**Simple Linear Regression Using Minitab 17**

Table 11.1 page 393 Walpole, Walpole, Myers, Ye

|  |  |
| --- | --- |
| X % solids reduction | Y % chemical oxygen demand |
| 3 | 5 |
| 7 | 11 |
| 11 | 21 |
| 15 | 16 |
| 18 | 16 |
| 27 | 28 |
| 29 | 27 |
| 30 | 25 |
| 30 | 35 |
| 31 | 30 |
| 31 | 40 |
| 32 | 32 |
| 33 | 34 |
| 33 | 32 |
| 34 | 34 |
| 36 | 37 |
| 36 | 38 |
| 36 | 34 |
| 37 | 36 |
| 38 | 38 |
| 39 | 37 |
| 39 | 36 |
| 39 | 45 |
| 40 | 39 |
| 41 | 41 |
| 42 | 40 |
| 42 | 44 |
| 43 | 37 |
| 44 | 44 |
| 45 | 46 |
| 46 | 46 |
| 47 | 49 |
| 50 | 51 |

|  |  |
| --- | --- |
| **Step 1**  **Graph the data.**  **Step 2**  **Evaluate the graph.**  **Step 3**  **If the assumption of a linear relationship seems reasonable, use Minitab 17 to estimate the equation.** |  |

**The Minitab 17 output is shown in the green block below:**

**Regression Analysis: y % chemical oxygen demand versus x % solids reduction**

Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Regression 1 3390.6 3390.55 325.08 0.000

x % solids reduction 1 3390.6 3390.55 325.08 0.000

Error 31 323.3 10.43

Lack-of-Fit 23 156.0 6.78 0.32 0.984

Pure Error 8 167.3 20.92

Total 32 3713.9

Model Summary

S R-sq R-sq(adj) R-sq(pred)

3.22954 91.29% 91.01% 90.02%

Coefficients

Term Coef SE Coef T-Value P-Value VIF

Constant 3.83 1.77 2.17 0.038

x % solids reduction 0.9036 0.0501 18.03 0.000 1.00

Regression Equation

y % chemical oxygen demand = 3.83 + 0.9036 x % solids reduction

**Interpretation:**

The estimated value for y equals b0 + b1 x where b0 is the Constant Coefficient and b1 is the X Coefficient.

*How good is our estimate?* Conduct a hypothesis test.

H0: b1 = 0

H1: b1 ≠ 0

Based on the Minitab 17 output, the T-value for x is 18.03. The p-value is less than our assumed alpha of 0.05%. Therefore we reject H0.

Since we rejected the null hypothesis (H0), we can conclude that the x variable explains a significant portion of the variability in y.

Thus, in our example, percent solids reduction (the x term) explains a significant portion of the variability in percent chemical oxygen demand.

Analysis of Variance ***original***

Source DF Adj SS Adj MS F-Value P-Value

Regression 1 3390.6 3390.55 325.08 0.000

x % solids reduction 1 3390.6 3390.55 325.08 0.000

Error 31 323.3 10.43

Lack-of-Fit 23 156.0 6.78 0.32 0.984

Pure Error 8 167.3 20.92

Total 32 3713.9

Analysis of Variance ***simplified***

Source DF Adj SS Adj MS F-Value P-Value

Regression x % solids reduction 1 3390.6 3390.55 325.08 0.000

Error 31 323.3 10.43

Total 32 3713.9

*How many observations were used to develop the model?*

33

*How do we calculate Adj MS?*

Adj SS / DF

*How do we calculate F?*

3390.55/10.43 = 325.08

*Is the regression significant? Why?*

Yes Regression P-value less than 0.05

Yes Regression P-value very close to 0

*What percent of the variability in Y is explained by our regression equation?*

The coefficient of determination (R2 )is a measure of the proportion of the variability in Y explained by our model. We will use the adjusted value. Therefore 91.01% of the variability in % chemical oxygen demand is explained by the model “% chemical oxygen demand = 3.83 + 0.9036 x % solids reduction”.