PROC IMPORT DATAFILE="/folders/myfolders/sasuser.v94/house\_price.xlsx"

   OUT=WORK.HOUSE\_PRICE

   DBMS=XLSX

   REPLACE;

RUN;

PROC IMPORT DATAFILE="/folders/myfolders/sasuser.v94/house\_price\_filtered\_test.xlsx"

   OUT=WORK.HOUSE\_PRICE\_TEST

   DBMS=XLSX

   REPLACE;

RUN;

/\* after you've imported the training and test sets \*/

data test;

set work.house\_price\_filtered\_test;

SalePrice = .;

RUN;

data train;

set work.log\_House\_price\_filter\_out\_train;

RUN;

data projection3;

set train test;

run;

proc glm data = projection3 plots = all;

class RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType;

model SalePrice = RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType LotArea BedroomAbvGr YearBuilt / cli;

output out = results p = Predict;

run;

/\* Can't have negative predictions because of RMSLE \*/

/\* Also must have only two columns with appropriate labels. \*/

data results3;

set results;

if Predict > 0 then Salerice = Predict;

if Predict < 0 then SalePrice = 100000;

keep id SalePrice;

where id> 1460;

;

proc means data = results3;

var SalePrice;

run;









