**MSDS 6371 Group Project - D&B All-Stars**

Vanessa Torres

Michael Landrum

Travis Deason

Brian Coari

# Introduction

As we come to the end of our first MSDS semester, this project allowed us to us the fundamental tools of EDA, expand our statistical knowledge, and gain experience with Kaggle. To complete this project we will be using 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa to predict the final price of each home. We will then evaluate our predictions and rank them using several metrics, ultimately choosing the best model.

# Data Description

The Ames Housing data set is a collection of individual residential properties that were sold between 2006 and 2010 in Ames, Iowa. Accumulated by Dean De Cock, this data set has over 2000 observations (2930 to be exact) which afforded us the opportunity to work with variables that included discrete, ordinal, nominal, and continuous. For our first analysis, we focused on the relationship between square footage and sale price for the properties within the North Ames, Eastwood, and Brookside neighborhoods. And for the second analysis we focused on all 28 neighborhoods.

# Analysis Question 1

Mr. Mason,

Thank you for contacting DABAS LLC; we feel confident we can answer your questions about the Ames, Iowa housing market. We were able to track down the most recent 2,930 home sales in Ames as well as 79 characteristics of each house. Since you are only interested in three specific neighborhoods, North Ames (NA), Eastwood (EW) and Brookside (BS), and how square footage (SF) relates to their sale price (SP, in dollars), we will restrict our model to only focus on those variables.

By examining the scatterplot *(see Appendix Plot 1A)*, we concluded we would need to take the log of the sales price and square footage to create a linear relationship. Also, the slope of the regression lines of different neighborhoods would be different, so we included interaction terms. This means our model would look like this:

ln(SP) = B0 + B1 \* ln(SF) + B2 \* BS + B3 \* EW + B4 \* NA + B5 \* ln(SF) \* BS + B6 \* ln(SF) \* EW + B7 \* ln(SF) \* NA

To test our model, we used a null hypothesis that all Bi are zero, or that there is no correlation between sales price and any of the other variables. And our alternative hypothesis is that at least one Bi is not zero, or that there is a correlation between at least one of the variables and sales price. We received a p-value well below 0.05 *(it actually was below 0.001, see Appendix Table 1B column 6)*, so we rejected the null hypothesis and concluded there was a relationship between the sales price and the other variables.

We then needed to test whether each individual variable was needed in our model. For each variable, our null hypothesis was that given the other variables contribution to the model, that particular variable did not help our model (or that the coefficient in front of that variable was 0). And our alternative hypothesis was that the variable contributed to the model (or that the coefficient in front of the variable was not 0). For each variable aside from BS, including the interaction terms, we received p-value below 0.05 (*see Appendix Table 1C column 6*), which means for each variable we rejected the null hypothesis and concluded that they were all beneficial to our model. Since the interaction term was significant that included BS, we kept BS in the model. The full regression equation is:

**ln(SP) = 5.16 + 0.95 ln(SF) + 1.04 BS + 2.40 EW + 3.33 NA – 0.17 BS ln(SF) - 0.36 EW ln(SF) - 0.48 NA ln(SF)**

Note: A zero for all categorical variables (BS, EW, NA) would represent the houses outside those three neighborhoods, but still in Ames. We decided to use them in an effort to reduce the overall variance to our model.

This model has an r-squared of 0.5649 and an adjusted r-squared of 0.5368. The r-squared means 56.49% of the variation in sales price, can be explained the variation in neighborhood and square footage. As always, these coefficients are just estimates. For instance, we are 95% confident the true coefficient for ln(Square Footage) is in the interval (0.914,0.984) *(see Appendix 1C, row 3 columns 7 and 8. The other confidence intervals for the rest of the coefficients can be found in the same table.).*

For us to use this model, we had to assume independence even though that likely isn't completely true. Also, this model gave us relatively normal residuals (slightly curved and skewed left, but with 2000+ observations, the Central Limit Theorem should kick in), but it did show some outliers and leverage points *(see Appendix Plot 1D)*. We concluded the outliers were all accurate measurements so we had to include them. The model created after removing the leverage point gave very similar results to our model, so we decided to keep our model and include the leverage points ***(****all confidence intervals for the coefficients contain the coefficients from our full model. See Appendix Table 1E)*.

We can simplify this model a little by separating it by neighborhood. This leaves us with three different equations you can look at depending on which neighborhood you are working with.

**{ ln(SP) | Brookside } = 6.20 + 0.78 ln(SF)**

**{ ln(SP) | Edwards } = 7.56 + 0.59 ln(SF)**

**{ ln(SP) | North Ames } = 8.49 + 0.47 ln(SF)**

In general, when you double the size of a house in Brookside, the sale price increases by a factor of 2.16. When you double the size of a house in Edwards, the sale price increases by a factor of 1.79. And when you double the size of a house in North Ames, the sale price increases by a factor of 1.56.

Note: The model created by removing the other neighborhoods produced the same regression equations for the individual neighborhoods *(See Appendix Table 1F and Reference 1G)*.

Due to the nature of this study (it was an observational study by nature), we cannot infer causation and say the square footage or the neighborhood causes a change in the sales price of the house; there are obviously many confounding variables that play a part. But, we can say there is a correlation between the sales price and the square footage and neighborhood.

# Analysis Question 2

**Restatement of Problem**

Using 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, we will build the most predictive model for sales prices of homes in all of Ames Iowa. This includes all neighborhoods.

Our basic strategy is to use regression statistics to evaluate all the variables included in the dataset, then remove erroneous variables and transform variables in order to provide the best model. We will then perform a Forward, Backward, Stepwise, and Custom Selection process to get the best model possible given the data, and evaluate the adjusted R-Squared, CV Press, and Kaggle scores on each model.

**Overall Data Observations:**

* Number of Observations Read: 1460
* Number of Observations with Missing Values: 339
* Number of Observations Used: 1121

**Checking Assumptions:**

First we will look at the distribution of our response variable, the Sale Price using all of our identified continuous and indicator variables through proc reg.

***(****All plots for the SalePrice are available in the Appendix. Please see Appendix Table 2A).*

In the SalePrice plot group we see a few worrying things. The residual plot seems clustered tightly together with a few outliers.

1. **Normality**: Judging from the qq plot, there seems to be a curvature at the tail ends of the data, which violates normality. This might be explainable by our classification variables, but in addition we see a disturbingly high peak in our histogram, and together they point to a need to transform our data if possible.
2. **Constant variance:** Judging from the scatter plots we do not see a large concern with variance. There are many outliers that are somewhat concerning, but with so many observations that itself wouldn’t be out of the question… still, given the violation in normality we should try a transform and hope this gets better.
3. **Linear trend:** From the predicted value line and an adjusted R-Square of .91 our linear trend looks good here, so we will compare this to our transformed data.
4. We will assume all observations are **independent**.

***(****All plots for the scatterplots for the individual variables are available in the Appendix. Please see Appendix Table 2B).*

Viewing the scatterplots for the individual variables we see many variables with uneven distributions and distant outliers. These data points might also benefit from a log transformation

After taking the logs we see a few improvements in the model:

***(****All plots for the scatterplots for the SalePrice and the individual variables after the log transformations are available in the Appendix. Please see Appendix Tables 2C and 2D).*

After the transformations we see many improvements in the shape of our data. In general the data has become less affected by outliers, and closer to a random cloud.

1. **Normality**: Judging from the qq plot, the curvature at both ends is still present, but far less pronounced. The tails are mostly going to be formed by our large number of categorical variables that have a small number of discrete values. In the middle we see a straight line indicating normality, and the histogram shows a much more normal distribution.
2. **Constant variance:** Across the board the scatter plots are showing data that are more cloud-like and normal in their variance. We still have a large number of variables with values of zero which negatively affects our constant variance, but we do not see a way around this, we will accept it and proceed.
3. **Linear trend:** From the predicted value line we see a slight improvement here, which is also reflected by an increase in our adjusted R-Square to .93. Our linear trend looks even better than before the transformation.
4. We will assume all observations are **independent**.

**Comparing Competing Models**

To begin this analysis, many variables were identified through proc reg as being linear combinations pf other variables, or having no variation that affected the outcome of the regressions. These variables were categorically removed from analysis. Here are the Categories for the elimination of variables, and the variables that fell into the categories:

**Omitted from analysis due to it being a "totals" row:**

* TotalBsmtSF ( = BsmtFinSF1 + BsmtFinSF2 + BsmtUnfSF)
* GrLivArea ( = 1stFlrSF + 2ndFlrSF + LowQualFinSF)

**Omitted for being always the same as another variable:**

* Exterior2nd: CBLOCK ( = Exterior1st CBLOCK)
* BsmtCond: NA ( = BsmtQual NA)
* BsmtFinType1: NA ( = BsmtQual NA)
* Electrical: Mix ( = BsmtCond Po)

**Omitted for being the only entry in valid rows (rows not containing an NA value)**

* Utilities: AllPub
* Condition2: RRAe
* Condition2: RRAn
* RoofStyle: Mansa
* RoofMatl: Metal
* Heating: Floor
* Functional: Sev
* MiscFeature: Gar2

**Forward**

***(****The Table output of the Forward Selection process is available in the Appendix. Please see Appendix Table 2E).*

We ran the Forward Selection process through proc glmselect using the cleaned up and transformed data created through insights from the proc reg function. For example: linear combinations or known non-impactful variables were removed. We used the Default Schwarz Bayesian Criteria (SBC), so the forward selection process starts with no variables, then continues adding variables until one lowers the SBC, then terminates and uses the existing model without the variable that would lower the SBC.

**Backward**

***(****The Table output of the Backward Selection process is available in the Appendix. Please see Appendix Table 2F).*

We ran the Backward Selection process through proc glmselect using the cleaned up and transformed data created through insights from the proc reg function. For example: linear combinations or known non-impactful variables were removed. We used the Default Schwarz Bayesian Criteria (SBC), so the backward selection process starts with all the variables and removes variable until one removed lowers the SBC, then terminated and used the current model with the variable that would have lowers the SBC if removed.

**Stepwise**

***(****The Table output of the Stepwise Selection process is available in the Appendix. Please see Appendix Table 2G).*

We ran the Stepwise Selection process through proc glmselect using the cleaned up and transformed data created through insights from the proc reg function. For example: linear combinations or known non-impactful variables were removed. We used the Default Schwarz Bayesian Criteria (SBC), so the backward Stepwise process starts with no variables, then continues adding variables until one lowers the SBC, each time checking to see if any variables can be removed now that the new variable is added. Once no more variables added would add value and non could be removed without lowering the SBC the process terminates and uses the existing model.

**CUSTOM**

***(****The Table output of the Stepwise Selection process is available in the Appendix. Please see Appendix Table 2H).*

For our Custom model we decided to switch to a Significance-Level elimination and re-run against all three methods. The Forward selection method looked most promising, so we started manually removing the variables that led to the significance level dipping below acceptable levels (.15) until we had the tightest-fitting significant model we could get. This resulted in a model with a higher adjusted R2 and lower CV Press than any other model.

**Conclusion**:

|  |  |  |  |
| --- | --- | --- | --- |
| **Predictive Models** | Adjusted R2 | CV PRESS | Kaggle Score |
| Forward | 0.8488 | 27.26 | .1487 |
| Backward | 0.9342 | 54.25 | .7005 |
| Stepwise | 0.8488 | 27.26 | .1487 |
| CUSTOM | 0.9372 | 17.72 | .4879 |

Looking at the models, it appears like we over-fit the data in our Custom model. The Forward and Stepwise Selection Processes provided the exact same models which proved to be excellent predictors of the test data, but while the Backward Selection Process and our Custom Selection Process looked like it fit our data extremely well, they under-performed against the test data. This is an indictor that we over-fit our train data using our models and the Forward and Stepwise models that had fewer datapoints were able to stay high-level enough to predict the data not in our datasets. This is an excellent example of how the simplest models often perform the best.

# Appendix

Plot 1A:

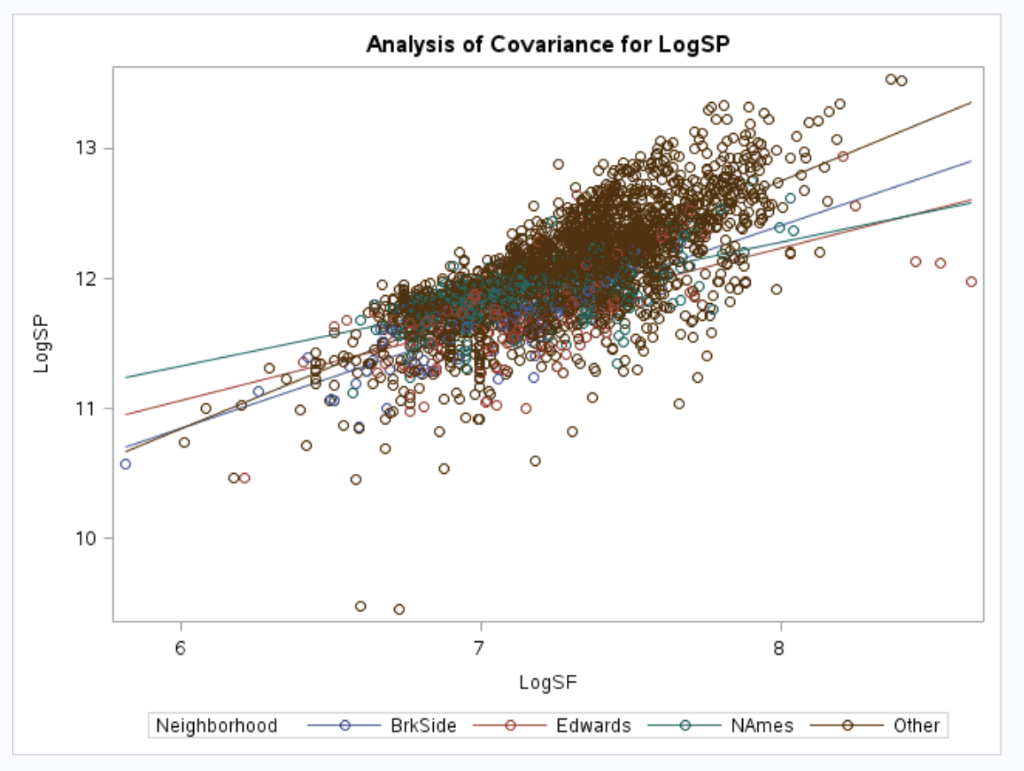


Table 1B:

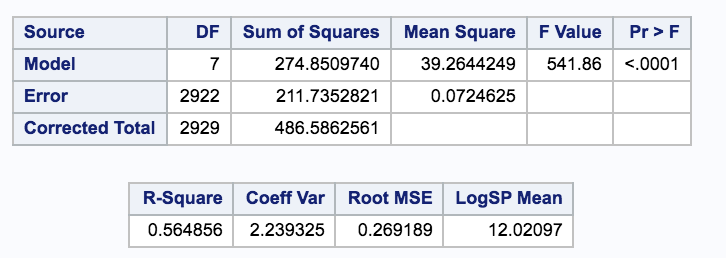
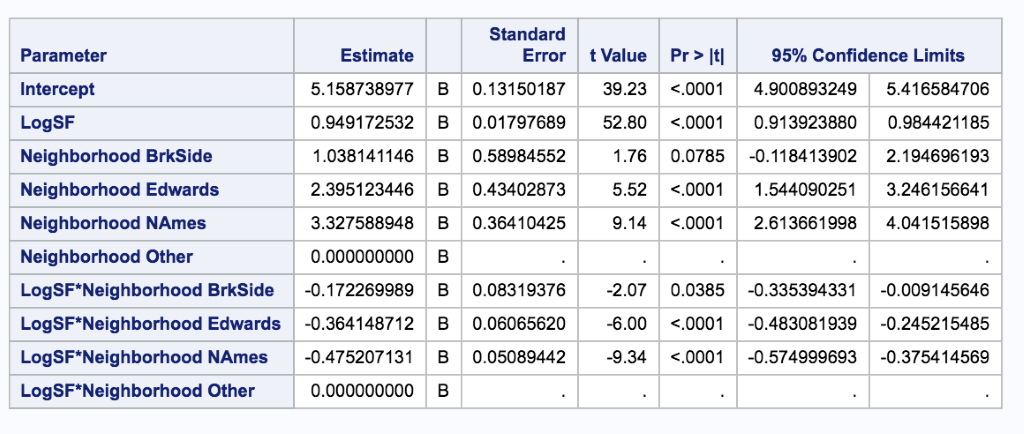


Table 1C:



Plot 1D:

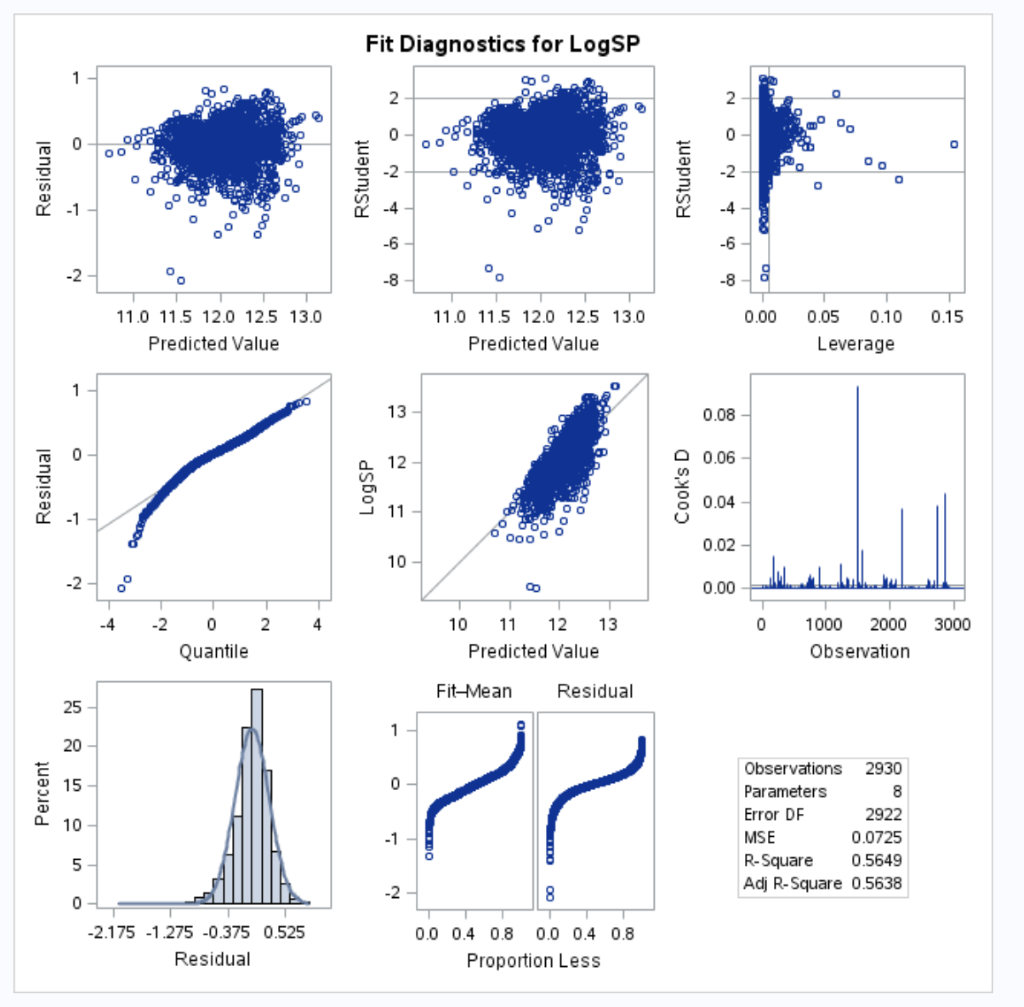


Table 1E:

