

MSDS 6371 Final Project

Team Members:

Morgan Nelson
Michael Hedge
Paul Huggins
Perry Gill

Introduction:

Our team was tasked with predicting home sale prices using the knowledge gained during our time in Statistical Foundations. To make our determination of predicted sales price, we are going to be evaluating 79 explanatory variables that describe nearly every aspect of homes in Ames, Iowa.

Data Description:

The “Ames Housing Dataset” is a publication by Dean De Cock that compiles data on 2,930 homes in the Ames, Iowa area that were sold between 2006 and 2010. Types of explanatory variables included and used in the analysis are continuous, discrete, nominal and ordinal data. The first analysis is utilizing the Neighborhood (Neighborhood) and Above ground living area square footage (GrLivArea) variables but filtering the neighborhoods down to the following three: NAmes, Edwards, and BrkSide. The second analysis performed utilizes the entire dataset to predict the sales price of the home. For more information on the dataset, please visit: [Ames Housing Dataset](#). For a full list of variable descriptions, refer to Appendix 1DESC.

Analysis Question #1:

Century 21 Team,

Our initial analysis created a model to predict the sales price of homes in the NAmes, Edwards, and BrkSide neighborhoods using only the Above ground living area square footage (GrLivArea) variable. The discussion and results will use living area in increments of 100 sq. ft. Our team created a model for a combination of the three neighborhoods as well as a model for each one.

The scatterplot (Appendix 1A) and fit diagnostic plots (Appendix 1B) for the original dataset showed some houses that we believe to skew the data and thus the model itself. This initial analysis provided insight on two questionable datapoints.

Questionable Homes:

1. Home ID: 524: This home was an outlier and we believe was an error or not in the target population. The GrLivArea of the house was very disproportionate to the Sale Price of the home compared to the rest of the data set. The cost per square foot is less than half the cost per square foot of even the lowest end of the data set.
2. Home ID 1299: This home had a Total Basement Square Feet (TotalBsmtSF) of 6,110 sq. ft which was nearly double the next closest value and more than the Greater Living Area (GrLivArea) of the whole house. A square footage this large we believe was either an error in the data or the makings of a massive underground compound which we believe does not represent to the vast majority of home buyers in the Ames area. If a potential homebuyer is

ever interested in something along these lines though, this would be on the short list of potential sales.

Before we built the model, we examined the assumptions for linear regression and determine if a transformation is necessary.

Original Data (Appendix 1A & 2A):

1. Linear relationship: The QQ plots show a curvature in the dataset at the extents of the data, and the QQ plot in the center of the diagnostics output shows a distinct deviation from the normal line. This assumption is not met.
2. Constant variance: The residual plots show a clustering effect around a central point instead of the desired random cloud effect. This assumption is not met.
3. Normal distribution: Judging by the histogram, there appears to be a right skew to the dataset. Given the large number of datapoints the central limit theorem would resolve this minor issue.
4. Independence: We assume all observations are independent.

Given that our original data did not meet 2 of the 4 key assumptions, we decided to perform a log-log transformation of the sales price and greater living area. The assumptions were then reevaluated for the transformed data below.

Log-Log Transformation (Appendix 3A)

1. Linear relationship: The QQ plots show a much straighter line that more closely mimics the normal line and the central QQ plot fits the line better. This assumption is now met.
2. Constant variance: The residual plots are resembling a cloud centered around the origin, indicating that the variances have a more equal spread. This assumption is now met.
3. Normal distribution: Judging by the histogram, there appears to be a slight left skew to the dataset. Given the large number of datapoints the central limit theorem would resolve this minor issue.
4. Independence: We assume all observations are independent.

The log-log transformation met the basic assumptions for linear regression. The Cook's D plots in Appendix 3A showed no troubling data points and the scales max out at 0.1, which is within the acceptable range. There were a few data points with high leverage but these will be resolved through the Central Limit Theorem.

The next step was building and fitting the model. We broke this analysis down into two separate models; one model for the combined three neighborhoods of interest and one model that fitted to each neighborhood. We did this to allow for more flexibility on your end. Please refer to Appendix 1DEF for a refresher on the equation of a line.

Simple Linear Regression Model (Appendix 3A):

The model below is the full simple linear regression model for the three neighborhoods of interest. To test this model, our null hypothesis is that the coefficients are zero and the alternative hypotheses is that at least one of the coefficients is not zero. Our p-value was <0.001 which indicates

that we reject the null hypothesis and that a relationship exists between the sales price and the above ground living area square footage (GrLivArea).

$$\text{Log}(SalePrice) = 10.234 + 0.623 * \text{Log}(GrLivArea)$$

In general, a doubling of the above ground living area square footage (GrLivArea) equates to a multiplicative change of $2^{0.623} = 1.54$ in the median distribution of the Sale Price. In other words, a doubling of the above ground living area square footage (GrLivArea) increases the Sale Price estimate by 54%.

The model has an r-squared of 0.453 and an adjusted r-squared of 0.4517. The r-squared indicated that 45% of the variation in sales price can be attributed to the variation in the above ground living area square footage (GrLivArea). The CV press statistic of this model is 15.734. While these coefficients are purely estimates, they provide some insight into the fitted model.

Model	Adjusted R-Squared	CV Press
Simple Linear Regression	0.4517	15.734

The 95% confidence interval for the log of the slope is (0.554, 0.692) and log of the y-intercept is (10.059, 10.408). The p-value for the model is <0.001 (0.05 cutoff) indicating that there is a statistically significant relationship between the sale price and greater living area of the home.

In conclusion, this model provides a quick and simple estimate of the sales price of a home in either of these neighborhoods. This might be useful if a potential homebuyer is looking at these 3 neighborhoods but hasn't decided on a certain one. This model does have its drawbacks, one being that it is too generic to be used in more definitive scenarios or applied to the dataset as a whole. Due to the nature of this analysis, we cannot infer causality and deduce that the greater living area square footage causes a direct change in the sales price. The Above ground living area square footage (GrLivArea) is one of 79 variables that might each play a role in predicting the final sales price and confounding variables can come into play. At the minimum, we can conclude that sales price and above ground living area square footage (GrLivArea) do have a correlation between them but we cannot determine to what extent when looking at the entirety of the dataset. The next model will address each neighborhood individually to generate a more accurate model for predicting sale price.

Multiple Linear Regression Model (Appendix 4A):

The model below is the full multiple linear regression model for the three neighborhoods of interest. To test this model, our null hypothesis is that all Bi are zero and the alternative hypotheses is that at least one of the coefficients is not zero. Our p-value was <0.001 (0.05 cutoff) indicating that we reject the null hypothesis and there a relationship exists between the sales price and above ground living area square footage (GrLivArea) of the home. The assumptions previously mentioned are still valid here as we did not change any of the parameters and are still using the log-log transformation.

$$\text{Log}(SalesPrice) = 9.65 + 0.90*\text{log}(GrLivArea) + 0.07(\text{BrkSide}_\text{GrLivArea}) + 0.39*(\text{Edwards}_\text{GrLivArea}) + 0.99*(\text{Names}_\text{GrLivArea}) - 0.09*\text{log}(\text{BrkSide}_\text{GrLivArea}) - 0.24*\text{log}(\text{Edwards}_\text{GrLivArea}) - 0.42*\text{log}(\text{Names}_\text{GrLivArea})$$

The equation above is the full multiple linear regression equation. The next step is to insert the coefficients for each neighborhood to determine an equation for each neighborhood individually since each neighborhood is going to have its' own slope (Appendix 4A). The model has an r-squared of 0.5807 and an adjusted r-squared of 0.5787 meaning that 58% of the variance in sales price can explain the variance in neighborhood and above ground living area square footage (GrLivArea). The CV press statistic for this model is 98.463.

Model	Adjusted R-Squared	CV Press
Multiple Linear Regression	0.5787	98.463

The 95% confidence interval for the above ground living area square footage (GrLivArea) coefficient is (0.855, 0.949). For more confidence intervals, please view Appendix 4A.

We simplified the model by separating out the neighborhoods, leaving us with three different equations to be used when looking in a particular neighborhoods.

$$\{\log(\text{SalesPrice}) \mid \text{BrkSide}\} = 9.72 + 0.81 * \log(\text{GrLivArea})$$

$$\{\log(\text{SalesPrice}) \mid \text{Edwards}\} = 10.04 + 0.67 * \log(\text{GrLivArea})$$

$$\{\log(\text{SalesPrice}) \mid \text{Names}\} = 10.64 + 0.349 * \log(\text{GrLivArea})$$

In conclusion, this model can be used to get an estimate of the Sales Price of a home in the BrkSide, Edwards and NAmes neighborhoods using only the above ground living area square footage (GrLivArea) in hundreds.

Analysis Question #2:

The second analysis utilized all 79 variables in the dataset to build a predictive model for sales price in all of Ames, Iowa. Unlike the first analysis, this model is not restricted to the three neighborhoods.

Our method of approach for this analysis was to evaluate all the variables in the dataset and either remove variables with no correlation or transform variables in order to meet the assumptions. Four models were created: forward selection, backwards elimination, stepwise selection and a custom model. We will evaluate each of these models from their adjusted R-squared, CV press and Kaggle scores.

Before we can build our predictive model, we needed to ensure that the data meets the assumptions for linear regression.

Original Data (Appendix 1B):

1. Linear relationship: The predicated value line looks to follow the linear line pretty well.
2. Constant variance: The residual plots show an unequal spread along the line indicating that a transformation may be necessary. This assumption is not met.
3. Normal distribution: Judging by the histogram, there residuals appear to be normally distributed. This assumption is met.
4. Independence: We assume all observations are independent.

The residual plots for each variable are shown in Appendix 2B and indicate that some of the variables could benefit from a transformation given their unequal distribution and outliers. We used a log transformation on the Sales Price and analyzed the residual plots for each variable.

Transformed Data (Appendix 2B):

1. Linear relationship: The predicated value line looks to follow the linear line relatively well.
2. Constant variance: The residual plots show an unequal spread along the line indicating that a transformation may be necessary. This assumption is not met.
3. Normal distribution: Judging by the histogram, there residuals appear to be normally distributed. This assumption is met.
4. Independence: We assume all observations are independent.

The Cook's D plots in Appendix 2B showed no troubling data points and the scales maxed out at 0.08 which is within the acceptable range. There were a few data points with high leverage but these will be resolved through the Central Limit Theorem. The residual and QQ plots looked much better on the transformed data compared to the original data.

The following variables were changed from their original state to run the models (Appendix 2B):

- TotalBsmtSF (Total Basement Square Feet): This is the summation of BsmtFinSF1, BsmtFinSF2 and BsmtUnfSF. This variable was removed to simplify the model and reduce redundancy.
- GrLivArea (Greater Living Area): This is the summation of 1stFlrSF, 2ndFlrSF and LowQualFinSF. This variable was removed to simplify the model and reduce redundancy.
- Log transformations:
 - Sales Price
 - Lot Area
 - BsmtFinSF1 (Finished Basement square feet)
 - BsmtFinSF2 (Finished Basement square feet)
 - BsmtUnfSF (Unfinished Basement square feet)
 - WoodDeckSF (Wood Deck square feet)
 - OpenPorchSF (Open Porch square feet)

Interpolated Data:

The test data set to be scored and submitted to Kaggle had missing values that if not addressed, would diminish the model. Our team tried removing the variables with missing data from the model but the model performed worse than desired due to the lack of highly corelative variables. We ended up imputing the missing values using the PROC MI function. This function replaces the missing value with a set of plausible values representing the uncertainty in a random sample of non-missing values. Following interpolation, the model performed significantly better than simply dropping the variables with missing data. The following variables along with the number of missing values that were interpolated.

Variable	Number of Missing Values	% of Total (1459 Observations)
BsmtFinSF1	1	0.06%
BsmtFinSF2	1	0.06%
BsmtUnfSF	1	0.06%
BsmtFullBath	2	0.13%

BsmtHalfBath	2	0.13%
GarageCars	1	0.06%
GarageArea	1	0.06%
MasVnrArea	15	1.03%

Building the Models:

Forward Model (Appendix 3B):

The forward model utilizes the process of adding variables. The model starts out as an empty equation and then variables are added one by one. The variables that are added in each step are ordered by significance. Meaning that the most significant variable is added first and the least significant is added last. Significance is determined by the Schwarz Bayesian Criteria (SBC). The forward model continues to add variables and monitoring the SBC value. Once the SBC value is negatively affected from its previous value the model stops and removes the variable that caused the change in SBC value. This point is where the model has reached its best fit. Our forward model reached 21 variables before the variable 'PoolArea' was added and started to diminish the model. The model terminated and dropped the 'PoolArea' variable. The table below holds the results of the model.

Model	Adjusted R-Squared	CV Press	Kaggle Score
Forward	0.9190	19.832	0.12629

Backwards Model (Appendix 4B):

The backwards model is the inverse of the forwards model. The model starts with all of the variables in the equation and then removes the least significant variable in each step. This process is repeated until the SBC is negatively impacted. At this point, the last variable that was added is removed and the model has reached its' best fit. The backwards model reached 10 variables before terminating on the variable 'Heating'. The table below holds the results of this model.

Model	Adjusted R-Squared	CV Press	Kaggle Score
Backwards	0.9207	20.111	0.12495

Stepwise Model (Appendix 5B):

The stepwise model utilizes the same premise as the previous two models in reference to the SBC values but the selection of variables is what separates this model from the other two. This model starts with no variables and adds one at a time. At each step, each variable is re-evaluated for significance. If adding a variable makes a current variable insignificant, the insignificant variable is removed the model adds the next the variable. This process is repeated until the SBC is negatively impacted, at which point the model removes the last added variable and arrives at the final output. The model reached 18 variables before terminating. The last variable included in the model was the log of 'BsmtUnfSF' The table below holds the results of the model.

Model	Adjusted R-Squared	CV Press	Kaggle Score
Stepwise	0.9182	19.971	0.12648

Custom Model (Appendix 6B, 7B & 8B):

For the custom model, we decided to run a stepwise model and define the significance level cutoff parameters to find the ideal significance levels. The three significance levels we tested were 0.10, 0.50 and 0.99. These values represent the value at which the model terminates adding variables. For example the 0.99 model will include all variables with a p-value below 0.99 and terminate once a variable goes over this value. The 0.99 model included the most variables while the 0.10 model used the least amount of variables. The 0.99 model displayed the highest CV press value which indicates that the model has the loosest fit and would. The 0.10 model shows the lowest CV press but highest Kaggle score which tells us that the model has the tightest fit but that it is potentially over-fitted to the test dataset. The best result of the three models was the 0.50 significance level which gave us the lowest Kaggle score out of all the models indicating that is it best fitted for the test data set.

Model	Adjusted R-Squared	CV Press	Kaggle Score
Custom – 0.99 SL	0.9204	22.01	0.12589
Custom – 0.50 SL	0.9211	20.28	0.12484
Custom – 0.10 SL	0.9207	19.75	0.12635

Summary and Conclusion:

Model	Adjusted R-Squared	CV Press	Kaggle Score
Forward	0.9190	19.832	0.12629
Backwards	0.9207	20.111	0.12495
Stepwise	0.9182	19.971	0.12648
Custom – 0.50 SL	0.9211	20.281	0.12484

In analysis 1, we attempted to predict the sales price of homes in Ames, Iowa using only the Greater Living Area of the home. Two models were run on the Names, BrkSide and Edwards neighborhoods. The first model was the simplest model and only used the three neighborhoods with the Greater Living Area. The second model utilized each neighborhood as its' own group. These models were simple in nature and had an average adjusted r-squared value of 0.5152 and an average press statistic of 57.099. Comparing these results to the results in the table above accentuate the difference in model types and variable selection.

In analysis 2, we attempted to predict the sales price of homes in Ames, Iowa using 4 different models; forward, backward, stepwise and custom. The forward and stepwise models performed slightly better than the backwards model from their CV and Kaggle scores but had a slightly lower R-Squared value. This could be a sign that our backwards model fit the training data the best but was too overfitted for the test dataset. In this case, the simplest model worked the best out of all the models.

Thank you for allowing our team to analyze this data for you, please feel free to reach out if you have any questions or would like further analysis done.

Appendix: SAS Code:

```
* QUESTION 1 -----;
* Simple Linear Regression Model;
data question1sr; * data name;
set WORK.TRAIN; * importing training data set;
    logsale = log(saleprice); * taking log of sale price;
    logsf = log(round(GrLivArea/100)); * taking log of GrLivArea;
    where (neighborhood = 'NAmes' or neighborhood = 'Edwards' or neighborhood = 'BrkSide')
        and (id <> 524 and id <> 1299); * filtering down to select neighborhoods and removing outliers;
run;

proc glm data = question1sr plots = all; * run regression on train data;
model logsale = logsf / cli clparm; * model y=x / confidence intervals and parameters;
run;

* Multiple Linear Regression Log-Log Model;
data question1mr; * data name;
    set WORK.TRAIN; * import training data;
    logsale = log(saleprice); * log of sale price;
    logsf = log(round(GrLivArea/100)); * log of GrLivArea;
    where (id <> 524 and id <> 1299); * remove outliers;
    if neighborhood = 'Edwards' then NeighborhoodDum = 'Edwards'; * code neighborhoods
    else if neighborhood = 'BrkSide' then NeighborhoodDum = 'BrkSide';
    else if neighborhood = 'NAmes' then NeighborhoodDum = 'NAmes';
    else NeighborhoodDum = 'Others'; * all other neighborhoods;
run;

proc glm data = question1mr plots=all; * run regression on train data;
class NeighborhoodDum; * set neighborhood to be reference variable;
model logsale = logsf | NeighborhoodDum / solution cli clparm; * model y=x / solution for
neighborhoods and confidence intervals with parameters;
run;

* QUESTION 2 -----;
data q2train; * data name;
    set WORK.TRAIN; * import training data;
    LogSalePrice = log(SalePrice); * log transform;
    LogLotArea = log(LotArea); * log transform;
    LogBsmtFinSF1 = log(BsmtFinSF1+1); * log transform. +1 is added to avoid log(0);
    LogBsmtFinSF2 = log(BsmtFinSF2+1); * log transform. +1 is added to avoid log(0);
    LogBsmtUnfSF = log(BsmtUnfSF+1); * log transform. +1 is added to avoid log(0);
    LogWoodDeckSF = log(WoodDeckSF+1); * log transform. +1 is added to avoid log(0);
    LogOpenPorchSF = log(OpenPorchSF+1); * log transform. +1 is added to avoid log(0);
    where (id <> 524 and id <> 1299); * remove outliers;
```

```

run;

data q2test; * data name;
  set WORK.TEST; * import test data;
  LogLotArea = log(LotArea); * log transform;
  LogBsmtFinSF1 = log(BsmtFinSF1+1); * log transform. +1 is added to avoid log(0);
  LogBsmtFinSF2 = log(BsmtFinSF2+1); * log transform. +1 is added to avoid log(0);
  LogBsmtUnfSF = log(BsmtUnfSF+1); * log transform. +1 is added to avoid log(0);
  LogWoodDeckSF = log(WoodDeckSF+1); * log transform. +1 is added to avoid log(0);
  LogOpenPorchSF = log(OpenPorchSF+1); * log transform. +1 is added to avoid log(0);
  if KitchenQual = "NA" then KitchenQual = "TA"; * 1 Missing value recoded to the mean of the
column. Low correlation value.;

run;

* impute the missing variables 1 time on a seed of 7;
proc mi data = q2test out = q2testimpute n impute= 1 seed = 7;
var      LogBsmtFinSF1 LogBsmtFinSF2 LogBsmtUnfSF BsmtFullBath BsmtHalfBath GarageCars
GarageArea; * variables to impute;

run;

%let int = MSSubClass LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1
LogBsmtFinSF2 LogBsmtUnfSF _1stFlrSF _2ndFlrSF LowQualFinSF BsmtFullBath BsmtHalfBath FullBath
HalfBath BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageCars GarageArea
LogWoodDeckSF LogOpenPorchSF _3SsnPorch ScreenPorch PoolArea MiscVal MoSold YrSold; *
assigning interger data to this holding variable;

%let cat = Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle
RoofMatl Exterior1st Exterior2nd ExterQual Heating HeatingQC CentralAir KitchenQual PavedDrive
SaleCondition Electrical; * assigning categorical data to this holding variable;

* Forward Model;
proc glmselect data = q2train; * regression model on training data;
  class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
  model LogSalePrice = &cat &int / selection = FORWARD(stop = Press) stats = adjrsq; * forward
model, y=x, stop at press stat and return press and adjrsqr;
  store out = work.q2forwardmodel; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2forwardmodel; * bring in model from above code;
  score data = work.q2testimpute
  out = work.q2forwardpredictions predicted; * run model on test data and save output with
predicted values;
run;

```

*** Backwards Model;**

```
proc glmselect data = q2train; * regression model on training data;
    class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
    model LogSalePrice = &cat &int / selection = BACKWARD(stop = Press) stats = adjrsq; *
backward model, y=x, stop at press stat and return press and adjrsq;
    store out = work.q2backwardmodel; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2backwardmodel; * bring in model from above code;
    score data = work.q2testimpute
    out = work.q2backwardpredictions predicted; * run model on test data and save output with
predicted values;
run;
```

*** Stepwise Model;**

```
proc glmselect data = q2train; * regression model on training data;
    class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
    model LogSalePrice = &cat &int / selection = STEPWISE(stop = Press) stats = adjrsq; * stepwise
model, y=x, stop at press stat and return press and adjrsqr;
    store out = work.q2stepmodel; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2stepmodel; * bring in model from above code;
    score data = work.q2testimpute
    out = work.q2steppredictions predicted; * run model on test data and save output with
predicted values;
run;
```

*** Custom models to show overfitting -----;**

```
%let int = MSSubClass LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1
LogBsmtFinSF2 LogBsmtUnfSF _1stFlrSF _2ndFlrSF LowQualFinSF BsmtFullBath BsmtHalfBath FullBath
HalfBath BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageCars GarageArea
LogWoodDeckSF LogOpenPorchSF _3SsnPorch ScreenPorch PoolArea MiscVal MoSold YrSold;
%let cat = Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle
RoofMatl ExterQual Heating HeatingQC CentralAir KitchenQual PavedDrive SaleCondition Electrical;
```

*** Stepwise model with a significance level of 0.1;**

```
proc glmselect data = q2train; * regression model on training data;
    class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
    model LogSalePrice = &cat &int / selection = STEPWISE(select = sl stop = sl slentry = 0.1 sls = 0.1)
stats = press;
```

```

store out = work.q2custom01model; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2custom01model; * bring in model from above code;
    score data = work.q2testimpute
    out = work.q2custom01predictions predicted; * run model on test data and save output with
predicted values;
run;

* Stepwise model with a significance level of 0.5;
proc glmselect data = q2train; * regression model on training data;
    class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
    model LogSalePrice = &cat &int / selection = STEPWISE(select = sl stop = sl slentry = 0.5 sls = 0.5)
stats = press;
    store out = work.q2custom05model; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2custom05model; * bring in model from above code;
    score data = work.q2testimpute
    out = work.q2custom05predictions predicted; * run model on test data and save output with
predicted values;
run;

* Stepwise model with a significance level of 0.99;
proc glmselect data = q2train; * regression model on training data;
    class &cat / param=glm ref=first; * SAS dummy code the categorical vars alphabetically;
    model LogSalePrice = &cat &int / selection = STEPWISE(select = sl stop = sl slentry = 0.99 sls =
0.99) stats = press;
    store out = work.q2custom099model; * output model to system to use for scoring test data;
run;

proc plm restore = work.q2custom099model; * bring in model from above code;
    score data = work.q2testimpute
    out = work.q2custom099predictions predicted; * run model on test data and save output with
predicted values;
run;

```

Appendix: 1DESC – Variable Descriptions

MSSubClass: Identifies the type of dwelling involved in the sale.

