USING THE "FIND NORMALIZING TRANSFORMATIONS" FEATURE
(Corresponds to Section 13.2 in textbook)

- A tool to help *you make decisions*.
- It won't "tell you the right answer."

Considerations to take into account:

1. The values found by the software are just estimates – it's silly to try to get too precise.

   Example: The software suggests exponent 0.504 for TeachTax .

   This is silly to use as an exponent, but it suggests trying exponent 0.5.

In other words:

- Do *not* stop with the values the software spits out.
- Instead, use them as *starting points* in your decisions.

2. One important consideration: What exponents make sense in context?

Examples:

- Exponent 1 is the easiest to interpret.
- "Exponent" 0 (logs) also can be interpreted (multiplicative rather than additive scale)
- Exponent 2 (respectively, 3) makes sense for a predictor when response is an area (respectively, volume) and the predictor is a linear measure.
- Exponent ½ (respectively, 1/3) makes sense for a predictor when response is a linear measure and the predictor is an area (respectively, volume). (Or maybe transform response instead of predictor, using exponent 2 or 3, respectively.)
- Exponent -1 makes sense when the variable is a ratio (e.g., gallons per mile instead of miles per gallon.)
- Other exponents might make sense in particular contexts.
- Integers or fractions of integers make more sense than other numbers.

3. Excessive detail can also lead to overfitting.

Remember: You just have a sample, not the population.

Parsimony is always one (not the only) consideration in model building.

4. Check out possible alternatives with a Likelihood Ratio Test (LRT). (Example later.)

5. Model selection methods are fairly robust as long as predictors are linearly related.

So look at the scatterplot matrix when p-values from tests are marginal.

Exercise: Find suitable transformations for the Big Mac data (using Big Mac, Bread, BusFare, TeachSal, and TeachTax) after seeing the example we do in class.