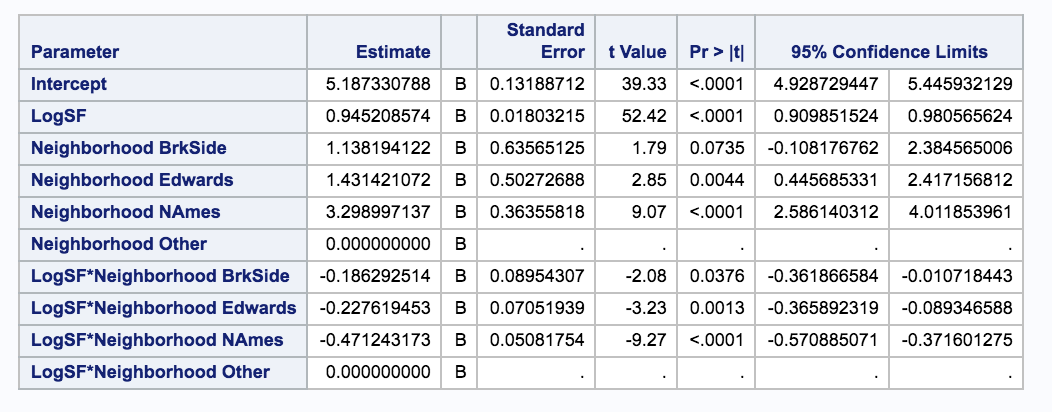
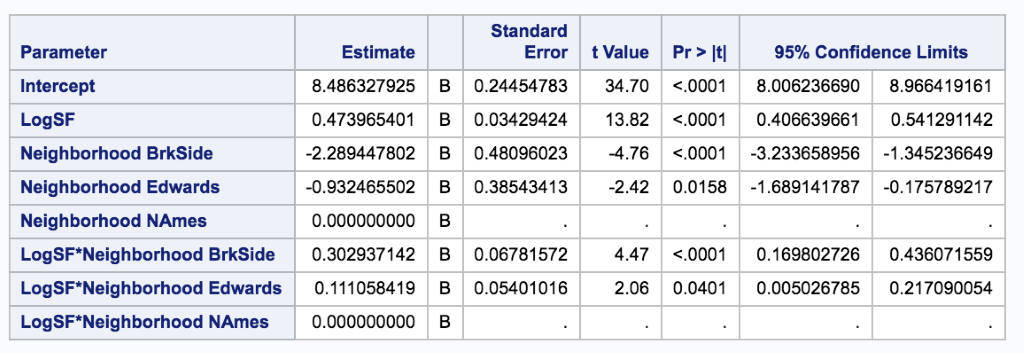


Table 1F:



Reference 1G:

Separated regression equations:

{ ln(SP) | Brookside } = 6.19 + 0.77 ln(SF)

{ ln(SP) | Edwards } = 7.55 + 0.58 ln(SF)

{ ln(SP) | North Ames } = 8.48 + 0.47 ln(SF)

**Question 1 Code:**

R code to turn all other neighborhoods into "other", removed unwanted variables to help with speed of SAS and created log variables:

ah <- read.csv("~/Documents/AmesHousing.csv")

q1 <- ah[,c("Neighborhood","Gr.Liv.Area","SalePrice")] ## removes unwanted variables

colnames(q1) <- c("Neighborhood","SquareFeet","SalePrice") ## Renamed Gr.Liv.Area to Sqaure Feet

q1$LogSF <- log( q1$SquareFeet )

q1$LogSP <- log( q1$SalePrice ) ## create log variables for square feet and sale price

q1$Neighborhood[ Q1$Neighborhood != "BrkSide" & Q1$Neighborhood != "Edwards" & q1$Neighborhood = "NAmes" ] <- "Other" ## turns all the other neighborhoods into "Other"

write.csv(q1,"q1all.csv")

R code to remove leverage points:

sorted <- q1[ order(q1$LogSF,decreasing=TRUE ), ]

row.names(sorted) <- 1:nrow(sorted) ## sorted and changed row names to easily find highest/lowest

sorted2 <- sorted[-c(1,2,3,4,5,2930),] ## 5 highest were above 4000 sq ft. Lowest was also a leverage pt

write.csv(sorted2,"nl.csv") ## that's an L, for "no leverage"

R code to remove all other neighborhoods:

q1some <- q1[ q1$Neighborhood=="NAmes" | q1$Neighborhood=="Edwards" | q1$Neighborhood == "BrkSide" , ]

write.csv(q1some,"q1some.csv")

All SAS code was the same, but with different DATANAME

proc glm data=DATANAME plots=all;

class Neighborhood;

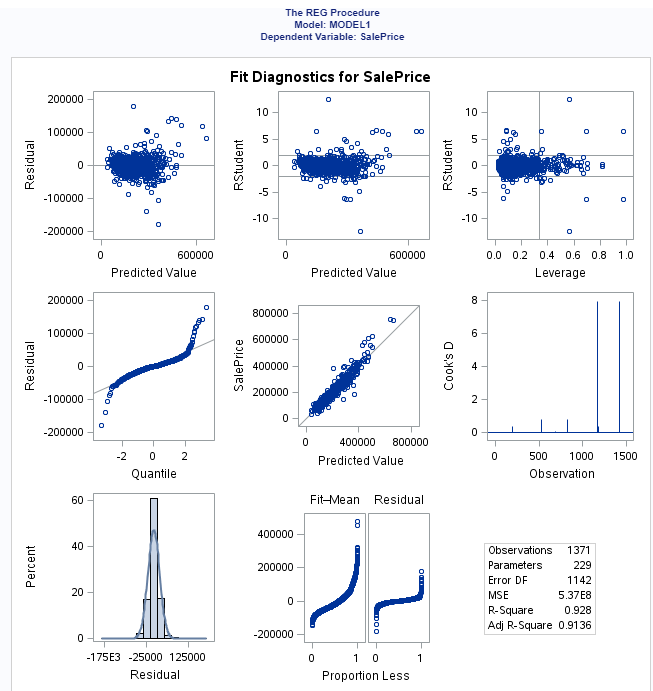
model LogSP = LogSF | Neighborhood / solution clparm;

run;

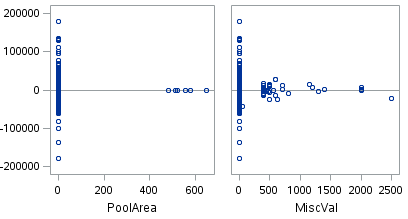
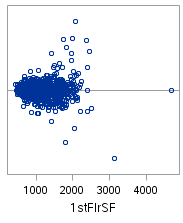
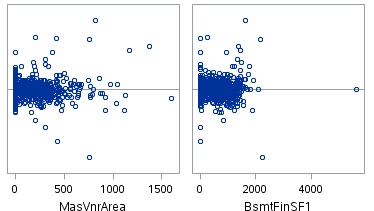
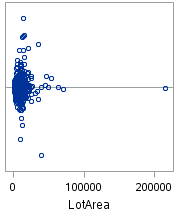
Plot 1A, Table 1B, Table 1C, and plot for 1D were generated using "q1all". Table 1E used "nl". Table 1F used "q1some".

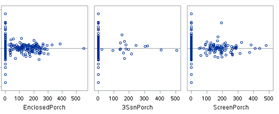
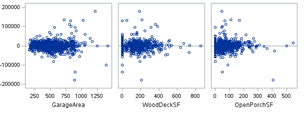
**Question 2 Prog Reg Output Before Log Transformations:**

Plot Group 2A:



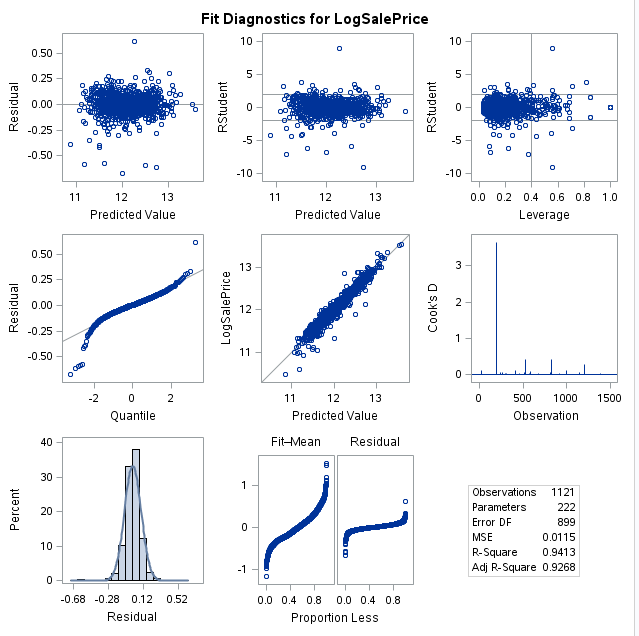
Plot Group 2B:



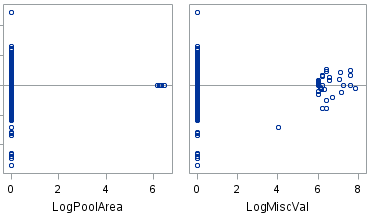
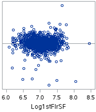
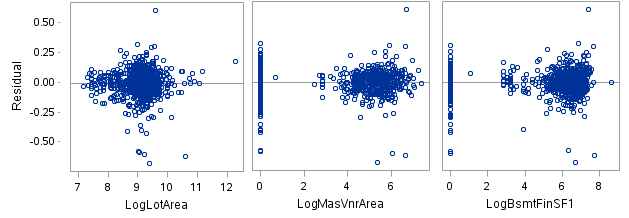


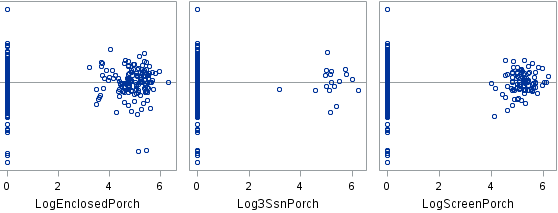
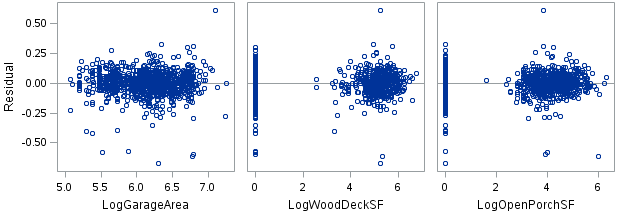
**Question 2 Prog Reg Output After Log Transformations:**

Plot Group 2C:



Plot Group 2D:





**Question 2 Forward Selection Output:**

Table 2E:

**The GLMSELECT Procedure**

|  |  |
| --- | --- |
| **Data Set** | WORK.TRAIN\_DUMMY |
| **Dependent Variable** | LogSalePrice |
| **Selection Method** | Forward |
| **Select Criterion** | SBC |
| **Stop Criterion** | SBC |
| **Effect Hierarchy Enforced** | None |

|  |  |
| --- | --- |
| **Number of Observations Read** | 1460 |
| **Number of Observations Used** | 1121 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **MSZoning** | 5 | C FV RH RL RM |
| **Street** | 2 | Grvl Pave |
| **Alley** | 3 | Gr NA Pa |
| **LotShape** | 4 | IR1 IR2 IR3 Reg |
| **LandContour** | 4 | Bnk HLS Low Lvl |
| **Utilities** | 1 | AllPub |
| **LotConfig** | 5 | Corner CulDSac FR2 FR3 Inside |
| **LandSlope** | 3 | Gtl Mod Sev |
| **Neighborhood** | 25 | Blmngtn Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker |
| **Condition1** | 9 | Artery Feedr Norm PosA PosN RRAe RRAn RRNe RRNn |
| **Condition2** | 6 | Artery Feedr Norm PosA PosN RRNn |
| **BldgType** | 5 | 1Fam 2fmCon Duplex Twnhs TwnhsE |
| **HouseStyle** | 8 | 1.5Fin 1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl |
| **RoofStyle** | 5 | Flat Gable Gambr Hip Mansa |
| **RoofMatl** | 7 | ClyTile CompShg Membran Roll Tar&Grv WdShake WdShngl |
| **Exterior1st** | 14 | AsbShng BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing |
| **Exterior2nd** | 16 | AsbShng AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng |
| **MasVnrType** | 4 | BrkCmn BrkFace None Stone |
| **ExterQual** | 4 | Ex Fa Gd TA |
| **ExterCond** | 4 | Ex Fa Gd TA |
| **Foundation** | 6 | BrkTil CBlock PConc Slab Stone Wood |
| **BsmtQual** | 5 | Ex Fa Gd NA TA |
| **BsmtCond** | 5 | Fa Gd NA Po TA |
| **BsmtExposure** | 5 | Av Gd Mn NA No |
| **BsmtFinType1** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **BsmtFinType2** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **Heating** | 5 | GasA GasW Grav OthW Wall |
| **HeatingQC** | 5 | Ex Fa Gd Po TA |
| **CentralAir** | 2 | N Y |
| **Electrical** | 6 | FuseA FuseF FuseP Mix NA SBrkr |
| **KitchenQual** | 4 | Ex Fa Gd TA |
| **Functional** | 6 | Maj1 Maj2 Min1 Min2 Mod Typ |
| **FireplaceQu** | 6 | Ex Fa Gd NA Po TA |
| **GarageType** | 6 | 2Types Attchd Basment BuiltIn CarPort Detchd |
| **GarageFinish** | 3 | Fin RFn Unf |
| **GarageQual** | 5 | Ex Fa Gd Po TA |
| **GarageCond** | 5 | Ex Fa Gd Po TA |
| **PavedDrive** | 3 | N P Y |
| **PoolQC** | 4 | Ex Fa Gd NA |
| **Fence** | 5 | GdPrv GdWo MnPrv MnWw NA |
| **MiscFeature** | 4 | NA Othr Shed TenC |
| **SaleType** | 6 | COD CWD Con New Oth WD |
| **SaleCondition** | 6 | Abnorml AdjLand Alloca Family Normal Partial |

| **Dimensions** | |
| --- | --- |
| **Number of Effects** | 78 |
| **Number of Parameters** | 285 |

**The GLMSELECT Procedure**

| **Forward Selection Summary** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Step** | **Effect Entered** | **Number Effects In** | **Number Parms In** | **Model R-Square** | **SBC** | **PRESS** |
| **\* Optimal Value of Criterion** | | | | | | |
| **0** | **Intercept** | 1 | 1 | 0.0000 | -2066.4335 | 176.6399 |
| **1** | **OverallQual** | 2 | 2 | 0.6851 | -3354.5767 | 55.7616 |
| **2** | **Log1stFlrSF** | 3 | 3 | 0.7412 | -3567.6573 | 45.9654 |
| **3** | **2ndFlrSF** | 4 | 4 | 0.7818 | -3751.9351 | 38.9323 |
| **4** | **YearBuilt** | 5 | 5 | 0.8147 | -3928.0564 | 33.1553 |
| **5** | **OverallCond** | 6 | 6 | 0.8382 | -4073.1053 | 29.0219 |
| **6** | **LogLotArea** | 7 | 7 | 0.8496 | -4147.7963 | 27.0661 |
| **7** | **RoofMatl** | 8 | 13 | 0.8656 | -4232.0952 | 24.1556 |
| **8** | **LogBsmtFinSF1** | 9 | 14 | 0.8744 | -4300.5435 | 22.6489 |
| **9** | **Neighborhood** | 10 | 38 | 0.8985 | -4371.1382 | 19.2971 |
| **10** | **GarageCars** | 11 | 39 | 0.9018 | -4400.6906 | 18.7174 |
| **11** | **KitchenAbvGr** | 12 | 40 | 0.9055 | -4436.9303 | 18.0567 |
| **12** | **MSZoning** | 13 | 44 | 0.9101 | -4465.5009 | 17.4914 |
| **13** | **Condition2** | 14 | 49 | 0.9154 | -4498.0609 | 18.8515 |
| **14** | **KitchenQual** | 15 | 52 | 0.9182 | -4514.9270 | 18.5146 |
| **15** | **SaleCondition** | 16 | 57 | 0.9214 | -4523.8525 | 18.2602 |
| **16** | **BsmtFullBath** | 17 | 58 | 0.9226 | -4533.9616 | 18.0625 |
| **17** | **CentralAir** | 18 | 59 | 0.9237 | -4542.9937 | 17.8060 |
| **18** | **GarageFinish** | 19 | 61 | 0.9250 | -4548.0730 | 17.5431 |
| **19** | **LogScreenPorch** | 20 | 62 | 0.9257 | -4552.6909 | 17.3707 |
| **20** | **LogOpenPorchSF** | 21 | 63 | 0.9263 | -4554.1894 | 17.2921 |
| **21** | **LogWoodDeckSF** | 22 | 64 | 0.9269 | -4555.8269\* | 17.1852\* |

|  |
| --- |
| Selection stopped at a local minimum of the SBC criterion. |

| **Stop Details** | | | | |
| --- | --- | --- | --- | --- |
| **Candidate For** | **Effect** | **Candidate SBC** |  | **Compare SBC** |
| **Entry** | LogEnclosedPorch | -4554.8188 | > | -4555.8269 |