- 20 1-STORY 1946 & NEWER ALL STYLES
- 30 1-STORY 1945 & OLDER
- 40 1-STORY W/FINISHED ATTIC ALL AGES

45 1-1/2 STORY - UNFINISHED ALL AGES
50 1-1/2 STORY FINISHED ALL AGES
60 2-STORY 1946 & NEWER
70 2-STORY 1945 & OLDER
75 2-1/2 STORY ALL AGES
80 SPLIT OR MULTI-LEVEL
85 SPLIT FOYER
90 DUPLEX - ALL STYLES AND AGES
120 1-STORY PUD (Planned Unit Development) - 1946 & NEWER
150 1-1/2 STORY PUD - ALL AGES
160 2-STORY PUD - 1946 & NEWER
180 PUD - MULTILEVEL - INCL SPLIT LEV/FOYER
190 2 FAMILY CONVERSION - ALL STYLES AND AGES

MSZoning: Identifies the general zoning classification of the sale.

A Agriculture
C Commercial
FV Floating Village Residential
I Industrial
RH Residential High Density
RL Residential Low Density
RP Residential Low Density Park
RM Residential Medium Density

LotFrontage: Linear feet of street connected to property

LotArea: Lot size in square feet

Street: Type of road access to property

Grvl Gravel
Pave Paved

Alley: Type of alley access to property

Grvl Gravel
Pave Paved
NA No alley access

LotShape: General shape of property

Reg Regular
IR1 Slightly irregular
IR2 Moderately Irregular
IR3 Irregular

LandContour: Flatness of the property

Lvl Near Flat/Level
Bnk Banked - Quick and significant rise from street grade to building
HLS Hillside - Significant slope from side to side
LowDepression

Utilities: Type of utilities available

AllPub All public Utilities (E,G,W,& S)
NoSewr Electricity, Gas, and Water (Septic Tank)

NoSeWa Electricity and Gas Only

ELO Electricity only

LotConfig: Lot configuration

Inside Inside lot

Corner Corner lot

CulDSac Cul-de-sac

FR2 Frontage on 2 sides of property

FR3 Frontage on 3 sides of property

LandSlope: Slope of property

Gtl Gentle slope

Mod Moderate Slope

Sev Severe Slope

Neighborhood: Physical locations within Ames city limits

Blmngtn Bloomington Heights

Blueste Bluestem

BrDale Briardale

BrkSide Brookside

ClearCr Clear Creek

CollgCr College Creek

Crawfor Crawford

Edwards Edwards

Gilbert Gilbert

IDOTRR Iowa DOT and Rail Road

MeadowV Meadow Village

Mitchel Mitchell

Names North Ames

NoRidge Northridge

NPkVill Northpark Villa

NridgHt Northridge Heights

NWAmes Northwest Ames

OldTown Old Town

SWISU South & West of Iowa State University

Sawyer Sawyer

SawyerW Sawyer West

Somerst Somerset

StoneBr Stone Brook

Timber Timberland

Veenker Veenker

Condition1: Proximity to various conditions

Artery Adjacent to arterial street

Feedr Adjacent to feeder street

Norm Normal

RRNn Within 200' of North-South Railroad

RRAn Adjacent to North-South Railroad

PosN	Near positive off-site feature--park, greenbelt, etc.
PosA	Adjacent to postive off-site feature
RRNe	Within 200' of East-West Railroad
RRAe	Adjacent to East-West Railroad

Condition2: Proximity to various conditions (if more than one is present)

Artery	Adjacent to arterial street
Feedr	Adjacent to feeder street
Norm	Normal
RRNn	Within 200' of North-South Railroad
RRAn	Adjacent to North-South Railroad
PosN	Near positive off-site feature--park, greenbelt, etc.
PosA	Adjacent to postive off-site feature
RRNe	Within 200' of East-West Railroad
RRAe	Adjacent to East-West Railroad

BldgType: Type of dwelling

1Fam	Single-family Detached
2FmCon	Two-family Conversion; originally built as one-family dwelling
Duplx	Duplex
TwnhsE	Townhouse End Unit
TwnhsI	Townhouse Inside Unit

HouseStyle: Style of dwelling

1Story	One story
1.5Fin	One and one-half story: 2nd level finished
1.5Unf	One and one-half story: 2nd level unfinished
2Story	Two story
2.5Fin	Two and one-half story: 2nd level finished
2.5Unf	Two and one-half story: 2nd level unfinished
SFoyer	Split Foyer
SLvl	Split Level

OverallQual: Rates the overall material and finish of the house

- 10 Very Excellent
- 9 Excellent
- 8 Very Good
- 7 Good
- 6 Above Average
- 5 Average
- 4 Below Average
- 3 Fair
- 2 Poor
- 1 Very Poor

OverallCond: Rates the overall condition of the house

- 10 Very Excellent
- 9 Excellent
- 8 Very Good

- 7 Good
- 6 Above Average
- 5 Average
- 4 Below Average
- 3 Fair
- 2 Poor
- 1 Very Poor

YearBuilt: Original construction date

YearRemodAdd: Remodel date (same as construction date if no remodeling or additions)

RoofStyle: Type of roof

- | | |
|---------|---------------|
| Flat | Flat |
| Gable | Gable |
| Gambrel | Gabrel (Barn) |
| Hip | Hip |
| Mansard | Mansard |
| Shed | Shed |

RoofMatl: Roof material

- | | |
|---------|------------------------------|
| ClyTile | Clay or Tile |
| CompShg | Standard (Composite) Shingle |
| Membran | Membrane |
| Metal | Metal |
| Roll | Roll |
| Tar&Grv | Gravel & Tar |
| WdShake | Wood Shakes |
| WdShngl | Wood Shingles |

Exterior1st: Exterior covering on house

- | | |
|---------|-------------------|
| AsbShng | Asbestos Shingles |
| AsphShn | Asphalt Shingles |
| BrkComm | Brick Common |
| BrkFace | Brick Face |
| CBlock | Cinder Block |
| CemntBd | Cement Board |
| HdBoard | Hard Board |
| ImStucc | Imitation Stucco |
| MetalSd | Metal Siding |
| Other | Other |
| Plywood | Plywood |
| PreCast | PreCast |
| Stone | Stone |
| Stucco | Stucco |
| VinylSd | Vinyl Siding |
| Wd Sdng | Wood Siding |
| WdShing | Wood Shingles |

Exterior2nd: Exterior covering on house (if more than one material)

AsbShng	Asbestos Shingles
AsphShn	Asphalt Shingles
BrkComm	Brick Common
BrkFace	Brick Face
CBlock	Cinder Block
CemntBd	Cement Board
HdBoard	Hard Board
ImStucc	Imitation Stucco
MetalSd	Metal Siding
Other	Other
Plywood	Plywood
PreCast	PreCast
Stone	Stone
Stucco	Stucco
VinylSd	Vinyl Siding
Wd Sdng	Wood Siding
WdShing	Wood Shingles

MasVnrType: Masonry veneer type

BrkCmn	Brick Common
BrkFace	Brick Face
CBlock	Cinder Block
None	None
Stone	Stone

MasVnrArea: Masonry veneer area in square feet

ExterQual: Evaluates the quality of the material on the exterior

Ex	Excellent
Gd	Good
TA	Average/Typical
Fa	Fair
Po	Poor

ExterCond: Evaluates the present condition of the material on the exterior

Ex	Excellent
Gd	Good
TA	Average/Typical
Fa	Fair
Po	Poor

Foundation: Type of foundation

BrkTil	Brick & Tile
CBlock	Cinder Block
PConc	Poured Concrete
Slab	Slab
Stone	Stone
Wood	Wood

BsmtQual: Evaluates the height of the basement

Ex Excellent (100+ inches)

Gd Good (90-99 inches)

TA Typical (80-89 inches)

Fa Fair (70-79 inches)

Po Poor (<70 inches)

NA No Basement

BsmtCond: Evaluates the general condition of the basement

Ex Excellent

Gd Good

TA Typical - slight dampness allowed

Fa Fair - dampness or some cracking or settling

Po Poor - Severe cracking, settling, or wetness

NA No Basement

BsmtExposure: Refers to walkout or garden level walls

Gd Good Exposure

Av Average Exposure (split levels or foyers typically score average or above)

Mn Minimum Exposure

No No Exposure

NA No Basement

BsmtFinType1: Rating of basement finished area

GLQ Good Living Quarters

ALQAverage Living Quarters

BLQBelow Average Living Quarters

Rec Average Rec Room

LwQ Low Quality

Unf Unfinished

NA No Basement

BsmtFinSF1: Type 1 finished square feet

BsmtFinType2: Rating of basement finished area (if multiple types)

GLQ Good Living Quarters

ALQAverage Living Quarters

BLQBelow Average Living Quarters

Rec Average Rec Room

LwQ Low Quality

Unf Unfinished

NA No Basement

BsmtFinSF2: Type 2 finished square feet

BsmtUnfSF: Unfinished square feet of basement area

TotalBsmtSF: Total square feet of basement area

Heating: Type of heating

Floor Floor Furnace

GasA Gas forced warm air furnace

GasW Gas hot water or steam heat

Grav Gravity furnace

OthW Hot water or steam heat other than gas

Wall Wall furnace

HeatingQC: Heating quality and condition

Ex Excellent

Gd Good

TA Average/Typical

Fa Fair

Po Poor

CentralAir: Central air conditioning

N No

Y Yes

Electrical: Electrical system

SBrkr Standard Circuit Breakers & Romex

FuseA Fuse Box over 60 AMP and all Romex wiring (Average)

FuseF 60 AMP Fuse Box and mostly Romex wiring (Fair)

FuseP 60 AMP Fuse Box and mostly knob & tube wiring (poor)

Mix Mixed

1stFlrSF: First Floor square feet

2ndFlrSF: 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

Bedroom: Bedrooms above grade (does NOT include basement bedrooms)

Kitchen: Kitchens above grade

KitchenQual: Kitchen quality

Ex Excellent

Gd Good

TA Typical/Average

Fa Fair

Po Poor

TotRmsAbvGrd: Total rooms above grade (does not include bathrooms)

Functional: Home functionality (Assume typical unless deductions are warranted)

Typ Typical Functionality

Min1 Minor Deductions 1

Min2 Minor Deductions 2

Mod Moderate Deductions

Maj1 Major Deductions 1

Maj2 Major Deductions 2

Sev Severely Damaged

Sal Salvage only

Fireplaces: Number of fireplaces

FireplaceQu: Fireplace quality

- Ex Excellent - Exceptional Masonry Fireplace
- Gd Good - Masonry Fireplace in main level
- TA Average - Prefabricated Fireplace in main living area or Masonry Fireplace in basement
- Fa Fair - Prefabricated Fireplace in basement
- Po Poor - Ben Franklin Stove
- NA No Fireplace

GarageType: Garage location

- 2Types More than one type of garage
- Attchd Attached to home
- Basment Basement Garage
- BuiltIn Built-In (Garage part of house - typically has room above garage)
- CarPort Car Port
- Detchd Detached from home
- NA No Garage

GarageYrBlt: Year garage was built

GarageFinish: Interior finish of the garage

- Fin Finished
- RFn Rough Finished
- Unf Unfinished
- NA No Garage

GarageCars: Size of garage in car capacity

GarageArea: Size of garage in square feet

GarageQual: Garage quality

- Ex Excellent
- Gd Good
- TA Typical/Average
- Fa Fair
- Po Poor
- NA No Garage

GarageCond: Garage condition

- Ex Excellent
- Gd Good
- TA Typical/Average
- Fa Fair
- Po Poor
- NA No Garage

PavedDrive: Paved driveway

- Y Paved
- P Partial Pavement
- N Dirt/Gravel

WoodDeckSF: Wood deck area in square feet

OpenPorchSF: Open porch area in square feet

EnclosedPorch: Enclosed porch area in square feet

3SsnPorch: Three season porch area in square feet

ScreenPorch: Screen porch area in square feet

PoolArea: Pool area in square feet

PoolQC: Pool quality

Ex Excellent

Gd Good

TA Average/Typical

Fa Fair

NA No Pool

Fence: Fence quality

GdPrv Good Privacy

MnPrv Minimum Privacy

GdWo Good Wood

MnWw Minimum Wood/Wire

NA No Fence

MiscFeature: Miscellaneous feature not covered in other categories

Elev Elevator

Gar2 2nd Garage (if not described in garage section)

Othr Other

Shed Shed (over 100 SF)

TenC Tennis Court

NA None

MiscVal: \$Value of miscellaneous feature

MoSold: Month Sold (MM)

YrSold: Year Sold (YYYY)

SaleType: Type of sale

WD Warranty Deed - Conventional

CWD Warranty Deed - Cash

VWD Warranty Deed - VA Loan

New Home just constructed and sold

COD Court Officer Deed/Estate

ConContract 15% Down payment regular terms

ConLw Contract Low Down payment and low interest

ConLI Contract Low Interest

ConLD Contract Low Down

Oth Other

SaleCondition: Condition of sale

Normal Normal Sale

Abnorml Abnormal Sale - trade, foreclosure, short sale

AdjLand Adjoining Land Purchase

Alloca Allocation - two linked properties with separate deeds, typically condo with a garage

unit

Family Sale between family members

Partial Home was not completed when last assessed (associated with New Homes)

Appendix: 1DEF – Review of equation for a line

Definitions and review:

Simple equation of a line:

$$y = mx + b$$

Where:

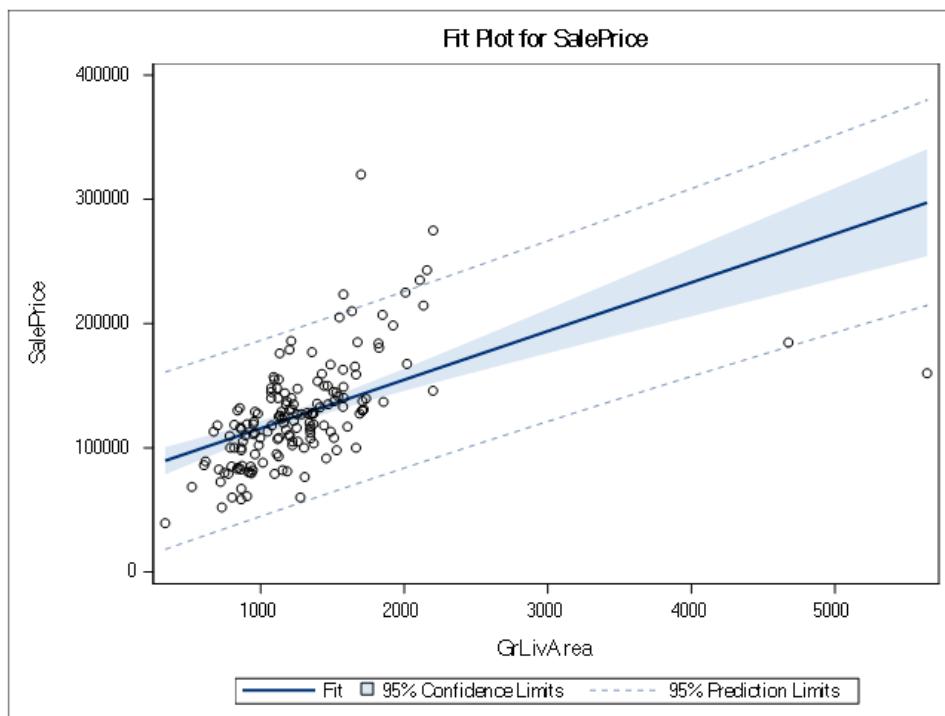
y = Variable we are solving for. Sales price in this case.

m = Slope of the line

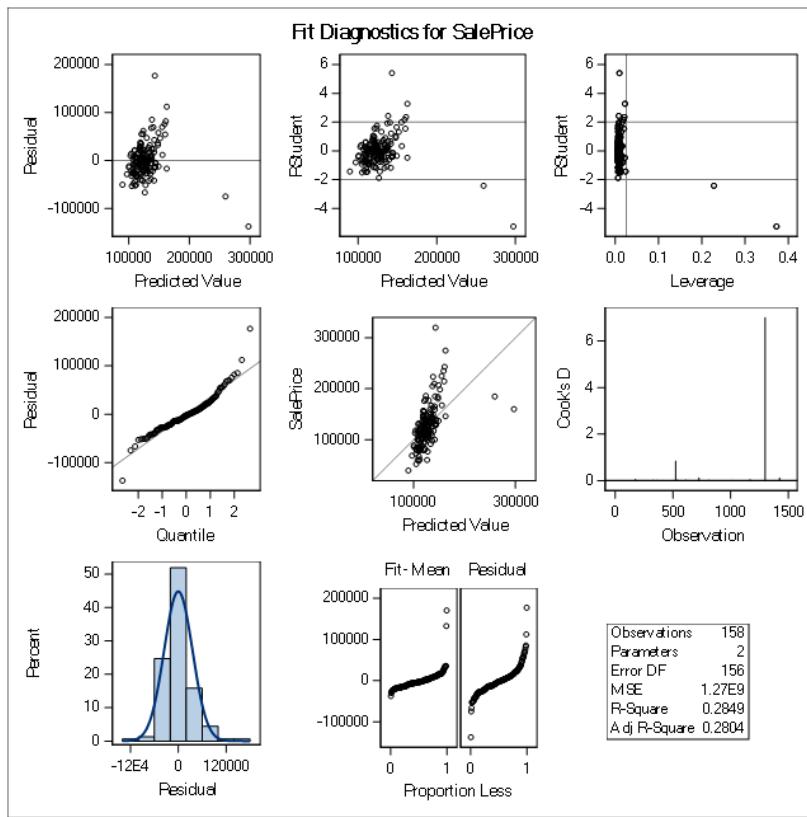
x = Explanatory variable. Greater living area in this case.

b = y Intercept

1A – Scatterplot for original data



2A – Fit diagnostic plots for original data



3A – Simple Linear Regression Model Output

Number of Observations Read	381
Number of Observations Used	381

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	12.87934798	12.87934798	314.01	<.0001
Error	379	15.54489640	0.04101556		
Corrected Total	380	28.42424438			

