

## Appendix B

### The Least Squares Regression Line

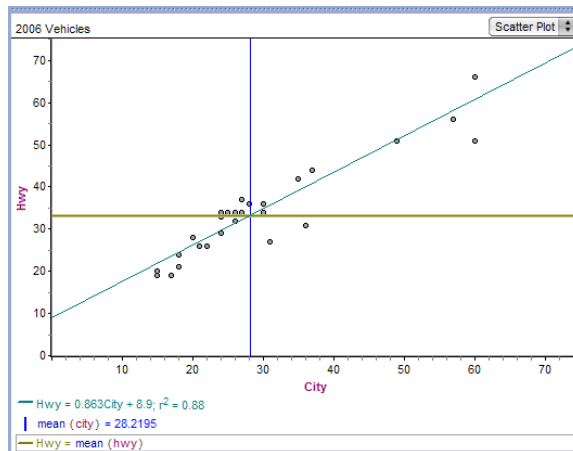
The least squares line is computed using means and standard deviations for each variable, and the correlation coefficient  $r$ . The following graph illustrates the scatterplot of the City and Hwy MPG from the 2006 vehicle data. In the scatterplot, the least squares regression line has been graphed, as well as vertical and horizontal lines representing the mean City MPG and mean Hwy MPG. Such a graph can help visualize the data and location of the least squares line in relation to the mean MPG for both City and Hwy.

Algebraically, the slope and y-intercept of the least squares line are:

$$\text{slope} = r \frac{s_y}{s_x} \quad \text{y-intercept} = \bar{y} - \text{slope} (\bar{x})$$

$r$  = correlation coefficient,  $s_y$  = Standard deviation of  $y$ ,  $s_x$  = Standard deviation of  $x$   
Thus the equation for the least squares line can be symbolically represented as

$$\hat{y} = r \frac{s_y}{s_x} (x) + (\bar{y} - r \frac{s_y}{s_x} (\bar{x})) \quad \text{Or an alternative form of} \quad \hat{y} = r \frac{s_y}{s_x} (x - \bar{x}) + \bar{y}.$$



#### FOCUS ON MATHEMATICS

Q33. The least squares regression line passes through the intersection of the mean City mpg and the mean Hwy mpg (see Figure). Will this always happen? Justify your answer algebraically.

#### FOCUS ON PEDAGOGY

Q36. If technologies like *Fathom*, as well as others such as Excel and graphing calculators, will compute and display the least squares line, would you choose to show students the algebraic form for computing the least squares line? Why or why not? Defend your position.