**The GLMSELECT Procedure**

**Selected Model**

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

|  |  |
| --- | --- |
| **Effects:** | Intercept MSZoning LogLotArea Neighborhood Condition2 OverallQual OverallCond YearBuilt RoofMatl LogBsmtFinSF1 CentralAir Log1stFlrSF 2ndFlrSF BsmtFullBath KitchenAbvGr KitchenQual GarageFinish GarageCars LogWoodDeckSF LogOpenPorchSF LogScreenPorch SaleCondition |

| **Analysis of Variance** | | | | |
| --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** |
| **Model** | 63 | 163.42779 | 2.59409 | 212.60 |
| **Error** | 1057 | 12.89713 | 0.01220 |  |
| **Corrected Total** | 1120 | 176.32491 |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 0.11046 |
| **Dependent Mean** | 12.04841 |
| **R-Square** | 0.9269 |
| **Adj R-Sq** | 0.9225 |
| **AIC** | -3754.23342 |
| **AICC** | -3746.10071 |
| **PRESS** | 17.18516 |
| **SBC** | -4555.82692 |

| **Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** |
| **Intercept** | 1 | 2.433351 | 0.664000 | 3.66 |
| **MSZoning C** | 1 | -0.326476 | 0.045596 | -7.16 |
| **MSZoning FV** | 1 | 0.097795 | 0.034641 | 2.82 |
| **MSZoning RH** | 1 | 0.040363 | 0.040920 | 0.99 |
| **MSZoning RL** | 1 | 0.042573 | 0.018581 | 2.29 |
| **MSZoning RM** | 0 | 0 | . | . |
| **LogLotArea** | 1 | 0.095983 | 0.011631 | 8.25 |
| **Neighborhood Blmngtn** | 1 | -0.045403 | 0.054670 | -0.83 |
| **Neighborhood Blueste** | 1 | 0.016407 | 0.093330 | 0.18 |
| **Neighborhood BrDale** | 1 | 0.021599 | 0.058062 | 0.37 |
| **Neighborhood BrkSide** | 1 | 0.039986 | 0.050699 | 0.79 |
| **Neighborhood ClearCr** | 1 | -0.059126 | 0.054649 | -1.08 |
| **Neighborhood CollgCr** | 1 | -0.039486 | 0.044405 | -0.89 |
| **Neighborhood Crawfor** | 1 | 0.112287 | 0.047981 | 2.34 |
| **Neighborhood Edwards** | 1 | -0.123455 | 0.045871 | -2.69 |
| **Neighborhood Gilbert** | 1 | -0.045361 | 0.046535 | -0.97 |
| **Neighborhood IDOTRR** | 1 | 0.013611 | 0.055717 | 0.24 |
| **Neighborhood MeadowV** | 1 | -0.100128 | 0.060629 | -1.65 |
| **Neighborhood Mitchel** | 1 | -0.081181 | 0.047478 | -1.71 |
| **Neighborhood NAmes** | 1 | -0.077424 | 0.044332 | -1.75 |
| **Neighborhood NPkVill** | 1 | -0.026667 | 0.062613 | -0.43 |
| **Neighborhood NWAmes** | 1 | -0.106910 | 0.046026 | -2.32 |
| **Neighborhood NoRidge** | 1 | 0.019621 | 0.047888 | 0.41 |
| **Neighborhood NridgHt** | 1 | 0.062033 | 0.045641 | 1.36 |
| **Neighborhood OldTown** | 1 | -0.032773 | 0.050600 | -0.65 |
| **Neighborhood SWISU** | 1 | -0.022331 | 0.052776 | -0.42 |
| **Neighborhood Sawyer** | 1 | -0.094850 | 0.046484 | -2.04 |
| **Neighborhood SawyerW** | 1 | -0.068615 | 0.046367 | -1.48 |
| **Neighborhood Somerst** | 1 | -0.002462 | 0.050162 | -0.05 |
| **Neighborhood StoneBr** | 1 | 0.096736 | 0.050136 | 1.93 |
| **Neighborhood Timber** | 1 | -0.015773 | 0.047616 | -0.33 |
| **Neighborhood Veenker** | 0 | 0 | . | . |
| **Condition2 Artery** | 1 | -0.037195 | 0.115706 | -0.32 |
| **Condition2 Feedr** | 1 | -0.056835 | 0.094788 | -0.60 |
| **Condition2 Norm** | 1 | 0.068464 | 0.080871 | 0.85 |
| **Condition2 PosA** | 1 | 0.214739 | 0.139874 | 1.54 |
| **Condition2 PosN** | 1 | -0.634447 | 0.114831 | -5.53 |
| **Condition2 RRNn** | 0 | 0 | . | . |
| **OverallQual** | 1 | 0.058897 | 0.004808 | 12.25 |
| **OverallCond** | 1 | 0.047338 | 0.003995 | 11.85 |
| **YearBuilt** | 1 | 0.002621 | 0.000304 | 8.61 |
| **RoofMatl ClyTile** | 1 | -1.628480 | 0.124099 | -13.12 |
| **RoofMatl CompShg** | 1 | -0.132856 | 0.047591 | -2.79 |
| **RoofMatl Membran** | 1 | 0.022952 | 0.123788 | 0.19 |
| **RoofMatl Roll** | 1 | -0.130134 | 0.122264 | -1.06 |
| **RoofMatl Tar&Grv** | 1 | -0.202109 | 0.065408 | -3.09 |
| **RoofMatl WdShake** | 1 | -0.204848 | 0.092597 | -2.21 |
| **RoofMatl WdShngl** | 0 | 0 | . | . |
| **LogBsmtFinSF1** | 1 | 0.011119 | 0.001503 | 7.40 |
| **CentralAir N** | 1 | -0.073293 | 0.017521 | -4.18 |
| **CentralAir Y** | 0 | 0 | . | . |
| **Log1stFlrSF** | 1 | 0.404321 | 0.018699 | 21.62 |
| **2ndFlrSF** | 1 | 0.000246 | 0.000011022 | 22.29 |
| **BsmtFullBath** | 1 | 0.031820 | 0.008593 | 3.70 |
| **KitchenAbvGr** | 1 | -0.086124 | 0.020168 | -4.27 |
| **KitchenQual Ex** | 1 | 0.097355 | 0.018262 | 5.33 |
| **KitchenQual Fa** | 1 | -0.010269 | 0.025318 | -0.41 |
| **KitchenQual Gd** | 1 | 0.014719 | 0.010234 | 1.44 |
| **KitchenQual TA** | 0 | 0 | . | . |
| **GarageFinish Fin** | 1 | 0.042068 | 0.011788 | 3.57 |
| **GarageFinish RFn** | 1 | 0.030232 | 0.010220 | 2.96 |
| **GarageFinish Unf** | 0 | 0 | . | . |
| **GarageCars** | 1 | 0.059185 | 0.007604 | 7.78 |
| **LogWoodDeckSF** | 1 | 0.004205 | 0.001469 | 2.86 |
| **LogOpenPorchSF** | 1 | 0.005996 | 0.001976 | 3.03 |
| **LogScreenPorch** | 1 | 0.008624 | 0.002427 | 3.55 |
| **SaleCondition Abnorml** | 1 | -0.118718 | 0.018883 | -6.29 |
| **SaleCondition AdjLand** | 1 | 0.159529 | 0.113486 | 1.41 |
| **SaleCondition Alloca** | 1 | -0.133985 | 0.042298 | -3.17 |
| **SaleCondition Family** | 1 | -0.107906 | 0.029625 | -3.64 |
| **SaleCondition Normal** | 1 | -0.055147 | 0.012949 | -4.26 |
| **SaleCondition Partial** | 0 | 0 | . | . |

**The GLMSELECT Procedure**

|  |  |
| --- | --- |
| **Data Set** | WORK.TRAIN\_DUMMY |
| **Dependent Variable** | LogSalePrice |
| **Selection Method** | Backward |
| **Select Criterion** | SBC |
| **Stop Criterion** | SBC |
| **Effect Hierarchy Enforced** | None |

|  |  |
| --- | --- |
| **Number of Observations Read** | 1460 |
| **Number of Observations Used** | 1121 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **MSZoning** | 5 | C FV RH RL RM |
| **Street** | 2 | Grvl Pave |
| **Alley** | 3 | Gr NA Pa |
| **LotShape** | 4 | IR1 IR2 IR3 Reg |
| **LandContour** | 4 | Bnk HLS Low Lvl |
| **Utilities** | 1 | AllPub |
| **LotConfig** | 5 | Corner CulDSac FR2 FR3 Inside |
| **LandSlope** | 3 | Gtl Mod Sev |
| **Neighborhood** | 25 | Blmngtn Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker |
| **Condition1** | 9 | Artery Feedr Norm PosA PosN RRAe RRAn RRNe RRNn |
| **Condition2** | 6 | Artery Feedr Norm PosA PosN RRNn |
| **BldgType** | 5 | 1Fam 2fmCon Duplex Twnhs TwnhsE |
| **HouseStyle** | 8 | 1.5Fin 1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl |
| **RoofStyle** | 5 | Flat Gable Gambr Hip Mansa |
| **RoofMatl** | 7 | ClyTile CompShg Membran Roll Tar&Grv WdShake WdShngl |
| **Exterior1st** | 14 | AsbShng BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing |
| **Exterior2nd** | 16 | AsbShng AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng |
| **MasVnrType** | 4 | BrkCmn BrkFace None Stone |
| **ExterQual** | 4 | Ex Fa Gd TA |
| **ExterCond** | 4 | Ex Fa Gd TA |
| **Foundation** | 6 | BrkTil CBlock PConc Slab Stone Wood |
| **BsmtQual** | 5 | Ex Fa Gd NA TA |
| **BsmtCond** | 5 | Fa Gd NA Po TA |
| **BsmtExposure** | 5 | Av Gd Mn NA No |
| **BsmtFinType1** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **BsmtFinType2** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **Heating** | 5 | GasA GasW Grav OthW Wall |
| **HeatingQC** | 5 | Ex Fa Gd Po TA |
| **CentralAir** | 2 | N Y |
| **Electrical** | 6 | FuseA FuseF FuseP Mix NA SBrkr |
| **KitchenQual** | 4 | Ex Fa Gd TA |
| **Functional** | 6 | Maj1 Maj2 Min1 Min2 Mod Typ |
| **FireplaceQu** | 6 | Ex Fa Gd NA Po TA |
| **GarageType** | 6 | 2Types Attchd Basment BuiltIn CarPort Detchd |
| **GarageFinish** | 3 | Fin RFn Unf |
| **GarageQual** | 5 | Ex Fa Gd Po TA |
| **GarageCond** | 5 | Ex Fa Gd Po TA |
| **PavedDrive** | 3 | N P Y |
| **PoolQC** | 4 | Ex Fa Gd NA |
| **Fence** | 5 | GdPrv GdWo MnPrv MnWw NA |
| **MiscFeature** | 4 | NA Othr Shed TenC |
| **SaleType** | 6 | COD CWD Con New Oth WD |
| **SaleCondition** | 6 | Abnorml AdjLand Alloca Family Normal Partial |

| **Dimensions** | |
| --- | --- |
| **Number of Effects** | 78 |
| **Number of Parameters** | 285 |

**The GLMSELECT Procedure**

| **Backward Selection Summary** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Step** | **Effect Removed** | **Number Effects In** | **Number Parms In** | **Model R-Square** | **SBC** | **PRESS** |
| **\* Optimal Value of Criterion** | | | | | | |
| **0** |  | 78 | 238 | 0.9491 | -3741.3243 | 19.6953 |
|  | **Utilities** | 77 | 238 | 0.9491 | -3741.3243 | 19.6953 |
| **1** | **Exterior2nd** | 76 | 224 | 0.9479 | -3813.1678 | 19.1923 |
| **2** | **Exterior1st** | 75 | 211 | 0.9465 | -3873.6103 | 18.4814 |
| **3** | **HouseStyle** | 74 | 204 | 0.9461 | -3915.2076 | 18.4250 |
| **4** | **BsmtFinType2** | 73 | 198 | 0.9458 | -3951.0012 | 18.3605 |
| **5** | **FireplaceQu** | 72 | 193 | 0.9455 | -3980.4038 | 18.1648 |
| **6** | **GarageType** | 71 | 188 | 0.9452 | -4009.1920 | 18.1249 |
| **7** | **Condition1** | 70 | 180 | 0.9437 | -4034.9265 | 18.6901 |
| **8** | **Fence** | 69 | 176 | 0.9436 | -4059.9306 | 18.6099 |
| **9** | **RoofStyle** | 68 | 172 | 0.9433 | -4083.3905 | 18.5748 |
| **10** | **GarageQual** | 67 | 168 | 0.9430 | -4105.9290 | 17.9521 |
| **11** | **GarageCond** | 66 | 164 | 0.9428 | -4129.3149 | 17.8495 |
| **12** | **Electrical** | 65 | 160 | 0.9426 | -4152.9985 | 17.6752 |
| **13** | **BsmtFinType1** | 64 | 155 | 0.9420 | -4176.4612 | 17.6145 |
| **14** | **BldgType** | 63 | 151 | 0.9416 | -4197.8929 | 17.6004 |
| **15** | **LandContour** | 62 | 148 | 0.9416 | -4218.4252 | 17.3774 |
| **16** | **ExterQual** | 61 | 145 | 0.9415 | -4238.1265 | 17.2663 |
| **17** | **PoolQC** | 60 | 142 | 0.9414 | -4257.3543 | 17.3000 |
| **18** | **LotShape** | 59 | 139 | 0.9413 | -4275.7642 | 17.2720 |
| **19** | **BsmtCond** | 58 | 136 | 0.9411 | -4293.9910 | 17.1940 |
| **20** | **SaleType** | 57 | 131 | 0.9403 | -4312.3393 | 17.3116 |
| **21** | **MiscFeature** | 56 | 128 | 0.9400 | -4327.6418 | 17.1233 |
| **22** | **MasVnrType** | 55 | 125 | 0.9396 | -4342.2036 | 16.8879\* |
| **23** | **BsmtQual** | 54 | 121 | 0.9388 | -4355.7520 | 16.9598 |
| **24** | **LotConfig** | 53 | 117 | 0.9380 | -4368.7718 | 17.0984 |
| **25** | **ExterCond** | 52 | 114 | 0.9376 | -4382.4912 | 17.0384 |
| **26** | **Foundation** | 51 | 109 | 0.9363 | -4395.2047 | 17.1120 |
| **27** | **PavedDrive** | 50 | 107 | 0.9362 | -4406.6127 | 17.1010 |
| **28** | **Alley** | 49 | 105 | 0.9359 | -4416.3882 | 17.0622 |
| **29** | **GarageFinish** | 48 | 103 | 0.9357 | -4425.8371 | 17.1176 |
| **30** | **HeatingQC** | 47 | 99 | 0.9345 | -4433.3414 | 17.3184 |
| **31** | **Heating** | 46 | 95 | 0.9333 | -4441.0348 | 17.6711 |
| **32** | **LandSlope** | 45 | 93 | 0.9329 | -4449.0608 | 17.4035 |
| **33** | **BsmtUnfSF** | 44 | 92 | 0.9329 | -4456.0320 | 17.3515 |
| **34** | **YrSold** | 43 | 91 | 0.9329 | -4462.9677 | 17.3242 |
| **35** | **LogMasVnrArea** | 42 | 90 | 0.9329 | -4469.8403 | 17.3057 |
| **36** | **BsmtHalfBath** | 41 | 89 | 0.9329 | -4476.5359 | 17.2867 |
| **37** | **IntLotFrontage** | 40 | 88 | 0.9328 | -4483.0655 | 17.2637 |
| **38** | **LowQualFinSF** | 39 | 87 | 0.9328 | -4489.4381 | 17.2378 |
| **39** | **LogMiscVal** | 38 | 86 | 0.9328 | -4495.7733 | 17.2248 |
| **40** | **BsmtFinSF2** | 37 | 85 | 0.9327 | -4501.7542 | 17.2312 |
| **41** | **GarageYrBlt** | 36 | 84 | 0.9326 | -4507.6361 | 17.1765 |
| **42** | **MSSubClass** | 35 | 83 | 0.9326 | -4513.3747 | 17.1732 |
| **43** | **LogPoolArea** | 34 | 82 | 0.9325 | -4519.0627 | 17.1502 |
| **44** | **Log3SsnPorch** | 33 | 81 | 0.9324 | -4524.3939 | 17.1419 |
| **45** | **MoSold** | 32 | 80 | 0.9323 | -4529.5626 | 17.1569 |
| **46** | **TotRmsAbvGrd** | 31 | 79 | 0.9321 | -4534.3710 | 17.1642 |
| **47** | **BedroomAbvGr** | 30 | 78 | 0.9321 | -4540.2189 | 17.0594 |
| **48** | **LogEnclosedPorch** | 29 | 77 | 0.9318 | -4543.5440 | 17.0665 |
| **49** | **LogGarageArea** | 28 | 76 | 0.9315 | -4545.3370 | 17.1443 |
| **50** | **HalfBath** | 27 | 75 | 0.9312 | -4547.3069 | 17.1776 |
| **51** | **BsmtExposure** | 26 | 71 | 0.9296 | -4549.1911 | 17.0857 |
| **52** | **Street** | 25 | 70 | 0.9293 | -4551.9157 | 17.0834 |
| **53** | **FullBath** | 24 | 69 | 0.9289 | -4553.2699 | 17.0938 |
| **54** | **Fireplaces** | 23 | 68 | 0.9285 | -4553.8875\* | 17.1811 |

**Note:**Effects dropped at step 0 are redundant.

|  |
| --- |
| Selection stopped at a local minimum of the SBC criterion. |

| **Stop Details** | | | | |
| --- | --- | --- | --- | --- |
| **Candidate For** | **Effect** | **Candidate SBC** |  | **Compare SBC** |
| **Removal** | Functional | -4553.0503 | > | -4553.8875 |