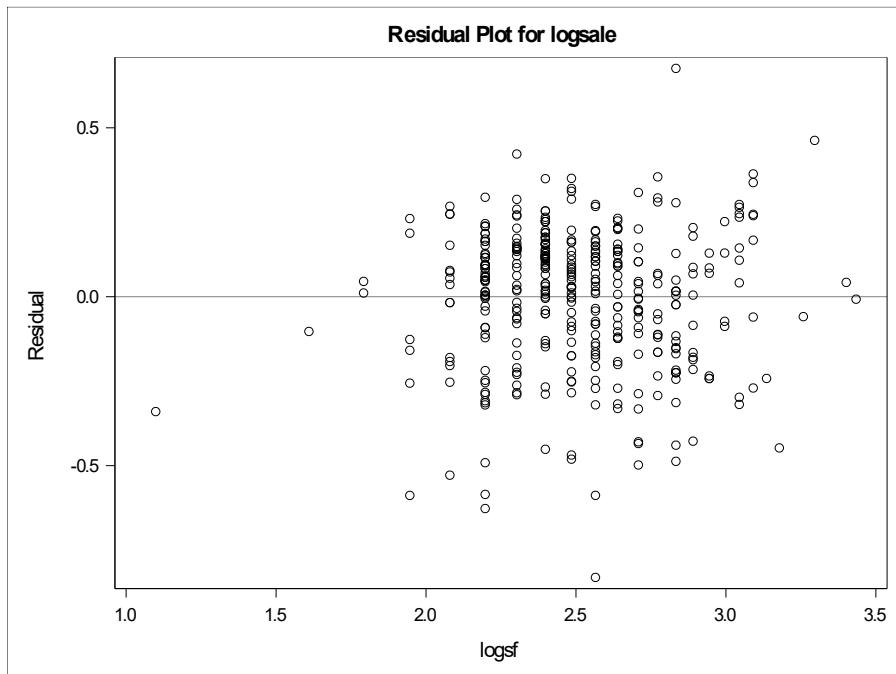
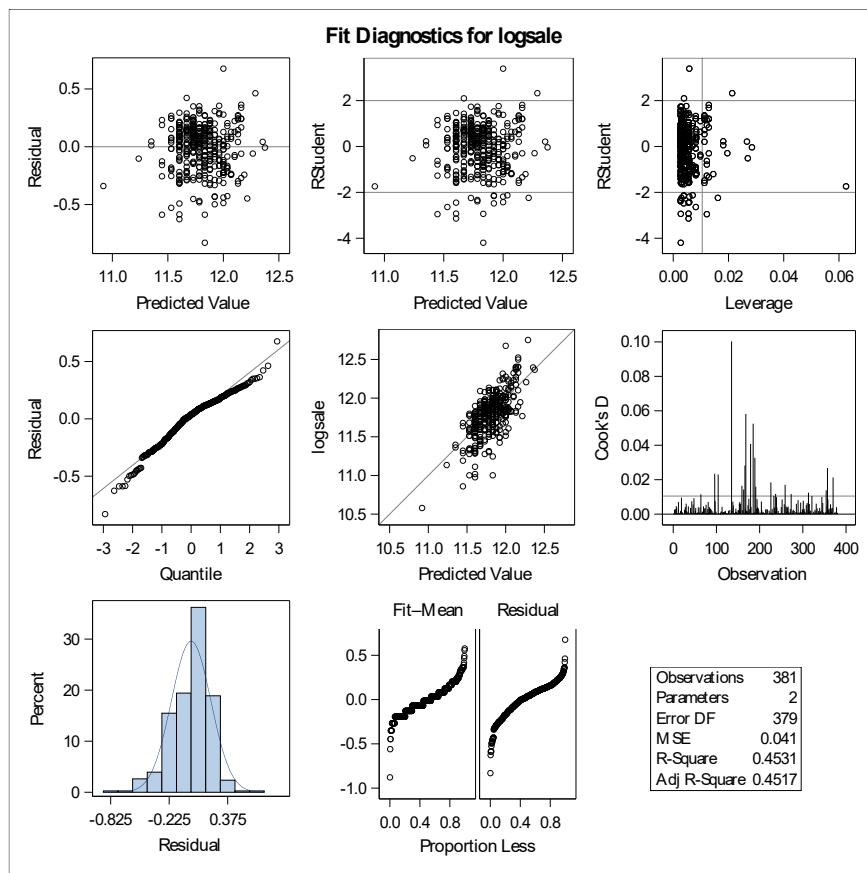
R-Square	Coeff Var	Root MSE	logscale Mean
0.453111	1.716657	0.202523	11.79752

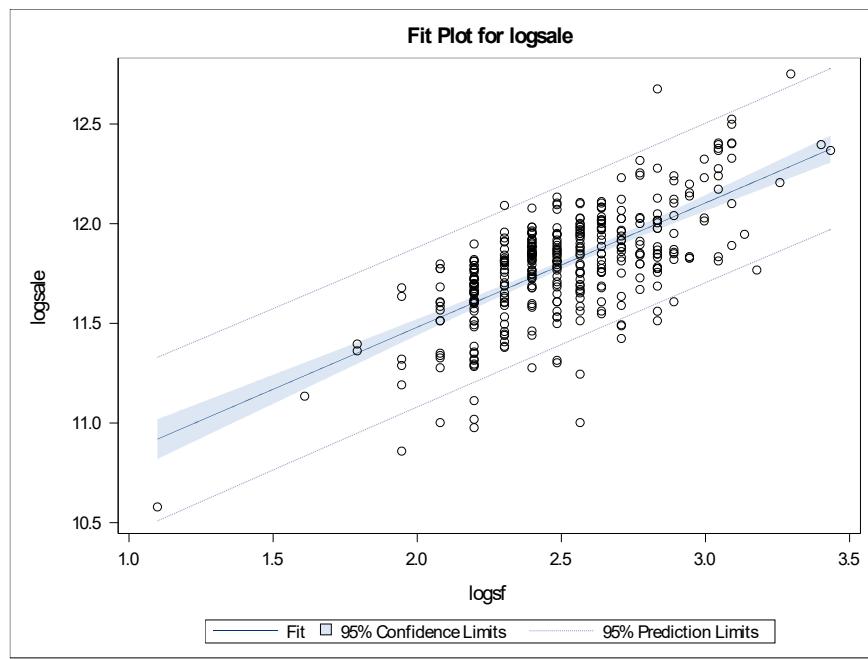
Source	DF	Type I SS	Mean Square	F Value	Pr > F
logsf	1	12.87934798	12.87934798	314.01	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
logsf	1	12.87934798	12.87934798	314.01	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > t	95% Confidence Limits
Intercept	10.23421558	0.08882891	115.21	<.0001	10.05955636 10.40887480
logsf	0.62330269	0.03517437	17.72	<.0001	0.55414132 0.69246405

Sum of Residuals	-0.00000000
Sum of Squared Residuals	15.54489640
Sum of Squared Residuals - Error SS	-0.00000000
PRESS Statistic	15.73478070
First Order Autocorrelation	-0.00077239
Durbin-Watson D	2.00107845





4A – Multiple Linear Regression Model

Class Level Information		
Class	Levels	Values
NeighborhoodDum	4	BrkSide Edwards NAmes Others

Number of Observations Read	1458
Number of Observations Used	1458

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	135.1822645	19.3117521	286.89	<.0001
Error	145 0	97.6061520	0.0673146		
Corrected Total	145 7	232.7884165			

R-Square	Coeff Var	Root MSE	logsale Mean
0.580709	2.157771	0.259451	12.02401

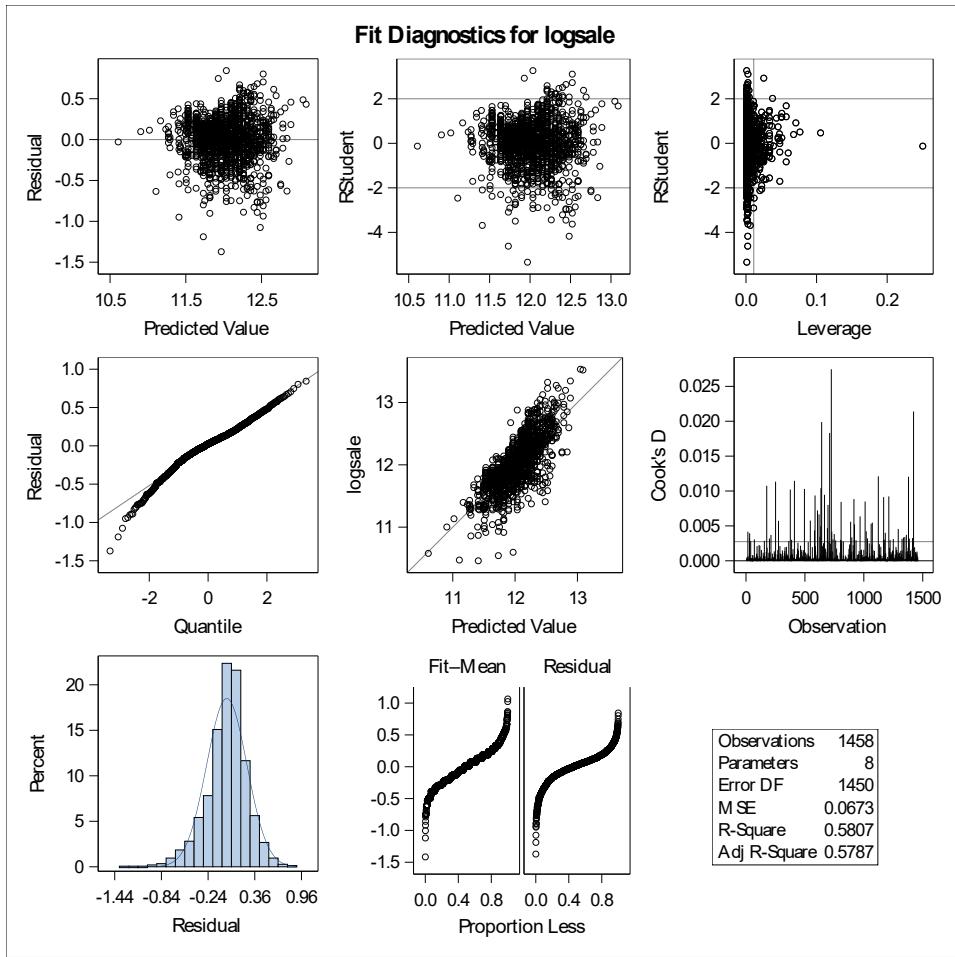
Source	DF	Type I SS	Mean Square	F Value	Pr > F
logsf	1	126.353440 1	126.3534401	1877.06	<.0001
NeighborhoodDum	3	5.9407069	1.9802356	29.42	<.0001
logsf*NeighborhoodDu	3	2.8881174	0.9627058	14.30	<.0001

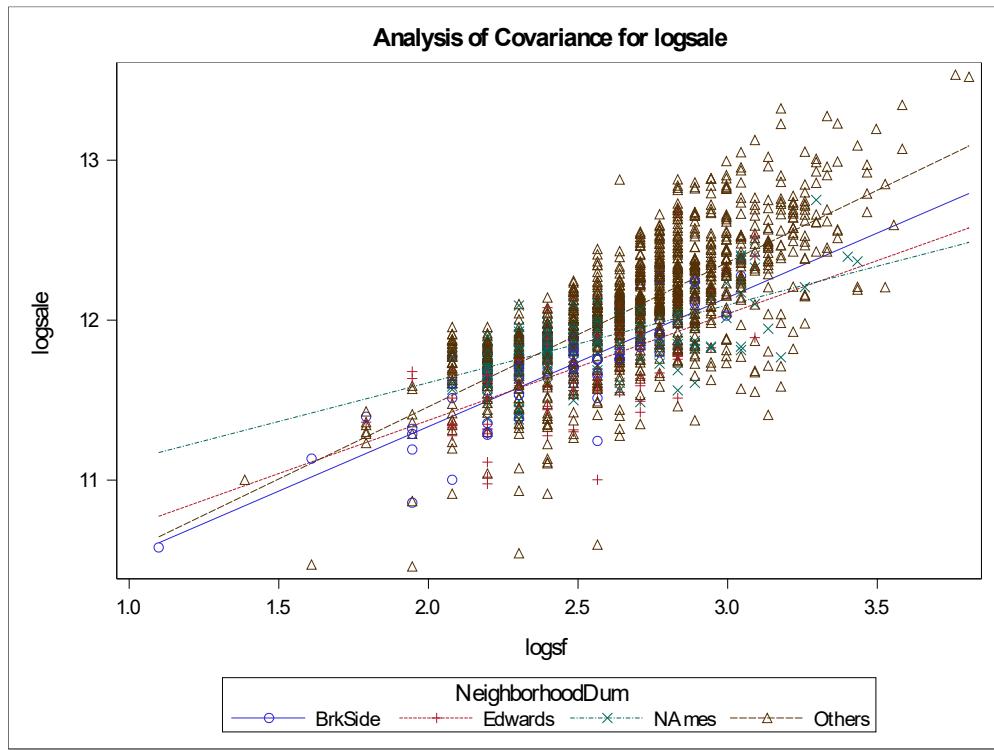
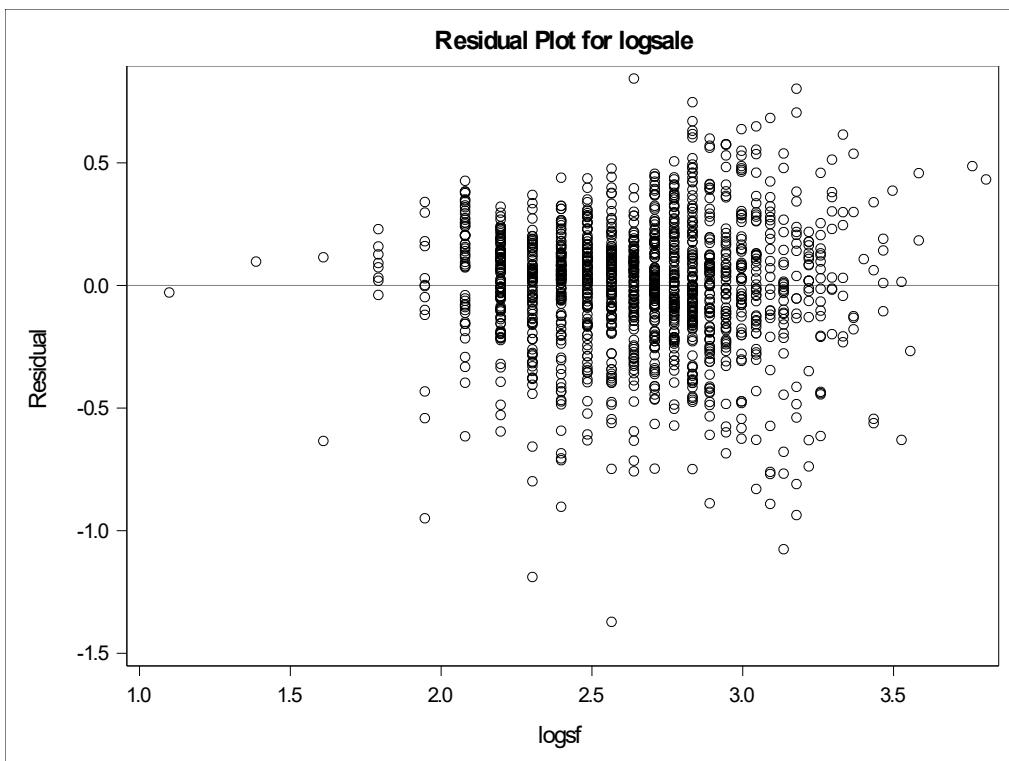
Source	DF	Type III SS	Mean Square	F Value	Pr > F
logsf	1	24.7364376 7	24.73643767	367.48	<.0001
NeighborhoodDum	3	2.30374292	0.76791431	11.41	<.0001
logsf*NeighborhoodDu	3	2.88811741	0.96270580	14.30	<.0001

Parameter	Estimate		Standard Error	t Value	Pr > t	95% Confidence Limits	
Intercept	9.653638182	B	0.06619704	145.83	<.0001	9.523785983	9.783490381
logsf	0.902392907	B	0.02420256	37.29	<.0001	0.854917132	0.949868683
NeighborhoodDum BrkSide	0.067805549	B	0.24064497	0.28	0.7782	-0.404243955	0.539855053
NeighborhoodDum Edwards	0.388874001	B	0.24672761	1.58	0.1152	-0.095107209	0.872855211
NeighborhoodDum NAmes	0.985375736	B	0.17127914	5.75	<.0001	0.649394331	1.321357141
NeighborhoodDum Others	0.000000000	B
logsf*NeighborhoodDu BrkSide	-0.095771889	B	0.09731694	-0.98	0.3252	-0.286668926	0.095125147
logsf*NeighborhoodDu Edwards	-0.236690893	B	0.09762269	-2.42	0.0154	-0.428187688	-0.045194098
logsf*NeighborhoodDu NAmes	-0.417239913	B	0.06654046	-6.27	<.0001	-0.547765779	-0.286714047
logsf*NeighborhoodDu Others	0.000000000	B

Note The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations.
 Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Sum of Residuals	0.00000000
Sum of Squared Residuals	97.60615198
Sum of Squared Residuals - Error SS	0.00000000
PRESS Statistic	98.46396722
First Order Autocorrelation	-0.01661957
Durbin-Watson D	2.03298942





1B – Original Data SalePrice Fit Diagnostics

Number of Observations Read	1458
Number of Observations Used	581
Number of Observations with Missing Values	877

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	73	4.498457E12	61622701566	67.73	<.0001
Error	507	4.612664E11	909795645		
Corrected Total	580	4.959724E12			

Root MSE	30163	R-Square	0.9070
Dependent Mean	222933	Adj R-Sq	0.8936
Coeff Var	13.53000		

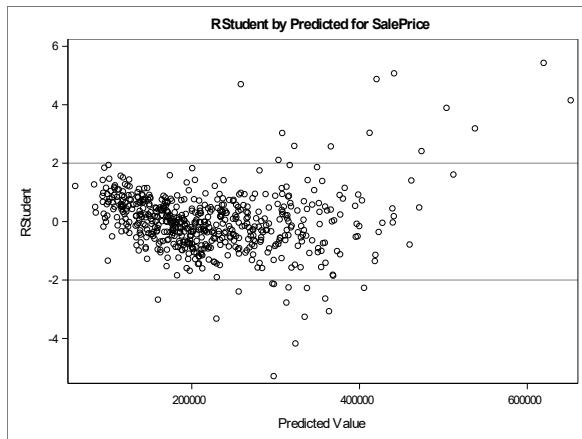
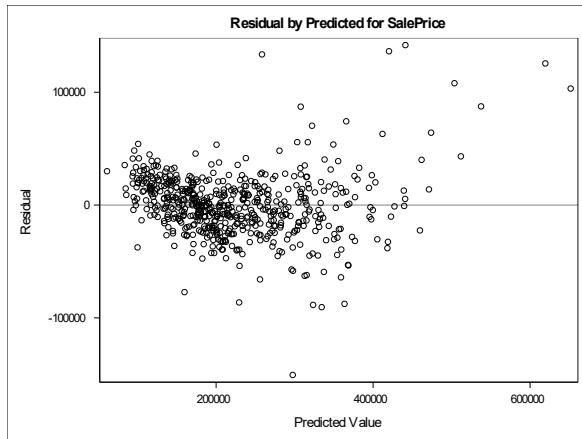
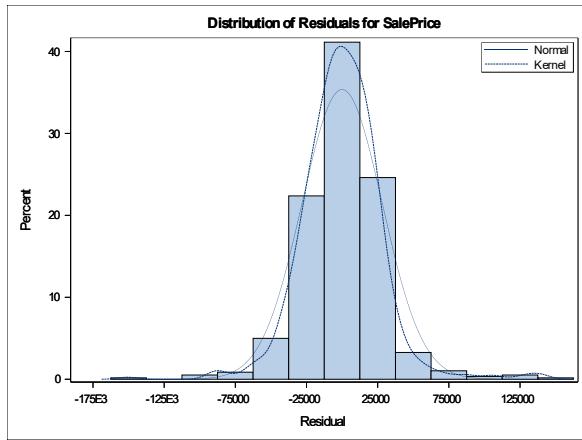
TotalBsmtSF	= -2.83E-7 * Intercept + BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF
GrLivArea	= 2.89E-7 * Intercept + X1stFlrSF + X2ndFlrSF + LowQualFinSF

