/\*ANALYSIS 1\*/

/\*Import train\_log.csv file\*/

**proc** **import** datafile="/home/u50452984/train\_log.csv"

out=train

dbms=csv

replace;

**run**;

/\*log-log\*/

**proc** **glm** data = train plots=ALL;

class Neighborhood (ref = "BrkSide");

model lSalePrice = lGrLivArea|Neighborhood /solution;

**run**;

/\*saleprice vs grlivarea -- getting internal CV\*/

**proc** **glmselect** data=train plots(stepaxis = number) = (criterionpanel ASEplot);

model lSalePrice = lGrLivArea

/ selection = stepwise(choose = cv stop=CV) cvdetails=all;

/\*neighborhood\*/

**proc** **glmselect** data=train plots(stepaxis = number) = (criterionpanel ASEplot);

class Neighborhood (ref = "BrkSide");

model lSalePrice = lGrLivArea Neighborhood

/ selection = stepwise(choose = cv stop=CV) cvdetails=all;

/\*neighborhood and interaction -- getting internal CV\*/

**proc** **glmselect** data=train plots(stepaxis = number) = (criterionpanel ASEplot);

class Neighborhood (ref = "BrkSide");

model lSalePrice = lGrLivArea|Neighborhood

/ selection = stepwise(choose = cv stop=CV) cvdetails=all;

/\*ANALYSIS 2\*/

/\*Import train\_q2.csv file\*/

**proc** **import** datafile="/home/u50452984/train\_q2.csv"

out=train\_q2

dbms=csv

replace;

**run**;

**proc** **glmselect** data=train\_q2 plots(stepaxis = number) = (criterionpanel ASEplot);

model lSalePrice = MSSubClass--lGrLivArea lTotalBsmtSF--lX2ndFlrSF

/ selection = stepwise(choose = cv stop=CV) cvdetails=all;

**proc** **glmselect** data=train\_q2 plots(stepaxis = number) = (criterionpanel ASEplot);

model lSalePrice = MSSubClass--lGrLivArea lTotalBsmtSF--lX2ndFlrSF

/ selection = forward(choose = cv stop=CV) cvdetails=all;

**proc** **glmselect** data=train\_q2 plots(stepaxis = number) = (criterionpanel ASEplot);

model lSalePrice = MSSubClass--lGrLivArea lTotalBsmtSF--lX2ndFlrSF

/ selection = backward(choose = cv stop=CV) cvdetails=all;