**The GLMSELECT Procedure**

**Selected Model**

**The selected model is the model at the last step (Step 54).**

|  |  |
| --- | --- |
| **Effects:** | Intercept MSZoning LogLotArea Neighborhood Condition2 OverallQual OverallCond YearBuilt YearRemodAdd RoofMatl LogBsmtFinSF1 CentralAir Log1stFlrSF 2ndFlrSF BsmtFullBath KitchenAbvGr KitchenQual Functional GarageCars LogWoodDeckSF LogOpenPorchSF LogScreenPorch SaleCondition |

| **Analysis of Variance** | | | | |
| --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** |
| **Model** | 67 | 163.72514 | 2.44366 | 204.22 |
| **Error** | 1053 | 12.59977 | 0.01197 |  |
| **Corrected Total** | 1120 | 176.32491 |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 0.10939 |
| **Dependent Mean** | 12.04841 |
| **R-Square** | 0.9285 |
| **Adj R-Sq** | 0.9240 |
| **AIC** | -3772.38187 |
| **AICC** | -3763.19062 |
| **PRESS** | 17.18111 |
| **SBC** | -4553.88747 |

| **Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** |
| **Intercept** | 1 | 0.924064 | 0.755859 | 1.22 |
| **MSZoning C** | 1 | -0.327061 | 0.045026 | -7.26 |
| **MSZoning FV** | 1 | 0.098583 | 0.034384 | 2.87 |
| **MSZoning RH** | 1 | 0.033519 | 0.040576 | 0.83 |
| **MSZoning RL** | 1 | 0.042862 | 0.018499 | 2.32 |
| **MSZoning RM** | 0 | 0 | . | . |
| **LogLotArea** | 1 | 0.094030 | 0.011642 | 8.08 |
| **Neighborhood Blmngtn** | 1 | -0.059404 | 0.054086 | -1.10 |
| **Neighborhood Blueste** | 1 | 0.035772 | 0.092322 | 0.39 |
| **Neighborhood BrDale** | 1 | 0.011598 | 0.057416 | 0.20 |
| **Neighborhood BrkSide** | 1 | 0.033468 | 0.050141 | 0.67 |
| **Neighborhood ClearCr** | 1 | -0.066868 | 0.054110 | -1.24 |
| **Neighborhood CollgCr** | 1 | -0.058847 | 0.043929 | -1.34 |
| **Neighborhood Crawfor** | 1 | 0.102033 | 0.047417 | 2.15 |
| **Neighborhood Edwards** | 1 | -0.130993 | 0.045369 | -2.89 |
| **Neighborhood Gilbert** | 1 | -0.054583 | 0.046118 | -1.18 |
| **Neighborhood IDOTRR** | 1 | 0.009742 | 0.055145 | 0.18 |
| **Neighborhood MeadowV** | 1 | -0.107620 | 0.060020 | -1.79 |
| **Neighborhood Mitchel** | 1 | -0.096081 | 0.046942 | -2.05 |
| **Neighborhood NAmes** | 1 | -0.079271 | 0.043874 | -1.81 |
| **Neighborhood NPkVill** | 1 | -0.047545 | 0.061855 | -0.77 |
| **Neighborhood NWAmes** | 1 | -0.111066 | 0.045547 | -2.44 |
| **Neighborhood NoRidge** | 1 | -0.000878 | 0.047400 | -0.02 |
| **Neighborhood NridgHt** | 1 | 0.040038 | 0.045142 | 0.89 |
| **Neighborhood OldTown** | 1 | -0.042151 | 0.050029 | -0.84 |
| **Neighborhood SWISU** | 1 | -0.028047 | 0.052213 | -0.54 |
| **Neighborhood Sawyer** | 1 | -0.098698 | 0.046039 | -2.14 |
| **Neighborhood SawyerW** | 1 | -0.088909 | 0.045757 | -1.94 |
| **Neighborhood Somerst** | 1 | -0.024130 | 0.049629 | -0.49 |
| **Neighborhood StoneBr** | 1 | 0.083604 | 0.049659 | 1.68 |
| **Neighborhood Timber** | 1 | -0.028047 | 0.047171 | -0.59 |
| **Neighborhood Veenker** | 0 | 0 | . | . |
| **Condition2 Artery** | 1 | -0.038382 | 0.114593 | -0.33 |
| **Condition2 Feedr** | 1 | -0.071294 | 0.093752 | -0.76 |
| **Condition2 Norm** | 1 | 0.055885 | 0.079951 | 0.70 |
| **Condition2 PosA** | 1 | 0.178098 | 0.138286 | 1.29 |
| **Condition2 PosN** | 1 | -0.662616 | 0.113545 | -5.84 |
| **Condition2 RRNn** | 0 | 0 | . | . |
| **OverallQual** | 1 | 0.058723 | 0.004789 | 12.26 |
| **OverallCond** | 1 | 0.038043 | 0.004371 | 8.70 |
| **YearBuilt** | 1 | 0.002556 | 0.000308 | 8.31 |
| **YearRemodAdd** | 1 | 0.000795 | 0.000275 | 2.90 |
| **RoofMatl ClyTile** | 1 | -1.660946 | 0.122995 | -13.50 |
| **RoofMatl CompShg** | 1 | -0.124295 | 0.047165 | -2.64 |
| **RoofMatl Membran** | 1 | 0.078112 | 0.127982 | 0.61 |
| **RoofMatl Roll** | 1 | -0.126267 | 0.121087 | -1.04 |
| **RoofMatl Tar&Grv** | 1 | -0.202707 | 0.064831 | -3.13 |
| **RoofMatl WdShake** | 1 | -0.201635 | 0.091645 | -2.20 |
| **RoofMatl WdShngl** | 0 | 0 | . | . |
| **LogBsmtFinSF1** | 1 | 0.011271 | 0.001498 | 7.52 |
| **CentralAir N** | 1 | -0.070558 | 0.017424 | -4.05 |
| **CentralAir Y** | 0 | 0 | . | . |
| **Log1stFlrSF** | 1 | 0.430404 | 0.019114 | 22.52 |
| **2ndFlrSF** | 1 | 0.000255 | 0.000011073 | 22.99 |
| **BsmtFullBath** | 1 | 0.031243 | 0.008552 | 3.65 |
| **KitchenAbvGr** | 1 | -0.097749 | 0.019993 | -4.89 |
| **KitchenQual Ex** | 1 | 0.090268 | 0.018258 | 4.94 |
| **KitchenQual Fa** | 1 | -0.007589 | 0.025224 | -0.30 |
| **KitchenQual Gd** | 1 | 0.007415 | 0.010725 | 0.69 |
| **KitchenQual TA** | 0 | 0 | . | . |
| **Functional Maj1** | 1 | -0.085603 | 0.036785 | -2.33 |
| **Functional Maj2** | 1 | -0.197489 | 0.057037 | -3.46 |
| **Functional Min1** | 1 | -0.071853 | 0.023495 | -3.06 |
| **Functional Min2** | 1 | -0.058743 | 0.023057 | -2.55 |
| **Functional Mod** | 1 | -0.080600 | 0.034424 | -2.34 |
| **Functional Typ** | 0 | 0 | . | . |
| **GarageCars** | 1 | 0.056376 | 0.007557 | 7.46 |
| **LogWoodDeckSF** | 1 | 0.004424 | 0.001459 | 3.03 |
| **LogOpenPorchSF** | 1 | 0.006393 | 0.001943 | 3.29 |
| **LogScreenPorch** | 1 | 0.008940 | 0.002405 | 3.72 |
| **SaleCondition Abnorml** | 1 | -0.115983 | 0.018737 | -6.19 |
| **SaleCondition AdjLand** | 1 | 0.165892 | 0.112528 | 1.47 |
| **SaleCondition Alloca** | 1 | -0.147734 | 0.041893 | -3.53 |
| **SaleCondition Family** | 1 | -0.111411 | 0.029362 | -3.79 |
| **SaleCondition Normal** | 1 | -0.052079 | 0.012849 | -4.05 |
| **SaleCondition Partial** | 0 | 0 | . | . |

**The GLMSELECT Procedure**

|  |  |
| --- | --- |
| **Data Set** | WORK.TRAIN\_DUMMY |
| **Dependent Variable** | LogSalePrice |
| **Selection Method** | Stepwise |
| **Select Criterion** | SBC |
| **Stop Criterion** | SBC |
| **Effect Hierarchy Enforced** | None |

|  |  |
| --- | --- |
| **Number of Observations Read** | 1460 |
| **Number of Observations Used** | 1121 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **MSZoning** | 5 | C FV RH RL RM |
| **Street** | 2 | Grvl Pave |
| **Alley** | 3 | Gr NA Pa |
| **LotShape** | 4 | IR1 IR2 IR3 Reg |
| **LandContour** | 4 | Bnk HLS Low Lvl |
| **Utilities** | 1 | AllPub |
| **LotConfig** | 5 | Corner CulDSac FR2 FR3 Inside |
| **LandSlope** | 3 | Gtl Mod Sev |
| **Neighborhood** | 25 | Blmngtn Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker |
| **Condition1** | 9 | Artery Feedr Norm PosA PosN RRAe RRAn RRNe RRNn |
| **Condition2** | 6 | Artery Feedr Norm PosA PosN RRNn |
| **BldgType** | 5 | 1Fam 2fmCon Duplex Twnhs TwnhsE |
| **HouseStyle** | 8 | 1.5Fin 1.5Unf 1Story 2.5Fin 2.5Unf 2Story SFoyer SLvl |
| **RoofStyle** | 5 | Flat Gable Gambr Hip Mansa |
| **RoofMatl** | 7 | ClyTile CompShg Membran Roll Tar&Grv WdShake WdShngl |
| **Exterior1st** | 14 | AsbShng BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing |
| **Exterior2nd** | 16 | AsbShng AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng |
| **MasVnrType** | 4 | BrkCmn BrkFace None Stone |
| **ExterQual** | 4 | Ex Fa Gd TA |
| **ExterCond** | 4 | Ex Fa Gd TA |
| **Foundation** | 6 | BrkTil CBlock PConc Slab Stone Wood |
| **BsmtQual** | 5 | Ex Fa Gd NA TA |
| **BsmtCond** | 5 | Fa Gd NA Po TA |
| **BsmtExposure** | 5 | Av Gd Mn NA No |
| **BsmtFinType1** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **BsmtFinType2** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **Heating** | 5 | GasA GasW Grav OthW Wall |
| **HeatingQC** | 5 | Ex Fa Gd Po TA |
| **CentralAir** | 2 | N Y |
| **Electrical** | 6 | FuseA FuseF FuseP Mix NA SBrkr |
| **KitchenQual** | 4 | Ex Fa Gd TA |
| **Functional** | 6 | Maj1 Maj2 Min1 Min2 Mod Typ |
| **FireplaceQu** | 6 | Ex Fa Gd NA Po TA |
| **GarageType** | 6 | 2Types Attchd Basment BuiltIn CarPort Detchd |
| **GarageFinish** | 3 | Fin RFn Unf |
| **GarageQual** | 5 | Ex Fa Gd Po TA |
| **GarageCond** | 5 | Ex Fa Gd Po TA |
| **PavedDrive** | 3 | N P Y |
| **PoolQC** | 4 | Ex Fa Gd NA |
| **Fence** | 5 | GdPrv GdWo MnPrv MnWw NA |
| **MiscFeature** | 4 | NA Othr Shed TenC |
| **SaleType** | 6 | COD CWD Con New Oth WD |
| **SaleCondition** | 6 | Abnorml AdjLand Alloca Family Normal Partial |

| **Dimensions** | |
| --- | --- |
| **Number of Effects** | 78 |
| **Number of Parameters** | 285 |

**The GLMSELECT Procedure**

| **Stepwise Selection Summary** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Step** | **Effect Entered** | **Effect Removed** | **Number Effects In** | **Number Parms In** | **Model R-Square** | **SBC** | **PRESS** |
| **\* Optimal Value of Criterion** | | | | | | | |
| **0** | **Intercept** |  | 1 | 1 | 0.0000 | -2066.4335 | 176.6399 |
| **1** | **OverallQual** |  | 2 | 2 | 0.6851 | -3354.5767 | 55.7616 |
| **2** | **Log1stFlrSF** |  | 3 | 3 | 0.7412 | -3567.6573 | 45.9654 |
| **3** | **2ndFlrSF** |  | 4 | 4 | 0.7818 | -3751.9351 | 38.9323 |
| **4** | **YearBuilt** |  | 5 | 5 | 0.8147 | -3928.0564 | 33.1553 |
| **5** | **OverallCond** |  | 6 | 6 | 0.8382 | -4073.1053 | 29.0219 |
| **6** | **LogLotArea** |  | 7 | 7 | 0.8496 | -4147.7963 | 27.0661 |
| **7** | **RoofMatl** |  | 8 | 13 | 0.8656 | -4232.0952 | 24.1556 |
| **8** | **LogBsmtFinSF1** |  | 9 | 14 | 0.8744 | -4300.5435 | 22.6489 |
| **9** | **Neighborhood** |  | 10 | 38 | 0.8985 | -4371.1382 | 19.2971 |
| **10** | **GarageCars** |  | 11 | 39 | 0.9018 | -4400.6906 | 18.7174 |
| **11** | **KitchenAbvGr** |  | 12 | 40 | 0.9055 | -4436.9303 | 18.0567 |
| **12** | **MSZoning** |  | 13 | 44 | 0.9101 | -4465.5009 | 17.4914 |
| **13** | **Condition2** |  | 14 | 49 | 0.9154 | -4498.0609 | 18.8515 |
| **14** | **KitchenQual** |  | 15 | 52 | 0.9182 | -4514.9270 | 18.5146 |
| **15** | **SaleCondition** |  | 16 | 57 | 0.9214 | -4523.8525 | 18.2602 |
| **16** | **BsmtFullBath** |  | 17 | 58 | 0.9226 | -4533.9616 | 18.0625 |
| **17** | **CentralAir** |  | 18 | 59 | 0.9237 | -4542.9937 | 17.8060 |
| **18** | **GarageFinish** |  | 19 | 61 | 0.9250 | -4548.0730 | 17.5431 |
| **19** | **LogScreenPorch** |  | 20 | 62 | 0.9257 | -4552.6909 | 17.3707 |
| **20** | **LogOpenPorchSF** |  | 21 | 63 | 0.9263 | -4554.1894 | 17.2921 |
| **21** | **LogWoodDeckSF** |  | 22 | 64 | 0.9269 | -4555.8269\* | 17.1852\* |

|  |
| --- |
| Selection stopped at a local minimum of the SBC criterion. |

| **Stop Details** | | | | |
| --- | --- | --- | --- | --- |
| **Candidate For** | **Effect** | **Candidate SBC** |  | **Compare SBC** |
| **Entry** | LogEnclosedPorch | -4554.8188 | > | -4555.8269 |
| **Removal** | GarageFinish | -4555.2554 | > | -4555.8269 |

**The GLMSELECT Procedure**

**Selected Model**

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

|  |  |
| --- | --- |
| **Effects:** | Intercept MSZoning LogLotArea Neighborhood Condition2 OverallQual OverallCond YearBuilt RoofMatl LogBsmtFinSF1 CentralAir Log1stFlrSF 2ndFlrSF BsmtFullBath KitchenAbvGr KitchenQual GarageFinish GarageCars LogWoodDeckSF LogOpenPorchSF LogScreenPorch SaleCondition |

| **Analysis of Variance** | | | | |
| --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** |
| **Model** | 63 | 163.42779 | 2.59409 | 212.60 |
| **Error** | 1057 | 12.89713 | 0.01220 |  |
| **Corrected Total** | 1120 | 176.32491 |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 0.11046 |
| **Dependent Mean** | 12.04841 |
| **R-Square** | 0.9269 |
| **Adj R-Sq** | 0.9225 |
| **AIC** | -3754.23342 |
| **AICC** | -3746.10071 |
| **PRESS** | 17.18516 |
| **SBC** | -4555.82692 |

| **Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** |
| **Intercept** | 1 | 2.433351 | 0.664000 | 3.66 |
| **MSZoning C** | 1 | -0.326476 | 0.045596 | -7.16 |
| **MSZoning FV** | 1 | 0.097795 | 0.034641 | 2.82 |
| **MSZoning RH** | 1 | 0.040363 | 0.040920 | 0.99 |
| **MSZoning RL** | 1 | 0.042573 | 0.018581 | 2.29 |
| **MSZoning RM** | 0 | 0 | . | . |
| **LogLotArea** | 1 | 0.095983 | 0.011631 | 8.25 |
| **Neighborhood Blmngtn** | 1 | -0.045403 | 0.054670 | -0.83 |
| **Neighborhood Blueste** | 1 | 0.016407 | 0.093330 | 0.18 |
| **Neighborhood BrDale** | 1 | 0.021599 | 0.058062 | 0.37 |
| **Neighborhood BrkSide** | 1 | 0.039986 | 0.050699 | 0.79 |
| **Neighborhood ClearCr** | 1 | -0.059126 | 0.054649 | -1.08 |
| **Neighborhood CollgCr** | 1 | -0.039486 | 0.044405 | -0.89 |
| **Neighborhood Crawfor** | 1 | 0.112287 | 0.047981 | 2.34 |
| **Neighborhood Edwards** | 1 | -0.123455 | 0.045871 | -2.69 |
| **Neighborhood Gilbert** | 1 | -0.045361 | 0.046535 | -0.97 |
| **Neighborhood IDOTRR** | 1 | 0.013611 | 0.055717 | 0.24 |
| **Neighborhood MeadowV** | 1 | -0.100128 | 0.060629 | -1.65 |
| **Neighborhood Mitchel** | 1 | -0.081181 | 0.047478 | -1.71 |
| **Neighborhood NAmes** | 1 | -0.077424 | 0.044332 | -1.75 |
| **Neighborhood NPkVill** | 1 | -0.026667 | 0.062613 | -0.43 |
| **Neighborhood NWAmes** | 1 | -0.106910 | 0.046026 | -2.32 |
| **Neighborhood NoRidge** | 1 | 0.019621 | 0.047888 | 0.41 |
| **Neighborhood NridgHt** | 1 | 0.062033 | 0.045641 | 1.36 |
| **Neighborhood OldTown** | 1 | -0.032773 | 0.050600 | -0.65 |
| **Neighborhood SWISU** | 1 | -0.022331 | 0.052776 | -0.42 |
| **Neighborhood Sawyer** | 1 | -0.094850 | 0.046484 | -2.04 |
| **Neighborhood SawyerW** | 1 | -0.068615 | 0.046367 | -1.48 |
| **Neighborhood Somerst** | 1 | -0.002462 | 0.050162 | -0.05 |
| **Neighborhood StoneBr** | 1 | 0.096736 | 0.050136 | 1.93 |
| **Neighborhood Timber** | 1 | -0.015773 | 0.047616 | -0.33 |
| **Neighborhood Veenker** | 0 | 0 | . | . |
| **Condition2 Artery** | 1 | -0.037195 | 0.115706 | -0.32 |
| **Condition2 Feedr** | 1 | -0.056835 | 0.094788 | -0.60 |
| **Condition2 Norm** | 1 | 0.068464 | 0.080871 | 0.85 |
| **Condition2 PosA** | 1 | 0.214739 | 0.139874 | 1.54 |
| **Condition2 PosN** | 1 | -0.634447 | 0.114831 | -5.53 |
| **Condition2 RRNn** | 0 | 0 | . | . |
| **OverallQual** | 1 | 0.058897 | 0.004808 | 12.25 |
| **OverallCond** | 1 | 0.047338 | 0.003995 | 11.85 |
| **YearBuilt** | 1 | 0.002621 | 0.000304 | 8.61 |
| **RoofMatl ClyTile** | 1 | -1.628480 | 0.124099 | -13.12 |
| **RoofMatl CompShg** | 1 | -0.132856 | 0.047591 | -2.79 |
| **RoofMatl Membran** | 1 | 0.022952 | 0.123788 | 0.19 |
| **RoofMatl Roll** | 1 | -0.130134 | 0.122264 | -1.06 |
| **RoofMatl Tar&Grv** | 1 | -0.202109 | 0.065408 | -3.09 |
| **RoofMatl WdShake** | 1 | -0.204848 | 0.092597 | -2.21 |
| **RoofMatl WdShngl** | 0 | 0 | . | . |
| **LogBsmtFinSF1** | 1 | 0.011119 | 0.001503 | 7.40 |
| **CentralAir N** | 1 | -0.073293 | 0.017521 | -4.18 |
| **CentralAir Y** | 0 | 0 | . | . |
| **Log1stFlrSF** | 1 | 0.404321 | 0.018699 | 21.62 |
| **2ndFlrSF** | 1 | 0.000246 | 0.000011022 | 22.29 |
| **BsmtFullBath** | 1 | 0.031820 | 0.008593 | 3.70 |
| **KitchenAbvGr** | 1 | -0.086124 | 0.020168 | -4.27 |
| **KitchenQual Ex** | 1 | 0.097355 | 0.018262 | 5.33 |
| **KitchenQual Fa** | 1 | -0.010269 | 0.025318 | -0.41 |
| **KitchenQual Gd** | 1 | 0.014719 | 0.010234 | 1.44 |
| **KitchenQual TA** | 0 | 0 | . | . |
| **GarageFinish Fin** | 1 | 0.042068 | 0.011788 | 3.57 |
| **GarageFinish RFn** | 1 | 0.030232 | 0.010220 | 2.96 |
| **GarageFinish Unf** | 0 | 0 | . | . |
| **GarageCars** | 1 | 0.059185 | 0.007604 | 7.78 |
| **LogWoodDeckSF** | 1 | 0.004205 | 0.001469 | 2.86 |
| **LogOpenPorchSF** | 1 | 0.005996 | 0.001976 | 3.03 |
| **LogScreenPorch** | 1 | 0.008624 | 0.002427 | 3.55 |
| **SaleCondition Abnorml** | 1 | -0.118718 | 0.018883 | -6.29 |
| **SaleCondition AdjLand** | 1 | 0.159529 | 0.113486 | 1.41 |
| **SaleCondition Alloca** | 1 | -0.133985 | 0.042298 | -3.17 |
| **SaleCondition Family** | 1 | -0.107906 | 0.029625 | -3.64 |
| **SaleCondition Normal** | 1 | -0.055147 | 0.012949 | -4.26 |
| **SaleCondition Partial** | 0 | 0 | . | . |