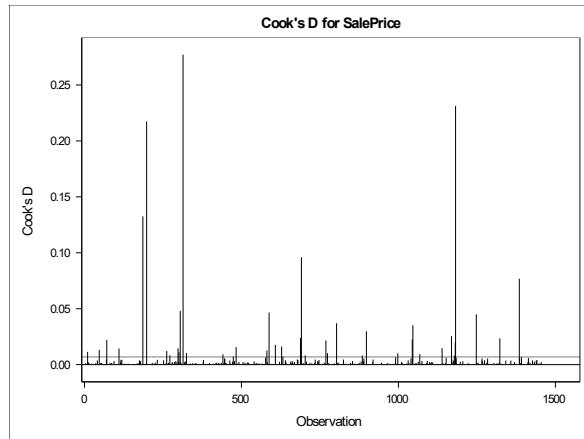
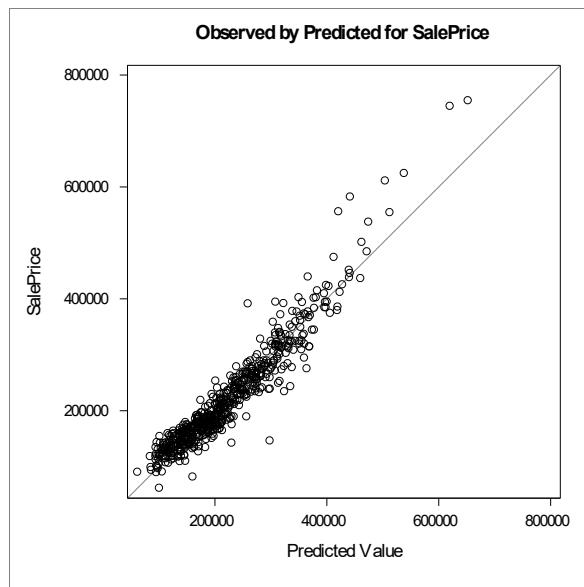
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	B	2904350	2098441	1.38	0.1670
MSSubClass	1	-110.14580	103.08891	-1.07	0.2858
LotArea	1	0.89138	0.15647	5.70	<.0001
OverallQual	1	11960	2057.97152	5.81	<.0001
OverallCond	1	8918.09765	1973.19342	4.52	<.0001
YearBuilt	1	192.62899	149.35072	1.29	0.1977
YearRemodAdd	1	27.91213	132.88459	0.21	0.8337
MasVnrArea	1	29.69519	7.62647	3.89	0.0001
BsmtFinSF1	B	58.41861	9.70167	6.02	<.0001
BsmtFinSF2	B	29.72473	14.46001	2.06	0.0403
BsmtUnfSF	B	23.44903	9.08966	2.58	0.0102

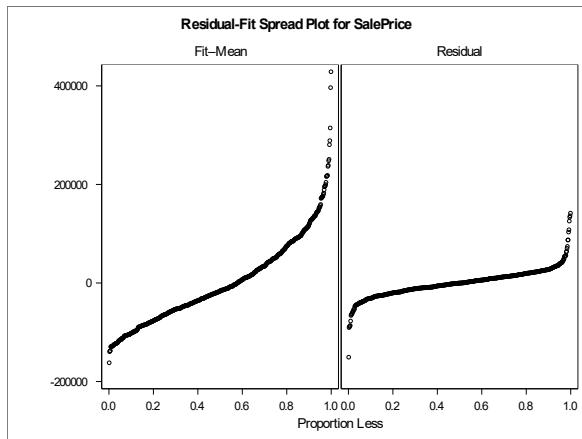
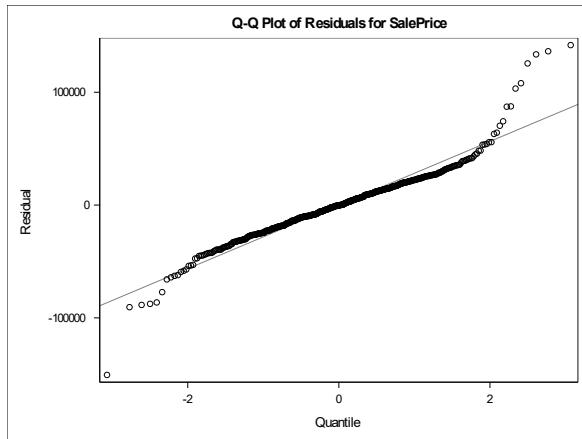
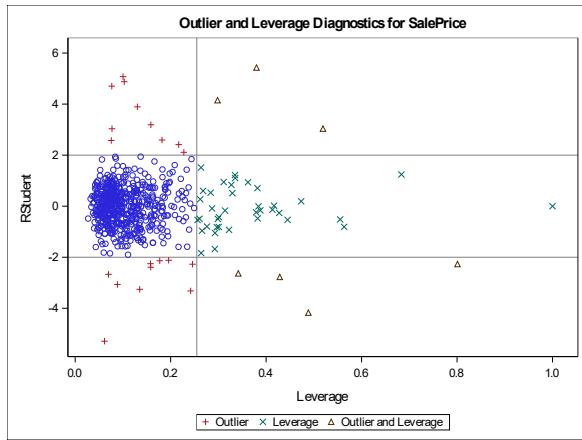
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
TotalBsmtSF	0	0	.	.	.
X1stFlrSF	B	11.76129	47.44936	0.25	0.8043
X2ndFlrSF	B	11.72768	47.05270	0.25	0.8033
LowQualFinSF	B	-63.70970	58.50359	-1.09	0.2767
GrLivArea	0	0	.	.	.
BsmtFullBath	1	-5812.44362	4068.43189	-1.43	0.1537
BsmtHalfBath	1	-5112.56244	5536.62740	-0.92	0.3562
FullBath	1	-1231.14233	4374.13558	-0.28	0.7785
HalfBath	1	-601.93606	4228.71058	-0.14	0.8869
BedroomAbvGr	1	-8938.96165	2820.40475	-3.17	0.0016
KitchenAbvGr	1	-34439	15948	-2.16	0.0313
TotRmsAbvGrd	1	-516.76996	1788.19161	-0.29	0.7727
Fireplaces	1	4282.45448	4098.58498	1.04	0.2966
GarageYrBlt	1	123.44322	130.33045	0.95	0.3440
GarageCars	1	10190	4450.65846	2.29	0.0225
GarageArea	1	-5.08729	14.50710	-0.35	0.7260
WoodDeckSF	1	1.03934	11.56593	0.09	0.9284
OpenPorchSF	1	23.81393	22.05428	1.08	0.2808
EnclosedPorch	1	-2.84981	25.20903	-0.11	0.9100
X3SsnPorch	1	45.21509	45.04215	1.00	0.3159
ScreenPorch	1	15.57054	20.50846	0.76	0.4481
PoolArea	1	44.51170	26.85334	1.66	0.0980
MiscVal	1	-0.78069	7.25569	-0.11	0.9144
MoSold	1	-542.68224	485.89102	-1.12	0.2646
YrSold	1	-1709.59756	1040.65709	-1.64	0.1010
MSZoning	1	-2087.76182	2986.94209	-0.70	0.4849
Street	1	27158	32988	0.82	0.4107
LotShape	1	265.03658	1034.11070	0.26	0.7978
LandContour	1	-2733.76860	2077.14800	-1.32	0.1887

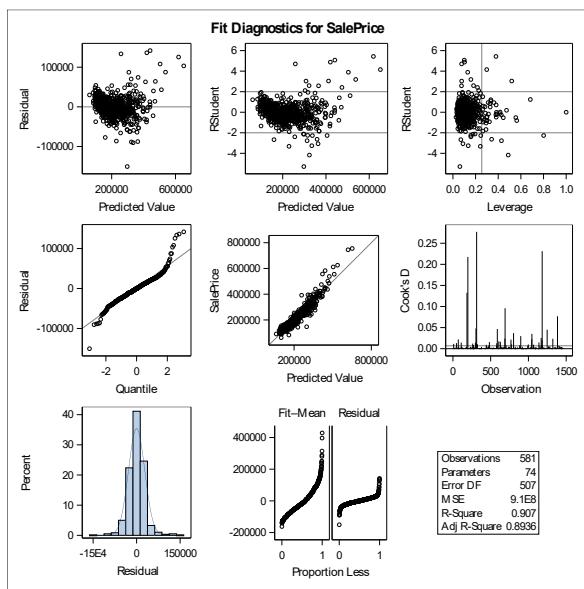
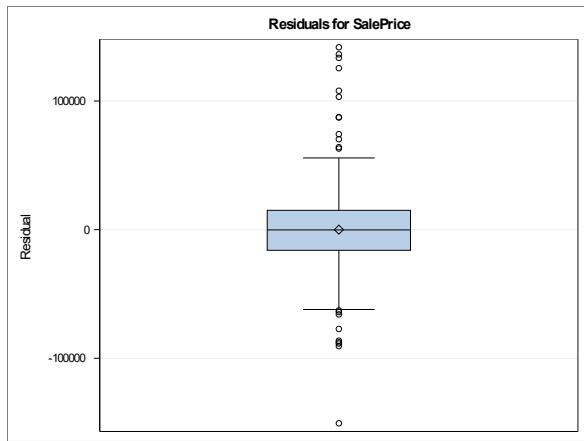
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
LotConfig	1	222.19729	878.90470	0.25	0.8005
LandSlope	1	-12391	6070.92947	-2.04	0.0418
Neighborhood	1	-35.04107	248.41206	-0.14	0.8879
Condition1	1	-358.76213	1590.55349	-0.23	0.8216
Condition2	1	-31831	11428	-2.79	0.0055
BldgType	1	-832.46849	2905.04531	-0.29	0.7746
HouseStyle	1	1509.85604	1199.44769	1.26	0.2087
RoofStyle	1	782.75164	1694.35229	0.46	0.6443
RoofMatl	1	1911.65253	2028.07813	0.94	0.3463
Exterior1st	1	-143.61708	858.83900	-0.17	0.8673
Exterior2nd	1	-637.96498	755.94271	-0.84	0.3991
MasVnrType	1	8568.78391	2164.53908	3.96	<.0001
ExterQual	1	-10406	2937.49840	-3.54	0.0004
ExterCond	1	2386.76699	2300.83707	1.04	0.3001
Foundation	1	3146.12360	3334.37122	0.94	0.3459
Heating	1	-38741	10245	-3.78	0.0002
HeatingQC	1	360.25269	1026.76736	0.35	0.7258
CentralAir	1	1105.54181	14021	0.08	0.9372
Electrical	1	-1951.66463	1928.73012	-1.01	0.3121
KitchenQual	1	-4014.93437	2142.00173	-1.87	0.0615
Functional	1	1850.43191	1675.50706	1.10	0.2699
GarageType	1	3523.10864	1230.02498	2.86	0.0044
GarageFinish	1	1896.59910	2237.08664	0.85	0.3970
GarageQual	1	-3318.15554	2863.50181	-1.16	0.2471
GarageCond	1	-4042.04227	4017.39200	-1.01	0.3148
PavedDrive	1	-1373.43963	4892.93369	-0.28	0.7791
SaleType	1	-1760.52412	1098.63932	-1.60	0.1097
SaleCondition	1	4332.93426	1550.93837	2.79	0.0054
sf_hundreds	1	6997.86411	4639.68355	1.51	0.1321

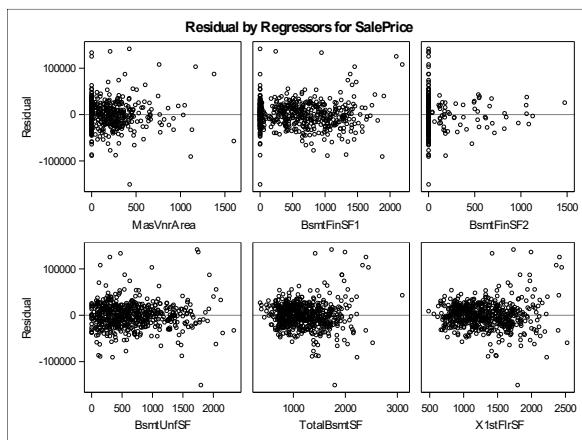
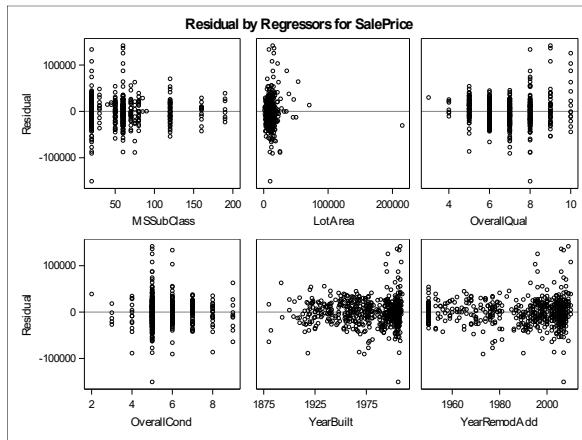
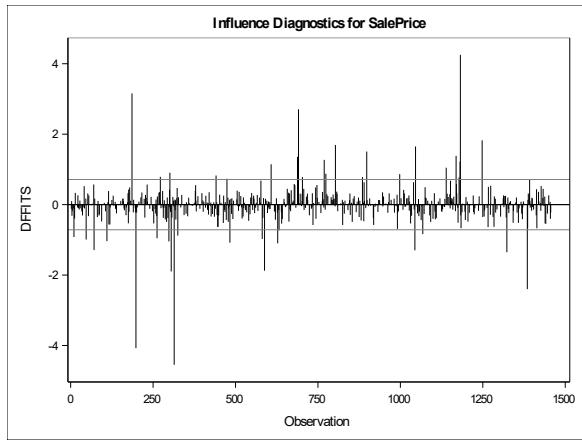
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
NumLotFrontage	1	-240.88856	57.21384	-4.21	<.0001
NumBsmtQual	1	-6935.05334	2078.98587	-3.34	0.0009
NumBsmtCond	1	2077.90477	2124.93613	0.98	0.3286
NumBsmtExposure	1	-3726.44861	1321.39823	-2.82	0.0050
NumBsmtFinType1	1	1646.16174	1007.25844	1.63	0.1028
NumBsmtFinType2	1	-1045.69729	2257.32547	-0.46	0.6434
NumFireplaceQu	1	-900.96545	1380.05023	-0.65	0.5141

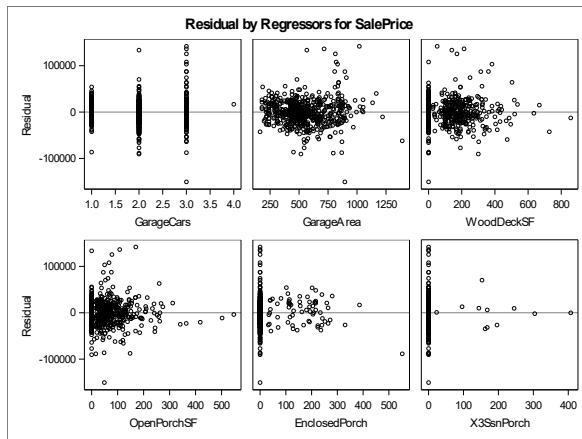
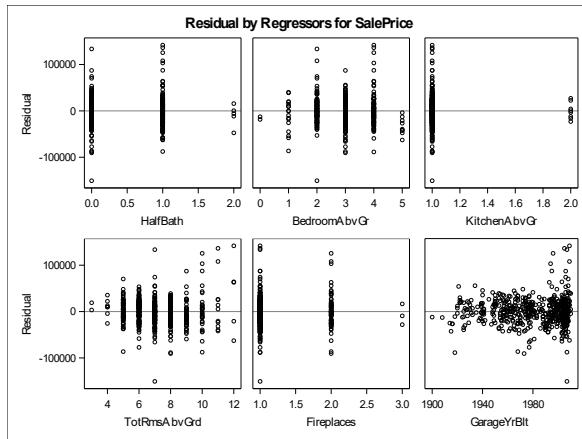
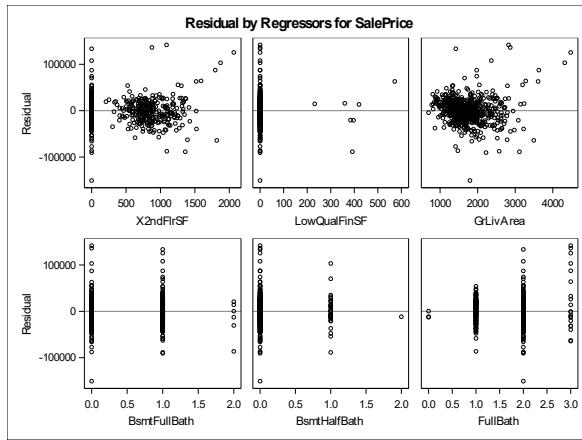


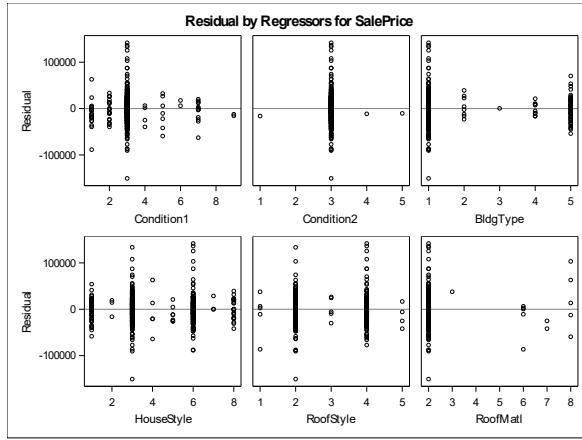
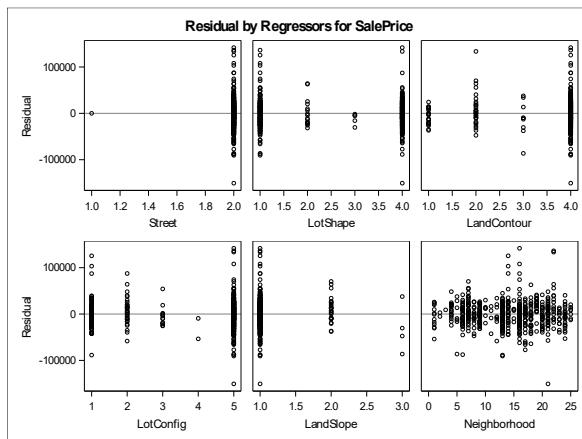
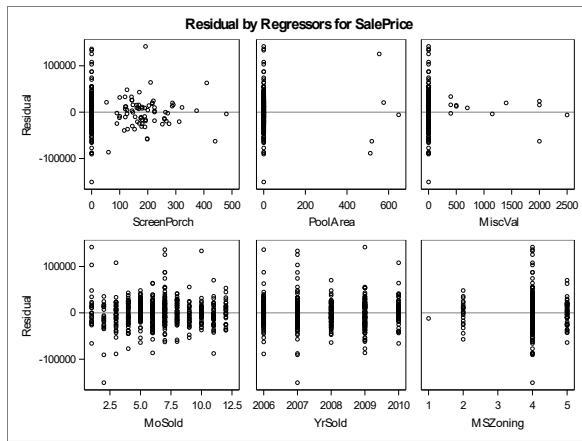


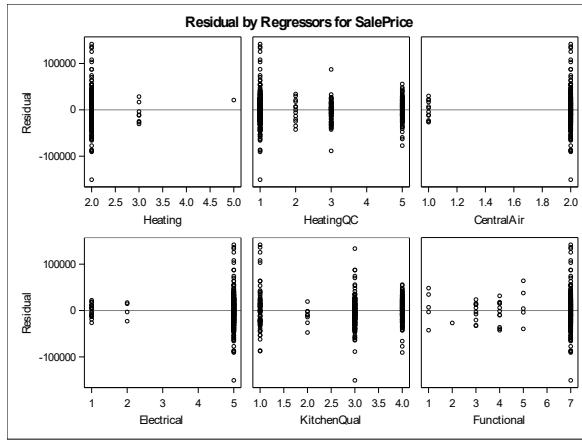
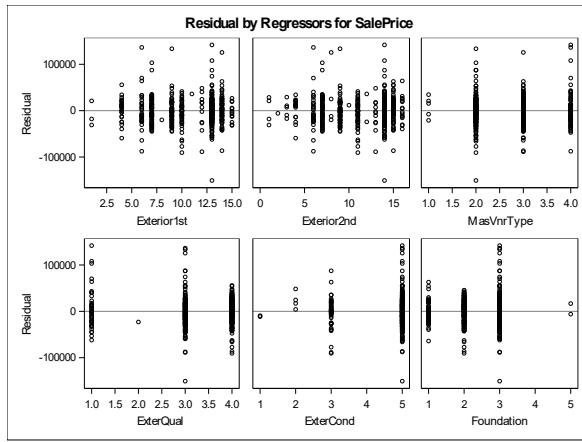


