4.181609e-02	3.627516e-01
## Fireplaces	4.166751e-01	0.000000e+00	1.216473e-07	2.685363e-01
## GarageYrBlt	1.771551e-01	0.000000e+00	3.498871e-06	3.997996e-03
## GarageCars	3.002845e-04	0.000000e+00	4.259910e-07	4.250739e-01
## GarageArea	9.772127e-03	0.000000e+00	5.316858e-12	3.488402e-01
## WoodDeckSF	3.312869e-01	0.000000e+00	1.526557e-11	1.250635e-01
## OpenPorchSF	4.859080e-01	0.000000e+00	1.005792e-02	3.335712e-01
## EnclosedPorch	1.959094e-02	7.278971e-01	5.656746e-02	7.439517e-01
## 3SsnPorch	8.697349e-01	4.305893e-01	9.967684e-01	1.799353e-01
## ScreenPorch	3.061609e-01	1.021511e-04	3.767857e-01	2.199645e-01
## PoolArea	1.753504e-02	5.918421e-11	9.756789e-03	4.445328e-01
## MiscVal	8.848674e-01	9.265212e-01	3.788657e-01	7.785284e-01
## MoSold	3.971935e-01	5.495607e-02	3.328625e-01	2.093598e-01
## YrSold	2.694414e-01	1.630401e-01	1.038816e-02	7.554937e-02
## SalePriceLog	1.471041e-01	0.000000e+00	0.000000e+00	8.441568e-01
## secondFloor	2.385267e-01	0.000000e+00	0.000000e+00	3.410671e-01

##	X1stFlrSF	5.866501e-01	0.000000e+00	0.000000e+00	9.404841e-01
##	X2ndFlrSF	1.547482e-02	0.000000e+00	7.124057e-11	3.623802e-01
##	X3SsnPorch	8.697349e-01	4.305893e-01	9.967684e-01	1.799353e-01
##		FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr
##	MSSubClass	4.502116e-07	8.788081e-12	3.708289e-01	0.000000e+00
##	LotFrontage	3.634426e-12	6.365788e-02	0.000000e+00	8.335889e-01
##	LotArea	1.360046e-06	5.861562e-01	4.519868e-06	4.971427e-01
##	OverallQual	0.000000e+00	0.000000e+00	9.949905e-05	1.435740e-12
##	OverallCond	7.238654e-14	2.022485e-02	6.202024e-01	8.754659e-04
##	YearBuilt	0.000000e+00	0.000000e+00	6.920572e-03	1.753175e-11
##	YearRemodAdd	0.000000e+00	1.677547e-12	1.211642e-01	9.270313e-09
##	MasVnrArea	0.000000e+00	9.325873e-15	8.671369e-05	1.520249e-01
##	BsmtFinSF1	2.529072e-02	8.707324e-01	3.949967e-05	1.950234e-03
##	BsmtFinSF2	3.470204e-03	2.195850e-01	5.481779e-01	1.196093e-01
##	BsmtUnfSF	0.000000e+00	1.163181e-01	1.485798e-10	2.506158e-01
##	TotalBsmtSF	0.000000e+00	6.227931e-02	5.394644e-02	8.449363e-03
##	LowQualFinSF	9.783903e-01	3.011218e-01	5.275460e-05	7.739872e-01
##	GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	1.282742e-04
##	BsmtFullBath	1.368378e-02	2.379416e-01	7.237188e-09	1.129357e-01
##	BsmtHalfBath	3.719872e-02	6.375543e-01	7.558104e-02	1.473008e-01
##	FullBath		NA	1.685132e-07	0.000000e+00
##	HalfBath	1.685132e-07		NA	0.000000e+00
##	BedroomAbvGr	0.000000e+00	0.000000e+00		NA
##	KitchenAbvGr	3.313170e-07	9.077592e-03	1.887379e-14	
##	TotRmsAbvGrd	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	Fireplaces	0.000000e+00	3.996803e-15	3.810705e-05	2.035535e-06
##	GarageYrBlt	0.000000e+00	1.665335e-13	1.656563e-02	3.589658e-06
##	GarageCars	0.000000e+00	0.000000e+00	9.897563e-04	5.307601e-02
##	GarageArea	0.000000e+00	3.252643e-10	1.263773e-02	1.379977e-02
##	WoodDeckSF	4.818368e-13	3.498384e-05	7.349737e-02	5.648744e-04
##	OpenPorchSF	0.000000e+00	1.332268e-14	3.315592e-04	7.380715e-03
##	EnclosedPorch	1.040475e-05	2.650862e-04	1.123474e-01	1.541632e-01
##	3SsnPorch	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
##	ScreenPorch	7.569616e-01	5.628841e-03	9.063268e-02	4.863598e-02
##	PoolArea	5.810529e-02	3.927890e-01	6.879729e-03	5.791972e-01
##	MiscVal	5.853585e-01	9.607167e-01	7.668285e-01	1.720427e-02
##	MoSold	3.278470e-02	7.297135e-01	7.542379e-02	3.099771e-01
##	YrSold	4.526681e-01	6.950290e-01	1.690210e-01	2.262661e-01
##	SalePriceLog	0.000000e+00	0.000000e+00	6.661338e-16	1.478820e-08
##	secondFloor	0.000000e+00	0.000000e+00	0.000000e+00	1.129716e-01
##	X1stFlrSF	0.000000e+00	4.334880e-06	1.041033e-06	9.243536e-03
##	X2ndFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	2.344382e-02
##	X3SsnPorch	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
##		TotRmsAbvGrd	Fireplaces	GarageYrBlt	GarageCars
##	MSSubClass	1.230183e-01	8.175109e-02	1.566879e-03	1.255479e-01
##	LotFrontage	0.000000e+00	0.000000e+00	1.834203e-02	0.000000e+00
##	LotArea	2.460254e-13	0.000000e+00	3.545897e-01	2.706817e-09
##	OverallQual	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	OverallCond	2.779336e-02	3.630812e-01	0.000000e+00	8.424372e-13
##	YearBuilt	2.544832e-04	1.423519e-08	0.000000e+00	0.000000e+00
##	YearRemodAdd	1.483258e-13	1.619633e-05	0.000000e+00	0.000000e+00
##	MasVnrArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	BsmtFinSF1	9.051660e-02	0.000000e+00	1.013706e-08	0.000000e+00
##	BsmtFinSF2	1.785383e-01	7.308663e-02	1.069187e-03	1.439250e-01

## BsmtUnfSF	0.000000e+00	4.880435e-02	9.254819e-13	0.000000e+00
## TotalBsmtSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## LowQualFinSF	4.904244e-07	4.166751e-01	1.771551e-01	3.002845e-04
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## BsmtFullBath	4.181609e-02	1.216473e-07	3.498871e-06	4.259910e-07
## BsmtHalfBath	3.627516e-01	2.685363e-01	3.997996e-03	4.250739e-01
## FullBath	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## HalfBath	0.000000e+00	3.996803e-15	1.665335e-13	0.000000e+00
## BedroomAbvGr	0.000000e+00	3.810705e-05	1.656563e-02	9.897563e-04
## KitchenAbvGr	0.000000e+00	2.035535e-06	3.589658e-06	5.307601e-02
## TotRmsAbvGrd	NA	0.000000e+00	3.283118e-08	0.000000e+00
## Fireplaces	0.000000e+00	NA	8.219437e-02	0.000000e+00
## GarageYrBlt	3.283118e-08	8.219437e-02	NA	0.000000e+00
## GarageCars	0.000000e+00	0.000000e+00	0.000000e+00	NA
## GarageArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## WoodDeckSF	1.758029e-10	1.221245e-14	0.000000e+00	0.000000e+00
## OpenPorchSF	0.000000e+00	7.290124e-11	0.000000e+00	0.000000e+00
## EnclosedPorch	8.740741e-01	3.432435e-01	0.000000e+00	6.067020e-09
## 3SsnPorch	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01
## ScreenPorch	2.326437e-02	1.195044e-12	5.077546e-03	5.373793e-02
## PoolArea	1.358848e-03	2.748854e-04	5.905625e-01	4.241329e-01
## MiscVal	3.443919e-01	9.571131e-01	2.289596e-01	9.987840e-02
## MoSold	1.586904e-01	7.660440e-02	8.430360e-01	1.217084e-01
## YrSold	1.874597e-01	3.575543e-01	9.699819e-01	1.351896e-01
## SalePriceLog	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## secondFloor	0.000000e+00	5.724116e-06	3.584860e-01	4.817098e-02
## X1stFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## X2ndFlrSF	0.000000e+00	6.394885e-14	8.506609e-03	1.417977e-12
## X3SsnPorch	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01
##	GarageArea	WoodDeckSF	OpenPorchSF	EnclosedPorch
## MSSubClass	1.592243e-04	6.310397e-01	8.158485e-01	6.458454e-01
## LotFrontage	0.000000e+00	2.136523e-03	1.210208e-07	7.110477e-01
## LotArea	3.802736e-12	4.001000e-11	1.185832e-03	4.837906e-01
## OverallQual	0.000000e+00	0.000000e+00	0.000000e+00	1.276996e-05
## OverallCond	5.945338e-09	8.987253e-01	2.133227e-01	7.159421e-03
## YearBuilt	0.000000e+00	0.000000e+00	3.623768e-13	0.000000e+00
## YearRemodAdd	0.000000e+00	1.998401e-15	0.000000e+00	7.771561e-14
## MasVnrArea	0.000000e+00	9.347851e-10	1.548562e-06	2.570285e-05
## BsmtFinSF1	0.000000e+00	3.108624e-15	1.867888e-05	9.005499e-05
## BsmtFinSF2	4.864925e-01	9.454565e-03	9.060156e-01	1.628382e-01
## BsmtUnfSF	1.690648e-12	8.391603e-01	7.585228e-07	9.228150e-01
## TotalBsmtSF	0.000000e+00	0.000000e+00	0.000000e+00	2.587670e-04
## LowQualFinSF	9.772127e-03	3.312869e-01	4.859080e-01	1.959094e-02
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	7.278971e-01
## BsmtFullBath	5.316858e-12	1.526557e-11	1.005792e-02	5.656746e-02
## BsmtHalfBath	3.488402e-01	1.250635e-01	3.335712e-01	7.439517e-01
## FullBath	0.000000e+00	4.818368e-13	0.000000e+00	1.040475e-05
## HalfBath	3.252643e-10	3.498384e-05	1.332268e-14	2.650862e-04
## BedroomAbvGr	1.263773e-02	7.349737e-02	3.315592e-04	1.123474e-01
## KitchenAbvGr	1.379977e-02	5.648744e-04	7.380715e-03	1.541632e-01
## TotRmsAbvGrd	0.000000e+00	1.758029e-10	0.000000e+00	8.740741e-01
## Fireplaces	0.000000e+00	1.221245e-14	7.290124e-11	3.432435e-01
## GarageYrBlt	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## GarageCars	0.000000e+00	0.000000e+00	0.000000e+00	6.067020e-09

## GarageArea	NA	0.000000e+00	0.000000e+00	3.064131e-06
## WoodDeckSF	0.000000e+00	NA	2.499829e-02	1.371106e-06
## OpenPorchSF	0.000000e+00	2.499829e-02	NA	3.691031e-04
## EnclosedPorch	3.064131e-06	1.371106e-06	3.691031e-04	NA
## 3SsnPorch	1.802697e-01	2.107784e-01	8.234948e-01	1.542416e-01
## ScreenPorch	4.952347e-02	4.569023e-03	4.502233e-03	1.529761e-03
## PoolArea	1.965912e-02	5.029192e-03	2.023972e-02	3.837510e-02
## MiscVal	2.954461e-01	7.153744e-01	4.779914e-01	4.832932e-01
## MoSold	2.854434e-01	4.224189e-01	6.454024e-03	2.699986e-01
## YrSold	2.958337e-01	3.951420e-01	2.769527e-02	7.050056e-01
## SalePriceLog	0.000000e+00	0.000000e+00	0.000000e+00	1.050805e-08
## secondFloor	9.467301e-01	2.184093e-01	5.305810e-08	5.719454e-04
## X1stFlrSF	0.000000e+00	0.000000e+00	4.440892e-16	1.258441e-02
## X2ndFlrSF	1.112965e-07	4.216853e-04	8.881784e-16	1.784426e-02
## X3SsnPorch	1.802697e-01	2.107784e-01	8.234948e-01	1.542416e-01
##	3SsnPorch	ScreenPorch	PoolArea	MiscVal
## MSSubClass	0.09414953	3.202578e-01	7.518384e-01	0.769269145
## LotFrontage	0.01520968	1.517836e-01	5.393463e-13	0.907190524
## LotArea	0.43552735	9.924777e-02	2.979905e-03	0.145989130
## OverallQual	0.24615822	1.314582e-02	1.275642e-02	0.230412263
## OverallCond	0.33014731	3.625216e-02	9.395944e-01	0.008568156
## YearBuilt	0.23118088	5.435519e-02	8.501179e-01	0.189169158
## YearRemodAdd	0.08367071	1.389957e-01	8.238849e-01	0.694533106
## MasVnrArea	0.47420644	1.916276e-02	6.553522e-01	0.256210660
## BsmtFinSF1	0.31250322	1.778536e-02	7.033176e-08	0.891545102
## BsmtFinSF2	0.25207571	6.748805e-04	1.111537e-01	0.850415539
## BsmtUnfSF	0.42789582	6.310420e-01	1.802008e-01	0.362745457
## TotalBsmtSF	0.15337655	1.232140e-03	1.354119e-06	0.480477906
## LowQualFinSF	0.86973491	3.061609e-01	1.753504e-02	0.884867376
## GrLivArea	0.43058934	1.021511e-04	5.918421e-11	0.926521232
## BsmtFullBath	0.99676835	3.767857e-01	9.756789e-03	0.378865709
## BsmtHalfBath	0.17993527	2.199645e-01	4.445328e-01	0.778528444
## FullBath	0.17698330	7.569616e-01	5.810529e-02	0.585358529
## HalfBath	0.84943689	5.628841e-03	3.927890e-01	0.960716747
## BedroomAbvGr	0.34997682	9.063268e-02	6.879729e-03	0.766828521
## KitchenAbvGr	0.34756881	4.863598e-02	5.791972e-01	0.017204270
## TotRmsAbvGrd	0.79860693	2.326437e-02	1.358848e-03	0.344391918
## Fireplaces	0.66735300	1.195044e-12	2.748854e-04	0.957113099
## GarageYrBlt	0.38232493	5.077546e-03	5.905625e-01	0.228959597
## GarageCars	0.17198411	5.373793e-02	4.241329e-01	0.099878400
## GarageArea	0.18026968	4.952347e-02	1.965912e-02	0.295446101
## WoodDeckSF	0.21077845	4.569023e-03	5.029192e-03	0.715374444
## OpenPorchSF	0.82349481	4.502233e-03	2.023972e-02	0.477991449
## EnclosedPorch	0.15424164	1.529761e-03	3.837510e-02	0.483293192
## 3SsnPorch	NA	2.299724e-01	7.602900e-01	0.989218187
## ScreenPorch	0.22997242	NA	4.998823e-02	0.222498354
## PoolArea	0.76029001	4.998823e-02	NA	0.257249528
## MiscVal	0.98921819	2.224984e-01	2.572495e-01	NA
## MoSold	0.26038906	3.753591e-01	1.976284e-01	0.804175819
## YrSold	0.47654325	6.830660e-01	2.256080e-02	0.851418697
## SalePriceLog	0.03594763	3.408972e-06	7.631787e-03	0.444619672
## secondFloor	0.33748531	5.241591e-01	1.622043e-01	0.840707205
## X1stFlrSF	0.03206515	6.856867e-04	4.578567e-07	0.420551089
## X2ndFlrSF	0.35234764	1.209302e-01	1.832441e-03	0.536316831

## X3SsnPorch	0.00000000	2.299724e-01	7.602900e-01	0.989218187
##	MoSold	YrSold	SalePriceLog	secondFloor
## MSSubClass	6.040063e-01	4.137257e-01	4.692451e-03	0.000000e+00
## LotFrontage	6.982029e-01	7.964813e-01	0.000000e+00	1.136398e-01
## LotArea	9.633078e-01	5.861053e-01	0.000000e+00	6.543549e-01
## OverallQual	6.790959e-03	2.963852e-01	0.000000e+00	1.513700e-11
## OverallCond	8.933753e-01	9.321243e-02	1.591324e-01	1.764981e-02
## YearBuilt	6.359583e-01	6.031268e-01	0.000000e+00	2.073907e-04
## YearRemodAdd	4.119179e-01	1.722487e-01	0.000000e+00	2.739984e-02
## MasVnrArea	8.203518e-01	7.548601e-01	0.000000e+00	1.313690e-02
## BsmtFinSF1	5.482075e-01	5.835463e-01	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.614175e-01	2.259961e-01	8.536300e-01	7.216583e-06
## BsmtUnfSF	1.827454e-01	1.150733e-01	0.000000e+00	1.797090e-01
## TotalBsmtSF	6.143911e-01	5.676653e-01	0.000000e+00	0.000000e+00
## LowQualFinSF	3.971935e-01	2.694414e-01	1.471041e-01	2.385267e-01
## GrLivArea	5.495607e-02	1.630401e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	3.328625e-01	1.038816e-02	0.000000e+00	0.000000e+00
## BsmtHalfBath	2.093598e-01	7.554937e-02	8.441568e-01	3.410671e-01
## FullBath	3.278470e-02	4.526681e-01	0.000000e+00	0.000000e+00
## HalfBath	7.297135e-01	6.950290e-01	0.000000e+00	0.000000e+00
## BedroomAbvGr	7.542379e-02	1.690210e-01	6.661338e-16	0.000000e+00
## KitchenAbvGr	3.099771e-01	2.262661e-01	1.478820e-08	1.129716e-01
## TotRmsAbvGrd	1.586904e-01	1.874597e-01	0.000000e+00	0.000000e+00
## Fireplaces	7.660440e-02	3.575543e-01	0.000000e+00	5.724116e-06
## GarageYrBlt	8.430360e-01	9.699819e-01	0.000000e+00	3.584860e-01
## GarageCars	1.217084e-01	1.351896e-01	0.000000e+00	4.817098e-02
## GarageArea	2.854434e-01	2.958337e-01	0.000000e+00	9.467301e-01
## WoodDeckSF	4.224189e-01	3.951420e-01	0.000000e+00	2.184093e-01
## OpenPorchSF	6.454024e-03	2.769527e-02	0.000000e+00	5.305810e-08
## EnclosedPorch	2.699986e-01	7.050056e-01	1.050805e-08	5.719454e-04
## 3SsnPorch	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
## ScreenPorch	3.753591e-01	6.830660e-01	3.408972e-06	5.241591e-01
## PoolArea	1.976284e-01	2.256080e-02	7.631787e-03	1.622043e-01
## MiscVal	8.041758e-01	8.514187e-01	4.446197e-01	8.407072e-01
## MoSold	NA	2.230277e-08	2.848930e-02	3.497548e-01
## YrSold	2.230277e-08	NA	1.547100e-01	6.406425e-01
## SalePriceLog	2.848930e-02	1.547100e-01	NA	7.446401e-09
## secondFloor	3.497548e-01	6.406425e-01	7.446401e-09	NA
## X1stFlrSF	2.309272e-01	6.034971e-01	0.000000e+00	0.000000e+00
## X2ndFlrSF	1.793059e-01	2.731188e-01	0.000000e+00	0.000000e+00
## X3SsnPorch	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
##	X1stFlrSF	X2ndFlrSF	X3SsnPorch	
## MSSubClass	0.000000e+00	0.000000e+00	0.09414953	
## LotFrontage	0.000000e+00	5.433170e-03	0.01520968	
## LotArea	0.000000e+00	5.144266e-02	0.43552735	
## OverallQual	0.000000e+00	0.000000e+00	0.24615822	
## OverallCond	3.126349e-08	2.690898e-01	0.33014731	
## YearBuilt	0.000000e+00	6.939294e-01	0.23118088	
## YearRemodAdd	0.000000e+00	7.777955e-08	0.08367071	
## MasVnrArea	0.000000e+00	2.121014e-11	0.47420644	
## BsmtFinSF1	0.000000e+00	1.455248e-07	0.31250322	
## BsmtFinSF2	2.020320e-04	1.453587e-04	0.25207571	
## BsmtUnfSF	0.000000e+00	8.645239e-01	0.42789582	
## TotalBsmtSF	0.000000e+00	1.894085e-11	0.15337655	

```
## LowQualFinSF 5.866501e-01 1.547482e-02 0.86973491
## GrLivArea 0.000000e+00 0.000000e+00 0.43058934
## BsmtFullBath 0.000000e+00 7.124057e-11 0.99676835
## BsmtHalfBath 9.404841e-01 3.623802e-01 0.17993527
## FullBath 0.000000e+00 0.000000e+00 0.17698330
## HalfBath 4.334880e-06 0.000000e+00 0.84943689
## BedroomAbvGr 1.041033e-06 0.000000e+00 0.34997682
## KitchenAbvGr 9.243536e-03 2.344382e-02 0.34756881
## TotRmsAbvGrd 0.000000e+00 0.000000e+00 0.79860693
## Fireplaces 0.000000e+00 6.394885e-14 0.66735300
## GarageYrBlt 0.000000e+00 8.506609e-03 0.38232493
## GarageCars 0.000000e+00 1.417977e-12 0.17198411
## GarageArea 0.000000e+00 1.112965e-07 0.18026968
## WoodDeckSF 0.000000e+00 4.216853e-04 0.21077845
## OpenPorchSF 4.440892e-16 8.881784e-16 0.82349481
## EnclosedPorch 1.258441e-02 1.784426e-02 0.15424164
## 3SsnPorch 3.206515e-02 3.523476e-01 0.00000000
## ScreenPorch 6.856867e-04 1.209302e-01 0.22997242
## PoolArea 4.578567e-07 1.832441e-03 0.76029001
## MiscVal 4.205511e-01 5.363168e-01 0.98921819
## MoSold 2.309272e-01 1.793059e-01 0.26038906
## YrSold 6.034971e-01 2.731188e-01 0.47654325
## SalePriceLog 0.000000e+00 0.000000e+00 0.03594763
## secondFloor 0.000000e+00 0.000000e+00 0.33748531
## X1stFlrSF NA 5.329071e-15 0.03206515
## X2ndFlrSF 5.329071e-15 NA 0.35234764
## X3SsnPorch 3.206515e-02 3.523476e-01 NA
```

```
# It is a list of matrices with names "r", "n" and "P"
c(typeof(res2),names(res2))
```

```
## [1] "list" "r" "n" "P"
```

```
# Identify the significant correlations
dim(res2$P)
```

```
## [1] 39 39
```

```
res2_price_p <- res2$P[response,]
res2_price_p
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond
## 4.692451e-03 0.000000e+00 0.000000e+00 0.000000e+00 1.591324e-01
## YearBuilt YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2
## 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 8.536300e-01
## BsmtUnfSF TotalBsmtSF LowQualFinSF GrLivArea BsmtFullBath
## 0.000000e+00 0.000000e+00 1.471041e-01 0.000000e+00 0.000000e+00
## BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr
## 8.441568e-01 0.000000e+00 0.000000e+00 6.661338e-16 1.478820e-08
## TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars GarageArea
## 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## WoodDeckSF OpenPorchSF EnclosedPorch 3SsnPorch ScreenPorch
## 0.000000e+00 0.000000e+00 1.050805e-08 3.594763e-02 3.408972e-06
## PoolArea MiscVal MoSold YrSold SalePriceLog
## 7.631787e-03 4.446197e-01 2.848930e-02 1.547100e-01 NA
## secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch
```

```
## 7.446401e-09 0.000000e+00 0.000000e+00 3.594763e-02
```

```
correlation_significant<-res2_price_p <= sig & !is.na(res2_price_p)
```

Subset to only fields significantly correlated with there response

```
# extract the correlations of the response to other fields
```

```
res2_price_r <- res2$r[response,]
```

```
res2_price_r
```

```
##      MSSubClass      LotFrontage      LotArea      OverallQual      OverallCond
## -0.073958993      0.355878472      0.257319897      0.817184448      -0.036867980
##      YearBuilt      YearRemodAdd      MasVnrArea      BsmtFinSF1      BsmtFinSF2
## 0.586570263      0.565607846      0.430808514      0.372023076      0.004832403
##      BsmtUnfSF      TotalBsmtSF      LowQualFinSF      GrLivArea      BsmtFullBath
## 0.221985057      0.612133980      -0.037962791      0.700926661      0.236224100
##      BsmtHalfBath      FullBath      HalfBath      BedroomAbvGr      KitchenAbvGr
## -0.005149092      0.594770551      0.313982278      0.209043682      -0.147548094
##      TotRmsAbvGrd      Fireplaces      GarageYrBlt      GarageCars      GarageArea
## 0.534422219      0.489449441      0.541072845      0.680624783      0.650887549
##      WoodDeckSF      OpenPorchSF      EnclosedPorch      3SsnPorch      ScreenPorch
## 0.334135085      0.321052969      -0.149050295      0.054900210      0.121207602
##      PoolArea      MiscVal      MoSold      YrSold      SalePriceLog
## 0.069797799      -0.020020811      0.057329483      -0.037262958      1.000000000
##      secondFloor      X1stFlrSF      X2ndFlrSF      X3SsnPorch
## 0.150549427      0.596981049      0.319299996      0.054900210
```

```
# Select the significant correlations
```

```
features_significant <- res2_price_r[correlation_significant]
```

```
features_significant
```

```
##      LotFrontage      LotArea      OverallQual      YearBuilt      YearRemodAdd
## 0.3558785      0.2573199      0.8171844      0.5865703      0.5656078
##      MasVnrArea      BsmtFinSF1      BsmtUnfSF      TotalBsmtSF      GrLivArea
## 0.4308085      0.3720231      0.2219851      0.6121340      0.7009267
##      BsmtFullBath      FullBath      HalfBath      BedroomAbvGr      KitchenAbvGr
## 0.2362241      0.5947706      0.3139823      0.2090437      -0.1475481
##      TotRmsAbvGrd      Fireplaces      GarageYrBlt      GarageCars      GarageArea
## 0.5344222      0.4894494      0.5410728      0.6806248      0.6508875
##      WoodDeckSF      OpenPorchSF      EnclosedPorch      ScreenPorch      secondFloor
## 0.3341351      0.3210530      -0.1490503      0.1212076      0.1505494
##      X1stFlrSF      X2ndFlrSF
## 0.5969810      0.3193000
```

```
# Sort by correlation
```

```
features_significant <- features_significant[order(abs(features_significant),decreasing=TRUE)]
```

```
features_significant
```

```
##      OverallQual      GrLivArea      GarageCars      GarageArea      TotalBsmtSF
## 0.8171844      0.7009267      0.6806248      0.6508875      0.6121340
##      X1stFlrSF      FullBath      YearBuilt      YearRemodAdd      GarageYrBlt
## 0.5969810      0.5947706      0.5865703      0.5656078      0.5410728
##      TotRmsAbvGrd      Fireplaces      MasVnrArea      BsmtFinSF1      LotFrontage
## 0.5344222      0.4894494      0.4308085      0.3720231      0.3558785
##      WoodDeckSF      OpenPorchSF      X2ndFlrSF      HalfBath      LotArea
## 0.3341351      0.3210530      0.3193000      0.3139823      0.2573199
##      BsmtFullBath      BsmtUnfSF      BedroomAbvGr      secondFloor      EnclosedPorch
```

```
##      0.2362241      0.2219851      0.2090437      0.1505494      -0.1490503
## KitchenAbvGr    ScreenPorch
##      -0.1475481      0.1212076
```

```
names(features_significant)
```

```
## [1] "OverallQual" "GrLivArea" "GarageCars" "GarageArea"
## [5] "TotalBsmtSF" "X1stFlrSF" "FullBath" "YearBuilt"
## [9] "YearRemodAdd" "GarageYrBlt" "TotRmsAbvGrd" "Fireplaces"
## [13] "MasVnrArea" "BsmtFinSF1" "LotFrontage" "WoodDeckSF"
## [17] "OpenPorchSF" "X2ndFlrSF" "HalfBath" "LotArea"
## [21] "BsmtFullBath" "BsmtUnfSF" "BedroomAbvGr" "secondFloor"
## [25] "EnclosedPorch" "KitchenAbvGr" "ScreenPorch"
```

```
remove_response <- !names(features_significant) %in% c('SalePriceLog',response)
remove_response
```

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

```
features_significant <- features_significant[remove_response]
features_significant
```

```
## OverallQual    GrLivArea    GarageCars    GarageArea    TotalBsmtSF
##      0.8171844    0.7009267    0.6806248    0.6508875    0.6121340
##      X1stFlrSF    FullBath    YearBuilt    YearRemodAdd    GarageYrBlt
##      0.5969810    0.5947706    0.5865703    0.5656078    0.5410728
## TotRmsAbvGrd    Fireplaces    MasVnrArea    BsmtFinSF1    LotFrontage
##      0.5344222    0.4894494    0.4308085    0.3720231    0.3558785
##      WoodDeckSF    OpenPorchSF    X2ndFlrSF    HalfBath    LotArea
##      0.3341351    0.3210530    0.3193000    0.3139823    0.2573199
## BsmtFullBath    BsmtUnfSF    BedroomAbvGr    secondFloor    EnclosedPorch
##      0.2362241    0.2219851    0.2090437    0.1505494    -0.1490503
## KitchenAbvGr    ScreenPorch
##      -0.1475481    0.1212076
```

```
#redo the t_selected
```

```
t_selected <- t_numeric %>% select(c(names(features_significant)[1:num_of_canadate_features],response))
names(t_selected)
```

```
## [1] "OverallQual" "GrLivArea" "GarageCars" "GarageArea"
## [5] "TotalBsmtSF" "X1stFlrSF" "FullBath" "YearBuilt"
## [9] "YearRemodAdd" "GarageYrBlt" "TotRmsAbvGrd" "Fireplaces"
## [13] "MasVnrArea" "BsmtFinSF1" "LotFrontage" "SalePriceLog"
```

```
# Note that the correlation for OverallQual is zero. It was the field used last time, so the correlati
#should be zero
```

```
features_significant
```

```
## OverallQual    GrLivArea    GarageCars    GarageArea    TotalBsmtSF
##      0.8171844    0.7009267    0.6806248    0.6508875    0.6121340
##      X1stFlrSF    FullBath    YearBuilt    YearRemodAdd    GarageYrBlt
##      0.5969810    0.5947706    0.5865703    0.5656078    0.5410728
## TotRmsAbvGrd    Fireplaces    MasVnrArea    BsmtFinSF1    LotFrontage
##      0.5344222    0.4894494    0.4308085    0.3720231    0.3558785
##      WoodDeckSF    OpenPorchSF    X2ndFlrSF    HalfBath    LotArea
##      0.3341351    0.3210530    0.3193000    0.3139823    0.2573199
```

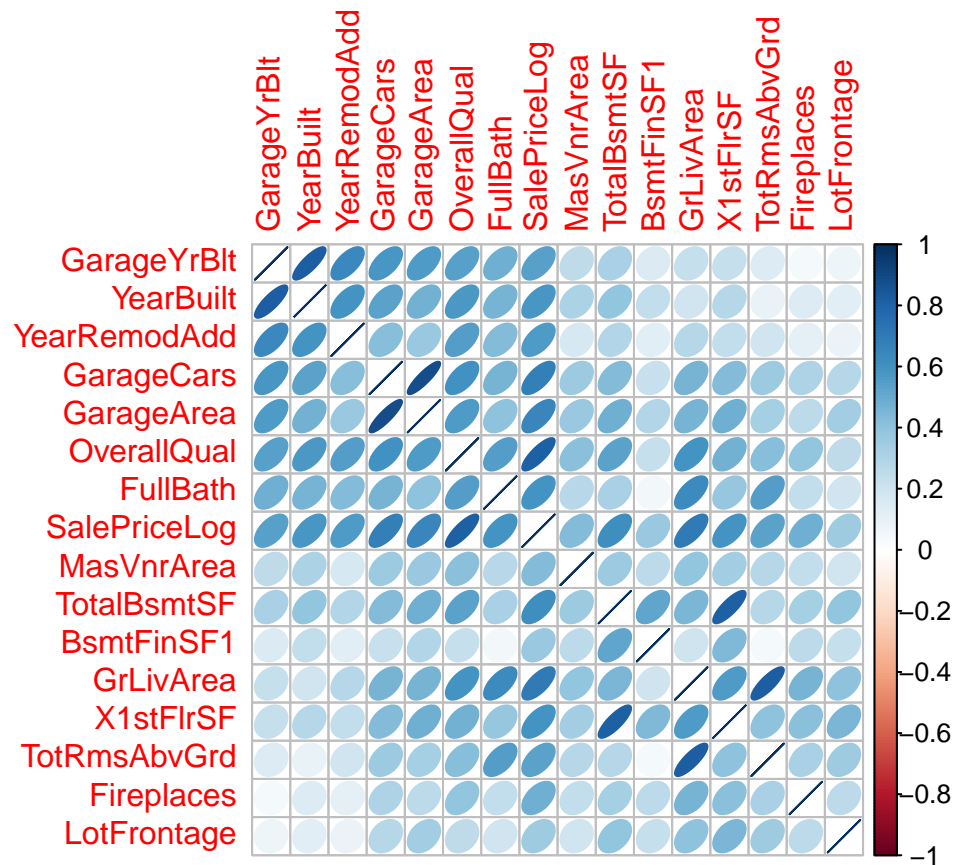
```
## BsmtFullBath BsmtUnfSF BedroomAbvGr secondFloor EnclosedPorch
## 0.2362241 0.2219851 0.2090437 0.1505494 -0.1490503
## KitchenAbvGr ScreenPorch
## -0.1475481 0.1212076
```

```
names(t_selected)
```

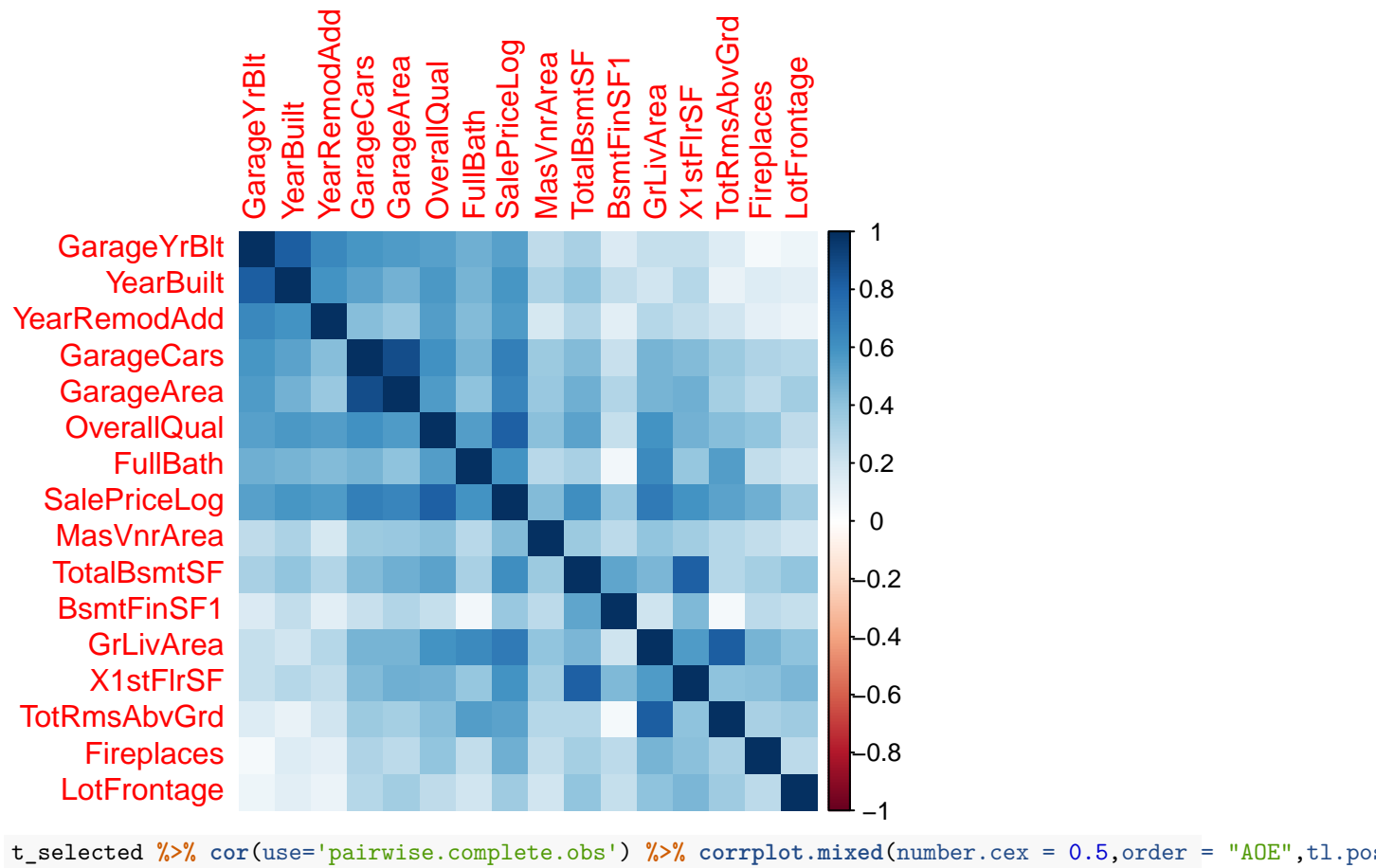
```
## [1] "OverallQual" "GrLivArea" "GarageCars" "GarageArea"
## [5] "TotalBsmtSF" "X1stFlrSF" "FullBath" "YearBuilt"
## [9] "YearRemodAdd" "GarageYrBlt" "TotRmsAbvGrd" "Fireplaces"
## [13] "MasVnrArea" "BsmtFinSF1" "LotFrontage" "SalePriceLog"
```

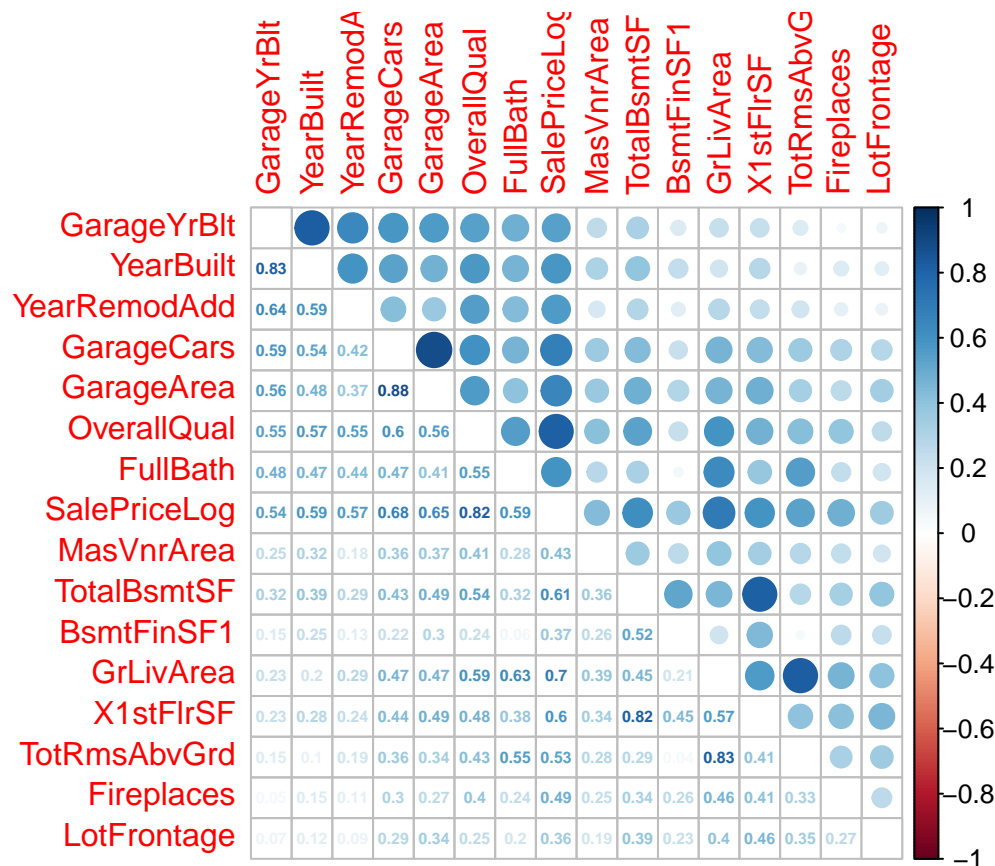
A few versions of the correlation matrix

```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='ellipse',order="AOE")
```



```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='shade',order="AOE")
```





```
#t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='circle',method='ellipse')
```

