

Section 4.3

Diagnostics on the Least-squares Regression Line

Objectives

- Compute and interpret the coefficient of determination
- Perform residual analysis on a regression model
- Identify influential observations

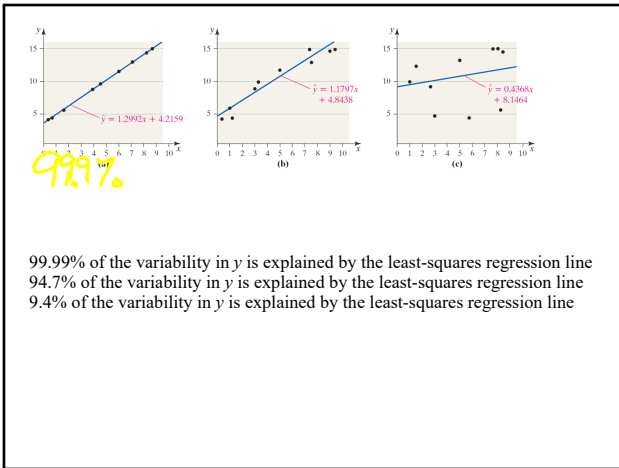
The **coefficient of determination, R^2** , measures the proportion of total variation in the response variable that is explained by the least-squares regression line.

$$0 \leq R^2 \leq 1$$

If $R^2 = 0$ the line has no explanatory value

If $R^2 = 1$ means the line explains 100% of the variation in the response variable.

To determine R^2 for the linear regression model simply square the value of the linear correlation coefficient.



Find and interpret the coefficient of determination for the drilling data.

$$r^2 = .597$$

$$59.7\%$$

Depth at Which Drilling Begins, x (in feet)	Time to Drill 5 Feet, y (in minutes)
35	5.88
50	5.99
75	6.74
95	6.1
120	7.47
130	6.93
145	6.42
155	7.97
160	7.92
175	7.62
185	6.89
190	7.9

So, 59.7% of the variability in drilling time is explained by the least-squares regression line.

- Scatter Plot Stat Plot
- Correlation Coefficient Stat, Calc, Lin Reg
 $r = .773$
- Is it Linear?
 $n=12$ $CV = .576$ $|.773| > .576$ Yes it is linear
- Find the least-squares Reg line.
 $\hat{y} = .012x + 5.527$
- Predict the drilling time for 75 feet.
 $.012(75) + 5.527$
6.427
Good estimate since 75 is within the scope of the data.

- At 75 feet we know the drilling time to be 6.74 minutes. Is our drilling time above or below the average drilling time at that depth?
6.427 (predicted)
Above
- Interpret the slope.
Slope: .012 For every foot increase in depth the drilling time increases by .012.
- Interpret the y-intercept.
y-int = 5.527 Drilling at ground level.
- What proportion of variability is explained in the time of drilling?
 $r^2 = .597$