Table 2H:

**The GLMSELECT Procedure**

|  |  |
| --- | --- |
| **Data Set** | WORK.TRAIN\_DUMMY |
| **Dependent Variable** | LogSalePrice |
| **Selection Method** | Forward |
| **Select Criterion** | Significance Level |
| **Stop Criterion** | Significance Level |
| **Entry Significance Level (SLE)** | 0.5 |
| **Effect Hierarchy Enforced** | None |

|  |  |
| --- | --- |
| **Number of Observations Read** | 1460 |
| **Number of Observations Used** | 1127 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **MSZoning** | 5 | C FV RH RL RM |
| **Street** | 2 | Grvl Pave |
| **Alley** | 3 | Gr NA Pa |
| **LotShape** | 4 | IR1 IR2 IR3 Reg |
| **Utilities** | 1 | AllPub |
| **LotConfig** | 5 | Corner CulDSac FR2 FR3 Inside |
| **LandSlope** | 3 | Gtl Mod Sev |
| **Neighborhood** | 25 | Blmngtn Blueste BrDale BrkSide ClearCr CollgCr Crawfor Edwards Gilbert IDOTRR MeadowV Mitchel NAmes NPkVill NWAmes NoRidge NridgHt OldTown SWISU Sawyer SawyerW Somerst StoneBr Timber Veenker |
| **Condition1** | 9 | Artery Feedr Norm PosA PosN RRAe RRAn RRNe RRNn |
| **Condition2** | 6 | Artery Feedr Norm PosA PosN RRNn |
| **BldgType** | 5 | 1Fam 2fmCon Duplex Twnhs TwnhsE |
| **RoofMatl** | 7 | ClyTile CompShg Membran Roll Tar&Grv WdShake WdShngl |
| **Exterior1st** | 14 | AsbShng BrkComm BrkFace CBlock CemntBd HdBoard ImStucc MetalSd Plywood Stone Stucco VinylSd Wd Sdng WdShing |
| **Exterior2nd** | 16 | AsbShng AsphShn Brk Cmn BrkFace CBlock CmentBd HdBoard ImStucc MetalSd Other Plywood Stone Stucco VinylSd Wd Sdng Wd Shng |
| **MasVnrType** | 5 | BrkCmn BrkFace NA None Stone |
| **ExterCond** | 4 | Ex Fa Gd TA |
| **Foundation** | 6 | BrkTil CBlock PConc Slab Stone Wood |
| **BsmtQual** | 5 | Ex Fa Gd NA TA |
| **BsmtCond** | 5 | Fa Gd NA Po TA |
| **BsmtExposure** | 5 | Av Gd Mn NA No |
| **BsmtFinType1** | 7 | ALQ BLQ GLQ LwQ NA Rec Unf |
| **Heating** | 5 | GasA GasW Grav OthW Wall |
| **HeatingQC** | 5 | Ex Fa Gd Po TA |
| **CentralAir** | 2 | N Y |
| **KitchenQual** | 4 | Ex Fa Gd TA |
| **Functional** | 6 | Maj1 Maj2 Min1 Min2 Mod Typ |
| **GarageType** | 6 | 2Types Attchd Basment BuiltIn CarPort Detchd |
| **GarageFinish** | 3 | Fin RFn Unf |
| **GarageQual** | 5 | Ex Fa Gd Po TA |
| **GarageCond** | 5 | Ex Fa Gd Po TA |
| **PavedDrive** | 3 | N P Y |
| **PoolQC** | 4 | Ex Fa Gd NA |
| **Fence** | 5 | GdPrv GdWo MnPrv MnWw NA |
| **SaleType** | 6 | COD CWD Con New Oth WD |
| **SaleCondition** | 6 | Abnorml AdjLand Alloca Family Normal Partial |

| **Dimensions** | |
| --- | --- |
| **Number of Effects** | 61 |
| **Number of Parameters** | 229 |

**The GLMSELECT Procedure**

| **Forward Selection Summary** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Step** | **Effect Entered** | **Number Effects In** | **Number Parms In** | **Model R-Square** | **PRESS** | **F Value** | **Pr > F** |
| **\* Optimal Value of Criterion** | | | | | | | |
| **0** | **Intercept** | 1 | 1 | 0.0000 | 177.8173 | 0.00 | 1.0000 |
| **1** | **OverallQual** | 2 | 2 | 0.6869 | 55.8077 | 2467.85 | <.0001 |
| **2** | **Neighborhood** | 3 | 26 | 0.7615 | 44.4759 | 14.34 | <.0001 |
| **3** | **TotRmsAbvGrd** | 4 | 27 | 0.7979 | 37.7749 | 198.01 | <.0001 |
| **4** | **Log1stFlrSF** | 5 | 28 | 0.8203 | 33.6662 | 137.54 | <.0001 |
| **5** | **2ndFlrSF** | 6 | 29 | 0.8403 | 30.0577 | 136.89 | <.0001 |
| **6** | **OverallCond** | 7 | 30 | 0.8570 | 27.0198 | 128.39 | <.0001 |
| **7** | **YearBuilt** | 8 | 31 | 0.8681 | 25.0848 | 91.88 | <.0001 |
| **8** | **LogBsmtFinSF1** | 9 | 32 | 0.8777 | 23.3271 | 86.23 | <.0001 |
| **9** | **RoofMatl** | 10 | 38 | 0.8889 | 21.2381 | 18.25 | <.0001 |
| **10** | **LogLotArea** | 11 | 39 | 0.8987 | 19.4361 | 106.16 | <.0001 |
| **11** | **Condition2** | 12 | 44 | 0.9040 | 20.8004 | 11.89 | <.0001 |
| **12** | **BsmtQual** | 13 | 48 | 0.9084 | 20.2373 | 12.96 | <.0001 |
| **13** | **MSZoning** | 14 | 52 | 0.9131 | 19.6959 | 14.35 | <.0001 |
| **14** | **GarageCars** | 15 | 53 | 0.9167 | 18.9419 | 46.87 | <.0001 |
| **15** | **KitchenAbvGr** | 16 | 54 | 0.9190 | 18.5103 | 30.98 | <.0001 |
| **16** | **SaleCondition** | 17 | 59 | 0.9216 | 18.3584 | 6.96 | <.0001 |
| **17** | **KitchenQual** | 18 | 62 | 0.9235 | 18.1484 | 8.98 | <.0001 |
| **18** | **Functional** | 19 | 67 | 0.9256 | 18.0153 | 6.05 | <.0001 |
| **19** | **CentralAir** | 20 | 68 | 0.9268 | 17.7483 | 17.09 | <.0001 |
| **20** | **Foundation** | 21 | 73 | 0.9285 | 17.7240 | 4.93 | 0.0002 |
| **21** | **BsmtFullBath** | 22 | 74 | 0.9294 | 17.5575 | 13.30 | 0.0003 |
| **22** | **Condition1** | 23 | 82 | 0.9314 | 17.3845 | 3.74 | 0.0003 |
| **23** | **Fireplaces** | 24 | 83 | 0.9321 | 17.1701 | 11.60 | 0.0007 |
| **24** | **BsmtExposure** | 25 | 87 | 0.9333 | 17.2649 | 4.52 | 0.0013 |
| **25** | **HeatingQC** | 26 | 91 | 0.9343 | 17.0792 | 4.28 | 0.0019 |
| **26** | **LogScreenPorch** | 27 | 92 | 0.9351 | 16.8892 | 11.49 | 0.0007 |
| **27** | **Heating** | 28 | 96 | 0.9362 | 16.6233 | 4.64 | 0.0010 |
| **28** | **LogGarageArea** | 29 | 97 | 0.9366 | 16.5165\* | 6.58 | 0.0105 |
| **29** | **SaleType** | 30 | 102 | 0.9375 | 16.7068 | 2.90 | 0.0131 |
| **30** | **Street** | 31 | 103 | 0.9379 | 16.6284 | 6.72 | 0.0097 |
| **31** | **Exterior1st** | 32 | 116 | 0.9395 | 17.4528 | 1.98 | 0.0191 |
| **32** | **YearRemodAdd** | 33 | 117 | 0.9399 | 17.4468 | 6.74 | 0.0096 |
| **33** | **LandSlope** | 34 | 119 | 0.9403 | 17.4321 | 3.88 | 0.0209 |
| **34** | **Log3SsnPorch** | 35 | 120 | 0.9406 | 17.3641 | 4.14 | 0.0421 |
| **35** | **LogWoodDeckSF** | 36 | 121 | 0.9408 | 17.3009 | 4.33 | 0.0377 |
| **36** | **LogOpenPorchSF** | 37 | 122 | 0.9411 | 17.2889 | 4.30 | 0.0384 |
| **37** | **LogEnclosedPorch** | 38 | 123 | 0.9413 | 17.2496 | 3.59 | 0.0583 |
| **38** | **LogPoolArea** | 39 | 124 | 0.9415 | 17.2134 | 3.49 | 0.0619 |
| **39** | **LotConfig** | 40 | 128 | 0.9420 | 17.1078 | 2.30 | 0.0575 |
| **40** | **Exterior2nd** | 41 | 142 | 0.9434 | 17.2497 | 1.73 | 0.0448 |
| **41** | **HalfBath** | 42 | 143 | 0.9436 | 17.2320 | 2.51 | 0.1135 |
| **42** | **FullBath** | 43 | 144 | 0.9438 | 17.2046 | 5.13 | 0.0237 |
| **43** | **MSSubClass** | 44 | 145 | 0.9440 | 17.1757 | 2.73 | 0.0985 |
| **44** | **PoolQC** | 45 | 148 | 0.9443 | 17.1313 | 1.78 | 0.1491 |
| **45** | **ExterCond** | 46 | 151 | 0.9446 | 17.1112 | 1.59 | 0.1904 |
| **46** | **BsmtFinType1** | 47 | 156 | 0.9450 | 17.2254 | 1.49 | 0.1908 |
| **47** | **BsmtFinSF2** | 48 | 157 | 0.9452 | 17.1923 | 4.39 | 0.0363 |
| **48** | **GarageQual** | 49 | 161 | 0.9456 | 17.2452 | 1.54 | 0.1893 |
| **49** | **GarageCond** | 50 | 165 | 0.9460 | 17.2800 | 2.02 | 0.0900 |
| **50** | **BsmtUnfSF** | 51 | 166 | 0.9462 | 17.3101 | 2.32 | 0.1277 |
| **51** | **IntLotFrontage** | 52 | 167 | 0.9463 | 17.3218 | 1.54 | 0.2145 |
| **52** | **Fence** | 53 | 171 | 0.9466 | 17.3172 | 1.34 | 0.2549 |
| **53** | **GarageFinish** | 54 | 173 | 0.9467 | 17.3222 | 1.29 | 0.2770 |
| **54** | **MasVnrType** | 55 | 177 | 0.9470 | 17.5583 | 1.17 | 0.3235 |
| **55** | **Alley** | 56 | 179 | 0.9471 | 17.6299 | 1.18 | 0.3086 |
| **56** | **PavedDrive** | 57 | 181 | 0.9472 | 17.6494 | 1.06 | 0.3464 |
| **57** | **GarageType** | 58 | 186 | 0.9475 | 17.7072 | 1.01 | 0.4097 |
| **58** | **LotShape** | 59 | 189 | 0.9477 | 17.6897 | 1.20 | 0.3073 |
| **59** | **BldgType** | 60 | 193 | 0.9479 | 17.7205 | 0.91 | 0.4602 |

|  |
| --- |
| Selection stopped because all candidate effects for entry are linearly dependent on effects in the model. |

**The GLMSELECT Procedure**

**Selected Model**

**The selected model is the model at the last step (Step 59).**

|  |  |
| --- | --- |
| **Effects:** | Intercept IntLotFrontage MSSubClass MSZoning LogLotArea Street Alley LotShape LotConfig LandSlope Neighborhood Condition1 Condition2 BldgType OverallQual OverallCond YearBuilt YearRemodAdd RoofMatl Exterior1st Exterior2nd MasVnrType ExterCond Foundation BsmtQual BsmtExposure BsmtFinType1 LogBsmtFinSF1 BsmtFinSF2 BsmtUnfSF Heating HeatingQC CentralAir Log1stFlrSF 2ndFlrSF BsmtFullBath FullBath HalfBath KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces GarageType GarageFinish GarageCars LogGarageArea GarageQual GarageCond PavedDrive LogWoodDeckSF LogOpenPorchSF LogEnclosedPorch Log3SsnPorch LogScreenPorch LogPoolArea PoolQC Fence SaleType SaleCondition |

| **Analysis of Variance** | | | | |
| --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** |
| **Model** | 192 | 168.25452 | 0.87633 | 88.51 |
| **Error** | 934 | 9.24736 | 0.00990 |  |
| **Corrected Total** | 1126 | 177.50188 |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 0.09950 |
| **Dependent Mean** | 12.05002 |
| **R-Square** | 0.9479 |
| **Adj R-Sq** | 0.9372 |
| **AIC** | -3897.95412 |
| **AICC** | -3816.77386 |
| **PRESS** | 17.72055 |
| **SBC** | -4056.68242 |

