# Thomas Jones – CS5530-0005 – Homework #2

**Commentary:**

*For the linear regression:*

We can see from the coefficients that price is:

* Highly correlated with carat
* Negatively correlated with depth, table, and width
* Correlated with length
* Negatively correlated with height

R^2 indicates a good percentage of the price variance is due to the input variables.

Given the mix of coefficients between height, width, and length plus the visualization of carat vs. price the polynomial regression may result in a better model.

*For the non-linear regression:*

The coefficients representing carat, carat^2, carat\*depth, depth, depth^2, etc. are shown above. These coefficients are less interpreable than the linear model because of the additional complexities associated with polynomial models. The linear components can be interpreted as general trends though the other components are less clear.

*Mean Squared Error Discussion:*

The regressions were run against normalized as non-normalized data. The normalized MSE data shows some improvement for non-linear vs linear, though scale is lost. For the non-linear data, we find the square root of MSE to be 1495 vs 1402 vs non-linear, i.e. the non-linear model produces a price that is $92 more accurate than the linear model.