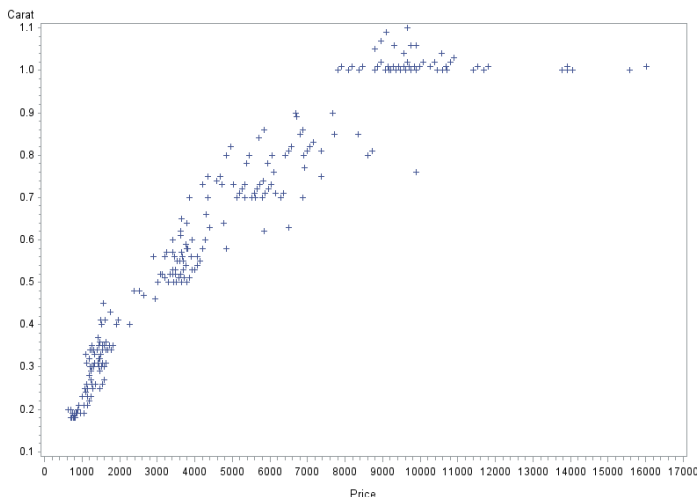


Diamond Write Up

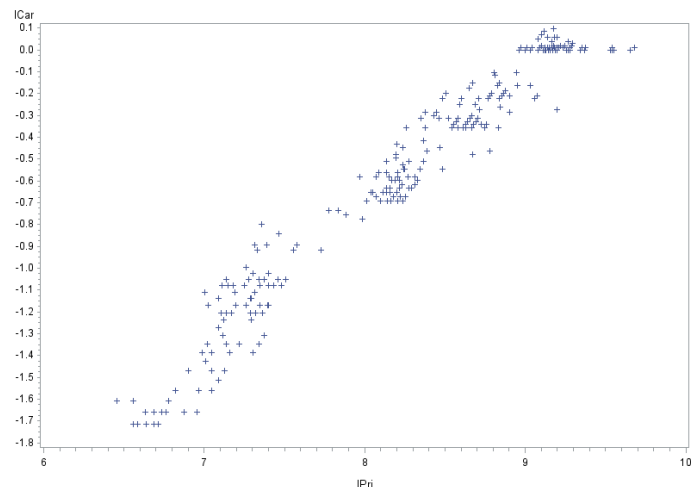
Introduction:

In March 2000 Dr. Chu asked his MBA students to predict the price of a diamond based off the four C's (Carat, Color, Clarity, and Cert). The focus was to predict Price (in Singapore \$) using Carat, where one carat is equal to 0.2 grams.

The first objective was to find the correlation coefficient between the Price and Carat of the diamond, and it shows a relatively strong correlation at 0.94473. However, upon looking at the graph (see Fig 1.1) there is a definite curve, which makes prediction difficult for heavier diamonds. To account for this, we perform a log transformation where we take the log of each carat and price to see how much better that is for predicting. Predicting power increases, as seen in the rise of the correlation coefficient, to 0.97846 (see Fig 1.2)



(Fig 1.1)



(Fig 1.2)

Analysis:

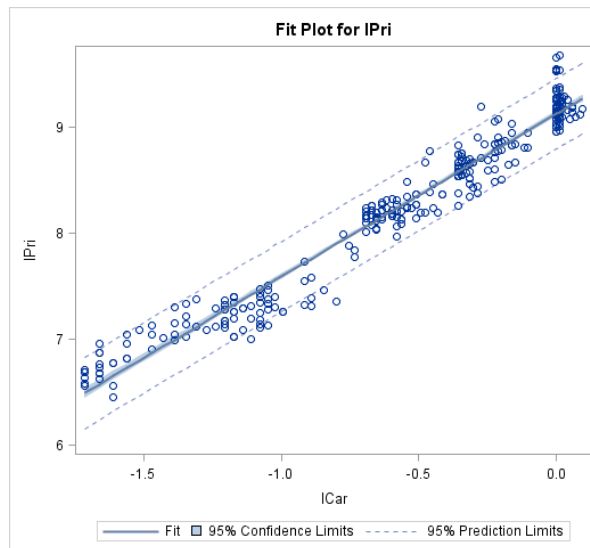
The response variable is the price (Singapore\$) of the diamond, and the explanatory variable is the weight of the diamond. The model we created is as follows:

$$\ln(\text{Price}) = \beta_0 + \beta_1 \ln(\text{Carat}) + \epsilon, \epsilon \sim N(0, \sigma^2)$$

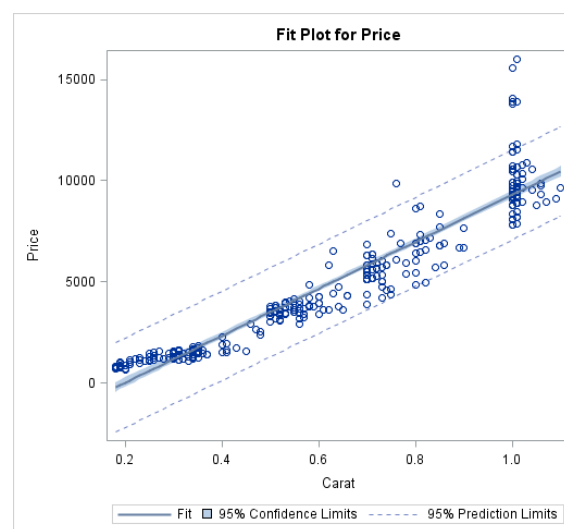
and our parameter estimates, and standard error are shown below

Variable	DF	Parameter Estimate	Standard Error
Intercept	1	9.12775	0.01440
ICar	1	1.53726	0.01854

In the parameter estimate for LCar (Log Carat), the value 1.53726 means that for a 1 log carat change in diamond weight, the price will increase 1.53726 log Singapore\$. Furthermore, Carat is a statistically significant variable in determining price. The following graph (see Fig 1.3) shows our 95% Confidence interval (represented by the thin blue shaded area) ignoring the dotted lines.



(Fig 1.3)



(Fig 1.4)

For a newly engaged couple, they can focus on Fig 1.4, where they can look between the dotted lines to get a 95% prediction interval for how much the price will be.

One thing is to be said about all of this though. While Carat is a good predictor for price, and log Carat is a slightly better predictor for price, this doesn't include the other C's of color, clarity, and cert. which also affects price. For example, the R^2 for the log price model is 0.9572, which is great, but shows room for improvement. Overall, there is a statistically significant weight effect on price.