Here is the top feature to try

```
feature_list <-list()
selected<-names(features_significant)[1]
selected

## [1] "OverallQual"
selected_label<-label(t_numeric[selected])
selected_label
```

```
## OverallQual
## "Overall material and finish quality"
feature_list[[model_num]] <- selected
str(feature_list)
```

```
## List of 1
## $ : chr "OverallQual"
```

Calculate a model using this variable

```
names(t_numeric)

## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
```



```
## [17] "FullBath"      "HalfBath"      "BedroomAbvGr"  "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces"    "GarageYrBlt"   "GarageCars"
## [25] "GarageArea"    "WoodDeckSF"    "OpenPorchSF"   "EnclosedPorch"
## [29] "3SsnPorch"     "ScreenPorch"   "PoolArea"      "MiscVal"
## [33] "MoSold"        "YrSold"        "SalePriceLog"  "secondFloor"
## [37] "X1stFlrSF"     "X2ndFlrSF"     "X3SsnPorch"
```

```
try(t_numeric <- t_numeric %>% select(-.hat,-.sigma,-.cooks, -.std.resid, -.se.fit, -.fitted, -.resid, -.inf)
names(t_numeric)
```

```
## [1] "MSSubClass"    "LotFrontage"   "LotArea"       "OverallQual"
## [5] "OverallCond"   "YearBuilt"     "YearRemodAdd"  "MasVnrArea"
## [9] "BsmtFinSF1"    "BsmtFinSF2"    "BsmtUnfSF"     "TotalBsmtSF"
## [13] "LowQualFinSF"  "GrLivArea"     "BsmtFullBath"  "BsmtHalfBath"
## [17] "FullBath"      "HalfBath"      "BedroomAbvGr"  "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces"    "GarageYrBlt"   "GarageCars"
## [25] "GarageArea"    "WoodDeckSF"    "OpenPorchSF"   "EnclosedPorch"
## [29] "3SsnPorch"     "ScreenPorch"   "PoolArea"      "MiscVal"
## [33] "MoSold"        "YrSold"        "SalePriceLog"  "secondFloor"
## [37] "X1stFlrSF"     "X2ndFlrSF"     "X3SsnPorch"
```

```
feature_list
```

```
## [[1]]
## [1] "OverallQual"
```

```
my_formula <- paste0('SalePriceLog ~ ',paste(feature_list,collapse = '+ ', sep=" "))
str(my_formula)
```

```
## chr "SalePriceLog ~ OverallQual"
```

```
model_1 <- t_numeric %>% lm(my_formula,data=.)
t_numeric<- model_1 %>% augment(type.predict = 'response',type.residuals = 'response',data=t_numeric) %>%
  mutate(model_1_resid = .resid) %>%
  mutate(influential = .cooks > 4/length(SalePriceLog))
head(t_numeric)
```

```
##   MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 1         60         65   8450           7           5      2003
## 2         20         80   9600           6           8      1976
## 3         60         68  11250           7           5      2001
## 4         70         60   9550           7           5      1915
## 5         60         84  14260           8           5      2000
## 6         50         85  14115           5           5      1993
##   YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1          2003         196        706         0         150         856
## 2          1976           0        978         0         284        1262
## 3          2002         162        486         0         434         920
## 4          1970           0        216         0         540         756
## 5          2000         350        655         0         490        1145
## 6          1995           0        732         0          64         796
##   LowQualFinSF GrLivArea BsmtFullBath BsmtHalfBath FullBath HalfBath
## 1             0     1710           1             0           2           1
## 2             0     1262           0             1           2           0
## 3             0     1786           1             0           2           1
## 4             0     1717           1             0           1           0
## 5             0     2198           1             0           2           1
```

```
## 6      0      1362      1      0      1      1
## BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars
## 1      3      1      8      0      2003      2
## 2      3      1      6      1      1976      2
## 3      3      1      6      1      2001      2
## 4      3      1      7      1      1998      3
## 5      4      1      9      1      2000      3
## 6      1      1      5      0      1993      2
## GarageArea WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch.1 ScreenPorch
## 1      548      0      61      0      0      0
## 2      460      298      0      0      0      0
## 3      608      0      42      0      0      0
## 4      642      0      35      272      0      0
## 5      836      192      84      0      0      0
## 6      480      40      30      0      320      0
## PoolArea MiscVal MoSold YrSold SalePriceLog secondFloor X1stFlrSF
## 1      0      0      2      2008      12.24769      TRUE      856
## 2      0      0      5      2007      12.10901      FALSE      1262
## 3      0      0      9      2008      12.31717      TRUE      920
## 4      0      0      2      2006      11.84940      TRUE      961
## 5      0      0      12      2008      12.42922      TRUE      1145
## 6      0      700      10      2009      11.87060      TRUE      796
## X2ndFlrSF X3SsnPorch .fitted .se.fit .resid .hat
## 1      854      0 12.23664 0.007193874 0.01105660 0.0009756337
## 2      0      0 12.00061 0.006043113 0.10840116 0.0006884660
## 3      866      0 12.23664 0.007193874 0.08052897 0.0009756337
## 4      756      0 12.23664 0.007193874 -0.38724002 0.0009756337
## 5     1053      0 12.47267 0.010246994 -0.04344948 0.0019794933
## 6      566     320 11.76458 0.007700847 0.10601809 0.0011179903
## .sigma .cooksd .std.resid model_1_resid influential
## 1 0.2303924 1.126442e-06 0.04803017 0.01105660 FALSE
## 2 0.2303750 7.636243e-05 0.47082991 0.10840116 FALSE
## 3 0.2303829 5.975442e-05 0.34982003 0.08052897 FALSE
## 4 0.2301689 1.381740e-03 -1.68218108 -0.38724002 FALSE
## 5 0.2303897 3.536513e-05 -0.18884060 -0.04344948 FALSE
## 6 0.2303758 1.187137e-04 0.46057827 0.10601809 FALSE
```

Check the fit using the Kaggle criteria

```
rmse(model_1,t_numeric)
```

```
## [1] 0.2301557
```

```
t_numeric[t_numeric$influential,]
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 30      30      60      6324      4      6      1927
## 31      70      50      8500      4      4      1920
## 57     160      24      2645      8      5      1999
## 75      50      60      5790      3      6      1915
## 113     60      77      9965      7      5      2007
## 126    190      60      6780      6      8      1935
## 156     50      60      9600      6      5      1924
## 179     20      63     17423      9      5      2008
## 232     60     174     15138      8      5      1995
## 265     30      30      5232      5      5      1925
```

## 314	20	150	215245	7	5	1965
## 324	20	49	5820	3	8	1955
## 336	190	NA	164660	5	6	1965
## 350	60	56	20431	9	5	2005
## 376	30	NA	10020	1	1	1922
## 384	45	60	9000	6	3	1928
## 399	30	60	8967	5	2	1920
## 406	20	NA	9991	4	4	1976
## 411	20	68	9571	5	3	1958
## 432	50	60	5586	6	7	1920
## 441	20	105	15431	10	5	2008
## 450	50	50	6000	3	7	1948
## 458	20	NA	53227	4	6	1954
## 459	70	NA	5100	8	7	1925
## 463	20	60	8281	5	5	1965
## 474	20	110	14977	8	5	2006
## 496	30	60	7879	4	5	1920
## 497	20	NA	12692	8	5	1992
## 524	60	130	40094	10	5	2007
## 528	60	67	14948	9	5	2008
## 534	20	50	5000	1	3	1946
## 535	60	74	9056	8	5	2004
## 584	75	75	13500	10	9	1893
## 599	20	80	12984	5	6	1977
## 608	20	78	7800	5	8	1948
## 633	20	85	11900	7	5	1977
## 662	60	52	46589	8	7	1994
## 665	20	49	20896	8	5	2005
## 689	20	60	8089	8	6	2007
## 692	60	104	21535	10	6	1994
## 706	190	70	5600	4	5	1930
## 711	30	56	4130	3	6	1935
## 770	60	47	53504	8	5	2003
## 775	20	110	14226	8	5	2006
## 799	60	104	13518	9	5	2008
## 804	60	107	13891	9	5	2008
## 808	70	144	21384	5	6	1923
## 813	20	66	8712	5	5	1952
## 875	50	52	5720	5	6	1941
## 899	20	100	12919	9	5	2009
## 917	20	50	9000	2	3	1949
## 943	90	42	7711	4	3	1977
## 964	20	122	11923	9	5	2007
## 969	50	50	5925	3	6	1910
## 992	70	121	17671	8	9	1882
## 1021	20	60	7024	4	5	2005
## 1046	20	NA	13680	3	5	1955
## 1047	60	85	16056	9	5	2005
## 1098	120	NA	3696	8	5	1986
## 1143	60	77	9965	8	5	2006
## 1149	50	NA	5700	7	7	1926
## 1156	20	90	10768	5	8	1976
## 1170	60	118	35760	10	5	1995
## 1182	120	64	5587	8	5	2008

## 1183	60	160	15623	10	5	1996
## 1191	190	NA	32463	4	4	1961
## 1192	160	24	2645	8	5	1999
## 1200	20	75	11235	4	5	1963
## 1212	50	152	12134	8	7	1988
## 1263	50	NA	11250	4	5	1957
## 1280	50	60	7500	4	4	1920
## 1299	60	313	63887	10	5	2008
## 1325	20	75	9986	8	5	2006
## 1326	30	40	3636	4	4	1922
## 1338	30	153	4118	4	4	1941
## 1350	70	50	5250	8	5	1872
## 1354	50	56	14720	8	5	1995
## 1379	160	21	1953	6	5	1973
## 1381	30	45	8212	3	3	1914
## 1405	50	60	10410	3	4	1915
## 1433	30	60	10800	4	6	1927
## 1438	20	96	12444	8	5	2008
## 1443	60	85	11003	10	5	2008
##	YearRemodAdd	MasVnrArea	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF
## 30	1950	0	0	0	520	520
## 31	1950	0	0	0	649	649
## 57	2000	456	649	0	321	970
## 75	1950	0	0	0	840	840
## 113	2007	220	984	0	280	1264
## 126	1982	0	490	0	30	520
## 156	1950	0	0	0	572	572
## 179	2009	748	1904	0	312	2216
## 232	1996	506	689	0	773	1462
## 265	2004	0	0	0	680	680
## 314	1965	0	1236	820	80	2136
## 324	2005	0	256	0	906	1162
## 336	1965	0	1249	147	103	1499
## 350	2006	870	1410	0	438	1848
## 376	1950	0	350	0	333	683
## 384	1950	0	0	0	784	784
## 399	1950	0	0	0	961	961
## 406	1993	0	1116	0	165	1281
## 411	1958	0	0	0	1276	1276
## 432	1998	0	0	0	901	901
## 441	2008	200	1767	539	788	3094
## 450	2002	0	331	0	318	649
## 458	1994	0	1116	0	248	1364
## 459	1996	0	0	0	588	588
## 463	1965	0	553	311	0	864
## 474	2007	304	1350	0	626	1976
## 496	1950	0	495	0	225	720
## 497	1993	0	1231	0	1969	3200
## 524	2008	762	2260	0	878	3138
## 528	2008	268	1330	0	122	1452
## 534	1950	0	0	0	0	0
## 535	2004	0	0	0	707	707
## 584	2000	0	0	0	1237	1237
## 599	1977	459	1283	147	0	1430

## 608	2002	0	603	0	293	896
## 633	1977	209	822	0	564	1386
## 662	2005	528	1361	180	88	1629
## 665	2006	0	1721	0	356	2077
## 689	2007	0	945	0	474	1419
## 692	1995	1170	1455	0	989	2444
## 706	1950	0	0	0	0	0
## 711	2003	0	0	0	270	270
## 770	2003	603	1416	0	234	1650
## 775	2006	375	0	0	1935	1935
## 799	2009	860	0	0	1926	1926
## 804	2009	424	0	0	1734	1734
## 808	2004	0	1309	0	15	1324
## 813	1952	0	0	0	540	540
## 875	1950	0	0	0	676	676
## 899	2010	760	2188	0	142	2330
## 917	1950	0	50	0	430	480
## 943	1977	0	1440	0	0	1440
## 964	2007	0	0	0	1800	1800
## 969	1950	0	0	0	600	600
## 992	1986	0	216	0	700	916
## 1021	2005	0	1024	0	108	1132
## 1046	1955	0	0	0	0	0
## 1047	2006	208	240	0	1752	1992
## 1098	1986	0	0	0	1074	1074
## 1143	2007	340	1150	0	316	1466
## 1149	1950	0	0	0	572	572
## 1156	2004	0	1157	0	280	1437
## 1170	1996	1378	1387	0	543	1930
## 1182	2008	186	1480	0	120	1600
## 1183	1996	0	2096	0	300	2396
## 1191	1975	149	1159	0	90	1249
## 1192	2000	456	0	0	776	776
## 1200	1979	51	547	0	504	1051
## 1212	2005	0	427	0	132	559
## 1263	1989	0	0	0	1104	1104
## 1280	1950	0	0	0	698	698
## 1299	2008	796	5644	0	466	6110
## 1325	2007	428	0	0	1795	1795
## 1326	1950	0	0	0	796	796
## 1338	1950	0	0	0	693	693
## 1350	1987	0	259	0	425	684
## 1354	1996	579	816	0	1217	2033
## 1379	1973	408	309	0	174	483
## 1381	1950	0	203	0	661	864
## 1405	1950	0	0	0	672	672
## 1433	2007	0	0	0	656	656
## 1438	2008	426	1336	0	596	1932
## 1443	2008	160	765	0	252	1017
##	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	FullBath	HalfBath
## 30	0	520	0	0	1	0
## 31	0	1317	0	0	1	0
## 57	0	1739	1	0	2	1
## 75	0	1605	0	0	2	0

## 113	0	2696	1	0	2	1
## 126	234	754	1	0	1	0
## 156	0	1096	0	0	1	0
## 179	0	2234	1	0	2	0
## 232	0	2794	1	0	2	1
## 265	0	764	0	0	1	0
## 314	0	2036	2	0	2	0
## 324	0	1163	1	0	1	0
## 336	0	1786	2	0	2	0
## 350	0	2728	1	0	2	1
## 376	0	904	1	0	0	1
## 384	0	784	0	0	1	0
## 399	0	1077	0	0	1	0
## 406	0	1620	1	0	2	0
## 411	0	1276	0	0	1	0
## 432	0	1198	0	0	1	0
## 441	0	2402	1	0	2	0
## 450	0	1183	0	0	1	1
## 458	0	1663	1	0	1	0
## 459	0	1666	0	0	1	0
## 463	0	864	0	0	1	0
## 474	0	1976	1	0	2	0
## 496	0	720	0	0	1	0
## 497	0	3228	1	0	3	0
## 524	0	4676	1	0	3	1
## 528	0	2713	1	0	2	1
## 534	0	334	0	0	1	0
## 535	0	1414	0	0	2	1
## 584	0	2775	0	0	3	1
## 599	0	1647	1	0	2	0
## 608	0	2008	1	0	3	0
## 633	0	1411	0	0	2	0
## 662	0	2448	1	0	2	1
## 665	0	2097	1	0	1	1
## 689	0	1419	1	0	2	0
## 692	0	4316	0	1	3	1
## 706	0	1092	0	0	2	0
## 711	0	729	0	0	1	0
## 770	0	3279	1	0	3	1
## 775	0	1973	0	0	2	0
## 799	0	3140	0	0	3	1
## 804	0	2822	0	0	3	1
## 808	0	1576	2	0	1	1
## 813	0	1044	0	0	1	0
## 875	0	1131	0	0	1	1
## 899	0	2364	1	0	2	1
## 917	0	480	1	0	0	0
## 943	0	1440	2	0	2	0
## 964	0	1800	0	0	2	0
## 969	0	968	0	0	1	0
## 992	0	1742	0	0	1	1
## 1021	0	1132	1	0	1	1
## 1046	0	1733	0	0	2	0
## 1047	0	2868	0	0	3	1

## 1098	0	1088	0	0	1	1
## 1143	0	2828	1	0	3	0
## 1149	0	1111	0	0	1	0
## 1156	0	1437	1	0	2	0
## 1170	0	3627	1	0	3	1
## 1182	0	1652	1	1	2	0
## 1183	0	4476	1	0	3	1
## 1191	0	1622	1	0	1	0
## 1192	0	1441	0	0	2	1
## 1200	0	1382	0	0	1	1
## 1212	0	1752	0	0	2	0
## 1263	0	1788	1	0	1	0
## 1280	0	1128	0	0	1	0
## 1299	0	5642	2	0	2	1
## 1325	0	1795	0	0	2	0
## 1326	0	796	0	0	1	0
## 1338	0	693	0	0	1	0
## 1350	205	2358	0	0	2	0
## 1354	0	3238	1	0	2	1
## 1379	0	987	0	0	1	1
## 1381	0	864	1	0	1	0
## 1405	0	1214	0	0	1	0
## 1433	0	968	0	0	2	0
## 1438	0	1932	1	0	2	0
## 1443	0	2007	1	0	2	1
##	BedroomAbvGr	KitchenAbvGr	TotRmsAbvGrd	Fireplaces	GarageYrBlt	
## 30	1	1	4	0	1920	
## 31	3	1	6	0	1920	
## 57	3	1	7	0	1999	
## 75	3	2	8	0	1915	
## 113	4	1	10	1	2007	
## 126	2	1	5	0	NA	
## 156	2	1	5	0	NA	
## 179	1	1	9	1	2009	
## 232	4	1	9	1	1995	
## 265	2	1	4	0	1965	
## 314	3	1	8	2	1965	
## 324	3	1	6	0	1955	
## 336	3	1	7	2	1965	
## 350	4	1	10	2	2006	
## 376	1	1	4	0	NA	
## 384	2	1	5	0	1950	
## 399	2	1	6	0	1920	
## 406	3	1	8	1	1993	
## 411	3	1	5	0	1958	
## 432	4	1	7	0	NA	
## 441	2	1	10	2	2008	
## 450	2	1	6	0	1981	
## 458	2	1	6	2	1954	
## 459	3	1	7	1	1925	
## 463	3	1	5	1	1965	
## 474	2	1	7	1	2006	
## 496	2	1	4	0	NA	
## 497	4	1	10	1	1992	

## 524	3	1	11	1	2007
## 528	3	1	11	1	2008
## 534	1	1	2	0	NA
## 535	3	1	6	1	2004
## 584	3	1	9	1	1988
## 599	3	1	7	1	1977
## 608	3	1	8	0	1948
## 633	3	1	6	1	1977
## 662	4	1	8	1	1994
## 665	1	1	8	1	2005
## 689	2	1	7	1	2007
## 692	4	1	10	2	1994
## 706	3	2	7	0	NA
## 711	2	1	5	0	NA
## 770	4	1	12	1	2003
## 775	3	1	9	1	2006
## 799	4	1	11	2	2009
## 804	4	1	12	1	2009
## 808	3	1	6	1	1923
## 813	2	1	4	0	1952
## 875	3	1	5	0	1941
## 899	2	1	11	2	2009
## 917	1	1	4	0	1958
## 943	4	2	8	0	NA
## 964	2	1	7	0	2007
## 969	2	1	6	0	NA
## 992	4	1	8	1	1925
## 1021	2	1	5	0	2005
## 1046	4	1	8	1	1955
## 1047	4	1	11	1	2005
## 1098	2	1	5	0	1987
## 1143	4	1	11	1	2006
## 1149	2	1	5	1	1982
## 1156	3	1	6	1	1976
## 1170	4	1	10	1	1995
## 1182	2	1	5	1	2008
## 1183	4	1	10	2	1996
## 1191	3	1	7	1	1975
## 1192	2	1	5	0	1999
## 1200	3	1	6	1	1974
## 1212	4	1	8	0	1988
## 1263	5	1	8	2	1957
## 1280	2	1	6	0	1980
## 1299	3	1	12	3	2008
## 1325	2	1	7	1	2007
## 1326	2	1	5	0	NA
## 1338	2	1	4	0	NA
## 1350	4	1	8	0	NA
## 1354	4	1	9	1	1996
## 1379	2	1	5	0	1973
## 1381	2	1	5	0	1938
## 1405	3	1	6	0	1998
## 1433	4	1	5	0	1928
## 1438	2	1	7	1	2008

## 1443	3	1	10	1	2008
##	GarageCars	GarageArea	WoodDeckSF	OpenPorchSF	EnclosedPorch
## 30	1	240	49	0	87
## 31	1	250	0	54	172
## 57	2	480	115	0	0
## 75	1	379	0	0	202
## 113	3	792	120	184	0
## 126	0	0	53	0	0
## 156	0	0	0	8	128
## 179	3	1166	0	60	0
## 232	3	810	0	146	202
## 265	2	504	0	0	0
## 314	2	513	0	0	0
## 324	1	220	142	98	0
## 336	2	529	670	0	0
## 350	3	706	0	0	0
## 376	0	0	0	0	0
## 384	2	360	0	0	91
## 399	1	338	0	0	0
## 406	2	490	120	78	0
## 411	1	350	0	0	0
## 432	0	0	0	98	0
## 441	3	672	0	72	0
## 450	1	308	0	176	0
## 458	2	529	224	137	0
## 459	1	228	192	63	0
## 463	1	360	0	0	236
## 474	3	908	250	63	0
## 496	0	0	0	523	115
## 497	2	546	264	75	291
## 524	3	884	208	406	0
## 528	3	858	126	66	0
## 534	0	0	0	0	0
## 535	2	403	100	35	0
## 584	2	880	105	502	0
## 599	2	621	0	0	0
## 608	1	230	103	0	0
## 633	2	544	192	0	0
## 662	3	711	517	76	0
## 665	3	1134	192	267	0
## 689	2	567	140	0	0
## 692	3	832	382	50	0
## 706	0	0	0	0	0
## 711	0	0	0	0	0
## 770	3	841	503	36	0
## 775	3	895	315	45	0
## 799	3	820	144	78	0
## 804	3	1020	52	170	0
## 808	2	528	0	312	0
## 813	2	504	0	0	0
## 875	1	200	26	0	0
## 899	3	820	0	67	0
## 917	1	308	0	0	0
## 943	0	0	321	0	0

## 964	2	702	288	136	0		
## 969	0	0	0	0	0		
## 992	2	424	0	169	0		
## 1021	2	451	252	64	0		
## 1046	2	452	0	0	0		
## 1047	3	716	214	108	0		
## 1098	2	461	0	74	137		
## 1143	3	1052	125	144	0		
## 1149	1	288	0	0	176		
## 1156	2	528	0	21	0		
## 1170	3	807	361	76	0		
## 1182	2	482	162	53	0		
## 1183	3	813	171	78	0		
## 1191	4	1356	439	0	0		
## 1192	2	492	206	0	0		
## 1200	2	459	0	82	0		
## 1212	2	492	325	12	0		
## 1263	1	304	120	0	0		
## 1280	2	528	30	0	164		
## 1299	2	1418	214	292	0		
## 1325	3	895	0	49	0		
## 1326	0	0	0	0	100		
## 1338	0	0	0	20	0		
## 1350	0	0	0	54	20		
## 1354	3	666	283	86	0		
## 1379	1	264	72	0	0		
## 1381	1	200	0	0	96		
## 1405	3	936	216	0	160		
## 1433	1	216	0	0	0		
## 1438	3	774	0	66	0		
## 1443	3	812	168	52	0		
##	X3SsnPorch.1	ScreenPorch	PoolArea	MiscVal	MoSold	YrSold	SalePriceLog
## 30	0	0	0	0	5	2008	11.13459
## 31	0	0	0	0	7	2008	10.59663
## 57	0	0	0	0	8	2009	12.05815
## 75	0	0	0	0	5	2010	11.58432
## 113	0	168	0	0	10	2007	12.85832
## 126	0	0	0	0	6	2006	11.34451
## 156	0	0	0	0	4	2008	11.27720
## 179	0	0	0	0	7	2009	13.12603
## 232	0	0	0	0	7	2009	12.90669
## 265	0	0	0	0	6	2008	11.19821
## 314	0	0	0	0	6	2009	12.83468
## 324	0	0	0	0	7	2006	11.74543
## 336	0	0	0	700	8	2008	12.34126
## 350	0	0	0	0	4	2006	12.98804
## 376	0	0	0	0	3	2009	11.01863
## 384	0	0	0	0	10	2009	11.23849
## 399	0	0	0	0	11	2007	11.11245
## 406	0	0	0	0	6	2009	11.91839
## 411	0	0	0	0	6	2009	11.00210
## 432	0	0	0	0	9	2008	11.28853
## 441	0	170	0	0	4	2009	13.22672
## 450	0	0	0	0	6	2007	11.69525

## 458	0	0	0	0	3	2008	12.45293
## 459	0	0	0	0	6	2008	11.98916
## 463	0	0	0	0	12	2009	11.04105
## 474	0	0	0	0	7	2007	12.99453
## 496	0	0	0	0	11	2009	10.46024
## 497	0	0	0	0	5	2007	12.97154
## 524	0	0	0	0	10	2007	12.12676
## 528	0	0	0	0	11	2008	13.00866
## 534	0	0	0	0	1	2007	10.57898
## 535	0	0	0	0	10	2006	12.08954
## 584	0	0	0	0	7	2008	12.69158
## 599	0	0	0	0	3	2006	12.28995
## 608	0	0	0	0	8	2006	12.32386
## 633	0	0	0	0	4	2009	11.32055
## 662	0	0	0	0	7	2009	12.90421
## 665	0	0	0	0	1	2006	12.95513
## 689	0	0	0	0	10	2007	12.87902
## 692	0	0	0	0	1	2007	13.53447
## 706	0	0	0	3500	7	2010	10.91509
## 711	0	0	0	0	7	2008	10.85900
## 770	0	210	0	0	6	2010	13.19561
## 775	0	0	0	0	7	2007	12.88664
## 799	0	0	0	0	7	2009	13.09190
## 804	0	192	0	0	1	2009	13.27583
## 808	0	0	0	0	5	2009	12.31717
## 813	0	0	0	54	6	2010	10.93298
## 875	0	0	0	0	8	2009	11.10496
## 899	0	0	0	0	3	2010	13.32393
## 917	0	0	0	0	10	2006	10.47195
## 943	0	0	0	0	8	2007	11.91839
## 964	0	0	0	0	5	2009	12.38422
## 969	0	0	0	0	5	2009	10.54271
## 992	0	0	0	0	11	2009	12.03172
## 1021	0	0	0	0	6	2008	12.07824
## 1046	0	0	0	0	6	2009	11.84654
## 1047	0	0	0	0	7	2006	13.22957
## 1098	0	0	0	0	10	2007	12.04355
## 1143	0	0	0	0	4	2007	12.95954
## 1149	0	0	0	0	8	2008	11.66907
## 1156	0	180	0	0	7	2007	12.29225
## 1170	0	0	0	0	7	2006	13.34551
## 1182	153	0	0	0	11	2008	12.88029
## 1183	0	0	555	0	7	2007	13.52114
## 1191	0	0	0	0	3	2007	12.03172
## 1192	0	0	0	0	11	2007	12.06681
## 1200	0	0	0	0	10	2006	11.90497
## 1212	0	0	0	0	6	2010	12.00762
## 1263	0	0	0	0	11	2009	11.99226
## 1280	0	0	0	0	4	2010	11.13313
## 1299	0	0	480	0	1	2008	11.98293
## 1325	0	0	0	0	2	2007	11.89819
## 1326	0	0	0	0	1	2008	10.91509
## 1338	0	0	0	0	3	2006	10.86857
## 1350	0	0	0	0	12	2008	11.71178

## 1354	0	0	0	0	3	2010	12.92391
## 1379	0	0	0	0	6	2006	11.32660
## 1381	0	0	0	0	6	2010	10.97678
## 1405	0	0	0	0	1	2006	11.56172
## 1433	0	0	0	0	8	2007	11.07442
## 1438	304	0	0	0	11	2008	12.88567
## 1443	0	0	0	0	4	2009	12.64433
##	secondFloor	X1stFlrSF	X2ndFlrSF	X3SsnPorch	.fitted	.se.fit	
## 30	FALSE	520	0	0	11.52855	0.010959165	
## 31	TRUE	649	668	0	11.52855	0.010959165	
## 57	TRUE	983	756	0	12.47267	0.010246994	
## 75	TRUE	840	765	0	11.29253	0.014795945	
## 113	TRUE	1282	1414	0	12.23664	0.007193874	
## 126	FALSE	520	0	0	12.00061	0.006043113	
## 156	TRUE	572	524	0	12.00061	0.006043113	
## 179	FALSE	2234	0	0	12.70869	0.014009510	
## 232	TRUE	1490	1304	0	12.47267	0.010246994	
## 265	FALSE	764	0	0	11.76458	0.007700847	
## 314	FALSE	2036	0	0	12.23664	0.007193874	
## 324	FALSE	1163	0	0	11.29253	0.014795945	
## 336	TRUE	1619	167	0	11.76458	0.007700847	
## 350	TRUE	1848	880	0	12.70869	0.014009510	
## 376	FALSE	904	0	0	10.82047	0.023034815	
## 384	FALSE	784	0	0	12.00061	0.006043113	
## 399	FALSE	1077	0	0	11.76458	0.007700847	
## 406	FALSE	1620	0	0	11.52855	0.010959165	
## 411	FALSE	1276	0	0	11.76458	0.007700847	
## 432	TRUE	1088	110	0	12.00061	0.006043113	
## 441	FALSE	2402	0	0	12.94472	0.018042957	
## 450	TRUE	679	504	0	11.29253	0.014795945	
## 458	FALSE	1663	0	0	11.52855	0.010959165	
## 459	TRUE	833	833	0	12.47267	0.010246994	
## 463	FALSE	864	0	0	11.76458	0.007700847	
## 474	FALSE	1976	0	0	12.47267	0.010246994	
## 496	FALSE	720	0	0	11.52855	0.010959165	
## 497	FALSE	3228	0	0	12.47267	0.010246994	
## 524	TRUE	3138	1538	0	12.94472	0.018042957	
## 528	TRUE	1476	1237	0	12.70869	0.014009510	
## 534	FALSE	334	0	0	10.82047	0.023034815	
## 535	TRUE	707	707	0	12.47267	0.010246994	
## 584	TRUE	1521	1254	0	12.94472	0.018042957	
## 599	FALSE	1647	0	0	11.76458	0.007700847	
## 608	TRUE	1112	896	0	11.76458	0.007700847	
## 633	FALSE	1411	0	0	12.23664	0.007193874	
## 662	TRUE	1686	762	0	12.47267	0.010246994	
## 665	FALSE	2097	0	0	12.47267	0.010246994	
## 689	FALSE	1419	0	0	12.47267	0.010246994	
## 692	TRUE	2444	1872	0	12.94472	0.018042957	
## 706	TRUE	372	720	0	11.52855	0.010959165	
## 711	FALSE	729	0	0	11.29253	0.014795945	
## 770	TRUE	1690	1589	0	12.47267	0.010246994	
## 775	FALSE	1973	0	0	12.47267	0.010246994	
## 799	TRUE	1966	1174	0	12.70869	0.014009510	
## 804	TRUE	1734	1088	0	12.70869	0.014009510	

## 808	TRUE	1072	504	0	11.76458	0.007700847
## 813	FALSE	1044	0	0	11.76458	0.007700847
## 875	TRUE	676	455	0	11.76458	0.007700847
## 899	FALSE	2364	0	0	12.70869	0.014009510
## 917	FALSE	480	0	0	11.05650	0.018861419
## 943	FALSE	1440	0	0	11.52855	0.010959165
## 964	FALSE	1800	0	0	12.70869	0.014009510
## 969	TRUE	600	368	0	11.29253	0.014795945
## 992	TRUE	916	826	0	12.47267	0.010246994
## 1021	FALSE	1132	0	0	11.52855	0.010959165
## 1046	FALSE	1733	0	0	11.29253	0.014795945
## 1047	TRUE	1992	876	0	12.70869	0.014009510
## 1098	FALSE	1088	0	0	12.47267	0.010246994
## 1143	TRUE	1466	1362	0	12.47267	0.010246994
## 1149	TRUE	572	539	0	12.23664	0.007193874
## 1156	FALSE	1437	0	0	11.76458	0.007700847
## 1170	TRUE	1831	1796	0	12.94472	0.018042957
## 1182	FALSE	1652	0	153	12.47267	0.010246994
## 1183	TRUE	2411	2065	0	12.94472	0.018042957
## 1191	FALSE	1622	0	0	11.52855	0.010959165
## 1192	TRUE	764	677	0	12.47267	0.010246994
## 1200	FALSE	1382	0	0	11.52855	0.010959165
## 1212	TRUE	1080	672	0	12.47267	0.010246994
## 1263	TRUE	1104	684	0	11.52855	0.010959165
## 1280	TRUE	698	430	0	11.52855	0.010959165
## 1299	TRUE	4692	950	0	12.94472	0.018042957
## 1325	FALSE	1795	0	0	12.47267	0.010246994
## 1326	FALSE	796	0	0	11.52855	0.010959165
## 1338	FALSE	693	0	0	11.52855	0.010959165
## 1350	TRUE	938	1215	0	12.47267	0.010246994
## 1354	TRUE	2053	1185	0	12.47267	0.010246994
## 1379	TRUE	483	504	0	12.00061	0.006043113
## 1381	FALSE	864	0	0	11.29253	0.014795945
## 1405	TRUE	694	520	0	11.29253	0.014795945
## 1433	FALSE	968	0	0	11.52855	0.010959165
## 1438	FALSE	1932	0	304	12.47267	0.010246994
## 1443	TRUE	1026	981	0	12.94472	0.018042957
##	.resid	.hat	.sigma	.cooksd	.std.resid	
## 30	-0.3939648	0.0022642065	0.2301607	0.003327599	-1.7124987	
## 31	-0.9319191	0.0022642065	0.2290924	0.018619692	-4.0508952	
## 57	-0.4145132	0.0019794933	0.2301360	0.003218717	-1.8015617	
## 75	0.2917895	0.0041271146	0.2302652	0.003339714	1.2695457	
## 113	0.6216820	0.0009756337	0.2298156	0.003561249	2.7006033	
## 126	-0.6561030	0.0006884660	0.2297500	0.002797408	-2.8497194	
## 156	-0.7234066	0.0006884660	0.2296112	0.003400767	-3.1420464	
## 179	0.4173370	0.0037000448	0.2301320	0.006119722	1.8154003	
## 232	0.4340262	0.0019794933	0.2301112	0.003528889	1.8863694	
## 265	-0.5663671	0.0011179903	0.2299137	0.003387948	-2.4604892	
## 314	0.5980436	0.0009756337	0.2298587	0.003295577	2.5979175	
## 324	0.4528992	0.0041271146	0.2300856	0.008045868	1.9705168	
## 336	0.5766771	0.0011179903	0.2298961	0.003512418	2.5052794	
## 350	0.2793472	0.0037000448	0.2302759	0.002741868	1.2151498	
## 376	0.1981591	0.0100030067	0.2303335	0.003777653	0.8647244	
## 384	-0.7621211	0.0006884660	0.2295252	0.003774504	-3.3101991	

## 399	-0.6521339	0.0011179903	0.2297575	0.004491737	-2.8330892
## 406	0.3898367	0.0022642065	0.2301656	0.003258228	1.6945544
## 411	-0.7624820	0.0011179903	0.2295240	0.006140446	-3.3124783
## 432	-0.7120786	0.0006884660	0.2296355	0.003295094	-3.0928443
## 441	0.2820018	0.0061372882	0.2302733	0.004657562	1.2282005
## 450	0.4027211	0.0041271146	0.2301498	0.006361780	1.7521972
## 458	0.9243789	0.0022642065	0.2291134	0.018319603	4.0181189
## 459	-0.4835060	0.0019794933	0.2300434	0.004379353	-2.1014193
## 463	-0.7235337	0.0011179903	0.2296106	0.005529150	-3.1432740
## 474	0.5218643	0.0019794933	0.2299857	0.005101777	2.2681326
## 496	-1.0683118	0.0022642065	0.2286824	0.024468765	-4.6437709
## 497	0.4988748	0.0019794933	0.2300208	0.004662184	2.1682153
## 524	-0.8179627	0.0061372882	0.2293876	0.039185183	-3.5624672
## 528	0.2999656	0.0037000448	0.2302580	0.003161556	1.3048392
## 534	-0.2414902	0.0100030067	0.2303048	0.005610389	-1.0538120
## 535	-0.3831268	0.0019794933	0.2301734	0.002749738	-1.6651502
## 584	-0.2531411	0.0061372882	0.2302965	0.003753014	-1.1025037
## 599	0.5253723	0.0011179903	0.2299806	0.002915245	2.2823941
## 608	0.5592739	0.0011179903	0.2299257	0.003303618	2.4296739
## 633	-0.9160841	0.0009756337	0.2291379	0.007732795	-3.9794942
## 662	0.4315417	0.0019794933	0.2301144	0.003488605	1.8755713
## 665	0.4824618	0.0019794933	0.2300449	0.004360457	2.0968808
## 689	0.4063514	0.0019794933	0.2301460	0.003093213	1.7660892
## 692	0.5897515	0.0061372882	0.2298707	0.020370086	2.5685402
## 706	-0.6134654	0.0022642065	0.2298300	0.008068563	-2.6666305
## 711	-0.4335269	0.0041271146	0.2301113	0.007372282	-1.8862301
## 770	0.7229482	0.0019794933	0.2296112	0.009790853	3.1420854
## 775	0.4139754	0.0019794933	0.2301366	0.003210371	1.7992244
## 799	0.3832105	0.0037000448	0.2301729	0.005159799	1.6669514
## 804	0.5671339	0.0037000448	0.2299112	0.011301318	2.4670112
## 808	0.5525849	0.0011179903	0.2299368	0.003225067	2.4006146
## 813	-0.8315999	0.0011179903	0.2293590	0.007304148	-3.6127496
## 875	-0.6596246	0.0011179903	0.2297428	0.004595518	-2.8656312
## 899	0.6152333	0.0037000448	0.2298260	0.013299569	2.6762418
## 917	-0.5845482	0.0067067147	0.2298796	0.021894071	-2.5466080
## 943	0.3898367	0.0022642065	0.2301656	0.003258228	1.6945544
## 964	-0.3244748	0.0037000448	0.2302351	0.003699302	-1.4114531
## 969	-0.7498195	0.0041271146	0.2295501	0.022053794	-3.2623860
## 992	-0.4409464	0.0019794933	0.2301022	0.003642318	-1.9164462
## 1021	0.5496854	0.0022642065	0.2299410	0.006478052	2.3893897
## 1046	0.5540106	0.0041271146	0.2299330	0.012039429	2.4104417
## 1047	0.5208744	0.0037000448	0.2299866	0.009532874	2.2657839
## 1098	-0.4291120	0.0019794933	0.2301176	0.003449431	-1.8650112
## 1143	0.4868728	0.0019794933	0.2300385	0.004440556	2.1160522
## 1149	-0.5675636	0.0009756337	0.2299118	0.002968211	-2.4655115
## 1156	0.5276685	0.0011179903	0.2299770	0.002940784	2.2923696
## 1170	0.4007854	0.0061372882	0.2301517	0.009407593	1.7455375
## 1182	0.4076261	0.0019794933	0.2301444	0.003112650	1.7716293
## 1183	0.5764179	0.0061372882	0.2298941	0.019459415	2.5104688
## 1191	0.5031654	0.0022642065	0.2300143	0.005427972	2.1871750
## 1192	-0.4058551	0.0019794933	0.2301466	0.003085661	-1.7639319
## 1200	0.3764137	0.0022642065	0.2301809	0.003037714	1.6362068
## 1212	-0.4650440	0.0019794933	0.2300696	0.004051298	-2.0211793
## 1263	0.4637066	0.0022642065	0.2300713	0.004610017	2.0156541

```

## 1280 -0.3954258 0.0022642065 0.2301590 0.003352324 -1.7188490
## 1299 -0.9617925 0.0061372882 0.2290020 0.054177329 -4.1888878
## 1325 -0.5744778 0.0019794933 0.2298995 0.006182338 -2.4968018
## 1326 -0.6134654 0.0022642065 0.2298300 0.008068563 -2.6666305
## 1338 -0.6599854 0.0022642065 0.2297414 0.009338663 -2.8688452
## 1350 -0.7608893 0.0019794933 0.2295269 0.010845491 -3.3069858
## 1354 0.4512468 0.0019794933 0.2300885 0.003814472 1.9612137
## 1379 -0.6740139 0.0006884660 0.2297144 0.002952225 -2.9275138
## 1381 -0.3157439 0.0041271146 0.2302434 0.003910567 -1.3737685
## 1405 0.2691897 0.0041271146 0.2302841 0.002842410 1.1712162
## 1433 -0.4541334 0.0022642065 0.2300845 0.004421635 -1.9740411
## 1438 0.4130053 0.0019794933 0.2301378 0.003195343 1.7950081
## 1443 -0.3003940 0.0061372882 0.2302573 0.005284907 -1.3083038
##      model_1_resid influential
## 30      -0.3939648          TRUE
## 31      -0.9319191          TRUE
## 57      -0.4145132          TRUE
## 75       0.2917895          TRUE
## 113     0.6216820          TRUE
## 126     -0.6561030          TRUE
## 156     -0.7234066          TRUE
## 179     0.4173370          TRUE
## 232     0.4340262          TRUE
## 265     -0.5663671          TRUE
## 314     0.5980436          TRUE
## 324     0.4528992          TRUE
## 336     0.5766771          TRUE
## 350     0.2793472          TRUE
## 376     0.1981591          TRUE
## 384     -0.7621211          TRUE
## 399     -0.6521339          TRUE
## 406     0.3898367          TRUE
## 411     -0.7624820          TRUE
## 432     -0.7120786          TRUE
## 441     0.2820018          TRUE
## 450     0.4027211          TRUE
## 458     0.9243789          TRUE
## 459     -0.4835060          TRUE
## 463     -0.7235337          TRUE
## 474     0.5218643          TRUE
## 496     -1.0683118          TRUE
## 497     0.4988748          TRUE
## 524     -0.8179627          TRUE
## 528     0.2999656          TRUE
## 534     -0.2414902          TRUE
## 535     -0.3831268          TRUE
## 584     -0.2531411          TRUE
## 599     0.5253723          TRUE
## 608     0.5592739          TRUE
## 633     -0.9160841          TRUE
## 662     0.4315417          TRUE
## 665     0.4824618          TRUE
## 689     0.4063514          TRUE
## 692     0.5897515          TRUE

```

```
## 706      -0.6134654      TRUE
## 711      -0.4335269      TRUE
## 770       0.7229482      TRUE
## 775       0.4139754      TRUE
## 799       0.3832105      TRUE
## 804       0.5671339      TRUE
## 808       0.5525849      TRUE
## 813      -0.8315999      TRUE
## 875      -0.6596246      TRUE
## 899       0.6152333      TRUE
## 917      -0.5845482      TRUE
## 943       0.3898367      TRUE
## 964      -0.3244748      TRUE
## 969      -0.7498195      TRUE
## 992      -0.4409464      TRUE
## 1021      0.5496854      TRUE
## 1046      0.5540106      TRUE
## 1047      0.5208744      TRUE
## 1098      -0.4291120      TRUE
## 1143      0.4868728      TRUE
## 1149      -0.5675636      TRUE
## 1156      0.5276685      TRUE
## 1170      0.4007854      TRUE
## 1182      0.4076261      TRUE
## 1183      0.5764179      TRUE
## 1191      0.5031654      TRUE
## 1192      -0.4058551      TRUE
## 1200      0.3764137      TRUE
## 1212      -0.4650440      TRUE
## 1263      0.4637066      TRUE
## 1280      -0.3954258      TRUE
## 1299      -0.9617925      TRUE
## 1325      -0.5744778      TRUE
## 1326      -0.6134654      TRUE
## 1338      -0.6599854      TRUE
## 1350      -0.7608893      TRUE
## 1354      0.4512468      TRUE
## 1379      -0.6740139      TRUE
## 1381      -0.3157439      TRUE
## 1405      0.2691897      TRUE
## 1433      -0.4541334      TRUE
## 1438      0.4130053      TRUE
## 1443      -0.3003940      TRUE
```

```
model_1_parameters <- model_1 %>% tidy
print(model_1_parameters)
```

```
##           term estimate std.error statistic p.value
## 1 (Intercept) 10.584442 0.027266623 388.18310      0
## 2 OverallQual  0.236028 0.004359842  54.13682      0
```

The R^2 is very good.