| **Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** |
| **Intercept** | 1 | 2.763297 | 0.964356 | 2.87 |
| **IntLotFrontage** | 1 | -0.000254 | 0.000174 | -1.46 |
| **MSSubClass** | 1 | -0.000169 | 0.000256 | -0.66 |
| **MSZoning C** | 1 | -0.395397 | 0.046848 | -8.44 |
| **MSZoning FV** | 1 | 0.085973 | 0.035835 | 2.40 |
| **MSZoning RH** | 1 | 0.078871 | 0.040631 | 1.94 |
| **MSZoning RL** | 1 | 0.040038 | 0.019618 | 2.04 |
| **MSZoning RM** | 0 | 0 | . | . |
| **LogLotArea** | 1 | 0.081825 | 0.014193 | 5.77 |
| **Street Grvl** | 1 | -0.133936 | 0.065710 | -2.04 |
| **Street Pave** | 0 | 0 | . | . |
| **Alley Gr** | 1 | -0.037105 | 0.029564 | -1.26 |
| **Alley NA** | 1 | -0.024209 | 0.023013 | -1.05 |
| **Alley Pa** | 0 | 0 | . | . |
| **LotShape IR1** | 1 | -0.007949 | 0.008426 | -0.94 |
| **LotShape IR2** | 1 | -0.001410 | 0.024213 | -0.06 |
| **LotShape IR3** | 1 | 0.073942 | 0.045983 | 1.61 |
| **LotShape Reg** | 0 | 0 | . | . |
| **LotConfig Corner** | 1 | 0.012825 | 0.008983 | 1.43 |
| **LotConfig CulDSac** | 1 | 0.042849 | 0.018303 | 2.34 |
| **LotConfig FR2** | 1 | -0.036476 | 0.020349 | -1.79 |
| **LotConfig FR3** | 1 | -0.063863 | 0.055380 | -1.15 |
| **LotConfig Inside** | 0 | 0 | . | . |
| **LandSlope Gtl** | 1 | 0.181234 | 0.059100 | 3.07 |
| **LandSlope Mod** | 1 | 0.197221 | 0.060930 | 3.24 |
| **LandSlope Sev** | 0 | 0 | . | . |
| **Neighborhood Blmngtn** | 1 | -0.032314 | 0.054986 | -0.59 |
| **Neighborhood Blueste** | 1 | 0.056905 | 0.092439 | 0.62 |
| **Neighborhood BrDale** | 1 | 0.033816 | 0.060887 | 0.56 |
| **Neighborhood BrkSide** | 1 | 0.019520 | 0.052030 | 0.38 |
| **Neighborhood ClearCr** | 1 | -0.037675 | 0.054796 | -0.69 |
| **Neighborhood CollgCr** | 1 | -0.069243 | 0.045060 | -1.54 |
| **Neighborhood Crawfor** | 1 | 0.094387 | 0.047995 | 1.97 |
| **Neighborhood Edwards** | 1 | -0.126565 | 0.046264 | -2.74 |
| **Neighborhood Gilbert** | 1 | -0.075332 | 0.047245 | -1.59 |
| **Neighborhood IDOTRR** | 1 | 0.015653 | 0.057498 | 0.27 |
| **Neighborhood MeadowV** | 1 | -0.116830 | 0.063516 | -1.84 |
| **Neighborhood Mitchel** | 1 | -0.080864 | 0.047940 | -1.69 |
| **Neighborhood NAmes** | 1 | -0.079463 | 0.044558 | -1.78 |
| **Neighborhood NPkVill** | 1 | -0.084092 | 0.086785 | -0.97 |
| **Neighborhood NWAmes** | 1 | -0.095962 | 0.046634 | -2.06 |
| **Neighborhood NoRidge** | 1 | 0.001377 | 0.047833 | 0.03 |
| **Neighborhood NridgHt** | 1 | 0.039540 | 0.046373 | 0.85 |
| **Neighborhood OldTown** | 1 | -0.035397 | 0.051846 | -0.68 |
| **Neighborhood SWISU** | 1 | -0.034323 | 0.053167 | -0.65 |
| **Neighborhood Sawyer** | 1 | -0.068996 | 0.046730 | -1.48 |
| **Neighborhood SawyerW** | 1 | -0.054541 | 0.046374 | -1.18 |
| **Neighborhood Somerst** | 1 | -0.023246 | 0.050534 | -0.46 |
| **Neighborhood StoneBr** | 1 | 0.103256 | 0.050631 | 2.04 |
| **Neighborhood Timber** | 1 | -0.053328 | 0.047521 | -1.12 |
| **Neighborhood Veenker** | 0 | 0 | . | . |
| **Condition1 Artery** | 1 | -0.120740 | 0.066718 | -1.81 |
| **Condition1 Feedr** | 1 | -0.075760 | 0.065737 | -1.15 |
| **Condition1 Norm** | 1 | -0.033825 | 0.063713 | -0.53 |
| **Condition1 PosA** | 1 | -0.105884 | 0.089507 | -1.18 |
| **Condition1 PosN** | 1 | -0.014134 | 0.076624 | -0.18 |
| **Condition1 RRAe** | 1 | -0.152492 | 0.075990 | -2.01 |
| **Condition1 RRAn** | 1 | -0.048827 | 0.065061 | -0.75 |
| **Condition1 RRNe** | 1 | -0.034937 | 0.120530 | -0.29 |
| **Condition1 RRNn** | 0 | 0 | . | . |
| **Condition2 Artery** | 1 | -0.029936 | 0.122425 | -0.24 |
| **Condition2 Feedr** | 1 | -0.042525 | 0.096765 | -0.44 |
| **Condition2 Norm** | 1 | 0.028020 | 0.078684 | 0.36 |
| **Condition2 PosA** | 1 | 0.248492 | 0.173605 | 1.43 |
| **Condition2 PosN** | 1 | -0.689070 | 0.117844 | -5.85 |
| **Condition2 RRNn** | 0 | 0 | . | . |
| **BldgType 1Fam** | 1 | 0.020815 | 0.031604 | 0.66 |
| **BldgType 2fmCon** | 1 | 0.027490 | 0.039007 | 0.70 |
| **BldgType Duplex** | 1 | -0.022346 | 0.042293 | -0.53 |
| **BldgType Twnhs** | 1 | -0.026069 | 0.025039 | -1.04 |
| **BldgType TwnhsE** | 0 | 0 | . | . |
| **OverallQual** | 1 | 0.043981 | 0.005024 | 8.75 |
| **OverallCond** | 1 | 0.040080 | 0.004547 | 8.81 |
| **YearBuilt** | 1 | 0.001906 | 0.000393 | 4.85 |
| **YearRemodAdd** | 1 | 0.000527 | 0.000276 | 1.91 |
| **RoofMatl ClyTile** | 1 | -1.764297 | 0.204234 | -8.64 |
| **RoofMatl CompShg** | 1 | -0.103710 | 0.050976 | -2.03 |
| **RoofMatl Membran** | 1 | 0.174668 | 0.136689 | 1.28 |
| **RoofMatl Roll** | 1 | 0.003212 | 0.125266 | 0.03 |
| **RoofMatl Tar&Grv** | 1 | -0.130411 | 0.070026 | -1.86 |
| **RoofMatl WdShake** | 1 | -0.155746 | 0.097026 | -1.61 |
| **RoofMatl WdShngl** | 0 | 0 | . | . |
| **Exterior1st AsbShng** | 1 | 0.010579 | 0.065188 | 0.16 |
| **Exterior1st BrkComm** | 1 | -0.389853 | 0.129444 | -3.01 |
| **Exterior1st BrkFace** | 1 | 0.039334 | 0.039900 | 0.99 |
| **Exterior1st CBlock** | 1 | -0.074869 | 0.110412 | -0.68 |
| **Exterior1st CemntBd** | 1 | -0.106103 | 0.086437 | -1.23 |
| **Exterior1st HdBoard** | 1 | -0.063039 | 0.036339 | -1.73 |
| **Exterior1st ImStucc** | 1 | -0.030076 | 0.116849 | -0.26 |
| **Exterior1st MetalSd** | 1 | 0.054808 | 0.052772 | 1.04 |
| **Exterior1st Plywood** | 1 | -0.068710 | 0.037242 | -1.84 |
| **Exterior1st Stone** | 1 | -0.068206 | 0.132437 | -0.52 |
| **Exterior1st Stucco** | 1 | 0.033188 | 0.049909 | 0.66 |
| **Exterior1st VinylSd** | 1 | 0.011817 | 0.044484 | 0.27 |
| **Exterior1st Wd Sdng** | 1 | -0.060473 | 0.034899 | -1.73 |
| **Exterior1st WdShing** | 0 | 0 | . | . |
| **Exterior2nd AsbShng** | 1 | -0.064467 | 0.060236 | -1.07 |
| **Exterior2nd AsphShn** | 1 | 0.112855 | 0.083119 | 1.36 |
| **Exterior2nd Brk Cmn** | 1 | 0.164038 | 0.095040 | 1.73 |
| **Exterior2nd BrkFace** | 1 | -0.016666 | 0.043553 | -0.38 |
| **Exterior2nd CBlock** | 0 | 0 | . | . |
| **Exterior2nd CmentBd** | 1 | 0.116620 | 0.083382 | 1.40 |
| **Exterior2nd HdBoard** | 1 | 0.045503 | 0.036118 | 1.26 |
| **Exterior2nd ImStucc** | 1 | 0.055730 | 0.049294 | 1.13 |
| **Exterior2nd MetalSd** | 1 | -0.038608 | 0.051790 | -0.75 |
| **Exterior2nd Other** | 1 | -0.120883 | 0.114013 | -1.06 |
| **Exterior2nd Plywood** | 1 | 0.041827 | 0.034495 | 1.21 |
| **Exterior2nd Stone** | 1 | 0.032556 | 0.073233 | 0.44 |
| **Exterior2nd Stucco** | 1 | -0.016035 | 0.047618 | -0.34 |
| **Exterior2nd VinylSd** | 1 | -0.000758 | 0.041252 | -0.02 |
| **Exterior2nd Wd Sdng** | 1 | 0.061000 | 0.031350 | 1.95 |
| **Exterior2nd Wd Shng** | 0 | 0 | . | . |
| **MasVnrType BrkCmn** | 1 | -0.075196 | 0.038738 | -1.94 |
| **MasVnrType BrkFace** | 1 | -0.015516 | 0.013378 | -1.16 |
| **MasVnrType NA** | 1 | -0.060777 | 0.043486 | -1.40 |
| **MasVnrType None** | 1 | -0.019547 | 0.014088 | -1.39 |
| **MasVnrType Stone** | 0 | 0 | . | . |
| **ExterCond Ex** | 1 | 0.019113 | 0.105670 | 0.18 |
| **ExterCond Fa** | 1 | -0.036632 | 0.029978 | -1.22 |
| **ExterCond Gd** | 1 | -0.026043 | 0.012099 | -2.15 |
| **ExterCond TA** | 0 | 0 | . | . |
| **Foundation BrkTil** | 1 | 0.218721 | 0.078846 | 2.77 |
| **Foundation CBlock** | 1 | 0.258104 | 0.077644 | 3.32 |
| **Foundation PConc** | 1 | 0.258482 | 0.076993 | 3.36 |
| **Foundation Slab** | 1 | 0.315905 | 0.093449 | 3.38 |
| **Foundation Stone** | 1 | 0.330759 | 0.090910 | 3.64 |
| **Foundation Wood** | 0 | 0 | . | . |
| **BsmtQual Ex** | 1 | 0.051969 | 0.019407 | 2.68 |
| **BsmtQual Fa** | 1 | -0.004436 | 0.022341 | -0.20 |
| **BsmtQual Gd** | 1 | 0.016774 | 0.012922 | 1.30 |
| **BsmtQual NA** | 1 | -0.020214 | 0.109293 | -0.18 |
| **BsmtQual TA** | 0 | 0 | . | . |
| **BsmtExposure Av** | 1 | 0.010286 | 0.010591 | 0.97 |
| **BsmtExposure Gd** | 1 | 0.066815 | 0.014509 | 4.60 |
| **BsmtExposure Mn** | 1 | 0.003236 | 0.012472 | 0.26 |
| **BsmtExposure NA** | 1 | -0.053228 | 0.100969 | -0.53 |
| **BsmtExposure No** | 0 | 0 | . | . |
| **BsmtFinType1 ALQ** | 1 | -0.108850 | 0.039050 | -2.79 |
| **BsmtFinType1 BLQ** | 1 | -0.105361 | 0.039628 | -2.66 |
| **BsmtFinType1 GLQ** | 1 | -0.086128 | 0.036711 | -2.35 |
| **BsmtFinType1 LwQ** | 1 | -0.126670 | 0.040834 | -3.10 |
| **BsmtFinType1 NA** | 0 | 0 | . | . |
| **BsmtFinType1 Rec** | 1 | -0.119039 | 0.040151 | -2.96 |
| **BsmtFinType1 Unf** | 0 | 0 | . | . |
| **LogBsmtFinSF1** | 1 | 0.028180 | 0.007314 | 3.85 |
| **BsmtFinSF2** | 1 | 0.000063953 | 0.000027619 | 2.32 |
| **BsmtUnfSF** | 1 | 0.000024454 | 0.000019112 | 1.28 |
| **Heating GasA** | 1 | -0.034565 | 0.076454 | -0.45 |
| **Heating GasW** | 1 | 0.084826 | 0.079303 | 1.07 |
| **Heating Grav** | 1 | -0.083981 | 0.112233 | -0.75 |
| **Heating OthW** | 1 | 0.006871 | 0.135507 | 0.05 |
| **Heating Wall** | 0 | 0 | . | . |
| **HeatingQC Ex** | 1 | 0.038287 | 0.010715 | 3.57 |
| **HeatingQC Fa** | 1 | 0.027651 | 0.023930 | 1.16 |
| **HeatingQC Gd** | 1 | 0.014087 | 0.011048 | 1.28 |
| **HeatingQC Po** | 1 | -0.078457 | 0.114728 | -0.68 |
| **HeatingQC TA** | 0 | 0 | . | . |
| **CentralAir N** | 1 | -0.088836 | 0.019646 | -4.52 |
| **CentralAir Y** | 0 | 0 | . | . |
| **Log1stFlrSF** | 1 | 0.346367 | 0.027336 | 12.67 |
| **2ndFlrSF** | 1 | 0.000209 | 0.000020450 | 10.22 |
| **BsmtFullBath** | 1 | 0.019871 | 0.009395 | 2.12 |
| **FullBath** | 1 | 0.022106 | 0.011209 | 1.97 |
| **HalfBath** | 1 | 0.026308 | 0.010443 | 2.52 |
| **KitchenAbvGr** | 1 | -0.055568 | 0.032921 | -1.69 |
| **KitchenQual Ex** | 1 | 0.071307 | 0.018367 | 3.88 |
| **KitchenQual Fa** | 1 | 0.000653 | 0.025117 | 0.03 |
| **KitchenQual Gd** | 1 | -0.003759 | 0.010697 | -0.35 |
| **KitchenQual TA** | 0 | 0 | . | . |
| **TotRmsAbvGrd** | 1 | 0.005667 | 0.004358 | 1.30 |
| **Functional Maj1** | 1 | -0.063272 | 0.036501 | -1.73 |
| **Functional Maj2** | 1 | -0.245460 | 0.059838 | -4.10 |
| **Functional Min1** | 1 | -0.055121 | 0.023432 | -2.35 |
| **Functional Min2** | 1 | -0.066112 | 0.022878 | -2.89 |
| **Functional Mod** | 1 | -0.074363 | 0.036531 | -2.04 |
| **Functional Typ** | 0 | 0 | . | . |
| **Fireplaces** | 1 | 0.017657 | 0.006812 | 2.59 |
| **GarageType 2Types** | 1 | -0.075196 | 0.054452 | -1.38 |
| **GarageType Attchd** | 1 | -0.010778 | 0.011429 | -0.94 |
| **GarageType Basment** | 1 | -0.028586 | 0.032345 | -0.88 |
| **GarageType BuiltIn** | 1 | -0.036827 | 0.019536 | -1.89 |
| **GarageType CarPort** | 1 | -0.019726 | 0.043611 | -0.45 |
| **GarageType Detchd** | 0 | 0 | . | . |
| **GarageFinish Fin** | 1 | 0.017795 | 0.012273 | 1.45 |
| **GarageFinish RFn** | 1 | 0.018071 | 0.010839 | 1.67 |
| **GarageFinish Unf** | 0 | 0 | . | . |
| **GarageCars** | 1 | 0.035750 | 0.011525 | 3.10 |
| **LogGarageArea** | 1 | 0.032566 | 0.018670 | 1.74 |
| **GarageQual Ex** | 1 | 0.317891 | 0.122978 | 2.58 |
| **GarageQual Fa** | 1 | -0.033383 | 0.021745 | -1.54 |
| **GarageQual Gd** | 1 | 0.023360 | 0.036856 | 0.63 |
| **GarageQual Po** | 1 | -0.107260 | 0.100091 | -1.07 |
| **GarageQual TA** | 0 | 0 | . | . |
| **GarageCond Ex** | 1 | -0.321564 | 0.143180 | -2.25 |
| **GarageCond Fa** | 1 | -0.021429 | 0.025405 | -0.84 |
| **GarageCond Gd** | 1 | 0.026731 | 0.050697 | 0.53 |
| **GarageCond Po** | 1 | 0.107092 | 0.069654 | 1.54 |
| **GarageCond TA** | 0 | 0 | . | . |
| **PavedDrive N** | 1 | -0.011264 | 0.019058 | -0.59 |
| **PavedDrive P** | 1 | -0.037773 | 0.025582 | -1.48 |
| **PavedDrive Y** | 0 | 0 | . | . |
| **LogWoodDeckSF** | 1 | 0.003105 | 0.001455 | 2.13 |
| **LogOpenPorchSF** | 1 | 0.003180 | 0.001924 | 1.65 |
| **LogEnclosedPorch** | 1 | 0.004931 | 0.002279 | 2.16 |
| **Log3SsnPorch** | 1 | 0.011069 | 0.004903 | 2.26 |
| **LogScreenPorch** | 1 | 0.008839 | 0.002403 | 3.68 |
| **LogPoolArea** | 1 | 0.839796 | 0.644726 | 1.30 |
| **PoolQC Ex** | 1 | -5.107219 | 4.050150 | -1.26 |
| **PoolQC Fa** | 1 | -5.357974 | 4.104926 | -1.31 |
| **PoolQC Gd** | 1 | -5.153240 | 4.099669 | -1.26 |
| **PoolQC NA** | 0 | 0 | . | . |
| **Fence GdPrv** | 1 | -0.022975 | 0.018218 | -1.26 |
| **Fence GdWo** | 1 | -0.030021 | 0.018164 | -1.65 |
| **Fence MnPrv** | 1 | -0.003539 | 0.011608 | -0.30 |
| **Fence MnWw** | 1 | -0.031698 | 0.035007 | -0.91 |
| **Fence NA** | 0 | 0 | . | . |
| **SaleType COD** | 1 | 0.021975 | 0.020988 | 1.05 |
| **SaleType CWD** | 1 | 0.067437 | 0.053746 | 1.25 |
| **SaleType Con** | 1 | 0.080819 | 0.029332 | 2.76 |
| **SaleType New** | 1 | 0.117770 | 0.066694 | 1.77 |
| **SaleType Oth** | 1 | 0.163032 | 0.102664 | 1.59 |
| **SaleType WD** | 0 | 0 | . | . |
| **SaleCondition Abnorml** | 1 | 0.010226 | 0.067035 | 0.15 |
| **SaleCondition AdjLand** | 1 | 0.313303 | 0.126678 | 2.47 |
| **SaleCondition Alloca** | 1 | 0.038659 | 0.078327 | 0.49 |
| **SaleCondition Family** | 1 | 0.030843 | 0.070340 | 0.44 |
| **SaleCondition Normal** | 1 | 0.072764 | 0.065950 | 1.10 |
| **SaleCondition Partial** | 0 | 0 | . | . |

**Question 2 SAS Code:**

\*import training data;

proc import datafile="/home/bcoari0/sasuser.v94/train.csv"

out= train dbms=csv replace;

getnames=yes;

run;

\*Creating Dummy Variables and performing log transforms, adding 1 to account for 0 values if necessary;

Data train\_dummy;

set train;

LogSalePrice = Log(SalePrice);

IntLotFrontage = int(LotFrontage);

LogLotArea = Log(LotArea);

LogMasVnrArea = Log(MasVnrArea+1);

LogBsmtFinSF1 = Log(BsmtFinSF1+1);

Log1stFlrSF = Log('1stFlrSF'n);

LogPoolArea = Log(PoolArea+1);

LogMiscVal = Log(MiscVal+1);

LogGarageArea = Log(GarageArea);

LogWoodDeckSF = Log(WoodDeckSF+1);

LogOpenPorchSF = Log(OpenPorchSF+1);

LogEnclosedPorch = Log(EnclosedPorch+1);

Log3SsnPorch = Log('3SsnPorch'n+1);

LogScreenPorch = Log(ScreenPorch+1);

Alley\_dum0 = 0;

Alley\_dum1 = 0;

LotShape\_dum0 = 0;

LotShape\_dum1 = 0;

LotShape\_dum2 = 0;

LandContour\_dum0 = 0;

LandContour\_dum1 = 0;

LandContour\_dum2 = 0;

Utilities\_dum0 = 0;

LotConfig\_dum0 = 0;

LotConfig\_dum1 = 0;

LotConfig\_dum2 = 0;

LotConfig\_dum3 = 0;

LandSlope\_dum0 = 0;

LandSlope\_dum1 = 0;

Neighborhood\_dum0 = 0;

Neighborhood\_dum1 = 0;

Neighborhood\_dum2 = 0;

Neighborhood\_dum3 = 0;

Neighborhood\_dum4 = 0;

Neighborhood\_dum5 = 0;

Neighborhood\_dum6 = 0;

Neighborhood\_dum7 = 0;

Neighborhood\_dum8 = 0;

Neighborhood\_dum9 = 0;

Neighborhood\_dum10 = 0;

Neighborhood\_dum11 = 0;

Neighborhood\_dum12 = 0;

Neighborhood\_dum13 = 0;

Neighborhood\_dum14 = 0;

Neighborhood\_dum15 = 0;

Neighborhood\_dum16 = 0;

Neighborhood\_dum17 = 0;

Neighborhood\_dum18 = 0;

Neighborhood\_dum19 = 0;

Neighborhood\_dum20 = 0;

Neighborhood\_dum21 = 0;

Neighborhood\_dum22 = 0;

Neighborhood\_dum23 = 0;

Condition1\_dum0 = 0;

Condition1\_dum1 = 0;

Condition1\_dum2 = 0;

Condition1\_dum3 = 0;

Condition1\_dum4 = 0;

