Stats Starter 12/10

3.1 (p. 159) #9

3.2a Regression Line

A line that describes how y changes as x changes

- Model for our data
- used to predict y values

Least Squares Regression Line

(LSRL) the line of best fit

Minimizes the sum of the squares of the residuals

residual: the vertical distance between an actual point and the LSRL

residual = Actual − Predicted

$= y - \hat{y}$

negative for points under the line.

$\hat{y} = a + bx$

- $\hat{y}$: y-hat (predicted y-value)
- a: y-intercept (the predicted value when x = 0)
- b: slope (amount y increases when x increases by 1)

ALWAYS INCLUDE CONTEXT!

Calculating the LSRL (and r):

**First time: Turn Diagnostics On.

one time thing — 2nd --> 0 --> Diagnostics On --> enter --> enter

Enter in 2 lists

Stat --> Calc --> 8:LinReg(a+bx) --> enter

_____ , _____

x-list y-list

ex: $\hat{y} = -10.62 + 9.92x$

$\hat{y}$ = the predicted time to write name 20x

x = # of letters

time to write name 20x $= -10.62 + 9.92$ (# of letters)

Interpreting:

slope: "For each additional ____ (unit) of ____ (explanatory var.) there is a predicted ____ (increase / decrease) of about ____ (slope) ____ (response var.)."

y-intercept: (often doesn't make sense) "When ____ (explanatory) is zero, the predicted ____ (response) is ____ (y-int. #)."

residual: "The regression line ____ (over/under) predicted the ____ (response var.) by ____ (length of residual)."

Predicting Y for a specific X

plug in an x, get out a $\hat{y}$.

$\hat{y} = -10.62 + 9.92x$

$\hat{y} = -10.62 + 9.92(7)$

$= 58.82$ seconds

Extrapolation:

using the LSRL to predict y values that are outside our data

* usually not accurate!