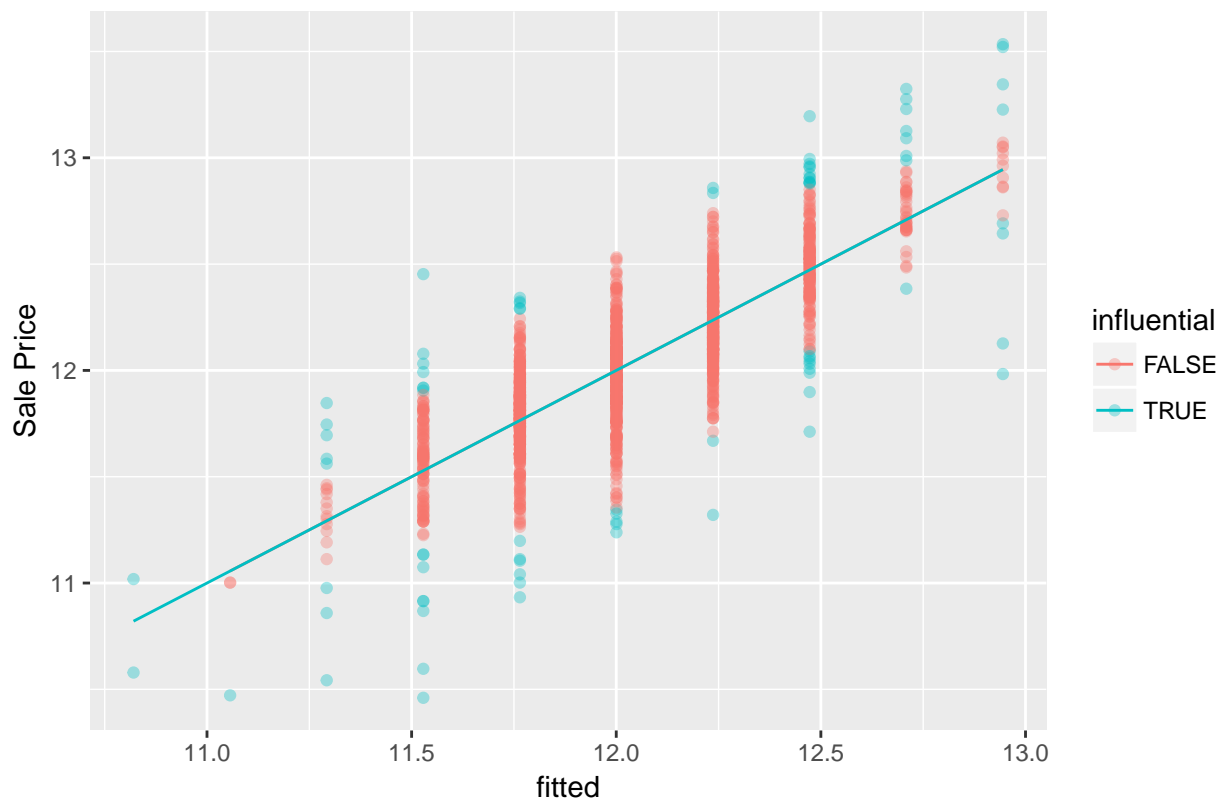
```
model_1_fit <- model_1 %>% glance
print(model_1_fit)
```



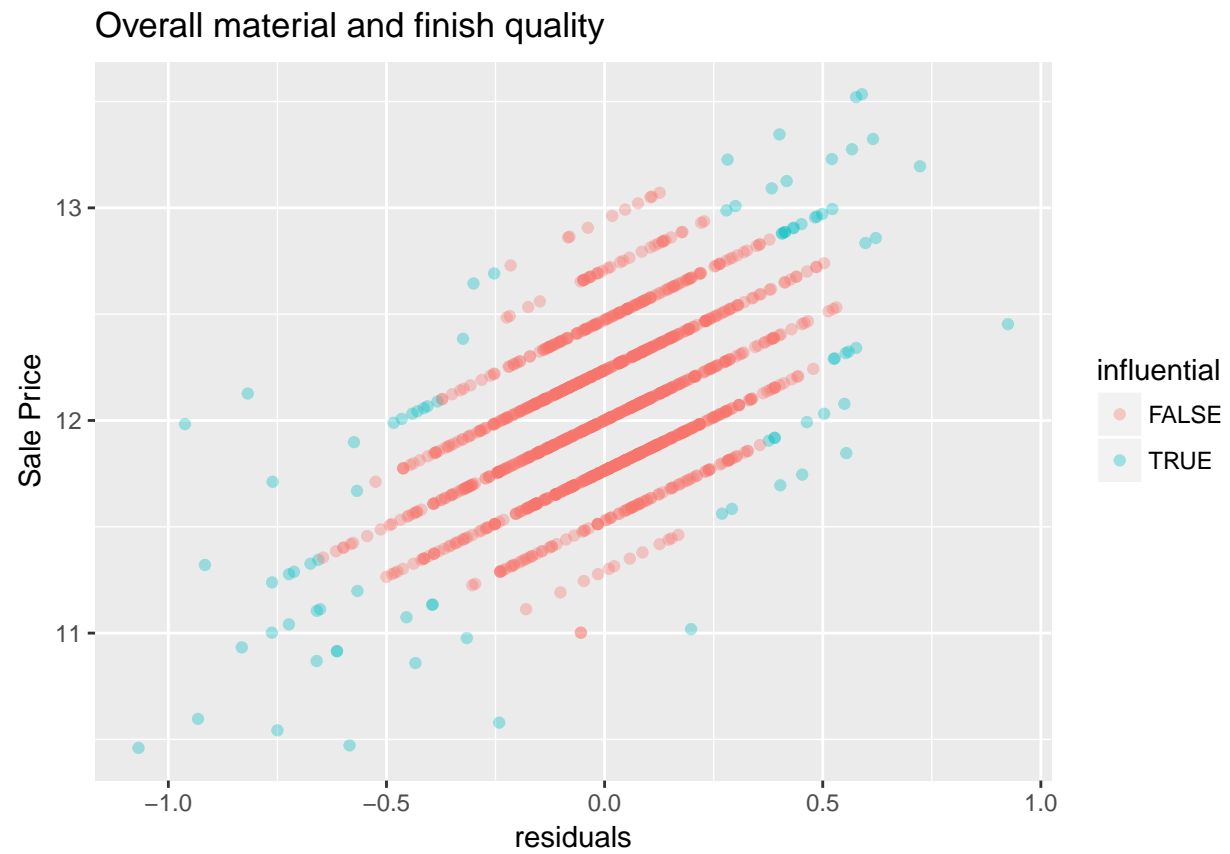
```
##      r.squared adj.r.squared      sigma statistic p.value df    logLik      AIC
## 1 0.6677904      0.6675625 0.2303135  2930.795      0 2 73.08848 -140.177
##      BIC deviance df.residual
## 1 -124.3184 77.33862      1458
```

```
t_numeric %>% ggplot() +
  aes(.fitted,SalePriceLog,color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
  geom_line(aes(.fitted,.fitted)) +
  ggtitle(selected_label) +
  xlab('fitted') +
  ylab("Sale Price")
```

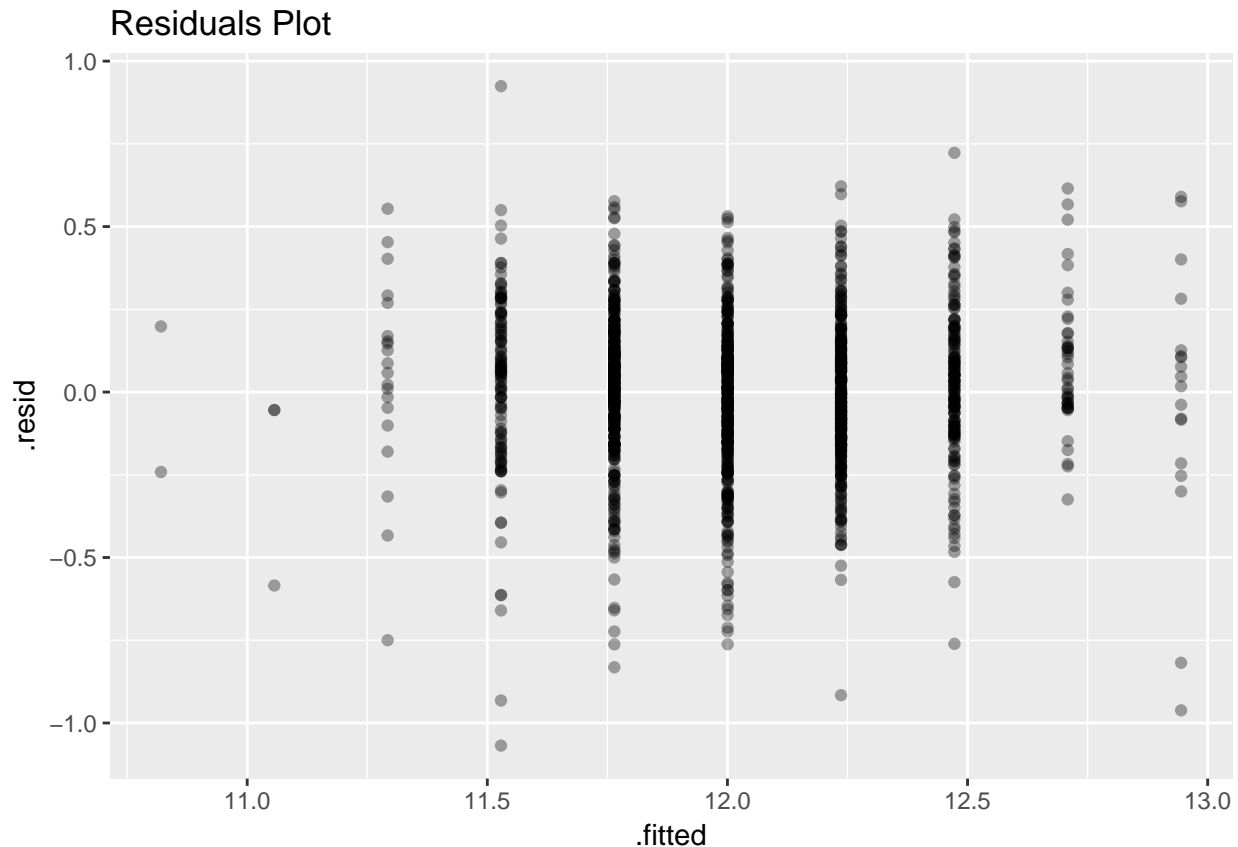
Overall material and finish quality



```
t_numeric %>% ggplot() +
  aes(.resid,SalePriceLog,color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
  ggtitle(selected_label) +
  xlab('residuals') +
  ylab("Sale Price")
```



```
t_numeric %>% ggplot() +
  aes(.fitted,.resid)+
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha=0.35) +
  ggtitle(paste("Residuals Plot"))
```



```
test <- read_data('test') %>%
  mutate(secondFloor = `2ndFlrSF` > 0) %>%
  mutate(X1stFlrSF = `1stFlrSF`) %>%
  mutate(X2ndFlrSF = `2ndFlrSF`) %>%
  mutate(X3SsnPorch = `3SsnPorch`) %>%
  select(`1stFlrSF`, `2ndFlrSF`, `3SsnPorch` )
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   Id = col_integer(),
##   MSSubClass = col_integer(),
##   LotFrontage = col_integer(),
##   LotArea = col_integer(),
##   OverallQual = col_integer(),
##   OverallCond = col_integer(),
##   YearBuilt = col_integer(),
##   YearRemodAdd = col_integer(),
##   MasVnrArea = col_integer(),
##   BsmtFinSF1 = col_integer(),
##   BsmtFinSF2 = col_integer(),
##   BsmtUnfSF = col_integer(),
##   TotalBsmtSF = col_integer(),
##   `1stFlrSF` = col_integer(),
##   `2ndFlrSF` = col_integer(),
##   LowQualFinSF = col_integer(),
##   GrLivArea = col_integer(),
```

```
## BsmtFullBath = col_integer(),
## BsmtHalfBath = col_integer(),
## FullBath = col_integer()
## # ... with 17 more columns
## )
```

See spec(...) for full column specifications.

```
p <- test %>%
  add_predictions(model_1) %>%
  select(Id,pred) %>%
  mutate(SalePrice = exp(pred)) %>%
  select(Id,SalePrice)
head(p)
```

```
## # A tibble: 6 x 2
##       Id SalePrice
##   <int>     <dbl>
## 1  1461  128615.4
## 2  1462  162854.1
## 3  1463  128615.4
## 4  1464  162854.1
## 5  1465  261101.8
## 6  1466  162854.1
```

```
write_csv(p,file.path('.', 'data', 'model_1.csv'))
```

Model 2

Corrrelation

So now we search for factors that are significantly correlated and order those by strength of correlation.

```
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" ".fitted"
## [41] ".se.fit" ".resid" ".hat" ".sigma"
## [45] ".cooksd" ".std.resid" "model_1_resid" "influential"
```

```
try(t_numeric <- t_numeric %>% select(-.hat,-.sigma,-.cooksd,-.std.resid,-.se.fit,-.fitted,-.resid,-inf)
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
```

```
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_1_resid"
```

```
model_num = 2
num_of_canadate_features = 15
sig <- 0.0010
response <- "model_1_resid"
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_1_resid"
```

```
# Calculate the correlation matrix
res2 <- rcorr(as.matrix(t_numeric))
res2$P
```

```
##          MSSubClass LotFrontage LotArea OverallQual
## MSSubClass      NA 0.000000e+00 8.194115e-08 2.127765e-01
## LotFrontage 0.000000e+00      NA 0.000000e+00 0.000000e+00
## LotArea      8.194115e-08 0.000000e+00      NA 5.105510e-05
## OverallQual 2.127765e-01 0.000000e+00 5.105510e-05      NA
## OverallCond 2.342026e-02 4.019570e-02 8.296279e-01 4.361720e-04
## YearBuilt    2.875793e-01 1.813803e-05 5.869923e-01 0.000000e+00
## YearRemodAdd 1.211632e-01 2.052218e-03 5.985903e-01 0.000000e+00
## MasVnrArea    3.824712e-01 1.536393e-11 6.995921e-05 0.000000e+00
## BsmtFinSF1    7.598818e-03 2.220446e-16 0.000000e+00 0.000000e+00
## BsmtFinSF2    1.210779e-02 8.388578e-02 2.068605e-05 2.388568e-02
## BsmtUnfSF     6.637585e-08 3.979765e-06 9.203746e-01 0.000000e+00
## TotalBsmtSF   0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## LowQualFinSF  7.586520e-02 1.827783e-01 8.552305e-01 2.452458e-01
## GrLivArea     4.213733e-03 0.000000e+00 0.000000e+00 0.000000e+00
## BsmtFullBath  8.939735e-01 4.587932e-04 1.230799e-09 2.094408e-05
## BsmtHalfBath  9.290423e-01 8.022392e-01 6.646032e-02 1.251676e-01
## FullBath      4.502116e-07 3.634426e-12 1.360046e-06 0.000000e+00
## HalfBath      8.788081e-12 6.365788e-02 5.861562e-01 0.000000e+00
## BedroomAbvGr  3.708289e-01 0.000000e+00 4.519868e-06 9.949905e-05
## KitchenAbvGr  0.000000e+00 8.335889e-01 4.971427e-01 1.435740e-12
## TotRmsAbvGrd  1.230183e-01 0.000000e+00 2.460254e-13 0.000000e+00
## Fireplaces    8.175109e-02 0.000000e+00 0.000000e+00 0.000000e+00
## GarageYrBlt   1.566879e-03 1.834203e-02 3.545897e-01 0.000000e+00
## GarageCars    1.255479e-01 0.000000e+00 2.706817e-09 0.000000e+00
## GarageArea    1.592243e-04 0.000000e+00 3.802736e-12 0.000000e+00
## WoodDeckSF    6.310397e-01 2.136523e-03 4.001000e-11 0.000000e+00
## OpenPorchSF   8.158485e-01 1.210208e-07 1.185832e-03 0.000000e+00
## EnclosedPorch 6.458454e-01 7.110477e-01 4.837906e-01 1.276996e-05
```

##	X3SsnPorch.1	9.414953e-02	1.520968e-02	4.355274e-01	2.461582e-01
##	ScreenPorch	3.202578e-01	1.517836e-01	9.924777e-02	1.314582e-02
##	PoolArea	7.518384e-01	5.393463e-13	2.979905e-03	1.275642e-02
##	MiscVal	7.692691e-01	9.071905e-01	1.459891e-01	2.304123e-01
##	MoSold	6.040063e-01	6.982029e-01	9.633078e-01	6.790959e-03
##	YrSold	4.137257e-01	7.964813e-01	5.861053e-01	2.963852e-01
##	SalePriceLog	4.692451e-03	0.000000e+00	0.000000e+00	0.000000e+00
##	secondFloor	0.000000e+00	1.136398e-01	6.543549e-01	1.513700e-11
##	X1stFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	X2ndFlrSF	0.000000e+00	5.433170e-03	5.144266e-02	0.000000e+00
##	X3SsnPorch	9.414953e-02	1.520968e-02	4.355274e-01	2.461582e-01
##	model_1_resid	1.861489e-11	0.000000e+00	0.000000e+00	1.000000e+00
##		OverallCond	YearBuilt	YearRemodAdd	MasVnrArea
##	MSSubClass	2.342026e-02	2.875793e-01	1.211632e-01	3.824712e-01
##	LotFrontage	4.019570e-02	1.813803e-05	2.052218e-03	1.536393e-11
##	LotArea	8.296279e-01	5.869923e-01	5.985903e-01	6.995921e-05
##	OverallQual	4.361720e-04	0.000000e+00	0.000000e+00	0.000000e+00
##	OverallCond	NA	0.000000e+00	4.816091e-03	9.716410e-07
##	YearBuilt	0.000000e+00	NA	0.000000e+00	0.000000e+00
##	YearRemodAdd	4.816091e-03	0.000000e+00	NA	5.400569e-12
##	MasVnrArea	9.716410e-07	0.000000e+00	5.400569e-12	NA
##	BsmtFinSF1	7.741099e-02	0.000000e+00	8.466438e-07	0.000000e+00
##	BsmtFinSF2	1.244256e-01	6.066982e-02	9.602934e-03	5.834234e-03
##	BsmtUnfSF	1.530091e-07	1.053188e-08	3.104628e-12	1.234288e-05
##	TotalBsmtSF	4.685274e-11	0.000000e+00	0.000000e+00	0.000000e+00
##	LowQualFinSF	3.303245e-01	1.475930e-12	1.706455e-02	8.467646e-03
##	GrLivArea	2.311138e-03	1.665335e-14	0.000000e+00	0.000000e+00
##	BsmtFullBath	3.580810e-02	4.964917e-13	4.707214e-06	1.138425e-03
##	BsmtHalfBath	6.367484e-06	1.449943e-01	6.376325e-01	3.097834e-01
##	FullBath	7.238654e-14	0.000000e+00	0.000000e+00	0.000000e+00
##	HalfBath	2.022485e-02	0.000000e+00	1.677547e-12	9.325873e-15
##	BedroomAbvGr	6.202024e-01	6.920572e-03	1.211642e-01	8.671369e-05
##	KitchenAbvGr	8.754659e-04	1.753175e-11	9.270313e-09	1.520249e-01
##	TotRmsAbvGrd	2.779336e-02	2.544832e-04	1.483258e-13	0.000000e+00
##	Fireplaces	3.630812e-01	1.423519e-08	1.619633e-05	0.000000e+00
##	GarageYrBlt	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	GarageCars	8.424372e-13	0.000000e+00	0.000000e+00	0.000000e+00
##	GarageArea	5.945338e-09	0.000000e+00	0.000000e+00	0.000000e+00
##	WoodDeckSF	8.987253e-01	0.000000e+00	1.998401e-15	9.347851e-10
##	OpenPorchSF	2.133227e-01	3.623768e-13	0.000000e+00	1.548562e-06
##	EnclosedPorch	7.159421e-03	0.000000e+00	7.771561e-14	2.570285e-05
##	X3SsnPorch.1	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
##	ScreenPorch	3.625216e-02	5.435519e-02	1.389957e-01	1.916276e-02
##	PoolArea	9.395944e-01	8.501179e-01	8.238849e-01	6.553522e-01
##	MiscVal	8.568156e-03	1.891692e-01	6.945331e-01	2.562107e-01
##	MoSold	8.933753e-01	6.359583e-01	4.119179e-01	8.203518e-01
##	YrSold	9.321243e-02	6.031268e-01	1.722487e-01	7.548601e-01
##	SalePriceLog	1.591324e-01	0.000000e+00	0.000000e+00	0.000000e+00
##	secondFloor	1.764981e-02	2.073907e-04	2.739984e-02	1.313690e-02
##	X1stFlrSF	3.126349e-08	0.000000e+00	0.000000e+00	0.000000e+00
##	X2ndFlrSF	2.690898e-01	6.939294e-01	7.777955e-08	2.121014e-11
##	X3SsnPorch	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
##	model_1_resid	1.118522e-02	1.776357e-15	1.021405e-14	3.433867e-10
##		BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF

## MSSubClass	7.598818e-03	1.210779e-02	6.637585e-08	0.000000e+00
## LotFrontage	2.220446e-16	8.388578e-02	3.979765e-06	0.000000e+00
## LotArea	0.000000e+00	2.068605e-05	9.203746e-01	0.000000e+00
## OverallQual	0.000000e+00	2.388568e-02	0.000000e+00	0.000000e+00
## OverallCond	7.741099e-02	1.244256e-01	1.530091e-07	4.685274e-11
## YearBuilt	0.000000e+00	6.066982e-02	1.053188e-08	0.000000e+00
## YearRemodAdd	8.466438e-07	9.602934e-03	3.104628e-12	0.000000e+00
## MasVnrArea	0.000000e+00	5.834234e-03	1.234288e-05	0.000000e+00
## BsmtFinSF1	NA	5.555041e-02	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.555041e-02	NA	6.661338e-16	6.010660e-05
## BsmtUnfSF	0.000000e+00	6.661338e-16	NA	0.000000e+00
## TotalBsmtSF	0.000000e+00	6.010660e-05	0.000000e+00	NA
## LowQualFinSF	1.369761e-02	5.718554e-01	2.821334e-01	2.042388e-01
## GrLivArea	8.881784e-16	7.128488e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	0.000000e+00	1.083816e-09	0.000000e+00	0.000000e+00
## BsmtHalfBath	9.972478e-03	6.687447e-03	2.463966e-04	9.904177e-01
## FullBath	2.529072e-02	3.470204e-03	0.000000e+00	0.000000e+00
## HalfBath	8.707324e-01	2.195850e-01	1.163181e-01	6.227931e-02
## BedroomAbvGr	3.949967e-05	5.481779e-01	1.485798e-10	5.394644e-02
## KitchenAbvGr	1.950234e-03	1.196093e-01	2.506158e-01	8.449363e-03
## TotRmsAbvGrd	9.051660e-02	1.785383e-01	0.000000e+00	0.000000e+00
## Fireplaces	0.000000e+00	7.308663e-02	4.880435e-02	0.000000e+00
## GarageYrBlt	1.013706e-08	1.069187e-03	9.254819e-13	0.000000e+00
## GarageCars	0.000000e+00	1.439250e-01	0.000000e+00	0.000000e+00
## GarageArea	0.000000e+00	4.864925e-01	1.690648e-12	0.000000e+00
## WoodDeckSF	3.108624e-15	9.454565e-03	8.391603e-01	0.000000e+00
## OpenPorchSF	1.867888e-05	9.060156e-01	7.585228e-07	0.000000e+00
## EnclosedPorch	9.005499e-05	1.628382e-01	9.228150e-01	2.587670e-04
## X3SsnPorch.1	3.125032e-01	2.520757e-01	4.278958e-01	1.533765e-01
## ScreenPorch	1.778536e-02	6.748805e-04	6.310420e-01	1.232140e-03
## PoolArea	7.033176e-08	1.111537e-01	1.802008e-01	1.354119e-06
## MiscVal	8.915451e-01	8.504155e-01	3.627455e-01	4.804779e-01
## MoSold	5.482075e-01	5.614175e-01	1.827454e-01	6.143911e-01
## YrSold	5.835463e-01	2.259961e-01	1.150733e-01	5.676653e-01
## SalePriceLog	0.000000e+00	8.536300e-01	0.000000e+00	0.000000e+00
## secondFloor	0.000000e+00	7.216583e-06	1.797090e-01	0.000000e+00
## X1stFlrSF	0.000000e+00	2.020320e-04	0.000000e+00	0.000000e+00
## X2ndFlrSF	1.455248e-07	1.453587e-04	8.645239e-01	1.894085e-11
## X3SsnPorch	3.125032e-01	2.520757e-01	4.278958e-01	1.533765e-01
## model_1_resid	0.000000e+00	4.194278e-04	4.796603e-02	0.000000e+00
##	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath
## MSSubClass	7.586520e-02	4.213733e-03	8.939735e-01	9.290423e-01
## LotFrontage	1.827783e-01	0.000000e+00	4.587932e-04	8.022392e-01
## LotArea	8.552305e-01	0.000000e+00	1.230799e-09	6.646032e-02
## OverallQual	2.452458e-01	0.000000e+00	2.094408e-05	1.251676e-01
## OverallCond	3.303245e-01	2.311138e-03	3.580810e-02	6.367484e-06
## YearBuilt	1.475930e-12	1.665335e-14	4.964917e-13	1.449943e-01
## YearRemodAdd	1.706455e-02	0.000000e+00	4.707214e-06	6.376325e-01
## MasVnrArea	8.467646e-03	0.000000e+00	1.138425e-03	3.097834e-01
## BsmtFinSF1	1.369761e-02	8.881784e-16	0.000000e+00	9.972478e-03
## BsmtFinSF2	5.718554e-01	7.128488e-01	1.083816e-09	6.687447e-03
## BsmtUnfSF	2.821334e-01	0.000000e+00	0.000000e+00	2.463966e-04
## TotalBsmtSF	2.042388e-01	0.000000e+00	0.000000e+00	9.904177e-01
## LowQualFinSF	NA	2.399685e-07	7.173345e-02	8.235244e-01

## GrLivArea	2.399685e-07	NA	1.834039e-01	4.700979e-01
## BsmtFullBath	7.173345e-02	1.834039e-01	NA	1.374508e-08
## BsmtHalfBath	8.235244e-01	4.700979e-01	1.374508e-08	NA
## FullBath	9.783903e-01	0.000000e+00	1.368378e-02	3.719872e-02
## HalfBath	3.011218e-01	0.000000e+00	2.379416e-01	6.375543e-01
## BedroomAbvGr	5.275460e-05	0.000000e+00	7.237188e-09	7.558104e-02
## KitchenAbvGr	7.739872e-01	1.282742e-04	1.129357e-01	1.473008e-01
## TotRmsAbvGrd	4.904244e-07	0.000000e+00	4.181609e-02	3.627516e-01
## Fireplaces	4.166751e-01	0.000000e+00	1.216473e-07	2.685363e-01
## GarageYrBlt	1.771551e-01	0.000000e+00	3.498871e-06	3.997996e-03
## GarageCars	3.002845e-04	0.000000e+00	4.259910e-07	4.250739e-01
## GarageArea	9.772127e-03	0.000000e+00	5.316858e-12	3.488402e-01
## WoodDeckSF	3.312869e-01	0.000000e+00	1.526557e-11	1.250635e-01
## OpenPorchSF	4.859080e-01	0.000000e+00	1.005792e-02	3.335712e-01
## EnclosedPorch	1.959094e-02	7.278971e-01	5.656746e-02	7.439517e-01
## X3SsnPorch.1	8.697349e-01	4.305893e-01	9.967684e-01	1.799353e-01
## ScreenPorch	3.061609e-01	1.021511e-04	3.767857e-01	2.199645e-01
## PoolArea	1.753504e-02	5.918421e-11	9.756789e-03	4.445328e-01
## MiscVal	8.848674e-01	9.265212e-01	3.788657e-01	7.785284e-01
## MoSold	3.971935e-01	5.495607e-02	3.328625e-01	2.093598e-01
## YrSold	2.694414e-01	1.630401e-01	1.038816e-02	7.554937e-02
## SalePriceLog	1.471041e-01	0.000000e+00	0.000000e+00	8.441568e-01
## secondFloor	2.385267e-01	0.000000e+00	0.000000e+00	3.410671e-01
## X1stFlrSF	5.866501e-01	0.000000e+00	0.000000e+00	9.404841e-01
## X2ndFlrSF	1.547482e-02	0.000000e+00	7.124057e-11	3.623802e-01
## X3SsnPorch	8.697349e-01	4.305893e-01	9.967684e-01	1.799353e-01
## model_1_resid	3.856256e-01	0.000000e+00	0.000000e+00	6.676885e-02
##	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr
## MSSubClass	4.502116e-07	8.788081e-12	3.708289e-01	0.000000e+00
## LotFrontage	3.634426e-12	6.365788e-02	0.000000e+00	8.335889e-01
## LotArea	1.360046e-06	5.861562e-01	4.519868e-06	4.971427e-01
## OverallQual	0.000000e+00	0.000000e+00	9.949905e-05	1.435740e-12
## OverallCond	7.238654e-14	2.022485e-02	6.202024e-01	8.754659e-04
## YearBuilt	0.000000e+00	0.000000e+00	6.920572e-03	1.753175e-11
## YearRemodAdd	0.000000e+00	1.677547e-12	1.211642e-01	9.270313e-09
## MasVnrArea	0.000000e+00	9.325873e-15	8.671369e-05	1.520249e-01
## BsmtFinSF1	2.529072e-02	8.707324e-01	3.949967e-05	1.950234e-03
## BsmtFinSF2	3.470204e-03	2.195850e-01	5.481779e-01	1.196093e-01
## BsmtUnfSF	0.000000e+00	1.163181e-01	1.485798e-10	2.506158e-01
## TotalBsmtSF	0.000000e+00	6.227931e-02	5.394644e-02	8.449363e-03
## LowQualFinSF	9.783903e-01	3.011218e-01	5.275460e-05	7.739872e-01
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	1.282742e-04
## BsmtFullBath	1.368378e-02	2.379416e-01	7.237188e-09	1.129357e-01
## BsmtHalfBath	3.719872e-02	6.375543e-01	7.558104e-02	1.473008e-01
## FullBath	NA	1.685132e-07	0.000000e+00	3.313170e-07
## HalfBath	1.685132e-07	NA	0.000000e+00	9.077592e-03
## BedroomAbvGr	0.000000e+00	0.000000e+00	NA	1.887379e-14
## KitchenAbvGr	3.313170e-07	9.077592e-03	1.887379e-14	NA
## TotRmsAbvGrd	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## Fireplaces	0.000000e+00	3.996803e-15	3.810705e-05	2.035535e-06
## GarageYrBlt	0.000000e+00	1.665335e-13	1.656563e-02	3.589658e-06
## GarageCars	0.000000e+00	0.000000e+00	9.897563e-04	5.307601e-02
## GarageArea	0.000000e+00	3.252643e-10	1.263773e-02	1.379977e-02
## WoodDeckSF	4.818368e-13	3.498384e-05	7.349737e-02	5.648744e-04

