1. Use the above dataset, calculate the following:

* 1. Means, standard deviations, sums of squares and cross products, and the three r’s (correlation between y and x1, y and x2 and x1 and x2).
  2. Regression equation of Y on X1 and X2. What method does linear regression use for parameter estimation? Please further elaborate the method.
  3. SSreg, SSres, R2y.12 (squared multiple correlation coefficient), F, S2y.12 (variance of estimate).
  4. t ratios for the two regression coefficients. How to interpret the two regression coefficients? Which variable, X1 or X2, is more important in predicting or explaining the outcome variable?
  5. Increment in the proportion of variance accounted for by x2, over and above x1, and the F ratio for the test of this increment. To what test statistic, calculated earlier, should this F be equal?

2. Use the above dataset, calculate the following (conti):

* 1. Increment in the proportion of variance accounted for by x1, over and above x2, and the F ratio for the test of this increment. To what test statistics, calculated earlier, should this F be equal?
  2. Using the regression equation, calculate the predicted y’s (ŷ), y-ŷ (residuls), Σ (y-ŷ ), and Σ (y-ŷ )2.
  3. Calculate the squared correlation between y and ŷ. To what statistic, calculated earlier, shoud this correlation be equal? How to interpret the squared multiple correlation coeffiefient?
  4. For all subjects, obtain the following: standardized residuals (zresid), studentized residuals (sresid), studentized deleted residuals (sdresid), leverage (h), Cook’s D, and DFBETAs (raw and standardized).
  5. Based on the stasticis you got from step (i), is there evidence for outliers and influential points?