Lesson 4-1 Modeling Nonlinear Data

When we attempt to model the relationship between two quantitative variables, we often find that the relationship is not linear. Whenever possible, we try to understand the underlying relationship between the variables, but sometimes the relationship is not obvious. In that case, we attempt to **re-express** the data by taking the logarithm, the square root, the reciprocal, or some other mathematical operation on the values in the data set.

# The *Ladder of Powers*

This list places in order the effects that many re-expressions have on a data set.

|  |  |  |  |
| --- | --- | --- | --- |
| **Power** | **Shape** | **Transformed Value** | **Name** |
| 3 |  | (Original value) 3 | Cube |
| 2 |  | (Original value) 2 | Square |
| 1 |  | (Original value) | No transformation |
| ½ |  |  | Square root |
| ⅓ |  |  | Cube Root |
| 0 |  | Log(Original value) | Logarithm |
| -½ |  |  | Reciprocal of square root |
| -1 |  |  | Reciprocal |

# What Can Go Wrong?

* Don’t expect your model to be perfect.
* Don’t choose your model based on the value of r2 alone.
* Don’t stray too far from the ladder; the most appropriate model is probably not too complicated.