Linear Regression

This will calculate the simple linear regression relationship between two variables given a set of paired observations on the two variables. The samples must be equal sizes. The method of least squares is used, to describe the relationship between the two variables and to predict a value of the dependent variable, y, given a value of the independent variable, x.

Method: The method of least squares amounts to finding the slope (m) and intercept (c) of the line y = mx + c such that the sum of (yi - m xi - c)^2 taken over all data points is a minimum, where (xi, yi) i=1,...,n denote the (paired) observations. It is usual practice to assume a model of the form:

y = a + bx + e, where the random errors, e, are assumed to be (approximately) normally distributed with mean zero and constant variance, s2. For this model, m and c are estimates of b and a respectively and an estimate of s2, is given by:

s2 = ∑(yi - mxi - c)^2 /(n-2).

Confidence intervals for a and b (at the level set in the Preferences dialog or overridden by the pop up at the bottom of the linear regression dialog box) are also presented. The Scatterplot option can also be used to calculate the linear regression equation also (and will place it on the plot). The Save Residuals check box near the top right will create a column of data in the spreadsheet with the residual values (magnitude of deviation from the linear regression equation closest value).

See the Statistics topic for instructions on selecting this test.