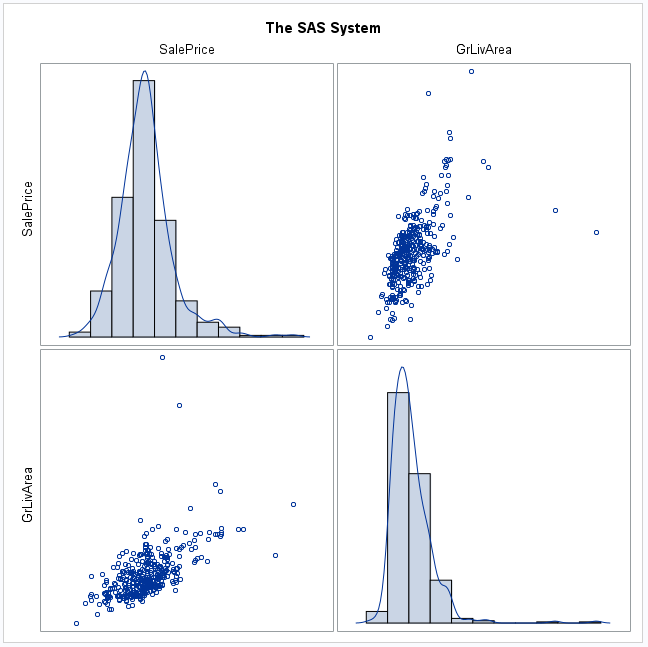
# Assumptions

## Linear Relationship & Normality

**proc** **sgscatter** data = neighborhoods;

matrix SalePrice GrLivArea / diagonal=(histogram kernel);

**run**;



Fail.

Both the independent and the dependent variables exhibit signs of right-skewness. Running a log-log model to correct for this. Also, coding neighborhoods as dummy variables for later use.

**data** loghood;

set neighborhoods;

logprice = log(SalePrice);

logarea = log(GrLivArea);

BrkSide = (Neighborhood = "BrkSide");

NAmes = (Neighborhood = "NAmes");

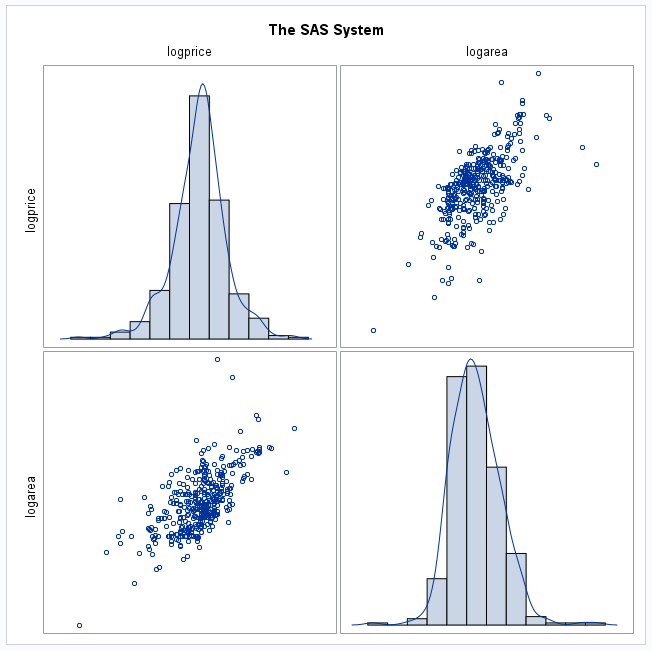
**run**;

## Linear Relationship

**proc** **sgscatter** data = loghood;

matrix logprice logarea / diagonal=(histogram kernel);

**run**;



Pass. The relationship between the log of the area and the log of the price does seem to exhibit a linear relationship.

## Multivariate Normality

Pass. Using the matrix above, it is evident that both inputs have normality.

## No Multicollinearity

## No auto-correlation

## Homoscedasticity