##	OpenPorchSF	0.000000e+00	1.332268e-14	3.315592e-04	7.380715e-03
##	EnclosedPorch	1.040475e-05	2.650862e-04	1.123474e-01	1.541632e-01
##	X3SsnPorch.1	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
##	ScreenPorch	7.569616e-01	5.628841e-03	9.063268e-02	4.863598e-02
##	PoolArea	5.810529e-02	3.927890e-01	6.879729e-03	5.791972e-01
##	MiscVal	5.853585e-01	9.607167e-01	7.668285e-01	1.720427e-02
##	MoSold	3.278470e-02	7.297135e-01	7.542379e-02	3.099771e-01
##	YrSold	4.526681e-01	6.950290e-01	1.690210e-01	2.262661e-01
##	SalePriceLog	0.000000e+00	0.000000e+00	6.661338e-16	1.478820e-08
##	secondFloor	0.000000e+00	0.000000e+00	0.000000e+00	1.129716e-01
##	X1stFlrSF	0.000000e+00	4.334880e-06	1.041033e-06	9.243536e-03
##	X2ndFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	2.344382e-02
##	X3SsnPorch	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
##	model_1_resid	0.000000e+00	1.609614e-09	0.000000e+00	8.571477e-01
##	TotRmsAbvGrd		Fireplaces	GarageYrBlt	GarageCars
##	MSSubClass	1.230183e-01	8.175109e-02	1.566879e-03	1.255479e-01
##	LotFrontage	0.000000e+00	0.000000e+00	1.834203e-02	0.000000e+00
##	LotArea	2.460254e-13	0.000000e+00	3.545897e-01	2.706817e-09
##	OverallQual	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	OverallCond	2.779336e-02	3.630812e-01	0.000000e+00	8.424372e-13
##	YearBuilt	2.544832e-04	1.423519e-08	0.000000e+00	0.000000e+00
##	YearRemodAdd	1.483258e-13	1.619633e-05	0.000000e+00	0.000000e+00
##	MasVnrArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	BsmtFinSF1	9.051660e-02	0.000000e+00	1.013706e-08	0.000000e+00
##	BsmtFinSF2	1.785383e-01	7.308663e-02	1.069187e-03	1.439250e-01
##	BsmtUnfSF	0.000000e+00	4.880435e-02	9.254819e-13	0.000000e+00
##	TotalBsmtSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	LowQualFinSF	4.904244e-07	4.166751e-01	1.771551e-01	3.002845e-04
##	GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	BsmtFullBath	4.181609e-02	1.216473e-07	3.498871e-06	4.259910e-07
##	BsmtHalfBath	3.627516e-01	2.685363e-01	3.997996e-03	4.250739e-01
##	FullBath	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	HalfBath	0.000000e+00	3.996803e-15	1.665335e-13	0.000000e+00
##	BedroomAbvGr	0.000000e+00	3.810705e-05	1.656563e-02	9.897563e-04
##	KitchenAbvGr	0.000000e+00	2.035535e-06	3.589658e-06	5.307601e-02
##	TotRmsAbvGrd		NA	0.000000e+00	0.000000e+00
##	Fireplaces	0.000000e+00		NA	8.219437e-02
##	GarageYrBlt	3.283118e-08	8.219437e-02		NA
##	GarageCars	0.000000e+00	0.000000e+00	0.000000e+00	
##	GarageArea	0.000000e+00	0.000000e+00	0.000000e+00	NA
##	WoodDeckSF	1.758029e-10	1.221245e-14	0.000000e+00	0.000000e+00
##	OpenPorchSF	0.000000e+00	7.290124e-11	0.000000e+00	0.000000e+00
##	EnclosedPorch	8.740741e-01	3.432435e-01	0.000000e+00	6.067020e-09
##	X3SsnPorch.1	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01
##	ScreenPorch	2.326437e-02	1.195044e-12	5.077546e-03	5.373793e-02
##	PoolArea	1.358848e-03	2.748854e-04	5.905625e-01	4.241329e-01
##	MiscVal	3.443919e-01	9.571131e-01	2.289596e-01	9.987840e-02
##	MoSold	1.586904e-01	7.660440e-02	8.430360e-01	1.217084e-01
##	YrSold	1.874597e-01	3.575543e-01	9.699819e-01	1.351896e-01
##	SalePriceLog	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	secondFloor	0.000000e+00	5.724116e-06	3.584860e-01	4.817098e-02
##	X1stFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
##	X2ndFlrSF	0.000000e+00	6.394885e-14	8.506609e-03	1.417977e-12
##	X3SsnPorch	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01

```

## model_1_resid 0.000000e+00 0.000000e+00 9.185973e-08 0.000000e+00
##          GarageArea  WoodDeckSF  OpenPorchSF  EnclosedPorch
## MSSubClass      1.592243e-04 6.310397e-01 8.158485e-01 6.458454e-01
## LotFrontage      0.000000e+00 2.136523e-03 1.210208e-07 7.110477e-01
## LotArea          3.802736e-12 4.001000e-11 1.185832e-03 4.837906e-01
## OverallQual      0.000000e+00 0.000000e+00 0.000000e+00 1.276996e-05
## OverallCond      5.945338e-09 8.987253e-01 2.133227e-01 7.159421e-03
## YearBuilt        0.000000e+00 0.000000e+00 3.623768e-13 0.000000e+00
## YearRemodAdd     0.000000e+00 1.998401e-15 0.000000e+00 7.771561e-14
## MasVnrArea       0.000000e+00 9.347851e-10 1.548562e-06 2.570285e-05
## BsmtFinSF1       0.000000e+00 3.108624e-15 1.867888e-05 9.005499e-05
## BsmtFinSF2       4.864925e-01 9.454565e-03 9.060156e-01 1.628382e-01
## BsmtUnfSF        1.690648e-12 8.391603e-01 7.585228e-07 9.228150e-01
## TotalBsmtSF      0.000000e+00 0.000000e+00 0.000000e+00 2.587670e-04
## LowQualFinSF     9.772127e-03 3.312869e-01 4.859080e-01 1.959094e-02
## GrLivArea        0.000000e+00 0.000000e+00 0.000000e+00 7.278971e-01
## BsmtFullBath     5.316858e-12 1.526557e-11 1.005792e-02 5.656746e-02
## BsmtHalfBath     3.488402e-01 1.250635e-01 3.335712e-01 7.439517e-01
## FullBath         0.000000e+00 4.818368e-13 0.000000e+00 1.040475e-05
## HalfBath         3.252643e-10 3.498384e-05 1.332268e-14 2.650862e-04
## BedroomAbvGr     1.263773e-02 7.349737e-02 3.315592e-04 1.123474e-01
## KitchenAbvGr     1.379977e-02 5.648744e-04 7.380715e-03 1.541632e-01
## TotRmsAbvGrd     0.000000e+00 1.758029e-10 0.000000e+00 8.740741e-01
## Fireplaces       0.000000e+00 1.221245e-14 7.290124e-11 3.432435e-01
## GarageYrBltd     0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## GarageCars       0.000000e+00 0.000000e+00 0.000000e+00 6.067020e-09
## GarageArea       NA 0.000000e+00 0.000000e+00 3.064131e-06
## WoodDeckSF       0.000000e+00 NA 2.499829e-02 1.371106e-06
## OpenPorchSF      0.000000e+00 2.499829e-02 NA 3.691031e-04
## EnclosedPorch    3.064131e-06 1.371106e-06 3.691031e-04 NA
## X3SsnPorch.1     1.802697e-01 2.107784e-01 8.234948e-01 1.542416e-01
## ScreenPorch      4.952347e-02 4.569023e-03 4.502233e-03 1.529761e-03
## PoolArea         1.965912e-02 5.029192e-03 2.023972e-02 3.837510e-02
## MiscVal          2.954461e-01 7.153744e-01 4.779914e-01 4.832932e-01
## MoSold           2.854434e-01 4.224189e-01 6.454024e-03 2.699986e-01
## YrSold           2.958337e-01 3.951420e-01 2.769527e-02 7.050056e-01
## SalePriceLog     0.000000e+00 0.000000e+00 0.000000e+00 1.050805e-08
## secondFloor      9.467301e-01 2.184093e-01 5.305810e-08 5.719454e-04
## X1stFlrSF        0.000000e+00 0.000000e+00 4.440892e-16 1.258441e-02
## X2ndFlrSF        1.112965e-07 4.216853e-04 8.881784e-16 1.784426e-02
## X3SsnPorch       1.802697e-01 2.107784e-01 8.234948e-01 1.542416e-01
## model_1_resid    0.000000e+00 0.000000e+00 4.967757e-06 2.038168e-04
##          X3SsnPorch.1  ScreenPorch  PoolArea  MiscVal
## MSSubClass      0.09414953 3.202578e-01 7.518384e-01 0.769269145
## LotFrontage      0.01520968 1.517836e-01 5.393463e-13 0.907190524
## LotArea          0.43552735 9.924777e-02 2.979905e-03 0.145989130
## OverallQual      0.24615822 1.314582e-02 1.275642e-02 0.230412263
## OverallCond      0.33014731 3.625216e-02 9.395944e-01 0.008568156
## YearBuilt        0.23118088 5.435519e-02 8.501179e-01 0.189169158
## YearRemodAdd     0.08367071 1.389957e-01 8.238849e-01 0.694533106
## MasVnrArea       0.47420644 1.916276e-02 6.553522e-01 0.256210660
## BsmtFinSF1       0.31250322 1.778536e-02 7.033176e-08 0.891545102
## BsmtFinSF2       0.25207571 6.748805e-04 1.111537e-01 0.850415539
## BsmtUnfSF        0.42789582 6.310420e-01 1.802008e-01 0.362745457

```

## TotalBsmtSF	0.15337655	1.232140e-03	1.354119e-06	0.480477906
## LowQualFinSF	0.86973491	3.061609e-01	1.753504e-02	0.884867376
## GrLivArea	0.43058934	1.021511e-04	5.918421e-11	0.926521232
## BsmtFullBath	0.99676835	3.767857e-01	9.756789e-03	0.378865709
## BsmtHalfBath	0.17993527	2.199645e-01	4.445328e-01	0.778528444
## FullBath	0.17698330	7.569616e-01	5.810529e-02	0.585358529
## HalfBath	0.84943689	5.628841e-03	3.927890e-01	0.960716747
## BedroomAbvGr	0.34997682	9.063268e-02	6.879729e-03	0.766828521
## KitchenAbvGr	0.34756881	4.863598e-02	5.791972e-01	0.017204270
## TotRmsAbvGrd	0.79860693	2.326437e-02	1.358848e-03	0.344391918
## Fireplaces	0.66735300	1.195044e-12	2.748854e-04	0.957113099
## GarageYrBlt	0.38232493	5.077546e-03	5.905625e-01	0.228959597
## GarageCars	0.17198411	5.373793e-02	4.241329e-01	0.099878400
## GarageArea	0.18026968	4.952347e-02	1.965912e-02	0.295446101
## WoodDeckSF	0.21077845	4.569023e-03	5.029192e-03	0.715374444
## OpenPorchSF	0.82349481	4.502233e-03	2.023972e-02	0.477991449
## EnclosedPorch	0.15424164	1.529761e-03	3.837510e-02	0.483293192
## X3SsnPorch.1	NA	2.299724e-01	7.602900e-01	0.989218187
## ScreenPorch	0.22997242	NA	4.998823e-02	0.222498354
## PoolArea	0.76029001	4.998823e-02	NA	0.257249528
## MiscVal	0.98921819	2.224984e-01	2.572495e-01	NA
## MoSold	0.26038906	3.753591e-01	1.976284e-01	0.804175819
## YrSold	0.47654325	6.830660e-01	2.256080e-02	0.851418697
## SalePriceLog	0.03594763	3.408972e-06	7.631787e-03	0.444619672
## secondFloor	0.33748531	5.241591e-01	1.622043e-01	0.840707205
## X1stFlrSF	0.03206515	6.856867e-04	4.578567e-07	0.420551089
## X2ndFlrSF	0.35234764	1.209302e-01	1.832441e-03	0.536316831
## X3SsnPorch	0.00000000	2.299724e-01	7.602900e-01	0.989218187
## model_1_resid	0.04616510	5.838197e-06	2.730217e-01	0.708524671
##	MoSold	YrSold	SalePriceLog	secondFloor
## MSSubClass	6.040063e-01	4.137257e-01	4.692451e-03	0.000000e+00
## LotFrontage	6.982029e-01	7.964813e-01	0.000000e+00	1.136398e-01
## LotArea	9.633078e-01	5.861053e-01	0.000000e+00	6.543549e-01
## OverallQual	6.790959e-03	2.963852e-01	0.000000e+00	1.513700e-11
## OverallCond	8.933753e-01	9.321243e-02	1.591324e-01	1.764981e-02
## YearBuilt	6.359583e-01	6.031268e-01	0.000000e+00	2.073907e-04
## YearRemodAdd	4.119179e-01	1.722487e-01	0.000000e+00	2.739984e-02
## MasVnrArea	8.203518e-01	7.548601e-01	0.000000e+00	1.313690e-02
## BsmtFinSF1	5.482075e-01	5.835463e-01	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.614175e-01	2.259961e-01	8.536300e-01	7.216583e-06
## BsmtUnfSF	1.827454e-01	1.150733e-01	0.000000e+00	1.797090e-01
## TotalBsmtSF	6.143911e-01	5.676653e-01	0.000000e+00	0.000000e+00
## LowQualFinSF	3.971935e-01	2.694414e-01	1.471041e-01	2.385267e-01
## GrLivArea	5.495607e-02	1.630401e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	3.328625e-01	1.038816e-02	0.000000e+00	0.000000e+00
## BsmtHalfBath	2.093598e-01	7.554937e-02	8.441568e-01	3.410671e-01
## FullBath	3.278470e-02	4.526681e-01	0.000000e+00	0.000000e+00
## HalfBath	7.297135e-01	6.950290e-01	0.000000e+00	0.000000e+00
## BedroomAbvGr	7.542379e-02	1.690210e-01	6.661338e-16	0.000000e+00
## KitchenAbvGr	3.099771e-01	2.262661e-01	1.478820e-08	1.129716e-01
## TotRmsAbvGrd	1.586904e-01	1.874597e-01	0.000000e+00	0.000000e+00
## Fireplaces	7.660440e-02	3.575543e-01	0.000000e+00	5.724116e-06
## GarageYrBlt	8.430360e-01	9.699819e-01	0.000000e+00	3.584860e-01
## GarageCars	1.217084e-01	1.351896e-01	0.000000e+00	4.817098e-02

## GarageArea	2.854434e-01	2.958337e-01	0.000000e+00	9.467301e-01
## WoodDeckSF	4.224189e-01	3.951420e-01	0.000000e+00	2.184093e-01
## OpenPorchSF	6.454024e-03	2.769527e-02	0.000000e+00	5.305810e-08
## EnclosedPorch	2.699986e-01	7.050056e-01	1.050805e-08	5.719454e-04
## X3SsnPorch.1	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
## ScreenPorch	3.753591e-01	6.830660e-01	3.408972e-06	5.241591e-01
## PoolArea	1.976284e-01	2.256080e-02	7.631787e-03	1.622043e-01
## MiscVal	8.041758e-01	8.514187e-01	4.446197e-01	8.407072e-01
## MoSold	NA	2.230277e-08	2.848930e-02	3.497548e-01
## YrSold	2.230277e-08	NA	1.547100e-01	6.406425e-01
## SalePriceLog	2.848930e-02	1.547100e-01	NA	7.446401e-09
## secondFloor	3.497548e-01	6.406425e-01	7.446401e-09	NA
## X1stFlrSF	2.309272e-01	6.034971e-01	0.000000e+00	0.000000e+00
## X2ndFlrSF	1.793059e-01	2.731188e-01	0.000000e+00	0.000000e+00
## X3SsnPorch	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
## model_1_resid	9.714902e-01	3.230896e-01	0.000000e+00	6.306428e-01
##	X1stFlrSF	X2ndFlrSF	X3SsnPorch	model_1_resid
## MSSubClass	0.000000e+00	0.000000e+00	0.09414953	1.861489e-11
## LotFrontage	0.000000e+00	5.433170e-03	0.01520968	0.000000e+00
## LotArea	0.000000e+00	5.144266e-02	0.43552735	0.000000e+00
## OverallQual	0.000000e+00	0.000000e+00	0.24615822	1.000000e+00
## OverallCond	3.126349e-08	2.690898e-01	0.33014731	1.118522e-02
## YearBuilt	0.000000e+00	6.939294e-01	0.23118088	1.776357e-15
## YearRemodAdd	0.000000e+00	7.777955e-08	0.08367071	1.021405e-14
## MasVnrArea	0.000000e+00	2.121014e-11	0.47420644	3.433867e-10
## BsmtFinSF1	0.000000e+00	1.455248e-07	0.31250322	0.000000e+00
## BsmtFinSF2	2.020320e-04	1.453587e-04	0.25207571	4.194278e-04
## BsmtUnfSF	0.000000e+00	8.645239e-01	0.42789582	4.796603e-02
## TotalBsmtSF	0.000000e+00	1.894085e-11	0.15337655	0.000000e+00
## LowQualFinSF	5.866501e-01	1.547482e-02	0.86973491	3.856256e-01
## GrLivArea	0.000000e+00	0.000000e+00	0.43058934	0.000000e+00
## BsmtFullBath	0.000000e+00	7.124057e-11	0.99676835	0.000000e+00
## BsmtHalfBath	9.404841e-01	3.623802e-01	0.17993527	6.676885e-02
## FullBath	0.000000e+00	0.000000e+00	0.17698330	0.000000e+00
## HalfBath	4.334880e-06	0.000000e+00	0.84943689	1.609614e-09
## BedroomAbvGr	1.041033e-06	0.000000e+00	0.34997682	0.000000e+00
## KitchenAbvGr	9.243536e-03	2.344382e-02	0.34756881	8.571477e-01
## TotRmsAbvGrd	0.000000e+00	0.000000e+00	0.79860693	0.000000e+00
## Fireplaces	0.000000e+00	6.394885e-14	0.66735300	0.000000e+00
## GarageYrBlt	0.000000e+00	8.506609e-03	0.38232493	9.185973e-08
## GarageCars	0.000000e+00	1.417977e-12	0.17198411	0.000000e+00
## GarageArea	0.000000e+00	1.112965e-07	0.18026968	0.000000e+00
## WoodDeckSF	0.000000e+00	4.216853e-04	0.21077845	0.000000e+00
## OpenPorchSF	4.440892e-16	8.881784e-16	0.82349481	4.967757e-06
## EnclosedPorch	1.258441e-02	1.784426e-02	0.15424164	2.038168e-04
## X3SsnPorch.1	3.206515e-02	3.523476e-01	0.00000000	4.616510e-02
## ScreenPorch	6.856867e-04	1.209302e-01	0.22997242	5.838197e-06
## PoolArea	4.578567e-07	1.832441e-03	0.76029001	2.730217e-01
## MiscVal	4.205511e-01	5.363168e-01	0.98921819	7.085247e-01
## MoSold	2.309272e-01	1.793059e-01	0.26038906	9.714902e-01
## YrSold	6.034971e-01	2.731188e-01	0.47654325	3.230896e-01
## SalePriceLog	0.000000e+00	0.000000e+00	0.03594763	0.000000e+00
## secondFloor	0.000000e+00	0.000000e+00	0.33748531	6.306428e-01
## X1stFlrSF	NA	5.329071e-15	0.03206515	0.000000e+00

```
## X2ndFlrSF      5.329071e-15      NA 0.35234764 2.233351e-07
## X3SsnPorch    3.206515e-02 3.523476e-01      NA 4.616510e-02
## model_1_resid 0.000000e+00 2.233351e-07 0.04616510      NA

# It is a list of matrices with names "r", "n" and "P"
c(typeof(res2),names(res2))

## [1] "list" "r"      "n"      "P"

# Identify the significant correlations
dim(res2$P)

## [1] 40 40

res2_price_p <- res2$P[response,]
res2_price_p

##      MSSubClass    LotFrontage      LotArea    OverallQual    OverallCond
## 1.861489e-11 0.000000e+00 0.000000e+00 1.000000e+00 1.118522e-02
##      YearBuilt    YearRemodAdd    MasVnrArea    BsmtFinSF1    BsmtFinSF2
## 1.776357e-15 1.021405e-14 3.433867e-10 0.000000e+00 4.194278e-04
##      BsmtUnfSF    TotalBsmtSF    LowQualFinSF    GrLivArea    BsmtFullBath
## 4.796603e-02 0.000000e+00 3.856256e-01 0.000000e+00 0.000000e+00
## BsmtHalfBath      FullBath      HalfBath    BedroomAbvGr    KitchenAbvGr
## 6.676885e-02 0.000000e+00 1.609614e-09 0.000000e+00 8.571477e-01
## TotRmsAbvGrd    Fireplaces    GarageYrBlt    GarageCars    GarageArea
## 0.000000e+00 0.000000e+00 9.185973e-08 0.000000e+00 0.000000e+00
##      WoodDeckSF    OpenPorchSF    EnclosedPorch    X3SsnPorch.1    ScreenPorch
## 0.000000e+00 4.967757e-06 2.038168e-04 4.616510e-02 5.838197e-06
##      PoolArea      MiscVal      MoSold      YrSold    SalePriceLog
## 2.730217e-01 7.085247e-01 9.714902e-01 3.230896e-01 0.000000e+00
##      secondFloor    X1stFlrSF    X2ndFlrSF    X3SsnPorch    model_1_resid
## 6.306428e-01 0.000000e+00 2.233351e-07 4.616510e-02      NA

correlation_significant<-res2_price_p <= sig & !is.na(res2_price_p)

Subset to only fields significantly correlated with there response

# extract the correlations of the response to other fields
res2_price_r <- res2$r[response,]
res2_price_r

##      MSSubClass    LotFrontage      LotArea    OverallQual    OverallCond
## -1.745767e-01 2.753770e-01 2.964332e-01 5.898617e-10 6.637624e-02
##      YearBuilt    YearRemodAdd    MasVnrArea    BsmtFinSF1    BsmtFinSF2
## 2.062489e-01 2.005591e-01 1.637771e-01 3.056541e-01 9.220241e-02
##      BsmtUnfSF    TotalBsmtSF    LowQualFinSF    GrLivArea    BsmtFullBath
## -5.176760e-02 2.995357e-01 -2.272208e-02 3.753281e-01 2.523295e-01
## BsmtHalfBath      FullBath      HalfBath    BedroomAbvGr    KitchenAbvGr
## 4.799120e-02 2.512752e-01 1.570442e-01 2.185297e-01 4.714991e-03
## TotRmsAbvGrd    Fireplaces    GarageYrBlt    GarageCars    GarageArea
## 3.211701e-01 2.866517e-01 1.432449e-01 3.292399e-01 3.324428e-01
##      WoodDeckSF    OpenPorchSF    EnclosedPorch    X3SsnPorch.1    ScreenPorch
## 2.409721e-01 1.191774e-01 -9.705961e-02 5.219138e-02 1.182968e-01
##      PoolArea      MiscVal      MoSold      YrSold    SalePriceLog
## 2.870572e-02 9.791946e-03 -9.361438e-04 -2.587835e-02 5.763763e-01
##      secondFloor    X1stFlrSF    X2ndFlrSF    X3SsnPorch    model_1_resid
## 1.259398e-02 3.605602e-01 1.350295e-01 5.219138e-02 1.000000e+00
```

```
# Select the significant correlations
```

```
features_significant <- res2_price_r[correlation_significant]
features_significant
```

```
##      MSSubClass    LotFrontage      LotArea      YearBuilt    YearRemodAdd
##      -0.17457670    0.27537698    0.29643324    0.20624892    0.20055911
##      MasVnrArea    BsmtFinSF1    BsmtFinSF2    TotalBsmtSF    GrLivArea
##      0.16377710    0.30565411    0.09220241    0.29953566    0.37532809
##      BsmtFullBath    FullBath      HalfBath    BedroomAbvGr    TotRmsAbvGrd
##      0.25232947    0.25127515    0.15704420    0.21852972    0.32117006
##      Fireplaces    GarageYrBlt    GarageCars    GarageArea    WoodDeckSF
##      0.28665167    0.14324492    0.32923988    0.33244279    0.24097209
##      OpenPorchSF    EnclosedPorch    ScreenPorch    SalePriceLog    X1stFlrSF
##      0.11917742    -0.09705961    0.11829682    0.57637626    0.36056021
##      X2ndFlrSF
##      0.13502951
```

```
# Sort by correlation
```

```
features_significant <- features_significant[order(abs(features_significant),decreasing=TRUE)]
features_significant
```

```
##      SalePriceLog    GrLivArea    X1stFlrSF    GarageArea    GarageCars
##      0.57637626    0.37532809    0.36056021    0.33244279    0.32923988
##      TotRmsAbvGrd    BsmtFinSF1    TotalBsmtSF    LotArea    Fireplaces
##      0.32117006    0.30565411    0.29953566    0.29643324    0.28665167
##      LotFrontage    BsmtFullBath    FullBath    WoodDeckSF    BedroomAbvGr
##      0.27537698    0.25232947    0.25127515    0.24097209    0.21852972
##      YearBuilt    YearRemodAdd    MSSubClass    MasVnrArea    HalfBath
##      0.20624892    0.20055911    -0.17457670    0.16377710    0.15704420
##      GarageYrBlt    X2ndFlrSF    OpenPorchSF    ScreenPorch    EnclosedPorch
##      0.14324492    0.13502951    0.11917742    0.11829682    -0.09705961
##      BsmtFinSF2
##      0.09220241
```

```
names(features_significant)
```

```
##      [1] "SalePriceLog" "GrLivArea"    "X1stFlrSF"    "GarageArea"
##      [5] "GarageCars"   "TotRmsAbvGrd" "BsmtFinSF1"    "TotalBsmtSF"
##      [9] "LotArea"      "Fireplaces"   "LotFrontage"   "BsmtFullBath"
##     [13] "FullBath"     "WoodDeckSF"   "BedroomAbvGr"  "YearBuilt"
##     [17] "YearRemodAdd" "MSSubClass"   "MasVnrArea"    "HalfBath"
##     [21] "GarageYrBlt"  "X2ndFlrSF"   "OpenPorchSF"   "ScreenPorch"
##     [25] "EnclosedPorch" "BsmtFinSF2"
```

```
remove_response <- !names(features_significant) %in% c('SalePriceLog',response)
remove_response
```

```
##      [1] FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
##     [12] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
##     [23] TRUE TRUE TRUE TRUE
```

```
features_significant <- features_significant[remove_response]
features_significant
```

```
##      GrLivArea    X1stFlrSF    GarageArea    GarageCars    TotRmsAbvGrd
##      0.37532809    0.36056021    0.33244279    0.32923988    0.32117006
##      BsmtFinSF1    TotalBsmtSF    LotArea    Fireplaces    LotFrontage
```

```
##      0.30565411      0.29953566      0.29643324      0.28665167      0.27537698
##      BsmtFullBath      FullBath      WoodDeckSF      BedroomAbvGr      YearBuilt
##      0.25232947      0.25127515      0.24097209      0.21852972      0.20624892
##      YearRemodAdd      MSSubClass      MasVnrArea      HalfBath      GarageYrBlt
##      0.20055911      -0.17457670      0.16377710      0.15704420      0.14324492
##      X2ndFlrSF      OpenPorchSF      ScreenPorch      EnclosedPorch      BsmtFinSF2
##      0.13502951      0.11917742      0.11829682      -0.09705961      0.09220241
```

```
#redo the t_selected
```

```
t_selected <- t_numeric %>% select(c(names(features_significant)[1:num_of_canadate_features],response))
names(t_selected)
```

```
##      [1] "GrLivArea"      "X1stFlrSF"      "GarageArea"      "GarageCars"
##      [5] "TotRmsAbvGrd"   "BsmtFinSF1"      "TotalBsmtSF"      "LotArea"
##      [9] "Fireplaces"      "LotFrontage"      "BsmtFullBath"      "FullBath"
##     [13] "WoodDeckSF"      "BedroomAbvGr"      "YearBuilt"      "model_1_resid"
```

```
# Note that the correlation for OverallQual is zero. It was the field used last time, so the correlation should be zero
```

```
features_significant
```

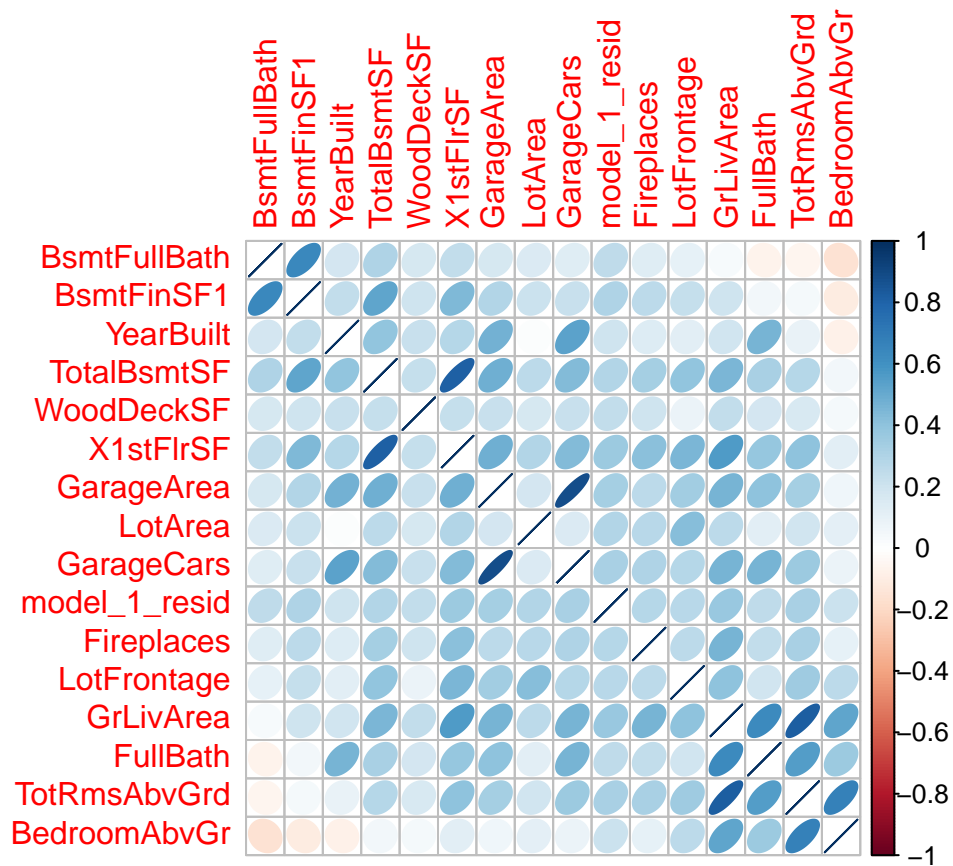
```
##      GrLivArea      X1stFlrSF      GarageArea      GarageCars      TotRmsAbvGrd
##      0.37532809      0.36056021      0.33244279      0.32923988      0.32117006
##      BsmtFinSF1      TotalBsmtSF      LotArea      Fireplaces      LotFrontage
##      0.30565411      0.29953566      0.29643324      0.28665167      0.27537698
##      BsmtFullBath      FullBath      WoodDeckSF      BedroomAbvGr      YearBuilt
##      0.25232947      0.25127515      0.24097209      0.21852972      0.20624892
##      YearRemodAdd      MSSubClass      MasVnrArea      HalfBath      GarageYrBlt
##      0.20055911      -0.17457670      0.16377710      0.15704420      0.14324492
##      X2ndFlrSF      OpenPorchSF      ScreenPorch      EnclosedPorch      BsmtFinSF2
##      0.13502951      0.11917742      0.11829682      -0.09705961      0.09220241
```

```
names(t_selected)
```

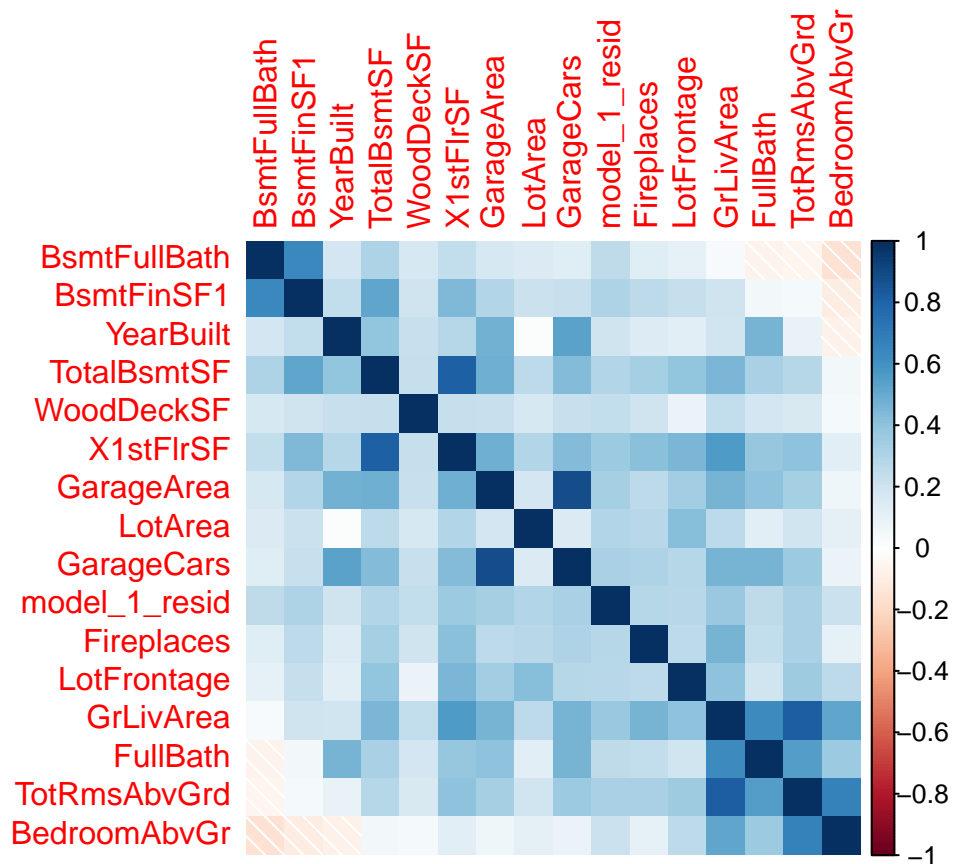
```
##      [1] "GrLivArea"      "X1stFlrSF"      "GarageArea"      "GarageCars"
##      [5] "TotRmsAbvGrd"   "BsmtFinSF1"      "TotalBsmtSF"      "LotArea"
##      [9] "Fireplaces"      "LotFrontage"      "BsmtFullBath"      "FullBath"
##     [13] "WoodDeckSF"      "BedroomAbvGr"      "YearBuilt"      "model_1_resid"
```

A few versions of the correlation matrix

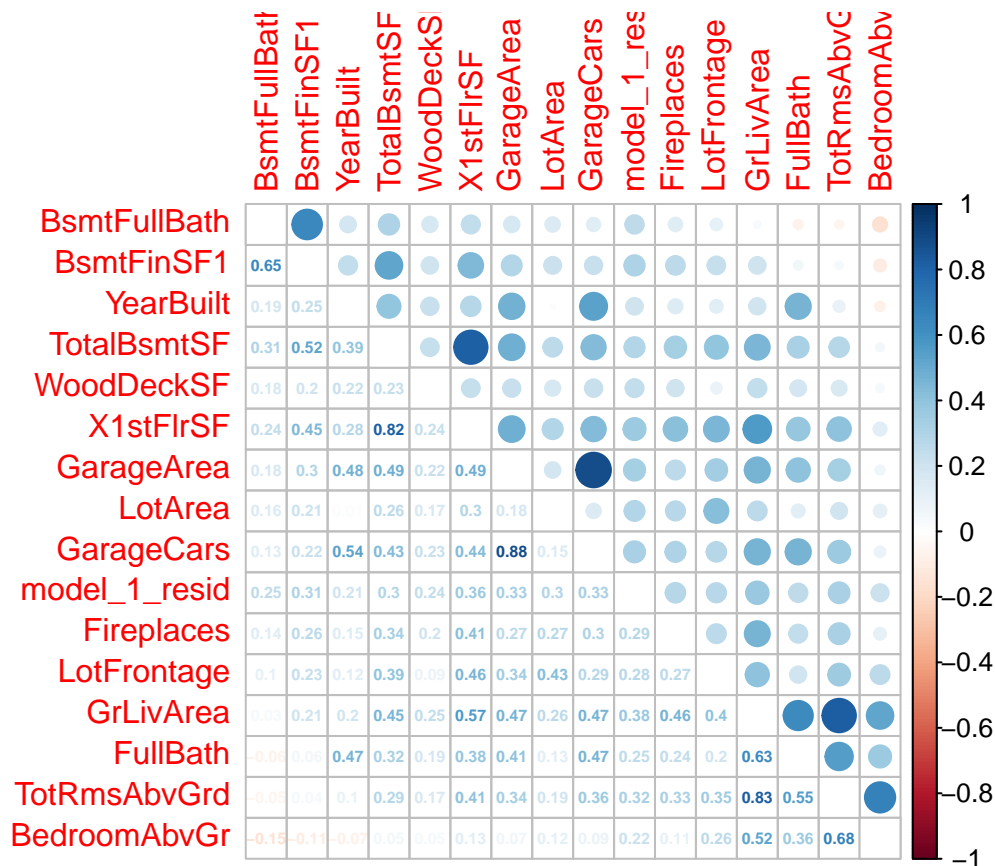
```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='ellipse',order="AOE")
```



```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='shade',order="AOE")
```

```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot.mixed(number.cex = 0.5,order = "AOE",tl.pos = "tl.pos")
```



```
#t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='circle',method='ellipse')
```

Here is the top feature to try

```
#feature_list <-list()
selected<-names(features_significant)[1]
selected
```

```
## [1] "GrLivArea"
```

```
selected_label<-label(t_numeric[selected])
selected_label
```

```
##
## GrLivArea
## "Above grade (ground) living area square feet"
```

```
feature_list[[model_num]] <- selected
str(feature_list)
```

```
## List of 2
## $ : chr "OverallQual"
## $ : chr "GrLivArea"
```

Calculate a model using this variable

```
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
```

```
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_1_resid"
```

```
feature_list
```

```
## [[1]]
## [1] "OverallQual"
##
## [[2]]
## [1] "GrLivArea"
```

```
my_formula <- paste0('SalePriceLog ~ ',paste(feature_list,collapse = '+ ', sep=" "))
str(my_formula)
```

```
## chr "SalePriceLog ~ OverallQual+ GrLivArea"
```

```
model_2 <- t_numeric %>% lm(my_formula,data=.)
```

```
t_numeric<- model_2 %>% augment(type.predict = 'response',type.residuals = 'response',data=t_numeric) %>%
```

```
mutate(model_2_resid = .resid) %>%
```

```
mutate(influential = .cooks > 4/length(SalePriceLog))
```

```
head(t_numeric)
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 1 60 65 8450 7 5 2003
## 2 20 80 9600 6 8 1976
## 3 60 68 11250 7 5 2001
## 4 70 60 9550 7 5 1915
## 5 60 84 14260 8 5 2000
## 6 50 85 14115 5 5 1993
## YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1 2003 196 706 0 150 856
## 2 1976 0 978 0 284 1262
## 3 2002 162 486 0 434 920
## 4 1970 0 216 0 540 756
## 5 2000 350 655 0 490 1145
## 6 1995 0 732 0 64 796
## LowQualFinSF GrLivArea BsmtFullBath BsmtHalfBath FullBath HalfBath
## 1 0 1710 1 0 2 1
## 2 0 1262 0 1 2 0
## 3 0 1786 1 0 2 1
## 4 0 1717 1 0 1 0
## 5 0 2198 1 0 2 1
## 6 0 1362 1 0 1 1
## BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars
## 1 3 1 8 0 2003 2
## 2 3 1 6 1 1976 2
## 3 3 1 6 1 2001 2
## 4 3 1 7 1 1998 3
## 5 4 1 9 1 2000 3
## 6 1 1 5 0 1993 2
## GarageArea WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch.1 ScreenPorch
```

```
## 1      548      0      61      0      0      0
## 2      460     298      0      0      0      0
## 3      608      0      42      0      0      0
## 4      642      0      35     272      0      0
## 5      836     192      84      0      0      0
## 6      480      40      30      0     320      0
##   PoolArea MiscVal MoSold YrSold SalePriceLog secondFloor X1stFlrSF
## 1      0      0      2   2008    12.24769      TRUE      856
## 2      0      0      5   2007    12.10901     FALSE     1262
## 3      0      0      9   2008    12.31717      TRUE      920
## 4      0      0      2   2006    11.84940      TRUE      961
## 5      0      0     12   2008    12.42922      TRUE     1145
## 6      0     700     10   2009    11.87060      TRUE      796
##   X2ndFlrSF X3SsnPorch model_1_resid .fitted .se.fit .resid
## 1      854      0    0.01105660 12.23451 0.006367596 0.01318822
## 2      0      0    0.10840116 11.94200 0.006090822 0.16701430
## 3      866      0    0.08052897 12.25378 0.006423538 0.06338382
## 4      756      0   -0.38724002 12.23628 0.006366738 -0.38688389
## 5     1053      0   -0.04344948 12.53716 0.009619118 -0.10794504
## 6      566     320    0.10601809 11.78848 0.006918237 0.08211702
##   .hat .sigma .cooksd .std.resid model_2_resid
## 1 0.0009759041 0.2039015 1.364466e-06 0.06473309 0.01318822
## 2 0.0008929106 0.2038547 2.001825e-04 0.81973925 0.16701430
## 3 0.0009931269 0.2038950 3.207447e-05 0.31111593 0.06338382
## 4 0.0009756412 0.2036493 1.173908e-03 -1.89898135 -0.38688389
## 5 0.0022270333 0.2038821 2.091238e-04 -0.53016971 -0.10794504
## 6 0.0011519857 0.2038904 6.246685e-05 0.40309887 0.08211702
##   influential
## 1      FALSE
## 2      FALSE
## 3      FALSE
## 4      FALSE
## 5      FALSE
## 6      FALSE
```

Check the fit using the Kaggle criteria

```
rmse(model_2,t_numeric)
```

```
## [1] 0.2036223
```

```
t_numeric[t_numeric$influential,]
```

```
##   MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 31         70         50     8500          4          4     1920
## 94        190         60     7200          6          6     1910
## 113        60         77     9965          7          5     2007
## 126        190         60     6780          6          8     1935
## 152        20        107    13891          8          5     2007
## 156        50         60     9600          6          5     1924
## 179        20         63    17423          9          5     2008
## 186        75         90    22950         10          9     1892
## 198        75        174    25419          8          4     1918
## 199        75         92     5520          6          6     1912
## 236       160         21     1680          6          3     1971
## 243        50         63     5000          5          4     1900
```

## 265	30	30	5232	5	5	1925
## 314	20	150	215245	7	5	1965
## 324	20	49	5820	3	8	1955
## 336	190	NA	164660	5	6	1965
## 384	45	60	9000	6	3	1928
## 399	30	60	8967	5	2	1920
## 411	20	68	9571	5	3	1958
## 431	160	21	1680	6	5	1971
## 432	50	60	5586	6	7	1920
## 441	20	105	15431	10	5	2008
## 450	50	50	6000	3	7	1948
## 458	20	NA	53227	4	6	1954
## 459	70	NA	5100	8	7	1925
## 463	20	60	8281	5	5	1965
## 474	20	110	14977	8	5	2006
## 496	30	60	7879	4	5	1920
## 524	60	130	40094	10	5	2007
## 534	20	50	5000	1	3	1946
## 584	75	75	13500	10	9	1893
## 608	20	78	7800	5	8	1948
## 633	20	85	11900	7	5	1977
## 636	190	60	10896	6	7	1914
## 656	160	21	1680	6	5	1971
## 665	20	49	20896	8	5	2005
## 667	60	NA	18450	6	5	1965
## 677	70	60	9600	4	2	1900
## 689	20	60	8089	8	6	2007
## 706	190	70	5600	4	5	1930
## 711	30	56	4130	3	6	1935
## 717	70	60	10800	7	8	1890
## 770	60	47	53504	8	5	2003
## 804	60	107	13891	9	5	2008
## 810	75	90	8100	5	5	1898
## 813	20	66	8712	5	5	1952
## 875	50	52	5720	5	6	1941
## 899	20	100	12919	9	5	2009
## 917	20	50	9000	2	3	1949
## 969	50	50	5925	3	6	1910
## 1021	20	60	7024	4	5	2005
## 1032	75	102	15863	7	3	1920
## 1046	20	NA	13680	3	5	1955
## 1047	60	85	16056	9	5	2005
## 1063	190	85	13600	5	5	1900
## 1069	160	42	3964	6	4	1973
## 1133	70	90	9900	6	4	1880
## 1143	60	77	9965	8	5	2006
## 1149	50	NA	5700	7	7	1926
## 1182	120	64	5587	8	5	2008
## 1191	190	NA	32463	4	4	1961
## 1212	50	152	12134	8	7	1988
## 1267	190	60	10120	7	4	1910
## 1280	50	60	7500	4	4	1920
## 1293	70	60	6600	5	4	1892
## 1299	60	313	63887	10	5	2008

