

### 3.3 Salary vs. GPA

- a. We consider Figure 3.16. From this figure, we find the regression equation to be,

$$Y = 14.8156 + 5.70657X$$

Thus, we see that the relevant coefficients are,

$$b_0 = 14.8156$$

$$b_1 = 5.70657$$

This value of  $b_0$  is equal to the initial salary value for a student with a GPA of 0. In this sense, it does not make sense. But, if we assume that this model is only used for students that have passed, i.e.  $GPA \neq 0$ , this begins to make more sense. In this way, we can call  $b_0$  a base salary that is then modified by higher GPA's

- b. Using the equation from Figure 3.16, we compute,

$$Y = 14.8156 + 5.70657 * 3.25 = 33.362$$

Because a GPA of 3.25 is within our experimental region, we note that this prediction of salary is also a point estimate for the value as well.

- c. We finally compute the Coefficients using the following equations,

$$SS_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$SS_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

Using R, we compute the values to be exactly as proposed in the figure.

### 3.7 Copier Data

- a. Using the figure, we identify the model to be,

$$Y = 11.4641 + 24.6022X$$

Thus, the relevant coefficients are,

$$b_0 = 11.4641$$

$$b_1 = 24.6022$$

This value of  $b_0$  corresponds to the initial amount of time taken to service 0 copiers. In this sense, this value doesn't make sense. If we make a few assumptions regarding how this job is performed, we can begin to extrapolate more value. For instance if we assume that the fixing of the copiers takes place in one location, we can call this value an amount of time required to set up for working, or acquiring more parts.

- b. Using the equation from Figure 3.18, we compute,

$$Y = 11.4641 + 24.6022 * 4 = 109.8729$$

Because 4 copiers is within our experimental region, we note that this prediction of time is also a point estimate as well.

- c. Finally, we compute the coefficients using the following equations,

$$SS_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$SS_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

Using R, we compute the values to be exactly as proposed in the figure.

### 3.9 Enterprise Industries

Using the scatter plot in Figure 3.19, we begin making some superficial remarks. First, we note that the overall shape of the data is roughly linear. This is a good indicator that a linear model will fit the data well. Next, we consider the nature of the relationship between  $y$  and  $x_4$ .  $x_4$  is the difference between the average price of detergent in the market and the cost of Fresh brand detergent. Hence,  $x_4 > 0 \Rightarrow$  The average cost is greater than the cost of Fresh, and  $x_4 < 0 \Rightarrow$