Input Data

The data set consisting of 15 observations (from 15 different sales territories), is given in Table 10.1. The dataset is referred to as Regdata 1.

Correlation

First, let us look at the correlations of all the variables with each other. The correlation table (output from the computer for the Pearson Correlation procedure) is shown in Table 10.2. The values in the correlation table are standardised, and range from 0 to 1 (+ ve and – ve). Looking at the last column, we find that except for COMPET (index of competitor activity), all other variables are highly correlated (ranging from .73 to .95) with sales. This means we may have chosen a fairly good set of independent variables (no. of dealers, sales potential, no. of customers, no. of service people, no. of salespeople) to try and correlate with sales. Only the index of competitor activity does not appear to be strongly correlated (correlation coefficient is – .05) with sales. But we must remember that these correlations in Table 10.2 are one-to-one correlations of each variable with the other. So we may still want to do a multiple regression with an independent variable showing low correlation with a dependent variable, because in the presence of other variables, this independent variable may become a good predictor of the dependent variable.

The other point to be noted in the correlation table is whether independent variables are highly correlated with each other. If they are, like in Table 10.2, this may indicate that they are not independent of each other, and we may be able to use only 1 or 2 of them to predict the dependent variables. As we will see later, our regression ends up eliminating some of the independent variables, because all six of them are not required. Some of them, being correlated with other variables, do not add any value to the regression model.

We now move on to the regression analysis of the same data.

Regression

We will first run the regression model of the following form, by entering all the 6 '*x*' variables in the model

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 \qquad \text{...Equation 1}$$

and determine the values of a, b_1, b_2, b_3, b_4, b_5, & b_6.

Regression Output

The results (output) of this regression model are in Table 10.4 in table form. Column 4 of the table, titled '*B*' lists all the coefficients for the model. According to this,