Lab 2 Group 8 Text File
Dan Williams, Amanda Spillane, Tony Tedesco

#Exercise 1 Question 1
Ames Variables(42)
[1] "Id" "LotFrontage" "LotArea"
[4] "OverallQual" "OverallCond" "YearBuilt"
[7] "YearRemodAdd" "BsmtFinSF1" "BsmtFinSF2"
[10] "BsmtUnfSF" "TotalBsmtSF" "X1stFlrSF"
[13] "X2ndFlrSF" "LowQualFinSF" "GrLivArea"
[16] "BsmtFullBath" "BsmtHalfBath" "FullBath"
[19] "HalfBath" "BedroomAbvGr" "KitchenAbvGr"
[22] "TotRmsAbvGrd" "Fireplaces" "GarageYrBlt"
[25] "GarageCars" "GarageArea" "WoodDeckSF"
[28] "OpenPorchSF" "EnclosedPorch" "ScreenPorch"
[31] "PoolArea" "MiscVal" "MoSold"
[34] "YrSold" "SalePrice" "GarageType2Types"
[37] "GarageTypeAttchd" "GarageTypeBasment" "GarageTypeBuiltIn"
[40] "GarageTypeCarPort" "GarageTypeDetchd" "GarageOutside"
No variables can be negative.
Each variable has positive integer values for their respective behavior
Many variables are binary with either 1 or 0 values for yes or no
Some variables have year int values (Ex. YearBuilt 2003)
Some variables hold either 0 or values such as 200 for sqft depending on if the value exists for the unit
Some variables are total count values
Each variable has its own use in describing the unit as a whole

#Exercise 1 Question 4
To identify and remove outliers that are 1.5 times outside the upper quadrant,
we would use the command boxplot.stats()$out on our data set for each variable.

#Exercise 2 Question 2

```
summary(multi_reg)
```

Call:
lm(formula = ameslist$SalePrice ~ ., data = z)

Residuals:
Min 1Q Median 3Q Max
-473767 -16943 -2030 14996 312486

Coefficients: (5 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.707e+05 4.611e+04 -5.871 4.32e-09 ***

42 NA NA NA NA

Id -2.227e+00 7.170e-02 -31.061 < 2e-16 ***

LotFrontage 1.798e+00 1.571e+00 1.145 0.252

LotArea 5.525e-01 4.261e-03 129.687 < 2e-16 ***

OverallQual 1.858e+04 3.983e+01 466.538 < 2e-16 ***

OverallCond 5.611e+03 3.708e+01 151.335 < 2e-16 ***

YearBuilt 3.504e+02 2.358e+00 148.586 < 2e-16 ***

YearRemodAdd 1.055e+02 2.342e+00 45.047 < 2e-16 ***

BsmtFinSF1 2.245e+01 1.570e-01 142.960 < 2e-16 ***

BsmtFinSF2 1.024e+01 2.377e-01 43.106 < 2e-16 ***

BsmtUnfSF 8.743e+00 1.424e-01 61.401 < 2e-16 ***

TotalBsmtSF NA NA NA NA

X1stFlrSF 4.832e+01 1.986e-01 243.305 < 2e-16 ***

X2ndFlrSF 4.124e+01 1.598e-01 258.103 < 2e-16 ***

LowQualFinSF 2.229e+01 7.571e-01 29.436 < 2e-16 ***

GrLivArea NA NA NA NA

BsmtFullBath 6.721e+03 8.635e+01 77.834 < 2e-16 ***

BsmtHalfBath 6.087e+02 1.371e+02 4.439 9.03e-06 ***

FullBath 3.566e+03 9.519e+01 37.459 < 2e-16 ***

HalfBath -6.171e+02 8.974e+01 -6.877 6.10e-12 ***

BedroomAbvGr -9.063e+03 5.801e+01 -156.225 < 2e-16 ***

KitchenAbvGr -3.497e+04 1.718e+02 -203.548 < 2e-16 ***

TotRmsAbvGrd 6.285e+03 4.011e+01 156.687 < 2e-16 ***

Fireplaces 3.796e+03 5.924e+01 64.079 < 2e-16 ***

GarageYrBlt -9.662e+01 2.465e+00 -39.198 < 2e-16 ***

GarageCars 1.758e+04 9.470e+01 185.679 < 2e-16 ***

GarageArea 1.347e+01 3.276e-01 41.119 < 2e-16 ***

WoodDeckSF 2.340e+01 2.706e-01 86.471 < 2e-16 ***

OpenPorchSF -7.656e+00 5.239e-01 -14.612 < 2e-16 ***

EnclosedPorch 7.684e+00 5.593e-01 13.739 < 2e-16 ***

ScreenPorch 5.815e+01 5.530e-01 105.146 < 2e-16 ***

PoolArea -8.297e+01 8.073e-01 -102.780 < 2e-16 ***

MiscVal -1.777e+00 1.885e-01 -9.427 < 2e-16 ***

MoSold -7.376e+01 1.140e+01 -6.469 9.87e-11 ***

YrSold -2.600e+02 2.293e+01 -11.340 < 2e-16 ***

GarageType2Types -5.214e-10 4.547e+02 0.000 1.000

GarageTypeAttchd -2.926e-09 6.753e+01 0.000 1.000

GarageTypeBasment -5.653e-10 2.597e+02 0.000 1.000

GarageTypeBuiltIn -5.095e-10 1.305e+02 0.000 1.000

GarageTypeCarPort -1.648e-09 3.727e+02 0.000 1.000
GarageTypeDetchd NA NA NA NA

GarageOutside NA NA NA NA

Signif. codes: 0 '*** *0.001* *' 0.01 '' 0.05 '.' 0.1 ' ' 1

Residual standard error: 37100 on 1554095 degrees of freedom
(459207 observations deleted due to missingness)
Multiple R-squared: 0.801, Adjusted R-squared: 0.801
F-statistic: 1.691e+05 on 37 and 1554095 DF, p-value: < 2.2e-16

In this data, we found very small, insignificant positive or negative relationships
with almost all of the predictors in response to the sale price
The variables with porches show significant positive relationships between an
added porch and added value. This is logically correct, as these additions would
tend to add value to the property. There are many other variables, such as extra floors
that are logically true, because these additions add value to the property.
Their coefficients tell as that with the addition of an EnclosedPorch, for example,
the added value of the house is $7,684 dollars.