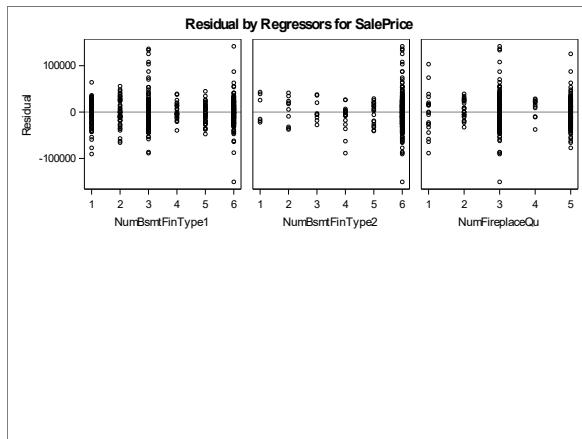
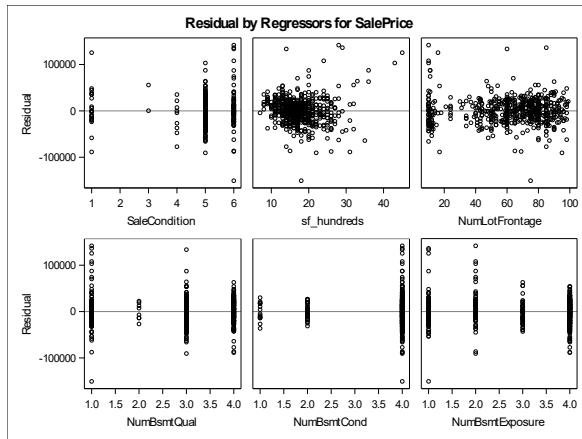
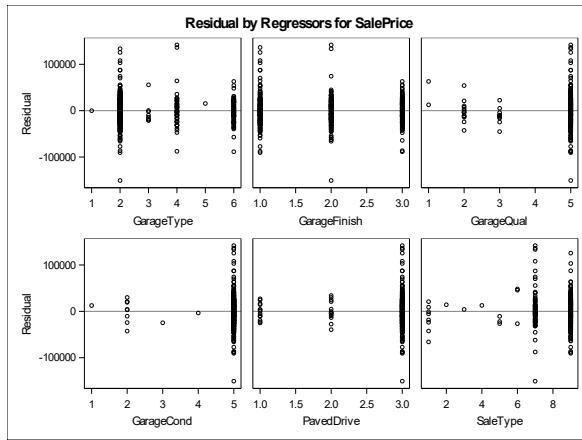












2B – Transformed Data LogSalePrice Fit Diagnostics

Number of Observations Read	145 8
Number of Observations Used	581
Number of Observations with Missing Values	877

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	74	74.9177 5	1.0124 0	82.88	<.0001
Error	50 6	6.18064	0.0122 1		
Corrected Total	58 0	81.0983 9			

Root MSE	0.11052	R-Square	0.923 8
Dependent Mean	12.2416 3	Adj R-Sq	0.912 6
Coeff Var	0.90282		

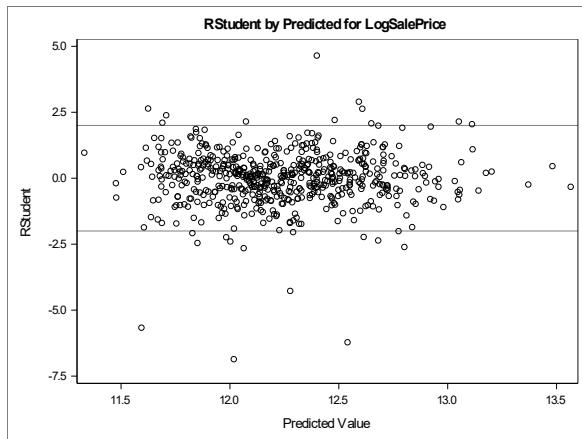
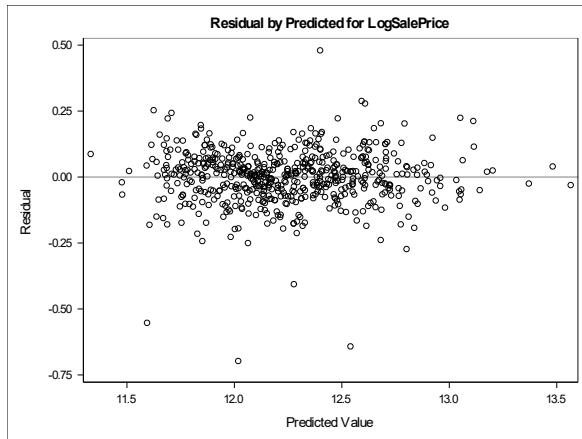
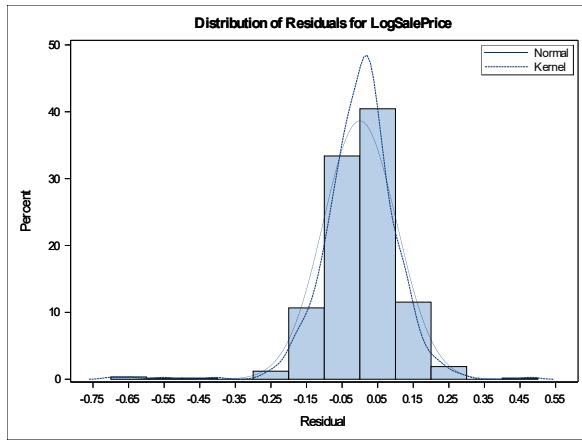
GrLivArea =	2.95E-7 * Intercept + X1stFlrSF + X2ndFlrSF + LowQualFinSF
--------------------	--

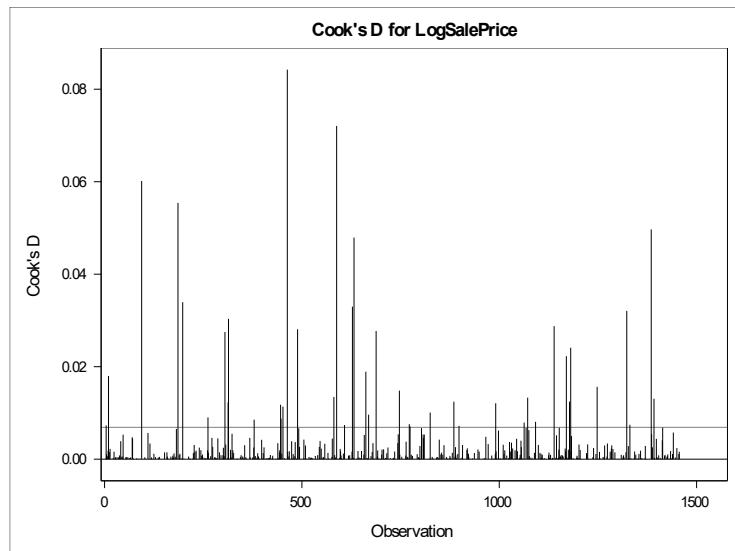
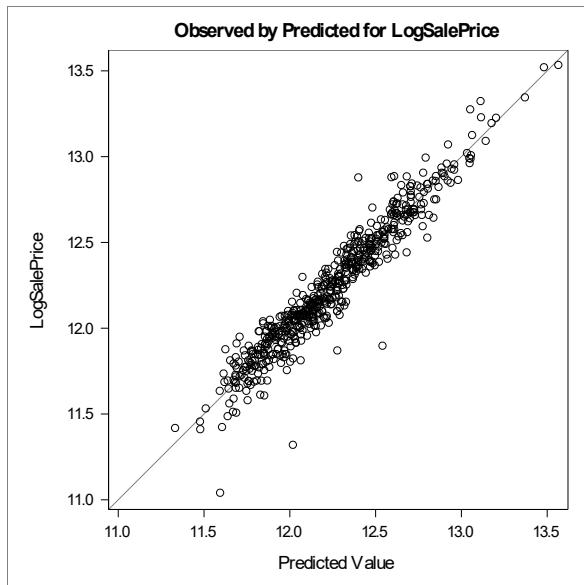
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	B	22.86983	7.71951	2.96	0.0032
MSSubClass	1	-0.00006455	0.00037982	-0.17	0.8651
LogLotArea	1	0.10726	0.01594	6.73	<.0001
OverallQual	1	0.05953	0.00757	7.86	<.0001
OverallCond	1	0.04870	0.00725	6.72	<.0001
YearBuilt	1	0.00108	0.00055412	1.95	0.0516
YearRemodAdd	1	0.00092467	0.00048689	1.90	0.0581
LogMasVnrArea	1	0.00512	0.00245	2.09	0.0369
LogBsmtFinSF1	1	0.00869	0.00362	2.40	0.0168

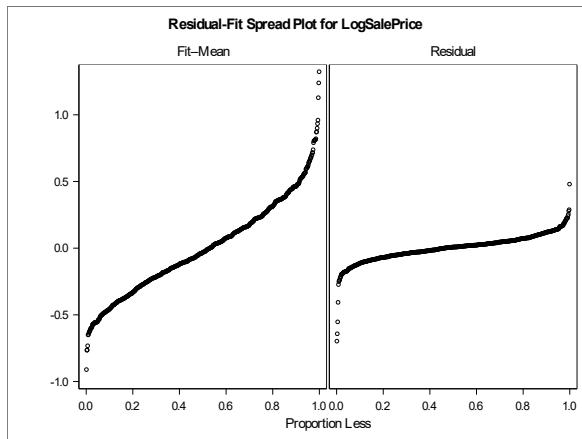
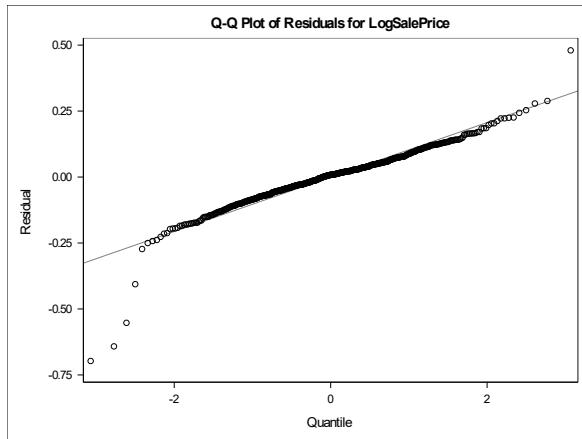
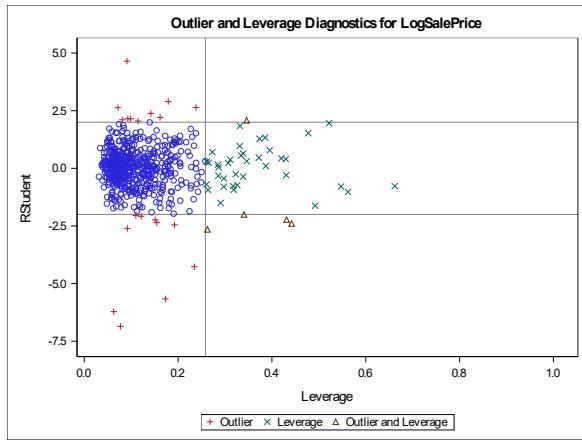
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
LogBsmtFinSF2	1	0.00122	0.00586	0.21	0.8359
LogBsmtUnfSF	1	-0.00105	0.00520	-0.20	0.8408
TotalBsmtSF	1	0.00009086	0.00003407	2.67	0.0079
X1stFlrSF	B	0.00034025	0.00017383	1.96	0.0509
X2ndFlrSF	B	0.00028139	0.00017240	1.63	0.1033
LowQualFinSF	B	0.00004858	0.00021399	0.23	0.8205
GrLivArea	0	0	.	.	.
BsmtFullBath	1	0.03092	0.01392	2.22	0.0268
BsmtHalfBath	1	-0.01010	0.02000	-0.51	0.6137
FullBath	1	0.01696	0.01617	1.05	0.2946
HalfBath	1	0.02540	0.01559	1.63	0.1038
BedroomAbvGr	1	-0.03695	0.01019	-3.62	0.0003
KitchenAbvGr	1	-0.09611	0.05862	-1.64	0.1017
TotRmsAbvGrd	1	0.00250	0.00654	0.38	0.7024
Fireplaces	1	0.01596	0.01503	1.06	0.2887
GarageYrBlt	1	-0.00023081	0.00048153	-0.48	0.6319
GarageCars	1	0.05964	0.01634	3.65	0.0003
GarageArea	1	0.00000477	0.00005357	0.09	0.9291
LogWoodDeckSF	1	0.00261	0.00226	1.16	0.2482
LogOpenPorchSF	1	0.00604	0.00287	2.10	0.0361
EnclosedPorch	1	0.00008733	0.00009253	0.94	0.3457
X3SsnPorch	1	0.00026593	0.00016613	1.60	0.1101
ScreenPorch	1	0.00016691	0.00007494	2.23	0.0264
PoolArea	1	-0.00002312	0.00009855	-0.23	0.8146
MiscVal	1	-0.00000666	0.00002665	-0.25	0.8027
MoSold	1	0.00064393	0.00178	0.36	0.7181
YrSold	1	-0.00823	0.00382	-2.15	0.0318
MSZoning	1	-0.03237	0.01113	-2.91	0.0038
Street	1	0.08767	0.12016	0.73	0.4660

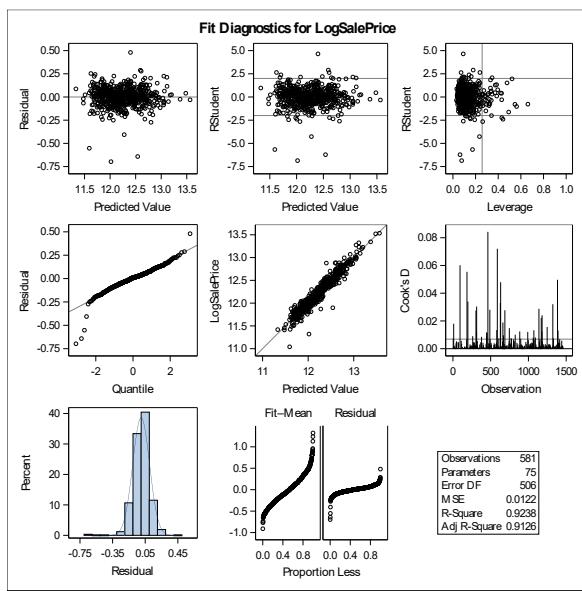
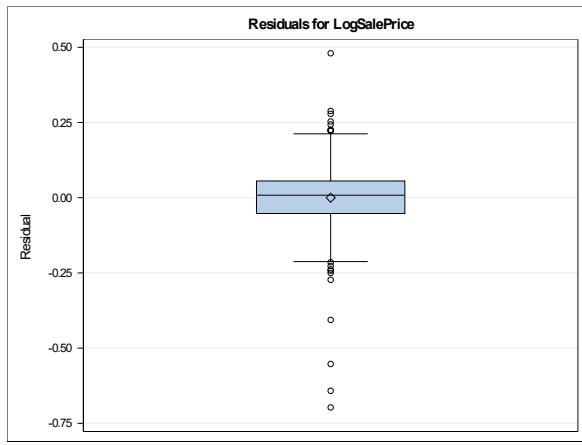
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
LotShape	1	0.00062827	0.00387	0.16	0.8710
LandContour	1	-0.01268	0.00769	-1.65	0.0997
LotConfig	1	-0.00029643	0.00325	-0.09	0.9273
LandSlope	1	-0.02092	0.02142	-0.98	0.3293
Neighborhood	1	-0.00095528	0.00091133	-1.05	0.2950
Condition1	1	0.00679	0.00586	1.16	0.2471
Condition2	1	-0.05331	0.04194	-1.27	0.2043
BldgType	1	0.00726	0.01090	0.67	0.5054
HouseStyle	1	0.00318	0.00439	0.72	0.4700
RoofStyle	1	0.00080501	0.00619	0.13	0.8966
RoofMatl	1	-0.01401	0.00746	-1.88	0.0612
Exterior1st	1	-0.00156	0.00316	-0.49	0.6213
Exterior2nd	1	-0.00006457	0.00279	-0.02	0.9816
MasVnrType	1	0.02627	0.00811	3.24	0.0013
ExterQual	1	-0.01618	0.01077	-1.50	0.1335
ExterCond	1	0.00474	0.00837	0.57	0.5716
Foundation	1	0.02548	0.01238	2.06	0.0401
Heating	1	-0.06166	0.03685	-1.67	0.0948
HeatingQC	1	-0.00281	0.00378	-0.74	0.4583
CentralAir	1	0.08340	0.05155	1.62	0.1063
Electrical	1	-0.00609	0.00709	-0.86	0.3907
KitchenQual	1	-0.00505	0.00785	-0.64	0.5204
Functional	1	0.01509	0.00610	2.47	0.0138
GarageType	1	0.00973	0.00457	2.13	0.0336
GarageFinish	1	-0.00391	0.00818	-0.48	0.6330
GarageQual	1	-0.01261	0.01050	-1.20	0.2303
GarageCond	1	-0.00775	0.01473	-0.53	0.5989
PavedDrive	1	-0.00066585	0.01816	-0.04	0.9708
SaleType	1	-0.00375	0.00403	-0.93	0.3529

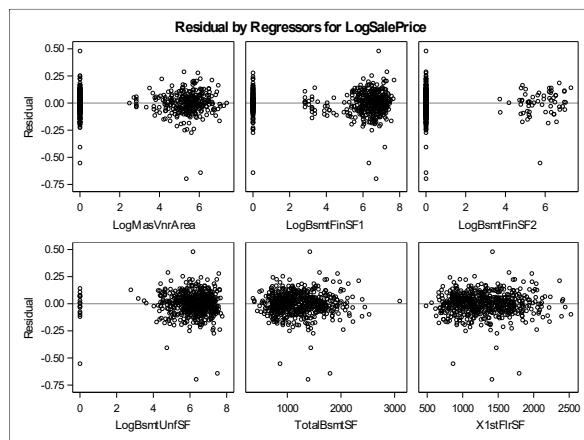
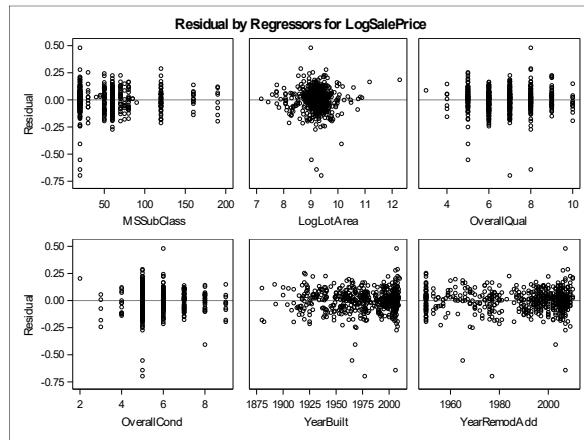
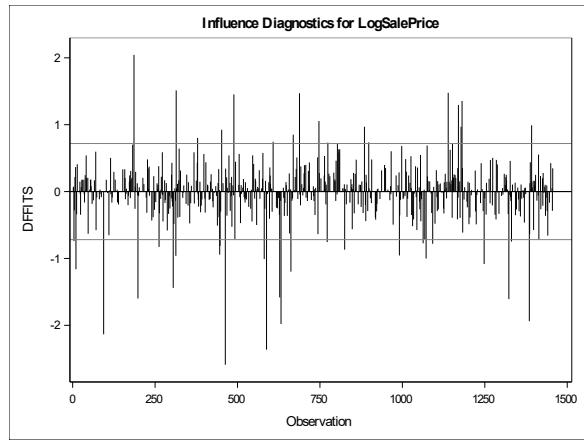
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
SaleCondition	1	0.02480	0.00571	4.34	<.0001
sf_hundreds	1	-0.00192	0.01698	-0.11	0.9098
NumLotFrontage	1	-0.00024060	0.00021044	-1.14	0.2534
NumBsmtQual	1	-0.02255	0.00758	-2.98	0.0031
NumBsmtCond	1	0.01137	0.00777	1.46	0.1440
NumBsmtExposure	1	-0.00567	0.00493	-1.15	0.2511
NumBsmtFinType1	1	-0.00086898	0.00518	-0.17	0.8669
NumBsmtFinType2	1	0.00877	0.01148	0.76	0.4454
NumFireplaceQu	1	-0.00315	0.00504	-0.62	0.5323

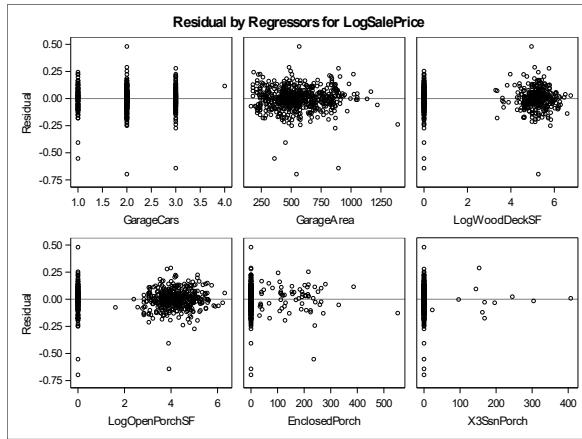
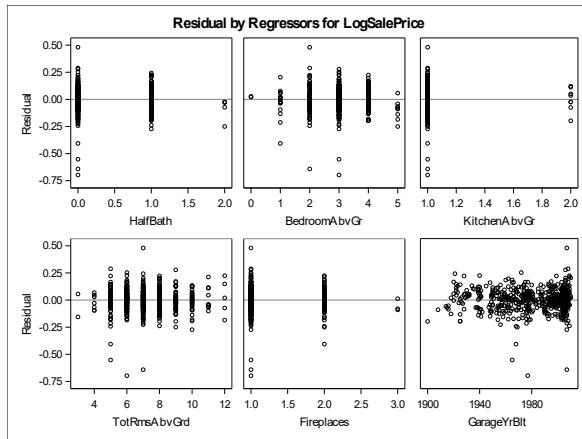
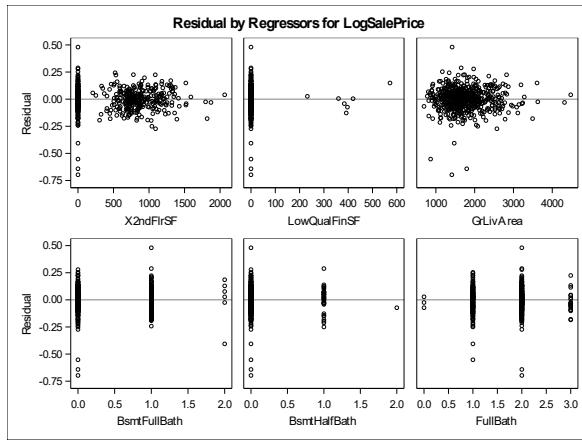