Condition1\_dum5 = 0;

Condition1\_dum6 = 0;

Condition1\_dum7 = 0;

Condition2\_dum0 = 0;

Condition2\_dum1 = 0;

Condition2\_dum2 = 0;

Condition2\_dum3 = 0;

Condition2\_dum4 = 0;

Condition2\_dum5 = 0;

Condition2\_dum6 = 0;

BldgType\_dum0 = 0;

BldgType\_dum1 = 0;

BldgType\_dum2 = 0;

BldgType\_dum3 = 0;

HouseStyle\_dum0 = 0;

HouseStyle\_dum1 = 0;

HouseStyle\_dum2 = 0;

HouseStyle\_dum3 = 0;

HouseStyle\_dum4 = 0;

HouseStyle\_dum5 = 0;

HouseStyle\_dum6 = 0;

RoofStyle\_dum0 = 0;

RoofStyle\_dum1 = 0;

RoofStyle\_dum2 = 0;

RoofStyle\_dum3 = 0;

RoofStyle\_dum4 = 0;

RoofMatl\_dum0 = 0;

RoofMatl\_dum1 = 0;

RoofMatl\_dum2 = 0;

RoofMatl\_dum3 = 0;

RoofMatl\_dum4 = 0;

RoofMatl\_dum5 = 0;

RoofMatl\_dum6 = 0;

Exterior1st\_dum0 = 0;

Exterior1st\_dum1 = 0;

Exterior1st\_dum2 = 0;

Exterior1st\_dum3 = 0;

Exterior1st\_dum4 = 0;

Exterior1st\_dum5 = 0;

Exterior1st\_dum6 = 0;

Exterior1st\_dum7 = 0;

Exterior1st\_dum8 = 0;

Exterior1st\_dum9 = 0;

Exterior1st\_dum10 = 0;

Exterior1st\_dum11 = 0;

Exterior1st\_dum12 = 0;

Exterior2nd\_dum0 = 0;

Exterior2nd\_dum1 = 0;

Exterior2nd\_dum2 = 0;

Exterior2nd\_dum3 = 0;

Exterior2nd\_dum4 = 0;

Exterior2nd\_dum5 = 0;

Exterior2nd\_dum6 = 0;

Exterior2nd\_dum7 = 0;

Exterior2nd\_dum8 = 0;

Exterior2nd\_dum9 = 0;

Exterior2nd\_dum10 = 0;

Exterior2nd\_dum11 = 0;

Exterior2nd\_dum12 = 0;

Exterior2nd\_dum13 = 0;

MasVnrType\_dum0 = 0;

MasVnrType\_dum1 = 0;

MasVnrType\_dum2 = 0;

ExterQual\_dum0 = 0;

ExterQual\_dum1 = 0;

ExterQual\_dum2 = 0;

ExterCond\_dum0 = 0;

ExterCond\_dum1 = 0;

ExterCond\_dum2 = 0;

Foundation\_dum0 = 0;

Foundation\_dum1 = 0;

Foundation\_dum2 = 0;

Foundation\_dum3 = 0;

Foundation\_dum4 = 0;

BsmtQual\_dum0 = 0;

BsmtQual\_dum1 = 0;

BsmtQual\_dum2 = 0;

BsmtQual\_dum3 = 0;

BsmtCond\_dum0 = 0;

BsmtCond\_dum1 = 0;

BsmtCond\_dum2 = 0;

BsmtExposure\_dum0 = 0;

BsmtExposure\_dum1 = 0;

BsmtExposure\_dum2 = 0;

BsmtExposure\_dum3 = 0;

BsmtFinType1\_dum0 = 0;

BsmtFinType1\_dum1 = 0;

BsmtFinType1\_dum2 = 0;

BsmtFinType1\_dum3 = 0;

BsmtFinType1\_dum4 = 0;

BsmtFinType1\_dum5 = 0;

BsmtFinType2\_dum0 = 0;

BsmtFinType2\_dum1 = 0;

BsmtFinType2\_dum2 = 0;

BsmtFinType2\_dum3 = 0;

BsmtFinType2\_dum4 = 0;

Heating\_dum0 = 0;

Heating\_dum1 = 0;

Heating\_dum2 = 0;

Heating\_dum3 = 0;

Heating\_dum4 = 0;

HeatingQC\_dum0 = 0;

HeatingQC\_dum1 = 0;

HeatingQC\_dum2 = 0;

HeatingQC\_dum3 = 0;

CentralAir\_dum0 = 0;

Electrical\_dum0 = 0;

Electrical\_dum1 = 0;

Electrical\_dum2 = 0;

Electrical\_dum3 = 0;

KitchenQual\_dum0 = 0;

KitchenQual\_dum1 = 0;

KitchenQual\_dum2 = 0;

Functional\_dum0 = 0;

Functional\_dum1 = 0;

Functional\_dum2 = 0;

Functional\_dum3 = 0;

Functional\_dum4 = 0;

Functional\_dum5 = 0;

FireplaceQu\_dum0 = 0;

FireplaceQu\_dum1 = 0;

FireplaceQu\_dum2 = 0;

FireplaceQu\_dum3 = 0;

FireplaceQu\_dum4 = 0;

GarageType\_dum0 = 0;

GarageType\_dum1 = 0;

GarageType\_dum2 = 0;

GarageType\_dum3 = 0;

GarageType\_dum4 = 0;

GarageFinish\_dum0 = 0;

GarageFinish\_dum1 = 0;

GarageQual\_dum0 = 0;

GarageQual\_dum1 = 0;

GarageQual\_dum2 = 0;

GarageQual\_dum3 = 0;

GarageCond\_dum0 = 0;

GarageCond\_dum1 = 0;

GarageCond\_dum2 = 0;

GarageCond\_dum3 = 0;

PavedDrive\_dum0 = 0;

PavedDrive\_dum1 = 0;

PoolQC\_dum0 = 0;

PoolQC\_dum1 = 0;

PoolQC\_dum2 = 0;

Fence\_dum0 = 0;

Fence\_dum1 = 0;

Fence\_dum2 = 0;

Fence\_dum3 = 0;

MiscFeature\_dum0 = 0;

MiscFeature\_dum1 = 0;

MiscFeature\_dum2 = 0;

MiscFeature\_dum3 = 0;

SaleType\_dum0 = 0;

SaleType\_dum1 = 0;

SaleType\_dum2 = 0;

SaleType\_dum3 = 0;

SaleType\_dum4 = 0;

if MSZoning = 'C' then MSZoning\_dum0 = 1;

if MSZoning = 'FV' then MSZoning\_dum1 = 1;

if MSZoning = 'RH' then MSZoning\_dum2 = 1;

if MSZoning = 'RL' then MSZoning\_dum3 = 1;

if Street = 'Grvl' then Street\_dum0 = 1;

if Alley = 'Gr' then Alley\_dum0 = 1;

if Alley = 'NA' then Alley\_dum1 = 1;

if LotShape = 'IR1' then LotShape\_dum0 = 1;

if LotShape = 'IR2' then LotShape\_dum1 = 1;

if LotShape = 'IR3' then LotShape\_dum2 = 1;

if LandContour = 'Bnk' then LandContour\_dum0 = 1;

if LandContour = 'HLS' then LandContour\_dum1 = 1;

if LandContour = 'Low' then LandContour\_dum2 = 1;

if Utilities = 'AllPub' then Utilities\_dum0 = 1;

if LotConfig = 'Corner' then LotConfig\_dum0 = 1;

if LotConfig = 'CulDSac' then LotConfig\_dum1 = 1;

if LotConfig = 'FR2' then LotConfig\_dum2 = 1;

if LotConfig = 'FR3' then LotConfig\_dum3 = 1;

if LandSlope = 'Gtl' then LandSlope\_dum0 = 1;

if LandSlope = 'Mod' then LandSlope\_dum1 = 1;

if Neighborhood = 'Blmngtn' then Neighborhood\_dum0 = 1;

if Neighborhood = 'Blueste' then Neighborhood\_dum1 = 1;

if Neighborhood = 'BrDale' then Neighborhood\_dum2 = 1;

if Neighborhood = 'BrkSide' then Neighborhood\_dum3 = 1;

if Neighborhood = 'ClearCr' then Neighborhood\_dum4 = 1;

if Neighborhood = 'CollgCr' then Neighborhood\_dum5 = 1;

if Neighborhood = 'Crawfor' then Neighborhood\_dum6 = 1;

if Neighborhood = 'Edwards' then Neighborhood\_dum7 = 1;

if Neighborhood = 'Gilbert' then Neighborhood\_dum8 = 1;

if Neighborhood = 'IDOTRR' then Neighborhood\_dum9 = 1;

if Neighborhood = 'MeadowV' then Neighborhood\_dum10 = 1;

if Neighborhood = 'Mitchel' then Neighborhood\_dum11 = 1;

if Neighborhood = 'NAmes' then Neighborhood\_dum12 = 1;

if Neighborhood = 'NPkVill' then Neighborhood\_dum13 = 1;

if Neighborhood = 'NWAmes' then Neighborhood\_dum14 = 1;

if Neighborhood = 'NoRidge' then Neighborhood\_dum15 = 1;

if Neighborhood = 'NridgHt' then Neighborhood\_dum16 = 1;

if Neighborhood = 'OldTown' then Neighborhood\_dum17 = 1;

if Neighborhood = 'SWISU' then Neighborhood\_dum18 = 1;

if Neighborhood = 'Sawyer' then Neighborhood\_dum19 = 1;

if Neighborhood = 'SawyerW' then Neighborhood\_dum20 = 1;

if Neighborhood = 'Somerst' then Neighborhood\_dum21 = 1;

if Neighborhood = 'StoneBr' then Neighborhood\_dum22 = 1;

if Neighborhood = 'Timber' then Neighborhood\_dum23 = 1;

if Condition1 = 'Artery' then Condition1\_dum0 = 1;

if Condition1 = 'Feedr' then Condition1\_dum1 = 1;

if Condition1 = 'Norm' then Condition1\_dum2 = 1;

if Condition1 = 'PosA' then Condition1\_dum3 = 1;

if Condition1 = 'PosN' then Condition1\_dum4 = 1;

if Condition1 = 'RRAe' then Condition1\_dum5 = 1;

if Condition1 = 'RRAn' then Condition1\_dum6 = 1;

if Condition1 = 'RRNe' then Condition1\_dum7 = 1;

if Condition2 = 'Artery' then Condition2\_dum0 = 1;

if Condition2 = 'Feedr' then Condition2\_dum1 = 1;

if Condition2 = 'Norm' then Condition2\_dum2 = 1;

if Condition2 = 'PosA' then Condition2\_dum3 = 1;

if Condition2 = 'PosN' then Condition2\_dum4 = 1;

if Condition2 = 'RRAe' then Condition2\_dum5 = 1;

if Condition2 = 'RRAn' then Condition2\_dum6 = 1;

if BldgType = '1Fam' then BldgType\_dum0 = 1;

if BldgType = '2fmCon' then BldgType\_dum1 = 1;

if BldgType = 'Duplex' then BldgType\_dum2 = 1;

if BldgType = 'Twnhs' then BldgType\_dum3 = 1;

if HouseStyle = '1.5Fin' then HouseStyle\_dum0 = 1;

if HouseStyle = '1.5Unf' then HouseStyle\_dum1 = 1;

if HouseStyle = '1Story' then HouseStyle\_dum2 = 1;

if HouseStyle = '2.5Fin' then HouseStyle\_dum3 = 1;

if HouseStyle = '2.5Unf' then HouseStyle\_dum4 = 1;

if HouseStyle = '2Story' then HouseStyle\_dum5 = 1;

if HouseStyle = 'SFoyer' then HouseStyle\_dum6 = 1;

if RoofStyle = 'Flat' then RoofStyle\_dum0 = 1;

if RoofStyle = 'Gable' then RoofStyle\_dum1 = 1;

if RoofStyle = 'Gambr' then RoofStyle\_dum2 = 1;

if RoofStyle = 'Hip' then RoofStyle\_dum3 = 1;

if RoofStyle = 'Mansard' then RoofStyle\_dum4 = 1;

if RoofMatl = 'ClyTile' then RoofMatl\_dum0 = 1;

if RoofMatl = 'CompShg' then RoofMatl\_dum1 = 1;

if RoofMatl = 'Membran' then RoofMatl\_dum2 = 1;

if RoofMatl = 'Metal' then RoofMatl\_dum3 = 1;

if RoofMatl = 'Roll' then RoofMatl\_dum4 = 1;

if RoofMatl = 'Tar&Grv' then RoofMatl\_dum5 = 1;

if RoofMatl = 'WdShake' then RoofMatl\_dum6 = 1;

if Exterior1st = 'AsbShng' then Exterior1st\_dum0 = 1;

if Exterior1st = 'BrkComm' then Exterior1st\_dum1 = 1;

if Exterior1st = 'BrkFace' then Exterior1st\_dum2 = 1;

if Exterior1st = 'CBlock' then Exterior1st\_dum3 = 1;

if Exterior1st = 'CemntBd' then Exterior1st\_dum4 = 1;

if Exterior1st = 'HdBoard' then Exterior1st\_dum5 = 1;

if Exterior1st = 'ImStucc' then Exterior1st\_dum6 = 1;

if Exterior1st = 'MetalSd' then Exterior1st\_dum7 = 1;

if Exterior1st = 'Plywood' then Exterior1st\_dum8 = 1;

if Exterior1st = 'Stone' then Exterior1st\_dum9 = 1;

if Exterior1st = 'Stucco' then Exterior1st\_dum10 = 1;

if Exterior1st = 'VinylSd' then Exterior1st\_dum11 = 1;

if Exterior1st = 'Wd Sdng' then Exterior1st\_dum12 = 1;

if Exterior2nd = 'AsbShng' then Exterior2nd\_dum0 = 1;

if Exterior2nd = 'AsphShn' then Exterior2nd\_dum1 = 1;

if Exterior2nd = 'Brk Cmn' then Exterior2nd\_dum2 = 1;

if Exterior2nd = 'BrkFace' then Exterior2nd\_dum3 = 1;

if Exterior2nd = 'CmentBd' then Exterior2nd\_dum5 = 1;

if Exterior2nd = 'HdBoard' then Exterior2nd\_dum6 = 1;

if Exterior2nd = 'ImStucc' then Exterior2nd\_dum7 = 1;

if Exterior2nd = 'MetalSd' then Exterior2nd\_dum8 = 1;

if Exterior2nd = 'Other' then Exterior2nd\_dum9 = 1;

if Exterior2nd = 'Plywood' then Exterior2nd\_dum10 = 1;

if Exterior2nd = 'Stone' then Exterior2nd\_dum11 = 1;

if Exterior2nd = 'Stucco' then Exterior2nd\_dum12 = 1;

if Exterior2nd = 'VinylSd' then Exterior2nd\_dum13 = 1;

if Exterior2nd = 'Wd Sdng' then Exterior2nd\_dum4 = 1;

if MasVnrType = 'BrkCmn' then MasVnrType\_dum0 = 1;

if MasVnrType = 'BrkFace' then MasVnrType\_dum1 = 1;

if MasVnrType = 'None' then MasVnrType\_dum2 = 1;

if ExterQual = 'Ex' then ExterQual\_dum0 = 1;

if ExterQual = 'Fa' then ExterQual\_dum1 = 1;

if ExterQual = 'Gd' then ExterQual\_dum2 = 1;

if ExterCond = 'Ex' then ExterCond\_dum0 = 1;

if ExterCond = 'Fa' then ExterCond\_dum1 = 1;

if ExterCond = 'Gd' then ExterCond\_dum2 = 1;

if Foundation = 'BrkTil' then Foundation\_dum0 = 1;

if Foundation = 'CBlock' then Foundation\_dum1 = 1;

if Foundation = 'PConc' then Foundation\_dum2 = 1;

if Foundation = 'Slab' then Foundation\_dum3 = 1;

if Foundation = 'Stone' then Foundation\_dum4 = 1;

if BsmtQual = 'Ex' then BsmtQual\_dum0 = 1;

if BsmtQual = 'Fa' then BsmtQual\_dum1 = 1;

if BsmtQual = 'Gd' then BsmtQual\_dum2 = 1;

if BsmtQual = 'NA' then BsmtQual\_dum3 = 1;

if BsmtCond = 'Fa' then BsmtCond\_dum0 = 1;

if BsmtCond = 'Gd' then BsmtCond\_dum1 = 1;

if BsmtCond = 'Po' then BsmtCond\_dum2 = 1;

if BsmtExposure = 'Av' then BsmtExposure\_dum0 = 1;

if BsmtExposure = 'Gd' then BsmtExposure\_dum1 = 1;

if BsmtExposure = 'Mn' then BsmtExposure\_dum2 = 1;

if BsmtExposure = 'NA' then BsmtExposure\_dum3 = 1;

if BsmtFinType1 = 'ALQ' then BsmtFinType1\_dum0 = 1;

if BsmtFinType1 = 'BLQ' then BsmtFinType1\_dum1 = 1;

if BsmtFinType1 = 'GLQ' then BsmtFinType1\_dum2 = 1;

if BsmtFinType1 = 'LwQ' then BsmtFinType1\_dum3 = 1;

if BsmtFinType1 = 'Rec' then BsmtFinType1\_dum4 = 1;

if BsmtFinType1 = 'NA' then BsmtFinType1\_dum5 = 1;

if BsmtFinType2 = 'ALQ' then BsmtFinType2\_dum0 = 1;

if BsmtFinType2 = 'BLQ' then BsmtFinType2\_dum1 = 1;

if BsmtFinType2 = 'GLQ' then BsmtFinType2\_dum2 = 1;

if BsmtFinType2 = 'LwQ' then BsmtFinType2\_dum3 = 1;

if BsmtFinType2 = 'Rec' then BsmtFinType2\_dum4 = 1;

if Heating = 'Floor' then Heating\_dum0 = 1;

if Heating = 'GasA' then Heating\_dum1 = 1;

if Heating = 'GasW' then Heating\_dum2 = 1;

if Heating = 'Grav' then Heating\_dum3 = 1;

if Heating = 'OthW' then Heating\_dum4 = 1;

if HeatingQC = 'Ex' then HeatingQC\_dum0 = 1;

if HeatingQC = 'Fa' then HeatingQC\_dum1 = 1;

if HeatingQC = 'Gd' then HeatingQC\_dum2 = 1;

if HeatingQC = 'Po' then HeatingQC\_dum3 = 1;

if CentralAir = 'N' then CentralAir\_dum0 = 1;

if Electrical = 'FuseA' then Electrical\_dum0 = 1;

if Electrical = 'FuseF' then Electrical\_dum1 = 1;

if Electrical = 'FuseP' then Electrical\_dum2 = 1;

if Electrical = 'NA' then Electrical\_dum3 = 1;

if KitchenQual = 'Ex' then KitchenQual\_dum0 = 1;