##	1325	20	75	9986	8	5	2006
##	1326	30	40	3636	4	4	1922
##	1338	30	153	4118	4	4	1941
##	1350	70	50	5250	8	5	1872
##	1377	30	52	6292	6	5	1930
##	1379	160	21	1953	6	5	1973
##	1381	30	45	8212	3	3	1914
##	1388	50	60	8520	6	7	1916
##	1433	30	60	10800	4	6	1927
##	1438	20	96	12444	8	5	2008
##	YearRemodAdd	MasVnrArea	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF	
##	31	1950	0	0	0	649	649
##	94	1998	0	1046	0	168	1214
##	113	2007	220	984	0	280	1264
##	126	1982	0	490	0	30	520
##	152	2008	436	1400	0	310	1710
##	156	1950	0	0	0	572	572
##	179	2009	748	1904	0	312	2216
##	186	1993	0	0	0	1107	1107
##	198	1990	0	1036	184	140	1360
##	199	1950	0	0	0	755	755
##	236	1971	604	358	0	125	483
##	243	1950	0	0	0	540	540
##	265	2004	0	0	0	680	680
##	314	1965	0	1236	820	80	2136
##	324	2005	0	256	0	906	1162
##	336	1965	0	1249	147	103	1499
##	384	1950	0	0	0	784	784
##	399	1950	0	0	0	961	961
##	411	1958	0	0	0	1276	1276
##	431	1971	232	387	0	96	483
##	432	1998	0	0	0	901	901
##	441	2008	200	1767	539	788	3094
##	450	2002	0	331	0	318	649
##	458	1994	0	1116	0	248	1364
##	459	1996	0	0	0	588	588
##	463	1965	0	553	311	0	864
##	474	2007	304	1350	0	626	1976
##	496	1950	0	495	0	225	720
##	524	2008	762	2260	0	878	3138
##	534	1950	0	0	0	0	0
##	584	2000	0	0	0	1237	1237
##	608	2002	0	603	0	293	896
##	633	1977	209	822	0	564	1386
##	636	1995	0	256	0	1184	1440
##	656	1971	381	0	0	525	525
##	665	2006	0	1721	0	356	2077
##	667	1979	113	187	723	111	1021
##	677	1950	0	0	0	1095	1095
##	689	2007	0	945	0	474	1419
##	706	1950	0	0	0	0	0
##	711	2003	0	0	0	270	270
##	717	1998	0	0	0	718	718
##	770	2003	603	1416	0	234	1650

## 804	2009	424	0	0	1734	1734
## 810	1965	0	0	0	849	849
## 813	1952	0	0	0	540	540
## 875	1950	0	0	0	676	676
## 899	2010	760	2188	0	142	2330
## 917	1950	0	50	0	430	480
## 969	1950	0	0	0	600	600
## 1021	2005	0	1024	0	108	1132
## 1032	1970	0	523	0	301	824
## 1046	1955	0	0	0	0	0
## 1047	2006	208	240	0	1752	1992
## 1063	1950	0	0	0	662	662
## 1069	1973	0	837	0	105	942
## 1133	1950	0	0	0	1008	1008
## 1143	2007	340	1150	0	316	1466
## 1149	1950	0	0	0	572	572
## 1182	2008	186	1480	0	120	1600
## 1191	1975	149	1159	0	90	1249
## 1212	2005	0	427	0	132	559
## 1267	1950	0	0	0	925	925
## 1280	1950	0	0	0	698	698
## 1293	1965	0	0	0	994	994
## 1299	2008	796	5644	0	466	6110
## 1325	2007	428	0	0	1795	1795
## 1326	1950	0	0	0	796	796
## 1338	1950	0	0	0	693	693
## 1350	1987	0	259	0	425	684
## 1377	1950	0	384	0	384	768
## 1379	1973	408	309	0	174	483
## 1381	1950	0	203	0	661	864
## 1388	1950	0	168	546	0	714
## 1433	2007	0	0	0	656	656
## 1438	2008	426	1336	0	596	1932
##	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	FullBath	HalfBath
## 31	0	1317	0	0	1	0
## 94	0	2291	0	1	2	0
## 113	0	2696	1	0	2	1
## 126	234	754	1	0	1	0
## 152	0	1710	1	0	2	0
## 156	0	1096	0	0	1	0
## 179	0	2234	1	0	2	0
## 186	572	3608	0	0	2	1
## 198	392	3112	1	1	2	0
## 199	371	2229	0	0	1	0
## 236	0	987	0	0	1	1
## 243	0	1440	0	0	1	0
## 265	0	764	0	0	1	0
## 314	0	2036	2	0	2	0
## 324	0	1163	1	0	1	0
## 336	0	1786	2	0	2	0
## 384	0	784	0	0	1	0
## 399	0	1077	0	0	1	0
## 411	0	1276	0	0	1	0
## 431	0	987	0	0	1	1

## 432	0	1198	0	0	1	0
## 441	0	2402	1	0	2	0
## 450	0	1183	0	0	1	1
## 458	0	1663	1	0	1	0
## 459	0	1666	0	0	1	0
## 463	0	864	0	0	1	0
## 474	0	1976	1	0	2	0
## 496	0	720	0	0	1	0
## 524	0	4676	1	0	3	1
## 534	0	334	0	0	1	0
## 584	0	2775	0	0	3	1
## 608	0	2008	1	0	3	0
## 633	0	1411	0	0	2	0
## 636	515	3395	0	0	2	0
## 656	0	1092	0	0	1	1
## 665	0	2097	1	0	1	1
## 667	0	2380	0	0	2	1
## 677	0	1774	1	0	2	0
## 689	0	1419	1	0	2	0
## 706	0	1092	0	0	2	0
## 711	0	729	0	0	1	0
## 717	0	2554	0	0	1	1
## 770	0	3279	1	0	3	1
## 804	0	2822	0	0	3	1
## 810	0	2138	0	0	2	0
## 813	0	1044	0	0	1	0
## 875	0	1131	0	0	1	1
## 899	0	2364	1	0	2	1
## 917	0	480	1	0	0	0
## 969	0	968	0	0	1	0
## 1021	0	1132	1	0	1	1
## 1032	397	3082	1	0	2	1
## 1046	0	1733	0	0	2	0
## 1047	0	2868	0	0	3	1
## 1063	0	2337	0	0	2	0
## 1069	0	2521	1	0	2	1
## 1133	0	2210	0	0	2	0
## 1143	0	2828	1	0	3	0
## 1149	0	1111	0	0	1	0
## 1182	0	1652	1	1	2	0
## 1191	0	1622	1	0	1	0
## 1212	0	1752	0	0	2	0
## 1267	0	1889	0	0	1	1
## 1280	0	1128	0	0	1	0
## 1293	0	2372	0	0	2	0
## 1299	0	5642	2	0	2	1
## 1325	0	1795	0	0	2	0
## 1326	0	796	0	0	1	0
## 1338	0	693	0	0	1	0
## 1350	205	2358	0	0	2	0
## 1377	0	790	0	0	1	0
## 1379	0	987	0	0	1	1
## 1381	0	864	1	0	1	0
## 1388	0	2526	0	0	2	0

## 1433	0	968	0	0	2	0
## 1438	0	1932	1	0	2	0
##	BedroomAbvGr	KitchenAbvGr	TotRmsAbvGrd	Fireplaces	GarageYrBlt	
## 31	3	1	6	0	1920	
## 94	4	2	9	1	1900	
## 113	4	1	10	1	2007	
## 126	2	1	5	0	NA	
## 152	2	1	6	1	2007	
## 156	2	1	5	0	NA	
## 179	1	1	9	1	2009	
## 186	4	1	12	2	1993	
## 198	4	1	8	1	1918	
## 199	5	1	8	0	NA	
## 236	2	1	5	0	1971	
## 243	3	1	6	0	1940	
## 265	2	1	4	0	1965	
## 314	3	1	8	2	1965	
## 324	3	1	6	0	1955	
## 336	3	1	7	2	1965	
## 384	2	1	5	0	1950	
## 399	2	1	6	0	1920	
## 411	3	1	5	0	1958	
## 431	2	1	4	0	1971	
## 432	4	1	7	0	NA	
## 441	2	1	10	2	2008	
## 450	2	1	6	0	1981	
## 458	2	1	6	2	1954	
## 459	3	1	7	1	1925	
## 463	3	1	5	1	1965	
## 474	2	1	7	1	2006	
## 496	2	1	4	0	NA	
## 524	3	1	11	1	2007	
## 534	1	1	2	0	NA	
## 584	3	1	9	1	1988	
## 608	3	1	8	0	1948	
## 633	3	1	6	1	1977	
## 636	8	2	14	0	NA	
## 656	3	1	6	0	1971	
## 665	1	1	8	1	2005	
## 667	3	1	7	1	1965	
## 677	4	2	8	0	1920	
## 689	2	1	7	1	2007	
## 706	3	2	7	0	NA	
## 711	2	1	5	0	NA	
## 717	3	1	8	0	1996	
## 770	4	1	12	1	2003	
## 804	4	1	12	1	2009	
## 810	2	3	11	0	1910	
## 813	2	1	4	0	1952	
## 875	3	1	5	0	1941	
## 899	2	1	11	2	2009	
## 917	1	1	4	0	1958	
## 969	2	1	6	0	NA	
## 1021	2	1	5	0	2005	

## 1032	5	1	12	2	1970
## 1046	4	1	8	1	1955
## 1047	4	1	11	1	2005
## 1063	5	2	10	0	1945
## 1069	5	1	10	1	1973
## 1133	5	1	8	0	1930
## 1143	4	1	11	1	2006
## 1149	2	1	5	1	1982
## 1182	2	1	5	1	2008
## 1191	3	1	7	1	1975
## 1212	4	1	8	0	1988
## 1267	4	2	9	1	1960
## 1280	2	1	6	0	1980
## 1293	4	2	11	0	1985
## 1299	3	1	12	3	2008
## 1325	2	1	7	1	2007
## 1326	2	1	5	0	NA
## 1338	2	1	4	0	NA
## 1350	4	1	8	0	NA
## 1377	2	1	4	0	1925
## 1379	2	1	5	0	1973
## 1381	2	1	5	0	1938
## 1388	5	1	10	1	1916
## 1433	4	1	5	0	1928
## 1438	2	1	7	1	2008
##	GarageCars	GarageArea	WoodDeckSF	OpenPorchSF	EnclosedPorch
## 31	1	250	0	54	172
## 94	2	506	0	0	0
## 113	3	792	120	184	0
## 126	0	0	53	0	0
## 152	3	866	0	102	0
## 156	0	0	0	8	128
## 179	3	1166	0	60	0
## 186	3	840	0	260	0
## 198	2	795	0	16	552
## 199	0	0	0	198	30
## 236	1	264	0	0	0
## 243	1	352	0	0	77
## 265	2	504	0	0	0
## 314	2	513	0	0	0
## 324	1	220	142	98	0
## 336	2	529	670	0	0
## 384	2	360	0	0	91
## 399	1	338	0	0	0
## 411	1	350	0	0	0
## 431	1	264	0	0	0
## 432	0	0	0	98	0
## 441	3	672	0	72	0
## 450	1	308	0	176	0
## 458	2	529	224	137	0
## 459	1	228	192	63	0
## 463	1	360	0	0	236
## 474	3	908	250	63	0
## 496	0	0	0	523	115

## 524	3	884	208	406	0		
## 534	0	0	0	0	0		
## 584	2	880	105	502	0		
## 608	1	230	103	0	0		
## 633	2	544	192	0	0		
## 636	0	0	0	110	0		
## 656	1	264	0	0	0		
## 665	3	1134	192	267	0		
## 667	2	596	0	265	0		
## 677	3	779	0	0	90		
## 689	2	567	140	0	0		
## 706	0	0	0	0	0		
## 711	0	0	0	0	0		
## 717	2	704	0	48	143		
## 770	3	841	503	36	0		
## 804	3	1020	52	170	0		
## 810	2	360	40	156	0		
## 813	2	504	0	0	0		
## 875	1	200	26	0	0		
## 899	3	820	0	67	0		
## 917	1	308	0	0	0		
## 969	0	0	0	0	0		
## 1021	2	451	252	64	0		
## 1032	2	672	136	63	0		
## 1046	2	452	0	0	0		
## 1047	3	716	214	108	0		
## 1063	2	560	0	57	0		
## 1069	2	576	728	20	0		
## 1133	1	205	0	48	0		
## 1143	3	1052	125	144	0		
## 1149	1	288	0	0	176		
## 1182	2	482	162	53	0		
## 1191	4	1356	439	0	0		
## 1212	2	492	325	12	0		
## 1267	1	308	0	0	264		
## 1280	2	528	30	0	164		
## 1293	1	432	0	287	0		
## 1299	2	1418	214	292	0		
## 1325	3	895	0	49	0		
## 1326	0	0	0	0	100		
## 1338	0	0	0	20	0		
## 1350	0	0	0	54	20		
## 1377	1	160	0	141	0		
## 1379	1	264	72	0	0		
## 1381	1	200	0	0	96		
## 1388	1	216	88	15	0		
## 1433	1	216	0	0	0		
## 1438	3	774	0	66	0		
##	X3SsnPorch.1	ScreenPorch	PoolArea	MiscVal	MoSold	YrSold	SalePriceLog
## 31	0	0	0	0	7	2008	10.59663
## 94	0	99	0	0	11	2007	11.80485
## 113	0	168	0	0	10	2007	12.85832
## 126	0	0	0	0	6	2006	11.34451
## 152	0	0	0	0	1	2008	12.82773

## 156	0	0	0	0	4	2008	11.27720
## 179	0	0	0	0	7	2009	13.12603
## 186	0	410	0	0	6	2006	13.07107
## 198	0	0	512	0	3	2006	12.36734
## 199	0	0	0	0	7	2009	11.55215
## 236	0	0	0	0	8	2008	11.40199
## 243	0	0	0	0	4	2006	11.27720
## 265	0	0	0	0	6	2008	11.19821
## 314	0	0	0	0	6	2009	12.83468
## 324	0	0	0	0	7	2006	11.74543
## 336	0	0	0	700	8	2008	12.34126
## 384	0	0	0	0	10	2009	11.23849
## 399	0	0	0	0	11	2007	11.11245
## 411	0	0	0	0	6	2009	11.00210
## 431	0	0	0	0	7	2008	11.35510
## 432	0	0	0	0	9	2008	11.28853
## 441	0	170	0	0	4	2009	13.22672
## 450	0	0	0	0	6	2007	11.69525
## 458	0	0	0	0	3	2008	12.45293
## 459	0	0	0	0	6	2008	11.98916
## 463	0	0	0	0	12	2009	11.04105
## 474	0	0	0	0	7	2007	12.99453
## 496	0	0	0	0	11	2009	10.46024
## 524	0	0	0	0	10	2007	12.12676
## 534	0	0	0	0	1	2007	10.57898
## 584	0	0	0	0	7	2008	12.69158
## 608	0	0	0	0	8	2006	12.32386
## 633	0	0	0	0	4	2009	11.32055
## 636	0	0	0	0	3	2007	12.20607
## 656	0	0	0	0	3	2010	11.38509
## 665	0	0	0	0	1	2006	12.95513
## 667	0	0	0	0	8	2007	11.76757
## 677	0	0	0	0	5	2006	11.37366
## 689	0	0	0	0	10	2007	12.87902
## 706	0	0	0	3500	7	2010	10.91509
## 711	0	0	0	0	7	2008	10.85900
## 717	0	0	0	0	7	2007	11.97980
## 770	0	210	0	0	6	2010	13.19561
## 804	0	192	0	0	1	2009	13.27583
## 810	0	0	0	0	11	2009	11.57119
## 813	0	0	0	54	6	2010	10.93298
## 875	0	0	0	0	8	2009	11.10496
## 899	0	0	0	0	3	2010	13.32393
## 917	0	0	0	0	10	2006	10.47195
## 969	0	0	0	0	5	2009	10.54271
## 1021	0	0	0	0	6	2008	12.07824
## 1032	0	0	0	0	8	2009	12.19096
## 1046	0	0	0	0	6	2009	11.84654
## 1047	0	0	0	0	7	2006	13.22957
## 1063	0	0	0	0	9	2007	11.40756
## 1069	0	0	0	0	6	2006	11.92768
## 1133	0	0	0	0	5	2007	11.67419
## 1143	0	0	0	0	4	2007	12.95954
## 1149	0	0	0	0	8	2008	11.66907

##	1182	153	0	0	0	11	2008	12.88029
##	1191	0	0	0	0	3	2007	12.03172
##	1212	0	0	0	0	6	2010	12.00762
##	1267	0	0	0	0	1	2007	11.71178
##	1280	0	0	0	0	4	2010	11.13313
##	1293	0	0	0	0	12	2009	11.58525
##	1299	0	0	480	0	1	2008	11.98293
##	1325	0	0	0	0	2	2007	11.89819
##	1326	0	0	0	0	1	2008	10.91509
##	1338	0	0	0	0	3	2006	10.86857
##	1350	0	0	0	0	12	2008	11.71178
##	1377	0	0	0	0	4	2008	11.41861
##	1379	0	0	0	0	6	2006	11.32660
##	1381	0	0	0	0	6	2010	10.97678
##	1388	0	0	0	0	8	2007	11.82041
##	1433	0	0	0	0	8	2007	11.07442
##	1438	304	0	0	0	11	2008	12.88567
##	secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch model_1_resid .fitted							
##	31	TRUE	649	668	0	-0.93191913	11.59819	
##	94	TRUE	1260	1031	0	-0.19576124	12.20299	
##	113	TRUE	1282	1414	0	0.62168198	12.48460	
##	126	FALSE	520	0	0	-0.65610296	11.81315	
##	152	FALSE	1710	0	0	0.35506352	12.41338	
##	156	TRUE	572	524	0	-0.72340664	11.89989	
##	179	FALSE	2234	0	0	0.41733702	12.72517	
##	186	TRUE	1518	1518	0	0.12634851	13.25255	
##	198	TRUE	1360	1360	0	-0.10532488	12.76899	
##	199	TRUE	929	929	0	-0.44846359	12.18727	
##	236	TRUE	483	504	0	-0.59861586	11.87225	
##	243	TRUE	889	551	0	-0.48737869	11.80827	
##	265	FALSE	764	0	0	-0.56636710	11.63681	
##	314	FALSE	2036	0	0	0.59804358	12.31719	
##	324	FALSE	1163	0	0	0.45289920	11.38025	
##	336	TRUE	1619	167	0	0.57667710	11.89603	
##	384	FALSE	784	0	0	-0.76212115	11.82076	
##	399	FALSE	1077	0	0	-0.65213392	11.71619	
##	411	FALSE	1276	0	0	-0.76248198	11.76667	
##	431	TRUE	483	504	0	-0.64550839	11.87225	
##	432	TRUE	1088	110	0	-0.71207864	11.92576	
##	441	FALSE	2402	0	0	0.28200182	12.94666	
##	450	TRUE	679	504	0	0.40272111	11.38533	
##	458	FALSE	1663	0	0	0.92437886	11.68595	
##	459	TRUE	833	833	0	-0.48350603	12.40222	
##	463	FALSE	864	0	0	-0.72353374	11.66217	
##	474	FALSE	1976	0	0	0.52186433	12.48085	
##	496	FALSE	720	0	0	-1.06831176	11.44677	
##	524	TRUE	3138	1538	0	-0.81796274	13.52344	
##	534	FALSE	334	0	0	-0.24149021	10.81223	
##	584	TRUE	1521	1254	0	-0.25314112	13.04127	
##	608	TRUE	1112	896	0	0.55927386	11.95234	
##	633	FALSE	1411	0	0	-0.91608415	12.15867	
##	636	TRUE	1440	1440	0	0.20546288	12.48301	
##	656	TRUE	525	567	0	-0.61551768	11.89888	
##	665	FALSE	2097	0	0	0.48246179	12.51154	

## 667	TRUE	1465	915	0	-0.23304209	12.22557
## 677	TRUE	1095	679	0	-0.15489047	11.71411
## 689	FALSE	1419	0	0	0.40635145	12.33957
## 706	TRUE	372	720	0	-0.61346540	11.54112
## 711	FALSE	729	0	0	-0.43352692	11.27017
## 717	TRUE	1576	978	0	-0.25683852	12.44858
## 770	TRUE	1690	1589	0	0.72294817	12.81135
## 804	TRUE	1734	1088	0	0.56713391	12.87431
## 810	TRUE	1075	1063	0	-0.19338744	11.98531
## 813	FALSE	1044	0	0	-0.83159986	11.70782
## 875	TRUE	676	455	0	-0.65962459	11.72989
## 899	FALSE	2364	0	0	0.61523332	12.75814
## 917	FALSE	480	0	0	-0.58454815	11.02814
## 969	TRUE	600	368	0	-0.74981952	11.33079
## 1021	FALSE	1132	0	0	0.54968541	11.55127
## 1032	TRUE	1687	998	0	-0.04567871	12.58250
## 1046	FALSE	1733	0	0	0.55401056	11.52483
## 1047	TRUE	1992	876	0	0.52087437	12.88598
## 1063	TRUE	1422	915	0	-0.35701687	12.03578
## 1069	TRUE	1291	1230	0	-0.07292915	12.26133
## 1133	TRUE	1178	1032	0	-0.32641616	12.18245
## 1143	TRUE	1466	1362	0	0.48687285	12.69696
## 1149	TRUE	572	539	0	-0.56756357	12.08257
## 1182	FALSE	1652	0	153	0.40762614	12.39867
## 1191	FALSE	1622	0	0	0.50316539	11.67555
## 1212	TRUE	1080	672	0	-0.46504397	12.42404
## 1267	TRUE	964	925	0	-0.52486140	12.27991
## 1280	TRUE	698	430	0	-0.39542576	11.55025
## 1293	TRUE	1378	994	0	-0.17933569	12.04466
## 1299	TRUE	4692	950	0	-0.96179248	13.76846
## 1325	FALSE	1795	0	0	-0.57447781	12.43494
## 1326	FALSE	796	0	0	-0.61346540	11.46604
## 1338	FALSE	693	0	0	-0.65998542	11.43992
## 1350	TRUE	938	1215	0	-0.76088935	12.57774
## 1377	FALSE	790	0	0	-0.58199498	11.82228
## 1379	TRUE	483	504	0	-0.67401388	11.87225
## 1381	FALSE	864	0	0	-0.31574388	11.30441
## 1388	TRUE	1664	862	0	-0.18019960	12.26260
## 1433	FALSE	968	0	0	-0.45413336	11.50967
## 1438	FALSE	1932	0	304	0.41300528	12.46969
##	.se.fit	.resid	.hat	.sigma	.cooksd	.std.resid
## 31	0.010298626	-1.0015563	0.002552788	0.2022009	0.020650026	-4.919925
## 94	0.011396250	-0.3981455	0.003125936	0.2036338	0.004000536	-1.956365
## 113	0.013876285	0.3737228	0.004634497	0.2036653	0.005241692	1.837750
## 126	0.010746735	-0.4686398	0.002779772	0.2035305	0.004925386	-2.302352
## 152	0.009535817	0.4143452	0.002188628	0.2036118	0.003027851	2.035008
## 156	0.007326987	-0.6226890	0.001292133	0.2032468	0.004030022	-3.056891
## 179	0.012425717	0.4008604	0.003716200	0.2036300	0.004826743	1.970288
## 186	0.022119625	-0.1814820	0.011776377	0.2038456	0.003186421	-0.895641
## 198	0.017301582	-0.4016491	0.007204897	0.2036279	0.009460999	-1.977630
## 199	0.010712048	-0.6351221	0.002761857	0.2032194	0.008987788	-3.120225
## 236	0.008327299	-0.4702512	0.001669031	0.2035284	0.002971052	-2.308983
## 243	0.007153182	-0.5310638	0.001231558	0.2034256	0.002793531	-2.607008
## 265	0.009317575	-0.4385904	0.002089594	0.2035769	0.003238411	-2.153979

## 314	0.007521925	0.5174880	0.001361803	0.2034496	0.002933819	2.540530
## 324	0.013802128	0.3651728	0.004585094	0.2036760	0.004950759	1.795661
## 336	0.009442901	0.4452319	0.002146185	0.2035669	0.003427999	2.186658
## 384	0.010420257	-0.5822672	0.002613443	0.2033285	0.007146043	-2.860347
## 399	0.007227609	-0.6037471	0.001257319	0.2032862	0.003686237	-2.963851
## 411	0.006816184	-0.7645698	0.001118250	0.2029138	0.005256297	-3.753083
## 431	0.008327299	-0.5171438	0.001669031	0.2034501	0.003593130	-2.539230
## 432	0.006515719	-0.6372324	0.001021835	0.2032160	0.003335796	-3.127865
## 441	0.015968645	0.2800634	0.006137512	0.2037688	0.003910095	1.378229
## 450	0.013883910	0.3099219	0.004639592	0.2037392	0.003608764	1.524018
## 458	0.012462921	0.7669816	0.003738487	0.2029049	0.017776816	3.769870
## 459	0.009721692	-0.4130641	0.002274783	0.2036136	0.003128152	-2.028804
## 463	0.008507737	-0.6211212	0.001742145	0.2032499	0.005411110	-3.049882
## 474	0.009077914	0.5136773	0.001983482	0.2034560	0.004215688	2.522607
## 496	0.010517150	-0.9865248	0.002662271	0.2022516	0.020898675	-4.846352
## 524	0.032909833	-1.3966827	0.026067990	0.2005001	0.430111143	-6.943230
## 534	0.020390362	-0.2332476	0.010007050	0.2038092	0.004456676	-1.150083
## 584	0.016674378	-0.3496880	0.006691993	0.2036943	0.006654017	-1.721341
## 608	0.011558868	0.3715202	0.003215783	0.2036684	0.003584134	1.825618
## 633	0.007454269	-0.8381136	0.001337415	0.2027137	0.007557349	-4.114543
## 636	0.024576034	-0.2769420	0.014537167	0.2037707	0.009211107	-1.368664
## 656	0.007361561	-0.5137854	0.001304356	0.2034561	0.002769671	-2.522280
## 665	0.009272533	0.4435840	0.002069440	0.2035694	0.003280493	2.178481
## 667	0.012398632	-0.4580005	0.003700017	0.2035469	0.006273217	-2.251122
## 677	0.013386455	-0.3404420	0.004313078	0.2037056	0.004045413	-1.673824
## 689	0.011226678	0.5394429	0.003033603	0.2034096	0.007125634	2.650533
## 706	0.009719177	-0.6260332	0.002273606	0.2032391	0.007181612	-3.074820
## 711	0.013141780	-0.4111728	0.004156852	0.2036156	0.005685469	-2.021422
## 717	0.012312458	-0.4687806	0.003648763	0.2035300	0.006480298	-2.304048
## 770	0.019127162	0.3842658	0.008805565	0.2036507	0.010617883	1.893566
## 804	0.014884394	0.4015159	0.005332348	0.2036286	0.006971126	1.975112
## 810	0.012919328	-0.4141146	0.004017317	0.2036116	0.005571966	-2.035742
## 813	0.007376603	-0.7748428	0.001309692	0.2028868	0.006325131	-3.803875
## 875	0.007030289	-0.6249344	0.001189605	0.2032422	0.003736287	-3.067757
## 899	0.012640141	0.5657832	0.003845564	0.2033598	0.009952719	2.781089
## 917	0.016752171	-0.5561872	0.006754581	0.2033766	0.016992691	-2.737920
## 969	0.013232207	-0.7880858	0.004214255	0.2028486	0.021177313	-3.874526
## 1021	0.009764601	0.5269719	0.002294908	0.2034325	0.005136541	2.588299
## 1032	0.018338408	-0.3915437	0.008094303	0.2036413	0.010118917	-1.928737
## 1046	0.017461259	0.3217083	0.007338500	0.2037261	0.006183929	1.584126
## 1047	0.015213057	0.3435888	0.005570436	0.2037017	0.005335241	1.690363
## 1063	0.015109606	-0.6282187	0.005494934	0.2032323	0.017591592	-3.090548
## 1069	0.014023972	-0.3336511	0.004733673	0.2037133	0.004268149	-1.640782
## 1133	0.010505105	-0.5082555	0.002656177	0.2034651	0.005534345	-2.496823
## 1143	0.014374364	0.2625829	0.004973172	0.2037850	0.002778637	1.291449
## 1149	0.009960922	-0.4135005	0.002388115	0.2036129	0.003291691	-2.031063
## 1182	0.009786694	0.4816190	0.002305304	0.2035098	0.004309978	2.365554
## 1191	0.012144868	0.3561674	0.003550110	0.2036872	0.003638934	1.750470
## 1212	0.009385603	-0.4164153	0.002120218	0.2036089	0.002962186	-2.045105
## 1267	0.006720437	-0.5681317	0.001087054	0.2033568	0.002821167	-2.788774
## 1280	0.009758895	-0.4171247	0.002292226	0.2036079	0.003214529	-2.048766
## 1293	0.015504850	-0.4594150	0.005786172	0.2035439	0.009912358	-2.260442
## 1299	0.043962059	-1.7855304	0.046517057	0.1981905	1.308747360	-8.970959
## 1325	0.009260723	-0.5367557	0.002064173	0.2034150	0.004791032	-2.636049

```

## 1326 0.010184944 -0.5509552 0.002496741 0.2033886 0.006111001 -2.706370
## 1338 0.010653453 -0.5713501 0.002731725 0.2033497 0.007193705 -2.806884
## 1350 0.010466247 -0.8659676 0.002636563 0.2026315 0.015946650 -4.254054
## 1377 0.010355385 -0.4036629 0.002581004 0.2036264 0.003391611 -1.982933
## 1379 0.008327299 -0.5456492 0.001669031 0.2033989 0.004000161 -2.679195
## 1381 0.013108023 -0.3276314 0.004135524 0.2037202 0.003591170 -1.610696
## 1388 0.014082287 -0.4421897 0.004773122 0.2035706 0.007559805 -2.174582
## 1433 0.009744412 -0.4352496 0.002285427 0.2035817 0.003489526 -2.137781
## 1438 0.009069986 0.4159784 0.001980019 0.2036096 0.002759736 2.042816
##      model_2_resid influential
## 31      -1.0015563          TRUE
## 94      -0.3981455          TRUE
## 113      0.3737228          TRUE
## 126     -0.4686398          TRUE
## 152      0.4143452          TRUE
## 156     -0.6226890          TRUE
## 179      0.4008604          TRUE
## 186     -0.1814820          TRUE
## 198     -0.4016491          TRUE
## 199     -0.6351221          TRUE
## 236     -0.4702512          TRUE
## 243     -0.5310638          TRUE
## 265     -0.4385904          TRUE
## 314      0.5174880          TRUE
## 324      0.3651728          TRUE
## 336      0.4452319          TRUE
## 384     -0.5822672          TRUE
## 399     -0.6037471          TRUE
## 411     -0.7645698          TRUE
## 431     -0.5171438          TRUE
## 432     -0.6372324          TRUE
## 441      0.2800634          TRUE
## 450      0.3099219          TRUE
## 458      0.7669816          TRUE
## 459     -0.4130641          TRUE
## 463     -0.6211212          TRUE
## 474      0.5136773          TRUE
## 496     -0.9865248          TRUE
## 524     -1.3966827          TRUE
## 534     -0.2332476          TRUE
## 584     -0.3496880          TRUE
## 608      0.3715202          TRUE
## 633     -0.8381136          TRUE
## 636     -0.2769420          TRUE
## 656     -0.5137854          TRUE
## 665      0.4435840          TRUE
## 667     -0.4580005          TRUE
## 677     -0.3404420          TRUE
## 689      0.5394429          TRUE
## 706     -0.6260332          TRUE
## 711     -0.4111728          TRUE
## 717     -0.4687806          TRUE
## 770      0.3842658          TRUE
## 804      0.4015159          TRUE

```



```
## 810      -0.4141146      TRUE
## 813      -0.7748428      TRUE
## 875      -0.6249344      TRUE
## 899       0.5657832      TRUE
## 917      -0.5561872      TRUE
## 969      -0.7880858      TRUE
## 1021     0.5269719      TRUE
## 1032     -0.3915437      TRUE
## 1046     0.3217083      TRUE
## 1047     0.3435888      TRUE
## 1063     -0.6282187      TRUE
## 1069     -0.3336511      TRUE
## 1133     -0.5082555      TRUE
## 1143     0.2625829      TRUE
## 1149     -0.4135005      TRUE
## 1182     0.4816190      TRUE
## 1191     0.3561674      TRUE
## 1212     -0.4164153      TRUE
## 1267     -0.5681317      TRUE
## 1280     -0.4171247      TRUE
## 1293     -0.4594150      TRUE
## 1299     -1.7855304      TRUE
## 1325     -0.5367557      TRUE
## 1326     -0.5509552      TRUE
## 1338     -0.5713501      TRUE
## 1350     -0.8659676      TRUE
## 1377     -0.4036629      TRUE
## 1379     -0.5456492      TRUE
## 1381     -0.3276314      TRUE
## 1388     -0.4421897      TRUE
## 1433     -0.4352496      TRUE
## 1438      0.4159784      TRUE
```

```
model_2_parameters <- model_2 %>% tidy
print(model_2_parameters)
```

```
##           term      estimate  std.error statistic      p.value
## 1 (Intercept) 1.054863e+01 2.419707e-02  435.94667 0.000000e+00
## 2 OverallQual 1.788779e-01 4.792047e-03   37.32809 1.439420e-214
## 3   GrLivArea 2.536418e-04 1.261205e-05   20.11107 1.390664e-79
```

The R^2 is very good.

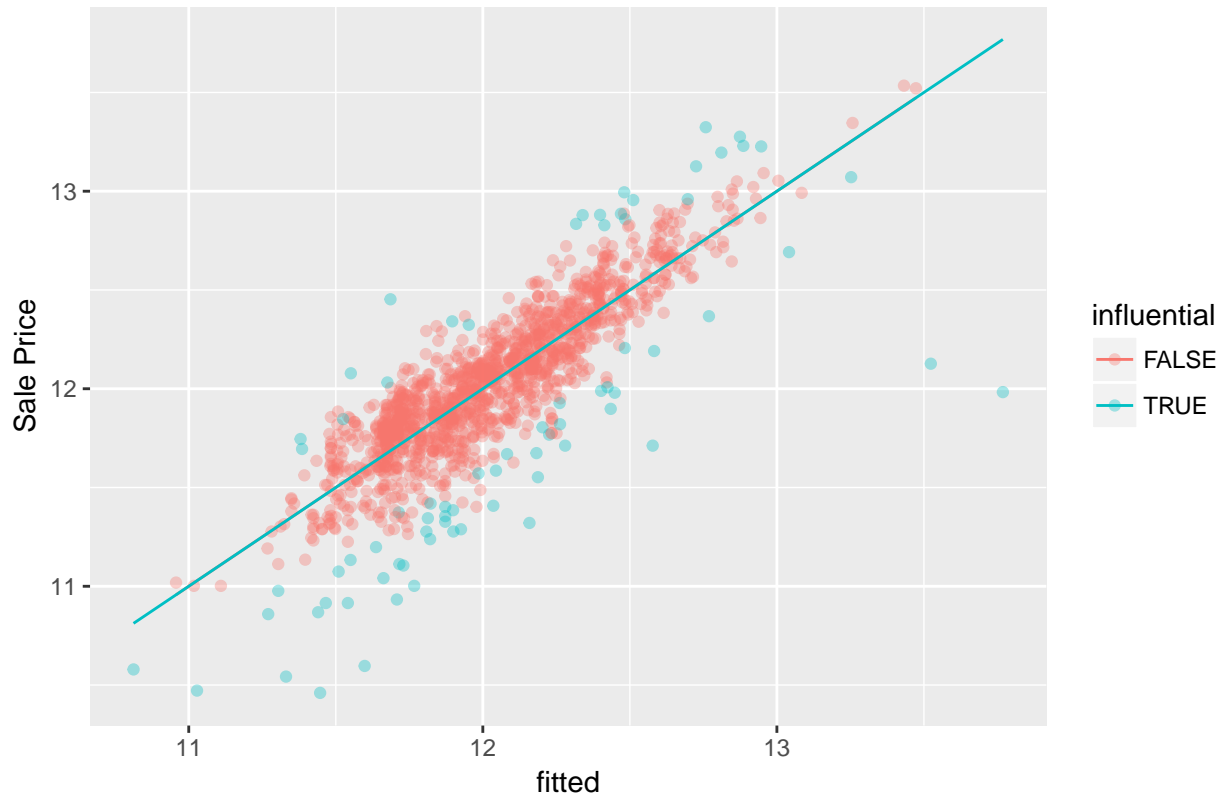
```
model_2_fit <- model_2 %>% glance
print(model_1_fit)
```

```
##    r.squared adj.r.squared    sigma statistic p.value df  logLik    AIC
## 1 0.6677904    0.6675625 0.2303135  2930.795      0  2 73.08848 -140.177
##           BIC deviance df.residual
## 1 -124.3184 77.33862      1458
```

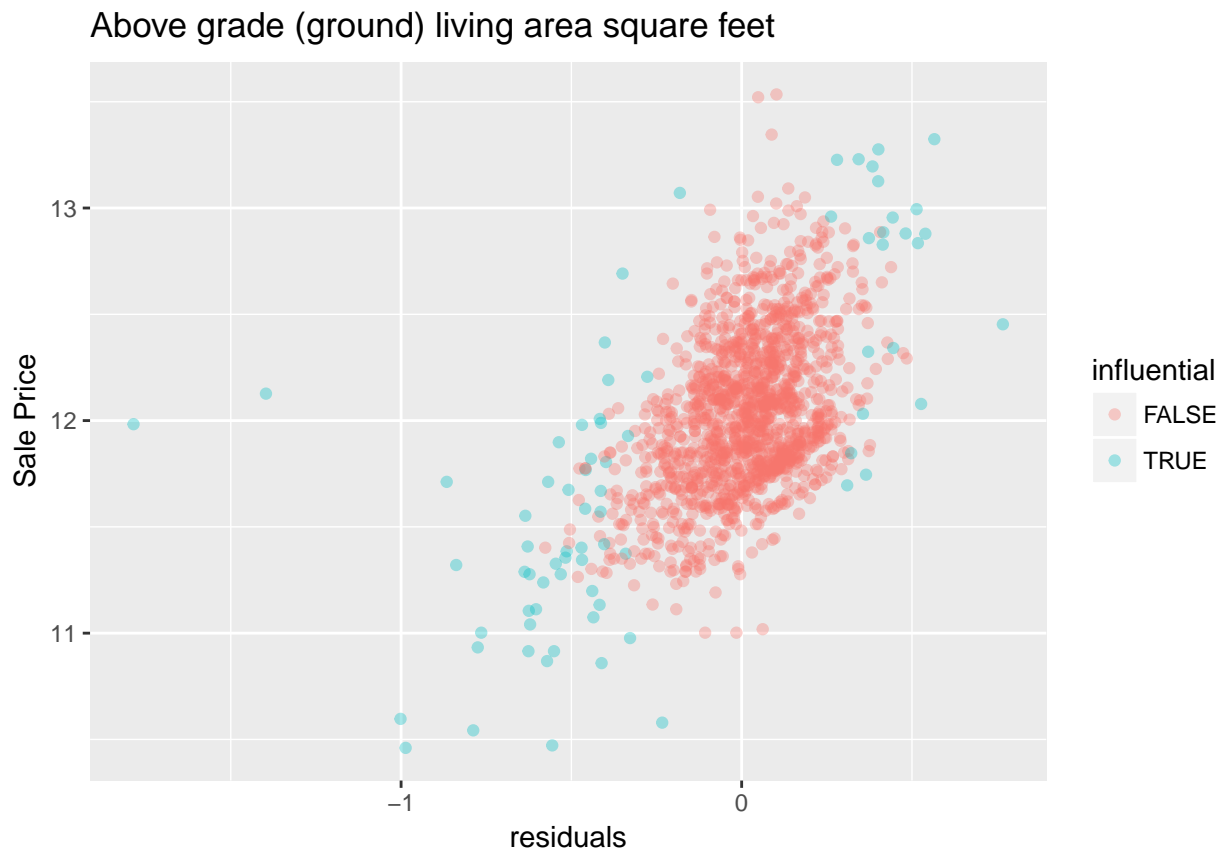
```
t_numeric %>% ggplot() +
  aes(.fitted, SalePriceLog, color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
```

```
geom_line(aes(.fitted,.fitted)) +
ggtitle(selected_label) +
xlab('fitted') +
ylab("Sale Price")
```