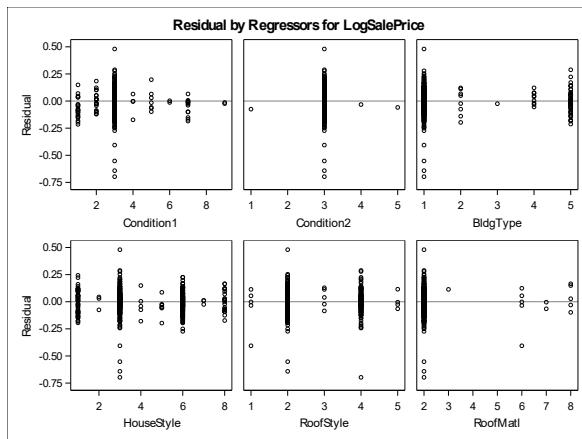
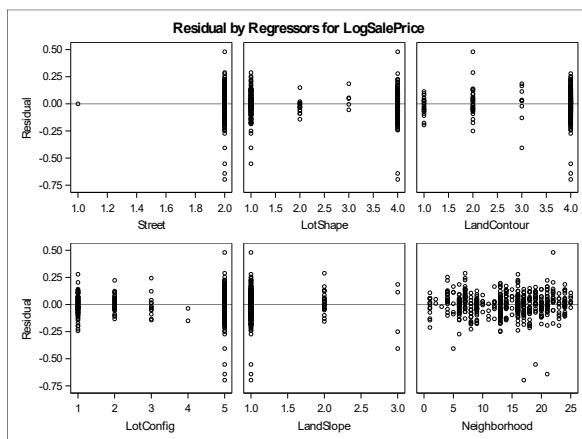
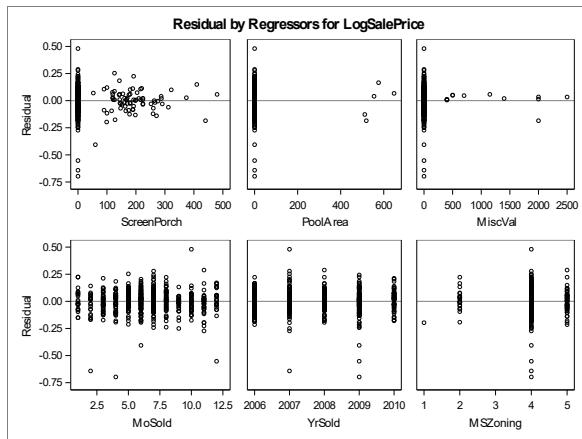


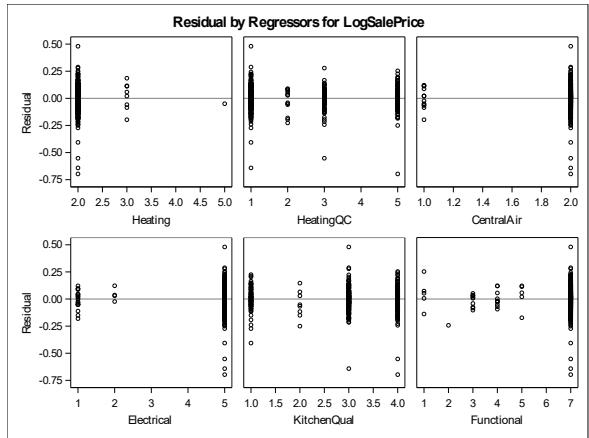
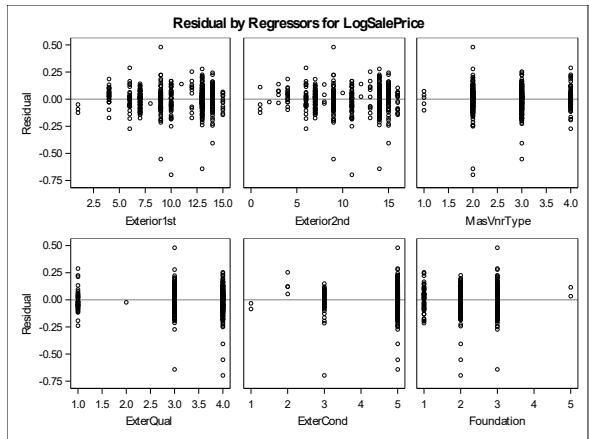


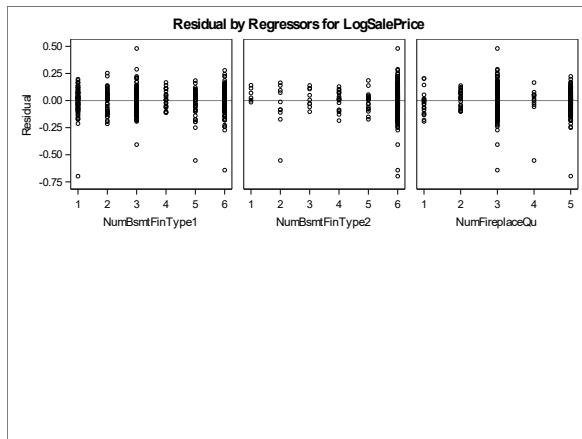
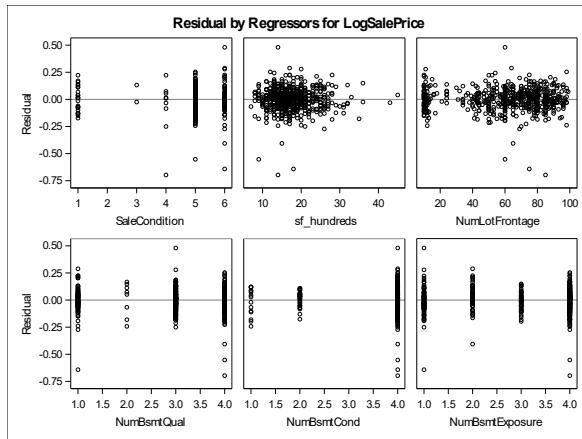
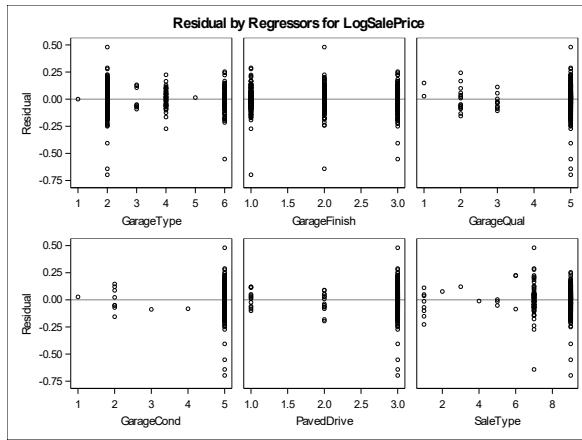












Appendix 3B – Forward Model Output

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Forward
Select Criterion	SBC
Stop Criterion	PRESS
Effect Hierarchy Enforced	None

Number of Observations Read	145 8
Number of Observations Used	145 0

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmngtn
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
Exterior1st	15	AsphShn BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing AsbShng
Exterior2nd	16	AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng AsbShng
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N

Class Level Information		
Class	Levels	Values
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	51
Number of Parameters	163

The GLMSELECT Procedure

Monday, April 13, 2020 10:48:57 PM 58

Forward Selection Summary						
Step	Effect Entered	Number Effects In	NumberParms In	Adjusted R-Square	SBC	PRESS
0	Intercept	1	1	0.0000	-2653.6575	231.7293
1	OverallQual	2	2	0.6733	-4269.4746	75.7775
2	LogLotArea	3	3	0.7460	-4628.0240	58.9592
3	GarageCars	4	4	0.7786	-4821.3924	51.4339
4	_1stFlrSF	5	5	0.7999	-4961.2707	46.5257
5	_2ndFlrSF	6	6	0.8291	-5184.3735	39.7536
6	YearBuilt	7	7	0.8557	-5422.5870	33.6158
7	OverallCond	8	8	0.8790	-5671.5479	28.2245
8	LogBsmtFinSF1	9	9	0.8889	-5789.1741	25.9416
9	Neighborhood	10	33	0.9040	-5850.6066	22.8265
10	CentralAir	11	34	0.9063	-5879.7846	22.3235
11	SaleCondition	12	39	0.9105	-5914.3712	21.6245
12	BsmtFullBath	13	40	0.9126	-5943.0824	21.0932
13	KitchenAbvGr	14	41	0.9138	-5957.3680	20.8528
14	ScreenPorch	15	42	0.9149	-5969.5548	20.6056
15	KitchenQual	16	45	0.9166	-5979.4391	20.2864
16	LogWoodDeckSF	17	46	0.9173	-5986.5238	20.1161
17	Street	18	47	0.9178	-5989.3993	20.0340
18	LogBsmtUnfSF	19	48	0.9182	-5989.8828*	19.9712
19	Fireplaces	20	49	0.9185	-5989.1581	19.9083
20	YearRemodAdd	21	50	0.9188	-5987.8336	19.8599
21	LogOpenPorchSF	22	51	0.9190*	-5985.4035	19.8317*
* Optimal Value of Criterion						

Selection stopped at a local minimum of the PRESS criterion.

The GLMSELECT Procedure

Monday, April 13, 2020 10:48:57 PM **59**

Stop Details				
Candidate For	Effect	Candidate PRESS	Compare PRESS	
Entry	PoolArea	19.8361	>	19.8317

The GLMSELECT Procedure
Selected Model

Monday, April 13, 2020 10:48:57 PM **60**

The selected model is the model at the last step (Step 21).

Effects:	Intercept Street Neighborhood CentralAir KitchenQual SaleCondition LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1 LogBsmtUnfSF _1stFlrSF _2ndFlrSF BsmtFullBath KitchenAbvGr Fireplaces GarageCars LogWoodDeckSF LogOpenPorchSF ScreenPorch
-----------------	--

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	50	213.31860	4.26637	329.92
Error	1399	18.09115	0.01293	
Corrected Total	1449	231.40975		

Root MSE	0.11372
Dependent Mean	12.02235
R-Square	0.9218
Adj R-Sq	0.9190
AIC	-4802.64880
AICC	-4798.70320
PRESS	19.83169
SBC	-5985.40354

The GLMSELECT Procedure
Selected Model

Monday, April 13, 2020 10:48:57 PM **61**

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	3.182792	0.640117	4.97
Street Pave	1	0.153126	0.049343	3.10
Street Grvl	0	0	.	.
Neighborhood Blueste	1	-0.051308	0.086396	-0.59
Neighborhood BrDale	1	-0.094538	0.042197	-2.24
Neighborhood BrkSide	1	-0.010998	0.036439	-0.30
Neighborhood ClearCr	1	-0.012600	0.039636	-0.32
Neighborhood CollgCr	1	-0.036273	0.031584	-1.15
Neighborhood Crawfor	1	0.091387	0.036462	2.51
Neighborhood Edwards	1	-0.095171	0.034171	-2.79
Neighborhood Gilbert	1	-0.047081	0.033558	-1.40
Neighborhood IDOTRR	1	-0.110324	0.038941	-2.83
Neighborhood MeadowV	1	-0.155891	0.041831	-3.73
Neighborhood Mitchel	1	-0.090364	0.035349	-2.56
Neighborhood NAmes	1	-0.062466	0.032769	-1.91
Neighborhood NPkVill	1	-0.023967	0.048313	-0.50
Neighborhood NWAmes	1	-0.093731	0.033951	-2.76
Neighborhood NoRidge	1	0.009965	0.036200	0.28
Neighborhood NridgHt	1	0.040208	0.032638	1.23
Neighborhood OldTown	1	-0.077726	0.035618	-2.18
Neighborhood SWISU	1	0.003475	0.040873	0.09
Neighborhood Sawyer	1	-0.082776	0.034563	-2.39
Neighborhood SawyerW	1	-0.071904	0.034071	-2.11
Neighborhood Somerst	1	0.031614	0.031971	0.99
Neighborhood StoneBr	1	0.080992	0.037341	2.17
Neighborhood Timber	1	-0.026691	0.036046	-0.74
Neighborhood Veenker	1	0.003737	0.046605	0.08
Neighborhood Blmngtn	0	0	.	.

The GLMSELECT Procedure
Selected Model

Monday, April 13, 2020 10:48:57 PM 62

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
CentralAir Y	1	0.077641	0.014836	5.23
CentralAir N	0	0	.	.
KitchenQual Fa	1	-0.105365	0.026565	-3.97
KitchenQual Gd	1	-0.068821	0.014467	-4.76
KitchenQual TA	1	-0.080832	0.016816	-4.81
KitchenQual Ex	0	0	.	.
SaleCondition AdjLand	1	0.138292	0.059737	2.32
SaleCondition Alloca	1	0.044396	0.038305	1.16
SaleCondition Family	1	0.042627	0.028228	1.51
SaleCondition Normal	1	0.084107	0.012144	6.93
SaleCondition Partial	1	0.140750	0.017279	8.15
SaleCondition Abnorml	0	0	.	.
LogLotArea	1	0.099506	0.009318	10.68
OverallQual	1	0.061903	0.004174	14.83
OverallCond	1	0.042376	0.003578	11.84
YearBuilt	1	0.002759	0.000262	10.51
YearRemodAdd	1	0.000533	0.000235	2.27
LogBsmtFinSF1	1	0.010343	0.001370	7.55
LogBsmtUnfSF	1	0.004516	0.001848	2.44
_1stFlrSF	1	0.000341	0.000013851	24.63
_2ndFlrSF	1	0.000256	0.000010004	25.59
BsmtFullBath	1	0.044287	0.007722	5.74
KitchenAbvGr	1	-0.054451	0.015941	-3.42
Fireplaces	1	0.015991	0.006058	2.64
GarageCars	1	0.047093	0.005847	8.05
LogWoodDeckSF	1	0.004806	0.001301	3.70
LogOpenPorchSF	1	0.003647	0.001685	2.16
ScreenPorch	1	0.000255	0.000055896	4.56

Store Information	
Item Store	WORK.Q2FORWARDMODEL
Data Set Created From	WORK.Q2TRAIN
Created By	PROC GLMSELECT
Date Created	14APR20:03:48:58
Response Variable	LogSalePrice
Class Variables	Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle RoofMatl ...
Model Effects	Intercept Street Neighborhood CentralAir KitchenQual SaleCondition LogLotArea OverallQual ...

Appendix 4B – Backward Model Output

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Backward
Select Criterion	SBC
Stop Criterion	PRESS
Effect Hierarchy Enforced	None

Number of Observations Read	1458
Number of Observations Used	1450

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmngtn
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam

Class Level Information		
Class	Levels	Values
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
Exterior1st	15	AsphShn BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing AsbShng
Exterior2nd	16	AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng AsbShng
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	51
Number of Parameters	163

Backward Selection Summary						
Step	Effect Removed	Number Effects In	NumberParms In	Adjusted R-Square	SBC	PRESS
0		51	143	0.9210	-5449.5729	22.7239
1	Exterior2nd	50	129	0.9212*	-5539.8297	22.3529
2	Exterior1st	49	115	0.9203	-5609.7533	21.9247
3	HouseStyle	48	108	0.9204	-5654.9626	21.6797
4	Electrical	47	103	0.9205	-5688.4649	21.5488
5	RoofMatl	46	97	0.9202	-5720.1809	21.0182
6	RoofStyle	45	92	0.9203	-5753.5502	20.4675
7	BldgType	44	88	0.9203	-5777.8807	20.4022
8	LotShape	43	85	0.9205	-5799.2915	20.3228
9	LandContour	42	82	0.9206	-5820.1579	20.2207
10	ExterQual	41	79	0.9207	-5840.7208*	20.1108*
* Optimal Value of Criterion						

Selection stopped at a local minimum of the PRESS criterion.

Stop Details				
Candidate For	Effect	Candidate PRESS	Compare PRESS	
Removal	Heating	20.1872	>	20.1108

The selected model is the model at the last step (Step 10).

Effects:	Intercept Street LandSlope Neighborhood Heating HeatingQC CentralAir KitchenQual PavedDrive SaleCondition MSSubClass LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1 LogBsmtFinSF2 LogBsmtUnfSF _1stFlrSF _2ndFlrSF LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageCars GarageArea LogWoodDeckSF LogOpenPorchSF _3SsnPorch ScreenPorch PoolArea MiscVal MoSold YrSold MasVnrArea
-----------------	--

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	78	214.04160	2.74412	216.61
Error	1371	17.36815	0.01267	
Corrected Total	1449	231.40975		

Root MSE	0.11255
Dependent Mean	12.02235
R-Square	0.9249
Adj R-Sq	0.9207
AIC	-4805.78695
AICC	-4796.32018
PRESS	20.11078
SBC	-5840.72076

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	10.590718	4.742308	2.23
Street Pave	1	0.147076	0.049973	2.94
Street Grvl	0	0	.	.
LandSlope Mod	1	0.010572	0.015623	0.68
LandSlope Sev	1	-0.048220	0.036951	-1.30
LandSlope Gtl	0	0	.	.
Neighborhood Blueste	1	-0.028067	0.086199	-0.33
Neighborhood BrDale	1	-0.094269	0.043196	-2.18
Neighborhood BrkSide	1	-0.008780	0.037438	-0.23
Neighborhood ClearCr	1	-0.001142	0.040369	-0.03
Neighborhood CollgCr	1	-0.041414	0.032632	-1.27
Neighborhood Crawfor	1	0.087379	0.037322	2.34
Neighborhood Edwards	1	-0.098536	0.034944	-2.82
Neighborhood Gilbert	1	-0.052656	0.034396	-1.53
Neighborhood IDOTRR	1	-0.115315	0.040087	-2.88
Neighborhood MeadowV	1	-0.144996	0.042416	-3.42
Neighborhood Mitchel	1	-0.091385	0.036178	-2.53
Neighborhood NAmes	1	-0.063357	0.034178	-1.85
Neighborhood NPkVill	1	-0.011142	0.048870	-0.23
Neighborhood NWAmes	1	-0.095388	0.035240	-2.71
Neighborhood NoRidge	1	0.007490	0.037956	0.20
Neighborhood NridgHt	1	0.036583	0.033456	1.09
Neighborhood OldTown	1	-0.078321	0.036456	-2.15
Neighborhood SWISU	1	-0.011635	0.042305	-0.28
Neighborhood Sawyer	1	-0.079827	0.035719	-2.23
Neighborhood SawyerW	1	-0.068686	0.034891	-1.97
Neighborhood Somerst	1	0.024325	0.032814	0.74
Neighborhood StoneBr	1	0.085492	0.037697	2.27
Neighborhood Timber	1	-0.024630	0.036598	-0.67

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Neighborhood Veenker	1	0.000809	0.047086	0.02
Neighborhood Blmngtn	0	0	.	.
Heating GasA	1	0.119503	0.115594	1.03
Heating GasW	1	0.177884	0.118587	1.50
Heating Grav	1	-0.029827	0.124436	-0.24
Heating OthW	1	0.066777	0.143427	0.47
Heating Wall	1	0.186595	0.128825	1.45
Heating Floo	0	0	.	.
HeatingQC Fa	1	-0.018084	0.020903	-0.87
HeatingQC Gd	1	-0.017042	0.009506	-1.79
HeatingQC Po	1	-0.073151	0.116148	-0.63
HeatingQC TA	1	-0.030814	0.009151	-3.37
HeatingQC Ex	0	0	.	.
CentralAir Y	1	0.065326	0.016393	3.99
CentralAir N	0	0	.	.
KitchenQual Fa	1	-0.086250	0.026977	-3.20
KitchenQual Gd	1	-0.066520	0.014611	-4.55
KitchenQual TA	1	-0.072727	0.016948	-4.29
KitchenQual Ex	0	0	.	.
PavedDrive P	1	0.023659	0.025179	0.94
PavedDrive Y	1	0.030858	0.015026	2.05
PavedDrive N	0	0	.	.
SaleCondition AdjLand	1	0.135761	0.060428	2.25
SaleCondition Alloca	1	0.042815	0.039001	1.10
SaleCondition Family	1	0.040405	0.028266	1.43
SaleCondition Normal	1	0.082517	0.012177	6.78
SaleCondition Partial	1	0.135349	0.017370	7.79
SaleCondition Abnorml	0	0	.	.
MSSubClass	1	-0.000032296	0.000107	-0.30

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
LogLotArea	1	0.102241	0.010195	10.03
OverallQual	1	0.061696	0.004208	14.66
OverallCond	1	0.042337	0.003611	11.72
YearBuilt	1	0.002502	0.000281	8.89
YearRemodAdd	1	0.000429	0.000242	1.78
LogBsmtFinSF1	1	0.009643	0.001419	6.80
LogBsmtFinSF2	1	-0.000880	0.001782	-0.49
LogBsmtUnfSF	1	0.004728	0.001962	2.41
_1stFlrSF	1	0.000319	0.000017648	18.08
_2ndFlrSF	1	0.000225	0.000017986	12.49
LowQualFinSF	1	0.000086509	0.000066864	1.29
BsmtFullBath	1	0.048855	0.008196	5.96
BsmtHalfBath	1	0.011766	0.013539	0.87
FullBath	1	0.020995	0.009852	2.13
HalfBath	1	0.021384	0.009139	2.34
BedroomAbvGr	1	-0.005530	0.005972	-0.93
KitchenAbvGr	1	-0.052122	0.018550	-2.81
TotRmsAbvGrd	1	0.002049	0.004280	0.48
Fireplaces	1	0.016501	0.006141	2.69
GarageCars	1	0.030462	0.009923	3.07
GarageArea	1	0.000050662	0.000034005	1.49
LogWoodDeckSF	1	0.004605	0.001306	3.52
LogOpenPorchSF	1	0.002739	0.001699	1.61
_3SsnPorch	1	0.000070865	0.000103	0.69
ScreenPorch	1	0.000244	0.000055953	4.36
PoolArea	1	0.000163	0.000081708	2.00
MiscVal	1	-0.000004737	0.000006063	-0.78
MoSold	1	-0.000875	0.001142	-0.77

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
YrSold	1	-0.003403	0.002347	-1.45
MasVnrArea	1	0.000019416	0.000020863	0.93

Store Information	
Item Store	WORK.Q2BACKWARDMODEL
Data Set Created From	WORK.Q2TRAIN
Created By	PROC GLMSELECT
Date Created	14APR20:03:51:46
Response Variable	LogSalePrice
Class Variables	Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle RoofMatl ...
Model Effects	Intercept Street LandSlope Neighborhood Heating HeatingQC CentralAir KitchenQual PavedDrive ...

