Lab 2

1. Values it can take, what each one means, could it have a negative value.

ID is none useful data just a classifier. Range 1:1460

MSSubClass is the type of neighborhood the house is in the higher the number the more expensive the house. Range 20:190

LotFrontage assuming the front yard (unknown). Range 21:313

LotArea is the permiter of the land of the house. Range 1300:215245

OverallQual is on a 1:10 scale the higher the number the nicer the house.

OverallCond is on a 1:10 scale the higher the number the nicer the house.

YearBuilt is the year the house was made value ranges from 1872:2010.

YearRemodAdd is the year the house was remodled, values range from 1950:2010.

MasVnrArea unknown. Range 0:1600

BsntFinSF1 finished a basement room square feet. Range 0:5644

BsmtFinSF2 finished another basement square feet. Range 0:1474

BsmtUnfSF uunfinshed room in basement square feet. Range 0:2336

TotalBsmtSF Is the total square feet of the basement, values range from 0:6110.

X1stFlrSF is the 1st floor square feet. Range 334:4692

X2ndFlrSF is the second floor square feet. Range 0:2065

LowQualfinSF unknown. Range 0:572

GrLivArea the living area square feet. Range 334:5642

BsmtFullBath is basement full bath. Range 0:3

BsmtHalfBath is basement half bath. Range 0:2

FullBath Full Bathroom at ground level. Range 0:3

HalfBath half bathroom at ground level. Range 0:2

BedroomAbvGr is bedroom above ground. Range 0:8

KitchenAbvGr is kitchen above ground. Range 0:3

TotRmsAbvGrd total rooms above ground. Range 2:14

Fireplaces number of fireplaces. Range 0:1

GarageYrBlt garage year buildt. Range 1900:2010

GarageCars garage car capacity. Range 0:4

GarageArea garage square feet. Range 0:1418

WoodDeckSF Wooddeck square feet. Range 0:857

OpenPorchSF open porch square feet. Range 0:547

EnclosedPorch enclosed porch square feet. Range 0:552

X3SsnPorch unknown. Range 0:508

ScreenPorch Screen porch square feet. Range 0:408

PoolArea Pool square feet. RAnge 0:738

MiscVal other values not listed above. RAnge 0:15500

MoSold month sold. Range 1:12

YrSold year sold. Range 2006:2010

SalePrice sold price. Range 34900:755000

None of the above can be negative.

3. This does match my beliefs that based on saleprice the more things you have the bigger the price. As for miscellaneous variables if that variable is high it has a neagtive effect with sale price.

4. the Id for this outlier is 1299 it has a lot area of 63887, quaility score of 10, condition is a 5, buildt in 2008, basement square footage of 6110, GrlivArea of 5642, 3 Fireplaces, sale price 160000.

2B. 1.There is a very small p value (2.2e-16 showing a relation between the two.

2.MSSubClass, LotArea, OverallQual, OverallCond, MasVnrArea,BsmtFinSf1, X1stFlrSF, X2ndFlrSF, BsmtFullBath,BedroomAbvGr,KitchenAbvGr,TotRmsAbvGrd, Fireplaces, GarageCars, WoodDecksSF, ScreenPorch, PoolArea.

3. Going by year Sold, for every year the house would go up in value by 0.769917.

3B. 1. The fit has some curvuture to it so it has some issues with the fit.

2. The graph dont seem to indicate any unusual large group of data outside of any of the curves.

3. The Levearge seems to be leveling right in the middle so no extremley high leverage.

4B. From what i was doing none seem to be to significant because the p value barely changed between new varations of implematations.

5B. To me the only one that makes most sense if the home owner wanted to know the exact perameter of the house then you could create a new variable and sqrt the total square feet.