Above grade (ground) living area square feet

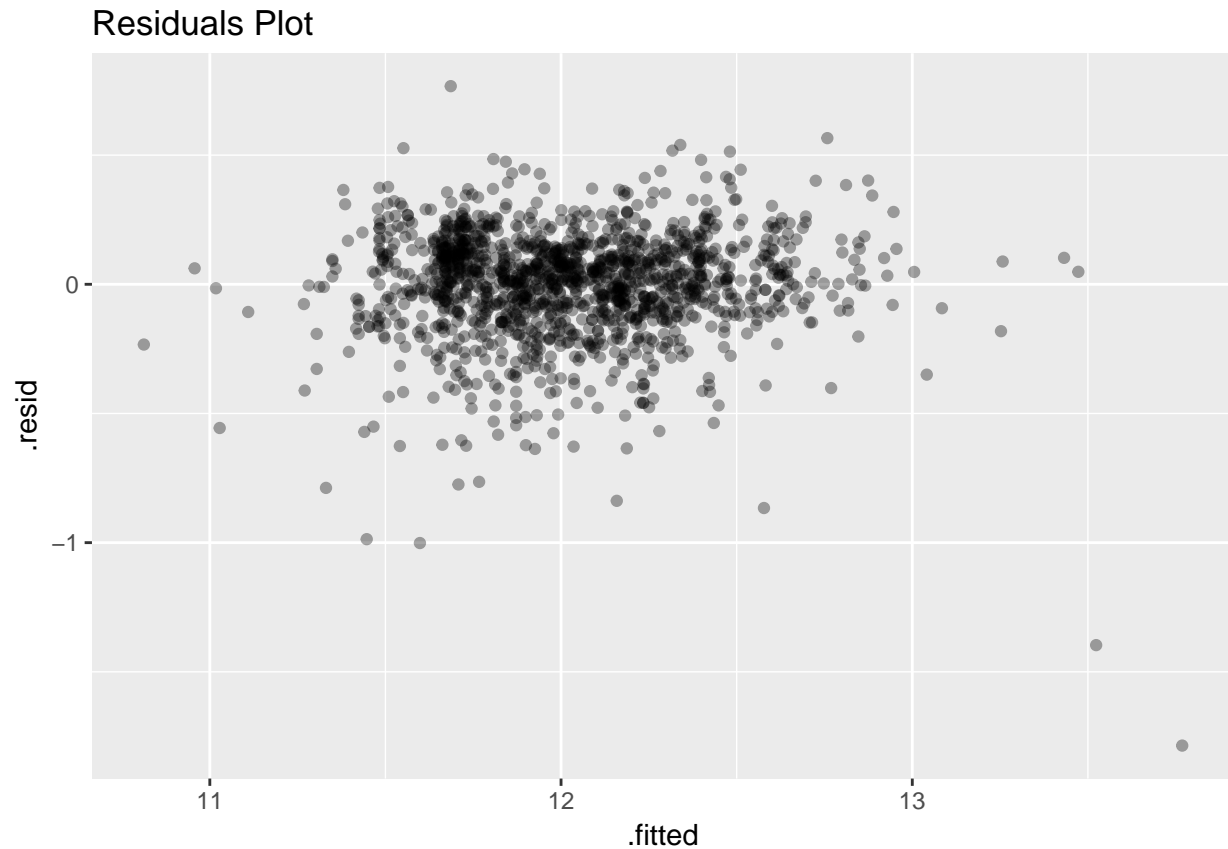


```
t_numeric %>% ggplot() +
  aes(.resid,SalePriceLog,color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
  ggtitle(selected_label) +
  xlab('residuals') +
  ylab("Sale Price")
```



This fitted by residual should show no pattern.

```
t_numeric %>% ggplot() +  
  aes(.fitted,.resid)+  
  scale_y_continuous(labels = comma) +  
  scale_x_continuous(labels = comma) +  
  geom_point(alpha=0.35) +  
  ggtitle(paste("Residuals Plot"))
```



```
p <- test %>%
  add_predictions(model_2) %>%
  select(Id,pred) %>%
  mutate(SalePrice = exp(pred)) %>%
  select(Id,SalePrice)
head(p)
```

```
## # A tibble: 6 x 2
##   Id SalePrice
##   <int>     <dbl>
## 1  1461 117041.7
## 2  1462 156215.3
## 3  1463 140956.3
## 4  1464 167500.6
## 5  1465 220646.7
## 6  1466 169681.4
```

```
write_csv(p,file.path('.', 'data', 'model_2.csv'))
```

Model 3

```
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
```

```
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_1_resid"
## [41] ".fitted" ".se.fit" ".resid" ".hat"
## [45] ".sigma" ".cooksd" ".std.resid" "model_2_resid"
## [49] "influential"
```

```
try(t_numeric <- t_numeric %>% select(-.hat,-.sigma,-.cooksd,-.std.resid,-.se.fit,-.fitted,-.resid,-inf)
try(t_numeric <- t_numeric %>% select(-model_1_resid))
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_2_resid"
```

Corrrelation

So now we search for factors that are significantly correlated and order those by strength of correlation.

```
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_2_resid"
```

```
model_num = 3
num_of_canadate_features = 15
sig <- 0.0010
response <- "model_2_resid"
names(t_numeric)
```

```
## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
## [9] "BsmtFinSF1" "BsmtFinSF2" "BsmtUnfSF" "TotalBsmtSF"
## [13] "LowQualFinSF" "GrLivArea" "BsmtFullBath" "BsmtHalfBath"
## [17] "FullBath" "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
## [21] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt" "GarageCars"
## [25] "GarageArea" "WoodDeckSF" "OpenPorchSF" "EnclosedPorch"
```

```
## [29] "X3SsnPorch.1" "ScreenPorch" "PoolArea" "MiscVal"
## [33] "MoSold" "YrSold" "SalePriceLog" "secondFloor"
## [37] "X1stFlrSF" "X2ndFlrSF" "X3SsnPorch" "model_2_resid"
```

```
# Calculate the correlation matrix
res2 <- rcorr(as.matrix(t_numeric))
res2$P
```

```
##          MSSubClass LotFrontage LotArea OverallQual
## MSSubClass      NA 0.000000e+00 8.194115e-08 2.127765e-01
## LotFrontage 0.000000e+00      NA 0.000000e+00 0.000000e+00
## LotArea      8.194115e-08 0.000000e+00      NA 5.105510e-05
## OverallQual 2.127765e-01 0.000000e+00 5.105510e-05      NA
## OverallCond 2.342026e-02 4.019570e-02 8.296279e-01 4.361720e-04
## YearBuilt   2.875793e-01 1.813803e-05 5.869923e-01 0.000000e+00
## YearRemodAdd 1.211632e-01 2.052218e-03 5.985903e-01 0.000000e+00
## MasVnrArea   3.824712e-01 1.536393e-11 6.995921e-05 0.000000e+00
## BsmtFinSF1   7.598818e-03 2.220446e-16 0.000000e+00 0.000000e+00
## BsmtFinSF2   1.210779e-02 8.388578e-02 2.068605e-05 2.388568e-02
## BsmtUnfSF    6.637585e-08 3.979765e-06 9.203746e-01 0.000000e+00
## TotalBsmtSF 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## LowQualFinSF 7.586520e-02 1.827783e-01 8.552305e-01 2.452458e-01
## GrLivArea    4.213733e-03 0.000000e+00 0.000000e+00 0.000000e+00
## BsmtFullBath 8.939735e-01 4.587932e-04 1.230799e-09 2.094408e-05
## BsmtHalfBath 9.290423e-01 8.022392e-01 6.646032e-02 1.251676e-01
## FullBath     4.502116e-07 3.634426e-12 1.360046e-06 0.000000e+00
## HalfBath     8.788081e-12 6.365788e-02 5.861562e-01 0.000000e+00
## BedroomAbvGr 3.708289e-01 0.000000e+00 4.519868e-06 9.949905e-05
## KitchenAbvGr 0.000000e+00 8.335889e-01 4.971427e-01 1.435740e-12
## TotRmsAbvGrd 1.230183e-01 0.000000e+00 2.460254e-13 0.000000e+00
## Fireplaces   8.175109e-02 0.000000e+00 0.000000e+00 0.000000e+00
## GarageYrBlt  1.566879e-03 1.834203e-02 3.545897e-01 0.000000e+00
## GarageCars   1.255479e-01 0.000000e+00 2.706817e-09 0.000000e+00
## GarageArea   1.592243e-04 0.000000e+00 3.802736e-12 0.000000e+00
## WoodDeckSF   6.310397e-01 2.136523e-03 4.001000e-11 0.000000e+00
## OpenPorchSF  8.158485e-01 1.210208e-07 1.185832e-03 0.000000e+00
## EnclosedPorch 6.458454e-01 7.110477e-01 4.837906e-01 1.276996e-05
## X3SsnPorch.1 9.414953e-02 1.520968e-02 4.355274e-01 2.461582e-01
## ScreenPorch  3.202578e-01 1.517836e-01 9.924777e-02 1.314582e-02
## PoolArea     7.518384e-01 5.393463e-13 2.979905e-03 1.275642e-02
## MiscVal      7.692691e-01 9.071905e-01 1.459891e-01 2.304123e-01
## MoSold       6.040063e-01 6.982029e-01 9.633078e-01 6.790959e-03
## YrSold       4.137257e-01 7.964813e-01 5.861053e-01 2.963852e-01
## SalePriceLog 4.692451e-03 0.000000e+00 0.000000e+00 0.000000e+00
## secondFloor  0.000000e+00 1.136398e-01 6.543549e-01 1.513700e-11
## X1stFlrSF    0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## X2ndFlrSF    0.000000e+00 5.433170e-03 5.144266e-02 0.000000e+00
## X3SsnPorch   9.414953e-02 1.520968e-02 4.355274e-01 2.461582e-01
## model_2_resid 0.000000e+00 5.930748e-07 3.552714e-15 1.000000e+00
##          OverallCond YearBuilt YearRemodAdd MasVnrArea
## MSSubClass 2.342026e-02 2.875793e-01 1.211632e-01 3.824712e-01
## LotFrontage 4.019570e-02 1.813803e-05 2.052218e-03 1.536393e-11
## LotArea 8.296279e-01 5.869923e-01 5.985903e-01 6.995921e-05
## OverallQual 4.361720e-04 0.000000e+00 0.000000e+00 0.000000e+00
## OverallCond      NA 0.000000e+00 4.816091e-03 9.716410e-07
```

## YearBuilt	0.000000e+00	NA	0.000000e+00	0.000000e+00
## YearRemodAdd	4.816091e-03	0.000000e+00	NA	5.400569e-12
## MasVnrArea	9.716410e-07	0.000000e+00	5.400569e-12	NA
## BsmtFinSF1	7.741099e-02	0.000000e+00	8.466438e-07	0.000000e+00
## BsmtFinSF2	1.244256e-01	6.066982e-02	9.602934e-03	5.834234e-03
## BsmtUnfSF	1.530091e-07	1.053188e-08	3.104628e-12	1.234288e-05
## TotalBsmtSF	4.685274e-11	0.000000e+00	0.000000e+00	0.000000e+00
## LowQualFinSF	3.303245e-01	1.475930e-12	1.706455e-02	8.467646e-03
## GrLivArea	2.311138e-03	1.665335e-14	0.000000e+00	0.000000e+00
## BsmtFullBath	3.580810e-02	4.964917e-13	4.707214e-06	1.138425e-03
## BsmtHalfBath	6.367484e-06	1.449943e-01	6.376325e-01	3.097834e-01
## FullBath	7.238654e-14	0.000000e+00	0.000000e+00	0.000000e+00
## HalfBath	2.022485e-02	0.000000e+00	1.677547e-12	9.325873e-15
## BedroomAbvGr	6.202024e-01	6.920572e-03	1.211642e-01	8.671369e-05
## KitchenAbvGr	8.754659e-04	1.753175e-11	9.270313e-09	1.520249e-01
## TotRmsAbvGrd	2.779336e-02	2.544832e-04	1.483258e-13	0.000000e+00
## Fireplaces	3.630812e-01	1.423519e-08	1.619633e-05	0.000000e+00
## GarageYrBlt	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## GarageCars	8.424372e-13	0.000000e+00	0.000000e+00	0.000000e+00
## GarageArea	5.945338e-09	0.000000e+00	0.000000e+00	0.000000e+00
## WoodDeckSF	8.987253e-01	0.000000e+00	1.998401e-15	9.347851e-10
## OpenPorchSF	2.133227e-01	3.623768e-13	0.000000e+00	1.548562e-06
## EnclosedPorch	7.159421e-03	0.000000e+00	7.771561e-14	2.570285e-05
## X3SsnPorch.1	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
## ScreenPorch	3.625216e-02	5.435519e-02	1.389957e-01	1.916276e-02
## PoolArea	9.395944e-01	8.501179e-01	8.238849e-01	6.553522e-01
## MiscVal	8.568156e-03	1.891692e-01	6.945331e-01	2.562107e-01
## MoSold	8.933753e-01	6.359583e-01	4.119179e-01	8.203518e-01
## YrSold	9.321243e-02	6.031268e-01	1.722487e-01	7.548601e-01
## SalePriceLog	1.591324e-01	0.000000e+00	0.000000e+00	0.000000e+00
## secondFloor	1.764981e-02	2.073907e-04	2.739984e-02	1.313690e-02
## X1stFlrSF	3.126349e-08	0.000000e+00	0.000000e+00	0.000000e+00
## X2ndFlrSF	2.690898e-01	6.939294e-01	7.777955e-08	2.121014e-11
## X3SsnPorch	3.301473e-01	2.311809e-01	8.367071e-02	4.742064e-01
## model_2_resid	4.646269e-04	0.000000e+00	0.000000e+00	6.821566e-04
##	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF
## MSSubClass	7.598818e-03	1.210779e-02	6.637585e-08	0.000000e+00
## LotFrontage	2.220446e-16	8.388578e-02	3.979765e-06	0.000000e+00
## LotArea	0.000000e+00	2.068605e-05	9.203746e-01	0.000000e+00
## OverallQual	0.000000e+00	2.388568e-02	0.000000e+00	0.000000e+00
## OverallCond	7.741099e-02	1.244256e-01	1.530091e-07	4.685274e-11
## YearBuilt	0.000000e+00	6.066982e-02	1.053188e-08	0.000000e+00
## YearRemodAdd	8.466438e-07	9.602934e-03	3.104628e-12	0.000000e+00
## MasVnrArea	0.000000e+00	5.834234e-03	1.234288e-05	0.000000e+00
## BsmtFinSF1	NA	5.555041e-02	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.555041e-02	NA	6.661338e-16	6.010660e-05
## BsmtUnfSF	0.000000e+00	6.661338e-16	NA	0.000000e+00
## TotalBsmtSF	0.000000e+00	6.010660e-05	0.000000e+00	NA
## LowQualFinSF	1.369761e-02	5.718554e-01	2.821334e-01	2.042388e-01
## GrLivArea	8.881784e-16	7.128488e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	0.000000e+00	1.083816e-09	0.000000e+00	0.000000e+00
## BsmtHalfBath	9.972478e-03	6.687447e-03	2.463966e-04	9.904177e-01
## FullBath	2.529072e-02	3.470204e-03	0.000000e+00	0.000000e+00
## HalfBath	8.707324e-01	2.195850e-01	1.163181e-01	6.227931e-02

```

## BedroomAbvGr 3.949967e-05 5.481779e-01 1.485798e-10 5.394644e-02
## KitchenAbvGr 1.950234e-03 1.196093e-01 2.506158e-01 8.449363e-03
## TotRmsAbvGrd 9.051660e-02 1.785383e-01 0.000000e+00 0.000000e+00
## Fireplaces 0.000000e+00 7.308663e-02 4.880435e-02 0.000000e+00
## GarageYrBlt 1.013706e-08 1.069187e-03 9.254819e-13 0.000000e+00
## GarageCars 0.000000e+00 1.439250e-01 0.000000e+00 0.000000e+00
## GarageArea 0.000000e+00 4.864925e-01 1.690648e-12 0.000000e+00
## WoodDeckSF 3.108624e-15 9.454565e-03 8.391603e-01 0.000000e+00
## OpenPorchSF 1.867888e-05 9.060156e-01 7.585228e-07 0.000000e+00
## EnclosedPorch 9.005499e-05 1.628382e-01 9.228150e-01 2.587670e-04
## X3SsnPorch.1 3.125032e-01 2.520757e-01 4.278958e-01 1.533765e-01
## ScreenPorch 1.778536e-02 6.748805e-04 6.310420e-01 1.232140e-03
## PoolArea 7.033176e-08 1.111537e-01 1.802008e-01 1.354119e-06
## MiscVal 8.915451e-01 8.504155e-01 3.627455e-01 4.804779e-01
## MoSold 5.482075e-01 5.614175e-01 1.827454e-01 6.143911e-01
## YrSold 5.835463e-01 2.259961e-01 1.150733e-01 5.676653e-01
## SalePriceLog 0.000000e+00 8.536300e-01 0.000000e+00 0.000000e+00
## secondFloor 0.000000e+00 7.216583e-06 1.797090e-01 0.000000e+00
## X1stFlrSF 0.000000e+00 2.020320e-04 0.000000e+00 0.000000e+00
## X2ndFlrSF 1.455248e-07 1.453587e-04 8.645239e-01 1.894085e-11
## X3SsnPorch 3.125032e-01 2.520757e-01 4.278958e-01 1.533765e-01
## model_2_resid 0.000000e+00 8.075468e-04 2.339482e-04 0.000000e+00
## LowQualFinSF GrLivArea BsmtFullBath BsmtHalfBath
## MSSubClass 7.586520e-02 4.213733e-03 8.939735e-01 9.290423e-01
## LotFrontage 1.827783e-01 0.000000e+00 4.587932e-04 8.022392e-01
## LotArea 8.552305e-01 0.000000e+00 1.230799e-09 6.646032e-02
## OverallQual 2.452458e-01 0.000000e+00 2.094408e-05 1.251676e-01
## OverallCond 3.303245e-01 2.311138e-03 3.580810e-02 6.367484e-06
## YearBuilt 1.475930e-12 1.665335e-14 4.964917e-13 1.449943e-01
## YearRemodAdd 1.706455e-02 0.000000e+00 4.707214e-06 6.376325e-01
## MasVnrArea 8.467646e-03 0.000000e+00 1.138425e-03 3.097834e-01
## BsmtFinSF1 1.369761e-02 8.881784e-16 0.000000e+00 9.972478e-03
## BsmtFinSF2 5.718554e-01 7.128488e-01 1.083816e-09 6.687447e-03
## BsmtUnfSF 2.821334e-01 0.000000e+00 0.000000e+00 2.463966e-04
## TotalBsmtSF 2.042388e-01 0.000000e+00 0.000000e+00 9.904177e-01
## LowQualFinSF NA 2.399685e-07 7.173345e-02 8.235244e-01
## GrLivArea 2.399685e-07 NA 1.834039e-01 4.700979e-01
## BsmtFullBath 7.173345e-02 1.834039e-01 NA 1.374508e-08
## BsmtHalfBath 8.235244e-01 4.700979e-01 1.374508e-08 NA
## FullBath 9.783903e-01 0.000000e+00 1.368378e-02 3.719872e-02
## HalfBath 3.011218e-01 0.000000e+00 2.379416e-01 6.375543e-01
## BedroomAbvGr 5.275460e-05 0.000000e+00 7.237188e-09 7.558104e-02
## KitchenAbvGr 7.739872e-01 1.282742e-04 1.129357e-01 1.473008e-01
## TotRmsAbvGrd 4.904244e-07 0.000000e+00 4.181609e-02 3.627516e-01
## Fireplaces 4.166751e-01 0.000000e+00 1.216473e-07 2.685363e-01
## GarageYrBlt 1.771551e-01 0.000000e+00 3.498871e-06 3.997996e-03
## GarageCars 3.002845e-04 0.000000e+00 4.259910e-07 4.250739e-01
## GarageArea 9.772127e-03 0.000000e+00 5.316858e-12 3.488402e-01
## WoodDeckSF 3.312869e-01 0.000000e+00 1.526557e-11 1.250635e-01
## OpenPorchSF 4.859080e-01 0.000000e+00 1.005792e-02 3.335712e-01
## EnclosedPorch 1.959094e-02 7.278971e-01 5.656746e-02 7.439517e-01
## X3SsnPorch.1 8.697349e-01 4.305893e-01 9.967684e-01 1.799353e-01
## ScreenPorch 3.061609e-01 1.021511e-04 3.767857e-01 2.199645e-01
## PoolArea 1.753504e-02 5.918421e-11 9.756789e-03 4.445328e-01

```


## MiscVal	8.848674e-01	9.265212e-01	3.788657e-01	7.785284e-01
## MoSold	3.971935e-01	5.495607e-02	3.328625e-01	2.093598e-01
## YrSold	2.694414e-01	1.630401e-01	1.038816e-02	7.554937e-02
## SalePriceLog	1.471041e-01	0.000000e+00	0.000000e+00	8.441568e-01
## secondFloor	2.385267e-01	0.000000e+00	0.000000e+00	3.410671e-01
## X1stFlrSF	5.866501e-01	0.000000e+00	0.000000e+00	9.404841e-01
## X2ndFlrSF	1.547482e-02	0.000000e+00	7.124057e-11	3.623802e-01
## X3SsnPorch	8.697349e-01	4.305893e-01	9.967684e-01	1.799353e-01
## model_2_resid	1.473062e-06	9.999999e-01	0.000000e+00	5.117504e-02
##	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr
## MSSubClass	4.502116e-07	8.788081e-12	3.708289e-01	0.000000e+00
## LotFrontage	3.634426e-12	6.365788e-02	0.000000e+00	8.335889e-01
## LotArea	1.360046e-06	5.861562e-01	4.519868e-06	4.971427e-01
## OverallQual	0.000000e+00	0.000000e+00	9.949905e-05	1.435740e-12
## OverallCond	7.238654e-14	2.022485e-02	6.202024e-01	8.754659e-04
## YearBuilt	0.000000e+00	0.000000e+00	6.920572e-03	1.753175e-11
## YearRemodAdd	0.000000e+00	1.677547e-12	1.211642e-01	9.270313e-09
## MasVnrArea	0.000000e+00	9.325873e-15	8.671369e-05	1.520249e-01
## BsmtFinSF1	2.529072e-02	8.707324e-01	3.949967e-05	1.950234e-03
## BsmtFinSF2	3.470204e-03	2.195850e-01	5.481779e-01	1.196093e-01
## BsmtUnfSF	0.000000e+00	1.163181e-01	1.485798e-10	2.506158e-01
## TotalBsmtSF	0.000000e+00	6.227931e-02	5.394644e-02	8.449363e-03
## LowQualFinSF	9.783903e-01	3.011218e-01	5.275460e-05	7.739872e-01
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	1.282742e-04
## BsmtFullBath	1.368378e-02	2.379416e-01	7.237188e-09	1.129357e-01
## BsmtHalfBath	3.719872e-02	6.375543e-01	7.558104e-02	1.473008e-01
## FullBath	NA	1.685132e-07	0.000000e+00	3.313170e-07
## HalfBath	1.685132e-07	NA	0.000000e+00	9.077592e-03
## BedroomAbvGr	0.000000e+00	0.000000e+00	NA	1.887379e-14
## KitchenAbvGr	3.313170e-07	9.077592e-03	1.887379e-14	NA
## TotRmsAbvGrd	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## Fireplaces	0.000000e+00	3.996803e-15	3.810705e-05	2.035535e-06
## GarageYrBlt	0.000000e+00	1.665335e-13	1.656563e-02	3.589658e-06
## GarageCars	0.000000e+00	0.000000e+00	9.897563e-04	5.307601e-02
## GarageArea	0.000000e+00	3.252643e-10	1.263773e-02	1.379977e-02
## WoodDeckSF	4.818368e-13	3.498384e-05	7.349737e-02	5.648744e-04
## OpenPorchSF	0.000000e+00	1.332268e-14	3.315592e-04	7.380715e-03
## EnclosedPorch	1.040475e-05	2.650862e-04	1.123474e-01	1.541632e-01
## X3SsnPorch.1	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
## ScreenPorch	7.569616e-01	5.628841e-03	9.063268e-02	4.863598e-02
## PoolArea	5.810529e-02	3.927890e-01	6.879729e-03	5.791972e-01
## MiscVal	5.853585e-01	9.607167e-01	7.668285e-01	1.720427e-02
## MoSold	3.278470e-02	7.297135e-01	7.542379e-02	3.099771e-01
## YrSold	4.526681e-01	6.950290e-01	1.690210e-01	2.262661e-01
## SalePriceLog	0.000000e+00	0.000000e+00	6.661338e-16	1.478820e-08
## secondFloor	0.000000e+00	0.000000e+00	0.000000e+00	1.129716e-01
## X1stFlrSF	0.000000e+00	4.334880e-06	1.041033e-06	9.243536e-03
## X2ndFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	2.344382e-02
## X3SsnPorch	1.769833e-01	8.494369e-01	3.499768e-01	3.475688e-01
## model_2_resid	1.085966e-03	6.589154e-01	3.687758e-02	4.604133e-07
##	TotRmsAbvGrd	Fireplaces	GarageYrBlt	GarageCars
## MSSubClass	1.230183e-01	8.175109e-02	1.566879e-03	1.255479e-01
## LotFrontage	0.000000e+00	0.000000e+00	1.834203e-02	0.000000e+00
## LotArea	2.460254e-13	0.000000e+00	3.545897e-01	2.706817e-09

## OverallQual	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## OverallCond	2.779336e-02	3.630812e-01	0.000000e+00	8.424372e-13
## YearBuilt	2.544832e-04	1.423519e-08	0.000000e+00	0.000000e+00
## YearRemodAdd	1.483258e-13	1.619633e-05	0.000000e+00	0.000000e+00
## MasVnrArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## BsmtFinSF1	9.051660e-02	0.000000e+00	1.013706e-08	0.000000e+00
## BsmtFinSF2	1.785383e-01	7.308663e-02	1.069187e-03	1.439250e-01
## BsmtUnfSF	0.000000e+00	4.880435e-02	9.254819e-13	0.000000e+00
## TotalBsmtSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## LowQualFinSF	4.904244e-07	4.166751e-01	1.771551e-01	3.002845e-04
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## BsmtFullBath	4.181609e-02	1.216473e-07	3.498871e-06	4.259910e-07
## BsmtHalfBath	3.627516e-01	2.685363e-01	3.997996e-03	4.250739e-01
## FullBath	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## HalfBath	0.000000e+00	3.996803e-15	1.665335e-13	0.000000e+00
## BedroomAbvGr	0.000000e+00	3.810705e-05	1.656563e-02	9.897563e-04
## KitchenAbvGr	0.000000e+00	2.035535e-06	3.589658e-06	5.307601e-02
## TotRmsAbvGrd	NA	0.000000e+00	3.283118e-08	0.000000e+00
## Fireplaces	0.000000e+00	NA	8.219437e-02	0.000000e+00
## GarageYrBlt	3.283118e-08	8.219437e-02	NA	0.000000e+00
## GarageCars	0.000000e+00	0.000000e+00	0.000000e+00	NA
## GarageArea	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## WoodDeckSF	1.758029e-10	1.221245e-14	0.000000e+00	0.000000e+00
## OpenPorchSF	0.000000e+00	7.290124e-11	0.000000e+00	0.000000e+00
## EnclosedPorch	8.740741e-01	3.432435e-01	0.000000e+00	6.067020e-09
## X3SsnPorch.1	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01
## ScreenPorch	2.326437e-02	1.195044e-12	5.077546e-03	5.373793e-02
## PoolArea	1.358848e-03	2.748854e-04	5.905625e-01	4.241329e-01
## MiscVal	3.443919e-01	9.571131e-01	2.289596e-01	9.987840e-02
## MoSold	1.586904e-01	7.660440e-02	8.430360e-01	1.217084e-01
## YrSold	1.874597e-01	3.575543e-01	9.699819e-01	1.351896e-01
## SalePriceLog	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## secondFloor	0.000000e+00	5.724116e-06	3.584860e-01	4.817098e-02
## X1stFlrSF	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## X2ndFlrSF	0.000000e+00	6.394885e-14	8.506609e-03	1.417977e-12
## X3SsnPorch	7.986069e-01	6.673530e-01	3.823249e-01	1.719841e-01
## model_2_resid	6.671082e-01	1.316058e-11	0.000000e+00	0.000000e+00
##	GarageArea	WoodDeckSF	OpenPorchSF	EnclosedPorch
## MSSubClass	1.592243e-04	6.310397e-01	8.158485e-01	6.458454e-01
## LotFrontage	0.000000e+00	2.136523e-03	1.210208e-07	7.110477e-01
## LotArea	3.802736e-12	4.001000e-11	1.185832e-03	4.837906e-01
## OverallQual	0.000000e+00	0.000000e+00	0.000000e+00	1.276996e-05
## OverallCond	5.945338e-09	8.987253e-01	2.133227e-01	7.159421e-03
## YearBuilt	0.000000e+00	0.000000e+00	3.623768e-13	0.000000e+00
## YearRemodAdd	0.000000e+00	1.998401e-15	0.000000e+00	7.771561e-14
## MasVnrArea	0.000000e+00	9.347851e-10	1.548562e-06	2.570285e-05
## BsmtFinSF1	0.000000e+00	3.108624e-15	1.867888e-05	9.005499e-05
## BsmtFinSF2	4.864925e-01	9.454565e-03	9.060156e-01	1.628382e-01
## BsmtUnfSF	1.690648e-12	8.391603e-01	7.585228e-07	9.228150e-01
## TotalBsmtSF	0.000000e+00	0.000000e+00	0.000000e+00	2.587670e-04
## LowQualFinSF	9.772127e-03	3.312869e-01	4.859080e-01	1.959094e-02
## GrLivArea	0.000000e+00	0.000000e+00	0.000000e+00	7.278971e-01
## BsmtFullBath	5.316858e-12	1.526557e-11	1.005792e-02	5.656746e-02
## BsmtHalfBath	3.488402e-01	1.250635e-01	3.335712e-01	7.439517e-01

## FullBath	0.000000e+00	4.818368e-13	0.000000e+00	1.040475e-05
## HalfBath	3.252643e-10	3.498384e-05	1.332268e-14	2.650862e-04
## BedroomAbvGr	1.263773e-02	7.349737e-02	3.315592e-04	1.123474e-01
## KitchenAbvGr	1.379977e-02	5.648744e-04	7.380715e-03	1.541632e-01
## TotRmsAbvGrd	0.000000e+00	1.758029e-10	0.000000e+00	8.740741e-01
## Fireplaces	0.000000e+00	1.221245e-14	7.290124e-11	3.432435e-01
## GarageYrBlt	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## GarageCars	0.000000e+00	0.000000e+00	0.000000e+00	6.067020e-09
## GarageArea	NA	0.000000e+00	0.000000e+00	3.064131e-06
## WoodDeckSF	0.000000e+00	NA	2.499829e-02	1.371106e-06
## OpenPorchSF	0.000000e+00	2.499829e-02	NA	3.691031e-04
## EnclosedPorch	3.064131e-06	1.371106e-06	3.691031e-04	NA
## X3SsnPorch.1	1.802697e-01	2.107784e-01	8.234948e-01	1.542416e-01
## ScreenPorch	4.952347e-02	4.569023e-03	4.502233e-03	1.529761e-03
## PoolArea	1.965912e-02	5.029192e-03	2.023972e-02	3.837510e-02
## MiscVal	2.954461e-01	7.153744e-01	4.779914e-01	4.832932e-01
## MoSold	2.854434e-01	4.224189e-01	6.454024e-03	2.699986e-01
## YrSold	2.958337e-01	3.951420e-01	2.769527e-02	7.050056e-01
## SalePriceLog	0.000000e+00	0.000000e+00	0.000000e+00	1.050805e-08
## secondFloor	9.467301e-01	2.184093e-01	5.305810e-08	5.719454e-04
## X1stFlrSF	0.000000e+00	0.000000e+00	4.440892e-16	1.258441e-02
## X2ndFlrSF	1.112965e-07	4.216853e-04	8.881784e-16	1.784426e-02
## X3SsnPorch	1.802697e-01	2.107784e-01	8.234948e-01	1.542416e-01
## model_2_resid	0.000000e+00	4.440892e-15	1.418879e-01	8.078760e-10
##	X3SsnPorch.1	ScreenPorch	PoolArea	MiscVal
## MSSubClass	0.09414953	3.202578e-01	7.518384e-01	0.769269145
## LotFrontage	0.01520968	1.517836e-01	5.393463e-13	0.907190524
## LotArea	0.43552735	9.924777e-02	2.979905e-03	0.145989130
## OverallQual	0.24615822	1.314582e-02	1.275642e-02	0.230412263
## OverallCond	0.33014731	3.625216e-02	9.395944e-01	0.008568156
## YearBuilt	0.23118088	5.435519e-02	8.501179e-01	0.189169158
## YearRemodAdd	0.08367071	1.389957e-01	8.238849e-01	0.694533106
## MasVnrArea	0.47420644	1.916276e-02	6.553522e-01	0.256210660
## BsmtFinSF1	0.31250322	1.778536e-02	7.033176e-08	0.891545102
## BsmtFinSF2	0.25207571	6.748805e-04	1.111537e-01	0.850415539
## BsmtUnfSF	0.42789582	6.310420e-01	1.802008e-01	0.362745457
## TotalBsmtSF	0.15337655	1.232140e-03	1.354119e-06	0.480477906
## LowQualFinSF	0.86973491	3.061609e-01	1.753504e-02	0.884867376
## GrLivArea	0.43058934	1.021511e-04	5.918421e-11	0.926521232
## BsmtFullBath	0.99676835	3.767857e-01	9.756789e-03	0.378865709
## BsmtHalfBath	0.17993527	2.199645e-01	4.445328e-01	0.778528444
## FullBath	0.17698330	7.569616e-01	5.810529e-02	0.585358529
## HalfBath	0.84943689	5.628841e-03	3.927890e-01	0.960716747
## BedroomAbvGr	0.34997682	9.063268e-02	6.879729e-03	0.766828521
## KitchenAbvGr	0.34756881	4.863598e-02	5.791972e-01	0.017204270
## TotRmsAbvGrd	0.79860693	2.326437e-02	1.358848e-03	0.344391918
## Fireplaces	0.66735300	1.195044e-12	2.748854e-04	0.957113099
## GarageYrBlt	0.38232493	5.077546e-03	5.905625e-01	0.228959597
## GarageCars	0.17198411	5.373793e-02	4.241329e-01	0.099878400
## GarageArea	0.18026968	4.952347e-02	1.965912e-02	0.295446101
## WoodDeckSF	0.21077845	4.569023e-03	5.029192e-03	0.715374444
## OpenPorchSF	0.82349481	4.502233e-03	2.023972e-02	0.477991449
## EnclosedPorch	0.15424164	1.529761e-03	3.837510e-02	0.483293192
## X3SsnPorch.1	NA	2.299724e-01	7.602900e-01	0.989218187

