5.26 a. Group is the qualitative independent variable. It must be coded into two dummy variables since it has three levels.

b. \[ E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \]
   where \[ x_1 = \begin{cases} 1 & \text{if group 2} \\ 0 & \text{otherwise} \end{cases} \]
   \[ x_2 = \begin{cases} 1 & \text{if group 3} \\ 0 & \text{otherwise} \end{cases} \]

Mean milk production of cows in group 1 (man-made shade structures): \[ \beta_0 = \mu_1 \]

Difference in mean milk production between cows in group 2 and group 1 (tree shade minus man-made shade structure): \[ \beta_1 = \mu_2 - \mu_1 \]

Difference in mean milk production between cows in group 3 and group 1 (no shade minus man-made shade structure): \[ \beta_2 = \mu_3 - \mu_1 \]
5.36  

a. \[ E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_1 x_2 + \beta_5 x_1^2 x_2 \]

b. For non-coached students: \( x_2 = 0 \)

\[ E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 (0) + \beta_4 x_1 (0) + \beta_5 x_1^2 (0) \]

y-intercept: \( \beta_0 \)

shift parameter: \( \beta_1 \)

rate of curvature: \( \beta_2 \)
c. For coached students:  \( x_2 = 1 \)

\[
E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 (1) + \beta_4 x_1 (1) + \beta_5 x_1^2 (1)
\]

\[
= (\beta_0 + \beta_3) + (\beta_1 + \beta_4) x_1 + (\beta_2 + \beta_5) x_1^2
\]

y-intercept:  \( \beta_0 + \beta_3 \)
shift parameter:  \( (\beta_1 + \beta_4) \)
rate of curvature:  \( (\beta_2 + \beta_5) \)

d. To determine if coaching has an effect on SAT-Math, we test:

\[
H_0 : \beta_3 = \beta_4 = \beta_5 = 0
\]

\[
H_a : \text{At least one } \beta_i \neq 0
\]
a. All of the independent variables are continuous except District. The variables involving money have been coded so that they represent thousands of dollars, rather than dollars as in the problem statement. Four dummy variables were created for District before using the stepwise regression. Also, note that several observations had missing values. Thus, only 217 observations were used in the analysis. Since Ratio was computed from the ratio of Price and DOT estimate, it was not used as a predictor variable. Using MINITAB and $\alpha = .10$ for keeping variables in the model, the Stepwise regression results are:

**Stepwise Regression: LOWBID versus DOTEST, LBERATIO, ...**

Alpha-to-Enter: 0.1  Alpha-to-Remove: 0.1

Response is LOWBID on 14 predictors, with N = 217
N(cases with missing observations) = 62  N(all cases) = 279

<table>
<thead>
<tr>
<th>Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>33176</td>
<td>-723253</td>
<td>-748001</td>
</tr>
<tr>
<td>DOTEST</td>
<td>0.9064</td>
<td>0.9147</td>
<td>0.8953</td>
</tr>
<tr>
<td>T-Value</td>
<td>94.11</td>
<td>105.86</td>
<td>62.92</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>LBERATIO</td>
<td>770762</td>
<td>763106</td>
<td></td>
</tr>
<tr>
<td>T-Value</td>
<td>7.58</td>
<td>7.53</td>
<td></td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>DAYSEST</td>
<td>260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-Value</td>
<td>1.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-Value</td>
<td>0.089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>278547</td>
<td>247868</td>
<td>246764</td>
</tr>
<tr>
<td>R-Sq</td>
<td>97.63</td>
<td>98.13</td>
<td>98.16</td>
</tr>
<tr>
<td>R-Sq(adj)</td>
<td>97.62</td>
<td>98.11</td>
<td>98.13</td>
</tr>
<tr>
<td>Mallows Cp</td>
<td>61.0</td>
<td>4.9</td>
<td>4.0</td>
</tr>
</tbody>
</table>

From the results, only three independent variables are selected to predict price using the stepwise regression and $\alpha = .10$. These variables are DOT estimate, Status, and Days.

b. $\hat{\beta}_0 = -59.79987$. This value has no meaning because an observation with all the independent variables equal to 0 is not in the observed range.

$\hat{\beta}_1 = .88616$. For each dollar increase in the DOT estimate, the mean price is estimated to increase by .88616 dollars, holding all other variables constant.

$\hat{\beta}_2 = 139.40952$. The mean price is estimated to be 139,409.52 dollars higher for fixed bids than for competitive bids, holding all other variables constant.

$\hat{\beta}_3 = .35761$. For each additional day to complete work, the mean price is estimated to increase by 357.61 dollars, holding all other variables constant.

c. Because of the large number of t tests performed in the stepwise regression, the probability is very high that one or more errors have been made in including and excluding variables. Also, we have not considered any higher order terms.