if KitchenQual = 'Fa' then KitchenQual\_dum1 = 1;

if KitchenQual = 'Gd' then KitchenQual\_dum2 = 1;

if Functional = 'Maj1' then Functional\_dum0 = 1;

if Functional = 'Maj2' then Functional\_dum1 = 1;

if Functional = 'Min1' then Functional\_dum2 = 1;

if Functional = 'Min2' then Functional\_dum3 = 1;

if Functional = 'Mod' then Functional\_dum4 = 1;

if Functional = 'Sev' then Functional\_dum5 = 1;

if FireplaceQu = 'Ex' then FireplaceQu\_dum0 = 1;

if FireplaceQu = 'Fa' then FireplaceQu\_dum1 = 1;

if FireplaceQu = 'Gd' then FireplaceQu\_dum2 = 1;

if FireplaceQu = 'NA' then FireplaceQu\_dum3 = 1;

if FireplaceQu = 'Po' then FireplaceQu\_dum4 = 1;

if GarageType = '2Types' then GarageType\_dum0 = 1;

if GarageType = 'Attchd' then GarageType\_dum1 = 1;

if GarageType = 'Basment' then GarageType\_dum2 = 1;

if GarageType = 'BuiltIn' then GarageType\_dum3 = 1;

if GarageType = 'CarPort' then GarageType\_dum4 = 1;

if GarageFinish = 'Fin' then GarageFinish\_dum0 = 1;

if GarageFinish = 'RFn' then GarageFinish\_dum1 = 1;

if GarageQual = 'Ex' then GarageQual\_dum0 = 1;

if GarageQual = 'Fa' then GarageQual\_dum1 = 1;

if GarageQual = 'Gd' then GarageQual\_dum2 = 1;

if GarageQual = 'Po' then GarageQual\_dum3 = 1;

if GarageCond = 'Ex' then GarageCond\_dum0 = 1;

if GarageCond = 'Fa' then GarageCond\_dum1 = 1;

if GarageCond = 'Gd' then GarageCond\_dum2 = 1;

if GarageCond = 'Po' then GarageCond\_dum3 = 1;

if PavedDrive = 'N' then PavedDrive\_dum0 = 1;

if PavedDrive = 'P' then PavedDrive\_dum1 = 1;

if PoolQC = 'Ex' then PoolQC\_dum0 = 1;

if PoolQC = 'Fa' then PoolQC\_dum1 = 1;

if PoolQC = 'Gd' then PoolQC\_dum2 = 1;

if Fence = 'GdPrv' then Fence\_dum0 = 1;

if Fence = 'GdWo' then Fence\_dum1 = 1;

if Fence = 'MnPrv' then Fence\_dum2 = 1;

if Fence = 'MnWw' then Fence\_dum3 = 1;

if MiscFeature = 'Gar2' then MiscFeature\_dum0 = 1;

if MiscFeature = 'NA' then MiscFeature\_dum1 = 1;

if MiscFeature = 'Othr' then MiscFeature\_dum2 = 1;

if MiscFeature = 'Shed' then MiscFeature\_dum3 = 1;

if SaleType = 'COD' then SaleType\_dum0 = 1;

if SaleType = 'CWD' then SaleType\_dum1 = 1;

if SaleType = 'Con' then SaleType\_dum2 = 1;

if SaleType = 'New' then SaleType\_dum3 = 1;

if SaleType = 'Oth' then SaleType\_dum4 = 1;

run;

\*Proc Reg to generate scatter, residual, Cook’s D plots and histograms;

\*This is the model after the log transforms were completed and erroneous variables were removed;

\*results and justifications for all adjustments are documented in the paper;

Proc reg data = train\_dummy;

Model LogSalePrice = IntLotFrontage

MSSUBCLASS

LogLotArea

OVERALLQUAL

OVERALLCOND

YEARBUILT

YEARREMODADD

LogMasVnrArea

LogBsmtFinSF1

BSMTFINSF2

BSMTUNFSF

Log1stFlrSF

'2ndFlrSF'n

LOWQUALFINSF

BSMTFULLBATH

BSMTHALFBATH

FULLBATH

HALFBATH

BEDROOMABVGR

KITCHENABVGR

TOTRMSABVGRD

FIREPLACES

GARAGEYRBLT

GARAGECARS

LogGarageArea

LogWoodDeckSF

LogOpenPorchSF

LogEnclosedPorch

Log3SsnPorch

LogScreenPorch

LogPoolArea

LogMiscVal

MOSOLD

YRSOLD

ALLEY\_DUM0

ALLEY\_DUM1

LOTSHAPE\_DUM0

LOTSHAPE\_DUM1

LOTSHAPE\_DUM2

LANDCONTOUR\_DUM0

LANDCONTOUR\_DUM1

LANDCONTOUR\_DUM2

LOTCONFIG\_DUM0

LOTCONFIG\_DUM1

LOTCONFIG\_DUM2

LOTCONFIG\_DUM3

LANDSLOPE\_DUM0

LANDSLOPE\_DUM1

NEIGHBORHOOD\_DUM0

NEIGHBORHOOD\_DUM1

NEIGHBORHOOD\_DUM2

NEIGHBORHOOD\_DUM3

NEIGHBORHOOD\_DUM4

NEIGHBORHOOD\_DUM5

NEIGHBORHOOD\_DUM6

NEIGHBORHOOD\_DUM7

NEIGHBORHOOD\_DUM8

NEIGHBORHOOD\_DUM9

NEIGHBORHOOD\_DUM10

NEIGHBORHOOD\_DUM11

NEIGHBORHOOD\_DUM12

NEIGHBORHOOD\_DUM13

NEIGHBORHOOD\_DUM14

NEIGHBORHOOD\_DUM15

NEIGHBORHOOD\_DUM16

NEIGHBORHOOD\_DUM17

NEIGHBORHOOD\_DUM18

NEIGHBORHOOD\_DUM19

NEIGHBORHOOD\_DUM20

NEIGHBORHOOD\_DUM21

NEIGHBORHOOD\_DUM22

NEIGHBORHOOD\_DUM23

CONDITION1\_DUM0

CONDITION1\_DUM1

CONDITION1\_DUM2

CONDITION1\_DUM3

CONDITION1\_DUM4

CONDITION1\_DUM5

CONDITION1\_DUM6

CONDITION1\_DUM7

CONDITION2\_DUM0

CONDITION2\_DUM1

CONDITION2\_DUM2

CONDITION2\_DUM3

CONDITION2\_DUM4

BLDGTYPE\_DUM0

BLDGTYPE\_DUM1

BLDGTYPE\_DUM2

BLDGTYPE\_DUM3

HOUSESTYLE\_DUM0

HOUSESTYLE\_DUM1

HOUSESTYLE\_DUM2

HOUSESTYLE\_DUM3

HOUSESTYLE\_DUM4

HOUSESTYLE\_DUM5

HOUSESTYLE\_DUM6

ROOFSTYLE\_DUM0

ROOFSTYLE\_DUM1

ROOFSTYLE\_DUM2

ROOFSTYLE\_DUM3

ROOFSTYLE\_DUM4

ROOFMATL\_DUM0

ROOFMATL\_DUM1

ROOFMATL\_DUM2

ROOFMATL\_DUM4

ROOFMATL\_DUM5

ROOFMATL\_DUM6

EXTERIOR1ST\_DUM0

EXTERIOR1ST\_DUM1

EXTERIOR1ST\_DUM2

EXTERIOR1ST\_DUM3

EXTERIOR1ST\_DUM4

EXTERIOR1ST\_DUM5

EXTERIOR1ST\_DUM6

EXTERIOR1ST\_DUM7

EXTERIOR1ST\_DUM8

EXTERIOR1ST\_DUM9

EXTERIOR1ST\_DUM10

EXTERIOR1ST\_DUM11

EXTERIOR1ST\_DUM12

EXTERIOR2ND\_DUM0

EXTERIOR2ND\_DUM1

EXTERIOR2ND\_DUM2

EXTERIOR2ND\_DUM3

EXTERIOR2ND\_DUM4

EXTERIOR2ND\_DUM5

EXTERIOR2ND\_DUM6

EXTERIOR2ND\_DUM7

EXTERIOR2ND\_DUM8

EXTERIOR2ND\_DUM9

EXTERIOR2ND\_DUM10

EXTERIOR2ND\_DUM11

EXTERIOR2ND\_DUM12

EXTERIOR2ND\_DUM13

MASVNRTYPE\_DUM0

MASVNRTYPE\_DUM1

MASVNRTYPE\_DUM2

EXTERQUAL\_DUM0

EXTERQUAL\_DUM1

EXTERQUAL\_DUM2

EXTERCOND\_DUM0

EXTERCOND\_DUM1

EXTERCOND\_DUM2

FOUNDATION\_DUM0

FOUNDATION\_DUM1

FOUNDATION\_DUM2

FOUNDATION\_DUM3

FOUNDATION\_DUM4

BSMTQUAL\_DUM0

BSMTQUAL\_DUM1

BSMTQUAL\_DUM2

BSMTQUAL\_DUM3

BSMTCOND\_DUM0

BSMTCOND\_DUM1

BSMTCOND\_DUM2

BSMTEXPOSURE\_DUM0

BSMTEXPOSURE\_DUM1

BSMTEXPOSURE\_DUM2

BSMTEXPOSURE\_DUM3

BSMTFINTYPE1\_DUM0

BSMTFINTYPE1\_DUM1

BSMTFINTYPE1\_DUM2

BSMTFINTYPE1\_DUM3

BSMTFINTYPE1\_DUM4

BSMTFINTYPE2\_DUM0

BSMTFINTYPE2\_DUM1

BSMTFINTYPE2\_DUM2

BSMTFINTYPE2\_DUM3

BSMTFINTYPE2\_DUM4

HEATING\_DUM1

HEATING\_DUM2

HEATING\_DUM3

HEATING\_DUM4

HEATINGQC\_DUM0

HEATINGQC\_DUM1

HEATINGQC\_DUM2

HEATINGQC\_DUM3

CENTRALAIR\_DUM0

ELECTRICAL\_DUM0

ELECTRICAL\_DUM1

ELECTRICAL\_DUM2

ELECTRICAL\_DUM3

KITCHENQUAL\_DUM0

KITCHENQUAL\_DUM1

KITCHENQUAL\_DUM2

FUNCTIONAL\_DUM0

FUNCTIONAL\_DUM1

FUNCTIONAL\_DUM2

FUNCTIONAL\_DUM3

FUNCTIONAL\_DUM4

FIREPLACEQU\_DUM0

FIREPLACEQU\_DUM1

FIREPLACEQU\_DUM2

FIREPLACEQU\_DUM3

FIREPLACEQU\_DUM4

GARAGETYPE\_DUM0

GARAGETYPE\_DUM1

GARAGETYPE\_DUM2

GARAGETYPE\_DUM3

GARAGETYPE\_DUM4

GARAGEFINISH\_DUM0

GARAGEFINISH\_DUM1

GARAGEQUAL\_DUM0

GARAGEQUAL\_DUM1

GARAGEQUAL\_DUM2

GARAGEQUAL\_DUM3

GARAGECOND\_DUM0

GARAGECOND\_DUM1

GARAGECOND\_DUM2

GARAGECOND\_DUM3

PAVEDDRIVE\_DUM0

PAVEDDRIVE\_DUM1

POOLQC\_DUM0

POOLQC\_DUM1

POOLQC\_DUM2

FENCE\_DUM0

FENCE\_DUM1

FENCE\_DUM2

FENCE\_DUM3

MISCFEATURE\_DUM1

MISCFEATURE\_DUM2

MISCFEATURE\_DUM3;

Run;

\*/

\*Forward Selection method using variables obtained through the proc reg cleanup;

\*stats =PRESS used to get CV Press Score ;

\*import training data;

Proc glmselect data=train\_dummy;

class MSZoning

Street

Alley

LotShape

LandContour

Utilities

LotConfig

LandSlope

Neighborhood

Condition1

Condition2

BldgType

HouseStyle

RoofStyle

RoofMatl

Exterior1st

Exterior2nd

MasVnrType

ExterQual

ExterCond

Foundation

BsmtQual

BsmtCond

BsmtExposure

BsmtFinType1

BsmtFinType2

Heating

HeatingQC

CentralAir

Electrical

KitchenQual

Functional

FireplaceQu

GarageType

GarageFinish

GarageQual

GarageCond

PavedDrive

PoolQC

Fence

MiscFeature

SaleType

SaleCondition;

model LogSalePrice = IntLotFrontage MSSubClass MSZoning LogLotArea Street Alley LotShape LandContour Utilities LotConfig LandSlope Neighborhood Condition1 Condition2 BldgType HouseStyle OverallQual OverallCond YearBuilt YearRemodAdd RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType LogMasVnrArea ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure BsmtFinType1 LogBsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF Heating HeatingQC CentralAir Electrical Log1stFlrSF '2ndFlrSF'n LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType GarageYrBlt GarageFinish GarageCars LogGarageArea GarageQual GarageCond PavedDrive LogWoodDeckSF LogOpenPorchSF LogEnclosedPorch Log3SsnPorch LogScreenPorch LogPoolArea PoolQC Fence MiscFeature LogMiscVal MoSold YrSold SaleType SaleCondition

/selection=Forward stats =PRESS;

Run;

\*Backward Selection method using variables obtained through the proc reg cleanup;

\*stats =PRESS used to get CV Press Score ;

Proc glmselect data=train\_dummy;

class MSZoning

Street

Alley

LotShape

LandContour

Utilities

LotConfig

LandSlope

Neighborhood

Condition1

Condition2

BldgType

HouseStyle

RoofStyle

RoofMatl

Exterior1st

Exterior2nd

MasVnrType

ExterQual

ExterCond

Foundation

BsmtQual

BsmtCond

BsmtExposure

BsmtFinType1

BsmtFinType2

Heating

HeatingQC

CentralAir

Electrical

KitchenQual

Functional

FireplaceQu

GarageType

GarageFinish

GarageQual

GarageCond

PavedDrive

PoolQC

Fence

MiscFeature

SaleType

SaleCondition;

model LogSalePrice = IntLotFrontage MSSubClass MSZoning LogLotArea Street Alley LotShape LandContour Utilities LotConfig LandSlope Neighborhood Condition1 Condition2 BldgType HouseStyle OverallQual OverallCond YearBuilt YearRemodAdd RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType LogMasVnrArea ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure BsmtFinType1 LogBsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF Heating HeatingQC CentralAir Electrical Log1stFlrSF '2ndFlrSF'n LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType GarageYrBlt GarageFinish GarageCars LogGarageArea GarageQual GarageCond PavedDrive LogWoodDeckSF LogOpenPorchSF LogEnclosedPorch Log3SsnPorch LogScreenPorch LogPoolArea PoolQC Fence MiscFeature LogMiscVal MoSold YrSold SaleType SaleCondition

/selection=Backward stats =PRESS;

Run;

\*Stepwise Selection method using variables obtained through the proc reg cleanup;

\*stats =PRESS used to get CV Press Score ;

Proc glmselect data=train\_dummy;

class MSZoning

Street

Alley

LotShape

LandContour

Utilities

LotConfig

LandSlope

Neighborhood

Condition1

Condition2

BldgType

HouseStyle

RoofStyle

RoofMatl

Exterior1st

Exterior2nd

MasVnrType

ExterQual

ExterCond

Foundation

BsmtQual

BsmtCond

BsmtExposure

BsmtFinType1

BsmtFinType2

Heating

HeatingQC

CentralAir

Electrical

KitchenQual

Functional

FireplaceQu

GarageType

GarageFinish

GarageQual

GarageCond

PavedDrive

PoolQC

Fence

MiscFeature

SaleType

SaleCondition;

model LogSalePrice = IntLotFrontage MSSubClass MSZoning LogLotArea Street Alley LotShape LandContour Utilities LotConfig LandSlope Neighborhood Condition1 Condition2 BldgType HouseStyle OverallQual OverallCond YearBuilt YearRemodAdd RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType LogMasVnrArea ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure BsmtFinType1 LogBsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF Heating HeatingQC CentralAir Electrical Log1stFlrSF '2ndFlrSF'n LowQualFinSF BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType GarageYrBlt GarageFinish GarageCars LogGarageArea GarageQual GarageCond PavedDrive LogWoodDeckSF LogOpenPorchSF LogEnclosedPorch Log3SsnPorch LogScreenPorch LogPoolArea PoolQC Fence MiscFeature LogMiscVal MoSold YrSold SaleType SaleCondition

/selection=Stepwise stats =PRESS;

Run;

/\*

Custom Function

Created by using select SL for Significance level-based elimination, and iterating the process removing one variable at a time until we have an ideal fit.

Resulted in a higher r-square and lower PRESS variable than any other model

FireplaceQu

HouseStyle

BsmtFinType2

LogMiscVal

GarageYrBlt

LandContour

BsmtHalfBath

MoSold

LogMasVnrArea

BsmtCond

Electrical

ExterQual

LowQualFinSF

BedroomAbvGr

RoofStyle

MiscFeature

YrSold

\*/

\*Forward Selection method using variables obtained through the proc reg cleanup;

\*(stop=CV) cvmethod=random(5) used to get CVPress Score per the Unit Learning presentations;

Proc glmselect data=train\_dummy;

class MSZoning

Street

Alley

LotShape

Utilities

LotConfig

LandSlope

Neighborhood

Condition1

Condition2

BldgType

RoofMatl

Exterior1st

Exterior2nd

MasVnrType

ExterCond

Foundation

BsmtQual

BsmtCond

BsmtExposure

BsmtFinType1

Heating

HeatingQC

CentralAir

KitchenQual

Functional

GarageType

GarageFinish

GarageQual

GarageCond

PavedDrive

PoolQC

Fence

SaleType

SaleCondition;

model LogSalePrice =

IntLotFrontage

MSSubClass

MSZoning

LogLotArea

Street

Alley

LotShape

Utilities

LotConfig

LandSlope

Neighborhood

Condition1

Condition2

BldgType

OverallQual

OverallCond

YearBuilt

YearRemodAdd

RoofMatl

Exterior1st

Exterior2nd

MasVnrType

ExterCond

Foundation

BsmtQual

BsmtExposure

BsmtFinType1

LogBsmtFinSF1

BsmtFinSF2

BsmtUnfSF

Heating

HeatingQC

CentralAir

Log1stFlrSF

'2ndFlrSF'n

BsmtFullBath

FullBath

HalfBath

KitchenAbvGr

KitchenQual

TotRmsAbvGrd

Functional

Fireplaces

GarageType

GarageFinish

GarageCars

LogGarageArea

GarageQual

GarageCond

PavedDrive

LogWoodDeckSF

LogOpenPorchSF

LogEnclosedPorch

Log3SsnPorch

LogScreenPorch

LogPoolArea

PoolQC

Fence

SaleType

SaleCondition

/selection=Forward(select=SL) stats =PRESS;