## ScreenPorch	0.22997242	NA	4.998823e-02	0.222498354
## PoolArea	0.76029001	4.998823e-02	NA	0.257249528
## MiscVal	0.98921819	2.224984e-01	2.572495e-01	NA
## MoSold	0.26038906	3.753591e-01	1.976284e-01	0.804175819
## YrSold	0.47654325	6.830660e-01	2.256080e-02	0.851418697
## SalePriceLog	0.03594763	3.408972e-06	7.631787e-03	0.444619672
## secondFloor	0.33748531	5.241591e-01	1.622043e-01	0.840707205
## X1stFlrSF	0.03206515	6.856867e-04	4.578567e-07	0.420551089
## X2ndFlrSF	0.35234764	1.209302e-01	1.832441e-03	0.536316831
## X3SsnPorch	0.00000000	2.299724e-01	7.602900e-01	0.989218187
## model_2_resid	0.02865667	4.035941e-04	4.043395e-02	0.985925981
##	MoSold	YrSold	SalePriceLog	secondFloor
## MSSubClass	6.040063e-01	4.137257e-01	4.692451e-03	0.000000e+00
## LotFrontage	6.982029e-01	7.964813e-01	0.000000e+00	1.136398e-01
## LotArea	9.633078e-01	5.861053e-01	0.000000e+00	6.543549e-01
## OverallQual	6.790959e-03	2.963852e-01	0.000000e+00	1.513700e-11
## OverallCond	8.933753e-01	9.321243e-02	1.591324e-01	1.764981e-02
## YearBuilt	6.359583e-01	6.031268e-01	0.000000e+00	2.073907e-04
## YearRemodAdd	4.119179e-01	1.722487e-01	0.000000e+00	2.739984e-02
## MasVnrArea	8.203518e-01	7.548601e-01	0.000000e+00	1.313690e-02
## BsmtFinSF1	5.482075e-01	5.835463e-01	0.000000e+00	0.000000e+00
## BsmtFinSF2	5.614175e-01	2.259961e-01	8.536300e-01	7.216583e-06
## BsmtUnfSF	1.827454e-01	1.150733e-01	0.000000e+00	1.797090e-01
## TotalBsmtSF	6.143911e-01	5.676653e-01	0.000000e+00	0.000000e+00
## LowQualFinSF	3.971935e-01	2.694414e-01	1.471041e-01	2.385267e-01
## GrLivArea	5.495607e-02	1.630401e-01	0.000000e+00	0.000000e+00
## BsmtFullBath	3.328625e-01	1.038816e-02	0.000000e+00	0.000000e+00
## BsmtHalfBath	2.093598e-01	7.554937e-02	8.441568e-01	3.410671e-01
## FullBath	3.278470e-02	4.526681e-01	0.000000e+00	0.000000e+00
## HalfBath	7.297135e-01	6.950290e-01	0.000000e+00	0.000000e+00
## BedroomAbvGr	7.542379e-02	1.690210e-01	6.661338e-16	0.000000e+00
## KitchenAbvGr	3.099771e-01	2.262661e-01	1.478820e-08	1.129716e-01
## TotRmsAbvGrd	1.586904e-01	1.874597e-01	0.000000e+00	0.000000e+00
## Fireplaces	7.660440e-02	3.575543e-01	0.000000e+00	5.724116e-06
## GarageYrBlt	8.430360e-01	9.699819e-01	0.000000e+00	3.584860e-01
## GarageCars	1.217084e-01	1.351896e-01	0.000000e+00	4.817098e-02
## GarageArea	2.854434e-01	2.958337e-01	0.000000e+00	9.467301e-01
## WoodDeckSF	4.224189e-01	3.951420e-01	0.000000e+00	2.184093e-01
## OpenPorchSF	6.454024e-03	2.769527e-02	0.000000e+00	5.305810e-08
## EnclosedPorch	2.699986e-01	7.050056e-01	1.050805e-08	5.719454e-04
## X3SsnPorch.1	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
## ScreenPorch	3.753591e-01	6.830660e-01	3.408972e-06	5.241591e-01
## PoolArea	1.976284e-01	2.256080e-02	7.631787e-03	1.622043e-01
## MiscVal	8.041758e-01	8.514187e-01	4.446197e-01	8.407072e-01
## MoSold	NA	2.230277e-08	2.848930e-02	3.497548e-01
## YrSold	2.230277e-08	NA	1.547100e-01	6.406425e-01
## SalePriceLog	2.848930e-02	1.547100e-01	NA	7.446401e-09
## secondFloor	3.497548e-01	6.406425e-01	7.446401e-09	NA
## X1stFlrSF	2.309272e-01	6.034971e-01	0.000000e+00	0.000000e+00
## X2ndFlrSF	1.793059e-01	2.731188e-01	0.000000e+00	0.000000e+00
## X3SsnPorch	2.603891e-01	4.765433e-01	3.594763e-02	3.374853e-01
## model_2_resid	8.053840e-01	5.422568e-01	0.000000e+00	0.000000e+00
##	X1stFlrSF	X2ndFlrSF	X3SsnPorch	model_2_resid
## MSSubClass	0.000000e+00	0.000000e+00	0.09414953	0.000000e+00

```
## LotFrontage 0.000000e+00 5.433170e-03 0.01520968 5.930748e-07
## LotArea 0.000000e+00 5.144266e-02 0.43552735 3.552714e-15
## OverallQual 0.000000e+00 0.000000e+00 0.24615822 1.000000e+00
## OverallCond 3.126349e-08 2.690898e-01 0.33014731 4.646269e-04
## YearBuilt 0.000000e+00 6.939294e-01 0.23118088 0.000000e+00
## YearRemodAdd 0.000000e+00 7.777955e-08 0.08367071 0.000000e+00
## MasVnrArea 0.000000e+00 2.121014e-11 0.47420644 6.821566e-04
## BsmtFinSF1 0.000000e+00 1.455248e-07 0.31250322 0.000000e+00
## BsmtFinSF2 2.020320e-04 1.453587e-04 0.25207571 8.075468e-04
## BsmtUnfSF 0.000000e+00 8.645239e-01 0.42789582 2.339482e-04
## TotalBsmtSF 0.000000e+00 1.894085e-11 0.15337655 0.000000e+00
## LowQualFinSF 5.866501e-01 1.547482e-02 0.86973491 1.473062e-06
## GrLivArea 0.000000e+00 0.000000e+00 0.43058934 9.999999e-01
## BsmtFullBath 0.000000e+00 7.124057e-11 0.99676835 0.000000e+00
## BsmtHalfBath 9.404841e-01 3.623802e-01 0.17993527 5.117504e-02
## FullBath 0.000000e+00 0.000000e+00 0.17698330 1.085966e-03
## HalfBath 4.334880e-06 0.000000e+00 0.84943689 6.589154e-01
## BedroomAbvGr 1.041033e-06 0.000000e+00 0.34997682 3.687758e-02
## KitchenAbvGr 9.243536e-03 2.344382e-02 0.34756881 4.604133e-07
## TotRmsAbvGrd 0.000000e+00 0.000000e+00 0.79860693 6.671082e-01
## Fireplaces 0.000000e+00 6.394885e-14 0.66735300 1.316058e-11
## GarageYrBlt 0.000000e+00 8.506609e-03 0.38232493 0.000000e+00
## GarageCars 0.000000e+00 1.417977e-12 0.17198411 0.000000e+00
## GarageArea 0.000000e+00 1.112965e-07 0.18026968 0.000000e+00
## WoodDeckSF 0.000000e+00 4.216853e-04 0.21077845 4.440892e-15
## OpenPorchSF 4.440892e-16 8.881784e-16 0.82349481 1.418879e-01
## EnclosedPorch 1.258441e-02 1.784426e-02 0.15424164 8.078760e-10
## X3SsnPorch.1 3.206515e-02 3.523476e-01 0.00000000 2.865667e-02
## ScreenPorch 6.856867e-04 1.209302e-01 0.22997242 4.035941e-04
## PoolArea 4.578567e-07 1.832441e-03 0.76029001 4.043395e-02
## MiscVal 4.205511e-01 5.363168e-01 0.98921819 9.859260e-01
## MoSold 2.309272e-01 1.793059e-01 0.26038906 8.053840e-01
## YrSold 6.034971e-01 2.731188e-01 0.47654325 5.422568e-01
## SalePriceLog 0.000000e+00 0.000000e+00 0.03594763 0.000000e+00
## secondFloor 0.000000e+00 0.000000e+00 0.33748531 0.000000e+00
## X1stFlrSF NA 5.329071e-15 0.03206515 0.000000e+00
## X2ndFlrSF 5.329071e-15 NA 0.35234764 2.074785e-12
## X3SsnPorch 3.206515e-02 3.523476e-01 NA 2.865667e-02
## model_2_resid 0.000000e+00 2.074785e-12 0.02865667 NA
```

```
# It is a list of matrices with names "r","n" and "P"
c(typeof(res2),names(res2))
```

```
## [1] "list" "r" "n" "P"
```

```
# Identify the significant correlations
dim(res2$P)
```

```
## [1] 40 40
```

```
res2_price_p <- res2$P[response,]
res2_price_p
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond
## 0.000000e+00 5.930748e-07 3.552714e-15 1.000000e+00 4.646269e-04
## YearBuilt YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2
```

```
## 0.000000e+00 0.000000e+00 6.821566e-04 0.000000e+00 8.075468e-04
## BsmtUnfSF TotalBsmtSF LowQualFinSF GrLivArea BsmtFullBath
## 2.339482e-04 0.000000e+00 1.473062e-06 9.999999e-01 0.000000e+00
## BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr
## 5.117504e-02 1.085966e-03 6.589154e-01 3.687758e-02 4.604133e-07
## TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars GarageArea
## 6.671082e-01 1.316058e-11 0.000000e+00 0.000000e+00 0.000000e+00
## WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch.1 ScreenPorch
## 4.440892e-15 1.418879e-01 8.078760e-10 2.865667e-02 4.035941e-04
## PoolArea MiscVal MoSold YrSold SalePriceLog
## 4.043395e-02 9.859260e-01 8.053840e-01 5.422568e-01 0.000000e+00
## secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch model_2_resid
## 0.000000e+00 0.000000e+00 2.074785e-12 2.865667e-02 NA
```

```
correlation_significant<-res2_price_p <= sig & !is.na(res2_price_p)
```

Subset to only fields significantly correlated with there response

```
# extract the correlations of the response to other fields
res2_price_r <- res2$r[response,]
res2_price_r
```

```
## MSSubClass LotFrontage LotArea OverallQual OverallCond
## -2.336442e-01 1.434919e-01 2.039489e-01 -1.177894e-09 9.149482e-02
## YearBuilt YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2
## 3.249822e-01 2.523246e-01 8.903859e-02 3.022656e-01 8.758509e-02
## BsmtUnfSF TotalBsmtSF LowQualFinSF GrLivArea BsmtFullBath
## -9.614879e-02 2.496139e-01 -1.256187e-01 -2.512603e-09 3.055243e-01
## BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr
## 5.104452e-02 8.542474e-02 1.156177e-02 -5.462834e-02 -1.314974e-01
## TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars GarageArea
## -1.126605e-02 1.758655e-01 2.225138e-01 2.994809e-01 2.869592e-01
## WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch.1 ScreenPorch
## 2.031764e-01 3.845889e-02 -1.598811e-01 5.726927e-02 9.246721e-02
## PoolArea MiscVal MoSold YrSold SalePriceLog
## -5.363965e-02 4.620578e-04 -6.453657e-03 -1.596148e-02 5.099288e-01
## secondFloor X1stFlrSF X2ndFlrSF X3SsnPorch model_2_resid
## -2.520517e-01 2.219602e-01 -1.825749e-01 5.726927e-02 1.000000e+00
```

```
# Select the significant correlations
features_significant <- res2_price_r[correlation_significant]
features_significant
```

```
## MSSubClass LotFrontage LotArea OverallCond YearBuilt
## -0.23364425 0.14349188 0.20394890 0.09149482 0.32498217
## YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2 BsmtUnfSF
## 0.25232464 0.08903859 0.30226561 0.08758509 -0.09614879
## TotalBsmtSF LowQualFinSF BsmtFullBath KitchenAbvGr Fireplaces
## 0.24961391 -0.12561868 0.30552432 -0.13149744 0.17586552
## GarageYrBlt GarageCars GarageArea WoodDeckSF EnclosedPorch
## 0.22251378 0.29948092 0.28695917 0.20317642 -0.15988106
## ScreenPorch SalePriceLog secondFloor X1stFlrSF X2ndFlrSF
## 0.09246721 0.50992882 -0.25205165 0.22196017 -0.18257485
```

```
# Sort by correlation
features_significant <- features_significant[order(abs(features_significant),decreasing=TRUE)]
features_significant
```

```
## SalePriceLog      YearBuilt  BsmtFullBath    BsmtFinSF1    GarageCars
##      0.50992882      0.32498217      0.30552432      0.30226561      0.29948092
##      GarageArea  YearRemodAdd    secondFloor    TotalBsmtSF    MSSubClass
##      0.28695917      0.25232464     -0.25205165      0.24961391     -0.23364425
##      GarageYrBlt    X1stFlrSF        LotArea      WoodDeckSF      X2ndFlrSF
##      0.22251378      0.22196017      0.20394890      0.20317642     -0.18257485
##      Fireplaces  EnclosedPorch    LotFrontage    KitchenAbvGr    LowQualFinSF
##      0.17586552     -0.15988106      0.14349188     -0.13149744     -0.12561868
##      BsmtUnfSF    ScreenPorch    OverallCond    MasVnrArea      BsmtFinSF2
##     -0.09614879      0.09246721      0.09149482      0.08903859      0.08758509
```

```
names(features_significant)
```

```
## [1] "SalePriceLog" "YearBuilt"      "BsmtFullBath"  "BsmtFinSF1"
## [5] "GarageCars"   "GarageArea"     "YearRemodAdd"  "secondFloor"
## [9] "TotalBsmtSF"  "MSSubClass"     "GarageYrBlt"   "X1stFlrSF"
## [13] "LotArea"      "WoodDeckSF"     "X2ndFlrSF"     "Fireplaces"
## [17] "EnclosedPorch" "LotFrontage"    "KitchenAbvGr"  "LowQualFinSF"
## [21] "BsmtUnfSF"    "ScreenPorch"    "OverallCond"   "MasVnrArea"
## [25] "BsmtFinSF2"
```

```
remove_response <- !names(features_significant) %in% c('SalePriceLog',response)
remove_response
```

```
## [1] FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [12] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [23] TRUE TRUE TRUE
```

```
features_significant <- features_significant[remove_response]
features_significant
```

```
##      YearBuilt  BsmtFullBath    BsmtFinSF1    GarageCars    GarageArea
##      0.32498217      0.30552432      0.30226561      0.29948092      0.28695917
##      YearRemodAdd    secondFloor    TotalBsmtSF    MSSubClass    GarageYrBlt
##      0.25232464     -0.25205165      0.24961391     -0.23364425      0.22251378
##      X1stFlrSF        LotArea      WoodDeckSF      X2ndFlrSF      Fireplaces
##      0.22196017      0.20394890      0.20317642     -0.18257485      0.17586552
##      EnclosedPorch    LotFrontage    KitchenAbvGr    LowQualFinSF      BsmtUnfSF
##     -0.15988106      0.14349188     -0.13149744     -0.12561868     -0.09614879
##      ScreenPorch    OverallCond    MasVnrArea      BsmtFinSF2
##      0.09246721      0.09149482      0.08903859      0.08758509
```

```
#redo the t_selected
```

```
t_selected <- t_numeric %>% select(c(names(features_significant)[1:num_of_canadate_features],response))
names(t_selected)
```

```
## [1] "YearBuilt"      "BsmtFullBath"  "BsmtFinSF1"    "GarageCars"
## [5] "GarageArea"     "YearRemodAdd"  "secondFloor"    "TotalBsmtSF"
## [9] "MSSubClass"     "GarageYrBlt"   "X1stFlrSF"     "LotArea"
## [13] "WoodDeckSF"     "X2ndFlrSF"     "Fireplaces"     "model_2_resid"
```

```
# Note that the correlation for OverallQual is zero. It was the field used last time, so the correlati
#should be zero
```

```
features_significant
```

```
##      YearBuilt  BsmtFullBath    BsmtFinSF1    GarageCars    GarageArea
##      0.32498217      0.30552432      0.30226561      0.29948092      0.28695917
```

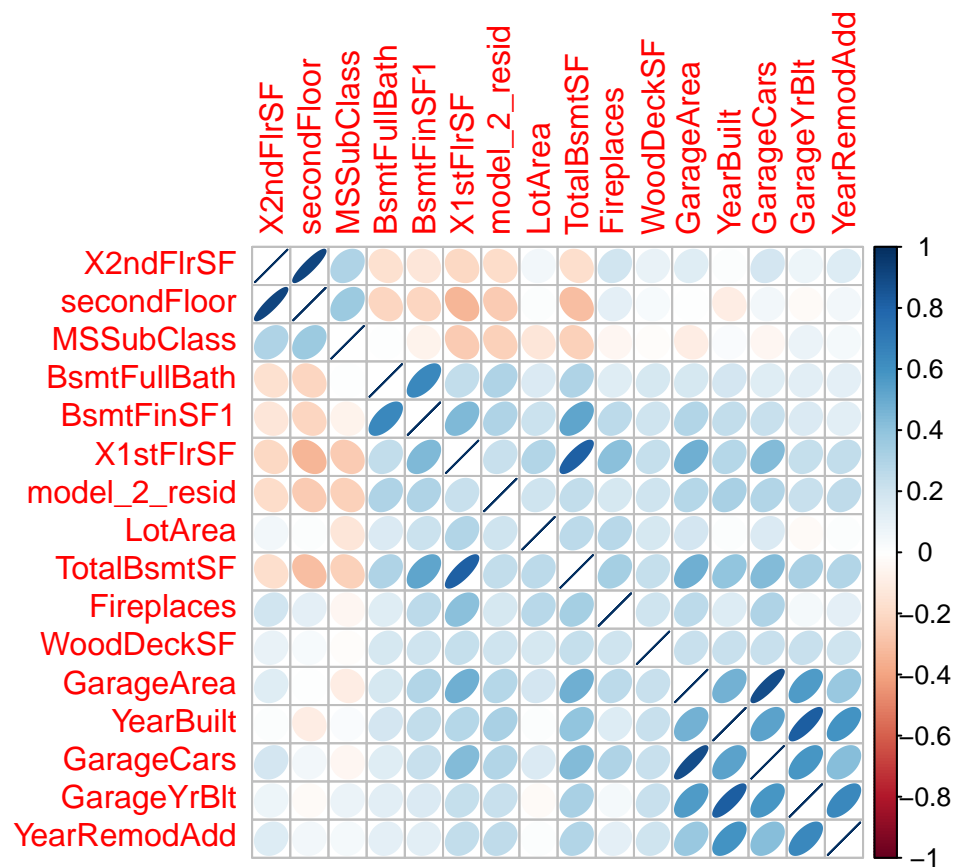
```
## YearRemodAdd secondFloor TotalBsmtSF MSSubClass GarageYrBlt
## 0.25232464 -0.25205165 0.24961391 -0.23364425 0.22251378
## X1stFlrSF LotArea WoodDeckSF X2ndFlrSF Fireplaces
## 0.22196017 0.20394890 0.20317642 -0.18257485 0.17586552
## EnclosedPorch LotFrontage KitchenAbvGr LowQualFinSF BsmtUnfSF
## -0.15988106 0.14349188 -0.13149744 -0.12561868 -0.09614879
## ScreenPorch OverallCond MasVnrArea BsmtFinSF2
## 0.09246721 0.09149482 0.08903859 0.08758509
```

```
names(t_selected)
```

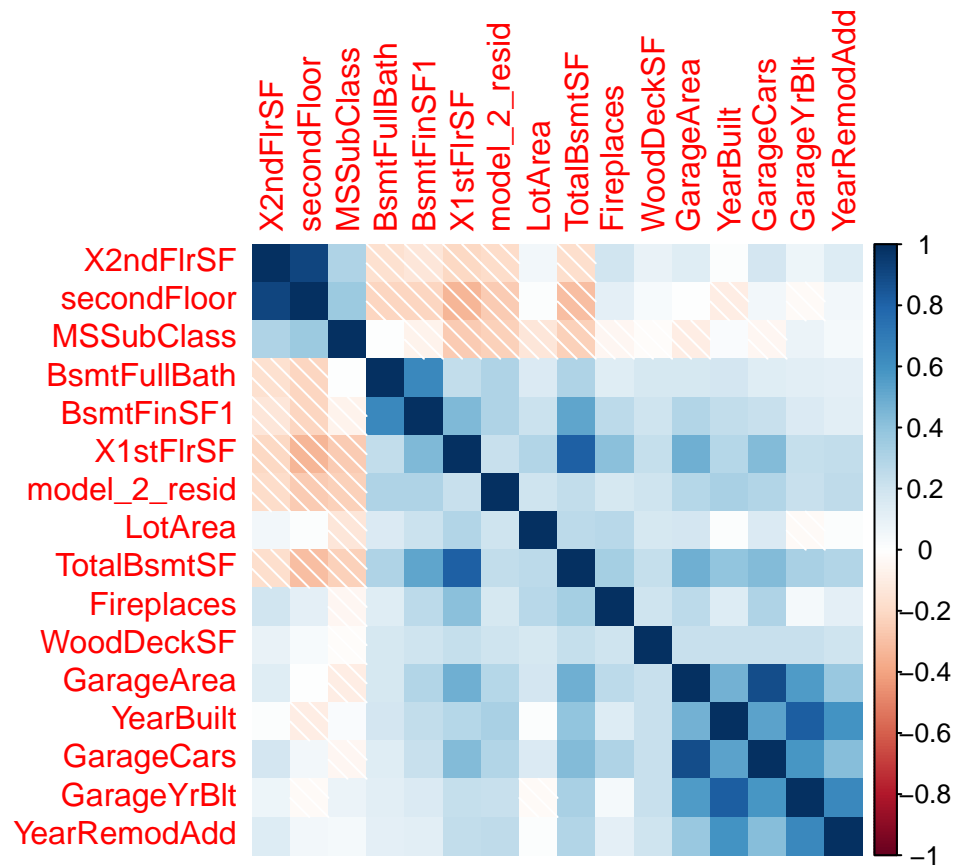
```
## [1] "YearBuilt" "BsmtFullBath" "BsmtFinSF1" "GarageCars"
## [5] "GarageArea" "YearRemodAdd" "secondFloor" "TotalBsmtSF"
## [9] "MSSubClass" "GarageYrBlt" "X1stFlrSF" "LotArea"
## [13] "WoodDeckSF" "X2ndFlrSF" "Fireplaces" "model_2_resid"
```

A few versions of the correlation matrix

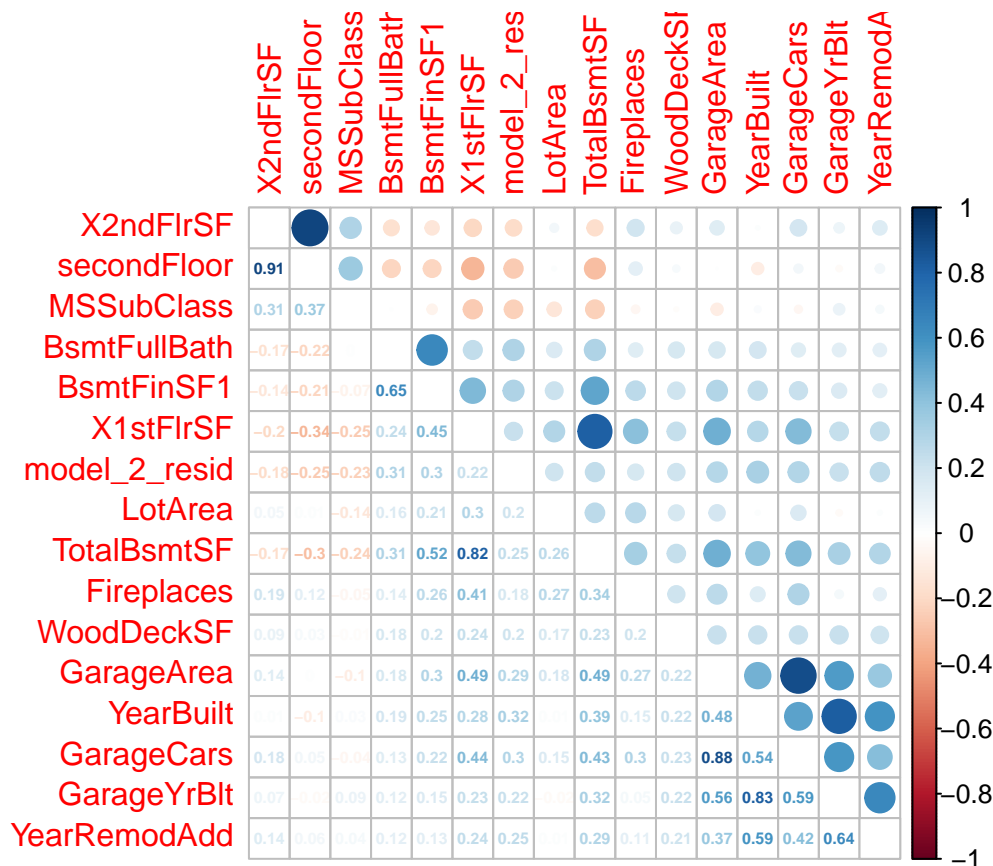
```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='ellipse',order="AOE")
```



```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='shade',order="AOE")
```

```
t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot.mixed(number.cex = 0.5, order = "AOE", tl.pos = "br")
```



```
#t_selected %>% cor(use='pairwise.complete.obs') %>% corrplot(method='circle',method='ellipse')
```

Here is the top feature to try

```
#feature_list <-list()
selected<-names(features_significant)[1]
selected

## [1] "YearBuilt"
selected_label<-label(t_numeric[selected])
selected_label
```

```
##
##      YearBuilt
## "Original construction date"
feature_list[[model_num]] <- selected
str(feature_list)
```

```
## List of 3
## $ : chr "OverallQual"
## $ : chr "GrLivArea"
## $ : chr "YearBuilt"
```

Calculate a model using this variable

```
names(t_numeric)

## [1] "MSSubClass" "LotFrontage" "LotArea" "OverallQual"
## [5] "OverallCond" "YearBuilt" "YearRemodAdd" "MasVnrArea"
```

```
## [9] "BsmtFinSF1"      "BsmtFinSF2"      "BsmtUnfSF"      "TotalBsmtSF"
## [13] "LowQualFinSF"    "GrLivArea"        "BsmtFullBath"    "BsmtHalfBath"
## [17] "FullBath"        "HalfBath"         "BedroomAbvGr"    "KitchenAbvGr"
## [21] "TotRmsAbvGrd"    "Fireplaces"       "GarageYrBlt"     "GarageCars"
## [25] "GarageArea"       "WoodDeckSF"       "OpenPorchSF"     "EnclosedPorch"
## [29] "X3SsnPorch.1"    "ScreenPorch"      "PoolArea"        "MiscVal"
## [33] "MoSold"          "YrSold"           "SalePriceLog"    "secondFloor"
## [37] "X1stFlrSF"       "X2ndFlrSF"       "X3SsnPorch"      "model_2_resid"
```

```
feature_list
```

```
## [[1]]
## [1] "OverallQual"
##
## [[2]]
## [1] "GrLivArea"
##
## [[3]]
## [1] "YearBuilt"
```

```
my_formula <- paste0('SalePriceLog ~ ',paste(feature_list,collapse = '+ ', sep=" "))
str(my_formula)
```

```
## chr "SalePriceLog ~ OverallQual+ GrLivArea+ YearBuilt"
```

```
model_3 <- t_numeric %>% lm(my_formula,data=.)
```

```
t_numeric<- model_3 %>% augment(type.predict = 'response',type.residuals = 'response',data=t_numeric) %>%
```

```
  mutate(model_3_resid = .resid) %>%
```

```
  mutate(influential = .cooks > 4/length(SalePriceLog))
```

```
head(t_numeric)
```

```
##   MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 1         60         65    8450           7           5      2003
## 2         20         80    9600           6           8      1976
## 3         60         68   11250           7           5      2001
## 4         70         60    9550           7           5      1915
## 5         60         84   14260           8           5      2000
## 6         50         85   14115           5           5      1993
##   YearRemodAdd MasVnrArea BsmtFinSF1 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1         2003         196         706          0        150         856
## 2         1976          0         978          0        284        1262
## 3         2002         162         486          0        434         920
## 4         1970          0         216          0        540         756
## 5         2000         350         655          0        490        1145
## 6         1995          0         732          0         64         796
##   LowQualFinSF GrLivArea BsmtFullBath BsmtHalfBath FullBath HalfBath
## 1             0     1710           1             0          2          1
## 2             0     1262           0             1          2          0
## 3             0     1786           1             0          2          1
## 4             0     1717           1             0          1          0
## 5             0     2198           1             0          2          1
## 6             0     1362           1             0          1          1
##   BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageYrBlt GarageCars
## 1             3             1             8             0        2003          2
## 2             3             1             6             1        1976          2
## 3             3             1             6             1        2001          2
```

```

## 4      3      1      7      1      1998      3
## 5      4      1      9      1      2000      3
## 6      1      1      5      0      1993      2
##   GarageArea WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch.1 ScreenPorch
## 1      548          0          61          0          0          0
## 2      460        298          0          0          0          0
## 3      608          0          42          0          0          0
## 4      642          0          35         272          0          0
## 5      836        192          84          0          0          0
## 6      480         40          30          0        320          0
##   PoolArea MiscVal MoSold YrSold SalePriceLog secondFloor X1stFlrSF
## 1          0          0      2   2008    12.24769         TRUE      856
## 2          0          0      5   2007    12.10901        FALSE     1262
## 3          0          0      9   2008    12.31717         TRUE      920
## 4          0          0      2   2006    11.84940         TRUE      961
## 5          0          0     12   2008    12.42922         TRUE     1145
## 6          0        700     10   2009    11.87060         TRUE      796
##   X2ndFlrSF X3SsnPorch model_2_resid .fitted .se.fit .resid
## 1      854          0    0.01318822 12.30404 0.007125142 -0.056347679
## 2          0          0    0.16701430 11.95257 0.005604198  0.156440206
## 3      866          0    0.06338382 12.31972 0.007047358 -0.002553542
## 4      756          0   -0.38688389 12.00572 0.014810167 -0.156318286
## 5     1053          0   -0.10794504 12.56495 0.008947297 -0.135733355
## 6      566        320    0.08211702 11.91358 0.009726502 -0.042975368
##   .hat .sigma .cooksd .std.resid model_3_resid
## 1 0.0014614888 0.1864367 3.349395e-05 -0.30255046 -0.056347679
## 2 0.0009041394 0.1863974 1.595388e-04  0.83974803  0.156440206
## 3 0.0014297532 0.1864425 6.728810e-08 -0.01371065 -0.002553542
## 4 0.0063143453 0.1863972 1.124600e-03 -0.84137474 -0.156318286
## 5 0.0023045814 0.1864085 3.069854e-04 -0.72910788 -0.135733355
## 6 0.0027234648 0.1864391 3.639806e-05 -0.23089578 -0.042975368
##   influential
## 1      FALSE
## 2      FALSE
## 3      FALSE
## 4      FALSE
## 5      FALSE
## 6      FALSE

```

Check the fit using the Kaggle criteria

```
rmse(model_2,t_numeric)
```

```
## [1] 0.2036223
```

```
t_numeric[t_numeric$influential,]
```

```

##   MSSubClass LotFrontage LotArea OverallQual OverallCond YearBuilt
## 31          70          50    8500          4          4     1920
## 54          20          68   50271          9          5     1981
## 93          30          80   13360          5          7     1921
## 113         60          77    9965          7          5     2007
## 126        190          60    6780          6          8     1935
## 146        160          24    2522          6          5     2004
## 156         50          60    9600          6          5     1924
## 179         20          63   17423          9          5     2008

```

## 186	75	90	22950	10	9	1892
## 194	160	24	2522	7	5	2004
## 198	75	174	25419	8	4	1918
## 199	75	92	5520	6	6	1912
## 219	50	NA	15660	7	9	1939
## 243	50	63	5000	5	4	1900
## 305	75	87	18386	7	9	1880
## 308	50	NA	7920	6	7	1920
## 314	20	150	215245	7	5	1965
## 324	20	49	5820	3	8	1955
## 384	45	60	9000	6	3	1928
## 399	30	60	8967	5	2	1920
## 411	20	68	9571	5	3	1958
## 431	160	21	1680	6	5	1971
## 432	50	60	5586	6	7	1920
## 441	20	105	15431	10	5	2008
## 458	20	NA	53227	4	6	1954
## 463	20	60	8281	5	5	1965
## 474	20	110	14977	8	5	2006
## 496	30	60	7879	4	5	1920
## 524	60	130	40094	10	5	2007
## 534	20	50	5000	1	3	1946
## 547	50	70	8737	6	7	1923
## 608	20	78	7800	5	8	1948
## 633	20	85	11900	7	5	1977
## 636	190	60	10896	6	7	1914
## 667	60	NA	18450	6	5	1965
## 689	20	60	8089	8	6	2007
## 706	190	70	5600	4	5	1930
## 711	30	56	4130	3	6	1935
## 717	70	60	10800	7	8	1890
## 748	70	65	11700	7	7	1880
## 770	60	47	53504	8	5	2003
## 804	60	107	13891	9	5	2008
## 808	70	144	21384	5	6	1923
## 810	75	90	8100	5	5	1898
## 813	20	66	8712	5	5	1952
## 842	70	60	10440	5	8	1904
## 849	50	75	45600	6	8	1908
## 875	50	52	5720	5	6	1941
## 898	90	64	7018	5	5	1979
## 899	20	100	12919	9	5	2009
## 916	160	21	2001	4	5	1970
## 917	20	50	9000	2	3	1949
## 969	50	50	5925	3	6	1910
## 1021	20	60	7024	4	5	2005
## 1032	75	102	15863	7	3	1920
## 1047	60	85	16056	9	5	2005
## 1063	190	85	13600	5	5	1900
## 1069	160	42	3964	6	4	1973
## 1133	70	90	9900	6	4	1880
## 1182	120	64	5587	8	5	2008
## 1217	90	68	8930	6	5	1978
## 1244	20	107	13891	10	5	2006

## 1245	70	NA	11435	8	7	1929
## 1267	190	60	10120	7	4	1910
## 1293	70	60	6600	5	4	1892
## 1299	60	313	63887	10	5	2008
## 1305	160	32	3363	7	5	2004
## 1311	20	100	17500	7	8	1959
## 1325	20	75	9986	8	5	2006
## 1326	30	40	3636	4	4	1922
## 1329	50	60	10440	6	7	1920
## 1338	30	153	4118	4	4	1941
## 1350	70	50	5250	8	5	1872
## 1379	160	21	1953	6	5	1973
## 1388	50	60	8520	6	7	1916
## 1454	20	90	17217	5	5	2006
##	YearRemodAdd	MasVnrArea	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF
## 31	1950	0	0	0	649	649
## 54	1987	0	1810	0	32	1842
## 93	2006	0	713	0	163	876
## 113	2007	220	984	0	280	1264
## 126	1982	0	490	0	30	520
## 146	2006	50	0	0	970	970
## 156	1950	0	0	0	572	572
## 179	2009	748	1904	0	312	2216
## 186	1993	0	0	0	1107	1107
## 194	2004	50	0	0	970	970
## 198	1990	0	1036	184	140	1360
## 199	1950	0	0	0	755	755
## 219	2006	312	341	0	457	798
## 243	1950	0	0	0	540	540
## 305	2002	0	0	0	1470	1470
## 308	1950	0	0	0	319	319
## 314	1965	0	1236	820	80	2136
## 324	2005	0	256	0	906	1162
## 384	1950	0	0	0	784	784
## 399	1950	0	0	0	961	961
## 411	1958	0	0	0	1276	1276
## 431	1971	232	387	0	96	483
## 432	1998	0	0	0	901	901
## 441	2008	200	1767	539	788	3094
## 458	1994	0	1116	0	248	1364
## 463	1965	0	553	311	0	864
## 474	2007	304	1350	0	626	1976
## 496	1950	0	495	0	225	720
## 524	2008	762	2260	0	878	3138
## 534	1950	0	0	0	0	0
## 547	1950	0	300	0	765	1065
## 608	2002	0	603	0	293	896
## 633	1977	209	822	0	564	1386
## 636	1995	0	256	0	1184	1440
## 667	1979	113	187	723	111	1021
## 689	2007	0	945	0	474	1419
## 706	1950	0	0	0	0	0
## 711	2003	0	0	0	270	270
## 717	1998	0	0	0	718	718

## 748	2003	0	0	0	1240	1240
## 770	2003	603	1416	0	234	1650
## 804	2009	424	0	0	1734	1734
## 808	2004	0	1309	0	15	1324
## 810	1965	0	0	0	849	849
## 813	1952	0	0	0	540	540
## 842	2002	0	0	0	650	650
## 849	1997	0	0	0	907	907
## 875	1950	0	0	0	676	676
## 898	1979	0	0	0	0	0
## 899	2010	760	2188	0	142	2330
## 916	1970	80	0	0	546	546
## 917	1950	0	50	0	430	480
## 969	1950	0	0	0	600	600
## 1021	2005	0	1024	0	108	1132
## 1032	1970	0	523	0	301	824
## 1047	2006	208	240	0	1752	1992
## 1063	1950	0	0	0	662	662
## 1069	1973	0	837	0	105	942
## 1133	1950	0	0	0	1008	1008
## 1182	2008	186	1480	0	120	1600
## 1217	1978	0	0	0	0	0
## 1244	2006	NA	1386	0	690	2076
## 1245	1950	0	0	0	792	792
## 1267	1950	0	0	0	925	925
## 1293	1965	0	0	0	994	994
## 1299	2008	796	5644	0	466	6110
## 1305	2004	117	0	0	976	976
## 1311	2002	0	1406	0	496	1902
## 1325	2007	428	0	0	1795	1795
## 1326	1950	0	0	0	796	796
## 1329	1950	0	493	0	1017	1510
## 1338	1950	0	0	0	693	693
## 1350	1987	0	259	0	425	684
## 1379	1973	408	309	0	174	483
## 1388	1950	0	168	546	0	714
## 1454	2006	0	0	0	1140	1140
##	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	FullBath	HalfBath
## 31	0	1317	0	0	1	0
## 54	0	1842	2	0	0	1
## 93	0	964	1	0	1	0
## 113	0	2696	1	0	2	1
## 126	234	754	1	0	1	0
## 146	0	1709	0	0	2	0
## 156	0	1096	0	0	1	0
## 179	0	2234	1	0	2	0
## 186	572	3608	0	0	2	1
## 194	0	1709	0	0	2	0
## 198	392	3112	1	1	2	0
## 199	371	2229	0	0	1	0
## 219	0	1954	0	1	1	1
## 243	0	1440	0	0	1	0
## 305	0	3493	0	0	3	0
## 308	0	1406	0	0	1	0

## 314	0	2036	2	0	2	0
## 324	0	1163	1	0	1	0
## 384	0	784	0	0	1	0
## 399	0	1077	0	0	1	0
## 411	0	1276	0	0	1	0
## 431	0	987	0	0	1	1
## 432	0	1198	0	0	1	0
## 441	0	2402	1	0	2	0
## 458	0	1663	1	0	1	0
## 463	0	864	0	0	1	0
## 474	0	1976	1	0	2	0
## 496	0	720	0	0	1	0
## 524	0	4676	1	0	3	1
## 534	0	334	0	0	1	0
## 547	0	1635	0	0	1	1
## 608	0	2008	1	0	3	0
## 633	0	1411	0	0	2	0
## 636	515	3395	0	0	2	0
## 667	0	2380	0	0	2	1
## 689	0	1419	1	0	2	0
## 706	0	1092	0	0	2	0
## 711	0	729	0	0	1	0
## 717	0	2554	0	0	1	1
## 748	0	2640	0	0	1	1
## 770	0	3279	1	0	3	1
## 804	0	2822	0	0	3	1
## 808	0	1576	2	0	1	1
## 810	0	2138	0	0	2	0
## 813	0	1044	0	0	1	0
## 842	0	1539	0	0	2	0
## 849	0	2358	0	0	3	0
## 875	0	1131	0	0	1	1
## 898	0	2240	0	0	2	0
## 899	0	2364	1	0	2	1
## 916	0	1092	0	0	1	1
## 917	0	480	1	0	0	0
## 969	0	968	0	0	1	0
## 1021	0	1132	1	0	1	1
## 1032	397	3082	1	0	2	1
## 1047	0	2868	0	0	3	1
## 1063	0	2337	0	0	2	0
## 1069	0	2521	1	0	2	1
## 1133	0	2210	0	0	2	0
## 1182	0	1652	1	1	2	0
## 1217	0	1902	0	0	2	0
## 1244	0	2076	1	0	2	1
## 1245	0	1517	0	0	1	0
## 1267	0	1889	0	0	1	1
## 1293	0	2372	0	0	2	0
## 1299	0	5642	2	0	2	1
## 1305	0	1708	0	0	2	0
## 1311	0	1902	1	0	2	0
## 1325	0	1795	0	0	2	0
## 1326	0	796	0	0	1	0

## 1329	0	2792	0	0	2	0
## 1338	0	693	0	0	1	0
## 1350	205	2358	0	0	2	0
## 1379	0	987	0	0	1	1
## 1388	0	2526	0	0	2	0
## 1454	0	1140	0	0	1	0
##	BedroomAbvGr	KitchenAbvGr	TotRmsAbvGrd	Fireplaces	GarageYrBltd	
## 31	3	1	6	0	1920	
## 54	0	1	5	1	1981	
## 93	2	1	5	0	1921	
## 113	4	1	10	1	2007	
## 126	2	1	5	0	NA	
## 146	3	1	7	0	2004	
## 156	2	1	5	0	NA	
## 179	1	1	9	1	2009	
## 186	4	1	12	2	1993	
## 194	3	1	7	0	2004	
## 198	4	1	8	1	1918	
## 199	5	1	8	0	NA	
## 219	3	1	8	2	1939	
## 243	3	1	6	0	1940	
## 305	3	1	10	1	2003	
## 308	3	1	6	0	NA	
## 314	3	1	8	2	1965	
## 324	3	1	6	0	1955	
## 384	2	1	5	0	1950	
## 399	2	1	6	0	1920	
## 411	3	1	5	0	1958	
## 431	2	1	4	0	1971	
## 432	4	1	7	0	NA	
## 441	2	1	10	2	2008	
## 458	2	1	6	2	1954	
## 463	3	1	5	1	1965	
## 474	2	1	7	1	2006	
## 496	2	1	4	0	NA	
## 524	3	1	11	1	2007	
## 534	1	1	2	0	NA	
## 547	3	1	6	1	1950	
## 608	3	1	8	0	1948	
## 633	3	1	6	1	1977	
## 636	8	2	14	0	NA	
## 667	3	1	7	1	1965	
## 689	2	1	7	1	2007	
## 706	3	2	7	0	NA	
## 711	2	1	5	0	NA	
## 717	3	1	8	0	1996	
## 748	4	1	8	1	1950	
## 770	4	1	12	1	2003	
## 804	4	1	12	1	2009	
## 808	3	1	6	1	1923	
## 810	2	3	11	0	1910	
## 813	2	1	4	0	1952	
## 842	3	1	8	1	1983	
## 849	5	1	10	1	1908	

## 875	3	1	5	0	1941
## 898	6	2	12	0	1979
## 899	2	1	11	2	2009
## 916	3	1	6	0	1970
## 917	1	1	4	0	1958
## 969	2	1	6	0	NA
## 1021	2	1	5	0	2005
## 1032	5	1	12	2	1970
## 1047	4	1	11	1	2005
## 1063	5	2	10	0	1945
## 1069	5	1	10	1	1973
## 1133	5	1	8	0	1930
## 1182	2	1	5	1	2008
## 1217	4	2	8	0	1978
## 1244	2	1	7	1	2006
## 1245	3	1	7	2	1931
## 1267	4	2	9	1	1960
## 1293	4	2	11	0	1985
## 1299	3	1	12	3	2008
## 1305	3	1	7	0	2004
## 1311	3	1	7	2	1959
## 1325	2	1	7	1	2007
## 1326	2	1	5	0	NA
## 1329	5	1	8	2	1920
## 1338	2	1	4	0	NA
## 1350	4	1	8	0	NA
## 1379	2	1	5	0	1973
## 1388	5	1	10	1	1916
## 1454	3	1	6	0	NA
##	GarageCars	GarageArea	WoodDeckSF	OpenPorchSF	EnclosedPorch
## 31	1	250	0	54	172
## 54	3	894	857	72	0
## 93	2	432	0	0	44
## 113	3	792	120	184	0
## 126	0	0	53	0	0
## 146	2	380	0	40	0
## 156	0	0	0	8	128
## 179	3	1166	0	60	0
## 186	3	840	0	260	0
## 194	2	380	0	40	0
## 198	2	795	0	16	552
## 199	0	0	0	198	30
## 219	2	431	0	119	150
## 243	1	352	0	0	77
## 305	3	870	302	0	0
## 308	0	0	0	144	0
## 314	2	513	0	0	0
## 324	1	220	142	98	0
## 384	2	360	0	0	91
## 399	1	338	0	0	0
## 411	1	350	0	0	0
## 431	1	264	0	0	0
## 432	0	0	0	98	0
## 441	3	672	0	72	0

