Modeling Summary – **ALL MODELS ARE WRONG, SOME MODELS ARE USEFUL!**

linear ($y = mx + b$),

We can expect to model a set of data with a linear model if for consecutive values of the independent variable, we find a *common difference* in dependent variable values. We can describe the slope as the expected or average change in the dependent variable for each unit change in the independent variable. The y-intercept may or may not have a meaningful interpretation.

power ($y = ax^b$)

Power functions can be difficult to recognize in modeling because they often look like exponential or logarithmic functions. The differences can be subtle and only emerge as the data accumulates. Power equations often contain (or potentially contain) the point (0, 0) and model increases or decreases that do not follow exponential or logarithmic data trends. Power trends will always eventually outgrow any logarithm. Similarly, powers will always eventually undergrow any exponential. However when performing a power regression on the TI, you cannot include this point because of the method the calculator uses to find the power equation. Areas and volumes (weights) are often modeled by a power equation.

Polynomial: quadratic ($y = ax^2 + bx + c$), cubic ($y = ax^3 + bx^2 + cx + d$), etc.

Polynomial functions are simply the sum of power functions. As a result, the end or long term behavior will resemble a power function but near the origin or in the short term it may wiggle up and down crossing the x-axis at several roots and having several maxima and minima. It is this wiggling that ultimately distinguishes polynomials in modeling situations. The stock market and demand for electricity are just two examples of situations of situations that may be modeled with a polynomial.

exponential ($y = ab^x$)

We can expect to model a set of data with an exponential model if for consecutive values of the independent variable, we find a *common ratio* in dependent variable values. If $b > 1$, the data exhibits exponential growth. If $0 < b < 1$, then the data exhibits exponential decay. Things that change over time are often modeled exponentially. The rate of growth of bacteria or the decay of radioactive material are often modeled exponentially.

logarithmic ($y = a + b \ln x$)

Logarithms increase by a constant interval over multiple intervals, in other words, you will find a constant change in the dependent variable for a multiple (common ratio) change in the independent variable. A logarithmic model may be used when something increases (or decreases) rapidly initially, then the rate of change gradually changes rather substantially. Sound intensity, earthquake intensity and pH levels use a logarithmic scale.

When determining the most useful model, one should carefully ***consider*** the following:

Does it make sense in context?

Compare the r values for competing models. The closer |r| is to 1, the better the model is theoretically.

Look at the residuals - they should be random. If you see curvature, try something else!

Remember - **ALL MODELS ARE WRONG, SOME MODELS ARE USEFUL!**

Choose the simplest model (as long as it makes sense in context).

Be careful of extrapolation. Using your model to predict for a value that is far from the data will often be inaccurate and unreliable.