Appendix 5B – Stepwise Model Output

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Stepwise
Select Criterion	SBC
Stop Criterion	PRESS
Effect Hierarchy Enforced	None

Number of Observations Read	1458
Number of Observations Used	1450

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmngtn
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
Exterior1st	15	AsphShn BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing AsbShng
Exterior2nd	16	AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng AsbShng
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	51
Number of Parameters	163

The GLMSELECT Procedure

Monday, April 13, 2020 10:54:35 PM 72

Stepwise Selection Summary							
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Adjusted R-Square	SBC	PRESS
0	Intercept		1	1	0.0000	-2653.6575	231.7293
1	OverallQual		2	2	0.6733	-4269.4746	75.7775
2	LogLotArea		3	3	0.7460	-4628.0240	58.9592
3	GarageCars		4	4	0.7786	-4821.3924	51.4339
4	_1stFlrSF		5	5	0.7999	-4961.2707	46.5257
5	_2ndFlrSF		6	6	0.8291	-5184.3735	39.7536
6	YearBuilt		7	7	0.8557	-5422.5870	33.6158
7	OverallCond		8	8	0.8790	-5671.5479	28.2245
8	LogBsmtFinSF1		9	9	0.8889	-5789.1741	25.9416
9	Neighborhood		10	33	0.9040	-5850.6066	22.8265
10	CentralAir		11	34	0.9063	-5879.7846	22.3235
11	SaleCondition		12	39	0.9105	-5914.3712	21.6245
12	BsmtFullBath		13	40	0.9126	-5943.0824	21.0932
13	KitchenAbvGr		14	41	0.9138	-5957.3680	20.8528
14	ScreenPorch		15	42	0.9149	-5969.5548	20.6056
15	KitchenQual		16	45	0.9166	-5979.4391	20.2864
16	LogWoodDeckSF		17	46	0.9173	-5986.5238	20.1161
17	Street		18	47	0.9178	-5989.3993	20.0340
18	LogBsmtUnfSF		19	48	0.9182*	-5989.8828*	19.9712*
* Optimal Value of Criterion							

Selection stopped as adding or dropping any effect does not improve the selection criterion.

The selected model is the model at the last step (Step 18).

Effects:	Intercept Street Neighborhood CentralAir KitchenQual SaleCondition LogLotArea OverallQual OverallCond YearBuilt LogBsmtFinSF1 LogBsmtUnfSF _1stFlrSF _2ndFlrSF BsmtFullBath KitchenAbvGr GarageCars LogWoodDeckSF ScreenPorch
-----------------	---

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	47	213.10072	4.53406	347.19
Error	1402	18.30903	0.01306	
Corrected Total	1449	231.40975		

Root MSE	0.11428
Dependent Mean	12.02235
R-Square	0.9209
Adj R-Sq	0.9182
AIC	-4791.29013
AICC	-4787.79013
PRESS	19.97124
SBC	-5989.88283

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	3.808827	0.531983	7.16
Street Pave	1	0.153038	0.049571	3.09
Street Grvl	0	0	.	.
Neighborhood Blueste	1	-0.054243	0.086744	-0.63
Neighborhood BrDale	1	-0.105634	0.042239	-2.50
Neighborhood BrkSide	1	-0.017398	0.036494	-0.48
Neighborhood ClearCr	1	-0.014261	0.039796	-0.36
Neighborhood CollgCr	1	-0.045302	0.031490	-1.44
Neighborhood Crawfor	1	0.088310	0.036597	2.41
Neighborhood Edwards	1	-0.103664	0.034139	-3.04
Neighborhood Gilbert	1	-0.045700	0.033695	-1.36
Neighborhood IDOTRR	1	-0.115171	0.038932	-2.96
Neighborhood MeadowV	1	-0.157603	0.041885	-3.76
Neighborhood Mitchel	1	-0.097813	0.035378	-2.76
Neighborhood NAmes	1	-0.072435	0.032736	-2.21
Neighborhood NPkVill	1	-0.024649	0.048477	-0.51
Neighborhood NWAmes	1	-0.099888	0.034012	-2.94
Neighborhood NoRidge	1	-0.000342	0.036215	-0.01
Neighborhood NridgHt	1	0.036494	0.032702	1.12
Neighborhood OldTown	1	-0.080767	0.035584	-2.27
Neighborhood SWISU	1	-0.001342	0.040932	-0.03
Neighborhood Sawyer	1	-0.090426	0.034624	-2.61
Neighborhood SawyerW	1	-0.078570	0.034093	-2.30
Neighborhood Somerst	1	0.029047	0.031880	0.91
Neighborhood StoneBr	1	0.069025	0.037341	1.85
Neighborhood Timber	1	-0.033374	0.036170	-0.92
Neighborhood Veenker	1	0.001398	0.046766	0.03
Neighborhood Blmgtn	0	0	.	.
CentralAir Y	1	0.079322	0.014794	5.36

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
CentralAir_N	0	0	.	.
KitchenQual_Fa	1	-0.114409	0.026418	-4.33
KitchenQual_Gd	1	-0.068581	0.014516	-4.72
KitchenQual_TA	1	-0.088408	0.016631	-5.32
KitchenQual_Ex	0	0	.	.
SaleCondition_AdjLand	1	0.143020	0.060000	2.38
SaleCondition_Alloca	1	0.046309	0.038436	1.20
SaleCondition_Family	1	0.044219	0.028350	1.56
SaleCondition_Normal	1	0.088196	0.012144	7.26
SaleCondition_Partial	1	0.146095	0.017278	8.46
SaleCondition_Abnorml	0	0	.	.
LogLotArea	1	0.101704	0.009288	10.95
OverallQual	1	0.063824	0.004158	15.35
OverallCond	1	0.044938	0.003350	13.41
YearBuilt	1	0.002958	0.000253	11.68
LogBsmtFinSF1	1	0.010262	0.001371	7.49
LogBsmtUnfSF	1	0.005077	0.001851	2.74
_1stFlrSF	1	0.000355	0.000013437	26.44
_2ndFlrSF	1	0.000266	0.000009746	27.26
BsmtFullBath	1	0.047331	0.007720	6.13
KitchenAbvGr	1	-0.061595	0.015885	-3.88
GarageCars	1	0.046402	0.005873	7.90
LogWoodDeckSF	1	0.004981	0.001301	3.83
ScreenPorch	1	0.000271	0.000055782	4.86

Store Information	
Item Store	WORK.Q2STEPMODEL
Data Set Created From	WORK.Q2TRAIN
Created By	PROC GLMSELECT
Date Created	14APR20:03:54:36
Response Variable	LogSalePrice
Class Variables	Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle RoofMatl ...
Model Effects	Intercept Street Neighborhood CentralAir KitchenQual SaleCondition LogLotArea OverallQual ...

Appendix 6B – Custom Model Output at 0.10 Significance Level

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Stepwise
Select Criterion	Significance Level
Stop Criterion	Significance Level
Entry Significance Level (SLE)	0.1
Stay Significance Level (SLS)	0.1
Effect Hierarchy Enforced	None

Number of Observations Read	1458
Number of Observations Used	1458

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmgtn

Class Level Information		
Class	Levels	Values
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	48
Number of Parameters	131

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
0	Intercept		1	1	0.0000	233.1081	0.00	1.0000
1	OverallQual		2	2	0.6747	75.9499	3019.94	<.0001
2	LogLotArea		3	3	0.7471	59.1344	416.22	<.0001
3	Neighborhood		4	27	0.7987	48.7034	15.29	<.0001
4	_1stFlrSF		5	28	0.8206	43.4469	174.28	<.0001
5	_2ndFlrSF		6	29	0.8565	34.7466	357.96	<.0001
6	LogBsmtFinSF1		7	30	0.8720	31.0605	172.98	<.0001
7	YearRemodAdd		8	31	0.8844	28.0961	153.19	<.0001
8	YearBuilt		9	32	0.8912	26.5171	89.04	<.0001
9	OverallCond		10	33	0.9027	23.7752	167.84	<.0001
10	GarageCars		11	34	0.9072	22.7194	69.69	<.0001
11	SaleCondition		12	39	0.9112	22.1234	12.73	<.0001
12	CentralAir		13	40	0.9137	21.5783	40.27	<.0001
13	BsmtFullBath		14	41	0.9156	21.1017	31.97	<.0001
14	KitchenQual		15	44	0.9173	20.7944	9.86	<.0001
15	ScreenPorch		16	45	0.9185	20.5304	20.50	<.0001
16	KitchenAbvGr		17	46	0.9195	20.3318	18.38	<.0001
17	LogWoodDeckSF		18	47	0.9203	20.1590	14.16	0.0002
18	Street		19	48	0.9208	20.0854	9.05	0.0027
19	Fireplaces		20	49	0.9213	20.0024	7.73	0.0055
20	Heating		21	54	0.9221	19.9677	3.01	0.0105
21	LogBsmtUnfSF		22	55	0.9225	19.9052	7.68	0.0056
22	HeatingQC		23	59	0.9232	19.8397	2.94	0.0197
23	LogOpenPorchSF		24	60	0.9234	19.8048	4.65	0.0313
24		YearRemodAdd	23	59	0.9233	19.8101	2.47	0.1166
25	PoolArea		24	60	0.9235	19.8029	4.09	0.0434
26	HalfBath		25	61	0.9237	19.7862	3.05	0.0810
27	FullBath		26	62	0.9239	19.7674	3.83	0.0507

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
28	GarageArea		27	63	0.9241	19.7587*	3.31	0.0690
* Optimal Value of Criterion								

Selection stopped because the candidate for entry has SLE > 0.1 and the candidate for removal has SLS < 0.1.

Stop Details				
Candidate For	Effect	Candidate Significance	Compare Significance	
Entry	LowQualFinSF	0.1130 >	0.1000	(SLE)
Removal	GarageArea	0.0690 <	0.1000	(SLS)

The selected model is the model at the last step (Step 28).

Effects:	Intercept Street Neighborhood Heating HeatingQC CentralAir KitchenQual SaleCondition LogLotArea OverallQual OverallCond YearBuilt LogBsmtFinSF1 LogBsmtUnfSF _1stFlrSF _2ndFlrSF BsmtFullBath FullBath HalfBath KitchenAbvGr Fireplaces GarageCars GarageArea LogWoodDeckSF LogOpenPorchSF ScreenPorch PoolArea
-----------------	---

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	62	215.11033	3.46952	273.78
Error	1395	17.67809	0.01267	
Corrected Total	1457	232.78842		

Root MSE	0.11257
Dependent Mean	12.02401
R-Square	0.9241
Adj R-Sq	0.9207
AIC	-4847.41762
AICC	-4841.44490
PRESS	19.75867
SBC	-5974.47390

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	4.317903	0.572712	7.54
Street Pave	1	0.150740	0.048985	3.08
Street Grvl	0	0	.	.
Neighborhood Blueste	1	-0.046260	0.085823	-0.54
Neighborhood BrDale	1	-0.099543	0.042351	-2.35
Neighborhood BrkSide	1	-0.008642	0.036218	-0.24
Neighborhood ClearCr	1	-0.004164	0.039342	-0.11
Neighborhood CollgCr	1	-0.040528	0.031418	-1.29
Neighborhood Crawfor	1	0.091322	0.036159	2.53
Neighborhood Edwards	1	-0.098832	0.033885	-2.92
Neighborhood Gilbert	1	-0.047561	0.033372	-1.43
Neighborhood IDOTRR	1	-0.117084	0.039080	-3.00
Neighborhood MeadowV	1	-0.157162	0.041712	-3.77
Neighborhood Mitchel	1	-0.086772	0.035100	-2.47
Neighborhood NAmes	1	-0.063209	0.032623	-1.94
Neighborhood NPkVill	1	-0.024482	0.048557	-0.50
Neighborhood NWAmes	1	-0.094373	0.034001	-2.78
Neighborhood NoRidge	1	0.009538	0.036004	0.26
Neighborhood NridgHt	1	0.038409	0.032544	1.18
Neighborhood OldTown	1	-0.077389	0.035405	-2.19
Neighborhood SWISU	1	-0.005221	0.040688	-0.13
Neighborhood Sawyer	1	-0.078959	0.034383	-2.30
Neighborhood SawyerW	1	-0.068767	0.033771	-2.04
Neighborhood Somerst	1	0.021474	0.031882	0.67
Neighborhood StoneBr	1	0.083564	0.037108	2.25
Neighborhood Timber	1	-0.025853	0.035761	-0.72
Neighborhood Veenker	1	0.002676	0.046368	0.06
Neighborhood Blmngtn	0	0	.	.
Heating GasA	1	0.120167	0.115409	1.04

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Heating GasW	1	0.172479	0.118221	1.46
Heating Grav	1	-0.030652	0.124138	-0.25
Heating OthW	1	0.079469	0.142842	0.56
Heating Wall	1	0.189398	0.128694	1.47
Heating Floo	0	0	.	.
HeatingQC Fa	1	-0.017998	0.020393	-0.88
HeatingQC Gd	1	-0.017170	0.009417	-1.82
HeatingQC Po	1	-0.079700	0.115621	-0.69
HeatingQC TA	1	-0.033977	0.008977	-3.78
HeatingQC Ex	0	0	.	.
CentralAir Y	1	0.067844	0.016268	4.17
CentralAir N	0	0	.	.
KitchenQual Fa	1	-0.099812	0.026469	-3.77
KitchenQual Gd	1	-0.071120	0.014393	-4.94
KitchenQual TA	1	-0.080547	0.016537	-4.87
KitchenQual Ex	0	0	.	.
SaleCondition AdjLand	1	0.137848	0.059595	2.31
SaleCondition Alloca	1	0.029481	0.037297	0.79
SaleCondition Family	1	0.042788	0.028009	1.53
SaleCondition Normal	1	0.082532	0.012079	6.83
SaleCondition Partial	1	0.138875	0.017034	8.15
SaleCondition Abnorml	0	0	.	.
LogLotArea	1	0.095834	0.009263	10.35
OverallQual	1	0.062017	0.004152	14.94
OverallCond	1	0.044745	0.003344	13.38
YearBuilt	1	0.002680	0.000266	10.08
LogBsmtFinSF1	1	0.010176	0.001363	7.47
LogBsmtUnfSF	1	0.004600	0.001869	2.46
_1stFlrSF	1	0.000326	0.000014942	21.79

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
_2ndFlrSF	1	0.000226	0.000013740	16.45
BsmtFullBath	1	0.045129	0.007714	5.85
FullBath	1	0.019843	0.009662	2.05
HalfBath	1	0.022362	0.009059	2.47
KitchenAbvGr	1	-0.060787	0.016358	-3.72
Fireplaces	1	0.015336	0.006000	2.56
GarageCars	1	0.031329	0.009856	3.18
GarageArea	1	0.000061043	0.000033543	1.82
LogWoodDeckSF	1	0.004959	0.001285	3.86
LogOpenPorchSF	1	0.003123	0.001677	1.86
ScreenPorch	1	0.000249	0.000055595	4.48
PoolArea	1	0.000173	0.000080419	2.16

Appendix 7B – Custom Model Output at 0.50 Significance Level

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Stepwise
Select Criterion	Significance Level
Stop Criterion	Significance Level
Entry Significance Level (SLE)	0.5
Stay Significance Level (SLS)	0.5
Effect Hierarchy Enforced	None

Number of Observations Read	1458
Number of Observations Used	1458

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmngtn
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	48
Number of Parameters	131

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
0	Intercept		1	1	0.0000	233.1081	0.00	1.0000
1	OverallQual		2	2	0.6747	75.9499	3019.94	<.0001
2	LogLotArea		3	3	0.7471	59.1344	416.22	<.0001
3	Neighborhood		4	27	0.7987	48.7034	15.29	<.0001
4	_1stFlrSF		5	28	0.8206	43.4469	174.28	<.0001
5	_2ndFlrSF		6	29	0.8565	34.7466	357.96	<.0001
6	LogBsmtFinSF1		7	30	0.8720	31.0605	172.98	<.0001
7	YearRemodAdd		8	31	0.8844	28.0961	153.19	<.0001
8	YearBuilt		9	32	0.8912	26.5171	89.04	<.0001
9	OverallCond		10	33	0.9027	23.7752	167.84	<.0001
10	GarageCars		11	34	0.9072	22.7194	69.69	<.0001
11	SaleCondition		12	39	0.9112	22.1234	12.73	<.0001
12	CentralAir		13	40	0.9137	21.5783	40.27	<.0001
13	BsmtFullBath		14	41	0.9156	21.1017	31.97	<.0001
14	KitchenQual		15	44	0.9173	20.7944	9.86	<.0001
15	ScreenPorch		16	45	0.9185	20.5304	20.50	<.0001
16	KitchenAbvGr		17	46	0.9195	20.3318	18.38	<.0001
17	LogWoodDeckSF		18	47	0.9203	20.1590	14.16	0.0002
18	Street		19	48	0.9208	20.0854	9.05	0.0027
19	Fireplaces		20	49	0.9213	20.0024	7.73	0.0055
20	Heating		21	54	0.9221	19.9677	3.01	0.0105
21	LogBsmtUnfSF		22	55	0.9225	19.9052	7.68	0.0056
22	HeatingQC		23	59	0.9232	19.8397	2.94	0.0197
23	LogOpenPorchSF		24	60	0.9234	19.8048	4.65	0.0313
24	PoolArea		25	61	0.9236	19.7998	4.17	0.0414
25	GarageArea		26	62	0.9238	19.7958	2.98	0.0847
26	HalfBath		27	63	0.9240	19.7775	3.16	0.0755
27	FullBath		28	64	0.9242	19.7604*	3.79	0.0518
28	LowQualFinSF		29	65	0.9243	19.7797	2.35	0.1259

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
29	PavedDrive		30	67	0.9245	19.8071	1.85	0.1569
30	RoofMatl		31	73	0.9250	19.8959	1.54	0.1620
31	YrSold		32	74	0.9251	19.9034	1.56	0.2113
32	LandSlope		33	76	0.9252	19.9951	1.30	0.2735
33	BsmtHalfBath		34	77	0.9253	20.0128	0.90	0.3416
34	BedroomAbvGr		35	78	0.9253	20.0366	0.76	0.3847
35	BldgType		36	82	0.9256	20.1165	1.07	0.3714
36	MiscVal		37	83	0.9256	20.2522	0.60	0.4370
37	MoSold		38	84	0.9256	20.2807	0.47	0.4947
* Optimal Value of Criterion								

Selection stopped because the candidate for entry has SLE > 0.5 and the candidate for removal has SLS < 0.5.

Stop Details				
Candidate For	Effect	Candidate Significance	Compare Significance	
Entry	_3SsnPorch	0.5275 >	0.5000	(SLE)
Removal	MoSold	0.4947 <	0.5000	(SLS)

The selected model is the model at the last step (Step 37).