## 458	2	529	224	137	0
## 463	1	360	0	0	236
## 474	3	908	250	63	0
## 496	0	0	0	523	115
## 524	3	884	208	406	0
## 534	0	0	0	0	0
## 547	2	440	0	38	0
## 608	1	230	103	0	0
## 633	2	544	192	0	0
## 636	0	0	0	110	0
## 667	2	596	0	265	0
## 689	2	567	140	0	0
## 706	0	0	0	0	0
## 711	0	0	0	0	0
## 717	2	704	0	48	143
## 748	4	864	181	0	386
## 770	3	841	503	36	0
## 804	3	1020	52	170	0
## 808	2	528	0	312	0
## 810	2	360	40	156	0
## 813	2	504	0	0	0
## 842	2	686	70	78	68
## 849	2	360	486	40	0
## 875	1	200	26	0	0
## 898	2	528	154	0	0
## 899	3	820	0	67	0
## 916	1	286	0	0	0
## 917	1	308	0	0	0
## 969	0	0	0	0	0
## 1021	2	451	252	64	0
## 1032	2	672	136	63	0
## 1047	3	716	214	108	0
## 1063	2	560	0	57	0
## 1069	2	576	728	20	0
## 1133	1	205	0	48	0
## 1182	2	482	162	53	0
## 1217	2	539	0	0	0
## 1244	3	850	216	229	0
## 1245	2	400	0	0	0
## 1267	1	308	0	0	264
## 1293	1	432	0	287	0
## 1299	2	1418	214	292	0
## 1305	2	380	0	40	0
## 1311	2	567	0	207	162
## 1325	3	895	0	49	0
## 1326	0	0	0	0	100
## 1329	2	520	0	547	0
## 1338	0	0	0	20	0
## 1350	0	0	0	54	20
## 1379	1	264	72	0	0
## 1388	1	216	88	15	0
## 1454	0	0	36	56	0
##	X3SsnPorch.1 ScreenPorch PoolArea MiscVal MoSold YrSold SalePriceLog				
## 31	0	0	0	0	7 2008 10.59663

## 54	0	0	0	0	11	2006	12.86100
## 93	0	0	0	0	8	2009	12.00457
## 113	0	168	0	0	10	2007	12.85832
## 126	0	0	0	0	6	2006	11.34451
## 146	0	0	0	0	4	2006	11.77529
## 156	0	0	0	0	4	2008	11.27720
## 179	0	0	0	0	7	2009	13.12603
## 186	0	410	0	0	6	2006	13.07107
## 194	0	0	0	0	5	2006	11.77529
## 198	0	0	512	0	3	2006	12.36734
## 199	0	0	0	0	7	2009	11.55215
## 219	0	0	0	0	5	2008	12.64915
## 243	0	0	0	0	4	2006	11.27720
## 305	0	0	0	0	5	2008	12.59473
## 308	0	0	0	0	3	2008	11.40199
## 314	0	0	0	0	6	2009	12.83468
## 324	0	0	0	0	7	2006	11.74543
## 384	0	0	0	0	10	2009	11.23849
## 399	0	0	0	0	11	2007	11.11245
## 411	0	0	0	0	6	2009	11.00210
## 431	0	0	0	0	7	2008	11.35510
## 432	0	0	0	0	9	2008	11.28853
## 441	0	170	0	0	4	2009	13.22672
## 458	0	0	0	0	3	2008	12.45293
## 463	0	0	0	0	12	2009	11.04105
## 474	0	0	0	0	7	2007	12.99453
## 496	0	0	0	0	11	2009	10.46024
## 524	0	0	0	0	10	2007	12.12676
## 534	0	0	0	0	1	2007	10.57898
## 547	144	0	0	0	5	2007	12.25486
## 608	0	0	0	0	8	2006	12.32386
## 633	0	0	0	0	4	2009	11.32055
## 636	0	0	0	0	3	2007	12.20607
## 667	0	0	0	0	8	2007	11.76757
## 689	0	0	0	0	10	2007	12.87902
## 706	0	0	0	3500	7	2010	10.91509
## 711	0	0	0	0	7	2008	10.85900
## 717	0	0	0	0	7	2007	11.97980
## 748	0	0	0	0	5	2009	12.49117
## 770	0	210	0	0	6	2010	13.19561
## 804	0	192	0	0	1	2009	13.27583
## 808	0	0	0	0	5	2009	12.31717
## 810	0	0	0	0	11	2009	11.57119
## 813	0	0	0	54	6	2010	10.93298
## 842	0	0	0	0	6	2008	11.96718
## 849	0	175	0	0	9	2008	12.38839
## 875	0	0	0	0	8	2009	11.10496
## 898	0	0	0	0	6	2009	11.87027
## 899	0	0	0	0	3	2010	13.32393
## 916	0	0	0	0	1	2007	11.22524
## 917	0	0	0	0	10	2006	10.47195
## 969	0	0	0	0	5	2009	10.54271
## 1021	0	0	0	0	6	2008	12.07824
## 1032	0	0	0	0	8	2009	12.19096

##	1047	0	0	0	0	7	2006	13.22957
##	1063	0	0	0	0	9	2007	11.40756
##	1069	0	0	0	0	6	2006	11.92768
##	1133	0	0	0	0	5	2007	11.67419
##	1182	153	0	0	0	11	2008	12.88029
##	1217	0	0	0	0	4	2010	11.62625
##	1244	0	0	0	0	9	2006	13.04979
##	1245	0	0	0	0	6	2006	12.34583
##	1267	0	0	0	0	1	2007	11.71178
##	1293	0	0	0	0	12	2009	11.58525
##	1299	0	0	480	0	1	2008	11.98293
##	1305	0	0	0	0	4	2006	11.77529
##	1311	0	0	0	0	5	2010	12.72189
##	1325	0	0	0	0	2	2007	11.89819
##	1326	0	0	0	0	1	2008	10.91509
##	1329	0	480	0	1150	6	2008	12.45293
##	1338	0	0	0	0	3	2006	10.86857
##	1350	0	0	0	0	12	2008	11.71178
##	1379	0	0	0	0	6	2006	11.32660
##	1388	0	0	0	0	8	2007	11.82041
##	1454	0	0	0	0	7	2006	11.34451
##		secondFloor	X1stFlrSF	X2ndFlrSF	X3SsnPorch	model_2_resid	.fitted	
##	31	TRUE	649	668	0	-1.00155634	11.52442	
##	54	FALSE	1842	0	0	0.23525594	12.52131	
##	93	FALSE	964	0	0	0.31703483	11.54994	
##	113	TRUE	1282	1414	0	0.37372277	12.60968	
##	126	FALSE	520	0	0	-0.46863978	11.66218	
##	146	TRUE	970	739	0	-0.28008480	12.18052	
##	156	TRUE	572	524	0	-0.62268896	11.72591	
##	179	FALSE	2234	0	0	0.40086037	12.72956	
##	186	TRUE	1518	1518	0	-0.18148198	12.86709	
##	194	TRUE	970	739	0	-0.45896273	12.30716	
##	198	TRUE	1360	1360	0	-0.40164907	12.55569	
##	199	TRUE	929	929	0	-0.63512210	12.02045	
##	219	TRUE	1137	817	0	0.35275992	12.15782	
##	243	TRUE	889	551	0	-0.53106382	11.61921	
##	305	TRUE	1675	1818	0	-0.09201883	12.41216	
##	308	TRUE	1035	371	0	-0.57652715	11.80405	
##	314	FALSE	2036	0	0	0.51748797	12.27086	
##	324	FALSE	1163	0	0	0.36517281	11.47165	
##	384	FALSE	784	0	0	-0.58226722	11.64717	
##	399	FALSE	1077	0	0	-0.60374707	11.57999	
##	411	FALSE	1276	0	0	-0.76456985	11.76863	
##	431	TRUE	483	504	0	-0.51714375	11.85407	
##	432	TRUE	1088	110	0	-0.63723243	11.74246	
##	441	FALSE	2402	0	0	0.28006337	12.90595	
##	458	FALSE	1663	0	0	0.76698158	11.74294	
##	463	FALSE	864	0	0	-0.62112118	11.67052	
##	474	FALSE	1976	0	0	0.51367725	12.51969	
##	496	FALSE	720	0	0	-0.98652480	11.34763	
##	524	TRUE	3138	1538	0	-1.39668268	13.57592	
##	534	FALSE	334	0	0	-0.23324758	10.94216	
##	547	TRUE	915	720	144	0.21825778	11.88210	
##	608	TRUE	1112	896	0	0.37152018	11.95126	

## 633	FALSE	1411	0	0	-0.83811362	12.12675
## 636	TRUE	1440	1440	0	-0.27694199	12.37256
## 667	TRUE	1465	915	0	-0.45800050	12.24609
## 689	FALSE	1419	0	0	0.53944286	12.35816
## 706	TRUE	372	720	0	-0.62603320	11.49192
## 711	FALSE	729	0	0	-0.41117276	11.27486
## 717	TRUE	1576	978	0	-0.46878059	12.16823
## 748	TRUE	1320	1320	0	0.02077965	12.15956
## 770	TRUE	1690	1589	0	0.38426580	12.89530
## 804	TRUE	1734	1088	0	0.40151588	12.90368
## 808	TRUE	1072	504	0	0.47440446	11.73799
## 810	TRUE	1075	1063	0	-0.41411457	11.81907
## 813	FALSE	1044	0	0	-0.77484283	11.67945
## 842	TRUE	958	581	0	0.13380325	11.66218
## 849	TRUE	1307	1051	0	0.16840613	12.04500
## 875	TRUE	676	455	0	-0.62493440	11.66766
## 898	TRUE	1120	1120	0	-0.14090922	12.12578
## 899	FALSE	2364	0	0	0.56578324	12.77147
## 916	TRUE	546	546	0	-0.31587827	11.62847
## 917	FALSE	480	0	0	-0.55618720	11.12228
## 969	TRUE	600	368	0	-0.78808576	11.26029
## 1021	FALSE	1132	0	0	0.52697194	11.75979
## 1032	TRUE	1687	998	0	-0.39154367	12.42699
## 1047	TRUE	1992	876	0	0.34358881	12.90706
## 1063	TRUE	1422	915	0	-0.62821871	11.88483
## 1069	TRUE	1291	1230	0	-0.33365106	12.31515
## 1133	TRUE	1178	1032	0	-0.50825547	11.90559
## 1182	FALSE	1652	0	153	0.48161901	12.43057
## 1217	TRUE	1318	584	0	-0.47807325	12.14892
## 1244	FALSE	2076	0	0	0.18581989	12.80258
## 1245	TRUE	792	725	0	-0.01859657	12.12092
## 1267	TRUE	964	925	0	-0.56813166	12.03958
## 1293	TRUE	1378	994	0	-0.45941500	11.86789
## 1299	TRUE	4692	950	0	-1.78553043	13.86539
## 1305	TRUE	976	732	0	-0.45870908	12.30686
## 1311	FALSE	1902	0	0	0.43868048	12.21070
## 1325	FALSE	1795	0	0	-0.53675572	12.46609
## 1326	FALSE	796	0	0	-0.55095522	11.37696
## 1329	TRUE	1584	1208	0	0.12286411	12.21448
## 1338	FALSE	693	0	0	-0.57135013	11.41132
## 1350	TRUE	938	1215	0	-0.86596760	12.17539
## 1379	TRUE	483	504	0	-0.54564925	11.86090
## 1388	TRUE	1664	862	0	-0.44218973	12.12206
## 1454	FALSE	1140	0	0	-0.38766759	11.89221
##	.se.fit	.resid	.hat	.sigma	.cooksd	.std.resid
## 31	0.010376094	-0.9277809	0.003099389	0.1848442	0.019320201	-4.9856717
## 54	0.013446134	0.3396888	0.005204791	0.1862286	0.004367639	1.8273367
## 93	0.010823793	0.4546316	0.003372619	0.1860599	0.005050906	2.4434157
## 113	0.014682119	0.2486442	0.006205629	0.1863278	0.002795746	1.3382406
## 126	0.013268987	-0.3176740	0.005068553	0.1862555	0.003718855	-1.7087921
## 146	0.009205961	-0.4052302	0.002439758	0.1861389	0.002897480	-2.1768899
## 156	0.012267278	-0.4487024	0.004332163	0.1860695	0.006332003	-2.4127115
## 179	0.011364710	0.3964721	0.003718134	0.1861515	0.004237739	2.1312073
## 186	0.030453098	0.2039769	0.026697547	0.1863637	0.008438904	1.1093312

## 194	0.007241905	-0.5318698	0.001509781	0.1859196	0.003083084	-2.8558645
## 198	0.020223526	-0.1883497	0.011773949	0.1863764	0.003078135	-1.0165785
## 199	0.013893014	-0.4683053	0.005556500	0.1860356	0.008868449	-2.5196673
## 219	0.010404675	0.4913301	0.003116488	0.1859957	0.005448434	2.6403126
## 243	0.012941157	-0.3420029	0.004821194	0.1862258	0.004097890	-1.8394305
## 305	0.026753283	0.1825739	0.020604523	0.1863798	0.005153140	0.9898372
## 308	0.011450523	-0.4020559	0.003774496	0.1861432	0.004424507	-2.1612836
## 314	0.007402268	0.5638203	0.001577386	0.1858548	0.003620254	3.0275249
## 324	0.013726371	0.2737754	0.005424002	0.1863036	0.002957885	1.4729217
## 384	0.013996415	-0.4086805	0.005639518	0.1861327	0.006855996	-2.1989543
## 399	0.010411430	-0.4675373	0.003120536	0.1860379	0.004939975	-2.5124598
## 411	0.006233619	-0.7665319	0.001118636	0.1853552	0.004741000	-4.1150722
## 431	0.007689580	-0.4989669	0.001702212	0.1859823	0.003060450	-2.6794514
## 432	0.012357790	-0.4539247	0.004396327	0.1860607	0.006577079	-2.4408706
## 441	0.014797998	0.3207763	0.006303973	0.1862516	0.004727813	1.7265523
## 458	0.011882446	0.7099947	0.004064621	0.1855073	0.014866762	3.8171897
## 463	0.007794888	-0.6294756	0.001749154	0.1857095	0.005005586	-3.3803623
## 474	0.008611757	0.4748384	0.002134970	0.1860256	0.003479272	2.5504343
## 496	0.011259005	-0.8873870	0.003649290	0.1849801	0.020833317	-4.7699205
## 524	0.030251116	-1.4491624	0.026344575	0.1824237	0.420013361	-7.8798617
## 534	0.020162077	-0.3631808	0.011702508	0.1861964	0.011373592	-1.9601224
## 547	0.010481738	0.3727596	0.003162824	0.1861854	0.003182970	2.0031843
## 608	0.010569320	0.3725969	0.003215900	0.1861856	0.003233904	2.0023631
## 633	0.007071935	-0.8061929	0.001439743	0.1852390	0.006754026	-4.3286854
## 636	0.023399510	-0.1664876	0.015762384	0.1863906	0.003245892	-0.9004011
## 667	0.011401588	-0.4785207	0.003742304	0.1860184	0.006213632	-2.5722849
## 689	0.010323934	0.5208534	0.003068307	0.1859403	0.006027641	2.7988975
## 706	0.009349963	-0.5768353	0.002516681	0.1858267	0.006057158	-3.0988689
## 711	0.012019691	-0.4158601	0.004159058	0.1861222	0.005219848	-2.2359208
## 717	0.020023118	-0.1884326	0.011541754	0.1863763	0.003018671	-1.0169068
## 748	0.021997449	0.3316103	0.013930061	0.1862369	0.011338145	1.7917539
## 770	0.018178721	0.3003125	0.009513384	0.1862746	0.006294092	1.6190242
## 804	0.013719996	0.3721476	0.005418966	0.1861857	0.005460280	2.0021625
## 808	0.009474324	0.5791744	0.002584074	0.1858216	0.006270748	3.1115399
## 810	0.015360690	-0.2478795	0.006792503	0.1863285	0.003044945	-1.3345191
## 813	0.006950077	-0.7464671	0.001390553	0.1854112	0.005591984	-4.0079012
## 842	0.012282746	0.3050039	0.004343095	0.1862702	0.002933182	1.6400410
## 849	0.015173202	0.3433971	0.006627700	0.1862236	0.005700074	1.8486080
## 875	0.007404898	-0.5627048	0.001578507	0.1858571	0.003608514	-3.0215365
## 898	0.014503746	-0.2555109	0.006055761	0.1863214	0.002880130	-1.3750948
## 899	0.011584581	0.5524585	0.003863394	0.1858769	0.008552223	2.9699178
## 916	0.010275877	-0.4032253	0.003039808	0.1861417	0.003578776	-2.1667713
## 917	0.016295660	-0.6503254	0.007644557	0.1856554	0.023628035	-3.5026865
## 969	0.012795662	-0.7175859	0.004713396	0.1854865	0.017633307	-3.8592597
## 1021	0.015211990	0.3184487	0.006661630	0.1862543	0.004927351	1.7143328
## 1032	0.019118918	-0.2360354	0.010522890	0.1863387	0.004309494	-1.2731469
## 1047	0.013966042	0.3225072	0.005615069	0.1862496	0.004250825	1.7352671
## 1063	0.016442923	-0.4772650	0.007783347	0.1860189	0.012960476	-2.5707540
## 1069	0.013211276	-0.3874703	0.005024559	0.1861642	0.005484014	-2.0841860
## 1133	0.018964856	-0.2313955	0.010353984	0.1863428	0.004073860	-1.2480133
## 1182	0.009144917	0.4497174	0.002407509	0.1860685	0.003521189	2.4158352
## 1217	0.007286201	-0.5226637	0.001528307	0.1859376	0.003013924	-2.8064586
## 1244	0.015486008	0.2472096	0.006903786	0.1863291	0.003078815	1.3309869
## 1245	0.017315760	0.2249133	0.008631604	0.1863485	0.003197422	1.2119977

```

## 1267 0.015467571 -0.3278050 0.006887357 0.1862430 0.005400515 -1.7649017
## 1293 0.017607055 -0.2826392 0.008924458 0.1862939 0.005223739 -1.5232921
## 1299 0.040603418 -1.8824623 0.047460639 0.1794546 1.334039461 -10.3487765
## 1305 0.007240595 -0.5315736 0.001509235 0.1859202 0.003078535 -2.8542736
## 1311 0.007530813 0.5111873 0.001632647 0.1859595 0.003080491 2.7449792
## 1325 0.008665327 -0.5679051 0.002161615 0.1858458 0.005039157 -3.0503515
## 1326 0.010696373 -0.4618734 0.003293680 0.1860476 0.005090275 -2.4822385
## 1329 0.017172734 0.2384540 0.008489601 0.1863368 0.003533868 1.2848729
## 1338 0.009886606 -0.5427514 0.002813863 0.1858972 0.005999302 -2.9161981
## 1350 0.025619261 -0.4636094 0.018894769 0.1860383 0.030364294 -2.5113000
## 1379 0.007643717 -0.5342997 0.001681967 0.1859147 0.003467352 -2.8691591
## 1388 0.015320297 -0.3016449 0.006756826 0.1862736 0.004485097 -1.6239488
## 1454 0.011417306 -0.5477059 0.003752629 0.1858867 0.008162900 -2.9442047
##      model_3_resid influential
## 31      -0.9277809          TRUE
## 54      0.3396888          TRUE
## 93      0.4546316          TRUE
## 113     0.2486442          TRUE
## 126     -0.3176740          TRUE
## 146     -0.4052302          TRUE
## 156     -0.4487024          TRUE
## 179     0.3964721          TRUE
## 186     0.2039769          TRUE
## 194     -0.5318698          TRUE
## 198     -0.1883497          TRUE
## 199     -0.4683053          TRUE
## 219     0.4913301          TRUE
## 243     -0.3420029          TRUE
## 305     0.1825739          TRUE
## 308     -0.4020559          TRUE
## 314     0.5638203          TRUE
## 324     0.2737754          TRUE
## 384     -0.4086805          TRUE
## 399     -0.4675373          TRUE
## 411     -0.7665319          TRUE
## 431     -0.4989669          TRUE
## 432     -0.4539247          TRUE
## 441     0.3207763          TRUE
## 458     0.7099947          TRUE
## 463     -0.6294756          TRUE
## 474     0.4748384          TRUE
## 496     -0.8873870          TRUE
## 524     -1.4491624          TRUE
## 534     -0.3631808          TRUE
## 547     0.3727596          TRUE
## 608     0.3725969          TRUE
## 633     -0.8061929          TRUE
## 636     -0.1664876          TRUE
## 667     -0.4785207          TRUE
## 689     0.5208534          TRUE
## 706     -0.5768353          TRUE
## 711     -0.4158601          TRUE
## 717     -0.1884326          TRUE
## 748     0.3316103          TRUE

```



```
## 770      0.3003125      TRUE
## 804      0.3721476      TRUE
## 808      0.5791744      TRUE
## 810     -0.2478795      TRUE
## 813     -0.7464671      TRUE
## 842      0.3050039      TRUE
## 849      0.3433971      TRUE
## 875     -0.5627048      TRUE
## 898     -0.2555109      TRUE
## 899      0.5524585      TRUE
## 916     -0.4032253      TRUE
## 917     -0.6503254      TRUE
## 969     -0.7175859      TRUE
## 1021     0.3184487      TRUE
## 1032    -0.2360354      TRUE
## 1047     0.3225072      TRUE
## 1063    -0.4772650      TRUE
## 1069    -0.3874703      TRUE
## 1133    -0.2313955      TRUE
## 1182     0.4497174      TRUE
## 1217    -0.5226637      TRUE
## 1244     0.2472096      TRUE
## 1245     0.2249133      TRUE
## 1267    -0.3278050      TRUE
## 1293    -0.2826392      TRUE
## 1299    -1.8824623      TRUE
## 1305    -0.5315736      TRUE
## 1311     0.5111873      TRUE
## 1325    -0.5679051      TRUE
## 1326    -0.4618734      TRUE
## 1329     0.2384540      TRUE
## 1338    -0.5427514      TRUE
## 1350    -0.4636094      TRUE
## 1379    -0.5342997      TRUE
## 1388    -0.3016449      TRUE
## 1454    -0.5477059      TRUE
```

```
model_3_parameters <- model_3 %>% tidy
print(model_3_parameters)
```

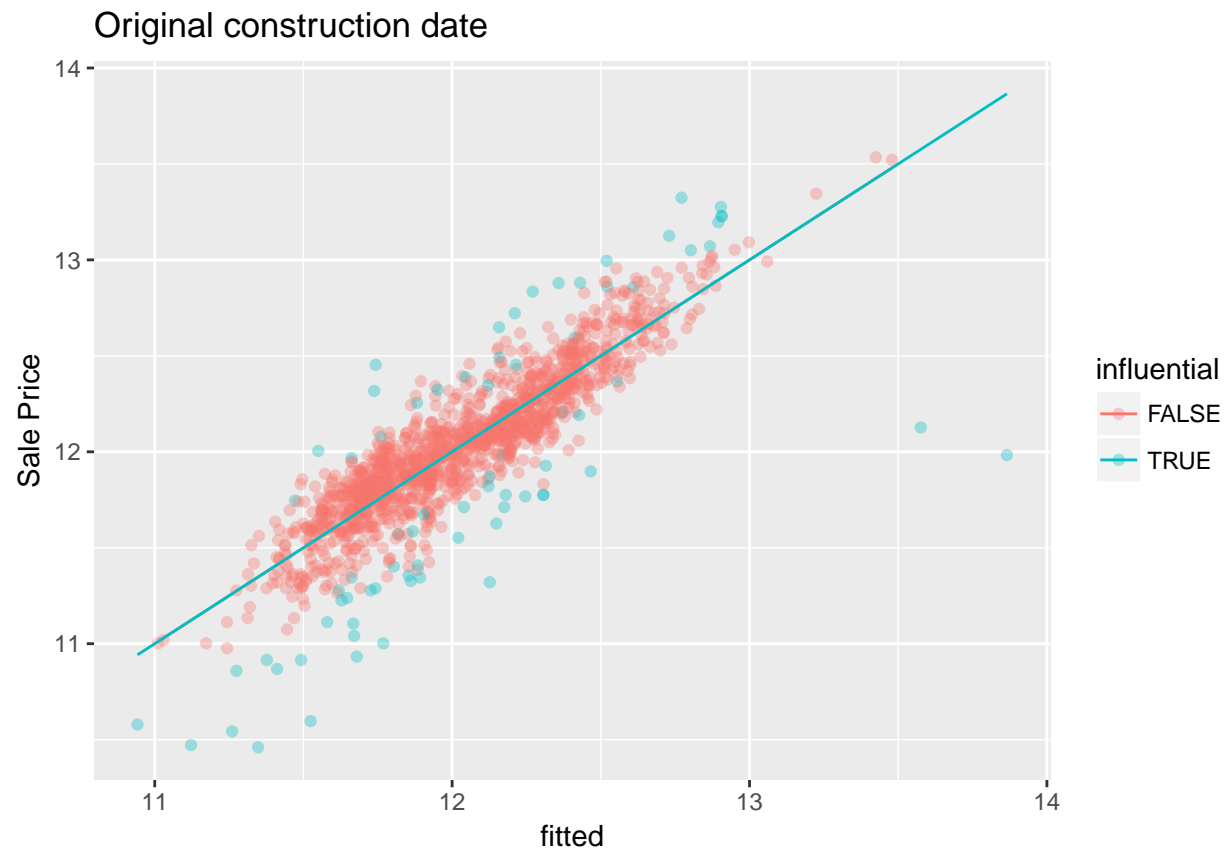
```
##      term      estimate  std.error statistic      p.value
## 1 (Intercept) 4.0737036664 3.830715e-01  10.63432 1.728826e-25
## 2 OverallQual 0.1266395300 5.359019e-03  23.63110 9.056264e-105
## 3  GrLivArea 0.0002961248 1.180195e-05  25.09117 9.247260e-116
## 4   YearBuilt 0.0034136237 2.016206e-04  16.93093 7.763065e-59
```

The R^2 is very good.

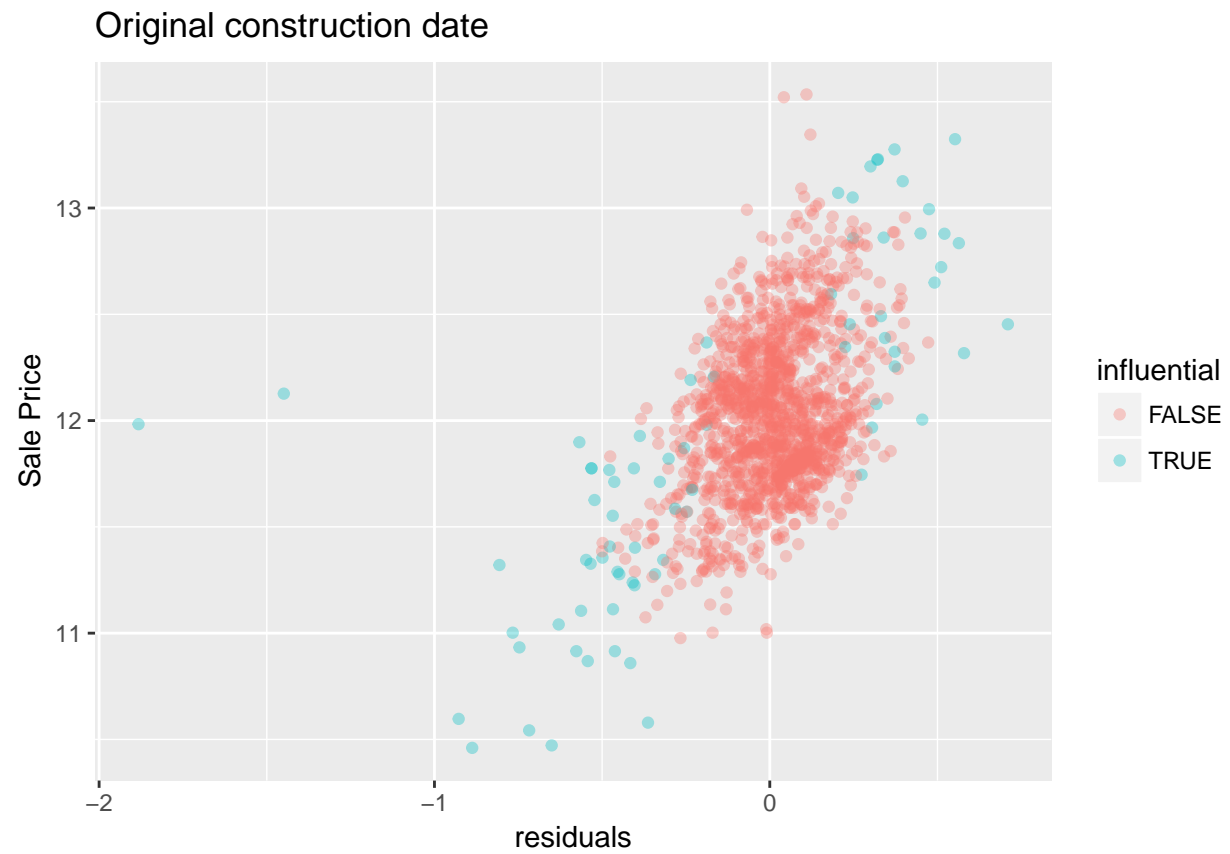
```
model_3_fit <- model_3 %>% glance
print(model_1_fit)
```

```
##      r.squared adj.r.squared      sigma statistic p.value df  logLik      AIC
## 1 0.6677904      0.6675625 0.2303135  2930.795      0  2 73.08848 -140.177
##      BIC deviance df.residual
## 1 -124.3184 77.33862      1458
```

```
t_numeric %>% ggplot() +
  aes(.fitted,SalePriceLog,color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
  geom_line(aes(.fitted,.fitted)) +
  ggtitle(selected_label) +
  xlab('fitted') +
  ylab("Sale Price")
```



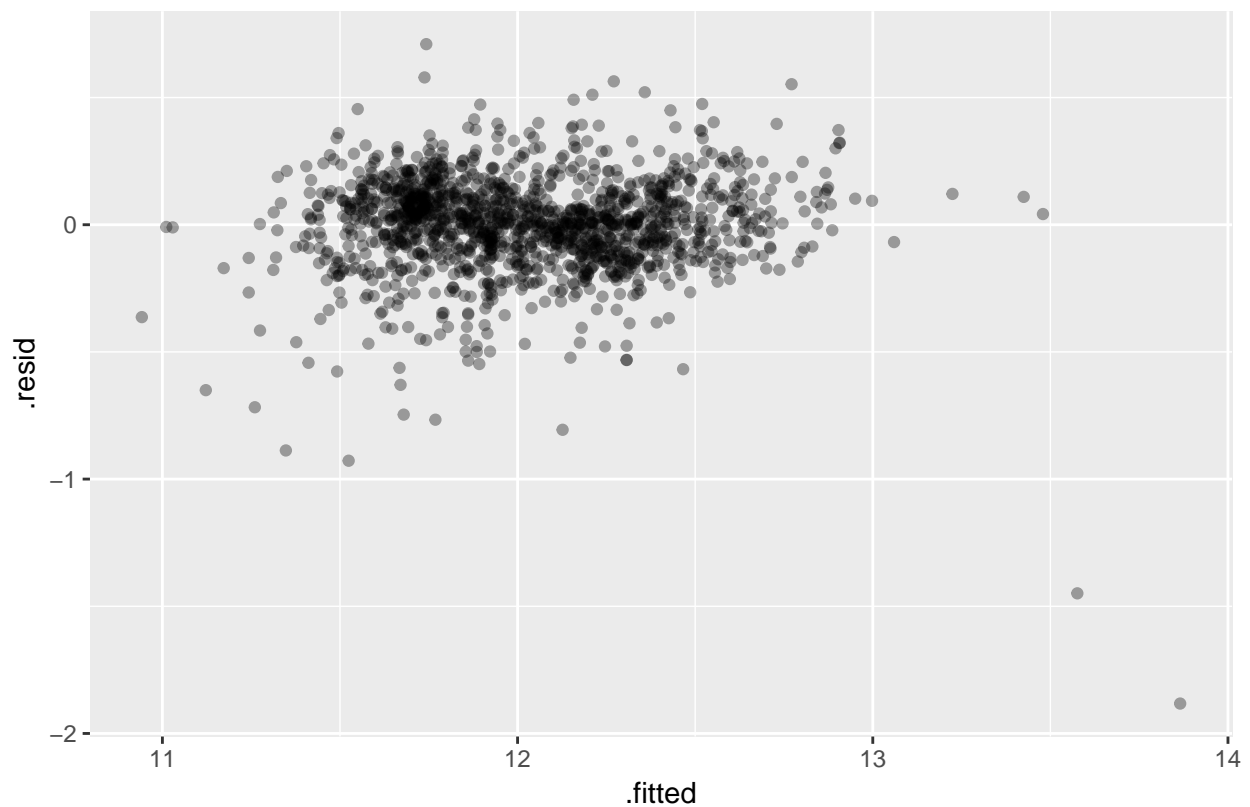
```
t_numeric %>% ggplot() +
  aes(.resid,SalePriceLog,color=influential) +
  scale_y_continuous(labels = comma) +
  scale_x_continuous(labels = comma) +
  geom_point(alpha = 0.35) +
  ggtitle(selected_label) +
  xlab('residuals') +
  ylab("Sale Price")
```



This fitted by residual should show no pattern.

```
t_numeric %>% ggplot() +  
  aes(.fitted,.resid)+  
  scale_y_continuous(labels = comma) +  
  scale_x_continuous(labels = comma) +  
  geom_point(alpha=0.35) +  
  ggtitle(paste("Residuals Plot"))
```

Residuals Plot



```
p <- test %>%
  add_predictions(model_3) %>%
  select(Id,pred) %>%
  mutate(SalePrice = exp(pred)) %>%
  select(Id,SalePrice)
head(p)
```

```
## # A tibble: 6 x 2
##   Id SalePrice
##   <int>     <dbl>
## 1  1461  116581.4
## 2  1462  148890.4
## 3  1463  163782.6
## 4  1464  185154.4
## 5  1465  212309.0
## 6  1466  184790.7
```

```
write_csv(p,file.path('.', 'data', 'model_3.csv'))
```

```
names(t)
```

```
## [1] "Id"           "MSSubClass"    "MSZoning"      "LotFrontage"
## [5] "LotArea"      "Street"        "Alley"         "LotShape"
## [9] "LandContour"  "Utilities"     "LotConfig"     "LandSlope"
## [13] "Neighborhood" "Condition1"    "Condition2"    "BldgType"
## [17] "HouseStyle"   "OverallQual"   "OverallCond"   "YearBuilt"
## [21] "YearRemodAdd" "RoofStyle"     "RoofMatl"      "Exterior1st"
## [25] "Exterior2nd"  "MasVnrType"    "MasVnrArea"    "ExterQual"
```

```
## [29] "ExterCond"      "Foundation"      "BsmtQual"        "BsmtCond"
## [33] "BsmtExposure"   "BsmtFinType1"    "BsmtFinSF1"      "BsmtFinType2"
## [37] "BsmtFinSF2"     "BsmtUnfSF"       "TotalBsmtSF"     "Heating"
## [41] "HeatingQC"      "CentralAir"      "Electrical"       "LowQualFinSF"
## [45] "GrLivArea"      "BsmtFullBath"    "BsmtHalfBath"    "FullBath"
## [49] "HalfBath"       "BedroomAbvGr"    "KitchenAbvGr"    "KitchenQual"
## [53] "TotRmsAbvGrd"   "Functional"       "Fireplaces"       "FireplaceQu"
## [57] "GarageType"     "GarageYrBlt"     "GarageFinish"     "GarageCars"
## [61] "GarageArea"     "GarageQual"       "GarageCond"       "PavedDrive"
## [65] "WoodDeckSF"     "OpenPorchSF"     "EnclosedPorch"    "3SsnPorch"
## [69] "ScreenPorch"    "PoolArea"        "PoolQC"           "Fence"
## [73] "MiscFeature"    "MiscVal"         "MoSold"           "YrSold"
## [77] "SaleType"       "SaleCondition"    "SalePrice"        "SalePriceLog"
## [81] "secondFloor"    "X1stFlrSF"       "X2ndFlrSF"       "X3SsnPorch"
```

```
m <- data.frame(model.matrix(Id ~ MSZoning,t))
names(m)
```

```
## [1] "X.Intercept." "MSZoningFV"     "MSZoningRH"     "MSZoningRL"
## [5] "MSZoningRM"
```

```
model <- t %>% lm(SalePriceLog ~ MSZoning,data=.)
model %>% tidy
```

```
##           term      estimate std.error statistic      p.value
## 1 (Intercept) 11.1182590 0.1148330  96.821120 0.000000e+00
## 2 MSZoningFV   1.1283568 0.1233504   9.147572 1.907631e-19
## 3 MSZoningRH   0.6315806 0.1463839   4.314549 1.707164e-05
## 4 MSZoningRL   0.9676260 0.1153308   8.390009 1.135445e-16
## 5 MSZoningRM   0.5746336 0.1174372   4.893112 1.102931e-06
```

```
model %>% glance
```

```
##    r.squared adj.r.squared      sigma statistic      p.value df    logLik
## 1 0.1758388    0.1735731 0.3631338  77.60784 1.021343e-59   5 -590.1895
##           AIC      BIC deviance df.residual
## 1 1192.379 1224.096 191.8653      1455
```

```
names(t)
```

```
## [1] "Id"           "MSSubClass"    "MSZoning"      "LotFrontage"
## [5] "LotArea"      "Street"        "Alley"         "LotShape"
## [9] "LandContour"  "Utilities"     "LotConfig"     "LandSlope"
## [13] "Neighborhood" "Condition1"    "Condition2"    "BldgType"
## [17] "HouseStyle"   "OverallQual"   "OverallCond"   "YearBuilt"
## [21] "YearRemodAdd" "RoofStyle"     "RoofMatl"      "Exterior1st"
## [25] "Exterior2nd"  "MasVnrType"    "MasVnrArea"    "ExterQual"
## [29] "ExterCond"    "Foundation"    "BsmtQual"      "BsmtCond"
## [33] "BsmtExposure" "BsmtFinType1"  "BsmtFinSF1"    "BsmtFinType2"
## [37] "BsmtFinSF2"   "BsmtUnfSF"     "TotalBsmtSF"   "Heating"
## [41] "HeatingQC"    "CentralAir"    "Electrical"     "LowQualFinSF"
## [45] "GrLivArea"    "BsmtFullBath"  "BsmtHalfBath"  "FullBath"
## [49] "HalfBath"     "BedroomAbvGr"  "KitchenAbvGr"  "KitchenQual"
## [53] "TotRmsAbvGrd" "Functional"     "Fireplaces"     "FireplaceQu"
## [57] "GarageType"   "GarageYrBlt"   "GarageFinish"   "GarageCars"
## [61] "GarageArea"   "GarageQual"     "GarageCond"     "PavedDrive"
## [65] "WoodDeckSF"   "OpenPorchSF"   "EnclosedPorch"  "3SsnPorch"
```

```
## [69] "ScreenPorch"    "PoolArea"      "PoolQC"        "Fence"
## [73] "MiscFeature"    "MiscVal"       "MoSold"        "YrSold"
## [77] "SaleType"       "SaleCondition" "SalePrice"     "SalePriceLog"
## [81] "secondFloor"    "X1stFlrSF"     "X2ndFlrSF"     "X3SsnPorch"
```

```
my_corr <- function(feature) {
  f_sub <- substitute(feature)
  my_f <- paste("SalePriceLog ~", f_sub)
  model <- t %>% lm(my_f, data=.)
  model %>% glance
}
```

```
my_corr(PavedDrive)
```

```
##      r.squared adj.r.squared      sigma statistic      p.value df    logLik
## 1 0.09322669    0.09198198 0.3806376   74.89815 1.090995e-31   3 -659.9237
##      AIC      BIC deviance df.residual
## 1 1327.847 1348.992 211.0974         1457
```