Effects:	Intercept Street LandSlope Neighborhood BldgType RoofMatl Heating HeatingQC CentralAir KitchenQual PavedDrive SaleCondition LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1 LogBsmtUnfSF _1stFlrSF _2ndFlrSF LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr Fireplaces GarageCars GarageArea LogWoodDeckSF LogOpenPorchSF ScreenPorch PoolArea MiscVal MoSold YrSold
-----------------	--

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	83	215.47276	2.59606	206.00
Error	1374	17.31566	0.01260	
Corrected Total	1457	232.78842		

Root MSE	0.11226
Dependent Mean	12.02401
R-Square	0.9256
Adj R-Sq	0.9211
AIC	-4835.61989
AICC	-4824.96392
PRESS	20.28072
SBC	-5851.69494

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	9.942060	4.724401	2.10
Street Pave	1	0.146096	0.050619	2.89
Street Grvl	0	0	.	.
LandSlope Mod	1	0.013094	0.015614	0.84
LandSlope Sev	1	-0.043953	0.039640	-1.11
LandSlope Gtl	0	0	.	.
Neighborhood Blueste	1	-0.026329	0.086232	-0.31
Neighborhood BrDale	1	-0.073651	0.045215	-1.63
Neighborhood BrkSide	1	-0.015136	0.037875	-0.40
Neighborhood ClearCr	1	-0.001448	0.041050	-0.04
Neighborhood CollgCr	1	-0.045054	0.032939	-1.37
Neighborhood Crawfor	1	0.081141	0.037443	2.17
Neighborhood Edwards	1	-0.099304	0.035354	-2.81
Neighborhood Gilbert	1	-0.056599	0.035004	-1.62
Neighborhood IDOTRR	1	-0.119476	0.040630	-2.94
Neighborhood MeadowV	1	-0.138332	0.042804	-3.23
Neighborhood Mitchel	1	-0.092200	0.036533	-2.52
Neighborhood NAmes	1	-0.063604	0.034403	-1.85
Neighborhood NPkVill	1	-0.003647	0.049187	-0.07
Neighborhood NWAmes	1	-0.097860	0.035478	-2.76
Neighborhood NoRidge	1	0.005492	0.037236	0.15
Neighborhood NridgHt	1	0.045456	0.033557	1.35
Neighborhood OldTown	1	-0.085585	0.036901	-2.32
Neighborhood SWISU	1	-0.017014	0.042498	-0.40
Neighborhood Sawyer	1	-0.082146	0.036041	-2.28
Neighborhood SawyerW	1	-0.070169	0.035004	-2.00
Neighborhood Somerst	1	0.023219	0.032917	0.71
Neighborhood StoneBr	1	0.088121	0.037479	2.35
Neighborhood Timber	1	-0.022833	0.037107	-0.62

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Neighborhood Veenker	1	-0.005023	0.047044	-0.11
Neighborhood Blmngtn	0	0	.	.
BldgType 2fmCon	1	0.023772	0.025555	0.93
BldgType Duplex	1	0.002077	0.027309	0.08
BldgType Twnhs	1	-0.046248	0.026053	-1.78
BldgType TwnhsE	1	-0.021774	0.016621	-1.31
BldgType 1Fam	0	0	.	.
RoofMatl Membran	1	0.105061	0.119810	0.88
RoofMatl Metal	1	0.115429	0.120491	0.96
RoofMatl Roll	1	-0.012032	0.115096	-0.10
RoofMatl Tar&Grv	1	-0.077264	0.036919	-2.09
RoofMatl WdShake	1	-0.045473	0.052986	-0.86
RoofMatl WdShngl	1	0.069830	0.048433	1.44
RoofMatl CompShg	0	0	.	.
Heating GasA	1	0.119602	0.115295	1.04
Heating GasW	1	0.180989	0.118213	1.53
Heating Grav	1	-0.034743	0.124214	-0.28
Heating OthW	1	0.068671	0.143160	0.48
Heating Wall	1	0.188608	0.128880	1.46
Heating Floo	0	0	.	.
HeatingQC Fa	1	-0.015137	0.020890	-0.72
HeatingQC Gd	1	-0.017154	0.009493	-1.81
HeatingQC Po	1	-0.077241	0.115848	-0.67
HeatingQC TA	1	-0.030780	0.009195	-3.35
HeatingQC Ex	0	0	.	.
CentralAir Y	1	0.067144	0.016464	4.08
CentralAir N	0	0	.	.
KitchenQual Fa	1	-0.086613	0.026895	-3.22
KitchenQual Gd	1	-0.066428	0.014532	-4.57

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
KitchenQual TA	1	-0.073076	0.016911	-4.32
KitchenQual Ex	0	0	.	.
PavedDrive P	1	0.024124	0.025168	0.96
PavedDrive Y	1	0.029836	0.015012	1.99
PavedDrive N	0	0	.	.
SaleCondition AdjLand	1	0.151455	0.061821	2.45
SaleCondition Allocacn	1	0.027648	0.037732	0.73
SaleCondition Family	1	0.042232	0.028196	1.50
SaleCondition Normal	1	0.080809	0.012158	6.65
SaleCondition Partial	1	0.131604	0.017291	7.61
SaleCondition Abnorml	0	0	.	.
LogLotArea	1	0.091653	0.011013	8.32
OverallQual	1	0.060948	0.004204	14.50
OverallCond	1	0.043105	0.003617	11.92
YearBuilt	1	0.002580	0.000281	9.18
YearRemodAdd	1	0.000400	0.000242	1.65
LogBsmtFinSF1	1	0.009447	0.001419	6.66
LogBsmtUnfSF	1	0.004802	0.001934	2.48
_1stFlrSF	1	0.000330	0.000015938	20.73
_2ndFlrSF	1	0.000234	0.000015321	15.29
LowQualFinSF	1	0.000081054	0.000066157	1.23
BsmtFullBath	1	0.048045	0.008151	5.89
BsmtHalfBath	1	0.013884	0.013494	1.03
FullBath	1	0.021813	0.009825	2.22
HalfBath	1	0.021887	0.009164	2.39
BedroomAbvGr	1	-0.006923	0.005518	-1.25
KitchenAbvGr	1	-0.063204	0.024218	-2.61
Fireplaces	1	0.015605	0.006136	2.54
GarageCars	1	0.032407	0.009933	3.26

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
GarageArea	1	0.000045599	0.000034125	1.34
LogWoodDeckSF	1	0.004523	0.001296	3.49
LogOpenPorchSF	1	0.002656	0.001702	1.56
ScreenPorch	1	0.000252	0.000055859	4.51
PoolArea	1	0.000187	0.000081646	2.29
MiscVal	1	-0.000004708	0.000006046	-0.78
MoSold	1	-0.000777	0.001138	-0.68
YrSold	1	-0.003071	0.002339	-1.31

Appendix 8B – Custom Model Output at 0.99 Significance Level

Data Set	WORK.Q2TRAIN
Dependent Variable	LogSalePrice
Selection Method	Stepwise
Select Criterion	Significance Level
Stop Criterion	Significance Level
Entry Significance Level (SLE)	0.99
Stay Significance Level (SLS)	0.99
Effect Hierarchy Enforced	None

Number of Observations Read	1458
Number of Observations Used	1458

Class Level Information		
Class	Levels	Values
Street	2	Pave Grvl
LotShape	4	IR2 IR3 Reg IR1
LandContour	4	HLS Low Lvl Bnk
LandSlope	3	Mod Sev Gtl

Class Level Information		
Class	Levels	Values
Neighborhood	25	Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker Blmgtn
BldgType	5	2fmCon Duplex Twnhs TwnhsE 1Fam
HouseStyle	8	1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl 1.5Fin
RoofStyle	6	Gable Gambr Hip Mansa Shed Flat
RoofMatl	7	Membran Metal Roll Tar&Grv WdShake WdShngl CompShg
ExterQual	4	Fa Gd TA Ex
Heating	6	GasA GasW Grav OthW Wall Floo
HeatingQC	5	Fa Gd Po TA Ex
CentralAir	2	Y N
KitchenQual	4	Fa Gd TA Ex
PavedDrive	3	P Y N
SaleCondition	6	AdjLand Alloca Family Normal Partial Abnorml
Electrical	6	FuseF FuseP Mix NA SBrkr FuseA

Dimensions	
Number of Effects	48
Number of Parameters	131

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
0	Intercept		1	1	0.0000	233.1081	0.00	1.0000
1	OverallQual		2	2	0.6747	75.9499	3019.94	<.0001
2	LogLotArea		3	3	0.7471	59.1344	416.22	<.0001
3	Neighborhood		4	27	0.7987	48.7034	15.29	<.0001
4	_1stFlrSF		5	28	0.8206	43.4469	174.28	<.0001
5	_2ndFlrSF		6	29	0.8565	34.7466	357.96	<.0001
6	LogBsmtFinSF1		7	30	0.8720	31.0605	172.98	<.0001
7	YearRemodAdd		8	31	0.8844	28.0961	153.19	<.0001
8	YearBuilt		9	32	0.8912	26.5171	89.04	<.0001
9	OverallCond		10	33	0.9027	23.7752	167.84	<.0001
10	GarageCars		11	34	0.9072	22.7194	69.69	<.0001
11	SaleCondition		12	39	0.9112	22.1234	12.73	<.0001
12	CentralAir		13	40	0.9137	21.5783	40.27	<.0001
13	BsmtFullBath		14	41	0.9156	21.1017	31.97	<.0001
14	KitchenQual		15	44	0.9173	20.7944	9.86	<.0001
15	ScreenPorch		16	45	0.9185	20.5304	20.50	<.0001
16	KitchenAbvGr		17	46	0.9195	20.3318	18.38	<.0001
17	LogWoodDeckSF		18	47	0.9203	20.1590	14.16	0.0002
18	Street		19	48	0.9208	20.0854	9.05	0.0027
19	Fireplaces		20	49	0.9213	20.0024	7.73	0.0055
20	Heating		21	54	0.9221	19.9677	3.01	0.0105
21	LogBsmtUnfSF		22	55	0.9225	19.9052	7.68	0.0056
22	HeatingQC		23	59	0.9232	19.8397	2.94	0.0197
23	LogOpenPorchSF		24	60	0.9234	19.8048	4.65	0.0313
24	PoolArea		25	61	0.9236	19.7998	4.17	0.0414
25	GarageArea		26	62	0.9238	19.7958	2.98	0.0847
26	HalfBath		27	63	0.9240	19.7775	3.16	0.0755
27	FullBath		28	64	0.9242	19.7604*	3.79	0.0518
28	LowQualFinSF		29	65	0.9243	19.7797	2.35	0.1259

Stepwise Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	PRESS	F Value	Pr > F
29	PavedDrive		30	67	0.9245	19.8071	1.85	0.1569
30	RoofMatl		31	73	0.9250	19.8959	1.54	0.1620
31	YrSold		32	74	0.9251	19.9034	1.56	0.2113
32	LandSlope		33	76	0.9252	19.9951	1.30	0.2735
33	BsmtHalfBath		34	77	0.9253	20.0128	0.90	0.3416
34	BedroomAbvGr		35	78	0.9253	20.0366	0.76	0.3847
35	BldgType		36	82	0.9256	20.1165	1.07	0.3714
36	MiscVal		37	83	0.9256	20.2522	0.60	0.4370
37	MoSold		38	84	0.9256	20.2807	0.47	0.4947
38	_3SsnPorch		39	85	0.9256	20.3048	0.40	0.5275
39	HouseStyle		40	92	0.9260	20.4854	0.84	0.5575
40	RoofStyle		41	97	0.9262	21.4694	0.79	0.5602
41	LogBsmtFinSF2		42	98	0.9262	21.4816	0.42	0.5150
42	TotRmsAbvGrd		43	99	0.9262	21.5428	0.24	0.6275
43	LandContour		44	102	0.9263	21.6539	0.51	0.6752
44	MSSubClass		45	103	0.9263	21.6857	0.13	0.7221
45	LotShape		46	106	0.9264	21.7586	0.43	0.7305
46	Electrical		47	111	0.9265	21.8847	0.48	0.7882
47	ExterQual		48	114	0.9265	22.0197	0.32	0.8135
* Optimal Value of Criterion								

Selection stopped because all effects are in the final model.

The selected model is the model at the last step (Step 47).

Effects:	Intercept Street LotShape LandContour LandSlope Neighborhood BldgType HouseStyle RoofStyle RoofMatl ExterQual Heating HeatingQC CentralAir KitchenQual PavedDrive SaleCondition Electrical MSSubClass LogLotArea OverallQual OverallCond YearBuilt YearRemodAdd LogBsmtFinSF1 LogBsmtFinSF2 LogBsmtUnfSF _1stFlrSF _2ndFlrSF LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr TotRmsAbvGrd Fireplaces GarageCars GarageArea LogWoodDeckSF LogOpenPorchSF _3SsnPorch ScreenPorch PoolArea MiscVal MoSold YrSold
-----------------	---

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	113	215.68979	1.90876	150.03
Error	1344	17.09863	0.01272	
Corrected Total	1457	232.78842		

Root MSE	0.11279
Dependent Mean	12.02401
R-Square	0.9265
Adj R-Sq	0.9204
AIC	-4794.00933
AICC	-4774.12855
PRESS	22.01972
SBC	-5651.53974

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	10.156853	4.807908	2.11
Street Pave	1	0.138551	0.053698	2.58
Street Grvl	0	0	.	.
LotShape IR2	1	0.015051	0.019656	0.77
LotShape IR3	1	0.037645	0.040788	0.92
LotShape Reg	1	-0.000226	0.007300	-0.03
LotShape IR1	0	0	.	.
LandContour HLS	1	0.020243	0.023861	0.85
LandContour Low	1	-0.010841	0.029294	-0.37
LandContour Lvl	1	0.003712	0.017418	0.21
LandContour Bnk	0	0	.	.
LandSlope Mod	1	0.016209	0.018189	0.89
LandSlope Sev	1	-0.040633	0.042748	-0.95
LandSlope Gtl	0	0	.	.
Neighborhood Blueste	1	-0.023320	0.087228	-0.27
Neighborhood BrDale	1	-0.078512	0.046253	-1.70
Neighborhood BrkSide	1	-0.021018	0.039332	-0.53
Neighborhood ClearCr	1	-0.003880	0.042738	-0.09
Neighborhood CollgCr	1	-0.044888	0.033821	-1.33
Neighborhood Crawfor	1	0.074910	0.038853	1.93
Neighborhood Edwards	1	-0.103413	0.036361	-2.84
Neighborhood Gilbert	1	-0.059989	0.036281	-1.65
Neighborhood IDOTRR	1	-0.122228	0.041860	-2.92
Neighborhood MeadowV	1	-0.152705	0.044256	-3.45
Neighborhood Mitchel	1	-0.094391	0.037749	-2.50
Neighborhood NAmes	1	-0.066890	0.035550	-1.88
Neighborhood NPkVill	1	-0.005860	0.049954	-0.12
Neighborhood NWAmes	1	-0.101152	0.036609	-2.76
Neighborhood NoRidge	1	0.006930	0.038630	0.18

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
Neighborhood NridgHt	1	0.044045	0.034489	1.28
Neighborhood OldTown	1	-0.093047	0.038110	-2.44
Neighborhood SWISU	1	-0.018712	0.044181	-0.42
Neighborhood Sawyer	1	-0.086451	0.037293	-2.32
Neighborhood SawyerW	1	-0.070258	0.035879	-1.96
Neighborhood Somerst	1	0.024325	0.033805	0.72
Neighborhood StoneBr	1	0.086913	0.038494	2.26
Neighborhood Timber	1	-0.025437	0.038646	-0.66
Neighborhood Veenker	1	-0.000690	0.047959	-0.01
Neighborhood Blmngtn	0	0	.	.
BldgType 2fmCon	1	0.047807	0.056609	0.84
BldgType Duplex	1	0.003291	0.033160	0.10
BldgType Twnhs	1	-0.032896	0.045333	-0.73
BldgType TwnhsE	1	-0.005818	0.040530	-0.14
BldgType 1Fam	0	0	.	.
HouseStyle 1.5Unf	1	0.027598	0.035841	0.77
HouseStyle 1Story	1	-0.004445	0.019208	-0.23
HouseStyle 2.5Fin	1	-0.039088	0.055552	-0.70
HouseStyle 2.5Unf	1	0.064280	0.039911	1.61
HouseStyle 2Story	1	0.006445	0.015772	0.41
HouseStyle SFoyer	1	0.032090	0.028065	1.14
HouseStyle SLvl	1	0.009264	0.024086	0.38
HouseStyle 1.5Fin	0	0	.	.
RoofStyle Gable	1	-0.043609	0.085252	-0.51
RoofStyle Gambr	1	-0.057012	0.092591	-0.62
RoofStyle Hip	1	-0.040281	0.085533	-0.47
RoofStyle Mansa	1	0.030807	0.098135	0.31
RoofStyle Shed	1	0.076330	0.126502	0.60
RoofStyle Flat	0	0	.	.

Parameter Estimates					
Parameter		DF	Estimate	Standard Error	t Value
RoofMatl	Membran	1	0.096473	0.145065	0.67
RoofMatl	Metal	1	0.079607	0.148562	0.54
RoofMatl	Roll	1	-0.012616	0.116638	-0.11
RoofMatl	Tar&Grv	1	-0.118415	0.084713	-1.40
RoofMatl	WdShake	1	-0.092914	0.060686	-1.53
RoofMatl	WdShngl	1	0.082146	0.049659	1.65
RoofMatl	CompShg	0	0	.	.
ExterQual	Fa	1	-0.035254	0.045199	-0.78
ExterQual	Gd	1	-0.018904	0.021429	-0.88
ExterQual	TA	1	-0.019811	0.023837	-0.83
ExterQual	Ex	0	0	.	.
Heating	GasA	1	0.117467	0.116702	1.01
Heating	GasW	1	0.173356	0.119754	1.45
Heating	Grav	1	-0.046169	0.126763	-0.36
Heating	OthW	1	0.044074	0.145740	0.30
Heating	Wall	1	0.185064	0.131481	1.41
Heating	Floo	0	0	.	.
HeatingQC	Fa	1	-0.010661	0.021355	-0.50
HeatingQC	Gd	1	-0.016664	0.009698	-1.72
HeatingQC	Po	1	-0.117026	0.122321	-0.96
HeatingQC	TA	1	-0.029130	0.009393	-3.10
HeatingQC	Ex	0	0	.	.
CentralAir	Y	1	0.073893	0.017598	4.20
CentralAir	N	0	0	.	.
KitchenQual	Fa	1	-0.078180	0.028576	-2.74
KitchenQual	Gd	1	-0.060045	0.015835	-3.79
KitchenQual	TA	1	-0.068746	0.018059	-3.81
KitchenQual	Ex	0	0	.	.
PavedDrive	P	1	0.023563	0.025705	0.92

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
PavedDrive Y	1	0.033006	0.015543	2.12
PavedDrive N	0	0	.	.
SaleCondition AdjLand	1	0.154199	0.066649	2.31
SaleCondition Allocac	1	0.020989	0.038805	0.54
SaleCondition Family	1	0.037834	0.028608	1.32
SaleCondition Normal	1	0.076843	0.012352	6.22
SaleCondition Partial	1	0.125019	0.017666	7.08
SaleCondition Abnorml	0	0	.	.
Electrical FuseF	1	0.015136	0.026806	0.56
Electrical FuseP	1	-0.039467	0.075520	-0.52
Electrical Mix	1	-0.127185	0.117443	-1.08
Electrical NA	1	0.064715	0.118219	0.55
Electrical SBrkr	1	-0.004576	0.013605	-0.34
Electrical FuseA	0	0	.	.
MSSubClass	1	-0.000152	0.000380	-0.40
LogLotArea	1	0.087960	0.011810	7.45
OverallQual	1	0.058360	0.004454	13.10
OverallCond	1	0.043012	0.003734	11.52
YearBuilt	1	0.002525	0.000294	8.58
YearRemodAdd	1	0.000397	0.000249	1.60
LogBsmtFinSF1	1	0.009318	0.001439	6.48
LogBsmtFinSF2	1	-0.001278	0.001823	-0.70
LogBsmtUnfSF	1	0.004741	0.002048	2.31
_1stFlrSF	1	0.000331	0.000018924	17.51
_2ndFlrSF	1	0.000228	0.000024121	9.44
LowQualFinSF	1	0.000120	0.000083362	1.45
BsmtFullBath	1	0.048206	0.008322	5.79
BsmtHalfBath	1	0.010879	0.013829	0.79
FullBath	1	0.021889	0.010031	2.18

Parameter Estimates				
Parameter	DF	Estimate	Standard Error	t Value
HalfBath	1	0.020464	0.009511	2.15
BedroomAbvGr	1	-0.006355	0.006237	-1.02
KitchenAbvGr	1	-0.067978	0.025600	-2.66
TotRmsAbvGrd	1	0.001962	0.004395	0.45
Fireplaces	1	0.015440	0.006217	2.48
GarageCars	1	0.029901	0.010056	2.97
GarageArea	1	0.000052688	0.000034666	1.52
LogWoodDeckSF	1	0.004483	0.001321	3.39
LogOpenPorchSF	1	0.003076	0.001755	1.75
_3SsnPorch	1	0.000065709	0.000104	0.63
ScreenPorch	1	0.000257	0.000056610	4.54
PoolArea	1	0.000191	0.000083768	2.28
MiscVal	1	-0.000006833	0.000006473	-1.06
MoSold	1	-0.000641	0.001157	-0.55
YrSold	1	-0.003066	0.002376	-1.29